ARL 72-0032
FEBRUARY 1972

Aerospace Research Laboratories

A MONTE CARLO COMPARISON OF FOUR ESTIMATORS OF THE DISPERSION MATRIX OF A BIVARIATE NORMAL POPULATION, USING INCOMPLETE DATA

J. N. SRIVASTAVA
M. K. ZAATAR
COLORADO STATE UNIVERSITY
FORT COLLINS, COLORADO

CONTRACT NO. F33615-67-C-1436
PROJECT NO. 7071

AIR FORCE SYSTEMS COMMAND
United States Air Force

29

NOTICES

Agencies of the Department of Defense, qualified contractors and other government agencies may obtain copies from the

Defense Documentation Center
Cameron Station
Alexandria, Virginia 22314

This document has been released to the

CLEARINGHOUSE
U. S. Department of Commerce
Springfield, Virginia 22151

for sale to the public.

AIR FORCE: 2-5-72/600

ARL 72-0032

A MONTE CARLO COMPARISON OF FOUR ESTIMATORS OF THE DISPERSION MATRIX OF A BIVARIATE NORMAL POPULATION, USING INCOMPLETE DATA

J. N. SRIVASTAVA AND M. K. ZAATAR
COLORADO STATE UNIVERSITY
FORT COLLINS, COLORADO

JANUARY 1972

CONTRACT NO. F33615-67-C-1436
PROJECT NO. 7071

AEROSPACE RESEARCH LABORATORIES
AIR FORCE SYSTEMS COMMAND
UNITED STATES AIR FORCE
WRIGHT-PATTERSON AIR FORCE BASE, OHIO

FOREWORD

This report constitutes the final report for the research work under Contract F33615-67-C-1436 of the Aerospace Research Laboratories, Air Force Systems Command, United States Air Force. The research work contained in this report was wholly supported by the above contract.

Below, we give a list* of all the technical papers published, and the ARL reports issued containing research work which was supported either wholly or partly by the above contract, or the previous contract (whose number was AF 33(615)-3231).

1. (1966). Some generalizations of multivariate analysis of variance. Multivariate Analysis. (edited by P. R. Krishnaiah), pp. 129-145.

2. (1967). On the extension of Gauss-Markov theorem to complex multivariate linear models. Ann. Inst. Stat. Math., 19, pp. 417-437.

3. (With R. L. Maik). (1967). On a new property of partially balanced association schemes useful in psychometric structural analysis. Psychometrika. 32, pp. 279-289.

4. (1968). On a general class of designs for multiresponse experiments. Ann. Math. Stat. pp. 1825-1843.

5. (1969). Some studies on intersection tests in multivariate analysis of variance. Multivariate Analysis II (edited by P. R. Krishnaiah), pp. 145-168. Academic Press, New York.

6. (With L. L. McDonald). (1969). On the costwise optimality of hierarchical multiresponse randomized block designs under the trace criterion. Ann. Inst. Stat. Math. 22.

7. (With D. A. Anderson). (1969). Fractional factorial designs for estimating main effects orthogonal to two-factor interactions: 3^n and $2^m \times 3^n$ series. (ARL Technical Report 69-0123).

8. (With D. A. Anderson). (1970). Optimal fractional factorial plans for main effects orthogonal to two-factor interactions: 2^m series. Journal Amer. Stat. Assoc., 65.

*List of papers and ARL reports authored or co-authored by J. N. Srivastava

9. (With L. L. McDonald). (1970). On the hierarchical two-response (cyclic PBIB) designs, costwise optimal under the trace criterion. Ann. Inst. Stat. Math. 22.

10. (With L. L. McDonald). (1971). Some results on the optimality of a class of hierarchical multiresponse models under the determinant criterion. J. Mult. Analysis.

11. (With L. L. McDonald). (1971). On the extension of Gauss-Markov theorem to a subset of a parameter space, under complex multivariate linear models. Ann. Inst. Stat. Math.

12. (With L. L. McDonald). (1971). On a large class of incomplete multivariate models, which can be transformed to make MANOVA applicable. Metron.

13. (With L. L. McDonald). (1971). Analysis of growth curves under hierarchical models and spline regression, 1. ARL Technical Report 71-0023.

14. (With M. K. Zaater). On the maximum likelihood classification rule for incomplete multivariate samples and its admissibility. To appear in J. Mult. Analysis.

15. (With M. K. Zaater). Incomplete multivariate designs, optimal with respect to Fisher's information matrix.

16. (With D. V. Chopra). (1971). Some new results in the combinatorial theory of balanced arrays of strength four with $2 \leq \mu_2 \leq 6$.

17. (With L. L. McDonald). Estimatibility in fractional factorial designs under the multiple design multiresponse model.

The above work was accomplished on Project 7071, (Research in Applied Mathematics), and was technically monitored by Dr. P. R. Krishnaiah of the Aerospace Research Laboratories. The interest of Dr. Krishnaiah in the work done under the above contract is greatly appreciated.

ABSTRACT

Consider a random vector (X_1,X_2) distributed as a bivariate normal with mean vector zero, and dispersion matrix $\Sigma = ((\sigma_{ij}))$. Suppose we are given samples of sizes n_1 and n_2, respectively, from the marginals of X_1,X_2, and a sample of size n_3 from the bivariate population of (X_1,X_2). Suppose the problem is to obtain a good estimator of Σ based on the above (incomplete) sample. In this paper, four estimators of Σ are compared using Monte Carlo methods, and it is found that a certain relatively simple estimator of Σ is the "best" or close to the best in almost all situations.

TABLE OF CONTENTS

1. Introduction. Consider the problem outlined in the summary. In the literature there are proposed many methods for dealing with this situation, which is also known as the case of missing observations in multivariate statistics. Four such methods are compared in this study. The first was proposed originally by Wilks (1932) for the case of two responses, and later generalized by Kleinbaum (1970) to an arbitrary number p of responses, and to a more general design for the location parameters. The second method is a variant of the first one. The The third method is due to Hocking and Smith (1968). The fourth one corresponds to the principle of maximum likelihood (m.l.). The theoretical evaluation of the optimality properties of these estimators, particularly the third and the fourth, seems to be cumbersome, except for large sample properties, such as consistency and asymptotic efficiency. But our concern here is with the more relevant situations when the sample sizes are not necessarily of the order needed for invoking asymptotic properties. For this reason we resorted to Monte Carlo simulation techniques.

2. <u>The Four Estimators of Σ</u>. The four estimators $E_i (i - 1,2,3,4)$ of Σ will now be spelled out in detail. Let $S_i (i = 1,2)$ denote the sample of size n_i available from the marginal distribution of the ith response, and let S_3 denote the (complete) bivariate sample whose size is n_3. Also, s_{111} and s_{113} will symbolize the mean square of the observations on the first response from S_1 and S_3, respectively. Similar is the definition of s_{222} and s_{223}. The mean cross-product of the first and second response over the units of S_3, is denoted by s_{123}. Finally, let r and ρ be equal to $s_{123}/\sqrt{s_{113}s_{223}}$ and $\sigma_{12}/\sqrt{\sigma_{11}\sigma_{22}}$, respectively.

(i) <u>Estimator E_1</u> $= ((\hat{\sigma}_{ij}))$ is given by

$$(2.1)\quad \hat{\sigma}_{11} = \frac{n_1 s_{111} + n_2 s_{222}}{n_1 + n_2}, \quad \hat{\sigma}_{12} = \hat{\sigma}_{21} = s_{123}, \; \hat{\sigma}_{22} = \frac{n_2 s_{222} + n_3 s_{223}}{n_2 + n_3}.$$

Note that E_1 is not necessarily nonnegative definite.

(ii) <u>Estimator E_2</u> $= ((\sigma_{ij}^{+}))$ is given by: $\sigma_{11}^{+} = \hat{\sigma}_{11}$, $\sigma_{22}^{+} = \hat{\sigma}_{22}$, and $\sigma_{12}^{+} = r\sqrt{\hat{\sigma}_{11}\hat{\sigma}_{22}}$. This is positive definite with probability 1.

