University of Southampton

Faculty of Mathematical Studies

Mathematics

Estimation of the Parameters in the Truncated Normal Distribution when the Truncation Point is known.

by

Mehdi Tazhibi

Thesis submitted for the degree of Doctor of Philosophy

July, 1996

To Zahra

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF MATHEMATICAL STUDIES

Doctor of Philosophy

Estimation of Parameters in Truncated Normal Distribution when the Truncation Point is Known

by Mehdi Tazhibi.

In various fields of science, such as biology, economics and medicine, scientific data frequently follow a truncated normal distribution. Measurement of variables in some parts of the population present difficulties. Because of the importance of this distribution, many statisticians have been involved with the estimation of the relevant parameters.

The problem with the estimation of the parameters is that the method of maximum likelihood gives rise to two equations which cannot be explicitly solved and, further, the results obtained are not acceptable due to the biases are large. Cox & Hinkley (1974) have presented an approximation formula based on a Taylor expansion, which can be used to find the expected value and variance of the maximum likelihood estimators. An alternative approach for estimating the parameters is by application of Shenton & Bowman's formula (1977).

In this thesis the method of Shenton & Bowman is extended to the two-parameter case to give the means, variances and covariances of the maximum likelihood estimators of the truncated normal distribution simultaneously.

The maximum product spacing method, which is asymptotically as efficient as the maximum likelihood and in some cases hyper-efficient, is used for the truncated normal distribution.

Finally, a comparison is made between the above methods and also with the method of estimation by means of simulation.

Contents

ACKNOWLEDGEMENTS

بسم الله الرحمن الرحيم

In the name of God, Most Gracious, Most Merciful.

ولا تقف ما ليس لك به علم ان السمع والبصر والفواد اوليك كان عنه مسئولا .

سورة الاسراء آيه ٣٦

And pursue not that Of which thou hast No knowledge; for Every act of hearing Or of seeing Or of (feeling in) the heart Will be enquired into (On the Day of Reckoning). The Holy Quran; Chapter 17 (AL. Isrā), Verse 36

I am grateful to my supervisor Mr. B.J.R. Bailey for his invaluable advice, guidance and support throughout this work.

I wish to thank Dr. S.M. Lewis for her encouragement and help.

I am also grateful to Dr. David Adams, Ray Brown and Peter Hubbard for their assistance in computing.

Many thanks also go to my friends Abdellatif, Rahim, Parviz, James, Masoud, Akbar, Dominique, Rory, Denise, and Pedro.

I cannot forget my wife Zahra, and my children; Yaser, Hajar, Hamed and Hadi for their patience, tolerance and encouragement.

My particular thanks are due to many, in particular my father, my mother, my father and mother-in-law and my brothers; Ali, Mojtaba, Mortaza, Abbas, Mohamad and my sister Mrs Waltham for their help and encouragement.

Finally, I would like to thank my sponsor, Ministry of Health and Medical Education, and my employer Isfahan University of Medical Sciences for enabling me to undertake this project.

Chapter 0

Notation and preliminary results:

0.1 Introduction and notation

In this chapter we give the preliminary assumptions, notations and intermediate results needed in later chapters. These results concern the properties of the truncated normal distribution.

0.1.1 Assumptions:

1. Let $X \sim N(\mu_1, \sigma_1^2)$.
2. Let $Y \sim N(\mu_2, \sigma_2^2)$.
3. Let $Z \sim N(\mu_3, \sigma_3^2)$.
4. Let $(X, Y) \sim N_2(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$. Then the conditional distribution of Y given $X = x$ is

$$(Y \mid X = x) \sim N\left(\mu_2 + \rho(\frac{\sigma_2}{\sigma_1})(x - \mu_1), \sigma_2^2(1 - \rho^2)\right).$$

5. Let
$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \sim N_3\left(\begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3 \\ \rho_{12}\sigma_1\sigma_2 & \sigma_2^2 & \rho_{23}\sigma_2\sigma_3 \\ \rho_{13}\sigma_1\sigma_3 & \rho_{23}\sigma_2\sigma_3 & \sigma_3^2 \end{bmatrix}\right).$$

6. Let $D = Y - X$ be the increment, then
$$(D \mid X = x) \sim N\left(\mu_2 - \mu_1 + \left(\rho(\frac{\sigma_2}{\sigma_1}) - 1\right)(x - \mu_1), \sigma_2^2(1 - \rho^2)\right).$$

0.1.2 Notation:

Throughout this thesis we shall adopt the following notations.

1. Let $\mu_x = E(X \mid a < X < b)$.
2. Let $\sigma_x^2 = \text{Var}(X \mid a < X < b)$.
3. Let $\mu_y = E(Y \mid a < X < b)$.
4. Let $\sigma_y^2 = \text{Var}(Y \mid a < X < b)$.
5. Let $\mu_c = E(X \mid -\infty < X < c)$.
6. Let $\sigma_c^2 = \text{Var}(X \mid -\infty < X < c)$.
7. Let $\phi(x')$ and $\Phi(x')$ denote the probability density function (p.d.f.) and the cumulative distribution function (c.d.f.) of the standard normal distribution.
8. Let $a' = \frac{a-\mu_1}{\sigma_1}$, $b' = \frac{b-\mu_1}{\sigma_1}$, $x' = \frac{x-\mu_1}{\sigma_1}$ and $c' = \frac{c-\mu_1}{\sigma_1}$.
9. Let $\delta_\Phi = \Phi(b') - \Phi(a')$.
10. Let $\delta_\phi = \phi(b') - \phi(a')$.
11. Let $\delta_{\phi'} = \phi'(b') - \phi'(a') = a'\phi(a') - b'\phi(b')$, since $\phi'(a') = -a'\phi(a')$ etc.
12. Let $\delta_{\phi''} = \phi''(b') - \phi''(a') = a'^2\phi(a') - b'^2\phi(b') - \phi(b') + \phi(a')$.
13. Let $\delta_{\phi'''} = \phi'''(b') - \phi'''(a') = a'^3\phi(a') - b'^3\phi(b') - 3[a'^2\phi(a') - b'^2\phi(b') - \phi(b') + \phi(a')]$.

14. Let $\tau(c') = \frac{\phi(c')}{\Phi(c')}$. Then

$$\tau'(c') = -c'\tau(c') - \tau^2(c').$$

$$\tau''(c') = -\tau(c') - c'\tau'(c') - 2\tau(c')\tau'(c').$$

$$\tau'''(c') = -2\tau'(c') - c'\tau''(c') - 2\tau'^2(c') - 2\tau(c')\tau''(c').$$

$$\tau^{(iv)}(c') = -3\tau''(c') - c'\tau'''(c') - 6\tau'(c')\tau''(c') - 2\tau(c')\tau'''(c').$$

15. Let $\psi(c') = \frac{\sigma\phi(c')}{\Phi(c')}$. Then

$$\psi'(c') = -c'\psi(c') - \psi^2(c')/\sigma.$$

$$\begin{aligned}\psi''(c') &= -\psi(c') - c'\psi'(c') - 2\psi'(c')\psi(c')/\sigma \\ &= \psi(c')\left((c'^2 - 1) + 3c'\psi(c')/\sigma + 2(\psi(c')/\sigma)^2\right).\end{aligned}$$

$$\psi'''(c') = -2\psi'(c') - c'\psi''(c') - 2\psi'^2(c')/\sigma - 2\psi''(c')\psi(c')/\sigma.$$

$$\psi^{(iv)}(c') = -3\psi''(c') - c'\psi'''(c') - 6\psi'(c')\psi''(c')/\sigma - 2\psi(c')\psi'''(c')/\sigma.$$

In trivariate normal distribution

16. Let $D_1 = Y - X$ and $D_2 = Z - Y$ denote two increments.

17. Let $\theta_1 = -\sigma_1 + \rho_{12}\sigma_2$, $\theta_2 = -\sigma_2\rho_{12} + \sigma_3\rho_{13}$ and $\Delta = \left(\frac{\delta_{\phi'}}{\delta_\Phi} - (\frac{\delta_\phi}{\delta_\Phi})^2\right)$.

0.1.3 Cumulative distribution function of $(X \mid a < X < b)$:

We want to find $F(x \mid a < X < b)$.

Now

$$
\begin{aligned}
F(x \mid a < X < b) &= \frac{P(X \le x, a < X < b)}{P(a < X < b)} \\
&= \begin{cases} 0 & \text{if } x \le a \\ \frac{P(a<X<x)}{\Phi(b')-\Phi(a')} & \text{if } a < x < b \\ 1 & \text{if } x \ge b \end{cases} \\
&= \frac{\int_{-\infty}^{x} f_X(t)dt - \int_{-\infty}^{a} f_X(t)dt}{\Phi(b') - \Phi(a')} \\
&= \frac{\Phi(x') - \Phi(a')}{\delta_\Phi}, \qquad a < x < b. \qquad (0.1)
\end{aligned}
$$

0.1.4 Cumulative distribution function of $(X \mid -\infty < X < c)$:

By putting $a = -\infty$ and $b = c$ in equation (0.1) we can find the cumulative distribution function of X truncated from the right at c.

$$F(x \mid -\infty < X < c) = \frac{\Phi(x')}{\Phi(c')}, \qquad -\infty < x < c.$$

0.1.5 Probability density function of $(X \mid a < X < b)$:

We want to find $f(x \mid a < X < b)$. If we take the first derivative of $F(x \mid a < X < b)$, we can find $f(x \mid a < X < b)$.

$$
\begin{aligned}
f(\, x \mid a < X < b \,) &= d\,(F(x \mid a < X < b))\,/dx \\
&= d\left(\frac{\Phi(x') - \Phi(a')}{\delta_\Phi}\right)/dx \\
&= \frac{\phi(x')}{\sigma_1 \delta_\Phi}
\end{aligned}
$$

$$= \begin{cases} \frac{e^{\frac{-(x-\mu_1)^2}{2\sigma_1^2}}}{\sigma_1\delta_\Phi\sqrt{2\pi}} & \text{if } a < x < b \\ 0 & \text{otherwise} \end{cases}$$

$$= \frac{f_X(x)}{\delta_\Phi} \tag{0.2}$$

where $f_X(x)$ is the conditional marginal p.d.f. of X.

0.1.6 Probability density function of $(X \mid -\infty < X < c)$:

By putting $a = -\infty$ and $b = c$ in equation (0.2) we can find the probability density function of X truncated from the right at c.

$$f(x, \mid -\infty < X < c) = \frac{\phi(\frac{x-\mu_1}{\sigma_1})}{\sigma_1\Phi(\frac{c-\mu_1}{\sigma_1})}.$$

0.1.7 Expected value of $(X \mid a < X < b)$:

Now we want to find $E(X \mid a < X < b)$. We know that

$$\begin{aligned} \mu_x = E(X \mid a < X < b) &= \int_{-\infty}^{+\infty} x f(x \mid a < X < b) dx \\ &= \int_a^b x \frac{e^{\frac{-(x-\mu_1)^2}{2\sigma_1^2}}}{\sigma_1\delta_\Phi\sqrt{2\pi}} dx. \end{aligned}$$

So with the change of variable $x' = (x - \mu_1)/\ \sigma_1$, we have

$$\begin{aligned} E(\ X \mid a < X < b\ ) &= \frac{1}{\delta_\Phi\sqrt{2\pi}} \int_{a'}^{b'} (\sigma_1 x' + \mu_1) e^{-\frac{x'^2}{2}}\, dx' \\ &= \frac{\sigma_1}{\delta_\Phi\sqrt{2\pi}} \int_{a'}^{b'} x' e^{-\frac{x'^2}{2}}\, dx' + \frac{\mu_1}{\delta_\Phi} \int_{a'}^{b'} \frac{1}{\sqrt{2\pi}} e^{-\frac{x'^2}{2}}\, dx' \\ &= \frac{-\sigma_1}{\delta_\Phi\sqrt{2\pi}} \int_{a'}^{b'} d\left(e^{-\frac{x'^2}{2}}\right) + \frac{\mu_1(\Phi(b') - \Phi(a'))}{\delta_\Phi} \\ &= \frac{-\sigma_1}{\delta_\Phi\sqrt{2\pi}} \left(e^{-\frac{b'^2}{2}} - e^{-\frac{a'^2}{2}}\right) + \frac{\mu_1\delta_\Phi}{\delta_\Phi}. \end{aligned}$$

Therefore

$$\mu_x = \mu_1 - \frac{\sigma_1 \delta_\phi}{\delta_\Phi}. \tag{0.3}$$

0.1.8 Expected value of $(X \mid -\infty < X < c)$:

By putting $a = -\infty$ and $b = c$ in equation (0.3) we can find the expected value of X truncated from the right at c

$$\begin{aligned} \mu_c &= \mu_1 - \psi(c') \\ &= \mu_1 - \sigma_1 \tau(c'). \end{aligned}$$

0.1.9 Expected value of $(Y \mid a < X < b)$:

We know that

$$\begin{aligned} \mu_y = E(Y \mid a < X < b) &= \int_{-\infty}^{+\infty} y f(y \mid a < X < b) dy \\ &= \int_{-\infty}^{+\infty} y \int_a^b f(x, y \mid a < X < b) dx dy \\ &= \int_a^b \int_{-\infty}^{+\infty} y f(y \mid x) f_X(x \mid a < X < b) dx dy \\ &= \int_a^b \left[\mu_2 + \rho(\frac{\sigma_2}{\sigma_1})(x - \mu_1) \right] f_X(x \mid a < X < b) dx \\ &= \mu_2 + \rho(\frac{\sigma_2}{\sigma_1}) \left[\mu_1 - \frac{\sigma_1 \delta_\phi}{\delta_\Phi} - \mu_1 \right] \\ &= \mu_2 - \frac{\rho \sigma_2 \delta_\phi}{\delta_\Phi}. \end{aligned} \tag{0.4}$$

To check this result we now consider what happens to $E(Y \mid a < X < b)$ as $b \to a$, or $b' \to a'$ by use of l'Hopital's rule.

Since

$$\lim_{b' \to a'} \left(\frac{\delta_\phi}{\delta_\Phi} \right) = \lim_{b' \to a'} \left(\frac{\frac{\partial(\delta_\phi)}{\partial b'}}{\frac{\partial(\delta_\Phi)}{\partial b'}} \right)$$

$$
\begin{aligned}
&= \lim_{b' \to a'} \left(\frac{\frac{\partial(\phi(b') - \phi(a'))}{\partial b'}}{\frac{\partial(\Phi(b') - \Phi(a'))}{\partial b'}} \right) \\
&= \lim_{b \to a} -b' \\
&= -a'. \qquad (0.5)
\end{aligned}
$$

Substituting $\lim_{b' \to a'} \left(\frac{\delta_\phi}{\delta_\Phi}\right)$ from equation (0.5) into equation (0.3) we obtain

$$E(Y \mid X = a) = \mu_2 + \rho a' \sigma_2 \qquad (0.6)$$

which is the true regression equation.

0.1.10 Probability density function of $(Y \mid a < X < b)$:

Now we want to find $f(y \mid a < X < b)$.

$$
\begin{aligned}
f(y \mid a < X < b) &= \int_a^b f(x, y \mid a < X < b) dx \\
&= \frac{\int_a^b f(x, y) dx}{P(a < X < b)} \\
&= \frac{\int_a^b f(x \mid y) f_Y(y) dx}{P(a < X < b)} \\
&= f_Y(y) \frac{\int_a^b f(x \mid y) dx}{P(a < X < b)}. \qquad (0.7)
\end{aligned}
$$

Since we know that

$$(X \mid Y = y) \sim N\left(\mu_1 + \rho(\frac{\sigma_1}{\sigma_2})(y - \mu_2), \sigma_1^2(1 - \rho^2)\right),$$

then, by substituting the p.d.f. $(x \mid y)$ into equation (0.7), we obtain

$$
\begin{aligned}
&f(\ y \mid a < X < b\ ) = \qquad (0.8) \\
&\frac{f_Y(y)}{\delta_\Phi} \left(\Phi\left(\frac{b'}{\sqrt{1 - \rho^2}} - \frac{\rho(y - \mu_2)}{\sigma_1^2 \sqrt{1 - \rho^2}} \right) - \Phi\left(\frac{a'}{\sqrt{1 - \rho^2}} - \frac{\rho(y - \mu_2)}{\sigma_1^2 \sqrt{1 - \rho^2}} \right) \right).
\end{aligned}
$$

0.1.11 Variance of $(X \mid a < X < b)$:

In this subsection we find the variance of X when $a < X < b$. We know that

$$\text{Var}(X \mid a < X < b) = E(X^2 \mid a < X < b) - (E(X \mid a < X < b))^2 . \tag{0.9}$$

Now we have to find $E(X^2 \mid a < X < b)$

$$\begin{aligned} E(X^2 \mid a < X < b) &= \int_{-\infty}^{+\infty} x^2 f(x \mid a < X < b) dx \\ &= \frac{1}{\delta_\Phi \sqrt{2\pi}} \int_{a'}^{b'} (\sigma_1 x' + \mu_1)^2 e^{-\frac{x'^2}{2}} \, dx' \\ &= \frac{1}{\delta_\Phi \sqrt{2\pi}} \int_{a'}^{b'} (\sigma_1^2 x'^2 + \mu_1^2 + 2\mu_1\sigma_1 x') e^{-\frac{x'^2}{2}} \, dx' \\ &= \frac{\sigma_1^2}{\delta_\Phi \sqrt{2\pi}} \int_{a'}^{b'} x'^2 e^{-\frac{x'^2}{2}} \, dx' + \\ & \quad \frac{\mu_1^2}{\delta_\Phi \sqrt{2\pi}} \int_{a'}^{b'} e^{-\frac{x'^2}{2}} \, dx' + \\ & \quad \frac{2\mu_1\sigma_1}{\delta_\Phi \sqrt{2\pi}} \int_{a'}^{b'} x' e^{-\frac{x'^2}{2}} \, dx'. \end{aligned} \tag{0.10}$$

Let

$$J_2 = \frac{1}{\delta_\Phi \sqrt{2\pi}} \int_{a'}^{b'} x'^2 e^{-\frac{x'^2}{2}} \, dx', \tag{0.11}$$

$$J_1 = \frac{1}{\delta_\Phi \sqrt{2\pi}} \int_{a'}^{b'} x' e^{-\frac{x'^2}{2}} \, dx', \tag{0.12}$$

$$J_0 = \frac{1}{\delta_\Phi \sqrt{2\pi}} \int_{a'}^{b'} e^{-\frac{x'^2}{2}} \, dx'. \tag{0.13}$$

The next task is to find J_0 , J_1 and J_2.

Now

$$\begin{aligned} J_0 &= \frac{1}{\delta_\Phi \sqrt{2\pi}} \int_{-\infty}^{b'} e^{-\frac{x'^2}{2}} \, dx' - \frac{1}{\delta_\Phi \sqrt{2\pi}} \int_{-\infty}^{a'} e^{-\frac{x'^2}{2}} \, dx' \\ &= \left(\Phi(b') - \Phi(a')\right) / \delta_\Phi \\ &= 1. \end{aligned} \tag{0.14}$$

Further,

$$\begin{aligned} J_1 &= -\frac{1}{\delta_\Phi\sqrt{2\pi}}\int_{a'}^{b'} -x'e^{x'^2/2}\,dx' \\ &= -\frac{1}{\delta_\Phi\sqrt{2\pi}}\int_{a'}^{b'} d\left(e^{-x'^2/2}\right) \\ &= -\frac{1}{\delta_\Phi\sqrt{2\pi}}\left(e^{-b'^2/2} - e^{-a'^2/2}\right) \\ &= -\left(\phi(b') - \phi(a')\right)/\delta_\Phi \\ &= -\delta_\phi/\delta_\Phi. \end{aligned} \tag{0.15}$$

By the use of partial integration, we have

$$\begin{aligned} J_2 &= -\frac{1}{\delta_\Phi\sqrt{2\pi}}\int_{a'}^{b'} x'^2 e^{x'^2/2}\,dx' \\ &= -\frac{1}{\delta_\Phi\sqrt{2\pi}}\int_{a'}^{b'} d\left(-x'e^{-x'^2/2}\right) \\ &- \frac{1}{\delta_\Phi\sqrt{2\pi}}\int_{a'}^{b'} \left(-x'e^{-x'^2/2}\right)dx' \\ &= \left(a'\phi(a') - b'\phi(b') + \phi(b') - \phi(a')\right)/\delta_\Phi \\ &= \left(\delta_{\phi'} + \delta_\Phi\right)/\delta_\Phi. \end{aligned} \tag{0.16}$$

If we substitute J_2, J_1 and J_0 into equation (0.10) we find

$$\begin{aligned} E(X^2 \mid a < X < b) &= \frac{\sigma_1^2}{\delta_\Phi}\left(\delta_{\phi'} + \delta_\Phi\right) + \mu_1^2 + \frac{2\mu_1\sigma_1}{\delta_\Phi}(-\delta_\phi) \\ &= \sigma_1^2 + \mu_1^2 + \sigma_1^2\left(\frac{\delta_{\phi'}}{\delta_\Phi}\right) - 2\mu_1\sigma_1\left(\frac{\delta_\phi}{\delta_\Phi}\right). \end{aligned} \tag{0.17}$$

If we substitute equation (0.3) and equation (0.17) into equation (0.9), then we obtain

$$\begin{aligned} \sigma_x^2 &= \sigma_1^2 + \mu_1^2 + \sigma_1^2\left(\frac{\delta_{\phi'}}{\delta_\Phi}\right) - 2\mu_1\sigma_1\left(\frac{\delta_\phi}{\delta_\Phi}\right) - \left(\mu_1 - \frac{\sigma_1\delta_\phi}{\delta_\Phi}\right)^2 \\ &= \sigma_1^2\left(1 - \left(\frac{\delta_\phi}{\delta_\Phi}\right)^2 + \frac{\delta_{\phi'}}{\delta_\Phi}\right). \end{aligned} \tag{0.18}$$

To check this result, we now consider what happens to $\mathrm{Var}(X \mid a < X < b)$ as $b \to a$, or $b' \to a'$ by use of l'Hopital's rule.

Since

$$\begin{aligned}\lim_{b'\to a'}\left(\frac{\delta_{\phi'}}{\delta_{\Phi}}\right) &= \lim_{b'\to a'}\left(\frac{\frac{\partial(\delta_{\phi'})}{\partial b'}}{\frac{\partial(\delta_{\Phi})}{\partial b'}}\right)\\ &= \lim_{b'\to a'}\left(\frac{\frac{\partial(\phi'(b')-\phi'(a'))}{\partial b'}}{\frac{\partial(\Phi(b')-\Phi(a'))}{\partial b'}}\right)\\ &= \lim_{b'\to a'}(b'^2-1)\\ &= a'^2-1, \qquad (0.19)\end{aligned}$$

using $\lim_{b'\to a'}\left(\frac{\delta_{\phi}}{\delta_{\Phi}}\right)$ from equation (0.5), and $\lim_{b'\to a'}\left(\frac{\delta_{\phi'}}{\delta_{\Phi}}\right)$ from equation (0.19) it follows that

$$\mathrm{Var}(X \mid X = a) = 0 \qquad (0.20)$$

as we expected.

0.1.12 Variance of $(X \mid -\infty < X < c)$:

By putting $a = -\infty$ and $b = c$ in equation (0.18) we can find the second moment of X in truncation point c

$$\begin{aligned}\sigma_c^2 &= \sigma_1^2\left(1-\frac{c'\psi(c')}{\sigma_1}-\frac{\psi^2(c')}{\sigma_1^2}\right) = \sigma_1^2\left(1+\psi'(c')/\sigma\right)\\ &= \sigma_1^2(1-c'\tau(c')-\tau^2(c')) = \sigma_1^2(1+\tau'(c')). \qquad (0.21)\end{aligned}$$

0.1.13 Variance of $(Y \mid a < X < b)$:

To find $\mathrm{Var}(Y \mid a < X < b)$, we use

$$\mathrm{Var}(Y \mid a < X < b) = E(Y^2 \mid a < X < b) - (E(Y \mid a < X < b))^2. \qquad (0.22)$$

$E(Y^2 \mid a < X < b)$ is worked out as follows

$$E(Y^2 \mid a < X < b) = \int_{-\infty}^{+\infty} y^2 f(y \mid a < X < b)dy$$

$$
\begin{aligned}
&= \int_{-\infty}^{+\infty} y^2 \left(\int_a^b f(x, y \mid a < X < b) dx \right) dy \\
&= \int_a^b \int_{-\infty}^{+\infty} y^2 f(y \mid x) f_X(y, x \mid a < X < b) dx dy \\
&= \frac{\int_a^b \int_{-\infty}^{+\infty} y^2 f(y \mid x) f_X(x) dx dy}{\delta_\Phi}. \qquad (0.23)
\end{aligned}
$$

On considering

$$
\begin{aligned}
\sigma_2^2(1-\rho^2) = \mathrm{Var}(Y \mid x) &= E(Y^2 \mid x) - (E(Y \mid x))^2 \\
&= E(Y^2 \mid x) - \left(\mu_2 + \rho(\frac{\sigma_2}{\sigma_1})(x - \mu_1) \right)^2, \qquad (0.24)
\end{aligned}
$$

it follows that

$$
E(Y^2 \mid x) = \sigma_2^2(1-\rho^2) + \left(\mu_2 + \rho(\frac{\sigma_2}{\sigma_1})(x - \mu_1) \right)^2. \qquad (0.25)
$$

However,

$$
\begin{aligned}
E(Y^2 \mid x) &= \int_{-\infty}^{+\infty} y^2 f(y \mid x) dy \\
&= \frac{\int_{-\infty}^{+\infty} y^2 f(y, x) dy}{f_X(x)}. \qquad (0.26)
\end{aligned}
$$

Therefore we have

$$
\int_{-\infty}^{+\infty} y^2 f(y \mid x) dy = E(Y^2 \mid x) f_X(x). \qquad (0.27)
$$

However, by the substitution of $E(Y^2 \mid x)$ from equation (0.25) into equation (0.27), we have

$$
\int_{-\infty}^{+\infty} y^2 f(y \mid x) dy = \left(\sigma_2^2(1-\rho^2) + \left(\mu_2 + \rho(\frac{\sigma_2}{\sigma_1})(x - \mu_1) \right)^2 \right) f_X(x). \qquad (0.28)
$$

So, if we substitute equation (0.28) into equation (0.23), we obtain

$$
\begin{aligned}
E(Y^2 \mid a < X < b) &= \frac{\int_a^b \left(\sigma_2^2(1-\rho^2) + \left(\mu_2 + \rho(\frac{\sigma_2}{\sigma_1})(x - \mu_1) \right)^2 \right) f_X(x) dx}{\delta_\Phi} \\
&= \sigma_2^2(1-\rho^2) + \mu_2^2 + 2\left(\frac{\rho\sigma_2\mu_2}{\sigma_1\delta_\Phi} \right) \int_a^b (x - \mu_1) f_X(x) dx + \\
&\quad \left(\frac{\rho^2\sigma_2^2}{\sigma_1^2\delta_\Phi} \right) \int_a^b (x - \mu_1)^2 f_X(x) dx. \qquad (0.29)
\end{aligned}
$$

If we let $I_1 = \int_a^b (x-\mu_1) f_X(x)dx$, then

$$\begin{aligned} I_1 &= \frac{\int_a^b (x-\mu_1) e^{\frac{-(x-\mu_1)^2}{2\sigma_1^2}} dx}{\sigma_1\sqrt{2\pi}} \\ &= -\sigma_1(\phi(b') - \phi(a')) \\ &= -\sigma_1 \delta_\phi. \end{aligned} \tag{0.30}$$

Similarly, for $I_2 = \int_a^b (x-\mu_1)^2 f_X(x)dx$, we have

$$\begin{aligned} I_2 &= \int_a^b (x-\mu_1)^2 f_X(x)dx \\ &= \frac{(-\sigma_1)\int_a^b (x-\mu_1) d\left(e^{-(x-\mu_1)^2/2}\right)}{\sqrt{2\pi}} \\ &= \frac{(-\sigma_1)\int_a^b d\left((x-\mu_1)\left(e^{-(x-\mu_1)^2/2}\right)\right)}{\sqrt{2\pi}} + \\ &\quad \frac{(-\sigma_1)\int_a^b \left(e^{-(x-\mu_1)^2/2}\right)}{\sqrt{2\pi}} \\ &= \sigma_1^2\left(a'\phi(a') - b'\phi(b') + \Phi(b') - \Phi(a')\right) \\ &= \sigma_1^2(\delta_{\phi'} + \delta_\Phi). \end{aligned} \tag{0.31}$$

Similarly, using integration by parts we obtain I_3 and I_4, as follows:

$$\begin{aligned} I_3 &= \int_a^b (x-\mu_1)^3 f_X(x)dx \\ &= \sigma_1^3(\delta_{\phi''} + 3\delta_\phi), \end{aligned} \tag{0.32}$$

$$\begin{aligned} I_4 &= \int_a^b (x-\mu_1)^4 f_X(x)dx \\ &= \sigma_1^4(\delta_{\phi'''} + 6\delta_\phi + 3\delta_\Phi). \end{aligned} \tag{0.33}$$

Now if we substitute I_1 and I_2 from equations (0.30) and (0.31) into equation (0.29) we obtain

$$\begin{aligned} E(Y^2 \mid a < X < b) &= \sigma_2^2(1-\rho^2) + \mu_2^2 + 2\left(\frac{\rho\sigma_2\mu_2}{\sigma_1\delta_\Phi}\right)(-\delta_\phi) \\ &+ \left(\frac{\rho^2\sigma_2^2}{\sigma_1^2\delta_\Phi}\right)\left(\sigma_1^2(\delta_{\phi'} + \delta_\Phi)\right). \end{aligned} \tag{0.34}$$

Further substituting equations (0.34) and (0.4) into (0.22) then we find that

$$\sigma_y^2 = \sigma_2^2 \left(1 - \left(\frac{\rho\delta_\phi}{\delta_\Phi}\right)^2 + \frac{\rho^2\delta_{\phi'}}{\delta_\Phi}\right). \quad (0.35)$$

To check this result we now consider what happens to $\mathrm{Var}(Y \mid a < X < b)$ as $b \to a$, or $b' \to a'$ by use of l'Hopital's rule. using $\lim_{b' \to a'} \left(\frac{\delta_\phi}{\delta_\Phi}\right)$ from equation (0.5), and $\lim_{b' \to a'} \left(\frac{\delta_{\phi'}}{\delta_\Phi}\right)$ from equation (0.19) it follows that

$$\begin{aligned} \mathrm{Var}(Y \mid X = a) &= \sigma_1^2(1 - \rho^2(-a')^2 + \rho^2(a'^2 - 1)) \\ &= \sigma_1^2(1 - \rho^2) \end{aligned} \quad (0.36)$$

as we expected.

0.1.14 Third moment of $(Y \mid a < X < b)$ about the mean:

$$\begin{aligned} E[(Y - \mu_y)^3 \mid a < X < b] &= \int_{-\infty}^{+\infty} (y - \mu_y)^3 f(y \mid a < X < b) dy \\ &= \int_{-\infty}^{+\infty} (y - \mu_y)^3 \left(\int_a^b f(x, y \mid a < X < b) dx\right) dy \\ &= \frac{\int_a^b (\int_{-\infty}^{+\infty} (y - \mu_y)^3 f(y \mid x) dy) f_X(x) dx}{\delta_\Phi} \\ &= \frac{\int_a^b f_X(x) [\int_{-\infty}^{+\infty} (y - \mu_y)^3 f(y \mid x) dy] dx}{\delta_\Phi} \\ &= \frac{\int_a^b f_X(x) \{\int_{-\infty}^{+\infty} [y - \mu_2 + \rho\sigma_2(\frac{x-\mu_1}{\sigma_1} + \frac{\delta_\phi}{\delta_\Phi})]^3 f(y \mid x) dy\} dx}{\delta_\Phi} \\ &= \frac{1}{\delta_\Phi} \int_a^b \Big\{ f_X(x) \Big(E[(y - \mu_2)^3 \mid X = x] + \\ &\quad 3\rho\sigma_2 (\frac{x - \mu_1}{\sigma_1} + \frac{\delta_\phi}{\delta_\Phi}) E[(Y - \mu_2)^2 \mid X = x] + \\ &\quad 3\rho^2\sigma_2^2 (\frac{x - \mu_1}{\sigma_1} + \frac{\delta_\phi}{\delta_\Phi})^2 E[(Y - \mu_2) \mid X = x] + \\ &\quad \rho^3\sigma_2^3 (\frac{x - \mu_1}{\sigma_1} + \frac{\delta_\phi}{\delta_\Phi})^3 \Big) \Big\} dx. \end{aligned} \quad (0.37)$$

Since $(Y \mid X = x) \sim N[\mu_2, \sigma_2^2(1-\rho^2)]$, the first and third central moments of $Y \mid X = x$ are zero, and we will therefore have

$$\begin{aligned}
E[(Y-\mu_y)^3 \mid a < X < b] &= \frac{3\rho\sigma_2(1-\rho^2)}{\delta_\Phi}\int_a^b (\frac{x-\mu_1}{\sigma_1} + \frac{\delta_\phi}{\delta_\Phi}) f_X(x)dx \\
&+ \frac{\rho^3\sigma_2^2}{\delta_\Phi}\int_a^b (\frac{x-\mu_1}{\sigma_1} + \frac{\delta_\phi}{\delta_\Phi})^3 f_X(x)dx \\
&= \frac{\rho^3\sigma_2^2}{\delta_\Phi}\int_a^b (\frac{x-\mu_1}{\sigma_1})^3 f_X(x)dx \\
&+ \frac{3\rho^3\sigma_2^2\delta_\phi}{\delta_\Phi^2}\int_a^b (\frac{x-\mu_1}{\sigma_1})^2 f_X(x)dx \\
&+ \frac{3\rho\sigma_2}{\delta_\Phi}[1-\rho^2+\rho^2(\frac{\delta_\phi}{\delta_\Phi})^2]\int_a^b (\frac{x-\mu_1}{\sigma_1}) f_X(x)dx \\
&+ \frac{\rho\sigma_2^3}{\delta_\Phi^2}[3(1-\rho^2)+\rho^2(\frac{\delta_\phi}{\delta_\Phi})^2]\int_a^b f_X(x)dx. \qquad (0.38)
\end{aligned}$$

Therefore, using I_1, I_2 and I_3 from equations (0.30), (0.31), and (0.32), we obtain

$$E[(Y-\mu_y)^3 \mid a < X < b] = \rho^3\sigma_2^3\left[-2(\frac{\delta_\phi}{\delta_\Phi})^3 + 3(\frac{\delta_\phi}{\delta_\Phi})(\frac{\delta_{\phi'}}{\delta_\Phi}) - \frac{\delta_{\phi''}}{\delta_\Phi}\right]. \qquad (0.39)$$

Consider now what happens to $E[(Y-\mu_y)^3 \mid a < X < b]$ as $b \to a$, or as $b' \to a'$. By the use of l'Hopital's rule, we have

Since

$$\begin{aligned}
\lim_{b'\to a'}\left(\frac{\delta_{\phi''}}{\delta_\Phi}\right) &= \lim_{b'\to a'}\left(\frac{\frac{\partial(\delta_{\phi''})}{\partial b'}}{\frac{\partial(\delta_\Phi)}{\partial b'}}\right) \\
&= \lim_{b'\to a'}\left(\frac{\frac{\partial(\phi''(b')-\phi''(a'))}{\partial b'}}{\frac{\partial(\Phi(b')-\Phi(a'))}{\partial b'}}\right) \\
&= \lim_{b'\to a'}[b'(3-b'^2)] \\
&= a'(3-a'^2). \qquad (0.40)
\end{aligned}$$

Using $\lim_{b'\to a'}\left(\frac{\delta_\phi}{\delta_\Phi}\right)$ from equation (0.5), and $\lim_{b'\to a'}\left(\frac{\delta_{\phi'}}{\delta_\Phi}\right)$ from equation (0.19) and $\lim_{b'\to a'}\left(\frac{\delta_{\phi''}}{\delta_\Phi}\right)$ from equation (0.40) it follows that

$$E[(Y-\mu_y)^3 \mid X = a] = \rho^3\sigma_2^3[2a'^3 + 3(-a')(a'^2) - a'(3-a'^2)] = 0 \qquad (0.41)$$

as expected.

0.1.15 Third moment of $(X \mid a < X < b)$ about the mean:

Similarly, we can replace $\rho\sigma_2$ by σ_1 and obtain the third moment of $(X \mid a < X < b)$ as

$$E[(X-\mu_x)^3 \mid a < X < b] = \sigma_1^3\left[-2(\frac{\delta_\phi}{\delta_\Phi})^3 + 3(\frac{\delta_\phi}{\delta_\Phi})(\frac{\delta_{\phi'}}{\delta_\Phi}) - \frac{\delta_{\phi''}}{\delta_\Phi}\right]. \tag{0.42}$$

0.1.16 Third moment of $(X \mid -\infty < X < c)$ about the mean:

By putting $a = -\infty$ and $b = c$ in equation (0.42), we can find the third moment of X truncated from the right at c.

$$\begin{aligned} E[(X-\mu_c)^3 \mid -\infty < X < c] &= -\sigma_1^3\left[(c'^2-1)\frac{\psi(c')}{\sigma_1} + 3c'\frac{\psi^2(c')}{\sigma_1^2} + 2\frac{\psi^3(c')}{\sigma_1^3}\right]. \\ &= -\sigma_1^3[(c'^2-1)\tau(c') + 3c'\tau^2(c') + 2\tau^3(c')]. \end{aligned} \tag{0.43}$$

0.1.17 Fourth moment of $(Y \mid a < X < b)$ about the mean:

$$\begin{aligned} E[(Y-\mu_y)^4 \mid a < X < b] &= \int_{-\infty}^{+\infty}(y-\mu_y)^4 f(y \mid a < X < b)dy \\ &= \int_{-\infty}^{+\infty}(y-\mu_y)^4\left(\int_a^b f(x,y \mid a < X < b)dx\right)dy \\ &= \frac{\int_a^b(\int_{-\infty}^{+\infty}(y-\mu_y)^4 f(y \mid x)dy)f_X(x)dx}{\delta_\Phi} \\ &= \frac{\int_a^b f_X(x)[\int_{-\infty}^{+\infty}(y-\mu_y)^4 f(y \mid x)dy]dx}{\delta_\Phi} \\ &= \frac{\int_a^b f_X(x)\{\int_{-\infty}^{+\infty}[y-\mu_2+\rho\sigma_2(\frac{x-\mu_1}{\sigma_1}+\frac{\delta_\phi}{\delta_\Phi})]^4 f(y \mid x)dy\}dx}{\delta_\Phi} \\ &= \frac{1}{\delta_\Phi}\int_a^b \Big\{f_X(x)\Big(E[(y-\mu_2)^4 \mid X = x] + \\ &\quad 4\rho\sigma_2(\frac{x-\mu_1}{\sigma_1}+\frac{\delta_\phi}{\delta_\Phi})E[(Y-\mu_2)^3 \mid X = x] + \\ &\quad 6\rho^2\sigma_2^2(\frac{x-\mu_1}{\sigma_1}+\frac{\delta_\phi}{\delta_\Phi})^2 E[(Y-\mu_2)^2 \mid X = x] + \\ &\quad 4\rho^3\sigma_2^3(\frac{x-\mu_1}{\sigma_1}+\frac{\delta_\phi}{\delta_\Phi})^3 E[(Y-\mu_2) \mid X = x] + \end{aligned}$$

$$\rho^4\sigma_2^4(\frac{x-\mu_1}{\sigma_1}+\frac{\delta_\phi}{\delta_\Phi})^4\Big)\Big\}\,dx \tag{0.44}$$

Since $(Y \mid X = x) \sim N[\mu_2, \sigma_2^2(1-\rho^2)]$, the first and third central moments of $Y \mid X = x$ about its mean are zero, and

$$E[(Y-\mu_2)^2 \mid X = x] = \sigma_2^2(1-\rho^2) \tag{0.45}$$

and

$$E[(Y-\mu_2)^4 \mid X = x] = 3[\sigma_2^2(1-\rho^2)]^2. \tag{0.46}$$

So we have

$$\begin{aligned} E[(Y-\mu_y)^4 \mid a < X < b] &= \frac{1}{\delta_\Phi}\int_a^b \left\{3\sigma_2^4(1-\rho^2)^2 + 6\rho^2\sigma_2^4(1-\rho^2)[\frac{x-\mu_1}{\sigma_1}+\frac{\delta_\phi}{\delta_\Phi}]^2\right\} f_X(x)dx \\ &+ \frac{1}{\delta_\Phi}\rho^4\sigma_2^4\int_a^b[\frac{x-\mu_1}{\sigma_1}+\frac{\delta_\phi}{\delta_\Phi}]^4 f_X(x)dx. \end{aligned} \tag{0.47}$$

Using I_1, I_2, I_3 and I_4 from equations (0.30), (0.31), (0.32) and (0.33) respectively, we obtain

$$\begin{aligned} E[(Y-\mu_y)^4 \mid a < X < b] &= 3\left\{\sigma_2^2\left[1-\rho^2(\frac{\delta_\phi}{\delta_\Phi})^2+\rho^2(\frac{\delta_{\phi'}}{\delta_\Phi})\right]\right\}^2 \\ &+ \rho^4\sigma_2^4\left[\frac{\delta_{\phi'''}}{\delta_\Phi} - 4(\frac{\delta_\phi}{\delta_\Phi})(\frac{\delta_{\phi''}}{\delta_\Phi}) + 12(\frac{\delta_\phi}{\delta_\Phi})^2(\frac{\delta_{\phi'}}{\delta_\Phi})\right. \\ &- \left. 6(\frac{\delta_\phi}{\delta_\Phi})^4 - 3(\frac{\delta_{\phi'}}{\delta_\Phi})^2\right]. \end{aligned} \tag{0.48}$$

Now consider what happens to $E[(Y-\mu_y)^4 \mid a < X < b]$ as $b \to a$, or as $b' \to a'$. Again, using l'Hopital's rule,

$$\begin{aligned} \lim_{b'\to a'}\left(\frac{\delta_{\phi'''}}{\delta_\Phi}\right) &= \lim_{b'\to a'}\left(\frac{\frac{\partial(\delta_{\phi'''})}{\partial b'}}{\frac{\partial(\delta_\Phi)}{\partial b'}}\right) \\ &= \lim_{b'\to a'}\left(\frac{\frac{\partial(\phi'''(b')-\phi'''(a'))}{\partial b'}}{\frac{\partial(\Phi(b')-\Phi(a'))}{\partial b'}}\right) \\ &= \lim_{b'\to a'} b'^4 - 6b'^2 + 3 \\ &= a'^4 - 6a'^2 + 3. \end{aligned} \tag{0.49}$$

Consequently, we obtain

$$E[(Y-\mu_y)^4 \mid X=a] = 3[\sigma_2^2(1-\rho^2)]^2 \qquad (0.50)$$

as expected.

0.1.18 Fourth moment of $(X \mid a < X < b)$ about the mean:

Similarly we can replace $\rho\sigma_2$ by σ_1 everywhere except in the first term of the first line of equation (0.48), where σ_2 is replaced by σ_1, to obtain

$$\begin{aligned} E[(X-\mu_x)^4 \mid a < X < b] &= 3\left\{\sigma_1^2\left[1-(\frac{\delta_\phi}{\delta_\Phi})^2+(\frac{\delta_{\phi'}}{\delta_\Phi})\right]\right\}^2 \\ &+ \sigma_1^4\left[\frac{\delta_{\phi'''}}{\delta_\Phi}-4(\frac{\delta_\phi}{\delta_\Phi})(\frac{\delta_{\phi''}}{\delta_\Phi})+12(\frac{\delta_\phi}{\delta_\Phi})^2(\frac{\delta_{\phi'}}{\delta_\Phi})\right. \\ &- \left. 6(\frac{\delta_\phi}{\delta_\Phi})^4-3(\frac{\delta_{\phi'}}{\delta_\Phi})^2\right]. \end{aligned} \qquad (0.51)$$

0.1.19 Fourth moment of $(X \mid -\infty < X < c)$ about the mean:

By putting $a = -\infty$ and $b = c$ in equation (0.51) we can find the fourth moment of X truncated from the right at c

$$\begin{aligned} E[(X-\mu_c)^4 \mid -\infty < X < c] &= 3\sigma_1^4\left(1-\frac{c'\psi(c')}{\sigma_1}-\frac{\psi^2(c')}{\sigma_1^2}\right)^2 \\ &+ \sigma_1^4\left[(3c'-c'^3)\frac{\psi(c')}{\sigma_1}+(4-7c'^2)\frac{\psi^2(c')}{\sigma_1^2}\right. \\ &- \left. 12c'\frac{\psi^3(c')}{\sigma_1^3}-6\frac{\psi^4(c')}{\sigma_1^4}\right] \\ &= 3\sigma_1^4(1-c'\tau(c')-\tau^2(c'))^2 \\ &+ \sigma_1^4[(3c'-c'^3)\tau(c')+(4-7c'^2)\tau^2(c') \\ &- 12c'\tau^3(c')-6\tau^4(c')]. \end{aligned} \qquad (0.52)$$

0.1.20 Moment generating function of $(Y \mid a < X < b)$:

In this section we find the moment generating function of $Y \mid a < X < b$.

$$\begin{aligned}
M_{Y|a<X<b}(t) &= \int_{-\infty}^{+\infty} e^{ty} f(y \mid a < X < b) dy \\
&= \int_{-\infty}^{+\infty} e^{ty} \left(\int_a^b f(y, x \mid a < X < b) dx \right) dy \\
&= \frac{\int_a^b (\int_{-\infty}^{+\infty} e^{ty} f(y \mid x) dy) f_X(x) dx}{\delta_\Phi} \\
&= \frac{\int_a^b f_X(x) [\int_{-\infty}^{+\infty} e^{ty} f(y \mid x) dy] dx}{\delta_\Phi} \\
&= \frac{\int_a^b f_X(x) \{ \int_{-\infty}^{+\infty} e^{[\mu_2 + \rho(\frac{\sigma_2}{\sigma_1})(x-\mu_1)]t + \frac{1}{2}[\sigma_2^2(1-\rho^2)]t^2} \} dx}{\delta_\Phi} \\
&= \frac{1}{\delta_\Phi} e^{\mu_2 t + \frac{1}{2}[\sigma_2^2(1-\rho^2)]t^2} \int_a^b e^{\rho \frac{\sigma_2}{\sigma_1}(x-\mu_1)t} f_X(x) dx \\
&= \frac{1}{\delta_\Phi} e^{\mu_2 t + \frac{1}{2}[\sigma_2^2(1-\rho^2)]t^2} \int_{a'}^{b'} e^{\rho\sigma_2 x' t} \frac{1}{\sqrt{2\pi}} e^{-\frac{x'^2}{2}} dx' \\
&= \frac{1}{\delta_\Phi} e^{\mu_2 t + \frac{1}{2}[\sigma_2^2(1-\rho^2)]t^2} \int_{a'}^{b'} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}[(x'-\rho\sigma_2 t)^2 - (\rho\sigma_2 t)^2]} dx' \\
&= \frac{1}{\delta_\Phi} e^{\mu_2 t + \frac{1}{2}[\sigma_2^2(1-\rho^2)]t^2 + \frac{\rho^2\sigma_2^2 t^2}{2}} \int_{a'}^{b'} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x'-\rho\sigma_2 t)^2} dx' \\
&= e^{\mu_2 t + \frac{\sigma_2^2 t^2}{2}} . \frac{\Phi(b' - \rho\sigma_2 t) - \Phi(a' - \rho\sigma_2 t)}{\Phi(b') - \Phi(a')} \qquad (0.53)
\end{aligned}$$

0.1.21 Moment generating function of $(X \mid a < X < b)$:

In equation (0.53) replacing σ_2 with σ_1 and μ_2 with μ_1 in the exponent of e and $\rho\sigma_2$ with σ_1 in $\Phi(.)$, we obtain

$$M_{X|a<X<b}(t) = e^{\mu_1 t + \frac{\sigma_1^2 t^2}{2}} . \frac{\Phi(b' - \sigma_1 t) - \Phi(a' - \sigma_1 t)}{\Phi(b') - \Phi(a')} \qquad (0.54)$$

0.1.22 Moment generating function of $(X \mid -\infty < X < c)$:

By putting $a = -\infty$ and $b = c$ in equation (0.54), we find the moment generating function of X truncated from the right at c to be

$$M_{X|-\infty<X<c}\ (t) = e^{\mu_1 t + \frac{\sigma_1^2 t^2}{2}} . \frac{\Phi(c' - \sigma_1 t)}{\Phi(c')}. \tag{0.55}$$

0.1.23 Cumulant generating function and first four cumulants of $(X \mid -\infty < X < c)$:

In this section we find the cumulant generating function of $(X \mid -\infty < X < c)$ and its first four cumulants. As $\kappa_{X|-\infty<X<c}\ (t) = \ln M_{X|-\infty<X<c}\ (t)$, the cumulant generating function of X truncated at c is, from (0.55),

$$\kappa_{X|-\infty<X<c}\ (t) = \mu_1 t + \frac{\sigma_1^2 t^2}{2} + \ln \Phi(c' - \sigma_1 t) - \ln \Phi(c').$$

The first cumulant of X is

$$\kappa_1(X) = \kappa'_{X|-\infty<X<c}\ (t) \mid_{t=0} = E(X) = \mu_1 - \sigma_1 \tau(c')$$

which is identical with the moment in section (0.1.7). The second cumulant of X is

$$\kappa_2(X) = \kappa''_{X|-\infty<X<c}\ (t) \mid_{t=0} = \mu_2(X) = \mathrm{Var}(X) = \sigma_1^2 \left[1 + \tau'(c')\right].$$

The third cumulant of X is

$$\kappa_3(X) = \kappa'''_{X|-\infty<X<c}\ (t) \mid_{t=0} = \mu_3(X) = -\sigma_1^3 \tau''(c') = -\sigma_1^2 \psi''(c').$$

The fourth cumulant of X is

$$\kappa_4(X) = \kappa^{(iv)}_{X|-\infty<X<c}\ (t) \mid_{t=0} = \mu_4(X) - 3\mu_2^2(X) = \sigma_1^4 \tau'''(c').$$

It follows that

$$\mu_4(X) = \sigma_1^4 \{\tau'''(c') + 3\left[1 + \tau'(c')\right]^2\}.$$

By using the formulae for $\tau'(c')$, $\tau''(c')$ and $\tau'''(c')$ from section (0.1.2) we can see that the second, third and fourth moments of X are identical with the moments in sections (0.1.12), (0.1.16) and (0.1.19).

0.1.24 Covariance of $(X, Y \mid a < X < b)$:

We know that

$$\begin{aligned}\mathrm{Cov}(X, Y \mid a < X < b) &= E(XY \mid a < X < b) - \\ &\quad E(X \mid a < X < b)E(Y \mid a < X < b).\end{aligned} \tag{0.56}$$

Since we know $E(X \mid a < X < b)$ and $E(Y \mid a < X < b)$, we only have to find $E(XY \mid a < X < b)$, and for this we use the following steps

$$\begin{aligned}E(XY \mid a < X < b) &= \int_{-\infty}^{+\infty}\int_a^b xyf(x,y)dydx \\ &= \int_{-\infty}^{+\infty} xy \int_a^b f(y, x \mid a < X < b)dxdy \\ &= \int_a^b \int_{-\infty}^{+\infty} xyf(y,x)dydx \\ &= \int_a^b \int_{-\infty}^{+\infty} xyf(y \mid x)f_X(x)dydx \\ &= \frac{\int_a^b xf_X(x)\left(\int_{-\infty}^{+\infty} yf(y \mid x)dy\right)dx}{\delta_\Phi} \\ &= \frac{\int_a^b xf_X(x)\left(\mu_2 + \rho(\frac{\sigma_2}{\sigma_1})(x-\mu_1)\right)dx}{\delta_\Phi}.\end{aligned} \tag{0.57}$$

Using equation (0.2), we can calculate $\frac{\int_a^b xf_X(x)\left(\mu_2 + \rho(\frac{\sigma_2}{\sigma_1})(x-\mu_1)\right)dx}{\delta_\Phi}$, as follows:

$$\begin{aligned}&\frac{\int_a^b xf_X(x)\left(\mu_2 + \rho(\frac{\sigma_2}{\sigma_1})(x-\mu_1)\right)dx}{\delta_\Phi} \\ &= \frac{\int_a^b (x - \mu_1 + \mu_1)f_X(x)\left(\mu_2 + \rho(\frac{\sigma_2}{\sigma_1})(x-\mu_1)\right)dx}{\delta_\Phi}\end{aligned}$$

$$= \frac{\mu_2 + \rho\mu_1(\frac{\sigma_2}{\sigma_1})}{\delta_\Phi} I_1 + \frac{\rho(\frac{\sigma_2}{\sigma_1})}{\delta_\phi} I_2 + \mu_1\mu_2$$
$$= \rho(\frac{\sigma_2}{\sigma_1})(-\sigma_1\delta_\phi) + \rho(\frac{\sigma_2}{\sigma_1})\sigma_1^2(\delta_{\phi'} + \delta_\Phi) + \mu_1\mu_2. \quad (0.58)$$

Using equation (0.58) in equation (0.57) we obtain

$$E(XY \mid a < X < b) = -\frac{\sigma_1\mu_2\delta_\phi}{\delta_\Phi} - \frac{\rho\sigma_1\mu_2\delta_\phi}{\delta_\Phi} + \frac{\rho\sigma_1\sigma_2\delta_{\phi'}}{\delta_\Phi} + \rho\sigma_1\sigma_2 + \mu_1\mu_2. \quad (0.59)$$

Now, substituting equations (0.59), (0.3), (0.4), into equation (0.56), and using assumption 5 of section (0.1.1), we have

$$\mathrm{Cov}(X, Y \mid a < X < b) = \rho_{12}\sigma_1\sigma_2 \left(1 - \left(\frac{\delta_\phi}{\delta_\Phi}\right)^2 + \frac{\delta_{\phi'}}{\delta_\Phi}\right). \quad (0.60)$$

Similarly, we obtain

$$\mathrm{Cov}(X, Z \mid a < X < b) = \rho_{13}\sigma_1\sigma_3 \left(1 - \left(\frac{\delta_\phi}{\delta_\Phi}\right)^2 + \frac{\delta_{\phi'}}{\delta_\Phi}\right). \quad (0.61)$$

0.1.25 Expected value of increment:

In this section we derive the mean of increment, D, which is defined by $Y - X$.

We know that

$$\begin{aligned} E(D \mid a < X < b) &= E(Y \mid a < X < b) - E(X \mid a < X < b) \\ &= \mu_2 - \frac{\rho\sigma_2\delta_\phi}{\delta_\Phi} - [\mu_1 - \frac{\rho\sigma_1\delta_\phi}{\delta_\Phi}] \\ &= \mu_2 - \mu_1 + \frac{(\sigma_1 - \rho\sigma_2)\delta_\phi}{\delta_\Phi}. \end{aligned} \quad (0.62)$$

To check this result we now consider what happens to $E(D \mid a < X < b)$ as $b \to a$, or $b' \to a'$ by use of l'Hopital's rule. Using $\lim_{b' \to a'} \left(\frac{\delta_\phi}{\delta_\Phi}\right)$ from equation (0.5) into equation (0.62)

$$\begin{aligned} E(D \mid X = a) &= \mu_2 - \mu_1 + \frac{(\sigma_1 - \rho\sigma_2)\delta_\phi}{\delta_\Phi} \\ &= \mu_2 - \mu_1 - a'(\sigma_1 - \rho\sigma_2) \end{aligned} \quad (0.63)$$

as we expected.

0.1.26 Variance of increment:

In this section we find Var($D \mid a < X < b$). Since we know that

$$\begin{aligned}\mathrm{Var}(D \mid a < X < b) &= \mathrm{Var}(Y \mid a < X < b) + \mathrm{Var}(X \mid a < X < b) \\ &- 2\mathrm{Cov}(X, Y \mid a < X < b). \qquad (0.64)\end{aligned}$$

Substituting $\mathrm{Var}(X \mid a < X < b)$, $\mathrm{Var}(Y \mid a < X < b)$ and $\mathrm{Cov}(X, Y \mid a < X < b)$ respectively from equations (0.18), (0.35), (0.60) into equation (0.64) we obtain

$$\mathrm{Var}(D \mid a < X < b) = (\sigma_1 - \rho\sigma_2)^2 \left(1 - (\frac{\delta_\phi}{\delta_\Phi})^2 + \frac{\delta_{\phi'}}{\delta_\Phi}\right) + \rho^2(1 - \sigma_1^2) + \sigma_2^2(1 - \rho^2). \qquad (0.65)$$

0.2 Conditional distribution of $[Y, Z]$ given $X = x$:

In this section we have made use of the following Lemma and Theorem.

Lemma 0.1 *If* $\mathbf{\Sigma}$ *is a matrix such that*

$$\mathbf{\Sigma} = \begin{bmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{bmatrix},$$

then

$$\mathbf{\Sigma}^{-1} = \begin{bmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{W}_{11} & \mathbf{W}_{12} \\ \mathbf{W}_{21} & \mathbf{W}_{22} \end{bmatrix}$$

where

$$\mathbf{\Sigma}_{11}^{-1} = \mathbf{W}_{11} - \mathbf{W}_{12}\mathbf{W}_{22}^{-1}\mathbf{W}_{12}' \qquad (0.66)$$

and

$$\mathbf{\Sigma}_{22}^{-1} = \mathbf{W}_{22} - \mathbf{W}_{12}'\mathbf{W}_{11}^{-1}\mathbf{W}_{12}. \qquad (0.67)$$

Theorem 0.1 *If* $\mathbf{X} \sim N_n(\mu, \Sigma)$, *and* $\mathbf{x}' = (\mathbf{x}_1', \mathbf{x}_2')$ *is a* $1 \times n$ *vector with* $\mathbf{x}_1' = [x_1, x_2, \ldots, x_k]$ *and* $\mathbf{x}_2' = [x_{k+1}, x_{k+2}, \ldots, x_n]$, *then the marginal distribution of* $\mathbf{x_2}$ *is*

$$g_{\mathbf{X_2}}(\mathbf{x_2}) = g(x_{k+1}, x_{k+2}, \ldots, x_n) = \frac{e^{-(\mathbf{x}-\mu)'\mathbf{\Sigma}_{22}^{-1}(\mathbf{x}-\mu)/2}}{(2\pi)^{(n-k)/2} \mid \mathbf{\Sigma}_{22} \mid^{1/2}}. \qquad (0.68)$$

If we want to find the conditional distribution of $\mathbf{x_1}$, given $\mathbf{x_2}$, we have

$$\begin{aligned} f(\mathbf{x_1} \mid \mathbf{x_2}) &= \frac{f_{\mathbf{X}}(\mathbf{x})}{g_{\mathbf{X}}(\mathbf{x_2})} \\ &= \frac{e^{-(\mathbf{x}-\mu)'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\mu)/2-(\mathbf{x}-\mu)'\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}-\mu)/2}}{(2\pi)^{k/2}\,[\frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{22}|}]^{1/2}} \end{aligned} \tag{0.69}$$

and

$$\mathbf{X_1} \mid \mathbf{X_2} \sim N_2(\mu_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x_2} - \mu_2), \mathbf{W}_{11}^{-1}), \tag{0.70}$$

where

$$\mathbf{W}_{11} = (\boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{12}')^{-1}$$

hence

$$\mathbf{W}_{11}^{-1} = \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{12}'.$$

Using the Lemma 0.1 and Theorem 0.1, for the partitioning of the covariance matrix of the following vector and also using the notation described in section (0.1.1), we have

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \sim N_3\left(\begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix}, \boldsymbol{\Sigma}\right) \tag{0.71}$$

where

$$\boldsymbol{\Sigma} = \begin{bmatrix} \mathrm{Var}(X) & \mathrm{Cov}(X,Y) & \mathrm{Cov}(X,Z) \\ \mathrm{Cov}(X,Y) & \mathrm{Var}(Y) & \mathrm{Cov}(Y,Z) \\ \mathrm{Cov}(X,Y) & \mathrm{Cov}(Y,Z) & \mathrm{Var}(Z) \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3 \\ \rho_{12}\sigma_1\sigma_2 & \sigma_2^2 & \rho_{23}\sigma_2\sigma_3 \\ \rho_{13}\sigma_1\sigma_3 & \rho_{23}\sigma_2\sigma_3 & \sigma_3^2 \end{bmatrix},$$

$$\boldsymbol{\Sigma}_{11} = \begin{bmatrix} \sigma_2^2 & \rho_{23}\sigma_2\sigma_3 \\ \rho_{23}\sigma_2\sigma_3 & \sigma_3^2 \end{bmatrix},$$

$$\boldsymbol{\Sigma}_{12} = \begin{bmatrix} \rho_{12}\sigma_1\sigma_2 \\ \rho_{13}\sigma_1\sigma_3 \end{bmatrix},$$

and

$$\boldsymbol{\Sigma}_{22} = [\sigma_1^2].$$

Then we find the conditional distribution of $\begin{bmatrix} Y \\ Z \end{bmatrix} \mid X = x$ is

$$\begin{bmatrix} Y \\ Z \end{bmatrix} |_{X=x} = N_2 \left(\begin{bmatrix} \mu_2 + \rho_{12}\sigma_2(x - \mu_1)/\sigma_1 \\ \mu_3 + \rho_{13}\sigma_3(x - \mu_1)/\sigma_1 \end{bmatrix}, \begin{bmatrix} \sigma_2^2(1 - \rho_{12}^2) & \sigma_2\sigma_3(\rho_{23} - \rho_{12}\rho_{13}) \\ \sigma_2\sigma_3(\rho_{23} - \rho_{12}\rho_{13}) & \sigma_2^2(1 - \rho_{12}^2) \end{bmatrix} \right). \tag{0.72}$$

0.2.1 Variance of increments in trivariate normal distribution:

As in the case of the bivariate normal distribution we are going to find the variance of each increment individually.

We know that

$$\begin{aligned} \mathrm{Var}(D_1) &= \mathrm{Var}(Y - X) \\ &= \mathrm{Var}(Y) + \mathrm{Var}(X) - 2\mathrm{Cov}(X, Y) \\ &= \sigma_2^2 + \sigma_1^2 - 2\rho_{12}\sigma_1\sigma_2. \end{aligned} \tag{0.73}$$

Similarly, we can find the variance of the second increment D_2.

$$\begin{aligned} \mathrm{Var}(D_2) &= \mathrm{Var}(Z - Y) \\ &= \mathrm{Var}(Z) + \mathrm{Var}(Y) - 2\mathrm{Cov}(Y, Z) \\ &= \sigma_3^2 + \sigma_2^2 - 2\rho_{23}\sigma_2\sigma_3. \end{aligned} \tag{0.74}$$

0.2.2 Covariance of increments in trivariate normal distribution:

In this section we find the covariance of D_1, D_2

$$\mathrm{Cov}(D_1, D_2) = \mathrm{Cov}(Y - X, Z - Y)$$

$$
\begin{aligned}
&= \operatorname{Cov}(Y,Z) - \operatorname{Var}(Y) - \operatorname{Cov}(X,Z) + \operatorname{Cov}(X,Y) \\
&= \rho_{23}\sigma_2\sigma_3 - \sigma_2^2 - \rho_{13}\sigma_1\sigma_3 + \rho_{12}\sigma_1\sigma_2. \qquad (0.75)
\end{aligned}
$$

0.2.3 Correlation between increments in trivariate normal distribution:

In this section we find the correlation coefficient of increments.

Using equations (0.73), (0.74) and (0.75) into the correlation coefficient of D_1, D_2 formula, will have

$$
\begin{aligned}
\rho_{(D_1,D_2)} &= \frac{\operatorname{Cov}(D_1,D_2)}{\sqrt{\operatorname{Var}(D_1)}\sqrt{\operatorname{Var}(D_2)}} \\
&= \frac{\rho_{23}\sigma_2\sigma_3 - \sigma_2^2 - \rho_{13}\sigma_1\sigma_3 + \rho_{12}\sigma_1\sigma_2}{\sqrt{\sigma_2^2 + \sigma_1^2 - 2\rho_{12}\sigma_1\sigma_2}\sqrt{\sigma_3^2 + \sigma_2^2 - 2\rho_{23}\sigma_2\sigma_3}}. \qquad (0.76)
\end{aligned}
$$

0.2.4 Correlation coefficient of D_1, D_2 given $X = x$:

In this section we will find the correlation between D_1 and D_2 denoted by $\rho_{(D_1,D_2|X=x)}$.

