THE THEORY OF SIGNAL DETECTABILITY:

ROC CURVES AND THEIR CHARACTER

by *(Theodore Gerald)*

Theodore G. Birdsall

COOLEY ELECTRONICS LABORATORY

Department of Electrical and Computer Engineering
The University of Michigan
Ann Arbor, Michigan

for

January 1973

ABSTRACT

The first problem in the theory of signal detectability deals with the decision between two alternative responses, corresponding to two possible classes of causes of an observation. When the goal of a decision process is to achieve the highest quality of terminal decision, the Receiver Operating Characteristic curve (ROC curve) contains all of the information necessary for the evaluation of the decision process. This present work introduces the ROC character, which is isomorphic to the ROC curve.

The formal development is based on two key facts. The first is the fundamental theorem: if $\ell(X)$ is the likelihood ratio of an observation, then the likelihood ratio of ℓ is ℓ itself. The second is the main theorem on ROC characters: each ROC character is isomorphic to a univariate probability distribution that possesses a moment generating function. The character convolution theorem and the character addition theorem follow directly from these.

Families of ROC curves are developed from the main theorem on ROC characters. The normal, binormal, Q-table, power, and several discrete families of ROC curve have appeared in the literature. The new families include the Pearson type III, Fisher-Tippett doubly exponential, H-type, Poisson, and the regular conics.

Additional families are generated from these by use of the meta-static relation, and the convolution and addition theorems.

ROC curves contain information about the performance of other two-cause decisions besides the two-response decision. Several are considered that are used in the testing of human perception; namely, the symmetric forced choice decision, type II decisions, the rating scale procedure, and the analysis of a decision based on the reports of multiple observers.

FOREWORD

This report contains foundational work in signal detection theory. The primary purpose of the work was to provide a technique for cataloging performance curves of detection-decision devices, and to expound upon properties of specific performance curves. Its usefulness lies in the exhaustive detailing, which the author hopes will provide detection theory researchers with both insight and detailed information about performance curves they encounter.

ACKNOWLEDGMENTS

The author wishes to thank the members of his committee, Professor Hok who was especially helpful in the initital development of the work, Professor Macnee and Professor Peterson who have aided and encouraged this and similar work over a number of years, and who patiently worked with the several drafts of this dissertation, Professor Tanner and Professor Pollack who have given guidance and example in psychophysical research extending over a number of years. Special thanks are also due to Professors R. V. Churchill, H. H. Goode, and A. H. Copeland, Sr., for encouragement in continuing study and research in applied mathematics.

This present study is based principally on the work of W. W. Peterson, W. P. Tanner, Jr., J. A. Swets, D. M. Green, and J. P. Egan, and the experience gained while working with these men.

The author is indebted to the U.S. Army Signal Corps, Countermeasures Branch, for support of research in signal detectability from 1952 to 1957, and to the U.S. Navy Office of Naval Research, Acoustics Programs, for continuing support of research in extending the theory of signal detectability since 1960, including the research and writing of this thesis.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS (Cont.)

LIST OF ILLUSTRATIONS(Cont.)

LIST OF ILLUSTRATIONS (Cont.)

LIST OF ILLUSTRATIONS (Cont.)

LIST OF TABLES

LIST OF SYMBOLS

The following letters are used repeatedly to denote constants or parameters: A, B, C, D, E, K, a, b, c, k, α, β, γ, δ.

The following is a list of special symbols:

Symbol	Page	Definition
"A"	5	the "alarm" response, correct for condition SN
"A_g", "A_h"	15	"A" due to decision devices g, h, et .
A_0	13	a designated set of N-measure zero
A		a set; index of pure power curves; "A"
$\mathcal{A}$	19	a σ-ring on the space X, Chapter I
A_D	20	a difference set
a. e.	46	read "almost everywhere"
"B"	5	the "background" response, correct for condition N
B		a set; "B"; an index
b		a special constant (or variable) in Chapter III
$\mathcal{B}$	22	a σ-ring on the space Y, Chapter I
C		a set; the event "correct"; a chance point
$\mathcal{C}$		a σ-ring on the space W, Chapter I
D(z)	Eq. 6.22	a function used only in Chapter VI
d(x)	13	a real function of the observation x, decision axis

LIST OF SYMBOLS (Cont.)

Symbol	Page	Definition
d	72	quality index, normal ROC
d_a	288	a parameter approximately equal to d
$\hat{d}$		a parameter approximately equal to d
d'	Eq. 5.68	a quality index for an ROC point, based on normal scale
d_{WWP}	Eq. 10.29	W. W. Peterson's approximate quality index
$\sqrt{d_e}$	143	a quality index for an ROC negative diagonal point
$E(u \mid v)$		expected value of u, conditional to v
$F(u \mid v)$		probability distribution function of u, conditional to v
$f(u \mid v)$		probability density function of u, conditional to v
$f_\pi(u)$	56	root product density function of u
$g(x, \beta, r)$	Eq. 1.6	a decision function based on likelihood ratio
$h(x)$		a designated decision function
$i(u)$	25	the identity map $i(u) = u$
iff		read "if and only if"
J		Jacobian of transformation
$J(1:0)$	Eq. 10.27	the "divergence between hypotheses"
K	Eq. 5.43	a specific constant for conic ROC characters
k	Eq. 10.61	a specific scale factor

Symbol	Page	Definition
K_0, K_1, K_2	Eq. 10.80	arbitrary scale factors
K_N, K_{SN}	Eq. 6.1	specific constants in the metastatic normal ROC
$\ell(x)$	Eq. 1.15	likelihood ratio of the observation x
ℓ		a likelihood ratio
$L(u)$	Eq. 5.66	the log-odds-ratio of u, the LOR scale
M		the integer multiplicity of the number of signals; the maximum of ℓ
M'	Eq. 5.70	a quality index for an ROC point, based on LOR scale
$M_F(\xi)$	Eq. 3.15	the moment generating function of the distribution F
m		the minimum of ℓ; a parameter
N		a cause class; the noise power
$P(u \mid v)$		the probability of u, conditional to v
Pr, Prob		probability
$p("A" \mid x)$	Eq. 1.1	abstract expression of a decision device
$P_M(\ )$	Eq. 5.13	principal choice probability in a forced choice experiment
Q_β	Eq. 1.7	a special formula
q	Eq. 7.2	quality index, binormal ROC
R		the reals; index for rectangular ROC character; a scale factor for pay-off matrices; the condition "right" for Type II ROC curves

Symbol	Page	Definition
r		a randomizing decision value
$r(u)$, $r_0(u)$		rating scale forcing function, Chapter X
r_a, r_b	215 208	constants related to metastatic normal
S		signal power
s	Eq. 7.2	binormal ROC slope parameter
$S_M(\ )$	Eq. 5.26	second choice probability in a forced choice experiment
SN		a cause class
$S_j N$	122	the j-th subset of SN
$(snr)_\ell$	Eq. 10.24	signal-to-noise ratio based on ℓ
$(snr)_z$	Eq. 10.28	signal-to-noise ratio based on z
T	234	point of tangency of ROC curve and tangent inferior
t		a real random variable
t_N, t_{SN}	319	values of t for the indicated cause
t_0	Eq. 7.10	value of t at z_0
t_T	239	value of t at the point of tangency
u, u_i		random variables
u	Eq. 10.61	observer's optimum rating scale report conditional to indicated decision
V_D, V_{FA}, V_M, V_{CR}	Fig. 10.5	pay-off matrix values

Symbol	Page	Definition
$V_A(\ell)$, $V_B(\ell)$		expected values conditional to designated ℓ value
$V_{n,k}$	Eq. 8.18	ROC vertex
W		a real subspace; the condition "wrong" for Type II ROC curves
w_i		an ROC point
X, $\mathscr{X}$		an observation space
X, x		an observation
X, x		the ROC abscissa, the false alarm probability $P("A"|N)$
$(x(\beta, r),\ y(\beta, r))$	49	ROC point corresponding to $g(x, \beta, r)$
Y	22	a space
Y, y		the ROC ordinate, the probability of detection $P("A"|SN)$
$Y(X)$, $y(x)$		formula for an ROC curve
Z, z		natural logarithm of the likelihood ratio, i.e., $\ln \ell$
z_0, z_0		minimum z; mode z
z_a, z_b	Eq. 6.2	maximum and minimum values of z, Chapter VI
z_c, z_d	Eq. 6.45	z values at $f(z\,|^{SN}_{\ N})$ modes
z_A, z_B	Eq. 10.45	z conditional to indicated response
z_T	Eq. 7.14	z values at point of tangency T
z	381	a cut value of z
1_B	23	the characteristic function of the set B

Symbol	Page	Definition
2WT	117	the number of independent variables
β		a cut in a decision axis, especially the ℓ-axis
Γ	Eq. 5.43	a conic ROC constant
Δ	Eq. 5.43	the conic ROC discriminant
$\lambda(\ \mid\ )$	27	an induced pair of measures
$\mu(\ \mid\ )$	Eq. 1.2	the observation statistics; a pair of measures
$\nu(\ \mid\ )$	23	an induced pair of measures
$\xi,\ \xi(x)$	39	a decision axis
$\Pi(z)$	55	the integrated ROC character
$\pi(z)$	53	the ROC character
$\pi_E(z)$		the exponential ROC character
$\pi_{MN}(z)$	Eq. 6.8	the metastatic normal ROC character
$\pi_N(z)$		the normal ROC character
$\pi_R(z)$		the rectangular ROC character
ρ	75	a Pearson III parameter
σ^2		variance
$\phi_{\pi\pi}(\tau)$	Eq. 10.14	the autocorrelation function of the ROC character
$\omega(\ \mid\ )$	46	a discrete probability pair; jumps in a pair of distribution functions
$\omega_\pi(z)$	54	the ROC character jump

The following functions are treated as well-known.

ln x	natural logarithm
e^x	exponential
sinh x, cosh x, sinh x, sech x $\Bigg\}$	hyperbolic trig
$I_0(x)$, $I_1(x)$	modified Bessel functions of the first kind
P(a, t)	incomplete gamma function
$\Gamma(x)$	complete gamma function
$\Phi(x)$	normal distribution function
$\phi(x)$	normal density function
$P(\chi^2 \mid \nu)$	chi-square distribution function with ν degrees of freedom
$Q(\chi^2 \mid \nu)$	$1 - P(\chi^2 \mid \nu)$
$Q(\alpha, \beta)$	Marcum's Q-function

LIST OF APPENDICES

CHAPTER I

INTRODUCTION

The theory of signal detectability has grown up in
an interdisciplinary way, combining parts of the traditional fields of
mathematics, physics, engineering, physiology, and psychology.
The skeleton on which the theory is built is the mathematics of
decision theory. The analytic description of the observation and per-
ception, the interferences and the cause-effect relationships are drawn
chiefly from physics and physiology. The subsequent realization of deci-
sion mechanisms is part of engineering when the realization is in hard-
ware, part of psychology when one is dealing with a human or animal
decision mechanism, and falls to other specialties when the realization
is a computer program or a finite automaton. The evaluation of the perfor-
mance of the decision mechanism has similarly fallen to various fields,
depending upon the application.

This present work concentrates on models for the third
aspect, evaluation of detectability. Consideration is restricted to those
situations where decision quality is the sole goal of the decision device,
and the sole consideration in the evaluation. In the past, evaluation
models have come principally from problems in electrical engineering.
Yet the theory is applicable to any information process involving both
observation and subsequent (binary) decision. Today's applications
involve such diverse areas as radar, sonar, and digital communications;

the psychophysics of human and animal perception; information retrieval and medical diagnoses by machines. One major purpose of this present work is to furnish a variety of evaluation models. A second major purpose is to develop sufficient mathematical structure to form a basis for classification of these models, and to form a basis for the generation of additional models. It is hoped that this work will be useful to those making mathematical analyses of detection devices by specifying the parameters that must be determined. It is hoped that it will be useful to those experimenters with performance data, that it will help classify the data and indicate appropriate models. For the mathematically inclined this paper develops a special L_2 function called the ROC character, that is one-to-one with the graphical presentation, the ROC curve. I hope those in communications sciences with a broad interest in observation-decision procedures will be able to learn what information an ROC curve furnishes, and what information it omits.

We shall concentrate on the simplest of all observation-decision problems -- the two-by-two problem. This work is directly applicable to the two-by-E^1 (rating) problem; however, no generalization to finite higher order decision problems will be made. In the two-by-two problem, it is assumed that two alternative decisions may be made, corresponding to two mutually exhaustive and exclusive classes of causes. In order to make any sensible decision a given cause (such as a specific signal) must belong to one class or the other, and not transfer back and forth. This is the reason for the word "exclusive." The assumption

"exhaustive" means that no other cause than those listed in the two classes

can occur.

This present work will deal principally with three subjects:

(1) the ROC curve, (2) the likelihood ratio and related functions, and

(3) a new function, $\pi(z)$, called the ROC character. In this first

chapter the basic relations between the ROC curve and the likelihood

ratio will be reviewed and the proof given for the fundamental theorem:

"The likelihood ratio of the likelihood ratio is the likelihood ratio. "

This theorem has been known and used by several specialists in the

field, but apparently has not previously been rigorously established.

In Chapter II the ROC character will be formally introduced and it will

be shown that it contains information that is unavailable in the likelihood

ratio itself. Even further it will be shown that the likelihood ratio and

the ROC character contain the two complementary and separate aspects

of detection; namely, that likelihood ratio is relevant to the design of

the decision device and that the ROC character is relevant to its perfor-

mance. In Chapter III the ROC character will be used to establish a

basis of classification of types of ROC curves. The primary analytic

result in Chapter III is that the class of all ROC characters corresponds

to the class of probability functions that are sufficiently smooth and

concentrated to have moment generating functions. This correspondence

yields the set of classical ROC curves.

Most of the ROC models available heretofore have come

from radar detection problems. In Chapter IV these few are collected

and analyzed in terms of their ROC character. Chapters V, VI and VII

each treat at length one broad class of ROC, the algebraic, the truncated

normal, and the binormal. Chapter VIII develops the Convolution

Theorem for ROC characters, and the Fisher-Tippett ROC character,

both arising in cases of multiple observations. Chapter IX presents

the ROC character Addition Theorem, and two double parameter charac-

ters, the Pearson Type III, and the H-type. Chapter X is devoted to

discussion of selected topics that have arisen in the literature.

1. 1 Review of Detection Theory

The procedure that will be followed in this section has

two gross steps. First, the graph called the receiver operating charac-

teristic (ROC) will be introduced and used to evaluate detection per-

formance. Secondly, a specific function of observation called the

"likelihood ratio" will be introduced and it will be explained why

optimum detection is based on likelihood ratio.

There are many proofs that the likelihood ratio for de-

tection problems is optimum. The most common proofs are based on

specific definitions of optimum, such as "maximum expected value,"

"minimum loss," or "Neyman-Pearson test." The proof used here

has been chosen to emphasize what is well known, but really unknown

to many -- that the optimumness of likelihood ratio does not require

that any specific quantity be maximized or minimized. All that is

required is that decision quality is the goal and that the definition of

decision quality reflects a preference for correct decisions over mistakes.

At this point, we must begin to introduce formal notation in order to be specific, and to be able to manipulate mathematically certain rather simple quantities. Many types of notation might be used, all of which appear in decision theory and detection theory. One might choose to emphasize the generality of the work by using abstract notation. I have chosen to utilize the notation common in radar and sonar signal detection, in the hopes that the nmemonic benefits gained will over-shadow the apparent loss of generality. Specifically, the two classes of causes will be signified by N (noise alone) and SN (there is a signal in the noise). Two alternative decisions that the decision device can make are to sound the alarm ("A", the decision that a signal was present) or to conclude that only background was present, the decision "B". When the cause is noise alone, N, the decisions correspond to a false alarm and a correct response, respectively; when the cause is SN the decisions correspond to a detection and a miss, respectively.

The formal description of a decision device is a probability function. For each possible observation the description is the probability that the device will response "A". One might feel that a description of a decision device would correspond to a listing of those observations for which the decision device would response "A". (On the remainder of the observations the decision device would respond "B".) Indeed, this has been the writer's standard approach in electronic problems (Ref. 1, p. 174). Although this description is quite adequate for a great many purposes, it does not cover those devices which may base their decision in part on extraneous quantities: for example devices whose decisions are

affected by internal noise or devices with some nonstationary behavior

in the decision process (such as a jittering decision threshold.) The

decision probability function is written p("A"|x), where x is a general

notation for the observation on which the device bases its decision.

In Fig. 1.1 the coordinate systems for the ROC are

drawn on a blank unit square. The horizontal axis indicated on the

bottom is called the probability of the false alarm. It represents the

probability of a mistake, the probability of making a decision "A" when

the correct decision would have been "B. " The probability of the corres-

ponding correct decision is shown on the horizontal axis above the ROC

and is simply one minus the probability of a false alarm. (The device is

forced to make either an "A" or "B" decision.) The left vertical scale is

the probability of making the "A" decision when it is correct, and

is called the probability of detection. Had the alternative decision

"B" been made the action would have been labeled a miss and its

probability is simply one minus the probability of detection. This

probability is indicated on the right-hand side of the graph on the
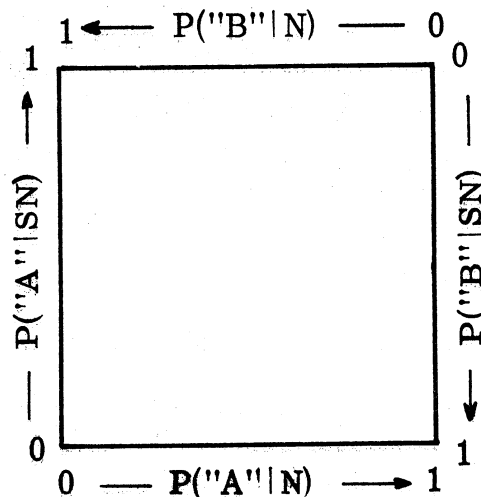
Fig. 1. 1. Blank ROC

vertical axis. Traditionally, one suppresses the upper and right-hand coordinates, listing only the probabilities of false alarm and detection, since the other two probabilities can be obtained readily from these. The term ROC first appeared in the paper by Peterson, Birdsall, and Fox (Ref. 1). Curves drawn on the ROC grid have also been referred to as contours of iso-effectiveness, or contours of equivalent performance. The traditional form of the curve that appears in statistics and some work in signal detection plots the probability of a miss on the vertical axis so that the ROC is the mirror image of the form used here. In statistics such a curve is called the power curve, where the probability of false alarm is called the critical level and the power is one minus the probability of a miss, that is, the probability of detection.

Our formalism for a decision device is the probability function describing the probability of deciding "A" for each observation x on which a decision is based.

$$\text{Decision Device: } p(\text{"A"} \mid x) \tag{1.1}$$

Our formalism for the observation statistics, describing a specific problem, are the measures of the probability of occurrence of the observation, x, for each of the cause classes.

$$\text{Observation Statistics: } \mu(x \mid SN), \ \mu(x \mid N) \tag{1.2}$$

These observation statistics refer specifically to those observations

on which a decision will be based. This is more restrictive than the

general use of the same term which might mean the statistics of the

entire random process of which the observation is just one portion.

A specific observation situation is defined by specifying the obser-

vations on which a decision may be based, and the observation

statistics, Eq. 1.2.

The evaluation of a specific decision device in a

specific observation situation is the ROC point given by the proba-

bility of a false alarm and probability of a detection.

ROC Evaluation Point: $P(A|N) = \int p(\text{"A"}|x)\, d\mu(x|N)$

$$(1.3)$$

$$P(A|SN) = \int p(\text{"A"}|x)\, d\mu(x|SN)$$

In Appendix A, taken from course notes for the Summer Intensive

Course on Random Process, 1961-62-63, it is shown that the ROC

plot of all possible decision devices in any specific decision situation

is convex, contains the chance diagonal, and is symmetric about the

center-point. The proofs are simple but tedious. The results are

shown in the following figures. In Fig. 1.2 the chance diagonal has

been added to the blank ROC. This is the locus of performance where

the detection probability is equal to the false alarm probability. Such

performance may be obtained with a decision device that reacts indepen-

dently of the observation, and hence is described as chance. Performance

falling below this diagonal is considered poorer than chance and per-

formance above the diagonal is considered better than chance.
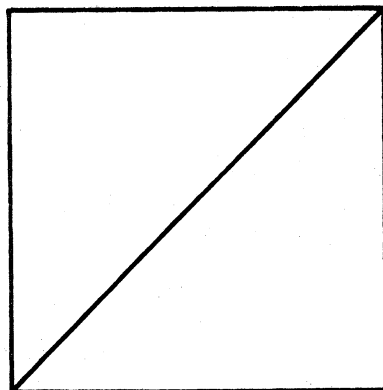
Fig. 1.2. The "chance diagonal "

We have made no specific definition of optimum. However,

let us consider some specific point below the chance diagonal and compare

it to the point directly above it on the chance diagonal. Both have the

same performance under the condition N. The point on the chance

diagonal has a higher detection probability than the point under discussion

below the chance diagonal. We did specify that correct decisions are

to be preferred to incorrect decision, that higher correct-decision

probabilities and correspondingly lower error probabilities will be

considered desirable. Therefore the chance diagonal is better than

any point plotted below it.

The convex closure of two points plotted on the ROC

grid consists of the straight line joining the two points. The implication

is that if two decisions devices are possible which plot at points

w_1 and w_2, then it is known that a continuum of decision devices

exist which will perform as indicated by the line between the points

w_1 and w_2.
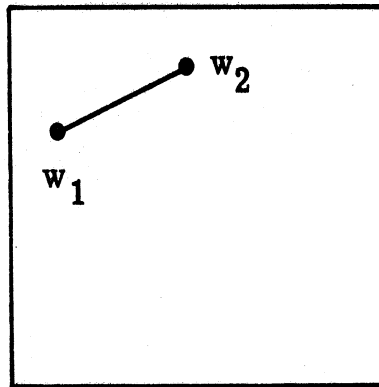
Fig. 1.3. The "convex closure" of two points

Consider two decision devices which disagree on each individual decision. The probability of "A" of the second device will be the probability of "B" of the first device.

$$P("A_2"|N) = P("B_1"|N)$$

$$(1.4)$$

$$P("A_2"|SN) = P("B_1"|SN)$$

Equation (1.4) is the reason the ROC region exhibits the midpoint symmetry that is indicated geometrically in Fig. 1.4.
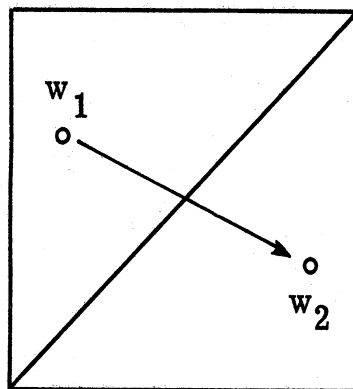
Fig. 1.4. Existence of midpoint symmetry

The purpose of this detailed geometric examination was to determine the implications of the existence of a single non-chance point. If there is a particular decision device with nonchance performance, plotted at w_1 , then it follows directly that the region of obtainable performance with the given observation statistics contains at least the parallelogram indicated in Fig. 1.5. The boundary of this parallelogram is the convex closure of the points $(0, 0)$, w_1, $(1, 1)$, and the opposite of w_1. Every point in the interior lies on many straight lines between boundary points.
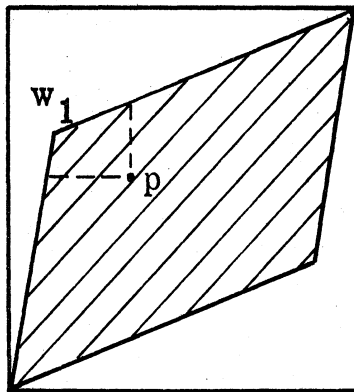
Fig. 1.5. Convex closure of a point, its opposite, and chance

The plot of all possible ROC performance points for a specific observation situation is called the ROC region. The ROC region is a function of the observation situation, not of any particular

decision device. Different observation situations may lead to different ROC regions. If two observation situations have the same ROC region, they are said to be equally detectable situations. Two observation situations may differ in all respects; the observations on which decisions are based, and the observation statistics. If two observation situations have the same set of observations, but differ in one of the measures in the observation statistics, they are considered as different situations.

For the remainder of this section (1. 1) we shall consider the observation situation fixed, and the ROC region known. If the ROC region is as shown in Fig. 1. 5, then one should desire to operate somewhere along the upper boundary. This is because each point p on the interior represents a higher miss probability than the upper boundary point directly above p, and a higher false alarm probability than the upper boundary point directly to the left of p. The upper boundary of the plot of the ROC region is called the optimum ROC curve. If one restricts himself to only the upper boundary instead of the entire region, he can still maximize any specific goal, that is, he can maximize every specific definition of performance.

The second step in this section is to establish that the ROC region is closed, meaning that the upper boundary actually exists and is obtainable, and that the widely acclaimed likelihood ratio will always lead one to performance on the upper boundary. Of special

interest are those decision devices which are based on the real quantity

likelihood ratio in a special way. A decision is said to be based on a

real decision axis $d(x)$ and a cut value β, if the decision "A" occurs

with probability one whenever $d(x)$ is greater than β and occurs

with probability zero whenever $d(x)$ is less than β. Whenever $d(x)$

equals β, the decision probability is some value, r, between zero

and one.

Because this present treatment is slightly more general

than the usual likelihood ratio treatment, such as is given in Appendix A,

we must go back to the foundations of the likelihood ratio principle.

The likelihood ratio of an observation is the connecting link between

two possible measures of occurrence of that observation. The Radon-

Nicodym Theorem (Ref. 2) establishes this connecting link. Consider

two observation statistics $\mu(x \mid N)$ and $\mu(x \mid SN)$. (These are assumed

to be measures on the same completely additive field of sets.) The

Radon-Nicodym Theorem states that there is a function, $\ell(x)$, and a

set of N-measure zero A_0, such that Eq. 1.5 holds for any mutually

measurable set.

$$
\left.
\begin{aligned}
\mu(C \mid N) &= \int_{C-A_0} d\mu(x \mid N) \quad, \quad \mu(A_0 \mid N) = 0 \\[2em]
\mu(C \mid SN) &\quad \int_{C-A_0} \ell(x)\, d\mu(x \mid N) + \mu(C \cap A_0 \mid SN)
\end{aligned}
\right\} \quad (1.5)
$$

To obtain the N-measure of some collection of observations, C, integrate

the number one with respect to the N-measure. To obtain the SN-measure

of the same collections of observations, integrate the weight $\ell(x)$ with res-

pect to the N-measure and add on the SN-measure of $C \cap A_0$. This allows

one to obtain both measures by integration with respect to $\mu(x \mid N)$,

except for sets of N-measure zero. The function $\ell(x)$ is never really

utilized when x is in the set of N-measure zero, A_0. We may

define $\ell(x)$ to be infinity on this set as a notational convenience. Such a

function $\ell(x)$ is called "the Radom-Nicodym derivative of the SN-measure

with respect to the N-measure," and is the most general form of a like-

lihood ratio.

The vast majority of the effort in the physical theory of

signal detectability is devoted to obtaining explicit equations for $\ell(x)$,

and to realizing practical equipment which will compute its value. In

contrast, this present work is devoted to evaluating performance and not

to obtaining performance. The meaning of the terminology "a decision

is based on likelihood ratio with cut value β and boundary value r" is

given by Eq. 1.6.

$$0 \leq \beta < \infty, \qquad 0 \leq r \leq 1$$

$$p(\text{"A"} \mid x) = g(x, \beta, r) = 1 \qquad \ell(x) > \beta$$
$$= r \qquad \ell(x) = \beta \qquad (1.6)$$
$$= 0 \qquad \ell(x) \quad \beta$$

To establish that such decision functions fall on the upper boundary

of the ROC region compare its performance to the performance of any other possible decision function, which is labeled h(x). The proof is shortest if we consider the contrived function Q_β.

$$Q_\beta = [\, P("A_g" \, | SN) - P("A_h" \, | SN) \,] - \beta[\, P("A_g" \, | N) - P("A_h" \, | N) \,]$$

$$(1.7)$$

Q_β is a comparison of the probabilities of detection for the g and h decision devices, together with a β-weighted comparison of the probabilities of false alarm of the two devices. The probabilities of false alarm are given by direct integration with respect to the N-measure,

$$P("A_g" | N) = \int_{X-A_0} g(x, \beta, r) \, d\mu \, (x \, | N)$$

$$P("A_h" | N) = \int_{X-A_0} h(x) \, d\mu \, (x | N)$$

$$(1.8)$$

where $X-A_0$ means those observations that are not in the set of N-measure zero, A_0. The probabilities of detection could be obtained by similar integration with respect to the SN-measure. However, utilizing the Radon-Nicodym Theorem and likelihood ratio we may evaluate the probabilities of detection by again using the N-measure.

$$P("A_g" | SN) = \int_{X-A_0} g(x, \beta, r) \, \ell(x) \, d\mu \, (x \, | N) + \int_{A_0} g(x, \beta, r) \, d\mu \, (x | SN)$$

$$(1.9)$$

$$P("A_h" | SN) = \int_{X-A_0} h(x) \, \ell(x) \, d\mu \, (x | N) + \int_{A_0} h(x) \, d\mu \, (x | SN)$$

The six integrals in Eqs. 1.8 and 1.9 can be collected into two integrals

for determining the value of the quantity Q_β.

$$Q_\beta = \int_{X-A_0} [g(x,\beta,r) - h(x)][\ell(x) - \beta] d\mu(x \mid N)$$

$$\text{(1.10)}$$

$$+ \int_{A_0} [g(x,\beta,r) - h(x)] d\mu(x \mid SN)$$

The first integrand is a product of two factors. When the likelihood

ratio is greater than β, the second factor will be positive. Since g

is one for these observations, the first factor will be zero or positive.

For those observations with likelihood ratio less than β, the second

factor will be negative. Since g is zero for these observations, the

first factor will be zero or negative.

$$\ell(x) > \beta, \qquad [g-h][\ell-\beta] = [1-h][\ell-\beta] \geq 0$$

$$\ell(x) = \beta, \qquad [g-h][\ell-\beta] = [r-h][0] = 0 \qquad \text{(1.11)}$$

$$\ell(x) < \beta, \qquad [g-h][\ell-\beta] = [0-h][\ell-\beta] \geq 0$$

Since the integrand of the first integral is never negative, the integral

will also be nonnegative. We now turn our attention to the second

integral, over the set of N-measure zero. The likelihood ratio is

infinite for these observations. Therefore, g is one and the

integrand is nonnegative.

$$x \in A_0, \qquad [w-h] = [1-h] \geq 0 \qquad \text{(1.12)}$$

We have therefore determined that a quantity Q_β is nonnegative for the decision device based on likelihood ratio whose cut level is the same number β. How does this establish that the decision based on likelihood ratio is better ? Pick any particular decision device. Compare it with a decision device based on likelihood ratio which has the same probability of false alarm. Proof that we can find this decision device based on likelihood ratio with matching false alarm probability is given in Appendix A. Granting that we can find such a matching device, it will have some value of β and r. The quantity Q_β evaluated for that specific value of β will be nonnegative. Since the false alarm probabilities are equal, the quantity Q_β is simply the difference between the two detection probabilities. Q_β being nonnegative means that the detection probabilities for the decision device based on likelihood ratio is greater than or equal to the detection probability for the decision device h.

$$P("A_g" | N) = P("A_h" | N)$$

$$(1.13)$$

$$Q_\beta \geq 0 \implies P("A_g" | SN) \geq P("A_h" | SN)$$

This establishes that the performance of the decision based on likelihood ratio falls on the optimum ROC curve for this ROC region.

The "onto" part of the proof establishes that for every possible probability of false alarm there is at least one decision

device based on likelihood ratio which yields this probability of false alarm. This part of the proof has been omitted here and given in Appendix A because it is tedious and not illuminating. The first part of the proof has been displayed to emphasize that the optimum ROC curve, the upper boundary, is directly related to likelihood ratio.

1.2 Fundamental Theorem

1.2.1 Likelihood Ratio.

A decision axis can be considered as an observable. Whatever the physical nature of the original observation and the relationship of the decision axis to these original observations, there will be certain probabilities of occurrence of the various decision values under each of the two classes of causes. The question immediately arises, "If we have two measures induced on this decision axis, what is the Radon-Nicodym relationship between these two measures?" That is, "What is the likelihood ratio of the decision value?"

We will consider the special decision axis which is the likelihood ratio of the observation, and will establish that the likelihood ratio of this decision axis is numerically the same as the value of the decision axis. That is, the likelihood ratio of the likelihood ratio is the likelihood ratio. For example, assume an observation space and a device which numerically evaluates the likelihood ratio of each observation. Let us further assume that for a given observation the value calculated by this device is the number 2.74. We will establish in this section that in this case the likelihood ratio of the number 2.74

is 2.74. That is, the numerical value 2.74 occurs under the cause

SN with probability 2.74 times the probability it will occur under

the cause N.

Let us consider an observation space, X, and a collection of measurable sets A which form a σ-ring $\mathscr{A}$ on X. A probability function is any nonnegative completely additive set function on $\mathscr{A}$, such that the probability of the entire observation space is unity. To introduce likelihood ratio we consider two such probability functions defined on the same σ-ring $\mathscr{A}$, and compare probabilities. Repeating the Radon-Nikodym Theorem (Ref.2),

If $\mu(A|SN)$ is a completely additive set function on a σ-ring of sets which are measurable with respect to $\mu(A|N)$, then there exists a unique decomposition consisting of the following:

A_0, a set of μ_N measure zero; $\mu(A_0|N) = 0$

f, a function integrable with respect to $\mu(x|N)$, unique almost everywhere

Q, a set function defined on $\mathscr{A} \cap A_0$

such that

$$\mu(A|SN) = \int_A f(x)\, d\mu(x|N) + Q(A \cap A_0) \qquad (1.14)$$

Here the Radon-Nikodym derivative relates $\mu(\cdot|SN)$ to $\mu(\cdot|N)$ except for a set of N-measure zero.

Using the above formulation define a likelihood ratio

$\ell(x)$ as

$$\ell(x) = f(x) \quad x \notin A_0$$

$$= \infty \quad x \in A_0 \qquad (1.15)$$

The definition of likelihood ratio is unique, up to a set of combined measure zero. That is, many functions and sets A_0 may satisfy the definition; however, for any two likelihood ratios let A_D be any measurable set on which they differ. Then

$$\mu(A_D | SN) + \mu(A_D | N) = 0 \qquad (1.16)$$

Corollary

$$P(A | SN) = \int_{A-A_0} \ell(x) d\mu(x | N) + Q(A \cap A_0)$$

$$P(A | N) = \int_{A-A_0} d\mu(x | N)$$

$$\left. \right\} \qquad (1.17)$$

Let us consider the implications of this definition to the familiar situations of probability density functions and point-mass probabilities.

For example, consider an n-dimensional space. Let $x = (x_1, x_2, \ldots, x_n)$ and let there be density functions. Then for any Borel set A

$$P(A | SN) = \int_A f(x | SN) dx \qquad P(A | N) = \int_A f(x | N) dx \qquad (1.18)$$

Let A_0 be those points for which the N density vanishes. Then for any Borel set

$$P(A \mid SN) = P(A - A_0 \mid SN) + P(A_0 \cap A \mid SN) \qquad (1.19)$$

$$P(A \mid SN) = \int_{A-A_0} f(x \mid SN)\,dx + P(A_0 \cap A \mid SN) \qquad (1.20)$$

Multiplying and dividing by $f(x \mid N)$

$$P(A \mid SN) = \int_{A-A_0} \left[\frac{f(x \mid SN)}{f(x \mid N)}\right] f(x \mid N)\,dx + P(A_0 \cap A \mid SN)$$

while

$$P(A \mid N) = \int_{A-A_0} f(x \mid N)\,dx$$

$$\left.\right\} \qquad (1.21)$$

Matching terms with Eq. 1.17

$$\ell(x) = \frac{f(x \mid SN)}{f(x \mid N)} \qquad \text{if } f(x \mid N) \neq 0$$
$$= \infty \qquad \text{if } f(x \mid N) = 0 \qquad (1.22)$$

Next consider a point space $\left\{x_i\right\}$, and A be any subset of these points. Then

$$P(A \mid SN) = \sum_{x_i \in A} P(x_i \mid SN) \quad, \quad P(A \mid N) = \sum_{x_i \in A} P(x_i \mid N) \qquad (1.23)$$

Let A_0 be those points having N-probability zero, then

$$P(A \mid SN) = \sum_{x_i \in A-A_0} P(x_i \mid SN) + P(A \cap A_0 \mid SN) \qquad (1.24)$$

$$P(A|SN) = \sum_{x_i \in A-A_0} \left[\frac{P(x_i|SN)}{P(x_i|N)}\right] P(x_i|N) + P(A \cap A_0|SN)$$

while

$$P(A|N) = \sum_{x_i \in A-A_0} P(x_i|N)$$

$$(1.25)$$

Matching terms with Eq. 1.17

$$\ell(x_i) = \frac{P(x_i|SN)}{P(x_i|N)} \qquad \text{if } P(x_i|N) \neq 0$$

$$= \infty \qquad \text{if } P(x_i|N) = 0$$

$$(1.26)$$

We shall not continue with the special cases by treating spaces with distribution functions, leading to a mixture of densities and point-mass probabilities. The purely mechanical difficulties involved in notation make the general Radon-Nikodym formulation attractive.

Before proving the basic theorem that "the likelihood ratio of the likelihood ratio is the likelihood ratio," let us set up the proper framework of "induced measures."

1.2.2 Induced Measures. Paraphrasing from Pitt, (Ref. 2, pp. 25-27) let α be any function mapping X onto Y, not necessarily one-to-one, $\mathcal{A}$ a σ-ring of sets in X and μ a measure on it. We will say a set B in Y is a simple set if $\alpha^{-1}(B)$ is in the σ-ring $\mathcal{A}$. Obviously, the collection of all such sets B form a σ-ring $\mathcal{B}$. The function ν is

called "the measure induced in Y by μ and α" if

$$B \epsilon \mathscr{B}, \quad \nu(B) = \mu\left\{\alpha^{-1}(B)\right\} \qquad (1.27)$$

Integration of a function in Y can be reflected back to the original space X. Quoting Pitt, Theorem 31:

<u>Theorem 31:</u> Suppose μ is a measure in X and ν is the measure induced in Y by $y = \alpha(x)$. Then for functions of y

$$\int_Y f(y) \, d\nu = \int_X f(\alpha(x)) \, d\mu \qquad (1.28)$$

in the sense that if one integral exists, then so does the other and the two are equal.

To say that "B is a ν-measurable set" is equivalent to saying that the characteristic function $1_B(y)$ is integrable with respect to ν.

$$1_B(y) = 1 \quad y \epsilon B$$
$$= 0 \quad y \notin B$$

$$\nu(B) = \int_Y 1_B(y) \, d\nu(y) = \int_B d\nu \qquad (1.30)$$

Since the product of integrable functions is also integrable, for any integrable f of Theorem 31, 1_B times f is also integrable. We therefore have the corollary:

<u>Corollary to Pitt's Theorem 31:</u> Suppose μ is a measure in X and ν is the measure induced in Y by $y = \alpha(x)$; then for any measurable set $B \epsilon \mathscr{B}$

$$\left.\begin{array}{l} \int_{B} f(y)\,d\nu = \int_{\alpha^{-1}(B)} f(\alpha(x)\,)\,d\mu \\[20pt] \text{if either} \quad \int_{Y} f(y)d\nu \quad \text{or} \quad \int_{X} f(\alpha(x)\,)d\mu \ \text{exist}. \end{array}\right\} \qquad (1.31)$$

Let us diagram our current status. We have assumed an observation space X, a σ-ring $\mathscr{A}$ of measurable sets, two measures $\mu(x \mid N)$ and $\mu(x \mid SN)$. For any map we can consider the induced measures $\nu(r \mid N)$ and $\nu(r \mid SN)$. If the map is the likelihood ratio $\ell(x)$, the image of X is some portion of the reals, together with ∞.

$$\ell: X \xrightarrow{\ \text{onto}\ } R \subseteq E' \cup \{\infty\}$$

$$\ell^{-1}: \ell^{-1}(B) \xleftarrow{\hspace{2cm}} B \ \text{measurable}$$
$$\text{measurable}$$

$$\ell: \mu(x \mid N) \longrightarrow \nu(r \mid N)$$

$$\ell: \mu(x \mid SN) \longrightarrow \nu(r \mid SN)$$

Right away let us single out $r = \infty$ and the special set A_0 and list their properties.

$$\ell: A_0 \longrightarrow \infty \ , \ \ell^{-1}(\,\{\infty\}\,) = A_0$$

$$\mu(A_0 \mid N) = 0 \implies \nu(\infty \mid N) = 0$$

$$\mu(A_0 \mid SN) \ \text{exists} \implies \nu_1(\infty \mid SN) \ \text{is some number}.$$

For any measurable B of reals

$$\ell^{-1}(B) \cap A_0 = \emptyset \text{ if } \infty \notin B$$

$$\ell^{-1}(B) \cap A_0 = A_0 \text{ if } \infty \in B$$

We now have a space R with σ-ring $\mathscr{B}$ of measurable sets, and two measures $\nu(r \mid N)$ and $\nu(r \mid SN)$. The question arises, What is the Radon-Nikodym derivative of $\nu(r \mid SN)$ with respect to $\nu(r \mid N)$? What is the likelihood ratio on R space? To answer this we consider the identify map $i(r)$ which maps each point into itself.

$$X \xrightarrow{\ \ell\ } R \xrightarrow{\ i\ } R$$

By definition the space X has probability one under both ν measures.

$$1 = \mu(X \mid SN) = \int_{X-A_0} \ell(x) d\mu(x \mid N) + \mu(A_0 \mid SN) \tag{1.32}$$

Because ℓ is integrable with respect to μ, for any set B measurable in R space

$$\int_{\ell^{-1}(B) - A_0} \ell(x) d\mu(x \mid N) \tag{1.33}$$

exists.

If we now consider the identity map of R onto itself, for any measurable B

$$\int_{B - \{\infty\}} i(r) d\nu(r \mid N) = \int_{\ell^{-1}(B - \{\infty\})} i(\ell(x)) d\mu(x \mid N) \tag{1.34}$$

Because $i(\ell) = \ell$

$$\int_{B-\infty} i(r)d\nu(r|N) = \int_{\ell^{-1}(B)-A_0} \ell(x)d\mu(x|N), \text{ which exists} \qquad (1.35)$$

$$= \mu(\ell^{-1}(B) - A_0 |SN) \qquad \text{by def. of } \ell \qquad (1.36)$$

$$= \nu(B-\{\infty\}|SN) \quad \text{by def. of } \nu \text{ being an} \qquad (1.37)$$
$$\text{induced measure}$$

Therefore we may write

$$\nu(B|SN) = \int_{B-\{\infty\}} i(r)d\nu(r|N) + \nu(B \cap \infty |SN) \qquad (1.38)$$

By direct identification $i(r)$ is a likelihood ratio. We have therefore

shown the following theorem:

Theorem: A likelihood ratio of r (which is a likelihood ratio of x) is

r itself.

We complete this formal abstract work with one further con-

sideration. Consider any one-to-one transformation of R which maps

infinity onto infinity. This means take any function

$$w = u(r) \qquad (1.39)$$

and its inverse

$$r = v(w) \qquad (1.40)$$

such that

$$u(\infty) = \infty, \ v(\infty) = \infty \qquad (1.41)$$

Of course

$$W = u(R) \qquad (1.42)$$

Let the induced measures on W be $\lambda(w|N)$ and $\lambda(w|SN)$. The λ-measurable sets $\mathscr{C}$ are the image of $\mathscr{B}$

$$\mathscr{C} = u(\mathscr{B}) \qquad (1.43)$$

For any measurable $C \epsilon \mathscr{C}$, let $B = v(C)$. Then

$$\lambda(C|SN) = \lambda(C - \{\infty\}|SN) + \lambda(C^{\cap}\{\infty\}|SN) \qquad (1.44)$$

$$\lambda(C|SN) = \nu(B - \{\infty\}|SN) + \lambda(C^{\cap}\{\infty\}|SN) \qquad (1.45)$$

$$\lambda(C|SN) = \int_{B-\{\infty\}} r \, d\nu(r|N) + \lambda(C^{\cap}\{\infty\}|SN) \qquad (1.46)$$

By the corollary to Pitt's Theorem 31

$$\int_{B-\{\infty\}} r \, d\nu(r|N) = \int_{C-\{\infty\}} v(w) \, d\lambda(w|N) \qquad (1.47)$$

So

$$\lambda(C|SN) = \int_{C-\{\infty\}} v(w) \, d\lambda(w|N) + \lambda(C^{\cap}\{\infty\}|SN) \qquad (1.48)$$

We see immediately that $v(w)$ is a likelihood ratio.

Example: If $w = \ln r$, then $r = e^w$. If r is a likelihood ratio, then e^w is a likelihood ratio for $w = \ln r$.

1. 2. 3 __Section Summary__. What has been considered is diagrammed in Fig. 1. 6. From any observation space **X** with measures μ and special set A_0 there is a likelihood ratio $\ell(x)$ such that

$$\mu(A|SN) = \int_{A-A_0} \ell(x)d\mu(x|N) + \mu(A^{\cap}A_0|SN) \qquad (1.49)$$

The induced measures on $R = \ell(X)$ are ν, so related that their likelihood ratio is simply the value of r itself.

$$\nu(B|SN) = \int_{B-\{\infty\}} rd\nu(r|N) + \nu(B^{\cap}\{\infty\}|SN) \qquad (1.50)$$

For any one-to-one transformation which leaves ∞ fixed, say

$$r = v(w)$$

$$\lambda(C|SN) = \int_{C-\{\infty\}} v(w)d\lambda(w|N) + \lambda(C^{\cap}\{\infty\}|SN) \qquad (1.51)$$

The likelihood ratio of w is simply the "substitution value" $v(w)$. In order to establish the proofs, care has been taken to avoid notation which is "suggestive." With the proofs complete we may now make use of more suggestive notation to summarize the two conclusions.

$$\ell(\ell) = \ell \qquad (1.52)$$

$$\ell(w(\ell)) = w^{-1}(w) \qquad (1.53)$$

and in particular if $z = \ln \ell$

$$r = \ell(x) \longrightarrow \qquad w = w(r) \longrightarrow$$
$$\longleftarrow r = v(w)$$

B

A

1-1 $\longrightarrow$ B $\longleftarrow$ 1-1 $\longrightarrow$ C

A_O

. ∞ $\longleftarrow$ $\longrightarrow$. ∞

X $\qquad\qquad$ R $\qquad\qquad$ W

$\mu(x \mid N), \ \mu(x \mid SN), \ \ell(x) \qquad \nu(r \mid N), \ \nu(r \mid SN), \ i(r) \qquad \lambda(w \mid N), \ \lambda(w \mid SN), \ v(w)$
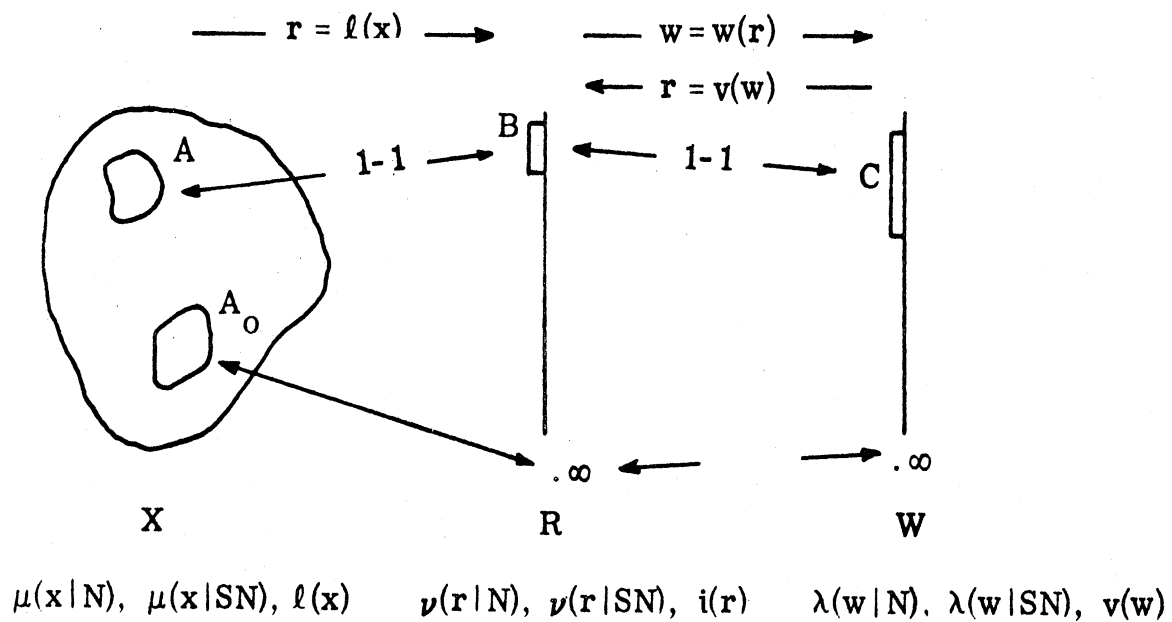
Fig. 1.6. Sketch of spaces and transformations

$$\ell(z) \qquad e^{z} \qquad\qquad (1.54)$$

When the measures $\nu(\ell \,|N)$, $\nu(\ell \,|SN)$ are given by distribution functions, we may write

$$F(\ell \,|SN) = \int_{0}^{\beta} \ell \, dF(\ell \,|N) \qquad\qquad (1.55)$$

This special case was given in Ref. 1 as Theorem 8, and was written

$$dF(\beta|SN) = \beta \, dF(\beta \,|N) \qquad\qquad (1.56)$$

1.3 Initiation of Present Study

In 1954 W. W. Peterson (Ref. 1, pp. 205-206) showed the rather amazing theorem that if the logarithm of the likelihood ratio is normally distributed under condition N, then it is also normally distributed under SN. Specifically, he started with the assumption $z = \ln \ell$ has a normal probability density function under N.

$$f(z\,|N) = \frac{1}{\sqrt{2\pi a}} \; e^{-\frac{(z-m)^{2}}{2d}} \qquad\qquad (1.57)$$

He was first able to show that the mean is related to the variance,

$$f(z\,|N) \qquad \frac{1}{\sqrt{2\pi a}} \; e^{-\frac{(z+.5d)^{2}}{2d}} \qquad\qquad (1.58)$$

and finally that the SN probability density of the logarithm of the likelihood ratio is

$$f(z \mid SN) = \frac{1}{\sqrt{2\pi d}} \, e^{- \frac{(z - .5d)^2}{2d}} \qquad (1.59)$$

Equations 1.58 and 1.59 may be expanded to show that the likelihood ratio of z is indeed e^z.

$$\ell(z) = \frac{f(z \, SN)}{f(z \, N)} = e^z \qquad (1.60)$$

This theorem of Peterson established the basis for the one parameter family of ROC curves, the normal ROC curves. The curves are indexed by the single real parameter, d, or as it is done in the psychophysical literature by the single parameter $d' = \sqrt{d}$. This one parameter family of ROC curves has been used extensively in both the electrical engineering and psychophysical work of detection.

The doubly truncated Halsted distribution is a five parameter class of probability density functions. The untruncated distribution has been used in the study of rapidly fading signals (Ref. 3). It arose again in a study of signal detection and learning in which the amplitude of the signal was initially unknown (Ref. 4). When it is used as the N-probability density function for the logarithm of the likelihood ratio, it is

$$f(z \mid N) = F_0 \, z^B \, e^{-Cz} \, e^{- \frac{z^2}{2D}} \qquad A < z < E \qquad (1.61)$$

It was discovered that the SN-probability density function for the

logarithm of the likelihood ratio fell in the same class

$$f(z|SN) = F_1 z^B e^{-(C-1)z} e^{-\frac{z^2}{2D}} \qquad A < z < E \qquad (1.62)$$

Since this class of functions contains as subclasses the exponential,

the normal, the Rayleigh, the chi-square, and the chi distributions

as special cases, it was evident that many theorems like Peterson's

normal theorem could be obtained from this relation. The key to all

these theorems is Eq. 1.54, that the likelihood ratio of the logarithm

of the likelihood ratio is e^z.

CHAPTER II

ROC CHARACTER

2.1 ROC Curve

In the introduction a simple decision process was charac-
terized as one which assigns a probability $p(\text{"A"} \mid x)$ to each obser-
vation x, denoting the probability that this decision mechanism will
give the response "A" whenever x is observed. The performance
of such a decision device in a specific observation situation plots as
a signal point on the ROC plot. The mechanism for decisions based
upon likelihood ratio considered a probability function $g(x, \beta, r)$
which yielded an entire ROC curve. The parameters, β and r ,
added sufficient dimensionality to the simple decision function to
produce this complete curve. In general, decision functions are
parameterized by some index which allows more than one performance
point. If the dimensionality of the index set is too small, the decision
mechanism may generate only a set of points, such as shown in Fig. 2.1(a),
or it may generate a continuum of points leaving some gaps in the curves,
or possibly multiple-values for y as a function of x, as shown in
Fig. 2.1(c).

These differences between the types of ROC curve given
by decisions based upon likelihood ratio and decisions which give incom-
plete or multiple-valued plots on the ROC curve may be removed by a
process of convex completion. The graphic form of convex completion
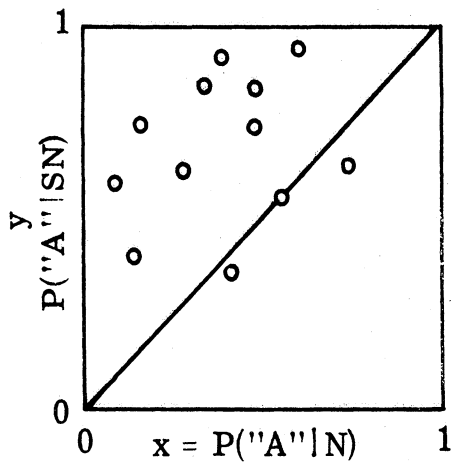
(a) Original arcs
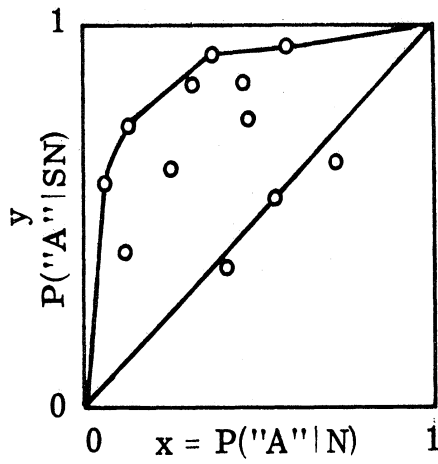
(b) Convex completion of (a)
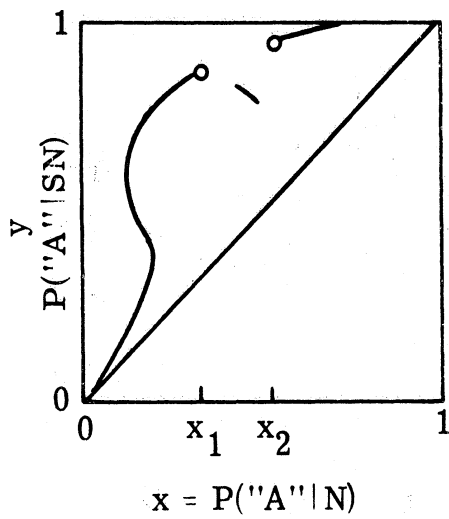
(c) Original arcs
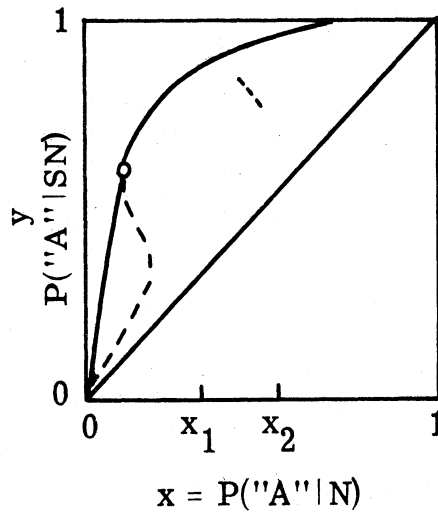
(d) Convex completion of (c)

Fig. 2.1. Illustration of complex completion of ROC
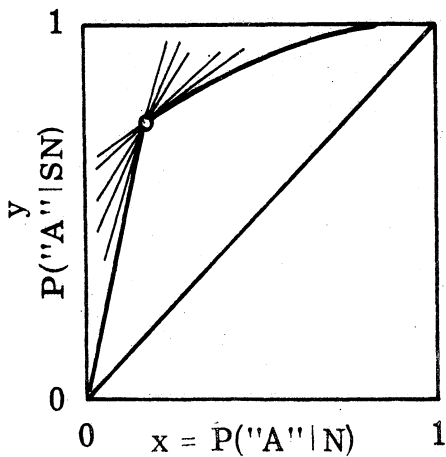
Fig. 2.2. Illustration of tangent cone at one point
of an ROC curve

was to add to the ROC region the line segment joining any two possible

points available from a given decision mechanism. By such a process,

a single point not on the chance diagonal yields a whole parallelogram of

points surrounding the chance diagonal.

Figure 2.1 illustrates two cases of convex completion.

The first, illustrated in Figs. 2.1 (a) and 2.1 (b), show the results of

convex completion for a finite set of points. In Fig. 2.1 (b), only the

upper bound of the ROC region has been filled in. This is sufficient,

since the lower bound of the ROC region lies below the chance diagonal

and could be obtained by symmetry through the midpoint if desired. This

upper bound lies on or above all of the internal secants forming the com-

plex completion between any two possible points. In Fig. 2.1 (c) a

somewhat different case is involved. Between the points with false alarm

probabilities x_1 and x_2, there is a gap which is filled in with the convex

completion of those two points. Since the result is better (that is, above

and to the left) than the short segment of arc in the interior of the region,

this short segment disappears from consideration. Secondly, the left-

hand side of the original arc shows a doubling back of the arc, making

the function $y(x)$ triple valued on a region of false alarms. A straight

line can be drawn from the point (0, 0) tangent to the original curve in

the triple valued section, that exceeds all of the original curve up to the

point of tangency. The resultant region shown in Fig. 2.1 (d) exhibits a

smooth upper bound, the ROC curve.

The ROC curve lies on or above any internal secants

since it is the upper bound of all convex completions. A convex ROC curve also lies on or below any tangents or tangent cones drawn to points on the ROC curve. For a tangent to one point on the ROC curve to contact another point of the ROC curve would require that the convex completion of those two points would fall above the ROC curve. Since the ROC curve is on or above all internal secants, this situation cannot occur. Figure 2.2 illustrates a tangent cone drawn at one point of an ROC curve. At that point the curve is continuous but has a break in slope, so that there is no unique tangent line to that point. The tangent cone consists of all of those lines contacting the ROC curve at only that one point, otherwise lying everywhere above it.

The upper bound of the convex completion of given ROC points is the ROC curve for those points or segments. Such an ROC curve is characterized by being complete and convex. The mathematical adjective "complete" means that for each value of the horizontal axis there is one value for the vertical axis. That is, the function $y(x)$ is a well-defined single value function for x between zero and one.

Definition

Complete: $y(x)$ has a single value for each

$$x \in [0, 1] \tag{2.1}$$

Convex means that the values of the curve lie on or above the internal secants.

Definition

$$\text{Convex:} \quad x_1 \leq x_2 \leq x_3 \, , \, 0 \leq \alpha \leq 1$$

$$\tag{2.2}$$

$$y(x_2) \geq \alpha y(x_1) + (1 - \alpha) y(x_3)$$

The term "singular" has been used in the detection literature to mean perfect performance. In terms of the ROC curve this means that the probability of detection is unity for all probabilities of false alarm. This is illustrated in Fig. 2.3(a).

| (a) Singular | (b) Nonsingular | (c) Regular |
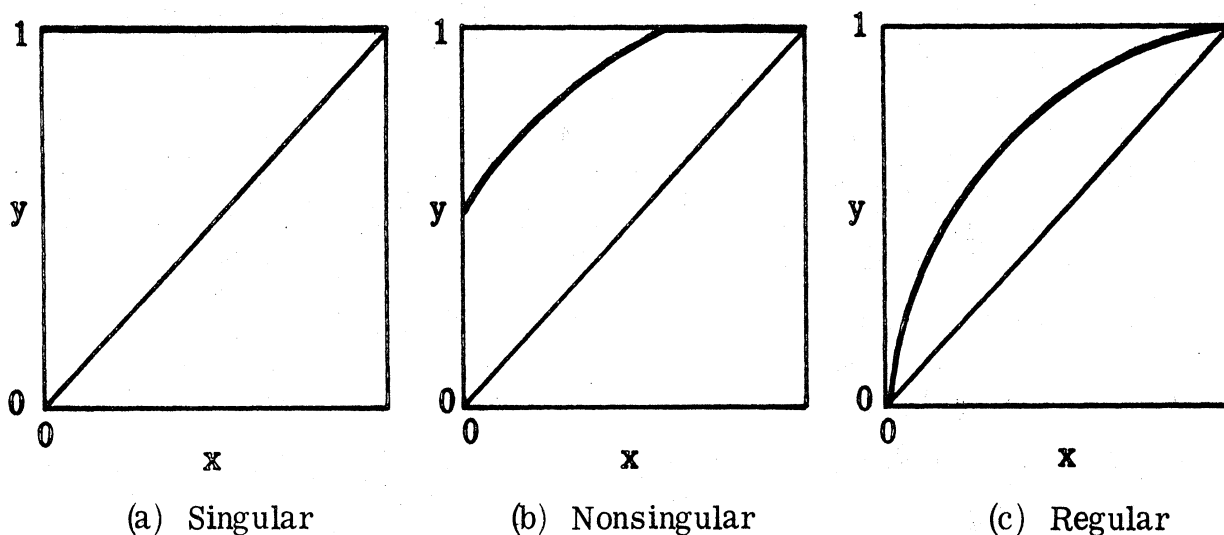
Fig. 2.3. Illustrating singularity for complete convex ROC curves

Definition

$$\text{Singular:} \quad y(x) = 1 \quad \text{for all} \quad x \in [0, 1] \tag{2.3}$$

Once it has been established that a given situation leads to singular detection there is no need for further description of its performance. A curve may be nonsingular and still reflect a certain degree of perfect detection. That is, there may be some observations yielding a positive probability of detection and a zero probability of false

alarm. Correspondingly there may be a set of observations for which one can unequivocably respond "B, " for which the probability of a miss is zero. The ROC for the hypothetical situation containing both of these possibilities is shown in Fig. 2. 3(b). Such a curve will be considered complete, convex, nonsingular, and nonregular. A regular ROC curve is complete, convex, and interior to the unit square except at the chance points (0, 0) and 1, 1).

At this point the precise statement of the type of ROC curves that will be considered can be made. This work will deal only with complete, convex, nonsingular ROC curves, or with nonsingular ROC curves that may be made complete and convex by the process of convex completion.

Definition[1]

$$\text{Regular: complete, convex, } x = 0 \Rightarrow y = 0$$
$$y = 1 \Rightarrow x = 1 \tag{2.4}$$

2. 2 ROC Models

In this section two propositions are proved. The first proposition is that an ROC curve contains insufficient information to specify the observation statistics that led to it, or even to specify the statistics on a real decision axis that lead to it. The second proposition is that an ROC curve does contain sufficient information to specify the statistics on the logarithm of the likelihood ratio.

2. 2. 1 One ROC Curve, Many Models. The purpose of this section is to demonstrate that a single ROC curve contains insuffi-

_____

[1] $\Rightarrow$ is read "implies. "

cient information to specify the observation statistics or decision axis from which it was obtained. Rather, there are many decision axes and observation statistics leading to the same ROC curve.

To begin the demonstration, start with a real variable, ξ, and choose a distribution function, $F_1(\xi)$, fairly arbitrarily. We shall also assume that a regular ROC curve, $Y(X)$, has been selected. The random variable ξ will be used as a decision axis, and $F_1(\xi)$ will be used as the N distribution function. The SN distribution function for ξ will be chosen to obtain an ROC curve identical to the given ROC curve.
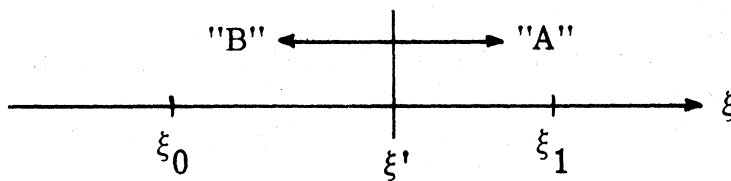
Fig. 2. 4. Arbitrary decision axis, ξ

To say that the ξ-axis is a decision axis means that whenever a particular value of ξ occurs which falls above a cut level, say ξ, the decision will be "A. " In contrast, when it falls below ξ', the decision will be "B. " When the observation falls right at the value

ξ ' a second (randomizing) random variable is introduced to determine how often the decision "A" is elicited. Since the purpose of this development is to show that there are many possible ways to set up observation statistics on ξ, let us introduce some convenient restrictions on these observation statistics. To avoid limiting considerations, it is assumed that the probable values of ξ fall in the open interval between two numbers ξ_0 and ξ_1. Therefore, the distribution function is zero below ξ_0 and is equal to one at and above ξ_1. It is also convenient to assume that the distribution function is strictly monotone increasing in the interval between ξ_0 and ξ_1. This guarantees that if a particular value of F_1 between zero and one is assumed for some ξ value that we can uniquely solve for this ξ value.

$$F_1(\xi) = 0 \qquad\qquad\qquad \xi \leq \xi_0$$

$$F_1(\xi) \text{ is strictly monotone increasing } \xi_0 < \xi < \xi_1 \qquad (2.5)$$

$$F_1(\xi) = 1 \qquad\qquad\qquad \xi_1 \leq \xi$$

With these minor restrictions established, let us set the N distribution equal to F_1.

$$F(\xi \mid N) = F_1(\xi) \qquad\qquad (2.6)$$

It will be convenient to have a special symbol for one minus the distribution.

$$X(\xi) = 1 - F_1(\xi) \tag{2.7}$$

X at the largest value of ξ, ξ_1, is zero. X at the smallest value of ξ, ξ_0, is one. X is strictly monotone decreasing as a function of ξ on the interval X between zero and one.

$$X(\xi_1) = 0 \qquad X(\xi_0) = 1$$
$$X \downarrow \xi \qquad \text{on} \quad 0 < X < 1 \tag{2.8}$$

(The notation $X \downarrow \xi$ means "strictly monotone decreasing.")

The decision axis and the distribution on ξ under N were arbitrary. We now choose the observation statistics for ξ under the condition SN, using the specified regular ROC curve $Y(X)$.

$$F(\xi \mid SN) = 1 - Y[X(\xi)] \qquad \xi \leq \xi_1$$
$$= 1 \qquad\qquad \xi > \xi_1 \tag{2.9}$$

The N Distribution function is certainly a legitimate distribution. Is the SN function as given a legitimate distribution function? For values of the variable less than the minimum, ξ_0, the X value is one and hence the Y value is one, since $Y(1) = 1$. Therefore, the distribution function, under the condition SN, is zero for all small values of the argument ξ.

$$\xi < \xi_0, \quad X(\xi) = 1, \quad Y = 1, \quad F(\xi \,|SN) = 0 \qquad (2.10)$$

For intermediate values of the argument ξ, X is a strictly monotone decreasing function of ξ. Y is a monotone increasing function (not necessarily strictly monotone increasing) function of X. The distribution function under the condition SN is a strictly monotone function of Y. Putting these all together we see that the distribution function for ξ under the condition SN is a monotone function of its argument, ξ.

$$\xi_0 < \xi < \xi_1, \quad X \downarrow \xi, \quad Y \updownarrow X, \quad F(\xi \,|SN) \downarrow y, \quad F(\xi \,|SN) \updownarrow \xi \qquad (2.11)$$

Now we have a decision axis, ξ, and two distribution functions on the decision axis for observation statistics. Let us determine the ROC curve. We will use x_2 to denote the values of the probability of false alarm plotted along the horizontal axis, and y_2 to denote the values of the probability of detection plotted along the vertical axis. The subscript two is used to indicate that this is the second ROC curve; the first one is the original or given ROC curve Y(X).

The probability of false alarm is simply the N-probability that the decision variable will exceed the threshold value ξ', which is simply one minus the distribution function for ξ' under the condition N.

$$x_2(\xi') = Pr(\xi > \xi' \,|N) = 1 - F(\xi' \,|N) \qquad (2.12)$$

The distribution function of ξ' under N condition may be written in terms of $X(\xi)$ by combining Eqs. 2.6 and 2.7. When these are

substituted into Eq. 2.12 we obtain

$$x_2(\xi') = 1 - [1 - X(\xi')]$$

$$x_2(\xi') = X(\xi') \tag{2.13}$$

The probability of detection is derived in a completely analogous manner.

$$y_2(\xi') = Pr(\xi > \xi'|SN) \tag{2.14}$$

$$y_2(\xi') = 1 - [1 - Y(X(\xi'))] = Y[X(\xi')] \tag{2.15}$$

Equation 2.13 can be used to justify replacing X by x_2 to obtain

$$y_2(\xi') = Y[x_2(\xi')] \tag{2.16}$$

The relationship between the vertical and horizontal axes in this second ROC is exactly the same as it was in the original ROC curve; that is, the two ROC curves are identical.

$$y_2(x_2) \equiv Y(X) \tag{2.17}$$

Let us review the above argument. Start with any complete convex ROC curve for which a model is desired. A model is a decision mechanism and observation statistics which yield that given ROC curve. The decision axis must meet only the minor restriction of boundedness. The distribution function for the variable under the condition N

must meet only the minor restriction of monotoneity. (Both restrictions were for convenience and are not strictly necessary.) Once a specific choice of these two has been made, a distribution function under the other condition, SN, can be specified to obtain the given ROC curve. This result has two implications. First, a regular ROC curve and a fixed decision axis together contain insufficient information to specify the observation statistics that led to the given ROC curve. Second, given a regular ROC curve for which a model is desired, one may exercise great freedom in assuming the nature of the decision axis and one distribution on it.

This above demonstration required that the ROC curve be regular. This could have been relaxed from "regular" to "complete, convex, nonsingular." The additional complexity did not seem warranted. In Section 3.5 a functional relation will be established between a complete, convex, nonsingular ROC curve, and an associated regular ROC curve. After that section, one may conclude that each model for the associated regular ROC curve implies a model for the complete convex, nonsingular ROC curve, thereby obtaining many models for the nonregular ROC curve.

A Notation Introduced. Many of the equations in this work will occur in pairs of equations, one for the condition N and the other for the condition SN. It will be convenient to have a notation that signifies both equations when they are nearly alike. The notation used is a slight modification of the ± type of notation commonly used. Its use is shown in the following two equations.

Consider two exponential density functions.

$$f(t \mid SN) = e^t \qquad |t| < .48125$$

$$f(t \mid N) = e^{-t} \qquad |t| < .48125 \tag{2.18}$$

In the double notation these would be written

$$f(t \mid {SN \atop N}) = e^{\pm t} \qquad |t| < .48125 \tag{2.19}$$

When the upper condition is used, the upper sign is used.

Consider the following equation in double notation.

$$\text{Prob } (z < \beta \mid {SN \atop N}) = \int_{-\infty}^{\beta} f(z \mid {SN \atop N}) \, dz \tag{2.20}$$

er condition is used throughout the whole equation, or the lower condition is used throughout the whole equation.

2.2.2 A Unique Model for Each ROC Curve. In the previous section it was shown that there are many models for each regular ROC curve. In this present section the relation between the decision axis and the likelihood ratio will be specified. Under this condition, the specification of the ROC curve will uniquely determine the distributions on the decision axis. The demonstration is complicated enough for regular ROC curves; the inclusion of consideration of the nonregular but complete, convex, nonsingular ROC curves adds no further complexity.

The specific decision axis considered is the z-axis, where z is the logarithm of the likelihood ratio of the observation.

$$z = \ln \ell(X) \tag{2.21}$$

We shall first analyze an ROC curve derived from distributions on the z-axis, then discuss the synthesis of z-axis distributions from an ROC curve.

Any distribution function on a real axis consists of two parts[1]: the jumps in the distribution function corresponding to values that have probability, and the smooth continuous parts of the distribution corresponding to those values that have zero probability individually, but which have a probability density function. For notation use F for distribution function, f for probability density function, $z(i)$ for a point on the decision axis for which the distribution function under N or SN has a discontinuity. The magnitude of these jumps will be denoted by $\omega(z)$.

$$F(z_0 \Big| {SN \atop N}) = \int_{-\infty}^{z_0} f(z \Big| {SN \atop N})\, dz + \sum_{\{z(i) \le z_0\}} \omega\, (z(i) \Big| {SN \atop N}) \quad \text{a. e.} \qquad (2.22)$$

The above equation represents the distribution function for the condition N if the lower conditions are read throughout, and the distribution function for the condition SN if the upper conditions are used throughout. As is true in all general distribution work, we can guarantee the existence of the probability density function almost everywhere, but not everywhere. This is because there may be a finite or countable number

---

[1] Strictly speaking, this may fail to be true at some points. However, the total set of points at which it fails is a set of zero probability. Thus, the statement is said to hold "almost everywhere," written a. e.

of points at which the probability density is not specified without affecting the value of the distribution function.

Concentrate on those values of the decision axis, z, that have probability under either condition N or SN. If the probability under N is positive but the point is a set of SN-measure zero the likelihood ratio will be zero.

$$\omega[z(i)|N] > 0, \quad \omega[z(i)|SN] = 0 \; , \; \ell[z(i)] = 0 \qquad (2.23)$$

If the converse situation holds with the point $z(i)$ having N-measure zero, then the likelihood ratio of that value is infinite.

$$\omega[z(i)|N] = 0, \quad \omega[z(i)|SN] > 0, \quad \ell[z(i)] = \infty \qquad (2.24)$$

In the situation where the point is of positive probability under both conditions N and SN the likelihood ratio is given by the ratio of these two probabilities.

$$\omega[z(i)|N] > 0, \quad \omega[z(i)|SN] > 0, \quad \ell[z(i)] = \frac{\omega[z(i)|SN]}{\omega[z(i)|N]} \qquad (2.25)$$

)roved in Section 1.2, the likelihood ratio of the logarithm of the likelihood ratio is the exponential of the log of the likelihood ratio.

$$\ell(z) = e^{z} \qquad (1.54)$$

Equation 1.54 is a direct result of the fundamental theorem. It is not obtained by inverting Eq. 2.21. The likelihood ratio in Eq. 2.21 is a

function which maps the observation space into the reals. The likelihood ratio given in Eq. 1.54 maps the reals into the reals. This distinction is the reason the fundamental theorem had to be established.

From Eq. 1.54, under the three conditions of Eq. 2.23 - Eq. 2.25, the values of $z(i)$ can be determined exactly at which the likelihood ratio takes on the values zero and infinity. When the likelihood ratio is zero, the decision axis value must be minus infinity, and similarly when the likelihood ratio is infinite, the decision axis value must be plus infinity. All other jumps in the distribution functions must be common to both of the distribution functions, and the ratio between the two magnitudes of the jump will be given by Eq. 1.54.

$$\ell[z(i)] = 0 \implies z(i) = -\infty$$

$$\ell[z(i)] = \infty \implies z(i) = +\infty \tag{2.26}$$

$$z(i) \neq \pm\infty, \quad \omega[z(i)|SN] = e^{z(i)} \omega[z(i)|N]$$

We will retain the form of Eq. 2.22 for the N distribution function but will utilize Eq. 1.54 and 2.26 to write the distribution function under SN.

$$F[z_0|SN] = \int_{-\infty}^{z_0} e^z f(z|N)\, dz + \sum_{\{z(i) \leq z_0\}} e^{z(i)} \omega[z(i)|N] \quad \text{a. e.} \tag{2.27}$$

In Section 2.2.1 the N distribution function could be chosen quite arbitrarily and the SN distribution function chosen to obtain the desired ROC curve. In contrast, the SN distribution function of z is determined once the N distribution function is chosen.

Let us analyze the ROC curve for a given N **distribution**

function for z. Let β stand for a particular cut value **along the z-axis.**

Whenever the value β is one for which the distributions have **jumped,**

representing a z-value with positive probability under both conditions, let

r be a number between zero and one inclusive. The simple decision

based on the z decision axis with parameters β and r will plot as

the point (x, y)

$$x(\beta, r) = \int_{\beta}^{\infty} f(z|N)\, dz + \sum_{\{z(i) > \beta\}} \omega(z(i)|N) + r\omega(\beta|N) \quad \text{a. e.} \quad (2.28)$$

$$y(\beta, r) = \int_{\beta}^{\infty} e^{z} f(z|N)\, dz + \sum_{\{z(i) > \beta\}} e^{z(i)} \omega(z(i)|N) + re^{\beta} \omega(\beta|N) \quad \text{a. e.}$$
$$(2.29)$$

The point moves up along the ROC curve with increasing r and moves

down the ROC curve with decreasing r, in the range of r from zero

to one.
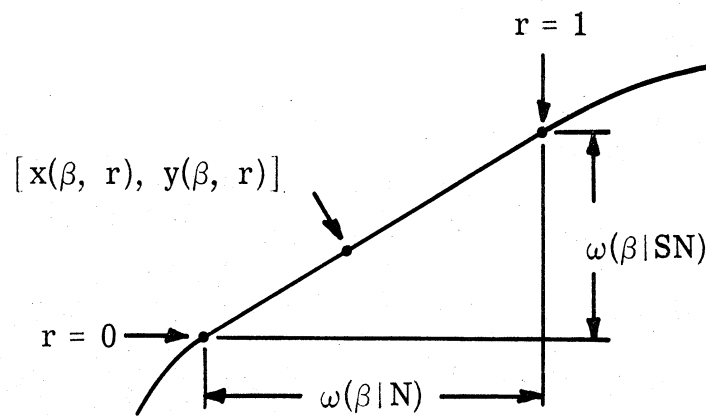
Fig. 2.5 Sketch of arc of ROC corresponding to a z value
β of positive probability

The derivative of the probability of false alarm (Eq. 2. 28) with respect to the parameter r is the magnitude of the jump under N. The derivative of the probability of detection (Eq. 2. 29) is the magnitude of the jump under SN.

$$\beta \in \{z(i)\}, \quad 0 < r < 1$$

$$\frac{dx(\beta, r)}{dr} = \omega(\beta \mid N) \qquad \frac{dy(\beta, r)}{dr} = e^{\beta} \omega(\beta \mid N) \qquad (2.30)$$

The ratio of these two derivatives is the slope of the ROC curve

$$e^{\beta} = \frac{dy}{dx}, \qquad \beta = \ln y'(x) \qquad (2.31)$$

where the prime denotes differentiation with respect to the argument. For any portion of an ROC curve where the slope is constant, the natural logarithm of this slope corresponds to the cut value on the z-axis. At the end points of the linear segment corresponding to the value r = 0 and r = 1 the value of the cut on the z-axis can be determined from the appropriate one-sided derivatives. The r value can also be determined for each point along a linear segment of the ROC curve, since r will increase linearly from the value zero at the left-hand point of the segment to one at the right-hand point of the segment.

$$r = 0 \qquad \qquad \beta = \ln y'(x + 0)$$
$$\qquad \qquad \qquad \qquad \qquad \qquad \qquad (2.32)$$
$$r = 1 \qquad \qquad \beta = \ln y'(x - 0)$$

For a cut at a z value which occurs with probability

zero under both conditions N and SN, the parameter r is irrelevant. A change in ROC point corresponds to a change in the cut value β.

If there are no jumps in a z-neighborhood of a cut β, one may differentiate the coordinates of the ROC curve with respect to β. From Eqs. 2.28 and 2.29, it follows directly that

$$\beta \notin \{z(i)\} \qquad \frac{dx(\beta, r)}{d\beta} = -f(z = \beta | N) \quad a.\, e.$$

$$\frac{dy(\beta, r)}{d\beta} = -e^{\beta} f(z = \beta | N) \quad a.\, e. \tag{2.33}$$

As before, the ratio of these two derivatives relates the value of the decision cut on the z-decision axis to the slope of the ROC curve.

$$e^{\beta} = \frac{dy}{dx}, \quad \beta = \ln y'(x) \quad a.\, e. \tag{2.34}$$

We have shown that when the ROC curve is obtained from the decision axis, which is the logarithm of the likelihood ratio of observation, that the curve is differentiable almost everywhere, and the logarithm of this derivative is numerically equal to the cut value on the decision axis.

Synthesis:

It was really unnecessary to hypothesize that the decision axis was the logartihm of the likelihood ratio of some (physical) observation space, $\mathscr{X}$. It was merely necessary for the variable z to be the logarithm of the likelihood ratio of something and, since it is the logarithm of the likelihood ratio of itself, the internal consistency would

have been sufficient to specify the problem. That particular circle was

avoided because the logic may have been questioned. Since the logic has

now been established by examples here and by the fundamental theorem,

Section 1. 2, we can adopt this seemingly circular logic from now on.

If we are given an ROC curve which is complete and con-

vex, we can hypothesize the existence of a decision axis which is the

logarithm of its own likelihood ratio and determine for each point on

the ROC the corresponding cut value on this decision axis; and, if

necessary, the value of the randomizing parameter r. If the value of

the derivative is constant over some portion of the arc of the ROC, the

horizontal extent of this arc indicates the probability $\omega(z \mid N)$, and if a

value of the derivative has been taken on only once, then Eq. 2. 34

relates the slope to the corresponding cut value and Eq. 2. 33 gives the

probability density function, $f(z \mid N)$.

It has been shown that every complete, convex, non-

singular ROC curve corresponds one-to-one with a particular distribu-

tion function for z under the condition N. It was also shown that the

distribution for z under N completely specified the distribution for

z under SN. Even more can be shown, since Eq. 2. 26 can be inverted

to read

$$\omega[z(i) \mid N] = e^{-z(i)} \omega[z(i) \mid SN] \qquad (2.35)$$

and the similar relation for probability density functions is

$$f(z \mid N) = e^{-z} f(z \mid SN) \qquad (2.36)$$

Therefore the distribution of z under SN completely specifies the distribution of z under N. It has therefore been shown that every complete, convex, nonsingular ROC curve corresponds one-to-one with a particular distribution function for z under the condition SN. Both correspondences relate to the same model; they are different ways of describing a sufficient part of the observation statistics, from which all of the remainder of the model may be obtained.

Another type of generalization is possible. Any class of functions that is in one-to-one correspondence with the distribution of z under N, will be in one-to-one correspondence with the set of complete, convex, and nonsingular ROC curves.

2. 3 ROC Character, $\pi(z)$

The ROC character $\pi(z)$ introduced in this section is the principal function that will be used in the organization of ROC curves into families. The only property of the ROC character and its several related functions discussed in this section is the one-to-one relation with the distribution functions for z.

Definition: ROC character $\pi(z)$.

Whenever the logarithm of the likelihood ratio, z, possesses probability density functions, the ROC character is defined as

$$\pi(z) = \sqrt{f(z \mid SN)\, f(z \mid N)} \qquad |z| < \infty \qquad (2.37)$$

The ROC character can be used to obtain the two probability densities used in its definition. Since the likelihood ratio of z is e^z, that is

$$e^z = \frac{f(z|SN)}{f(z|N)} \qquad (2.38)$$

it follows that

$$e^{+.5z} \pi(z) = \sqrt{\frac{f(z|SN)}{f(z|N)}} \; \sqrt{f(z|SN)\, f(z|N)} = f(z|SN) \qquad (2.39)$$

In a similar manner

$$e^{-.5z} \pi(z) = \sqrt{\frac{f(z|N)}{f(z|SN)}} \; \sqrt{f(z|SN)\, f(z|N)} = f(z|N) \qquad (2.40)$$

These are summarized in the double notation equation

$$e^{\pm.5z} \pi(z) = f(z \left| \begin{matrix} SN \\ N \end{matrix} \right) \qquad (2.41)$$

When the distribution functions for z contain discontinuities, the jump functions (discrete probabilities) are also needed to specify the distributions.

<u>Definition:</u> ROC character jump $\omega_\pi(z)$

Let $\omega(z|SN)$ and $\omega(z|N)$ be the jump functions for the distribution of z. The ROC character jump is defined as

$$\omega_\pi(z) = \sqrt{\omega(z|SN)\, \omega(z|N)} \qquad |z| < \infty \qquad (2.42)$$

Utilizing Eq. 2.26, it follows that

$$e^{\pm.5z} \omega_\pi(z) = \omega(z \left| \begin{matrix} SN \\ N \end{matrix} \right) \qquad (2.43)$$

The ROC character is obtained from probability densities and, as will be formally established in Section 3. 2, has many of the properties of a probability density. The function that is analogous to a probability distribution function is the integrated ROC character $\Pi(z)$.

Definition: Integrated ROC character $\Pi(z)$

$$\Pi(z_0) = \int_{-\infty}^{z_0} \pi(z)\, dz + \sum_{z \leq z_0} \omega_\pi(z) \qquad (2.44)$$

The integrated ROC character will be useful in Stieltjes integration, where it will appear written $d\Pi(z)$. As a memory aid, one may think of Eq. 2. 44 as reading

$$\text{memory aid: } d\Pi(z) = \pi(z)\, dz + \omega_\pi(z) \qquad (2.45)$$

The distribution functions for z can be expanded in density and jump function form.

$$F(z_0 \Big| {}^{SN}_{N}) = \int_{-\infty}^{z_0} f(z \Big| {}^{SN}_{N})\, dz + \sum_{z \leq z_0} \omega(z \Big| {}^{SN}_{N}) \qquad (2.46)$$

Using 2. 41 and 2. 44

$$F(z_0 \Big| {}^{SN}_{N}) = \int_{-\infty}^{z_0} e^{\pm.5z} \pi(z)\, dz + \sum_{z \leq z_0} e^{\pm.5z} \omega_\pi(z) \qquad (2.47)$$

The precise relation of these distribution functions to the integrated ROC character is given by Eq. 2. 48

$$F(z_0 \Big| {SN \atop N}) = \int\limits_{z = -\infty}^{z = z_0} e^{\pm. 5z} \, d\Pi(z) \qquad (2.48)$$

Although almost all of this work will be devoted to the distributions of z (meaning, as always herein, the logarithm of the likelihood ratio), it will be convenient to have a function for arbitrary probability density functions that parallels the ROC character.

Definition: Root product density function $f_\pi(\;)$

If t is a random variable with probability density functions f(t|SN) and f(t|N), then the root product density function for t is defined as

$$f_\pi(t) = \sqrt{f(t|SN) \, f(t|N)} \qquad (2.49)$$

2.4. Remarks on the Distribution of z

This section contains four short remarks about the distribution of z and their correspondingly short proofs. The first three have to do with the expected value of z under the two conditions N and SN.

Remark: $E(z|N) \leq 0$

Proof:

(1) $E(\ell|N) = \int \ell \, dF(\ell|N)$

$= \int \ell(\ell) \, dF(\ell|N) \qquad (2.50)$

$= \int dF(\ell|SN) = 1$

(2) ln is a convex function, and hence lies on or below a

tangent

$$\therefore z \leq z(1) + z'(1)(\ell - 1) = \ell - 1$$

$$E(z|N) \leq E(\ell - 1|N) = 0$$

Remark: $E(z|SN) \geq \ln E(\ell|SN)$ (2.51)

Proof:

ln is a convex function

Remark: $E(z|SN) \geq 0$ (2.52)

Proof:

Consider $g(\ell) = \ell \ln \ell; \quad g(1) = 0$

$$g'(\ell) = 1 + \ln \ell; \, g'(1) = 1$$

$$g''(\ell) = \ell^{-1}; \quad\quad g'' > 0$$

$g(\ell)$ is "convex upward" and lies on or above any tangent

$$\therefore g(\ell) \geq g(1) + g'(1)(\ell - 1)$$

$$\ell \ln \ell \geq \ell - 1$$

$$E(\ell \ln \ell | N) \geq E(\ell - 1 | N) = 0$$

$$E(\ell \ln \ell | N) = E(\ln \ell | SN)$$

$$\therefore E(\ln \ell | SN) \geq 0 \quad\quad \text{Q. E. D.}$$

Whenever there are no jumps in the probability distribution functions for z, the ROC character contains all of the information, and there is no need to use the integrated ROC character. In that case, we can say something about the symmetry of the ROC character.

Remark: $\pi(z)$ is symmetric iff it is symmetric about $z = 0$

Proof:

Assume $\pi(z) = \pi(c - z)$

$$1 = \int_{-\infty}^{\infty} f(z \mid N) \, dz = \int_{-\infty}^{\infty} e^{-.5z} \pi(z) \, dz$$

$$= \int_{-\infty}^{\infty} e^{-.5z} \pi(c-z) \, dz = e^{-.5c} \int_{-\infty}^{\infty} e^{.5(c-z)} \pi(c-z) \, dz$$

$$= e^{-.5c} \int_{-\infty}^{\infty} e^{.5\lambda} \pi(\lambda) \, d\lambda = e^{-.5c} \int_{-\infty}^{\infty} f(z \mid SN) \, dz$$

$$= e^{-.5c}$$

$$\therefore \ c = 0$$

CHAPTER III

ROC FAMILIES

3.1 Philosophical Orientation

The objective of the present work is to organize **ROC** curves into families. To do this, some restriction has to be made on the system of **ROC** curves to be considered. The first intent is to treat as large a system of ROC curves as possible. The second intent is to develop the **ROC** families which have some generality while still requiring that each curve in the family can be specified by only one or two parameters. In this way, the determination of an **ROC** curve can be broken into a two-step process of determining the family, and then determining the one or two parameters necessary to specify the particular curve.

In this work I have been influenced by Karl Pearson, who wished to unify and codify the treatment of probability density functions of one real variable. Restricting himself to the class of unimodal functions he worked with those probability densities described by the range of the random variable and the first four moments. His results were strongly related to the function which has come to be known as the Pearson ratio. For a probability density function $f(t)$, the Pearson ratio is

$$\text{Pearson ratio} = \frac{d}{dt} \left(\ln f(t) \right) = \frac{f'(t)}{f(t)} \qquad (3.1)$$

Whenever the Pearson ratio is a rational function of degree no higher than a monomial over a quadratic, the density function is within the Pearson system. The specific type is based on the roots of the quadratic. The root of the numerator, the monomial, is the position of the mode of the density function. Pearson later expanded this work by the consideration of a quadratic over a cubic, which allowed inclusion of more complicated unimodal functions as well as bimodal functions. The Pearson ratio plays two roles of special interest to this present work. First, the Pearson ratio defines the area of applicability of the Pearson classification system. That is, the mathematician dealing with a given probability density function could decide definitely whether that function falls within the Pearson system or not. Secondly, within the system the ratio provides a classification into families.

In attempting to codify ROC curves the present work is restricted to complete convex nonsingular ROC curves. The formal function that is proposed as a basis for classification is the ROC character.

The classification of families of ROC curves would be a huge undertaking if it were not true that a link can be found that relates this classification to our heritage in probability theory and statistics. In a moment the relation of the ROC character to the Pearson ratio will be determined to see if the Pearson system can be used directly for classifying ROC characters. In the next section,

3.2, the relationship of the ROC character to all univariate proba-

bility density functions is obtained. This second link provides the

basis for the classification of ROC families by ROC characters.

The relationship of the ROC character, $\pi(z)$, to the

Pearson classes is obtained from the relationship of the SN and N

density functions to the ROC character.

$$f\left(z \middle| \begin{matrix} SN \\ N \end{matrix}\right) = e^{\pm .5z} \pi(z) \tag{2.41}$$

Divide through by the exponential

$$\pi(z) = e^{\mp .5z} f\left(z \middle| \begin{matrix} SN \\ N \end{matrix}\right) \tag{3.2}$$

and take natural logarithms.

$$\ln \pi(z) = \ln f\left(z \middle| \begin{matrix} SN \\ N \end{matrix}\right) \mp .5z \tag{3.3}$$

The functions are in convenient form to take the derivative with respect

to the variable, z.

$$\frac{\pi'(z)}{\pi(z)} = \frac{f'\left(z \middle| \begin{matrix} SN \\ N \end{matrix}\right)}{f\left(z \middle| \begin{matrix} SN \\ N \end{matrix}\right)} \mp .5 \tag{3.4}$$

The Pearson ratio for the ROC character differs from the Pearson

ratio for either density function by either plus or minus one-half. It

is not enough, however, that each of these Pearson ratios be a rational

function. If the denominator is of higher order than the numerator, then the inclusion of the added constant in obtaining a new rational function will generally lead to numerator and denominator having the same order. It is quite common to find that the distribution for the logarithm of the likelihood ratio under one of the conditions does fall in a Pearson class, while the distribution of the logarithm of the likelihood ratio under the other condition falls in either a different Pearson class or outside the Pearson system. Only when the denominator is of equal or lower order than the numerator will the two distributions of likelihood ratio and the ROC character stay within the Pearson system. In this special case, all three will be in the same Pearson class.

The Pearson ratio for a normal probability density is

$$\text{Normal Pearson ratio} = \frac{\text{mean} - z}{\text{variance}} \tag{3.5}$$

Section 1.3 contains W. W. Peterson's result, "When the logarithm of the likelihood ratio is normally distributed under one condition, it is also normally distributed under the other condition." We can add to this at this time, by adjoining "and the ROC character will be proportional to a normal probability density function and centered halfway between the two conditional density functions."

3.2 Fundamental Theorem on Regular ROC Characters

In this section it will be shown that ROC characters for regular ROC curves correspond to those probability distributions on the real line which possess moment generating functions. Such distributions have received considerable study in the history of probability theory.

All probability distributions on the real line possess characteristic functions (Fourier transforms), but not all possess moment generating functions (double sided LaPlace transforms).[1] Those distributions that possess both will have finite moments of all orders, and their transforms will be analytic in some neighborhood of the origin.

In order to consider both discrete probability functions and probability density functions, we work with probability distribution functions, denoted by F, and the integrated ROC character, $\Pi(z)$. For discrete random variables the individual probabilities are the jumps in the distribution. For continuous random variables the probability density function is the derivative of the distribution, and correspondingly the ROC character $\pi(z)$ is the derivative of $\Pi(z)$.

