DOCTORAL DISSERTATION SERIES

TITLE Sample Criteria for Testing Outlying Observations

AUTHOR Frank E. Grubbs

UNIVERSITY Of Michigan DATE 1948

DEGREE Ph.D. PUBLICATION NO. 1244

UNIVERSITY MICROFILMS
ANN ARBOR · MICHIGAN

SAMPLE CRITERIA FOR TESTING OUTLYING OBSERVATIONS

by

Frank E. Grubbs

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in the University of Michigan.

Committee in Charge:

Professor Cecil C. Craig, Chairman
Professor S. S. Attwood
Professor H. C. Carver
Professor A. H. Copeland
Professor G. Y. Rainich

December 1948

Table of Contents

Table of Contents
(Cont'd.)

Sample Criteria For Testing Outlying Observations

1. Introduction. Scientific data are collected usually for purposes of interpretation and if proper use is to be made of the information thus obtained then some decision should be reached or some action taken as a result of analyzing the data. In many cases a critical examination of the data collected is necessary in order to insure that the results of sampling are representative of the thing or process we are examining. Quite frequently our observations do not appear to be consistent with one another, i.e. the data may seem to display non-homogeneities and the group of observations as a whole may not appear to represent a random sample from, say, a single normal population or universe. In particular, one or more of the observations may have the appearance of being "outliers" and we are interested here in determining once and for all whether such observations should be retained in the sample for interpreting results or whether they should be regarded as being inconsistent with the remaining observations. It is clear that rejection of the "outliers" in a sample will in a great number of cases lead to a different course of action than would have been taken had such observations been retained in the sample. Actually, the rejection of "outlying" observations may be just as much a practical (or common sense) problem as a statistical one and sometimes the practical or experimental viewpoint may naturally outweigh any statistical contributions. In this connection, the concluding remarks of Rider's survey [2] are pertinent: "In the final analysis it would seem that the question of the rejection or the retention of a discordant observation reduces to a question of common sense. Certainly the judgement of an experienced observer should be allowed considerable influence in reaching a decision. This judgement can undoubtedly be aided by the application of one or more tests based on the theory of probability, but any test which requires an inordinate amount of calculation seems hardly to be worth while, and the testimony of any criterion which is based upon a complicated hypothesis should be accepted with extreme caution." Hence, it would

appear that statistical tests of significance for judging or testing "outliers" come into importance either in supporting doubtful practical viewpoints or in providing a basis for action in the absence of sufficient experimental knowledge of underlying causes in an investigation. Indeed, the latter two situations are met quite frequently in practice.

In the present treatment, we intend to throw some light beyond the work that has already been done [1], [2], [3], [4], [11], [12] on the problem of testing outlying observations statistically and to see just where our contributions fit into this corner of mathematical statistics. In the course of our investigation, we take into consideration the idea of efficient tests for various hypotheses according to the theory of Neyman and Pearson [17] and present also derivations of several sampling distributions for statistics which have an important bearing on the problem of testing outlying observations. First, however, we give a very brief history of the problem.

2. Historical Comments. A survey of statistical literature indicates that the problem of testing the significance of outlying observations received considerable attention prior to 1937. Since this date, however, published literature on the subject seems to have been unusually scant—perhaps because of inherent difficulties in the problem as pointed out by E.S. Pearson and C. Chandra-Sekar [1]. These authors made some important contributions to the problem of outlying observations by bringing clearly into the foreground the concept of efficiency of tests which may be used in view of admissible alternative hypotheses (to be discussed later).

In 1933, P.R. Rider [2] published a rather comprehensive survey of work on the problem of testing the significance of outlying observations up to that date. The test criteria surveyed by Rider appear to impose as an initial condition that the standard deviation, σ, of the population from which the items were drawn should be known accurately. In connection with such tests requiring accurate knowledge of σ, we mention (1) Irwin's criteria [3] which utilize the difference between the first two individuals or the difference between the second and third individuals in random samples from a normal population and (2) the range* or maximum dispersion [4], [5], [6], [7], [8], [9], [10], [18] of a sample which has been advocated by "Student" [4] and others for testing the significance of outlying observations. We remark further that a natural statistic to use for testing an "outlier" is the difference between such an extreme observation and the sample mean. In 1935, McKay [11] published a note on the distribution of the last-mentioned statistic and by means of a rather elaborate procedure obtained a recurrence relation between the distribution of the extreme minus the mean in samples of n from a normal universe and the distribution of this statistic in samples of n-1 from the same parent. McKay gave also an approximate expression for the upper percentage points of the distribution but did not tabulate the exact distribution due to the complicity of the multiple integrals involved. McKay pointed out that if K_p denotes the p-th semi-invariant of the distribution of $x_n-\bar{x}$ (where x_n is the largest observation) and K_p' refers similarly to the distribution of x_n, then $K_1=K_1'-\mu$, $K_2=K_2'-\frac{1}{n}$ and $K_p=K_p'$ $(p \geq 3)$ where $\mu = E(x_i)$. The author was not aware of the work of A.T. McKay when the simplified derivation for the distribution of $x_n-\bar{x}$ given in section 3 below was worked out, McKay's result being called to his attention by C.C. Craig.

Under certain circumstances, accurate knowledge concerning σ may be available as, for example, in using "daily control" tests [4], [18] the population standard deviation may be estimated in some cases with sufficient precision from past data. In general, however, an accurate estimate of σ may not be available and it becomes necessary to estimate the population standard deviation from the single sample involved or "Studentize" [18] the statistic to be used, thus providing a true measure of the risks involved in the significance test advocated for testing outlying observations.

---

* The derivation for the exact distribution of the range is given in reference [9], 1942; however, Dr. L.S. Dederick of the Ballistic Research Laboratory also derived the exact distribution of the range in an unpublished Aberdeen Proving Ground Report (1926).

W.R. Thompson [12] apparentlyhad this very point in mind when he devised an exact test in his paper, "On a Criterion for the Rejection of Observations and the Distribution of the Ratio of the Deviation to the Sample Standard Deviation", which appeared in 1935. Thompson showed that if

$$T_i = \frac{x_i - \bar{x}}{s}$$

where $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$, $s^2 = \frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x})^2$ and x_i is an observation selected arbitrarily from a random sample of n items drawn from a normal parent, then the probability density function of

$$t = \frac{T\sqrt{n-2}}{\sqrt{n-1-T^2}}$$

is given by

$$p(t) = \frac{\Gamma(\frac{n-1}{2})}{\sqrt{(n-2)\pi}\ \Gamma(\frac{n-2}{2})} \left(1 + \frac{t^2}{n-2}\right)^{-\frac{n-1}{2}}$$

i.e. "Student's" t-distribution with $f = n-2$ degrees of freedom. For

$$P = P\left\{|T| > T_o\right\} = P\left\{|T| > \frac{t_o\sqrt{n-1}}{n-2+t_o^2}\right\} ,$$

Thompson fixes $\varphi = nP$, <u>the frequency of rejection per sample</u> and tabulates the limits T_o and t_o for values of $\varphi = 0.2$, 0.1 and 0.05. It is to be noted that for a given value of φ and criterion of rejection,

$$|x_i - \bar{x}| \geq T_o\, s ,$$

then Thompson's criterion rejects on the average one observation in every $\frac{1}{\varphi}$ samples unnecessarily and this figure will be the same regardless of the sample size n. A more customary practice would have been to fix P at the values, say, 0.2, 0.1, 0.05, etc., in order that the probability of rejection under the null-hypothesis of sampling a single normal parent remains fixed regardless of sample size.

Pearson and Chandra Sekar have given a rather comprehensive study of Thompson's criterion in an interesting and important paper [1] which appeared in 1936. They discussed also some very important viewpoints which should be taken into consideration when dealing with the problem of testing outlying observations. It consequently appears appropriate at this point to summarize briefly the comments of Pearson and Chandra Sekar on Thompson's paper. These authors show that if $|T_1|$, $|T_2|$, ... $|T_n|$ represent the n values of

$$T_i = \frac{x_i - \bar{x}}{s}$$

in a sample arranged in descending order of absolute magnitude, then for certain sample sizes the maximum values of $|T_i|$ for $i \geq 2$ lie within the significance levels of Thompson's test! This indicates a weakness of Thompson's criterion to detect outlying observations in some circumstances. By setting up alternatives to the null-hypothesis H_0 that all items in the sample come from the same population, Pearson and Chandra Sekar point out that if only one of the observations actually came from a population with divergent mean, then Thompson's criterion would be very useful, whereas if two or more of the observations are truly outlying then the criterion $|x_i - \bar{x}| \geq T_0$ s may be quite ineffective, particularly if the sample contains less than about 30 or 40 observations.

A point of major interest concerning Thompson's work nevertheless is that he proposed an exact test for the hypothesis that all of the observations came from the same normal population. With regard to the use of an arbitrary observation in Thompson's test, however, it should be borne in mind that the problem of finding the probability that an arbitrary observation will be outlying is different from that of finding the probability that a single observation from a population with mean $\mu + \lambda\sigma$ will be outlying with respect to a set of n-1 observations from a normal universe with mean μ.

Returning now to the paper of Pearson and Chandra Sekar [1], we find that for the n values of T_i arranged in order of magnitude taking account of sign, say

$$T^{(1)}, \ T^{(2)}, \ \ldots \ T^{(n)},$$

then

$$T^{(1)} \geq T^{(2)} \geq T^{(3)} \ \ldots \ \geq T^{(n)}$$

and the above authors show that the form of the total distribution of all the T_i at its extremes depend only on $T^{(1)}$ and $T^{(n)}$. This is because for some combinations of sample size and percentage points the algebraic upper limit for $T^{(2)}$ and algebraic lower limit for $T^{(n-1)}$ do not extend into the "tails" of the total distribution. Hence, the following probability law holds for $T^{(1)}$ when $T^{(1)} \geq$ the algebraic maximum of $T^{(2)}$:

$$p\left\{T^{(1)}\right\} = Np(T) \quad .$$

Likewise,

$$p\left\{T^{(n)}\right\} = Np(T)$$

for $T^{(n)} \leq$ algebraic minimum of $T^{(n-1)}$. Therefore, since $\varphi = NP$, Pearson and Chandra Sekar were able to use Thompson's table [12] and give (for some sample sizes) upper probability limits for $T^{(1)} = \dfrac{x_i - \bar{x}}{s}$ for the highest observation and lower probability limits for $T^{(n)} = \dfrac{x_i - \bar{x}}{s}$ for the lowest observation without actually obtaining the exact probability distribution of $T^{(1)}$ and $T^{(n)}$. Hence, the appearance of the table of percentage points on page 318 of their paper [1] was a substantial contribution to the problem of testing outlying observations since an exact test for the significance of a single outlying observation was provided for the case where an accurate estimate of σ is not available. [The exact distribution of $T^{(1)}$ or $T^{(n)}$ is derived later in this work.]

With the above highlights of historical background in mind, we turn now to a consideration of the types of problems the experimenter may be faced with in testing "outlying" observations.

3. <u>Statement of Hypotheses in Tests of Outliers</u>. Once the sample results of an experiment are available, the practicing statistician is confronted with one or more of the following distinct situations as regards discordant observations: (a) To begin with, a very frequent or perhaps prevalent situation is that either the greatest observation or the least observation in a sample may have the appearance of belonging to a different population than the one from which the remaining observations were drawn. Here we are confronted with tests for a single outlying observation. (b) Then again, both the largest and the smallest observations may appear to be "different" from the remaining items in the sample. Here we are interested in testing the hypothesis that both the largest and the smallest observations are truly "outliers". (c) Another frequent situation is that either the two largest or the two smallest observations may have the appearance of being discordant. Here we are interested in reaching a decision as to whether we should reject the two largest or the two smallest observations as not being representative of the thing we are sampling.

In this work, we will not be concerned generally with more than two outlying observations (although some of our theory may be extended to more than two outliers) since for such cases in dealing with small samples the data will probably be so heterogeneous that it would be desirable to conduct a fundamental investigation of our experimental technique or product under investigation.

As to why the discordant observations in a sample may be outliers, this may be due to errors of measurement in which case we would naturally want to reject or at least "correct" such observations. On the other hand, it may be that the population we are sampling is not homogeneous in the uni-modal sense and it will consequently be desirable to know this so that we may carry out further development work on our product if possible or desirable.

In what follows, we will examine the problem from the standpoint of testing whether the observations appear to have been drawn from more than one normal population. As a useful practical assumption, we will postulate that each of the observations is subject to the same standard error even though they may come from different normal populations. We shall consider two cases: one in which the standard deviation σ of the normal populations sampled is known accurately and the other in which σ must in effect be inferred from the sample.

4. Sample Criteria Based Upon the Principle of Likelihood and the Theory of Best Tests in the Neyman-Pearson Sense [17]. We now turn to the problem of determining optimum sample statistics to use in testing "outlying" observations for the situations described in the preceding section. As mentioned above, we will restrict our investigation to normal populations which have the same variance (known or unknown) but which may have different true means as specified below.

The general probability function for normal samples takes on the form

$$P = (\sqrt{2\pi}\sigma)^{-n} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-a_i)^2}$$

where we may or may not have $a_i = a_j$ for $i \neq j$. No generality will be lost by specifying in this section that $x_1 \leq x_2 \leq x_3 \leq \ldots \leq x_n$ since the likelihood function of such a sample is simply $n!$ times P.

(a) Sample Criteria for a Single Outlier— σ Assumed Known $= \sigma_o$. We are interested here in testing the composite hypothesis H_o that $a_i=a$ for $i=1, 2, \ldots n$ against the hypothesis H_1 that $a_n \neq a$. In particular, we would like to know whether $a_n = a + \lambda\sigma_o$, $\lambda > 0$. In order to find sample criteria based upon the principle of likelihood, we obtain (1) P_ω, by substituting in P those values of the a_i which maximize P with respect to H_o and (2) P_Ω, by substituting in P those values of the a_i which maximize P with respect to all admissible values of the population parameters under the more general hypothesis H_1 and then take the ratio of P_ω to P_Ω.

Under H_o,

$$P = P_{H_o} = (\sqrt{2\pi}\,\sigma_o)^{-n} e^{-\frac{1}{2\sigma_o^2}\sum_{i=1}^{n}(x_i-a)^2}$$

$$\frac{\partial P_{H_o}}{\partial a} = 0 \quad \text{gives} \quad a = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x}$$

so that

$$P_{\omega} = (\sqrt{2\pi}\,\sigma_o)^{-n}\, e^{-\frac{1}{2\sigma_o^2}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

To determine P_{Ω}, we have

$$P_{H_1} = (\sqrt{2\pi}\,\sigma_o)^{-n}\, e^{-\frac{1}{2\sigma_o^2}\left\{\sum_{i=1}^{n-1}(x_i - a)^2 + (x_n - a_n)^2\right\}}$$

Now

$$\frac{\partial P_{H_1}}{\partial a} = 0 \quad \text{gives} \quad a = \frac{1}{n-1}\sum_{i=1}^{n-1} x_i = \bar{x}_n, \quad \text{say}$$

$$\frac{\partial P_{H_1}}{\partial a_n} = 0 \quad \text{gives } a_n = x_n.$$

so that

$$P_{\Omega} = (\sqrt{2\pi}\,\sigma_o)^{-n}\, e^{-\frac{1}{2\sigma_o^2}\sum_{i=1}^{n-1}(x_i - \bar{x}_n)^2}$$

$$= (\sqrt{2\pi}\,\sigma_o)^{-n}\, e^{-\frac{1}{2\sigma_o^2}\left\{\sum_{i=1}^{n}(x_i - \bar{x})^2 - \frac{n}{n-1}(x_n - \bar{x})^2\right\}}$$

Hence,

$$\nu = \frac{P_{\omega}}{P_{\Omega}} = e^{-\frac{n}{2(n-1)\sigma_o^2}(x_n - \bar{x})^2}$$

Therefore, it follows that the sample statistic which provides an efficient test of H_o against H_1 for the case of a single outlying observation is the difference between this observation and the sample mean (in terms of the population σ_o). Of course, if in the above approach we had under H_1 put $a_n = a + \lambda\,\sigma_o$, $\lambda \neq 0$ and had taken partial derivatives with respect to a and λ, we would have arrived at the same result. Moreover, it follows that if we test the null-hypothesis H_o that $a_1 = a_2 = a_3 = \ldots = a_n = a$ against the alternative H_1 that $a_2 = a_3 = \ldots = a_n = a$ but that $a_1 = a - \lambda\sigma_o$, then an efficient test would be based on the sample statistic

$$\frac{\bar{x} - x_1}{\sigma_o} ,$$

i.e. the difference between the sample mean and the smallest observation.

We will now demonstrate that the above test provides the most efficient criteria for comparing H_o against H_1. Following the theory of Neyman-Pearson [17], we are testing a composite hypothesis with one degree of freedom, i.e. H_o: $a_n = a$, against alternatives of the form $a_n - a = b > 0$. Here the true mean a is unspecified.