Let $\mathbf{H}$ denote the covariance matrix of $\begin{bmatrix} D_1 \\ D_2 \\ X \end{bmatrix}$. Then we can write

$$
\mathbf{H} = \begin{bmatrix} \operatorname{Var}(D_1) & \operatorname{Cov}(D_1,D_2) & \operatorname{Cov}(D_1,X) \\ \operatorname{Cov}(D_1,D_2) & \operatorname{Var}(D_2) & \operatorname{Cov}(D_2,X) \\ \operatorname{Cov}(D_1,X) & \operatorname{Cov}(D_1,X) & \operatorname{Var}(X) \end{bmatrix}.
$$

Subsequently

$$
\begin{aligned}
\operatorname{Cov}(D_1,X) &= E(D_1X) - E(D_1)E(X) \\
&= E[(Y-X)X] - E(Y-X)E(X) \\
&= E(YX) - E(X^2) - E(Y)E(X) + [E(X)]^2
\end{aligned}
$$

$$\begin{aligned} &= E(YX) - E(Y)E(X) - [E(X^2) - [E(X)]^2] \\ &= \text{Cov}(X,Y) - \text{Var}(X) \\ &= \rho_{12}\sigma_1\sigma_2 - \sigma_1^2 \\ &= \sigma_1(\rho_{12}\sigma_2 - \sigma_1) \\ &= \sigma_1\theta_1. \end{aligned} \tag{0.77}$$

Similarly,

$$\begin{aligned} \text{Cov}(D_2, X) &= E(D_2X) - E(D_2)E(X) \\ &= E[(Z-Y)X] - E(Z-Y)E(X) \\ &= E(ZX) - E(YX) - E(Z)E(X) + E(Y)E(X) \\ &= E(ZX) - E(Z)E(X) - [E(YX) - E(Y)E(X)] \\ &= \text{Cov}(Z,X) - \text{Cov}(Y,X) \\ &= \rho_{13}\sigma_1\sigma_3 - \rho_{12}\sigma_1\sigma_2 \\ &= \sigma_1(\rho_{13}\sigma_3 - \rho_{12}\sigma_2) \\ &= \sigma_1\theta_2. \end{aligned} \tag{0.78}$$

To find the distribution of $\begin{bmatrix} D_1 \\ D_2 \\ X \end{bmatrix}$ we partition $\mathbf{H}$ as

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_{11} & \mathbf{H}_{12} \\ \mathbf{H}_{12}' & \mathbf{H}_{22} \end{bmatrix},$$

where

$$\mathbf{H}_{11} = \begin{bmatrix} \text{Var}(D_1) & \text{Cov}(D_1, D_2) \\ \text{Cov}(D_1, D_2) & \text{Var}(D_2) \end{bmatrix},$$

$$\mathbf{H}_{12} = \begin{bmatrix} \sigma_1\theta_1 \\ \sigma_1\theta_2 \end{bmatrix},$$

$$\mathbf{H}'_{12} = [\sigma_1\theta_1 \qquad \sigma_1\theta_2],$$

and

$$\mathbf{H}_{22} = [\sigma_1^2].$$

By the use of Lemma 0.1 and Theorem 0.1 we can write

$$\begin{bmatrix} D_1 \\ D_2 \\ X \end{bmatrix} |_{X=x} \sim N_2 \left(\begin{bmatrix} \mu_2 - \mu_1 \\ \mu_3 - \mu_2 \end{bmatrix} + \mathbf{H}_{12}\mathbf{H}_{22}^{-1}(x - \mu_1), \mathbf{U}_{11}^{-1} \right)$$

where $\mathbf{U}_{11}^{-1}$ is the covariance matrix of $\begin{bmatrix} D_1 \\ D_2 \end{bmatrix} |_{X=x}$.

Now, using the partition, we have

$$\begin{aligned} \mathbf{U}_{11}^{-1} &= \mathbf{H}_{11} - \mathbf{H}_{12}\mathbf{H}_{22}^{-1}\mathbf{H}'_{12} \\ &= \begin{bmatrix} \mathrm{Var}(D_1) & \mathrm{Cov}(D_1, D_2) \\ \mathrm{Cov}(D_1, D_2) & \mathrm{Var}(D_2) \end{bmatrix} - \begin{bmatrix} \sigma_1\theta_1 \\ \sigma_1\theta_2 \end{bmatrix} [\sigma_1^2][\sigma_1\theta_1 \qquad \sigma_1\theta_2] \\ &= \begin{bmatrix} \mathrm{Var}(D_1) - \theta_1^2 & \mathrm{Cov}(D_1, D_2) - \theta_1\theta_2 \\ \mathrm{Cov}(D_1, D_2) - \theta_1\theta_2 & \mathrm{Var}(D_2) - \theta_2^2 \end{bmatrix}. \end{aligned} \tag{0.79}$$

In matrix (0.79) we have

$$\mathrm{Var}(D_1 \mid X = x) = \mathrm{Var}(D_1) - \theta_1^2, \tag{0.80}$$

$$\mathrm{Var}(D_2 \mid X = x) = \mathrm{Var}(D_2) - \theta_2^2 \tag{0.81}$$

and

$$\mathrm{Cov}(D_1, D_2 \mid X = x) = \mathrm{Cov}(D_1, D_2) - \theta_1\theta_2. \tag{0.82}$$

Therefore, we obtain

$$\begin{aligned} \rho(D_1, D_2 \mid X = x) &= \frac{\mathrm{Cov}(D_1, D_2 \mid X = x)}{\sqrt{\mathrm{Var}(D_1 \mid X = x)}\sqrt{\mathrm{Var}(D_2 \mid X = x)}} \\ &= \frac{\mathrm{Cov}(D_1, D_2) - \theta_1\theta_2}{\sqrt{\mathrm{Var}(D_1) - \theta_1^2}\sqrt{\mathrm{Var}(D_2) - \theta_2^2}}. \end{aligned} \tag{0.83}$$

0.2.5 Covariance of Z, Y given $a < X < b$:

In this section we set out $\mathrm{Cov}(Z, Y \mid a < X < b)$, for which we can use the following formula

$$\mathrm{Cov}(Z, Y \mid a < X < b) = E(ZY \mid a < X < b) - E(Z \mid a < X < b)E(Y \mid a < X < b). \quad (0.84)$$

Now we have to find $E(ZY \mid a < X < b)$. Using equation (0.72) and

$$\begin{aligned} E(ZY \mid a < X < b) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{a}^{b} zyf(z, y \mid a < X < b)dxdydz \\ &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} zy \int_{a}^{b} f(x, y, z)dxdydz/\delta_{\Phi} \\ &= \int_{a}^{b} f_X(x) \int_{-\infty}^{+\infty} yzf(y, z \mid x)dydzdx/\delta_{\Phi} \quad (0.85) \end{aligned}$$

we can write

$$\mathrm{Cov}(Z, Y \mid X = x) = E(ZY \mid X = x) - E(Z \mid X = x)E(Y \mid X = x),$$

that is,

$$\sigma_2\sigma_3(\rho_{23} - \rho_{12}\rho_{13}) = E(ZY \mid X = x) - [\mu_3 + \rho_{13}\frac{\sigma_3}{\sigma_1}(x - \mu_1)][\mu_2 + \rho_{12}\frac{\sigma_2}{\sigma_1}(x - \mu_1)]. \quad (0.86)$$

Therefore, we have

$$E(ZY \mid X = x) = \int_{-\infty}^{+\infty} yzf(y, z \mid x)dydz = \sigma_2\sigma_3(\rho_{23} - \rho_{12}\rho_{13}) + [\mu_3 + \rho_{13}\frac{\sigma_3}{\sigma_1}(x - \mu_1)][\mu_2 + \rho_{12}\frac{\sigma_2}{\sigma_1}(x - \mu_1)]. \quad (0.87)$$

Substituting $\int_{-\infty}^{+\infty} yzf(y, z \mid x)dydz$ from equation (0.87) into equation (0.85) gives

$$\begin{aligned} E(ZY \mid a < X < b) &= [\sigma_2\sigma_3(\rho_{23} - \rho_{12}\rho_{13}) + \mu_2\mu_3] \int_{a}^{b} f_X(x)dx/\delta_{\Phi} \\ &+ [\mu_3\rho_{12}\frac{\sigma_2}{\sigma_1} + \mu_2\rho_{13}\frac{\sigma_3}{\sigma_1}] \int_{a}^{b} (x - \mu_1)f_X(x)dx/\delta_{\Phi} \\ &+ [\rho_{12}\frac{\sigma_2}{\sigma_1}][\rho_{13}\frac{\sigma_3}{\sigma_1}] \int_{a}^{b} (x - \mu_1)^2 f_X(x)dx/\delta_{\Phi}. \quad (0.88) \end{aligned}$$

Substituting I_1 and I_2 from equations (0.30), (0.31), we obtain

$$\begin{aligned} E(ZY \mid a < X < b) &= [\sigma_2\sigma_3\rho_{23} + \mu_2\mu_3] \\ &- [\mu_3\rho_{12}\frac{\sigma_2}{\sigma_1} + \mu_2\rho_{13}\frac{\sigma_3}{\sigma_1}]\delta_\phi/\delta_\Phi \\ &+ \sigma_2\sigma_3\rho_{12}\rho_{13}\delta_{\phi'}/\delta_\Phi. \end{aligned} \tag{0.89}$$

Similar to equation (0.4) in trivariate normal distribution, we can write

$$E(Y \mid a < X < b) = \mu_2 - \frac{\rho_{12}\sigma_2\delta_\phi}{\delta_\Phi} \tag{0.90}$$

and

$$E(Z \mid a < X < b) = \mu_3 - \frac{\rho_{13}\sigma_3\delta_\phi}{\delta_\Phi}. \tag{0.91}$$

Substituting equations (0.90), (0.91) and (0.85) into equation (0.84) we have

$$\text{Cov}(Z, Y \mid a < X < b) = \sigma_2\sigma_3\left[\rho_{23} + \rho_{12}\rho_{13}\left(-(\frac{\delta_\phi}{\delta_\Phi})^2 + \frac{\delta_{\phi'}}{\delta_\Phi}\right)\right]. \tag{0.92}$$

0.2.6 Covariance of D_1, D_2 given $a < X < b$:

We know that

$$\begin{aligned} \text{Cov}(D_1, D_2 \mid a < X < b) &= \text{Cov}(Y - X, Z - Y \mid a < X < b) \\ &= \text{Cov}(Y, Z \mid a < X < b) - \text{Cov}(X, Z \mid a < X < b) \\ &- \text{Var}(Y \mid a < X < b) + \text{Cov}(X, Y \mid a < X < b) \\ &= \sigma_2\sigma_3\left[\rho_{23} + \rho_{12}\rho_{13}\left(-(\frac{\delta_\phi}{\delta_\Phi})^2 + \frac{\delta_{\phi'}}{\delta_\Phi}\right)\right] \\ &- \rho_{13}\sigma_1\sigma_3\left(1 - (\frac{\delta_\phi}{\delta_\Phi})^2 + \frac{\delta_{\phi'}}{\delta_\Phi}\right) \\ &- \sigma_2^2\left(1 - (\frac{\rho_{12}\delta_\phi}{\delta_\Phi})^2 + \frac{\rho_{12}^2\delta_{\phi'}}{\delta_\Phi}\right) \\ &+ \rho_{12}\sigma_1\sigma_2\left(1 - (\frac{\delta_\phi}{\delta_\Phi})^2 + \frac{\delta_{\phi'}}{\delta_\Phi}\right). \end{aligned} \tag{0.93}$$

Assuming $\rho_{13} = \rho_{12}\rho_{23}$, then we can write

$$\text{Cov}(D_1, D_2 \mid a < X < b) = (\sigma_2 - \sigma_1\rho_{12})(\sigma_3\rho_{23} - \sigma_2) - \theta_1\theta_2\Delta. \tag{0.94}$$

and

$$\text{Cov}(D_1, D_2 \mid a < X < b) = \text{Cov}(D_1, D_2) - \theta_1\theta_2\Delta. \tag{0.95}$$

0.2.7 Correlation coefficient of D_1, D_2 given $a < X < b$:

In this section we want to find $\rho(D_1, D_2 \mid a < X < b)$. To do this we use the $\text{Cov}(D_1, D_2 \mid a < X < b)$. We also have to find $\text{Var}(D_1 \mid a < X < b)$ and $\text{Var}(D_2 \mid a < X < b)$.

$$\begin{aligned}
\text{Var}(D_1 \mid a < X < b) &= \text{Var}(Y - X \mid a < X < b) \\
&= \text{Var}(Y \mid a < X < b) + \text{Var}(X \mid a < X < b) \\
&- 2\text{Cov}(Y, X \mid a < X < b) \\
&= \sigma_2^2 \left(1 - (\frac{\rho_{12}\delta_\phi}{\delta_\Phi})^2 + \frac{\rho_{12}^2\delta_{\phi'}}{\delta_\Phi}\right) \\
&+ \sigma_1^2 \left(1 - (\frac{\delta_\phi}{\delta_\Phi})^2 + \frac{\delta_{\phi'}}{\delta_\Phi}\right) \\
&- 2\rho_{12}\sigma_1\sigma_2 \left(1 - (\frac{\delta_\phi}{\delta_\Phi})^2 + \frac{\delta_{\phi'}}{\delta_\Phi}\right). \\
&= \sigma_1^2 + \sigma_2^2 - 2\rho_{12}\sigma_1\sigma_2 + (\rho_{12}\sigma_2 - \sigma_1)^2 \left(-(\frac{\delta_\phi}{\delta_\Phi})^2 + \frac{\delta_{\phi'}}{\delta_\Phi}\right) \\
&= \text{Var}(D_1) + \theta_1^2\Delta.
\end{aligned} \tag{0.96}$$

Similarly, we can find

$$\begin{aligned}
\text{Var}(D_2 \mid a < X < b) &= \text{Var}(Z - Y \mid a < X < b) \\
&= \text{Var}(Z \mid a < X < b) + \text{Var}(Y \mid a < X < b) \\
&- 2\text{Cov}(Z, Y \mid a < X < b) \\
&= \sigma_3^2 \left(1 - (\frac{\rho_{13}\delta_\phi}{\delta_\Phi})^2 + \frac{\rho_{13}^2\delta_{\phi'}}{\delta_\Phi}\right)
\end{aligned}$$

$$
\begin{aligned}
&+ \quad \sigma_2^2 \left(1 - (\frac{\rho_{12}\delta_\phi}{\delta_\Phi})^2 + \frac{\rho_{12}^2\delta_{\phi'}}{\delta_\Phi}\right) \\
&- \quad 2\sigma_2\sigma_3 \left[\rho_{23} + \rho_{12}\rho_{13}\left(-(\frac{\delta_\phi}{\delta_\Phi})^2 + \frac{\delta_{\phi'}}{\delta_\Phi}\right)\right] \\
&= \quad \sigma_3^2 + \sigma_2^2 - 2\rho_{23}\sigma_2\sigma_3 + (\rho_{13}\sigma_3 - \rho_{12}\sigma_2)^2 \left(-(\frac{\delta_\phi}{\delta_\Phi})^2 + \frac{\delta_{\phi'}}{\delta_\Phi}\right) \\
&= \quad \mathrm{Var}(D_2) + \theta_2^2\Delta. \qquad (0.97)
\end{aligned}
$$

Therefore we have

$$
\begin{aligned}
\rho(D_1, D_2 \mid a < X < b) &= \frac{\mathrm{Cov}(D_1, D_2 \mid a < X < b)}{\sqrt{\mathrm{Var}(D_1 \mid a < X < b)}\sqrt{\mathrm{Var}(D_2 \mid a < X < b)}} \\
&= \frac{\mathrm{Cov}(D_1, D_2) - \theta_1\theta_2\Delta}{\sqrt{\mathrm{Var}(D_1) + \theta_1^2\Delta}\sqrt{\mathrm{Var}(D_2) + \theta_2^2\Delta}}. \qquad (0.98)
\end{aligned}
$$

Chapter 1

Introduction:

1.1 History and background:

Sir Francis Galton (1897) was the first researcher to investigate the singly truncated normal distribution. He came across this distribution while he was analysing the registered speeds of American trotting horses. He had extracted data of 5705 appropriate horses, (stallions, geldings and mares, which are equally efficient trotters) from Wallace's year book, Vol. 8-12 (1892-1896) (Sample sizes varied from 982 to 1324 observations each.). These data consisted of running times of horses that qualified for registration by trotting around a one-mile course in not more than 2 minutes and 30 seconds while harnessed to a two-wheeled cart carrying a weight of not less than 150 pounds. Galton added "The object of my inquiry was to test the suitability of these trotting (and pacing) records for investigations into the laws of heredity." He was concerned with the estimation of the joint influences of different ancestors. Thus, he raised the question whether the arithmetical mean of the speeds was the most appropriate estimate of the mean of the complete distribution. After going through a troublesome and tedious investigation, as he called it, he concluded "It would be a strong presumption in the affirmative, if the relative frequency of the various speeds should correspond approximately by the normal law of frequency, because if they do so they would fall into line with numerous anthropometric and other measures which have been often discussed, and which, when treated by methods in which the arithmetic mean was employed, have yielded results that accord with observed facts." So by connecting the mid points of a histogram and drawing the normal curve with mean of the observations which were in the same path he realized that, these data, up to recorded points, followed the normal distribution (This method was reasonably satisfactory for Galton's purposes.). Since records were not usually kept of the slower unsuccessful trotters, their number remained unknown. In today's terminology, the samples were drawn from singly truncated normal distributions. To sum up Galton's work, we can say that he assumed the underlying distribution to be normal. According to Pearson, Galton determined the position of the mode of the full normal distribution (i.e. μ) by inspection of plotted figures of data. By this method he estimated the parameters of the complete distribution from the observations of the registered speeds of American trotting horses.

Pearson (1902) noted that a histogram or a frequency polygon gives us a certain numbers of values of y and x from which to fit the curve

$$y = y_0 e^{-(ax^2+b)} \ . \tag{1.1}$$

At first he suggested finding y_0, a and b by the method of least squares or moments. But he stated that this works be rather tedious and "unmanageable". Later he wrote the probability density function of the truncated normal distribution in the form

$$y = y_0 e^{-\frac{(x-\mu)^2}{2\sigma^2}} \ . \tag{1.2}$$

Using the transformation $y = e^Y$, equation (1.2) can be written as the form

$$Y = a'x^2 + b'x + c' \tag{1.3}$$

where $a' = -\frac{1}{\sigma^2}$, $b' = \frac{\mu}{\sigma^2}$ and $c' = \ln(y_0) - \frac{\mu^2}{\sigma^2}$.

Now using, the method of moments, Pearson fitted a parabola of the second order estimating a', b', c' and eventually μ and σ^2 for Galton's data. His result was almost identical with Galton's result. Pearson thought his method considerably improved things.

A number of workers have studied the maximum likelihood estimation of the parameters of the truncated normal distribution (Cohen, 1950 a, 1957; Raj, 1953; Thompson, 1951) as applied to quality control in biological and medical studies. Moreover, Cohen (1950 a), with the aid of standard tables of the areas and ordinates of the normal distribution, found the asymptotic variances and covariance of the estimators of the singly and doubly truncated distributions. Votaw, Rafferty and Deemer (VRD) (1950) found the maximum likelihood estimators for certain parameters of a truncated trivariate normal, and their asymptotic variances and covariances, when the values of other parameters are known. Raj (1953) discussed the problem of estimating the parameters of the complete bivariate normal population from linearly truncated random samples, with a known truncation point. He showed that the method of moments and the method of maximum likelihood are identical.

The method of moments was considered by Lee (1983) and also by Fisher (1931). Moreover Cohen (1950 b) suggested the method of moments. From 1950-1988 Cohen's publications were concerned with various aspects of truncated distributions. Cohen (1986) and Schneider (1989) published books to discuss various aspects of truncated distributions.

1.2 Definition of the truncated normal distribution:

According to Pearson's definition, a frequency distribution, which is normal, but of which only a portion can be known or observed, is called a truncated normal distribution.

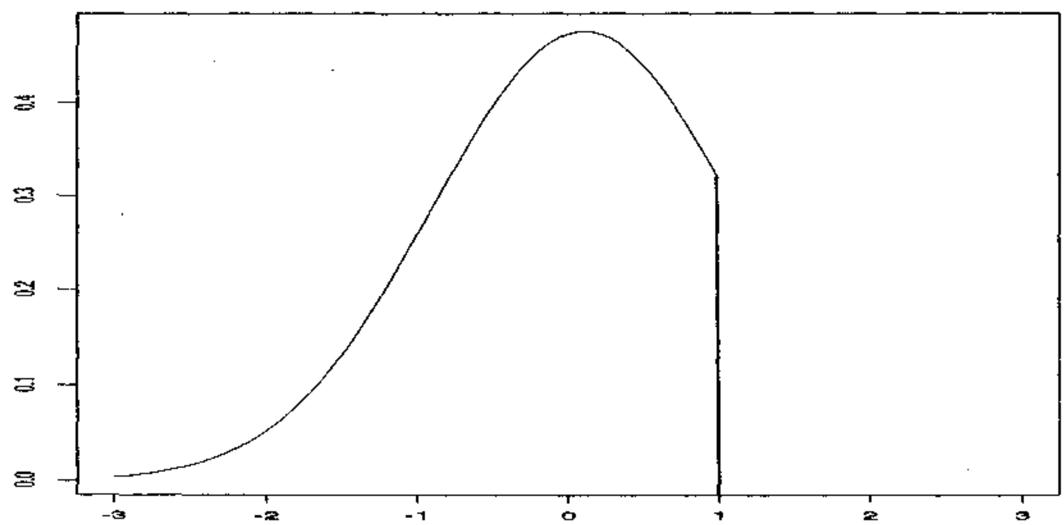
1.2.1 Singly truncated normal distribution from the right:

As we found in section (0.1.6) the p.d.f. of X truncated at c on the right is

$$f(x) = \frac{\phi(\frac{x-\mu}{\sigma})}{\sigma\Phi(\frac{c-\mu}{\sigma})}, \quad -\infty < X < c. \tag{1.4}$$

Figure 1.1 shows the p.d.f. of a standard normal distribution truncated from the right at $c = 1$.

Figure 1.1: p.d.f. of the standard normal distribution truncated from the right at $c = 1$



1.2.2 Singly truncated normal distribution from the left:

Using the notation of section (0.1.2) the p.d.f. of X at truncation point c from the left is

$$f(x) = \frac{\phi(\frac{x-\mu}{\sigma})}{\sigma(1-\Phi(\frac{c-\mu}{\sigma}))}, \quad c < X < \infty. \tag{1.5}$$

1.2.3 Doubly truncated normal distribution:

As we found in section (0.1.5) the p.d.f. of X in truncation points a and b is

$$f(x) = \frac{\phi(x')}{\sigma\delta_\Phi}, \quad a < X < b. \tag{1.6}$$

1.3 The importance of the subject under study:

In various fields of science, such as biology, psychology, medicine, economics and engineering, scientific data are frequently observed from a truncated distribution. The measurement of variables in some parts of the population often presents difficulties in collecting data, or data are preserved only from a part of the sample space. Such a case obtains when monitoring children's heights and only those heights below, say, the 3^{rd} centile survive the initial screening. From these it may be necessary to estimate the parameters of the entire population from which the children were selected.

Pearson (1902) has examples, "The marks of candidates in a competitive examination, wherein candidates below a certain grade have been rejected by a preliminary examination, or are cast out without placing. Or again, the statures of the soldiers in a regiment with a minimum admissible height."

Schneider (1989) stated that " detection limits are another field in which the truncated normal distribution is suitable. Quite often instruments measuring data from a normal population have detection limits; i.e., small values $X \leq a$ and/or large values $X \geq b$ are not observable, and their existence is not even reported by instrument." A second type of truncation appears when the limits are unknown, but the proportion truncated from the population is known.

In this thesis we are considering cases in which the truncation points are known.

1.4 Description of data:

In order to illustrate the methods of estimating the mean and the variance we use two types of data: empirical and simulated. We find the maximum likelihood (ML) estimates and, the maximum product spacing (MPS) estimates of the mean and variance for these two types of data.

1.4.1 Empirical data:

Two sets of data were drawn from the Wessex Growth study on the heights of 1287 school boys and school girls between 4.5 and 5.5 years of age. In these cases, by selecting the heights of short children, we find ML and MPS estimates of the mean and variance of the population from which they were drawn.

1. **Data set 1:**

   The heights of 634 boys were measured. After standardization of the height for age, we identified those children whose heights were particularly short, in other words, those children whose standard deviation score (SDS), $(x - \mu)/\sigma$, did not exceed its third centile (-1.88), or standards published in (1966). It was found that 12 of them were below the third centile compared with 19 that would have been expected if μ and σ had not changed, and the mean, variance and the standard deviation of their scores were, respectively, -2.2833, 0.1963 and 0.4627. From these data, we estimated the mean and variance of the population, using two different methods. Computer programs are presented in the Appendix.

2. **Data set 2:**

   A sample of 653 girls was selected and their heights were measured. After standardizing their heights, nine (compared with 19 expected) of them did not exceed the third centile. The mean, variance and the standard deviation of the scores are, respectively, -2.26, 0.0512 and 0.24. Again we used these data in the programs to estimate the mean and variance of the population.

1.4.2 Ideal samples:

In this section we construct what we might call the ideal sample in which the observations are placed at the expected positions of the order statistics from the distribution (much as in the construction of normal scores from a complete normal distribution).

Let y_i be the i^{th} component of the ideal sample of size n. Then using equation (0.1) from Chapter 0 we have

$$F(y_i) = \frac{\Phi(y_i')}{\Phi(c')} = \frac{\Phi(\frac{y_i-\mu}{\sigma})}{\Phi(\frac{c-\mu}{\sigma})} = \frac{i}{n+1}.$$

From which

$$y_i' = \Phi^{-1}\left(\frac{i}{n+1}\Phi(\frac{c-\mu}{\sigma})\right), \tag{1.7}$$

where c is the truncation point.

Using the formula, for the standard normal distribution with the various truncation points $c = -1.88$, $c = -1$, $c = 0$, $c = 1$ and $c = 3$ we construct ideal samples $y_i' = y_i$, (see Appendix Program 1) for two different sample sizes 5 and 10 as below:

Table 1.1: The ideal sample of size 5 and 10 and its mean, variance and standard deviation for different truncation points

c	-1.88		-1		0		1		3	
n	5	10	5	10	5	10	5	10	5	10
y_i	-2.5752	-2.7783	-1.9359	-2.1856	-1.3830	-1.6906	-1.0793	-1.4291	-0.9683	-1.3359
	-2.3257	-2.5450	-1.6175	-1.8980	-0.9674	-1.3352	-0.5815	-1.0238	-0.4319	-0.9094
	-2.1694	-2.4000	-1.4096	-1.7139	-0.6745	-1.0968	-0.2002	-0.7406	-0.0016	-0.6057
	-2.0530	-2.2928	-1.2493	-1.5744	-0.4307	-0.9084	0.1532	-0.5073	0.4282	-0.3501
	-1.9591	-2.2068	-1.1160	-1.4602	-0.2104	-0.7478	0.5276	-0.2991	0.9629	-0.1157
		-2.1347		-1.3624		-0.6045		-0.1032		0.1418
		-2.0722		-1.2761		-0.4727		0.0888		0.3465
		-2.0168		-1.1984		-0.3487		0.2824		0.6016
		-1.9670		-1.1273		-0.2299		0.4912		0.9043
		-1.9217		-1.0615		-0.1142		0.7220		1.3277
$\bar{y}$	-2.2165	-2.2335	-1.4657	-1.4858	-0.7332	-0.7549	-0.2360	-0.2517	-0.0022	-0.0024
$\text{Var}(y_i)$	0.0471	0.0677	0.0832	0.1171	0.1691	0.2306	0.3133	0.4214	0.4470	0.6190
$\text{sd}(y_i)$	0.2428	0.2742	0.3226	0.3606	0.4597	0.5062	0.6258	0.6843	0.7475	0.8293

By using these data we find the ML and MPS estimates of the mean and variance of the complete normal distribution.

1.4.3 Simulated data:

We took samples from truncated normal distributions, simulations the process $R = 10000$ times for each sample size and each truncation point. Using these simulations we estimated $E(\hat{\mu})$, $\sigma(\hat{\mu})$, $E(\hat{\sigma}^2)$, $\sigma(\hat{\sigma}^2)$, $E(\tilde{\mu})$, $\sigma(\tilde{\mu})$, $E(\tilde{\sigma^2})$ and $\sigma(\tilde{\sigma^2})$ and compared these with the theoretical expansions derived in Chapter 2 where appropriate.

For $c > \mu$ we simulated from the normal distribution (using a NAG routine) and rejected unwanted (i.e. truncated) values.

For $c \leq \mu$ we used an exponential envelope function for the random deviate generation, which gave a more efficient method than initially generating from the normal distribution.

Chapter 2

The one parameter case of maximum likelihood estimator for the truncated normal distribution:

2.1 Introduction:

The purpose of this chapter is to describe the Maximum Likelihood (ML) method of estimating separately the mean and variance of a singly truncated normal population from a sample.

Theoretical results and simulations for the different methods are also presented and comparisons are made. In section 2.2 the case of estimating the mean, when the variance is known, is considered. In section 2.3 two methods of solving the log likelihood equation are described. In section 2.4 the theoretical formulae for $E(\hat{\mu})$ and $\mathrm{Var}(\hat{\mu})$ based on Cox & Hinkley's and Shenton & Bowman's methods are derived when σ is known. In section 2.8, we consider the ML estimator of σ^2 when the mean is known. In section 2.9 the theoretical formulae for $E(\hat{\sigma}^2)$ and $\mathrm{Var}(\hat{\sigma}^2)$ based on Cox & Hinkley's and Shenton & Bowman's methods are derived when μ is known. Sections 2.5, 2.6 and 2.10 are concerned with simulation studies to investigate the properties of $\hat{\mu}$ and $\hat{\sigma}^2$.

2.2 Likelihood equation when variance is known:

We begin by deriving the likelihood equation for the case where the variance is known. Suppose X has a normal distribution with mean μ and variance σ^2 truncated at the point c, i.e. $x \leq c$. Let $\Phi(x)$ be as defined in Chapter 0. Then the probability density function of the singly truncated random variable X is

$$f(x,\mu) = \frac{1}{\sigma\sqrt{2\pi}\Phi(\frac{c-\mu}{\sigma})} e^{\frac{-(x-\mu)^2}{2\sigma^2}} = \frac{\phi(\frac{x-\mu}{\sigma})}{\sigma\Phi(\frac{c-\mu}{\sigma})}, \; -\infty < x < c. \tag{2.1}$$

Suppose a random sample of size n is selected from the population with distribution as in (2.1). Then the likelihood function is

$$L(\mathbf{x},\mu) = (\frac{1}{2\pi\sigma^2})^{\frac{n}{2}} \frac{e^{-\sum_{i=1}^{n}(x_i-\mu)^2/2\sigma^2}}{[\Phi((c-\mu)/\sigma)]^n}.$$

The natural logarithm of the likelihood function is

$$l(\mathbf{x},\mu) = \frac{-n}{2}\ln(2\pi\sigma^2) - \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2} - n\ln\Phi(\frac{c-\mu}{\sigma}). \tag{2.2}$$

If μ is unknown and σ^2 is known, the estimate $\hat{\mu}$ is obtained by solving the following equation:

$$\frac{\partial l(\mathbf{x},\mu)}{\partial\mu}\,|_{\mu=\hat{\mu}} = \frac{\sum_{i=1}^{n}(x_i-\hat{\mu})}{\sigma^2} + \frac{n\phi(\frac{c-\hat{\mu}}{\sigma})}{\sigma\Phi(\frac{c-\hat{\mu}}{\sigma})} = 0. \tag{2.3}$$

To simplify the above equation, let $c' = \frac{(c-\mu)}{\sigma}$, $\hat{c}' = \frac{(c-\hat{\mu})}{\sigma}$ and $\psi(c') = \frac{\sigma\phi(c')}{\Phi(c')}$.

We have

$$\frac{\partial l(\mathbf{x},\mu)}{\partial\mu} = \frac{n\,(\bar{x}-\mu+\psi(c'))}{\sigma^2}, \tag{2.4}$$

which, on being equated to zero, gives

$$\hat{\mu} = \bar{x} + \psi(\hat{c}'). \tag{2.5}$$

It is impossible to solve (2.5) algebraically. Therefore two different iterative methods are used to solve it. These are described in the following section.

Note: We noted that if $X \sim N(0,1)$; suppose $\phi(x)$ and $\Phi(x)$ are p.d.f. and c.d.f. of x, then using Abramowitz & Stegun (1965, p. 932) we can write

$$\Phi(x) = \begin{cases} 1 - \phi(x)u(x) & ; \quad x > 0 \\ \phi(-x)u(-x) = \phi(x)u(-x) & ; \quad x < 0 \end{cases}$$

where $u(x)$ is the continued fraction expansion:

$$u(x) = \frac{1}{x+}\frac{1}{x+}\frac{2}{x+}\frac{3}{x+}\frac{4}{x+}\cdots.$$

Theorem 2.1 *For a fixed value of σ, the function $l(\mathrm{x},\mu)$ has a local maximum.*

Proof: We prove this theorem for the two cases $c' > 0$ ($\mu \to -\infty$) and $c' < 0$ ($\mu \to \infty$), separately.

Since $l(\mathrm{x},\mu)$ is continuous and differentiable, if we show that $l(\mathrm{x},\mu) \to -\infty$ as $\mu \to -\infty$ and $l(\mathrm{x},\mu) \to -\infty$ as $\mu \to \infty$, then we conclude that $l(\mathrm{x},\mu)$ has a local maximum.

1. For $c' > 0$ ($\mu < c$):

   Consider $\Phi(c') = \Phi(\frac{c}{\sigma} - \frac{\mu}{\sigma}) \to 1$ as $\mu \to -\infty$ and $\sum_{i=1}^{n}(x_i - \mu)^2 \to \infty$ as $\mu \to -\infty$. Therefore, implies that $l(\mathrm{x},\mu) \to -\infty$ as $\mu \to -\infty$.

2. For $c' < 0$ ($\mu > c$):

   Using the above note, $l(\mathrm{x},\mu)$ can be written as

$$\begin{aligned} l(\mathrm{x},\mu) &= -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) \\ &- n\ln u(-c') - \frac{\sum_{i=1}^{n}[(c-x_i)(2\mu - c - x_i)]}{2\sigma^2}. \end{aligned}$$

   Since $c > x_i$ for all i and $\mu > c$, then $(c - x_i) > 0$ and $(2\mu - c - x_i) > 0$, therefore $-\frac{\sum_{i=1}^{n}[(c-x_i)(2\mu-c-x_i)]}{2\sigma^2} \to -\infty$ as $\mu \to \infty$. Also we know that $u(-c')$ is bounded and $-n\ln u(-c')$ is constant as $\mu \to \infty$.

Therefore $l(\mathbf{x}, \mu) \to -\infty$ as $\mu \to \infty$ and the theorem is proved.

Theorem 2.2 *The equation (2.5) is satisfied by $\hat{\mu} = \mu$ if $\bar{x} = \mu_c$ where $\bar{x} = \sum_{i=1}^{n} x_i/n$ and $\mu_c = E(X)$.*

Proof: Suppose we know that the solution of $\frac{\partial l}{\partial \mu} = 0$ is $\mu = \hat{\mu}$ and we can solve equation (2.5) to find the ML estimate.

Also from section (0.1.8) we know that the first moment of X is unique and can be written as

$$E(X) = \mu_c = \mu - \psi(c').$$

By using the above equation and the fact that equation (2.5) can be written in the form

$$\hat{\mu} - \psi(\hat{c}') = \bar{x},$$

and using the assumption of the theorem, the above equation can be expressed as

$$\hat{\mu} - \psi(\hat{c}') - \mu + \psi(c') = 0.$$

In view of the fact that

$$\mu - \psi(c') - \mu + \psi(c') = 0,$$

it can be seen that $(\hat{\mu} = \mu)$ is a solution of equation (2.5). This is equivalent to the method of moments.

2.3 Methods of solving equation (2.5):

In this section we outline a simple method and the false position method of solving $\frac{\partial l}{\partial \mu} = 0$ and use them on the samples described in Chapter 1.

2.3.1 Simple method:

Let μ_n be the n^{th} iterate of $\hat{\mu}$. We find μ_{n+1} from equation (2.5) by

$$\mu_{n+1} = \bar{x} + \psi[(c - \mu_n)/\sigma]. \tag{2.6}$$

To find the ML estimate of μ, we assume an initial value for μ_0, say $\mu_0 = \bar{x}$, then find μ_1 from (2.6) with $n = 0$ and continue iteratively.

According to a theorem in numerical analysis (Jacques & Judd (1957)) for an algorithm of $\mu_{n+1} = g(\mu_n)$ to converge on an interval $I = [\alpha - A, \alpha + A]$, it is sufficient for $|g'(\mu)| \leq k$, on the interval where $k < 1$.

Using the above theorem we can see whether $\bar{x}+\psi[(c-\mu_n)/\sigma]$ satisfies the above condition of the theorem or not.

Taking derivative of both sides of equation (2.6) with respect to μ we have $\frac{\partial\psi(c')}{\partial\mu} = 1$. Therefore we cannot certainly say that the method is convergent. It might be convergent or divergent.

A program has been written in Fortran (see Appendix , Program 2) which finds the value of $\hat{\mu}$ to a derived accuracy by stopping when the absolute error $|\mu_n - \mu_{n-1}| < \varepsilon$, for a suitable ε, say 10^{-4} or 10^{-5}, specified by the user. This has been used on the data sets described in Chapter 1 and the results are given below.

2.3.1.1 Estimates of μ in data sets 1 and 2:

1. Using the data set 1 (boys) and letting $\sigma = 1$ and $\varepsilon = 10^{-5}$, $\hat{\mu}$ is estimated on the 70^{th} iteration of (2.6) and is found to be

$$\hat{\mu} = -0.1266.$$

2. Using the data set 2 (girls) and letting $\sigma = 1$ and $\varepsilon = 10^{-5}$, $\hat{\mu}$ is estimated on the 78^{th}

iteration of (2.6) and is found to be

$$\hat{\mu} = 0.0615.$$

Because the simple method takes a large number of iterations to reach convergence, we try find a method that converges faster.

2.3.2 False position method:

The false position method is used to solve $f(x) = 0$ for cases where the derivative of the function is not readily available or where the evaluation of the second derivative of $f(x)$ requires considerable computational effort. When the method does converge, its rate of convergence is not as fast as for the scoring method (used in Chapter 3), but it is considerably faster than the simple method.

According to Conte (1965), the proper solution of the equation by the False position method is obtained as follows:

1. Choose two approximations x_0 and x_1 such that $f(x_0)f(x_1) < 0$; i.e $f(x_0)$ and $f(x_1)$ are of opposite signs.

2. Find another approximation from the following formula

$$x_2 = \frac{x_0 f(x_1) - x_1 f(x_0)}{f(x_1) - f(x_0)}. \tag{2.7}$$

3. If $|x_2 - x_1| < \varepsilon$ or $|x_2 - x_0| < \varepsilon$ for the prescribed ε, x_2 is accepted as the answer. If not, go to step 4.

4. If $f(x_2)f(x_0) < 0$, replace x_1 by x_2, leave x_0 unchanged, and compute the next approximation from (2.7), otherwise replace x_0 by x_2, leave x_1 unchanged, and compute the next approximation.

Figure 2.1: The false position method for solving f(x)=0 when f is concave.

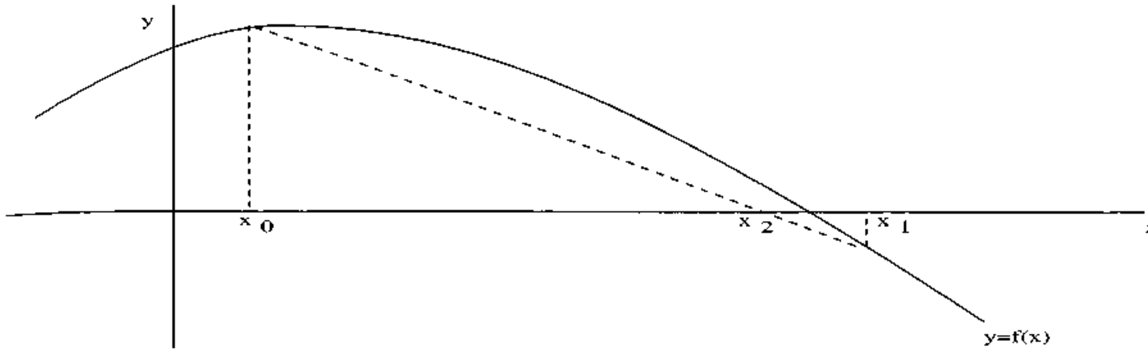
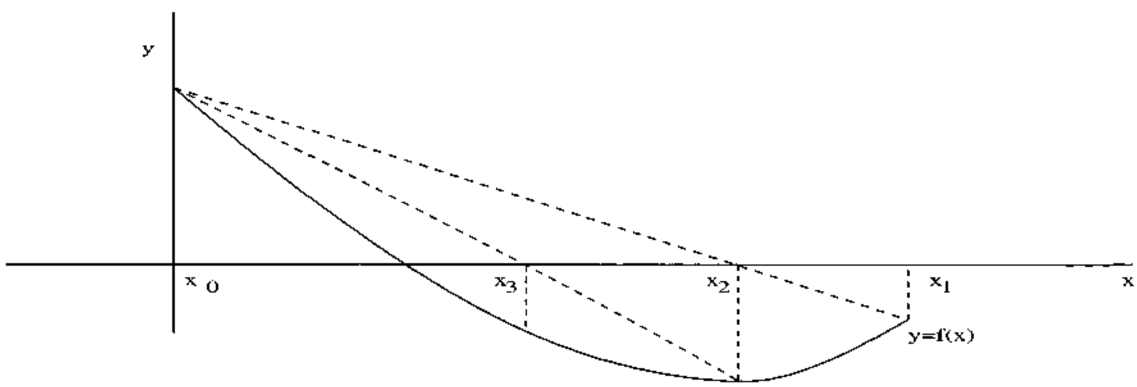


Figure 2.2: The false position method for solving f(x)=0, when f is convex.



The Figures 2.1 and 2.2 show this method of solution for both concave and convex functions.

The NAG routine C05ADF follows the above procedure and can be used to find the root of the equation

$$\hat{\mu} - \bar{x} - \psi((c - \hat{\mu})/\sigma) = 0. \tag{2.8}$$

The method is now used on the different samples of Chapter 1. The computer program is given in Appendix, Program 3.

2.3.2.1 Estimates of μ in data sets 1 and 2:

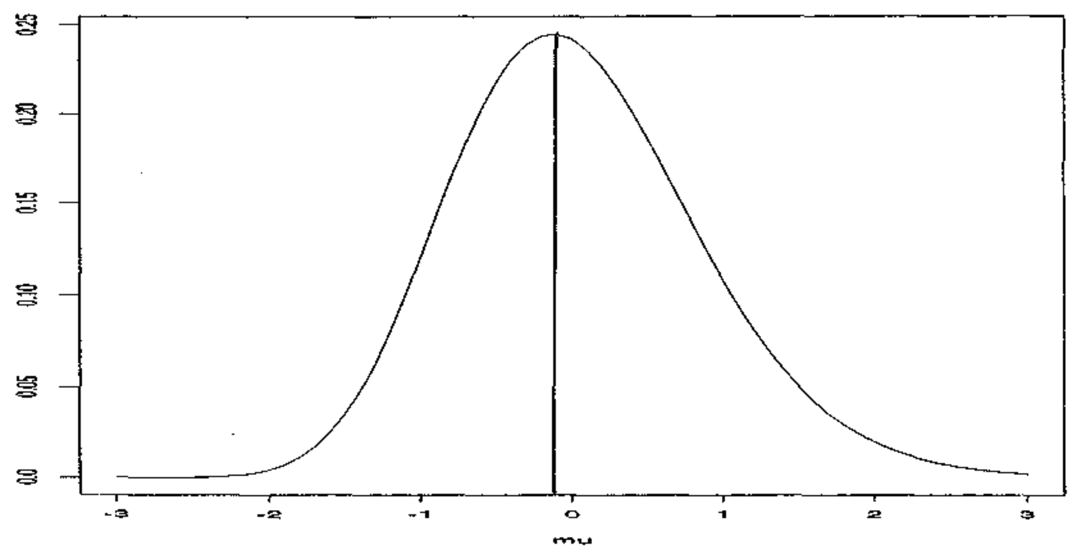
1. Using the data set 1 (boys) and letting $\sigma = 1$ and $\varepsilon = 10^{-5}$ we find that the ML estimate of μ is

$$\hat{\mu} = -0.1266,$$

after only 11 iteration.

On plotting the likelihood against μ, when $\sigma = 1$, we get Figure 2.3 (see Appendix Program 4).

Figure 2.3: likelihood versus μ for data set 1 (boys) ($\sigma = 1$)



2. Using the data set 2 (girls) and letting $\sigma = 1$ and $\varepsilon = 10^{-5}$ we find that the ML estimate of μ is

$$\hat{\mu} = 0.0615,$$

after 10 iteration.

2.3.2.2 Estimates of μ for ideal samples of size 5 and 10:

Using the ideal samples of size 5 and 10 and letting $\sigma = 1$ and $\varepsilon = 10^{-5}$, $\hat{\mu}$ is estimated for various truncation points in Table 2.1.

By running this program for various cases of data we have got that, for $\varepsilon = 10^{-5}$ four decimal places of results are valid.

Table 2.1: The estimate of the μ in ideal samples of size 5 and 10 for different truncation points

c	-1.88		-1		0		1		3	
n	5	10	5	10	5	10	5	10	5	10
$\bar{y}$	-2.2165	-2.234	-1.4657	-1.4858	-0.7332	-0.7549	-0.2360	-0.2517	-0.0022	-0.0024
$\hat{\mu}$	0.4709	0.2940	0.3281	0.2100	0.1883	0.1227	0.0836	0.0578	0.0023	0.0021

From Table 2.1 we can see that the bias of $\hat{\mu}$ is positive for all truncation points.

Moreover, the estimate of μ in sample size 5 in comparison with its counterpart in sample size 10 is rather high.

We see that these two methods give the same solutions for the data sets. But the first method for data set 1 takes 71 iterations and for data set 2 takes 78 iterations whereas second method reaches the same convergence points after at most 12 iterations. In oyher words the False Position method improves rate of convergence.

2.4 Theoretical results:

The aim of this section is to find the expected value and variance of the maximum likelihood estimator of μ when σ is assumed known. Formulae for these are given in Cox & Hinkley (1974, p. 310) and Shenton & Bowman (1977, p. 15).

2.4.1 Cox & Hinkley method:

In this section we derive the formulae using the following notation.

1. For a single observation x, let $S(\mu)$ be the score defined by

$$S(\mu) = \frac{\partial \ln f(x,\mu)}{\partial \mu}.$$

2. Let $S'(\mu)$ and $S''(\mu)$ be the first and second derivatives of $S(\mu)$ with respect to μ.

3. Let

$$S_j(\mu) = \frac{\partial \ln f(x_j,\mu)}{\partial \mu}.$$

4. Let

$$S_.(\mu) = \frac{\partial \ln L(\mathbf{x},\mu)}{\partial \mu} = \Sigma_{j=1}^{n} S_j(\mu).$$

5. Let

$$I(\mu) = [S(\mu)]^2 = -S'(\mu).$$

6. Let

$$i(\mu) = E[S(\mu)]^2 = E[-S'(\mu)],$$

since regularity conditions hold.

7. Let

$$i_.(\mu) = E[S_.(\mu)]^2 = E[-S'_.(\mu)] = ni(\mu). \qquad (2.9)$$

8. Let

$$\kappa_{ij}(\mu) = E\left([S(\mu)]^i[S'(\mu) + i(\mu)]^j\right).$$

Now it is well known that

$$E[S_.(\mu)] = E[S(\mu)] = 0.$$

Before going further we write the following definitions from Bisop, Fienberg and Holland (1975) and Mann and Wald (1943) about the order relationships and stochastic limit.

Definition 2.1 *$a_n = O(b_n)$ (Read: a_n is big O of b_n) if the ratio $|a_n/b_n|$ is bounded for large n; in detail, if there exists a number K and an integer $n(K)$ such that if n exceeds $n(K)$ then $|a_n| < K|b_n|$.*

Definition 2.2 *$a_n = o(b_n)$ (Read: a_n is little o of b_n) if the ratio $|a_n/b_n|$ converges to zero; in detail, if for any $\varepsilon > 0$, there exists an integer $n(\varepsilon)$ such that if n exceeds $n(\varepsilon)$ then $|a_n| < \varepsilon|b_n|$.*

Definition 2.3 *We write $\hat{\mu}_n = O_p[f(n)]$ ($\hat{\mu}_n$ is of probability order $O[f(n)]$) if for each $\varepsilon > 0$ there exists an $A_\varepsilon > 0$ such that $P(|\hat{\mu}_n| \leq A_\varepsilon f(n)) \geq 1 - \varepsilon$ for all values of n.*

Definition 2.4 *We write $\hat{\mu}_n = o_p[f(n)]$ ($\hat{\mu}_n$ is of probability order $o[f(n)]$) if $plim_{n\to\infty} \frac{\hat{\mu}_n}{f(n)} = 0$.*

Using Taylor's expansion, we can expand the function $S_.(\hat{\mu})$ about μ.

Then, since $S_.(\hat{\mu}) = 0$, we have, from Cox & Hinkley

$$0 = S_.(\hat{\mu}) = S_.(\mu) + (\hat{\mu} - \mu)S'_.(\mu) + \frac{1}{2}(\hat{\mu} - \mu)^2 S''_.(\mu) + O_p(n^{-\frac{1}{2}}). \tag{2.10}$$

where $O_p(n^{-\frac{1}{2}}) = \lim_{n\to\infty} P(n^{\frac{1}{2}}\hat{\mu}_n)$. Thus, to first order, solution to (2.10) can be written as

$$i(\mu)(\hat{\mu} - \mu)\sqrt{n} = \frac{S_.(\mu)}{\sqrt{n}}.$$

Now as $n \to \infty$, using the central limit theorem the limiting distribution of $i(\mu)(\hat{\mu} - \mu)\sqrt{n}$ is $N(0, i(\mu))$. In other words, we have

$$\hat{\mu} - \mu = \frac{S_.(\mu)}{ni(\mu)} \sim N(0, \frac{1}{i(\mu)}) \tag{2.11}$$

asymptotically.

Taking expectations through the equation (2.10), we have

$$E[S(\hat{\mu})] = E[(\hat{\mu}-\mu)S'_{\cdot}(\mu)] + \frac{1}{2}E[(\hat{\mu}-\mu)^2 S''_{\cdot}(\mu)] + O(n^{-\frac{1}{2}}) = 0. \tag{2.12}$$

Since

$$E[S_{\cdot}(\mu)] = 0,$$

and

$$\mathrm{Cov}(X,Y) = E(XY) - E(X)E(Y),$$

we can write from equation (2.12)

$$E[(\hat{\mu}-\mu)]E[S'_{\cdot}(\mu)] + \mathrm{Cov}[\hat{\mu}, S'_{\cdot}(\mu)] + \frac{1}{2}E[(\hat{\mu}-\mu)^2]E[S''_{\cdot}(\mu)] + \frac{1}{2}\mathrm{Cov}[(\hat{\mu}-\mu)^2, S''_{\cdot}(\mu)] + O(n^{-1}) = 0. \tag{2.13}$$

If we substitute $\hat{\mu}$ from (2.11) into $\mathrm{Cov}[\hat{\mu}, S'_{\cdot}(\mu)]$, then we have

$$\begin{aligned}
\mathrm{Cov}[\hat{\mu}, S'_{\cdot}(\mu)] &= \mathrm{Cov}[\frac{S_{\cdot}(\mu)}{ni(\mu)}, S'_{\cdot}(\mu)] \\
&= \frac{\mathrm{Cov}[S_{\cdot}(\mu), S'_{\cdot}(\mu)]}{ni(\mu)} \\
&= \frac{E[S_{\cdot}(\mu)S'_{\cdot}(\mu)]}{ni(\mu)} \\
&= \frac{nE[S(\mu)S'(\mu)]}{ni(\mu)} = \frac{\kappa_{11}(\mu)}{i(\mu)}.
\end{aligned} \tag{2.14}$$

Now if we square both sides of equation (2.11), and take the expected value, we have

$$\begin{aligned}
E(\hat{\mu}-\mu)^2 &= \frac{E[S_{\cdot}(\mu)]^2}{[ni(\mu)]^2} \\
&= \frac{E\left\{\sum_{j=1}^{n}[S_j(\mu)]\right\}^2}{[ni(\mu)]^2} \\
&= \frac{E\left\{\sum_{j=1}^{n}[S_j(\mu)]^2 + 2\sum_{i<j} S_i(\mu)S_j(\mu)\right\}}{[ni(\mu)]^2} \\
&= \frac{\sum_{j=1}^{n} E[S_j(\mu)]^2 + 0}{[ni(\mu)]^2},
\end{aligned} \tag{2.15}$$

as the observations are independent.

Substituting equation (2.9) into equation (2.16), we can write

$$E(\hat{\mu}-\mu)^2 = \frac{ni(\mu)}{[ni(\mu)]^2} = \frac{1}{ni(\mu)} \tag{2.16}$$

to first order.

Now we want to find the value of $E[S''(\mu)]$.

According to the definition of the probability density function $f(x,\mu)$, we have

$$\int_{-\infty}^{\infty} f(x_i,\mu)\,dx_i = 1, \qquad i = 1,2\ldots n.$$

If we differentiate this equation with respect to μ, we obtain

$$\int_{-\infty}^{\infty} \frac{\partial f(x_i,\mu)}{\partial\mu}\,dx_i = \int_{-\infty}^{\infty} \frac{\partial \ln f(x_i,\mu)}{\partial\mu} f(x_i,\mu)\,dx_i = E[S(\mu)] = 0. \tag{2.17}$$

Taking the derivative of equation (2.17) with respect to μ, we have

$$\int_{-\infty}^{\infty} \frac{\partial^2 \ln f(x_i,\mu)}{\partial\mu^2}\,dx_i + \int_{-\infty}^{\infty} \frac{\partial\left[\frac{\partial \ln f(x_i,\mu)}{\partial\mu} f(x_i,\mu)\right] dx_i}{\partial\mu} = 0,$$

from which we can find the following formula

$$\int_{-\infty}^{\infty} \frac{\partial^2 \ln f(x_i,\mu)}{\partial\mu^2}\,dx_i + \int_{-\infty}^{\infty} \left[\frac{\partial \ln f(x_i,\mu)}{\partial\mu}\right]^2 f(x_i,\mu)\,dx_i = 0. \tag{2.18}$$

Therefore, we obtain

$$E[-S'(\mu)] = E[S(\mu)]^2.$$

Again, taking the derivative of both sides of the equation (2.18) with respect to μ, we have

$$\begin{aligned}\int_{-\infty}^{\infty} \frac{\partial^3 \ln f(x_i,\mu)}{\partial\mu^3}\,dx_i &+ 3\int_{-\infty}^{\infty} \left[\frac{\partial^2 \ln f(x_i,\mu)}{\partial\mu^2} \times \frac{\partial \ln f(x_i,\mu)}{\partial\mu}\right] f(x_i,\mu)\,dx_i \\ &+ \int_{-\infty}^{\infty} \left[\frac{\partial \ln f(x_i,\mu)}{\partial\mu}\right]^3 f(x_i,\mu)\,dx_i = 0. \end{aligned} \tag{2.19}$$

Therefore, from equation (2.19), we can write

$$E[S''(\mu)] \;=\; -3E[S(\mu)S'(\mu)] - E[S(\mu)]^3. \tag{2.20}$$

The second covariance in equation (2.13) is $o(n^{-1})$, [see Cox & Hinkley (1974), p. 309], which means that

$$\lim_{n\to\infty} n\mathrm{Cov}[(\hat{\mu}-\mu)^2, S''(\mu)] = 0. \tag{2.21}$$

From equation (2.13) we have

$$E(\hat{\mu}-\mu) = -\frac{\mathrm{Cov}[\hat{\mu}, S'(\mu)] + \frac{1}{2}E[(\hat{\mu}-\mu)^2]E[S''(\mu)] + \frac{1}{2}\mathrm{Cov}[(\hat{\mu}-\mu)^2, S''(\mu)]}{E[S'(\mu)]} + \frac{O(n^{-\frac{1}{2}})}{E[S'(\mu)]}. \tag{2.22}$$

Substituting equations (2.14), (2.15), and (2.21) into equation (2.13), we obtain

$$E(\hat{\mu}-\mu) = -\frac{\kappa_{11}(\mu) + \kappa_{30}(\mu)}{2n[i(\mu)]^2} + O(n^{-2}). \tag{2.23}$$

In particular, we are interested in deriving the expected value and the variance of the maximum likelihood estimator of μ in the truncated normal case.

2.4.1.1 Expected value of $\hat{\mu}$ for the truncated normal distribution:

The aim of this section is to find the expected value of the maximum likelihood estimator of μ in the distribution given at (2.1). Following the notation introduced in the last section, we have

$$S(\mu) = \frac{X - \mu + \psi(c')}{\sigma^2} \tag{2.24}$$

and

$$S'(\mu) = \frac{1}{\sigma^2}\left[-1 + \frac{\partial\psi(c')}{\partial\mu}\right], \tag{2.25}$$

where

$$\frac{\partial\psi(c')}{\partial\mu} = \frac{c'\psi(c')}{\sigma} + \left[\frac{\psi(c')}{\sigma}\right]^2. \tag{2.26}$$

Substituting equation (2.26) into (2.25), and also using $\frac{\partial \psi(c')}{\partial \mu} = \psi'(c')\frac{\partial c'}{\partial \mu}$ we obtain

$$\begin{aligned} S'(\mu) &= \frac{1}{\sigma^2}\left\{-1 + \frac{c'\psi(c')}{\sigma} + \left[\frac{\psi(c')}{\sigma}\right]^2\right\} \\ &= -\frac{1}{\sigma^2}\left[1 + \psi'(c')/\sigma\right]. \end{aligned} \tag{2.27}$$

Since $S'(\mu)$ is independent of X and $E[S(\mu)] = 0$, we can write

$$\begin{aligned} \kappa_{11}(\mu) &= E\{S(\mu)[S'(\mu) + i(\mu)]\} \\ &= E[S(\mu)S'(\mu)] + i(\mu)E[S(\mu)] \\ &= S'(\mu)E[S(\mu)] + i(\mu)E[S(\mu)] \\ &= 0. \end{aligned} \tag{2.28}$$

Since from Chapter 0, section (0.1.8), we know that

$$\mu_c = E(X) = \mu - \psi(c')$$

and $S'(\mu)$ does not depend on X, we have

$$\begin{aligned} \kappa_{30}(\mu) &= E\left[S(\mu)\right]^3 \\ &= E\left[\frac{X - \mu + \psi(c')}{\sigma^2}\right]^3 \\ &= \frac{E(X-\mu_c)^3}{\sigma^6} \\ &= \frac{\mu_3(X)}{\sigma^6}. \end{aligned} \tag{2.29}$$

From Chapter 0, section (0.1.16), we have that

$$\mu_3(X) = -\sigma^3\left[(c'^2 - 1)\frac{\psi(c')}{\sigma} + 3c'\frac{\psi^2(c')}{\sigma^2} + 2\frac{\psi^3(c')}{\sigma^3}\right]. \tag{2.30}$$

Now we can find $i(\mu)$,

$$\begin{aligned} i(\mu) &= E[-S'(\mu)] = -\frac{1}{\sigma^2}\left\{-1 + \frac{c'\psi(c')}{\sigma} + \left[\frac{\psi(c')}{\sigma}\right]^2\right\} \\ &= \frac{1}{\sigma^2}\left[1 + \psi'(c')/\sigma\right]. \end{aligned} \tag{2.31}$$

Thus we obtain

$$\begin{aligned} E(\hat{\mu}) &= \mu - \frac{\kappa_{11}(\mu) + \kappa_{30}(\mu)}{2n[i(\mu)]^2} + O(n^{-2}) \\ &= \mu + \frac{\psi''(c)}{2n[1 + \psi'(c')/\sigma]^2} + O(n^{-2}) \\ &= \mu + \frac{b(\mu)}{n} + O(n^{-2}) \end{aligned} \tag{2.32}$$

where

$$b(\mu) = \frac{\psi''(c)}{2[1 + \psi'(c')/\sigma]^2}. \tag{2.33}$$

2.4.1.2 Variance of $\hat{\mu}$ for the truncated normal distribution:

The aim of this section is to find the variance of the maximum likelihood estimator of μ in the distribution given at (2.1). According to Cox & Hinkley (1974), we know that

$$\begin{aligned} \mathrm{Var}(\hat{\mu}) &= \frac{1}{ni(\mu)} + \frac{2b'(\mu)}{n^2 i(\mu)} + \\ & \frac{2[\kappa_{20}(\mu)\kappa_{02}(\mu) - \kappa_{11}^2(\mu)] + [\kappa_{11}(\mu) + \kappa_{30}(\mu)]^2}{2n^2 i^4(\mu)} + O(n^{-3}). \end{aligned} \tag{2.34}$$

Using the second derivative of $\psi(c')$ and equation (2.30), it can be shown that third moment of X

$$\mu_3(X) = -\sigma^2\psi''(c'). \tag{2.35}$$

From equation (2.34) we can obtain

$$b'(\mu) = \frac{\psi''^2(c')}{\sigma^2[1 + \psi'(c')/\sigma]^3} - \frac{\psi'''(c')}{2\sigma[1 + \psi'(c')/\sigma]^2}. \tag{2.36}$$

Using equation (2.31) for $i(\mu)$, we can find the second term of equation (2.34) as

$$\frac{2b'(\mu)}{n^2 i(\mu)} = \frac{1}{n^2}\left\{\frac{2\psi''^2(c')}{\sigma[1+\psi'(c')/\sigma]^4} - \frac{\sigma\psi''(c')}{[1+\psi'(c')/\sigma]^3}\right\}. \tag{2.37}$$

To find the third term of equation (2.34) we have to find $\kappa_{20}(\mu)$ and $\kappa_{02}(\mu)$.

Now, since $S'(\mu)$ is independent of X, we have

$$\begin{aligned}\kappa_{02}(\mu) &= E[S'(\mu)+i(\mu)]^2 \\ &= E\{S'(\mu)-E[S'(\mu)]\}^2 \\ &= \mathrm{Var}[S'(\mu)] = 0 \end{aligned} \tag{2.38}$$

and

$$\begin{aligned}\kappa_{20}(\mu) &= E[S(\mu)]^2 \\ &= E[-S'(\mu)] \\ &= \frac{[1+\psi'(c')/\sigma]}{\sigma^2}. \end{aligned} \tag{2.39}$$

Therefore, substituting $i(\mu)$, $\kappa_{11}(\mu)$, $\kappa_{30}(\mu)$, $\kappa_{02}(\mu)$ and $\kappa_{20}(\mu)$ respectively from equations (2.31), (2.28), (2.29), (2.38) and (2.39), we obtain the variance of $\hat{\mu}$ as

$$\mathrm{Var}(\hat{\mu}) = \frac{\sigma^2}{n[1+\psi'(c')/\sigma]} + \frac{1}{n^2}\left\{\frac{5\psi''^2(c')}{2[1+\psi'(c')/\sigma]^4} - \frac{\sigma\psi'''(c')}{[1+\psi'(c')/\sigma]^3}\right\} + O(n^{-3}). \tag{2.40}$$

Table 2.2 gives a summary of the values of $E(\hat{\mu})$ and $\sigma(\hat{\mu})$ given in equations (2.32) and (2.40) for the truncated standard normal distribution with truncation points $c = -1.88, -1, 0, 1$ and sample sizes $n = 5, 10, 20, 50, 100$. Note that, as the value of c increases, the truncated normal distribution tends to the standard normal distribution (see Program 5 in the Appendix) as we would expected.

Table 2.2: The theoretical results for the expected value and standard deviation of the maximum likelihood estimator of μ for different values of n and c, using the Cox and Hinkley method.

$E(\hat{\mu})$ and $\sigma(\hat{\mu})$ are calculated in $O(n^{-1})$.

$\mu = 0$, $\sigma = 1$

n	$c = -1.88$		$c = -1$		$c = 0$		$c = 1$	
	$E(\hat{\mu})$	$\sigma(\hat{\mu})$	$E(\hat{\mu})$	$\sigma(\hat{\mu})$	$E(\hat{\mu})$	$\sigma(\hat{\mu})$	$E(\hat{\mu})$	$\sigma(\hat{\mu})$
5	0.435	1.608	0.295	1.214	0.165	0.853	0.074	0.611
10	0.217	1.028	0.147	0.787	0.082	0.565	0.037	0.416
20	0.108	0.685	0.074	0.529	0.041	0.386	0.019	0.287
50	0.043	0.417	0.029	0.324	0.016	0.238	0.007	0.180
100	0.021	0.291	0.015	0.227	0.008	0.167	0.004	0.126

2.4.2 Shenton & Bowman methods:

The aim of this section is to find the expected value and the variance of the maximum likelihood estimator μ by using the formulae given in Shenton & Bowman (1977).

