A HYBRID PROBLEM ON THE EXPONENTIAL FAMILY[1]

By Arthur Cohen

Rutgers—The State University

1. Introduction and summary. It is frequently of interest to an investigator to decide whether an unknown parameter lies in some interval and also to obtain a point estimate of the parameter. There are many instances in practice where one does this by first testing a hypothesis and then estimating. Sometimes the investigator will even allow the result of the hypothesis test to effect his method of estimation. Theoretical work on the procedure, where one follows a test of hypothesis by estimation, has been generally reviewed by Kitagawa [10]. Included in Kitagawa's review are references and discussions of other approaches to what we will call a hybrid problem. That is, we describe a hybrid problem as follows: Suppose we take n observations on a random variable whose distribution is known to belong to a one-parameter family. Then, on the basis of these observations we decide whether the parameter (or some function of it) lies in a given interval of the parameter space (or the range space of the function of the parameter). Secondly, we estimate the parameter (or some function of it).

In the next section we will explicitly state a decision theory formulation for a hybrid problem. Following the statement of the problem, we will give examples of situations where a hybrid problem is defined and for which the given formulation is appropriate.

The formulation will be given for a hybrid problem when a single observation is made on a random variable whose distribution is of the exponential type, and when the decisions are made with regard to the expected value of the random variable. We decide whether the expected value, which is a function of the underlying parameter of the exponential distribution, lies in some given interval and we also estimate the expected value. For a large subclass of the exponential family, including the normal distribution with unknown mean, complete classes of procedures are found for the case when the given interval consists of a single point. These complete classes consist of procedures called interval-monotone procedures defined as follows: If the observation falls in a particular interval of the sample space, then decide that the expected value of the random variable lies in the given interval consisting of a single point and estimate it to be precisely that point. If the observation lies outside that particular interval of the sample space, decide that the expected value lies outside the given interval and estimate it by some monotone analytic function of the observation.

Another result for the given formulation is concerned with the case when the distribution is normal with unknown mean and the given fixed interval is any symmetric interval about the origin. Here, a class of admissible procedures is

Received 30 November 1964.

[1] This paper is part of the thesis submitted in partial fulfillment of the requirements for the Ph.D. degree, Columbia University, September, 1963.

found, where each procedure is a symmetric interval-monotone procedure, the symmetry pertaining both to the interval and the estimate. Furthermore, included in this class is the following intuitively appealing procedure: If the observation falls in some particular symmetric interval about the origin, then decide that the parameter lies in the given fixed interval of the parameter space and estimate the parameter by zero; while if the observation falls outside the particular interval of the sample space, then decide that the parameter lies outside the given interval and estimate it by the value of the observation. It is interesting to note that the estimate resulting from the above procedure was proved to be an inadmissible estimate for the squared error loss function. (See [5]).

The analogue of the admissibility result for the normal case is next found for the cases where the probability distribution of the random variable is binomial, Poisson and gamma, and the left end point of the given interval of the parameter space is zero.

In the next section then, we state the problem and give examples. In Section 3 we give complete class theorems. The admissibility results for the normal case and binomial, Poisson, and gamma cases follow in Sections 4 and 5, respectively.

2. Statement of the problem. Let y be a single observation on a random variable whose density is of the exponential type $dP_\theta(y) = \beta(\theta)e^{y\theta}\, d\mu(y)$, where μ is a σ-finite measure defined on the real line, whose spectrum is denoted by Y. To avoid trivialities we assume Y is non-degenerate. The parameter space is

$$\Omega = \{\theta: \textstyle\int_Y e^{y\theta}\, d\mu(y) < \infty\},$$

and it is known to be an interval of the real line ([6], p. 67), while $\beta(\theta)^{-1} = \int_Y e^{y\theta}\, d\mu(y)$, is positive and analytic at each interior point of Ω. Let $h(\theta) = E_\theta(y)$, and denote the range space of $h(\theta)$ by R_h. Also let $I = [\theta_1, \theta_2]$ be an arbitrary but fixed interval in Ω and let $I_h = [h(\theta_1), h(\theta_2)]$. From the well known relations ([6], p. 67),

$$h(\theta) = -\beta'(\theta)/\beta(\theta), \text{ and } h'(\theta) = \beta'(\theta)^2 - \beta(\theta)\beta''(\theta)/\beta^2(\theta) = \sigma_y^2 > 0,$$

$h(\theta)$ is seen to be a continuous increasing function of θ and therefore $\theta \in I$ if and only if $h(\theta) \in I_h$.

Now on the basis of the single observation, (note that there is no loss of generality if we restrict ourselves to a single observation), we would like to

$$\text{decide whether } h(\theta) \in I_h, \tag{2.1}$$

$$\text{estimate } h(\theta). \tag{2.2}$$

We require that if it is decided that $h(\theta) \in I_h$, then $h(\theta)$ must be estimated by some constant h^*, where h^* is some preassigned point in I_h. We choose a loss function for this problem to satisfy the following:

(1) If it is decided that $h(\theta) \in I_h$ and $h(\theta)$ is estimated by h^*, then the loss is zero when $\theta \in I$; and $(h^* - h(\theta))^2$ when $\theta \notin I$.

(2) If it is decided that $h(\theta) \not\in I_h$ and $h(\theta)$ is estimated by some function of y, say $a(y)$, then the loss is $(a - h(\theta))^2 + M$, for M a positive constant, when $\theta \varepsilon I$; and $(a - h(\theta))^2$ when $\theta \not\in I$.
Before justifying the formulation for this problem we find it convenient to make some definitions and to present the loss function in tabular form. Therefore, suppose we let

$$\chi(h(\theta)) = 0 \quad \text{if} \quad h(\theta) \, \varepsilon \, I_h$$
$$= 1 \quad \text{if} \quad h(\theta) \not\in I_h .$$

Also, let us define a function $b(y)$ such that

$$b(y) = 0, \quad \text{if after observing} \quad y \quad \text{we decide} \quad h(\theta) \, \varepsilon \, I_h ,$$
$$b(y) = 1, \quad \text{if after observing} \quad y \quad \text{we decide} \quad h(\theta) \not\in I_h .$$

If we let $a(y)$ denote a nonrandomized estimate of $h(\theta)$, and set $t(y) = (a(y), b(y))$, the loss function $W(\theta, t)$ for this problem is presented in Table 2.1.

Note that the requirement that $h(\theta)$ be estimated by h^* whenever it is decided that $\theta \, \varepsilon \, I_h$, and the table of losses, Table 2.1, could both be derived if we stated the problem (that is, to be (2.1) and (2.2)) and defined the loss function to be

$$(2.3) \quad \begin{aligned} W(\theta, t) = {} & (a - bh(\theta) - (1 - b)h^*)^2 \\ & + \max [M(b - \chi(h(\theta)), (\chi(h(\theta)) - b)(h^* - h(\theta))^2]. \end{aligned}$$

The table of losses derived from (2.3) is Table 2.2.
Now it is obvious that any procedure where $b(y) = 0$ and $a(y) \neq h^*$ can be improved by choosing $a(y) = h^*$. Therefore, the problem can be reduced to one

TABLE 2.1
Loss function in tabular form

$\chi(h(\theta))$ \ $b(y)$	0	1
0	0	$[a - h(\theta)]^2 + M$
1	$[h^* - h(\theta)]^2$	$[a - h(\theta)]^2$

TABLE 2.2
Loss function derived from (2.3) *in tabular form*

$\chi(h(\theta))$ \ $b(y)$	0	1
0	$[a - h^*]^2$	$[a - h(\theta)]^2 + M$
1	$[a - h^*]^2 + [h^* - h(\theta)]^2$	$[a - h(\theta)]^2$

in which we need only consider procedures with $a(y) = h^*$ whenever it is decided that $\theta \varepsilon I_h$, and where the table of losses is Table 2.1.

We now turn to the motivation and justification of the hybrid problem stated above. We will cite three examples for which the formulation is appropriate. Then we will briefly discuss the idea of using h^*, and some variations of the loss function.

Let us suppose then that an investigator is interested in some unknown parameter. If he decides that the parameter lies in some interval (which he has chosen before making any observations) it might be easier, more convenient, and no less beneficial for him to estimate the parameter by some constant in that interval. If he decides that the parameter lies outside that interval, then he wants a good estimate of the parameter. Furthermore, if he decides that the parameter lies outside the interval, when in fact the parameter lies inside the interval, he incurs a loss that consists of his error due to estimation plus an additional penalty factor. We cite three examples.

EXAMPLE 1. Suppose a pharamcologist is investigating a drug that might effect the cholesterol count of an individual. From past experience the average cholesterol count of an individual belonging to a certain group is known. The pharmacologist wishes to decide if the drug changes the average cholesterol count and he would also like to measure the amount of change. He feels that if he decides that the drug does have an effect and he estimates the amount of change in the cholesterol count, when in fact the drug does not have any effect, then he should incur a penalty factor in addition to his loss due to estimation. The penalty factor is due to the inconvenience of notifying physicians, preparing literature, alerting the public, and perhaps injuring the reputation of the pharmacologist's company. If the pharmacologist correctly decides the drug has an effect, then his loss is due only to estimation. If he decides correctly that the drug has no effect, he loses nothing, whereas if he decides that the drug has no effect, when in fact it does, then he feels that the loss due to poor estimation is severe enough.

