Proceedings of the World Congress on Engineering 2012 Vol III
WCE 2012, July 4 - 6, 2012, London, U.K.

Optimal Prediction Intervals for Future Order Statistics from Extreme Value Distributions

Nicholas A. Nechval, *Member, IAENG*, Maris Purgailis, Konstantin N. Nechval, and Inta Bruna

Abstract—**Prediction intervals for future order statistics are widely used for reliability problems and other related problems. The determination of these intervals has been extensively investigated. But the optimality property of these intervals has not been fully explored. In this paper we discuss this problem for extreme value distributions. Introducing a risk function to compare prediction intervals, the interval which minimizes it among the class of invariant prediction intervals is obtained. The technique used here for optimization of prediction intervals based on censored data emphasizes pivotal quantities relevant for obtaining ancillary statistics and factors. It allows one to solve the optimization problems in a simple way.**

Index Terms — **Extreme value distribution, future order statistic, prediction interval, risk function, optimization**

I. Introduction

PREDICTION of an unobserved random variable is a fundamental problem in statistics. Patel [1] provides an extensive survey of literature on this topic. In the areas of reliability and life-testing, this problem translates to obtaining prediction intervals for life distributions such as the Exponential and the Weibull. One of the earlier works on prediction for the Weibull distribution is by Mann and Saunders [2]. They considered prediction intervals for the smallest of a set of future observations, based on a small (two or three) preliminary sample of past observations. An expression for the warranty period (time before the failure of the first ordered observation from a set of future observations or a lot) was derived as a function of the ordered past observations. Mann [3] extended the results for lot sizes n = 10 (5) 25 and sample sizes m = 2 (1) $n-3$ for a specified assurance level of 0.95. This method requires numerical integration. In addition, the tables provided are limited to sample sizes less than 25 and are given only for the assurance level of 0.95. Antle and Rademaker [4] provided a method of obtaining a prediction bound for the largest observation from a future sample of the Type I extreme value distribution, based on the maximum likelihood estimates of the parameters. They used Monte Carlo simulations to obtain the prediction intervals. Using the well-known relationship between the Weibull distribution and the Type I extreme value distribution one can use their method to construct an upper prediction limit for the largest among a set of future Weibull observations. However this method is valid only for complete samples and limited to constructing an upper prediction limit for the largest among a set of future observations. Lawless [5] proposed a method for constructing prediction intervals for the smallest ordered observation among a set of k future observations based on a Type II censored sample of past observations. These results are based on the conditional distribution of the maximum likelihood estimates given a set of ancillary statistics. This procedure is exact, but it requires numerical integration, for each new sample obtained, to determine the prediction bounds. Mee and Kushary [6] provided a simulation based procedure for constructing prediction intervals for Weibull populations for Type II censored case. This procedure is based on maximum likelihood estimation and requires an iterative process to determine the percentile points.

To develop appropriate probabilistic models and assess the risks caused by stochastic events, business analysts and engineers frequently use the extreme value distributions (EVD). In this paper we assume that the parent EVD are the Gumbel distribution,

$$\Pr\{X > x\} = \exp\left[-\exp\left(\frac{x-\mu}{\sigma}\right)\right], \quad -\infty < x < \infty, \tag{1}$$

where μ is the location parameter, and σ is the scale parameter ($\sigma > 0$), and the Weibull distribution,

$$\Pr\{Y > y\} = \exp\left[-\left(\frac{y}{\beta}\right)^{\delta}\right], \quad y \geq 0, \tag{2}$$

where both distribution parameters (δ – shape, β – scale) are positive.

Let Y be a random variable with the Weibull distribution (2), and define $X = \ln Y$. Then X becomes a random variable with the Gumbel distribution (1) where $\mu = \ln\beta$ and $\sigma = \delta^{-1}$. Therefore it is enough to consider only the Gumbel distribution, because the results for the Weibull distribution are easily obtained from those for the Gumbel distribution.

Manuscript received March 06, 2012. This work was supported in part by Grant No. 06.1936, Grant No. 07.2036, Grant No. 09.1014, and Grant No. 09.1544 from the Latvian Council of Science and the National Institute of Mathematics and Informatics of Latvia.

Nicholas A. Nechval is with the Statistics Department, EVF Research Institute, University of Latvia, Riga LV-1050, Latvia (e-mail: nechval@junik.lv).

Maris Purgailis is with the Cybernetics Department, University of Latvia, Riga LV-1050, Latvia (e-mail: marispur@lanet.lv).

Konstantin N. Nechval is with the Applied Mathematics Department, Transport and Telecommunication Instutute, Riga LV-1019, Latvia (e-mail: konstan@tsi.lv).

Inta Bruna is with the Cybernetics Department, University of Latvia, Riga LV-1050, Latvia (e-mail: inta.bruna@lu.lv).

