

Division of Research
Graduate School of Business Administration
The University of Michigan

January 1980

A GENERAL BAYESIAN MODEL
FOR COMPLIANCE TESTING

Working Paper No. 204

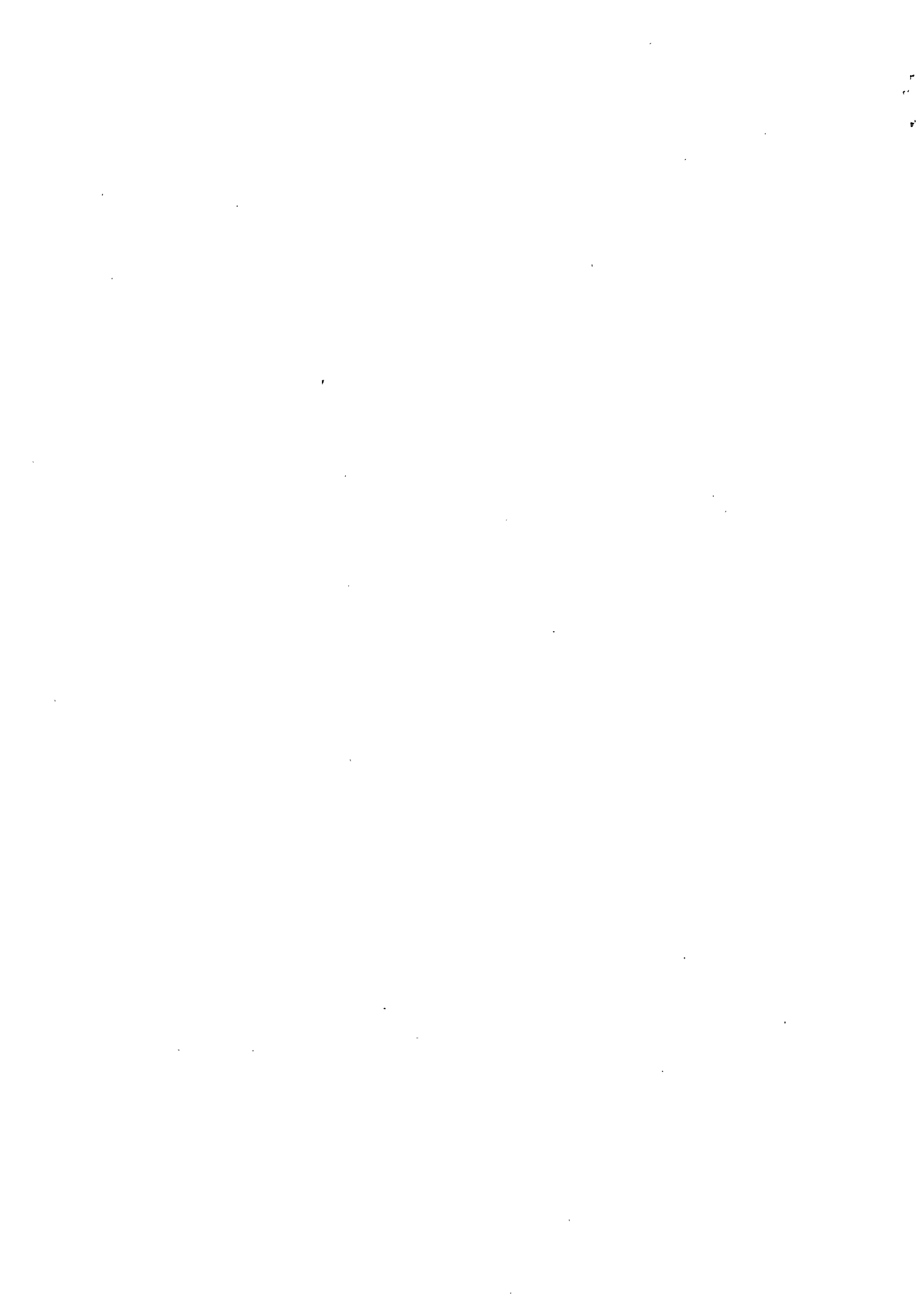
J. T. Godfrey

R. W. Andrews

The University of Michigan

FOR DISCUSSION PURPOSES ONLY

None of this material is to be quoted or
reproduced without the express permission
of the Division of Research.



A General Bayesian Model for Compliance Testing

J. T. Godfrey and R. W. Andrews
The University of Michigan

Introduction

Statistical sampling techniques are accepted as procedures that may be used by an auditor. In particular, in testing for compliance with internal accounting control procedures, attribute sampling methods are widely used.¹ For these tests the objective is to conclude whether or not a rate of non-compliance with a specific attribute of an internal control system is greater than or equal to a predetermined level.

In this paper we develop a Bayesian decision model for compliance testing by auditors. We describe the process that may be followed by an auditor in using our model. This process begins with an elicitation of a prior probability distribution for the error rate and ends with a determination of the sample size required to accomplish a set of auditing and statistical objectives. We show that for a suitable set of alternative prior distributions from which an auditor may choose, it is possible to develop tables of required sample sizes for accomplishing alternative statistical objectives. The primary emphasis of this paper is the statistical developments that lead from a very general and flexible prior distribution on a population error rate to a discrete posterior distribution on the number of errors in a finite population. Through the posterior distribution we are able to calculate the required sample size.

¹The AICPA has provided support in the statistical sampling area under the Individual Study Program of its Continuing Professional Education. In this program, Volume 2--Sampling for Attributes, Volume 4--Discovery Sampling, and Volume 6--Field Manual of Statistical Sampling relate to statistical methods for use in the study and evaluation of internal control systems.

Support for this research came from the Graduate School of Business Administration at The University of Michigan and is gratefully appreciated.

Current Compliance Test Methodology

In compliance test applications a sample is usually taken without replacement from a finite population. Thus the hypergeometric distribution is the appropriate probability distribution to use; in the case of large populations, however, the binomial distribution is often used instead.² We use the binomial distribution given below to illustrate current attribute test methodology:

$$P[r \leq r^*] = \sum_{i=0}^{r^*} \binom{n}{i} p^i (1 - p)^{n-i},$$

where r = number of errors observed in a sample,

n = sample size, and

p = probability that an item will be in error.

If p^* is the maximum tolerable error rate (desired upper precision limit), REL^* is the desired reliability level, and r^* is the acceptance criterion for errors in a sample, then we can solve the following for n , the required sample size.

$$P[r \leq r^*] = \sum_{i=0}^{r^*} \binom{n}{i} p^{*i} (1 - p^*)^{n-i} = 1 - REL^*$$

If we already have sample results (n, r^*) we can either insert p^* into the equation and solve for the achieved reliability or insert REL^* and solve for the achieved upper precision limit at the given reliability level. Thus, if we have any three components of the set (n, r^*, p^*, REL^*) , we can solve for the fourth. This is the procedure commonly followed in various attribute sampling plans used by auditors. It is clear that for given r^* , p^* , and REL^* the required sample size is the same, despite any experience or knowledge acquired in prior audits.

²The Poisson Distribution is sometimes used as an approximation of the binomial for the purpose of calculational ease.

The upper precision limit is a value for the unknown population error rate such that if it were the true error rate, the probability of obtaining the observed sample results (or worse) is equal to REL*, the reliability level. The upper precision limit, p*, is found by solving the following for p*,

$$\sum_{i=r^*+1}^n \binom{n}{i} (p^*)^i (1-p^*)^{n-i} = \text{REL}^*,$$

where r* = number of errors for acceptance of sample results. Implicit in this, the classical approach, is the view that the population error rate is an unknown constant. In the Bayesian approach the population error rate is viewed as a random variable. A statistical conclusion resulting from the Bayesian approach is a probabilistic statement about the population error rate, while a conclusion resulting from the classical approach is a statement about the probability of getting results worse than those actually obtained, given that p* is the population error rate.

