LINEAR ESTIMATION IN CENSORED SAMPLES FROM MULTIVARIATE NORMAL POPULATIONS

By G. A. Watterson

Australian National University

1. Summary. In this paper, the known methods of linear estimation are extended to various cases of censored samples from multivariate normal populations. The two estimators considered correspond to the minimum variance and the 'alternative' estimators treated by Gupta [3] and Sarhan and Greenberg [4], [5] for univariate samples. It is found that the 'alternative' estimator has important applications in multivariate samples, being easy of computation and of low variance.

2. Introduction. During recent years, the estimation of population parameters from censored samples has received considerable attention. Gupta [3] found both maximum likelihood and linear estimators for the mean and standard deviation of a univariate normal distribution using a sample from which a number of the largest observations had been censored. The maximum likelihood method of estimation was used in the more general case of censoring from a multivariate distribution by Cohen [1], who laid emphasis on samples restricted to a fixed region of possible population values rather than on samples with a fixed number of observations missing; but his results also apply to the latter case after minor modifications (Watterson [8]).

The advantages of maximum likelihood estimators are well known, the most important being the properties of asymptotic efficiency and unbiasedness. But for estimation from small censored samples neither the bias nor the exact variance can be calculated for these estimators, and the actual computing of the estimates is considerable, involving iterative solution of the likelihood equations.

In contrast, linear estimation has the following advantages. Firstly, for most parameters it is possible to obtain unbiased estimators and to calculate their variances, and secondly, linear estimates are easy to compute once the coefficients have been found. Further, for certain special cases Gupta has shown that linear estimators are not substantially less efficient than those obtained by maximum likelihood.

These reasons have given a motivation for further study of linear estimators, and Gupta's original methods have been generalised to doubly censored samples (having both large and small values missing) from univariate normal and exponential distributions (Sarhan and Greenberg [4], [5], [6]). The theory is here further extended to linear estimation from multivariate populations with censoring effective on interior as well as extreme variates in the sample.

Received June 27, 1958; revised December 18, 1958.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to The Annals of Mathematical Statistics. www.jstor.org

3. Censored samples. Consider a k-variate normal population, the moments of the variates being determined by the parameters

$$\mu_i = E(x_i), \qquad \sigma_{ij} = \text{Cov}\,(x_i, x_j), \qquad i, j = 1, 2, \cdots, k.$$

From this population a sample of size n is drawn, containing in all $n \times k$ variables. Suppose the sample is ordered with respect to one variate, say x_1, so that it may be represented by

$$\begin{array}{llll} x_{11} & < x_{12} & < \cdots & < x_{1n} \\ x_{2(1)}, & x_{2(2)}, & \cdots, & x_{2(n)} \\ \vdots & & & \\ x_{k(1)}, & x_{k(2)}, & \cdots, & x_{k(n)}, \end{array}$$

where a bracket on the second subscript indicates the association of the variable with the corresponding ordered variable having no such bracket. Note that the associated variables are not necessarily in increasing order of magnitude.

We define a censored sample as one having some or all observations missing from a number of the vectors of the ordered sample. For definiteness, censoring will be subdivided into three distinct types, each of which have practical applications.

Type A: Censoring effective on all variates of certain sample vectors,

Type B: Censoring of associate variates only,

Type C: Censoring of the ordered variate only.

We shall not restrict the missing observations to belonging to the first or last vectors, but of course these cases are included in our formulation. Samples censored from one end only are chosen to typify the three types of sample from a singly censored bivariate distribution.

A. The 5 tallest trees out of a group of 20 are removed for milling. Their heights and mean diameters are measured so that an estimate of the volume of timber remaining in the group may be made. The two measured variates form a type A sample.

B. The examination scores of 17 students are known, but only the best 12 students are allowed to proceed to the next year of the course. The associated variate, the examination scores after one year's study, is thus censored, but the ordered variate is completely known.

C. Densities of several metal alloy specimens are measured, and each specimen is then subjected to fatigue testing. Supposing that the specimens are set in operation simultaneously, the population parameters may be estimated sequentially as the specimens fail, but at each stage the sample available will be censored with respect to the ordered variate 'time to failure', whilst the associated 'density' is completely known. In fact, the first estimates may be made after only two failures.

4. Linear estimation. For these types of samples, the estimation of all possible parameters may be inferred from the univariate and bivariate cases, which will therefore be treated in detail.

