Confidence Bounds for Extreme Quantiles

By

Leo Breiman
Charles J. Stone
and
Charles Kooperberg*

Technical Report No. 167
August 1988

Research supported by ONR under contract N00014-82-K-0054
and NSF Grant DMS-8600409.

Department of Statistics
University of California
Berkeley, California

Confidence Bounds for Extreme Quantiles

LEO BREIMAN, CHARLES J. STONE, and CHARLES KOOPERBERG*

Let y_p be the upper pth quantile of the distribution of a random variable Y, so that $\Pr(Y \geq y_p) = p$. We consider four related methods for obtaining confidence bounds for y_p when p is very small ($p \approx 0$) that are obtained by first fitting a parametric model to the m upper order statistics based on a random sample of size $n > m$ from the distribution of Y. The exponential-tail (ET) method corresponds to the assumption that the upper tail of the distribution of Y is approximately exponential or, equivalently, that y_p is approximately linear in $\log(1/p)$ for $p \approx 0$. The quadratic-tail (QT) method corresponds to the assumption that y_p is approximately quadratic in $\log(1/p)$ for $p \approx 0$. Associated with these two methods are two other methods, ETP and QTP, which involve the use of a preliminary power transformation to make the upper tail more nearly exponential. We also consider the multisample problem, in which the tails of the distributions corresponding to various samples are assumed to have approximately the same shape. When the ETP and QTP methods are applied to the multisample problem, a common power transformation is made to all samples. The confidence bounds we obtain depend on a parameter t that must be adjusted to yield a given nominal coverage probability. We make this adjustment by adaption, via simulation, to the exponential distribution. An extensive simulated study is described, which compares the performance of 90% upper confidence bounds corresponding to the four methods over a wide range of distributions "centered at the exponential;" that is, which are neither too heavy tailed nor too light tailed.

KEY WORDS: Quantile estimation; Exponential-tail model; Quadratic-tail model; Power transformation; Adaption; Tail heaviness.

---

*This research has been supported by the Office of Naval Research under contract N00014-82-K-0054 and by National Science Foundation Grant DMS-8600409.

1. INTRODUCTION

Often, in applications, we want to estimate extreme quantiles from sample data. For example, we might have 30 years of annual high-water levels on a river and want to estimate the 100-year flood level $y_{.01}$, defined by the requirement that the probability of annual high-water level exceeding $y_{.01}$ should be .01.

As usual, it is important that the estimates we obtain be accompanied by some indication of accuracy. The distributions of estimates of extreme quantiles are far from normal; in particular, they are typically quite skewed. Thus "standard errors" are inappropriate measures of accuracy. Much better are confidence intervals or, equivalently, upper and lower confidence bounds. Since a $100(1-c)\%$ lower confidence bound is a $100c\%$ upper confidence bound (UCB), we can restrict our attention to UCB's.

Let Y be a random variable (whose distribution function is continuous) and let y_p denote the *upper* pth quantile of Y for $0 < p < 1$; so that $\Pr(Y \geq y_p) = p$. Let n be a positive integer and let $Y_1, \cdots, Y_n$ be a random sample of size n from the distribution of Y. Then $Y_1, \cdots, Y_n$ are independent and identically distributed random variables. Let $Y_{(1)}, \cdots, Y_{(n)}$ denote the corresponding upper order statistics, obtained by writing $Y_1, \cdots, Y_n$ in decreasing order; thus $Y_{(1)} \geq \cdots \geq Y_{(n)}$.

Let U be a statistic based on the random sample, which is thought of as an upper confidence bound (UCB) for y_p. The corresponding coverage probability is $\Pr(y_p \leq U)$. Let $0 < c < 1$. If U is derived as a $100c\%$ UCB for y_p by making various assumptions and approximations, then we refer to c as the nominal coverage probability of U and to $\Pr(y_p \leq U)$ as its actual coverage probability.

Consider, for example, the maximum value, $Y_{(1)}$, in the sample as an UCB for y_p. Since

$$\Pr(Y_{(1)} < y_p) = \Pr(Y_1 < y_p, \cdots, Y_n < y_p) = [\Pr(Y_1 < y_p)]^n = (1-p)^n,$$

we see that the actual coverage probability of $Y_{(1)}$ is given by

$$\Pr(Y_{(1)} \geq y_p) = 1 - (1-p)^n.$$

In particular, $Y_{(1)}$ is a 90% UCB for y_p if and only if

$$p = 1 - (.1)^{1/n} = 1 - \exp(-\log(10)/n) \approx \frac{\log(10)}{n} \approx \frac{2.3}{n}, \quad n \gg 1.$$

Thus (for $n \geq 8$) if $p \leq 2/n$, there is no order statistic that serves as a 90% UCB for y_p.

When $p \leq 2/n$, we can obtain a nominal 90% UCB for y_p in a standard manner by assuming a Weibull, gamma, lognormal or other classical parametric model for the distribution of Y. But if our assumption is even mildly inaccurate in a given application, the actual coverage probability can differ substantially from .9. In other words, the actual coverage probability of the nominal 90% UCB for an extreme quantile is very sensitive to model departures.

It is better to fit a more flexible parametric model such as the logspline model discussed in Stone and Koo (1986). But we will not pursue this approach in the present paper.

Another approach is to obtain a nominal 90% UCB for y_p by first fitting a parametric model to the upper tail of the data; that is, to the m upper order statistics $Y_{(1)}, \cdots, Y_{(m)}$, where $m < n$. This is the approach that will be followed here.

In Section 2 we describe several such methods for obtaining confidence bounds for extreme quantiles. The well-known exponential-tail method is described in Section 2.1. The quadratic-tail method, briefly described in Section 2.2, was introduced in Breiman, et al. (1981), which is a precursor to the present work. Further details for this method are given in Section 5 and Appendix A. The preliminary power transformation is discussed in Section 2.3; adaption of the parameter t in a confidence bound to the exponential distribution in Section 2.4; and the multisample problem in Section 2.5. In Section 4 we discuss a reasonably extensive simulation study of 90% UCB's, in which four specific methods are compared when the actual distribution of Y is Weibull, generalized gamma, or lognormal. The power parameter of each of these distributions is chosen so that the corresponding tail heaviness, as defined in Section 3, ranges from –.2 to .4 (the tail heaviness of an exponential distribution is zero). The results of the simulation study are presented in graphical form in Section 4.4.

We are unaware of previous work on confidence bounds for extreme quantiles (other than Breiman, et al. (1981)). But there have been many studies of exponential-tail and related methods of estimation for tail probabilities and extreme quantiles. These (mainly theoretical) studies have focussed on methods that are appropriate when the tail is (I) in the domain of attraction of some extreme-value distribution; (II) approximately algebraically decreasing; or (III) approximately exponentially decreasing. These three conditions are very closely related. For example, the upper tail of Y is approximately algebraically decreasing if and only if that of $\log(Y)$ is approximately exponentially decreasing; so methods appropriate to approximately exponentially decreasing tails may be applied to data having approximately algebraically decreasing tails by first taking logs. In category (I) are Maritz and Munro (1967), Pickands (1975), Weissman (1978), Boos (1984), Davis and Resnick (1984), and Smith (1987); in category (II) are Hill (1975), DuMouchel and Olshen (1975), DuMouchel (1983), Hall and Welsh (1985), and Csörgő, et al. (1985); and in category (III) are Breiman, et al. (1978, 1979, and 1981) and Crager (1982). See Smith (1987) for a recent and thorough review of this literature.

2. CONFIDENCE BOUNDS

2.1 Exponential-tail Model

Let $0 < p_0 < 1$. Consider the *exponential-tail model*, in which there is an $\alpha > 0$ such that

$$\Pr(Y \geq y \mid Y \geq y_{p_0}) = \exp(-(y-y_{p_0})/\alpha) \tag{2.1}$$

for $y \geq y_{p_0}$ or, equivalently, in which y_p is a linear function of $\log(1/p)$ as p ranges over $(0, p_0]$. Let m be a positive integer with $m/n \leq p_0$. Then

$$y_p = y_{m/n} + \alpha \log\left[\frac{m}{np}\right]$$

for $0 < p \leq p_0$. It is reasonable to estimate $y_{m/n}$ by $Y_{(m)}$ and to estimate α by

$$\hat{\alpha} = \frac{1}{m-1} \sum_{i=1}^{m-1} [Y_{(i)} - Y_{(m)}]. \tag{2.2}$$

This leads to the quantile estimate

$$\hat{y}_p = Y_{(m)} + \hat{\alpha} \log\left[\frac{m}{np}\right] \tag{2.3}$$

for $0 < p \leq p_0$.

Suppose that Y has a (two-parameter) exponential distribution or, equivalently, that y_p is a linear function of $\log(1/p)$ as p ranges over $(0, 1)$. Let $\hat{\alpha}$ and $\hat{y}_p$ be given by (2.2) and (2.3) respectively. Then the constant $t = t_{p,c,n,m}$ can be obtained numerically from the incomplete beta function so that $U = Y_{(m)} + t\hat{\alpha}$ is an exact $100c\%$ UCB for y_p. We refer to U as the exponential-tail $100c\%$ UCB for y_p.

Under the more general exponential-tail model, the actual coverage probability of the exponential-tail UCB is close to its nominal coverage probability if $\Pr(Y_{(m)} \geq y_{p_0}) \approx 1$. But it is more realistic to consider the exponential-tail model as being a reasonably accurate approximation. Hopefully, the actual coverage probability of the exponential-tail UCB will be close to its nominal coverage probability if p is not too small. But if p is extremely small, then the actual and nominal coverage probabilities may well be considerably different.

2.2 Quadratic-tail Model

Let $0 < p_0 < 1$. In the corresponding exponential-tail model, y_p is a linear function of $\log(1/p)$ as p ranges over $(0, p_0]$. In order to obtain a more accurate approximation, we consider the *quadratic-tail model*, in which y_p is assumed to be a quadratic function of $\log(1/p)$ as p ranges over $(0, p_0]$. Let m be a positive integer with $m/n \le p_0$. Then

$$y_p = y_{m/n} + \alpha[\log(1/p) - \log(n/m)] + \frac{\beta}{2}[\log^2(1/p) - \log^2(n/m)], \quad 0 < p \le p_0. \quad (2.4).$$

Here α and β are unknown parameters with $\alpha \ge 0$. The quadratic model can be exactly valid if $\beta > 0$ or if $\alpha > 0$ and $\beta = 0$. It cannot be exactly valid when $\beta < 0$; for, in that case, the quadratic function in (2.4) tends to $-\infty$ as $p \to 0$. But, even when $\beta < 0$, the quadratic-tail model can provide a good approximation to y_p for values of p that are not too close to zero.

Given m and p, set

$$L = \log(1/p) - \log(n/m) \quad \text{and} \quad M = \frac{1}{2}[\log^2(1/p) - \log^2(n/m)]. \tag{2.5}$$

It follows from (2.4) that $y_p = y_{m/n} + \tau$, where $\tau = L\alpha + M\beta$. Corresponding to an estimate $\hat{\tau}$ of τ is the estimate $\hat{y}_p = Y_{(m)} + \hat{\tau}$ of y_p.

More generally, let L and M be arbitary constants and set $\tau = L\alpha + M\beta$. Consider an estimate $\hat{\tau}$ of τ of the form

$$\hat{\tau} = \sum_{i=1}^{m-1} w_i[Y_{(i)} - Y_{(i+1)}]. \tag{2.6}$$

It is shown in Section 5 and Appendix A that

$$\mathrm{Var}(\hat{\tau}) = c_1\alpha^2 + c_2\alpha\beta + c_3\beta^2, \tag{2.7}$$

where c_1, c_2 and c_3 are given explicitly in terms of L, M and the weights $w_1, \cdots, w_{m-1}$. If the exponential-tail model is reasonably accurate and, in particular, if $\beta \approx 0$, then

$$\mathrm{Var}(\hat{\tau}) \approx c_1\alpha^2. \tag{2.8}$$

In light of (2.8) it is reasonable to choose the weights to minimize c_1 subject to the constraint that $\hat{\tau}$ be unbiased; that is, that, for all values of L and M,

$$E\hat{\tau} = L\alpha + M\beta. \tag{2.9}$$

It is shown in Section 5 that this minimization problem has a unique solution, which is given explicitly. (The quadratic model should be thought of as an approximation. In Section 5, the error of approximation is ignored. Thus (2.7) and the solution to the indicated minimization problem should be thought of as informal approximations. The minimization problem itself is reasonable, since it is not possible to choose the weights to minimize Var($\hat{\tau}$) for all values of α and β.)

In particular, by choosing $L = 1$ and $M = 0$, we obtain an unbiased estimate of α having the form

$$\hat{\alpha} = \sum_{i=1}^{m-1} w_{1i}[Y_{(i)} - Y_{(i+1)}];$$

and by choosing $L = 0$ and $M = 1$, we obtain an unbiased estimate of β having the form

$$\hat{\beta} = \sum_{i=1}^{m-1} w_{2i}[Y_{(i)} - Y_{(i+1)}].$$

As shown in Section 5, for arbitrary values of L and M the unbiased estimate of $\tau = L\alpha + M\beta$ for which c_1 is minimized is given by $\hat{\tau} = L\hat{\alpha} + M\hat{\beta}$; so the corresponding quantile estimate is given by

$$\hat{y}_p = Y_{(m)} + L\hat{\alpha} + M\hat{\beta} \tag{2.10}$$

for $0 < p \le p_0$.

It is shown in Section 5 and Appendix A that

$$\mathrm{Var}(\hat{y}_p) = C_1\alpha^2 + C_2\alpha\beta + C_3\beta^2, \tag{2.11}$$

where C_1, C_2 and C_3 are given explicitly in terms of n, m, L, and M. The corresponding standard error is given by

$$\mathrm{SE}(\hat{y}_p) = (C_1\hat{\alpha}^2 + C_2\hat{\alpha}\hat{\beta} + C_3\hat{\beta}^2)^{1/2}.$$

Presumably, under suitable conditions,

$$\mathrm{Dist}\left[\frac{\hat{y}_p - y_p}{\mathrm{SE}(\hat{y}_p)}\right] \approx \mathrm{N}(0,\ 1),$$

in which case $\hat{y}_p + z_{1-c}\mathrm{SE}(\hat{y}_p)$ is an approximate $100c\%$ UCB for y_p; here $\Pr(Z \ge z_{1-c}) = 1-c$, where Z has the standard normal distribution.

