TESTING A STRAGGLER MEAN IN A TWO-WAY CLASSIFICATION USING THE RANGE

By Jack Moshman

Oak Ridge National Laboratory

1. Summary and introduction. The use of the range in place of the standard deviation as a measure of dispersion has long been recognized as a convenient and easily calculated statistic. Possibly its most notable employment has been in the industrial statistician's quality control charts. A statistic based upon the range is described below which may be used to test whether one of a group of means may be considered to be a straggler from all or some of the others in a two-way analysis of variance.

2. Previous literature. Nair [1] derived the distribution of $u_\nu = | x_k - \bar{x} | / s_\nu$ and $u_\nu' = | \bar{x} - x_1 | / s_\nu$, where $x_1 \leq x_2 \leq \cdots \leq x_k$, $\bar{x} = \sum_{i=1}^{k} x_i/k$ and s_ν is an independent estimate of the standard deviation based on ν degrees of freedom from a normal population.

If x_1 and x_k are considered row or column means in a two-way analysis of variance, $\bar{x}$ the grand mean, and s_ν the error root-mean-square estimate of the standard deviation, u_ν or u_ν' may be adapted to examine an individual sample mean as a possible straggler from the grand mean. In this case $\nu = (r - 1)(k - 1)$ if there exist r rows and k columns and the statistic takes the form $u_\nu = (x_{.k} - \bar{x})\sqrt{r}/s_\nu$, where $x_{.k}$ is the largest of the k column means. Nair provided critical values of u_ν and u_ν' for $k = 3(1)\ 9$ and $\nu = 10(1)\ 20, 24, 30, 40, 60, 120$ and ∞.

Tukey [2] suggested an empirical approximation to Nair's procedure which does not involve the use of any special tables. He showed that

$$w = \frac{u_\nu - \frac{6}{5}\log_{10}k}{3\left(\frac{1}{4} + \frac{1}{\nu}\right)}, \qquad k > 3;$$

or

$$w = \frac{u_\nu - \frac{1}{2}}{3\left(\frac{1}{4} + \frac{1}{\nu}\right)}, \qquad k = 3,$$

may be treated as a normal deviate and will test the most deviant mean from the grand mean as a straggler.

3. The *g*-statistic. Hartley [3] recently prepared, as an alternative to the F ratio, a test based solely on the range to examine row and column homogeneity in a two-way classification. It is proposed here to develop a statistic, also based on the range, to examine one of a group of means as a possible straggler from the grand mean.

Consider Nair's statistic applied to column means:

$$u_\nu \doteq \frac{(x_{.k} - \bar{x})\sqrt{r}}{s_\nu}.$$

It is well known that s_ν may be approximated by the range. Let us consider the usual probability model

$$x_{ij} = \alpha_i + \beta_j + \epsilon_{ij},$$

where α_i and β_j are the effects of the ith row and jth column respectively, and the ϵ_{ij} are error variates distributed as $N(0, \sigma^2)$. Then

$$x_{i.} = \alpha_i + \bar{\beta} + \epsilon_{i.},$$

whence

$$x_{ij} - x_{i.} = (\beta_j - \bar{\beta}) + (\epsilon_{ij} - \epsilon_{i.}),$$

which we shall call row residuals. In the jth column, the range of the row residuals is equal to the range of the r independent normal deviates $(\epsilon_{ij} - \epsilon_{i.})$, distributed as $N(0, (k - 1)\sigma^2/k)$. We may define $\bar{w}_{k,r}$ as the mean of the k ranges (k columns) with r observations in each. Patnaik [4] showed that the distribution of $\bar{w}_{k,r}/(\sigma c_{k,r})$ may be approximated by that of $\chi/\sqrt{\nu'}$, where $c_{k,r}$ is an appropriate scale factor and ν' is the "equivalent number of degrees of freedom."

