Some consideration on estimation of population variance due to the use of pooling data

Asano, Choichiro
Kyushu University | Shionogi Pharamaceutical Co.

Sugimura, Masahiko
Kumamoto Women's University

https://doi.org/10.5109/12999

出版情報：統計数理研究. 10 (1/2), pp.33-44, 1961-10. Research Association of Statistical Sciences

SOME CONSIDERATION ON ESTIMATION OF POPULATION VARIANCE DUE TO THE USE OF POOLING DATA

By

Chooichiro ASANO and Masahiko SUGIMURA

(*Received* December 20, 1960)

§ 1. Introduction

The statistical inferences based on sometimes pooling after a preliminary test of validity, that is, the inference problems on T.E. type have been given in a thought of successive procedure of statistical inferences by many authors, Bancroft [1], Bennett [1], Kitagawa [1]~[4] and others. And recently Asano [1], one of authors of this paper, discussed seven typical statistical procedures and their meaning from a view-point of making some inferences on the validity test of observations and some uses of previous information, and proposed in each procedure that the switching constants λ and Λ may be chosen with a minimax principle for values of a mean square error like in game theory.

While, Huntsberger [1] has proposed an extensive and special method of always pooling data independently of the thought of the successive process of statistical inferences. Namely he dealt an inference based on a following conception.

Let $N(\theta_i, \sigma_i^2)$ be a normal population with a known value of σ_i^2, $(i=1, 2)$, and as an estimator of θ_1 he gave $W(T)$ defined by

$$W(T)=\phi(T)\hat{\theta}_1+\{1-\phi(T)\}\frac{\sigma_2^2\hat{\theta}_1+\sigma_1^2\hat{\theta}_2}{\sigma_1^2+\sigma_2^2}, \tag{1.1}$$

where the statistic T and $\hat{\theta}_i$ were an unbiased estimate of $(\theta_1-\theta_2)/\sqrt{\sigma_1^2+\sigma_2^2}$ and θ_i respectively and $\phi(T)$ was called a weighting function. And there the weighted method was corresponding with the sometimes pooling method on the basis of giving such a definition as that (i) $\phi(T)=0$ in case when $T\subset A_\alpha$ (acceptance region), (ii) $\phi(T)=1$ in case when $T\subset R_\alpha$ (rejection region) and (iii) $0\leq\phi(T)\leq 1$ for all T.

Thus in order to determine whether or not $W(T)$ offers any advantages over sometimes pooling method as an estimator for θ_1, the mean square deviation was used as a criterion of goodness and was illustrated by some numerical examples.

The object of this paper is to discuss the inference problems of population variance in the light of both principle of successive process of inferences and principle of weighted estimator and to reconsider the advantages of each principle and is to propose some recommendations from the view-point of the efficiency of each estimator.

Through this paper, the weighted estimator of a population variance is defined in the following way. Let u_i^2 be a sample unbiased variance given by the i-th random observation O_{n_i} of size n_i drawn from a normal population $N(\mu_i, \sigma_i^2)$, $i=1, 2$, and let ρ be a ratio σ_2^2/σ_1^2, where the value of ρ is unknown to us but is assumed that $\rho \geq 1$. Then a pooling estimator $W(u_1^2, F)$ for σ_1^2 is defined by

$$W(u_1^2, F) = \phi(F)u_1^2 + \{1-\phi(F)\}\frac{\nu_1+\nu_2 F}{\nu_1+\nu_2}u_1^2, \tag{1.2}$$

where the weighted function $\phi(F)$ is only a function of a statistic $F=u_2^2/u_1^2$ and $\nu_i=N_i-1$ $(i=1,2)$. Section 2 of this paper is concerned with the inference having a previous information and Section 3 is concerned with the inference having not any previous information.

In conclusion, the authors wish to express their heartiest thanks to Prof. T. Kitagawa for his kind advice and valuable suggestions in connection with this work.

§ 2. Case 1: The inference having a certain previous information about ρ

In this section, let us consider the cases when we have a certain previous information concerning to a non-negative value of ρ.

2.1. A trivial case

If a value of ρ is known exactly to us and is assured of holding the truth on the basis of physical science, it is well-known that we ought to take a following statistic W_0 as a best estimate of σ_1^2,

$$W_0 = \frac{\nu_1 u_1^2 + \nu_2 u_2^2/\rho}{\nu_1+\nu_2}. \tag{2.1}$$

In a situation of the weighted estimator (1.2), this is easily obtained from a result of $\phi_0(F)=0$ as a most desirable estimator of σ_1^2. In this situation the sometimes pooling method may be senseless to apply.

2.2. A case having an uncertain information about ρ

Now suppose that we have a previous information with uncertainty and assume that ρ_0 is put as a value of ρ. That is to say, in spite of being assumed that $\rho=\rho_0$ like in a case when a value of ρ is known to us, we may be confronted very often and usually with the cases that we are unable to assure ourself of that a value of ρ is exactly true in the statistical situation of inferences. Then the sometimes pooling procedure may have an active meaning.

Under such a circumstance, let us suppose that ρ_0 is assumed as a value of ρ and now is put less than or equal to the true value ρ without the loss of generality, that is, $\rho/\rho_0 \geqq 1$.

