Ann. Inst. Statist. Math.
Vol. 47, No. 1, 119–128 (1995)

ADMISSIBILITY OF PREDICTION INTERVALS

YOSHIKAZU TAKADA

Department of Mathematics, Faculty of Science, Kumamoto University, Kumamoto 860, Japan

(Received July 15, 1993; revised June 10, 1994)

Abstract. Admissibility of prediction intervals is considered in a specified family. It is shown that the best invariant prediction interval is strongly admissible in a location family and in a scale family. Though the similar result has not been obtained for a location and scale family, the best invariant prediction interval for a normal distribution is shown to be weakly admissible.

Key words and phrases: Prediction interval, admissibility, invariance, location family, scale family, location and scale family.

1. Introduction

A prediction set is a set determined by the observed sample and having the property that it contains the result of a future sample with a specified probability. A prediction interval corresponds to the special case that the set is an interval. Prediction intervals are widely used for reliability problems and other related problems. Patel (1989) provided a review on the construction of prediction intervals. The present paper deals with admissibility of prediction sets.

Suppose that the observed variable $\boldsymbol{X}$ and the predicted variable Y have a joint distribution, which depends on an unknown parameter θ. A prediction set $S(\boldsymbol{X})$ is examined by its coverage probability $P_\theta\{Y \in S(\boldsymbol{X})\}$ and its volume $\xi(S(\boldsymbol{X}))$ with respect to some measure ξ on the sample space of Y. The larger its coverage probability and the smaller its volume are, the better the prediction set is. Hence a prediction set $S(\boldsymbol{X})$ is said to be strongly admissible if there exists no other prediction set $S'(\boldsymbol{X})$ such that

$$P_\theta\{Y \in S'(\boldsymbol{X})\} \geq P_\theta\{Y \in S(\boldsymbol{X})\} \quad \text{for all } \theta \tag{1.1}$$

and

$$E_\theta\{\xi(S'(\boldsymbol{X}))\} \leq E_\theta\{\xi(S(\boldsymbol{X}))\} \quad \text{for all } \theta \tag{1.2}$$

and the strict inequality holds for at least one θ either in (1.1) or in (1.2). Weak admissibility is defined by replacing (1.2) by

$$\xi(S'(\boldsymbol{x})) \leq \xi(S(\boldsymbol{x})) \quad \text{for almost all } \boldsymbol{x}$$

and requiring the strict inequality to hold for at least one θ in (1.1). See Joshi (1970). Strong admissibility of a prediction set implies its weak admissibility.

Section 2 deals with a strong admissibility of the best invariant prediction set in a location family and in a scale family. It has not been known if the similar result holds in a location and scale family or not. In Section 3 we shall take a normal distribution with unknown mean and unknown variance as an example of a location and scale family and show that the best invariant prediction interval is weakly admissible among prediction intervals.

2. Strong admissibility

Given a prediction set $S(\boldsymbol{X})$, consider a function ϕ defined by

$$\phi(\boldsymbol{x}, y) = \begin{cases} 1 & \text{if } y \in S(\boldsymbol{x}), \\ 0 & \text{otherwise.} \end{cases}$$

Then it holds that

$$P_\theta\{Y \in S(\boldsymbol{X})\} = E_\theta\{\phi(\boldsymbol{X}, Y)\} \tag{2.1}$$

and

$$E_\theta\{\xi(S(\boldsymbol{X}))\} = E_\theta\left\{\int \phi(\boldsymbol{X}, y)\xi(dy)\right\}. \tag{2.2}$$

Conversely, every function ϕ with $0 \le \phi(\boldsymbol{x}, y) \le 1$ defines a prediction procedure by which a randomized prediction set is constructed such that (2.1) and (2.2) are satisfied. In this section, prediction sets considered are randomized and identified with such a function ϕ. Further, strong admissibility is discussed among the class of randomized prediction sets.

First we treat a location family. Suppose that the observed sample $\boldsymbol{X}$ consists of n random variables $X_1, \dots, X_n$ and $\boldsymbol{X}$ and Y have a joint probability density function

$$f(x_1 - \theta, \dots, x_n - \theta, y - \theta) \tag{2.3}$$

with respect to Lebesgue measure, where f is known and θ is unknown.