From Character to Distribution: As is customary in probability notation the limiting values for large negative and large positive values

---

[1] For example, Ref. 30, p. 12.

are indicated by the arguments $-\infty$ and $+\infty$. The four requirements

on the integrated ROC character for regular ROC curves are:

(1) the lower limit is zero $\qquad \Pi(-\infty) = 0 \qquad$ (3.6)

(2) monotone growth $\qquad d\Pi(z) \geq 0 \qquad$ (3.7)

(3) unit "N" value $\qquad \int e^{-.5z} \, d\Pi(z) = 1 \qquad$ (3.8)

(4) unit "SN" value $\qquad \int e^{+.5z} \, d\Pi(z) = 1 \qquad$ (3.9)

Add the equalities for the third and fourth conditions to obtain

$$\int \left[e^{.5z} + e^{-.5z} \right] d\Pi(z) = 2 \qquad (3.10)$$

Since the sum $(e^{.5z} + e^{-.5z})$ is greater than or equal to 2 for all z

values,

$$\Pi(+\infty) = \int_{-\infty}^{\infty} d\Pi(z) \leq 1 \qquad (3.11)$$

Let c be reciprocal of $\Pi(+\infty)$.

$$c\Pi(+\infty) = 1 \qquad (3.12)$$

This has provided the basis for the distribution function

that is proportional to the integrated ROC character, and the proof

that the corresponding moment generating function exists in some

neighborhood of the origin. Consider the function

$$F(z) = c\Pi(z) \qquad (3.13)$$

From (3.6), (3.7) and (3.13) it follows that

$$F(-\infty) = 0 \qquad F(+\infty) = 1 \qquad (3.14)$$

F is monotone nondecreasing, $dF(z) \geq 0$

Therefore F defined by (3.13) is a distribution function.

The definition of the moment generating function is usually given using the expected value operator, $E( )$.

$$M_F(\xi) = E(e^{\xi t}) \qquad (3.15)$$

This may be evaluated using the distribution function

$$M_F(\xi) = \int_0^1 e^{\xi t} \, dF(t) \qquad (3.16)$$

If the distribution function contains no jumps, the density function may be used.

$$M_F(\xi) = \int_{-\infty}^{\infty} e^{\xi t} f(t) \, dt \qquad (3.17)$$

$M_F(\xi)$ is called the moment generating function because

$$\left. \frac{d^n M_F(\xi)}{d\xi^n} \right|_{\xi=0} = E(t^n) \qquad (3.18)$$

Only those random variables which are sufficiently "concentrated" will possess a moment generating function and finite moments of all orders.

A simple dominance method can be used to show that the moment generating function exists near the origin $\xi = 0$. For any real ξ with absolute value not greater than one-half,

$$M_F(\xi) = \int_{-\infty}^{\infty} e^{\xi z} \, dF(z) \tag{3.19}$$

$$= c \int_{-\infty}^{\infty} e^{\xi z} \, d\Pi(z) \tag{3.20}$$

$$\leq c \int_{-\infty}^{\infty} e^{\xi z} + e^{-\xi z} \, d\Pi(z) \tag{3.21}$$

$$\leq c \int_{-\infty}^{\infty} e^{.5z} + e^{-.5z} \, d\Pi(z) = 2c \tag{3.22}$$

From Distribution to Character. To satisfy the requirements on the integrated ROC character, specifically the unit N and unit SN values listed in Eqs. 3.8 and 3.9, $\Pi(z)$ may differ from a parent distribution by a shift and a scale factor. Start with any probability distribution function, $F(t)$, which possesses a moment generating function, $M_F(\xi)$. Later, values will be specified for the constants a, b, and c, the latter two positive. For the moment consider

$$z = a + bt \tag{3.23}$$

The moment generating function for the random variable z is

$$E[e^{\xi z}] = E[e^{\xi(a+bt)}] = e^{a\xi} M_F(b\xi) \tag{3.24}$$

that is

$$\int e^{\xi z} dF\left(\frac{z-a}{b}\right) = e^{a\xi} M_F(b\xi) \tag{3.25}$$

The specific values of interest are

$$\int e^{\pm.5z} dF\left(\frac{z-a}{b}\right) = e^{\pm.5a} M_F(\pm.5b) \tag{3.26}$$

For any positive value of b which is sufficiently small so that the

function $M_F(\pm.5b)$ exists, set

$$a = \ln M_F(-.5b) - \ln M_F(+.5b) \tag{3.27}$$

$$c = [M_F(-.5b) M_F(+.5b)]^{.5} \tag{3.28}$$

The decimal exponent $.5$, the fractional exponent $1/2$, and the square

root symbol all mean the same thing. The decimal exponent is con-

venient for the type of manipulations in this present work. If the

constants have been chosen correctly, then the following is an integra-

ted ROC character

$$\Pi(z) = c^{-1} F\left(\frac{z-a}{b}\right) \tag{3.29}$$

If the distribution function F is the integral of a probability density function, f, then

$$\pi(z) = (bc)^{-1} f(\frac{z-a}{b}) \qquad (3.30)$$

The final step is to determine that Eqs. 3.8 and 3.9 hold. From (3.26) and (3.27),

$$\int e^{\pm.5z} \, dF(\frac{z-a}{b}) = [M_F(-.5b)]^{\pm.5} [M_F(+.5b)]^{\mp.5} M_F(\pm.5b) \qquad (3.31)$$

$$\int e^{\pm.5z} \, dF(\frac{z-a}{b}) = [M_F(-.5b) M_F(+.5b)]^{.5} = c \qquad (3.32)$$

From (3.29),

$$\int e^{\pm.5z} \, d\Pi(z) = c^{-1} \int e^{\pm.5z} \, dF(\frac{z-a}{b}) \qquad (3.33)$$

From (3.32) and (3.33),

$$\int e^{\pm.5z} \, d\Pi(z) = c^{-1} c = 1 \qquad (3.34)$$

There is a continuous range of b values. Corresponding to each b value are appropriate constants a and c. Using these values, the simple scaling, Eq. 3.23, generates a family of (integrated) ROC characters.

The key relations are repeated for later reference.

Given a distribution function $F(t)$, or a moment generating function $M_F(\xi)$, for any positive value of b sufficiently small

that $M_F(\pm. 5b)$ exists, let

$$a = \ln M_F(-. 5b) - \ln M_F(+. 5b) \tag{3.27}$$

$$c = [M_F(-. 5b) \, M_F(+. 5b)]^{. 5} \tag{3.28}$$

Then

$$\Pi(z) \quad c^{-1} \; F(\frac{z - a}{b}) \tag{3.29}$$

$$\pi(z) \quad (bc)^{-1} \; f(\frac{z - a}{b}) \tag{3.30}$$

are valid integrated ROC character, and ROC character, respectively.

3.3 Classical ROC Characters

The development of a system of ROC families will begin with probability density functions of a real variable which have moment generating functions. Each such probability density function corresponds to some ROC character. If the probability density function is well known and has a universally accepted name, that same name will be used for both the ROC character and the resultant ROC family of curves. Certain classes of probability density function are known by several different names, because they were independently developed in several fields of applications. An example is the normal, Gaussian, or error function. An attempt has been made to relate this present work to each of the names used in the literature.

Armed with a table of probability density functions and

their moment generating functions, one may mechanically generate classes of ROC characters. The ROC curves are found by direct integration. The basic equations for the ROC curve in terms of the z-axis cut value β, and the randomizing parameter r are given by Eqs. 2.28 and 2.29. When β is a point of continuity, or otherwise when r = 0, the ROC curve equations can be written compactly using the form of Eq. 2.48.

$$\frac{Y(\beta)}{X(\beta)} = \int_{\beta}^{\infty} e^{\pm .5z} \, d\Pi(z) \qquad (3.35)$$

The following sections list the results for the major classical probability density functions. Some mechanical comments are in order before proceeding with these.

(1) The fundamental theorem of ROC characters allows for a translation and scaling of the variable.

$$z = a + bt \qquad (3.23)$$

An entire probability density function class may usually be considered by using the normalized form for the density function, having zero mean and unit variance, since Eq. 3.23 rescales the mean and the variance.

(2) Bounded probability density functions on a bounded range will always have moment generating functions. Many of these correspond

to sections of probability density functions with infinite or semi-infinite range. Section 3.4 will treat these separately. The beta and rectangular density functions are not sections of other probability density functions, and are included in the following sections.

(3) The table upon which the following sections are based is Table 26.1 of Ref. 6.

3.3.1 Normal, Gaussian, Error Function. For

$-\infty < t < \infty$

$$f(t) = \frac{1}{\sqrt{2\pi}}\ e^{-\frac{t^2}{2}}\ , \quad M_F(\xi) = e^{\frac{\xi^2}{2}}$$

$$M_F(\pm .5b) = e^{\frac{b^2}{8}} \quad \text{exist for all } b > 0$$

$$a = \frac{b^2}{8} - \frac{b^2}{8} = 0$$

$$c = e^{\frac{b^2}{8}}$$

so

$$\Pi(z) = e^{\frac{b^2}{8}}\ \Phi(\frac{z}{b}) = e^{-\frac{b^2}{8}} \int_{-\infty}^{\frac{z}{b}} \frac{1}{\sqrt{2\pi}}\ e^{-\frac{u^2}{2}}\ du$$

and

$$\pi(z) = e^{-\frac{b^2}{8}} \frac{1}{\sqrt{2\pi}\, b} \; e^{-\frac{z^2}{2b^2}}$$

For standard notation for the normal ROC we let $b^2 = d$

$$\pi(z) = e^{-\frac{d}{8}} \frac{1}{\sqrt{2\pi d}} \; e^{-\frac{z^2}{2d}}$$

The ROC curve is $y = \Phi(t + \sqrt{d})$ when $x = \Phi(t)$. $\Phi(\ )$ is the normal distribution function.

3.3.2 Exponential, Pearson X. For $t > 0$

$$f(t) = e^{-t}$$

$$F(t) = 1 - e^{-t}$$

$$M_F(\xi) = (1-\xi)^{-1}$$

$$M_F(\pm .5b) = (1 \pm .5b)^{-1} \quad \text{exists for } 0 < b < 2$$

$$a = \ln\left(\frac{1 - .5b}{1 + .5b}\right)$$

$$c^2 = (1 - .25b^2)^{-1}$$

$$\Pi(z) = c^{-1}\left[1 - e^{-\frac{z-a}{b}}\right] \qquad z > a$$

$$\pi(z) = \frac{1}{bc} \, e^{\frac{a}{b}} \, e^{-\frac{z}{b}}$$

$$\pi(z) = \frac{\sqrt{1 - .25b^2}}{b} \left(\frac{1 - .5b}{1 + .5b}\right)^{\frac{1}{b}} e^{-\frac{z}{b}} \quad z > \ln\left(\frac{1 - .5b}{1 + .5b}\right)$$

If we let

$$b = 2\frac{A - 1}{A + 1}$$

that is, let

$$A = \frac{1 + .5b}{1 - .5b} \qquad A > 1$$

then

$$\pi(z) = (A-1)^{-1} A^{-\frac{1}{A-1}} e^{-\frac{z}{2}\frac{A+1}{A-1}}, \qquad z > -\ln A$$

or

$$\pi(z) = \frac{A^{.5}}{(A-1)} e^{-\frac{A+1}{A-1}\frac{z-z_0}{2}} \qquad z > z_0 = -\ln A$$

The ROC curve is

$$X = Y^A$$

3.3.3 LaPlace, Double Exponential. For $-\infty < t < \infty$

$$f(t) = \frac{1}{2} e^{-|t|} \qquad M_F(\xi) = (1 - \xi^2)^{-1}$$

$$M_F(\pm .5b) = (1 - .25b^2)^{-1} \qquad \text{exists for } 0 < b < 2$$

$$a = 0$$

$$c = (1 - .25b^2)^{-1}$$

$$\pi(z) = \frac{1}{bc} \frac{1}{2} e^{-\frac{|z|}{b}}$$

$$\pi(z) = \frac{1 - .25b^2}{2b} e^{-\frac{|z|}{b}}$$

$$b = 2 \frac{B-1}{B+1}$$

that is, let

$$B = \frac{2+b}{2-b} \qquad B > 1$$

then

$$\pi(z) = \frac{B}{B^2 - 1} e^{-\frac{|z|}{2} \frac{B+1}{B-1}}$$

The ROC curve is

$$X = \frac{(B+1)^{B-1}}{B^B} Y^B$$

from $(0, 0)$ to the negative diagonal, and

$$Y = 1 - \frac{(B+1)^{B-1}}{B^B} (1-X)^B$$

from the negative diagonal to $(1, 1)$. The negative diagonal point is

$$X = \frac{1}{B+1} \qquad Y = \frac{B}{B+1}$$

3.3.4 Pearson Type III, Gamma, Chi-Square. For

$t > 0, \rho > -1$

$$f(t) = t^{\rho} \, e^{-t} / \Gamma(\rho + 1)$$

$$M_F(\xi) = (1 - \xi)^{-(\rho+1)}$$

$$M_F(\pm .5b) = (1 \pm .5b)^{-(\rho+1)} \qquad \text{exists for } 0 < b < 2$$

$$a = (\rho + 1) \ln\left(\frac{1 - .5b}{1 + .5b}\right)$$

$$c^2 = (1 - .25b)^{-(\rho+1)}$$

so

$$\pi(z) = \frac{(1 - .25b^2)^{\frac{\rho+1}{2}}}{b^{\rho+1} \, \Gamma(\rho+1)} \left[z - (\rho+1) \ln\left(\frac{1 - .5b}{1 + .5b}\right)\right]^{\rho} e^{-\frac{z-a}{b}}, \quad z > a$$

As before, let

$$B = \frac{1 + .5b}{1 - .5b} ; \qquad B > 1$$

$$\pi(z) = \frac{\sqrt{B}}{B-1}^{\rho+1} \frac{1}{\Gamma(\rho+1)} \left(z + (\rho+1) \ln B\right)^{\rho} e^{-\frac{z-a}{2}\frac{B+1}{B-1}},$$

$$z > -(\rho+1) \ln B$$

or

$$\pi(z) = \frac{B^{.5(\rho+1)}}{(B-1)\,\Gamma(\rho+1)} \left(\frac{z-z_0}{B-1}\right)^{\rho} e^{-\frac{B+1}{B-1}\frac{z-z_0}{2}} \quad , \quad z > z_0 = -(\rho+1)\ln B$$

3.3.5 Fisher-Tippett Type I. For $-\infty < t < \infty$

$$F(t) = e^{-e^{-t}} \qquad M_F(\xi) = \Gamma(1-\xi)$$

and

$$f(t) = e^{-t-e^{-t}}$$

$$M_F(\pm.5b) = \Gamma(1\mp.5b) \quad \text{exists for } 0 < b < 2$$

$$a = \ln\Gamma(1+.5b) - \ln\Gamma(1-.5b)$$

$$c^2 = \Gamma(1+.5b)\,\Gamma(1-.5b)$$

$$\pi(z) = \frac{1}{b\,\sqrt{\Gamma(1+.5b)(1-.5b)}} \left[\frac{\Gamma(1+.5b)}{\Gamma(1-.5b)}\right]^{\frac{1}{b}} e^{-\frac{z}{b}\left[\frac{\Gamma(1+.5b)}{\Gamma(1-.5b)}\right]^{\frac{1}{b}}} e^{-\frac{z}{b}}$$

3.3.6 Beta, Pearson I and XII. For $0 < t < 1$;

$m, n \geq 1$

$$f(t) = \frac{1}{B(m,n)}\, t^{m-1}\,(1-t)^{n-1}$$

$$M_F(\xi) = M(m, m+n, \xi)$$

This is Kummers confluent hypergeometric function, and yields a

power series form for all b for positive n, m

$$M_F(\pm . 5b) = \sum_{k=0}^{\infty} (\pm 1)^k \frac{(m+k-1)! \ (m+n-1)!}{(m+n+k-1)! \ (m-1)!} \frac{b^k}{2^k k!}$$

This is so unwieldy that the general form is retained

$$\pi(z) = (bc \ B(m, n))^{-1} (\frac{z-a}{b})^{m-1} (\frac{a+b-z}{b})^{n-1}, \quad a \leq z \leq b$$

where

$$a = \ln M_F(-. 5b) - \ln M_F(+. 5b)$$

$$c = [M_F(-. 5b) \ M_F(+. 5b)]^{.5}$$

For m = 1 or n = 1 the infinite series becomes a finite sum, but no

real simplification results. ROC curve formulae have not been obtained.

3.3.7 Rectangular. For $0 \leq t \leq 1$

$$f(t) = 1$$

$$M_F(\xi) = \xi^{-1}(e^{\xi} - 1)$$

$$M_F(\xi) = (\pm . 5b) = \pm 2b^{-1}(e^{\pm . 5b} - 1)$$

$$a = -. 5b$$

$$bc = (2 \ e^{.25b} - e^{-.25b})$$

If we let

$$b = 2 \ln(1 + R) \qquad R > 1$$

then

$$bc = 2R \Big/ \sqrt{1 + R}$$

$$\pi(z) = \frac{\sqrt{1+R}}{2R} , \qquad z| < \ln(1+R)$$

The ROC curve is a symmetric rectangular hyperbola

$$(1 + Rx)(1 + R - Ry) = 1 + R \qquad R > 1$$

The asymptotes lie outside the ROC unit square at a distance R^{-1}

3.3.8 Geometric. For integer $t \geq 0; 0 < q < 1$

$$\omega(t) = (1 - q) q^t$$

$$M(\xi) = (1 - q)/(1 - qe^\xi)$$

$$M(\pm .5b) = (1 - q)/(1 - qe^{\pm .5b}) , \qquad 0 < b < 2|\ln q|$$

$$a = \ln \left(\frac{1 - qe^{.5b}}{1 - qe^{-.5b}} \right)$$

$$c = (1 - q) \Big/ (1 + q^2 - q(e^{.5b} + e^{-.5b}))^{.5}$$

$$\omega_\pi(z) = c^{-1}(1 - q) q^{\left[\frac{z - a}{b} \right]} , \qquad z = a, a + b, a + 2b, \ldots$$

$$\omega_\pi(z) = \left(1+q^2 - qe^{\cdot 5b} - qe^{-\cdot 5b}\right)^{\cdot 5}\left(\frac{1-qe^{\cdot 5b}}{1-qe^{-\cdot 5b}}\right)^{\left|\frac{\ln q}{b}\right|} e^{-\left|\frac{\ln q}{b}\right|z}$$

write

$$z_0 = \ln\left(\frac{1-qe^{\cdot 5b}}{1-qe^{-\cdot 5b}}\right)$$

and

$$z_k = z_0 + kb, \qquad k = 0,1,2,\ \ldots$$

Let

$$A = \frac{\ln q - .5b}{\ln q + .5b}$$

The ROC curve subscribes $X = Y^A$ and touches it at

$$X_k = \left(qe^{-\cdot 5b}\right)^k, \qquad Y_k = \left(qe^{+\cdot 5b}\right)^k, \qquad k = 0,1,2,\ \ldots$$

3.3.9 Binomial. For integer $t \geq 0;\ 0 < p < 1$

$$\omega(t) = C_t^n\, p^t(1-p)^{n-t}$$

where

$$C_t^n = \text{binomial coefficients} = \frac{n!}{t!(n-t)!}$$

Let

$$u = \ln \frac{p}{1-p}$$

$$\omega(t) = (1 + e^u)^{-n} \; C_t^n \; e^{ut}$$

$$M(\xi) = (1 + e^u)^{-n} \; (e^{u+\xi} + 1)^n$$

$$a = n \ln \frac{e^{u-.5b}+1}{e^{u+.5b}+1} = n \ln \frac{(e^u + e^{.5b})}{(e^{u+.5b}+1)} - .5bn$$

$$c = (1 + e^u)^{-n} \; (e^{2u} + e^{u-.5b} + 1)^{.5n}$$

$$c = (1 + e^u)^{-n} \; 2^n \; e^{.5un}(.5 \cosh u + .5 \cosh .5b)^{.5n}$$

$$\omega_\pi(z) = c^{-1} \; \omega \, (t = \frac{z-a}{b})$$

$$= \left[e^u(.5 \cosh u + .5 \cosh .5b) \right]^{-.5n} 2^{-n} C_t^n \; e^{ut}, \qquad z = a + bt$$

The term $2^{-n}C_t^n$ has a "bell-shaped" graph and sums over t to unity. The symmetric case of $p = .5$, or $u = 0$, yields $a = -.5bn$ and

$$\omega_\pi(z) = [1 + .5 \cosh .5b]^{-.5n} \; 2^{-n} C_t^n, \qquad z = b(t - .5n)$$

The ROC curve consists of $n+1$ straight segments, with slope change of e^b at each vertex. The equation of the ROC must be obtained from the two parameter table of binomial sums.

3.3.10 Specific Binomial.

$n = 0$, one point: from the general binomial

$a = 0$, $c = 1$, and $t = 0$

$\therefore \omega(z) = c^{-1} \omega(t) = 1$ and $z = a = 0$

This is the chance diagonal, $Y = X$.

$n = 1$, two points

$$z_0 = z_1 - b \qquad \omega(z_0) = \left[e^{2u} + e^{u+5b} + e^{u-.5b} + 1 \right]^{-.5}$$

$$z_1 = \ln \frac{e^{.5b} + e^u}{e^{-.5b} + e^u} \qquad \omega(z_1) = e^u \omega(z_0)$$

The ROC curve is called a "Luce-ROC" after Duncan Luce. (Ref. 35)

$$Y = e^{z_1} X \qquad \text{from } (0,0) \text{ to } \left(\frac{e^u}{e^u + e^{.5b}}, \frac{e^u}{e^u + e^{-.5b}} \right)$$

$$Y = 1 - e^{z_0}(1-X) \qquad \text{from the vertex to } (1,1)$$

For this special case it is easier to write

ROC Curve is line segment from $(0,0)$ to (X_0, Y_0)

from (X_0, Y_0) to $(1,1)$

and

$$z_0 = \ln \frac{1 - Y_0}{1 - X_0} \qquad \omega(z_0) = \sqrt{(1 - Y_0)(1 - X_0)}$$

$$z_1 = \ln \frac{Y_0}{X_0} \qquad \omega(z_1) = \sqrt{Y_0 X_0}$$

3.3.11 Negative Binomial, Pascal, Po'lya. For integer $t \geq 0$; $0 \leq p \leq 1$. and $q = 1 - p$.

$$\omega(t) = C_{r-1}^{t+r-1} \; p^r (1-p)^t$$

$$M(\xi) = p^r \; e^{\pm .5br} \; (1 - qe^{\pm .5b})^{-r} \qquad 0 < b < 2|\ln q|$$

$$a = r \ln \frac{e^{-.5b} - q}{e^{.5b} - q}$$

$$c = p^r \; \left[1 + q^2 - qe^{+.5b} - qe^{-.5b} \right]^{-.5r}$$

$$\omega_\pi(z) = \left(1 + q^2 - qe^{.5b} - qe^{-.5b} \right)^{.5r} C_{r-1}^{t+r-1} (1-p)^t$$

$$t = 0, 1, 2, \ldots$$

and

$$z = a + bt$$

The extensive use of tables is necessary to plot the ROC curves.

3.3.12 Poisson. For integer $t \geq 0$; $m > 0$.

$$\omega(t) = \frac{e^{-m} m^t}{t!}$$

$$M(\xi) = e^{m(e^{\xi} - 1)}$$

$$M(\pm.5b) = e^{m(e^{\pm 5b} - 1)}$$

$$a = -2m \sinh .5b$$

$$c = e^{2m[\sinh .25b]^2}$$

$$\omega_{\pi}(z) = e^{-m \cosh .5b} \, m^t/t! \quad t = 0, 1, 2, \ldots$$

$$z = bt - 2m \sinh .5b$$

The **ROC** curves can be obtained from chi-square tables, using the identities

$$\sum_{k}^{\infty} \frac{(me^{\pm .5b})^t}{t!} = P(\chi^2 = 2me^{\pm .5b} \mid \nu = 2k)$$

Many of the classical distribution functions for continuous and discrete variables do not possess moment generating functions. Among these are the Cauchy distribution and generalizations of the Cauchy distribution, the Pearson Types IV, VI, and XI, the F distribution and the Student's t distribution. These cover a semi-infinite or an infinite range, and have "tails" which go to zero as some power of the random variable.

3.4 Truncated ROC Characters and the Metastatic Transformation

ROC curves which approach the origin at some slope other than infinity or approach the point (1, 1) at some slope greater than zero correspond to ROC characters whose z-ranges are bounded away from $+\infty$ or $-\infty$, respectively. There is a functional relationship between many of these characters and those corresponding to ROC curves covering the complete range of slope from infinity to zero. There is similarly a geometric relation between the two types of ROC curves. This correspondence shall be called the metastatic relation. The word comes from the medical term metastasis, meaning a portion of something which has broken off and grows in a new place, yet retains its original character.

The metastatic transformation will be used in Section 3.5 to relate each nonregular but complete and convex. ROC curve and character to a regular ROC curve and character. This will extend the classification system based on associated distributions to these nonregular ROC.

3.4.1 The Transformation. The algebra of the transformation, and the geometric interpretation will be discussed at each step in the following development. Figure 3.1 displays a particular regular ROC curve; Eq. 3.36 is a formal representation of the corresponding functional relationship.

First ROC curve: $Y_1 = Y_1(X_1)$ (3.36)
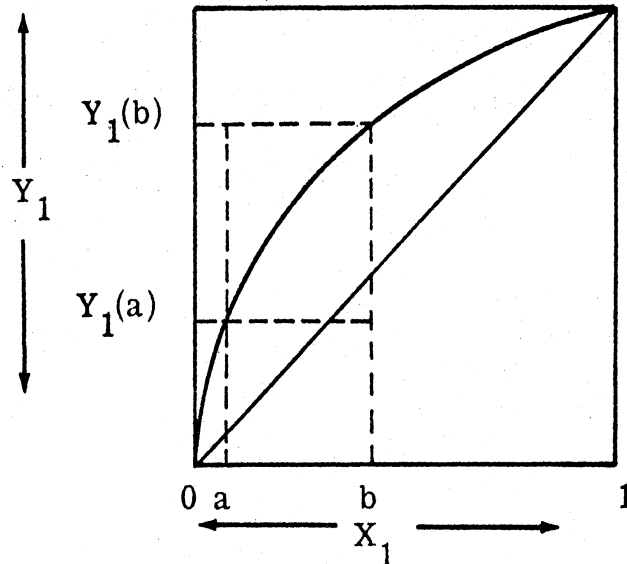
Fig. 3.1. A particular ROC curve

Also sketched in Fig. 3.1 are the coordinate lines at

$X_1 = a$, $X_1 = b$ and $Y_1 = Y_1(a)$, $Y_1 = Y_1(b)$. These form a rectangle,

with the first ROC curve passing through the lower left and upper

right corners. Between these two corners the arc of the first ROC

curve is interior to the rectangle, because the first ROC curve is

convex.

The metastatic transformation is the mapping of this

rectangle onto a unit ROC square. Geometrically, the rectangle is

removed, and the coordinate axes uniformly stretched to extend from

zero to one. The transformed arc of the first ROC curve is the sec-

ond ROC curve, the metastatic image. Is the second ROC regular?

Yes, it is complete, convex, and interior to the unit ROC square

except at $(0,0)$ and $(1,1)$. An equation for the new ROC curve, (X_2, Y_2), is

$$X_2 = (X_1 - a)/(b - a)$$

$$a \leq X_1 \leq b \qquad (3.37)$$

$$Y_2 = (Y_1(X_1) - Y_1(a))/(Y_1(b) - Y_1(a))$$

When a point (X_1, Y_1) with slope e^{z_1} maps onto a point (X_2, Y_2) the slope is multiplied by the ratio of the Y to X axis expansion. Formally

$$e^{z_2} = \frac{dY_2}{dX_2} = \frac{b - a}{Y_1(b) - Y_1(a)} e^{z_1} \qquad (3.38)$$

Therefore

$$z_2 = z_1 + \ln \frac{b - a}{Y_1(b) - Y_1(a)} \qquad (3.39)$$

The second z-axis is simply a translation or horizontal shift of the first z-axis.

The ROC character for the second ROC curve can be obtained as follows. The formal relationship between the ROC character and the probability density function of the log likelihood ratio, conditional to the condition SN is

$$\pi_2(z_2) = e^{-.5z_2} f(z_2|SN) \qquad (3.40)$$

This probability density function can be found by differentiating the

corresponding distribution function, $1 - Y_2$.

$$\left. \begin{aligned} f(z_2 | SN) &= \frac{d}{d z_2} \, (1 - Y_2) \\ \\ &= \frac{d(1 - Y_2)}{d(1 - Y_1)} \; \frac{d(1 - Y_1)}{d z_1} \; \frac{d z_1}{d z_2} \\ \\ &= \frac{1}{Y_1(b) - Y_1(a)} \; f(z_1 | SN) \end{aligned} \right\} \qquad (3.41)$$

The probability density function of the first log likelihood ratio in

terms of its ROC character is

$$f(z_1 | SN) = e^{.5 z_1} \, \pi_1(z_1) \qquad (3.42)$$

From Eq. 3.39 which relates the two z-axes,

$$e^{.5 z_2} = e^{.5 z_1} \left[\frac{b - a}{Y_1(b) - Y_1(a)} \right]^{.5} \qquad (3.43)$$

Equations 3.40 through 3.43 lead to the equation for the second ROC

character

$$\pi_2(z) = \frac{\pi_1 \left(z + \ln \dfrac{Y_1(b) - Y_1(a)}{b - a} \right)}{\left[(b - a) \, (Y_1(b) - Y_1(a)) \right]^{.5}} \qquad (3.44)$$

Let us draw a parallel between the geometric trans-
formation of the ROC curves and the algebraic changes in the ROC
character. Both a horizontal and vertical scaling convert the rec-
tangle to the unit square. In Eq. 3.44 there is a vertical scaling given
by the denominator, and a horizontal translation, not scaling, given by
the added constant in the argument of the numerator. The second par-
allel has not been explicit. Only an arc of the original ROC curve
was transformed, not the entire curve. The bounds on the z_2 axis
correspond to the upper and lower bounds on the z_1 variable at the
two ends of the arc by Eq. 3.39. In summary, the ROC curve was
cut, and a small portion selected. This parallels the condition that
only a portion of the original character is used by the metastatic char-
acter. The ROC curve was re-scaled to fit the unit square; the ROC
character has been re-scaled according to its necessary properties.

In Fig. 3.2 are two specific ROC curves and their cor-
responding characters. The first is a normal ROC curve with index
d of 2.828. Above it is the normal ROC character, a curve propor-
tional to the normal density function, centered symmetrically about
zero and extending from $-\infty$ to $+\infty$. The metastatic transformation
chosen to illustrate the process was the part of the arc from $(0, 0)$ to
the negative diagonal. This corresponds to the z_1 axis from zero to
infinity. A metastatic transformation maps the dotted rectangle to the
unit square. The transformation on the ROC character truncates the
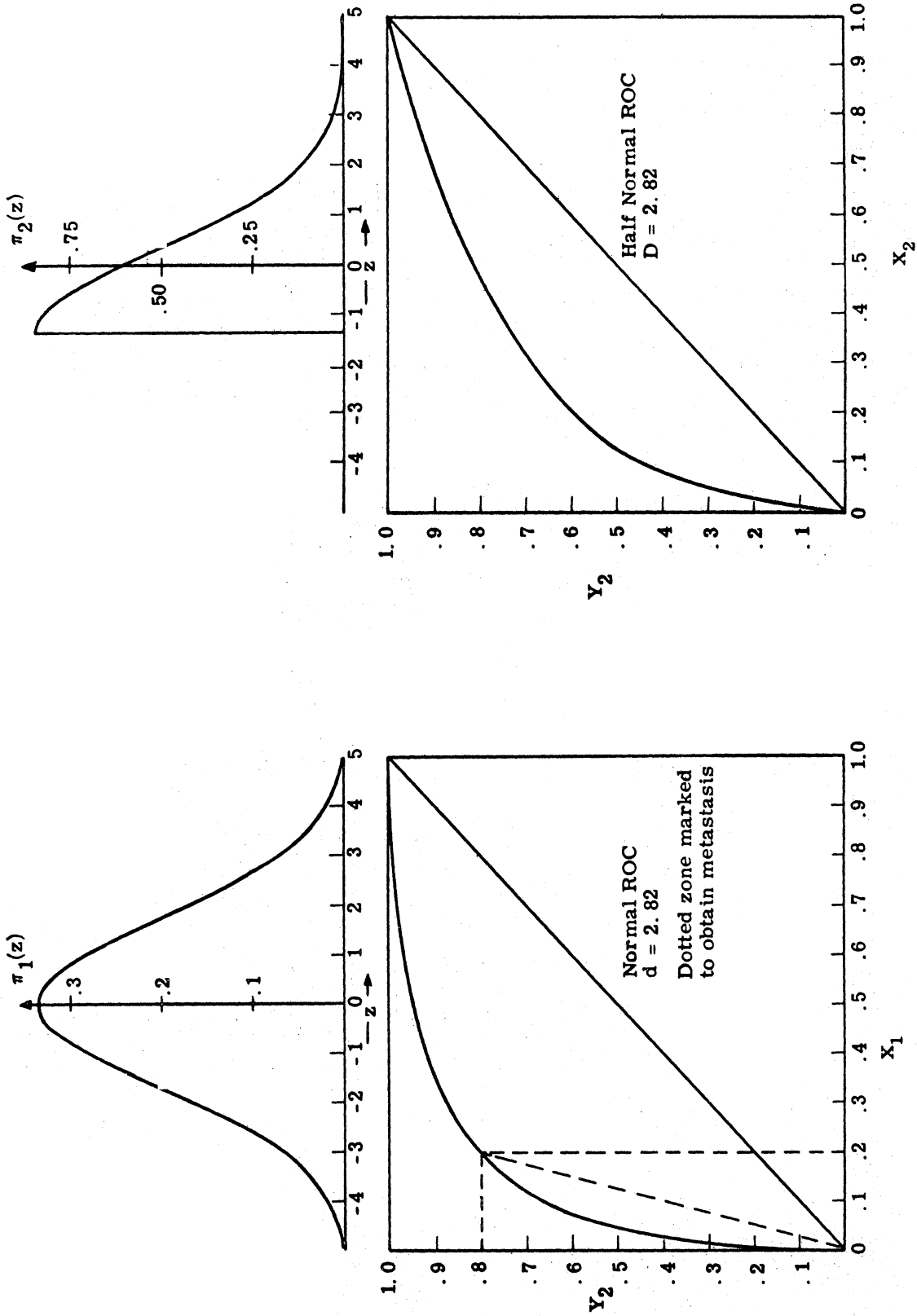
Fig. 3. 2. Example of metastatic transformation

ROC character on the left, re-scales it vertically (the scaling is a factor of 2. 5 for this specific example), and translates the axis horizontally the appropriate amount. The resulting ROC curve and character are shown on the right in Fig. 3. 2. The ROC curves are repeated in Fig. 3. 3 on normal-normal probability paper.

Figure 3. 4 shows ROC curves before and after a metastatic transformation from normal with d = 6. 55. The metastatic ROC curve is the image of the original ROC curve from the (0, 0) point to the negative diagonal. Both of these metastatic "half-normal" ROC curves appear to be binormal ROC curves when plotted on normal-normal paper.

3. 4. 2 Self-Metastatic Families. Certain families of ROC curves are self-metastatic; that is, a metastatic transformation of one character will yield a second character which is in the same class. The most startling of these cases is the pure power ROC curve with exponential ROC character (see Section 3. 3. 2). This is the type corresponding to the ROC curve

$$x = y^A \tag{3.45}$$

A metastatic transformation of the power ROC curve containing the point (0, 0) to any point along the curve is the same as the original. The formal proof is in Appendix B.

The rectangular hyperbola ROC curve has a rectangular
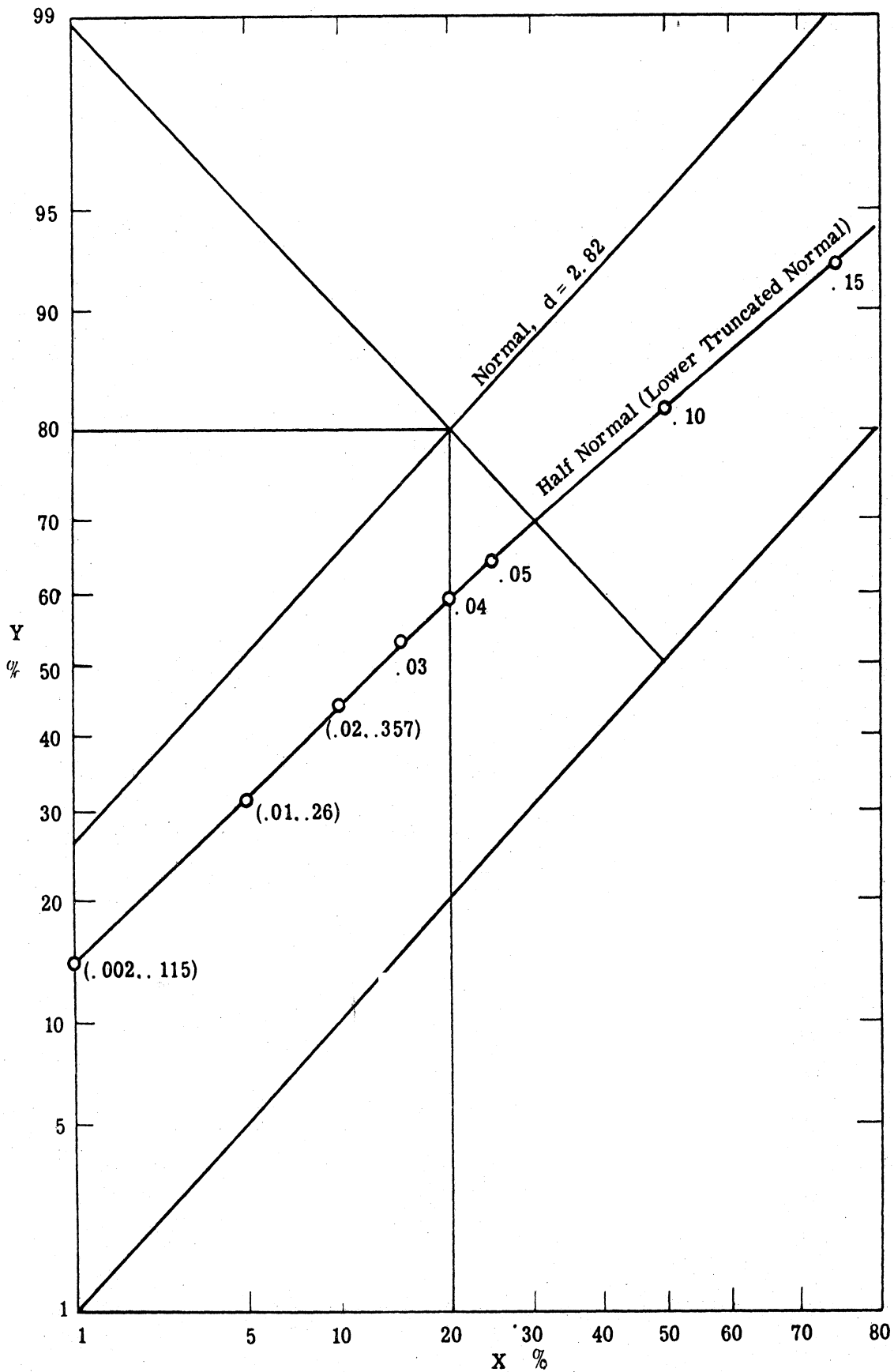
Fig. 3.3. Replot of Fig. 3.2
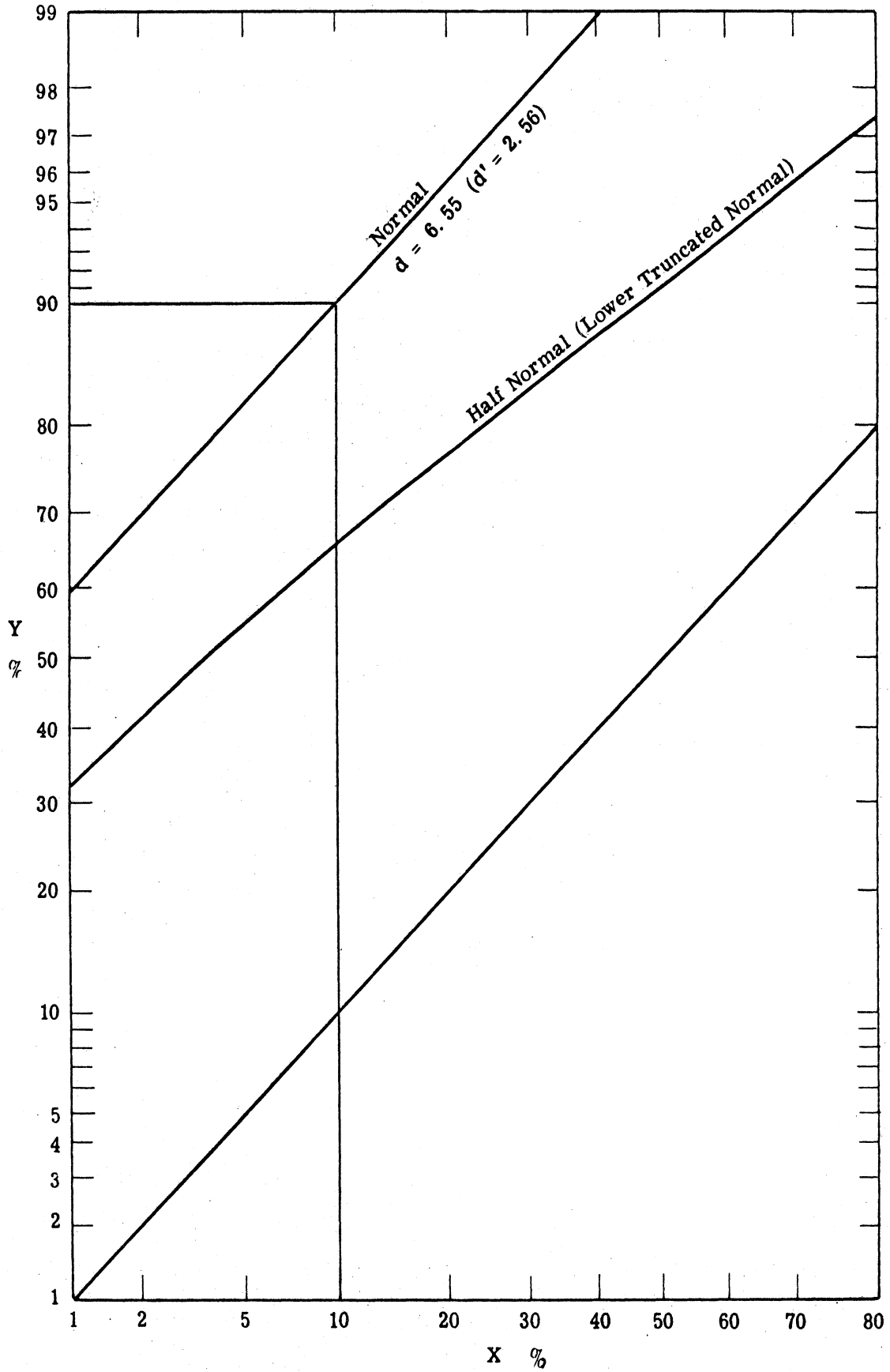
Fig. 3.4. Second example of metastatic transformation

(flat) ROC character. Such a character, being a simple rectangle,
is symmetric about its midpoint. The midpoint must be $z = 0$
since this is the symmetry point for all symmetric ROC characters
(see Section 2. 4). If a flat character is truncated either on the right
or on the left, or both, and then the translation and rescaling dicta-
ted by the metastatic transform rules performed, the character will
again be flat and symmetrically placed about $z = 0$. Since the char-
acter will not span as much of the z axis as the original character
did, the new ROC curve will be different from the first; it will still
be a rectangular hyperbola. The more that is removed from the ori-
ginal character, the poorer the resulting right hyperbola ROC curve.
As a family, rectangular ROC are self-metastatic.

It will be shown in Chapter V that the ROC character
for the conic class is

$$\pi(z) = \pm K(Ae^{Z} + 2B + Ce^{-Z})^{-1.5} \tag{3.46}$$

where the range of the log likelihood ratio, z, depends on the specific
curve being considered. The coefficients A, B, and C are the first
three coefficients in the classical form of a conic section. The trans-
lation of the z axis, Eq. 3.39, will be absorbed into the new coeffi-
cients A and C, magnifying one and diminishing the other propor-
tionately. The discriminant for a conic section in classic form is
$B^{2} - AC$. Any translation of z, although changing the coefficients A

and C , will leave the value of the discriminant unchanged. Not only will the individual subtypes of conic section (hyperbola, parabola, ellipse), be self-metastatic families, but each value of the discriminant will correspond to a self-metastatic subfamily.

3.5 Nonsingular, Nonregular ROC

The only difference between the complete convex nonsingular ROC curve and a regular ROC curve is that the initial lefthand point of the ROC curve may begin at some nonzero value and/or attain unity probability of detection at some value of false alarm probability less than one. Such a complete convex nonsingular ROC curve is illustrated in Fig. 3.5. It should be emphasized that both conditions of starting above zero and terminating before one are not necessary, either one is sufficient to make the curve nonregular. The singular ROC curve is simply the upper edge. Nonsingular means that the initial point is not as high as one, nor is the value of false alarm at which it reaches unity detection as small as zero.

$$\text{Nonsingular} \Rightarrow Y(0) \neq 1, \quad X(1) \neq 0 \qquad (3.47)$$

On Fig. 3.5, a metastatic transformation of the rectangle from $Y(0)$ to one, and from zero to $X(1)$ would produce a regular metastatic image. For this metastatic image the ROC character would be
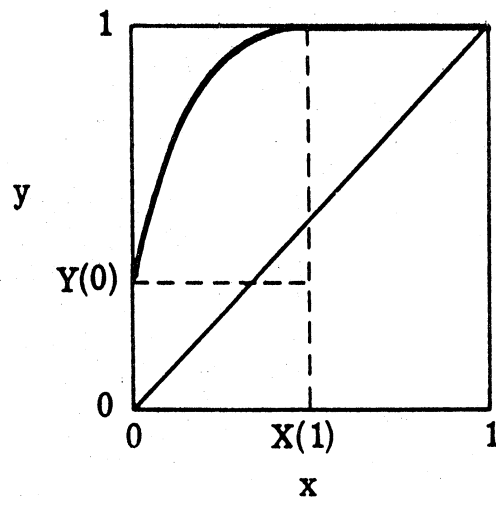
Fig. 3.5. A complete, convex, nonsingular ROC
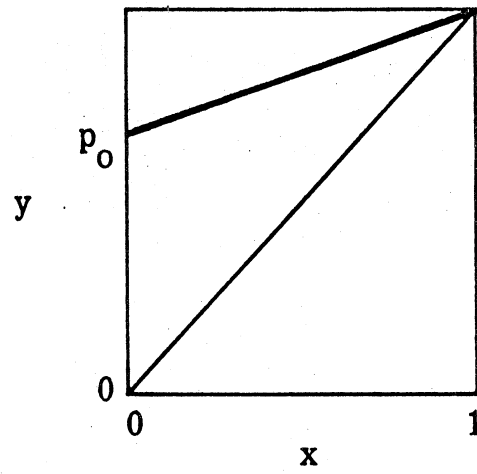
Fig. 3.6. Traditional threshold ROC
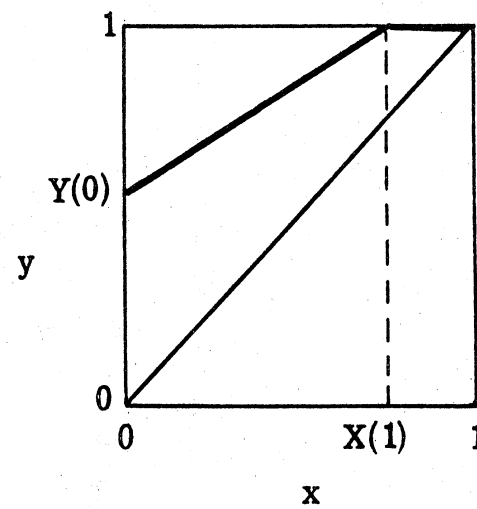
Fig. 3.7. Green ROC

$$\pi_2(z) = \frac{\pi_1\left(z + \ln\frac{1-Y(0)}{X(1)}\right)}{[X(1)(1-Y(0))]^{.5}} \tag{3.48}$$

Consider any ROC character, $\pi_2(z)$ with either bounded or unbounded range denoted by $\underline{z_2}$ to $\overline{z_2}$.

$$\text{Given:} \quad \pi_2(z), \quad \underline{z_2} < z < \overline{z_2} \tag{3.49}$$

If this were the metastatic image of some arc, such as in Fig. 3.5, the ROC character for the regular part of the original arc would be

$$\left.\begin{array}{c} \pi_1(z) = [X(1)(1-Y(0))]^{.5} \; \pi_2\left(z - \ln\frac{1-Y(0)}{X(1)}\right) \\[3mm] \underline{z_2} - \ln\frac{1-Y(0)}{X(1)} < z < \overline{z_2} - \ln\frac{1-Y(0)}{X(1)} \end{array}\right\} \tag{3.50}$$

The complete description of the nonregular ROC curve contains the boundary conditions.

$$P(z_1 = -\infty \mid N) = 1 - X(1), \quad P(z_1 = +\infty \mid SN) = Y(0) \tag{3.51}$$

Such boundary conditions cannot be included in the ROC character since probabilities are obtained by multiplying the character value by $e^{\pm .5z}$. This multiplier is either zero or infinite at the extreme boundary values.

If the values for the boundary were unknown, they could be obtained by integrating the nonregular ROC character over the

range between $-\infty$ and $+\infty$.

$$\int_{-\infty}^{\infty} e^{+.5z} \, d\Pi_1(z) = 1 - Y(0) \tag{3.52}$$

$$\int_{-\infty}^{\infty} e^{-.5z} \, d\Pi_1(z) = X(1) \tag{3.53}$$

The above equations are the essential formal difference between the nonregular and regular ROC characters. For the nonregular ROC the value of one or both of these integrals will be less than one, indicating that portion of the decision which is error free. The regular ROC character always has a value of unity for the complete integral (Eqs. 3.8, 3.9).

Two special ROC curves with discrete ROC character have been of importance in the psychophysical literature. Both are related to nonregular ROC characters with just one value of log likelihood ratio other than $\pm\infty$.

The pure-threshold or traditional-threshold ROC has only two values of log likelihood ratio that play any role, $+\infty$ and some negative value Z_0. The ROC character jump is

$$\omega_\pi(Z_0) = e^{.5Z_0} \tag{3.54}$$

Equations 3.52 and 3.53 are applied to this nonregular ROC character to determine that $X(1)$ is one, meaning that the ROC curve first

contacts the upper edge at the corner point $(1, 1)$. The ROC curve is given by Eq. 3.55, and shown in Fig. 3.6.

$$Y = 1 - e^{Z_0} + e^{Z_0} X \qquad (3.55)$$

The more common form of this equation may be given by relabeling $Y(0)$ as p_0, sometimes called the "true probability of perception."

$$Y = p_0 + (1 - p_0) X \qquad (3.56)$$

The value Z_0 and its ROC character jump are

$$Z_0 = \ln(1 - p_0), \quad \omega_\pi(Z_0) = (1 - p_0)^{.5} \qquad (3.57)$$

A second nonregular ROC curve related to a single point on the log likelihood ratio axis is the Green ROC. This is the nonregular limit of D. M. Green's double threshold ROC curve. For any point z_0 and any ROC character jump value small enough that

$$\omega_\pi(z_0) < e^{.5 z_0} \qquad (3.58)$$

The formal expressions for the Green ROC curve are

$$\left. \begin{array}{ll} Y(0) = 1 - e^{.5Z_0} \omega(Z_0) & X = 0 \\[2mm] Y = Y(0) + e^{Z_0} X & 0 \le X \le X(1) = e^{-.5Z_0} \omega(Z_0) \\[2mm] Y = 1 & X(1) \le X \le 1 \end{array} \right\} \qquad (3.59)$$

The ROC curve is sketched in Fig. 3.7 for a slightly negative value of z_0.

3.6 Summary of Chapters I, II, and III

The purpose of this research was to furnish a variety of ROC curves, to develop sufficient analytic structure to classify ROC curves into families and to provide a means of generating new families. In Chapter II it was shown that the distribution of a decision variable may be quite arbitrary under one cause condition; therefore decision models may be developed with considerable freedom.

The ROC curve contains information sufficient to specify a decision model iff the decision axis bears a specific functional relation to the likelihood ratio ℓ. This research has concentrated on z, the logarithm of the likelihood ratio, because the distributions of z are necessarily sufficiently concentrated to possess moment generating functions. It has been demonstrated that the ROC character can be used to structure ROC curves into families, and to provide both inter-family relationships and a basis for generating new ROC families.

This completes the development of the research into the structure of ROC curves. The following chapters deal with specific ROC families, in order to determine properties peculiar to each, and to demonstrate the techniques for working with ROC curves and ROC families.

CHAPTER IV

TRADITIONAL ROC FAMILIES

This chapter will deal with those ROC families that appeared in Ref. 1.

4.1 Normal ROC

The normal ROC has been used extensively in both the electronic and psychophysical literature (Refs. 1, 4, 7 through 14). This type of ROC curve was used so extensively in the original psychophysical work of Tanner, Swets, and Green, that many thought it was a necessary part of their perception theory. Tables of the normal ROC curve are available (Refs. 17 and 18).

The ROC character for the normal ROC curve is

$$\pi(z) = e^{-\frac{d}{8}} \frac{1}{\sqrt{2\pi d}} e^{-\frac{z^2}{2d}} \qquad -\infty < z < \infty$$

$$d > 0$$

(4.1)

Multiply the character by $e^{\pm .5z}$ to obtain the probability density functions for the logarithm of the likelihood ratio under the two conditions N and SN. Both of these are normal,

$$f(z \mid N) = \frac{1}{\sqrt{2\pi d}} \; e^{- \frac{(z + .5d)^2}{\sqrt{2d}}} \tag{4.2}$$

$$f(z \mid SN) = \frac{1}{\sqrt{2\pi d}} \; e^{- \frac{(z - .5d)^2}{\sqrt{2d}}} \tag{4.3}$$

with equal variance and with mean values shifted plus and minus .5d from the character mode of zero. The ROC curve is obtained by direct integration of Eqs. 4.2 and 4.3.

$$P("A" \mid N) = x = \Phi \left(\frac{-.5d - z}{\sqrt{d}} \right)$$

$$P("A" \mid SN) = y = \Phi \left(\frac{.5d - z}{\sqrt{d}} \right) \tag{4.4}$$

The symbol Φ stands for the normal distribution function. From this point on, the probability notation for the ROC curve will usually be omitted and only the description of the ROC curve as a real function, y, of a real variable, x, retained. While Eq. 4.4 explicitly indicates the relation between the coordinate values and the logarithm of the likelihood ratio, this relation is not always desired. z may be suppressed, and the equation for the ROC curve written as

$$y = \Phi (\lambda + \sqrt{d}) \text{ when } x = \Phi (\lambda) \tag{4.5}$$

where the variable λ is the dummy parameter along the ROC curve. In

this form the natural parameterization is by $\sqrt{d}$, which is normally

called d' in the psychophysical literature.

Graphs of the normal ROC are displayed in Figs. 4.1

to 4.3. In 4.1 they are displayed on ordinary linear coordinate paper,

with curves given for d values of zero (the chance diagonal) and

d = 1, d = 4, and d = 9. This reflects the natural stepping shown in

Eq. 4.5. The coordinate paper for Fig. 4.2 has been called "normal-

normal paper" indicating that it is related to the normal distribution

function on both axes, and is also called "double-probability paper"

and "z-scale paper." The linear distance in each coordinate direction

on this paper is the argument of the normal distribution function. Refer-

ring to Eq. 4.5, on such paper a normal ROC point plots with the

vertical coordinate $\lambda + \sqrt{d}$, and horizontal coordinate λ . These

lie along a straight line with slope one and separated by a difference

$\sqrt{d}$. Transformation to normal coordinates spreads the unit square

over the entire infinite two-dimensional plane. The region displayed in

Fig. 4.2 is that used for medium probabilities, between one percent and

ninety-nine percent. The region displayed in Fig. 4.3 is more common

in machine application where very small probabilities can be measured,

and extends from 10^{-6} to $1 - 10^{-6}$

The main diagonal with slope plus one is the chance

diagonal. The minor diagonal, with slope minus one, is usually

referred to as the "negative diagonal" and corresponds to those points

where the probability of miss, 1 - y, is equal to the probability of false

alarm, x.
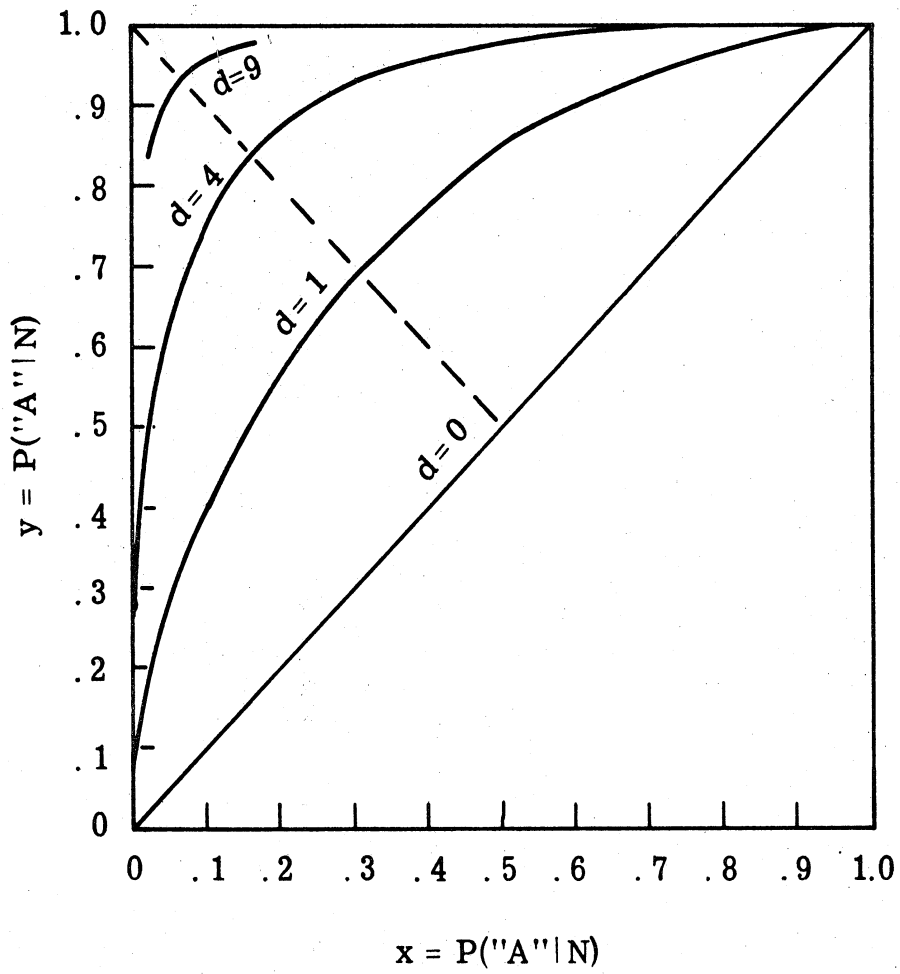
Fig. 4. 1. Normal ROC
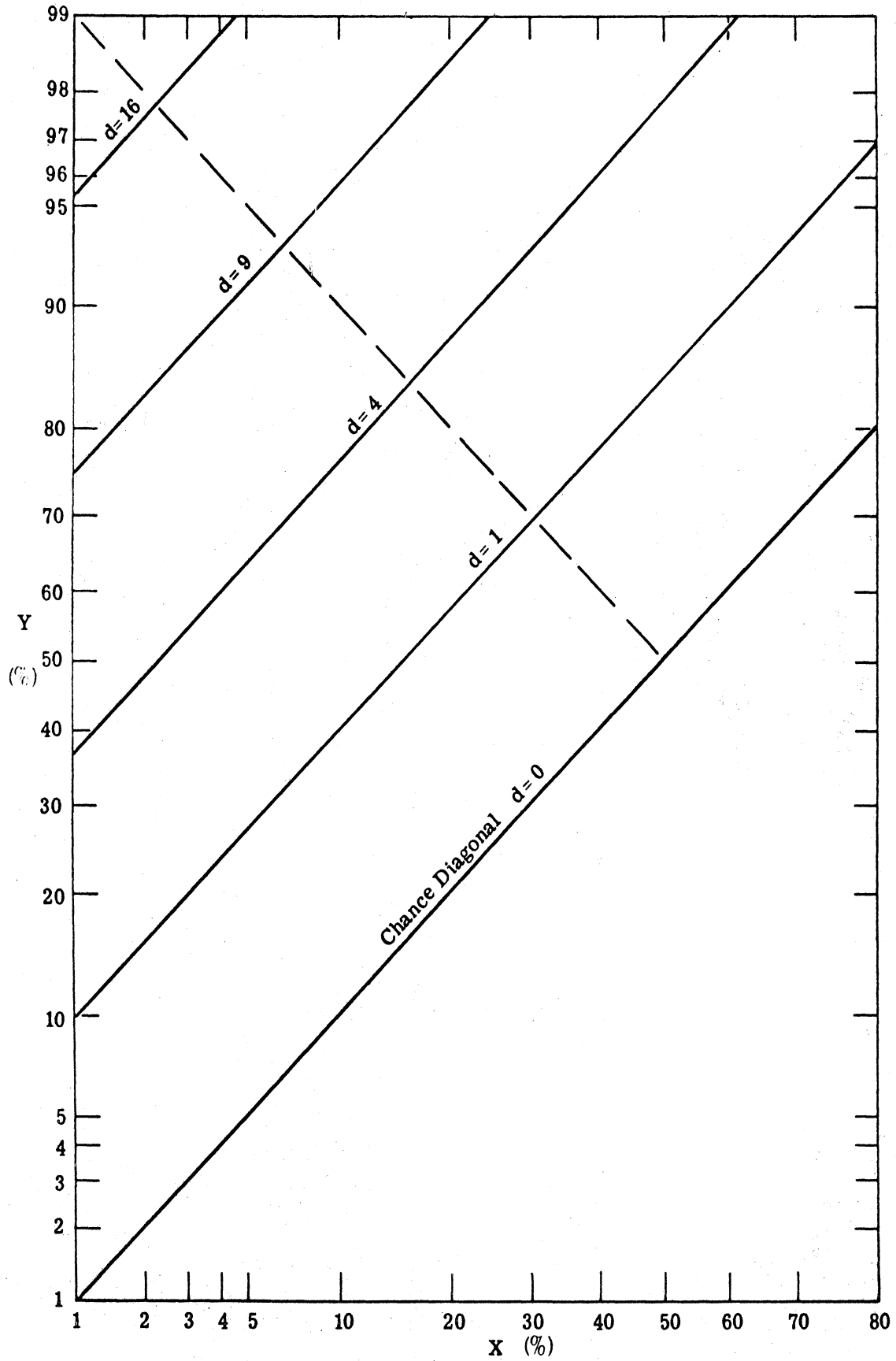
Fig. 4. 2. Normal ROC on "normal-normal" paper
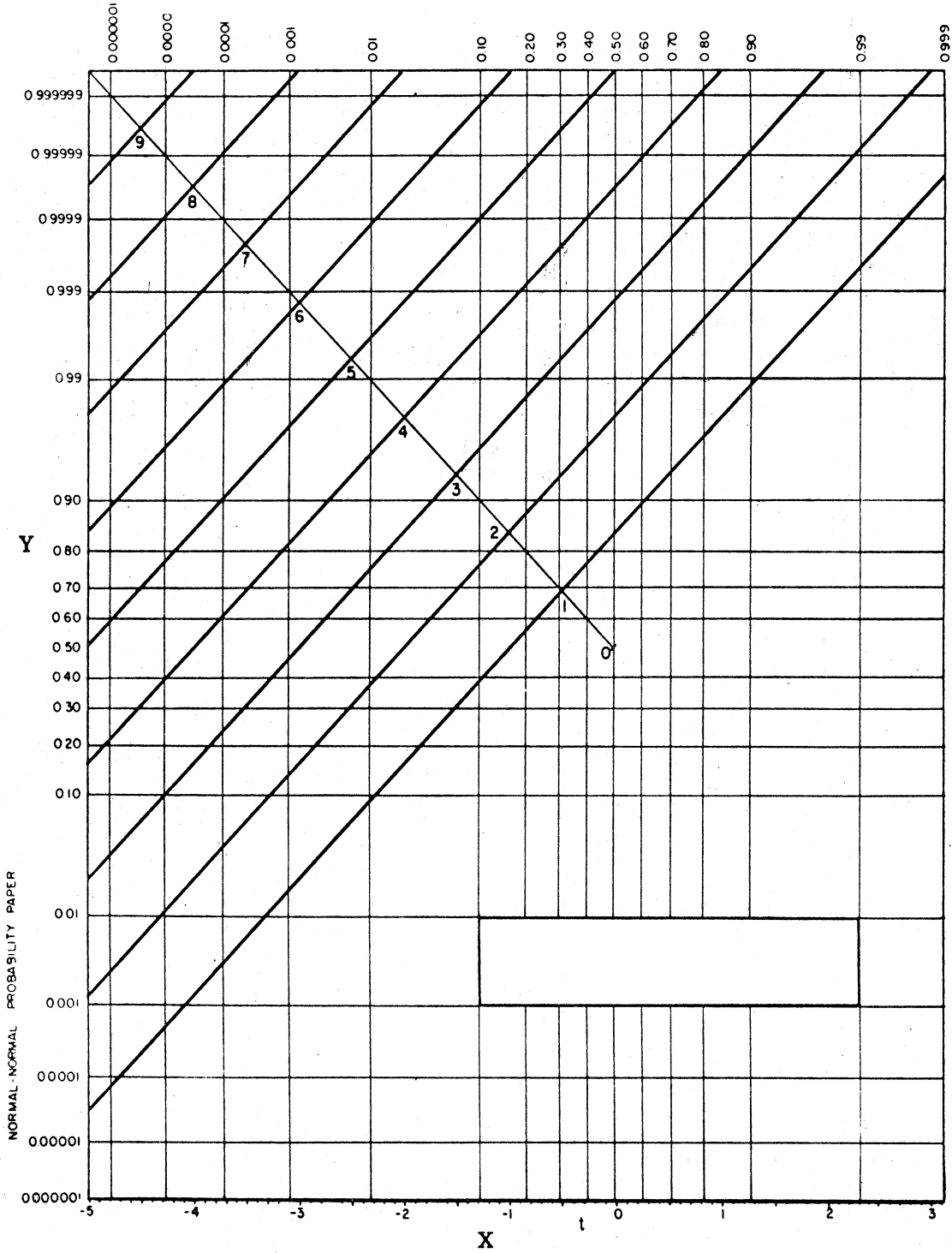
Fig. 4.3. Normal receiver operating characteristics

Four normal ROC characters are plotted in Fig. 4.4

on the equivalent of semi-log paper; that is, ln $\pi(z)$ is plotted against

z . These are simple parabolas with center at $z = 0$ and opening

downward. As the detectability increases, d increases, the ROC

characters broaden out, giving more weight to large magnitude z

values. The locus of the one percent false alarm probability and 10^{-6}

false alarm probability have been indicated. Since the ROC curves,

and ROC characters, are symmetric, the negative diagonal corresponds

to $z = 0$. The regions of interest for the psychologist will normally lie

in the neighborhood of the negative diagonal with possible extension to

high positive z values out to the one percent point, $x = .01$. The

region of interest for most radar and sonar applications is between the

one percent and 10^{-6} false alarm probabilities. (In the past, some have

been interested in false alarm probabilities as low as 10^{-14}.) These

correspond to regions of high positive z value.

4.2 "Case II", Detection of a Sine Wave in Added Normal Noise

A problem occurring frequently in the electronic literature

is the detection of a stable sine wave with uniformly uncertain phase in

the presence of added white Gaussian noise. This has been treated by

Rice, (Ref. 19), by Marcum (Ref. 16), by Middleton (Ref. 7), by Helstrom

(Ref. 15), was included in Ref. 1 as Case II, and has been recently

investigated extensively for its usefulness in hearing by Lloyd Jeffries

(Ref. 20).
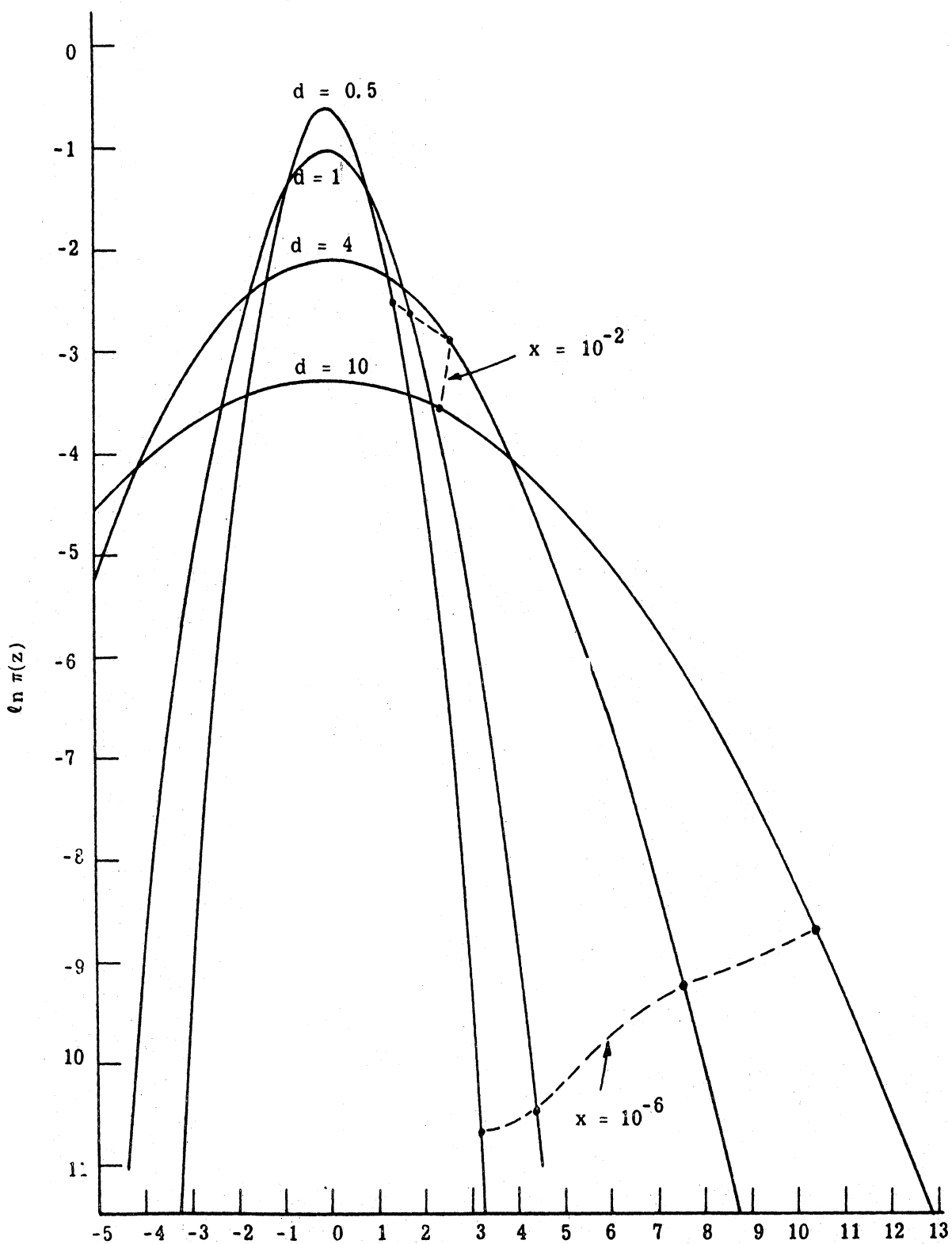
Fig. 4.4. Four normal ROC characters

The normal ROC case has been used so easily by so many. Yet its equation, Eq. 4.5, is really quite complicated. The normal distribution function $\Phi(\ )$ cannot be determined in terms of polynomials or the elementary transcendental functions. It owes whatever simplicity it possesses to the ease with which tables can be obtained, and familiarity with this specific function. In the case at hand, the coordinates of the ROC curve cannot be simply written in terms of each other, and the parametric form of the curve again involves a tabulated function. However, this particular function is one with which few people have great familiarity and for which the tables are not as readily available as one would wish. Tables have been calculated by Marcum (Ref. 16), specifically for the evaluation of this ROC curve. These are called Q Tables since that is the symbol that he uses.

$$\text{Q Tables: } x = e^{-\beta^2 / 4} = Q(0, \beta) = \int_{\beta}^{\infty} t\, e^{-t^2/2}\, dt$$

$$y = Q(\alpha, \beta) = \int_{\beta}^{\infty} t\, e^{-t^2/2}\, e^{-\alpha^2/2}\, I_0(\alpha t)\, dt$$

$$(4.6)$$

The false alarm probability is a very simple function of the parameter of the curve, β, but the detection probability involves integration of a function whose integrand contains the modified Bessel function of order zero, itself a non-simple tabulated function. The parameter α, a

positive real number, indicates the quality of detection. As α increased, the ROC curves become better. The Q tables were used to plot the ROC family of Fig. 4. 5.

α	√d Values			Slope	Formulae	
	At Y = 50; q	At X = 1 - Y	At X = .50	s	$\alpha - .5 \sqrt{\ln I_0(\alpha)}$	
0	0	0	0	1	-	0
.5	-	-	0.108	0.970	0	-
1.0	0.412	0.40	0.381	0.925	0.5	0.486
1.5	0.869	-	0.761	0.877	1.0	-
2.0	1.397	1.30	1.188	0.851	1.5	1.56
2.5	1.940	-	1.640	0.845	2.0	-
3.0	2.476	2.30	2.106	0.852	2.5	2.65
3.5	3.003	-	2.581	0.861	3.0	-
4.0	3.512	3.35	3.544	0.872	3.5	3.70
4.5	4.001	-	4.036	0.885	4.0	-
5.0	-	4.40	-	-	4.5	4.75
5.5	-	Graph Table	-	-	5.0	-
6.0	-	5.45	-	-	5.5	5.78
7.0	-	6.47	-	-	6.5	6.80
8.0	-	7.51	-	-	7.5	7.82
9.0	-	8.55	-	-	8.5	8.84

Note: q and s are binormal approximation numbers. See Chapter VII.

Table 4. 1. Numbers relating Q-Table ROC to normal

The √d and slope readings for a range of α are given in Table 4. 1 and plotted in Fig. 4. 6.

Consider a random variable t which ranges from zero to infinity with probability density functions given by the integrand in Eq. 4. 6. It is immediately evident that the likelihood ration of t is the product of the third and fourth terms in the integrand of the equation
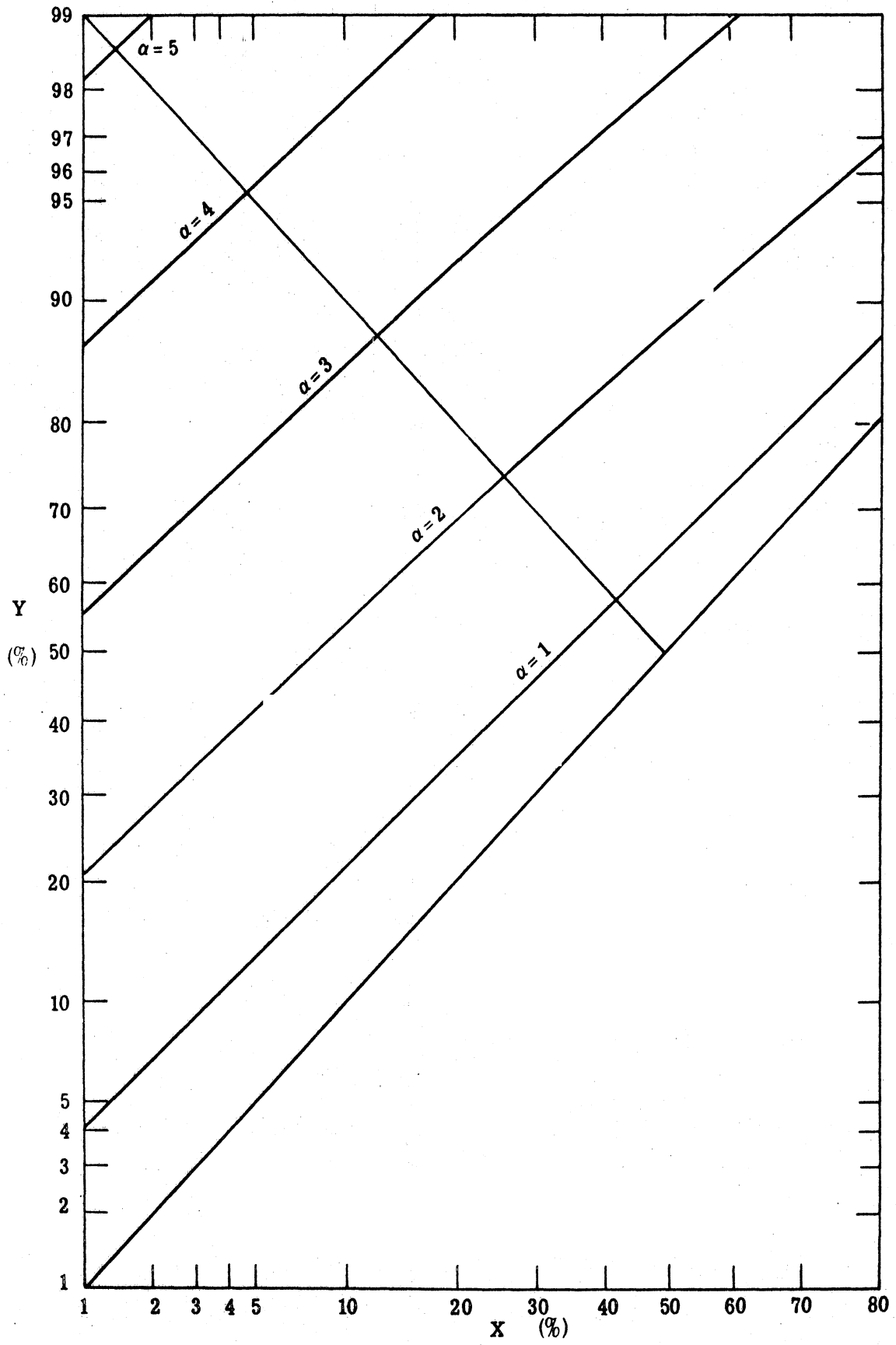
Fig. 4.5. ROC based on Q tables
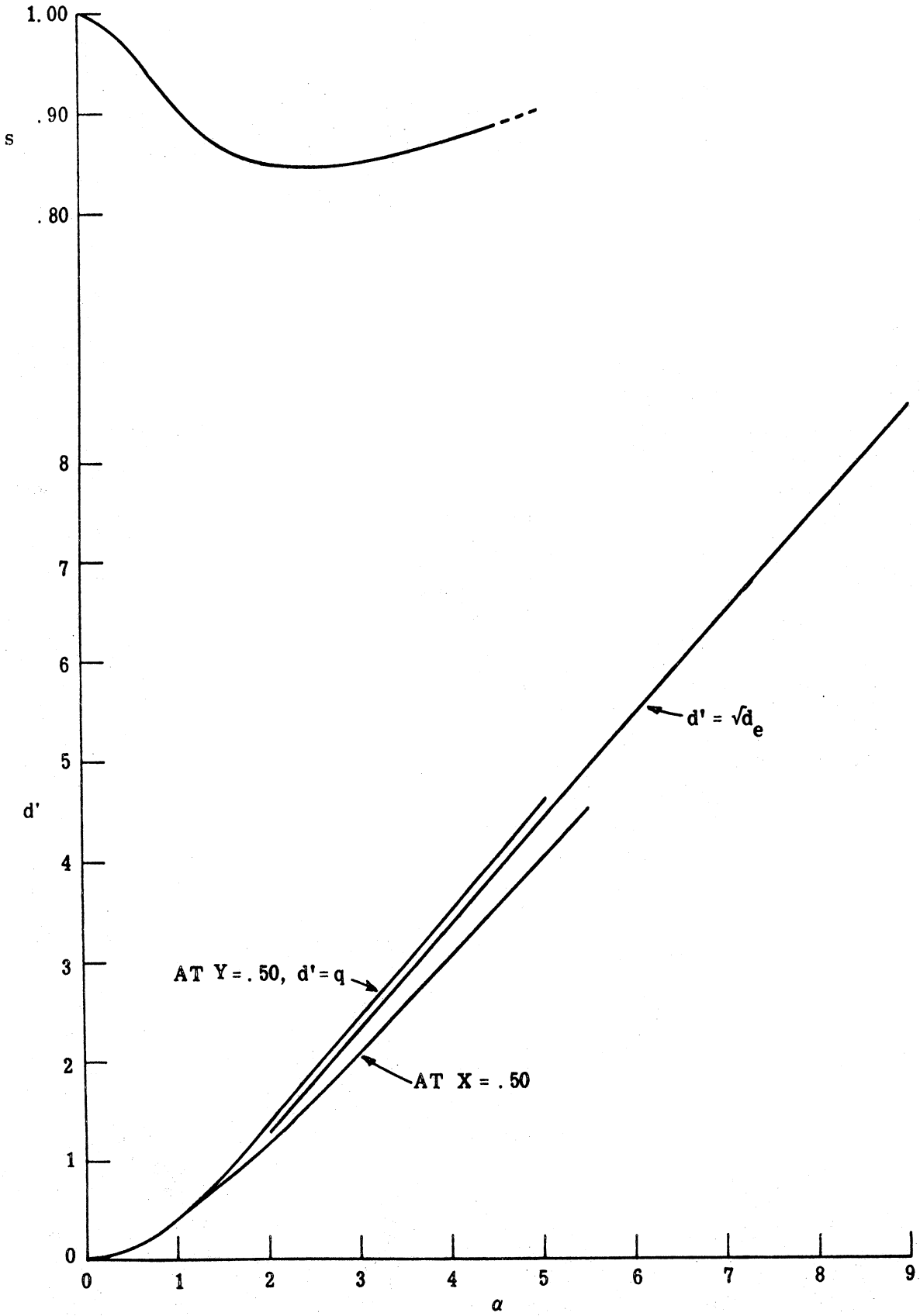
Fig. 4.6. Slope and quality values relating
Q-Table ROC to normal and binormal

for y (Eq. 4.6). Therefore, the logarithm for the likelihood ratio of t is

$$z = \ln I_0 (\alpha t) - .5\alpha^2 \qquad (4.7)$$

and is strictly monotone increasing with the variable t. Those familiar with the chi-square distribution will recognize that under the N condition t is the square root of chi-square with two degrees of freedom while under the SN condition t is the square root of a noncentral chi-square with two degrees of freedom.

$$t = \sqrt{X^2}_{2d.f.} \qquad \text{vs.} \qquad t = \sqrt{X'^2}_{2d.f.} \qquad (4.8)$$

--- general problem of detecting a shift from central to noncentral chi-square distribution with a known number of degrees of freedom has been treated in Ref. 21. The variable t rather than chi-square has been chosen for two reasons; the first is that Marcum's tables use t instead of t^2, the second is that the Bessel function $\ln I_0(\alpha t)$ is nearly linear with its argument, αt, when the argument is large. This means that the density function for z should be of the same type as the density function for t if large values of αt are the relevant values.

The ROC character for z can be obtained from the root likelihood product of t multiplied by the Jacobian of transformation from t to z. The root likelihood product of t is simply the N density function multiplied by $e^{.5z}$.

$$\pi(z) \quad f(t \mid N) \, e^{.5z} \left| \frac{dt}{dz} \right| \qquad (4.9)$$

It would be desirable to write t in terms of z and determine the form of the ROC character as a function of its argument, z. We cannot invert Eq. 4.7 in any practical fashion. Therefore, the best procedure is to write the ROC character in terms of the parameter t, and to attempt to graph or approximate the ROC character.

$$\pi(z) = t e^{-t^2/2} \, e^{-\alpha^2/2} \, I_0^{.5} (\alpha t) \, \frac{I_0(\alpha t)}{\alpha I_1(\alpha t)} \qquad 4.10)$$

Table 9.8 of the NBS Handbook (Ref. 6) gives not the modified Bessel function, but the more slowly varying quantity.

$$T_\nu(\lambda) = e^{-\lambda} \, I_\nu(\lambda) \qquad (4.11)$$

(T_ν has been used to indicate "tabulated function.") In order to use the tables, Eq. 4.10 is rewritten in terms of the table entries. The logarithm of the Bessel functions can be written as

$$\ln I_\nu(\lambda) = \ln T_\nu(\lambda) + \lambda \qquad (4.12)$$

and inserted in Eq. 4.7 and 4.10.

$$z = [\ln T_0(\alpha t) + \alpha t]_1 - [.5\alpha^2] \qquad (4.13)$$

$$\ln \pi (z) = \left[.5(\alpha t) + \ln(\alpha t) + 1.5 \ln T_0 (\alpha t) - \ln T_1 (\alpha t) \right]_2 - .5t^2$$

$$- \left[.25 \alpha^2 + 2 \ln \alpha \right] \quad (4.14)$$

The argument of the Bessel function is αt, and therefore the equations

have been written in terms of a quantity which is a function of αt alone,

and then modifiers depending upon t or α. Since α is the parameter

of an ROC curve, it will be constant over any computations of the ROC

or its character. These equations were utilized to calculate ROC

characters with table values and slide rule. The resultant ROC

character for three values of the parameter is shown in Fig. 4.7. This

is shown on the same coordinate system that was used for the normal

ROC character Fig. 4. 4. The range of the random variable z is

bounded below by $.5\alpha^2$ as can be seen in Eq. 4.13, and hence the ROC

curve will have a nonzero minimum slope.

Since the ROC character is bounded below but unbounded

above, it cannot possibly be symmetric. The mode of the ROC character

for the cases shown is slightly negative. In the case $\alpha = 1$ the mode

is at the extreme negative value, $z = -.5$. For the cases $\alpha = 2$ and

$\alpha = \sqrt{10}$, the mode values are approximately $z = -.73$, and $z = -.79$,

respectively. As detectability increases (α increases) the logarithm of

the ROC character covers a wider and wider range and begins to appear

similar to those for the normal case. It is therefore very tempting to

attempt to fit the ROC curve with a normal ROC curve, to fit the ROC

character with a normal ROC character. In Ref. 22, the approximation
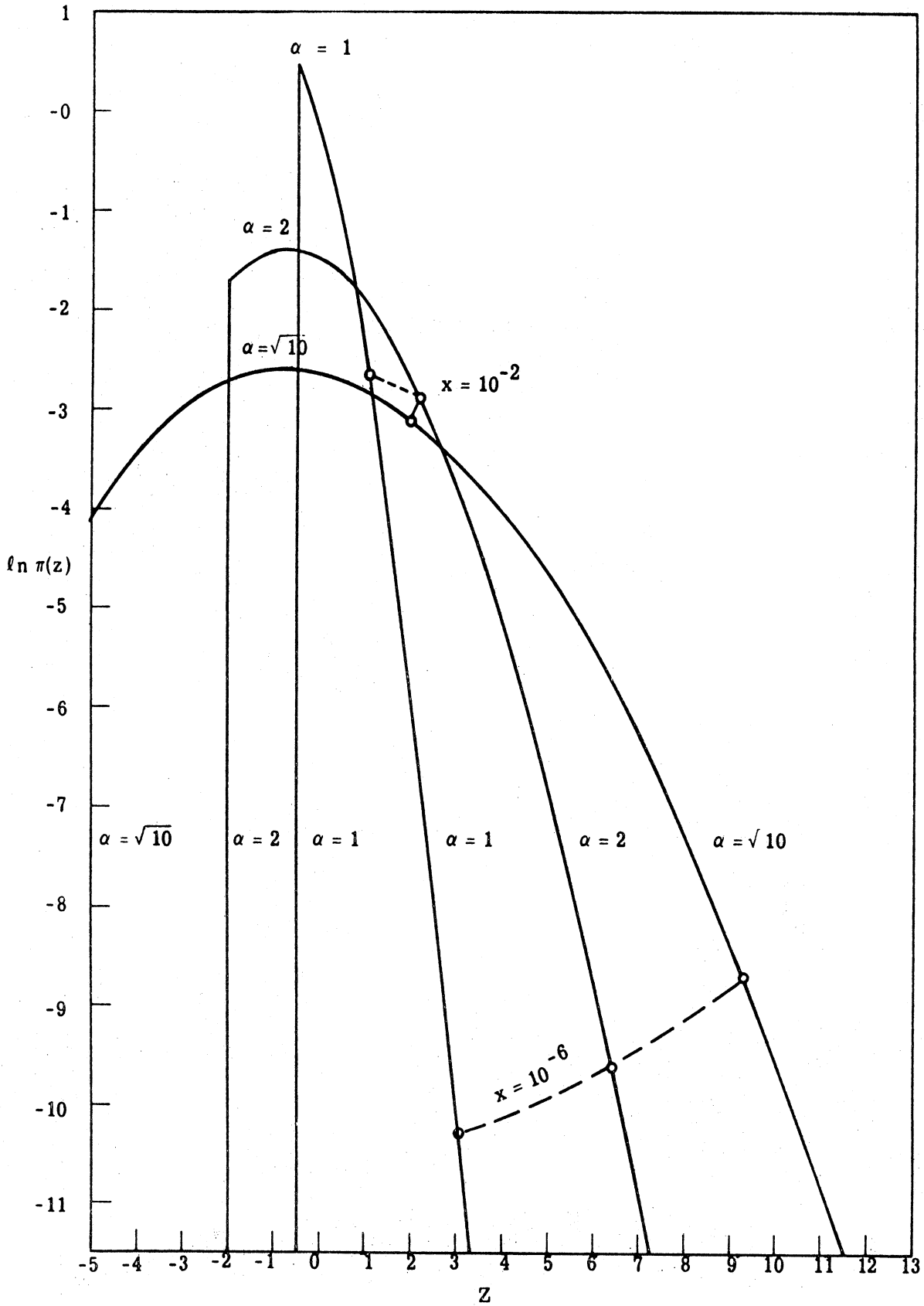
Fig. 4.7. Q-table ROC characters

$d = (\alpha - .5)^2$ has been given. Some have used the approximation

$d = \ln I_0 (\alpha^2)$. Both of these have been useful in engineering application

to indicate that performance is very poor for α less than .5, and

essentially normal for α greater than $\sqrt{10}$. For those interested in

the ROC curve between false alarm probabilities of 10^{-6} and 10^{-2} the

picture is simpler, since it is evident from Fig. 4.7 that over this

region the arcs of the logarithm of the ROC character could fit quite

well with a parabola. Appendix C contains the numerical work in an

attempt to fit the entire $\alpha = 2$, $\ln \pi(z)$ curve over the portion shown

in Fig. 4.7. This means that the same expression must fit near the

mode as well as along the skirt. The result is summarized in

Eq. 4.15.

$$\alpha = 2 \qquad \pi(z) \approx .25 \, e^{-\frac{1}{2} \left| \frac{z + .725}{1.625} \right|^{1.885}} \tag{4.15}$$

The numbers .725 and .25 come directly from the graphically determined

mode. The power 1.885 is close to 2.000. This accounts for the normal

appearance of the ROC curve and ROC character over short segments.

The difference between 2.000 and 1.885 accounts for the failure of

attempts to match the ROC curve or its character over major portions.

4.3 Noise-in-Noise, Same Spectrum

There are two basically different problems called noise-

in-noise in the detection literature. The first, and the one that we shall

consider here, is the case which is an increase in the power level

without any other change in the statistics of the observation. The

contrasting case, not considered here, is when the two noise pro-

cesses differ in spectrum and in autocorrelation function. This

latter case has caused considerable controversy in the literature

because one can obtain singular detection at low power levels from

seemingly innocent assumptions.

The standard problem assumes 2WT independent

samples of a Gaussian noise process are observed. Each sample

has the probability density function given by Eq. 4.16.

$$f(u_i \mid N) = \frac{1}{\sqrt{2\pi N}} e^{-\frac{u_i^2}{2N}} \tag{4.16}$$

(I apologize for the confusing notation which uses N to indicate the

cause on the left-hand side of the equation, and to represent the noise

power or variance on the right-hand side of the equation; however, this

is the standard notation in the field.) The signal to be detected, if it

were observed noise-free, would be of the same nature

$$f(u_i \mid S) = \frac{1}{\sqrt{2\pi S}} e^{-\frac{u_i^2}{2S}} \tag{4.17}$$

except that it has a power S. When the signal is added to the noise, the

resultant observations will have the probability density function given

by Eq. 4.18.

$$f(u_i \mid SN) = \frac{1}{\sqrt{2\pi(N+S)}} \; e^{\frac{u_i^2}{2(N+S)}} \tag{4.18}$$

The logarithm of the likelihood ratio of an individual point observation is obtained by dividing and taking logarithms.

$$z(u_i) = \frac{u_i^2}{2} \left(\frac{1}{N} - \frac{1}{N+S} \right) + \frac{1}{2} \ln \left(\frac{N}{N+S} \right) \tag{4.19}$$

If a number of independent observations are to be used reach a single decision, then the likelihood ratio of the total observation is the product of the likelihood ratios of the individual parts. This means that the logarithm of the likelihood ratio of the total observation will be the sum of the logarithm of the likelihood ratios of the individual observations. It follows directly from Eq. 4.19 that the sum of the z values corresponds to the sum of the squares of the individual observations.

$$u = \sum_{i=1}^{2WT} u_i^2 \tag{4.20}$$

Thus, u is a sufficient statistic. It follows directly that the log of the likelihood ratio of u is the sum of the logarithms of the likelihood ratios of the individual observations.

$$z(u) = \sum_{i=1}^{2WT} z(u_i) = \frac{u}{2}\left(\frac{1}{N} - \frac{1}{N+S}\right) + WT \ln\left(\frac{N}{N+S}\right)$$

$$z = u \frac{S}{2N(N+S)} + WT \ln\left(\frac{N}{N+S}\right) \tag{4.21}$$

Since u is the sum of the square of Gaussian random variables, its distribution is closely related to the chi-square distribution with $2WT$ degrees of freedom. Specifically, the variable u/N will have a chi-square distribution with $2WT$ degrees of freedom. This can be used to obtain the distribution for u.

$$f(u|N) = \frac{u^{WT-1} e^{-\frac{u}{2}}}{2^{WT}\,\Gamma(WT)N^{WT}} \qquad u \geq 0 \tag{4.22}$$

In a similar manner, the distribution for u under condition SN is obtained from $u/(S+N)$ being a chi-square random variable. The equation is exactly like Eq. 4.22, except that N is replaced by $(N+S)$.