2.3 Preliminary Power Transformation

Suppose that Y is a positive random variable. The approximation errors of the exponential-tail and quadratic-tail models can be substantially reduced by a preliminary power transformation. Given a positive constant γ, set $W = Y^\gamma$ and $W_i = Y_i^\gamma$ for $1 \le i \le n$. The upper pth quantile of W is given by $w_p = y_p^\gamma$. Let $\hat{w}_p$ be an estimator of w_p based on the random sample $W_1, \cdots, W_n$. By applying the inverse power transformation, we obtain the estimate $\hat{y}_p = \hat{w}_p^{1/\gamma}$ of y_p based on the original random sample. Similarly, let

$\hat{w}_p + t\,\mathrm{SE}(\hat{w}_p)$ be an approximate 100c% UCB for w_p. Then $[\hat{w}_p + t\,\mathrm{SE}(\hat{w}_p)]^{1/\gamma}$ is an approximate 100c% UCB for y_p.

Let $0 < p_0 < 1$. We would like to choose γ so that the conditional distribution of $W - w_{p_0}$ given that $W \geq w_{p_0}$ is approximately exponential. Note that if V has an exponential distribution, then $E(V^2) = 2(EV)^2$. This suggests choosing $\gamma > 0$ to satisfy

$$\frac{E[(Y^{\gamma} - y_{p_0}^{\gamma})^2 \mid Y \geq y_{p_0}]}{\left[E[Y^{\gamma} - y_{p_0}^{\gamma} \mid Y \geq y_{p_0}]\right]^2} = 2. \tag{2.12}$$

In practice, the power transformation must be determined from the random sample; so we denote the corresponding parameter by $\hat{\gamma}$. We are led to $\hat{y}_p = \hat{w}_p^{1/\hat{\gamma}}$ as an estimate of y_p and to $[\hat{w}_p + t\,\mathrm{SE}(\hat{w}_p)]^{1/\hat{\gamma}}$ as an approximate 100c% UCB for y_p. Let $p_0 = m/n$, where $2 \leq m \leq n$. The obvious sample version of (2.12) is to choose $\hat{\gamma} > 0$ to satisfy

$$\frac{\frac{1}{m-1}\sum_{i=1}^{m-1}[Y_{(i)}^{\hat{\gamma}} - Y_{(m)}^{\hat{\gamma}}]^2}{\left[\frac{1}{m-1}\sum_{i=1}^{m-1}[Y_{(i)}^{\hat{\gamma}} - Y_{(m)}^{\hat{\gamma}}]\right]^2} = 2. \tag{2.13}$$

For a refinement of (2.13), let $Z_1, \cdots, Z_n$ be independent and identically distributed exponential random variables and let $Z_{(1)}, \cdots, Z_{(n)}$ be the corresponding decreasing order statistics. It is shown in Appendix B that

$$\frac{m}{m-1} E\left[\frac{\frac{1}{m-1}\sum_{i=1}^{m-1}[Z_{(i)} - Z_{(m)}]^2}{\left[\frac{1}{m-1}\sum_{i=1}^{m-1}[Z_{(i)} - Z_{(m)}]\right]^2}\right] = 2. \tag{2.14}$$

This suggests choosing $\hat{\gamma} > 0$ to satisfy

$$\frac{m}{m-1}\left[\frac{\frac{1}{m-1}\sum_{i=1}^{m-1}[Y_{(i)}^{\hat{\gamma}} - Y_{(m)}^{\hat{\gamma}}]^2}{\left[\frac{1}{m-1}\sum_{i=1}^{m-1}[Y_{(i)}^{\hat{\gamma}} - Y_{(m)}^{\hat{\gamma}}]\right]^2}\right] = 2. \tag{2.15}$$

Suppose that $m \geq 3$ and that $Y_{(1)} > \cdots > Y_{(m)} > 0$. Then the left side of (2.15) is a continuous function of $\hat{\gamma} \in (0, \infty)$, which has limit

$$\hat{A} = \frac{m}{m-1}\left[\frac{\frac{1}{m-1}\sum_{i=1}^{m-1}[\log(Y_{(i)})-\log(Y_{(m)})]^2}{\left[\frac{1}{m-1}\sum_{i=1}^{m-1}[\log(Y_{(i)})-\log(Y_{(m)})]\right]^2}\right]$$

as $\hat{\gamma} \to 0$ and limit m as $\hat{\gamma} \to \infty$; and, as shown in Appendix B, it is a strictly increasing function of $\hat{\gamma}$. Thus there is at most one value of $\hat{\gamma} \in (0, \infty)$ that satisfies (2.15); and there is such a value if and only if $\hat{A} < 2$. It is an elementary numerical computation to determine $\hat{\gamma}$ when it exists. (When $\hat{\gamma}$ fails to exist, it is reasonable to consider the logarithmic transformation: $W = \log(Y)$ and $W_i = \log(Y_i)$, $1 \leq i \leq n$.)

The use of (2.12)–(2.15) to select the parameter of the power transformation was suggested by the test for exponentiality due to Shapiro and Wilk (1972). The use of a power transformation before applying the ET method was suggested by a similar estimator in Weinstein (1973).

2.4 Adaption to the Exponential Distribution

Consider a confidence bound $U(t)$ for y_p that involves a constant t in its definition, where t is to be chosen to yield the coverage probability c. Usually this is done by means of an appeal to some central limit theorem to justify normal approximation. We have found that in the context of obtaining confidence bounds for quantiles, the resulting confidence bounds are very unreliable; that is, that the actual coverage probability can differ substantially from c even when Y has an exponential distribution.

At least for distributions of Y similar to those in the simulation study discussed in Section 4, it is more reliable to choose t by adaption to the exponential distribution; that is, so that $\Pr(y_p \leq U(t)) = c$ when Y has an exponential distribution. In practice this must be done by Monte Carlo simulation (as in the implementation of bootstrap methods of obtaining confidence bounds).

Consider, for example, a confidence bound for y_p of the form

$$(\hat{w}_p + t\,\mathrm{SE}(\hat{w}_p))^{1/\hat{\gamma}}.$$

Let t be chosen so that, when Y has an exponential distribution,

$$\Pr\left[y_p \le [\hat{w}_p + t\,\mathrm{SE}(\hat{w}_p)]^{1/\hat{\gamma}}\right] = c$$

or, equivalently, so that

$$\Pr\left[\frac{\hat{w}_p - y_p^{\hat{\gamma}}}{\mathrm{SE}(\hat{w}_p)} \ge -t\right] = c.$$

Then $-t$ is the upper cth quantile of the distribution of

$$\frac{\hat{w}_p - y_p^{\hat{\gamma}}}{\mathrm{SE}(\hat{w}_p)}$$

when Y has an exponential distribution; so t is easily found by Monte Carlo simulation.

2.5 Multisample Problem

For $1 \le k \le K$, let Y_k be a positive random variable having a density f_k that is continuous and positive on $(0, \infty)$ but otherwise unknown. Let y_{kp} denote the upper pth quantile of Y_k. Let n_k be a positive integer; let $Y_{k1}, \cdots, Y_{kn_k}$ be a random sample of size n_k from the distribution of Y_k; and let $Y_{k(1)}, \cdots, Y_{k(n_k)}$ denote the corresponding decreasing order statistics. It is assumed that the $n_1 + \cdots + n_K$ random variables obtained by combining the K random samples are independent. We can obtain separate estimates and confidence intervals for the quantiles y_{kp}, $1 \le k \le K$.

In many practical applications, the upper tails of the distributions of $Y_1, \cdots, Y_K$ have approximately the same shape. If so, it is reasonable to use a common power transformation for the K samples. To this end, we choose positive integers m_k, $1 \le k \le K$, such that $2 \le m_k \le n_k$ for $1 \le k \le K$ and then determine $\hat{\gamma}$ as the unique positive number such that

$$\sum_{k=1}^{K} \frac{m_k}{m_k - 1} \left[\frac{\frac{1}{m_k - 1} \sum_{i=1}^{m_k - 1} [Y_{k(i)}^{\hat{\gamma}} - Y_{k(m_k)}^{\hat{\gamma}}]^2}{\left[\frac{1}{m_k - 1} \sum_{i=1}^{m_k - 1} [Y_{k(i)}^{\hat{\gamma}} - Y_{k(m_k)}^{\hat{\gamma}}] \right]^2} \right] = 2K. \tag{2.16}$$

3. TAIL HEAVINESS

Let G denote the tail distribution function of Y, which is given by $G(y) = \Pr(Y \geq y)$, $y \in \mathbb{R}$; and let G^{-1} be the inverse function to G, which is assumed to be continuous and strictly decreasing on $(0, 1)$. Then $y_p = G^{-1}(p)$ for $0 < p < 1$. In this section, it is assumed, in addition, that Y has a density that is positive and continuously differentiable on the range of G^{-1}.

Let $0 < p < 1$. The *tail heaviness* of the distribution of Y at y_p is defined by

$$H(p) = H_Y(p) = \frac{d^2 y_p}{d(\log(1/p))^2} \Bigg/ \frac{dy_p}{d(\log(1/p))}. \tag{3.1}$$

The tail heaviness is invariant under location and scale transformations; that is,

$$H_{a+bY}(p) = H_Y(p), \quad a \in \mathbb{R} \text{ and } b > 0.$$

The effects of power and logarithmic transformations on the tail heaviness of a positive random variable Y are given by

$$H_{Y^b}(p) = H_Y(p) + \frac{(b-1)}{y_p} \frac{dy_p}{d(\log(1/p))}, \quad b > 0,$$

and

$$H_{\log(Y)}(p) = H_Y(p) - \frac{1}{y_p} \frac{dy_p}{d(\log(1/p))}.$$

An exponential random variable has tail heaviness zero, since its pth quantile is a constant multiple of $\log(1/p)$. A random variable is said to be *heavy-tailed* if its tail heaviness is positive and *light-tailed* if its tail heaviness is negative.

It follows from (3.1) and elementary calculus that

$$H(p) = -p\left[\frac{d^2 y_p}{dp^2} \Bigg/ \frac{dy_p}{dp}\right] - 1 = \frac{pG''(y_p)}{[G'(y_p)]^2} - 1. \tag{3.2}$$

Suppose, for example, that Y has a Weibull distribution with positive power parameter β; so that Y has the same distribution as W^β, where W has an exponential distribution. Then $y_p = \alpha\,[\log(1/p)]^\beta$ for some positive constant α, which is a scale parameter. By (3.1), the tail heaviness of Y at y_p is given by

$$H(p) = \frac{\beta-1}{\log(1/p)}.$$

When $\beta = 1$, Y is exponentially distributed and hence it has tail heaviness zero; when $\beta > 1$, it is heavy-tailed; and when $0 < \beta < 1$, it is light-tailed. For fixed $\beta \neq 1$, the tail heaviness converges very slowly to zero as $p \to 0$.

The tail heaviness corresponding to the quadratic-tail model is given by

$$H(p) = \frac{\beta}{\alpha + \beta \log(1/p)}$$

for $0 < p \leq p_0$. If α and β are positive, the tail heaviness is positive; and it converges to zero as $p \to 0$ at the same rate as for heavy-tailed Weibull distributions. If α is positive and β is negative, however, the tail heaviness is negative for $p > \exp(\alpha/\beta)$ and it decreases to $-\infty$ as $p \to \exp(\alpha/\beta)$. This is another indication that the quadratic-tail model is unrealistic for sufficiently extreme quantiles of light-tailed distributions.

Suppose next that Y has a lognormal distribution; so that $\log(Y)$ is normally distributed. Then $y_p = \alpha \exp(\beta z_p)$ and

$$G(y) = 1 - \Phi\left[\frac{\log(y/\alpha)}{\beta}\right], \quad y > 0.$$

Here Φ denotes the standard normal distribution function, whose density is denoted by φ; and the scale parameter α and power parameter β are both positive. The random variable $\log(Y)$ has mean $\log(\alpha)$ and standard deviation β. According to (3.2), the tail heaviness of Y at y_p is given by

$$H(p) = \frac{p(z_p+\beta)}{\varphi(z_p)} - 1.$$

It follows by straightforward asymptotics that

$$\lim_{p \to 0} H(p)\sqrt{\log(1/p)} = \frac{\beta}{\sqrt{2}}.$$

Thus the tail heaviness is positive for p sufficiently close to zero and it converges extremely slowly to zero as $p \to 0$.

Suppose, finally, that Y has a Pareto distribution; so that $y_p = \alpha \exp(\beta \log(1/p))$ for $0 < p < 1$, where the scale parameter α and power parameter β are both positive. By (3.1), $H(p) = \beta$ for $0 < p < 1$. In particular, Y is heavy-tailed and its tail heaviness fails to converge to zero as $p \to 0$.

4. SIMULATION STUDY

4.1 Distributions

The simulation study involves Weibull, generalized gamma(5), and lognormal distributions. In each case, the power parameter β takes on values corresponding to seven values, $-.2$, $-.1$, 0, .1, .2, .3, and .4, of the tail heaviness, $H(.1)$, at the upper decile. For the Weibull distributions the seven values of β are (approximately) .54, .77, 1.00, 1.23, 1.46, 1.69, and 1.92. For the lognormal distributions the seven values of β are .12, .30, .47, .65, .82, 1.00, and 1.18.

The generalized gamma(5) distributions considered are distributions of W^{β}, where W is distributed as the sum of five independent and identically distributed exponential random variables (so that W has a gamma distribution). The seven values of β are .68, 1.14, 1.60, 2.06, 2.52, 2.97, and 3.43.

Table 1. Selected Quantiles of Simulated Distributions

	$H(.1)$						
p	$-.2$	$-.1$	0	.1	.2	.3	.4
	Weibull distributions						
.2	1.6	1.9	2.3	2.8	3.4	4.2	5.0
.02	2.5	3.8	5.6	8.4	12.5	18.7	27.8
.002	3.3	5.4	9.0	14.9	24.6	40.8	67.6
.0002	3.9	6.9	12.3	21.9	39.0	69.5	123.9
.00002	4.4	8.3	15.6	29.4	55.3	104.2	196.1
	Generalized gamma(5) distributions						
.2	1.3	1.5	1.8	2.1	2.5	3.0	3.5
.02	1.7	2.5	3.7	5.4	7.8	11.4	16.6
.002	2.1	3.4	5.7	9.4	15.4	25.4	41.9
.0002	2.4	4.3	7.8	14.1	25.4	45.8	82.7
.00002	2.7	5.2	10.0	19.5	37.8	73.4	142.4
	Lognormal distributions						
.2	1.1	1.3	1.5	1.7	2.0	2.3	2.7
.02	1.3	1.8	2.6	3.8	5.4	7.8	11.2
.002	1.4	2.4	3.9	6.5	10.7	17.8	29.5
.0002	1.5	2.9	5.3	9.9	18.5	34.5	64.1
.00002	1.7	3.4	7.0	14.4	29.5	60.8	124.9

Table 1 shows for the distributions being simulated, the dependence on $H(.1)$ of $y_{.2}$, $y_{.02}$, $y_{.002}$, $y_{.0002}$, and $y_{.00002}$ (selected in light of the discussion later on in this section). In this table, the scale parameter is chosen so that the median, $y_{.5}$, of each distribution is one.