Now $$p_o = (\sqrt{2\pi}\,\sigma_o)^{-n}\, e^{-\frac{1}{2\sigma_o^2}\sum_{i=1}^{n}(x_i - a)^2}$$

and

$$\phi = \frac{\partial \log p_o}{\partial a} = \frac{n(\bar{x} - a)}{\sigma_o^2} ,$$

$$\phi' = \frac{\partial \phi}{\partial a} = -\frac{n}{\sigma_o^{\,o}}$$

so that $\phi' = A + B\phi$ where A and B are independent of the x_i (in fact, B=0). Consequently, for the test of H_o critical regions, w, similar to the sample space exist [17]. Further, the inequality

$$p_t \geq k\,(\phi)\,p_o \qquad [(105) \text{ of ref. } 17]$$

where $k(\phi)$ is a constant (which may depend upon ϕ), will determine the best critical region $w_o(\phi)$ subject the condition for similarity

$$P_o\{w\,(\phi)\} = \epsilon\, P_o\{W(\phi)\} ,$$

where ϵ is the size of the critical region and $P_o\{w\,(\phi)\}$ and $P_o\{W\,(\phi)\}$ represent respectively the integral of p_o taken over the critical region $w(\phi)$ and the entire sample space $W(\phi)$ whatever be the value of ϕ_1 for the family of hypersurfaces $\phi = \phi_1 =$ constant. Now under the simple alternative hypothesis $H_1 = H_t$, we have

$$p_t = (\sqrt{2\pi}\,\sigma_o)^{-n} e^{-\frac{1}{2\sigma_o^2}\left\{\sum_{i=1}^{n-1}(x_i - a_t)^2 + (x_n - a_t - b)^2\right\}}$$

and since we are dealing with regions similar to the sample space with regard to the parameter a, the condition

$$p_t \geq k(\phi)\, p_o$$

means that $x_n \geq \frac{\sigma_o^2}{b} \log k(\phi) + \frac{b}{2} + a_t$

i.e. $x_n \geq k'(\phi)$, say

Thus, we wish to determine $k'(\phi)$ so that the condition

$$P_o\left\{w(\phi)\right\} = \epsilon P_o\left\{W(\phi)\right\}$$

holds. That is to say, the best critical region will be built up from pieces of the family of hyper-surfaces $\phi = \phi_1$ ($\bar{x}$ is constant on each ϕ_1) for which $x_n \geq k'(\phi)$.

It can be seen from section 5 below that $p(x_n, \bar{x})$ can be represented in the form

$$\frac{n\sqrt{n}}{\sqrt{n-1}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\sqrt{\frac{n}{n-1}}\, u_n\right)^2} F_{n-1}\left(\frac{n}{n-1} u_n\right) \cdot p(\bar{x})$$

where $u_n = \frac{x_n - \bar{x}}{\sigma_o}$, $p(\bar{x})$ is the probability density function of $\bar{x}$

and F_{n-1} is defined in section 5 below. Thus $\bar{x}$ and u_n are in effect distributed independently* of one another and the conditions,

(a) $P_o\left\{w(\phi_1)\right\} = \epsilon P_o\left\{W(\phi_1)\right\}$

i.e. $\int_{k'}^{\infty} p_o(x_n, \bar{x})\, dx_n = \epsilon\, p(\bar{x})$ where $\bar{x}$ is constant on $\phi = \phi_1$,

---

*As a matter of fact, $\bar{x}$ and u_n are orthogonal linear forms of normally distributed variates and therefore independent.

and

(b) $x_n \geq k'(\alpha)$,

really imply

$$\frac{n\sqrt{n}}{\sqrt{2\pi}\sqrt{n-1}} \int_{u_0}^{\infty} e^{-\frac{1}{2}\frac{n}{n-1}u_n^2} F_{n-1}\left(\frac{n}{n-1}u_n\right) du_n = \epsilon ,$$

where $u_o = \frac{k'-\bar{x}}{\sigma_o}$. Hence, $x_n \geq k'(\alpha)$ is equivalent to $u_n \geq u_o = \frac{k'-\bar{x}}{\sigma_o}$.

It follows that u_o can depend only on n and ϵ and whatever be $\bar{x}$ the element $w(\alpha)$ of the best critical region is given by the sample criterion,

$$\frac{x_n-\bar{x}}{\sigma_o} \geq \text{a const.}$$

(b) Sample Criteria When the Two Largest or the Two Smallest Observations are Outliers — σ Assumed Known = σ_o. For this case, we test the hypothesis H_o that $a_i=a$ for i=1, 2, ..., n against the hypothesis H_1 that $a_1=a_2= \ldots a_{n-2}= a$ but that $a_{n-1}= a + \lambda_{n-1}\sigma_o$ and $a_n= a + \lambda_n\sigma_o$ where $\lambda_{n-1}, \lambda_n \neq 0$.

As before,

$$P_\omega = (\sqrt{2\pi}\,\sigma_o)^{-n} e^{-\frac{1}{2\sigma_o^2}\sum_{i=1}^{n}(x_i-\bar{x})^2}$$

To determine P_Ω, we have

$$P_{H_1} = (\sqrt{2\pi}\,\sigma_o)^{-n} e^{-\frac{1}{2\sigma_o^2}\left\{\sum_{i=1}^{n-2}(x_i-a)^2 + (x_{n-1}-a-\lambda_{n-1}\sigma_o)^2+(x_n-a-\lambda_n\sigma_o)^2\right\}}$$

Moreover,

$$\frac{\partial P_{H_1}}{\partial \lambda_{n-1}} = 0 \quad \text{gives} \quad \lambda_{n-1} = \frac{x_{n-1}-a}{\sigma_o}$$

$$\frac{\partial P_{H_1}}{\partial \lambda_n} = 0 \quad \text{gives} \quad \lambda_n = \frac{x_n - a}{\sigma_o}$$

$$\frac{\partial P_{H_1}}{\partial a} = 0 \quad \text{gives} \quad a = \frac{1}{n-2} \sum_{i=1}^{n-2} x_i = \bar{x}_{n-1,n} \text{ , say}$$

and thus

$$\lambda_{n-1} = \frac{1}{\sigma_o}(x_{n-1} - \bar{x}_{n-1,n}) \text{ ,} \qquad \lambda_n = \frac{1}{\sigma_o}(x_n - \bar{x}_{n-1,n})$$

so that

$$P_\Omega = (\sqrt{2\pi}\,\sigma_o)^{-n}\, e^{-\frac{1}{2\sigma_o^2}\sum_{i=1}^{n-2}(x_i - \bar{x}_{n-1,n})^2}$$

Hence, $\surd = \frac{P_\omega}{P_\Omega}$ leads to the test

$$\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2 - \sum_{i=1}^{n-2}\left\{x_i - \frac{1}{n-2}\sum_{i=1}^{n-2} x_i\right\}^2}{\sigma_o^2} \geq \text{ a const.}$$

The similar statistic for the two low outliers is therefore obvious.

If in the above treatment we had put $\lambda_{n-1} = \lambda_n = \lambda$, then

$$P_\Omega = (\sqrt{2\pi}\,\sigma_o)^{-n}\, e^{-\frac{1}{2\sigma_o^2}\left\{\sum_{i=1}^{n}(x_i - \bar{x})^2 - \frac{2n}{(n-2)}\left[\frac{x_n + x_{n-1}}{2} - \bar{x}\right]^2\right\}}$$

$\therefore \quad \surd = \frac{P_\omega}{P_\Omega}$ leads to a critical region based on

$$\frac{1}{\sigma_o}\left\{\frac{x_{n-1} + x_n}{2} - \bar{x}\right\} \geq \text{ a const.}$$

Analogously, if the two lowest observations x_1 and x_2 are equally outlying the likelihood principle would give a test based on

$$\frac{1}{\sigma_o}\left\{\bar{x} - \frac{x_1 + x_2}{2}\right\} \geq \text{a const.}$$

(c) Sample Criteria When Both the Largest and the Smallest Observations are Outliers — σ Assumed Known = σ_o. Here, we test the composite hypothesis H_o that $a_i = a$ for $i=1, 2, \ldots, n$ against the alternative hypothesis H_1 that $a_1 = a-\lambda_1 \sigma_o$, $a_2 = a_3 = \ldots = a_{n-1} = a$ and $a_n = a + \lambda_n \sigma_o$ where $\lambda_1, \lambda_n \neq 0$.

As before,

$$P_\omega = (\sqrt{2\pi}\,\sigma_o)^{-n} e^{-\frac{1}{2\sigma_o^2}\sum_{i=1}^{n}(x_i-\bar{x})^2}$$

To determine P_Ω, we have

$$P_{H_1} = (\sqrt{2\pi}\,\sigma_o)^{-n} e^{-\frac{1}{2\sigma_o^2}\left\{(x_1-a+\lambda_1\sigma_o)^2 + \sum_{i=2}^{n-1}(x_i-a)^2 + (x_n-a-\lambda_n\sigma_o)^2\right\}}$$

Now

$$\frac{\partial P_{H_1}}{\partial \lambda_1} = 0 \quad \text{gives} \quad \lambda_1 = -\frac{x_1 - a}{\sigma_o}$$

$$\frac{\partial P_{H_1}}{\partial \lambda_n} = 0 \quad \text{gives} \quad \lambda_n = \frac{x_n - a}{\sigma_o}$$

$$\frac{\partial P_{H_1}}{\partial a} = 0 \quad \text{gives} \quad a = \frac{1}{n}\sum_{i=1}^{n} x_i + \lambda_1 \frac{\sigma_o}{n} - \lambda_n \frac{\sigma_o}{n}$$

$$= \frac{1}{n-2}\sum_{i=2}^{n-1} x_i = \bar{x}_{1n}, \text{ say}$$

and thus

$$\lambda_1 = -\frac{1}{\sigma_o}(x_1-\bar{x}_{1n}) \ , \quad \lambda_n = \frac{1}{\sigma_o}(x_n-\bar{x}_{1n})$$

so that

$$P_\Omega = (\sqrt{2\pi}\,\sigma_o)^{-n}\, e^{-\frac{1}{2\sigma_o^2}\sum_{i=2}^{n-1}(x_i-\bar{x}_{1n})^2} .$$

Hence, $\nu = \frac{P_\omega}{P_\Omega}$ leads to the test

$$\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2-\sum_{i=2}^{n-1}\left\{x_i-\frac{1}{n-2}\sum_{i=2}^{n-1}x_i\right\}^2}{\sigma_o^2} \geq \text{a constant.}$$

We remark that if above we put $\lambda_1=\lambda_n = \lambda$ so that x_i and x_n are equally outlying (in opposite directions), then

$$P_\Omega = (\sqrt{2\pi}\,\sigma_o)^{-n}\, e^{-\frac{1}{2\sigma_o^2}\left\{\sum_{i=1}^{n}(x_i-\bar{x})^2-\frac{1}{2}(x_n-x_1)^2\right\}}$$

Hence, $\nu = \frac{P_\omega}{P_\Omega}$ gives as an efficient test,

$$\frac{x_n-x_1}{\sigma_o} \geq \text{a constant},$$

i.e. the sample range, which has been proposed by "Student" [4] and others.

(d) <u>Sample Criteria for a Single Outlier -- σ Unknown.</u> Our purpose here is to test the composite hypothesis H_o that $a_i=a$ for $i=1, \ldots n$ against the simple alternative hypothesis H_1 that $a_i=a$ for $i=1, 2, \ldots, n-1$ and $a_n = a+\lambda\sigma$, $\lambda > 0$, no knowledge of σ being available. In fact, a and σ both are unspecified.

Maximizing

$$P_{H_o} = (\sqrt{2\pi}\,\sigma)^{-n}\, e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-a)^2}$$

with respect to a and σ, we have

$$\frac{\partial P_{H_o}}{\partial a} = 0 \qquad \text{gives} \qquad a = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x}$$

$$\frac{\partial P_{H_o}}{\partial \sigma} = 0 \qquad \text{gives } \sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-a)^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2 = s^2$$

so that

$$P_\omega = (\sqrt{2\pi}\, s)^{-n}\, e^{-\frac{n}{2}}$$

To determine P_Ω, we take into account all admissible values of the parameters by setting

$$P_{H_1} = (\sqrt{2\pi}\,\sigma)^{-n}\, e^{-\frac{1}{2\sigma^2}\left\{\sum_{i=1}^{n-1}(x_i-a)^2 + (x_n-a_n)^2\right\}}$$

Taking partial derivatives, we have

$$\frac{\partial P_{H_1}}{\partial a} = 0 \qquad \text{gives} \qquad a = \frac{1}{n-1}\sum_{i=1}^{n-1} x_i = \bar{x}_n, \text{ say}$$

$$\frac{\partial P_{H_1}}{\partial a_n} = 0 \qquad \text{gives} \qquad a_n = x_n$$

$$\frac{\partial P_{H_1}}{\partial \sigma} = 0 \qquad \text{gives} \qquad \sigma^2 = \frac{1}{n}\left\{\sum_{i=1}^{n-1}(x_i-a)^2 + (x_n-a_n)^2\right\}$$

$$= \frac{1}{n}\sum_{i=1}^{n-1}(x_i-\bar{x}_n)^2 = s_n^2, \text{ say,}$$

so that

$$P_\Omega = (\sqrt{2\pi}\, s_n)^{-n}\, e^{-\frac{1}{2s_n^2}\sum_{i=1}^{n-1}(x_i-\bar{x}_n)^2} = (\sqrt{2\pi}\, s_n)^{-n}\, e^{-\frac{n}{2}}$$

Hence,

$$\nu = \frac{P_\omega}{P_\Omega} = \left(\frac{s_n^2}{s^2}\right)^{\frac{n}{2}}$$

and we will use the $\frac{2}{n}$ th root of ν as our test criterion, i.e. $\frac{s_n^2}{s^2}$.

Now $s_n^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2 - \frac{1}{n-1}(x_n-\bar{x})^2$

and thus

$$\frac{s_n^2}{s^2} = 1 - \frac{1}{n-1}\left(\frac{x_n-\bar{x}}{s}\right)^2 = 1 - \frac{1}{n-1}[T^{(1)}]^2$$

where $T^{(1)}$ is the statistic suggested by Pearson and Chandra Sekar [1] referred to in the Historical Comments above. We see therefore that the use of $T^{(1)}$ as a test of the hypothesis concerning a single outlier is equivalent to the test based on the most efficient (see proof below) sample criterion,

$$\frac{S_n^2}{S^2} = \frac{\sum_{i=1}^{n-1}\left(x_i - \frac{1}{n-1}\sum_{i=1}^{n-1} x_i\right)^2}{\sum_{i=1}^{n}\left(x_i - \frac{1}{n}\sum_{i=1}^{n} x_i\right)^2} \leq \text{a const.}$$

where $S_n^2 = n\, s_n^2$ and $S^2 = n\, s^2$.

The best test for the case in which the lowest observation is a single outlier is, correspondingly,

$$\frac{S_1^2}{S^2} = \frac{\sum_{i=2}^{n}\left(x_i - \frac{1}{n-1}\sum_{i=2}^{n} x_i\right)^2}{\sum_{i=1}^{n}\left(x_i - \frac{1}{n}\sum_{i=1}^{n} x_i\right)^2} \leq \text{a const.}$$

Percentage points for $\frac{S_n^2}{S^2}$ (or $\frac{S_1^2}{S^2}$) are given in Table I and were computed from the exact distribution derived in Section 8 of this work. Again, we emphasize that we have available here an exact test for the situation where σ is unknown and we are dealing with a single outlier.

We will now show that the sample criteria $\frac{S_n^2}{S^2} \leqslant$ a const. is the most efficient test of the statistical hypothesis H_o: $a_i = a$ for $i=1, 2, \ldots, n$ against alternatives of the type H_1: $a_1=a_2=\ldots=a_{n-1}=a$, $a_n = a+b$ where $b>0$, i.e. for the case of a single outlying observation, assuming that all the x_i are subject to the same but unknown standard error, σ. Since both a and σ are unspecified we are interested in testing the composite hypothesis with two degrees of freedom that $b=0$. That is to say, following the Neyman-Pearson theory of best tests [17], the simple alternative $H_1 = H_t$ specifies

$$\alpha_t^{(1)} = a_t, \quad \alpha_t^{(2)} = \sigma_t \quad \text{and} \quad \alpha_t^{(3)} = a_n - a_t = b$$

while H_o specifies only the single parameter, $\alpha_0^{(3)} = b = 0$.