EXAMPLE 2. Suppose a jobber is interested in the proportion of good items in a lot. If the proportion of good items is less than some number p_0, say, then the proportion is too small to be of practical significance to him, and as far as he is concerned, the proportion might just as well be zero. On the other hand, if the proportion exceeds p_0, then the jobber would like to purchase the lot and he would be willing to pay according to the proportion of good items. Hence, he would be interested in a good estimate of the proportion. The jobber feels that if he decides that the proportion exceeds p_0, when in fact it does not, his loss should consist of an error due to estimation and an additional penalty factor. The penalty factor would account for the inconvenience of unnecessarily informing storekeepers, and perhaps the adverse effect on the jobber's relation with these storekeepers. As in the previous example, the jobber feels that if he correctly decides that the proportion of good items exceeds p_0, then his loss is due only to estimation. If he decides correctly that the proportion is less than p_0, then he loses nothing, whereas if he decides incorrectly that the proportion is less than p_0, then his loss due to underestimating the proportion is severe enough.

EXAMPLE 3. Suppose we consider a polynomial regression model. That is, we assume that y_i, $i = 1, 2, \cdots, n$ are observations which are independent and normally distributed, such that

$$E(y_i) = \sum_{j=1}^{p} \beta_j \varphi_j(x_i),$$

where the β_j are parameters, the x_i are fixed, and the φ_j are orthogonal polynomials of degree j. A problem of interest then is to simultaneously decide on the degree of the polynomial and to predict a future observation at a given value of the fixed independent variable. A loss function for such a problem can take into account both errors in predicting and errors in deciding on the degree of the polynomial. In fact, it is reasonable to assign a penalty factor whenever one overestimates the degree of a polynomial regression since "it is an advantage to represent the data by a polynomial of low degree because the curve is smoother, the presumed 'explanation' is simpler, and the function is more economical." ([1], p. 256). Now one can define, in some sense, the degree of a polynomial by some function of the parameters $\beta_1, \beta_2, \cdots, \beta_p$, given in the regression model. Such a function, say, would equal j, if j were the largest index for which β_j was "large". In the author's dissertation [4] such a function is defined when formulating a problem of simultaneously deciding on the degree of a polynomial and predicting. The idea of a penalty factor is incorporated in this formulation. When $p = 1$ and $\varphi_1(x) \equiv 1$, the loss function in tabular form for the polynomial problem reduces to Table 2.1 with $h(\theta) = \theta$, $I = I_h = (-\theta_0, \theta_0)$, and $h^* = 0$. Here M plays the role of the penalty factor, while $\chi(\theta)$ is the function, defining in some sense, the degree of the polynomial. We recognize that it would not be difficult to alter Example 3 so that it defines the interesting problem of including or deleting the pth variate of a pth order regression to be used for prediction.

The idea of using h^* is explained by the requirement that whenever it is decided that $h(\theta) \varepsilon I_h$, then $h(\theta)$ must be estimated by some constant in the interval I_h. The only restriction then on h^* is that $h^* \varepsilon I_h$. Hence, the investigator, in setting up his problem, may choose an $h^* \varepsilon I_h$ that makes the practical aspect of the problem more meaningful or more convenient.

Although the ensuing development of this problem in Sections 3 thru 5 will be based on the loss function given in Table 2.1, we remark that no major changes in methods or results would be required if the following changes in Table 2.1 were made:

(1) In the cell $(b(y) = 0, \chi(h(\theta)) = 1)$, replace $(h^* - h(\theta))^2$ by $(h^* - h(\theta))^2 + K$, for K a positive constant.

(2) In the cell $(b(y) = 0, \chi(h(\theta)) = 0)$, put C, a positive constant or put $(h(\theta))^2$.

Other variations of the losses could be made so that the development would not require major changes. We also remark that if $M = 0$, no major changes will be required. If $M = \infty$, by always choosing $b(y) = 0$ we are assured of a finite, but not necessarily a bounded risk function, whereas if we do not always choose $b(y) = 0$, the risk function is infinite for all $\theta \varepsilon I$.

We make one final remark before turning to the development of the hybrid problem. The problem defined here is concerned with decisions regarding $h(\theta) = E_\theta(y)$. However, many of the results hold if we wish to make decisions regarding θ itself, or perhaps other functions of θ. The problem defined in terms of $h(\theta)$ however, is felt to be a problem of considerable interest.

3. Complete classes. In this section we will start out by noting that the procedures with nonrandomized estimates of $h(\theta)$ form a complete class. Then we go on to characterize the Bayes procedures. After this, we find a complete class of procedures for a large subfamily of the exponential family. This complete class is B^*, where B^* consists of procedures which have nonrandomized and finite estimates of $h(\theta)$ and are also Bayes or the limit of a sequence of Bayes procedures. We give a characterization of those procedures in B^* for the normal case (that is $h(\theta) = \theta$) and for other distributions (that is, for other functions $h(\theta)$). Finally we finish this section by describing a complete class for the case when the interval I consists of a single point. This complete class consists of interval-monotone procedures only.

Before proceeding, we remark that the decision theory terminology and definitions are more or less those of Wald [13] and LeCam [11]. Now we note that the action space, which we denote by T, for this problem, is a subset of Euclidean 2-space; namely the line segment where $b = 1$, and a ranges over R_h, and also the point $(a = h^*, b = 0)$. Hence, the most general type of decision procedure is one which assigns a probability distribution $\psi(t \mid y)$ on T for every y. Considering such a $\psi(t \mid y)$, let $(1 - \varphi(y))$ be the probability which $\psi(t \mid y)$ assigns to the point $(a = h^*, b = 0)$. That is, $(1 - \varphi(y)) = \Pr\{b = 0 \mid y\}$. If $\varphi(y) = 0$, let $\psi_1(t \mid y)$ be the conditional distribution function on R_h, given $b = 1$. We will consider only those $\psi(t \mid y)$ such that $\int_T u(t)\, d\psi(t \mid y)$ is a measurable function of y, where $u(t)$ is any bounded continuous function which is zero outside a compact set. Suppose then that we are given a procedure $\psi(t \mid y)$ such that for every y in some set of positive μ measure, $\varphi(y)$ is positive and $\psi_1(t \mid y)$ is positive on some non-degenerate set in R_h. It is easy then to prove that the procedure $\psi^*(t \mid y)$ is better than $\psi(t \mid y)$, where $\psi^*(t \mid y)$ is defined as follows: let $\varphi^*(y) = \varphi(y)$, and for all those points y where $\varphi(y) \neq 0$, let $\psi_1^*(t \mid y) = 1$ at the point $a^*(y) = \int_{R_h} a\, d\psi_1(t \mid y)$. The proof follows immediately from convexity by consideration of the risk functions which for the procedure $\psi(t \mid y)$ is

$$\rho(\theta, \psi(t \mid y)) = \int_Y \{[\int_{R_h} [a - h(\theta)]^2\, d\psi_1(t \mid y) + M[1 - \chi(h(\theta))]]\varphi(y) + [\chi(h(\theta))(h^* - h(\theta))^2](1 - \varphi(y))\}\beta(\theta)e^{y\theta}\, d\mu(y).$$

The fact that $\psi^*(t \mid y)$ is better than $\psi(t \mid y)$ implies

THEOREM 3.1. *The class of procedures whose estimates of $h(\theta)$ are nonrandomized is complete.*

Theorem 3.1 implies that the most general type of procedure that we need consider is $t(y)$, where we write

$$t(y) = (a(y), 1) \text{ with probability } \varphi(y)$$
$$= (h^*, 0) \text{ with probability } (1 - \varphi(y)),$$

to denote that the procedure takes action $(a, 1)$ with probability $\varphi(y)$ and action $(h^*, 0)$ with probability $(1 - \varphi(y))$. Sometimes, for convenience, we write $t(y) = (a(y), \varphi(y))$ to denote such a procedure, since $t(y)$ is completely determined by the pair $(a(y), \varphi(y))$. Now we characterize the Bayes procedures in

THEOREM 3.2. *A procedure* $t(y) = (a(y), \varphi(y))$ *is Bayes with respect to the a priori distribution* $\xi(\theta)$, *if and only if*

$$t(y) = \left(\int_\Omega h(\theta)\beta(\theta)e^{y\theta}\, d\xi(\theta) / \int_\Omega \beta(\theta)e^{y\theta}\, d\xi(\theta), 1\right)$$

when $w(y) < 0$, *or with probability* λ *if* $w(y) = 0$, *for any* λ, $0 \leqq \lambda \leqq 1$, *and*

$$t(y) = (h^*, 0)$$

when $w(y) > 0$, *or with probability* $(1 - \lambda)$ *if* $w(y) = 0$; *where*

$$w(y) = -\left[\int_\Omega (h(\theta) - h^*)\beta(\theta)e^{y\theta}\, d\xi(\theta) / \int_\Omega \beta(\theta)e^{y\theta}\, d\xi(\theta)\right]^2 + \int_{\theta_1}^{\theta_2} [(h(\theta) - h^*)^2 + M]\beta(\theta)e^{y\theta}\, d\xi(\theta) / \int_\Omega \beta(\theta)e^{y\theta}\, d\xi(\theta).$$

We omit the proof.

From here on we will restrict the class of distributions to those for which R_h is closed and bounded, half-closed and open at infinity, (for example $[c, \infty)$ for some real c), or the entire real line, that is $(-\infty, \infty)$. Then for these distributions, if we let B^* be a subset of the class of all decision procedures (those with randomized estimates of $h(\theta)$ included), such that each procedure in B^* has a finite estimate of $h(\theta)$ and each procedure lies in the closure, according to the topology $\mathfrak{J}$ (See [11], p. 74), of the class of Bayes procedures, then it follows that the class of procedures B^* is complete. Furthermore it follows that the class of procedures, which is the intersection of B^* and the class of procedures with nonrandomized estimates of $h(\theta)$ is complete. For convenience we retain the symbol B^* to designate this latter class. We now set out to characterize the procedures in B^*. We do so by generalizing a result of Sacks [12] in giving a necessary and sufficient condition for a procedure to lie in B^*. Although this characterization is proved below for the case where $h(\theta) = \theta$ (that is when y is normal) and θ_1 and θ_2 are finite, indications of other cases for which the characterization is true are given. Before stating the theorem giving the necessary and sufficient condition, we need to state some definitions.

