The Annals of Mathematical Statistics
1970, Vol. 41, No. 1, 59–69

BAYES AND FIDUCIAL EQUIVARIANT ESTIMATORS OF THE COMMON MEAN OF TWO NORMAL DISTRIBUTIONS[1]

BY S. ZACKS

University of New Mexico

1. Introduction. The problem of pooling observations which are distributed according to different distribution laws, for the purpose of estimating a common parameter has been studied in many papers. The more specific problem of estimating the common mean of two normal distributions, with an emphasis on small sample estimators, was previously studied by the author [9], and more recently by Gurland and Mehta [4]. The reader can find in these articles relevant reference lists. Both Zacks, and Gurland and Mehta, studied procedures of combining the sample mean, by some weighted average estimators of the form:

$$\tilde{\mu} = \bar{X}\phi(S_2/S_1) + \bar{Y}(1-\phi(S_2/S_1)), \tag{1.1}$$

where $\bar{X}$ and $\bar{Y}$ are the sample means; S_1 and S_2 are the sample sum of squares of deviations, respectively. These authors confined their attention to the cases of equal sample sizes. We notice that for all choices of weighing functions $\phi(S_2/S_1)$ the above estimators are *unbiased* and invariant with respect to translation and change of scale. Gurland and Mehta showed in [4] that if it is known which one of the two distributions has the smaller variance, although the actual variance ratio is unknown, certain of the estimators suggested by Zacks in [9] can be improved upon uniformly. In the present study we investigate the whole problem more systematically in a decision theoretic framework. We start by characterizing the class of all estimators of the common mean, μ, which are translation invariant and scale preserving. Following Wijsmann [8] and Berk [1], we call these estimators *equivariant estimators*. The sample variance ratio S_2/S_1 is not the maximal invariant statistic for the group of translations and change of scale. Thus, the class of all estimators of the form (1.1) is only a subclass of the class of all equivariant estimators, and many of the estimators of the form (1.1) discussed in the previous studies are inadmissible even among the equivariant estimators. The general form of all equivariant estimators is:

$$\hat{\mu} = \bar{X} + (\bar{Y}-\bar{X})\psi(S_1(\bar{X}-\bar{Y})^{-2}, S_2(\bar{Y}-\bar{X})^{-2}). \tag{1.2}$$

The problem of choosing an equivariant estimator is equivalent to the problem of choosing a (properly measurable) function $\psi(\cdot,\cdot)$ of the maximal invariant $(S_1(\bar{Y}-\bar{X})^{-2}, S_2(\bar{Y}-\bar{X})^{-2})$. Adopting a quadratic loss function, which yields a risk function proportional to the mean-square-error risk function, in Section 3 we determine the class of all *Bayes equivariant* estimators. These are the equivariant estimators of the form (1.2), which minimize the prior risk functions corresponding

Received December 21, 1968.

[1] Supported by the National Science Foundation, Grant GP-9007.

to proper prior distributions of the parameters of the model. We shall also verify that all the locally minimum variance unbiased estimators are particular Bayes equivariant estimators; namely:

$$\tilde{\mu}(\rho_0) = \rho_0(1+\rho_0)^{-1}\bar{X}+(1+\rho_0)^{-1}\bar{Y}, \tag{1.3}$$

where $0 < \rho_0 < \infty$ is a chosen constant. We have $\tilde{\mu}(\rho_0)$ as a trivial Bayes equivariant estimator for which the prior distribution of the variance ratio $\rho = {\sigma_2}^2/{\sigma_1}^2$ is concentrated on the single point, ρ_0. Available examples of non-trivial Bayes equivariant estimators yield quite complicated *weighing functions*: $\psi(V_1, V_2)$; $V_i = S_i(\bar{X}-\bar{Y})^{-2}$; $i = 1, 2$. It is therefore doubtful whether estimators of the form (1.1) are among the Bayes equivariant estimators.

In Section 4 we consider the class of quasi-Bayes estimators of the common mean μ, for which the variance ratio ρ is assigned a prior distribution $H(\rho)$, independent of μ and σ^2; while μ and σ^2 are assigned the Jeffrey's improper prior $K(d\mu, d\sigma^2) = d\mu \cdot d\sigma^2$. These estimators of μ are called *fiducial estimators*. It is shown that all such fiducial estimators are equivariant. The unique (up to an additive constant) invariant Haar measure for the group of transformations we consider is $K^*(d\mu, d\sigma^2) = d\mu \cdot d\sigma^2/\sigma^3$ (see Buehler and Hora [5] and Stein [7]). However, we also obtain equivariant fiducial estimators with the Jeffrey's improper prior $K(d\mu, d\sigma^2)$. This is due to the special structure of the estimator. As shown in Section 4, the fiducial estimators are generally different from the Bayes equivariant estimators, although there is an interesting relationship between the two estimators. In Section 5 we apply a criterion given by Karlin in [6] to show that the Bayes equivariant and the fiducial estimators are weakly admissible when the prior distribution $H(\rho)$ is absolutely continuous and has a positive density. It is conjectured that this result can be improved, and all Bayes equivariant and all fiducial estimators can be shown to be admissible (possibly by using the criterion given by Stein in [7]).