2.2.1. Weighted method

On the basis of the principle of (1.1) and (1.2), we may and now shall define a pooling estimator $W_0(u_1^2, F)$ for σ_1^2 in the following way:

$$W(u_1^2, u_2^2) = \phi(u_2^2/u_1^2)u_1^2 + \{1-\phi(u_2^2/u_1^2)\}\frac{\nu_1 u_1^2 + \nu_2 u_2^2/\rho_0}{\nu_1+\nu_2}. \tag{2.2}$$

Thus we shall consider an estimate W_A defined by

$$W_A = A u_1^2 + (1-A)\frac{\nu_1 u_1^2 + \nu_2 u_2^2/\rho_0}{\nu_1+\nu_2}, \tag{2.3}$$

where a constant A is considered to be a function of ρ/ρ_0 fixed. Then the mean square deviation D_A^2 about σ_1^2 is given by

$$\frac{\sigma_1^4}{(\nu_1+\nu_2)^2}\left\{\frac{\nu_2}{\nu_1}K_0A^2 - 2\frac{\nu_2}{\nu_1}H_0A + G_0\right\}, \tag{2.4}$$

where we put that

$$\begin{cases} K_0 = \nu_1(\nu_2+2)(\rho/\rho_0)^2 - 2\nu_1\nu_2(\rho/\rho_0) + (\nu_1+2)\nu_2, \\ H_0 = \nu_1(\nu_2+2)(\rho/\rho_0)^2 - 2\nu_1\nu_2(\rho/\rho_0) + \nu_1(\nu_2-2), \\ G_0 = \nu_2(\nu_2+2)(\rho/\rho_0)^2 - 2\nu_2^2(\rho/\rho_0) + (2\nu_1+\nu_2^2), \end{cases} \tag{2.5}$$

and where the value of D_A^2 is minimized in case when $A=A_0$ given by

$$A_0 = 1 - \frac{2(\nu_1+\nu_2)}{\nu_1\nu_2(\rho/\rho_0-1)^2 + 2(\nu_1(\rho/\rho_0)^2+\nu_2)}. \tag{2.6}$$

In the present case, however, there may be no way that $\rho=\rho_0$ like in **2.1**, because of that the true value ρ is unknown. Hence we may and now shall adopt a following weighted estimator W_0 for a case $A_0=0$.

$$W_0 = \frac{\nu_1 u_1^2 + \nu_2 u_2^2/\rho_0}{\nu_1+\nu_2}. \tag{2.7}$$

Then the mean square deviation D_0^2 of W_0 about σ_1^2 is obtained as follows:

$$D_0^2 = \frac{\sigma_2^4}{(\nu_1+\nu_2)^2}\{\nu_2^2(1-\rho/\rho_0)^2 + 2(\nu_1+\nu_2(\rho/\rho_0)^2)\}. \tag{2.8}$$

2.2.2. Sometimes pooling method

Under the same circumstances, we are now concerned with the rule of the inference procedure for σ_1^2 formulated in the following way:

(i) Let u_i^2 be sample unbiassed variance mutually independently for $i=1, 2$.

(ii) Let the statistic F_0 be defined by $F_0 = u_2^2/u_1^2\rho_0$.

(iii) Let us define the statistic W_{sp} in the following way,

(a) $W_{sp} = \dfrac{\nu_1 u_1^2 + \nu_2 u_2^2/\rho_0}{\nu_1+\nu_2}$, if $F_0 \leqq \lambda$,

(b) $W_{sp} = u_1^2$, if $F_0 > \lambda$,

where the switching constant λ is certain prescribed constant corresponding to a value of a certain α-point value of F-distribution with the pair of the degrees of freedom (ν_2, ν_1).

Then the mean square deviation of W_{sp} about σ_1^2 is given by

$$(2.9)\quad D_{sp}^2 = \sigma_1^4 \Big[\frac{2}{\nu_1} + \frac{1}{(\nu_1+\nu_2)^2} \Big\{ \nu_1(\nu_1+2) I_\varphi\Big(\frac{\nu_2}{2}, \frac{\nu_1}{2}+2\Big)$$

$$+ 2\nu_1\nu_2 I_\varphi\Big(\frac{\nu_2}{2}+1, \frac{\nu_1}{2}+1\Big)(\rho/\rho_0) + \nu_2(\nu_2+2) I_\varphi\Big(\frac{\nu_2}{2}+2, \frac{\nu_1}{2}\Big)(\rho/\rho_0)^2 \Big\}$$

$$-\Big(1+\frac{2}{\nu_1}\Big) I_\varphi\Big(\frac{\nu_2}{2}, \frac{\nu_1}{2}+2\Big) - \frac{2\nu_2}{\nu_1+\nu_2} \Big\{ I_\varphi\Big(\frac{\nu_2}{2}+1, \frac{\nu_1}{2}\Big)(\rho/\rho_0)$$

$$- I_\varphi\Big(\frac{\nu_2}{2}, \frac{\nu_1}{2}+1\Big)\Big\}\Big] ,$$

where we put that $\varphi = \nu_2\lambda/\{(\nu_1\rho/\rho_0)+\nu_2\lambda\}$ and $I_\varphi(m, n) = B_\varphi(m, n)/B(m, n)$.

