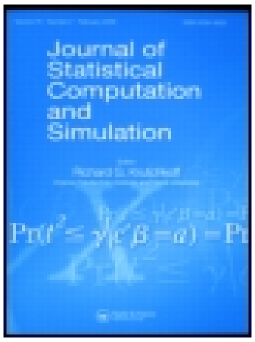

Journal of Statistical Computation and Simulation

ISSN: 0094-9655 (Print) 1563-5163 (Online) Journal homepage: https://www.tandfonline.com/loi/gscs20

Estimation of the common mean of a bivariate normal distribution

K. Krishnamoorthy & Vijay K. Rohatgi

To cite this article: K. Krishnamoorthy & Vijay K. Rohatgi (1989) Estimation of the common mean of a bivariate normal distribution, Journal of Statistical Computation and Simulation, 31:3, 187-194, DOI: 10.1080/00949658908811142

To link to this article: https://doi.org/10.1080/00949658908811142

Published online: 20 Mar 2007.

J. Statist. Comput. Simul., Vol. 31, pp. 187–194
Reprints available directly from the publisher
Photocopying permitted by license only

ESTIMATION OF THE COMMON MEAN OF A BIVARIATE NORMAL DISTRIBUTION

K. KRISHNAMOORTHY and VIJAY K. ROHATGI

Bowling Green State University, Bowling Green, Ohio 43403-0221, USA

(Received 8 February 1988; in final form 9 August 1988)

The problem of unbiased estimation of the common mean in sampling from a bivariate normal distribution is considered. Performance of several estimators is compared.

KEY WORDS: Bivariate normal distribution, unbiased estimation, maximum likelihood estimator, testimator.

1. INTRODUCTION

The problem of estimation of the common mean of two or more normal populations has attracted, and continues to attract, a great deal of attention. We refer to Graybill and Deal (1959), Mehta and Gurland (1969), and Zacks (1966) for some early papers, and to Kubokawa (1987a, b, c) for some recent work. See also Lehmann (1983, pp. 88–89) for further references and comments. In this paper we consider the more general problem of estimation of the common mean of a bivariate normal population.

Let $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$ be a random sample of size n from a bivariate normal population with unknown parameters

$$\mathscr{E}X = \mathscr{E}Y = \mu, \quad \text{var}(X) = \sigma_1^2, \quad \text{var}(Y) = \sigma_2^2, \quad \text{and} \quad \text{cov}(X, Y) = \sigma_{12} = \rho\sigma_1\sigma_2.$$

The problem of estimation of μ has received little attention. In the general case of sampling from a p-variate ($p \geqq 2$) normal population with common mean μ. Halperin (1961) obtained maximum likelihood estimators (MLE) of μ and the variance–covariance matrix. He showed that the MLE of μ is almost linearly optimum combination of the sample means. In the bivariate case Rastogi and Rohatgi (1974) made efficiency comparisons between the MLE ($\hat{\mu}_2$) derived by Halperin and the conventional estimator ($\hat{\mu}_1$), namely, the average of the two sample means. They found that $\hat{\mu}_2$ generally performs better than $\hat{\mu}_1$ for moderate sample sizes and values of $k = \sigma_2^2/\sigma_1^2$ not too close to one. (For $k = 1$, $\hat{\mu}_1$ is the best linear combination of the sample means.)

In this paper we consider two additional unbiased estimators $\hat{\mu}_3$, $\hat{\mu}_4$ of μ and compare their performance to that of $\hat{\mu}_2$. Our numerical computations of their variances indicate that the estimator $\hat{\mu}_3$ proposed in Section 2 has the least variance for a certain range of values of k and ρ whereas $\hat{\mu}_2$ performs best for the rest of the parameter space excepting values of k near one. In Section 2 we introduce the estimators and compute their variances and in Section 3, we present numerical computations and our recommendations.