Now $$p_o = (\sqrt{2\pi}\,\sigma)^{-n} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-a)^2}$$

and the condition of indefinite differentiability (see (A), page 317 of reference [17]) with respect to both a and σ is obviously satisfied. Also

$$\phi_1 = \frac{\partial \log p_o}{\partial a} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-a)$$

$$\phi_1^1 = \frac{\partial \phi_1}{\partial a} = -\frac{n}{\sigma^2}$$

so that

$$\phi_1^1 = A_1 + B_1\phi_1 \quad \text{where } A_1 \text{ and } B_1 \text{ are independent of the } x_i \text{ ;}$$

in fact, $A = -\frac{n}{\sigma^2}$, $B = 0$,

and

$$\phi_2 = \frac{\partial \log p_o}{\partial \sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{n}(x_i-a)^2 = -\frac{n}{\sigma} + \frac{n}{\sigma^3}\left\{(\bar{x}-a)^2 + s^2\right\}$$

$$\phi_2^1 = \frac{n}{\sigma^2} - \frac{3}{\sigma^4} \sum_{i=1}^{n} (x_i-a)^2 = A_2 + B_2 \phi_2$$

where A_2 and B_2 are independent of the x_i also. Moreover, the equation, ϕ_1 = a constant, is equivalent to $\bar{x} = c_1$ and the family of hypersurfaces $S(\alpha^{(2)}, C_1)$ — see condition (C), page 327 of [17] — corresponding to $\bar{x} = C_1$ is clearly independent of $\alpha^{(2)} = \sigma$. Therefore critical regions similar to the sample space with regard to a and σ exist for testing H_0. In order to find the best critical region w_0 of size ϵ we first find the hypersurface $W(\phi_1, \phi_2)$ determined by the loci of points such that ϕ_1 = const., ϕ_2 = const. and then find that part $w_0(\phi_1, \phi_2)$ of $W(\phi_1, \phi_2)$ satisfying the conditions:

$$p_t \geqslant k(\phi_1, \phi_2)\, p_o ,$$

$$\int \cdots \int_{w_o} p_o \, dw_o(\phi_1, \phi_2) = \epsilon \int \cdots \int_W p_o \, dW(\phi_1, \phi_2)$$

$W(\phi_1, \phi_1)$ is determined also by $\bar{x} = C_1$ and $s^2 = C_2 \geqslant 0$.

Now

$$p_t = (\sqrt{2\pi}\,\sigma_t)^{-n} e^{-\frac{1}{2\sigma_t^2} \sum_{i=1}^{n-1} \left\{(x_i-a_t) + (x_n - a_t - b)^2\right\}}$$

and since the best critical region is independent of a and σ, then $a=a_t$ and $\sigma=\sigma_t$ may be substituted in p_o. Hence the condition

$$p_t \geqslant k(\phi_1, \phi_2) p_o$$

means that

$$2b(x_n-a_t) \geqslant b^2 + 2\sigma_t^2 \log k(\phi_1, \phi_2)$$

or since for the single outlier x_n we are interested in $b > 0$, then $p_t \geqslant k\, p_o$ means simply that, say,

$$x_n \geqslant k'(\phi_1, \phi_2), \quad \text{a const.}$$

Further, we can determine k' by considering the hypersurface $W(\phi_1, \phi_2)$ on which $\bar{x}$ and s^2(or S^2) are constant, i.e. from

$$\text{(i)} \quad \int_{k'}^{\infty} p_o(\bar{x}, s^2, x_n) \cdot dx_n = \epsilon \int_{-\infty}^{\infty} p_o(\bar{x}, s^2, x_n)\, dx_n$$

We complete the argument by saying that it can be seen from sections 5 and 8 below, specifically the transformation (2) of section 5 and the transformation (11) of section 8, that $x_n - \bar{x} = \sqrt{\frac{n-1}{n}}\, \eta_n$, $\eta_n = r \cos \theta_n$, $\frac{\eta_{n+1}}{\sqrt{n}} = \bar{x}$ and $r = S$. Also, $\bar{x}$ and S are distributed independently of each other and can be factored out of (i) without affecting the variable $\sin \theta_n$ of section 8. Upon making the above transformations in (i), it will be found that

$$k' = \bar{x} + \sqrt{\frac{n-1}{n}}\, S \cos \theta_n'$$

(where θ_n' is that value of θ for a given n which makes the integral (17) below $= \epsilon$) so that $x_n \geq k'$ means

$$\frac{\frac{n}{n-1}(x_n - \bar{x})^2}{S^2} \geq \cos^2 \theta_n'$$

or

$$\frac{S^2 - \frac{n}{n-1}(x_n - \bar{x})^2}{S^2} \leq \sin^2 \theta_n'$$

That is to say, the best critical region is given by

$$\frac{S_n^2}{S^2} = \frac{\sum_{i=1}^{n-1} (x_i - \bar{x}_n)^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2} \leq \text{a const., where } \bar{x}_n = \frac{1}{n-1} \sum_{i=1}^{n-1} x_i$$

It follows, therefore, that the best critical region for testing x_1 will be given by $\frac{S_1^2}{S^2}$

(e) <u>Sample Criteria When the Two Largest or the Two Smallest Observations Are Outliers — σ Unknown</u>. This case is similar to that in 4(b) above except that σ must in effect be estimated from the sample.

It is clear from preceding developments that

$$P_\omega = (\sqrt{2\pi}\, s)^{-n}\, e^{-\frac{n}{2}}$$

For P_Ω, we have

$$P_{H_1} = (\sqrt{2\pi}\,\sigma)^{-n}\, e^{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n-2}\left\{(x_i-a)^2\right\} + \left\{(x_{n-1}-a_{n-1})^2 + (x_n-a_n)^2\right\}\right]}$$

$$\frac{\partial P_{H_1}}{\partial a} = 0 \quad \text{gives} \quad a = \frac{1}{n-2}\sum_{i=1}^{n-2} x_i = \bar{x}_{n-1,n}$$

$$\frac{\partial P_{H_1}}{\partial a_{n-1}} = 0 \quad \text{gives} \quad a_{n-1} = x_{n-1}$$

$$\frac{\partial P_{H_1}}{\partial a_n} = 0 \quad \text{gives} \quad a_n = x_n$$

$$\frac{\partial P_{H_1}}{\partial \sigma} = 0 \quad \text{gives} \quad \sigma^2 = \frac{1}{n}\sum_{i=1}^{n-2}(x_i - \bar{x}_{n-1,n})^2 = s^2_{n-1,n}\text{, say}$$

Hence, $P_\Omega = (\sqrt{2\pi}\, s_{n-1,n})^{-n}\, e^{-\frac{n}{2}}$ and the principle of likelihood gives

$$\frac{P_\omega}{P_\Omega} = \left(\frac{s^2_{n-1,n}}{s^2}\right)^{\frac{n}{2}}$$

We propose to use the $\frac{2}{n}$ th root of P_ω / P_Ω as the sample criterion, i.e. the test based on the principle of likelihood will involve the critical region

$$\frac{s^2_{n-1,n}}{s^2} = \frac{S^2_{n-1,n}}{S^2} \leqslant \text{ a const.}$$

The distribution of this sample statistic is derived in Section 9.

(f) <u>Sample Criteria When Both the Largest and the Smallest Observations are Outliers — σ Unknown.</u> We test here the composite hypothesis H_o that $a_1 = a_2 = \dots = a_n = a$ against the simple alternative hypothesis H_1 that $a_2 = a_3 = \dots = a_{n-1} = a$ but that $a_1 = a - \lambda_1\sigma$, say, and $a_n = a + \lambda_n\sigma$, where $\lambda_1, \lambda_n \neq 0$. This case is similar to 4(c) above except that σ must in effect be estimated from the sample.

As previously, we have

$$P_\omega = (\sqrt{2\pi}\, s)^{-n} e^{-\frac{n}{2}}$$

For P_Ω, we have

$$P_{H_1} = (\sqrt{2\pi}\,\sigma)^{-n} e^{-\frac{1}{2\sigma^2}\left\{(x_1 - a_1)^2 + \sum_{i=2}^{n-1}(x_i - a) + (x_n - a_n)^2\right\}}$$

$$\frac{\partial P_{H_1}}{\partial a_1} = 0 \quad \text{gives} \quad a_1 = x_1$$

$$\frac{\partial P_{H_1}}{\partial a} = 0 \quad \text{gives} \quad a = \frac{1}{n-2}\sum_{i=2}^{n-1} x_i = \bar{x}_{1n} \text{ , say}$$

$$\frac{\partial P_{H_1}}{\partial a_n} = 0 \quad \text{gives} \quad a_n = x_n$$

$$\frac{\partial P_{H_1}}{\partial \sigma} = 0 \quad \text{gives} \quad \sigma^2 = \frac{1}{n}\sum_{i=2}^{n-1}\left\{x_i - \bar{x}_{1n}\right\}^2 = s^2_{1n} \text{ , say}$$

Hence, $P_\Omega = (\sqrt{2\pi}\, s_{1n})^{-n} e^{-\frac{n}{2}}$ and the principle of likelihood renders

$$\frac{P_\omega}{P_\Omega} = \left(\frac{s_{1n}^2}{s^2}\right)^{\frac{n}{2}}$$

Thus, an efficient test would be based on the critical region

$$\frac{s_{1n}^2}{s^2} \leq \text{a const.}$$

It follows from this section that if the population standard deviation, σ, is known then efficient sample criteria for testing "outlying observations" should consist of the difference between two sums of squares: the first sum of squares is based on the entire sample and the second is determined from only the observations not suspected of being discordant. On the other hand, if σ is in effect to be inferred from the sample, efficient tests should include the ratio of one sum of squares to a second sum of squares: the first sum of squares is based on the observations considered as being harmonious and the second is made up from all the items in the sample. If we are concerned with the hypotheses of 4(a) our test should be based on the difference between the largest observation and the sample mean (or the sample mean minus the smallest observation in testing whether the smallest observation is outlying). We explore this distribution in the following sections. If we are interested in the hypotheses of 4(d), then we have an exact test and Table I gives significance levels for some useful sample sizes, the exact distribution being derived in Section 8 below. The exact distribution of the sample statistic for testing the hypotheses of 4(e) is given in Section 9 below. For the general cases of either 4(b), 4(c) or 4(f) further developments in the field of order statistics are apparently needed. (See Section 10 below for comments on the distributions called for in 4(f), however.) For the special case of 4(b) where the largest and the smallest observations are so to speak equally outlying and σ is assumed known, an efficient test in the Neyman-Pearson sense is already available; namely, the sample range [5], [6], [8], [9], [10]. The author believes that the tests proposed in 4(b), 4(c), 4(e) and 4(f) are "best" tests, however, he has not succeeded in proving this. We conclude this section by remarking that efficient tests for outlying observations depend apparently on just what situation we are faced with. The appropriate test to employ will depend upon the particular observations in the ordered sample considered to be outlying. No single test provides a uniformly most powerful test for all conditions likely to be encountered in practice. The type of criteria giving best tests for any number of outliers in a sample appears to involve two sums of squares—one based on all the observations and the other only on the "harmonious" observations in the sample.

5. Distribution of the Difference Between the Extreme and Mean in Samples of n From a Normal Population.* The simultaneous density function of n independent observations from a normal parent with zero mean and variance σ^2 which are arranged in order of magnitude is given by

$$(1) \quad dF(x_1,x_2,\ldots x_n)= \frac{n!}{(\sqrt{2\pi}\ \sigma)^n} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n} x_i^2} dx_1\, dx_2 \ldots dx_n$$

$$\text{subject to } x_1 \leq x_2 \leq \ldots \leq x_n$$

Since

$$\sum_{i=1}^{n}(x_i-\bar{x})^2= \frac{n}{n-1}(x_n-\bar{x})^2+ \sum_{i=1}^{n-1}(x_i-\bar{x}_n)^2 \text{ where } \bar{x}_n= \frac{1}{n-1}\sum_{i=1}^{n-1} x_i$$

then

$$\sum_{i=1}^{n}x_i^2=n\bar{x}^2+ \frac{n}{n-1} (x_n-\bar{x})^2+ \frac{n-1}{n-2}(x_{n-1}-\bar{x}_n)^2+ \frac{n-2}{n-3}(x_{n-2}- \bar{x}_{n,n-1})^2$$

$$+ \ldots + \frac{3}{2}(x_3- \frac{x_1+x_2+x_3}{3})^2+ \frac{2}{1}(x_2-\frac{x_1+x_2}{2})^2$$

where

$$\bar{x}_{n,n-1}= \frac{1}{n-2}\sum_{i=1}^{n-2} x_i, \text{ etc.}$$

and consequently we find that we are particularly interested in the following Helmert orthogonal transformation:

*It has been noted that K. R. Nair published the same derivation of this distribution in Biometrika, Vol. 35 (May, 1948)- see Ref. [20]. However, the author arrived at the derivation in the spring of 1945 and hence includes it as a basic part of this work in connection with the derivations given in Sections 8 and 9. Our Table II is considerable more extensive than Nair's table.

$$\sqrt{2\cdot 1}\ \eta_2 = -x_1 + x_2$$

$$\sqrt{3\cdot 2}\ \eta_3 = -x_1 - x_2 + 2x_3$$

$$\sqrt{4\cdot 3}\ \eta_4 = -x_1 - x_2 - x_3 + 3x_4$$

$$\vdots$$

(2)

$$\sqrt{r(r-1)}\ \eta_r = -x_1 - x_2 - x_3 - x_4 - \dots + (r-1)x_r$$

$$\vdots$$

$$\sqrt{n(n-1)}\ \eta_n = -x_1 - x_2 - x_3 - x_4 - \dots - x_r - \dots - x_{n-1} + (n-1)\, x_n$$

$$\sqrt{n}\qquad \eta_{n+1} = x_1 + x_2 + x_3 + x_4 + \dots + x_r + \dots + x_{n-1} + x_n$$

The above transformation solved for the x_i gives

$$x_n = \frac{\sqrt{n}}{n}\eta_{n+1} + \frac{\sqrt{n(n-1)}}{n}\eta_n$$

$$x_{n-1} = \frac{\sqrt{n}}{n}\eta_{n+1} - \frac{\sqrt{n(n-1)}}{n(n-1)}\eta_n + \frac{\sqrt{(n-1)(n-2)}}{n-1}\eta_{n-1}$$

$$\vdots$$

$$x_{n-r} = \frac{\sqrt{n}}{n}\eta_{n+1} - \frac{\sqrt{n(n-1)}}{n(n-1)}\eta_n - \frac{\sqrt{(n-1)(n-2)}}{(n-1)(n-2)}\eta_{n-1} - \dots + \frac{\sqrt{(n-r)(n-r-1)}}{(n-r)}\eta_{n-r}$$

(3)

$$\vdots$$

$$x_3 = \frac{\sqrt{n}}{n}\eta_{n+1} - \frac{\sqrt{n(n-1)}}{n(n-1)}\eta_n - \frac{\sqrt{(n-1)(n-2)}}{(n-1)(n-2)}\eta_{n-1} - \dots - \frac{\sqrt{4\cdot 3}}{4\cdot 3}\eta_4 + \frac{\sqrt{3\cdot 2}}{3}\eta_3$$

$$x_2 = \frac{\sqrt{n}}{n}\eta_{n+1} - \frac{\sqrt{n(n-1)}}{n(n-1)}\eta_n - \frac{\sqrt{(n-1)(n-2)}}{(n-1)(n-2)}\eta_{n-1} - \dots - \frac{\sqrt{4\cdot 3}}{4\cdot 3}\eta_4 - \frac{\sqrt{3\cdot 2}}{3\cdot 2}\eta_3 + \frac{\sqrt{2\cdot 1}}{2}\eta_2$$

$$x_1 = \frac{\sqrt{n}}{n}\eta_{n+1} - \frac{\sqrt{n(n-1)}}{n(n-1)}\eta_n - \frac{\sqrt{(n-1)(n-2)}}{(n-1)(n-2)}\eta_{n-1} - \dots - \frac{\sqrt{4\cdot 3}}{4\cdot 3}\eta_4 - \frac{\sqrt{3\cdot 2}}{3\cdot 2}\eta_3 - \frac{\sqrt{2\cdot 1}}{2}\eta_2$$

Thus, the conditions

$$x_1 \leq x_2 \leq x_3 \leq \cdots \leq x_n$$

give the following relations for the η_i

$$\infty \geq \eta_2 \geq 0$$

$$\sqrt{3}\,\eta_3 \geq \eta_2$$

$$\sqrt{\tfrac{4}{2}}\,\eta_4 \geq \eta_3$$

$$\vdots$$

(4) $$\sqrt{\frac{r}{r-2}}\,\eta_r \geq \eta_{r-1}$$

$$\vdots$$

$$\sqrt{\frac{n-1}{n-3}}\,\eta_{n-1} \geq \eta_{n-2}$$

$$\sqrt{\frac{n}{n-2}}\,\eta_n \geq \eta_{n-1}$$

whereas the region of integration for η_{n+1} is $-\infty \leq \eta_{n+1} \leq +\infty$

The Jacobian of the transformation (3) is unity,

i.e. $$\left| J\left(\frac{x_1, x_2, \ldots x_n}{\eta_2, \eta_3, \ldots \eta_{n+1}}\right) \right| = 1$$

and due to orthogonality we have

$$\sum_{i=1}^{n} x_i^2 = \sum_{i=2}^{n+1} \eta_i^2$$

Hence, the density function of the η_i is given by

$$dF(\eta_2, \eta_3, \cdots \eta_{n+1}) = \frac{n!}{(\sqrt{2\pi}\,\sigma)^n}\, e^{-\frac{1}{2\sigma^2}\sum_{i=2}^{n+1}\eta_i^2}\, d\eta_2\, d\eta_3 \cdots d\eta_{n+1}$$

subject to the conditions or regions of integration (4).