II. Within – Sample Prediction

A. Mathematical Preliminaries

Theorem 1. Let $X_1 \le ... \le X_k$ be the first k ordered observations (order statistics) in a sample of size m from a continuous distribution with some probability density function $f_\theta(x)$ and distribution function $F_\theta(x)$, where θ is a parameter (in general, vector). Then the joint probability density function of $X_1 \le ... \le X_k$ and the lth order statistics X_l $(1 \le k < l \le m)$ is given by

$$f_\theta(x_1,...,x_k,x_l) = f_\theta(x_1,...,x_k)\, f_\theta(x_l \mid x_k), \quad (3)$$

where

$$f_\theta(x_1,...,x_k) = \frac{m!}{(m-k)!}\prod_{i=1}^{k} f_\theta(x_i)[1-F_\theta(x_k)]^{m-k}, \quad (4)$$

$$f_\theta(x_l \mid x_k) = \frac{(m-k)!}{(l-k-1)!(m-l)!}\left[\frac{F_\theta(x_l)-F_\theta(x_k)}{1-F_\theta(x_k)}\right]^{l-k-1}$$

$$\times\left[1-\frac{F_\theta(x_l)-F_\theta(x_k)}{1-F_\theta(x_k)}\right]^{m-l}\frac{f_\theta(x_l)}{1-F_\theta(x_k)}$$

$$=\frac{(m-k)!}{(l-k-1)!(m-l)!}\sum_{j=0}^{l-k-1}\binom{l-k-1}{j}(-1)^j$$

$$\times\left[\frac{1-F_\theta(x_l)}{1-F_\theta(x_k)}\right]^{m-l+j}\frac{f_\theta(x_l)}{1-F_\theta(x_k)}$$

$$=\frac{(m-k)!}{(l-k-1)!(m-l)!}\sum_{j=0}^{m-l}\binom{m-l}{j}(-1)^j$$

$$\times\left[\frac{F_\theta(x_l)-F_\theta(x_k)}{1-F_\theta(x_k)}\right]^{l-k-1+j}\frac{f_\theta(x_l)}{1-F_\theta(x_k)} \quad (5)$$

represents the conditional probability density function of X_l given X_k=x_k.

Proof. The joint density of $X_1 \le ... \le X_k$ and X_l is given by

$$f_\theta(x_1,...,x_k,x_l) = \frac{m!}{(l-k-1)!(m-l)!}\prod_{i=1}^{k} f_\theta(x_i)$$

$$\times[F_\theta(x_l)-F_\theta(x_k)]^{l-k-1} f_\theta(x_l)[1-F_\theta(x_l)]^{m-l}$$

$$= f_\theta(x_1,...,x_k)\, f_\theta(x_l \mid x_k). \quad (6)$$

It follows from (4) and (6) that

$$f_\theta(x_l \mid x_1,...,x_k) = \frac{f_\theta(x_1,...,x_k,x_l)}{f_\theta(x_1,...,x_k)} = f_\theta(x_l \mid x_k), \quad (7)$$

i.e., the conditional distribution of X_l, given $X_i = x_i$ for all i = 1,…, k, is the same as the conditional distribution of X_l, given only $X_k = x_k$, which is given by (5). This ends the proof. □

Theorem 2. Let $X_1 \le ... \le X_k$ be the first k ordered observations from a sample of size m, which follow the Gumbel distribution (1) with the density

$$f_\theta(x) = \frac{1}{\sigma}\exp\left(\frac{x-\mu}{\sigma}\right)\exp\left(-\exp\left(\frac{x-\mu}{\sigma}\right)\right) \quad (-\infty < x < \infty), \quad (8)$$

where $\theta = (\mu, \sigma)$. Then the joint probability density function of the pivotal quantities

$$S_1 = \frac{\hat{\mu}-\mu}{\sigma}, \quad V_2 = \frac{\hat{\sigma}}{\sigma}, \quad (9)$$

conditional on fixed

$$\mathbf{z}^{(k)} = (z_1,...,z_k), \quad (10)$$

where

$$Z_i = \frac{X_i - \hat{\mu}}{\hat{\sigma}}, \quad i = 1,...,k, \quad (11)$$

are ancillary statistics, any k–2 of which form a functionally independent set, $\hat{\mu}$ and $\hat{\sigma}$ are the maximum likelihood estimates for μ and σ based on the first k ordered observations ($X_1 \le ... \le X_k$) from a sample of size m from the Gumbel distribution (1), which can be found from solution of

$$\hat{\mu} = \hat{\sigma}\ln\left(\left[\sum_{i=1}^{k} e^{x_i/\hat{\sigma}} + (m-k)e^{x_k/\hat{\sigma}}\right]\Big/ k\right), \quad (12)$$

and

$$\hat{\sigma} = \left(\sum_{i=1}^{k} x_i e^{x_i/\hat{\sigma}} + (m-k)x_k e^{x_k/\hat{\sigma}}\right)$$

