Ann. Inst. Statist. Math.
Vol. 52, No. 4, 646–657 (2000)

SEQUENTIAL ESTIMATION OF THE MAXIMUM IN A MODEL FOR CORROSION DATA

ADAM T. MARTINSEK

Department of Statistics, University of Illinois at Urbana-Champaign, 101 Illini Hall, 725 S. Wright Street, Champaign, IL 61820, U.S.A.

(Received January 4, 1999; revised May 24, 1999)

Abstract. One method of monitoring corrosion in an underground storage tank involves placing a sensor in the tank and running it around the tank's interior. As it runs, the sensor records the local thickness of the tank. In this paper we consider the problem of estimating the maximum pit depth by providing a confidence interval that achieves both a specified confidence level and a specified degree of precision. A particular model, the three-parameter beta, is considered, and a stopping rule for determining the sample size is proposed. It is shown that the stopping rule achieves the desired confidence level and precision, asymptotically as the precision requirement becomes increasingly stringent. Moreover, the stopping rule is asymptotically efficient in terms of sample size. The limiting distribution of the stopping rule is derived, and simulation results are presented to supplement the asymptotics with finite sample size behavior.

Key words and phrases: Corrosion data, precise estimation, extreme value theory, stopping rule.

1. Introduction

Monitoring corrosion in underground storage tanks is an important environmental concern, with significant budgetary and public health implications. A relatively new method of assessing the degree of corrosion involves placing a sensor in the tank and running it around the tank's interior. As it runs, the sensor records the local thickness of the tank. The tank's original wall thickness at the time of manufacture is known, and we will denote it by t. Current standards (J. Carnahan, personal communication) call for sampling 15% of the tank's interior surface area and replacing the tank if the current average wall thickness is too small or the current maximum pit depth is too great (i.e., the minimum wall thickness is too small).

The 15% sampling requirement is clearly intended to produce sufficiently precise estimates of the average wall thickness and maximum pit depth, yet the precision requirement is not explicitly stated. Whatever the precision requirement may be, there is no guarantee that it will be achieved using the 15% sampling rate, or any preassigned sampling rate. Moreover, even if the preassigned sampling rate happens to achieve the desired degree of precision, it is highly unlikely that it does so efficiently, i.e., with as few observations as possible. Both precision and efficiency are important considerations. Imprecise estimates may result in needless and expensive replacement of a good tank, or else leaving in place a tank that presents a significant threat of near-term leakage. On the other hand, running the sensor involves significant expense and one would like to minimize this expense, subject to the precision requirement. A 15% sampling rate can produce sample sizes well over 100,000.

In this paper we will consider precise estimation of the maximum, since this problem has received much less attention in the literature than the analogous problem for the population mean. We will formulate a stopping rule that is designed to provide a sufficiently precise confidence interval for the maximum pit depth. The efficiency of this procedure will also be examined.

One approach to the problem of estimating the maximum pit depth is to assume that one has a sample of maxima and to use the sample to estimate the form and parameters of the limiting distribution of the maximum. This approach has been used successfully in many areas of application, for example in estimating the distribution of record flood levels of a river. However, in the present situation only one maximum will be observed in a given tank. While one could subdivide the observations into batches and then consider the set of batch maxima as a sample, this seems rather artificial (note that in the flood level case, by contrast, it is natural to group the observations by year). Instead, we will assume a parametric model for the pit depths, derive the limiting distribution from it, and then estimate the parameters of the model sequentially as sampling proceeds.

Let Y_i denote the i-th pit depth (i.e., original wall thickness minus current wall thickness) recorded by the sensor. Because the original wall thickness t is known, we will work with the random variables $X_i = Y_i/t$, which represent the pit depths as a fraction of the original wall thickness. We will assume as a first approximation that the X_i are independent and identically distributed (i.i.d.), even though in reality some spatial correlation may be present. The support of the X_i is an interval $[0, \theta]$, where $0 \le \theta \le 1$; here θ is the unknown maximum pit depth in the tank. Our goal is to construct a sufficiently precise confidence interval for θ, i.e., a confidence interval I such that the width of I is at most $2d$, where d is a prespecified positive number, and such that $P(\theta \in I) \approx 1 - \gamma$, where $1 - \gamma$ is the desired confidence level. The idea is to determine θ to within the desired tolerance d and to do so with sufficiently high confidence in the determination.

We will assume that the X_i have a three-parameter beta distribution with density

$$f_{\theta,\alpha,\beta}(x) = C(\alpha, \beta)(x/\theta)^{\alpha-1}(1 - x/\theta)^{\beta-1}/\theta$$

for $0 \le x \le \theta$, where $\alpha, \beta > 0$ and $C(\alpha, \beta) = \Gamma(\alpha + \beta)/\Gamma(\alpha)\Gamma(\beta)$. This is a fairly flexible family for modeling the pit depths: as α and β vary, a wide variety of density shapes and endpoint behavior can be achieved. Both α and β will be assumed to be unknown. Previous research on precise point and interval estimation of a maximum has been carried out by Graybill and Connell (1964), Ghosh and Mukhopadhyay (1975), and Alvo (1978), all of whom consider the special case when $\alpha = \beta = 1$ (i.e., the uniform distribution) and by Mukhopadhyay *et al.* (1983), who investigate the case when $\beta = 1$ and α is known (a "power distribution"). In related work, Basawa *et al.* (1990) consider sequential estimation of an autoregressive parameter using a first order statistic.