(iii) <u>Estimator E_3</u> $= ((\sigma_{ij}^{*}))$. This is given by

$$(2.2)\qquad \sigma_{ij}^{*} = \tilde{\sigma}_{ij} + \beta_{ij}(\tilde{\sigma}_{22} - s_{222}), \quad \text{where}$$

$$(2.3)\qquad \tilde{\sigma}_{ij} = s_{ij3} + \alpha_{ij}(s_{113} - s_{111}),$$

$$(2.4)\qquad (\alpha_{11}, \alpha_{12}, \alpha_{22}) = -\left(\frac{n_1}{n_1 + n_3}\right)\left(1, \frac{s_{123}}{s_{113}}, \frac{s_{123}^2}{s_{113}^2}\right),$$

and

$$(2.5)\quad \begin{bmatrix}\beta_{11}\\ \beta_{12}\\ \beta_{22}\end{bmatrix} = \left[2\frac{\tilde\sigma_{22}^2}{n_3} + 2\frac{\tilde\sigma_{22}^2}{n_2} - \frac{2n_1}{n_3(n_1+n_3)}\cdot\frac{\tilde\sigma_{12}^4}{\tilde\sigma_{11}^2}\right]^{-1} \times \begin{bmatrix} \frac{2\tilde\sigma_{12}^2}{n_1+n_3} \\ \frac{2\tilde\sigma_{12}\tilde\sigma_{22}}{n_3} - \frac{2n_1}{n_3(n_1+n_3)}\frac{\tilde\sigma_{12}^3}{\tilde\sigma_{11}} \\ \frac{2\tilde\sigma_{22}^2}{n_3} - \frac{2n_1}{n_3(n_1+n_3)}\frac{\tilde\sigma_{12}^4}{\tilde\sigma_{11}^2}\end{bmatrix}$$

(iv) <u>Estimator $E_4 = ((\hat\sigma_{ij}))$</u>. This is the maximum likelihood estimator. To derive it, let L denote the likelihood function of the total sample. Then

$$(2.6)\quad \frac{\partial \log L}{\partial\Sigma} = -\frac{n_1}{2}\begin{bmatrix}\sigma_{11}^{-1} & 0\\ 0 & 0\end{bmatrix} - \frac{n_2}{2}\begin{bmatrix}0 & 0\\ 0 & \sigma_{22}^{-1}\end{bmatrix}$$

$$-\frac{n_3}{2}\frac{1}{|\Sigma|}\begin{bmatrix}\sigma_{22} & -\sigma_{12}\\ -\sigma_{12} & \sigma_{11}\end{bmatrix} + \frac{n_1}{2}\begin{bmatrix}\sigma_{11}^{-2}S_{111} & 0\\ 0 & 0\end{bmatrix}$$

$$+\frac{n_2}{2}\begin{bmatrix}0 & 0\\ 0 & \sigma_{22}^{-1}S_{222}\end{bmatrix} + \frac{1}{2}\frac{n_3}{|\Sigma|^2}$$

$$\times\begin{bmatrix}\sigma_{22}^2S_{113} - 2\sigma_{12}\sigma_{22}S_{123} + \sigma_{12}^2S_{223} & -\sigma_{12}\sigma_{22}S_{113} + (\sigma_{12}^2 + \sigma_{11}\sigma_{22})S_{123} - \sigma_{12}\sigma_{11}S_{223}\\ -\sigma_{12}\sigma_{22}S_{113} + (\sigma_{12}^2 + \sigma_{11}\sigma_{22})S_{123} - \sigma_{12}\sigma_{11}S_{223} & \sigma_{11}^2S_{223} - 2\sigma_{12}\sigma_{11}S_{123} + \sigma_{12}^2S_{113}\end{bmatrix}.$$

Let $\theta_{11} = s_{113}\sigma_{11}^{-1}$, $\theta_{12} = s_{123}\sigma_{12}^{-1}$, $\theta_{22} = s_{223}\sigma_{22}^{-1}$, $f_1 = s_{111}S_{113}^{-1}$, $f_2 = s_{222}S_{223}^{-1}$. Substituting these in the likelihood equation obtained by equating $\partial \log L/\partial\Sigma$ to zero, and doing some simplification, we arrive at the following set of three equations that are linear in the parameters θ_{ij}

$$[n_1(1-\rho^2)^2 f_1+n_3]\theta_{11}-2n_3\rho^2\theta_{12}+n_3\rho^2\theta_{22} = (1-\rho^2)[n_1(1-\rho^2) + n_3] ,$$

$$n_3\rho^2\theta_{11} - 2n_3\rho^2\theta_{12} + [n_2(1-\rho^2)^2 f_2 + n_3]\theta_{22} = (1-\rho^2)[n_2(1-\rho^2) + n_3],$$

$$\theta_{11} - (1 + \rho^2)\theta_{12} + \theta_{22} = 1 -\rho^2.$$

These give

$$(2.7) \quad \hat{\hat{\theta}}_{11} = \Delta^{-1}[n_1n_2(1 -\rho^4)f_2 + n_2n_3(1 -\rho^2)f_2 + n_2n_3\rho^2 + n_1n_3 + n_3^2],$$

$$(2.8) \quad \hat{\hat{\theta}}_{22} = \Delta^{-1}[n_1n_2(1 - \rho^4)f_1 + n_1n_3(1 - \rho^2)f_1 + n_1n_3\rho^2 + n_2n_3 + n_3^2],$$

$$(2.9) \quad \hat{\hat{\theta}}_{12} = (-\Delta^{-1}) \{n_1n_2(1 - \rho^2)[(1 - \rho^2)f_1f_2 - f_1 - f_2] - n_3n\} ,$$ where

$$n = n_1 + n_2 + n_3, \text{ and } \Delta = n_1n_2(1 - \rho^4)f_1f_2 + n_1n_3f_1 + n_2n_3f_2 + n_3^2 .$$

Thus, we obtain $\hat{\hat{\sigma}}_{ij} = s_{ij3} / \hat{\hat{\theta}}_{ij}$, $(i,j = 1,2)$. Also, by invoking the invariance property of maximum likelihood (m.l.) estimation we find that the m.l. estimator $\hat{\rho}^2$ of ρ^2 must satisfy

$$(2.10) \quad \hat{\hat{\rho}}^2 = (\hat{\hat{\sigma}}_{12})^2(\hat{\hat{\sigma}}_{11}\hat{\hat{\sigma}}_{22})^{-1} = (\hat{\hat{\theta}}_{12})^{-2} (\hat{\hat{\theta}}_{11}\hat{\hat{\theta}}_{22})r^2 .$$

The above implies that the m.l. estimator of Σ is a positive definite matrix with probability one. We also get

$$\rho^2\hat{\hat{\theta}}_{12}^2 - \hat{\hat{\theta}}_{11}\hat{\hat{\theta}}_{22} r^2 = 0.$$

