CHAPTER :-II

GENERAL THEORY OF β-EXPECTATION TOLERANCE INTERVAL

2.1. INTRODUCTION

As explained in the earlier chapter, a statistical tolerance interval is an interval determined from observed values of a random sample for the purpose of drawing an inference about the proportion of the distribution contained within that interval. Usually tolerance intervals are designated to capture at least a given proportion of some distribution. Therefore a straight forward way to get an interval is to find the estimates of parameters based on observed values and an interval I so that it covers, say 90% of the observations from this estimated distribution. Thus if the estimates of the parameters are close to the parameters of the distribution then the probability that the interval I covers the observations from that distribution is close to 90% . Such interval is called as an estimative interval by Aitchson(1975).

The main result of this chapter is to obtain an improvement on this interval. Atwood(1984) gave a general theory for constructing β-expectation tolerance interval when underlying parameters are to be estimated. To approximate the expected coverage he used second order Taylor series expansion. Also the large sample properties of the maximum likelihood estimators are used to get a correction term of order 1/n, which gives expected coverage closer to β.

Section-2.2 contains the terminology and formulation of the problem. Section-2.3 presents assumptions on random variables used in later sections. In Section-2.4 we state and prove the result due to Atwood(1984) regarding the derivation of an approximation to expected coverage of the proposed tolerance interval. Two applications of this theorem are given in later chapters.

2.2. TERMINOLOGY

Let $X_1, X_2, \ldots, X_n$ be independent random variables with distributions depending on θ, where θ and X_i's may be multidimensional, X_i's neednot be identically distributed. However, regularity conditions hold so that the maximum likelihood estimator is asymptotically normal with mean θ and non-singular variance-covariance matrix Σ.

Let Y be a one dimensional random variable whose distribution also depends on θ. An interval $I = I(X_1, \ldots, X_n)$ is to be obtained such that

$$P\,[Y \in I(X_1, X_2, \ldots, X_n)] \approx \beta, \qquad (2.2.1)$$

where β is some desired probability. If we rewrite the above probability statement in an equivalent form as

$$E\left[P(Y \in I(X_1, X_2, \ldots, X_n) \mid X_1, X_2, \ldots, X_n)\right] = \beta, \qquad (2.2.2)$$

where the expected value is taken with respect to the joint distribution function of $X_1, X_2, \ldots, X_n$. Then the interval I is called a tolerance interval with expected content β.

We use the following notation for computation of expected coverage of β-expectation tolerance interval.

Let $F(Y; \theta) = P[Y \le y \mid \theta]$, be the c.d.f. of Y. We define β^{th} percentile of the distribution as

$$F[a_\beta(\theta); \theta] = \beta.$$

That is $a_\beta(\theta)$ is the β^{th} percentile of the distribution. Let

$F_{10} = \partial F(y; \theta)/\partial y$

$F_{20} = \partial^2 F(y; \theta)/\partial y^2$

F_{01} = The vector having $\partial F(y; \theta)/\partial\theta_i$ as i^{th} element.

F_{11} = The vector having $\partial^2 F(y; \theta)/\partial\theta_i\partial y$ as i^{th} element.

F_{02} = The matrix having $\partial^2 F(y; \theta)/\partial\theta_i\partial\theta_j$ as ij^{th} element.

Σ = Non-singular variance-covariance matrix of $\hat{\theta}$, where $\hat{\theta}$ is the maximum likelihood estimator of θ.

Note that the derivatives of F are all evaluated at $y = a_\beta(\theta)$

and a superscript T will be used to denote the transpose and tr denote the trace of a matrix.

2.3. ASSUMPTIONS

2.3.1. Assumptions on $X_1, X_2, \ldots, X_n$

i) The random variables $X_1, X_2, \ldots, X_n$ are independent. Suppose that the maximum likelihood estimator is asymptotically normal with mean θ and non-singular variance-covariance matrix Σ. Conditions which ensure asymptotic normality are given in Lehmann(1983).

ii) Enough regularity conditions must be assumed to give the expansion of $E_\theta(\hat{\theta} - \theta)$ to order 1/n.

iii) The exact mean squared error matrix of $\hat{\theta}$ satisfies

$$E_\theta[(\hat{\theta} - \theta)(\hat{\theta} - \theta)^T] = \Sigma + O(1/n),$$

where $\quad E_\theta[\,O_p(1/n)] = O(1/n), \qquad$ for all θ.