4.2 Confidence Bounds

Four methods of obtaining upper confidence bounds for extreme quantiles are evaluated in the simulation study: exponential-tail (ET), quadratic-tail (QT), exponential-tail with power transformation (ETP), and quadratic-tail with power transformation (QTP). In all four methods, the parameter t is chosen by adaption to the exponential distribution, as described in Section 2.4.

The ET and QT methods, as described in Sections 2.1 and 2.2, respectively, depend on a single integer-valued parameter m. The ETP and QTP methods depend on two integer-valued parameters, m_1 and m_2. Here m_1 is the value of m used in the preliminary power transformation and m_2 is the value of m that is used in the exponential-tail or quadratic-tail method applied to the transformed data. In the one-sample problem the parameter $\hat{\gamma}$ of the preliminary power transformation is chosen to satisfy (2.15).

In the simulation study, we also consider the multisample problem with $K = 10$. The ET and QT methods treat the ten samples separately. The ETP and QTP methods involve a preliminary power transformation, as described in Section 2.5. In the simulation, the ten sample sizes coincide; the integers $m_1, \cdots, m_{10}$ introduced in Section 2.5 also coincide, the common value being denoted by m_1. The exponential-tail method with parameter m_2 or quadratic-tail method with parameter m_2 is then separately applied to each of the ten transformed samples to yield upper confidence bounds for the various quantiles of interest. Finally, the inverse power transformation is applied to these upper confidence bounds to yield upper confidence bounds corresponding to the ten original samples.

Consider a confidence bound obtained by using the ETP or QTP method in the context of the one-sample or the multisample problem, with t being chosen by adaption to the exponential distribution (which takes the preliminary power transformation into account). Its actual coverage probability does not depend on the power parameter of the underlying Weibull, generalized gamma(5), or lognormal distribution. In particular, for underlying Weibull distributions, its actual coverage probability is equal to its nominal coverage probability.

Consider, instead, a confidence bound obtained from the ET or QT method with t being chosen by adaption to the exponential distribution. Its coverage probability does depend on the power parameter of the underlying distribution. In particular, for underlying Weibull distributions, its actual coverage probability is equal to its nominal coverage probability when the tail heaviness is zero but not when the tail heaviness is nonzero.

4.3 Parameter Selection

Let Med(U) denote the median of a random variable U; so that

$$\Pr(U \geq \mathrm{Med}(U)) = .5.$$

The *excess* of an upper confidence bound U for a quantile y_p is defined as

$$\frac{\mathrm{Med}(U) - y_p}{y_p} \times 100\%.$$

The excesses of the confidence bounds obtained from any of the four methods under investigation depend on the power parameter of the underlying Weibull, generalized gamma(5), or lognormal distribution.

Excesses and coverage probabilities of the nominal 90% upper confidence bounds obtained by using the four methods described in Section 4.2 will be used to compare these methods. Two sample sizes are considered: n = 50 and n = 500. Attention is restricted to three quantiles for each sample size: $y_{1/n}$, $y_{.1/n}$, and $y_{.01/n}$. Thus, for n = 50, we compare nominal 90% upper confidence bounds for $y_{.02}$, $y_{.002}$, and $y_{.0002}$; and, for n = 500, we compare nominal 90% upper confidence bounds for $y_{.002}$, $y_{.0002}$, and $y_{.00002}$.

In order to determine reasonable values of the pairs m_1,m_2 of parameters of the ETP and QTP methods, we conducted a preliminary simulation, in which we used 2000 trials to determine values of t for adaption to the exponential distribution, 1000 trials to determine actual coverage probabilities and excesses for Weibull distributions, 1000 trials to determine those for generalized gamma(5) distributions, and 1000 trials for lognormal distributions. For each sample size and each of the three families of distributions, the same (pseudo) random numbers were used for both methods, all three quantiles, and all pairs of parameters under investigation.

It is necessary to make tradeoffs between coverage probabilities and excesses. Consider a given sample size and the nominal 90% upper confidence bound for $y_{1/n}$ obtained by either method. We settled on choosing pairs m_1,m_2 to meet the objective that the estimated actual coverage probability be at least 88% for each of the three families of distributions. Our second objective is, subject to the constraint of the first objective, to minimize the sum of the excesses of the confidence bounds for the three distributions with zero tail heaviness. For confidence bounds for $y_{.1/n}$ and $y_{.01/n}$, we settled on 85% and 82%, respectively, instead of 88% in the first objective.

We actually made two passes at parameter selection. In the first pass we looked at pairs m_1,m_2 broadly spread out and in the second pass, we concentrated on regions that looked most promising in the first pass. Upon reflection, it occurred to us that we could do just about as well by choosing the same pair m_1,m_2 for all three quantiles. Since doing this has obvious conceptual and practical advantages, we decided to restrict attention to such quantile-invariant choices. We are satisfied that the pairs we ended up with, shown in Table 2, come close to meeting our two objectives.

Table 2. Values of m_1,m_2 for the ETP and QTP Methods

ETP		QTP	
n		n	
50	500	50	500
One-sample problem			
25, 7	150, 6	30,20	450,100
Ten-sample problem			
19,19	40,30	35,20	300,135

The ET and QT methods have a single parameter m. For these methods, the coverage probability depends on the power parameter of the underlying distribution.

Here, we initially modified our first objective by restricting the power parameter for each of the three families to the seven values indicated in Section 4.1, which correspond to values of the tail heaviness at the upper decile that range from –.2 to .4.

It became apparent, however, that it is not always possible to realize this objective, especially for the lognormal distribution, the highest values of the tail heaviness, and the most extreme quantiles. For the QT method, we ended up by choosing m to give (approximately) the best coverage probabilities for the heavy-tailed distributions for all three quantiles: for $n = 50$ we chose $m = 30$ and for $n = 500$ we chose $m = 45$. These values of m yield reasonably good coverage probabilities and excesses that are not unreasonably large. For the ET method we chose $m = 3$ for both sample sizes, since the coverage probabilities dipped too low for $m \geq 4$ and the excesses were unreasonably large for $m = 2$.

4.4 Results

In the final simulation, the performance of the four methods for obtaining 90% upper confidence bounds, with m or m_1,m_2 as described in Section 4.3, was re-evaluated. Now, 10,000 trials were used to determine values of t for adaption to the exponential distribution, 5000 trials to determine actual coverage probabilities and excesses for Weibull distributions, and 5000 trials each for generalized gamma(5) and lognormal distributions.

Figures 1, 2, and 3 show the results for the one-sample problem with $p = 1/n$, $.1/n$ and $.01/n$ respectively, while Figures 4–6 show those for the ten-sample problem and the same values of p. Although all three families were used in the determination of m or m_1,m_2, as described in Section 4.3, the final results for the generalized gamma(5) family are omitted from Figures 1–6 in order to save space. These results are intermediate between those for the Weibull family and those for the lognormal family. This is compatible with the quantile values for the three families shown in Table 1 in Section 4.1.

Corresponding to each of the four methods for obtaining an upper confidence bound for y_p is an estimate $\hat{y}_p$ (which does not involve adaption to the exponential distribution). The *bias* of such an estimate is defined as

$$E\left[\frac{\hat{y}_p - y_p}{y_p}\right] \times 100\%.$$

In all six figures, the Monte Carlo estimate of the bias of the corresponding estimate is shown along with the coverage probability (expressed as a percentage) and excess of each method. The correspondence between linetypes and methods is as follows:

- - - - - - - - -	ET
— — — — —	QT
· · · · · · · · · · · · · · ·	ETP
————————	QTP

The ET and QT methods treat separately the ten samples in the ten-sample problem. Thus the results for these methods that are shown in Figures 4–6 coincide with the corresponding results shown in Figures 1–3.

In each plot on each figure, the tail heaviness ranges from –.2 to .4 along the horizontal axis. Note that the coverage probabilities of confidence bounds obtained by the ETP and QTP methods do not depend on tail heaviness, while those obtained by the ET and QT methods do depend on tail heaviness.

According to the coverage and excess plots in Figures 1–3, in the one-sample problem the QT method is best for all three quantiles when $n = 50$ and the QTP method is best for all three quantiles when $n = 500$. According to Figures 4–6, in the ten-sample problem the ETP method is best for both sample sizes and all three quantiles. (Letting the shapes of the distributions in the ten samples differ somewhat from each other would have favored QTP over ETP.) In particular, the ET method is never the best method.

As of now, someone wanting to apply the QT, ETP or QTP method to real data would need to do a simulation to select t as described in Section 2.4. Perhaps further simulation studies would be required to select m or m_1, m_2 for the sample sizes, quantiles and hypothetical distributions of interest. Although such computer simulations are increasingly feasible because of the ever greater prevalence of powerful workstations, they are still not easy to perform. But nobody has claimed that getting reliable confidence bounds for extreme quantiles would be easy!

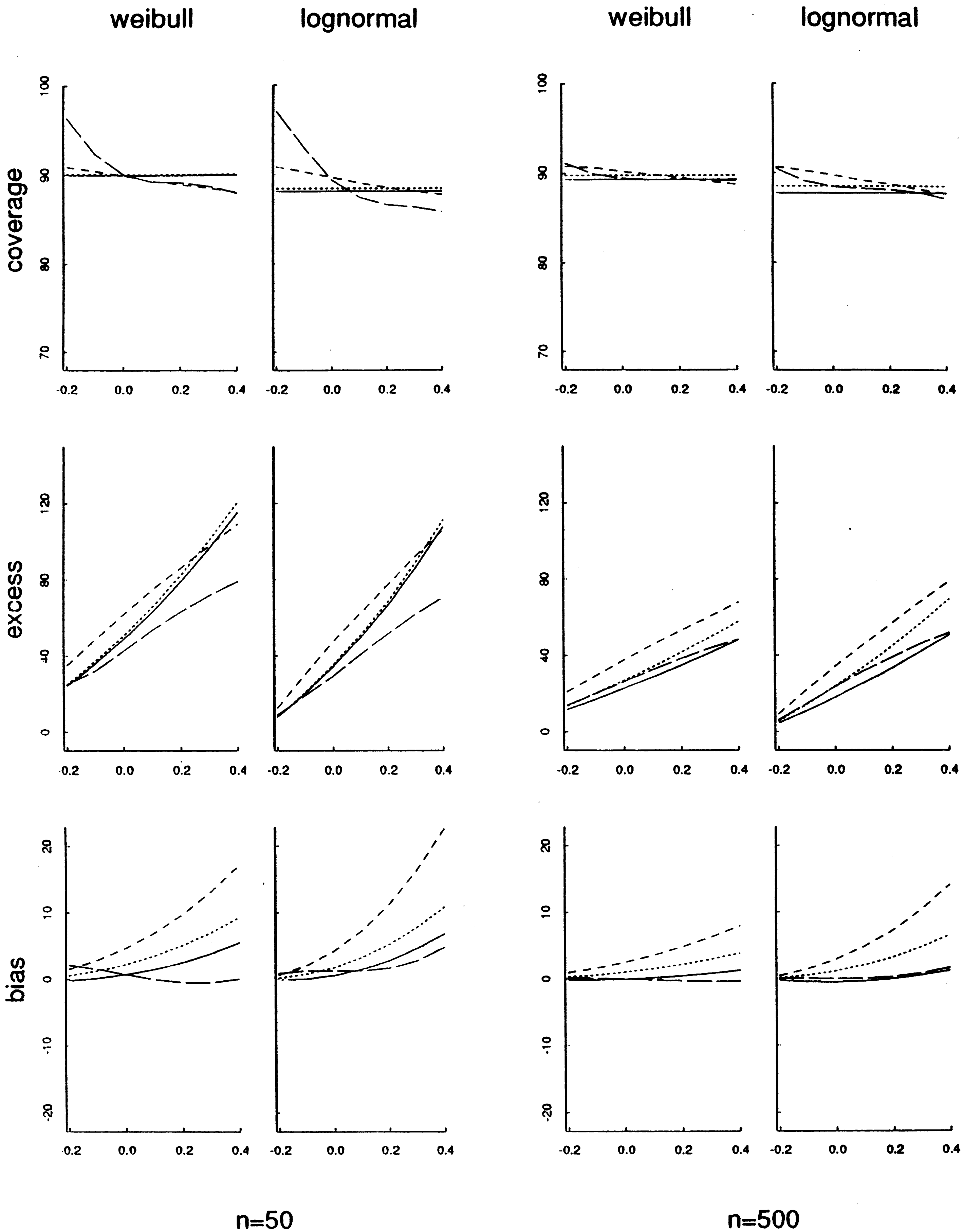

Figure 1. One-sample Problem With $p = 1/n$.

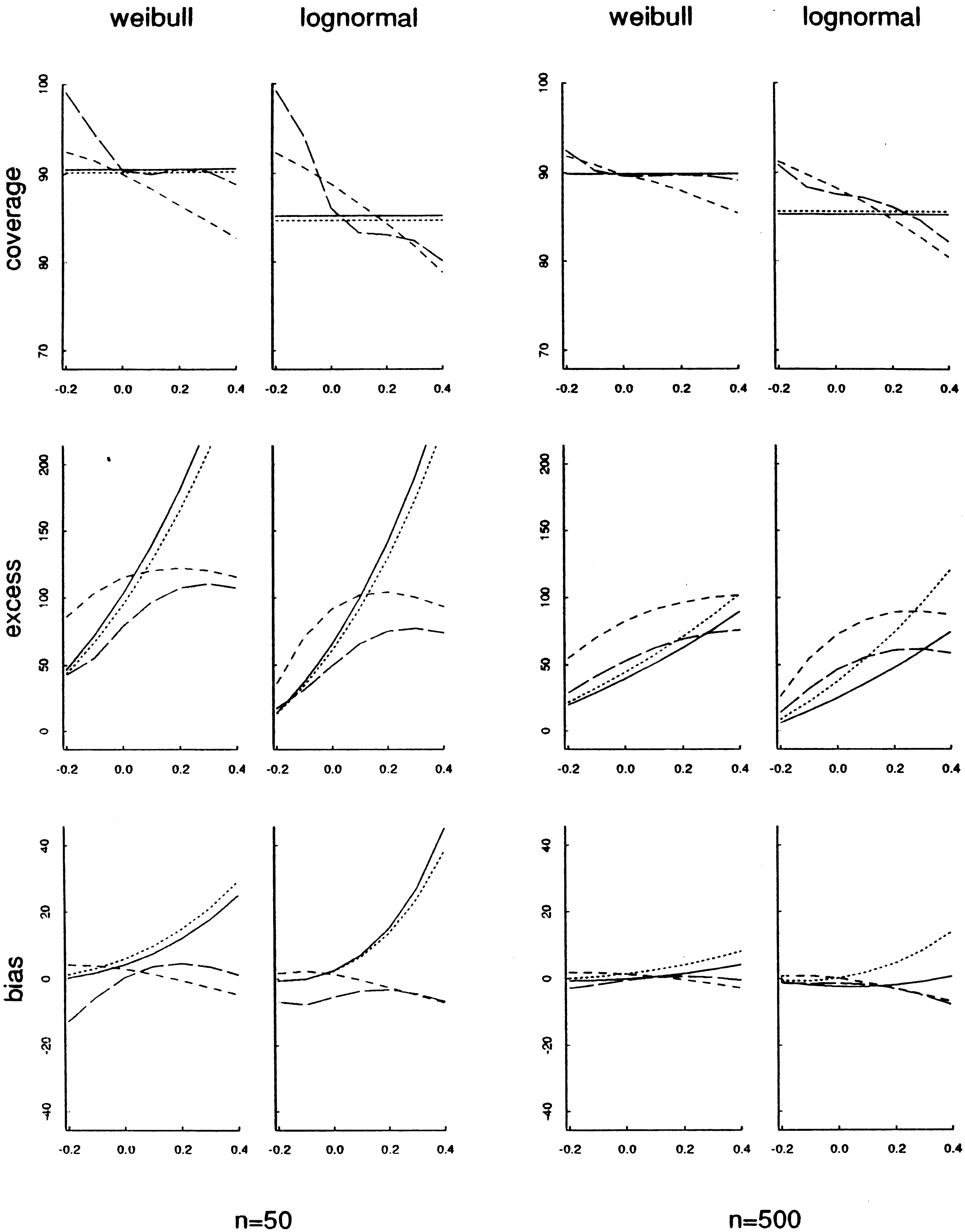

Figure 2. One-sample Problem With $p = .1/n$.