Substituting the values of $\hat{\hat{\theta}}_{11}$, $\hat{\hat{\theta}}_{22}$ and $\hat{\hat{\theta}}_{12}$ from (2.7-2.9), we obtain, after some simplification, the equation:

$$
\begin{aligned}
(2.11)\quad & \rho^{10}(n_1^2n_2^2f_1^2f_2^2) + \rho^8[n_1^2n_2^2f_1f_2(-4f_1f_2+2f_1+2f_2-r^2)] + \rho^6[n_1^2n_2^2(6f_1^2f_2^2-6f_1^2f_2 \\
& -6f_1f_2^2 + f_1^2 + f_2^2 + 2f_1f_2) - 2n_1n_2n_3nf_1f_2 - n_1n_2n_3r^2(n_1f_1f_2 + n_2f_1f_2 - n_2f_1-n_1f_2)] \\
& + \rho^4 \{n_1^2n_2^2(-4f_1^2f_2^2 + 6f_1^2f_2 + 6f_1f_2^2 - 2f_1^2 - 2f_2^2 - 4f_1f_2) + 2n_1n_2n_3n(2f_1f_2-f_1-f_2) \\
& - r^2n_1n_2[(-2n_1n_2 - n_1n_3 -n_2n_3 + n_3^2)f_1f_2 + n_3(-n_1 - 2n_3)f_1 + n_3(-n_2 - 2n_3)f_2+n_3^2\} \\
& \rho^2 \{n_1^2n_2^2(f_1^2f_2^2 - 2f_1^2f_2 - 2f_1f_2^2 + f_1^2 + f_2^2 + 2f_1f_2) - 2n_1n_2n_3n(f_1f_2 - f_1 - f_2) \\
& + n_3^2n^2 - r^2[n_1n_2n_3f_1f_2(-n_1 -n_2 -2n_3) + n_1n_3f_1(n_2^2 + n_2n_3 - n_1n_3 - n_3^2) + n_2n_3f_2(n_1^2 \\
& + n_1n_3 - n_2n_3 - n_3^2) + n_3^2(n_1^2 + n_2^2) + n_3^3(n_1 + n_2)]\} - r^2(n_1n_2 + n_1n_3 + n_2n_3 + n_3^2) \\
& (n_1n_2f_1f_2 + n_1n_3f_1 + n_2n_3f_2 + n_3^2) = 0.
\end{aligned}
$$

Let $f(\rho^2)$ denote the left hand side of (2.11). Then $f(\rho^2) = 0$, is a fifth degree polynomial equation in ρ^2 with stochastic coefficients, and every root it has between zero and one can be substituted in (2.7-2.9), leading to a set of maximum likelihood estimators $\hat{\sigma}_{11}$, $\hat{\sigma}_{22}$ and $\hat{\sigma}_{12}$. Finding (closed) exact solutions for the above polynomial does not seem to be an easy task. However, our program of the Monte Carlo study (which we shall explain later) is designed to evaluate, by a simple iterative method. the roots of this equation which lie in the interval (0,1), to any pre-assigned degree of precision. Nevertheless, at this point we can make the following observations. If we put $\rho^2 = 1$, we find after some simplification that $f(1) = n_3^2n^2[1 - r^2] > 0$, with probability one. Similarly, we find that

$$f(0) = -r^2(n_1n_2 + n_1n_3 + n_2n_3 + n_3^2)(n_1n_2f_1f_2 + n_1n_3f_1 + n_2n_3f_2 + n_3^2) < 0.$$

From this we conclude that (2.11) has at least one root in the interval (0,1) with probability one.

3. <u>The Simulation Technique</u>. The Monte Carlo study conducted here involves, as a first step, the generation of both univariate and bivariate normal random samples. For this purpose, we first generate independent random variables U_1, U_2 distributed uniformly over the interval (0,1). Let $X_1 = (-2 \log_e U_1)^{\frac{1}{2}} \cos(2\pi U_2)$, $X_2 = (-2 \log_e U_1)^{\frac{1}{2}} \sin(2\pi U_2)$. Then it is well known that X_1 and X_2 are independent standard normal random variables. To generate a sample $(Y_1, Y_2)'$ from a normal population with dispersion matrix Σ, we take $(Y_1, Y_2)' = T(X_1, X_2)$, where $\Sigma = TT'$. To examine how accurately the computer is approximating sampling from a uniform population, 87 samples of 10,000 observations each, were generated in a first run, and 44 samples of 10,000 observations each, were generated in a second run. A chi square goodness of fit test, with 9 degrees of freedom, was carried out for each of the samples, and the probabilities of the X^2 were about 0.45 in both cases.

4. <u>Description of Computer Input and Output</u>. The set of input parameters consists of n_1, n_2, n_3, σ_{11}, σ_{22} and ρ. As noted earlier, σ_{11} was always taken equal to 1, and interchanging the values of n_1 and n_2, serves to avoid any loss of generality that may stem from this choice of σ_{11}. The following table shows the different choices of the set of values of the sample sizes (n_1, n_2, n_3) in our Monte Carlo study.

Table 1. Sample size values.

Group size	n_1	n_2	n_3	n_1	n_2	n_3	Group size
$\frac{1}{2}$ C	3	7	5	28	12	10	2C
	7	3	5	8	8	32	
	2	2	8	32	32	8	
	8	8	2	40	0	20	
	10	0	5	42	18	30	3C
	0	10	5	12	12	48	
C	6	14	10	48	48	12	
	14	6	10	60	0	30	
	4	4	6	56	24	40	4C
	16	16	4	16	16	64	
	20	0	10	64	64	16	
	0	20	10	80	0	40	
				70	30	50	5C
				20	20	80	
				80	80	20	
				100	0	50	

For each chosen set of values for the input parameters, 500 samples were drawn, and at the end of each sample the estimators $\tilde{\tilde{\sigma}}_{ij}$ $(i,j = 1,2)$ are calculated according to the formulas of each of the four methods under consideration. Then a mean cross-product deviation matrix V was calculated and printed out for each method over the 500 samples. Thus the (3 x 3) matrix $V_r = ((v_{r,(i,j),(i',j')}))$, where (i,j), $(i',j') = (1,1)$, (2.2), and $(1,2)$, and $v_{r,(i,j),(i',j')} = (500)^{-1} \sum_{u=1}^{500} \{ (\tilde{\tilde{\sigma}}_{r,i,j} - \sigma_{ij})(\tilde{\tilde{\sigma}}_{r,i',j'} - \sigma_{i',j'}) \}$. We also print out $|V|^{1/3}$ and (1/3 trV, for each V. In the beginning of the study and for values of (n_1,n_2,n_3) of the order C and (1/2)C we chose four values of σ_{22}, namely $\sigma_{22} = 1,2,5$ and 10, and fine values of ρ, namely $\rho = -.7,-.3,+.1,+.5,+.9$; which means that each fixed choice of (n_1, n_2, n_3) was repeated twenty times. However, after careful examination of the output matrix V, and in order to study the effect of the values of σ_{22}, each V was transformed to DVD, where $D = (1,\sigma_{22}^{-1},\sigma_{22}^{-1/2})$. The matrix DVD corresponding to the estimator E_1,E_2 and E_4 clearly exhibited its stability or invariance with respect to changes in the (scaling) parameter σ_{22}. However, the determinant and the trace of the matrix DVD for the estimator E_3, tended to increase monotonically with σ_{22}. It was decided then, that only the case $\sigma_{22} = 1$, should be considered in the input, and additional values of ρ were introduced as follows: σ_{22} and σ_{11} were taken to equal 1 in all cases; and for values of (n_1,n_2,n_3) in the group sizes 1/2C and C, ρ took the values -.7,-.3,.1,.5,.7 and .9; and for values of (n_1,n_2,n_3) in the group sizes 2C, 3C, 4C and 5C, the values -.7,-.3,-.1,.1,.3,.5 and .9 were assumed by ρ.

It may be useful to make a few remarks on the number of solutions of the equation $f(\rho^2) = 0$. In most cases, we found only one root inside the interval (0,1). In a few cases, we found that $f(\rho^2)$ had exactly three roots in the unit interval. This indicated that the solution of the maximum likelihood equations was not unique (this situation occured only for very small values of ρ_3). In this case we computed the logarithm of the likelihood function L, and took the root that maximized log L.