In this case the weighted method given by **2.2.1** is corresponding to the sometimes pooling method in case that $\lambda=\infty$.

2.2.3. Numerical comparison between W_0 and W_{sp} on efficiency

In order to investigate the properties of these two estimators from a view-point of the efficiency, now we shall attempt to obtain numerical considerations by evaluations of D_0^2/σ_1^4 and D_{sp}^2/σ_1^4 as a function of ρ/ρ_0 for certain values of the pair of degrees of freedom (ν_2, ν_1) : Example 1 $(\nu_2 = \nu_1 = 9)$ and Example 2 $(\nu_2 = 20,\ \nu_1 = 4)$. The figures of efficiencies of each example are tabulated respectively and the behaviours of those are given in Fig. 1 and Fig. 2.

Example 1: Efficiencies for $\nu_2=\nu_1=9$

ρ/ρ_0		1	2	3	4	5
D_0^2/σ_1^4		.111	.528	1.556	3.194	5.444
D_{sp}^2/σ_1^4	$\lambda=\infty$	.111	.528	1.556	3.194	5.444
	$\lambda_{.01}=5.35$	.116	.480	1.047	1.504	1.833
	$\lambda_{.05}=3.18$	.126	.374	.576	.645	.648
	$\lambda_{.25}=1.59$	.139	.185	.253	.237	.236
	$\lambda=1$	.138	.184	.206	.215	.218
	$\lambda=0$	.222	.222	.222	.222	.222

Example 2: Efficiencies for ($\nu_2=20$, $\nu_1=4$)

ρ/ρ_0		1	2	3	4	5
D_0^2/σ_1^4		.083	.986	3.417	7.375	12.861
D^2_{sp}/σ_1^4	$\lambda=\infty$	.083	.986	3.417	7.375	12.861
	$\lambda_{.01}=14.02$	.088	.964	3.153	6.450	10.025
	$\lambda_{.05}=5.8$	.112	.897	2.430	3.050	5.312
	$\lambda_{.25}=2.08$	.185	.429	.832	.877	.908
	$\lambda=1$	.197	.331	.426	.469	.487
	$\lambda=0$	.500	.500	.500	.500	.500

Fig. 1: The behaviours of efficiencies for Example 1 ($\nu_2=\nu_1=9$)

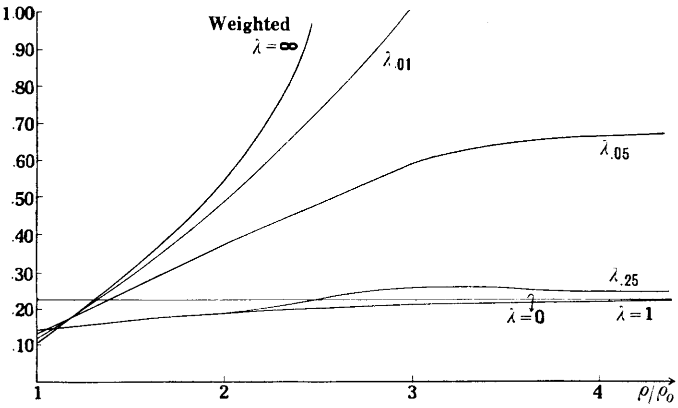
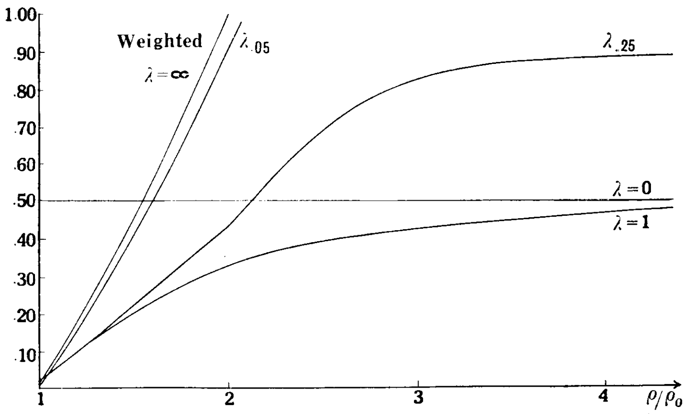


Fig. 2: The behaviours of efficiencies for Example 2 ($\nu_2=20$, $\nu_1=4$)



So far as these numerical data are concerned, we may be able to make the following observations which might be suggestive to our statistical method in more general situation.

(1) In considerable wide range of ρ/ρ_0, that is to say, supposing that there may be somewhat a large difference between an applied ρ_0 and the

true value of ρ, the sometimes pooling method is more efficient than the weighted estimator method. Above all, it may be found in our examples that a sometimes pooling method defined by $\lambda=1$, which is nearly corresponding to 50 percent point of F-distribution in our examples, is more efficiency than others.

In contrast to the trivial case of **2. 1**, this may be concluded that there is unconsciously a reasonable consideration for us under the circumstance of this section in case when a sometimes pooling method is applied regardless of that a value of ρ is assumed.

