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 smallest of a set of future observations, based on a small sample size. The tables provided are specified for a given assurance level of 0.95. This method requires lot sizes ordered past observations. Mann [3] extended the results for two or three preliminary sample of past observations. An expression for the warranty period (time before the failure of the smallest ordered observation among a set of future observations) was derived as a function of the well-known relationship between the Weibull distribution and the Type I extreme value distribution, because the results for the Weibull distribution are easily obtained from those for the Gumbel distribution (1) where

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

where both distribution parameters (\(\delta\)– shape, \(\beta\) – 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@janik.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 Institute, Riga LV-1019, Latvia (e-mail: konstan@sti.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 $X_1 \leq \ldots \leq X_k$ 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 $\theta$ is a parameter (in general, vector). Then the joint probability density function of $X_1 \leq \ldots \leq X_k$ and the $l$th order statistics $X_i$ ($1 \leq k < l \leq m$) is given by

$$f_\theta(x_1, \ldots, x_k) = f_\theta(x_1, \ldots, x_k) f_\theta(x_l | x_k),$$

where

$$f_\theta(x_1, \ldots, x_k) = \frac{m!}{(m-k)!} \prod_{i=1}^k f_\theta(x_i) | F_\theta(x_k) |^{m-k},$$

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

$$= \frac{(m-k)!}{(l-k-1)!(m-l)!} \sum_{j=0}^{l-k-1} \left[ \frac{F_\theta(x_l) - F_\theta(x_k)}{1 - F_\theta(x_k)} \right]^{m-l-j} x_k^{m-l-j} \frac{f_\theta(x_l)}{1 - F_\theta(x_k)}$$

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

Proof. The joint density of $X_1 \leq \ldots \leq X_l$ and $X_l$ is given by

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

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

$$= f_\theta(x_1, \ldots, x_k, x_l) f_\theta(x_l | x_k).$$

It follows from (4) and (6) that

$$f_\theta(x_l | x_1, \ldots, x_k) = \frac{f_\theta(x_1, \ldots, x_k, x_l)}{f_\theta(x_1, \ldots, x_k)} = f_\theta(x_l | x_k),$$

i.e., the conditional distribution of $X_l$, given $X_k=x_k$ for all $i=1,\ldots, k$, 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. \hfill \Box

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

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

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

$$S_i = \frac{\mu - \mu_k}{\sigma}, \quad V_i = \frac{\sigma}{\sigma_k},$$

conditional on fixed

$$Z^{(k)} = (z_1, \ldots, z_k),$$

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

$$\hat{\mu} = \hat{\sigma} \ln \left( \sum_{i=1}^k e^{Z_i/\hat{\sigma}} + (m-k)e^{Z_k/\hat{\sigma}} \right)/k,$$

and

$$\hat{\sigma} = \sqrt{\sum_{i=1}^k Z_i e^{Z_i/\hat{\sigma}} + \sum_{i=1}^k (m-k) e^{Z_k/\hat{\sigma}}} - \frac{1}{k} \sum_{i=1}^k Z_i,$$

is given by

$$f(s_1, v_2 | Z^{(k)}) = \theta^*(Z^{(k)}) v_2^{k-2} \exp \left( v_2 \sum_{i=1}^k z_i \right)$$

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

$$= f(v_2 | Z^{(k)}) f(s_1 | v_2, Z^{(k)}), \quad s_1 \in (-\infty, \infty), \quad v_2 \in (0, \infty),$$

where

$$\theta^*(Z^{(k)}) = \left( \Gamma(k) \right)^{1/k} \left( v_2^{k-2} \exp \left( v_2 \sum_{i=1}^k z_i \right) \right)^{-1} \left( \frac{\sum \exp(z_i v_2) + (m-k) \exp(z_k v_2)}{v_2} \right)^{-1},$$

is the normalizing constant.
\[ f(v_2 \mid z^{(k)}) = \vartheta(z^{(k)}) v_2^{-k} \exp \left( v_2 \sum_{i=1}^{k} z_i \right) \]

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

\[ \vartheta(z^{(k)}) = \left\{ \begin{array}{ll}
\int_{0}^{v_2} \exp \left( v_2 \sum_{i=1}^{k} z_i \right) \\
\int_{v_2}^{\infty} \exp \left( v_2 \sum_{i=1}^{k} z_i \right)
\end{array} \right. \]