Assume we can write

$$\hat{\mu} = \sum_{j=0}^{\infty} A_j(\bar{x} - \mu_1')^j/j! \quad , \tag{2.41}$$

where the A_j are constants. Then the A_j can be found from

$$A_j = \frac{\partial^j \hat{\mu}}{\partial \bar{x}^j} \tag{2.42}$$

as $\bar{x} \to \mu_1'$ and $\hat{\mu} \to \mu$, where μ_1' is the mean of X. In particular $A_0 = \mu$. After expanding the right hand side of equation (2.41) we take expectation. We obtain

$$\begin{aligned}
E(\hat{\mu}) &= A_0 E(\bar{X} - \mu_1')^0 + \frac{A_1}{1!}E(\bar{X} - \mu_1')^1 + \frac{A_2}{2!}E(\bar{X} - \mu_1')^2 + \\
&\quad \frac{A_3}{3!}E(\bar{X} - \mu_1')^3 + \frac{A_4}{4!}E(\bar{X} - \mu_1')^4 \\
&= A_0 + \frac{A_2}{2!}\mu_2(\bar{X}) + \frac{A_3}{3!}\mu_3(\bar{X}) + \frac{A_4}{4!}\mu_4(\bar{X}) \\
&= A_0 + \frac{A_2\mu_2(X)}{2!n} + \frac{A_3\mu_3(X)}{3!n^2} + \\
&\quad \frac{A_4}{4!}\left(\frac{3\mu_2^2(X)}{n^2} + \frac{\mu_4(X) - 3\mu_2^2(X)}{n^3}\right) + \cdots \\
&= A_0 + \frac{A_2\mu_2(X)}{2!n} + \frac{1}{n^2}\left(\frac{A_3\mu_3(X)}{3!} + \frac{3A_4\mu_2^2(X)}{4!}\right) + O(n^{-3}) \qquad (2.43)
\end{aligned}$$

2.4.2.1 Expected value of $\hat{\mu}$ for the truncated normal distribution:

From equation (2.5), we have

$$\hat{\mu} = \bar{x} + \psi(\frac{c - \hat{\mu}}{\sigma}). \qquad (2.44)$$

Taking the derivative of $\hat{\mu}$ with respect to $\bar{x}$, we have

$$\frac{\partial\hat{\mu}}{\partial\bar{x}} = 1 + \frac{\partial\psi(\frac{c-\hat{\mu}}{\sigma})}{\partial\hat{\mu}}.\frac{\partial\hat{\mu}}{\partial\bar{x}}. \qquad (2.45)$$

Hence

$$\frac{\partial\hat{\mu}}{\partial\bar{x}} = \frac{1}{1 - \frac{\partial\psi(\frac{c-\hat{\mu}}{\sigma})}{\partial\hat{\mu}}} = \frac{1}{1 + \psi'(\hat{c})/\sigma}, \qquad (2.46)$$

so that

$$A_1 = \frac{\partial\hat{\mu}}{\partial\bar{x}}\,|_{\hat{\mu}=\mu} = \frac{1}{1 - \frac{\partial\psi(\frac{c-\mu}{\sigma})}{\partial\mu}} = \frac{1}{1 + \psi'(c')/\sigma} \quad , \qquad (2.47)$$

where

$$\psi'(c') = -c'\psi(c') - \psi^2(c')/\sigma. \qquad (2.48)$$

To find A_2, we have to take the derivative of $\frac{\partial \hat{\mu}}{\partial \bar{x}}$ with respect to $\bar{x}$. Hence we find

$$\frac{\partial^2 \hat{\mu}}{\partial \bar{x}^2} = \frac{\psi''(\frac{c-\hat{\mu}}{\sigma})}{\sigma^2[1+\psi'(\frac{c-\hat{\mu}}{\sigma})/\sigma]^3}. \tag{2.49}$$

Therefore

$$A_2 = \frac{\partial^2 \hat{\mu}}{\partial \bar{x}^2} \mid_{\hat{\mu}=\mu} = \frac{\psi''(c')}{\sigma^2[1+\psi'(c')/\sigma]^3}, \tag{2.50}$$

where

$$\begin{aligned} \psi''(c') &= -\psi(c') - c'\psi'(c') - 2\psi'(c')\psi(c')/\sigma \\ &= \psi(c')\left\{(c'^2-1) + 3c'\psi(c')/\sigma + 2[\psi(c')/\sigma]^2\right\}. \end{aligned} \tag{2.51}$$

To find A_3, we have to take the derivative of $\frac{\partial^2 \hat{\mu}}{\partial \bar{x}^2}$ with respect to $\bar{x}$. We can demonstrate that

$$\frac{\partial^3 \hat{\mu}}{\partial \bar{x}^3} = \frac{-\psi'''(\frac{c-\hat{\mu}}{\sigma})}{\sigma^3[1+\psi'(\frac{c-\hat{\mu}}{\sigma})/\sigma]^4} + 3\frac{\psi''^2(\frac{c-\hat{\mu}}{\sigma})}{\sigma^4[1+\psi'(\frac{c-\hat{\mu}}{\sigma})/\sigma]^5}. \tag{2.52}$$

Therefore

$$A_3 = \frac{\partial^3 \hat{\mu}}{\partial \bar{x}^3} \mid_{\hat{\mu}=\mu} = \frac{-\psi'''(c')}{\sigma^3[1+\psi'(c')/\sigma]^4} + 3\frac{\psi''^2(c')}{\sigma^4[1+\psi'(c')/\sigma]^5}, \tag{2.53}$$

where

$$\psi'''(c') = -2\psi'(c') - c'\psi''(c') - 2\psi'^2(c')/\sigma - 2\psi''(c')\psi(c')/\sigma. \tag{2.54}$$

Similarly differentiating $\frac{\partial^3 \hat{\mu}}{\partial \bar{x}^3}$ with respect to $\bar{x}$ gives

$$\frac{\partial^4 \hat{\mu}}{\partial \bar{x}^4} = \frac{\psi^{(iv)}(\frac{c-\hat{\mu}}{\sigma})}{\sigma^4[1+\psi'(\frac{c-\hat{\mu}}{\sigma})/\sigma]^5} - 10\frac{\psi''(\frac{c-\hat{\mu}}{\sigma})\psi'''(\frac{c-\hat{\mu}}{\sigma})}{\sigma^5[1+\psi'(\frac{c-\hat{\mu}}{\sigma})/\sigma]^5} + 15\frac{\psi''^3(\frac{c-\hat{\mu}}{\sigma})}{\sigma^6[1+\psi'(\frac{c-\hat{\mu}}{\sigma})/\sigma]^7}. \tag{2.55}$$

which, after substituting μ for $\hat{\mu}$, leads to

$$A_4 = \frac{\partial^4 \hat{\mu}}{\partial \bar{x}^3} \mid_{\hat{\mu}=\mu} = \frac{\psi^{(iv)}(c')}{\sigma^4[1+\psi'(c')/\sigma]^5} - 10\frac{\psi''(c')\psi'''(c')}{\sigma^5[1+\psi'(c')/\sigma]^6} + 15\frac{\psi''^3(c')}{\sigma^6[1+\psi'(c')/\sigma]^7}, \tag{2.56}$$

where

$$\psi^{(iv)}(c') = -3\psi''(c') - c'\psi'''(c') - 6\psi'(c')\psi''(c')/\sigma - 2\psi(c')\psi'''(c')/\sigma. \tag{2.57}$$

Substituting A_0, A_2 and μ_2(X) (the second moment of the truncated normal which we derived in Chapter 0, section (0.1.12)) in equation (2.43), we get

$$E(\hat{\mu}) = \mu + \frac{\psi''(c)}{2n[1+\psi'(c')/\sigma]^2} + O(n^{-2}) \tag{2.58}$$

which is identical to the Cox & Hinkley result derived in section (2.4.1) equation (2.32).

To obtain a more accurate expression for $E(\hat{\mu})$, we use $\mu_2(X)$ and $\mu_3(X)$ (the second and third moments of the truncated normal derived in Chapter 0, sections (0.1.12) and (0.1.22)) and A_1, A_2, A_3 and A_4 in equation (2.43). We then find the expected value of $\hat{\mu}$ to be

$$\begin{aligned} E(\hat{\mu}) &= \mu + \frac{\psi''(c)}{2n[1+\psi'(c')/\sigma]^2} + \frac{1}{n^2}\left\{\frac{\psi^{(iv)}(c')}{8[1+\psi'(c')/\sigma]^3} - \frac{13\psi''(c')\psi'''(c')}{12\sigma[1+\psi'(c')/\sigma]^4} + \right. \\ &\left. \frac{11\psi''^3(c')}{8\sigma^2[1+\psi'(c')/\sigma]^5}\right\} + O(n^{-3}). \end{aligned} \tag{2.59}$$

2.4.2.2 Variance of $\hat{\mu}$ for the truncated normal distribution:

The aim of this section is to find the variance of the maximum likelihood estimator of μ in the distribution (2.1), to make a comparison of this method with that of Cox & Hinkley and to compare the theoretical results in both cases with the simulation results.

According to Shenton & Bowman (1977), we have

$$\text{Var}(\hat{\mu}) = \frac{A_1^2\mu_2(X)}{n} + \frac{1}{n^2}\left[A_1A_2\mu_3(X) + (A_1A_3 + \frac{A_2^2}{2})\mu_2^2(X)\right] + O(n^{-3}) \tag{2.60}$$

Using $\mu_2(X)$ and $\mu_3(X)$ (the second and third moments of the truncated normal derived in Chapter 0, sections (0.1.12) and (0.1.22)) and A_1, A_2, A_3 and A_4 in equations (2.47), (2.50), (2.53) and (2.56), we find the variance of $\hat{\mu}$ to be:

$$\text{Var}(\hat{\mu}) = \frac{\sigma^2}{n[1+\psi'(c')/\sigma]} + \frac{1}{n^2}\left\{\frac{5\psi''^2(c')}{2[1+\psi'(c')/\sigma]^4} - \frac{\sigma\psi'''(c')}{[1+\psi'(c')/\sigma]^3}\right\} + O(n^{-3}). \tag{2.61}$$

in which the coefficients of $\frac{1}{n}$ and $\frac{1}{n^2}$ agree with their counterparts in Cox & Hinkley. A Fortran program was written to calculate the expected value and the standard deviation of the maximum likelihood estimator. A listing of the program is given in Appendix Program 6, the results obtained are shown in Table 2.3.

Table 2.3: The theoretical results for the expected value and standard deviation of the maximum likelihood estimator, of μ for different values of n and c, using the Shenton and Bowman methods. $E(\hat{\mu})$ is calculated in $O(n^{-2})$ and $\sigma(\hat{\mu})$ is calculated in $O(n^{-1})$.

$\mu = 0,\ \sigma = 1$

n	$c = -1.88$		$c = -1$		$c = 0$		$c = 1$	
	$E(\hat{\mu})$	$\sigma(\hat{\mu})$	$E(\hat{\mu})$	$\sigma(\hat{\mu})$	$E(\hat{\mu})$	$\sigma(\hat{\mu})$	$E(\hat{\mu})$	$\sigma(\hat{\mu})$
5	0.531	1.608	0.352	1.214	0.192	0.853	0.084	0.611
10	0.241	1.028	0.162	0.787	0.089	0.565	0.040	0.416
20	0.115	0.685	0.077	0.529	0.043	0.386	0.019	0.287
50	0.044	0.417	0.030	0.324	0.017	0.238	0.007	0.180
100	0.022	0.291	0.015	0.227	0.008	0.167	0.004	0.126

Concentrating in Tables 2.2 and 2.3, we can see that for fixed c, $E(\hat{\mu})$ in Table 2.3 is bigger than its counterpart in Table 2.2. The reason is that in this table $E(\hat{\mu})$ is calculated in $O(n^{-2})$ whereas in Table 2.2 is in $O(n^{-1})$. But in both theoretical methods $\sigma(\hat{\mu})$ is calculated in $O(n^{-1})$.

2.5 Simulation to estimate the mean when the variance is known:

The purpose of this section is to compare the results of a small simulation study with the theoretical results for $E(\hat{\mu})$ and $\sigma(\hat{\mu})$ to see whether the high order terms of the theoretical results are negligible or not.

In the study, $R = 10000$ samples were simulated for each value of the sample size $n = 5, 10, 20, 50, 100$ and each value of c shown in Table 2.4. In each case, the mean value of $\hat{\mu}$ was calculated.

We used the NAG routines G05DDF (0, 1), which generates random numbers from a standard normal distribution with mean zero and variance one, and G05CCF, which changes the seed of the random generator for each combination of n and c.

Program 7 (see Appendix) was used to carry out the calculations. The results are given in Table 2.4.

Table 2.4: The simulation results of the expected value and standard deviation of the maximum likelihood estimator, for different values of n and c when $R = 10000$

R	n	$c = -1.88$		$c = -1$		$c = 0$		$c = 1$	
		$E(\hat{\mu})$	$\sigma(\hat{\mu})$	$E(\hat{\mu})$	$\sigma(\hat{\mu})$	$E(\hat{\mu})$	$\sigma(\hat{\mu})$	$E(\hat{\mu})$	$\sigma(\hat{\mu})$
10000	5	0.537	1.589	0.381	1.354	0.199	0.912	0.076	0.621
	10	0.234	1.062	0.162	0.808	0.089	0.576	0.036	0.416
	20	0.105	0.687	0.084	0.536	0.043	0.387	0.023	0.290
	50	0.038	0.417	0.024	0.326	0.016	0.240	0.005	0.181
	100	0.025	0.290	0.015	0.228	0.009	0.168	0.005	0.126

The results for simulation, Shenton & Bowman and Cox & Hinkley are almost the same for different values of n. Moreover all theoretical values lie within 95% confidence interval obtained for the simulation.

2.6 Distribution of $\hat{\mu}$ when the variance is known:

In this section we investigate the distribution of $\hat{\mu}$, using Program 7 given in the Appendix, when sample size $n = 100$. The number of simulation runs $R = 10000$. Using the S-PLUS software, we can obtain the histogram, density, qqnorm and qqline plots for $\hat{\mu}$. The descriptions of these plots, from the S-PLUS manual (1993) are as follows:

Application of density:

Density plots are essentially smooth versions of histograms, which provide smooth estimates of population frequency or probability density curves. The kernel method is used to estimate the density function.

Application of qqnorm and qqline:

To check a hypothesized distribution is normal, use the function qqnorm, for example a plot from qqnorm that is bent up on the left and bent up on the right. Also the qqline function gives the highly robust straight line fit, which is not much influenced by outliers in other words this function fits and plots a line through a normal qqnorm.

To find how close the distribution of $\hat{\mu}$ is to the normal distribution, we used the tests for skewness and kurtosis, Senedecor and Cochran (1967, p. 86). The assumption of normality of $g_1(\hat{\mu})$ is accurate for $R \geq 150$. Also we know that, in very large samples, the measure of skewness $g_1(\hat{\mu})$ is defined by

$$g_1(\hat{\mu}) = \frac{m_3(\hat{\mu})}{m_2(\hat{\mu})\sqrt{m_2(\hat{\mu})}}$$

and the measure of kurtosis defined by

$$g_2(\hat{\mu}) = \frac{m_4(\hat{\mu})}{m_2^2(\hat{\mu})},$$

where $m_2(\hat{\mu})$, $m_3(\hat{\mu})$ and $m_4(\hat{\mu})$ are the second, third and fourth moments of $\hat{\mu}$. Also, if the sample comes from a normal distribution, with the sample size R, then measures of skewness and kurtosis are, respectively

$$g_1(\hat{\mu}) \sim N(0, \frac{6}{R})$$

and

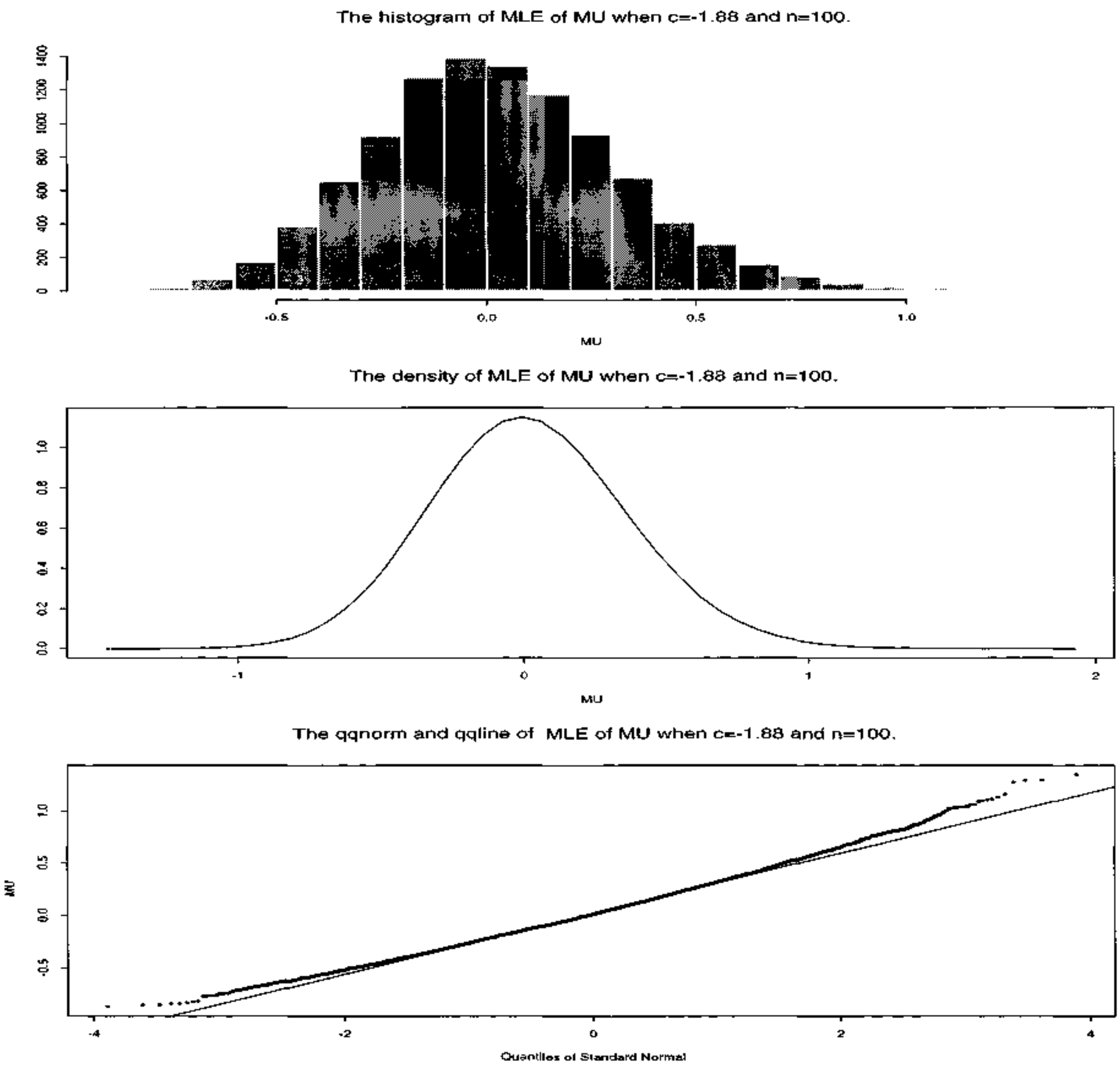
$$g_2(\hat{\mu}) \sim N(0, \frac{24}{R}).$$

The section (2.6.1) shows the distribution of $\hat{\mu}$ using plots obtained by the Program 8 (see Appendix) and the section (2.6.2) shows the skewness and kurtosis of the distribution of $\hat{\mu}$.

2.6.1 Graphs of data when $c = -1.88$:

Figure 2.4 are based on simulation runs $R = 10000$ sample of size $n = 100$.

Figure 2.4: The distribution of $\hat{\mu}$, when $c = -1.88$



2.6.2 Test for skewness and kurtosis:

In this section we give the results of testing the skewness and kurtosis of the distribution of $\hat{\mu}$ for different truncation points.

Using the above property, the moments, the measures skewness, and kurtosis and the corresponding z ratios were calculated for the various truncation points and are tabulated in Table 2.5.

Table 2.5: The moments and measures of skewness and kurtosis of the distribution of $\hat{\mu}$ for various truncation points when $n = 100$

c	-1.88	-1	0	1	3
$m_2(\hat{\mu})$	0.0856	0.0490	0.0288	0.0158	0.0103
$m_3(\hat{\mu})$	0.0075	0.0026	0.0010	0.0002	-0.00004
$m_4(\hat{\mu})$	0.0235	0.0076	0.0026	0.0007	0.0003
$g_1(\hat{\mu})$	0.3019	0.2430	0.2050	0.1010	-0.0040
$g_2(\hat{\mu})$	0.2135	0.1500	0.0979	0.0177	-0.0076
z_{sk}	12.322***	9.918***	8.367***	4.122***	-0.163
z_{ku}	4.366***	3.067**	2.002*	0.362	-0.155

($*$ p-value < 0.05, $**$ $0.05 < p$-value < 0.01 and $***$ p-value < 0.001)

From the above table, we observe large values for the skewness and kurtosis of the distribution of $\hat{\mu}$ for truncation points $c \leq 0$ and, as c increases, the values of the test statistics z_{sk} and z_{ku} decrease. We conclude that, there is significant evidence of both skewness and kurtosis for $c \leq 0$.

2.7 Conclusion:

Firstly, in Table 2.4 we can see that for each fixed value of c, the $E(\hat{\mu})$ decrease, as n increases. This means that the $E(\hat{\mu})$ has a relationship with the n and c.

Secondly, the comparison of Table 2.3 and Table 2.2, shows that the expected value of the maximum likelihood estimator $\hat{\mu}$ in Table 2.3 differs noticeably from Table 2.2. This shows that the second term on the right hand side equation (2.23) is very important. Therefore we used the Shenton & Bowman method to evaluate $E(\hat{\mu})$ up to $O(n^{-2})$ which was discussed in section (2.4.2). The comparison of Table 2.3 with the simulation results Table 2.4 shows that the expected value and standard deviation of $\hat{\mu}$ for the large sample sizes $n = 50$ and $n = 100$ are almost identical.

Finally, considering Tables 2.2-2.4 all the results obtained for various truncation points we can make the following comments.

1. In general as we increase the value of the truncation point, we see that the $E(\hat{\mu})$ and $\sigma(\hat{\mu})$ tend to the values for the full normal distribution i.e $E(\hat{\mu}) = \mu$ and $\sigma(\hat{\mu}) = \frac{\sigma}{\sqrt{n}}$ and this is of course in accordance with our expectations.

2. Since the simulation results, Table 2.4 agreeing with the Shenton & Bowman results, Table 2.3 we conclude that the higher order terms of equations (2.59) and (2.61) are negligible.

2.8 Likelihood equation when mean is known:

In this section, using the probability density function of the singly truncated normal random variable X given in equation (2.1), we find the maximum likelihood estimator of σ^2, when μ is known. Then the likelihood function is

$$L(\mathrm{x},\sigma^2) = (\frac{1}{2\pi\sigma^2})^{\frac{n}{2}} \frac{e^{-\sum_{i=1}^{n}(x_i-\mu)^2/2\sigma^2}}{[\Phi((c-\mu)/\sigma)]^n}.$$

Therefore the derivative of the natural logarithm of the likelihood function with respect to σ^2 is

$$\frac{\partial l(\mathrm{x},\sigma^2)}{\partial\sigma^2} = -\frac{n}{2\sigma^2} + \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^4} + \frac{n(c-\mu)\phi(\frac{c-\mu}{\sigma})}{2\sigma^3\Phi(\frac{c-\mu}{\sigma})}. \tag{2.62}$$

Since using $\bar{x}$ and s^2 for computation purposes, and letting $\hat{c}' = \frac{c-\mu}{\hat{\sigma}}$ and $\tau(c') = \frac{\phi(c')}{\Phi(c')}$ and using notation of section 2.2 we split equation (2.62) as:

$$\frac{\partial l(\mathrm{x},\sigma^2)}{\partial\sigma^2}\,|_{\sigma=\hat{\sigma}} = -\frac{n}{2\hat{\sigma}^2} + \frac{ns^2+n(\bar{x}-\mu)^2}{2\hat{\sigma}^4} + \frac{n\hat{c}'\tau(\hat{c}')}{2\hat{\sigma}^2} = 0, \tag{2.63}$$

which equation (2.63) can be written as:

$$-\hat{\sigma}^2 + s^2 + (\bar{x}-\mu)^2 + \hat{c}'\hat{\sigma}^2\tau(\hat{c}') = 0. \tag{2.64}$$

The algebraic solution of equation (2.64) is impossible. Therefore, in the following section we use the iterative methods to solve it.

Theorem 2.3 *For a fixed value of μ, the function $l(\mathrm{x},\sigma^2)$ has a local maximum.*

Proof: We prove this theorem for the two cases $c' > 0$ and $c' < 0$.

Since $l(\mathrm{x},\sigma^2)$ is continuous and differentiable, if we show that $l(\mathrm{x},\sigma^2) \to -\infty$ as $\sigma \to 0$ and $l(\mathrm{x},\sigma^2) \to -\infty$ as $\sigma \to \infty$, then we conclude that $l(\mathrm{x},\sigma^2)$ has a local maximum.

1. For $c' > 0$ $(\mu < c)$:

   (a) For $\sigma \to 0$:

Consider $\Phi(c') = \Phi(\frac{c-\mu}{\sigma}) \to 1$ as $\sigma \to 0$. Also $-\frac{n}{2}\ln(\sigma^2) \to -\infty$ and $-\frac{1}{2}\sum_{i=1}^{n}(\frac{x_i-\mu}{\sigma})^2 \to -\infty$ as $\sigma \to 0$. Therefore, implies that $l(\mathbf{x}, \sigma^2) \to -\infty$ as $\sigma \to 0$.

(b) For $\sigma \to \infty$:

Consider $\Phi(c') = \Phi(\frac{c-\mu}{\sigma}) \to 0.5$ as $\sigma \to \infty$. Then $-n\ln\Phi(c')$ and $-\frac{1}{2}\sum_{i=1}^{n}(\frac{x_i-\mu}{\sigma})^2$ are constant as $\sigma \to \infty$. But $-\frac{n}{2}\ln(\sigma^2) \to -\infty$ as $\sigma \to \infty$. Therefore, implies that $l(\mathbf{x}, \sigma^2) \to -\infty$ as $\sigma \to \infty$.

2. For $c' < 0$ $(\mu > c)$:

(a) For $\sigma \to 0$:

Using the note of section (2.2) $l(\mathbf{x}, \sigma^2)$ can be written as

$$\begin{aligned} l(\mathbf{x}, \sigma^2) &= -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - n\ln u(-c') \\ &- \frac{\sum_{i=1}^{n}[(c-x_i)(2\mu-c-x_i)]}{2\sigma^2}. \end{aligned}$$

Since $u(-c')$ is bounded, $-n\ln u(-c')$ is constant as $\sigma \to 0$. We know that $c > x_i$ for all i and $\mu > c$, then $(c - x_i) > 0$ and $(2\mu - c - x_i) > 0$, therefore $-\frac{\sum_{i=1}^{n}[(c-x_i)(2\mu-c-x_i)]}{2\sigma^2} \to -\infty$ as $\sigma \to 0$. Also $-\frac{n}{2}\ln(\sigma^2) \to -\infty$ as $\sigma \to 0$. Therefore $l(\mathbf{x}, \sigma^2) \to -\infty$ as $\sigma \to 0$.

(b) For $\sigma \to \infty$:

Again since $u(-c')$ is bounded $-n\ln u(-c')$ is constant. Also $c > x_i$ for all i and $\mu > c$, then $(c - x_i) > 0$ and $(2\mu - c - x_i) > 0$, therefore $-\frac{\sum_{i=1}^{n}[(c-x_i)(2\mu-c-x_i)]}{2\sigma^2}$ is constant as $\sigma \to \infty$. But $-\frac{n}{2}\ln(\sigma^2) \to -\infty$ as $\sigma \to \infty$. Therefore $l(\mathbf{x}, \sigma^2) \to -\infty$ as $\sigma \to \infty$ and the theorem is proved.

2.8.1 False position method:

The false position method which was described in section (2.3.2) is used to calculate $\hat{\sigma}^2$ from equation (2.64).

To find the ML estimate of σ, $(\hat{\sigma})$, we run the Fortran program (see Appendix Program 9), to solve the equation (2.64) numerically.

2.8.1.1 Estimates of standard deviation in data sets 1 and 2:

1. Using the data set 1 and letting $\mu = 0$ and $\varepsilon = 10^{-5}$, we obtain

$$\hat{\sigma} = 1.1250.$$

2. Using the data set 2 and letting $\mu = 0$ and $\varepsilon = 10^{-5}$, we obtain

$$\hat{\sigma} = 1.0266.$$

2.8.1.2 Estimates of standard deviation in ideal samples size 5 and 10:

1. Using the ideal sample size 5 and letting $\mu = 0$, when $c = -1.88$ and $\varepsilon = 10^{-5}$, after 14 iteration $\hat{\sigma}$ is estimated as

$$\hat{\sigma} = 0.9657.$$

2. Using the ideal sample size 10 and letting $\mu = 0$ and $c = -1.88$, $\varepsilon = 10^{-5}$, afterin 15 iteration $\hat{\sigma}$ is estimated as

$$\hat{\sigma} = 1.0009.$$

   On plotting the likelihood against σ when $\mu = 0$ and $c = -1.88$ for ideal samples of size 5 and 10 we get the Figures 2.5 and 2.6 (see Appendix Program 10).

The estimates of σ for different ideal sample sizes, constructed for different truncation points (see Chapter 1, section (1.4.2)), are given in Table 2.6.

Figure 2.5: Likelihood versus σ for ideal sample of size 5 ($\mu = 0, c = -1.88$)

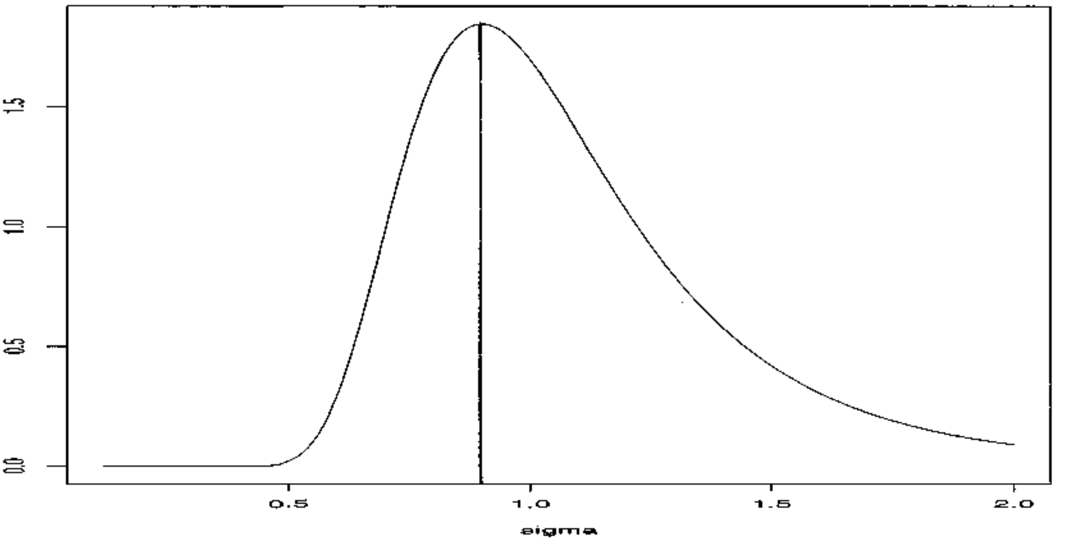


Figure 2.6: Likelihood versus σ for ideal sample of size 10 ($\mu = 0, c = -1.88$)

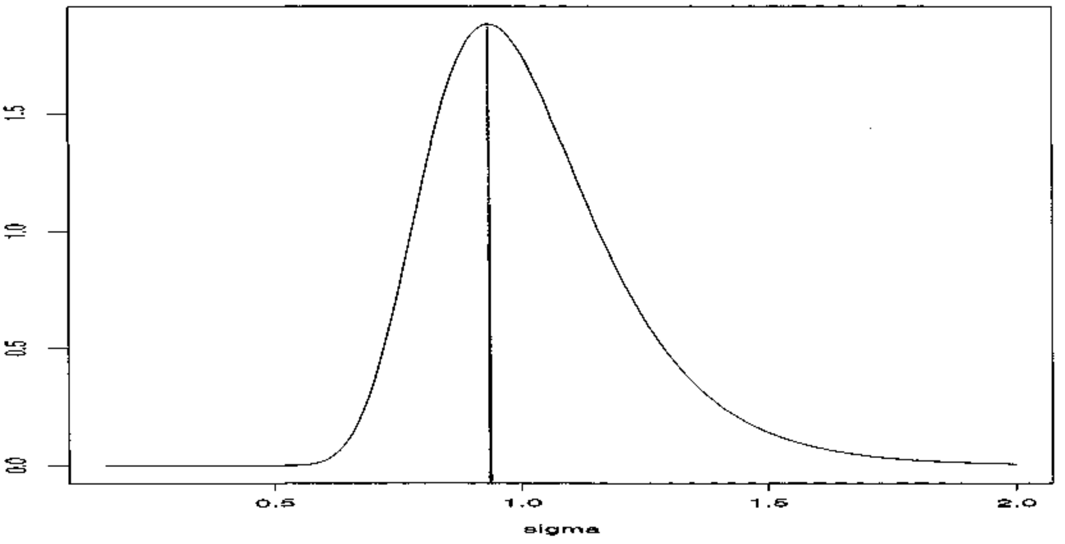


Table 2.6: The estimate of σ in ideal sample of size 5 and 10 for different truncation points

c	-1.88		-1		0		1		3	
n	5	10	5	10	5	10	5	10	5	10
$\bar{y}$	-2.2165	-2.234	-1.4657	-1.4858	-0.7332	-0.7549	-0.2360	-0.2517	-0.0022	-0.0024
$\hat{\sigma}$	0.9557	1.0009	0.9797	1.0158	0.9988	1.0373	0.9766	1.0258	0.8662	0.9134

As we can see from Table 2.6 the estimate of σ in almost every cell is close to the exact value of $\sigma = 1$. Also, for both ideal samples of size 5 and 10, $\hat{\sigma}$ increases as c increases up to $c = 0$. After this point, with increase in c, the value of $\hat{\sigma}$ decreases.

2.9 Theoretical results:

The aim of this section is to find the expected value and variance of the maximum likelihood estimator of σ^2 when μ is known. In a similar manner to section (2.4), we derive the formulae of Cox & Hinkley (1974), and Shenton & Bowman (1977).

2.9.1 Shenton & Bowman methods:

The aim of this section is to find the expected value and the variance of the maximum likelihood estimator of σ^2 by using the formula given in Shenton & Bowman (1977).

According to the factorization theorem given by Hogg & Craig (1970), since $L(\mathbf{x}, \sigma^2)$ can be written as the product of $k_1[t; \sigma^2]$ and $k_2(x_1, x_2, \ldots, x_n)$, it follows that T is a sufficient statistic for σ^2.

Assuming that

$$\hat{\sigma}^2 = \sum_{j=0}^{\infty} B_j[T - \mu(T)]^j / j! \quad , \tag{2.65}$$

where

$$T = \frac{\sum_{i=1}^{n}(X_i - \mu)^2}{n} \tag{2.66}$$

then it follows that

$$B_j = \frac{\partial^j \hat{\sigma}^2}{\partial T^j} \,|_{T=\mu(T), \hat{\sigma}^2 = \sigma^2} \tag{2.67}$$

for $j = 0, 1, 2, \ldots$. Note that $\hat{\sigma}^2 \to \sigma^2$ as $T \to \mu(T)$.

2.9.1.1 Expected value of $\hat{\sigma}^2$ for the truncated normal distribution:

Using the notations of sections (2.2) and (2.8), equation (2.64) can be written as

$$\hat{\sigma}^2 = T + \hat{c}' \hat{\sigma}^2 \tau(\hat{c}'). \tag{2.68}$$

Since, as $T \to \mu(T)$, $\hat{\sigma}^2 \to \sigma^2$, it can be shown that $B_0 = \sigma^2$.

Using repeated differentiation of the equation (2.68) B_1 and B_2, can be found. Now, using the following formulae, we can find the expected value of $\hat{\sigma}^2$ in the truncated normal distribution. According to the formula given in Shenton & Bowman (1977), we have

$$E(\hat{\sigma}^2) = \sigma^2 + \frac{B_2[\mu_2(X-\mu)^2]}{2!n} + \frac{1}{n^2}\left\{\frac{B_3[\mu_3(X-\mu)^2]}{3!} + \frac{3B_4[\mu_2(X-\mu)^2]^2}{4!}\right\} + \cdots. \quad (2.69)$$

In equation (2.68), taking the derivative of $\hat{\sigma}^2$ with respect to T, we have

$$\frac{\partial\hat{\sigma}^2}{\partial T} = 1 + \hat{c}'\left[\frac{1}{2}\tau(\hat{c}') + \frac{\partial\tau(\hat{c}')}{\partial\hat{\sigma}^2}\hat{\sigma}^2\right].\frac{\partial\hat{\sigma}^2}{\partial T} \quad (2.70)$$

Hence

$$\frac{\partial\hat{\sigma}^2}{\partial T} = \frac{1}{1 - \frac{\hat{c}'}{2}\tau(\hat{c}') + \frac{\hat{c}'^2}{2}\tau'(\hat{c}')}. \quad (2.71)$$

We then from (2.67) obtain

$$B_1 = \frac{1}{1 - \frac{c'}{2}\tau(c') + \frac{c'^2}{2}\tau'(c')} \quad ,$$

where

$$\tau'(c') = \frac{\partial\tau(c')}{\partial c'} = -c'\tau(c') - \tau^2(c'). \quad (2.72)$$

Let $D = 1 - \frac{c'}{2}\tau(c') + \frac{c'^2}{2}\tau'(c')$, then we can express B_1 as

$$B_1 = \frac{1}{D}. \quad (2.73)$$

To find B_2, we have to take the derivative of $\frac{\partial\hat{\sigma}^2}{\partial T}$ with respect to T. From (2.71) we find

$$\frac{\partial^2\hat{\sigma}^2}{\partial T^2} = \frac{\hat{c}'\left[\hat{c}'^2\tau''(\hat{c}') + \hat{c}'\tau'(\hat{c}') - \tau(\hat{c}')\right]}{4\hat{\sigma}^2\left[1 - \frac{\hat{c}'}{2}\tau(\hat{c}') + \frac{\hat{c}'^2}{2}\tau'(\hat{c}')\right]^3}. \quad (2.74)$$

Therefore

$$\begin{aligned} B_2 = \frac{\partial^2\hat{\sigma}^2}{\partial T^2}\Big|_{\hat{\sigma}^2=\sigma^2} &= \frac{c'\left[c'^2\tau''(c') + c'\tau'(c') - \tau(c')\right]}{4\sigma^2\left[1 - \frac{c'}{2}\tau(c') + \frac{c'^2}{2}\tau'(c')\right]^3} \\ &= \frac{2D + c'^3\tau''(c') - 2}{4\sigma^2 D^3}, \end{aligned} \quad (2.75)$$

where

$$\tau''(c') = -\tau(c') - c'\tau'(c') - 2\tau(c')\tau'(c'). \qquad (2.76)$$

To find B_3, we have to take the derivative of $\frac{\partial^2\hat{\sigma}^2}{\partial T^2}$ with respect to T. Hence we can demonstrate that

$$\begin{aligned} \frac{\partial^3\hat{\sigma}^2}{\partial T^3} &= -\frac{\hat{c}'^4\tau'''(\hat{c}') + 6\hat{c}'^3\tau''(\hat{c}') + 3\hat{c}'^2\tau'(\hat{c}') - 3\hat{c}'\tau(\hat{c}')}{8\hat{\sigma}^4\left[1 - \frac{\hat{c}'}{2}\tau(\hat{c}') + \frac{\hat{c}'^2}{2}\tau'(\hat{c}')\right]^4} \\ &+ \frac{3\left[\hat{c}'^3\tau''(\hat{c}') + \hat{c}'^2\tau'(\hat{c}') - \hat{c}'\tau(\hat{c}')\right]^2}{16\hat{\sigma}^4\left[1 - \frac{\hat{c}'}{2}\tau(\hat{c}') + \frac{\hat{c}'^2}{2}\tau'(\hat{c}')\right]^5}. \end{aligned} \qquad (2.77)$$

Therefore

$$\begin{aligned} B_3 = \frac{\partial^3\hat{\sigma}^2}{\partial T^3}\Big|_{\hat{\sigma}^2=\sigma^2} &= -\frac{c'^4\tau'''(c') + 6c'^3\tau''(c') + 3c'^2\tau'(c') - 3c'\tau(c')}{8\sigma^4 D^4} \\ &+ \frac{3\left[c'^3\tau''(c') + c'^2\tau'(c') - c'\tau(c')\right]^2}{16\sigma^4 D^5} \\ &= -\frac{c'^4\tau'''(c') + 6}{8\sigma^4 D^4} + \frac{3\left(c'^3\tau''(c') - 2\right)^2}{16\sigma^4 D^5} \end{aligned} \qquad (2.78)$$

where

$$\tau'''(c') = -2\tau'(c') - c'\tau''(c') - 2\tau'^2(c') - 2\tau(c')\tau''(c'). \qquad (2.79)$$

Similarly differentiating $\frac{\partial^3\hat{\sigma}^2}{\partial T^3}$ with respect to T gives B_4, after substituting σ^2 for $\hat{\sigma}^2$. We find

$$\begin{aligned} B_4 = \frac{\partial^4\hat{\sigma}^2}{\partial T^4}\Big|_{\hat{\sigma}^2=\sigma^2} &= \frac{c'^4[c'\tau^{(iv)}(c') + 4\tau'''(c')]}{16\sigma^6 D^5} \\ &- \frac{5[c'^3\tau''(c') - 2][2c'^4\tau'''(c') + 3c'^3\tau''(c') + 6]}{32\sigma^6 D^6} \\ &+ \frac{15[c'^3\tau''(c') - 2]^3}{64\sigma^6 D^7}, \end{aligned} \qquad (2.80)$$

where

$$\tau^{(iv)}(c') = -3\tau''(c') - c'\tau'''(c') - 6\tau'(c')\tau''(c') - 2\tau(c')\tau'''(c'). \qquad (2.81)$$

Now we find the second and third moments of the random variable $(X - \mu)^2$.

In order to find the second and third moments of $(X-\mu)^2$, it is easiest to find the moment generating function of $(\frac{X-\mu}{\sigma})^2$.

First, as defined in general in Chapter 0 (0.54), for variable $a < X < b$, we can obtain

$$\begin{aligned} M_{(\frac{X-\mu}{\sigma})^2}(t) &= E(e^{t(\frac{X-\mu}{\sigma})^2}) \\ &= \int_a^b e^{t(\frac{x-\mu}{\sigma})^2} \frac{e^{\frac{-(x-\mu)^2}{2\sigma^2}}}{\sigma\delta_\phi\sqrt{2\pi}} dx \\ &= \int_a^b e^{(\frac{x-\mu}{\sigma})^2(t-1/2)} \frac{1}{\sigma\delta_\phi\sqrt{2\pi}} dx \\ &= \frac{\Phi(b'\sqrt{1-2t}) - \Phi(a'\sqrt{1-2t})}{\sqrt{1-2t}(\Phi(b') - \Phi(a'))}. \end{aligned} \tag{2.82}$$

From equation (2.82), as $a' \to -\infty$ and $b' = c'$, we get the moment generating function of the random variable $(\frac{X-\mu}{\sigma})^2$ truncated at c' as

$$M_{(\frac{X-\mu}{\sigma})^2}(t) = \frac{\Phi(c'\sqrt{1-2t})}{\sqrt{1-2t}\Phi(c')}. \tag{2.83}$$

Note that, for $c' = \infty$, $M_{(\frac{X-\mu}{\sigma})^2}(t)$ is the moment generating function of a $\chi^2(1)$ variate.

Using the Maple software we found the first, second and third derivatives of equation (2.83) (see Appendix, Program 11), and substituting $t = 0$ we obtain the first, second and third moments of the random variable $(\frac{X-\mu}{\sigma})^2$ about the origin.

So

$$\mu(\frac{X-\mu}{\sigma})^2 = 1 - c'\tau(c'), \tag{2.84}$$

$$\mu_2'(\frac{X-\mu}{\sigma})^2 = 3 - c'\left(3 + c'^2\right)\tau(c'), \tag{2.85}$$

$$\mu_3'(\frac{X-\mu}{\sigma})^2 = 15 - c'[15 + 5c'^2 + c'^4\tau(c')]. \tag{2.86}$$

Consequently, the moments of $(X - \mu)^2$ about its mean are

$$\mu_2(X-\mu)^2 = E\left[(X-\mu)^2 - E(X-\mu)^2\right]^2$$

$$
\begin{aligned}
&= \sigma^4 \left\{2 - c'\tau(c')\left[1 + c'^2 - c'\tau(c')\right]\right\} \\
&= 2\sigma^4 \left[1 - \frac{c'}{2}\tau(c') + \frac{c'^2}{2}\tau'(c')\right] = 2\sigma^4 D
\end{aligned}
\tag{2.87}
$$

and

$$
\begin{aligned}
\mu_3(X-\mu)^2 &= E\left[(X-\mu)^2 - E(X-\mu)^2\right]^3 \\
&= \sigma^6\left[8 - c'(3 + 2c'^2 + c'^4)\tau(c') - 3c'^2(1 + c'^2)\tau^2(c') - 2c'^3\tau(c')\right]. \\
&= 8\sigma^6\left[1 + \frac{-c'^3\tau''(c') + 3c'^2\tau'(c') - 3c'\tau(c')}{8}\right] \\
&= 8\sigma^6\left[1 - \frac{c'^3\tau''(c') - 6D + 6}{8}\right].
\end{aligned}
\tag{2.88}
$$

By substituting B_2 and $\mu_2(X-\mu)^2$ from equations (2.75) and (2.87) into equation (2.69) we obtain the expected value of $\hat{\sigma}^2$ up to $O(n^{-1})$ as

$$
E(\hat{\sigma}^2) = \sigma^2 + \frac{\sigma^2}{n}\left[\frac{c'^3\tau''(c') + 2D - 2}{4D^2}\right] + O(n^{-2}). \tag{2.89}
$$

By substituting $\mu_2(X-\mu)^2$, $\mu_3(X-\mu)^2$, B_3 and B_4 respectively from equations (2.87), (2.88), (2.78) and (2.80) in second term of equation (2.69), we obtain

$$
\begin{aligned}
E(\hat{\sigma}^2) \;=\; & \sigma^2 + \frac{\sigma^2}{n}\left[\frac{c'^3\tau''(c') + 2D - 2}{4D^2}\right] + \\
& \frac{1}{n^2}\left\{\frac{\sigma^2}{3!}\left[-\frac{c'^4\tau'''(c') + 6}{D^4} + \frac{3\left[c'^3\tau''(c') - 2\right]^2}{2D^5}\right]\left[1 - \frac{c'^3\tau''(c') - 6D + 6}{8}\right] + \right. \\
& \frac{3}{4!}\left[-\frac{c'^4\left[c'\tau^{(iv)}(c') + 4\tau'''(c')\right]}{16\sigma^6 D^3} - \frac{5\left[c'^3\tau''(c') - 2\right]\left[2c'^4\tau'''(c') + 3c'^3\tau''(c') + 6\right]}{32\sigma^6 D^4} + \right. \\
& \left.\left. \frac{15\left[c'^3\tau''(c') - 2\right]^3}{64\sigma^6 D^5}\right]\right\} + O(n^{-3}).
\end{aligned}
\tag{2.90}
$$

2.9.1.2 Variance of $\hat{\sigma}^2$ for the truncated normal distribution:

The aim of this section is to find the variance of the maximum likelihood estimator of σ^2 in the distribution (2.1) to make a comparison of the theoretical results with the simulation results.

According to Shenton & Bowman (1977), we have

$$\mathrm{Var}(\hat{\sigma}^2) = \frac{1}{n}[B_1^2\mu_2(X-\mu)^2]+\frac{1}{n^2}[B_1B_2\mu_3(X-\mu)^2+(B_1B_3+\frac{B_2^2}{2})\mu_2^2(X-\mu)^2]+O(n^{-3}) \quad (2.91)$$

Substituting $\mu_2(X-\mu)^2$, $\mu_3(X-\mu)^2$ (the second and third moments of $(X-\mu)^2$), B_1, B_2, B_3 and B_4 respectively from equations (2.87), (2.88), (2.73), (2.75), (2.78) and (2.80) into equation (2.91), we find the variance of $\hat{\sigma}^2$ up to $O(n^{-2})$. Since

$$B_1B_2\mu_3(X-\mu)^2 = \sigma^4\left\{-\frac{[c'^3\tau''(c')-2]^2}{4D^4}+\frac{c'^3\tau''(c')-2}{D^3}+\frac{3}{D^2}\right\} \quad (2.92)$$

and

$$(B_1B_3+\frac{B_2^2}{2})\mu_2^2(X-\mu)^2 = \sigma^4\left\{\frac{5[c'^3\tau''(c')-2]^2}{8D^4}+\frac{-c'^4\tau'''(c')+3c'^3\tau''(c')-12}{2D^3}+\frac{1}{2D^2}\right\} \quad (2.93)$$

therefore variance of $\hat{\sigma}^2$ becomes

$$\begin{aligned}\mathrm{Var}(\hat{\sigma}^2) &= \frac{2\sigma^4}{nD}+\frac{\sigma^4}{n^2}\left\{\frac{5[c'^3\tau''(c')-2]^2}{8D^4}+\frac{-c'^4\tau'''(c')+3c'^3\tau''(c')-12}{2D^3}+\frac{7}{2D^2}\right\}\\ &+ O(n^{-3}) \quad (2.94)\end{aligned}$$

A Fortran program was written to calculate the expected value and the variance of the maximum likelihood estimator of σ^2 in $O(n^{-1})$ and $O(n^{-2})$. Program 12, given in the Appendix, evaluated results for $E(\hat{\sigma}^2)$ up to order n^{-1} and n^{-2} and for $\sigma(\hat{\sigma}^2)$ up to order $n^{-\frac{1}{2}}$ and n^{-1}. The equations (2.89), (2.90) and (2.94) were used to calculate the results presented in Table 2.7.

Table 2.7: The first and second order approximations of the expected value and standard deviation* of the maximum likelihood estimator of σ^2, for different values of n and c, when $\mu = 0$

n	order	$c = -1.88$		$c = -1$		$c = 0$		$c = 1$	
		$E(\hat{\sigma}^2)$	$\sigma(\hat{\sigma}^2)$	$E(\hat{\sigma}^2)$	$\sigma(\hat{\sigma}^2)$	$E(\hat{\sigma}^2)$	$\sigma(\hat{\sigma}^2)$	$E(\hat{\sigma}^2)$	$\sigma(\hat{\sigma}^2)$
5	first	1.0146	0.5033	1.0164	0.5419	1.0000	0.6325	0.9598	0.7721
	second	1.0133	0.5128	1.0140	0.5507	1.0000	0.6325	0.9737	0.7277
10	first	1.0073	0.3559	1.0082	0.3832	1.0000	0.4472	0.9799	0.5459
	second	1.0070	0.3593	1.0076	0.3863	1.0000	0.4472	0.9834	0.5305
20	first	1.0037	0.2516	1.0041	0.2710	1.0000	0.3162	0.9899	0.3860
	second	1.0036	0.2528	1.0039	0.2721	1.0000	0.3162	0.9908	0.3806
50	first	1.0015	0.1592	1.0016	0.1714	1.0000	0.2000	0.9960	0.2442
	second	1.0015	0.1595	1.0016	0.1716	1.0000	0.2000	0.9961	0.2428
100	first	1.0007	0.1125	1.0008	0.1212	1.0000	0.1414	0.9980	0.1726
	second	1.0007	0.1126	1.0008	0.1213	1.0000	0.1414	0.9980	0.1722

[* In this table first and second are used for the calculation of $E(\hat{\sigma}^2)$ and $\mathrm{Var}(\hat{\sigma}^2)$. Therefore $\sigma(\hat{\sigma}^2)$ is calculated in $O(n^{-\frac{1}{2}})$ and $O(n^{-1})$.]

As we can see from Table 2.7 the expected value of $\hat{\sigma}^2$ in almost every cell is very close to the exact value of σ^2, but it is interesting that for $c < 0$ the $E(\hat{\sigma}^2) > \sigma^2$ whereas for $c > 0$ the $E(\hat{\sigma}^2) < \sigma^2$. It also shows a very important fact that when we choose $c = 0$, then half of the normal distribution is considered and the $E(\hat{\sigma}^2)$ is equal to the exact value of the normal distribution. Considering the standard deviation of $\hat{\sigma}^2$ we realize that they are related to the truncation points as well.

2.9.2 Cox & Hinkley methods:

In this section we derive the formulae for σ^2 using the method described in section (2.4.1). For convenience , we set $\gamma = \sigma^2$.

1. Let
$$S_j(\gamma) = \frac{\partial \ln f(x_j, \gamma)}{\partial \gamma}.$$
2. Let
$$S_.(\gamma) = \frac{\partial \ln L(\mathbf{x}, \gamma)}{\partial \gamma}.$$
3. Let
$$I(\gamma) = [S(\gamma)]^2 = -S'(\gamma).$$
4. Let
$$i(\gamma) = E[S(\gamma)]^2 = E[-S'(\gamma)].$$
5. Let
$$i_.(\gamma) = E[S_.(\gamma)]^2 = E[-S_.'(\gamma)] = ni(\gamma). \tag{2.95}$$
6. Let
$$\kappa_{ij}(\gamma) = E\left\{[S(\gamma)]^i[S'(\gamma) + i(\gamma)]^j\right\}.$$

2.9.2.1 Expected value of $\hat{\gamma}$ for the truncated normal distribution:

The aim of this section is to find the expected value of the maximum likelihood estimator of γ in the distribution given at (2.1). We know that the score of γ for a single observation, when $\mu = 0$, is

$$\begin{aligned} S(\gamma) &= \frac{X^2 - \gamma + c\sqrt{\gamma}\tau(c/\sqrt{\gamma})}{2\gamma^2} \\ &= -\frac{1}{2\gamma} + \frac{X^2}{2\gamma^2} + \frac{c'\tau(c')}{2\gamma} \end{aligned}$$

$$= \frac{1}{2\gamma^2}[X^2 - \gamma + c'\gamma\tau(c')]. \tag{2.96}$$

Since we know that $E[S(\gamma)] = E[S_.(\gamma)] = 0$, we can find

$$E(X^2) = \gamma - c'\gamma\tau(c'). \tag{2.97}$$

To use the Cox & Hinkley formula for the expected value and variance of $\hat{\gamma}$, we have to find the following expressions.

$$i(\gamma) = E[S(\gamma)]^2 = E[-S'(\gamma)], \tag{2.98}$$

$$\kappa_{11}(\gamma) = E\{[S(\gamma)][S'(\gamma) + i(\gamma)]\}, \tag{2.99}$$

$$\kappa_{30}(\gamma) = E\{[S(\gamma)]^3[S'(\gamma) + i(\gamma)]^0\} = E[S(\gamma)]^3. \tag{2.100}$$

As in equation (2.20), we can write

$$E[S(\gamma)]^3 = -3E[S(\gamma)S'(\gamma)] - E[S''(\gamma)]. \tag{2.101}$$

Therefore, we have to find $S'(\gamma)$ and $S''(\gamma)$.

Taking the derivative of $S(\gamma)$ with respect to γ, we have

$$S'(\gamma) = \frac{1}{2\gamma^2} - \frac{X^2}{\gamma^3} - \frac{c'^2\tau'(c')}{4\gamma^2} - \frac{3c'\tau(c')}{4\gamma^2}, \tag{2.102}$$

and taking derivative of $S'(\gamma)$ with respect to γ we obtain

$$S''(\gamma) = -\frac{1}{\gamma^3} + \frac{3X^2}{\gamma^4} + \frac{c'^3\tau''(c')}{8\gamma^3} + \frac{9c'^2\tau'(c')}{8\gamma^3} + \frac{15c'\tau(c')}{8\gamma^3}. \tag{2.103}$$

Taking the expected value of $S''(\gamma)$, and substituting $E(X^2)$ from equation (2.97), we obtain

$$E[S''(\gamma)] = \frac{1}{\gamma^3}\left\{2 + \frac{c'}{8}[-9\tau(c') + 9c'\tau'(c') + c'^2\tau''(c')]\right\}. \tag{2.104}$$

Now using (2.98) and (2.102), we obtain

$$\begin{aligned} i(\gamma) &= E[-S'(\gamma)] \\ &= E\left[-\frac{1}{2\gamma^2} + \frac{X^2}{\gamma^3} + \frac{c'^2\tau'(c')}{4\gamma^2} + \frac{3c'\tau(c')}{4\gamma^2}\right] \\ &= \frac{1}{2\gamma^2}\left[1 - \frac{c'\tau(c')}{2} + \frac{c'^2\tau'(c')}{2}\right]. \end{aligned} \tag{2.105}$$

Using the notation $1 - \frac{c'\tau(c')}{2} + \frac{c'^2\tau'(c')}{2} = D$, as before we have

$$i(\gamma) = \frac{D}{2\gamma^2}. \tag{2.106}$$

To find the $\kappa_{11}(\gamma)$, we have to find $S'(\gamma) + i(\gamma)$. From equation (2.97), we obtain

$$\begin{aligned} S'(\gamma) + i(\gamma) &= -\frac{1}{\gamma^3}[X^2 - \gamma + c'\gamma\tau(c')] \\ &= -\frac{1}{\gamma^3}[X^2 - E(X^2)]. \end{aligned} \tag{2.107}$$

Now using equation (2.99) we obtain

$$\begin{aligned} \kappa_{11}(\gamma) &= E\{[S(\gamma)][S'(\gamma) + i(\gamma)]\} \\ &= E\left\{\left[\frac{1}{2\gamma^2}[X^2 - E(X^2)]\right]\left[-\frac{1}{\gamma^3}[X^2 - E(X^2)]\right]\right\} \\ &= -\frac{1}{2\gamma^5}\mathrm{Var}(X^2). \end{aligned} \tag{2.108}$$

By using $\sigma^2 = \gamma$, and noting from section (2.9.1) that the second moment of $(X-\mu)^2$ about its mean is

$$\begin{aligned} \mu_2(X-\mu)^2 = \mathrm{Var}(X-\mu)^2 &= 2\gamma^2\left[1 - \frac{c'\tau(c')}{2} + \frac{c'^2\tau'(c')}{2}\right] \\ &= 2\gamma^2 D, \end{aligned} \tag{2.109}$$

we have, when $\mu = 0$,

$$\begin{aligned} \mathrm{Var}(X^2) &= 2\gamma^2\left[1 - \frac{c'\tau(c')}{2} + \frac{c'^2\tau'(c')}{2}\right] \\ &= 2\gamma^2 D. \end{aligned} \tag{2.110}$$

Substituting $\mathrm{Var}(X^2)$ from equation (2.110) into equation (2.108), we have

$$\kappa_{11}(\gamma) = -\frac{1}{\gamma^3}D. \tag{2.111}$$

Now we want to find the $\kappa_{30}(\gamma)$. To use equation (2.101) we must find the value of $E[S(\gamma)S'(\gamma)]$.

$$\begin{aligned} E[S(\gamma)S'(\gamma)] &= E\left\{\left[\frac{1}{2\gamma^2}(X^2-\gamma+c'\gamma\tau(c'))\right]\left[\frac{1}{2\gamma^3}-\frac{X^2}{\gamma^3}-\frac{c'^2\tau'(c')}{4\gamma^2}-\frac{3c'\tau(c')}{4\gamma^2}\right]\right\} \\ &= -\frac{1}{2\gamma^5}E\left\{\left[X^2-\gamma+c'\gamma\tau(c')\right]\left[X^2-\frac{\gamma}{2}+\frac{c'^2\gamma\tau(c')}{4}+\frac{c'\gamma\tau(c')}{4}\right]\right\} \\ &= -\frac{1}{2\gamma^5}\mathrm{Var}(X^2) \\ &= -\frac{1}{\gamma^3}D \end{aligned} \tag{2.112}$$

Now substituting $E[S(\gamma)S'(\gamma)]$ and $E[S''(\gamma)]$ from equation (2.112) and (2.104) into (2.101), we obtain

$$\begin{aligned} E[S(\gamma)]^3 &= -3E[S(\gamma)S'(\gamma)] - E[S''(\gamma)] \\ &= \frac{3}{\gamma^3}\left\{\left[1-\frac{c'\tau(c')}{2}+\frac{c'^2\tau'(c')}{2}\right]\right\} \\ &- \frac{1}{\gamma^3}\left\{2+\frac{c'}{8}[-9\tau(c')+9c'\tau'(c')-c'^2\tau''(c')]\right\} \\ &= \frac{1}{\gamma^3}\left\{1+\frac{c'}{8}[-3\tau(c')+3c'\tau'(c')-c'^2\tau''(c')]\right\}. \end{aligned} \tag{2.113}$$

Further we find

$$\kappa_{30}(\gamma)+\kappa_{11}(\gamma) = \frac{c'}{8\gamma^3}[\tau(c')-c'\tau'(c')-c'^2\tau''(c')]. \tag{2.114}$$

To find the bias term in the expected value of $\hat{\gamma}$ we have to use the following formula:

$$b(\gamma) = -\frac{\kappa_{30}(\gamma)+\kappa_{11}(\gamma)}{2i^2(\gamma)}. \tag{2.115}$$

Substituting $\kappa_{30}(\gamma)+\kappa_{11}(\gamma)$ and $i^2(\gamma)$ from equations (2.114) and (2.106) into equation (2.115) we obtain

$$b(\gamma) = \frac{\gamma c'[c'^2\tau''(c')+c'\tau'(c')-\tau(c')]}{4[1-\frac{c'\tau(c')}{2}+\frac{c'^2\tau'(c')}{2}]^2}. \tag{2.116}$$

Therefore

$$E(\hat{\gamma}) = \gamma + \frac{\gamma c'[c'^2\tau''(c') + c'\tau'(c') - \tau(c')]}{4n\left[1 - \frac{c'\tau(c')}{2} + \frac{c'^2\tau'(c')}{2}\right]^2} + O(n^{-2})$$

$$= \gamma + \frac{\gamma}{n}\left[\frac{c'^3\tau''(c') + 2D - 2}{4D^2}\right] + O(n^{-2}), \qquad (2.117)$$

which is identical to the formula for $E(\hat{\sigma}^2)$ derived by the Shenton & Bowman's method , see equation (2.89) in section (2.9.1.1).

2.9.3 Variance of $\hat{\gamma}$ for the truncated normal distribution:

The aim of this section is to find the variance of the maximum likelihood estimator of γ in the distribution given in equation (2.1). According to Cox & Hinkley (1974), we know that

$$\text{Var}(\hat{\gamma}) = \frac{1}{ni(\gamma)} + \frac{2b'(\gamma)}{n^2 i(\gamma)} + \frac{2[\kappa_{20}(\gamma)\kappa_{02}(\gamma) - \kappa_{11}^2(\gamma)] + [\kappa_{11}(\gamma) + \kappa_{30}(\gamma)]^2}{2n^2 i^4(\gamma)} + O(n^{-3}) \qquad (2.118)$$

In the expression for $b(\gamma)$, let $c'^3\tau''(c') + c'^2\tau'(c') - c'\tau(c') = M$. Then we have

$$b(\gamma) = \frac{\gamma M}{4D^2}. \qquad (2.119)$$

Taking the derivative of $b(\gamma)$ with respect to γ, we have

$$b'(\gamma) = \frac{1}{4}\left\{\left[M + (\frac{\partial M}{\partial c'} \times \frac{\partial c'}{\partial \gamma})\gamma\right]/D^2 - \left[\left(\frac{\partial(4D^2)}{\partial c'} \times \frac{\partial c'}{\partial \gamma}\right)\gamma M\right]/4D^4\right\}. \quad (2.120)$$

Now

$$\frac{\partial M}{\partial c'} = c'^3\tau'''(c') + 4c'^2\tau''(c') + c'\tau'(c') - \tau(c'), \quad (2.121)$$

$$\begin{aligned}\frac{\partial(4D^2)}{\partial c'} &= 4D[c'^2\tau''(c') + c'\tau'(c') - \tau(c')] \\ &= \frac{4DM}{c'} \end{aligned} \quad (2.122)$$

and

$$\frac{\partial c'}{\partial \gamma} = -\frac{c'}{2\gamma}. \quad (2.123)$$

Substituting these last three expressions into equation (2.120) we obtain

$$b'(\gamma) = \frac{1}{8D^2}[-c'^4\tau'''(c') - 2c'^3\tau''(c') + c'^2\tau'(c') - c'\tau(c')] + \frac{M^2}{8D^3}. \quad (2.124)$$

According to Cox & Hinkley, the first part of the coefficient of $\frac{1}{n^2}$ in $\mathrm{Var}(\hat{\gamma})$ is

$$\frac{2b'(\gamma)}{i(\gamma)} = \frac{\gamma^2}{2D^4}\{D[-c'^4\tau'''(c') - 2c'^3\tau''(c') + c'^2\tau'(c') - c'\tau(c')] + M^2\}. \quad (2.125)$$

To find the second part of the coefficient of $\frac{1}{n^2}$ in $\mathrm{Var}(\hat{\gamma})$, we have to find $\kappa_{20}(\gamma)$ and $\kappa_{02}(\gamma)$.

