Improved Estimators of Normal Tail Probabilities

by

Manohari Nalliah
Purdue University

Technical Report #86-26

Department of Statistics
Purdue University

July 1986

Improved Estimators of Normal Tail Probabilities

Abstract

The estimation problem of normal tail probabilities is considered. This study shows that the best unbiased estimator is generalized Bayes and it is asymptotically admissible within the class of estimators whose risk functions depend only on the ratio of mean and standard deviation. Improved estimators are found for large values and small values of the parameter.

1. Introduction

Let $x_1, x_2, \ldots, x_n, n \geq 2$ be a random sample from $N(\mu, \sigma^2)$ where μ and σ^2 are unknown. In many practical situations arising in reliability theory, quality control, insurance problems, etc., it is important to estimate tail probability $\theta = P(x_1 < a) = \Phi((a-\mu)/\sigma)$, where a is a given constant. It is sufficient to consider $a = o$.

Let $X = \sum_{1}^{n} x_j/n,\ S^2 = \sum_1^n (x_j - X)^2$ and $Z = X/S$. Different forms of the uniformly minimum variance unbiased (UMVU) estimator δ_U has been given in Kolmogorov (1950), Folks, Pierce and Stewart (1965) and Rukhin (1985). Generalized Bayes rules for a family of prior densities are obtained, and some of them substantially improve upon UMVU estimator for small values of the parameter and the others for large values of the parameter. The specific form (1.1) is used in Theorem (2.1) to show that δ_U is generalized Bayes estimator, for $n \geq 3$.

$$\delta_U(X,S) = \begin{cases} \int_V^\infty \frac{dt}{(1+t^2)^{\frac{n-1}{2}}} / B\left(\frac{n-2}{2}, \frac{1}{2}\right) & \text{if } |Z| < \left(\frac{n-1}{n}\right)^{1/2} \\ 1 & \text{if } Z < -\left(\frac{n-1}{n}\right)^{1/2} \\ 0 & \text{otherwise.} \end{cases} \tag{1.1}$$

where $V = \frac{\sqrt{n}Z}{(n-1-nZ^2)^{1/2}}$.

2. Generalized Bayes Estimators

Let $\lambda(\mu,\sigma) = \exp\{\frac{c}{2}\left(\frac{\mu}{\sigma}\right)^2\}\sigma^{-\alpha}, 0 < c < n$, be the generalized prior density for parameters μ and σ. The corresponding generalized Bayes estimator $\delta_{Bc}(X,S)$ for quadratic loss has the form

$$\delta_{Bc}(X,S) = \frac{\int_{-\infty}^{\infty} \int_0^\infty \Phi(-\mu/\sigma) f(X,S)\lambda(\mu,\sigma)d\mu\frac{d\sigma}{\sigma}}{\int_{-\infty}^{\infty} \int_0^\infty f(X,S)\lambda(\mu,\sigma)d\mu\frac{d\sigma}{\sigma}}$$

where

$$f(X,S) \propto \frac{1}{\sigma^n} \exp\{-[n(X-\mu)^2 + S^2]/2\sigma^2\}S^{n-2}.$$

Let $\eta = \mu/\sigma$ and $y = S/\sigma$. Then,

$$\delta_{Bc}(X,S) = \frac{\int_{-\infty}^{\infty}\int_0^{\infty} \Phi(-\eta)\exp\{-[n(yZ-\eta)^2+y^2]/2\}\exp\{\frac{c}{2}\eta^2\}y^{n+\alpha-2}d\eta dy}{\int_{-\infty}^{\infty}\int_0^{\infty} \exp\{-[n(yZ-\eta)^2+y^2]/2\}\exp\{\frac{c}{2}\eta^2\}y^{n+\alpha-2}d\eta\, dy}. \quad (2.1)$$

$$\int_{-\infty}^{\infty} \exp\left\{-n(yZ-\eta)^2/2\right\}\exp\left\{\frac{c}{2}\eta^2\right\}d\eta = (2\pi)^{1/2}(n-c)^{-1/2}\exp\left\{\frac{1}{2}\frac{nc}{n-c}y^2Z^2\right\}. \quad (2.2)$$

The denominator of (2.1) can be written using (2.2) as

$$(2\pi)^{1/2}(n-c)^{-1/2}\int_0^{\infty}\exp\left\{\frac{1}{2}\left(\frac{nc}{n-c}y^2Z^2-y^2\right)\right\}y^{n+\alpha-2}dy$$

$$= (2\pi)^{1/2}2^{\frac{n+\alpha-3}{2}}(n-c)^{-1/2}\Gamma\left(\frac{n+\alpha-1}{2}\right)\left(1-\frac{ncZ^2}{n-c}\right)^{-\frac{n+\alpha-1}{2}} \quad (2.3)$$

for $|Z| < \left(\frac{n-c}{nc}\right)^{1/2}$.

