The Annals of Mathematical Statistics
1969, Vol. 40, No. 6, 2149–2155

OPTIMUM ESTIMATORS FOR LINEAR FUNCTIONS OF LOCATION AND SCALE PARAMETERS[1]

BY NANCY R. MANN

Rocketdyne, A Division of North American Rockwell Corporation

0. Summary. In this paper, loss is taken to be proportional to squared error with the constant of proportionality equal to the square of the inverse of a scale parameter, and an invariant estimator is defined to be one with risk invariant under transformations of location and scale.

For certain classes of estimators, best (minimum-mean-squared-error) invariant estimators are found for specified linear functions of an unknown scale parameter and one or more unknown location parameters. Even when the specified function is equal to a single location parameter, the best invariant estimator is not equal to the best unbiased estimator in the class except for complete samples from certain distributions such as the Gaussian.

1. Introduction. In the following we consider invariant estimators of estimable functions of the form $\Phi = \boldsymbol{l}'\boldsymbol{\mu} + m\sigma$, where $\boldsymbol{\mu}$ is a p-dimensional column vector of location parameters, σ is a scale parameter, and $\boldsymbol{l}(p \times 1)$ and m are known. It is assumed that unique (with probability one) minimum-variance unbiased estimators exist in some class of estimators for $\mu_1, \mu_2, \cdots, \mu_p$ and σ. Then a unique minimum-expected-loss estimator of Φ among estimators in that class with risk invariant under transformations of location and scale can be determined as a function of these unbiased estimators. Loss is taken to be proportional to squared error, with the constant of proportionality equal to σ^{-2}.

The best (minimum-risk) invariant estimator of Φ is shown to be equal to the best unbiased estimator of Φ only when $\Phi = \boldsymbol{l}'\boldsymbol{\mu}$ and the density function for which $\mu_1, \mu_2, \cdots$, and μ_p are location parameters and σ is a scale parameter is symmetric about $\boldsymbol{l}'\boldsymbol{\mu}$. The best invariant estimators of $\mu_1, \mu_2, \cdots$ and μ_p for σ known are derived as functions of the best unbiased estimators of the μ's and of σ for σ unknown.

The results here generalize results of Goodman in [2] and [3] which apply to scale parameters only. Theorems in [7] which deal with general linear estimators of Φ are also generalized in this paper.

2. Best invariant estimators of Φ and σ. The following two assumptions are made.

A1. The r-dimensional random column vector $\boldsymbol{x}$ is such that the distribution of $(\boldsymbol{x} - \boldsymbol{Q}\boldsymbol{\mu})/\sigma$ is independent of $\boldsymbol{\theta}' = (\boldsymbol{\mu}', \sigma)$, with $\sigma > 0$, $\boldsymbol{\mu}' = (\mu_1, \mu_2, \cdots, \mu_p)$, and $\boldsymbol{Q}$ a known $r \times p$ matrix.

A2. Let $\boldsymbol{l}' = (l_1, l_2, \cdots, l_p)$ and m be known. For $(\boldsymbol{\mu}', \sigma)$ unknown, there exists

Received 13 December 1968.

[1] This research was supported by the Aerospace Research Laboratories, Office of Aerospace Research, United States Air Force, under Contract AF33(615)-2818.

a vector of unique[2] (with probability one) minimum-variance joint unbiased estimators (Φ^*, σ^*) of $(\boldsymbol{l}'\boldsymbol{\mu} + m\sigma, \sigma) \equiv (\Phi, \sigma)$ in a specified class ξ[3] of estimators of (Φ, σ). Let the covariance matrix of (Φ^*, σ^*) be

$$\sigma^2 \begin{bmatrix} A & B \\ B & C \end{bmatrix}.$$

We let loss be proportional to squared error, with the constant of proportionality equal to $1/\sigma^2$, and under the preceding assumptions, Theorem 1 below applies.