Because maximum likelihood estimation for the three-parameter beta distribution is problematic (note that the likelihood is neither everywhere differentiable nor monotone in θ), we will estimate θ by the sample maximum $M_n = \max(X_1, \ldots, X_n)$ and subject to this choice we will estimate α and β by the method of moments. In order to formulate a sensible procedure for precise estimation of θ, it is first necessary to know the form of the limiting distribution of M_n. This can be obtained using well known asymptotic results for maxima (see Galambos (1978) and Leadbetter *et al.* (1983) and the numerous references they contain). In what follows, let $F_{\theta,\alpha,\beta}$ denote the distribution function of the three-parameter beta distribution and let $F_{\alpha,\beta}$ denote this distribution function when $\theta = 1$ (i.e., the distribution function of a standard two-parameter beta distribution on [0, 1]). Using l'Hôpital's Rule we have for $x > 0$,

$$(1.1)\quad \lim_{t\to\infty}[1 - F_{\theta,\alpha,\beta}(\theta - (tx)^{-1})]/[1 - F_{\theta,\alpha,\beta}(\theta - t^{-1})]$$
$$= \lim_{t\to\infty}[1 - F_{\alpha,\beta}(1 - (\theta tx)^{-1})]/[1 - F_{\alpha,\beta}(1 - (\theta t)^{-1})]$$
$$= \lim_{t\to\infty}\left[\int_{1-(\theta tx)^{-1}}^{1} w^{\alpha-1}(1-w)^{\beta-1}dw\right] \Big/ \left[\int_{1-(\theta t)^{-1}}^{1} w^{\alpha-1}(1-w)^{\beta-1}dw\right]$$
$$= \lim_{t\to\infty}[-(1-(\theta tx)^{-1})^{\alpha-1}(\theta tx)^{1-\beta}(\theta x)^{-1}t^{-2}]/[-(1-(\theta t)^{-1})^{\alpha-1}(\theta t)^{1-\beta}\theta^{-1}t^{-2}]$$
$$= x^{-\beta}.$$

It follows from (1.1) and Theorem 2.1.2 of Galambos (1978) that as $n \to \infty$,

$$P[(M_n - \theta)/b_n \le y] \to e^{-(-y)^\beta},$$

for $y < 0$, where $b_n = \theta - \theta F_{\alpha,\beta}^{-1}(1 - 1/n)$. This is an example of a Type III extreme value distribution. Another application of l'Hôpital's Rule yields

$$(1.2)\quad \lim_{x\to 1}[1 - F_{\alpha,\beta}(x)]/[C(\alpha,\beta)x^{\alpha-1}(1-x)^\beta]$$
$$= \lim_{x\to 1}\left[\int_x^1 w^{\alpha-1}(1-w)^{\beta-1}dw\right] \Big/ [x^{\alpha-1}(1-x)^\beta]$$
$$= \lim_{x\to 1}[-x^{\alpha-1}(1-x)^{\beta-1}]/[(\alpha-1)x^{\alpha-2}(1-x)^\beta - \beta x^{\alpha-1}(1-x)^{\beta-1}] = 1/\beta.$$

Since $1 - b_n/\theta \to 1$ as $n \to \infty$, (1.2) implies

$$1 = \lim_{n\to\infty}(1/n)/[C(\alpha,\beta)(1-b_n/\theta)^{\alpha-1}(b_n/\theta)^\beta/\beta]$$
$$= \lim_{n\to\infty}(1/n)/[C(\alpha,\beta)(b_n/\theta)^\beta/\beta].$$

Hence we can take

$$b_n = \theta[\beta/C(\alpha,\beta)]^{1/\beta}n^{-1/\beta}$$

and therefore

$$(1.3)\quad P\left[\frac{n^{1/\beta}(M_n - \theta)}{\theta[\beta/C(\alpha,\beta)]^{1/\beta}} \le y\right] \to e^{-(-y)^\beta},$$

for $y < 0$. Note that the rate of convergence depends on β in the way one would expect: the larger β is, the flatter the density is near θ and the longer one needs to wait to get an observation that is close to θ.

(1.3) can be used to construct confidence intervals for θ, provided that α and β can be estimated consistently. The mean and variance of the three-parameter beta are

$$\mu = \frac{\theta\alpha}{\alpha+\beta}$$

and

$$\sigma^2 = \frac{\theta^2\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}.$$

