2001

Reliability studies of the skew normal distribution

Nicole Dawn Brown

RELIABILITY STUDIES OF THE SKEW NORMAL DISTRIBUTION

By

Nicole Dawn Brown

A.B. Bowdoin College, **1997**

A THESIS

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Arts

(in Mathematics)

The Graduate School

The University of Maine

May, 2001

Advisory Committee:

Ramesh Gupta, Professor of Mathematics, Advisor

David Bradley, Assistant Professor of Mathematics

Pushpa Gupta, Professor of Mathematics

RELIABILITY STUDIES OF THE SKEW NORMAL DISTRIBUTION

By Nicole Dawn Brown

Thesis Advisor: Dr. Ramesh Gupta

An Abstract of the Thesis Presented

in Partial Fulfillment of the Requirements for the

Degree of Master of Arts

(in Mathematics)

May, 2001

It has been observed in various practical applications that data do not conform to the normal distribution, which is symmetric with no skewness. The skew normal distribution proposed by Azzalini (1985) is appropriate for the analysis of data which is unimodal but exhibits some skewness. The skew normal distribution includes the normal distribution as a special case where the skewness parameter is zero.

In this thesis we study the structural properties of the skew normal distribution, with an emphasis on the reliability properties of the model. More specifically, we obtain the failure rate, the mean residual life function, and the reliability function of a skew normal random variable. We also compare it with the normal distribution with respect to certain

stochastic orderings. Appropriate machinery is developed to obtain the reliability of a component when the strength and stress follow the skew normal distribution. Finally, IQ score data from Roberts (1988) is analyzed to illustrate the procedure.

ACKNOWLEDGEMENTS

I would like to express my gratitude to Professor Ramesh Gupta for his guidance and encouragement throughout this project. I would also like to thank my other committee members, Professor Pushpa Gupts and Professor David Bradley, for their participation. Finally, I would like to thank my family for their support in all my endeavors.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS ii

LIST OF TABLES iv

Chapter

1. INTRODUCTION 1

2. THE UNIVARIATE SKEW NORMAL DISTRIBUTION 5

   2.1. The Model 5

   2.2. Representations of the Skew Normal Distribution 11

   2.3. Properties of the Skew Normal Distribution 14

   2.4. Moments of the Skew Normal 15

3. RELIABILITY FUNCTIONS OF THE SKEW NORMAL DISTRIBUTION 17

   3.1. Reliability Properties of the Skew Normal 19

   3.2. Comparison with the Normal Distribution 22

4. APPLICATION TO STRENGTH-STRESS MODEL 25

5. ANALYSIS OF THE ROBERTS DATA 32

   5.1. Maximum Likelihood Estimates 35

   5.2. Estimates for IQ Data 37

6. CONCLUSIONS AND RECOMMENDATIONS 40

BIBLIOGRAPHY 42

APPENDIX. Fisher Information Matrix 44

BIOGRAPHY OF THE AUTHOR 49

LIST OF TABLES

Table 5.1. Otis IQ scores for non-whites 38
Table 5.2. Otis IQ scores for whites 38
Table 5.3. Parameter estimates for Otis IQ scores for whites 39
Table 5.4. Parameter estimates for Otis IQ scores for non-whites 39

Chapter 1

INTRODUCTION

The celebrated Gaussian (Normal) distribution has been known for centuries. Its popularity has been driven by its analytical simplicity and the associated Central Limit Theorem. The multivariate extension is straightforward because the marginals and conditionals are both normal, a property rarely found in most of the other multivariate distributions. Yet there have been doubts, reservations, and criticisms about the unqualified use of normality. There are numerous situations when the assumption of normality is not validated by the data. In fact Geary (1947) remarked, "Normality is a myth; there never was and never will be a normal distribution." As an alternative, many near normal distributions have been proposed. Some families of such near normal distributions, which include the normal distribution and to some extent share its desirable properties, have played a crucial role in data analysis. For description of some such families of distributions, see Mudholkar and Hutson (2000). See also Azzalini (1985), Turner (1960) and Prentice (1975). Many of the near normal distributions described above deal with effects of asymmetry. These families of asymmetrical distributions are analytically tractable, accommodate practical values of skewness and kurtosis, and strictly include the normal distribution. These distributions can be quite useful for data modeling and statistical analysis.

In this thesis we are concerned with a skew normal distribution, proposed by Azzalini(1985), whose probability density function is given by

$$\phi(z;\lambda) = 2\phi(z)\,\Phi(\lambda z) \tag{1.1}$$

where $\phi(z)$ and $\Phi(z)$ denote the standard normal density and distribution function, respectively. The parameter λ varies in $(-\infty, \infty)$ and regulates the skewness and $\lambda = 0$ corresponds to the standard normal case. The density given by (1.1) enjoys a number of formal properties which resemble those of the normal distribution, for example if Z has the pdf of (1.1), then Z^2 has a chi-square distribution with one degree of freedom. From a practical point of view, the density (1.1) is suitable for the analysis of data exhibiting a unimodal empirical distribution but with some skewness present, a structure often occurring in data analysis. Arnold et al. (1993) provided the following motivation for the skew normal model. Suppose students admitted to a college are screened with respect to their SAT scores and their progress is monitored with respect to their grade point average (GPA). Let (X,Y) denote their (GPA, SAT). Assuming that (X,Y) follows a bivariate normal distribution and assuming that only those students whose SAT scores are above average are admitted to the college, the distribution of X follows a non-standard skew normal distribution and its standardized version is given by (1.1).

A multivariate version of (1.1) has been recently studied by Azzalini and Dalla Valle (1996) and Azzalini and Capitanio (1999). This distribution represents a mathematically tractable extension of the multivariate normal density with the addition of a parameter to regulate skewness. These authors demonstrate that the multivariate skew normal distribution has a reasonable flexibility in real data fitting, while it maintains some convenient formal properties of the normal density.

The purpose of this present work is to study, in detail, the model given by (1.1) and investigate some of its properties useful in reliability. We also study the maximum likelihood estimation of the parameters and present an application to the strength-stress model useful in reliability. The strength-stress model consists in estimating $R = P(Y < X)$, which has been studied extensively in the literature. The problem originated in the context of the reliability of a component of strength X subjected to a stress Y. The component fails if at any time the applied stress is greater than its strength and there is no failure when $X > Y$. Thus $P(Y<X)$ is a measure of the reliability of the component. More specifically, in Chapter **2** we present the basic properties of the model including several representations, the moment generating function, and moments. Chapter **3** deals with the failure rate and other reliability functions of the aforementioned model. We also compare it with the normal distribution with respect to certain stochastic orderings and prove that

the failure rate of a skew normal distribution is increasing, a property enjoyed by the symmetric normal distribution. In Chapter 4 appropriate machinery is developed to obtain an expression for the $P(Y < X)$, where X and Y each have a skew normal distribution. In Chapter 5, the data of Roberts (1988) dealing with Otis IQ scores is analyzed to illustrate the procedure. Finally, we give some conclusions and recommendations justifying the skew normal distribution. We also point out several directions for future research.

Chapter 2

THE UNIVARIATE SKEW NORMAL DISTRIBUTION

2.1 The Model

In this chapter we shall define the univariate skew normal distribution, first proposed by Azzalini (1985). The distribution will be defined by its density function and three representations in terms of the normal distribution. We shall also discuss several properties and the moments of the distribution.

First, we introduce two lemmas which can be used to prove that the skew normal is a proper density and to derive the moment generating function.

Lemma 2.1 *Let Y be a standard normal random variable and let h and k be real numbers. Then*

$$E\left\{\Phi\left(hY+k\right)\right\}=\Phi\left\{\frac{k}{\sqrt{1+h^2}}\right\} \qquad \text{for all } h \text{ and } k, \tag{2.1}$$

where $\Phi(\cdot)$ is the standard normal distribution function.

Proof. Let X and Y be independent random variables where $X \sim N\,(p, \sigma^2)$ and $Y \sim N\,(0,1)$. Let $Z = X - Y$. Then Z is a normal random variable with mean μ and variance $\sigma^2 + 1$. We have

$$
\begin{aligned}
P\,(Z < 0) &= P\,(X < Y) = E_Y\left[P\,(X < Y | Y = y)\right] \\
& \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma}\left(\int_{-\infty}^{y} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} dx\right) \phi\,(y)\,dy \\
&= \int_{-\infty}^{\infty} \left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{y-\mu}{\sigma}} e^{-\frac{t^2}{2}} dt\right) \phi\,(y)\,dy \\
&= \int_{--M}^{\infty} \Phi\left(\frac{y-\mu}{\sigma}\right) \phi\,(y)\,dy \\
&= E_Y\left\{\Phi\left(\frac{Y-\mu}{\sigma}\right)\right\} \\
&= E\left\{\Phi\,(hY + k)\right\} \qquad \text{where } h = \frac{1}{\sigma} \text{ and } k = -\frac{\mu}{\sigma}.
\end{aligned}
$$

We now derive an alternative expression for $P\,(Z < 0)$ directly from the distribution of Z. We have

$$
\begin{aligned}
P\,(Z < 0) &= \int_{-\infty}^{0} \frac{1}{\sqrt{2\pi}\sqrt{\sigma^2+1}} \exp\left\{-\frac{(z-\mu)^2}{2\,(\sigma^2+1)}\right\} dz \\
&= \frac{1}{\sqrt{2\pi}} \int_{--M}^{\left(\frac{-\mu}{\sqrt{\sigma^2+1}}\right)} e^{-\frac{t^2}{2}} dt \\
&= \Phi\left\{\frac{-\mu}{\sqrt{\sigma^2+1}}\right\} \\
&= \Phi\left\{\frac{k}{\sqrt{1+h^2}}\right\}, \qquad \text{where } h = \frac{1}{\sigma} \text{ and } k = \frac{-\mu}{\sigma}.
\end{aligned}
$$

Equating the two expressions for $P\,(Z < 0)$, we have the desired result. ■

We now present an **alternative proof** of Lemma 2.1.