2. The statistical model and the equivariant estimators. Let $X_1, \cdots, X_n$; $Y_1 \cdots, Y_n$ be independent random variables. Assume that $X_i \sim N(\mu, \sigma^2)$ $(i = 1, \cdots, n)$, and $Y_i \sim N(\mu, \rho\sigma^2)$ $(i = 1, \cdots, n)$. The common mean μ is unknown, $-\infty < \mu < \infty$. The variance ratio ρ is also unknown. We assume that $0 < \sigma^2 < \infty; 0 < \rho < \infty$. The parameter space Θ is the set of all $\theta = (\mu, \sigma^2, \rho)$. It is well known that a minimal sufficient statistic for the present model is $(\bar{X}, \bar{Y}, S_1, S_2)$, where $\bar{X} = \sum_{i=1}^n X_i/n$, $\bar{Y} = \sum_{i=1}^n Y_i/n$, $S_1 = \sum_{i=1}^n (X_i - \bar{X})^2$ and $S_2 = \sum_{i=1}^n (Y_i - \bar{Y})^2$. This minimal sufficient statistic is, however, incomplete. We further assume that each sample is of size $n \geqq 2$, since if $n = 1$ then the minimal sufficient statistic reduces to (X, Y), and the problem of estimating μ becomes relatively simple. We study in the present paper estimators of the common mean; μ, which are translation invariant and scale preserving. The loss function adopted is a quadratic loss function. We therefore consider only estimators which are functions $\hat{\mu}(\bar{X}, \bar{Y}, S_1, S_2)$ of the minimal sufficient statistic.

Consider the group of transformations

$$\mathscr{G} = \{g_{\alpha\beta}(X) = \alpha + \beta X, \quad -\infty < \alpha < \infty, \quad \beta \neq 0\}. \tag{2.1}$$

An estimator $\hat{\mu}(\bar{X}, \bar{Y}, S_1, S_2)$ is called *equivariant* with respect to $\mathscr{G}$ if, and only if, for every $g_{\alpha\beta} \in \mathscr{G}$

$$\hat{\mu}(\bar{g}_{\alpha\beta}\bar{X}, \bar{g}_{\alpha\beta}\bar{Y}, \bar{g}_{\alpha\beta}S_1, \bar{g}_{\alpha\beta}S_2) = \alpha + \beta\hat{\mu}(\bar{X}, \bar{Y}, S_1, S_2), \tag{2.2}$$

where $\bar{g}_{\alpha\beta}$ is the transformation of $(\bar{X}, \bar{Y}, S_1, S_2)$ corresponding to $g_{\alpha\beta}$. It is immediate then to verify that every equivariant estimator of μ, based on the minimal sufficient statistic, is of the general form:

$$\hat{\mu}(\bar{X}, \bar{Y}, S_1, S_2) = \bar{X} + (\bar{Y} - \bar{X})\psi(S_1(\bar{Y} - \bar{X})^{-2}, S_2(\bar{Y} - \bar{X})^{-2}), \tag{2.3}$$

where $(S_1(\bar{Y} - \bar{X})^{-2}, S_2(\bar{Y} - \bar{X})^{-2})$ is the maximal invariant function of the sufficient statistic.

THEOREM 2.1. *Every sufficient-equivariant estimator of the form* (2.3) *is an unbiased estimator of* μ.

PROOF.
Since $E_\theta\{\bar{X}\} = \mu$ for all θ, we have to show that

$$E_\theta\{(\bar{Y} - \bar{X})\psi(S_1(\bar{Y} - \bar{X})^{-2}, S_2(\bar{Y} - \bar{X})^{-2})\} = 0, \quad \text{for all} \quad \theta. \tag{2.4}$$

But (2.4) follows immediately from the independence of $\bar{X}$, $\bar{Y}$, S_1, S_2. Indeed,

$$\begin{aligned} &E_\theta\{(\bar{Y} - \bar{X})\psi(S_1(\bar{Y} - \bar{X})^{-2}, S_2(Y - X)^{-2})\} \\ &\quad = E_\theta\{\psi(S_1(\bar{Y} - \bar{X})^{-2}, S_2(\bar{Y} - \bar{X})^{-2})E_\theta\{\bar{Y} - \bar{X} \,\big|\, |\bar{Y} - \bar{X}|\}\}. \end{aligned} \tag{2.5}$$

Due to the symmetry of the distribution of $\bar{X} - \bar{Y}$, $E_\theta\{\bar{Y} - \bar{X} \,\big|\, |\bar{Y} - \bar{X}|\} = 0$ a.s., for all θ. Substituting this in the R.H.S. of (2.5) we obtain (2.4).

THEOREM 2.2. *The variance of an equivariant estimator of the form* (2.3) *is:*

$$\begin{aligned} D^2(\psi; \sigma^2, \rho) = {} & n^{-1}\sigma^2 + 2\sigma^2(1 - (2n)^{-1}) \cdot E_\rho\{(\psi^2(V_1, V_2) \\ & - 2(1+\rho)^{-1}\psi(V_1, V_2))(1+\rho)(1 + n^{-1}(1+\rho)V_1 + \rho^{-1}n^{-1}(1+\rho)V_2)^{-1}\}, \end{aligned} \tag{2.6}$$

where $V_i = S_i/(\bar{Y} - \bar{X})^2$, $i = 1, 2$.

PROOF. The variance of an equivariant estimator is:

$$\begin{aligned} \operatorname{Var}\{\bar{X} + (\bar{Y} - \bar{X})\psi(V_1, V_2)\} = {} & n^{-1}\sigma^2 + 2E_{(\sigma^2,\rho)}\{(\bar{X} - \mu)(\bar{Y} - \bar{X})\psi(V_1, V_2)\} \\ & + E_{(\sigma^2,\rho)}\{(\bar{Y} - \bar{X})^2\psi^2(V_1, V_2)\}. \end{aligned} \tag{2.7}$$