We know that

$$\begin{aligned}\kappa_{20}(\gamma) &= E\{[S(\gamma)]^2[S'(\gamma) + i(\gamma)]^0\} = E[S(\gamma)]^2 \\ &= E[-S'(\gamma)] \\ &= i(\gamma) \\ &= \frac{D}{2\gamma^2} \end{aligned} \quad (2.126)$$

and

$$\begin{aligned}\kappa_{02}(\gamma) &= E\{[S(\gamma)]^0[S'(\gamma)+i(\gamma)]^2\} \\ &= E[S'(\gamma)+i(\gamma)]^2 \\ &= E\left\{-\frac{1}{\gamma^3}[X^2-\gamma+c'\gamma\tau(c')]\right\}^2 \\ &= \frac{1}{\gamma^6}E[X^2-E(X^2)]^2 \\ &= \frac{1}{\gamma^6}\mathrm{Var}(X^2). \end{aligned} \quad (2.127)$$

Hence, on substituting $\mathrm{Var}(X^2)$ from equation (2.110) into equation (2.127) we obtain

$$\kappa_{02}(\gamma) = \frac{2D}{\gamma^4}. \quad (2.128)$$

Therefore, using equations (2.126), (2.128) and (2.111) we have

$$2\{\kappa_{20}(\gamma)\kappa_{02}(\gamma) - [\kappa_{11}(\gamma)]^2\} = 2\left\{\frac{D}{2\gamma^2}\frac{2D}{\gamma^4} - \left[-\frac{D}{\gamma^3}\right]^2\right\} = 0. \quad (2.129)$$

Moreover, from equation (2.114) we know that

$$\begin{aligned}[\kappa_{11}(\gamma)+\kappa_{30}(\gamma)]^2 &= \frac{c'^2}{64\gamma^6}[\tau(c') - c'\tau'(c') - c'^2\tau''(c')]^2 \\ &= \frac{M^2}{64\gamma^6}. \end{aligned} \quad (2.130)$$

Using equation (2.129), (2.130) and (2.106) we obtain the second term of the coefficient of $\frac{1}{n^2}$ in $\mathrm{Var}(\hat{\gamma})$ as

$$\begin{aligned}\frac{2\{\kappa_{20}(\gamma)\kappa_{02}(\gamma) - [\kappa_{11}(\gamma)]^2\} + [\kappa_{11}(\gamma)+\kappa_{30}(\gamma)]^2}{2i^4(\gamma)} &= \frac{\gamma^2 c'^2[\tau(c') - c'\tau'(c') - c'^2\tau''(c')]^2}{8D^4} \\ &= \frac{\gamma^2 M^2}{8D^4} \end{aligned} \quad (2.131)$$

Substituting equations (2.106), (2.125) and (2.131) into equation (2.118) and again using $\gamma = \sigma^2$, we obtain

$$\begin{aligned}\mathrm{Var}(\hat{\sigma}^2) &= \frac{2\sigma^4}{nD} + \frac{\sigma^4}{n^2}\left\{\frac{5[c'^3\tau''(c')-2]^2}{8D^4} + \frac{-c'^4\tau'''(c') + 3c'^3\tau''(c') - 12}{2D^3} + \frac{7}{2D^2}\right\} \\ &+ O(n^{-3}) \end{aligned} \quad (2.132)$$

which is identical to the formula for $\mathrm{Var}(\hat{\sigma}^2)$ derived by Shenton & Bowman's method in equation (2.94) in section (2.9.1.2).

2.10 Simulation to estimate the variance when the mean is known:

The purpose of this section is to compare the results of a small simulation study with the theoretical results for $E(\hat{\sigma}^2)$.

In the study, $R = 10000$ samples were simulated for each value of the sample size $n = 5, 10, 20, 50, 100$ and each value of c shown in Table 2.8. The mean value and the standard deviation of $\hat{\sigma}^2$ was calculated in each case.

We used the NAG routine G05DDF(0,1), which generates random deviates from a standard normal distribution, and NAG routine G05CCF which changes the seeds of the random generator for each combination of n and c.

Program 13 (see Appendix) was used to compute the expected value and the standard deviation of the maximum likelihood estimator ($\hat{\sigma}^2$).

Table 2.8: The simulation results of the expected value and standard deviation of the maximum likelihood estimator of σ^2, for different values of n and c when $R = 10000$ and $\mu = 0$

R	n	$c = -1.88$		$c = -1$		$c = 0$		$c = 1$	
		$E(\hat{\sigma}^2)$	$\sigma(\hat{\sigma}^2)$	$E(\hat{\sigma}^2)$	$\sigma(\hat{\sigma}^2)$	$E(\hat{\sigma}^2)$	$\sigma(\hat{\sigma}^2)$	$E(\hat{\sigma}^2)$	$\sigma(\hat{\sigma}^2)$
10000	5	1.0186	0.5148	1.0179	0.5432	1.0079	0.7016	1.0090	0.8782
	10	1.0061	0.3583	1.0115	0.3840	0.9963	0.4412	0.9835	0.5271
	20	1.0040	0.2554	1.0030	0.2744	1.0004	0.3160	0.9917	0.3792
	50	1.0037	0.1589	1.0015	0.1720	1.0010	0.2014	0.9957	0.2420
	100	1.0005	0.1119	1.0002	0.1214	1.0005	0.1415	0.9988	0.1728

2.11 Conclusion:

In this section we compare the simulation results with the theoretical results as a check on accuracy of the theory and the algebra.

The results given in Tables 2.7 and 2.8, shown that, although the simulation does not give up to $O(n^{-3})$, the expectation of the estimators of the parameters are in agreement with the Shenton & Bowman and Cox & Hinkley methods and the sample size increase, the agreement becomes closer. We also conclude that expressions $E(\hat{\mu})$, $\mathrm{Var}(\hat{\mu})$, $E(\hat{\sigma}^2)$ and $\mathrm{Var}(\hat{\sigma}^2)$ in both theoretical methods; Shenton & Bowman and Cox & Hinkley are identical. Moreover all theoretical values lie within 95% confidence interval obtained for the simulation.

Chapter 3

The two parameter case of maximum likelihood estimator for the truncated normal distribution:

3.1 Introduction:

In this chapter we consider the Maximum Likelihood method of estimating the parameters μ and σ^2 simultaneously. Results are presented that extend the method of Shenton & Bowman to the two-parameter case to give the means, variances and covariance of $\hat{\mu}$ and $\hat{\sigma}^2$.

3.2 Likelihood equation when two parameters are unknown:

Suppose X has a normal distribution truncated at c, i.e. $x \leq c$, and that both the mean μ and variance σ^2 are unknown. Then its probability density function and the log likelihood function for a random sample of n observations are as defined in equations (2.1) and (2.2).

Since both μ and σ^2 are unknown, the estimates $\hat{\mu}$ and $\hat{\sigma^2}$ are simultaneous solutions of the following likelihood equations.

$$\frac{\partial l(\mathbf{x},\mu,\sigma^2)}{\partial\mu}\Big|_{\mu=\hat{\mu},\sigma^2=\hat{\sigma}^2} = \frac{\sum_{i=1}^{n}(x_i-\hat{\mu})}{\hat{\sigma}^2} + \frac{n\phi(\frac{c-\hat{\mu}}{\hat{\sigma}})}{\hat{\sigma}\Phi(\frac{c-\hat{\mu}}{\hat{\sigma}})} = 0 \tag{3.1}$$

$$\frac{\partial l(\mathbf{x},\mu,\sigma^2)}{\partial\sigma^2}\Big|_{\mu=\hat{\mu},\sigma^2=\hat{\sigma}^2} = -\frac{n}{2\hat{\sigma}^2} + \frac{\sum_{i=1}^{n}(x_i-\hat{\mu})^2}{2\hat{\sigma}^4} + \frac{n\phi(\frac{c-\hat{\mu}}{\hat{\sigma}})}{2\hat{\sigma}^3\Phi(\frac{c-\hat{\mu}}{\hat{\sigma}})} = 0. \tag{3.2}$$

To simplify the equations let

1. $l(\mathbf{x},\mu,\sigma^2) = l$
2. $c' = \frac{(c-\mu)}{\sigma}$.
3. $\hat{c}' = \frac{(c-\hat{\mu})}{\hat{\sigma}}$.
4. $\psi(c') = \frac{\sigma\phi(c')}{\Phi(c')}$.
5. $\psi(\hat{c}') = \frac{\hat{\sigma}\phi(\hat{c}')}{\Phi(\hat{c}')}$.
6. $\sum_{i=1}^{n}(x_i-\hat{\mu})^2 = \sum_{i=1}^{n}(x_i-\bar{x})^2 + n(\bar{x}-\hat{\mu})^2 = ns^2 + n(\bar{x}-\hat{\mu})^2$

where $s^2 = \sum_{i=1}^{n}(x_i-\bar{x})^2/n$.

Then we have

$$\frac{\partial l}{\partial\mu} = n\left(\bar{x}-\mu+\psi(c')\right),$$

$$\frac{\partial l}{\partial\sigma^2} = -\frac{n}{2\sigma^2} + \frac{ns^2+n(\bar{x}-\mu)^2}{2\sigma^4} + \frac{nc'\psi(c')}{2\sigma^3}.$$

The solution of $\frac{\partial l}{\partial\mu} = 0$ and $\frac{\partial l}{\partial\sigma^2} = 0$ can be written as

$$\hat{\mu} = \bar{x} + \psi(\hat{c}'), \tag{3.3}$$

$$\hat{\sigma}^2 = s^2 + (\bar{x}-\hat{\mu})^2 + \hat{c}'\hat{\sigma}\psi(\hat{c}'). \tag{3.4}$$

The algebraic simultaneous solution of the above equations is impossible, so an iterative numerical method is required.

We first show that $\bar{X}$ and s^2 are jointly sufficient statistics for μ and σ^2.

From a random sample of size n, the likelihood equation can be written as

$$L(x,\mu,\sigma^2) = \exp\left[\frac{-1}{2\sigma^2}\sum_{i=1}^{n} x_i^2 + \frac{\mu}{\sigma^2}\sum_{i=1}^{n} x_i - \frac{n\mu^2}{2\sigma^2} - n\ln\sqrt{2\pi\sigma^2}\Phi(\frac{c-\mu}{\sigma})\right].$$

Therefore, the statistics $Y_1 = \sum_{i=1}^{n} X_i^2$ and $Y_2 = \sum_{i=1}^{n} X_i$ are jointly sufficient for μ and σ^2.

Since

$$\bar{X} = \frac{Y_2}{n}$$

and

$$s^2 = \frac{Y_1 - Y_2^2/n}{n} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n}$$

are one-to-one functions of Y_1 and Y_2, satisfying the requirement of the factorization theorem, it follows that they are also jointly sufficient for the parameters μ and σ^2.

Theorem 3.1 *The function l has a local maximum.*

Proof: We prove this theorem for the two cases $c' > 0$ and $c' < 0$. From Theorems (2.1) and (2.3) we see what happens for fixed μ or for fixed σ.

We need to establish that there does not exist a path in the (μ, σ) plane along which l becomes unbounded.

Since l is continuous and differentiable, if we show that $l \to -\infty$ as $\mu \to -\infty$ and $l \to -\infty$ as $\mu \to \infty$, then we conclude that l has a local maximum.

1. For $c' > 0$ $(\mu < c)$:

   In this case l can be written as

$$l \;=\; -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2)$$

$$- \quad \frac{1}{2}\sum_{i=1}^{n}(\frac{x_i}{\sigma} - \frac{\mu}{\sigma})^2 - n\ln\Phi(\frac{c}{\sigma} + \frac{-\mu}{\sigma}).$$

If $\frac{-\mu}{\sigma} = \gamma$, a constant, then $-\frac{1}{2}\sum_{i=1}^{n}(\frac{x_i}{\sigma} - \frac{\mu}{\sigma})^2$ and $-n\ln\Phi(\frac{c}{\sigma} + \frac{-\mu}{\sigma})$ are constant as $\mu \to -\infty$ and $-\frac{n}{2}\ln(\sigma^2) \to -\infty$ as $\mu \to -\infty$. Therefore $l \to -\infty$ as $\mu \to -\infty$.

2. For $c' < 0$ ($\mu > c$):

   Using the note of section (2.2) l can be written as

$$\begin{aligned} l &= -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - n\ln u(-c') \\ &+ \frac{\sum_{i=1}^{n}(c - x_i)(c + x_i)}{2\sigma^2} - \frac{\mu}{\sigma^2}\sum_{i=1}^{n}(c - x_i). \end{aligned}$$

   If $\frac{\mu}{\sigma^2} = \gamma$, a constant, and $u(-c')$ is bounded then $-n\ln u(-c')$, $\frac{\sum_{i=1}^{n}(c-x_i)(c+x_i)}{2\sigma^2}$ and $-\frac{\mu}{\sigma^2}\sum_{i=1}^{n}(c - x_i)$ are constant as $\mu \to \infty$. But $-\frac{n}{2}\ln(\sigma^2) \to -\infty$ as $\mu \to \infty$. Therefore $l \to -\infty$ as $\mu \to \infty$ and the theorem is proved.

3.3 Scoring method:

Let $\hat{\theta} = (\mu, \sigma^2)$ be the simultaneous solution of equations (3.3) and (3.4) and θ_n be the n^{th} iteration with this method, then $\theta_{n+1} = \theta_n + \mathbf{I}^{-1}\mathbf{S}$, where $\mathbf{S}$ is the first derivative of the log likelihood at θ_n, i.e. the score vector, and $\mathbf{I}$ is the information matrix.

As Cohen (1986) described "One advantage of this method is that the inverse of the matrix $\mathbf{I}$, which is computed in each iteration step, can be used as an estimate of the covariance matrix". It is noted that this method is a modification of the Newton Raphson method.

3.3.1 Score and information matrix:

In the context of this thesis, let

$$\theta = \begin{bmatrix} \mu \\ \sigma^2 \end{bmatrix},$$

$$\mathbf{S} = \begin{bmatrix} \frac{\partial l}{\partial \mu} \\ \frac{\partial l}{\partial \sigma^2} \end{bmatrix},$$

$$\mathbf{I} = \begin{bmatrix} I_{11} & I_{12} \\ I_{21} & I_{22} \end{bmatrix} = \begin{bmatrix} -E(\frac{\partial^2 l}{\partial \mu^2}) & -E(\frac{\partial^2 l}{\partial \mu \partial \sigma^2}) \\ -E(\frac{\partial^2 l}{\partial \mu \partial \sigma^2}) & -E(\frac{\partial^2 l}{\partial (\sigma^2)^2}) \end{bmatrix}.$$

For any appropriate starting value θ_0 of θ, we can find $\theta_1 = \theta_0 + \mathbf{I}^{-1}\mathbf{S}$ which is closer to real solution of equation. This procedure can be continued until two consecutive values are the same to some given accuracy.

To obtain each element of the information matrix we have the following.

$$I_{11} = -E(\frac{\partial^2 l}{\partial \mu^2}) = \frac{n}{\sigma^2}\left[1 - \frac{\partial \psi(c')}{\partial \mu}\right],$$

where

$$\frac{\partial \psi(c')}{\partial \mu} = \frac{c'\psi(c')}{\sigma} + \left[\frac{\psi(c')}{\sigma}\right]^2.$$

$$I_{12} = -E(\frac{\partial^2 l}{\partial \mu \partial \sigma^2}) = \frac{-\Sigma_{i=1}^n E(X_i - \mu)}{\sigma^4} + \frac{n}{\sigma^4}\left[\sigma^2 \frac{\partial \psi(c')}{\partial \sigma^2} - \psi(c')\right].$$

Since from Chapter 0, section (0.1.8), we have $E(X) = \mu - \psi(c')$, we can write

$$I_{12} = \frac{-n\psi(c')\left[c'\psi(c')/\sigma + c'^2 + 1\right]}{2\sigma^4}.$$

Similarly

$$I_{21} = -E(\frac{\partial^2 l}{\partial \sigma^2 \partial \mu}) = \frac{-n\psi(c')\left[c'\psi(c')/\sigma + c'^2 + 1\right]}{2\sigma^4} = I_{12}.$$

Now we want to find I_{22}. For this we have

$$\frac{\partial^2 l}{\partial (\sigma^2)^2} = \frac{n}{2\sigma^4} - \frac{\Sigma_{i=1}^n (x_i - \mu)^2}{\sigma^6} + \frac{nc'\sigma}{2} \frac{\partial \left[\frac{\psi(c')}{\sigma^4}\right]}{\partial \sigma^2}.$$

Since we obtained

$$\frac{\partial \psi(c')}{\partial \sigma^2} = \frac{\psi(c')}{2\sigma^2}\left[1 + c'^2 + c'\psi(c')\right],$$

by taking the derivative of $\frac{\psi(c')}{\sigma^4}$ with respect to σ^2, we can find

$$\frac{\partial\left[\frac{\psi(c')}{\sigma^4}\right]}{\partial \sigma^2} = \frac{\psi(c')}{2\sigma^6}\left[1 + c'^2 + c'\psi(c') - 4\right].$$

Therefore

$$I_{22} = -E\left[\frac{\partial^2 l}{\partial(\sigma^2)^2}\right] = \frac{-n}{2\sigma^4} + \frac{E\left[\Sigma_{i=1}^n (X_i - \mu)^2\right]}{\sigma^6} - \frac{nc'\sigma}{2}\frac{\psi(c')}{2\sigma^6}\left[c'^2 + c'\psi(c') - 3\right].$$

Now using Chapter 0, sections (0.1.8) and (0.1.12) we obtain that

$$E(X - \mu)^2 = \sigma^2\left[1 - \frac{c'\psi(c')}{\sigma}\right]. \tag{3.5}$$

Therefore we can write

$$\begin{aligned}
I_{22} &= \frac{-n}{2\sigma^4} + \frac{n\sigma^2}{\sigma^6}\left[1 - \frac{c'\psi(c')}{\sigma}\right] - \frac{nc'\psi(c')}{4\sigma^5}\left[c'^2 + c'\psi(c') - 3\right] \\
&= -\frac{n}{2\sigma^4} + \frac{n}{\sigma^4} - \frac{nc'\psi(c')}{4\sigma^5}\left[4 + c'^2 + c'\psi(c') - 3\right] \\
&= \frac{n}{2\sigma^4} + \frac{c'}{2\sigma}\left\{\frac{-n\psi(c')\left[c'\psi(c')/\sigma + c'^2 + 1\right]}{2\sigma^4}\right\} \\
&= \frac{n}{2\sigma^4} + \frac{c'I_{12}}{2\sigma}.
\end{aligned}$$

We can now write the score vector and information matrix of the parameters as follows:

$$\mathbf{S} = \begin{bmatrix} \frac{\partial l}{\partial \mu} \\ \frac{\partial l}{\partial \sigma^2} \end{bmatrix} = \begin{bmatrix} \frac{n[\bar{x} - \mu + \psi(c')]}{\sigma^2} \\ \frac{n\left[-\sigma^2 + s^2 + (\bar{x} - \mu)^2 + (c - \mu)\psi(c')\right]}{\sigma^4} \end{bmatrix}$$

$$\mathbf{I} = \begin{bmatrix} \frac{n\left\{1 - c'\psi(c')/\sigma - [\psi(c')/\sigma]^2\right\}}{\sigma^2} & \frac{-n\psi(c')\left[c'\psi(c')/\sigma + c'^2 + 1\right]}{2\sigma^4} \\ \frac{-n\psi(c')\left[c'\psi(c')/\sigma + c'^2 + 1\right]}{2\sigma^4} & \frac{n}{2\sigma 4} + \frac{c'}{2\sigma}\left\{\frac{-n\psi(c')\left[c'\psi(c')/\sigma + c'^2 + 1\right]}{2\sigma^4}\right\} \end{bmatrix}$$

3.3.2 Estimation of the mean and variance:

To find the maximum likelihood estimators $\hat{\mu}$ and $\hat{\sigma}^2$ of the parameters of the population, we have to find the score vector and information matrix. Since both of them involve $\phi(c')$ and $\Phi(c')$ through $\psi(c')$, we need to calculate $\psi(c')$.

For this, we use Normal routine from the NAG library. The computer program is provided in Program 14 the Appendix.

This program was tested for two data sets that were introduced in Chapter 1. In these the heights of children, 3^{rd} centile, were given. We used these data in the program to estimate the mean and variance of the population by the method of scoring.

There are problems in estimating $\hat{\mu}$ and $\hat{\sigma}^2$ simultaneously by the method of scoring. Cohen (1957) stated " Newton's method tends to produce rather slowly converging iterants during the first few cycles of computation, unless initial approximations are in a close neighbourhood of the solution. "He mentioned that this difficulty has been recognized and discussed by Norton (1956). As we cannot guarantee that our starting values are near enough to the solution, we have tried to evaluate the estimate of the parameters by the scoring method. But this method failed and we have tried to evaluate them by the other method.

Using the NAG routine C05NBF, which find a solution of a system of nonlinear equations by a modification of the Powell hybrid method (1970), to solve equations (3.3) and (3.4) simultaneously, we obtained the following results for data described in Chapter 1 (see Appendix Program 15). Moreover, to illustrate our results, we wrote a program using S-PLUS software, to draw a three dimensional plot of the likelihood against μ and σ (see Appendix Program 16). It should be noted that the Program 15 failed for data set 1 (boys) and we used the method of gride (used in Chapter 5) to maximize the l. (see Appendix Program 15a). This program stops when $|\mu_{i+1} - \mu_i| < 10^{-5}$ and $|\log(\sigma_{i+1}) - \log(\sigma_i)| < 10^{-5}$.

3.3.3 Estimates of (μ, σ) in data sets 1 and 2:

1. Using the data set 1, (boys), we obtain the maximum likelihood estimates

$$(\hat{\mu}, \hat{\sigma}) = (1.3377, 1.2871).$$

   On plotting the loglikelihood against μ [-2, 2] and σ [0.1, 2] we get Figure 3.1. By changing the range of μ and σ to [1, 1.5] and [1, 1.5] we get Figure 3.2.

2. Using the data set 2, with $\varepsilon = 10^{-5}$, we obtain after 8 iterations, the maximum likelihood estimates

$$(\hat{\mu}, \hat{\sigma}) = (-2.1932, 0.2767).$$

In view of the exact values of μ and σ, the difference of $\hat{\mu}$ and $\hat{\sigma}$ from the exact values of μ and σ in the two-parameter case are high. In particular in the extreme case of truncation point $c = -1.88$, the value $\hat{\mu} < c$ whereas in the one-parameter case we obtained $\hat{\mu} > c$.

3.3.4 Estimates of (μ, σ) in the ideal samples sizes 5 and 10:

In this section we find the ML estimates of μ and σ, simultaneously for the ideal samples.

Using the ideal samples of size 5 and 10 and letting $\varepsilon = 10^{-5}$, we now obtain estimates of $\hat{\mu}$ and $\hat{\sigma}$ for the various truncation points.

Figure 3.1: Loglikelihood versus μ and σ data set 1 (boys) ($c = -1.88$)

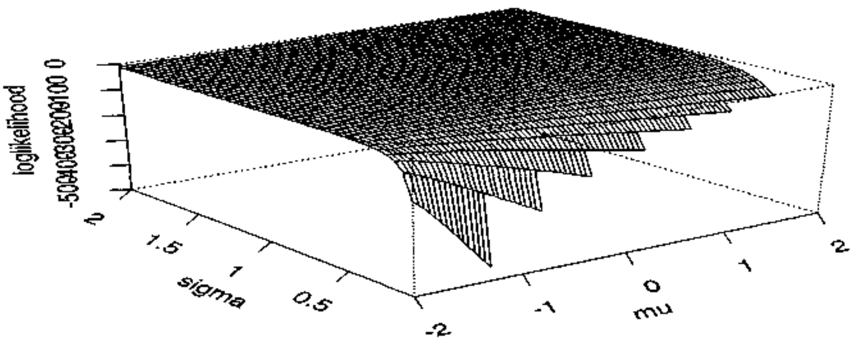


Figure 3.2: Loglikelihood versus μ and σ data set 1 (boys) ($c = -1.88$)

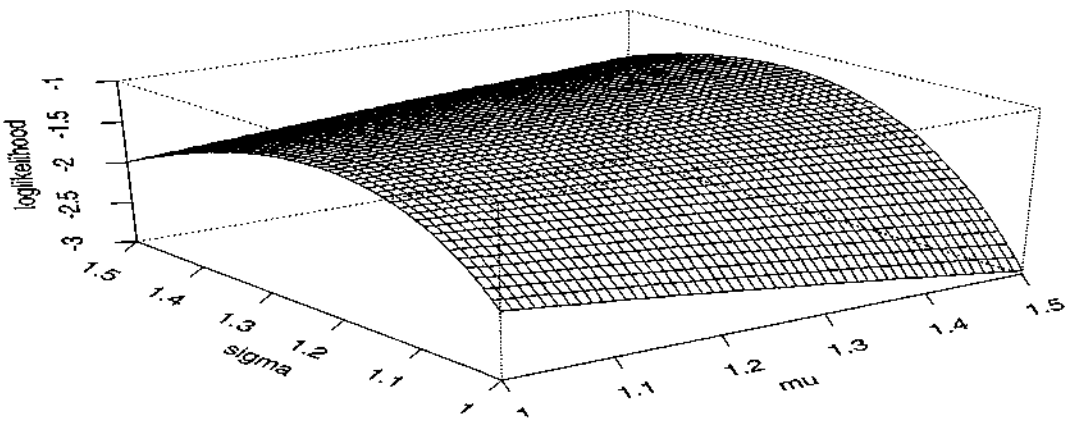


Table 3.1: The estimate of (μ, σ) in the ideal sample of size 5 and 10 for different truncation points

c	-1.88		-1		0		1		3	
n	5	10	5	10	5	10	5	10	5	10
$\bar{y}$	-2.2165	-2.234	-1.4657	-1.4858	-0.7332	-0.7549	-0.2360	-0.2517	-0.0022	-0.0024
s	0.2170	0.2602	0.2884	0.3422	0.4112	0.4802	0.5597	0.6492	0.6686	0.7868
$\hat{\mu}$	-2.1126	-1.9509	-1.3576	-1.2033	-0.6483	-0.5443	-0.2056	-0.1681	-0.0022	-0.0022
$\hat{\sigma}$	0.2865	0.4099	0.3655	0.5043	0.4809	0.6242	0.5924	0.7253	0.6686	0.7872

1. Using the ideal sample of size 5, when $c = 1$ with $\varepsilon = 10^{-5}$, after 9 iterations the following maximum likelihood estimate is obtained:

$$(\hat{\mu}, \hat{\sigma}) = (-0.2056, 0.5924).$$

2. Using the ideal sample of size 10, when $c = 1$ with $\varepsilon = 10^{-5}$, we obtain after 10 iterations the maximum likelihood estimates:

$$(\hat{\mu}, \hat{\sigma}) = (-0.1681, 0.7253).$$

On plotting the likelihood against μ and σ together with the contour plot, we obtain Figures 3.3 and 3.4 which confirms the calculated values.

Similar to data set 2, we can see that for all the truncation points considered the ML estimates of μ in the two-parameter case are less than their counterpart truncation points, whereas in the one-parameter case they are not. Also, as we see from Table 3.1, the greater the truncation points the more sensible are the estimates of μ and σ. We can see that in two-parameter case all the values $\hat{\mu} < c$ whereas in the one-parameter case we obtained $\hat{\mu} > c$ for $c < 0$ and $\hat{\mu} < c$ for $c > 0$. Therefore we conclude that lack of constraint on μ and σ in two-parameter case allow the ML estimate of μ be less than c.

Figure 3.3: Likelihood versus μ and σ for the ideal sample of size 10 ($c = 1$)

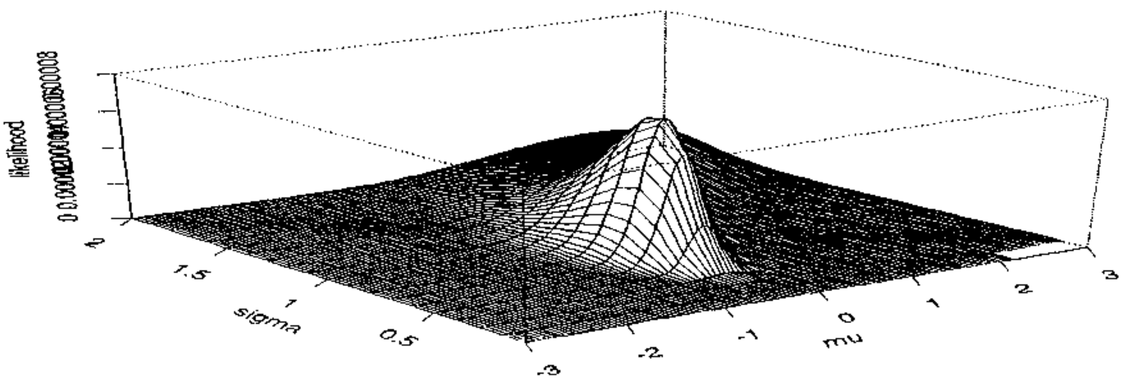
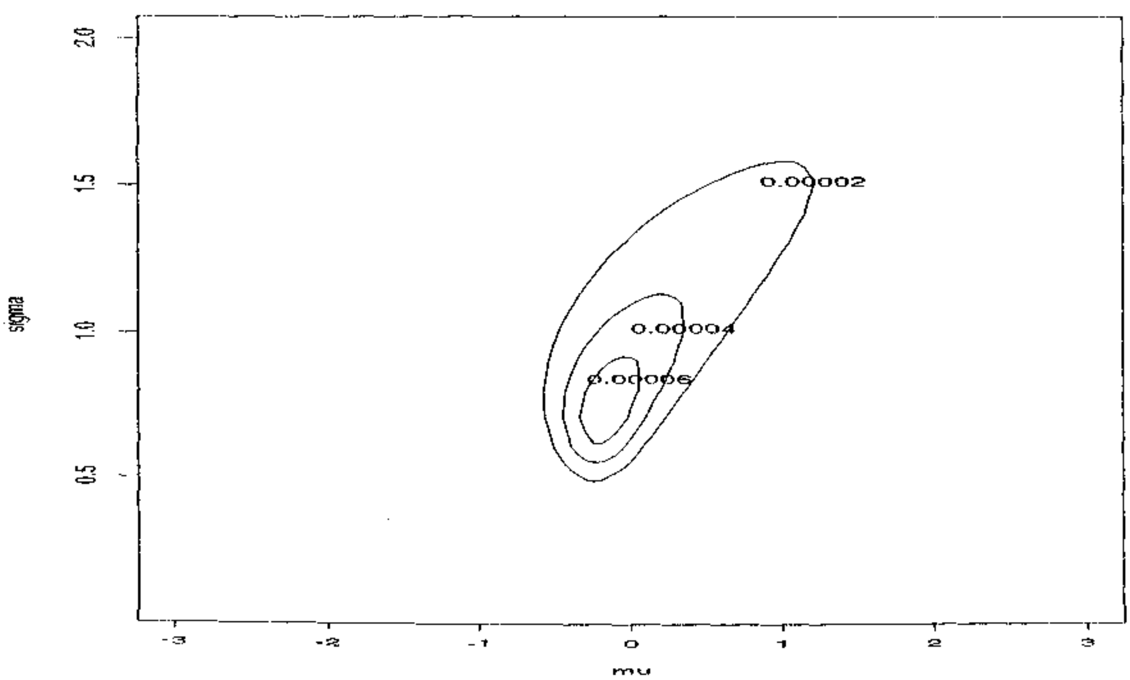


Figure 3.4: Contour plot of the likelihood versus μ and σ for the ideal sample of size 10 ($c = 1$)



3.4 Simulation to estimate mean and variance simultaneously:

The purpose of this section is to compare the results of a small simulation study with the theoretical results for $E(\hat{\mu})$, $E(\hat{\sigma}^2)$, $\sigma(\hat{\mu})$, $\sigma(\hat{\sigma}^2)$ and $\rho(\hat{\mu}, \hat{\sigma}^2)$.

In the study, $R = 10000$ sample were simulated from a $N(0,1)$ truncated distribution. Values of $E(\hat{\mu})$, $\sigma(\hat{\mu})$, $E(\hat{\sigma}^2)$, $\sigma(\hat{\sigma}^2)$ and $\rho(\hat{\mu}, \hat{\sigma}^2)$ were calculated for each value of the sample sizes $n = 5, 10, 20, 50, 100$ and each truncation point $c = -1.88, -1, 0, 1, 3$ and 10. The results were obtained from the solutions of the likelihood equations using Program 17, given in the Appendix, they are tabulated in the Tables 3.2 − 3.7.

Table 3.2: The simulation results for the expected value and standard deviation of $\hat{\mu}$ and $\hat{\sigma}^2$ and $\rho(\hat{\mu}, \hat{\sigma}^2)$, for different values of n when $R = 10000$ and $c = -1.88$

n	Mean		Variance		$\rho(\hat{\mu}, \hat{\sigma}^2)$
	$E(\hat{\mu})$	$\sigma(\hat{\mu})$	$E(\hat{\sigma}^2)$	$\sigma(\hat{\sigma}^2)$	
5	-2.211	0.149	0.121	0.132	-0.345
10	-2.198	0.107	0.136	0.100	-0.308
20	-2.188	0.076	0.147	0.075	-0.228
50	-2.184	0.049	0.151	0.047	-0.217
100	-2.182	0.034	0.153	0.035	-0.168

Concentrating on Table 3.2, we see that the value of $E(\hat{\mu})$ is always less than c and increases as the sample size n increases, whereas the value of $\sigma(\hat{\mu})$ decreases. Again by increasing n the value of $E(\hat{\sigma}^2)$ increases and the value $\sigma(\hat{\sigma}^2)$ decreases. It seems that the lack of constraint on μ and σ cause that the ML estimate $\hat{\mu}$ be alwayse less than c.

Table 3.3: The simulation results for the expected value and standard deviation of $\hat{\mu}$ and $\hat{\sigma}^2$ and $\rho(\hat{\mu}, \hat{\sigma}^2)$, for different values of n when $R = 10000$ and $c = -1$

n	Mean		Variance		$\rho(\hat{\mu}, \hat{\sigma}^2)$
	$E(\hat{\mu})$	$\sigma(\hat{\mu})$	$E(\hat{\sigma}^2)$	$\sigma(\hat{\sigma}^2)$	
5	-1.457	0.197	0.196	0.190	-0.320
10	-1.445	0.141	0.222	0.148	-0.283
20	-1.433	0.100	0.237	0.110	- 0.254
50	-1.429	0.064	0.246	0.072	-0.239
100	-1.426	0.046	0.249	0.052	-0.209

Table 3.4: For $c = 0$

n	Mean		Variance		$\rho(\hat{\mu}, \hat{\sigma}^2)$
	$E(\hat{\mu})$	$\sigma(\hat{\mu})$	$E(\hat{\sigma}^2)$	$\sigma(\hat{\sigma}^2)$	
5	-0.724	0.274	0.347	0.298	-0.257
10	-0.705	0.199	0.397	0.233	-0.216
20	-0.697	0.142	0.426	0.170	- 0.207
50	-0.690	0.091	0.440	0.112	-0.196
100	-0.689	0.064	0.446	0.079	-0.178

Table 3.5: The simulation results for the expected value and standard deviation of $\hat{\mu}$ and $\hat{\sigma}^2$ and $\rho(\hat{\mu}, \hat{\sigma}^2)$, for different values of n when $R = 10000$ and $c = 1$

n	Mean		Variance		$\rho(\hat{\mu}, \hat{\sigma}^2)$
	$E(\hat{\mu})$	$\sigma(\hat{\mu})$	$E(\hat{\sigma}^2)$	$\sigma(\hat{\sigma}^2)$	
5	-0.215	0.375	0.589	0.448	-0.119
10	-0.208	0.268	0.674	0.344	-0.108
20	-0.202	0.193	0.707	0.250	- 0.076
50	-0.201	0.122	0.730	0.161	-0.054
100	-0.199	0.087	0.736	0.113	-0.042

Table 3.6: For $c = 3$

n	Mean		Variance		$\rho(\hat{\mu}, \hat{\sigma}^2)$
	$E(\hat{\mu})$	$\sigma(\hat{\mu})$	$E(\hat{\sigma}^2)$	$\sigma(\hat{\sigma}^2)$	
5	0.0039	0.452	0.818	0.612	0.051
10	0.0012	0.315	0.909	0.440	0.042
20	0.0057	0.223	0.957	0.321	0.029
50	0.0014	0.141	0.982	0.207	0.025
100	-0.0016	0.100	0.992	0.144	0.021

Table 3.7: The simulation results for the expected value and standard deviation of $\hat{\mu}$ and $\hat{\sigma}^2$ and $\rho(\hat{\mu}, \hat{\sigma}^2)$, for different values of n when $R = 10000$ and $c = 10$

n	Mean		Variance		$\rho(\hat{\mu}, \hat{\sigma}^2)$
	$E(\hat{\mu})$	$\sigma(\hat{\mu})$	$E(\hat{\sigma}^2)$	$\sigma(\hat{\sigma}^2)$	
5	0.0010	0.455	0.804	0.565	0.00739
10	-0.0020	0.317	0.903	0.422	-0.00523
20	-0.00020	0.224	0.948	0.307	-0.00145
50	-0.0013	0.142	0.981	0.199	-0.00120
100	0.0018	0.100	0.989	0.141	-0.00080

3.4.1 Conclusion:

Figure 3.1 contains a rather flat area so that many points have likelihood almost equal to the maximum, Figure 3.2 shows a close up of the likelihood around its maximum.

Concentrating on the above tables we can see that $E(\hat{\mu})$ and $E(\hat{\sigma}^2)$ generally increase and become closer to the true values as the truncation point c increases.

Note that we include $c = 10$, representing the complete normal distribution, in order to check the results of the simulation study.

Using the t statistic, $t = \frac{\rho(\hat{\mu},\hat{\sigma}^2)\sqrt{R-2}}{\sqrt{1-\rho^2(\hat{\mu},\hat{\sigma}^2)}}$ with $R-2$ degrees of freedom, we tested the null hypothesis $H_0 : \rho = 0$ and found that the tests are significant for all values of n and c (except $c = 10$) in other words $\rho \neq 0$. We conclude that $\hat{\mu}$ and $\hat{\sigma}^2$ are dependent in the truncated case when truncation is present (see Tables 3.2 $-$ 3.7).

3.5 Theoretical results based on expansions in terms of $(\bar{x} - \mu_c)$ and $(s^2 - \sigma_c^2)$ (Method A):

The aim of this section is to extend the results of the Shenton & Bowman formula to the joint estimation of the two parameters. To do this, letting $\tau(c') = \frac{\phi(c')}{\Phi(c')}$ and $\tau(\hat{c}') = \frac{\phi(\hat{c}')}{\Phi(\hat{c}')}$, we note that (3.1) and (3.2) can be written as

$$\begin{cases} \hat{\mu} = \bar{x} + \hat{\sigma}\tau(\hat{c}') \\ \hat{\sigma}^2 = s^2 + (\bar{x} - \hat{\mu})^2 + \hat{c}'\hat{\sigma}^2\tau(\hat{c}'). \end{cases} \tag{3.6}$$

Firstly, we have to prove the following theorem and, secondly, expand the equations (3.6).

Theorem 3.2 *The equations (3.6) are satisfied by $\hat{\mu} = \mu$ and $\hat{\sigma}^2 = \sigma^2$ if $\bar{x} = \mu_c$ and $s^2 = \sigma_c^2$, where $\bar{x} = \sum_{i=1}^n x_i/n$, $s^2 = \sum_{i=1}^n (x_i - \bar{x})^2/n$, $\mu_c = E(X)$ and $\sigma_c^2 = Var(X)$.*

Proof: Assume that we have equations (3.6), and the conditions $\bar{x} = \mu_c$ and $s^2 = \sigma_c^2$. We want to prove that $\hat{\mu} = \mu$ and $\hat{\sigma}^2 = \sigma^2$.

The equations (3.6) can be written as

$$\begin{cases} \hat{\mu} - \hat{\sigma}\tau(\hat{c}') = \bar{x} \\ \hat{\sigma}^2 - (\bar{x} - \hat{\mu})^2 - \hat{c}'\hat{\sigma}^2\tau(\hat{c}') = s^2. \end{cases} \tag{3.7}$$

As we know from Chapter 0, sections (0.1.8) and (0.1.12)

$$\mu_c = \mu - \sigma\tau(c')$$

and

$$\sigma_c^2 = \sigma^2[1 - c'\tau(c') - \tau^2(c')].$$

By the assumption of the theorem, the equations (3.7) can be expressed as

$$\begin{cases} \hat{\mu} - \hat{\sigma}\tau(\hat{c}') - \mu + \sigma\tau(c') = 0 \\ \hat{\sigma}^2 - (\mu_c - \hat{\mu})^2 - \hat{c}'\hat{\sigma}^2\tau(\hat{c}') - \sigma^2[1 - c'\tau(c') - \tau^2(c')] = 0. \end{cases} \tag{3.8}$$

If, in equation (3.8), $\hat{\mu} = \mu$ and $\hat{\sigma}^2 = \sigma^2$, then

$$\begin{cases} \mu - \sigma\tau(c') - \mu + \sigma\tau(c') = 0 \\ \sigma^2 - [\mu - \sigma\tau(c') - \mu]^2 - c'\sigma^2\tau(c') - \sigma^2[1 - c'\tau(c') - \tau^2(c')] = 0. \end{cases} \tag{3.9}$$

In view of the fact that

$$\begin{cases} \mu - \sigma\tau(c') - \mu + \sigma\tau(c') = 0 \\ \sigma^2[1 - c'\tau(c') - \tau^2(c')] - \sigma^2[1 - c'\tau(c') - \tau^2(c')] = 0, \end{cases} \tag{3.10}$$

it can be seen that $(\hat{\mu} = \mu, \hat{\sigma}^2 = \sigma^2)$ is a solution of (3.6).

Expansion of Equations: Using the extension of Shenton & Bowman's results we can expand each equation of (3.6) in the following way

$$\begin{aligned} \hat{\mu} &= A_{00} + A_{10}(\bar{x} - \mu_c)/1! + A_{01}(s^2 - \sigma_c^2)/1! + \\ &\quad A_{20}(\bar{x} - \mu_c)^2/2! + A_{02}(s^2 - \sigma_c^2)^2/2! + \\ &\quad A_{11}(\bar{x} - \mu_c)(s^2 - \sigma_c^2)/1!1! + \cdots \end{aligned} \tag{3.11}$$

and

$$\begin{aligned} \hat{\sigma}^2 &= B_{00} + B_{10}(\bar{x} - \mu_c)/1! + B_{01}(s^2 - \sigma_c^2)/1! + \\ &\quad B_{20}(\bar{x} - \mu_c)^2/2! + B_{02}(s^2 - \sigma_c^2)^2/2! + \\ &\quad B_{11}(\bar{x} - \mu_c)(s^2 - \sigma_c^2)/1!1! + \cdots. \end{aligned} \tag{3.12}$$

This expansion can be continued in a similar manner. For our present estimation purposes we have given the expansion of $\hat{\mu}$ and $\hat{\sigma}^2$ (equations (3.11) and (3.12)) only up to the second term. Then it follows that we must have

1. $A_{10} = \frac{\partial\hat{\mu}}{\partial\bar{x}} \big|_{(\bar{x}=\mu_c, s^2=\sigma_c^2)}$,
2. $A_{01} = \frac{\partial\hat{\mu}}{\partial s^2} \big|_{(\bar{x}=\mu_c, s^2=\sigma_c^2)}$,

3. $A_{20} = \frac{\partial^2 \hat{\mu}}{\partial(\bar{x})^2} \mid_{(\bar{x}=\mu_c, s^2=\sigma_c^2)}$,
4. $A_{02} = \frac{\partial^2 \hat{\mu}}{\partial(s^2)^2} \mid_{(\bar{x}=\mu_c, s^2=\sigma_c^2)}$,
5. $A_{11} = \frac{\partial^2 \hat{\mu}}{\partial\bar{x}\partial s^2} \mid_{(\bar{x}=\mu_c, s^2=\sigma_c^2)}$,
6. $B_{10} = \frac{\partial \hat{\sigma}^2}{\partial \bar{x}} \mid_{(\bar{x}=\mu_c, s^2=\sigma_c^2)}$,
7. $B_{01} = \frac{\partial \hat{\sigma}^2}{\partial s^2} \mid_{(\bar{x}=\mu_c, s^2=\sigma_c^2)}$,
8. $B_{20} = \frac{\partial^2 \hat{\sigma}^2}{\partial(\bar{x})^2} \mid_{(\bar{x}=\mu_c, s^2=\sigma_c^2)}$,
9. $B_{02} = \frac{\partial^2 \hat{\sigma}^2}{\partial(s^2)^2} \mid_{(\bar{x}=\mu_c, s^2=\sigma_c^2)}$,
10. $B_{11} = \frac{\partial^2 \hat{\sigma}^2}{\partial\bar{x}\partial s^2} \mid_{(\bar{x}=\mu_c, s^2=\sigma_c^2)}$ etc.

According to Theorem 3.2, if $\bar{x} \to \mu_c$ and $s^2 \to \sigma_c^2$, then $\hat{\mu} \to \mu$ and $\hat{\sigma}^2 \to \sigma^2$. Therefore from equations (3.11) and (3.12) it can be concluded that $A_{00} = \mu$ and $B_{00} = \sigma^2$.

By taking the first and second partial derivatives of each equation and using Theorem 3.2, we can find the remaining coefficients.

3.5.1 Preliminary calculation:

In this section we calculate some of the expressions which we will need later on.

3.5.1.1 Calculation of $\frac{\partial \hat{c}'}{\partial s^2}$:

Let

$$\hat{c}' = (c - \hat{\mu})(\hat{\sigma}^2)^{-1/2} . \tag{3.13}$$

Taking the partial derivative of $\hat{c}'$ with respect to s^2 we obtain

$$\frac{\partial \hat{c}'}{\partial s^2} = \frac{\partial(c-\hat{\mu})}{\partial \hat{\mu}} \cdot \frac{\partial \hat{\mu}}{\partial s^2}(\hat{\sigma}^2)^{-1/2} + \frac{\partial(\hat{\sigma}^2)^{-1/2}}{\partial \hat{\sigma}^2} \cdot \frac{\partial \hat{\sigma}^2}{\partial s^2}(c - \hat{\mu})$$

$$
\begin{aligned}
&= -(\hat{\sigma}^2)^{-1/2}\,\frac{\partial\hat{\mu}}{\partial s^2} - \frac{1}{2}(\hat{\sigma}^2)^{-3/2}\,(c-\hat{\mu})\frac{\partial\hat{\sigma}^2}{\partial s^2} \\
&= [-\frac{1}{\hat{\sigma}}\frac{\partial\hat{\mu}}{\partial s^2} - \frac{\hat{c}'}{2\hat{\sigma}^2}\frac{\partial\hat{\sigma}^2}{\partial s^2}]. \qquad (3.14)
\end{aligned}
$$

3.5.1.2 Calculation of $\frac{\partial\hat{c}'}{\partial\bar{x}}$:

Similarly, taking partial derivative of $\hat{c}'$ with respect to $\bar{x}$ we obtain

$$
\frac{\partial\hat{c}'}{\partial\bar{x}} = [-\frac{1}{\hat{\sigma}}\frac{\partial\hat{\mu}}{\partial\bar{x}} - \frac{\hat{c}'}{2\hat{\sigma}^2}\frac{\partial\hat{\sigma}^2}{\partial\bar{x}}]. \qquad (3.15)
$$

3.5.2 Calculation of A_{10} and B_{10}:

Now, taking the partial derivative of the first equation in (3.6) with respect to $\bar{x}$, we have

$$
\begin{aligned}
\frac{\partial\hat{\mu}}{\partial\bar{x}} &= 1 + \frac{\partial[(\hat{\sigma}^2)^{1/2}\tau(\hat{c}')]}{\partial\bar{x}} \\
&= 1 + \frac{\partial(\hat{\sigma}^2)^{1/2}}{\partial\hat{\sigma}^2}.\frac{\partial\hat{\sigma}^2}{\partial\bar{x}}\tau(\hat{c}') + \frac{\partial\tau(\hat{c}')}{\partial\hat{c}'}.\frac{\partial\hat{c}'}{\partial\bar{x}}.(\hat{\sigma}^2)^{1/2} \qquad (3.16)
\end{aligned}
$$

which , on replacing $\frac{\partial\hat{c}'}{\partial\bar{x}}$ from equation (3.15) writing $\tau'(\hat{c}')$ for $\frac{\partial\tau(\hat{c}')}{\partial\hat{c}'}$, gives

$$
[1+\tau'(\hat{c}')]\frac{\partial\hat{\mu}}{\partial\bar{x}} + [\frac{\hat{c}'\tau'(\hat{c}') - \tau(\hat{c}')}{2\hat{\sigma}}]\frac{\partial\hat{\sigma}^2}{\partial\bar{x}} = 1. \qquad (3.17)
$$

Substituting $\hat{\mu}$ from first equation (3.6) into second equation, we obtain

$$
\hat{\sigma}^2 = s^2 + (c-\bar{x})\hat{\sigma}\tau(\hat{c}'). \qquad (3.18)
$$

Again, substituting $\hat{\sigma}\tau(\hat{c}')$ from first equation (3.6) into equation (3.18), we obtain

$$
\hat{\sigma}^2 = s^2 + (c-\bar{x})(\hat{\mu}-\bar{x}). \qquad (3.19)
$$

Taking the partial derivative of the equation (3.19) with respect to $\bar{x}$, we have

$$
(\bar{x}-c)\frac{\partial\hat{\mu}}{\partial\bar{x}} + \frac{\partial\hat{\sigma}^2}{\partial\bar{x}} = 2\bar{x} - c - \hat{\mu}. \qquad (3.20)
$$

Equations (3.17) and (3.20) can be written as

$$\begin{cases} [1+\tau'(\hat{c}')]\frac{\partial\hat{\mu}}{\partial\bar{x}} + [\frac{\hat{c}'\tau'(\hat{c}')-\tau(\hat{c}')}{2\hat{\sigma}}]\frac{\partial\hat{\sigma}^2}{\partial\bar{x}} = 1, \\ (\bar{x}-c)\frac{\partial\hat{\mu}}{\partial\bar{x}} + \frac{\partial\hat{\sigma}^2}{\partial\bar{x}} = 2\bar{x}-c-\hat{\mu}. \end{cases} \tag{3.21}$$

The simultaneous solution of equation (3.21), gives

$$\frac{\partial\hat{\mu}}{\partial\bar{x}} = \frac{1-(2\bar{x}-c-\hat{\mu})[\hat{c}'\tau'(\hat{c}')-\tau(\hat{c}')]/2\hat{\sigma}}{\hat{E}} \tag{3.22}$$

and

$$\frac{\partial\hat{\sigma}^2}{\partial\bar{x}} = \frac{c-\bar{x}+(2\bar{x}-c-\hat{\mu})[1+\tau'(\hat{c}')]}{\hat{E}}, \tag{3.23}$$

where

$$\hat{E} = 1+\tau'(\hat{c}')-(\bar{x}-c)[\hat{c}'\tau'(\hat{c}')-\tau(\hat{c}')]/2\hat{\sigma}. \tag{3.24}$$

Now, according to Theorem 3.2, as $\bar{x} \to \mu_c$ and $s^2 \to \sigma_c^2$, then $\hat{\mu} \to \mu$ and $\hat{\sigma} \to \sigma$. Therefore we can find

$$A_{10} = \frac{\partial\hat{\mu}}{\partial\bar{x}}\,|_{(\bar{x}=\mu_c,s^2=\sigma_c^2)} = \frac{1+[c'\tau'(c')-\tau(c')][c'+2\tau(c')]/2}{E} \tag{3.25}$$

and

$$B_{10} = \frac{\partial\hat{\sigma}^2}{\partial\bar{x}}\,|_{(\bar{x}=\mu_c,s^2=\sigma_c^2)} = \frac{-\sigma[\tau(c')+\tau'(c')(c'+2\tau(c'))]}{E}, \tag{3.26}$$

where

$$E = 1+\tau'(c')+[c'\tau'(c')-\tau(c')][c'+\tau(c')]/2. \tag{3.27}$$

3.5.3 Calculation of A_{01} and B_{01}:

Now, taking the partial derivative of the first equation in (3.6) with respect to s^2, we have

$$\begin{aligned} \frac{\partial\hat{\mu}}{\partial s^2} &= \frac{\partial[(\hat{\sigma}^2)^{1/2}\tau(\hat{c}')]}{\partial s^2} \\ &= \frac{\partial(\hat{\sigma}^2)^{1/2}}{\partial\hat{\sigma}^2}.\frac{\partial\hat{\sigma}^2}{\partial s^2}\tau(\hat{c}') + \frac{\partial\tau(\hat{c}')}{\partial\hat{c}'}.\frac{\partial\hat{c}'}{\partial s^2}.(\hat{\sigma}^2)^{1/2} \end{aligned} \tag{3.28}$$

which , with replacing $\frac{\partial \hat{c}'}{\partial s^2}$ from equation (3.14) and $\tau'(\hat{c}')$ for $\frac{\partial \tau(\hat{c}')}{\partial \hat{c}'}$ gives

$$[1+\tau'(\hat{c}')]\frac{\partial \hat{\mu}}{\partial s^2}+[\frac{\hat{c}'\tau'(\hat{c}')-\tau(\hat{c}')}{2\hat{\sigma}}]\frac{\partial \hat{\sigma}^2}{\partial s^2}=0. \tag{3.29}$$

Taking the partial derivative of the equation (3.19) with respect to s^2, we have

$$(\bar{x}-c)\frac{\partial \hat{\mu}}{\partial s^2}+\frac{\partial \hat{\sigma}^2}{\partial s^2}=1. \tag{3.30}$$

Solving the (3.29) and (3.30) simultaneously

gives

$$\frac{\partial \hat{\mu}}{\partial s^2}=-\frac{\hat{c}'\tau'(\hat{c}')-\tau(\hat{c}')}{2\hat{\sigma}\hat{E}} \tag{3.31}$$

and

$$\frac{\partial \hat{\sigma}^2}{\partial s^2}=\frac{1+\tau'(\hat{c}')}{\hat{E}} \tag{3.32}$$

where $\hat{E}$ is as defined in equation (3.24). Now, using Theorem 3.2 as $s^2 \to \sigma_c^2$, then $\hat{\sigma} \to \sigma$, we find

$$A_{01}=\frac{\partial \hat{\mu}}{\partial s^2}\mid_{(\bar{x}=\mu_c, s^2=\sigma_c^2)}=-\frac{c'\tau'(c')-\tau(c')}{2\sigma E}, \tag{3.33}$$

and

$$B_{01}=\frac{\partial \hat{\sigma}^2}{\partial s^2}\mid_{(\bar{x}=\mu_c, s^2=\sigma_c^2)}=\frac{1+\tau'(c')}{E} \tag{3.34}$$

where E is as defined in equation (3.27).

3.5.4 Calculation of A_{20} and B_{20}:

In this section we are going to calculate A_{20} and B_{20}. Taking partial derivative of equations (3.21) with respect to $\bar{x}$, we obtain

$$\begin{cases} [1+\tau'(\hat{c}')]\frac{\partial^2 \hat{\mu}}{\partial \bar{x}^2}+[\frac{\hat{c}'\tau'(\hat{c}')-\tau(\hat{c}')}{2\hat{\sigma}}]\frac{\partial^2 \hat{\sigma}^2}{\partial \bar{x}^2}=\frac{\tau''(\hat{c}')}{\hat{\sigma}}\left(\frac{\partial \hat{\mu}}{\partial \bar{x}}+\frac{\hat{c}'}{2\hat{\sigma}}\frac{\partial \hat{\sigma}^2}{\partial \bar{x}}\right)^2+\frac{\hat{c}'\tau'(\hat{c}')-\tau(\hat{c}')}{4\hat{\sigma}^3}(\frac{\partial \hat{\sigma}^2}{\partial \bar{x}})^2 \\ (\bar{x}-c)\frac{\partial^2 \hat{\mu}}{\partial \bar{x}^2}+\frac{\partial^2 \hat{\sigma}^2}{\partial \bar{x}^2}=2(1-\frac{\partial \hat{\mu}}{\partial \bar{x}}). \end{cases} \tag{3.35}$$

Solving the equations (3.35) simultaneously, we find

$$\frac{\partial^2\hat{\mu}}{\partial\bar{x}^2} = \frac{1}{\hat{\sigma}\hat{E}}\left\{\tau''(\hat{c}')\left(\frac{\partial\hat{\mu}}{\partial\bar{x}} + \frac{\hat{c}'}{2\hat{\sigma}}\frac{\partial\hat{\sigma}^2}{\partial\bar{x}}\right)^2 + [\hat{c}'\tau'(\hat{c}') - \tau(\hat{c}')]\left[\frac{1}{4\hat{\sigma}^2}\left(\frac{\partial\hat{\sigma}^2}{\partial\bar{x}}\right)^2 + \frac{\partial\hat{\mu}}{\partial\bar{x}} - 1\right]\right\} \tag{3.36}$$

and

$$\begin{aligned}\frac{\partial^2\hat{\sigma}^2}{\partial\bar{x}^2} &= \frac{1}{\hat{E}}\left\{2(1 - \frac{\partial\hat{\mu}}{\partial\bar{x}})[1 + \tau'(\hat{c}')]\right. \\ &- \left.(\bar{x} - c)\left[\frac{\tau''(\hat{c}')}{\hat{\sigma}}\left(\frac{\partial\hat{\mu}}{\partial\bar{x}} + \frac{\hat{c}'}{2\hat{\sigma}}\frac{\partial\hat{\sigma}^2}{\partial\bar{x}}\right)^2 + \frac{\hat{c}'\tau'(\hat{c}') - \tau(\hat{c}')}{4\hat{\sigma}^3}(\frac{\partial\hat{\sigma}^2}{\partial\bar{x}})^2\right]\right\}\end{aligned} \tag{3.37}$$

Now, using Theorem 3.2 we find

$$A_{20} = \frac{1}{\sigma E}\left\{\tau''(c')\left(A_{10} + \frac{c'}{2\sigma}B_{10}\right)^2 + [c'\tau'(c') - \tau(c')]\left(\frac{1}{4\sigma^2}B_{10}^2 + A_{10} - 1\right)\right\} \tag{3.38}$$

and

$$\begin{aligned}B_{20} &= \frac{1}{E}\left\{2(1 - A_{10})[1 + \tau'(c')] + [c' + \tau(c')]\left[\tau''(c')\left(A_{10} + \frac{c'}{2\sigma}B_{10}\right)^2\right.\right. \\ &+ \left.\left.\frac{1}{4\sigma^2}[c'\tau'(c') - \tau(c')]B_{10}^2\right]\right\},\end{aligned} \tag{3.39}$$

or in terms of c', $\tau(c')$, $\tau'(c')$, $\tau''(c')$,

$$\begin{aligned}A_{20} &= \frac{1}{4\sigma E^3}\left\{4\tau''(c')[1 - \tau(c')(c' + \tau(c'))]^2\right. \\ &+ \left.[c'\tau'(c') - \tau(c')][\tau(c') + \tau'(c')(c' + 2\tau(c'))]^2\right\} \\ &- \frac{1}{2\sigma E^2}\left\{[c'\tau'(c') - \tau(c')][2\tau'(c') - \tau(c')(c'\tau'(c') - \tau(c'))]\right\}\end{aligned}$$

and

$$\begin{aligned}B_{20} &= \frac{1}{E^3}\left\{\tau''(c')[c' + \tau(c')][1 - \tau(c')(c' + \tau(c'))]^2\right. \\ &+ \left.[c' + \tau'(c')][c'\tau'(c') - \tau(c')][\tau(c') + \tau'(c')(c' + 2\tau(c'))]^2/4\right\} \\ &+ \frac{1}{E^2}\left\{[1 + \tau(c')][2\tau'(c') - \tau(c')(c'\tau'(c') - \tau(c'))]\right\}\end{aligned}$$

3.5.5 Calculation of A_{02} and B_{02}:

In this section we calculate A_{02} and B_{02}. Taking the partial derivatives of equations (3.29) and (3.30) with respect to s^2 we obtain

$$\begin{cases} [1+\tau'(\hat{c}')]\frac{\partial^2\hat{\mu}}{\partial(s^2)^2} + [\frac{\hat{c}'\tau'(\hat{c}')-\tau(\hat{c}')}{2\hat{\sigma}}]\frac{\partial^2\hat{\sigma}^2}{\partial(s^2)^2} = \frac{\tau''(\hat{c}')}{\hat{\sigma}}\left(\frac{\partial\hat{\mu}}{\partial s^2}+\frac{\hat{c}'}{2\hat{\sigma}}\frac{\partial\hat{\sigma}^2}{\partial s^2}\right)^2 + \frac{\hat{c}'\tau'(\hat{c}')-\tau(\hat{c}')}{4\hat{\sigma}^3}(\frac{\partial\hat{\sigma}^2}{\partial s^2})^2 \\ (\bar{x}-c)\frac{\partial^2\hat{\mu}}{\partial(s^2)^2} + \frac{\partial^2\hat{\sigma}^2}{\partial(s^2)^2} = 0. \end{cases} \tag{3.40}$$

Solving the equations (3.40) simultaneously,

$$\frac{\partial^2\hat{\mu}}{\partial(s^2)^2} = \frac{1}{\hat{E}}\left\{\frac{\tau''(\hat{c}')}{\hat{\sigma}}\left(\frac{\partial\hat{\mu}}{\partial s^2}+\frac{\hat{c}'}{2\hat{\sigma}}\frac{\partial\hat{\sigma}^2}{\partial s^2}\right)^2 + \frac{\hat{c}'\tau'(\hat{c}')-\tau(\hat{c}')}{4\hat{\sigma}^3}(\frac{\partial\hat{\sigma}^2}{\partial s^2})^2\right\} \tag{3.41}$$

and

$$\frac{\partial^2\hat{\sigma}^2}{\partial(s^2)^2} = -\frac{(\bar{x}-c)}{\hat{E}}\left\{\frac{\tau''(\hat{c}')}{\hat{\sigma}}\left(\frac{\partial\hat{\mu}}{\partial s^2}+\frac{\hat{c}'}{2\hat{\sigma}}\frac{\partial\hat{\sigma}^2}{\partial s^2}\right)^2 + \frac{\hat{c}'\tau'(\hat{c}')-\tau(\hat{c}')}{4\hat{\sigma}^3}(\frac{\partial\hat{\sigma}^2}{\partial s^2})^2\right\}. \tag{3.42}$$

By Theorem 3.2 we therefore can find

$$A_{02} = \frac{1}{\sigma E}\left\{\tau''(c')\left(A_{01}+\frac{c'}{2\sigma}B_{01}\right)^2 + \frac{1}{4\sigma^2}[c'\tau'(c')-\tau(c')]B_{01}^2\right\} \tag{3.43}$$

and

$$B_{02} = \frac{1}{E}\left\{[c'+\tau(c')]\left[\tau''(c')\left(A_{01}+\frac{c'}{2\sigma}B_{01}\right)^2 + \frac{1}{4\sigma^2}[c'\tau'(c')-\tau(c')]B_{01}^2\right]\right\}. \tag{3.44}$$

Equivalently, in terms of c', $\tau(c')$, $\tau'(c')$, $\tau''(c')$, we have

$$A_{02} = \frac{\tau''(c')[c'+\tau(c')]^2 + [c'\tau'(c')-\tau(c')][1+\tau'(c')]^2}{4\sigma^3E^3}$$

and

$$B_{02} = \frac{\tau''(c')[c'+\tau(c')]^3 + [c'\tau'(c')-\tau(c')][c'+\tau(c')][1+\tau'(c')]^2}{4\sigma^2E^3}.$$

3.5.6 Calculation of A_{11} and B_{11}:

In this section we are going to calculate A_{11} and B_{11}. Taking the derivatives of equations (3.21) with respect to s^2 we obtain

$$\begin{cases} [1+\tau'(\hat{c}')]\frac{\partial^2\hat{\mu}}{\partial s^2\partial\bar{x}} + [\frac{\hat{c}'\tau'(\hat{c}')-\tau(\hat{c}')}{2\hat{\sigma}}]\frac{\partial^2\hat{\sigma}^2}{\partial s^2\partial\bar{x}} = \frac{\tau''(\hat{c}')}{\hat{\sigma}}\left(\frac{\partial\hat{\mu}}{\partial\bar{x}} + \frac{\hat{c}'}{2\hat{\sigma}}\frac{\partial\hat{\sigma}^2}{\partial\bar{x}}\right)\left(\frac{\partial\hat{\mu}}{\partial s^2} + \frac{\hat{c}'}{2\hat{\sigma}}\frac{\partial\hat{\sigma}^2}{\partial s^2}\right) + \\ \qquad \frac{\hat{c}'\tau'(\hat{c}')-\tau(\hat{c}')}{4\hat{\sigma}^3}\frac{\partial\hat{\sigma}^2}{\partial s^2}\frac{\partial\hat{\sigma}^2}{\partial\bar{x}} \\ (\bar{x}-c)\frac{\partial^2\hat{\mu}}{\partial s^2\partial\bar{x}} + \frac{\partial^2\hat{\sigma}^2}{\partial s^2\partial\bar{x}} = -\frac{\partial\hat{\mu}}{\partial s^2}. \end{cases} \tag{3.45}$$

Solving the equations (3.45) simultaneously,

$$\begin{aligned} \frac{\partial^2\hat{\mu}}{\partial s^2\partial\bar{x}} &= \frac{\tau''(\hat{c}')}{\hat{\sigma}\hat{E}}\left(\frac{\partial\hat{\mu}}{\partial s^2} + \frac{\hat{c}'}{2\hat{\sigma}}\frac{\partial\hat{\sigma}^2}{\partial s^2}\right)\left(\frac{\partial\hat{\mu}}{\partial\bar{x}} + \frac{\hat{c}'}{2\hat{\sigma}}\frac{\partial\hat{\sigma}^2}{\partial\bar{x}}\right) \\ &+ \left(\frac{\hat{c}'\tau'(\hat{c}')-\tau(\hat{c}')}{2\hat{\sigma}\hat{E}}\right)\left(\frac{1}{2\hat{\sigma}^2}\frac{\partial\hat{\sigma}^2}{\partial s^2}\frac{\partial\hat{\sigma}^2}{\partial\bar{x}} + \frac{\partial\hat{\mu}}{\partial s^2}\right) \end{aligned} \tag{3.46}$$

and

$$\begin{aligned} \frac{\partial^2\hat{\sigma}^2}{\partial s^2\partial\bar{x}} &= -\frac{1}{\hat{E}}\left\{\frac{(\bar{x}-c)\tau''(\hat{c}')}{\hat{\sigma}}\left(\frac{\partial\hat{\mu}}{\partial s^2} + \frac{\hat{c}'}{2\hat{\sigma}}\frac{\partial\hat{\sigma}^2}{\partial s^2}\right)\left(\frac{\partial\hat{\mu}}{\partial\bar{x}} + \frac{\hat{c}'}{2\hat{\sigma}}\frac{\partial\hat{\sigma}^2}{\partial\bar{x}}\right)\right. \\ &+ \left.\frac{(\bar{x}-c)[\hat{c}'\tau'(\hat{c}')-\tau(\hat{c})]}{4\hat{\sigma}^3}\left(\frac{\partial\hat{\sigma}^2}{\partial s^2}\frac{\partial\hat{\sigma}^2}{\partial\bar{x}}\right) + [1+\tau'(\hat{c}')]\frac{\partial\hat{\mu}}{\partial s^2}\right\}. \end{aligned} \tag{3.47}$$

By Theorem 3.2, therefore we can find

$$\begin{aligned} A_{11} &= \frac{\tau''(c')}{\sigma E}\left(A_{01} + \frac{c'}{2\sigma}B_{01}\right)\left(A_{10} + \frac{c'}{2\sigma}B_{10}\right) \\ &+ \left(\frac{c'\tau'(c')-\tau(c')}{2\sigma E}\right)\left(\frac{1}{2\sigma^2}B_{01}B_{10} + A_{01}\right) \end{aligned} \tag{3.48}$$

and

$$B_{11} = \frac{\tau''(c')[c'+\tau(c')]}{E}\left(A_{01} + \frac{c'}{2\sigma}B_{01}\right)\left(A_{10} + \frac{c'}{2\sigma}B_{10}\right) \tag{3.49}$$

$$+ \frac{[c'+\tau(c')][c'\tau'(c')-\tau(c')]}{4\sigma^2 E}B_{01}B_{10} - \frac{1+\tau'(c')}{E}A_{01}. \tag{3.50}$$

Expressions in terms of c', $\tau(c')$, $\tau'(c')$, $\tau''(c')$, are

$$
\begin{aligned}
A_{11} &= \frac{1}{4\sigma^2 E^3}\{2\tau''(c')[c'+\tau(c')][1-\tau(c')(c'+\tau(c'))] \\
&- [c'\tau'(c')-\tau(c')][1+\tau'(c')][\tau(c')+\tau'(c')(c'+2\tau(c'))]\} \\
&- \frac{1}{4\sigma^2 E^2}\left\{[c'\tau'(c')-\tau(c')]^2\right\}
\end{aligned}
$$

and

$$
\begin{aligned}
B_{11} &= \frac{c'+\tau(c')}{4\sigma E^3}\{2[\tau''(c')(c'+\tau(c'))[1-\tau(c')(c'+\tau(c'))] \\
&- [c'\tau'(c')-\tau(c')][1+\tau'(c')][\tau(c')+\tau'(c')(c'+2\tau(c'))]]\} \\
&+ \frac{1}{2\sigma E^2}\{[1+\tau(c')][c'\tau'(c')-\tau(c')]\}.
\end{aligned}
$$

To make sure that our calculations are correct, A_{11} and B_{11} were obtained by the alternative method of differentiation firstly with respect to s^2.

