J. Statist. Comput. Simul., 1997, Vol. 58, pp. 145–153
Reprints available directly from the publisher
Photocopying permitted by license only



# COMBINING INDEPENDENT NORMAL SAMPLE MEANS BY WEIGHTING WITH THEIR STANDARD ERRORS

BRETT MOORE and K. KRISHNAMOORTHY

*Department of Statistics, University of Southwestern Louisiana, Lafayette, LA 70504-1006, USA*

*(Received 6 September 1996; In final form 21 January 1997)*

The problem of estimating the common mean of several normal populations based on independent samples available from all the populations is considered. Two combined estimators are compared. One is the weighted average of the individual sample means with weights inversely proportional to their estimated standard errors ($\hat{\mu}_s$) and another is the traditional one which combines the individual sample means with weights inversely proportional to their estimated variances ($\hat{\mu}_v$). Exact expressions for the variances of $\hat{\mu}_v$ and $\hat{\mu}_s$ are derived. These estimators are compared numerically with respect to their variances. Comparison studies indicate that $\hat{\mu}_s$ is superior to $\hat{\mu}_v$ if the sample sizes are small and/or the population variances are not far apart from each other. Further, applications of these point estimates and the confidence intervals centered at these estimates to practical examples are discussed.

*Keywords:* Affine invariant; unbiased; combining data; confidence interval

## 1. INTRODUCTION

Suppose that we have independent samples from $k$ normal populations with unknown common mean $\mu$ but possibly different unknown variances $\sigma_1^2, \ldots, \sigma_k^2$. We consider the problem of estimating the common mean based on all the samples. This problem can be stated in its canonical forms as follows: We have $x_1, \ldots, x_k, s_1^2, \ldots, s_k^2$ are all statistically independent with

$$x_i \sim N(\mu, \sigma_i^2) \quad \text{and} \quad m_i s_i^2/\sigma_i^2 \sim \chi_{m_i}^2, \tag{1.1}$$

where $N(\mu, \sigma_i^2)$ denotes the normal distribution with mean $\mu$ and variance $\sigma_i^2$, and $\chi^2_{m_i}$ denotes the chi-squared random variable with degrees of freedom $m_i, i = 1, \ldots, k$. Note that $x_i$'s are independent unbiased estimates of $\mu$, and $s_i^2$'s are unbiased estimates $\sigma_i^2$'s. One commonly used estimator combines the $x_i$'s with weights inversely proportional to the estimated variances, $s_i^2$'s. That is,

$$\hat{\mu}_v = \frac{\sum_{i=1}^k x_i/s_i^2}{\sum_{j=1}^k 1/s_j^2} \tag{1.2}$$

This estimator is a descendent of the best linear unbiased estimator of $\mu$ (when $\sigma_i^2$'s are known), obtained by replacing the variances with their estimated values and is known in the literature as the Graybill-Deal (1959) estimator. Note that the weights are directly proportional to the estimated precision of the individual estimators $x_i$'s. Since standard error is also commonly used to measure the uncertainty of an estimator, another intuitively reasonable estimator of $\mu$ is found by weighting the individual estimators with the inverse of their estimated standard errors. That is,

$$\hat{\mu}_s = \frac{\sum_{i=1}^k x_i/s_i}{\sum_{j=1}^k 1/s_j}. \tag{1.3}$$

Both estimators have some appealing features. They are unbiased, affine invariant, and symmetric with respect to $(x_i, s_i)$'s. Further, it is observed by Fairweather (1972) that $\hat{\mu}_s$ is the center of the confidence interval ($I_s$) obtained by inverting the absolute sum of independent $t$-test statistics for testing $H_0 : \mu = \mu_0$. Jordan and Krishnamoorthy (1996) pointed out that by inverting the sum of independent $F$-test statistics one can obtain a confidence set ( $I_v$) for $\mu$. This confidence set $I_v$ could be empty when the sample means are especially different and is a finite interval centered at $\hat{\mu}_v$ when it is nonempty. Further, Jordan (1994) investigated several exact combined tests given in Zhou and Mathew (1993) and Mathew *et al.* (1993) to develop exact confidence sets for the common mean $\mu$. His study indicates that these confidence sets are not always finite intervals; they could be either empty or a union of disjoint intervals. Recently, Johnson and

Krishnamoorthy (1995) observed similar results in the context of a calibration problem. Since $I_s$ is the only exact confidence set which is always a finite interval, one may want to consider the center $\hat{\mu}_s$ as a point estimate of $\mu$. In particular, it is of interest to see how good $\hat{\mu}_s$ is in comparison to $\hat{\mu}_v$ and the individual sample means. This is what motivates the present work.