THEOREM 1. *Let $(\boldsymbol{\mu}, \sigma)$ be unknown and $\boldsymbol{l} \neq \mathbf{0}$. Consider estimators of $(\boldsymbol{l}'\mu + m\sigma, \sigma)$ with expected loss invariant under transformations of location and scale. $\tilde{\Phi} = \Phi^* - [B/(1 + C)]\sigma^*$ and $\tilde{\sigma} = \sigma^*/(1 + C)$ are the joint unique estimators of Φ and σ, respectively, with smallest expected squared deviation among such estimators in ξ. The mean squared errors of $\tilde{\Phi}$ and $\tilde{\sigma}$ are $[A - B^2/(1 + C)]\sigma^2$ and $[C/(1 + C)]\sigma^2$, respectively, and $E[(\tilde{\Phi} - \Phi)(\tilde{\sigma} - \sigma)]$ is $[B/(1 + C)]\sigma^2$.*

PROOF OF THEOREM 1. In the class ξ of estimators, consider any estimator $\bar{\Psi}$ of $\Psi = a(\boldsymbol{l}'\boldsymbol{\mu} + m\sigma) + c\sigma \equiv a\Phi + c\sigma$ (a and c not both zero) with invariant expected loss (or risk), that is, one such that the risk function of $\bar{\Psi}$ is independent of $(\mu_1, \mu_2, \cdots, \mu_p, \sigma)$. The risk function $R(\Psi, \bar{\Psi})$ of $\bar{\Psi}$ is equal to

$$\text{Var}(\bar{\Psi})/\sigma^2 + [E(\bar{\Psi} - \Psi)^2]/\sigma^2,$$

where Var (X) denotes the variance of the estimator X and $E(X - x)$ is the bias of X in estimating x.

Thus, each of Var $(\bar{\Psi})/\sigma^2$ and $[E(\bar{\Psi} - \Psi)]^2/\sigma^2$, being nonnegative, must also be independent of $\boldsymbol{\theta}' = (\boldsymbol{\mu}', \sigma)$ for all $\boldsymbol{\theta}$. The bias of $\bar{\Psi}$ is $E[\bar{\Psi} - a(\boldsymbol{l}'\boldsymbol{\mu} + m\sigma) + c\sigma]$. Therefore, $\bar{\Psi}$, in order to have risk independent of $(\boldsymbol{\mu}', \sigma)$, must be of the form

$$a\boldsymbol{l}'\hat{\boldsymbol{\mu}} + \kappa\hat{\sigma} \equiv a(\boldsymbol{l}'\hat{\boldsymbol{\mu}} + m\hat{\sigma}) + k\hat{\sigma} \equiv a\hat{\Phi} + k\hat{\sigma}$$

for some κ and some $k = \kappa - m$, where $\hat{\boldsymbol{\mu}}$, $\hat{\sigma}$, and $\hat{\Phi}$ are unbiased estimators of $\boldsymbol{\mu}$, σ, and Φ, respectively, and k is independent of $\mu_1, \mu_2, \cdots, \mu_p$, and σ for all $\boldsymbol{\theta}'$.

The risk of $\bar{\Psi}$ is thus given by

$$(1)\quad [a^2 \text{Var}(\hat{\Phi}) + 2ak \text{Cov}(\hat{\Phi}, \hat{\sigma}) + k^2 \text{Var}(\hat{\sigma})]/\sigma^2 + \{[E(a\hat{\Phi} + k\hat{\sigma} - a\Phi - c\sigma)]^2\}/\sigma^2,$$

where Cov (X, Y) denotes the covariance of X and Y, and the second term of (1) is equal to $(k - c)^2$. Furthermore,

$$(2)\quad (1/\sigma^2) \text{Var}(\bar{\Psi}) = (1/\sigma^2)[a^2 \text{Var}(\hat{\Phi}) + 2ak \text{Cov}(\hat{\Phi}, \hat{\sigma}) + k^2 \text{Var}(\hat{\sigma})]$$

must be independent of $\boldsymbol{\theta}$ for all $\boldsymbol{\theta}$.