The present objective is not to obtain the individual density functions to take their ratio, since we already know the likelihood ratio. The objective is to obtain the individual density functions in order to obtain their product. Since, by Eq. 4.21, u is a linear translation of Z, the ROC character is obtained from the root likelihood product for u by multiplying it by the Jacobian of transformation (the derivative of u with respect to Z.)

$$\pi(z) = \left[f(u\,N)\,f(u\,SN) \right]^{.5} \frac{du}{dz} \qquad z - Z_0 \tag{4.23}$$

From Eq. 4.21, z is bounded below by a quantity we shall call Z_0

because u is bounded below by zero.

$$Z_0 = WT \ln \left(\frac{N}{N+S}\right) = -WT \ln \left(1 + \frac{S}{N}\right) \qquad (4.24)$$

From (4.22) and (4.23)

$$\pi(z) = \frac{N^{.5WT}(S+N)^{.5WT}}{\Gamma(WT) \, S^{WT}} \left[\frac{u \, S}{2N(N+S)}\right]^{WT-1} e^{-\frac{u \, S}{2N(N+S)} \frac{2N+S}{2S}}$$

$$(4.25)$$

The next step simply is the change of variables from u to z.

$$\pi(z) = \frac{1}{\Gamma(WT)} \left[\frac{N(S+N)}{S^2}\right]^{.5WT} (z - Z_0)^{WT-1} e^{-\left(\frac{N}{S} + .5\right)(z - Z_0)}$$

$$(4.26)$$

This ROC character is a Pearson III character, a two-parameter class. When WT is small, or S/N is large, the ROC curve can be obtained by direct integration. Tables of the chi-square distribution may be used to plot the ROC curves.

$$X = 1 - P(\chi^2 = \frac{u}{N} \, \Big| \, \nu = 2WT)$$

$$Y = 1 - P(\chi^2 = \frac{u}{S+N} \, \Big| \, \nu = 2WT) \qquad (4.27)$$

When S/N is very small, but WT is sufficiently large

so that the resulting performance corresponds to normal indices of

the order of 0.5 to 50, a normal approximation to the ROC curves may

be made. The formula for the matching normal index

$$\hat{d} = (WT - 1) \left(\frac{S}{N}\right)^2 \qquad (4.28)$$

is derived in Appendix D, using the ROC character. This approximation

is not new; the use of the ROC character to obtain it is new.

4.4 Signal One of M Orthogonal Signals

Chapter IV on ROC models derived from electronics

cases concludes with a review of the M orthogonal signals case. The

situation is this. An observation

$$X = (X_1, X_2, \ldots, X_M) \qquad (4.29)$$

consists of M similar parts. If the condition is noise alone, N, the

individual parts are similarly distributed and are statistically independ-

ent

$$f_M(X \mid N) = \prod_{i=1}^{M} f_1(X_i \mid N) \qquad (4.30)$$

Therefore, the probability density for the occurrence of the total

observation is given by a product of the individual N probability

density functions. These M parts of the observation correspond to

M different subcauses, possibly M specific signals. Each possible

subcause affects only one part of the observation. Thus, if the j-th

signal is present, the probability density for the total observation is

given by

$$f_M(X \mid S_j N) = f_1(X_j \mid SN) \prod_{i \neq j} f_1(X_i \mid N) \tag{4.31}$$

The j-th signal affects only the j-th part of the observation and the rest

of the observation is distributed the same way it was under the condition

N. Rewrite Eq. 4.31 in terms of Eq. 4.30 by multiplying and dividing

by the missing factor in the large product and obtain

$$f_M(X \mid S_j N) = \ell_1(X_j) f_M(X \mid N) \tag{4.32}$$

for the probability density under the special subcondition of the j-th SN

cause. The actual SN cause is an ensemble of these subcauses, each

occurring with known probability $P(S_j N)$. The probability density

function for the total observation under the condition SN is therefore

the average probability density function (averaged over the various sub-

causes.)

$$f_M(X \mid SN) = \sum_j P(S_j N) f_M(X \mid S_j N) \tag{4.33}$$

Using Eq. 4.32,

$$f_M(X \mid SN) = \sum_j P(S_j N) \ell_1(X_j) f_M(X \mid N) \tag{4.34}$$

Since the term $f_M(X|N)$ is independent of the summation, divide

through by it to obtain the likelihood ratio for the total observation.

$$\ell(X) = \sum_j P(S_j N) \, \ell_1(X_j) \tag{4.35}$$

The logarithm of the likelihood ratio is simply the logarithm of

Eq. 4.35

$$z = \ln \sum_j P(S_j N) \, e^{z_j} \tag{4.36}$$

The physical picture of the receiver for such a situation

is one of M parallel branches, each of which computes a log likelihood

ratio on part of the observation. Each of the log likelihood ratios is

passed through an exponential nonlinearity which greatly emphasizes

large positive values. The output of these nonlinearities are averaged

together to form the total likelihood ratio of the observation. If this

total likelihood ratio is passed through a logarithmic nonlinearity, the

output will again be on a z axis. Although the value of z on any

observation will, in general, be less than the maximum individual log

likelihood ratio, the effect of the peaking function before the averaging

means that the output z will, in general, be greater than the average

value over all of the M parallel channels.

$$\sum_j P(S_j N) \, z_j \leq z \leq \max_j z_j \tag{4.37}$$

The evaluation of reasonable parameter cases for the M orthogonal signal problem has taxed many people. The difficulty is the presence of the nonlinearities before the averaging process. Some authors have felt that the presence of the averaging process would tend to make the output normally distributed. However, if one realizes that ℓ ranges over positive values, has an expected value of one for condition N, a good normal fit could be obtained only if the standard deviation of the fit is small compared to one. The direct route to obtain the distribution of z, or ℓ , is to carry through the M-fold convolution to obtain the results. This could be done in the transform plane by finding the characteristic function for ℓ, raising to a power, and transforming back. Since the likelihood ratio often fails to possess a moment generating function one suspects that a great deal of care will be necessary if any approximation or numerical work is done in obtaining the characteristic function or its inverse.

An Example : Let us consider a particular example of ROC character for the individual parts of the observations that simplify the process of obtaining the ROC character for the total observation. Consider the ROC equation given in Eq. 4.38

$$y_1 = x_1 - x_1 \ln x_1 \tag{4.38}$$

and graphed on normal-normal paper in Fig. 4.8. This does not represent particularly good observation quality, or any physical case of interest, but has been chosen because of the specific form of its

$$Y = X - X \ln X$$

$$\pi(z) = e^{1.5z - e^{z}} \quad -\infty < z < \infty$$
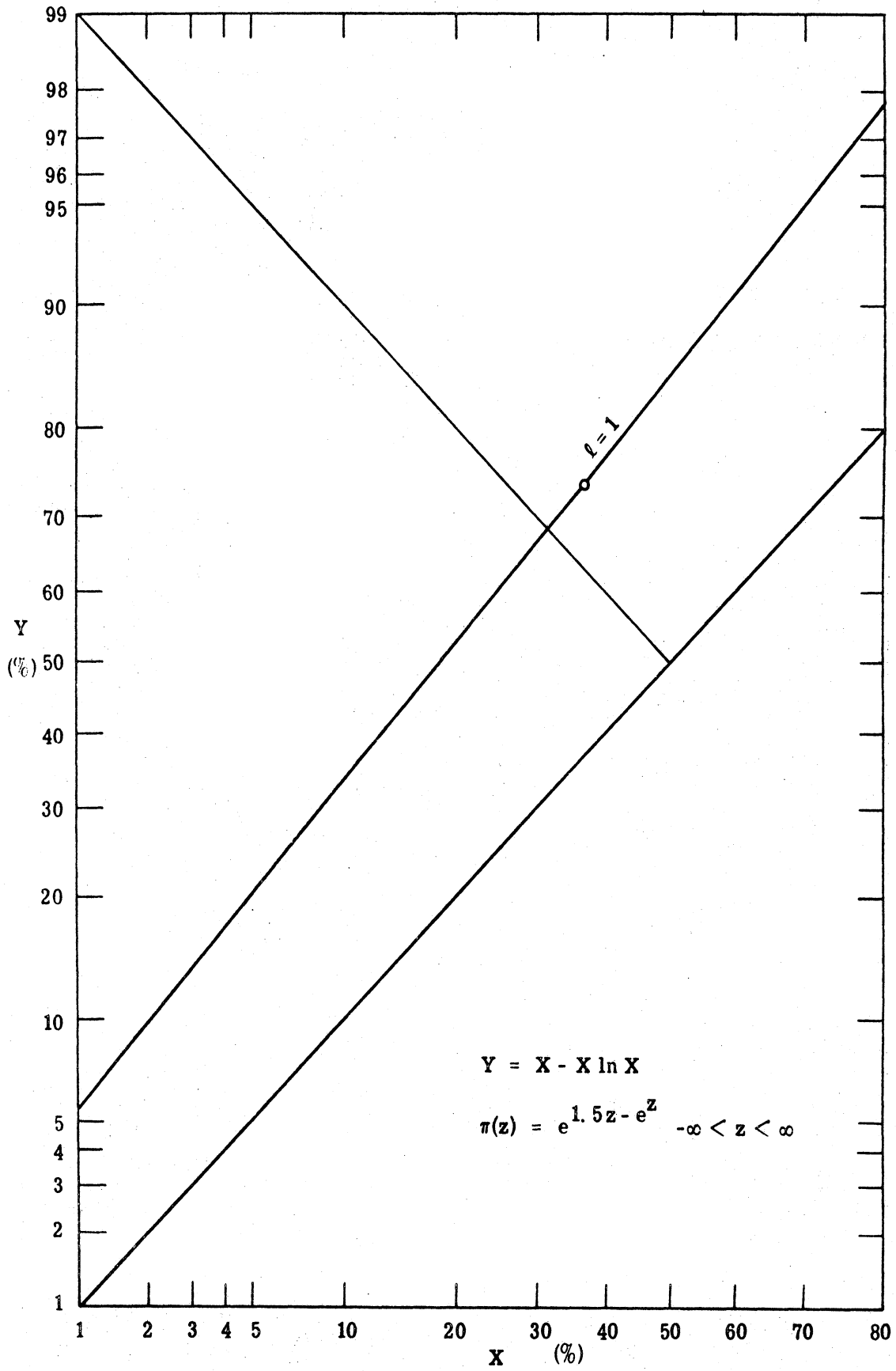
Fig. 4.8. A particular ROC curve

ROC character.

$$\pi_1(z_i) = e^{1.5 z_i} e^{-e^{z_i}} \qquad -\infty < z_i < \infty \qquad (4.39)$$

Multiply the character by $e^{-.5z}$ to obtain the probability density function for z under the condition N,

$$f_1(z_i | N) = e^{z_i} e^{-e^{z_i}} \qquad (4.40)$$

and by substitution, obtain the probability density function for the individual part-of-the-observation's likelihood ratio.

$$f_1(\ell_1 | N) = \ell e^{-\ell_i} \ell^{-1} = e^{-\ell_i} \qquad 0 < \ell_i < \infty \qquad (4.41)$$

Let

$$u = \sum \ell_i \qquad (4.42)$$

The distribution of u is obtained by simple use of the characteristic function. The resulting density function is

$$f(u | N) = \frac{u^{M-1} e^{-u}}{(M-1)!} \qquad (4.43)$$

The likelihood ratio desired is not the sum of the individual likelihood ratios, but the average likelihood ratio. Assume that the subcauses are equally likely. From Eq. 4.43 obtain

$$u = M\ell, \qquad f(\ell \mid N) = \frac{M^M}{(M-1)!} \ell^{M-1} e^{-M\ell} \qquad (4.44)$$

From the distribution for the likelihood ratio, transform to the log likelihood ratio axis

$$f(z \mid N) = \frac{M^M}{(M-1)!} e^{Mz} e^{-Me^z} \qquad (4.45)$$

and multiply by $e^{.5z}$ to obtain the ROC character for the total observation procedure.

$$\pi_M(z) = \frac{M^M}{(M-1)!} e^{(M+.5)z} e^{-Me^z} \qquad -\infty < z < \infty \qquad (4.46)$$

This ROC character is functionally similar to the single observation character.

$$\pi_1(z_i) = e^{1.5 z_i} e^{-e^{z_i}} \qquad -\infty < z < \infty \qquad (4.47)$$

Normal Case: To obtain an accurate estimate of the ROC curve when the individual observations are normally distributed, Jaarsma (Ref. 24) programmed the IBM 7090 to obtain the distributions using the characteristic function. Reproduced here are some of his ROC curves and the ROC characters for the particular case when the original observation was normally distributed with parameter $d = 4$. The ROC character for this individual observation is given by Eq. 4.48.

$$\pi_1(z_i) = e^{-\frac{1}{2}} \frac{1}{\sqrt{8\pi}} e^{-\frac{z_i^2}{16}} \qquad -\infty < Z < \infty \qquad (4.48)$$

ROC curves are for 5, 20 and 100 parallel channels of observation. The resulting performance is displayed on normal-normal paper in Fig. 4.9. The ROC characters for these three cases are drawn on linear paper in Fig. 4.10. The three characters all have the same general form. They are unimodal with the mode being at a slightly negative value of Z. The decrease from the mode is faster in the negative direction than in the positive direction. As the number of parallel channels, M, increases and the detection performance falls off the ROC character is essentially confined to a narrower range of z values and becomes correspondingly more peaked. To compare these ROC characters with those shown for Cases I and II the logarithm of the ROC character has been plotted in Fig. 4.11 to the same scale as for Fig. 4.4 and Fig. 4.7. Because of the specific mechanics of the computer program, the computation of the character was terminated at the neighborhood of the one percent false alarm values; that is, x = .01. In order to look at one of these in more detail, the axes were changed and the ROC character replotted for the case of M = 20 in Fig. 4.12. In Fig. 4.12, the z-range shown extends only from -2 to +2. To the left of the mode, the appearance of the ROC character is quite parabolic (or at least power law). The fall-off to the positive z direction is shallower, as has already been noted, and
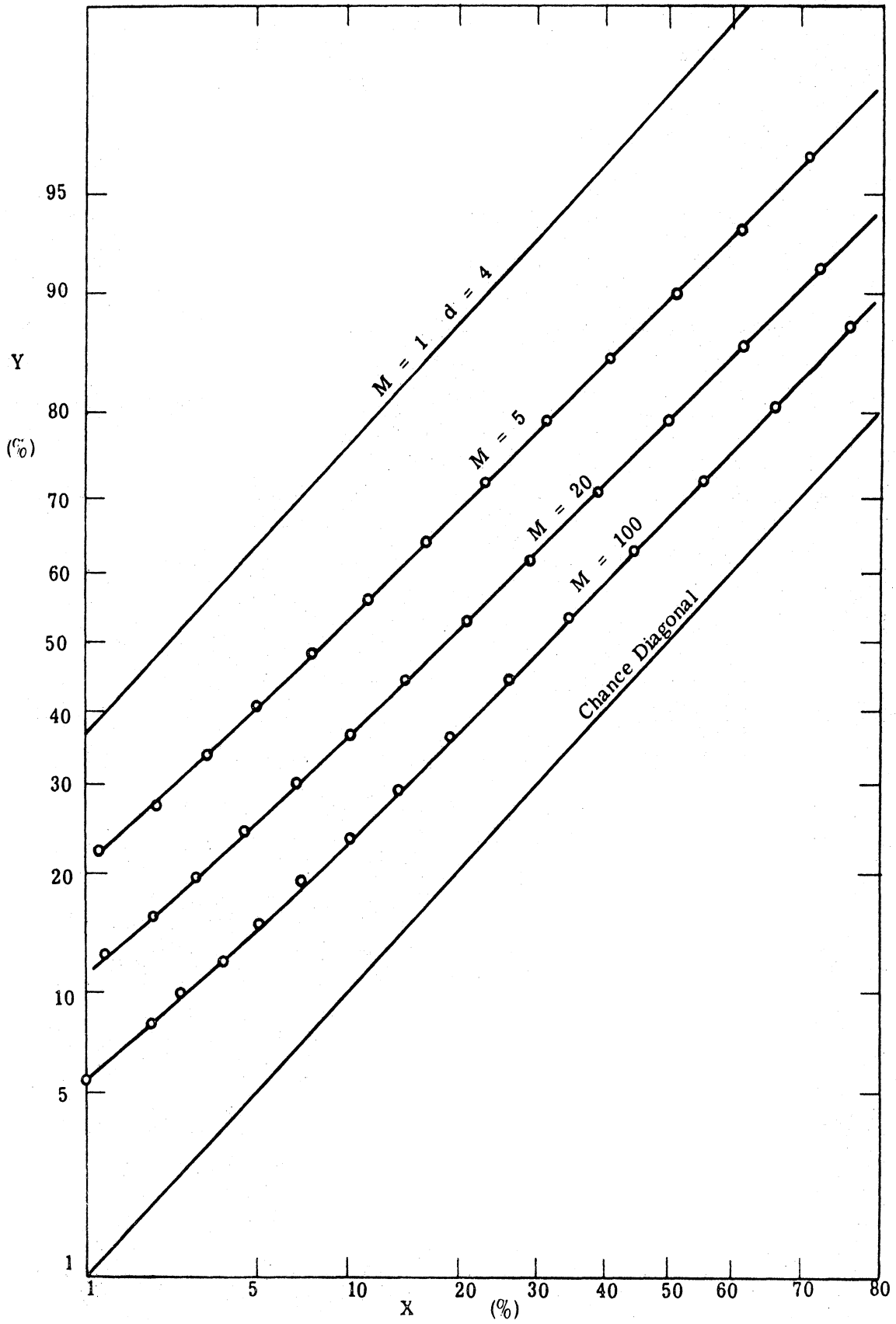
Fig. 4.9. ROC curves for M orthogonal normal signals
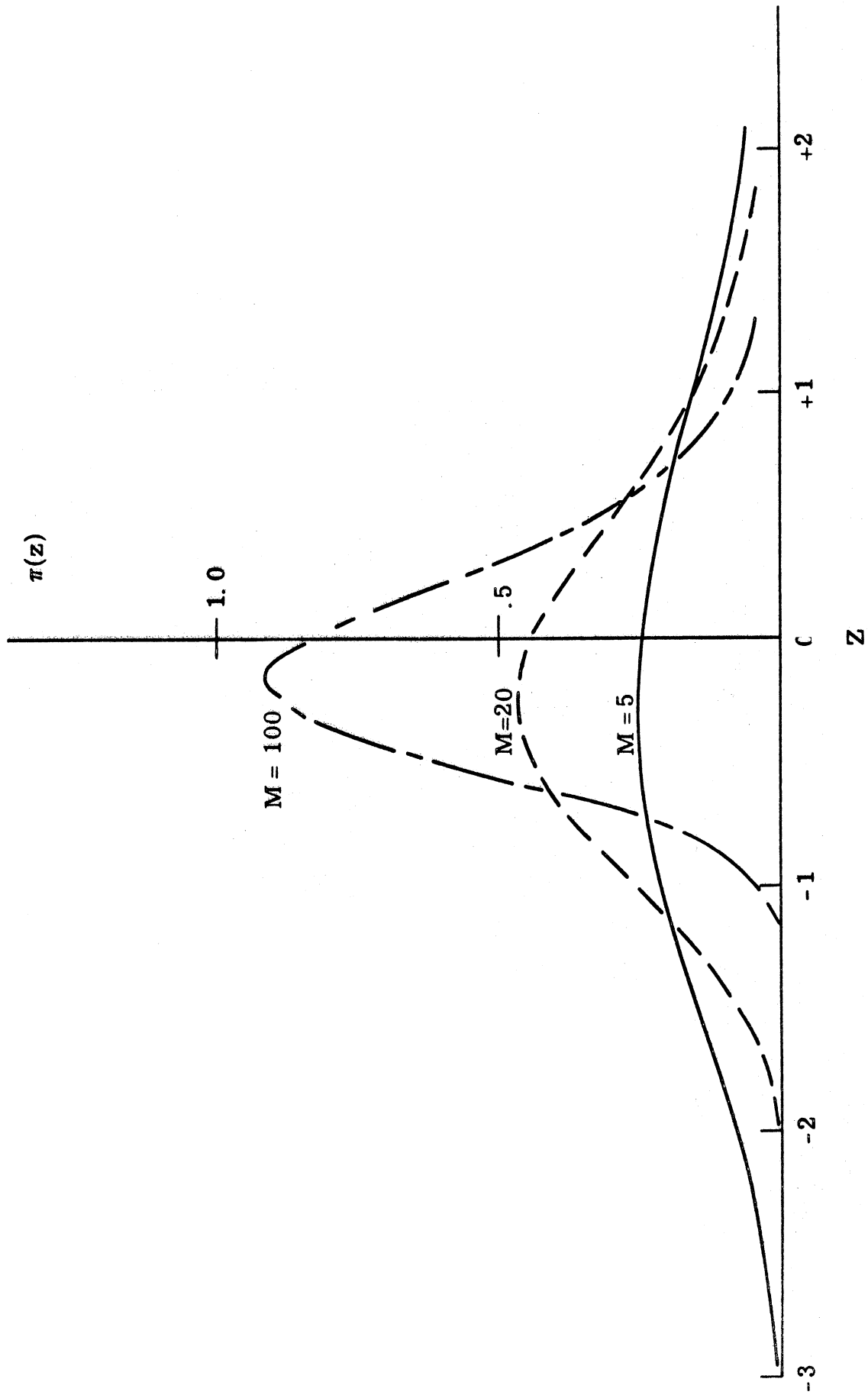
Fig. 4.10. ROC characters for M orthogonal
normal signals
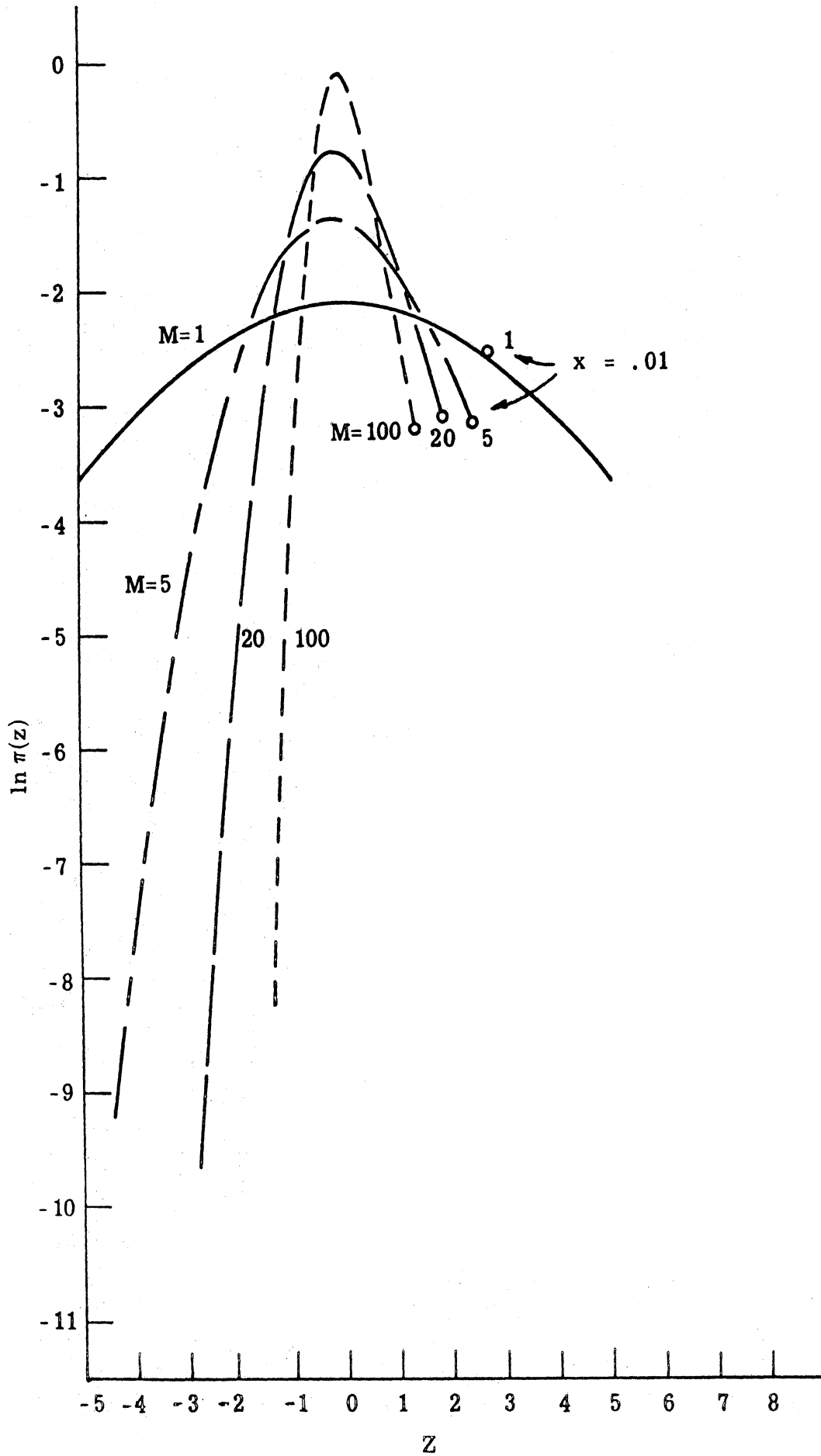
Fig. 4.11. ROC characters for M
orthogonal normal signals (log scale)

M Orthogonal Normal Signals

M = 20

$d_1 = 4$
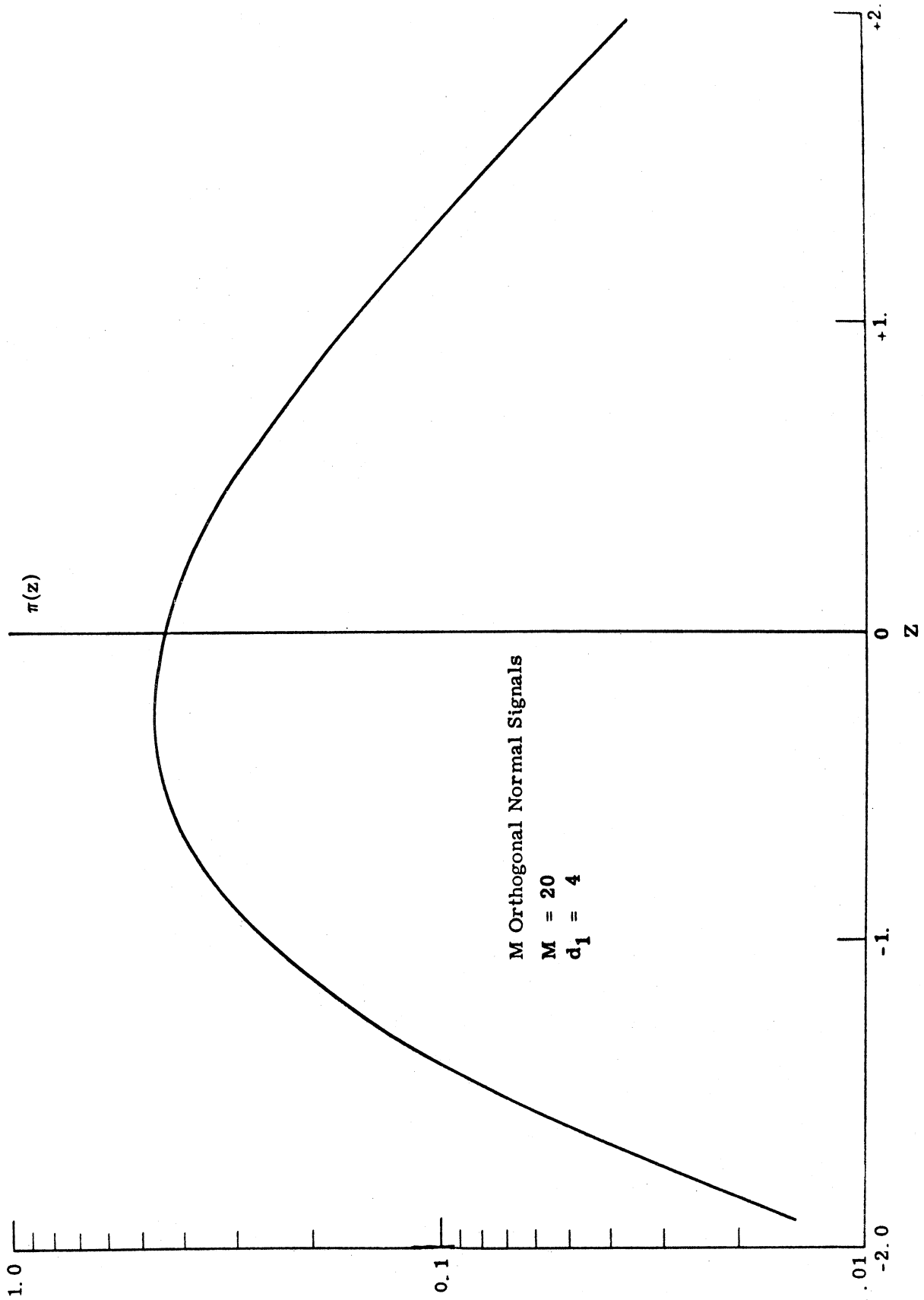
Fig. 4.12. ROC character

appears to be more linear.

In order to investigate this nearly linear fall-off of the log of the ROC character, consider what the effect would be if the ROC character behaves exponentially for large positive values. Assume

$$\pi(z) = C_1 \, e^{-C_2 z} \qquad C_1 > 0, \; C_2 > .5, \; z > C_3 \qquad (4.49)$$

holds for z greater than some value, say C_3, and with coefficients C_1 and C_2. For this simple ROC character, the ROC curve can be obtained by direct integration

$$\left(\frac{Y}{X}\right) = \int_z^\infty e^{\pm .5t} \, \pi(t) \, dt = \frac{C_1}{C_2 \mp .5} \, e^{-(C_2 \mp .5) z} \qquad (4.50)$$

Taking logarithms of both sides

$$\ln\left(\frac{Y}{X}\right) = \ln C_1 - \ln(C_2 \mp .5) - (C_2 \mp .5) z \qquad (4.51)$$

Eliminate the parameter z to obtain a nonparametric form for the ROC curve.

$$\ln X = \left(\frac{C_2 + .5}{C_2 - .5}\right) \ln Y - \left(\frac{1}{C_2 - .5}\right) \ln C_1 + \left(\frac{C_2 + .5}{C_2 - .5}\right) - \ln(C_2 + .5)$$

$$(4.52)$$

This shows that the ROC curve would plot as a straight line on log-log paper for that range for which the ROC character is essentially a simple exponential. The general form for such a curve is given by

$$\ln X = A \ln Y + \ln B \Rightarrow X = BY^A \qquad (4.53)$$

The ROC curves of Fig. 4.9 are replotted on log-log paper in Fig. 4.13. The straight-line portions of the ROC continue well beyond the initial part of the curve up to the general region of $Y = .55$. There must be a departure from this straight line form as the ROC curve approaches the point $(1, 1)$ since the ROC curve is regular, and the ROC character is definitely not exponential over its complete extent. Figure 4.13 does indicate that the exponential portion of the ROC character extends amazingly close to the mode.

A detection situation where the likelihood ratio is an average likelihood ratio, corresponding to the signal affecting only one of many parts of the observation, leads to great difficulty in treating the ROC analytically. This is in contrast to the happier situation when the logarithm of the likelihood ratio is the sum of the logarithm of the likelihood ratio for various parts of the observation, corresponding to the situation where the signal affects all of the individual parts of the observation. In this latter case, one will be dealing with the distribution for z, a variable which has concentrated density function, moment generating functions, and tends to be more amenable to transform methods than ℓ, which has wildly skewed probability distributions.
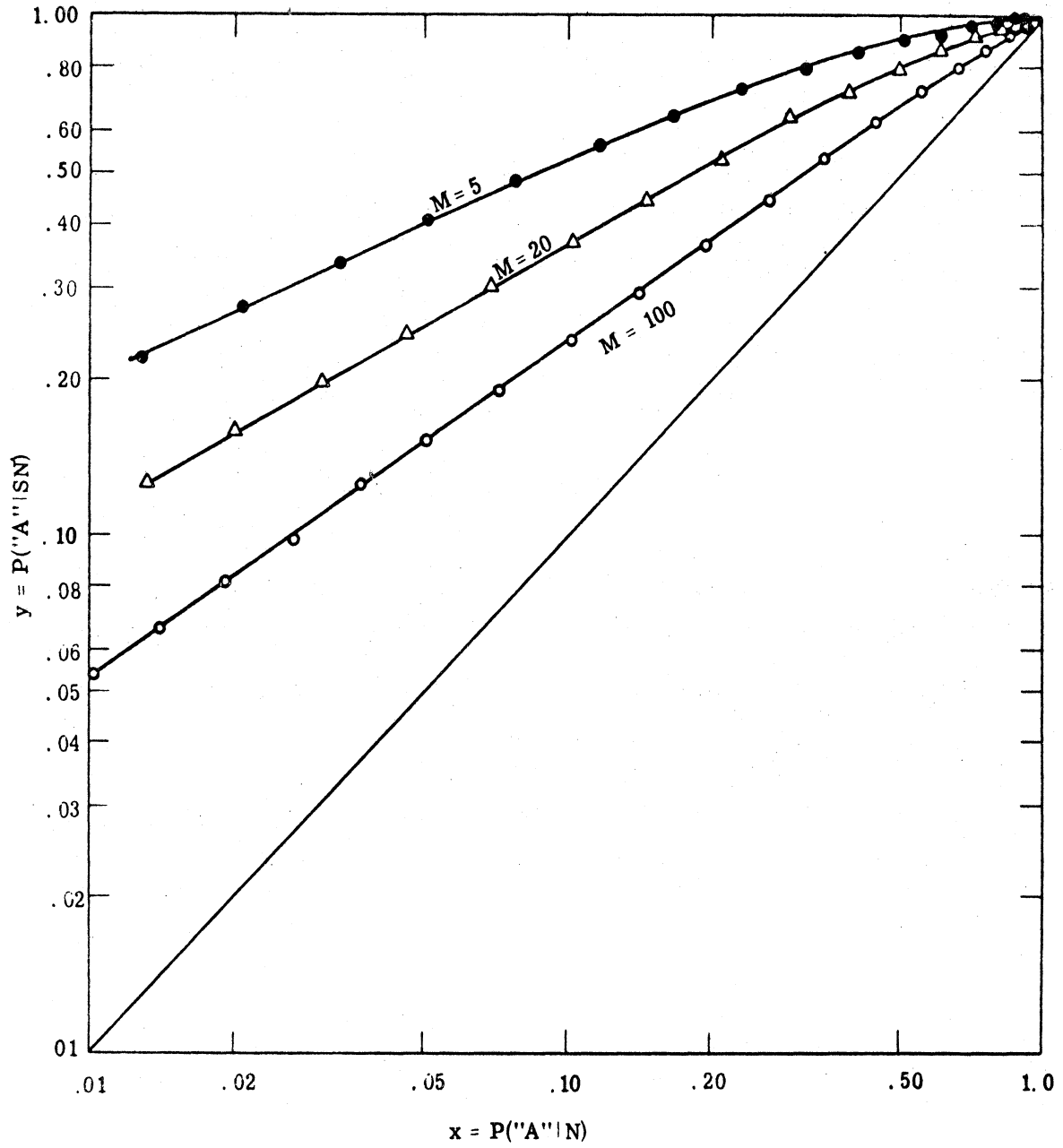
Fig. 4.13. M orthogonal normal signals, ROC for $d_1 = 4$

CHAPTER V

ALGEBRAIC ROC

This chapter is dedicated to Professor James Egan. His work with the type of ROC known as the Power ROC (Ref. 23) suggested the use of simple algebraic formulas for use in ROC models. Such models can be expected to be used to fit data, or approximate the behavior of a decision mechanism, and not necessarily arise from physical phenomena. Egan noted the nonnormal nature of ROC curves that he obtained for human observers. The region of the ROC near the (0, 0) point appeared to be fairly normal, sharply rising and curved. However, the region of the ROC near the upper corner (1, 1) was more like the pure threshold ROC curve of traditional psychophysics, with the curve approaching the upper corner at some nonzero slope. ROC curves of this nature are shown in Fig. 5. 1. These curves are Egan's pure power ROC curves, detailed in Section 5. 1.

The description of the ROC curve as an algebraic formula will mean that most of the mathematics will be simpler than that necessary when the ROC is described by higher transcendental equations. The calculus that is involved is generally elementary. The algebraic ROC curves are recommended as classroom examples since little higher mathematics is required of the student.

In this chapter Egan's pure power ROC is examined in Section 5. 1, obtaining a one parameter family of ROC curves. In 5. 2

136
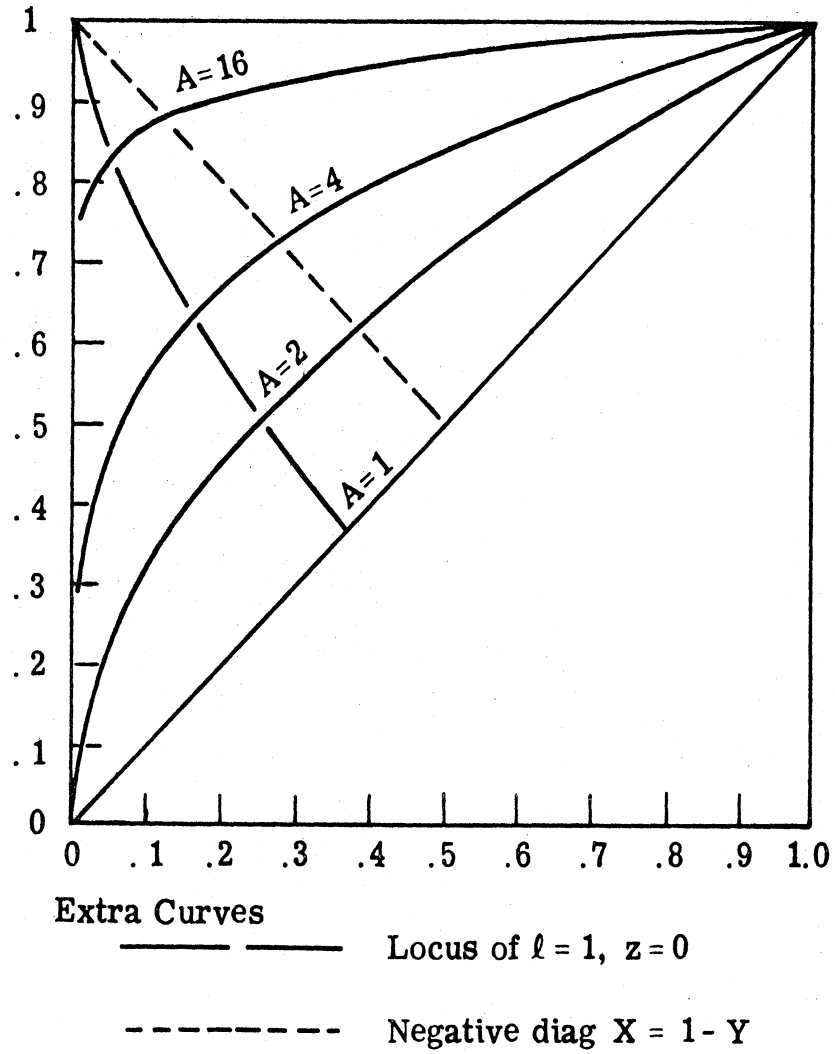
Fig. 5.1. Power ROC curves on linear graph paper;

$$X = Y^A$$

the general conic ROC curve is considered. These ROC curves are arcs of parabolas, circles, ellipses, or hyperbolas. The **Luce ROC Curve** will be considered a conic section, it being the two straight line asymptotes of a hyperbola. The regular conic ROC is a three parameter family. Section 5.3 is devoted to a nonlinear scaling which can be used to produce a special type of probability paper, called lor-lor graph paper. The Chapter concludes with Section 5.4 on curve-fitting using the conic ROC curves. While the three parameters of the conic ROC family are too cumbersome to yield a simple set of ROC, they become three degrees of freedom in fitting curves to data.

5.1 Power ROC: $X = Y^A$

The ROC curve of the pure power type, $X = Y^A$, where A is greater than one, has already been shown in Fig. 5.1 on linear graph paper. In this graphical form it is similar to many ROC curves that have been measured. This type of ROC has one property (Ref. 23) which is extremely convenient in situations where the probabilities, X or Y, can be determined except for a fixed, unknown, constant of proportionality.

For example, in an experimental situation in which the number of false alarms is measured, but the number of opportunities for false alarm can only be estimated, the false alarm probability, **X**, can not be determined exactly. If two similar experimental runs are made, or if an experimental technique is used so that two points on

the ROC can be determined simultaneously, then one can obtain the

value for A without determining the exact number of opportunities for

false alarm. If each of these two operating points fall on the power

ROC, then so does their ratio

$$X = Y^A \implies \left(\frac{X_1}{X_2}\right) = \left(\frac{Y_1}{Y_2}\right)^A \tag{5.1}$$

Once it is known that the ROC curve is a power law ROC curve, and

the parameter A determined from Eq. 5.1 using two operating points,

then the number of opportunities for false alarm can be obtained from

either one of the original ROC data points.

ROC Curves. Equation 5.1 suggests the use of log-log paper, since

the equation will plot as a straight line on such paper. Power ROC

curves with A parameter values of 1, 2, 3, 4, 8, 16, 32, 64, are

shown in Fig. 5.2, on log-log paper. In addition to the ROC curves,

the locus of the proper operating point for a symmetric bet, $z = 0$,

and the position of the negative diagonal points are shown. Log-log

paper may be useful when A is less than eight or ten, but the crowd-

ing on the graph will negate its usefulness for higher values of A. If

extremely low false alarm rates are being investigated, then larger

A values will be feasible. Probability of detection values less than

.80 can be plotted well on this paper.

Pure power ROC curves have been drawn on Fig. 5.3
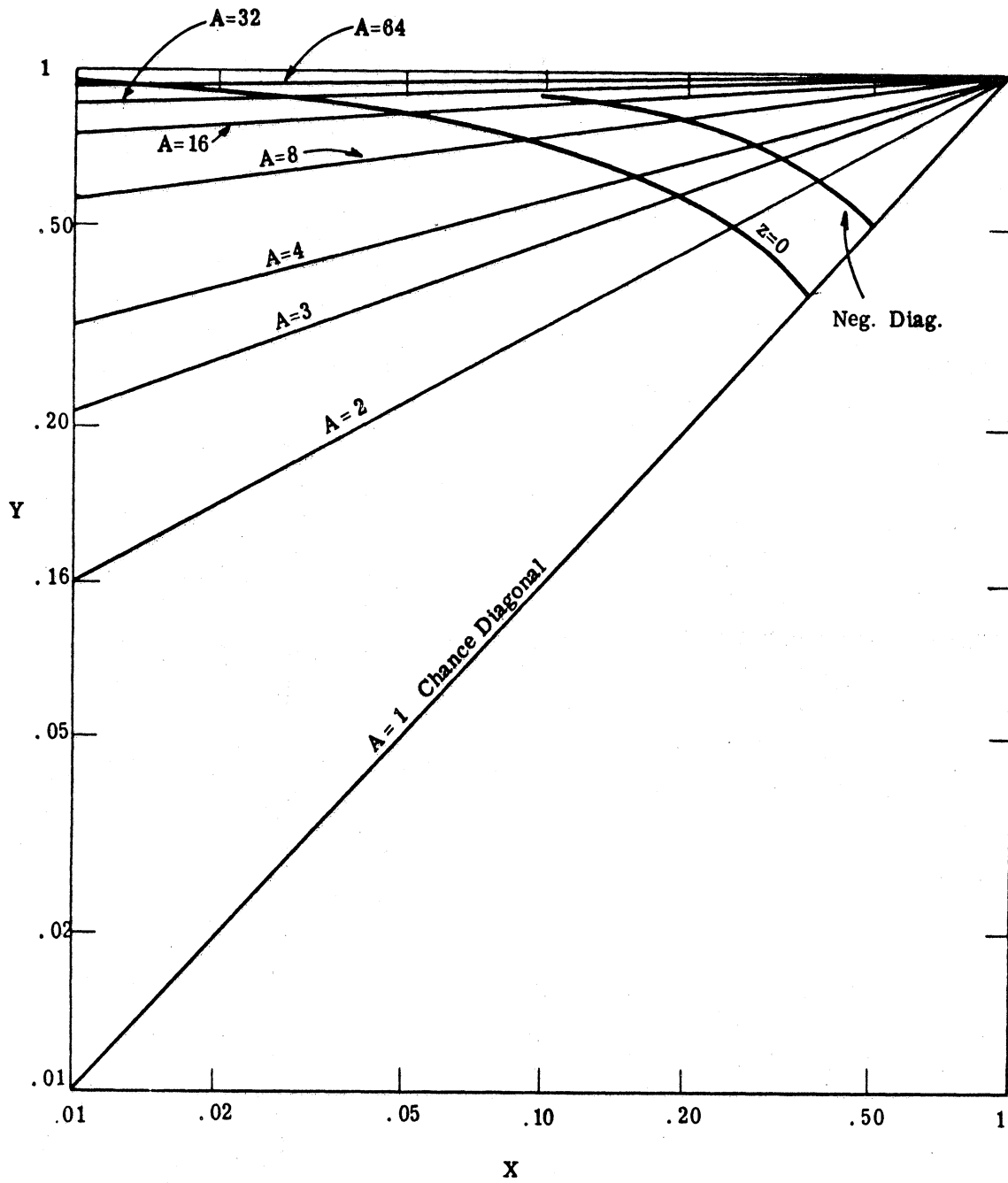
Fig. 5.2. Power ROC curves on log-log paper; $X = Y^A$
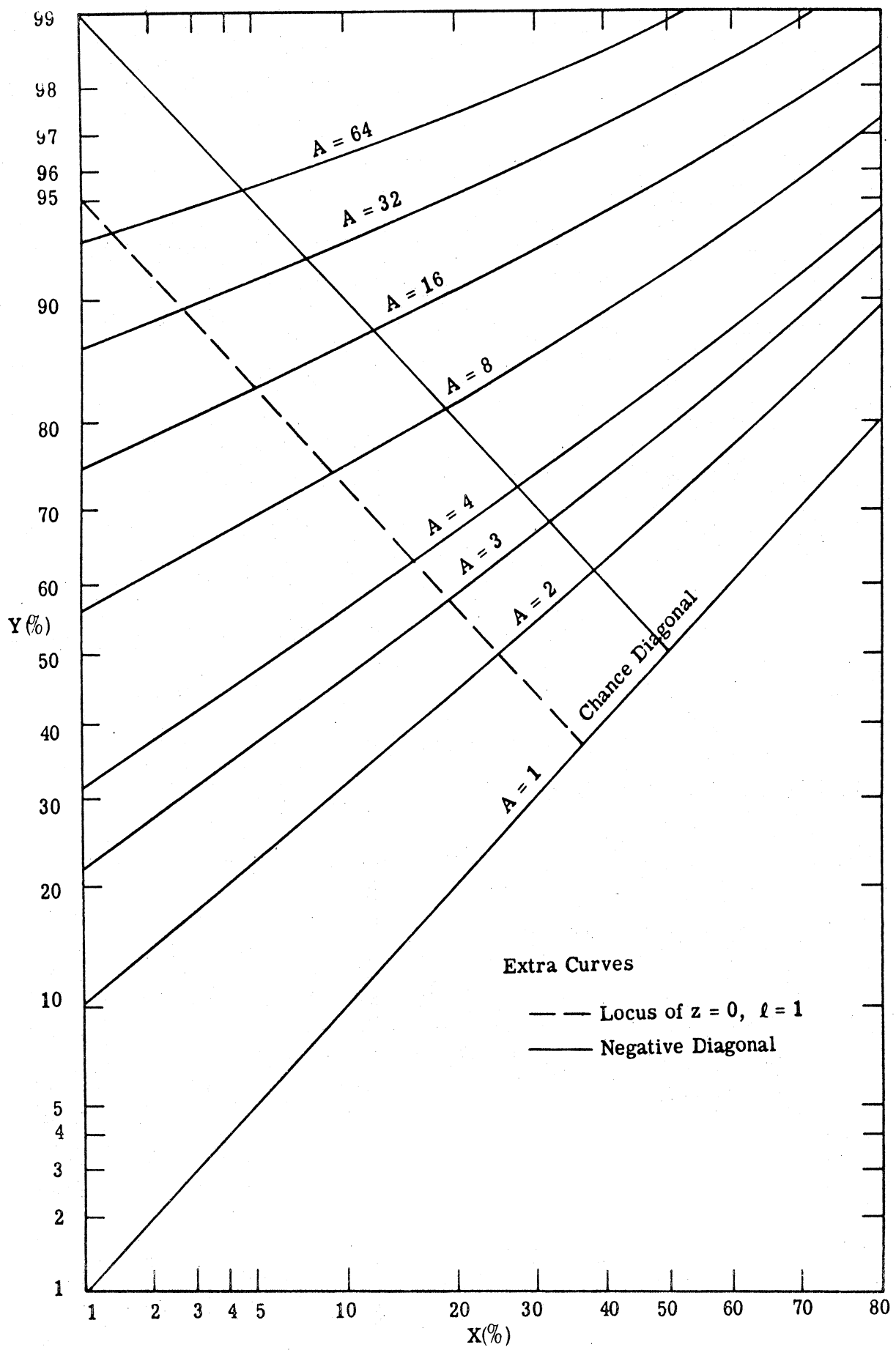
Fig. 5.3. Power ROC curves on normal-normal paper; $X = Y^A$

on normal-normal paper. These curves are "amazingly straight."

The crowding as the value of A is doubled is not as pronounced as

on log-log paper. If one had experimental data lying along one of

these pure power ROC curves plotted on normal-normal paper in the

range shown, one might well come to the conclusion that the "true"

ROC curve was a straight line on normal-normal paper, but with

slope less than one. The strict interpretation of such truly straight

lines on normal-normal paper will be considered in Chapter VI on

binormal ROC curves. The negative diagonal on linear paper is also

the negative diagonal on normal-normal paper. The locus of the unity

slope values, where $z = 0$, plot along a line almost straight and par-

allel to the negative diagonal. If a human observer had a pure power

ROC curve, and operated optimumly in a fair bet yes-no detection

experiment, he would operate somewhere along this $z = 0$ line and

the experimenter would conclude that the observer "had a bias toward

the 'B' decision."

A common procedure in both psychophysical work and

in binary communication work is to relate performance along the neg-

ative diagonal to either the index of the curve, to the signal to noise

ratio of the presentation, or to some other indicator of goodness of

performance. Along the negative diagonal the probability of a miss,

$1 - Y$, is the same as the probability of a false alarm, X. For the

pure power ROC curve with index A the probability of detection along

the negative diagonal will satisfy Eq. 5. 2.

$$1 - Y = Y^A \qquad\qquad (5.2)$$

Equation 5.2 relates the index A to the ROC coordinates on the negative diagonal. This was solved numerically and the results plotted in Fig. 5.4.

Another common practice is to convert the negative diagonal point to a normal index of detectability. This index is labeled d_e because it is related to the point of equal error. Figure 5.5 is a plot of the power ROC index A vs. $\sqrt{d_e}$. Although there is some curvature in this graph, a first approximation is that the logarithm of A is roughly proportional to the square root of d_e. For low index of detection, one might use the approximation that $\sqrt{d_e} = 2 \log_{10} A$. Although such practices of reporting a normal index of detection for nonnormal ROC curves may be commendable because it increases general communication among experimenters using different models, it carries with it the risk that some uninformed people will assume that the ROC curve is truly normal. The index d_e should always be accompanied by the statement that d_e applies only to the single point on the negative diagonal. This is the practice of most user's of the index d_e.

ROC Character. Two basic properties of any ROC curve are (1) the ROC curve gives the complement of the distribution functions of the decision axis.
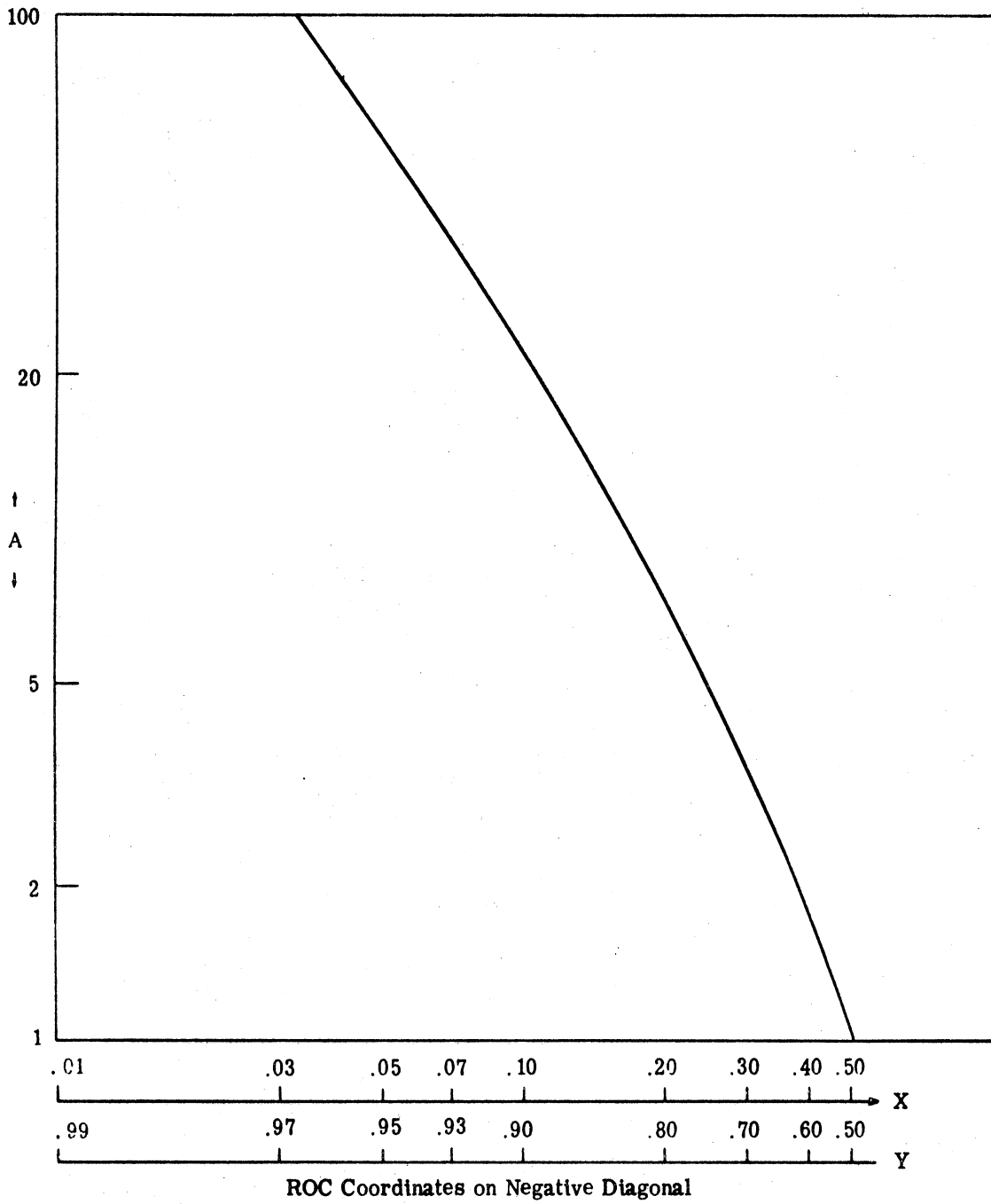
Fig. 5.4. Pure power A vs. negative diagonal
coordinates (X, Y)
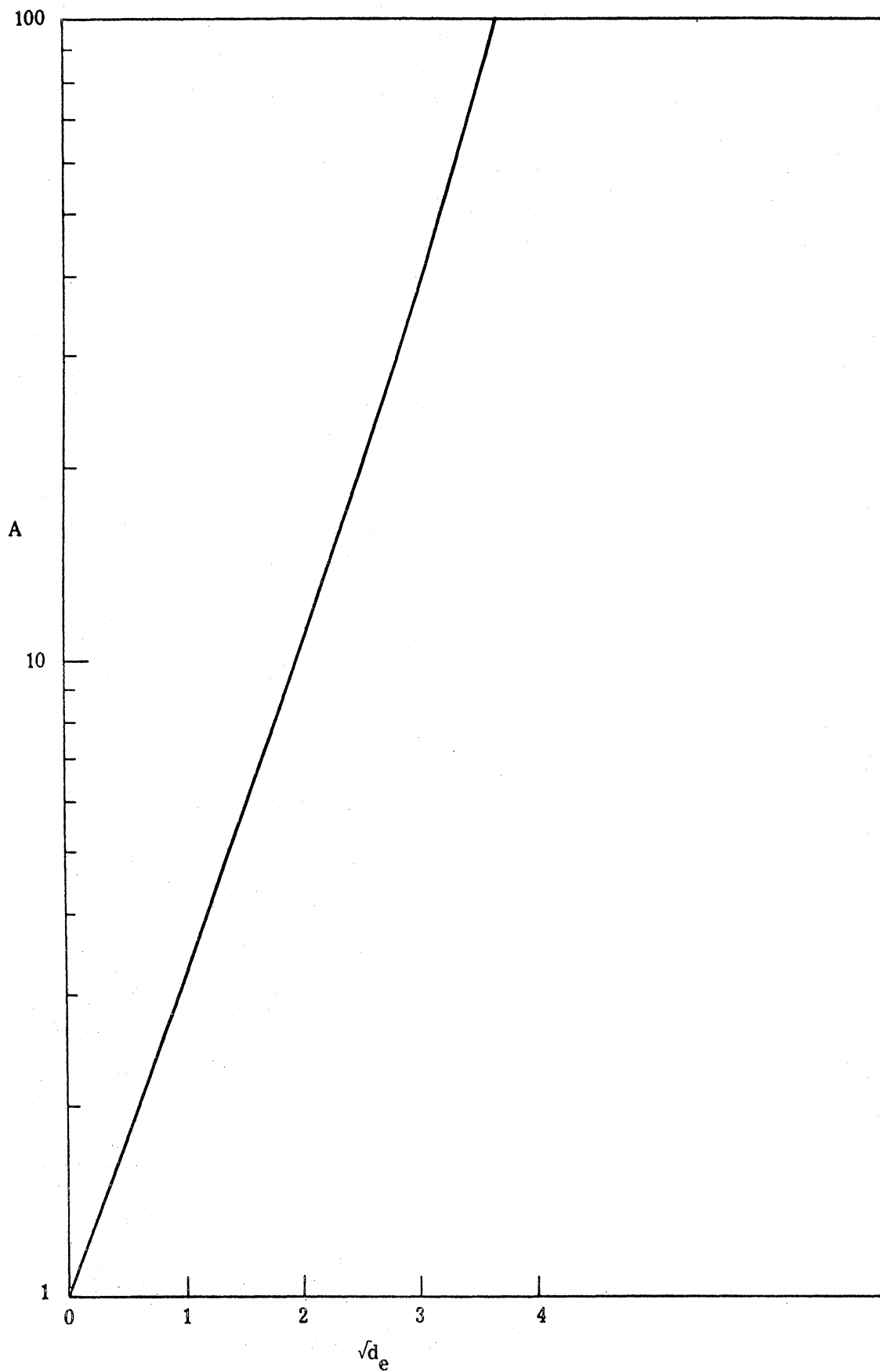
Fig. 5.5. Pure power A vs. negative diagonal normal $\sqrt{d}_e$

$$F(z|N) = 1 - X \qquad F(z|SN) = 1 - Y \qquad (5.3)$$

(2) the slope of a differentiable ROC curve is equal to the likelihood ratio cut value corresponding to that point.

$$\frac{dY}{dX} = \ell = e^z \qquad (5.4)$$

The equation for the pure power ROC is

$$X \quad Y^A, \qquad A > 1 \qquad (5.5)$$

To determine the ROC character, differentiate Eq. 5.5 with respect to X to obtain

$$1 = A \ Y^{A-1} \ e^z \qquad (5.6)$$

Since the slope of the ROC curve decreases along the ROC from $(0,0)$ to $(1,1)$, Eq. 5.6 can be used to determine the minimum and maximum values of z. The maximum value is infinity when Y is zero, and the minimum value is $-\ln A$ when Y is one. Next, solve Eq. 5.6 for Y

$$Y = \left(\frac{1}{A}\right)^{\frac{1}{A-1}} e^{-\frac{z}{A-1}} \qquad (5.7)$$

Differentiate the result with respect to z, to obtain the probability density function for z under the condition SN.

$$f(z \mid SN) = \left(\frac{1}{A-1}\right)\left(\frac{1}{A}\right)^{\frac{1}{A-1}} e^{-\frac{z}{A-1}} \tag{5.8}$$

The final step is to convert the probability density function to the ROC character

$$\pi(z) = e^{-.5z} f(z \mid SN) = \left(\frac{1}{A-1}\right)\left(\frac{1}{A}\right)^{\frac{1}{A-1}} e^{-\frac{z}{2}\frac{A+1}{A-1}} \qquad z \geq -\ln A \tag{5.9}$$

The probability density function for z under the condition N is

$$f(z \mid N) = e^{-.5z} \pi(z) = \frac{1}{A-1}\left(\frac{1}{A}\right)^{\frac{1}{A-1}} e^{-\frac{Az}{A-1}} \tag{5.10}$$

The ROC character is a simple exponential. The pure power ROC curves may also be called the simple exponential ROC curves. If z is replaced by $-z$ in Eq. 5.9 to yield

$$\pi(z) = (A-1)^{-1} A^{-1/(A-1)} e^{+\frac{z}{2}\frac{A+1}{A-1}} \qquad z < \ln A \tag{5.11}$$

another simple exponential ROC curve is obtained. It is

$$1 - Y = (1 - X)^A \tag{5.12}$$

Forced Choice and Second Choice Equations. The distributions of the decision-axis variable serve as a model for many performance measures besides the ROC curve. The following situation is used in

experimental psychophysics (Ref. 8, pp. 30-36), and corresponds also
to the digital communication case known as M-ary coding using orthog-
onal waveforms. The description will use the psychophysical ter-
minology. An observation consists of M similar portions. The SN
condition holds for one and only one of these, the other M-1 are sta-
tistically independent and due to N. When the SN condition is
equally probable in each portion of the observation and response
utility is uniform, the experiment is considered "symmetric". In a
symmetric forced-choice experiment, the observer is asked to indi-
cate his principal choice as to the portion of the observation due to
SN. In a symmetric second choice experiment he is asked to indicate
his principal and second choices.

Equations for the probability of these choices being
correct, for an observer with pure power ROC curve with index A,
can be obtained in closed form. Let $P_M(A)$ denote the probability
that the principal choice is correct. This probability will be a func-
tion of the index A. Let $S_M(A)$ denote the probability that the second
choice is correct, when the first choice is wrong.

If the principal choice is correct, then the part of the
observation due to condition SN yielded a higher decision value than
all of the M-1 values due to condition M.

$$P_M(A) = \Pr\left[\text{all (M-1) values of } z_N \text{ fall below } z_{SN}\right] \qquad (5.13)$$

The distribution function for the decision value under condition N is 1-X, and the probability density for the decision value under condition SN is the derivative of 1-Y with respect to the decision axis.

$$P_M(A) = \int_0^1 [1-X]^{M-1} \, dY \qquad (5.14)$$

When the number of presentations is two, Eq. 5.14 takes on a very simple form

$$P_2 = \int_0^1 (1-X) \, dY = \text{Area under the ROC} \qquad (5.15)$$

Since for each value of Y , the length of the horizontal line under the ROC curve at that value of Y is 1-X , the integral yields the area under the ROC curve. Equation 5.15 is known as Green's theorem, and will be discussed in Section 10.1.1. For the power ROC the two-choice probability, Green's theorem value, is simply evaluated

$$P_2(A) = \int_0^1 [1-Y^A] \, dY = 1 - \frac{1}{A+1} = \frac{A}{A+1} \qquad (5.16)$$

To obtain an expression for M > 2, factor one of the 1-X terms out of the integrand in Eq. 5.14

$$P_M(A) = \int_0^1 [1-Y^A]^{M-2} [1-Y^A] \, dY \qquad (5.17)$$

$$P_M(A) = \int_0^1 [1 - Y^A]^{M-2} dY - \int_0^1 Y^A [1 - Y^A]^{M-2} dY \qquad (5.18)$$

The first integral is simply the probability that the principal choice is correct had there been one fewer alternative. The second integral is evaluated using integration by parts. The general formula for integration by parts is

$$\int u\, dv = uv \; \Big| \; - \int v\, du$$

where the bar after the term uv indicates that this product is evaluated at the lower and upper limits as are the integrals. The appropriate choice of u and v for the second integral in Eq. 5.18 are given in Eq. 5.19.

$$u = Y \qquad dv = -Y^{A-1} [1 - Y^A]^{M-2} dY$$

$$(5.19)$$

$$du = dY \qquad v = [1 - Y^A]^{M-1} \Big/ A(M-1)$$

This choice is partly appropriate because the value of u at the lower limit zero, is zero; the value of z at the upper limit one, is zero. Therefore the term uv contributes zero to the integral.

$$\int_0^1 Y^A [1 - Y^A]^{M-2} dY = -\int v\, du = \frac{1}{A(M-1)} \int_0^1 [1 - Y^A]^{M-1} dY$$

$$(5.20)$$

Equation 5.20 substituted in (5.18) yields a recursion relation between

probabilities of the principal choice being correct for **M** and **M-1**.

$$P_M(A) = P_{M-1}(A) - \frac{1}{A(M-1)} P_M(A) \tag{5.21}$$

Collect the two coefficients of P_M

$$\left[1 + \frac{1}{A(M-1)}\right] P_M(A) = P_{M-1}(A) \tag{5.22}$$

and rewrite

$$P_M(A) = \frac{A}{A + 1/(M-1)} P_{M-1}(A) \tag{5.23}$$

Repeating Eq. 5.16

$$P_2(A) = \frac{A}{A+1} \tag{5.16}$$

Applying (5.23) once

$$P_3(A) = \frac{A^2}{(A+1)(A + 1/2)} \tag{5.24}$$

and in general

$$P_{M+1}(A) = \frac{A^M}{(A+1)(A + 1/2) \dots (A + 1/M)} \tag{5.25}$$

Each of these expressions is a simple algebraic formula which takes

on the value 1/M for A = 1 (chance) and grows monotonically to one as

A approaches infinity.

The second choice measure is

$$S_M(A) = \text{Pr} \left[\text{Second choice is correct} \mid \text{First choice is wrong} \right] \quad (5.26)$$

A conditional probability is the ratio of the joint probability divided by the probability of the condition. The probability that the first choice is wrong is simply one minus the probability that the first choice is right.

$$S_M(A) = \text{Pr} \left[\text{Second choice is correct AND First choice is wrong} \right] /$$

$$[1 - P_M(A)] \quad (5.27)$$

The event "second choice is correct and first choice is wrong" is the event that one, but only one, of the N based decision values exceeds the SN based decision value.

$$S_M(A) = \text{Pr} \left[(M-2) \text{ values of } z_N \text{ fall below } z_{SN} \text{ and one } z_N \text{ exceeds } z_{SN} \right] /$$

$$[1 - P_M(A)] \quad (5.28)$$

$$S_M(A) = \int_0^1 (M-1) [1-X]^{M-2} X \, dY / [1 - P_M(A)] \quad (5.29)$$

where the multiplying term $(M-1)$ is the number of ways of choosing which N based decision value exceeds the SN based decision value. For the power ROC curve

$$S_M(A) = (M-1) \int_0^1 [1-Y^A]^{M-2} Y^A \, dY / [1-P_M(A)] \qquad (5.30)$$

This integral was evaluated previously, Eq. 5.20. Therefore

$$S_M(A) = \frac{1}{A} P_M(A) / [1-P_M(A)] \qquad (5.31)$$

Using Eq. 5.25

$$S_M(A) = \frac{A^{M-2}}{\left[(A+1)(A+\frac{1}{2}) \cdots A + \frac{1}{M-1} \right] - A^{M-1}} \qquad (5.32)$$

It is somewhat simpler to evaluate this expression for small values of M if both numerator and denominator are multiplied by (M-1) factorial.

$$S_M(A) = \frac{(M-1)! \, A^{M-2}}{\left[(A+1)(2A+1) \cdots [(M-1)A+1] \right] - (M-1)! \, A^{M-1}} \qquad (5.33)$$

If there were an experiment with only two presentations and the signals had to be in one of them, then if the first choice is incorrect, the second choice must be correct. Equation 5.33 for M = 2 is

$$S_2(A) = \frac{1!}{[A+1] - A} = 1 \qquad (5.34)$$

In a three-choice experiment

$$S_3(A) = \frac{2A}{3A+1} \qquad \frac{1}{2} \leq S_3 \leq \frac{2}{3} \qquad (5.35)$$

As one would expect, when the value of A is one (chance) the second

choice probability is one-half. (Since the observer knows his first

choice was wrong, he is guessing between two intervals with no basis

for decision.) The nonintuitive result is that as the ROC curve be-

comes better and better (A approaches infinity) the second choice

probability does not approach unity; instead, its maximum value is

two-thirds. The probability of the principal choice being right does

approach one when the ROC index approaches infinity, however, the

second choice probability does not approach one. In a similar manner,

the second choice in a four-choice experiment is correct with proba-

bility

$$S_4(A) = \frac{6A^2}{11A^2 + 6A + 1} \qquad \frac{1}{3} \leq S_4 \leq \frac{6}{11} \qquad (5.36)$$

which takes on the chance value of one-third when the ROC is the

chance ROC, but takes on the nonunity limiting value of 6/11 when the

ROC index approaches infinity.

Although a great many equations have been displayed here,

they are all quite simple and easily obtained. This is the nice

feature of an algebraic ROC curve. Such computations could proceed

ad infinitum, since once the ROC character has been specified, per-

formance of the decision device under many different circumstances

can be obtained.

5.2 Conic ROC

This section considers a somewhat more complicated algebraic ROC curve. Since the formula for the ROC will be an algebraic formula, the same general procedure for determining the ROC character from the equation for the ROC curves shall be used. Although this will be considerably more difficult than was the case for the pure power ROC, the work is generally much easier than when the equation for the ROC is not algebraic. The purpose of this section is twofold: (1) To demonstrate again the relation between the ROC curve and the ROC character by obtaining the character from the curve, and (2) to develop the three-parameter class of conic ROC curves for their own right.

5.2.1 From Curve to Character.

The general technique for obtaining the ROC character from the equation for the ROC curve has basically five steps.

(1) Obtain an explicit formula for $y(x)$.

(2) Differentiate this expression with respect to x, substituting e^z for the derivative of y with respect to x.

(3) Invert the equation to find $x(z)$; the distribution function for z under condition N is $1 - x(z)$.

(4) Differentiate this expression with respect to z to obtain the probability density function for z under condition N.

(5) Multiply this density function by $e^{.5z}$ to obtain the ROC character.

The conic section in the plane is the locus of those points which can be described by a second order algebraic equation.

$$Ay^2 + 2Bxy + Cx^2 + 2Dy + 2Ex + F = 0 \qquad (5.37)$$

A regular conic section ROC curve must go through the point $(0, 0)$ and the point $(1, 1)$. This establishes two conditions on the coefficients in Eq. 5.37

$$(0, 0) \Rightarrow F = 0, \qquad (1, 1) \Rightarrow A + 2B + C + 2D + 2E = 0 \qquad (5.38)$$

It is convenient to use the first condition explicitly but to reserve the use of the second condition until the end of the development. In order to solve for y explicitly in terms of x, rewrite Eq. 5.37 as a quadratic equation in y

$$Ay^2 + 2(Bx + D) y + (Cx^2 + 2Ex) = 0 \qquad (5.39)$$

A form of the quadratic equation solution, which may be unfamiliar to many, is

$$a \neq 0, \qquad a\xi^2 + 2b\xi + c = 0 \Rightarrow a\xi = -b \pm \sqrt{b^2 - ac} \qquad (5.40)$$

This is used in the present development because it eliminates the factors of two and four involved in the more common equation for the solution of a quadratic equation. For the moment assume that the

coefficient A is nonzero and apply the standard solution.

$$A \neq 0, \quad Ay = -Bx - D \pm [B^2x^2 + 2B\,Dx + D^2 - ACx^2 - 2AEx]^{.5} \quad (5.41)$$

$$Ay = -Bx - D \pm [(B^2 - AC)\,x^2 + 2(BD - AE)\,x + D^2]^{.5} \quad (5.42)$$

Two collections of terms appear in Eq. 5.42. The first is the geo-
metric discriminant $B^2 - AC$, the second is the coefficient of x.

Another collection of terms which will appear later in the development

is called simply K. Let

$$\Delta = B^2 - AC, \quad \Gamma = BD - AE, \quad K = CD^2 + AE^2 - 2BDE \quad (5.43)$$

Using these new symbols Eq. 5.42 becomes

$$Ay = -Bx - D \pm [\Delta x^2 + 2\Gamma x + D^2]^{.5} \quad (5.44)$$

The first basic step is complete, an expression for y explicitly in

terms of x. Differentiate both sides of the equation with respect to

x and set dy/dx equal to e^Z.

$$Ae^Z = -B \pm .5[\Delta x^2 + 2\Gamma x + D^2]^{-.5}[2\Delta x + 2\Gamma] \quad (5.45)$$

The third basic step is to obtain an equation for x explicitly in terms

of z. Cancel some two's and move B to the other side of the equation.

$$(Ae^Z + B) = \pm[\Delta x^2 + 2\Gamma x + D^2]^{-.5}[\Delta x + \Gamma] \quad (5.46)$$

Square both sides and multiply through by the denominator of the

right-hand side

$$(Ae^Z + B)^2 [\Delta x^2 + 2\Gamma x + D^2] = \Delta^2 x^2 + 2\Delta \Gamma x + \Gamma^2 \qquad (5.47)$$

The resultant expression is a quadratic equation in x. Collect the co-

efficients of the various powers of x.

$$\Delta x^2 [(Ae^Z + B)^2 - \Delta] + 2\Gamma x [(Ae^Z + B)^2 - \Delta]$$

$$+ D^2 (Ae^Z + B)^2 - \Gamma^2 = 0 \qquad (5.48)$$

The large factor in the coefficient for both x^2 and x is common.

This factor may be zero; however, it is a function of the variable z

and hence will be zero for only one (or two) values of the argument.

We shall exclude such values and later obtain the results for those

values by a continuity argument. The result of division by $[(Ae^Z + B)^2 - \Delta]$

is

$$\Delta x^2 + 2\Gamma x + \frac{D^2 (Ae^Z + B)^2 - \Gamma^2}{(Ae^Z + B)^2 - \Delta} = 0 \qquad (5.49)$$

The third term in Eq. 5.49 can be rewritten to obtain

$$\Delta x^2 + 2\Gamma x + \left[D^2 + \frac{D^2 \Delta - \Gamma^2}{(Ae^Z + B)^2 - \Delta} \right] = 0 \qquad (5.50)$$

The new term can be simplified by utilizing the definitions of Eq. 5.43.

$$D^2 \Delta - \Gamma^2 = D^2(B^2 - AC) - (B^2 D^2 - 2ABDE + A^2 E^2)$$

$$D^2 \Delta - \Gamma^2 = D^2 AC + 2ABDE - A^2 E^2 \qquad (5.51)$$

$$D^2 \Delta - \Gamma^2 = A(CD^2 - 2BDE + AE^2) = -AK$$

The numerator is simply $-AK$. Next examine the denominator,

$$(Ae^z + B)^2 - \Delta = A^2 e^{2z} + 2ABe^z + B^2 - (B^2 - AC)$$

$$= A(Ae^{2z} + 2Be^z + C) \qquad (5.52)$$

The denominator also contains a factor A. Utilizing Eqs. 5.51 and 5.52, canceling out the common factor of A, Eq. 5.50 becomes

$$\Delta x^2 + 2\Gamma x + [D^2 - K(Ae^{2z} + 2Be^z + C)^{-1}] = 0 \qquad (5.53)$$

This is a sufficiently simple form to apply the standard quadratic equation, (5.40), and obtain

$$\Delta x = -2\Gamma \pm [\Gamma^2 - \Delta D^2 + \Delta K(Ae^{2z} + 2Be^z + C)^{-1}]^{.5} \qquad (5.54)$$

Use Eq. 5.51 again to simplify the term $\Gamma^2 - \Delta D^2$

$$\Delta x = -2\Gamma \pm [AK + \Delta K(Ae^{2z} + 2Be^z + C)^{-1}]^{.5} \qquad (5.55)$$

This completes the third basic step, an explicit expression for x in terms of the logarithm of the likelihood ratio z. Differentiate with respect to z to obtain the probability density function for z under the condition N.

$$-\Delta\ f(z\,|\,N)\ =\ \pm\ .5\ \sqrt{K}\ [A + \Delta\ (Ae^{2z} + 2Be^z + C)^{-1}]^{-.5}\ \Delta(-1)$$

$$\cdot\ (Ae^{2z} + 2Be^z + C)^{-2}\ (2Ae^{2z} + 2Be^z) \tag{5.56}$$

The differentiation has been accomplished by the standard chain rule so that the manipulations may be traced. Cleaning up the minus signs (5.56) becomes

$$f(z\,|\,N)\ =\ \sqrt{K}\ (Ae^{2z} + Be^z)\,(Ae^{2z} + 2Be^z + C)^{-2}$$

$$\cdot\ [A + \Delta(Ae^{2z} + 2Be^z + C)^{-1}]^{-.5} \tag{5.57}$$

This expression must be simplified if it is going to be of much use. The first maneuver is to take a square root of the factor $(Ae^{2z} + \dots)$ from the third factor into the fourth factor. This eliminates the awkward expression containing the reciprocal of the square root of a factor containing another reciprocal.

$$f(z\,|\,N)\ =\ \sqrt{K}\ e^z\,(Ae^z + B)\,(Ae^{2z} + 2Be^z + C)^{-1.5}$$

$$\cdot\ [A(Ae^{2z} + 2Be^z + C) + \Delta]^{-.5} \tag{5.58}$$

The last factor, under the square root, is

$$A(Ae^{2z} + 2Be^z + C) + \Delta\ =\ A^2e^{2z} + 2ABe^z + AC + B^2 - AC$$

$$=\ (Ae^z + B)^2 \tag{5.59}$$

Eliminating the discriminant Δ yields a perfect square, and nothing could be nicer under a square root than a perfect square. The density function for z then becomes simply

$$f(z \mid N) = \sqrt{K}\, e^{z}\, (Ae^{2z} + 2Be^{z} + C)^{-1.5} \tag{5.60}$$

The fifth basic step is to convert to the ROC character by multiplying by $e^{.5z}$

$$\pi(z) = K^{.5}\, e^{1.5z}\, (Ae^{2z} + 2Be^{z} + C)^{-1.5} \tag{5.61}$$

Since both terms are raised to the 1.5 power, absorb the simple exponential into the more complicated term.

$$\pi(z) = K^{.5}\, (Ae^{z} + 2B + Ce^{-z})^{-1.5} \tag{5.62}$$

Let us summarize what has been done. A regular ROC curve was formed by passing a general conic through the point $(0,0)$ and the point $(1,1)$. The form of the equation was

$$Ay^{2} + 2Bxy + Cx^{2} + 2Dy + 2Ex = 0 \tag{5.37}$$

with the side condition

$$A + 2B + C + 2D + 2E = 0 \tag{5.38}$$

We defined two particular constants, the geometric discriminant and a coefficient K

$$\Delta = B^2 - AC \qquad K = CD^2 + AE^2 - 2BDE \qquad (5.43)$$

After straightforward manipulation the formula for the ROC character was obtained.

$$\pi(z) = K^{.5} (Ae^z + 2B + Ce^{-z})^{-1.5} \qquad (5.63)$$

In order to obtain the solution the assumption was made that A was nonzero. One could return to the original equation, set A equal to zero, and repeat the process. Indeed, the author has done just that, but the work was as involved as that above, and did obtain the same solution. Why? A continuity argument gives the answer. We assume that if A is zero either B or D is nonzero other-wise y would not appear in (5.37). Consider two curves, one for A zero and another with the same coefficients but with A nonzero. These can be made to lie arbitrarily close to each other. The char-acter, Eq. 5.63, is also a continuous function of A when either B or C is nonzero. We may therefore omit the special case A = 0 from our derivation and fill in from the solution by this continuity argument. The same argument holds for those values of z for which the expression in the denominator of Eq. 5.49 is zero.

5.2.2 Conic ROC Curves. A conic has six original parameters. Two are eliminated by the regularity conditions, and one more could be eliminated by dividing Eq. 5.37 through by any one of the constants known to be nonzero. The conics are essentially a

three-parameter system of ROC curves, and (basically) three parameters appear in the ROC character, A, B, C.

The coefficients, A, B, C, are not necessarily the most appropriate coefficients to manipulate in detection work. A more appropriate set of parameters might be based on the range of the logarithm of the likelihood ratio. Let M be the slope of the ROC curve as it approaches the point (0,0), and let m be the slope of the ROC curve as it approaches (1,1). The range of z is

$$\ln m \; < \; z \; < \; \ln M \tag{5.64}$$

M = +∞ and/or m = 0 will be allowed. Since the slope of the chance diagonal is one, the value for m is less than one, and the value for M is greater than one for nonchance regular ROC curves. Given an ROC curve in an expression such as Eq. 5.37, these terminal values for ROC slope can be obtained from Eq. 5.45. The work is trivial but tedious, the results are given below.

$$M \; = \; -E/D \qquad m \; = \; -(B + C + E)/(A + B + D) \tag{5.65}$$

In Table 5.1 are displayed parameters for various conditions of terminal slope. Table 5.1 also contains equations for the ROC curves and characters. The fifth case in Table 5.1 is the symmetric case, because the election of symmetric bounds on this type of ROC character means that the ROC curves will be symmetric.

$$\underline{M = \infty}$$

$$D = 0, \ E = -1, \ A + 2B + C = 2$$

ROC: $\quad Ay^2 + 2Bxy + Cx^2 - 2x = 0$

$$\pi(z) = \sqrt{A} \ (Ae^Z + 2B + Ce^{-Z})^{-1.5}$$

$$K = \sqrt{A}, \ \pi(0) = \sqrt{A/8}, \ m = 1 - (A + B)^{-1}$$

$$\underline{M = \infty, \quad m = 0}$$

$$B = 1 - A, \ C = A, \ D = 0, \ E = -1, \ \Delta = 1 - 2A$$

$$\pi(z) = \sqrt{A/8} \ (1 + 2A \ \sinh^2 .5z)^{-1.5}$$

ROC: $\quad A(y-x)^2 - 2x(1-y) = 0$

$$\underline{M \neq \infty, \quad m = 0}$$

$$B = -(1 + A), \ C = 2M + A, \ D = 1, \ E = -M, \ \Delta = 1 - 2A(M-1)$$

ROC: $\quad Ay^2 - 2(1 + A) \ xy + (A + 2M) \ x^2 + 2y - 2Mx = 0$

$$\underline{M \neq \infty, \quad m \neq 0}$$

$$D = 1, \ E = -M, \ A + 2B + C = 2(M-1)$$

ROC: $\quad Ay^2 + 2Bxy + Cx^2 + 2y - 2Mx = 0$

$$K = AM^2 - 2MB + C, \ m = 1 - (M-1)(A + B + 1)$$

$$\underline{M \neq \infty, \quad mM = 1}$$

$$B = M - 1 - A, \ C = A, \ D = 1, \ E = -M; \ A > -2M/(M-1)$$

$$K = A(M-1)^2 + 2M(M-1), \ \Delta = (M-1)(M-1-2A)$$

Table 5. 1.   Canonical conic equations

Relevant equations and parameter values are collected

in Table 5.2 for the case of zero and infinite terminal slopes.

	ROC Curve	A=C	Δ	B	Y on neg. diag.
Chance Diagonal	y = x	∞	$-\infty$	$-\infty$	.50
Ellipse (Long Edge)					
Circle	$y^2 + (1-x)^2 = 1$	1	-1	0	.707
Ellipse (Pointed End)					
Parabola	$2(y-x) = 1 - (x+y-1)^2$	.5	0	.5	.75
Hyperbola					
Singular	y = 1	.0	1	1	1

Table 5.2. Symmetric conics, unbounded z

When the ROC is symmetric but with bounded range the geometric

names and the corresponding values for A, B, C, D and Δ are displayed

in Table 5.3.

Figures 5.6 through 5.12 display on linear paper the

ROC curves and the ROC characters for bounded symmetric conics.

Each figure displays the complete range of geometric type from a

chance diagonal to the Luce ROC curve, all with the same terminal

slopes.