$$\times\left(\sum_{i=1}^{k} e^{x_i/\hat{\sigma}} + (m-k)e^{x_k/\hat{\sigma}}\right)^{-1} - \frac{1}{k}\sum_{i=1}^{k} x_i, \quad (13)$$

is given by

$$f(s_1, v_2 \mid \mathbf{z}^{(k)}) = \vartheta^{\bullet}(\mathbf{z}^{(k)})v_2^{k-2}\exp\left(v_2\sum_{i=1}^{k} z_i\right)$$

$$\times e^{ks_1}\exp\left(-e^{s_1}\left[\sum_{i=1}^{k}\exp(z_i v_2) + (m-k)\exp(z_k v_2)\right]\right)$$

$$= f(v_2 \mid \mathbf{z}^{(k)}) f(s_1 \mid v_2, \mathbf{z}^{(k)}), \quad s_1 \in (-\infty, \infty), \quad v_2 \in (0, \infty), \quad (14)$$

where

$$\vartheta^{\bullet}(\mathbf{z}^{(k)}) = \left(\Gamma(k)\int_0^\infty v_2^{k-2}\exp\left(v_2\sum_{i=1}^{k} z_i\right)\right.$$

$$\left.\times\left[\sum_{i=1}^{k}\exp(z_i v_2) + (m-k)\exp(z_k v_2)\right]^{-k} dv_2\right)^{-1} \quad (15)$$

is the normalizing constant,

ISBN: 978-988-19252-2-0
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

$$f(v_2 \mid \mathbf{z}^{(k)}) = \vartheta(\mathbf{z}^{(k)}) v_2^{k-2} \exp\left(v_2 \sum_{i=1}^{k} z_i \right)$$

$$\times \left[\sum_{i=1}^{k} \exp(z_i v_2) + (m-k)\exp(z_k v_2) \right]^{-k}, \quad v_2 \in (0,\infty), \quad (16)$$

$$\vartheta(\mathbf{z}^{(k)}) = \left(\int_0^\infty v_2^{k-2} \exp\left(v_2 \sum_{i=1}^{k} z_i \right) \right.$$

$$\left. \times \left[\sum_{i=1}^{k} \exp(z_i v_2) + (m-k)\exp(z_r v_2) \right]^{-k} dv_2 \right)^{-1}, \quad (17)$$

$$f(s_1 \mid v_2, \mathbf{z}^{(k)}) = \frac{1}{\Gamma(k)} \left[\sum_{i=1}^{k} \exp(z_i v_2) + (m-k)\exp(z_k v_2) \right]^{k}$$

$$\times e^{ks_1} \exp\left(-e^{s_1} \left[\sum_{i=1}^{k} \exp(z_i v_2) + (m-k)\exp(z_r v_2) \right] \right),$$

$$s_1 \in (-\infty, \infty). \quad (18)$$

Proof. The joint density of $X_1 \le ... \le X_k$ is given by

$$f(x_1,...,x_k \mid \mu,\sigma) = \frac{m!}{(m-k)!} \prod_{i=1}^{k} \frac{1}{\sigma} \exp\left(\frac{x_i - \mu}{\sigma} - \exp\left(\frac{x_i - \mu}{\sigma} \right) \right)$$

$$\times \exp\left(-(m-k)\exp\left(\frac{x_k - \mu}{\sigma} \right) \right). \quad (19)$$

Using the invariant embedding technique [7-14], we then find in a straightforward manner, that the probability element of the joint density of S_1, V_2, conditional on fixed $\mathbf{z}^{(k)} = (z_1, ..., z_k)$, is

$$f(s_1, v_2 \mid \mathbf{z}^{(k)}) ds_1 dv_2$$

$$= \vartheta^{\bullet}(\mathbf{z}^{(k)}) v_2^{k-2} \exp\left(v_2 \sum_{i=1}^{k} z_i \right) e^{ks_1}$$

$$\times \exp\left(-e^{s_1} \left[\sum_{i=1}^{k} \exp(z_i v_2) + (m-k)\exp(z_k v_2) \right] \right) ds_1 dv_2,$$

$$s_1 \in (-\infty, \infty), \quad v_2 \in (0, \infty). \quad (20)$$

This ends the proof. □

Theorem 3. Let $X_1 \le ... \le X_k$ be the first k ordered observations (order statistics) in a sample of size m from the Gumbel distribution (1). Then the joint probability density function of the pivotal quantities

$$V_1 = \frac{X_l - X_k}{\sigma}, \quad S_2 = \frac{X_k - \mu}{\sigma}, \quad (21)$$

where X_l ($1 \le k < l \le m$) is the lth order statistic from the same sample, is given by

$$f^-(v_1, s_2) = f(v_1 \mid s_2) f(s_2), \quad (22)$$

where

$$f(v_1 \mid s_2) = \frac{1}{\mathrm{B}(l-k, m-l+1)} \sum_{j=0}^{l-k-1} \binom{l-k-1}{j} (-1)^j \frac{1}{m-l+j+1}$$