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

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

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

\[ s_1 \in (\infty, \infty). \]

\[ \text{Proof.} \] The joint density of \( X_1 \leq \ldots \leq X_k \) is given by

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

\[ \times \exp \left( - (m-k) \exp \left( \frac{x_k - \mu}{\sigma} \right) \right). \]

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 \( z^{(k)} = (z_1, \ldots, z_k) \), is

\[ f(s_1, v_2 \mid z^{(k)}) ds_1 dv_2 \]

\[ = \vartheta(z^{(k)}) v_2^{-k} \exp \left( v_2 \sum_{i=1}^{k} z_i \right) e^{v_2} \]

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

\[ s_1 \in (-\infty, \infty), \quad v_2 \in (0, \infty). \]

This ends the proof. \[ \square \]

**Theorem 3.** Let \( X_1 \leq \ldots \leq X_k \) 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

\[ V_1 = X_j - X_k \]

\[ S_2 = X_k - \mu, \]

where \( X_j (1 \leq k < l \leq m) \) is the \( l \)th order statistic from the same sample, is given by

\[ f^{-1}(v_1, s_2) = f(v_1 \mid s_2) f(s_2), \]

where

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

\[ \times e^{-(m-l+j) e^{v_2} (e^{s_2} - 1)} (m-l+j+1) e^{v_2} e^{s_2}, \quad 0 < v_1 < \infty, \]

\[ f(s_2) = \frac{1}{B(k, m-k+1)} \sum_{j=0}^{k-1} \frac{(k-1)}{j} \frac{1}{e^{(m-k+j) e^{v_2} e^{s_2}}}, \quad -\infty < s_2 < \infty, \]

\[ \text{Proof.} \] The joint density function of the order statistics \( X_k, X_l \) (1 \leq k < l \leq m) is given by

\[ f^\pi(x_k, x_l) = f^\pi(x_k \mid x_l) f^\pi(x_l). \]

It will be noted that

\[ f^\pi^\pi(x_k \mid x_l) dx_l = \frac{m!}{(l-k)! (m-l)!} \sum_{j=0}^{l-k-1} \frac{(l-k-1)}{j} \frac{1}{m-l+j+1} \]

\[ \times e^{-(m-l+j) e^{v_2} (e^{s_2} - 1)} (m-l+j+1) e^{v_2} e^{s_2} \]

\[ = f(v_1 \mid s_2) dv_2, \quad 0 < v_1 < \infty, \]

\[ f^\pi^\pi(x_k) dx_k = \frac{m!}{(k-1)! (m-k)!} \]

\[ \times \left[ (F(x_k))^{k-1} (1 - F(x_k))^{m-k} f^\pi(x_k) \right] dx_k \]

\[ = \frac{1}{B(k, m-k+1)} \sum_{j=0}^{k-1} \frac{(k-1)}{j} \frac{1}{e^{(m-k+j) e^{v_2} e^{s_2}}}, \quad -\infty < s_2 < \infty, \]

This ends the proof. \[ \square \]

**Corollary 3.1.** The probability density function of the pivotal quantity \( V_1 \) is given by

\[ f^{-1}(v_1) = \int_{-\infty}^{v_1} f(v_1 \mid s_2) ds_2 = \int_{-\infty}^{v_1} f(v_1 \mid s_2) f(s_2) ds_2. \]

**Corollary 3.2.** The joint probability density function of the pivotal quantities
\[ V_1 = \frac{X_1 - X_k}{\sigma}, \quad V_2 = \frac{\sigma}{\tilde{\sigma}} \]  \hspace{1cm} (29)

is given by

\[ f(v_1, v_2) = f^{-1}(v_1) f(v_2 | z^{(k)}), \]  \hspace{1cm} (30)

### B. Piecewise-Linear Loss Function

We shall consider the interval prediction problem for the \( l \)th order statistic \( X_l, k \leq l \leq m \), in the same sample of size \( m \) for the situation where the first \( k \) observations \( X_j < X_j < \cdots < X_k \), \( 1 \leq k \leq m \), have been observed. Suppose that we assert that an interval \( d = (d_1, d_2) \) 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 \( c_1(X_l - X_1) \) 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 \( \mu \) 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 \( d = (d_1, d_2) \) 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, \( \sigma \) 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

\[
r(\theta, d) = \begin{cases} 
\frac{c_1(X_l - X_1)}{\sigma} + \frac{c_2(d_2 - d_1)}{\sigma} & (X_l < d_1), \\
\frac{c_2(X_l - d_2)}{\sigma} & (d_1 \leq X_l \leq d_2), \\
\frac{c_1(X_l - X_1)}{\sigma} + \frac{c_2(d_2 - d_1)}{\sigma} & (X_l > d_2).
\end{cases} \hspace{1cm} (31)
\]