Because we are estimating θ by M_n, the method of moments estimates of α and β are

$$\tilde{\beta}_n = (M_n - \bar{X}_n)[\bar{X}_n(M_n - \bar{X}_n)/s_n^2 - 1]/M_n$$

and

$$\tilde{\alpha}_n = \bar{X}_n[\bar{X}_n(M_n - \bar{X}_n)/s_n^2 - 1]/M_n,$$

where the sample mean $\bar{X}_n = n^{-1}\sum_{i=1}^n X_i$ estimates μ and the sample variance $s_n^2 = n^{-1}\sum_{i=1}^n (X_i - \bar{X}_n)^2$ estimates σ^2. We will modify these to

$$\hat{\beta}_n = (M_n^* - \bar{X}_n)[\bar{X}_n(M_n^* - \bar{X}_n)/s_n^2 - 1]/M_n^* \tag{1.4}$$

and

$$\hat{\alpha}_n = \bar{X}_n[\bar{X}_n(M_n^* - \bar{X}_n)/s_n^2 - 1]/M_n^*, \tag{1.5}$$

where

$$M_n^* = M_n + M_n[-\log(1/2)\tilde{\beta}_n/C(\tilde{\alpha}_n,\tilde{\beta}_n)]^{1/\tilde{\beta}_n} n^{-1/\tilde{\beta}_n}.$$

M_n^* is an asymptotically median-unbiased version of M_n. It turns out that this modification produces better simulation results than those resulting from use of $\tilde{\beta}_n$ and $\tilde{\alpha}_n$. Because these estimates as well as M_n are consistent and C is continuous, we have

$$P\left[\frac{n^{1/\beta}(M_n - \theta)}{M_n[\hat{\beta}_n/C(\hat{\alpha}_n,\hat{\beta}_n)]^{1/\hat{\beta}_n}} \le y\right] \to e^{-(-y)^\beta} \tag{1.6}$$

for $y < 0$. Moreover, from (1.2), for n sufficiently large and $\delta < 1/\beta$ we have

$$\begin{aligned} P[M_n \le \theta - \theta n^{-\delta}] &= P[M_n/\theta \le 1 - n^{-\delta}] \\ &\le [1 - C(\alpha,\beta)(1 - n^{-\delta})^{\alpha-1} n^{-\delta\beta}/(2\beta)]^n \\ &\le [1 - C(\alpha,\beta) n^{-\delta\beta}/(4\beta)]^n, \end{aligned} \tag{1.7}$$

so that the series

$$\sum_1^\infty P[M_n \le \theta - \theta n^{-\delta}]$$

converges and hence by the Borel-Cantelli Lemma

$$P[M_n \le \theta - \theta n^{-\delta} \text{i.o.}] = 0. \tag{1.8}$$

Using this result and the Law of the Iterated Logarithm for $\bar{X}_n$ and s_n^2, it is easy to check that

$$\hat{\beta}_n - \beta = o(n^{-\epsilon}) \quad \text{a.s.} \tag{1.9}$$

for every $\epsilon \in (0, \min(1/2, 1/\beta))$, and it follows that

$$n^{1/\hat{\beta}_n - 1/\beta} = \exp[\log(n)(\beta - \hat{\beta}_n)/(\beta\hat{\beta}_n)] \to 1 \quad \text{a.s.} \tag{1.10}$$

Combining (1.6) and (1.10) yields

$$P\left[\frac{n^{1/\hat{\beta}_n}(M_n - \theta)}{M_n[\hat{\beta}_n/C(\hat{\alpha}_n,\hat{\beta}_n)]^{1/\hat{\beta}_n}} \le y\right] \to e^{-(-y)^\beta} \tag{1.11}$$

for $y < 0$, and hence for any $\gamma \in (0,1)$, if $0 < \gamma_1 < \gamma_2 \leq 1$ satisfy $\gamma_2 - \gamma_1 = 1 - \gamma$,

$$I_n = [M_n + M_n n^{-1/\hat{\beta}_n}[\hat{\beta}_n/C(\hat{\alpha}_n, \hat{\beta}_n)]^{1/\hat{\beta}_n}(-\log(\gamma_2))^{1/\hat{\beta}_n}, \quad M_n + M_n n^{-1/\hat{\beta}_n}[\hat{\beta}_n/C(\hat{\alpha}_n, \hat{\beta}_n)]^{1/\hat{\beta}_n}(-\log(\gamma_1))^{1/\hat{\beta}_n}] \tag{1.12}$$

is an approximate $100(1-\gamma)\%$ confidence interval for θ. Two special cases are $\gamma_1 = \gamma/2, \gamma_2 = 1 - \gamma/2$ (equal tail probabilities), and $\gamma_1 = \gamma, \gamma_2 = 1$ (lower endpoint M_n, reflecting $M_n \leq \theta$a.s.).

Because the width of the interval in (1.12) is random, there is no nonrandom sample size that can ensure that the width is no larger than $2d$. We must therefore resort to a stopping rule to determine the sample size. Based on (1.12), define the stopping rule T_d by

$$T_d = \text{first } n \geq 2 \text{ such that } M_n[\hat{\beta}_n/C(\hat{\alpha}_n, \hat{\beta}_n)]^{1/\hat{\beta}_n}[(-\log(\gamma_1))^{1/\hat{\beta}_n} - (-\log(\gamma_2))^{1/\hat{\beta}_n}] + n^{-1} \leq 2dn^{1/\hat{\beta}_n}, \tag{1.13}$$

where the n^{-1} term is added to ensure that sampling does not stop too soon (see Chow and Robbins (1965), Chow and Yu (1981) and Sriram (1987) for similar considerations). Then the confidence interval