A prediction set ϕ is said to be invariant if

$$\phi(x_1 + a, \dots, x_n + a, y + a) = \phi(\boldsymbol{x}, y) \qquad \text{for all } \boldsymbol{x},\ y,\ a. \tag{2.4}$$

An invariant prediction set is said to be the best invariant with respect to ξ if it uniformly minimizes (2.2) among the class of invariant prediction sets with coverage probability not less than a specified value.

Assuming $\xi(dy) = dy$, a prediction set given by

$$\phi_0(\boldsymbol{x}, y) = \begin{cases} 1 & \text{if } F(\boldsymbol{x}, y) \ge b \\ 0 & \text{otherwise} \end{cases} \tag{2.5}$$

is shown to be the best invariant, where $b > 0$ is a suitable constant,

$$F(\boldsymbol{x}, y) = \frac{\int f(x_1 - \theta, \ldots, x_n - \theta, y - \theta) d\theta}{\int h(x_1 - \theta, \ldots, x_n - \theta) d\theta}, \tag{2.6}$$

and $h(x_1 - \theta, \ldots, x_n - \theta)$ is the probability density function of $\boldsymbol{X}$. The proof is given in Section 4 of Takeuchi (1975). See also Takada (1983) and Hooper (1986). We now consider the strong admissibility of the prediction set (2.5).

Let $X = X_1$ and $Z_i = X_{i+1} - X_1$ $(i = 1, \ldots, n-1)$. Then the probability density function of $\boldsymbol{Z} = (Z_1, \ldots, Z_{n-1})$ becomes

$$u(\boldsymbol{z}) = \int h(t, t + z_1, \ldots, t + z_{n-1}) dt,$$

and the joint probability density function of X, $\boldsymbol{Z}$ and Y is written as

$$p(x - \theta, \boldsymbol{z}, y - \theta),$$

where

$$p(x - \theta, \boldsymbol{z}, y - \theta) = f(x - \theta, x - \theta + z_1, \ldots, x - \theta + z_{n-1}, y - \theta).$$

It follows from (2.6) that

$$\begin{aligned} F(\boldsymbol{x}, y) &= u(\boldsymbol{z})^{-1} \int p(t, \boldsymbol{z}, y - x + t) dt \\ &= g(y - x, \boldsymbol{z}) \qquad \text{(say)}. \end{aligned} \tag{2.7}$$

Let $k(x - \theta)$ be the probability density function of X_1.

THEOREM 2.1. *Suppose that for a location family* (2.3)

$$\int |t| k(t) dt < \infty, \tag{2.8}$$

$$\iint_{[g(y,\boldsymbol{z})=b]} u(\boldsymbol{z}) d\boldsymbol{z} dy = 0. \tag{2.9}$$

Then the best invariant prediction set (2.5) *is strongly admissible.*

PROOF. The proof follows the method used by Joshi (1970) who proved the strong admissibility of the best invariant confidence intervals in a location family.

Suppose that there exists a prediction set $\phi_1(X, Y)$ such that

$$E_\theta\{\phi_1(\boldsymbol{X}, Y)\} \geq E_\theta\{\phi_0(\boldsymbol{X}, Y)\} \qquad \text{for all } \theta$$

and

$$E_\theta\left\{\int \phi_1(\boldsymbol{X}, y) dy\right\} \leq E_\theta\left\{\int \phi_0(\boldsymbol{X}, y) dy\right\} \qquad \text{for all } \theta.$$

Write

$$\phi(\boldsymbol{x}, \cdot) = \int \phi(\boldsymbol{x}, y) dy$$

and let

$$q(\boldsymbol{x}, y) = [b\phi_0(\boldsymbol{x}, \cdot) - \phi_0(\boldsymbol{x}, y)] - [b\phi_1(\boldsymbol{x}, \cdot) - \phi_1(\boldsymbol{x}, y)].$$

Then it holds that for all θ

$$\iiint q(x, \boldsymbol{z}, y) p(x - \theta, \boldsymbol{z}, y - \theta) dx d\boldsymbol{z} dy \geq 0,$$

so that for any $L > 0$

$$\int_{-L}^{L} \left\{ \iiint q(x, \boldsymbol{z}, y) p(x - \theta, \boldsymbol{z}, y - \theta) dx d\boldsymbol{z} dy \right\} d\theta \geq 0. \tag{2.10}$$

