Journal of the Korean Statistical Society 45 (2016) 342–357

Contents lists available at ScienceDirect

Journal of the Korean Statistical Society

journal homepage: www.elsevier.com/locate/jkss

Statistical inference on Gumbel distribution using record values

Jung In Seo [a], Yongku Kim [b,*]

[a] Department of Statistics, Daejeon University, South Korea

[b] Department of Statistics, Kyungpook National University, South Korea

ARTICLE INFO

Article history:
Received 24 February 2015
Accepted 7 December 2015
Available online 28 December 2015

AMS 2000 subject classifications:
62F15
62F10

Keywords:
Gumbel distribution
Objective Bayesian analysis
Upper record value
Unbiased estimator

ABSTRACT

In this study, we address inference problems for Gumbel distribution when the available data are lower record values. We first derive unbiased estimators of unknown parameters, and then, we construct an exact confidence interval for the scale parameter and a predictive interval for the next lower value by deriving certain properties and pivotal quantities. These are compared with the results for existing inference. For Bayesian inference, we derive noninformative priors such as the Jeffreys and reference priors for unknown parameters and examine whether they satisfy the probability matching criteria; then, we apply them to develop objective Bayesian analysis.

1. Introduction

Let $\{X_1, X_2, \ldots\}$ be a sequence of independent and identically distributed (i.i.d.) random variables (RVs) with the cumulative distribution function (CDF) and probability density function (PDF). If an observation X_j is lower than all previous observations, then it is called a lower record value. That is, X_j is a lower record value if $X_j < X_i$ for every $i < j$. The indexes for which lower record values occur are given by the record times $\{L(k), k \geq 1\}$, where $L(k) = \min\{j | j > L(k-1), X_j < X_{L(k-1)}\}$, $k > 1$, with $L(1) = 1$. Therefore, a sequence of lower record values is denoted by $\{x_{L(k)}, k = 1, 2, \ldots\}$ from the original sequence $\{X_1, X_2, \ldots\}$. Since these record values arise in many real-life situations, such as weather, sports, economics, and life tests, they are widely used in statistical applications and modeling. Chandler (1952) examined record values and documented a number of basic properties of records. Baklizi (2008) discussed the likelihood and Bayesian estimation of stress–strength reliability in generalized exponential distribution based on record values. Asgharzadeh and Abdi (2012) derived an exact confidence interval and an exact joint confidence region for the shape parameters of the Burr Type XII distribution based on record values.

We discuss estimation and prediction for Gumbel distribution based on lower record values from both classical and Bayesian points of view. The PDF and the CDF of X with the Gumbel distribution are given by

$$f(x) = \frac{1}{\sigma} e^{-(x-\mu)/\sigma} e^{-e^{-(x-\mu)/\sigma}}$$

* Corresponding author.
E-mail address: kim.1252@knu.ac.kr (Y. Kim).

http://dx.doi.org/10.1016/j.jkss.2015.12.002
1226-3192/

and

$$F(x) = e^{-e^{-(x-\mu)/\sigma}}, \quad -\infty < x < \infty, \ -\infty < \mu < \infty, \ \sigma > 0$$

where μ is the location parameter and σ is the scale parameter. This distribution was introduced by Gumbel (1958). Ahsanullah (1990, 1991) derived the maximum likelihood estimators (MLEs) and minimum variance unbiased estimators (MVUEs) of location and scale parameters in the Gumbel distribution based on record data, and provided future lower record values. Balakrishnan, Ahsanullah, and Chan (1992) established some recurrence relationships for single and double moments of lower record values from the Gumbel distribution. Asgharzadeh, Abdi, and Nadarajah (in press) provided an exact confidence interval for the scale parameter and a joint confidence region for the location and scale parameters in the Gumbel distribution based on record values. We derive another exact confidence interval for σ as well as unbiased estimators of μ and σ. Our derived confidence interval is relatively shorter than that of Asgharzadeh et al. (in press). In addition, we derive the predictive interval for the next lower value based on a pivotal quantity. In the Bayesian inference, Ali Mousa, Jaheen, and Ahmad (2002) and Vidal (2014) considered a bivariate prior distribution (μ, σ) given by

$$\pi(\mu, \sigma) \propto \frac{\beta^{\alpha}}{\sigma^{\alpha+2}\Gamma(\alpha)} e^{-\beta/\sigma}, \quad \alpha, \ \beta > 0. \tag{1}$$

Based on this prior, and developed estimation methods for unknown parameters and prediction methods for future record values in the Gumbel distribution based on record values and complete data, respectively. We develop noninformative priors for (μ, σ), which may be a good choice in situations in which little or no prior information is available, and apply them for objective Bayesian analysis. In the prediction, the Bayesian approach has the advantage that the Bayesian predictive distribution does not depend on an unknown parameter, unlike those provided in Ahsanullah (1991). Jeffreys (1961) introduced a noninformative prior, the Jeffreys prior, to address the problem of the uniform prior, which does not remain invariant under one-to-one reparameterization. The Jeffreys prior works well in the one-parameter case. However, in the presence of nuisance parameters, it has some problems, such as a marginalization paradox and a Neyman–Scott problem. To overcome the drawbacks of the multiparameter case, Bernardo (1979) proposed a reference prior that maximizes the Kullback–Leibler divergence between the prior and posterior while Berger and Bernardo (1989, 1992) provided a general algorithm to derive a reference prior. In addition to these priors, we investigate the probability matching prior of Welch and Peers (1963), which is another popular noninformative prior, because it has frequentist properties in which the posterior probabilities of certain regions coincide with their coverage probabilities. Finally, we develop a Bayesian prediction method whose key advantage is that its predictive distribution does not depend on unknown parameters, unlike the prediction results provided in Ahsanullah (1991).

The remainder of the paper is structured as follows. First, related classical inferences are revisited in Section 2. In Section 3, we develop an objective Bayesian analysis. In the following section, the performance of the proposed methods are compared using real data and simulation. Finally, the study findings are discussed in Section 5.

2. Revisit to classical inference

We introduce two lemmas that are used to obtain unbiased estimators of unknown parameters, an exact confidence interval for the scale parameter σ, and a predictive interval for the next lower record value.

Lemma 2.1. *Let $X_{L(k)}$ be the kth lower record value from the Gumbel distribution. Then*

$$E\left(X_{L(k)}\right) = \mu - \sigma\psi(k) \tag{2}$$

and

$$Var\left(X_{L(k)}\right) = \sigma^2\psi_1(k),$$

where $\psi(\cdot)$ is the digamma function and $\psi_1(\cdot)$ is the trigamma function.

Proof. Based on Ahsanullah (1995), the PDF of the kth lower record value $X_{L(k)}$ with a continuous probability distribution is defined as

$$f_{X_{L(k)}}(x) = \frac{1}{\Gamma(k)}\left[-\log F(x)\right]^{k-1} f(x), \quad -\infty < x < \infty,$$

where $\Gamma(\cdot)$ is a gamma function. Then, since the PDF of $X_{L(k)}$ with the Gumbel distribution is given by

$$f_{X_{L(k)}}(x) = \frac{1}{\sigma\Gamma(k)} e^{-k(x-\mu)/\sigma} e^{-e^{-(x-\mu)/\sigma}}, \tag{3}$$