Since η_{n+1} is unrestricted, we integrate this variable out arriving at the density function

$$dF(\eta_2, \eta_3, \ldots \eta_n) = \frac{n!}{(\sqrt{2\pi}\,\sigma)^{n-1}}\, e^{-\frac{1}{2\sigma^2}\sum_{i=2}^{n} \eta_i^2}\, d\eta_2\, d\eta_3 \cdots d\eta_n$$

Recall that we seek the distribution of the statistic

$$u_r = \frac{x_r - \bar{x}}{\sigma} \qquad \text{where } \bar{x} = \frac{1}{r}(x_1 + x_2 + \ldots + x_r)$$

Hence, we make the transformation

(5) $$\frac{\sqrt{r(r-1)}}{r\,\sigma}\,\eta_r = u_r \qquad r = 2, 3, \ldots n$$

The Jacobian of this transformation is given by

$$J\left(\frac{\eta_2, \eta_3, \ldots \eta_n}{u_2, u_3, \ldots u_n}\right) = \frac{2\sigma}{\sqrt{2.1}} \cdot \frac{3\sigma}{\sqrt{3.2}} \cdots \frac{n\sigma}{\sqrt{n(n-1)}} = \sqrt{n}\,\sigma^{n-1}$$

and the conditions on the η_r (4) determine the limits of integration for the u_r as

(6) $$\begin{aligned} \infty \geq\ & u_2 \geq 0 \\ \tfrac{3}{2} u_3 &\geq u_2 \\ \tfrac{4}{3} u_4 &\geq u_3 \\ &\vdots \\ \tfrac{n}{n-1} u_n &\geq u_{n-1} \end{aligned}$$

Integrating the u_r, $(r=2, 3, \ldots n-1)$ over their appropriate ranges, we have the probability density function of u_n:

$$F(u_n) = \frac{n!\sqrt{n}}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{n}{n-1}u_n^2\right)} du_n \int_0^{\frac{n}{n-1}u_n} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{n-1}{n-2}u_{n-1}^2\right)} du_{n-1} \quad .$$

$$(7) \int_0^{\frac{n-1}{n-2}u_{n-1}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{n-2}{n-3}u_{n-2}^2\right)} du_{n-2} \cdots \int_0^{\frac{5}{4}u_5} \frac{1}{\sqrt{2\pi}} {}^{-\frac{1}{2}\left(\frac{4}{3}u_4^2\right)} du_4 \quad .$$

$$\int_0^{\frac{4}{3}u_4} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{3}{2}u_3^2\right)} du_3 \int_0^{\frac{3}{2}u_3} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(2u_2^2\right)} du_2$$

Thus, the distribution of $x_n - \bar{x}$ or of $\bar{x} - x_1$ resolves into the problem of computing a transformed multiple normal probability integral for n-1 variables.

Defining $F_n(u) = \int_0^u dF(u_n)$ we have

$$F_2(u) = 2\sqrt{2} \int_0^u \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(2x^2)} dx$$

$$(8) \qquad F_3(u) = \frac{3\sqrt{3}}{\sqrt{2}} \int_0^u \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{3}{2}x^2\right)} F_2\left(\tfrac{3}{2}x\right) dx$$

$$F_4(u) = \frac{4\sqrt{4}}{\sqrt{3}} \int_0^u \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{4}{3}x^2\right)} F_3\left(\tfrac{4}{3}x\right) dx$$

$$\vdots$$

$$F_n(u) = \frac{n\sqrt{n}}{\sqrt{n-1}} \int_0^u \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{n}{n-1}x^2\right)} F_{n-1}\left(\tfrac{n}{n-1}x\right) dx$$

This is equivalent to the result of McKay (11), although the above derivation is a considerably simpler one.

Now $F_{n-1}(u)$ increases from 0 to 1 as u increases from 0 to ∞. Hence, if $F_{n-1}(\frac{n}{n-1}u)$ is practically unity, i.e. for $\frac{n}{n-1}u$ numerically large, the upper percentage points of u_n may be approximated by the normal integral

$$(9)\int_{u_n}^{\infty} dF(u_n) = \frac{n}{\sqrt{2\pi}}\int_{u_n}^{\infty} e^{-\frac{1}{2}\frac{n}{n-1}u_n^2} \frac{\sqrt{n}}{\sqrt{n-1}}\, du_n$$

$$= \frac{n}{\sqrt{2\pi}}\int_{\sqrt{\frac{n}{n-1}}u_n}^{\infty} e^{-\frac{t^2}{2}}\, dt$$

Formula (9) was found to be particularly useful in checking the higher probabilities in Table II.

It may be noted that for $n=2$ in (7) and (8), we have the well-known result

$$F_2(u) = 2\sqrt{2}\int_0^{u} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\sqrt{2}u_2)^2}\, du_2$$

$$= \frac{2}{\sqrt{\pi}}\int_0^{u} e^{-u_2^2}\, du_2$$

where u_2 is either the sample standard deviation, the difference between the largest and mean, the mean deviation or the semi-range for a sample of two items.

The cumulative distribution functions (8) may be put into another form by setting

$$u_r = \frac{1}{r}v_r, \quad r = 2, 3, \ldots, n$$

Then $F_n(u)$ becomes

$$(10) \quad F_n(u) = \frac{\sqrt{n}}{(\sqrt{2\pi})^{n-1}} \int_0^{nu} \int_0^{v_n} \int_0^{v_{n-1}} \cdots \int_0^{v_4} \int_0^{v_3} e^{-\frac{1}{2} \sum_{i=2}^{n} \frac{v_i^2}{i(i-1)}} \; .$$

$$\cdot dv_2 \, dv_3 \ldots dv_n$$

Define the following functions:

$$H_1(x) = 1$$

$$H_2(x) = \sqrt{2} \int_0^x \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} \cdot \frac{t^2}{2 \cdot 1}} H_1(t)\, dt$$

$$H_3(x) = \sqrt{\frac{3}{2}} \int_0^x \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} \cdot \frac{t^2}{3 \cdot 2}} H_2(t)\, dt$$

$$H_4(x) = \sqrt{\frac{4}{3}} \int_0^x \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} \cdot \frac{t^2}{4 \cdot 3}} H_3(t)\, dt$$

$$\vdots$$

$$H_n(x) = \sqrt{\frac{n}{n-1}} \int_0^x \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} \cdot \frac{t^2}{n(n-1)}} H_{n-1}(t)\, dt$$

Hence, the probability that the difference between the extreme and the mean in samples of n from a normal population is less than $u\sigma$ is given by the alternative forms

$$P\{u_n \leq u\,\sigma\} = F_n(u) = H_n(n\,u)$$

Of course, $H_n(n\,u) \to 1$ as $u \to \infty$ for any given n.

In the November 1945 issue of Biometrika, Godwin [13] arrived at a series of functions closely related to the $H_r(x)$ in connection with the distribution of the mean deviation in samples of n from a normal parent. In Godwin's work, he defines functions $G_r(x)$ which are related to the $H_r(x)$ by the equation

$$(2\pi)^{\frac{r}{2}} \; H_{r+1}(x) = G_r(x)$$

The $G_r(x)$ functions which were computed by H.O. Hartley [15] for r=2, 3, ... 9 only have not as yet been published. Computations on the functions $F_n(u)$, i.e. (8), were well under way by the author before Godwin's article on the mean deviation appeared. The $H_r(x)$ or $G_r(x)$ can be used to obtain both the distribution of the difference between the extreme and mean and also the probability integral of the mean deviation. Indeed, it is believed that these functions will have a useful place in tabulating distributions of order statistics.

6. The Tabulation of the Distribution Function, $F_n(u)$.

The tabulation of the $F_n(u)$ with ordinary computing equipment is quite laborious. However, a table model computing machine was used initially to obtain the $F_n(u)$ for n=2 to n=15 with the aid of formulae (8) in the following fashion.

For a sample of size n=2, the probabilities $F_2(u)$ may be obtained directly from the WPA "Tables of Probability Functions", Volume I [14], since

$$F_2(u) = \frac{2}{\sqrt{\pi}} \int_0^{u} e^{-x^2} \, dx$$

For sample sizes greater than n=2, a process of numerical quadrature was used to obtain the $F_n(u)$. Thus,

$$dF_3(u) = \frac{3\sqrt{3}}{\sqrt{2}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{3}{2}u^2\right)} du \cdot \int_{-\frac{3}{\sqrt{2}}u}^{\frac{3}{\sqrt{2}}u} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

$$dF_4(u) = \frac{4\sqrt{4}}{\sqrt{3}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{4}{3}u^2\right)} du \cdot \int_0^{\frac{4}{3}u} dF_3(u)$$

$$\vdots$$

$$dF_r(u) = \frac{r\sqrt{r}}{r-1} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{r}{r-1}u^2\right)} du \int_0^{\frac{r}{r-1}u} dF_{r-1}(u)$$

The functions $\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{r}{r-1}u^2\right)}$ were obtained to seven or eight decimal accuracy by interpolation on column 2, of the WPA Tables [14] and the integrals

$$\int_0^{\frac{r}{r-1}u} dF_{r-1}(u) = F_{r-1}\left(\frac{r}{r-1}u\right)$$

were recorded to seven decimal places using six-point Lagrangian interpolation formulae on the distribution functions

$$F_{r-1}(u)$$

In view of the fact that $F_{r-1}(u)$ must be interpolated on to find one of the factors of $dF_r(u)$, checks were made at each stage, $F_r(u)$, by differencing, calculating moments in the final tables and by formula (9).

Each $F_r(u)$ was produced by numerical quadrature (Weddle's Rule) on the computed ordinates. The following table gives an approximate summary of the intervals used for the tabulation:

n	Range of $F_n(u)$
3	u = 0 (.025) 1.00 (.05) 5.00
4	u = 0 (.025) 1.00 (.05) 5.20
5	u = 0 (.025) 1.00 (.05) 2.00 (.10) 5.30
6	u = 0 (.025) 0.90 (.05) 2.00 (.10) 5.40
7	u = 0 (.025) 0.80 (.05) 2.00 (.10) 5.50
8	u = 0 (.025) 0.70 (.05) 2.10 (.10) 5.60
9	u = 0 (.05) 2.20 (.10) 5.70
10	u = 0 (.05) 2.30 (.10) 5.80
11	u = 0 (.05) 2.40 (.10) 5.90
12	u = 0 (.05) 2.50 (.10) 6.00
13	u = 0 (.05) 2.60 (.10) 6.10
14	u = 0 (.05) 2.70 (.10) 6.20
15	u = 0 (.05) 2.80 (.10) 6.30

About five decimal accuracy was available for $F_{15}(u)$ using the above procedure.

In view of the possible general usefulness of the $H_r(x)$, these functions were also computed as a sample problem on a high-speed computing device, the ENIAC (Electronic Numerical Intergrator and Computor) of the Ballistic Research Laboratories of the Ordnance Department*. In this connection, the $H_r(x)$ have been computed for r=2 to r=25 at the Ballistic Research Laboratories. For n=2, the functions $H_r(x)$ were computed to nine decimal places of accuracy on the ENIAC and at n=25 about five decimal places of accuracy were obtained. In Table II we have tabulated $F_n(u)$ or $H_n(nu)$, i.e. the probability integral of the extreme minus the mean, at intervals of $u = .05\sigma$. Values computed on the table model computing machine agreed to five decimal places at n=15 with values from the ENIAC. Percentage Points of the distribution are given in Table III and the moment constants may be found in Table IV. Moment constants for n=60, 100, 200, 500 and 1000 were obtained by use of McKay's formulae [11] (which relate the semi-invariants of $x_n - \bar{x}$ with those of x_n) and Tippett's moments [5] for the largest observation x_n.

---

* The author suggested the problem of tabulating the functions $F_n(u)$ or $H_n(nu)$ to the Computing Laboratory of the Ballistic Research Laboratories in the fall of 1945; however, due to problems of higher priority, these functions were not computed on the ENIAC until March, 1948.

7. <u>Relation Between the Distribution of the Largest Minus the Mean of All n Observations and the Largest Minus the Mean of the Remaining n-1 Items</u>. The following relation is of interest concerning these two statistics:

Let $$u_n = x_n - \frac{x_1 + x_2 + \ldots + x_n}{n}$$

$$= \frac{1}{n}\left\{(n-1)x_n - x_1 - x_2 - \ldots - x_{n-1}\right\}$$

Let $$v_n = x_n - \frac{x_1 + x_2 + \ldots + x_{n-1}}{n-1}$$

$$= \frac{1}{n-1}\left\{(n-1)x_n - x_1 - x_2 - \ldots - x_{n-1}\right\}$$

Hence, $$v_n = \frac{n}{n-1} u_n$$

or

$$P(v_n \leq t_o) = P\left(\frac{n}{n-1} u_n \leq t_o\right) = P\left\{u_n \leq \frac{n-1}{n} t_o\right\}$$

i.e. the probability integral of the largest minus the mean of the other observations may be obtained by interpolation on the distribution of the largest minus the mean of all n items in the sample.

8. <u>The Distribution of S_n^2/S^2 and S_1^2/S^2.</u> As shown in Section 4 above the best test in the Neyman-Pearson sense and the likelihood principle leads to the sample criterion

$$\frac{S_n^2}{S^2} = \frac{\sum_{i=1}^{n-1}(x_i - \bar{x}_n)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \leq k \qquad \bar{x}_n = \frac{1}{n-1}\sum_{i=1}^{n-1} x_i$$

for testing the largest observation and

$$\frac{S_1^2}{S^2} = \frac{\sum_{i=2}^{n}(x_i-\bar{x}_1)^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2} \leq k \qquad \bar{x}_1 = \frac{1}{n-1}\sum_{i=2}^{n} x_i$$

for testing whether the smallest observation is outlying. We now find the probability distribution of S_n^2/S^2; hence, also that of S_1^2/S^2.

Returning to the density function

$$dF(\eta_2, \eta_3, \ldots \eta_n) = \frac{n!}{(\sqrt{2\pi})^{n-1}} e^{-\frac{1}{2}\sum_{i=2}^{n}\eta_i^2} d\eta_2 d\eta_3 \cdots d\eta_n$$

of Section 4 (but neglecting σ which will not effect what follows), we make the polar transformation

$$\begin{aligned}
\eta_2 &= r \sin\theta_n \sin\theta_{n-1} \cdots \sin\theta_4 \sin\theta_3 \\
\eta_3 &= r \sin\theta_n \sin\theta_{n-1} \cdots \sin\theta_4 \cos\theta_3 \\
\eta_4 &= r \sin\theta_n \sin\theta_{n-1} \cdots \sin\theta_4 \cos\theta_3 \\
&\vdots \\
\eta_{n-2} &= r \sin\theta_n \sin\theta_{n-1} \cos\theta_{n-2} \\
\eta_{n-1} &= r \sin\theta_n \cos\theta_{n-1} \\
\eta_n &= r \cos\theta_n
\end{aligned} \tag{11}$$

Now

$$\sum_{i=2}^{n}\eta_i^2 = \sum_{i=1}^{n}(x_i-\bar{x})^2 = r^2$$

and

$$\sum_{i=2}^{n-1}\eta_i^2 = \sum_{i=1}^{n-1}(x_i-\bar{x}_n)^2 = r^2\sin^2\theta_n$$

Hence,

$$\sin^2 \theta_n = \frac{\sum_{i=1}^{n-1}(x_i-\bar{x}_n)^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}$$

The Jacobian of the above transformation is

$$r^{n-2} \sin^{n-3}\theta_n \sin^{n-4}\theta_{n-1} \cdots \sin^3\theta_6 \sin^2\theta_5 \sin\theta_4$$

and since $0 \leq r \leq \infty$

(12) $$dF(\theta_n, \theta_{n-1}, \ldots, \theta_5, \theta_4, \theta_3)$$

$$= \frac{n!}{(2\pi)^{\frac{n-1}{2}}} 2^{\frac{n-3}{2}} \Gamma\left(\frac{n-1}{2}\right) \sin^{n-3}\theta_n \ldots \sin^2\theta_5 \sin\theta_4 \, d\theta_n \ldots d\theta_5 \, d\theta_4 \, d\theta_3$$

Recalling the restrictions

$$\eta_2 \geq 0 \quad , \quad \sqrt{\frac{r}{r-2}}\, \eta_r \geq \eta_{r-1} \qquad r \geq 3$$

we have

$$\tan \theta_n \cos \theta_{n-1} = \frac{\eta_{n-1}}{\eta_n} \qquad n \geq 4$$

or

$$\tan \theta_n \leq \sqrt{\frac{n}{n-2}} \sec \theta_{n-1} \qquad n \geq 4$$

and

$$0 \leq \theta_3 \leq \frac{\pi}{3}$$

Thus, letting $K_n = \frac{n!}{(2\pi)^{\frac{n-1}{2}}} 2^{\frac{n-3}{2}} \Gamma\left(\frac{n-1}{2}\right)$, we see that

