Title	A combined estimator of a common parameter
Sub Title	
Author	赤井, 豊秋(Akai, Toyoaki)
Publisher	慶應義塾大学理工学部
Publication year	1982
Jtitle	Keio Science and Technology Reports Vol.35, No.5 (1982. 4) ,p.93- 104
JaLC DOI	
Abstract	Suppose that two probability distributions have parameters (θ, σx) and (θ, σy) respectively. To estimate the common parameter θ based on independent samples from each distribution, a weighted mean of unbiased estimators ■x and ■y is used. In this paper, we give necessary and sufficient conditions for the combined estimator to have smaller variance than ■x. And we give those for the uniform distribution and the variance of the estimator for it is computed. Also we give those for the inverse gaussian distribution.
Notes	
Genre	Departmental Bulletin Paper
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=KO50001004-00350005-0093

KEIO SCIENCE AND TECHNOLOGY REPORTS
VOL. 35, NO. 5, pp. 93-104, 1982

A COMBINED ESTIMATOR OF A COMMON PARAMETER

TOYOAKI AKAI

Dept. of Mathematics Keio University
Yokohama 223, Japan

(*Received March 6, 1982*)

ABSTRACT

Suppose that two probability distributions have parameters (θ, σ_x) and (θ, σ_y) respectively. To estimate the common parameter θ based on independent samples from each distribution, a weighted mean of unbiased estimators $\hat{\theta}_x$ and $\hat{\theta}_y$ is used.

In this paper, we give necessary and sufficient conditions for the combined estimator to have smaller variance than $\hat{\theta}_x$. And we give those for the uniform distribution and the variance of the estimator for it is computed. Also we give those for the inverse gaussian distribution.

1. Introduction

Suppose that two probability distributions P_x and P_y on the real line have parameters (θ, σ_x) and (θ, σ_y) respectively. The common parameter θ is estimated by a linear combination $\hat{\theta}^*$ of unbiased estimators $\hat{\theta}_x$ and $\hat{\theta}_y$ of samples from P_x and P_y respectively. The weights for $\hat{\theta}^*$ are determined by estimators of variances of $\hat{\theta}_x$ and $\hat{\theta}_y$ so that $\hat{\theta}^*$ is unbiased and has small variance. If the estimators of the variances $V(\hat{\theta}_x)$ and $V(\hat{\theta}_y)$ are not accurate enough, the combined estimator $\hat{\theta}^*$ is not necessarily better than $\hat{\theta}_x$ and $\hat{\theta}_y$.

In this paper, two cases are studied. Firstly P_x and P_y are identically distributed and have their densities which are symmetric about θ, and σ_x and σ_y are their scale parameters. Independent random samples of sizes m and n are observed from P_x and P_y respectively. The estimator $\hat{\theta}_x$ is covariant with respect to the location-scale transformation, that is,

$$\hat{\theta}_x(ax_1+b, \cdots, ax_m+b) = a\hat{\theta}_x(x_1, \cdots, x_m)+b \tag{1.1}$$

for every real a and every real b. The estimator $\hat{\sigma}_x$ is invariant and covariant with respect to the location and the scale transformation respectively, that is,

$$\hat{\sigma}_x(ax_1+b, \cdots, ax_m+b) = |a|\hat{\sigma}_x(x_1, \cdots, x_m) \tag{1.2}$$

for every real a and every real b. The estimators $\hat{\theta}_y$ and $\hat{\sigma}_y$ have the corresponding properties.

In the second case, the parameter θ is a common mean value, not always a location parameter, of P_x and P_y. All the statistics $\hat{\theta}_x, \hat{\theta}_y$ and the estimators of $V(\hat{\theta}_x)$

and $V(\hat{\theta}_y)$ are assumed to be independently distributed. For a few families of distributions such an estimator of the variance is available from a single sample, otherwise the estimator is obtained from another independent sample.

In both cases, we try to find good combined estimators and determine such limits on m and n that $\hat{\theta}^*$ has smaller variance than $\hat{\theta}_x$ or than both $\hat{\theta}_x$ and $\hat{\theta}_y$. The normal distributions and the two-parameter rectangular distributions belong to the both cases.

In the first case, Cohen (1976), using Hogg's result (1960), gave necessary and sufficient conditions for his combined unbiased estimator to have smaller variance than $\hat{\theta}_x$. He also showed situations when his estimator is better than both of the individual estimators for the uniform distribution. Bhattacharya (1981) improved the sufficient condition for Cohen's estimator to have smaller variance than the first sample's. The combined estimator of this paper is a slight modification of Cohen's and of the same type as Khatri and Shah's (1974) of the normal case.

In section 2, we give necessary and sufficient conditions for the combined estimator to have smaller variance than $\hat{\theta}_x$ for the first case. In section 3, we give a necessary and sufficient condition for the estimator to be better than $\hat{\theta}_x$ for the uniform distribution by applying the result of section 2. The condition is simpler than those by Cohen (1976) and Bhattacharya (1981): Our combined estimator is better than both of the individual estimators for all $m \geqq 19$ if $m=n$. In section 4, the variance of the estimator for the uniform distribution is computed and a numerical table is shown.