Proof. Let Y be a standard normal random variable. For any real h and k, define a function $\Psi(h,k)$ as follows:

$$\Psi(h,k) = \int_{-\infty}^{\infty} \Phi(hy+k)\,\phi(y)\,dy. \tag{2.2}$$

Then $\Psi(h,k) = E\{\Phi(hY+k)\}$. We now differentiate (2.2) with respect to k:

$$\begin{aligned}
\frac{\partial\Psi(h,k)}{\partial k} &= \int_{--m}^{\infty} \phi(hy+k)\,\phi(y)\,dy \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2}\left[(hy+k)^2+y^2\right]\right\}dy \\
&= \frac{1}{2\pi}e^{-\frac{k^2}{2(1+h^2)}}\int_{-\infty}^{\infty} \exp\left\{-\frac{(1+h^2)}{2}\left(y+\frac{hk}{1+h^2}\right)^2\right\}dy.
\end{aligned}$$

Letting $u = \sqrt{1+h^2}\left(y+\frac{hk}{1+h^2}\right)$, we have

$$\begin{aligned}
\frac{\partial\Psi(h,k)}{\partial k} &= \frac{1}{\sqrt{2\pi}\sqrt{1+h^2}}e^{-\frac{k^2}{2(1+h^2)}}\int_{-\infty}^{\infty} e^{-\frac{u^2}{2}}du \\
&= \frac{1}{\sqrt{1+h^2}}\phi\left(\frac{k}{\sqrt{1+h^2}}\right).
\end{aligned}$$

Now, integrating with respect to k, we have

$$\Psi(h,k) = \Phi\left\{\frac{k}{\sqrt{1+h^2}}\right\},$$ which proves the lemma.

■

Lemma 2.2 *Let f be a density function which is symmetric about 0 and let* G *be a distribution function which is absolutely continuous and whose derivative is symmetric about 0. Then*

$$2G(\lambda y)\, f(y)\,, \qquad (-\infty < y < \infty) \tag{2.3}$$

is a proper density function for any $\lambda \in$ **R.**

Proof. Let X and Y be independent random variables where X has density function G' and Y has density function f. Since X and Y are both symmetric about 0, then $X - XY$ must also be symmetric about 0. So we have $P(X - XY < 0) = \frac{1}{2}$. Conditioning on Y, we also have

$$\begin{aligned} P(X - XY < 0) &= E_Y\left[P\left(X < \lambda Y \mid Y = y\right)\right] \\ &= E_Y\left[G(\lambda Y) \mid Y = y\right] \\ &= \int_{-\infty}^{\infty} G(\lambda y)\, f(y)\, dy \quad . \end{aligned}$$

It follows that $\int_{-cQ}^{\infty} 2G(\lambda y)\, f(y)\, dy = 1.$ ∎

We now define the skew normal probability density function (pdf) and prove that it is a proper density.

Definition **2.1** *Let $\lambda \in$ **R.** A random variable Z is distributed skew normal with parameter λ if Z has the density function*

$$\phi(z;\lambda) = 2\,\phi(z)\;\Phi(\lambda z)\,, \qquad (-\infty < z < \infty)\,, \tag{2.4}$$

where $\phi(\cdot)$ is the standard normal density function.

If Z has the skew normal density, we write $Z \sim SN(\lambda)$. The fact that $\phi(z;\lambda)$ is a proper **pdf** can be verified by applying Lemma 2.1 or Lemma 2.2 as shown in the following theorems.

Theorem **2.1** *The skew normal density $\phi(z;\lambda)$ is a proper density for each $\lambda \in$ **R.***

Proof. Let Y be a standard normal random variable and fix $\lambda \in$ **R.** Apply Lemma 2.1 with $k = 0$. Then we have

$$\int_{--m}^{\infty} 2a)\,(Xy)\phi(y)dy = 2\,\mathrm{E}\,\{a)\,(\lambda Y)\} = 1.$$

■

We now present an alternative proof of Theorem 2.2.

Proof. Let X and Y be independent standard normal random variables and fix $\lambda \in \mathbf{R}$. Define $Z = X - XY$. Then Z is a normal random variable with expected value 0. Then we have

$$
\begin{aligned}
\frac{1}{2} &= P(Z < 0) = P(X - XY < 0) = P(X < XU) \\
&= E_Y\left[P(X < \lambda Y \mid Y = y)\right] \\
&= \int_{-\infty}^{\infty} \Phi(\lambda y)\,\phi(y)\,dy.
\end{aligned}
$$

It follows from lemma **2.2** that

$$\int_{-\infty}^{\infty} 2\,\Phi(\lambda y)\,\phi(y)\,dy = 1$$

and therefore $\phi(z;\lambda)$ is a proper density. ■

2.2 Representations of the Skew Normal Distribution

In this section we present some useful representations of the skew normal distribution in terms of normal random variables.

Theorem 2.2 *Let U and V be independent standard normal random variables and let*

$$Z = \frac{\lambda}{\sqrt{1+\lambda^2}}\,|U| + \frac{1}{\sqrt{1+\lambda^2}}\,V. \tag{2.5}$$

Then $Z \sim SN(\lambda)$

Proof. Let $a = \frac{\lambda}{\sqrt{1+\lambda^2}}$ and let $b = \frac{1}{\sqrt{1+\lambda^2}}$. Then

$$\begin{aligned} P(Z \le z) &= E_{|U|}\left[P(Z \le z \mid |U| = u)\right] \\ &= 2\int_0^\infty P\left(V \le \frac{-au}{b}\right)\phi(u)du \\ &= 2\int_0^\infty \Phi\left(\frac{z-au}{b}\right)\phi(u)\,du. \end{aligned}$$

Differentiation yields the density of Z as follows:

$$\frac{d}{dz}P(Z \le z) = 2\int_0^\infty \frac{1}{b}\phi\left(\frac{z-au}{b}\right)\phi(u)\,du.$$

Using the fact that $a^2 + b^2 = 1$, we obtain

$$\begin{aligned} \frac{d}{dz}P(Z \le z) &= \phi(z)\int_0^\infty \frac{1}{\sqrt{2\pi}b}\exp\left\{-\frac{(u-az)^2}{2b^2}\right\}du \\ &= 2\phi(z)\int_{-az/b}^\infty \frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{t^2}{2}\right\}dt \\ &= 2\phi(z)\left\{1-\Phi(-az/b)\right\} \end{aligned}$$

$= 2\phi(z)\,\Phi(\lambda z)$, the skew normal density.

■

Theorem 2.3 *Let X and Y be independent standard normal random variables and $\lambda \in \mathbf{R}$. The distribution of Y conditionally on $X < \lambda Y$ is $SN(\lambda)$.*

Proof. If X and Y are independent standard normal random variables and $\lambda \in \mathbf{R}$, then

$$\begin{aligned} P(Y \leq t \mid X < \lambda Y) &= \frac{P(Y \leq t, X < \lambda Y)}{P(X - \lambda Y < 0)} \\ &= \frac{\int_{-\infty}^{t} \Phi(\lambda y)\,\phi(y)\,dy}{\int_{-\infty}^{0} \frac{1}{\sqrt{2\pi}\sqrt{1+\lambda^2}} e^{-\frac{u^2}{2(1+\lambda^2)}}\,du} \\ &= \frac{\int_{-\infty}^{t} \Phi(\lambda y)\,\phi(y)\,dy}{\Phi(0)} \\ &= 2\int_{-\infty}^{t} \Phi(\lambda y)\,\phi(y)\,dy. \end{aligned}$$

Differentiating with respect to t, we have the skew normal pdf:

$$\frac{d}{dt} P(Y \leq t \mid X < \lambda Y) = 2\Phi(\lambda t)\,\phi(t).$$

■

Theorem 2.4 *Let Y and W be independent standard normal random variables and $\lambda \in \mathbf{R}$. Define $X = (\lambda Y - W)/(1+\lambda^2)^{1/2}$. Then*

(i)(X,Y) has a standard bivariate normal distribution with correlation coefficient $\frac{\lambda}{\sqrt{1+\lambda^2}}$, and

(ii)the distribution of Y conditionally on $X > 0$ is $SN(\lambda)$.

Proof. Let Y and W be independent standard normal random variables and fix $\lambda \in \mathbf{R}$. Define the random variable $X = (\lambda Y - W)/(1+\lambda^2)^{1/2}$. Then X has a standard normal distribution and

$$\begin{aligned} corr(X,Y) &= E(XY) \\ &= E\left[\frac{(\lambda Y - W)Y}{(1+\lambda^2)^{1/2}}\right] \\ &= \frac{\lambda}{(1+\lambda^2)^{1/2}}. \end{aligned}$$

Thus (X,Y) has a standard bivariate normal distribution. Then we have

$$\begin{aligned} P[Y \le y | X > 0] &= P\left[Y \le y \left| \frac{(\lambda Y - W)}{(1+\lambda^2)^{1/2}} > 0\right.\right] \\ &= P[Y \le y | \lambda Y > W] \\ &= \Phi(y;\lambda) \qquad \text{by Theorem 2.4.} \end{aligned}$$

■

There is one further representation of the skew normal distribution which will be discussed in a later section.

2.3 Properties of the Skew Normal Distribution

The following are some useful properties of the skew normal density.

Property I. The standard normal distribution is a special case of the skew-normal distribution when $\lambda = 0$.