$$\times e^{-(m-l+j+1)e^{s_2}(e^{v_1}-1)} (m-l+j+1) e^{s_2} e^{v_1}, \quad 0 < v_1 < \infty, \quad (23)$$

$$f(s_2) = \frac{1}{\mathrm{B}(k, m-k+1)} \sum_{j=0}^{k-1} \binom{k-1}{j} (-1)^j e^{-(m-k+j+1)e^{s_2}} e^{s_2},$$

$$-\infty < s_2 < \infty. \quad (24)$$

Proof. The joint density function of the order statistics X_k, X_l ($1 \le k < l \le m$) is given by

$$f_\theta^+(x_k, x_l) = f_\theta^+(x_l \mid x_k) f_\theta^+(x_k). \quad (25)$$

It will be noted that

$$f_\theta^+(x_l \mid x_k) dx_l = \frac{(m-k)!}{(l-k-1)!(m-l)!} \sum_{j=0}^{l-k-1} \binom{l-k-1}{j} (-1)^j$$

$$\times \left[\frac{1 - F_\theta(x_l)}{1 - F_\theta(x_k)} \right]^{m-l+j} \frac{f_\theta(x_l)}{1 - F_\theta(x_k)} dx_l$$

$$= \frac{1}{\mathrm{B}(l-k, m-l+1)} \sum_{j=0}^{l-k-1} \binom{l-k-1}{j} (-1)^j \frac{1}{m-l+j+1}$$

$$\times e^{-(m-l+j+1)e^{s_2}(e^{v_1}-1)} (m-l+j+1) e^{s_2} e^{v_1} dv_1$$

$$\equiv f(v_1 \mid s_2) dv_1, \quad 0 < v_1 < \infty, \quad (26)$$

$$f_\theta^+(x_k) dx_k = \frac{m!}{(k-1)!(m-k)!}$$

$$\times [F(x_k)]^{k-1} [1 - F(x_k)]^{m-k} f_\theta(x_k) dx_k$$

$$= \frac{1}{\mathrm{B}(k, m-k+1)} \sum_{j=0}^{k-1} \binom{k-1}{j} (-1)^j e^{-(m-k+j+1)e^{s_2}} e^{s_2} ds_2$$

$$\equiv f(s_2) ds_2, \quad -\infty < s_2 < \infty. \quad (27)$$

This ends the proof. □

Corollary 3.1. The probability density function of the pivotal quantity V_1 is given by

$$f^-(v_1) = \int_{-\infty}^{\infty} f^-(v_1, s_2) ds_2 = \int_{-\infty}^{\infty} f(v_1 \mid s_2) f(s_2) ds_2. \quad (28)$$

Corollary 3.2. The joint probability density function of the pivotal quantities

$$V_1 = \frac{X_l - X_k}{\sigma}, \quad V_2 = \frac{\widehat{\sigma}}{\sigma} \tag{29}$$

is given by

$$f(v_1, v_2) = f^{-}(v_1) f(v_2 \,|\, \mathbf{z}^{(k)}). \tag{30}$$

B. Piecewise-Linear Loss Function

We shall consider the interval prediction problem for the lth order statistic X_l, $k<l\leq m$, in the same sample of size m for the situation where the first k observations $X_1 < X_2 < \cdots < X_k$, $1\leq k<m$, have been observed. Suppose that we assert that an interval $\mathbf{d}=(d_1,d_2)$ contains X_l. If, as is usually the case, the purpose of this interval statement is to convey useful information we incur penalties if d_1 lies above X_l or if d_2 falls below X_l. Suppose that these penalties are $c_1(d_1- X_l)$ and $c_2(X_l-d_2)$, losses proportional to the amounts by which X_l escapes the interval. Since c_1 and c_2 may be different the possibility of differential losses associated with the interval overshooting and undershooting the true μ is allowed. In addition to these losses there will be a cost attaching to the length of interval used. For example, it will be more difficult and more expensive to design or plan when the interval $\mathbf{d}=(d_1,d_2)$ is wide. Suppose that the cost associated with the interval is proportional to its length, say $c(d_2-d_1)$. In the specification of the loss function, σ is clearly a 'nuisance parameter' and no alteration to the basic decision problem is caused by multiplying all loss factors by $1/\sigma$. Thus we are led to investigate the piecewise-linear loss function

$$r(\boldsymbol{\theta}, \mathbf{d}) = \begin{cases} \dfrac{c_1(d_1 - X_l)}{\sigma} + \dfrac{c(d_2 - d_1)}{\sigma} & (X_l < d_1), \\ \dfrac{c(d_2 - d_1)}{\sigma} & (d_1 \leq X_l \leq d_2), \\ \dfrac{c(d_2 - d_1)}{\sigma} + \dfrac{c_2(X_l - d_2)}{\sigma} & (X_l > d_2). \end{cases} \tag{31}$$