(2) Under the consideration of the restricted neighborhood of $\rho/\rho_0=1$, it may be prefer to apply the weighted method, which is equivalent to a special case defined by sometimes pooling method for $\lambda=\infty$, is more efficient than the sometimes pooling method for general values of λ.

§ 3. Case 2: The inference having not any previous information to be assumed about ρ

Let us consider in this section a case when we have nothing of the previous information for a value of ρ, and we are principally interested in a subject of this case.

3.1. Weighted method

Let us consider the weighted estimator of σ_1^2 on the basis of the principle of (1.2) in the following way:

$$W(u_1^2, F)=\phi(F)u_1^2+\{1-\phi(F)\}\frac{\nu_1+\nu_2 F}{\nu_1+\nu_2}u_1^2\ , \tag{3.1}$$

where the weighted function $\phi(F)$ is only a function of F defined by u_2^2/u_1 and we assume for convenience that ν_1 and ν_2 of the degrees of freedom are two even integers.

Thus we shall consider an estimate W_A given by

$$W_A=Au_1^2+(1-A)\frac{\nu_1+\nu_2 F}{\nu_1+\nu_2}u_1^2\ , \tag{3.2}$$

where a constant A is considered to be a function of ρ fixed.

Then the mean square deviation D_A^2 about σ_1^2 is given by

$$D_A^2=\frac{\sigma_1^2}{(\nu_1+\nu_2)^2}\left\{\frac{\nu_2}{\nu_1}K_0'A^2-2\frac{\nu_2}{\nu_1}H_0'A^2+G_0'\right\}\ , \tag{3.3}$$

where we put

$$\left\{\begin{array}{l} K_0'=\nu_1(\nu_2+2)\rho^2-2\nu_1\nu_2\rho+(\nu_1+2)\nu_2, \\ H_0'=\nu_1(\nu_2+2)\rho^2-2\nu_1\nu_2\rho+\nu_1(\nu_2-2), \\ G_0'=\nu_2(\nu_2+2)\rho^2-2\nu_2^2\rho+(2\nu_1+\nu_2^2)\ , \end{array}\right. \tag{3.4}$$

and the value of D_A^2 is minimized in case when $A=A_0=H_0'/K_0'$.

And A_0 is summarized as follows,

$$(3.5)\qquad A_0 = 1 - \frac{2(\nu_1+\nu_2)}{\nu_1(\nu_2+2)(\rho-\nu_2/(\nu_2+2))^2+2\nu_2(\nu_1+\nu_2+2)/(\nu_2+2)}\ .$$

In the present case when a value of ρ is to be unknown, it may be desirable for us to choose an unbiassed estimate of A_0 for the purpose of giving explicit $\phi(F)$, but this is so complicate that we shall avoid the obstacle. Now we shall substitute $(\nu_1-2)F/\nu_1$ for ρ of (3.5) as a result of some trials, where the statistic $(\nu_1-2)F/\nu_1$ is an unbiassed estimate of ρ, and define an estimate $\phi_0(F)$ of A_0. And there we shall add a restriction given by that $\phi_0(F)=0$ in case when the value of $\phi_0(F)$ is negative on the basis of a non-negative property of weighting function. This restriction is also equivalent to that $\phi_0(F)=0$ in case when $(\nu_1-2)F/\nu_1<1$.

Thus the weighting function $\phi_0(F)$ as an estimate of A_0 shall be defined in conclusion as follows:

$$(3.6)\qquad \begin{aligned}\phi_0(F) &= 1-\frac{2\nu_1(\nu_1+\nu_2)}{b_1(F+b_2)^2+b_3}\ , && \text{if } (\nu_1-2)F/\nu_1 \geqq 1\ ,\\ &=0 && \text{, if } (\nu_1-2)F/\nu_1<1\ ,\end{aligned}$$

where $0 \leqq \phi_0(F) \leqq 1$ and we put

$$(3.7)\qquad \begin{cases} b_1=(\nu_1-2)^2(\nu_2+2),\ b_2=-\nu_1\nu_2/(\nu_1-2)(\nu_2-2),\\ b_3=2\nu_1\nu_2(\nu_1+\nu_2+2)/(\nu_2+2).\end{cases}$$

Thus our weighted estimator $W_0(u_1^2, F)$ of σ_1^2 is obtained by substituting $\phi_0(F)$ for $\phi(F)$ of (3.1) in the following way.

$$(3.8)\qquad W_0(u_1^2,F)=\phi_0(F)u_1^2+\{1-\phi_0(F)\}\frac{\nu_1+\nu_2 F}{\nu_1+\nu_2}u_1^2\ .$$

(i) The expectation of $W_0(u_1^2, F)$

Since u_1^2 and u_2^2 are independently distributed, the joint distribution of u_1^2 and u_2^2 is given by

$$(3.9)\qquad K(u_1^2)^{p_1-1}(u_2^2)^{p_2-1}\exp.\{-(\alpha_1u_1^2+\alpha_2u_2^2)\}du_1^2du_2^2\ ,$$

where we put that $K=\alpha_1^{p_1}\alpha_2^{p_2}/\Gamma(p_1)\Gamma(p_2)$, $\alpha_i=\nu_i/2\sigma_i^2$ and $p_i=\nu_i/2$ for $i=1, 2$.