It is easy to see that the order of integration can be changed. Hence the left hand side of (2.10) is equal to

$$\begin{aligned}
(2.11)\quad & \int d\boldsymbol{z} \int dy \left\{ \int dx \int_{-L}^{L} q(x, \boldsymbol{z}, y) p(x - \theta, \boldsymbol{z}, y - \theta) d\theta \right\} \\
&= \int d\boldsymbol{z} \int dy \left\{ \int dx \int_{x-L}^{x+L} q(x, \boldsymbol{z}, y) p(t, \boldsymbol{z}, y - x + t) dt \right\} \\
&= \int d\boldsymbol{z} \int dy \left\{ \int dx \int_{x-L}^{x+L} q(x, \boldsymbol{z}, y + x) p(t, \boldsymbol{z}, y + t) dt \right\} \\
&= \int d\boldsymbol{z} \int dy \left\{ \int_{-L/2}^{L/2} dx \int_{-L/2}^{L/2} dt + \int_{-3L/2}^{-L/2} dx \int_{-L/2}^{L+x} dt \right. \\
&\qquad + \int_{L/2}^{3L/2} dx \int_{x-L}^{L/2} dt + \int_{L/2}^{\infty} dt \int_{t-L}^{t+L} dx \\
&\qquad \left. + \int_{-\infty}^{-L/2} dt \int_{t-L}^{t+L} dx \right\} q(x, \boldsymbol{z}, y + x) p(t, \boldsymbol{z}, y + t) \\
&= T_1 + T_2 + T_3 + T_4 + T_5 \qquad \text{(say)}.
\end{aligned}$$

See (21) of Joshi (1970). From (2.5) and (2.7)

$$1 = \int g(y - x, \boldsymbol{z}) dy \geq b \int_{[y; g(y-x, \boldsymbol{z}) \geq b]} dy = b\phi_0(x, \boldsymbol{z}, \cdot),$$

so that we have

$$q(x, \boldsymbol{z}, y) \leq 2.$$

Then

$$\begin{aligned}
T_4 + T_5 &\le 2\int dz \int dy \left\{ \int_{L/2}^{\infty} dt \int_{t-L}^{t+L} dx + \int_{-\infty}^{-L/2} dt \int_{t-L}^{t+L} dx \right\} p(t, z, y+t) \\
&= 4L \int_{|t|>L/2} k(t)dt \\
&\le 8 \int_{|t|>L/2} |t|k(t)dt.
\end{aligned}$$

Hence from (2.8)

$$\limsup_{L\to\infty} (T_4 + T_5) \le 0. \tag{2.12}$$

The proof of the results that

$$\limsup_{L\to\infty} T_2 \le 0 \quad \text{and} \quad \limsup_{L\to\infty} T_3 \le 0 \tag{2.13}$$

are similar to those of (66) and (67) of Joshi (1970) and are therefore omitted. Combining (2.12) and (2.13) with (2.11) and using (2.10), we have

$$\liminf_{L\to\infty} T_1 \ge 0. \tag{2.14}$$

It can been shown that

$$\liminf_{L\to\infty} T_1 \le \int dx \left\{ \int dz \int dy \int q(x, z, y+x)p(t, z, y+t)dt \right\}. \tag{2.15}$$

See (32) of Joshi (1970). Hence it follows from (2.14) and (2.15) that

$$\int dx \left\{ \int dz \int dy \int q(x, z, y)p(x-\theta, z, y-\theta)d\theta \right\} \ge 0. \tag{2.16}$$

Since

$$\begin{aligned}
&\iint \{b\phi(x, z, \cdot) - \phi(x, z, y)\} p(x-\theta, z, y-\theta)d\theta dy \\
&\qquad = bu(z)\phi(x, z, \cdot) - u(z)\int \phi(x, z, y)g(y-x, z)dy \\
&\qquad = u(z)\int \{b - g(y-x, z)\}\phi(x, z, y)dy,
\end{aligned}$$

it follows from (2.5) and (2.7) that

$$\iint q(x, z, y)p(x-\theta, z, y-\theta)d\theta dy = -u(z)W(x, z),$$

where

$$W(x, \boldsymbol{z}) = \int |\phi_0(x, \boldsymbol{z}, y) - \phi_1(x, \boldsymbol{z}, y)||g(y - x, z) - b|dy.$$