2. RESULTS

Let $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$ be a random sample from a bivariate normal population (X, Y) with unknown parameters

$$\mathscr{E}X = \mathscr{E}Y = \mu, \quad \operatorname{var}(X) = \sigma_1^2, \quad \operatorname{var}(Y) = \sigma_2^2, \quad \operatorname{cov}(X, Y) = \sigma_{12} = \rho\sigma_1\sigma_2. \tag{2.1}$$

Let us write

$$\left.\begin{aligned} \bar{X} = \sum_{i=1}^{n} X_i/n, \quad \bar{Y} = \sum_{i=1}^{n} Y_i/n, \quad S_1^2 = \sum_{i=1}^{n} (X_i - \bar{X})^2/(n-1), \\ S_2^2 = \sum_{i=1}^{n} (Y_i - \bar{Y})^2/(n-1), \quad S_{12} = \sum_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y})/(n-1) \end{aligned}\right\}. \tag{2.2}$$

For convenience we use the notation of Rastogi and Rohatgi (1974) and set

$$\hat{\mu}_1 = (\bar{X} + \bar{Y})/2, \tag{2.3}$$

$$\hat{\mu}_2 = \bar{Y} + (\bar{X} - \bar{Y})(S_2^2 - S_{12})/(S_1^2 + S_2^2 - 2S_{12}). \tag{2.4}$$

Clearly $\hat{\mu}_1$ is unbiased for μ with variance

$$\operatorname{var}(\hat{\mu}_1) = (\sigma_1^2 + \sigma_2^2 + 2\sigma_{12})/(4n). \tag{2.5}$$

The estimator $\hat{\mu}_2$ was shown to be the MLE of μ by Halperin (1961) with variance

$$\operatorname{var}(\hat{\mu}_2) = (n-2)(\sigma_1^2\sigma_2^2\sigma_{12}^2)/[n(n-3)(\sigma_1^2 + \sigma_2^2 - 2\sigma_{12})]. \tag{2.6}$$

provided that $n > 3$. It is also known to be unbiased for μ.

We note that

$$\operatorname{var}(\hat{\mu}_2) < \operatorname{var}(\bar{X}) \quad \text{if and only if} \quad n > 3 + [k(1-\rho^2)/(1-\rho\sqrt{k})^2]. \tag{2.7}$$

where $k = \sigma_2^2/\sigma_1^2$, and a similar statement holds for $\bar{Y}$.

Following Rastogi and Rohatgi (1974) we let $U = X - Y$ and $V = X + Y$. Then (U, V) has a bivariate normal distribution with

$$\left.\begin{aligned} \mathscr{E}U = 0, \quad \mathscr{E}V = 2\mu, \quad \operatorname{var}(U) = \sigma_U^2 = \sigma_1^2 + \sigma_2^2 - 2\sigma_{12}, \\ \operatorname{var}(V) = \sigma_V^2 = \sigma_1^2 + \sigma_2^2 + 2\sigma_{12} \quad \text{and} \quad \operatorname{cov}(U, V) = \sigma_{UV} = \sigma_1^2 - \sigma_2^2 \end{aligned}\right\}. \tag{2.8}$$

Writing $\bar{U}$, $\bar{V}$ for the sample means, S_U^2, S_V^2 for the corresponding sample variances, and S_{UV} for the sample covariance we see that

$$\hat{\mu}_2 = \bar{V}/2 - (\bar{U}/2)(S_{UV}/S_U^2). \tag{2.9}$$

Let $\beta = \sigma_{UV}/\sigma_U^2$, $v = \sigma_U/\sigma_{V|U}$, where $\sigma_{V|U}^2 = \sigma_V^2\{1 - \sigma_{UV}^2/(\sigma_U^2\sigma_V^2)\}$. Then $W = S_{UV}/S_U^2$ was shown (Rastogi and Rohatgi, 1974) to have the density function

$$f(w) = \frac{v\Gamma(n/2)}{\sqrt{\pi}\Gamma((n-1)/2)}[1 + v^2(w-\beta)^2]^{-n/2}. \tag{2.10}$$