Then the expectation of $W_0(u_1^2, F)$ is denoted by

$$(3.10)\qquad E\{W_0(u_1^2,F)\}$$
$$=K\Big[\frac{1}{\nu_1+\nu_2}\int_0^\theta\int_0^\infty(\nu_1+\nu_2F)F^{p_2-1}(u_1^2)^{p_1+p_2}\exp.\{-(\alpha_1+\alpha_2F)u_1^2\}du_1^2dF$$
$$+\int_\theta^\infty\int_0^\infty\Big\{F^{p_2-1}-\frac{2\nu_1\nu_2(1-F)F^{p_2-1}}{b_1(F+b_2)^2+b_3}\Big\}(u_1^2)^{p_1+p_2}\exp.\{-(\alpha_1+\alpha_2F)u_1^2\}du_1^2dF\Big],$$

where $\theta=\nu_1/(\nu_1-2)$, and integrating out u_1^2 and F partially, we obtain

(3.11) $E\{W_0(u_1^2, F)\}$

$$=\sigma_1^2\left\{1+\frac{\nu_2}{\nu_1+\nu_2}\left(I_\kappa\left(\frac{\nu_2}{2}+1, \frac{\nu_1}{2}\right)\rho-I_\kappa\left(\frac{\nu_2}{2}, \frac{\nu_1}{2}+1\right)\right)\right\}$$

$$-2\nu_1\nu_2K\Gamma\left(\frac{\nu_1}{2}+\frac{\nu_2}{2}+1\right)\int_\theta^\infty\frac{(1-F)F^{p_2-1}\,dF}{(\alpha_1+\alpha_2F)^{p_1+p_2+1}\{b_1(F+b_2)^2+b_3\}},$$

where we put that $I_\kappa(m, n)=B_\kappa(m, n)/B(m, n)$, $\kappa=\nu_2/((\nu_1-2)\rho+\nu_2)$.

Furthermore, for a purpose of numerical calculations, using the following asymptotic formulae

$$(3.12)\quad\begin{cases} J(0,-1)=\dfrac{2}{\sqrt{4a-b^2}}\arctan\dfrac{2at+b}{\sqrt{4a-b^2}}, \\ J(1,-1)=\dfrac{1}{2a}\left\{\log|at^2+bt+1|-bJ(0,-1)\right\}, \\ J(m,-1)=\dfrac{1}{(m-1)a}\left[t^{m-1}-(m-1)\{bJ(m-1,-1)+J(m-2,-1)\}\right], \\ \qquad (m\geqq 2), \\ J(l,-2)=\dfrac{1}{4a-b^2}\left\{\dfrac{2at+b}{at^2+bt+1}t^l-2(l-1)aJ(l,-1)-lbJ(l-1,-1)\right\}, \\ \qquad (l\geqq 1), \end{cases}$$

we obtain that

(3.13) $E\{W_0(u_1^2, F)\}$

$$=\sigma_1^2\left[1+\frac{\nu_2}{\nu_1+\nu_2}\left\{I_\kappa\left(\frac{\nu_2}{2}+1, \frac{\nu_1}{2}\right)\rho-I_\kappa\left(\frac{\nu^2}{2}, \frac{\nu_1}{2}+1\right)\right\}\right]$$

$$-\frac{2\nu_1(\nu_1/\nu_2)^{p_1}(\nu_1+\nu_2)\rho^{p_1+1}}{(\nu_1-2)^2(\nu_2+2)B\left(\frac{\nu_1}{2}, \frac{\nu_2}{2}\right)}\left\{\sum_{i=0}^{r_2-1}\binom{p_2-1}{i}(-\nu_1\rho/\nu_2)^i J_\tau\left(\frac{\nu_1}{2}+2+i,-1\right)\right.$$

$$\left.-\sum_{i=0}^{\nu_2}\binom{p_2}{i}(-\nu_1\rho/\nu_2)^i J_\tau\left(\frac{\nu_1}{2}+1+i,-1\right)\right\},$$

where we put $\tau=(\nu_1-2)\nu_2/\{\nu_1\nu_2+\nu_1(\nu_1-2)\rho\}$ and $J_\tau(m,n)=[J(m,n)]_0^\tau$

$$(3.14)\quad\begin{cases} a=\dfrac{2\nu_1\nu_2(\nu_1+\nu_2+2)}{(\nu_1-2)^2(\nu_2+2)}-\nu_1^2\left(\dfrac{\rho}{\nu_2}+\dfrac{\nu_2}{(\nu_1-2)(\nu_2+2)}\right)^2, \\ b=-2\nu_1\left(\dfrac{\rho}{\nu_2}+\dfrac{\nu_2}{(\nu_1-2)(\nu_2+2)}\right),\quad 4a-b^2=\dfrac{8\nu_1\nu_2(\nu_1+\nu_2+2)}{(\nu_1-2)^2(\nu_2+2)^2}. \end{cases}$$

