NOTES

This section is devoted to brief research and expository articles on methodology and other short items.

ON THE STUDENTIZATION OF SEVERAL VARIANCES

By B. L. Welch

University of Leeds, England

1. Introduction. In a recent paper [1] the author considered the problem of eliminating several variances simultaneously from probability statements concerning the mean of a normally distributed variable. The general situation envisaged was as follows. We supposed that we had an observed quantity y which could be assumed to be normally distributed about a population mean η with variance $\sigma_y^2 = \sum_{i=1}^{k} \lambda_i \sigma_i^2$, where the λ_i are known positive numbers and the σ_i^2 unknown population variances. It was supposed further that the data provided estimates s_i^2 of the σ_i^2 based on f_i degrees of freedom, and having the sampling distributions

$$p(s_i^2)\, ds_i^2 = \frac{1}{\Gamma(\frac{1}{2}f_i)} \exp\left\{-\frac{1}{2}\frac{f_i s_i^2}{\sigma_i^2}\right\}\left(\frac{1}{2}\frac{f_i s_i^2}{\sigma_i^2}\right)^{\frac{1}{2}f_i - 1} d\left(\frac{1}{2}\frac{f_i s_i^2}{\sigma_i^2}\right) \tag{1}$$

and that these estimates were distributed independently of each other and of y. The problem was to make statements about the magnitude of the difference $y - \eta$ which would involve explicitly only the observed variances s_i^2. The probability of the truth of the statements was also to be entirely independent of the population values σ_i^2.

The solution was given implicitly in a formal mathematical expression and a general process of developing successive terms in a series expansion was described. In the present communication a slightly different way of reaching this development is provided.

2. General method. If the f_i are large enough the ratio

$$v = \frac{y - \eta}{\sqrt{\Sigma \lambda_i s_i^2}} \tag{2}$$

can be taken to be normally distributed with mean zero and standard deviation unity. This suggests that, when the f_i are not necessarily large, we might approach the matter by seeking some other function

$$x = g\{s_1^2, s_2^2, \cdots, s_k^2, y - \eta\} \tag{3}$$

which will still be normally distributed with the same mean and standard deviation. We shall see that such a function can be found, although the method to be followed leads us first to another expression

$$y - \eta = h(s_1^2, s_2^2, \cdots, s_k^2, x) \tag{4}$$

which is simply the transposed form of (3). Once we have obtained h we **can** solve out from (4) to obtain x.

Since the distribution of y is independent of s_i^2 we have

$$p(y \mid s^2)\, dy = \frac{1}{\sqrt{2\pi\Sigma\lambda_i\sigma_i^2}} \exp\left\{-\tfrac{1}{2}\frac{(y-\eta)^2}{\Sigma\lambda_i\sigma_i^2}\right\} dy. \tag{5}$$

Transforming therefore to the new variable x we have for *given* s_i^2

$$\begin{aligned} p(x \mid s^2)\, dx &= \frac{1}{\sqrt{2\pi\Sigma\lambda_i\sigma_i^2}} \exp\left\{-\tfrac{1}{2}\frac{h^2(s^2, x)}{\Sigma\lambda_i\sigma_i^2}\right\} \frac{\partial h(s^2, x)}{\partial x}\, dx \\ &= j\{s^2, x, \Sigma\lambda_i\sigma_i^2\}\, dx \qquad \text{(say).} \end{aligned} \tag{6}$$

The unrestricted distribution of x is then obtained by averaging over the joint distribution of the s_i^2. In order that x should be a unit normal deviate we must therefore have

$$p(x) = \underset{s^2}{\int \cdots \int} j\{s^2, x, \Sigma\lambda_i\sigma_i^2\} \prod_i \{p(s_i^2)\, ds_i^2\} = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}. \tag{7}$$

We have to substitute from (1) and (6) into (7) and then choose the function $h(s^2, x)$ in such a manner that the equation is satisfied whatever may be the values of the unknown σ_i^2. To evaluate the function by the methods of numerical integration is probably impracticable except perhaps in some simple special cases. A series development is, however, quite feasible.