Now,

$$\int_{-\infty}^{\infty}\Phi(-\eta)\exp\left\{-n(yZ-\eta)^2/2\right\}\lambda(\eta)d\eta$$

$$= (2\pi)^{-1/2}\int_{-\infty}^{\infty}\int_0^{\infty}\exp\left\{-[(t+\eta)^2+n(yZ-\eta)^2-c\eta^2]/2\right\}d\eta\, dt$$

$$= (n+1-c)^{-1/2}\int_0^{\infty}\exp\left\{-[t^2+n(yZ)^2-\frac{(t-nyZ)^2}{n+1-c}]/2\right\}dt. \quad (2.4)$$

Let $t = y\ell$. Then the numerator of (2.1) becomes, using (2.4),

$$(n+1-c)^{-1/2}\int_0^{\infty}\int_0^{\infty}\exp\left\{-y^2[1+\ell^2+nZ^2-\frac{(\ell-nZ)^2}{n+1-c}]/2\right\}y^{n+\alpha-1}\,d\ell\, dy$$

$$= (n+1-c)^{-1/2}\int_0^{\infty}\int_0^{\infty}\exp\left\{-y^2[(1-\frac{nc}{n-c}Z^2)+\frac{n-c}{n+1-c}(\ell+\frac{nZ}{n-c})^2]/2\right\}y^{n+\alpha-1}d\ell\, dy$$

$$= \frac{2^{\frac{n+\alpha-2}{2}}\Gamma(\frac{n+\alpha}{2})}{(n+1-c)^{1/2}}\int_0^{\infty}\frac{d\ell}{[(1-\frac{nc}{n-c}Z^2)+\frac{n-c}{n+1-c}(\ell+\frac{nZ}{n-c})^2]^{\frac{n+\alpha}{2}}} \quad (2.5)$$

for $|Z| < \left(\frac{n-c}{nc}\right)^{1/2}$.

From (2.3) and (2.5),

$$\delta_{Bc}(X,S) = \begin{cases} \int_{V_1}^{\infty}\frac{dt}{(1+t^2)^{\frac{n+\alpha}{2}}}/B(\frac{n+\alpha-1}{2},\frac{1}{2}) & \text{if } |Z| < \left(\frac{n-c}{nc}\right)^{1/2} \\ 1 & \text{if } Z < -\left(\frac{n-c}{nc}\right)^{1/2} \\ 0 & \text{otherwise} \end{cases} \quad (2.6)$$

where

$$V_1 = \frac{nZ}{(n+1-c)^{1/2}(n-c-ncZ^2)^{1/2}}. \tag{2.7}$$

Theorem 2.1

δ_U is generalized Bayes estimator with respect to the prior

$$\lambda(\mu,\sigma) = \exp\left\{1/2(\mu/\sigma)^2\right\}/\sigma, \text{ for } n \geq 3.$$

Proof

$\alpha = 1, c = 1$ in (2.6) gives δ_U.

If $c \to 0$, which corresponds to the uniform, non-informative prior for μ, one obtains

$$\delta_{Bc}(X,S) \longrightarrow \int_h^\infty \frac{dt}{(1+t^2)^{\frac{n+\alpha}{2}}} / B\left(\frac{n+\alpha-1}{2}, \frac{1}{2}\right) \tag{2.8}$$

where $h = \left(\frac{n}{n+1}\right)^{1/2} Z$.

The risk of (2.8) has been studied in Rukhin (1985). For $c \to n$, one obtains

$$\delta_{Bc}(X,S) \longrightarrow \begin{cases} 1 & \text{if } X/S \leq 0, \\ 0 & \text{if } X/S > 0. \end{cases}$$

3. Admissibility of Generalized Bayes Estimators

In this section, estimators of θ as a function of Z are considered. Let $\lambda(\eta) = \exp\left\{\frac{c}{2}\eta^2\right\}, 0 < c < n$. Then the corresponding generalized Bayes estimator is

$$\tilde{\delta}_{Bc}(Z) = \frac{\int_{-\infty}^{\infty} \Phi(-\eta)\tilde{\pi}_\eta(\lambda(\eta)d\eta}{\int_{-\infty}^{\infty} \tilde{\pi}_\eta(Z)\lambda(\eta)d\eta}$$

where

$$\tilde{\pi}_\eta(Z) \propto \int_0^\infty \exp\left\{-[n(SZ-\eta)^2 + S^2]/2\right\} S^{n-1}dS.$$