Let us write $S_{UV}^* = (n-1)S_{UV}$, $S_{UU}^* = (n-1)S_U^2$ and $S_{VV}^* = (n-1)S_V^2$. Then $W = S_{UV}/S_U^2 = S_{UV}^*/S_{UU}^*$. We note that $\mathscr{E}W = \beta$ so that S_{UV}^*/S_{UU}^* is an unbiased estimator of σ_{UV}/σ_U^2. Since $\mathscr{E}U = 0$ it is more reasonable to estimate σ_{UV}/σ_U^2 by $S_{UV}^*/\sum_{i=1}^n U_i^2$ and μ by

$$\hat{\mu}_3 = \bar{V}/2 - (\bar{U}/2)\left(S_{UV}^* \Big/ \sum_{i=1}^n U_i^2\right). \tag{2.11}$$

It is easy to see that $\mathscr{E}\hat{\mu}_3 = \mu$. In order to compute the variance we use the following well-known facts.

i) The conditional distribution of S_{UV}^* given S_{UU}^* is normal with mean βS_{UU}^* and variance $\sigma_{V|U}^2 S_{UU}^{*2}$.
ii) If X and Y are independent rvs with χ_a^2 and χ_b^2 distributions respectively then $X/(X+Y)$ has a beta distribution with parameters $a/2$ and $b/2$.
iii) If X has a χ_n^2 distribution then for any real function ψ for which $\mathscr{E}[X\psi(X)]$ exists,

$$\mathscr{E}[X\psi(X)] = n\mathscr{E}\psi(X^*)$$

where X^* has a χ_{n+2}^2 distribution.

Using (i)–(iii) we show that

$$\operatorname{var}(\hat{\mu}_3) = \frac{1}{4n}\left\{\sigma_V^2 + \frac{n-1}{n+2}\left[\frac{\sigma_{V|U}^2}{n} - \frac{\sigma_{UV}^2}{\sigma_U^2}\left(\frac{n+7}{n+4}\right)\right]\right\}. \tag{2.12}$$

Indeed

$$\operatorname{var}(\hat{\mu}_3) = \frac{1}{4n}\left\{\sigma_V^2 + n\mathscr{E}\left[\frac{\bar{U}^2 S_{UV}^{*2}}{(S_{UU}^* + n\bar{U}^2)^2}\right] - 4n\mathscr{E}\left[\frac{S_{UV}^*}{S_{UU}^* + n\bar{U}^2}\bar{U}(\bar{V}/2 - \mu)\right]\right\} \tag{2.13}$$

We first evaluate the second term on the right-hand side of (2.13). Since

$n\bar{U}^2/\sigma_U^2 = h_1^*$ has a χ_1^2 distribution independently of S_{UV}^* and S_{UU}^* we have in view of (iii), and (i)

$$n\mathscr{E}\left[\frac{\bar{U}^2 S_{UV}^{*2}}{(S_{UU}^* + n\bar{U}^2)^2}\right] = \frac{1}{\sigma_U^2}\mathscr{E}\left[\frac{h_1^* S_{UV}^{*2}}{(S_{UU}^*/\sigma_U^2 + h_1^*)^2}\right] = \frac{1}{\sigma_U^2}\mathscr{E}\left[\frac{S_{UV}^{*2}}{(S_{UU}^*/\sigma_U^2 + h_1)^2}\right]$$

$$= \frac{1}{\sigma_U^2}\mathscr{E}\left\{\mathscr{E}\left\{\frac{S_{UV}^{*2}}{(S_{UU}^*/\sigma_U^2 + h_1)^2}\middle| S_{UU}^*\right\}\right\}$$

$$= \frac{1}{\sigma_U^2}\mathscr{E}\left\{\frac{\sigma_{V|U}^2 S_{UU}^* + (\sigma_{UV}^2/\sigma_U^4) S_{UU}^{*2}}{(S_{UU}^*/\sigma_U^2 + h_1)^2}\right\}$$

$$= \mathscr{E}\left\{\frac{\sigma_{V|U}^2 h_2 + (\sigma_{UV}^2/\sigma_U^2) h_2^2}{(h_2 + h_1)^2}\right\}$$

where h_1 is a χ_3^2 rv and $h_2 = S_{UU}^*/\sigma_U^2$ is a χ_{n-1}^2 rv. Finally using (ii) we get

$$n\mathscr{E}\left[\frac{\bar{U}^2 S_{UV}^{*2}}{(S_{UU}^* + n\bar{U}^2)^2}\right] = \sigma_{V|U}^2\left(\frac{n-1}{n(n+2)}\right) + \frac{\sigma_{UV}^2}{\sigma_U^2}\left(\frac{n^2-1}{(n+2)(n+4)}\right). \tag{2.14}$$