2.3.2. Assumptions on Y

i) The random variable Y is independent of $X_1, X_2, \ldots, X_n$.

ii) The range R of Y is a possiblly infinite interval that does not depend on θ.

iii) The c.d.f. F_θ has continuous derivatives with respect to y and all components of θ of order three, for all

$Y \in R$ and $\theta \in \Theta$.

iv) $\partial F(y;\theta)/\partial y$ is strictly positive for $y \in R$ and $\theta \in \Theta$.

2.4. PROCEDURE TO OBTAIN EXPECTED COVERAGE:

Let β be some probability of interest such as 0.05 or 0.95. Let F_θ be the c.d.f. of Y, for arbitrary θ and β we define $a_\beta(\theta)$ as $F_\theta^{-1}(\beta)$ so that

$$F[\,a_\beta(\theta);\ \theta] = \beta \qquad (2.4.1)$$

The estimative interval would use $a_\beta(\hat{\theta})$ as one end point of the interval I. If $a_\beta(\hat{\theta})$ is the upper tolerance limit, then the coverage probability of the interval I can be found from

$$P_\theta[\,Y \le a_\beta(\hat{\theta})\,] = E_\theta\left\{ P_\theta[\,Y \le a_\beta(\hat{\theta}) \mid \hat{\theta}\,]\right\}$$

$$= E_\theta\left\{ F[\,a_\beta(\hat{\theta});\ \theta\,]\right\}$$

Therefore, if $\hat{\theta}$ is approximately equal to θ, Equation-(2.4.1) is expected to imply that

$$E_\theta\left\{ F[a_\beta(\hat{\theta});\ \theta] \right\} \approx \beta \qquad (2.4.2)$$

To improve on this approximation Atwood(1984) has obtained an expression for expected coverage of the form

$$E_\theta\left\{ F[a_\beta(\hat{\theta});\ \theta] \right\} = \beta + C_n(\theta) + O_p(1/n), \qquad (2.4.3)$$

where $C_n(\theta)$ is called the correction term, by using second order Taylor series expansion. Also he has suggested that, to get the corresponding improvement on the estimative interval, iterate on β until β plus the correction equals the desired probability, such as 0.95. For this β, $a_\beta(\hat{\theta})$ is a one end point of a tolerance interval with approximately the desired expected content.

In the following, We state and prove the result due to Atwood(1984) regarding the approximation to expected coverage of the proposed tolerance interval.

THEOREM(2.4.1) Under the assumptions on $X_1, X_2, \ldots, X_n$ and Y given in Section-2.3 we have

$$E_\theta\left\{ F\left[a_\beta(\hat{\theta});\ \theta\right] \right\} = \beta - E\,[\hat{\theta} - \theta\,]^T F_{01} - 0.5\ \mathrm{tr}(\,F_{02}\Sigma\,) + F_{01}^T \Sigma\, F_{11} / F_{10} + O\,(1/n). \quad (2.4.4)$$

PROOF: By Taylor's Theorem

$$F\left[a_\beta(\hat{\theta});\ \theta\right] = F[a_\beta(\theta);\ \theta\,] + [a_\beta(\hat{\theta}) - a_\beta(\theta)]\ F^1[a_\beta(\theta);\ \theta] + 0.5\ [a_\beta(\hat{\theta}) - a_\beta(\theta)]^2\ F^{11}[a_\beta(\theta);\ \theta] + \ldots$$

$$= \beta + \delta\, F_{10} + 0.5\ \delta^2 F_{20} + O\,(\delta^2), \quad (2.4.5)$$

where $\delta = [a_\beta(\hat{\theta}) - a_\beta(\theta)]$

$$F_{10} = F^{1}[a_\beta(\theta);\ \theta] = \partial F(y;\ \theta)/\partial y$$

and $$F_{20} = F^{11}[a_\beta(\theta);\ \theta] = \partial^2 F(y;\ \theta)/\partial y^2.$$

Note that here and below, F_{10}, F_{20}, F_{01}, F_{02} and F_{11} are always evaluated at $a_\beta(\theta)$. To evaluate other terms, an expression for δ is required.