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 \( \mu \) (the location parameter) and \( \sigma \) (the scale) into \( cz + b \) and \( c\sigma \), respectively, where \( b \) lies in the range of \( \mu, c > 0 \). This group acts transitively on the parameter space. Thus, the problem is to find the best invariant interval predictor of \( X_l \),

\[ d^* = \arg \min_{d \in D} R(\theta, d), \hspace{1cm} (32) \]

where \( D \) is a set of invariant interval predictors of \( X_l \).

**R(\theta, d) = E_{\theta} [r(\theta, d)] \** is a risk function.

### C. Transformation of the Loss Function

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

\[ r(\theta, d) = \tilde{r}(V, \eta), \hspace{1cm} (33) \]

where

\[
\tilde{r}(V, \eta) = \begin{cases} 
\frac{c_1(V_1 + \eta V_2)}{\eta^2} + \frac{c_2(\eta_2 - \eta_2) V_2}{\eta^2} & (V_1 < \eta V_2), \\
\frac{c_2(\eta_2 - \eta_2) V_2}{\eta^2} & (\eta V_2 \leq V_1 \leq \eta_2 V_2), \\
\frac{c_2(\eta_2 - \eta_2) V_2}{\eta^2} + \frac{c_2(\eta_2 - \eta_2) V_2}{\eta^2} & (V_1 > \eta_2 V_2),
\end{cases} \hspace{1cm} (34)
\]

and

\[
\frac{\partial E[\tilde{r}(V, \eta)]}{\partial \eta_2} = \int_0^\infty v f(v, \eta_2) dv - c, \hspace{1cm} (40)
\]

and

\[
\frac{\partial E[\tilde{r}(V, \eta)]}{\partial \eta_1} = \int_0^\infty v f(v, \eta_1) dv [c(Q(\eta_1)) - c], \hspace{1cm} (41)
\]

D. Risk Function

It follows from (34) that the risk associated with \( d \) and \( \theta \) can be expressed as

\[
R(\theta, d) = E_{\theta} [r(\theta, d)] = E[\tilde{r}(V, \eta)]
\]

\[
= c_1 \int_0^{\infty} v f(v_1, v_2) dv_1 dv_2 + c_2 \int_0^{\infty} (\eta_1 - \eta_2) f(v_1, v_2) dv_1 dv_2 + c(\eta_1 - \eta_2) \int_0^{\infty} v f(v_1, v_2) dv_1 dv_2,
\]

which is constant on orbits when an invariant predictor (decision rule) \( d \) 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

\[ u_2 f(u_1, u_2) \int_0^{\infty} u_2 f(u_1, u_2) du_1 du_2 \]

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 \( c_1/c_2 < 1 \) then the optimal invariant linear-loss interval predictor of \( X_l \) based on \( X \) is \( d^*(x_l + \eta S_l, x_l + \eta S_l) \), where

\[ Q(\eta_1) = c_1/c_2, \quad Q(\eta_2) = 1 - c_2/c_2. \hspace{1cm} (38) \]

(ii) If \( c_1 + c_2 \geq 1 \) then the optimal invariant linear-loss interval predictor of \( X_l \) based on \( X \) degenerates into a point predictor \( X_l + \eta S_l \), where

\[ Q(\eta_1) = c_2/(c_1 + c_2). \hspace{1cm} (39) \]

**Proof.** From (36)

\[
\frac{\partial E[\tilde{r}(V, \eta)]}{\partial \eta_1} = c_1 \int_0^{\infty} v f(v_1, v_2) dv_1 dv_2 - c \int_0^{\infty} v f(v_1, v_2) dv_1 dv_2,
\]

\[
= \int_0^{\infty} v f(v_1, v_2) dv_1 dv_2 [c(Q(\eta_1)) - c], \hspace{1cm} (40)
\]

and

\[
\frac{\partial E[\tilde{r}(V, \eta)]}{\partial \eta_2} = \int_0^{\infty} v f(v_1, v_2) dv_1 dv_2 [-c_2 (1 - Q(\eta_2)) + c], \hspace{1cm} (41)
\]
where

\[ Q(\eta) = \int_{0}^{\eta} q(w)dw, \quad (42) \]

\[ q(w) = \frac{\int_{0}^{w} f(wv_{2}, v_{2})dv_{2}}{\int_{0}^{w} v_{2}f(v_{1}, v_{2})dv_{1}dv_{2}}, \quad \text{for case (i)} \]

\[ W = V_{1} / V_{2}, \quad (44) \]