To develop this variance formula we notice that, since

$$\begin{pmatrix} \bar{X} - \mu \\ \bar{Y} - \bar{X} \end{pmatrix} \sim N\left(\mathbf{0}, n^{-1}\sigma^2 \begin{bmatrix} 1 & -1 \\ -1 & 1+\rho \end{bmatrix}\right),$$

the conditional distribution of $\bar{X} - \mu$, given $\bar{X} - \bar{Y}$, is the normal distribution $N(-(1+\rho)^{-1}(\bar{Y} - \bar{X}), n^{-1}\sigma^2 \cdot \rho(1+\rho)^{-1})$. Hence,

$$\begin{aligned} &E_{(\sigma^2,\rho)}\{(\bar{X} - \mu)(\bar{Y} - \bar{X})\psi(V_1, V_2) \mid \bar{Y} - \bar{X}, S_1, S_2\} \\ &\qquad = -(1+\rho)^{-1}(\bar{Y} - \bar{X})^2\psi(V_1, V_2) \quad \text{a.s.} \end{aligned} \tag{2.8}$$

Routine manipulations yield that the conditional distribution of $W=(\bar{Y}-\bar{X})^2$, given (V_1, V_2), is like that of a gamma random variable $G(2\sigma^{-2}(n(1+\rho)^{-1}+V_1+\rho^{-1}V_2, n-\frac{1}{2})$; where $G(\lambda, p)$ has the density function

$$g(x \mid \lambda, p) = (\Gamma(p))^{-1}\lambda^p x^{p-1} e^{-\lambda x}, \qquad 0 \leqq x \leqq \infty,$$

and $0 < \lambda < \infty$, $0 < p < \infty$. Thus, we deduce that

$$(2.9) \qquad E_{(\sigma^2,\rho)}\{W \mid V_1, V_2\} = 2(n-\tfrac{1}{2})\sigma^2(n(1+\rho)^{-1}+V_1+\rho^{-1}V_2)^{-1}, \qquad \text{a.s.}$$

Substituting (2.8) and (2.9) in (2.7) we obtain (2.6).

3. Bayes equivariant estimators. We shall derive in the present section the Bayes equivariant estimators for the quadratic loss function:

$$(3.1) \qquad L(\hat{\mu}, \mu) = (\hat{\mu}-\mu)^2/\sigma^2.$$

According to Theorems 2.1 and 2.2 the risk function associated with an equivariant estimator of the form (2.3) is:

$$(3.2) \qquad \begin{aligned} R(\psi; \rho) = n^{-1} + 2(1-(2n)^{-1})E_\rho\{[\psi^2(V_1, V_2) - 2(1+\rho)^{-1}\psi(V_1, V_2)] \\ \cdot [(1+\rho)^{-1} + n^{-1}V_1 + (n\rho)^{-1}V_2]^{-1}\}. \end{aligned}$$

We notice that the expectation on the R.H.S. of (3.2) is a function of ρ only, since the joint distribution of (V_1, V_2) depends only on the variance-ratio ρ. Hence, the prior risk associated with any equivariant estimator depends only on the assumed prior distribution of the nuisance parameter ρ. This result fits well into the well-known fact that, when the nuisance parameter ρ is known, there exists an essentially unique uniformly best equivariant estimator for which $\psi(V_1, V_2) = (1+\rho)^{-1}$ a.s. This estimator is also the essentially unique uniformly best unbiased estimator. When ρ is unknown there exists no uniformly best equivariant estimator, but the Bayes equivariant estimator depends only on the assumed prior distribution of ρ. To be more specific, given any prior distribution of ρ, $H(\rho)$ say, the prior risk associated with an equivariant estimator is defined as:

$$(3.3) \qquad R(\psi; H) = \int_0^\infty R(\psi; \rho)H(d\rho).$$

An equivariant estimator μ_H is called Bayes against $H(\rho)$ if it minimizes the prior risk function $R(\psi; H)$. To obtain this Bayes equivariant estimator one has to minimize the function

$$(3.4) \qquad \begin{aligned} Q(\psi, H) = \int_0^\infty H(d\rho)E_\rho\{[\psi^2(V_1, V_2) - 2(1+\rho)^{-1}\psi(V_1, V_2)] \\ \cdot [(1+\rho)^{-1} + n^{-1}V_1 + (n\rho)^{-1}V_2]^{-1}\}. \end{aligned}$$

Let $H(\rho \mid V_1, V_2)$ designate the posterior distribution of ρ, given (V_1, V_2), and let $p(V_1, V_2 \mid H)$ designate the mixture of the densities of (V_1, V_2) given ρ, with respect to $H(\rho)$. We further notice that for each (V_1, V_2),

$$(3.5) \qquad \begin{aligned} S_H(V_1, V_2; \psi) = \int_0^\infty |\psi^2(V_1, V_2) - 2(1+\rho)^{-1}\psi(V_1, V_2)| \\ \cdot [(1+\rho)^{-1} + n^{-1}V_1 + (n\rho)^{-1}V_2]^{-1} H(d\rho \mid V_1, V_2) < \infty \qquad \text{a.s.} \end{aligned}$$

and is a measurable function of (V_1, V_2). Thus, if the function $\psi(V_1, V_2)$ is such that $S_H(V_1, V_2; \psi)$ is integrable with respect to the mixed distribution of (V_1, V_2), i.e., if

$$\int_0^\infty \int_0^\infty S_H(V_1, V_2; \psi) p(V_1, V_2 \mid H) \, dV_1 \, dV_2 < \infty \tag{3.6}$$

then, by the Fubini theorem, we can write:

$$\begin{aligned} Q(\psi, H) = \int_0^\infty \int_0^\infty p(V_1, V_2 \mid H) \, dV_1 \, dV_2 \int_0^\infty H(d\rho \mid V_1, V_2)[\psi^2(V_1, V_2) \\ -2(1+\rho)^{-1}\psi(V_1, V_2)] \cdot [(1+\rho)^{-1} + n^{-1}V_1 + (n\rho)^{-1}V_2]^{-1}. \end{aligned} \tag{3.7}$$

The function which minimizes (3.5) is

$$\begin{aligned} \psi_H(V_1, V_2) = E_{\rho \mid V_1 V_2}\{[1+(1+\rho)n^{-1}V_1 + \rho^{-1}(1+\rho)\cdot n^{-1}V_2]^{-1}\} \\ \div E_{\rho \mid V_1, V_2}\{(1+\rho)[1+(1+\rho)n^{-1}V_1 + \rho^{-1}(1+\rho)\cdot n^{-1}V_2]^{-1}\} \quad \text{a.s.}, \end{aligned} \tag{3.8}$$

where $E_{\rho \mid V_1, V_2}\{\cdot\}$ designates the posterior expectation of the term in brackets, given (V_1, V_2).