[2] In the discussion that follows, "unique" will always be taken to mean unique with probability one.

[3] The class ξ may consist, for example, of all linear combinations of sample observations or all possible estimators based on sample observations.

Now, one may make use of the fact that the absolute value of the correlation coefficient ρ of $\hat{\Phi}$ and $\hat{\sigma}$ is less than or equal to 1 to demonstrate that $(1/\sigma^2)$ Var $(\hat{\Phi})$, $(1/\sigma^2)$ Var $(\hat{\sigma})$, and $(1/\sigma^2)$ Cov $(\hat{\Phi}, \hat{\sigma})$ must each be independent of $\boldsymbol{\theta}$ if (2) is to be independent of $\boldsymbol{\theta}$. Let Var $(\hat{\Phi}) = \alpha\sigma^2$, Var $(\hat{\sigma}) = \gamma\sigma^2$, and Cov $(\hat{\Phi}, \hat{\sigma}) = \beta\sigma^2$, where α and γ are both positive. Then, since $\rho^2 = \beta^2/(\alpha\gamma)$ and $\rho^2 \leqq 1$, it follows that $\beta^2 \leqq \alpha\gamma$ (since $\alpha > 0$ and $\gamma > 0$) and

$$(3) \qquad (1/\sigma^2) \text{ Var } (\bar{\Psi}) \leqq a^2\alpha + |2ak(\alpha\gamma)^{\frac{1}{2}}| + k^2\gamma.$$

Since Var $(\bar{\Psi})$ is not identically zero unless $\bar{\Psi} = a\hat{\Phi} + k\hat{\sigma}$ is identically zero, and since the left side of (3) is independent of $\boldsymbol{\theta}$ for all $\boldsymbol{\theta}$, it must be true that the right side is also independent of $\boldsymbol{\theta}$ for all $\boldsymbol{\theta}$ for a and k not both zero. Also, each term on the right side of (3) is nonnegative. Therefore, for a and k not both zero, each term must be independent of $\boldsymbol{\theta}$, and α, γ, and hence β must also be independent of $\boldsymbol{\theta}$, for all $\boldsymbol{\theta}$.

In order to demonstrate that for $\tilde{\Psi} = a\Phi + c\tilde{\sigma}$, a and k are not both zero, it is necessary to determine the form of $R(\Psi, \Psi')$, where $\Psi' = a\Phi' + c\sigma'$ has minimum risk among invariant estimators of Ψ in ξ based on a given combination of $\hat{\Phi}$ and $\hat{\sigma}$. This is accomplished by minimizing k, which is equal to $(c - a\beta)/(1 + \gamma)$, which is not zero unless $a = c/\beta$. Hence, $\bar{\Psi} = a\hat{\Phi} + k\hat{\sigma}$ is not identically zero unless a and c are both zero, contrary to assumption. Hence α, β, and γ are each independent of $\boldsymbol{\theta}$ for all $\boldsymbol{\theta}$ and Ψ' is equal to $a\hat{\Phi} + [(c - a\beta)/(1 + \gamma)]\hat{\sigma} \equiv a[\hat{\Phi} - \beta\hat{\sigma}/(1 + \gamma)] + c\hat{\sigma}/(1 + \gamma) \equiv a\Phi' + c\sigma'$. $R(\Psi, \Psi')$, with $\Psi' = a\Phi' + c\sigma'$ is then equal to

$$(4) \qquad a^2\alpha + 2ac\beta + c^2\gamma - (c\gamma + a\beta)^2/(1 + \gamma)$$

or

$$a^2[\alpha - \beta^2/(1 + \gamma)] + 2ac\beta/(1 + \gamma) + c^2\gamma/(1 + \gamma).$$

It can now be established that $(\hat{\Phi}, \hat{\sigma})$ must be the unique best unbiased estimator of (Φ', σ), in the class ξ in order that (Φ', σ') be equal to $(\tilde{\Phi}, \tilde{\sigma})$, the unique best invariant estimator of (Φ, σ) in ξ. Let a equal 1 and c equal 0 so that $\Psi = \Phi$ and $\Psi' = \Phi'$. The change in $R(\Phi, \Phi')$ induced by changes in α, β, and γ corresponding to change in $\hat{\Phi}$ and $\hat{\sigma}$ is now considered.