In the following we derive exact expressions for the variances of $\hat{\mu}_v$ and $\hat{\mu}_s$ for $k \geq 2$. To examine practically the general merit of $\hat{\mu}_s$, its variance is compared to that of $\hat{\mu}_v$ numerically. The comparison study shows that if one or both sample sizes are small, or the ratio of the population variances is not large, then $\hat{\mu}_s$ is preferable to $\hat{\mu}_v$ otherwise the latter performs better than the former. Further, we show that $\hat{\mu}_s$ has larger variance than that of either of the independent estimators $x_i$'s if the difference between the population variances is too large, and thus proving the conjecture of Jordon and Krishnamoorthy(1996). In Section 3 we discuss the application of $(\hat{\mu}_s, I_s)$ and $(\hat{\mu}_v, I_v)$ to the practical examples considered in Jordon and Krishnamoorthy (1996) in the context of the present problem.

## 2. MAIN RESULT

We derive exact expressions for the variance of $\hat{\mu}_v$ and $\hat{\mu}_s$ by extending the method of Nair (1980) who derived a variance formula for $\hat{\mu}_v$ when $k=2$. For simplicity and ease of presentation, we first derive the variance of $\hat{\mu}_v$ for $k = 3$.

Let $\alpha_1 = \sigma_1^2 m_2/(\sigma_2^2 m_1)$ and $\alpha_2 = \sigma_1^2 m_3/(\sigma_3^2 m_1)$. Assume without loss of generality that $0 < \alpha_i \leq 1$, $i = 1, 2$. Let $v_i = m_i s_i^2/\sigma_i^2$ be independent chi-squared random variables with degrees of freedom $m_i, i = 1, \ldots, 3$. Further, let $w_1 = v_1/(v_1 + v_2)$ and $w_2 = (v_1 + v_2)/(v_1 + v_2 + v_3)$. Note that $w_1$ and $w_2$ are independent beta random variables with $w_1 \sim B(m_1/2, m_2/2)$ and $w_2 \sim B((m_1 + m_2)/2, m_3/2)$. It is easy to show that

$$\mathrm{Var}(\hat{\mu}_v) = \sigma_1^2 E\left[\frac{1 + \alpha_1 m_2 v_1/(m_1 v_2) + \alpha_2 m_3 v_1/(m_1 v_3)}{(1 + \alpha_1 v_1/v_2 + \alpha_2 v_1/v_3)^2}\right].$$

Using the identities $v_1/v_2 = w_1/(1-w_1)$ and $v_1/v_3 = w_1w_2/(1-w_2)$ and after some minor simplification, it can be shown that

$$\mathrm{Var}(\hat{\mu}_v) = \sigma_1^2 E\left[\frac{(1-w_1)(1-w_2)+\alpha_1 m_2 w_1(1-w_2)/m_1+\alpha_2 m_3 w_1 w_2(1-w_1)/m_1}{[1-[1-(1-w_1)(1-w_2)-\alpha_1 w_1(1-w_2)-\alpha_2 w_1 w_2(1-w_1)]]^2}\right]. \tag{2.1}$$