This is done by equating the first two moments of $\bar{w}_{k,r}/(\sigma \,.\, c_{k,r})$ to those of $\chi/\sqrt{\nu'}$:

$$(1) \qquad E\left(\frac{\bar{w}_{k,r}}{\sigma . c_{k,r}}\right) = \sqrt{\frac{k-1}{k}}\,\frac{d_r}{c_{k,r}} = \sqrt{\frac{2}{\nu'}}\,\frac{\Gamma\left(\frac{\nu'+1}{2}\right)}{\Gamma\left(\frac{\nu'}{2}\right)},$$

$$(2) \qquad \mathrm{Var}\left(\frac{\bar{w}_{k,r}}{\sigma . c_{k,r}}\right) = \frac{V_r(k-1)[1 + (k-1)\rho_w]}{k^2 . c_{k,r}^2} = \frac{1}{\nu'}\left[\nu' - 2\left\{\frac{\Gamma\left(\frac{\nu'+1}{2}\right)}{\Gamma\left(\frac{\nu'}{2}\right)}\right\}^2\right],$$

where d_r and V_r are the population mean and variance respectively of the range in samples of r from a normal population with unit variance and ρ_w is the correlation between any two column ranges, derived from the row residuals. Hartley ([3], p. 276) provides a table of ρ_w, and V_r is tabulated in Pearson [5]. But since $s_{\nu'}/\sigma$ is distributed as $\chi/\sqrt{\nu'}$ we may let $s_{\nu'} = \bar{w}_{k,r}/c_{k,r}$ after determining ν' and $c_{k,r}$ from the solution of equations (1) and (2).

From the theory of the analysis of variance, it is known that

$$(\epsilon_{ij} - \epsilon_{i.} - \epsilon_{.j} + \bar{\epsilon}),$$

$(\epsilon_{i.} - \bar{\epsilon})$ and $(\epsilon_{.j} - \bar{\epsilon})$ are independently distributed ([6], p. 344). Since $\bar{w}_{k,r}$ is a function of $(\epsilon_{ij} - \epsilon_{i.} - \epsilon_{.j} + \bar{\epsilon})$ only and $(x_{.k} - \bar{x})$ and $(x_{r.} - \bar{x})$ are functions only of $(\epsilon_{.j} - \bar{\epsilon})$ and $(\epsilon_{i.} - \bar{\epsilon})$ respectively, it follows that $\bar{w}_{k,r}$, $(x_{.k} - \bar{x})$ and $(x_{r.} - \bar{x})$ are also independently distributed.

Returning to Nair's statistic we have, after substituting $\bar{w}_{k,r}/c_{k,r}$ for s_ν,

$$u_\nu = \frac{|x_{.k} - \bar{x}|\sqrt{r}}{s_\nu} \approx \frac{|x_{.k} - \bar{x}|\sqrt{r}c_{k,r}}{\bar{w}_{k,r}} = u_{\nu'}.$$

By replacing all means with summations we have

$$g = \frac{\sqrt{r}u_{\nu'}}{c_{k,r}} = \frac{\left|k\sum_{i=1}^{r} x_{ik} - \sum_{j=1}^{k}\sum_{i=1}^{r} x_{ij}\right|}{W_{k,r}}, \tag{3}$$

where $W_{k,r} = k\,\bar{w}_{k,r} = \sum_{j=1}^{k} w_r^{(j)}$, and $w_r^{(j)}$ is the range of the r residuals in the jth column. The g-statistic is now in its simplest form. Obviously the distribution of g is the same for $x_{.k}$ or $x_{.1}$.

$W_{k,r}$ is the sum of the k column ranges of the row residuals of r observations each. Alternately one may use $W_{r,k}$ by first determining the column residuals and then summing the r row ranges. In order to enjoy the maximum number of degrees of freedom it is advisable to use $W_{k,r}$ or $W_{r,k}$ according as r or k is greater respectively [3]. Letting $l = \max(r, k)$ and $s = \min(r, k)$, then $W_{s,l}$ is the preferred denominator. Since the range of the residuals is independent of the numerator, there will be no advantage in preferentially selecting $W_{r,k}$ or $W_{k,r}$ on the basis of the numerator. In some cases, however, the arithmetic processes will be simpler when using $W_{l,s}$ and outweigh the advantages of slight increase in the equivalent number of degrees of freedom.

Four alternate forms of equation (3) are possible by interchanging r and k in the numerator and the denominator independently. When $k > r$ we define

$$g_{||} = \frac{\left|r\sum_{j=1}^{k} x_{rj} - \sum_{i=1}^{r}\sum_{j=1}^{k} x_{ij}\right|}{w_{s,l}}$$

to test whether the rth row mean is a straggler, and

$$g_{\perp} = \frac{\left|k\sum_{i=1}^{r} x_{ik} - \sum_{i=1}^{r}\sum_{j=1}^{k} x_{ij}\right|}{w_{s,l}}$$

to test whether the kth column mean is a straggler from the grand mean.