We verify now that the conditions of the Fubini theorem are satisfied by $\psi_H(V_1, V_2)$. Substituting (3.8) in (3.5) we obtain, since $0 \leqq \psi_H(V_1, V_2) \leqq 1$ a.s.,

$$\begin{aligned} &\int_0^\infty \int dV_1 \, dV_2 \, p(V_1, V_2; H) \cdot S_H(V_1, V_2; \psi_H) \\ &\leqq \int_0^\infty \int dV_1 \, dV_2 \, p(V_1, V_2; H) \int_0^\infty H(d\rho \mid V_1, V_2) \, |\psi_H(V_1, V_2)(1+\rho) - 2| \\ &\qquad \cdot (1+(1+\rho)n^{-1}V_1 + \rho^{-1}(1+\rho)n^{-1}V_2)^{-1}. \end{aligned} \tag{3.9}$$

It is a straightforward matter to prove that the R.H.S. of (3.9) is bounded by $2^{\frac{1}{2}}$. Hence, the Fubini theorem is satisfied, and we obtain:

THEOREM. 3.1. *If $H(\rho)$ is any prior distribution function of ρ, the Bayes equivariant estimator against $H(\rho)$ is given by*:

$$\begin{aligned} \hat{\mu}_H = \bar{X} + [(\bar{Y}-\bar{X}) \cdot E_{\rho \mid V_1, V_2}\{(1+(1+\rho)n^{-1}V_1 + \rho^{-1}(1+\rho)\cdot n^{-1}V_2)^{-1}\} \\ \div E_{\rho \mid V_1, V_2}\{(1+\rho)(1+(1+\rho)n^{-1}V_1 + \rho^{-1}(1+\rho)\cdot n^{-1}V_2)^{-1}\}]. \end{aligned} \tag{3.10}$$

We can further develop the Bayes invariant coefficient $\psi_H(V_1, V_2)$ in the following manner. Let $Z_i = n^{-1}/V_i$ $(i = 1, 2)$. Then, it is easy to verify that the joint density function of (Z_1, Z_2) given ρ is:

$$\begin{aligned} p(Z_1, Z_2 \mid \rho) = \Gamma(n-\tfrac{1}{2})\Gamma^{-2}(\tfrac{1}{2}(n-1))\pi^{-\frac{1}{2}}\rho^{\frac{1}{2}(n-1)}(1+\rho)^{n-1} \\ \cdot (Z_1 Z_2)^{\frac{1}{2}(n-3)}(1+(1+\rho)Z_1 + \rho^{-1}(1+\rho)Z_z)^{-(n-\frac{1}{2})}, \end{aligned} \tag{3.11}$$

$0 \leqq Z_1, Z_2 \leqq \infty$. Finally, using Bayes formula for the posterior distribution of ρ given (Z_1, Z_2), we obtain

$$\begin{aligned} \phi_H(Z_1, Z_2) &= E_{\rho \mid Z_1, Z_2}\{(1+(1+\rho)Z_1 + \rho^{-1}(1+\rho)Z_2)^{-1}\} \\ &\quad \div E_{\rho \mid Z_1, Z_2}\{(1+\rho)(1+(1+\rho)Z_1 + \rho^{-1}(1+\rho)Z_2)^{-1}\} \\ &= \int_0^\infty \rho^{-\frac{1}{2}(n-1)}(1+\rho)^{n-1}[1+(1+\rho)Z_1 + \rho^{-1}(1+\rho)Z_2]^{-(n+\frac{1}{2})}H(d\rho) \\ &\quad \div \int^\infty \rho^{-\frac{1}{2}(n-1)}(1+\rho)^{n}[1+(1+\rho)Z_1 + \rho^{-1}(1+\rho)Z_2]^{-(n+\frac{1}{2})}H(d\rho). \end{aligned} \tag{3.12}$$

We notice that $\psi_H(V_1, V_2) = \phi_H(Z_1, Z_2)$ for all $(Z_1, Z_2) = (n^{-1}V_1, n^{-1}V_2)$.

In the special case, where the prior distribution $H(\rho)$ is concentrated on a single point ρ_0, the Bayes equivariant estimator is

$$\hat{\mu}(\rho_0) = \bar{X} + (\bar{Y} - \bar{X})(1+\rho_0)^{-1}, \tag{3.13}$$

which is the locally best unbiased estimator of μ, at $\rho = \rho_0$(μ and σ^2 are arbitrary). From (3.13) we can draw immediately the following:

COROLLARY 3.1. *The sample mean $\bar{X}$ is a minimax estimator of μ.*

PROOF. First, obviously the risk function of $\bar{X}$ is n^{-1} for all ρ. Fix $\rho = \rho_0$ and $\sigma = \sigma_0$. Then the Bayes estimator against a prior normal $N(\mu_0, \tau^2)$ for μ is $(\bar{X}n/\sigma_0{}^2 + \bar{Y}n/\rho_0\sigma_0{}^2 + \mu_0/\tau^2)/(n/\sigma_0{}^2 + n/\rho\sigma_0{}^2 + \tau^{-2})$. This Bayes esimator has a prior risk of $n^{-1}\rho_0(1+\rho_0)/(1+\rho_0+\rho_0\sigma_0{}^2/n\tau^2)^2$. Letting $\tau^2 \to \infty$ and $\rho_0 \to \infty$ the above sequence of Bayes estimator approach $\bar{X}$, and the associated sequence of prior risks approaches n^{-1}. Hence, applying a result of Blyth [2], $\bar{X}$ is a minimax estimator, for the loss function (3.1), in the class of all estimators.

