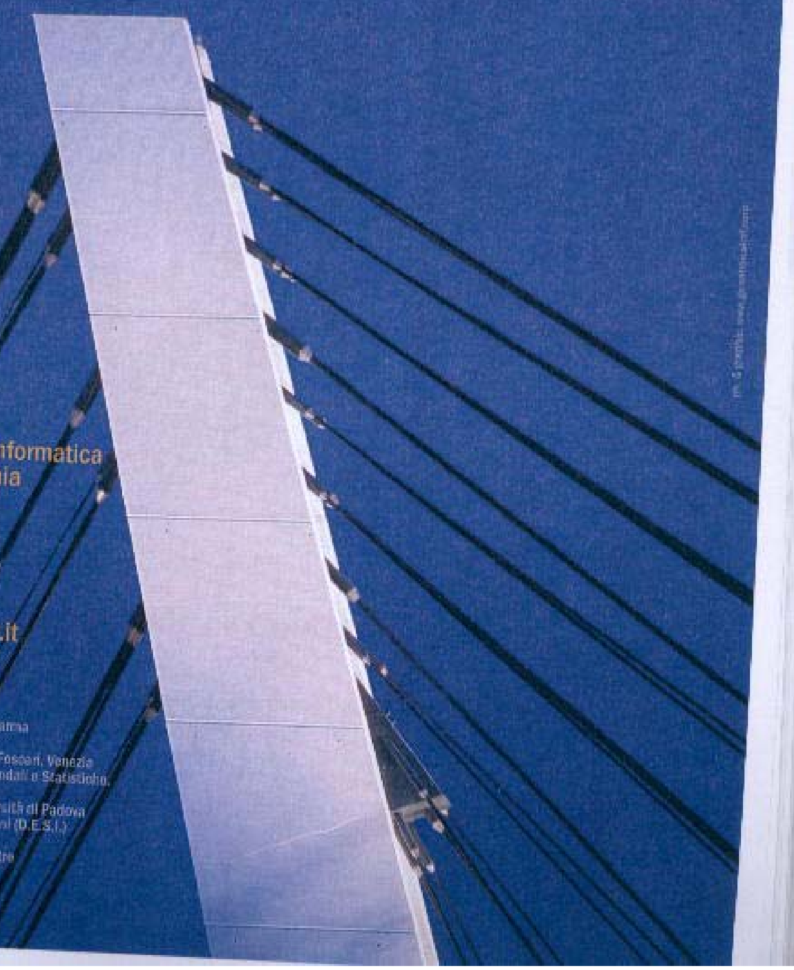
The Minerva Research Foundation
Princeton, New Jersey

icors09 INTERNATIONAL CONFERENCE ON ROBUST STATISTICS

Università degli Studi di Parma
Facoltà di Economia
www.icors2009.unipr.it

14-19 JUNE 2009
Parma, Italy



Sezione di Statistica e Informatica
Dipartimento di Economia
Università di Parma
Via J.F. Kennedy 6,
43100 PARMA - Italia

fax +39 0521 032375
e-mail icors2009@unipr.it

The Minerva Research Foundation
Dipartimento di Economia, Università di Parma
Università degli Studi di Parma
Dipartimento di Statistica, Università Ca' Foscari, Venezia
Dipartimento di Scienze Economiche, Aziendali e Statistiche,
Università di Milano
Dipartimento di Scienze Statistiche, Università di Padova
Dipartimento di Economie Società Istituzioni (D.E.S.I.)
Università di Verona
European Commission, Joint Research Centre
CARIPARMA Crédit Agricole
Comune di PARMA
Institute of Mathematical Statistics

Optimal statistical procedures on the basis of records in a two-parameter exponential distribution

M. Doostparast

[1] Department of Statistics, School of Mathematical Sciences, Ferdowsi University of Mashhad, P. O. Box 91775-1159, Mashhad, Iran

Keywords: Generalized likelihood ratio, Invariant test, LINEX-unbiased estimator, Monotone likelihood ratio, Robustness, Shortest-length confidence interval, Unbiased interval, Uniform minimum variance unbiased estimation, Uniformly most powerful test.

1 Abstract

In data-processing standpoint, an efficient algorithm for identifying the minimum value among a set of measurements are record statistics. In this paper, point estimates are derived for the two parameters of the Exponential distribution based on record statistics with respect to the squared error and Linear-Exponential loss functions and then compared with together. The admissibility of some estimators is discussed. Optimal confidence intervals as well as uniformly most powerful test functions are derived for the two parameters of the Exponential distribution on the basis of such data. Since the assumption of known location parameter tend to be unreliable, we discussed how sensitive the procedure is to the known value of location parameter. We will consider the robustness of the test concerning the behavior of the significance level as well as power function when the location parameter is known. As an illustration, the times between consecutive telephone calls to a company's switchboard is analyzed using the proposed procedures. Also, some open problems are described.

2 Introduction

Let $X_1, X_2, X_3, \cdots$ be a sequence of continuous random variables. X_k is a lower record value if its value is smaller than all preceding values $X_1, X_2, \cdots, X_{k-1}$. By definition, X_1 is a lower record value. The data may be represented as $(\mathbf{r}, \mathbf{k}) := (r_1, k_1, r_2, k_2, \cdots, r_m, k_m)$, where r_i is the value of the ith observed minimum, and k_i is the number of trials required to obtain the next new minimum. There are variety of situations in which the experimental outcomes available for analysis of data are record statistics. Examples include industrial stress testing, meteorological analysis, sporting and athletic events. For more details see Samaniego and Whitaker (1986), Ahmadi, et. al. (2005) and Doostparast (2009). A random variable X is said to have an Exponential distribution, denoted by $X \sim Exp(\theta, \sigma)$, if its cumulative distribution function is $F(x; \theta, \sigma) = 1 - e^{-\frac{1}{\sigma}(x-\theta)}$, $x \geq \theta$. The exponential distribution is applied in a wide variety of statistical procedures, especially in life testing problems (Balakrishnan and Basu, 1995).

