COLLOQUIUM MATHEMATICUM

VOL. XIII 1965 FASC. 2

ON TESTING THE EQUALITY OF PARAMETERS IN RECTANGULAR POPULATIONS FOR UNEQUAL SAMPLE SIZES

BY

W. KLONECKI (POZNAŃ)

Khatri [1] has given a statistic for testing the equality of ranges in k rectangular populations when the lower limit of each population is known to be zero. He has also found the distribution of that statistic for equal sample sizes and tabulated the $5\,^0/_0$ points. In this paper the distribution of the Khatri's statistic for unequal sample sizes is given. The $\alpha\,^0/_0$ points can be then computed when sample sizes are known.

For $i = 1, 2, \ldots, k$ let X_{ij}, where $j = 1, 2, \ldots, n_i$, be independent observations from rectangular populations with range Θ_i and lower limit equal to zero. Let D_i be the sample maximum value. The density function of D_i is given by

$$\frac{n_i}{\Theta_i}\left(\frac{d_i}{\Theta_i}\right)^{n_i-1} \quad \text{for} \quad 0 \leqslant d_i \leqslant \Theta_i. \tag{1}$$

Let $D_{\max}$ be the maximum value and $D_{\min}$ be the minimum value of the random variables $D_1, D_2, \ldots, D_k$.

The hypothesis H_0: $\Theta_1 = \Theta_2 = \ldots = \Theta_k$ can be tested by Khatri's statistic $D_{\min}/D_{\max}$ and is accepted if

$$\frac{D_{\min}}{D_{\max}} > u. \tag{2}$$

An $\alpha\,^0/_0$ test of H_0 is then obtained by taking u as the lower $\alpha\,^0/_0$ point of $D_{\min}/D_{\max}$.

The distribution of $D_{\min}/D_{\max}$ under the null hypothesis is obtained as follows.

Let $\Theta_1 = \Theta_2 = \ldots = \Theta_k = \Theta$. Without any restriction we can assume that $\Theta = 1$. Let $(t_1, t_2, \ldots, t_k)$ be a permutation of the indices

[1] C. G. Khatri, *On testing the equality of parameters in k rectangular populations*, Journal of the American Statistical Association 55 (1960), p. 144-147.

$(1, 2, \ldots, k)$ and

$$X_{(t_1, t_2, \ldots, t_k)} \tag{3}$$

be the region in the sample space leading to $0 < d_{t_1} < d_{t_2} < \ldots < d_{t_k} < 1$. Using the above given nomenclature, we have

$$P\left\{\frac{D_{\min}}{D_{\max}} > u\right\} = \sum_{(t_1, t_2, \ldots, t_k)} P\left\{\frac{D_{\min}}{D_{\max}} > u \,\middle|\, X_{(t_1, t_2, \ldots, t_k)}\right\} P\{X_{(t_1, t_2, \ldots, t_k)}\}, \tag{4}$$

where the summation is over the $k!$ permutations corresponding to the disjoint $k!$ sets (3). Note that

$$P\left\{\frac{D_{\min}}{D_{\max}} > u \,\middle|\, X_{(t_1, t_2, \ldots, t_k)}\right\} = P\left\{\frac{D_{t_1}}{D_{t_k}} > u \,\middle|\, X_{(t_1, t_2, \ldots, t_k)}\right\}. \tag{5}$$

First take $(t_1, t_2, \ldots, t_k) = (1, 2, \ldots, k)$. According to (5) we have

$$P\left\{\frac{D_{\min}}{D_{\max}} > u \,\middle|\, X_{(1, 2, \ldots, k)}\right\} = P\left\{\frac{D_1}{D_k} > u \,\middle|\, X_{(1, 2, \ldots, k)}\right\}. \tag{6}$$

To evaluate the conditional probability (6) write

$$f(d_1, d_2, \ldots, d_k) = \begin{cases} \prod_{i=1}^{k} n_i d_i^{n_i - 1} & \text{for} \quad 0 < d_1 < d_2 < \ldots < d_k < 1, \\ 0 & \text{otherwise.} \end{cases} \tag{7}$$

When divided by $p_{(1, 2, \ldots, k)} = P\{X_{(1, 2, \ldots, k)}\}$ function (7) represents the density of a random variable, say $(D^{(1)}, D^{(2)}, \ldots, D^{(k)})$. Thus

$$P\left\{\frac{D_{\min}}{D_{\max}} > u \,\middle|\, X_{(1, 2, \ldots, k)}\right\} = P\left\{\frac{D^{(1)}}{D^{(k)}} > u\right\}. \tag{8}$$

Now, putting $U_1 = D^{(1)}/D^{(k)}$, and $U_i = D^{(i)}$ for $i = 2, 3, \ldots, k$, we obtain the joint distribution of $(U_1, U_2, \ldots, U_k)$ whence, after integration, the distribution of U_1 can be found.