Symbolically we can write

$$j\,\{s^2, x, (\Sigma\lambda_i\sigma_i^2)\} = e^{\Sigma(s_i^2 - \sigma_i^2)\partial_i} j\{w, x, \Sigma\lambda_i\sigma_i^2\} \tag{8}$$

where ∂_i denotes differentiation with respect to w_i and subsequent equation to σ_i^2. Equation (7) then integrates out to give

$$\prod_i e^{-\sigma_i^2\partial_i}\left\{1 - \frac{2\sigma_i^2\partial_i}{f_i}\right\}^{-\frac{1}{2}f_i} j\{w, x, \Sigma\lambda_i\sigma_i^2\} = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \tag{9}$$

i.e.

$$\Theta j\{w, x, \Sigma\lambda_i\sigma_i^2\} = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \qquad \text{(say).} \tag{10}$$

The operator Θ must be expanded in powers of ∂_i before it can be interpreted When this is done we find

$$\Theta = \exp\left\{\Sigma\frac{\sigma_i^4\partial_i^2}{f_i} + \tfrac{4}{3}\Sigma\frac{\sigma_i^6\partial_i^3}{f_i^2} + 2\Sigma\frac{\sigma_i^8\partial_i^4}{f_i^3} + \cdots\right\} \tag{11}$$

$$= 1 + \Sigma\frac{\sigma_i^4\partial_i^2}{f_i} + \left\{\tfrac{4}{3}\Sigma\frac{\sigma_i^6\partial_i^3}{f_i^2} + \tfrac{1}{2}\left(\Sigma\frac{\sigma_i^4\partial_i^2}{f_i}\right)^2\right\} + \cdots. \tag{12}$$

Our procedure now is to find successive approximations to $h(s^2, x)$. It will be convenient to denote by $h_r(s^2, x)$ an expression which equals $h(s^2, x)$ to terms of order $1/f_i^r$. Further let $c_{r+1}(s^2, x)$ be a corrective term which when added on to $h_r(s^2, x)$ will give a result correct to terms in $1/f_i^{r+1}$. Then to this order we shall have from (6)

$$(13)\quad \begin{aligned}\sqrt{2\pi}\, j\{w, x, \Sigma\lambda_i\sigma_i^2\} &= \frac{1}{\sqrt{\Sigma\lambda_i\sigma_i^2}} \exp\left\{-\tfrac{1}{2}\frac{h_r^2(w, x)}{\Sigma\lambda_i\sigma_i^2}\right\}\frac{\partial h_r(w, x)}{\partial x} \\ &\quad + \frac{1}{\sqrt{\Sigma\lambda_i\sigma_i}} \exp\left\{-\tfrac{1}{2}\frac{x^2(\Sigma\lambda_i w_i)}{\Sigma\lambda_i\sigma_i^2}\right\}\left\{\frac{\partial c_{r+1}(w, x)}{\partial x} - \frac{x(\Sigma\lambda_i w_i)c_{r+1}(w, x)}{\Sigma\lambda_i\sigma_i^2}\right\}\end{aligned}$$

remembering that the leading term in $h(w, x)$ is $x\sqrt{\Sigma\lambda_i w_i}$.

Hence from (10) we find

$$(14)\quad \begin{aligned}&\Theta\frac{1}{\sqrt{\Sigma\lambda_i\sigma_i^2}} \exp\left\{-\tfrac{1}{2}\frac{h_r^2(w, x)}{\Sigma\lambda_i\sigma_i^2}\right\}\frac{\partial h_r(w, x)}{\partial x} \\ &\qquad + \frac{1}{\sqrt{\Sigma\lambda_i\sigma_i^2}}\, e^{-\frac{1}{2}x^2}\left\{\frac{\partial c_{r+1}(\sigma^2, x)}{\partial x} - x c_{r+1}(\sigma^2, x)\right\} = e^{-\frac{1}{2}x^2}\end{aligned}$$

i.e.

$$(15)\quad \frac{\partial}{\partial x}\left\{e^{-\frac{1}{2}x^2}\frac{c_{r+1}(\sigma^2, x)}{\sqrt{\Sigma\lambda_i\sigma_i^2}}\right\} + \Theta \exp\left\{-\tfrac{1}{2}\frac{h_r^2(w, x)}{\Sigma\lambda_i\sigma_i^2}\right\}\frac{1}{\sqrt{\Sigma\lambda_i\sigma_i^2}}\frac{\partial h_r(w, x)}{\partial x} = e^{-\frac{1}{2}x^2}.$$

Given h_r we can therefore proceed directly to c_{r+1} and hence to h_{r+1}.