In order to evaluate the third term on the right side of (2.13) we note that $\bar{V}$ given $\bar{u}$ has a normal $(2\mu + \beta\bar{u}, \sigma_{V|U}^2)$ distribution. Hence

$$n\mathscr{E}\left\{\frac{S_{UV}^*\bar{U}(\bar{V}/2-\mu)}{S_{UU}^* + n\bar{U}^2}\right\} = \frac{n}{2}\mathscr{E}\left\{\frac{\beta S_{UV}^*\bar{U}^2}{S_{UU}^* + n\bar{U}^2}\right\} = (\beta/2)\mathscr{E}\left(\frac{S_{UV}^* h_1^*}{S_{UU}^*/\sigma_U^2 + h_1^*}\right).$$

Using an argument similar to the one used above in deriving (2.14) we see that

$$\mathscr{E}\left(\frac{S_{UV}^* h_1^*}{S_{UU}^*/\sigma_U^2 + h_1^*}\right) = \sigma_{UV}\mathscr{E}\left(\frac{h_2}{h_2 + h_1}\right) = \sigma_{UV}\left(\frac{n-1}{n+2}\right)$$

in view of (ii) so that

$$n\mathscr{E}\left\{\frac{S_{UV}^*\bar{U}(\bar{V}/2-\mu)}{S_{UU}^* + n\bar{U}^2}\right\} = \frac{\sigma_{UV}^2}{2\sigma_U^2}\left(\frac{n-1}{n+2}\right). \tag{2.15}$$

Substituting the expressions in (2.14) and (2.15) in (2.13) and simplifying we get (2.12).

In order to compare $\mathrm{var}(\hat{\mu}_2)$ and $\mathrm{var}(\hat{\mu}_3)$ we rewrite (2.6) in terms of moments of U and V as

$$\operatorname{var}(\hat{\mu}_2)=\frac{n-2}{4n(n-3)}\sigma^2_{V|U}. \tag{2.16}$$

Simple algebraic manipulations show that

$$\operatorname{var}(\hat{\mu}_3)<\operatorname{var}(\hat{\mu}_2) \quad \text{if and only if}$$

$$\sigma^2_{UV}/(\sigma^2_V\sigma^2_U)<(2n-1)(n+4)/(7n^2-8n-4) \tag{2.17}$$

or, in terms of ρ and $k=\sigma^2_2/\sigma^2_1$

$$\operatorname{var}(\hat{\mu}_3)<\operatorname{var}(\hat{\mu}_2) \quad \text{if and only if}$$

$$\rho^2<1-5n(n-1)(1-k)^2/[4k(2n^2+7n-4)]. \tag{2.18}$$

Yet another unbiased estimator of μ can be based on the fact that if $\sigma_1=\sigma_2$ then the best unbiased estimator of μ is $\hat{\mu}_1$. One can therefore construct an unbiased testimator using a test of $H_0{:}\sigma^2_1=\sigma^2_2$. Under H_0, $\sigma_{UV}=0$ so that $\beta=0$ and W has p.d.f.

$$f_0(w)=\frac{\Gamma(n/2)}{(\pi)^{1/2}\Gamma((n-1)/2)}\left(\frac{1-\rho}{1+\rho}\right)^{1/2}[1+(1-\rho)w^2/(1+\rho)]^{-n/2}. \tag{2.19}$$

Let $c_{\alpha/2}=c$ be the unique solution of $\alpha/2=\int_c^\infty f_0(w)\,dw$, and let $A=\{|W|\leq c\}$, $\bar{A}=\{|W|>c\}$. Consider the estimator

$$\hat{\mu}_4=\hat{\mu}_1 I(A)+\hat{\mu}_2 I(\bar{A}), \tag{2.20}$$

where $I(B)$ is the indicator function of event B. Clearly $\mathscr{E}\hat{\mu}_4=\mu$ and

$$\operatorname{var}\{\hat{\mu}_4|W\}=I(A)(\sigma^2_V/(4n))+I(\bar{A})(\sigma^2_V+W^2\sigma^2_V-2W\sigma_{UV})/(4n).$$