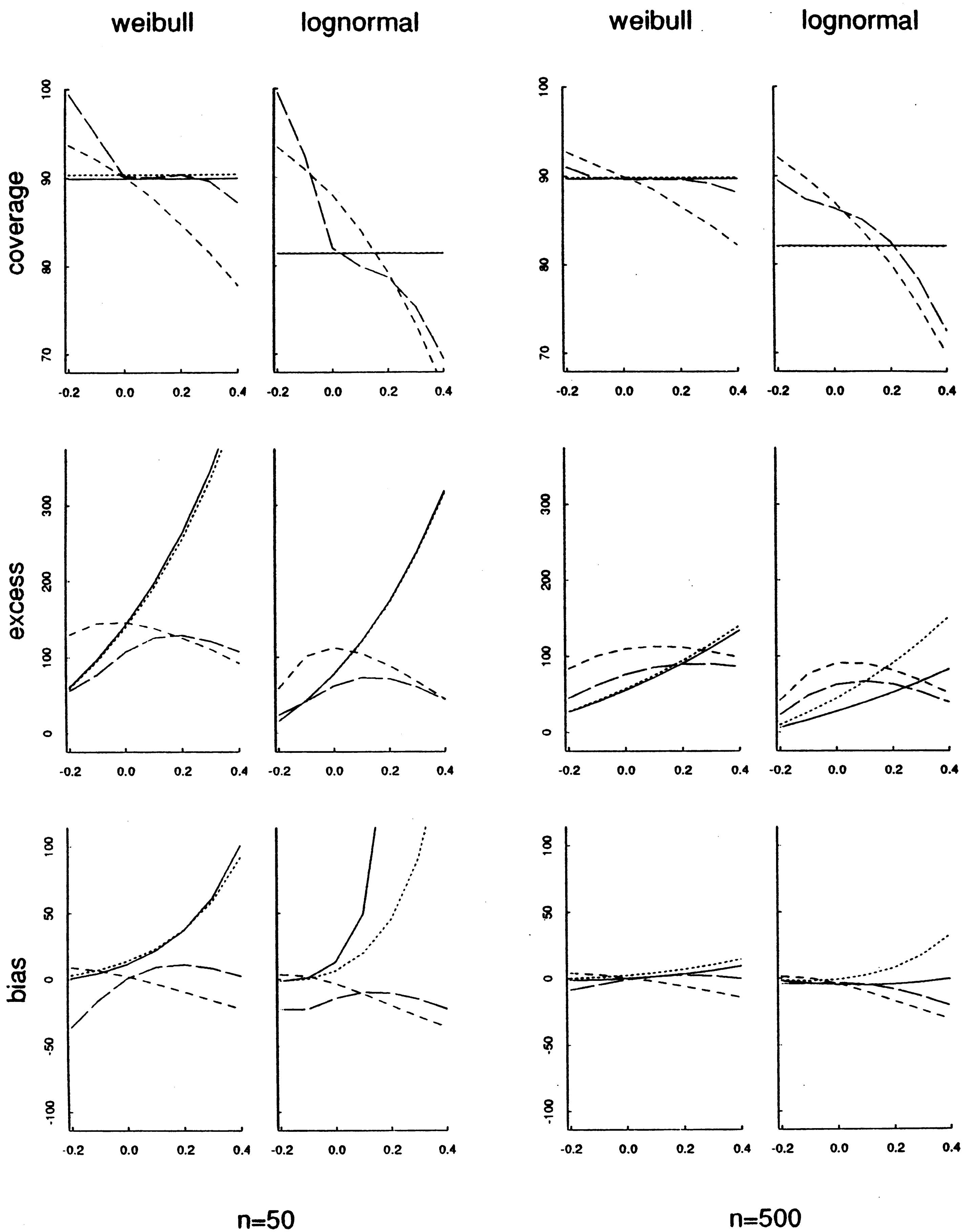

Figure 3. One-sample Problem With $p = .01/n$.

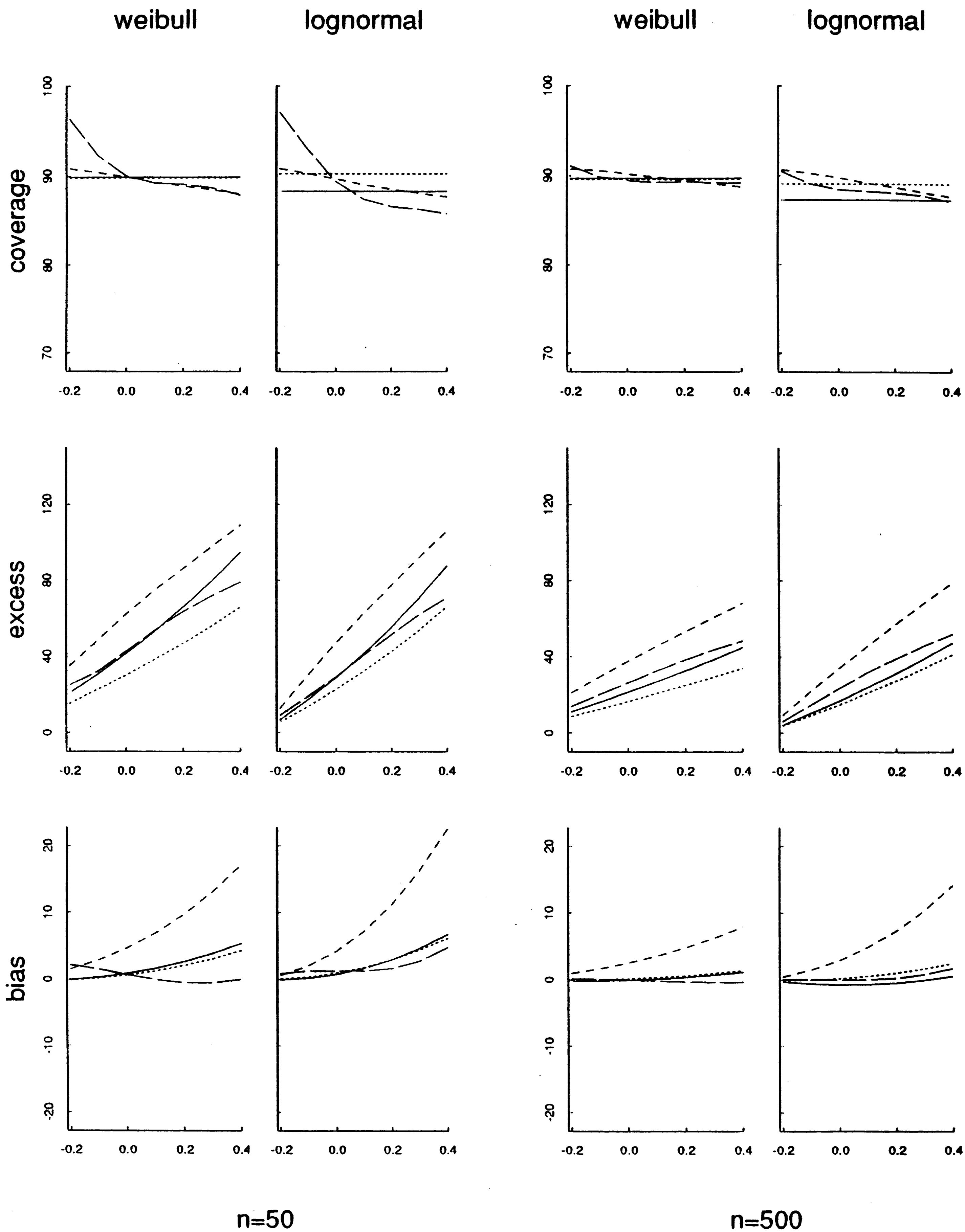

Figure 4. Ten-sample Problem With $p = 1/n$.

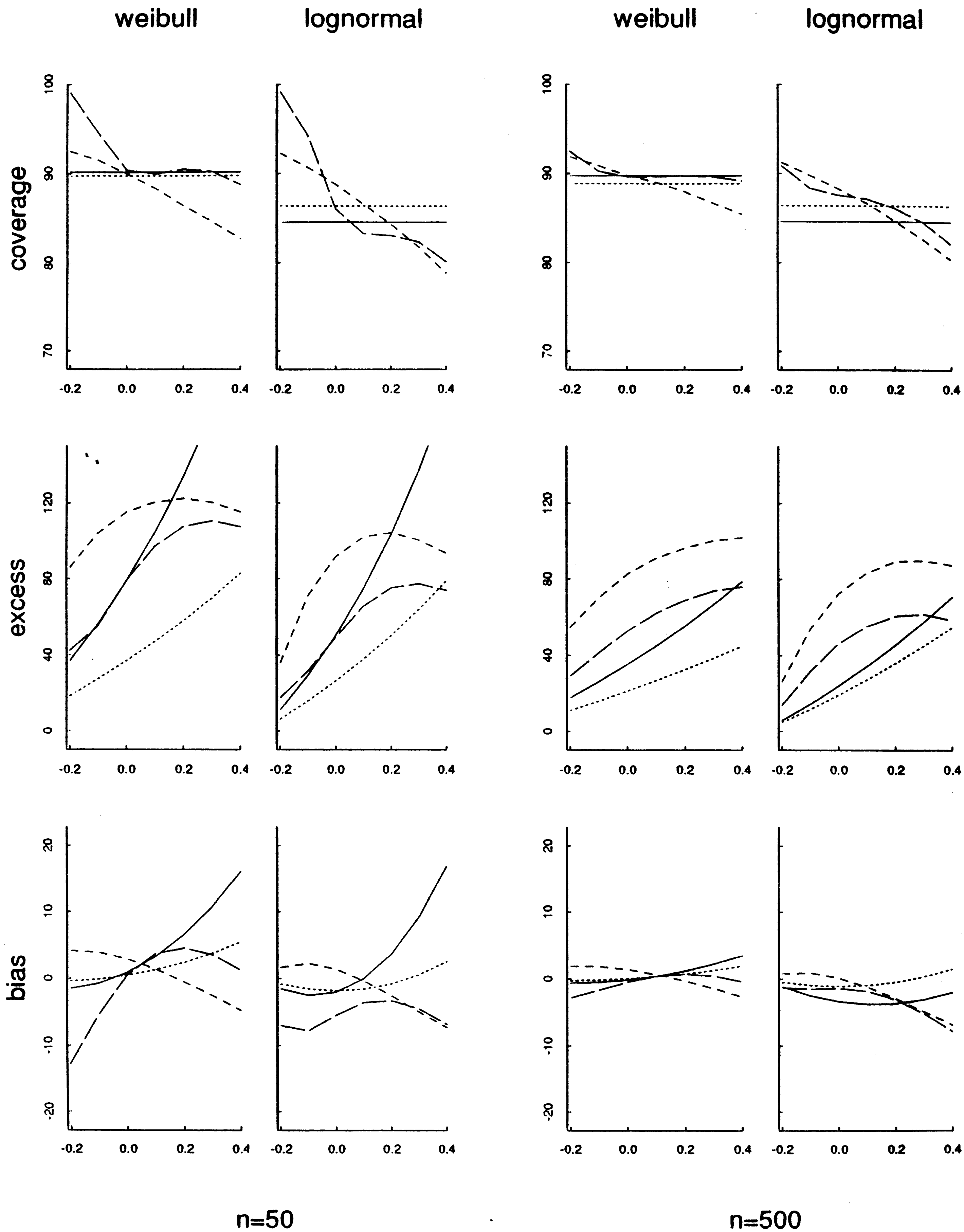

Figure 5. Ten-sample Problem With $p = .1/n$.