$$I_{T_d} = [M_{T_d} + M_{T_d}T_d^{-1/\hat{\beta}_{T_d}}[\hat{\beta}_{T_d}/C(\hat{\alpha}_{T_d}, \hat{\beta}_{T_d})]^{1/\hat{\beta}_{T_d}}(-\log(\gamma_2))^{1/\hat{\beta}_{T_d}}, \quad M_{T_d} + M_{T_d}T_d^{-1/\hat{\beta}_{T_d}}[\hat{\beta}_{T_d}/C(\hat{\alpha}_{T_d}, \hat{\beta}_{T_d})]^{1/\hat{\beta}_{T_d}}(-\log(\gamma_1))^{1/\hat{\beta}_{T_d}}] \tag{1.14}$$

has width at most $2d$. It remains to show that the confidence level of this interval is approximately $1-\gamma$ and that the sample size T_d is efficient. These issues are addressed in the following theorem.

THEOREM 1. *Define*

$$n_d^* = (2d)^{-\beta}\theta^\beta[\beta/C(\alpha,\beta)][(-\log(\gamma_1))^{1/\beta} - (-\log(\gamma_2))^{1/\beta}]^\beta.$$

Then as $d \to 0$,

$$T_d/n_d^* \to 1 \quad a.s. \tag{1.15}$$

$$P[\theta \in I_{T_d}] \to 1-\gamma \tag{1.16}$$

and

$$E(T_d)/n_d^* \to 1. \tag{1.17}$$

Remarks.

1. (1.16) states that the proposed procedure achieves the desired confidence level as well as the desired precision, asymptotically as $d \to 0$, i.e., as one requires more and more precision. Rounding n_d^* to the nearest integer yields the smallest n such that $width(I_n)/(2d) \to 1$ a.s. as $d \to 0$. It is therefore the smallest nonrandom sample size that would, if it were known, provide the desired precision and confidence level (asymptotically). (1.15) and (1.17) state that the stopping rule T_d is asymptotically efficient for small values of d in the sense that it is equivalent to the ideal but unavailable n_d^*.

2. One interesting feature of the procedure is that the power of n in the normalizing constant b_n, namely $n^{-1/\beta}$, is unknown and must therefore be estimated. This is different from the usual situation in sequential, and for that matter nonsequential estimation: typically the power of n is known and in most cases it is $n^{-1/2}$, as one would expect.

3. de Haan (1981) gives a general, nonparametric approach to constructing (nonsequential) confidence intervals for the minimum of a function. The present approach is parametric, with the attendant advantages and disadvantages. It makes fuller use of the model assumptions but at the same time is valid only for the three-parameter beta distribution.

It is also of interest to determine the limiting distribution of the stopping time T_d and its rate of convergence to n_d^*. To do this it is first necessary to derive the limiting distribution of $\hat{\beta}_n$. We know from the Central Limit Theorem that

$$n^{1/2}(\bar{X}_n - \mu, s_n^2 - \sigma^2)^T \to_d N(\mathbf{0}, \mathbf{\Sigma}_{\theta,\alpha,\beta}) \tag{1.18}$$

as $n \to \infty$, where the entries in the asymptotic covariance matrix $\mathbf{\Sigma}_{\theta,\alpha,\beta}$ are $\sigma_{11} = \sigma^2$, $\sigma_{12} = \sigma_{21} = \mathrm{Cov}(X_1, (X_1 - \mu)^2)$, $\sigma_{22} = \mathrm{Var}((X_1 - \mu)^2)$. We can write

$$\frac{\bar{X}_n}{M_n^*} - \frac{\mu}{\theta} = \frac{(\bar{X}_n - \mu)}{M_n^*} + \frac{\mu(\theta - M_n^*)}{\theta M_n^*} \tag{1.19}$$

and

$$\frac{s_n^2}{(M_n^*)^2} - \frac{\sigma^2}{\theta^2} = \frac{(s_n^2 - \sigma^2)}{(M_n^*)^2} + \frac{\sigma^2(\theta^2 - (M_n^*)^2)}{\theta^2 (M_n^*)^2}. \tag{1.20}$$

Let H denote a random variable whose distribution is the limiting distribution of $n^{1/\beta}(M_n^* - \theta)$. In what follows we may take H to be independent of the multivariate normal distribution in (1.18) (see Tiago de Oliveira (1961) and Rosengard (1962)). From (1.18)–(1.20),

$$\begin{aligned} &n^{\min(1/2,1/\beta)}\left(\frac{\bar{X}_n}{M_n^*} - \frac{\mu}{\theta}, \frac{s_n^2}{(M_n^*)^2} - \frac{\sigma^2}{\theta^2}\right)^T \\ &\quad \to_d I_{(1/\beta \geq 1/2)} N(\mathbf{0}, \mathbf{\Sigma}_{\mathbf{1},\alpha,\beta}) - I_{(1/\beta \leq 1/2)} \left(\frac{\mu H}{\theta^2}, \frac{2\sigma^2 H}{\theta^3}\right)^T, \end{aligned} \tag{1.21}$$

where I denotes indicator function. Let $\boldsymbol{W}$ denote a random vector with components W_1 and W_2 whose distribution is specified by the right side of (1.21). (1.4) and the delta method yield

$$\begin{aligned} &n^{\min(1/2,1/\beta)}(\hat{\beta}_n - \beta) \\ &\quad \to_d [(\theta - \mu)(\theta - 3\mu)/\sigma^2 + 1]W_1 - [\mu\theta(\mu - \theta)^2/\sigma^4]W_2. \end{aligned} \tag{1.22}$$

Let S denote a random variable with distribution given by the right side of (1.22). Then we have the following theorem, in which the need to estimate the rate of convergence of M_n to θ accounts for the $\log(n_d^*)$ term (see (2.21)–(2.27) below).