(ii) The expectation of $\{W_0(u_1^2, F)\}^2$

By making use the same manner, we obtain that

$$(3.15)\quad E[\{W_0(u_1^2,F)\}^2]$$

$$=\sigma_1^4\Big[\frac{1}{(\nu_1+\nu_2)^2}\Big\{\nu_1(\nu_1+2)\,I_\kappa\Big(\frac{\nu_2}{2},\ \frac{\nu_1}{2}+2\Big)+2\rho\nu_1\nu_2 I_\kappa\Big(\frac{\nu_2}{2}+1,\ \frac{\nu_1}{2}+1\Big)$$

$$+\nu_2(\nu_2+2)\,I_\kappa\Big(\frac{\nu_2}{2}+2,\ \frac{\nu_1}{2}\Big)\rho^2\Big\}+\Big(1+\frac{2}{\nu_1}\Big)\Big\{1-I_\kappa\Big(\frac{\nu_2}{2},\ \frac{\nu_1}{2}+2\Big)\Big\}\Big]$$

$$+4\nu_1\nu_2K\Gamma\Big(\frac{\nu_1}{2}+\frac{\nu_2}{2}+2\Big)\Big[\nu_1\nu_2\int_\theta^\infty\frac{(1-F)^2F^{p_2-1}}{(\alpha_1+\alpha_2F)^{p_1+p_2+2}\{b_1(F+b_2)^2+b_3\}^2}dF$$

$$-\int_\theta^\infty\frac{(1-F)\ F^{p_2-1}}{(\alpha_1+\alpha_2F)^{p_1+p_2+2}\{b_1(F+b_2)^2+b_3\}}dF\Big]\ .$$

(iii) The mean square deviation D_0^2 of $W_0(u_1^2,F)$ about σ_1^2

We obtain by combining (i) and (ii) that

$$(3.16)\quad D_0^2=\ \sigma_1^4\Big[\frac{2}{\nu_1}+\frac{1}{(\nu_1+\nu_2)^2}\Big\{\nu_1(\nu_1+2)\,I_\kappa\Big(\frac{\nu_2}{2},\ \frac{\nu_1}{2}+2\Big)$$

$$+\ 2\nu_1\nu_2\ I_\kappa\Big(\frac{\nu_2}{2}+1,\ \frac{\nu_1}{2}+1\Big)\rho+\nu_2(\nu_2+2)\,I_\kappa\Big(\frac{\nu_2}{2}+2,\ \frac{\nu_1}{2}\Big)\rho^2\Big\}$$

$$-\Big(1+\frac{2}{\nu_1}\Big)\,I_\kappa\Big(\frac{\nu_2}{2},\ \frac{\nu_1}{2}+2\Big)-\frac{2\nu_2}{\nu_1+\nu_2}\Big\{I_\kappa\Big(\frac{\nu_2}{2}+1,\frac{\nu_1}{2}\Big)\,\rho-I_\kappa\Big(\frac{\nu_2}{2},\frac{\nu_1}{2}+1\Big)\Big\}\Big]$$

$$+4\nu_1\nu_2K\Gamma\Big(\frac{\nu_1}{2}+\frac{\nu_2}{2}+2\Big)\Big[\nu_1\nu_2\int_\theta^\infty\frac{(1-F)^2F^{p_2-1}}{(\alpha_1+\alpha_2F)^{p_1+p_2+2}\{b_1(F+b_2)^2+b_3\}^2}\ dF$$

$$-\int_\theta^\infty\frac{(1-F)F^{p_2-1}}{(\alpha_1+\alpha_2F)^{p_1+p_2+2}\{b_1(F+b_2)^2+b_3\}}\ dF\Big]$$

$$+4\nu_1\nu_2K\Gamma\Big(\frac{\nu_1}{2}+\frac{\nu_2}{2}+1\Big)\,\sigma_1^2\int_\theta^\infty\frac{(1-F)F^{p_2-1}}{(\alpha_1+\alpha_2F)^{p_1+p_2+1}\{b_1(F+b_2)^2+b_3\}}\quad dF.$$

For the purpose of numerical calculations, (3.16) is denoted for the sake of the asymptotic formulae (3.12) as follows:

$$(3.17)\quad D_0^2/\sigma_1^4=\frac{2}{\nu_1}+\frac{1}{(\nu_1+\nu_2)^2}\Big\{\nu_1(\nu_1+2)\ I_\kappa\Big(\frac{\nu_2}{2},\ \frac{\nu_1}{2}\Big)+2\nu_1\nu_2 I_\kappa\Big(\frac{\nu_2}{2}+1,\frac{\nu_1}{2}+1\Big)\rho$$

$$+\nu_2(\nu_2+2)I_\kappa\Big(\frac{\nu_2}{2}+2,\frac{\nu_1}{2}\Big)\rho^2\Big\}-\Big(1+\frac{2}{\nu_1}\Big)\ I_\kappa\Big(\frac{\nu_2}{2},\ \frac{\nu_1}{2}+2\Big)$$

$$-\frac{2\nu_2}{\nu_1+\nu_2}\Big\{I_\kappa\Big(\frac{\nu_2}{2}+1,\frac{\nu_1}{2}\Big)\ \rho\ -I_\kappa\Big(\frac{\nu_2}{2},\ \frac{\nu_1}{2}+1\Big)\Big\}$$