(a) *Univariate case.* The estimation of the parameters μ_1 and $\sqrt{\sigma_{11}}$ for the ordered variate has been carried out by Gupta [3] and Sarhan and Greenberg [4], [5] in the most usual cases of single and double censoring. For the ordered sample

$$x_{11} < x_{12} < \cdots < x_{1n}$$

define

$$(1) \qquad u_l = \sigma_{11}^{-\frac{1}{2}} E(x_{1l} - \mu_1), \qquad v_{lm} = \sigma_{11}^{-1} \operatorname{Cov}(x_{1l}, x_{1m})$$

as being the means and covariances of ordered standard normal variables. The values of u_l and v_{lm} are tabulated by Teichroew [7] and Sarhan and Greenberg [4] respectively, for samples of size $n \leqq 20$. Suppose now that some of the ordered sample variables are missing; then a linear estimator will have the form $\sum \alpha_l x_{1l}$ where the summation extends over all values of l for which observations are available. (This summation convention will be adhered to throughout the paper). The mean and variance of this estimator are, from (1),

$$(2) \qquad E(\textstyle\sum \alpha_l x_{1l}) = \mu_1 \sum \alpha_l + \sqrt{\sigma_{11}} \sum \alpha_l u_l,$$

$$(3) \qquad \operatorname{Var}(\textstyle\sum \alpha_l x_{1l}) = \sigma_{11} \sum \sum \alpha_l \alpha_m v_{lm}.$$

If we choose α_l so that $\sum \alpha_l = 1$, $\sum \alpha_l u_l = 0$, then $\sum \alpha_l x_{1l}$ is an unbiased estimator for μ_1; similarly, if $\sum \alpha_l = 0$, $\sum \alpha_l u_l = 1$, then $\sum \alpha_l x_{1l}$ is an unbiased estimator for $\sqrt{\sigma_{11}}$. Obviously, these conditions do not determine the coefficients, and we may impose the further restriction that the variance (3) be made a minimum. Sarhan and Greenberg [4], [5] tabulate two sets of values α_{1l} and α_{2l} such that $\sum \alpha_{1l} x_{1l}$ and $\sum \alpha_{2l} x_{1l}$ are unbiased linear estimators for μ_1 and $\sqrt{\sigma_{11}}$ of minimum variance, in the particular cases of singly and doubly censored samples of size $n \leqq 15$.

For the more general case where observations may be missing from the interior of the sample, no tables have been constructed since the number of possible sample types is of order 2^n for each n, and the task of computing and tabulating the coefficients even for reasonably small n would be excessive. Instead, an alternative estimator can be constructed with simple computational properties and which is not very less efficient than the optimal one. Gupta [3] suggested for single censoring that instead of minimising the variance (3), the coefficients obtained by minimising $\sum \alpha_l^2$ subject to the unbiasedness conditions gave an estimator of low variance. In our case, suppose there are p out of the n variates observable, and write $\bar{u} = p^{-1} \sum u_l$. Then with

$$(4) \qquad \beta_{1l} = p^{-1} - \frac{\bar{u}(u_l - \bar{u})}{\sum_m (u_m - \bar{u})^2}, \qquad \beta_{2l} = \frac{u_l - \bar{u}}{\sum_m (u_m - \bar{u})^2},$$

we find that $\sum \beta_{1l} x_{1l}$ is an unbiased estimator of μ_1, $\sum \beta_{2l} x_{1l}$ is an unbiased estimator of $\sqrt{\sigma_{11}}$, and $\sum \beta_{1l}^2$ and $\sum \beta_{2l}^2$ are minimum subject to $\sum \beta_{1l} = 1$, $\sum \beta_{1l} u_l = 0$ and $\sum \beta_{2l} = 0$, $\sum \beta_{2l} u_l = 1$. The relative efficiences of these 'alternative' estimators compared with the best linear estimate can be found by comparing variances calculated according to (3). In the case of singly and doubly censored samples, Sarhan and Greenberg [4], [5] have tabulated these variances and efficiencies for samples up to size $n = 15$, and the worst relative efficiency so obtained was for single censoring with $n = 15$, when for the mean the efficiency was not lower than 84.66% and for the standard deviation not lower than 86.75%. Clearly, the extra efficiency of the optimal estimator is hardly worth the effort of computation, and this may be even more pronounced in the general case.