The decision problem specified by the informative experiment density function (1) and the loss function (31) is invariant under the group of transformations, which takes μ (the location parameter) and σ (the scale) into $c\mu + b$ and $c\sigma$, respectively, where b lies in the range of μ, $c > 0$. This group acts transitively on the parameter space. Thus, the problem is to find the best invariant interval predictor of X_l,

$$\mathbf{d}^* = \arg\min_{\mathbf{d}\in\mathscr{D}} R(\boldsymbol{\theta}, \mathbf{d}), \tag{32}$$

where $\mathscr{D}$ is a set of invariant interval predictors of X_l, $R(\boldsymbol{\theta},\mathbf{d})=E_{\boldsymbol{\theta}}\{r(\boldsymbol{\theta},\mathbf{d})\}$ is a risk function.

C. Transformation of the Loss Function

It follows from (31) that the invariant loss function, $r(\boldsymbol{\theta},\mathbf{d})$, can be transformed as follows:

$$r(\boldsymbol{\theta}, \mathbf{d}) = \ddot{r}(\mathbf{V}, \boldsymbol{\eta}), \tag{33}$$

where

$$\ddot{r}(\mathbf{V}, \boldsymbol{\eta}) = \begin{cases} c_1(-V_1 + \eta_1 V_2) + c(\eta_2 - \eta_1) V_2 & (V_1 < \eta_1 V_2), \\ c(\eta_2 - \eta_1) V_2 & (\eta_1 V_2 \leq V_1 \leq \eta_2 V_2), \\ c_2(V_1 - \eta_2 V_2) + c(\eta_2 - \eta_1) V_2 & (V_1 > \eta_2 V_2), \end{cases} \tag{34}$$

$$\mathbf{V}=(V_1,V_2), \quad V_1=(X_l - X_k)/\sigma, \quad V_2=\widehat{\sigma}/\sigma;$$

$$\boldsymbol{\eta}=(\eta_1,\eta_2), \quad \eta_1=(d_1 - X_k)/\widehat{\sigma}, \quad \eta_2=(d_2 - X_k)/\widehat{\sigma}. \tag{35}$$

D. Risk Function

It follows from (34) that the risk associated with $\mathbf{d}$ and $\boldsymbol{\theta}$ can be expressed as

$$R(\boldsymbol{\theta}, \mathbf{d}) = E_{\boldsymbol{\theta}}\{r(\boldsymbol{\theta}, \mathbf{d})\} = E\{\ddot{r}(\mathbf{V}, \boldsymbol{\eta})\}$$

$$= c_1 \int_0^\infty \int_0^{\eta_1 v_2} (-v_1 + \eta_1 v_2) f(v_1, v_2) dv_1 dv_2$$

$$+ c_2 \int_0^\infty \int_{\eta_2 v_2}^\infty (v_1 - \eta_2 v_2) f(v_1, v_2) dv_1 dv_2$$

$$+ c(\eta_2 - \eta_1) \int_0^\infty \int_0^\infty v_2 f(v_1, v_2) dv_1 dv_2, \tag{36}$$

which is constant on orbits when an invariant predictor (decision rule) $\mathbf{d}$ is used, where $f(v_1,v_2)$ is defined by (30).

E. Risk Minimization and Invariant Prediction Rules

The following theorem gives the central result in this section.

Theorem 4. Suppose that (U_1,U_2) is a random vector having density function

$$u_2 f(u_1, u_2) \left[\int_0^\infty \int_0^\infty u_2 f(u_1, u_2) du_1 du_2 \right]^{-1} \quad (u_1, u_2 > 0), \tag{37}$$

where f is defined by $f(v_1,v_2)$, and let Q be the probability distribution function of U_1/U_2.

(i) If $c/c_1+c/c_2<1$ then the optimal invariant linear-loss interval predictor of X_l based on $\mathbf{X}$ is $\mathbf{d}^*=(X_k+\eta_1 S_k,\ X_k+\eta_2 S_k)$, where

$$Q(\eta_1) = c/c_1, \quad Q(\eta_2) = 1 - c/c_2. \tag{38}$$

(ii) If $c/c_1+c/c_2\geq 1$ then the optimal invariant linear-loss interval predictor of X_l based on $\mathbf{X}$ degenerates into a point predictor $X_k+\eta_\bullet S_k$, where

$$Q(\eta_\bullet) = c_2/(c_1 + c_2). \tag{39}$$

Proof. From (36)

$$\frac{\partial E\{\ddot{r}(\mathbf{V}, \boldsymbol{\eta})\}}{\partial \eta_1}$$

$$= c_1 \int_0^\infty \int_0^{\eta_1 v_2} v_2 f(v_1, v_2) dv_1 dv_2 - c \int_0^\infty \int_0^\infty v_2 f(v_1, v_2) dv_1 dv_2$$

$$= \int_0^\infty \int_0^\infty v_2 f(v_1, v_2) dv_1 dv_2 [c_1 Q(\eta_1) - c], \tag{40}$$

and

$$\frac{\partial E\{\ddot{r}(\mathbf{V}, \boldsymbol{\eta})\}}{\partial \eta_2} = \int_0^\infty \int_0^\infty v_2 f(v_1, v_2) dv_1 dv_2 [-c_2(1 - Q(\eta_2)) + c], \tag{41}$$