$$+\frac{4(\nu_1/\nu_2)^{p_1}(\nu_1+\nu_2)\,\rho^{p_1+1}}{(\nu_1-2)^2\,(\nu_2+2)B(p_1,\,p_2)}\Bigg(\nu_1(\nu_1+\nu_2+2)\,\rho\Bigg[\frac{\nu_1}{(\nu_1-2)^2(\nu_2+2)}$$

$$\Bigg\{\sum_{i=0}^{p_2-1}\binom{p_2-1}{i}(-\nu_1\rho/\nu_2)^i J_\tau\Big(\frac{\nu_1}{2}+5+i,\,-2\Big)-2\sum_{i=0}^{p_2}\binom{p_2}{i}(-\nu_1\rho/\nu_2)^i$$

$$J_\tau\Big(\frac{\nu_1}{2}+4+i,\,-2\Big)+\sum_{i=0}^{p_2+1}\binom{p_2+1}{i}\,(-\nu_1\rho/\nu_2)^i\,J_\tau\Big(\frac{\nu_1}{2}+3+i,\,-2\Big)\Bigg\}$$

$$-\frac{1}{\nu_2}\Bigg\{\sum_{i=0}^{p_2-1}\binom{p_2-1}{i}\,(-\nu_1\rho/\nu_2)^i J_\tau\Big(\frac{\nu_1}{2}+3+i,\,-1\Big)$$

$$-\sum_{i=0}^{p_2}\binom{p_2}{i}\,(-\nu_1\rho/\nu_2)^i\,J_\tau\Big(\frac{\nu_1}{2}+2+i,\,-1\Big)\Bigg\}\Bigg]$$

$$+\nu_1\Bigg[\sum_{i=0}^{p_2-1}\binom{p_2-1}{i}\,(-\nu_1\rho/\nu_2)^i\,J_\tau\Big(\frac{\nu_1}{2}+2+i,\,-1\Big)$$

$$-\sum_{i=0}^{p_2}\binom{p_2}{i}\,(-\nu_1\rho/\nu_2)^i\,J_\tau\Big(\frac{\nu_1}{2}+1+i,\,-1\Big)\Bigg]\Bigg).$$

3.2. Sometimes pooling method

Under the same circumstances as **3.1.**, we are now concerned with the rule of the inference procedure for σ_1^2 formulated in the following way:

(i) Let u_i^2 be sample unbiassed variance mutually independently for $i=1,2$.

(ii) Let the statistic F_0 be defined by $F_0=u_2^2/u_1^2$.

(iii) Let us define the statistic W_{sp} in the following way,

(a) $W_{sp}=(\nu_1+\nu_2F_0)u_1^2/(\nu_1+\nu_2)$, if $F_0\leqq\lambda$,

(b) $W_{sp}=u_1^2$, if $F_0>\lambda$,

where the switching constant λ is defined by the same as **2.2.2.**

Then the expectation $E\{W_{sp}\}$ and the mean square deviation D_{sp}^2 are obtained from the results reported by Bancroft [1] as follows:

$$(3.18)\quad E\{W_{sp}\}=\sigma_1^2\Big[1+\frac{\nu_2}{\nu_1+\nu_2}\Big\{I_{\varphi'}\Big(\frac{\nu_2}{2}+1,\,\frac{\nu_1}{2}\Big)\rho-I_{\varphi'}\Big(\frac{\nu_2}{2}\,,\,\frac{\nu_1}{2}+1\Big)\Big\}\Big],$$

$$(3.19)\quad D_{sp}^2=\sigma_1^4\Big[\frac{2}{\nu_1}+\frac{1}{(\nu_1+\nu_2)^2}\Big\{\nu_1(\nu_1+2)I_{\varphi'}\Big(\frac{\nu_2}{2}\,,\,\frac{\nu_1}{2}+2\Big)$$

$$+2\nu_1\nu_2I_{\varphi'}\Big(\frac{\nu_2}{2}+1,\,\frac{\nu_1}{2}+1\Big)\rho+\nu_2(\nu_2+2)I_{\varphi'}\Big(\frac{\nu_2}{2}+2,\,\frac{\nu_1}{2}\Big)\rho^2\Big\}$$

$$-\Big(1+\frac{2}{\nu_1}\Big)I_{\varphi'}\Big(\frac{\nu_2}{2}\,,\,\frac{\nu_1}{2}+2\Big)-\frac{2\nu_2}{\nu_1+\nu_2}\Big\{I_{\varphi'}\Big(\frac{\nu_2}{2}+1,\,\frac{\nu_1}{2}\Big)\rho$$

$$-I_{\varphi'}\Big(\frac{\nu_2}{2}\,,\,\frac{\nu_1}{2}+1\Big)\Big\}\Big],$$

where $\varphi' = \nu_2\lambda/(\nu_1\rho + \nu_2\lambda)$.