Property II. **As** $\lambda \rightarrow \infty$, $\phi(z;\ \mathsf{A})$ tends to the half normal density.

Property III. If $Z \sim SN(\mathsf{A})$, then $-Z \sim SN(-\lambda)$.

Property IV. $\Phi(z;\ -\mathsf{A}) = 1 - \Phi(-z;\ \mathsf{A})$, where $\Phi(z;\ \mathsf{A})$ is the distribution function of the skew normal.

Property V. $\Phi(z;\ 1) = \{\Phi(z)\}^2$

Property VI. If $Z \sim SN(\mathsf{A})$, then Z^2 is a chi-square random variable with one degree of freedom. It is known that the square of a standard normal random variable is a chi-square random variable with one degree of freedom. This property of the skew normal implies that the converse is not true. **A** chi-square random variable is not necessarily the square of a standard normal.

2.4 Moments of the Skew Normal

In this section we derive the moment generating function and the moments of a skew normal random variable.

Theorem 2.5 *Let $Z \sim SN(\lambda)$. The moment generating function* of Z *is*

$$M_Z(t) = 2e^{\frac{t^2}{2}}\Phi\left(\frac{\lambda t}{\sqrt{1+\lambda^2}}\right).$$

Proof.

$$\begin{aligned} M_Z(t) &= E\left[e^{tz}\right] = 2\int_{-\infty}^{\infty} e^{tz}\phi(z)\,\Phi(\lambda z)\,dz \\ &= 2e^{\frac{t^2}{2}}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(z-t)^2}\Phi(\lambda t)\,dt \\ &= 2e^{\frac{t^2}{2}}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{u^2}{2}}\Phi(\lambda(u+t))\,du \\ &= 2e^{\frac{t^2}{2}}E\{\Phi(\lambda(u+t))\} \qquad \text{where } U \sim N(0,1). \end{aligned}$$

Applying Lemma 2.1, $M_Z(t) = 2e^{\frac{t^2}{2}}\Phi\left(\frac{\lambda t}{\sqrt{1+\lambda^2}}\right)$ ■

The first moment of a skew normal is given by

$$\begin{aligned} E(Z) &= M'(0) = 2\left[e^{\frac{t^2}{2}}\frac{\lambda}{\sqrt{1+\lambda^2}}\phi\left(\frac{\lambda t}{\sqrt{1+\lambda^2}}\right) + te^{\frac{t^2}{2}}\Phi\left(\frac{\lambda t}{\sqrt{1+\lambda^2}}\right)\right]\Bigg|_{t=0} \\ &= 2\frac{\lambda}{\sqrt{1+\lambda^2}}\phi(0) \\ &= \sqrt{\frac{2}{\pi}}\left(\frac{\lambda}{\sqrt{1+\lambda^2}}\right). \end{aligned}$$

The second moment of a skew normal is given by

$$
\begin{aligned}
E\left(Z^2\right) &= M"(0) \\
&= 2\left[e^{\frac{t^2}{2}}\left(\frac{\lambda}{\sqrt{1+\lambda^2}}\right)^2 \phi'\left(\frac{\lambda t}{\sqrt{1+\lambda^2}}\right)\right. \\
&\quad +2te^{\frac{t^2}{2}}\left(\frac{\lambda}{\sqrt{1+\lambda^2}}\right)\phi\left(\frac{\lambda t}{\sqrt{1+\lambda^2}}\right) \\
&\quad \left.\left.+\left(t^2+1\right)e^{\frac{t^2}{2}}\Phi\left(\frac{\lambda}{\sqrt{1+\lambda^2}}\right)\right]\right|_{t=0} \\
&= 2\Phi(0) \\
&= 1.
\end{aligned}
$$

It follows that

$$
E(Z) = \sqrt{\frac{2}{\pi}}\left(\frac{\lambda}{\sqrt{1+\lambda^2}}\right)
$$

and

$$
Var(Z) = 1-\frac{2\lambda^2}{\pi\left(1+\lambda^2\right)}.
$$

Chapter **3**

RELIABILITY FUNCTIONS OF THE SKEW NORMAL DISTRIBUTION

In this chapter we discuss the reliability properties of the skew normal distribution and compare it with the normal distribution with respect to some stochastic orderings. Before proceeding further, we present the following definitions:

Let X be a random variable having absolutely continuous distribution function F and pdf f. Then

1. The survival function of X is defined as $\bar{F}(t) = P(X > t) = 1 - F(t)$.

2. The failure rate (hazard rate) of X is defined as

$$r_F(t) = \lim_{\Delta t \to 0} \frac{P(t \leq X \leq t + \Delta t \mid X > t)}{\Delta t} = \frac{f(t)}{\bar{F}(t)}.$$

3. The mean residual life function or life expectancy is defined as

$$\mu_F(t) = E(X - t \mid X > t) = \int_t^\infty \frac{(x-t) f(x)}{\bar{F}(t)} dx.$$

It is well known that $\bar{F}(t)$, $r_F(t)$, and $\mu_F(t)$ are equivalent in the sense that given one of them, the other two can be determined. They also characterize the distribution uniquely; see Gupta (1981).

We now define the following criteria used in reliability:

1. F is said to be Polya frequency of order 2 (PF_2) if $\ln f(x)$ is concave.

2. F is said to have increasing (decreasing) failure rate, IFR (DFR), if $r_F(t)$ is increasing (decreasing).

3. F is said to have decreasing (increasing) mean residual life, DMRL (IMRL), if $\mu_F(t)$ is decreasing (increasing), assuming that the mean exists.

It is well known that $PF_2 \Longrightarrow IFR \Longrightarrow DMRL$. The reverse implications are not necessarily true.

3.1 Reliability Properties of the Skew Normal

In order to derive the failure rate of a skew normal random variable Z, we must first define its distribution function, which is given by

$$\begin{aligned}\Phi(z;\lambda) &= 2\int_{-\infty}^{z}\int_{-\infty}^{\lambda t}\phi(u)\,\phi(t)\,du\,dt \\ &= \Phi(z) - 2\int_{z}^{\infty}\int_{0}^{\lambda t}\phi(u)\,\phi(t)\,du\,dt \\ &= \Phi(z) - 2T(z;\lambda), \qquad \lambda > 0 \end{aligned} \tag{3.1}$$

where

$$T(z;\mathrm{A}) = \int_{z}^{\infty}\int_{0}^{\lambda t}\phi(u)\,\phi(t)\,du\,dt;$$

see Azzalini (1985) for details. The function $T(z;\mathrm{A})$ is an integral over **a** polygonal region. Its derivation in closed form is not feasible. Owen (1956) gives tables of values of $T(z;\mathrm{A})$. Computer routines which evaluate $T(z;\mathbf{A})$ are also available. The following is an expression for $T(z;\mathbf{A})$ in terms of an infinite series:

$$T(z;\lambda) = \frac{\arctan\lambda}{2\pi} - \frac{1}{2\pi}\sum_{j=0}^{\infty} c_j\lambda^{2j+1}, \tag{3.2}$$

where

$$c_j = (-1)^j\frac{1}{2j+1}\left[1 - \exp\left(-\frac{1}{2}z^2\right)\sum_{i=0}^{j}\frac{z^{2i}}{2^i i!}\right], \tag{3.3}$$

see Owen (1956). It is known that $T(z;\mathrm{A})$ is a decreasing function of h and

1. $-T(z;\lambda) = T(z;-\lambda)$,

2. $T(-z;\lambda) = T(z;\lambda)$, and

3. $2T(z;1) = \Phi(z)\Phi(-z)$, see equations (2.4) and (2.5) of Owen (1956).

Therefore, the reliability function and the failure rate of a skew normal random variable Z are given by

$$R(t) = P(Z > t) = 1 - \Phi(t) + 2T(t;\lambda) \tag{3.4}$$

and

$$r(t) = \frac{\phi(t;\lambda)}{1-\Phi(t;\lambda)} = \frac{2\phi(t)\,\Phi(\lambda t)}{1-\Phi(t) + 2T(t;\lambda)}. \tag{3.5}$$

The expressions for $\lambda < 0$ can be similarly obtained.

We now derive the mean residual life function (MRLF) of Z, which is given by

$$\mu(t) = E(Z - t \mid Z > t) = E(Z \mid Z > t) - t.$$

Now

$$\begin{aligned}
E(Z \mid Z > t) &= \frac{2}{R(t)} \int_t^\infty x\phi(x)\,\Phi(\lambda x)\,dx \\
&= \frac{-2}{R(t)} \int_t^\infty \phi'(x)\,\Phi(\lambda x)\,dx \\
&= \frac{2}{R(t)} \left[\Phi(\lambda t)\,\phi(t) + \lambda \int_t^\infty \phi(x)\,\phi(\lambda x)\,dx\right] \\
&= \frac{2}{R(t)} \left\{\Phi(\lambda t)\,\phi(t) + \frac{\lambda}{\sqrt{2\pi}\sqrt{1+\lambda^2}}\left[1 - \Phi\left(t\sqrt{1+\lambda^2}\right)\right]\right\}.
\end{aligned}$$

Hence

$$\mu(t) = \frac{2\Phi(\lambda t)\phi(t) + \frac{2\lambda}{\sqrt{2\pi}\sqrt{1+\lambda^2}}\left[1 - \Phi\left(t\sqrt{1+\lambda^2}\right)\right]}{R(t)} - t. \qquad (3.6)$$

Due to the complicated nature of the expressions for the reliability function and the failure rate, the usual derivative methods are cumbersome if we are interested in studying the monotonicity of the failure rate **or** the MRLF. Accordingly, we take an alternative approach. In the following, we shall examine the monotonicity of the failure rate and the mean residual life function. First, we prove the following result.