It follows that

$$\operatorname{var}(\hat{\mu}_4)=\mathscr{E}(\operatorname{var}\{\hat{\mu}_4|W\})=\sigma^2_V/(4n)+(1/(4n))\int_{|w|>c}(w^2\sigma^2_U-2w\sigma_{UV})f(w)\,dw \tag{2.21}$$

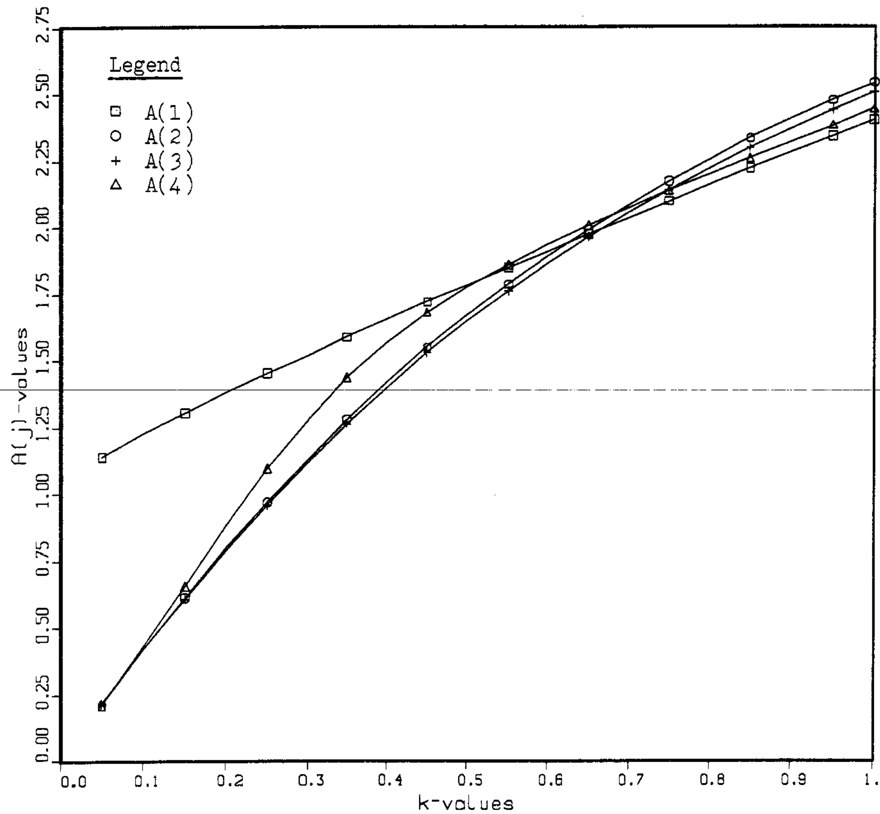

Figure 1 Graph of $A(j)$ vs. k for $n=20$ and $\rho=0.2$.

where $f(w)$ is given in (2.10). It is clear that $\mathrm{var}(\hat{\mu}_4)$ has to be evaluated numerically.

3. CONCLUSION

Figures 1–3 show the graphs of $A(j)=4n\,\mathrm{var}(\hat{\mu}_j)/\sigma_1^2$, $j=1,2,3,4$, for $n=20$ vs. k $(0<k\leqq 1)$ for values of $\rho=0.2$, 0.6 and 0.9 respectively. The computations for A_4 were done at significance level $\alpha=0.05$. Computations were also done for values of $k>1$, and other values of ρ $(=-0.2, -0.4, -0.6, -0.8, -0.9)$ and $n=10$, 30, 40, 50 with similar results.