We also carried out an (indirect) overall test for the normality assumption concerning the samples obtained in this study. This consisted of looking at the first and the second diagonal elements (say, v_1, v_2) of V_1. Using the assumption of normality of the various samples, one can easily calculate the mean and variance of both v_1 and v_2 was computed. Next, we note that v_i (i=1,2) is the average of 500 independent random variables, and hence using $v_i^* = [v_i - E(v_i)]/\sqrt{\text{var } v_i}$, may be considered as an observation from a standard normal population. The values of v_1 considered in this test arose from all choices of (n_1,n_2,n_3). Using the cases $\sigma_{11} = \sigma_{22} = 1$ and ρ - -.7, -.3, +.1, +.5, +.9, we got 120 observations on each of v_1^* and v_2^*. These 240 observations were then tested for normality using a X^2; the probability of the X^2 exceeding its observed value was less than 0.45.

In the above test, we could have similarly used various other elements of the matrix V_i(i=1,2,3,4), but we felt that the two elements actually used would be sufficient. Although the above is not a conclusive proof of the normality of the data, it seems to be one of the best procedures that could be used under the present circumstances. A direct check of the normality of all the samples would clearly have been too expensive. Furthermore, even if each sample were tested individually, we could not meaningfully conclude the presence of normality in general.

5. <u>The Results of the Study and the Associated Plots</u>. The main results of this study are depicted in a series of plots. Plots 1 through 4 are constructed as follows.

The different values of the correlation coefficient ρ are indicated on the horizontal axis. The vertical axis exhibits the values of the determinant of the matrix $V(3 \times 3)$ corresponding to the four different estimators and the different input values of n_1, n_2, n_3 and ρ, with $\sigma_{11} = \sigma_{22} = 1$. The horizontal axis for plot 5 corresponds to the total number of observations $(n_1 + n_2 + 2n_3)$ taken for size orders (1/2)C, C, 2C, 3C, 4C and 5C. The points on the vertical axis represent the value of $|V|$ (averaged over the different values of ρ) for the various $E_i (i=1,2,3,4)$ and the different designs D_i, where (D_1, D_2, D_3, D_4) correspond to the values of $\underline{n} = (n_1, n_2, n_3)$ being (7,3,5), (10,0,5), (8,8,2) and (2,2,8), respectively. These values are for $\underline{n}$ of the order (1/2)C. Values of $\underline{n}$ of the form, say, (7,3,5) and 3,7,5), belong to the same design and $|V|$ was averaged over them.

By analyzing the computer output directly and examining the preceding plots, one can make the following observations.

(1) The comparative efficiency of the estimators is the same under both the determinant and the trace criterion for V. Thus we restricted attention in drawing the plots to the determinant criterion. However, the determinant criterion is also more meaningful here, since the parameters being estimated may not necessarily be in the same scale.

(2) All estimators possessed a maximum in the neighborhood of the point $\rho = .1$ (one expects this point to be $\rho = 0$), except for the designs (8k,8k,2k); k=1,2,4,6,8,10) for which E_1 had a minimum there. However, under trV, E_1 always possessed a minimum at that point, while E_2, E_3 and E_4 still had a maximum there.

(3) For the HM design, E_3 and E_4 coincided, as expected theoretically.

(4) For almost all cases, E_1 and E_2 coincided at $\rho = .1$.

(5) E_3 behaved extremely badly for small values of n_3, and thus should not be considered in the designs (8k,8k,2k). This estimator was never the best, and it was always the worst, except for very large values of n_3 and $\rho > .7$, where it, sometimes, became the second best behind E_4.

(6) E_2 is the best for $.3 < \rho < .5$, for $\rho < .8$ in the case of samples of order 1/2C, and for $\rho < .85$ in (16,16,4).

(7) E_4 is the best for very large values of n_3 and large values of ρ.

(8) E_1 performs its best for small values of ρ.

(9) Aside from E_3,E_4 had the largest range, being larger than E_1 and E_2 for small ρ, and having the smallest value for $\rho = .9$.

(10) E_1 is the most stable as a function of ρ, followed by E_2.

(11) Differences among the estimators are negligible for sample sizes of order 5C, except E_3 in (80,80,20), where it joins the others only for $\rho > .6$.

In the final analysis, one may conclude that for samples of sizes 1/2C and C, E_2 is highly recommended for all values of $\rho < .8$. For sample sizes of order kC(k=2,3,4,5), E_2 is to be used except for $|\rho| > .5$, where E_4 becomes the most efficient. However, the simplicity of E_2 should count heavily in its favor, particularly when a small gain in efficiency by using E_4 is not very crucial.

REFERENCES

1. Hartley, H. O. and Hocking, R. R. (1971). Incomplete data analysis. Presidential invited lecture, ENAR and IMS Meeting, College Park, Penn.

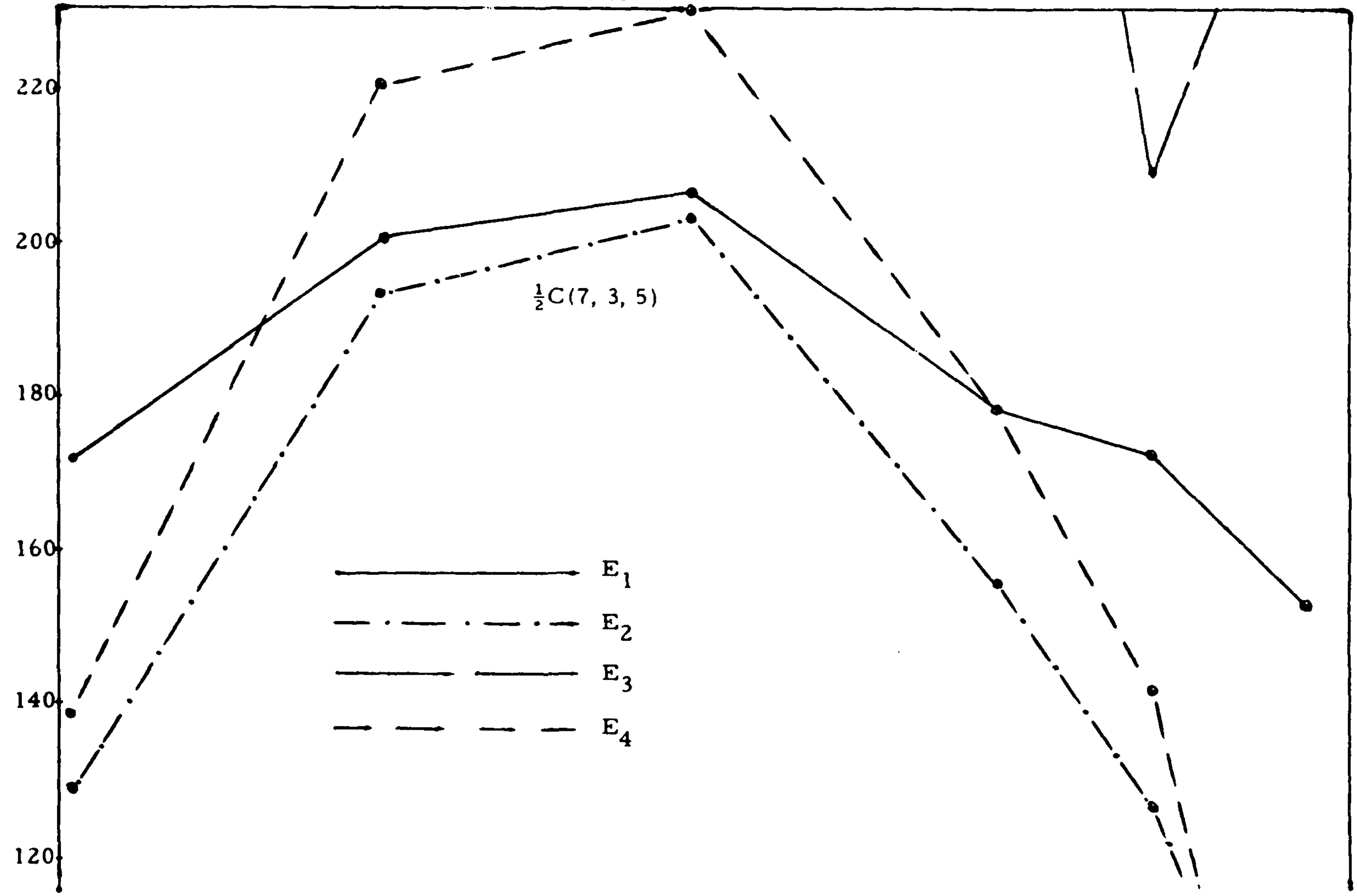
2. Hocking, R. R. and Smith, W. B. (1968). Estimation of parameters in the multivariate normal distribution with missing observations. J. Amer. Stat. Assoc. 63, 154-173.