Therefore from (2.16)

$$\iint W(x, \boldsymbol{z})u(\boldsymbol{z})dxd\boldsymbol{z} = 0.$$

Let $B = \{(x, \boldsymbol{z}, y); g(y - x, \boldsymbol{z}) = b\}$. Then by (2.9)

$$\begin{aligned}\iiint_B u(\boldsymbol{z})dxd\boldsymbol{z}dy &= \int dx \left\{ \iint_{[g(y-x,\boldsymbol{z})=b]} u(\boldsymbol{z})d\boldsymbol{z}dy \right\} \\ &= \int dx \left\{ \iint_{[g(y,\boldsymbol{z})=b]} u(\boldsymbol{z})d\boldsymbol{z}dy \right\} \\ &= 0.\end{aligned}$$

Hence we have

$$\phi_0(\boldsymbol{x}, y) = \phi_1(\boldsymbol{x}, y) \qquad \text{a.e.}$$

with respect to $u(\boldsymbol{z})dxd\boldsymbol{z}dy$, so that the best invariant prediction set is strongly admissible, which completes the proof.

Example 1. Let $X_1, \ldots, X_n$ be an observed sample from a normal distribution $N(\mu, \sigma^2)$ and let $Y_1, \ldots, Y_m$ be a future sample from the same distribution. Based on $\boldsymbol{X} = (X_1, \ldots, X_n)$, we consider a prediction of the sample mean $\bar{Y} = (y_1 + \cdots + Y_m)/m$. Assume that $\theta = \mu$ is unknown but σ is known. Then it easily follows from (2.5) that

$$I(\boldsymbol{X}) = (\bar{X} - c\sigma, \bar{X} - c\sigma)$$

is the best invariant prediction interval, where $c(> 0)$ is a suitable constant and $\bar{X} = (X_1 + \cdots + X_n)/n$. It is easy to see that the conditions of Theorem 2.1 are satisfied, so that $I(\boldsymbol{X})$ is strongly admissible.

Next we turn our attention to a scale family. Suppose that $\boldsymbol{X} = (X_1, \ldots, X_n)$ and Y have a joint probability density function

$$\text{(2.17)} \qquad \theta^{-(n+1)} f(x_1/\theta, \ldots, x_n/\theta, y/\theta), \qquad x_1 > 0, \ldots, x_n > 0, \quad y > 0$$

with respect to Lebesgue measure, where f is known and $\theta > 0$ is unknown.

A prediction set $\phi(\boldsymbol{x}, y)$ is said to be invariant if

$$\text{(2.18)} \qquad \phi(ax_1, \ldots, ax_n, ay) = \phi(\boldsymbol{x}, y) \qquad \text{for all} \quad \boldsymbol{x},\ y,\ a > 0.$$

Let $X_i' = \log X_i$ $(i = 1, \ldots, n)$, $Y' = \log Y$ and $\theta' = \log \theta$. Then it is easy to see that the joint probability distribution of $\boldsymbol{X}' = (X_1', \ldots, X_n')$ and Y' belongs to a location family (2.3) with location parameter θ'. Let

$$\phi'(\boldsymbol{x}', y') = \phi(\exp(x_1'), \ldots, \exp(x_n'), \exp(y')).$$

Then ϕ satisfies (2.18) if and only if ϕ' satisfies (2.4). Further,

$$E_\theta\{\phi(\boldsymbol{X}, Y)\} = E_{\theta'}\{\phi'(\boldsymbol{X}', Y')\}.$$

Assuming $\xi(dy) = y^{-1}dy$, $y > 0$,

$$E_\theta\left\{\int \phi(\boldsymbol{X}, y)\xi(dy)\right\} = E_{\theta'}\left\{\int \phi'(\boldsymbol{X}', y')dy'\right\}.$$