The notation $g_{||}$ (read g-parallel) indicates the means are compared in a manner parallel to the determination of residual ranges and $g_{\perp}$ (read g-perpendicular) indicates means are compared perpendicular to the array of residuals when their range is determined. Specifically if $k > r$, then $W_{s,l}$ is determined by subtracting column means and determining ranges horizontally along the rows. Hence to test row means we work parallel to the range determinations and perpendicular for column means.

In Table 1 are tabulated the 5 per cent and 1 per cent critical values of $g_{||}$ and $g_{\perp}$ when $W_{s,l}$, the preferred measurement, is used in the denominator. $g_{||}$ is found above the main diagonal, $g_{\perp}$ below; the two coincide on the main diagonal. Table 2 is similar to Table 1, but $W_{l,s}$ is the denominator and $g_{||}$ is now on and below the main diagonal and $g_{\perp}$ is on and above. In each case s is found along the top stub and l along the side when the statistic is $g_{\perp}$. Their positions are reversed for $g_{||}$.

TABLE 1

1 per cent and 5 per cent critical values of $g_{\perp}$† and $g_{||}$† using denominator $W_{s,l}$*

$l(g_{\perp})$ $s(g_{\|\|})$ \ $s(g_{\perp})$ $l(g_{\|\|})$	3	4	5	6	7	8	9
3	3.94 3.21	3.90 3.05	3.94 3.05	3.96 3.10	4.00 3.15	4.08 3.22	4.16 3.32
4	3.38 2.64	3.46 2.61	3.46 2.65	3.51 2.74	3.57 2.81	3.64 2.90	3.73 3.00
5	3.05 2.40	3.10 2.37	3.21 2.47	3.27 2.57	3.35 2.66	3.44 2.76	3.53 2.85
6	2.80 2.19	2.87 2.23	2.98 2.34	3.13 2.47	3.22 2.57	3.30 2.67	3.40 2.77
7	2.62 2.06	2.70 2.13	2.83 2.25	2.98 2.38	3.14 2.51	3.22 2.61	3.33 2.72
8	2.50 1.97	2.58 2.05	2.72 2.18	2.86 2.31	3.02 2.45	3.17 2.56	3.27 2.67
9	2.40 1.92	2.48 2.00	2.63 2.13	2.77 2.76	2.93 2.40	3.08 2.51	3.22 2.64

* Upper figure for $\alpha = 1$ per cent, lower figure for $\alpha = 5$ per cent.

† g on$_{\perp}$ and below main diagonal, $g_{||}$ on and above main diagonal.

4. Application. As an illustration of the use of the g-statistic, we may consider an example given by Rider ([7], p. 147), which appears in Table 3.

We wish to test whether hazel does actually flower significantly earlier than at least some of the other plants. We first form the row residuals by subtracting each station mean from the plants at that station and compute the column ranges as in Table 4.

TABLE 2

1 per cent and 5 per cent critical values of $g_\perp$† and $g_\parallel$† using denominator $W_{l,s}$*

$s(g_\perp)$ $l(g_\parallel)$ / $l(g_\perp)$ $s(g_\parallel)$	3	4	5	6	7	8	9
3	3.94 3.21	4.10 3.10	4.23 3.10	4.36 3.19	4.49 3.30	4.64 3.42	4.77 3.55
4	3.55 2.69	3.46 2.61	3.58 2.69	3.69 2.81	3.85 2.93	3.99 3.06	4.15 3.20
5	3.28 2.41	3.20 2.40	3.21 2.47	3.35 2.60	3.50 2.73	3.64 2.85	3.80 3.00
6	3.08 2.25	3.02 2.29	3.06 2.37	3.13 2.47	3.28 2.60	3.43 2.73	3.57 2.85
7	2.94 2.16	2.91 2.22	2.95 2.31	3.04 2.41	3.14 2.51	3.27 2.64	3.42 2.75
8	2.84 2.10	2.82 2.16	2.88 2.26	2.97 2.37	3.06 2.47	3.17 2.56	3.31 2.69
9	2.75 2.05	2.77 2.13	2.83 2.24	2.92 2.33	3.02 2.43	3.12 2.54	3.22 2.64

* Upper figure for α = 1 per cent, lower figure for α = 5 per cent.
† $g_\perp$ on and above main diagonal, $g_\parallel$ on and below main diagonal.