THEOREM 2. *As* $d \to 0$,

$$\beta(n_d^*)^{\min(1/2,1/\beta)-1}(T_d - n_d^*)/\log(n_d^*) \to_d S \tag{1.23}$$

The proofs of Theorems 1 and 2 are given in Section 2. Section 3 gives simulation results for the procedure.

2. Proofs

PROOF OF (1.15). From the definition of T_d,

$$M_{T_d}[\hat{\beta}_{T_d}/C(\hat{\alpha}_{T_d},\hat{\beta}_{T_d})]^{1/\hat{\beta}_{T_d}}[(-\log(\gamma_1))^{1/\hat{\beta}_{T_d}} - (-\log(\gamma_2))^{1/\hat{\beta}_{T_d}}] \le 2dT_d^{1/\hat{\beta}_{T_d}}. \tag{2.1}$$

Because $\hat{\alpha}_n \to \alpha$, $\hat{\beta}_n \to \beta$, $M_n \to \theta$ a.s. and $T_d \to \infty$ a.s. as $d \to 0$,

$$\hat{\alpha}_{T_d} \to \alpha, \qquad \hat{\beta}_{T_d} \to \beta, \qquad M_{T_d} \to \theta \text{ a.s.} \tag{2.2}$$

It follows from (2.1) and (2.2) that

$$\liminf_{d\to 0}(dT_d^{1/\hat{\beta}_{T_d}}) \ge \{\theta\beta^{1/\beta}[(-\log(\gamma_1))^{1/\beta} - (-\log(\gamma_2))^{1/\beta}]\}/2C(\alpha,\beta)^{1/\beta} \quad \text{a.s.} \tag{2.3}$$

Similarly, the inequality

$$\begin{aligned} &M_{T_d-1}[\hat{\beta}_{T_d-1}/C(\hat{\alpha}_{T_d-1},\hat{\beta}_{T_d-1})]^{1/\hat{\beta}_{T_d-1}} \\ &\quad\cdot [(-\log(\gamma_1))^{1/\hat{\beta}_{T_d-1}} - (-\log(\gamma_2))^{1/\hat{\beta}_{T_d-1}}] + (T_d-1)^{-1} \ge 2d(T_d-1)^{1/\hat{\beta}_{T_d-1}} \end{aligned} \tag{2.4}$$

yields

$$\begin{aligned} &\limsup_{d\to 0}[d(T_d-1)^{1/\hat{\beta}_{T_d-1}}] \\ &\qquad \le \{\theta\beta^{1/\beta}[(-\log(\gamma_1))^{1/\beta} - (-\log(\gamma_2))^{1/\beta}]\}/2C(\alpha,\beta)^{1/\beta} \text{a.s.} \end{aligned} \tag{2.5}$$

Because $(T_d - 1)/T_d \to 1$ a.s., (1.15) now follows from (2.3), (2.5) and (1.10), noting that since $T_d \to \infty$ a.s. (1.10) holds a.s. with n replaced by T_d or $T_d - 1$.

PROOF OF (1.16). It suffices to show that $T_d^{1/\hat{\beta}_{T_d}}(M_{T_d} - \theta)/M_{T_d}(\hat{\beta}_{T_d}/C(\hat{\alpha}_{T_d}, \hat{\beta}_{T_d}))^{1/\hat{\beta}_{T_d}}$ has the same limiting distribution as $n^{1/\beta}(M_n - \theta)/\theta(\beta/C(\alpha,\beta))^{1/\beta}$, i.e.,

$$P\left[\frac{T_d^{1/\hat{\beta}_{T_d}}(M_{T_d}-\theta)}{M_{T_d}(\hat{\beta}_{T_d}/C(\hat{\alpha}_{T_d},\hat{\beta}_{T_d}))^{1/\hat{\beta}_{T_d}}} \le y\right] \to e^{-(-y)^\beta} \tag{2.6}$$

for $y < 0$. Fix $\epsilon \in (0,1)$. In what follows we will for convenience of notation treat $(1-\epsilon)n_d^*$ and $(1+\epsilon)n_d^*$ as integers. Because $M_n \le M_{n+1} \le \theta$ for all n, we have

$$\begin{aligned} &P\left[\frac{T_d^{1/\beta}(M_{T_d}-\theta)}{\theta(\beta/C(\alpha,\beta))^{1/\beta}} \le y;\ (1-\epsilon)n_d^* \le T_d \le (1+\epsilon)n_d^*\right] \\ &\qquad \le P\left[\frac{(1+\epsilon)^{1/\beta}(n_d^*)^{1/\beta}(M_{(1-\epsilon)n_d^*}-\theta)}{\theta(\beta/C(\alpha,\beta))^{1/\beta}} \le y\right] \\ &\qquad \to e^{-(-y)^\beta(1-\epsilon)/(1+\epsilon)}. \end{aligned} \tag{2.7}$$