(13) $$K_n \int_0^{\frac{\pi}{3}} \int_0^{\tan^{-1}\sqrt{\frac{4}{2}}\sec\theta_3} \cdots \int_0^{\tan^{-1}\sqrt{\frac{n-1}{n-3}}\sec\theta_{n-2}} \int_0^{\tan^{-1}\sqrt{\frac{n}{n-2}}\sec\theta_{n-1}} \sin^{n-3}\theta_n$$

$$\cdots \sin^2\theta_5 \sin\theta_4 \, d\theta_n \ldots d\theta_4 \, d\theta_3 = 1$$

Upon reversing the order of integration (the variable limits are monotonic) we get for n=3

$$K_3 \int_0^{\frac{\pi}{3}} d\theta_3 = 1$$

so that

(14) $$P(\theta_3 \leq \theta) = K_3 \int_0^{\theta} d\theta_3 \qquad 0 \leq \theta \leq \tan^{-1}\sqrt{3\cdot 1}$$

When n=4, we obtain

$$K_4 \int_0^{\tan^{-1}\sqrt{\frac{4}{2}}} \int_0^{\frac{\pi}{3}} \sin\theta_4 \, d\theta_3 \, d\theta_4 + K_4 \int_{\tan^{-1}\sqrt{\frac{4}{2}}}^{\tan^{-1}\sqrt{4\cdot 2}} \int_{\sec^{-1}\sqrt{\frac{2}{4}}\tan\theta_4}^{\frac{\pi}{3}} \sin\theta_4 \, d\theta_3 \, d\theta_4 = 1$$

so that

(15a) $$P(\theta_4 \leq \theta) = \frac{K_4}{K_3} \int_0^{\theta} \sin\theta_4 \, d\theta_4 \qquad \text{when } 0 \leq \theta \leq \tan^{-1}\sqrt{\frac{4}{2}}$$

and

(15b) $P(\theta_4 \leq \theta)$

$$= \frac{K_4}{K_3} \int_0^{\tan^{-1}\sqrt{\frac{4}{2}}} \sin\theta_4 \, d\theta_4 + K_4 \int_{\tan^{-1}\sqrt{\frac{4}{2}}}^{\theta} \int_{\sec^{-1}\sqrt{\frac{2}{4}}\tan\theta_4}^{\frac{\pi}{3}} \sin\theta_4 \, d\theta_3 \, d\theta_4$$

$$\text{when } \tan^{-1}\sqrt{\frac{4}{2}} \leq \theta \leq \tan^{-1}\sqrt{4\cdot 2}$$

When n=5, we get

$$K_5 \int_0^{\tan^{-1}\sqrt{\frac{5}{3}}} \int_0^{\tan^{-1}\sqrt{\frac{4}{2}}} \int_0^{\frac{\pi}{3}} \sin^2\theta_5 \sin\theta_4 \, d\theta_3 \, d\theta_4 \, d\theta_5$$

$$+K\int_{0}^{\tan^{-1}\sqrt{\frac{5}{3}}}\int_{\tan^{-1}\sqrt{\frac{4}{2}}}^{\tan^{-1}\sqrt{4\cdot 2}}\int_{\sec^{-1}\sqrt{\frac{2}{4}}\tan\theta_4}^{\frac{\pi}{3}} \sin^2\theta_5 \sin\theta_4\, d\theta_3\, d\theta_4\, d\theta_5$$

$$+K\int_{\tan^{-1}\sqrt{\frac{5}{3}}}^{\tan^{-1}\sqrt{5\cdot 3}}\int_{\sec^{-1}\sqrt{\frac{3}{5}}\tan\theta_5}^{\tan^{-1}\sqrt{4\cdot 2}}\int_{\sec^{-1}\sqrt{\frac{2}{4}}\tan\theta_4}^{\frac{\pi}{3}} \sin^2\theta_5 \sin\theta_4\, d\theta_3\, d\theta_4\, d\theta_5 = 1$$

(where $\sec^{-1}\sqrt{\frac{2}{4}}\tan\theta_4$ is to be taken as 0 whenever $\theta_4 \leq \tan^{-1}\sqrt{\frac{4}{2}}$)

so that

(16a) $$P(\theta_5 \leq \theta) = \frac{K_5}{K_4}\int_0^{\theta} \sin^2\theta_5\, d\theta_5 \qquad \text{when} \quad 0 \leq \theta \leq \tan^{-1}\sqrt{\frac{5}{3}}$$

and

(16b) $P(\theta_5 \leq \theta)$

$$= \frac{K_5}{K_4}\int_0^{\tan^{-1}\sqrt{\frac{5}{3}}} \sin^2\theta_5\, d\theta_5 + K\int_{\tan^{-1}\sqrt{\frac{5}{3}}}^{\theta}\int_{\sec^{-1}\sqrt{\frac{3}{5}}\tan\theta_5}^{\tan^{-1}\sqrt{4\cdot 2}}$$

$$\int_{\sec^{-1}\sqrt{\frac{2}{4}}\tan\theta_4}^{\frac{\pi}{3}} \sin^2\theta_5 \sin\theta_4\, d\theta_3\, d\theta_4\, d\theta_5$$

when $\tan^{-1}\sqrt{\frac{5}{3}} \leq \theta \leq \tan^{-1}\sqrt{5\cdot 3}$

and we put $\sec^{-1}\sqrt{\frac{2}{4}}\tan\theta_4 = 0$ whenever $\theta_4 \leq \tan^{-1}\sqrt{\frac{4}{2}}$.

For a sample of n items

(17a) $$P(\theta_n \leq \theta) = \frac{n}{\sqrt{\pi}} \frac{\Gamma(\frac{n-1}{2})}{\Gamma(\frac{n-2}{2})} \int_0^{\theta} \sin^{n-3}\theta_n \, d\theta_n$$

$$= \frac{n}{2} I_{\sin^2\theta} \left(\frac{n-2}{2}, \frac{1}{2}\right) \qquad \text{when } 0 \leq \theta \leq \tan^{-1}\sqrt{\frac{n}{n-2}}$$

and

(17b) $$P(\theta_n \leq \theta)$$

$$= \frac{n}{2} I_{\frac{n}{2(n-1)}} \left(\frac{n-2}{2}, \frac{1}{2}\right)$$

$$+K_n \int_{\tan^{-1}\sqrt{\frac{n}{n-2}}}^{\theta} \int_{\sec^{-1}\sqrt{\frac{n-2}{n}}\,\theta_n}^{\tan^{-1}\sqrt{(n-1)(n-3)}} \int_{\sec^{-1}\sqrt{\frac{n-3}{n-1}}\tan\theta_{n-1}}^{\tan^{-1}\sqrt{(n-2)(n-4)}}$$

$$\cdots \int_{\sec^{-1}\sqrt{\frac{2}{4}}\tan\theta_4}^{\frac{\pi}{3}} \sin^{n-3}\theta_n \ldots \sin\theta_4 \, d\theta_3 \, d\theta_4 \ldots d\theta_n$$

for $\tan^{-1}\sqrt{\frac{n}{n-2}} \leq \theta \leq \tan^{-1}\sqrt{n(n-2)}$

where $I_x(p,q)$ is K. Pearson's Incomplete Beta Function Ratio [19]. It is to be understood in (17) that

$$\sec^{-1}\sqrt{\frac{i-2}{i}}\tan\theta_i \text{ for } i=4, 5, \ldots n-1$$

is to be taken as zero when $$\theta_i \leq \tan^{-1}\sqrt{\frac{i}{i-2}}$$

Percentage points for the sample statistic

$$\sin^2 \theta_n = \frac{S_n^2}{S^2} = \frac{\sum_{i=1}^{n-1} (x_i - \bar{x}_n)^2}{\sum_{i=1}^{n} (x_i - \bar{x})}$$

are given in Table I and were obtained by inverse interpolation on the tabulation of the probability integral (17) above. The job of tabulating the multiple integral (17) was carried out on the Bell Relay Computers at the Ballistic Research Laboratories.

9. <u>The Distribution of $S^2_{n-1,n}/S^2$ and $S^2_{1,2}/S^2$.</u> As indicated in Section 4, the Neyman-Pearson likelihood criterion for judging the significance of the two largest observations is

$$S^2_{n-1,n}/S^2 = \frac{\sum_{i=1}^{n-2} (x_i - \bar{x}_{n-1,n})^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2} \leq k \qquad \text{where} \quad \bar{x}_{n-1,n} = \frac{1}{n-2} \sum_{i=1}^{n-2} x_i$$

and that for testing the two smallest observations is

$$S^2_{1,2}/S^2 = \frac{\sum_{i=3}^{n} (x_i - \bar{x}_{1,2})^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2} \leq k \qquad \text{where} \quad \bar{x}_{1,2} = \frac{1}{n-2} \sum_{i=3}^{n} x_i$$

From the preceding section, we note that

$$\sum_{i=2}^{n} \eta_i^2 = r^2 \qquad \sum_{i=2}^{n-2} \eta_i^2 = r^2 \sin^2 \theta_n \quad \sin^2 \theta_{n-1}$$

Hence,

$$\sin^2\theta_n \sin^2\theta_{n-1} = \frac{\sum_{i=1}^{n-2} (x_i - \bar{x}_{n-1,n})^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$$

so that if we find the distribution of

$$\sin^2\theta_n \; \sin^2\theta_{n-1} = \sin^2\Delta_n \quad , \text{ say,}$$

then we have the distribution of $S^2_{n-1,n}/S^2$ and hence also that of $S^2_{1,2}/S^2$, i.e.

$$P\left\{\sin^2\Delta_n \leq k\right\} = P\left\{\Delta_n \leq \sin^{-1}\sqrt{k}\right\}$$

Returning to the multiple integral (13), let

$$\sin\Delta_n = \sin\theta_n \, \sin\theta_{n-1}$$

$$\Delta_i = \theta_i \qquad\qquad 3 \leq i \leq n-1$$

The Jacobian of this transformation is given by

$$\frac{\partial(\theta_n,\ldots,\theta_3)}{\partial(\Delta_n,\ldots,\Delta_3)} = \frac{\cos\Delta_n}{\sqrt{\sin^2\Delta_{n-1} - \sin^2\Delta_n}}$$

The limits of integration for Δ_n are given by

$$0 \leq \Delta_n \leq \sin^{-1}\frac{\sqrt{n}\;\sin\Delta_{n-1}}{\sqrt{2(n-1)-(n-2)\sin^2\Delta_{n-1}}}$$

and, of course, those for $\Delta_{n-1},\ldots,\Delta_3$ are the same as the limits for $\theta_{n-1},\ldots,\theta_3$ respectively. Hence, substituting in (13), we obtain

$$(20) \qquad K_n\int_0^{\frac{\pi}{3}}\int_0^{\tan^{-1}\sqrt{\frac{4}{2}}\sec\Delta_3}\cdots\int_0^{\tan^{-1}\sqrt{\frac{n-1}{n-3}}\sec\Delta_{n-2}}$$

$$\int_0^{\sin^{-1}\frac{\sqrt{n}\;\sin\Delta_{n-1}}{\sqrt{2(n-1)-(n-2)\sin^2\Delta_{n-1}}}} \frac{\sin^{n-3}\Delta_n\sin^{n-4}\Delta_{n-1}\cdots\sin^2\Delta_5\sin\Delta_4\cos\Delta_n d\Delta_n\cdots d\Delta_3}{\sin^{n-3}\Delta_{n-1}\sqrt{\sin^2\Delta_{n-1}-\sin^2\Delta_n}} = 1$$

Reversing the order of integration, we have

$$(21)\quad K_n \int_0^{\sin^{-1}\sqrt{\frac{n(n-3)}{(n-1)(n-2)}}} \int_{\sin^{-1}\frac{\sqrt{2(n-1)}\sin\Delta_n}{\sqrt{n+(n-2)\sin^2\Delta_n}}}^{\tan^{-1}\sqrt{(n-1)(n-3)}} \int_{\sec^{-1}\sqrt{\frac{n-3}{n-1}}\tan\Delta_{n-1}}^{\tan^{-1}\sqrt{(n-2)(n-4)}}$$

$$\cdots \int_{\sec^{-1}\sqrt{\frac{2}{4}}\tan\Delta_4}^{\frac{\pi}{3}} \frac{\sin^{n-3}\Delta_n \sin^{n-4}\Delta_{n-1}\cdots\sin\Delta_4\cos\Delta_n\, d\Delta_3\cdots d\Delta_n}{\sin^{n-3}\Delta_{n-1}\sqrt{\sin^2\Delta_{n-1}-\sin^2\Delta_n}} = 1$$

(For $\Delta_i \leq \tan^{-1}\sqrt{\frac{i}{i-2}}$, then $\sec^{-1}\sqrt{\frac{i-2}{i}}\tan\Delta_i$ is for $i \geq 4$ to be put equal to zero)
so that for n=4,

$$(22)\quad P(\Delta_4 \leq \Delta) = K_4 \int_0^{\Delta} \int_{\sin^{-1}\frac{\sqrt{3}\sin\Delta_3}{\sqrt{2+\sin^2\Delta_3}}}^{\frac{\pi}{3}} \frac{\sin\Delta_4\cos\Delta_4\, d\Delta_3\, d\Delta_4}{\sin\Delta_3\sqrt{\sin^2\Delta_3-\sin^2\Delta_4}}$$

where $0 \leq \Delta \leq \sin^{-1}\frac{2}{3}$,

and for n=5,

$$(23)\quad P(\Delta_5 \leq \Delta) = K_5 \int_0^{\Delta} \int_{\sin^{-1}\frac{\sqrt{4\cdot 2}\sin\Delta_5}{\sqrt{5+3\sin^2\Delta_5}}}^{\tan^{-1}\sqrt{4\cdot 2}}$$

$$\cdot \int_{\sin^{-1}\sqrt{\frac{2}{4}}\tan\Delta_4}^{\frac{\pi}{3}} \frac{\sin^2\Delta_5\cos\Delta_5\, d\Delta_3\, d\Delta_4\, d\Delta_5}{\sin\Delta_4\sqrt{\sin^2\Delta_4-\sin^2\Delta_5}}$$

where $0 \leq \Delta \leq \sin^{-1}\sqrt{\frac{5}{6}}$,
etc.

We remark that an obvious extension of the above principles should lead to the distributions of

$$S^2_{n-2,n-1,n}/S^2 \quad \text{and} \quad S^2_{1,2,3}/S^2 \ ,$$

$$S^2_{n-3,n-2,n-1,n}/S^2 \quad \text{and} \quad S^2_{1,2,3,4}/S^2 \ ,$$

etc. although the tabulation of such probability integrals may be exceedingly difficult.

The problem of tabulating the probability integral (21) involves a double quadrature process and has already been suggested to the Computing Laboratory of the Ballistic Research Laboratories. It is hoped that percentage points for

$$S^2_{n-1,n}/S^2 \quad \text{and} \quad S^2_{1,2}/S^2$$

will be available in due course; however, this problem will have to wait its turn along with other computing problems. For $n=4$, the 1% point for the above statistic is .0055 and the 5% point is .0138.

10. Comment on the Distribution of $S^2_{1,n}/S^2$. In connection with the distribution of the statistic

$$\frac{S^2_{1,n}}{S^2} = \frac{\sum_{i=2}^{n-1}(x_i-\bar{x}_{1,2})^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2} \quad ; \quad \bar{x}_{1,n} = \frac{1}{n-2}\sum_{i=2}^{n-1} x_i$$

for testing simultaneously whether the smallest and largest observations are outlying, an investigation indicates that since

$$\Sigma x_i^2 = n\bar{x} + \frac{n}{n-1}(x_n-\bar{x})^2 + \frac{n-1}{n-2}(x_1-\bar{x}_n)^2 + \frac{n-2}{n-3}(x_{n-1}-\bar{x}_{1,n})^2$$

$$+ \dots + \frac{3}{2}\left(x_4 - \frac{x_2+x_3+x_4}{3}\right)^2 + 2\left(x_3 - \frac{x_2+x_3}{2}\right)^2$$

then the transformation

$$\begin{aligned}
\sqrt{2\cdot 1}\, v_2 &= -x_2 + x_3 \\
\sqrt{3\cdot 2}\, v_3 &= -x_2 - x_3 + 2x_4 \\
\sqrt{4\cdot 3}\, v_4 &= -x_2 - x_3 - x_4 + 3x_5 \\
&\vdots \\
\sqrt{(n-2)(n-3)}\, v_{n-2} &= -x_2 - x_3 - \dots - x_{n-2} + (n-3)x_{n-1} \\
\sqrt{(n-1)(n-2)}\, v_{n-1} &= -(n-2)x_1 + x_2 + x_3 + \dots + x_{n-1} \\
\sqrt{n(n-1)}\, v_n &= -x_1 - x_2 - x_3 - \dots - x_{n-1} + (n-1)x_n \\
\sqrt{n}\, v_{n+1} &= x_1 + x_2 + \dots + x_n
\end{aligned} \tag{24}$$

followed by transformations of the type (11) and that of Section 9 will lead to the distribution of $S^2_{1\,n}/S^2$. However, the limits of integration do not turn out to be functions of single variables and the task of computing the resulting multiple integral may be exceedingly difficult.