3. Application to give terms in $1/f_i$. It will be sufficient illustration of the method, if we show here how to obtain h_1 from h_0. We have from (15)

$$(16)\quad \frac{\partial}{\partial x}\left\{e^{-\frac{1}{2}x^2}\frac{c_1(\sigma^2,x)}{\sqrt{\Sigma\lambda_i\sigma_i^2}}\right\}, + \left\{1 + \Sigma\frac{\sigma_i^4\partial_i^2}{f_i}\right\}\exp\left\{-\frac{x^2}{2}\frac{(\Sigma\lambda_i w_i)}{(\Sigma\lambda_i\sigma_i^2)}\right\}\sqrt{\frac{(\Sigma\lambda_i w_i)}{(\Sigma\lambda_i\sigma_i^2)}} = e^{-\frac{1}{2}x^2}$$

i.e.

$$(17)\quad \frac{\partial}{\partial x}\left\{e^{-\frac{1}{2}x^2}\frac{c_1(\sigma^2,x)}{\sqrt{\Sigma\lambda_i\sigma_i^2}}\right\} + \frac{(\Sigma\lambda_i^2\sigma_i^4/f_i)}{(\Sigma\lambda_i\sigma_i^2)^2}\, d^2\exp\left\{-\frac{x^2u}{2}\right\}\sqrt{u} = 0$$

where d now denotes differentiation with respect to u and subsequent equation to unity

i.e.

$$(18)\quad \frac{\partial}{\partial x}\left\{e^{-\frac{1}{2}x^2}\frac{c_1(\sigma^2, x)}{\sqrt{\Sigma\lambda_i\sigma_i^2}}\right\} = \frac{(\Sigma\lambda_i^2\sigma_i^4/f_i)}{(\Sigma\lambda_i\sigma_i^2)^2}\frac{1}{4}e^{-\frac{1}{2}x^2}(1 + 2x^2 - x^4)$$

$$(19)\quad = \frac{1}{4}\frac{(\Sigma\lambda_i^2\sigma_i^4/f_i)}{(\Sigma\lambda_i\sigma_i^2)^2}\frac{\partial}{\partial x}\left\{e^{-\frac{1}{2}x^2}(x + x^3)\right\}$$

whence

$$(20)\quad c_1(\sigma^2, x) = x\sqrt{\Sigma\lambda_i\sigma_i^2}\left[\frac{(1 + x^2)}{4}\frac{(\Sigma\lambda_i^2\sigma_i^4/f_i)}{(\Sigma\lambda_i\sigma_i^2)^2}\right].$$

Hence to the terms in $1/f_i$ we have

$$y - \eta = h(s^2, x) = x\sqrt{\Sigma\lambda_i s_i^2}\left[1 + \frac{(1 + x^2)}{4}\frac{(\Sigma\lambda_i^2 s_i^4/f_i)}{(\Sigma\lambda_i s_i^2)^2}\right]. \tag{21}$$

Solving this out for x we obtain to the same order

$$x = v\left[1 - \frac{(1 + v^2)}{4}\frac{(\Sigma\lambda_i^2 s_i^4/f_i)}{(\Sigma\lambda_i s_i^2)^2}\right]. \tag{22}$$

where v equals $(y - \eta)/\sqrt{\Sigma\lambda_i s_i^2}$. To order $1/f_i$ we may regard x as a unit normal deviate and hence determine the probability level corresponding to the observed ratio v. On the other hand if we wish to determine the value of $y - \eta$ which will lie on a given percentage level the expression (21) is the appropriate one to use.

4. Further discussion. The present development is of course basically equivalent to that given in the previous paper. Indeed if we integrate (10) or (15) out with respect to x we arrive immediately at the formulae which were then obtained and which were illustrated by calculating terms to order $1/f_i^2$. In fact when calculating higher order terms it seems best to do this integration before carrying out the operation Θ. The object of the present note is really to stress the fact that we are simply finding a function of the observations and of $y - \eta$ which is distributed as a unit normal deviate, whatever the values the true σ_i^2 may chance to possess.