Theorem 3.1 *The skew normal density function* is *log concave.*

Proof. To prove that $\log \phi(z;\lambda)$ is a concave function of z, it suffices to show that the second derivative of $\log \phi(z;\lambda)$ is negative for all z. Differentiating log $\phi(z;\lambda)$ we have

$$\begin{aligned}
\frac{d^2}{dz^2}\log \phi(z;\lambda) &= \frac{d^2}{dz^2}\left[\log 2 + \log\phi(z) + \log\Phi(\lambda z)\right] \\
&= \frac{d}{dz}\left[\frac{\frac{d}{dz}\phi(z)}{\phi(z)}\right] + \frac{d}{dz}\left[\frac{\frac{d}{dz}\Phi(\lambda z)}{\Phi(\lambda z)}\right] \\
&= \frac{d}{dz}\left[\frac{-z\phi(z)}{\phi(z)}\right] + \frac{d}{dz}\left[\frac{\lambda\phi(\lambda z)}{\Phi(\lambda z)}\right] \\
&= -1 + \lambda\left[\frac{-z\lambda^2\phi(\lambda z)\,\Phi\lambda z}{\{\Phi(\lambda z)\}^2} - \frac{\lambda\phi^2(\lambda z)}{\{\Phi(\lambda z)\}}\right] \\
&= -1 + \frac{-\lambda^2\phi(\lambda z)}{\Phi(\lambda z)}\left[\frac{\phi(\lambda z)}{\Phi(\lambda z)} + \lambda z\right].
\end{aligned}$$

We will show that the above quantity is negative. Since $\phi(z)$ and $\Phi(\lambda z)$ are positive for all z, it is sufficient to show that $\frac{\phi(\lambda z)}{\Phi(\lambda z)} + \lambda z$ is positive for all λz.

Case I: If $\lambda z \geq 0$, then $\frac{\phi(\lambda z)}{\Phi(\lambda z)} + \lambda z$ is clearly positive.

Case 11: If $\lambda z < 0$, let $t = -\lambda z$. Then $\phi(\lambda z) = \phi(-\lambda z) = \phi(t)$ and $\Phi(\lambda z) = 1 - \Phi(-\lambda z) = 1 - \Phi(t)$. Thus $\frac{\phi(\lambda z)}{\Phi(\lambda z)} + \lambda z = \frac{\phi(t)}{1-\Phi(t)} - t = h(t) - t$, where $h(t)$ is the failure rate of the standard normal distribution. Since it is known (see Azzalini (1986)) that $h(t) > t$ for all t, the assertion is proved.

■

Corollary 3.1 *The skew normal random variable Z has increasing failure rate (IFR) for all values* of λ *and hence decreasing mean residual life (DMRL).*

3.2 Comparison with the Normal Distribution

We shall now compare the skew normal distribution with the normal distribution with respect to some stochastic relations. First we present the definitions of some stochastic order relations.

Let X and Y be two absolutely continuous random variables with probability density functions f and g and survival functions $\bar{F}$ and G. Then

1. X is said to be larger than Y in likelihood ratio ordering, written as $X \underset{LR}{\geq} Y$, if $f(x)/g(x)$ is nondecreasing as x increases.

2. X is said to be larger than Y in failure rate ordering, written as $X \underset{FR}{\geq} Y$, if $r_F(x) \leq r_G(x)$ for all x.

3. X is said to be larger than Y in stochastic ordering, written as $X \underset{ST}{\geq} Y$, if $\bar{F}(x) \geq \bar{G}(x)$ for all x.

4. X is said to be larger than Y in mean residual life ordering, written as $X \underset{MRL}{\geq} Y$, if $\mu_F(x) \geq \mu_G(x)$ for all x.

It is well known that $X \underset{LR}{\geq} Y \Longrightarrow X \underset{FR}{\geq} Y \Longrightarrow X \underset{ST}{\geq} Y$ and $X \underset{FR}{\geq} Y \Longrightarrow X \underset{MRL}{\geq} Y$; see Gupta and Kirmani (1998).

Now suppose X is a random variable having a skew normal distribution function F given by (3.1) and pdf given by $f(x) = 2\phi(x)\Phi(\lambda x)$. Also, suppose Y is a random variable with a standard normal distribution function $G(x) = \Phi(x)$ and pdf $g(x) = \phi(x)$. Then

$$\begin{aligned} \frac{f(x)}{g(x)} &= \frac{2\phi(x)\;\Phi(\lambda x)}{\phi(x)} \\ &= 2\Phi(\lambda x), \end{aligned}$$

which is increasing if $\lambda > 0$ and decreasing if $\lambda < 0$. Thus if $\lambda > 0$, it follows that

$X \underset{LR}{\geq} Y$. This implies that:

1. r_F (z) $\leq r_G$ (z) for all x,

2. $\bar{F}(x) \geq$ G (z) for all z, and

3. $\mu_F(x) \geq \mu_G$ (z) for all x.

Similarly, if $\lambda < 0$, it follows that $X \underset{LR}{\leq} Y$ and hence:

1. r_F **(z)**$\geq r_G$ *(x)* for all x,

2. $\bar{F}$ (z) $\leq$ G (z) for all x, and

3. $\mu_F(x) \leq \mu_G(x)$ for all x.

Chapter 4

APPLICATION TO STRENGTH-STRESS MODEL

In this section we are interested in estimating the $P(Z_1 < Z_2)$ when Z_1 and Z_2 are independent skew normal random variables with parameters λ_1 and x_2, respectively. Before proceeding further we obtain the distribution function of $a_1Z_1 + a_2Z_2$, where a_1 and a_2 are constants.

Theorem 4.1 *Let Z_1 and Z_2 be independent where $Z_1 \sim SN(\lambda_1)$ and $Z_2 \sim SN(\lambda_2)$. For any real numbers a_1 and a_2, the distribution function of $a_1Z_1 + a_2Z_2$ is given by*

$$P(a_1Z_1 + a_2Z_2 < u) = 2\int_0^\infty \Phi\left(\frac{u - (a_1\delta_1 + a_2\delta_2)\,t}{\sigma}\right)\phi(t)\,dt, \qquad (4.1)$$

where $\delta_i = \frac{\lambda_i}{\sqrt{1+\lambda_i^2}}$, $i = 1,2$ and $\sigma^2 = a_1^2\left(1-\delta_1^2\right) + a_2^2\left(1-\delta_2^2\right)$.

Proof. Using Theorem 2.2, with $\delta_1 = \frac{\lambda_1}{\sqrt{1+\lambda_1^2}}$ and $\delta_2 = \frac{\lambda_2}{\sqrt{1+\lambda_2^2}}$, Z_1 and Z_2 can be written as

$$Z_1 = \delta_1 |Y_0| + \left(1 - \delta_1^2\right)^{\frac{1}{2}} Y_1$$

and

$$Z_2 = \delta_2 |Y_0| + \left(1 - \delta_2^2\right)^{\frac{1}{2}} Y_2,$$

where Y_0, Y_1, and Y_2 are independent standard normal random variables. Thus

$$\begin{aligned} P(a_1 Z_1 + a_2 Z_2 < u) &= P\Big(a_1 \Big(\delta_1 |Y_0| + \left(1 - \delta_1^2\right)^{\frac{1}{2}} Y_1\Big) \\ &\quad + a_2 \Big(\delta_2 |Y_0| + \left(1 - \delta_2^2\right)^{\frac{1}{2}} Y_2\Big) < u) \\ &= P\Big(a_1 \left(1 - \delta_1^2\right)^{\frac{1}{2}} Y_1 \\ &\quad + a_2 \left(1 - \delta_2^2\right)^{\frac{1}{2}} Y_2 < -(a_1\delta_1 + a_2\delta_2) |Y_0| + u\Big) \\ &= \int_0^\infty P(A_1 Y_1 + A_2 Y_2 < u - (a_1\delta_1 + a_2\delta_2)\, t)\, 2 4 (t) dt, \end{aligned}$$

where $A_1 = a_1 \left(1 - \delta_1^2\right)^{\frac{1}{2}}$ and $A_2 = a_2 \left(1 - \delta_2^2\right)^{\frac{1}{2}}$. Let V $= A_1 Y_1 + A_2 Y_2$. Then V has a normal distribution with E (V) $= 0$ and Var (V) $= A_1^2 + A_2^2 = \sigma^2$. Therefore,

$$\begin{aligned} P(a_1 Z_1 + a_2 Z_2 < u) &= \int_0^\infty P\left(\frac{V - 0}{\sigma} < \frac{u - (\text{a161} + a_2\delta_2)\, t}{\sigma}\right) 2\phi(t)\, dt \\ &= 2 \int_0^\infty \Phi\left(\frac{u - (a_1\delta_1 + a_2\delta_2)\, t}{\sigma}\right) \phi(t)\, dt. \end{aligned}$$

■

Remark 4.2 *In general, if for* $i = 1, 2, ..n$*, the* Z_i *are independent, and each* $Z_i \sim SN(\lambda_i)$*, then*

$$P\left(\sum_{i=1}^{n} a_i Z_i < u\right) = 2\int_{0}^{\infty} \Phi\left(\frac{u - t\sum_{i=1}^{n} a_i\delta_i}{\sigma}\right)\phi(t)\,dt \tag{4.2}$$

where $\delta_i = \frac{\lambda_i}{\sqrt{1+\lambda_i^2}}$ *and* $\sigma^2 = \sum_{i=1}^{n} a_i^2 \left(1 - \delta_i^2\right)$.

The next theorem deals with the pdf of $a_1 Z_1 + a_2 Z_2$.

