CHAPTER 9

ESTIMATION OF THE PARAMETERS OF DOUBLY TRUNCATED NORMAL DISTRIBUTION FROM FIRST FOUR SAMPLE MOMENTS

9.1 Introduction

A.C. Cohen [11] has estimated the parameters of singly and doubly truncated normal distributions by the method of maximum likelihood. Using the method of moments, he [12] has also obtained the estimators of singly truncated normal distribution from first three sample moments. In this chapter, the method of moments has been applied to estimate the parameters of doubly truncated normal distribution. The estimators are obtained from the first four sample moments. Further, comparison of these estimators has been made with the maximum likelihood estimators. The study reveals that the method of moments provides simple estimators of the parameters of doubly truncated normal distribution without much loss of efficiency, while the maximum likelihood estimators are complicated and laborious.

9.2 Doubly truncated normal distribution

The density function of doubly truncated normal distribution is given by

(9.2.1) $$f(x) = \phi\left(\frac{x-u}{\sigma}\right)/\sigma\left[\Phi(x_1) - \Phi(x_0)\right],$$

$$x_0 < x < x_1,$$

where $\phi(t) = (\sqrt{2\pi})^{-1}e^{-t^2/2}$ and $\Phi(t) = \int_{-\infty}^{t} \phi(x)dx$.

Let $(x_0 - \mu)/\sigma = k_0$ and $(x_1 - \mu)/\sigma = k_1$. Change the origin to x_0 and put $x - x_0 = y$. Then, the density function of y is

(9.2.2) $$g(y) = \phi\left(\frac{y}{\sigma} + k_0\right)/G\sigma, \qquad 0 < y < d$$

where $G = \Phi(k_1) - \Phi(k_0)$ and $d = x_1 - x_0$. The moments m_r of the distribution of y about its origin are given by

(9.2.3) $$m_1 = \sigma(z_0 - z_1) - \sigma k_0$$

(9.2.4) $$m_r = -\sigma d^{r-1}z_1 + (r-1)\sigma^2 m_{r-2} - k_0\sigma m_{r-1},$$

$$r = 2, 3, \ldots\ldots$$

where $z_0 = \phi(k_0)/G$ and $z_1 = \phi(k_1)/G$.

9.3 Estimators of the parameters

From (9.2.4), by taking $r = 2, 3$ and 4, we obtain

(9.3.1) $$m_2 = -\sigma d z_1 + \sigma^2 - \sigma k_0 m_1 ,$$

(9.3.2) $$m_3 = -\sigma d^2 z_1 + 2\sigma^2 m_1 - \sigma k_0 m_2 ,$$

(9.3.3) $$m_4 = -\sigma d^3 z_1 + 3\sigma^2 m_2 - \sigma k_0 m_3 .$$

It should be observed that equations (9.3.1), (9.3.2), (9.3.3) are the same as equations (19) obtained by Cohen([13], p. 259) for the case of doubly truncated normal population when $b_1 = b_2 = 0$. Considering equations (9.3.1), (9.3.2), (9.3.3) in three unknowns σz_1, σ^2, and $h = \sigma k_0$ and solving them, we see that the solutions for h and σ^2 are

(9.3.4) $$h = [(3m_2m_3-2m_1m_4)+d(m_4-3m_2^2)+d^2(2m_1m_2-m_3)]/P,$$

(9.3.5) $$\sigma^2 = [(m_3^2-m_2m_4)+d(m_1m_4-m_2m_3)+d^2(m_2^2-m_1m_3)]/P,$$

where $P = (2m_1m_3-3m_2^2)+d(3m_1m_2-m_3)+d^2(m_2-2m_1^2)$.

This suggests that we can obtain estimators h^* and σ^{2*} of h and σ^2 from (9.3.4) and (9.3.5) by substituting for m_r the corresponding sample moments $v_r = \Sigma(x - x_0)^r/N$. The estimator μ^* of μ is then given by $\mu^* = x_0 - h^*$. Thus, we have

$$(9.3.6) \qquad h^* = [(3v_2v_3-2v_1v_4)+d(v_4-3v_2^2)+d^2(2v_1v_2-v_3)]/P^*,$$

$$(9.3.7) \qquad \sigma^{2*} = [(v_3^2-v_2v_4)+d(v_1v_4-v_2v_3)+d^2(v_2^2-v_1v_3)]/P^*,$$

where P^* is the value of P obtained by substituting for m_r the corresponding sample moments v_r in it.

