*The Annals of Statistics*
2000, Vol. 28, No. 6, 1561–1569

# THE PROBLEM OF LOW COUNTS IN A SIGNAL PLUS NOISE MODEL

BY MICHAEL WOODROOFE[1] AND HSIUYING WANG[2]

*University of Michigan and Academia Sinica*

Consider the model $X = B + S$, where $B$ and $S$ are independent Poisson random variables with means $\mu$ and $\nu$, $\nu$ is unknown, but $\mu$ is known. The model arises in particle physics and some recent articles have suggested conditioning on the observed bound on $B$; that is, if $X = n$ is observed, then the suggestion is to base inference on the conditional distribution of $X$ given $B \leq n$. This conditioning is non-standard in that it does not correspond to a partition of the sample space. It is examined here from the view point of decision theory and shown to lead to admissible formal Bayes procedures.

**1. Introduction.** In some problems a signal $S$ may be superimposed on a background $B$, leaving an observed count $X = B+S$. Here we suppose that $B$ and $S$ are independent Poisson random variables with means $\mu$ and $\nu$, so that $X$ has a Poisson distribution with mean $\theta = \mu+\nu$; $\nu$ is regarded as an unknown parameter, but $\mu$ is assumed to be known, as might be appropriate if there were historical data on the background. What special techniques, if any, are appropriate if the observed count $X$ is smaller than the expected background? A problem of this nature has arisen recently in physics. The KARMEN 2 Group has been searching for a neutrino oscillation, reported earlier from an experiment at the Los Alamos Neutrino Detector. As of Summer 1998, they had expected to see about $3 \pm .1$ background events and at least one signal event, based on the earlier findings (and using rounded values), but had observed nothing. See Zeitnitz et al. (1998). What inference about $\nu$ is appropriate here? Can the hypothesis $H_0 : \nu \geq 1$ be rejected? A naive analysis suggests that it can. If $\mu = 3$ is regarded as a known quantity and the hypothesis is rewritten as $H_0 : \theta \geq 4$, then the $p$ value is $P_\theta\{X \leq 0\} \leq e^{-4}$ for $\theta \geq 4$, and this is less than the usual levels of significance. But this analysis is suspect, because if $X = 0$, then both $B$ and $S$ must be zero; and $P_4\{S \leq 0\} = e^{-1}$, which is not less than the usual levels of significance. The second value is obtained by conditioning on the ancillary variable $B = 0$ and seems right in this context.

The problem becomes much more interesting when the observed count is non-zero but smaller than the expected background, since then it is no longer possible to recover the value of $B$. Roe and Woodroofe (1999) argue that if $X = n$, then it is appropriate to base inferences on the conditional distribu-

Received October 1999; revised September 2000.

[1]Supported by the U.S. Army and National Security Agency.

[2]Supported by the National Science Council of Taiwan.

*AMS* 2000 *subject classifications*. Primary 62C15; secondary 62F03, 62P35.

*Key words and phrases*. Admissibility, ancillary statistic, Bayesian solutions, confidence intervals, neutrino oscillations, $p$-values, risk.

tion of $X$ given $B \leq n$. In effect, they argue that $B \leq n$ provides a more appropriate reference set, since "it seems unfair to include smaller than expected background radiation as evidence against $H_0$." The Particle Data Group (1992) [PDG] and Read (2000) have made related suggestions. This type of conditioning does not correspond to a partition of the sample space and, so, there is no obvious relation between conditional and unconditional properties. The purpose of this paper is to study such conditional inference from a decision theoretic point of view. In Section 2, it is shown that the usual $p$-value for testing $\theta \geq \theta_0$ is inadmissible, in the terminology of Hwang et al. (1992), when $X \sim \text{Poisson}(\theta)$ and $\theta \geq \mu > 0$. Also, an admissible modification is constructed using conditional probability given that $B$ is at most the observed value of $X$. Given the close association between hypothesis tests and confidence sets, the (in)admissibility results for $p$-values have implications for confidence sets, but it is possible to formulate the entire question in terms of confidence sets. This is done in Section 3, where the problem of finding an upper confidence bound is formulated as a decision problem. It is shown that the bound obtained from the modified $p$-value is admissible formal Bayes in the latter problem.