In section 5, in the second case, we give a necessary and sufficient condition for the combined estimator to be better than $\hat{\theta}_x$. We give that for the uniform distribution. It is interesting to see the condition be applicable to the estimation of the common mean of two inverse gaussian distributions.

2. A combined estimator of a common location

In order to study our first case, put $T_x=(\hat{\theta}_x-\theta)/\sigma_x$, $T_y=(\hat{\theta}_y-\theta)/\sigma_y$, $S_x=\hat{\sigma}_x/\sigma_x$, $S_y=\hat{\sigma}_y/\sigma_y$, $W=S_x^2/S_y^2$, $K=V(\hat{\theta}_x)/\sigma_x^2=E(T_x^2)$ and $L=V(\hat{\theta}_y)/\sigma_y^2=E(T_y^2)$. The distributions of T_x, T_y, S_x and S_y are independent of the unknown parameters. K and L depend on the density and the sample sizes. Assume that $E(1/S_y^4)<\infty$, which holds when $n \geqq 6$ as shown by Cohen (1976).

Cohen (1976) suggested the following unbiased estimator;

$$\hat{\theta}_a=\hat{\theta}_x+a(\hat{\theta}_y-\hat{\theta}_x)/(1+Z) \tag{2.1}$$

where $Z=\hat{\sigma}_y^2/\hat{\sigma}_x^2$ and $a \geqq 0$ is a constant to be suitably chosen. Since this estimator is not symmetric in $\hat{\theta}_x$ and $\hat{\theta}_y$, the following modification is suggested;

$$\hat{\theta}^*=\frac{c_1L\hat{\sigma}_y^2}{c_2K\hat{\sigma}_x^2+c_1L\hat{\sigma}_y^2}\hat{\theta}_x+\frac{c_2K\hat{\sigma}_x^2}{c_2K\hat{\sigma}_x^2+c_1L\hat{\sigma}_y^2}\hat{\theta}_y \tag{2.2}$$

where c_1 and c_2 are positive constants to be suitably chosen. The estimator is unbiased because of Theorem in Appendix by Hogg (1960). Note that $\hat{\theta}_a$ is a special form of $\hat{\theta}^*$ if $a=1$.

Rewriting Theorem 2.2 of Cohen (1976), we can obtain the following;

Theorem 2.1 Let $c=c_2/c_1$ and $\rho=K\sigma_x^2/L\sigma_y^2$. Suppose that

$$[d/d\rho]E[(T_x^2/K+T_y^2\rho c^2W^2/L)(1+\rho cW)^{-2}] \\ =E[d/d\rho][(T_x^2/K+T_y^2\rho c^2W^2/L)(1+\rho cW)^{-2}]. \tag{2.3}$$

Then a necessary condition for the estimator $\hat{\theta}^*$ to have uniformly smaller variance than $\hat{\theta}_x$ is

$$c\leqq 2E(T_x^2W/K)/E(T_y^2W^2/L). \tag{2.4}$$

Note. The assumption (2.3) is equivalent to the following; Let $f(T_x, T_y, S_x, S_y)$ denote the joint density of T_x, T_y, S_x and S_y and put

$$g(T_x, T_y, S_x, S_y, \rho)=(T_x^2/K+T_y^2\rho c^2W^2/L)(1+\rho cW)^{-2}f(T_x, T_y, S_x, S_y).$$

We assume that g satisfies the following conditions:
(i) g is integrable with respect to T_x, T_y, S_x and S_y.
(ii) g is differential with respect to ρ.
(iii) There exists an integrable function $\phi(T_x, T_y, S_x, S_y)$ such that

$$|g_\rho(T_x, T_y, S_x, S_y)|\leqq\phi(T_x, T_y, S_x, S_y),$$

where $g_\rho(T_x, T_y, S_x, S_y)$ denotes the partial derivative of g.

Proof. The proof is essentially the same as given in Graybill and Deal (1959). The variance of $\hat{\theta}^*$ is equal to

$$E[(1+\rho cW)^{-2}(\hat{\theta}_x-\theta)^2]+E[2\rho cW(1+\rho cW)^{-2}(\hat{\theta}_x-\theta)(\hat{\theta}_y-\theta)] \\ +E[\rho^2c^2W^2(1+\rho cW)^{-2}(\hat{\theta}_y-\theta)^2].$$

Theorem in Appendix can be used to show that

$$E[2\rho cW(1+\rho cW)^{-2}(\hat{\theta}_x-\theta)(\hat{\theta}_y-\theta)]=0.$$

Therefore

$$V(\hat{\theta}^*)=K\sigma_x^2E[(T_x^2/K+T_y^2\rho c^2W^2/L)(1+\rho cW)^{-2}]. \tag{2.5}$$

The value of (2.5) is $K\sigma_x^2$ at $\rho=0$. The derivative of (2.5) with respect to ρ at $\rho=0$ must be equal to 0 or negative. Therefore we have (2.4).

Next we give a sufficient condition for $\hat{\theta}^*$ to be better than $\hat{\theta}_x$.