A Bayesian statistical model explicitly incorporates into a solution a decision maker's past knowledge about a decision problem. This knowledge takes the form of a prior probability distribution on a range of possible error rates that may exist. For example, if control had been excellent in the past, then most of the probability could be located near an error rate equal to zero; on the other hand, if controls had been only fair, most of the probability could be located on error rates greater than zero. Figure 1 depicts these two situations by means of two different probability density functions, f₁(p) and f₂(p), respectively. It is reasonable to expect that a sampling plan for a current period will be influenced by the results from past sampling plans, as well as by other knowledge gained from a preliminary review of a client's

internal control system. The Bayesian model we propose has this characteristic. We believe that a Bayesian decision model is a more realistic representation of the decision problem and provides a more logical conclusion than does the classical model.

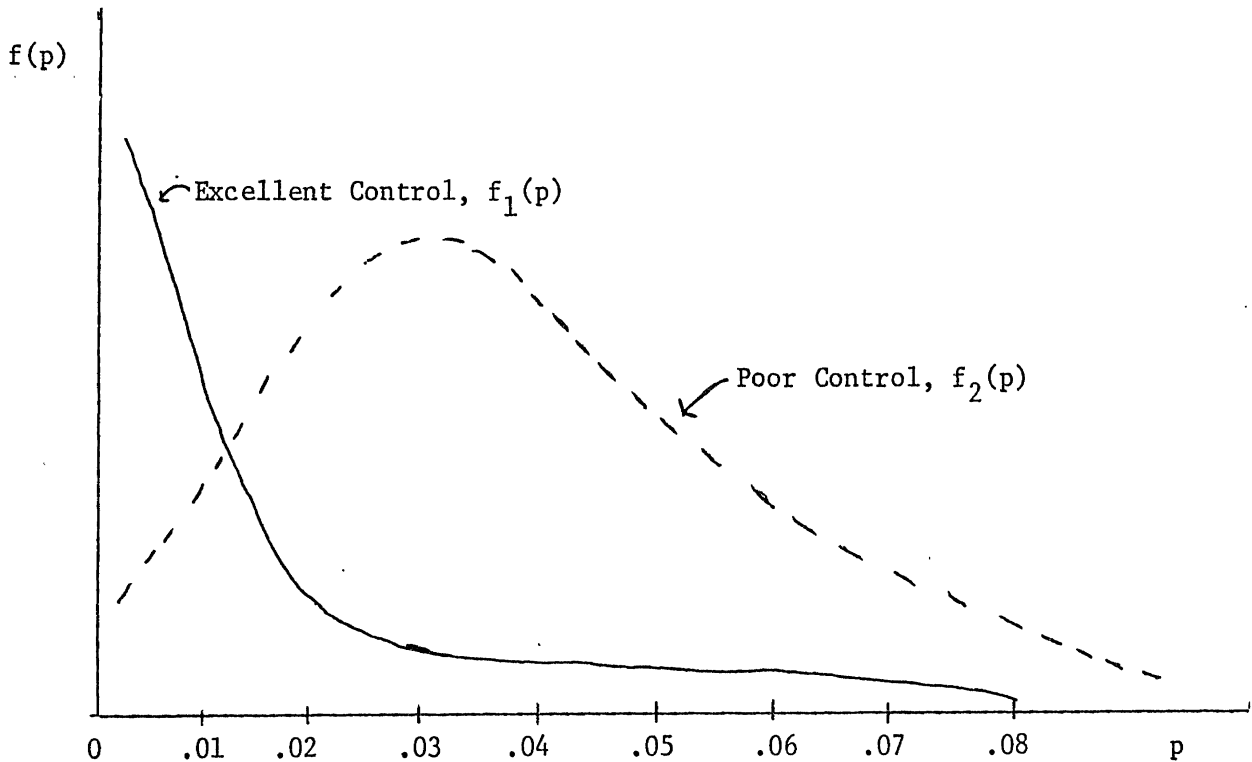


Figure 1

Bayesian Model Development

The development of Bayesian decision models in auditing has been proposed by other authors. In the compliance testing area, Kraft's analysis [1968] used three specific prior distributions on six possible error rates and three sample sizes. Kraft demonstrated how, for different numbers of errors observed in a constant sample size, the cumulative posterior distribution changes, given a specific prior distribution. His results also demonstrated how, for a fixed number of errors observed and fixed prior distribution, the cumulative posterior distribution changes, given different sample sizes. Our model is more general than Kraft's

in that our prior distribution provides greater flexibility than the three specific prior distributions Kraft considered and is not limited to six possible error rates.

Tracy [1969] also suggested a Bayesian model for compliance testing and used the hypergeometric distribution to produce a discrete posterior distribution on the number of errors in a finite population. Our model also produces a discrete posterior distribution, but we also show how such a distribution may be systematically linked to a general class of prior distributions and used to determine required sample sizes.

Francisco [1972] investigated the choice between discrete and continuous prior distributions on the error rate and proposed the beta distribution as a reasonable continuous distribution to use. Our model uses the beta distribution but, as will be seen below, we also propose a discrete prior distribution on the parameters of the beta distribution in order to produce greater generality.

We view our model as an extension and formalization of the previously developed models discussed above. Our formalized model provides an opportunity to implement the Bayesian approach in compliance testing by systematically developing tables analogous to those currently used for compliance testing.

As mentioned above, we select the beta distribution,

$$f'(p|r;n') = \frac{\Gamma(n')}{\Gamma(r')\Gamma(n'-r')} p^{r'-1} (1-p)^{n'-r'-1}; 0 < p < 1$$
$$n' > r' > 0$$

as a prior distribution on p , the population error rate. The beta distribution provides a rich family of possible distributions for p which can reflect a wide variety of auditor beliefs. Felix and Grimlund [1977] also selected the beta distribution, and we agree with them that "an auditor can adequately express

his prior judgmental uncertainty for the population error rate (p) using the beta distribution." Corless [1972] used the beta distribution as an instrument in his experiments on the elicitation of a prior distribution on p, and we have already referenced Francisco[1972] and his use of the beta distribution.

In the most common case, where an auditor believes the error rate is very close to zero, it is simple to alter the values of r' and n' to obtain a distribution for p with most of the probability mass between, say, 0 and 5 percent error rates. As an example of this, consider the three prior distributions used by Kraft [1968] given in Table 1. These distributions were chosen by

Table 1

Kraft Prior Probability Distributions on Population Error Rate

Error Rate p*	Cumulative Distributions		
	[Prob (p ≤ p*)]		
	1	2	3
.001	.60	.25	.25
.01	.90	.75	.90
.02	.95	.90	.95
.03	.98	.95	.98
.04	.99	.98	.99
.05	1.00	1.00	1.00

Kraft to reflect typical auditing situations where the population error rate is expected to be low. Kraft suggested that if an auditor could choose one of the three distributions, then predetermined tables could be developed to give the appropriate sample size for whichever prior distribution is selected.³ Table 2 compares Kraft's three prior distributions with three specific beta distributions. It should be apparent from Table 2 that a choice

³Required sample size also depends on the maximum tolerable error rate (upper precision limit) and the reliability level.

of values for the parameters (r', n') of the beta distribution provides a wide range of possible distributions for p . The flexibility of the beta distribution makes it an appealing form for an auditor's prior distribution on the population error rate.

Table 2

Comparison of Kraft Distributions with
Three Different Specific Beta Distributions

Error Rate p^*	Prior Probability Distributions on Population Error Rates					
	Cumulative Distributions					
	[Prob ($p \leq p^*$)]					
	Kraft 1	Beta $r'=.2$ $n'=56.2$	Kraft 2	Beta $r'=.6$ $n'=84.6$	Kraft 3	Beta $r'=.84$ $n'=195.84$
.001	.60	.606	.25	.245	.25	.246
.01	.90	.891	.75	.759	.90	.893
.02	.95	.956	.90	.914	.95	.987
.03	.98	.981	.95	.968	.98	.998
.04	.99	.991	.98	.988	.99	.9998
.05	1.00	.996	1.00	.995	1.00	.99997

Continuing with the example of three prior distributions, suppose an auditor was indifferent with respect to a choice of one of the three distributions. In that case the parameters (r', n') could be considered as joint random variables with the prior distribution given below.

