FIXED-SIZE CONFIDENCE REGIONS FOR THE MEAN

VECTOR OF A MULTINORMAL DISTRIBUTION

By

JAFFAR SELMAN AL-MOUSAWI

Bachelor of Science
University of Baghdad
Baghdad, Iraq
1975

Master of Science
Mississippi State University
Mississippi State, Mississippi
1980

Submitted to the Faculty of the Graduate College
of the Oklahoma State University
in partial fulfillment of the requirements
for the Degree of
DOCTOR OF PHILOSOPHY
July, 1984
FIXED-SIZE CONFIDENCE REGIONS FOR THE MEAN
VECTOR OF A MULTINORMAL DISTRIBUTION

Thesis Approved:

[Signatures]

Thesis Adviser
Ronald W. McNemar
William H. Stewart
Lavel M. Verhagen

Dean of the Graduate College
1.1 Introduction

Let $X_1, X_2, \ldots, X_n, \ldots$ be a sequence of independent and identically distributed (i.i.d.) random variables with each $X$ being distributed as $N_p(\mu, \sigma^2I)$, where $\mu \in \mathbb{R}^p$ is the unknown mean vector, $\sigma \in (0, \infty)$ is an unknown scale parameter, and $I$ is a known $p \times p$ positive definite matrix.

The problem of constructing a fixed-size confidence region for $\mu$ is formulated as follows. Given two preassigned numbers $d \in (0, \infty)$ and $\alpha \in (0, 1)$ and having recorded $n (\geq 2)$ samples $X_1, X_2, \ldots, X_n$, we propose the following ellipsoidal confidence region for $\mu$:

$$R = \{\omega \in \mathbb{R}^p: (\overline{X}_n - \omega)'H^{-1}(\overline{X}_n - \omega) \leq d^2\},$$

where $\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$. Let us use the notations $P(\cdot)$ and $E(\cdot)$ instead of $P_{\mu, \sigma}(\cdot)$ and $E_{\mu, \sigma}(\cdot)$, respectively, from this point onward. Now, the confidence coefficient associated with the region $R_n$ is given by

$$P(\mu \in R_n) = P((\overline{X}_n - \mu)'H^{-1}(\overline{X}_n - \mu) \leq d^2)$$

$$= P((n/\sigma^2)(\overline{X}_n - \mu)'H^{-1}(\overline{X}_n - \mu) \leq nd^2/\sigma^2)$$

$$= P(nd^2/\sigma^2),$$

$$\ldots (1.2)$$
where $F(u) = P(U \leq u)$ with $U$ being distributed as $\chi^2$ with $p$ degrees of freedom. The region $R_n$ is constructed in a way such that the length of its maximum diameter is at the most $2d$. This fact is referred to as the region $R_n$ being an ellipsoid of "fixed-size".

We also require that the confidence coefficient be at least $1 - \alpha$, and thus the sample size $n$ should be at least $a\sigma^2/d^2 = C$, say, where $F(a) = 1 - \alpha$. This number "$a$" can easily be found from the chi-square tables. Our "$C$" is referred to as the optimal fixed sample size required had $\sigma^2$ been known. However, $C$ is unknown since $\sigma^2$ is unknown, and thus no fixed-sample-size approach is feasible for our use.

For the sake of completeness, we now state definitions of some properties for any particular procedure giving rise to the stopping time, say, $N$.

**Definitions:**

(a) A procedure is called **consistent** in the Chow-Robbins (1965) sense if

$$P\{\mu \in R_N(d)\} \geq 1 - \alpha,$$  \hspace{1cm} (1.3)

for all $\mu \in \mathbb{R}^p$ and $\sigma \in (0, \infty)$. The property (1.3) is also referred to as **exact consistency** in Mukhopadhyay (1982).

(b) A procedure is called **asymptotically consistent** in the Chow-Robbins (1965) sense if

$$\lim_{d \to 0} P\{\mu \in R_N(d)\} = 1 - \alpha,$$  \hspace{1cm} (1.4)

for all $\mu \in \mathbb{R}^p$ and $\sigma \in (0, \infty)$.

(c) A procedure is called **asymptotically efficient** in the Chow-Robbins (1965) sense if
\[
\lim_{d \to 0} E[N(d)/C] = 1, \quad \ldots (1.5)
\]

for all \( \mu \in \mathbb{R}^p \) and \( \sigma \in (0, \infty) \). The equation (1.5) is now referred to as \textit{asymptotically first-order efficiency} property in Ghosh and Mukhopadhyay (1981).