Theorem 2.2 Suppose that

$$E(T_x^2W^2/K)E(T_y^2W/L)<5E(T_x^2W/K)E(T_y^2W^2/L). \tag{2.6}$$

Then a sufficient condition for $\hat{\theta}^*$ to be uniformly better than $\hat{\theta}_x$ is

$$c\leqq 2E(T_x^2W/K)/E(T_y^2W^2/L)$$

and

$$c\leqq 5E(T_x^2W/K)/2E(T_y^2W^2/L) \\ -E(T_x^2W^2/K)E(T_y^2W/L)/2[E(T_y^2W^2/L)]^2. \tag{2.7}$$

Note. The right hand side of (2.7) is positive under the condition (2.6).

Proof. Let

$$h_1(W,\rho)=h_1(W)=1/(1+\rho cW)^2$$

and

$$h_2(W,\rho)=h_2(W)=\rho c^2W^2/(1+\rho cW)^2.$$

Then

$$h_1(0)=1,\ h_1(1/cd)=1/(1+\rho/d)^2,$$

and

$$h_2(0)=0,\ h_2(1/cd)=(\rho/d^2)/(1+\rho/d)^2,$$

where $d=E(T_y^2W^2/L)/E(T_x^2W^2/K)$.

We approximate the curve $h_1(W)$ by the parabola $f_1(W)$ passing at the point $(0,1)$ and touching at the point $(1/cd, 1/(1+\rho/d)^2)$. The equation for the parabola is

$$f_1(W)=\frac{\rho}{(1+\rho/d)^3}\left\{\left(\frac{\rho^2}{d}+3\rho\right)c^2W^2-2\left(\frac{\rho^2}{d^2}+\frac{3\rho}{d}+1\right)cW\right\}+1.$$

And we approximate similarly the curve $h_2(W)$ by the parabola $f_2(W)$:

$$f_2(W)=\frac{\rho}{(1+\rho/d)^3}\left\{\left(1-\frac{\rho}{d}\right)c^2W^2+\frac{2\rho}{d^2}cW\right\}.$$

It can be easily shown that $h_1(W)\leqq f_1(W)$ and $h_2(W)\leqq f_2(W)$ for all values of W and ρ satisfying $0<W<\infty$ and $0<\rho\leqq d$. Thus from (2.5), we have

$$E\left\{\frac{T_x^2/K+T_y^2\rho c^2W^2/L}{(1+\rho cW)^2}\right\}\leqq\frac{\rho cd_1}{(1+t)^3}\left\{t^2+2\left(1+\frac{d_2}{d_1}\right)t+1\right\}+1 \tag{2.8}$$

where

$$d_1=cE(T_y^2W^2/L)-2E(T_x^2W/K),$$
$$d_2=E(T_x^2W^2/K)E(T_y^2W/L)/E(T_y^2W^2/L)-E(T_x^2W/K)$$

and

$$t=\rho/d.$$

By the assumption, it is clear that $\rho cd_1/(1+t)^3\leqq 0$. Hence from (2.8), it suffices to show that $g_1(t)=t^2+2(1+d_2/d_1)t+1\geqq 0$ for any t with $0<t\leqq 1$ under the conditions (2.4) and (2.7). If $t=0$, then $g_1(0)=1$. Therefore it is clear that $g_1(t)\geqq 0$ for all t in the interval $0<t\leqq 1$ if $g_1(1)\geqq 0$. By the condition that $g_1(1)\geqq 0$, we have (2.7). Hence we can show that the conditions (2.4), (2.6) and (2.7) are sufficient for $E[(T_x^2/K+T_y^2\rho c^2W^2/L)(1+\rho cW)^{-2}]\leqq 1$ for all $0<\rho\leqq d$.

Next we prove the case $\rho\geqq d$. The curve $h_1(W)$ is nonincreasing in ρ. The maximum value of $h_2(W)$ is equal to the value of $f_2(W)$ with $\rho=d$. Hence the proof for $\rho=d$ implies that for $\rho\geqq d$ and the proof is complete.

Corollary 2.3 Suppose that

$$E(T_x^2W^2/K)E(T_y^2W/L)\leqq E(T_x^2W/K)E(T_y^2W^2/L). \tag{2.9}$$

Then a sufficient condition for $\hat{\theta}^*$ to be uniformly better than $\hat{\theta}_x$ is

$$c\leqq 2E(T_x^2W/K)/E(T_y^2W^2/L).$$

Proof. From (2.9), it is clear that d_2 in the right hand side of (2.8) is negative or zero. Hence we have $g_1(t)\geqq 0$ for all $0<\rho<\infty$ and the proof is complete.

Remark 2.1 In fact it can be easily shown that the inequality (2.9) in Corollary 2.3 holds for the estimation of a common location for the normal distribution which was discussed in Khatri and Shah (1974). By Theorem 2.1 and Corollary 2.3, a necessary and sufficient condition for $\hat{\theta}^*$ to be better than $\hat{\theta}_x$ is

$$c\leqq 2(m-1)(n-5)/(m+1)(n-1).$$

We can show that the inequality (2.9) holds for the estimation of a common location for the uniform distribution in section 3.