The change $dR(\Phi, \Phi')$ in $R(\Phi, \Phi')$ is equal to

$$[\partial R(\Phi, \Phi')/\partial\alpha]\, d\alpha + [\partial R(\Phi, \Phi')/\partial\beta]\, d\beta + [\partial R(\Phi, \Phi')/\partial\gamma]\, d\gamma$$

and from (4), $\partial R(\Phi, \Phi')/\partial\alpha = 1$, $\partial R(\Phi, \Phi')/\partial\beta = -2\beta/(1 + \gamma)$, and $\partial R(\Phi, \Phi')/\partial\gamma = \beta^2/(1 + \gamma)^2$. Clearly, for fixed β and increasing α and γ, $dR(\Phi, \Phi')$ is positive and $R(\Phi, \tilde{\Phi})$ is increasing. It can be shown also, by once again using the fact the $\beta^2 \leqq \alpha\gamma$, that $dR(\Phi, \Phi')$ is positive for positive $d\alpha$ and $d\gamma$ regardless of the value of $d\beta$. Since $\beta^2 \leqq \alpha\gamma$, $d\beta \leqq |(\gamma/2\beta)\, d\alpha|$ and $|(\partial R(\Phi, \Phi')/\partial\beta)\, d\beta| = |[-2\beta/(1 + \gamma)]\, d\beta| \leqq |[\gamma/(1 + \gamma)]\, d\alpha|$. Then $dR(\Phi, \Phi') \geqq d\alpha - [\gamma/(1 + \gamma)]|d\alpha| + [\beta^2/(1 + \gamma)^2]\, d\gamma$ or

$$(5) \qquad dR(\Phi, \Phi') \geqq [1 - \gamma/(1 + \gamma)]\, d\alpha + [\beta^2/(1 + \gamma)]\, d\gamma$$

for $d\alpha$ positive. Note that the right side of (5) is positive, and hence $dR(\Phi, \Phi')$ is positive, for $d\alpha > 0$ and $d\gamma > 0$. Also, if $a = 0$ and $c = 1$ so that $\Psi = \sigma$ and $\Psi' = \sigma'$, then $\partial R(\sigma, \sigma')/\partial\beta = \partial R(\sigma, \sigma')/\partial\alpha = 0$ and $dR(\sigma, \sigma') = [\partial R(\sigma, \sigma')/\partial\gamma]\, d\gamma = d\gamma/(1 + \gamma)^2$, which is positive for $d\gamma$ positive.

Therefore, for the class ξ of estimators for which a vector (Φ^*, σ^*) of unique best unbiased estimators of (Φ, σ) exists, $(\tilde{\Phi}, \tilde{\sigma}) = (\Phi^* - B\sigma^*/(1 + C), \sigma^*/(1 + C))$ is the unique best invariant estimator of (Φ, σ), where $B\sigma^2$ is the covariance between Φ^* and σ^* and $C\sigma^2$ is the variance of σ^*. The uniqueness of $(\tilde{\Phi}, \tilde{\sigma})$ follows from the uniqueness of (Φ^*, σ^*). The expressions for the mean squared errors of $\tilde{\Phi}$ and $\tilde{\sigma}$ can be calculated directly from (4) and shown to be equal to $[A - B^2/(1 + C)]\sigma^2$ and $C\sigma^2/(1 + C)$, respectively, where $A\sigma^2$ is the variance of Φ^*. In the same manner it can be shown that $E[\tilde{\Phi} - \Phi)(\tilde{\sigma} - \sigma)]$ is equal to $B\sigma^2/(1 + C)$. The proof of Theorem 1 is now complete.