The density of $(U_1, U_2, \ldots, U_k)$ is given by

$$h(u_1, u_2, \ldots, u_k) \tag{9}$$
$$= \begin{cases} \dfrac{1}{p_{(1, 2, \ldots, k)}} \left(\prod_{i=1}^{k} n_i\right) \left(\prod_{i=1}^{k-1} u_i^{n_i - 1}\right) u_k^{n_1 + n_k - 1} & \text{for} \quad (u_1, u_2, \ldots, u_k) \in A, \\ 0 & \text{otherwise,} \end{cases}$$

and the region A by the inequalities $0 < u_k < 1$, $0 < u_r < u_{r+1}$, when $r = 2, 3, \ldots, k-1$, and $0 < u_1 u_k < u_2$. We obtain by integrating the following expression for the density of U_1:

$$(10)\qquad g(u_1) = \frac{1}{p_{(1,2,\dots,k)}}\left(\prod_{i=1}^{k} n_i\right) u_1^{n_1-1}\int_0^1 du_k\, u_k^{n_1+n_k-1}\times$$

$$\times\int_{u_1u_k}^{u_k} du_{k-1}\, u_{k-1}^{n_{k-1}-1}\dots\int_{u_1u_k}^{u_4} du_3\, u_3^{n_3-1}\int_{u_1u_k}^{u_3} u_2^{n_2-1}du_2$$

for $0 < u_1 < 1$ and $g(u_1) = 0$ for $u_1 \leqslant 0$ or $u_1 \geqslant 1$. To find a convenient formula for the integral (10) first note that

$$(11)\qquad \int_{u_1u_k}^{u_s} du_{s-1}\, u_{s-1}^{n_{s-1}-1}\dots\int_{u_1u_k}^{u_3} u_2^{n_2-1}du_2$$

$$= \frac{1}{n_2(n_2+n_3)\dots(n_2+n_3+\dots+n_{s-1})}\, u_s^{n_2+n_3+\dots+n_{s-1}}+$$

$$+\sum_{i=3}^{s-1}(-1)^i\frac{1}{n_i(n_i+n_{i+1})\dots(n_i+n_{i+1}+\dots+n_{s-1})}\times$$

$$\times\frac{1}{n_{i-1}(n_{i-1}+n_{i-2})\dots(n_{i-1}+n_{i-2}+\dots+n_2)}\, u_s^{n_i+\dots+n_{s-1}}(u_1u_k)^{n_{i-1}+\dots+n_2}+$$

$$+(-1)^s\frac{1}{n_{s-1}(n_{s-1}+n_{s-2})\dots(n_{s-1}+n_{s-2}+\dots+n_2)}\,(u_1u_k)^{n_2+\dots+n_{s-1}}.$$

Formula (11) is obtained by induction. In the proof of (11) the formula

$$(12)\qquad \frac{(-1)^{r+1}}{n_r(n_r+n_{r-1})\dots(n_r+n_{r-1}+\dots+n_1)}$$

$$= \frac{1}{n_1(n_1+n_2)\dots(n_1+n_2++\dots+n_r)}+$$

$$+\sum_{i=2}^{r}(-1)^{i-1}\frac{1}{n_i(n_i+n_{i+1})\dots(n_i+n_{i+1}+\dots+n_r)}\times$$

$$\times\frac{1}{n_{i-1}(n_{i-1}+n_{i-2})\dots(n_{i-1}+n_{i-2}+\dots+n_1)}$$

is used, which holds whenever $n_\beta+\dots+n_\gamma \neq 0$ for $1 \leqslant \beta \leqslant \gamma \leqslant r$.

To prove (12) multiply both sides by $n_r(n_r+n_{r-1})\dots(n_r+n_{r-1}+\dots$ $\dots+n_1)$. Consider the left side as a function of n_r, say $h(n_r)$. This is a polynomial. Moreover, the degree of this polynomial is lower than r.

On the other hand, there are r points such that

$$h(0) = h(-n_r) = h(-n_r - n_{r-1}) = h(-n_r - n_{r-1} - n_{r-2}) = \ldots$$
$$= h(-n_{r-1} - \ldots - n_1) = (-1)^{r+1}.$$

Hence $h(n_r) \equiv (-1)^{r+1}$, and (12) follows.