It is interesting to notice that due to the loss function (3.1), which gives an unsymmetric role to $\bar{X}$ and $\bar{Y}$, $\bar{Y}$ is *not* a minimax estimator of μ. In certain cases this non-symmetric structure is unwarranted, since one may find himself in a situation in which he has no preference for the X observations over the Y observations, and one would like to consider only equivariant estimators which are symmetric functions of the sufficient statistics. That is, to satisfy the condition (a.s.)

$$\bar{X} + (\bar{Y} - \bar{X})\phi_H(Z_1, Z_2) = \bar{Y} + (\bar{X} - \bar{Y})\phi_{H^*}(Z_2, Z_1) \tag{3.14}$$

where, if H is a given prior distribution of ρ then H^* is the induced prior distribution of ρ^{-1}. Equivalently, the symmetric weighing function should satisfy

$$\phi_H(Z_1, Z_2) + \phi_{H^*}(Z_2, Z_1) = 1 \qquad \text{a.s.} \tag{3.15}$$

One way to attain (3.15) is to restrict attention to prior distributions of ρ, which are "symmetric" in the sense that

$$\int_{a-0}^{b} H(d\rho) = \int_{b^{-1}-0}^{a^{-1}} H(d\rho), \quad \text{for all} \quad 0 \leq a < b \leq 1. \tag{3.16}$$

This is, however, quite a severe restriction on the class of prior distributions. Property (3.15) can be attained, however, without the restriction (3.16) by considering Bayes equivariant estimators for the quadratic loss function

$$L^*(\hat{\mu}, \mu) = (\hat{\mu} - \mu)^2/[\sigma^2 \max(1, \rho)]. \tag{3.17}$$

This loss function makes the roles of $\bar{X}$, S_1 and of $\bar{Y}, S_2$ symmetric in the above sense. Using the above derivation of the Bayes equivariant estimators where, for a given prior distribution $H(\rho)$ we substitute the defective[2] prior distribution $H^*(\rho) = \int_0^\rho H(dt)/\max(1, t)$. We then obtain:

[2] Following Feller [3], page 112, a distribution function $F(x)$ is called defective if $F(+\infty) < 1$.

THEOREM 3.2. *For any given prior distribution* $H(\rho)$, *the Bayes equivariant estimator for the quadratic loss function* (3.17) *is* $\mu_H = \bar{X}+(\bar{Y}-\bar{X})\phi_H{}^*(Z_1, Z_2)$, *where*:

$$\phi_H{}^*(Z_1,Z_2) = \left(\int_0^1 \rho^{-\frac{1}{2}(n-1)}+\int_1^\infty \rho^{-\frac{1}{2}(n+1)}\right)(1+\rho)^{n-1}$$

$$\cdot[1+(1+\rho)Z_1+\rho^{-1}(1+\rho)Z_2]^{-(n+\frac{1}{2})}H(d\rho) \tag{3.18}$$

$$\div \left(\int_0^1 \rho^{-\frac{1}{2}(n-1)}+\int_1^\infty \rho^{-\frac{1}{2}(n+1)}\right)(1+\rho)^{n}$$

$$\cdot[1+(1+\rho)Z_1+\rho^{-1}(1+\rho)Z_2]^{-(n+\frac{1}{2})}H(d\rho).$$

Furthermore, a minimax estimator for the quadratic loss function (3.17) *is* $\tilde{\mu} = \frac{1}{2}(\bar{X}+\bar{Y})$.

One can easily check that $\phi_H{}^*(Z_1, Z_2)$ given by (3.18) satisfies the symmetry condition (3.15). The minimaxity of $\tilde{\mu} = \frac{1}{2}(\bar{X}+\bar{Y})$ under (3.17) can be verified on the basis of Corollary 3.1 and the symmetry structure of the decision problem. Another way to verify it is to notice that the risk of $\tilde{\mu}$ under (3.17) is

$$R(\tilde{\mu},\rho) = (4n)^{-1}(1+\rho), \qquad 0 \leqq \rho \leqq 1; \tag{3.19}$$

$$= (4n)^{-1}(1+\rho^{-1}), \qquad 1 \leqq \rho \leqq \infty.$$

Hence, $\sup_{0\leqq\rho\leqq\infty} R(\tilde{\mu},\rho) = R(\tilde{\mu},1) = (2n)^{-1}$. This, is, however, the limit when $\tau^2 \to \infty$ of the prior risk of the Bayes estimators, with $\rho_0 = 1$, $\tilde{\mu} = (\bar{X}_n/\sigma_0{}^2 + \bar{Y}_n/\sigma_0{}^2+\mu/\tau^2)/(2n/\sigma_0{}^2+1/\tau^2)$. Hence $\tilde{\mu}$ is minimax.