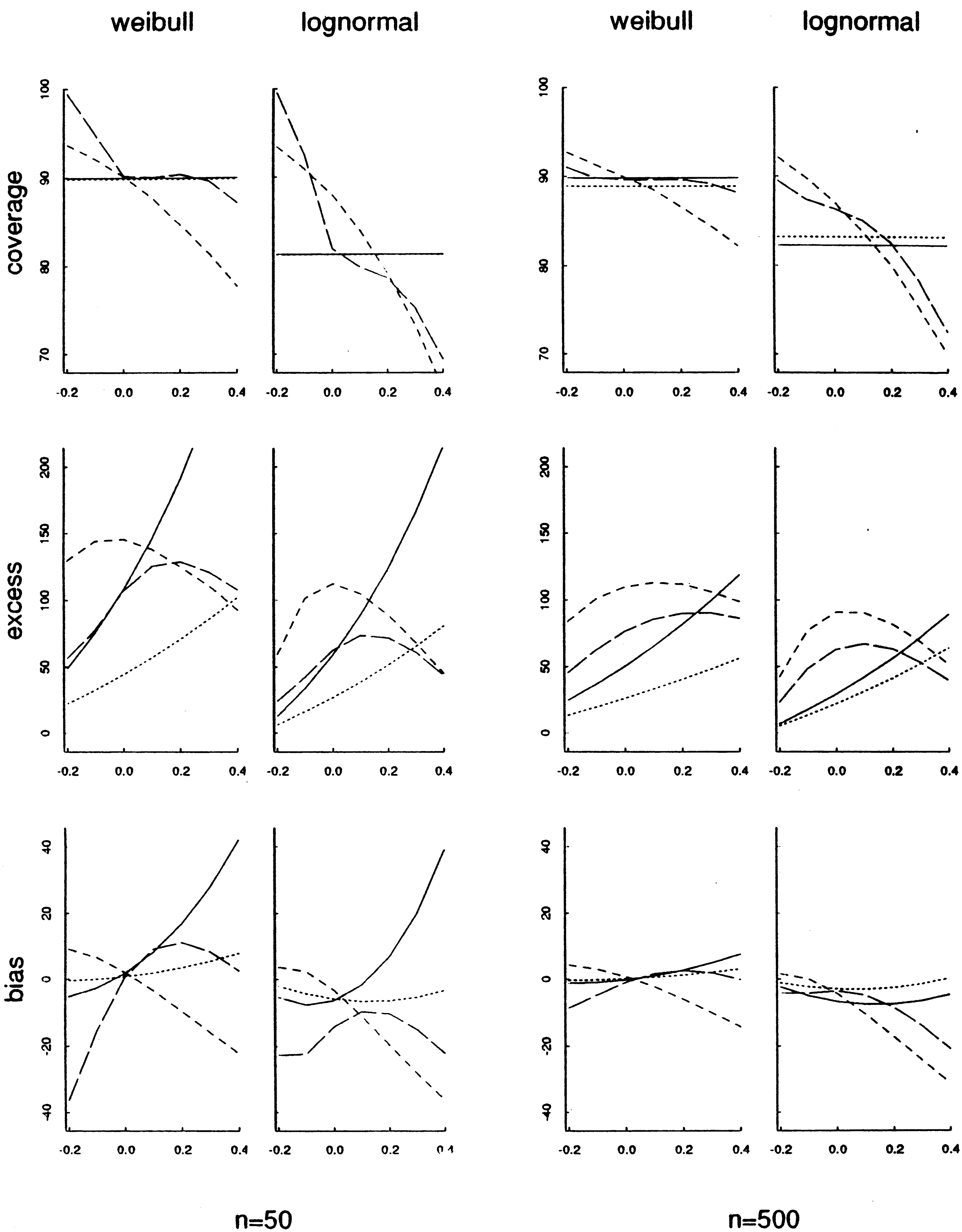

Figure 6. Ten-sample Problem With $p = .01/n$.

5. QUADRATIC-TAIL MODEL

We now develop the informal properties of the quadratic-tail model that were used in Section 2 to obtain the corresponding upper confidence bounds. To this end, let $U_{(1)},\cdots,U_{(n)}$ be the decreasing order statistics based on a random sample of size n from the uniform distribution on [0, 1]. Then $G(Y_{(1)}),\cdots,G(Y_{(n)})$ have the same joint distribution as $1-U_{(1)},\cdots,1-U_{(n)}$. Let $Z_1,\cdots,Z_n$ be a random sample of size n from the exponential distribution with mean one and let $Z_{(1)},\cdots,Z_{(n)}$ be the corresponding decreasing order statistics. Then $1-U_{(1)},\cdots,1-U_{(n)}$ have the same joint distribution as $\exp(-Z_{(1)}),\cdots,\exp(-Z_{(n)})$; and $Y_{(1)},\cdots,Y_{(n)}$ have the same joint distribution as $G^{-1}(\exp(-Z_{(1)})),\cdots,G^{-1}(\exp(-Z_{(n)}))$. In particular, $Y_{(1)},\cdots,Y_{(m)}$ have the same joint distribution as $G^{-1}(\exp(-Z_{(1)})),\cdots,G^{-1}(\exp(-Z_{(m)}))$.

It follows from (2.4) with $p = e^{-y}$ that, for $y \geq \log(n/m)$,

$$G^{-1}(e^{-y}) = y_{m/n} + \alpha\,[y-\log(n/m)] + \frac{\beta}{2}\,[y^2-\log^2(n/m)]$$

for $y \geq \log(1/p_0)$. Thus if $Z_{(m)} \geq \log(1/p_0)$, then $G^{-1}(\exp(-Z_{(i)}))$, $i = 1,\cdots,m$ coincide respectively with

$$y_{m/n} + \alpha[Z_{(i)}-\log(n/m)] + \frac{\beta}{2}[Z_{(i)}^2-\log^2(n/m)], \quad i = 1,\cdots,m.$$

Ignoring the error in the quadratic-tail model and the possibility that $Y_{(m)} < \log(1/p_0)$, we conclude that $Y_{(i)}$, $i = 1,\cdots,m$ have the same joint distribution as

$$y_{m/n} + \alpha\,[Z_{(i)}-\log(n/m)] + \frac{\beta}{2}\,[Z_{(i)}^2-\log^2(n/m)], \quad i = 1,\cdots,m.$$

In particular, $Y_{(i)}-Y_{(i+1)}$, $i = 1,\cdots,m-1$, have the same joint distribution as

$$\alpha\,[Z_{(i)}-Z_{(i+1)}] + \frac{\beta}{2}\,[Z_{(i)}^2-Z_{(i+1)}^2], \quad i = 1,\cdots,m.$$

Let L and M be known constants. Consider the parameter $\tau = L\alpha + M\beta$. Let $v_1,\cdots,v_{m-1}$ be known constants and consider the estimate

$$\hat{\tau} = \sum_{i=1}^{m-1} iv_i[Y_{(i)}-Y_{(i+1)}]$$

of τ. Observe that $\hat{\tau}$ has the same distribution as

$$\sum_{i=1}^{m-1} iv_i\left[\alpha\,[Z_{(i)}-Z_{(i+1)}] + \frac{\beta}{2}\,[Z_{(i)}^2-Z_{(i+1)}^2]\right]$$

and hence that

$$E\hat{\tau} = \alpha \sum_{i=1}^{m-1} iv_i E(Z_{(i)} - Z_{(i+1)}) + \frac{\beta}{2} \sum_{i=1}^{m-1} iv_i E(Z_{(i)}^2 - Z_{(i+1)}^2). \tag{5.1}$$

As is well known (see page 37 of Galambos 1978 or page 21 of David 1981), $Z_{(i)}$, $i = 1,\cdots,n$, have the same joint distribution as

$$\sum_{j=i}^{n} \frac{Z_j}{j}, \quad i = 1,\cdots,n.$$

Consequently, for $1 \le i \le n-1$,

$$E(Z_{(i)} - Z_{(i+1)}) = E\left[\frac{Z_i}{i}\right]$$

and hence

$$E(Z_{(i)} - Z_{(i+1)}) = \frac{1}{i}. \tag{5.2}$$

Now

$$Z_{(i)}^2 - Z_{(i+1)}^2 = [Z_{(i)} - Z_{(i+1)}]^2 + 2Z_{(i+1)}[Z_{(i)} - Z_{(i+1)}]$$

and hence

$$\mathrm{E}\left[Z_{(i)}^2 - Z_{(i+1)}^2\right] = E\left(\left[\frac{Z_i}{i}\right]^2\right) + 2E\left[\frac{Z_i}{i}\left[\sum_{j=i+1}^{n} \frac{Z_j}{j}\right]\right] = \frac{2}{i^2} + \frac{2}{i} \sum_{j=i+1}^{n} \frac{1}{j}.$$

Therefore,

$$E\left[Z_{(i)}^2 - Z_{(i+1)}^2\right] = \frac{2u_i}{i}, \tag{5.3}$$

where

$$u_i = \sum_{j=i}^{n} \frac{1}{j}.$$

We conclude from (5.1)–(5.3) that

$$E\hat{\tau} = \alpha \sum_{i=1}^{m-1} v_i + \beta \sum_{i=1}^{m-1} u_i v_i. \tag{5.4}$$

Thus $\hat{\tau}$ is unbiased if and only if

$$L = \sum_{i=1}^{m-1} v_i \quad \text{and} \quad M = \sum_{i=1}^{m-1} u_i v_i. \tag{5.5}$$

The variance of $\hat{\tau}$ is derived by a simple but lengthy computation given in Appendix A. To state the result, set

$$u_k^{(2)} = \sum_{j=i}^{n} \frac{1}{j^2} \quad \text{and} \quad \bar{v}_i = \frac{1}{i} \sum_{j=1}^{i} v_j.$$

Then

$$\mathrm{Var}(\hat{\tau}) = \sum_{i=1}^{m-1}(\alpha v_i+\beta\bar{v}_i+\beta u_iv_i)^2 + \beta^2\left[\sum_{i=1}^{m-1}u_i^{(2)}v_i^2 + (m-1)u_m^{(2)}\bar{v}_{m-1}^2\right]. \tag{5.6}$$

It follows from (5.6) that (2.7) holds with $w_i/i = v_i$ for $1 \le i \le m-1$,

$$c_1 = \sum_{i=1}^{m-1}v_i^2, \tag{5.7}$$

$$c_2 = \sum_{i=1}^{m-1}v_i(\bar{v}_i+u_iv_i),$$

and

$$c_3 = \sum_{i=1}^{m-1}[(\bar{v}_i+u_iv_i)^2+u_i^{(2)}v_i^2] + (m-1)^2u_m^{(2)}\bar{v}_{m-1}^2.$$

Consider the problem of choosing $v_1,\cdots,v_{m-1}$ to minimize the right side of (5.7) subject to (5.5). It is geometrically clear that there is a unique solution to this minimization problem and that the solution is given by $v_i = \lambda_1 + \lambda_2u_i$, $1 \le i \le m-1$, where λ_1 and λ_2 are chosen to satisfy (5.5). It is easily seen that

$$\lambda_1 = \frac{S_2L-S_1M}{D} \quad \text{and} \quad \lambda_2 = \frac{(m-1)M-S_1L}{D},$$

where

$$S_1 = \sum_{i=1}^{m-1}u_i, \quad S_2 = \sum_{i=1}^{m-1}u_i^2,$$

and $D = (m-1)S_2 - S_1^2$. Thus, for $1 \le i \le m-1$,

$$v_i = \frac{L}{D}[S_2-S_1u_i] + \frac{M}{D}[(m-1)u_i-S_1]. \tag{5.8}$$

By choosing $L = 1$ and $M = 0$, we obtain the unbiased estimate of α given by

$$\hat{\alpha} = \sum_{i=1}^{m-1}w_{1i}[Y_{(i)}-Y_{(i+1)}] = \sum_{i=1}^{m-1}iv_{1i}[Y_{(i)}-Y_{(i+1)}],$$

where

$$\frac{w_{1i}}{i} = v_{1i} = \frac{S_2-S_1u_i}{D}$$

for $1 \le i \le m-1$. By choosing $L = 0$ and $M = 1$, we obtain the unbiased estimate of β given by

$$\hat{\beta} = \sum_{i=1}^{m-1}w_{2i}[Y_{(i)}-Y_{(i+1)}] = \sum_{i=1}^{m-1}iv_{2i}[Y_{(i)}-Y_{(i+1)}],$$

where

$$\frac{w_{2i}}{i} = v_{2i} = \frac{(m-1)u_i - S_1}{D}$$

for $1 \le i \le m-1$. It now follows from (5.8) that, for arbitrary values of L and M, the unbiased estimate of τ for which c_1 is minimized is given by $\hat{\tau} = L\hat{\alpha} + M\hat{\beta}$.

The variance of

$$\hat{y}_p = Y_{(m)} + \hat{\tau} = Y_{(m)} + \sum_{i=1}^{m-1} iv_i[Y_{(i)} - Y_{(i+1)}]$$

is the same as the variance of

$$\alpha\, Z_{(m)} + \frac{\beta}{2} Z^2_{(m)} + \sum_{i=1}^{m-1} iv_i\left[\alpha\, [Z_{(i)} - Z_{(i+1)}] + \frac{\beta}{2}\, [Z^2_{(i)} - Z^2_{(i+1)}]\right].$$

This variance is clearly a quadratic function of α and β; so that (2.11) holds. It follows from (5.8) that the constants C_1, C_2 and C_3 in (2.11) depend only on n, m, L, and M. These constants are determined explicitly in Appendix A.

APPENDIX A: QUADRATIC-TAIL MODEL

We now derive (5.6) and determine the constants in (2.11).

Recall that $Z_1, \cdots, Z_n$ are independent random variables, each having an exponential distribution with mean one. The following facts are easily checked: $\mathrm{Var}(Z_1) = 1$; $\mathrm{Var}(Z_1^2) = 20$; $\mathrm{Var}(Z_1Z_2) = 3$; $\mathrm{Cov}(Z_1, Z_1^2) = 4$; $\mathrm{Cov}(Z_1, Z_1Z_2) = 1$; $\mathrm{Cov}(Z_1^2, Z_1Z_2) = 4$; and $\mathrm{Cov}(Z_1Z_2, Z_1Z_3) = 1$. It can be assumed that, for $1 \le i \le n$,

$$Z_{(i)} = \sum_{j=i}^{n} \frac{Z_j}{j}.$$

Until further notice, unless otherwise indicated, the variables i and j range over $1, \cdots, m-1$. Set $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ if $i \ne j$; and set $\psi_{ij} = 1$ if $i > j$ and $\psi_{ij} = 0$ if $i \le j$. The following formulas are easily verified:

$$\mathrm{Cov}\left[Z_i, Z_jZ_{(j+1)}\right] = \delta_{ij}u_j + \frac{\psi_{ij}}{i} - \frac{\delta_{ij}}{i};$$

$$\mathrm{Cov}\left[Z_i^2, Z_jZ_{(j+1)}\right] = 4\left[\delta_{ij}u_j + \frac{\psi_{ij}}{i} - \frac{\delta_{ij}}{i}\right];$$

$$\mathrm{Var}\left[Z_iZ_{(i+1)}\right] = u_i^2 - 2\frac{u_i}{i} + 2u_i^{(2)} - \frac{1}{i^2};$$

and

$$\mathrm{Cov}\left[Z_iZ_{(i+1)}, Z_jZ_{(j+1)}\right] = u_i^{(2)} + \frac{u_i}{i} - \frac{2}{i^2}, \quad i > j.$$

In verifying (5.6), it can be assumed that

$$\begin{aligned}\hat{\tau} &= \Sigma i v_i\left[\alpha\,[Z_{(i)} - Z_{(i+1)}] + \frac{\beta}{2}\,[Z_{(i)}^2 - Z_{(i+1)}^2]\right] \\ &= \alpha\,\Sigma v_iZ_i + \beta\,\Sigma v_i\left[\frac{Z_i^2}{2i} + Z_iZ_{(i+1)}\right].\end{aligned}$$