Thus,

$$\tilde{\delta}_{Bc}(Z) = \frac{\int_{-\infty}^{\infty}\int_0^\infty \Phi(-\eta)\exp\left\{-[n(SZ-\eta)^2 + S^2 - c\eta^2]/2\right\} S^{n-1}d\eta\, dS}{\int_{-\infty}^{\infty}\int_0^\infty \exp\{-[n(SZ-\eta)^2 + S^2 - c\eta^2]/2\} S^{n-1}d\eta\, dS}. \tag{3.1}$$

Note that, (3.1) is equivalent to (2.1) for $\alpha = 1$ and $y \equiv S$. Therefore,

$$\tilde{\delta}_{Bc}(Z) = \begin{cases} \int_{V_1}^{\infty} \frac{dt}{(1+t^2)^{\frac{n+1}{2}}} / B(\frac{n}{2}, \frac{1}{2}) & \text{if } |Z| < \left(\frac{n-c}{nc}\right)^{1/2} \\ 1 & \text{if } Z < -\left(\frac{n-c}{nc}\right)^{1/2} \\ 0 & \text{otherwise} \end{cases} \tag{3.2}$$

where V_1 is as in (2.7), and $V_1 = V$ for $c = 1$.

Theorem 3.1

$$|\tilde{\delta}_{B1}(Z) - \delta_U(Z)| \longrightarrow 0 \quad \text{as } n \longrightarrow \infty.$$

Proof

Integration by parts gives,

$$\begin{aligned} \tilde{\delta}_{B1}(Z) &= \int_V^{\infty} \frac{dt}{(1+t^2)^{\frac{n+1}{2}}} / B(\frac{n}{2}, \frac{1}{2}) \\ &= \left[\frac{n-2}{n-1} \int_0^{\infty} \frac{dt}{(1+t^2)^{\frac{n-1}{2}}} - \frac{1}{n-1} \frac{V}{(1+V^2)^{\frac{n+1}{2}}} \right] / B(\frac{n}{2}, \frac{1}{2}) \\ &= \left[\int_V^{\infty} \frac{dt}{(1+t^2)^{\frac{n-1}{2}}} / B(\frac{n-2}{2}, \frac{1}{2}) \right] - \left[\frac{1}{n-1} \frac{V}{(1+V^2)^{\frac{n+1}{2}}} / B(\frac{n}{2}, \frac{1}{2}) \right] \\ &= \delta_U(Z) - \frac{1}{n-1} \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})\pi^{1/2}} \frac{V}{(1+V^2)^{\frac{n+1}{2}}}. \end{aligned}$$

$$\frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})} \sim \left(\frac{n}{2}\right)^{1/2} (1 - \frac{1}{8n}) \text{ as } n \longrightarrow \infty.$$

Thus,

$$\left| \frac{1}{n-1} \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})\pi^{1/2}} \frac{V}{(1+V^2)^{\frac{n+1}{2}}} \right| \longrightarrow 0 \text{ as } n \rightarrow \infty.$$

Hence the result.

Remark

It is known from Rukhin (1985) that

$$R(0, \delta_U) \sim [2\pi(n-3)]^{-1}[1 - \frac{5}{2n} + o(n^{-1})] \text{ as } n \rightarrow \infty.$$

Therefore by Theorem (3.1), $R(0, \delta_{B1})$ has the same asymptotic behavior as $n \rightarrow \infty$.

Theorem 3.2

$\tilde{\delta}_{Bc}$ is admissible within the class of estimators whose risk functions depend only on η.

Proof

From (A6),

$$R(\eta, \tilde{\delta}_{Bc}) \sim F_{n1} \exp\left\{-\frac{c}{2}\eta^2\right\} / \eta^{n+3} \text{ as } |\eta| \to \infty$$

where F_{n1} is a constant depends only on n. Thus,

$$\int_{-\infty}^{\infty} R(\eta, \tilde{\delta}_{Bc})\lambda(\eta) d\eta = \int_{-\infty}^{\infty} R(\eta, \tilde{\delta}_{Bc}) \exp\left\{\frac{c}{2}\eta^2\right\} d\eta$$
$$< \infty.$$

Therefore, the Bayes risk is finite for $\tilde{\delta}_{Bc}$. Hence $\tilde{\delta}_{Bc}$ is admissible.