Implications of Theorem 1. If $p = 1$, $l = l_1 = 1$, and $m = 0$ then $\Phi = \mu$ and $\Phi^* = \mu^*$. In the case of a single unknown location parameter, $\boldsymbol{Q}$ is a column vector of 1's and $\boldsymbol{x}$ may be a vector of ordered observations. Note that if $\boldsymbol{y} = (\boldsymbol{x} - \mu)/\sigma$, then $E(\boldsymbol{x}) = \mu + \sigma E(\boldsymbol{y})$. For this model Φ may also be simply σ or the 100 P percent point x_P of the distribution of the random variable X when $\boldsymbol{l} = l_1 = 1$, $\boldsymbol{\mu} = \mu$, and $m = y_P$, the corresponding percentile of the distribution of the reduced parameter-free variate $Y = (X - \mu)/\sigma$. Thus, x_P is given by $\mu + y_P\sigma$, where y_P can usually be calculated and for some distributions is tabulated. If

$$\sigma^2\begin{bmatrix} \alpha & \beta \\ \beta & \gamma \end{bmatrix}$$

is the covariance matrix of (μ^*, σ^*), then the variance of $X_P^* = \mu^* + y_P\sigma^*$ is $[\alpha + 2y_P\beta + y_P^2\gamma]\sigma^2$, and the mean squared error of the best invariant estimator of x_P in ξ is $[\alpha + 2y_P\beta + y_P^2\gamma]\sigma^2 - [(y_P\gamma + \beta)^2/(1 + \gamma)]\sigma^2$ if X_P^* is the best unbiased estimator of x_P in ξ.

The covariance between Φ^* and σ^* will not in general be zero. An exception is, of course, the case in which Φ and σ are respectively equal to μ and σ, the mean and standard deviation of a Gaussian distribution, and there is no censoring of the sample (or location and scale parameters of any distribution with density symmetric about μ). For an uncensored sample from a Gaussian distribution, best unbiased joint estimators μ^* and σ^* exist for μ and σ, respectively, and have covariance zero. If, however, $\Phi = \mu + y_P\sigma$, where y_P is a standard normal deviate and Φ is the 100 P percent point of a normal distribution with mean μ and standard deviation σ, then for $P \neq .50$, the covariance between Φ^* and σ^*, the joint best unbiased estimators of Φ and σ, respectively, is not zero.

Suppose p is equal to 1 and (μ, σ) is a general location-scale parameter, as defined in Assumption 1. Suppose either that the distribution of X is such that no complete sufficient statistics exist or that absence of complete sufficient statistics is due to censoring of the sample. In either case no best unbiased estimator of (μ, σ) exists. Let (μ^*, σ^*) be the unique vector of best linear unbiased estimators of (μ, σ) as specified by the generalized Gauss-Markov Theorem (see Lloyd [6]). Then the unique best linear invariant estimator of $x_P = \mu + y_P\sigma$, the 100 P percent point of X, is $\mu^* + (y_P - \beta)\sigma^*/(1 + \gamma)$ where β is the covariance of μ^* and σ^* and γ is the variance of σ^*.

Suppose that $\Phi = \boldsymbol{l}'\boldsymbol{\mu}$ where $\boldsymbol{Q}$ is $p \times n$. Then if $\boldsymbol{e} = (\boldsymbol{x} - \boldsymbol{Q\mu})/\sigma$, $E(\boldsymbol{x}) = \boldsymbol{Q\mu} + \sigma E(\boldsymbol{e})$ and if $E(\boldsymbol{e}) = 0$ and Φ is estimable, as discussed by Scheffé [9], the ordinary Gauss-Markov Theorem applies (since, by Assumption 1, $E(\boldsymbol{ee}')$ is independent of σ^2). We therefore let $\boldsymbol{\mu}^*$ be the vector of unique Gauss-Markov estimators of the elements of $\boldsymbol{\mu}$ and σ^* be $k[\sum_{i=1}^n (x_i - \sum_{j=1}^p q_{ji}\mu_j^*)^2]^{\frac{1}{2}} = k(s^2)^{\frac{1}{2}}$, where the optimality properties of s^2 have been demonstrated by Hsu [4].