Formulas (10) and (11) yield

$$(13)\quad g(u_1) = \frac{1}{p_{(1,2,\ldots,k)}} \frac{\prod_{i=1}^{k} n_i}{\sum_{i=1}^{k} n_i} u_1^{n_1-1} \Bigg[\frac{1}{n_2(n_2+n_3)\ldots(n_2+n_3+\ldots+n_{k-1})} +$$
$$+ \sum_{i=3}^{k-1} (-1)^i \frac{1}{n_i(n_i+n_{i+1})\ldots(n_i+n_{i+1}+\ldots+n_{k-1})} \times$$
$$\times \frac{1}{n_{i-1}(n_{i-1}+n_{i-2})\ldots(n_{i-1}+n_{i-2}+\ldots+n_2)} u_1^{n_{i-1}+\ldots+n_2} +$$
$$+ (-1)^k \frac{1}{n_{k-1}(n_{k-1}+n_{k-2})\ldots(n_{k-1}+n_{k-2}+\ldots+n_2)} u_1^{n_{k-1}+\ldots+n_2} \Bigg].$$

Finally, by integrating and using (12) we conclude that

$$(14)\quad P\left\{ \frac{D_{\min}}{D_{\max}} > u \,\middle|\, X_{(1,2,\ldots,k)} \right\} = \int_u^1 g(t)\,dt$$
$$= \frac{1}{p_{(1,2,\ldots,k)}} \frac{\prod_{i=1}^{k} n_i}{\sum_{i=1}^{k} n_i} \Bigg[\frac{1}{n_1(n_1+n_2)\ldots(n_1+n_2+\ldots+n_{k-1})} +$$
$$+ \sum_{i=2}^{k-1} (-1)^{i-1} \frac{1}{n_i(n_i+n_{i+1})\ldots(n_i+n_{i+1}+\ldots+n_{k-1})} \times$$
$$\times \frac{1}{n_{i-1}(n_{i-1}+n_{i-2})\ldots(n_{i-1}+n_{i-2}+\ldots+n_1)} u^{n_{i-1}+\ldots+n_1} +$$
$$+ (-1)^{k-1} \frac{1}{n_{k-1}(n_{k-1}+n_{k-2})\ldots(n_{k-1}+n_{k-2}+\ldots+n_1)} u^{n_{k-1}+\ldots+n_1} \Bigg].$$

Now, using the fact that by appropriate permutations of indices $(1, 2, \ldots, k)$ in formula (14) the expressions for the particular terms in the sum (4) are obtained, we get

$$(15)\quad P\left\{\frac{D_{\min}}{D_{\max}} > u\right\}$$

$$= \frac{\prod_{i=1}^{k} n_i}{\sum_{i=1}^{k} n_i} \sum_{(t_1, t_2, \ldots, t_k)} \left[\frac{1}{n_{t_1}(n_{t_1}+n_{t_2})\ldots(n_{t_1}+n_{t_2}+\ldots+n_{t_{k-1}})} + \right.$$

$$+ \sum_{i=2}^{k-1} (-1)^{i-1} \frac{1}{n_{t_i}(n_{t_i}+n_{t_{i+1}})\ldots(n_{t_i}+n_{t_{i+1}}+\ldots+n_{t_{k-1}})} \times$$

$$\times \frac{1}{n_{t_{i-1}}(n_{t_{i-1}}+n_{t_{i-2}})\ldots(n_{t_{i-1}}+n_{t_{i-2}}+\ldots+n_{t_1})} u^{n_{t_{i-1}}+\ldots+n_{t_2}} +$$

$$\left. +(-1)^{k-1} \frac{1}{n_{t_{k-1}}(n_{t_{k-1}}+n_{t_{k-2}})\ldots(n_{t_{k-1}}+n_{t_{k-2}}+\ldots+n_{t_1})} u^{n_{t_{k-1}}+\ldots+n_{t_1}}\right],$$

where the summation is over the $k!$ permutations of the indices.

It is also easy to show that formula (15) can be simplified and presented in the form

$$(16)\quad P\left\{\frac{D_{\min}}{D_{\max}} > u\right\}$$

$$= 1 - \frac{1}{n_1+n_2+\ldots+n_k} \left[\sum_{i=1}^{k} (n_1+\ldots+n_{i-1}+n_{i+1}+\ldots+n_k) u^{n_i} - \right.$$

$$- \sum_{1 \leqslant i < j \leqslant k} (n_1+\ldots+n_{i-1}+n_{i+1}+\ldots+n_{j-1}+n_{j+1}+\ldots+n_k) u^{n_i+n_j} + \ldots \pm$$

$$\left. \pm \sum_{i=1}^{k} n_i u^{n_1+\ldots+n_{i-1}+n_{i+1}+\ldots+n_k}\right].$$

It may be remarked that (16) agrees for $n_1 = n_2 = \ldots = n_k$ with the formula given by Khatri.

When the lower limits are not known to be zero, the distribution of the Khatri's statistic for unequal sample sizes may be obtained in the same manner. Unfortunately, the formula is too complicated to be presented here.

Reçu par la Rédaction le 9. 10. 1964