Our computations show that $\hat{\mu}_1$ has the least variance for values of k near one $(0.85\leqq k\leqq 1.18)$. The estimator $\hat{\mu}_2$ is preferred for very small values of k $(\leqq 0.05)$ or very large values of k $(\geqq 20)$. For moderate values of k, $0.20\leqq k\leqq 0.65$ (and hence also for $0.20\leqq 1/k\leqq 0.65$) and values of $|\rho|\leqq 0.8$, the estimator $\hat{\mu}_3$ has the least variance. For $|\rho|$ near one, $\hat{\mu}_3$ has the least variance for $0.3\leqq k\leqq 0.8$ (and

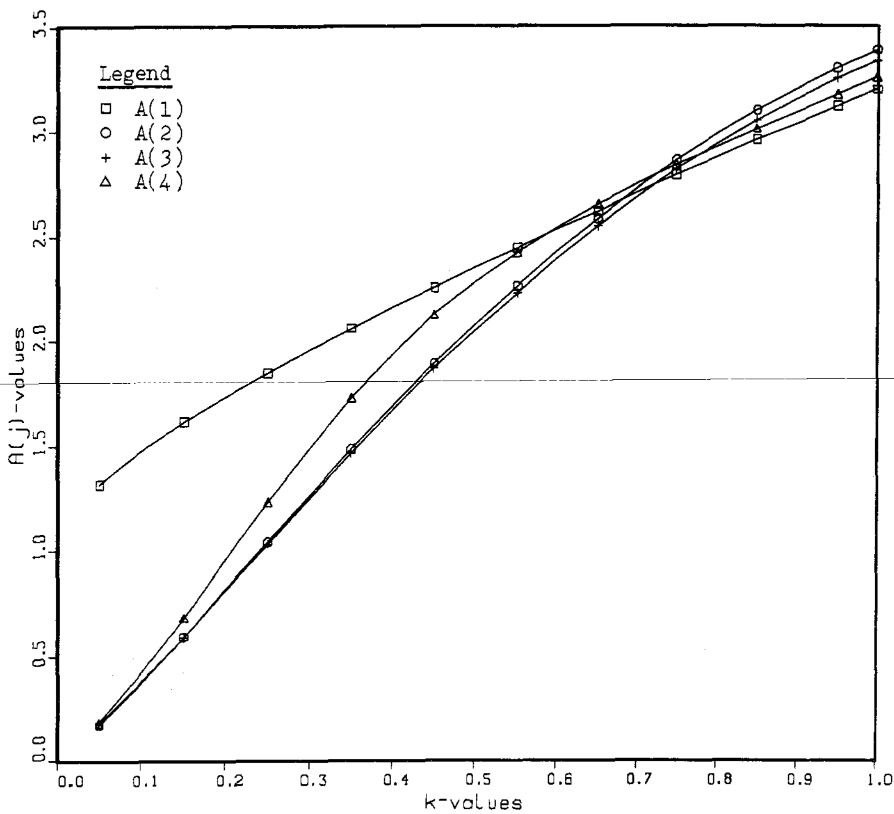

Figure 2 Graph of $A(j)$ vs. k for $n=20$ and $\rho=0.6$.

hence also for $0.3 \leqq 1/k \leqq 0.8$). Moreover, the estimator $\hat{\mu}_3$ dominates $\hat{\mu}_2$ except when k is small. It should be noted that the estimator $\hat{\mu}_4$ is not the preferred estimator in any region of the parameter space. These results are summarized in Table 1 below.

Table 1 Preferred estimator

k (or 1/k)	$\lvert\rho\rvert$				
	0.2	*0.4*	*0.6*	*0.8*	*0.9*
0.05	$\hat{\mu}_2$	$\hat{\mu}_2$	$\hat{\mu}_2$	$\hat{\mu}_2$	$\hat{\mu}_2$
0.10	$\hat{\mu}_3$	$\hat{\mu}_2$	$\hat{\mu}_2$	$\hat{\mu}_2$	$\hat{\mu}_2$
0.15	$\hat{\mu}_3$	$\hat{\mu}_3$	$\hat{\mu}_3$	$\hat{\mu}_2$	$\hat{\mu}_2$
0.20, 0.25	$\hat{\mu}_3$	$\hat{\mu}_3$	$\hat{\mu}_3$	$\hat{\mu}_3$	$\hat{\mu}_2$
0.30 to 0.65	$\hat{\mu}_3$	$\hat{\mu}_3$	$\hat{\mu}_3$	$\hat{\mu}_3$	$\hat{\mu}_3$
0.70	$\hat{\mu}_1$	$\hat{\mu}_1$	$\hat{\mu}_3$	$\hat{\mu}_3$	$\hat{\mu}_3$
0.75	$\hat{\mu}_1$	$\hat{\mu}_1$	$\hat{\mu}_1$	$\hat{\mu}_3$	$\hat{\mu}_3$
0.80	$\hat{\mu}_1$	$\hat{\mu}_1$	$\hat{\mu}_1$	$\hat{\mu}_1$	$\hat{\mu}_3$
0.85 to 1.0	$\hat{\mu}_1$	$\hat{\mu}_1$	$\hat{\mu}_1$	$\hat{\mu}_1$	$\hat{\mu}_1$