Prior Distribution on the Parameters of
a Beta Distribution

(r', n')	$f(r', n')$
$(.2, 56.2)$	$1/3$
$(.6, 84.6)$	$1/3$
$(.84, 195.84)$	$1/3$
	<u><u>1.0</u></u>

Our general Bayesian model development will use prior distribution on (r', n') which will provide model generality. For example, if an auditor could be precise enough to limit his prior distribution to one specific beta distribution, then only one set of values for (r', n') would be required. In another situation, if an auditor were less certain about the cumulative prior distribution of p , then wider ranges of values for (r', n') would be necessary. Later we will suggest a form for a specific prior distribution that presents an efficient way of incorporating an auditor's prior belief into a distribution.

Let $P'(r', n')$ ⁴ be a prior distribution on (r', n') . A prior distribution on the population error rate p is a conditional beta distribution,

$$f'(p|r', n') = \frac{\Gamma(n')}{\Gamma(r')\Gamma(n'-r')} p^{r'-1} (1-p)^{n'-r'-1}$$

$$n' > r' > 0$$

$$0 < p < 1.$$

Next, let N be the population size (known) and R the total number of errors in the population (unknown). Then the prior distribution on R is binomial, conditional on p ,

$$P'(R|p) = \binom{N}{R} p^R (1-p)^{N-R}$$

$$\text{for } R = 0, 1, \dots, N.$$

The prior distribution on R , given (r', n') is beta-binomial and can be found by the following,

⁴We adopt the usual notational convention of a single prime to indicate a prior distribution and a double prime to indicate a posterior distribution. Also, we use $P(\cdot)$ to indicate a probability density function for a discrete random variable and $f(\cdot)$ for a continuous random variable.

$$\begin{aligned}
 P'(R|r',n') &= \int_0^1 P'(R|p)f'(p|r',n') dp = \\
 &= \binom{N}{R} \frac{\Gamma(n')}{\Gamma(r')\Gamma(n'-r')} \int_0^1 p^{r'+R-1} (1-p)^{n'-r'+N-R-1} dp \\
 &= \binom{N}{R} \frac{\Gamma(n')\Gamma(r'+R)\Gamma(n'-r'+N-R)}{\Gamma(r')\Gamma(n'-r')\Gamma(n'+N)} ; \\
 &R = 0, 1, \dots, N.
 \end{aligned}$$

If n is the sample size and r the number of errors in the sample, then the distribution of r , given R , is hypergeometric.

$$P(r|R) = \frac{\binom{R}{r} \binom{N-R}{n-r}}{\binom{N}{n}},$$

$$r = 0, 1, \dots, \min(n, R).$$

In that case, the distribution of r , given (r',n') (the likelihood function for our sample data), is beta-binomial and can be derived as follows:

$$\begin{aligned}
 P(r|r',n') &= \sum_{R=0}^N P(r|R)P'(R|r',n') \\
 &= \frac{\Gamma(n')}{\Gamma(r')\Gamma(n'-r')} \sum_{R=0}^N \binom{N}{R} \frac{\binom{R}{r} \binom{N-R}{n-r} \Gamma(r'+R)\Gamma(n'-r'+N-R)}{\Gamma(n'+N)} \\
 &= \binom{n}{r} \frac{\Gamma(n')\Gamma(r'+r)\Gamma(n'-r'+n-r)}{\Gamma(r')\Gamma(n'-r')\Gamma(n'+n)}
 \end{aligned}$$

$$r = 0, 1, \dots, n.$$

Now when a sample of size n is taken and r errors are observed, we can combine the sample information with the distributions derived above to obtain certain posterior distributions of interest. First, the posterior distribution of (r',n') , given r , is derived as follows.

$$P''(r',n'|r) = \frac{P'(r|r',n')P'(r',n')}{\sum_{r',n'} P'(r|r',n')P'(r',n')}.$$

where $\sum_{r',n'}$ is the sum over all points (r',n') in the support of the prior distribution on (r',n') . If we assume a uniform discrete prior distribution for (r',n') , then $P'(r',n')$ is a constant and

$$P''(r',n'|r) = \frac{K(r',n',r,n)}{\sum_{r',n'} K(r',n',r,n)},$$

$$\text{where } K(r',n',r,n) = \frac{\Gamma(n')\Gamma(r'+r)\Gamma(n'-r'+n-r)}{\Gamma(r')\Gamma(n'-r')\Gamma(n'+n)}$$

Next, the posterior distribution of p , given r',n' and the sample information r , is

$$f''(p|r',n',r) = \frac{\Gamma(n'+r)}{\Gamma(r'+r)\Gamma(n'-r'+n-r)} p^{r'+r-1} (1-p)^{n'-r'+n-r-1};$$

$$0 < p < 1$$

$$n'+n > r'+r > 0$$

i.e., a beta distribution with parameters $(r'+r, n'+n)$.

Once a sample is taken and r errors are observed, the posterior distribution of R (the total number of errors in the population) will be on the values $R = r, r+1, \dots, N-n+r$. This leads us to the posterior distribution of R , given p and r :

$$P''(R|p,r) = \binom{N-n}{R-r} p^{R-r} (1-p)^{N-n-R+r}; \quad R = r, r+1, \dots, N-n+r.$$

Then the posterior distribution of R , given r',n' , and r can be derived as follows:

$$\begin{aligned} P''(R|r',n',r) &= \int_0^1 P''(R|p,r) f''(p|r',n',r) dp \\ &= \binom{N-n}{R-r} \frac{\Gamma(n'+n)}{\Gamma(r'+r)\Gamma(n'-r'+n-r)} \int_0^1 p^{R+r'-1} (1-p)^{N+n'-r'-R-1} dp \\ &= \binom{N-n}{R-r} \frac{\Gamma(n'+n)\Gamma(R+r')\Gamma(N+n'-r'-R)}{\Gamma(r'+r)\Gamma(n'-r'+n-r)\Gamma(N+n')} \\ &= \binom{N-n}{R-r} Q(r',n', R, N, r, n) \end{aligned}$$

where $Q(r', n', R, N, r, n) = \frac{\Gamma(n'+n)\Gamma(R+r')\Gamma(N+n'-r'-R)}{\Gamma(r'+r)\Gamma(n'-r'+n-r)\Gamma(N+n')}$,

$$R = r, r+1, \dots, N-n+r.$$

Finally, the posterior distribution of primary interest is that of R , given the sample information r .

$$\begin{aligned} P''(R|r) &= \sum_{r', n'} [P''(R|r', n', r)P''(r', n'|r)] \\ &= \sum_{r', n'} \binom{N-n}{R-r} Q(r', n', R, N, r, n) \frac{K(r', n', r, n)}{\sum_{r', n'} K(r', n', r, n)} \\ &= \frac{\binom{N-n}{R-r}}{\sum_{r', n'} K(r', n', r, n)} \sum_{r', n'} Q(r', n', R, N, r, n) K(r', n', r, n) \end{aligned}$$

where $K(r', n', r, n)$ and $Q(r', n', R, N, r, n)$ are as defined above, and $R = r, r+1, \dots, N-n+r$.

The posterior distribution, $P''(R|r)$, is a discrete distribution with $N-n+1$ points to be evaluated. Evaluation is accomplished by inserting the known values of N , n , and r into $P''(R|r)$ and then, for each possible value of R , summing over the domain of (r', n') . Since in auditing situations most of the probability mass will be on low values of R relative to N , it will be reasonable to truncate the evaluation of $P''(R|r)$ for large values of R in order to provide computational efficiency. This will be illustrated in an example below.