Finally, the remarks following equation (7) above should be somewhat amplified. The equation asserts that the distribution of any arbitrary function x, defined by (3), is

$$p(x) = \int_{s^2}\cdots\int \frac{1}{\sqrt{2\pi\Sigma\lambda_i\sigma_i^2}}\left\{\exp - \frac{1}{2}\frac{h^2(s^2, x)}{\Sigma\lambda_i\sigma_i^2}\right\}\frac{\partial h(s^2, x)}{\partial x}\prod_i \{p(s_i^2)\, ds_i^2\}, \tag{23}$$

where $h(s, x)$ is the function obtained by solving out (3) for $y - \eta$. On carrying out the integrations in (23) we shall in general obtain $p(x)$ as a function of x and σ_i^2. Our argument is that if h be chosen properly the σ_i^2 will disappear from $p(x)$, and x will appear only in the form of the unit normal probability function.

To find $h(s, x)$ by a direct process of numerical integration would appear to involve in the first instance the choice of a net-work of points for x and s_i^2. Suppose the range of x is covered by n_x points and the range of s_i^2 by n_i points. We may then as an approximation look on our task as that of finding the $(n_x\pi_i n_i)$ values of $h(s^2, x)$ corresponding to this network. Since (23) is to be true for all x and σ_i^2, we can take in turn n_i values of σ_i^2, and then (23) can be replaced by $(n_x\pi_i n_i)$ simultaneous equations (it would be necessary to use some formula expressing $\partial h(s^2, x)/\partial x$ in terms of values of $h(s^2, x)$ at discrete values of x or conceivably this may be avoided if we work with the integrated form). With a proper choice of the points for x, s_i^2, and σ_i^2, we might expect to evaluate the series $h(s^2, x)$ to any required degree of accuracy, but clearly as a general process to be used over a whole range of values f_i this approach would be too laborious.

It may indeed be queried whether theoretically, with an indefinitely fine network of points, we shall be led to a unique function $h(s^2, x)$ with the common sense properties, which, from general statistical considerations, we know it should have in order to be acceptable. As with integral equations of a simpler character, the passage from a discrete network to a continuum may raise problems, but it is the author's opinion that the infinite ranges of x and s_i^2 give us the freedom which we require in the solution.

The author, however, prefers to approach the problem from the numerical behavior of the series, of which (15) gives the general terms. Here the practical issue appears to be to investigate the relation between the magnitude of the last term retained and the f_i. The author hopes in a further paper to give some results of an investigation of this character and also some tables facilitating the calculation of $h(s^2, x)$.

REFERENCE

[1] B. L. WELCH, "The generalization of 'Student's' problem when several different population variances are involved". *Biometrika*, Vol. 34 (1947), pp. 28–35.

PROBABILITY SCHEMES WITH CONTAGION IN SPACE AND TIME[1]

BY FÉLIX CERNUSCHI[2] AND LOUIS CASTAGNETTO

Harvard University

1. Summary. In many natural assemblies of elements, the probability of an event for a given element depends not only on the intrinsic nature of that particular element, but also on the states of some or all of the rest of the elements belonging to the same assembly. On the basis of this general idea of "contagion" some urn schemes are developed in this paper in which one has contagious influence in space and time. The most interesting result found is that in general the points of convergence of the probability of the assembly are given by some of the roots of an equation $p = f(p)$ and that some of these roots, between zero and one, represent stable states of the assembly, or points of convergence, and others represent unstable ones, or points of divergence. The two neighboring roots, (if they are single), of a root representing a point of convergence are unstable values of the probability. Consequently, under certain conditions, the limiting probability may be made to have a finite jump by changing the initial probability by an arbitrarily small amount. The concrete cases developed in this paper can be considerably extended by similar methods by assuming more complicated and general assemblies and laws of contagion.

[1] On the suggestion of the referee, some parts of the original paper were deleted and some mathematical simplifications were introduced.

[2] Research Associate at Harvard Astronomical Observatory and Guggenheim Fellow.