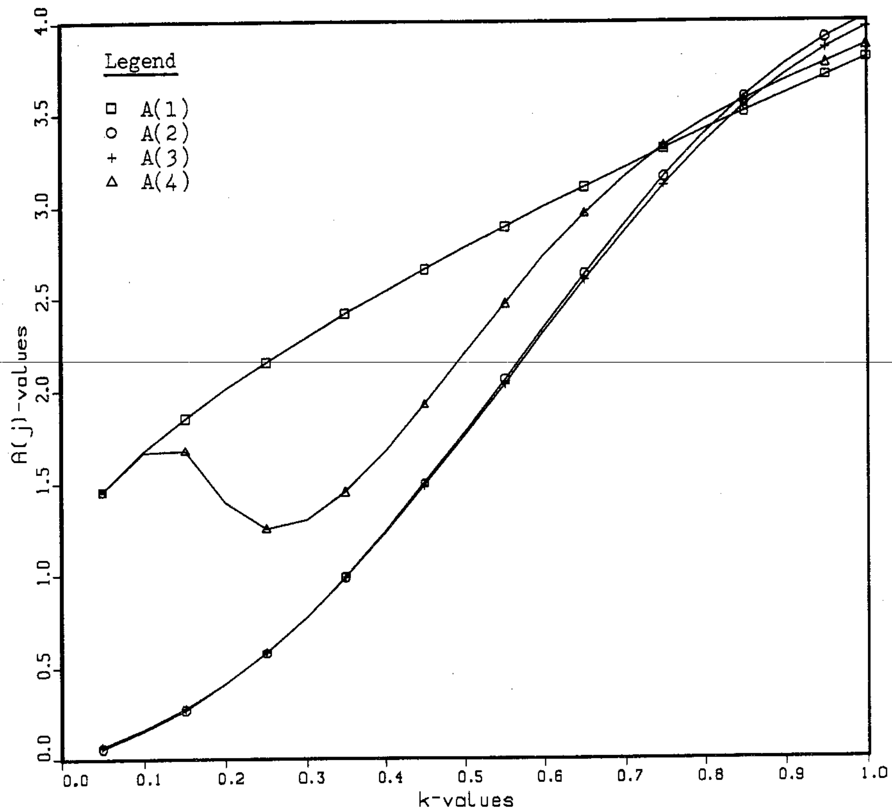

Figure 3 Graph of $A(j)$ vs. k for $n=20$ and $\rho=0.9$.

References

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

Halperin, M. (1961). Almost linearly-optimum combination of unbiased estimates, *Amer. Statist. Assoc.* **56,** 36–43.

Kubokawa, T. (1987a). Estimation of a common mean of two normal distributions, *Tsukuba J. Math.* **11,** 157–175.

Kubokawa, T. (1987b). Estimation of a common mean with symmetric loss, *J. Japan Statist. Soc.* **17,** 75–79.

Kubokawa, T. (1987c). Admissible minimax estimation of a common mean of two normal populations, *Ann. Stat.* **15,** 1245–1256.

Lehmann, E. (1983). *Theory of Point Estimation*, John Wiley, New York.

Mehta, J. S. and Gurland, J. (1969). On combining unbiased estimators of the mean. *Trabajos de Estadistica* **20,** 173–185.

Rastogi, S. C. and Rohatgi, V. K. (1974). Unbiased estimation of the common mean of a bivariate normal distribution. *Biom. Z.* **16,** 155–166.

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.