(b) *Bivariate case.* For a bivariate normal population, there are five parameters requiring estimation, namely μ_1, μ_2, $\sqrt{\sigma_{11}}$ $\sqrt{\sigma_{22}}$ and σ_{12}. The estimation of μ_1 and $\sqrt{\sigma_{11}}$ can be accomplished as in the univariate case, using all the x_{1l} observations available; for type A and C samples only $p(<n)$ such observations can be used, whilst for type B samples where the ordered variate is not censored n observations are available. In the latter case, μ_1 is estimated by the arithmetic mean $n^{-1}\sum x_{1l}$, and $\sqrt{\sigma_{11}}$ can be linearly estimated by either a minimum variance or an alternative estimator. The method of estimating the remaining parameters μ_2, $\sqrt{\sigma_{22}}$, and σ_{12} depends on the type of sample considered.

Type A *or* B *sample.* In a sample where the associated variates are censored, we cannot re-arrange them into increasing order of magnitude because the ranks of the missing observations are not known.

From conditional expectations, or by direct evaluation of the moment generating function of an ordered sample, it may be shown that

$$
\begin{aligned}
E(x_{2(l)}) &= \mu_2 + \sigma_{12}\sigma_{11}^{-\frac{1}{2}}u_l, \\
\text{Var}\,(x_{2(l)}) &= (1 - \sigma_{12}^2\sigma_{11}^{-1}\sigma_{22}^{-1})\sigma_{22} + \sigma_{12}^2\sigma_{11}^{-1}v_{ll}, \\
\text{Cov}\,(x_{2(l)}, x_{2(m)}) &= \sigma_{12}^2\sigma_{11}^{-1}v_{lm}, \qquad l \neq m, \\
\text{Cov}\,(x_{1l}, x_{2(m)}) &= \sigma_{12}v_{lm}.
\end{aligned} \tag{5}
$$

Therefore a linear combination of the available observations has the moments

$$E\{\textstyle\sum \alpha_l x_{2(l)}\} = \mu_2 \sum \alpha_l + \sigma_{12}\sigma_{11}^{-\frac{1}{2}} \sum \alpha_l u_l, \tag{6}$$

$$\text{Var}\,\{\textstyle\sum \alpha_l x_{2(l)}\} = \sigma_{12}^2\sigma_{11}^{-1} \sum\sum \alpha_l\alpha_m v_{lm} + (1 - \sigma_{12}^2\sigma_{11}^{-1}\sigma_{22}^{-1})\sigma_{22}\sum \alpha_l^2. \tag{7}$$

If the α_l are chosen to satisfy $\sum \alpha_l = 1$, $\sum \alpha_l u_l = 0$, then the linear combination is an unbiased estimator for μ_2, and alternatively, if $\sum \alpha_l = 0$, $\sum \alpha_l u_l = 1$, then the combination is an unbiased estimator for $\sigma_{12}\sigma_{11}^{-\frac{1}{2}}$. This latter quantity becomes $\rho_{12}\sqrt{\sigma_{22}}$ on the introduction of the correlation coefficient $\rho_{12} = \sigma_{12}\sigma_{11}^{-\frac{1}{2}}\sigma_{22}^{-\frac{1}{2}}$. Once more, additional restrictions may be made to determine the coefficients. The obvious criterion would be to minimise the variance in (7), which may be rewritten

$$\text{Var}\,(\textstyle\sum \alpha_l x_{2(l)}) = \sigma_{22}\{\rho_{12}^2 \sum\sum \alpha_l\alpha_m v_{lm} + (1 - \rho_{12}^2)\sum \alpha_l^2\}. \tag{8}$$

However, the resulting coefficients would be functions of ρ_{12}^2; for example, when $\rho_{12}^2 = 1$, the variance in (8) reduces to that of (3), so that $\sum \alpha_{1l} x_{2(l)}$ and $\sum \alpha_{2l} x_{2(l)}$

are the best linear unbiased estimators for μ_2 and $\rho_{12}\sqrt{\sigma_{22}}$. By contrast, when $\rho_{12} = 0$ the minimum variance estimate of μ_2 will be simply the arithmetic mean of the observed values, and the standard deviation will be best estimated by treating the missing values as if at randon, ordering the remaining x_2 variables, and applying the univariate theory for a complete sample of size p.