Thus $\mathrm{Var}(\hat{\tau}) = c_1\alpha^2 + c_2\alpha\beta + c_3\beta^2$, where

$$c_1 = \mathrm{Var}\left[\Sigma v_iZ_i\right] = \Sigma v_i^2$$

and

$$\begin{aligned}c_2 &= 2\,\mathrm{Cov}\left[\Sigma v_iZ_i, \Sigma v_i\left[\frac{Z_i^2}{2i} + Z_iZ_{(i+1)}\right]\right] \\ &= 4\,\Sigma\frac{v_i^2}{i} + 2\,\Sigma\Sigma v_iv_j\left[\delta_{ij}u_j + \frac{\psi_{ij}}{i} - \frac{\delta_{ij}}{i}\right] \\ &= 2\,\Sigma v_i(\bar{v}_i + u_iv_i).\end{aligned}$$

Also,

$$c_3 = \mathrm{Var}\left[\frac{1}{2}\,\Sigma\frac{v_i}{i}Z_i^2 + \Sigma v_i Z_i Z_{(i+1)}\right] = c_4 + c_5 + c_6.$$

Here

$$c_4 = \frac{1}{4}\,\mathrm{Var}\left[\Sigma\frac{v_i}{i}Z_i^2\right] = 5\,\Sigma\frac{v_i^2}{i^2}.$$

Next,

$$c_5 = \mathrm{Cov}\left[\Sigma\frac{v_i}{i}Z_i^2,\ \Sigma v_i Z_i Z_{(i+1)}\right]$$
$$= 4\,\Sigma\frac{u_i v_i^2}{i} + 4\,\Sigma\frac{v_i \bar{v}_i}{i} - 8\,\Sigma\frac{v_i^2}{i^2}.$$

Moreover,

$$c_6 = \mathrm{Var}\left[\Sigma v_i Z_i Z_{(i+1)}\right]$$
$$= \Sigma v_i^2\left[u_i^2 - 2\frac{u_i}{i} + 2u_i^{(2)} - \frac{1}{i^2}\right] + 2\,\Sigma v_i\left[u_i^{(2)} + \frac{u_i}{i} - \frac{2}{i^2}\right]\left[i\bar{v}_i - v_i\right]$$
$$= \Sigma u_i^2 v_i^2 - 4\,\Sigma\frac{u_i v_i^2}{i} + 3\,\Sigma\frac{v_i^2}{i^2} + 2\,\Sigma i u_i^{(2)} v_i \bar{v}_i - 4\,\Sigma\frac{v_i\bar{v}_i}{i} + 2\,\Sigma u_i v_i \bar{v}_i.$$

Consequently,

$$c_3 = \Sigma u_i^2 v_i^2 + 2\,\Sigma i u_i^{(2)} v_i \bar{v}_i + 2\,\Sigma u_i v_i \bar{v}_i.$$

Observe that

$$\Sigma \bar{v}_i^2 = \Sigma\left[u_i^{(2)} - u_{i+1}^{(2)}\right]\left[\Sigma_{i\geq j} v_j\right]^2 = 2\,\Sigma i u_i^{(2)} v_i \bar{v}_i - \Sigma u_i^{(2)} v_i^2 - (m-1)^2 u_m^{(2)} \bar{v}_{m-1}^2$$

and hence that

$$c_3 = \Sigma(\bar{v}_i + u_i v_i)^2 + \Sigma u_i^{(2)} v_i^2 + (m-1)^2 u_m^{(2)} \bar{v}_{m-1}^2.$$

The last formula for c_3 and the previous formulas for c_1 and c_2 together show that (5.6) is valid.

Writing $\hat{y}_p$ as $Y_{(m)} + \hat{\tau}$, we see that

$$\mathrm{Var}\left[\hat{y}_p\right] = \mathrm{Var}\left[Y_{(m)}\right] + 2\,\mathrm{Cov}\left[Y_{(m)}, \hat{\tau}\right] + \mathrm{Var}\left[\hat{\tau}\right].$$

Now $\mathrm{Var}(\hat{\tau})$ is given explicitly in (5.6). Thus to determine the constants in (2.11), we need to determine explicit formulas for $\mathrm{Var}(Y_{(m)})$ and $\mathrm{Cov}(Y_{(m)}, \hat{\tau})$.

Now

$$\mathrm{Var}(Y_{(m)}) = \mathrm{Var}\left[\alpha\, Z_{(m)} + \frac{\beta}{2} Z_{(m)}^2\right]$$

and hence

$$\mathrm{Var}\left[Y_{(m)}\right] = \alpha^2 \mathrm{Var}\left[Z_{(m)}\right] + \alpha\beta\, \mathrm{Cov}\left[Z_{(m)}, Z_{(m)}^2\right] + \frac{\beta^2}{4} \mathrm{Var}\left[Z_{(m)}^2\right]. \tag{A.1}$$

It will be shown below that

$$\mathrm{Var}\left[Z_{(m)}\right] = u_m^{(2)}, \tag{A.2}$$

$$\mathrm{Cov}\left[Z_{(m)}, Z_{(m)}^2\right] = 2\left[u_m^{(3)} + u_m^{(2)} u_m\right], \tag{A.3}$$

and

$$\mathrm{Var}\left[Z_{(m)}^2\right] = 6u_m^{(4)} + 8u_m^{(3)} u_m + 2(u_m^{(2)})^2 + 4u_m^{(2)} u_m^2, \tag{A.4}$$

where

$$u_m^{(3)} = \sum_{i=m}^{n} \frac{1}{i^3} \quad \text{and} \quad u_m^{(4)} = \sum_{i=m}^{n} \frac{1}{i^4}.$$

Equations (A.1)–(A.4) together yield an explicit formula for $\mathrm{Var}(Y_{(m)})$. Also,

$$\begin{aligned}
\mathrm{Cov}\left[Y_{(m)}, \hat{\tau}\right] &= \mathrm{Cov}\left[\alpha Z_{(m)} + \frac{\beta}{2} Z_{(m)}^2, \beta\, \Sigma i v_i Z_{(i+1)}[Z_{(i)} - Z_{(i+1)}]\right] \\
&= \mathrm{Cov}\left[\alpha Z_{(m)} + \frac{\beta}{2} Z_{(m)}^2, \beta\left[\Sigma v_i\right] Z_{(m)}\right]
\end{aligned}$$

and hence

$$\mathrm{Cov}\left[Y_{(m)}, \hat{\tau}\right] = (m-1)\bar{v}_{m-1} \alpha\beta \mathrm{Var}\left[Z_{(m)}\right] + \frac{\beta^2}{2} \mathrm{Cov}\left[Z_{(m)}, Z_{(m)}^2\right]. \tag{A.5}$$

Equations (A.2), (A.3), and (A.5) determine an explicit formula for $\mathrm{Cov}(Y_{(m)}, \hat{\tau})$.

It remains to verify (A.2)–(A.4). To this end, let i,j,k,l range from m to n. Then

$$\mathrm{Var}\left[Z_{(m)}\right] = \mathrm{Var}\left[\Sigma \frac{Z_i}{i}\right] = \Sigma \frac{1}{i^2} = u_m^{(2)},$$

so (A.2) holds. Observe next that

$$\begin{aligned}
\mathrm{Cov}\left[Z_{(m)}, Z_{(m)}^2\right] &= \mathrm{Cov}\left[\Sigma \frac{Z_i}{i}, \left[\Sigma\left[\frac{Z_i}{i}\right]\right]^2\right] \\
&= \Sigma\Sigma\Sigma \frac{1}{ijk} \mathrm{Cov}\left[Z_i, Z_j Z_k\right] \\
&= \Sigma \frac{1}{i^3} \mathrm{Cov}\left[Z_i, Z_i^2\right] + 2\, \Sigma \frac{1}{i^2} \sum_{i \neq j} \frac{1}{j} \mathrm{Cov}(Z_i, Z_i Z_j) \\
&= \Sigma 4u_m^{(3)} + 2\, \Sigma \frac{1}{i^2}\left[\Sigma \frac{1}{j} - \frac{1}{i}\right] \\
&= 2\left[u_m^{(3)} + u_m^{(2)} u_m\right],
\end{aligned}$$

so (A.3) holds.

Finally,

$$\mathrm{Var}\left[Z_{(m)}\right] = \mathrm{Var}\left[\left[\Sigma\frac{Z_i}{i}\right]^2\right] = \Sigma\Sigma\Sigma\Sigma\frac{1}{ijkl}\,\mathrm{Cov}\left[Z_iZ_j,\, Z_kZ_l\right].$$

The total contribution of all terms for which $i = j = k = l$ is

$$\mathrm{Var}\left[Z_1^2\right]\ \Sigma\frac{1}{i^4} = 20u_m^{(4)}.$$

The total contribution of all terms for which exactly three of the four quantities i,j,k,l coincide is

$$4\ \mathrm{Cov}\left[Z_1^2,\, Z_1Z_2\right]\ \Sigma\frac{1}{i^3}\sum_{i\neq j}\frac{1}{j} = 16\left[u_m^{(3)}u_m - u_m^{(4)}\right].$$

The total contribution of all terms for which i and j are distinct and exactly one of the pair k,l equals i or j is

$$4\ \mathrm{Cov}\left[Z_1Z_2,\, Z_1Z_3\right]\ \sum_{i\neq j\neq k}\Sigma\Sigma\frac{1}{i^2jk} = 4\,\Sigma\frac{1}{i^2}\sum_{i\neq j}\frac{1}{j}\left[u_m - \frac{1}{i} - \frac{1}{j}\right]$$

$$= 8u_m^{(4)} - 8u_m^{(3)}u_m - 4(u_m^{(2)})^2 + 4u_m^{(2)}u_m^2;$$

here $i\neq j\neq k$ means that i,j and k are distinct. The total contribution of all terms for which i and j are distinct and (k, l) is either (i, j) or (j, i) is

$$2\ \mathrm{Var}(Z_1Z_2)\ \Sigma\frac{1}{i^2}\sum_{i\neq j}\frac{1}{j^2} = 6\left[(u_m^{(2)})^2 - u_m^{(4)}\right].$$

Equation (A.4) follows by adding up these four totals.

APPENDIX B: POWER TRANSFORMATION

We first verify (2.14). Let $Z_1, \cdots, Z_n$ be independent random variables each having an exponential distribution with (say) mean 1 and let $Z_{(1)}, \cdots, Z_{(n)}$ be the corresponding decreasing order statistics. Let $2 \le m \le n$ and let i range over $1, \cdots, m$. Then the conditional distribution of

$$\frac{\Sigma\, [Z_{(i)} - Z_{(m)}]^2}{\left[\Sigma\, [Z_{(i)} - Z_{(m)}]\right]^2}$$

given $Z_{(m)}$ is the same as the distribution of $\Sigma Z_i^2 \,/\, (\Sigma Z_i)^2$. Thus, to verify (2.14), it suffices to verify that the latter random variable has mean $2/m$ or, equivalently, that $Z_1^2/(\Sigma Z_i)^2$ has mean $2/[m(m-1)]$.

Set $V = Z_1$ and $W = Z_2 + \cdots + Z_{m-1}$. We need to verify that

$$E\left[\left[\frac{V}{V+W}\right]^2\right] = \frac{2}{m(m-1)}. \tag{B.1}$$

Now V and W are independent random variables; V has the gamma distribution with parameters 1 and 1; W has the gamma distribution with parameters $m-2$ and 1; and $V+W$ has the gamma distribution with parameters $m-1$ and 1. It follows form the change of variables formula involving Jacobians that W/V and $V+W$ are independent and hence that $V/(V+W)$ and $V+W$ are independent. Consequently,

$$2 = E(V^2) = E\left[\left[\frac{V}{V+W}\right]^2 (V+W)^2\right] = [(m-1)^2 + m-1]E\left[\left[\frac{V}{V+W}\right]^2\right] = m(m-1)E\left[\left[\frac{V}{V+W}\right]^2\right]$$

and hence (B.1) holds.

We will now show that the left side of (2.15) is a strictly increasing function of $\hat{\gamma} \in (0, \infty)$. It is enough to verify the following result: Let W be a nonconstant discrete random variable having a finite number of possible values, each of which is greater than one. Define the function g on $(0, \infty)$ by

$$\frac{1}{2} g(\gamma) = \frac{E[(W^\gamma - 1)^2]}{[E(W^\gamma - 1)]^2}.$$

Then g is a strictly increasing function.

To prove this result, we observe that

$$\frac{g'(\gamma)}{2} = \frac{E[W^\gamma (W^\gamma - 1)\log(W)]}{[E(W^\gamma - 1)]^2} - \frac{E[(W^\gamma - 1)^2]\, E[W^\gamma \log(W)]}{[E(W^\gamma - 1)]^3}.$$

It suffices to verify that $g' > 0$ or, equivalently, that

$$E[W^{\gamma}-1]\; E[W^{\gamma}(W^{\gamma}-1)\log(W)] > E[(W^{\gamma}-1)^2]\; E[W^{\gamma}\log(W)] \tag{B.2}$$

for $\gamma > 0$. Set $V = W^{\gamma}$. After noting that $\log(W^{\gamma}) = \gamma \log(W)$, we see that (B.2) is equivalent to

$$E[V-1]\; E[V(V-1)\log(V)] > E[(V-1)^2]\; E[V\log(V)],$$

which we can rewrite as

$$E\left[\frac{V-1}{V\log(V)}\,\frac{V\log(V)}{E[V\log(V)]}\right] E\left[V-1\,\frac{V\log(V)}{E[V\log(V)]}\right] > E\left[\frac{(V-1)^2}{V\log(V)}\,\frac{V\log(V)}{E[V\log(V)]}\right]. \tag{B.3}$$

Let P denote the distribution of V, let P^* be the distribution on $\mathbb{R}$ defined by

$$P^*(dv) = \frac{v\log(v)}{E[V\log(V)]}\,P(dv),$$

and let U be a random variable having distribution P^*. Then (B.3) can be written as

$$E\left[\frac{U-1}{U\log(U)}\right] E(U-1) > E\left[\frac{(U-1)^2}{U\log(U)}\right]. \tag{B.4}$$

But (B.4) follows from Schwarz's inequality, provided we can show that $U-1$ is not, with probability one, a constant multiple of

$$\frac{U-1}{U\log(U)}.$$

To this end, it is enough to verify that the function h on $(1, \infty)$ defined by

$$h(u) = \frac{u\log(u)}{(u-1)^2}$$

is strictly decreasing. But

$$h'(u) = \frac{1+\log(u)}{(u-1)^2} - \frac{2u\log(u)}{(u-1)^2},$$

so it suffices to show that

$$(u-1)[1+\log(u)] < 2u\log(u)$$

for $u > 1$ or, equivalently, that

$$u\log(u) + \log(u) - u + 1 > 0 \tag{B.5}$$

for $u > 1$. But (B.5) is clearly valid, since the function defined by the left side of (B.5) equals zero at $u = 1$ and its derivative is positive on $(1, \infty)$.