Let K be the class of real valued, bounded, continuous functions defined on T, which are zero outside a compact set. Denote the elements of K by u. Let L be the class of all measurable functions $f(y)$ on Y to the real line such that they are μ-integrable. Let $\|f\| = \int_Y |f(y)|\, d\mu(y)$ be the norm of f. We now restrict slightly the class of distributions to which the development applies by requiring that L be separable. (L is separable whenever the probability distribution is discrete or admits a density). This condition on L implies that convergence with respect to the topology $\mathfrak{I}$ is equivalent to regular convergence defined in Wald [13]. (See [11], p. 78). A sequence of decision procedures $\psi^n(t \mid y)$ is said to converge regularly to a limit $\psi(t \mid y)$, if for every $u \in K, f \in L$,

$$(3.1) \quad \lim_{n\to\infty} \int_Y \int_T f(y)u(t)\, d\psi^n(t \mid y)\, d\mu(y) = \int_Y \int_T f(y)u(t)\, d\psi(t \mid y)\, d\mu(y).$$

It is clear then that if $t_n(y) = (a_n(y), \varphi_n(y))$ is a sequence of procedures such that $\varphi_n(y)$ converges to $\varphi(y)$ a.e. μ and whenever $\varphi(y) = 0$, $a_n(y)$ converges to $a(y)$ a.e. μ, then $t_n(y)$ converges regularly to $t(y) = (a(y), \varphi(y))$.

Now we are ready to state

THEOREM 3.3. *Assume $h(\theta) = \theta$ and that θ_1 and θ_2 are finite. Then a necessary and sufficient condition for a procedure $t(y) = (a(y), \varphi(y))$ to lie in B^* is that there exists a non-decreasing, non-constant function $F(\theta)$ such that*

$$\int_\Omega \beta(\theta)e^{y\theta}\, dF(\theta) < \infty,$$

and

$$t(y) = \left(\int_\Omega \theta\beta(\theta)e^{y\theta}\, dF(\theta) / \int_\Omega \beta(\theta)e^{y\theta}\, dF(\theta), 1\right)$$

if $w(y) < 0$, and with probability λ, for any $0 \leqq \lambda \leqq 1$, if $w(y) = 0$;

$$t(y) = (\theta^*, 0)$$

if $w(y) > 0$, or with probability $(1 - \lambda)$, if $w(y) = 0$; where

$$(3.2) \quad w(y) = -\left[\int_\Omega (\theta - \theta^*)\beta(\theta)e^{y\theta}\, dF(\theta) / \int_\Omega \beta(\theta)e^{y\theta}\, dF(\theta)\right]^2$$
$$\int_{\theta_1}^{\theta_2} [(\theta - \theta^*)^2 + M]\beta(\theta)e^{y\theta}\, dF(\theta) / \int_\Omega \beta(\theta)e^{y\theta}\, dF(\theta).$$

Before proving this theorem, we give the following two lemmas.

LEMMA 3.1. *The function $w(y)$ given in (3.2) is not identically zero.*

PROOF. If $dF(\theta)$ is positive only outside the interval $[\theta_1, \theta_2]$ then $w(y)$ is negative except possibly for a single value of y for which $w(y) = 0$. This follows, since from a lemma due to Karlin ([8], p. 116) the function $\int_\Omega (\theta - \theta^*)\beta(\theta)e^{y\theta}\, dF(\theta)$ has at most one zero. Now consider

$$g(y) = w(y)\int_\Omega \beta(\theta)e^{y\theta}\, dF(\theta) / \int_{\theta_1}^{\theta_2} [(\theta - \theta^*)^2 + M]\beta(\theta)e^{y\theta}\, dF(\theta)$$
$$(3.3) \quad = -\left[\int_\Omega (\theta - \theta^*)\beta(\theta)e^{y\theta}\, dF(\theta)\right]^2 / \int_\Omega \beta(\theta)e^{y\theta}\, dF(\theta)$$
$$\cdot \int_{\theta_1}^{\theta_2} [(\theta - \theta^*)^2 + M]\beta(\theta)e^{y\theta}\, dF(\theta) + 1$$

and call the first term of the right-hand side of (3.3) $-g_1(y)$. Hence $w(y)$ is identically zero if and only if $g_1(y)$ is identically one. Suppose next then that $dF(\theta)$ is positive only on $[\theta_1, \theta_2]$. Then it follows by applying Schwartz's inequality that $g_1(y) < 1$ for all y. Therefore the only remaining case to consider is when $dF(\theta)$ is positive on $[\theta_1, \theta_2]$ and positive outside the interval. Suppose then that $dF(\theta)$ is positive within $[\theta_1, \theta_2]$ and also positive on some set (θ', θ'') where

$\theta_2 < \theta' \leqq \theta''$. We show that $g_1(y)$ tends to ∞ as y tends to ∞. For we may write

$$(3.4) \quad g_1(y) = \{\int_\Omega (\theta - \theta^*)\beta(\theta)e^{y\theta}\, dF(\theta)/\int_\Omega \beta(\theta)e^{y\theta}\, dF(\theta)\} \cdot\{\int_\Omega (\theta - \theta^*)\beta(\theta)e^{y\theta}\, dF(\theta)/\int_{\theta_1}^{\theta_2} [(\theta - \theta^*)^2 + M]\beta(\theta)e^{y\theta}\, dF(\theta)\}.$$

Now if we write the numerator of the second factor on the right-hand side of (3.4) as the sum of two integrals, one of which has $(-\infty, \theta_2)$ as limits of integration, it is then possible to show that the entire second factor tends to ∞ as y tends to ∞. The first factor on the right-hand side of (3.4) can be shown to approach a positive value as y tends to ∞ by showing that the ratio

$$\int_{\theta^*}^{\infty} (\theta - \theta^*)\beta(\theta)e^{y\theta}\, dF(\theta)/\int_{-\infty}^{\theta^*} (\theta - \theta^*)\beta(\theta)e^{y\theta}\, dF(\theta)$$

tends to ∞ as y tends to ∞. Thus, it follows that $g_1(y)$ tends to ∞ as y tends to ∞. For the case $dF(\theta)$ positive on values less than θ_1 a similar argument would show that $g_1(y)$ tends to ∞ as y tends to $-\infty$. These facts then prove that $g_1(y)$ cannot be identically one which suffices to prove the lemma.

LEMMA 3.2. *The function $w(y)$ given in* (3.2) *equals zero on at most a set of μ-measure zero.*

PROOF. Consider the two expressions

$$(3.5) \qquad \int_\Omega \beta(\theta)e^{y\theta}\, dF(\theta), \qquad \int_{\theta_1}^{\theta_2} [(\theta - \theta^*)^2 + M]\beta(\theta)e^{y\theta}\, dF(\theta).$$

By the conditions on $F(\theta)$ each is a finite LaPlace transform with a non-decreasing determining function. Therefore, by a theorem of Widder ([14], p. 273) each is analytic in the entire complex plane. Also since the first expression of (3.5) is positive, it follows from elementary properties of analytic functions that $w(y)$ is analytic. Hence, by the identity theorem of complex variables, if $w(y) = 0$ on some subset of the real line, where this subset is of positive Lebesgue measure, then $w(y)$ is identically zero. But from Lemma 3.1, $w(y)$ is not identically zero and thus it is zero at real values on at most a set of Lebesgue measure zero. Now since μ for Theorem 3.3 and the generalizations to follow is absolutely continuous, this implies that $w(y) = 0$ on at most a set of μ-measure zero. This completes the lemma.

PROOF OF THEOREM 3.3. First sufficiency. That is, suppose there exists an $F(\theta)$, with the stated properties, associated with given $t(y)$ and $w(y)$. Then define $t_n(y) = (a_n(y), \varphi_n(y)$ according to

$$a_n(y) = \int_{-n}^{n} \theta\beta(\theta)e^{y\theta}\, dF(\theta)/\int_{-n}^{n} \beta(\theta)e^{y\theta}\, dF(\theta),$$

and

$$w_n(y) = -[\int_{-n}^{n} (\theta - \theta^*)\beta(\theta)e^{y\theta}\, dF(\theta)/\int_{-n}^{n} \beta(\theta)e^{y\theta}\, dF(\theta)]^2 + \int_{(-n,n)\cap(\theta_1,\theta_2)} [(\theta - \theta^*)^2 + M]\beta(\theta)e^{y\theta}\, dF(\theta)/ \int_{-n}^{n} \beta(\theta)e^{y\theta}\, dF(\theta).$$

Now from Theorem 3.2, $t_n(y)$ is Bayes. Also Lemma 3.2 implies that if $w_n(y)$ converges to $w(y)$ a.e. μ then $\varphi_n(y)$ converges to $\varphi(y)$ a.e. μ. Hence it is obvious that $a_n(y)$ converges to $a(y)$ a.e. μ and $\varphi_n(y)$ converges to $\varphi(y)$ a.e. μ, which implies that $t(y)$ is the regular limit of a sequence of Bayes procedures and therefore $t(y)$ lies in B^*.

The necessity part of the proof is essentially the same as the proof given by Sacks [12]. That is, the function $F(\theta)$ can be constructed as in Sacks [12]. Since only minor modifications of that argument are needed for the proof here, we omit the details. This complete Theorem 3.3.