	A = C	Δ	B	Y on neg. diag.
Chance Diagonal	∞	-∞	-∞	.50
Ellipse (Long Edge)				
Circle	$M-1$	$-(M-1)^2$	0	$\sqrt{.5}\,\dfrac{\sqrt{M^2+1}-1}{M-1}$
Ellipse (Pointed End)				
Parabola	$.5(M-1)$	0	$.5(M-1)$	$\dfrac{3M+1}{4M+4}$
Hyperbola				
Rt. Hyperbola	0	$(M-1)^2$	$(M-1)$	$\dfrac{\sqrt{M}}{\sqrt{M}+1}$
Hyperbola				
Luce	$-\dfrac{2M}{M-1}$	$(M+1)^2$	$\dfrac{M^2+1}{M-1}$	$\dfrac{M}{M+1}$

Table 5.3. Symmetric conics, bounded z

$\pi(z)$

(1) Chance diagonal
(2) Circle
(3) Parabola
(4) Rt. hyperbola
(5) Luce

$y = P(A|SN)$

$x = P(A|N)$
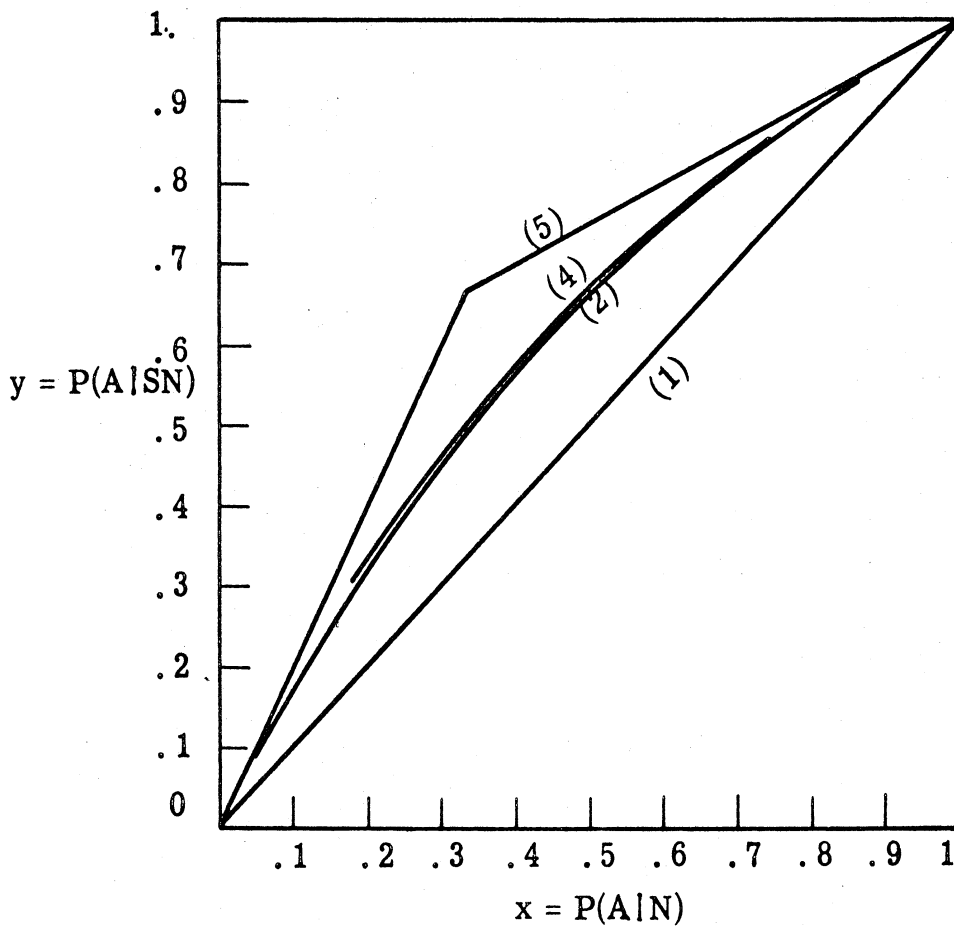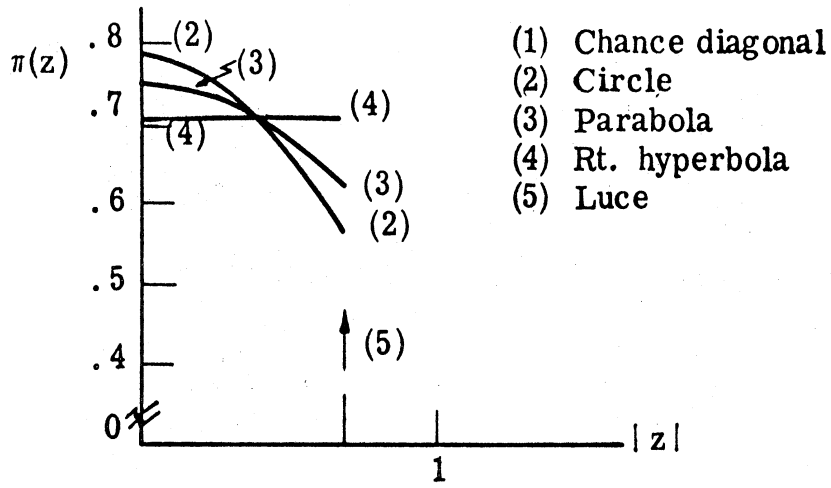
Fig. 5.6. Symmetric conics $.5 < \ell < 2$, $|z| < .693$
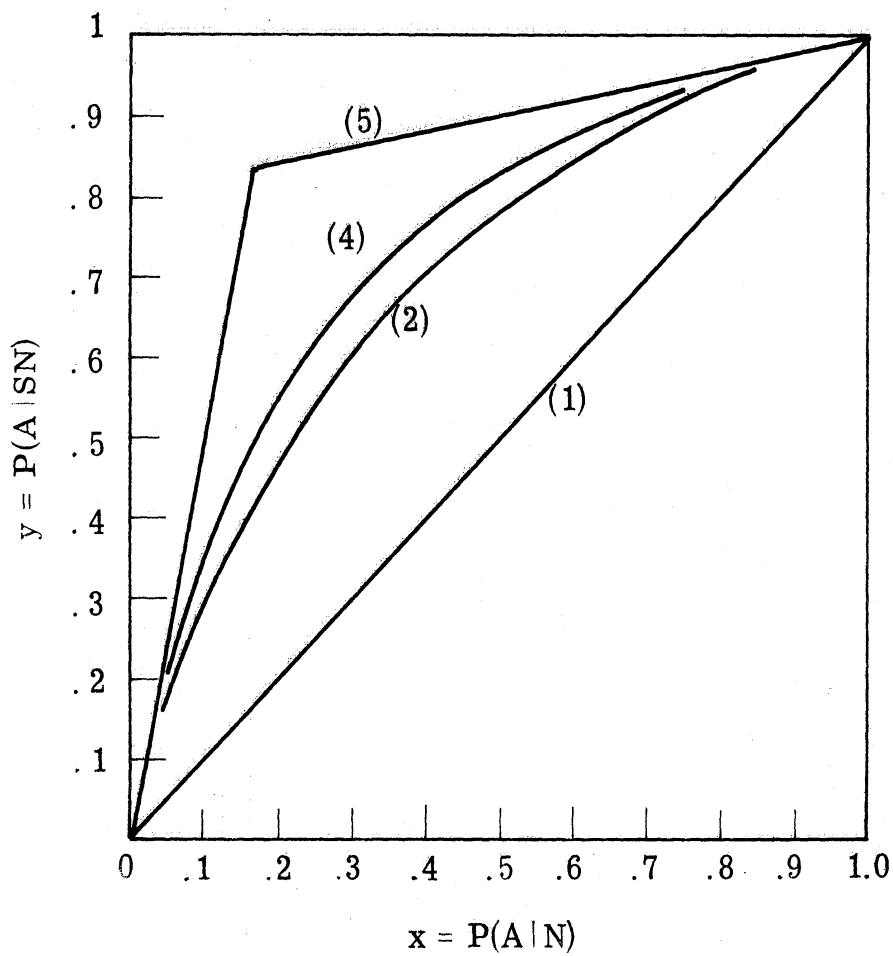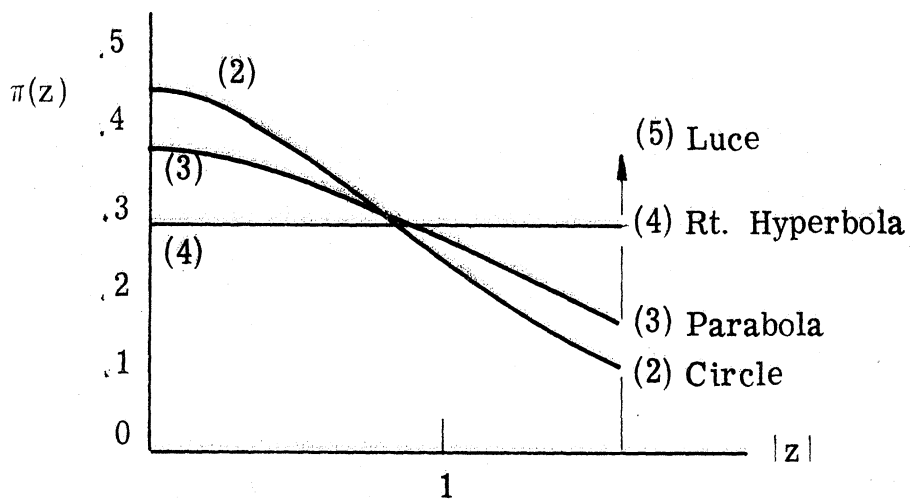
Fig. 5.7. Symmetric conics

$.2 \leq \ell \leq 5 \quad |z| < 1.61$
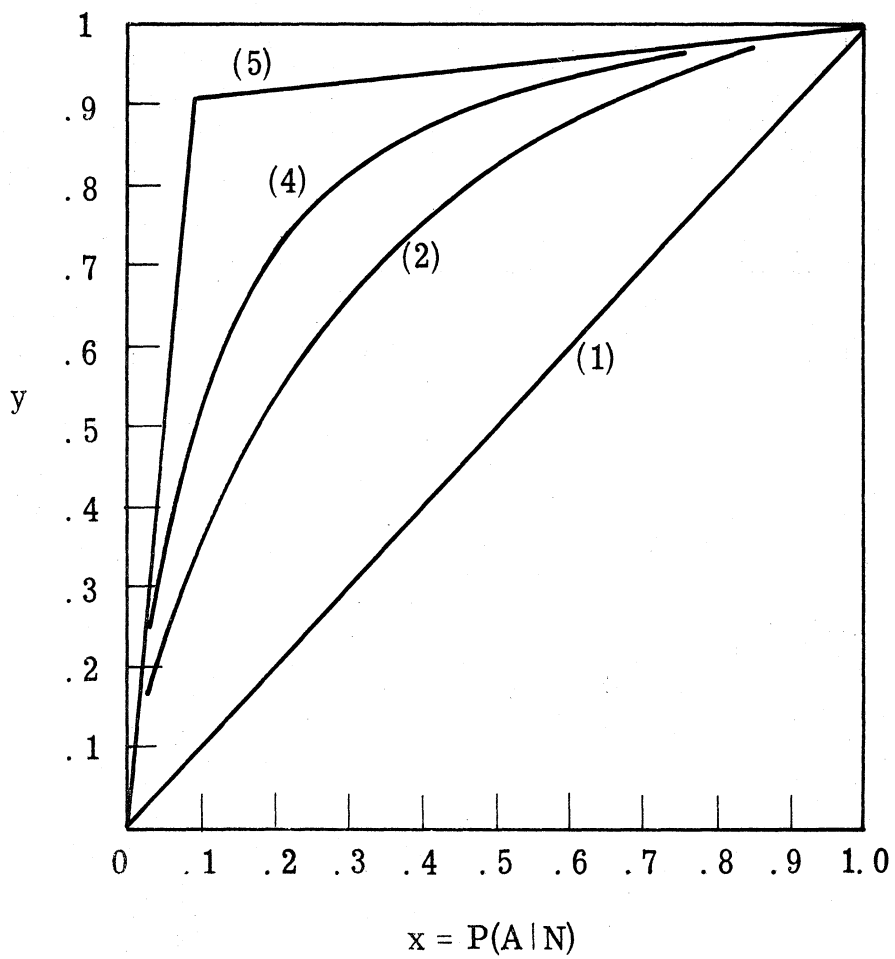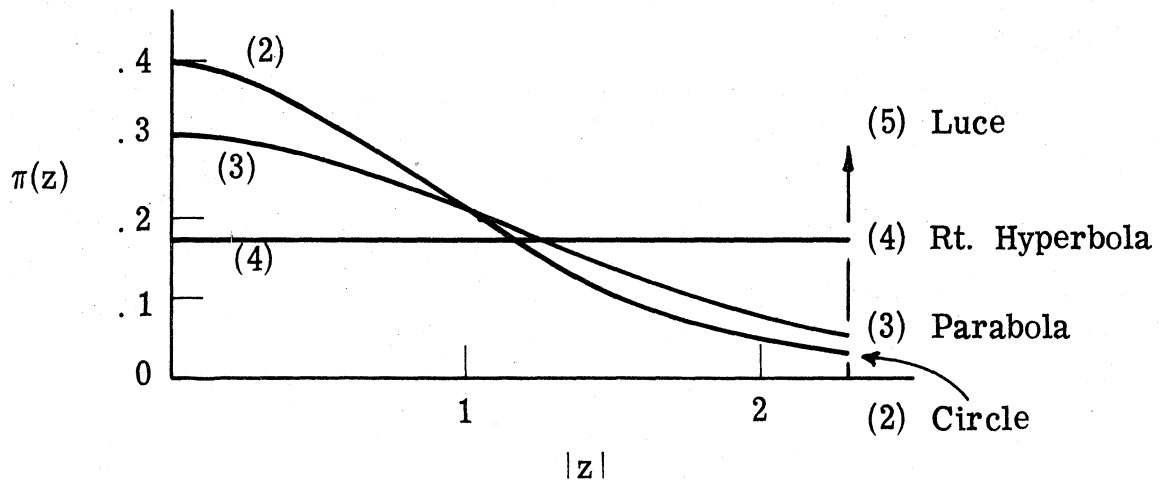
Fig. 5. 8. Symmetric conics, $.1 \leq \ell \leq 10$, $|z| < 2.30$
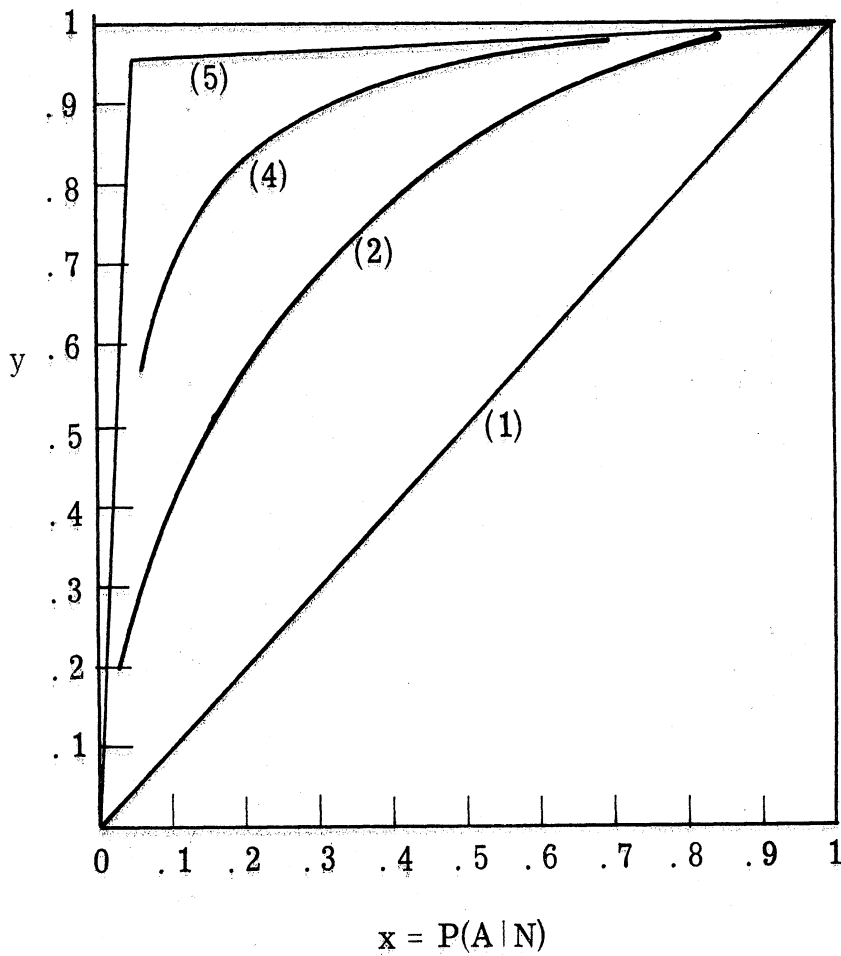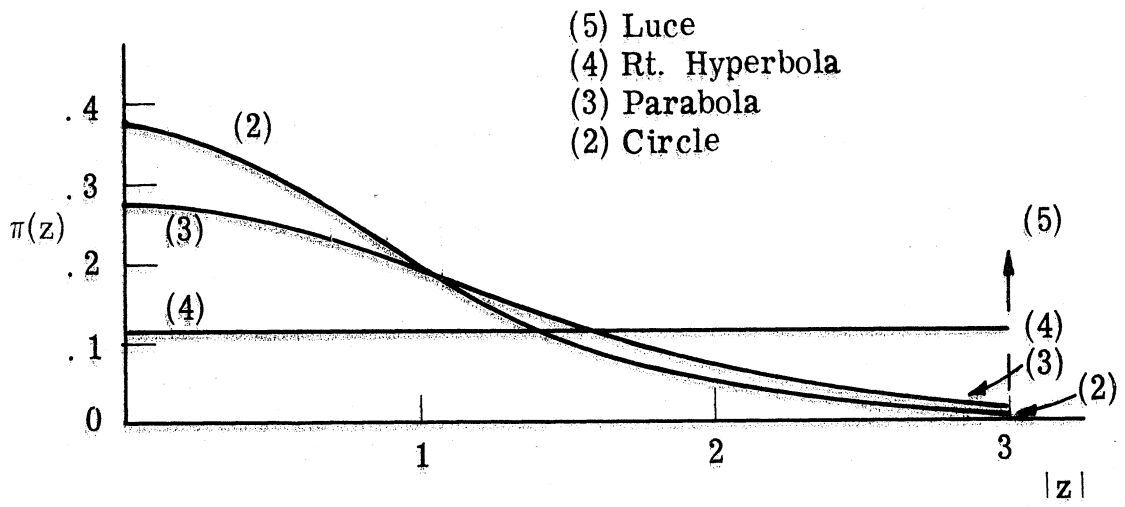
Fig. 5.9. Symmetric conics, $.05 \leq \ell \leq 20$, $|z| < 3.00$
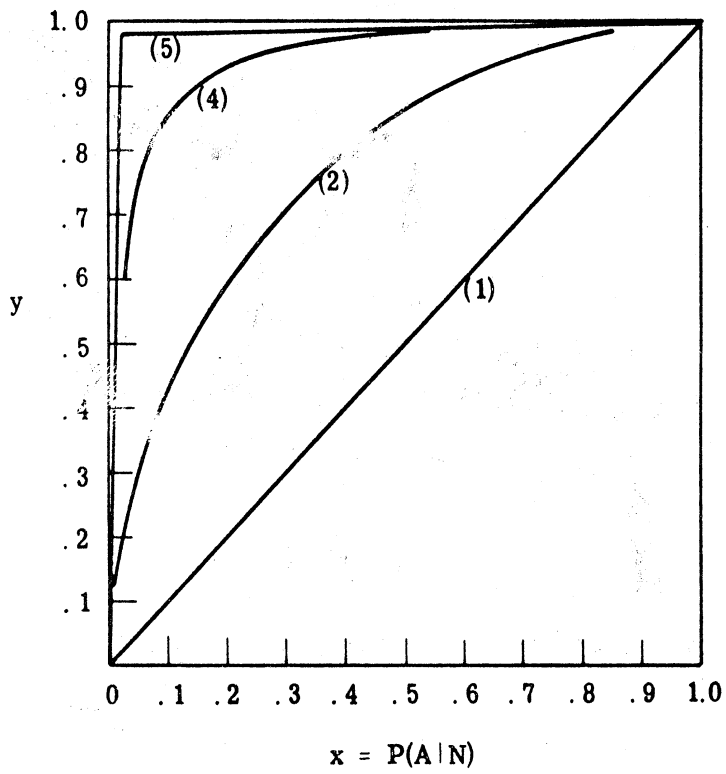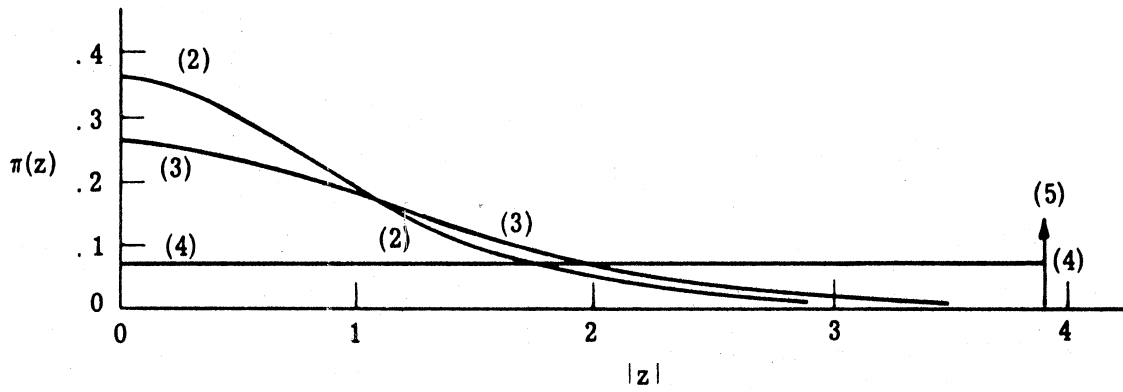
Fig. 5.10. Symmetric conics,
.02 ≤ ℓ ≤ 50 |z| < 3.91
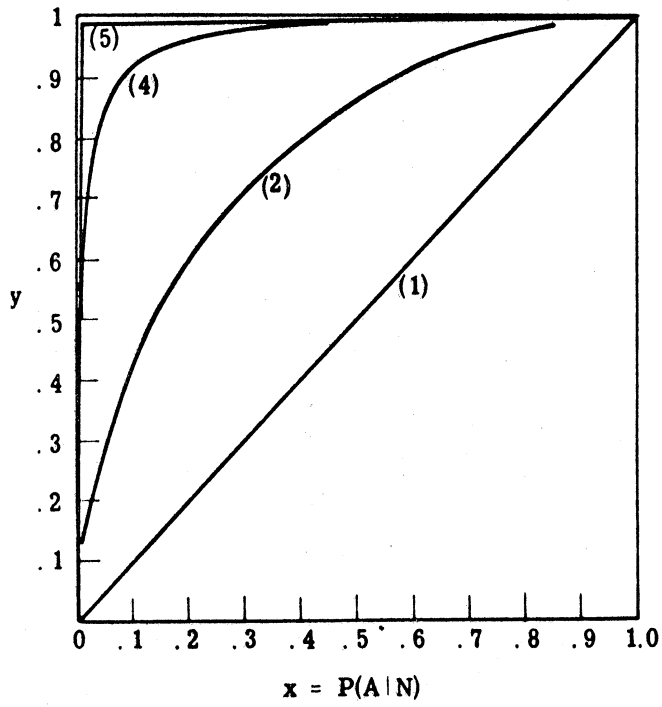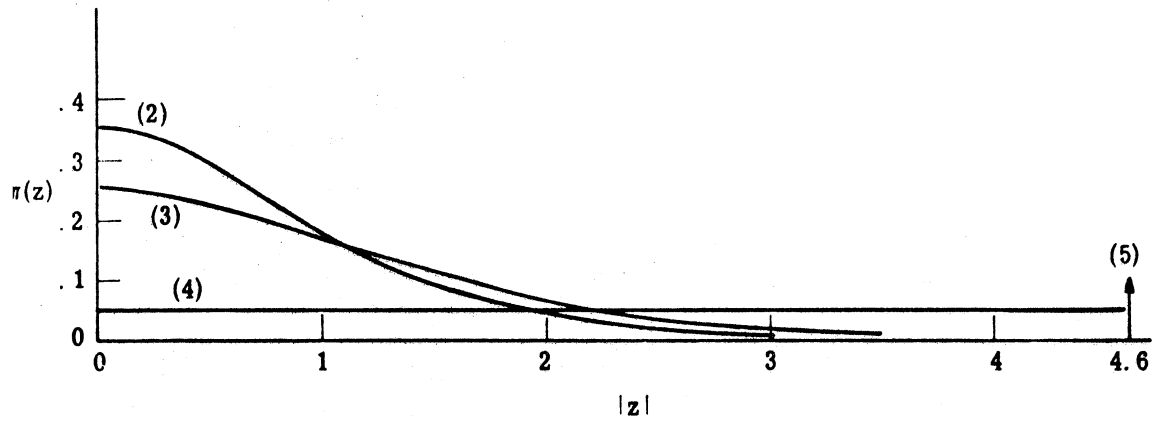
Fig. 5.11. Symmetric conics,
.0101 $\leq \ell \leq$ 99 |z| < 4.59

(1) Chance
(2) Circle
(3) Parabola
(4) Rt. Hyperbola
(5) Luce

x = P(A | N)
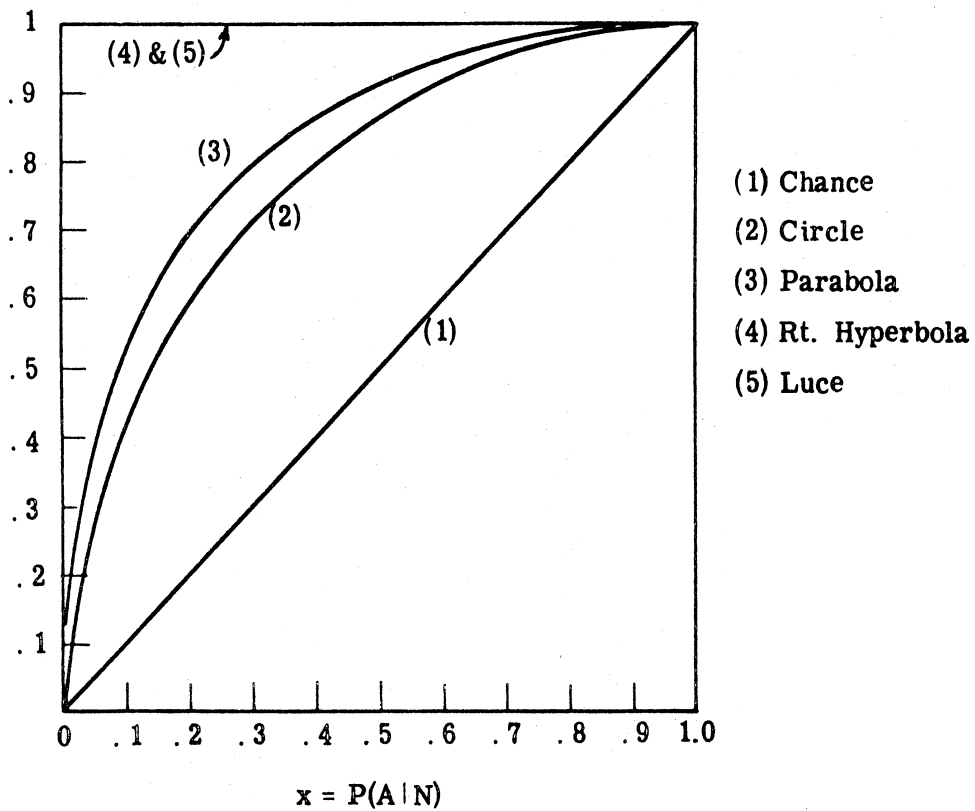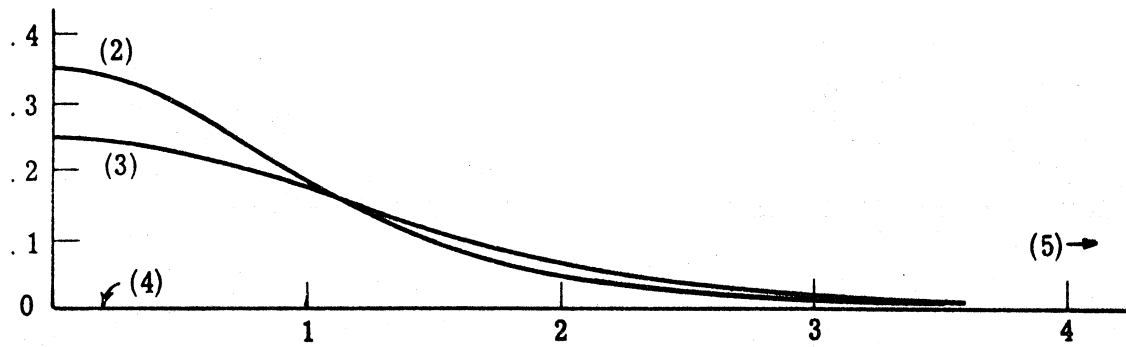
Fig. 5.12. Symmetric conics, unbounded ℓ, $|z| < \infty$

Two ROC characters, one concave and the other convex, are compared to a rectangular character. The specific cases considered are shown in Table 5. 4.

(1)	$A = C = 1,\ B = 14, D = 1,\ E = -16,\ K = 705,\ \Delta = 195,\ \Gamma = 30$
	ROC: $y^2 + 28\,xy + x^2 + 2y - 32x = 0$
	$\pi(z) = (705)^{\cdot\,5}\ (28 + e^z + e^{-z})^{-1.5}$ $\quad |z| < 2.77$
(2)	$A = C = 0,\ B = 15,\ D = 1,\ E = -16,\ K = 480,\ \Delta = 225,\ \Gamma = 15$
	ROC: $30xy + 2y - 32x = 0$
	$\pi(z) = (480)^{\cdot\,5}\ (30)^{-1.5}$ $\quad\quad |z| < 2.77$
(3)	$A = C = -1,\ B = 16,\ D = 1,\ E = -16,\ K = 255,\ \Delta = 255,\ \Gamma = 0$
	ROC: $-y^2 + 32xy - x^2 + 2y - 32x = 0$
	$\pi(z) = (255)^{\cdot\,5}\ (32 - e^z - e^z)^{-1.5}$ $\quad |z| < 2.77$

Table 5. 4. Collected equations

For the middle case the coefficients A and C are zero. Therefore the ROC character is a constant over its full range. The range has been chosen to correspond to a maximum slope of 16 and a minimum slope of 1/16. For Curve 1 and 3 the z range is utilized and therefore the minimum and maximum slopes are the same. However for Curve 1, A and C are plus one, while for Curve 3, A and C are minus one. This means that as the parameter z increases in absolute value from zero, the denominator of the ROC character

for Case 1 will increase, and the ROC character will decrease. In just the opposite manner as z increases in absolute value, the denominator for the ROC character for Curve 3 will decrease in magnitude and the ROC character will increase for large z. The ROC characters and the corresponding ROC curves are plotted on linear paper in Fig. 5.13. The cusp nature of the third ROC character concentrates more probability at extreme values of z, and therefore places more emphasis on those arcs of the ROC with extreme slope. This means that it spends more "time" rising sharply from the origin than does Curve 2 with a flat ROC character; the ROC curve is superior to the rectangular hyperbola with the same limiting slopes. In contrasting manner, the first ROC character has the more normal convex appearance with mode at z = 0. This emphasizes slope in the neighborhood of one (z = 0). If one were to use an approximation procedure for data by using rectangular hyperbola for the first order approximation, it seems fairly evident how one might adjust the ROC character compared to the flat character to obtain better fits in the neighborhood of the negative diagonal.

The three ROC characters of Table 5.4 and their ROC curves, can be compared to normal by plotting the curves on normal-normal paper. This is done in Fig. 5.14. These hyperbolas differ from normal ROC curves by indicating higher detection indices, d, along the negative diagonal and lower values off the diagonal. That is, they have sharper "elbows" along the negative diagonal. For
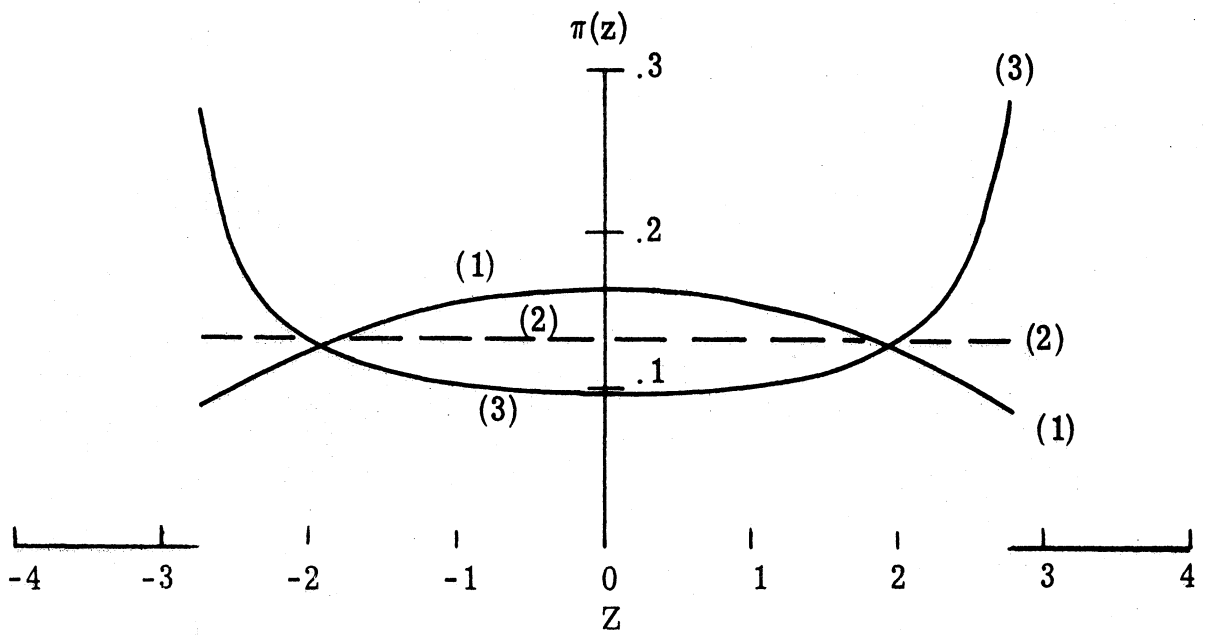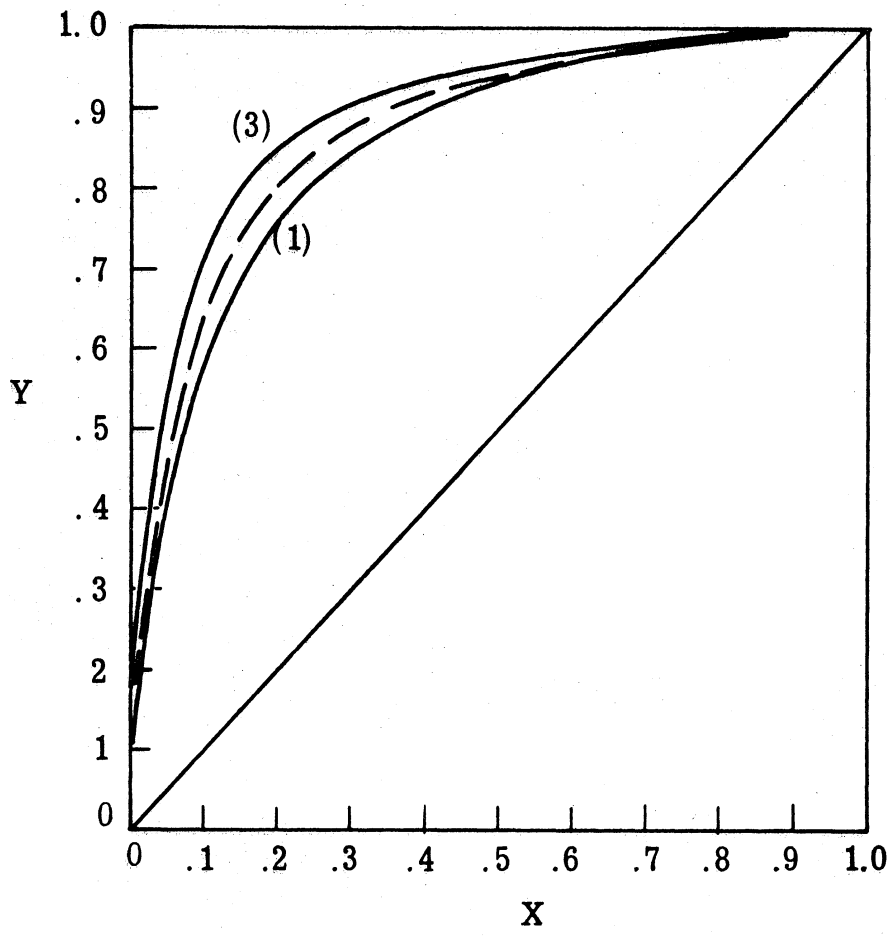
Fig. 5.13. ROC and character for three hyperbolas
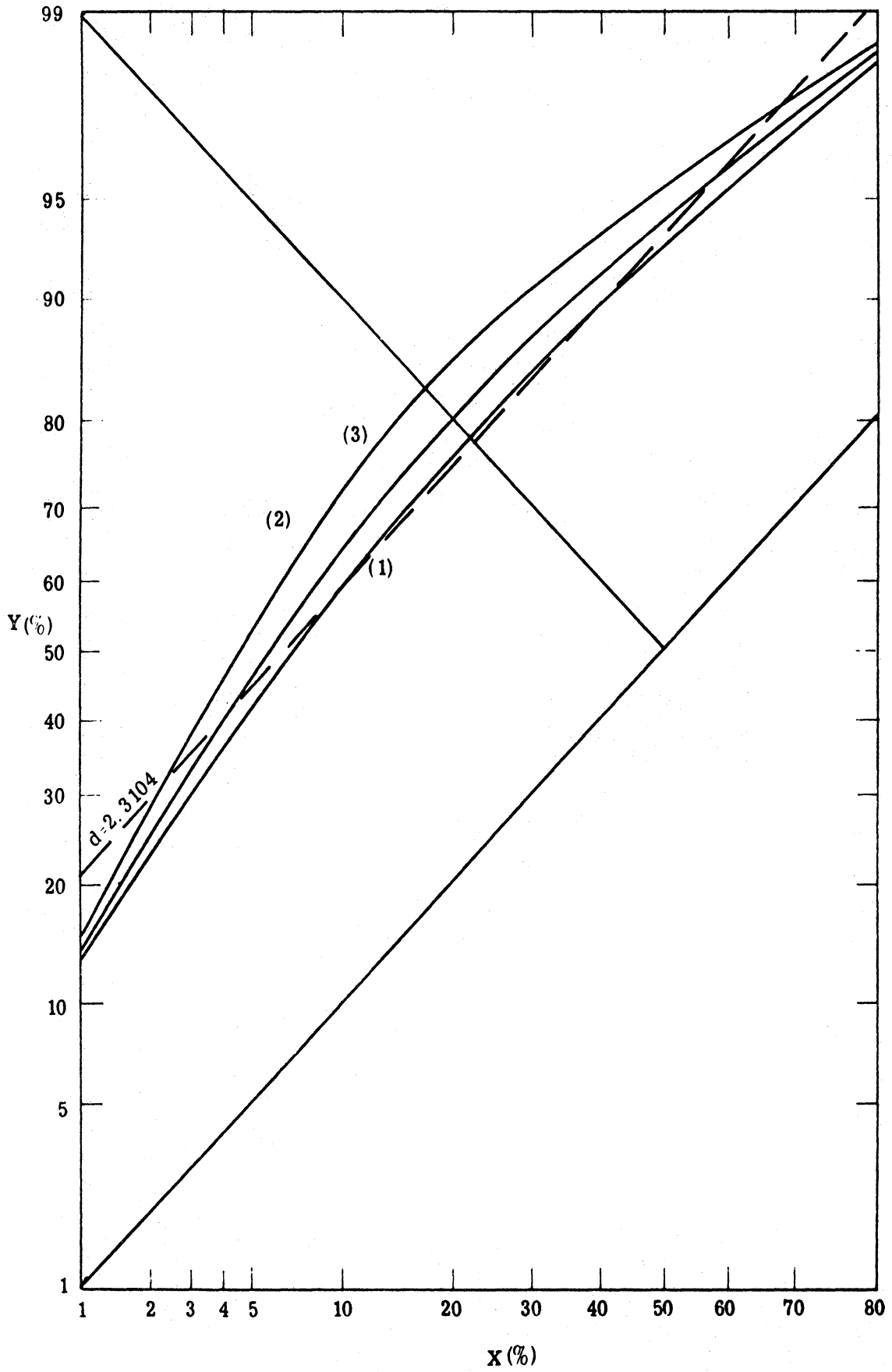
Fig. 5.14. Three hyperbolas compared to normal

comparison, a normal curve with d = 2.3104 has been chosen to approximate Curve 1 in the .05 to .50 false alarm probability region. The character and ROC curve for the first hyperbola have been repeated in Fig. 5.15 on linear paper, along with the approximating normal curve and character. The normal ROC character extends to plus infinity, therefore we would expect it to have far superior performance at very low false alarm rates. Indeed, for x = .01 the normal ROC has y = .21 while the hyperbolic ROC shown has only y = .132. This discrepancy might be exceedingly important in some situations. If the data for which an approximating ROC curve is desired are based on fewer than 500 observations, then determination of an accurate false alarm rate of the order of .01 would not be possible, and this region of the ROC would be of little interest. A feature of the chosen fit is that the one percent false alarm rates occur at roughly the same z value. As can be seen from Fig. 5.15, from the maximum z value of the conic ROC down to about z = 1 the ROC character for the conic is larger than the character of the normal; the conic ROC rises proportionately faster. By the time the false alarm rate is ten percent the two curves cross. If one's definition of "close" means "as points appear when plotted on linear paper," then these two ROC characters yield ROC curves which are quite close.

5.3 Lor-Lor Paper

The previous section developed the three parameter
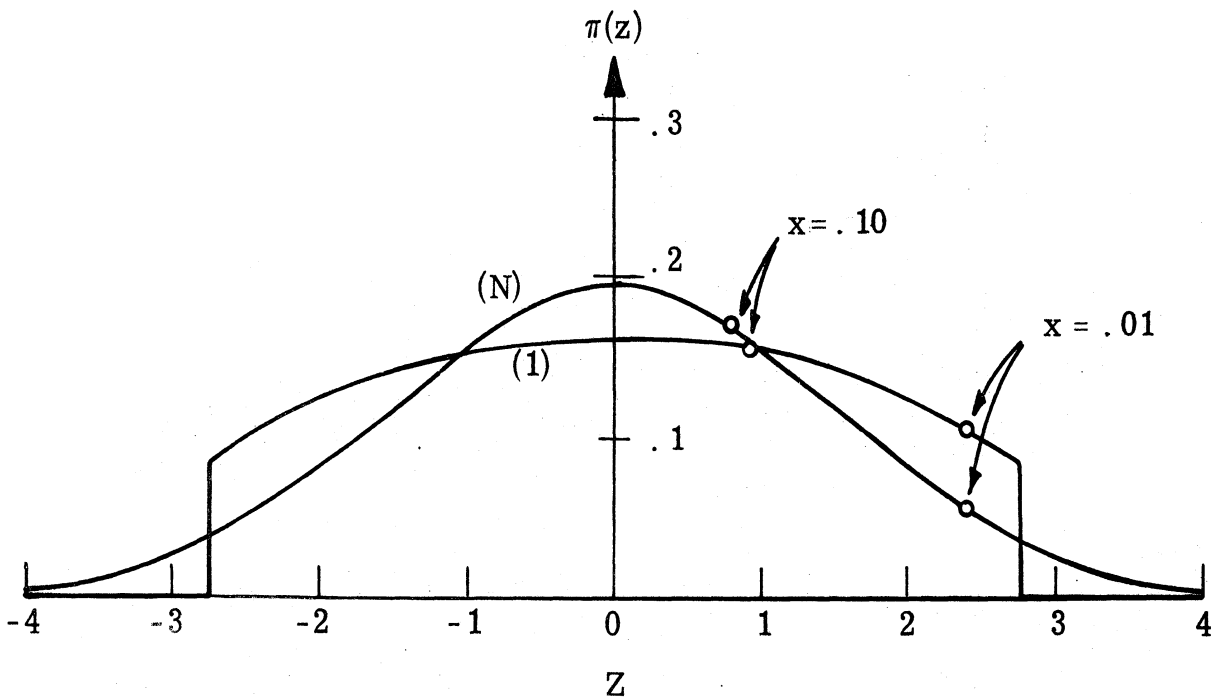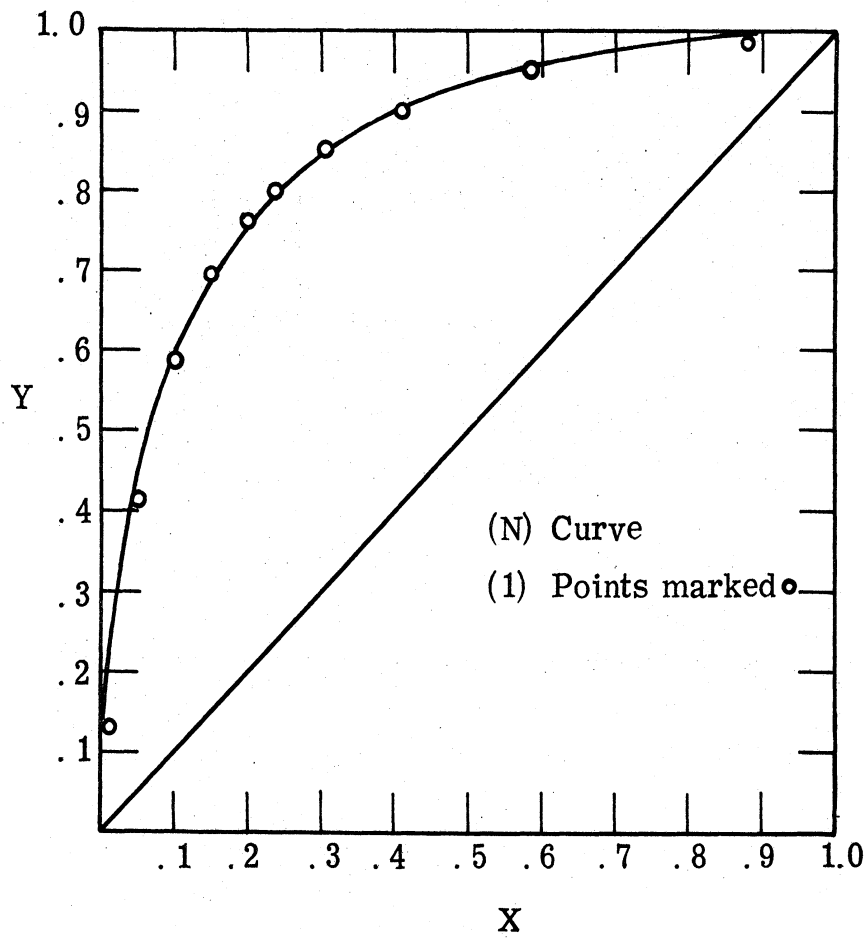
Fig. 5. 15 Comparison of curves and characters
for a conic and a normal ROC

class of regular conic ROC curves. If one wishes to use this form of ROC curve and character to fit data, one would appreciate some graphical aid to obtaining the fit. Those who have worked with normal ROC approximations are familiar with the usefulness of normal-normal paper. In some experimental work the data have considerable scatter, and the availability of various types of special paper, a straight-edge, and the experimenter's eye can be as powerful as a digital computer.

This section will develop the formulae for the coordinates of special paper, called lor-lor paper on which the rectangular hyperbolas plot as straight lines with slope one. This is followed with the general equation for the symmetric conic ROC curves as they would plot on this paper. Finally, graphs are presented of the hyperbolas, other conics, and normal ROC curves, on lor-lor paper. The coordinate transformation which shall be used was used by R. A. Fisher in his study of the correlation coefficient, and is sometimes known as the Fisher-z transformation. The coordinate transformation has also been used in the study of sequential observation (Ref. 25). I am unaware of the commercial availability of such paper.

The ratio of the probability of an event occurring to the probability of that same event not occurring is sometimes called the "odds ratio." Quite naturally, the logarithm of this ratio is referred to as the logarithm of the odds ratio, or simply log-odds-ratio. We shall reserve the symbol L for the log-odds-ratio transformation (simply, the lor transformation).

$L(p)$ vs. P
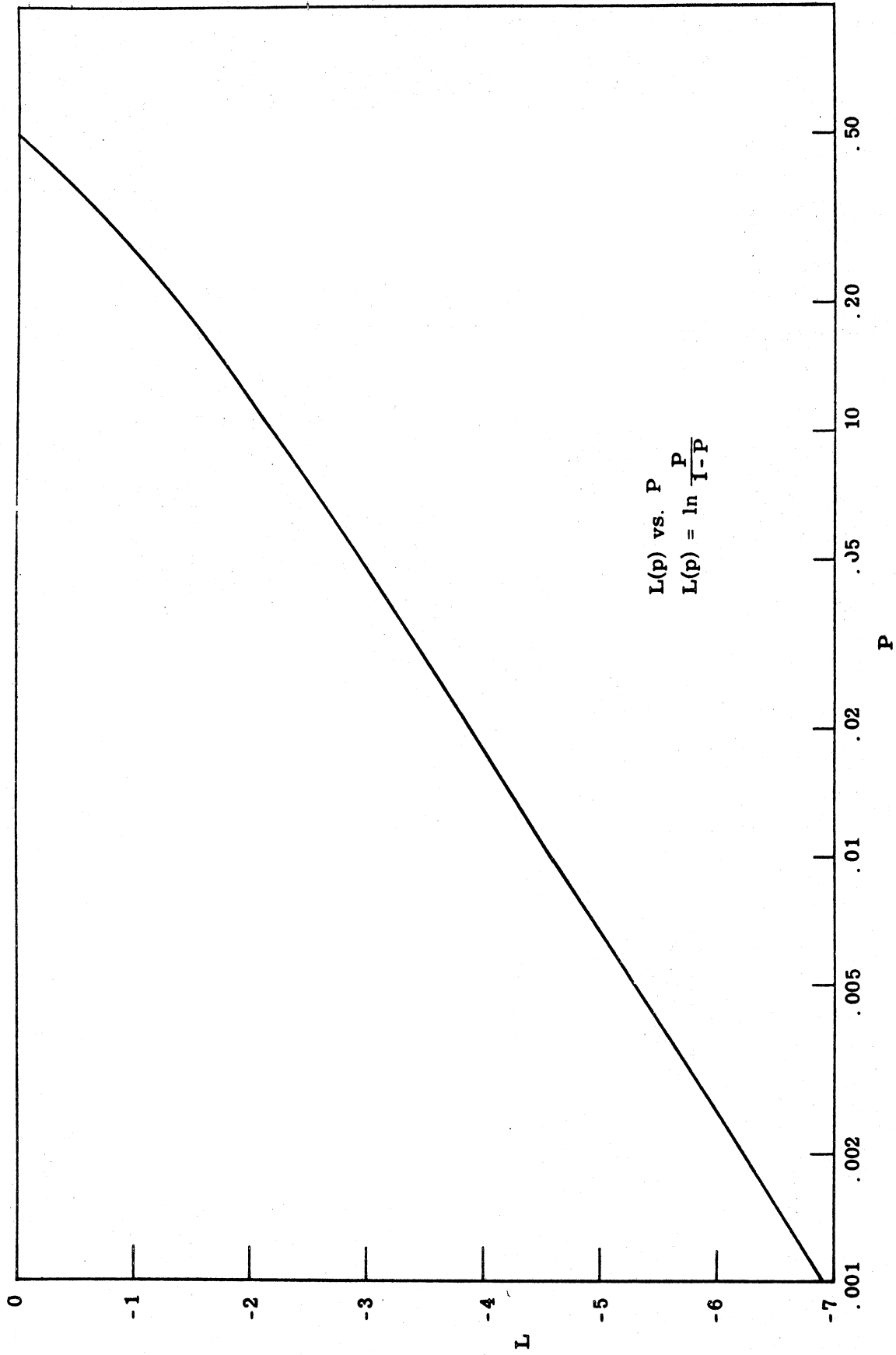
$L(p) = \ln \dfrac{P}{1-P}$

Fig. 5.16. Comparison of lor and log scales

<u>Definition</u> of log-odds-ratio, L( ):

$$0 < P < 1, L(P) = \ln \frac{P}{1-P} \qquad (5.66)$$

For very small values of its argument, the log-odds-ratio is very close to being a simple logarithm. To demonstrate how close, the log-odds-ratio has been plotted against probability on semi-log paper in Fig. 5.16. For values of probability below something like three percent to ten percent L(P) is graphically equivalent to $\ln P$. L is an odd symmetric function about 50 percent probability. This may be seen by substituting (1-P) for P in Eq. 5.66. Therefore when P becomes very close to unity, L(P) will again become nearly logarithmic, but with the logarithm of (1-P).

$$L(1-P) = -L(P)$$
$$P \to 0, \quad L(P) \to \ln P \qquad (5.67)$$
$$P \to 1, \quad L(P) \to -\ln(1-P)$$

This symmetry property makes the lor scale more useful than the ordinary log scale for plotting ROC curves. In Fig. 5.17 the log, lor, and normal scales are displayed. The placement was chosen to align the one percent and fifty percent points on the scales. The major contrast between the lor and log scales is in the symmetry. The log scale expands the region below .50, but greatly compresses the region between .50 and 1.00, while the lor scale expands both regions
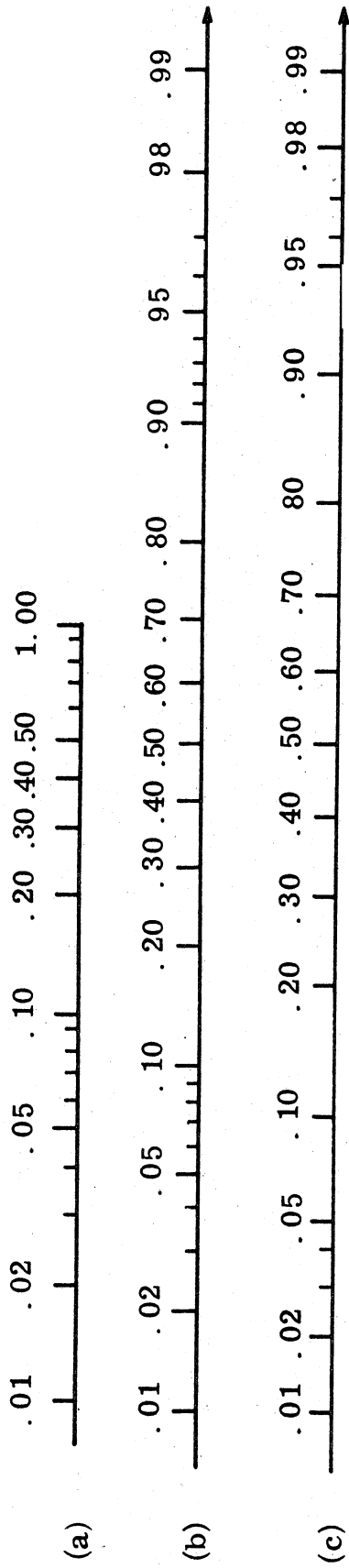
(a) Logarithmic Scale $s = \ln p$, $p(s) = e^s$

(b) Lor Scale $L(p) = \ln \dfrac{p}{1-p}$, $p(L) = e^L/(1+e^L)$

(c) Normal (Probability) Scale, $p(t) = \displaystyle\int_{-\infty}^{t} \dfrac{1}{\sqrt{2\pi}} e^{-u^2/2} \, du$

Fig. 5. 17. Comparison of lor, normal and log scales

symmetrically. The normal scale is also a symmetric scale about
.50 probability. On all three scales the marked positions for proba-
bility are not uniformly spaced. The scale variables, L, t, and s,
are uniformly spaced on the axes.

The quantity d' advanced by Tanner, Swets, and Green
as an index of detectability is directly related to the normal scale when
the experimental procedure is a YES-NO test (Ref. 13, 10). For each
point (x, y) on the ROC curve the detection index is given by Eq. 5.68.

$$d' = \Phi^{-1}(y) - \Phi^{-1}(x) \tag{5.68}$$

The function Φ is the normal distribution function and the symbol
Φ^{-1} indicates the inverse function. The inverse of the normal proba-
bility distribution function is the linear distance scale on normal-
normal paper. Therefore the value of d' for any ROC point is its
vertical scale value minus the horizontal scale value on normal-
normal graph paper. The normal ROC curves plot as straight lines
with a slope of one on normal-normal paper. The difference between
the vertical and horizontal scale values is constant over the normal
ROC. The numerical value of the constant is the square root of the
index d appearing in the ROC character. Therefore, for the normal
ROC the value of d' is the square root of d over the whole ROC curve.

$$\text{Normal ROC: } d' = \sqrt{d} \tag{5.69}$$

In an analogous manner we shall define a lor index of detectability

of a point as the difference between the vertical scale coordinate and

the horizontal scale coordinate on lor-lor graph paper.

Definition: $M' = L(y) - L(x)$ (5. 70)

      The coefficients of the equation for the regular rectan-

gular hyperbola are repeated in Eq. 5. 71.

Regular Rt. Hyperbola: $A = 0,\ B = M-1,\ C = 0,\ D = 1,\ E = -M$ (5. 71)

M is the maximum slope of the ROC curve, and the reciprocal of the

minimum slope of the ROC curve. Substitute these values in the gen-

eral equation, and divide through by a common factor of two.

$$(M-1)\,xy + y - Mx = 0$$ (5. 72)

Such an equation can be explicitly solved for y. The solution is

$$y = \frac{Mx}{Mx + 1 - x}$$ (5. 73)

Subtract y from one

$$1 - y = \frac{1 - x}{Mx + 1 - x}$$ (5. 74)

and take the ratio of y to (1-y).

$$\frac{y}{1-y} = M\,\frac{x}{1-x}$$ (5. 75)

In logarithmic form,

$$\ln \frac{y}{1-y} = \ln \frac{x}{1-x} + \ln M \qquad (5.76)$$

Equation 5.76 can be simply rewritten in terms of the lor scale.

$$L(y) = L(x) + \ln M \qquad (5.77)$$

If graph paper is used so that the true linear scale along the two axes is the lor of the probability marked on those axes, then a regular rectangular hyperbola will plot as a straight line with slope of one. Along the entire ROC curve the lor index of detectability, M', will be constant and equal to the logarithm of M.

$$\text{Rt. Hyp. ROC:} \quad M' = \ln M \qquad (5.78)$$

In Fig. 5.18 curves of constant M' are plotted on lor-lor paper.

A subtle distinction which is very necessary is the difference between M and M'. The same difference must be made between d and d' (Ref. 10). The index d applies only to the normal ROC curve, and the normal ROC character. The index M applies only to the regular rectangular hyperbola ROC curve, the rectangular ROC character. The indices d' and M' apply to any point on the ROC graph. In general, the indices d' and M' will vary along a given ROC curve. Only when the ROC curves are the special cases of normal (for d') and rectangular hyperbola (for M') will these indices
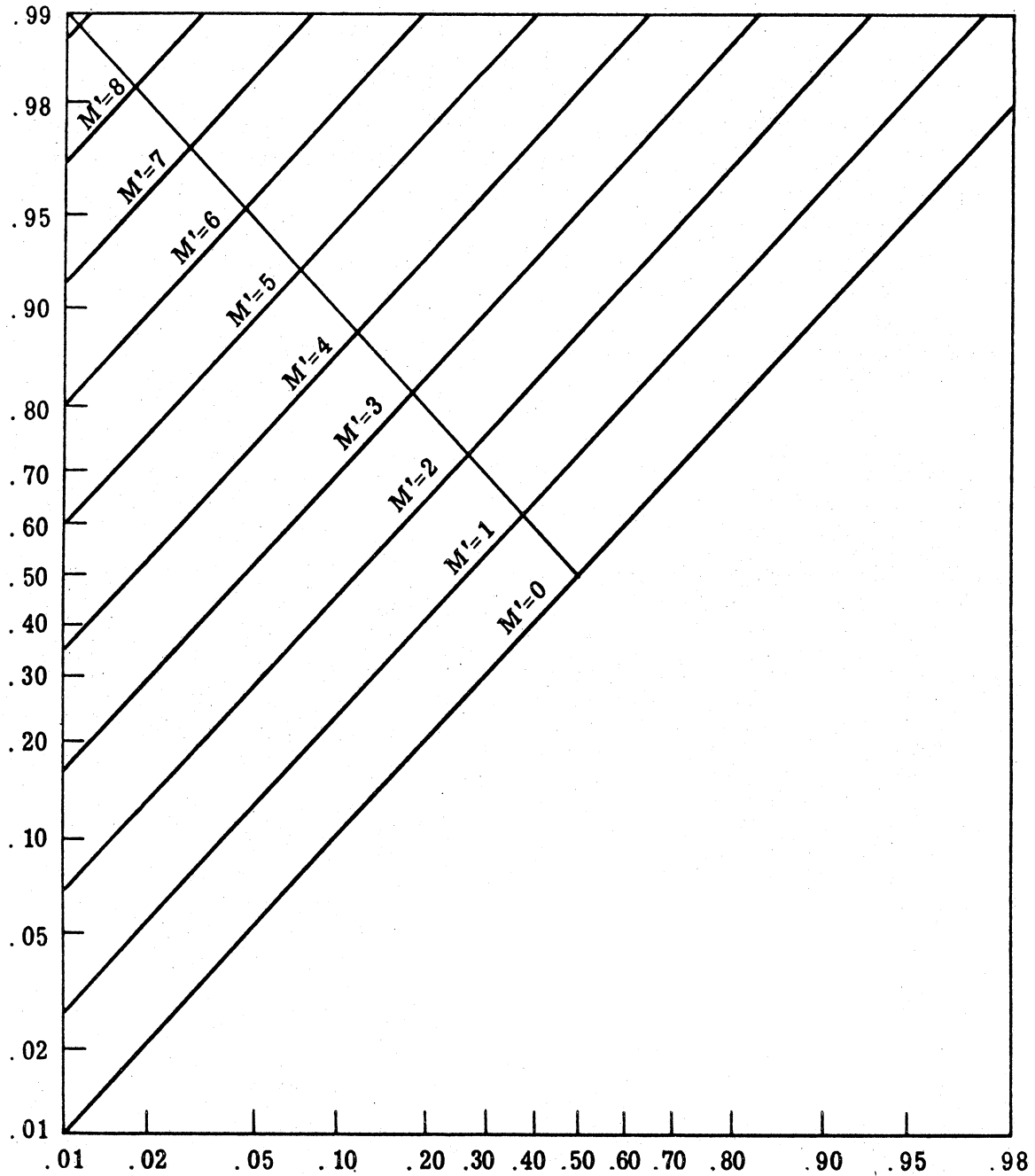
Fig. 5.18. Lor-Lor paper with lines of equal M'

be constant over the entire ROC curve, and directly related to the parameter for the entire curve.

In Fig. 5.19, a family of normal ROC curves are plotted on lor-lor paper. The index M' varies along each curve, although not radically. Along the negative diagonal, the value for M' is a minimum, and M' increases along a given curve as the point moves away from the negative diagonal. This indicates that the normal character concentrates more probability at large magnitude z values than would a rectangular character, which yielded the same negative diagonal point.

When working with ROC curves with nonzero minimum and noninfinite maximum slope, then one may obtain information about the ROC character from the ROC curve plotted on lor-lor paper. In order to demonstrate this interpretation procedure, a set of conic ROC curves with the same initial and slope values were selected. The parameters for these are given in Table 5.5.

Neg. Diag. Point	A	B	C	D	E	Δ
(.30, .70)	$5\frac{1}{8}$	$3\frac{7}{8}$	$5\frac{1}{8}$	1	-10	-11.25
(.2405, .7595)	0	9	0	1	-10	81.
(.20, .80)	$-1\frac{1}{3}$	$10\frac{1}{3}$	$-1\frac{1}{3}$	1	-10	105
(.10, .90)	-2.21875	11.21875	-2.21875	1	-10	120.94

Table 5.5. Table of parameters
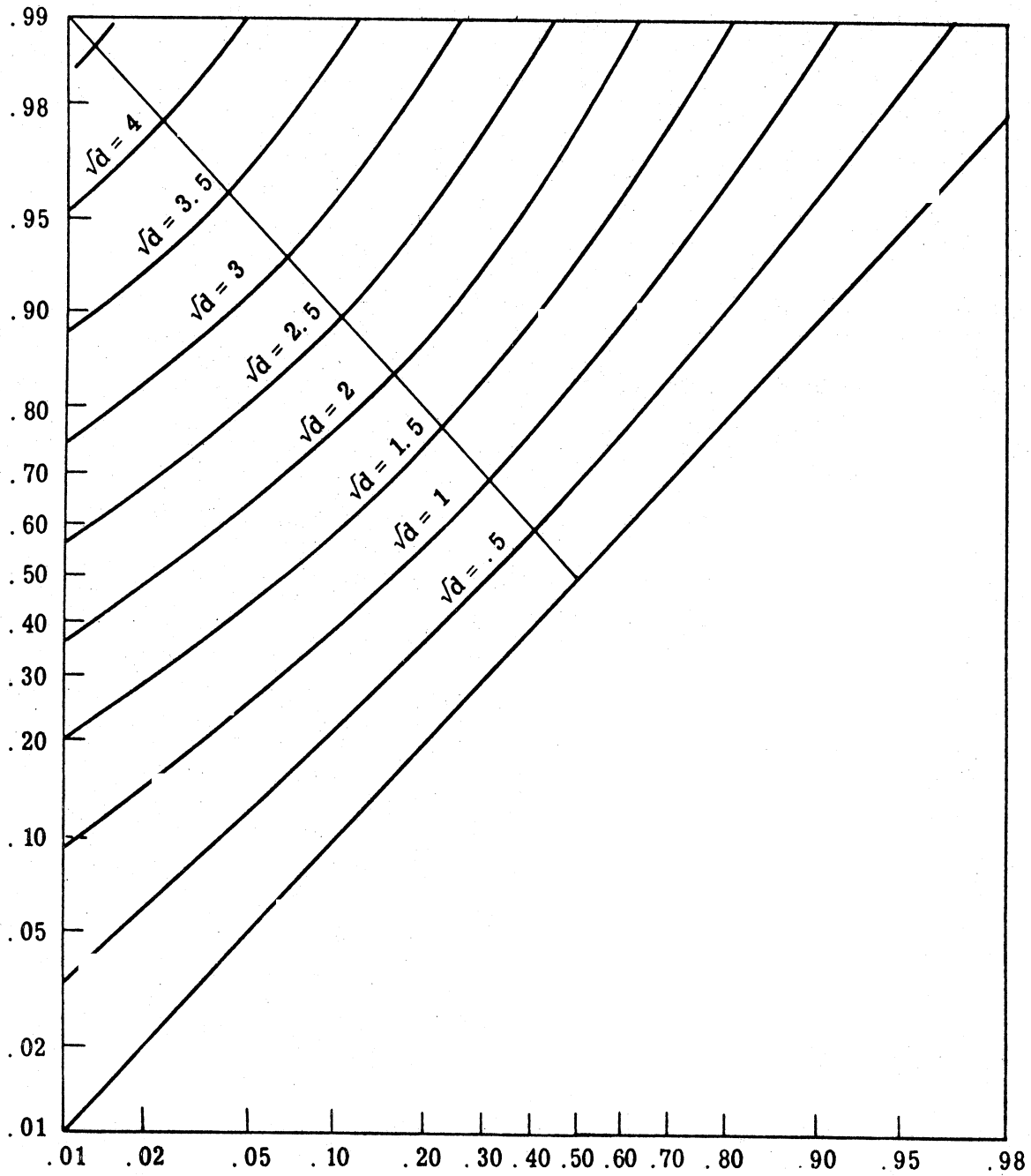
Fig. 5.19. Normal ROC curves on lor-lor paper

Each of these curves has maximum slope of ten and minimum slope of $.1$. A rectangular hyperbola with index $M' = \log 10 = 2.30$ also has these same terminal slope values. These ROC curves are shown on linear paper in Fig. 5.20. The Luce ROC curve with slopes 10 and $.1$ is also drawn. This is the limiting value for conic ROC curves with these slopes. The hyperbola going through the negative diagonal point $(.10, .90)$ lies very close to the Luce ROC curve. The hyperbola going through the negative diagonal point $(.20, .80)$ is superior to the rectangular hyperbola, and the elipse going through the negative diagonal point $(.30, .70)$ is inferior to the rectangular hyperbola ROC curve.

Figure 5.21 displays the ROC characters for the same curves on semi-log paper. Figure 5.22 displays the ROC curves on lor-lor paper. The rectangular hyperbola has flat character, and plots as a straight line on Fig. 5.22. Since all of the curves have the same minimum and maximum slope values, the z range is common. This is shown as a common range in Fig. 5.21 from -2.3 to $+2.3$. This appears on the ROC curve in Fig. 5.22 as the common asymptotic M' value. All of the curves are asymptotic to $M' = 2.30$. The curve going through the negative diagonal point $(.30, .70)$ was inferior to the rectangular hyperbola, while the other two were superior. The same relationship holds for large magnitude z values on the ROC character. The $(.30, .70)$ curve has lower character value than the flat character while the other two have higher character value at
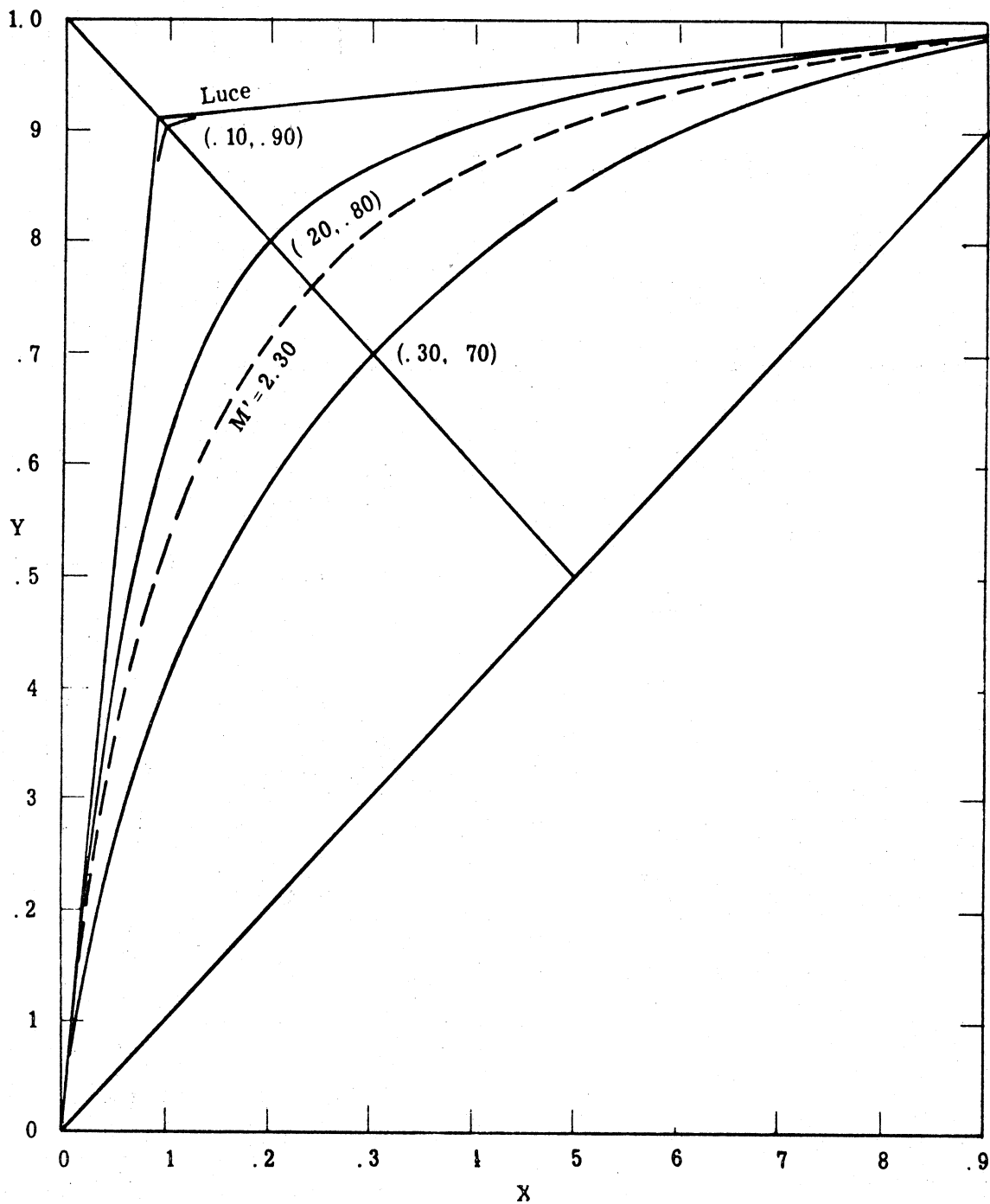
Fig. 5.20. Five ROC curves with common terminal slope

(112.4)   (112.4)

(.30,.70)

M' = 2.30

(.20,.80)

(.10,.90)

Neg Diag Pt Given

● Indicates X = .01 for z > 0
             Y = .99 for z < 0
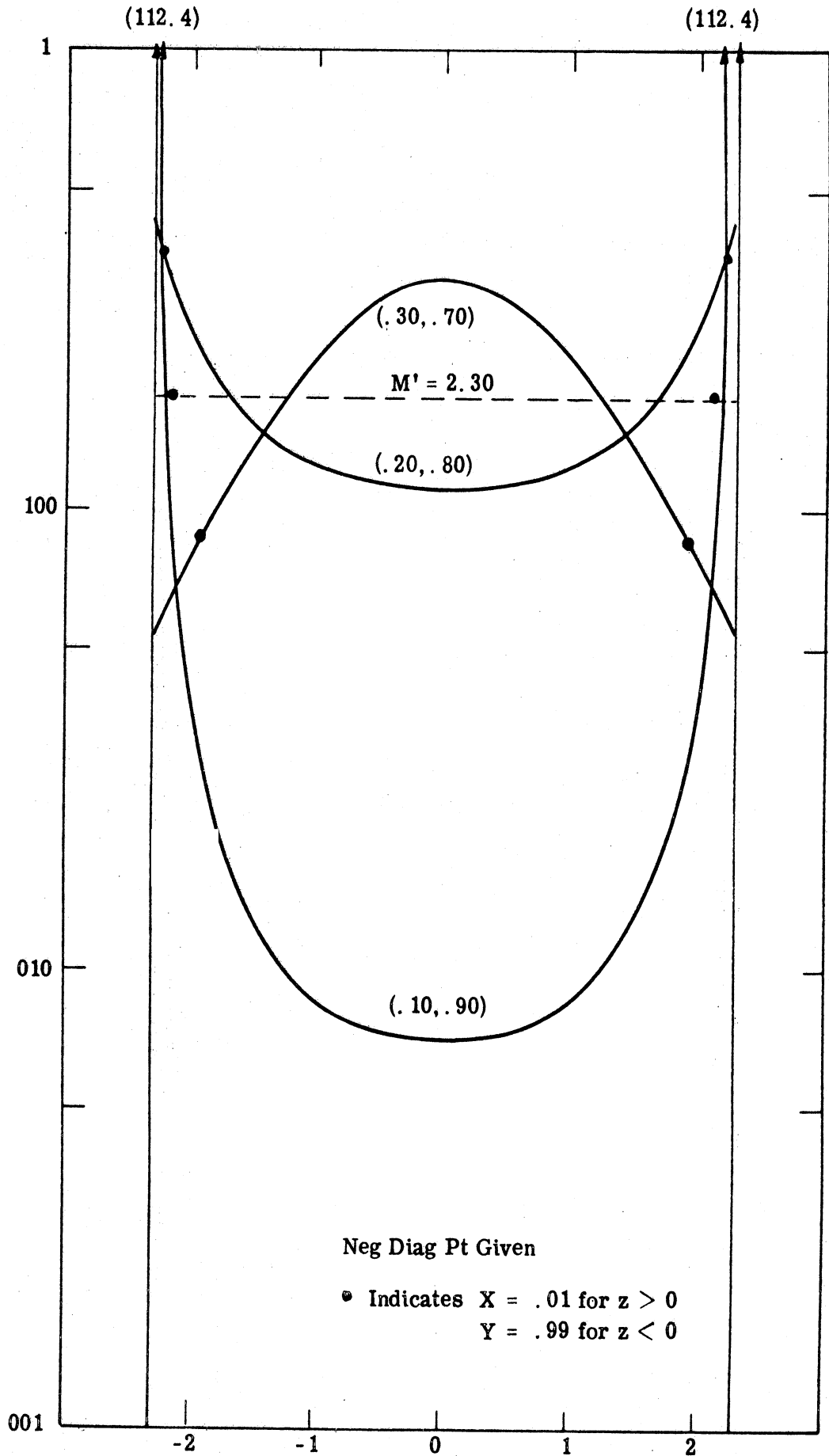
Fig. 5.21. ROC characters -2.3 ≤ z ≤ +2.3
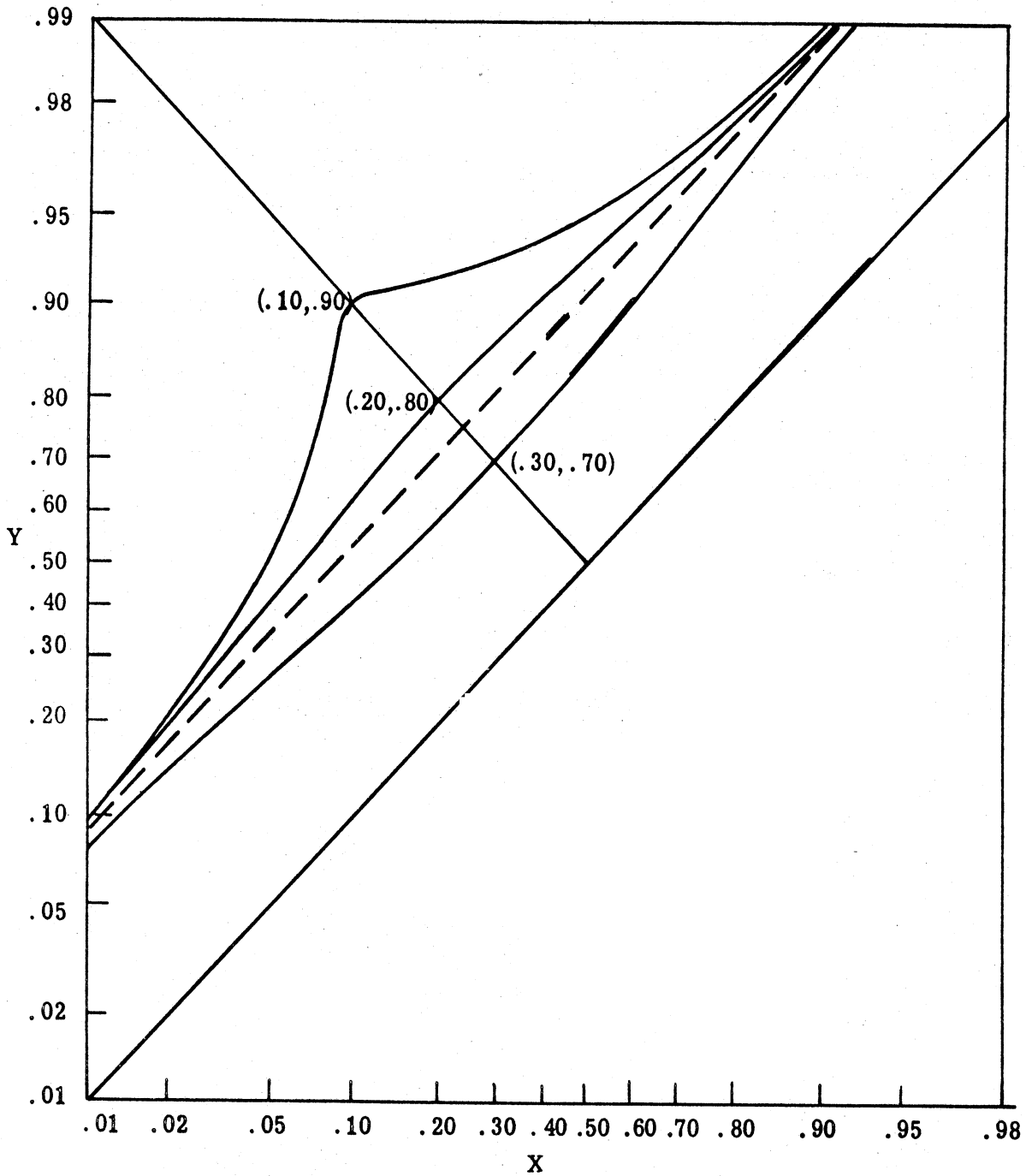
Fig. 5.22. Four ROC curves with common
terminal slope

large magnitude z. Of course, just the opposite relation holds when

z is close to zero. Since the ROC curve is obtained by multiplying

the ROC character by $e^{\pm .5z}$, and integrating, the effect of the ROC

character for large magnitudes of z overshadows the effect at small

magnitudes of z. The ROC character for the (.10, .90) case rises

from very small values to the value .16 at z = ±2.2, to the value

112.4 at z = ±2.3. If the peak character value of 112 were main-

tained constant over a z range of only .0026, it would accumulate

an area of .288. The integrated ROC character for the Luce ROC has

jumps of .288 at z = ±M'; that is, the character has delta functions

of magnitude .288 at ±M'. The two ROC curves and characters are

close.

The above example compared symmetric conic ROC

with the same limiting slope values. The next example compares non-

symmetric conic ROC's with the same limiting slope values. The lim-

iting slopes chosen are a maximum slope of 100 and a minimum slope

of .1. This means that the z range will be from -2.3 to +4.6.

The coefficients for the five cases considered are given in Table 5.6.

The ROC curves are displayed on linear coordinate in Fig. 5.23.

Since the initial and final slope values are not reciprocal values, there

is no rectangular hyperbola with constant M' value to compare with

these. In Fig. 5.23 are shown portions of two constant M' curves.

In the high slope region, to the left of the negative diagonal, is drawn

the rectangular hyperbola with index M' = 4.6. To the right of the
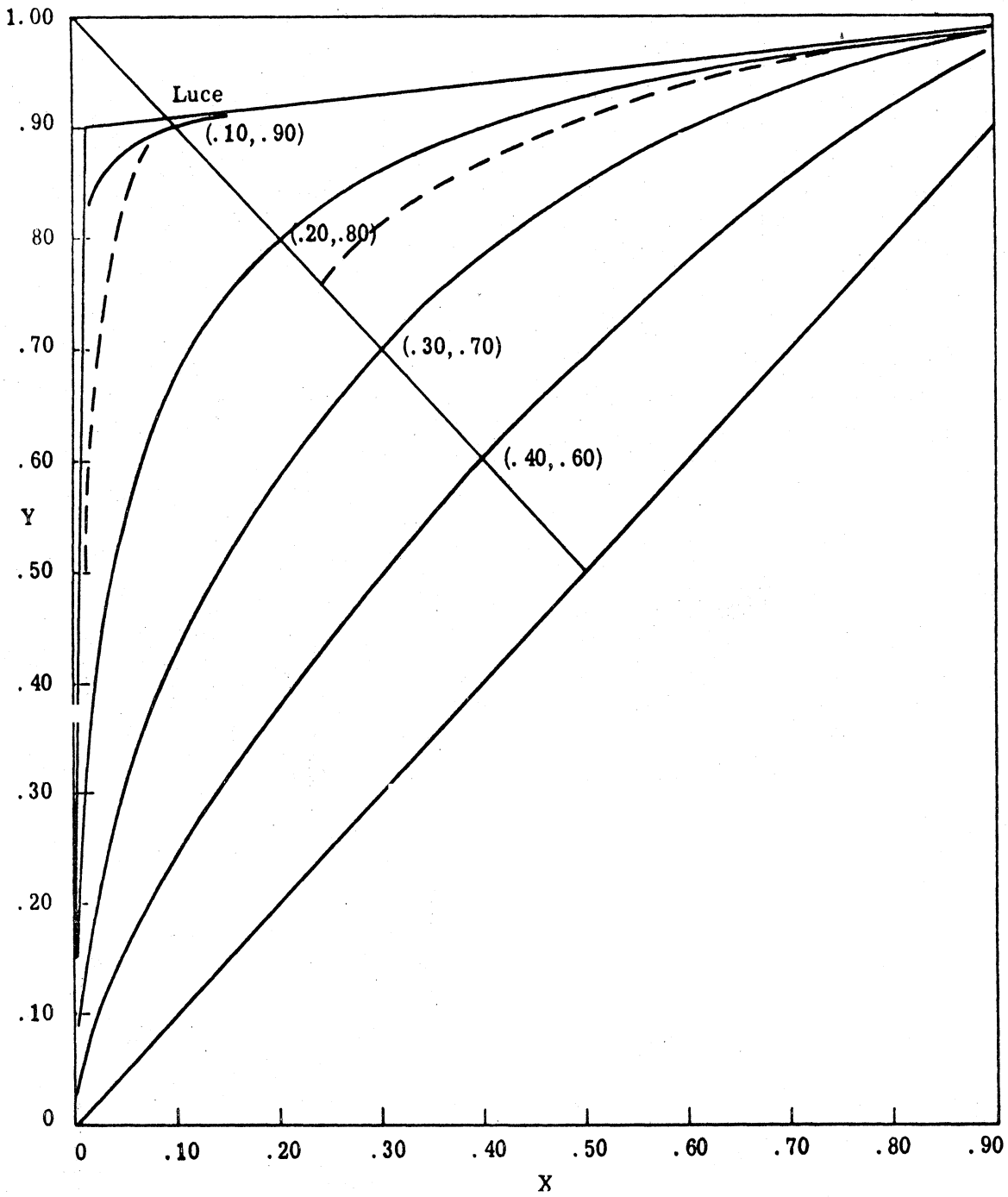
Fig. 5.23. Five asymmetric ROC curves with common terminal slope

(.4, .6)	A = 742	B = -633	C = 722	D = 1 E = -100	Δ = -135,035
(.3, .7)	A = $91\frac{3}{8}$	B = $17\frac{5}{8}$	C = $71\frac{3}{8}$	D = 1 E = -100	Δ = -6211. +
(.2, .8)	A = 12	B = 97	C = 8	D = 1 E = -100	Δ = +9505
(.1, .9)	A = $-1\frac{29}{32}$	B = $110\frac{29}{32}$	C = $-21\frac{29}{32}$	D = 1 E = -100	Δ = +12,258[+]
$(\frac{1}{11}, \frac{10}{11})$	$y - 100x = 0$, $x \leq \frac{1}{111}$;	$10y - x = 9$	$x \geq \frac{1}{111}$		Luce ROC

Table 5.6. Table of coefficients

negative diagonal, in the low slope region, is drawn the rectangular

hyperbola with $M' = 2.3$. In the neighborhood of the corners points

of the ROC all of the curves will come together because they approach

the corners with the terminal slope values.

The corresponding ROC characters are displayed on

semi-log paper in Fig. 5.24. The ROC curves on lor-lor paper are

displayed on Fig. 5.25. The two curves which are inferior to both of

the constant M' curves display the same general peak character as in

the previous example. The curve marked (.1, .9) hugs the Luce ROC

for the low slope values, but not for high slope values. Therefore the

ROC character displays a nearly delta-like behavior at the minimum

z value, -2.3, but not nearly so sharp a cusp at the maximum z

value, 4.6. The curve marked (.2, .8) has a monotone decreasing

ROC character. The value of this character falls between that of the

two limiting constant M' values between $z = -1.4$ and $z = 2.6$. This

(13, 6)

1.000

(.4, .6)

(.3, .7)

π(z)

(.2, .8)

100

(.1, .9)

010

Negative Diag Pt Given

∘ Indicates X = .01 for z > 0
           Y = .99 for z < 0

.001

-2.3 -2   -1   0   1   2   3   4   4.6

z
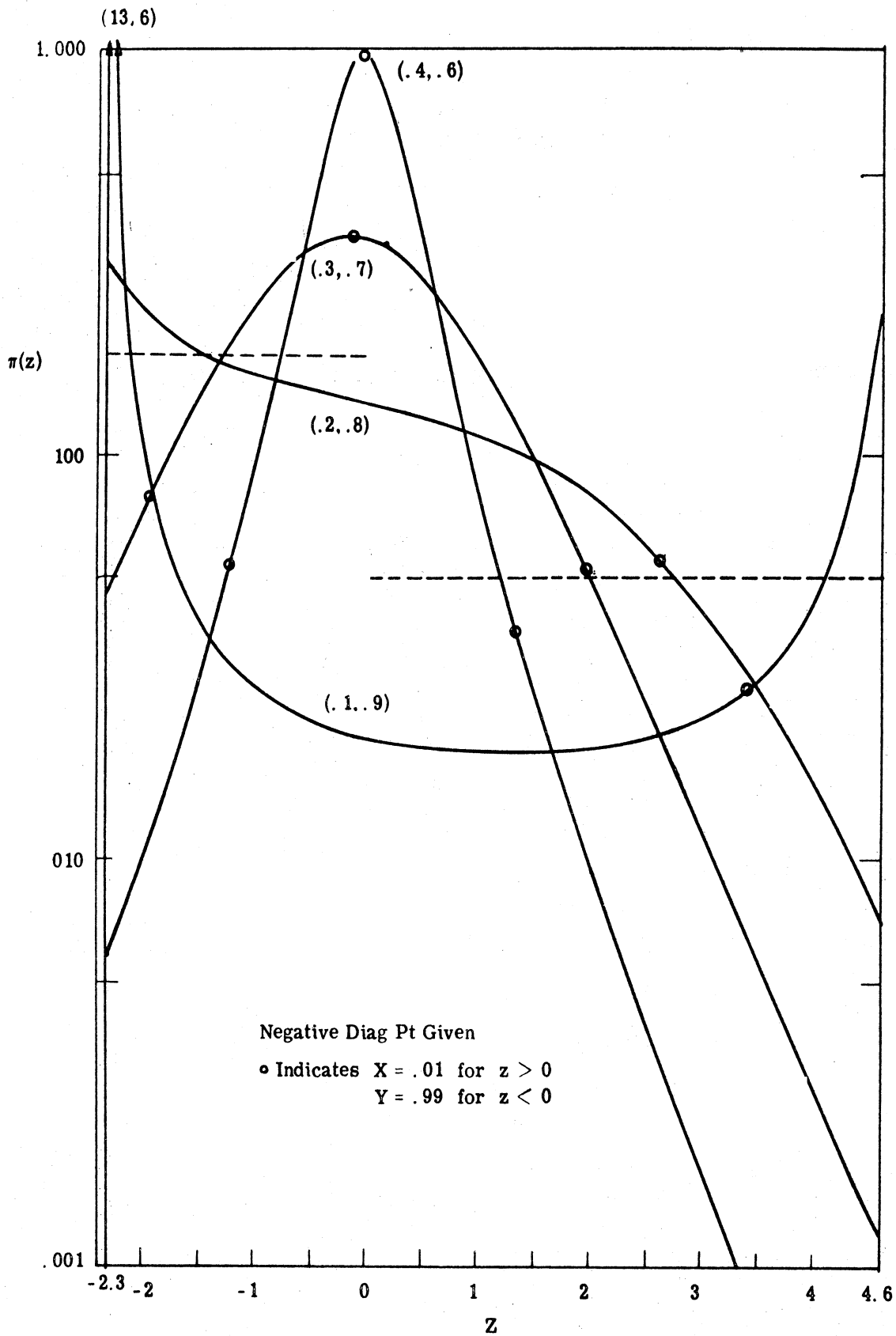
Fig. 5.24. ROC characters -2.3 ≤ z ≤ +4.6
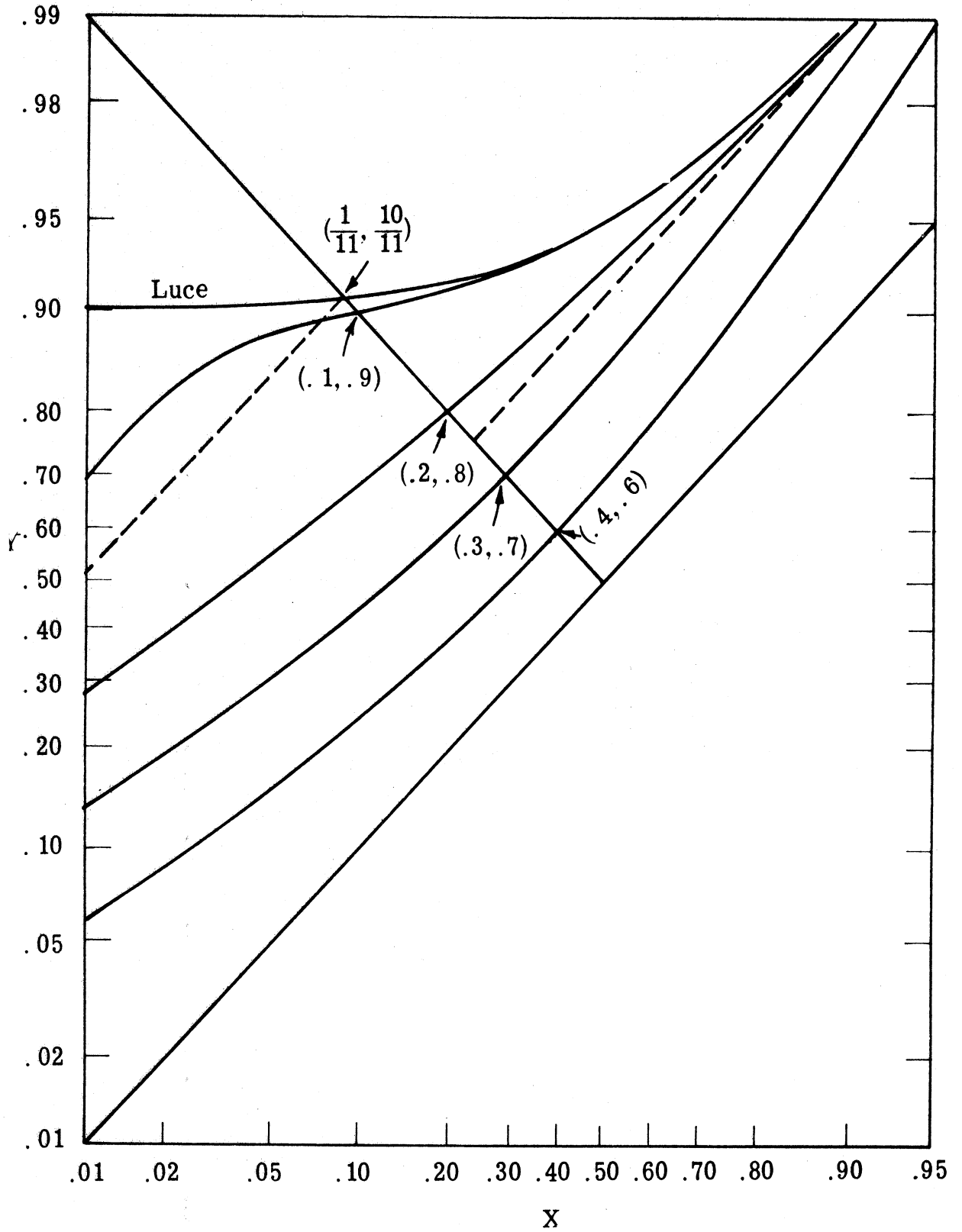
Fig. 5.25. Five asymmetric ROC with common terminal slope on lor-lor paper

corresponds to the range of slopes from .25 to 13.5. Since the ROC character for (.2, .8) falls below the limiting M' character at z = 4.6, the ROC curve will be inferior to the M' = 4.6 level near the origin. Since the ROC character for (.2, .8) is larger than that for M' = 2.3, the ROC curve will fall above the M' = 2.3 value as the curve approaches (1, 1).

In Fig. 5.26, a pure power law ROC is plotted on lor-lor paper. Two asymptotes are drawn to this curve. The left-hand asymptote is obtained from the logarithmic nature of the lor for small values of argument. Since the pure power ROC plots as a straight line on log-log paper, with slope A (in this case, 2) the pure power ROC will plot as a straight line with slope A on lor-lor paper when both x and y are small. Since the pure power ROC curves approach the (1, 1) point with slope 1/A, the ROC curve will become asymptotic to the rectangular hyperbola with index M' = ln A.

5.4 Choosing a Fitting Conic

The conic equation contains six coefficients, A, B, C, D, E, and F. However, restriction to regular ROC curves places three side conditions on these coefficients, leaving three degrees of freedom. It would be desirable to have a procedure to obtain the "best" fitting ROC to any amount of data. Such a procedure has not been developed in this present work.

If three representative points are selected, the procedure
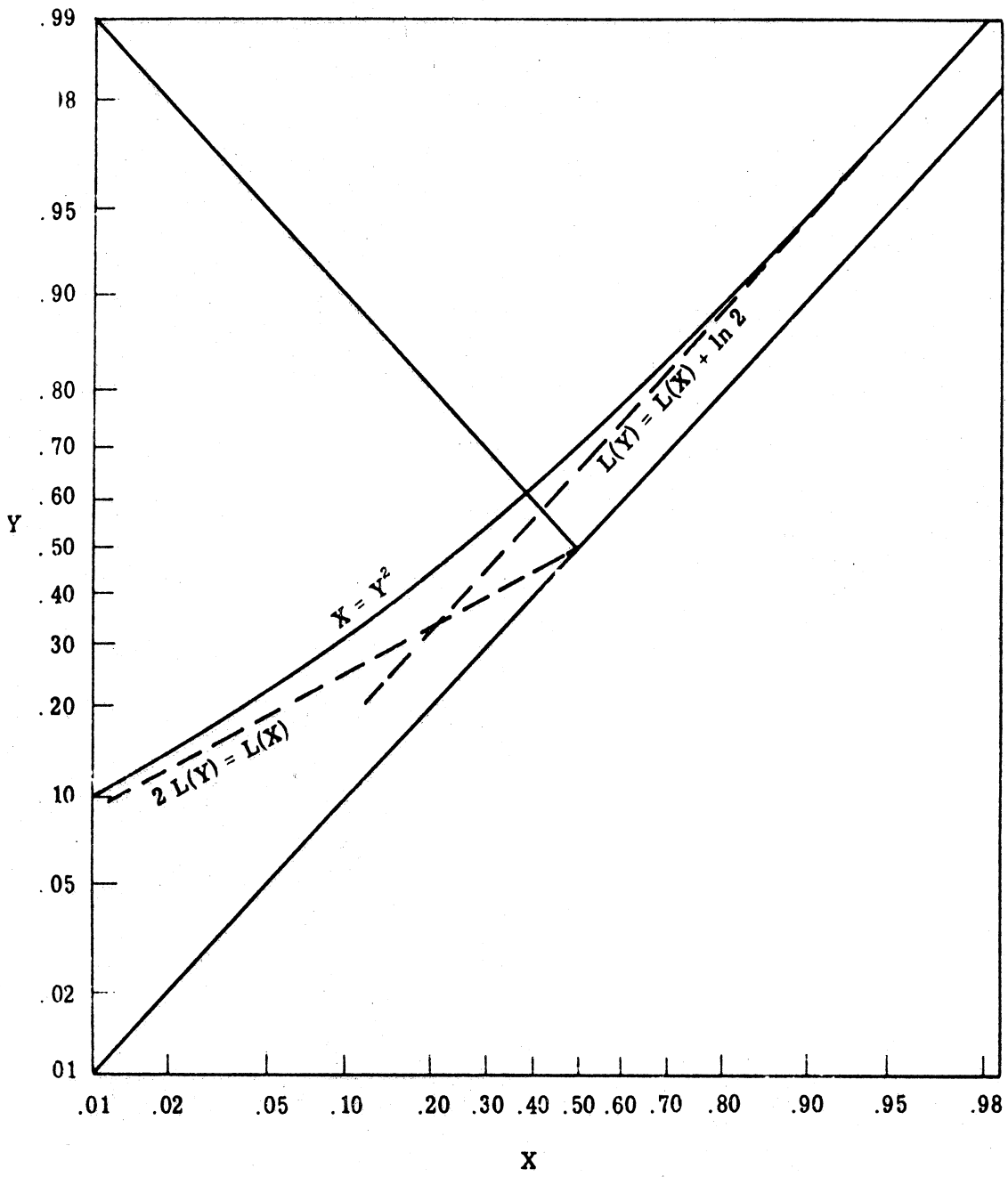
Fig. 5.26. Power ROC on lor-lor paper

to fit the regular conic through these three points is fairly simple and straightforward. One must first check that the broken line segment that goes from the point $(0, 0)$ to the point with lowest false alarm rate, then to the point with next false alarm rate, to the point with the highest false alarm rate, and then to the upper corner $(1, 1)$, is convex and that no three points are collinear. This last means that there are really four distinct straight line segments going from $(0, 0)$ to $(1, 1)$. The only problem is to determine the coefficients so that Eq. 5.79 is satisfied for these three points.

$$Ay_i^2 + 2Bx_iy_i + Cx_i^2 + 2Dy_i + 2Ex_i = 0, \qquad i = 1, 2, 3 \qquad (5.79)$$

Utilizing the equation that makes the ROC go through $(1, 1)$, and momentarily making the assumption that E is nonzero, the set of equations (5.79) becomes Eq. 5.80.

$$\left. \begin{array}{l} Ay_i(1-y_i) + By_i(1-x_i) + C(y_i-x_i^2) = (y_i-x_i) \quad i = 1, 2, 3 \\ \\ E = -1 \qquad D = 1 - B - .5(A+C) \end{array} \right\} \qquad (5.80)$$

If the determinant of the three equations in three variables is not zero, then the assumption that E was nonzero is valid and the ROC has been determined. If the determinant is zero, we may assume that $E = 0$ and then resolve the equations in the first part of Eq. 5.80 after setting the right-hand side to zero.

At a meeting of the Acoustical Society of America in

1965, students of Lloyd Jeffries presented some ROC curves with very distinct and measurable slopes from the origin and the (1, 1) point. If one wishes to match a conic ROC to such data, these two slopes and any one point are sufficient to define the conic. The only restriction is that the point must be chosen to lie above the chance diagonal and interior to the Luce ROC curve with the given slopes. The Luce ROC curve with these given slopes is given by

$$0 < m < 1 < M, \quad y = Mx \quad 0 \le x \le \frac{1-m}{M+m}$$

$$\text{(5.81)}$$

$$y = 1 - m + mx \quad \frac{1-m}{M+m} \le x \le 1$$

program was written to do this matching and is contained in Appendix E. The input data for such a program are the two slopes and the additional point. The program is designed to check that the point lies above the chance diagonal and below the Luce ROC curve, to determine the coefficients of the equations, and in addition to print out the ROC curve in linear and lor coordinates and determine the ROC character for the resultant conic.

CHAPTER VI

CHARACTER FITTING BY TRUNCATED NORMAL CHARACTER

6. 1 Metastatic Normal

The purpose of this chapter is to present various methods

of approximating ROC characters by sections of normal character, in

order to yield approximate ROC curves through the use of normal

tables. Each technique assumes a certain limited amount of knowledge

about the true ROC character, and based on this information achieves

some sort of fit in terms of a truncated normal character.