Remark:

δ_U is asymptotically admissible, because of Theorems 3.1 and 3.2.

4. Comparison of Risks

From (A4) and (A5), as $|\eta| \longrightarrow \infty$,

$$R(\eta, \delta_{Bc}) \sim F_{n\alpha} \exp\{-\frac{c}{2}\eta^2\}/\eta^{n+2\alpha+1}$$

and

$$R(\eta, \delta_U) \sim F_{n(-1)} \exp\{-\frac{1}{2}\eta^2\}/\eta^{n-1}.$$

Thus,

$$R(\eta, \delta_{Bc}) < R(\eta, \delta_U) \quad \text{for } c \geq 1,$$

and

$$R(\eta, \delta_{Bc}) > R(\eta, \delta_U) \quad \text{for } c < 1.$$

Because of the fact that the risk at $\eta = 0$ increases as α increases (fixed h) for estimators of the form

$$\int_h^{\infty} \frac{dt}{(1+Z^2)^{\frac{n+\alpha}{2}}} / B\left(\frac{n+\alpha-1}{2}, \frac{1}{2}\right),$$

$$R(0, \tilde{\delta}_{B1}) > R(0, \delta_U) \quad \text{and}$$

$$R(0, \delta_{B1}) > R(0, \delta_U) \quad \forall\, \alpha > -1.$$

5. Numerical Results

It is shown in Rukhin (1986) that for any integrable function $g(y)$, $y = n^{1/2}|X|/(nX^2 + S^2)^{1/2}$,

$$E_\xi g(y) = 2\exp\{-\xi^2/2\} \int_0^1 g(y)(1-y^2)^{\frac{n-3}{2}} \pi_{n-1}(y\xi)dy/d_{n-2} \tag{5.1}$$

where $\xi = n^{1/2}|\eta|$ and,

$$\begin{aligned}\pi_m(y) &= (2\pi)^{1/2} \int_0^\infty v^m \exp\{-v^2/2\} \cosh\,(vy)\; dv, \quad y \geq 0 \\ &= \begin{cases} .5\frac{d^m}{dy^m} \exp\{y^2/2\} & \text{m}-\text{even} \\ .5\frac{d^m}{dy^m}(\exp\{y^2/2\}\ \text{erf}(y2^{-1/2})) & \text{m}-\text{odd}\end{cases}\end{aligned} \tag{5.2}$$

All estimators δ considered in previous sections are bounded functions of Z such that $\delta(-Z) = 1 - \delta(Z)$. The quadratic risk depends only on η and is symmetric:

$$R(\eta, \delta) = R(-\eta, \delta) \tag{5.3}$$

Therefore, it is sufficient to take $\eta \geq 0$.

Hence from (5.1) and (5.2),

$$\begin{aligned} R(\delta,\eta) = D\{ &\int_0^1 \delta^2(y)(1-y^2)^{\frac{n-3}{2}} \pi_{n-1}(\xi y)\; dy \\ &- \int_0^1 \delta(y)(1-y^2)^{\frac{n-3}{2}} (\pi_{n-1}(\xi y) - \ \text{erfc}(\eta 2^{-1/2})\pi'_{n-2}(\xi y))\; dy\} \\ &+ \frac{1}{4}\ \text{erfc}^2(\eta 2^{-1/2}) + \frac{1}{2}\ \text{erfc}(\xi 2^{-1/2}) - \frac{1}{2}\ \text{erfc}(\xi 2^{-1/2})\ \text{erfc}(\eta 2^{-1/2}) \end{aligned} \tag{5.4}$$

where $\text{erfc}(y) = 1-\ \text{erf}(y)$and $D = \frac{2}{2^{\frac{n-3}{2}}\Gamma\frac{(n-1)}{2}} \exp\{-\frac{1}{2}\xi^2\}$.

The representation (5.4) facilitate numerical evaluation of risks of δ_U and $\tilde{\delta}_{Bc}$.

Figures I, II and III respectively show the risks for n =3, 4 and 5. These figures show that for $\eta \geq 1.0$,

$$\tilde{\delta}_{B1} \text{ is better than } \delta_U.$$

When $c = .1$, $\tilde{\delta}_{Bc}$ is better than δ_U for $\eta \leq 1.7$. The same result holds when $c = .2$ for $n = 3$ but not for $n = 4$ or 5.

$R(0, \tilde{\delta}_{B.2}) > R(0, \delta_U)$ from Figures II and III. Thus the optimal choice of c is .1 for small values of η and 1.0 for large values of η.