As noted above, the problem is of current interest to particle physicists, and there is some controversy. A competing method is the unified method of Feldman and Cousins (1996), which consists of regions of high relative likelihood. Roe and Woodroofe (1999) criticized the unified method, because the intervals may depend on $\mu$ when $B = 0$; and Cousins has criticized the Roe Woodroofe approach, finding that it produces a lower boundary that is much too large when applied to a related problem. By focusing on the one-sided case, we have changed the formulation of Roe and Woodroofe (1999), whose interest was in confidence intervals, and we do not address Cousins' criticism. Our results do support the non-standard conditioning, however, and have clear implications for the PDG Method.

**2. (In)Admissible P-Values.** It is convenient to let $f_\theta$ and $F_\theta$ denote the probability mass function and distribution function of the Poisson distribution with mean $\theta$, so that $f_\theta(n) = (1/n!)\theta^n e^{-\theta}$ and $F_\theta(n) = f_\theta(0) + \cdots + f_\theta(n)$ for $n = 0, 1, 2, \ldots$. As in the Introduction, suppose that $X \sim F_\theta$, where $\mu \leq \theta < \infty$ and $\mu \geq 0$ is known. Thus, the parameter space is $\Omega = [\mu, \infty)$. For testing $H_0 : \theta \geq \theta_0$, where $\theta_0 > \mu$, the usual $p$-value is

$$\phi_0(n) = F_{\theta_0}(n) = P_{\theta_0}\{X \leq n\}.$$

Hwang et al. (1992) suggest regarding $\phi_0$ as an estimator of the indicator function of $\Omega_0 = [\theta_0, \infty)$ and introduce the risk function

$$r(\psi; \theta) = E_\theta\big\{[\psi(X) - \mathbf{1}_{\Omega_0}(\theta)]^2\big\}$$

for estimators $\psi$ of this indicator function. Below, the phrase "admissible $p$-value" means an admissible estimator of $\mathbf{1}_{\Omega_0}$ with respect to squared error loss. If $\pi$ is a not necessarily proper prior distribution, then the integrated risk of an estimator $\psi$ with respect to $\pi$ is denoted by

$$\bar{r}(\psi;\pi) = \int_{\Omega} r(\psi;\theta)\pi\{d\theta\}, \tag{1}$$

and a $\psi_0$ is said to be *Bayes* with respect to $\pi$ if $\bar{r}(\psi_0;\pi) = \inf_{\psi} \bar{r}(\psi;\pi) = \bar{r}(\pi)$, say. It is easily seen that if $\psi_0$ is Bayes with respect to $\pi \neq 0$, and $\bar{r}(\pi) < \infty$, then $\psi_0$ is admissible. See, for example, Berger (1985), page 255. The main result of Hwang et al. provides a converse. When specialized to the present context, it asserts: *Let $\psi_0$ be an admissible estimator. Suppose that $\psi_0(n)$ is non-decreasing in $n$ and that there are $-1 \leq n_1 < n_2 \leq \infty$ for which $\psi_0(n) = 0$ for $n \leq n_1$, $0 < \psi_0(n) < 1$ for $n_1 < n < n_2$, and $\psi_0(n) = 1$ for $n \geq n_2$. Then there are sigma-finite measures $\pi_0$ and $\pi_1$ on $\Omega_0 = [\theta_0, \infty)$ and $\Omega_1 = [\mu, \theta_0]$ for which*

$$\int_{\Omega_0} f_\theta(n)\pi_0\{d\theta\} + \int_{\Omega_1} f_\theta(n)\pi_1\{d\theta\} < \infty \tag{2}$$

*and*

$$\psi_0(n) = \frac{\int_{\Omega_0} f_\theta(n)\pi_0\{d\theta\}}{\int_{\Omega_0} f_\theta(n)\pi_0\{d\theta\} + \int_{\Omega_1} f_\theta(n)\pi_1\{d\theta\}} \tag{3}$$