Prior Distribution on (r', n')

Earlier we discussed the flexibility provided by the beta distribution in describing an auditor's prior belief about an uncertain error rate. While it is true that the beta distribution can describe a wide assortment of distributions of p , it is also a fact that auditors never directly observe realizations of p ; rather, they observe $\hat{p} = r/n$, where n is sample size and r is the number of errors observed in a sample. Only if an auditor observed an entire population of N items would a realization of p be observed. Thus we believe that an assessment of a prior distribution on (r', n') must

begin with an assessment of a distribution on \hat{p} or on r , given a specific sample size. That is, we would like to develop a methodology whereby a direct assessment of a distribution on \hat{p} or r would lead to an indirect assessment of a prior distribution on (r', n') . Our previously developed conditional distribution on r , $P(r|r', n')$, provides a starting point for the development of an indirect assessment of a prior distribution on (r', n') .

We let D represent the support of a prior distribution on (r', n') ; that is,

$$D = \{(r', n') : P'(r', n') > 0\}.$$

Then we can define another conditional prior distribution on r ,

$$P(r|D) = \sum_{r', n' \in D} P(r|r', n') P'(r', n').$$

We assume that $P'(r', n')$ is uniform on D and set $w = P'(r', n')$. Therefore,

$$P(r|D) = w \binom{n}{r} \sum_{r', n' \in D} \frac{\Gamma(n') \Gamma(r'+r) \Gamma(n'-r'+n-r)}{\Gamma(r') \Gamma(n'-r') \Gamma(n'+n)},$$

where $0 < w \leq 1$ and w is a constant if we assume that $P'(r', n')$ is a uniform distribution on D .

Recall that earlier we determined three pairs of values for r', n' based on the three alternative prior distributions suggested by Kraft. Those three points can be viewed as a support with equal probability ($1/3 = w$) on each point. This provides an example of $P'(r', n')$, which then could be used to develop the conditional distribution of r , $P(r|D)$

It should be apparent that diverse sets of points (r', n') may be contained in D and can lead to diverse distributions of r . A proper selection of points in D has the potential to provide a distribution of r (or \hat{p}) that appears reasonable to an auditor. This leads to our suggested approach which begins

with an indirect assessment of a discrete uniform prior distribution on (r', n') and culminates with a determination of the sample size required to meet auditing and statistical objectives.

1. We construct alternative discrete sets of points (r', n') , say D_s , $s = 1, \dots, S$. We construct each D_s in such a way that the resulting S distributions, $P(r|D_s)$, represent S alternative distributions of r that are reasonable alternatives for auditors to consider.
2. We assume a sample size, of 100 ($n = 100$) and calculate the S distributions, $P(r|D_s)$. From these distributions we produce S graphs.
3. Next, in a specific compliance testing situation where an attribute is to be statistically tested, we present the S graphs to the auditor in charge and ask the following question. If a sample of $n = 100$ items were taken from the population of interest, which one of the S graphs best reflects your belief about the number of errors you would see?
4. If graph s is selected, we know that the graph was generated by support D_s .
5. The prior distribution support D_s can then be used in making calculations with a posterior distribution, $P''(R|r)$. Given a specification of statistical objectives for our audit test, we can calculate the required sample size.

An important question not addressed in this paper concerns the determination of what are "reasonable" distributions for r or \hat{p} . It is possible to generate almost any distribution that would seem to be of interest to auditors, i.e., low expected error rates with most of the probability mass near zero. Further research would be required to determine an array of prior distributions

on \hat{p} rich enough to capture a broad range of auditor expectations.⁵ Indeed, it may ultimately be appropriate to develop prior distribution arrays that are specialized according to client or industry peculiarities.

A Suggested Strategy for Selection of $P'(r',n')$

We will now discuss a strategy for selecting sets of points (r',n') that could serve as the basis for development of reasonable prior distributions. Consider the points (r',n') are depicted in Figure 2. The shaded area consists

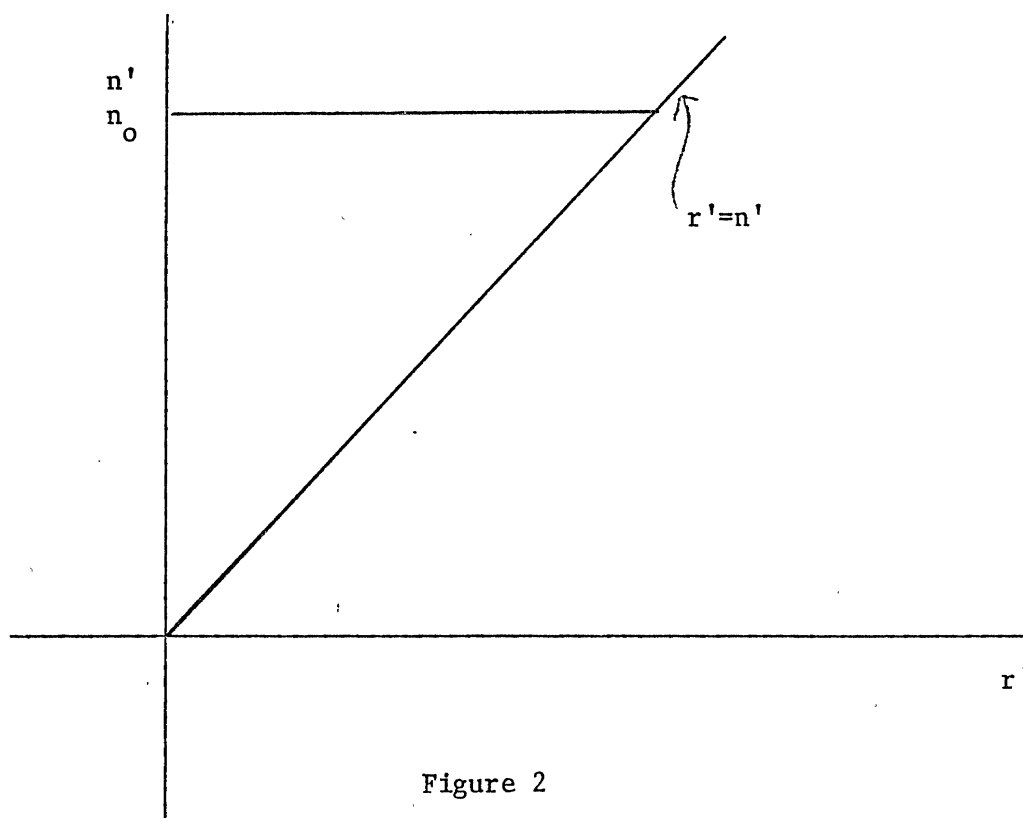


Figure 2

of points such that r' and $n' > 0$, $r' < n'$, and $n' \leq n_0 < \infty$. It is not likely that points near the $r' = n'$ line would be included in a prior distribution support D_S , since $E(p|r',n') = r'/n' = 1.0$ for points on the line. In fact, we

⁵Francisco [1972, p. 62] suggested using sketches of different beta distributions for assessment by auditors. Also, in Chapter VI, he discussed the problems associated with the assessment process and the many factors that can influence an auditor's beliefs.

can also make the following observations about the minimum and maximum values of $E(p|r',n')$ and $\text{Var}(p|r',n')$ for the points included in the triangle depicted in Figure 2:

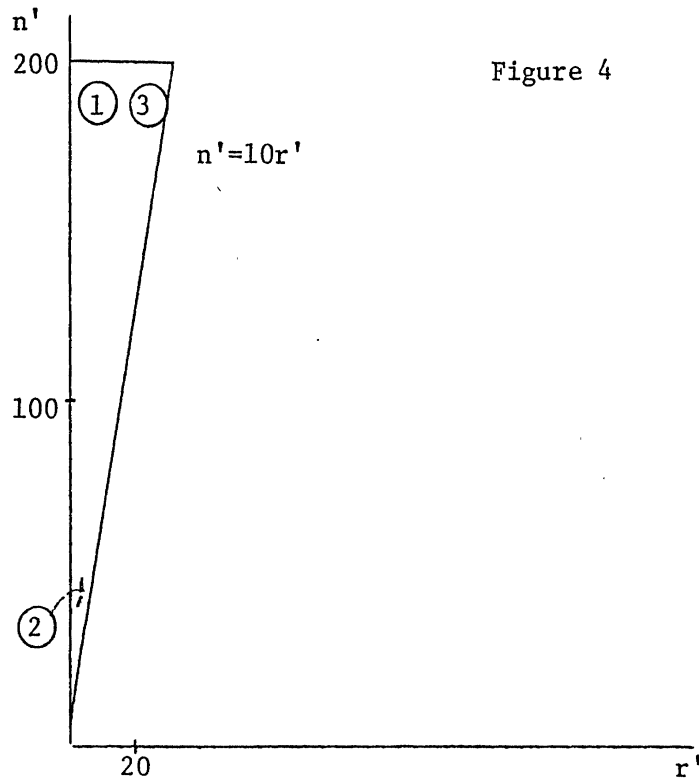
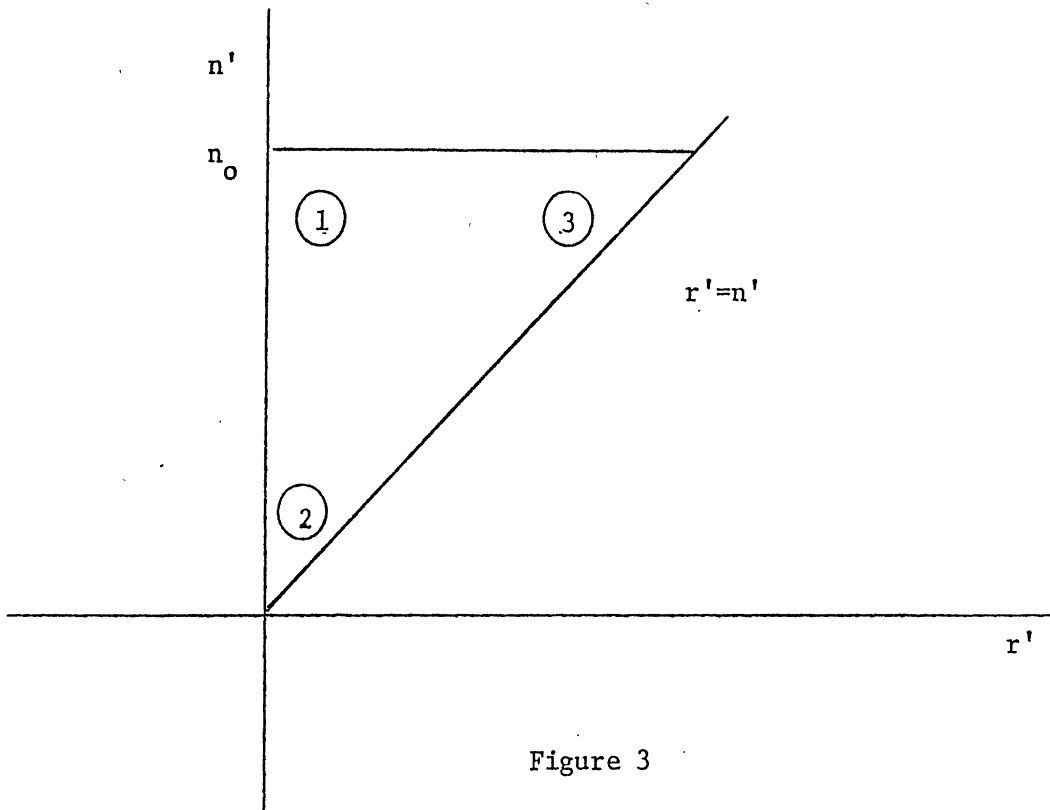
	<u>Points (r',n')</u>	
	<u>Minimum</u>	<u>Maximum</u>
$E(p r',n')$	$(0, n_0)^*$	$(r', n'=r')^{**}$
$\text{Var}(p r',n')$	$(0, n_0)^*$	$(0, 0)^*$

*actually the positive neighborhood of $r' = 0$ and/or $n'=0$ since we cannot have $r=0$ or $n'=0$.

**Maximum for $E(p|r',n')$ is for any point on line $r'=n'$.

Given the relationships indicated above, we can make some observations about subsets of points in the triangle. Consider the triangle in Figure 3, in which the three numbered circles depict subsets of points. The points in circle 1 would generate beta distributions with relatively low values for $E(p|r',n')$ and $\text{Var}(p|r',n')$. The points in circle 2 would generate beta distributions with larger values for both $E(p|r',n')$ and $\text{Var}(p|r',n')$ than for points in circle 1. The points in circle 3 would generate beta distributions with values of $E(p|r',n')$ similar to those for points in circle 2 but greater than values of $E(p|r',n')$ for points in circle 1. $\text{Var}(p|r',n')$ for points in circle 3 would be less than for points in circle 2 but greater than for points in circle 1.

The shape of the distribution $P(r|D_s)$ would be heavily influenced by selection of a support, D_s , in the different areas described above. To illustrate the way in which a choice of D_s can influence $P(r|D_s)$, we further restrict the feasible region for selection of D_s as indicated in Figure 4. Since $n' = 10r'$, the maximum value of $E(p|r',n') = .1$ for points on the line.



The minimum value of $E(p|r',n') = \frac{r'}{200}$ for points r' as close to zero as is reasonable. The three circled numbers represent diverse sets of points (r',n') that make up three alternative prior distribution supports, i.e., $S = 3$.

For our illustration, we will arbitrarily select five points for each D_s as follows:

D_1	D_2	D_3
(r',n')	(r',n')	(r',n')
(.1, 200)	(.1, 1)	(20, 200)
(.2, 200)	(.2, 1)	(18, 200)
(.2, 180)	(.2, 21)	(18, 180)
(.1, 180)	(.1, 21)	(16, 180)
(.15, 190)	(.15, 11)	(18, 190)

Each of these clusters of points corresponds to the circle with the same number in Figure 4. Table 3 lists the probability density (PDF) and cumulative (CDF) distributions of r for a sample size of $n=100$ given D_1 , D_2 , D_3 , and, for comparison purposes, the three-point Kraft distribution discussed earlier. For calculational ease we truncated calculations when the $CDF \geq .99$.

In Table 3 we can see that the diversity of values of $E(p|r',n')$ and $Var(p|r',n')$ across D_1 , D_2 , and D_3 is manifested in the distributions of r . With D_1 we have a "tight" distribution of r with the mode at zero, while with D_2 we still have the mode at zero but the variability is much greater (99th percentile at $r = 91$). D_3 produces a distribution for r with a mode at $r = 9$ ($\hat{p} = .09$) and with more or less variability, respectively, than with D_1 or D_2 . $E(r|D_s)$ and $Var(r|D_s)$ are given below.

	$E(r D_s)$	$Var(r D_s)$
D_1	.079	.121
D_2	6.558	314.133
D_3	9.473	13.250
Kraft	.501	1.075

Table 3

Prior Probability Density (PDF) and Cumulative (CDF)
Distributions on r , Given $n = 100$ and
Four Different Prior Supports on (r', n')

r	Prior Supports							
	Kraft		D_1		D_2		D_3	
	PDF	CDF	PDF	CDF	PDF	CDF	PDF	CDF
0	.715	.715	.938	.938	.634	.634	.0004	.0004
1	.170	.885	.049	.987	.082	.716	.003	.0034
2	.064	.949	.010	.997	.043	.759	.009	.012
3	.027	.976			.028	.787	.021	.033
4	.012	.988			.020	.807	.039	.072
5	.006	.994			.015	.822	.061	.133
6					.012	.834	.082	.215
7					.010	.844	.100	.315
8					.008	.852	.110	.425
9					.007	.859	.111	.536
10					.006	.865	.104	.640
11					.005	.870	.091	.731
12					.005	.875	.097	.806
13					.004	.879	.059	.865
14					.004	.883	.045	.910
15					.004	.887	.032	.942
16					.003	.890	.022	.964
17					.003	.893	.015	.979
18					.003	.896	.009	.988
19					.003	.899	.006	.994
20					.002	.901		
⋮					⋮	⋮		
91					.001	.990		

The Kraft-induced distribution is most like that induced by D_1 , although it still is not as concentrated around $r = 0$.