In Section 3. 4, the metastatic relation for general ROC

curves was given. The two key equations are Eq. 3. 39 relating the

ROC characters, and Eq. 3. 37, relating the ROC curves.

$$\pi_2(z) = \frac{\pi_1\left(z + \ln \dfrac{Y_1(b) - Y_1(a)}{b - a}\right)}{((b - a)(Y_1(b) - Y_1(a))^{.5}} \tag{3.39}$$

$$X_2 = \frac{X_1 - a}{b - a} \qquad\qquad Y_2 = \frac{Y_1 - Y_1(a)}{Y_1(b) - Y_1(a)} \tag{3.37}$$

The original ROC curve is subscripted by a one and the metastatic

derivative subscripted by a two. Two expressions that repeat in

these equations are the range of the X axis that is mapped onto the

metastatic derivative, and the portion of the Y axis that is mapped

onto the metastatic derivative. It will be convenient to have single

letters for these positive quantities. Let

$$K_N = b - a \qquad K_{SN} = Y_1(b) - Y_1(a) \qquad (6.1)$$

Two of the salient characteristics of an ROC character are the minimum and maximum values of z. Let the maximum value be z_a and the minimum value be z_b. The subscript was chosen to correspond to the X_1 values on the original ROC curve.

$$z_a \text{ is value of } z_2 \text{ when } X_1 = a, \ Y_1 = Y_1(a)$$

$$z_b \text{ is value of } z_2 \text{ when } X_1 = b, \ Y_1 = Y_1(b)$$

$$(6.2)$$

It will also be convenient to have a notation for the shift between the z values on the first (the original) and the second (the metastatic) decision axes. This is the logarithm of the ratio of K_N and K_{SN}, herein denoted by z_0.

$$z_0 = \ln (K_N / K_{SN}) \qquad (6.4)$$

$$z_1 = z_2 - z_0 \qquad (6.5)$$

This chapter is devoted to the metastatic normal ROC curve. Therefore $\pi_1(z)$ will always be assumed to be a normal ROC character. The general form of the ROC character with index d is

$$\pi_N(z) = e^{-\frac{d}{8}} \frac{1}{\sqrt{2\pi d}} \ e^{-\frac{z^2}{2d}} \qquad (6.6)$$

In order to avoid confusion between the normal detection index d, and

the parameter in an original ROC character which carries over as a parameter of the metastic character, write m^2 for d. The notation that ϕ is the normal probability density function in standardized form and Φ is the normal distribution function is as used throughout this work. Rewriting the normal ROC character with these changes,

$$\pi_N(z_1) = \frac{1}{m} e^{-\frac{m^2}{8}} \phi\left(\frac{z_1}{m}\right) \tag{6.7}$$

The ROC character for a metastatic normal ROC curve is

$$\pi_{MN}(z) = \frac{1}{m\sqrt{K_N K_{SN}}} e^{-\frac{m^2}{8}} \phi\left(\frac{z - z_0}{m}\right) \qquad z_b < z < z_a \tag{6.8}$$

In Fig. 6.1 are sketched two complete normal density functions on the original z_1 axis, and underneath this the z_2 axis. The
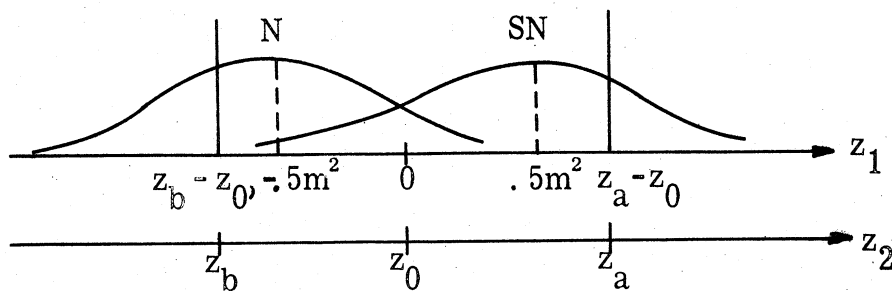
Fig. 6.1. Sketch of two z axes

three key z_2 values are the minimum value, z_b, the mode value, z_0, and the maximum value, z_a. Since the mean of the N distribution is at $-.5m^2$ the values for the false alarm probabilities, X_1, can be written in terms of the normal distribution function as given below.

$$X_1(z_1) = 1 - \Phi\left(\frac{z_1 + .5m^2}{m}\right) = 1 - \Phi\left(\frac{z_1}{m} + .5m\right) \quad (6.9)$$

At the extremes of the z_2 range,

$$a = 1 - \Phi\left(\frac{z_a - z_0}{m} + .5m\right) \qquad b = 1 - \Phi\left(\frac{z_b - z_0}{m} + .5m\right) \quad (6.10)$$

Because the mean of the SN distribution is $+.5m^2$, the formula for the Y_1 probabilities will be quite similar to Eq. 6.9.

$$Y_1(z_1) = 1 - \Phi\left(\frac{z_1 - .5m^2}{m}\right) = 1 - \Phi\left(\frac{z_1}{m} - .5m\right) \quad (6.11)$$

The areas remaining under the normal density curves inside the abrupt truncation point are K_N and K_{SN}.

$$K_N = b - a = \Phi\left(\frac{z_a - z_0}{m} + .5m\right) - \Phi\left(\frac{z_b - z_0}{m} + .5m\right) \quad (6.12)$$

$$K_{SN} = \Phi\left(\frac{z_a - z_0}{m} - .5m\right) - \Phi\left(\frac{z_b - z_0}{m} - .5m\right) \quad (6.13)$$

The general equation for the metastatic ROC, Eq. 3.37, becomes the following for the metastatic normal ROC.

$$X_{MN}(z) = \left[\Phi\left(\frac{z_a - z_0}{m} + .5m\right) - \Phi\left(\frac{z - z_0}{m} + .5m\right) \right] \Big/ K_N$$

$$(6.14)$$

$$Y_{MN}(z) = \left[\Phi\left(\frac{z_a - z_0}{m} - .5m\right) - \Phi\left(\frac{z - z_0}{m} - .5m\right) \right] \Big/ K_{SN}$$

$$(6.15)$$

In the following sections, the problem of fitting a given

ROC character with the metastatic normal ROC character of Eq. 6.8

is attacked. The metastatic normal is given by three parameters.

Given a specific ROC character, three salient characteristics must be

selected to determine the matching metastic normal. It is assumed

that the ROC character to be fit contains a simple mode and has the general

appearance of being normal about this mode. This can be most easily

seen by looking at the logarithm of the ROC character, since the log-

arithm of the normal character is a simple parabola. It is also

assumed that the character to be fit covers a bounded z-range.

6.2 First Procedure

In this section the equations necessary to fit a meta-

static normal ROC character in a given ROC character are based on

the minimum value and mode value of that given character, and the

ratio of the character values at the mode and the minimum. It will be

assumed that the maximum value of z is above the mode, but no other

use will be made of the maximum value. This information is summarized

in Eq. 6.16.

Known: z_b, z_0, that $z_a > z_0$, $\quad \pi(z_0)/\pi(z_b) = r_b > 0$ (6.16)

This procedure is based on only the information given

in Eq. 6.16. The metastatic normal ROC character, Eq. 6.8, evaluated

at the mode, is

$$\pi_{MN}(z_0) = \frac{1}{m\sqrt{K_N K_{SN}}} \, e^{-\frac{m^2}{8}} \frac{1}{\sqrt{2\pi}} \qquad (6.17)$$

The metastatic normal ROC character differs from this value by only

an exponential term. Equation 6.17 can be used to simplify the writing

of the off-mode values of π_{NM}. The known information relates the

character values at the mode and at the minimum value,

$$\pi_{MN}(z_b) = \pi_{MN}(z_0) \, e^{-\frac{(z_b - z_0)^2}{2m^2}} \qquad (6.18)$$

Use this equation to determine the value for the parameter m

$$m = \frac{z_b - z_0}{\sqrt{2 \ln r_b}} \qquad (6.19)$$

The detection index, m^2, of the original normal ROC curve, has been

determined. The lower z_1 cut point $z_b - z_0$ was contained in the

original information. To determine the upper z_1 cut point, $z_a - z_0$,

requires a devious route which is not always successful. The only

clue to the value of the upper cut point lies in the normalizing constants

K_N and K_{SN}. Rewriting Eq. 6.4 in an equivalent form

$$e^{z_0} K_{SN} = K_N \qquad (6.20)$$

the task of obtaining z_a begins. In the equation for the K values,

Eqs. 6.12 and 6.13, z_a and z_b appear in both equations. Expand

Eq. 6.20 using Eqs. 6.12 and 6.13. All the terms containing z_a can

to moved to one side of the equation and all the terms containing z_b to

the other side of the equation.

$$\Phi\left(\frac{z_a - z_0}{m} + .5m\right) - e^{z_0} \Phi\left(\frac{z_a - z_0}{m} - .5m\right)$$

$$\qquad (6.21)$$

$$\equiv \Phi\left(\frac{z_b - z_0}{m} + .5m\right) - e^{z_0} \Phi\left(\frac{z_b - z_0}{m} - .5m\right)$$

The right-hand side of this equation is identical with the left-hand

side, except that z_b appears in place of z_a. Let

$$D(z) = \Phi\left(\frac{z - z_0}{m} + .5m\right) - e^{z_0} \Phi\left(\frac{z - z_0}{m} - .5m\right) \qquad (6.22)$$

Equation 6.21 is simply

$$D(z_a) = D(z_b) \qquad (6.23)$$

The value of $D(z_b)$ is known. If the function $D(z)$ takes on each of its

possible values exactly twice, then there would always be a solution

for z_a. For very small arguments nearing $-\infty$, the normal distribution function approaches zero. Therefore, for very large negative values of z, D will be nearly zero.

$$D(-\infty) = 0 \qquad (6.24)$$

The sign of the derivative of $D(z)$ is positive for z values less than $.5z_0$ and negative for z values greater than $.5z_0$.

$$\text{Sign } D'(z) = \text{sign } (.5z_0 - z) \qquad (6.25)$$

This means that D takes on its values at most twice. For large arguments $D(z)$ approaches $1 - e^{z_0}$.

$$D(+\infty) = 1 - e^{z_0} \qquad (6.26)$$

If the mode value, z_0, is positive the D function takes on negative as well as positive values to the right of $.5z_0$, while it takes on only positive values to the left of $.5z_0$. Therefore, if the known value z_b, which is below the mode, is also below half the mode, then we are guaranteed that for a positive z_0 a solution exists for Eq. 6.23.

If z_0 is negative, the range of values to the right of the peak of D is less than the range of values to the left of the peak. Therefore if z_b is too far below the mode, z_0, there will be no value z_a that satisfies Eq. 6.23. This is summarized below.

$$z_0 < 0, \quad D(z_b) < 1 - e^{z_0} \implies \text{no solution for } z_a$$

$$(6.27)$$

Given the information in Eq. 6.16, to use Procedure I

for determining the fitting ROC character, one must proceed partly

through the solution before knowing whether the procedure will be

successful or not. The solution consists essentially of Eq. 6.19 for

the parameter m describing the strength of the original normal ROC

and Eq. 6.23 from which one determines the upper cut value.

First Example: Consider the ROC character of the conic section

which is Curve 1 of Table 5.4 and Fig. 5.13. This ROC curve and

character were matched to an extent by a complete normal curve, in

Fig. 5.15. The equation for this character is

$$\pi(z) = (705)^{.5} (28 + e^z + e^{-z})^{-1.5} \quad , |z| < 2.77$$

$$(6.28)$$

Although the complete ROC character is known, pick out simply the

lowest value of the range, the mode value, and the ratio of the mode

character to the character at the minimum value.

$$\text{Given: } z_b = -2.77 \quad z_0 = 0 \quad \pi(z_0)/\pi(z_b) = 1.772$$

$$(6.29)$$

Based on this information alone, find the fitting metastatic ROC char-

acter. The parameter m is found from using Eq. 6.19. Numerically

this results in

$$m = (2.77)/\sqrt{2 \ln 1.772} = 2.60; \quad m^2 = 6.76 \quad (6.30)$$

(The work leading to Fig. 5.15 utilized a normal ROC with parameter

d = 2.31 as an approximation.) The present procedure will use a normal ROC

with strength 6.76, but will not utilize the entire curve. Since the

mode is zero, the equation for the function d is symmetric and the

solution for the upper cut value is immediate.

$$D(z_a) = D(z_b) \implies z_a = +2.77 \tag{6.31}$$

All that remains is to evaluate the K parameter using Eqs. 6.12 and 6.13,

and to write out the expressions for the metastatic normal ROC using

Eqs. 6.14 and 6.15. This results in

$$K_N = K_{SN} = .400152 \tag{6.32}$$

$$X = [.991106 - \Phi(t + 1.3)] / .400152$$

$$\tag{6.33}$$

$$Y = [.409046 - \Phi(t - 1.3)] / .400152$$

A character with known ROC curves has been used in

order that we might evaluate this procedure of fitting an ROC curve by

fitting three parameters of the ROC character. The true ROC curve

for this case has been computed with great accuracy on the IBM 7090

computer for use in Chapter V. In order to test the usefulness of the

procedure for quick hand calculations, all of the parameters, and the

ROC curves of Eq. 6.33, were calculated using a handbook of normal

tables and an ordinary ten-inch slide rule. The result of the ROC curve

evaluation is given in Table 6.1, along with the true value of the ROC

curve as obtained by its computer.

X	.0125	.0224	.0496	.0890	.1450	.2200
Y_{MN}	.160	.249	.416	.560	.685	.782
Y_{TRUE}	.161	.251	.418	.563	.683	.781

Table 6.1. Comparison of true and approximate ROC curves

The combination of errors in the procedure and in the slide rule technique resulted in a discrepancy that can not be observed when the ROC is plotted in the one percent to 99 percent range, as it is in Fig. 6.2.

In Fig. 6.2 the hyperbola and the fitting metanormal plot on top of each other. Also on that graph is plotted original normal ROC from which the metastatic normal was cut out. The points on the original normal leading to the metastatic normal are indicated.

To summarize this example, the first fitting procedure has been used on a shallow symmetric ROC character. The resultant fit utilized only forty percent of an original normal ROC curve. The resultant fit to the ROC curve was excellent.

Second Example: For a second example consider the nonsymmetric ROC character that appeared in Fig. 5.24, and whose ROC passes through the negative diagonal point (.3, .7). We shall use only the mode value, the minimum value, and the ratio of the ROC character values at the mode and the minimum.

Given: $z_0 = -0.2 \quad z_b = -2.3 \quad \pi(z_0) / \pi(z_b) = 7.5$ (6.34)

The value of the constant m is found by Eq. 6.19.

$$m = (2.1)/\sqrt{2 \ln 7.5} = 1.04 \qquad (6.35)$$
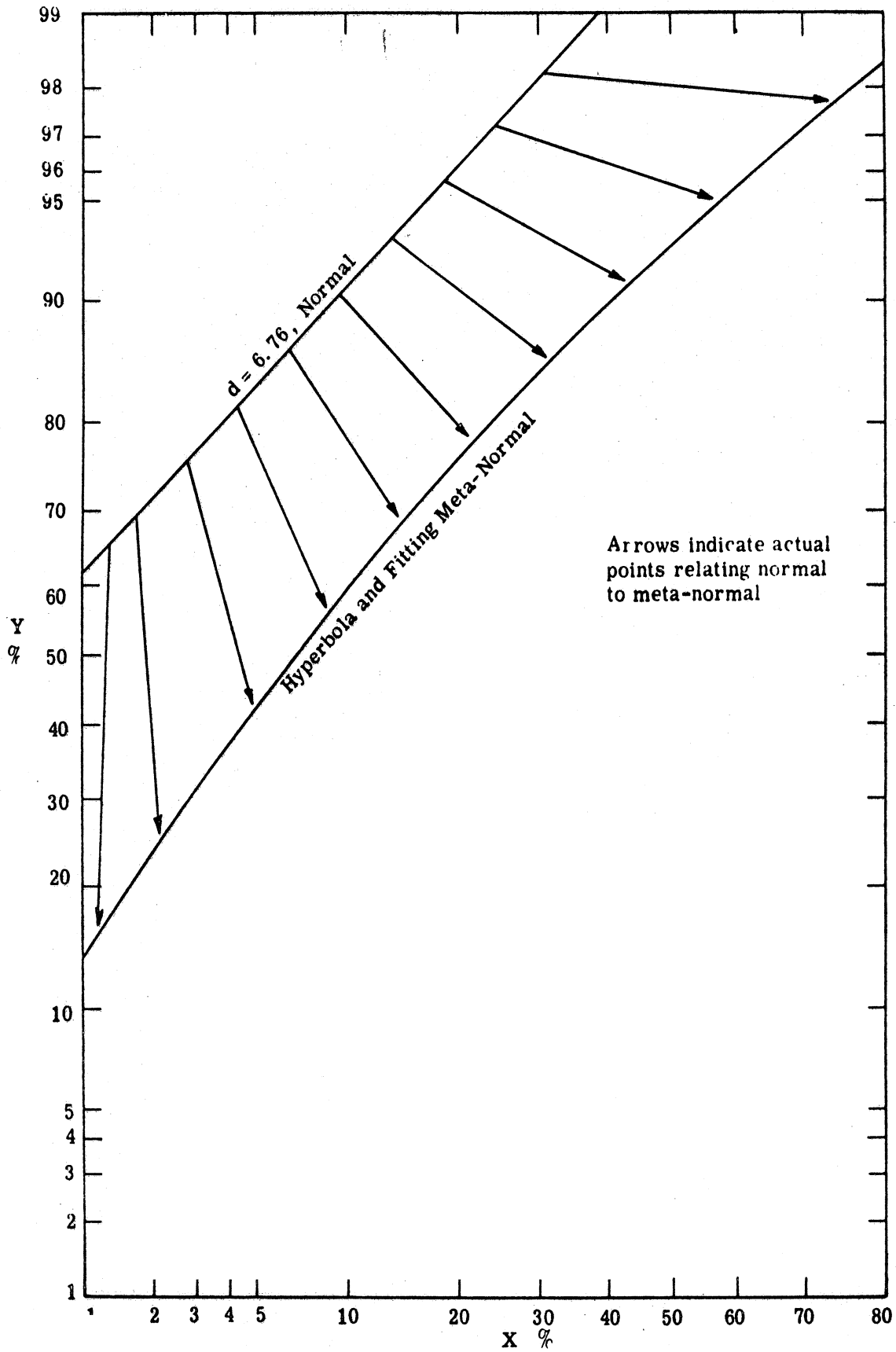
Fig. 6.2. ROC curves showing Normal ROC
and its metastatic image

In order to utilize the D function of Eq. 6.22, e^{z_0} is determined and then the D function evaluated at the lower limit. To check for the existence of a solution, $D(+\infty)$ is calculated by Eq. 6.26. The results are

$$e^{z_0} = .81873 \qquad D(z_b) = .06345 \qquad D(+\infty) = .18127 \qquad (6.36)$$

The value .06345 is taken on only once by D, and Procedure I yields no solution.

There are many ways to explain the existence of this nonsolution. The most apparent is that for this value of mode, the function D takes on more values to the left of its peak than to the right of its peak.

6.3 Second Procedure

The second procedure is similar to the first, and is utilized when the first fails. The first procedure's failure meant the D function took on more values to the left of its peak than to the right of its peak. The second procedure also uses the mode z_0, and hence the same function D. However, if the known value is z_a, a solution is guaranteed. As in the first procedure the ratio of the character values at the mode and the known cut are given. This is formalized below.

$$\text{Known: } z_a, \quad z_0, \quad \text{that } z_b < z_0, \quad \pi(z_0)/\pi(z_a) = r_a \qquad (6.37)$$

As in the first procedure this amount of knowledge is sufficient to determine the parameter m. The expression is

$$m = \frac{\left| z_a - z_0 \right|}{\sqrt{2 \ln r_a}} \qquad (6.38)$$

The second part of the solution is to utilize the D function defined by

Eq. 6.22 and to solve Eq. 6.23. In contrast to Procedure I, the value

of the upper cut z_a is known, and Eq. 6.23 is utilized to determine z_b.

Third Example: Let us return to the example for which the first

procedure failed. This was the ROC character given in Fig. 5.24

with the negative diagonal ROC point $(.3, .7)$. The information neces-

sary for the second procedure is the mode value of -0.2 and the upper

cut value of +4.6, and the ratio of the ROC character at the upper cut

and the mode. The values are given in Eq. 6.39.

$$\text{Given:} \quad z_0 = -0.2 \qquad z_a = +4.6 \qquad \pi(z_0)/\pi(z_a) = 288$$

$$(6.39)$$

The value of the parameter m, this time using Eq. 6.38, is

$$m = (4.8)/\sqrt{2 \ln 288} = 1.426 \qquad (6.40)$$

In this second procedure m is 1.426, whereas it was 1.04 in the first

unsuccessful attempt. First step numerical values are shown in Eq. 6.41.

$$e^{z_0} = .81873 \qquad D(z_a) = .18452 \qquad D(-\infty) = 0 \qquad (6.41)$$

$D(-\infty) < D(z_a)$ means a solution to Eq. 6.23 exists. The numerical

values obtained here, and the trial-and-error solution for z_b was

carried out using a slide rule. The solution only took five iterations,

and the accuracy is good to the second decimal place in z_b.

$$z_b = -2.65 \tag{6.42}$$

$$K_N = 1 - .000023 - .193522 = .806455$$

$$K_{SN} = 1 - .004001 - .010983 = .985016 \tag{6.43}$$

$$X = [.999977 - \Phi(t + .713)]/.806455$$

$$Y = [.995999 - \Phi(t - .713)]/.985016 \tag{6.44}$$

An intuitive feel for the determination of the metastatic ROC curve is obtained from the probabilities of occurrence on the original z_1 axis. This is sketched in Fig. 6.3.

N   .193522 | 3455 | .000023

          -1.578   3.365

SN   .010983 | .985016 | .004001 z_1
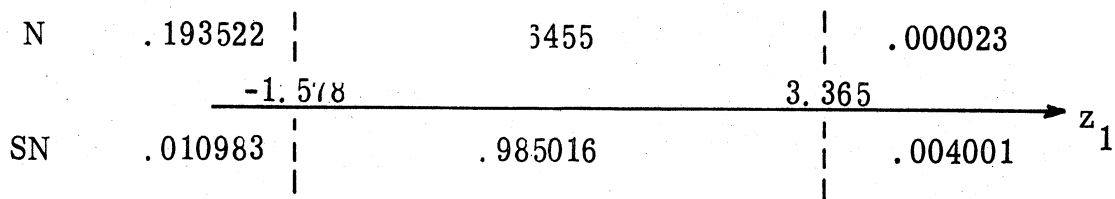
Fig. 6.3. Sketch of probabilities on original axis

The constant $K_N = .806455$ means that this amount of N-probability falls between the two cut values used in obtaining the metastatic ROC. Almost all of the remaining N-probability fell below the lower cut, with only a small amount being chopped off above the upper cut. The fact that the mode value is negative means the value of K_{SN} is bigger than K_N; more of the SN-probability was retained between the two cut values than was N-probability. Although more of the SN-probability

fell below the lower than fell above the upper cut, the proportion is nowhere near as large as for the N case.

The metastatic normal ROC curve of Eq. 6.44 and the true ROC curve cross in the neighborhood of .001 false alarm and .037 detection probability. By the time the curves get into the ROC region usually displayed on normal-normal paper, the ROC of Eq. 6.44 is about .1 in d' above the desired ROC curve.

The second procedure is based on establishing a metastatic ROC character that has the same maximum z value, the same mode value, and the same ratio of ROC character at mode and maximum z values. The match should be relatively good in the region of the maximum. If the match were perfect at the maximum, it would be perfect at the mode value. The ROC character of the second example is shown in Fig. 6.5. The fit falls somewhat below the true value at the maximum, $z = 4.6$. This is relatively unimportant because of the extremely small value of the character here. However, it is reflected in a corresponding discrepancy in the neighborhood of the mode, where more probability is concentrated. The original ROC character was hyperbolic, and the limiting slope on semi-log paper behaves like $e^{-1.5z}$. A parabola cannot match such a straight line for much of a z-range. Therefore, the approximation developed in Procedure II is first below, then above, and then below this straight line section. This accounts for the ROC curve falling above the desired locus.
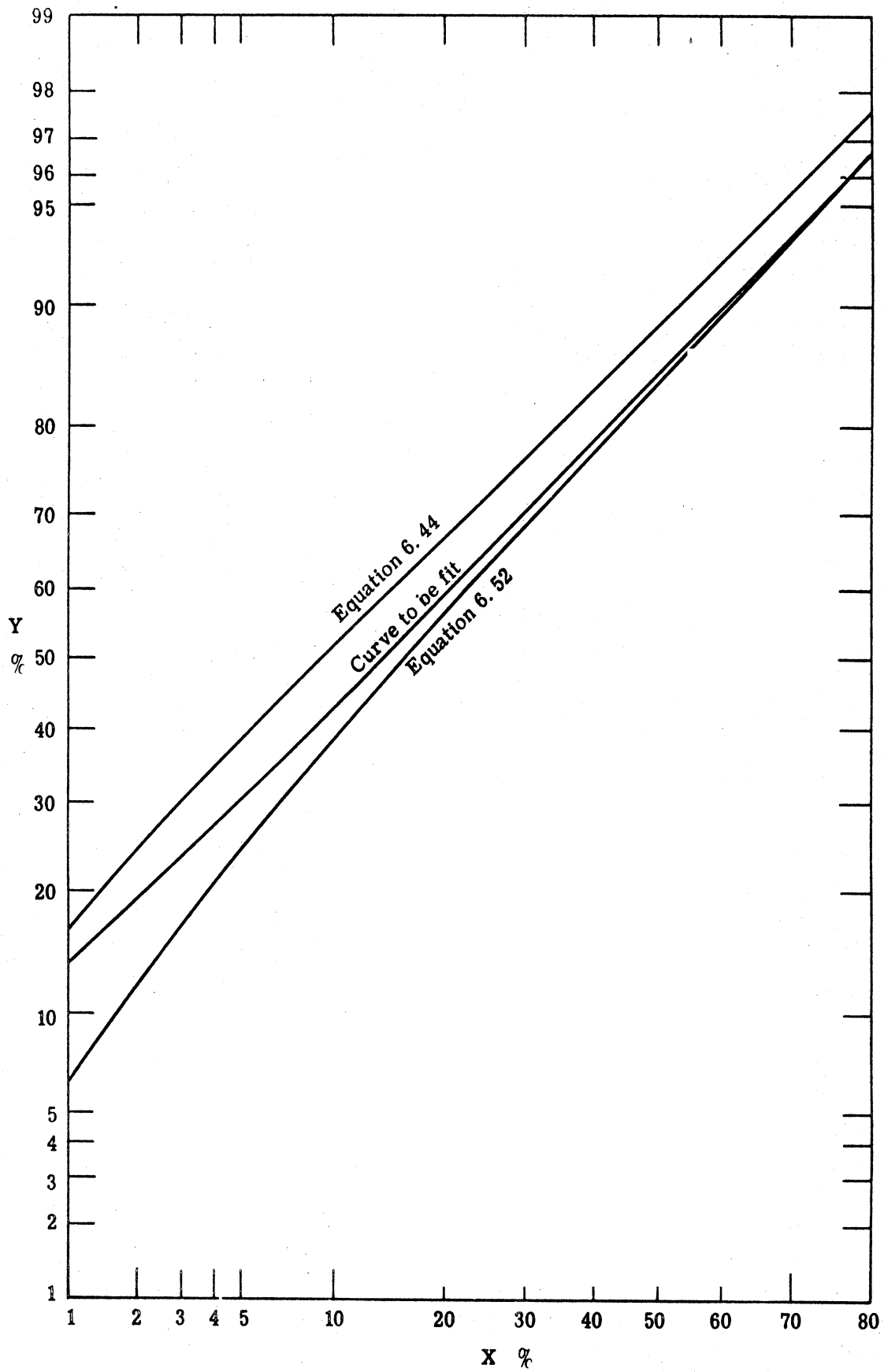
Fig. 6.4. Comparison of third and fourth examples
with true curve
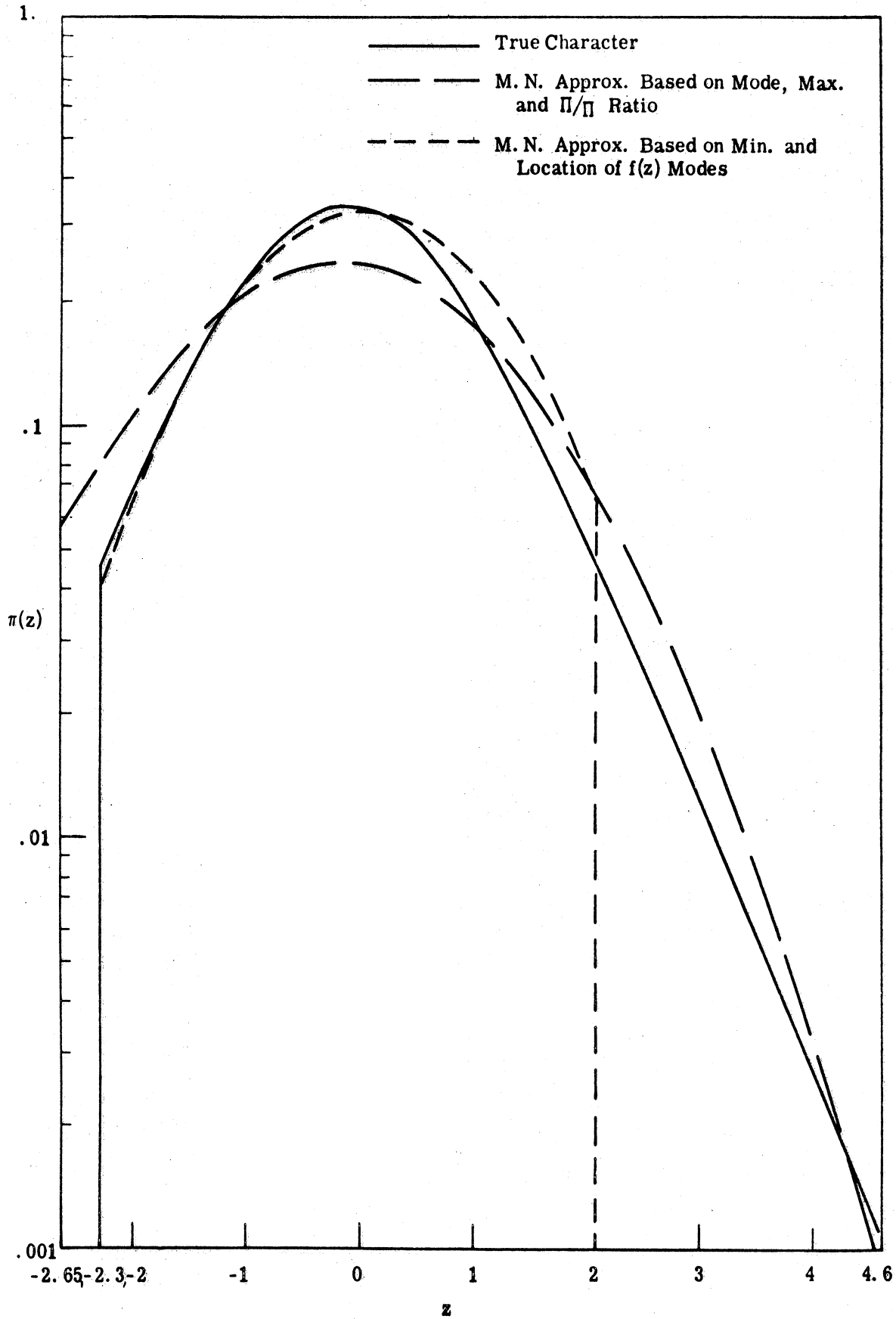
Fig. 6.5. Comparison of characters of third
and fourth examples

6.4 Third Procedure

The first two procedures utilized the ratio of the ROC character at the mode and at an extreme point. The matching meta-static normal ROC character covered the same relative range of $\pi(z)$ values. For the medium probability range between one and 99 percent, the behavior of the ROC character when its magnitude is one percent of the mode magnitude is of little importance. Another characteristic of any ROC character which appears somewhat parabolic on semi-log paper is the "width" of the hump. Sharply peaked characters correspond to poor detection, broad characters correspond to better detection. Another item of interest about the ROC character are the points on the character which correspond to the mode of the z distribution under condition N and condition SN. When the slope of the derivative of the logarithm of the character is +. 5, the derivative of the N density function for z will be zero. Correspondingly, the slope -. 5 at a somewhat larger z value identifies the mode of the SN density function. The third procedure shall assume knowledge of one of the extreme values of the z-range, and the position of the two modes of the z density functions. If these two modal values are widely separated, then the logarithm of $\pi(z)$ must be very slowly changing slope, and be very broad. Conversely, if the two points are close together, the ROC curve corresponds to poor detection. Formalizing what is known for Procedure III we write

Known: z_a or z_b , $\left.\dfrac{d\ln\pi(z)}{dz}\right|_{z=z_c} = +0.5$, $\left.\dfrac{d\ln\pi(z)}{dz}\right|_{z=z_d} = -0.5$

$$(6.45)$$

$$z_b < z_c < z_d < z_a$$

The solution for this procedure is fairly straightforward. The mode of the matching metastatic normal ROC character will lie midway between the two points with slope $+1/2$ and $-1/2$. These two points also determine parameter m.

$$z_0 = .5(z_c + z_d) \quad , \quad m = \sqrt{z_d - z_c} \qquad (6.46)$$

The third step is the same as the third step in the first two procedures, namely, to solve Eq. 6.23 for the unknown cut value, once z_0 and one cut value are known.

Fourth Example: As an example of this third procedure consider the same character for which the first procedure failed, and for which the second procedure was overly generous. From the graph of the character the points with required slope were estimated to be -0.6 for the N mode and +0.7 for the SN mode.

$$\text{Given: } z_b = -2.3 \qquad z_c = -0.6 \qquad z_d = +0.7 \qquad (6.47)$$

In rapid order solve for z_0 and m.

$$z_0 = +0.05 \qquad e^{z_0} = 05127 \qquad m = \sqrt{1.3} = 1.14 \qquad (6.48)$$

Check to see if a solution for the D equation exists.

$$D(z_b) = +.06362 \qquad D(+\infty) = -.05127 \qquad (6.49)$$

A solution of the D equation is possible, and in a few iterations was determined to be

$$z_a = 2.044 \qquad (6.50)$$

The next steps are to determine the K constants and to obtain the detailed equation for the ROC curves.

$$K_N = 1 - .010170 - .068112 = .921718$$

$$(6.51)$$

$$K_{SN} = 1 - .119000 - .004269 = .876731$$

$$X = [.989830 - \Phi(t + .57)]/.921718$$

$$(6.52)$$

$$Y = [.881000 - \Phi(t - .57)]/.876731$$

A sketch very much like Fig. 6.3 has been drawn for this case and is shown in Fig. 6.6.

N	.068112		.921718		.010170
		-2.06		1.75	
SN	.004269		.876731		.119000
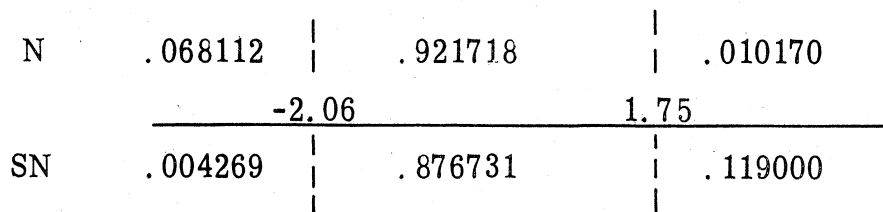
Fig. 6.6. Sketch of probabilities on original axis

This approximate solution leaves less of the SN-probability in the middle than N-probability, and that far more of the high z values are chopped off than with the second procedure.

The ROC curve and the ROC character were plotted in Figs. 6.4 and 6.5 respectively. In Fig. 6.5 the match is very good all along the ROC character from the minimum point up to the neighborhood of the mode. The corresponding segment of the ROC curve is that dealing with the high probabilities. However, the maximum value z of the metastatic normal is only a little over two compared to the value of 4.6 for the given ROC. On the ROC curve, as a point proceeds from high probability to low probability this metastatic normal drops away from the desired ROC curve.

6.5 Summary

The purpose of these procedures is to determine the parameters of a metastatic normal ROC curve whose ROC character matches three salient distinguishing items of a given character. The corresponding ROC is obtained with the aid of normal tables and simple calculation.

CHAPTER VII

BINORMAL ROC

7.1 General Development

The binormal ROC curves have appeared in the literature as a by-product of the popular use of normal ROC curves, and use of normal-normal graph paper for plotting empirical ROC data. Normal ROC curves plot as straight lines on normal-normal graph paper, with a slope of one, that is, parallel to the chance diagonal. A binormal ROC curve plots as a straight line on normal-normal paper but with slope different from one.

The use of binormal ROC curves adds a second parameter to the class of ROC curves based on normal distribution functions. They have been very useful, and investigators have learned to interpret the slope as indicative of additional variance under one of the causes, N or SN, compared to the other. However, a strict interpretation of the binormal ROC curve is as a nonregular ROC. The equations for the relation of the logarithm of the likelihood ratio to the points along the ROC curve will be developed in Section 7.2. The ROC curve can be made regular by either one of two (usually slight) modifications. The first of these is treated in Section 7.3, and is called external rectification. External rectification corresponds to adjoining the convex closure to the regular part of the

binormal ROC. The second change, made in Section 7.4, is called internal rectification. Internal rectification corresponds to deriving an ROC based on the likelihood ratio of the decision axis that led to the original nonregular binormal ROC curve. The effects of these modifications on data fitting procedures, and upon the ROC character, will be discussed in Section 7.5.

In an article in Science (Ref. 27), J.A. Swets proposed the ROC curve as a basis for measuring effectiveness of information retrieval systems. Figure 7 of that article was a particular type of binormal ROC family; it is redrawn on normal-normal paper in Fig. 7.1. The letter E is Swets' index for effectiveness. Although the curve with effectiveness $E = 0.5$ is nearly parallel to the chance diagonal, $E = 0.0$, as the effectiveness increases the slope is decreasing slightly with each successive step. These curves are binormal ROC curves. Figure 7.2 is Swets' curve[1] comparing empirical data to the binormal ROC curves for two information retrieval techniques. Figure 7.3 is the same data drawn on normal-normal paper, with best-fitting straight lines through the data.[1] The values for s marked in this figure, 0.79 and 0.86, are the slopes of the best-fitting straight line on this paper. The binormal ROC curves fit the data very much better than the best normal fit.

---

[1] By permission of J. A. Swets. These appeared in an unpublished work.
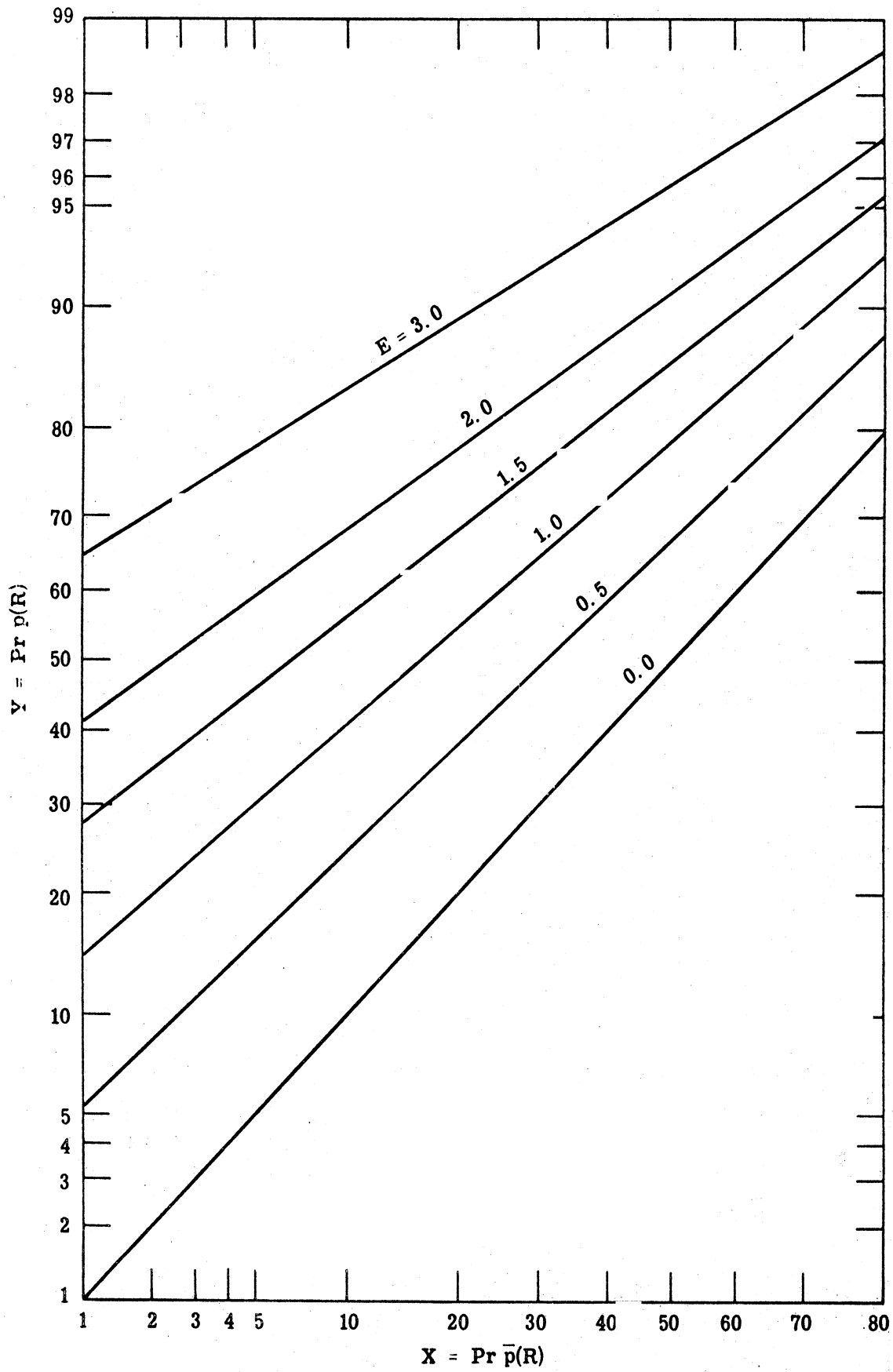
Fig. 7.1. A fan of binomial ROC curves
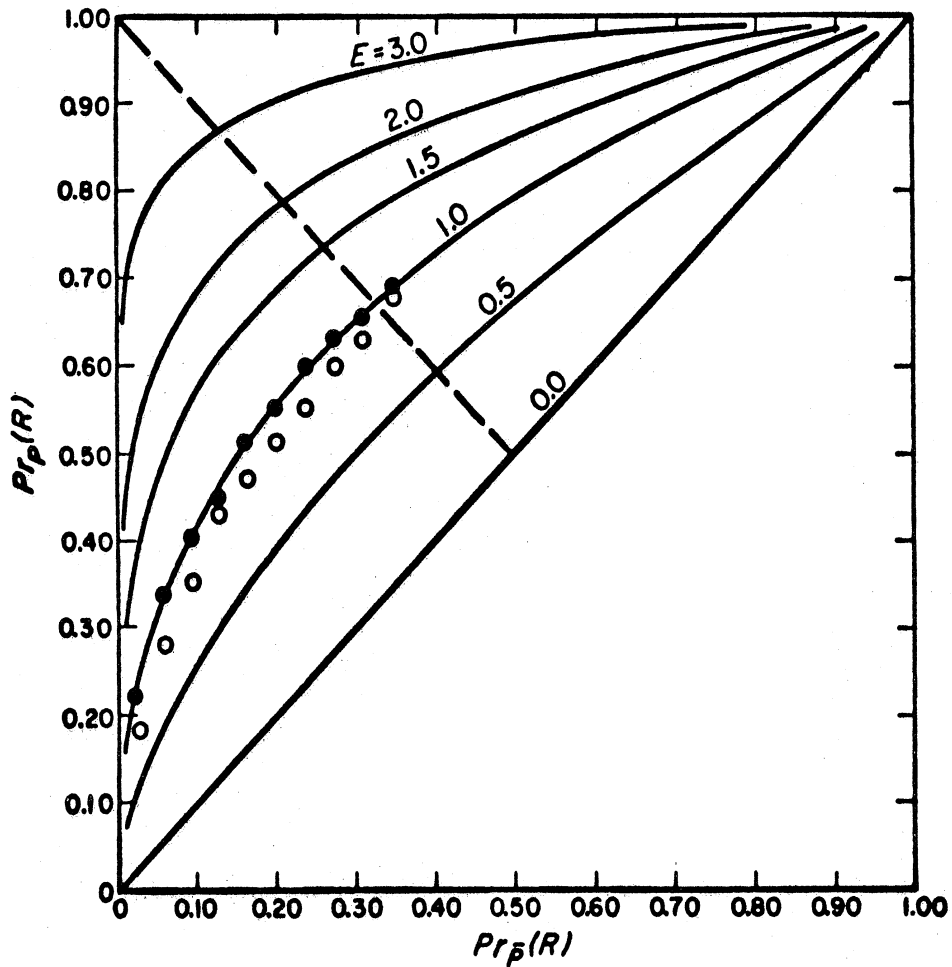
The empirical operating characteristics for two retrieval
techniques and various theoretical operating characteristics
based on the assumption of Gaussian probability distributions
with a variance proportional to the mean, specifically, with a
ratio of mean to standard deviation of 4. 0.

Fig. 7. 2. Swets' first curves
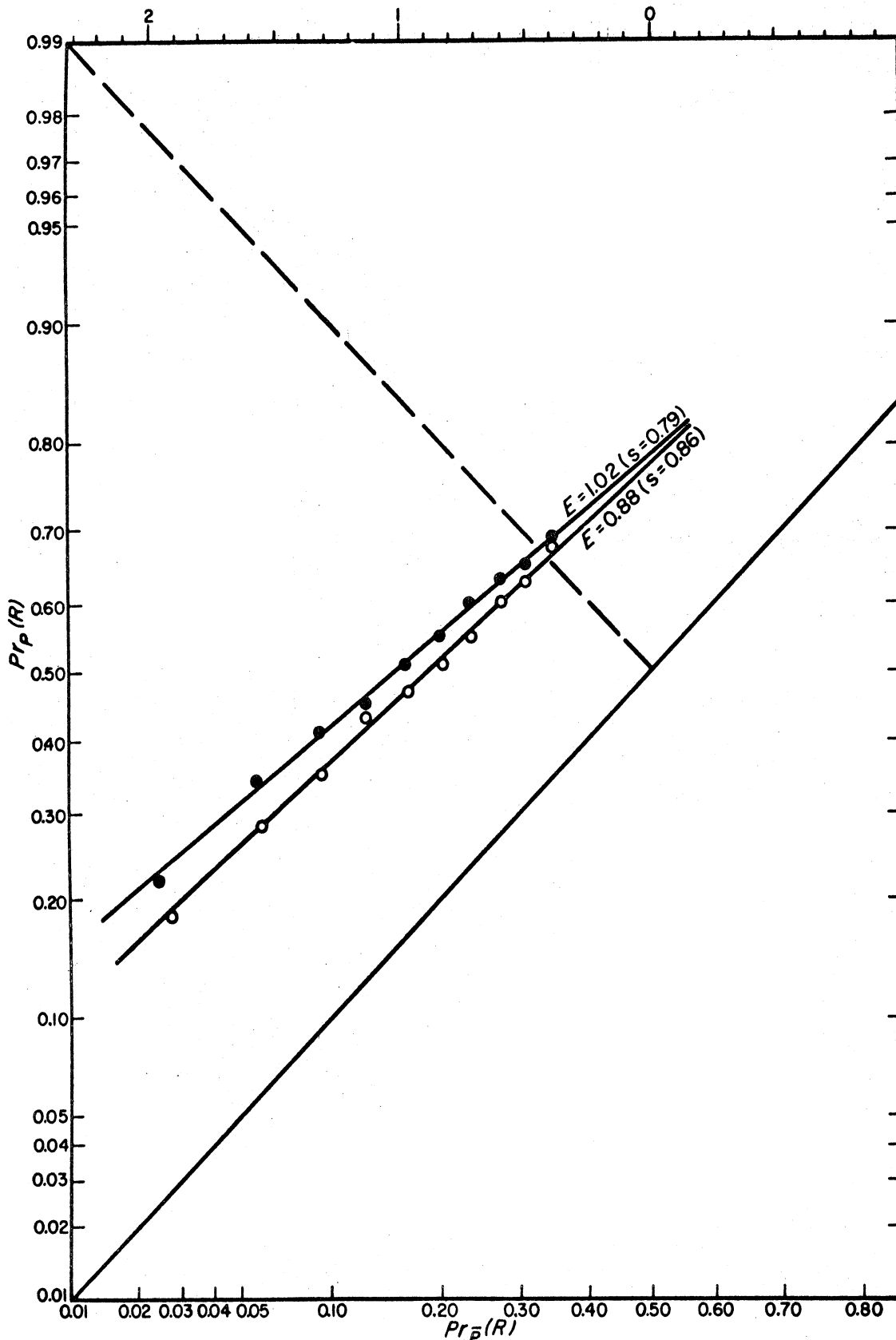
The data reproduced on probability scales, with the best-fitting theoretical curves assuming Gaussian probability distributions. The slopes of the curves are the best-fitting slopes.

Fig. 7.3. Swets' curves on normal-normal paper

This example has two features typical of applications of the binormal ROC. The first is that the slope is less than one. The second is that only a portion of a binormal ROC curve has been utilized to fit the data. In Fig. 7.2 the data points lie near or to the left of the negative diagonal.

7.2 Formal Development

One form of the equation for the normal ROC curve is

$$X = \Phi(-t) \qquad Y = \Phi(\sqrt{d} - t) \qquad (7.1)$$

The function $\Phi(\ )$ is the normal distribution function. When a normal ROC curve is plotted on normal-normal paper, the linear horizontal coordinate of a point is $-t$ when the linear vertical coordinate is $\sqrt{d} - t$. This is a straight line, with slope one, displaced vertically above the chance diagonal by a distance $\sqrt{d}$. The letter s is often used to denote the slope on normal-normal paper. E is used so often in probability theory for expected value, that the letter q (for quality) shall be used as the index of effectiveness. The binormal ROC curves can be written in a manner analagous to the normal ROC as

$$X = \Phi(-t) \qquad Y = \Phi(s(q-t)) \qquad 0 < s \neq 1 \qquad (7.2)$$

The geometric interpretation of s is the slope of the ROC curve on normal-normal graph paper. The geometric interpretation of the quality index, q, is the distance between the binormal and the chance diagonal at the fifty percent detection point. This distance may be

measured either horizontally or vertically. Table 7.1 gives the

separation from the chance diagonal, the d' value (Ref. 10) at

several locations along the binormal ROC curve.

Location	Value of d' $= \Phi^{-1}(Y) - \Phi^{-1}(X)$	$\dfrac{dY}{dX}$
Y = .50	q	$se^{.5q^2}$
Neg. Diag.	$2sq/(1+s)$	s
X = .50	sq	$se^{-.5s^2 q^2}$

Table 7.1. Special Values of d'

Let us first look at a particular binormal graphically, to see what is

objectionable about the binormal ROC. Figure 7.4 shows a binormal

ROC with quality 1 and the rather small slope of 1/2 drawn on normal-

normal paper. In this form, there seems to be very little that is

obviously wrong with the ROC curve. The same ROC curve is displayed

on linear coordinates in Fig. 7.5(a). On this scale it is evident that

the curve is an S-shaped curve, and not convex. From the point (0, 0)

toward the point (1, 1) the curve convex is a regular ROC, then has a

nearly linear portion where the slope is a minimum, and then the slope

increases again, in fact to infinity, as the ROC goes to the point (1, 1).

This same ROC has been reproduced in Fig. 7.5(b),

together with the chance diagonal and one other straight line. The point

marked C in Fig. 7.5(b) is the point at which the binormal ROC

curve crosses the chance diagonal. From this point on to the right
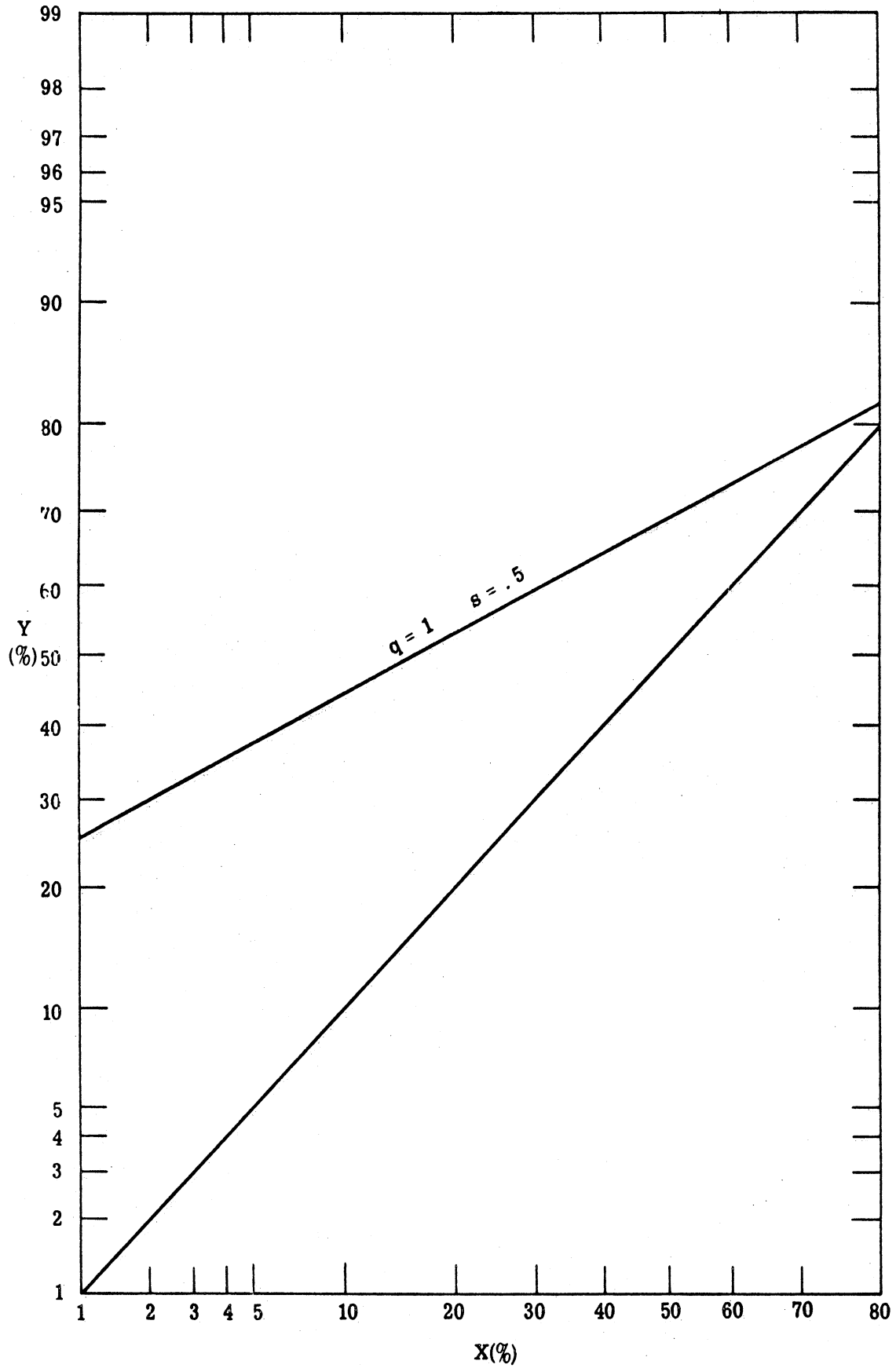
Fig. 7.4. A binormal ROC curve

(a)

(b)
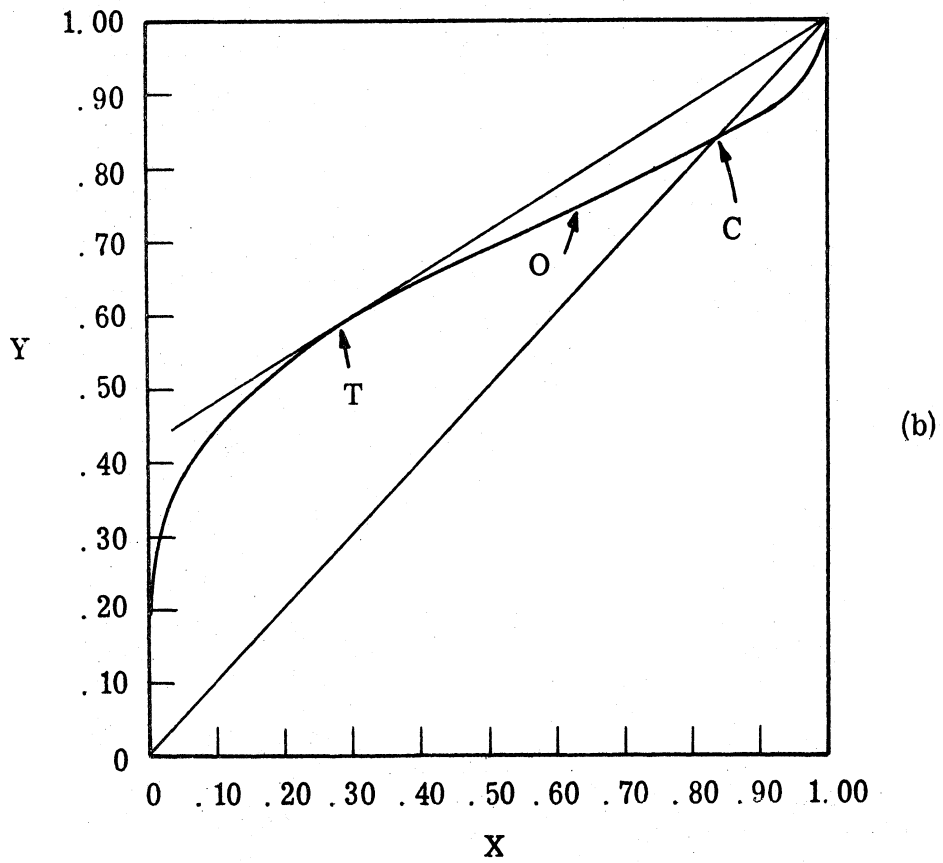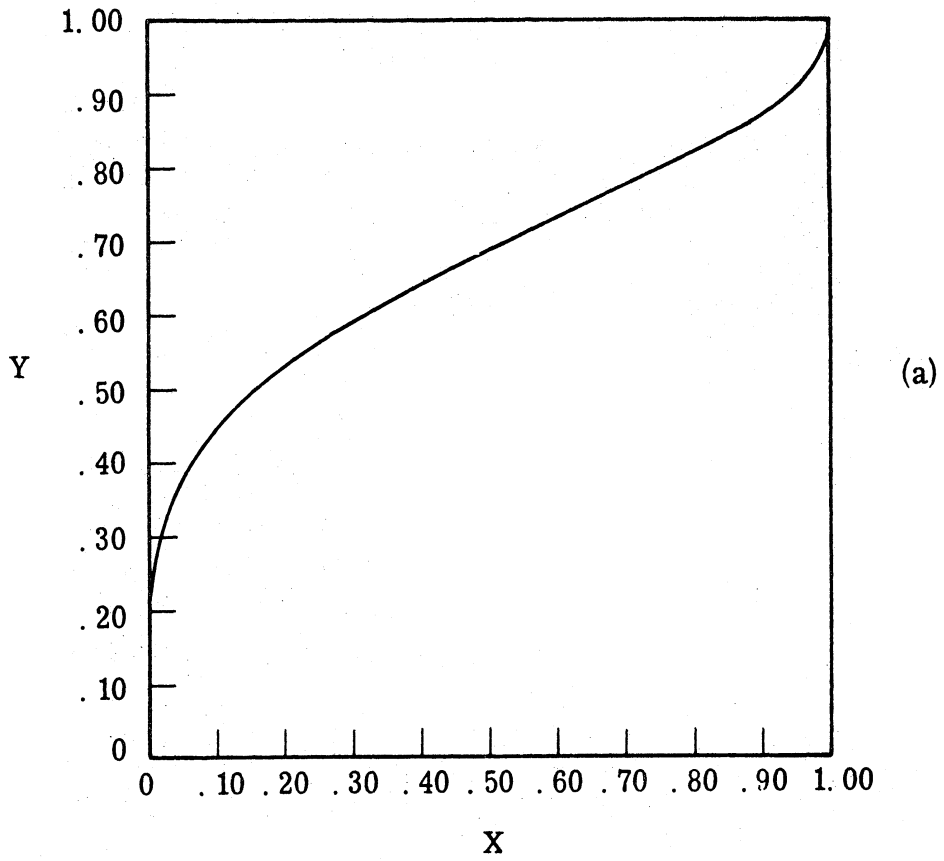
Fig. 7. 5. Binormal ROC curve, q = 1, s = . 5

the binormal ROC curve falls below the chance diagonal. Does this always happen? The binormal ROC curve is a straight line on normal-normal paper, and not parallel to the chance diagonal; it must intersect the chance diagonal. For $s < 1$ the part of the binormal below chance is near $(1, 1)$; for $s > 1$ the part of the binormal below chance is near $(0, 0)$. The point on the ROC curve marked with an arrow and zero indicates the point of minimum slope to the ROC curve. The point of minimum slope of a regular ROC curve always occurs at the upper right-hand corner $(1, 1)$. Whenever the minimum slope occurs before the upper right-hand corner, the ROC curve must be nonregular. The additional line drawn on this figure is a tangent to the ROC curve that goes through the upper right-hand corner. The point of tangency is indicated with the arrow marked T. The ROC curve formed by piecing together that part of the binormal ROC curve from the origin to the point of tangency (marked T) and the straight line from the point T to the upper right-hand corner, is a regular ROC curve. That part of the binormal ROC that is used, between the origin and the point of tangency, is called "the regular part" of the binormal ROC. The straight line between the point of tangency and the upper right-hand corner is called the "tangent inferior." (The tangent line to any regular ROC curve that has the smallest slope of any tangent is the "tangent inferior"; the tangent line to a regular ROC curve that has maximum slope is the "tangent superior.") In Fig. 7.6 the binormal ROC curve with quality $q=2$ and slope $s=.5$ is drawn on normal-normal paper together with the tangent inferior. The point of tangency, the
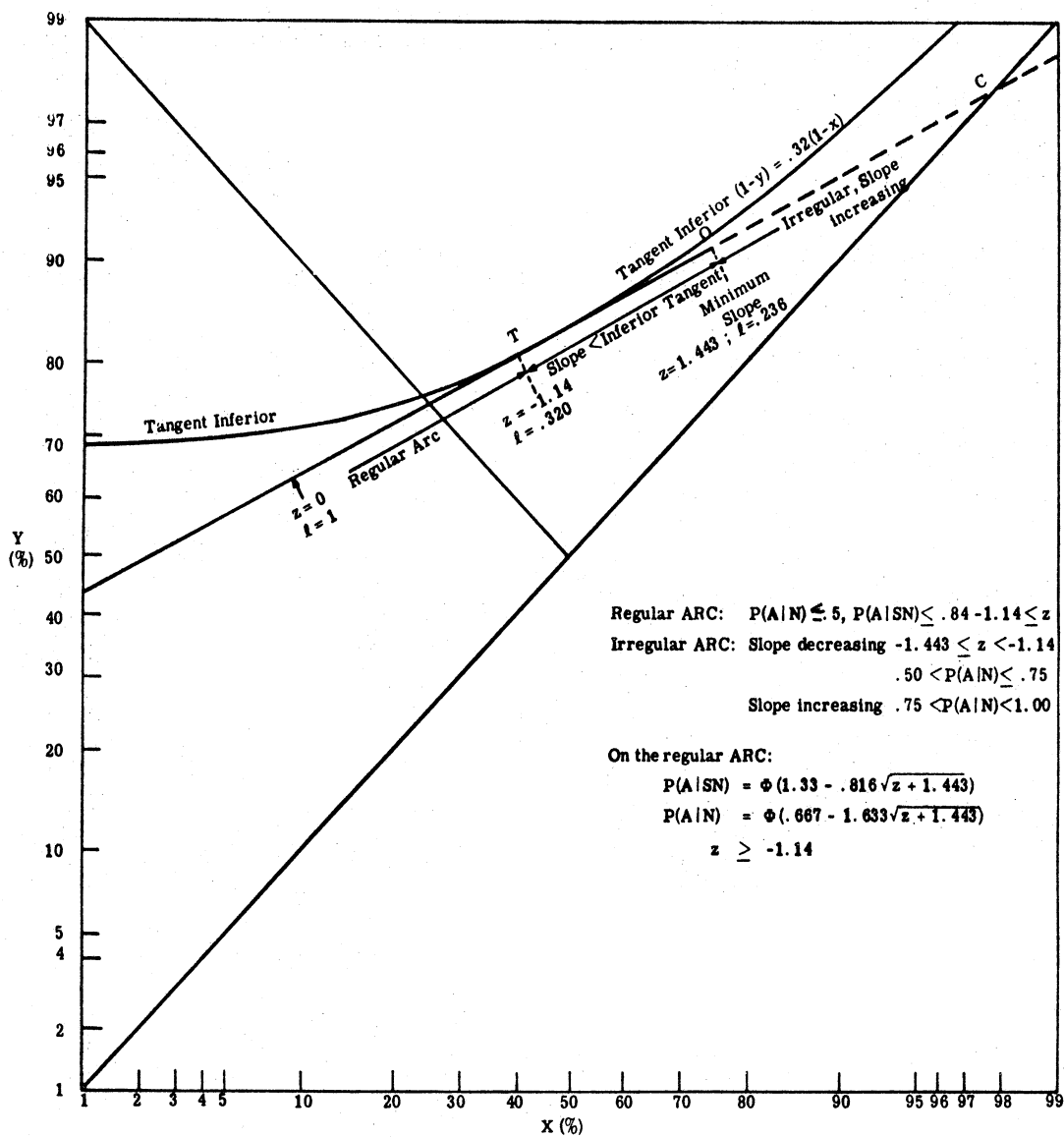
Fig. 7.6. Binormal ROC q = 2 s = .5

point of minimum slope, and the chance diagonal point are indicated.

In the formal development that follows, the restriction that the slope falls between zero and one is made. An alternative restriction would be that the slope is greater than one. The two cases must be treated separately, but the results are similar since the two cases correspond to reflections about the negative diagonal.

The equation for the binormal ROC curve is

$$X = \Phi(-t) \qquad Y = \Phi(s(q - t)) \qquad 0 < s < 1 \qquad (7.3)$$

Since X and Y are one minus the distribution functions for the dummy random variable t, the probability density function for t is obtained by differentiation.

$$f(t|N) = \frac{d(1 - X)}{dt} = \phi(-t) \qquad (7.4)$$

$$f(t|SN) = \frac{d(1 - Y)}{dt} = s\phi[s(q-t)] \qquad (7.5)$$

The function $\phi(\ )$ is the normal probability density function. In order to determine the logarithm of the likelihood ratio of t, divide Eq. 7.5 by Eq. 7.4 and take logarithms.

$$z(t) = \ln s - .5s^2 (q-t)^2 + .5(-t)^2 \qquad (7.6)$$

When (7.6) is expanded in terms of the variable t

$$z(t) = \ln s + .5(1-s^2) t^2 + s^2 qt - .5s^2 q^2 \qquad (7.7)$$

Up to this point no use has been made of the restriction that s falls between zero and one. In fact all of the equations above hold for all values of s greater than zero, including one. Were s equal to one the quadratic term $.5(1-s^2)t^2$ would drop out of the Eq. 7.7, and z would be linear with t. When s is not one, z is quadratic in t. It is basically this quadratic term which causes all the trouble for the binormal ROC; this quadratic will always appear when the curve is binormal and not strictly normal.

Use of the assumption that s falls between zero and one is made in order to factor out the term $1 - s^2$, complete the square, and obtain the final expression for z.

$$z(t) = .5(1 - s^2)\left(t + \frac{s^2 q}{1 - s^2}\right)^2 + \left(\ln s - .5 \frac{s^2}{1 - s^2} q^2\right) \qquad (7.8)$$

The minimum value of the quadratic term is zero. Therefore the minimum value for z is the second term of Eq. 7.8. Call this value z_0.

$$z_0 = \ln s - .5 \frac{s^2}{1 - s^2} q^2 \qquad (7.9)$$

The corresponding value for the dummy variable t shall be called t_0.

$$t_0 = \frac{-s^2}{1 - s^2} q \qquad (7.10)$$

In Figures 7.5 and 7.6 this point on the ROC was indicated with the index zero. The coordinates of the point on the ROC are given by Eq. 7.11.

$$X_0 = \Phi(-t_0) = \Phi\left(\frac{s^2}{1-s^2}\, q\right)$$

$$Y_0 = \Phi[s(q-t_0)] = \Phi\left(\frac{s^1}{1-s^2}\, q\right)$$

(7.11)

From this point of the binomal ROC curve to the point $(1,1)$ the slope will be increasing. The slope of the ROC curve is the likelihood ratio, e^z. In Eq. 7.8, as the quadratic term increases either because t is becoming greater than t_0 or because t is becoming less than t_0, the value of z will increase. Since the slope along the arc from (X_0, Y_0) to $(1,1)$ is increasing it is greater than e^{z_0}. This formal proof of the concave nature of the binormal ROC curve near $(1,1)$ is summarized in Eq. 7.13.

$$X_0 < X < 1 \implies \frac{dY}{dX} > e^{z_0} = \frac{dY}{dX}\bigg|_{X_0}$$

(7.13)

It would be very nice to be able to solve for the point of tangency of the tangent inferior as simply as for the point of minimum slope. In order to do this consider the logarithm of the slope of the

tangent inferior.

$$z(t_T) = \ln \frac{1 - Y_T}{1 - X_T} = \ln \frac{\Phi(s(t_T-q))}{\Phi(t_T)} \qquad (7.14)$$

The value of t_T is the value of t at the point of tangency. An explicit

solution by equating the right-hand sides of Eq. 7.8 and Eq. 7.14 is

not possible. The graphic solution used in Fig. 7.5 when the ROC

is drawn on linear coordinate paper is straight-forward although not

extremely accurate. A similar procedure can be used on normal-

normal paper by drawing a set of curves on that paper which would

be straight lines from the point $(1,1)$ on linear paper. The point of

tangency between the tangent inferior and the binormal can again be

judged rapidly but not extremely accurately by this graphical method.

The only guaranteed accurate method known to the writer is an iterative

solution for t_T, by tables and hand computation, or by computer. The

values obtained in the present paper were obtained using an IBM 7090

computer.

Before proceeding to various methods of rectifying this

irregular ROC curve, the ROC character for the regular part of the arc

is obtained. The equations also hold for the irregular part of the arc,

but the intended application is to the regular part from the origin to

the point of tangency with the tangent inferior.

Equation 7.8 is simplified by using Eqs. 7.9 and 7.10.

$$z = .5(1 - s^2)(t - t_0)^2 + z_0 \tag{7.15}$$

Solve Eq. 7.15 for t

$$t = t_0 + \left[\frac{2}{1 - s^2} (z - z_0) \right]^{.5} \tag{7.16}$$

The equation for the ROC curve, in terms of the logarithm of the likelihood ratio, is therefore

$$\left. \begin{array}{l} X = \Phi\left(\dfrac{s^2 q}{1 - s^2} + \sqrt{2/(1 - s^2)}\sqrt{z - z_0} \right) \\[4mm] Y = \Phi\left(\dfrac{sq}{1 - s^2} + s\sqrt{2/(1 - s^2)}\sqrt{z - z_0} \right) \end{array} \right\} \tag{7.17}$$

The Jacobian of transformation from t to z is the derivative of t with respect to z.

$$\left| \frac{dt}{dz} \right| = \left[2(1 - s^2)(z - z_0) \right]^{-.5} \tag{7.18}$$

The density function for z under condition N is obtained by the substitution of Eq. 7.16 into Eq. 7.4, and then multiplying by Eq. 7.18

$$f(z \mid N) = \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{2(1-s^2)}}$$

$$(z - z_0)^{-.5} e^{-\left[\frac{1}{2} t_0^2 + \frac{2\sqrt{2}}{\sqrt{1-s^2}} t_0 \sqrt{z-z_0} \frac{2}{1-s^2} (z-z_0)\right]}$$

(7.19)

It will be convenient to allow the term t_0 to remain for several more steps. The ROC character is obtained from the N density function by multiplying by $e^{.5z}$. The form used is

$$\pi(z) = e^{.5(z-z_0)} e^{.5z_0} f(z \mid N)$$

(7.20)

Equation 7.9 yields

$$e^{.5z_0} = s^{.5} e^{.25qt_0}$$

(7.21)

This leads to the following form of the ROC character

$$\pi(z) = \frac{1}{2} \left[\frac{s}{\pi(1-s^2)}\right]^{.5} e^{.25qt_0 - .5t_0^2} (z-z_0)^{-.5}$$

(7.22)

$$\cdot e^{-\frac{\sqrt{2}t_0}{\sqrt{1-s^2}} \sqrt{z-z_0}} e^{-\left(\frac{1+s^2}{1-s^2}\right)\left(\frac{z-z_0}{2}\right)}$$

The final step is to eliminate the constant t_0 by using its definition, Eq. 7.10. The final form of the ROC character is

$$\pi(z) = \frac{1}{2} \left[\frac{s}{\pi(1-s^2)} \right]^{.5} e^{-q^2 \left[\frac{s^2(1+s^2)}{4(1-s^2)^2} \right]} (z-z_0)^{-.5}$$

$$e^{q \left[\frac{\sqrt{2}\,s^2}{(1-s^2)^{1.5}} \right] \sqrt{z-z_0}} \; e^{-\left(\frac{1+s^2}{1-s^2}\right)\left(\frac{z-z_0}{2}\right)}$$

(7. 23)

Equation 7. 23 is so complicated with clusters of coefficients that it is difficult to envision its behavior. It is rewritten to focus on those places where the variable $z - z_0$ appears.

$$\pi(z) = A(z-z_0)^{-.5} \; e^{B\sqrt{z-z_0} - C(z-z_0)}$$

(7. 24)

If the quality is positive, as we may assume it to be in most interesting cases, the coefficients A, B and C will be positive. For a very large magnitude z the final exponential term will dominate, and the ROC curve will behave similarly to one with an exponential character, a pure power ROC. If the constant B were zero (it is not), then the character would be that of a Pearson III with $\rho = -.5$. Such a Pearson character is monotone decreasing in z. The term $B\sqrt{z-z_0}$ tends to make the character increase as z increases. This means that the character may exhibit a mode, or may be monotone decreasing in z, depending on the specific values of the parameters.

In this section the ROC character for the regular part of the binormal ROC curve, Eq. 7. 23 was obtained. The equation

for the ROC curve itself in terms of the slope and quality parameters,

and the logarithm of the likelihood ratio z was also determined. This

is Eq. 7.17.

7.3 External Rectification

The process of external rectification consists of replacing

the irregular part of the ROC curve with the tangent inferior. The corres-

ponding action on the ROC character is to limit the applicability of

Eq. 7.23 to those z which fall above z_T, (the logarithm of the

slope of the tangent inferior) and to add a jump in the integrated ROC

character at z_T of appropriate magnitude.

No analytic solution for the tangent inferior has been

found. Therefore a computer program was written that did the following

search. The binormal ROC curve was traced starting at the origin. At

each point along the ROC curve, the value of z was compared with the

(logarithm of) the slope from the point on the ROC to the point $(1,1)$.

Near the origin the line connecting the ROC point with the point $(1,1)$

is a secant, interior to the ROC. As soon as the two slopes agree, the

secant has become a tangent, and the program halts. This is a straight-

forward, albeit brute force, method for determining the tangent inferior

and z_T. This program is listed in Appendix F.

Figure 7.7 presents five ROC characters which were

plotted as part of one run of this program. For this run, the slope

parameter s was fixed at 0.8. The qualities examined were q = 1,2,3,4,5.
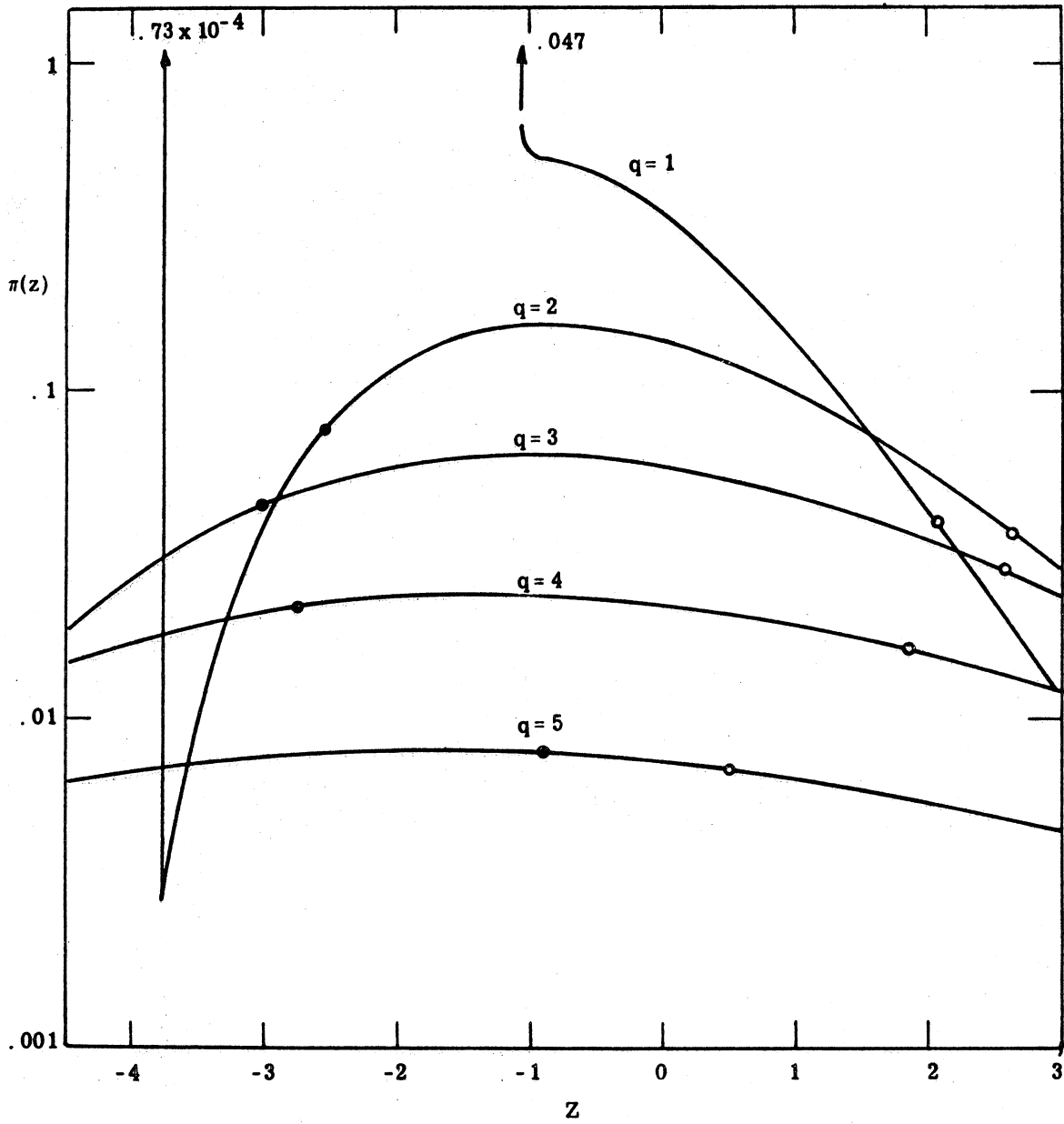
Fig. 7.7. Five binormal characters,
external rectification

In Fig. 7.7 the points on the ROC character with positive z marked with a large open dot indicate the point at which the ROC curve passed through .01 false alarm probability. The points marked with the solid dot indicate the points at which the ROC curve passed through the .99 detection probability. When the quality parameter is one, the ROC character is strictly monotone decreasing. In the neighborhood of z_0 the ROC character is very steep. The other four ROC characters present a more normal appearance. Each is terminated at z_T with an ω_π term. For $q = 2$ and greater, z_T was sufficiently to the right of z_0 to prevent the appearance of a sharp rise in the ROC character due to the term $(z-z_0)^{-.5}$.

The corresponding ROC curves are plotted on normal-normal paper in Fig. 7.8. The effect of the external rectification is evident only in the $q = 1$ curve. As was seen on the ROC character, all of the other curves contain the one percent to 99 percent region in their regular portion. Although it is important in theory to make sure that the ROC curves are regular, it may very well turn out in practice that this careful consideration has been unnecessary. In order to exhibit binormal ROC curves which evidenced need for rectification on the usual plot on normal-normal paper, a number of cases with slope equal to one-half were considered. These are presented in Fig. 7.9. The points where these would cross the chance diagonal are marked with C and the points where these would take on their minimum slope
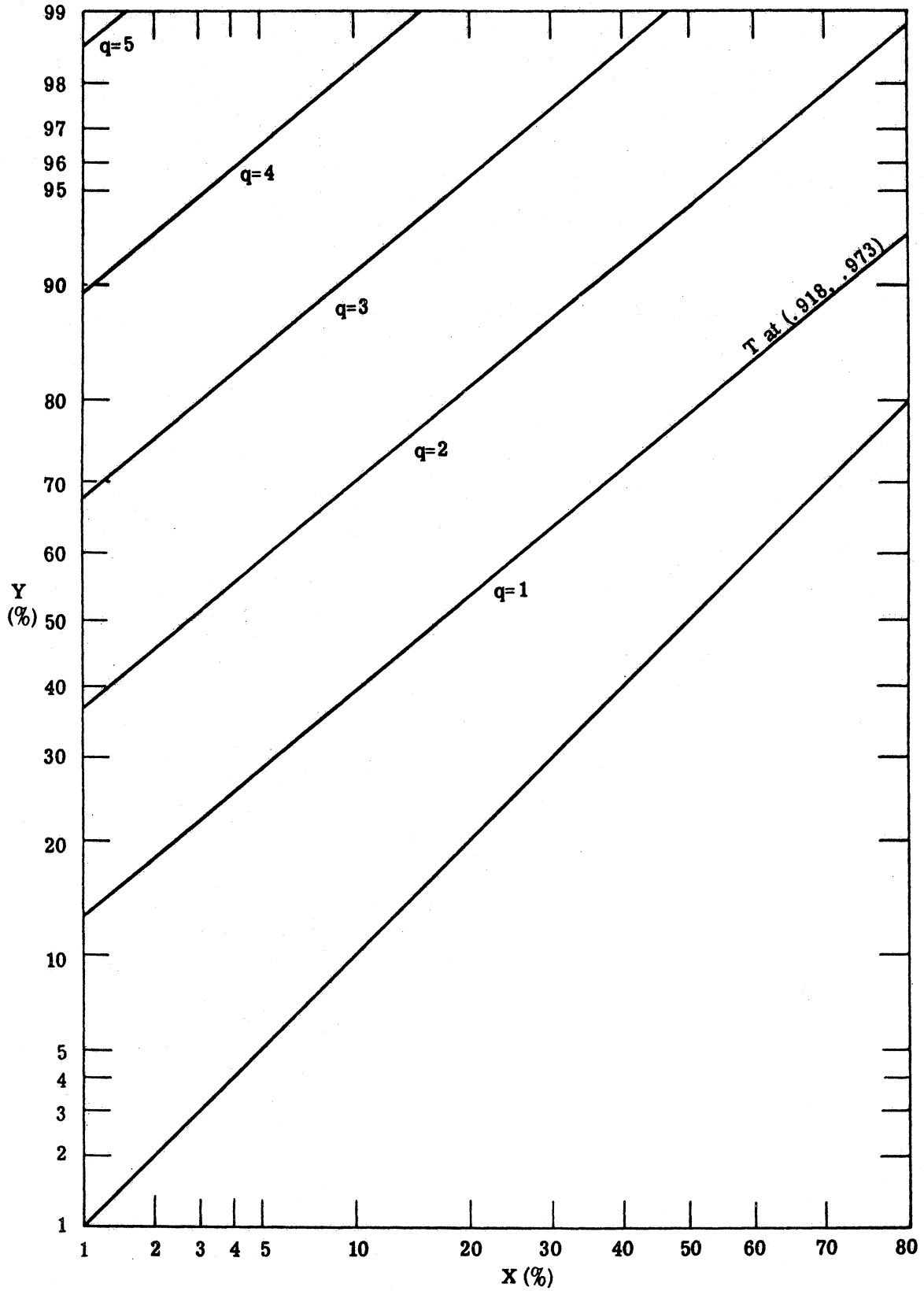
Fig. 7.8. Binormal ROC with slope s = .8 q = 1
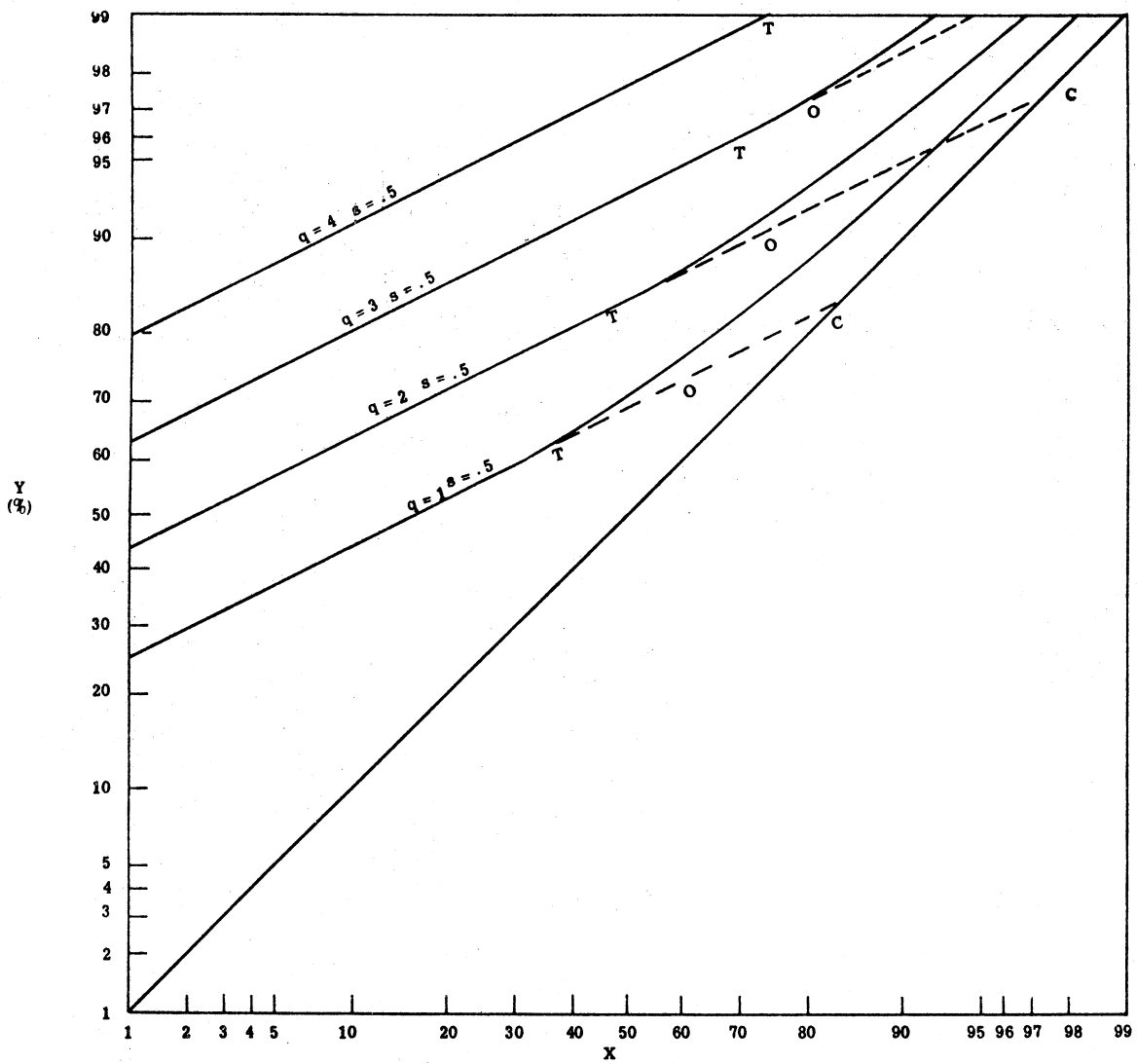
Fig. 7.9. Externally rectified binormal ROC curves

e^{z_0} are marked with 0. The point at which the tangent inferior is adjoined to the curve to make it regular is marked with T.

The curves shown in Fig. 7.9 are regular ROC curves, derived from binormal ROC curves by external rectification. The corresponding integrated character has a jump at the value z_T, and is differentiable from z_T to infinity.

7.4 Internal Rectification

The binormal ROC curves are irregular because they are based on a decision axis, t , which is not monotone with its own likelihood ratio. External rectification removed part of the axis from consideration in order to maintain a monotone relationship between the remaining range of t and the likelihood ratio. A decision mechanism that allows access to the t axis can undergo a superior rectification to the simple external rectification. Without access to the decision axis there is no other choice but to stay with external rectification.

Consider a decision device with a decision axis, t, which led to a specific binormal ROC curve with quality q and slope s between zero and one. The probability density functions for this decision axis are both normal, and are given in Eq. 7.4 and 7.5. Since the logarithm of the likelihood ratio is quadratic in t, then the response "A" corresponding to z being greater than some level will be equivalent to a two-tailed test on the t axis.

The probability density function for z under both of the causes, N and SN, is based on the sum of the probability densities for each of two appropriate t values (the solution of 7.15). The ROC character may be determined as the probability density for z under one of the causes, multiplied by either $e^{\pm .5z}$.

Another approach is to find the root likelihood product, and to multiply this by the Jacobian of transformation, dt/dz.

The root likelihood product, using Eq. 7.3 and 7.4, and taking the square root, is

$$[\, f(t\,|\,N)\, f(t\,|\,SN)]^{.5} = [\, \frac{s}{2\pi}\,]^{.5}\, e^{\,-\,\frac{1}{4}\,[\,t^2\,+\,s^2\,(q-t)^2\,]} \tag{7.25}$$

In the previous section the relation between z and t was derived

$$t = t_0 \pm \left[\,\frac{2}{1-s^2}\,(z-z_0)\right]^{.5} \tag{7.26}$$

This differs from the previous equation, Eq. 7.16, in the range of application. Previously t had been limited to those values t greater than t_0, in fact, those values of t greater than t_T. However, direct use of the likelihood ratio allows use of the entire t axis. A transformation of variables effects the probability density functions by the absolute value of the Jacobian of transformation and not the signed value. Therefore, Eq. 7.18 holds independently of whether the plus or the minus sign in Eq. 7.26 is used.

$$\left|\frac{dt}{dz}\right| = \left[2(1-s^2)(z-z_0)\right]^{-.5} \tag{7.18}$$

The only remaining step is to substitute t into the expression in the exponent of Eq. 7.25 and do some careful algebra. Expand the exponent of Eq. 7.25 and use Eq. 7.26 to obtain

$$t^2 + s^2(q-t)^2 = \left[2\frac{1+s^2}{1-s^2}\right](z-z_0) \mp \left[\frac{4\sqrt{2}s^2 q}{(1-s^2)^{1.5}}\right](z-z_0)^{.5} \tag{7.27}$$

$$+ \left[\frac{s^2 q^2(1+s^2)}{(1-s^2)^2}\right]$$

The ROC character is obtained by adding together the two values for the root likelihood product that correspond to the same value of z, and multiply the result by Eq. 7.18. Two of the factors in the exponent, Eq. 7.27, are common. That is, two of the terms in Eq. 7.27 are independent of whether the plus sign or the minus sign in Eq. 7.26 is used. The Jacobian of transformation is also independent of which sign has been used. Therefore, the only effect of the summation is on the middle term in Eq. 7.27. Since this middle term appears in an exponent, the adding of the two terms will yield a hyperbolic cosine term. The result of the transformation is shown below.

$$\pi(z) = \left[\frac{s}{\pi(1-s^2)}\right]^{.5} e^{-\frac{s^2 q^2 (1+s^2)}{4(1-s^2)}} (z-z_0)^{-.5} e^{-\left[\frac{1+s^2}{1-s^2}\right]\frac{z-z_0}{2}}$$

$$\cdot \cosh\left[\frac{\sqrt{2} s^2 q}{(1-s^2)^{1.5}} (z-z_0)^{.5}\right] \tag{7.28}$$

$$z > z_0 = \ln s - \frac{s^2 q^2}{2(1-s^2)}$$

The main change between Eq. 7.28 and Eq. 7.23 appears to be that the single exponential term involving a square root of z has become a hyperbolic cosine term involving the square root of z. Another basic difference which is not so evident is that Eq. 7.28 is valid for z right down to z_0, whereas Eq. 7.23 is valid only for z as low as z_T, which is greater than z_0. This means that the minimum slope of the internally rectified binormal ROC curve will be the same as the minimum slope of the original irregular binormal ROC curve. The integrated ROC character for the internally rectified case contains no jump, that is, there is no ω_π term at the lower end of the ROC character. However, there is an action in the ROC character which behaves very similar to this "delta-function." Since z is no longer bounded away from z_0, in the immediate neighborhood of z_0, the ROC character will be dominated by the term $(z-z_0)^{-.5}$. As z gets

smaller and smaller the character will suddenly become strongly

influenced by this term and will approach infinity.

$$z \rightarrow z_0 \ , \ \pi(z) \rightarrow \left[\frac{s}{\pi(1-s^2)} \right]^{.5} e^{- \frac{s^2 q^2 (1+s^2)}{4(1-s^2)^2}} (z - z_0)^{-.5} \rightarrow \infty$$

$$(7.29)$$

An example of this behavior in the neighborhood of

z_0 is shown in Fig. 7.10, a plot of the logarithm of the ROC charac-

ter for the case $q = 1$, $s = .8$. The ROC character for the exter-

nally rectified case was shown in Fig. 7.7. The characters for both

the internal and external rectification appear very similar. The only

difference is that the ROC character for the internally rectified case

is continuous and has no ω_π term adjoined to it. However, the

appearance is very much the same for both cases.