From the proof of Sacks' Lemma 3 and his Corollary 1, Theorem 3.3 can be generalized to distributions other than the normal. For example, one such generalization is

THEOREM 3.4. *Let* $Y = (-\infty, \infty)$, (*that is the spectrum or support of* μ *is* $(-\infty, \infty)$), $\Omega = (-\infty, \infty)$, $R_h = (-\infty, \infty)$, *and* θ_1, θ_2 *finite. Consider those distributions for which*

$$\lim_{\theta\to\infty} h(\theta)e^{-r\theta} = 0 \quad \text{for every} \quad r > 0,$$

$$\lim_{\theta\to-\infty} h(\theta)e^{r\theta} = 0 \quad \text{for every} \quad r > 0.$$

Then a necessary and sufficient condition for a procedure $t(y)$ *to lie in* B^* *is that there exist a non-decreasing, non-constant function* $F(\theta)$ such that $\int_\Omega \beta(\theta)e^{y\theta}\, dF(\theta)$ *is finite and*

$$t(y) = (\textstyle\int_\Omega h(\theta)\beta(\theta)e^{y\theta}\, dF(\theta)/\int_\Omega \beta(\theta)e^{y\theta}\, dF(\theta), \qquad 1)$$

if $w(y) < 0$, *and with probability* λ *for any* $0 \leqq \lambda \leqq 1$, *if* $w(y) = 0$;

$$t(y) = (h^*, 0),$$

if $w(y) > 0$, *or with probability* $(1 - \lambda)$, *if* $w(y) = 0$; *where*

$$(3.6) \quad w(y) = -[\textstyle\int_\Omega (h(\theta) - h^*)\beta(\theta)e^{y\theta}\, dF(\theta)/\int_\Omega \beta(\theta)e^{y\theta}\, dF(\theta)]^2 + \int_{\theta_1}^{\theta_2} [(h(\theta) - h^*)^2 + M]\beta(\theta)e^{y\theta}\, dF(\theta)/\int_\Omega \beta(\theta)e^{y\theta}\, dF(\theta).$$

At the end of Section 2 we remarked that it may be of interest to make decisions with regard to the parameter θ and not $h(\theta)$. If this were the case, then of course the class of distributions to which Theorem 3.3 is applicable is the same as given in Sacks [12].

We conclude this section by describing complete classes of procedures for the case when I consists of a single point. We will again assume the conditions of Theorem 3.3 since generalizations will either be obvious or will be indicated. The following lemma is needed.

LEMMA 3.3. *If I is a single point, say* θ^*, *then* $w(y)$ *of Theorem* 3.3 *crosses the y axis at most twice, and if it does cross twice, it is positive between crossings.*

PROOF. Let us consider the following function $g(y)$ which has the same sign as $w(y)$,

$$g(y) = w(y)[\textstyle\int_\Omega \beta(\theta)e^{y\theta}\, dF(\theta)]e^{y\theta^*}.$$

Let ν be the saltus of $F(\theta)$ at $\theta = \theta^*$. Then

$$(3.7) \quad g(y) = -\{[\textstyle\int_\Omega (\theta - \theta^*)\beta(\theta)e^{y\theta}\, dF(\theta)]^2/e^{y\theta^*} \int_\Omega \beta(\theta)e^{y\theta}\, dF(\theta)\} + M\beta(\theta^*)\nu.$$

Call the first term on the right-hand side of (3.7) $-g_1(y)$ and notice that $M\beta(\theta^*)\nu$ is a constant. Hence, if we show that the derivative of $g_1(y)$ has at most one zero, then by virtue of Rolle's theorem it follows that $g(y)$ and hence $w(y)$ has at most two zeros. But the numerator of the derivative of $g_1(y)$ can be shown to equal

$$e^{y\theta^*} \int_\Omega (\theta - \theta^*)\beta(\theta)e^{y\theta}\, dF(\theta)\{2 \int_\Omega \beta(\theta)e^{y\theta}\, dF(\theta) \int_\Omega \theta^2\beta(\theta)e^{y\theta}\, dF(\theta) - 2[\int_\Omega \theta\beta(\theta)e^{y\theta}\, dF(\theta)]^2 + [\int_\Omega (\theta - \theta^*)\beta(\theta)e^{y\theta}\, dF(\theta)]^2\}.$$

Since, by an application of Schwarz's inequality the term in brackets is positive, it follows that the derivative of $g_1(y)$ has as many zeros as

$$\int_\Omega (\theta - \theta^*)\beta(\theta)e^{y\theta}\, dF(\theta). \tag{3.8}$$

But from the lemma and remarks due to Karlin [8], (3.8) has at most one zero and it changes sign in the same order as $(\theta - \theta^*)$ does. This proves that $g(y)$ and hence $w(y)$ has at most two zeros, and when it does have two, it is positive between them and negative otherwise. This completes the proof of Lemma 3.3.

Now we prove

THEOREM 3.5. *If I consists of a single point, then any procedure in B^* is an interval-monotone procedure.*

PROOF. The proof is a consequence of Theorem 3.3 and Lemma 3.3.

An immediate corollary to Theorem 3.5 is that the class of interval-monotone procedure is a complete class. Also we make the following remarks.

(1) Lemma 3.3 is true for $w(y)$ as given in (3.6) provided for every y,

$$\int_\Omega \theta h(\theta)\beta(\theta)e^{y\theta}\, dF(\theta) - \int_\Omega \theta\beta(\theta)e^{y\theta}\, dF(\theta) \int_\Omega h(\theta)\beta(\theta)e^{y\theta}\, dF(\theta) \geqq 0.$$

(2) When I reduces to a single point the hybrid problem is analogous in some sense to "testing" a simple hypothesis and estimating.

(3) The class of all interval-monotone procedures is a relatively large class. Many of these procedures of course will not lie in B^*. For any particular interval there may be several estimates associated with it, such that the resulting pair lies in B^*. However, with each estimate there can exist at most one interval associated with it. This latter fact follows since the estimate

$$a(y) = \int_\Omega \theta\beta(\theta)e^{y\theta}\, dF(\theta) / \int_\Omega \beta(\theta)e^{y\theta}\, dF(\theta) = d[\log \int_\Omega \beta(\theta)e^{y\theta}\, dF(\theta)],$$

and by the uniqueness theorem for the bilateral Laplace transform, there exists only one such $F(\theta)$ that will give such an $a(y)$.

(4) In selecting a procedure for use one needs to also consider the fact that the class of procedures that lie in B^* and that are Bayes in the wide sense is a smaller complete class than B^*.

(5) Suppose we define a procedure, $t(y) = (a(y), \varphi(y))$ to be symmetric if $a(y) = -a(-y)$ and $\varphi(y) = \varphi(-y)$. Then for the special case where y is normal and $I = \theta^* = 0$, we can prove

(i) For any given symmetric interval, there is a symmetric interval-monotone procedure with the given interval that is admissible.

(ii) The class of symmetric interval procedures is complete among the class of all symmetric procedures.

The proof of (i) follows by the construction of an *a priori* distribution for which the given interval is part of an interval-monotone procedure that is a unique Bayes procedure. The proof of (ii) is an immediate consequence of some invariance properties as outlined in [2], pp. 223–224. We omit the details of the proofs of (i) and (ii).

4. Admissible procedures for the normal distribution. In this section we assume that y is normal with unknown mean θ, and known variance, which without loss of generality, is taken to be one. The interval I and hence I_h is taken to be $[-\theta_0, \theta_0]$ for $0 < \theta_0 < \infty$ and $h^* = \theta^*$ is chosen to be 0. Now consider the procedures $G_\gamma(y) = (a_\gamma(y), \varphi_\gamma(y))$, for $0 < \gamma \leqq 1$, where

$$a_\gamma(y) = 0, \varphi_\gamma(y) = 0 \quad \text{if} \quad |y| < k_\gamma$$
$$a_\gamma(y) = \gamma y, \varphi_\gamma(y) = 1 \quad \text{if} \quad |y| > k_\gamma,$$

and k_γ is the unique positive solution of the following equation.

$$\begin{aligned}(4.1)\qquad & -(\gamma y)^2[1 - \Phi((\theta_0 - \gamma y)/\gamma^{\frac{1}{2}}) + \Phi(-(\theta_0 + \gamma y)/\gamma^{\frac{1}{2}}))] \\ & + (M + \gamma)[\Phi((\theta_0 - \gamma y)/\gamma^{\frac{1}{2}}) - \Phi(-(\theta_0 + \gamma y)/\gamma^{\frac{1}{2}}))] \\ & - \gamma^{\frac{3}{2}} y/(2\pi)^{\frac{1}{2}}\{\exp[-(\theta_0 - \gamma y)^2/2\gamma] - \exp[-(\theta_0 + \gamma y)^2/2\gamma]\} \\ & - \theta_0\gamma^{\frac{1}{2}}/(2\pi)^{\frac{1}{2}}\{\exp[-(\theta_0 - \gamma y)^2/2\gamma] + \exp[-(\theta_0 + \gamma y)^2/2\gamma]\} = 0.\end{aligned}$$

In (4.1) $\Phi(x) = \int_{-\infty}^{x} \exp(-z^2/2)/(2\pi)^{\frac{1}{2}}\, dz$.

We now state

THEOREM 4.1. *The procedures $G_\gamma(y)$ are admissible.*

Before giving the proof we need the following two lemmas.

LEMMA 4.1. *There is exactly one positive value of y that satisfies equation* (4.1).