(d) A procedure is called \textit{asymptotically second-order efficient} in the Ghosh-Mukhopadhyay (1981) sense if

\[
\lim_{d \to 0} E[N(d) - C] = k, \quad \ldots (1.6)
\]

for all \( \mu \in \mathbb{R}^p \) and \( \sigma \in (0, \infty) \), where \( k \) is a bounded constant.

From this point onward, we will write \( N \) instead of \( N(d) \).

1.2 Review of Literature

We begin this literature review with the univariate normal theory of fixed-width interval estimation of the mean. The literature dealing with the corresponding multivariate normal theory for the mean vector is then considered next. Finally, we mention some of the work done on point estimation problems for the mean or the mean vector.

Stein (1945, 1949) developed a two-stage procedure for constructing a fixed-width confidence interval for the mean \( \mu \) of a univariate normal distribution when the variance \( \sigma^2 \) is unknown. This procedure satisfies the properties (1.3) and (1.4), but it does not satisfy the property (1.5). See, for example, Chow and Robbins (1965) and Simons (1968).

Ray (1957) developed a purely sequential procedure to estimate the mean of a normal population with a confidence interval of preassigned width and confidence coefficient when the variance \( \sigma^2 \) is unknown. However, only the small sample approach was really discussed. More elabor-
ate and thorough treatments came from Chow and Robbins (1965). This purely sequential procedure is known to satisfy the properties (1.4) and (1.5). The basic reason behind going through a sequential scheme was to achieve property (1.5). In achieving that goal, however, the sequential procedure lost the exact property of (1.3).

Recently, Mukhopadhyay (1980) proposed a two-stage procedure (this is now called the "modified two-stage procedure") for constructing a fixed-width confidence interval for the mean \( \mu \) of a normal distribution when the variance \( \sigma^2 \) is unknown. This procedure has all the properties of (1.3), (1.4), and (1.5).

A natural question then arises. If the asymptotic efficiency property (1.5) can also be achieved by suitably modifying Stein's (1945, 1949) two-stage procedure, then exactly in what sense is the purely sequential procedure superior? Ghosh and Mukhopadhyay (1981) settled this issue by introducing a concept known as the second-order efficiency property. The sequential procedure satisfies property (1.6), whereas the modified two-stage procedure satisfies only the weaker property, namely (1.5).

Mukhopadhyay (1982) also showed that a fixed-width confidence interval for the mean of a univariate population can be constructed in a fairly reasonable way so as to achieve exact consistency even without the normality assumption. In Stein's construction, normality assumption is not crucial, and this was replaced by independence of some estimators of a pivotal nature. Modified two-stage procedures were also proposed along the lines of Mukhopadhyay (1980), and they were shown to be asymptotically first-order efficient.

Woodroofe (1977) obtained the second-order approximations of the
expected sample size and the risk associated with sequential procedures of the Ray-Chow-Robbins type. Woodroofe (1977) considered both point and interval estimation of the mean of a normal distribution when the variance is unknown.

Hall (1981) studied a three-stage procedure for constructing a fixed width confidence interval for the mean $\mu$ of a univariate normal distribution when $\sigma^2$ is unknown. If a third stage was appended to Stein's two-stage procedure, it lost its exactness (property (1.3)) but it became strongly competitive with the Ray-Chow-Robbins procedure from the efficiency point of view (properties (1.4) and (1.5)). Hall (1981) considered the asymptotic theory of triple sampling as it pertained to the estimation of the mean of a univariate normal distribution. He obtained various limit theorems and expansions, and his results showed in turn that a suitable triple sampling procedure actually combines the simplicity of Stein's double sampling techniques with that of the Ray-Chow-Robbins sequential procedure.

In the area of multivariate sequential estimation, Chatterjee (1959, 1960) extended the works of Stein (1945, 1949) for developing suitable two-stage procedures in the multivariate normal case with unknown mean vector $\mu$ and completely unknown positive definite dispersion matrix $\Sigma$. It was demonstrated how that procedure could be used to obtain a fixed-size ellipsoidal confidence region for $\mu$.

Srivastava (1967) extended Chow and Robbins' (1965) sequential procedure to construct ellipsoidal or spherical confidence regions with pre-assigned confidence coefficients for (i) the linear regression parameters and (ii) the mean vector of a multivariate population. No assumptions regarding the population distribution were made; and as a result, all