A detailed comparison of the behavior near z_0 for

internally and externally rectified ROC characters is shown in

Fig. 7.11. The case considered is $q = 2$, $s = .8$. The value

of z_T is about .01 greater than the value z_0 . For the internally

rectified case the ROC character is monotone decreasing as it

approaches z_T from higher values, the then $\omega_\pi(z_T)$ is abruptly

adjoined. For the internally rectified case the ROC character de-

creases, passes through a minimum, and rises abruptly to infinity

as it approaches z_0 from above.
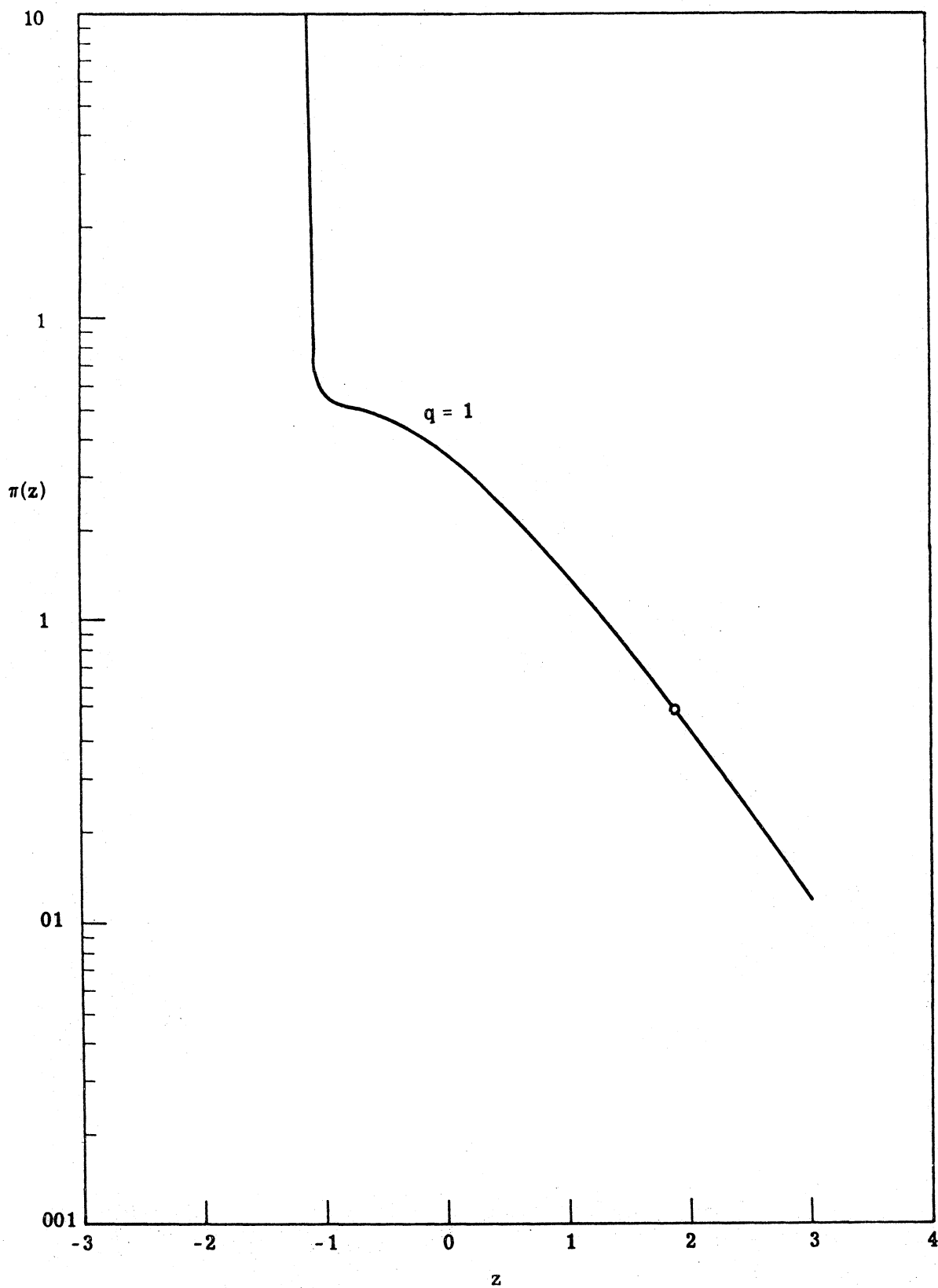
Fig. 7.10. ROC character for internally
rectified binormal, q=1, s=.8
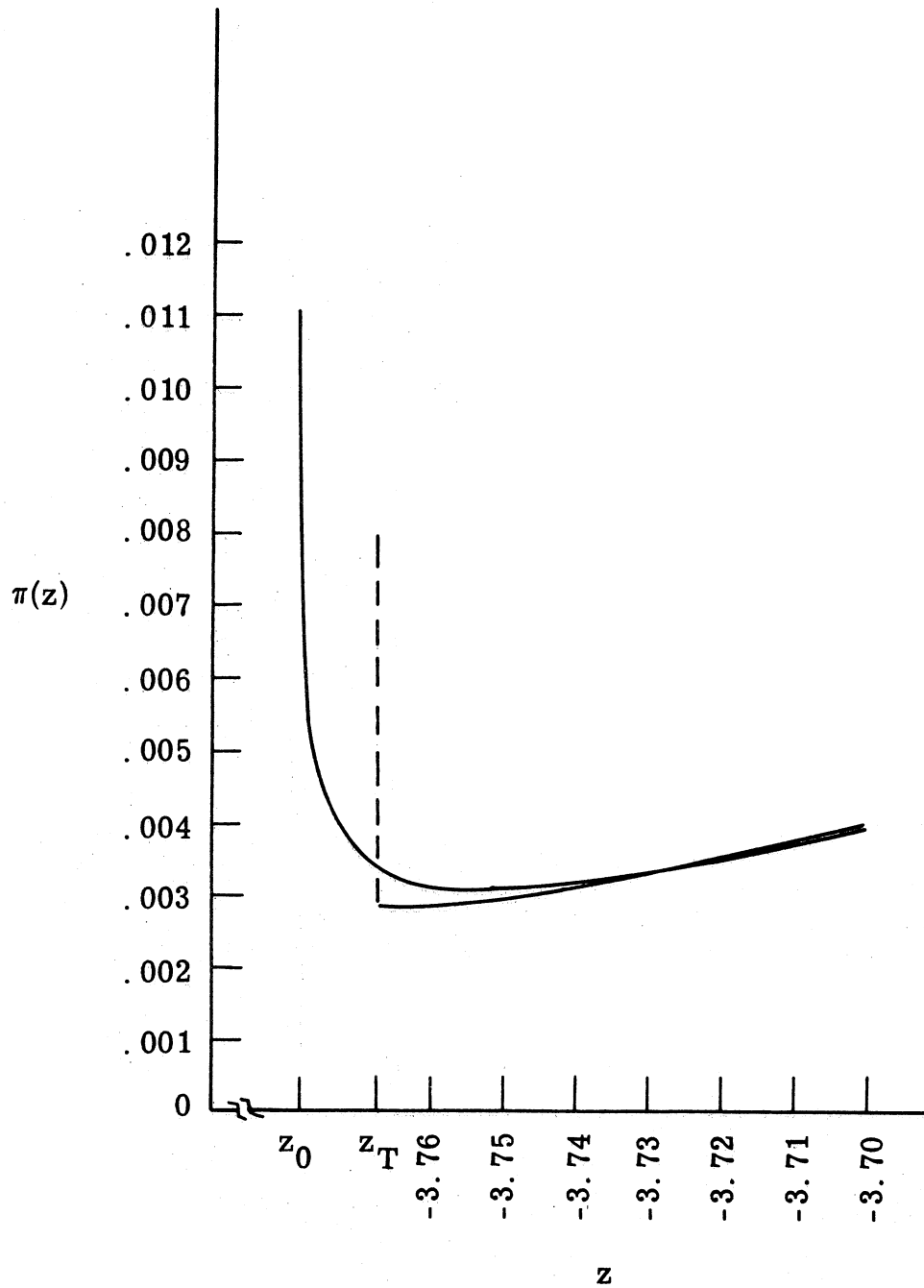
Fig. 7.11 Detail of ROC character,
q=2, s=.8
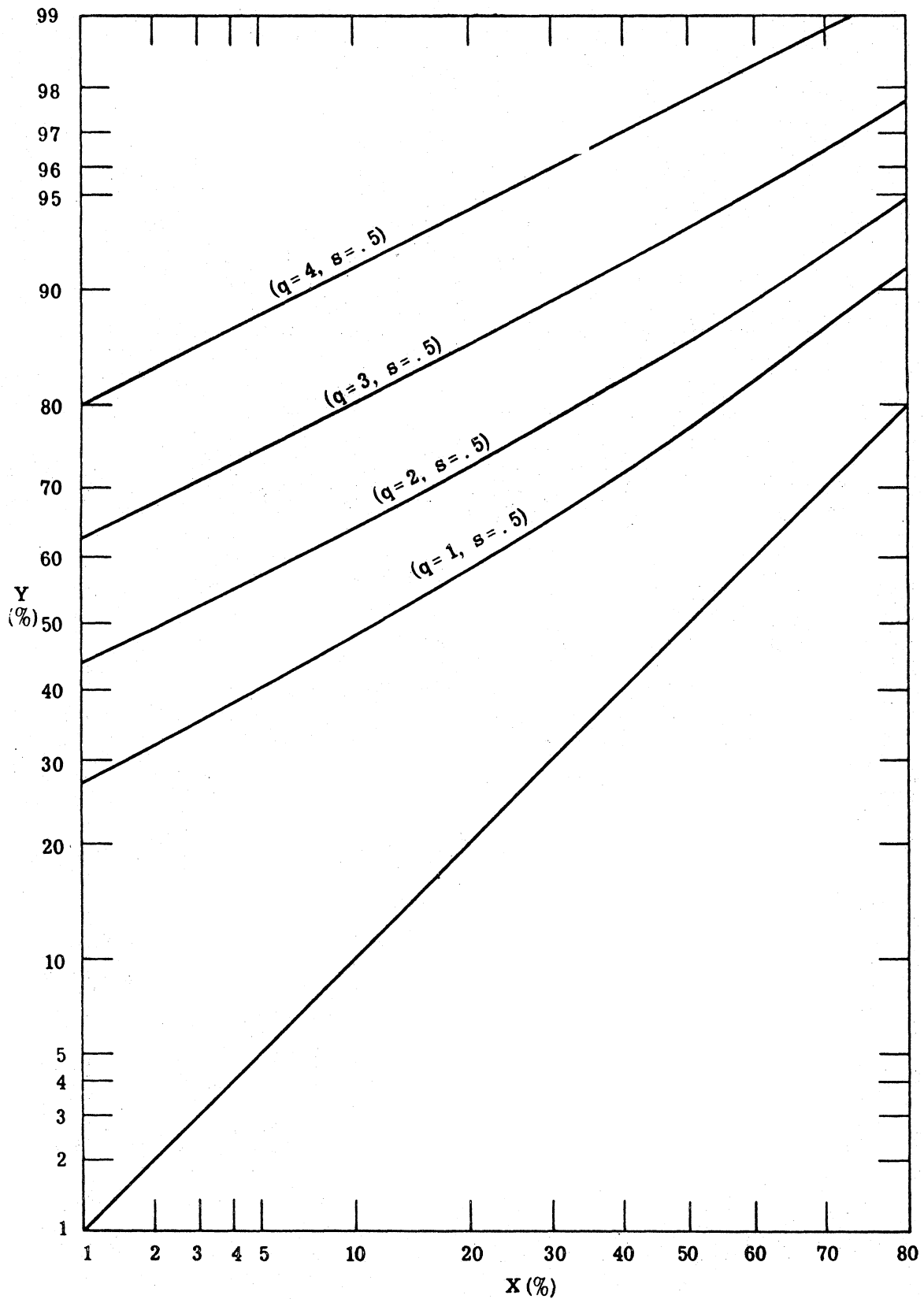
Fig. 7.12. Binormal ROC curves,
internally rectified

The computer program that accomplishes internal rectification is also in Appendix F. The case s = .5 that was analyzed for external rectification was also analyzed for internal rectified ROC curves. Figure 7.12 shows the results of internal rectification when the original quality was q = 1, 2, 3, and 4. These may be compared with the externally rectified solution given in Fig. 7.9. Curves for internal rectification are smoother, and somewhat better than for the externally rectified case. The appearance of the q = 4 curve is virtually unchanged, however, the additional increases being in the third or fourth decimal place.

7.5 Discussion

This chapter considered binormal ROC curves, which plot as straight lines at slope less than one on normal-normal paper. These are irregular ROC curves, but we can associate with any binormal curve a regular ROC curve by either external or internal rectification. Such rectification changes the shape of the curve, and changes the ROC character.

A very important question is the actual magnitude of the changes in the ROC curve, especially if one is interested in a limited region of the ROC curve. In the first section, ROC curves were taken from some work of J. A. Swets. For the family of curves used, the slope is related to the quality by the equation

$$s = \frac{1}{1 + .25q} \qquad (7.30)$$

It can be shown that all of the ROC curves in this family pass through

a common point on the chance diagonal. This is the point (.999968,

.999968). The data (Fig. 7.2, 7.3) fell in the medium percentage

levels, in fact, to the left of the negative diagonal, and were quite

distant from probabilities such as .999968. The program developed

for this work internally rectified six curves from this family. The

result is shown on extended normal-normal paper in Fig. 7.13.

The resulting rectified ROC curves do not cross the chance diagonal,

and do not cross each other. These curves and their corrections on

the usual range of normal-normal paper is Fig. 7.14. In the region

above .97 detection the rectified ROC curve is slightly better than

the original binormal curve. One can certainly conclude that Swets

is in no trouble from having chosen to operate with a nonregular

ROC curve. The rectification is formally called for if a rigorous

treatment is desired for the ROC curves. As a descriptive model

of empirical data, a way of fitting and describing data, the curves

are perfectly valid. In fact, one may undoubtedly manipulate the

model mathematically so long as the operations stay concentrated

on the region far from z_0.

Binormal ROC curves have arisen in detection problems

with slopes as drastic as $s = .5$. When such curves are to be used

in a region anywhere close to the chance diagonal, one or the other

type of rectification must be considered to render the ROC curves

regular. The final figure of this section, Fig. 7.15, shows two
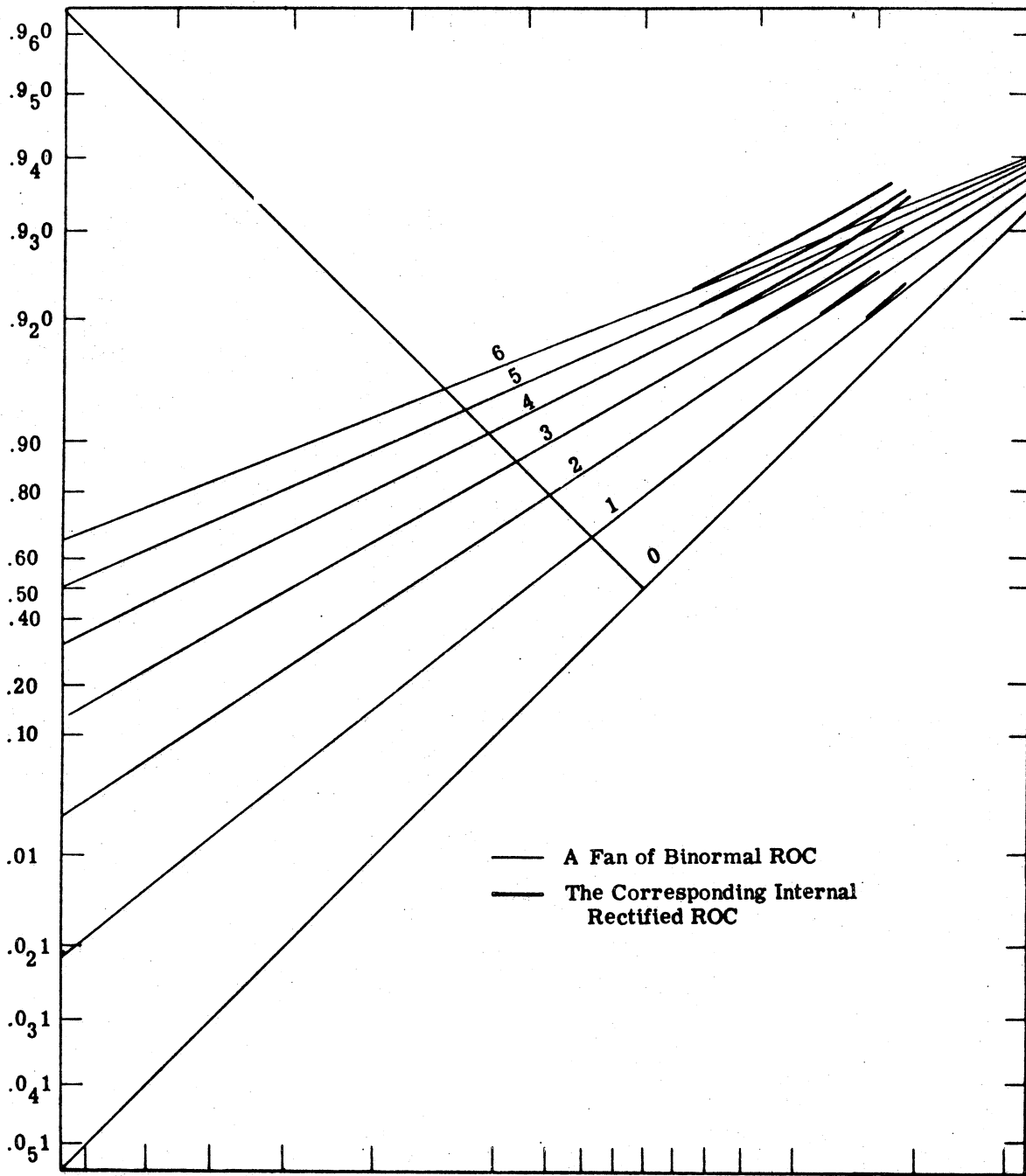
Fig. 7.13. Example of internal rectification
of binormal ROC curves
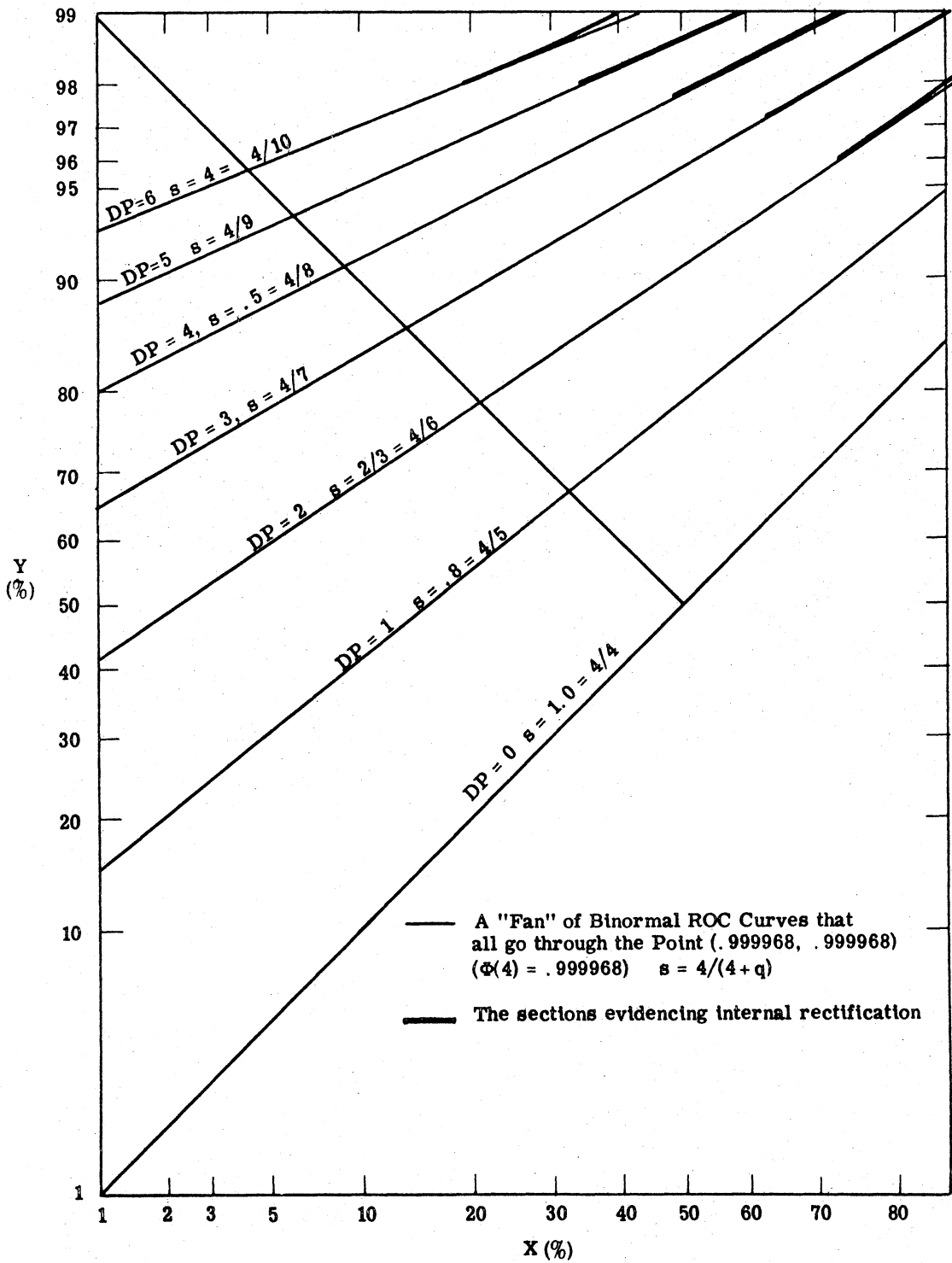
Fig. 7.14. A fan of binormal ROC curves
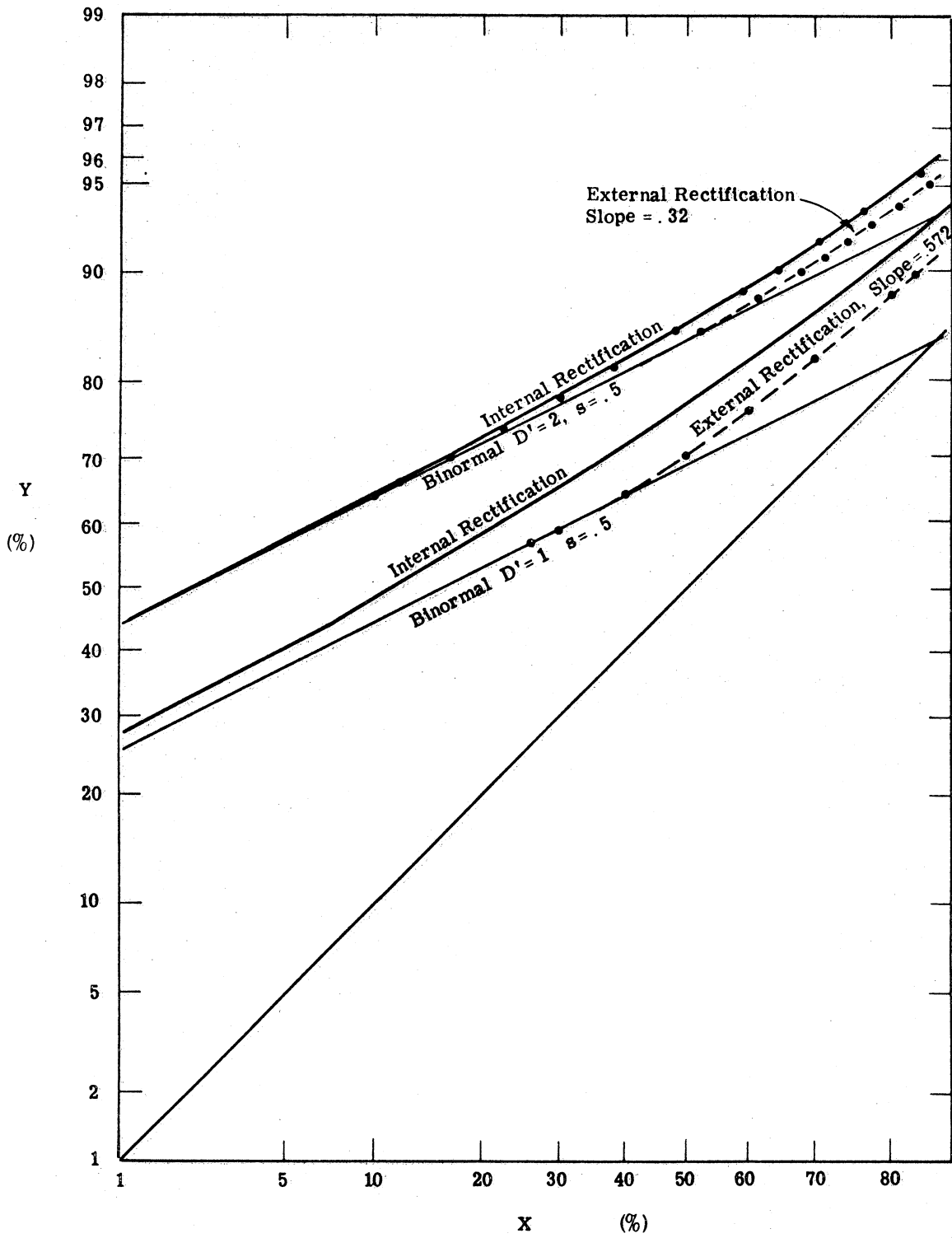
Fig. 7.15. Binormal ROC, s = .5 showing
two types of rectification

binormal ROC curves with slope equal to .5. Both the internal

and external rectification have been shown on the same figure.

If one is dealing with slopes as drastic as this, he may very well

wish to incorporate rectification into his model.

CHAPTER VIII

ROC FOR MULTIPLE OBSERVATIONS

8.1 Convolution Theorem

The distribution for the sum of a number of independent random variables can be determined by convolution of the distributions for the individual random variables. In this section, the convolution theorem for the ROC characters is developed for characters without jump functions. The convolution concept is important in detection situations in which the observation upon which a single decision is based consists of several independent parts. The logarithm of the likelihood ratio for such an observation is the sum of the logarithm of the likelihood ratios of the independent parts. Let z_i denote the logarithm of the likelihood ratio for the i-th part of the observation, and $\pi_i(z_i)$ denote its ROC character. Let z be the logarithm of the likelihood ratio for the complete observation.

$$z = \sum_{i=1}^{n} z_i \qquad (8.1)$$

Theorem: If $\pi_i(z_i)$ are the ROC characters for statistically independent observations, then

$$\pi(z) = \pi_1(z_1) * \pi_2(z_2) * \ldots * \pi_n(z_n) \qquad (8.2)$$

where the star indicates convolution.

Proof: The proof is inductive, and depends on the convolution theorem for probability density functions.

Initial Step: $n = 2$ $\qquad z = z_1 + z_2$

$$f(z \mid SN) = \int f_1(z_1 \mid SN) \, f_2(z_2 = z - z_1 \mid SN) \, dz_1$$

$$= \int e^{.5z_1} \, \pi_1(z_1) \, e^{.5(z - z_1)} \, \pi_2(z - z_1) \, dz_1$$

$$= e^{.5z} \int \pi_1(z_1) \, \pi_2(z - z_1) \, dz_1$$

$$\pi(z) = e^{-.5z} \, f(z \mid SN)$$

$$\therefore \pi(z) = \int \pi_1(z_1) \, \pi_2(z - z_1) \, dz$$

Inductive Step: Assume that the theorem is true for $n = r \geq 2$.

$$z = \sum_{i=1} z_i$$

Let

$$Z = \sum_{i=1}^{r} z_i$$

$$z = Z + z_{r+1}$$

Proof (Cont.)

by the initial step, and the inductive hypothesis

$$\pi(z) = [\pi_1(z_1) * \pi_2(z_2) * \ldots \pi_r(z_r)] * \pi_{r+1}(z_{r+1})$$

Convolution is associative, so we may omit the [ ].

Q. E. D.

The convolution of the classical distributions are well known. Brief comments about the normal, the rectangular, and the exponential, in order to call attention to certain specific ideas.

The convolution of a finite number of normal density functions is again normal, with the mean being the sum of the means and the variance, the sum of the variances. The normal character is proportional to a zero mean normal density, with variance d. Therefore, the convolution of a number of normal characters will yield a normal character, symmetric about zero, and with detection index d being the sum of the detection indices for the individual z_i.

$$d = \sum d_i \tag{8.3}$$

The rectangular ROC character is

$$\pi_i(z_i) = \frac{\sqrt{1+R}}{2R} \quad , \left| z_i \right| < \ln(1 + R) \tag{8.4}$$

The convolution of two identical such rectangular characters is given in Eq. 8.5

$$\pi(z = z_1 + z_2) = \frac{1 + R}{4R^2} \left| z - 2 \ln (1 + R) \right|$$

$$|z| < 2 \ln (1 + R) \tag{8.5}$$

This is a triangular ROC character. The convolution of n identical rectangular characters will be expressed in n-th degree polynomials in z. The expression will consist of $n + 1$ sections along the z axis, and become very cumbersome. The result has a central tendency to the normal character. It would be useful to have a measure of the normal parameter, d, which could be used as the number of individual rectangular characters is large. The usual procedure is to equate variances. In order to do this, we would have to re-normalize the ROC character to have unit area. An equivalent calculation for symmetric characters is to define a quantity such as

$$\hat{d} = \frac{\int z^2 \, \pi(z) \, dz}{\int \pi(z) \, dz} \tag{8.6}$$

For the rectangular character of Eq. 8.4

$$\hat{d}_1 = \frac{1}{3} \left(\ln (1 + R) \right)^2 \tag{8.7}$$

Eq. 8.6 evaluated for the character resulting from n convolutions of similar ROC characters will be n times as large as for the single

ROC character.

$$\hat{d}_n = \frac{n}{3} \left(\ln(1 + R) \right)^2 \qquad (8.8)$$

The gamma, Pearson III, and chi-square distributions all correspond to the same ROC character. It is well-known that the sum of chi-square random variables has a chi-square distribution, and similarly for the other distributions. These, in turn, are strongly related to the simple exponential. Consider the simple exponential ROC character given in Eq. 8.9.

$$\pi_i(z_i) = \frac{A^{.5}}{A - 1} e^{-\frac{A + 1}{A - 1} \frac{z_i - z_0}{2}} \qquad z_i > z_0 = -\ln A \qquad (8.9)$$

If n such characters, all with the same index A, are convolved, the result is the Pearson III character with the same index.

$$\pi(z) = \frac{1}{\Gamma(n)} \left[\frac{A^{.5}}{A - 1} \right]^n (z - n z_0)^{n - 1} e^{-\frac{A + 1}{A - 1} \frac{z - n z_0}{2}} \qquad z > n z_0$$

$$(8.10)$$

This result does demand that the indices A be identical for all of the individual characters.

8.2 Discrete Characters

The proof used in 8.1 for the convolution of ROC characters was given in terms of probability density function. A similar

result holds for discrete ROC characters. It will be assumed without proof. The simplest discrete character is the Luce character with only two points. The ROC curve, such as shown in Fig. 8.1(a), consists of two straight lines meeting at the point (x_1, y_1), also marked as the vertex $V_{1,0}$. In the general binomial ROC curve, let the vertices be marked $V_{n,k}$ indicating that n observations were made, and k of these or fewer were indicative that the condition was N. As with any polygon, it is easiest to write out the total polygon by naming the vertices. The Luce ROC curve is specified by

$$(0,0) \qquad V_{1,0} = (x_1, y_1) \qquad V_{1,1} = (1,1) \qquad (8.11)$$

If two independent observations are made, and each has the same Luce ROC character, then three likelihood ratios are possible. Because of the statistical independence and similarity of the observation, order is unimportant. The ROC curve resulting from Fig. 8.1(a) is shown in Fig. 8.1(b). This is a Green double threshold ROC. What has previously been referred to as the Green ROC is the limiting form of the Green double threshold ROC, where the vertices $V_{2,0}$ and $V_{2,1}$ lie arbitrarily close to the edges of the graph. The vertex $V_{2,0}$ corresponds to the response "A" made whenever neither observation was indicative of condition N. The vertex $V_{2,1}$ corresponds to the response "A" whenever one or fewer observations were indicative of condition N. The description of this ROC by its vertices, together with the relation of these vertices to the Luce vertex is given in Eq. 8.12.
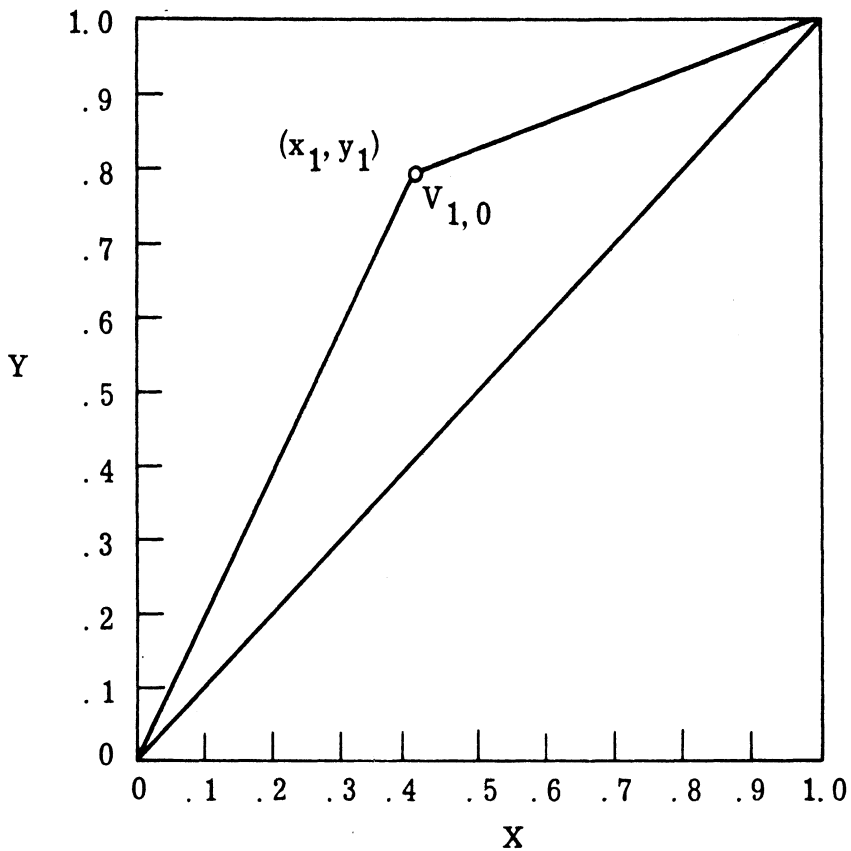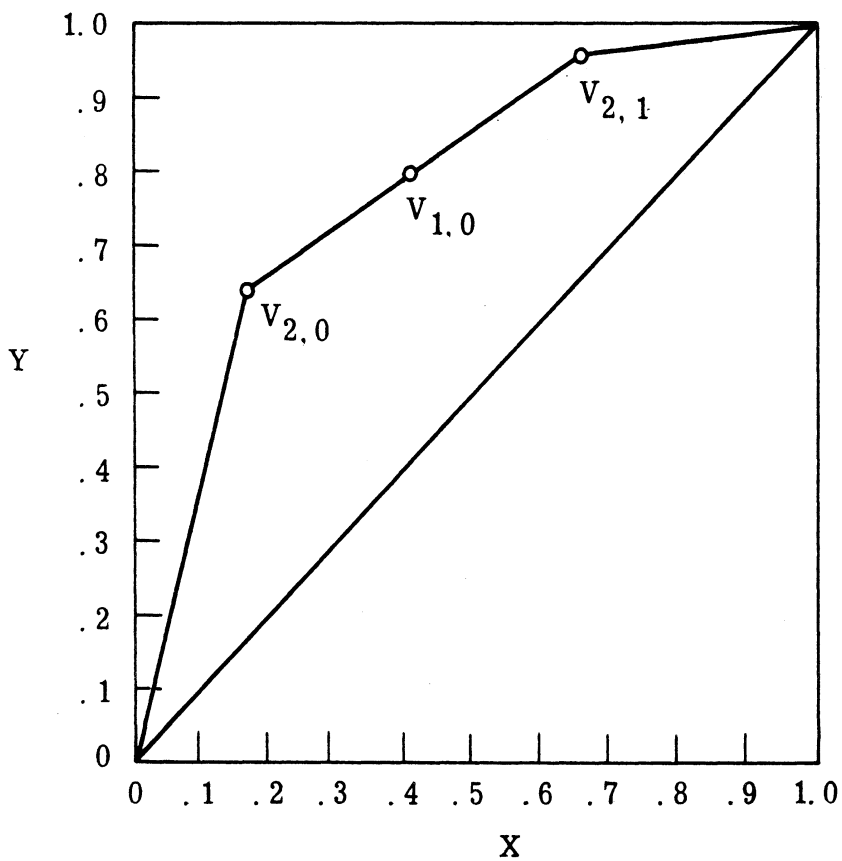
Fig. 8. 1.   (a) A Luce ROC, (b) a Green double threshold ROC

$$(0,0) \qquad V_{2,0} = (x_1^2, y_1^2) \qquad V_{2,1} = (2x_1 - x_1^2, 2y_1 - y_1^2)$$

$$V_{2,2} = (1,1) \tag{8.12}$$

In the particular case graphed in Fig. 8.1, the Luce vertex $V_{1,0}$ falls on the Green double threshold ROC curve. It falls exactly in the middle of the line segment between the two middle vertices. This will always happen. The notation that half of one vertex plus half of another vertex means the point lying halfway between the two vertices indicates both the geometric relation, and the algebra that must be performed on the coordinates to obtain this geometric relation. Equations 8.12 for the Green double threshold ROC and 8.11 for the Luce ROC can be combined to obtain:

$$\frac{1}{2} V_{2,0} + \frac{1}{2} V_{2,1} = (x_1, y_1) = V_{1,0} \tag{8.13}$$

If three independent observations are taken, and each has the same Luce ROC, the resultant ROC will have three nonchance vertices. These are listed below

$$(0,0) \qquad V_{3,0} \qquad V_{3,1} \qquad V_{3,2} \qquad V_{3,3} = (1,1) \tag{8.14}$$

and the equations for the nonchance vertices in terms of the Luce vertex are

$$V_{3,0} = (x_1{}^3, y_1{}^3)$$

$$V_{3,1} = (3x_1{}^2 - 2x_1{}^3, 3y_1{}^2 - 2y_1{}^3)$$

$$V_{3,2} = (3x_1 - 3x_1{}^2 + x_1{}^3, 3y_1 - 3y_1{}^2 + y_1{}^3) \qquad (8.15)$$

Examination of Eq. 8.15 and the equations for the Green double threshold ROC vertices, Eq. 8.12, reveals that the Green double threshold vertices lie along the segments of the three observation ROC.

$$\frac{2}{3} V_{3,0} + \frac{1}{3} V_{3,1} = (x_1{}^2, y_1{}^2) = V_{2,0} \qquad (8.16)$$

$$\frac{1}{3} V_{3,1} + \frac{2}{3} V_{3,2} = (2x_1 - x_1{}^2, 2y_1 - y_1{}^2) = V_{3,0} \qquad (8.17)$$

It appears that each successive ROC curve, due to one additional observation, is better than the previous ROC, but not strictly better. Each vertex of the prior ROC touches a segment on the new ROC. This is indeed the case. The theorem will be stated in terms of the vertex notation; however, the proof itself will deal with the coordinates of the vertex. The meaning of the vertex notation is written out explicitly

$$V_{n,k} = (X_{n,k}, Y_{n,k}) \longleftrightarrow \left\{ n \text{ observation}, \leq k \text{ "B"} \right\} \qquad (8.18)$$

Theorem:

$$(1 - \frac{k+1}{n}) V_{n,k} + (\frac{k+1}{n}) V_{n,k+1} = V_{n-1,k} \qquad 0 < k < n$$

Proof:

Assume that n observations have been taken, and that each

has a Luce ROC character associated with the Luce point

(x_1, y_1) . The k-th ROC point corresponds to responding "A"

whenever there have been k or less observations indicative

of condition N . The probability of a false alarm is there-

fore given by a partial sum over the binomial distribution

$$X_{n,k} = \sum_{j=0}^{k} C_j^{n} (1 - x_1)^j x_1^{n-j} \qquad (8.19)$$

The probability of a detection is given by a similar sum

$$Y_{n,k} = \sum_{j=0}^{k} C_j^{n} (1 - y_1)^j y_1^{n-j} \qquad (8.20)$$

The only difference between Eqs. 8.19 and 8.20 is the use of

x_1 for the condition N and y_1 for the condition SN . The

following manipulations will deal with the N condition. Simi-

lar results for the SN condition may be obtained by inter-

changing x_1 and y_1 . Most of the specific manipulations

have been placed in Appendix G and only the key results dis-

played here. Equation 8.19 can be expanded and rewritten

as a polynomial in x_1 .

<u>Proof</u> (Cont.)

$$X_{n,k} = (-1)^k (n - k) C_k^n \sum_{m=0}^{k} (-1)^m \frac{1}{n - m} C_m^k x_1^{n-m}$$

$$(8, 21)$$

If the "A" decision is relaxed slightly to admit $k + 1$ indications of conditions N, the additional probability of false alarm is a probability of exactly $k + 1$ indications of condition N.

$$X_{n,k+1} - X_{n,k} = C_{k+1}^n (1 - x_1)^{k+1} x_1^{n-k-1}$$

$$(8.22)$$

This can also be written in terms of a polynomial in x_1.

$$X_{n,k+1} - X_{n,k} = (-1)^{k+1} (n - k) C_k^n \sum_{m=0}^{k} (-1)^m \frac{1}{n-m}$$

$$\cdot C_m x_1^n + C_{k+1}^n x_1^{n-k-1}$$

$$(8.23)$$

The exact form in Eq. 8.23 has been chosen to make it as similar as possible to Eq. 8.21. In order to establish the theorem, the highest power of x_1 in Eq. 8.21 must be eliminated by canceling it with the highest power in Eq. 8.23. This highest power corresponds to the summation index $m = 0$.

Proof (Cont.)

This dictates that we try adding $(k+1)/n$ of Eq. 8.23 to Eq. 8.21.

$$X_{n,k} + (\frac{k+1}{n}) (X_{n,k+1} - X_{n,k})$$
$$= (-1)^k (n-k-1) C_k^{n-1} \sum_{m=0}^{k} (-1)^m \frac{1}{n-1-m} C_m^k x_1^{n-1-m}$$

(8.24)

The right-hand side of Eq. 8.24 has the form of Eq. 8.21 when n is replaced by $n-1$, so the X part of the vertex equation is proven.

Since the manipulation on $Y_{n,k}$ is identical to those for $X_{n,k}$, the theorem has been established.

One additional fact about the binomial ROC curves that develop from repeated Luce observation may be noted. Pass a pure power ROC through the Luce vertex. The parameter value for this pure power ROC is

$$x_1 = y_1^A \Rightarrow A = (\ln x_1)/(\ln y_1)$$

(8.25)

The first vertex above the point (0,0) on each of the binomial ROC curves corresponds to saying "A" if and only if there have been no indications of condition N. The coordinates of this vertex are

therefore the n-th power of the coordinates for the Luce vertex.

$$V_{n,0} = (x_1^n, y_1^n)$$ (8.26)

Therefore, this vertex also falls on the pure power ROC curve with index A.

$$X_{n,0} = Y_{n,0}^A$$ (8.27)

Consider an observer whose complete one-observation ROC curve is a pure power ROC. If this observer records decisions made on a number of independent observations, and forgets everything else about the observation, then his individual decisions act as if they came from a Luce ROC. If he then uses these several decisions to reach a terminal decision, his terminal decision will be no better than one he could obtain from a single observation, if his final criterion demands confirmation of the condition SN on each observation. In order to use the multiple decisions to reach a better terminal decision, he must employ not-too-lax a criterion for the individual decisions, and not-too-strict a criterion for the terminal decision. This behavior of the initial point on the poly-gon ROC curves is shown in Fig. 8.2 for the same Luce and Green double threshold ROC used in Fig. 8.1. The first part of the polygon ROC falls below the pure power ROC, while the remainder falls above the pure power ROC.

The work above has been very detailed, and is the type of analysis one would use when the number of observations leading to a
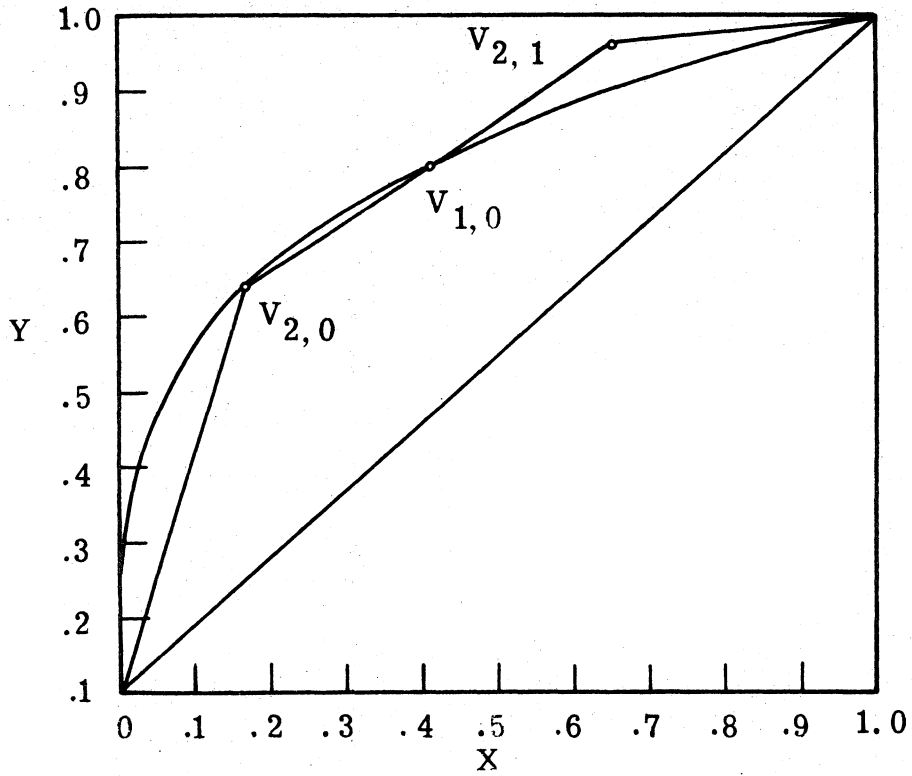
Fig. 8.2 Initial point power law

terminal decision is very small. When the number of observations

leading to a terminal decision is large, a description of a continuous

ROC character that approximates the binomial ROC character would be

convenient. The two possible log likelihood ratios that occur in the

Luce ROC character, and the corresponding jumps in the integrated

ROC character are sketched in Fig. 8.3.
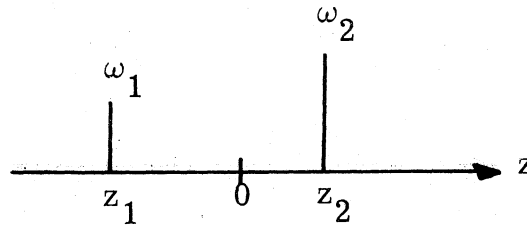
Fig. 8.3. Sketch for Luce ROC
character

The equations for these quantities in terms of the Luce ROC point

(x_1, y_1) are given in Eq. 8.28.

$$z_1 = \ln\left((1 - y_1)/(1 - x_1)\right), z_2 = \ln(y_1/x_1)$$

$$\omega_1 = \left((1 - y_1)(1 - x_1)\right)^{.5}, \omega_2 = (x_1 y_1)^{.5} \tag{8.28}$$

The binomial ROC character for n observations is

$$d\Pi_n(z) = C_k^n \omega_1^{n-k} \omega_2^k \quad \text{at } z = (n - k) z_1 + kz_2 \tag{8.29}$$

Equation 8.29 is not the familiar binomial probability equation unless

the sum of ω_1 and ω_2 is one. Since an expression that is the binomial

probability function would be useful, multiply and divide by the sum

of ω_1 plus ω_2 to obtain Eq. 8.30

$$d\Pi_n(z) = [\omega_1 + \omega_2]^n \; C_k^n \; (\frac{\omega_1}{\omega_1 + \omega_2})^{n-k} \; (\frac{\omega_2}{\omega_1 + \omega_2})^k \qquad (8.30)$$

The standard approximation for the binomial probability function is a

normal probability density with mean value and variance given by

familiar equations. Since we want to use this approximation for z,

and not for the integer value k, we must multiply by the Jacobian of

transformation. This is indicated here

$$[\omega_1 + \omega_2]^{-n} \; \pi_n(z) \approx \frac{1}{\sqrt{2\pi}\,\sigma} \; e^{-\frac{(k-m)^2}{2\sigma^2}} \; \left| \frac{dk}{dz} \right| \qquad (8.31)$$

where the values for the parameters m and σ^2 are given in 8.32.

$$m = n \, (\frac{\omega_2}{\omega_1 + \omega_2}) \qquad \sigma^2 = n \frac{\omega_1 \, \omega_2}{(\omega_1 + \omega_2)^2} \qquad (8.32)$$

This is the continuous ROC character which will be used to approximate

the discrete character of the actual binomial ROC. To obtain this

equation in terms of the variable z, use Eq. 8.29 for z in terms of n

and k to do the transformation. After the very straightforward algebra,

the resulting form for the approximating ROC character is given in

Eq. 8.33.

$$\pi_n(z) \simeq \lfloor \omega_1 + \omega_2 \rfloor^n \frac{1}{\sqrt{2\pi n D}} e^{-\frac{(z - n\beta)^2}{2nD}} \qquad nz_1 \leq z \leq nz_2$$

$$D = (z_2 - z_1)^2 \frac{\omega_1 \omega_2}{(\omega_1 + \omega_2)^2}$$

$$\beta = \frac{z_1 \omega_1 + z_2 \omega_2}{(\omega_1 + \omega_2)^2} \qquad (8.33)$$

The parameter D has the role of a variance term, while the parameter β will be interpreted as a skewness parameter. If the range of z were infinite, the normal ROC character would have to have a mode of zero, which would preclude the use of Eq. 8.33 except for β values of zero. However, the approximating ROC character of Eq. 8.33 applies only to the bounded range of z values that can be obtained from the original Luce character in n observations. If the original Luce character is nonsymmetric, this resulting range will also be nonsymmetric. Therefore, Eq. 8.33 is interpreted as a metastatic normal ROC character.

If the original Luce ROC character is symmetric, the skewness parameter β will be zero, and the range of the approximating normal character will also be symmetric. This is shown in Eq. 8.34.

$$x_1 = 1 - y_1 \Rightarrow \beta = 0, \quad D = (\ln \frac{y_1}{1 - y_1})^2 = (z_2)^2 \qquad (8.34)$$

The parameters D and β apply to the individual distribution; the factor n appears separately in Eq. 8.33. These quantities D and β are the parameters of the original Luce ROC curve that are relevant in studying the limiting form of the binomial ROC resulting from many similar observations.

Table 8.1 presents three examples of Luce ROC curves

x_1	y_1	z_1	z_2	ω_1	ω_2	β	$\dfrac{\beta}{\sqrt{D}}$	D
.0100	.0918	-.0862	2.215	.949	.030	-.164	-.042	.157
.1586	.5000	-.5200	1.150	.649	.273	-.026	-.034	.580
.3085	.6915	-.806	.806	.462	.462	0	0	.650

Table 8.1. Limiting binomial
examples

and the values of D and β evaluated for them. The three points chosen for the Luce ROC all fall on a normal ROC with index d of unity. The first was picked as x_1 = .01, a very strict criterion. The second was chosen at y_1 = .50. The third was chosen at the negative diagonal. A column for the ratio of β to the square root of D has been included. The parameter β is the skewness parameter, giving the amount that the z mode shifts for each observation. The ratio of β to the square root of D indicates the mode shift in standard deviation units. The highest value of the parameter D results from the symmetric Luce ROC. The fact that this is only 65 percent (roughly $\frac{2}{\pi}$) of the original d is familiar to those who have worked

with detectors based upon clipped correlation devices sometimes known as polarity coincidence comparators (Ref. 36). The performance is fairly insensitive to the departure from the negative diagonal until the departure becomes extreme. For very strict criteria, the parameter D is very much lower than the original normal ROC parameter d, and the resultant binormal ROC will be poorer than if a more symmetric original position had been used.

8.3 Fisher-Tippett, Doubly Exponential Character

In Section 3.3.5, the Fisher-Tippett distribution was used to obtain an ROC character. The form of this character is repeated here.

$$\pi(z) = A\ e^{\alpha z\ -\ \beta e^{\alpha z}} \qquad \alpha >.5, \quad \beta > 0 \qquad (8.35)$$

The probability distribution associated with this character arises in studies of the distribution of the largest of a number of independent random variables. In that application, the two multipliers of z are identical. This section will treat a slight generalization of this character, where the multipliers of z need not be identical. This character is shown in Eq. 8.36.

$$\pi(z) = A\ e^{\alpha z\ -\ \beta e^{\gamma z}} \qquad \alpha\gamma > 0, \quad |\alpha| >.5 \qquad (8.36)$$

The coefficients α and γ may be either positive or negative, but they must both have the same sign. Otherwise, the ROC character would not approach zero for large magnitude z. The coefficient β is always positive.

Examination of the first two derivatives of the exponent show

that this ROC character is unimodal. Let $\xi(z)$ denote the exponent

in Eq. 8.36.

$$\xi(z) = \alpha z - \beta e^{\gamma z}$$

$$\xi'(z) = \alpha - \beta \gamma e^{\gamma z}$$

$$\xi^{(n)}(z) = -\beta \gamma^n e^{\gamma z} \qquad n \geq 2 \qquad (8.37)$$

The mode of the ROC character is at the z value for which the first

derivative is zero.

$$\text{mode } z : \frac{1}{\gamma} \ln \frac{\alpha}{\beta \gamma} \qquad (8.38)$$

The second derivative of the exponent is everywhere negative, confirm-

ing that the Eq. 8.38 yields a mode, and not a relative minimum.

$$\xi''(z) = -\beta \gamma^2 e^{\gamma z} < 0 \qquad (8.39)$$

From ROC Curve to Character. The easiest demonstration of the ROC

curves corresponding to the characters of Eq. 8.36 is to exhibit the

ROC curves and then to show that they have this character. The incom-

plete gamma function is defined in Eq. 8.37.

$$P(a,t) = \frac{1}{\Gamma(a)} \int_0^t \lambda^{a-1} e^{-\lambda} d\lambda \qquad a > 0 \qquad (8.40)$$

The ROC curve will be a comparison of two incomplete gamma functions

at different parameter values a, but with the same cut t. The ROC

curves are given parametrically in Eq. 8.41. The notation used gives

the two parameter values in terms of their mean and difference.

$$
\left.
\begin{array}{l}
X(t) = 1 - P(m - \delta, t) \\[2em]
Y(t) = 1 - P(m + \delta, t) \\[2em]
0 < \delta < m
\end{array}
\right\}
\qquad (8.41)
$$

To obtain the ROC character first find the probability

density functions for the dummy variable t, and then transform to the

log likelihood ratio. Since the ROC curve is given in terms of integrals,

the probability density functions are simply the integrands.

$$
f(t \mid N) = \Gamma^{-1}(m - \delta)t^{m - \delta - 1}e^{-t} \qquad (8.42)
$$

$$
f(t \mid SN) = \Gamma^{-1}(m + \delta)t^{m + \delta - 1}e^{-t} \qquad (8.43)
$$

These are Pearson III probability density functions in the variable t. The

likelihood ratio is formed by dividing Eq. 8.43 by 8.42.

$$
e^{z} = t^{2\delta}\, \Gamma(m - \delta)/\, \Gamma(m + \delta) \qquad (8.44)
$$

The likelihood ratio has a Pearson III probability density function, but z

will not. To obtain the probability density functions for z, solve Eq. 8.44

for t.

$$t = \left[\frac{\Gamma\,(m\,+\,\delta)}{\Gamma\,(m\,-\,\delta)}\right]^{\frac{1}{2\delta}} \; e^{\frac{z}{2\delta}} \tag{8.45}$$

It will be convenient to have a symbol for the constant factor in Eq. 8.45.

Call this constant β

$$\beta = \left[\frac{\Gamma\,(m\,+\,\delta)}{\Gamma\,(m\,-\,\delta)}\right]^{\frac{1}{2\delta}} \qquad \beta > 0 \tag{8.46}$$

Rewrite Eq. 8.45 as

$$t = \beta\,e^{\frac{z}{2\delta}} \tag{8.47}$$

e to obtain the Jacobian of transformation.

$$\frac{dt}{dz} = \frac{\beta}{2\delta}\,e^{\frac{z}{2\delta}} = \frac{t}{2\delta} \tag{8.48}$$

The root product density function for t is

$$[\,f(t\,|N)\,f(t\,|SN)\,]^{\cdot\,5} = [\,\Gamma\,(m\,-\,\delta)\;\Gamma(m\,+\,\delta)\,]^{-\cdot\,5}\;\; t^{m-1}\,e^{-t} \tag{8.49}$$

Since the Jacobian of transformation, Eq. 8.48, contains a single power

of t, absorb this into the root product density before substitution for

t according to Eq. 8.44. The result is

$$\pi(z) = \frac{1}{2\delta}\,[\,\Gamma\,(m\,-\,\delta)\;\Gamma(m\,+\,\delta)\,]^{\cdot\,5}\;\beta^{\,m}\,e^{\frac{mz}{2\delta}}\,e^{-\beta e^{\frac{z}{2\delta}}} \tag{8.50}$$

To simplify this equation to the general form of Eq. 8.36, merely identify the constant factor, A, and the two multipliers of z, α and γ.

$$A = \frac{1}{2} \left[\Gamma(m - \delta) \Gamma(m + \delta) \right]^{-.5} \beta^{m}$$

$$\alpha = \frac{m}{2\delta} \qquad \gamma = \frac{1}{2\delta} \qquad \alpha \geq .5, \gamma > 0 \qquad (8.51)$$

The restriction that the parameter δ be between zero and m converts to the restriction that the coefficient α be greater than one-half and that γ be positive.

Whenever an ROC character is nonsymmetric, a second ROC character and ROC curve are obtained by reversing the z-axis. Specifically, replace z by -z everywhere in the first ROC character, and replace X by 1 - Y and Y by 1 - X in the equation for the first ROC curve. When this is done for the doubly exponential ROC character of Eq. 8.36, another doubly exponential ROC character is obtained

$$z \to -z \qquad X \to 1 - Y \qquad Y \to 1 - X \qquad (8.52)$$

$$\pi(z) = A\, e^{-\alpha z - \beta^{-\gamma z}} \qquad \alpha > .5, \beta > 0, \gamma > 0 \qquad (8.53)$$

The coefficients are the same as those given in Eq. 8.51, and are positive. The two minus signs can be absorbed into the terms α and γ, giving an ROC character for which the new α and γ parameters are both negative, but the coefficient β is still positive. The ROC curve is given parametrically in terms of the incomplete gamma function.

$$X(t) = P(m + \delta, t)$$

$$Y(t) = P(m - \delta, t)$$

$$0 < \delta < m$$

(8. 54)

From ROC Character to ROC Curve. Whenever a doubly exponential ROC character is encountered, the parameters necessary for the incomplete gamma functions used in the ROC curve formulae are easily obtained. The m parameter is the ratio of the two multipliers of z, α and γ. These coefficients will either be both positive or both negative, and hence their ratio is positive. The difference parameter, δ, is given solely in terms of the γ parameter. Although there are only two degrees of freedom in the Fisher-Tippett Doubly Exponential Character, there are three coefficients in the exponential. One must therefore make sure that the β parameter has the appropriate value. This is summarized in Eq. 8. 55.

$$m = \frac{\alpha}{\gamma} \qquad \delta = \frac{1}{2|\gamma|}$$

$$\text{Check:} \quad \beta = [\Gamma(m + \delta) / \Gamma(m - \delta)]^{\frac{1}{2\delta}}$$

(8. 55)

Tables of the incomplete gamma function are not always available. The incomplete gamma function is related to the chi-square distribution whenever the parameter values $m \pm \delta$ are integers or half integers. The relation is that the chi-square value is $2t$ and the

degrees of freedom are twice the incomplete gamma parameter.

$$P(m \pm \delta, t) = P(\chi^2 = 2t \mid \nu = 2m \pm 2\delta)$$

m, δ are both odd quarter positive integer

or are half positive integer or positive integer   (8.56)

Whenever the degrees of freedom of the chi-square distribution are

even integers, the probability may be written in terms of a finite

power series, a polynomial, times a simple exponential.

$$1 - P(m \pm \delta, t) = e^{-t} \sum_{n=0}^{\frac{m-1 \pm \delta}{}} \frac{t^n}{n!}$$

m, δ are both half positive integer or both

positive integer   (8.57)

Tables for the incomplete gamma function and the chi-square are nor-

mally available for only a low number of degrees of freedom. Typically,

the largest tables available will terminate at thirty degrees of freedom.

Approach to Normal.   As the number of degrees of freedom become

large, the Chi-Square distribution approaches a normal distribution.

The following approximation is usually suggested

$$1 - P(\chi^2 \mid \nu) \approx \Phi\left(-\sqrt{2\chi^2} + \sqrt{2\nu - 1}\right), \quad \nu \text{ large}   (8.58)$$

The table in Cramér (Ref. 5) suggests ν of thirty is sufficiently large

to use this approximation, while the NBS Handbook (Ref. 6, 26.4.13)

suggests ν larger than 100. ROC curves deal with a comparison of two distribution functions. Our experience has been that such approximations may be used at much lower values of parameter in a comparison than would be allowed were one evaluating a single distribution. The approximation of Eq. 8.58 used in evaluating the ROC curves of Eq. 8.51 will necessarily yield a normal ROC curve. The parameter d will be given by the square of the difference of the two terms involving the two degrees of freedom. When both of the degrees of freedom are large, and their difference is small compared to their average value, some standard approximations make a further simplification in determining the normal ROC parameter. The manipulations are omitted but the result is given below.

$$\hat{d} = (\sqrt{4m - 1 + 2\delta} - \sqrt{4m - 1 - 2\delta}) \cong 4\delta^2 \, m^{-1} = (\alpha\gamma)^{-1} \quad (8.59)$$

Equation 8.59 strongly suggests that no single parameter in the double exponential character can be interpreted as the quality parameter for this distribution. Rather, detectability will be inversely proportional to the product of the two multipliers of z, α and γ.

Since the ROC character is nonsymmetric, the exact ROC curve will be nonsymmetric. When the ROC curve is plotted on normal-normal paper, the nonsymmetry should be manifested as a slope different from one. It would be advantageous to have a value obtainable from the character which would be simply related to this slope. In Carver's tables (Ref. 28) for the Pearson III distribution, the skewness of the incomplete gamma function is given in terms of the parameter a.

$$\alpha_3 = \text{skewness of } P(a,t) = 2\,a^{-.5} \tag{8.60}$$

Although the ROC equations use two incomplete gamma distributions with different skewness, under the condition that m is much larger than the difference, δ, α_3 is of the order of $2/\sqrt{m}$. No quantitative measure is implied. It is implied that the skewness should disappear as m is increased.

The approximation given in Eq. 8.59 groups together those ROC curves with the same ratio of δ^2 to m. If the role of δ is replaced by a new parameter, d_a, using the substitution $\delta = .5\sqrt{md_a}$ the ROC character is rewritten as

$$\pi(z) = A\,e^{\sqrt{\frac{m}{d_a}}\,z} - \beta\,e^{\frac{z}{\sqrt{md_a}}} \tag{8.61}$$

Collecting Eqs. 8.41 and 8.56, the ROC curve for the above ROC character in terms of the chi-square distribution is

$$\frac{Y}{X} = 1 - P(\chi^2 = 2t \mid \nu = 2m \pm \sqrt{md_a}) \tag{8.62}$$

For tables which give the upper tail area, which is one minus the chi-square probability,

$$\frac{Y}{X} = Q(\chi^2 = 2t \mid \nu = 2m \pm \sqrt{md_a}) \tag{8.63}$$

Holding d_a constant, it can be shown that β behaves like m minus one-eighth d_a as m approaches infinity.

$$m \to \infty, \quad \beta \to m - \frac{d_a}{8} \tag{8.64}$$

When this is used in the expression for the parameter t, Eq. 8.47 and

the limiting normal form of the chi-square distribution, Eq. 8.58, the

locus of the $z = 0$ point is the negative diagonal.

$$\text{at } z = 0, \quad t = \beta, \quad \frac{Y}{X} \to \Phi(\pm .5 \sqrt{d_a}) \tag{8.65}$$

For m values much larger than the detection index d_a, β may be

considered equal to m. In this limiting form, the double exponential

ROC character approaches

$$m \to \infty, \quad \pi(z) \to A \, e^{\frac{\sqrt{m}}{d_a} z - m e^{\frac{z}{\sqrt{m d_a}}}} \tag{8.66}$$

Examination of the derivatives of the exponent for this limiting form,

shows that the mode has shifted to zero and that the derivatives of order

three and larger rapidly approach zero. The second derivative of the

exponent is the only derivative that does not approach zero.

$$\left. \begin{array}{l} \xi'(0) = \alpha - \beta\gamma = \dfrac{\sqrt{m}}{d_a} - m \dfrac{1}{\sqrt{m d_a}} = 0 \\[3mm] \xi''(0) = -\beta\gamma^2 = -\dfrac{1}{d_a} \\[3mm] \xi'''(0) = -\beta\gamma^3 = \dfrac{1}{d_a \sqrt{m d_a}} \to 0 \end{array} \right\} \tag{8.67}$$

The power series approximation to the exponent is dominated by the second term.

$$\xi(z) \to \xi''(0) \frac{z^2}{2} = -\frac{z^2}{2d_a} \tag{8.68}$$

This is the exponent for the normal ROC character.

Using Eq. 8.63, ROC curves were evaluated for $\sqrt{d_a} = 1, 2,$ and 3, and all possible small m values yielding integer degrees of freedom. These are shown in Fig. 8.4. As can be seen from the plots, the grouping of the parameters m and δ provided by Eq. 8.59 was fairly effective. That grouping was based on an approximation to the chi-square distribution normally considered valid for degrees of freedom of the order of 30 to 100 or greater. The actual values for the degrees of freedom used in plotting Fig. 8.4 are given 'n Table 8.2.

d_a	m	δ	ν_{SN}	ν_N
1	1	0.50	3	1
	2.25	0.75	6	3
	4	1.00	10	6
	6.25	1.25	15	10
	9	1.50	21	15
4	4	1	12	4
	9	2	24	12
9	4	3	14	2
	9	4.5	27	9

Table 8.2. Parameters for $X = Q(\chi^2 | \nu_N)$, $Y = Q(\chi^2 | \nu_{SN})$

Y $= Q(\chi^2 | \nu = 2m + \sqrt{dm})$

X $= Q(\chi^2 | \nu = 2m - \sqrt{dm})$

$$\pi(z) = A e^{-\sqrt{\frac{m}{d}} z - \beta e^{-\frac{z}{md}}}$$

where $\beta = \left[\dfrac{\Gamma(m+.5\sqrt{dm})}{\Gamma(m-.5\sqrt{dm})} \right]^{\frac{1}{\sqrt{dm}}}$
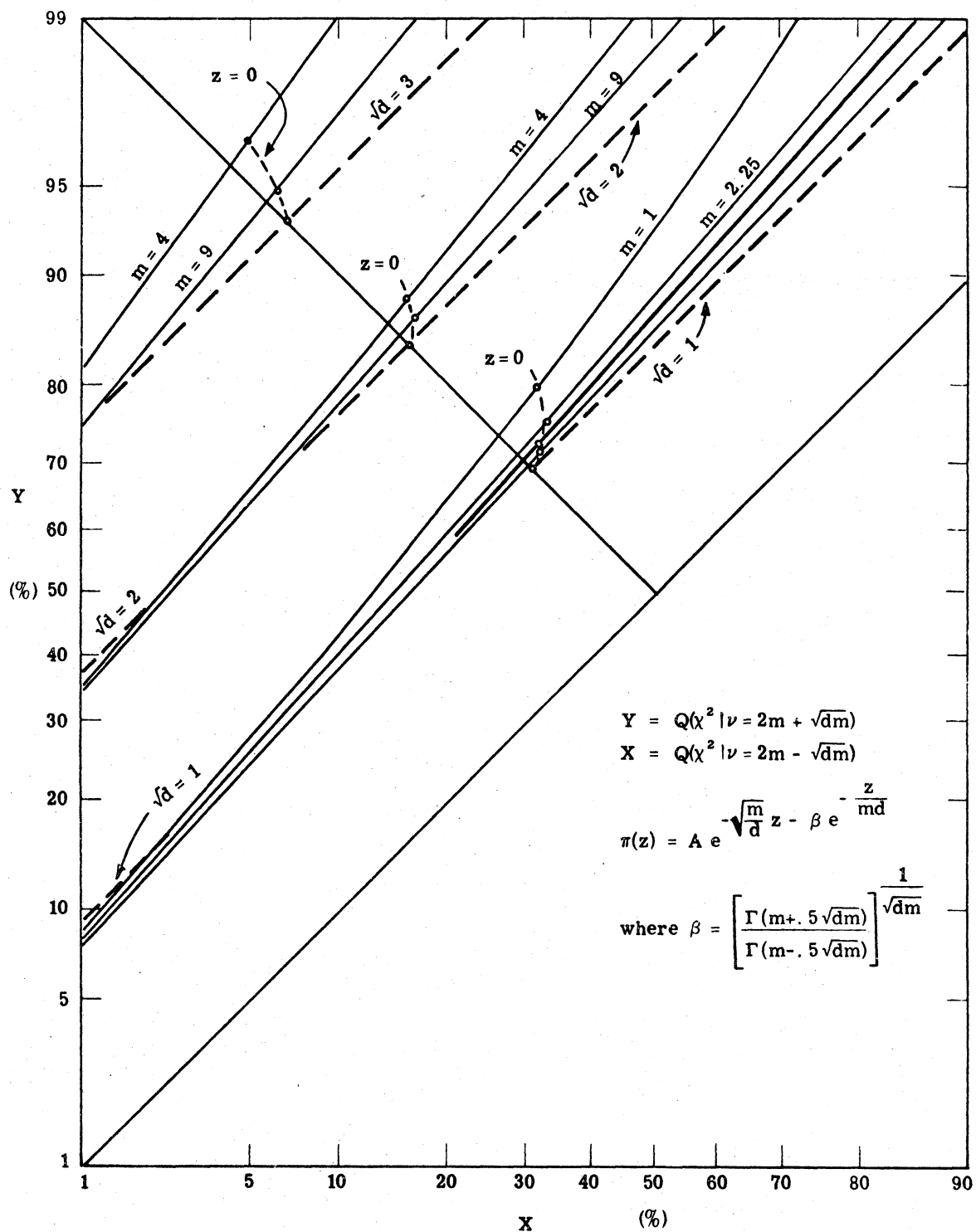
Fig. 8.4. Fisher-Tippett ROC curves

The normal ROC curves with parameter $\sqrt{d}$ = 1, 2, and 3 lie slightly below the double exponential character ROC curves, crossing them in the neighborhood of x = .02. The development in this subsection was suggested by the standard approximation to the chi-square distribution for large degrees of freedom. This approximation indicated that the proper way to consider the ROC character

$$\pi(z) = A \ e^{\alpha z - \beta \ e^{\gamma z}} \tag{8.36}$$

is to consider the product of the two z multipliers, $\alpha\gamma$, as the reciprocal of the approximating normal index, and to consider the ratio of these two z multipliers, α/γ, as a skewness index. By a sequence of approximations, it was established that in the limit, as m = α/γ approached infinity, that the ROC curves do become symmetrical (indicated by the z = 0 point falling on the negative diagonal), and the ROC character does approach normality with the index equal to the reciprocal of $\alpha\gamma$. ROC curves based on available tables indicated that this same viewpoint of the parameters in the ROC character is valid for small integer m, as well as when m approaches infinity.

Other Cases. The cases considered numerically showed a slope greater than one on normal-normal paper, contrary to most of the examples that we have considered. The reverse double exponential ROC character of Eq. 8.50 would show the corresponding slope less than one, since these ROC curves would be the mirror image of those shown in Fig. 8.4.

In the above subsection, the concern was with cases where both parameters representing degrees of freedom became large. A particular example that appeared earlier, in Section 4. 4, was the ROC curve

$$Y = X(1 + |\ln X|) \tag{8.69}$$

This is a particular case of the double exponential character; namely, when α is $\gamma + .5$. This fixes m - δ at unity, and the basic equations for the ROC curves are

$$X(t) = 1 - P(1, t) \tag{8.70}$$

$$Y(t) = 1 - P(2m - 1, t) \tag{8.71}$$

This is a case where Eq. 8. 57 can be used effectively, if 2m - 1 is an integer.

$$X(t) = e^{-t}, \quad Y(t) = e^{-t} \sum_{n=0}^{2m-2} \frac{t^n}{n!} \tag{8.72}$$

The dummy variable t can be removed by solving the simple X(t) equation, and substituting into the second equation. The result is

$$Y = X\left(1 + \left|\ln X\right| + \frac{1}{2!}\left|\ln X\right|^2 + \ldots + \frac{1}{(2m-2)!}\left|\ln X\right|^{2m-2}\right) \tag{8.73}$$

The first nine such ROC curves are plotted in Fig. 8. 5.
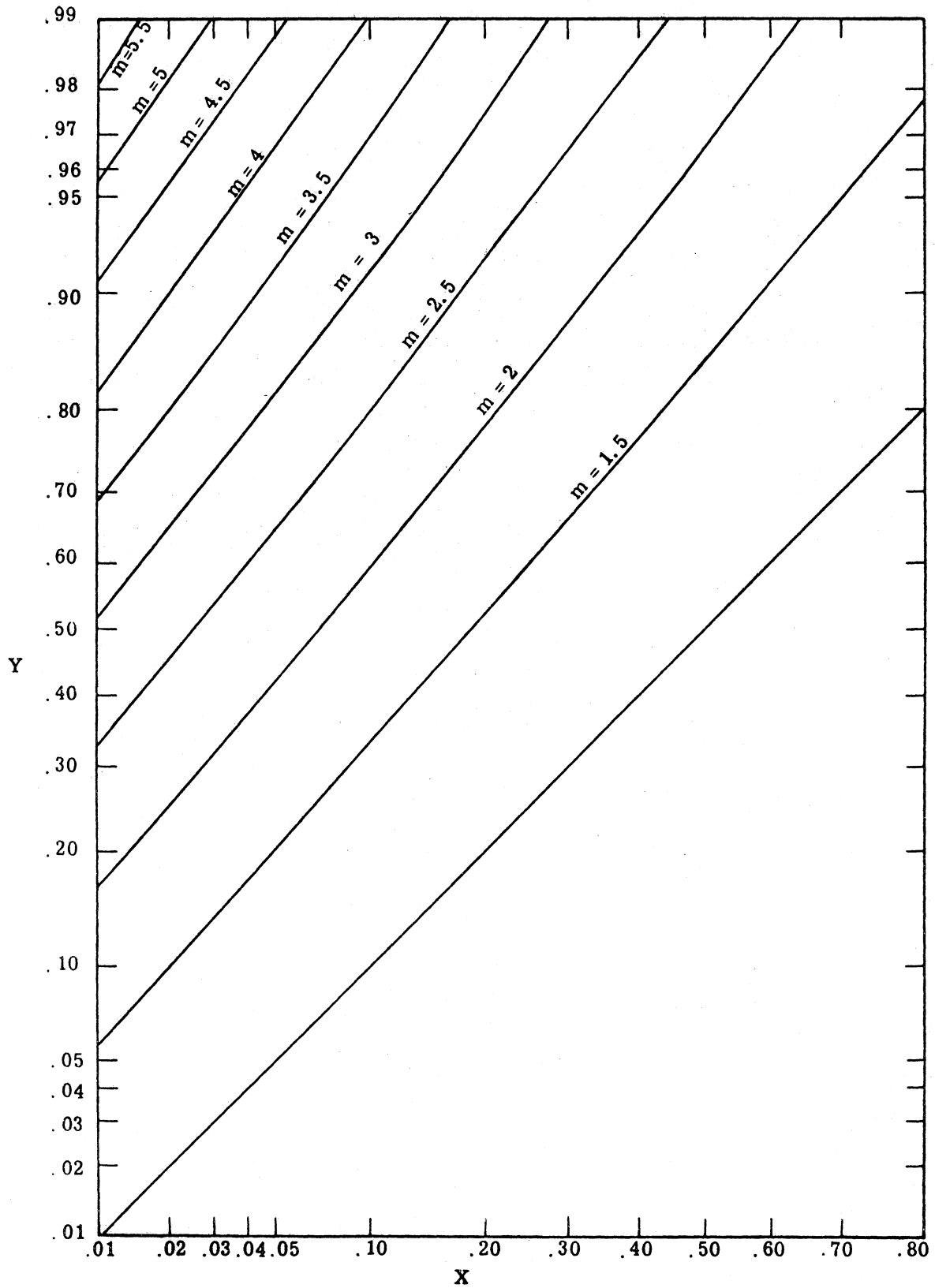
Fig. 8.5. ROC curves of Eq. 8.73

CHAPTER IX

ADDITION OF ROC CHARACTERS

This chapter deals with the ROC character addition theorem, and with two families of two-parameter ROC characters. The character addition theorem says that if a number of regular ROC characters are combined in a weighted sum, then the resultant ROC curve is formed by using the same type of weighted sum over the original regular ROC curves.

$$\pi(z) = \sum c_k \pi_k(z) \Rightarrow X(z) = \sum c_k X_k(z)$$
$$Y(z) = \sum c_k Y_k(z)$$

$$(9.1)$$

The ROC characters treated in this chapter are the type containing a power of z, and an exponential term in z. The two parameters in each family are the power of z, and the coefficient in the exponential. The first family considered in Section 9.2 is the **Pearson III.**

Pearson III Type: $\pi(z) = C z^\rho e^{-z/k}$ $\qquad$ (9.2)

The power parameter, ρ, may be any real value greater than -1. The exponential is a simple exponential, that is, it is linear in z. The second family considered in Section 9.3 is a form of the Halsted family mentioned in Chapter I. These latter ROC characters are related to the Hermite polynomials, and derivatives of the normal density function.

$$\text{H Type: } C z^n e^{-(z-z_0)^2/2k} \tag{9.3}$$

This type of ROC character is just being developed at this time, and the power of z is restricted to being an integer greater than, or equal to, one. The H-type ROC character differs from the Pearson III character in that the exponential term is quadratic in z. Both types of ROC character are bounded below, and unbounded above; that is, the z-range is from some negative value up to plus infinity.

The addition theorem allows the expansion of either or both of these types to ROC characters which consists of polynomials in z, times either a linear or quadratic exponential term. When the equations for the ROC curves in terms of the variable z have been developed, one may immediately obtain the ROC curve for such a generalized character.

There is no restriction of the use of the character addition theorem to combining characters of the same general family. Any types of character may be added together.

9.1 Character Addition Theorem

This section states and proves the character addition theorem. The statement and the proof are in terms of the integrated ROC character, to obtain the maximum generally. It will therefore apply to ROC curves that are differentiable, and have ordinary ROC characters; it will also apply to polygon type ROC curves, with straight lines joining the vertices, as well as mixtures. The key mathematical point in the proof is the ability to interchange a finite summation with

integration.

<u>ROC Character Addition Theorem</u>: Let $\Pi_k(z)$ be the integrated ROC

characters for the regular ROC curves $(X_k(z),\ Y_k(z)), k = 1,\ 2,\ \ldots\ n.$

Let c_k be any positive numbers whose sum is one. Then

$$\Pi(z) = \sum_{k=1}^{n} c_k \Pi_k(z) \tag{9.4}$$

is an integrated ROC character, the ROC curve is regular and is given by

$$X(z) = \sum_{k=1}^{n} c_k X_k(z) \qquad Y(z) = \sum_{k=1}^{n} c_k Y_k(z) \tag{9.5}$$

<u>Proof</u>:

The four requirements for a nonnegative function to be

an integrated ROC character of a regular ROC curve were

given in Chapter III, and are

(1) the lower limit is zero $\qquad \Pi(-\infty) = 0$
$$\tag{3.6}$$

(2) monotone growth $\qquad d\Pi(z) \geq 0$
$$\tag{3.7}$$

(3) unit "N" value $\qquad \int_{-\infty}^{\infty} e^{-.5z}\ d\Pi(z) = 1$
$$\tag{3.8}$$

(4) unit "SN" value $\qquad \int_{-\infty}^{\infty} e^{+.5z}\ d\Pi(z) = 1$
$$\tag{3.9}$$

Each of the given individual integrated ROC characters,

Π_k, have these four properties. Does the newly formed

sum of Eq. 9.4 also have these properties? If the individ-

ual integrated ROC characters are all zero at $z = -\infty$,

then the sum will also be zero. Similarly, if each of the

individual integrated ROC characters exhibits monotone growth (not necessarily strict monotone growth), then a finite sum does also. To verify Eqs. 3.8 and 3.9 for the newly formed sum one need merely note that the definite Stieltjes integral is a linear operator to both functions. Therefore the integral with respect to the sum is the sum of the integrals. Since the coefficients c_k add to one, the sum of Eq. 9.4 will satisfy the properties 3.8 and 3.9.

The notation for the ROC curve in terms of the single parameter z implies the dual parameter of a decision based on likelihood ratio. Specifically, whenever the integrated ROC character has a jump, the associated straight segment of the ROC curve is given in terms of a randomizing parameter, r. The complete equation for the ROC curve in terms of the parameter z is actually given in terms of a cut along the z axis, β, and the randomizing parameter, r, whenever it is needed.

$$\left\{ \begin{array}{l} y(\beta, \ r) \\ x(\beta, \ r) \end{array} \right\} = \int_{\beta}^{\infty} e^{\pm .5z} \ \pi(z) \ dz$$

$$+ re^{\pm .5\beta} \omega_{\pi}(\beta) + \sum_{z \,>\, \beta} e^{\pm .5z} \omega_{\pi}(z) \tag{9.6}$$

Most of this present work has dealt with the first integral alone, since for most ROC characters the jump function

ω_π is zero everywhere. Each of the original n ROC

curves are given by an equation such as Eq. 9.6. To

determine the ROC curve for the integrated ROC character

of Eq. 9.4 substitute the equation for the new integrated

character into Eq. 9.6.

$$\left\{\begin{matrix} Y(\beta, r) \\ X(\beta, r) \end{matrix}\right\} = \int_\beta^\infty e^{\pm.5z} \sum_k c_k \pi_k(z)\, dz$$

$$+ re^{\pm.5\beta} \sum_k c_k \omega_{\pi k}(\beta) + \sum_{z > \beta} e^{\pm.5z} \sum_k c_k \omega_{\pi k}(z)$$

$$(9.7)$$

In the first term, the interchange of the order of summa-

tion and integration is allowed because the sum is finite.

The second term dealing with the jump at the point β con-

tains only one summation and, therefore, any reordering

of the terms is valid. The third summation deals with one

finite sum and one possibly infinite sum. However, since

all of the terms involved are positive, if the sums converge

in any order they will converge in any other order. There-

fore, the summation over k may be taken outside of the

expression to obtain Eq. 9.8.

$$\left\{\begin{matrix} Y(\beta, r) \\ X(\beta, r) \end{matrix}\right\} = \sum_k c_k \left[\int_\beta^\infty e^{\pm.5z} \pi_k(z)\, dz + re^{\pm.5\beta} \omega_{\pi k}(\beta) \right.$$

$$+ \sum_{z > \beta} e^{\pm.5z} \, \omega_{\pi k}(z) \Bigg] \tag{9.8}$$

The term inside each bracket is the pair of symmetric expressions for the vertical and horizontal coordinates of the ROC curve, Y_k and X_k. The proof is complete.

9.2 Pearson III ROC Characters

In Section 3.3.4 the Pearson III character was obtained from the Pearson III probability distribution. The form of the ROC character is

$$\pi(z) = \frac{1}{\Gamma(p+1)} \left(\frac{\sqrt{B}}{B-1}\right)^{p+1} (z-z_0)^p \, e^{-\frac{B+1}{B-1}\left(\frac{z-z_0}{2}\right)} , z > z_0 \tag{9.9}$$

$$B > 1, \ p > -1, \ z_0 = -(p+1) \ln B$$

In this section explicit expressions for the ROC curve are obtained when the parameter p is either an integer or when twice p is an integer. This latter is usually referred to as half-integer order. In each case the steps will be to display an ROC curve in terms of a dummy parameter t, obtain the relation between t and z, and check to be sure the ROC character is truly Pearson III.

Integer Order. Consider any nonnegative integer n, and the following ROC curve.

$$\rho = n \geq 0 \qquad X = e^{-Bt} \sum_{k=0}^{n} \frac{(Bt)^k}{k!} \qquad t > 0$$

(9.10)

$$Y = e^{-t} \sum_{k=0}^{n} \frac{t^k}{k!}$$

Both X and Y are unity when t is zero, and decrease toward zero as t approaches infinity. Therefore, X and Y, as usual, are treated as one minus the distribution function for t. The probability density functions for t are the negative of the derivative of X and Y with respect to t. Proceeding formally, the derivative of X with respect to t is

$$\frac{dX}{dt} = e^{-Bt} \sum_{k=1}^{n} \frac{B^k t^{k-1}}{(k-1)!} - B e^{-Bt} \sum_{k=0}^{n} \frac{B^k t^k}{k!}$$

(9.11)

Collect the two sums, first changing indices so that the coefficients of a single power may be determined, to obtain

$$\frac{dX}{dt} = e^{-Bt} \left[\sum_{r=0}^{n-1} \frac{B^{r+1} t^r}{r!} - \sum_{k=0}^{n} \frac{B^{k+1} t^k}{k!} \right] = - \frac{B^{n+1} t^n}{n!} e^{-Bt}$$

(9.12)

Therefore, the density function for t under condition N is

$$f(t \mid N) = \frac{B^{n+1}}{n!} t^n e^{-Bt}$$

(9.13)

Since the equation for Y is similar to that for X with B set equal to one,

$$f(t \mid SN) = \frac{1}{n!} \, t^n \, e^{-t} \qquad (9.14)$$

The ratio of these two equations gives the likelihood ratio, e^z.

$$e^z = B^{-(n+1)} \, e^{(B-1)t} \qquad (9.15)$$

Solving for z,

$$z = (B-1) \, t - (n+1) \ln B \qquad (9.16)$$

By Eqs. 9.13 and 9.14, the variable t has a Pearson III density function. Since z is a linear transformation of t, it too will have Pearson III density functions under both condition N and SN. Explicitly, t is

$$t = \left(\frac{z - z_0}{B-1} \right), \qquad z_0 = -(n+1) \ln B \qquad (9.17)$$

Any of the several possible routes to obtain the ROC character will reobtain Eq. 9.9.

It is therefore established that Eq. 9.10 is the ROC curve for a Pearson III ROC character of integer order. Eq. 9.16 gives the explicit relation between z and the variable t.

Half-Integer Order. Consider any value for ρ which is one-half larger than an integer. The claim is that Eq. 9.18 is the corresponding ROC curve.

$$\rho = n + .5 \geq -.5 \qquad X = 2 - 2\,\Phi\,(\sqrt{2Bt}\,) + \sum_{k=0}^{n} \frac{(Bt)^{k+.5}}{\Gamma(k+1.5)}\,e^{-Bt}$$

$$(9.18)$$

$$Y = 2 - 2\,\Phi\,(\sqrt{2t}\,) + \sum_{k=0}^{n} \frac{t^{k+.5}}{\Gamma(k+1.5)}\,e^{-t}$$

If the value for ρ is $-.5$, so that n would take on the value -1, the summation is omitted. The function $\Phi(\ )$ is the normal distribution function, and the range of t is from zero to infinity. Differentiate X with respect to t in two steps. First consider the term involving the normal distribution function. The detailing of differentiation is given in Eq. 9.19

$$\left. \begin{aligned} \frac{d}{dt}\,2\,\Phi\,(\sqrt{2Bt}\,) &= 2\,\phi\,(\sqrt{2Bt}\,)\,\sqrt{2B}\,\frac{1}{2}\,t^{-.5} \\[2mm] &= \sqrt{2Bt}^{\,-.5}\,\frac{1}{\sqrt{2\pi}}\,e^{-\frac{(\sqrt{2Bt}\,)^{2}}{2}} \\[2mm] &= \sqrt{\frac{B}{\pi}}\,t^{-.5}\,e^{-Bt} \end{aligned} \right\} \qquad (9.19)$$

The square root is squared in the normal density function, leaving a linear exponential term multiplied by t to the power minus one-half. The differentiation of the sum is subject to the same collapsing, or telescoping that occurred for integer order, with the exception that two terms are left after the cancellation instead of one.

$$\frac{d}{dt} \left\{ -e^{-Bt} \sum_{k=0}^{n} \frac{(Bt)^{k+.5}}{\Gamma(k+1.5)} = Be^{-Bt} \sum_{k=0}^{n} \frac{(Bt)^{k+.5}}{\Gamma(k+1.5)} - e^{-Bt} \sum_{k=0}^{n} \frac{B^{k+.5}t^{k-.5}}{\Gamma(k+.5)} \right.$$

$$= e^{-Bt} \left[\sum_{k=0}^{n} \frac{B^{k+1.5}t^{k+.5}}{\Gamma(k+1.5)} - \sum_{k=0}^{n} \frac{B^{k+.5}t^{k-.5}}{\Gamma(k+.5)} \right]$$

$$\left. = e^{-Bt} \left[\frac{B^{n+1.5}t^{n+.5}}{\Gamma(n+1.5)} - \frac{B^{.5}t^{-.5}}{\Gamma(.5)} \right] \right\} \quad (9.20)$$

The two remaining terms contain the desired large power of t, and a term looking very much like that in the last line in Eq. 9.19. Indeed, if the gamma function is evaluated at one-half, it is simply the square root of π.

$$\Gamma(.5) = \sqrt{\pi} \quad (9.21)$$

In the sum of Eqs. 9.19 and 9.20, these similar terms will cancel.

$$f(t|N) = \frac{d}{dt}(1 - X) = \frac{B^{n+1.5}t^{n+.5}}{\Gamma(n+1.5)} e^{-Bt} \quad (9.22)$$

The density function for t under the condition SN is obtained by setting B equal to one.

$$f(t|SN) = \frac{t^{n+.5}}{\Gamma(n+1.5)} e^{-t} \quad (9.23)$$

Divide Eq. 9.23 by 9.22 to obtain the likelihood ratio.

$$e^z = B^{-(n+1.5)} e^{(B-1)t} \quad (9.24)$$

The logarithm of the likelihood ratio, z, is linear with the variable, t

and, therefore, will have the same type of density function as t. Since these are both Pearson III, there is no need to detail the familiar route and reobtain the ROC character, Eq. 9.9, for half-integer order.

This has established the equations for the ROC curves for Pearson III character for integer and half integer order. Other forms of these equations may be obtained in terms of the chi-square distribution, and the incomplete gamma function. The ROC curve is given in terms of a comparison of two points along either of these distributions, using the same number of degrees of freedom. This is contrasted to the case of the Fisher-Tippett Doubly Exponential Character, which was obtained from the incomplete gamma function, or chi-square distribution, by reading at the same argument but for two different degrees of freedom.

The most skewed of the characters considered in this section is that with $\rho = -.5$. The ROC curves for several B values are shown on normal-normal paper in Fig. 9.1. There is a special coordinate system that makes these curves plot as straight lines. It is based on the error function, erf t, and is used to display the same ROC curves as before in Fig. 9.2.