4. Fiducial equivariant estimators. Let $\varphi(u)$ designate the standard normal density, and let $q(x)$ designate the density function of $\chi^2[n-1]$, i.e., a chi-square with $(n-1)$ degrees of freedom. Then, the joint density function of the sufficient statistic $(\bar{X}, \bar{Y}, S_1, S_2)$, given $\theta = (\mu, \sigma^2, \rho)$ is:

$$f(\bar{X},\bar{Y},S_1,S_2) \propto \rho^{-\frac{3}{2}}\sigma^{-6}\varphi(n^{\frac{1}{2}}\sigma^{-1}(\bar{X}-\mu))\varphi(n^{\frac{1}{2}}\sigma^{-1}\rho^{-\frac{1}{2}}(\bar{Y}-\mu))q(\sigma^{-2}S_1)q(\rho^{-1}\sigma^{-2}S_2). \tag{4.1}$$

If $H(\rho\,|\,\mu,\sigma^2)$ designates the conditional prior distribution of ρ given (μ,σ^2), and $K(\mu,\sigma^2)$ designates the joint prior distribution of (μ,σ^2) then the general formula of the Bayes estimator of the common mean μ, for the quadratic loss functon $L(\tilde{\mu},\mu) = (\hat{\mu}-\mu)^2/\sigma^2$, is:

$$\hat{\mu}(X,Y,S_1,S_2)$$

$$= \int_{-\infty}^{\infty}\int_0^\infty K(d\mu,d\sigma^2)\mu\sigma^{-8}\varphi(n^{\frac{1}{2}}\sigma^{-1}(\bar{X}-\mu))q(\sigma^{-2}S_1)f_H(n^{\frac{1}{2}}\sigma^{-1}(\bar{Y}-\mu),\sigma^{-2}S_2) \tag{4.2}$$

$$\div \int_{-\infty}^{\infty}\int_0^\infty K(d\mu,d\sigma^2)\sigma^{-8}\varphi(n^{\frac{1}{2}}\sigma^{-1}(\bar{X}-\mu))q(\sigma^{-2}S_2)f_H(n^{\frac{1}{2}}\sigma^{-1}(\bar{Y}-\mu),\sigma^{-2}S_2),$$

where

$$f_H(n^{\frac{1}{2}}\sigma^{-1}(\bar{Y}-\mu),\sigma^{-2}S_2) = \int_0^\infty H(d\rho\,|\,\mu,\sigma^2)\rho^{-\frac{1}{2}}\varphi(n^{\frac{1}{2}}\sigma^{-1}\rho^{-\frac{1}{2}}(\bar{Y}-\mu))q(\rho^{-1}\sigma^{-2}S_2.) \tag{4.3}$$

If we substitute for ρ any prior distribution $H(\rho)$, independently of μ and σ^2 and the Jeffrey's prior $K(d\mu, d\sigma^2) = d\mu\, d\sigma^2$, we obtain the *formal* Bayes estimator:

$$\begin{aligned}\tilde{\mu}_H(\bar{X}, \bar{Y}, S_1, S_2) = \textstyle\int_0^\infty H(d\rho)\rho^{-\frac{3}{2}} \int_0^\infty d\sigma^2 \cdot \sigma^{-8} q(\sigma^{-2}S_1) q(\rho^{-1}\sigma^{-2}S_2) \\ \cdot \textstyle\int_{-\infty}^{\infty} d\mu \cdot \mu\varphi(n^{\frac{1}{2}}\sigma^{-1}(\bar{X}-\mu))\varphi(n^{\frac{1}{2}}\sigma^{-1}\rho^{-\frac{1}{2}}(\bar{Y}-\mu)) \\ \div \textstyle\int_0^\infty H(d\rho)\rho^{-\frac{3}{2}} \int_0^\infty d\sigma^2 \cdot \sigma^{-8} q(\sigma^{-2}S_1) q(\rho^{-1}\sigma^{-2}S_2) \\ \cdot \textstyle\int_{-\infty}^{\infty} d\mu \cdot \varphi(n^{\frac{1}{2}}\sigma^{-1}(\bar{X}-\mu))\varphi(n^{\frac{1}{2}}\sigma^{-1}\rho^{-\frac{1}{2}}(\bar{Y}-\mu)).\end{aligned} \tag{4.4}$$

This estimator will be called a *fiducial estimator* of μ. It is an equivariant estimator of μ, with respect to $\mathcal{G}$. We develop now the form of the fiducial estimators. It is easily verified that:

$$\begin{aligned}\textstyle\int_{-\infty}^{\infty} \varphi(n^{\frac{1}{2}}\sigma^{-1}(\bar{X}-\mu))\varphi(n^{\frac{1}{2}}\sigma^{-1}\rho^{-\frac{1}{2}}(\bar{Y}-\mu))\, d\mu \\ = [n\sigma^{-2}\cdot\rho^{-1}(1+\rho)]^{\frac{1}{2}}\varphi(n^{\frac{1}{2}}\sigma^{-1}(1+\rho)^{-\frac{1}{2}}(\bar{X}-\bar{Y})),\end{aligned} \tag{4.5}$$

and

$$\begin{aligned}\textstyle\int_{-\infty}^{\infty} \mu\varphi(n^{\frac{1}{2}}\sigma^{-1}(\bar{X}-\mu))\varphi(n^{\frac{1}{2}}\sigma^{-1}\rho^{-\frac{1}{2}}(\bar{Y}-\mu))\, d\mu \\ = (1+\rho)^{-1}(\rho\bar{X}+\bar{Y})[n(1+\rho)\sigma^{-2}\rho^{-1}]^{\frac{1}{2}}\varphi(n^{\frac{1}{2}}\sigma^{-1}(1+\rho)^{-\frac{1}{2}}(\bar{X}-\bar{Y})).\end{aligned} \tag{4.6}$$