*for $n_1 < n < n_2$.* Hwang et al. (1992) show that the usual $p$-value for testing $\theta \leq \theta_0$ is admissible when $\mu = 0$ by showing that it is formal Bayes' with respect to $d\theta/\theta$. The modified $p$-value

$$\phi_\mu(n) = \frac{F_{\theta_0}(n)}{F_\mu(n)} = P_{\theta_0}\{X \leq n | B \leq n\} \tag{4}$$

for testing $\theta \geq \theta_0$ may be analyzed in a similar manner. Repeated use is made of the following identity below: If $G_n$ is the gamma distribution function with shape parameter $n$ and unit scale parameter, then

$$1 - G_n(\theta) = \int_\theta^\infty \frac{\omega^{n-1}}{(n-1)!} e^{-\omega} d\omega = F_\theta(n-1) \tag{5}$$

for all $n \geq 1$ and $\theta > 0$. The identity may be derived by repeated integrations by parts.

PROPOSITION 1. *For each $0 \leq \mu < \theta_0$, $\phi_\mu$ is an admissible $p$-value for the parameter space $\Omega = [\mu, \infty)$.*

PROOF. Let $\pi$ be the restriction of Lebesgue measure to $[\mu, \infty)$. Then the marginal mass function of $X$ is

$$\bar{f}(n) = \int_\mu^\infty \frac{1}{n!}\theta^n e^{-\theta} d\theta = F_\mu(n),$$

by (5), and the posterior probability of $[\theta_0, \infty)$ given $X = n$ is

$$\frac{1}{\bar{f}(n)} \int_{\theta_0}^\infty \frac{1}{n!}\theta^n e^{-\theta} d\theta = \frac{F_{\theta_0}(n)}{F_\mu(n)} = \phi_\mu(n).$$

So, $\phi_\mu$ is Bayes with respect to $\pi$. Moreover, the total Bayes risk is finite, since $\bar{r}(\pi) = \sum_{n=0}^{\infty} \phi_\mu(n)[1-\phi_\mu(n)]F_\mu(n) \le \sum_{n=0}^{\infty}[1-F_{\theta_0}(n)] \le \theta_0$.

Slightly more was proved than claimed: if $0 < \mu < \mu' < \theta_0$, then $\phi_{\mu'}$ is admissible for $\Omega = [\mu, \infty)$. However, $\phi_0$ is not admissible when $\mu > 0$.

THEOREM 1. *If* $0 < \mu < \theta_0 < \infty$, *then* $\phi_0$ *is an inadmissible p-value for testing* $H_0 : \theta \ge \theta_0$ *when* $\Omega = [\mu, \infty)$.

PROOF. It is shown that assuming admissibility of $\phi_0$ leads to a contradiction. If $\phi_0$ were admissible, then there would be sigma-finite measures $\pi_0$ and $\pi_1$ for which (2) and (3) hold for all $0 \le n < \infty$. Let $\pi = \pi_0 + \pi_1$,

$$H\{d\theta\} = \frac{e^{-\theta}\pi\{d\theta\}}{c},$$
$$H_i\{d\theta\} = \frac{e^{-\theta}\pi_i\{d\theta\}}{c},$$

$i = 0, 1$, where $c$ is so chosen that $H$ is a probability distribution. Then $H_0\{\Omega\} = \phi_0(0)H\{\Omega\} = \phi_0(0)$ from (3) with $n = 0$, and $H_1\{\Omega\} = 1 - \phi_0(0)$. Define $\tilde{H}_0$ and $\tilde{H}_1$ by $\tilde{H}_0 = H_0/\phi_0(0)$ and $\tilde{H}_1 = H_1/[1-\phi_0(0)]$. Then $\tilde{H}_0$ and $\tilde{H}_1$ are probability distributions and $H = \phi_0(0)\tilde{H}_0 + [1-\phi_0(0)]\tilde{H}_1$. With this notation,

$$n!\frac{1-\phi_0(n)}{1-\phi_0(0)}\int_\mu^\infty \theta^n H\{d\theta\} = n!\int_\mu^{\theta_0} \theta^n \tilde{H}_1\{d\theta\} \tag{6}$$