Appendix

Asymptotic behavior of $R(\eta, \delta)$ as $|\eta| \longrightarrow \infty$ is considered. Because of (5.3), it is sufficient to assume $\eta \longrightarrow \infty$.

Let

$$
\begin{aligned}
t &= \sqrt{n}Z/\sqrt{1+nZ^2} \\
\xi &= \sqrt{n}\eta \\
A_n &= 1/\left[\sqrt{2\pi}2^{\frac{n-3}{2}}\Gamma\left(\frac{n-1}{2}\right)\right] \\
B_{n\alpha} &= A_n(n-c+1)^{\frac{n+\alpha-1}{2}}(n+\alpha-1)^{-1}/B\left(\frac{n+\alpha-1}{2},\frac{1}{2}\right) \\
D_{n\alpha} &= B_{n\alpha}\frac{\sqrt{2\pi}}{2}\left(\frac{n}{n-c}\right)^{\frac{n+\alpha-1}{2}} \\
D^1_{n\alpha} &= B^2_{n\alpha}\frac{\sqrt{2\pi}}{2}\left(\frac{n}{n-c}\right)^{n+\alpha-1}/A_n \\
E_{n\alpha} &= D_{n\alpha}\left(\frac{c}{n}\right)^{\frac{n-3}{2}}\left(1-\frac{c}{n}\right)^{\frac{n-2}{2}}\Gamma\left(\frac{n+\alpha+1}{2}\right)2^{\frac{n+\alpha-1}{2}} \\
F_{n\alpha} &= D^1_{n\alpha}\left(\frac{c}{n}\right)^{\frac{n-3}{2}}(1-\frac{c}{n})^{\frac{n-2}{2}}\Gamma(n+\alpha)2^{n+\alpha}
\end{aligned}
$$

$$
\begin{aligned}
\delta_{Bc}(t) &\sim B_{n\alpha}(1-\frac{n}{n-c}t^2)^{\frac{n+\alpha-1}{2}} \quad \text{as } t\uparrow(1-\frac{c}{n})^{1/2} \text{ and} \\
&= 0 \qquad \text{if } t > (1-\frac{c}{n})^{1/2}
\end{aligned}
$$

$$
\begin{aligned}
E(\delta_{Bc}) &= A_n\exp\left\{-\frac{\xi^2}{2}\right\}\int_{-1}^{1}\delta_{Bc}(t)(1-t^2)^{\frac{n-3}{2}}\left[\int_0^\infty\exp\left\{-\frac{1}{2}(\ell^2+2\xi t\ell)\right\}\ell^{n-1}d\ell\right]dt \\
&\sim B_{n\alpha}\sqrt{2\pi}\exp\left\{-\frac{\xi^2}{2}\right\}\int_0^{\sqrt{1-c/n}}\left(1-\frac{n}{n-c}t^2\right)^{\frac{n+\alpha-1}{2}}(1-t^2)^{\frac{n-3}{2}}\exp\left\{\frac{(\xi t)^2}{2}\right\}(\xi t)^{n-1}dt \\
&\sim D_{n\alpha}\xi^{n-1}\exp\left\{-c\frac{\xi^2}{2n}\right\}\int_0^{1-c/n}y^{\frac{n+\alpha-1}{2}}\left(y+\frac{c}{n}\right)^{\frac{n-3}{2}}\left(1-\frac{c}{n}\right)^{\frac{n-2}{2}}\exp\left\{-\frac{\xi^2}{2}y\right\}dy \\
&\sim E_{n\alpha}\exp\left\{-c\frac{\xi^2}{2n}\right\}/\xi^{\alpha+1} \qquad (A1)
\end{aligned}
$$

$$E\delta_{Bc}^2 \sim D_{n\alpha}^1 \left(\frac{n}{n-c}\right)^{\frac{n+\alpha-1}{2}} \xi^{n-1} \exp\left\{-c\frac{\xi^2}{2n}\right\} \int_0^{1-c/n} y^{n+\alpha-1}$$

$$\left(y+\frac{c}{n}\right)^{\frac{n-3}{2}} \left(1-\frac{c}{n}\right)^{\frac{n-2}{2}} \exp\left\{-\frac{\xi^2}{2}y\right\} dy$$

$$\sim F_{n\alpha} \exp\left\{-c\frac{\xi^2}{2n}\right\} / \xi^{n+2\alpha+1} \tag{A2}$$

$$\theta \sim \frac{\sqrt{n}}{\sqrt{2\pi}} \exp\left\{-\frac{\xi^2}{2n}\right\} / \xi \tag{A3}$$

From (A1), (A2) and (A3)

$$R(\eta, \delta_{Bc}) \sim E\delta_{Bc}^2 \quad \text{as } |\eta| \longrightarrow \infty$$

$$\sim F_{n\alpha} \exp\left\{-c\frac{\xi^2}{2n}\right\} / \xi^{n+2\alpha+1} \tag{A4}$$

$\alpha = -1, \quad c = 1$ gives $R(\eta, \delta_U)$ and $\alpha = 1$ gives $R(\eta, \tilde{\delta}_{Bc})$ in $(A4)$.