9.3 H-Type ROC Character

The character under study in this section consists of a power of z, times a quadratic term in z in the exponential. The general form is
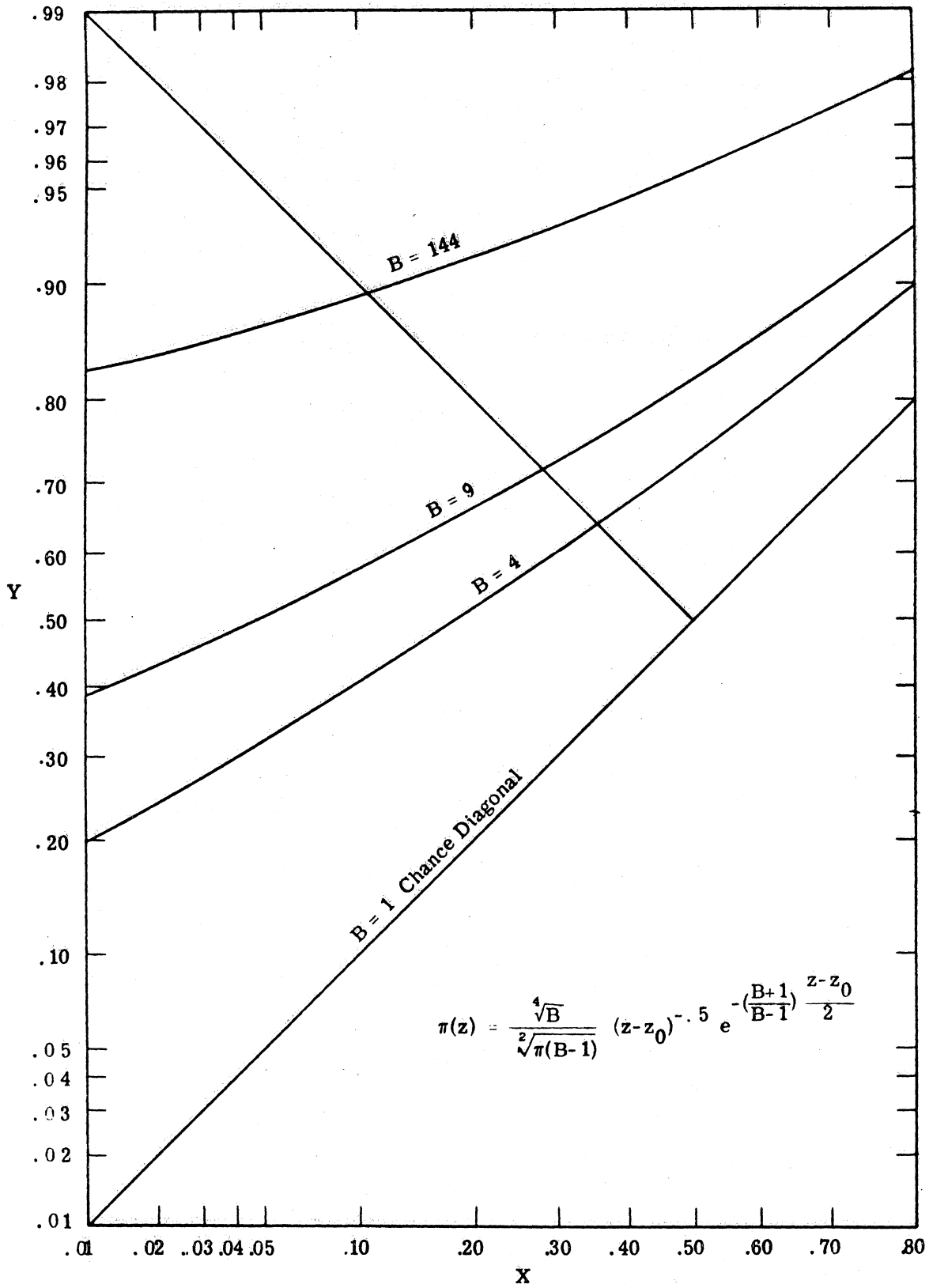
Fig. 9.1. ROC curves Pearson III,
$\rho = -0.5$, $z_0 = -.5 \ln B$

$$\pi(z) = \frac{\sqrt[4]{B}}{\sqrt[2]{\pi(B-1)}} \, (z-z_0)^{-.5} \, e^{-(\frac{B+1}{B-1}) \frac{z-z_0}{2}}$$
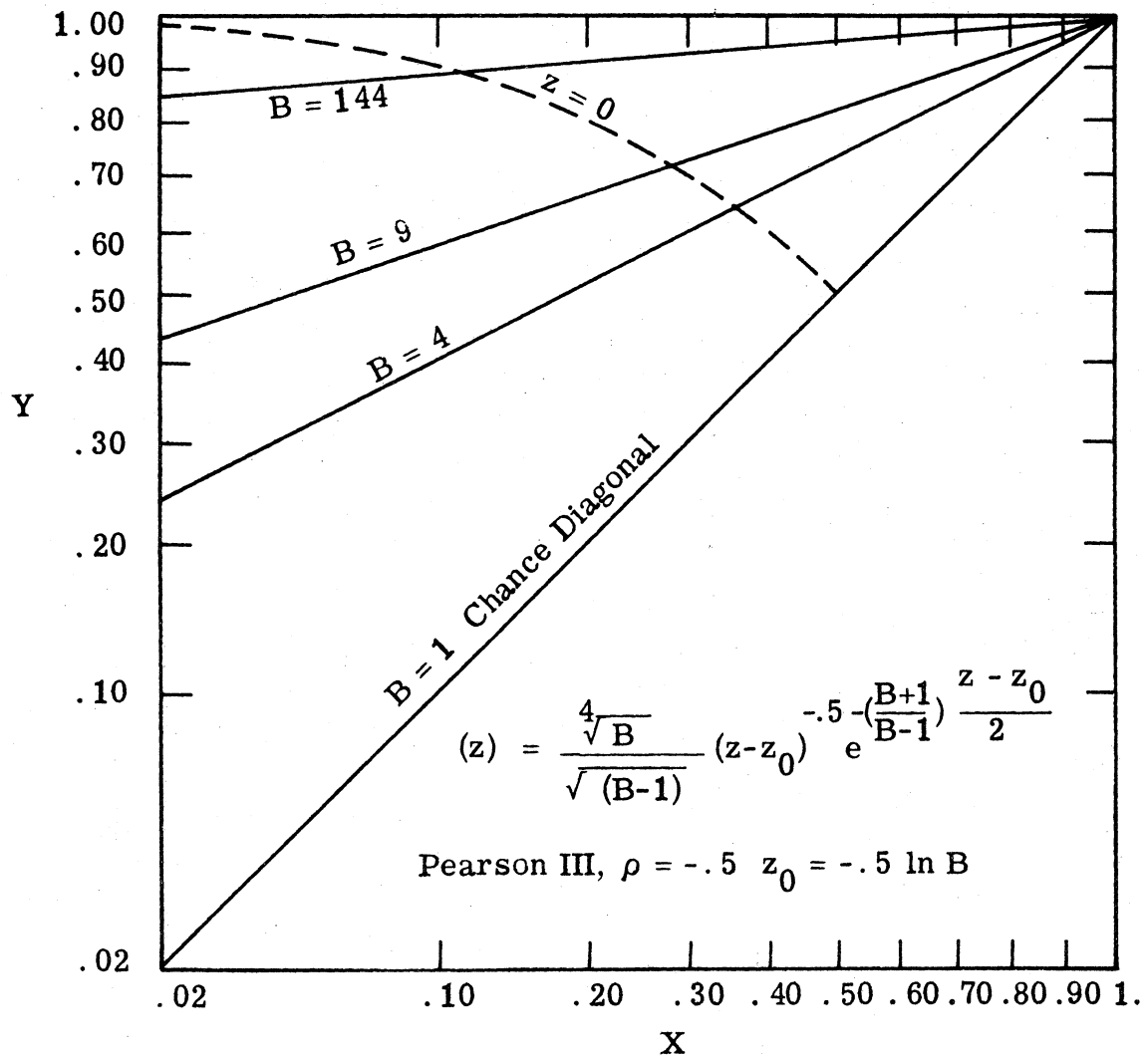
Fig. 9.2. Pearson III, $\rho = -.5$, ROC curves
on erf-erf paper

$$\pi(z) \quad C \, (z - z_0)^n \, e^{-(z - z_0)^2 / \, 8\alpha^2} \quad , \; z \geq z_0, \; n > 0 \qquad (9.25)$$

The term z_0 is the minimum value for z and will be negative. If the quadratic in the exponential were expanded, it would contain both a term in z^2 and a term in z. The constant value, C, is a function of both the power n and the detection index α. No general expression for this constant, C, or the value of z_0 in terms of n and α is known at this writing. The exponential with the quadratic term can be expressed as a normal density function, and the form of the constant C changed to make subsequent manipulation simpler. The following form of the ROC character will be used

$$\pi(z) = A \frac{1}{2\alpha} e^{-.5\alpha^2} \left(\frac{z - z_0}{2\alpha} \right)^n \phi \left(\frac{z - z_0}{2\alpha} \right) \qquad (9.26)$$

To obtain the equations for the ROC curve, multiply the character by $e^{\pm.5z}$, and integrate. The integration will be simplest with the change variable

$$t = \frac{z - z_0}{2\alpha}, \quad \frac{dt}{dz} = \frac{1}{2\alpha}, \quad t \geq 0 \qquad (9.27)$$

The multiplier necessary for the integration can be expressed in terms of α, t, and z_0.

$$e^{\pm.5z} = e^{\pm \alpha t \, \pm \, .5 z_0} \qquad (9.28)$$

Reversing the usual route, the density functions for t' are obtained from the corresponding density functions for z, by multiplying Eq. 9.28 by 9.26, and by the Jacobian of transformation.

$$f\left(t\middle|\begin{matrix}SN\\N\end{matrix}\right) = A\,e^{\pm\alpha t\,\pm\,.5z_0\,-\,.5\alpha^2}\,t^n\,e^{.5t^2}\,/\,\sqrt{2\pi} \qquad (9.29)$$

Of the four terms in the exponential, the three not involving z_0 form a perfect square. When this is done, the densities can again be expressed in terms of a normal density function.

$$f\left(t\middle|\begin{matrix}SN\\N\end{matrix}\right) = A\,e^{\pm.5z_0}\,t^n\,\phi(t\mp\alpha) \qquad (9.30)$$

The equations for the ROC curve can then be formally expressed by integrating the density functions.

$$\left.\begin{matrix}Y = Ae^{.5z_0}\int_T^\infty t^n\,\phi(t-\alpha)\,dt\\[2em]X = Ae^{-.5z_0}\int_T^\infty t^n\,\phi(t+\alpha)\,dt\end{matrix}\right\} \qquad (9.31)$$

At this point the symmetric form for these equations has to be abandoned in order to obtain simpler integrals. Make a change of variables in each of the integrals of Eq. 9.31 to obtain a single letter argument for the normal density function.

$$Y = Ae^{.5z_0} \int_{T-\alpha}^{\infty} (t+\alpha)^n \phi(t) \, dt$$

$$\text{(9.32)}$$

$$X = Ae^{-.5z_0} \int_{T+\alpha}^{\infty} (t-\alpha)^n \phi(t) \, dt$$

The derivatives of the normal density function can be expressed in terms of a polynomial times the normal density function. These polynomials are known as Hermite polynomials. Using tables of Hermite polynomials one may develop formulae for the integral of a single power times the normal density function. Such a table is given here.

$$\int \phi(t) \, dt = \Phi(t)$$

$$\int t\phi(t) \, dt = -\phi(t)$$

$$\int t^2\phi(t) \, dt = \Phi(t) - t\phi(t)$$

$$\int t^3\phi(t) \, dt = -(t^2 + 2)\phi(t)$$

$$\int t^4\phi(t) \, dt = 3\Phi(t) - (t^3 + 3t)\phi(t)$$

$$\int t^5\phi(t) \, dt = -(t^4 + 4t^2 + 8)\phi(t)$$

$$\int t^6\phi(t) \, dt = 15\Phi(t) - (t^5 + 5t^3 + 15t)\phi(t)$$

Table 9.1 Special Integrals

The equations of 9.32 may be integrated by expanding the term $(t\pm\alpha)^n$ as a polynomial in t. The coefficients will depend upon α, and whether the sign is plus or minus. By using the table of special integrals, the equations for each ROC curve may be reduced to a term involving the

normal distribution function, and a polynomial times the normal density function. The values of the polynomial times the normal density function at the upper limit, plus infinity, is zero. The value for the normal distribution at the upper limit is unity. Therefore, the upper limit will contribute at most a constant term. The rest of the equation for the ROC curve will depend upon the evaluation of the indefinite integral at the lower limits, $t \pm \alpha$. The resulting form for the ROC curve will be

$$Y = C_1 [1 - \Phi(t - \alpha)] + P_{n-1}(t - \alpha) \phi(t - \alpha)$$

$$X = C_0 [1 - \Phi(t + \alpha)] + Q_{n-1}(t + \alpha) \phi(t + \alpha) \qquad (9.33)$$

$$\alpha > 0 \ , \ t > 0$$

The ROC curve and ROC character have been graphed for the simplest case, when n is one. The ROC character is shown in Fig. 9.3. The lower scale is the dummy parameter t. Five z scales are shown. As the parameter α increases, the ROC character will be effectively spread over much more of the z axis. The mode of the ROC character is negative, and effectively moves farther to the left as detectability increases. The equations for the ROC curve in terms of the normal distribution and density for $n = 1$ are given below

$$n = 1 \quad \begin{aligned} Y &= \left(\alpha[1 - \Phi(t - \alpha)] + \phi(t - \alpha) \right) \Big/ [\alpha \Phi(\alpha) + \phi(\alpha)] \\ X &= \left(-\alpha[1 - \Phi(t + \alpha)] + \phi(t + \alpha) \right) \Big/ [-\alpha \Phi(-\alpha) + \phi(\alpha)] \end{aligned} \qquad (9.34)$$

The ROC curves for $\alpha = 1$ and 2 are displayed on normal-normal
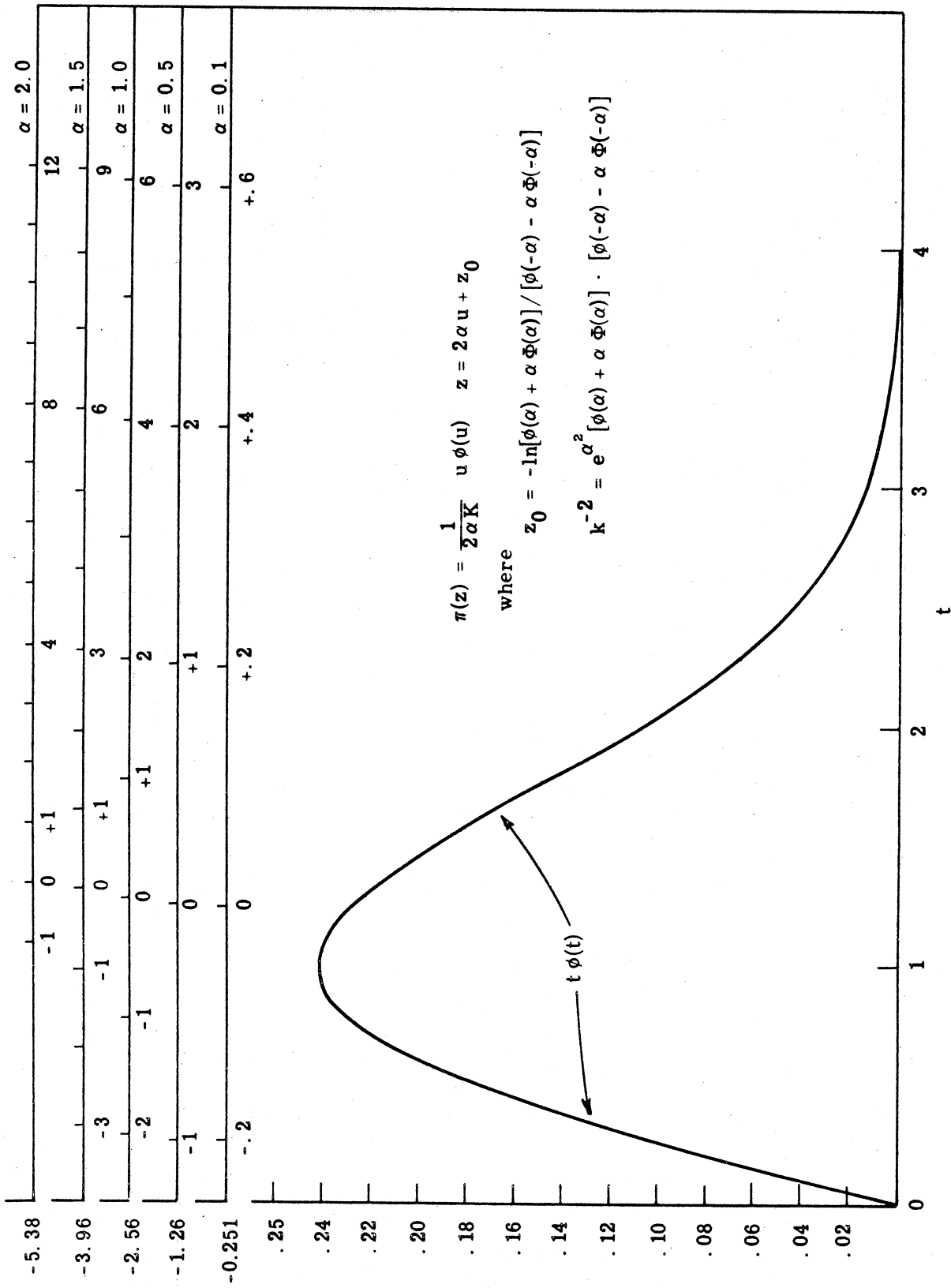
Fig. 9.3. H-Type ROC character Z scales and shape of ROC character

$$\pi(z) = C\left(\frac{z-z_0}{2\alpha}\right) e^{-\frac{1}{2}\left(\frac{z-z_0}{2\alpha}\right)^2} \quad z \geq z_0$$
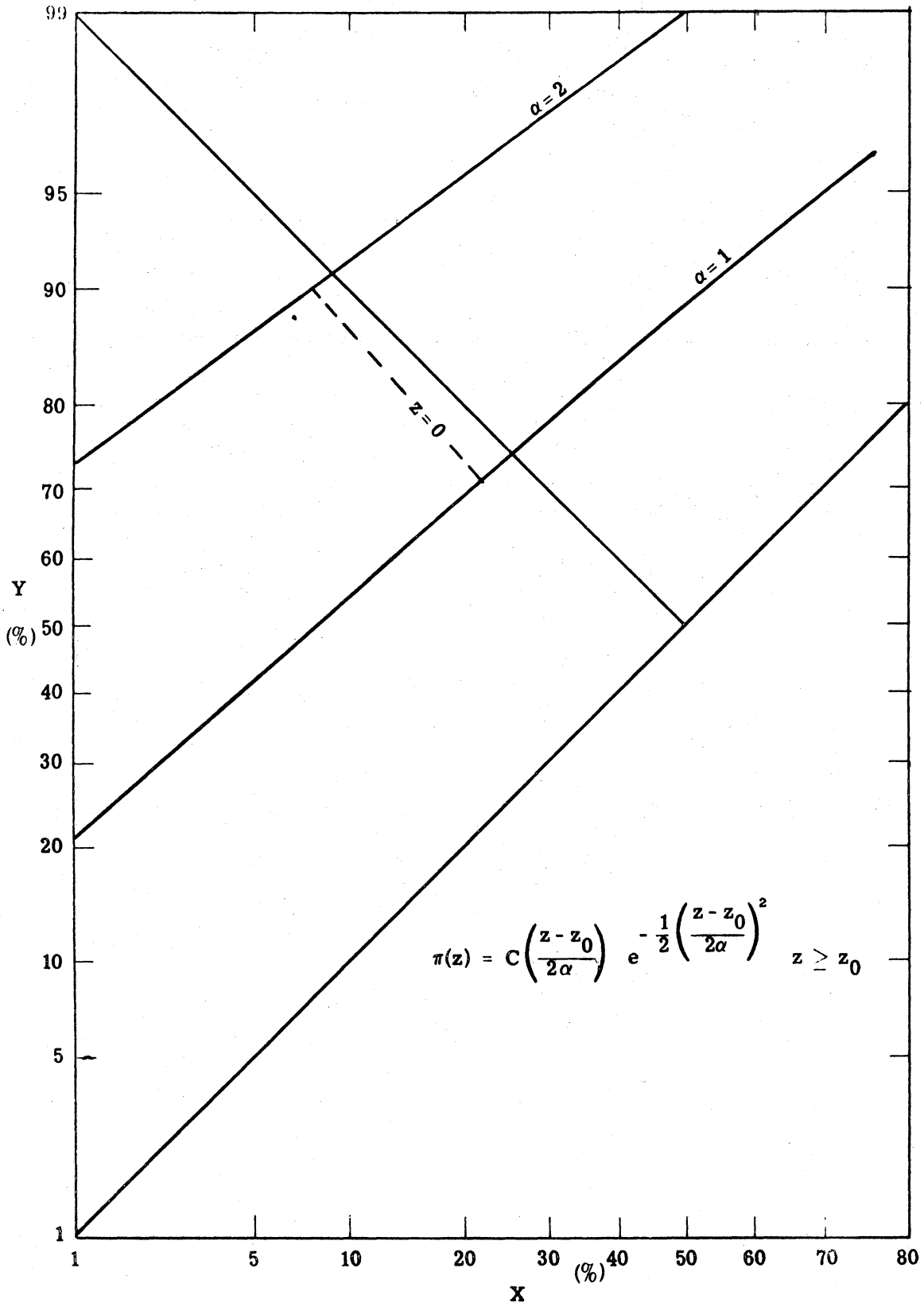
Fig. 9.4. H-type ROC, n = 1

paper in Fig. 9. 4. They appear to be nearly straight lines, with slope less than one. For $\alpha = 1$, the slope is $s = .85$ and for $\alpha = 2$, the slope is .74.

9. 4 Discussion on Approximations

The addition theorem may be used to advantage to obtain an ROC curve, for an ROC character by fitting the given character with a weighted sum of characters with known ROC curves. If the nature of the tail portion of the given character is known precisely, it may be matched with ROC characters with similar tail behavior. The Pearson III class, which includes the simple exponential of the power ROC, all exhibit a linear exponential tail. The conic ROC class also has a linear exponential tail, with coefficient fixed at -1. 5. The H-type ROC character, as well as the normal and metastatic normal, have a quadratic exponential tail behavior.

The field of approximations has been barely studied in the present work. It has been the purpose of this present work to provide the relation between the ROC curve and an analytic form, the ROC character, and to present a variety of families of ROC curves. The author intends to expand the ideas presented in this chapter to further enrich the collection of ROC characters with known ROC curve equations.

The area of obtaining ROC curves and characters from experimental data is worthy of research. Especially useful would be the development of techniques for successive approximations, whereby

limited data could be used to obtain a first ROC curve and character, and subsequent data could be used to refine the first approximation.

CHAPTER X

SELECTED TOPICS

10.1 Character-Free Measures of ROC Quality

When the description and evaluation of an ROC curve is just one part of an experiment, or one part of a mathematical analysis, there is a strong practical demand to reduce the description of detection performance from the single function ROC to a single number. When the purpose of the detection is well defined, this can be done by stating the degree to which the detection performance achieved its purpose. For example, in binary communications operating with a symmetric channel, the likelihood ratio cut is uniquely established at one, and the probability of error is the single numerical quantity of interest.

One of the major purposes of this present work was to reduce the description of the ROC curve to a two-step description. The first step is the determination of the family to which the ROC curve belongs. The second step is the evaluation of the one or two parameters necessary to describe the particular curve. The specific parameters to be evaluated depends on the nature of the character for the particular ROC family.

In this section various measures of ROC quality are considered that have been proposed for use independent of the particular character of the ROC curve. We call such measures

"character-free." The definition of the measure is independent of the character of the ROC. However, the values obtained for such measures will depend on the particular character of the ROC.

10.1.1 D. M. Green's Theorem. A particular type of psychophysical experiment is known as the multiple presentation, multiple alternative, forced choice experiment, or simply, "forced choice" experiment. The description which will be given here is the forced choice in time experiment. There are other forms of experiment with equivalent analysis. In the two alternative forced choice experiment, two intervals in time are marked for the observers attention. A signal will appear in one but not both of these time intervals. The observer is required to decide which interval contains the signal. In the general M-alternative forced choice situation, M intervals are marked off, and the signal will appear in one and only one of these intervals. A completely symmetric forced choice experiment presents the signals in each interval with equal probability, and the subject is asked to obtain the maximum number of correct identifications. The single number score for his performance is the fraction of identifications that are correct. The corresponding analytical measure for a theoretical situation is the maximum probability of correct identification.

Green's Theorem (Ref. 29) relates the score for a two-alternative symmetric forced choice experiment to the ROC curve.

Experimentally, the ROC curve is obtained in a single presentation two-alternative situation. Green's Theorem is that the two-alternative forced choice score will be the area under the ROC curve plotted on linear paper.

$$P_2(C) = \text{Area under the ROC curve} \qquad (10.1)$$

A general formulation for the probability of correct identification for the M alternative symmetric forced choice situation in terms of the ROC curve can be readily obtained. Consider an observer with a given ROC curve which measures his ability to distinguish between signal in noise, SN, from a background of noise alone, N. If this ROC curve is regular, then it corresponds to any one of a number of decision axes. Consider any such decision axis, and label it t. When the cause is N, denote the decision axis value by t_N; when the cause is SN, denote the random variable by t_{SN}. Consider the observation upon which the subject bases his identification. It consists of M-1 values of the decision axis due to N, and only one value resulting from SN. To maximize the number of correct identifications, the subject should choose the largest value of t as corresponding to the interval with the signal in it. The probability that he will be correct is the probability that the value of t due to SN, t_{SN}, is greater than all of the other observed decision axis values.

$$P_M(C) = \text{Prob (all M-1 } t_N < t_{SN}) \qquad (10.2)$$

Assume for convenience that t has a probability density function

under the condition SN; then $P_M(C)$ can be written in terms of an

integral as follows: The probability of correct identification condi-

tional to any specific value of t_{SN} is the product of the probabilities

that each of the t_N is below the t_{SN} value. This individual proba-

bility is the distribution function for t under condition N. For the

symmetric presentation case, all t_N variables have the same distri-

bution. $P_M(C)$ is obtained by averaging over all possible t_{SN} values.

$$P_M(C) = \int_{-\infty}^{\infty} F^{M-1}(t \mid N) \, f(t \mid SN) \, dt \tag{10.3}$$

It is convenient to convert this integral to the more general Stieltjes

integral.

$$P_M(C) = \int_0^1 F^{M-1}(t \mid N) \, dF(t \mid SN) \tag{10.4}$$

No matter which of the decision axes we hypothesized for the observer,

the distribution function under N is simply $1 - X$; the distribution

function under SN is $1 - Y$.

$$P_M(C) = \int_0^1 (1 - X)^{M-1} \, d(1 - Y) \tag{10.5}$$

A change of variables from $1 - Y$ to Y involves a minus sign which

is absorbed in reversing the limits of integration.

$$P_M(C) = \int_0^1 (1-X)^{M-1} \, dY \qquad (10.6)$$

Setting M equal to two in Eq. 10.6 yields the two expressions

$$P_2(C) = \int_0^1 (1-X) \, dY = 1 - \int_0^1 X \, dY \qquad (10.7)$$

Interpreting the integral as the area under the curve of the integrand,

$$P_2(C) = \text{Area under ROC curve}$$

$$= 1 - \text{Area above ROC curve} \qquad (10.8)$$

Another integral which is equivalent, since it also computes the area under the ROC curve, is

$$P_2(C) = \int_0^1 Y \, dX \qquad (10.9)$$

Green's Theorem gives a strong geometric relation, easily understood, which relates two psychophysical experiments. It also emphasizes the insensitive nature of the forced choice score on the detail structure of the ROC curve.

The two-alternative measures applied to the Luce ROC curve with vertex (x_1, y_1), to the normal ROC curve with index d,

and to the exponential ROC curve with index A, are given in Eqs.

10. 10, 10. 11, and 10. 13. An excellent approximation to the M-alter-

native probability for the normal ROC curve is given in Eq. 10. 12

(Ref. 17).

$$\text{Luce } (x_1, y_1): \quad P_2(C) = .5 + .5(y_1 - x_1) \tag{10.10}$$

$$\text{Normal, d:} \quad P_2(C) = \Phi(\sqrt{.5d}) \tag{10.11}$$

$$\text{Normal, d:} \quad P_M(C) = \Phi(a_M \sqrt{d} - b_M), \quad \Phi(-b_M) = \frac{1}{M} \tag{10.12}$$

$$\text{Exponential (Power), A:} \quad P_2(C) = \frac{A}{A+1} \tag{10.13}$$

Green's Theorem related the probability of correct

identification in the two-alternative test to the ROC curve. Every

quantity related to the ROC curve is also related to the ROC character.

The following theorem will establish the corresponding relationship

between $P_2(C)$ and the ROC character.

Any real L_2 function has an autocorrelation function.

The form chosen for the (unnormalized) autocorrelation function of the

ROC character is

$$\phi_{\pi\pi}(\tau) = \int_{-\infty}^{\infty} \pi(z + .5\tau) \, \pi(z - .5\tau) \, dz \tag{10.14}$$

It will be shown that the two-alternative probability of correct identi-

fication is the area under the autocorrelation on the positive side of

zero, weighted with an exponential factor which emphasizes large

values.

Theorem:

$$P_2(C) = \int_0^\infty e^{-.5\tau} \phi_{\pi\pi}(\tau) \, d\tau \qquad (10.15)$$

Proof:

Equation 10.9 is the starting point. In order to apply

the integration, the explicit relation between corres-

ponding values of **X** and **Y** is needed.

$$y(\beta) = \int_\beta^\infty e^{+.5z} \pi(z) \, dz, \quad -dx(\beta) = e^{-.5\beta} \pi(\beta) \, d\beta \qquad (10.16)$$

When these are formally utilized in Eq. 10.9, the minus

sign on the differential for x disappears in the Jacobian

of transformation.

$$P_2(C) = \int_{-\infty}^\infty \int_\beta^\infty e^{.5z} \pi(z) \, dz \; e^{-.5\beta} \pi(\beta) \, d\beta \qquad (10.17)$$

$$P_2(C) = \int\int_{-\infty < \beta < z < \infty} e^{.5(z-\beta)} \pi(z) \, \pi(\beta) \, dz \, d\beta \qquad (10.18)$$

The region of integration is over the half-plane of those

values of z which are greater than β. Any alternative

description of the same half-plane can be used as well.

A change of variables which singles out $z - \beta$ has been selected.

$$z = u + .5\tau \qquad u = .5(z + \beta) \qquad \left| J \left(\frac{u, \ \tau}{z, \ \beta} \right) \right| = 1 \qquad (10.\ 19)$$

$$\beta = u - .5\tau \qquad \tau = z - \beta$$

The half-plane of integration is now described as the half-plane with positive τ, and u ranging from $-\infty$ to $+\infty$.

$$P_2(C) = \int_0^\infty \int_{-\infty}^\infty e^{+.5\tau} \ \pi(u + .5\tau) \ \pi(u - .5\tau) \ du \ d\tau \qquad (10.\ 20)$$

The integral over u is the autocorrelation function of the character.

$$P_2(C) = \int_0^\infty e^{.5\tau} \ \phi_{\pi\pi}(\tau) \ d\tau \qquad (10.\ 15)$$

The proof of the theorem is complete.

10. 1. 2 **W. W. Peterson's Theorem.** The basic relation between the moments of the likelihood ratio under the two distributions, N and SN, was first pointed out to this writer by **W. W. Peterson.**

Theorem!

$$E(\ell^n | N) = E(\ell^{n-1} | SN) \qquad n \geq 1 \qquad (10.21)$$

This is a direct result of the fundamental theorem that the likelihood ratio of the likelihood ratio is the likelihood ratio. The proof of Eq. 10.21 can be accomplished simply in one line by writing down the integral definition for the n-th moment. Factoring out one ℓ term and associating it with the differential yields the desired (n-1)st moment under condition SN.

$$\int \ell^n \, dF(\ell | N) = \int \ell^{n-1} \ell \, dF(\ell | N) = \int \ell^{n-1} \, dF(\ell | SN)$$

This theorem leads to a basic understanding of some calculations on likelihood ratio used to obtain measures of ROC quality. The average value of likelihood ratio under noise is unity because the average value of the constant one under condition SN is unity. The average value of likelihood ratio under condition SN is the second moment under condition N.

$$E(\ell | N) = 1 \qquad E(\ell^2 | N) = E(\ell | SN) \qquad (10.22)$$

This means that the variance of the likelihood ratio under condition N will be equal to the difference of the means under the two conditions.

$$\sigma^2(\ell | N) = E(\ell | SN) - 1 = E(\ell | SN) - E(\ell | N) \qquad (10.23)$$

The signal-to-noise ratio, snr, for a linear filter is defined as the square of the peak output shift caused by the presence of signal, compared to the average noise power. In similar fashion, many authors including the present one, will define the snr for any kind of device as the square of the shift in the expected value, divided by the variance. When applied to linear filters, for situations in which a signal is added to stationary noise, the definitions agree. In general detection theory, any receiver is considered a likelihood ratio receiver whose output is a fixed order-preserving transformation of likelihood ratio. Consider for the moment a receiver whose output is precisely the likelihood ratio of the input. The signal-to-noise ratio for this output is

$$(snr)_{\ell} = \frac{[E(\ell \mid SN) - E(\ell \mid N)]^2}{\sigma^2(\ell \mid N)} \qquad (10.24)$$

By virtue of Eq. 10.23, this signal-to-noise ratio is also equal to those quantities in Eq. 10.23. If this signal-to-noise ratio could be interpreted in usual engineering fashion, and if this likelihood ratio is the one that leads to a terminal decision, then an order of magnitude of the quantities involved can be inferred. In most engineering applications, the signal-to-noise ratio desired is of the order of ten, or a hundred, or a thousand, or greater. If the variance of the likelihood ratio under N is any of these numbers, while the expected value is

fixed at unity, then the distribution of ℓ must be skewed to the right with tremendous upper tail. A description based on only the first two moments is hardly an adequate description for such a statistic. Therefore internal consistency dictates that Eq. 10.24 be considered a valid description of ROC quality only when its value is smaller than one.

In an attempt to compensate for the highly skewed nature of the distribution of likelihood ratio for cases with large detection quality, W. W. Peterson presented the following argument. For a normal ROC curve, the measures of Eqs. 10.23 and 10.24 are all related to the index d in exponential manner.

$$\text{Normal ROC} \quad E(\ell \,|\, SN) = e^d \tag{10.25}$$

The index d is therefore

$$d = \ell n \left[1 + \sigma^2 \, (\ell \,|\, N) \right] \tag{10.26}$$

When a signal is added to white Gaussian noise, and applied to the matched linear filter, the distribution of likelihood ratio is log normal, and d agrees with the accepted value of signal-to-noise ratio. He therefore suggested that if one wishes to base a quality statement on the first two moments of likelihood ratio, that one should do so by computing the variance of ℓ under N and applying Eq. 10.26.

10.1.3 Comparison of Quality Measures. In this section four measures of detectability are applied to four types of ROC:

normal, rectangular character, two-line symmetric Luce character, and the exponential character.

Kulback, Leibler, and Jeffries (Ref. 31) have used a measure which they call "the divergence between hypotheses," or "the difference in mean information for decision per observation available from the hypothesis space." It is the difference between the means of the logarithm of the likelihood ratio.

$$J(1:0) = E(z|SN) - E(z|N) \qquad (10.27)$$

The second measure of performance will be found used in many papers in the physical theory of detectability, called output signal-to-noise ratio of a decision device. It is the difference of the means of the logarithm of the likelihood ratio, squared, divided by the variance of the log likelihood ratio.

$$(snr)_z = J^2(1:0)/\sigma^2(z|N) \qquad (10.28)$$

This form is used when dealing with a decision axis which is known to be a linear translation of the z variable, but in which the constant has been lost during the analysis manipulations. It is justified because the difference of the means removes any translational constant, and the division by the variance will remove any multiplicative factor. A third measure of detectability was suggested by W. W. Peterson.

$$d_{WWP} = \ln(1 + \sigma^2(\ell|N)) = \ln E(\ell^2|N) \qquad (10.29)$$

The fourth measure of detectability is the variance of the logarithm

of the likelihood ratio taken under either hypothesis.

$$\sigma^2(z \mid N) \qquad \text{or} \qquad \sigma^2(z \mid SN) \qquad\qquad (10.30)$$

The four ROC characters to be considered are:

Normal, d > 0: $\pi_N(z) = e^{-\frac{d}{8}} \frac{1}{\sqrt{2\pi d}} e^{-\frac{z^2}{2d}}$

Rectangular, R > 1: $\pi_R(z) = \frac{\sqrt{R}}{2(1+R)}$ $\qquad |z| < \ln R$

Symmetric Luce, a > 0: $\omega_\pi(z = \pm a) = (e^a + e^{-a})^{-1}$

Exponential, A > 0: $\pi_E(z) = \frac{1}{A-1} \left(\frac{1}{A}\right)^{\frac{1}{A-1}} e^{-\frac{A+1}{A-1}\frac{z}{2}}$ $\qquad z > -\ln A$

Table 10.1 lists expected values and exact equations for

the four measures of detectability. All four measures of detectability

are the same for the normal ROC curve. The four measures of de-

tectability are not the same for the other types of ROC curves. Let us

examine the approximate value of these measures at the two extremes

--when the detectability is very low, and when the detectability is high.

Item Character	Normal	Rectangular	Sym Luce	Exponential
$E(z \mid \frac{SN}{N})$	$\pm\,.5d$	$\pm\left[\dfrac{R+1}{R-1}\ln R - 2\right]$	$\pm\,2a\,\tanh a$	$\left.\begin{array}{l}(A-1) - \ln A \\[4pt] \left(\dfrac{A-1}{A}\right) - \ln A\end{array}\right\}$
$E(z^2 \mid \frac{SN}{N})$	$d + .25d^2$	$\ln^2 R + 8 - 4\,\dfrac{R+1}{R-1}\ln R$	$4a^2$	$\dfrac{1}{A(2-A)}\,; \quad A < 2$
$E(\ell^2 \mid N)$	e^d	$\dfrac{1}{3R}[R^2 + R + 1]$	$2\cosh 2a - 1$	$(A-1)^2/A$
$J(1\!:\!0)$	d	$2\,\dfrac{R+1}{R-1}\ln R - 4$	$4a\,\tanh a$	$\left.\begin{array}{l}(A-1)^2 \\[4pt] \dfrac{(A-1)^2}{A}\end{array}\right\}$
$\sigma^2(z \mid \frac{SN}{N})$	d	$4 - \dfrac{4R}{(R-1)^2}\ln^2 R$	$4a^2\,\text{sech}^2\,a$	$(A-1)^2$
$(snr)_z$	d	$\dfrac{[(R+1)\ln R - 2(R-1)]^2}{[(R-1)^2 - R\ln^2 R]}$	$4\sinh^2 a$	
d_{WWP}	d	$\ln[1 + R + R^{-1}] - \ln 3$	$\ln[2\cosh 2a - 1]$	$-\ln A(2-A),\quad A < 2$

Table 10.1. Exact formulae

Character	Index	Limiting Performance
Rectangular	R → 1	$\frac{1}{3}$ (ln R)^2
Sym Luce	a → 0	$4a^2$
Exponential	A → 1	$(A-1)^2$

Table 10. 2. Low detectability limiting formulae

All four measures differ in the way the limits are approached. However, for small signals all four measures agree. If one's interest in detectability is limited to a problem in which repeated, independent and similar observations will be taken, then any one of these four measures of individual detectability will be appropriate for determining the ultimate detectability. In predetermined observation procedures or in sequential observation procedures, the logarithm of the likelihood ratio is the sum of the individual z values. The resultant ROC curve for fixed size decision procedures will be approximately normal, with index d equal to the sum of the individual detection "d_i" values. For example, one may validly conclude that if a final decision is based on n individual decisions, each decision yielding an observable output with log likelihood ratio either +2a or -2a, then the final ROC will be nearly normal and the detectability index will be

$$d = 4n \, a^2.$$

We have demonstrated that each of these four measures applies to the normal ROC curve independent of the level of detectability,

and to other types of ROC curve (at least, for a small class of ROC curves) at low levels of detectability.

If one could similarly conclude that for asymptotically large levels of detectability all four measures were nearly the same, one would have great confidence in using any of them. However, for each character these measures will approach different limits and approach them at different rates.

Item/Character	Rectangular	Sym Luce	Exponential
$J(1:0)$	$2 \ln R - 4 \to 2 \ln R$	$\to 4a$	$\to A$
$\sigma^2(z \mid {}^{SN}_{N})$	$4 - 4\dfrac{\ln^2 R}{R-1} \to R$	$\to 16a^2 e^{-2a} \to 0$	$\left\{ {\to A \atop \to 1} \right\}$
$(\text{snr})_z$	$\to \ln^2 R$	$\to e^{2a}$	$\to A^2$
d_{WWP}	$\ln((1+R)/3) \to \ln R$	$\to 2a$	∞ for $A > 2$

Table 10. 3. High detectability limiting formulae

In summary, several measures of detectability have been examined, and have been shown to lead to common numerical evaluations when the detectability is low. The normal ROC curve is unique in that all four measures of detectability yield the same value, d, at all levels of detectability. When the ROC curves near perfection, the four measures of detectability become drastically different. Not only do the measures disagree for the same type of character, but the behavior of each measure is different for different characters.

Anyone using a single number characterization for the ROC curve level without mention of the form of the ROC character must be very careful to know the use that will be made of his one number in order to have an appropriate measure of detectability.

10.2 A Special Case of Type-II ROC

In certain psychophysical experiments two quality measures have been obtained simultaneously, one of which is an ROC curve (Ref. 14). The first such experiment known to this writer is an experiment of Egan's for articulation score testing (Ref. 32). A word was randomly selected from a known list of fifty words, mixed with noise, and presented to a subject. The subject responded to each utterance by selecting one of the words from the known set as the most likely transmitted. The novel part of the experiment was that the subject was asked to score himself; that is, to indicate which responses he felt were right and which responses he felt were likely to be wrong. An ROC curve was then obtained by plotting on the horizontal axis the proportion of words accepted as correct, which were, however, wrong, and on the vertical axis, the proportion of words accepted as correct, and indeed were correct. This was a measure of the listener's ability to detect his own correct responses in a background of incorrect responses. This ROC curve is a "type-two" ROC curve, the distinction being made that the subject was indicating something about his first response. The first response was measured with a percent

correct or articulation score.

In this section a somewhat simpler experiment leading to two responses from the device, or subject, is analyzed. Both responses will be evaluated on ROC curves. Consider any two-alternative single-presentation experiment as the primary experiment. As always, it is not important in ROC analysis that one causal hypothesis be background noise and the other a signal introduced into the noise; any two causes may be used as long as there is a correct response "A" corresponding to the cause SN, and alternative response "B" corresponding to the second cause, N. The response to the first detection experiment will be subscripted throughout by "sub one"; that is, the ROC curve will be Y_1 as a function of X_1, with log likelihood z_1, and ROC character $\pi_1(z)$. The second response, leading to a type-two ROC curve, will be a judgment on the part of the observer as to whether the first response was correct or incorrect. That is, the observer is asked to place his responses into two categories, the high category indicating those quite likely to be right, and the low category being those which are less likely to be right.

What is the observation on which the subject bases this second response? It is exactly the same observation on which he based his first response. The observation itself is not different; the type of response is different. The relevant probability densities will also be different. The only feature of the observation that one need consider is the log likelihood ratio z_1. Although this is the logarithm of the

likelihood ratio for the first decision, it may help to de-emphasize this nature of the observation; therefore replace z_1 by the simple letter "t" as the observation for the second decision.

The type-two ROC curve conditions are right, R, or wrong, W, corresponding to whether the first decision was right or wrong. Two ways the subject would be right would be detection, A · SN, and correct rejection, B · N. Two type-two ROC curves are to be developed, the first conditional to the response A, and the second conditional to the response B. This is done for three reasons; one, the subject knows the response; two, the experimenter or grader knows the response; and three, these conditions may be significantly different. In order to develop type-two ROC curves a specific point must be picked as operating point on the type-one ROC curve. Let β designate a cut in the z_1 axis, and the corresponding ROC point be (X_1, Y_1).

The type-two ROC curve conditional to the response "A" is relevant whenever the observation t is greater than the cut value β. The density function for t, conditional to being right, is the probability density of t, given the joint condition A · SN. Care is taken here to indicate the conditions, so that the proper normalization will be included.

$$t > \beta \qquad f(t|R) = f(t|A, SN)$$

$$= f(t, A|SN)/P(A|SN) \qquad (10.31)$$

$$= e^{+.5t} \pi_1(t)/Y_1$$

In similar fashion, the probability density for t, given that the response was wrong, is the probability density for t, conditional to the joint condition $A \cdot N$.

$$t > \beta \qquad f(t|W) = f(t|A, N)$$

$$= f(t, A|N)/P(A|N) \qquad (10.32)$$

$$= e^{-.5t} \pi_1(t)/X_1$$

The likelihood ratio relevant for the second type of response is

$$\ell(t) = f(t|R)/f(t|W) \qquad (10.33)$$

Using (10.31) and (10.32), Eq. 10.33 becomes

$$t > \beta \qquad z_2 = t + \ln(X_1/Y_1) \qquad (10.34)$$

The ROC character is

$$\pi_2(z|A) = \pi_1\left(z + \ln\frac{Y_1}{X_1}\right) \Big/ \sqrt{X_1 Y_1}$$

$$(10.35)$$

$$\beta + \ln\frac{X_1}{Y_1} < z < \max z_1 + \ln\frac{X_1}{Y_1}$$

The type-two ROC curve conditional to the response "A" is the meta-static ROC curve derived from the type-one ROC considered from the point $(0, 0)$ to the point (X_1, Y_1).

The type-two ROC curve conditional to the response "B" is similarly derived by considering t values less than β. The probability density for t conditional to a wrong response is the probability density for t, condition to the joint occurrence of the response B and the condition SN.

$$t < \beta \qquad f(t\,|\,W) = f(t\,|\,B, SN)$$

$$= f(t,\ B\,|\,SN) / P(B\,|\,SN) \qquad (10.36)$$

$$= e^{+.5t}\,\pi_1(t)/(1 - Y_1)$$

Similarly, the probability of t given the right response is

$$t < \beta \qquad f(t\,|\,R) = f(t\,|\,B,\ N)$$

$$= f(t,\ B\,|\,N) / P(B\,|\,N) \qquad (10.37)$$

$$= e^{-.5t}\,\pi_1(t)/(1 - X_1)$$

When the observable t is less than β, the relevant likelihood ratio is

$$\ell(t) = f(t\,|\,R)/f(t\,|\,W) \qquad (10.33)$$

The value of z_2 and character are

$$t < \beta \qquad z_2 = -t + \ln \frac{1-Y_1}{1-X_1} \qquad\qquad (10.38)$$

$$\pi_2(z \mid B) = \pi_1\left(-z - \ln \frac{1-X_1}{1-Y_1}\right) \Big/ \sqrt{(1-X_1)(1-Y_1)} \qquad (10.39)$$

$$-\beta + \ln(1-Y_1)/(1-X_1) < z < -\min z_1$$

$$+ \ln(1-Y_1)/(1-X_1)$$

This type-two ROC curve is the metastatic ROC curve obtained from the type-one ROC by considering the arc from (X_1, Y_1) to the extreme $(1, 1)$ point.

These two type-two ROC characters are both obtained from the type-one ROC character. Graphically, the type-one character is cut apart at the point corresponding to the operating point β, and the left portion reflected about the cut point, that is, the $-\infty$ tail swung around toward $+\infty$. The appropriate horizontal translations and vertical magnifications are applied to yield a legitimate pair of ROC characters.

It appears that three conditions are both necessary and sufficient to make the two resulting type-two ROC characters equal. The magnifications will have to be equal, so the original operating point must fall on the negative diagonal. The domains of z must be the same (and utilizing the first point that the operating point is on the negative diagonal), the cut point β must be zero. Third, and most

important, the original ROC character must be symmetric in order that the mirror image $\pi_2(z \mid B)$ be equal to $\pi_2(z \mid A)$.

For many reasons an analyst may wish to lump two conditional type-two ROC curves together as a single curve. This may be done formally by simply averaging the type-two ROC characters. These two characters may be averaged if and only if the corresponding cuts in the type-two responses are not conditional to the type-one response. That is, in reaching a type-two decision the observer must cut the z_2-axis at a single point, and not one cut conditional to the A response and a second cut conditional to a B response. The ROC character is zero outside of the ranges displayed in Eqs. 10.35 and 10.39. Using p and q to indicate $p = P(SN)$ and $q = P(N)$, the a priori probabilities for the type-one conditions, the lumped data type-two ROC curve will have character

$$\pi_2(z) = P(A) \, \pi_2(z \mid A) + P(B) \, \pi_2(z \mid B)$$

$$= \frac{pY_1 + qX_1}{X_1 Y_1} \, \pi_1 \left(z + \ln \frac{Y_1}{X_1} \right) + \frac{1 - pY_1 - qX_1}{(1 - Y_1)(1 - X_1)} \tag{10.40}$$

$$\cdot \, \pi_1 \left(-z - \ln \frac{1 - X_1}{1 - Y_1} \right)$$

For a symmetric case, where the type-one character is symmetric and the observer has selected the symmetric cut, $\beta = 0$, falling on the negative diagonal, the type-two ROC curve is

$$\pi_2(z) = \pi_1\left(z + \ln\frac{Y_1}{X_1}\right)\Big/\sqrt{X_1 Y_1} \quad , z > -\ln\frac{Y_1}{X_1} \qquad (10.41)$$

The experimental procedure determines if a specific decision device operates according to this optimum theory would require the responding device (subject) to operate at a fixed operating point for the type-one decision. Either a number of fixed operating points, or a rating scale, may be used to obtain the type-two ROC curves.

10.3 The Multiple Observer

The title, "Multiple Observer," comes from psychophysics. It refers to the situation where the final decision maker is aided by several subordinate decision makers. In the context used in this present work, the multiple observer has been generalized to include situations when the subordinate processors may or may not reach decisions themselves. There is no essential restriction to situations involving human observers. An electronic system with inputs from several types of sensors is a multiple observer.

Formally, the multiple observer is considered as a terminal decision maker whose input observation consists of a vector V containing the outputs of the n subordinate observers. The emphasis in multiple observer studies is usually on those situations for which the outputs of the subordinate observers are of different nature. However, all outputs are influenced simultaneously by the presence or absence of signal. The terminal decision maker may also be one

of his own subordinate observers.

Such situations are prevalent in our everyday life. A person who must reach a decision as to whether it will rain this afternoon or not, may listen to a weather forecast, talk to his friends, and observe the sky himself. However, he must reach the final decision "A", to act as if there will be no rain, or "B" to act as if there will be bad weather. He takes his own observation into account as well as those of his friends and weather specialists. He must process his own visual information in the appropriate manner, and combine that processed output appropriately with the information from his other sources.

We shall make two assumptions: first, that all subordinate observers are operating under the same condition, N or SN; second, that inputs to the final decision maker are statistically independent. Likelihood ratio theory yields what the proper processing by the final decision maker should be, and what his ROC character is. He should compute the logarithm of the likelihood ratio for each component of his input, and add these to obtain the logarithm of the likelihood ratio of his total observation. His ROC character is the convolution of the characters of the subordinate observers.

$$V = (t_1, t_2, \ldots, t_n)$$

(10. 42)

$$z(V) = \sum_{k=1}^{n} z_k(t_k) , \qquad \pi_z = \pi_1 * \pi_2 * \ldots * \pi_n$$

The likelihood ratio of a decision can be considered as easily as the likelihood ratio of any other kind of observable. If the k-th subordinate observer makes the decision "A", and his ROC operating point (X_k, Y_k) is known, then the likelihood ratio of the observation is the ratio of Y_k to X_k. This is the ratio of the likelihood of the decision "A" under the condition SN compared to the likelihood of the decision "A" under the condition N. Therefore, each subordinate observer who is a decision maker will behave as if his ROC is a Luce ROC.

$$t_k \in \{\text{"A"}, \text{"B"}\}, \quad z_k(\text{"A"}) = \ln(Y_k/X_k)$$
$$z_k(\text{"B"}) = \ln((1 - Y_k)/(1 - X_k))$$

<div align="right">(10.43)</div>

No matter how different the types of observations the subordinate observers make, no matter how different their ROC curves, and style of reporting are, the formal description of the multiple observer treats each of the subordinate observers similarly. That is, the terminal decision maker calculates the log likelihood ratio of the subordinate to know how to process his information, and considers the ROC character of the subordinate in order to determine his own ROC character.

While the problem of the final decision maker is fairly straightforward, the problem of the subordinate observer who must report a decision is much more difficult. The overall optimum behavior for the subordinate observer is to report his decision value,

his z value, to his superior. (This is obviously optimum since the superior could do no worse than process this decision value in the same way that the subordinate observer does.) The problem of selecting the optimum operating point, the cut value on the z axis, for the subordinate observer forced to make a decision is unsolved except for a few special cases. Several workers have suggested that the cut be chosen to maximize the Shannon information measure. Certainly further research is necessary to arrive at a completely satisfactory technique. It is hoped that the description of the final decision maker's performance in terms of his ROC character may provide a new approach; the subordinate observer may set his goal as the choice of the Luce ROC character which, convoluted with the other ROC characters, will yield the best final character.

An observer who reports a continuous decision axis is called a continuous observer, and one who reports a discrete number of decision values is called a discrete observer. The phrase "Luce observer" is reserved for a subordinate observer who reports a binary decision. The resulting ROC character for situations involving continuous observers, or discrete observers who report many distinct values, is best handled by direct application of convolution. In this section two examples will be presented. The first is a combination of one Normal and one Luce observer; this is treated numerically. The second example is of one exponential observer and one Luce observer; it is possible to treat this situation analytically.

Example: Normal Plus Luce. The normal ROC character, and the

ROC curve parameterized by the variable z are given in Eq. 10.44.

$$\pi_N(z) = e^{-\frac{d}{8}} \frac{1}{\sqrt{2\pi d}} e^{-\frac{z^2}{2d}}$$

$$X_N(z) = \Phi(-.5\sqrt{d} - z/\sqrt{d}) \qquad (10.44)$$

$$Y_N(z) = \Phi(+.5\sqrt{d} - z/\sqrt{d})$$

sible z values and the corresponding discrete character

values for a Luce observer are

$$z_A = \ln(y_1/x_1) \qquad , \qquad \omega_\pi(z_A) = \sqrt{x_1 y_1}$$

$$\qquad (10.45)$$

$$z_B = \ln(1-y_1)/(1-x_1) \ , \qquad \omega_\pi(z_B) = \sqrt{(1-x_1)(1-y_1)}$$

The convolution of these ROC characters is simple because only two

possible z values occur with the Luce observer. The convolution is

therefore the sum of two products.

$$\pi(z) = \omega_\pi(z_A) \pi_N(z - z_A) + \omega_\pi(z_B) \pi_N(z - z_B) \qquad (10.46)$$

One can obtain the ROC curve by one of two routes. The formal meth-

od is to multiply the ROC character by $e^{\pm .5z}$ and to integrate from

some point to infinity. The other is to verbalize the equation as fol-

lows: whenever the Luce observer reports "A" and the terminal

observer's cut value is z, the decision maker will say yes whenever the normal observer gives a z_N value greater than $z - z_A$. Conversely whenever the Luce observer reports "B", the normal observer must give a much stronger indication, namely, $z - z_B$, to generate a terminal "A" decision. Either route yields precisely the same equations.

$$Y(z) = y_1 \, \Phi(+.5\sqrt{d} + z_A/\sqrt{d} - z/\sqrt{d}) + (1-y_1) \, \Phi(+.5\sqrt{d} + z_B/\sqrt{d} - z/\sqrt{d})$$

$$X(z) = x_1 \, \Phi(-.5\sqrt{d} + z_A/\sqrt{d} - z/\sqrt{d}) + (1-x_1) \, \Phi(-.5\sqrt{d} + z_B/\sqrt{d} - z/\sqrt{d})$$

$$(10.47)$$

When the Luce observer is operating on the negative diagonal, the ROC curve and the ROC character are symmetric. This allows a slight simplification of the ROC equations and some further analytic work.

$$\text{Symmetric Luce:} \quad x_1 = 1 - y_1, \quad z_B = -z_A \qquad (10.48)$$

Whenever the normal observer is sufficiently superior to the Luce observer, the ROC character of Eq. 10.46 will have a single maximum at zero. Under the converse situation, when the Luce observer is sufficiently superior to the normal observer, the ROC character will be bimodal. The precise definition of "superior" depends on whether z_A is smaller or larger than the square root of d. When the Luce observer is superior, the twin modes of the ROC character

are given by the solution of Eq. 10.49.

$$z_A > \sqrt{d} \ , \qquad z = z_A \tanh \frac{z \, z_A}{d} \qquad (10.49)$$

This is demonstrated in Fig. 10.1, where the ROC characters for the final decision maker are plotted on semi-log paper. The continuous normal observer had an index of $d = 1$. Four Luce observers were considered. The characteristics of these four are listed in Table 10.4.

z_A	0	1.00	1.45	2.95
x_1	0.50	0.31	0.19	0.05
y_1	0.50	0.69	0.81	0.95

Table 10.4. Luce z_A for selected points

When the Luce observer operates at the $(.19, .81)$ point, the ROC character is bimodal, but on semi-log paper appears to be fairly flat topped. The ROC curves for this situation are plotted in Fig. 10.2. The Luce ROC is dotted, since the only point used is the single operating point. The ROC curve for this observer before he was forced to take a fixed position may have been anything going through the point $(.19, .81)$. In the neighborhood of the negative diagonal the combined Luce and normal observer is only slightly better than the Luce operating point. However, the combination is definitely better than either away from the Luce operating point. For
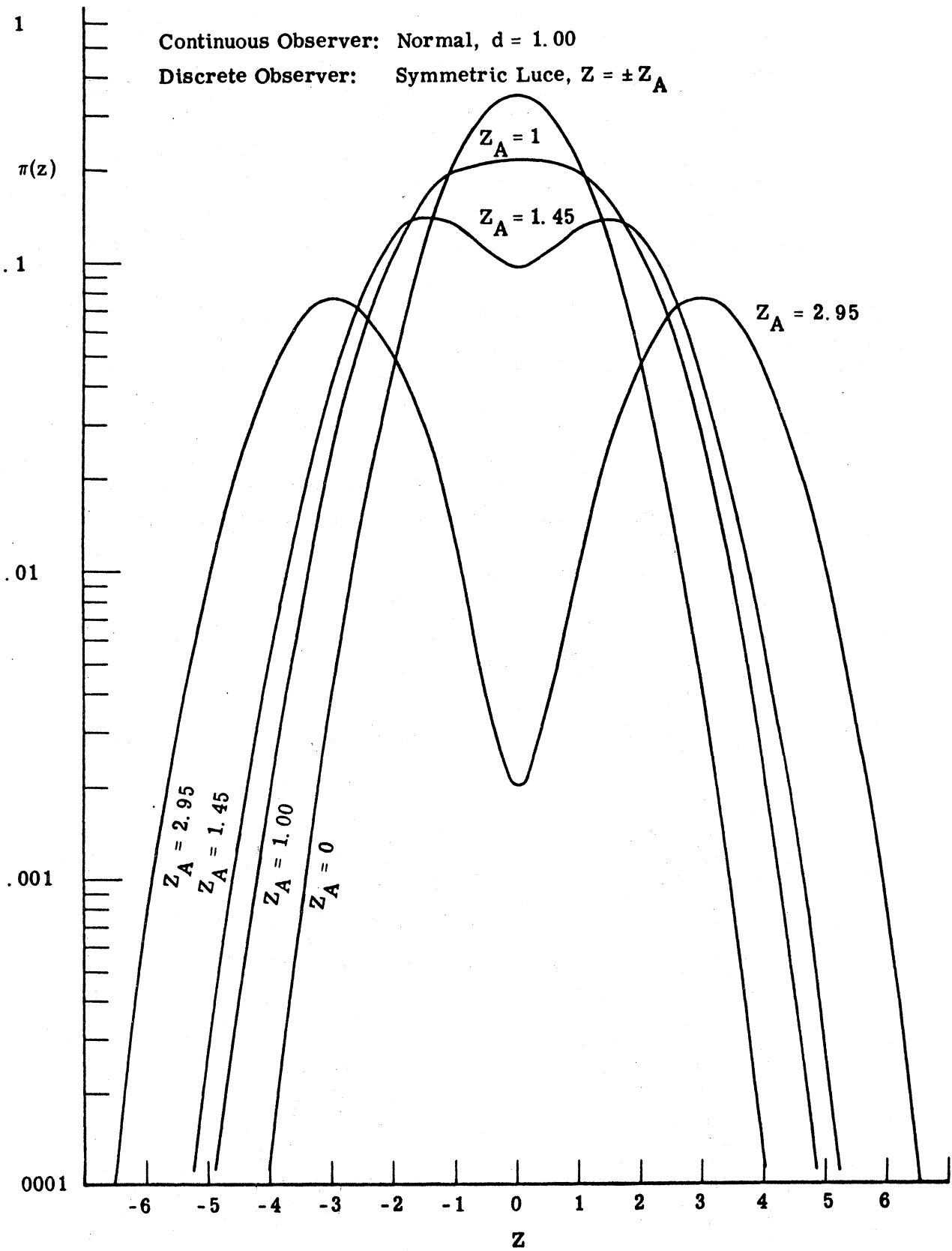
Fig. 10.1. ROC characters for four multiple observers
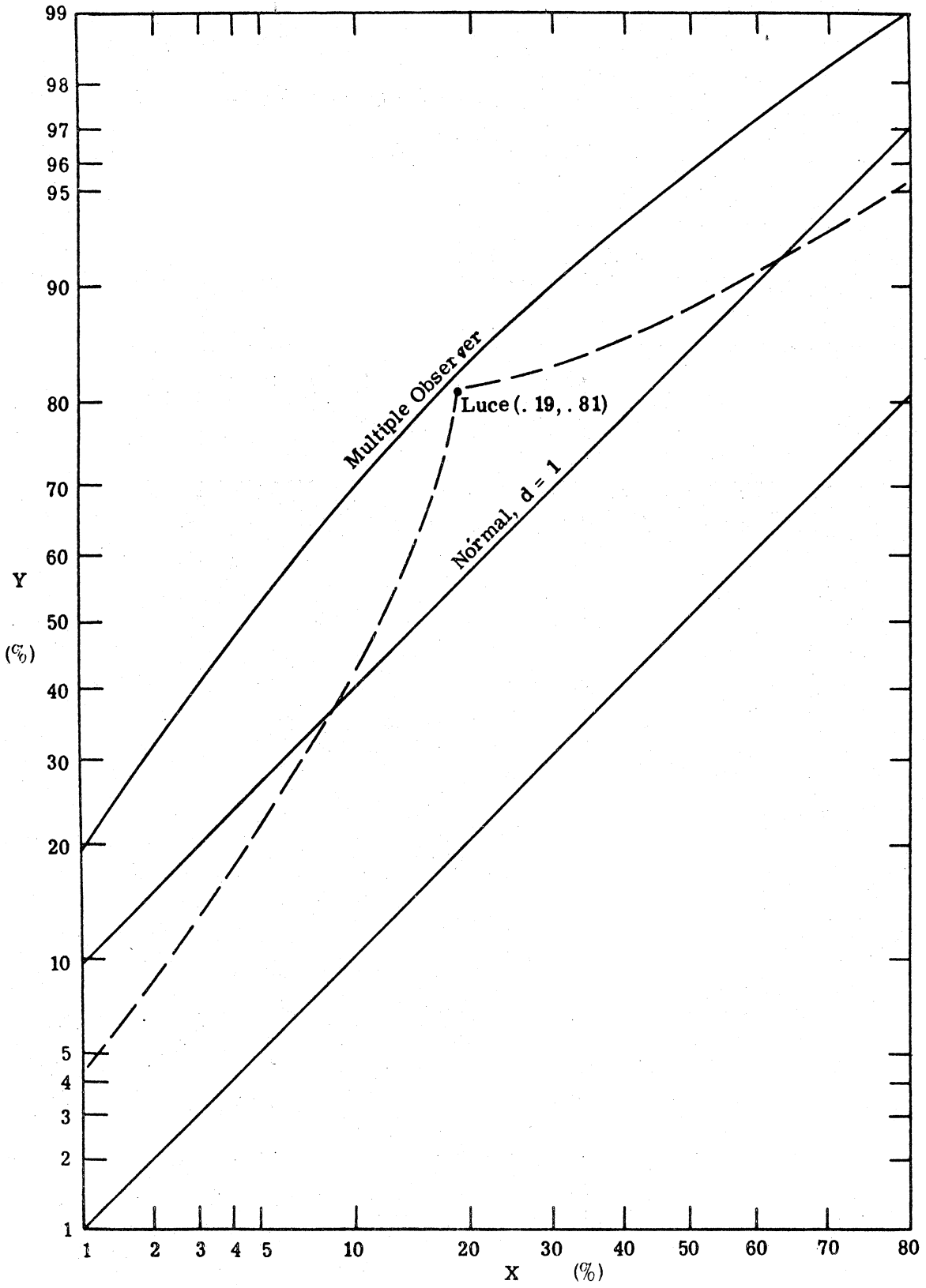
Fig. 10.2. Multiple observer ROC curve

comparison the normal character for the curve that would pass through the Luce operating point is drawn together with the multiple observer ROC character in Fig. 10.3(a). That part of the character associated with the ROC region appearing on one percent to ninety-nine percent normal-normal paper is also indicated in Fig. 10.3(a).

When the Luce observer is considerably better than his co-worker, the normal observer, the ROC character is strongly bimodal. This is best seen in Fig. 10.3(b), plotted on linear paper. This strongly suggests its origin as the sum of two displaced normal density functions. When the ROC curve is calculated, that region of the character affecting the part displayed on usual normal-normal paper is basically the part between the two modes. That is the concave section of the character, and much more like the extreme hyperbolas considered in Section 5.2 than like the normal character. The corresponding ROC curve is plotted in Fig. 10.4. This ROC displays much more of a "corner" than has been displayed before except for the Luce ROC. This occurs because of the relative minimum on the ROC character, which places very little probability on those z values near zero. As a result the slope drops from five ($z = 1.6$) to one ($z = 0$) with only a 0.01 change in false alarm probability.

If this Luce observer represents the operating point on a normal ROC curve, the index would be $\sqrt{d} = 3.29$. If a continuous normal character with this index is convolved with the $d = 1$ continuous normal character, the result would be a normal ROC with index
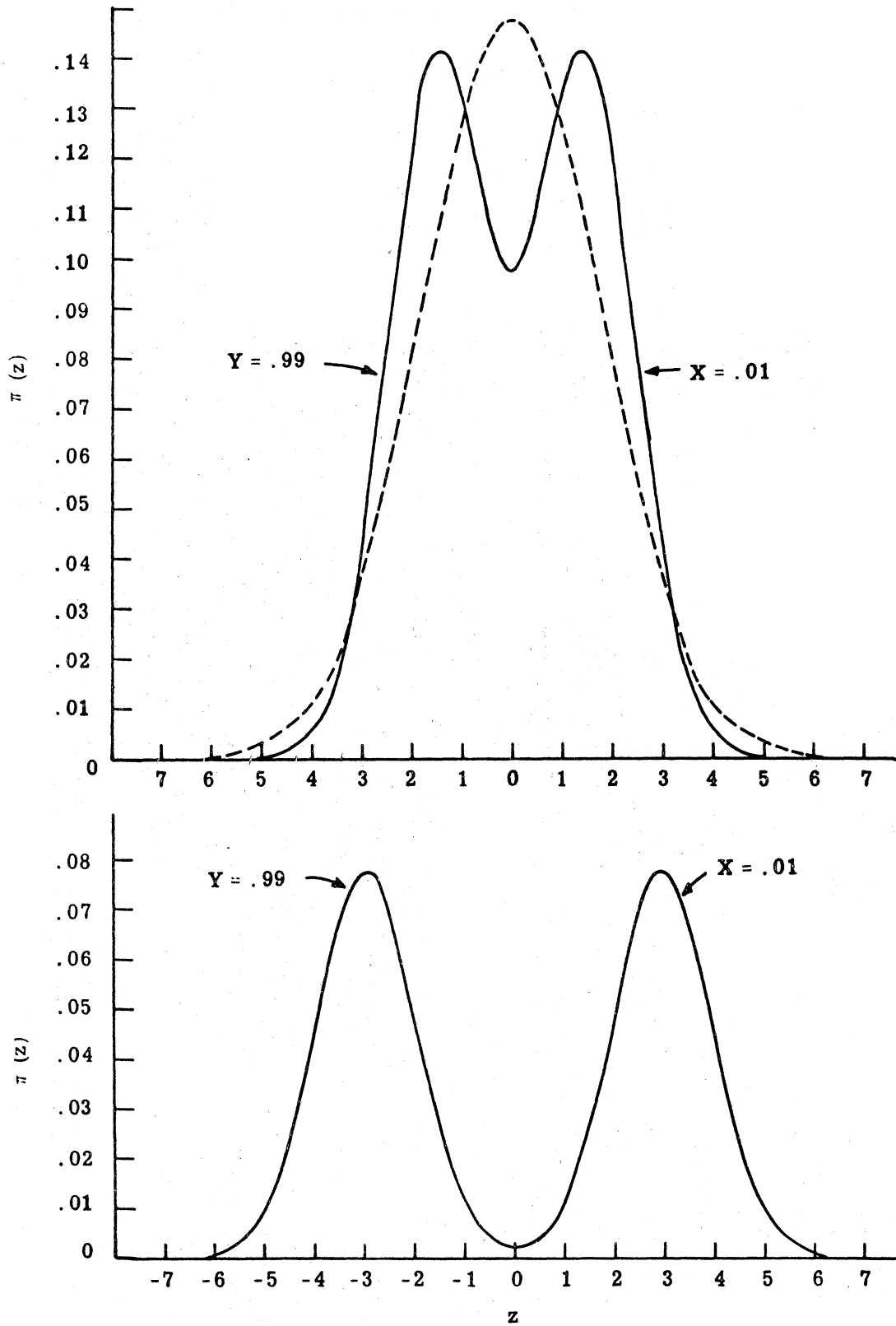
Fig. 10.3. Two multiple observer ROC characters
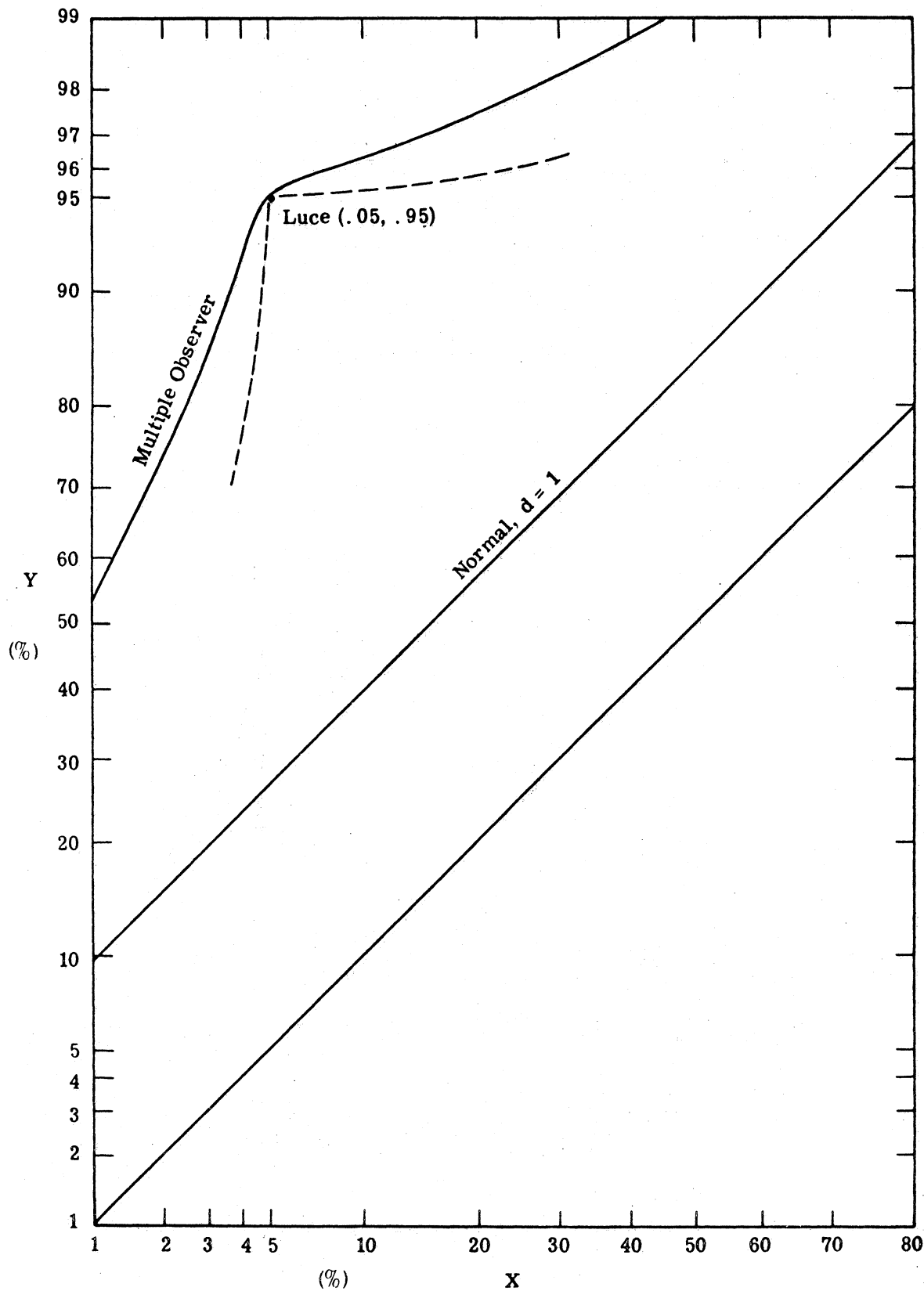
Fig. 10.4. Multiple observer ROC curve

$\sqrt{d} = 3.44$. The $d = 1$ observer is completely overshadowed by the better observer. However, when the better observer reports only a decision and not a continuous measure, the normal observer with $d = 1$ becomes important in improving the ROC curve away from the better observer's operating point.

Example: Exponential Plus Luce. The following work is restricted to considering a symmetric Luce character. Similar analysis can be performed with nonsymmetric Luce character, but the generality seems to obscure the nature of the result. The exponential character and corresponding ROC curve are

$$\pi_E(z) = \frac{\sqrt{A}}{A-1} e^{-\frac{A+1}{A-1}\frac{z-z_0}{2}} \qquad z > z_0 = \ln A$$

(10.50)

$$Y_E(z) = e^{-\frac{z-z_0}{A-1}}, \qquad X_E(z) = e^{-A\frac{z-z_0}{A-1}} = Y_E^A(z)$$

The terminal observer's character is formed by convolution

$$\pi(z) = \omega_\pi(z_A)\left[\pi_E(z - z_A) + \pi_E(z + z_A)\right]$$

(10.51)

Each of the exponential terms in this convolution can be simplified. Using the double notation these are considered simultaneously. Each i written explicitly in the exponential form of Eq. 10.50, and a common exponential factored from each.

$$\pi_E(z \mp z_A) = \frac{\sqrt{A}}{A-1} e^{-\frac{A+1}{A-1} \frac{z - z_0 \mp z_A}{2}} = e^{\mp \frac{A+1}{A-1} \frac{z_A}{2}} \pi_E(z), \quad z > z_0 \pm z_A$$

$$= 0 \qquad\qquad\qquad z < z_0 \pm z_A$$

$$(10.52)$$

Care must be taken to keep track of the lower limit of each of these equations, since they differ. Rewrite the ROC character as Eq. 10.53.

$$\pi(z) = \omega_\pi(z_A) \left[e^{-\frac{A+1}{A-1} \frac{z_A}{2}} + e^{\frac{A+1}{A-1} \frac{z_A}{2}} \right] \pi_E(z) \qquad z > z_0 + z_A$$

$$= \omega_\pi(z_A) \left[0 + e^{\frac{A+1}{A-1} \frac{z_A}{2}} \right] \pi_E(z) \qquad z_0 - z_A < z < z_0 + z_A$$

$$= 0 \qquad\qquad\qquad z < z_0 - z_A$$

$$(10.53)$$

The important feature of this ROC character is that in each of the two active regions, the character is proportional to the original exponential character. The coefficients of proportionality depend on the quality of the Luce observer and the index for the exponential observer, as do the ranges. However, they are constants as far as the ROC character is concerned. Let us emphasize this constant nature by using the two symbols, B and C.

$$\pi(z) = B \pi_E(z) \qquad z - z_0 > z_A$$

$$= C \pi_E(z + z_A) \qquad |z - z_0| < z_A \tag{10.54}$$

Integrate Eq. 10.54 formally.

$$X(z) = B X_E(z) \qquad Y(z) = B Y_E(z) \qquad z > z_0 + z_A \tag{10.55}$$

Use Eq. 10.55, and the power law relation on the original exponential ROC curve, to obtain a power law relation for this ROC also.

$$X(z) = B Y_E^A(z) = B^{1-A} Y^A(z) \qquad z > z_0 + z_A \tag{10.56}$$

This is not a pure power ROC. This equation only holds for the initial part of the ROC curve. To obtain the remainder of the ROC curve integrate the second part of Eq. 10.54. This is most simply done by integrating in the direction opposite the usual, by integrating from the lower limit up to some value of z, and obtaining the probability of miss and the probability of correct rejection.

$$1 - X(z) = C \left[1 - X_E(z + z_A) \right], \qquad 1 - Y(z) = C \left[1 - Y_E(z + z_A) \right]$$

$$|z - z_0| < z_A \tag{10.57}$$

The expression for the original power law is used to simplify (10.57) to

$$\left[1 - \frac{1 - X(z)}{C}\right] = \left[1 - \frac{1 - Y(z)}{C}\right]^A \qquad |z - z_0| < z_A \qquad (10.58)$$

Evaluate B and C by comparing Eqs. 10.53 and 10.54.

$$B = \cosh\left(\frac{A+1}{A-1} \frac{z_A}{2}\right)\Big/\cosh\left(\frac{z_A}{2}\right) \quad , \quad B > 1 \qquad (10.59)$$

$$C = \omega_\pi(z_A) = .5\Big/\cosh\left(\frac{z_A}{2}\right) \quad , \quad C < .25 \qquad (10.60)$$

The constant B is always greater than one, and is much greater than one when the Luce observer is far superior to the original exponential observer. The portion of the ROC curve beginning at the origin is a modified power curve. It has the same power as the original exponential. However, the false alarm rate will be less. To obtain the numerical idea of how much less, consider the power $A = 2.5$, which is in the neighborhood of the normal $\sqrt{d} = 1$. For $z_A = 1.45$, $B = 2.22$, and the false alarm ratio between original exponential and final modified power is $.302$. When $z_A = 2.95$, $B = 6.79$, and the false alarm factor is $.063$.

The primary purpose of these examples has been to show how the multiple observer can be treated by using the ROC character. Examples could be similarly given to show how a completely discrete set of subordinate observers could be quickly convolved, and the resulting polygon ROC curve calculated. A second purpose has

also been to show the role of a relatively inferior continuous observer operating in conjunction with a superior discrete observer.

10.4 Rating Scale Pay-Off Matrices

This section deals with experimental techniques used to obtain ROC curves from subjects who can be motivated by reward and punishment. In the basic procedure for determining the observer's ROC curve (the yes-no experiment), a single decision is obtained from the observer on each observation. An operating point is determined by making a sufficiently large number of observations to obtain the accuracy desired. In order to obtain the shape of the ROC curve, the observer must repeat this observation-decision experiment, operating at a number of different operating points along the ROC curve. Instructions by the experimenter can be used to motivate the observer to shift his decision mechanism to obtain these different points. Another technique that may be used is to establish a bonus system, or pay-off matrix, for the experiment. The subject is given a four-entry table which shows him exactly how much he will be rewarded for each correct decision, and how much he will be fined (or, rewarded less) for each incorrect decision. The observer also knows the a priori probability of each of the conditions N and SN. He is assumed to operate in a manner which will tend to maximize his total bonus. The true motivating utility to the observer may not be precisely what the pay-off matrix describes, since he may

take into account many other factors in choosing his operating point.

However, when the pay-off matrix is changed, almost all observers

will respond with a change in the operating point that is in the same

general direction as the operating point for an observer motivated

solely by the pay-off matrix. The bonus system also gives a form

of feedback to the observer, who can try to adjust his own internal

description of a cut on a decision axis so as to maximize the bonus.

This type of feedback may tend to stabilize the experimental situation

and aid in the repeatability of particular operating points.

This procedure for determining an ROC curve point by

point can take considerable time. A procedure which might establish

four or six or even fifty operating points on an ROC curve in a single

experimental session is a tremendous boon. Such a procedure is

the rating technique. In the fixed category rating technique, the ob-

server may make one of a fixed number of responses. One end of

the response scale indicates "A" with a great deal of confidence, and

the other extreme represents "B" with a great deal of confidence,

while intermediate ratings represent intermediate preferences for the

"A" and "B" decisions. This experimental technique, a discussion

of its advantages, and experiments verifying that it yields essentially

the same results as the longer yes-no procedure are given by Egan,

Schulman, and Greenburg (Ref. 34, and Ref. 8, p. 172.). They

used a four-category technique. Watson and Rilling (Ref. 33) have

used an essentially continuous rating technique to obtain ROC curves.

In a study by Swets, Tanner and Birdsall (Ref. 26, also Ref. 8, page 130) a six-category rating scale was used to obtain an ROC curve. However, the intention of that experiment was to determine more than just the ROC curve. The observers were instructed as to the meaning of each category. Specifically, they were informed that the highest category should mean an a posteriori probability of signal presence of roughly ninety-five percent or better, that the next lower category must indicate between eighty and ninety-five percent a posteriori probability, etc. The technique was satisfactory in obtaining a six point ROC curve, but the experimenters were basically dissatisfied with the relation of the responses to the assigned a posteriori categories.

In personal communications, Tanner has argued that the pay-off matrix, with its monetary entries, is a simple, clearly understood, and efficient means of explaining the experimenter's desires so that the observer can understand them. This is not a reflection on the observer's intelligence, but on his experience. The pay-off matrix deals with quantities more familiar to the observer than such quantities as a posteriori probability, likelihood, "strict" versus "medium" criterion, or "nine percent false alarm rate." A bonus system also affords the observer a specific measure that he can check after an experimental run, to see to what extent he is holding a particular criterion.

The purpose of this section is to present a type of pay-off matrix to use in conjunction with a continuous rating scale report by the observer. It also presents a formula for tailoring this rating scale pay-off so that the ideal observer, "motivated solely by the pay-off matrix," would report a rating proportional to any one of a number of functions which the experimenter specifies.

The pay-off matrix for a standard yes-no experiment, and the appropriate cut value on the likelihood ratio decision axis, is shown in Fig. 10. 5.

R: \ C: SN   N

A	V_D	V_{FA}
B	V_M	V_{CR}

$$\beta = \frac{P(N)}{P(SN)} \frac{V_{CR} - V_{FA}}{V_D - V_M}$$

Fig. 10. 5.   Fixed pay-off and ℓ-cut matrix

The only restriction on the values in the table is that the value of a correct response must be greater than the value of the corresponding incorrect response. Thus, when noise alone is present the value of a correct rejection must be greater than the value of a false alarm, when signal plus noise is present, the value of a detection must be greater than the value of a miss. This section will start with an example of one type of rating scale pay-off matrix and then generalize in two steps.

The observer's reporting procedure will be as follows.

He makes the customary binary decision "A", "B", and reports a magnitude u. If he is correct, he will be paid proportional to u. If he is incorrect, he will be penalized proportional to u^2. Thus, he will wish to make the value of his rating response, u, large in order to obtain a large reward, but this is done at the risk of a very much larger penalty. In addition to the description in terms of the observer's rating, u, a constant k is fixed by the experimenter to scale the positive pay-offs (to keep the experimenter from financial ruin) and a factor R is used to give a relative weight of the N pay-off to the SN pay-off.

$$k > 0 \quad \text{a fixed scale factor for positive pay-off}$$

$$R > 0 \quad \text{a fixed factor for N pay-off over SN pay-off} \qquad (10.61)$$

$$u > 0 \quad \text{magnitude of observer's rating response}$$

Using these constants the pay-off matrix for what will be called the ℓ-rating scale is given in Fig. 10.6.

R: \ C:	SN	N
A	$V_D = +uk$	$V_{FA} = -.5u^2 R$
B	$V_M = -.5u^2$	$V_{CR} = uk R$

Fig. 10.6. Pay-off matrix for ℓ-rating scale

Consider for a moment any fixed value for u. The proper cut-off

on the likelihood ratio axis can be obtained by using the pay-off ma-

trix of Fig. 10.6 in the optimum cut equation of Fig. 10.5.

$$\beta = \frac{P(N)}{P(SN)} \frac{ukR + .5u^2R}{uk + .5u^2} = \frac{P(N)}{P(SN)} R \qquad (10.62)$$

The cut on the likelihood ratio axis is independent of the value of the

rating, u. Therefore, before the observer has chosen the magnitude

of rating for a particular response, he may determine whether the

"A" or "B" decision is appropriate.

In the analysis the sole motivation of the observer is to

maximize his average pay-off. Whenever his response is "A", he will

wish to choose the rating value u to maximize the expected pay-off

for that particular single response. This value is given in Eq. 10.63.

$$\ell > \beta \qquad V_A(\ell) = P(SN|x) uk - .5u^2R P(N|x) \qquad (10.63)$$

The a posteriori probability appears in this equation because the

choice of the rating will be made after the observation x. The rat-

ing u is the only variable in this equation. The proper u value is

obtained by differentiating the value V. The first and second deriva-

tives of the value with respect to u are

$$\frac{\partial V_A(\ell)}{\partial u} = P(SN|x) k - uR P(N|x) \qquad (10.64)$$

$$\frac{\partial^2 V_A(\ell)}{\partial u^2} = -R \ P(N|x) \tag{10.65}$$

Since the second derivative is everywhere negative, the value has a single maximum at the zero in the first derivative. From Eq. 10.64, the derivative is zero whenever

$$u = \frac{k}{R} \ \frac{P(SN|x)}{P(N|x)} = \frac{k}{R} \ \frac{P(SN)}{P(N)} \ \ell(x) = \frac{k}{\beta} \tag{10.66}$$

The constants β and R are related to the a priori probability ratio, and are used to eliminate this a priori probability from the final expression in Eq. 10.66. The same type of analysis is used to determine the proper magnitude of the rating response when the decision response has been "B".

$$\ell < \beta \qquad V_B(\ell) = P(N|x) \ ukR - .5u^2 \ P(SN|x) \tag{10.67}$$

$$\frac{\partial V_B(\ell)}{\partial u} = P(N|x) \ kR - u \ P(SN|x) \tag{10.68}$$

$$\frac{\partial^2 V_B(\ell)}{\partial u^2} = -P(SN|x) \tag{10.69}$$

$$u = k R \ \frac{P(N|x)}{P(SN|x)} = k R \ \frac{P(N)}{P(SN)} \ \frac{1}{\ell(x)} = \frac{k\beta}{\ell} \tag{10.70}$$

Let us summarize the effect of this type of pay-off matrix on the observer who is motivated solely by maximizing his

expected return. His binary decision is the same as with a fixed
pay-off matrix with value ratio equal to R. Whenever his response
has been "A", his rating is directly proportional to the likelihood ra-
tio. Whenever his decision has been "B", his rating response is
inversely proportional to the likelihood ratio. The constants of pro-
portionality are established by the experimenter.

What can be learned from such an experimental tech-
nique? As long as the observer's rating response is order-preserv-
ing with likelihood ratio, a valid ROC curve can be obtained with this
method as with any rating scale method. If enough observations taken
under suitably stable conditions have been used to obtain this ROC
curve, an analytic fit can be used to approximate the curve. From
the analytic fit the distribution function can be obtained for the likeli-
hood ratio, as in any of the analytic models in this present work.
These distributions can be compared with the distributions of the
observer's u values. The experimenter can thus obtain informa-
tion about the degree to which the observer has been able to match
his u-scale with the ℓ-scale.

The first step in generalization of this type of pay-off
matrix is to determine if an ideal observer can be forced to report a
rating of any other nature than one proportional to likelihood ratio.
In this generalization it may be necessary to restrict the range of the
observer's available rating responses. A function $r(u)$, a rating
forcing function, will play the same role as u^2 in the ℓ-rating pay-off

matrix.

$u \geq u_0$ magnitude of observer's rating response

$R > 0$ a fixed factor for N pay-off over SN pay-off

$r(u) > 0$ a smooth convex upward function for negative pay-offs

$$(10.71)$$

The four-fold pay-off matrix is shown in Fig. 10.7.

R: \ C:	SN	N
A	$V_D = +u$	$V_{FA} = -R\, r_0(u)$
B	$V_M = -r_0(u)$	$V_{CR} = +R\, u$

Fig. 10.7. Base pay-off matrix for general rating scale

At this stage in the development the rating forcing function $r(u)$ will be given some additional properties to make the solution unique. After this has been completed and a single solution obtained for each situation, the rating forcing function may be generalized to make it more appropriate for particular practical situations. These restrictions are given in Eq. 10.72.

$$r_0(u_0) = u_0 \, , \quad r_0'(u_0) = 1 \, , \quad r_0''(u) > 0 \, , \quad u \geq u_0 \qquad (10.72)$$

The first restriction is used so that the pay-off matrix will be "balanced" when the observer chooses the minimum rating value.

That is, the value for detection will be as much positive as the value for miss is negative, and the value for correct rejection as much positive as the value for false alarm will be negative. The second restriction will insure that a solution exists. The third will always hold, the function must be a convex upper function; this is the characteristic that restrains the observer from using extremely large scale values.

The first step is to determine a proper operating point by using the β equation of Fig. 10.5 for the pay-off matrix of Fig. 10.7.

$$\beta = \frac{P(N)}{P(SN)} \frac{R \ u + R \ r_0(u)}{u + r_0(u)} = \frac{P(N)}{P(SN)} \ R \tag{10.73}$$

As expected, the observer may determine the proper response threshold to cut between the "A" and "B" decisions. As before, the derivative of the value function conditional to the binary decision response yields the appropriate u value. The second derivatives are indeed negative and are not displayed here. The values' first derivatives and the equation controlling the optimum rating response are given in Eqs. 10.74 - 10.79.

$$\ell > \beta \qquad V_A(\ell) = P(SN|x) \ u - R \ r_0(u) \ P(N|x) \tag{10.74}$$

$$\frac{\partial V_A(\ell)}{\partial u} = P(SN|x) - R \ r_0'(u) \ P(N|x) \tag{10.75}$$

$$r_0'(u_A) = \frac{1}{R} \frac{P(SN|x)}{P(N|x)} = (\frac{\ell}{\beta}) \tag{10.76}$$

$$\ell < \beta \qquad V_B(\ell) = P(N|x) R u - r_0(u) P(SN|x) \tag{10.77}$$

$$\frac{\partial V_B(\ell)}{\partial u} = P(N|x) R - r_0'(u) P(SN|x) \tag{10.78}$$

$$r_0'(u_B) = R \frac{P(N|x)}{P(SN|x)} = (\frac{\beta}{\ell}) \tag{10.79}$$

Equations 10.76 and 10.79 appear to repeat virtually

the same situation as (10.66) and (10.70), namely, that some function

of the rating response be directly proportional to the ℓ over β ratio

under condition A and inversely proportional to the same ratio under

condition B. This is the basic restriction on this present technique.

The rating may be forced to follow some particular function of the

experimenter's choice only if the experimenter's function is in turn

a function of the ℓ to β ratio. However, in the study of decision

processes based on likelihood ratio there are a number of functions

of this type of interest.

One more step of generalization is called for before

displaying particular cases of rating scale forcing functions. The

pay-off matrix of Fig. 10.7 can be changed by the addition of a prop-

erly selected set of four numbers which will change the pay-off to

the observer but will not change any of the equations. That is, the

decisions and the rating scale reports will be exactly the same, but

the monetary return to the observer will be different. Assume that the experimenter has been successful in solving Eqs. 10.76 and 10.79 for his particular rating function. Once he knows the value of the minimum rating, u_0, he may choose any three constants, K_0, K_1, and K_2, so long as K_2 is less than $2u_0$.

$$K_0, \; K_1, \; \text{arbitrary}, \qquad K_2 < 2u_0 \qquad (10.80)$$

The role of these three arbitrary constants is shown in Fig. 10.8.

R: \ C:	SN	N
A	u	$-R \, r_0(u)$
B	$-r_0(u)$	$R u$

+

K_1	$K_0 + R K_2$
$K_1 + K_2$	K_0

=

$K_1 + u$	$K_0 + R K_2 - R r_0(u)$
$K_1 + K_2 - r_0(u)$	$K_0 + R u$

Fig. 10.8. Formation of complete pay-off matrix
from base matrix

The resultant pay-off matrix must be a valid pay-off matrix, an affect the optimization equations. To be a valid pay-off matrix, the value paid for detection must be greater than the value paid for a miss.

$$V_D > V_M : \; K_1 + u > K_1 + K_2 - r_0(u) \qquad (10.81)$$

The constant K_1 appears on both sides of the equation and therefore

disappears. This leaves the restriction

$$r_0(u) + u > K_2 \qquad (10.82)$$

Because the forcing function $r_0(u)$ is monotone increasing, it takes on its smallest value at u_0.

$$r_0(u) + u > r_0(u_0) + u_0 = 2u_0 \qquad (10.83)$$

By definition K_2 is smaller than this value, and the value of detection is indeed greater than the value of a miss. Similarly the value for a correct rejection is greater than the value paid for false alarm.

$$V_{CR} > V_{FA}: \quad K_0 + R\,u > K_0 + R\,K_2 - R\,r_0(u) \qquad (10.84)$$

Subtract K_0 from both sides of this equation, divide by R, and obtain Eq. 10.82 again. The final pay-off matrix in the β equation is

$$\beta = \frac{P(N)}{P(SN)} \frac{K_0 + R\,u - \left[K_0 + R\,K_2 - R\,r_0(u)\right]}{K_1 + u - \left[K_1 + K_2 - r_0(u)\right]} \qquad (10.85)$$

$$\beta = \frac{P(N)}{P(SN)} \frac{R\,u - R\,K_2 + R\,r_0(u)}{u - K_2 + r_0(u)} = \frac{P(N)}{P(SN)}\,R \qquad (10.86)$$

the same β equation as always. Although K_0 and K_1 disappeared from the β equation, they will appear in the value equation under the conditions A and B.

$$\ell > \beta \quad V_A(\ell) = P(SN|x) [K_1 + u] + P(N|x)$$

$$\cdot [K_0 + R K_2 - R r_0(u)] \qquad (10.87)$$

However, when one differentiates to find the maximum, these constants disappear and the same equation as before appears.

$$\frac{\partial V_A(\ell)}{\partial u} = P(SN|x) - R r_0'(u) P(N|x) \qquad (10.74)$$

Similarly, the constants appear in the equation

$$\ell < \beta \quad V_B(\ell) = P(N|x) [K_0 + Ru] + P(SN|x)$$

$$\cdot [K_0 + RK_2 - R r_0(u)] \qquad (10.88)$$

but disappear upon differentiation.

$$\frac{\partial V_B(\ell)}{\partial u} = P(N|x) R - r_0'(u) P(SN|x) \qquad (10.77)$$

The experimenter therefore has a great deal of freedom in choosing the actual pay-off matrix that he will use in his practical experimental situation. To determine the rating scale forcing function we may consider the simpler base pay-off matrix of Fig. 10.7.

In the remainder of this section, three rating scale forcing functions will be developed, both for their own sake and to demonstrate the technique for determining such a forcing function.