Since $|1-(1-w_1)(1-w_2)-\alpha_1 w_1(1-w_2)-\alpha_2 w_1 w_2(1-w_1)| < 1$, using the Taylor series expansion and then taking term by term expectations we get

$$\begin{aligned}
\mathrm{Var}(\hat{\mu}_v) = & \sigma_1^2\left[B\left(\frac{m_1}{2},\frac{m_2}{2}\right)B\left(\frac{m_1+m_2}{2},\frac{m_3}{2}\right)\right]^{-1} \\
& \sum_{r=0}^{\infty}(r+1)\sum_R \frac{r!}{r_1!r_2!r_3!r_4!}(-1)^{r+r_4}\alpha_1^{r_2}\alpha_2^{r_3} \\
& \times\left[B\left(\frac{m_1}{2}+r_2+r_3,\frac{m_2}{2}+r_1+r_3+2\right)\right. \\
& B\left(\frac{m_1+m_2}{2}r_3,\frac{m_3}{2}+r_1+r_2+2\right) \\
& +\frac{\alpha_1 m_2}{m_1}B\left(\frac{m_1}{2}+r_2+r_3+2,\frac{m_2}{2}+r_1+r_3\right) \\
& B\left(\frac{m_1+m_2}{2}+r_3,\frac{m_3}{2}+r_1+r_2+2\right) \\
& +\frac{\alpha_2 m_3}{m_1}B\left(\frac{m_1}{2}+r_2+r_3+2,\frac{m_2}{2}+r_1+r_3+2\right) \\
& \left.B\left(\frac{m_1+m_2}{2}+r_3+2,\frac{m_3}{2}+r_1+r_2\right)\right]
\end{aligned} \tag{2.2}$$

where $R = \{(r_1,r_2,r_3,r_4) : \sum_{i=1}^{4} r_i = r, r_i \geq 0, i = 1,\ldots,4\}$.

To evaluate Var $(\hat{\mu}_v)$for a general $k$, let $w_i = \sum_{j=1}^{i} v_j / \sum_{j=1}^{i+1} v_j$, $i = 1,\ldots,k-1$. Then $w_1,\ldots,w_{k-1}$ are all independent with $w_i \sim B$ $(.5\sum_{j=1}^{i} m_j, .5m_{i+1})$ for $i = 1,\ldots,k-1$. The variance can be expressed as

$$\mathrm{Var}(\hat{\mu}_v) = \sigma_1^2 E\left(\frac{\prod_{i=1}^{k-1}(1-w_i)+\sum_{i=1}^{k-1}\frac{\alpha_i m_{i+1}}{m_1}\left(\prod_{j=1}^{i} w_j \prod_{l=1,l\neq 1}^{k-1}(1-w_l)\right)}{[1-(1-\prod_{i=1}^{k-1}(1-w_i)-\sum_{i=1}^{k-1}\alpha_i(\prod_{j=1}^{i} w_j\prod_{l=1,l\neq i}^{k-1}(1-w_l)))]^2}\right). \tag{2.3}$$

It is shown in the appendix that $| 1 - \Pi_{i=1}^{k-1}(1 - w_i) - \sum_{i=1}^{k-1} \alpha_i(\Pi_{j=1}^{i} w_j \Pi_{l=1,l\neq i}^{k-1}(1 - w_l)) | < 1$. Thus, using the Taylor series expansion and taking term by term expectations, we get

$$\begin{aligned}\mathrm{Var}\,(\hat{\mu}_v) = \sigma_1^2 \left( \prod_{i=1}^{k-1} B\left( .5\sum_{j=1}^{i} m_j,\ .5m_{i+1} \right) \right)^{-1} \sum_{r=0}^{\infty} \sum_{R} \frac{(r+1)r!(-1)^{r+r_{k+1}}}{\prod_{i=1}^{k+1} r_i!} \\ \times \left( \prod_{i=1}^{k-1} \alpha_{\alpha_i}^{r_i+1} \right) \sum_{j=1}^{k} \frac{m_j \alpha_{j-1}}{m_1} \prod_{l=1}^{k-1} B\Bigg( .5\sum_{i=1}^{l} m_i + \sum_{i=1}^{k} I_{[i>l]} r_i \\ + 2I_{[j>l]},\ .5m_{l+1} + \sum_{i=1}^{k} I_{[i\neq l+1]} r_i + 2I_{[j\neq l+1]} \Bigg) \end{aligned} \tag{2.4}$$

where $R = \left\{ (r_1, \ldots, r_{k+1}) : \sum_{i=1}^{k+1} r_i = r,\ r_i \geq 0, i = 1, \ldots, k+1 \right\}$, $\alpha_i = \sigma_1^2 m_{i+1}/(\sigma_{i+1}^2 m_1), i = 0, 1, \ldots, k-1$ and $I_{[\cdot]}$ is the indicator function.