Figure 5 presents graphical representations of the data contained in Table 3. We suggest that these graphs could be presented to an auditor for selection of the graph that best represents the auditor's prior beliefs about r or \hat{p} . The selection of a graph will imply a support for (r', n') ; we can then proceed to plan the sample size requirement for our statistical test.

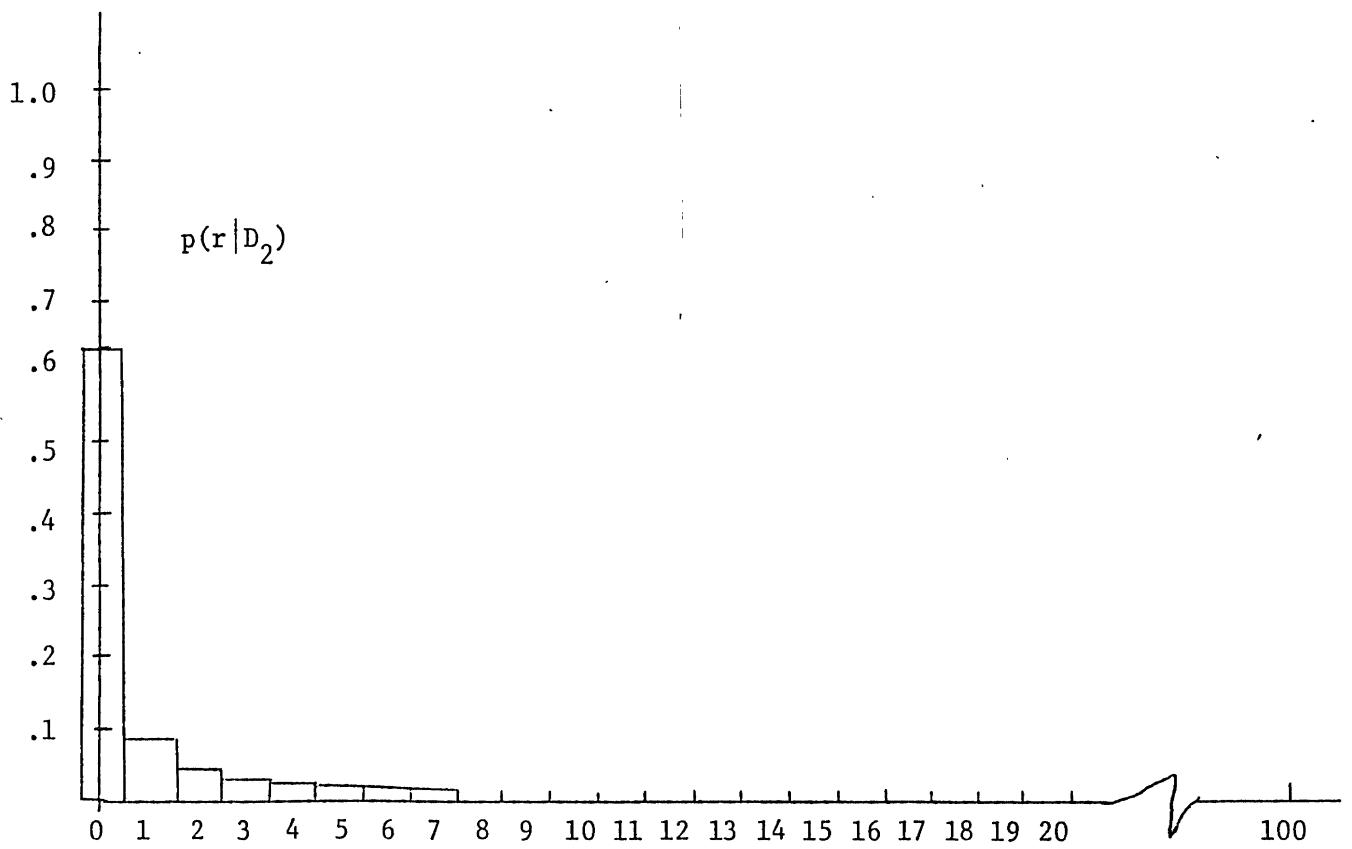
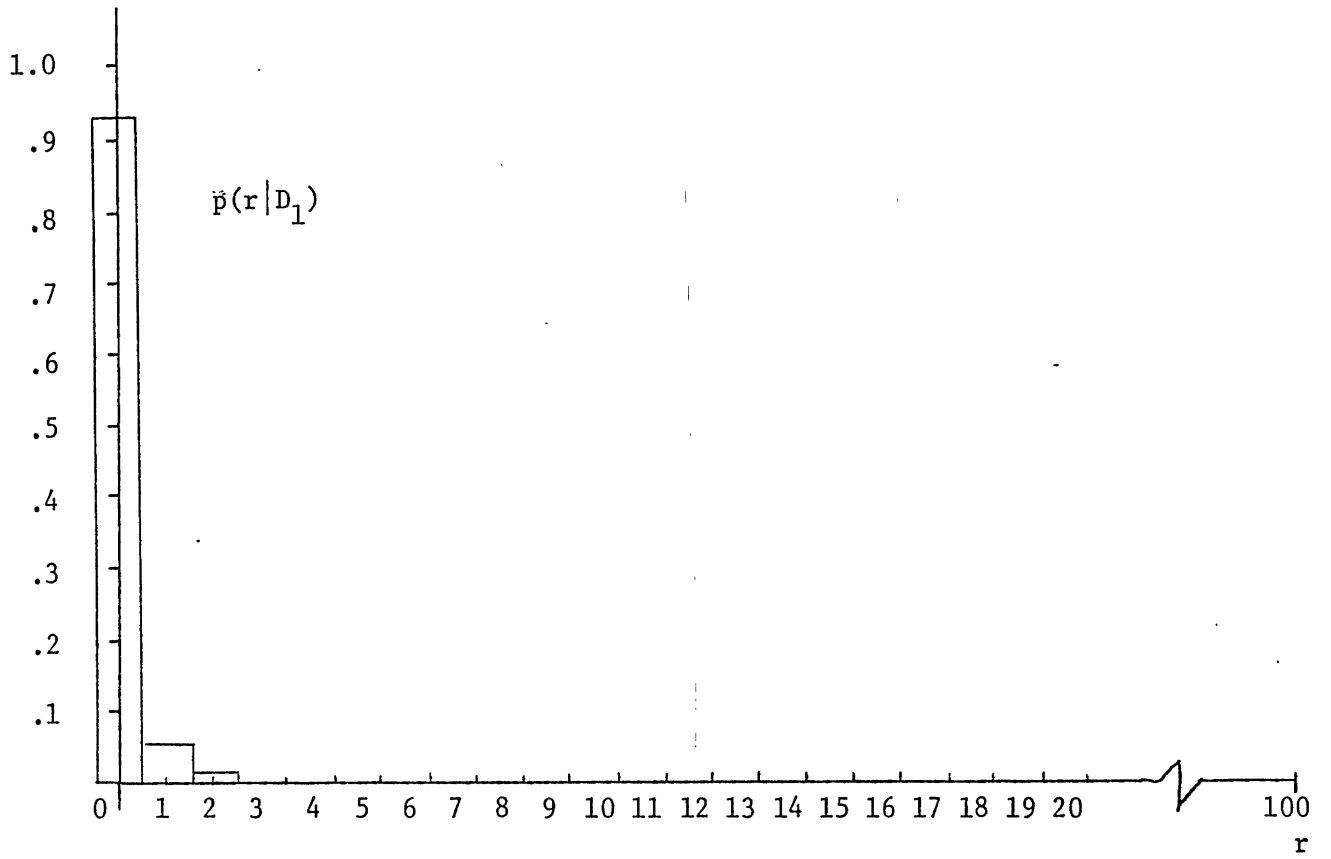