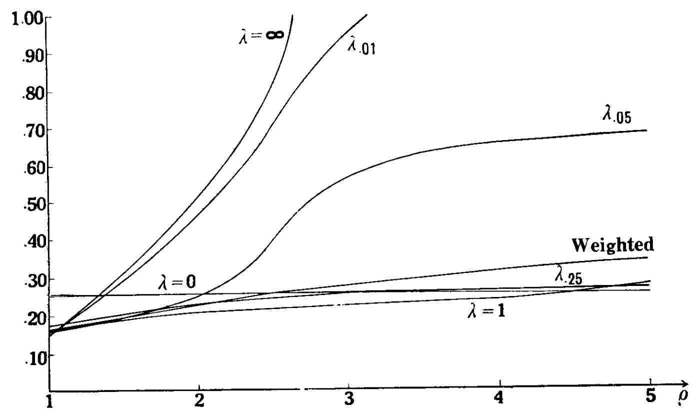
3.3. Numerical comparison between W_0 and W_{sp} on efficiency

In order to investigate the properties of these two estimators from a view-point of the efficiency, now we shall attempt to give numerical considerations like in section 2 by evaluations of D_0^2/σ_1^4 and D_{sp}^2/σ_1^4 as the functions of ρ for certain value of the pair of degrees of freedom (ν_2, ν_1) : Example 3 $(\nu_2=6, \nu_1=8)$. The figures of efficiencies of this example are tabulated and the behaviours of those are given in Fig. 3.

Example 3: Efficiencies for $(\nu_2=6,\ \nu_1=8)$

	ρ	1	2	3	4	5
D_0^2/σ_1^4		.157	.230	.278	.317	.344
D_{sp}^2/σ_1^4	$\lambda=\infty$	.143	.510	1.366	2.712	4.548
	$\lambda_{.01}=6.37$	.147	.464	.948	1.467	1.736
	$\lambda_{.05}=3.58$	.157	.246	.564	.653	.681
	$\lambda_{.25}=1.65$	.167	.231	.256	.267	.265
	$\lambda=1$	.164	.205	.223	.236	.282
	$\lambda=0$	.250	.250	.250	.250	.250

Fig. 3: The behaviours of efficiencies for Example 3 $(\nu_2=6,\ \nu_1=8)$



So far as the numerical data are concerned, we may be able to make the following observations which might be suggestive to our statistical method in more general situation.

(1) From a general view of Fig. 3, it seems to us that the efficiencies of the present weighted estimator at $\rho<3$ are nearly those of a sometimes pooling estimation defined by $\lambda=\lambda_{0.25}$ and that the slope of the efficiency curve of the weighted estimator is also gentle.

For further details, however, the present weighted estimator is preferable to the sometimes pooling estimation defined by $\lambda=\lambda_{0.25}$ and $\lambda=1$ in case when ρ is near 1, but is inferior to the sometimes pooling estimation defined by $\lambda=\lambda_{0.01}$ or $\lambda=\infty$. While, the weighted estimator is so much preferable to those defined by $\lambda=\lambda_{0.01}$ or $\lambda=\infty$ in case when ρ is distant from 1, but is inferior to those defined by $\lambda=\lambda_{0.25}$ or $\lambda=1$.

Under these circumstances, we may propose on the whole that the present weighted estimator may be applied instead of the sometimes pooling estimation on the basis of the thought of successive inference, in order to improve the efficiency of estimation in case when we have nothing of the previous information about a value of ρ. That is to say, we may suggest a property of certain robustness of the present weighted estimator defined by (3.8) so far as the value of ρ is not so large.

(2) Specially if we attempt to compare the efficiencies between the present weighted estimator and the sometimes pooling estimation defined by $\lambda=\lambda_{0.05}$, like in Huntsberger [1], their efficiencies coincide at $\rho=1$ each other and the weighted estimator is also more efficient than another in our numerical example.

§ 4. Summary

This paper discusses some properties for efficiencies of the weighted estimator method proposed by Huntsberger [1] and the sometimes pooling estimation based on a thought of successive procedure of inferences, on a subject of the inference of population variance. There, on the basis of giving concretely their weighted estimator and of computing their numerical examples, the authors give certain propositions concerning to the cases that each method is to be chosen.

KYUSHU UNIVERSITY and SHIONOGI PHARMACEUTICAL Co.
KUMAMOTO WOMEN'S UNIVERSITY

Literatures

Asano, C. [1] : Estimations after preliminary test of significance and their applications to biometrical researches, Bull. Math. Stat. (1960), 1–23.

Bancroft, T. A. [1] : On biases in estimation due to the use of preliminary tests of significance, Ann. Math. Stat., **15** (1944), 190–204.

Bennett, B. D. [1] : Estimation of means on the basis of preliminary tests of significance, Ann. Inst. Stat. Math., Tokyo, **4** (1952), 31–43.

Huntsberger, D. V. [1] : A generalization of a preliminary testing procedure for pooling data, Ann. Math. Stat., **26** (1955), 734–743.

Kitagawa, T. [1] : Successive process of statistical inferences, Mem. Faculty of Sci., Kyushu Univ., Ser. A, **15** (1950), 139–180.

———— [2] : Some contributions to the design of sample surveys, Parts IV–VI, Sankhyā **17** (1956), 1–36.

———— [3] : Empirical functions and interpenetrating sampling procedures, Mem. Faculty of Sci., Kyushu Univ., Ser. A, **8** (1954), 109–152.

———— [4] : Some aspects of stochastically approximative analysis, Bull. Math. Stat., **6** (1956), 109–129.