Let $\boldsymbol{A}\sigma^2$ be the $p \times p$ covariance matrix of $\boldsymbol{\mu}^*$, $\boldsymbol{b}\sigma^2$ the $p \times 1$ covariance matrix of $\boldsymbol{\mu}^*$ and σ^* and $C\sigma^2$ the variance of σ^*. Then, by Theorem 1, the estimator $\tilde{\Phi}$ given by

$$\boldsymbol{l}'\boldsymbol{\mu}^* - \boldsymbol{l}'\boldsymbol{b}\sigma^*/(1 + C) \equiv \boldsymbol{l}'\tilde{\boldsymbol{\mu}}$$

has smaller mean squared error than $\boldsymbol{l}'\boldsymbol{\mu}^*$. The mean squared error of $\tilde{\Phi}$ is

$$[\boldsymbol{l}'\boldsymbol{A}\boldsymbol{l} - \boldsymbol{l}'\boldsymbol{bb}'\boldsymbol{l}/(1 + C)]\sigma^2.$$

If X is from a Gaussian or any other distribution with density symmetric about $\boldsymbol{\mu}$, $\boldsymbol{b}$ is equal to $\boldsymbol{0}$.

The theorem thus extends both the Gauss-Markov Theorem and the Lehmann-Scheffé Theorem, as given in [1] and [5], respectively, indicating how unique estimators with uniformly smaller risk may be obtained from those specified by these classical theorems. Theorem 2 and Theorem 3, which follow and which also extend these two well-known theorems, apply to cases in which either $\boldsymbol{\mu}$ or σ is known.

3. Examples of use of Theorem 1. As an illustration of the manner in which Theorem 1 can be used to obtain estimates which will, on the average, lie closer to the parameters estimated than those obtained by using best unbiased estimators, we consider the exponential distribution with location parameter θ and scale parameter σ. The best unbiased estimators of the two parameters and the variances and covariances of these estimators are, for a complete sample:

$$\sigma^* = (n-1)^{-1}n(\bar{x} - x_1), \qquad V(\sigma^*) = (n-1)^{-1}\sigma^2,$$
$$\theta^* = (n-1)^{-1}(nx_1 - \bar{x}), \qquad V(\theta^*) = n^{-1}(n-1)^{-1}\sigma^2,$$
$$\text{Cov}\ (\theta^*, \sigma^*) = n^{-1}(n-1)^{-1}\sigma^2.$$

These estimators are inadmissible; uniformly better estimators are given below with their risk functions.

$$\tilde{\sigma} = n^{-1}(n-1)(n-1)^{-1}n(\bar{x} - x_1) = \bar{x} - x_1, \qquad R(\sigma, \tilde{\sigma}) = n^{-1}\sigma^2,$$
$$\tilde{\theta} = (n-1)^{-1}(nx_1 - \bar{x}) + n^{-1}(n-1)^{-1}(\bar{x} - x_1) = n^{-1}((n+1)x_1 - \bar{x}),$$
$$R(\theta, \tilde{\theta}) = n^{-1}(n-1)^{-1}\sigma^2 - [-n(n-1)]^{-2}n^{-1}(n-1) = n^{-3}(n+1)\sigma^2.$$

Suppose that one wishes to estimate the 90% point of such an exponential population. Then, since $P = \Phi(u) = 1 - e^{-u_P}$, $u_P = \ln[1/(1 - P)] = \ln 10$. If the sample size is 4, then

$$R(x_{.90}, \tilde{X}_{.90}) = .08\bar{3}\sigma^2 + [2(-.08\bar{3})\ln 10 + .33\bar{3}(\ln 10)^2 - (.08\bar{3})^2].75\sigma^2 = 1.12\sigma^2.$$

The risk of $X^*_{.90} = \mu^* + (\ln 10)\sigma^*$ is $1.47\sigma^2$, which is more than 30% greater than that of $\tilde{X}_{.90}$.