PROOF. Call the left-hand side of (4.1) $w(y)$. Notice that, by applying the identity $\Phi(x) = 1 - \Phi(-x)$, it is easy to check that $w(y)$ is symmetric in y. Also, since $w(0) > 0$, and $w(y)$ tends to $-\infty$, we need only verify that $w(y)$ crosses the y axis at most once for some $y > 0$. But now consider the derivative of $w(y)$, which is

$$\begin{aligned} w'(y) = & -(\gamma y)^2(\gamma)^{\frac{1}{2}}/(2\pi)^{\frac{1}{2}}\cdot\{\exp[-(\theta_0 - \gamma y)^2/2\gamma] - \exp[-(\theta_0 + \gamma y)^2/2\gamma]\} \\ & -2\gamma^2 y[1 - \Phi((\theta_0 - \gamma y)/\gamma^{\frac{1}{2}}) + \Phi(-(\theta_0 + \gamma y)/\gamma^{\frac{1}{2}})] \\ & -\gamma^{\frac{1}{2}}(M + \gamma)/(2\pi)^{\frac{1}{2}}\cdot\{\exp[-(\theta_0 - \gamma y)^2/2\gamma] - \exp[-(\theta_0 + \gamma y)^2/2]\} \\ & -\gamma^{\frac{3}{2}} y/(2\pi)^{\frac{1}{2}}\cdot\{(\theta_0 - \gamma y)\exp[-(\theta_0 - \gamma y)^2/2\gamma] \\ & \quad + (\theta_0 + \gamma y)\exp[-(\theta_0 + \gamma y)^2/2\gamma]\} \\ & -\gamma^{\frac{3}{2}}/(2\pi)^{\frac{1}{2}}\cdot\{\exp[-(\theta_0 - \gamma y)^2/2\gamma] - \exp[-(\theta_0 + \gamma y)^2/2\gamma]\} \\ & -\theta_0\gamma^{\frac{1}{2}}/(2\pi)^{\frac{1}{2}}\cdot\{(\theta_0 - \gamma y)\exp[-(\theta_0 - \gamma y)^2/2\gamma] \\ & \quad - (\theta_0 + \gamma y)\exp[-(\theta_0 + \gamma y)^2/2\gamma]\} \end{aligned}$$

$$= -2\gamma^2 y[1 - \Phi((\theta_0 - \gamma y)/\gamma^{\frac{1}{2}}) + \Phi(-(\theta_0 + \gamma y)/\gamma^{\frac{1}{2}})]$$
$$-\gamma^{\frac{1}{2}}(M + \gamma)/(2\pi)^{\frac{1}{2}} \cdot \{\exp[-(\theta_0 - \gamma y)^2/2\gamma] - \exp[-(\theta_0 + \gamma y)^2/2\gamma]\}$$
$$-\gamma^{\frac{3}{2}}/(2\pi)^{\frac{1}{2}} \cdot \{\exp[-(\theta_0 - \gamma y)^2/2\gamma] - \exp[-(\theta_0 + \gamma y)^2/2\gamma]\}$$
$$-\theta_0^2\gamma^{\frac{1}{2}}/(2\pi)^{\frac{1}{2}} \cdot \{\exp[-(\theta_0 - \gamma y)^2/2\gamma] - \exp[-(\theta_0 + \gamma y)^2/2\gamma]\}$$
$$= -2\gamma^2 y[1 - \Phi((\theta_0 - \gamma y)/\gamma^{\frac{1}{2}}) + \Phi(-(\theta_0 + \gamma y)/\gamma^{\frac{1}{2}})]$$
$$-\gamma^{\frac{1}{2}}/(2\pi)^{\frac{1}{2}} \cdot (M + \theta_0^2 + 2\gamma)$$
$$\cdot \{\exp[-(\theta_0 - \gamma y)^2/2\gamma] - \exp[-(\theta_0 + \gamma y)^2/2\gamma]\}.$$

Clearly for every $y > 0$, $w'(y)$ is negative, which implies that $w(y)$ will cross the y axis exactly once for $y \geqq 0$ and thus the proof of the lemma is complete.

LEMMA 4.2. *Let $f_n(y)$ be a sequence of functions defined for $y \geqq 0$, such that for each n, $f_n(y)$ is continuous, strictly decreasing, and $f_n(y)$ has a zero denoted by k_n. Also let $f(y)$ be a function defined for $y \geqq 0$, such that $f(y)$ is continuous, strictly decreasing, and with a zero denoted by k. Then if $\lim_{n\to\infty} f_n(y) = f(y)$, it follows that $\lim_{n\to\infty} k_n = k$.*

PROOF. Suppose that the sequence k_n does not converge to k. Then there exists a subsequence k_{n_i} and a positive number δ such that for every n_i, either

$$(k_{n_i} - k) > \delta, \tag{4.2}$$

or

$$(k - k_{n_i}) > \delta. \tag{4.3}$$

Suppose then that (4.2) is true. We have by hypothesis $\lim_{n\to\infty} f_n(k + \delta) = f(k + \delta) < 0$, yet $f_{n_i}(k + \delta) \geqq 0$ and so $\lim_{n\to\infty} f_{n_i}(k + \delta) \geqq 0$. Therefore $\lim_{n\to\infty} f_{n_i}(k + \delta) \neq f(k + \delta)$. This is a contradiction. Similarly for (4.3), and so we conclude that the lemma is true.

PROOF OF THEOREM 4.1. From Lemma 4.1 we may conclude that the procedures $G_\gamma(y)$ are well defined. Now for $0 < \gamma < 1$, $G_\gamma(y)$ can be shown, by use of Theorem 3.2, to be the essentially unique Bayes solution with respect to the *a priori* distribution $d\xi(\theta) = \lambda^{\frac{1}{2}} \exp(-\lambda\theta^2/2)/(2\pi)^{\frac{1}{2}}\, d\theta$, for $\lambda = (1 - \gamma)/\gamma$. This implies that for $0 < \gamma < 1$, $G_\gamma(y)$ is admissible. For $\gamma = 1$, we need to use an argument due to Blyth [3]. That is, suppose $G_1(y)$ is inadmissible. Then there exists a procedure $H(y) = (h(y), \varphi_h(y))$ which is better. That is,

$$\rho(\theta, G_1(y)) - \rho(\theta, H(y)) \geqq 0, \tag{4.4}$$

for every θ, with strict inequality for some θ. Now the risk function for any procedure $t(y) = (a(y), \varphi(y))$ is

$$\rho(\theta, t(y)) = \int_Y \{[(a(y) - \theta)^2 + M(1 - \chi(\theta))]\varphi(y) + \theta^2\chi(\theta)(1 - \varphi(y))\}$$
$$\cdot (2\pi)^{-\frac{1}{2}} \exp[-(y - \theta)^2/2]\, dy,$$

and so, although it is not continuous in θ for every fixed $t(y)$, still every point θ is a point of either right or left continuity. Therefore, since (4.4) must be strictly positive for at least one value of θ, there must exist some interval, (θ', θ'') say, such that $\rho(\theta, G_1(y)) - \rho(\theta, H(y)) \geqq \epsilon$ for $\epsilon > 0$, for every θ in (θ', θ''). This means that, if $d\xi(\theta)$ denotes an *a priori* density for θ, then the difference in expected risks is such that $\rho(\xi, G_1(y)) - \rho(\xi, H(y)) \geqq \epsilon \int_{\theta'}^{\theta''} d\xi(\theta)$. If in fact we let

$$d\xi_n(\theta) = (2\pi n)^{-\frac{1}{2}} e^{-\theta^2/2n}\, d\theta,$$

then

$$\rho(\xi_n, G_1(y)) - \rho(\xi_n, H(y)) \geqq K/n^{\frac{1}{2}}, \tag{4.5}$$

for K a positive constant. Now if we denote the Bayes solution with respect to $\xi_n(\theta)$ by $t_n(y) = (a_n(y), \varphi_n(y))$ and show that

$$\rho(\xi_n, G_1(y)) - \rho(\xi_n, t_n(y)) = o(1/n^{\frac{1}{2}}), \tag{4.6}$$

then it follows by the reasoning used by Blyth that the theorem is proved. That is, if (4.6) is true then the ratio of the left-hand side of (4.6) to the left-hand side of (4.5) would tend to zero, contradicting the fact that $t_n(y)$ is a Bayes solution with respect to $\xi_n(\theta)$. But from Theorem 3.2 the Bayes procedures $t_n(y)$ are as follows:

$$\begin{aligned} a_n(y) &= 0,\ \varphi_n(y) = 0 && \text{if } |y| < k_n \\ a_n(y) &= ny/n+1,\ \varphi_n(y) = 0 && \text{if } |y| \geqslant k_n \end{aligned}$$

where k_n is the unique non-negative solution of

$$\begin{aligned} &-(ny/n+1)^2[1 - \Phi((\theta_0 - ny/n+1)(n/n+1)^{-\frac{1}{2}}) \\ &\quad + \Phi(-(\theta_0 + ny/n+1)(n/n+1)^{-\frac{1}{2}})] \\ &+(M + n/n+1)[\Phi((\theta_0 - ny/n+1)(n/n+1)^{-\frac{1}{2}}) \\ &\quad - \Phi(-(\theta_0 + ny/n+1)(n/n+1)^{-\frac{1}{2}})] \\ &-(n/n+1)^{\frac{1}{2}}(2\pi)^{-\frac{1}{2}}y\{\exp[-(n+1/2n)(\theta_0 - ny/n+1)^2 \\ &\quad - \exp[-(n+1/2n)(\theta_0 + ny/n+1)^2]\} \\ &-(n\theta_0/n+1)\{\exp[-(n+1/2n)(\theta_0 - ny/n+1)^2] \\ &\quad + \exp[-(n+1/2n)(\theta_0 + ny/n+1)^2]\} = 0. \end{aligned} \tag{4.7}$$