11. Examples on Testing Outlying Observations for Rejection.

We now turn to the problem of applying our theory to particular practical examples of data which appear to have outlying observations. Apparently, in the following examples there were not sufficient practical or experimental grounds to reject the suspected outliers and hence some statistical judgement became necessary either to support retaining the "outliers" in the sample or leave little doubt that certain of the observations should be rejected.

Example 1. Our first example has almost become a classical one as Irwin [3], Rider [2], and other writers on the subject including Chauvenet, Peirce, Gould, etc. (see Rider's survey [2]) all refer to it, applying their various tests. The example consists of a sample of 15 observations of the vertical semi-diameters of Venus made by Lieut. Herndon in 1846 and is given in William Chauvenet's, "A Manual of Spherical and Practical Astronomy", II (5th ed., 1876), p.562. The individual residuals or deviations from the mean are:

-0.30"	0.48	0.63	-0.22	0.18
-0.44	-0.24	-0.13	-0.05	0.38
1.01	0.06	-1.40	0.20	0.10

Arranging the observations in increasing order of magnitude, we have:

-1.40"	-0.24	-0.05	0.18	0.48
-0.44	-0.22	0.06	0.20	0.63
-0.30	-0.13	0.10	0.39	1.01

and it is seen that two of the residuals, -1.40 and 1.01, appear to be outliers. Rider [2] indicates that the above observations have been referred to by previous writers as "residuals"; nevertheless their sum is 0.27, so that the sample mean, $\bar{x}=.018$. Let us apply the exact test, i.e. S_1^2/S^2 as developed in Sections 4 and 5 for a single outlier to the least observation, -1.40. We find $S^2= 4.2496$ using all 15 observations and $S_1^2= 2.0953$ which is based on 14 observations, the suspected outlier -1.40 not being considered. Further, $S_1^2/S^2= 0.4931$ and from Table I we see that $0.01 < P < 0.025$ so that we would reject the observation -1.40 when using the 5% level of significance. Having rejected -1.40, we now have left a sample of 14 observations and test the greatest one, i.e. 1.01. Upon computing new sums of squares, we find $S_n^2 = 1.2409$ leaving out 1.01 and $S^2 = 2.0953$ including the observation 1.01. Hence, $S_n^2/S^2 = 0.5922$ and from Table I, we find P slightly greater than .10, so that we decide to retain the observation 1.01.

It would be interesting to see whether or not the test $S^2_{1,n}/S^2$ would reject simultaneously the observations -1.40 and 1.01 if percentage points for the distribution of this statistic were available.

It is of interest to remark that for this particular example Irwin [3, page 245], using the difference between the first two individuals divided by an estimate of σ, i.e. $\frac{x_2 - x_1}{\sigma}$, concluded also that -1.40 but not 1.01 should be rejected. In testing both of these observations, Irwin used the single biased estimate for σ,

$$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2} = .5326 \qquad (\text{assuming } \bar{x}=0) \quad ,$$

based on all 15 observations. It is a mere coincidence, of course, that for this example Irwin's test gives the same results as the exact test based on the ratio S_1^2/S^2. In this connection, Irwin rightly calls attention to the fact that in dealing with a sample of only 15 observations the standard deviation of the sample is a very unreliable estimate of the population standard deviation.

It is remarked that here we would, of course, hesitate to apply the test $\frac{\bar{x}-x_1}{\sigma}$ of Section 5 above to the observation -1.40 as we do not have available an accurate estimate of σ from past data.

Example 2. We next consider an example of Pearson and Hartley [18] which is a modification of one used by "Student" [4] concerning the control of accuracy in routine chemical analyses. The problem of "Student" was one of checking day after day the conformance of some solution or substance to a standard by use of a small sample. We quote Pearson and Hartley [18], p.90: "The characteristic measured, for example the acidity of a solution, is estimated from the mean of a few (say n) observations, and a routine check on the consistency of these determinations is required to ensure that accuracy is maintained. Discordant observations will be repeated and, if necessary, rejected, and 'Student' pointed out that it would obviously be of advantage to work on a regular system; for this purpose he proposed the use of the range of the n determinations. The situation he considered was one in which the standard deviation of the within-day error of analysis had been found from experience to remain constant and could be assigned a known value, σ. It is clear, however, that situations will occur in which the standard error of analysis appropriate on a given day can only be estimated from the determinations of a few previous days, ...". In [18] Pearson and Hartley give percentage points for $q = (x_n - x_1)/s$, where x_n is the largest observation, x_1 the smallest and

$$s^2 = \frac{1}{\nu} \sum_{i=1}^{\nu+1} (x_i - \bar{x})^2$$

is determined from an <u>independent</u> sample (ν = d.f.)., i.e. in "Student's" problem s^2 based on ν d.f. is found from previous days and $x_1, x_2, \ldots x_n$ represent the sample values on the particular day the acidity of a solution is checked.

Pearson and Hartley cite as an example the four test results 22.8, 23.5, 26.0 and 26.6 on a particular day and give s=.675 as an estimate of σ which is determined from tests of five previous days and based on $\nu = 15$ degrees of freedom. Using the statistic, q, and the tables of percentage points for q in [18] Pearson and Hartley proceed according to "Student" and find that $x_4 - x_1 = 26.6 - 22.8 = 3.8$ is greater than can be ascribed to chance (5% level of significance). Consequently, an additional observation is called for and it is hypothesized that a retest gives the new result 23.9. (This new observation is called for since the estimate of a day's mean is to be based on at least four observations.) With a sample now of five observations, Pearson and Hartley now find that $x_5 - x_1 = 3.8$ is still too great and decide to reject 22.8 thus leaving 23.5, 26.0, 26.6 and 23.9. But now for the remaining four observations the sample range 26.6 - 23.5 = 3.1 is found to be significantly large as that a new test result is called for: suppose 23.5 is obtained. Finally, using <u>all six</u> observations and testing their significance with the statistic $q = (x_n - x_1)/S$, Pearson and Hartley following "Student's" procedure reject 22.8, reach a sample of five items, then reject 26.6 and wind up with a sample of four observations, i.e. 23.5, 26.0, 23.9, 23.5, whose range (2.5) is not too great and accept as the day's mean 1/4(23.5 + 26.0 + 23.9 + 23.5) = 24.2.

Our interest in the above example stems from the fact that the observation 26.0 in the final sample appears to be a single outlier and moreover it is of interest to compare the above procedure with the findings of this paper. Let us first apply the exact test for a single outlier to the final sample, 23.5, 26.0, 23.9, 23.5 without using any previous knowledge of σ. Here, we have $S^2 = 4.3075$ and $S_4^2 = 0.0107$ (leaving out $x_4 = 26.0$) so that $S_4^2/S^2 = 0.025$ which is significant at the 2.5% probability level—see Table I below for n=4. Thus, the observation 26.0 would be rejected under this test. On the other hand, using the value s=.675 based on past data for 15 d.f. and tables of the x^2-distribution, we find the $P\{\sigma \geq .970\} = 0.05$. Thus, taking a fair upper limit for σ and using the test $\frac{x_n - \bar{x}}{\sigma} = (26.0-24.2)/.970 = 1.86$, for which P = .06 from Table II and we may even say there is indicative significance. If we had used Irwin's criterion for a single outlier, $(x_n - x_{n-1})/\sigma$, we would obtain (26.0-23.9)/.970 = .216 and find from Table II of Irwin's paper [3] that $P < .05$; hence, we are inclined to doubt that the observation 26.0 is consistent with the other observations, even using the somewhat liberal maximum for $\sigma = .970$.

Suppose next we consider the entire sample of six observations: 22.8, 23.5, 26.0, 26.6, 23.9 and 23.5 ($\bar{x} = 24.38$). In view of the findings of the present work, we should probably want to apply a test based on the sums of squares about the mean of the observations 23.5, 23.9, and 23.5 divided by the sums of squares about the mean for all six items since we may suspect a "low" observation (22.8) and perhaps two "high" observations (26.0 and 26.6). If we suspected only the observations 22.8 and 26.6, Section 4 above indicates that a statistic based on the range provides an efficient test for the case in which both the greatest and least observations are equally outlying—this is the sort of test used by Pearson and Hartley in this example. Presumably, however, in using a test based on range we should reject simultaneously both the largest and smallest observations. Suppose, nevertheless, that instead of rejecting first the observation 22.8 and then the observation 26.6 (as Pearson and Hartley [18] did), we first reject using $(x_n - x_1)/s$ the observation 26.6 since $x_n - \bar{x}$ turns out to be greater than $\bar{x} - x_1$, i.e. 26.6 -24.38 = 2.22 and 24.38 - 22.8 = 1.58. We are then left with the five observations: 22.8, 23.5, 26.0, 23.9 and 23.5. Now employing the exact test S_n^2/S^2 to the suspected single outlier 26.0 where now $S^2 = 5.932$ and $S_n^2 = 0.6275$, we obtain $S_n^2/S^2 = 0.106$ and find from Table I that $0.025 < P < 0.05$; hence, 26.0 should be rejected, leaving 22.8, 23.5, 23.9 and 23.5 which appear "consistent" by using any of the tests. Moreover, using the figure of .970 as a maximum for the past σ and Irwin's criterion for the difference between the first two individuals we would also reject 26.0 from the sample of five (22.8, 23.5, 26.0, 23.9 and 23.5) or with σ = .970 and using Irwin's criterion for the difference between the second and third individuals [3] we would reject both 26.0 and 26.6 from the sample of six observations. Hence, our final sample would again consist of the four observations 22.8, 23.5, 23.9 and 23.5 with $\bar{x}$ = 23.4.

We remark that Pearson and Hartley were interested primarily in demonstrating "Student's" procedure using the statistic $(x_n - x_1)/s$, in the above example. Consequently, it is not the intention here to question the judgement of outstanding authorities in the field of applied statistics, but rather to point up the difficulties encountered with presently available tools and to indicate that certain combinations of appropriate tests may be valuable, depending on just what information is available.

<u>Example 3</u>. The following ranges (horizontal distances from gun muzzle to point of impact) were obtained in firing projectiles from a weapon at a constant angle of elevation and at the same weight of charge of propellant powder:

Distances in Yards

4782	4420
4838	4803
4765	4730
4549	4833

It is desired to know whether the projectiles exhibit uniformity in ballistic behavior or if some of the ranges, such as 4549 and 4420, are not consistent with the others.

Arranging the distances in increasing order of magnitude,

4420	4782
4549	4803
4730	4833
4765	4838

we suspect the presence of two outliers, i.e. 4420 and 4549. Having no available knowledge of σ from past data for this example, an efficient test to apply would be that of Section 9, i.e. $S^2_{1,2}/S^2$; however, not yet having available the percentage points of this distribution, we apply the test, S^2_1/S^2, suspecting that at least the observation 4420 is an outlier. From the entire sample of eight items we get $S^2 = 158592$ and leaving out 4420 we obtain $S^2_1 = 59134.9$ so that $S^2_1/S^2 = 0.3729$ and from Table I we find $0.05 < P < 0.10$. Thus, we do not necessarily reject the lowest observation, 4420. We are still not satisfied, however, for we know that the test, S^2_1/S^2, may be weak in the presence of more than a single outlier [1]. Hence, the following treatment is mentioned in passing as a temporary practical expedient (until the distribution of $S^2_{1,2}/S^2$ has been tabulated). Consider the last seven of the ordered observations, i.e.

4549	4782
4730	4803
4765	4833
	4838

and apply the exact test, S^2_1/S^2, to the smallest observation, 4549. We find then our new

$$S^2_1 = 8590.8 \qquad S^2 = 59134.9$$

and $S^2_1/S^2 = 0.1453$ so that $0.01 < P < .025$ from Table I and we should thus reject 4549 from the sample of seven. Moreover, we should now surely reject 4420 as being outlying. Hence, following this scheme as a practical expedient we would be inclined to reject both the observations, 4420 and 4549, leaving a sample of six items:

4782	4803
4838	4730
4765	4833

Apparently then, the two projectiles giving the short ranges of 4420 yards and 4549 yards behaved differently in ballistic characteristics from the others.

12. Summary. The problem of testing outlying observations, although an old one, is of considerable importance in applied statistics. Many and various types of significance tests have been proposed by statisticians interested in this field of application. In this connection, we bring out in the Historical Comments notable advances toward a clear formulation of the problem and important points which should be considered in attempting a complete solution. In Section 3 we state the situations the experimental statistician will very likely encounter in practice, these considerations being based on experience, and in Section 4 we proceed in a systematic manner to set up the statistical hypotheses to be tested when engaged in the problem of outlying observations and determine just what sample criteria provide efficient or best tests in the sense of Neyman-Pearson. Efficient tests for outlying observations appear to be based on (1) the sum of squares about the mean of all the observations in the sample and (2) the sum of squares about the mean of only the observations in the sample which appear consistent with one another. In case an accurate estimate of population standard deviation is available, say from past experience, we should employ a test of the difference between (1) and (2). On the other hand, for an exact test having no knowledge of σ we obtain for efficient tests the ratio of (2) to (1). For the case of a single outlying observation, σ assumed known, the best sample statistic consists of the difference between the extreme and mean. The distribution of this statistic is derived in Section 5 and its probability integral

is tabulated in Table II. It is apparent that the set of functions necessary for tabulating the probability integral of the difference between the extreme and sample mean may prove of wide use in Mathematical Statistics - probably in the field of order statistics. The set of functions referred to have been tabulated at the Ballistic Research Laboratories on the Electronic Numerical Integrator and Computer.

For the case in which the population standard deviation is unknown, we derive exact distributions for efficient sample statistics in testing the largest or the smallest observation in the sample (Section 8) and the two largest or two smallest observations (Section 9). The distribution of S_n^2/S^2 or S_1^2/S^2 in Section 8 has been tabulated on the Bell Relay Computors at the Ballistic Research Laboratories and it is hoped that priorities will permit the tabulation of the distribution called for in Section 9 in due course. In Section 10, we indicate some theoretical difficulties connected with determining in suitable computational form the distribution of efficient sample criteria for testing simultaneously the largest and smallest observation.

We conclude the dissertation with applications of our theory to practical problems and point out some useful methods of attack in testing outlying observations.

13. Acknowledgement. The author is greatly indebted to Prof. C.C. Craig under whose most competent guidance this work was carried out. Indeed, the stimulus and encouragement received by the author as a result of Prof. Craig's interest in the problem were paramount in orienting various phases of the subject matter and in accomplishing the results given herein. In connection with the computing, debts of gratitude are owed to several members of the staff of the Ballistic Research Laboratories. It is desired to express appreciation to Col. Leslie E. Simon, Director of the Ballistic Research Laboratories, for recognizing the importance and desirability of carrying out the computations on high-speed computing devices. In this connection, appreciation is also expressed to Dr. L.S. Dederick, Chief, Computing Laboratory, Ballistic Research Laboratories. The

programming of the $H_n(u)$ functions on the Electronic Numerical Integrator and Computor was done by Dr. Derrick Lehmer, University of California, who was with the Computing Laboratory, BRL, during the latter part of World War II and Miss Ruth Lichterman. The author is particularly indebted to Dr. Franz Alt, Dr. Bernard Dimsdale, Miss R. Lichterman, Mr. John Holberton, Miss H. Marks, Mr. F. Spence, and others of the Computing Laboratory during the period the functions $H_n(u)$ were computed on the ENIAC. The computing of the distribution of the statistic S_n^2/S^2 was done on the Bell Relay Computers under the direction of Mr. J.O. Harrison assisted by Miss M. Masincup. The author also desires to express appreciation to Miss Helen J. Coon for considerable computation and checking carried out on a table-model computing machine. Many thanks are due to Mr. Vernon Zink for expert typing and preparation of the manuscript.