In [8] the near optimality of best linear invariant and maximum-likelihood estimators with respect to known estimators of the parameters of the first asymptotic distribution of smallest (extreme) values is demonstrated.

4. Best invariant estimators of σ for μ known. In order to derive the best invariant estimator of σ when the vector $\boldsymbol{\mu}$ is known the following assumption is made.

A2$'$. A unique best unbiased estimator $\sigma_\mu{}^*$ with variance $C_\mu\sigma^2$ exists in ξ for σ.

Theorem 2 below holds under Assumptions 1 and 2$'$ when the loss function is squared error divided by σ^2, as defined earlier. The term "invariant" in Theorem 2 and those following should be taken to mean invariant under transformations of location and scale.

THEOREM 2. *For $\boldsymbol{\mu}$ known and σ unknown, the unique minimum-risk invariant estimator of σ in ξ is $\tilde{\sigma}_\mu = \sigma_\mu{}^*/(1 + C_\mu)$, with expected squared error equal to $[C_\mu/(1 + C_\mu)]\sigma^2$.*

PROOF OF THEOREM 2. The proof of Theorem 2 can take the form of that of Theorem 1 with $\boldsymbol{l} = \mathbf{0}$, $m = 1$, and $\Phi = \sigma$. In such a case, $\sigma_\mu{}^*$ is substituted for both σ^* and Φ^* and C_μ is substituted for A, B, and C. Then $A - B^2/(1 + C)$, $B/(1 + C)$, and $C/(1 + C)$ are replaced in the proof by $C_\mu{}^*/(1 + C_\mu)$.

5. Best invariant estimation of Φ for σ known. Theorem 3 applies to the case where σ is known, but $\boldsymbol{\mu}$ is unknown. It holds under Assumptions 1 and 2.

THEOREM 3. *For σ known, $\boldsymbol{\mu}$ unknown, and $\boldsymbol{l} \neq \mathbf{0}$, the unique minimum-risk invariant estimator of $\Phi = \boldsymbol{l}'\boldsymbol{\mu} + m\sigma$ in ξ is $\tilde{\Phi}_\sigma = \Phi^* - (B/C)(\sigma^* - \sigma)$, which is unbiased. The variance of $\tilde{\Phi}_\sigma$ is $(A - B^2/C)\sigma^2$.*

PROOF OF THEOREM 3. The form of the proof of Theorem 3 is similar to that of Theorem 1. For σ known, an estimator $\bar{\Phi}_\sigma$ of Φ in ξ, in order to be independent of $\boldsymbol{\theta}$ for all $\boldsymbol{\theta}$, must be of the form $\boldsymbol{l}'\hat{\boldsymbol{\mu}} + m\hat{\sigma} + (k_1 - m)\hat{\sigma} + (k_2 + m)\sigma \equiv \hat{\Phi} + (k_1 - m)\hat{\sigma} + (k_2 + m)\sigma$, where $\hat{\mu}$ is a vector of unbiased estimators of the respective elements of $\boldsymbol{\mu}$ and $(\hat{\Phi}, \hat{\sigma})$ is a vector of unbiased estimators of (Φ, σ) for σ not known. Again, as in the proof of Theorem 1, it can be demonstrated that for $\bar{\Phi}_\sigma$ any invariant estimator of Φ, the variance of $\bar{\Phi}_\sigma$ is a function of $\alpha\sigma^2$, $\beta\sigma^2$, and $\gamma\sigma^2$, with α, β, and γ independent of $\boldsymbol{\theta}$ for all $\boldsymbol{\theta}$, and $\alpha\sigma^2$ and $\gamma\sigma^2$ the variances of $\hat{\Phi}$ and $\hat{\sigma}$, respectively, and $\beta\sigma^2$ their covariance. $R(\Phi, \bar{\Phi}_\sigma)$ is thus of the form