3 Point estimation

It can be shown that, the associated likelihood function of records coming from $Exp(\theta, \sigma)$ is given by $L(\theta, \sigma; \mathbf{r}, \mathbf{k}) = \sigma^{-m} \exp\left\{-\sigma^{-1} \sum_{i=1}^{m} k_i (r_i - \theta)\right\}$. The best unbiased estimator and unique uniformly minimum risk unbiased estimation of σ under LINEX loss function, i.e. $L(\theta, \delta) = e^{a(\delta-\theta)} - a(\delta - \theta) - 1$ as well as maximum likelihood estimation of σ are derived.

4 Confidence intervals

We prove that when θ is known, minimum length confidence interval is given by $I = (2\sum_{i=1}^{m} K_i(R_i - \theta)/b,$ $2\sum_{i=1}^{m} K_i(R_i - \theta)/a)$, where a and b are obtained from the equations $\int_a^b h_{2m}(x)dx = 1-\alpha$, and $h_{2m+4}(a) =$

$h_{2m+4}(b)$. Having obtained a and b from the equations, $\int_a^b h_{2m}(x)dx = 1-\alpha$, and $ah_{2m}(a) = bh_{2m}(b)$, we have an unbiased $(1-\alpha)$-level confidence interval. When θ is unknown, solving the above equations we derive minimum length and unbiased intervals replacing m and $\sum_{i=1}^{m} K_i(R_i-\theta)$ by $m-1$ and $\sum_{i=1}^{m} K_i(R_i-R_m)$, respectively.

5 Testing Hypotheses

When θ is known , we show that uniformly most powerful (UMP) test size α of $H_0 : \sigma = \sigma_0$ against the alternative $H_1 : \sigma > \sigma_0$ rejects H_0 if $2\sum_{i=1}^{m} K_i(R_i-\theta) > \sigma_0\chi^2_{2m,1-\alpha}$. We prove that UMP size α $(0<\alpha<1)$ test of $H_0 : \sigma = \sigma_0$ against the alternative $H_1 : \sigma \neq \sigma_0$ does not exist. Therefore, we consider a classical procedure for constructing tests for $H_0 : \sigma = \sigma_0$ against the alternative $H_1 : \sigma \neq \sigma_0$ well known as *generalized likelihood ratio* (GLR) that has some intuitive appeal and that frequently, though not necessarily, leads to optimal tests. Applying GLR procedure leads to the critical region at level α as $C_{GLR,1} = \{(\mathbf{r},\mathbf{k}) : z^m \exp\{-z/2\} < a\}$ where $Z = 2\sum_{i=1}^{m} K_i(R_i-\theta)/\sigma_0$ and a is chosen such that $\alpha = P(Z^m \exp\{-Z/2\} < a)$ and $Z \sim \chi^2_{(2m)}$ under H_0.

When θ is known we could find UMP test for one-sided hypotheses. But when θ is unknown, there is no UMP test for one-sided hypotheses (Open problem). We show that to test $H_0 : \sigma \geq \sigma_0$, against $H_1 : \sigma < \sigma_0$, a UMP invariant size α test rejects H_0 when $2\sum_{i=1}^{m} K_i(R_i-R_m) < \sigma_0\chi^2_{2(m-1),\alpha}$. There is no UMP test of $H_0 : \sigma = \sigma_0$ against the alternative $H_0 : \sigma \neq \sigma_0$. But GLR procedure leads to the critical region at level α as $C_{GLR,2} = \{(\mathbf{r},\mathbf{k}) : y^m \exp\{-y/2\} < c\},\}$ where $Y = 2\sum_{i=1}^{m} K_i(R_i-R_m)/\sigma_0$ and c is chosen such that $\alpha = P(Y^m \exp\{-Y/2\} < c)$ and $Y \sim \chi^2_{(2(m-1))}$ under H_0.

6 Telephone Call Data

Here, the times between 48 (in minutes) consecutive telephone calls to a company's switchboard presented by Castillo *et. al.* (2005)is considered. Assuming that the times between telephone calls have one parameter exponential distribution $Exp(\theta = 0, \sigma)$, Castillo *et. al.* (2005) obtained the maximum likelihood estimation of σ based on the complete data as 0.934. Assuming $m = 4$, the maximum likelihood estimate of σ on the basis of record data is given by $\sum_{i=1}^{m} K_i(R_i-0)/m = 4.67/4 = 1.1675$. The 95% minimum length, unbiased and equal tail confidence intervals for σ are given by $[0.3749, 3.4558]$, $[0.4952, 3.8693]$ and $[0.5327, 4.2850]$, respectively. In this example, the results from different approaches are very similar.

References

J. Ahmadi, M. Doostparast and A. Parsian (2005) Estimation and prediction in a two-parameter exponential distribution based on k-record values under LINEX loss function, *Commun. Statist. Theor. Meth.* **34**(4), 795–805.

N. Balakrishnan and A. P. Basu (Eds.) (1995) *Exponential Distribution; Theory, Methods and Applications*, Taylor & Francis, Newark, New Jersey.

F. Castillo, A. S. Hadi, N. Balakrishnan and J. M. Sarabia (2005) *Extreme Value and Related Models with Applications in Engineering and Science*, John Wiley & Sons, Inc.

M. Doostparast (2009) A Note on estimation based on record data, *Metrika* **69**, 69–80.

F. J. Samaniego and L. R. Whitaker (1986) On estimating population characteristics from record breaking observations I. Parametric results. *Naval Research Logistics Quarterly*. **33**, 531-543.