For more general results along these lines, see Breiman, et al. (1979).

REFERENCES

Boos, D. D. (1984), "Using Extreme Value Theory to Estimate Large Percentiles," *Technometrics*, 26, 33–39.

Breiman, L., Gins, J. D., and Stone, C. J. (1978), "Statistical Analysis and Interpretation of Peak Air Pollution Measurements," TSC-PD-A190-10, Technology Service Corporation, Santa Monica, California.

Breiman, L., Stone, C. J., and Gins, J. D. (1979), "New Methods for Estimating Tail Probabilities and Extreme Value Distributions," TSC-PD-A226-1, Santa Monica, CA: Technology Service Corporation.

Breiman, L., Stone, C. J., and Gins, J. D. (1981), "Further Development of New Methods for Estimating Tail Probabilities and Extreme Value Distributions," TSC-PD-A243-1, Santa Monica, CA: Technology Service Corporation.

Crager, M. R. (1982), "Exponential Tail Quantile Estimators for Air Quality Data," Technical Report Nos. 4 (Part I) and 5 (Part II), San Francisco, CA: Bay Area Air Quality Management District.

Csörgő, S., Deheuvels, P. and Mason, D. (1985), "Kernel Estimates of the Tail Index of a Distribution," *The Annals of Statistics*, 13, 1050–1077.

David, H. A. (1981), *Order Statistics* (2nd ed.), New York: John Wiley.

Davis, R. and Resnick, S. (1984), "Tail Estimates Motivated by Extreme Value Theory," *The Annals of Statistics*, 12, 1467–1487.

DuMouchel, W. H. (1983), "Estimating the Stable Index α in Order to Measure Tail Thickness: a Critique," *The Annals of Statistics*, 11, 1019–1031.

DuMouchel, W. H. and Olshen, R. A. (1975), "On the Distribution of Claims Costs," in *Credibility*, ed. P. M. Kahn, New York: Academic Press, pp. 23–46.

Galambos, J. (1978), *The Asymptotic Theory of Extreme Order Statistics*, New York: John Wiley.

Hall, P. and Welsh, A. H. (1985), "Adaptive Estimates of Parameters of Regular Variation," *The Annals of Statistics*, 13, 331–341.

Hill, B. M. (1975), "A Simple General Approach to Inference about the Tail of a Distribution," *The Annals of Statistics*, 3, 1163–1174.

Maritz, J. S. and Munro, A. H. (1967), "On the Use of the Generalized Extreme-value Distribution in Extinating Extreme Percentiles," *Biometrics*, 23, 79–103.

Pickands, J. III (1975), "Statistical Inference Using Extreme Order Statistics," *The Annals of Statistics*, 3, 119–131.

Shapiro, S. S., and Wilk, M. B. (1972), "An Analysis of Variance Test for the Exponential Distribution (Complete Samples)," *Technometrics*, 20, 33–35.

Smith, R. L. (1987), "Estimating Tails of Probability Distributions," *The Annals of Statistics*, 15, 1174–1207.

Stone, C. J. and Koo C.-Y. (1986), "Logspline Density Estimation," *Contemporary Mathematics*, 59, 1–15. Providence, RI: American Mathematical Society.

Weinstein, S. B. (1973), "Theory and Application of Some Classical and Generalized Asymptotic Distributions of Extreme Values," *IEEE Transactions on Information Theory*, IT–19, 148–154.

Weissman, I. (1978), "Estimation of Parameters and Large Quantiles Based on the *k* Largest Observations," *Journal of the American Statistical Association*, 73, 812–815.

TECHNICAL REPORTS

Statistics Department

University of California, Berkeley

1. BREIMAN, L. and FREEDMAN, D. (Nov. 1981, revised Feb. 1982). How many variables should be entered in a regression equation? Jour. Amer. Statist. Assoc., March 1983, 78, No. 381, 131-136.

2. BRILLINGER, D. R. (Jan. 1982). Some contrasting examples of the time and frequency domain approaches to time series analysis. Time Series Methods in Hydrosciences, (A. H. El-Shaarawi and S. R. Esterby, eds.) Elsevier Scientific Publishing Co., Amsterdam, 1982, pp. 1-15.

3. DOKSUM, K. A. (Jan. 1982). On the performance of estimates in proportional hazard and log-linear models. Survival Analysis, (John Crowley and Richard A. Johnson, eds.) IMS Lecture Notes - Monograph Series, (Shanti S. Gupta, series ed.) 1982, 74-84.

4. BICKEL, P. J. and BREIMAN, L. (Feb. 1982). Sums of functions of nearest neighbor distances, moment bounds, limit theorems and a goodness of fit test. Ann. Prob., Feb. 1982, 11. No. 1, 185-214.

5. BRILLINGER, D. R. and TUKEY, J. W. (March 1982). Spectrum estimation and system identification relying on a Fourier transform. The Collected Works of J. W. Tukey, vol. 2, Wadsworth, 1985, 1001-1141.

6. BERAN, R. (May 1982). Jackknife approximation to bootstrap estimates. Ann. Statist., March 1984, 12 No. 1, 101-118.

7. BICKEL, P. J. and FREEDMAN, D. A. (June 1982). Bootstrapping regression models with many parameters. Lehmann Festschrift, (P. J. Bickel, K. Doksum and J. L. Hodges, Jr., eds.) Wadsworth Press, Belmont, 1983, 28-48.

8. BICKEL, P. J. and COLLINS, J. (March 1982). Minimizing Fisher information over mixtures of distributions. Sankhyā, 1983, 45, Series A, Pt. 1, 1-19.

9. BREIMAN, L. and FRIEDMAN, J. (July 1982). Estimating optimal transformations for multiple regression and correlation.

10. FREEDMAN, D. A. and PETERS, S. (July 1982, revised Aug. 1983). Bootstrapping a regression equation: some empirical results. JASA, 1984, 79, 97-106.

11. EATON, M. L. and FREEDMAN, D. A. (Sept. 1982). A remark on adjusting for covariates in multiple regression.

12. BICKEL, P. J. (April 1982). Minimax estimation of the mean of a mean of a normal distribution subject to doing well at a point. Recent Advances in Statistics, Academic Press, 1983.

14. FREEDMAN, D. A., ROTHENBERG, T. and SUTCH, R. (Oct. 1982). A review of a residential energy end use model.

15. BRILLINGER, D. and PREISLER, H. (Nov. 1982). Maximum likelihood estimation in a latent variable problem. Studies in Econometrics, Time Series, and Multivariate Statistics, (eds. S. Karlin, T. Amemiya, L. A. Goodman). Academic Press, New York, 1983, pp. 31-65.

16. BICKEL, P. J. (Nov. 1982). Robust regression based on infinitesimal neighborhoods. Ann. Statist., Dec. 1984, 12, 1349-1368.

17. DRAPER, D. C. (Feb. 1983). Rank-based robust analysis of linear models. I. Exposition and review. Statistical Science, 1988, Vol.3 No. 2 239-271.

18. DRAPER, D. C. (Feb 1983). Rank-based robust inference in regression models with several observations per cell.

19. FREEDMAN, D. A. and FIENBERG, S. (Feb. 1983, revised April 1983). Statistics and the scientific method, Comments on and reactions to Freedman, A rejoinder to Fienberg's comments. Springer New York 1985 Cohort Analysis in Social Research, (W. M. Mason and S. E. Fienberg, eds.).

20. FREEDMAN, D. A. and PETERS, S. C. (March 1983, revised Jan. 1984). Using the bootstrap to evaluate forecasting equations. J. of Forecasting. 1985, Vol. 4, 251-262.

21. FREEDMAN, D. A. and PETERS, S. C. (March 1983, revised Aug. 1983). Bootstrapping an econometric model: some empirical results. JBES, 1985, 2, 150-158.

22. FREEDMAN, D. A. (March 1983). Structural-equation models: a case study.

23. DAGGETT, R. S. and FREEDMAN, D. (April 1983, revised Sept. 1983). Econometrics and the law: a case study in the proof of antitrust damages. Proc. of the Berkeley Conference, in honor of Jerzy Neyman and Jack Kiefer. Vol I pp. 123-172. (L. Le Cam, R. Olshen eds.) Wadsworth, 1985.

24. DOKSUM, K. and YANDELL, B. (April 1983). Tests for exponentiality. Handbook of Statistics, (P. R. Krishnaiah and P. K. Sen, eds.) 4, 1984, 579-611.

25. FREEDMAN, D. A. (May 1983). Comments on a paper by Markus.

26. FREEDMAN, D. (Oct. 1983, revised March 1984). On bootstrapping two-stage least-squares estimates in stationary linear models. Ann. Statist., 1984, 12, 827-842.

27. DOKSUM, K. A. (Dec. 1983). An extension of partial likelihood methods for proportional hazard models to general transformation models. Ann. Statist., 1987, 15, 325-345.

28. BICKEL, P. J., GOETZE, F. and VAN ZWET, W. R. (Jan. 1984). A simple analysis of third order efficiency of estimate Proc. of the Neyman-Kiefer Conference, (L. Le Cam, ed.) Wadsworth, 1985.

29. BICKEL, P. J. and FREEDMAN, D. A. Asymptotic normality and the bootstrap in stratified sampling. Ann. Statist. 12 470-482.

30. FREEDMAN, D. A. (Jan. 1984). The mean vs. the median: a case study in 4-R Act litigation. JBES. 1985 Vol 3 pp. 1-13.

31. STONE, C. J. (Feb. 1984). An asymptotically optimal window selection rule for kernel density estimates. Ann. Statist., Dec. 1984, 12, 1285-1297.

32. BREIMAN, L. (May 1984). Nail finders, edifices, and Oz.

33. STONE, C. J. (Oct. 1984). Additive regression and other nonparametric models. Ann. Statist., 1985, 13, 689-705.

34. STONE, C. J. (June 1984). An asymptotically optimal histogram selection rule. Proc. of the Berkeley Conf. in Honor of Jerzy Neyman and Jack Kiefer (L. Le Cam and R. A. Olshen, eds.), II, 513-520.

35. FREEDMAN, D. A. and NAVIDI, W. C. (Sept. 1984, revised Jan. 1985). Regression models for adjusting the 1980 Census. Statistical Science. Feb 1986, Vol. 1, No. 1, 3-39.

36. FREEDMAN, D. A. (Sept. 1984, revised Nov. 1984). De Finetti's theorem in continuous time.

37. DIACONIS, P. and FREEDMAN, D. (Oct. 1984). An elementary proof of Stirling's formula. Amer. Math Monthly. Feb 1986, Vol. 93, No. 2, 123-125.

38. LE CAM, L. (Nov. 1984). Sur l'approximation de familles de mesures par des familles Gaussiennes. Ann. Inst. Henri Poincaré, 1985, 21, 225-287.

39. DIACONIS, P. and FREEDMAN, D. A. (Nov. 1984). A note on weak star uniformities.

40. BREIMAN, L. and IHAKA, R. (Dec. 1984). Nonlinear discriminant analysis via SCALING and ACE.

41. STONE, C. J. (Jan. 1985). The dimensionality reduction principle for generalized additive models.

42. LE CAM, L. (Jan. 1985). On the normal approximation for sums of independent variables.

43. BICKEL, P. J. and YAHAV, J. A. (1985). On estimating the number of unseen species: how many executions were there?

44. BRILLINGER, D. R. (1985). The natural variability of vital rates and associated statistics. Biometrics, to appear.

45. BRILLINGER, D. R. (1985). Fourier inference: some methods for the analysis of array and nonGaussian series data. Water Resources Bulletin, 1985, 21, 743-756.

46. BREIMAN, L. and STONE, C. J. (1985). Broad spectrum estimates and confidence intervals for tail quantiles.

47. DABROWSKA, D. M. and DOKSUM, K. A. (1985, revised March 1987). Partial likelihood in transformation models with censored data. Scandinavian J. Statist., 1988, 15, 1-23.

48. HAYCOCK, K. A. and BRILLINGER, D. R. (November 1985). LIBDRB: A subroutine library for elementary time series analysis.

49. BRILLINGER, D. R. (October 1985). Fitting cosines: some procedures and some physical examples. Joshi Festschrift, 1986. D. Reidel.

50. BRILLINGER, D. R. (November 1985). What do seismology and neurophysiology have in common? - Statistics! Comptes Rendus Math. Rep. Acad. Sci. Canada. January, 1986.

51. COX, D. D. and O'SULLIVAN, F. (October 1985). Analysis of penalized likelihood-type estimators with application to generalized smoothing in Sobolev Spaces.

52. O'SULLIVAN, F. (November 1985). A practical perspective on ill-posed inverse problems: A review with some new developments. To appear in Journal of Statistical Science.

53. LE CAM, L. and YANG, G. L. (November 1985, revised March 1987). On the preservation of local asymptotic normality under information loss.

54. BLACKWELL, D. (November 1985). Approximate normality of large products.

55. FREEDMAN, D. A. (June 1987). As others see us: A case study in path analysis. Journal of Educational Statistics. 12, 101-128.

56. LE CAM, L. and YANG, G. L. (January 1986). Replaced by No. 68.

57. LE CAM, L. (February 1986). On the Bernstein - von Mises theorem.

58. O'SULLIVAN, F. (January 1986). Estimation of Densities and Hazards by the Method of Penalized likelihood.

59. ALDOUS, D. and DIACONIS, P. (February 1986). Strong Uniform Times and Finite Random Walks.

60. ALDOUS, D. (March 1986). On the Markov Chain simulation Method for Uniform Combinatorial Distributions and Simulated Annealing.

61. CHENG, C-S. (April 1986). An Optimization Problem with Applications to Optimal Design Theory.

62. CHENG, C-S., MAJUMDAR, D., STUFKEN, J. & TURE, T. E. (May 1986, revised Jan 1987). Optimal step type design for comparing test treatments with a control.

63. CHENG, C-S. (May 1986, revised Jan. 1987). An Application of the Kiefer-Wolfowitz Equivalence Theorem.

64. O'SULLIVAN, F. (May 1986). Nonparametric Estimation in the Cox Proportional Hazards Model.

65. ALDOUS, D. (JUNE 1986). Finite-Time Implications of Relaxation Times for Stochastically Monotone Processes.

66. PITMAN, J. (JULY 1986, revised November 1986). Stationary Excursions.

67. DABROWSKA, D. and DOKSUM, K. (July 1986, revised November 1986). Estimates and confidence intervals for median and mean life in the proportional hazard model with censored data. Biometrika, 1987, 74, 799-808.

68. LE CAM, L. and YANG, G.L. (July 1986). Distinguished Statistics, Loss of information and a theorem of Robert B. Davies (Fourth edition).