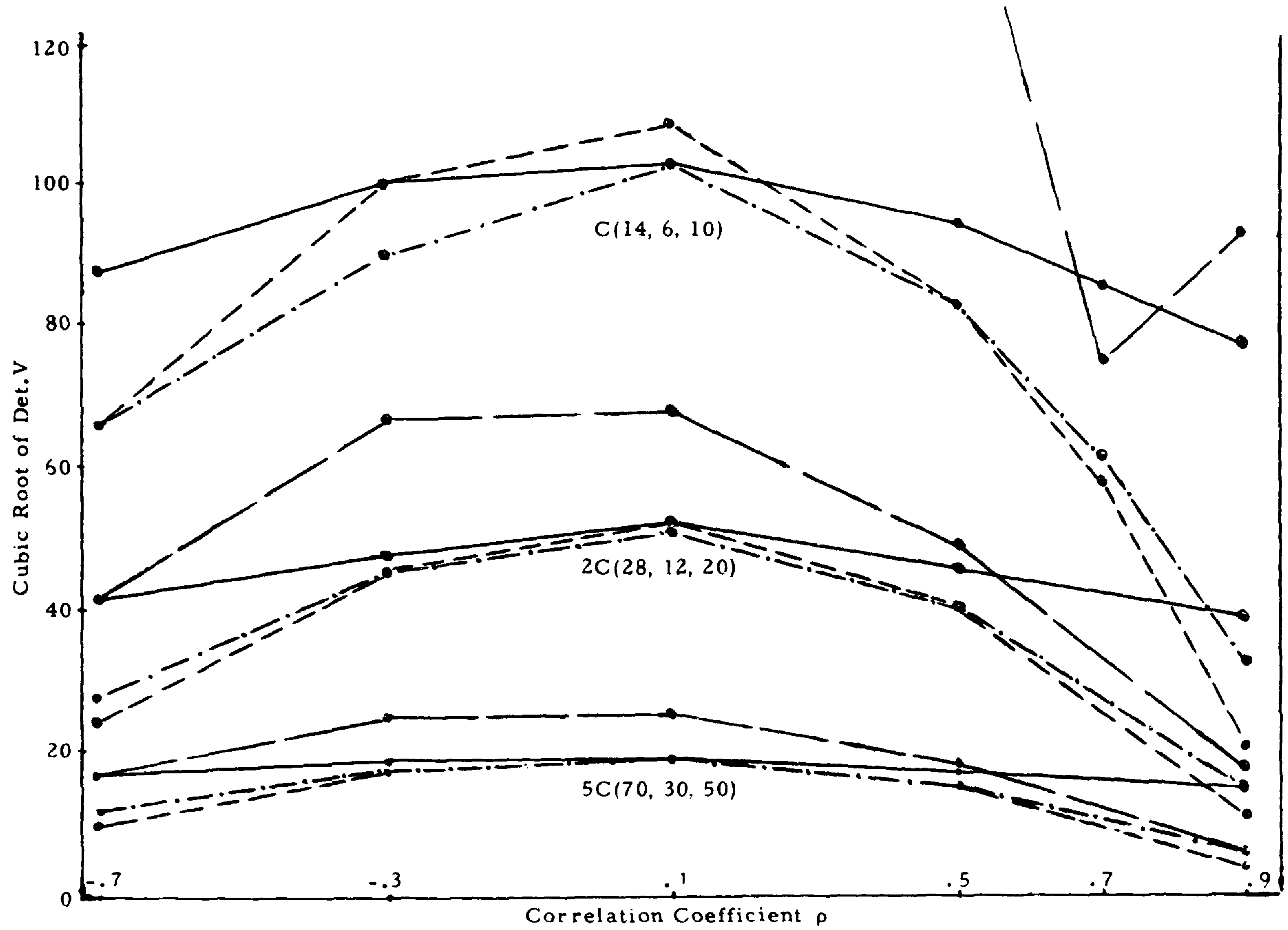
3. Kleinbaum, D. G. (1970). Estimation and hypothesis testing for generalized multivariate linear models. Unpublished thesis, University of North Carolina, Chapel Hill, N.C.

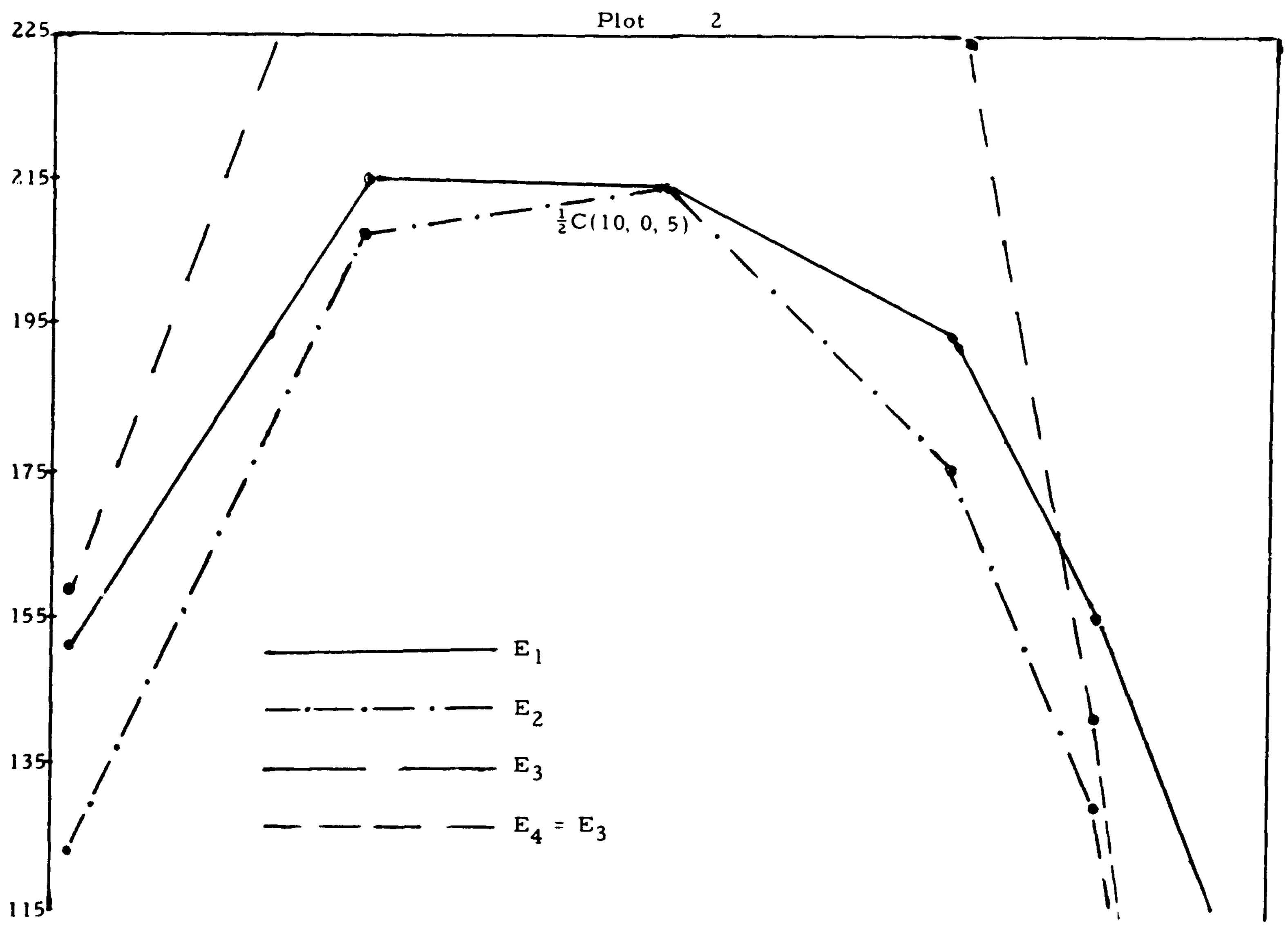
4. Roy, S. M., Gnanadesikan, R. and Srivastava, J. N. (1970). Analysis and Design of Certain Multiresponse Experiments. Pergamon Press, London.