Table 1

Table of Percentage Points for $\frac{S_n^2}{S^2}$ or $\frac{S_1^2}{S^2}$

Percentage Points

n	1%	2.5%	5%	10%
3	.0001	.0007	.0027	.0109
4	.0100	.0248	.0494	.0975
5	.0442	.0808	.1270	.1984
6	.0928	.1453	.2032	.2826
7	.1447	.2066	.2696	.3503
8	.1948	.2616	.3261	.4050
9	.2410	.3101	.3742	.4502
10	.2831	.3526	.4154	.4881
11	.3211	.3901	.4511	.5204
12	.3554	.4232	.4822	.5483
13	.3864	.4528	.5097	.5727
14	.4145	.4792	.5340	.5942
15	.4401	.5030	.5559	.6134
16	.4634	.5246	.5755	.6306
17	.4848	.5442	.5933	.6461
18	.5044	.5621	.6095	.6601
19	.5225	.5785	.6243	.6730
20	.5393	.5937	.6379	.6848
21	.5548	.6076	.6504	.6958
22	.5692	.6206	.6621	.7058
23	.5827	.6327	.6728	.7151
24	.5953	.6439	.6829	.7238
25	.6071	.6544	.6923	.7319

$$x_1 \leq x_2 \leq x_3 \leq \cdots \leq x_n$$

$$S^2 = \sum_{i=1}^{n} (x_i - \bar{x})^2 \quad \text{where} \quad \bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

$$S_n^2 = \sum_{i=1}^{n-1} (x_i - \bar{x}_n)^2 \quad \text{where} \quad \bar{x}_n = \frac{1}{n-1} \sum_{i=1}^{n-1} x_i$$

$$S_1^2 = \sum_{i=2}^{n} (x_i - \bar{x}_1)^2 \quad \text{where} \quad \bar{x}_1 = \frac{1}{n-1} \sum_{i=2}^{n} x_i$$

Table II

Probability Integral of the Extreme Minus the Mean, u_n, in Normal Samples of n Observations (Pop. S.D. as unit)

$P(u_n \leq u)$

u \ n	2	3	4	5	6	7	8	9	u
.00	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00
.05	.05637	.00309	.00017	.00001	.00000	.00000	.00000	.00000	.05
.10	.11246	.01251	.00134	.00015	.00002	.00000	.00000	.00000	.10
.15	.16800	.02745	.00445	.00072	.00012	.00002	.00000	.00000	.15
.20	.22270	.04817	.01033	.00221	.00047	.00010	.00002	.00000	.20
.25	.27633	.07403	.01966	.00520	.00137	.00036	.00010	.00003	.25
.30	.32863	.10450	.03292	.01033	.00324	.00101	.00032	.00010	.30
.35	.37938	.13896	.05040	.01820	.00656	.00236	.00085	.00031	.35
.40	.42839	.17677	.07218	.02935	.01191	.00482	.00195	.00079	.40
.45	.47548	.21724	.09816	.04416	.01982	.00889	.00398	.00178	.45
.50	.52050	.25968	.12807	.06288	.03080	.01507	.00737	.00360	.50
.55	.56332	.30344	.16152	.08559	.04525	.02390	.01261	.00665	.55
.60	.60386	.34788	.19801	.11219	.06344	.03583	.02022	.01140	.60
.65	.64203	.39243	.23697	.14246	.08547	.05121	.03067	.01836	.65
.70	.67780	.43656	.27781	.17602	.11130	.07030	.04437	.02800	.70
.75	.71116	.47983	.31992	.21242	.14076	.09318	.06164	.04076	.75
.80	.74210	.52185	.36274	.25113	.17353	.11978	.08263	.05698	.80
.85	.77067	.56230	.40571	.29160	.20920	.14993	.10739	.07688	.85
.90	.79691	.60095	.44835	.33325	.24727	.18329	.13578	.10055	.90
.95	.82089	.63761	.49021	.37555	.28721	.21945	.16757	.12791	.95
1.00	.84270	.67214	.53093	.41795	.32847	.25791	.20240	.15877	1.00
1.05	.86244	.70448	.57020	.45999	.37050	.29815	.23980	.19280	1.05
1.10	.88021	.73459	.60777	.50125	.41276	.33961	.27927	.22957	1.10
1.15	.89612	.76248	.64346	.54136	.45478	.38173	.32025	.26858	1.15
1.20	.91031	.78817	.67713	.58001	.49611	.42401	.36220	.30931	1.20
1.25	.92290	.81174	.70870	.61697	.53638	.46595	.40457	.35117	1.25
1.30	.93401	.83325	.73812	.65205	.57525	.50712	.44685	.39362	1.30
1.35	.94376	.85280	.76540	.68513	.61249	.54716	.48857	.43613	1.35
1.40	.95229	.87049	.79055	.71612	.64788	.58574	.52933	.47822	1.40
1.45	.95970	.88644	.81364	.74497	.68129	.62263	.56878	.51945	1.45
1.50	.96611	.90075	.83472	.77170	.71261	.65762	.60663	.55944	1.50
1.55	.97162	.91355	.85390	.79632	.74180	.69058	.64265	.59789	1.55
1.60	.97635	.92495	.87127	.81890	.76885	.72143	.67668	.63456	1.60
1.65	.98038	.93506	.88693	.83949	.79378	.75013	.70862	.66925	1.65
1.70	.98379	.94400	.90099	.85820	.81664	.77666	.73839	.70184	1.70

Table II (Cont'd.)

Probability Integral of the Extreme Minus the Mean, u_n, in Normal Samples of n Observations (Pop. S.D. as unit)

$P(u_n \leq u)$

u \ n	2	3	4	5	6	7	8	9	u
1.75	.98667	.95187	.91358	.87513	.83750	.80107	.76597	.73225	1.75
1.80	.98909	.95877	.92480	.89037	.85646	.82341	.79139	.76046	1.80
1.85	.99111	.96480	.93476	.90405	.87360	.84376	.81469	.78647	1.85
1.90	.99279	.97005	.94358	.91628	.88903	.86220	.83593	.81032	1.90
1.95	.99418	.97461	.95135	.92716	.90288	.87885	.85522	.83207	1.95
2.00	.99532	.97854	.95818	.93682	.91526	.89381	.87264	.85183	2.00
2.05	.99626	.98193	.96416	.94536	.92627	.90721	.88832	.86968	2.05
2.10	.99702	.98483	.96938	.95289	.93605	.91916	.90236	.88574	2.10
2.15	.99764	.98731	.97392	.95949	.94468	.92977	.91490	.90012	2.15
2.20	.99814	.98942	.97785	.96527	.95229	.93917	.92604	.91296	2.20
2.25	.99854	.99121	.98125	.97032	.95897	.94746	.93591	.92438	2.25
2.30	.99886	.99273	.98418	.97470	.96482	.95476	.94462	.93448	2.30
2.35	.99911	.99400	.98669	.97850	.96992	.96114	.95229	.94340	2.35
2.40	.99931	.99507	.98883	.98178	.97435	.96672	.95900	.95125	2.40
2.45	.99947	.99596	.99066	.98461	.97819	.97158	.96487	.95812	2.45
2.50	.99959	.99670	.99222	.98703	.98151	.97580	.96999	.96412	2.50
2.55	.99969	.99732	.99353	.98911	.98436	.97944	.97443	.96935	2.55
2.60	.99976	.99782	.99464	.99088	.98681	.98259	.97827	.97389	2.60
2.65	.99982	.99824	.99557	.99238	.98891	.98529	.98158	.97781	2.65
2.70	.99987	.99858	.99635	.99365	.99070	.98761	.98443	.98120	2.70
2.75	.99990	.99886	.99701	.99473	.99223	.98959	.98688	.98411	2.75
2.80	.99992	.99909	.99755	.99564	.99352	.99128	.98897	.98661	2.80
2.85	.99994	.99928	.99800	.99640	.99461	.99272	.99075	.98874	2.85
2.90	.99996	.99943	.99838	.99704	.99553	.99393	.99227	.99056	2.90
2.95	.99997	.99955	.99868	.99757	.99631	.99496	.99355	.99211	2.95
3.00	.99998	.99964	.99894	.99801	.99696	.99582	.99464	.99342	3.00
3.05	.99998	.99972	.99914	.99838	.99750	.99655	.99555	.99453	3.05
3.10	.99999	.99978	.99931	.99868	.99795	.99716	.99632	.99546	3.10
3.15	.99999	.99983	.99945	.99893	.99832	.99766	.99697	.99625	3.15
3.20	.99999	.99987	.99956	.99913	.99863	.99808	.99750	.99690	3.20
3.25	1.00000	.99990	.99965	.99930	.99889	.99843	.99795	.99745	3.25
3.30		.99992	.99972	.99944	.99910	.99872	.99832	.99791	3.30
3.35		.99994	.99978	.99955	.99927	.99896	.99863	.99829	3.35
3.40		.99995	.99983	.99964	.99941	.99916	.99889	.99860	3.40
3.45		.99996	.99986	.99971	.99953	.99932	.99910	.99886	3.45

Table II (Cont'd.)

Probability Integral of the Extreme Minus the Mean, u_n, in Normal Samples of N Observations (Pop. S.D. as unit)

$P(u_n \leq u)$

u \ n	2	3	4	5	6	7	8	9	u
3.50		.99997	.99989	.99977	.99962	.99945	.99927	.99908	3.50
3.55		.99998	.99992	.99982	.99970	.99956	.99941	.99925	3.55
3.60		.99998	.99994	.99986	.99976	.99965	.99952	.99940	3.60
3.65		.99999	.99995	.99989	.99981	.99972	.99962	.99951	3.65
3.70		.99999	.99996	.99991	.99985	.99977	.99969	.99961	3.70
3.75		.99999	.99997	.99993	.99988	.99982	.99976	.99969	3.75
3.80		1.00000	.99998	.99995	.99991	.99986	.99981	.99975	3.80
3.85			.99998	.99996	.99993	.99989	.99985	.99980	3.85
3.90			.99999	.99997	.99994	.99991	.99988	.99984	3.90
3.95			.99999	.99997	.99995	.99993	.99990	.99987	3.95
4.00			.99999	.99998	.99996	.99995	.99992	.99990	4.00
4.05			.99999	.99999	.99997	.99996	.99994	.99992	4.05
4.10			1.00000	.99999	.99998	.99997	.99995	.99994	4.10
4.15				.99999	.99998	.99997	.99996	.99995	4.15
4.20				.99999	.99999	.99998	.99997	.99996	4.20
4.25				.99999	.99999	.99998	.99998	.99997	4.25
4.30				1.00000	.99999	.99999	.99998	.99998	4.30
4.35					.99999	.99999	.99999	.99998	4.35
4.40					1.00000	.99999	.99999	.99999	4.40
4.45						.99999	.99999	.99999	4.45
4.50						1.00000	.99999	.99999	4.50
4.55							1.00000	.99999	4.55
4.60								1.00000	4.60

Table II (Cont'd.)

Probability Integral of the Extreme Minus the Mean, u_n, in Normal Samples of n Observations (Pop. S.D. as unit)

$P(u_n < u)$

u \ n	10	11	12	13	14	15	16	17	u
.25	.00001	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.25
.30	.00003	.00001	.00000	.00000	.00000	.00000	.00000	.00000	.30
.35	.00011	.00004	.00001	.00001	.00000	.00000	.00000	.00000	.35
.40	.00032	.00013	.00005	.00002	.00001	.00000	.00000	.00000	.40
.45	.00080	.00036	.00016	.00007	.00003	.00001	.00001	.00000	.45
.50	.00176	.00086	.00042	.00021	.00010	.00005	.00002	.00001	.50
.55	.00351	.00185	.00098	.00051	.00027	.00014	.00008	.00004	.55
.60	.00643	.00363	.00204	.00115	.00065	.00037	.00021	.00012	.60
.65	.01098	.00657	.00393	.00235	.00141	.00084	.00050	.00030	.65
.70	.01766	.01113	.00702	.00443	.00279	.00176	.00111	.00070	.70
.75	.02694	.01780	.01177	.00777	.00514	.00339	.00224	.00148	.75
.80	.03928	.02707	.01865	.01285	.00886	.00610	.00420	.00289	.80
.85	.05503	.03938	.02818	.02016	.01442	.01031	.00738	.00527	.85
.90	.07444	.05510	.04077	.03017	.02232	.01652	.01222	.00904	.90
.95	.09761	.07448	.05682	.04334	.03305	.02521	.01922	.01466	.95
1.00	.12452	.09763	.07655	.06000	.04703	.03687	.02889	.02265	1.00
1.05	.15497	.12454	.10008	.08041	.06460	.05190	.04169	.03348	1.05
1.10	.18867	.15503	.12737	.10464	.08595	.07060	.05799	.04762	1.10
1.15	.22520	.18879	.15825	.13263	.11116	.09315	.07806	.06541	1.15
1.20	.26407	.22542	.19240	.16420	.14013	.11957	.10203	.08706	1.20
1.25	.30475	.26442	.22941	.19901	.17263	.14973	.12987	.11264	1.25
1.30	.34666	.30525	.26876	.23662	.20830	.18336	.16140	.14207	1.30
1.35	.38924	.34734	.30992	.27650	.24667	.22005	.19629	.17509	1.35
1.40	.43196	.39011	.35229	.31810	.28721	.25931	.23411	.21135	1.40
1.45	.47430	.43302	.39529	.36082	.32934	.30058	.27433	.25036	1.45
1.50	.51583	.47555	.43838	.40408	.37244	.34327	.31636	.29156	1.50
1.55	.55615	.51726	.48104	.44733	.41595	.38676	.35960	.33434	1.55
1.60	.59495	.55774	.52282	.49004	.45930	.43046	.40342	.37807	1.60
1.65	.63196	.59668	.56332	.53178	.50199	.47384	.44726	.42216	1.65
1.70	.66699	.63380	.60221	.57216	.54358	.51641	.49058	.46602	1.70
1.75	.69991	.66892	.63925	.61086	.58370	.55773	.53289	.50915	1.75
1.80	.73063	.70189	.67424	.64763	.62204	.59744	.57380	.55108	1.80
1.85	.75912	.73264	.70704	.68229	.65838	.63528	.61297	.59144	1.85
1.90	.78538	.76113	.73758	.71472	.69254	.67102	.65016	.62992	1.90
1.95	.80945	.78737	.76584	.74486	.72443	.70453	.68516	.66630	1.95

Table II (Cont'd.)

Probability Integral of the Extreme Minus the Mean, u_n, in Normal Samples of n Observations (Pop. S.D. as unit)

$P(u_n < u)$

u \ n	10	11	12	13	14	15	16	17	u
2.00	.83141	.81140	.79183	.77269	.75399	.73571	.71786	.70042	2.00
2.05	.85133	.83330	.81560	.79824	.78121	.76453	.74819	.73218	2.05
2.10	.86932	.85314	.83721	.82155	.80614	.79101	.77614	.76153	2.10
2.15	.88550	.87105	.85678	.84271	.82885	.81519	.80174	.78849	2.15
2.20	.89998	.88713	.87440	.86183	.84941	.83715	.82505	.81311	2.20
2.25	.91290	.90151	.89021	.87902	.86795	.85699	.84616	.83545	2.25
2.30	.92437	.91431	.90432	.89441	.88458	.87484	.86518	.85563	2.30
2.35	.93453	.92568	.91688	.90812	.89943	.89081	.88224	.87375	2.35
2.40	.94348	.93572	.92799	.92030	.91264	.90504	.89748	.88997	2.40
2.45	.95134	.94457	.93781	.93106	.92435	.91766	.91101	.90440	2.45
2.50	.95823	.95233	.94644	.94055	.93468	.92883	.92300	.91720	2.50
2.55	.96424	.95912	.95400	.94887	.94376	.93866	.93357	.92850	2.55
2.60	.96948	.96504	.96060	.95616	.95172	.94728	.94285	.93844	2.60
2.65	.97401	.97019	.96635	.96251	.95866	.95482	.95098	.94715	2.65
2.70	.97793	.97464	.97134	.96802	.96471	.96139	.95807	.95475	2.70
2.75	.98131	.97849	.97565	.97280	.96995	.96709	.96423	.96137	2.75
2.80	.98422	.98180	.97937	.97693	.97448	.97203	.96957	.96712	2.80
2.85	.98671	.98464	.98257	.98048	.97839	.97629	.97418	.97208	2.85
2.90	.98883	.98708	.98531	.98353	.98174	.97995	.97816	.97636	2.90
2.95	.99064	.98915	.98765	.98614	.98462	.98309	.98156	.98003	2.95
3.00	.99218	.99092	.98965	.98837	.98708	.98578	.98448	.98318	3.00
3.05	.99348	.99242	.99134	.99026	.98917	.98807	.98697	.98587	3.05
3.10	.99458	.99369	.99278	.99187	.99095	.99002	.98909	.98816	3.10
3.15	.99551	.99476	.99400	.99323	.99245	.99167	.99089	.99010	3.15
3.20	.99628	.99566	.99502	.99437	.99372	.99307	.99241	.99175	3.20
3.25	.99694	.99641	.99588	.99534	.99479	.99424	.99369	.99314	3.25
3.30	.99748	.99704	.99660	.99615	.99569	.99523	.99477	.99431	3.30
3.35	.99793	.99757	.99720	.99682	.99644	.99606	.99568	.99529	3.35
3.40	.99831	.99801	.99770	.99739	.99707	.99676	.99644	.99611	3.40
3.45	.99862	.99837	.99812	.99786	.99760	.99733	.99707	.99680	3.45
3.50	.99888	.99867	.99846	.99825	.99803	.99781	.99759	.99737	3.50
3.55	.99909	.99892	.99875	.99857	.99839	.99821	.99803	.99785	3.55
3.60	.99926	.99912	.99898	.99884	.99869	.99854	.99839	.99824	3.60
3.65	.99940	.99929	.99917	.99906	.99894	.99881	.99869	.99857	3.65
3.70	.99952	.99943	.99933	.99924	.99914	.99904	.99894	.99883	3.70

Table II (Cont'd.)