3. The uniform distribution

We estimate a common location for the two-parameter uniform distribution. we assume that the density is

$$\begin{aligned} f(x\,;\,\theta,\sigma_x)&=1/\sigma_x, \quad \text{for} \quad \theta-\sigma_x/2<x<\theta+\sigma_x/2,\\ &=0, \qquad\qquad \text{otherwise.}\end{aligned}$$

The estimators of θ and σ_x are

$$\hat{\theta}_x=(X_{(m)}+X_{(1)})/2, \hat{\sigma}'_x=X_{(m)}-X_{(1)}$$

where $X_{(m)}$ and $X_{(1)}$ are the maximum and the minimum of the random samples of sizes m from P_x respectively. The estimators $\hat{\theta}_y$ and $\hat{\sigma}'_y$ are defined similarly. Also let $\hat{\sigma}_x^2=(m+2)(m+1)\hat{\sigma}'^2_x/m(m-1)$ and $\hat{\sigma}_y^2=(n+2)(n+1)\hat{\sigma}'^2_y/n(n-1)$, and note that $\hat{\sigma}_x^2$ and $\hat{\sigma}_y^2$ are the unbiased estimators of σ_x^2 and σ_y^2 respectively. The purpose of this section is to evaluate a necessary and sufficient condition for the estimator (2.2) to be better than $\hat{\theta}_x$ using Theorem 2.1 and Corollary 2.3.

Let $U_i=(X_i-\theta)/\sigma_x$ and $V_i=(Y_i-\theta)/\sigma_y$, then

$$T_x=(U_{(m)}+U_{(1)})/2,\ T_y=(V_{(n)}+V_{(1)})/2,\ S_x=U_{(m)}-U_{(1)}$$

and $S_y=V_{(n)}-V_{(1)}$. For convenience, put $L_x=2T_x$ and $L_y=2T_y$ and note that the joint density of L_x and S_x is

$$p(L_x, S_x)=[m(m-1)/2]S_x^{m-2}, \tag{3.1}$$
$$-1+S_x<L_x<1-S_x,\ 0<S_x<1.$$

Using the joint density (3.1), it is easy to see that

$$K=1/2(m+2)(m+1),$$
$$L=1/2(n+2)(n+1),$$

$$E(T_x^2W^2/K)=\frac{M^2N(m+2)(m+1)}{(m+6)(m+5)(m+4)(m+3)(n-4)(n-5)},$$

$$E(T_x^2W/K)=\frac{MN}{(m+4)(m+3)(n-2)(n-3)}, \tag{3.2}$$

$$E(T_y^2W^2/L)=\frac{M^2N(n+2)(n+1)}{(m+4)(m+3)(n-2)(n-3)(n-4)(n-5)}$$

and

$$E(T_y^2W/L)=\frac{MN(n+2)(n+1)}{(m+2)(m+1)n(n-1)(n-2)(n-3)},$$

where

$$M=\frac{(m+2)(m+1)n(n-1)}{m(m-1)(n+2)(n+1)}$$

and $N=m(m-1)n(n-1)$. Hence from (3.2), it can be easily shown that the inequality (2.9) holds. By Theorem 2.1 and Corollary 2.3, a necessary and sufficient condition for the estimator (2.2) to be better than $\hat{\theta}_x$ is

$$c\leqq\frac{2m(m-1)(n-4)(n-5)}{(m+2)(m+1)n(n-1)}. \tag{3.3}$$

Remark 3.1 By combining the above condition (3.3) with a necessary and sufficient condition for the estimator (2.2) to be better than $\hat{\theta}_y$, a necessary and sufficient condition for the estimator (2.2) to be better than both $\hat{\theta}_x$ and $\hat{\theta}_y$ is

$$\frac{m(m-1)(n+2)(n+1)}{2(m-4)(m-5)n(n-1)}\leqq c\leqq\frac{2m(m-1)(n-4)(n-5)}{(m+2)(m+1)n(n-1)}. \tag{3.4}$$

Remark 3.2 The conditions on m and n that there exists c such that the inequality (3.4) holds are $m=11, n\geqq 164$; $m=12, n\geqq 60$; $m=13, n\geqq 40$; $m=14, n\geqq 32$; $m=15, n\geqq 27$; $m=16, n\geqq 24$; $m=17, n\geqq 22$; $m=18, n\geqq 21$; $m\geqq 19, n\geqq 19$; $m\geqq 164, n=11$; $m\geqq 60, n=12$; $m\geqq 40, n=13$; $m\geqq 32, n=14$; $m\geqq 27, n=15$; $m\geqq 24, n=16$; $m\geqq 22, n=17$ and $m\geqq 21, n=18$.

4. An approximation for the variance of $\hat{\theta}^*$

Exact expressions and computation for variances in the normal distribution were given in Khatri and Shah (1974). In this section we give an appoximation of the variance of the estimator $\hat{\theta}^*$, given by (2.2), of the common location of two uniform populations and a numerical table for it.