Figure 5

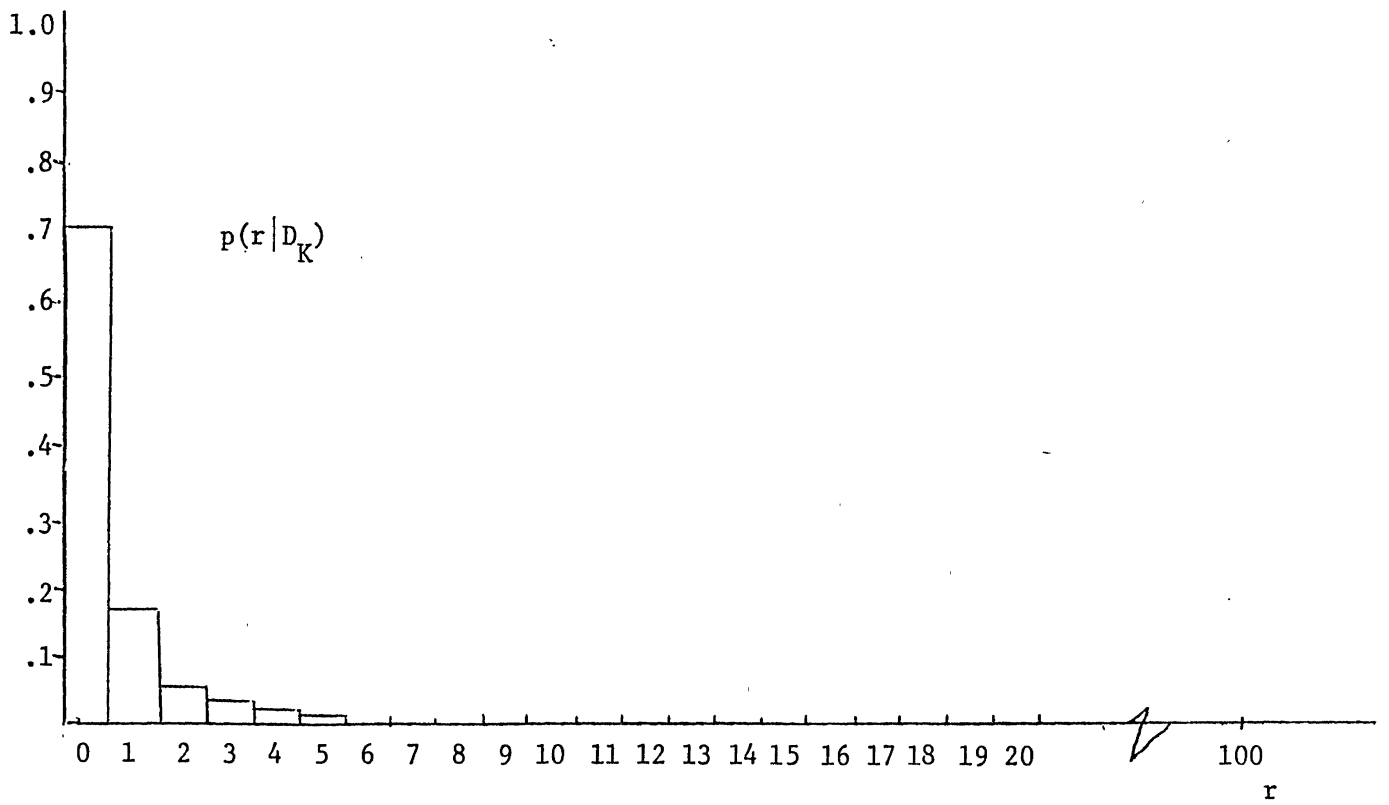
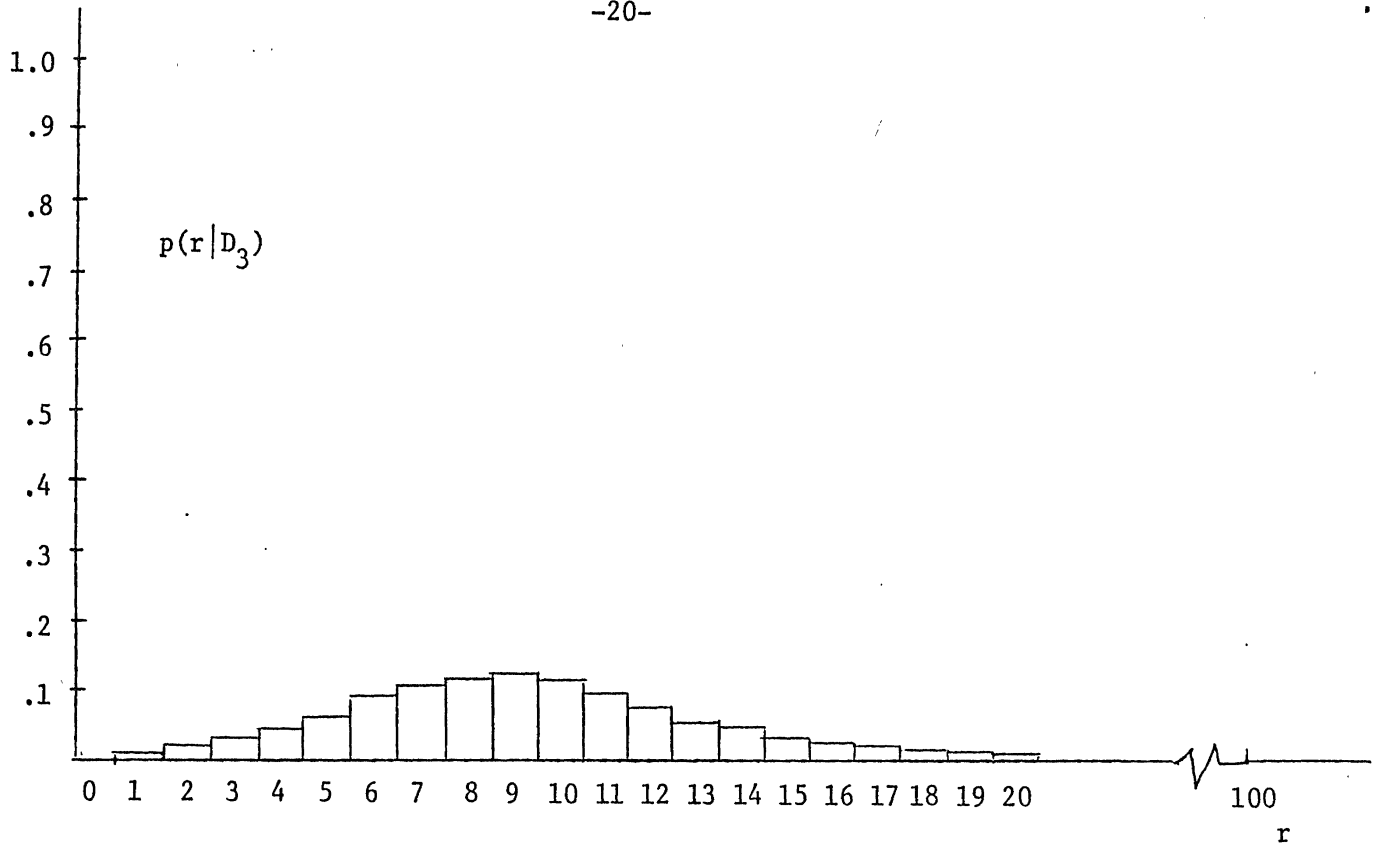


Figure 5, continued

Posterior Distribution and Audit Planning

By specifying values of N, n, r , and a set D_g , we can calculate the posterior distribution, $P''(R|r)$, for values of $R = r, r+1, \dots, N-n+r$. We assume N (population size) will be known. Determination of n will depend on the auditing and statistical objectives of a specific audit plan.