ISBN: 978-988-19252-2-0
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

where

$$Q(\eta)=\int_0^{\eta} q(w)dw, \tag{42}$$

$$q(w)=\frac{\int_0^{\infty} v_2^2 f(wv_2,v_2)dv_2}{\int_0^{\infty}\int_0^{\infty} v_2 f(v_1,v_2)dv_1dv_2}, \tag{43}$$

$$W=V_1/V_2. \tag{44}$$

Now $\partial E\{\ddot{r}(\mathbf{V},\boldsymbol{\eta})\}/\partial\eta_1=\partial E\{\ddot{r}(\mathbf{V},\boldsymbol{\eta})\}/\partial\eta_2=0$ if and only if (38) hold. Thus, $E\{\ddot{r}(\mathbf{V},\boldsymbol{\eta})\}$ provided (38) has a solution with $\eta_1<\eta_2$ and this is so if $1-c/c_2>c/c_1$. It is easily confirmed that this $\boldsymbol{\eta}=(\eta_1,\eta_2)$ gives the minimum value of $E\{\ddot{r}(\mathbf{V},\boldsymbol{\eta})\}$. Thus (i) is established.

If $c/c_1+c/c_2\geq 1$ then the minimum of $E\{\ddot{r}(\mathrm{v},\eta)\}$ in the region $\eta_2\geq\eta_1$ occurs where $\eta_1=\eta_2=\eta_\bullet$, $\eta_\bullet$ being determined by setting

$$\partial E\{\ddot{r}(V,(\eta_\bullet,\eta_\bullet))\}/\partial\eta_\bullet=0 \tag{45}$$

and this reduces to

$$c_1Q(\eta_\bullet)-c_2[1-Q(\eta_\bullet)]=0, \tag{46}$$

which establishes (ii). □

Corollary 4.1. The minimum risk of the optimal invariant predictor of X_l based on $\mathbf{X}$ is given by

$$R(\boldsymbol{\theta},\mathbf{d}^*)=E_{\boldsymbol{\theta}}\{r(\boldsymbol{\theta},\mathbf{d}^*)\}=E\{\ddot{r}(\mathbf{V},\boldsymbol{\eta})\}$$

$$=-c_1\int_0^{\infty}\int_0^{\eta_1 v_2} v_1 f(v_1,v_2)dv_1dv_2+c_2\int_0^{\infty}\int_{\eta_2 v_2}^{\infty} v_1 f(v_1,v_2)dv_1dv_2 \tag{47}$$

for case (i) with $\boldsymbol{\eta}=(\eta_1,\eta_2)$ as given by (38) and for case (ii) with $\eta_1=\eta_2=\eta_\bullet$ as given by (39).

Proof. These results are immediate from (36) when use is made of $\partial E\{\ddot{r}(\mathbf{V},\boldsymbol{\eta})\}/\partial\eta_1=\partial E\{\ddot{r}(\mathbf{V},\boldsymbol{\eta})\}/\partial\eta_2=0$ in case (i) and $\partial \mathrm{E}\{\ddot{r}(\mathbf{V},(\eta_\bullet,\eta_\bullet))\}/\partial\eta_\bullet=0$ in case (ii). □

The underlying reason why $c/c_1+c/c_2$ acts as a separator of interval and point prediction is that for $c/c_1+c/c_2\geq 1$ every interval predictor is inadmissible, there existing some point predictor with uniformly smaller risk.

Theorem 5. Suppose that $\mu=0$ and

$$\mathbf{V}=(V_1,V_2),\quad V_1=(X_l-X_k)/\sigma,\quad V_2=X_k/\sigma;$$

$$\boldsymbol{\eta}^\circ=(\eta_1^\circ,\eta_2^\circ),\quad \eta_1^\circ=(d_1-X_k)/X_k,\quad \eta_2^\circ=(d_2-X_k)/X_k. \tag{48}$$

Let us assume that (U_1,U_2) is a random vector having density function

$$u_2 f_0(u_1,u_2)\left[\int_0^{\infty}\int_0^{\infty} u_2 f_0(u_1,u_2)du_1du_2\right]^{-1}\quad (u_1,u_2>0), \tag{49}$$

where f_0 is defined by $f_0(v_1,v_2)$, and let Q_0 be the probability distribution function of u_1/u_2.