By (3.1), the conditional expectation of L_x^2 given S_x is $(1-S_x)^2/3$ and that of L_y^2 is $(1-S_y)^2/3$. Put $S_x=U$ and $S_y=V$ and from (2.5),

$$(4.1)\qquad V(\hat{\theta}^*)=K\sigma_x^2\left\{\frac{n(n-1)}{(n+4)(n+3)}E\frac{1}{(V^2+\rho cU^2)^2}+\frac{\rho c^2m(m+1)}{(m+4)(m+3)}E\frac{1}{(V^2+\rho cU^2)^2}\right\}$$

where U and V in the first term follow the Beta distributions with parameters $(m-1,4)$ and $(n+3,2)$ respectively, and those in the second term the Beta distributions with parameters $(m+3,2)$ and $(n-1,4)$ respectively. Rewrite the random variable as

$$(4.2)\qquad \frac{1}{(V^2+\rho cU^2)^2}=\frac{1}{(v_0^2+\rho cu_0^2)^2[1+((V^2+\rho cU^2)-(v_0^2+\rho cu_0^2))/(v_0^2+\rho cu_0^2)]^2}$$

where u_0 and v_0 are expectations of U and V respectively. Expanding the right hand side of (4.2) and using the central moments of Beta distributions, we get

$$\begin{aligned}
&\frac{n(n-1)}{(n+4)(n+3)}E\frac{1}{(V^2+\rho cU^2)^2}\\
&\quad\simeq\frac{n(n-1)}{(n+4)(n+3)(v_0^2+\rho cu_0^2)^2}\left[1+\frac{1}{(v_0^2+\rho cu_0^2)^2}\left\{\frac{16(\rho^4c^4u_0^6+\rho^3c^3u_0^4v_0^2+\rho^2c^2u_0^2v_0^4)}{(v_0^2+\rho cu_0^2)^2}\right.\right.\\
&\quad\left.-2(3\rho^2c^2u_0^2+\rho cv_0^2)\right\}u_2+\frac{1}{(v_0^2+\rho cu_0^2)^2}\left\{\frac{16(\rho^2c^2u_0^4v_0^2+\rho cu_0^2v_0^4+v_0^6)}{(v_0^2+\rho cu_0^2)^2}\right.\\
&\quad\left.-2(3v_0^2+\rho cu_0^2)\right\}v_2+\frac{4}{(v_0^2+\rho cu_0^2)^2}\left\{\frac{2(3\rho^4c^4u_0^5+4\rho^3c^3u_0^3v_0^2+\rho^2c^2u_0v_0^4)}{(v_0^2+\rho cu_0^2)^2}\right.\\
&\quad\left.-\frac{16(\rho^6c^6u_0^9+\rho^3c^3u_0^3v_0^6)}{(v_0^2+\rho cu_0^2)^4}-\rho^2c^2v_0\right\}u_3+\frac{4}{(v_0^2+\rho cu_0^2)^2}\left\{\frac{2(3v_0^5+4\rho cu_0^2v_0^3+\rho^2c^2u_0^4v_0)}{(v_0^2+\rho cu_0^2)^2}\right.\\
&\quad\left.\left.-\frac{16(v_0^9+\rho^3c^3u_0^6v_0^3)}{(v_0^2+\rho cu_0^2)^4}-v_0\right\}v_3+O(m^{-4})+O(n^{-4})\right]
\end{aligned}$$

where

$$\begin{aligned}
&u_0=E(U)=(m-1)/(m+3),\ v_0=E(V)=(n+3)/(n+5),\\
&u_2=E(U-u_0)^2=4(m-1)/(m+4)(m+3)^2,\\
&v_2=E(V-v_0)^2=2(n+3)/(n+6)(n+5)^2,\\
&u_3=E(U-u_0)^3=-8(m-1)(m-5)/(m+5)(m+4)(m+3)^3
\end{aligned}$$

and

$$v_3=E(V-v_0)^3=-4(n+3)(n+1)/(n+7)(n+6)(n+5)^3.$$

The second term is computed similarly, but the moments of Beta distributions are different. That is,

$$\begin{aligned}
&u_0=E(U)=(m+3)/(m+5),\ v_0=E(V)=(n-1)/(n+3),\\
&u_2=E(U-u_0)^2=2(m+3)/(m+6)(m+5)^2,
\end{aligned}$$

Table 4.1 Approximate values of the order $1/m^3$ and $1/n^3$ and of the order $1/m^2$ and $1/n^2$ of variances for the estimator in (2.2) with $c_1=1$ and $c_2=n/m$.