Probability Integral of the Extreme Minus the Mean, u_n, in Normal Samples of n Observations (Pop. S.D. as unit)

$P(u_n \leq u)$

u \ n	10	11	12	13	14	15	16	17	u
3.75	.99961	.99954	.99946	.99938	.99930	.99922	.99914	.99905	3.75
3.80	.99969	.99963	.99957	.99950	.99944	.99937	.99930	.99923	3.80
3.85	.99975	.99970	.99965	.99960	.99955	.99949	.99944	.99938	3.85
3.90	.99980	.99976	.99972	.99968	.99964	.99959	.99955	.99950	3.90
3.95	.99984	.99981	.99978	.99974	.99971	.99967	.99964	.99960	3.95
4.00	.99988	.99985	.99982	.99980	.99977	.99974	.99971	.99968	4.00
4.05	.99990	.99988	.99986	.99984	.99982	.99979	.99977	.99974	4.05
4.10	.99992	.99991	.99989	.99987	.99985	.99983	.99981	.99979	4.10
4.15	.99994	.99993	.99991	.99990	.99988	.99987	.99985	.99984	4.15
4.20	.99995	.99994	.99993	.99992	.99991	.99990	.99988	.99987	4.20
4.25	.99996	.99995	.99995	.99994	.99993	.99992	.99991	.99990	4.25
4.30	.99997	.99996	.99996	.99995	.99994	.99993	.99993	.99992	4.30
4.35	.99998	.99997	.99997	.99996	.99996	.99995	.99994	.99993	4.35
4.40	.99998	.99998	.99997	.99997	.99996	.99996	.99995	.99995	4.40
4.45	.99999	.99998	.99998	.99998	.99997	.99997	.99996	.99996	4.45
4.50	.99999	.99999	.99998	.99998	.99998	.99998	.99997	.99997	4.50
4.55	.99999	.99999	.99999	.99999	.99998	.99998	.99998	.99997	4.55
4.60	.99999	.99999	.99999	.99999	.99999	.99998	.99998	.99998	4.60
4.65	1.00000	.99999	.99999	.99999	.99999	.99999	.99999	.99998	4.65
4.70		1.00000	.99999	.99999	.99999	.99999	.99999	.99999	4.70
4.75			1.00000	.99999	.99999	.99999	.99999	.99999	4.75
4.80				1.00000	1.00000	.99999	.99999	.99999	4.80
4.85						1.00000	.99999	.99999	4.85
4.90							.99999	.99999	4.90
4.95							1.00000	.99999	4.95
5.00								.99999	5.00

Table II (Cont'd.)

Probability Integral of the Extreme Minus the Mean, u_n, in Normal Samples of n Observations (Pop. S.D. as unit)

$P(u_n \leq u)$

u \ n	18	19	20	21	22	23	24	25	u
.50	.00001	.00000	.0000	.0000	.0000	.0000	.0000	.0000	.50
.55	.00002	.00001	.0000	.0000	.0000	.0000	.0000	.0000	.55
.60	.00007	.00004	.0000	.0000	.0000	.0000	.0000	.0000	.60
.65	.00018	.00011	.0001	.0000	.0000	.0000	.0000	.0000	.65
.70	.00044	.00028	.0002	.0001	.0001	.0000	.0000	.0000	.70
.75	.00098	.00065	.0004	.0003	.0002	.0001	.0001	.0001	.75
.80	.00199	.00137	.0009	.0007	.0004	.0003	.0002	.0001	.80
.85	.00377	.00270	.0019	.0014	.0010	.0007	.0005	.0004	.85
.90	.00669	.00494	.0037	.0027	.0020	.0015	.0011	.0008	.90
.95	.01118	.00855	.0065	.0049	.0038	.0029	.0022	.0017	.95
1.00	.01775	.01391	.0109	.0085	.0067	.0052	.0041	.0032	1.00
1.05	.02690	.02161	.0174	.0139	.0112	.0090	.0072	.0058	1.05
1.10	.03911	.03212	.0264	.0217	.0178	.0146	.0120	.0099	1.10
1.15	.05481	.04592	.0385	.0322	.0270	.0226	.0190	.0159	1.15
1.20	.07428	.06338	.0541	.0461	.0394	.0336	.0287	.0244	1.20
1.25	.09769	.08472	.0735	.0637	.0553	.0479	.0416	.0360	1.25
1.30	.12504	.11005	.0969	.0853	.0750	.0660	.0581	.0512	1.30
1.35	.15618	.13930	.1242	.1108	.0988	.0882	.0786	.0701	1.35
1.40	.19080	.17225	.1555	.1404	.1267	.1144	.1033	.0932	1.40
1.45	.22848	.20851	.1903	.1736	.1585	.1446	.1320	.1204	1.45
1.50	.26869	.24761	.2282	.2103	.1938	.1786	.1646	.1516	1.50
1.55	.31084	.28899	.2687	.2498	.2322	.2159	.2007	.1866	1.55
1.60	.35430	.33202	.3111	.2916	.2732	.2560	.2399	.2248	1.60
1.65	.39845	.37607	.3549	.3349	.3162	.2984	.2816	.2658	1.65
1.70	.44269	.42052	.3994	.3794	.3604	.3424	.3252	.3089	1.70
1.75	.48645	.46476	.4440	.4242	.4053	.3872	.3699	.3534	1.75
1.80	.52924	.50827	.4881	.4687	.4502	.4323	.4152	.3987	1.80
1.85	.57065	.55058	.5312	.5125	.4945	.4771	.4603	.4441	1.85
1.90	.61031	.59130	.5729	.5549	.5377	.5209	.5047	.4890	1.90
1.95	.64796	.63011	.6127	.5958	.5794	.5634	.5479	.5328	1.95
2.00	.68340	.66678	.6506	.6348	.6193	.6042	.5895	.5752	2.00
2.05	.71650	.70114	.6861	.6714	.6570	.6429	.6291	.6156	2.05
2.10	.74719	.73311	.7193	.7058	.6924	.6793	.6665	.6540	2.10
2.15	.77545	.76262	.7500	.7375	.7254	.7133	.7015	.6899	2.15
2.20	.80132	.78971	.7782	.7670	.7558	.7448	.7340	.7234	2.20

Table II (Cont'd.)

Probability Integral of the Extreme Minus the Mean, u_n, in Normal Samples of n Observations (Pop. S.D. as unit)

$P(u_n \leq u)$

u \ n	18	19	20	21	22	23	24	25	u
2.25	.82486	.81440	.8041	.7938	.7838	.7738	.7640	.7543	2.25
2.30	.84616	.83679	.8275	.8184	.8093	.8003	.7914	.7827	2.30
2.35	.86533	.85699	.8487	.8405	.8324	.8244	.8164	.8085	2.35
2.40	.88251	.87511	.8678	.8605	.8533	.8461	.8390	.8319	2.40
2.45	.89783	.89129	.8848	.8784	.8720	.8656	.8593	.8530	2.45
2.50	.91142	.90568	.9000	.8943	.8887	.8831	.8775	.8719	2.50
2.55	.92345	.91842	.9134	.9084	.9035	.8985	.8936	.8888	2.55
2.60	.93404	.92965	.9253	.9209	.9166	.9123	.9080	.9037	2.60
2.65	.94332	.93951	.9357	.9319	.9282	.9244	.9207	.9169	2.65
2.70	.95144	.94814	.9448	.9416	.9382	.9351	.9318	.9286	2.70
2.75	.95852	.95567	.9528	.9500	.9472	.9444	.9415	.9387	2.75
2.80	.96466	.96220	.9598	.9573	.9549	.9524	.9500	.9476	2.80
2.85	.96997	.96787	.9658	.9637	.9616	.9595	.9574	.9553	2.85
2.90	.97456	.97275	.9710	.9692	.9674	.9656	.9638	.9620	2.90
2.95	.97850	.97696	.9754	.9739	.9724	.9709	.9693	.9678	2.95
3.00	.98187	.98057	.9793	.9780	.9767	.9753	.9741	.9728	3.00
3.05	.98476	.98365	.9825	.9814	.9803	.9793	.9781	.9771	3.05
3.10	.98722	.98629	.9853	.9844	.9835	.9826	.9816	.9807	3.10
3.15	.98931	.98852	.9877	.9869	.9862	.9853	.9846	.9838	3.15
3.20	.99108	.99042	.9898	.9891	.9884	.9878	.9871	.9865	3.20
3.25	.99258	.99202	.9915	.9909	.9904	.9898	.9893	.9887	3.25
3.30	.99384	.99337	.9929	.9924	.9920	.9915	.9911	.9906	3.30
3.35	.99490	.99451	.9941	.9937	.9933	.9930	.9926	.9922	3.35
3.40	.99579	.99546	.9951	.9948	.9945	.9942	.9939	.9936	3.40
3.45	.99653	.99626	.9960	.9957	.9955	.9952	.9949	.9947	3.45
3.50	.99715	.99693	.9967	.9965	.9963	.9961	.9958	.9956	3.50
3.55	.99766	.99748	.9973	.9971	.9969	.9968	.9966	.9964	3.55
3.60	.99809	.99794	.9978	.9976	.9975	.9973	.9972	.9971	3.60
3.65	.99844	.99832	.9982	.9981	.9979	.9978	.9977	.9976	3.65
3.70	.99873	.99863	.9985	.9984	.9983	.9982	.9982	.9981	3.70
3.75	.99897	.99889	.9988	.9987	.9986	.9986	.9985	.9984	3.75
3.80	.99917	.99910	.9990	.9990	.9989	.9988	.9988	.9988	3.80
3.85	.99933	.99927	.9992	.9992	.9991	.9991	.9990	.9990	3.85
3.90	.99946	.99941	.9994	.9993	.9993	.9993	.9992	.9992	3.90
3.95	.99956	.99953	.9995	.9995	.9994	.9994	.9994	.9994	3.95

Table II (Cont'd.)

Probability Integral of the Extreme Minus the Mean, u_n, in Normal Sample of n Observations (Pop. S.D. as unit)

$P(u_n \leq u)$

u \ n	18	19	20	21	22	23	24	25	u
4.00	.99965	.99962	.9996	.9996	.9995	.9995	.9995	.9995	4.00
4.05	.99972	.99969	.9997	.9996	.9996	.9996	.9996	.9996	4.05
4.10	.99977	.99975	.9997	.9997	.9997	.9997	.9997	.9997	4.10
4.15	.99982	.99980	.9998	.9998	.9998	.9998	.9998	.9998	4.15
4.20	.99986	.99984	.9998	.9998	.9998	.9998	.9998	.9998	4.20
4.25	.99989	.99987	.9999	.9999	.9999	.9999	.9999	.9999	4.25
4.30	.99991	.99990	.9999	.9999	.9999	.9999	.9999	.9999	4.30
4.35	.99993	.99992	.9999	.9999	.9999	.9999	.9999	.9999	4.35
4.40	.99994	.99994	.9999	.9999	.9999	.9999	.9999	.9999	4.40
4.45	.99995	.99995	1.0000	.9999	.9999	.9999	.9999	.9999	4.45
4.50	.99996	.99996		1.0000	1.0000	1.0000	.9999	.9999	4.50
4.55	.99997	.99997					1.0000	1.0000	4.55
4.60	.99998	.99997							4.60
4.65	.99998	.99998							4.65
4.70	.99998	.99998							4.70
4.75	.99999	.99998							4.75
4.80	.99999	.99999							4.80
4.85	.99999	.99999							4.85
4.90	1.00000	1.00000							4.90

Table III

Percentage Points for Extreme Minus Mean

n	90%	95%	99%	99.5%
2	1.163	1.386	1.821	1.985
3	1.497	1.738	2.215	2.396
4	1.696	1.941	2.431	2.618
5	1.835	2.080	2.574	2.764
6	1.939	2.184	2.679	2.870
7	2.022	2.267	2.761	2.952
8	2.091	2.334	2.828	3.019
9	2.150	2.392	2.884	3.074
10	2.200	2.441	2.931	3.122
11	2.245	2.484	2.973	3.163
12	2.284	2.523	3.010	3.199
13	2.320	2.557	3.043	3.232
14	2.352	2.589	3.072	3.261
15	2.382	2.617	3.099	3.287
16	2.409	2.644	3.124	3.312
17	2.434	2.668	3.147	3.334
18	2.458	2.691	3.168	3.355
19	2.480	2.712	3.188	3.375
20	2.500	2.732	3.207	3.393
21	2.519	2.750	3.224	3.409
22	2.538	2.768	3.240	3.425
23	2.555	2.784	3.255	3.439
24	2.571	2.800	3.269	3.453
25	2.587	2.815	3.282	3.465

Table IV

Moment Constants for Extreme Minus Mean

n	Mean	Std. Dev.	α_3	α_4
2	.5642	.4263	.9953	3.8692
3	.8463	.4755	.8296	3.7135
4	1.0294	.4916	.7675	3.6717
5	1.1630	.4974	.7372	3.6560
6	1.2672	.4993	.7165	3.6511
7	1.3522	.4991	.7042	3.6503
8	1.4236	.4979	.6959	3.6518
9	1.4850	.4962	.6900	3.6546
10	1.5388	.4943	.6857	3.6582
11	1.5864	.4923	.6827	3.6622
12	1.6292	.4902	.6804	3.6663
13	1.6680	.4881	.6788	3.6705
14	1.7034	.4861	.6777	3.6746
15	1.7359	.4841	.6770	3.6787
20	1.867	.475	.677	3.700
60	2.319	.436	.699	3.801
100	2.508	.418	.712	3.855
200	2.746	.395	.737	3.932
500	3.037	.368	.771	4.033
1000	3.241	.350	.794	4.105

References

[1] E.S. Pearson and C. Chandra Sekar, "The Efficiency of Statistical Tools and a Criterion for the Rejection of Outlying Observations", Biometrika, Vol. 28 (1936), pp. 308-320.

[2] P.R. Rider, "Criteria for Rejection of Observations", Washington University Studies - New Series, Science and Technology - No. 8, St. Louis (1933).

[3] J.O. Irwin, "On a Criterion for the Rejection of Outlying Observations", Biometrika, Vol. 17 (1925), pp. 238-250.

[4] "Student", Biometrika, Vol. 19 (1927), pp. 151-164.

[5] L.H.C. Tippett, "The Extreme Individuals and the Range of Samples Taken From a Normal Population", Biometrika, Vol. 17 (1925), pp. 151-164.

[6] E.S. Pearson, "A Further Note on the Distribution of Range in Samples Taken From a Normal Population", Biometrika, Vol. 18 (1926), pp. 173-194.

[7] "Tables for Statisticians and Biometricians", Part II, edited by Karl Pearson, pp. CX-CXIX.

[8] E.S. Pearson, "The Percentage Limits for the Distribution of Range in Samples From a Normal Population", Biometrika, Vol. 24 (1932), pp. 404-417.

[9] H.O. Hartley, "The Range in Random Samples", Biometrika, Vol. 32 (1942), pp. 334-348.

[10] E.S. Pearson and H.O. Hartley, "The Probability Integral of the Range in Samples of n Observations From a Normal Population", Biometrika, Vol. 32 (1942), pp. 301-310.

[11] A.T. McKay, "The Distribution of the Difference Between the Extreme Observation and the Sample Mean in Samples of n From a Normal Universe", Biometrika, Vol. 27 (1935), pp.466-471.

[12] W.R. Thompson, "On a Criterion for the Rejection of Observations and the Distribution of the Ratio of the Deviation to the Sample Standard Deviation", Annals of Math. Stat., Vol. 6 (1935), pp. 214-219.

[13] H.J. Godwin, "On the Distribution of the Estimate of Mean Deviation Obtained From Samples From a Normal Population", Biometrika, Vol. 33 (1945), pp. 254-256.

[14] W.P.A., "Tables of Probability Functions", Vols. I and II, New York, N.Y. (1942).

[15] H.O. Hartley, "Note on the Calculation of the Distribution of the Estimate of Mean Deviation in Normal Samples", Biometrika, Vol. 33 (1945), pp. 257-258.

[16] "Tables of the Probability Integral of the Mean Deviation in Normal Samples", Biometrika, Vol. 33 (1945), pp. 259-265.

[17] J. Neyman and E.S. Pearson, "On the Problem of the Most Efficient Tests of Statistical Hypotheses", Phil. Trans. Royal Society (London), Vol. 231 (1933), pp. 289-337.

[18] E.S. Pearson and H.O. Hartley, "Tables of the Probability Integral of the Studentized Range,", Biometrika, Vol. 33 (1943), pp. 89-99.

[19] K. Pearson, "Tables of the Incomplete Beta-Function", published by the Biometrika Office, University College, London.

[20] K. R. Nair, "The Distribution of the Extreme Deviate From the Sample Mean and Its Studentized Form," Biometrika, Vol. 35 (1948), pp. 118-144.