(i) If $c/c_1+c/c_2<1$ then the optimal invariant linear-loss interval predictor of X_l based on X_k is $\mathbf{d}^*=((1+\eta_1^\circ)X_k,(1+\eta_2^\circ)X_k)$, where

$$Q_0(\eta_1^\circ)=c/c_1,\quad Q_0(\eta_2^\circ)=1-c/c_2. \tag{50}$$

(ii) If $c/c_1+c/c_2\geq 1$ then the optimal invariant linear-loss interval predictor of X_l based on X_k degenerates into a point predictor $(1+\eta_\bullet^\circ)X_k$, where

$$Q_0(\eta_\bullet^\circ)=c_2/(c_1+c_2). \tag{51}$$

Proof. For the proof we refer to Theorem 1. □

Corollary 5.1. The minimum risk of the optimal invariant predictor of X_l Based on X_k is given by

$$R(\boldsymbol{\theta},\mathbf{d}^*)=E_{\boldsymbol{\theta}}\{r(\boldsymbol{\theta},\mathbf{d}^*)\}=E\{\ddot{r}(\mathbf{V},\boldsymbol{\eta}^\circ)\}$$

$$=-c_1\int_0^{\infty}\int_0^{\eta_1^\circ v_2} v_1 f_0(v_1,v_2)dv_1dv_2+c_2\int_0^{\infty}\int_{\eta_2^\circ v_2}^{\infty} v_1 f_0(v_1,v_2)dv_1dv_2 \tag{52}$$

for case (i) with $\boldsymbol{\eta}^\circ=(\eta_1^\circ,\eta_2^\circ)$ as given by (50) and for case (ii) with $\eta_1^\circ=\eta_2^\circ=\eta_\bullet^\circ$ as given by (51).

Proof. For the proof we refer to Corollary 4.1. □

III. Equivalent Confidence Coefficient

For case (i) when we obtain an interval predictor for X_l we may regard the interval as a confidence interval in the conventional sense and evaluate its confidence coefficient. The general result is contained in the following theorem.

Theorem 6. Suppose that $\mathbf{V}=(V_1,V_2)$ is a random vector having density function $f(v_1,v_2)$ $(v_1,v_2>0)$ where f is defined by (30) and let H be the distribution function of $W=V_1/V_2$, i.e., the probability density function of W is given by

$$h(w)=\int_0^{\infty} v_2 f(wv_2,v_2)dv_2. \tag{53}$$

Then the confidence coefficient based on $\mathbf{X}$ and associated with the optimum prediction interval $\mathbf{d}^*=(d_1,d_2)$, where $d_1=X_k+\eta_1\widehat{\sigma}$, $d_2=X_k+\eta_2\widehat{\sigma}$, is

$$\Pr\{\mathbf{d}^*: d_1<X_l<d_2\,|\,\mu,\sigma\}$$

$$=H[Q^{-1}(1-c/c_2)]-H[Q^{-1}(c/c_1)]. \tag{54}$$

ISBN: 978-988-19252-2-0
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

Proof. The confidence coefficient for $\mathbf{d}^*$ corresponding to (μ,σ) is given by

$$\Pr\{(X_k,\hat{\sigma}): X_k+\eta_1\hat{\sigma} < X_l < X_k+\eta_2\hat{\sigma} \mid \mu,\sigma\}$$

$$= \Pr\{(v_1,v_2): \eta_1 < v_1/v_2 < \eta_2\}$$

$$= H(\eta_2) - H(\eta_1) = H[Q^{-1}(1-c/c_2)] - H[Q^{-1}(c/c_1)]. \quad (55)$$

This is independent of (μ,σ). □

Theorem 7. Suppose that $\mathbf{V}=(V_1,V_2)$ is a random vector having density function $f_0(v_1,v_2)$ (v_1 real, $v_2>0$), where f_0 is defined by

$$f_\theta(x_k,x_l) = \frac{1}{\mathrm{B}(k,l-k)\mathrm{B}(l,m-l+1)}$$

$$\times[F_\theta(x_k)]^{k-1}[F_\theta(x_l)-F_\theta(x_k)]^{l-k-1}$$

$$\times[1-F_\theta(x_l)]^{m-l} f_\theta(x_k) f_\theta(x_l), \quad (56)$$

where $\mu=0$, and let H_0 be the distribution function of $W=V_1/V_2$, i.e., the probability density function of W is given by

$$h_0(w) = \int_0^\infty v_2 f_0(wv_2,v_2)dv_2. \quad (57)$$

Then the confidence coefficient based on X_k and associated with the optimum prediction interval $\mathbf{d}^*=(d_1,d_2)$, where $d_1=(1+\eta_1^\circ)X_k$, $d_2=(1+\eta_2^\circ)X_k$, is

$$\Pr\{\mathbf{d}^*: d_1 < X_l < d_2 \mid \mu,\sigma\}$$

$$= H_0[Q_0^{-1}(1-c/c_2)] - H_0[Q_0^{-1}(c/c_1)]. \quad (58)$$

Proof. For the proof we refer to Theorem 6. □

The way in which (54) (or (58)) varies with c, c_1 and c_2, and the fact that c_1 and c_2 are the factors of proportionality associated with losses from overshooting and undershooting relative to loss involved in increasing the length of interval, provides an interesting interpretation of confidence interval prediction.