3.5.7 Calculation of $E(\hat{\mu})$:

In this section we derive $E(\hat{\mu})$. Taking the expectation of both sides of equation (3.11) and letting $\bar{x} \to \mu_c$, $s^2 \to \sigma_c^2$, $\hat{\mu} \to \mu$ and $\hat{\sigma}^2 \to \sigma^2$, we have

$$
\begin{aligned}
E(\hat{\mu}) &= A_{00} + A_{10}E(\bar{X}-\mu_c)/1! + A_{01}E(s^2-\sigma_c^2)/1! + \\
&\quad A_{20}E(\bar{X}-\mu_c)^2/2! + A_{02}E(s^2-\sigma_c^2)^2/2! + \\
&\quad A_{11}E(\bar{X}-\mu_c)(s^2-\sigma_c^2)/1!1! + \cdots \\
&= \mu + \Big\{A_{01}[E(s^2-\sigma_c^2)] + A_{20}E(\bar{X}-\mu_c)^2/2 + \\
&\quad A_{02}E(s^2-\sigma_c^2)^2/2 + A_{11}E(\bar{X}-\mu_c)(s^2-\sigma_c^2)\Big\} + \cdots, \qquad (3.51)
\end{aligned}
$$

since $E(\bar{X}-\mu_c) = 0$.

3.5.7.1 Calculation of $E(s^2 - \sigma_c^2)$:

Define the bias $b(s^2) = E(s^2 - \sigma_c^2) = E(s^2) - \sigma_c^2$. Since

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}, \tag{3.52}$$

and we know that

$$E[\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}] = \sigma_c^2 = \mathrm{Var}(X), \tag{3.53}$$

we obtain

$$E(s^2) = \frac{n-1}{n}\sigma_c^2 = \frac{n-1}{n}\mathrm{Var}(X) = \frac{n-1}{n}\mu_2(X). \tag{3.54}$$

The bias term is

$$\begin{aligned} b(s^2) &= E(s^2) - \sigma_c^2 = \frac{n-1}{n}\sigma_c^2 - \sigma_c^2 = -\frac{1}{n}\sigma_c^2 \\ &= -\frac{1}{n}\mu_2(X) = -\frac{\sigma^2}{n}\{1 - \tau(c')[c' + \tau(c')]\} \\ &= -\frac{\sigma^2}{n}\{[1 + \tau'(c')]\}. \end{aligned} \tag{3.55}$$

3.5.7.2 Calculation of $E(\bar{X} - \mu_c)^2$:

From Chapter 0, sections (0.1.23) and (0.1.12)

$$\begin{aligned} \mu_2(X) &= \sigma^2\{1 - \tau(c')[c' + \tau(c')]\} \\ &= \sigma^2\{[1 + \tau'(c')]\}. \end{aligned} \tag{3.56}$$

Also we have the following conditions

$$\mu_2(\bar{X}) = \frac{\mu_2(X)}{n}. \tag{3.57}$$

Therefore, we have

$$\begin{aligned} E(\bar{X}-\mu_c)^2 &= \frac{\mu_2(X)}{n} \\ &= \frac{\sigma^2}{n}\{1-\tau(c')[c'+\tau(c')]\} \\ &= \frac{\sigma^2}{n}\{[1+\tau'(c')]\}. \end{aligned} \tag{3.58}$$

3.5.7.3 Calculation of $E(s^2-\sigma_c^2)^2$:

We can write

$$\begin{aligned} E(s^2-\sigma_c^2)^2 &= E\{[s^2-E(s^2)]+[E(s^2)-\sigma_c^2]\}^2 \\ &= \text{Var}(s^2)+2b(s^2)E[s^2-E(s^2)]+b^2(s^2) \\ &= \text{Var}(s^2)+b^2(s^2). \end{aligned} \tag{3.59}$$

It is well-known (Kendall & Stuart (1952), p. 233) that

$$\text{Var}\left(\frac{\sum_{i=1}^n(x_i-\bar{x})^2}{n-1}\right) = \frac{\kappa_4(X)}{n}+\frac{2\kappa_2^2(X)}{n-1} \tag{3.60}$$

then

$$\text{Var}(s^2) = \left(\frac{n-1}{n}\right)^2\left[\frac{\mu_4(X)}{n}-\frac{(n-3)\mu_2^2(X)}{n(n-1)}\right]. \tag{3.61}$$

Substituting equations (3.55) and (3.61) into equation (3.59) we obtain

$$\begin{aligned} E(s^2-\sigma_c^2)^2 &= \frac{1}{n^3}[(n-1)^2\mu_4(X)-(n^2-5n+3)\mu_2^2(X)] \\ &= \sigma^4\left\{\frac{(n-1)^2}{n^3}\tau'''(c')+\frac{2n-1}{n^2[1+\tau'(c')]^2}\right\}, \end{aligned} \tag{3.62}$$

where from Chapter 0, sections (1.1.16) and (0.1.23) fourth moment of X,

$$\begin{aligned} \mu_4(X) &= 3\sigma^4(1-c'\tau(c')-\tau^2(c'))^2 \\ &+ \sigma^4[(3c'-c'^3)\tau(c')+(4-7c'^2)\tau^2(c') \\ &- 12c'\tau^3(c')-6\tau^4(c')] \\ &= \tau'''(c')+3[1+\tau'(c')]^2. \end{aligned} \tag{3.63}$$

3.5.7.4 Calculation of $E(\bar{X} - \mu_c)(s^2 - \sigma_c^2)$:

In order to obtain $E(\bar{X} - \mu_c)(s^2 - \sigma_c^2)$ we use the following Lemma

Lemma 3.1 *For $\{a_i\}_1^n$ and $\{b_i\}_1^n$ sequences of real numbers,*

$$[\sum_{i=1}^{n} a_i][\sum_{i=1}^{n} b_i] = \sum_{i=1}^{n} a_i b_i + \sum_{i=1}^{n} \sum_{j \neq i} a_i b_j. \tag{3.64}$$

Let $y_i = x_i - \mu_c$ and $\bar{y} = \bar{x} - \mu_c$ and the y_i's are independent random variables. In the following argument use equation (3.52) and apply Lemma 3.1.

$$\begin{aligned} E(\bar{X} - \mu_c)(s^2 - \sigma_c^2) &= E[\bar{X} - \mu_c]s^2 \\ &= \frac{1}{n^2} E[\sum_{i=1}^{n} y_i (\sum_{i=1}^{n} y_i^2 - n\bar{y}^2)] \\ &= \frac{1}{n^2} E[\sum_{i=1}^{n} y_i^3 + \sum_{i=1}^{n} \sum_{i \neq j} y_i y_j^2 - (n\bar{y})(n\bar{y}^2)] \\ &= \frac{1}{n^2} E[\sum_{i=1}^{n} y_i^3 + \sum_{i=1}^{n} \sum_{i \neq j} y_i y_j^2 - (n^2 \bar{y}^3)] \\ &= \frac{1}{n^2}[n\mu_3(X) + n(n-1)\mu_1(X)\mu_2(X) - n^2 \mu_3(\bar{X})] \\ &= \frac{1}{n^2}[n\mu_3(X) + n(n-1)\mu_1(X)\mu_2(X) - \mu_3(X)] \\ &= \frac{n-1}{n^2}\mu_3(X) = \frac{n-1}{n^2}\left[-\sigma^3 \tau''(c')\right]. \end{aligned} \tag{3.65}$$

Now the first moment of the variable X about its mean is zero and

$$\mu_3(\bar{X}) = \frac{\mu_3(X)}{n^2}. \tag{3.66}$$

Hence, using $\mu_3(X)$ from Chapter 0, section (0.1.23), we find

$$\begin{aligned} E(\bar{X} - \mu_c)(s^2 - \sigma_c^2) &= \frac{-(n-1)\sigma^3[(c'^2 - 1)\tau(c') + 3c'\tau^2(c') + 2\tau^3(c')]}{n^2}. \\ &= \frac{-(n-1)\sigma^3 \tau''(c')}{n^2}. \end{aligned} \tag{3.67}$$

Substituting $E(s^2-\sigma_c^2)$, $E(\bar{X}-\mu_c)^2$, $E(s^2-\sigma_c^2)^2$ and $E(\bar{X}-\mu_c)(s^2-\sigma_c^2)$ from equations (3.55), (3.56), (3.62) and (3.65) into equation (3.51), we obtain

$$\boxed{\begin{aligned} E(\hat{\mu}) &= \mu+\tfrac{1}{n}\Big\{(\tfrac{A_{20}}{2}-A_{01})\mu_2(X)+\tfrac{A_{02}}{2}[\mu_4(X)-\mu_2^2(X)]+A_{11}\mu_3(X)\Big\} \\ &+ O(n^{-2}) \\ &= \mu+\tfrac{\sigma^2}{n}\Big\{(\tfrac{A_{20}}{2}-A_{01})[1+\tau'(c')]+\tfrac{A_{02}}{2}[\tau'''(c')+2[1+\tau'(c')]^2] \\ &\quad -A_{11}\sigma\tau''(c')\Big\}+O(n^{-2}). \end{aligned}} \tag{3.68}$$

3.5.8 Calculation of $E(\hat{\sigma}^2)$:

In this section we derive the theoretical formula for $E(\hat{\sigma}^2)$. Taking the expectation of both sides of equation (3.12) and letting $\bar{x}\to\mu_c$ and $s^2\to\sigma_c^2$ so that $\hat{\mu}\to\mu$ and $\hat{\sigma}^2\to\sigma^2$, we have

$$\begin{aligned} E(\hat{\sigma}^2) &= B_{00}+B_{10}E(\bar{X}-\mu_c)/1!+B_{01}E(s^2-\sigma_c^2)/1!+ \\ &\quad B_{20}E(\bar{X}-\mu_c)^2/2!+B_{02}E(s^2-\sigma_c^2)^2/2!+ \\ &\quad B_{11}E(\bar{X}-\mu_c)(s^2-\sigma_c^2)/1!1!+\cdots. \end{aligned} \tag{3.69}$$

Setting $B_{00}=\sigma^2$ and substituting $E(s^2-\sigma_c^2)$, $E(\bar{X}-\mu_c)^2$, $E(s^2-\sigma_c^2)^2$ and $E(\bar{X}-\mu_c)(s^2-\sigma_c^2)$ from equations (3.55), (3.56), (3.62) and (3.65) into equation (3.69), we obtain

$$\boxed{\begin{aligned} E(\hat{\sigma}^2) &= \sigma^2+\tfrac{1}{n}\Big\{(\tfrac{B_{20}}{2}-B_{01})\mu_2(X)+\tfrac{B_{02}}{2}[\mu_4(X)-\mu_2^2(X)]+B_{11}\mu_3(X)\Big\} \\ &+ O(n^{-2}) \\ &= \sigma^2+\tfrac{\sigma^2}{n}\Big\{(\tfrac{B_{20}}{2}-B_{01})[1+\tau'(c')]+\tfrac{B_{02}}{2}[\tau'''(c')+2[1+\tau'(c')]^2] \\ &- B_{11}\sigma\tau''(c')\Big\}+O(n^{-2}). \end{aligned}} \tag{3.70}$$

3.5.9 Calculation of Var$(\hat{\mu})$:

In this section we find the variance of $\hat{\mu}$, by using the formula

$$\text{Var}(\hat{\mu}) = E(\hat{\mu}-\mu_c)^2-[E(\hat{\mu}-\mu_c)]^2. \tag{3.71}$$

Using $E(\hat{\mu})$, $E(s^2 - \sigma_c^2)$, $E(\bar{X} - \mu_c)^2$, $E(s^2 - \sigma_c^2)^2$ and $E(\bar{X} - \mu_c)(s^2 - \sigma_c^2)$ from equations (3.68), (3.55), (3.56), (3.62) and (3.65) we find the variance of $\hat{\mu}$ up to $O(n^{-1})$,

$$\boxed{\begin{aligned} \mathrm{Var}(\hat{\mu}) &= \tfrac{1}{n}\{A_{10}^2\mu_2(X) + A_{01}^2[\mu_4(X) - \mu_2^2(X)] + 2A_{10}A_{01}\mu_3(X)\} \\ &+ O(n^{-2}) \\ &= \tfrac{\sigma^2}{n}\{A_{10}^2[1+\tau'(c')] + A_{01}^2[\tau'''(c') + 2[1+\tau'(c')]^2] \\ &- 2A_{10}A_{01}\sigma\tau''(c')\} + O(n^{-2}). \end{aligned}} \qquad (3.72)$$

3.5.10 Calculation of Var($\hat{\sigma}^2$):

In this section we derive the variance of $\hat{\sigma}^2$. For this we are using the formula

$$\mathrm{Var}(\hat{\sigma}^2) = E(\hat{\sigma}^2 - \sigma_c^2)^2 - [E(\hat{\sigma}^2 - \sigma_c^2)]^2. \qquad (3.73)$$

Using $E(\hat{\sigma}^2)$, $E(s^2 - \sigma_c^2)$, $E(\bar{X} - \mu_c)^2$, $E(s^2 - \sigma_c^2)^2$ and $E(\bar{X} - \mu_c)(s^2 - \sigma_c^2)$ from equations (3.70), (3.55), (3.56), (3.62) and (3.65), we find the variance of $\hat{\sigma}^2$ up to $O(n^{-1})$,

$$\boxed{\begin{aligned} \mathrm{Var}(\hat{\sigma}^2) &= \tfrac{1}{n}\{B_{10}^2\mu_2(X) + B_{01}^2[\mu_4(X) - \mu_2^2(X)] + 2B_{10}B_{01}\mu_3(X)\} \\ &+ O(n^{-2}) \\ &= \tfrac{\sigma^2}{n}\{B_{10}^2[1+\tau'(c')] + B_{01}^2[\tau'''(c') + 2[1+\tau'(c')]^2] \\ &- 2B_{10}B_{01}\sigma\tau''(c')\} + O(n^{-2}). \end{aligned}} \qquad (3.74)$$

3.5.11 Calculation of Cov($\hat{\mu}, \hat{\sigma}^2$):

In this section we derive the covariance of $\hat{\mu}$ and $\hat{\sigma}^2$. For this we use the formula

$$\mathrm{Cov}(\hat{\mu}, \hat{\sigma}^2) = E\{[(\hat{\mu} - \mu_c) - E(\hat{\mu} - \mu_c)][(\hat{\sigma}^2 - \sigma_c^2) - E(\hat{\sigma}^2 - \sigma_c^2)]\}. \qquad (3.75)$$

Equations (3.68), (3.70), (3.55), (3.56), (3.62) and (3.65) give expressions for $E(\hat{\mu})$, $E(\hat{\sigma}^2)$, $E(s^2 - \sigma_c^2)$, $E(\bar{X} - \mu_c)^2$, $E(s^2 - \sigma_c^2)^2$ and $E(\bar{X} - \mu_c)(s^2 - \sigma_c^2)$ the covariance of $\hat{\mu}$ and $\hat{\sigma}^2$ up to $O(n^{-1})$, namely

$$
\begin{array}{rcl}
\operatorname{Cov}(\hat{\mu},\hat{\sigma}^2) & = & \frac{1}{n}\{A_{10}B_{10}\mu_2(X) + A_{01}B_{01}\left[\mu_4(X) - \mu_2^2(X)\right] \\
 & + & (A_{10}B_{01} + A_{01}B_{10})\mu_3(X)\} + O(n^{-2}) \\
 & = & \frac{\sigma^2}{n}\{A_{10}B_{10}[1+\tau'(c')] + A_{01}B_{01}[\tau'''(c') + 2[1+\tau'(c')]^2]\} \\
 & - & (A_{10}B_{01} + A_{01}B_{01})\sigma\tau''(c')\} + O(n^{-2}).
\end{array} \quad (3.76)
$$

To compare the theoretical results of $E(\hat{\mu})$, $\sigma(\hat{\mu})$, $E(\hat{\sigma}^2)$, $\sigma(\hat{\sigma}^2)$ and $\rho(\hat{\mu},\hat{\sigma}^2)$ with the simulation study, we use a computer program (see Appendix Program 19) to calculate the expected values for different sample sizes $n = 5, 10, 20, 50, 100$ and different truncation points $c = -1.88, -1, 0, 1, 3, 10$. The results are presented in Tables 3.8 $-$ 3.13.

Table 3.8: The theoretical results for the expected value and standard deviation of $\hat{\mu}$ and $\hat{\sigma}^2$ and $\rho(\hat{\mu},\hat{\sigma}^2)$, for different values of n when $c = -1.88$, using Shenton & Bowman methods, $\mu = 0$, $\sigma = 1$

n	Mean		Variance		$\rho(\hat{\mu},\hat{\sigma}^2)$
	$E(\hat{\mu})$	$\sigma(\hat{\mu})$	$E(\hat{\sigma}^2)$	$\sigma(\hat{\sigma}^2)$	
5	35.07	11.58	2261.23	4.55	0.99
10	17.54	8.19	1131.11	3.21	0.99
20	8.77	5.79	566.06	2.27	0.99
50	3.51	3.66	227.02	1.44	0.99
100	1.75	2.59	114.01	1.02	0.99

Table 3.9: The theoretical results for the expected value and standard deviation of $\hat{\mu}$ and $\hat{\sigma}^2$ and $\rho(\hat{\mu}, \hat{\sigma}^2)$, for different values of n when $c = -1$, using Shenton & Bowman methods, $\mu = 0$, $\sigma = 1$

n	Mean		Variance		$\rho(\hat{\mu}, \hat{\sigma}^2)$
	$E(\hat{\mu})$	$\sigma(\hat{\mu})$	$E(\hat{\sigma}^2)$	$\sigma(\hat{\sigma}^2)$	
5	9.62	5.52	252.76	2.99	0.98
10	4.81	3.91	126.88	2.11	0.98
20	2.40	2.76	63.94	1.49	0.98
50	0.96	1.74	26.18	0.94	0.98
100	0.48	1.23	13.59	0.67	0.98

Table 3.10: For $c = 0$

n	Mean		Variance		$\rho(\hat{\mu}, \hat{\sigma}^2)$
	$E(\hat{\mu})$	$\sigma(\hat{\mu})$	$E(\hat{\sigma}^2)$	$\sigma(\hat{\sigma}^2)$	
5	1.83	2.10	14.54	1.80	0.93
10	0.91	1.49	7.77	1.27	0.93
20	0.46	1.05	4.39	0.90	0.93
50	0.18	0.66	2.35	0.57	0.93
100	0.091	0.47	1.68	0.40	0.93

Comparing Tables 3.8-3.10 with 3.2-3.4 we notice that for small values of n the bias is large while for increasing n the bias is reduced. This may be due to the fact that our Taylor expansions provide us with good approximations when n is large, but, when n is small, converge rather slowly so that obtain good accuracy we require more terms in the Cox & Hinkley and Shenton & Bowman expansions.

Table 3.11: The theoretical results for the expected value and standard deviation of $\hat{\mu}$ and $\hat{\sigma}^2$ and $\rho(\hat{\mu}, \hat{\sigma}^2)$, for different values of n when $c = 1$, using Shenton & Bowman methods, $\mu = 0$, $\sigma = 1$

n	Mean		Variance		$\rho(\hat{\mu}, \hat{\sigma}^2)$
	$E(\hat{\mu})$	$\sigma(\hat{\mu})$	$E(\hat{\sigma}^2)$	$\sigma(\hat{\sigma}^2)$	
5	0.34	0.81	1.54	0.22	0.72
10	0.17	0.57	1.27	0.78	0.72
20	0.084	0.40	1.13	0.55	0.72
50	0.033	0.26	1.05	0.35	0.72
100	0.017	0.18	1.02	0.25	0.72

Table 3.12: For $c = 3$

n	Mean		Variance		$\rho(\hat{\mu}, \hat{\sigma}^2)$
	$E(\hat{\mu})$	$\sigma(\hat{\mu})$	$E(\hat{\sigma}^2)$	$\sigma(\hat{\sigma}^2)$	
5	0.016	0.45	0.85	0.66	0.033
10	0.0079	0.32	0.92	0.46	0.033
20	0.0039	0.23	0.96	0.33	0.033
50	0.0016	0.14	0.98	0.21	0.033
100	0.00079	0.10	0.99	0.15	0.033

Table 3.13: The theoreticals results for the expected value and standard deviation of $\hat{\mu}$ and $\hat{\sigma}^2$ and $\rho(\hat{\mu}, \hat{\sigma}^2)$, for different values of n when $c = 10$, using Shenton & Bowman methods, $\mu = 0$, $\sigma = 1$

n	Mean		Variance		$\rho(\hat{\mu}, \hat{\sigma}^2)$
	$E(\hat{\mu})$	$\sigma(\hat{\mu})$	$E(\hat{\sigma}^2)$	$\sigma(\hat{\sigma}^2)$	
5	0.38×10^{-19}	0.45	0.80	0.63	0.55×10^{-20}
10	0.19×10^{-19}	0.32	0.90	0.45	0.55×10^{-20}
20	0.94×10^{-20}	0.22	0.95	0.32	0.55×10^{-20}
50	0.38×10^{-20}	0.14	0.98	0.20	0.55×10^{-20}
100	0.19×10^{-20}	0.10	0.99	0.14	0.55×10^{-20}

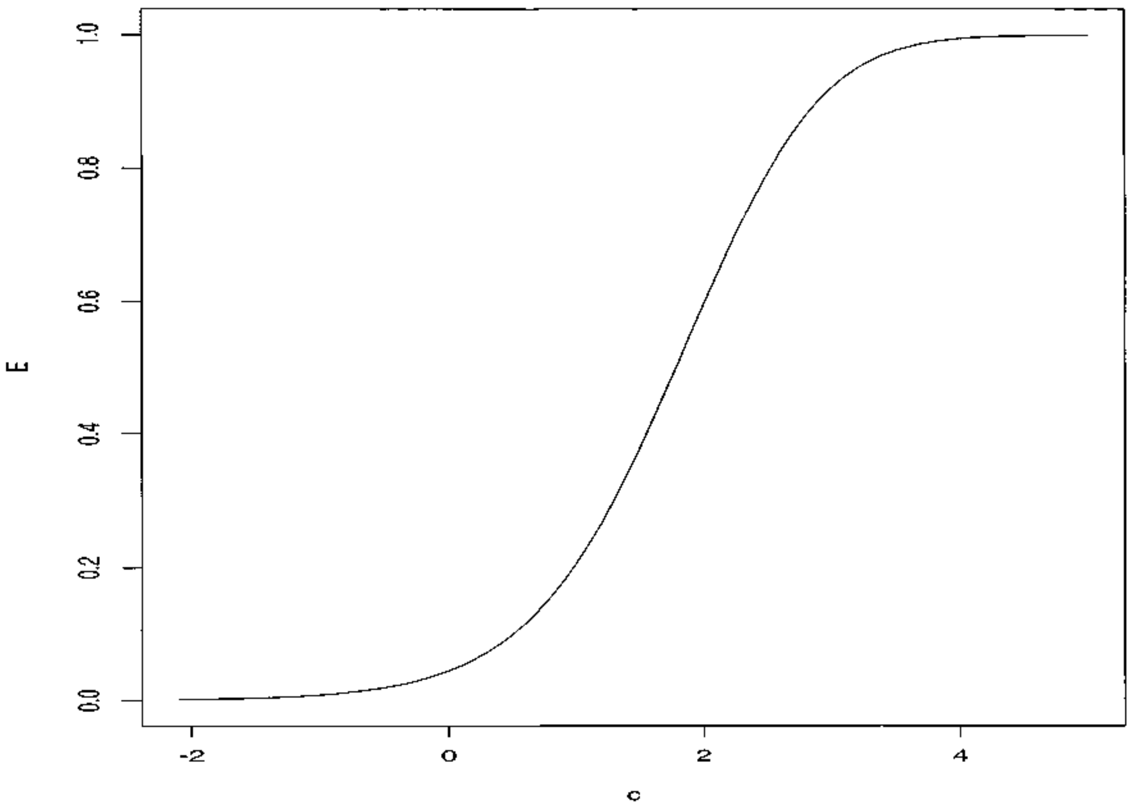
3.5.12 Conclusion:

The Figure 3.5 shows a plot of E (calculated in equation 3.27) against c, from which we see that E is a monotonic increasing function of c. E becomes small when $c \leq 0$ and approaches 1 as $c \to \infty$. From the Figure 3.5 we see that the bias in $\hat{\mu}$ and $\hat{\sigma}^2$ may will be large for $c \leq 0$ and possibly for some positive values (Program 20 which plot the E against c is given in Appendix).

The expected values of $\hat{\mu}$ and $\hat{\sigma}^2$, given in equations (3.68) and (3.70), involve A and B terms (defined in section (3.5.2)) whose denominators are functions of E defined in equation (3.27). Consequently, E should have an important bearing on the biasedness of $\hat{\mu}$ and $\hat{\sigma}^2$.

As noted, the amount of bias is large for small values of c. The smaller the value of c, the larger is the difference between the values in the two tables 3.2 and 3.8. For large c

Figure 3.5: The plot of E against the truncation point c



for example, when $c = 3$ the Tables 3.6 and 3.12 have almost identical values. For $c = 10$, Tables 3.7 and 3.13 show that the two methods provide identical values for $E(\hat{\mu})$, $\sigma(\hat{\mu})$ and $E(\hat{\sigma}^2)$, and that the values for $\sigma(\hat{\sigma}^2)$ are very close. Furthermore, for the same value of c, Table 3.7 shows that the simulated values are quite close to the theoretical ones. Hence, for small values of c, to obtain better approximations of the theoretical results expansion to further terms is needed, which requires the evaluation of A_{30}, A_{03}, B_{30}, B_{03} etc.

3.6 Theoretical results based on expansions in terms of $(\bar{x} - \mu_c)$ and $(s - \sigma_c)$ (Method B):

An alternative approach is the expansion of $\hat{\mu}$ and $\hat{\sigma}$ based on $(\bar{x} - \mu_c)$ and $(s - \sigma_c)$. Its difference from Method A is that in Method B, $(\bar{x}-\mu_c)$ and $(s-\sigma_c)$ are of the same dimension whereas in Method A, $(\bar{x} - \mu_c)$ and $(s^2 - \sigma_c^2)$ are not. For the joint estimation of the two parameters μ and σ, we extend the results of Shenton & Bowman formula as following:

Expansion of Equations: Using the extension of Shenton & Bowman's results, we can expand the equations (3.6) in the following way.

Assume that

$$\begin{aligned}\hat{\mu} &= A'_{00} + A'_{10}(\bar{x} - \mu_c)/1! + A'_{01}(s - \sigma_c)/1! + \\ &\quad A'_{20}(\bar{x} - \mu_c)^2/2! + A'_{02}(s - \sigma_c)^2/2! + \\ &\quad A'_{11}(\bar{x} - \mu_c)(s - \sigma_c)/1!1! + \cdots \end{aligned} \tag{3.77}$$

and

$$\begin{aligned}\hat{\sigma} &= B'_{00} + B'_{10}(\bar{x} - \mu_c)/1! + B'_{01}(s - \sigma_c)/1! + \\ &\quad B'_{20}(\bar{x} - \mu_c)^2/2! + B'_{02}(s - \sigma_c)^2/2! + \\ &\quad B'_{11}(\bar{x} - \mu_c)(s - \sigma_c)/1!1! + \cdots . \end{aligned} \tag{3.78}$$

We need to know

1. $A'_{10} = \frac{\partial \hat{\mu}}{\partial \bar{x}} \mid_{(\bar{x}=\mu_c, s=\sigma_c)}$,
2. $A'_{01} = \frac{\partial \hat{\mu}}{\partial s} \mid_{(\bar{x}=\mu_c, s=\sigma_c)}$,
3. $A'_{20} = \frac{\partial^2 \hat{\mu}}{\partial (\bar{x})^2} \mid_{(\bar{x}=\mu_c, s=\sigma_c)}$,
4. $A'_{02} = \frac{\partial^2 \hat{\mu}}{\partial (s)^2} \mid_{(\bar{x}=\mu_c, s=\sigma_c)}$,

5. $A'_{11} = \frac{\partial^2 \hat{\mu}}{\partial \bar{x} \partial s} \mid_{(\bar{x}=\mu_c, s=\sigma_c)}$,

6. $B'_{10} = \frac{\partial \hat{\sigma}}{\partial \bar{x}} \mid_{(\bar{x}=\mu_c, s=\sigma_c)}$,

7. $B'_{01} = \frac{\partial \hat{\sigma}}{\partial s} \mid_{(\bar{x}=\mu_c, s=\sigma_c)}$,

8. $B'_{20} = \frac{\partial^2 \hat{\sigma}}{\partial (\bar{x})^2} \mid_{(\bar{x}=\mu_c, s=\sigma_c)}$,

9. $B'_{02} = \frac{\partial^2 \hat{\sigma}}{\partial (s)^2} \mid_{(\bar{x}=\mu_c, s=\sigma_c)}$,

10. $B'_{11} = \frac{\partial^2 \hat{\sigma}}{\partial \bar{x} \partial s} \mid_{(\bar{x}=\mu_c, s=\sigma_c)}$ etc.

According to Theorem 3.2 as $\bar{x} \to \mu_c$ and $s \to \sigma_c$, then $\hat{\mu} \to \mu$ and $\hat{\sigma} \to \sigma$. Therefore from the equations (3.77) and (3.78), it can be concluded that $A'_{00} = A_{00} = \mu$ and $B'_{00} = \sigma$.

By taking the first and second partial derivatives of each equation and using Theorem 3.2, we can find the remaining coefficients.

3.6.1 Preliminary calculation:

In this section we calculate some of the expressions which we will need later on.

3.6.1.1 Calculation of $\frac{\partial \hat{c}'}{\partial s}$:

Let

$$\hat{c}' = (c - \hat{\mu})\hat{\sigma}^{-1}. \tag{3.79}$$

Taking the partial derivative of $\hat{c}'$ with respect to s we obtain

$$\begin{aligned} \frac{\partial \hat{c}'}{\partial s} &= \frac{\partial (c-\hat{\mu})}{\partial \hat{\mu}} . \frac{\partial \hat{\mu}}{\partial s} \hat{\sigma}^{-1} + \frac{\partial \hat{\sigma}^{-1}}{\partial \hat{\sigma}} . \frac{\partial \hat{\sigma}}{\partial s} (c - \hat{\mu}) \\ &= [-\hat{\sigma}^{-1} \frac{\partial \hat{\mu}}{\partial s} - \hat{\sigma}^{-2} (c - \hat{\mu}) \frac{\partial \hat{\sigma}}{\partial s}]. \end{aligned} \tag{3.80}$$

3.6.1.2 Calculation of $\frac{\partial \hat{c}'}{\partial \bar{x}}$:

Similarly, taking the partial derivative of $\hat{c}'$ with respect to $\bar{x}$ we obtain

$$\frac{\partial \hat{c}'}{\partial \bar{x}} = -\hat{\sigma}^{-1}\left(\frac{\partial \hat{\mu}}{\partial \bar{x}} + \hat{c}'\frac{\partial \hat{\sigma}}{\partial \bar{x}}\right). \tag{3.81}$$

3.6.2 Calculation of A'_{10} and B'_{10}:

Now taking the partial derivative of the first equation in (3.6) with respect to $\bar{x}$, we have

$$\begin{aligned} \frac{\partial \hat{\mu}}{\partial \bar{x}} &= 1 + \frac{\partial \hat{\sigma}}{\partial \bar{x}}\tau(\hat{c}') + \hat{\sigma}\frac{\partial \tau(\hat{c}')}{\partial \bar{x}} \\ &= 1 + \tau(\hat{c}')\frac{\partial \hat{\sigma}}{\partial \bar{x}} + \hat{\sigma}\frac{\partial \tau(\hat{c}')}{\partial \hat{c}'}\cdot\frac{\partial \hat{c}'}{\partial \bar{x}} \end{aligned} \tag{3.82}$$

which , on replacing $\frac{\partial \hat{c}'}{\partial \bar{x}}$ from equation (3.81) and $\tau'(\hat{c}')$ for $\frac{\partial \tau(\hat{c}')}{\partial \hat{c}'}$, gives

$$[1 + \tau'(\hat{c}')]\frac{\partial \hat{\mu}}{\partial \bar{x}} + [\hat{c}'\tau'(\hat{c}') - \tau(\hat{c}')]\frac{\partial \hat{\sigma}}{\partial \bar{x}} = 1. \tag{3.83}$$

Taking the partial derivative of the equation (3.19) with respect to $\bar{x}$, we have

$$(\bar{x} - c)\frac{\partial \hat{\mu}}{\partial \bar{x}} + 2\hat{\sigma}\frac{\partial \hat{\sigma}}{\partial \bar{x}} = 2\bar{x} - c - \hat{\mu}. \tag{3.84}$$

Solving the (3.83) and (3.84) simultaneously

gives

$$\frac{\partial \hat{\mu}}{\partial \bar{x}} = \frac{2\hat{\sigma} - (2\bar{x} - c - \hat{\mu})[\hat{c}'\tau'(\hat{c}') - \tau(\hat{c}')]}{2\hat{\sigma}\hat{E}} \tag{3.85}$$

and

$$\frac{\partial \hat{\sigma}}{\partial \bar{x}} = \frac{c - \bar{x} + (2\bar{x} - c - \hat{\mu})[1 + \tau'(\hat{c}')]}{2\hat{\sigma}\hat{E}}. \tag{3.86}$$

Now, according to Theorem 3.2, as $\bar{x} \to \mu_c$ and $s \to \sigma_c$, then $\hat{\mu} \to \mu$ and $\hat{\sigma} \to \sigma$. Therefore we find

$$\begin{aligned} A'_{10} = \frac{\partial \hat{\mu}}{\partial \bar{x}} \mid_{(\bar{x}=\mu_c, s=\sigma_c)} &= \frac{1 + [c'\tau'(c') - \tau(c')][c' + 2\tau(c')]/2}{E} \\ &= A_{10} \end{aligned} \tag{3.87}$$

and

$$B'_{10} = \frac{\partial \hat{\sigma}}{\partial \bar{x}} \mid_{(\bar{x}=\mu_c, s=\sigma_c)} = \frac{-\{\tau(c')[1 + \tau'(c')] + \tau'(c')[c' + \tau(c')]\}}{2E} = \frac{B_{10}}{2\sigma} \tag{3.88}$$

3.6.3 Calculation of A'_{01} and B'_{01}:

Now, taking the partial derivative of the first equation in (3.6) with respect to s, we have

$$\begin{aligned} \frac{\partial \hat{\mu}}{\partial s} &= \tau(\hat{c}')\frac{\partial \hat{\sigma}}{\partial s} + \hat{\sigma}\frac{\partial \tau(\hat{c}')}{\partial s} \\ &= \tau(\hat{c}')\frac{\partial \hat{\sigma}}{\partial s} + \hat{\sigma}\frac{\partial \tau(\hat{c}')}{\partial \hat{c}'}.\frac{\partial \hat{c}'}{\partial s} \end{aligned} \tag{3.89}$$

which , on replacing $\frac{\partial \hat{c}'}{\partial s}$ from equation (3.80) and $\tau'(\hat{c}')$ for $\frac{\partial \tau(\hat{c}')}{\partial \hat{c}'}$, gives

$$[1 + \tau'(\hat{c}')]\frac{\partial \hat{\mu}}{\partial s} + [\hat{c}'\tau'(\hat{c}') - \tau(\hat{c}')]\frac{\partial \hat{\sigma}}{\partial s} = 0. \tag{3.90}$$

Taking the partial derivative of the equation (3.19) with respect to s, we have

$$-\hat{\sigma}(\tau(\hat{c}') + \hat{c}')\frac{\partial \hat{\mu}}{\partial s} + 2\hat{\sigma}\frac{\partial \hat{\sigma}}{\partial s} = 2s. \tag{3.91}$$

Solving (3.90) and (3.91) simultaneously gives

$$\frac{\partial \hat{\mu}}{\partial s} = -\frac{s[\hat{c}'\tau'(\hat{c}') - \tau(\hat{c}')]}{\hat{\sigma}\hat{E}} \tag{3.92}$$

and

$$\frac{\partial\hat{\sigma}}{\partial s} = \frac{s[1+\tau'(\hat{c}')]}{\hat{\sigma}\hat{E}}, \tag{3.93}$$

where $\hat{E}$ is as defined in equation (3.24).

Now, using Theorem 3.2, as $\bar{x} \to \mu_c$ and $s \to \sigma_c$, then $\hat{\mu} \to \mu$ and $\hat{\sigma} \to \sigma$, we find

$$\begin{aligned} A'_{01} = \frac{\partial\hat{\mu}}{\partial s}\mid_{(\bar{x}=\mu_c, s=\sigma_c)} &= -\frac{[c'\tau'(c') - \tau(c')][1+\tau'(c')]^{1/2}}{E} \\ &= 2\sigma[1+\tau'(c')]^{1/2} A_{01} \end{aligned} \tag{3.94}$$

and

$$B'_{01} = \frac{\partial\hat{\sigma}}{\partial s}\mid_{(\bar{x}=\mu_c, s=\sigma_c)} = \frac{[1+\tau'(c')]^{3/2}}{E}, \tag{3.95}$$

where E is as defined in equation (3.27).

3.6.4 Calculation of A'_{20} and B'_{20}:

In this section we calculate A'_{20} and B'_{20}. Taking the partial derivatives of equations (3.83) and (3.84) with respect to $\bar{x}$, we obtain

$$\begin{cases} [1+\tau'(\hat{c}')]\frac{\partial^2\hat{\mu}}{\partial\bar{x}^2} + [\hat{c}'\tau'(\hat{c}') - \tau(\hat{c}')]\frac{\partial^2\hat{\sigma}}{\partial\bar{x}^2} = \frac{\tau''(\hat{c}')}{\hat{\sigma}}[\frac{\partial\hat{\mu}}{\partial\bar{x}} + \hat{c}'\frac{\partial\hat{\sigma}}{\partial\bar{x}}]^2 \\ (\bar{x}-c)\frac{\partial^2\hat{\mu}}{\partial\bar{x}^2} + 2\hat{\sigma}\frac{\partial^2\hat{\sigma}}{\partial\bar{x}^2} = 2[1 - \frac{\partial\hat{\mu}}{\partial\bar{x}} - (\frac{\partial\hat{\sigma}}{\partial\bar{x}})^2]. \end{cases} \tag{3.96}$$

Solving the equations (3.96) simultaneously, we find

$$\frac{\partial^2\hat{\mu}}{\partial\bar{x}^2} = \frac{\tau''(\hat{c}')[\frac{\partial\hat{\mu}}{\partial\bar{x}} + \hat{c}'\frac{\partial\hat{\sigma}}{\partial\bar{x}}]^2 - [\hat{c}'\tau'(\hat{c}') - \tau(\hat{c}')]\left[1 - \frac{\partial\hat{\mu}}{\partial\bar{x}} - (\frac{\partial\hat{\sigma}}{\partial\bar{x}})^2\right]}{\hat{\sigma}\hat{E}} \tag{3.97}$$

and

$$\frac{\partial^2\hat{\sigma}}{\partial\bar{x}^2} = \frac{2[1+\tau'(\hat{c}')][1 - \frac{\partial\hat{\mu}}{\partial\bar{x}} - (\frac{\partial\hat{\sigma}}{\partial\bar{x}})^2] - \frac{(\bar{x}-c)\tau''(\hat{c}')}{\hat{\sigma}}[\frac{\partial\hat{\mu}}{\partial\bar{x}} + \hat{c}'\frac{\partial\hat{\sigma}}{\partial\bar{x}}]^2}{2\hat{\sigma}\hat{E}}. \tag{3.98}$$

Now, using Theorem 3.2,

$$A'_{20} = \frac{1}{\sigma E}\left\{\tau''(c')[A'_{10} + c'B'_{10}]^2 - [c'\tau'(c') - \tau(c')][1 - A'_{10} - (B'_{10})^2]\right\} \quad (3.99)$$

and

$$B'_{20} = \frac{1}{2\sigma E}\left\{\tau''(c')[c' + \tau(c')][A'_{10} + c'B'_{10}]^2 + 2[1 + \tau'(c')][1 - A'_{10} - (B'_{10})^2]\right\} \quad (3.100)$$

or, in terms of c', $\tau(c')$, $\tau'(c')$ and $\tau''(c')$,

$$\begin{aligned} A'_{20} &= \frac{1}{2\sigma E^3}\left\{4\tau''(c')[1 + \tau'(c')]^2 + [c'\tau'(c') - \tau(c')][\tau(c')(1 + \tau'(c')) + \tau'(c')(c' + \tau(c'))]^2\right\} \\ &+ \frac{[c'\tau'(c') - \tau(c')]}{2\sigma E^2}\{[2 + (c'\tau'(c') - \tau(c'))(c' + 2\tau(c'))]\} - \frac{[c'\tau'(c') - \tau(c')]}{\sigma E} \end{aligned}$$

and

$$\begin{aligned} B'_{20} &= \frac{[1 + \tau'(c')]}{2\sigma E^3}\left\{2\tau''(c')[c' + \tau(c')][1 + \tau'(c')] - [\tau(c')(1 + \tau'(c')) + \tau'(c')(c' + \tau(c'))]^2\right\} \\ &- \frac{[1 + \tau'(c')]}{2\sigma E^2}\{[2 + (c'\tau'(c') - \tau(c'))(c' + 2\tau(c'))]\} + \frac{[1 + \tau'(c')]}{\sigma E}. \end{aligned}$$

3.6.5 Calculation of A'_{02} and B'_{02}:

In this section we calculate A'_{02} and B'_{02}. Taking the partial derivatives of equations (3.90) and (3.91) with respect to s we obtain

$$\begin{cases} [1 + \tau'(\hat{c}')]\frac{\partial^2\hat{\mu}}{\partial(s)^2} + [\hat{c}'\tau'(\hat{c}') - \tau(\hat{c}')]\frac{\partial^2\hat{\sigma}}{\partial(s)^2} = \frac{\tau''(\hat{c}')}{\hat{\sigma}}[\frac{\partial\hat{\mu}}{\partial s} + \hat{c}'\frac{\partial\hat{\sigma}}{\partial s}]^2 \\ -\hat{\sigma}(\tau(\hat{c}') + \hat{c}')\frac{\partial^2\hat{\mu}}{\partial(s)^2} + 2\hat{\sigma}\frac{\partial^2\hat{\sigma}}{\partial(s)^2} = 2[1 - (\frac{\partial\hat{\sigma}}{\partial s})^2]. \end{cases} \quad (3.101)$$

Solving the equations (3.101) simultaneously, gives

$$\frac{\partial^2\hat{\mu}}{\partial(s)^2} = \frac{1}{\hat{\sigma}\hat{E}}\left\{\tau''(\hat{c}')[\frac{\partial\hat{\mu}}{\partial s} + \hat{c}'\frac{\partial\hat{\sigma}}{\partial s}]^2 - [\hat{c}'\tau'(\hat{c}') - \tau(\hat{c}')][1 - (\frac{\partial\hat{\sigma}}{\partial s})^2]\right\} \quad (3.102)$$

and

$$\frac{\partial^2\hat{\sigma}}{\partial(s)^2} = \frac{1}{2\hat{\sigma}\hat{E}}\left\{2[1 + \tau'(\hat{c}')][1 - (\frac{\partial\hat{\sigma}}{\partial s})^2] - (\bar{x} - c)\frac{\tau''(c')}{\sigma}[\frac{\partial\hat{\mu}}{\partial s} + \hat{c}'\frac{\partial\hat{\sigma}}{\partial s}]^2\right\}. \quad (3.103)$$

Using Theorem 3.2 we can find

$$A'_{02} = \frac{1}{\sigma E}\left\{\tau''(c')(A'_{01} + c'B'_{01})^2 - [c'\tau'(c') - \tau(c')][1 - (B'_{01})^2]\right\} \tag{3.104}$$

and

$$B'_{02} = \frac{1}{2\sigma E}\left\{2[1 + \tau'(c')][1 - (B'_{01})^2] + \tau''(c')[c' + \tau(c')](A'_{01} + c'B'_{01})^2\right\} \tag{3.105}$$

or, expressed in c', $\tau(c')$, $\tau'(c')$ and $\tau''(c')$,

$$A'_{02} = \frac{1}{\sigma E}\left\{\frac{\tau''(c')[1 + \tau'(c')][c' + \tau(c')]^2}{E^2} - [c'\tau'(c') - \tau(c')]\left[1 - \frac{[1 + \tau'(c')]^3}{E^2}\right]\right\}$$

and

$$B'_{02} = \frac{[1 + \tau'(c')]}{2\sigma E}\left\{2\left[1 - \frac{[1 + \tau'(c')]^3}{E^2}\right] + \frac{\tau''(c')[c' + \tau(c')]^3}{E^2}\right\}.$$

3.6.6 Calculation of A'_{11} and B'_{11}:

To calculate A'_{11} and B'_{11} we take the derivatives of equations (3.83) and (3.84) with respect to s and obtain

$$\begin{cases} [1 + \tau'(\hat{c}')]\frac{\partial^2 \hat{\mu}}{\partial s \partial \bar{x}} + [\hat{c}'\tau'(\hat{c}') - \tau(\hat{c}')]\frac{\partial^2 \hat{\sigma}}{\partial s \partial \bar{x}} = \frac{\tau''(\hat{c}')}{\hat{\sigma}}[\frac{\partial \hat{\mu}}{\partial \bar{x}} + \hat{c}'\frac{\partial \hat{\sigma}}{\partial \bar{x}}][\frac{\partial \hat{\mu}}{\partial s} + \hat{c}'\frac{\partial \hat{\sigma}}{\partial s}] \\ -\hat{\sigma}(\tau(\hat{c}') + \hat{c}')\frac{\partial^2 \hat{\mu}}{\partial s \partial \bar{x}} + 2\hat{\sigma}\frac{\partial^2 \hat{\sigma}}{\partial s \partial \bar{x}} = -\frac{\partial \hat{\mu}}{\partial s} - 2\frac{\partial \hat{\sigma}}{\partial s}\frac{\partial \hat{\sigma}}{\partial \bar{x}}. \end{cases} \tag{3.106}$$

Solving the equations (3.106) simultaneously, gives

$$\begin{aligned} \frac{\partial^2 \hat{\mu}}{\partial s \partial \bar{x}} &= \frac{1}{2\hat{\sigma}\hat{E}}\left\{2\tau''(\hat{c}')\left[\frac{\partial \hat{\mu}}{\partial \bar{x}} + \hat{c}'\frac{\partial \hat{\sigma}}{\partial \bar{x}}\right]\left[\frac{\partial \hat{\mu}}{\partial s} + \hat{c}'\frac{\partial \hat{\sigma}}{\partial s}\right]\right. \\ &+ \left.[\hat{c}'\tau'(\hat{c}') - \tau(\hat{c}')]\left[\frac{\partial \hat{\mu}}{\partial s} + 2\frac{\partial \hat{\sigma}}{\partial s}\frac{\partial \hat{\sigma}}{\partial \bar{x}}\right]\right\} \end{aligned} \tag{3.107}$$

and

$$\begin{aligned} \frac{\partial^2 \hat{\sigma}}{\partial s \partial \bar{x}} &= -\frac{1}{2\hat{\sigma}\hat{E}}\left\{-\hat{\sigma}(\tau(\hat{c}') + \hat{c}')\frac{\tau''(\hat{c}')}{\hat{\sigma}}\left[\frac{\partial \hat{\mu}}{\partial \bar{x}} + \hat{c}'\frac{\partial \hat{\sigma}}{\partial \bar{x}}\right]\left[\frac{\partial \hat{\mu}}{\partial s} + \hat{c}'\frac{\partial \hat{\sigma}}{\partial s}\right]\right. \\ &+ \left.[1 + \tau'(\hat{c}')]\left[\frac{\partial \hat{\mu}}{\partial s} + 2\frac{\partial \hat{\sigma}}{\partial s}\frac{\partial \hat{\sigma}}{\partial \bar{x}}\right]\right\}. \end{aligned} \tag{3.108}$$

Using Theorem 3.2, we find

$$\begin{aligned} A'_{11} &= \frac{1}{2\sigma E}\{2\tau''(c')[A'_{01} + c'B'_{01}][A'_{10} + c'B'_{10}] \\ &+ [c'\tau'(c') - \tau(c')][A'_{01} + 2B'_{10}B'_{01}]\} \end{aligned} \tag{3.109}$$

and

$$\begin{aligned} B'_{11} &= \frac{1}{2\sigma E}\{\tau''(c')[c' + \tau(c')][A'_{01} + c'B'_{01}][A'_{10} + c'B'_{10}] \\ &- [1 + \tau'(\hat{c}')][A'_{01} + 2B'_{10}B'_{01}]\}. \end{aligned} \tag{3.110}$$

Rewriting, in terms of c', $\tau(c')$, $\tau'(c')$ and $\tau''(c')$, we have

$$\begin{aligned} A'_{11} &= \frac{[1 + \tau'(c')]^{3/2}}{2\sigma E^3}\{2\tau''(c')[c' + \tau(c')] \\ &- [c'\tau'(c') - \tau(c')][\tau(c')(1 + \tau'(c')) + \tau'(c')(c' + \tau(c'))]\} \\ &- \frac{[c'\tau'(c') - \tau(c')]^2[1 + \tau'(c')]^{1/2}}{2\sigma E^2} \end{aligned}$$

and

$$\begin{aligned} B'_{11} &= \frac{[1 + \tau'(c')]^{3/2}}{2\sigma E^3}\left\{\tau''(c')[c' + \tau(c')]^2\right. \\ &+ [1 + \tau'(c')][\tau(c')(1 + \tau'(c')) + \tau'(c')(c' + \tau(c'))]\} \\ &+ \frac{[1 + \tau'(c')]^{3/2}[c'\tau'(c') - \tau(c')]}{2\sigma E^2}. \end{aligned}$$

To make sure that our calculations are correct, A'_{11} and B'_{11} were calculated by the alternative method of differentiating firstly with respect to s.

3.6.7 Calculation of $E(\hat{\mu})$:

In this section we derive $E(\hat{\mu})$. Taking the expectation of both sides of equation (3.77) and letting $\bar{x} \to \mu_c$, $s \to \sigma_c$, $\hat{\mu} \to \mu$ and $\hat{\sigma} \to \sigma$, we have

$$E(\hat{\mu}) = A'_{00} + A'_{10}E(\bar{X} - \mu_c)/1! + A'_{01}E(s - \sigma_c)/1! +$$

$$A'_{20}E(\bar{X}-\mu_c)^2/2! + A'_{02}E(s-\sigma_c)^2/2! +$$

$$A'_{11}E(\bar{X}-\mu_c)(s-\sigma_c)/1!1! + \cdots$$

$$= \mu + \Big\{A'_{01}[E(s-\sigma_c)] + A'_{20}E(\bar{X}-\mu_c)^2/2 +$$

$$A'_{02}E(s-\sigma_c)^2/2 + A'_{11}E(\bar{X}-\mu_c)(s-\sigma_c)\Big\} + \cdots, \quad (3.111)$$

since $E(\bar{X}-\mu_c) = 0$.

3.6.7.1 Calculation of $E(s-\sigma_c)^2$:

It is well-known (Kendall & Stuart (1952, p. 233)) that up to $O(n^{-1})$

$$\text{Var}(s) = \frac{\mu_4(X) - \mu_2^2(X)}{4n\mu_2(X)}. \quad (3.112)$$

Now the bias in s is in $O(n^{-1})$ (see the next section). Therefore

$$E(s-\sigma_c)^2 = \frac{\mu_4(X) - \mu_2^2(X)}{4n\mu_2(X)} + O(n^{-2}). \quad (3.113)$$

Substituting $\mu_2(X)$ and $\mu_4(X)$ from equations (3.56) and (3.63) into equation (3.113), we obtain

$$E(s-\sigma_c)^2 = \frac{\tau'''(c') + 2[1+\tau'(c')]^2}{4n[1+\tau'(c')]} + O(n^{-2}). \quad (3.114)$$

3.6.7.2 Calculation of $E(s-\sigma_c)$:

We know that

$$\text{Var}(s) = E(s^2) - [E(s)]^2. \quad (3.115)$$

Substituting $\text{Var}(s)$ and $E(s^2)$ from equations (3.112) and (3.54) into equation (3.115) we obtain

$$E(s) = \left[\frac{(4n-3)\mu_2^2(X) - \mu_4(X)}{4n\mu_2(X)}\right]^{1/2}. \quad (3.116)$$

On expanding $E(s)$, we find

$$E(s) = \sqrt{\mu_2(X)} + \frac{1}{8n\sqrt{\mu_2(X)}}\left(-3\mu_2(X) - \frac{\mu_4(X)}{\mu_2(X)}\right) + O(n^{-2}). \tag{3.117}$$

Since, in this case, $\sigma_c = \sqrt{\mu_2(X)}$, we have

$$\begin{aligned} E(s-\sigma_c) &= -\frac{1}{8n\sqrt{\mu_2(X)}}\left(3\mu_2(X) + \frac{\mu_4(X)}{\mu_2(X)}\right) + O(n^{-2}) \\ &= -\frac{\sigma\left\{\tau'''(c') + 6[1+\tau'(c')]^2\right\}}{8n[1+\tau'(c')]^{3/2}} + O(n^{-2}). \end{aligned} \tag{3.118}$$

3.6.7.3 Calculation of $E(\bar{X}-\mu_c)(s-\sigma_c)$:

From Kendall & Stuart (1952, p. 233), up to $O(n^{-1})$,

$$\mathrm{Cov}(\bar{X}, s) = \frac{\mu_3(X)}{2n\sqrt{\mu_2(X)}}. \tag{3.119}$$

Therefore, we can find

$$E(\bar{X}-\mu_c)(s-\sigma_c) = \frac{\mu_3(X)}{2n\sqrt{\mu_2(X)}} + O(n^{-2}). \tag{3.120}$$

From the results in Chapter 0, section (0.1.23), we have

$$E(\bar{X}-\mu_c)(s-\sigma_c) = \frac{-\sigma^2\tau''(c')}{2n\sqrt{1+\tau'(c')}} + O(n^{-2}). \tag{3.121}$$

Substituting $E(s-\sigma_c)$ from (3.118), $E(\bar{X}-\mu_c)^2$ from (3.58), $E(s-\sigma_c)^2$ from (3.114) and $E(\bar{X}-\mu_c)(s-\sigma_c)$ from (3.121) into equation (3.111), we obtain

$$\boxed{\begin{aligned} E(\hat{\mu}) &= \mu + \frac{1}{n}\left\{-\frac{A'_{01}[\mu_4(X)+3\mu_2^2(X)]}{8\mu_2^{\frac{3}{2}}(X)} + \frac{A'_{20}\mu_2(X)}{2}\right. \\ &+ \left.\frac{A'_{02}[\mu_4(X)-\mu_2^2(X)]}{8\mu_2(X)} + \frac{A'_{11}\mu_3(X)}{2\sqrt{\mu_2(X)}}\right\} + O(n^{-2}) \\ &= \mu + \frac{\sigma^2}{n}\left\{-\frac{A'_{01}[\tau'''(c')+6(1+\tau'(c'))^2]}{8\sigma[1+\tau'(c')]^{\frac{3}{2}}} + \frac{A'_{20}[1+\tau'(c')]}{2}\right. \\ &+ \left.\frac{A'_{02}[\tau'''(c')+2(1+\tau'(c'))^2]}{8[1+\tau'(c')]} - \frac{A'_{11}\tau''(c')}{2\sqrt{1+\tau'(c')}}\right\} + O(n^{-2}). \end{aligned}} \tag{3.122}$$

3.6.8 Calculation of $E(\hat{\sigma})$:

In this section we derive the theoretical formula for $E(\hat{\sigma})$. Take the expectation of both sides of equation (3.78) and let $\bar{x} \to \mu_c$, $s \to \sigma_c$, $\hat{\mu} \to \mu$ and $\hat{\sigma} \to \sigma$. Then setting $B'_{00} = \sigma$ and using $E(s - \sigma_c)$ from (3.118), $E(\bar{X} - \mu_c)^2$ from (3.58), $E(s - \sigma_c)^2$ from (3.114) and $E(\bar{X} - \mu_c)(s - \sigma_c)$ from (3.121), we obtain

$$\boxed{\begin{aligned} E(\hat{\sigma}) &= \sigma + \frac{1}{n}\left\{-\frac{B'_{01}[\mu_4(X)+3\mu_2^2(X)]}{8\mu_2^{\frac{3}{2}}(X)} + \frac{B'_{20}\mu_2(X)}{2}\right. \\ &+ \left.\frac{B'_{02}[\mu_4(X)-\mu_2^2(X)]}{8\mu_2(X)} + \frac{B'_{11}\mu_3(X)}{2\sqrt{\mu_2(X)}}\right\} + O(n^{-2}) \\ &= \sigma + \frac{\sigma^2}{n}\left\{-\frac{B'_{01}[\tau'''(c')+6(1+\tau'(c'))^2]}{8\sigma[1+\tau'(c')]^{\frac{3}{2}}} + \frac{B'_{20}[1+\tau'(c')]}{2}\right. \\ &+ \left.\frac{B'_{02}[\tau'''(c')+2(1+\tau'(c'))^2]}{8[1+\tau'(c')]} - \frac{B'_{11}\tau''(c')}{2\sqrt{1+\tau'(c')}}\right\} + O(n^{-2}). \end{aligned}} \tag{3.123}$$

3.6.9 Calculation of Var$(\hat{\mu})$:

In this section we find the variance of $\hat{\mu}$. To do this using the formula:

$$\mathrm{Var}(\hat{\mu}) = E(\hat{\mu} - \mu_c)^2 - [E(\hat{\mu} - \mu_c)]^2. \tag{3.124}$$

Taking $E(\hat{\mu})$ from (3.122), $E(s - \sigma_c)$ from (3.118), $E(\bar{X} - \mu_c)^2$ from (3.58), $E(s - \sigma_c)^2$ from (3.114) and $E(\bar{X} - \mu_c)(s - \sigma_c)$ from (3.121), we find the variance of $\hat{\mu}$, up to $O(n^{-1})$.

$$\boxed{\begin{aligned} \mathrm{Var}(\hat{\mu}) &= \frac{1}{n}\left\{(A'_{10})^2\mu_2(X) + \frac{(A'_{01})^2[\mu_4(X)-\mu_2^2(X)]}{4\mu_2(X)} + \frac{A'_{10}A'_{01}\mu_3(X)}{\sqrt{\mu_2(X)}}\right\} + O(n^{-2}) \\ &= \frac{1}{n}\left\{(A'_{10})^2[1+\tau'(c')] + \frac{(A'_{01})^2\{\tau'''(c')+2[1+\tau'(c')]^2\}}{4[1+\tau'(c')]} - \frac{A'_{10}A'_{01}\tau''(c')}{\sqrt{1+\tau'(c')}}\right\} \\ &+ O(n^{-2}). \end{aligned}} \tag{3.125}$$

Using the relationships between A'_{10} and A_{10} from equation (3.87) and between A'_{01} and A_{01} from equation (3.94) it can be shown that formula for Var$(\hat{\mu})$ in equation (3.125) (Method B) is identical with that given in equation (3.72) (Method A).

3.6.10 Calculation of Var($\hat{\sigma}$):

In this section we derive the variance of $\hat{\sigma}$, through the formula

$$\mathrm{Var}(\hat{\sigma}) = E(\hat{\sigma}-\sigma_c)^2 - [E(\hat{\sigma}-\sigma_c)]^2. \tag{3.126}$$

Using $E(\hat{\sigma})$ from (3.123), $E(s-\sigma_c)$ from (3.118), $E(\bar{X}-\mu_c)^2$ from (3.58), $E(s-\sigma_c)^2$ from (3.114) and $E(\bar{X}-\mu_c)(s-\sigma_c)$ from (3.121), we find the variance of $\hat{\sigma}$ up to $O(n^{-1})$

$$\begin{aligned}
\mathrm{Var}(\hat{\sigma}) &= \frac{1}{n}\left\{(B'_{10})^2\mu_2(X) + \frac{(B'_{01})^2[\mu_4(X)-\mu_2^2(X)]}{4\mu_2(X)} + \frac{B'_{10}B'_{01}\mu_3(X)}{\sqrt{\mu_2(X)}}\right\} + O(n^{-2}) \\
&= \frac{1}{n}\left\{(B'_{10})^2[1+\tau'(c')] + \frac{(B'_{01})^2\{\tau'''(c')+2[1+\tau'(c')]^2\}}{4[1+\tau'(c')]} - \frac{B'_{10}B'_{01}\tau''(c')}{\sqrt{1+\tau'(c')}}\right\} \\
&+ O(n^{-2}).
\end{aligned} \tag{3.127}$$

3.6.11 Calculation of Cov($\hat{\mu}, \hat{\sigma}$):

To derive the covariance of $\hat{\mu}$ and $\hat{\sigma}$, we apply the formula

$$\mathrm{Cov}(\hat{\mu},\hat{\sigma}) = E\{[(\hat{\mu}-\mu_c) - E(\hat{\mu}-\mu_c)][(\hat{\sigma}-\sigma_c) - E(\hat{\sigma}-\sigma_c)]\}. \tag{3.128}$$

Using $E(\hat{\mu})$ from (3.122), $E(\hat{\sigma})$ from (3.123), $E(s-\sigma_c)$ from (3.118), $E(\bar{X}-\mu_c)^2$ from (3.58), $E(s-\sigma_c)^2$ from (3.114) and $E(\bar{X}-\mu_c)(s-\sigma_c)$ from (3.121) , we obtain the following form for covariance of $\hat{\mu}$ and $\hat{\sigma}$, up to $O(n^{-1})$:

$$\begin{aligned}
\mathrm{Cov}(\hat{\mu},\hat{\sigma}) &= \frac{1}{n}\left\{A'_{10}B'_{10}\mu_2(X) + \frac{A'_{01}B'_{01}[\mu_4(X)-\mu_2^2(X)]}{4\mu_2(X)}\right. \\
&+ \left.\frac{[A'_{10}B'_{01}+A'_{01}B'_{10}]\mu_3(X)}{2\sqrt{\mu_2(X)}}\right\} + O(n^{-2}) \\
&= \frac{1}{n}\left\{A'_{10}B'_{10}[1+\tau'(c')] + \frac{A'_{01}B'_{10}\{\tau'''(c')+2[1+\tau'(c')]^2\}}{4[1+\tau'(c')]}\right. \\
&- \left.\frac{[A'_{10}B'_{01}+A'_{01}B'_{10}]\tau''(c')}{2\sqrt{1+\tau'(c')}}\right\} + O(n^{-2}).
\end{aligned} \tag{3.129}$$

To compare the theoretical results of $E(\hat{\mu})$, $\sigma(\hat{\mu})$, $E(\hat{\sigma})$, $\sigma(\hat{\sigma})$ and $\rho(\hat{\mu},\hat{\sigma})$ with the results from a simulation study, we use a computer program (see Appendix Program 21) to calculate the expected values for sample sizes $n = 5, 10, 20, 50, 100$ and truncation points $c = -1.88, -1, 0, 1, 3, 10$.