Hence it easily follows from (2.5) that a prediction set given by

$$\phi_0(\boldsymbol{x}, y) = \begin{cases} 1 & \text{if } F(\boldsymbol{x}, y) \geq b, \\ 0 & \text{otherwise,} \end{cases} \tag{2.19}$$

is the best invariant with respect to ξ, where

$$F(\boldsymbol{x}, y) = \frac{y\int_0^\infty \theta^{-(n+2)} f(x_1/\theta, \ldots, x_n/\theta, y/\theta)d\theta}{\int_0^\infty \theta^{-(n+1)} h(x_1/\theta, \ldots, x_n/\theta)d\theta} \tag{2.20}$$

and $\theta^{-n}h(x_1/\theta, \ldots, x_n/\theta)$ is the probability density of $\boldsymbol{X}$. Let

$$g(y, \boldsymbol{z}) = \frac{y\int_0^\infty t^n f(t, z_1 t, \ldots, z_{n-1}t, yt)dt}{\int_0^\infty t^{n-1} h(t, z_1 t, \ldots, z_{n-1}t)dt},$$

with $\boldsymbol{z} = (z_1, \ldots, z_{n-1})$. Then it follows from (2.20) that

$$F(\boldsymbol{x}, y) = g(y/x, \boldsymbol{z}), \tag{2.21}$$

where $z_i = x_{i+1}/x_1$ $(i = 1, \ldots, n-1)$ and $x = x_1$. The probability density function of $\boldsymbol{Z} = (Z_1, \ldots, Z_{n-1})$ becomes

$$u(\boldsymbol{z}) = \int_0^\infty t^{n-1} h(t, z_1 t, \ldots, z_{n-1}t)dt, \tag{2.22}$$

and let $\theta^{-1}k(x/\theta)$ be the probability density function of X_1. Now the strong admissibility of the prediction set (2.19) follows from Theorem 2.1.

THEOREM 2.2. *Suppose that for a scale family* (2.17)

$$\int_0^\infty |\log t| k(t)dt < \infty, \tag{2.23}$$

and

$$\iint_{[g(y,\boldsymbol{z})=b]} u(\boldsymbol{z})y^{-1}d\boldsymbol{z}dy = 0. \tag{2.24}$$

Then the best invariant prediction set (2.19) *is strongly admissible.*

Remark. The case that $\xi(dy) = dy$ can also be similarly treated (cf. Remark 5.1 in Cohen (1972)).

Example 2. Let $X_{(1)} \le X_{(2)} \le \cdots \le X_{(n)}$ be an ordered sample of size n from an exponential distribution with density

$$\theta^{-1}\exp(-x/\theta), \qquad x > 0, \quad \theta > 0.$$

Suppose that only the first r of these order statistics are observed. Let $Y_{(1)} \le Y_{(2)} \le \cdots \le Y_{(m)}$ be a future ordered sample of size m from the same distribution. Based on $\boldsymbol{X} = (X_{(1)}, X_{(2)}, \ldots, X_{(r)})$, we consider a prediction of $Y = Y_{(s)}$ $(1 \le s \le m)$. The joint probability density function of $\boldsymbol{X}$ and Y is given by

$$(\text{const.}) \times \theta^{-(r+1)}\exp(-v/\theta)\{1 - \exp(-y/\theta)\}^{s-1}\exp\{-(m-s+1)y/\theta\},$$

where $v = \sum_{i=1}^{r} x_{(i)} + (n-r)x_{(r)}$. Then from (2.20) we find that $F(\boldsymbol{x}, y)$ is proportional to

$$g*(y/v) = (y/v)\int_0^\infty t^r\{1 - \exp(-(y/v)t)\}^{s-1}\exp\{-t[1 + (m-s+1)y/v]\}dt.$$

It can be shown that $g*(y) > b$ is equivalent to $c_1 < y < c_2$, where $g*(c_1) = g*(c_2)$ (e.g. Takada (1979)). Hence the best invariant prediction set becomes an interval given by

$$I(\boldsymbol{X}) = (c_1 V, c_2 V).$$

We check the conditions of Theorem 2.2 to show its strong admissibility. Take $X_1 = X_{(r)}$. Then it is easy to see that (2.23) is satisfied. From (2.22) $u(\boldsymbol{z})$ is proportional to $\{z_1 + \cdots + z_{r-1} + n - r + 1\}^{-r}$, $0 < z_i < 1$ $(i = 1, \ldots, r-1)$. It follows from (2.21) that $g(y, \boldsymbol{z})$ is proportional to $g*(y/v')$ with $v' = z_1 + \cdots + z_{r-1} + n - r + 1$. Hence (2.24) is satisfied, so that $I(\boldsymbol{X})$ is strongly admissible.