5. Wilks, S. S. (1932). Moments and distributions of estimates of population parameters from fragmentary samples. Ann. Math. Stat. 3, 167-195.

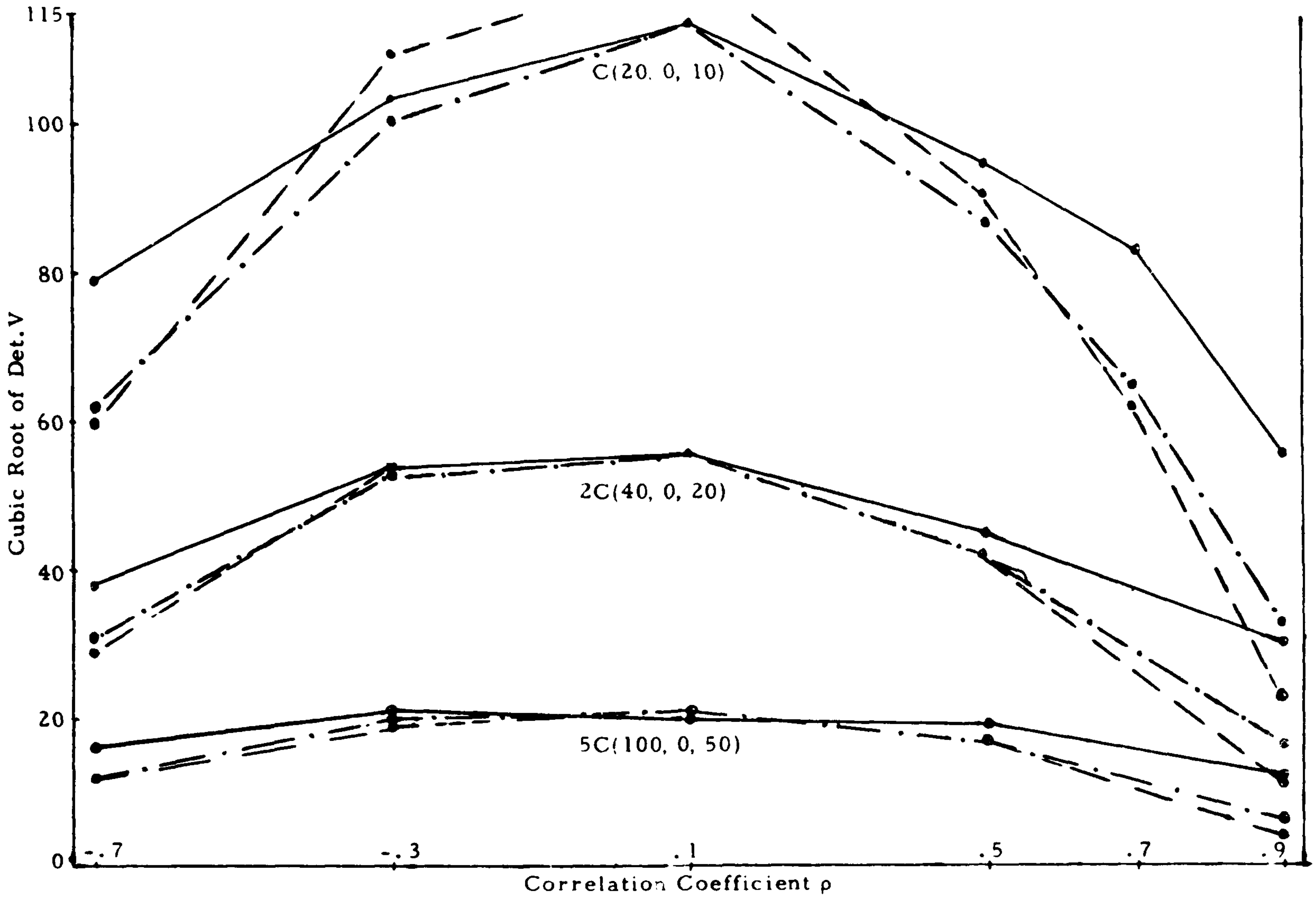
Plot 1




120
100
80
60
40
20
0
Cubic Root of Det.V
C(14, 6, 10)
2C(28, 12, 20)
5C(70, 30, 50)
-.7
-.3
.1
.5
.7
.9
Correlation Coefficient ρ