But in general the correlation will not be known and we must therefore relax the restriction of minimum variance, and instead seek estimators which have reasonably small variances for all possible values of the correlation. One such set of estimators is generated by the α_{1l} and α_{2l} considered before, because they are unbiased, and of minimum variance when $\rho_{12}^2 = 1$. The variances of these possible estimators are

$$(9)\quad \begin{aligned} \operatorname{Var}\{\textstyle\sum\alpha_{1l}x_{2(l)}\} &= \sigma_{22}\{\rho_{12}^2\textstyle\sum\sum\alpha_{1l}\alpha_{1m}v_{lm} + (1-\rho_{12}^2)\sum\alpha_{1l}^2\}, \\ \operatorname{Var}\{\textstyle\sum\alpha_{2l}x_{2(l)}\} &= \sigma_{22}\{\rho_{12}^2\textstyle\sum\sum\alpha_{2l}\alpha_{2m}v_{lm} + (1-\rho_{12}^2)\sum\alpha_{2l}^2\}. \end{aligned}$$

The alternative estimates based on the coefficients (4) will likewise be unbiased for μ_2 and $\rho_{12}\sqrt{\sigma_{22}}$, and will have variances

$$(10)\quad \begin{aligned} \operatorname{Var}\{\textstyle\sum\beta_{1l}x_{2(l)}\} &= \sigma_{22}\{\rho_{12}^2\textstyle\sum\sum\beta_{1l}\beta_{1m}v_{lm} + (1-\rho_{12}^2)\sum\beta_{1l}^2\} \\ \operatorname{Var}\{\textstyle\sum\beta_{2l}x_{2(l)}\} &= \sigma_{22}\{\rho_{12}^2\textstyle\sum\sum\beta_{2l}\beta_{2m}v_{lm} + (1-\rho_{12}^2)\sum\beta_{2l}^2\} \end{aligned}$$

Comparing (9) with (10), the relative efficiency of $\sum\beta_{1l}x_{2(l)}$ to $\sum\alpha_{1l}x_{2(l)}$ as an estimate of μ_2 is

$$E = \frac{\rho_{12}^2\sum\sum\alpha_{1l}\,\alpha_{1m}\,v_{lm} + (1-\rho_{12}^2)\sum\alpha_{1l}^2}{\rho_{12}^2\sum\sum\beta_{1l}\,\beta_{1m}\,v_{lm} + (1-\rho_{12}^2)\sum\beta_{1l}^2},$$

and a similar expression holds for the estimates of $\rho_{12}\sqrt{\sigma_{22}}$. The minimum and maximum values of E are given when $\rho_{12}^2 = 1$ and $\rho_{12} = 0$ respectively, and E then takes the values

$$(11)\quad E_{\min} = \frac{\sum\sum\alpha_{1l}\,\alpha_{1m}\,v_{lm}}{\sum\sum\beta_{1l}\,\beta_{1m}\,v_{lm}}, \qquad E_{\max} = \frac{\sum\alpha_{1l}^2}{\sum\beta_{1l}^2}.$$

For equal efficiency ($E = 1$), ρ_{12}^2 has the value

$$(12)\quad \rho_{12}^2 = \left\{1 + \frac{1-E_{\min}}{E_{\max}-1}\cdot\frac{E_{\max}}{E_{\min}}\cdot\frac{\sum\sum\alpha_{1l}\,\alpha_{1m}\,v_{lm}}{\sum\alpha_{1l}^2}\right\}^{-1}.$$

Of course, $E_{\min}$ is also the relative efficiency of $\sum\beta_{1l}x_{1l}$ compared with $\sum\alpha_{1l}x_{1l}$ as an estimate of μ_1, and as such has been tabulated for doubly censored samples by Sarhan and Greenberg [4], [5] for values of n up to 15. Using their tabulations and also some further calculations, we have found the values of $E_{\min}$, $E_{\max}$ and $|\rho_{12}|$ for equal efficiency in the case of doubly censored samples of size $n = 10$, where r_1 observations are missing from the left of the sample, r_2 from the right, leaving $p = n - r_1 - r_2$ central values of x_2 observable. Table 1 shows these

TABLE 1

Relative efficiency of $\sum \beta_{1l} x_{2(l)}$ *compared with* $\sum \alpha_{1l} x_{2(l)}$ *as estimates of* μ_2, *showing minimum and maximum values, and the correlation required for equal efficiency. Doubly censored bivariate normal sample of type* A *or* B, *size* $n = 10$