The corresponding formula for $\mathrm{Var}(\hat{\mu}_s)$ is derived similarly. For $k \geq 2$,

$$\mathrm{Var}\,(\hat{\mu}_s) = \sigma_1^2 E\left( \frac{\prod_{i=1}^{k-1}(1 - w_i) + \sum_{i=1}^{k-1}\left( \prod_{j=1}^{i} w_j \prod_{l=1,\ l\neq i}^{k-1}(1 - w_l) \right)}{[1 - (1 - \prod_{i=1}^{k-1}(1 - w_i)^5 - \sum_{i=1}^{k-1}(\alpha_i \prod_{j=1}^{i} w_j \prod_{l=1,\ l\neq i}^{k-1}(1 - w_l))^{.5})]^2} \right), \tag{2.5}$$

and after expanding the denominator using the Taylor series and taking expectations,

$$\begin{aligned}\mathrm{Var}\,(\hat{\mu}_s) = \sigma_1^2 \left( \prod_{i=1}^{k-1} B\left( .5\sum_{j=1}^{i} m_j, .5m_{i+1} \right) \right)^{-1} \sum_{r=0}^{\infty} \sum_{R} \frac{(r+1)r!(-1)^{r+r_{k+1}}}{\prod_{i=1}^{k+1} r_i!} \\ \times \prod_{i=1}^{k-1} \alpha_i^{r_{i+1/2}} \sum_{j=1}^{k} \frac{m_j}{m_1} \prod_{l=1}^{k-1} B\Bigg( .5\sum_{i=1}^{l} m_i + .5\sum_{i=1}^{k} I_{[i>l]} r_{i1} + I_{[j>l]}, \\ .5m_{l+1} + .5\sum_{i=1}^{k} I_{[i\neq l+1]} r_i + I_{[j\neq l+1]} \Bigg). \end{aligned} \tag{2.6}$$

It is shown in the appendix that this series converges when $\alpha_i$'s lie between 0 and 1.

In the special case $\alpha_i$'s $\to 0$, the most unfavorable case for $\hat{\mu}_s$, we find from (2.5) that $\mathrm{Var}(\hat{\mu}_s) = \sigma_1^2(1 + \sum_{i=2}^{k} \frac{m_i}{m_i - 2})$ which is equal to $(3k-2)\sigma_1^2$ when $m_1 = \ldots = m_k = 3$ and decreases to $k\sigma_1^2$ when

$m_i$'s $\to \infty$. This means that $\hat{\mu}_s$ is worse than $x_1$ when $\sigma_1^2$ is especially smaller than $\sigma_1^2, i = 2, \ldots, k$; by symmetry $\hat{\mu}_s$ is worse than $x_i$ when $\sigma_i^2$ is much smaller than $\sigma_j^2$ for $j = 1, \ldots, i-1, i+1, \ldots, k$. Since $\text{Var}(\hat{\mu}_v) \to \sigma_1^2$ as $\alpha_i$'s$\to 0$ the same relation holds between $\hat{\mu}_s$ and $\hat{\mu}_v$ as between $\hat{\mu}_s$ and $x_1$. On the other hand, when $\alpha_i$'s $= 1$ and $m_i$'s $= m$ direct comparison of (2.3) and (2.5) shows that $\text{Var}(\hat{\mu}_s) < \text{Var}(\hat{\mu}_v)$ for all $k$ and $m$. When computed numerically, the variance ratio $\text{Var}(\hat{\mu}_s)/\text{Var}(\hat{\mu}_v)$ equals 0.85 when $m = 1$, $k = 2$ and equals 0.76 when $m = 1$, $k = 3$ and climbs towards 1 for larger values of $m$.

We computed the variance ratios using (2.4) and (2.6) for various parameter values and present them in Table I. We used the Fortran IMSL function subroutine DBETA (double precision) to compute the variance ratios.