Applying Taylor's theorem for $a_\beta(\hat{\theta})$, we get

$$a_\beta(\hat{\theta}) = a_\beta(\theta) + [\hat{\theta} - \theta]^T a^{1}_\beta(\theta)$$
$$+ 0.5\ [\hat{\theta} - \theta]^T a^{11}_\beta(\theta)[\hat{\theta} - \theta] + O_p[|\hat{\theta} - \theta|^2].$$

This implies that

$$\delta = [\hat{\theta} - \theta]^T a^{1}_\beta(\theta)$$
$$+ 0.5\ [\hat{\theta} - \theta]^T a^{11}_\beta(\theta)\ [\hat{\theta} - \theta] + O_p[|\hat{\theta} - \theta|^2]. \quad (2.4.6)$$

Here $a^{1}_\beta(\theta)$ is the vector with i^{th} element $\partial[a_\beta(\theta)]/\partial\theta_i$ and $a^{11}_\beta(\theta)$ is the matrix with ij^{th} element $\partial^2[a_\beta(\theta)]/\partial\theta_i\partial\theta_j$.

To evaluate $a^{1}_\beta(\theta)$ and $a^{11}_\beta(\theta)$ we differentiate the identity

$$F[a_\beta(\theta);\ \theta] = \beta$$

with respect to θ; so that we have

$$[\partial F(y;\ \theta)/\partial y\ ]\partial y/\partial\theta + [\partial F(y;\ \theta)/\partial\theta]\partial\theta/\partial\theta = 0. \quad (2.4.7)$$

When $y = a_\beta(\theta)$, we have

$$F_{10}\ a^{1}_{\beta}(\theta) + F_{01} = 0.$$

That is $$a^{1}_{\beta}(\theta) = - F_{01}/ F_{10} \quad . \qquad (2.4.8)$$

Now differentiating Equation-(2.4.7) with respect to θ we get

$$[\{ \partial^2 F(y;\theta)/\partial y^2 \} \partial y/\partial\theta + \{ \partial^2 F(y;\theta)/\partial y\partial\theta \} \partial\theta/\partial\theta] \partial y/\partial\theta$$
$$+ \{ \partial^2 y / \partial\theta\ \partial\theta^T \} \partial F(y;\theta)/\partial y + \{ \partial^2 F(y;\theta)/\partial y\partial\theta \} \partial y/\partial\theta$$
$$+ \{ \partial^2 F(y;\theta)/\partial\theta\ \partial\theta^T \} \partial\theta/\partial\theta = 0.$$

When $y = a_{\beta}(\theta)$, we get

$$[F_{20} a^{1}_{\beta}(\theta) + F_{11}]\ a^{1}_{\beta}(\theta) + a^{11}_{\beta}(\theta)\ F_{10} + F_{11} a^{1}_{\beta}(\theta) + F_{02} = 0$$

So That

$$F_{10} a^{11}_{\beta}(\theta) = - F_{02} - F_{11} [a^{1}_{\beta}(\theta)]^T$$
$$- F_{20} [a^{1}_{\beta}(\theta)]\ [a^{1}_{\beta}(\theta)]^T - F_{11} [a^{1}_{\beta}(\theta)]^T \qquad (2.4.9)$$

Substituting the expression for $a^{1}_{\beta}(\theta)$ and after some simplification we get

$$a^{11}_{\beta}(\theta) = F^{-1}_{10} [- F_{02} + (F_{11} F^T_{01} + F_{01} F^T_{11})/ F_{10}$$
$$- (F_{20}/F^2_{10})\ F_{01}\ F^T_{01}]. \qquad (2.4.10)$$

Substitute expresssion for $a^{1}_{\beta}(\theta)$ and $a^{11}_{\beta}(\theta)$ in Equation-(2.4.6), which gives

$$\delta = [\hat{\theta} - \theta]^T (- F_{01}/F_{10})$$

$$+ 0.5 \; [\hat{\theta} - \theta]^T \Big[F_{10}^{-1} \{ - F_{02} + (F_{11}F_{01}^T + F_{01} F_{11}^T)/F_{10}$$

$$- (F_{20}/F_{10}^2) \; F_{01}F_{01}^T \} \Big] \; [\hat{\theta} - \theta] \; + O_p\big[|\hat{\theta} - \theta|^2\big].$$