r_1		r_2 = 0	1	2	3	4	5	6	7	8
0	E_{min}	100.00	99.43	98.06	96.03	93.54	91.13	89.83	91.50	100.00
	E_{max}	100.00	105.09	119.22	136.17	148.27	150.79	142.58	124.64	100.00
	$\lvert\rho_{12}\rvert$	—	.9541	.9655	.9712	.9744	.9762	.9768	.9763	—
1	E_{min}		99.04	98.29	97.28	96.29	95.89	97.06	100.00	
	E_{max}		108.56	118.93	130.77	134.34	126.13	110.76	100.00	
	$\lvert\rho_{12}\rvert$		.9559	.9689	.9773	.9823	.9852	.9862	—	
2	E_{min}			98.20	97.95	97.85	98.57	100.00		
	E_{max}			123.03	128.46	124.13	108.84	100.00		
	$\lvert\rho_{12}\rvert$			.9745	.9814	.9867	.9896	—		
3	E_{min}				98.43	99.03	100.00			
	E_{max}				127.32	115.24	100.00			
	$\lvert\rho_{12}\rvert$				.9861	.9908	—			
4	E_{min}					100.00				
	E_{max}					100.00				
	$\lvert\rho_{12}\rvert$					—				

Note: The non-entries (—) in the above table correspond to the fact that $E = 1$ for all values of ρ_{12}.

quantities for estimates of μ_2 whilst Table 2 gives the similar quantities for estimates of $\rho_{12}\sqrt{\sigma_{22}}$. Clearly, unless $|\rho_{12}|$ is very near unity, the alternative estimators are more efficient than the original, and in any case are never substantially less efficient.

To investigate the absolute efficiencies of μ_2 against all linear alternatives, we see from (9) and (10) that the estimators $\sum \alpha_{1l} x_{2(l)}$ and $\sum \beta_{1l} x_{2(l)}$, will be least efficient when $\rho_{12} = 0$, and in this case the best estimator is the arithmetic mean with variance $p^{-1}\sigma_{22}$. Thus the least efficiency that the original estimator $\sum \alpha_{1l} x_{2(l)}$ can have is $p^{-1}(\sum \alpha_{1l}^2)^{-1}$, and for the alternative estimator, $\sum \beta_{1l} x_{2(l)}$ the least efficiency is $p^{-1}(\sum \beta_{1l}^2)^{-1}$. In Table 3, these quantities are tabulated (as percentages) for the case $n = 10$ and all possible doubly censored samples.

Clearly, neither estimator is satisfactory when $\rho_{12} = 0$ unless the sample is almost complete (r_1 and r_2 small) or unless it is nearly symmetrically censored ($r_1 \doteqdot r_2$). But without knowledge of ρ_{12} no simple method of improving the estimators seems possible without going into the more complicated estimates deduced by maximum likelihood.

TABLE 2

Relative efficiency of $\sum \beta_{2l} x_{2(l)}$ compared with $\sum \alpha_{2l} x_{2(l)}$ as estimates of $\rho_{12}\sqrt{\sigma_{22}}$, showing minimum and maximum values, and the correlation required for equal efficiency. Doubly censored bivariate normal sample of type A *or* B, *size* $n = 10$

r_1		r_2 = 0	1	2	3	4	5	6	7	8
	E_{min}	99.87	96.92	94.07	92.03	90.72	90.17	90.66	92.97	100.00
0	E_{max}	100.39	112.41	127.78	139.47	145.40	144.69	136.81	121.53	100.00
	$\|\rho_{12}\|$	.9319	.9569	.9672	.9724	.9752	.9767	.9771	.9763	—
	E_{min}		97.08	96.11	95.64	95.80	96.65	98.22	100.00	
1	E_{max}		113.45	119.22	122.02	120.27	114.02	104.78	100.00	
	$\|\rho_{12}\|$		.9719	.9790	.9834	.9862	.9877	.9863	—	
	E_{min}			96.32	96.88	98.02	99.56	100.00		
2	E_{max}			118.00	114.99	108.91	101.18	100.00		
	$\|\rho_{12}\|$			.9843	.9881	.9909	.9896	—		
	E_{min}				96.16	99.96	100.00			
3	E_{max}				103.11	100.11	100.00			
	$\|\rho_{12}\|$				.9544	.9913	—			
	E_{min}					100.00				
4	E_{max}					100.00				
	$\|\rho_{12}\|$					—				

Note: The non-entries (—) in the above table correspond to the fact that $E = 1$ for all values of ρ_{12}.