It is clear from Table I that for $k = 2$, if $m_1$ or $m_2$ are small, or the ratios of the variances are not large, then $\hat{\mu}_s$ is more efficient than $\hat{\mu}_v$; otherwise $\hat{\mu}_v$ performs better. For the three sample case the same trends seem to hold. However, the improvement of $\hat{\mu}_s$ in the case of equal variances and equal sample sizes is more substantial.

*Remark* 2.1 We found that the variance ratios estimated using simulation (100 000 runs) are as accurate as (up to two decimal) the ones computed using the series expressions (2.4) and (2.6) for $k = 2$ and 3. For ease of computation, we recommend simulation to estimate the variance ratios for $k \geq 4$. We computed the variance ratios using simulation for $k = 4$ and 5 (not repeated here), and found that $\hat{\mu}_s$ becomes more efficient when the sample sizes are small and the differences among the $\sigma_i^2$'s are not too large; otherwise $\hat{\mu}_s$ can be much worse than $\hat{\mu}_v$.

## 3. DISCUSSION

The only desirable propertly $\hat{\mu}_v$ has, which $\hat{\mu}_s$ does not have, is that the former is better than each sample mean provided $m_i \geq 10$ for $i = 1, \ldots, k$ or $m_1 = 9$ and $m_2 \geq 19$ when $k = 2$. It is to be noted that both estimators are better than the individual sample means (even when the sample sizes do not satisfy the above conditions) if the population variances are not far apart from each other. This seems to

TABLE I $\text{Var}(\hat{\mu}_s)/\text{Var}(\hat{\mu}_v)$ for Several values of $m_i$ and $\sigma_i^2$

| | $m_1 m_2 (k=2)$ | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\sigma_2^2/\sigma_1^2$ | 1,1 | 2,2 | 3,3 | 5,5 | 10,10 | 30,30 | 30,1 | 1,30 | 30,5 | 5,30 | 30,15 | 15,30 |
| 1 | .85 | .86 | .87 | .90 | .94 | .98 | .86 | .86 | .93 | .93 | .97 | .97 |
| 1.25 | .85 | .86 | .88 | .91 | .94 | .98 | .86 | .87 | .93 | .93 | .97 | .97 |
| 1.5 | .85 | .86 | .88 | .91 | .95 | .99 | .86 | .88 | .95 | .94 | .98 | .98 |
| 1.75 | .85 | .87 | .89 | .92 | .96 | 1.00 | .86 | .80 | .95 | .95 | .99 | .99 |
| 2 | .86 | .87 | .90 | .93 | .97 | 1.01 | .86 | .90 | .96 | .96 | 1.00 | 1.00 |
| 2.5 | .86 | .89 | .92 | .96 | 1.00 | 1.03 | .87 | .92 | .99 | .99 | 1.03 | 1.02 |
| 3 | .87 | .90 | .93 | .98 | 1.02 | 1.06 | .87 | .94 | 1.01 | 1.01 | 1.05 | 1.05 |
| 4 | .88 | .93 | .97 | 1.02 | 1.07 | 1.10 | .88 | .98 | 1.06 | 1.06 | 1.10 | 1.09 |
| 5 | .89 | .95 | 1.01 | 1.06 | 1.11 | 1.14 | .90 | 1.01 | 1.10 | 1.09 | 1.14 | 1.13 |
| 10 | .93 | 1.06 | 1.15 | 1.23 | 1.27 | 1.27 | .94 | 1.14 | 1.27 | 1.23 | 1.28 | 1.27 |
| 20 | .98 | 1.20 | 1.34 | 1.48 | 1.44 | 1.42 | 1.00 | 1.29 | 1.48 | 1.38 | 1.44 | 1.41 |