This implies

$$\delta^2 = [\hat{\theta} - \theta]^T (F_{01}F_{01}^T / F_{10}^2) \; [\hat{\theta} - \theta] + O_p\big[\; |\hat{\theta} - \theta|^2\big].$$

Substitute the expression for δ and δ^2 in Equation-(2.4.5) after simplification, this yields

$$F\big[a_{\beta}(\hat{\theta}); \; \theta\big] = \beta - [\hat{\theta} - \theta]^T F_{01} - 0.5 \; [\hat{\theta} - \theta]^T F_{02} [\hat{\theta} - \theta]$$

$$+ 0.5 \; [\hat{\theta} - \theta]^T \{ (F_{11}F_{01}^T + F_{01}F_{11}^T)/F_{10} \} [\hat{\theta} - \theta]$$

$$+ O_p\big[\; |\hat{\theta} - \theta|^2\big]. \qquad (2.4.11)$$

But $\quad [\hat{\theta} - \theta]^T F_{02} [\hat{\theta} - \theta] = tr\{[\hat{\theta} - \theta]^T F_{02} [\hat{\theta} - \theta]\}$

$$= tr\{F_{02}[\hat{\theta} - \theta][\hat{\theta} - \theta]^T\}$$

and $\qquad F_{11}F_{01}^T = F_{01}F_{11}^T$

Therefore Equation-(2.4.11) reduces to

$$F\big[a_{\beta}(\hat{\theta}); \; \theta\big] = \beta - [\hat{\theta} - \theta]^T F_{01} - 0.5 \; tr\{F_{02}[\hat{\theta} - \theta][\hat{\theta} - \theta]^T\},$$

$$+ (1/F_{10}) \; F_{01}^T [\hat{\theta} - \theta][\hat{\theta} - \theta]^T F_{11} + O_p\big[| \; \hat{\theta} - \theta|^2\big].$$

$$(2.4.12)$$

Taking expectation on both the sides of Equation-(2.4.12) and using the assumption on $X_1, X_2, \ldots, X_n$ we have

$$E_\theta\left\{F[\, a_\beta(\hat{\theta});\ \theta]\right\} = \beta - E_\theta[\hat{\theta} - \theta]^T F_{01} - 0.5\ \mathrm{tr}(F_{02}\Sigma)$$

$$+ F_{01}^T \Sigma\, F_{11} / F_{10} + O_p(1/n)$$

Hence the theorem. □

This theorem suggests a way to estimate $E_\theta\left\{F[a_\beta(\hat{\theta});\ \theta]\right\}$. If the expected values in the expressions for Σ and $E_\theta[\hat{\theta} - \theta]$ can be written as functions of θ, then we can estimate them by substituting $\hat{\theta}$ for θ. Otherwise replace each expected sum by corresponding sum of observed values to obtain estimates of Σ and $E_\theta[\hat{\theta} - \theta]$. Note that both the estimators $n^{-1}\Sigma$ and $nE_\theta[\hat{\theta} - \theta]$ are consistent. But, since $\hat{\theta}$ is a sufficient statistic, we use the first estimators because they only depend on the sufficient statistic.

Similarly estimate F_{10}, F_{01}, F_{11} and F_{02} by evaluating them at $\theta = \hat{\theta}$ and $y = a_\beta(\hat{\theta})$. These estimators are all consistent. Use them to estimate the quantity

$$\beta - E\,[\hat{\theta} - \theta]^T F_{01} - 0.5\ \mathrm{tr}\{\, F_{02}\Sigma\,\} + F_{01}^T \Sigma\, F_{11} / F_{10}. \tag{2.4.13}$$

The quantity (2.4.13) is of the form (β + correction). Note that each term in the correction is of order $1/n$. The whole expression (2.4.13) equals

$$P_\theta[Y \le a_\beta(\hat{\theta})] = E_\theta\left\{ F[\, a_\beta(\hat{\theta});\ \theta] \right\},$$

to within $O(1/n)$. The estimate of (2.4.13) equals $P_\theta[Y \le a_\beta(\hat{\theta})]$ to within $O_p(1/n)$. Iterate on β until the estimate of (2.4.13) equals a desired probability, such as 0.95. Then use $a_\beta(\hat{\theta})$ as one end point of the interval I.

---0---