Table 3.14: The theoretical results for the expected value and standard deviation of $\hat{\mu}$ and $\hat{\sigma}$ and $\rho(\hat{\mu}, \hat{\sigma}^2)$, for different values of n when $c = -1.88$, using Shenton & Bowman's method, $\mu = 0$, $\sigma = 1$

n	$E(\hat{\mu})$	$\sigma(\hat{\mu})$	$E(\hat{\sigma})$	$\sigma(\hat{\sigma})$	$\rho(\hat{\mu}, \hat{\sigma})$	$E(\hat{\sigma}^2)$
5	781.06	11.58	5.11	2.27	0.71	31.26
10	390.53	8.19	3.06	1.61	0.71	11.95
20	195.26	5.79	2.02	1.13	0.71	5.35
50	78.11	3.66	1.41	0.71	0.71	2.49
100	39.05	2.59	1.21	0.51	0.71	1.72

Table 3.15: For $c = -1$

n	$E(\hat{\mu})$	$\sigma(\hat{\mu})$	$E(\hat{\sigma})$	$\sigma(\hat{\sigma})$	$\rho(\hat{\mu}, \hat{\sigma})$	$E(\hat{\sigma}^2)$
5	109.93	5.52	2.31	1.49	0.59	7.56
10	54.96	3.91	1.65	1.06	0.59	3.58
20	27.48	2.76	1.32	0.75	0.59	2.30
50	10.99	1.74	1.13	0.47	0.59	1.50
100	5.49	1.23	1.06	0.33	0.59	1.23

Table 3.16: The theoretical results for the expected value and standard deviation of $\hat{\mu}$ and $\hat{\sigma}$ and $\rho(\hat{\mu}, \hat{\sigma}^2)$, for different values of n when $c = 0$, using Shenton & Bowman methods, $\mu = 0$, $\sigma = 1$

n	$E(\hat{\mu})$	$\sigma(\hat{\mu})$	$E(\hat{\sigma})$	$\sigma(\hat{\sigma})$	$\rho(\hat{\mu}, \hat{\sigma})$	$E(\hat{\sigma}^2)$
5	9.35	2.10	1.22	0.89	0.44	2.28
10	4.67	1.49	1.11	0.63	0.44	1.63
20	2.33	1.05	1.05	0.45	0.44	1.31
50	0.93	0.66	1.02	0.28	0.44	1.12
100	0.46	0.47	1.01	0.20	0.44	1.06

Table 3.17: For $c = 1$

n	$E(\hat{\mu})$	$\sigma(\hat{\mu})$	$E(\hat{\sigma})$	$\sigma(\hat{\sigma})$	$\rho(\hat{\mu}, \hat{\sigma})$	$E(\hat{\sigma}^2)$
5	0.74	0.81	0.96	0.55	0.35	1.22
10	0.37	0.57	0.98	0.39	0.35	1.11
20	0.19	0.40	0.99	0.27	0.35	1.05
50	0.074	0.26	0.99	0.17	0.35	1.00
100	0.037	0.18	0.99	0.12	0.35	0.99

Comparing Tables 3.14- 3.16 with 3.2-3.4 shows that their disparities are high for small values of n and c. It is possibly due to slow convergence of underlying Taylor expansions.

Table 3.18: The theoretical results for the expected value and standard deviation of $\hat{\mu}$ and $\hat{\sigma}$ and $\rho(\hat{\mu}, \hat{\sigma}^2)$, for different values of n when $c = 3$, using Shenton & Bowman methods, $\mu = 0$, $\sigma = 1$

n	$E(\hat{\mu})$	$\sigma(\hat{\mu})$	$E(\hat{\sigma})$	$\sigma(\hat{\sigma})$	$\rho(\hat{\mu}, \hat{\sigma})$	$E(\hat{\sigma}^2)$
5	0.016	0.45	0.87	0.32	0.032	0.86
10	0.0080	0.32	0.93	0.23	0.032	0.92
20	0.0040	0.23	0.97	0.16	0.032	0.97
50	0.0016	0.14	0.99	0.10	0.032	0.99
100	0.00080	0.10	0.99	0.073	0.032	99

Table 3.19: For $c = 10$

n	$E(\hat{\mu})$	$\sigma(\hat{\mu})$	$E(\hat{\sigma})$	$\sigma(\hat{\sigma})$	$\rho(\hat{\mu}, \hat{\sigma})$	$E(\hat{\sigma}^2)$
5	0.38×10^{-19}	0.45	0.85	0.32	0.54×10^{-20}	0.82
10	0.19×10^{-19}	0.32	0.92	0.22	0.54×10^{-20}	0.89
20	0.94×10^{-20}	0.22	0.96	0.16	0.54×10^{-20}	0.95
50	0.38×10^{-20}	0.14	0.98	0.10	0.54×10^{-20}	0.97
100	0.19×10^{-20}	0.10	0.99	0.071	0.54×10^{-20}	0.98

3.6.12 Conclusion:

In comparison of Tables 3.14-3.19 related to the results of expansion method B with the Tables 3.8-3.13 in method A we can say that the $E(\hat{\mu})$ in method A is less than $E(\hat{\mu})$ in method B. Moreover the comparing of the $E(\hat{\sigma}^2)$ with its counterpart we can say that $E(\hat{\sigma}^2)$ in method B is less than method A.

In comparison Tables 3.2-3.7 with Tables 3.8-3.13 (Method A) and Tables 3.14-3.19 (Method B) we can see that for $c < 3$ the $E(\hat{\mu})$ in simulation has a significant difference with its counterpart in methods A and B. But by increasing the truncation points c and sample size n, the values of $E(\hat{\mu})$, $\sigma(\hat{\mu})$ and $E(\hat{\sigma}^2)$ are approximately the same for the simulation method and methods A and B.

This can be explained as follows.

(a) The $\hat{\mu}$, $\hat{\sigma}^2$ and $\hat{\sigma}$ expansion were performed only up to second term.

(b) The values of $E(\hat{\mu})$, $\text{Var}(\hat{\mu})$, $E(\hat{\sigma}^2)$, $\text{Var}(\hat{\sigma}^2)$, $E(\hat{\sigma})$, $\text{Var}(\hat{\sigma})$, $\text{Cov}(\hat{\mu}, \hat{\sigma})$ and $\text{Cov}(\hat{\mu}, \hat{\sigma}^2)$ are calculated up to $O(n^{-1})$. But, by increasing the truncation points c and sample size n these differences disappear.

In equations (3.122) and (3.123), giving the expected values of $\hat{\mu}$ and $\hat{\sigma}$ respectively, we see that the denominators of the coefficients A' and B' terms, are all functions of the term E defined in section (3.5.2). Consequently, E should have an important bearing on the biasedness, the theoretical results for $\hat{\mu}$ and $\hat{\sigma}$ that we have derived.

Chapter 4

The one parameter case of maximum product spacing estimator for the truncated normal distribution:

4.1 Introduction:

The method of maximum likelihood estimation offers estimators which are sufficient , consistent and efficient, rendering it one of the best methods of parameter estimation. Cheng & Amin (1982) suggested the maximum product spacing (MPS) method for some distributions, such as the uniform, lognormal etc. They pointed out that the main properties of MPS estimation are:

1. "MPS estimation gives consistent estimation under more general conditions than ML estimation. In particular it gives consistent estimators when the underlying distribution is J-shaped (parameter is shifted origin to the right in lognormal, Weibull and gamma distribution), a situation were ML estimation is bound to

fail.

2. MPS estimators are asymptotically normal and asymptotically are as efficient as ML estimators. In some situations ML estimates are known to exist that are "hyper-efficient" (in the sense of having variance less than the usual order n^{-1}) e.g. the maximum likelihood estimator of θ in the uniform distribution over $(0, \theta)$. MPS estimators are then also hyper-efficient. In J-shaped distributions, where ML estimation breaks down, MPS estimation still gives efficient estimators."

They suggested that the MPS method can be applied to any continuous univariate distribution with density function $f(x, \theta)$ and cumulative distribution function $F(x, \theta)$ (It is assumed that $f(x, \theta)$ is strictly positive in the interval (α_1, α_2)). They showed that, if θ is the true parameter value and $y_1 < y_2 < \ldots < y_n$ is an ordered sample of size n drawn from the density function $f(x, \theta)$, by using the transformation $z_i = F(y_i, \theta)$, $i = 0, 1, \ldots, n+1$, where $y_0 = \alpha_1$ and $y_{n+1} = \alpha_2$, and maximizing the geometric mean of the spacings

$$G = \{\Pi_{i=1}^{n+1} D_i\}^{\frac{1}{n+1}} \tag{4.1}$$

or its equivalent

$$H = (n+1)^{-1} \ln(G), \tag{4.2}$$

where

$$D_i = z_i - z_{i-1} = \int_{y_{i-1}}^{y_i} f(y, \theta) dy; \qquad i = 1, 2, \ldots, n+1, \tag{4.3}$$

the MPS estimator can be found. Thus the formal definition of MPS is

Definition 4.1 *The estimator, $\tilde{\theta}$, which maximizes the geometric mean G or its equivalent $H = \ln(G)$, is called the MPS estimator of θ.*

Two years later Ranneby (1984) showed that if the distribution function $F(x, \theta)$ is not too "heavy-tailed", and MPS estimator is $\tilde{\theta}$, $\sqrt{n}(\tilde{\theta}-\theta)$ convergences to a normal distribution with mean zero and a variance which is given by the lower bound in the Cramer-Rao inequality. They showed that, in some situations, the MPS method gives consistent estimates but the ML method does not. Also by the simulation, they demonstrated that the MPS estimate converges faster than ML estimate. To find the rules for choosing between the MPS estimate and the ML estimate when they are asymptotically equivalent, one needs to know more about the small sample properties which are discussed in Ranneby (1984).

4.2 Estimation of μ when σ is known:

In this section, we estimate μ by the MPS method , for the truncated normal distribution when σ is known. We then compare $\tilde{\mu}$ with $\hat{\mu}$ found in section (2.2).

From Chapter 2, equation (2.1) we know that

$$f(x, \mu) = \frac{\phi(\frac{x-\mu}{\sigma})}{\sigma\Phi(\frac{c-\mu}{\sigma})}. \tag{4.4}$$

Putting $f(x, \mu)$ from equation (4.4) into equation (4.3) we have, for $y_0, y_1, \ldots, y_{n+1}$

$$\begin{aligned} D_i &= z_i - z_{i-1} = \int_{y_{i-1}}^{y_i} f(y, \mu)dy; \qquad i = 1, 2, \ldots, n+1 \\ &= \frac{1}{\sigma}\int_{-\infty}^{y_i} \frac{\phi(\frac{t-\mu}{\sigma})}{\Phi(\frac{c-\mu}{\sigma})}dt - \frac{1}{\sigma}\int_{-\infty}^{y_{i-1}} \frac{\phi(\frac{t-\mu}{\sigma})}{\Phi(\frac{c-\mu}{\sigma})}dt \\ &= \frac{\Phi(\frac{y_i-\mu}{\sigma}) - \Phi(\frac{y_{i-1}-\mu}{\sigma})}{\Phi(\frac{c-\mu}{\sigma})}, \end{aligned} \tag{4.5}$$

where $y_0, y_1, \ldots, y_{n+1}$ are the order statistics of sample $x_0, x_1, \ldots, x_{n+1}$, and $\Phi(y_0), \Phi(y_1), \ldots, \Phi(y_{n+1})$ are their corresponding cumulative distribution functions.

Therefore

$$\ln(D_i) = \ln[\Phi(\frac{y_i - \mu}{\sigma}) - \Phi(\frac{y_{i-1} - \mu}{\sigma})] - \ln\Phi(\frac{c - \mu}{\sigma}). \tag{4.6}$$

Hence, from (4.1) and (4.2),

$$\begin{aligned} H = \ln(G) &= (n+1)^{-1}\{\ln(\Pi_{i=1}^{n+1} D_i)\} \\ &= (n+1)^{-1}\{\sum_{i=1}^{n+1} \ln(D_i)\} \\ &= (n+1)^{-1}\{\sum_{i=1}^{n+1} \ln[\Phi(\frac{y_i-\mu}{\sigma}) - \Phi(\frac{y_{i-1}-\mu}{\sigma})]\} - \ln\Phi(\frac{c-\mu}{\sigma}). \end{aligned} \tag{4.7}$$

Taking the derivative of H with respect to μ, we obtain

$$\frac{\partial H}{\partial \mu} = -\frac{1}{\sigma(n+1)}\{\sum_{i=1}^{n+1}[\frac{\phi(\frac{y_i-\mu}{\sigma}) - \phi(\frac{y_{i-1}-\mu}{\sigma})}{\Phi(\frac{y_i-\mu}{\sigma}) - \Phi(\frac{y_{i-1}-\mu}{\sigma})}]\} + \frac{\phi(\frac{c-\mu}{\sigma})}{\sigma\Phi(\frac{c-\mu}{\sigma})}. \tag{4.8}$$

The algebraic solution of equation (4.8) is impossible, so it has to be found numerically. For this we use the NAG routine C05AGF in our program (see Appendix Program 22).

Theorem 4.1 *The MPS estimator ($\tilde{\mu}$) is asymptotically a sufficient, consistent and efficient estimator of μ.*

Proof: Consider

$$\begin{aligned} \ln(D_i) &= \ln[\int_{y_{i-1}}^{y_i} f(y,\mu)dy] \qquad ,i = 1,2,\ldots,n+1, \\ &= \ln[f(y_i,\mu)(y_i - y_{i-1})] + R(y_i, y_{i-1}, \mu) \\ &= \ln f(y_i,\mu) + \ln(y_i - y_{i-1}) + R(y_i, y_{i-1}, \mu) \end{aligned} \tag{4.9}$$

where

$$R(y_i, y_{i-1}, \mu) = \ln\left\{\frac{\sigma[\Phi(\frac{y_i-\mu}{\sigma}) - \Phi(\frac{y_{i-1}-\mu}{\sigma})]}{(y_i - y_{i-1})[\phi(\frac{y_i-\mu}{\sigma})]}\right\}.$$

Therefore, substituting $\ln(D_i)$ into H we obtain

$$\begin{aligned} H = \ln(G) &= (n+1)^{-1}\{\ln(\Pi_{i=1}^{n+1} D_i)\} \\ &= (n+1)^{-1}\{\sum_{i=1}^{n+1} \ln(D_i)\} \\ &= (n+1)^{-1}\{\sum_{i=1}^{n+1}[\ln f(y_i,\mu) + \ln(y_i - y_{i-1}) + R(y_i, y_{i-1}, \mu)]\}. \end{aligned} \tag{4.10}$$

Since, $R(y_i, y_{i-1}, \mu)$ is dependent on μ, using the proof of Cheng & Amin (1982), MPS and ML estimation are asymptotically equal and have the same asymptotic sufficiency, consistency and efficiency properties.

It can be shown that for $n = 1$ and $n = 2$ the MPS estimator $\tilde{\mu}$ approaches the sample mean $\bar{x}$ as $c \to \infty$.

Theorem 4.2 *Let $X_1, X_2, \ldots X_n$ be identically and independently distributed random variables with p.d.f. $f(x, \mu)$, and let the transformation ϕ be one-to-one. Then a MPS estimate is invariant under one-to-one transformation ϕ.*

Proof: Suppose $\tilde{\mu}$ is a MPS estimate of μ, we now prove that $\phi(\tilde{\mu})$ is a MPS estimate of $\phi(\mu)$.

Let us set $\mu^* = \phi(\mu)$, hence $\mu = \phi^{-1}(\mu^*)$. Then

$$H(\mu) = H(\phi^{-1}(\mu^*)).$$

Let $H(\mu) = H(\phi^{-1}(\mu^*)) = H^*(\mu^*)$. It follows that

$$\max_{\mu}[H(\mu)] = \max_{\mu^*}[H^*(\mu^*)]. \tag{4.11}$$

If we assume that a MPS estimate exists, then the term on the left-hand side of equation (4.11) attains its maximum at $\tilde{\mu}$. It follows then that the right-hand side attains its maximum at $\tilde{\mu^*}$, where $\tilde{\mu^*} = \phi(\tilde{\mu})$. Therefore $\phi(\tilde{\mu})$ is a MPS estimate of $\phi(\mu)$.

4.2.1 The MPS estimator of mean in data sets 1 and 2:

1. Consider the data set 1 and let $\sigma = 1$ and $\varepsilon = 10^{-5}$, where ε is the maximum of the absolute value of the difference between the iterated value and the solution of the equation. Then we find that the MPS estimate of μ is

$$\tilde{\mu} = -0.4248.$$

Let $HP = \frac{\partial H}{\partial \mu}$. On plotting H and HP against a certain range of values μ in [-2, 2], we get Figures 4.1 and 4.2 (see Appendix Program 23). We are plotting HP against c to see that $HP = 0$ when H is maximum.

2. Using the data set 2 and letting $\sigma = 1$ and $\varepsilon = 10^{-5}$, we find that the MPS estimate of μ is

$$\tilde{\mu} = -0.2032.$$

When we compare the MPS estimator with the ML estimator, we can see that for the two data sets of Chapter 1, the difference between the MPS estimate and the exact value ($\mu = 0$) is bigger than corresponding difference for the ML estimate.

4.2.2 The MPS estimator of the mean in ideal samples:

In this section we prove that the MPS estimators of μ for the ideal samples are zero and also plot H and HP against different values of μ.

Theorem 4.3 *In ideal samples, for every truncation point c the MPS estimator $\tilde{\mu}$ is zero.*

Proof: From Chapter 1 we know that, for y_i a variable of ideal sample,

$$F(y_i) = \frac{\Phi(\frac{y_i-\mu}{\sigma})}{\Phi(\frac{c-\mu}{\sigma})} = \frac{i}{n+1},$$

and hence

$$\Phi(\frac{y_i - \mu}{\sigma}) = \frac{i}{n+1}\Phi(\frac{c-\mu}{\sigma}). \tag{4.12}$$

Figure 4.1: H versus μ for data set 1 (boys) ($\sigma = 1$)

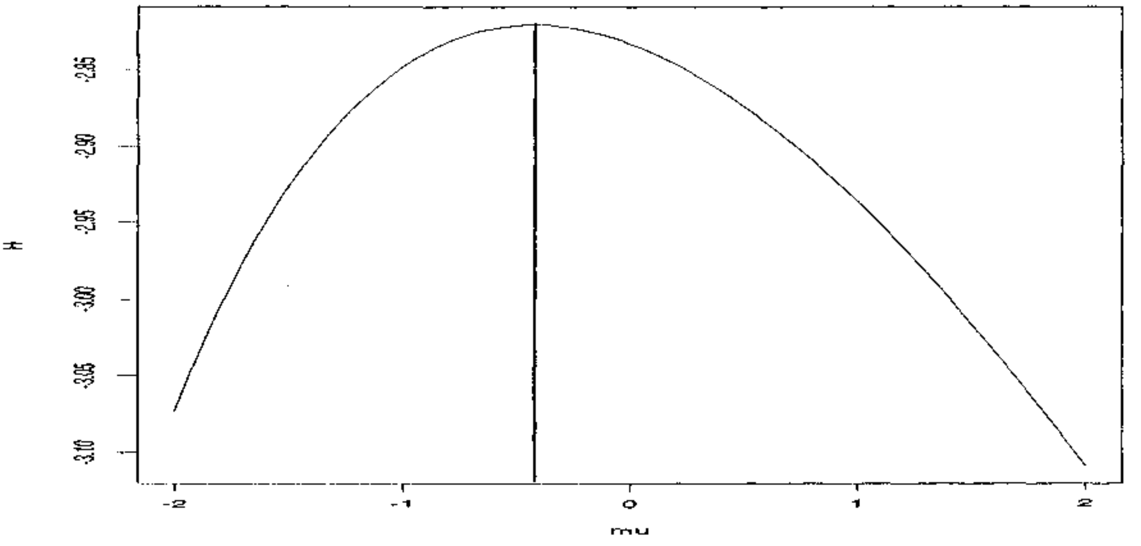
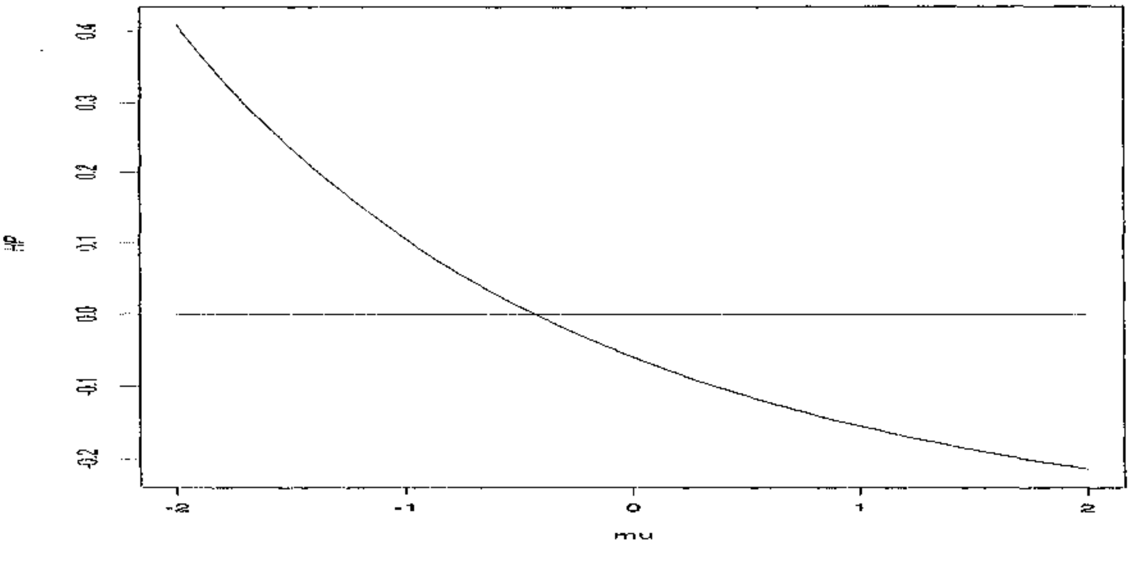


Figure 4.2: HP versus μ for data set 1 (boys) ($\sigma = 1$)



Now, with the application of the formula (4.12), consider

$$\begin{aligned}
\sum_{i=1}^{n+1}[\frac{\phi(\frac{y_i-\mu}{\sigma})-\phi(\frac{y_{i-1}-\mu}{\sigma})}{\Phi(\frac{y_i-\mu}{\sigma})-\Phi(\frac{y_{i-1}-\mu}{\sigma})}] &= (n+1)\sum_{i=1}^{n+1}[\frac{\phi(\frac{y_i-\mu}{\sigma})-\phi(\frac{y_{i-1}-\mu}{\sigma})}{\Phi(\frac{c-\mu}{\sigma})}] \\
&= \frac{n+1}{\Phi(\frac{c-\mu}{\sigma})}[\phi(\frac{y_1-\mu}{\sigma})-\phi(\frac{y_0-\mu}{\sigma})+\cdots+\phi(\frac{y_{n+1}-\mu}{\sigma})-\phi(\frac{y_n-\mu}{\sigma})] \\
&= \frac{n+1}{\Phi(\frac{c-\mu}{\sigma})}[\phi(\frac{y_{n+1}-\mu}{\sigma})-\phi(\frac{y_0-\mu}{\sigma})] \\
&= \frac{(n+1)\phi(\frac{c-\mu}{\sigma})}{\Phi(\frac{c-\mu}{\sigma})}. \qquad (4.13)
\end{aligned}$$

Using (4.13), it follows that

$$\frac{\partial H}{\partial \mu}\mid_{\mu=\tilde{\mu}}=-\frac{1}{(n+1)}\{\sum_{i=1}^{n+1}[\frac{\phi(\frac{y_i-\tilde{\mu}}{\sigma})-\phi(\frac{y_{i-1}-\tilde{\mu}}{\sigma})}{\Phi(\frac{y_i-\tilde{\mu}}{\sigma})-\Phi(\frac{y_{i-1}-\tilde{\mu}}{\sigma})}]\}+\frac{\phi(\frac{c-\tilde{\mu}}{\sigma})}{\Phi(\frac{c-\tilde{\mu}}{\sigma})}=0.$$

Consequently, $\tilde{\mu}=0$ maximizes H, and the theorem is proved.

For the ideal sample of size 5 the plot of H and HP against a certain range of values μ [-2, 2] are shown in Figures 4.3 and 4.4.

4.3 Simulation study to estimate the mean when the variance is known:

In order to compare the expected value, standard deviation and variance of the MPS estimator with those of the ML estimator we embark upon a simulation study.

4.3.1 The simulation study:

The Program 24, given in the Appendix, has been written to calculate $E(\hat{\mu})$, Var$(\tilde{\mu})$ and $\sigma(\hat{\mu})$, for the number of iterations $R = 10000$ and different sample sizes n = 5, 10, 20, 50 and 100. In this program we use the NAG routine G05DDF(0,1) to generate random deviates from the normal distribution with mean zero and variance one and also use the Program 22

Figure 4.3: H versus μ for the ideal sample of size 5 ($c = -1.88$, $\sigma = 1$)

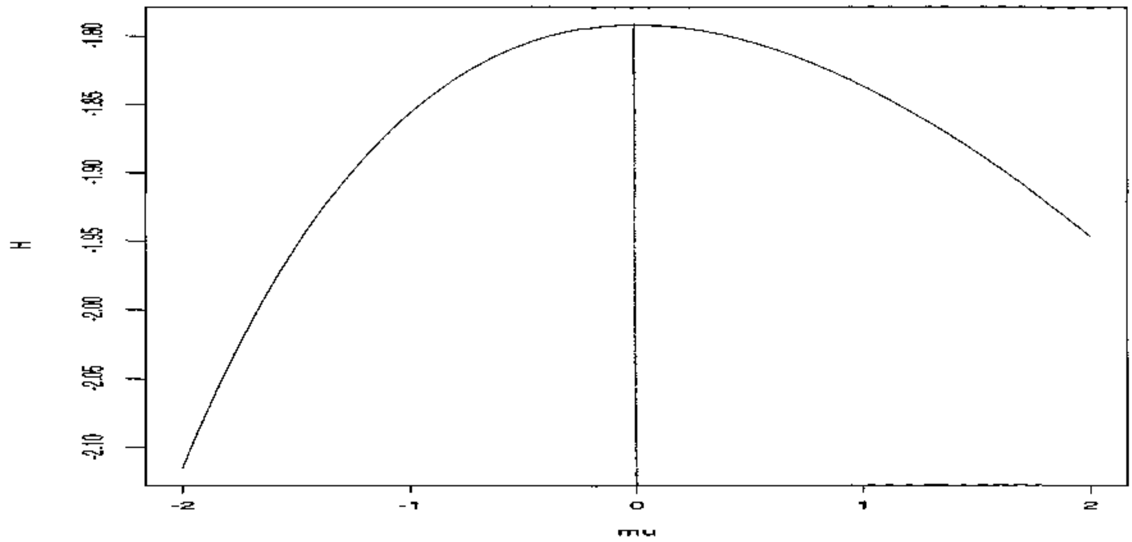
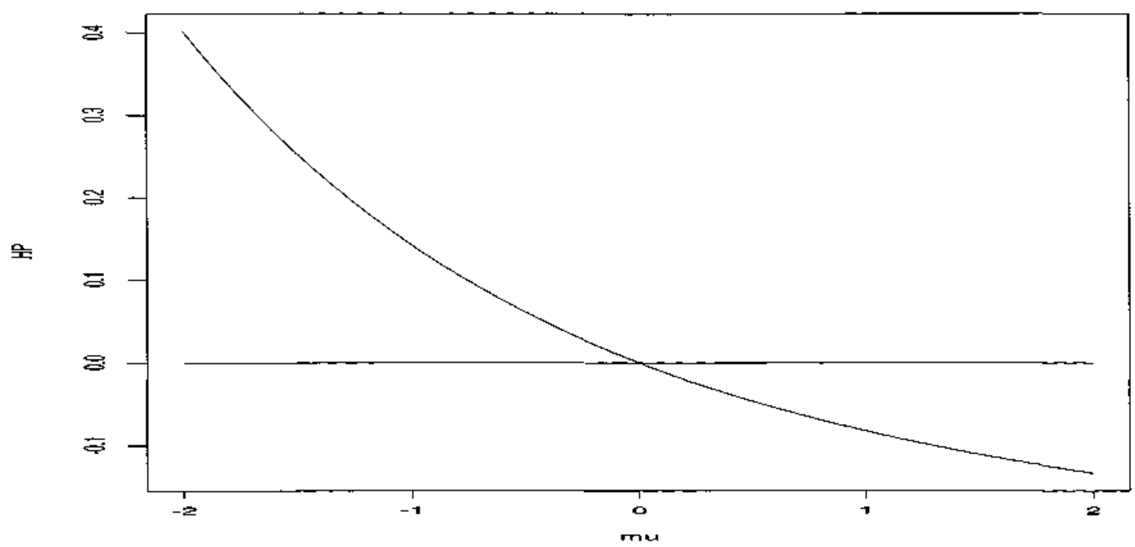


Figure 4.4: HP versus μ for the ideal sample of size 5 ($c = -1.88$, $\sigma = 1$)



as a subroutine to solve the equation (4.8). The results are given, for values of $c = -1.88$, -1, 0, 1,3 and 10, in Tables 4.1.

Table 4.1: The simulation results for the MPS estimate of μ, for different values of n and c when $\mu = 0$, $\sigma = 1$

n	$c = -1.88$			$c = -1$		
	$E(\tilde{\mu})$	$\sigma(\tilde{\mu})$	$\text{Var}(\tilde{\mu})$	$E(\tilde{\mu})$	$\sigma(\tilde{\mu})$	$\text{Var}(\tilde{\mu})$
5	0.051	1.57	2.46	0.037	1.13	1.28
10	-0.051	0.96	0.92	-0.046	0.73	0.53
20	-0.061	0.66	0.44	-0.041	0.50	0.25
50	-0.043	0.40	0.16	-0.031	0.32	0.10
100	-0.036	0.28	0.078	-0.021	0.22	0.048

Table 4.1: Continued

n	$c = 0$			$c = 1$		
	$E(\tilde{\mu})$	$\sigma(\tilde{\mu})$	$\text{Var}(\tilde{\mu})$	$E(\tilde{\mu})$	$\sigma(\tilde{\mu})$	$\text{Var}(\tilde{\mu})$
5	-0.0021	0.80	0.64	-0.0083	0.59	0.35
10	-0.040	0.54	0.29	-0.0023	0.40	0.16
20	-0.034	0.37	0.14	-0.018	0.28	0.078
50	-0.018	0.23	0.053	-0.0098	0.18	0.032
100	-0.016	0.16	0.025	-0.0067	0.13	0.016

Table 4.1: Continued

n	$c = 3$			$c = 10$		
	$E(\tilde{\mu})$	$\sigma(\tilde{\mu})$	$\mathrm{Var}(\tilde{\mu})$	$E(\tilde{\mu})$	$\sigma(\tilde{\mu})$	$\mathrm{Var}(\tilde{\mu})$
5	0.0066	0.45	0.20	-0.0047	0.44	0.18
10	-0.0049	0.32	0.10	-0.0017	0.31	0.097
20	-0.0038	0.22	0.050	-0.0027	0.22	0.049
50	-0.0020	0.14	0.020	-0.0011	0.14	0.020
100	-0.0020	0.10	0.010	-0.00030	0.10	0.010

The results in the extended Table 4.1 show that, for each truncation point, the value of $\mathrm{Var}(\tilde{\mu})$ decreases as n increases, as is to be expected. Moreover, it can be seen that, as the truncation point c increases, so $\mathrm{Var}(\tilde{\mu})$ decreases. But the comparison of $E(\tilde{\mu})$ and $\sigma(\tilde{\mu})$ from Table 4.1 with the corresponding values of $E(\hat{\mu})$ and $\sigma(\hat{\mu})$ from Table 2.3 shows that $E(\tilde{\mu})$ is closer to the exact value of $\mu(= 0)$ than $E(\hat{\mu})$, for example for $n = 5$ and $c = -1.88$ we have $E(\tilde{\mu} = 0.051)$. Also, we can see that for almost every truncation point, $\sigma(\tilde{\mu})$ is less than $\sigma(\hat{\mu})$, for example for $n = 5$ and $c = -1.88$ we have $\sigma(\tilde{\mu}) = 1.57$ whereas for $\sigma(\mu) = 1.589$. Therefore we conclude that the MPS estimator is better than the ML estimator.

4.3.2 Simulation study using the rejection method:

In practical situations we are sometimes dealing with the extreme left tail of the standard normal. For example, if we need $n = 9$ random deviates from the truncated normal, with truncation point $c = -1.88$, since this point represent the 3^{rd} centile, we have to generate about 300 random deviates from the normal distribution, the method of generating random number which we have used so far, is then time consuming. Can we find a more efficient method? In this section, we investigate the rejection method more efficient than generating from the normal distribution. We find $E(\tilde{\mu})$, $\text{Var}(\tilde{\mu})$ and $\sigma(\tilde{\mu})$, for the rejected method and demonstrate the relative speeds of the two methods. Morgan (1984) expressed "If it were possible to choose $h(x)$ to be of a roughly similar shape to $f(x)$ and then to envelop $f(x)$ by $h(x)$, we would obtain the desired scatter of points under $f(x)$ by first obtaining a scatter of points under $h(x)$ but not under $f(x)$." By using Morgan (1984) and Gallagher (1993), we try to find an envelope function $g(x)$ for the density $f(x,\mu)$.

From Chapter 2, equation (2.1) we know that

$$f(x,\mu) = \frac{\phi(\frac{x-\mu}{\sigma})}{\sigma\Phi(\frac{c-\mu}{\sigma})} \qquad ; -\infty \le x \le c \tag{4.14}$$

By setting $\mu = 0$, $\sigma = 1$ and knowing the shape of the density, we guess that the envelope function can be of the form

$$g(x) = de^{ax} \qquad ; -\infty \le x \le c \tag{4.15}$$

where d and a are constants, to be determined.

Let the envelope satisfy the following equations

$$\begin{cases} f(c) = g(c) \\ f'(c) = g'(c). \end{cases} \tag{4.16}$$

Then

$$a = -c \tag{4.17}$$

and

$$d = \frac{e^{\frac{c^2}{2}}}{\sqrt{2\pi}\Phi(c)}. \tag{4.18}$$

Having found a and d, we can write

$$g(x) = \frac{e^{\frac{c^2}{2} - cx}}{\sqrt{2\pi}\Phi(c)}. \tag{4.19}$$

We now prove the following Theorem.

Theorem 4.4 *For all value of $c \leq 0$, every value of x, $-\infty \leq x \leq c \leq 0$, we have*

$$g(x) \geq f(x). \tag{4.20}$$

Proof: From (4.14) and (4.19) with $\mu = 0$, $\sigma = 1$, we have

$$\frac{g(x)}{f(x)} = e^{c^2/2+x^2/2-cx} \quad . \tag{4.21}$$

Now $\frac{c^2}{2} + \frac{x^2}{2} - cx$ can be written as $\frac{1}{2}(x-c)^2$, which is positive for all x and all c. Therefore we have

$$\frac{g(x)}{f(x)} \geq 1, \qquad -\infty \leq x \leq c \tag{4.22}$$

which proves the theorem.

Now we define $h(x) = kg(x)$ such that $h(x)$ is a density function over $-\infty \leq x \leq c \leq 0$. To find the normalising constant k, we require

$$\int_{-\infty}^{c} h(x)dx = 1. \tag{4.23}$$

Hence we find $k = -ce^{\frac{c^2}{2}}\sqrt{2\pi}\Phi(c)$ which, on substitution into the equation $h(x) = kg(x)$, gives

$$h(x) = -ce^{c^2-cx}, -\infty \leq x \leq c \leq 0. \tag{4.24}$$

4.3.3 Simulating data from $h(x)$:

Now we have to simulate data from the density function $h(x)$. To do this we first have to generate a random variable R_1 from the uniform (0,1) distribution. The cumulative distribution function $H(x)$ can be written as

$$\begin{aligned} H(x) &= P(X \leq x) \\ &= \int_{-\infty}^{x} h(t)dt \\ &= e^{c^2-cx}, \end{aligned} \tag{4.25}$$

which satisfies $0 \leq H(x) \leq 1$. Therefore, by inversion of the equation

$$R_1 = e^{c^2 - cx} \tag{4.26}$$

we can find

$$x = c - \frac{1}{c} \ln(R_1). \tag{4.27}$$

To select the random deviate x from $f(x)$, we generate another random variable R_2 from the uniform distribution and then accept x if $R_2 \leq \frac{f(x)}{g(x)}$.

This process continues up to the required sample size. Using the above method we calculate $E(\tilde{\mu})$, $\sigma(\tilde{\mu})$ and $\text{Var}(\tilde{\mu})$ for $R = 10000$ iterations and truncation points $c = -1.88$ and $c = -1$ and sample sizes $n = 5$, 10, 20, 50 and 100 (see Appendix Program 24). If we now combine the above results with those of Table 4.1 for $c = -1.88$ and $c = -1$, we obtain Table 4.2 for $R = 20000$.

Table 4.2: The simulation results for the MPS estimator of μ, for different values of n and c when $\mu = 0$, $\sigma = 1$

n	$c = -1.88$			$c = -1$		
	$E(\tilde{\mu})$	$\sigma(\tilde{\mu})$	$\text{Var}(\tilde{\mu})$	$E(\tilde{\mu})$	$\sigma(\tilde{\mu})$	$\text{Var}(\tilde{\mu})$
5	0.065	1.56	2.44	0.033	1.14	1.30
10	-0.055	0.96	0.92	-0.041	0.74	0.55
20	-0.057	0.65	0.43	-0.045	0.50	0.25
50	-0.043	0.40	0.16	-0.033	0.32	0.10
100	-0.035	0.28	0.081	-0.021	0.22	0.048

4.3.3.1 The comparison of the speeds of the two methods:

Generating $R = 10000$ random deviates from the truncated normal with truncation point $c = -1.88$ and calculating their mean, variance and standard deviation by the first method takes 58 seconds of CPU time. The rejection method is much faster, taking only 6 seconds. The calculations were performed using Fortran Programs 25 and 26 given in the Appendix.

4.4 Relationship between $E(\tilde{\mu})$, sample size and truncation point:

In this section we investigate the relationship of $E(\tilde{\mu})$ with the sample size n and truncation point c.

Our approach is to fit a regression model, of $E(\tilde{\mu})$ on n and c.

We have to find $\widetilde{E(\tilde{\mu})}$ for different values of c and n. The values have been obtained and are shown in Table 4.3. Also in Figure 4.5 we plot the $\widetilde{E(\tilde{\mu})}$ against c for different values of n.

Table 4.3: The simulated expected value of $\tilde{\mu}$, found from $R = 100000$ simulation runs, for different values of n and c when $\mu = 0$, $\sigma = 1$

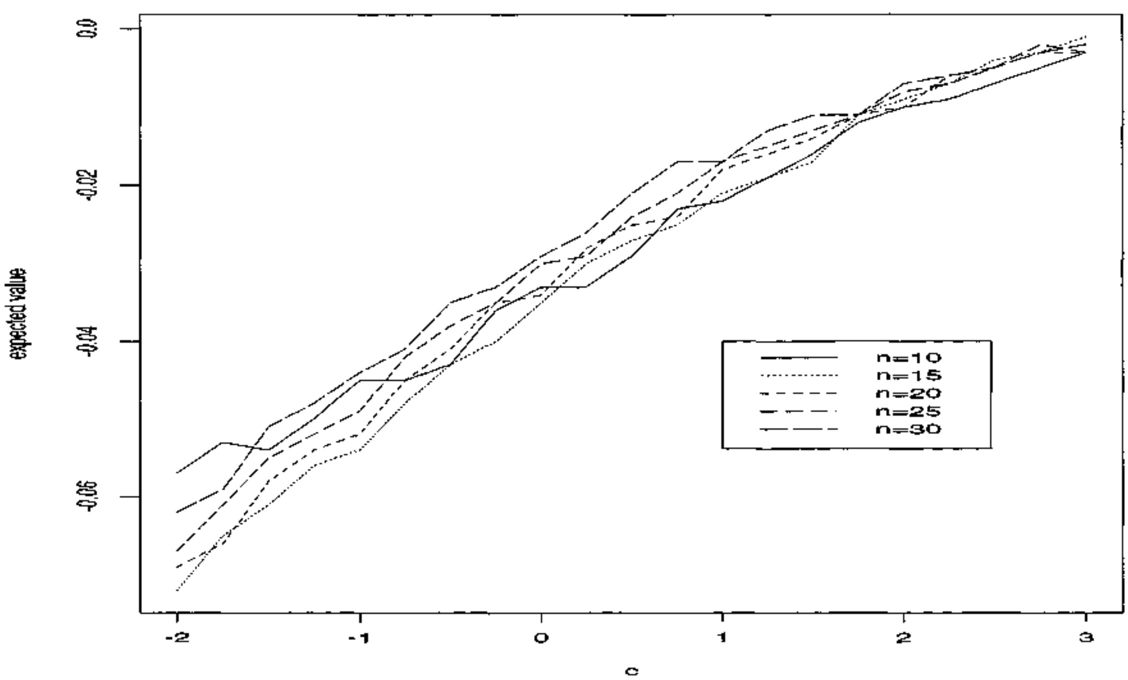
c	$\widetilde{E(\tilde{\mu})}$				
	$n = 10$	$n = 15$	$n = 20$	$n = 25$	$n = 30$
-2	-0.057	-0.072	-0.069	-0.067	-0.062
-1.75	-0.053	-0.065	-0.066	-0.061	-0.059
-1.5	-0.054	-0.061	-0.058	-0.055	-0.051
-1.25	-0.050	-0.056	-0.054	-0.052	-0.048
-1	-0.045	-0.054	-0.052	-0.049	-0.044
-0.75	-0.045	-0.048	-0.045	-0.042	-0.041
-0.5	-0.043	-0.043	-0.041	-0.038	-0.035
-0.25	-0.036	-0.040	-0.035	-0.035	-0.033
0	-0.033	-0.035	-0.034	-0.030	-0.029
0.25	-0.033	-0.030	-0.028	-0.029	-0.026
0.5	-0.029	-0.027	-0.025	-0.024	-0.021
0.75	-0.023	-0.025	-0.024	-0.021	-0.017
1	-0.022	-0.021	-0.018	-0.017	-0.017
1.25	-0.019	-0.019	-0.016	-0.015	-0.013
1.5	-0.016	-0.017	-0.014	-0.013	-0.011
1.75	-0.012	-0.011	-0.011	-0.011	-0.011
2	-0.010	-0.009	-0.010	-0.008	-0.007
2.25	-0.009	-0.007	-0.006	-0.007	-0.006
2.5	-0.007	-0.004	-0.005	-0.005	-0.005
2.75	-0.005	-0.003	-0.003	-0.002	-0.003
3	-0.003	-0.001	-0.003	-0.003	-0.002

The following plot shows that, for sample sizes $n = 10, 15, 20, 25$ and 30, as the truncation point c increases then $\widetilde{E(\tilde{\mu})} \to \mu(= 0)$. In Table 4.3, since the differences of $\widetilde{E(\tilde{\mu})}$ from zero are significant and $\tilde{\mu}$ is consistent then an appropriate model for $E(\tilde{\mu})$ can be written as

$$E(\tilde{\mu}) = \mu + \frac{1}{n} g_1(c) + \frac{1}{n^2} g_2(c) + O(n^{-3}). \tag{4.28}$$

Our approach is to choose two values of n and obtain two different equations. Then, by the simultaneous solution of the two sets of equations, we find the corresponding $\widetilde{g_1(c)}$ and $\widetilde{g_2(c)}$ (see Appendix Program 27). By choosing two values of n it is possible to solve (4.28) with the *simulated* mean of $\tilde{\mu}$ for $E(\tilde{\mu})$. The values of $n = 10$ and $n = 20$ were chosen since as not

Figure 4.5: The plot of $\widetilde{E(\tilde{\mu})}$ against c for different n



to give so small a value that cause $E(\tilde{\mu})$ become large nor too large a value that it become so small. Therefore $\widetilde{g_1(c)}$ and $\widetilde{g_2(c)}$ which calculated for $n = 10$ and $n = 20$, are shown in Table 4.4.

Table 4.4: The estimated $\widetilde{g_1(c)}$ and $\widetilde{g_2(c)}$ for different values of c which $n = 10$ and $n = 20$

c	$n = 10$, $n = 20$		c	$n = 10$, $n = 20$		c	$n = 10$, $n = 20$	
	$\widetilde{g_1(c)}$	$\widetilde{g_2(c)}$		$\widetilde{g_1(c)}$	$\widetilde{g_2(c)}$		$\widetilde{g_1(c)}$	$\widetilde{g_2(c)}$
-2.00	-2.19	16.20	-0.25	-1.04	6.80	1.50	-0.40	2.40
-1.75	-2.11	15.80	0.00	-1.03	7.0	1.75	-0.32	2.00
-1.50	-1.78	12.40	0.25	-0.79	4.60	2.00	-0.30	2.00
-1.25	-1.66	11.60	0.50	-0.71	4.20	2.25	-0.15	0.60
-1.00	-1.63	11.80	0.75	-0.73	5.00	2.50	-0.13	0.60
-0.75	-1.35	9.00	1.00	-0.50	2.80	2.75	-0.07	0.20
-0.50	-1.21	7.80	1.25	-0.45	2.60	3.00	-0.09	0.60

The plots of $\widetilde{g_1(c)}$ and $\widetilde{g_2(c)}$ against c are shown in Figures 4.6 and Figure 4.7.

By looking at Figure 4.6 we can see that all the values of $\widetilde{g_1(c)}$ are negative and $\widetilde{g_1(c)} \to 0$ as c increases. But in Figure 4.7 all the values of $\widetilde{g_2(c)}$ are positive and $\widetilde{g_2(c)} \to 0$ as c increases. Therefore we choose models $g_1(c) = -\alpha_1 e^{-\beta_1 c}$ and $g_2(c) = \alpha_2 e^{-\beta_2 c}$. Note that these functions have the multiplicative error terms. Now, we are interested to find the functions $g_1(c)$ and $g_2(c)$ in terms of c.

Applying the logarithm transformation and linear regression model by the help of Minitab software the following models are obtained:

$$g_1(c) = -0.832e^{-0.658c}$$

where the corresponding $t-$ratios of the coefficients -0.832 and -0.658 are -2.90 and -16.47. Similarly,

$$g_2(c) = 5.3122e^{-0.755c}$$

Figure 4.6: The plot of $\widetilde{g_1(c)}$ against c

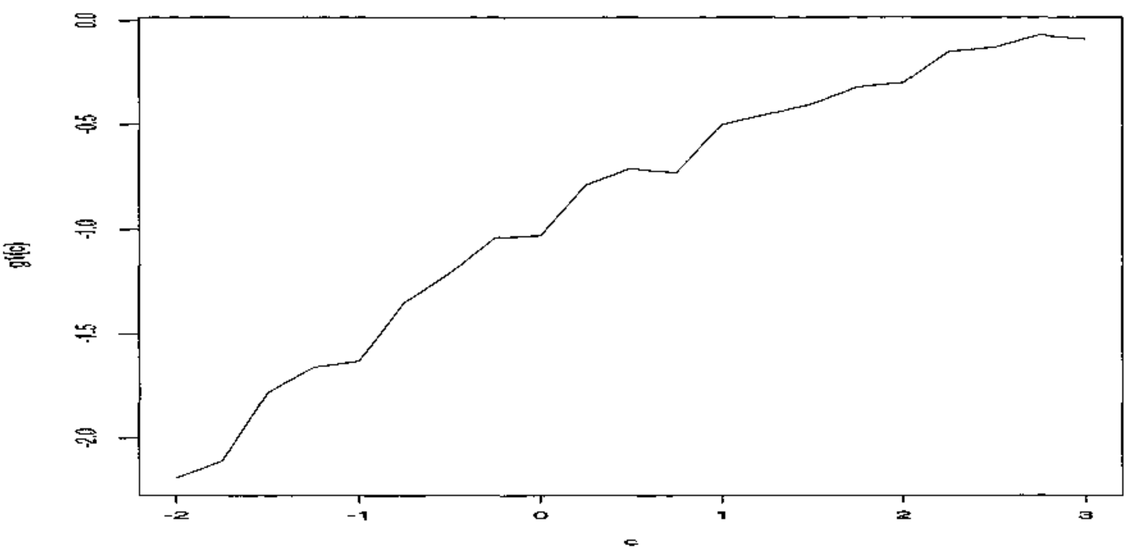
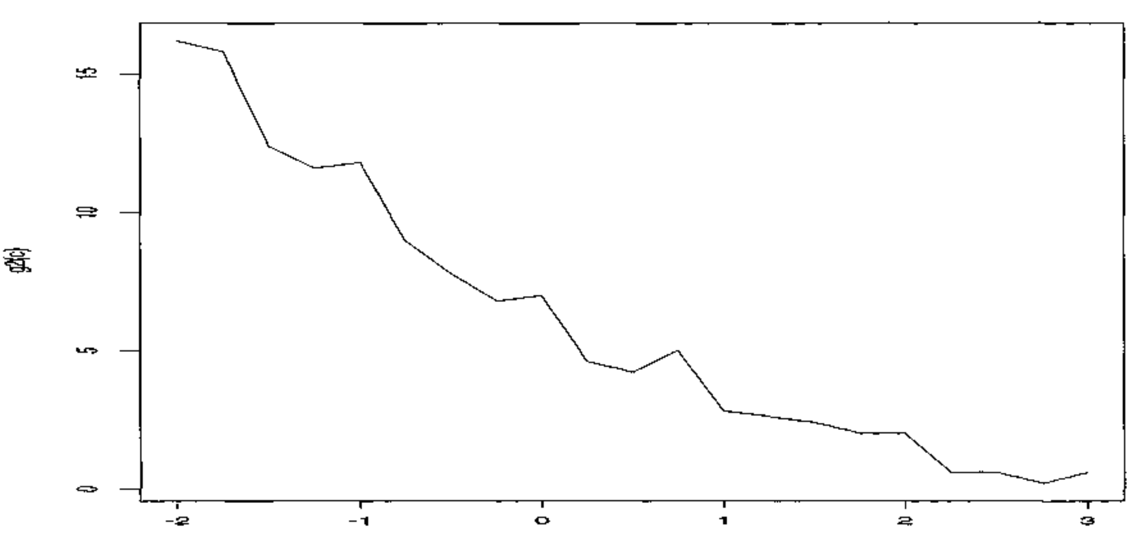


Figure 4.7: The plot of $\widetilde{g_2(c)}$ against c



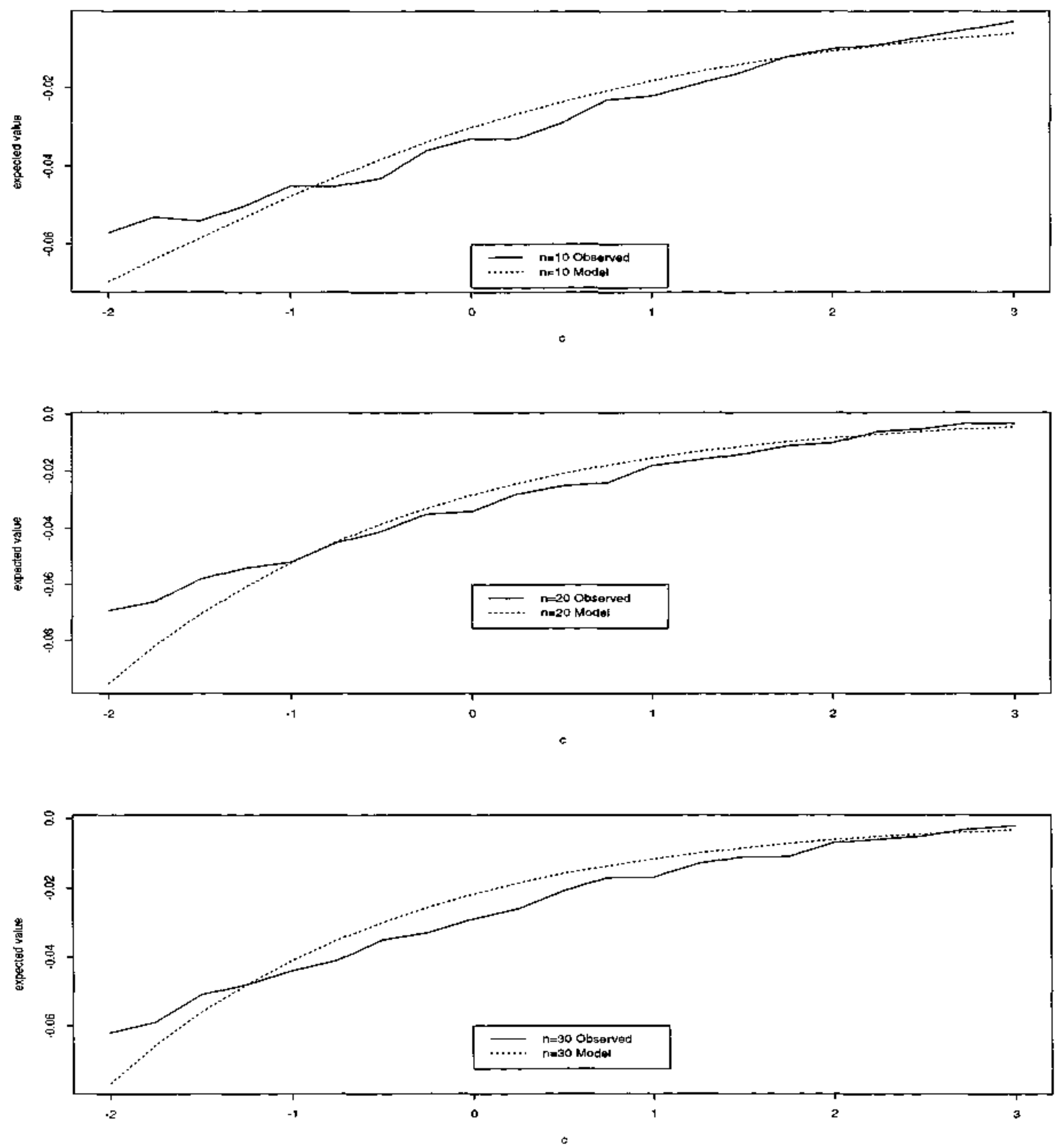
where the corresponding $t-$ratios of the coefficients 5.3122 and -0.755 are 17.44 and -12.60.

Substituting the estimated $g_1(c)$ and $g_2(c)$ functions into equation (4.28) we obtain

$$E(\tilde{\mu}) = \mu + \frac{-0.832e^{-0.658c}}{n} + \frac{5.3122e^{-0.755c}}{n^2} + O(n^{-3}).$$

To compare $E(\tilde{\mu})$ calculated from the model together with the $\widetilde{E(\tilde{\mu})}$ calculated from the simulation, we plot them in Figure 4.8.

Figure 4.8: The plot of $\widetilde{E(\hat{\mu})}$ and the model $E(\tilde{\mu})$ against c



4.5 Relationship between Var($\tilde{\mu}$) and sample size and truncation point:

In this section we investigate the relationship of Var($\tilde{\mu}$) with sample size n and truncation point c.

To find the appropriate regression model, we have to find $\widetilde{\mathrm{Var}(\tilde{\mu})}$ for different values of c and n. The values have been obtained and are shown in Table 4.5 and Figure 4.9.

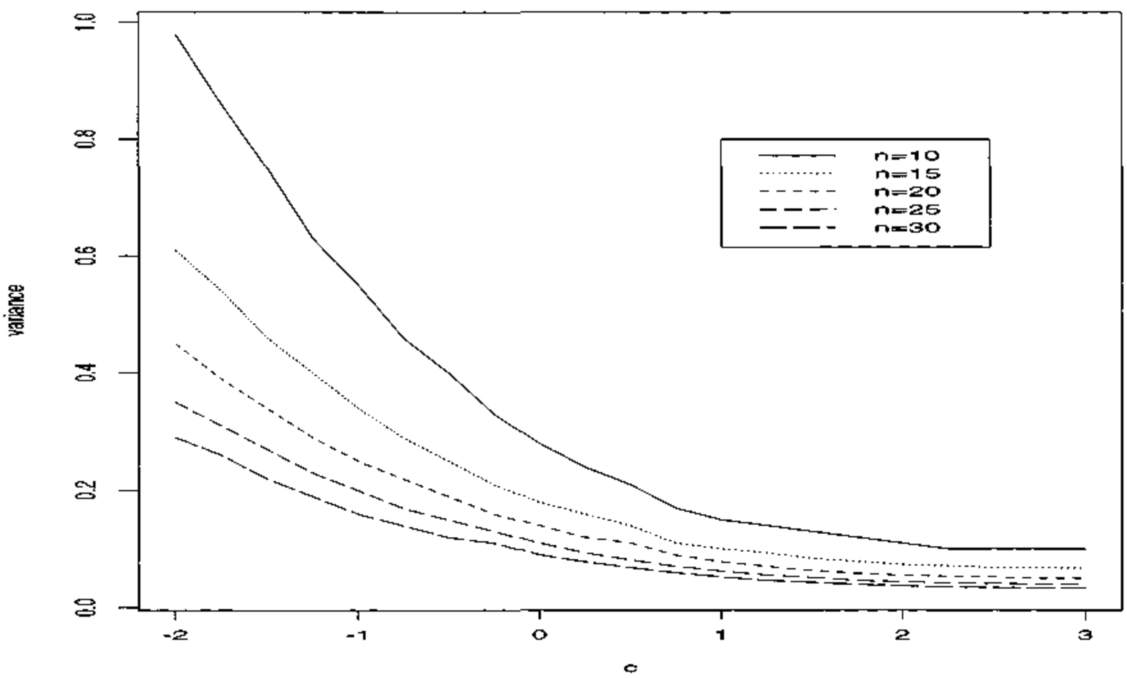
Table 4.5: The variance of $\tilde{\mu}$, by simulation for different values of n and c when $\mu = 0$, $\sigma = 1$ and simulation run $R = 100000$

c	$\widetilde{\mathrm{Var}(\tilde{\mu})}$				
	$n = 10$	$n = 15$	$n = 20$	$n = 25$	$n = 30$
-2	0.985	0.615	0.450	0.351	0.293
-1.75	0.862	0.535	0.391	0.308	0.258
-1.50	0.746	0.463	0.338	0.269	0.224
-1.25	0.633	0.403	0.294	0.233	0.191
-1.00	0.546	0.343	0.253	0.199	0.165
-0.75	0.462	0.296	0.217	0.172	0.143
-0.50	0.397	0.252	0.187	0.148	0.123
-0.25	0.334	0.215	0.161	0.126	0.105
0.00	0.284	0.185	0.136	0.109	0.090
0.25	0.239	0.159	0.117	0.093	0.078
0.50	0.207	0.137	0.101	0.081	0.068
0.75	0.179	0.117	0.089	0.070	0.059
1.00	0.157	0.104	0.078	0.062	0.052
1.25	0.141	0.093	0.070	0.055	0.046
1.50	0.127	0.084	0.063	0.051	0.042
1.75	0.117	0.078	0.056	0.046	0.039
2.00	0.110	0.073	0.055	0.044	0.037
2.25	0.105	0.070	0.053	0.042	0.035
2.50	0.103	0.068	0.051	0.041	0.034
2.75	0.102	0.067	0.050	0.040	0.033
3.00	0.101	0.067	0.050	0.040	0.033

Since $\tilde{\mu}$ is consistent then Var($\tilde{\mu}$) can be written as

$$\mathrm{Var}(\tilde{\mu}) = \frac{1}{n} g_1(c) + \frac{1}{n^2} g_2(c) + \frac{1}{n^3} g_3(c) + O(n^{-4}). \tag{4.29}$$

Figure 4.9: The plot of $\widetilde{\text{Var}(\tilde{\mu})}$ against c for different n



Since we know that $\hat{\mu}$ and $\tilde{\mu}$ are asymptotically equivalent Cheng & Amin (1982), it follows that $\text{Var}(\tilde{\mu}) = \text{Var}(\hat{\mu})$ to $O(n^{-1})$. Hence $g_1(c) = \frac{\sigma^2}{1+\psi'(c')/\sigma}$ and equation (4.29) can be written as

$$\text{Var}(\tilde{\mu}) = \frac{1}{n}\left[\frac{\sigma^2}{1+\psi'(c')/\sigma}\right] + \frac{1}{n^2}g_2(c) + \frac{1}{n^3}g_3(c) + O(n^{-4}). \tag{4.30}$$

As in section 4.4 by choosing two values of n it is possible to solve equation (4.30) with the simulated variance $\widetilde{\text{Var}(\tilde{\mu})}$ the $\text{Var}(\tilde{\mu})$. Similar to section 4.4 the values of $n = 10$ and $n = 20$ were chosen.

The calculated $\widetilde{g_2(c)}$ and $\widetilde{g_3(c)}$ are shown in Table 4.6 (see Appendix Program 28).

Table 4.6: The estimated $\widetilde{g_2(c)}$ and $\widetilde{g_3(c)}$ for different values of c when $n = 10$ and $n = 20$

c	$n = 10$, $n = 20$		c	$n = 10$, $n = 20$		c	$n = 10$, $n = 20$	
	$\widetilde{g_2(c)}$	$\widetilde{g_3(c)}$		$\widetilde{g_2(c)}$	$\widetilde{g_3(c)}$		$\widetilde{g_2(c)}$	$\widetilde{g_3(c)}$
-2.00	-0.92	119.06	-0.25	-0.29	17.41	1.50	-0.46	2.71
-1.75	-3.32	128.88	0.00	-1.55	24.42	1.75	-0.64	4.50
-1.50	-4.40	121.73	0.25	-1.22	14.98	2.00	-0.51	3.14
-1.25	-1.89	71.57	0.50	-1.09	12.58	2.25	-0.46	2.64
-1.00	-2.41	68.74	0.75	-0.37	3.15	2.50	-0.20	0.55
-0.75	-2.57	55.53	1.00	-0.90	8.18	2.75	-0.30	1.42
-0.50	-1.30	38.31	1.25	-0.72	6.47	3.00	0.06	-0.70

The plots of $\widetilde{g_2(c)}$ and $\widetilde{g_3(c)}$ against c are shown in Figures 4.10 and Figure 4.11.

By looking at Figure 4.10, we can see that all the values of $\widetilde{g_2(c)}$ except for $c = 3$ are negative and as c increases $\widetilde{g_2(c)} \to 0$. But in Figure 4.11 all the values of $\widetilde{g_3(c)}$ except $c = 3$ are positive and as c increases $\widetilde{g_3(c)} \to 0$.

To obtain smoother curves of $\widetilde{g_2(c)}$ and $\widetilde{g_3(c)}$ against c we use three-point moving averages, which whose plots are shown in Figures 4.12 and 4.13. The former is not very smooth but smooth enough to suggest a formula of the form $g_2(c) = -\alpha_3 e^{-\beta_3 c}$. Similarly we take $g_3(c) = \alpha_4 e^{-\beta_4 c}$.

Now, we are interested in finding the functions $g_2(c)$ and $g_3(c)$ in terms of c.