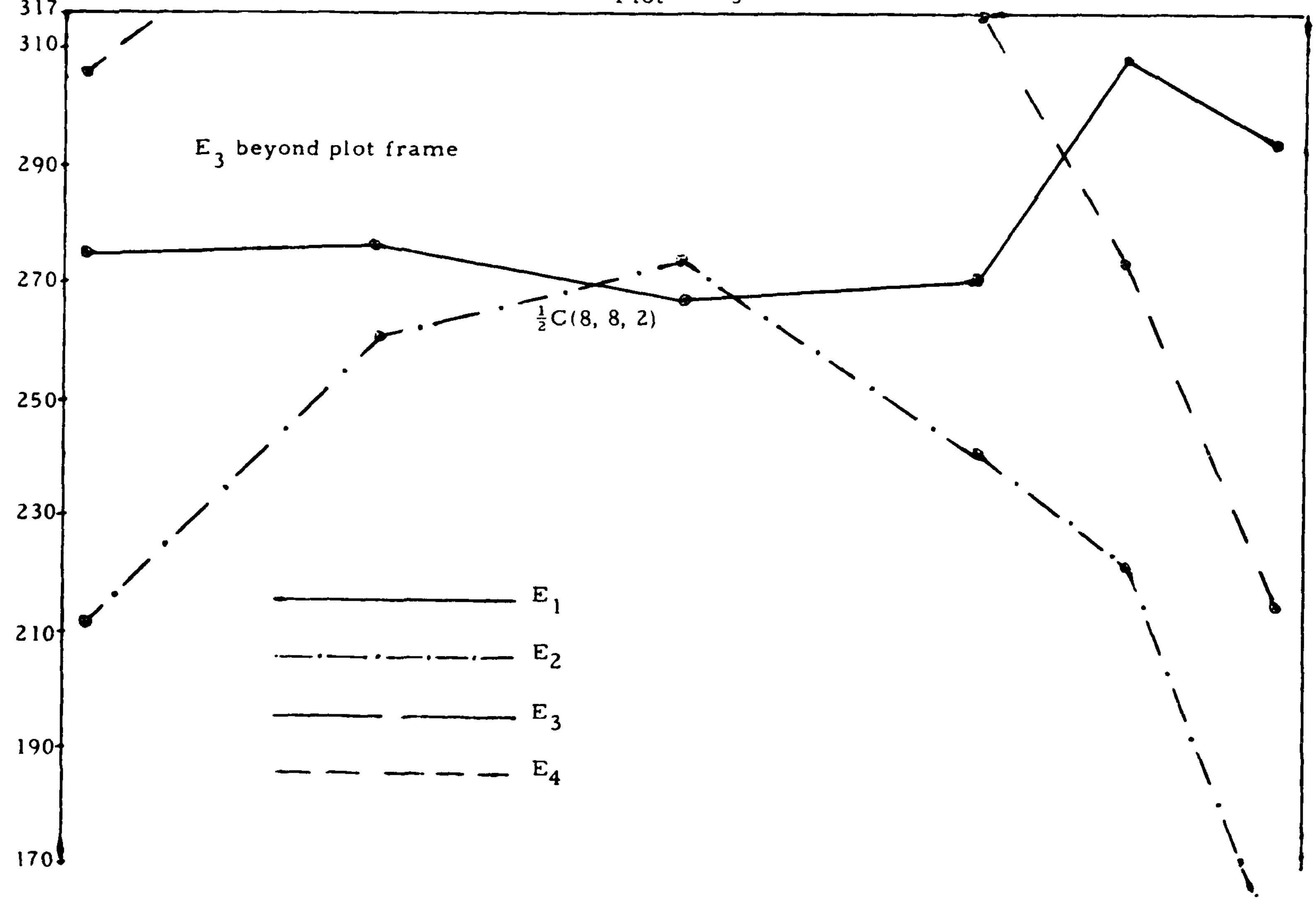


Plot 2
225
215
195
175
155
135
115
½C(10, 0, 5)
E1
E2
E3
E4 = E3

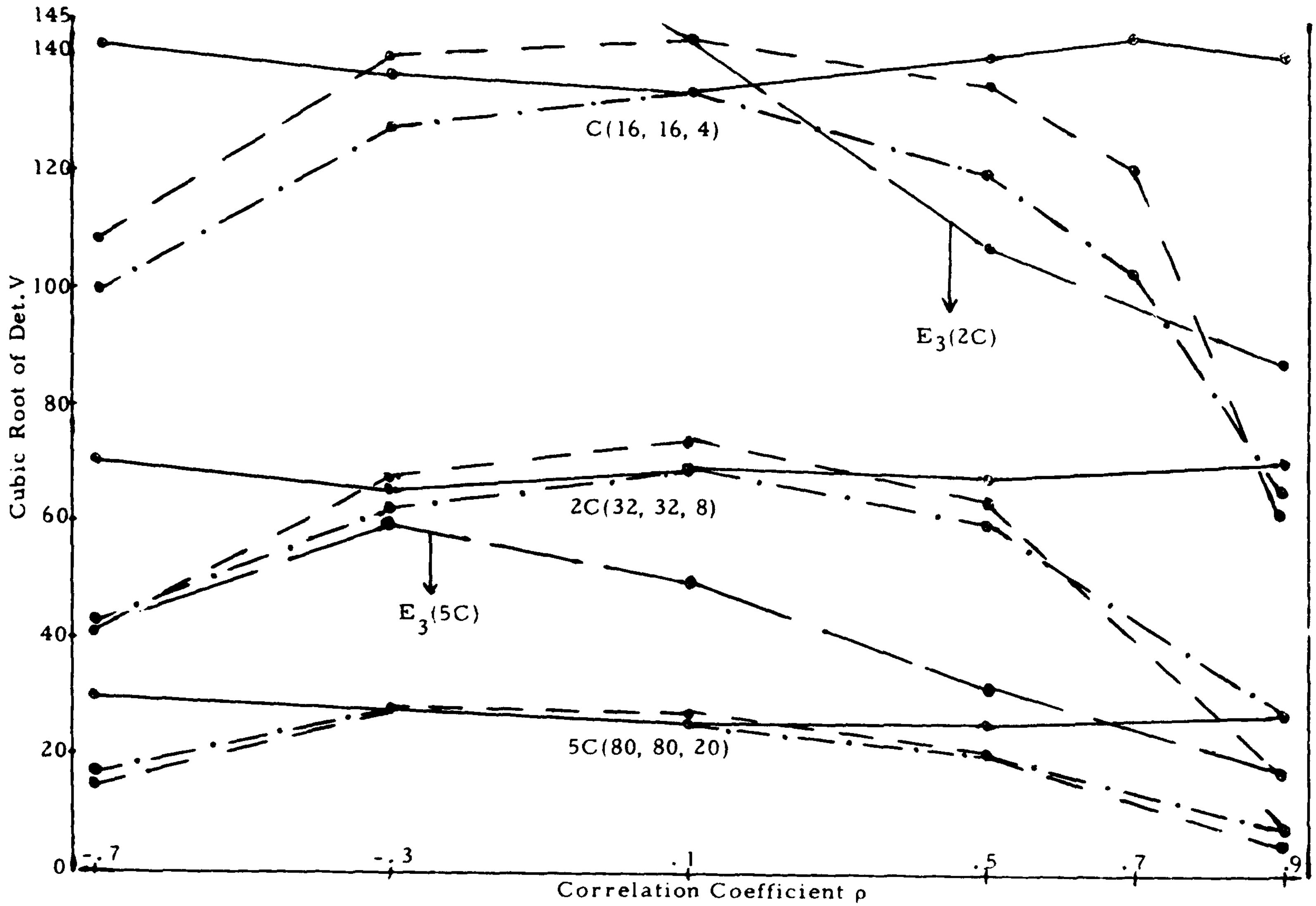


115
100
80
60
40
20
0
Cubic Root of Det.V
C(20, 0, 10)
2C(40, 0, 20)
5C(100, 0, 50)
-.7
-.3
.1
.5
.7
.9
Correlation Coefficient ρ

