19/10/34.

F.Yates B.A.,
Rothamsted Experimental Station,
Harpenden,
Herts.

Dear Yates,

I'am interested in your paper Bias in sampling and think it would be suitable for the Annals if you can reiterate the application to Man in the summary and perhaps make some illusion to Neyman's exellent paper to the Statistical Society last summer.

I return the paper herewith with a few pencil notes and hope to see it again soon.

Yours Sincerly,

(α) $\frac{1}{(p-1)}\left\{(R_1-a_1)^2+(R_2-a_2)^2+\dots\right\}-\frac{1}{p(p-1)}(G-T_a)^2$ R − a

(β) $\frac{1}{p(p-1)(p-2)}\left[\left\{(p-1)(C_1-a_1)+(R_1-a_1)\right\}^2+\dots\dots\right]-\frac{1}{(p-1)(p-2)}(G-T_a)^2$

(γ) $\frac{1}{p}\left[T_b^2+T_c^2+\dots\dots\right]-\frac{1}{p(p-1)}(G-T_a)^2$

(δ) $\frac{1}{p(p-2)}\left[\left\{p\,a_1-R_1-C_1\right\}^2+\dots\dots\right]-\frac{1}{p^2(p-2)}(p\,T_a-2G)^2$

$2(p-1)-(p-2)^2$

The key to the interconnections between the two sets of degrees of freedom is that rows and columns are completely contained within part rows, part columns and "within a". Hence the remaining (p-1) D.F. in the latter group must be pure error in the customary division, and the remaining (p-3) (p-1) D.F. for customary error must be my error.

In the case of the 4 x 4 square above one set of single error degrees of freedom is as follows.

<u>Involving a.</u>

(1) $\begin{array}{cccc} + & \cdot & \cdot & - \\ \cdot & - & + & \cdot \\ \cdot & + & - & \cdot \\ - & \cdot & \cdot & + \end{array}$ (2) $\begin{array}{cccc} + & \cdot & - & \cdot \\ \cdot & - & \cdot & + \\ - & \cdot & + & \cdot \\ \cdot & + & \cdot & - \end{array}$ (3) $\begin{array}{cccc} + & - & \cdot & \cdot \\ - & + & \cdot & \cdot \\ \cdot & \cdot & - & + \\ \cdot & \cdot & + & - \end{array}$

<u>Not involving a.</u>

(4) $\begin{array}{cccc} \cdot & + & - & \cdot \\ - & \cdot & \cdot & + \\ + & \cdot & \cdot & - \\ \cdot & - & + & \cdot \end{array}$ (5) $\begin{array}{cccc} \cdot & + & \cdot & - \\ + & \cdot & - & \cdot \\ \cdot & - & \cdot & + \\ - & \cdot & + & \cdot \end{array}$ (6) $\begin{array}{cccc} \cdot & \cdot & + & - \\ \cdot & \cdot & - & + \\ - & + & \cdot & \cdot \\ + & - & \cdot & \cdot \end{array}$

(1), (3), (4), (6) can be rehashed into:

$$
\begin{array}{cccc}
 & (1)+(3)-(4)+(6) & & (1)-(3)-(4)-(6) & & (1)+(3)-(4)-(6) & & (1)-(3)-(4)+(6) \\
(7) & \begin{array}{cccc} + & - & + & - \\ \cdot & \cdot & \cdot & \cdot \\ - & + & - & + \\ \cdot & \cdot & \cdot & \cdot \end{array} &
(8) & \begin{array}{cccc} \cdot & \cdot & \cdot & \cdot \\ + & - & + & - \\ \cdot & \cdot & \cdot & \cdot \\ - & + & - & + \end{array} &
(9) & \begin{array}{cccc} + & \cdot & - & \cdot \\ - & \cdot & + & \cdot \\ + & \cdot & - & \cdot \\ - & \cdot & + & \cdot \end{array} &
(10) & \begin{array}{cccc} \cdot & + & \cdot & - \\ \cdot & - & \cdot & + \\ \cdot & + & \cdot & - \\ \cdot & - & \cdot & + \end{array}
\end{array}
$$

These are various combinations of the differences $a + c - b - d$, $a - c$, and $b - d$ for the individual rows. The other combinations give the column degrees of freedom. The system is:

Rows	$a + c - b - d$	$a - c$	$b - d$
1 - 3	(7)	x_1	y_2
2 - 4	- (8)	y_1	x_2
1 + 3 - 2 - 4	Columns	(2)	(5)

$$
\left.\begin{array}{l} x_1 + x_2 \\ y_1 + y_2 \end{array}\right\} \text{Cols.} \qquad \begin{array}{l} x_1 - x_2 = (9) \\ y_1 - y_2 = -(10) \end{array}
$$

The connection with the randomised block system is now clear.

If a and c are phosphate and b and d no phosphate one might test (7) and (8) and / or (9) and (10) (which are composed of $a + c - b - d$ for columns) against (2) and (5), but it doesn't seem to be possible to confine differential phosphate effect to 2 degrees of freedom. Why should it ? Perhaps this test might be generalised for squares whose sides are even (at least for some transformation sets).

The differential effect of <u>a</u> against the rest appears to be fairly tested by (1), (2), (3) against (4), (5), (6).

The above partitioning is peculiar to the transformation set of the square. There is for instance no error degree of freedom depending only on comparisons of a and b. In the other transformation set there are six orthogonal degrees of freedom depending one each on comparisons of the pairs of treatments.

The process of splitting up the error degrees of freedom into sets of $p - 1$ can be continued for more than one treatment by the subtraction method, but there do not appear to be any simple expressions for the sums of squares due to part rows and part columns when more than one treatment is missing. $(p - 1)^2$ linear equations must be solved whose coefficients depend on the pattern of the square.

Yours sincerely,

F Yates