Theorem 4.3 *Let* Z_1 *and* Z_2 *be independent where* $Z_1 \sim SN(\lambda_1)$ *and* $Z_2 \sim SN(\lambda_2)$. *For any real numbers* a_1 *and* a_2,

$$\frac{a_1 Z_1 + a_2 Z_2}{\sqrt{a^2 + \sigma^2}} \sim SN\left(\frac{a}{\sigma}\right), \tag{4.3}$$

where $\delta_i = \frac{\lambda_i}{\sqrt{1+\lambda_i^2}}, i = 1, 2;\ a = a_1\delta_1 + a_2\delta_2$ *and* $\sigma^2 = a_1^2\left(1 - \delta_1^2\right) + a_2^2\left(1 - \delta_2^2\right)$.

Proof. From **(4.1)** the pdf of $a_1 Z_1 + a_2 Z_2$ is given by

$$
\begin{aligned}
\frac{d}{du}P\left(a_1Z_1+a_2Z_2<u\right) &= \frac{d}{du}2\int_0^{\infty}\Phi\left(\frac{u-\left(a_1\delta_1+a_2\delta_2\right)t}{\sigma}\right)\phi\left(t\right)dt \\
&= \frac{2}{\sigma}\int_0^{\infty}\phi\left(\frac{u-\left(a_1\delta_1+a_2\delta_2\right)t}{\sigma}\right)\phi\left(t\right)dt \\
&= \frac{1}{\pi\sigma}\int_0^{\infty}\exp\left\{-\frac{1}{2}\left[t^2+\frac{1}{\sigma^2}\left(u-at\right)^2\right]\right\}dt \\
&= \frac{e^{-\frac{1}{2}\frac{u^2}{\left(a^2+\sigma^2\right)}}}{\pi\sigma}\int_0^{\infty}\exp\left\{-\frac{1}{2}\frac{\left(t-\frac{ua}{a^2+\sigma^2}\right)^2}{\left(\sqrt{\frac{\sigma^2}{a^2+\sigma^2}}\right)^2}\right\}dt \\
&= \frac{e^{-\frac{1}{2}\frac{u^2}{\left(a^2+\sigma^2\right)}}}{\pi\sqrt{a^2+\sigma^2}}\int_c^{\infty}e^{-\frac{1}{2}s^2}ds \quad \text{where } c=\frac{-ua}{\sigma\sqrt{a^2+\sigma^2}} \\
&= \frac{\sqrt{2\pi}e^{-\frac{1}{2}\frac{u^2}{\left(a^2+\sigma^2\right)}}}{\pi\sqrt{a^2+\sigma^2}}\Phi\left(\frac{ua}{\sigma\sqrt{a^2+\sigma^2}}\right) \\
&= \frac{2}{\sqrt{2\pi}}e^{-\frac{1}{2}w^2}\Phi\left(\frac{a}{\sigma}w\right) \quad \text{where w}=\frac{U}{\sqrt{a^2+\sigma^2}} \\
&= 2\phi\left(\text{w}\right)\Phi\left(\frac{a}{\sigma}w\right).
\end{aligned}
$$

Thus

$$\frac{a_1Z_1+a_2Z_2}{\sqrt{a^2+\sigma^2}}\sim SN\left(\frac{a}{\sigma}\right).$$

■

Remark 4.4 *In general if* Z_i *are independent* SN (Xi) , $i=1,2,..n$, *then*

$$\frac{\sum_{i=1}^{n}a_iZ_i}{\sqrt{a^2+\sigma^2}}\sim SN\left(\frac{a}{\sigma}\right) \tag{4.4}$$

where $\delta_i=\frac{\lambda_i}{\sqrt{1+\lambda_i^2}}$, $a=\sum_{i=1}^{n}a_i\delta_i$ *and* $\sigma^2=\sum_{i=1}^{n}a_i^2\left(1-\delta_i^2\right)$.

We now apply the skew normal to a strength-stress model. If Z_1 is the stress on a component and Z_2 is the strength of the component, we use the previous theorem to evaluate $P(Z_1 < Z_2)$, where $Z_1 \sim SN(\lambda_1)$ and $Z_2 \sim SN(\lambda_2)$. First we must prove the following lemma.

Lemma 4.1 *For any real* h,

$$\int_0^\infty \Phi(hu)\phi(u)\,du = \frac{1}{2\pi}\arctan h + \frac{1}{4}. \tag{4.5}$$

Proof. Consider a more general integral, given by

$$\Psi(h,k) = \int_0^\infty \Phi(hu+k)\,\phi(u)\,du. \tag{4.6}$$

Taking the derivative of (4.6) with respect to h, we get

$$\begin{aligned}
\frac{\partial}{\partial h}\Psi(h,k) &= \int_0^\infty u\phi(hu+k)\,\phi(u)\,du \\
&= \frac{1}{2\pi}\int_0^\infty u\exp\left\{-\frac{1}{2}\left[(hu+k)^2+u^2\right]\right\}du \\
&= \frac{1}{2\pi}e^{-\frac{k^2}{2(1+h^2)}}\int_0^\infty u\exp\left\{-\frac{(1+h^2)}{2}\left(u+\frac{hk}{1+h^2}\right)^2\right\}du \\
&= \frac{1}{2\pi}\frac{e^{-\frac{k^2}{2(1+h^2)}}}{\sqrt{1+h^2}}\left[\frac{1}{\sqrt{1+h^2}}\int_{\frac{hk}{\sqrt{1+h^2}}}^\infty te^{-\frac{t^2}{2}}dt\right. \\
&\qquad\left.-\frac{\sqrt{2\pi}hk}{1+h^2}\left\{1-\Phi\left(\frac{hk}{\sqrt{1+h^2}}\right)\right\}\right] \\
&= \frac{1}{2\pi}\frac{e^{-\frac{k^2}{2(1+h^2)}}}{\sqrt{1+h^2}}\left[\frac{e^{-\frac{h^2k^2}{2(1+h^2)}}}{\sqrt{1+h^2}}-\frac{\sqrt{2\pi}(hk)}{1+h^2}\left\{1-\Phi\left(\frac{hk}{\sqrt{1+h^2}}\right)\right\}\right] \\
&= \frac{e^{-\frac{k^2}{2}}}{2\pi(1+h^2)}-\frac{(hk)\,e^{-\frac{k^2}{2(1+h^2)}}}{\sqrt{2\pi}(1+h^2)^{\frac{3}{2}}}\left[1-\Phi\left(\frac{hk}{\sqrt{1+h^2}}\right)\right]
\end{aligned} \tag{4.7}$$

Strictly speaking, (4.7) has to be integrated back with respect to h. However, the value of this cannot be obtained in a closed form. It can be obtained in an infinite series form as it is the same form of integral involved in the distribution function of Z, see Theorem 1 of Henze (1986).

For our purposes, we can consider the case $k = 0$ and get

$$\frac{\partial}{\partial h}\Psi(h,0) = \frac{1}{2\pi(1+h^2)}, \tag{4.8}$$

where

$$\Psi(h,0) = \frac{1}{2\pi}\arctan h + B, \text{ for some real constant } B.$$

Since $\Psi(0,0) = \frac{1}{4}$, it follows that $B = \frac{1}{4}$. Hence

$$\int_0^\infty \Phi(hu)\phi(u)du = \frac{1}{2\pi}\arctan h + \frac{1}{4}.$$

■

We are now in a position to evaluate $P(Z_1 < Z_2)$.

Theorem 4.5 *Let Z_1 and Z_2 be independent where $Z_1 \sim SN(\lambda_1)$ and $Z_2 \sim SN(\lambda_2)$. Then*

$$P(Z_1 < Z_2) = \frac{1}{\pi}\arctan\left(\frac{\delta_2 - \delta_1}{\sqrt{2-\delta_1^2-\delta_2^2}}\right) + \frac{1}{2}, \tag{4.9}$$

where $\delta_i = \frac{\lambda_i}{\sqrt{1+\lambda_i^2}}, i = 1, 2.$

Proof. We apply Lemma **4.1** in the case where $a_1 = 1, a_2 = -1$, and $u = 0$. We get $\sigma^2 = (1-\delta_1^2) + (1-\delta_2^2) = 2 - \delta_1^2 - \delta_2^2$ and hence using **(4.1)** we have

$$\begin{aligned} P(Z_1 < Z_2) &= 2\int_0^{\infty} \Phi\left(\frac{(\delta_2-\delta_1)\,t}{\sqrt{2-\delta_1^2-\delta_2^2}}\right)\phi(t)\,dt \\ &= 2\left[\frac{1}{2\pi}\arctan\left(\frac{\delta_2-\delta_1}{\sqrt{2-\delta_1^2-\delta_2^2}}\right)+\frac{1}{4}\right] \\ &= \frac{1}{\pi}\arctan\left(\frac{\delta_2-\delta_1}{\sqrt{2-\delta_1^2-\delta_2^2}}\right)+\frac{1}{2}. \end{aligned}$$

■

Chapter 5

ANALYSIS OF THE ROBERTS DATA

Before proceeding further, we present the following motivation, due to Arnold et al. (1993), for the skew normal distribution involving location and scale parameters. Let (X, Y) have a bivariate normal density with mean vector (μ_1, μ_2), variance vector (σ_1^2, σ_2^2) and correlation ρ. Let $f(x, y)$ denote the joint density of X and Y. If Y is truncated below at its mean, μ_2, then the joint density of X and Y is given by

$$f_{X,Y}(x,y) = \begin{cases} 2f(x,y), & -\infty < x < \infty, \quad y > \mu_2 \\ 0, & \text{elsewhere,} \end{cases} \tag{5.1}$$

which is a truncated bivariate normal distribution. We shall now derive a non-standard skew normal random variable X.