			τ						
m	n	c	.1	.2	.5	1	2	5	10
20	20	c_1	0.971 (0.966)	0.925 (0.923)	0.780 (0.780)	0.595 (0.595)	0.390 (0.390)	0.185 (0.185)	0.097 (0.097)
		c_2	0.971 (0.966)	0.925 (0.923)	0.780 (0.780)	0.595 (0.595)	0.390 (0.390)	0.185 (0.185)	0.097 (0.097)
	25	c_1	0.961 (0.978)	0.911 (0.934)	0.767 (0.783)	0.586 (0.589)	0.386 (0.382)	0.184 (0.180)	0.097 (0.095)
		c_2	0.980 (0.978)	0.935 (0.934)	0.782 (0.783)	0.588 (0.589)	0.381 (0.382)	0.181 (0.180)	0.096 (0.095)
	30	c_1	0.955 (0.993)	0.903 (0.952)	0.759 (0.792)	0.581 (0.589)	0.384 (0.378)	0.184 (0.179)	0.097 (0.094)
		c_2	0.995 (0.993)	0.952 (0.952)	0.791 (0.792)	0.589 (0.589)	0.378 (0.378)	0.179 (0.179)	0.095 (0.094)
25	20	c_1	0.969 (0.950)	0.921 (0.902)	0.772 (0.763)	0.586 (0.589)	0.383 (0.391)	0.182 (0.187)	0.096 (0.098)
		c_2	0.955 (0.950)	0.905 (0.902)	0.763 (0.763)	0.588 (0.589)	0.391 (0.391)	0.187 (0.187)	0.098 (0.098)
	25	c_1	0.958 (0.955)	0.907 (0.905)	0.758 (0.759)	0.577 (0.577)	0.379 (0.379)	0.181 (0.181)	0.096 (0.096)
		c_2	0.958 (0.955)	0.907 (0.905)	0.758 (0·759)	0.577 (0.577)	0.379 (0.379)	0.181 (0.181)	0.096 (0.096)
	30	c_1	0.952 (0.963)	0.898 (0.914)	0.750 (0.761)	0.571 (0.573)	0.376 (0.373)	0.181 (0.178)	0.096 (0.094)
		c_2	0.965 (0.963)	0.914 (0.914)	0.760 (0.761)	0.573 (0.573)	0.373 (0.373)	0.179 (0.178)	0.095 (0.094)
30	20	c_1	0.968 (0.943)	0.919 (0.893)	0.767 (0.757)	0.581 (0.589)	0.379 (0.396)	0.181 (0.190)	0.095 (0.099)
		c_2	0.949 (0.943)	0.896 (0.893)	0.757 (0.757)	0.589 (0.589)	0.396 (0.396)	0.190 (0.190)	0.099 (0.099)
	25	c_1	0.956 (0.945)	0.903 (0.891)	0.753 (0.747)	0.571 (0.573)	0.375 (0.380)	0.180 (0.183)	0.095 (0.096)
		c_2	0.948 (0.945)	0.893 (0.891)	0·747 (0.747)	0.573 (0.573)	0.380 (0.380)	0.183 (0.183)	0.096 (0.096)
	30	c_1	0.950 (0.948)	0.894 (0.894)	0.743 (0.744)	0.565 (0.565)	0.372 (0.372)	0.179 (0.179)	0.095 (0.095)
		c_2	0.950 (0.948)	0.894 (0.894)	0.743 (0.744)	0.565 (0.565)	0.372 (0.372)	0.179 (0.179)	0.095 (0.095)

$$v_2=E(V-v_0)^2=4(n-1)/(n+4)(n+3)^2,$$
$$u_3=E(U-u_0)^3=-4(m+3)(m+1)/(m+7)(m+6)(m+5)^3$$

and

$$v_3=E(V-v_0)^3=-8(n-1)(n-5)/(n+5)(n+4)(n+3)^3.$$

Table 4.1 shows the variance of the combined estimator of the common location θ of two uniform populations $U(\theta-\sigma_x/2, \theta+\sigma_x/2)$ and $U(\theta-\sigma_y/2, \theta+\sigma_y/2)$ on the basis of samples of sizes m and n respectively. We consider estimators $\hat{\theta}^*$ with (i) $c=1$ and (ii) $c=n/m$. Table 4.1 was computed by approximations of the order $1/m^3$ and $1/n^3$ and of the order $1/m^2$ and $1/n^2$ in the parenthesis for

$$2(m+2)(m+1)V(\hat{\theta}^*)/\sigma_x^2$$

for $m=20, 25, 30, n=20, 25, 30$ and $\tau=(n+2)(n+1)\sigma_x^2/(m+2)(m+1)\sigma_y^2$. In fact it is checked by numerical integration that at least two digits of the approximation of (4.1) are accurate. If $c=n/m$ does not satisfy with the condition (3.4), that is, $m=20, n=25$; $m=20, n=30$; $m=25, n=20$ and $m=30, n=20$ in Table 4.1, $c=n/m$ will not give $V(\hat{\theta}^*)\leqq \min(\sigma_x^2/2(m+2)(m+1), \sigma_y^2/2(n+2)(n+1))$, while $c=1$ will give $\hat{\theta}^*$ with that property, unless both m and n are less than 18.

The pattern of the values in Table 4.1 is similar to that for the normal distribution in Khatri and Shah (1974).

5. The case where variance estimators are independent

Let $\hat{\theta}_x$ be the unbiased estimator of a common mean and $\hat{V}_x$ be the estimator of $\eta(\theta)V(\hat{\theta}_x)$ which is independent of $\hat{\theta}_x$ and the parameter θ, where $\eta(\theta)$ is a function of θ. Let $\hat{\theta}_y$ and $\hat{V}_y$ be defined similarly.

