A further note on multivariate selection

By A. C. Aitken, University of Edinburgh.

(*Received 24th February*, 1936. *Read 6th March*, 1936).

In an earlier "Note on Selection from a Normal Multivariate Population" the author considered[1] the transformation of an n-variate normal population induced by "selection" operating to alter the parameters of frequency in a subset of p variates, while preserving normal frequency in the subset. The results were new in notation only; in substance they were first found[2] by K. Pearson.

Pearson observed, in a footnote to a later paper (*Biom.* 6 (1909), 111) that the formulae as developed in his first memoir on the subject "are true far beyond the range of the Gaussian distribution for which they were proved. They are universally true provided we start with a generalised notion of correlation as involving the maximum dependence of one variable on an arbitrary linear function of $n - 1$ other variables." A subsequent paper (*Biom.* 8 (1911-12), 437-443) "On the General Theory of the Influence of Selection on Correlation and Variation" provided mathematical justification for this remark.

This identity of the results under the much wider formulation is of great interest. It recalls, and may suitably be compared with, the related situation in Least Squares, where the same normal equations for a set of linear unknowns can be derived from the quite different postulates (i) of maximum probability of the observations under a normal law, or (ii) of minimum variance of consistent linear estimate under any law.

Our present purpose, however, is to show that it is not necessary to extend the notion of correlation in order that this particular transformation under selection may hold; that, in fact, the transformation of the vector of means and of the matrix of variance (or of the binary product moment quadric) holds for any multivariate population, when the subset of p variates is selected to conform to any distribution, not necessarily of the same type as that before selection.

---

[1] *Proc. Edin. Math. Soc.* (2) 4 (1934), 106-110.

[2] *Phil. Trans. Roy. Soc. London*, 200A (1902), 1-66.

General selection in means, variances and bivariances.

A shortened notation of vectors and matrices is used. The population is characterised by the differential of frequency

$$dp = \phi(x, y)\, dx\, dy, \tag{1}$$

where x is used to distinguish the p selected, y the $n-p$ unselected variates, $\phi(x, y)$ denotes the frequency function

$$\phi(x_1, x_2, \ldots, x_p, y_{p+1}, \ldots, y_p),$$

dx and dy indicate products of differentials dx_i and dy_j, and all variates are measured with respect to their means as origin.

The moment generating function (exponential transform) of the distribution is

$$\begin{aligned} G(\alpha, \beta) &= \int \phi(x, y) \exp(\alpha' x + \beta' y)\, dx\, dy \\ &= 1 + [\alpha' : \beta']\, V\, \{\alpha : \beta\}/2! + \ldots, \end{aligned} \tag{2}$$

where α', β' are row vectors of p and $n-p$ elements, x and y are corresponding column vectors, and V is the matrix of variance, having the n variances for its diagonal elements and the $\frac{1}{2}n(n-1)$ bivariances[1] for its non-diagonal elements. It is assumed that the integral is convergent within a certain range of values of each of the α and β.

Let the differential of frequency in the x alone, variation in the y being ignored, be $\phi(x)\, dx$. It may be obtained by integrating the differential in (1) over the whole range of the y. The effect of selection will be to substitute for $\phi(x)\, dx$ another differential $\psi(x-h)\, dx$, with a new vector of means h and a new variance matrix U_{pp} in the selected variates. To prepare for this substitution we partition the original matrix V according to selected and unselected variates, and segregate all terms in α in the variance quadric of (2) into a separate quadratic form, thus:

$$\begin{aligned} [\alpha' : \beta'] \left[\begin{array}{c:c} V_{pp} & V_{p,n-p} \\ \hdashline V_{n-p,p} & V_{n-p,n-p} \end{array}\right] \left[\begin{array}{c} \alpha \\ \beta \end{array}\right] &= \alpha'\, V_{pp}\, \alpha + 2\alpha'\, V_{p,n-p}\, \beta + \beta'\, V_{n-p,n-p}\, \beta \\ = (\alpha + V_{pp}^{-1} V_{p,n-p}\beta)'\, V_{pp}\, (\alpha + V_{pp}^{-1}\, V_{p,n-p}\beta) &+ \beta'(V_{pp} - V_{n-p,p}\, V_{pp}^{-1} V_{p,n-p})\, \beta. \end{aligned} \tag{3}$$

[1] I have ventured to introduce this neologism in order to avoid the confusion resulting from the use of the same word with different meanings in other branches of mathematics. "Covariance" is fully appropriated elsewhere in another sense.

https://doi.org/10.1017/S0013091500008245 Published online by Cambridge University Press

We next express $G(\alpha, \beta)$ as an integral taken over elements of frequency in x, thus:

$$G(\alpha, \beta) = F(\alpha, \beta) \int \phi(x) \exp\{(\alpha + V_{pp}^{-1} V_{p, n-p} \beta)' x\} dx \qquad (4)$$

where in view of (3) the partial generating function $F(\alpha, \beta)$ does not contain α in its terms of first and second degree.

Substituting $\psi(x - h)\, dx$ for $\phi(x)\, dx$ in (4) and integrating again, we obtain the moment generating function after selection in the shape

$$\exp\{(\alpha + V_{pp}^{-1} V_{p, n-p} \beta) h\}.\ [1 + (\alpha + V_{pp}^{-1} V_{p, n-p} \beta)'\, U_{pp} (\alpha + V_{pp}^{-1} V_{p, n-p} \beta) + \beta' (V_{pp} - V_{n-p, p}\, V_{pp}^{-1}\, V_{p, n-p}) \beta + \ldots .]. \qquad (5)$$

The vector of means in the modified population is therefore

$$\{h:\ V_{pp}^{-1}\, V_{p, n-p}\, h\}, \qquad (6)$$

which shows that the means of the variates y are given by the $n - p$ elements of the vector $V_{n-p, p}\, V_{pp}^{-1} h$. By expressing the variance quadric in (5) as the sum of a quadratic form in α, a quadratic form in β and twice a bilinear form in α and β, we obtain also the transformation induced by selection on the original matrix V:

$$\begin{bmatrix} V_{pp} & V_{p, n-p} \\ V_{n-p, p} & V_{n-p, n-p} \end{bmatrix}$$

$$\to \begin{bmatrix} U_{pp} & U_{pp}\, V_{pp}^{-1}\, V_{p, n-p} \\ V_{n-p, p}\, V_{pp}^{-1}\, U_{pp} & V_{n-p, n-p} - V_{n-p, p} (V_{pp}^{-1} - V_{pp}^{-1}\, U_{pp}\, V_{pp}^{-1})\, V_{p, n-p} \end{bmatrix}. \qquad (7)$$

The results (6) and (7) are those obtained under the much more restricted hypotheses of the earlier note.

The form of (7) will reveal, on examination, a reflexive property which is to be expected, namely that a second selection operating on the already selected population to restore U_{pp} to V_{pp} would at the same time restore the *whole* original matrix V. This reciprocity of formulae between the parent population and the selected one is a well-known fact for two and three varieties. More generally, the operations $V_{pp} \to U_{pp}$, $U_{pp} \to W_{pp}$, performed in succession according to (7), may be verified as equivalent to the single operation $V_{pp} \to W_{pp}$; in other words the operation (7) is transitive.

Selection in a bivariate Poisson population.

The correlation function of two variates which are separately distributed according to Poisson laws has been considered from quite

https://doi.org/10.1017/S0013091500008245 Published online by Cambridge University Press

different standpoints by Wicksell[1], McKendrick[2] and Campbell[3]. The effect on such a distribution of selection in one variate could be ascertained by applying the general result of the preceding section; but it is instructive to use the factorial moment generating function (binomial transform) and to segregate α.

The factorial moment generating function is

$$\begin{aligned} G(\alpha, \beta) &= \exp(m_{10}\alpha + m_{01}\beta + m_{11}\alpha\beta) \\ &= \exp(m_{00}\beta)\exp\{\alpha(m_{10} + m_{11}\beta)\} \\ &= \exp(m_{01}\beta)\sum_{x=0}^{\infty} e^{-m_{10}}\frac{m_{10}^x}{x!}(1+\alpha)^x e^{-m_{11}\beta}\left(1+\frac{m_{11}}{m_{10}}\beta\right)^x. \end{aligned}$$

The generating function has thus been expressed as a sum over elements of frequency in x only. Let selection in x now operate to produce Poisson frequency with a different mean m'_{10}.

The new generating function is then

$$\begin{aligned} &\exp(m_{01}\beta)\sum_{0}^{\infty} e^{-m'_{10}}\frac{m'^x_{10}}{x!}(1+\alpha)^x e^{-m_{11}\beta}\left(1+\frac{m_{11}}{m_{10}}\beta\right)^x \\ &= \exp(m_{01}\beta - m'_{10} - m_{11}\beta)\exp\left\{m'_{10}(1+\alpha)\left(1+\frac{m_{11}}{m_{10}}\beta\right)\right\} \\ &= \exp\left\{m'_{10}\alpha + \left[m_{01} - m_{11}\left(1-\frac{m'_{10}}{m_{10}}\right)\right]\beta + \frac{m'_{10}}{m_{10}}m_{11}\alpha\beta\right\}. \end{aligned}$$

By inspection, the mean of y in the modified population is seen to be

$$m_{01} - m_{11}(1-s),$$

and the new bivariance is $m_{11}s$, where $s = m'_{10}/m_{10}$.

These formulae are the analogues of those giving the variance of y,

$$1 - \rho^2(1-s^2),$$

and the bivariance, ρs^2, for two normally correlated variables x and y originally in standard measure, when x is selected to have variance s^2.

---

[1] *Svenska Aktuarieföreningens Tidskr.* (1916), 192.

[2] *Proc. Edin. Math. Soc.* (1) **44** (1925), 106.

[3] *Proc. Edin. Math. Soc.* (2) **4** (1934), 18-26.

https://doi.org/10.1017/S0013091500008245 Published online by Cambridge University Press