These objectives are characterized in terms of how precise and how confident we must be in a final conclusion about an internal control procedure. We let R^* be the maximum number of errors that can be tolerated and REL^* be the required minimum probability in the posterior distribution that $R \leq R^*$, i.e., $P''(R \leq R^*|r) \geq REL^*$. We also let r^* be the maximum number of errors that we are willing to see in our sample while still accepting that $P''(R \leq R^*|r^*) \geq REL^*$. Thus r^* is an acceptance criterion, and, in the low error rate populations often encountered by auditors, we expect that r^* will be set equal to zero. Also, selecting $r^* = 0$ will minimize the required sample size.

Now, with specifications of N, R^*, r^* , and D_g , we can proceed to find a value for n such that,

$$P''(R \leq R^*|r^*) \geq REL^*, \text{ or}$$

$$\sum_{R=r^*}^{R^*} P''(R|r^*) \geq REL^*.$$

The form of $P''(R|r)$ does not allow a direct solution for n and therefore n must be determined by an iterative procedure. Although the iterative procedure can be costly in terms of computer time, once completed the results can be tabled for use in subsequent audit situations. That is, if we can agree on a set of reasonable alternative prior distribution supports (D_g) on (r', n') , then tables of required sample sizes can be developed for each prior distribution, given different specifications of N, r^*, R^* , and REL^* . This is illustrated

in an example below using the four alternative prior distributions contained in Table 3.

Continuing our numerical example, we let $N = 1,000$, $REL^* = .99$, $r^* = 0$, and $R^* = 10$. Thus, our conclusion of acceptance requires that $P''(R \leq 10 | r^* = 0) \geq .99$, or equivalently, $P''(R/N \leq .01 | r^* = 0) \geq .99$. In Table 4 we present the minimum required sample sizes that will allow us to reach our acceptance conclusion with $r^* = 0$, given each of the alternative prior distributions in Table 3.

Table 4
Required Sample Sizes

$N = 1,000$
 $R^* = 10$
 $r^* = 0$
 $REL^* = .99$

Prior Distribution Domain	Required Sample Size
D_1	10
D_2	154
D_3	759
D_K (Kraft)	186

It is clear from Table 4 that the less assurance reflected by the prior distribution that the population error rate is less than .01, the more sample evidence is required to allow the acceptance conclusion to be made. The diversity of the three prior distributions selected as examples is reinforced by the relative sizes of the required samples.

It is useful to compare the data in Table 4 with the calculation of the required sample size in a classical attribute sampling plan. With the same auditing objectives as before, the hypergeometric distribution is required and we must solve the following equation for n .

$$\frac{\binom{10}{0} \binom{990}{n-0}}{\binom{1,000}{n}} = .01;$$

$$\frac{(1,000-n) \dots (991-n)}{(1,000) \dots (991)} = .01.$$

An iterative procedure must also be used in this case, and we find $n = 367$. Thus, in a classical sampling plan, if no errors are observed in a sample of size $n = 367$, the statistical conclusion is that we are 99 percent confident that the true population error rate does not exceed .01, or equivalently, that there are no more than 10 errors in the population of $N = 1,000$.⁶ The calculation of $n = 367$ is the same, regardless of how strongly an auditor feels about the likelihood of seeing errors. In comparing this result with the sample sizes required by our Bayesian model (Table 4), we see that it ($n = 367$) exceeds all but one of the same sizes contained in Table 4. Recall that for prior distribution support D_3 , the distribution of r , given $n = 100$, was unimodal with the mode at $r = 9$. In that case an auditor has a strong prior belief that errors will be encountered, and it would require a large amount of sample evidence ($n = 759$) with no errors observed ($r^* = 0$) for him to conclude there are no more than ten errors in the entire population ($N = 1,000$).

Summary and Conclusions

In this paper we have proposed a Bayesian statistical model for use in the testing and evaluation of an attribute in an internal accounting control system. Our model is a generalization and extension of approaches suggested

⁶The literal interpretation of these sample results in the classical setting is that $P(r > 0 | N=1,000, R=10) = .99$, i.e., the chance of getting sample results worse than were obtained ($r=0$) is .99, given there are actually 10 errors in the population. It cannot be interpreted as the probability that $R \leq 10$, as is the case in the Bayesian model.

by Kraft [1968], Tracy [1969], and Francisco [1972]. The primary objective of the Bayesian model is to allow the experience and judgment of an auditor to become an explicit part of the model and therefore to have a direct effect on the planning and execution of an attribute sampling plan. We have used a numerical example to show how our model allows the required sample size to be directly related to an auditor's prior knowledge. Furthermore, our example shows that for the low error rate populations often encountered by auditors the sample size can be considerably less than is required for classical attribute sampling plans. The potential exists for the development of tables of required sample sizes for our Bayesian model, just as tables have been developed for classical attribute sampling plans.

An area for further research is the assessment of prior distributions of (r', n') . We have suggested a direction for developing prior distributions by illustrating how selectively restricting (r', n') to subsets of points in (r', n') space can reflect a diverse range of auditors' beliefs. We also suggested that an auditor might be presented with visual representations, such as graphs, of different prior distributions of r and/or \hat{p} which are related to specific subsets of points (r', n') . This suggestion was based on the fact that much of an auditor's experience consists of observations of \hat{p} . However, more research is obviously required to determine a most efficient and consistent way of assessing a prior distribution on the (r', n') space.

Another area of research is to extend the model proposed in this paper to include the case where an attribute of interest exists at more than one site. In many auditing situations a client has multiple sites, geographically separated, but uses the same internal accounting control procedures for all sites. An interesting statistical question is whether a specific attribute which exists at multiple sites can be statistically evaluated by taking samples at fewer than all sites. We are currently examining this problem.

References

- Andrews, R. W. and J. T. Godfrey, "Testing Compliance at Multiple Sites: A Sequential Probability Ratio Model," Working Paper, The University of Michigan, 1978.
- Corless, J. C., "Assessing Prior Distributions for Applying Bayesian Statistics in Auditing," The Accounting Review, July 1972, pp. 556-66.
- Cushing, B. E., "A Mathematical Approach to the Analysis and Design of Internal Control Systems," The Accounting Review, January 1974, pp. 24-32.
- Felix, W. L. and R. A. Grimlund, "A Sampling Model for Audit Tests of Composite Accounts," Journal of Accounting Research, Spring 1977, pp. 23-41.
- Kinney, W. R., Jr., "A Decision Theory Approach to the Sampling Problem in Auditing," Journal of Accounting Research, Spring 1975, pp. 117-32.
- _____, "Decision Theory Aspects of Internal Control System Design/Compliance and Substantive Tests," Journal of Accounting Research.
- Knoblett, J. A., "The Applicability of Bayesian Statistics in Auditing," Decision Sciences, July-October 1970, pp. 423-40.
- Kraft, W. H., Jr., "Statistical Sampling for Auditors: A New Look," Journal of Accountancy, August 1968, pp. 49-56.
- Francisco, A. K., "The Application of Bayesian Statistics to Auditing: Discrete Versus Continuous Prior Distributions," Ph.D. Dissertation, Michigan State University, 1972.
- Scott, W. R., "A Bayesian Approach to Asset Valuation and Audit Size," Journal of Accounting Research, Autumn 1973, pp. 304-330.
- Sorensen, J. E., "Bayesian Analysis in Auditing," The Accounting Review, July 1969, pp. 555-61.
- Tracy, J. A., "Bayesian Statistical Methods in Auditing," The Accounting Review, January 1969, pp. 90-98.
- _____, "Bayesian Statistical Confidence Intervals for Auditors," Journal of Accountancy, July 1969, p. 41.