Applying the logarithm transformation and linear regression model by the help of Minitab software the following models are obtained:

$$g_2(c) = -1.18e^{-0.530c},$$

Figure 4.10: The plot of $\widetilde{g_2(c)}$ against c

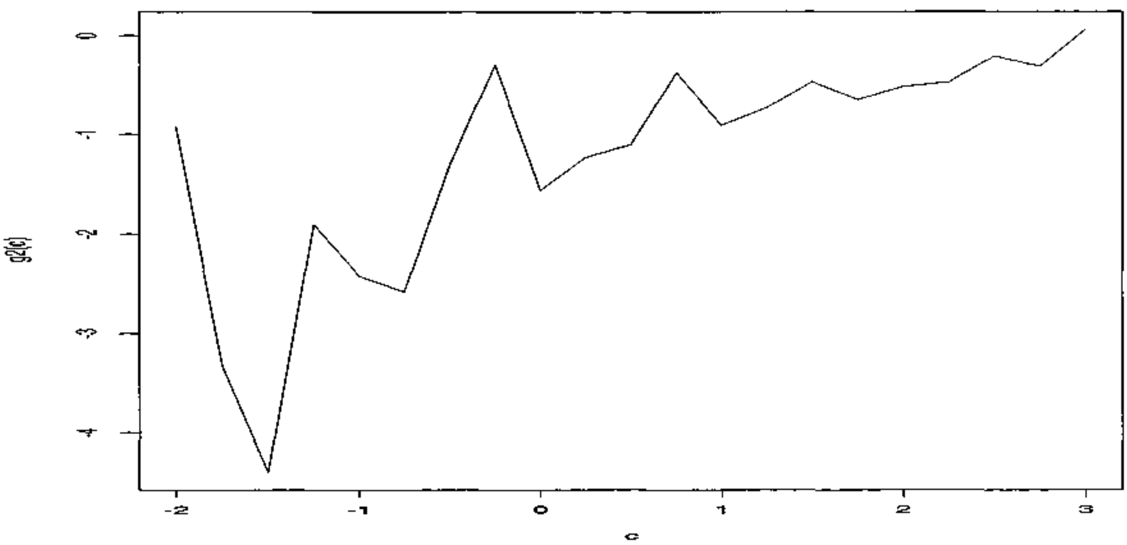


Figure 4.11: The plot of $\widetilde{g_3(c)}$ against c

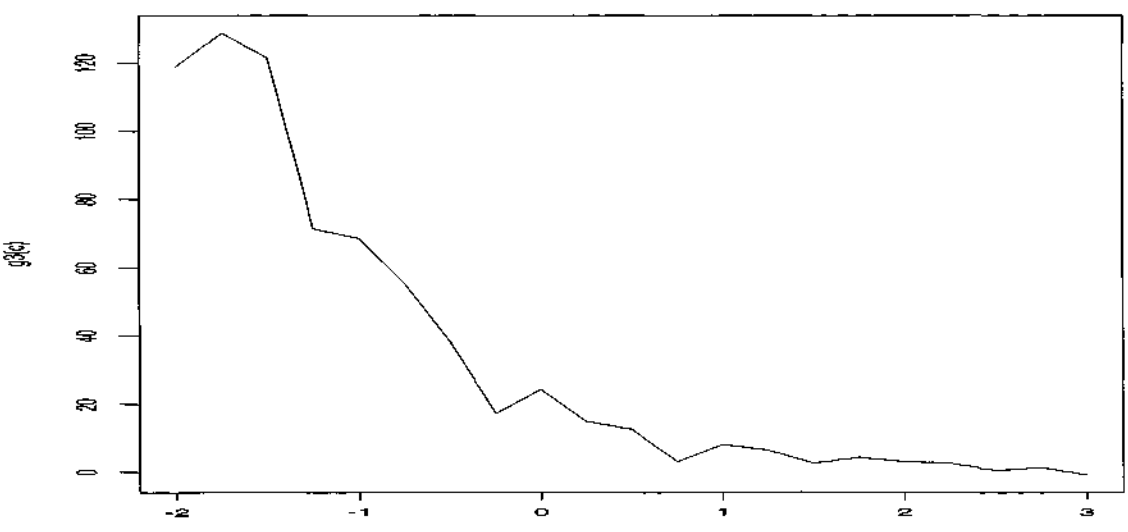


Figure 4.12: The three-point smooth plot of $\widetilde{g_2(c)}$ against c

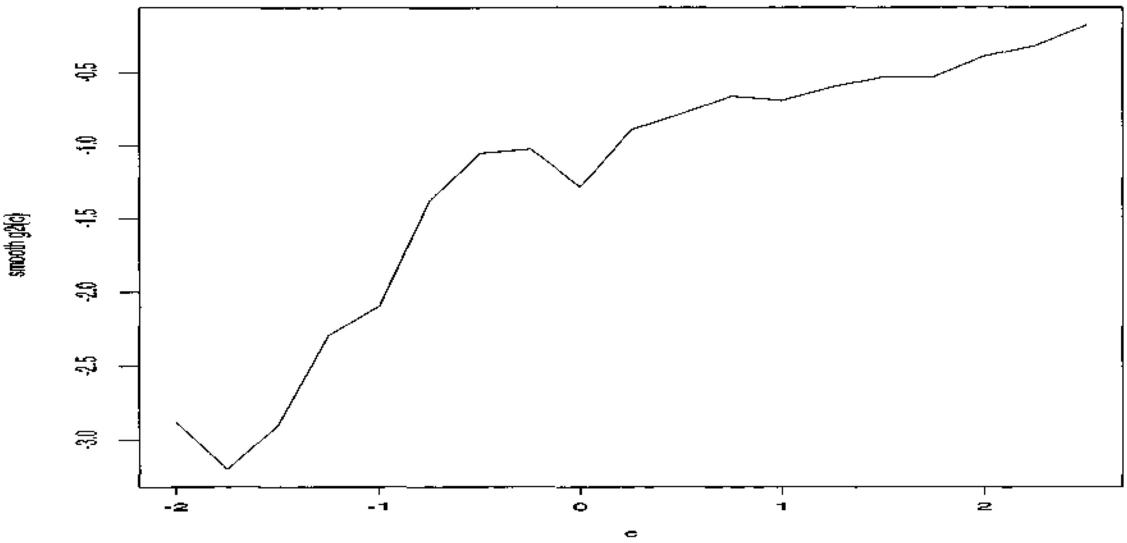
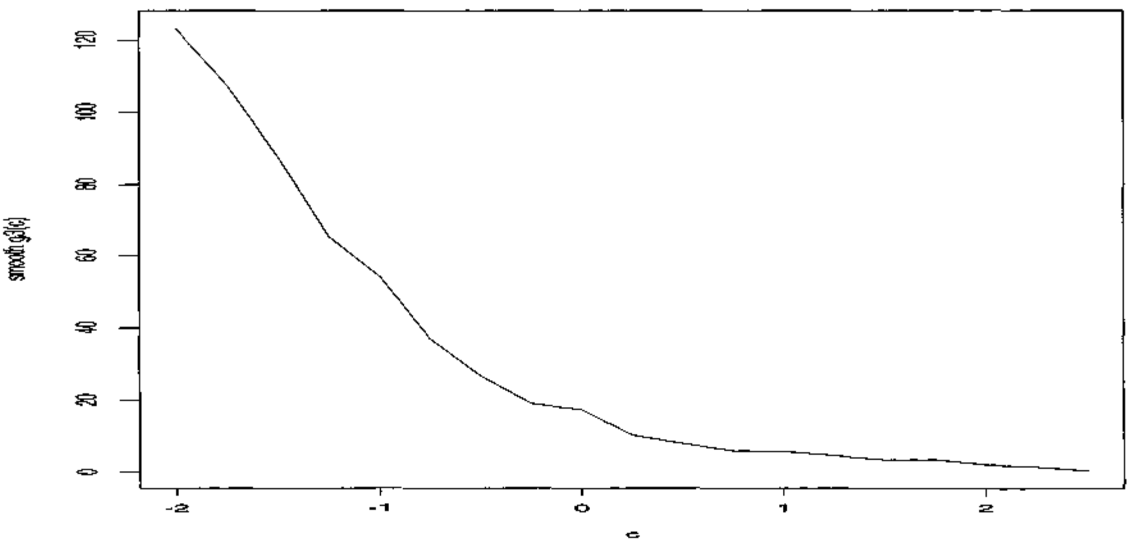


Figure 4.13: The three-point smooth plot of $\widetilde{g_3(c)}$ against c



where the corresponding $t-$ratios of the parameter estimators -1.18 and -0.530 are 1.99 and -6.42,

and

$$g_3(c) = 19.5e^{-1.07c}$$

where the corresponding $t-$ratios of coefficients 19.5 and -1.07 are 31.17 and -16.66.

From these equations, we can estimate $g_2(c)$ and $g_3(c)$ for any values of c.

Substituting the estimated $g_2(c)$ and $g_3(c)$ in equation (4.30) we obtain

$$\text{Var}(\tilde{\mu}) = \frac{1}{n}\left[\frac{\sigma^2}{1+\psi'(c')/\sigma}\right] + \frac{1}{n^2}\left[-0.96e^{-0.527c}\right] + \frac{1}{n^3}\left[12.9e^{-1.06c}\right] + O(n^{-4}).$$

We can easily evaluate $\text{Var}(\tilde{\mu})$ for different values of n and c. To compare $\text{Var}(\tilde{\mu})$ calculated from the model together with the $\widetilde{\text{Var}(\tilde{\mu})}$ calculated from the simulation, we plot them in Figure 4.14.

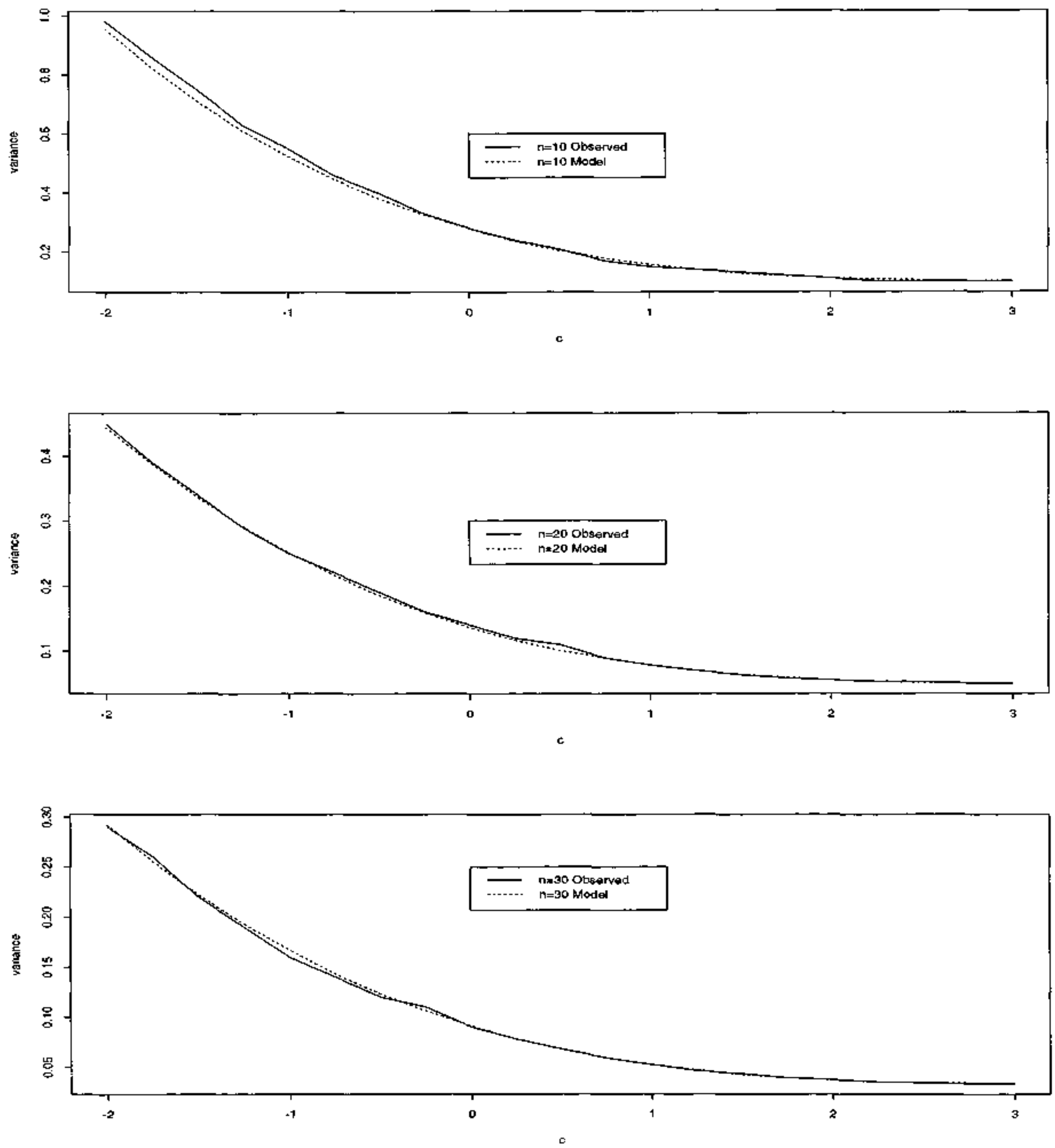
Figures 4.8 and 4.14 shows that these models are reliable and we can easily find $E(\tilde{\mu})$ and $\text{Var}(\tilde{\mu})$ for various values of n and c.

4.6 Distribution of $\tilde{\mu}$ when the variance is known:

We investigate the distribution of $\tilde{\mu}$ when the truncation point is $c = -1.88$. This truncation point is chosen because it is seen that in many investigations involving children's growth the third percentile of the distribution is important. Also this point is the worst case of truncation points. Using the Program 7, for various values of $n = 5, 10, 20, 50$ and 100, we have constructed $R = 10000$ observations of $\tilde{\mu}$. By use of the S-PLUS software, we have plotted the histogram, density plot, qqnorm and qqline of $\tilde{\mu}$. If the sample comes from a normal distribution, with sample size R, the sample estimate of the coefficient of skewness $g_1(\tilde{\mu})$ is given by

$$g_1(\tilde{\mu}) = \frac{m_3(\tilde{\mu})}{m_2(\tilde{\mu})\sqrt{m_2(\tilde{\mu})}}$$

Figure 4.14: The plot of $\widetilde{\mathrm{Var}(\tilde{\mu})}$ and the model $\mathrm{Var}(\tilde{\mu})$ against c



is asymptotically

$$g_1(\tilde{\mu}) \sim N(0, \frac{6}{R}).$$

The assumption of normality of $g_1(\tilde{\mu})$ is accurate for $R \geq 150$. (Snedecor & Cochran (1967), p. 68) Also we know that, in very large samples from the normal distribution, the measure of kurtosis $g_2(\tilde{\mu})$ defined as

$$g_2(\tilde{\mu}) = \frac{m_4(\tilde{\mu})}{m_2^2(\tilde{\mu})}$$

has asymptotically a normal distribution such that

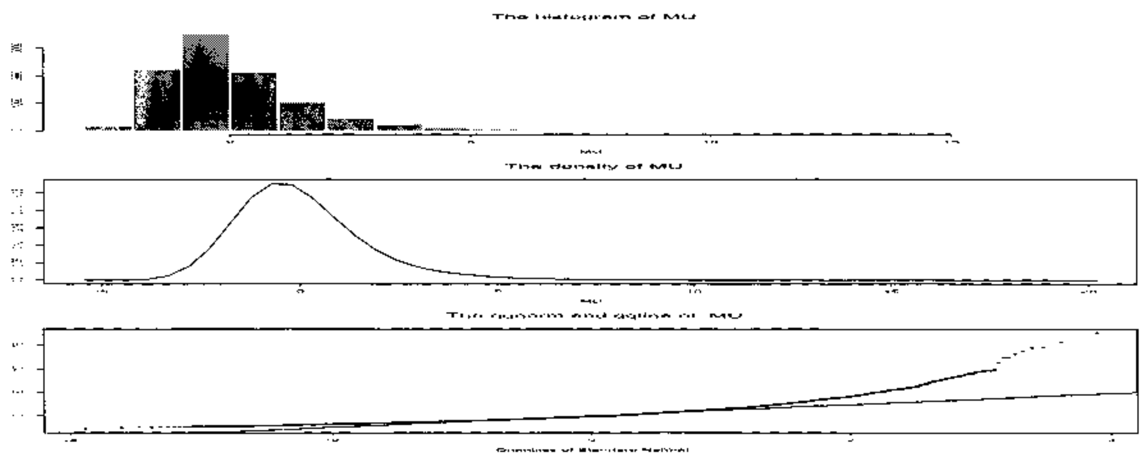
$$g_2(\tilde{\mu}) \sim N(0, \frac{24}{R}),$$

where $m_2(\tilde{\mu})$, $m_3(\tilde{\mu})$ and $m_4(\tilde{\mu})$ are the second, third and fourth moments of $\tilde{\mu}$.

4.6.1 Description of data when $n = 5$:

The histogram, density, qqnorm and qqline of $\tilde{\mu}$ are shown in Figure 4.15.

Figure 4.15: The distribution of $\hat{\mu}$, when variance is known and n=5



From this we can see that the distribution is not close to the normal.

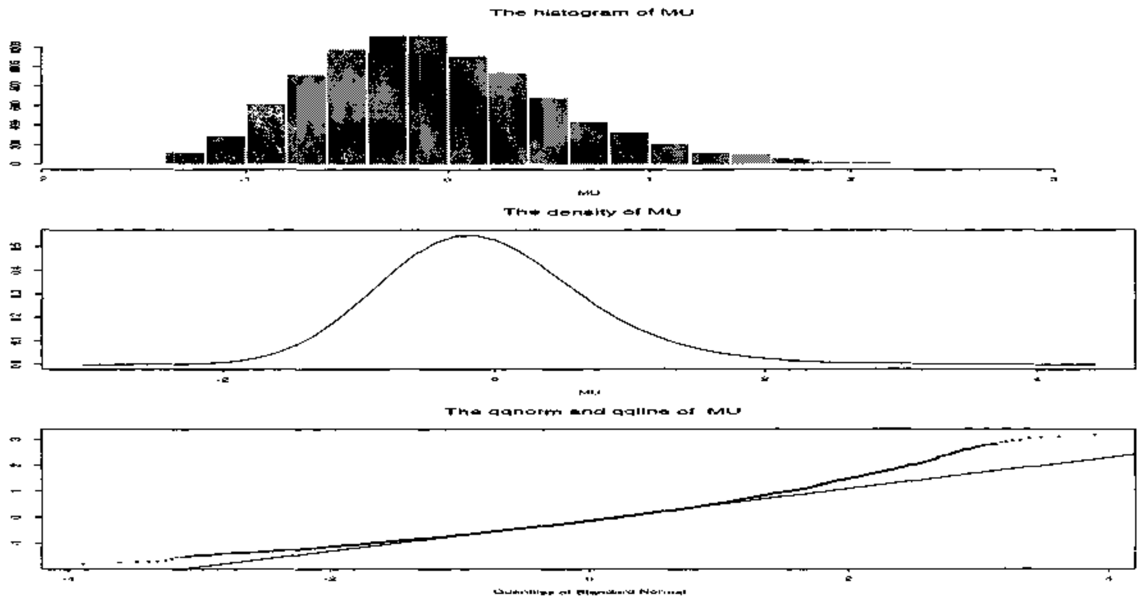
In this set of data, the calculated value $g_1(\tilde{\mu}) = 2.16258$ is over 88 times its standard deviation $(s.d = 0.02449)$ from zero, and the positive skewness is confirmed. The second, third and fourth moments of $\tilde{\mu}$ are $m_2(\tilde{\mu}) = 2.4059$, $m_3(\tilde{\mu}) = 8.0704$, $m_4(\tilde{\mu}) = 75.6026$ and the measure for kurtosis $g_2(\tilde{\mu}) = 10.0614$.

Since $g_2(\tilde{\mu}) = 10.0614$ is over 205 times its standard deviation $(s.d = 0.04898)$, the large kurtosis of this distribution is confirmed.

4.6.2 Description of data when $n = 20$:

As we increase the number of observations to $n = 20$, we can draw the various plots of $\tilde{\mu}$ in Figure 4.16

Figure 4.16: The distribution of $\tilde{\mu}$, when variance is known and n=20

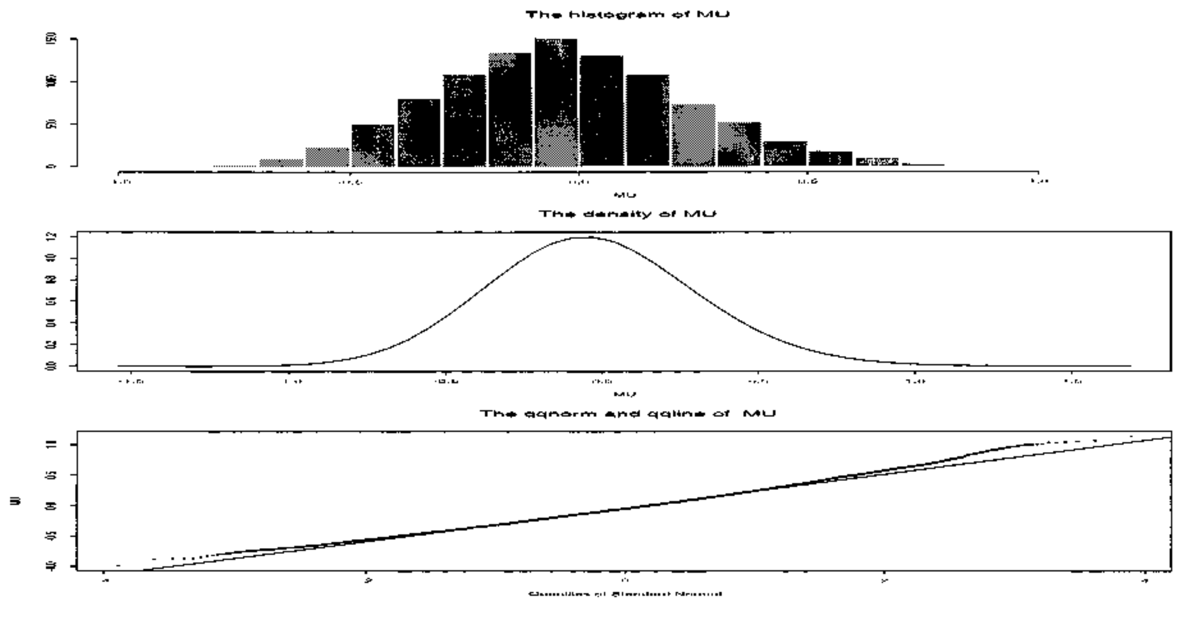


In this case $m_2(\tilde{\mu}) = 0.419485$, $m_3(\tilde{\mu}) = 1.20869$, $m_4(\tilde{\mu}) = 0.7281$, the measure of skewness $g_1(\tilde{\mu}) = 0.7681$ $(s.d = 0.02449)$ and the measure for kurtosis $g_2(\tilde{\mu}) = 1.1375$ $(s.d = 0.04898)$. Therefore the skewness and large kurtosis are again confirmed.

4.6.3 Description of data when $n = 100$:

Finally, for $n = 100$ observations, the various plots of $\tilde{\mu}$ are shown in Figure 4.17.

Figure 4.17: The distribution of $\tilde{\mu}$, when variance is known and n=100



In this case $m_2(\tilde{\mu}) = 0.07927$, $m_3(\tilde{\mu}) = 0.006680$, $m_4(\tilde{\mu}) = 0.2993$, the measure of skewness $g_1(\tilde{\mu}) = 1.4347$ $(s.d = 0.02449)$ and the measure for kurtosis $g_2(\tilde{\mu}) = 0.20864$ $(s.d = 0.04898)$; therefore the skewness and kurtosis are confirmed.

These Figures show that even in case $n = 100$, although the shapes are very close to the normal distribution, but its measure of skewness and kurtosis are still significantly high.

4.7 The comparison of the different estimators of μ:

In this section we compare the estimators of μ based on the ML and MPS methods with the exact value of μ. Since MSE is a good criteria for the comparison of two estimator we used them in Tables 4.7-4.10. By using the results of the simulation study of the maximum likelihood estimator, the theoretical results of Chapter 2 and the results of the simulation study of the maximum spacing methods, described in this chapter, we have produced Tables 4.7-4.10.

Table 4.7: The comparison of the ML and MPS estimators of μ for $c = -1.88$ and different values of n

n	$\hat{\mu}$ Sim.mean	$\hat{\mu}$ Sim.var	$\hat{\mu}$ MSE	$\hat{\mu}$ Theory mean	$\hat{\mu}$ Theory var.	$\tilde{\mu}$ Sim.mean	$\tilde{\mu}$ Sim.var	$\tilde{\mu}$ MSE
5	0.53	2.52	2.81	0.53	2.59	0.053	2.46	2.47
10	0.23	1.13	1.18	0.24	1.06	-0.051	0.96	0.96
20	0.10	0.47	0.48	0.11	0.47	-0.061	0.44	0.44
50	0.040	0.17	0.18	0.044	0.17	-0.043	0.16	0.16
100	0.025	0.084	0.085	0.020	0.084	-0.036	0.078	0.081

Table 4.8: The comparison of the ML and MPS estimators of μ for $c = -1$ and different values of n

n	$\hat{\mu}$ Sim.mean	$\hat{\mu}$ Sim.var	$\hat{\mu}$ MSE	$\hat{\mu}$ Theory mean	$\hat{\mu}$ Theory var.	$\tilde{\mu}$ Sim.mean	$\tilde{\mu}$ Sim.var	$\tilde{\mu}$ MSE
5	0.38	1.83	1.98	0.35	1.47	0.014	1.28	1.29
10	0.16	0.65	0.68	0.16	0.63	-0.046	0.53	0.53
20	0.084	0.29	0.29	0.077	0.26	-0.041	0.25	0.26
50	0.024	0.11	0.11	0.030	0.10	-0.031	0.10	0.10
100	0.015	0.052	0.052	0.015	0.050	-0.021	0.048	0.048

Table 4.9: The comparison of the ML and MPS estimators of μ for $c = 0$ and different values of n

n	$\hat{\mu}$ Sim.mean	$\hat{\mu}$ Sim.var	$\hat{\mu}$ MSE	$\hat{\mu}$ Theory mean	$\hat{\mu}$ Theory var.	$\tilde{\mu}$ Sim.mean	$\tilde{\mu}$ Sim.var	$\tilde{\mu}$ MSE
5	0.20	0.83	0.87	0.19	0.73	-0.0021	0.64	0.64
10	0.089	0.33	0.34	0.089	0.32	-0.040	0.29	0.29
20	0.043	0.15	0.15	0.043	0.15	-0.034	0.14	0.14
50	0.016	0.058	0.058	0.017	0.056	-0.0023	0.053	0.054
100	0.0087	0.028	0.028	0.0083	0.028	-0.016	0.025	0.027

Table 4.10: The comparison of the ML and MPS estimators of μ for $c = 1$ and different values of n

n	$\hat{\mu}$ Sim.mean	$\hat{\mu}$ Sim.var	$\hat{\mu}$ MSE	$\hat{\mu}$ Theory mean	$\hat{\mu}$ Theory var.	$\tilde{\mu}$ Sim.mean	$\tilde{\mu}$ Sim.var	$\tilde{\mu}$ MSE
5	0.076	0.39	0.39	0.084	0.37	-0.083	0.35	0.35
10	0.036	0.17	0.17	0.040	0.17	-0.0023	0.16	0.16
20	0.023	0.084	0.084	0.019	0.083	-0.021	0.078	0.079
50	0.0054	0.033	0.033	0.0075	0.032	-0.0098	0.032	0.032
100	0.0046	0.016	0.016	0.0037	0.016	-0.0067	0.016	0.016

From Tables 4.7 – 4.10 we can see that, as n increases the variances of $\hat{\mu}$ and $\tilde{\mu}$, become identical, and also they are equivalent to the theoretical variance of $\hat{\mu}$. Moreover, we can see that the MSE of $\tilde{\mu}$ is less than that of $\hat{\mu}$ for all sample sizes. Therefore the MPS estimator is more efficient than the ML estimator.

4.8 Estimation of σ when μ is known:

We now estimate σ by the MPS estimator $\tilde{\sigma}$ when μ is known and then compare $\tilde{\sigma}$ with $\hat{\sigma}$.

Taking the derivative of H in equation (4.7) with respect to σ^2 we obtain

$$\frac{\partial H}{\partial \sigma^2} = -\frac{1}{2\sigma^3(n+1)}\{\sum_{i=1}^{n+1}[\frac{(y_i-\mu)\phi(\frac{y_i-\mu}{\sigma}) - (y_{i-1}-\mu)\phi(\frac{y_{i-1}-\mu}{\sigma})}{\Phi(\frac{y_i-\mu}{\sigma}) - \Phi(\frac{y_{i-1}-\mu}{\sigma})}]\} + \frac{(c-\mu)\phi(\frac{c-\mu}{\sigma})}{2\sigma^3\Phi(\frac{c-\mu}{\sigma})}. \quad (4.31)$$

The algebraic solution of $\frac{\partial H}{\partial \sigma^2} = 0$ gives $\tilde{\sigma}$, but it is impossible to solve this equation analytically. Therefore to find the MPS estimator, $\tilde{\sigma}$, we use the NAG routine C05AGF, in Program 29, to get the solution iteratively.

Theorem 4.5 *The MPS estimator ($\tilde{\sigma}$) is asymptotically sufficient, consistent and efficient estimator of σ.*

Proof: By the definition of integration, $\ln(D_i)$ can be written as

$$\begin{aligned} \ln(D_i) &= \ln[\int_{y_{i-1}}^{y_i} f(y,\sigma)dy] \qquad ;i=1,2,\ldots,n+1 \\ &= \ln[f(y_i,\sigma)(y_i - y_{i-1})] + R(y_i, y_{i-1}, \sigma) \\ &= \ln f(y_i,\sigma) + \ln(y_i - y_{i-1}) + R(y_i, y_{i-1}, \sigma). \qquad (4.32) \end{aligned}$$

where

$$R(y_i, y_{i-1}, \sigma) = \ln\left\{\frac{\sigma[\Phi(\frac{y_i-\mu}{\sigma}) - \Phi(\frac{y_{i-1}-\mu}{\sigma})]}{(y_i - y_{i-1})[\phi(\frac{y_i-\mu}{\sigma})]}\right\}.$$

Since, $R(y_i, y_{i-1}, \sigma)$ is dependent on σ, using the proof of Cheng & Amin (1982), the MPS and ML estimators are asymptotically equal and have the same asymptotic sufficiency, consistency and efficiency properties.

Theorem 4.6 *Let $X_1, X_2, \ldots X_n$ be identically and independently distributed random variables with p.d.f. $f(x,\sigma)$, and let the transformation ϕ be one-to-one. Then a MPS estimate is invariant under one-to-one transformation ϕ.*

Proof: The proof is similar to Theorem 4.2 (therefore $\phi(\tilde{\sigma})$ is a MPS estimator of $\phi(\sigma)$).

4.8.1 The MPS estimate of σ in data sets 1 and 2:

1. Using the data set 1 and letting $\mu = 0$ and $\varepsilon = 10^{-5}$, we find that the MPS estimate of σ is

$$\tilde{\sigma} = 1.1366.$$

On plotting H and HP against a certain range of σ [0.1, 3], we get Figures 4.18 and 4.19 (see Appendix Program 30).

2. Using the data set 2 and letting $\mu = 0$ and $\varepsilon = 10^{-5}$, we find that the MPS estimate of σ is

$$\tilde{\sigma} = 1.0320.$$

4.8.2 The MPS estimate of σ in the ideal sample:

In this section we prove that the MPS estimate of σ, for the ideal sample is one. We also plot the H and HP against different values of σ.

Theorem 4.7 *In ideal samples, for every truncation point c the MPS estimator $\tilde{\sigma}$ is one.*

Proof: From Chapter 1, section (1.4.2) we know that

$$F(y_i) = \frac{\Phi(\frac{y_i-\mu}{\sigma})}{\Phi(\frac{c-\mu}{\sigma})} = \frac{i}{n+1}.$$

Hence

$$\Phi(\frac{y_i - \mu}{\sigma}) = \frac{i}{n+1}\Phi(\frac{c - \mu}{\sigma}). \tag{4.33}$$

Figure 4.18: H **versus** σ **for data set 1 (boys)** ($\mu = 0$)

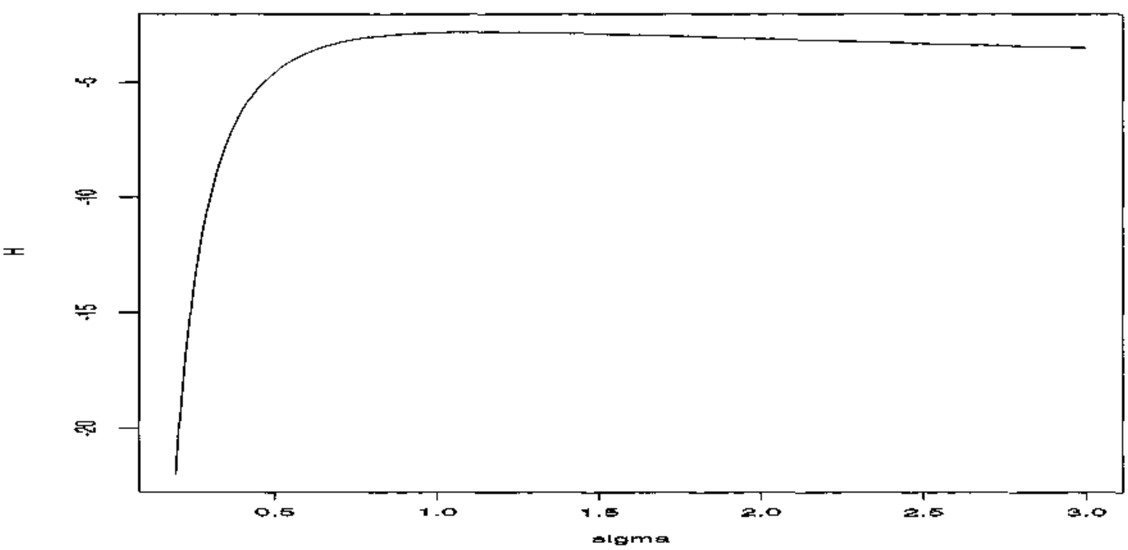
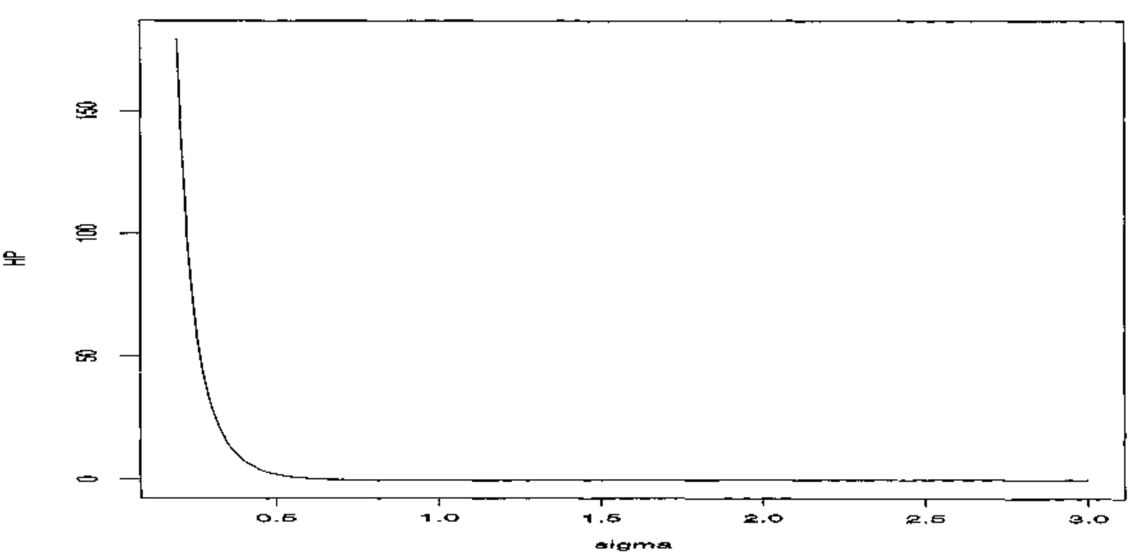


Figure 4.19: HP **versus** σ **for data set 1 (boys)** ($\mu = 0$)



Now, applying formula (4.33), consider

$$\begin{aligned}\sum_{i=1}^{n+1}\left[\frac{(y_i-\mu)\phi(\frac{y_i-\mu}{\sigma})-(y_{i-1}-\mu)\phi(\frac{y_{i-1}-\mu}{\sigma})}{\Phi(\frac{y_i-\mu}{\sigma})-\Phi(\frac{y_{i-1}-\mu}{\sigma})}\right] &= (n+1)\sum_{i=1}^{n+1}\left[\frac{(y_i-\mu)\phi(\frac{y_i-\mu}{\sigma})-(y_{i-1}-\mu)\phi(\frac{y_{i-1}-\mu}{\sigma})}{\Phi(\frac{c-\mu}{\sigma})}\right]\\ &= \frac{n+1}{\Phi(\frac{c-\mu}{\sigma})}[(y_1-\mu)\phi(\frac{y_1-\mu}{\sigma})-(y_0-\mu)\phi(\frac{y_0-\mu}{\sigma})+\cdots\\ &+ (y_{n+1}-\mu)\phi(\frac{y_{n+1}-\mu}{\sigma})-(y_n-\mu)\phi(\frac{y_n-\mu}{\sigma})]\\ &= \frac{n+1}{\Phi(\frac{c-\mu}{\sigma})}[(y_{n+1}-\mu)\phi(\frac{y_{n+1}-\mu}{\sigma})-(y_0-\mu)\phi(\frac{y_0-\mu}{\sigma})]\\ &= \frac{(n+1)(c-\mu)\phi(\frac{c-\mu}{\sigma})}{\Phi(\frac{c-\mu}{\sigma})}.\end{aligned}$$

Using the above equation, it follows that

$$\frac{\partial H}{\partial \sigma^2}\Big|_{\sigma=\tilde{\sigma}} = -\frac{1}{2(n+1)}\left\{\sum_{i=1}^{n+1}\left[\frac{(y_i-\mu)\phi(\frac{y_i-\mu}{\tilde{\sigma}})-(y_{i-1}-\mu)\phi(\frac{y_{i-1}-\mu}{\tilde{\sigma}})}{\Phi(\frac{y_i-\mu}{\tilde{\sigma}})-\Phi(\frac{y_{i-1}-\mu}{\tilde{\sigma}})}\right]\right\}+\frac{(c-\mu)\phi(\frac{c-\mu}{\tilde{\sigma}})}{2\Phi(\frac{c-\mu}{\tilde{\sigma}})}=0.$$

Consequently, $\tilde{\sigma}=1$ maximizes H, and the theorem is proved.

4.9 Simulation study to estimate the variance when the mean is known:

In order to compare the expected value, standard deviation and the variance of the MPS estimate of σ with the ML estimator, we embark upon a simulation study.

4.9.1 The simulation study:

The Program 31 given in Appendix, was written to calculate $E(\tilde{\sigma^2})$, Var$(\tilde{\sigma^2})$ and $\sigma(\tilde{\sigma^2})$ for the $R = 10000$ iterations and sample sizes $n = 5$, 10, 20, 50 and 100. In this program we use NAG routine G05DDF(0,1) to generate random deviates from the normal distribution with mean zero and variance one, and use the Program 29 as a subroutine to solve the equation $\frac{\partial H}{\partial \sigma^2} = 0$. The numerical results are tabulated in the Table 4.11.

Table 4.11: The simulation results for the MPS estimator $\tilde{\sigma^2}$, for different values of n and c, when $\mu = 0$, $\sigma = 1$

n	$c = -1.88$			$c = -1$		
	$E(\tilde{\sigma^2})$	$\sigma(\tilde{\sigma^2})$	$\text{Var}(\tilde{\sigma^2})$	$E(\tilde{\sigma^2})$	$\sigma(\tilde{\sigma^2})$	$\text{Var}(\tilde{\sigma^2})$
5	1.193	0.799	0.638	1.205	0.797	0.636
10	1.179	0.535	0.287	1.117	0.472	0.223
20	1.113	0.350	0.129	1.147	0.364	0.132
50	1.051	0.170	0.029	1.042	0.180	0.032
100	1.028	0.116	0.013	1.034	0.126	0.016

Table 4.11: Continued

n	$c = 0$			$c = 1$		
	$E(\tilde{\sigma^2})$	$\sigma(\tilde{\sigma^2})$	$\text{Var}(\tilde{\sigma^2})$	$E(\tilde{\sigma^2})$	$\sigma(\tilde{\sigma^2})$	$\text{Var}(\tilde{\sigma^2})$
5	1.398	0.897	0.804	1.677	1.135	1.2886
10	1.248	0.554	0.306	1.406	0.710	0.504
20	1.142	0.364	0.132	1.233	0.449	0.202
50	1.073	0.216	0.047	1.116	0.266	0.071
100	1.045	0.148	0.022	1.066	0.182	0.033

Table 4.11: Continued

n	$c = 3$			$c = 10$		
	$E(\tilde{\sigma^2})$	$\sigma(\tilde{\sigma^2})$	$\text{Var}(\tilde{\sigma^2})$	$E(\tilde{\sigma^2})$	$\sigma(\tilde{\sigma^2})$	$\text{Var}(\tilde{\sigma^2})$
5	1.812	1.648	2.717	1.199	0.807	0.652
10	1.429	0.865	0.748	1.184	0.538	0.290
20	1.230	0.484	0.235	1.137	0.357	0.127
50	1.103	0.243	0.059	1.095	0.219	0.048
100	1.067	0.158	0.025	1.064	0.151	0.023

From the Table 4.11 which extended, we can see that for each truncation point c, when the sample size n is increased, all the values of $\sigma(\tilde{\sigma^2})$ and $\text{Var}(\tilde{\sigma^2})$ decrease. We can also see that for all values of c, the bias of $\tilde{\sigma^2}$ is rather high. To study the bias further, we concentrate on the values $c = -1.88$ and $c = 3$. We find the biases of other functions of $\tilde{\sigma}$.

Running the Program 31 for these two values of the truncation point, with $R = 1000$ and with $\sigma = 1$, $\sigma = 2$ and $\sigma = 3$ calculated four functions of $(\tilde{\sigma})$, to see which function has the smallest $\frac{f(\tilde{\sigma})}{f(\sigma)}$. The functions are $\tilde{\sigma}$, $\tilde{\sigma^2}$, $\frac{1}{\tilde{\sigma}}$ and $\frac{1}{\tilde{\sigma^2}}$ and the results are tabulated in Tables 4.12 and 4.13.

Table 4.12: The bias of different functions of the MPS estimator σ, for different values of σ when $n = 5$, $c = -1.88$ and $R = 1000$

$f(\tilde{\sigma})$	$\sigma = 1$			$\sigma = 2$			$\sigma = 3$		
	Mean	Exact	$\frac{f(\tilde{\sigma})}{f(\sigma)}$	Mean	Exact	$\frac{f(\tilde{\sigma})}{f(\sigma)}$	Mean	Exact	$\frac{f(\tilde{\sigma})}{f(\sigma)}$
$\tilde{\sigma}$	0.964	1	0.964	2.191	2	1.095	3.285	3	1.095
$\tilde{\sigma^2}$	1.369	1	1.369	5.092	4	1.0273	11.990	9	1.332
$\frac{1}{\tilde{\sigma}}$	0.9510	1	0.9510	0.519	0.5	1.038	0.332	0.333	0.996
$\frac{1}{\tilde{\sigma^2}}$	0.989	1	0.989	0.279	0.25	1.116	0.122	0.111	1.099

Table 4.13: For $c = 3$

$g(\tilde{\sigma})$	$\sigma = 1$			$\sigma = 2$			$\sigma = 3$		
	Mean	Exact	$\frac{f(\tilde{\sigma})}{f(\sigma)}$	Mean	Exact	$\frac{f(\tilde{\sigma})}{f(\sigma)}$	Mean	Exact	$\frac{f(\tilde{\sigma})}{f(\sigma)}$
$\tilde{\sigma}$	1.235	1	1.235	2.557	2	1.278	3.693	3	1.231
$\tilde{\sigma^2}$	1.870	1	1.870	7.302	4	1.825	15.230	9	1.692
$\frac{1}{\tilde{\sigma}}$	1.034	1	1.034	0.463	0.5	0.926	0.325	0.333	0.975
$\frac{1}{\tilde{\sigma^2}}$	1.437	1	1.437	0.260	0.25	1.040	0.144	0.111	1.297

Concentrating on Table 4.13 for $c = 3$, we can see that $\frac{f(\tilde{\sigma})}{f(\sigma)}$ of $\tilde{\sigma^2}$ is more than $\frac{f(\tilde{\sigma})}{f(\sigma)}$ of its counterpart, $\tilde{\sigma}$, whereas $\frac{f(\tilde{\sigma})}{f(\sigma)}$ of $\frac{1}{\tilde{\sigma}}$ is less than that of its counterpart, $\frac{1}{\tilde{\sigma^2}}$. We also see in the fourth, seventh and tenth columns of Table 4.13, that for each value of σ, $\frac{f(\tilde{\sigma})}{f(\sigma)} > 1$, except for the function $\frac{1}{\tilde{\sigma}}$ with $\sigma = 2$, $\sigma = 3$.

Since, according to Theorem 4.6, $\tilde{\sigma}$ is invariant and for $f(\tilde{\sigma}) = \frac{1}{\tilde{\sigma}}$, we can see $\frac{f(\tilde{\sigma})}{f(\sigma)} \approx 1$. Therefore we suggest that to have a less bias in our estimator, we prefer to estimate $\frac{1}{\sigma}$.

4.10 Relationship between $E(\tilde{\sigma^2})$ and sample size and truncation point:

In this section we investigate the relationship of $E(\tilde{\sigma^2})$ with sample size n and truncation point c.

To find an appropriate regression model, we have to find $\widetilde{E(\tilde{\sigma^2})}$ for different values of c and n. Since we are interested to see what happens to $E(\tilde{\sigma^2})$ as $c \to \infty$, we calculate $\widetilde{E(\tilde{\sigma^2})}$ for values of c up to $c = 10$ inclusive. The values have been obtained and are shown in Table 4.14. By looking at Figure 4.20 we can see that the bias of $\tilde{\sigma^2}$ reduces with increasing the sample size n, and decreases as the sample size n increases.

Table 4.14: The expected value of $\tilde{\sigma^2}$, by simulation for different values of n and c when $\mu = 0$, $\sigma = 1$ and simulation run $R = 100000$

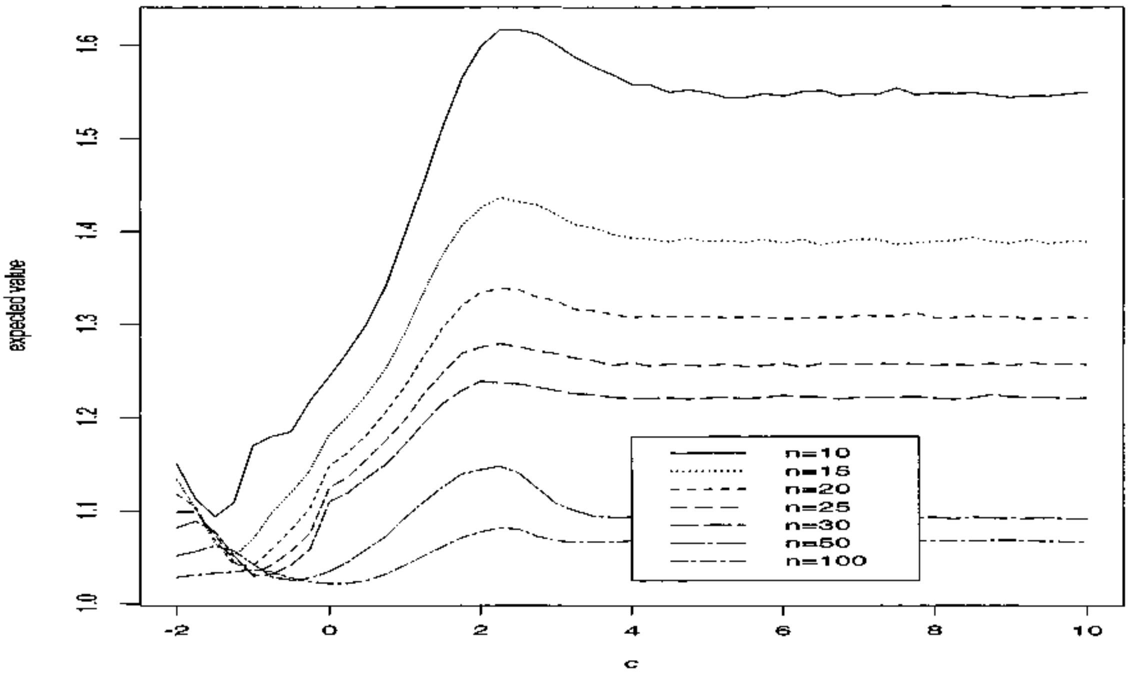
c	$\widetilde{E(\tilde{\sigma^2})}$					c	$\widetilde{E(\tilde{\sigma^2})}$				
	$n=10$	$n=15$	$n=20$	$n=25$	$n=30$		$n=10$	$n=15$	$n=20$	$n=25$	$n=30$
-2	1.150	1.134	1.118	1.099	1.082	4.25	1.557	1.392	1.309	1.255	1.221
-1.75	1.113	1.102	1.104	1.099	1.089	4.50	1.549	1.389	1.309	1.258	1.222
-1.5	1.094	1.065	1.069	1.075	1.078	4.75	1.552	1.393	1.309	1.256	1.220
-1.25	1.109	1.051	1.042	1.045	1.052	5.00	1.548	1.390	1.309	1.255	1.221
-1	1.170	1.070	1.041	1.032	1.030	5.25	1.543	1.390	1.307	1.258	1.222
-0.75	1.180	1.098	1.059	1.042	1.032	5.50	1.544	1.388	1.308	1.255	1.220
-0.5	1.185	1.120	1.080	1.057	1.042	5.75	1.548	1.392	1.308	1.256	1.221
-0.25	1.218	1.143	1.103	1.076	1.059	6.00	1.545	1.388	1.306	1.258	1.224
0	1.244	1.181	1.148	1.124	1.109	6.25	1.550	1.391	1.307	1.254	1.222
0.25	1.270	1.201	1.162	1.136	1.119	6.50	1.551	1.386	1.307	1.258	1.222
0.5	1.301	1.225	1.182	1.155	1.135	6.75	1.545	1.389	1.307	1.257	1 219
0 75	1.341	1.252	1.206	1.175	1.149	7.00	1.547	1.392	1.309	1.257	1.222
1	1.397	1.397	1.290	1.233	1.197	7.25	1.547	1.391	1.308	1.257	1.221
1.25	1.455	1.335	1.265	1.225	1.194	7.50	1.553	1.386	1.308	1.257	1.221
1.5	1.513	1.375	1.296	1.247	1.215	7.75	1.546	1.387	1.312	1.257	1.222
1.75	1.565	1.406	1.320	1.269	1.229	8.00	1.548	1.388	1.307	1.256	1.220
2	1.596	1.425	1.334	1.275	1.239	8.25	1.547	1.390	1.307	1.256	1.219
2.25	1.616	1.436	1.338	1.279	1.237	8.50	1.548	1.393	1.308	1.256	1.220
2.5	1.616	1.432	1.337	1.276	1.236	8.75	1.545	1.389	1.308	1.258	1.224
2.75	1.619	1.428	1.329	1.271	1.233	9.00	1.543	1.387	1.307	1.255	1.222
3	1.600	1.418	1.325	1.269	1.229	9.25	1.545	1.391	1.306	1.259	1.222
3.25	1.587	1.408	1.316	1.264	1.226	9.50	1.545	1.387	1.306	1.257	1.222
3.50	1.577	1.405	1.315	1.261	1.225	9.75	1.548	1.389	1.307	1.257	1.220
3.75	1.568	1.397	1.311	1.256	1.222	10.00	1.549	1.389	1.307	1.257	1.221
4.00	1.558	1.393	1.308	1.259	1.221						

The following plot shows that, for sample sizes $n = 10, 15, 20, 25$, 30, 50 and 100 as the truncation point c increases, then $\widetilde{E(\tilde{\sigma^2})} \to k$, a constant. Since the differences of values $\widetilde{E(\tilde{\sigma^2})}$ in Table 4.14 are substantially different from one and $\tilde{\sigma^2}$ is consistent. Then appropriate model for $E(\tilde{\sigma^2})$ can be written as

$$E(\tilde{\sigma^2}) = \sigma^2 + \frac{1}{n} g_1(c) + \frac{1}{n^2} g_2(c) + O(n^{-3}). \tag{4.34}$$

If we use two different values of n, we obtain two different equations. Then by simultaneous solution of the equations, we find their corresponding $\widetilde{g_1(c)}$ and $\widetilde{g_2(c)}$. In this section similar

Figure 4.20: The plot of $\widetilde{E(\tilde{\sigma}^2)}$ against c for different n



to section 4.4 we choose $n = 10$ and $n = 20$. Then the calculated $\widetilde{g_1(c)}$ and $\widetilde{g_2(c)}$ are plotted against c in Figure 4.21.

By looking at Figure 4.21 we can see that all the values of $\widetilde{g_1(c)}$ are positive and, as c increases, $\widetilde{g_1(c)}$ approaches to the constant $k \approx 6$. But in Figure 4.22 most of the values of $\widetilde{g_2(c)}$ are positive and, as c increases, $\widetilde{g_2(c)}$ approaches to the constant $k \approx -11$. Now, we are interested to find the functions $g_1(c)$ and $g_2(c)$ in terms of c. From the Figures 4.21, and 4.22 we guess that the models should follow the $\alpha_5 ce^{-c} + \beta_5$. Note that the error term in this model is additive. Using a Macro in GLIM 4 software (see Appendix Program 33), Ekholm & Green (1993), which used numerical derivatives for fitting nonlinear models and assumes normality of error terms, by fitting the model and entering the c^2, c^3 etc terms in

Figure 4.21: The plot of $\widetilde{g_1(c)}$ against c

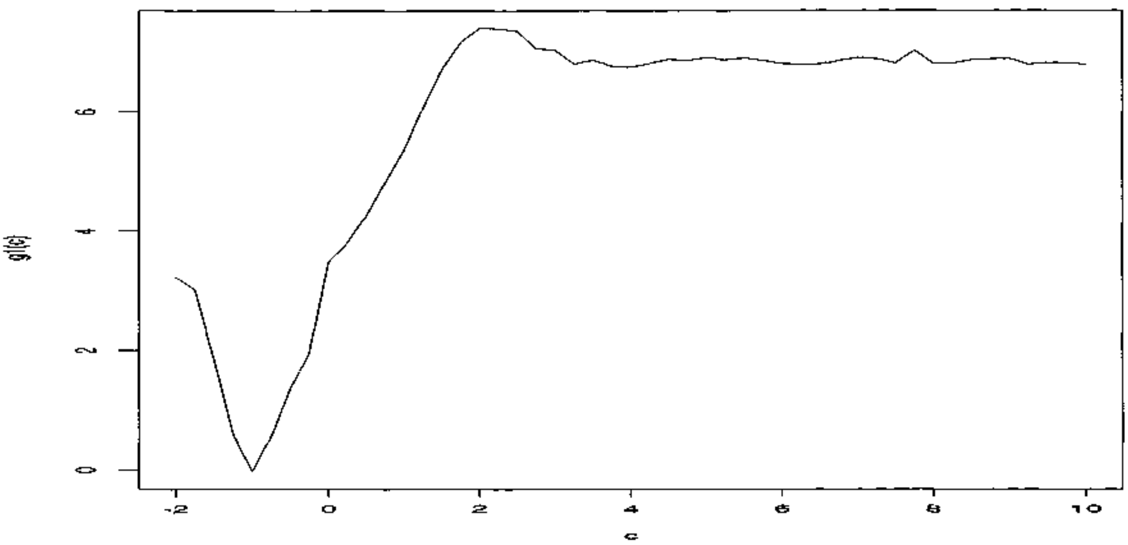
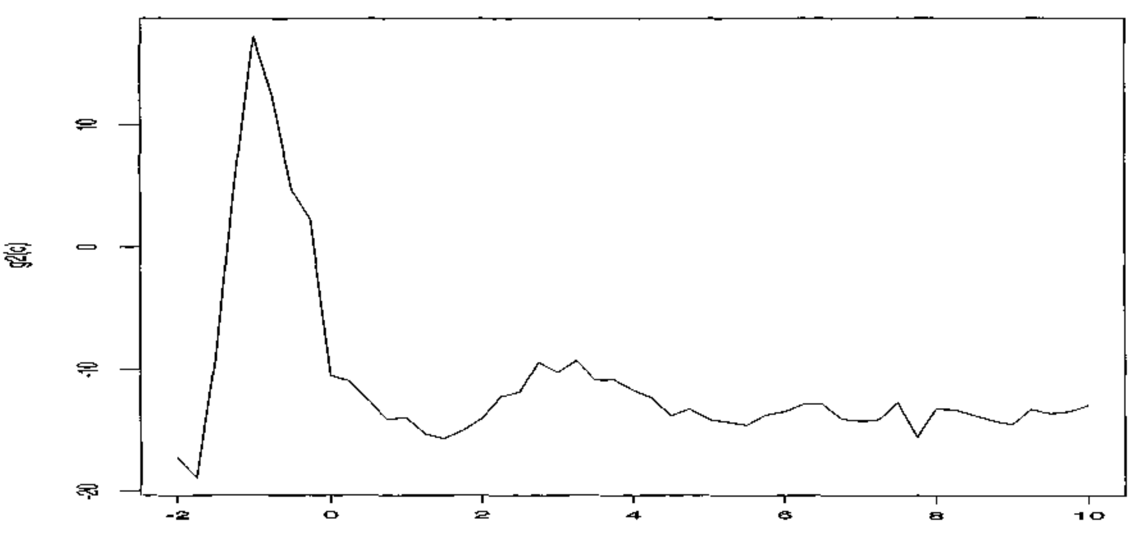


Figure 4.22: The plot of $\widetilde{g_2(c)}$ against c



the model we found the following models

$$g_1(c) = (1.345c^2 + 2.814c)e^{-c} + 5.829$$

where the corresponding $t-$ratios of the coefficients 1.345, 2.814 and 5.829 are 6.23, 7.14 and 28.85.

Further

$$g_2(c) = (-5.492c^2 - 9.961c)e^{-c} - 10.20$$

where the corresponding $t-$ratios of the coefficients -5.492, -9.961 and -10.20 are -6.49, -6.48 and -12.97. (see Appendix Programs 32 and 33). We also, checked the residual plots, and they confirmed the validity of the models.

Now, using these equations we can estimate $g_1(c)$ and $g_2(c)$ for any values of c.

Substituting the estimated $g_1(c)$ and $g_2(c)$ into equation (4.34) we obtain

$$\begin{aligned} E(\tilde{\sigma^2}) &= \sigma^2 + \frac{1}{n}\left\{(1.345c^2 + 2.814c)e^{-c} + 5.829\right\} \\ &+ \frac{1}{n^2}\left\{(-5.492c^2 - 9.961c)e^{-c} - 10.20\right\} + O(n^{-3}). \end{aligned}$$

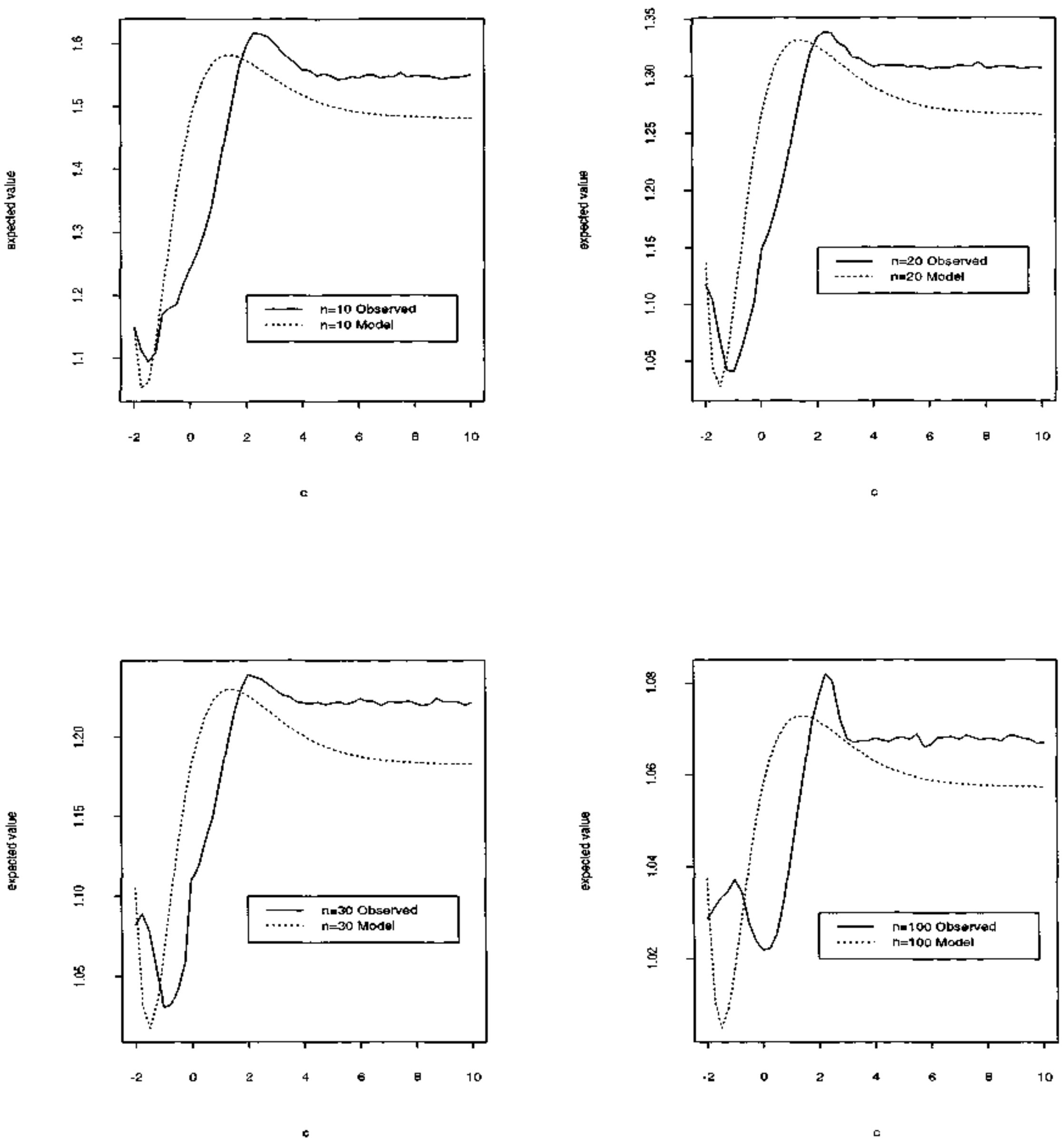
To compare $E(\tilde{\sigma^2})$ calculated from the model together with the $\widetilde{E(\tilde{\sigma^2})}$, calculated from the simulation, we plot them in Figure 4.23.

Figures 4.23 shows that if we shift c to $+0.75$ the values of $E(\tilde{\sigma^2})$ approach to the observe values $\widetilde{E(\tilde{\sigma^2})}$, as the sample size n increases. We conclude that these models are reliable and we can easily find $E(\tilde{\sigma^2})$ for various values of n and c.

4.11 Relationship between Var($\tilde{\sigma^2}$) and sample size and truncation point:

In this section we investigate the relationship of Var($\tilde{\sigma^2}$) with sample size n and truncation point c.

Figure 4.23: The plot of $E(\tilde{\sigma^2})$ against c for different n



To find the appropriate regression model, we have to find $\widetilde{\mathrm{Var}}(\tilde{\sigma}^2)$ for different values of c and n. Since we are interested to see what happens to $\widetilde{\mathrm{Var}}(\tilde{\sigma}^2)$ as $c \to \infty$ we have calculated $\widetilde{\mathrm{Var}}(\tilde{\sigma}^2)$ for values of c up to $c = 10$ inclusive. The values have been obtained and are shown in Table 4.15 and Figure 4.24.