Similarly, since $T_d/n_d^* \to 1$ a.s.,

$$(2.8)\ P\left[\frac{T_d^{1/\beta}(M_{T_d}-\theta)}{\theta(\beta/C(\alpha,\beta))^{1/\beta}} \le y;\ (1-\epsilon)n_d^* \le T_d \le (1+\epsilon)n_d^*\right]$$
$$\ge P\left[\frac{(1-\epsilon)^{1/\beta}(n_d^*)^{1/\beta}(M_{(1+\epsilon)n_d^*}-\theta)}{\theta(\beta/C(\alpha,\beta))^{1/\beta}} \le y;\ (1-\epsilon)n_d^* \le T_d \le (1+\epsilon)n_d^*\right]$$
$$\ge P\left[\frac{(1-\epsilon)^{1/\beta}(n_d^*)^{1/\beta}(M_{(1+\epsilon)n_d^*}-\theta)}{\theta(\beta/C(\alpha,\beta))^{1/\beta}} \le y\right] + o(1)$$
$$\to e^{-(-y)^\beta(1+\epsilon)/(1-\epsilon)}.$$

Since

$$P\left[\frac{T_d^{1/\beta}(M_{T_d}-\theta)}{\theta(\beta/C(\alpha,\beta))^{1/\beta}} \le y\right]$$
$$= P\left[\frac{T_d^{1/\beta}(M_{T_d}-\theta)}{\theta(\beta/C(\alpha,\beta))^{1/\beta}} \le y; (1-\epsilon)n_d^* \le T_d \le (1+\epsilon)n_d^*\right] + o(1),$$

it follows that

$$(2.9)\qquad \limsup_{d\to 0} P\left[\frac{T_d^{1/\beta}(M_{T_d}-\theta)}{\theta(\beta/C(\alpha,\beta))^{1/\beta}} \le y\right] \le e^{-(-y)^\beta(1-\epsilon)/(1+\epsilon)}$$

and

$$(2.10)\qquad \liminf_{d\to 0} P\left[\frac{T_d^{1/\beta}(M_{T_d}-\theta)}{\theta(\beta/C(\alpha,\beta))^{1/\beta}} \le y\right] \ge e^{-(-y)^\beta(1+\epsilon)/(1-\epsilon)}.$$

Letting $\epsilon \to 0$ yields

$$P\left[\frac{T_d^{1/\beta}(M_{T_d}-\theta)}{\theta(\beta/C(\alpha,\beta))^{1/\beta}} \le y\right] \to e^{-(-y)^\beta}.$$

Combining this result with $T_d^{-1/\beta+1/\hat\beta_{T_d}} \to 1$, $\hat\alpha_{T_d} \to \alpha$, $\hat\beta_{T_d} \to \beta$, $M_{T_d} \to \theta$ a.s. and continuity of the gamma function finishes the proof of (1.16).

PROOF OF (1.17). In view of (1.15), it suffices to show that $\{d^\beta T_d : d \le 1\}$ is uniformly integrable (u.i.). Let $G(a,b)$ denote

$$[b/C(a,b)]^{1/b}[(-\log(\gamma_1))^{1/b} - (-\log(\gamma_2))^{1/b}]$$

and note that T_d is the first n such that

$$M_n G(\hat\alpha_n, \hat\beta_n) + n^{-1} \le 2dn^{1/\hat\beta_n}.$$

By the Marcinkiewicz-Zygmund inequality, for $\epsilon > 0$, $s \ge 2$, and $0 \le \delta < 1/2$,

$$(2.11)\qquad P[|\bar X_n - \mu| \ge \epsilon n^{-\delta}] \le (\text{const})n^{-s(1/2-\delta)}$$

and

$$P[|s_n^2 - \sigma^2| \geq \epsilon n^{-\delta}] \leq (\text{const}) n^{-s(1/2-\delta)}. \tag{2.12}$$

From (2.11), (2.12) and (1.7) it is easy to show that for $0 \leq \delta < \min(1/2, 1/\beta)$,

$$P[|\hat{\alpha}_n - \alpha| \geq \epsilon n^{-\delta}] \leq (\text{const}) n^{-s(\min(1/2,1/\beta)-\delta)} \tag{2.13}$$

and

$$P[|\hat{\beta}_n - \beta| \geq \epsilon n^{-\delta}] \leq (\text{const}) n^{-s(\min(1/2,1/\beta)-\delta)}. \tag{2.14}$$

Because G has continuous derivatives of all orders at (α, β), it follows that for $0 \leq \delta < \min(1/2, 1/\beta)$,

$$P[|G(\hat{\alpha}_n, \hat{\beta}_n) - G(\alpha, \beta)| \geq \epsilon n^{-\delta}] \leq (\text{const}) n^{-s(\min(1/2,1/\beta)-\delta)}. \tag{2.15}$$