TABLE 3

The Minimum Efficiencies of Original and Alternative Estimators of μ_2 Against all Linear Estimators. Doubly censored bivariate normal samples of type A *or* B, *size* $n = 10$, $\rho_{12} = 0$

r_1	r_2 = 0	1	2	3	4	5	6	7	8
0	100.00	90.64	68.99	47.38	31.34	20.57	13.39	8.35	4.28
	100.00	95.26	82.25	64.52	46.47	31.01	19.09	10.41	4.28
1		92.11	78.55	56.41	36.08	21.32	11.26	4.16	
		100.00	93.42	73.77	48.47	26.89	12.47	4.16	
2			81.28	68.83	44.93	22.30	6.87		
			100.00	88.41	55.77	24.27	6.87		
3				78.54	62.93	20.50			
				100.00	72.52	20.50			
4					100.00				
					100.00				

It is interesting to consider the product of the estimators for $\sqrt{\sigma_{11}}$ and $\sigma_{12}\sigma_{11}^{-\frac{1}{2}}$, namely

$$\sum \beta_{2l}x_{1l} \cdot \sum \beta_{2m}x_{2(m)} = \sum \sum \beta_{2l}\beta_{2m}x_{1l}x_{2(m)} .$$

From (5) we have

$$E\{\sum \sum \beta_{2l}\beta_{2m}x_{1l}x_{2(m)}\} = \sum \sum \beta_{2l}\beta_{2m} \{\sigma_{12}v_{lm} + (\mu_1 + u_l\sqrt{\sigma_{11}})(\mu_2 + u_m\sigma_{12}\sigma_{11}^{-\frac{1}{2}})\}. \tag{13}$$

But also $\sum \beta_{2l} = 0$, $\sum \beta_{2l}u_l = 1$, so that on summation (13) becomes

$$E\{\sum \sum \beta_{2l}\beta_{2m}x_{1l}x_{2(m)}\} = \sigma_{12}(1 + \sum \sum \beta_{2l}\beta_{2m}v_{lm}).$$

We have thus found an explicit unbiased estimator for σ_{12}, namely

$$\sum \beta_{2l}x_{1l} \cdot \sum \beta_{2m}x_{2(m)} \cdot (1 + \sum \sum \beta_{2l}\beta_{2m}v_{lm})^{-1},$$

but, of course, this is not strictly a linear estimator. The coefficients β_{2l} may be replaced by the original ones α_{2l}.

In summary, the above theory allows us to find unbiased estimators for μ_1, μ_2, $\sqrt{\sigma_{11}}$, $\sigma_{12}\sigma_{11}^{-\frac{1}{2}} = \rho_{12}\sqrt{\sigma_{22}}$, and σ_{12}. It should be noted that the quantities

$$\sum \sum \alpha_{1l}\alpha_{1m}v_{lm} , \sum \alpha_{2l}\alpha_{2m}v_{lm} , \sum \beta_{1l}\beta_{1m}v_{lm} , \sum \beta_{2l}\beta_{2m}v_{lm}$$

which occur in many of the equations are tabulated by Sarhan and Greenberg [4], [5] for some doubly censored samples of size $n \leqq 15$, and this facilitates the calculation of the variances of the estimators.

Type C *samples.* A bivariate type C sample will have p observed variables x_1, and n associated variables x_2. There are two distinct cases possible here, because there are $n - p$ x_2-variables associated with missing x_1 values and it might or might not be known what rank these have in the ordered sample. An example of the latter case was given in §3, example C where ranks cannot be assigned to the specimens which have not as yet failed. By slightly changing example A of §3 we can illustrate the former case, for, supposing that as well as measuring the heights and mean diameters of the 5 tallest trees we also know the ordering (but not the exact value) of the heights of the remaining 15 trees and their associated exact diameters, then clearly the positions of the associated variables in the ordered sample are known. The estimation of the parameters $\rho_{12}\sqrt{\sigma_{22}}$ or σ_{12} will depend on which case is available.

Consider first the estimation of μ_2 and $\sqrt{\sigma_{22}}$. Because all variables $x_{2(l)}$ are known, we may re-order them into increasing order of magnitude, say

$$x_{21} < x_{22} < \cdots < x_{2n}$$

where we have now dropped the bracket from the second subscript to indicate strict ordering. Obviously the best estimator for μ_2 is the arithmetic mean $n^{-1}\sum x_{2l}$, but there are two possibilities $\sum \alpha_{2l}x_{2l}$ and $\sum \beta_{2l}x_{2l}$ for estimating $\sqrt{\sigma_{22}}$, the former being of minimum variance, whilst the latter has easily calculated coefficients. Note that we could not estimate $\sqrt{\sigma_{22}}$ for either type A or B samples.

Coming to the problem of estimating $\rho_{12}\sqrt{\sigma_{22}}$, if we do know the position of all the variables $x_{2(l)}$ in the ordered sample then we can proceed as before and estimate $\rho_{12}\sqrt{\sigma_{22}}$ by $\sum \alpha_{2l} x_{2(l)}$ or $\sum \beta_{2l} x_{2(l)}$, and estimate σ_{12} by $\sum \beta_{2m} x_{1m} \cdot \sum \beta_{2l} x_{2(l)} \cdot (1 + \sum\sum \beta_{2m} \beta_{2l} v_{lm})^{-1}$ or with α_{2l} instead of β_{2l}. Here, the summation over m is for the p uncensored x_1 values. As we can also estimate $\sqrt{\sigma_{22}}$ for this type of sample, we can estimate ρ_{12} directly by

$$\sum \alpha_{2l} x_{2(l)} \{ \sum \alpha_{2l} x_{2l} \}^{-1} \quad \text{or} \quad \sum \beta_{2l} x_{2(l)} \{ \sum \beta_{2l} x_{2l} \}^{-1},$$

but these estimators are biased. On the other hand, if not all positions for the x_2 variables can be assigned in the ordered sample, we can proceed by disregarding those of doubtful rank and treat the sample as if it were of type A. Thus $\sum' \alpha'_{2l} x_{2(l)}$ and $\sum' \beta'_{2l} x_{2(l)}$ are unbiased estimators for $\rho_{12}\sqrt{\sigma_{22}}$, and

$$\frac{\sum' \alpha'_{2l}\, x_{2l} \sum \alpha_{2m}\, x_{1m}}{1 + \sum' \sum \alpha'_{2l}\, \alpha_{2m}\, v_{lm}} \quad \text{and} \quad \frac{\sum' \beta'_{2l}\, x_{2l} \sum \beta_{2m}\, x_{1m}}{1 + \sum' \sum \beta'_{2l}\, \beta_{2m}\, v_{lm}}$$

are unbiased estimators for σ_{12}. Here, the dash on the summation and the coefficients indicates that variates of unknown rank are disregarded. Also,

$$\sum' \alpha'_{2l} x_{2(l)} \{ \sum \alpha_{2l} x_{2l} \}^{-1} \quad \text{and} \quad \sum' \beta'_{2l} x_{2(l)} \{ \sum \beta_{2l} x_{2l} \}^{-1}$$

will be (biased) estimates of ρ_{12}.

(c) *Multivariate case.* For a multivariate sample of type A or B, the theory deduced for bivariate samples may be applied to each pair of variates x_1, x_i, and most of the parameters may be estimated by linear estimators or their combinations. In addition, for type C samples the bivariate theory may be applied to all variate pairs x_i, x_j; when two associated variates are considered they form a complete bivariate sample and this may be ordered with respect to each variate in turn, thus providing estimates for all parameters of the population.

It is clear from the efficiencies given in Tables 1 and 2 that for multivariate (as well as bivariate) populations, the alternative estimators based on the coefficients β_{1l}, β_{2l} defined in (4) are simply calculated, are generally more efficient than the original ones except for estimating the parameters μ_1 and $\sqrt{\sigma_{11}}$, and have a high absolute efficiency (compared with maximum likelihood estimators) when not many sample elements are censored (see Table 3 and Sarhan and Greenberg [4], [5]). Therefore they can be recommended as a satisfactory solution to the problem of estimation from a multivariate normal censored sample.

5. Example. We illustrate the above methods of estimation applied to a type C censored sample drawn from a bivariate normal population with parameters

$$\mu_1 = \mu_2 = 0, \qquad \sigma_{11} = \sigma_{22} = 1, \qquad \rho_{12} = \sigma_{12} = 0.6.$$

A sample of size $n = 10$ was drawn from the tables of 'Correlated Random Normal Deviates' of Fieller, Lewis and Pearson [2], and when ordered with respect to x_1 is

l	1	2	3	4	5	6	7	8	9	10
x_{1l}	−0.87*	−0.86	−0.73	−0.15	0.39	0.41	0.48	0.64	1.20*	2.13*
$x_{2(l)}$	−0.16	−1.48	0.60	0.30	1.40	−0.49	2.40	0.65	2.03	1.01.

The starred variables will be assumed missing for the purposes of the example, and thus a doubly censored sample of type C results. We assume that we know the ranking of the associated variables −0.16, 2.03, 1.01. The ordered values of x_2 are

l	1	2	3	4	5	6	7	8	9	10
x_{2l}	−1.48	−0.49	−0.16	0.30	0.60	0.65	1.01	1.40	2.03	2.40.

In Table 4 we show the original and alternative estimates for the various parameters, and in the case of strictly linear estimators their variances calculated according to (3) and (7).

TABLE 4

Parameter Estimates for a Type C, *Double Censored Bivariate Sample*

Parameter	Original Estimates		Alternative Estimates	
	Estimator	Variance (No. of Obs.)	Estimator	Variance (No. of Obs.)
$\mu_1 = 0$	$\sum \alpha_{1l} x_{1l} = 0.1298$	0.1085 (7)	$\sum \beta_{1l} x_{1l} = 0.1682$	0.1103 (7)
$\mu_2 = 0$	$n^{-1} \sum x_{2l} = 0.6260$	0.1000 (10)	$n^{-1} \sum x_{2l} = 0.6260$	0.1000 (10)
$\sqrt{\sigma_{11}} = 1$	$\sum \alpha_{2l} x_{1l} = 0.9263$	0.1014 (7)	$\sum \beta_{2l} x_{1l} = 0.9961$	0.1055 (7)
$\sqrt{\sigma_{22}} = 1$	$\sum \alpha_{2l} x_{2l} = 1.2391$	0.0576 (10)	$\sum \beta_{2l} x_{2l} = 1.2369$	0.0577 (10)
$\rho_{12}\sqrt{\sigma_{22}} = 0.6$	$\sum \alpha_{2l} x_{2(l)} = 0.7189$	0.1019 (10)	$\sum \beta_{2l} x_{2(l)} = 0.7461$	0.1016 (10)
$\sigma_{12} = 0.6$	$\dfrac{\sum \alpha_{2l} x_{2(l)} \sum \alpha_{2m} x_{1m}}{1 + \sum\sum \alpha_{2l} \alpha_{2m} v_{lm}} = 0.6297$		$\dfrac{\sum \beta_{2l} x_{2(l)} \sum \beta_{2m} x_{1m}}{1 + \sum\sum \beta_{2l} \beta_{2m} v_{lm}} = 0.7019$	
$\rho_{12} = 0.6$	$\dfrac{\sum \alpha_{2l} x_{2(l)}}{\sum \alpha_{2l} x_{2l}} = 0.5802$		$\dfrac{\sum \beta_{2l} x_{2(l)}}{\sum \beta_{2l} x_{2l}} = 0.6032$	

As is expected, the original and alternative estimates are similar both in values and in variances, but the sample seems rather extreme with respect to deviations of x_2 from its mean zero.

Acknowledgement. I wish to thank Professor H. A. David for suggesting this problem and for much helpful discussion and criticism. I am also grateful to the referee for suggesting the generalization from doubly censored samples to samples with interior variables missing, and for other helpful remarks.

REFERENCES

[1] A. C. Cohen, Jr. "Restriction and selection in multinormal distributions," *Ann. Math. Stat.*, Vol. 28 (1957), pp. 731–741.

[2] E. C. Fieller, T. Lewis, and E. S. Pearson, "Correlated random normal deviates," *Tracts for Computers, No. XXVI*, Cambridge University Press, Cambridge, 1955.
[3] A. K. Gupta, "Estimation of the mean and standard deviation of a normal population from a censored sample." *Biometrika*, Vol. 39 (1952), pp. 260–273.
[4] A. E. Sarhan and B. G. Greenberg, "Estimation of location and scale parameters by order statistics from singly and doubly censored samples Pt. I," *Ann. Math. Stat.*, Vol. 27 (1956), pp. 427–451.
[5] A. E. Sarhan and B. G. Greenberg, "Estimation of location and scale parameters by order statistics from singly and doubly censored samples Pt. II," *Ann. Math. Stat.*, Vol. 29 (1958), pp. 79–105.
[6] A. E. Sarhan and B. G. Greenberg, "Tables for best estimates by order statistics of the parameters of single exponential distributions from singly and doubly censored samples," *J. Amer. Stat. Assn.*, Vol. 52 (1957), pp. 58–87.
[7] D. Teichroew, "Tables of expected values of order statistics and products of order statistics for samples of size twenty and less from the normal distribution," *Ann. Math. Stat.*, Vol. 27 (1956), pp. 410–426.
[8] G. A. Watterson, "Ordered and censored samples from a multivariate normal population," Master's thesis, University of Melbourne, unpublished, 1958.