We find that $l = 6$, $s = 5$ and $W_{s,l} = 150.8$. To test whether hazel is a straggler, we run parallel to the range layouts and have

$$g_{\parallel} = \frac{|(5)(650) - 7176|}{150.8} = 26.03.$$

Since $W_{s,l}$ was used in the denominator, we refer to Table 1 and above the main diagonal. Now l runs along the top and s along the side stub. The 1 per cent critical value for $l = 6$ and $s = 5$ is 3.27 and we conclude that hazel is indeed a straggler with level of significance $\alpha << .01$.

If we wished to test whether Bratton is a straggler from the other stations, we would be running perpendicular to the range layout and would have

$$g_{\perp} = \frac{|(6)(1145) - 7176|}{150.8} = 2.03.$$

Referring to Table 1 below the main diagonal, we see that the 5 per cent critical value is 2.34 and the difference is not significant.

TABLE 3

Day of year of flowering of five plants at six stations

Station	Plant					Totals	Means
	Hazel	Colts-foot	Ane-mone	Black-thorn	Mus-tard		
Broadchalke	131	205	274	299	337	1246	249.2
Bratton	84	176	276	291	318	1145	229.0
Lenham	131	196	262	299	333	1221	244.2
Dorstone	106	194	239	317	344	1200	240.0
Coaley	77*	190	275	298	332	1172	234.4
Ipswich	121	179	271	293	328	1192	238.4
Totals	650	1140	1597	1797	1992	7176	

* This figure was misprinted as 777 in Rider [7].

TABLE 4

Station residuals and variety ranges from Table 3

Station	Plant					Totals
	Hazel	Coltsfoot	Ane-mone	Black-thorn	Mustard	
Broadchalke	−118.2	−44.2	24.8	49.8	87.8	
Bratton	−145.0	−53.0	47.0	62.0	89.0	
Lenham	−113.2	−48.2	17.8	54.8	88.8	
Dorstone	−134.0	−46.0	−1.0	77.0	104.0	
Coaley	−157.4	−44.4	40.6	63.6	97.6	
Ipswich	−117.4	−59.4	32.6	54.6	89.6	
Range	44.2	15.2	48.0	27.2	16.2	150.8

5. Remarks. In the example, if station residuals and plant ranges were computed, Table 2 should have been used. $W_{5,6}$ has 18.5 degrees of freedom associated with it; $W_{6,5}$ has 18.2 degrees of freedom, a slight difference. No different in significance of the test results would have been obtained. As a check on the procedure it was found that $\bar{w}_{5,6}/c_{5,6} = 13.11$ and $\bar{w}_{6,5}/c_{6,5} = 14.93$. The value of σ estimated by s in the illustration was 13.99, a reasonably good check.

In the event that $r = k$, one has the same number of degrees of freedom, but in general $W_{r,k} \neq W_{k,r}$. The difference will be a sampling fluctuation of the ϵ_{ij} and will ordinarily make little difference except when g lies close to one of the critical values, but in practice, one makes little differentiation between levels of significance of 6 per cent and 4 per cent.

Acknowledgment. The author wishes to express his indebtedness to Professor John W. Tukey for indirectly suggesting this problem, and for some direct recommendations concerning the notation used in this paper.

REFERENCES

[1] K. R. NAIR, "The distribution of the extreme deviate from the sample mean and its studentized form," *Biometrika*, Vol. 35 (1948), pp. 118–144.
[2] J. W. TUKEY, "Comparing individual means in the analysis of variance," *Biometrics*, Vol. 5 (1949), pp. 99–114.
[3] H. O. HARTLEY, "The use of range in analysis of variance," *Biometrika*, Vol. 37 (1950), pp. 271–280.
[4] P. B. PATNAIK, "The use of mean range as an estimator of variance in statistical tests," *Biometrika*, Vol. 37 (1950), pp. 78–87.
[5] E. S. PEARSON, "The percentage limits for the distribution of range in samples from a normal population," *Biometrika*, Vol. 24 (1932), pp. 404–409.
[6] A. M. MOOD, *An Introduction to the Theory of Statistics*, McGraw-Hill Book Co., 1950.
[7] P. R. RIDER, *An Introduction to Statistical Methods*, John Wiley and Sons, 1939.