For $\lambda \geq 1$, $d \leq 1$, and $n(d, \lambda) = [\lambda d^{-\beta}]$, where [ ] denotes greatest integer, we have

$$\begin{aligned} P[T_d > n(d,\lambda)] \leq P[T_d > n(d,\lambda), G(\hat{\alpha}_{n(d,\lambda)}, \hat{\beta}_{n(d,\lambda)}) \leq 2G(\alpha,\beta), \\ |1/\hat{\beta}_{n(d,\lambda)} - 1/\beta| < n(d,\lambda)^{-\delta}] \\ + P[G(\hat{\alpha}_{n(d,\lambda)}, \hat{\beta}_{n(d,\lambda)}) > 2G(\alpha,\beta)] \\ + P[|1/\hat{\beta}_{n(d,\lambda)} - 1/\beta| \geq n(d,\lambda)^{-\delta}] \\ = I + II + III, \end{aligned} \tag{2.16}$$

say. It follows from (2.14) and (2.15) that

$$\sup_{d \leq 1} II = O(n(d,\lambda)^{-s \cdot \min(1/2,1/\beta)}) = O(\lambda^{-s \cdot \min(1/2,1/\beta)}) \tag{2.17}$$

and

$$\sup_{d \leq 1} III = O(n(d,\lambda)^{-s(\min(1/2,1/\beta)-\delta)}) = O(\lambda^{-s(\min(1/2,1/\beta)-\delta)}) \tag{2.18}$$

as $\lambda \to \infty$, for every $s \geq 2$ and $0 \leq \delta < \min(1/2, 1/\beta)$. Moreover, since $M_n \leq \theta$ for all n,

$$\sup_{d \leq 1} I \leq P[2\lambda^{1/\beta - n(d,\lambda)^{-\delta}} d^{1-[1-\beta n(d,\lambda)^{-\delta}]} < 2\theta G(\alpha,\beta) + n(d,\lambda)^{-1}]. \tag{2.19}$$

It is easy to check that

$$\lambda^{1/\beta - n(d,\lambda)^{-\delta}} d^{1-[1-\beta n(d,\lambda)^{-\delta}]} \to \infty$$

as $\lambda \to \infty$, uniformly in $d \leq 1$, so we have

$$\sup_{d \leq 1} I = 0 \tag{2.20}$$

for λ sufficiently large. It follows from (2.16)–(2.20) that

$$\{d^\beta T_d : d \leq 1\} \quad \text{is u.i.},$$

proving (1.17).

Remark. Note that this argument actually shows $E(T_d^p)/(n_d^*)^p \to 1$ for all $p > 0$.

PROOF OF THEOREM 2. We have

$$(2.21) \qquad (2d)^\beta T_d^{\beta/\hat\beta_{T_d}} - (2d)^\beta n_d^* \geq M_{T_d}^\beta G^\beta(\hat\alpha_{T_d}, \hat\beta_{T_d}) - \theta^\beta G^\beta(\alpha, \beta)$$

and

$$(2.22) \qquad \begin{aligned} &(2d)^\beta (T_d - 1)^{\beta/\hat\beta_{T_d - 1}} - (2d)^\beta n_d^* \\ &\leq [M_{T_d - 1} G(\hat\alpha_{T_d - 1}, \hat\beta_{T_d - 1}) + (T_d - 1)^{-1}]^\beta - \theta^\beta G^\beta(\alpha, \beta). \end{aligned}$$

It follows from arguments similar to those used to prove (1.16), together with various uniform continuity in probability results (see Woodroofe (1982), Section 1.3), that (1.18)–(1.22) hold with n replaced by either T_d or $T_d - 1$. It follows from this observation and the delta method that

$$(2.23) \quad T_d^{\min(1/2, 1/\beta)}(M_{T_d}^\beta G^\beta(\hat\alpha_{T_d}, \hat\beta_{T_d}) - \theta^\beta G^\beta(\alpha, \beta)) \text{ converges in distribution}$$

and

$$(2.24) \quad (T_d - 1)^{\min(1/2, 1/\beta)}(M_{T_d - 1}^\beta G^\beta(\hat\alpha_{T_d - 1}, \hat\beta_{T_d - 1}) - \theta^\beta G^\beta(\alpha, \beta)) \text{ converges in distribution.}$$

We also have

$$(2.25) \qquad \begin{aligned} T_d - T_d^{\beta/\hat\beta_{T_d}} &= T_d[1 - \exp(\log(T_d)(\beta/\hat\beta_{T_d} - 1))] \\ &\approx T_d \log(T_d)(\hat\beta_{T_d} - \beta)/\hat\beta_{T_d}, \end{aligned}$$

so that

$$(2.26) \qquad \beta T_d^{\min(1/2, 1/\beta) - 1}(T_d - T_d^{\beta/\hat\beta_{T_d}})/\log(T_d) \to_d S.$$

Combining (1.15), (2.21)–(2.24) and (2.26) yields

$$(2.27) \qquad \beta T_d^{\min(1/2, 1/\beta) - 1}(T_d - n_d^*)/\log(T_d) \to_d S.$$

In view of (1.15), this completes the proof of Theorem 2.