<u>Example:</u>   Desire $u_A = k\left(\frac{\ell}{\beta}\right)$ and $u_B = k\left(\frac{\beta}{\ell}\right)$

The basic tools are Eqs. 10.76 and 10.79. These are symmetric equations. Equation 10.76 shall be solved first in each case, and then Eq. 10.79 checked to see if it is satisfied. Equation 10.76 says that the first derivative of r evaluated at the proper value of u will be equal to the ℓ to β ratio. The desired result is to have the proper value of u equal to $\frac{k\ell}{\beta}$. This requires that r' evaluated at $\frac{k\ell}{\beta}$ equals $\frac{\ell}{\beta}$.

$$r_0'\left(k\frac{\ell}{\beta}\right) = \frac{\ell}{\beta} \Rightarrow r_0'(u) = \frac{u}{k}$$

This determines the equation for the derivative of r. The side conditions of Eq. 10.72 set the value of r' at the minimum. Since the minimum of the ℓ to β ratio is one, u_0 is equal to k.

$$r_0'(u_0) = 1 \Rightarrow u_0 = (k)$$

The function $r_0(u)$ is obtained by integration of $r_0'(u)$. Integration introduces an arbitrary additive constant, normally suppressed but displayed here

$$r_0(u) = \int r_0'(u) \, du = \int \frac{u}{k} \, du = \frac{u^2}{2k} + c$$

The first side condition of Eq. 10.72 determines this arbitrary constant.

$$r_0(u_0) = u_0 \Rightarrow C = \frac{k}{2}$$

Gathering these all together, the final form of the basic rating forcing function which will make the response proportional to the likelihood ratio under the A condition is

$$r_0(u) = \frac{u^2}{2k} + \frac{k}{2} = \frac{k}{2} \left[\left(\frac{u}{k}\right)^2 + 1 \right] \qquad u \geq k$$

Although not displayed here, this does check with Eq. 10.79. This function differs somewhat from that given in the original example. The difference is that the original example had a constant added to $r_0(u)$ to form the simpler appearing function

$$r_0(u) - \frac{k}{2} = \frac{u^2}{2k} = r(u)$$

This forcing function is displayed in Fig. 10.9(a).

Example: Desire $u = k|z - \ln \beta|$

Writing z as $\ln \ell$, we insert the desired form of u in the equation for the derivative of r.

A: $r_0'(k(\ln \ell - \ln \beta)) = \frac{\ell}{\beta} \Rightarrow r_0'(u) = e^{u/k}$

The side condition on the derivative of r at the minimum value determines that the minimum value is zero.
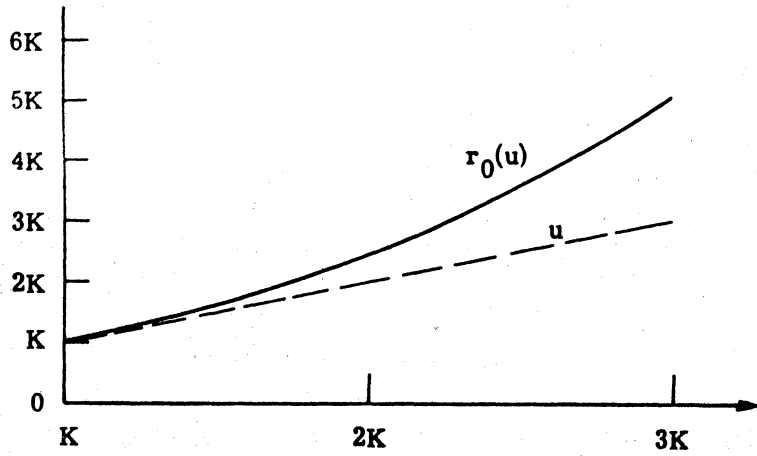
(a)

$$r_0(u) = \frac{K}{2}\left[\left(\frac{u}{K}\right)^2 + 1\right]$$

Opt $u_A = K \ell/\beta$

$u_B = K \beta/\ell$

$u_0 = K$
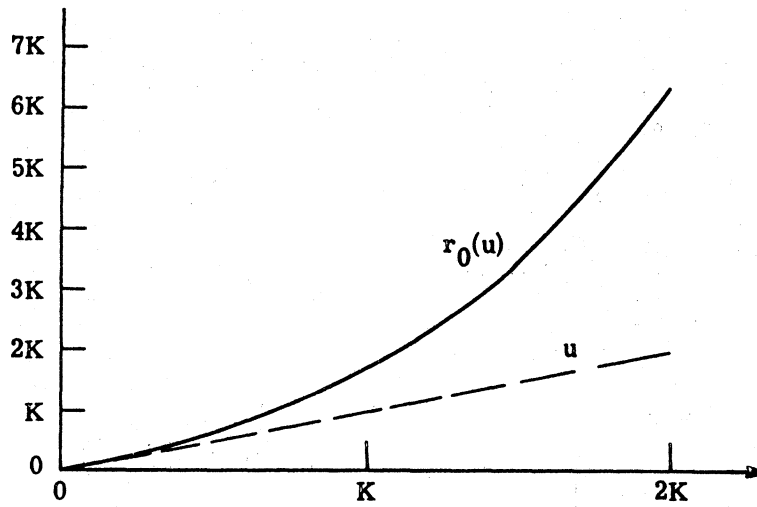
(b)

$$r_0(u) = K\left[e^{u/K} - 1\right]$$

Opt $u = K|z - \ln \beta|$

$u_0 = 0$
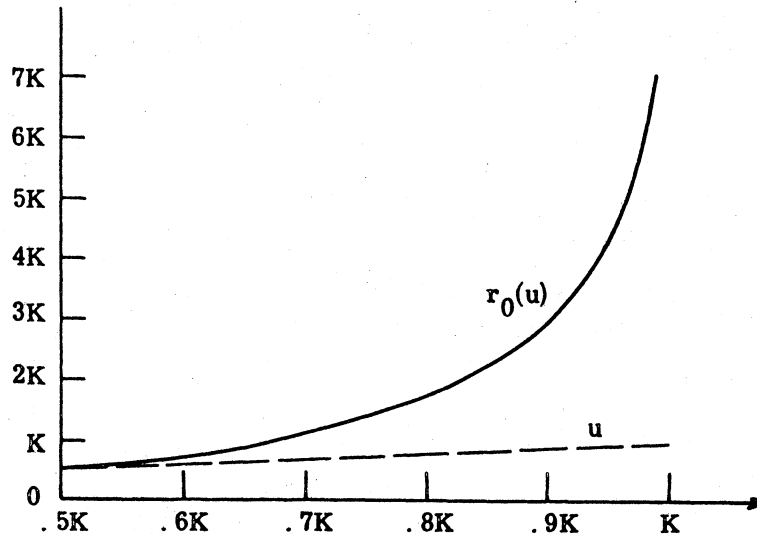
(c)

$R = \beta = 1$

$$r_0(u) = K\left[1 - \ln 2 - \frac{u}{K} - \ln\right.$$
$$\left.(1 - \frac{u}{K})\right]$$

Opt $u_A = K\ P(SN|x)$

$u_B = K\ P(N|x)$

$u_0 = K/2$

Fig. 10.9. Rating forcing functions

$$r_0'(u_0) = 1 \Rightarrow u_0 = 0$$

Integrating the derivative to obtain the function

$$r_0(u) = k\, e^{u/k} + C$$

and using the side condition to determine the constant of integration

$$r_0(0) = 0 \Rightarrow C = -k$$

the final form of the forcing function is determined.

$$r_0(u) = k\left[e^{u/k} - 1\right] \qquad u \geq 0$$

This forcing function is shown in Fig. 10.9(b).

Example: Desire $u_A = k\, P(SN|x)$ and $u_B = k\, P(N|x)$

The goal is to choose the forcing function, $r_0(u)$, so that the observer will report u values that are proportional to a posteriori probability, $P(SN|x)$. The formal solution for the "A" condition has a solution, but it differs from the "B" solution unless R and $P(SN)/P(N)$ are both one.

Restriction: $R = 1$, $P(SN) = .5 = P(N)$

If this restriction is made, the solution is

$$r_0(u) = k \left[1 - \ln 2 - \frac{u}{k} - \ln(1 - \frac{u}{k}) \right] \quad \frac{k}{2} \leq u < k$$

This is drawn in Fig. 10.9(c).

10.5 Summary of Chapters IV-X

Many different features of ROC curves have been developed in the above chapters. The purpose of these developments has been twofold: (1) to obtain specific useful results, (2) to demonstrate techniques for working with ROC curves and ROC characters.

The author hopes that others working with simple decision situations may find some useful results among these developments, and techniques that will be helpful for their own needs. If the radar analyst can determine the appropriate ROC family for his device, he will find which parameters he needs to evaluate to determine a specific ROC curve. Those in psychophysics will find a variety of curves and characters that may be more like their experimental data than the previously available ROC curves.

Every result obtained suggests further research. Especially needed is research in fitting ROC curves (and characters) to experimental data so that one may extract what information is truly present, by placing appropriate weighting on each datum. This work has no conclusion, only a continuation in the development of our understanding of the evaluation of decision devices.

APPENDIX A

TWO THEOREMS ON ROC CURVES AND LIKELIHOOD RATIO

This appendix is part of the notes for one lecture issued to the students enrolled in "Random Processes," a University of Michigan Engineering Summer Conference, in the years 1959 through 1963.

It is included in this present work because it contains the proofs of theorems herein called one and two. A more general proof of one-half of the second theorem is given in Chapter I.

Theorem 1: Given two simple hypotheses, the ROC image of all randomized decisions is convex, and contains the chance diagonal $P(A|SN) = P(A|N)$. It is symmetric through the center $(.5, .5)$, that is, if (x, y) is in the image, then so is $(1 - x, 1 - y)$.

Proof:

(1) The chance diagonal: for any real number r on the closed unit interval, consider the randomized decision probability $g(x) = r$. The expected value of any constant is that constant.

$$P("A"|N) = E(r|N) = r$$

and

$$P("A"|SN) = E(r|SN) = r$$

(2) Convex: Consider any two distinct randomized decisions, say, $g_1(x)$ and $g_2(x)$. The ROC images are $P("A_i"|N)$, $P("A_i"|SN)$. We wish to show that for any real number a

374

Proof (Cont.)

on the closed unit interval there is a randomized decision probability $g_3(x)$ whose image is

$$[aP(\text{"}A_1\text{"}|N) + (1-a) P(\text{"}A_2\text{"}|N),$$

$$aP(\text{"}A_1\text{"}|SN) + (1-a) P(\text{"}A_2\text{"}|SN)]$$

The obvious choice is to try

$$g_3(x) = a\,g_1(x) + (1-a)\,g_2(x)$$

Since g_1 and g_2 range from 0 to 1, as does a, g_3 will also; the sum of integrable functions is integrable. Therefore the above $g_3(x)$ is admissible as a randomized decision probability, and the linearity of integration assures that the image of $g_3(x)$ is the desired point (on a straight line between the images of g_1 and g_2).

(3) Symmetry: If the image of some $g_1(x)$ is the ROC point (x, y), then consider $g_2(x) = 1 - g_1(x)$. This has the range 0 to 1, is integrable, and has the desired image at $(1 - x, 1 - y)$. With respect to the ROC, g_2 is called "the opposite of g_1."

Q. E. D.

Theorem 2: The collection of all criteria based on likelihood ratio maps onto the upper boundary of the ROC region.

Proof:

The two steps in this proof are the "into" and "onto all"

Proof (Cont.)

parts, i. e., to show that any criterion based on likeli-
hood ratio maps into some point on the upper boundary,
and then that any point on the upper boundary is the
image of some criterion based on likelihood ratio.

(1) Consider any decision based on likelihood ratio
$g(x; \beta, r)$ and any other randomized decision probability,
$h(x)$, with not greater expected value under N, i. e.,

$$E(h(x)|N) \leq E(g(x; \beta, r)|N)$$

Whenever

$$f(x|SN) - \beta f(x|N) > 0, \quad \ell(x) > \beta, \quad g(x; \beta, r) = 1$$

and

$$g(x; \beta, r) - h(x) \geq 0;$$

whenever

$$f(x|SN) - f(x|N) < 0, \quad \ell(x) < \beta, \quad g(x; \beta, r) = 0$$

and

$$g(x; \beta, r) - h(x) \leq 0$$

That is

$$[f(x|SN) - \beta f(x|N)] [g(x; \beta, r) - h(x)] \geq 0$$

Therefore

$$\int [f(x|SN) - \beta f(x|N)] [g(x; \beta, r) - h(x)] dx \geq 0$$

Expanding

Proof (Cont.)

$$E(g(x; \; \beta, \; r)|SN) - E(h(x)|SN) - \beta[E(g(x; \; \beta, \; r)|N)$$

$$- E(h(x)|N)] \geq 0$$

Since the last two expected values have a positive dif-
ference by original assumption and $\beta \geq 0$,

$$E(g(x; \; \beta, \; r)|SN) \geq E(h(x)|SN)$$

Thus one concludes that, given any criterion based on
likelihood ratio, its $P("A"|SN)$ value will be greater the
or equal to the $P("A"|SN)$ value for any randomized deci-
sion device with equal to smaller $P("A"|N)$ value. Con-
sidering only equal values of $P("A"|N)$ is all that is
needed to show that any criterion based on likelihood
ratio maps into the upper boundary of the ROC, concluding
step one.

(2) One wishes first to show that for any k, $0 < k < 1$, that
there is a nonnegative β_k such that

$$E(g(x; \; \beta_k, \; 0)|N) \leq k \leq E(g(x; \; \beta_k, \; 1)|N)$$

since if one of the equalities holds, that strict-criterion
will have the appropriate $P(A|N)$ value, and if neither
equality holds, the value

$$r_k = \frac{k - E(g(x; \; \beta_k, \; 0)|N)}{E(g(x; \; \beta_k, \; 1)|N) - E(g(x; \; \beta_k, \; 0)|N)}$$

<u>Proof</u> (Cont.)

will yield the appropriate probability

$$E(g(x; \beta_k, r_k) \mid N) = k$$

Although it may be obvious to many that $E(g(x, \beta, 1) \mid N)$

is continuous on the left as a function of β, and that

$E(g(x; \beta, 0) \mid N)$ is continuous on the right, and that these

functions having common points of continuity and discon-

tinuity by their very definitions establish the proof,

the following simple steps may help others to that con-

clusion.

At the extreme left of $\beta = 0$, $E(g(x; 0, 1) \mid N) = P(\ell \geq 0 \mid N) = 1$

Toward the right extreme of arbitrarily large β note that

$$E(g(x; \beta, r) \mid SN) = \int g(x; \beta, r) \, \ell(x) f(x \mid N) dx \geq \beta \int g(x; \beta, r) f(x \mid N) \, dx$$

that is

$$E(g(x; \beta, r) \mid N) \leq \frac{1}{\beta} \ E(g(x; \beta, r) \mid SN) \leq \frac{1}{\beta}$$

so that

$$\lim_{\beta \to \infty} E(g(x; \infty, r) \mid N) \leq \lim_{\beta \to \infty} \frac{1}{\beta} = 0$$

For any point β_0 not zero or infinity, if $\beta < \beta_0$

$$P(\ell(x) \geq \beta \mid N) = P(\ell(x) \geq \beta_0 \mid N) + P(\beta \leq \ell(x) < \beta_0) \mid N)$$

Since as β approaches β_0, it exceeds all values less

than β_0.

Proof (Cont.)

$$\lim_{\beta \to \infty} P(\beta \leq \ell(x) < \beta_0 | N) = 0$$

and so

$$\lim_{\substack{\beta \to \beta_0 \\ \beta > \beta_0}} E(g(x; \beta, 1) | N) = E(g(x; \beta_0, 1) | N)$$

if $\beta > \beta_0$

$$P(\ell(x) \geq \beta | N) + P(\beta > \ell(x) > \beta_0 | N) = P(\ell > \beta_0 | N)$$

Since as β approaches β_0 it excludes all values greater

than β_0

$$\lim_{\beta - \beta_0} P(\ell(x) \geq \beta) | N) = P(\ell > \beta_0 | N)$$

and so

$$\lim_{\substack{\beta \to \beta_0 \\ \beta > \beta_0}} E(g(x; \beta, 1) | N) = E(g(x; \beta_0, 0) | N)$$

Thus the monotone decreasing (nonincreasing) function

$E(g(x; \beta, 1) | N)$ has extreme limiting values of 1 and 0,

and at any β_0 between zero and infinity has a left-hand

limit of $E(g(x; \beta_0, 1) | N)$ and a right-hand limit of

$E(g(x; \beta_0, 0) | N)$.

The extreme value of 1 corresponds to the strict cri-

terion that is the whole observation space (which is based

on likelihood ratio in a trivial sense), and the extreme value

of 0 corresponds to the strict criterion of all observations

<u>Proof</u> (Cont.)

with infinite likelihood ratio ($f(x|N) = 0$ and $f(x|SN) \neq 0$).

The extreme limiting values serve one other purpose; for

any value of k between but not equal to 0 or 1, there

will be a set of β yielding $E(g(x; \beta, 1)|N)$ values be-

tween k and 0, and a set of β's yielding $E(g(x; \beta, 1)|N)$

values between k and 1. If the value k is not taken on

by the function, then these two sets of β's are disjoint,

and the cut value is such that the left-hand limit is

greater than k, and the right-hand limit is less than or

equal to k. This is the appropriate β_k, and the value of

r_k was given at the beginning of step 2.

<div align="right">Q. E. D.</div>

APPENDIX B

A LOWER TRUNCATED SIMPLE EXPONENTIAL IS
LOWER METASTATIC INVARIANT

In Section 3.4.2 the statement was made that the metastatic image of a pure power ROC curve would be the same as the original if the point $(0, 0)$ was included in the transformation. This statement is proved in this appendix.

Denote the original ROC as

$$X_1 = Y_1^A \qquad\qquad A > 1$$

and its character is

$$\pi_1(z) = (A - 1)^{-1} A^{-\frac{1}{A-1}} e^{-\frac{A+1}{A-1}\frac{z}{2}} \qquad , z > z_0 = -\ln A$$

The equations for the coordinates, as functions of z, are

$$X_1(z) = e^{-A^{\frac{z-z_0}{A-1}}}$$

$$Y_1(z) = e^{-\frac{z-z_0}{A-1}}$$

Choose any cut, z_c, greater than z_0. The metastatic transformation will be to use only that part of $\pi_1(z)$ for which z is greater than z_c.

Let $a = 0 \qquad\qquad Y_1(a) = 0$

Let $b = e^{-A\frac{z_c-z_0}{A-1}} \qquad Y_1(b) = e^{-\frac{z_c-z_0}{A-1}}$

381

The second ROC will be the image of the rectangle,

$$0 \leq X_1 \leq b \qquad\qquad 0 \leq Y_1 \leq Y_1(b)$$

The shift in the the z-axis is given by Eq. 3.39,

$$z_2 = z_1 + \ln \frac{b-a}{Y_1(b) - Y_1(a)} \qquad z_1 \geq z_c$$

$$z_2 = z_1 + \ln \left(e^{-A\frac{z_c - z_0}{A-1}} \Big/ e^{-\frac{z_c - z_0}{A-1}} \right)$$

$$z_2 = z_1 - (z - z_0) \qquad z_1 \geq z_c$$

The minimum z_2 value occurs for z_1 at the cut value z_c.

$$\min z_2 = z_c - (z_c - z_0) = z_0$$

The new ROC character is given by applying Eq. 3.34,

$$\pi_2(z_2) = \frac{\pi_1(z_1(z_2))}{\sqrt{(b-a)(Y(b) - Y(a))}}$$

$$= \frac{\pi_1(z_2 + z_c - z_0)}{\left[e^{-A\frac{z_c - z_0}{A-1}} \quad e^{-\frac{z_c - z_0}{A-1}} \right]^{.5}} \qquad , z_2 \geq z_0$$

$$= e^{\frac{A+1}{A-1}\frac{z_c - z_0}{2}} (A-1)^{-1} A^{-\frac{1}{A-1}} e^{-\frac{A+1}{A-1}\frac{z_2 + z_c - z_0}{2}}$$

$$= (A-1)^{-1} A^{-\frac{1}{A-1}} e^{-\frac{A+1}{A-1}\frac{z_2}{2}}$$

$$\pi_2(z_2) = \pi_1(z_2) \qquad\qquad z_2 \geq z_0$$

Q. E. D.

APPENDIX C

APPROXIMATION TO A CASE II CHARACTER

The purpose of this appendix is to obtain Eq. 4. 15, an almost normal appearing character to approximate the $\alpha = 2$ Case II character.

The equations for the ROC character for Case II are written in terms of a parameter of quality, α, and a dummy, t. The tabulated function

$$T_\nu(\lambda) = e^{-\lambda} I_\nu(\lambda) \qquad (4.11)$$

is used in computation. The computational equations are

$$z = [\ln T_0(\alpha t) + \alpha t]_1 - .5\alpha^2 \qquad (4.13)$$

$$\ln \pi(z) = [.5\alpha t + \ln(\alpha t) + 1.5 \ln T_0(\alpha t) - \ln T_1(\alpha t)]_2$$

$$- .5t^2 - [.25\alpha^2 + 2\ln \alpha] \qquad (4.14)$$

Let

$$B_1(\alpha t) = \ln T_0(\alpha t) + \alpha t \qquad (C.1)$$

$$B_2(\alpha t) = .5\alpha t + \ln(\alpha t) + 1.5\ln T_0(\alpha t) - \ln T_1(\alpha t) \qquad (C.2)$$

These were calculated over the range of αt from zero to twenty. The results appear in Table C. 1, and B_2 is plotted against B_1 with integer αt points indicated in Fig. C. 1.

These equations were used to calculate the log of the ROC character for $\alpha = 2$. This is tabulated in the fourth and fifth

383

αt	$B_1'(\alpha t)$	$B_2(\alpha t)$	z	$-\ln \pi(z)$	$X(t)$	$z+.725$	$-(\ln \pi(z)+1.385)$
					$\sigma = 2$		
0.	0.	.6935	-2.000	1.6935	1.000	-1.275	.3085
.2	.013	.709	-1.987	1.683	.995	-1.262	.298
.4	.039	.734	-1.941	1.673	.980	-1.216	.288
.6	.088	.792	-1.912	1.640	.956	-1.187	.255
.8	.155	.848	-1.845	1.619	.923	-1.120	.234
1.0	.235	.923	-1.765	1.589	.883	-1.040	.204
1.2	.334	1.021	-1.666	1.546	.835	-0.941	.161
1.4	.441	1.118	-1.559	1.514	.811	-0.834	.129
1.6	.560	1.230	-1.440	1.477	.726	-0.715	.092
1.8	.687	1.344	-1.313	1.448	.667	-0.588	.063
2.0	.823	1.466	-1.177	1.421	.606	-0.453	.036
2.2	.965	1.586	-1.035	1.406	.546	-0.310	.021
2.4	1.113	1.712	-0.887	1.395	.487	-0.162	.010
2.6	1.269	1.846	-0.731	1.386	.430	-0.006	0.001
2.8	1.426	1.974	-0.574	1.393	.375	+0.151	.008
3.0	1.585	2.105	-0.415	1.407	.325	+0.310	.022
4.0	2.425	2.744	+0.425	1.643	.135	+1.150	.258
5.0	3.304	3.374	1.304	2.138	.048	+2.029	.753
6.0	4.209	3.987	2.209	2.900	.0111	2.934	1.515
7.0	5.126	4.508	3.126	4.004	.00218	3.851	2.619
8.0	6.058	5.177	4.058	5.210	.000335	4.783	3.825
9.0	7.000	5.760	5.000	6.752	.00004	5.725	5.367
10.0	7.942	6.323	5.942	8.564	.00000373	6.667	7.179
12.0	9.850	7.452	7.850	12.935	.0000000152	8.572	11.550
15.0	12.739	9.117	10.739	21.395	6.1×10^{-13}	11.464	20.010
18.0	15.641	10.731	13.641	32.156	-	14.366	30.771
20.0	17.59	11.795	15.59	40.592	1.93×10^{-22}	16.315	39.207

Table C. 1 Calculated number
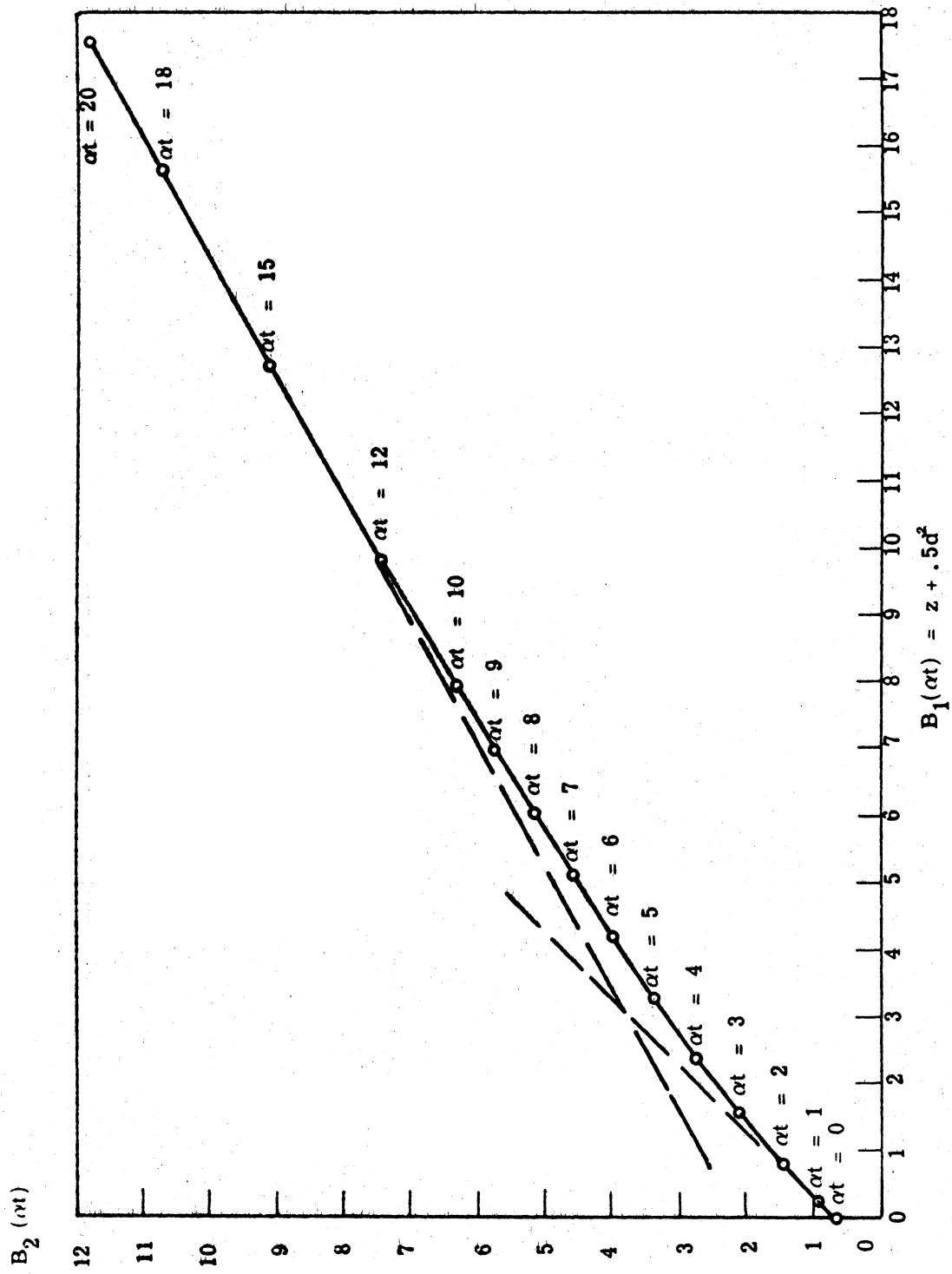
Fig. C. 1. Comparison of intermediate calculation functions

columns of Table C. 1, and plotted in Fig. C. 2 for the z range of -2 to +5. This range of z corresponds to the range of false alarm probabilities from x = .00004 to x = 1.00.

This arc appears so parabolic that the calculations were continued to obtain a power-law fit to the arc. From Table C. 1, columns four and five, it was estimated that the mode of ln $\pi(z)$ occurred at z = -.725 at which point the ln $\pi(z)$ value was -1.385. If the character were truly parabolic, then it should obey the law

$$\ln \pi(z) = -1.385 - A(z + .725)^2 \qquad (C.3)$$

If so, then

$$-(\ln \pi(z) + 1.385) = A(z + 725)^2 \qquad (C.4)$$

A plot of $-(\ln \pi(z) + 1.385)$ against $|z + .725|$ should plot as a straight line of slope 2 on log-log paper. These last two quantities were calculated; they are listed in columns seven and eight of Table C. 1 and plotted in Fig. C. 3.

The plotable ranges for Fig. C. 2 are -2. to -.887 for z below the mode of z = -.725, and -.415 to 15.59 for z above the mode. Slight differences in the fit near the mode are inconsequential. The character above the mode is the more important arc, because it covers the false alarm range below x = .325. All plotted points above the mode are well fit by a single straight line. However, the slope is not two. The end points of this straight line are given in Table C. 2.
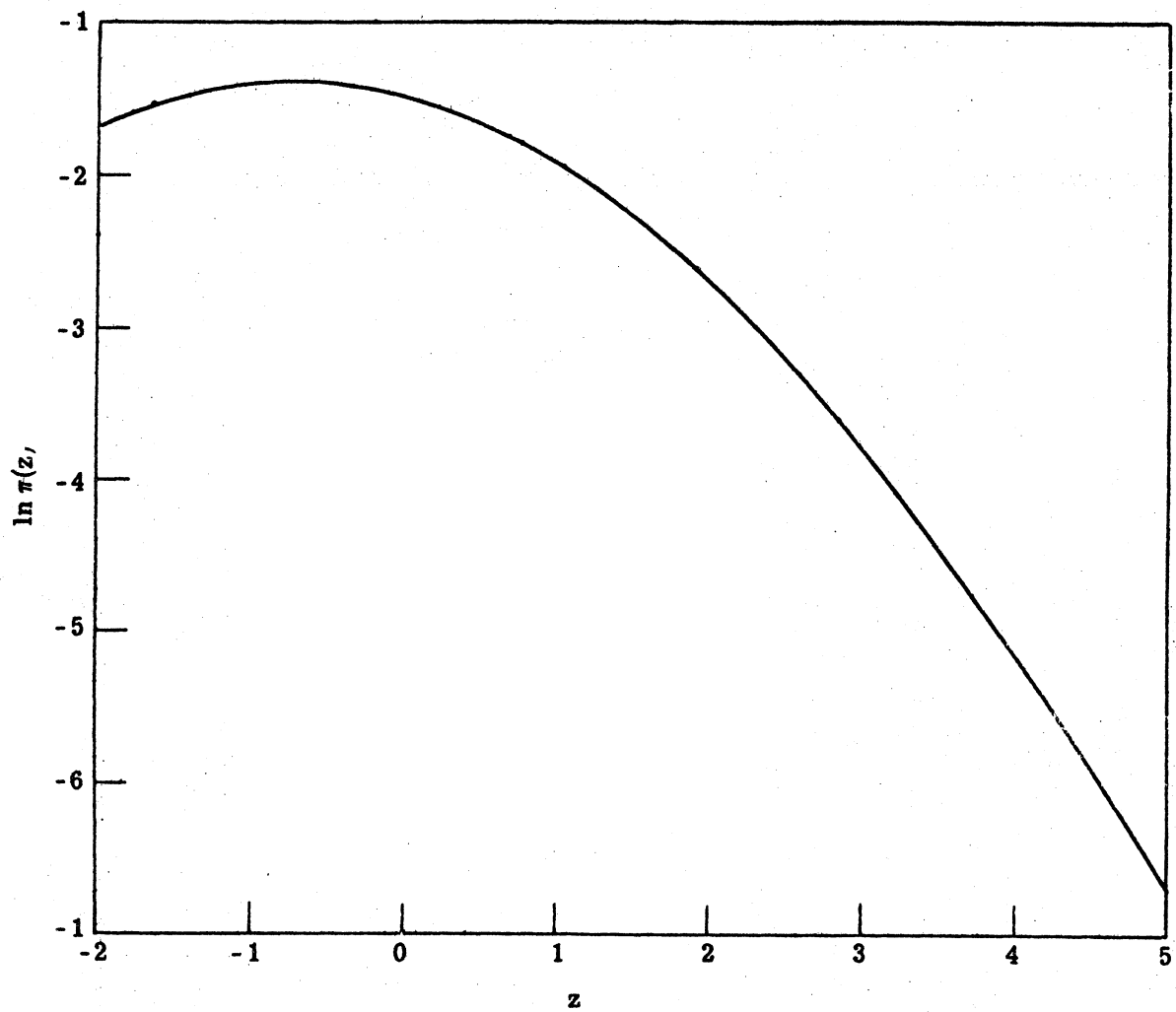
Fig. C.2. ROC character Case II, $\alpha = 2$
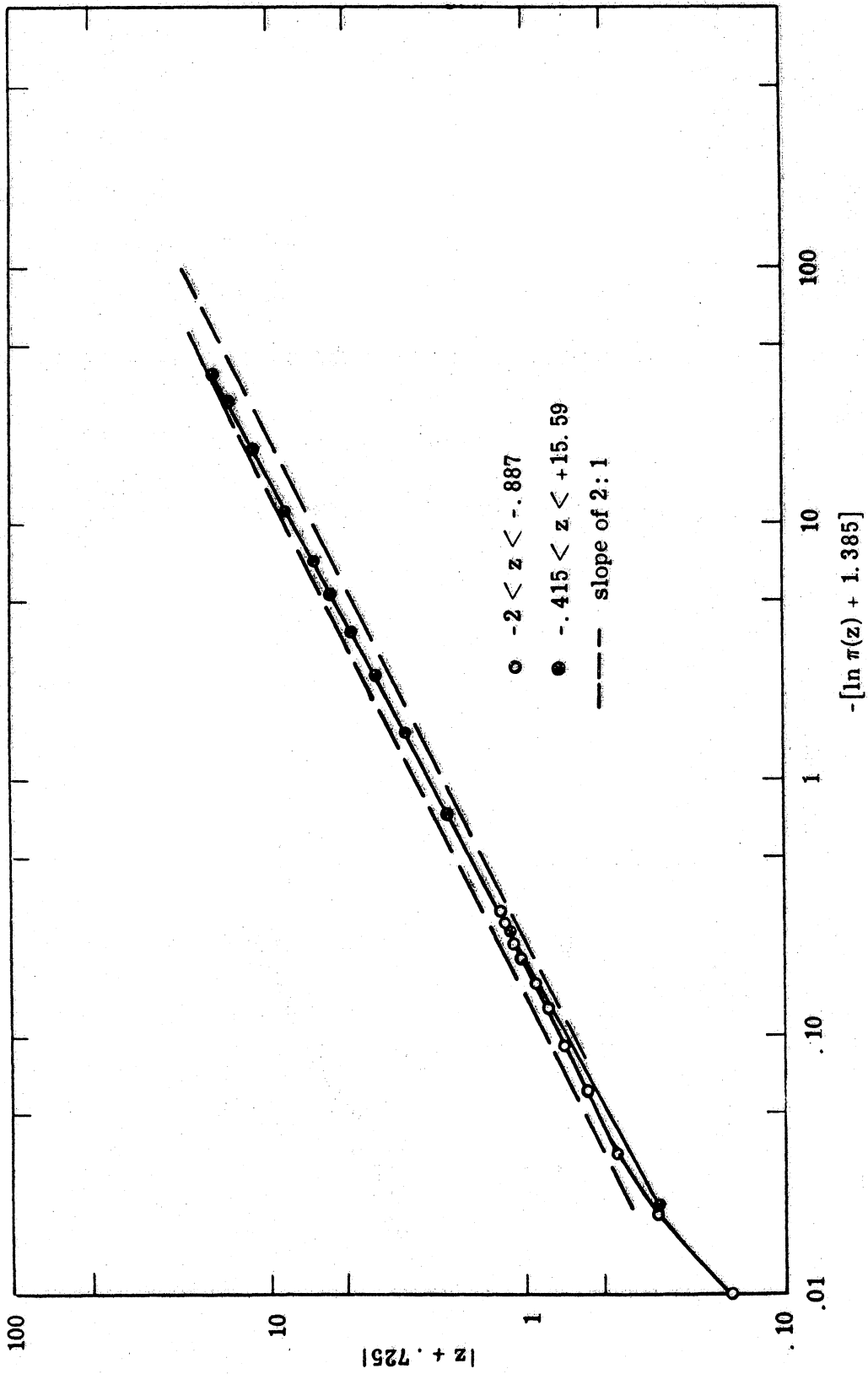
Fig. C.3. Graphic determination of a power law relation

z + .725	-(ln π(z) + 1.385)	x(z)
0.310	0.22	0.325
16.315	39.207	1.93×10^{-22}

Table C.2 End points of approximation

This straight line has the equation

$$-\ln \pi(z) \quad -1.385 = .20 \, |z + .725|^{1.885} \qquad (C.5)$$

From which it follows that

$$\pi(z) = e^{-1.385} \, e^{-.2|z + .725|^{1.885}} \qquad (C.6)$$

Equation C.6 may also be written

$$\pi(z) = .25 \; e^{-\frac{1}{2} \frac{|z + .725|^{1.885}}{2.5}} \qquad (C.7)$$

This may reasonably be considered an approximation to the $\alpha = 2$

Case II ROC character, since it also fits the mode region well.

The conclusion is the Eq. 4.15.

$$z > -2$$

$$(C.8)$$

APPENDIX D

DETAILS OF AN APPROXIMATION TECHNIQUE FOR SECTION 4.3

This appendix contains a demonstration of how the ROC character can be used to obtain an approximation for the "Case 3: Noise in Noise" ROC curve. The special restrictions are that WT is large, and the signal-to-noise ratio, S/N, is small.

Some idea of the magnitude of the various terms involved in the ROC character must be obtained. The first of these is the lower bound, z_0. The first several terms of the standard approximation for the logarithm yield

$$z_0 = -WT \ln(1 + \frac{S}{N}) \approx -WT \frac{S}{N} (1 - \frac{S}{2N} + \frac{S^2}{3N^2}) \quad (D.1)$$

The lower bound z_0 is negative, and its magnitude is the product of the large quantity WT times a small quantity S/N. For detectability to be reasonably good the range of the nonzero part of the ROC character must extend to large values of z. Although large positive values might appear to be sufficient to have a good ROC curve for low values of false alarm, one cannot have large positive values for the ROC character without having large negative values also. (The expected value of z is negative under the condition N.) Make the further assumption that z_0 has large magnitude. The logarithm of the ROC character

$$\ln \pi(z) = (WT - 1) \ln(z - z_0) - (\frac{N}{S} + .5)(z - z_0) + WT \ln(\frac{N}{S} \sqrt{1 + \frac{S}{N}}) - \ln \Gamma(WT) \quad (D.2)$$

and its derivative

$$\frac{d}{dz} \ln\pi(z) = \frac{WT - 1}{z - z_0} - (\frac{N}{S} + .5) = (\frac{N}{S} + .5) \frac{z_m - z}{z_0 - z} \qquad (D.3)$$

are going to be the tools in determining the nature of approximation to make for this ROC curve. The quantity z_m in Eq. D.3 is the mode of the ROC character.

$$z_m = z_0 + \frac{S}{N} \frac{WT - 1}{\left(1 + \frac{S}{2N}\right)} = z_0 + (WT - 1) (\frac{S}{N}) (1 - \frac{S}{2N} + \frac{S^2}{4N^2} \cdots)$$

$$(D.4)$$

Since z_0 is roughly $-WT\, S/N$ and the mode z_m is $(WT - 1)\, S/N$ above it, z_m must be in the neighborhood of S/N, that is, very close to zero. Equation D.3, with mode set at zero is very similar to a normal ROC character if the denominator changes very little. To formalize this, expand Eq. D.3 is a power series in z about the mode, z_m.

$$\frac{z_m - z}{z - z_0} = - \frac{z - z_m}{z - z_0} = - \frac{z - z_m}{(z_m - z_0) + (z - z_m)} = \frac{z - z_m}{z_m - z_0} \frac{1}{\left(1 + \frac{z - z_m}{z_m - z_0}\right)}$$

$$(D.5)$$

$$= \left(\frac{z_m - z}{z_m - z_0}\right) \sum_{k=0}^{\infty} (-1)^k \left(\frac{z - z_m}{z_m - z_0}\right)^k \qquad \left|z - z_m\right| < z_m - z_0$$

For z very close to z_m, neglect all except the first term

$$\frac{d}{dz} \ln\pi(z) \cong (\frac{N}{S} + .5) \frac{z_m - z}{z_m - z_0} = \frac{\left(\frac{N}{S} + .5\right)^2}{WT - 1} (z_m - z) \qquad (D.6)$$

Since $S \ll N$

$$z_0 \approx -WT \frac{S}{N} \tag{D. 7}$$

$$z_m \approx z_0 + (WT - 1)(\frac{S}{N}) \left(1 - \frac{S}{2N}\right) < |z_0| \tag{D. 8}$$

$$\frac{d}{dz} \ln \pi(z) \approx \frac{(N/S)^2}{WT - 1} \frac{(z_m - z)}{z_m - z_0} = \frac{\left(\frac{N}{S} + .5\right)^2}{WT - 1} (z_m - z) \tag{D. 9}$$

For a normal ROC character

$$\frac{d}{dz} \ln \pi_N(z) = -z/d \tag{D. 10}$$

Can we treat the small S/N case as a normal ROC? The matching

index is

$$\hat{d} = (WT - 1) \ (\frac{S}{N})^2 \tag{D. 11}$$

To answer this we must determine how well the entire ROC matches;

specifically, how the values at the mode match for the same index.

Calculate $\ln \pi(z_m)$.

$$\ln \pi(z_m) = (WT - 1) \ln(z_m - z_0) - (\frac{N}{S} + .5) (z_m - z_0) + WT \ln\left(\frac{N}{S} \sqrt{1 + \frac{S}{N}}\right) - \ln \Gamma(WT) \tag{D. 2}$$

$$(WT - 1) \ln(z_m - z_0) = (WT - 1) \ln \left[(WT - 1)\frac{S}{N} \left(1 + \frac{S}{2N}\right)^{-1}\right] \tag{D. 12}$$

$$= (WT - 1) \ln(WT - 1) + (WT - 1) \left[\ln \frac{S}{N} - \ln \left(1 + \frac{S}{2N}\right)\right]$$

Now $(WT - 1) \ln(WT - 1)$ is closely related to $\ln \Gamma$. Specifically,

Eq. 6.1.40 of Ref. 6 gives

$$\ln \Gamma(\lambda) \sim (\lambda - .5)\ln \lambda - \lambda + \ln\sqrt{2\pi} + \ldots$$

hence   (D. 13)

$$(WT - 1)\ln(WT - 1) = .5\ln(WT - 1) + (WT - 1.5)\ln(WT - 1)$$

$$\approx .5\ln(WT - 1) + \ln\Gamma(WT - 1) + (WT - 1) - \ln\sqrt{2\pi}$$

(D. 14)

Collecting

$$(WT - 1)\ln(z_m - z_0) \approx .5\ln(WT - 1) + \ln\Gamma(WT - 1) + (WT - 1) - \ln\sqrt{2\pi}$$

$$+ (WT - 1) \cdot \left[\ln \frac{S}{N} - \ln\left(1 + \frac{S}{2N}\right)\right]$$ (D. 15)

Next

$$-\left(\frac{N}{S} + .5\right)(z_m - z_0) = -(WT - 1)$$ (D. 16)

Hence

$$\ln\pi(z_m) \approx .5\ln(WT - 1) + \ln\Gamma(WT - 1) - \ln\Gamma(WT) - \ln\sqrt{2\pi} + (WT - 1)$$

$$\left[\ln\frac{S}{N} - \ln\left(1 + \frac{S}{2N}\right)\right] + WT\left[-\ln\frac{S}{N} + .5\ln\left(1 + \frac{S}{N}\right)\right]$$ (D. 17)

Now

$$\Gamma(WT) = (WT - 1)\Gamma(WT - 1)$$ (D. 18)

So

$$-\ln\Gamma(WT) = -\ln(WT - 1) - \ln\Gamma(WT - 1)$$ (D. 19)

$$\ln \pi(z_m) \approx -.5(WT - 1) - \ln\sqrt{2\pi} - \ln\left(\frac{S}{N}\right) - (WT - 1)\ln\left(1 + \frac{S}{2N}\right) + .5WT\ln\left(1 + \frac{S}{N}\right)$$

(D. 20)

The first three terms can be written as

$$-\ln \sqrt{2\pi(WT - 1)\left(\frac{S}{N}\right)^2} = -\ln \sqrt{2\pi \hat{d}}$$

(D. 21)

Expand the last two terms in power series

$$-(WT - 1)\ln\left(1 + \frac{S}{2N}\right) = -(WT - 1)\left[\frac{S}{2N} - \frac{1}{2} \circ \frac{S^2}{4N^2} + \frac{1}{3} \cdot \frac{S^3}{8N^3} \cdots\right]$$

(D. 22)

$$+ .5WT\ln\left(1 + \frac{S}{N}\right) = .5 WT\left[\frac{S}{N} - \frac{1}{2}\frac{S^2}{N^2} + \frac{1}{3} \circ \frac{S^3}{N^3}\right]$$

(D. 23)

Adding, and collecting in powers of S/N

$$\frac{S}{2N} \quad [-(WT - 1) + WT] \qquad = \frac{S}{2N}$$

$$\frac{S^2}{8N^2} \quad [(WT - 1) - 2WT] \qquad = -(WT + 1)\frac{S^2}{8N^2}$$

$$\frac{S^3}{24N^3} \quad [-(WT - 1) + 4WT] \qquad = (3WT + 1)\frac{S^3}{24N^3}$$

$$\frac{S^4}{64N^4} \quad [(WT - 1) - 8WT] \qquad = -(7WT + 1)\frac{S^4}{64N^4}$$

(D. 24)

Carefully review the assumed order of magnitude of the variables.

(1) S/N is quite small, $S/N \ll 1$

If $\hat{d}$ is to be the detectability index, it should range from $.5$ to 50.

Hence

$\quad$ (2) $(WT - 1)\,(S/N)^2$ is between $.5$ and 50

(1) and (2) mean

$\quad$ (3) $(WT - 1)\,S/N$ is very large

Putting these all together means the second term in the above collection from the two power series is outstandingly large compared to the rest. It is

$$-(WT + 1)\,\frac{S^2}{8N^2} = -(WT - 1)\,\frac{S^2}{8N^2} - \frac{S^2}{4N^2} \approx -(WT - 1)\,\frac{S^2}{8N^2} = -\frac{1}{8}\,\hat{d}$$

$$(D.\,25)$$

Gathering all together

$$\ln \pi(z_m) \approx -\frac{1}{8}\,\hat{d} - \ln \sqrt{2\,\pi\,\hat{d}} \qquad (D.\,26)$$

To summarize the approximation technique; the Pearson ratio, the derivative of the logarithm of the ROC character, under the assumption that the range of the variable was substantial, suggested a match with a normal ROC curve. By matching values for the derivation of the logarithm of the ROC character, a relation between what would have to be the matching normal index and the parameters of this problem was obtained (Eq. 4.38). Since a metastatic transformation of a normal ROC curve would also give this same matching value, it was necessary to calculate an actual value on the height of the ROC character, or correspondingly, its logarithm. Using the standard power series and approximation-piled-on-approximation technique common in applied

mathematics, such a calculation of the magnitude of the ROC character at its mode value was obtained. This is Eq. D. 26. This maximum value of the ROC character matched that for the normal. Since the mode is extremely close to zero, the logarithm of the ROC character is parabolic in the neighborhood of zero and the magnitude of the ROC character corresponds to the width parameter in the parabola in exactly the same way that a normal character behaves. We conclude that the Pearson III characters are essentially normal in the limiting case studied.

APPENDIX E

A CONIC ROC COMPUTER PROGRAM

This appendix contains the details of a computer program mentioned in Section 5.4. A regular conic ROC is specified by its minimum slope (called LL0), maximum slope (called LHI), and one super-chance point (X_0, Y_0). These three items of information are the input information to the program.

A simplified computer flow diagram is shown in Fig. E.1. The equations for each computation block constitute the remainder of this appendix.

SETUP   Certain quantities are used repeatedly and do not depend on the input data. These are computed first.

$$Z(I) = .1(I-100)$$
$$L(I) = e^{Z(I)} \qquad \left.\right\} \quad 1 \leq I \leq 200$$

$$\left.\begin{array}{l} X(1) = 10^{-4} \\ X(2) = 10^{-3} \\ X(J \geq 3) = .02J-.05 \\ H(J) = \ln \dfrac{X(J)}{1-X(J)} \end{array}\right\} \quad 1 \leq J \leq 52$$

CHECK   The point (X_0, Y_0) must lie between the chance diagonal and the Luce ROC. The program is written for the ROC curve slope range (.00005, 22000). If any of the following inequalities is TRUE, a transfer to "GOOF" occurs.

397
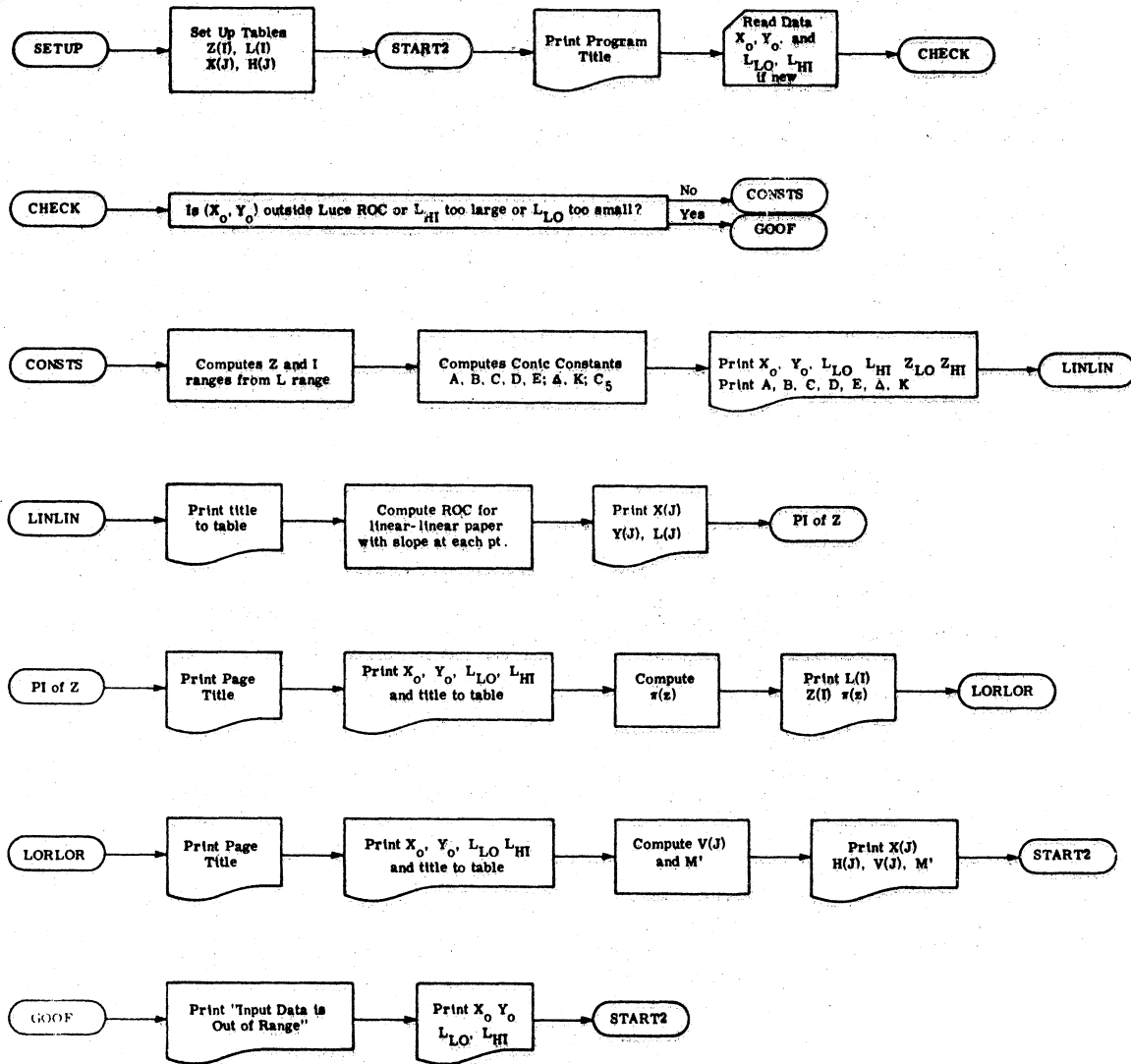
Fig. E.1. Simplified flow diagram for a conic ROC program

$$X_0 \geq Y_0$$

$$Y_0 > (LHI)X_0$$

$$(1-Y_0) < (LL0)(1-X_0)$$

$$LHI > 22000$$

$$LL0 < .00005$$

CONSTS   The constants in the algebraic equation for the ROC and several auxiliary constants are computed.

$$ZL0 = \ln (LL0)$$

$$ZHI = \ln (LHI)$$

ILO = (integer part of) $10(ZL0) + 101$

       whenever this is less than 1, set IL0 = 1

IHI = (integer part of) $10(ZHI) + 100$

       whenever this is greater than 200, set IHI = 200

$$C_1 = (LHI + LL0 - 2)/(1 - LL0)$$

$$C_2 = (1 - (LL0)(LHI))/(1 - LL0)$$

$$C_3 = -2(Y_0 - X_0)^{-2}$$

$$A = C_3(C_1 X_0 Y_0 + C_2 X_0^2 + Y_0 - (LHI)X_0)$$

$$B = C_1 - A$$

$$C = 2C_2 + A$$

$$D = 1$$

$$E = -(LHI)$$

$$\Delta = B^2 - AC$$

$$K = \sqrt{AE^2 - 2BDE + CD^2}$$

$$C_5 = 2(B + A(LHI))$$

The ROC curve appropriate for plotting on linear-linear graph paper is computed, for the x points precomputed in SETUP. The slope of the ROC curve, the likelihood ratio, is also computed.

$$X = X(J)$$

$$Y = (\sqrt{\Delta X^2 + C_5 X + 1} - BX - 1)/A$$

$$Y(J) = Y$$

$$L = (-E - AC - BY)/(D + BX + AY)$$

$$1 \leq J \leq 52$$

The horizontal coordinate **X** is called from memory to compute the vertical coordinate **Y** and the slope L. **X**, **Y** and **L** are then printed out and Y is stored for later use. L is not stored, nor does it correspond to the stored list L(I).

PIOFZ   This computes the ROC character. The general equation for the conic ROC character is

$$\pi(z) = k(Ae^{Z} + 2B + Ce^{-Z})^{-1.5}$$

The computational equations are

$$T = AL + 2B + C/L$$

$$CHAR = k/(T\sqrt{T})$$

for

$$L = LL0; \; L(I) \qquad IL0 \leq I \leq IHI; \; LHI$$

After each character value is computed, L, Z, and the character value are printed.

LORLOR   The log-odds-ratio coordinate for any probability p is ln (p/(1 - p)). This was computed for the x coordinate in SETUP. The y values were calculated in LINLIN.

$$H = H(J)$$
$$V = \ln \frac{Y(J)}{1 - Y(J)}$$
$$MPRIME = V - H$$

$$1 \leq J \leq 52$$

The program terminates (on the University of Michigan machine) when the READ DATA statement fails to encounter new data.

APPENDIX F

BINORMAL COMPUTER PROGRAMS

This appendix contains two programs for finding regular ROC curves related to an irregular binormal ROC curve. Both assume that the slope, S, is between zero and one and that the quality Q, (the d' value at $Y = .50$) is positive.

The first program searches for the point of tangency of the ROC curve and the tangent inferior. It then performs "external rectification" by adjoining the tangent inferior to the regular part of the ROC curve. The printout for the regular part of the ROC curve consists of the ROC linear coordinates, the slope L, its natural logarithm z, the slope of the secants from $(0, 0)$ and $(1, 1)$ to the ROC point, and the ROC character.

The second program operates on the decision axis inferred by the original ROC. The likelihood ratio decision is essentially a two-tail test whereas the original ROC is a one-tail test. The program computes the original ROC, the lower tail contributions, the total two-tail ROC curve, and the ROC character.

Figure F. 1 is the computer flow diagram for the external rectification program. The details of the printed comments and equations are as follows:
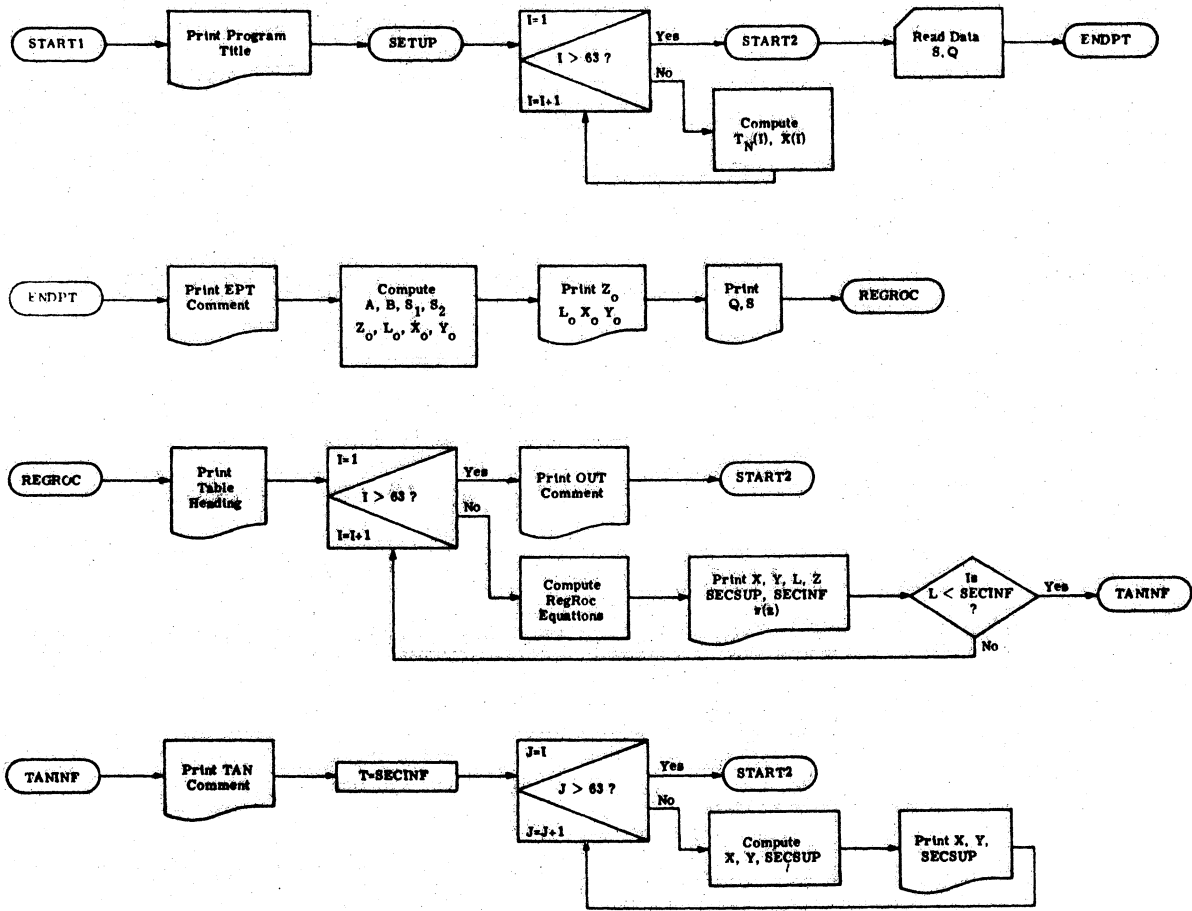
Fig. F.1. Flow diagram for external rectification

START1   The program title reads:

> REGULARIZED BINORMAL ROC PROGRAM NUMBER ONE
> BINORMAL ROCS PLOT AS STRAIGHT LINES WITH
> SLOPE LESS THAN ONE ON NORMAL-NORMAL PAPER.
> IF NOT REGULARIZED THEY WOULD STOP SHORT OF
> (1, 1) AT Z = Z_0. TO BE REGULARIZED THEY ARE
> CONTINUED ALONG THE TANGENT INFERIOR TO (1, 1).
> THE INPUT DATA ARE IN TERMS OF THE ROC ON
> NORMAL-NORMAL PAPER. Q IS THE VALUE OF
> DPRIME AT . 50 DETECTION, AND S IS THE SLOPE.

SETUP

$$T_N(I) = .1(I - 25)$$

$$X(I) = \Phi(T_N(I))$$

$$1 \leq I \leq 63$$

This computes the normal argument from -2. 4 to +3. 8, and the assoc-

iated false alarm probability range is . 0082 to . 999928. These points

are uniformly spaced on normal-normal paper. The University of

Michigan IBM 7090 MAD Language does not have a $\Phi(\ )$ subroutine.

The actual equation used is

$$\Phi(t) = {}_{\circ}5 + . 5*ERF. (\sqrt{.5}\, t)$$

ENDPT   The EPT comment reads:

> THE UNREGULARIZED BINORMAL ROC WITH SLOPE
> S AND . 50 DETECTION DPRIME Q WOULD RUN FROM
> (0, 0) TO THE POINT OF MINIMUM SLOPE.

Some constants, and the coordinates of the point of minimum slope are

then calculated.

Whenever $S > 1$, $S = 1/S$

$A = S Q/(1 - S^2)$

$B = (1 - S^2)/2S^2$

$S_1 = \sqrt{S}$

$S_2 = 1/S_1$

$Z_0 = \ln S - QSA/2$

$L_0 = e^{Z_0}$

$X_0 = \Phi(SA)$

$Y_0 = \Phi(A)$

REGROC   The slope of the ROC curve, L, will usually fall below the

secant inferior before x = .999928. This program assumes that 0.1 steps

in normal argument is sufficiently fine quantization that the tangent inferior

corresponds to the secant inferior which first falls above L. ROC curve

points are computed until this occurs, or until I = 63.

$T_N = T_N(I)$

$T_{SN} = S(T_N + Q)$

$C = .39894228 \text{ EXP. } (-(T_N^2 + T_{SN}^2)/4)$

$PI = C/(S_1 T_{SN} - S_2 T_N)$

$X = X(I)$

$Y = \Phi(T_{SN})$

$Z = Z_0 + B(A - T_{SN})^2$

$L = e^Z$

$SECSUP = Y/X$

$SECINF = (1 - Y)/(1 - X)$

If the slope of the ROC curve is still above the secant inferior by

x = .999928, then the computation terminates and the following OUT

comment is printed:

IN THE ABOVE COMPUTATION, NO REGULARIZATION WAS
USED.

TANINF The TAN comments read:

      BINORMAL ROC IS IRREGULAR FROM HERE ON.

      TANGENT INFERIOR PORTION FOLLOWS.

The loop index, J, for this computation is initiated at the last used

value for the REGROC loop index, I. This repeats this computation

for that one point, and then continues until x = .999928.

     $X = X(J)$

     $Y = 1 - T + TX$

     $SECSUP = Y/X$

The program terminates (on the University of Michigan machine)

when the READ DATA statement fails to encounter new data.

    The flow diagram for the internal rectification program

is given in Fig. F.2.

FRONT The front page of each program bears the following

program title:

      REGULARIZED BINORMAL ROC PROGRAM NUMBER TWO.

      BINORMAL ROCS PLOT AS STRAIGHT LINES WITH

      SLOPE LESS THAN ONE ON NORMAL-NORMAL PAPER.

      IF NOT REGULARIZED THEY HAVE A POINT OF

      MINIMUM SLOPE SHORT OF (1, 1). TO BE REGULARIZED

      A NEW DECISION AXIS BASED ON LIKELIHOOD RATIO

      IS USED. THE INPUT DATA ARE IN TERMS OF THE

      ROC ON NORMAL-NORMAL PAPER. Q IS THE VALUE

      OF DPRIME AT .50 DETECTION, AND S IS THE SLOPE.

      TANGENT INFERIOR P(A/SN) IS GIVEN VS P(A/N)

      OF THE FINAL ROC.
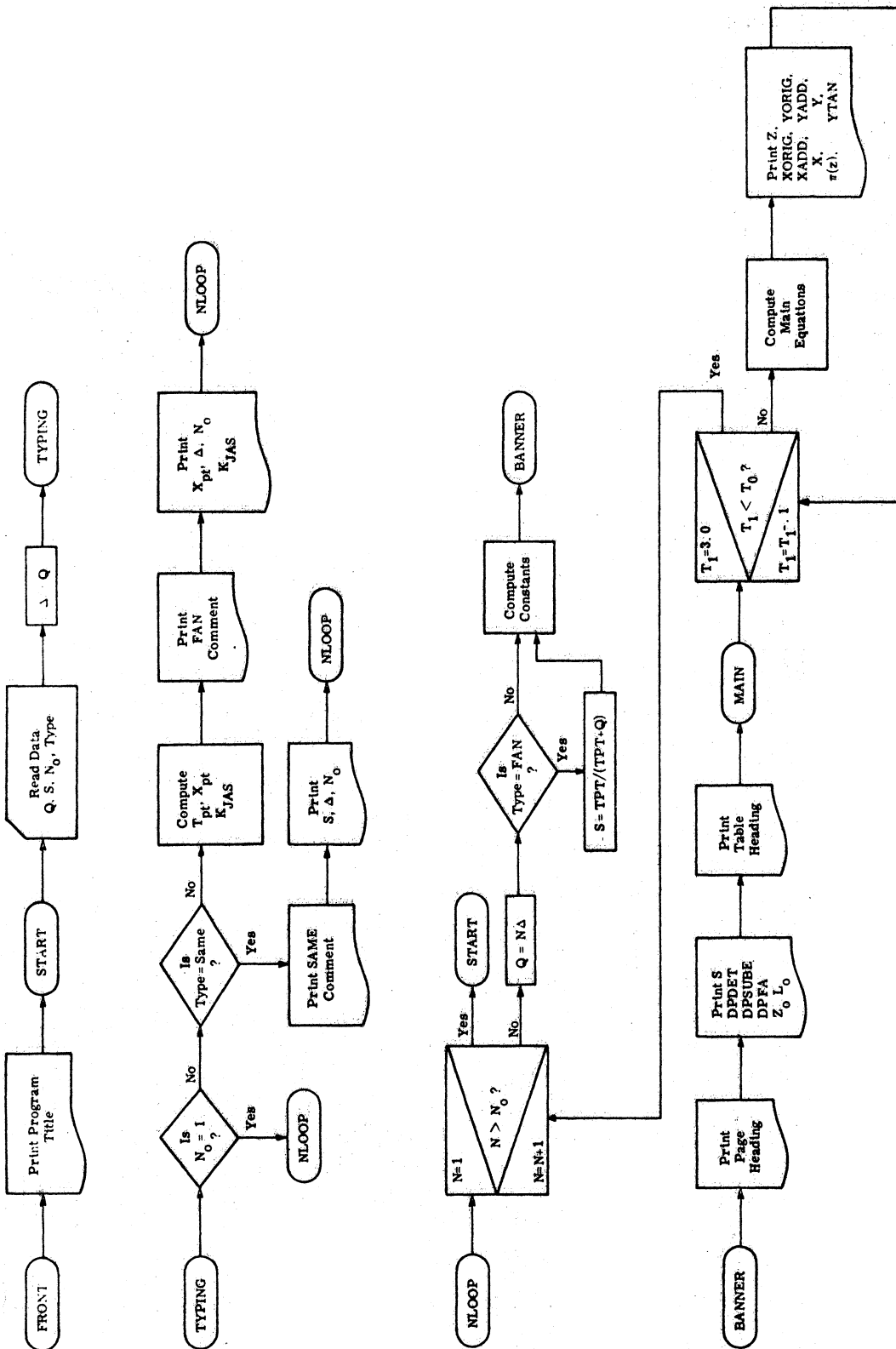
Fig. F.2. Flow diagram for internal rectification

<u>START</u>   In addition to the data Q and S, the input data contains

an integer, No., and a word, type. To obtain one curve, set No. = 1.

To obtain N parallel curves, set No. = N, and type = SAME. To

obtain N ROC curves that have a common chance diagonal point set No.

= N, type = FAN. (Such a "fan" is a family of ROC curves modeled as

a zero-mean unit-variance noise, with SN distribution normal with

mean Q and standard deviation, $1 + k_{JAS}Q$). If a number of curves are

calculated, Q will be used both as the initial value and as the incre-

mental Q value.

<u>TYPING</u>   The SAME comment reads: "I SHALL NOW COMPUTE

A NUMBER OF CURVES WITH IDENTICAL NORMAL-NORMAL SLOPE."

It is followed by the values of S, Δ, No.

If the type is FAN

$$T_{pt} = SQ/(1 - S)$$
$$X_{pt} = \Phi(T_{pt})$$
$$k_{JAS} = 1/T_{pt}$$

The FAN comment reads: "I SHALL NOW COMPUTE A FAN OF

CURVES THROUGH THE COMMON CHANCE DIAGONAL POINT." It

is followed by the values of X_{pt}, Δ, No., k_{JAS}.

<u>NLOOP</u>   The equations in the "compute constants" block are

$$B = 1 - S^2$$
$$Z_0 = \ln S - S^2 Q^2 / 2B$$
$$L_0 = e^{Z_0}$$
$$DPDET = Q$$
$$DPSUBE = 2SQ/(1 + S)$$

$$DPFA = QS$$

$$T_0 = S^2 Q/B$$

$$C = 2T_0$$

BANNER   The page heading reads: "THE SLOPE ON NORMAL-NORMAL PAPER, AND THE DPRIME VALUES OF THE UNREGULARIZED ROC AT .50 DETECTION, NEGATIVE DIAGONAL, AND .50 FALSE ALARM ARE   ." This is followed by the values of S, DPDET, DPSUBE, DPFA, then followed by Z_0 and L_0.

MAIN   The MAIN equations are

$$T_2 = -T_1 - C$$

$$T_3 = S(T_1 - Q)$$

$$T_4 = S(T_2 - Q)$$

$$T_5 = T_0 + T_1$$

$$XORIG = \Phi(-T_1)$$

$$YORIG = \Phi(-T_3)$$

$$XADD = \Phi(T_2)$$

$$YADD = \Phi(T_4)$$

$$X = XORIG = XADD$$

$$Y = YORIG + YADD$$

$$Z = Z_0 + BT_5^2 \qquad -.5T_1^2 \qquad -.5T_2^2$$

$$PI = .39894228 \, e^{-.5} \, e \qquad + e \qquad / BT_5$$

$$YTAN = 1 - L_0 + L_0 X$$

These programs were used to obtain regularized binormal examples for Chapter VIII.

MANIPULATION OF BINOMIAL SUMS

__Remark 1__

$$X_{n,k} = \sum_{j=0}^{k} C_j^n (1-X)^j X^{n-j} \Rightarrow X_{n,k} = (-1)^k (n-k) C_k^n \sum_{m=0}^{K} (-1)^m \frac{1}{n-m} C_m^k X^{n-m}$$

$$(G.1)$$

__Proof__:

$$\sum_{j=0}^{k} C_j^n (1-X)^j X^{n-j} = \sum_{j=0}^{k} C_j^n X^{n-j} \sum_{m=0}^{j} (-1)^{j-m} C_m^j X^{j-m}$$

$$= \sum_{0 \le m \le j \le k} (-1)^{j-m} C_j^n C_m^j X^{n-m}$$

Now

$$C_j^n C_m^j = \frac{n!}{j!(n-j)!} \frac{j!}{(j-m)!m!} = \frac{n!}{m!} \frac{1}{(j-m)!((n-m)-(j-m))!}$$

$$C_m^n C_{j-m}^{n-m} \qquad\qquad (G.2)$$

Let $r = j - m$ replace j

$$0 \le m \le j \le k \equiv 0 \le m \le k \text{ and } 0 \le r \le k-m$$

So

$$X_{n,k} = \sum_{m=0}^{k} C_m^n X^{n-m} \sum_{r=0}^{k-m} (-1)^r C_r^{n-m}$$

But Eq. 24.1.4 of Ref. 6 is

$$\sum_{r=0}^{k-m} (-1)^r \, C_r^{n-m} = (-1)^{k-m} \, C_{k-m}^{n-m-1}$$

$$\therefore \quad X_{n,k} = \sum_{m=0}^{k} (-1)^{k-m} \, C_{k-m}^{n-m-1} \, C_m^n \, x^{n-m}$$

We may use Eq. G. 2 in the following

$$C_m^n \, C_{k-m}^{n-m-1} = C_m^n \, C_{k-m}^{n-m} \, \frac{n-k}{n-m} = C_k^n \, C_m^k \, \frac{n-k}{n-m}$$

Hence

$$X_{n,k} = \sum_{m=0}^{k} (-1)^{k-m} \, \frac{n-k}{n-m} \, C_k^n \, C_m^k \, x^{n-m}$$

$$X_{n,k} = (-1)^k \, (n-k) \, C_k^n \, \sum_{m=0}^{k} (-1)^m \, \frac{1}{n-m} \, C_m^k \, x^{n-m}$$

$$\tag{G. 3}$$

Remark 2

$$X_{n,k+1} - X_{n,k} = (-1)^{k+1} \, (n-k) \, C_k^n \, \sum_{m=0}^{k} (-1)^m \, \frac{1}{k+1-m} \, C_m^k \, x_1^{n-m} + C_{k+1}^n \, x_1^{n-k-1}$$

$$\tag{G. 4}$$

Proof :

$$X_{n,k+1} - X_{n,k} = C_{k+1}^n \, (1 - x_1)^{k+1} \, x_1^{n-k-1}$$

$$= C_{k+1}^n \, \sum_{m=0}^{k+1} (-1)^{k+1-m} \, C_m^{k+1} \, x_1^{n-m}$$

$$= (-1)^{k+1} \, \sum_{m=0}^{k} (-1)^m \, C_{k+1}^n \, C_m^{k+1} \, x_1^{n-m} + C_{k+1}^n \, x_1^{n-k-1}$$

Since

$$C_{k+1}^{n} \, C_{m}^{k+1} = C_{k}^{n} \, \frac{n-k}{k+1} \, C_{m}^{k} \, \frac{k+1}{k+1-m} = C_{k}^{n} \, C_{m}^{k} \, \frac{n-k}{k+1-m}$$

$$X_{n,\,k+1} - X_{n,\,k} = (-1)^{k+1} \, (n-k) \, C_{k}^{n} \sum_{m=0}^{k} (-1)^{m} \, C_{m}^{k} \, X_{1}^{n-m} \, \frac{1}{k+1-m}$$

$$+ \, C_{k+1}^{n} \, X_{1}^{n-1-k} \qquad\qquad \text{Q. E. D}$$

Remark 3

$$X_{n,\,k} + \frac{k+1}{n} (X_{n,k+1} - X_{n,\,k}) = (-1)^{k} (n-1-k) \, C_{k}^{n-1} \sum_{m=0}^{k} (-1)^{m} \, \frac{1}{n-1-r} \, C_{m}^{k} \, X_{1}^{n-1-m}$$

$$(\, \text{G. 5})$$

Proof:

$$X_{n,\,k} + \frac{k+1}{n} (X_{n,\,k+1} - X_{n,\,k}) = (-1)^{k} (n-k) \, C_{k}^{n} \sum_{m=0}^{k} (-1)^{m} X_{1}^{n-m}$$

$$\cdot C_{m}^{k} \left\{ \frac{1}{n-m} - \frac{k+1}{n(k+1-m)} \right\} + \frac{k+1}{n} \, C_{k+1}^{n} \, X_{1}^{n-k-1}$$

$$\left\{ \frac{1}{n-m} - \frac{k+1}{n(k+1-m)} \right\} = \frac{n(k+1) - nm - n(k+1) + m(k+1)}{n(n-m)(k+1-m)}$$

$$= \frac{-m(n-k-1)}{n(n-m)(k+1-m)}$$

For $m = 0$ this term is zero; the sum now runs from one to k.

$$X_{n,k} + \frac{k+1}{n}(X_{n,k+1} - X_{n,k}) = (-1)^k (n-k-1) \, C_k^n \sum_{m=1}^{k} (-1)^{m+1}$$

$$\cdot \frac{(n-k)m}{(n-m)(k+1-m)n} \, C_m^k \, X_1^{n-m} + \frac{k+1}{n} \, C_{k+1}^n \, X_1^{n-k-1}$$

Now

$$C_k^n \frac{(n-k)m}{(n-m)(k+1-m)n} \, C_m^k = \frac{n!(n-k)m \; k!}{k!(n-k)!(n-m)(k+1-m)n \, m!(k-m)!}$$

$$= \frac{(n-1)! \; k!}{k!(n-k-1)!(n-m)(k+1-m)!(m-1)!} = C_k^{n-1} \frac{1}{(n-1)-(m-1)} C_{m-1}^k$$

$$\frac{k+1}{n} \, C_{k+1}^n = \frac{(k+1)n!}{n(n-k-1)!(k+1)!} = \frac{(n-1)!}{k!(n-1-k)!} = C_k^{n-1}$$

$$X_{n,k} + \frac{k+1}{n}(X_{n,k+1} - X_{n,k}) = (-1)^k (n-k-1) \, C_k^{n-1} \sum_{m=1}^{k} (-1)^{m+1}$$

$$\cdot \frac{1}{(n-1)-(m-1)} \, C_{m-1}^k X_1^{(n-1)-(m-1)} + C_k^{n-1} X_1^{n-1-k}$$

$$= (-1)^k (n-1-k) \, C_k^{n-1} \sum_{r=0}^{k-1} (-1)^r \frac{1}{(n-1)-r} \, C_r^k \, X_1^{n-1-r}$$

$$+ C_k^{n-1} X_1^{n-1-k}$$

$$= (-1)^k (n-1-k) \, C_k^{n-1} \sum_{r=0}^{k} (-1)^r \frac{1}{(n-1)-r} \, C_r^k \, X_1^{n-1-r}$$

Q. E. D.

REFERENCES

1. Peterson, W. W., Birdsall, T. G., and Fox, W. C., The theory of signal detectability. Trans. IRE Professional Group on Information Theory, 1954, PGIT-4, pp. 171-211. Also Peterson, W.W. and Birdsall, T. G. The theory of signal detectability. Electronic Defense Group Technical Report No. 13. Ann Arbor: Electronic Defense Group, The University of Michigan, 1953.

2. Pitt, H. R., Integration, measure and probability. Edinburgh and London: Oliver and Boyd, 1963.

3. Halsted, L. R, Birdsall, T. G., and Nolte, L. W., On the detection of a randomly distorted signal in Gaussian noise. Cooley Electronics Technical Report No. 129. Ann Arbor: Cooley Electronics Laboratory, The University of Michigan, 1962.

4. Roberts, R. A., Theory of signal detectability: composite deferred decision theory. Cooley Electronics Technical Report No. 161. Ann Arbor: Cooley Electronics Laboratory, The University of Michigan, 1965.

5. Cramér, H., Mathematical methods of statistics. Princeton University Press, 1946. First published in Sweden, Almquist and Wiksells, Uppsala, 1945.

6. Abramowitz, M. and Stegun, I. A. (Ed.), Handbook of mathematical functions with formulas, graphs, and mathematical tables. National Bureau of Standards, Department of Commerce, 1964.

7. Middleton, D., An introduction to statistical communication theory. New York, Toronto, London: McGraw-Hill, 1960.

8. Swets, J. A. (Ed.), Signal detection and recognition by human observers. Contemporary readings. New York, London, Sydney: John Wiley and Sons, 1964.

9. Quastler, H. (Ed.), Information theory in psychology. Problems and methods. Glencoe, Illinois: The Free Press, 1955.

10. Tanner, W. P., Jr., and Birdsall, T. G., Definitions of d' and η as psychophysical measures. JASA. 1958, Vol. 30, pp. 922-928. Also Ref. 8, p. 147.

11. Tanner, W. P., Jr., and Swets, J. A., The human use of information I. Trans. IRE Professional Group on Information Theory, 1954, PGIT-4, pp. 213-221.

12. Green, D. M., Psychoacoustics and detection theory. JASA, 1960, Vol. 32, pp. 1189-1203.

13. Tanner, W. P., Jr., Swets, J. A. and Green, D. M., Some general properties of the hearing mechanism. Electronic Defense Group Technical Report No. 30. Ann Arbor: Electronic Defense Group, The University of Michigan, 1956.

14. Pollack, I., On indices of signal and response discriminability. JASA, 1959, Vol. 31, pp. 1031.

15. Helstrom, C. W., Statistical theory of signal detection. New York: Pergamon Press, 1960.

16. Marcum, J. I., A statistical theory of target detection by pulsed radar. RAND Research Memo RM-754. Santa Barbara, California: The RAND Corporation, 1947. Also reprinted in IRE Trans. on Information Theory, 1960, Vol. IT-6, pp. 59-144.

17. Elliot, P. B., Tables of d'. Electronic Defense Group Technical Report No. 97. Ann Arbor: Electronic Defense Group, The University of Michigan, 1959. Also Ref. 8, pp. 651-684.

18. Blosser, A. B., A performance-oriented approach to detection: tables for detection, discrimination, and decision theory. Tracor Inc. Report No. 65-267-U. Austin, Texas: Tracor, Inc., 1965.

19. Rice, S. O., Mathematical analysis of random noise. BSTJ, 1945-1946, Vol. 23, pp. 282-332, Vol. 24, pp. 46-156.

20. Jeffress, L. A., Stimulus-oriented approach to detection. JASA, 1964, Vol. 36, pp. 766-774.

21. Birdsall, T. G. and Lamphiear, D. E., Approximations to the non-central chi-square distribution with applications to signal detection models. Electronic Defense Group Technical Report No. 101. Ann Arbor: Electronic Defense Group, The University of Michigan, 1960.

22. Ristenbatt, M. P. and Birdsall, T. G., ROC curves. Notes for EE 534. Ann Arbor: The University of Michigan, 1958.

23.   Egan, J. P. , Greenberg, G. Z. , and Schulman, A. I., Operating characteristics, signal detectability, and the method of free response. JASA, 1961, Vol. 33, pp. 993-1007. Also Ref. 8, pp. 316-347.

24.   Nolte, L. W. , and Jaarsma, D. , Detectability of recurrence phenomena. Cooley Electronics Laboratory Technical Report No. 179. Ann Arbor: Cooley Electronics Laboratory, The University of Michigan, 1966.

25.   Birdsall, T. G. and Roberts, R. A. , Theory of signal detectability: observation-decision procedures. Cooley Electronics Laboratory Technical Report No. 136. Ann Arbor: Cooley Electronics Laboratory, The University of Michigan, 1964.

26.   Swets, J. A. , Tanner, W. P. , Jr. , and Birdsall, T. G., The evidence for a decision-making theory of visual detection. Electronic Defense Group Technical Report No. 40. Ann Arbor: Electronic Defense Group, The University of Michigan, 1955.

27.   Swets, J. A., Information retrieval systems. Science, 1963, Vol. 141, No. 3577, pp. 245-250.

28.   Carver, H. C. (Ed.), Mathematical statistical tables. Ann Arbor: Edwards Brothers, Inc. , 1950.

29.   Green, D. M. , General prediction relating yes-no and forced choice results. JASA, 1964, Vol. 36, p. 1042(A).

30.   Parzen, E. , Stochastic processes. San Francisco: Holden Day, Inc. , 1962, p. 12.

31.   Kullback, S. and Leibler, R. A. , On information and sufficiency. Annals of Mathematical Statistics, 1951, Vol. 22, No. 1, pp. 79-86.

32.   Egan, J. P. , Message reception, operating characteristics, and confusion matrices in speech communication. Hearing and Communications Laboratory AFCRC TR-57-50. Bloomington: Hearing and Communications Laboratory, Indiana University, 1957. (AD-110064).

33.   Watson, C. S. , Rilling, M. E. , and Bourbon, W. T. , Receiver operating characteristics determined by a mechanical analog to the rating scale. JASA, 1964, Vol. 36, pp. 283-288.

34.   Egan, J. P., Schulman, A. E. and Greenberg, G. Z., Operating
        characteristics determined by binary decisions and by ratings.
        JASA, 1959, Vol. 31, pp. 768-773.

35.   Luce, R. D., A threshold theory for simple detection experiments.
        Psychol. Review, 1963, Vol. 70, pp. 61-79.

36.   Faran, J. J., Jr. and Hills, R., Jr., Correlators for signal reception.
        Harvard Acoustical Laboratory Technical Memo No. 27. Cambridge:
        Harvard Acosutical Laboratory, Harvard University. 1952, pp. 58-65.

DISTRIBUTION LIST

No. of Copies

Office of Naval Research (Code 468)	1
(Code 102-OS)	1
(Code 480)	1
Navy Department	
Washington, D. C. 20360	
Director, Naval Research Laboratory	6
Technical Information Division	
Washington, D. C. 20390	
Director	1
Office of Naval Research Branch Office	
1030 East Green Street	
Pasadena, California 91101	
Office of Naval Research	1
San Francisco Area Office	
1076 Mission Street	
San Francisco, California 94103	
Director	1
Office of Naval Research Branch Office	
495 Summer Street	
Boston, Massachusetts 02210	
Office of Naval Research	1
New York Area Office	
207 West 24th Street	
New York, New York 10011	
Director	1
Office of Naval Research Branch Office	
536 S. Clark Street	
Chicago, Illinois 60605	
Director	8
Naval Research Laboratory	
Attn: Library, Code 2029 (ONRL)	
Washington, D. C. 20390	

DISTRIBUTION LIST (Cont.)

No. of Copies

Commander Naval Ordnance Laboratory
Acoustics Division
White Oak, Silver Spring, Maryland 20907
    Attn: Dr. Zaka Slawsky

1

Commanding Officer
Naval Ship Research & Development Center
Annapolis, Maryland 21401

1

Commander
Naval Undersea Research & Development Center
San Diego, California 92132
    Attn: Dr. Dan Andrews
          Mr. Henry Aurand

2

Chief Scientist
Navy Underwater Sound Reference Division
P. O. Box 8337
Orlando, Florida 32800

1

Commanding Officer and Director
Navy Underwater Systems Center
Fort Trumbull
New London, Connecticut 06321

1

Commander
Naval Air Development Center
Johnsville, Warminster, Pennsylvania 18974

1

Commanding Officer and Director
Naval Ship Research and Development Center
Washington, D. C. 20007

1

Superintendent
Naval Postgraduate School
Monterey, California 93940

1

Commanding Officer & Director
Naval Ship Research & Development Center*
Panama City, Florida 32402

1

---

*Formerly Mine Defense Lab.

DISTRIBUTION LIST (Cont.)

No. of Copies

Naval Underwater Weapons Research & Engineering
      Station   1
Newport, Rhode Island 02840

Superintendent   1
Naval Academy
Annapolis, Maryland 21401

Scientific and Technical Information Center   2
4301 Suitland Road
Washington, D. C. 20390
         Attn: Dr. T. Williams
               Mr. E. Bissett

Commander   1
Naval Ordnance Systems Command
Code ORD-03C
Navy Department
Washington, D. C. 20360

Commander   1
Naval Ship Systems Command
Code SHIPS 037
Navy Department
Washington, D. C. 20360

Commander   2
Naval Ship Systems Command
Code SHIPS 00V1
Washington, D. C. 20360
        Attn: CDR Bruce Gilchrist
              Mr. Carey D. Smith

Commander   1
Naval Undersea Research & Development Center
3202 E. Foothill Boulevard
Pasadena, California 91107

Commanding Officer   1
Fleet Numerical Weather Facility
Monterey, California 93940

DISTRIBUTION LIST (Cont.)

	No. of Copies
Defense Documentation Center Cameron Station Alexandria, Virginia 22314	12
Dr. James Probus Office of the Assistant Secretary of the Navy (R&D) Room 4E741, The Pentagon Washington, D. C. 20350	1
Mr. Allan D. Simon Office of the Secretary of Defense DDR&E Room 3E1040, The Pentagon Washington, D. C. 20301	1
Capt. J. Kelly Naval Electronics Systems Command Code EPO-3 Washington, D. C. 20360	1
Chief of Naval Operations Room 5B718, The Pentagon Washington, D. C. 20350 Attn: Mr. Benjamin Rosenberg	1
Chief of Naval Operations Rm 4C559, The Pentagon Washington, D. C. 20350 Attn: CDR J. M. Van Metre	1
Chief of Naval Operations 801 No. Randolph St. Arlington, Virginia 22203	1
Dr. Melvin J. Jacobson Rensselaer Polytechnic Institute Troy, New York 12181	1
Dr. Charles Stutt General Electric Co. P. O. Box 1088 Schenectady, New York 12301	1

DISTRIBUTION LIST (Cont.)

No. of Copies

Dr. Alan Winder   1
EDO Corporation
College Point, New York 11356

Dr. T. G. Birdsall   1
Cooley Electronics Lab.
The University of Michigan
Ann Arbor, Michigan 48105

Mr. Morton Kronengold   1
Director, Institute for Acoustical Research
615 S.W. 2nd Avenue
Miami, Florida 33130

Mr. Robert Cunningham   1
Bendix Corporation
11600 Sherman Way
North Hollywood, California 91606

Dr. H. S. Hayre   1
University of Houston
Cullen Boulevard
Houston, Texas 77004

Dr. Robert R. Brockhurst   1
Woods Hole Oceanographic Institute
Woods Hole, Massachusetts 02543

Dr. Stephen Wolff   1
Johns Hopkins University
Baltimore, Maryland 21218

Dr. M. A. Basin   1
Litton Industries
8000 Woodley Avenue
Van Nuys, California 91409

Dr. Albert Nuttall   1
Navy Underwater Systems Center
Fort Trumbull
New London, Connecticut 06320

DISTRIBUTION LIST (Cont.)

<div align="right">No. of Copies</div>

Dr. Philip Stocklin
Raytheon Company
P. O. Box 360
Newport, Rhode Island 02841

<div align="right">1</div>

Dr. H. W. Marsh
Navy Underwater Systems Center
Fort Trumbull
New London, Connecticut 06320

<div align="right">1</div>

Dr. David Middleton
35 Concord Ave., Apt. #1
Cambridge, Massachusetts 02138

<div align="right">1</div>

Mr. Richard Vesper
Perkin-Elmer Corporation
Electro-Optical Division
Norwalk, Connecticut 06852

<div align="right">1</div>

Dr. Donald W. Tufts
University of Rhode Island
Kingston, Rhode Island 02881

<div align="right">1</div>

Dr. Loren W. Nolte
Dept. of Electrical Engineering
Duke University
Durham, North Carolina 27706

<div align="right">1</div>

Dr. Thomas W. Ellis
Texas Instruments, Inc.
13500 North Central Expressway
Dallas, Texas 75231

<div align="right">1</div>

Mr. Robert Swarts
Honeywell, Inc.
Marine Systems Center
5303 Shilshole Ave., N.W.
Seattle, Washington, 98107

<div align="right">1</div>

Mr. Charles Loda
Institute for Defense Analyses
400 Army-Navy Drive
Arlington, Virginia 22202

<div align="right">1</div>

DISTRIBUTION LIST (Cont.)

No. of Copies

Mr. Beaumont Buck   1
General Motors Corporation
Defense Research Division
6767 Holister Ave.
Goleta, California 93017

Dr. M. Weinstein   1
Underwater Systems, Inc.
8121 Georgia Avenue
Silver Spring, Maryland 20910

Dr. Harold Saxton   1
1601 Research Blvd.
TRACOR, Inc.
Rockville, Maryland 20850

Dr. Thomas G. Kincaid   1
General Electric Company
P. O. Box 1088
Schenectady, New York 12305

Applied Research Laboratories   3
The University of Texas at Austin
Austin, Texas 78712
    Attn: Dr. Loyd Hampton
          Dr. Charles Wood

Dr. Paul McElroy   1
Woods Hole Oceanographic Institution
Woods Hole, Massachusetts 02543

Dr. John Bouyoucos   1
Hydroacoustics, Inc.
P. O. Box 3818
Rochester, New York 14610

Hydrospace Research Corporation   1
5541 Nicholson Lane
Rockville, Maryland 20852
    Attn: CDR Craig Olson

Cooley Electronics Laboratory   100
University of Michigan
Ann Arbor, Michigan 48105

DISTRIBUTION LIST (Cont.)

No. of Copies

Mr. Ray Veenkant   1
Texas Instruments, Inc.
North Central Expressway
Dallas, Texas 75222
Mail Station 208

Dr. Joseph Lapointe   1
Systems Control, Inc.
260 Sheridan Ave.
Palo Alto, Calif. 94306

Dr. Bruce P. Bogert   1
Bell Telephone Laboratories
Whippany Road
Whippany, New Jersey 07981

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)*	2a. REPORT SECURITY CLASSIFICATION
Cooley Electronics Laboratory	Unclassified
The University of Michigan	2b. GROUP
Ann Arbor, Michigan 48105	

3. REPORT TITLE

The Theory of Signal Detectability: ROC Curves and Their Character

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*

Technical Report No. 177 - January 1973

5. AUTHOR(S) *(First name, middle initial, last name)*

Theodore G. Birdsall

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
January 1973	450	36

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
N00014-67-A-9181-0032	036040-16-T
b. PROJECT NO.	
c.	9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*
d.	TR177

10. DISTRIBUTION STATEMENT

Approved for public release; distribution unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	Office of Naval Research
	Department of the Navy
	Arlington, Va. 22217

13. ABSTRACT The first problem in the theory of signal detectability deals with the decision between two alternative responses, corresponding to two possible classes of causes of an observation. When the goal of a decision process is to achieve the highest quality of terminal decision, the Receiver Operating Characteristic curve (ROC curve) contains all of the information necessary for the evaluation of the decision process. This present work introduces the ROC character, which is isomorphic to the ROC curve. The formal development is based on two key facts. The first is the fundamental theorem: if $\ell(X)$ is the likelihood ratio of an observation, then the likelihood ratio of ℓ is ℓ itself. The second is the main theorem on ROC characters: each ROC character is isomorphic to a univariate probability distribution that possesses a moment generating function. The character convolution theorem and the character addition theorem follow directly from these. Families of ROC curves are developed from the main theorem on ROC characters. The normal, binormal, Q-table, power, and several discrete families of ROC curve have appeared in the literature. The new families include the Pearson type III, Fisher-Tippett doubly exponential, H-type, Poisson, and the regular conics. Additional families are generated from these by use of the metastatic relation, and the convolution and addition theorems. ROC curves contain information about the performance of other two-cause decisions besides the two-response decision. Several are considered that are used in the testing of human perception; namely, the symmetric forced choice decision, type II decisions, the rating scale procedure, and the analysis of a decision based on the reports of multiple observers.

DD FORM 1473
NOV 65

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Receiver operating characteristic curve Likelihood ratio Signal detection theory ROC characters Detection-decision devices						