Now \( \partial E[\hat{r}(V, \eta)]/\partial \eta_{1} = \partial E[\hat{r}(V, \eta)]/\partial \eta_{2} = 0 \) if and only if (38) holds. Thus, \( E[\hat{r}(V, \eta)] \) provided (38) has a solution with \( \eta_{1} < \eta_{2} \) and this is so if \( 1/c_{2} > c/c_{1} \). It is easily confirmed that this \( \eta^{*} = (\eta_{1}^{*}, \eta_{2}^{*}) \) gives the minimum value of \( E[\hat{r}(V, \eta)] \). Thus (i) is established.

If \( c/c_{1} + c/c_{2} \geq 1 \) then the minimum of \( E[\hat{r}(V, \eta)] \) in the region \( \eta_{2} \geq \eta_{1} \) occurs where \( \eta^{*} = (\eta_{1}^{*}, \eta_{2}^{*}) \), \( \eta^{*} \) being determined by setting

\[ \partial E[\hat{r}(V, \eta_{*}, \eta_{*})]/\partial \eta_{2} = \partial E[\hat{r}(V, \eta_{*}, \eta_{*})]/\partial \eta_{1} = 0, \quad (45) \]

and this reduces to

\[ c_{1}Q(\eta_{*}) - c_{2}[1 - Q(\eta_{*})] = 0, \quad (46) \]

which establishes (ii). \( \Box \)

**Corollary 4.1.** The minimum risk of the optimal invariant predictor of \( X_{1} \) based on \( V_{1} \) is given by

\[ R(\theta, d^{*}) = E_{\theta}[\hat{r}(\theta, d^{*})] = E[\hat{r}(V, \eta)] \]

\[ = -c_{1}\int_{0}^{\eta_{2}^{*}} \int_{0}^{w} f(v_{1}, v_{2})dv_{1}dv_{2} + c_{2}\int_{0}^{\eta_{2}^{*}} \int_{\eta_{1}^{*}}^{w} f(v_{1}, v_{2})dv_{1}dv_{2} \]

\[ \text{for case (i) with } \eta^{*} = (\eta_{1}^{*}, \eta_{2}^{*}) \text{ as given by (38) and for case (ii) with } \eta^{*} = (\eta_{1}^{*}, \eta_{2}^{*}) \text{ as given by (39)}. \]

**Proof.** These results are immediate from (36) when use is made of \( \partial E[\hat{r}(V, \eta)]/\partial \eta_{1} = \partial E[\hat{r}(V, \eta)]/\partial \eta_{2} = 0 \) in case (i) and \( \partial E[\hat{r}(V, \eta_{1}, \eta_{2})]/\partial \eta_{2} = 0 \) in case (ii). \( \Box \)

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

\[ V_{1} = (V_{1}, V_{2}), \quad V_{1} = (X_{1} - X_{2})/\sigma, \quad V_{2} = X_{2} / \sigma; \]

\[ \eta^{*} = (\eta_{1}^{*}, \eta_{2}^{*}), \quad \eta_{1}^{*} = (d_{2} - X_{2}) / X_{k}, \quad \eta_{2}^{*} = (d_{2} - X_{2}) / X_{k}. \]

Let us assume that \( (U_{1}, U_{2}) \) is a random vector having density function

\[ u_{2}f_{0}(u_{1}, u_{2}) \left[ \int_{0}^{u_{1}} f_{0}(u_{1}, u_{2})du_{1}du_{2} \right]^{-1} (u_{1}, u_{2} > 0), \quad (49) \]

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 \( c/c_{1} + c/c_{2} \leq 1 \), then the optimal invariant linear-loss interval predictor of \( X_{1} \) based on \( X_{2} \) is \( d^{*} = (1 + \eta_{1}^{*})X_{2} \), (i.e., the probability density function of \( W/V_{2} \), where

\[ Q_{0}(\eta_{1}^{*}) = c / c_{1}, \quad Q_{0}(\eta_{2}^{*}) = 1 - c / c_{2}. \]