Theorem 5.1 *Let X and Y have the joint density* ***(5.1).*** *The marginal distribution of the untruncated variable X is given* by

$$f_X(x) = \frac{2}{\sigma_1}\phi\left(\frac{x-\mu_1}{\sigma_1}\right)\Phi(\lambda), \tag{5.2}$$

where $\lambda = \frac{\rho}{\sqrt{1-\rho^2}}$. *The random variable* X *is said to have a non-standard skew normal distribution.*

Proof. Integrating the joint density of X and Y with respect to Y, we have

$$\begin{aligned} f_X(x) &= 2\int_{\mu_2}^{\infty} f(x,y)\,dy \\ &= 2\int_{\mu_2}^{\infty} \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{y-\mu_2}{\sigma_2}\right)\left(\frac{x-\mu_1}{\sigma_1}\right) + \left(\frac{y-\mu_2}{\sigma_2}\right)^2\right]} dy \\ &= \frac{2e^{-\frac{1}{2}\left(\frac{x-\mu_1}{\sigma_1}\right)^2}}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \int_{\mu_2}^{\infty} e^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{y-\mu_2}{\sigma_2}\right) - \rho\left(\frac{x-\mu_1}{\sigma_1}\right)\right]^2} dy \\ &= \frac{2\phi\left(\frac{x-\mu_1}{\sigma_1}\right)}{\sqrt{2\pi}\sigma_1\sigma_2\sqrt{1-\rho^2}} \int_{\mu_2}^{\infty} e^{-\frac{1}{2(1-\rho^2)\sigma_2^2}\left[y-\mu_2-\rho\sigma_2\left(\frac{x-\mu_1}{\sigma_1}\right)\right]^2} dy. \end{aligned}$$

Letting V $= y - \mu_2 - pa2\left(\frac{x-\mu_1}{\sigma_1}\right) \Big/ \sigma_2\sqrt{1-\rho^2}$, we have

$$\begin{aligned} f_X(x) &= \frac{2\phi\left(\frac{x-\mu_1}{\sigma_1}\right)}{\sigma_1} \int_{c}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{v^2}{2}} dv, \quad \text{where } c = \frac{-\rho\left(\frac{x-\mu_1}{\sigma_1}\right)}{\sqrt{1-\rho^2}}, \\ &= \frac{2}{\sigma_1}\phi\left(\frac{x-\mu_1}{\sigma_1}\right)\Phi\left(\frac{\rho\left(\frac{x-\mu_1}{\sigma_1}\right)}{\sqrt{1-\rho^2}}\right). \end{aligned}$$

■

If X has the pdf given by (5.2), it is called a non-standard skew normal with parameter $\lambda = \frac{\rho}{\sqrt{1-\rho^2}}$.

The moments of the non-standard skew normal are given by

$$\begin{aligned} E(X) &= \mu_1 + \rho\sqrt{\frac{2}{\pi}}\sigma_1 \\ Var(X) &= \left[1-\rho^2\left(\frac{2}{\pi}\right)\right]\sigma_1^2 \end{aligned}$$

and

$$Skewness(X) = \left(\frac{4}{\pi}-1\right)\left(\frac{2}{\pi}\right)^{\frac{1}{2}}\rho^3\left(1-\frac{2\rho^2}{\pi}\right)^{-\frac{3}{2}}.$$

Let $X_1, X_2, ...X_n$ be a random sample of size n from the pdf (5.2). Then the moment estimators are given by

$$\begin{aligned} \bar{\mu} &= \bar{x} - \sqrt{\frac{2}{\pi}}\left(\frac{m_3}{\left(\frac{4}{\pi}-1\right)\sqrt{\frac{2}{\pi}}}\right)^{\frac{1}{3}} \\ \bar{\sigma}^2 &= s^2 + \left(\frac{2}{\pi}\right)\left(\frac{m_3}{\left(\frac{4}{\pi}-1\right)\sqrt{\frac{2}{\pi}}}\right)^{\frac{2}{3}} \\ \bar{\rho} &= \left(\frac{2}{\pi} + s^2\left(\frac{\left(\frac{4}{\pi}-1\right)\sqrt{\frac{2}{\pi}}}{m_3}\right)^{\frac{2}{3}}\right)^{-\frac{1}{2}} \\ \bar{\lambda} &= \frac{\bar{\rho}}{\sqrt{1-\bar{\rho}^2}}, \end{aligned}$$

where $\bar{x}$ is the sample mean, s^2 is the sample variance, and m_3 is the third central sample moment. These moment estimators will be useful in choosing the initial values for solving the nonlinear likelihood equations given in the next section.

5.1 Maximum Likelihood Estimates

The log likelihood function is given by

$$\ln\ L = C - \frac{n}{2}\ln\sigma^2 - \frac{1}{2}\sum_{i=1}^{n}\frac{(x_i-\mu)^2}{\sigma^2} + \sum_{i=1}^{n}\ln\Phi\left[\lambda\left(\frac{x_i-\mu}{\sigma}\right)\right]. \tag{5.3}$$

The likelihood equations are given by

$$\begin{aligned}
\frac{\partial \ln L}{\partial \mu} &= \sum_{i=1}^{n}\left(\frac{x_i-\mu}{\sigma^2}\right) - \frac{\lambda}{\sigma}\sum_{i=1}^{n}\frac{\phi\left[\lambda\left(\frac{x_i-\mu}{\sigma}\right)\right]}{\Phi\left[\lambda\left(\frac{x_i-\mu}{\sigma}\right)\right]} = 0,\\
\frac{\partial \ln L}{\partial \sigma^2} &= -\frac{n}{2\sigma^2} + \frac{1}{2}\sum_{i=1}^{n}\frac{(x_i-\mu)^2}{\sigma^4} - \frac{\lambda}{2\sigma^3}\sum_{i=1}^{n}\frac{\phi\left[\lambda\left(\frac{x_i-\mu}{\sigma}\right)\right]}{\Phi\left[\lambda\left(\frac{x_i-\mu}{\sigma}\right)\right]}(x_i-\mu) = 0,
\end{aligned}$$

and

$$\frac{\partial \ln L}{\partial \lambda} = \sum_{i=1}^{n}\frac{\phi\left[\lambda\left(\frac{x_i-\mu}{\sigma}\right)\right]}{\Phi\left[\lambda\left(\frac{x_i-\mu}{\sigma}\right)\right]}\left(\frac{(x_i-\mu)}{\sigma}\right) = 0.$$

If we let $W(x_i) = \phi\,[\mathrm{A}\left(\frac{x_i-\mu}{\sigma}\right)] / \Phi\left[\lambda\left(\frac{x_i-\mu}{\sigma}\right)\right]$, the above equations become

$$\sum_{i=1}^{n}\left(\frac{x_i-\mu}{\sigma^2}\right) = \frac{\lambda}{\sigma}\sum_{i=1}^{n} W(x_i), \tag{5.4}$$

$$\frac{n}{\sigma^2} = \sum_{i=1}^{n}\frac{(x_i-\mu)^2}{\sigma^4} - \frac{\lambda}{2\sigma^3}\sum_{i=1}^{n} W(x_i)(x_i-\mu), \tag{5.5}$$

and

$$\sum_{i=1}^{n} W(x_i)(x_i-\mu) = 0. \tag{5.6}$$

Solving (5.3), (5.4), and (5.5), the maximum likelihood estimators are given by

$$\hat{\mu} = \frac{\sum_{i=1}^{n} x_i W(x_i)}{\sum_{i=1}^{n} W(x_i)}, \tag{5.7}$$

$$\hat{\sigma}^2 = \sum_{i=1}^{n}\frac{(x_i-\hat{\mu})^2}{n} \tag{5.8}$$

and

$$\hat{\lambda} = \frac{\sum_{i=1}^{n}\left(\frac{x_i-\hat{\mu}}{\sigma}\right)}{\sum_{i=1}^{n} W(x_i)} \tag{5.9}$$

The inverse of the Fisher Information Matrix can be used to find the variance-covariance matrix for the estimates. The Fisher Information Matrix is given by

$\hat{I}_{3\times 3} =$

$$\begin{bmatrix} \frac{n\left(1+\lambda^2 a_0\right)}{\sigma^2} & \frac{nb\lambda\left(1+2\lambda^2\right)}{\sigma^2\left(1+\lambda^2\right)^{3/2}} + \frac{n\lambda^2 a_1}{\sigma^2} & \frac{1}{\sigma}n\left[\frac{b}{\left(1+\lambda^2\right)^{3/2}} - \lambda a_1\right] \\ \frac{nb\lambda\left(1+2\lambda^2\right)}{\sigma^2\left(1+\lambda^2\right)^{3/2}} + \frac{n\lambda^2 a_1}{\sigma^2} & \frac{n}{\sigma^2}\left(1+\lambda^2 a_2\right) & -\frac{n\lambda a_2}{\sigma} \\ \frac{1}{\sigma}n\left[\frac{b}{\left(1+\lambda^2\right)^{3/2}} - \lambda a_1\right] & -\frac{n\lambda a_2}{\sigma} & na_2 \end{bmatrix}$$

where

$$\begin{aligned} b &= \sqrt{\frac{2}{\pi}}, \quad \text{and} \\ a_k &= E\left\{Z^k\left(\frac{\phi(\lambda z)}{\Phi(\lambda z)}\right)^2\right\}, (k = 0, 1, 2), \quad \text{with} \\ Z &= \frac{X-\mu}{\sigma}. \end{aligned}$$

The derivation of the matrix can be found in the appendix.