69. STONE, C.J. (July 1986). Asymptotic properties of logspline density estimation.

71. BICKEL, P.J. and YAHAV, J.A. (July 1986). Richardson Extrapolation and the Bootstrap.

72. LEHMANN, E.L. (July 1986). Statistics - an overview.

73. STONE, C.J. (August 1986). A nonparametric framework for statistical modelling.

74. BIANE, PH. and YOR, M. (August 1986). A relation between Lévy's stochastic area formula, Legendre polynomial, and some continued fractions of Gauss.

75. LEHMANN, E.L. (August 1986, revised July 1987). Comparing Location Experiments.

76. O'SULLIVAN, F. (September 1986). Relative risk estimation.

77. O'SULLIVAN, F. (September 1986). Deconvolution of episodic hormone data.

78. PITMAN, J. & YOR, M. (September 1987). Further asymptotic laws of planar Brownian motion.

79. FREEDMAN, D.A. & ZEISEL, H. (November 1986). From mouse to man: The quantitative assessment of cancer risks. To appear in Statistical Science.

80. BRILLINGER, D.R. (October 1986). Maximum likelihood analysis of spike trains of interacting nerve cells.

81. DABROWSKA, D.M. (November 1986). Nonparametric regression with censored survival time data.

82. DOKSUM, K.J. and LO, A.Y. (Nov 1986, revised Aug 1988). Consistent and robust Bayes Procedures for Location based on Partial Information.

83. DABROWSKA, D.M., DOKSUM, K.A. and MIURA, R. (November 1986). Rank estimates in a class of semiparametric two-sample models.

84. BRILLINGER, D. (December 1986). Some statistical methods for random process data from seismology and neurophysiology.

85. DIACONIS, P. and FREEDMAN, D. (December 1986). A dozen de Finetti-style results in search of a theory. Ann. Inst. Henri Poincaré, 1987, 23, 397-423.

86. DABROWSKA, D.M. (January 1987). Uniform consistency of nearest neighbour and kernel conditional Kaplan - Meier estimates.

87. FREEDMAN, D.A., NAVIDI, W. and PETERS, S.C. (February 1987). On the impact of variable selection in fitting regression equations.

88. ALDOUS, D. (February 1987, revised April 1987). Hashing with linear probing, under non-uniform probabilities.

89. DABROWSKA, D.M. and DOKSUM, K.A. (March 1987, revised January 1988). Estimating and testing in a two sample generalized odds rate model. J. Amer. Statist. Assoc., 1988, 83, 744-749.

90. DABROWSKA, D.M. (March 1987). Rank tests for matched pair experiments with censored data.

91. DIACONIS, P and FREEDMAN, D.A. (April 1988). Conditional limit theorems for exponential families and finite versions of de Finetti's theorem. To appear in the Journal of Applied Probability.

92. DABROWSKA, D.M. (April 1987, revised September 1987). Kaplan-Meier estimate on the plane.

92a. ALDOUS, D. (April 1987). The Harmonic mean formula for probabilities of Unions: Applications to sparse random graphs.

93. DABROWSKA, D.M. (June 1987, revised Feb 1988). Nonparametric quantile regression with censored data.

94. DONOHO, D.L. & STARK, P.B. (June 1987). Uncertainty principles and signal recovery.

95. CANCELLED

96. BRILLINGER, D.R. (June 1987). Some examples of the statistical analysis of seismological data. To appear in Proceedings, Centennial Anniversary Symposium, Seismographic Stations, University of California, Berkeley.

97. FREEDMAN, D.A. and NAVIDI, W. (June 1987). On the multi-stage model for carcinogenesis. To appear in Environmental Health Perspectives.

98. O'SULLIVAN, F. and WONG, T. (June 1987). Determining a function diffusion coefficient in the heat equation.

99. O'SULLIVAN, F. (June 1987). Constrained non-linear regularization with application to some system identification problems.

100. LE CAM, L. (July 1987, revised Nov 1987). On the standard asymptotic confidence ellipsoids of Wald.

101. DONOHO, D.L. and LIU, R.C. (July 1987). Pathologies of some minimum distance estimators. Annals of Statistics, June, 1988.

102. BRILLINGER, D.R., DOWNING, K.H. and GLAESER, R.M. (July 1987). Some statistical aspects of low-dose electron imaging of crystals.

103. LE CAM, L. (August 1987). Harald Cramér and sums of independent random variables.

104. DONOHO, A.W., DONOHO, D.L. and GASKO, M. (August 1987). Macspin: Dynamic graphics on a desktop computer. IEEE Computer Graphics and applications, June, 1988.

105. DONOHO, D.L. and LIU, R.C. (August 1987). On minimax estimation of linear functionals.

106. DABROWSKA, D.M. (August 1987). Kaplan-Meier estimate on the plane: weak convergence, LIL and the bootstrap.

107. CHENG, C-S. (Aug 1987, revised Oct 1988). Some orthogonal main-effect plans for asymmetrical factorials.

108. CHENG, C-S. and JACROUX, M. (August 1987). On the construction of trend-free run orders of two-level factorial designs.

109. KLASS, M.J. (August 1987). Maximizing $E \max_{1 \le k \le n} S_k^+ / ES_n^+$: A prophet inequality for sums of I.I.D. mean zero variates.

110. DONOHO, D.L. and LIU, R.C. (August 1987). The "automatic" robustness of minimum distance functionals. Annals of Statistics, June, 1988.

111. BICKEL, P.J. and GHOSH, J.K. (August 1987, revised June 1988). A decomposition for the likelihood ratio statistic and the Bartlett correction — a Bayesian argument.

112. BURDZY, K., PITMAN, J.W. and YOR, M. (September 1987). Some asymptotic laws for crossings and excursions.

113. ADHIKARI, A. and PITMAN, J. (September 1987). The shortest planar arc of width 1.

114. RITOV, Y. (September 1987). Estimation in a linear regression model with censored data.

115. BICKEL, P.J. and RITOV, Y. (Sept. 1987, revised Aug 1988). Large sample theory of estimation in biased sampling regression models I.

116. RITOV, Y. and BICKEL, P.J. (Sept.1987, revised Aug. 1988). Achieving information bounds in non and semiparametric models.

117. RITOV, Y. (October 1987). On the convergence of a maximal correlation algorithm with alternating projections.

118. ALDOUS, D.J. (October 1987). Meeting times for independent Markov chains.

119. HESSE, C.H. (October 1987). An asymptotic expansion for the mean of the passage-time distribution of integrated Brownian Motion.

120. DONOHO, D. and LIU, R. (Oct. 1987, revised Mar. 1988, Oct. 1988). Geometrizing rates of convergence, II.

121. BRILLINGER, D.R. (October 1987). Estimating the chances of large earthquakes by radiocarbon dating and statistical modelling. To appear in Statistics a Guide to the Unknown.

122. ALDOUS, D., FLANNERY, B. and PALACIOS, J.L. (November 1987). Two applications of urn processes: The fringe analysis of search trees and the simulation of quasi-stationary distributions of Markov chains.

123. DONOHO, D.L., MACGIBBON, B. and LIU, R.C. (Nov.1987, revised July 1988). Minimax risk for hyperrectangles.

124. ALDOUS, D. (November 1987). Stopping times and tightness II.

125. HESSE, C.H. (November 1987). The present state of a stochastic model for sedimentation.

126. DALANG, R.C. (December 1987, revised June 1988). Optimal stopping of two-parameter processes on nonstandard probability spaces.

127. Same as No. 133.

128. DONOHO, D. and GASKO, M. (December 1987). Multivariate generalizations of the median and trimmed mean II.

129. SMITH, D.L. (December 1987). Exponential bounds in Vapnik-Červonenkis classes of index 1.

130. STONE, C.J. (Nov.1987, revised Sept. 1988). Uniform error bounds involving logspline models.

131. Same as No. 140

132. HESSE, C.H. (December 1987). A Bahadur - Type representation for empirical quantiles of a large class of stationary, possibly infinite - variance, linear processes

133. DONOHO, D.L. and GASKO, M. (December 1987). Multivariate generalizations of the median and trimmed mean, I.

134. DUBINS, L.E. and SCHWARZ, G. (December 1987). A sharp inequality for martingales and stopping-times.

135. FREEDMAN, D.A. and NAVIDI, W. (December 1987). On the risk of lung cancer for ex-smokers.

136. LE CAM, L. (January 1988). On some stochastic models of the effects of radiation on cell survival.

137. DIACONIS, P. and FREEDMAN, D.A. (April 1988). On the uniform consistency of Bayes estimates for multinomial probabilities.

137a. DONOHO, D.L. and LIU, R.C. (1987). Geometrizing rates of convergence, I.

138. DONOHO, D.L. and LIU, R.C. (January 1988). Geometrizing rates of convergence, III.

139. BERAN, R. (January 1988). Refining simultaneous confidence sets.

140. HESSE, C.H. (December 1987). Numerical and statistical aspects of neural networks.

141. BRILLINGER, D.R. (January 1988). Two reports on trend analysis: a) An Elementary Trend Analysis of Rio Negro Levels at Manaus, 1903-1985 b) Consistent Detection of a Monotonic Trend Superposed on a Stationary Time Series

142. DONOHO, D.L. (Jan. 1985, revised Jan. 1988). One-sided inference about functionals of a density.

143. DALANG, R.C. (Feb. 1988, revised Nov. 1988). Randomization in the two-armed bandit problem.

144. DABROWSKA, D.M., DOKSUM, K.A. and SONG, J.K. (February 1988). Graphical comparisons of cumulative hazards for two populations.

145. ALDOUS, D.J. (February 1988). Lower bounds for covering times for reversible Markov Chains and random walks on graphs.

146. BICKEL, P.J. and RITOV, Y. (Feb.1988, revised August 1988). Estimating integrated squared density derivatives.

147. STARK, P.B. (March 1988). Strict bounds and applications.

148. DONOHO, D.L. and STARK, P.B. (March 1988). Rearrangements and smoothing.

149. NOLAN, D. (March 1988). Asymptotics for a multivariate location estimator.

150. SEILLIER, F. (March 1988). Sequential probability forecasts and the probability integral transform.

151. NOLAN, D. (March 1988). Limit theorems for a random convex set.

152. DIACONIS, P. and FREEDMAN, D.A. (April 1988). On a theorem of Kuchler and Lauritzen.

153. DIACONIS, P. and FREEDMAN, D.A. (April 1988). On the problem of types.

154. DOKSUM, K.A. (May 1988). On the correspondence between models in binary regression analysis and survival analysis.

155. LEHMANN, E.L. (May 1988). Jerzy Neyman, 1894-1981.

156. ALDOUS, D.J. (May 1988). Stein's method in a two-dimensional coverage problem.

157. FAN, J. (June 1988). On the optimal rates of convergence for nonparametric deconvolution problem.

158. DABROWSKA, D. (June 1988). Signed-rank tests for censored matched pairs.

159. BERAN, R.J. and MILLAR, P.W. (June 1988). Multivariate symmetry models.

160. BERAN, R.J. and MILLAR, P.W. (June 1988). Tests of fit for logistic models.

161. BREIMAN, L. and PETERS, S. (June 1988). Comparing automatic bivariate smoothers (A public service enterprise).

162. FAN, J. (June 1988). Optimal global rates of convergence for nonparametric deconvolution problem.

163. DIACONIS, P. and FREEDMAN, D.A. (June 1988). A singular measure which is locally uniform. (Revised by Tech Report No. 180).

164. BICKEL, P.J. and KRIEGER, A.M. (July 1988). Confidence bands for a distribution function using the bootstrap.

165. HESSE, C.H. (July 1988). New methods in the analysis of economic time series I.

166. FAN, JIANQING (July 1988). Nonparametric estimation of quadratic functionals in Gaussian white noise.

167. BREIMAN, L., STONE, C.J. and KOOPERBERG, C. (August 1988). Confidence bounds for extreme quantiles.

168. LE CAM, L. (August 1988). Maximum likelihood an introduction.

169. BREIMAN, L. (August 1988). Submodel selection and evaluation in regression-The conditional case and little bootstrap.

170. LE CAM, L. (September 1988). On the Prokhorov distance between the empirical process and the associated Gaussian bridge.

171. STONE, C.J. (September 1988). Large-sample inference for logspline models.

172. ADLER, R.J. and EPSTEIN, R. (September 1988). Intersection local times for infinite systems of planar brownian motions and for the brownian density process.

173. MILLAR, P.W. (October 1988). Optimal estimation in the non-parametric multiplicative intensity model.

174. YOR, M. (October 1988). Interwinings of Bessel processes.

175. ROJO, J. (October 1988). On the concept of tail-heaviness.

176. ABRAHAMS, D.M. and RIZZARDI, F. (September 1988). BLSS - The Berkeley interactive statistical system: An overview.

177. MILLAR, P.W. (October 1988). Gamma-funnels in the domain of a probability, with statistical implications.

178. DONOHO, D.L. and LIU, R.C. (October 1988). Hardest one-dimensional subfamilies.

179. DONOHO, D.L. and STARK, P.B. (October 1988). Recovery of sparse signals from data missing low frequencies.

180. FREEDMAN, D.A. and PITMAN, J.A. (Nov. 1988). A measure which is singular and uniformly locally uniform. (Revision of Tech Report No. 163).

181. DOKSUM, K.A. and HOYLAND, ARNLJOT (Nov. 1988, revised Jan. 1989). A model for step-stress accelerated life testing experiments based on Wiener processes and the inverse Gaussian distribution.

182. DALANG, R.C., MORTON, A. and WILLINGER, W. (November 1988). Equivalent martingale measures and no-arbitrage in stochastic securities market models.

183. BERAN, R. (November 1988). Calibrating prediction regions.

184. BARLOW, M.T., PITMAN, J. and YOR, M. (Dec. 1988). On Walsh's Brownian Motions.

185. DALANG, R.C. and WALSH, J.B. (Dec. 1988). Almost-equivalence of the germ-field Markov property and the sharp Markov property of the Brownian sheet.

186. HESSE, C.H. (Dec. 1988). Level-Crossing of integrated Ornstein-Uhlenbeck processes

187. PITMAN, J.W. (Dec. 1988). Neveu's branching process in Brownian excursions.

188. PITMAN, J.W. (Dec. 1988). The equilibrium law of Brownian extrema.

189. BARLOW, M.T., PITMAN, J.W. and Yor, M. (Jan. 1989). Some extensions of the arcsine law.

190. STARK, P.B. (Dec. 1988). Duality and discretization in linear inverse problems.

191. LEHMANN, E.L. and SCHOLZ, F.W. (Jan. 1989). Ancillarity.

Copies of these Reports plus the most recent additions to the Technical Report series are available from the Statistics Department technical typist in room 379 Evans Hall or may be requested by mail from:

Department of Statistics
University of California
Berkeley, California 94720

Cost: $1 per copy.