Therefore,

$$R(\eta, \delta_U) \sim F_{n(-1)} \exp\left\{-\frac{\xi^2}{2}\right\} / \xi^{n-1} \tag{A5}$$

$$R(\eta, \tilde{\delta}_{Bc}) \sim F_{n(1)} \exp\left\{-\frac{\xi^2}{2}\right\} / \xi^{n+3} \tag{A6}$$

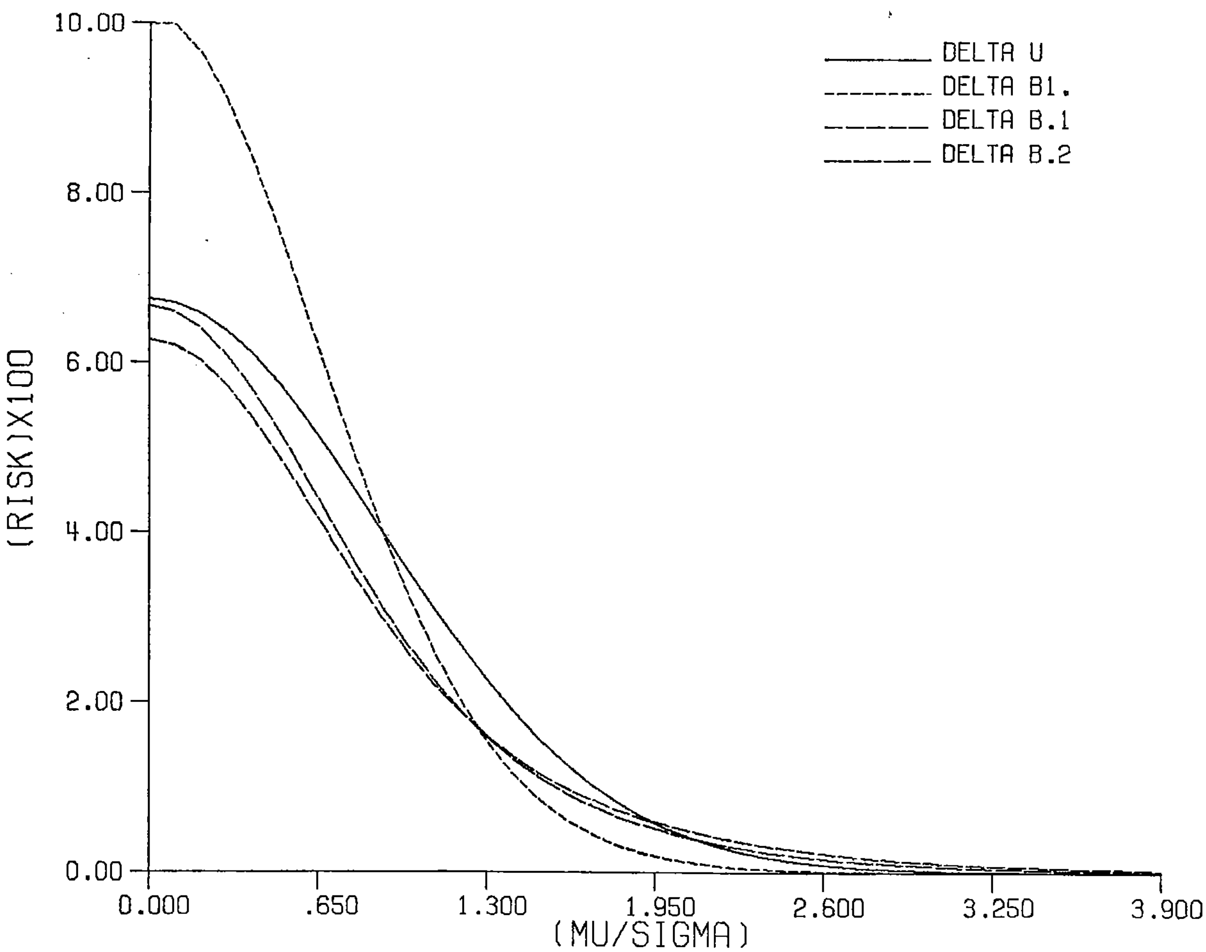
References

Erdelyi, A. (1956), Asymptotic Expansions, New York: Dover.

