INTERRELATIONS BETWEEN CERTAIN LINEAR SYSTEMATIC STATISTICS OF SAMPLES FROM ANY CONTINUOUS POPULATION

By G. P. SILLITTO

Research Department, Imperial Chemical Industries Ltd., Nobel Division

1. Statistics which consist of linear combinations of the observations in a sample have been called by Mosteller (1946) 'systematic statistics'. In one class of these the law of combination of the observations is always the same for samples of a given size. It is evident, therefore, that linear systematic statistics of this class must be linear combinations of each other. Their expectations will be linearly related, and the variance of one will be expressible in terms of variances and co-variances of others. These relations will hold for samples from all continuous populations. In this paper, expressions are obtained for the expectations of some linear systematic statistics of this class in terms of expectations of others.

It is found that certain linear systematic statistics which measure dispersion, such as Gini's coefficient of mean difference, have expectations which are independent of the size of the sample from which they are calculated, and linear systematic statistics having this property are derived which may be used as indices of skewness and kurtosis for samples from any continuous population. Statistics with this property have advantages, for instance when estimates from samples of different sizes are to be combined.

2. The expectation of the range, $\overline{w}_n$, in a sample of n from any continuous population of variate-values z, which has distribution function F, is (cf. Kendall, 1943, p. 223)

$$\overline{w}_n = \int \{1-(1-F)^n - F^n\}\,dz. \tag{1}$$

Similarly,

$$\overline{w}_{n-1} = \int \{1-(1-F)^{n-1} - F^{n-1}\}\,dz,$$

the integrals being taken over the whole range of z. Therefore

$$\begin{aligned}\overline{w}_n &= \int \{(1-F)^{n-1}F + F^{n-1}(1-F)\}\,dz + \overline{w}_{n-1} \\ &= (\chi_{n,1}+\chi_{n,n-1})/n + \overline{w}_{n-1},\end{aligned} \tag{2}$$

where $\chi_{n,p}$ is the expectation of the difference between the $(p+1)$th and the pth value of z when the sample members are arranged in ascending order of magnitude. It is known (K. Pearson, 1902) that

$$\chi_{n,p} = \binom{n}{p}\int F^{n-p}(1-F)^p\,dz. \tag{3}$$

Obviously $\overline{w}_{n-1}$ in (2) can be expressed in terms of $\chi_{n-1,1}$, $\chi_{n-1,n-2}$ and $\overline{w}_{n-2}$, and so on.

3. From (3)

$$\begin{aligned}\frac{n}{p}\chi_{n-1,p-1} - \frac{n-p+1}{p}\chi_{n,p-1} &= \int \left\{\binom{n}{p}F^{n-p}(1-F)^{p-1} - \binom{n}{p}F^{n-p+1}(1-F)^{p-1}\right\}dz \\ &= \binom{n}{p}\int F^{n-p}(1-F)^p\,dz = \chi_{n,p},\end{aligned} \tag{4}$$