Therefore the risk function for $t_n(y)$ reduces to

$$\begin{aligned} \rho(\theta, t_n(y)) = n/n+1 - \int_{-k_n}^{k_n} &[(ny/(n+1) - \theta)^2 \\ &+ M(1 - \chi(\theta)) - \chi(\theta)\theta^2](2\pi)^{-\frac{1}{2}} \exp[-(y-\theta)^2/2]\, dy, \end{aligned}$$

and hence after we reduce and interchange the order of integration,

$$\begin{aligned}
&\rho(\xi_n, G_1(y)) - \rho(\xi_n, t_n(y)) = 1/(n+1) \\
&\quad - \textstyle\int_{-k_1}^{k_1} \int_{-\infty}^{\infty} (y/n+1)^2 (2\pi n/n+1)^{-\frac{1}{2}} \exp[-(n+1/2n)(\theta_0 - ny/n+1)^2] \\
&\qquad \cdot [2\pi(n+1)]^{-\frac{1}{2}} \exp[-y^2/2(n+1)]\, d\theta\, dy \\
&\quad - 2\textstyle\int_{-k_1}^{k_1} \int_{-\infty}^{\infty} (y/n+1)(ny/(n+1) - \theta)(2\pi n/n+1)^{-\frac{1}{2}} \\
&\qquad \cdot \exp[-(n+1/2n)(\theta_0 - ny/n+1)^2][2\pi(n+1)]^{-\frac{1}{2}} \\
&\qquad \cdot \exp[-y^2/2(n+1)]\, d\theta\, dy - \textstyle\int_{-k_1}^{k_1} \int_{-\infty}^{\infty} [(ny/(n+1) - \theta)^2 \\
&\quad + M(1 - \chi(\theta)) - \theta^2 \chi(\theta)](2\pi n/n+1)^{-\frac{1}{2}} \\
&\qquad \cdot \exp[-(n+1/2n)(\theta_0 - ny/n+1)^2][2\pi(n+1)]^{-\frac{1}{2}} \\
&\qquad \cdot \exp[-y^2/2(n+1)]\, d\theta\, dy \\
&\quad + \textstyle\int_{-k_n}^{k_n} \int_{-\infty}^{\infty} [(ny/(n+1) - \theta)^2 + M(1 - \chi(\theta)) - \theta^2 \chi(\theta)] \\
&\quad \cdot (2\pi n/n+1)^{-\frac{1}{2}} \exp[-(n+1/2n)(\theta_0 - ny/n+1)^2][2\pi(n+1)]^{-\frac{1}{2}} \\
&\quad \cdot \exp[-y^2/2(n+1)]\, d\theta\, dy.
\end{aligned} \tag{4.8}$$

We notice that

$$\begin{aligned}
&\textstyle\int_{-\infty}^{\infty} [(ny/(n+1) - \theta)^2 + M(1 - \chi(\theta)) - \theta^2\chi(\theta)](2\pi n/n+1)^{-\frac{1}{2}} \\
&\qquad\qquad \cdot \exp[-(n+1/2n)(\theta_0 - ny/(n+1))^2]\, d\theta
\end{aligned}$$

is the same as the left-hand side of equation (4.7), which we now call $w_n(y)$. Also notice that the third term on the right-hand side of (4.8) is zero. Therefore we reduce the right-hand side of (4.8) to

$$\begin{aligned}
&1/(n+1) - \textstyle\int_{-k_1}^{k_1} (y/n+1)^2 [2\pi(n+1)]^{-\frac{1}{2}} \exp[-y^2/2(n+1)]\, dy \\
&\qquad - \textstyle\int_{-k_1}^{k_1} w_n(y)[2\pi(n+1)]^{-\frac{1}{2}} \exp[-y^2/2(n+1)]\, dy \\
&\qquad\qquad + \textstyle\int_{-k_n}^{k_n} w_n(y)[2\pi(n+1)]^{-\frac{1}{2}} \exp[-y^2/2(n+1)]\, dy.
\end{aligned} \tag{4.9}$$

Clearly the second term in (4.9) is bounded by $1/(n+1)$. Also if we denote $\bar{k}_n = \max(k_n, k_1)$, $\underline{k}_n = \min(k_n, k_1)$, C_1 and C_2 positive constants, it follows from (4.8), (4.9), the properties of $w_n(y)$, and Lemma 4.2 that

$$\begin{aligned}
&\rho(\xi_n, G_1(y)) - \rho(\xi_n, t_n(y)) \\
&\qquad \leqq 2/(n+1) + 2C_1 \textstyle\int_{\underline{k}_n}^{\bar{k}_n} [2\pi(n+1)]^{-\frac{1}{2}} \exp[-y^2/2(n+1)]\, dy \\
&\qquad = 2/(n+1) + 2C_1C_2(\bar{k}_n - \underline{k}_n)/(n+1)^{\frac{1}{2}} = o(1/n^{\frac{1}{2}}).
\end{aligned}$$

This shows (4.6) and hence the theorem is proved.

5. Admissible procedures for the binomial, Poisson, and gamma distributions. In this section we obtain the analogue of Theorem 4.1 for the cases where the distribution of y is binomial, Poisson and gamma. For some of these cases we modify the loss function given in Table 2.1 by replacing every squared error term by a term which is squared error divided by a factor proportional to the variance of y. These modified loss functions correspond to loss functions considered for estimation problems by Hodges and Lehmann [7]. Analogues of the results given here however, can be obtained even when the loss function is not so modified. The fixed interval of the parameter space in all cases will be denoted by I and I will be taken to be an interval $[0, \theta_0]$ where θ_0 is a positive finite number, and in the binomial case, $\theta_0 < 1$. When we decide $\theta \varepsilon I$, then we will always estimate θ by zero. For each of the three distributions considered in this section we find specific admissible procedures. Some of these procedures may be such that they allow θ to be estimated by zero even when the observation is positive. At first, this may seem unreasonable since if θ is zero, then the probability of any observation being positive is zero. However, we agree to estimate θ by zero whenever we decide $\theta \varepsilon I$, since this is a convenient way of stating that θ is too small to be of practical significance.

For the binomial distribution, that is $p(y, \theta) = \binom{n}{y}\theta^y(1 - \theta)^{n-y}$, $0 < \theta < 1$; $y = 1, 2, \cdots, n$, let us first consider the case where the table of losses is as given in Table 5.1. Now consider the procedures $G_\gamma(y) = (a_\gamma(y), \varphi_\gamma(y))$, $0 < \gamma \leqq 1$, where

$$a_\gamma(y) = 0, \qquad \varphi_\gamma(y) = 0 \qquad \text{if } y < k_\gamma \text{, and with probability } (1-\gamma) \text{ if } y = k_\gamma$$

$$a_\gamma(y) = \gamma y/n, \qquad \varphi_\gamma(y) = 1 \qquad \text{if } y > k_\gamma \text{, and with probability } \nu \text{ if } y = k_\gamma;$$

for each given γ, ν can take on every value between zero and one; k_γ is the integer lying between one and n for which $v(y) = 0$, if such an integer exists; otherwise k_γ is any non-integral number, if it exists, such that $0 < k_\gamma < n$, and $v([k_\gamma]) > 0$, and $v([k_\gamma] + 1) < 0$, where $[k_\gamma]$ is the smallest integer less than or equal to k_γ; otherwise k_γ is any number greater than n; here

$$v(y) = -\gamma y/n + [(n - \gamma y)/(n + \gamma)]MI_{\theta_0}(y + 1, (n/\gamma) - y + 1) + [\gamma(y + 1)/(n + \gamma)]I_{\theta_0}(y + 2, (n/\gamma) - y), \tag{5.1}$$

TABLE 5.1
Table of losses

$x(\theta)$ \ $b(y)$	0	1
0	0	$[(a - \theta)^2/\theta(1 - \theta)] + M$
1	$\theta/(1 - \theta)$	$(a - \theta)^2/\theta(1 - \theta)$

where

$$I_x(s, t) = [\Gamma(s + t)/\Gamma(s)\Gamma(t)]\int_0^x p^{s-1}(1 - p)^{t-1}\, dp,$$

for $0 < p < 1$, $s > 0$, $t > 0$, $0 \leqq x \leqq 1$; that is $I_x(s, t)$ is the ratio of the incomplete beta function to the complete beta function.

Note that $v(0) > 0$, and in Lemma 5.1 below we show $v(y)$ decreases as y varies over the integers from 1 to n, thus insuring that k_γ is uniquely defined. Finally we can state

THEOREM 5.1. *The procedures $G_\gamma(y)$ are admissible.*

We need the following lemmas.

LEMMA 5.1. *As y increases over the integers 1 to n, the function $v(y)$ in (5.1) decreases.*

PROOF. From (5.1) we can write

$$\begin{aligned} v(y) &= -\gamma y/n + [\gamma y/(n + \gamma)]I_{\theta_0}(y + 2, (n/\gamma) - y) \\ &\quad + [(n - \gamma y)/(n + \gamma)]MI_{\theta_0}(y + 1, (n/\gamma) - y + 1) \\ &\quad + [\gamma/(n + \gamma)]I_{\theta_0}(y + 2, (n/\gamma) - y). \end{aligned} \tag{5.2}$$

Now by use of partial integration we can derive the following recursion relationship,

$$I_x(s + 1, t - 1) = I_x(s, t) - [\Gamma(s + t)/\Gamma(s + 1)\Gamma(t - 1)]x^s(1 - x)^{t-1}.$$

This relationship enables us to conclude that as y increases one unit, the terms $I_{\theta_0}(y + 1, (n/\gamma) - y + 1)$ and $I_{\theta_0}(y + 2, (n/\gamma) - y)$ decrease. Hence the last two terms on the right-hand side of (5.2) decrease. Furthermore the decrease in the first term on the right-hand side of (5.2), which is γ/n, outweighs the maximum increase of the second term which is $\gamma/(n + \gamma)$. Therefore $v(y)$ does decrease as y increases and the lemma is proved.