(ii) If \( c/c_{1} + c/c_{2} \geq 1 \), then the optimal invariant linear-loss interval predictor of \( X_{1} \) based on \( X_{2} \) degenerates into a point predictor \( (1 + \eta_{1}^{*})X_{2} \), where

\[ Q_{0}(\eta_{1}^{*}) = c_{2} / (c_{1} + c_{2}). \]

**Proof.** For the proof we refer to Theorem 1. \( \Box \)

**Corollary 5.1.** The minimum risk of the optimal invariant predictor of \( X_{1} \) based on \( X_{2} \) is given by

\[ R(\theta, d^{*}) = E_{\theta}[\hat{r}(\theta, d^{*})] = E[\hat{r}(V, \eta^{*})] \]

\[ = -c_{1}\int_{0}^{\eta_{2}^{*}} \int_{0}^{w} f(v_{1}, v_{2})dv_{1}dv_{2} + c_{2}\int_{0}^{\eta_{2}^{*}} \int_{\eta_{1}^{*}}^{w} f(v_{1}, v_{2})dv_{1}dv_{2} \]

\[ \text{for case (i) with } \eta^{*} = (\eta_{1}^{*}, \eta_{2}^{*}) \text{ as given by (50) and for case (ii) with } \eta_{1}^{*} = \eta_{2}^{*} = \eta_{*} \text{ as given by (51)}. \]

**Proof.** For the proof we refer to Corollary 4.1. \( \Box \)

III. EQUIVALENT CONFIDENCE COEFFICIENT

For case (i) when we obtain an interval predictor for \( X_{1} \) 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 \( V = (V_{1}, V_{2}) \) is a random vector having density function \( f(v_{1}, v_{2}) \) \((v_{1}, v_{2}) \geq 0) \) which 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

\[ h(w) = \int_{0}^{\infty} f(v_{1}, v_{2})dv_{1}dv_{2}. \]

Then the confidence coefficient based on \( X \) and associated with the optimum prediction interval \( d^{*} = (d_{1}, d_{2}) \), where \( d_{1} = X_{2} + \eta_{1}^{*}\sigma, \quad d_{2} = X_{2} + \eta_{2}^{*}\sigma \), is

\[ \Pr\{d^{*} : d_{1} < X_{1} < d_{2} | \mu, \sigma\} = H[Q^{-1}(1 - c / c_{2})] - H[Q^{-1}(c / c_{1})]. \]
Proof. The confidence coefficient for \( d' \) corresponding to \((\mu, \sigma)\) is given by

\[
\Pr\{ |X_1, \sigma|: X_k + \eta_1 \sigma < X_1 < X_k + \eta_2 \sigma | \mu, \sigma \} = \Pr\{ v_1, v_2: \eta_1 < v_1 / v_2 < \eta_2 \} = H(\eta_2) - H(\eta_1) = H[Q^{-1}(1-c/c_2)] - H[Q^{-1}(c/c_1)].
\]

This is independent of \((\mu, \sigma)\). \(\square\)

**Theorem 7.** Suppose that \( V=(V_1, V_2) \) is a random vector having density function \( f_\theta(v_1, v_2) \) \((v_1, v_2 \geq 0)\), where \( f_0 \) is defined by

\[
f_\theta(x_k, x_1) = \frac{1}{B(k, l-k)B(l, m-l+1)} \times [F_\theta(x_k)]^{k-1} [F_\theta(x_1) - F_\theta(x_k)]^{l-1-1} \times [1 - F_\theta(x_1)]^{m-l} f_\theta(x_k) f_\theta(x_1),
\]

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

\[
h_0(w) = \int_0^\infty v_2 f_0(wv_2, v_2) dv_2.
\]

Then the confidence coefficient based on \( X_1 \) and associated with the optimum prediction interval \( d=(d_1, d_2) \), where \( d_1=(1+\eta_1) X_1 \), \( d_2=(1+\eta_2) X_1 \), is

\[
\Pr\{ d' : d_1 < X_1 < d_2 | \mu, \sigma \} = H_0[Q^{-1}(1-c/c_2)] - H_0[Q^{-1}(c/c_1)].
\]

**Proof:** For the proof we refer to Theorem 6. \(\square\)

IV. CONCLUSION

In many statistical decision problems it is reasonable to 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 transversely 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