Furthermore,

$$\begin{aligned}\textstyle\int_0^\infty d\sigma^2\cdot\sigma^{-9}q(\sigma^{-2}S_1)q(\rho^{-1}\sigma^{-2}S_2)\exp\{-\frac{1}{2}n\sigma^{-2}(1+\rho)^{-1}(\bar{X}-\bar{Y})^2\} \\ \propto \rho^{-\frac{1}{2}(n-3)}(S_1 S_2)^{\frac{1}{2}(n-3)} \textstyle\int_0^\infty d\sigma^2\cdot\sigma^{-9}\cdot(\sigma^{-2})^{n-3} \\ \cdot\exp\{-\frac{1}{2}n\sigma^{-2}(1+\rho)^{-1}(\bar{X}-\bar{Y})^2[1+(1+\rho)Z_1+\rho^{-1}(1+\rho)Z_2]\} \\ \propto (\bar{X}-\bar{Y})^{-(n+\frac{1}{2})}(S_1 S_2)^{\frac{1}{2}(n-3)}\cdot\rho^{-\frac{1}{2}(n-3)}(1+\rho)^{n+\frac{1}{2}} \\ \cdot[1+(1+\rho)Z_1+\rho^{-1}(1+\rho)Z_2]^{-(n+\frac{1}{2})}.\end{aligned} \tag{4.7}$$

Substituting (4.5)–(4.7) in (4.4) we obtain that the general form of a fiducial estimator, with a prior distribution $H(\rho)$ is

$$\tilde{\mu}_H(\bar{X}, \bar{Y}, S_1, S_2) = \bar{X} + (\bar{Y}-\bar{X})\tilde{\phi}_H(Z_1, Z_2), \tag{4.8}$$

where the invariant coefficient is

$$\begin{aligned}\tilde{\phi}_H(Z_1, Z_2) = \textstyle\int_0^\infty H(d\rho)\rho^{-\frac{1}{2}(n+1)}(1+\rho)^n[1+(1+\rho)Z_1+(1+\rho)\rho^{-1}Z_2]^{-(n+\frac{1}{2})} \\ \div \textstyle\int_0^\infty H(d\rho)\rho^{-\frac{1}{2}(n+1)}(1+\rho)^{n+1}[1+(1+\rho)Z_1+(1+\rho)\rho^{-1}Z_2]^{-(n+\frac{1}{2})}.\end{aligned} \tag{4.9}$$

We notice from (4.9) that $H(\rho)$ for the fiducial estimators does not have to be necessarily a prior distribution. It could be a defective distribution or a σ-finite measure for which the denominator of (4.9) is finite. The comparison of (3.12) and (4.9) shows that if one chooses a prior distribution $H(\rho)$ for the Bayes equivariant estimator, and a defective prior distribution $\tilde{H}(\rho)$ for the fiducial estimator, such that $\tilde{H}(d\rho) = (1+\rho)^{-1}\rho H(d\rho)$ then $\tilde{\phi}_{\tilde{H}}(Z_1, Z_2) = \phi_H(Z_1, Z_2)$ a.s. That is, in that case the Bayes equivariant and the fiducial estimators coincide. If we use $\tilde{H}(\rho) = H(\rho)$ then the corresponding Bayes equivariant and the fiducial estimators are different.

5. Weak admissibility. An estimator of a function $g(\theta)$ of the parameter $\theta = (\mu, \sigma^2, \rho)$, $\hat{g}(T)$ say, where T is the minimal sufficient statistic $T = (\bar{X}, \bar{Y}, S_1, S_2)$, is called weakly admissible, with respect to the risk function $R(\hat{g}, \theta)$ if, for any other estimator $\hat{g}(T)$ such that

(5.1) $$R(\hat{g}, \theta) \leqq R(\hat{g}, \theta) \quad \text{for all} \quad \theta,$$

the set of parameter points over which strict inequality holds is a subset of a set having a Lebesgue measure zero. Karlin gave in [6] a simple criterion to check weak admissibility in the case of a quadratic loss function. As in the previous section, let $f_\theta(T)$ denote the density function of T. We have in the case of (5.1)

(5.2) $$\int (\tilde{g}(T) - g(\theta))^2 f_\theta(T)\, dT \leqq \int (\hat{g}(T) - g(\theta))^2 f_\theta(T)\, dT, \quad \text{all} \quad \theta.$$

This is equivalent to the inequality:

(5.3) $$\int (\tilde{g}(T) - \hat{g}(T))^2 f_\theta(T)\, dT \leqq 2 \int (\hat{g}(T) - \tilde{g}(T))(\hat{g}(T) - g(\theta)) f_\theta(T)\, dT,$$

for all θ.

Let $\xi(d\theta) = h(\theta)\, d\theta$, with $h(\theta) > 0$ for all $\theta = (\mu, \sigma^2, \rho)$ be a sigma-finite positive measure on the sigma-field generated by θ. $\xi(d\theta)$ is equivalent to the Lebesgue measure, i.e. $\xi(d\theta) \ll d\theta$ and $d\theta \ll \xi(d\theta)$. Furthermore, assume that $\xi(d\theta)$ satisfies:

(5.4) $$f_\xi(T) = \int_\Theta f_\theta(T) \xi(d\theta) < \infty, \qquad \text{a.s.,}$$

and

(5.5) $$\int_\Theta R(\hat{g}, \theta) \xi(d\theta) < \infty.$$

Then, integrating both sides of (5.3) with respect to $\xi(d\theta)$, we obtain

(5.6) $$\int \xi(d\theta) \int (\tilde{g}(T) - \hat{g}(T))^2 f_\theta(T)\, dT$$
$$\leqq 2 \int \xi(d\theta) \int (\hat{g}(T) - \tilde{g}(T))(\hat{g}(T) - g(\theta)) f_\theta(T)\, dt.$$

On the set of T points, Λ, for which $f_\xi(T) > 0$, define

(5.7) $$\xi(d\theta \mid T) = f_\theta(T) \xi(d\theta) / f_\xi(T) \qquad \text{a.s.}$$

We have $P_\xi(\Lambda) = 1$. We impose also the condition:

(5.8) $$S(T; \xi) = \int_\Theta |\hat{g}(T) - g(\theta)|\, \xi(d\theta \mid T) < \infty \qquad \text{a.s.}$$

Then, (5.1), (5.5), (5.8) and the Schwartz inequality imply that

(5.9) $$\int |\hat{g}(T) - \tilde{g}(T)|\, S(T, \xi) f_\xi(T)\, dT < \infty.$$

Thus, the Fubini theorem holds, and (5.6) can be written as

(5.10) $$\int \xi(d\theta) \int (\hat{g}(T) - \tilde{g}(T))^2 f_\theta(T)\, dt \leqq 2 \int (\hat{g}(T) - \tilde{g}(T)) G(T; \xi) f_\xi(T)\, dT,$$

where

$$G(T;\xi) = \int (\hat{g}(T) - g(\theta))\xi(d\theta \mid T) \qquad \text{a.s.} \tag{5.11}$$

Finally, if

$$\hat{g}(T) f_{\xi}(T) = \int g(\theta) f_{\theta}(T) \xi(d\theta) \qquad \text{a.s.} \tag{5.12}$$

then $G(T, \xi) = 0$ a.s. Since the L.H.S. of (5.10) is a.s. non-negative, $G(T, \xi) = 0$ a.s. implies that $\int (\hat{g}(T) - \tilde{g}(T))^2 f_{\theta}(T)\, dT = 0$ on every set of θ values of a positive Lebesgue measure. Hence, $G(T, \xi) = 0$ a.s. implies that $R(\hat{g}, \theta) = R(\tilde{g}, \theta)$ on every set of θ values having a positive Lebesgue measure. Thus, under the above conditions, $G(T, \xi) = 0$ a.s. is a sufficient condition for the weak-admissibility of $\hat{g}(T)$.

Consider the fiducial estimators of Section 4. There we used the sigma-finite measures $\xi(d\theta) = H(d\rho)\, d\mu\, d\sigma^2$ and $g(\theta) = \mu$. Assume that $H(d\rho) = h(\rho)\, d\rho$, with $h(\rho) > 0$ for all $0 < \rho < \infty$. T is the minimal sufficient statistic $(\bar{X}, \bar{Y}, S_1, S_2)$. $f_{\xi}(T)$ is given by the denominator of (4.4), and $H(d\rho)$ is such that $f_{\xi}(T)$ is finite a.s. Hence (5.4) is satisfied. Furthermore, if $H(\rho)$ is a prior distribution and from the results of Section 3, if $\hat{g}(T)$ is either the fiducial estimator $\tilde{\mu}_H(\bar{X}, \bar{Y}, S_1, S_2)$ or the Bayes equivariant estimator $\hat{\mu}_H(\bar{X}, \bar{Y}, S_1, S_2)$ both conditions (5.5) and (5.8) are satisfied. Moreover, the definition of the fiducial estimator (4.4) implies that (5.12) is satisfied a.s. Hence, the sufficient condition for the weak-admissibility of a fiducial estimator of μ, for any prior distribution $H(\rho)$ satisfying the above conditions, has been established. Furthermore, considering $\tilde{\xi}(d\theta) = H(d\rho) \cdot \rho(1+\rho)^{-1} \cdot d\mu \cdot d\sigma^2$ where $H(\rho)$ is a prior distribution satisfying the above conditions, we imply that the fiducial estimator $\tilde{\mu}_{\tilde{H}}(\bar{X}, \bar{Y}, S_1, S_2)$ is weakly-admissible, where $\tilde{H}(\rho) = \rho(1+\rho)^{-1} H(\rho)$. But, since $\tilde{\mu}_{\tilde{H}}(\bar{X}, \bar{Y}, S_1, S_2) = \hat{\mu}_H(\bar{X}, \bar{Y}, S_1, S_2)$ we obtain:

THEOREM 5.1. *All Bayes equivariant and fiducial estimators, with respect to absolutely continuous prior distributions $H(\rho)$, with positive densities for all $0 < \rho < \infty$, are weakly-admissible.*

Acknowledgment. The author wishes to express his gratitude to his former student Dr. Sung Wook Lee, who checked the derivations of various formulae. Thanks are also extended to the referee and the Associate Editor for their helpful criticism of the manuscript.

REFERENCES

[1] BERK, R. H. (1967). A special group structure and equivariant estimation. *Ann. Math. Statist.* **38** 1436–1445.

[2] BLYTH, C. (1951). On minimax statistical decision procedures and their admissibility. *Ann. Math. Statist.* **22** 22–42.

[3] FELLER, W. (1966). *An Introduction to Probability Theory and its Applications* **2**. Wiley, New York.

[4] GURLAND, J. and MEHTA, J. S. Combinations of unbiased estimators of the mean which consider inequality of unknown variances. To be published in the *J. Amer. Statist. Assoc.*

[5] HORA, R. B. and BUEHLER, R. J. (1966). Fiducial theory and invariant estimation. *Ann. Math. Statist.* **37** 643–656

[6] KARLIN, S. (1958). Admissibility of estimation with quadratic loss. *Ann. Math. Statist.* **29** 406–436.

[7] STEIN, C. M. (1965). Approximation of improper prior measures by prior probability measures. *Bernulli* (1713) *Bayes* (1763) *Laplace* (1813), ed. J. Neyman and L. M. LeCam, eds. Springer-Verlag, New York, 217–270.

[8] WIJSMAN, R. A. (1967). Cross-sections or orbits and their application to densities of maximal invariants. *Proc. Fifth Berkeley Symp. Math. Statist. Prob.* **1** 389–400.

[9] ZACKS, S. (1966). Unbiased estimation of the common mean of two normal distributions based on small samples of equal size. *J. Amer. Statist. Assoc.* **61** 467–476.