| | | $m_1 m_2 m_3 (k=3)$ | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $\sigma_2^2/\sigma_1^2$ | $\sigma_3^2/\sigma_1^2$ | 1,1,1 | 2,2,2 | 3,3,3 | 5,5,5 | 10,10,10 | 15,10,5 | 30,15,10 | 45,15,5 |
| 1 | 1 | .76 | .76 | .79 | .85 | .91 | .89 | .94 | .91 |
| 1 | 2 | .77 | .78 | .81 | .87 | .94 | .92 | .97 | .93 |
| 1 | 5 | .78 | .83 | .89 | .96 | 1.03 | 1.02 | 1.06 | 1.04 |
| 1 | 10 | .81 | .90 | .97 | 1.05 | 1.11 | 1.13 | 1.15 | 1.16 |
| 1 | 20 | .83 | .98 | 1.07 | 1.15 | 1.19 | 1.24 | 1.24 | 1.28 |
| 2 | 2 | .77 | .78 | .82 | .88 | .94 | .92 | .97 | .94 |
| 2 | 3 | .77 | .79 | .84 | .90 | .97 | .95 | 1.00 | .97 |
| 2 | 5 | .78 | .82 | .88 | .95 | 1.02 | 1.01 | 1.06 | 1.03 |
| 2 | 10 | .80 | .88 | .96 | 1.04 | 1.11 | 1.11 | 1.15 | 1.14 |
| 5 | 5 | .79 | .85 | .95 | 1.02 | 1.10 | 1.08 | 1.13 | 1.10 |
| 5 | 10 | .81 | .90 | 1.00 | 1.11 | 1.20 | 1.19 | 1.24 | 1.22 |
| 5 | 20 | .83 | .97 | 1.10 | 1.23 | 1.30 | 1.33 | 1.35 | 1.36 |

be more likely in practical situations. Jordan and Krishnamoorthy (1996) computed $(\mu_s, I_s)$ and $(\mu_v, I_v)$ for two real data sets (one is concerned with the estimation of Selenium in non-fat milk powder by combining the results of four different analytical methods, and the other is concerned with the estimation of the percentage of albumin in blood plasma protein by combining the results of four different experiments) and found that there was only a minute difference between $\hat{\mu}_s$, and $\hat{\mu}_v$, and that $I_s$ was narrower than $I_v$ in one case and almost equal to $\hat{\mu}_v$ in the other case. In light of these results, $(\hat{\mu}_s, I_s)$, seems to be a simple alternative to $(\hat{\mu}_v, I_v)$ for practical usage and may be preferable to $(\hat{\mu}_v, I_v)$ when the population variances are close to each other.

### *References*


Fairweather, W.R. (1972) A Method of Obtaining an Exact Confidence Interval for the Common Mean of Several Normal Populations. *Applied Statistics*, **21**, 229–233.

Graybill, F.A. and Deal, R.B. (1959) Combining Unbiased Estimators. *Biometrics*, **15**, 543–550.

Jordan, D. and Krishnamoorthy, K. (1996) Combining Independent Studies in a Calibration Problem. *Journal of the American Statiscal Association*, **91**, 1707–1718.

Jordan, S. J. (1994) Interval Estimation and Hypothesis Testing of the Common Mean of Several Normal Populations. Ph. D. Dissertation, Dept. Statistics, University of Southwestern Louisiana.

Jordan, S. J. and Krishnamoorthy, K. (1996) Exact Confidence Intervals for the Common Mean of Several Normal Populations, *Biometrics*, **52**, 77–86.

Mathew, T., Sinha, B. K. and Zhou, L. (1993) Some Statistical Procedure for Combining Independent Tests. *Journal of the American Statistical Association*, **88**, 912–919.

Nair, A.K. (1980) Variance and Distribution of the Graybill-Deal Estimator of the common Mean of Two Normal Populations. *The Annals of Statistics*, **8**, 212–216.

Zhou, L. and Mathew, T. (1993) Combining Independent Tests in Linear Models. *Journal of the American Statiscal Association*, **88**, 650–655.