LEMMA 5.2. *Given γ, $0 < \gamma \leqq 1$ and the function $v(y)$ of (5.1), suppose k_γ is an integer, $1 \leqq k_\gamma < n$.* For a given value of ν, $0 \leqq \nu \leqq 1$, *let $G_\gamma^\nu(y)$ denote the procedure of Theorem 5.1 with these given values of γ and ν. Then given any pair of numbers, ν_1, ν_2 say, where $0 \leqq \nu_1 < \nu_2 \leqq 1$, $G_\gamma^{\nu_1}(y)$ is not better than $G_\gamma^{\nu_2}(y)$, nor is $G_\gamma^{\nu_2}(y)$ better than $G_\gamma^{\nu_1}(y)$.*

PROOF. Since $G_\gamma^{\nu_1}(y)$ and $G_\gamma^{\nu_2}(y)$ agree for every y except $y = k_\gamma$, the difference in their risks is

$$\begin{aligned} \rho(\theta, G_\gamma^{\nu_1}(y)) - \rho(\theta, G_\gamma^{\nu_2}(y)) &= \binom{n}{k_\gamma}\theta^{k_\gamma}(1 - \theta)^{n-k_\gamma} \\ &\cdot\{[(\gamma k_\gamma - \theta)^2/\theta(1 - \theta)] + M(1 - \chi(\theta)) - \theta\chi(\theta)/(1 - \theta)\}(\nu_1 - \nu_2). \end{aligned} \tag{5.3}$$

Now for $\theta \varepsilon I$, (5.3) < 0 indicating that $G_\gamma^{\nu_1}(y)$ is better than $G_\gamma^{\nu_2}(y)$ for all $\theta \varepsilon I$. If $\theta \not\varepsilon I$, that is $\theta > \theta_0$, then (5.3) > 0, for at least all $\theta > \gamma k_\gamma/n$. Hence, on these values of θ, $G_\gamma^{\nu_2}(y)$ is better than $G_\gamma^{\nu_1}(y)$. Therefore, neither procedure can be better than the other for every θ and the proof of the lemma is complete.

PROOF OF THEOREM 5.1. Lemma 5.1 assures us that the procedures $G_\gamma(y)$

are well defined. Now for a fixed γ, it is easy to verify that $G_\gamma(y)$ is a Bayes solution with respect to the *a priori* distribution

$$d\xi_\gamma(\theta) = (n\lambda + 1)(1 - \theta)^{n\lambda}\, d\theta$$

where $\lambda = (1 - \gamma)/\gamma$. That is, the procedures $G_\gamma^{\nu}(y)$, $0 \leqq \nu \leqq 1$, are the only Bayes solution with respect to $\xi_\gamma(\theta)$. This fact, and Lemma 5.2 insure that each $G_\gamma^{\nu}(y)$ is admissible. This completes the proof of the theorem.

If we consider the case where the losses are those of Table 2.1, with $h(\theta) = \theta$, $h^* = 0$, instead of those of Table 5.1, a theorem analogous to Theorem 5.1 could be proved. The proof for $0 < \gamma < 1$ would be essentially the same as the proof of Theorem 5.1. For $\gamma = 1$ however, it is not true that $G_1(y)$ is a Bayes solution and a limiting argument is necessary to prove admissibility. The limiting argument that is used is one given by Karlin [9]. However, since this type of argument is easier to apply to the binomial case than to the Poisson case, to which it is applied below, we omit further discussion of the binomial case.

TABLE 5.2

Table of losses

$x(\theta)$ \ $b(y)$	0	1
0	0	$[(a - \theta)^2/\theta] + M$
1	θ	$(a - \theta)^2/\theta$

For the Poisson distribution, that is

$$p(y, \theta) = e^{-\theta}\theta^y/y!, \qquad 0 < \theta < \infty;\ y = 0, 1, 2, \cdots,$$

consider the table of losses to be as given in Table 5.2. Now consider the procedures $G_\gamma(y) = (a_\gamma(y), \varphi_\gamma(y))$, $0 < \gamma \leqq 1$, where

$$a_\gamma(y) = 0, \qquad \varphi_\gamma(y) = 0 \quad \text{if } y < k_\gamma, \quad \text{and with probability } (1 - \nu) \text{ if } y = k_\gamma,$$

$$a_\gamma(y) = \gamma y, \qquad \varphi_\gamma(y) = 1 \quad \text{if } y > k_\gamma, \quad \text{and with probability } \nu \text{ if } y = k_\gamma;$$

for each given γ, ν can take on every value between zero and one; k_γ is the integer greater than or equal to one for which $w(y) = 0$, if such an integer exists, otherwise k_γ is any non-integral number, such that $0 < k_\gamma < \infty$, and such that $w([k_\gamma]) > 0$ and $w([k_\gamma] + 1) < 0$; and

$$\begin{aligned} w(y) &= -\gamma y[\exp(-\theta_0/\gamma)(1 + \theta_0/\gamma + \theta_0^2/2\gamma^2 \\ &\quad + \cdots + \theta_0^{y+1}/(y + 1)!\gamma^{y+1})] \\ &\quad + M[1 - \exp(-\theta_0/\gamma)(1 + \theta_0/\gamma + \theta_0^2/2\gamma^2 + \cdots + \theta_0^y/y!\gamma^y)], \end{aligned} \tag{5.4}$$

for $y = 0, 1, 2, \cdots$. Now we can state

THEOREM 5.2. *The procedures $G_\gamma(y)$ are admissible.*

For $0 < \gamma < 1$ the proof of Theorem 5.2 follows the same line of reasoning as the proof of Theorem 5.1. In this case $G_\gamma(y)$ are Bayes solutions with respect to $d\xi_\gamma = \lambda e^{-\lambda\theta}\, d\theta$, for $\lambda = (1 - \gamma)/\gamma$. We omit the details for the cases where $0 < \gamma < 1$. For $\gamma = 1$, we need the following lemmas.

LEMMA 5.3. *Suppose a procedure $t(y) = (a(y), \varphi(y))$ is such that*

$$\liminf_{y\to\infty} a^2(y)\varphi(y) < \infty.$$

Then the risk for $t(y)$ is unbounded.

PROOF. Let Y^* be an infinite set of points on which $a(y)\varphi(y) \leqq K$, for some $K \geqq 0$. Clearly such a K exists. Now for $\theta > \theta_0$,

$$\begin{aligned}\rho(\theta, t(y)) &= \textstyle\sum_{y=0}^{\infty} \{[(a(y) - \theta)^2/\theta]\varphi(y) + \theta(1 - \varphi(y))\}e^{-\theta}\theta^y/y! \\ &\geqq \textstyle\sum_{Y^*} (a^2(y)\varphi(y)/\theta - 2a(y)\varphi(y))e^{-\theta}\theta^y/y! + \theta\sum_{Y^*}e^{-\theta}\theta^y/y! \\ &\geqq \theta\textstyle\sum_{Y^*} e^{-\theta}\theta^y/y! - 2K.\end{aligned} \tag{5.5}$$

But since Y^* is an infinite set, within Y^* there is a monotone increasing sequence of integers, say y_n, with $y_n \to \infty$, as $n \to \infty$. Therefore, if we let θ tend to ∞ through the sequence $\theta_n = y_n$, we see from (5.5) that

$$\rho(\theta_n, t(y)) \geqq [\theta_n e^{-\theta_n}\theta_n^{\theta_n}/\theta_n!] - 2K. \tag{5.6}$$

Using Stirling's formula, we see that as $n \to \infty$ the right-hand side of (5.6) tends to $(\theta_n/2\pi)^{\frac{1}{2}} - 2K$, which tends to ∞. Hence $\rho(\theta, t(y))$ is unbounded and the lemma is proved.

LEMMA 5.4. *For any procedure $t(y) = (a(y), \varphi(y))$ of bounded risk such that $\varphi(y) < 1$ on an infinite set, there exists a procedure $t^*(y) = (a^*(y), \varphi^*(y))$ such that $\varphi^*(y) < 1$ on at most a finite set and $t^*(y)$ is better than $t(y)$.*

PROOF. We show that there exists an integer $\tilde{y}$, $1 \leqq \tilde{y} < \infty$, such that if

$$t^*(y) = t(y) \qquad \text{for} \quad y = 0, 1, 2, \cdots, \tilde{y} - 1$$

and

$$t^*(y) = (a(y)\varphi(y), 1) \quad \text{for} \quad y \geqq \tilde{y},$$

then $t^*(y)$ is better than $t(y)$.

First notice that for any $\tilde{y} \geqq 1$, when $\theta > \theta_0$, $t^*(y)$ is as good as $t(y)$, for

$$\begin{aligned}\rho(\theta, t(y)) - \rho(\theta, t^*(y)) = \textstyle\sum_{y=\tilde{y}}^{\infty} \{&[(a(y) - \theta)^2\varphi(y)/\theta] \\ &+ \theta(1 - \varphi(y)) - (a(y)\varphi(y) - \theta)^2/\theta\}e^{-\theta}\theta^y/y!,\end{aligned}$$

and the term in brackets is always non-negative. For all $\theta \leqq \theta_0$, we seek a $\tilde{y}$ such that

$$\begin{aligned}\rho(\theta, t(y)) - \rho(\theta, t^*(y)) = \textstyle\sum_{y=\tilde{y}}^{\infty} \{&[[(a(y) - \theta)^2/\theta] + M]\varphi(y) \\ &- [[(a(y)\varphi(y) - \theta)^2/\theta] + M]\}e^{-\theta}\theta^y/y! > 0.\end{aligned} \tag{5.7}$$

Rewriting (5.7) we seek a $\tilde{y}$ such that for all $\theta \leqq \theta_0$,

$$(5.8) \quad \textstyle\sum_{\tilde{y}}^{\infty} a^2(y)\varphi(y)(1 - \varphi(y))e^{-\theta}\theta^y/y! \geqq \sum_{\tilde{y}}^{\infty} (1 - \varphi(y))(\theta^2 + M\theta)e^{-\theta}\theta^y/y!.$$

But from Lemma 5.3 and the hypothesis of this lemma we know that

$$\liminf_{y\to\infty} a^2(y)\varphi(y) = \infty,$$

which implies that there exists a sufficiently large integer y, say $\tilde{y}$, such that

$$(5.9) \qquad a^2(y)\varphi(y) > \theta_0^2 + M\theta_0 > \theta^2 + M\theta,$$

for every $y \geqq \tilde{y}$ and every $\theta < \theta_0$. Now (5.9) implies (5.8) holds and thus the proof of the lemma is complete.