3. Simulation results

A small simulation study was carried out to examine the performance of the stopping rule T_d for nonasymptotic values of d. Observations were generated from a three-parameter beta with $\alpha = \beta = 2$ and $\theta = 1$ (note that the value of θ does not actually matter, as the performance is invariant under rescaling). γ was set at 0.05, corresponding to 95% confidence, γ_1 was 0.025, γ_2 was 0.975, and the values of d were 0.0100, 0.0095, 0.0090, 0.0085, 0.0080, 0.0075, 0.0070, 0.0065, 0.0060, 0.0055, 0.0050. This represents a range of desired precision from 1% down to 0.5%. These values seem appropriate given the serious consequences of imprecise estimates. Due to the rather large sample sizes that

Table 1. Coverage frequencies, average sample sizes, and n_d^*, with standard errors in parentheses.

d	Coverage Freq.	Average Sample Size	n_d^*
0.0100	0.8825 (0.016)	2408.82 (34.53)	2586
0.0095	0.8800 (0.016)	2678.81 (33.72)	2865
0.0090	0.8725 (0.017)	3000.35 (35.69)	3192
0.0085	0.8850 (0.016)	3331.33 (40.26)	3579
0.0080	0.9125 (0.014)	3827.40 (40.75)	4040
0.0075	0.8825 (0.016)	4387.94 (48.87)	4597
0.0070	0.8800 (0.016)	5048.52 (50.24)	5277
0.0065	0.9025 (0.015)	5929.11 (53.64)	6120
0.0060	0.9350 (0.012)	6969.06 (52.28)	7183
0.0055	0.8900 (0.016)	8231.75 (65.17)	8548
0.0050	0.9075 (0.014)	10076.48 (72.61)	10343

resulted, only four hundred repetitions were conducted for each value of d. Table 1 summarizes the coverage frequencies, average sample sizes (i.e., estimated expected values of T_d) and ideal sample sizes n_d^*. Standard errors of the estimates are given in parentheses.

Several things stand out from the simulations. The average sample sizes are reasonably close to the n_d^* values, indicating that the stopping rule is doing a good job of choosing the sample size. The average sample size is in all cases below n_d^*, but the discrepancy is small: the ratio of the two quantities ranges from 0.931 to 0.974, with the higher ratios occurring for smaller values of d, as one would predict from Theorem 1. All of the coverage frequencies are below the desired 95% value. They range from 87.25% to 93.5%. Even when the standard error is factored in, it is clear that the actual coverage probability is somewhat lower than the nominal value. Still, the degree of confidence that is attained is fairly high. Preliminary simulations with larger values of d and hence smaller expected sample sizes showed much worse performance; in particular, the method does not seem to work well unless it produces sample sizes in the thousands or tens of thousands. These sample sizes are still well below those that can occur when the 15% sampling rate is used, so the procedure produces definite gains in efficiency.

Acknowledgements

I would like to thank the referees for helpful suggestions that yielded substantial improvements in the paper.

References

Alvo, M. (1978). Sequential estimation of a truncation parameter, *J. Amer. Statist. Assoc.*, **73**, 404–407.

Basawa, I. V., McCormick, W. P. and Sriram, T. N. (1990). Sequential estimation for dependent observations with an application to non-standard autoregressive processes, *Stochastic Process Appl.* **35**, 149–168.

Chow, Y. S. and Robbins, H. (1965). On the asymptotic theory of fixed-width confidence intervals for the mean, *Ann. Math. Statist.*, **36**, 457–462.

Chow, Y. S. and Yu, K. F. (1981). On the performance of a sequential procedure for estimation of the mean, *Ann. Statist.*, **9**, 184–189.

de Haan, L. (1981). Estimation of the minimum of a function using order statistics, *J. Amer. Statist. Assoc.*, **76**, 467–469 (Correction: *ibid.* (1983), **78**, p. 1008).

Galambos, J. (1978). *The Asymptotic Theory of Extreme Order Statistics*, Krieger, Malabar, Florida.

Ghosh, M. and Mukhopadhyay, N. (1975). Sequential point estimation of the parameter of a rectangular distribution, *Calcutta Statist. Assoc. Bull.*, **24**, 117–122.

Graybill, F. A. and Connell, T. L. (1964). Sample size required to estimate the parameter in the uniform density within d units, *J. Amer. Statist. Assoc.*, **59**, 550–556.

Leadbetter, M. R., Lindgren, G. and Rootzén, H. (1983). *Extremes and Related Properties of Random Sequences and Processes*, Springer, New York.

Mukhopadhyay, N., Ghosh, M., Hamdy, H. and Wackerly, D. (1983). Sequential and two-stage point estimation for the range in a power family distribution, *Sequential Anal.*, **2**, 259–288.

Rosengard, A. (1962). Etude des lois-limites jointes et marginales de la moyenne et des valeurs extréme d'un enchantillon, *Publications de l'Institut de Statistique de l'Université de Paris*, **11**, 3–53.

Sriram, T. N. (1987). Sequential estimation of the mean of a first-order stationary autoregressive process, *Ann. Statist.*, **15**, 1079–1090.

Tiago de Oliveira, J. (1961). The asymptotic independence of the sample mean and extremes, *Revista da Faculdade de Ciencias, University Lisbon*, **A8**, 299–309.

Woodroofe, M. (1982). *Nonlinear Renewal Theory in Sequential Analysis*, Society for Industrial and Applied Mathematics, Philadelphia.