We consider the following estimator;

$$\hat{\theta}^{**}=\frac{c_1\hat{V}_y}{c_2\hat{V}_x+c_1\hat{V}_y}\hat{\theta}_x+\frac{c_2\hat{V}_x}{c_2\hat{V}_x+c_1\hat{V}_y}\hat{\theta}_y \tag{5.1}$$

where c_1 and c_2 are positive constants to be suitably chosen. A necessary and sufficient condition for $\hat{\theta}^{**}$ to have smaller variance than $\hat{\theta}_x$ is given.

Proposition 5.1 Let $c=c_2/c_1$, $\rho=V_x/V_y$, $S_x^2=\hat{V}_x/V_x$, $S_y^2=\hat{V}_y/V_y$ and $W=S_x^2/S_y^2$. Suppose that

$$\begin{aligned}&[d/d\rho]E[(1+\rho c^2W^2)(1+\rho cW)^{-2}]\\&=E[d/d\rho][(1+\rho c^2W^2)(1+\rho cW)^{-2}].\end{aligned} \tag{5.2}$$

Then a necessary and sufficient condition for $\hat{\theta}^{**}$ to be uniformly better then $\hat{\theta}_x$ is

$$c\leqq 2E(W)/E(W^2).$$

Note. The assumption (5.2) is equivalent to the following; Let $f(W)$ denote the density of W and put $g_2(W, \rho)=(1+\rho c^2 W^2)(1+\rho c W)^{-2} f(W)$. We assume that g_2 satisfies the following conditions:

(i) g_2 is integrable with respect to W.

(ii) g_2 is differential with respect to ρ.

(iii) There exists an integrable function $\phi_1(W)$ such that $|g_{2\rho}(W, \rho)| \leqq \phi_1(W)$, where $g_{2\rho}(W, \rho)$ denotes the partial derivative of g_2.

Proof. This proposition can be proved along the same line as in the proof of Theorem 2.1 and 2.2.

Remark 5.1 A necessary and sufficient condition for $\hat{\theta}^{**}$ to have smaller than both $\hat{\theta}_x$ and $\hat{\theta}_y$ is

$$E(W^{-2})/2E(W^{-1}) \leqq c \leqq 2E(W)/E(W^2).$$

At first we estimate a common location for the uniform distribution. Let $\hat{\theta}_x = X, \hat{\theta}_y = Y, \hat{\sigma}_x' = X_{(m)} - X_{(1)}, \hat{\sigma}_y' = Y_{(n)} - Y_{(1)}, \hat{V}_x = (m+2)(m+1)\hat{\sigma}_x'^2/m(m-1)$ and $\hat{V}_y = (n+2)(n+1)\hat{\sigma}_y'^2/n(n-1)$. A necessary and sufficient condition for $\hat{\theta}^{**}$ to be better than $\hat{\theta}_x$ is given. Using the density (3.1), it is easy to see that

$$E(W^2) = \frac{M^2 N}{(m+4)(m+3)(n-4)(n-5)}$$

and

$$E(W) = \frac{MN}{(m+2)(m+1)(n-2)(n-3)}$$

where

$$M = \frac{(m+2)(m+1)n(n-1)}{m(m-1)(n+2)(n+1)}$$

and $N = m(m-1)n(n-1)$. By Proposition 5.1, a necessary and sufficient condition for $\hat{\theta}^{**}$ to be better than $\hat{\theta}_x$ is

$$c \leqq \frac{2(m+4)(m+3)m(m-1)(n+2)(n+1)(n-4)(n-5)}{(m+2)^2(m+1)^2 n(n-1)(n-2)(n-3)}.$$

Remark 5.2 A necessary and sufficient condition for $\hat{\theta}^{**}$ to have smaller variance than both $\hat{\theta}_x$ and $\hat{\theta}_y$ is

$$\frac{m(m-1)(m-2)(m-3)(n+2)^2(n+1)^2}{2(m+2)(m+1)(m-4)(m-5)(n+4)(n+3)n(n-1)}$$
$$\leqq c \leqq \frac{2(m+4)(m+3)m(m-1)(n+2)(n+1)(n-4)(n-5)}{(m+2)^2(m+1)^2 n(n-1)(n-2)(n-3)}.$$

The conditions on m and n that there exists c such that the above inequality holds are $m=6, n\geqq 20$; $m\geqq 20, n=6$; $m\geqq 7, n\geqq 8$ and $m\geqq 8, n\geqq 7$. Comparing it with the result in section 3, the latter has a wider range on m and n.

Now we consider the problem of estimating the common mean of the inverse gaussian distribution by using Proposition 5.1. We assume that the density is

$$f(x\,;\theta,\lambda_x)=\exp(-\lambda_x(x-\theta)^2/2\theta^2 x)(\lambda_x/2\pi x^3)^{1/2},\quad \text{if}\quad x>0,$$
$$=0,\ \text{otherwise},$$

where θ and λ_x are both positive. The estimators are

$$\hat{\theta}_x=\bar{X},\ \frac{1}{\hat{\lambda}_x}=\frac{1}{(m-1)}\sum_{i=1}^{m}(X_i-\bar{X}^{-1}).$$

The estimators $\hat{\theta}_y$ and $1/\hat{\lambda}_y$ are defined similarly. It is shown that $\lambda_x\sum_{i=1}^{m}(X_i-\bar{X}^{-1})$ is distributed as a chi-square variable with $m-1$ degrees of freedom and that $\bar{X}$ is independent of $1/\hat{\lambda}_x$ in Folks and Chhihara (1978). And note that the variance of $\hat{\theta}_x$ is $\theta^3/m\lambda_x$.