When the upper truncation point x_1 is infinity, i.e., $d = \infty$, then we obtain the estimators as

$$h^* = (2v_1v_2 - v_3)/(v_2 - 2v_1^2),$$

$$\sigma^{2*} = (v_2^2 - v_1v_3)/(v_2 - 2v_1^2),$$

which are the same as obtained by Cohen [12] for singly truncated normal distribution.

In practical situation, rather computing h^* and σ^{2*} from (9.3.6) and (9.3.7), it might be easier to solve some pair of linear equations obtained from (9.3.1), (9.3.2) and (9.3.3).

9.4 Comparison with the maximum likelihood estimators

The maximum likelihood estimators $\hat{h}$ and $\hat{\sigma}$ of the parameters h and σ of doubly truncated normal distribution were obtained by Cohen [11]. In the present notation, they are given as

(9.4.1) $$v_1 = \hat{\sigma}(z_0 - z_1) - \hat{h} \quad ,$$

(9.4.2) $$v_2 = -\hat{\sigma} d z_1 + \hat{\sigma}^2 - \hat{h} v_1 \ .$$

It should be noted that z_0 and z_1 depend on $\hat{h}$ and $\hat{\sigma}$ and hence the estimators $\hat{h}$ and $\hat{\sigma}$ are to be found by numerical methods. As can be seen from (9.2.3) and (9.3.1), these estimators are the usual moment estimators using only first two moments. The asymptotic variances of $\hat{h}$ and $\hat{\sigma}$ are given below.

(9.4.3) $$(N/\sigma^{2})\mathrm{Var}(\hat{h}) = F_3/(F_1F_3 - F_2^2),$$

(9.4.4) $$(N/\sigma^{2})\mathrm{Var}(\hat{\sigma}) = F_1/(F_1F_3 - F_2^2),$$

where

$$F_1 = \left[-(z_1-z_0)^2-k_1z_1+k_0z_0+1\right],$$

$$F_2 = \left[(k_1^2z_1-k_0^2z_0)+(z_1-z_0)(k_1z_1-k_0z_0+1)\right],$$

$$F_3 = \left[-3(k_1z_1-k_0z_0-1)-(k_1z_1-k_0z_0-1)^2-(k_1^3z_1-k_0^3z_0)\right]$$

and N = size of the truncated sample.

The asymptotic variances of h^* and σ^* have been obtained by the δ-method (Kendall and Stuart [23], §10.6). Let

$$A_j = \frac{\partial h^*}{\partial v_j}\bigg|_{(v_1,v_2,v_3,v_4)=(m_1,m_2,m_3,m_4)}, \quad (j=1,2,3,4),$$

$$B_j = \frac{\partial \sigma^*}{\partial v_j}\bigg|_{(v_1,v_2,v_3,v_4)=(m_1,m_2,m_3,m_4)}, \quad (j=1,2,3,4).$$

Then, the asymptotic variances of h^* and σ^* are given by

(9.4.5) $$\text{Var}(h^*) = \sum_{i,\ j=1}^{4}\sum A_iA_j\text{Cov}(v_i,\ v_j),$$

(9.4.6) $$\text{Var}(\sigma^*) = \sum_{i,\ j=1}^{4}\sum B_iB_j\text{Cov}(v_i,\ v_j),$$

where it is to be understood that $\mathrm{Cov}(v_i, v_i)=\mathrm{Var}(v_i)$. Now if we write

$$m_j = \sigma^{-j} w_j, \qquad (j = 1, 2, 3, \ldots, 8),$$

then we get

$$w_1 = z_0 - z_1 - k_0,$$

$$w_2 = -z_1(k_1-k_0)+1-k_0 w_1,$$

$$w_j = -z_1(k_1-k_0)^{j-1}+(j-1)w_{j-2}-k_0 w_{j-1}, \qquad (j=3,4,\ldots,8).$$

From the definitions of A_i and B_j we obtain

$$A_1=\left[-2(w_4+k_0w_3)-3k_0w_2(k_1-k_0)+2(k_1-k_0)^2(2k_0w_1+w_2)\right]/\Delta \ ,$$