5.2 Estimates for IQ Data

Arnold et al. **(1993)** applied the skew normal distribution to a portion of an IQ score data set from Roberts **(1988).** In this section we expand the application to the full data set. The Roberts IQ data gives the Otis IQ scores for 87 white males and 52 non-white males hired by a large insurance company in **1971.** The data is given in the following tables:

Table 5.1: Otis IQ scores for whites

124, 106, 108, 112, 113, 122, 100, 108, 108, 94, 102, 120, 101, 118,
113, 117, 100, 106, 111, 107, 112, 120, 102, 135, 125, 98, 121, 117,
124, 114, 103, 122, 122, 113, 113, 104, 103, 113, 120, 106, 132, 106,
112, 118, 113, 112, 112, 121, 112, 85, 117, 109, 104, 129, 140, 106,
115, 109, 122, 108, 119, 121, 113, 107, 122, 103, 97, 116, 114, 131,
94, 112, 108, 118, 112, 116, **113, 111,** 122, 112, 136, 116, 108, 112,
108, 116, 103

Table 5.2: Otis IQ scores for non-whites

91, 102, 100, 117, 122, 115, 97, 109, 108, 104, 108, 118, 103, 123,
123, 103, 106, 102, 118, 100, **103,** 107, 108, 107, 97, 95, 119, 102,
108, 103, 102, 112, 99, 116, 114, 102, 111, 104, 122, 103, 111, 101,
91, 99, 121, 97, 109, 106, 102, 104, 107, 95

To apply the skew normal as a truncated normal according to the motivation of the model given by Arnold et al. (1993), we assume that these individuals were screened with respect to some variable Y, which is unknown. We further assume that only individuals who scored above average with respect to the screening variable were hired.

Let X represent the IQ scores of the individuals hired. The variable X is the unscreened variable, and only this variable is observed. We assume that (X, Y) has a bivariate normal distribution with mean vector (μ_1, μ_2), variance vector (σ_1^2, σ_2^2) and correlation p. Therefore the observed IQ scores represent a sample from a non-standard skew normal distribution.

We apply the non-standard skew normal maximum likelihood estimators to the IQ score sample to estimate the mean and variance of IQ scores for the unscreened population. We now let X_1 represent the score for whites and let X_2 represent the scores for non-whites. The two data sets displayed above are analyzed separately, each under the assumptions of normality and skew normality. The estimates are given in the following tables:

Table 5.3: Parameter estimates for Otis IQ scores for whites

paramer	normal	skew normal	(95% Confidence Interval)
μ_1	112.86	105.78	(100.38, 111.26)
σ_1	9.58	11.94	(8.81, 15.86)
λ_1		1.14	(-0.09, 1.77)

Table 5.4: Parameter estimates for Otis IQ scores for non-whites

paramer	normal	skew normal	(95% Confidence Interval)
μ_1	106.65	98.79	(93.11, 104.47)
σ_1	8.23	11.38	(8.17, 11.71)
λ_1		1.71	(0.4, 2.02)

For both data sets, under the assumption of normality the mean is overestimated and the standard deviation is underestimated.

Using the estimates from above, we first transform the data sets, given in Tables 5.1 and 5.2, to the data sets on standard skew-normal random variables Z_1 and Z_2. We then estimate λ_1' and λ_2' for the standard skew-normal random variables. The resulting estimates are $\lambda_1' = 1.15$ and $\lambda_2' = 1.84$. We now employ the machinery developed in section **4** to estimate the probability that the IQ score for a white employee is less than the IQ score for a non-white employee. From the estimates of λ_1' and λ_2' and (4.9), we have $\delta_1 = .755$ and $\delta_2 = .878$ and

$$\begin{aligned} P(Z_1 < Z_2) &= \frac{1}{\pi}\arctan\left(\frac{\delta_2 - \delta_1}{\sqrt{2 - \delta_1^2 - \delta_2^2}}\right) + \frac{1}{2} \\ &= .5473. \end{aligned}$$

Chapter 6

CONCLUSIONS AND RECOMMENDATIONS

In many real life applications, it has been observed that the unrestricted use of the normal distribution to model data can yield erroneous results. For the Roberts (1988) data analyzed in chapter 5, the application of normality resulted in overestimates of the mean IQ scores in both cases. This is due to the fact that the scores were obtained by screening on some other variable which is unknown, giving rise to skewness in the data. For this reason, we have used the skew normal distribution to estimate the desired probability.

There are many possible extensions of the skew normal model. Azzalini and Dalla Valle (1996) and Azzalini and Capitanio (1999) have investigated the properties of the multivariate skew normal. The reliability functions examined in this paper may also be extended to the multivariate case. Also many other reliability properties of skew normal or multivariate skew normal models which are true for a normal distribution can be investigated, see for example Gupta and Gupta **(1997,** 2000).

Another area of interest is the application of the univariate skew normal to linear models. In the recent presidential election, there was considerable discussion about the measurement error of the machine and hand recounts. In linear models, the error term is assumed to be normal with mean zero. However, if we consider each Florida county separately, with each showing a significant margin of victory for one of the two candidates, then the measurement error will have skewness in favor of the winner. Therefore, it is appropriate to investigate the nature of a linear model with a skew normal error term.

BIBLIOGRAPHY

[1] Arnold, B. C., Beaver, R. J., Groeneveld, R. A. and Meeker, W. Q. **(1993).** The nontruncated marginal of a truncated bivariate **normal** distribution. *Psychometrika,* 58(3), **471-488.**

[2] Azzalini, A. **(1985).** A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics,* **12,** 171-178.

[3] Azzalini, A. **(1986).** Further results on a class of distributions which includes the normal ones. *Statistica, 46*(2), **199-208.**

[4] Azzalini, A, and Capitanio, A. (1999). Statistical applications of the multivariate skew normal distribution. *Journal of the Royal Statistical Society,* **81**(3), **579-602.**

[5] Azzalini, **A.** and Dalla Valle, A. **(1996).** The multivariate skew-normal distribution. *Biornetrika,* 83(4), **715-726.**

[6] Geary, **R.** C., **(1947).** Testing for Normality. *Biometrika,* **34, 209-242.**

[7] Gupta, P. L., and Gupta, R. C. (1997). On the multivariate normal hazard. *Journal of Multivariate Analysis,* 62(1), 64-73.

[8] Gupta, P. L., and Gupta, R. C. (2001). Failure rate of the minimum and maximum of a multivariate normal distribution. To appear in *Metrika.*

[9] Gupta, R. C., (1981). On the mean residual life function in survival studies. *Statistical Distributions in Scientific Work,* 5, D-Reidel Publishing Company, Boston, 327-334.

[10] Gupta, R. C. and Kirmani, S. N. U. A. (1998). Residual life function in reliability studies. *Frontiers in Reliability,* 4, 175-190.

[11] Mudholkar, G. S. and Hutson, A. D. (2000). The epsilon-skew-normal distribution for analyzing near-normal data. *Journal of Statistical Planning and Inference,* 83, 291-309.

[12] Owen, D. B. (1956). Tables for computing bivariate normal probabilities. *Annals of Mathematical Statistics,* 27, 1075-1090.

[13] Prentice, R. L. (1975). Discrimination among some parametric models. *Biometrika,* 62, 607-614.

[14] Roberts, H. **V.** (1988). *Data Analysis for Managers with Minitab.* Redwood City, CA:Scientific Press.

[15] Turner, M. E. (1960). On heuristic estimation methods. *Biometrics,*16, 299-401.

APPENDIX

FISHER INFORMATION MATRIX

The elements of the Fisher Information Matrix are derived as follows. We let

$$
\begin{aligned}
b &= \sqrt{\frac{2}{\pi}}, \\
a_k &= E\left\{ Z^k \left(\frac{\phi(\lambda z)}{\Phi(\lambda z)} \right)^2 \right\}, (k = 0, 1, 2), \quad \text{and} \\
Z_i &= \frac{X_i - \mu}{\sigma}.
\end{aligned}
$$

The log likelihood function is given by

$$
\ln L = C - n \ln \sigma^2 - \sum_{i=1}^{n} \frac{Z_i^2}{2} + \sum_{i=1}^{n} \ln \Phi(\lambda Z_i), \quad \text{where } C \text{ is a constant.}
$$

We now derive the elements of the matrix. First, we have

$$
\begin{aligned}
I_{1,1} &= -E\left(\frac{\partial^2}{\partial\mu^2}\ln L\right) = -E\left\{\frac{\partial}{\partial\mu}\left(\sum_{i=1}^{n}\frac{Z_i}{\sigma} - \frac{\lambda}{\sigma}\sum_{i=1}^{n}\frac{\phi(\lambda Z_i)}{\Phi(\lambda Z_i)}\right)\right\} \\
&= -E\left\{-\frac{n}{\sigma^2} - \frac{\lambda^2}{\sigma^2}\sum_{i=1}^{n}\left[\frac{(\lambda Z_i)\,\phi(\lambda Z_i)}{\Phi(\lambda Z_i)} + \left(\frac{\phi(\lambda Z_i)}{\Phi(\lambda Z_i)}\right)^2\right]\right\} \\
&= \frac{n}{\sigma^2} + \frac{na_0\lambda^2}{\sigma^2} + \frac{\lambda^2}{\sigma^2}\sum_{i=1}^{n}\int_{-\infty}^{\infty}\frac{(\lambda Z_i)\,\phi(\lambda Z_i)}{\Phi(\lambda Z_i)}2\phi(Z_i)\,\Phi(\lambda Z_i)\,dZ_i \\
&= \frac{n}{\sigma^2} + \frac{na_0\lambda^2}{\sigma^2} + \frac{2\lambda^3}{\sigma^2}\sum_{i=1}^{n}\int_{-\infty}^{\infty}\frac{1}{2\pi}Z_i\exp\left\{-\frac{Z_i^2\left(1+\lambda^2\right)}{2}\right\}dZ_i \\
&= \frac{n}{\sigma^2} + \frac{na_0\lambda^2}{\sigma^2} + \frac{2\lambda^3}{\sigma^2}\sum_{i=1}^{n}\int_{-\infty}^{\infty}\frac{1}{2\pi\left(1+\lambda^2\right)}ue^{-\frac{u^2}{2}}du.
\end{aligned}
$$