for all $n = 0, 1, 2, \dots,$ by (3). Next, let $Y,\ Z,\ \Theta_0$ and $\Theta_1$ be independent random variables for which $Y = 0$ or $1$ with probabilities $\phi_0(0)$ and $1-\phi_0(0)$, $Z$ has the standard exponential distribution, $\Theta_0 \sim \tilde{H}_0$ and $\Theta_1 \sim \tilde{H}_1$, and let $\Theta = (1-Y)\Theta_0 + Y\Theta_1$. Then $\Theta \sim H$. Now $E(Z^n) = n!$ and

$$E(Z^n|Z \le \theta_0) = \frac{1}{1-e^{-\theta_0}}\int_0^{\theta_0} z^n e^{-z}dz = n!\frac{1-\phi_0(n)}{1-\phi_0(0)}$$

for all $n$, by (5). So, (6) may be rewritten as $E(Z^n|Z \le \theta_0)E(\Theta^n) = E(Z^n)E(\Theta^n|Y=1)$ or, equivalently,

$$E[(Z\Theta)^n|Z \le \theta_0] = E[(Z\Theta)^n|Y = 1]$$

for all $n$. That is, the moments of the two conditional distributions are the same. The magnitude of these moments is at most $n!\theta_0^n$ from the right side of (6). So, the moments uniquely determine the conditional distributions and, therefore, $P\{Z\Theta \le z|Z \le \theta_0\} = P\{Z\Theta \le z|Y = 1\}$ for all $z$. See, for example, Billingsley [(1995), pages 388-389]. Since $P\{Y = 1\} = 1 - e^{-\theta_0} = P\{Z \le \theta_0\}$ and $P\{Z\Theta \le z\} = P\{Z\Theta \le z|A\}P(A) + P\{Z\Theta \le z|A'\}P(A')$ for any event $A$, it then follows that $P\{Z\Theta \le z|Z > \theta_0\} = P\{Z\Theta \le z|Y = 0\}$ for all $z$. This, however, is impossible, since the right side is positive for all $z > 0$ and the left side vanishes for $0 < z < \mu\theta_0$.

From Proposition 1, the usual $p$-value is Bayes with respect to the uniform distribution on $[0, \infty)$. So, freely mixing points of view, the following derivation of $\phi_\mu$ suggests itself. Use of $\phi_0$ is consistent with a uniform prior; if there is a known lower bound for $\theta$, then the prior distribution of $\theta$ should be conditioned to reflect this knowledge, and that leads to $\phi_\mu$. That the modified and usual $p$-values may be very different was illustrated in the Introduction.

For testing $H_0 : \theta \le \theta_0$, the usual $p$-value $\tilde{\phi}_0(n) = 1 - F_{\theta_0}(n-1)$ is inadmissible for $\Omega = [\mu, \infty)$ for any $0 < \mu < \theta_0$, by an argument similar to the proof of Theorem 1. If $\mu > 0$, then the $p$-value obtained by restricting the prior $\pi\{d\theta\} = d\theta/\theta$ to $[\mu, \infty)$ is $\tilde{\phi}_\mu(n) = 1 - F_{\theta_0}(n-1)/F_\mu(n-1)$ for $n \ge 1$ and $\tilde{\phi}_\mu(0) < 1$. This is admissible for $\Omega = [\mu, \infty)$, as is $1 - \phi_\mu$. For testing $H_0 : \theta = \theta_0$, the usual $p$-value is complicated and inadmissible, even when $\mu = 0$. See Theorem 4.3 of Hwang et al. (1992)

Of course, it does not follow from Theorem 1 that $\phi_\mu$ dominates $\phi_0$; and it does not, since $\phi_\mu(n) > \phi_0(n)$ for all $n$. The risk functions of the two estimators are compared in Figure 1 below for $b = 5$ and $\theta_0 = 2$.

**3. Confidence bounds.** Let us write $\phi_\mu^{\theta_0}(n)$ for the modified $p$-value in (4). Then formally applying the relationship between one sided tests and confidence bounds [e.g., Lehmann (1986), pages 89–95] to the modified $p$-values

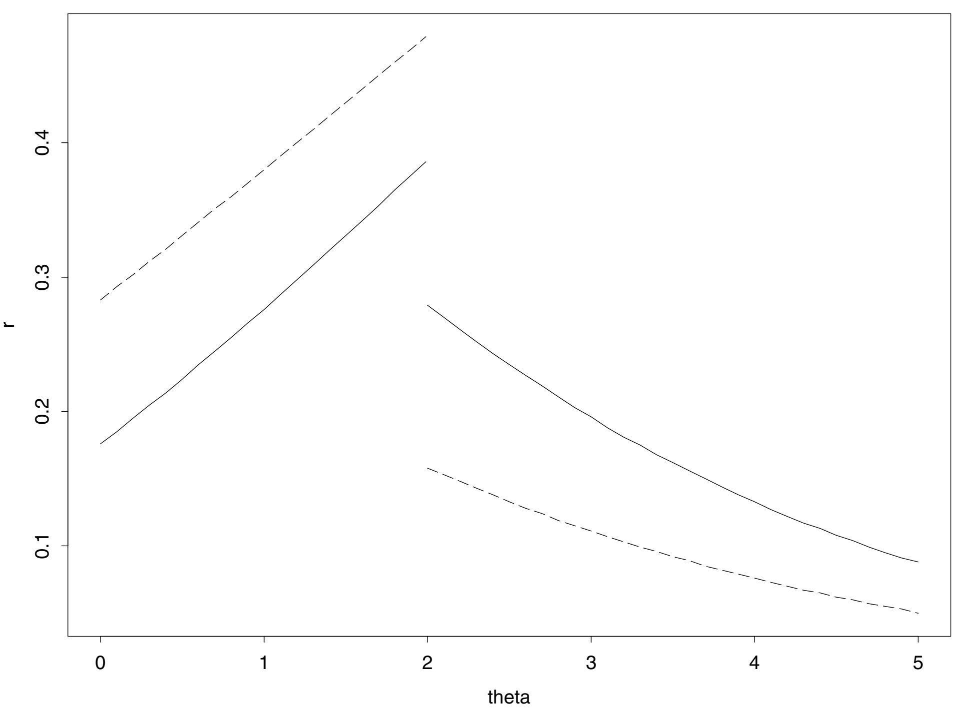


FIG. 1. *For $b = 5$ and $\theta_0 = 2$; solid line is risk of $\phi_0$; dotted line is risk of $\phi_\mu$*

leads to sets of the form $\{\theta : \phi_\mu^\theta(n) \geq \alpha\}$, where $0 < \alpha < 1$. The latter set is easily seen to be the interval $[\mu, b_n^\mu]$, where $b_n^\mu$ solves the equation $F_b(n) = \alpha F_\mu(n)$. Equivalently, $G_{n+1}(b) = 1 - \alpha F_\mu(n)$, where $G_n$ is the gamma distribution function with shape parameter $n$ and scale parameter 1, as in (5). This is essentially the argument used by Roe and Woodroofe (1999).

The confidence interval $[\mu, b_n^\mu]$ can be derived directly too. Consider a decision problem in which the action space is $\mathscr{A} = \Omega$ and the loss function is

$$L(\theta; b) = (\theta - b)_+ + \alpha b, \tag{7}$$

where $0 < \alpha < 1$ and $x_+ = \max(0, x)$. If $\xi^*$ is a prior density, then it is easily seen that $\int_\Omega L(\theta; b)\xi^*(\theta)d\theta = \int_b^\infty [1 - \Xi^*(\theta)]d\theta + \alpha b$, where $\Xi^*$ is the distribution function of $\xi^*$. The latter is minimized when $\Xi^*(b) = 1 - \alpha$, and with this choice of $b$

$$\int_\Omega L(\theta; b)\xi^*(\theta)d\theta = \int_b^\infty \theta\xi^*(\theta)d\theta.$$

This provides the solution to a no data problem. If $\xi$ is a (possibly improper) prior density and $X = n$ is observed, then the (formal) Bayes solution is obtained by letting $\xi^* = \xi_n^*$ be the posterior density. In particular, if $\xi(\theta) = 1$ for $\mu \leq \theta < \infty$, then $1 - \Xi_n^*(\theta) = F_\theta(n)/F_\mu(n)$, and the solution to the equation $1 - \Xi_n^*(b) = \alpha$ is $b_n^\mu$. That is, the confidence interval $[\mu, b_n^\mu]$ is formal Bayes with respect to the uniform distribution on $[\mu, \infty)$. The overall Bayes risk for the this problem is not finite, however, and admissibility is not clear.

Admissibility can be shown if the loss is changed to lessen the influence of large $\theta$, or equivalently large $n$. (Recall that the paper is about small $n$.) Let

$$L(\theta; n, b) = c_n L(\theta; b), \tag{8}$$

where $c_n$ are positive constants. This change does not affect Bayesian solutions to the decision problem, but can affect admissibility. The risk function and Bayes risk of a decision function $b$ for the problem (8) are defined by

$$r(b; \theta) = \sum_{n=0}^{\infty} L(\theta; n, b_n) f_\theta(n)$$

and (1) (with $\psi$ replaced by $b$).

LEMMA 1. *If $r(b; \theta) < \infty$ for all $\theta$, then $r(b; \theta)$ is continuous in $\theta$.*

PROOF. Fix a $\theta_0 \in \Omega$. To see that $r(b; \theta)$ is continuous at $\theta = \theta_0$, it suffices to show that

$$g_n := \sup_{|\theta - \theta_0| \leq 1} L(\theta; n, b_n) f_\theta(n)$$

is summable over $n \geq 0$, since continuity then follows from the Dominated Convergence Theorem. Clearly, $g_n \leq c_n[(\theta_0 + 1 - b_n)_+ + \alpha b_n] \sup_{|\theta - \theta_0| \leq 1} f_\theta(n)$ for all $n$; and this is summable, since $\sup_{|\theta-\theta_0|\le 1} f_\theta(n) \le f_{\theta_0+1}(n)$ for all $n > \theta_0+1$ and $r(b;\theta_0+1) < \infty$.

THEOREM 2. *If*

$$\sum_{n=0}^{\infty} c_n^2 < \infty,$$

*then the formal Bayes solution* $b_n^\mu$ *(with respect to the uniform distribution) is admissible with respect to the loss* (8) *when* $\Omega = [\mu, \infty)$.

PROOF. First it is shown that there are finite priors $\pi_h$, $h > 0$, for which $\bar{r}(\pi_h) < \infty$ for all $h$, $\pi_h$ converges vaguely to Lebesgue measure as $h \to 0$, and

$$\lim_{h\to 0} \bar{r}(b^\mu; \pi_h) - \bar{r}(\pi_h) = 0. \tag{9}$$

Admissibility will follow easily. Let $\pi_h$ have density $\xi_h(\theta) = e^{-h\theta}$ for $\theta \ge \mu$. Then it is easily seen that the marginal mass function of $X$ and the posterior distribution function of $\theta$ given $X = n$ are

$$\bar{f}_h(n) = \left(\frac{1}{1+h}\right)^{n+1} F_{\mu(1+h)}(n)$$

and

$$1 - \Xi_{h,n}^*(\theta) = \frac{F_{\theta(1+h)}(n)}{F_{\mu(1+h)}(n)}.$$

It follows easily that the Bayes solution for $\pi_h$ is

$$b_{h,n}^* = \frac{1}{1+h} G_{n+1}^{-1}[1 - \alpha F_{\mu(1+h)}(n)]. \tag{10}$$

This simple relation has important consequences. First, since $G_n^{-1}(u) = O(n)$ for fixed $0 < u < 1$, it follows that $b_{h,n}^* = O(n)$ as $n \to \infty$ and, therefore, that $\bar{r}(\pi_h) < \infty$ for each fixed $h > 0$. It also follows that $b_{h,n}^* \le b_n^\mu$, $b_{h,n}^* \ge b_n^\mu/(1+h)$ and, therefore, that $b_{h,n}^* \le b_n^\mu \le (1+h) b_{h,n}^*$.

To verify the crucial relation (9), write

$$\begin{aligned}
\int_\Omega L(\theta; b_n^\mu)\xi_{h,n}^*(\theta)d\theta &= \int_{b_n^\mu}^\infty (\theta - b_n^\mu)\xi_{h,n}^*(\theta)d\theta + \alpha b_n^\mu \\
&= \int_{b_{h,n}^*}^\infty (\theta - b_n^\mu)\xi_{h,n}^*(\theta)d\theta + \int_{b_{h,n}^*}^{b_n^\mu} (b_n^\mu - \theta)\xi_{h,n}^*(\theta)d\theta + \alpha b_n^\mu \\
&= \int_{b_{h,n}^*}^\infty \theta\xi_{h,n}^*(\theta)d\theta + \int_{b_{h,n}^*}^{b_n^\mu} (b_n^\mu - \theta)\xi_{h,n}^*(\theta)d\theta.
\end{aligned}$$

Thus,

$$\bar{r}(b^{\mu};\pi_h)-\bar{r}(\pi_h)=\sum_{n=1}^{\infty}c_n\int_{b^*_{h,n}}^{b^{\mu}_n}(b^{\mu}_n-\theta)\xi^*_{h,n}(\theta)d\theta\times\bar{f}_h(n)$$
$$=\sum_{n=1}^{\infty}c_n\int_{b^*_{h,n}}^{b^{\mu}_n}(b^{\mu}_n-\theta)\frac{1}{n!}\theta^n\exp\{-(1+h)\theta\}d\theta$$

Since $b^*_{h,n}\to b^{\mu}_n$ as $h\downarrow 0$ for fixed $n$, it is clear that

$$\lim_{h\to 0}\sum_{n=1}^{m}c_n\int_{b^*_{h,n}}^{b^{\mu}_n}(b^{\mu}_n-\theta)\frac{1}{n!}\theta^n\exp\{-(1+h)\theta\}d\theta=0$$

for each fixed $m$. Moreover, for all large $n$ and small $h$, $\sup_{\theta}\theta^n e^{-\theta}/n!\le 1/\sqrt{n}$, $0\le b^{\mu}_n-b^*_{h,n}\le 2hn$, and $\min[b^{\mu}_n,b^*_{h,n}]\ge n/2$. So,

$$\sum_{n=m+1}^{\infty}c_n\int_{b^*_{h,n}}^{b^{\mu}_n}(b^{\mu}_n-\theta)\frac{1}{n!}\theta^n\exp\{-(1+h)\theta\}d\theta$$
$$\le\sum_{n=m+1}^{\infty}c_n(b^{\mu}_n-b^*_{h,n})^2\times\frac{1}{\sqrt{n}}\exp\left\{-\frac{1}{2}hn\right\}$$
$$\le\sqrt{\sum_{n=m+1}^{\infty}c_n^2}\times\sqrt{(2h)^4\sum_{n=1}^{\infty}n^3\exp\{-hn\}}$$

for all sufficiently large $m$ and small $h$. The second factor on the right is independent of $m\ge 1$ and bounded in $0<h<1$; the first is independent of $h$ and approaches 0 as $m\to\infty$. Relation (9) follows.

To complete the proof, suppose that $b^{\mu}$ were inadmissible. Then there would be a $b$ for which $r(b;\theta)\le r(b^{\mu};\theta)$ for all $\theta$ with strict inequality for some $\theta$, say $\theta=\theta_0$. Since $r(b^{\mu};\theta)$ is everywhere finite, $r(b;\theta)$ and $r(b^{\mu};\theta)$ are both continuous in $\theta$. Thus, there would be $\varepsilon>0$ and $\delta>0$ for which $r(b;\theta)\le r(b^{\mu};\theta)-\varepsilon$ for $|\theta-\theta_0|\le\delta$, and this is incompatible with (9). □

As with $p$-values, $b^{\mu}_n$ is obtained from $b^0_n$, by conditioning the prior distribution to account from the known bound, and the two can be quite different.