$$A_2=\left[3(w_3+2k_0w_2)-3(k_1-k_0)(2w_2+k_0w_1)+(k_1-k_0)^2(2w_1-k_0)\right]/\sigma\Delta ,$$

$$A_3=\left[(3w_2-2k_0w_1)+k_0(k_1-k_0)-(k_1-k_0)^2\right]/\sigma^2\Delta \ ,$$

$$A_4=\left[-2w_1+(k_1-k_0)\right]/\sigma^3\Delta \ ,$$

$$B_1=\left[-2w_3+(k_1-k_0)(w_4-3w_2)+(k_1-k_0)^2(4w_1-w_3)\right]/2\Delta \ ,$$

$$B_2=\left[(6w_2-w_4)-(k_1-k_0)(w_3+3w_1)+(k_1-k_0)^2(2w_2-1)\right]/2\sigma\Delta ,$$

$$B_3=\left[2(w_3-w_1)+(k_1-k_0)(1-w_2)-w_1(k_1-k_0)^2\right]/2\,\sigma^2\Delta\ ,$$

$$B_4=\left[-w_2+w_1(k_1-k_0)\right]/2\,\sigma^3\Delta\ ,$$

where

$$\Delta=(2w_1w_3-3w_2^2)+(3w_1w_2-w_3)(k_1-k_0)+(w_2-2w_1^2)(k_1-k_0)^2.$$

If we put $A_j=a_j/\sigma^{j-1}$ and $B_j=b_j/\sigma^{j-1}$, then from (9.4.5) and (9.4.6), observing that $Var(v_r)=\sigma^{2r}(w_{2r}-w_r^2)/N$, and $Cov(v_r,\ v_s)=\sigma^{r+s}(w_{r+s}-w_r\cdot w_s)/N$, we obtain

(9.4.7) $$(N/\sigma^2)Var(h^*)=\sum_{i,\ j=1}^{4}\sum(w_{i+j}-w_iw_j)a_ia_j,$$

(9.4.8) $$(N/\sigma^2)Var(\sigma^*)=\sum_{i,\ j=1}^{4}\sum(w_{i+j}-w_iw_j)b_ib_j.$$

The asymptotic efficiencies of h^* and σ^* are then obtained by dividing expressions in (9.4.3) and (9.4.4) respectively by (9.4.7) and (9.4.8). It should be noted that the asymptotic efficiencies of the estimators depend on k_0 and k_1, the lower and the upper points of truncation in standard units.

9.5 Example

To illustrate the foregoing results, a truncated sample of size 35 from a normal population with $\mu = 0$ and $\sigma = 1$ was selected from Mahalanobis' tables [26] by discarding observations less than $x_0 = -1$ and greater than $x_1 = 2$. The first four sample moments about $x_0 = -1$ were found to be $v_1 = 1.26508$, $v_2 = 2.18548$, $v_3 = 4.46599$, $v_4 = 10.000867$. Substituting these values in equations (9.3.6) and (9.3.7), we obtain $h^* = -0.834$ and $\sigma^{2*} = 1.591$. These give us $\mu^* = -0.166$ and $\sigma^* = 1.261$. The maximum likelihood estimates of μ and σ were obtained by numerical methods from equations (9.4.1) and (9.4.2). They are $\hat{\mu} = -0.052$ and $\hat{\sigma} = 1.192$.

Applying formulas (9.4.3), (9.4.4), (9.4.7) and (9.4.8), we find that

$$(N/\sigma^2)\mathrm{Var}(h^*)=3.49372, \qquad (N/\sigma^2)\mathrm{Var}(\hat{h})=3.42971$$

$$(N/\sigma^2)\mathrm{Var}(\sigma^*)=3.25404, \qquad (N/\sigma^2)\mathrm{Var}(\hat{\sigma})=3.14663.$$

Hence, the estimated asymptotic efficiencies of h^* and σ^* are respectively 98 % and 97 %.