3. Normal distribution

The arguments used to prove the strong admissibility in Section 2 can not be applied to a location and scale family. In this section, we consider the problem of Example 1 in Section 2 in which both μ and σ are unknown.

The usual prediction interval of $\bar{Y}$ based on $\boldsymbol{X}$ is

$$I(\boldsymbol{X}) = (\bar{X} - cS, \bar{X} + cS) \tag{3.1}$$

where $S^2 = (X_1 - \bar{X})^2 + \cdots + (X_n - \bar{X})^2$ and $c(> 0)$ is a suitable constant. It can be shown that (3.1) is the best invariant with respect to Lebesgue measure (e.g. Takeuchi (1975), Section 4). Now we shall show that (3.1) is weakly admissible among the class of prediction intervals. For the proof we need the following lemma.

LEMMA 3.1. *Suppose that a prior distribution of* μ *is* $N(0,\tau^2)$. *Then the conditional distribution of* $\bar{Y}$ *given* $\boldsymbol{X} = \boldsymbol{x}$ *is* $N(\mu(\boldsymbol{x}), \Delta^2)$, *where*

$$\mu(\boldsymbol{x}) = \frac{n\bar{x}/\sigma^2}{n/\sigma^2 + 1/\tau^2}$$

and

$$\Delta^2 = \frac{\sigma^2/m}{n/\sigma^2 + 1/\tau^2}(m/\sigma^2 + n/\sigma^2 + 1/\tau^2).$$

PROOF. It is straightforward to see that given $\boldsymbol{X} = \boldsymbol{x}$, the conditional distribution of μ is $N(\mu(\boldsymbol{x}), 1/(n/\sigma^2 + 1/\tau^2))$. Then the result follows from the fact that the conditional distribution of $\bar{Y}$ given $\boldsymbol{X} = \boldsymbol{x}$ is obtained by integrating the density of $\bar{Y}$ with respect to the conditional distribution of μ given $\boldsymbol{X} = \boldsymbol{x}$.

Given a statistic $\delta(\boldsymbol{X})$, define its risk by

$$R(\theta,\delta) = 1 - P_\theta\{\delta(\boldsymbol{X}) - cS < \bar{Y} < \delta(\boldsymbol{X}) + cS\},$$

where $\theta = (\mu, \sigma)$. Consider a prior distribution $N(0,\tau^2)$ for μ and denote the average risk by $R_\sigma(\tau,\delta)$, that is

$$R_\sigma(\tau,\delta) = \int R(\theta,\delta)(2\pi\tau^2)^{-1/2}\exp\{-\mu^2/(2\tau^2)\}d\mu.$$

Then by the lemma we get

$$R_\sigma(\tau,\delta) = 1 - E^*\{\Phi[(\delta(\boldsymbol{X}) + cS - \mu(\boldsymbol{X}))/\Delta)] - \Phi[(\delta(\boldsymbol{X}) - cS - \mu(\boldsymbol{X}))/\Delta)]\},$$

where the expectation E^* is taken with respect to the marginal distribution of $\boldsymbol{X}$. It is easy to see that

$$\begin{aligned} R_\sigma(\tau,\mu*) &= \operatorname*{Inf}_\delta R_\sigma(\tau,\delta) \\ &= 1 - E^*\{\Phi(cS/\Delta) - \Phi(-cS/\Delta)\} \\ &= 1 - E\{\Phi(cS/\Delta) - \Phi(-cS/\Delta)\}, \end{aligned} \tag{3.2}$$

where $\mu* = \mu(\boldsymbol{X})$ and we used the fact that the marginal distribution of S does not depend on the prior distribution of μ.