It is known that the first moment of a standard normal random variable is zero. Thus

$$
I_{1,1} = \frac{n}{\sigma^2} + \frac{na_0\lambda^2}{\sigma^2}.
$$

Next, we have

$$\begin{aligned}
I_{1,2} &= I_{2,1} = -E\left(\frac{\partial^2}{\partial\sigma\partial\mu}\ln L\right) = -E\left\{\frac{\partial}{\partial\sigma}\left(\sum_{i=1}^{n}\frac{Z_i}{\sigma} - \frac{\lambda}{\sigma}\sum_{i=1}^{n}\frac{\phi(\lambda Z_i)}{\Phi(\lambda Z_i)}\right)\right\} \\
&= -E\left\{\sum_{i=1}^{n}\frac{-2Z_i}{\sigma^2} - \frac{\lambda}{\sigma}\sum_{i=1}^{n}\left[\left(Z_i^2\lambda^2 - 1\right)\frac{\phi(\lambda Z_i)}{\Phi(\lambda Z_i)} + \lambda Z_i\left(\frac{\phi(\lambda Z_i)}{\Phi(\lambda Z_i)}\right)^2\right]\right\} \\
&= \frac{2n\lambda b}{\sigma^2\sqrt{1+\lambda^2}} + \frac{\lambda^3}{\sigma^2}\sum_{i=1}^{n}\int_{-\infty}^{\infty}\frac{Z_i^2\phi(\lambda Z_i)}{\Phi(\lambda Z_i)}2\phi(Z_i)\Phi(\lambda Z_i)\,dZ_i \\
&\quad -\frac{\lambda}{\sigma^2}\sum_{i=1}^{n}\int_{-\infty}^{\infty}\frac{\phi(\lambda Z_i)}{\Phi(\lambda Z_i)}2\phi(Z_i)\Phi(\lambda Z_i)\,dZ_i + \frac{\lambda^2 n a_1}{\sigma^2} \\
&= \frac{2n\lambda b}{\sigma^2\sqrt{1+\lambda^2}} + \frac{2\lambda^3}{\sigma^2}\sum_{i=1}^{n}\int_{-\infty}^{\infty}\frac{1}{2\pi}Z_i^2\exp\left\{-\frac{Z_i^2\left(1+\lambda^2\right)}{2}\right\}dZ_i \\
&\quad -\frac{2\lambda}{\sigma^2}\sum_{i=1}^{n}\int_{-\infty}^{\infty}\frac{1}{2\pi}\exp\left\{-\frac{Z_i^2\left(1+\lambda^2\right)}{2}\right\}dZ_i + \frac{\lambda^2 n a_1}{\sigma^2} \\
&= \frac{2n\lambda b}{\sigma^2\sqrt{1+\lambda^2}} + \frac{2\lambda^3}{\sigma^2}\sum_{i=1}^{n}\int_{-\infty}^{\infty}\frac{1}{2\pi\left(1+\lambda^2\right)^{\frac{3}{2}}}u^2 e^{-\frac{u^2}{2}}\,du \\
&\quad -\frac{2\lambda}{\sigma^2}\sum_{i=1}^{n}\int_{-\infty}^{\infty}\frac{1}{2\pi\sqrt{1+\lambda^2}}e^{-\frac{u^2}{2}}\,du + \frac{\lambda^2 n a_1}{\sigma^2}.
\end{aligned}$$

It is known that the even moments of a standard normal random variable equal one, thus

$$\begin{aligned}
I_{1,2} &= I_{2,1} = \frac{2n\lambda b}{\sigma^2\sqrt{1+\lambda^2}} + \frac{\lambda^3 n b}{\sigma^2\left(1+\lambda^2\right)^{\frac{3}{2}}} - \frac{n\lambda b}{\sigma^2\sqrt{1+\lambda^2}} + \frac{\lambda^2 n a_1}{\sigma^2} \\
&= \frac{\lambda n b\left(1+2\lambda^2\right)}{\sigma^2\left(1+\lambda^2\right)^{\frac{3}{2}}} + \frac{\lambda^2 n a_1}{\sigma^2}.
\end{aligned}$$

Next, we have

$$
\begin{aligned}
I_{1,3} &= I_{3,1} = -E\left(\frac{\partial^2}{\partial\mu\partial\lambda}\ln L\right) = -E\left\{\frac{\partial}{\partial\mu}\left(\sum_{i=1}^{n}\frac{Z_i\phi(\lambda Z_i)}{\Phi(\lambda Z_i)}\right)\right\} \\
&= -E\left\{\frac{1}{\sigma}\sum_{i=1}^{n}\left[\left(Z_i^2\lambda^2-1\right)\frac{\phi(\lambda Z_i)}{\Phi(\lambda Z_i)} + \lambda Z_i\left(\frac{\phi(\lambda Z_i)}{\Phi(\lambda Z_i)}\right)^2\right]\right\} \\
&= -\frac{\lambda^2}{\sigma}\sum_{i=1}^{n}\int_{-\infty}^{\infty}\frac{Z_i^2\phi(\lambda Z_i)}{\Phi(\lambda Z_i)}2\phi(Z_i)\,\Phi(\lambda Z_i)\,dZ_i \\
&\quad +\frac{1}{\sigma}\sum_{i=1}^{n}\int_{-\infty}^{\infty}\frac{\phi(\lambda Z_i)}{\Phi(\lambda Z_i)}2\phi(Z_i)\,\Phi(\lambda Z_i)\,dZ_i - \frac{\lambda n a_1}{\sigma} \\
&= -\frac{\lambda^2 bn}{\sigma\left(1+\lambda^2\right)^{\frac{3}{2}}} + \frac{bn}{\sigma\sqrt{1+\lambda^2}} - \frac{\lambda n a_1}{\sigma} \\
&= \frac{n}{\sigma}\left(\frac{b}{\left(1+\lambda^2\right)^{\frac{3}{2}}} - a_1\lambda\right).
\end{aligned}
$$

Next, we have

$$
\begin{aligned}
I_{2,2} &= -E\left(\frac{\partial^2}{\partial\sigma^2}\ln L\right) = -E\left\{\frac{\partial}{\partial\sigma}\left(-\frac{2n}{\sigma} + \sum_{i=1}^{n}\frac{Z_i^2}{\sigma} - \frac{\lambda}{\sigma}\sum_{i=1}^{n}\frac{Z_i\phi(\lambda Z_i)}{\Phi(\lambda Z_i)}\right)\right\} \\
&= -E\left\{-\frac{2n}{\sigma} - \sum_{i=1}^{n}\frac{3Z_i^2}{\sigma^2}\right. \\
&\quad \left. -\frac{\lambda}{\sigma^2}\sum_{i=1}^{n}\left[\left(\lambda^2 Z_i^3 - 2Z_i\right)\frac{\phi(\lambda Z_i)}{\Phi(\lambda Z_i)} + \lambda Z_i^2\left(\frac{\phi(\lambda Z_i)}{\Phi(\lambda Z_i)}\right)^2\right]\right\}.
\end{aligned}
$$

Again using the fact that the odd moments of a standard normal are zero and the fact that the second moment is one, we have

$$
I_{2,2} = \frac{n}{\sigma^2}\left(1+\lambda^2 a_2\right).
$$

Next, we have

$$\begin{aligned} I_{2,3} &= I_{3,2} = -E\left(\frac{\partial^2}{\partial\sigma\partial\lambda}\ln L\right) = -E\left\{\frac{\partial}{\partial\sigma}\left(\sum_{i=1}^{n}\frac{Z_i\phi(\lambda Z_i)}{\Phi(\lambda Z_i)}\right)\right\} \\ &= -E\left\{\frac{1}{\sigma}\sum_{i=1}^{n}\left[\left(Z_i^3\lambda^2 - Z_i\right)\frac{\phi(\lambda Z_i)}{\Phi(\lambda Z_i)} + \lambda Z_i^2\left(\frac{\phi(\lambda Z_i)}{\Phi(\lambda Z_i)}\right)^2\right]\right\} \\ &= \frac{-n\lambda a_2}{\sigma}. \end{aligned}$$

Finally, we have

$$\begin{aligned} I_{3,3} &= -E\left(\frac{\partial^2}{\partial\lambda^2}\ln L\right) = -E\left\{\frac{\partial}{\partial\lambda}\left(\sum_{i=1}^{n}\frac{Z_i\phi(\lambda Z_i)}{\Phi(\lambda Z_i)}\right)\right\} \\ &= -E\left\{-\sum_{i=1}^{n}\left[\frac{\lambda Z_i^2\phi(\lambda Z_i)}{\Phi(\lambda Z_i)} + Z_i^2\left(\frac{\phi(\lambda Z_i)}{\Phi(\lambda Z_i)}\right)^2\right]\right\} \\ &= na_{2.} \end{aligned}$$

BIOGRAPHY OF THE AUTHOR

Nicole Brown was born in Waterville, Maine on April 11, 1975. She was raised in Clinton, Maine and graduated from Lawrence High School in 1993. She attended Bowdoin College in Brunswick, Maine and graduated in 1997 with a Bachelor's degree in Mathematics and Economics. She entered the Mathematics and Statistics graduate program at The University of Maine in the fall of 1998. Nicole is a candidate for **a** Master of Arts degree in Mathematics from The University of Maine in May, **2001.**