PROOF OF THEOREM 5.2. Suppose $G_1(y)$ is not admissible. Then there exists a procedure $H(y) = (h(y), \varphi_h(y))$ which is better, that is

$$(5.10) \qquad \rho(\theta, G_1(y)) - \rho(\theta, H(y)) \geqq 0$$

for every θ, with strict inequality for some θ. Now let us partition the sample space into the following disjoint sets,

$$Y^{(1)} = \{y : \varphi_1(y) = 1, \varphi_h(y) = 1\} \qquad Y^{(0)} = \{y : \varphi_1(y) = 0, \varphi_h(y) = 0\}$$
$$Y^{+} = \{y : \varphi_1(y) = 1, \varphi_h(y) < 1\} \qquad Y^{-} = \{y : \varphi_1(y) = 0, \varphi_h(y) > 0\}.$$

(We assume k_1 is not an integer, since if it is, the proof is one step longer.) If we write out the difference of the risk functions for the two procedures and use (5.10) we are led to

$$(5.11) \quad \begin{aligned} &\textstyle\sum_{Y^{(1)}} (h(y) - y)^2 e^{-\theta}\theta^{y-1}/y! \leqq 2\sum_{Y^{(1)}}(y - h(y))(y - \theta)e^{-\theta}\theta^{y-1}/y! \\ &\quad + \textstyle\sum_{Y^+} \{[[(y - \theta)^2/\theta] + M(1 - \chi(\theta)) - \theta\chi(\theta)] \\ &\quad - \varphi_h(y)[[(h(y) - \theta)^2/\theta] + M(1 - \chi(\theta)) - \theta\chi(\theta)]\}e^{-\theta}\theta^y/y! \\ &\quad - \textstyle\sum_{Y^-} [[(h(y) - \theta)^2/\theta] + M(1 - \chi(\theta)) - \theta\chi(\theta)]\varphi_h(y)e^{-\theta}\theta^y/y! \end{aligned}$$

Now consider the function $F(\theta)$ where $dF(\theta) = d\theta$, let r, s be two reals numbers such that $0 < r < s < \infty$, and call the left-hand side of (5.11) $T(\theta)$. If we then integrate both sides of (5.11) with respect to θ over the interval (r, s), and use the same steps in Karlin [9], p. 413, we are led to

$$(5.12) \quad \begin{aligned} &\textstyle\int_r^s T(\theta)\, d\theta \leqq 2(sT(s))^{\frac{1}{2}} + 2(rT(r))^{\frac{1}{2}} \\ &\quad + \textstyle\int_r^s \sum_{Y^+} \{[[(y - \theta)^2/\theta] + M(1 - \chi(\theta)) - \theta\chi(\theta)] \\ &\quad - \varphi_h(y)[[(h(y) - \theta)^2/\theta] + M(1 - \chi(\theta)) - \theta\chi(\theta)]\}e^{-\theta}\theta^y/y!\, d\theta \\ &\quad - \textstyle\int_r^s \sum_{Y^-} [[(h(y) - \theta)^2/\theta] + M(1 - \chi(\theta)) - \theta\chi(\theta)]\varphi_h(y)e^{-\theta}\theta^y/y!\, d\theta. \end{aligned}$$

Now fix r and notice that the last two expressions on the right-hand side of (5.12) are bounded for every s. For Y^- is surely a finite set, it lies in the set 0 to $[k_1]$, and Y^+ must be a finite set by virtue of the fact that the risk function for $G_1(y)$ is bounded and Lemma 5.4. Since Y^+ and Y^- are finite sets the last two terms on the right-hand side represent a finite sum of integrals, each of which is bounded. Thus, for sufficiently large s we can find some constant K, such that, by (5.12)

$$F(s) = \int_r^s T(\theta)\, d\theta \leqq K(sT(s))^{\frac{1}{2}}.$$

Again, if we follow Karlin [9], p. 413, we find that

$$\liminf_{s\to\infty} (sT(s))^{\frac{1}{2}} = 0. \tag{5.13}$$

Now consider $Q(r) = \int_r^\infty T(\theta)\, d\theta$. By virtue of (5.12), (5.13) and following Karlin again we are led to the conclusion that $Q(r) \equiv 0$, which implies $T(\theta) \equiv 0$, which in turn implies $h(y) = y$ for every $y \in Y^{(1)}$. This means that if $H(y)$ is better than $G(y)$ we must have

$$\begin{aligned} 0 \leqq \textstyle\sum_{Y^+} \int_0^\infty \{&[[(y - \theta)^2/\theta] + M(1 - \chi(\theta)) - \theta\chi(\theta)] \\ &- \varphi_h(y)[[(h(y) - \theta)^2/\theta] + M(1 - \chi(\theta)) - \theta\chi(\theta)]\}e^{-\theta}\theta^y/y!\, d\theta \\ &- \textstyle\sum_{Y^-} \int_0^\infty [[(h(y) - \theta)^2/\theta] + M(1 - \chi(\theta)) - \theta\chi(\theta)] \\ &\cdot \varphi_h(y)e^{-\theta}\theta^y/y!\, d\theta. \end{aligned} \tag{5.14}$$

But now note that Y^+ is contained in the sets of points for which $w(y)$ in (5.4) is negative and Y^- is contained in the sets of points for which $w(y)$ is positive. Furthermore, $w(y)$ has the same sign as

$$\int_0^\infty [[(y - \theta)^2/\theta] + M(1 - \chi(\theta)) - \theta\chi(\theta)]e^{-\theta}\theta^y/y!\, d\theta \tag{5.15}$$

and if $h(y)$ replaces y in (5.15) the resulting integral is greater than (5.15) for every $y \in Y^+$. But these facts contradict (5.14) unless Y^+ and Y^- are empty. Thus $G(y)$ cannot be beaten by $H(y)$ and we conclude $G(y)$ is admissible. This completes the proof of Theorem 5.2.

For the gamma distribution, that is

$$p(y, \theta) = y^{\alpha-1}e^{-y/\theta}/(\alpha - 1)\theta^\alpha, \qquad 0 < \theta < \infty, y > 0, \alpha > 0,$$

consider the table of losses to be given in Table 5.3. Now consider the procedure $G(y) = (a(y), \varphi(y))$, where

$$\begin{aligned} a(y) &= 0, \quad \varphi(y) = 0 && \text{if } y < k \\ a(y) &= y/\alpha + 1, \quad \varphi(y) = 1 && \text{if } y \geqq k, \end{aligned}$$

TABLE 5.3
Table of losses

$\chi(\theta)$ \ $b(y)$	0	1
0	0	$[(a - \theta)^2/\theta^2] + M$
1	1	$(a - \theta)^2/\theta^2$

where k is the unique solution of the following equation,

$$-(1/(\alpha + 1))^2 + [(M + 1)/\alpha(\alpha + 1)]\int_{y/\theta_0}^{\infty} e^{-w}w^{\alpha-1}/(\alpha - 1)!\,dw = 0.$$

It is then possible to prove

THEOREM 5.3. *The procedure $G(y)$ is admissible.*

The proof is omitted.

6. Acknowledgment. I am deeply grateful to Professor T. W. Anderson for his guidance throughout the course of this work. His suggestions, comments, and criticisms were of great value to me. This work was done while under a grant from the Biostatistics Training Program of Columbia University, which is partially supported by the National Institutes of Health. I would like to take this opportunity to thank Professor J. W. Fertig, chairman of the biostatistics department, for his continual encouragement.

REFERENCES

[1] ANDERSON, T. W. (1962). The choice of the degree of a polynomial regression as a multiple decision problem. *Ann. Math. Statist.* **33** 255–265.

[2] BLACKWELL, D. and GIRSHICK, M. A. (1954). *Theory of Games and Statistical Decisions.* Wiley, New York.

[3] BLYTH, C. R. (1951). On minimax statistical decision problems and their admissibility. *Ann. Math. Statist.* **22** 22–42.

[4] COHEN, ARTHUR (1963). Some questions of predicting from a regression function and a hybrid problem on the exponential family. Ph.D. Thesis, Columbia Univ.

[5] COHEN, ARTHUR (1965). Estimates of linear combinations of the parameters in the mean vector of a multivariate distribution. *Ann. Math. Statist.* **36** 78–87.

[6] GIRSHICK, M. A. and SAVAGE, L. J., (1951). Bayes and minimax estimates for quadratic loss functions. *Proc. Second Berkely Symp. Math. Statist. Prob.*, 53–73 University of California Press.

[7] HODGES, J. L., JR. and LEHMANN, E. L. (1951). Some applications of the Cramer-Rào inequality. *Proc. Second Berkely Symp. Math. Statist. Prob.* 13–22. University of California Press.

[8] KARLIN, SAMUEL (1955). Decision theory of Polya type distributions. Case of two actions, I. *Proc. Third Berkely Symp. Math. Statist. Probab.* **1** 115–128. University of California Press.

[9] KARLIN, SAMUEL (1958). Admissibility for estimation with quadratic loss, *Ann. Math. Statist.* **29** 406–436.

[10] KITAGAWA, TOSIO (1963). Estimation after preliminary tests of significance, Univ. of California Publ. Statist. **3** 147–186.

[11] LECAM, L. (1955). An extension of Wald's theory of statistical decision functions. *Ann. Math. Statist.* **26** 69–81.

[12] SACKS, JEROME (1963). Generalized Bayes solutions in estimation problems. *Ann. Math. Statist.* **34** 751–768.

[13] WALD, ABRAHAM (1950). *Statistical Decision Functions.* Wiley, New York.

[14] WIDDER, D. V. (1941). *The Laplace Transform.* Princeton.