IV. Conclusion

In many statistical decision problems it is reasonable co confine attention to rules that are invariant with respect to a certain group of transformations. If a given decision problem admits a sufficient statistic, it is well known that the class of invariant rules based on the sufficient statistic is essentially complete in the class of all invariant rules under some assumptions. This result may be used to show that if there exists a minimax invariant rule among invariant rules based on sufficient statistic, it is minimax among all invariant rules. In this paper, we consider statistical prediction problems which are invariant with respect to a certain group of transformations and construct the optimal invariant interval predictors. The method used is that of the invariant embedding of sample statistics in a loss function in order to form pivotal quantities which allow one to eliminate unknown parameters from the problem. This method is a special case of more general considerations applicable whenever the statistical problem is invariant under a group of transformations, which acts transitively on the parameter space.

Acknowledgment

This research was supported in part by Grant No. 06.1936, Grant No. 07.2036, Grant No. 09.1014, and Grant No. 09.1544 from the Latvian Council of Science and the National Institute of Mathematics and Informatics of Latvia.

References

[1] J.K. Patel, "Prediction intervals – a review," *Communications in Statistics –Theory and Methods*, vol. 13, pp. 2393 –2465, 1989.

[2] N.R. Mann and S.C. Saunders, "On evaluation of warranty assurance when life has a Weibull distribution," *Biometrika*, vol. 56, pp. 615–625, 1969.

[3] N.R. Mann, "Warranty periods based on three ordered sample observations from a Weibull population," *IEEE Transactions on Reliability*, vol. R–19, pp. 167–171, 1970.

[4] C.E. Antle and F. Rademaker, "An upper confidence limit on the maximum of *m* future observations from a type I extreme-value distribution," *Biometrika*, vol. 59, pp. 475–477, 1972.

[5] J.F. Lawless, "On estimation of safe life when the underlying distribution is Weibull," *Technometrics*, vol. 15, pp. 857–865, 1973.

[6] R.W. Mee and D. Kushary, "Prediction limits for the Weibull distribution utilizing simulation," *Computational Statistics & Data Analysis*, vol. 17, pp. 327–336, 1994.

[7] N.A. Nechval and K.N. Nechval, "Invariant embedding technique and its statistical applications," in *Conference Volume of Contributed Papers of the 52nd Session of the International Statistical Institute*, Finland, Helsinki: ISI–International Statistical Institute,1999, pp.1–2. Available:http://www.stat.fi/isi99/proceedings/arkisto/varasto/nech0902.pdf.

[8] N.A. Nechval and E.K. Vasermanis, *Improved Decisions in Statistics*. Riga: SIA "Izglitibas soli", 2004.

[9] N.A. Nechval, G. Berzins, M. Purgailis, and K.N. Nechval, "Improved estimation of state of stochastic systems via invariant embedding technique", *WSEAS Transactions on Mathematics*, vol. 7, pp. 141–159, 2008.

[10] N.A. Nechval, M. Purgailis, G. Berzins, K. Cikste, J. Krasts, and K. N. Nechval, "Invariant embedding technique and its applications for improvement or optimization of statistical decisions," in *Analytical and Stochastic Modeling Techniques and Applications*, K. Al-Begain, D. Fiems, and W. Knottenbelt (Eds.). LNCS, vol. 6148, Berlin, Heidelberg: Springer-Verlag, 2010, pp. 306 –320.

[11] N.A. Nechval, M. Purgailis, K. Cikste, G. Berzins, U. Rozevskis, and K.N. Nechval, "Prediction model selection and spare parts ordering policy for efficient support of maintenance and repair of equipment," in *Analytical and Stochastic Modeling Techniques and Applications*, K. Al-Begain, D. Fiems, and W. Knottenbelt (Eds.). LNCS, vol. 6148, Berlin, Heidelberg: Springer-Verlag, 2010, pp. 321–338.

[12] N.A. Nechval, M. Purgailis, K. Cikste, G. Berzins, and K.N. Nechval, "Optimization of statistical decisions via an invariant embedding technique," in *Lecture Notes in Engineering and Computer Science: World Congress on Engineering 2010*, pp. 1776 –1782.

[13] N.A. Nechval, M. Purgailis, K. Cikste, and K.N. Nechval, "Planning inspections of fatigued aircraft structures via damage tolerance approach," in *Lecture Notes in Engineering and Computer Science: World Congress on Engineering 2010*, pp. 2470 –2475.

[14] N.A. Nechval and M. Purgailis, "Improved state estimation of stochastic systems via a new technique of invariant embedding," in *Stochastic Control*, Chris Myers (Ed.). Croatia, India, Publisher: Sciyo, 2010, pp. 167 –193.