Rukhin, A. L. (1985), "Estimating Normal Tail Probabilities", Technical Report #84-46.

Rukhin, A. L. (1986), "How Much Better are Better Estimators of a Normal Variance", Technical Report #86-9.

FIGURE I



10.00
8.00
6.00
4.00
2.00
0.00
(RISK)X100
0.000
.650
1.300
1.950
2.600
3.250
3.900
(MU/SIGMA)
DELTA U
DELTA B1.
DELTA B.1
DELTA B.2

FIGURE II

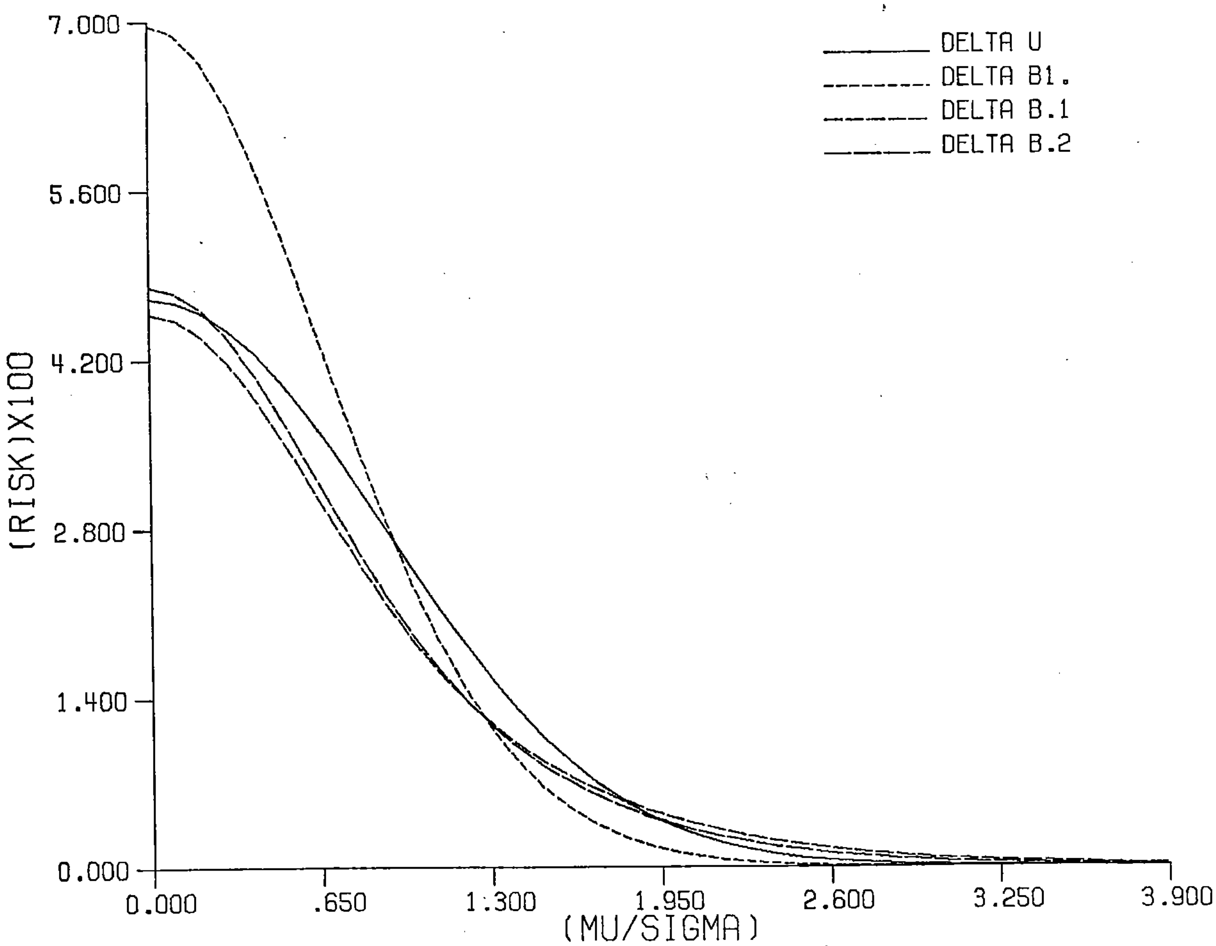
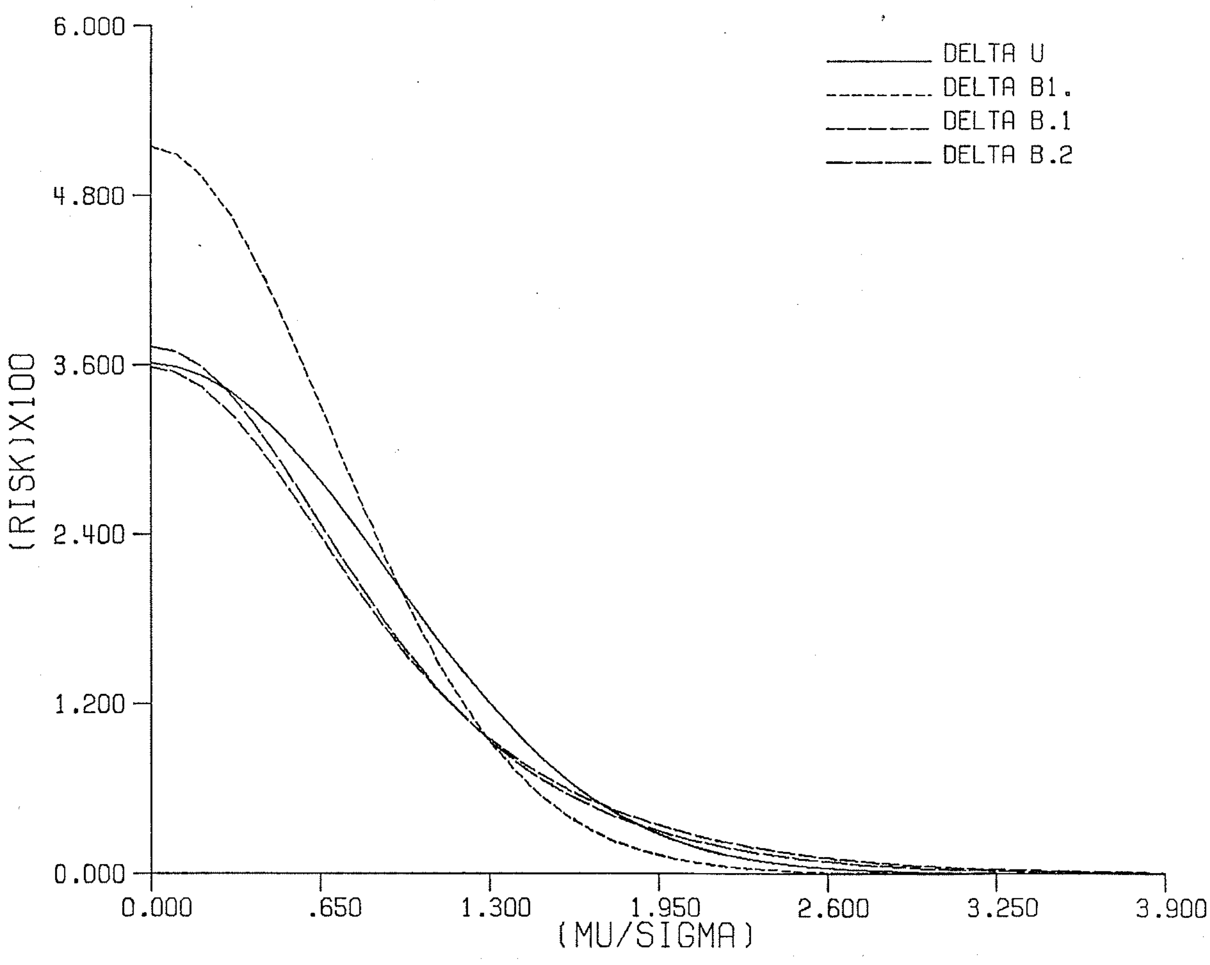

DELTA U
DELTA B1.
DELTA B.1
DELTA B.2
7.000
5.600
4.200
2.800
1.400
0.000
(RISK)X100
0.000
.650
1.300
1.950
2.600
3.250
3.900
(MU/SIGMA)

FIGURE III



6.000
4.800
3.600
2.400
1.200
0.000
(RISK)X100
0.000
.650
1.300
1.950
2.600
3.250
3.900
(MU/SIGMA)
DELTA U
DELTA B1.
DELTA B.1
DELTA B.2