Table 4.15: The variance of $\tilde{\sigma^2}$, by simulation for different values of n and c when $\mu = 0$, $\sigma = 1$ and simulation run $R = 100000$

c	$\mathrm{Var}(\tilde{\sigma^2})$					c	$\mathrm{Var}(\tilde{\sigma^2})$				
	$n=10$	$n=15$	$n=20$	$n=25$	$n=30$		$n=10$	$n=15$	$n=20$	$n=25$	$n=30$
-2	0.213	0.110	0.081	0.063	0.052	4.25	0.514	0.264	0.175	0.128	0.100
-1.75	0.187	0.093	0.072	0.060	0.051	4.50	0.503	0.262	0.172	0.128	0.100
-1.5	0.195	0.097	0.068	0.056	0.048	4.75	0.498	0.263	0.173	0.127	0.099
-1.25	0.203	0.112	0.076	0.059	0.049	5.00	0.488	0.259	0.171	0.125	0.100
-1	0.215	0.126	0.086	0.066	0.053	5.25	0.486	0.259	0.172	0.126	0.100
-0.75	0.222	0.133	0.094	0.072	0.059	5.50	0.484	0.259	0.171	0.127	0.100
-0.5	0.249	0.148	0.103	0.079	0.063	5.75	0.484	0.260	0.171	0.126	0.099
-0.25	0.281	0.165	0.115	0.087	0.071	6.00	0.482	0.259	0.171	0.127	0.100
0	0.309	0.185	0.132	0.101	0.082	6.25	0.482	0.260	0.171	0.126	0.100
0.25	0.350	0.211	0.149	0.113	0.091	6.50	0.486	0.259	0.171	0.128	0.099
0.5	0.400	0.239	0.168	0.128	0.104	6.75	0.483	0.259	0.172	0.127	0.098
0.75	0.442	0.265	0.185	0.143	0.115	7.00	0.485	0.260	0.173	0.126	0.100
1	0.498	0.292	0.206	0.158	0.126	7.25	0.481	0.259	0.171	0.128	0.100
1.25	0.535	0.318	0.218	0.168	0.135	7.50	0.489	0.257	0.171	0.126	0.099
1.5	0.570	0.335	0.232	0.176	0.141	7.75	0.481	0.259	0.173	0.128	0.100
1.75	0.614	0.346	0.237	0.179	0.141	8.00	0.485	0.257	0.171	0.126	0.100
2	0.632	0.355	0.237	0.175	0.139	8.25	0.482	0.260	0.173	0.127	0.100
2.25	0.651	0.352	0.233	0.169	0.132	8.50	0.481	0.261	0.171	0.126	0.100
2.5	0.653	0.341	0.222	0.160	0.125	8.75	0.481	0.260	0.172	0.126	0.100
2.75	0.627	0.330	0.212	0.152	0.118	9.00	0.479	0.256	0.171	0.126	0.100
3	0.610	0.312	0.201	0.144	0.112	9.25	0.480	0.258	0.171	0.128	0.100
3.25	0.590	0.296	0.189	0.138	0.107	9.50	0.484	0.255	0.171	0.126	0.100
3.50	0.564	0.289	0.183	0.134	0.104	9.75	0.485	0.257	0.171	0.128	0.099
3.75	0.549	0.275	0.180	0.131	0.103	10.00	0.484	0.260	0.170	0.126	0.099
4.00	0.525	0.271	0.174	0.130	0.101						

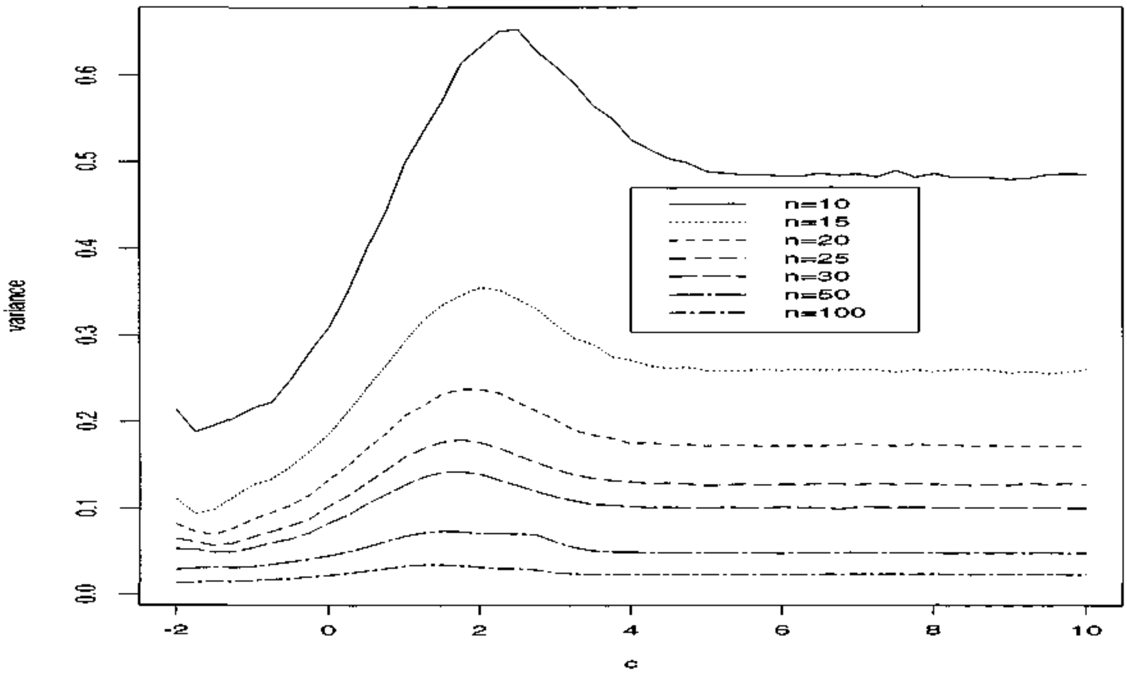
Let us assume that $\mathrm{Var}(\tilde{\sigma^2})$ can be written as

$$\mathrm{Var}(\tilde{\sigma^2}) = \frac{1}{n}g_1(c) + \frac{1}{n^2}g_2(c) + \frac{1}{n^3}g_3(c) + O(n^{-4}). \tag{4.35}$$

Since we know that $\hat{\sigma}^2$ and $\tilde{\sigma^2}$ are asymptotically equivalent, therefore $\mathrm{Var}(\tilde{\sigma^2}) = \mathrm{Var}(\hat{\sigma}^2)$ for $O(n^{-1})$. Hence $g_1(c) = \frac{2\sigma^4}{D}$, and equation (4.35) can be written as

$$\mathrm{Var}(\tilde{\sigma^2}) = \frac{1}{n}\left[\frac{2\sigma^4}{D}\right] + \frac{1}{n^2}g_2(c) + \frac{1}{n^3}g_3(c) + O(n^{-4}). \tag{4.36}$$

Figure 4.24: The plot of $\widetilde{\text{Var}(\tilde{\sigma}^2)}$ against c for different n



Similar to section 4.4 choosing $n = 10$ and $n = 20$. We find the corresponding $\widetilde{g_2(c)}$ and $\widetilde{g_3(c)}$ and plot them against c in Figures 4.25 and 4.26.

By looking at Figure 4.25, we can see that almost all the values of $\widetilde{g_2(c)}$ are positive and that $\widetilde{g_2(c)}$ increases as c increases. It has a maximum point at $c = 3$, then $\widetilde{g_2(c)}$ decreases as c increases and we can say $\widetilde{g_2(c)} \to\approx 27$ as $c \to \infty$. But in Figure 4.26, $\widetilde{g_3(c)}$ has two obvious local maxima and one minimum in its domain. Also $\widetilde{g_3(c)} \to\approx -30$ as $c \to \infty$.

Now we find the functions $g_2(c)$ and $g_3(c)$ in terms of c. Again we use similar procedure of section 4.10 for entering the c, c^2 etc terms in the model.

Using the GLIM 4 software we have found the following models

$$g_2(c) = (7.279c^2 + 15.71c)e^{-c} + 26.31$$

Figure 4.25: The plot of $\widetilde{g_2(c)}$ against c

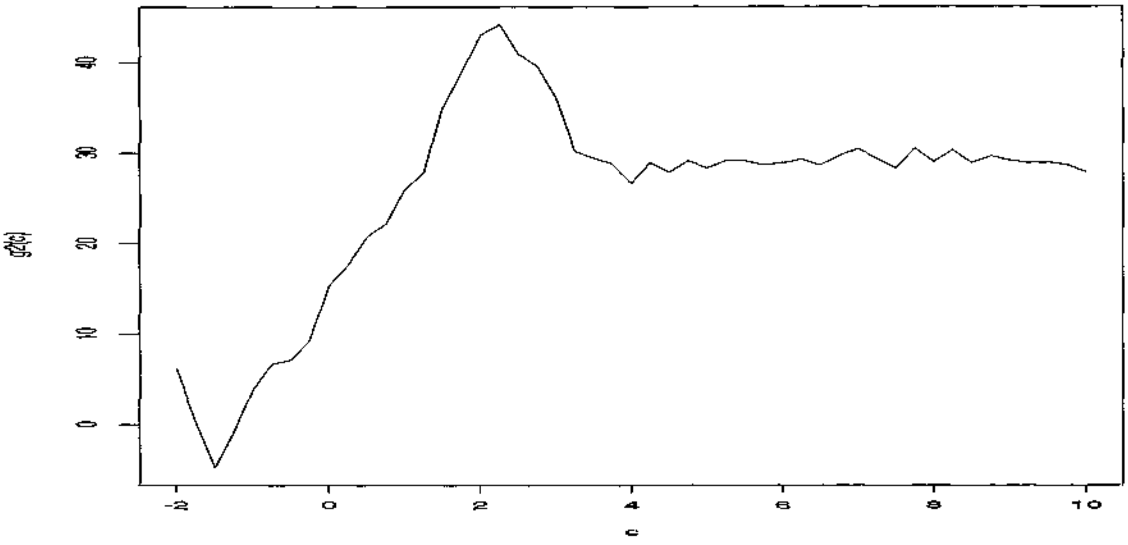
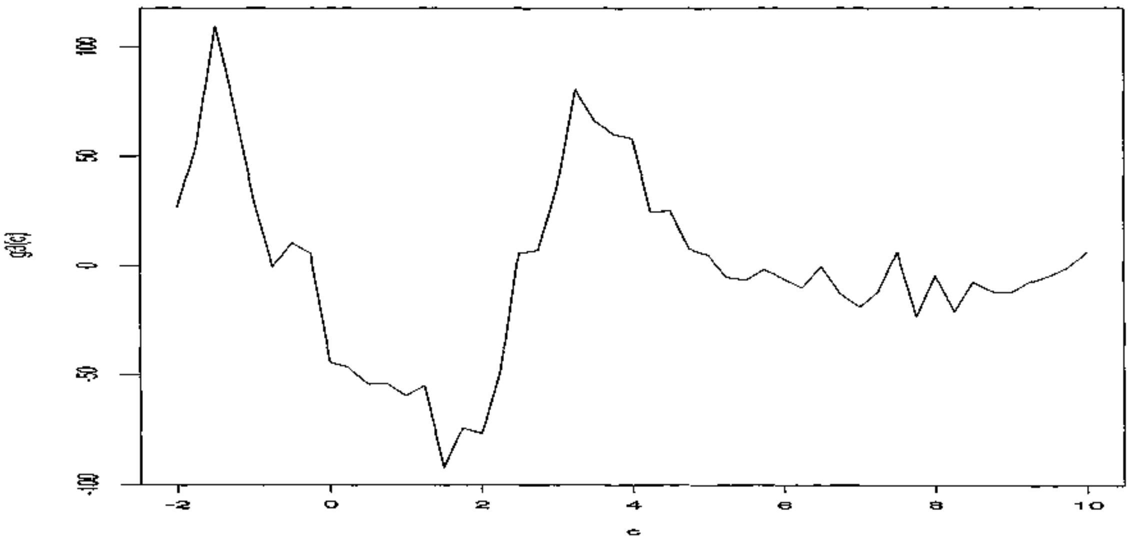


Figure 4.26: The plot of $\widetilde{g_3(c)}$ against c



where the corresponding $t-$ratios of coefficients 7.279, 15.71 and 26.31 are 8.109, 9.632 and 31.49. We also, checked the residual plot, and they confirmed the validity of the models.

Further

$$g_3(c) = (1.207c^5 + 6.234c^4 + 17.90c^3 - 45.04c)e^{-c} - 42.21$$

where the corresponding $t-$ratios of the coefficients 1.207, 6.234, 17.90, -45.04 and -42.21 are 1.91, 3.76, 2.65, 5.12 and 6.17.

Now, using these equations we can find $g_2(c)$ and $g_3(c)$ for any values of c.

Substituting the $g_2(c)$ and $g_3(c)$ in equation (4.36) we obtain

$$\begin{aligned} \text{Var}(\tilde{\sigma^2}) &= \frac{1}{n}\left[\frac{2\sigma^4}{D}\right] \\ &+ \frac{1}{n^2}\left[(7.279c^2 + 15.71c)e^{-c} + 26.31\right] \\ &+ \frac{1}{n^3}\left[(1.207c^5 + 6.234c^4 + 17.90c^3 - 45.04c)e^{-c} - 42.21\right] + O(n^{-4}). \end{aligned}$$

To compare $\text{Var}(\tilde{\sigma^2})$ calculated from the model together with the $\widetilde{\text{Var}(\tilde{\sigma^2})}$ calculated from the simulation, we plot them in Figure 4.27.

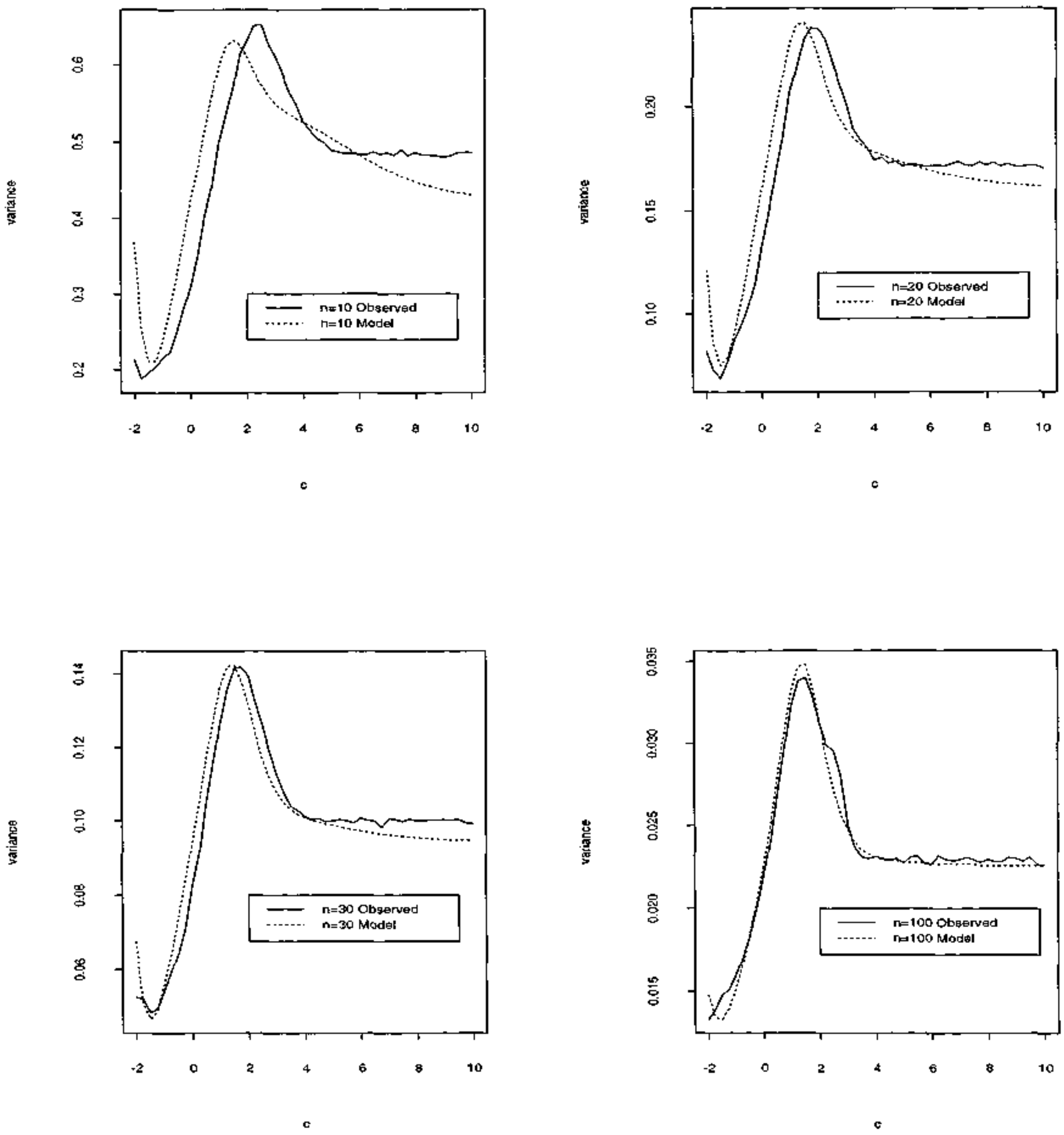
Figure 4.27 shows that if we shift c to $\frac{1.25}{n}$ the value of $\text{Var}(\tilde{\sigma^2})$ approach to the observe value $\widetilde{\text{Var}(\tilde{\sigma^2})}$, as the sample size n increases. We conclude that these models are reliable and we can easily find $\text{Var}(\tilde{\sigma^2})$ for various values of n and c.

4.12 The comparison of the different estimators of σ^2:

In this section we compare the two estimators of σ^2 based on the MLE and the MPS methods with the theoretical results for MLE, and the exact value of σ^2.

By using the results of the simulation study for the maximum likelihood estimator, the

Figure 4.27: The plot of $\text{Var}(\tilde{\sigma^2})$ and $\widetilde{\text{Var}}(\tilde{\sigma^2})$ against c for different n



theoretical results of Chapter 2 and the results of the simulation study for the maximum spacing methods, described in this chapter, we obtained Tables 4.16–4.19.

Table 4.16: The comparison of the MLE and the MPS estimators of σ^2 for $c = -1.88$ and different values of n

n	$\hat{\sigma}^2$ Sim.mean	$\hat{\sigma}^2$ Sim.var.	$\hat{\sigma}^2$ MSE	$\hat{\sigma}^2$ Theory mean	$\hat{\sigma}^2$ Theory var.	$\tilde{\sigma}^2$ Sim.mean	$\tilde{\sigma}^2$ Sim.var.	$\tilde{\sigma}^2$ MSE
5	1.019	0.497	0.497	1.092	0.497	1.320	0.856	0.958
10	1.006	0.188	0.188	1.026	0.188	1.138	0.194	0.213
20	1.004	0.078	0.078	1.008	0.078	1.118	0.079	0.093
50	1.004	0.028	0.028	1.002	0.028	1.051	0.028	0.030
100	1.001	0.013	0.0131	1.001	0.013	1.028	0.013	0.014

Table 4.17: The comparison of the MLE and the MPS estimators of σ^2 for $c = -1$ and different values of n

n	$\hat{\sigma}^2$ Sim.mean	$\hat{\sigma}^2$ Sim.var.	$\hat{\sigma}^2$ MSE	$\hat{\sigma}^2$ Theory mean	$\hat{\sigma}^2$ Theory var.	$\tilde{\sigma}^2$ Sim.mean	$\tilde{\sigma}^2$ Sim.var.	$\tilde{\sigma}^2$ MSE
5	1.018	0.295	0.295	1.012	0.32	1.548	0.462	0.762
10	1.011	0.147	0.147	1.007	0.153	1.117	0.222	0.0235
20	1.003	0.075	0.075	1.004	0.075	1.042	0.088	0.089
50	1.001	0.030	0.030	1.002	0.030	1.038	0.032	0.033
100	1.000	0.015	0.015	1.001	0.015	1.034	0.015	0.016

Table 4.18: The comparison of the MLE and the MPS estimators of σ^2 for $c = 0$ and different values of n

n	$\hat{\sigma}^2$ Sim.mean	$\hat{\sigma}^2$ Sim.var.	$\hat{\sigma}^2$ MSE	$\hat{\sigma}^2$ Theory mean	$\hat{\sigma}^2$ Theory var.	$\tilde{\sigma}^2$ Sim.mean	$\tilde{\sigma}^2$ Sim.var.	$\tilde{\sigma}^2$ MSE
5	1.008	0.492	0.492	1.000	0.400	1.397	0.778	0.936
10	0.996	0.195	0.195	1.000	0.200	1.248	0.310	0.371
20	1.000	0.100	0.100	1.000	0.100	1.142	0.124	0.144
50	1.001	0.041	0.041	1.000	0.040	1.073	0.044	0.049
100	1.000	0.020	0.020	1.000	0.020	1.045	0.021	0.023

Table 4.19: The comparison of the MLE and the MPS estimators of σ^2 for $c = 1$ and different values of n

n	$\hat{\sigma}^2$ Sim.mean	$\hat{\sigma}^2$ Sim.var.	$\hat{\sigma}^2$ MSE	$\hat{\sigma}^2$ Theory mean	$\hat{\sigma}^2$ Theory var.	$\tilde{\sigma}^2$ Sim.mean	$\tilde{\sigma}^2$ Sim.var.	$\tilde{\sigma}^2$ MSE
5	1.009	0.771	0.771	0.977	0.532	1.677	1.240	1.690
10	0.983	0.278	0.278	0.984	0.282	1.406	0.484	0.648
20	0.992	0.144	0.144	0.990	0.145	1.233	0.201	0.255
50	0.996	0.058	0.058	0.996	0.059	1.116	0.067	0.080
100	0.999	0.030	0.030	0.998	0.030	1.066	0.031	0.035

4.13 Conclusion:

Comparing the estimated value of σ^2 from the two estimation methods, we see that in Table 4.16 for $n = 5$ and $c = -1.88$, the MPS method has a mean value of $\tilde{\sigma}^2 = 1.320$, while the MLE method has a mean $\hat{\sigma}^2 = 1.019$. We can see also that $\sigma(\tilde{\sigma}^2)$ for the MPS method is bigger in almost every cell than the $\sigma(\hat{\sigma}^2)$ from the MLE method.

Comparing the variances for ML estimator (0.497) with the MPS method (0.856), show that the variance of ML estimator is high, but by increasing the sample size n, for example $n = 100$, the variances of $\tilde{\sigma}^2$ and $\hat{\sigma}^2$ become the same. Therefore we can conclude that the variance of MPS estimator is asymptotically equivalent to the variance of ML estimator for estimating the variance of the distribution. We can make similar comparisons for other values of c in Tables 4.16-4.19. As we see from Tables 4.16-4.19 the larger the truncation points, the closer the variance of two estimators. Moreover, for large values of n these two estimators are almost identical, which is in line with Theorem 4.5.

Chapter 5

The two parameter case of maximum product spacing in the truncated normal distribution:

5.1 Introduction:

The purpose of this chapter is to describe the maximum product spacing method of estimating the parameters of a distribution, simultaneously. The application of maximum product spacing in the truncated normal distribution with both parameters unknown is considered. Further, a simulation study is curried out to investigate the expected value and variance of the MPS estimators of the parameters.

5.2 Estimation of μ and σ when both are unknown:

In this section, we are going to find the MPS estimator $(\tilde{\mu}, \tilde{\sigma})$ of (μ, σ) for the truncated normal distribution when the mean and variance are both unknown, and also to make a comparison with $(\hat{\mu}, \hat{\sigma})$. Taking the derivative of H in equation (4.7) with respect to μ and σ^2, we obtain

$$\frac{\partial H}{\partial \mu} = -\frac{1}{\sigma(n+1)}\left\{\sum_{i=1}^{n+1}\left[\frac{\phi(\frac{y_i-\mu}{\sigma}) - \phi(\frac{y_{i-1}-\mu}{\sigma})}{\Phi(\frac{y_i-\mu}{\sigma}) - \Phi(\frac{y_{i-1}-\mu}{\sigma})}\right]\right\} + \frac{\phi(\frac{c-\mu}{\sigma})}{\sigma\Phi(\frac{c-\mu}{\sigma})} \tag{5.1}$$

and

$$\frac{\partial H}{\partial \sigma^2} = -\frac{1}{2\sigma^3(n+1)}\left\{\sum_{i=1}^{n+1}\left[\frac{(y_i-\mu)\phi(\frac{y_i-\mu}{\sigma}) - (y_{i-1}-\mu)\phi(\frac{y_{i-1}-\mu}{\sigma})}{\Phi(\frac{y_i-\mu}{\sigma}) - \Phi(\frac{y_{i-1}-\mu}{\sigma})}\right]\right\} + \frac{(c-\mu)\phi(\frac{c-\mu}{\sigma})}{2\sigma^3\Phi(\frac{c-\mu}{\sigma})}. \tag{5.2}$$

The simultaneous solution of the equations

$$\begin{cases} \frac{\partial H}{\partial \mu} = 0 \\ \frac{\partial H}{\partial \sigma^2} = 0 \end{cases} \tag{5.3}$$

or

$$\begin{cases} -\frac{1}{\sigma(n+1)}\left\{\sum_{i=1}^{n+1}\left[\frac{\phi(\frac{y_i-\mu}{\sigma})-\phi(\frac{y_{i-1}-\mu}{\sigma})}{\Phi(\frac{y_i-\mu}{\sigma})-\Phi(\frac{y_{i-1}-\mu}{\sigma})}\right]\right\} + \frac{\phi(\frac{c-\mu}{\sigma})}{\sigma\Phi(\frac{c-\mu}{\sigma})} = 0 \\ -\frac{1}{2\sigma^3(n+1)}\left\{\sum_{i=1}^{n+1}\left[\frac{(y_i-\mu)\phi(\frac{y_i-\mu}{\sigma})-(y_{i-1}-\mu)\phi(\frac{y_{i-1}-\mu}{\sigma})}{\Phi(\frac{y_i-\mu}{\sigma})-\Phi(\frac{y_{i-1}-\mu}{\sigma})}\right]\right\} + \frac{(c-\mu)\phi(\frac{c-\mu}{\sigma})}{2\sigma^3\Phi(\frac{c-\mu}{\sigma})} = 0. \end{cases} \tag{5.4}$$

with respect to μ and σ gives $(\tilde{\mu}, \tilde{\sigma})$.

The algebraic solution of equation (5.4) is impossible. Therefore we use the NAG routine C05NBF in Program 34 given in Appendix to solve it iteratively.

5.2.1 The MPS estimator of (μ, σ) in data sets 1 and 2:

1. Using the data set 1 and $\varepsilon = 10^{-5}$, we find that the MPS estimates of (μ, σ) is $(\tilde{\mu}, \tilde{\sigma}) = (0.0603, 1.1079)$.

2. Using the data set 2 and $\varepsilon = 10^{-5}$, we find that the MPS estimates of (μ, σ) is $(\tilde{\mu}, \tilde{\sigma}) = (-0.0158, 1.0396)$.

The ML estimates for data set 1 is $(\hat{\mu}, \hat{\sigma}) = (1.3377, 1.2871)$ and for data set 2 is $(\hat{\mu}, \hat{\sigma}) = (-2.1932, 0.2767)$. When we compare the MPS estimates $(\tilde{\mu}, \tilde{\sigma})$ with the ML estimates $(\hat{\mu}, \hat{\sigma})$, we see that in the two data sets, the difference of the MPS estimates from the true values $(\mu = 0, \sigma = 1)$ are less than their counterparts in the ML estimates.

5.2.2 The MPS estimator of (μ, σ) in ideal samples:

In this section we prove a desirable property of the MPS estimator, namely that it gives the correct answer for an ideal sample. Also we plot H against different values of μ and σ (see Appendix Program 35).

Theorem 5.1 *In ideal samples, for every truncation point c the MPS estimator $(\tilde{\mu}, \tilde{\sigma})$ is* $(0, 1)$.

Proof: In Chapter 4, Theorem 4.3 and Theorem 4.7 as we proved for ideal samples,

$$\begin{aligned}\sum_{i=1}^{n+1}\left[\frac{\phi(\frac{y_i-\mu}{\sigma})-\phi(\frac{y_{i-1}-\mu}{\sigma})}{\Phi(\frac{y_i-\mu}{\sigma})-\Phi(\frac{y_{i-1}-\mu}{\sigma})}\right] &= (n+1)\sum_{i=1}^{n+1}\left[\frac{\phi(\frac{y_i-\mu}{\sigma})-\phi(\frac{y_{i-1}-\mu}{\sigma})}{\Phi(\frac{c-\mu}{\sigma})}\right] \\ &= \frac{(n+1)\phi(\frac{c-\mu}{\sigma})}{\Phi(\frac{c-\mu}{\sigma})}. \end{aligned} \tag{5.5}$$

and

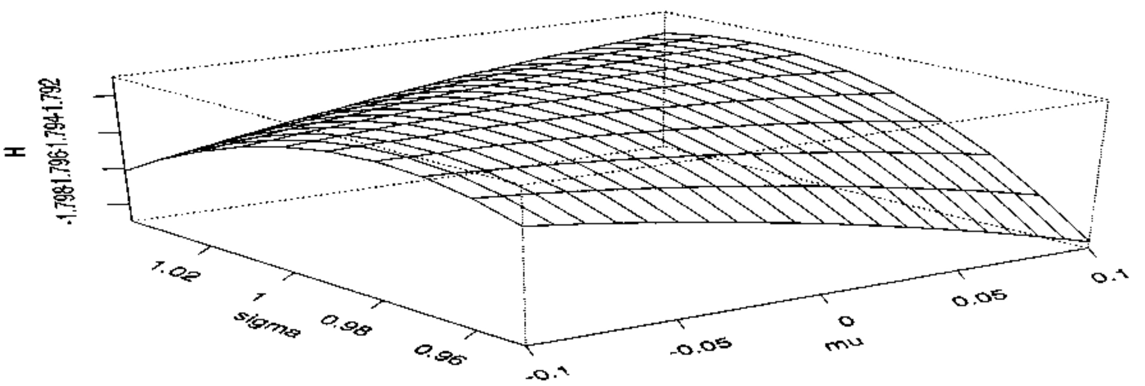
$$\begin{aligned}\sum_{i=1}^{n+1}\left[\frac{(y_i-\mu)\phi(\frac{y_i-\mu}{\sigma})-(y_{i-1}-\mu)\phi(\frac{y_{i-1}-\mu}{\sigma})}{\Phi(\frac{y_i-\mu}{\sigma})-\Phi(\frac{y_{i-1}-\mu}{\sigma})}\right] &= (n+1)\sum_{i=1}^{n+1}\left[\frac{(y_i-\mu)\phi(\frac{y_i-\mu}{\sigma})-(y_{i-1}-\mu)\phi(\frac{y_{i-1}-\mu}{\sigma})}{\Phi(\frac{c-\mu}{\sigma})}\right] \\ &= \frac{(n+1)(c-\mu)\phi(\frac{c-\mu}{\sigma})}{\Phi(\frac{c-\mu}{\sigma})}. \end{aligned} \tag{5.6}$$

Therefore, using equations (5.5) and (5.6), it follows

$$\begin{cases} \frac{\partial H}{\partial \mu}\big|_{\mu=\tilde{\mu},\sigma=\tilde{\sigma}} = -\frac{1}{(n+1)}\left\{\sum_{i=1}^{n+1}\left[\frac{\phi(\frac{y_i-\tilde{\mu}}{\tilde{\sigma}})-\phi(\frac{y_{i-1}-\tilde{\mu}}{\tilde{\sigma}})}{\Phi(\frac{y_i-\tilde{\mu}}{\tilde{\sigma}})-\Phi(\frac{y_{i-1}-\tilde{\mu}}{\tilde{\sigma}})}\right]\right\} + \frac{\phi(\frac{c-\tilde{\mu}}{\tilde{\sigma}})}{\Phi(\frac{c-\tilde{\mu}}{\tilde{\sigma}})} = 0. \\ \frac{\partial H}{\partial \sigma^2}\big|_{\mu=\tilde{\mu},\sigma=\tilde{\sigma}} = -\frac{1}{2(n+1)}\left\{\sum_{i=1}^{n+1}\left[\frac{(y_i-\tilde{\mu})\phi(\frac{y_i-\tilde{\mu}}{\tilde{\sigma}})-(y_{i-1}-\tilde{\mu})\phi(\frac{y_{i-1}-\tilde{\mu}}{\tilde{\sigma}})}{\Phi(\frac{y_i-\tilde{\mu}}{\tilde{\sigma}})-\Phi(\frac{y_{i-1}-\tilde{\mu}}{\tilde{\sigma}})}\right]\right\} + \frac{(c-\tilde{\mu})\phi(\frac{c-\tilde{\mu}}{\tilde{\sigma}})}{2\Phi(\frac{c-\tilde{\mu}}{\tilde{\sigma}})} = 0. \end{cases} \tag{5.7}$$

Consequently, $(\tilde{\mu}, \tilde{\sigma}) = (0, 1)$ maximizes H, and the theorem is proved. To see the obtained results are in agreement with the corresponding graphs, we draw the following graphs in case $c = -1.88$ against μ [-0.1, 0.1] and σ [0.94, 1.02] which is shown in Figures 5.1.

Figure 5.1: H versus μ and σ for ideal sample when $n = 5$ and $c = -1.88$



5.3 Simulation to estimate the mean and variance simultaneously

The purpose of this section is to compare the MPS estimate of a simulation study with the ML estimate. Several attempts were made to calculate $E(\breve{\mu})$, $E(\widetilde{\sigma^2})$, $\sigma(\breve{\mu})$ and $\sigma(\widetilde{\sigma^2})$ for simulation run $R = 10000$, truncation points $c = -1.88, -1, 0, 1$ and sample sizes $n = 5, 10, 20, 50, 100$ through a simulation study.

At first, we used NAG routine C05NBF to solve the equations (5.3). This program failed in a number of cases for some data sets. (see Appendix Program 36). It seems that the routine is very sensitive to the starting value. To tackle this sensitivity problem we wrote a subroutine to find the local maximum approximately using a grid search and having found this we used it as starting value to start the simulation. Again this program failed for some data sets. At this point we embarked upon a new method to maximize equation 4.7, H. The logic of this method is to find the maximum of H, in a certain boundary of μ [-1, 1] and $\log(\sigma)$ [-1, 1]. The proceture is as follow:

1. find the maximum of H, if the two coordinates of the maximum are within the range, then shrink the range, and find the maximum again.
2. If one of the coordinates of the maximum point H is on the boundary, then shift the boundary, and find the maximum.
3. Stop the program, if two successive coordinates of the maximum point H are the same to some accuracy significant figures (see Appendix Programs 37 and 38).

This program also failed in a number of cases.

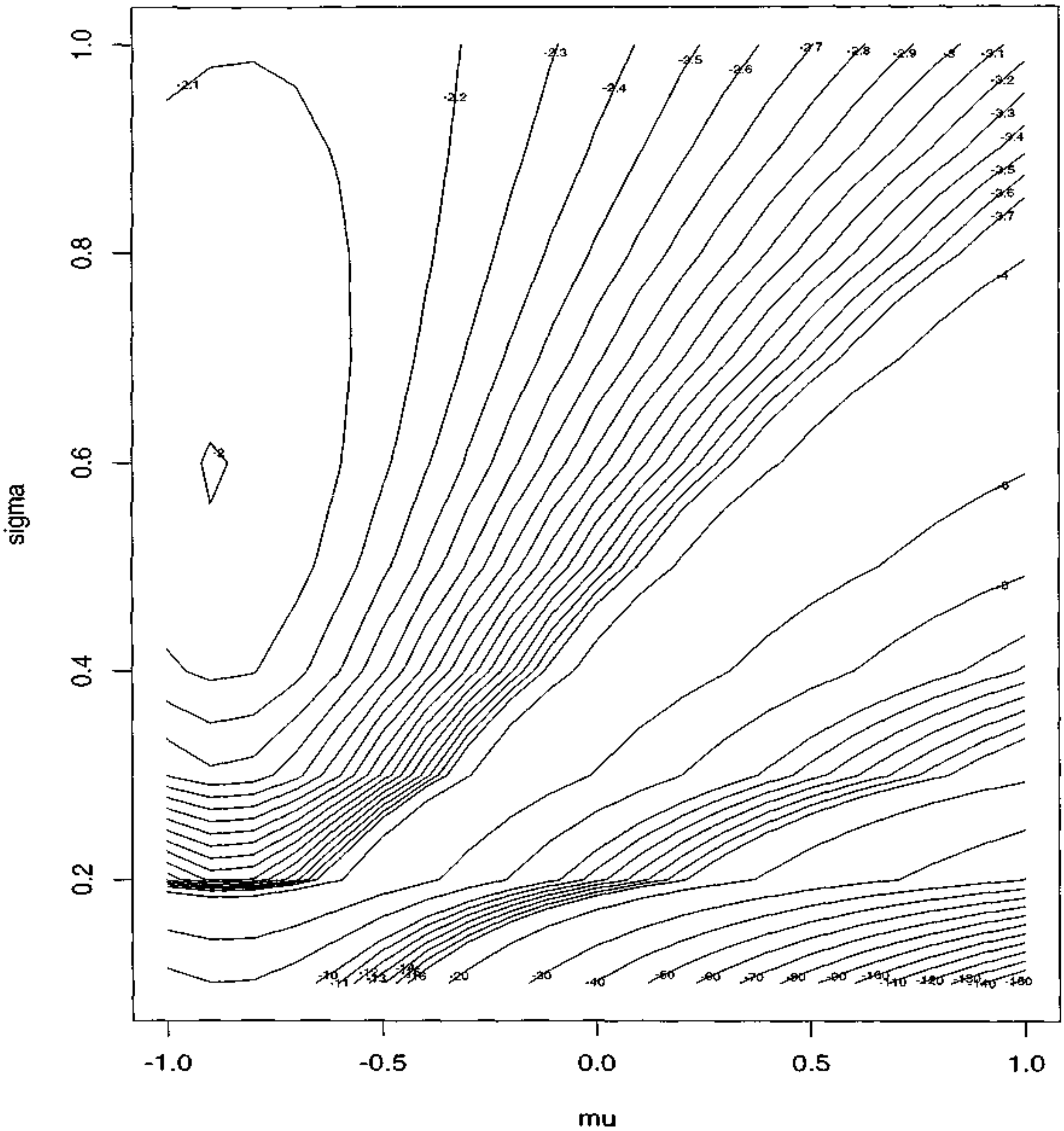
To investigate the problem further, we choose one example of each case for the truncation point $c = 1$.

1. For the random deviates of size $n = 5$, -1.59912, -0.88362, -0.757980, -0.658480 and -0.468060, we obtain $\tilde{\mu} = -0.8828$ and $\tilde{\sigma} = 0.5700$. The value of H at this point is -1.9948.

   Plotting the contour of H against a range of values of μ [-1, 1] and σ [0.1, 1] we get Figure 5.2, which shows the same values for the maximum of H, and its coordinates $\tilde{\mu}$ and $\tilde{\sigma}$. As we can see this Figure fails to show the maximum point of H.

2. However for random deviates of size $n = 5$, -1.89937, -0.55113, 0.039112, 0.33803 and 0.80620, the program fails to give maximize of H.

Figure 5.2: The contour plot of H versus μ and σ when H has maximum



Plotting the contour of H against two ranges of values μ [-1, 2] and [40, 100] and σ [0.1, 2] and [5, 10] we obtain Figure 5.3, which shows the same behaviour of H, $\tilde{\mu}$ and $\tilde{\sigma}$.

Finally, we use NAG routine E04CCF, simplex method to find the maximum of H. The simplex in two dimensions is a triangle which by providing the starting value, we specify the first vertex of the simplex and the remaining vertices are generated by the routine (NAG Manual, 1993)(see Program 39 in Appendix).

Running the Program 39 for the two data sets mentioned above confirmed the results of Programs 37 and 38. Therefore we conclude that we failed to find the maximum point of function H for some data sets. Hence, obtaining $E(\tilde{\mu})$, $E(\tilde{\sigma^2})$, $\sigma(\tilde{\mu})$ and $\sigma(\tilde{\sigma^2})$ for the simulation study are impossible.

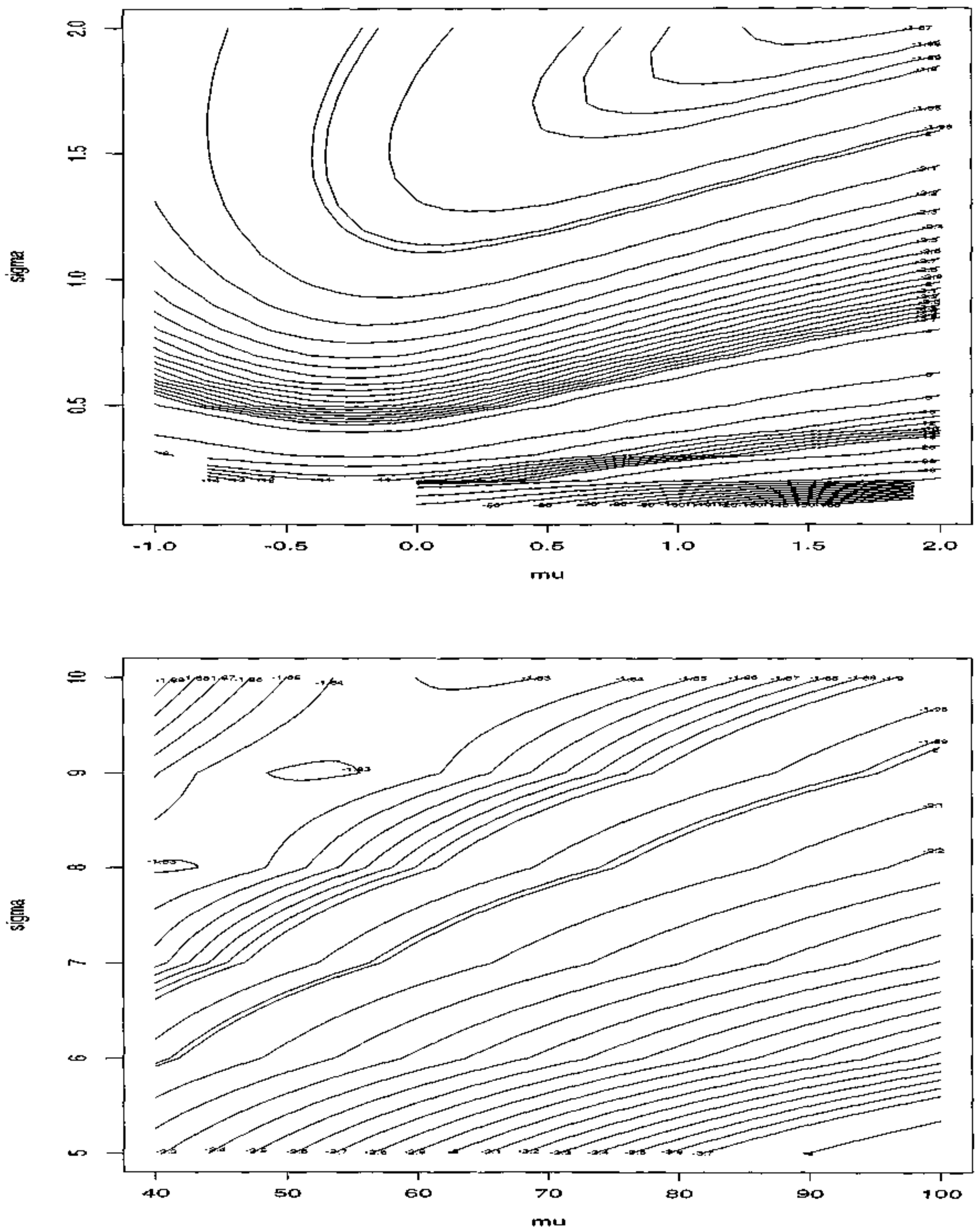
5.4 Conclusion:

In this section we compare the MPS estimator with MLE by three means.

1. For data set 1 the MPS estimator $(\tilde{\mu}, \tilde{\sigma}) = (0.0603, 1.1079)$ where in MLE $(\hat{\mu}, \hat{\sigma}) = (1.3377, 1.2871)$, but for data set 2 the MPS estimator $(\tilde{\mu}, \tilde{\sigma}) = (-0.0158, 1.0396)$ where as in MLE method $(\hat{\mu}, \hat{\sigma}) = (-2.1932, 0.2767)$.
2. For the ideal sample, the MPS estimator $(\tilde{\mu}, \tilde{\sigma}) = (0, 1)$ whereas in MLE for sample size 5 is $(\hat{\mu}, \hat{\sigma}) = (-2.1126, 0.2865)$ and for sample size 10 is $(\hat{\mu}, \hat{\sigma}) = (-1.9509, 0.4099)$.
3. For simulation study; the MPS estimator for some data sets cannot be found , whereas in MLE method we have got the results for all cases.

Concentrating in these examples we see that the MPS estimators are closer to the exact value of μ and σ than MLE. Therefore we conclude that although for the MPS method,

Figure 5.3: The contour plot of H versus μ and σ



simulation failed for some data sets, the MPS method results are more sensible than MLE.

Chapter 6

Modified score functions of the maximum likelihood estimators:

6.1 Introduction and notation:

We know that for the parameter θ the asymptotic bias of the maximum likelihood estimator $\hat{\theta}$ can be written as

$$b(\theta) = \frac{b_1(\theta)}{n} + \frac{b_2(\theta)}{n^2} + \cdots \tag{6.1}$$

where $b_1(\theta)$ and $b_2(\theta)$ are the first and second order terms of $b(\theta)$.

The basis of the present chapter is the idea that the bias in $\hat{\theta}$ can be reduced by introducing a small bias into the score function, Firth (1993).

By employing the notation and methods of McCullagh (1987) for log likelihood derivatives and their null cumulants, the derivatives are denoted by

$$S_r(\theta) = \frac{\partial l}{\partial \theta^r}, \qquad S_{rs}(\theta) = \frac{\partial^2 l}{\partial \theta^r \partial \theta^s}, \tag{6.2}$$

and so on, where $\theta = (\theta^1, \ldots, \theta^p)$ is the parameter vector. The joint null cumulants are

$$\kappa_{r,s} = n^{-1} E\{S_r S_s\}, \quad \kappa_{r,s,t} = n^{-1} E\{S_r S_s S_t\} \quad \text{and} \quad \kappa_{r,st} = n^{-1} E\{S_r S_{st}\}. \tag{6.3}$$

If $\hat{\theta}$ is a solution of score function $S_r(\theta) = 0$, we define a modified score function $S_r^*(\theta)$ such that

$$S_r^*(\theta) = S_r(\theta) + A_r(\theta). \tag{6.4}$$

The solution of $S_r^*(\theta) = 0$ gives a modified estimator θ^* of θ whose its bias is less than the bias of $\hat{\theta}$. According to Firth (1993), in matrix notation, the vector $A(\theta)$ should be such that

$$E\left[A(\theta)\right] = -i(\theta)b_1(\theta)/n + O(n^{-\frac{1}{2}}). \tag{6.5}$$

Therefore, in equation (6.4), $A_r(\theta)$ can be substituted by the $A^{(E)}(\theta) = -i(\theta)b_1(\theta)/n$ or $A^{(O)}(\theta) = -I(\theta)b_1(\theta)/n$, where $I(\theta) = -S_{rs}(\theta)$ and $i(\theta) = E[-S_{rs}(\theta)]$ are called the observed and expected information matrices respectively. The first part of the bias term can be written as

$$b_1^r(\theta) = -\kappa^{r,s}\kappa^{t,u}(\kappa_{s,t,u} + \kappa_{s,tu})/2, \tag{6.6}$$

where $\kappa^{r,s}$ denoted the inverse of the Fisher information matrix $\kappa_{r,s}$.

Finally, the application of these modifications removes the $O(n^{-1})$ bias term. In other words, in equation (6.4) substituting either $A^{(E)}(\theta)$ or $A^{(O)}(\theta)$ for $A_r(\theta)$ removes the $O(n^{-1})$ terms of the equation (6.1).

6.2 The modified score function of μ in the truncated normal distribution, when σ^2 is known:

In this section, using the above explanation, we derived the modified estimator μ^* of μ in the truncated normal distribution.

Using the score function of Chapter 2 and the notations, of this chapter, for $r = s = t = 1$, we have

$$S_1(\mu) = \frac{\partial l}{\partial \mu} = \frac{n\left(\bar{x} - \mu + \psi(c')\right)}{\sigma^2}, \tag{6.7}$$

$$S_{11}(\mu) = \frac{\partial^2 l}{\partial \mu^2} = \frac{-n}{\sigma^2}\left[1 + \psi'(c')/\sigma\right], \tag{6.8}$$

$$\begin{aligned}
\kappa_{1,11}(\mu) &= \frac{1}{n}E[S_1(\mu)S_{11}(\mu)] \qquad (6.9)\\
&= \frac{1}{n}E\{[\frac{n\left(\bar{X} - \mu + \psi(c')\right)}{\sigma^2}][\frac{-n}{\sigma^2}\left(1 + \psi'(c')/\sigma\right)]\}\\
&= \frac{-n[1 + \psi'(c')/\sigma]}{\sigma^4}E[\bar{X} - \mu + \psi(c')]\\
&= 0,
\end{aligned}$$

$$\begin{aligned}
\kappa_{1,1,1}(\mu) &= \frac{1}{n}E[S_1(\mu)]^3\\
&= \frac{1}{n}E\left\{\frac{n}{\sigma^2}[\bar{X} - \mu + \psi(c')]\right\}^3\\
&= \frac{n^3}{n\sigma^6}E(\bar{X} - \mu_c)^3\\
&= \frac{n^2}{\sigma^6}\mu_3(\bar{X})\\
&= \frac{\mu_3(X)}{\sigma^6} \qquad (6.10)
\end{aligned}$$

and

$$\begin{aligned}
\kappa_{11}(\mu) = i(\mu) &= \frac{1}{n}E[S_1^2(\mu)] = \frac{1}{n}E[-S_{11}(\mu)]\\
&= \frac{1}{n}E\left\{\frac{n}{\sigma^2}[1 + \psi'(c')/\sigma]\right\}\\
&= \frac{1}{\sigma^2}[1 + \psi'(c')/\sigma]. \qquad (6.11)
\end{aligned}$$

Therefore we obtain

$$\kappa^{11}(\mu) = \frac{\sigma^2}{[1 + \psi'(c')/\sigma]}. \tag{6.12}$$

Now, using the equation (6.6) and $\mu_3(X) = -\sigma^2\psi''(c')$ from section (0.1.23), we can find the $b_1(\mu)$

$$\begin{aligned} b_1(\mu) &= -\left\{\frac{\sigma^2}{[1+\psi'(c')/\sigma]}\right\}^2 \left[\frac{\mu_3(X)}{\sigma^6}+0\right]/2 \\ &= -\frac{\mu_3(X)}{2\sigma^2[1+\psi'(c')/\sigma]^2} = \frac{\psi''(c')}{2[1+\psi'(c')/\sigma]^2}, \end{aligned} \tag{6.13}$$

which is identical with $b(\mu)$ in Chapter 2. Substituting $i(\mu)$ and $b_1(\mu)$ from equations (6.11) and (6.12) we obtain

$$E[A_1(\mu)] = \frac{\mu_3(X)}{2\sigma^4[1+\psi'(c')/\sigma]} + O(n^{-\frac{1}{2}}). \tag{6.14}$$

Now, the modified score function can be found as

$$\begin{aligned} S_1^*(\mu) &= S_1(\mu) + A_1^{(E)}(\mu) \\ &= \frac{n}{\sigma^2}[\bar{x} - \mu + \psi(c')] + \frac{\mu_3(X)}{2\sigma^4[1+\psi'(c')/\sigma]}, \end{aligned} \tag{6.15}$$

or

$$S_1^*(\mu) = \frac{n}{\sigma^2}[\bar{x} - \mu + \psi(c')] - \frac{\psi''(c')}{2\sigma^2[1+\psi'(c')/\sigma]}. \tag{6.16}$$

Now, we expect the solution of the equation (6.16), μ^* of μ, to have a smaller bias than $\hat{\mu}$.

This would be an interesting problem to study in the future.

6.3 The modified score function of the σ^2 in the truncated normal distribution, when μ is known:

In this section again using the $\sigma^2 = \gamma$, we find the modified estimate γ^* of γ.

Using the score function and information matrix of Chapter 2 (equation 2.62) with the notation defined in this chapter for $\mu = 0$, we have

$$S_1(\gamma) = \frac{\partial l}{\partial \gamma} = \frac{n}{2\gamma^2}\left[\frac{\sum_1^n x_i^2}{n} - \gamma + c'\gamma\tau(c')\right]. \tag{6.17}$$

Since we want to find $i(\gamma)$, and also we know that $E[S_1(\gamma)] = 0$, we obtain

$$E\left[\frac{\sum_1^n X_i^2}{n}\right] = \gamma - c'\gamma\tau(c'). \tag{6.18}$$

Since we know that $i(\gamma) = E[I(\gamma)] = E[-S_{11}(\gamma)]$, we have to obtain

$$S_{11}(\gamma) = \frac{\partial^2 l}{\partial\gamma^2} = \frac{n}{2\gamma^2}\left[\frac{1}{2} - \frac{\sum_1^n x_i^2}{n\gamma} - \frac{c'^2\tau'(c')}{4\gamma^2} + \frac{3c'\tau(c')}{4\gamma^2}\right]. \tag{6.19}$$

Using equation (6.17), into $E[-S_{11}(\gamma)]$ we obtain

$$i(\gamma) = \frac{n}{2\gamma^2}\left[1 - \frac{c'}{2}\tau(c') + \frac{c'^2}{2}\tau'(c')\right] = \frac{nD}{2\gamma^2}. \tag{6.20}$$

Now, let us find $\kappa_{1,11}(\gamma)$ and $\kappa_{1,1,1}(\gamma)$. From equation (6.3) we have

$$\begin{aligned}
\kappa_{1,11}(\gamma) &= \frac{1}{n}E[S_1(\gamma)S_{11}(\gamma)] \\
&= \frac{1}{n}E\left\{\left[\frac{n}{2\gamma^2}\left(\frac{\sum_1^n X_i^2}{n} - \gamma + c'\gamma\tau(c')\right)\right]\left[\frac{-n}{\gamma^3}\left(\frac{\sum_1^n X_i^2}{n} + \frac{c'^2\gamma\tau(c') + c'\gamma\tau(c') - 2\gamma}{4}\right)\right]\right\} \\
&= -\frac{n^2}{2n\gamma^5}\mathrm{Var}\left(\frac{\sum_1^n X_i^2}{n}\right) \\
&= -\frac{D}{\gamma^3}
\end{aligned} \tag{6.21}$$

and

$$\begin{aligned}
\kappa_{1,1,1}(\gamma) &= \frac{1}{n}E[S_1(\gamma)]^3 \\
&= \frac{1}{n}E\left[\frac{n}{2\gamma^2}\left(\frac{\sum_1^n X_i^2}{n} - \gamma + c'\gamma\tau(c')\right)\right]^3 \\
&= \frac{n^2}{8\gamma^6}\mu_3\left(\frac{\sum_1^n X_i^2}{n}\right).
\end{aligned} \tag{6.22}$$

Since finding the third moment of $\left(\frac{\sum_1^n X_i^2}{n}\right)$ is cumbersome, we use the formula (2.20) in Chapter 2. Therefore we have to obtain $\kappa_{111}(\gamma)$

$$\begin{aligned}
\kappa_{111}(\gamma) &= \frac{1}{n}E[S_{111}(\gamma)] \\
&= \frac{1}{n}E\left[n\left(-\frac{1}{\gamma^3} + \frac{3}{\gamma^4}\left(\frac{\sum_1^n X_i^2}{n}\right) + \frac{c'^3\tau''(c')}{8\gamma^3} + \frac{9c'^2\tau'(c')}{8\gamma^3} + \frac{15c'\tau(c')}{8\gamma^3}\right)\right]
\end{aligned} \tag{6.23}$$

Using the equation (6.23) we obtain

$$\kappa_{111}(\gamma) = \frac{1}{\gamma^3}\left\{2 + \frac{c'}{8}\left[-9\tau(c') + c'^2\tau''(c') + 9c'\tau'(c')\right]\right\}. \tag{6.24}$$

Now, we have

$$\begin{aligned}\kappa_{1,1,1}(\gamma) &= -3\kappa_{1,11}(\gamma) - \kappa_{111}(\gamma) \\ &= \frac{3}{\gamma^3}\left[1 - \frac{c'\tau(c')}{2} + \frac{c'^2\tau'(c')}{2}\right] - \\ &\quad \frac{1}{\gamma^3}\left\{2 + \frac{c'}{8}\left[-9\tau(c') + c'^2\tau''(c') + 9c'\tau'(c')\right]\right\}.\end{aligned} \tag{6.25}$$

To find the $\kappa^{1,1}$ we have

$$\kappa_{1,1} = \frac{D}{2\gamma^2}. \tag{6.26}$$

Therefore

$$\kappa^{1,1} = \frac{2\gamma^2}{D}. \tag{6.27}$$

Substituting $\kappa_{1,11}$, $\kappa_{1,1,1}$ and $\kappa^{1,1}$ from equations (6.21), (6.25) and (6.27) into equation (6.6) we obtain $b_1(\gamma)$

$$\begin{aligned}b_1(\gamma) &= \frac{\gamma c'\left[c'^2\tau''(c') + c'\tau'(c') - \tau(c')\right]}{4\left[1 - \frac{c'\tau(c')}{2} + \frac{c'^2\tau'(c')}{2}\right]^2} \\ &= \frac{\gamma M}{4D^2}\end{aligned} \tag{6.28}$$

which is identical with the $b(\gamma)$ in Chapter 2.

Using the $i(\gamma)$ and $b_1(\gamma)$ from equations (6.20) and (6.28) into equation (6.5) we obtain

$$\begin{aligned}E[A_1(\gamma)] &= -i(\gamma)\frac{b(\gamma)}{n} + O(n^{-\frac{1}{2}}) \\ &= \frac{c'[\tau(c') - c'\tau'(c') - c'^2\tau''(c')]}{8\gamma[1 - \frac{c'\tau(c')}{2} + \frac{c'^2\tau''(c')}{2}]} + O(n^{-\frac{1}{2}}).\end{aligned} \tag{6.29}$$

Now the modified score function in terms of expected information is

$$\begin{aligned} S_1^*(\gamma) &= S_1(\gamma) + A_1^{(E)}(\gamma) \\ &= \frac{n}{2\gamma^2}\left[\frac{\sum_1^n x_i^2}{n} - \gamma + c'\gamma\tau(c')\right] + \frac{c'[\tau(c') - c'\tau'(c') - c'^2\tau''(c')]}{8\gamma[1 - \frac{c'\tau(c')}{2} + \frac{c'^2\tau''(c')}{2}]}. \end{aligned} \tag{6.30}$$

and in terms of observed information is

$$\begin{aligned} S_1^*(\gamma) &= S_1(\gamma) + A_1^{(O)}(\gamma) \\ &= \frac{n}{2\gamma^2}\left[\frac{\sum_1^n x_i^2}{n} - \gamma + c'\gamma\tau(c')\right] + \\ &\quad \left[\frac{\sum_1^n x_i^2}{n\gamma} + \frac{3c'\tau(c') + c'^2\tau'(c') - 2\gamma}{4}\right]\frac{c'[\tau(c') - c'\tau'(c') - c'^2\tau''(c')]}{4\gamma[1 - \frac{c'\tau(c')}{2} + \frac{c'^2\tau''(c')}{2}]^2}. \end{aligned} \tag{6.31}$$

We expect the solution of the equation (6.30) or (6.31), γ^* of γ to have a smaller bias than $\hat{\gamma}$. This also could be an interesting problem to study in the future.

Chapter 7

Summary and recommendations for further research:

7.1 Summary of the work:

The work in this thesis concerns the truncated normal distribution. Specifically we consider the singly truncated normal from the right and its parameter estimation. In Chapter 2, we use the maximum likelihood method to estimate one parameter when the other is known. For both actual data and simulated data, we work out the estimates of the parameters. Two theoretical methods, those of Cox & Hinkley and Shenton & Bowman, are investigated and found the $E(\hat{\mu})$, $\text{Var}(\hat{\mu})$, $E(\hat{\sigma}^2)$ and $\text{Var}(\hat{\sigma}^2)$ to give identical results. Our simulated estimates are compared with the theoretical methods and the results are almost the same.

In Chapter 3, we estimate two parameters of the distribution simultaneously. Moreover, we extend the Shenton & Bowman formula for the two parameters. We make a comparison between theory and simulation and finds that they give identical results.

We see that the maximum product spacing method for the one parameter case in Chapter 4, is asymptotically as efficient as the maximum likelihood method. We consider a model of the variance in terms of the truncation point and the sample size. The distribution and moments of the estimators $\hat{\mu}$ and $\tilde{\mu}$ have been determined.

In Chapter 5, by the method of maximum product spacing, we consider the simultaneous estimation the two parameters of the truncated normal distribution and we make a comparison with the maximum likelihood method.

7.2 Recommendations for further research:

There are various possible areas of further research that evaluate for this work. We list them as follows:

1. In Chapter 2, we have shown that the ML estimate of μ exists. It would be worthwhile to prove theoretically that it is unique.
2. In Chapter 4, we investigate the distribution of $\tilde{\mu}$ when σ is known. Further research could be done to investigate the distribution of $\tilde{\sigma^2}$ when μ is known.
3. In Chapter 5, in the simulation to estimate the mean and variance simultaneously the computer program fails for some samples, in other words the routine C05NBF fails to converge. Further research, especially in modifying the NAG routine C05NBF could be done to produce a better routine for solving the nonlinear equations. It would be useful if the recommended routine incorporate of the choice of a suitable starting value for the iterations that would lead to convergence.
4. In Chapter 6, we found the modified score functions of μ, $S_1^*(\mu)$ and γ, $S_1^*(\gamma)$. Further research could be done to find the properties of μ^* and γ^*.

Appendix

The diskette contained all the programs is available in the following address.

(a) M.Tazhibi, Faculty of Health, Isfahan University of Medical Sciences, Hezar Gerib Street, Isfahan, Iran.

(b) Mr B. J. R. Bailey, Faculty of Mathematical Studies, University of Southampton, Highfield, SO17 1BJ, U.K.

Bibliography

Abramowitz, M & Stegun, A. (1965), *Handbook of mathematical functions* , New York: Dover publications, Inc.

Adams, R. A. (1995) *A complete course calculus*, New york: Addison-Wesley .

Bishop, M., Fienberg, S. E. and Holland, P. W.(1975).*Discrete multivariate analysis : theory and practice*, Cambridge, Mass.: M. I. T. Press

Cheng, R. C. H and Amin, N. A. K (1983) Estimating parameters in continuous univariate distributions with a shifted origin. *The J. Roy. Statist. Soc. B* **45** *3, 394-403.*

Cohen, A. C (1950 a), Estimating the mean and variance of normal population from singly and doubly truncated samples. Annals of Mathematical Statistics, **21**, *557-569.*

Cohen, A. C (1950 b), On estimating the mean and variance of singly truncated normal frequency distributions from the first three sample moments, Annals of the Institute of Statistical Mathematics, **1-5**, *37-44.*

Cohen, A. C (1950), On the solution of estimating equations for truncated and censored samples from normal population Biometrika **44**, *225-236.*

Cohen, A. C (1957), On the solution of estimating equations for truncated and censored samples from normal population, Biometrika **44**,*225-236.*

Cohen, A. C (1986) Truncated and censored samples from normal population, New York: Marcel Dekker.

Conte, S. D. (1965), Elementary numerical analysis, New York: Mc Graw Hill.

Cox, D. R. and Hinkley, D. V. (1974), Theoretical statistics, London: Chapman and Hall.

Dahiya, R. C, and Hegde, L. M. (1989) Estimation of the parameters in a truncated normal distribution, Commun. Statist. Theory Math., **18** *, 4177-4195.*

Deemer, W. L. and Votaw, D.F. (1955), Estimation of parameter of truncated or censored exponential distribution. Ann. Math. Statist. **26***, 498-504*

Ekholm, A & Green, M (1995) Fitting nonlinear models in GLIM 4 using numerical derivatives GLIM Newsletter **23***, 12-20.*

*Firth, D. (1993) Bias reduction of maximum likelihood estimates, Biometrika,***80***, 27-38.*

Fisher, R. A (1936), Statistical methods for research workers 6th ed. Oliver and Boyd.

*Fisher, R. A (1931), Properties and applications of the functions in mathematical tables,***1***, XXV-XXXV, British Association for the Advancement of Science.*

Galton, F. (1897), An examination into registered speeds of American trotting horses, with remarks on their value as hereditary data, Proc. Royal Soc. of London **62***, 310-314*

Gallagher, J (1993), Simulating from continuous probability distributions, MSc dissertation in Statistics with Application in Medicine, University of Southampton.

Hogg, R. V. and Craig, A. T. (1970) Introduction to mathematical statistics New York: Macmillan.

Jacques, I and Judd, C (1957) Numerical analysis London: Chapman and Hall.

Kendall, M. G. and Stuart, A (1952) The advanced theory of statistics, Vol. I London: Griffin.

Lee, L. F. (1983), The determination of moments of the doubly truncated multivariate normal Tobit Model, Economic Letters **11** *245-250.*

Larson, H. J (1982), Introduction to probability theory and statistical inference, New York: John Wiley.

Lehman, E. L and D'abrera, H. J. M. (1975) Nonparametric statistical methods based on ranks San Francisco:, Holden-Day.

Lindgren, B. W (1962), Statistical theory, New York: Mac Millan.

Mann,s H.B and Wald, A. (1943). On stochastic limit and order relationships. Ann. Math. Statist., **14**, *217-226.*

McCullagh, P. (1987), Tensor methods in statistics, London: Chapman and Hall.

Morgan, B. J. T. (1984) Elements of simulation, Cambridge: CUP.

Norton, H.W. (1956). One likelihood adjustment may be inadequate. Biometrika, **12**, *79.*

Pearson, K. (1902), On the systematic fitting of frequency curves. Biometrika, **2**, *2-7.*

Powell, M.J.D. Rabinowitz, P.(ed) (1970) A hybrid method for nonlinear algebraic equations. In Numerical methods for nonlinear algebraic equations, New york: Gordon and Breach.

Raj, D. (1953) On estimation of the parameters of bivariate normal population from doubly and singly linearly truncated samples, Sankhya, **12**, *277-290.*

Ranneby, B. (1984) The maximum spacing method: an estimating method related to the maximum likelihood method. Scand. J. Statist., **11**, *93-112.*

Schneider H. (1989), Truncated and censored samples from normal populations, New York: Marcel Dekker.

Shenton, L. R. and Bowman, K. O. (1977) Maximum likelihood estimation in small samples, London: Griffin

Snedecor, G. W. and Cochran, W. G. (1967) Statistical methods Iowa: The University Press

StatSci (1993), Statistical Science, S-PLUS Programer's Manual Version 3.2 , Seattle: a division of MathSoft, Inc.

Tanner, J. M , Whitehouse, R. H and Takaishi, M (1966) Standards from Birth to Maturity for Height, Weight, Height Velocity, and Weight Velocity: British Children, 1965, Part I Arch. Dis. Childh., **41**, *454-471.*

The Numerical Algorithms Group Limited (1993), The NAG Fortran Library Manual, Mark 16, Oxford: NAG Ltd.

Thompson, H. R. (1951) Truncated lognormal distribution I. Solution by moments, Biometrika, **38**, *414-422.*

Votaw, D. F, Rafferty, J. A, and Deemer, W. L. (1950) Estimation of parameters in truncated trivariate normal distribution, Psychometrika, **15**, *339-347.*

World Health Organization (1986) The Growth Chart A tool for use in infant and child health care Geneva: