

## CORRESPONDENCE

The Joint Editors,  
*The Journal of the Institute of  
Actuaries Students' Society*

12 December 1949

### *A Proof of Certain Statistical Distributions*

Sirs,

Dr J. Wishart, following an admirable lecture<sup>(1)</sup>, recently gave your readers a simple proof of certain fundamental theorems in mathematical statistics<sup>(2, 3)</sup>.

What I believe to be an even simpler proof of some of the results, based on a method given by Prof. J. F. Steffensen<sup>(4)</sup>, may be stated as follows:

$x_i$ ,  $i=1$  to  $n$ , are  $n$  independent observations from a normal distribution with mean  $\xi$  and variance  $\sigma^2$ , so that the probability that  $x_1 x_2 \dots x_n$  fall within  $dx_1 dx_2 \dots dx_n$  is given by

$$p(x_1 \dots x_n) dx_1 \dots dx_n \\ = (2\pi\sigma^2)^{-\frac{1}{2}n} \exp\left[-\frac{1}{2}\left\{\sum_{i=1}^n (x_i - \xi)^2\right\} / \sigma^2\right] dx_1 \dots dx_n.$$

Note that 
$$\sum_{i=1}^n (x_i - \xi)^2 = \sum_{i=1}^n (x_i - \bar{x})^2 + n(\bar{x} - \xi)^2,$$

where  $\bar{x} = \sum_{i=1}^n x_i/n$ .

Let us make the transformation

$$x_i = z + u_i \sqrt{S} \quad (i=1 \text{ to } n)$$

subject to 
$$\sum_{i=1}^n u_i = 0 \quad \text{and} \quad \sum_{i=1}^n u_i^2 = 1.$$

Summing the  $n$  equations shows that

$$z = \sum_{i=1}^n x_i/n = \bar{x},$$

since  $\sum_{i=1}^n u_i = 0$ , and summing the  $(x_i - \bar{x})^2$  shows that

$$S = \sum_{i=1}^n (x_i - \bar{x})^2,$$

since

$$\sum_{i=1}^n u_i^2 = I.$$

Further, since two equations hold between the  $u_i$ , any two of them, say  $u_{n-1}$  and  $u_n$ , can in theory be expressed in terms of the remaining  $u_i$ .

We are thus transforming from  $n$  functionally independent variables  $x_i$  to  $n$  functionally independent variables  $\bar{x}, S, u_1 \dots u_{n-2}$ . The Jacobian of the transformation is then  $\partial(x_1 \dots x_n) / \partial(\bar{x} S u_1 \dots u_{n-2})$ , which is

I	$\frac{u_1}{2\sqrt{S}}$	$\sqrt{S}$	0	0	:	:	0
I	$\frac{u_2}{2\sqrt{S}}$	0	$\sqrt{S}$	0	:	:	0
:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:
I	$\frac{u_{n-2}}{2\sqrt{S}}$	0	0	0	:	:	$\sqrt{S}$
I	$\frac{u_{n-1}}{2\sqrt{S}}$	$\sqrt{S} \frac{\partial u_{n-1}}{\partial u_1}$	:	:	:	:	$\sqrt{S} \frac{\partial u_{n-1}}{\partial u_{n-2}}$
I	$\frac{u_n}{2\sqrt{S}}$	$\sqrt{S} \frac{\partial u_n}{\partial u_1}$	:	:	:	:	$\sqrt{S} \frac{\partial u_n}{\partial u_{n-2}}$

This formidable looking affair immediately factorizes into

$S^{\frac{1}{2}(n-3)}$	I	$\frac{1}{2}u_1$	I	0	0	:	:	0
	I	$\frac{1}{2}u_2$	0	I	0	:	:	0
	:	:	:	:	:	:	:	:
	:	:	:	:	:	:	:	:
	I	$\frac{1}{2}u_{n-2}$	0	0	0	:	:	I
	I	$\frac{1}{2}u_{n-1}$	$\frac{\partial u_{n-1}}{\partial u_1}$	:	:	:	:	$\frac{\partial u_{n-1}}{\partial u_{n-2}}$
	I	$\frac{1}{2}u_n$	$\frac{\partial u_n}{\partial u_1}$	:	:	:	:	$\frac{\partial u_n}{\partial u_{n-2}}$

and all we need note about this determinant is that it is a function of the  $u_i$  only, say  $f(u_1 \dots u_{n-2})$ , and does not contain  $\bar{x}$  or  $S$ .

Performing the transformation, remembering that  $dx_1 dx_2 \dots dx_n$  transforms into  $|J| dS d\bar{x} du_1 \dots du_{n-2}$ , we find that the probability that  $S, \bar{x}, u_1 \dots u_{n-2}$  fall within  $dS d\bar{x} du_1 \dots du_{n-2}$  is given by

$$p(S, \bar{x}, u_1 \dots u_{n-2}) dS d\bar{x} du_1 \dots du_{n-2} \\ = c S^{\frac{1}{2}(n-3)} \exp\left[-\frac{1}{2}\{S + n(\bar{x} - \xi)^2\}/\sigma^2\right] f(u_1 \dots u_{n-2}) dS d\bar{x} du_1 \dots du_{n-2},$$

where  $c$  is a constant.

This factorizes at sight and shows that

$$p(S) dS = c_1 S^{\frac{1}{2}(n-3)} \exp(-\frac{1}{2}S/\sigma^2) dS,$$

$$p(\bar{x}) d\bar{x} = c_2 \exp\{-\frac{1}{2}n(\bar{x} - \xi)^2/\sigma^2\} d\bar{x},$$

$$p(u_1 \dots u_{n-2}) du_1 \dots du_{n-2} = c_3 f(u_1 u_2 \dots u_{n-2}) du_1 du_2 \dots du_{n-2}.$$

It also shows that  $S, \bar{x}$  and the  $u_i$  taken as a group are stochastically independent. (Among themselves the  $u_i$  are correlated; but this is a fact without interest to us here.)  $c_1$  and  $c_2$  are rapidly obtained from

$$\int_0^\infty p(S) dS = \int_{-\infty}^\infty p(\bar{x}) d\bar{x} = 1.$$

The  $u_i$  we cheerfully ignore since we are not interested in them, and the method has yielded the two distributions that we do want in a delightfully 'sudden' manner. The method can be applied again, with a slight modification, in the first part of Dr Wishart's third paper(3), but I do not wish to use up any more of your valuable space.

Yours faithfully,

J. E. KERRICH

University of the Witwatersrand

#### REFERENCES

- (1) WISHART, J. (1947). The variance ratio test in statistics. *J.S.S.* Vol. VI, p. 172.
- (2) WISHART, J. (1947). Proof of the distribution of  $\chi^2$ , of the estimate of the variance and of the variance ratio. *J.S.S.* Vol. VII, p. 98.
- (3) WISHART, J. (1948). Tests of significance in the simple regression problem. *J.S.S.* Vol. VIII, p. 38.
- (4) STEFFENSEN, J. F. (1936). Free functions and the Student-Fisher theorem. *Skand. AktuarTidskr.* Vol. XIX, p. 108.

[The method used by our correspondent may be compared with R. A. Fisher's 'geometrical' method, of which it appears to be the analytical equivalent. The subject is discussed in Chapter 10 of M. G. Kendall's *Advanced Theory of Statistics*.—Eds. *J.S.S.*]