Plot 3

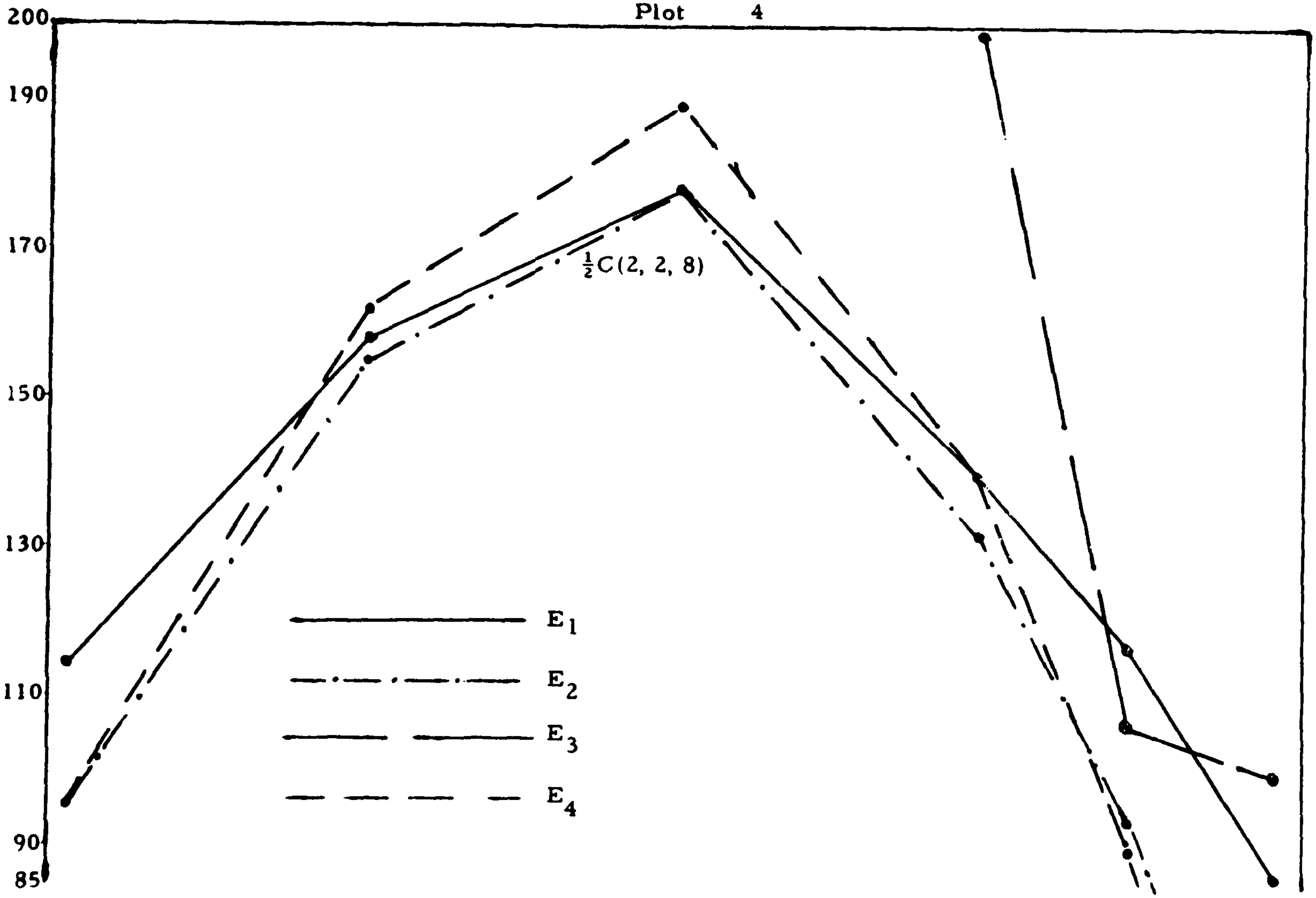


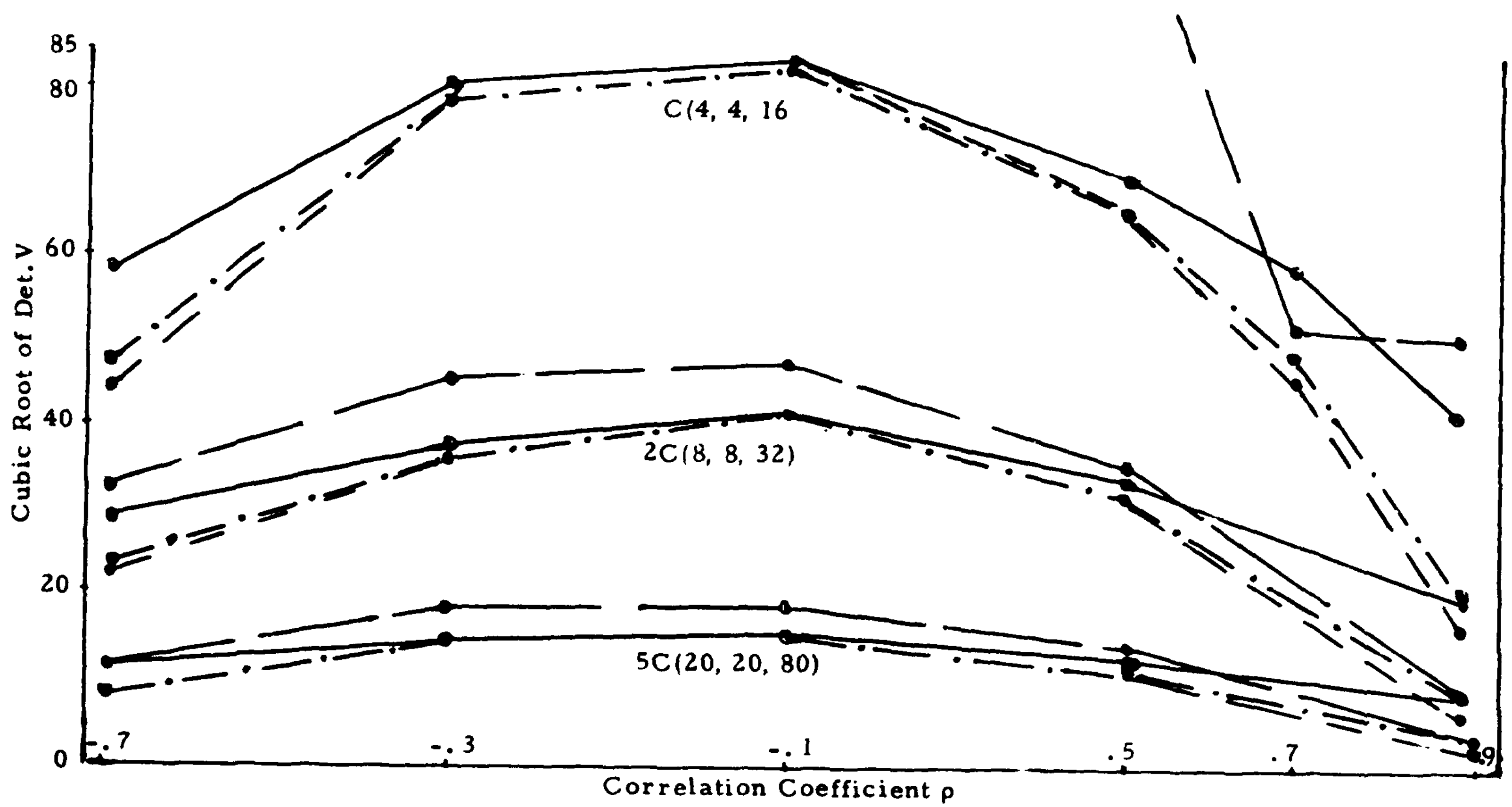
317
310
290
270
250
230
210
190
170
E_3 beyond plot frame
$\frac{1}{2}C(8, 8, 2)$
E_1
E_2
E_3
E_4



Cubic Root of Det.V
145
140
120
100
80
60
40
20
0
C(16, 16, 4)
E$_3$(2C)
2C(32, 32, 8)
E$_3$(5C)
5C(80, 80, 20)
-.7
-.3
.1
.5
.7
.9
Correlation Coefficient ρ

Plot 4





85
80
60
40
20
0
Cubic Root of Det. V
C(4, 4, 16
2C(8, 8, 32)
5C(20, 20, 80)
-.7
-.3
-.1
.5
.7
.9
Correlation Coefficient ρ

Unclassified
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Department of Statistics Colorado State University Fort Collins, Colorado 80521	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP M-2

3. REPORT TITLE

A Monte Carlo Comparison of Four Estimators of the Dispersion Matrix of a Bivariate Normal Population, Using Incomplete Data

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Scientific Final

5. AUTHOR(S) (First name, middle initial, last name)

J. N. Srivastava and M. K. Zaatar

6. REPORT DATE February 1972	7a. TOTAL NO. OF PAGES 26	7b. NO. OF REFS 5
8a. CONTRACT OR GRANT NO. F 33615-67-C-1436 b. PROJECT NO. 7071-00-12 c. DoD Element 61102F d. DoD Subelement 681304	9a. ORIGINATOR'S REPORT NUMBER(S) 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) ARL 72-0032	

10. DISTRIBUTION STATEMENT

Approved for public release; distribution unlimited

11. SUPPLEMENTARY NOTES TECH OTHER	12. SPONSORING MILITARY ACTIVITY Aerospace Research Laboratories (LB) Wright-Patterson AFB, Ohio 45433

13. ABSTRACT

Consider a random vector (X_1,X_2) distributed as a bivariate normal with mean vector zero, and dispersion matrix $\Sigma = ((\sigma_{ij}))$. Suppose we are given samples of sizes n_1 and n_2, respectively, from the marginals of X_1,X_2, and a sample of size n_3 from the bivariate population of (X_1,X_2). Suppose the problem is to obtain a good estimator of Σ based on the above (incomplete) sample. In this paper, four estimators of Σ are compared using Monte Carlo methods, and it is found that a certain relatively simple estimator of Σ is the "best" or close to the best in almost all situations.

DD FORM 1 NOV 65 1473

Unclassified
Security Classification

Unclassified
Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Bivariate Normal						
Estimators						
Incomplete Data						
Dispersion Matrix						

Unclassified
Security Classification

U.S. Government Printing Office: 1972 — 759-085/532