Sequential point estimation of the powers of a normal scale parameter

Chikara Uno, Eiichi Isogai

Department of Mathematics, Akita University, Akita 010-8502, Japan
(e-mail: uno@math.akita-u.ac.jp)
Department of Mathematics, Niigata University, Niigata 950-2181, Japan

Received August 2000

Abstract. We consider the sequential point estimation problem of the powers of a normal scale parameter $\sigma^r$ with $r \neq 0$ when the loss function is squared error plus linear cost. It is shown that the regret due to using our fully sequential procedure in ignorance of $\sigma$ is asymptotically minimized for estimating $\sigma^{-2}$. We also propose a bias-corrected procedure to reduce the risk and show that the larger the distance between $r$ and $-2$ is, the more effective our bias-corrected procedure is.

Key words: Second order approximation, fully sequential procedure, regret, uniform integrability, bias-correction

1. Introduction

Let $X_1, X_2, \ldots$ be a sequence of independent observations drawn from a normal population with both mean $\mu \in (-\infty, \infty)$ and standard deviation $\sigma \in (0, \infty)$ unknown. We consider the problem of point estimation of the powers of standard deviation $\sigma^r$ with $r \neq 0$. Suppose we terminate sampling after $n$ observations and estimate $\sigma^r$ by $S_n^r$, where $\bar{X}_n = n^{-1} \sum_{i=1}^{n} X_i$ and $S_n^2 = (n-1)^{-1} \sum_{i=1}^{n} (X_i - \bar{X}_n)^2$. As a measure of loss incurred when sampling terminates after $n$ observations we choose the quantity

$$L_n = (S_n^r - \sigma^r)^2 + cn,$$

where $c > 0$ is the known cost per unit sample. The risk is given by $R_n = E(L_n)$. The goal here is to find an appropriate sample size that will minimize the risk.

As for other loss structures, see, for instance, Chattopadhyay (1998), Takada (2000) and Chattopadhyay et al. (2000).

For $r = 2$, that is, estimating the variance by $S_n^2$, Starr and Woodrofe
(1972) proposed a sequential procedure and Woodroofe (1977) investigated its second order properties. For \( r = 1 \), namely, estimating the standard deviation by \( S_n \), Ghosh and Mukhopadhyay (1989) proposed a fully sequential procedure and gave a second order approximation to the risk. In the sequential point estimation problem, the cost \( c \) is assumed to be sufficiently small, so that we suppose that the sample size \( n \) is sufficiently large. It is well known that \( K_{n-1, r}, S_k^r \) is the uniformly minimum variance unbiased estimator of \( \sigma^r \) provided \( n > \max \{ 1, -2r + 1 \} \), where \( K_{n, r} = (n/2)^{r/2} \Gamma(n/2)/\Gamma((n + r)/2). \) \( K_{n-1, r}, S_k^r \) is approximately equal to \( S_k^r \) for large \( n \) and the calculation of \( S_k^r \) is easier than that of \( K_{n-1, r}. \) For this reason we restrict attention on the estimator \( S_k^r \). As seen on page 85 in Lehmann (1983), sometimes it is of interest to measure \( \mu \) in \( \sigma \)-units and hence to estimate \( \mu/\sigma. \) In fact, Sriram (1990) considered sequential point estimation of ratio \( \mu/\sigma \) and estimated \( 1/\sigma \) by \( S_n^{-1}. \) Further, Takada (1997) constructed sequential confidence intervals for some function of \( \mu \) and \( \sigma^2 \) by restricting attention to the statistic \( S_k^r \) for estimating \( \sigma^r. \)

Now, we shall compute \( E(L_n). \) We have, for \( n > \max \{ 1, -2r + 1 \}, \)

\[
E(S_k^r - \sigma^r)^2 = \sigma^{-2}E[(S_n^r/\sigma^2)^r - 1 - 2((S_n/\sigma)^r - 1)].
\]

By Taylor’s theorem,

\[
\left( \frac{S_n^2}{\sigma^2} \right)^r = 1 + r \left( \frac{S_n^2}{\sigma^2} - 1 \right) + \beta_x \left( \frac{S_n^2}{\sigma^2} - 1 \right)^2 + \gamma_x \left( \frac{S_n^2}{\sigma^2} - 1 \right)^3 \eta_n^{-3} \]

where

\[
\beta_x = \frac{x(x - 1)}{2}, \quad \gamma_x = \frac{x(x - 1)(x - 2)}{6}
\]

and \( \eta_n \) is a random variable such that \( |\eta_n - 1| < |(S_n^2/\sigma^2) - 1| \). From the Hölder inequality with \( u > 1 \) and \( u^{-1} + v^{-1} = 1 \), and the Marcinkiewicz-Zygmund inequality (see Chow and Teicher (1978), Corollary 10.3.2),

\[
E \left| \left( \frac{S_n^2}{\sigma^2} - 1 \right)^{3/2} \right| \leq \left\{ E \left| \frac{S_n^2}{\sigma^2} - 1 \right|^{3/2} \right\}^{1/\nu} \{ E(\eta_n^{-3})^{\nu} \}^{1/\nu}
\]

is \( O(n^{-3/2}) \cdot \{ E(\eta_n^{-3})^{\nu} \}^{1/\nu} \)

as \( n \to \infty. \) Let \( M \) be a generic positive constant. By \( c_r \)-inequality (see Loève (1977), p. 157), for \( r - 3 > 0 \)

\[
E(\eta_n^{-3})^{\nu} \leq M \left\{ 1 + E \left| \frac{S_n^2}{\sigma^2} - 1 \right|^{(r-3)\nu} \right\} = O(1) \quad \text{as } n \to \infty.
\]

By the maximal inequality for a reverse sub-martingale and the fact that \( (n - 1)S_n^2/\sigma^2 \) has the chi-square distribution with \( n - 1 \) degrees of freedom, for every \( q > 1, \)