**4. Remarks.** As of Summer 1999, the KARMEN 2 Group had expected to see 7.8 background events and had seen 8 events total. Based on this data alone, it is still impossible to confirm or deny the existence of a signal.

There is a weak connection between the main results here and the functional model of Dawid and Stone (1982). If the signal $S$ of the Introduction is represented as $S=F^{\#}_{\nu}(U)$, where $F^{\#}_{\nu}$ denotes the Poisson quantile function, and $U$ is a uniformly distributed random variable that is independent of $B$, then $X=B+S$ is a function of $\theta=\mu+\nu$ and $E=(B,U)$. This is an instance of the functional model. For a given $n\ge 0$, the set of $E$ for which $X=n$ for some $\theta \geq \mu$ is the set where $B \leq n$. Dawid and Stone (1982) suggest conditioning on observed constraints in their own context but require a condition called partitionability that is not satisfied in our model.

It is likely that the inadmissibility results of Section 2 have extensions to other one-parameter exponential families, when the natural parameter space is restricted. Different mathematical arguments may be necessary, however, because the use of the moment problem is special to the discrete case. For example, if $X$ is normally distributed with mean $\theta$ and variance one, where $\theta \leq 1$, and if $H_0 : \theta \leq 0$, then the usual $p$-value $\phi(x) = P_0\{X \geq x\}$ is easily seen to be inadmissible by a simple argument using Theorem 3.3 of Hwang et al (1992) and analysis of $\phi(x)$ as $x \to \infty$. We have not pursued such extensions here, because our interest is in examining the suggestion of Roe and Woodroofe (1999).

The $c_n$ in Section 3 may seem a mystery. A more familiar way of accomplishing the same objective is to divide the loss by a function of $\theta$, for example, forming relative square error loss to keep the minimax risk bounded. That is possible in our context too, but leads to slightly different answers.

**Acknowledgment.** Thanks to Byron Roe for helpful conversations.

## REFERENCES

BERGER, J. (1985). *Statistical Decision Theory and Bayesian Analysis*, 2nd ed. Springer, New York.

BILLINGSLEY, P. (1995). *Probability and Measure*, 3rd ed. Wiley, New York.

COUSINS, R. (2000). Application of conditioning to the Gaussian with boundary problem in the unified approach to confidence intervals. *Phys. Rev. D*. To appear.

DAWID, A. P. and STONE, M. (1982). The functional model basis of fiducial inference (with discussion). *Ann. Statist.* **10** 1054–1067.

FELDMAN, G. and COUSINS, R. (1998). Unified approach to the classical statistical analysis of small signals. *Phys. Rev. D* **57** 3873–3889.

HWANG, J. T., CASELLA, G., ROBERT, C. WELLS, M. and FARRELL, R. (1992). Estimation of accuracy in testing. *Ann. Statist.* **20** 490–509.

LEHMANN, E. (1986). *Testing Statistical Hypotheses*, 2nd ed. Wadsworth, Belmont, CA.

READ, A. L. (2000). Modified frequentist analysis of search methods (the $CL_s$ method). Unpublished manuscript.

THE PARTICLE DATA GROUP (1992). Probability, statistics, and Monte Carlo. *Phys. Rev. D* **45** III.32-III.42.

ROE, B. P. and WOODROOFE, M. (1999). Improved probability method for estimating signal in the presence of background. *Phys. Rev. D* **60** 3009–3015.

ZEITNITZ, B. et al. (1998). Neutrino oscillation results from KARMEN. *Progress in Particle and Nuclear Physics* **40** 169–181.

DEPARTMENT OF MATHEMATICS AND STATISTICS
UNIVERSITY OF MICHIGAN
ANN ARBOR, MICHIGAN 48109
E-MAIL: michaelw@umich.edu

INSTITUTE OF STATISTICAL SCIENCE
ACADEMIA SINICA
TAIPEI
TAIWAN
E-MAIL: hywang@stat.sinica.edu.tw