$$\alpha + 2(k_1 - m)\beta + (k_1 - m)^2\gamma + (k_1 + k_2)^2.$$

Minimizing $R(\Phi, \bar{\Phi}_\sigma)$ with respect to k_1 and k_2, one obtains $k_2 = -k_1 = \beta/\gamma - m$. Thus Φ_σ', the invariant estimator of Φ with minimum risk among those based on $\hat{\Phi}$ and $\hat{\sigma}$, is given by $\Phi_\sigma' = \hat{\Phi} - (\beta/\gamma)(\hat{\sigma} - \sigma)$ with $R(\Phi, \Phi_\sigma')$ equal to $\alpha - \beta^2/\gamma$. Then since $\partial R(\Phi, \Phi_\sigma')/\partial\alpha$ is equal to 1, $dR(\hat{\Phi}, \Phi_\sigma')\partial\beta$ is equal to $-2\beta/\gamma$, and $\partial R(\Phi, \Phi_\sigma')/\partial\gamma$ is equal to β^2/γ^2, $\partial R(\Phi, \Phi_\sigma') = d\alpha - 2\beta/\gamma\, d\beta + \beta^2/\gamma^2 d\gamma$.

Making use of the fact that $\beta^2 \leqq \alpha\gamma$, one can then demonstrate that

$$dR(\Phi, \Phi_\sigma') \geqq d\alpha - |d\alpha| + \beta^2/\gamma^2\, d\gamma$$

and that $R(\Phi, \Phi_\sigma')$ is increasing for $d\alpha$ and $d\gamma$ both positive. Thus if Φ^* and σ^* are the unique best unbiased joint estimators in ξ of Φ and σ, respectively, when σ is not known,

$$\tilde{\Phi}_\sigma = \Phi^* - (B/C)(\sigma^* - \sigma),$$

with variance $[A - (B^2/C)]\sigma^2$, is the unique minimum-risk invariant estimator of Φ when σ is known. This completes the proof of Theorem 3.

Implications of Theorem 3. Theorem 3 is of practical interest when a distribution scale parameter is known, there is censoring of a sample, and coefficients for obtaining best linear unbiased (or best linear invariant) estimates of both the distribution location and scale parameters are available. One can then convert these coefficients to new coefficients from which one can obtain the best linear invariant estimate of the location parameter for a known scale parameter.

REFERENCES

[1] Gauss, K. F. (1809). *Theoria Motus Corporum Coelestium In Sectionibus Cononicus Solem Ambientium.* Perthes and Besser, Hamburg.

[2] Goodman, Leo A. (1953). A simple method for improving some estimators. *Ann. Math. Statist.* **24** 114–117.

[3] Goodman, Leo. (1960). A note on the estimation of variance. *Sankhyā Ser. A* **22** 221.

[4] Hsu, P. L. (1938). On the best quadratic estimate of the variance. *Statistical Research Memoirs,* **2** 91–104.

[5] Lehmann, E. L. and Scheffé, H. (1950). Completeness, similar regions and unbiased estimation. *Sankhyā* **10** 305–340.

[6] Lloyd, E. H. (1952). Least-squares estimation of location and scale parameters using order statistics. *Biometrika* **39** 88–95.

[7] Mann, N. R. (1965). Point and interval estimates for reliability parameters when failure times have the two-parameter Weibull distribution. Ph.D. dissertation. Univ. of California at Los Angeles.

[8] Mann, Nancy R. (1968). Point and interval estimation procedures for the two-parameter Weibull and extreme-value distributions. *Technometrics* **10** 231–256.

[9] Scheffé, Henry (1959). *The Analysis of Variance.* Wiley, New York.