THEOREM 3.1. *For the normal distribution with unknown mean and unknown variance, the prediction interval* (3.1) *is weakly admissible among the class of prediction intervals.*

PROOF. Suppose that the prediction interval $(X - \bar{c}S, X + \bar{c}S)$ is not weakly admissible so that there exists a prediction interval $(L(\boldsymbol{X}), U(\boldsymbol{X}))$ such that

$$U(\boldsymbol{X}) - L(\boldsymbol{X}) \le 2cS \quad \text{a.e.} \tag{3.3}$$

and

(3.4) $\quad P_\theta\{L(\boldsymbol{X}) < Y < U(\boldsymbol{X})\} \geq P_\theta\{\bar{X} - cS < \bar{Y} < \bar{X} + cS\} \quad$ for all θ

and the strict inequality holds for at least one θ in (3.4). Let

$$\delta(\boldsymbol{X}) = (L(\boldsymbol{X}) + U(\boldsymbol{X}))/2.$$

Then from (3.3)

$$(\delta(\boldsymbol{X}) - cS, \delta(\boldsymbol{X}) + cS) \supseteq (L(\boldsymbol{X}), U(\boldsymbol{X})) \quad \text{a.e.},$$

so that from (3.4) we get

$$R(\theta, \delta) \leq R(\theta, \bar{X}) \quad \text{for all } \theta \tag{3.5}$$

and the strict inequality holds for at least one θ (cf. (2) of Joshi (1966)). Observing

$$R_\sigma(\tau, \bar{X}) = 1 - E\{\Phi(cS/\lambda) - \Phi(-cS/\lambda)\},$$

where $\lambda = \{\sigma^2(1/n + 1/m)\}^{1/2}$, it follows from (3.2) that

$$\tau\{R_\sigma(\tau, \bar{X}) - R_\sigma(\tau, \mu*)\} = 2\tau E\{\Phi(-cS/\lambda) - \Phi(-cS/\Delta)\}. \tag{3.6}$$

It is easy to see that the right hand side of (3.6) converges to zero as $\tau \to \infty$. Hence we get

$$\tau\{R_\sigma(\tau, \bar{X}) - R_\sigma(\tau, \mu*)\} \to 0 \quad \text{as} \quad \tau \to \infty. \tag{3.7}$$

Since $R(\theta, \delta)$ is a continuous function of μ, it follows from (3.5) that there exist $\epsilon > 0$, σ and $\mu_1 < \mu_2$ such that

$$R(\theta, \delta) < R(\theta, \bar{X}) - \epsilon$$

for $\theta = (\mu, \sigma)$ with $\mu_1 < \mu < \mu_2$. Then we get

$$\frac{R_\sigma(\tau, \bar{X}) - R_\sigma(\tau, \delta)}{R_\sigma(\tau, \bar{X}) - R_\sigma(\tau, \mu*)} > \frac{\epsilon \int_{\mu_1}^{\mu_2} (2\pi)^{-1/2} \exp\{-\mu^2/(2\tau^2)\} d\mu}{\tau\{R_\sigma(\tau, \bar{X}) - R_\sigma(\tau, \mu*)\}}.$$

By (3.7) the above right hand side converges to ∞ as $\tau \to \infty$. Hence there exists τ such that

$$R_\sigma(\tau, \delta) < R_\sigma(\tau, \mu*),$$

which contradicts (3.2), so that the proof is completed.

Acknowledgements

The author wishes to thank the referees for their suggestions.

References

Cohen, A. (1972). Improved confidence intervals for the variance of a normal distribution, *J. Amer. Statist. Assoc.*, **67**, 382–385.

Hooper, P. M. (1986). Invariant prediction regions with smallest expected measure, *J. Multivariate Anal.*, **18**, 117–126.

Joshi, V. M. (1966). Admissibility of confidence intervals, *Ann. Math. Statist.*, **37**, 629–638.

Joshi, V. M. (1970). Admissibility of invariant confidence procedures for estimating a location parameter, *Ann. Math. Statist.*, **41**, 1568–1581.

Patel, J. K. (1989). Prediction intervals—a review, *Comm. Statist. Theory Methods*, **18**, 2393–2465.

Takada, Y. (1979). The shortest invariant prediction interval from the largest observation from the exponential distribution, *J. Japan Statist. Soc.*, **9**, 87–91.

Takada, Y. (1983). Invariant prediction regions of future observations, *Kumamoto Journal of Science (Mathematics)*, **15**, 79–89.

Takeuchi, K. (1975). *Statistical Prediction Theory*, Baihūkan, Tokyō (in Japanese).