It is natural to construct the following estimator;

$$\hat{\theta}^{**}=\frac{c_1 m\hat{\lambda}_x}{c_1 m\hat{\lambda}_x+c_2 n\hat{\lambda}_y}\hat{\theta}_x+\frac{c_2 n\hat{\lambda}_y}{c_1 m\hat{\lambda}_x+c_2 n\hat{\lambda}_y}\hat{\theta}_y \tag{5.4}$$

where c_1 and c_2 are positive constants to be suitably chosen. A necessary and sufficient condition for the estimator (5.4) to have smaller variance than $\hat{\theta}_x$ is given. Let

$$S_x=\frac{\lambda_x}{(m-1)}\sum_{i=1}^{m}(X_i-\bar{X}^{-1}),\ S_y=\frac{\lambda_y}{(n-1)}\sum_{i=1}^{n}(Y_i-\bar{Y}^{-1})$$

$W=S_x/S_y$ and $\rho=n\lambda_y/m\lambda_x$, and note that W is distributed as Snedecor's F with $m-1$ degrees of freedom in the numerator and $n-1$ degrees of freedom in the denominator. Therefore we get

$$E(W^2)=\frac{(m+1)(n-1)^2}{(m-1)(n-3)(n-5)}$$

and

$$E(W)=(n-1)/(n-3).$$

By Proposition 5.1, a necessary and sufficient condition for the estimator (5.4) to be better than $\hat{\theta}_x$ is

$$c\leqq 2(m-1)(n-5)/(m+1)(n-1).$$

Acknowledgement

The author would like to thank Prof. M. Sibuya, Prof. Y. Washio, and Dr. N. Shinozaki of Keio University for their helpful suggestions and criticisms.

REFERENCES

[1] Bhattacharya, C. G. (1981) Estimation of a common location, *Commun. Statist.*, **A 10(10)**, 955-961.
[2] Cohen, Arthur (1976) Combining estimates of location, *J. Amer. Statist. Assoc.*, **71**, 172-175.
[3] Folks, J. L., and Chhikara, R. S. (1978) The inverse gaussian distribution and its application- a review, *Journal of the Royal Statistical Society, Ser.* **B**, **40**, 263-289.
[4] Graybill, F. A. and Deal, R. B. (1959) Combining unbiased estimators, *Biometrics* **15**, 543-550.
[5] Khatri, C. G. and Shah, K. R. (1974) Estimation of location parameters from two linear models under normality, *Commun. Statist.*, **3(7)**, 647-663.
[6] Hogg, R. V. (1960) On conditional expectation of location statistics, *J. Amer. Statist. Assoc.*, **55**, 714-717.

Appendix

We state a version of Hogg's theorem (1960) for completeness. The proof is similar to Hogg's (1960) and omitted. Let $X_1, X_2, \cdots, X_n$ be a random sample from a distribution. The statistic $T(X_1, X_2, \cdots, X_n)$ is an odd location statistic in the sense that, for all real $x_1, x_2, \cdots, x_n$,

(a) $$T(x_1+h, \cdots, x_n+h)=T(x_1, \cdots, x_n)+h, \text{ for every } h, \text{ and}$$

(b) $$T(-x_1, \cdots, -x_n)=-T(x_1, \cdots, x_n).$$

The statistic $S(X_1, X_2, \cdots, X_n)$ is an even location-free statistic in the sense that, for all real $x_1, x_2, \cdots, x_n$,

(c) $$S(x_1+h, \cdots, x_n+h)=S(x_1, \cdots, x_n), \text{ for every } h, \text{ and}$$

(d) $$S(-x_1, \cdots, -x_n)=S(x_1, \cdots, x_n).$$

And put
$$T=t(X_1, X_2, \cdots, X_n),$$
$$S=s(X_1, X_2, \cdots, X_n)$$

and

$$U_i=u_i(X_1, X_2, \cdots, X_n), \quad i=1, 2, \cdots, n-2.$$

In this terminology, we state the following theorem.

Theorem. Let $X_1, X_2, \cdots, X_n$ be a random sample from a distribution that is symmetric about a point θ. If $T(X_1, X_2, \cdots, X_n)$ is an odd location statistic and $S(X_1, X_2, \cdots X_n)$ is an even location-free statistic, and if they are regular in the sense that there exist statistics $U_1, U_2, \cdots, U_{n-2}$ such that Jacobian

$$|\partial(t, s, u_1, \cdots, u_{n-2})/\partial(x_1, x_2, \cdots, x_n)|$$

is not zere almost surely, then the conditional distribution of T given $S=s$ is symmetric about the parameter θ.