## APPENDIX

To show that the infinite series in (2.4) for $\mathrm{Var}(\hat{\mu}_v)$ converges, we need to prove that $|1 - \prod_{i=1}^{k-1}(1-w_i) - \sum_{i=1}^{k-1}\alpha_i(\prod_{j=1}^{i} w_j \prod_{l=1,l\neq i}^{k-1}(1-w_l))| < 1$. Since $\alpha_i$'s lie between 0 and 1, this inequality holds if $\prod_{i=1}^{k-1}(1-w_i) + \sum_{i=1}^{k-1}(\prod_{j=1}^{i} w_j \prod_{l=1,l\neq i}^{k-1}(1-w_l)) < 2$ The left-hand side (lhs) of the inequality can be factored as

$$\begin{aligned}
&\prod_{i=2}^{k-1}(1-w_i) + w_1(1-w_1)\sum_{i=2}^{k-1}\left(\prod_{j=2}^{i} w_j \prod_{l=2,\, l\neq i}^{k-1}(1-w_l)\right)\\
&\leq 1 + (1/4)\sum_{i=2}^{k-1}\left(\prod_{j=2}^{i} w_j \prod_{l=2,\, l\neq i}^{k-1}(1-w_l)\right)\\
&\leq 1 + (1/4)\left[w_2\prod_{j=3}^{k-1}(1-w_j) + w_2(1-w_2)\sum_{i=3}^{k-1}\left(\prod_{j=3}^{i} w_j \prod_{l=3,\, l\neq i}^{k-1}(1-w_l)\right)\right]\\
&\leq 1 + (1/4)\left[1 + (1/4)\sum_{i=3}^{k-1}\left(\prod_{j=3}^{i} w_j \prod_{l=3,\, l\neq i}^{k-1}(1-w_l)\right)\right] \qquad \text{(A.1)}
\end{aligned}$$

since $w_i$'s lie between 0 and 1, $\prod(1-w_i) \leq 1$ and $w_i(1-w_i) \leq 0.25$. Proceeding this way, it is seen that the right-hand side (rhs) of (A.1) is less than $\sum_{i=0}^{\infty}(1/4)^i = 4/3$.

The series expression (2.6) of $\mathrm{Var}(\hat{\mu}_s)$ converges if

$$\prod_{i=1}^{k-1}(1-w_i)^{.5}+\sum_{i=1}^{k-1}\left(\prod_{j=1}^{i}w_j\prod_{l=1,l\neq i}^{k-1}(1-w_l)\right)^{.5}<2.$$

Factoring the lhs of this inequality, and using the facts that $\sqrt{w_i}+\sqrt{1-w_i}\leq\sqrt{2}$ and $\sqrt{2}\sqrt{1-w_i}+.5\sqrt{w_i}\leq 3/2$, we see that

$$\left(\sqrt{w_1}+\sqrt{1-w_1}\right)\prod_{i=2}^{k-1}\sqrt{1-w_i}+\sqrt{w_1}\sqrt{1-w_1}$$

$$\sum_{i=2}^{k-1}\prod_{j=2}^{i}\sqrt{w_j}\prod_{l=2,l\neq i}^{k-1}\sqrt{1-w_l}$$

$$\leq\left(\sqrt{2(1-w_2)}+\sqrt{w_2}/2\right)\prod_{i=3}^{k-1}\sqrt{1-w_i}$$

$$+.5\sqrt{w_2(1-w_2)}\sum_{i=3}^{k-1}\prod_{j=3}^{i}\sqrt{w_j}\prod_{l=3,l\neq i}^{k-1}\sqrt{1-w_l}$$

$$\leq 3/2+(1/4)\left(\sqrt{w_3}\prod_{j=4}^{k-1}\sqrt{1-w_j}\right.$$

$$\left.+\sqrt{w_3(1-w_3)}\sum_{i=4}^{k-1}\prod_{j=4}^{i}\sqrt{w_j}\prod_{l=4,l\neq i}^{k-1}\sqrt{1-w_l}\right)$$

$$\leq 3/2+(1/4)\left(1+(1/2)\sum_{i=4}^{k-1}\prod_{j=4}^{i}\sqrt{w_j}\prod_{j=4,l\neq i}^{k-1}\sqrt{1-w_l}\right). \qquad (\text{A}.2)$$

Continuing this way, it is clear that the rhs of (A.2) is less than $3/2+\sum_{i=0}^{\infty}(1/2)^i/4=2$.