A LAGUERRE PRODUCT SERIES APPROXIMATION TO ONE-WAY CLASSIFICATION VARIANCE RATIO DISTRIBUTION

By M. L. Tiku

Indian Statistical Institute

Introduction

An approximation to the sampling distribution of the variance ratio of one-way classification for analysis of variance and of the general non-normal variance ratio was given by Gayen (1950). He started with a Gram Charlier series expansion of the probability density function of the population and ignored all population cumulants of order greater than four and square and higher powers of the fourth cumulant. An alternative approach is presented in this paper. The probability density function of the variances has been formally expanded as a series of the product of Laguerre polynomials and Gamma-density functions. The coefficients of the first few terms are worked out in terms of population cumulants, up to the eighth order. The distribution of the ratio of variance components of one-way classification is given explicitly. However, the method works on the same lines for the general non-normal variance ratio. Gayen's expression agrees with this expansion to the order of approximation used by him.

1. *Laguérre Polynomials*

For $m > 0$, a Laguerre polynomial of degree r in x, $L_r^{(m)}(x)$ is defined as: see Szego (1939), Chapter V,

$$L_r^{(m)}(x) = \sum_{t=0}^{r} c_{r,t}^{(m)} \frac{(-x)^t}{t!} \tag{1.1}$$

where

$$c_{r,t}^{(m)} = \begin{cases} (m+t)(m+t+1)\ldots\dfrac{(m+r-1)}{(r-t)!} & \text{for } t = 0, 1, \ldots, r-1 \\ 1 & \text{for } t = r \\ & \text{for } r = 0, 1, 2, \ldots \end{cases}$$

In particular,

$$L_0^{(m)}(x) = 1$$

$$L_1^{(m)}(x) = m - x$$

$$L_2^{(m)}(x) = \frac{1}{2!} m(m+1) - (m+1)x + \frac{x^2}{2!}$$

$$L_3^{(m)}(x) = \frac{1}{3!} m(m+1)(m+2) - \frac{1}{2!}(m+1)(m+2)x + (m+2)\frac{x^2}{2!} - \frac{x^3}{3!}$$

$$L_4^{(m)}(x) = \frac{1}{4!} m(m+1)(m+2)(m+3) - \frac{1}{3!}(m+1)(m+2)(m+3)x + \frac{1}{2!}(m+2)(m+3)\frac{x^2}{2!} - (m+3)\frac{x^3}{3!} + \frac{x^4}{4!}. \tag{1.2}$$

If we write

$$\psi_m(x) = \frac{1}{\Gamma(m)} e^{-x} x^{m-1},\ x \geqslant 0$$

for the Gamma-density function with parameter m, the orthogonality property of Laguerre polynomials can be stated as

$$\int_0^\infty L_r^{(m)}(x)\, L_s^{(m)}(x)\, \psi_m(x)\, dx = \begin{cases} 0 & \text{if } r \neq s \\ c_{r,o}^{(m)} & \text{if } r = s \end{cases} \tag{1.3}$$

2. *One-way Classification*

In a one-way lay-out let x_{ij}; $i = 1, 2, \ldots, b$, $j = 1, 2, \ldots, 1\, n_i$; be the j-th observation in the i-th block. We assume that x_{ij} have the same distribution which we specify by the finite cumulants $x_1, x_2, \ldots, x_r, \ldots$. Consider the identity

$$\sum_i \sum_j (x_{ij} - \bar{x}_{..})^2 = \sum_i n_i (\bar{x}_{i.} - \bar{x}_{..})^2 + \sum_i \sum_j (x_{ij} - \bar{x}_{i.})^2$$

$$= S_1 \qquad + S_2, \text{ say} \tag{2.1}$$

where

$$\bar{x}_{i.} = \sum_j \frac{x_{ij}}{n_i} \text{ and } \bar{x}_{..} = \sum_i \sum_j \frac{x_{ij}}{N},\ N = \sum_i n_i$$

S_1 and S_2 are said to have $v_1 = (b-1)$ and $v_2 = (N-b)$ degrees of freedom respectively. Our object here is to obtain the distribution of

$$F = \frac{\frac{S_1}{v_1}}{\frac{S_2}{v_2}} \tag{2.2}$$

under the assumption of no block effect. To be concise we take

$$n_1 = n_2 = \ldots = n_b = n \tag{2.3}$$

otherwise the method is general.

3. *Distribution of S_1 and S_2*

Define the r-th population standard cumulant Λ_r as

$$\Lambda_r = k_r k_2^{\frac{1}{2}r},\ r = 3, 4, \ldots\ldots \tag{3.1}$$

Alternatively we write $x_2 = \sigma^2$. Let $X = S_1/2\sigma^2$ and $Y = S_2/2\sigma^2$. Denote the probability density function of X and Y by $\phi(X, Y)$. The quotient $\phi(X, Y)/\psi_k(X)\psi_p(Y)$; $k = v_1/2$, $p = v_2/2$; can be formally expanded as a series of the product of Laguerre polynomials as

$$\frac{\phi(X, Y)}{\psi_k(X)\psi_p(Y)} = \sum_{r=0}^{\infty}\sum_{s=0}^{\infty} \beta_{rs} L_r^{(k)}(X)\, L_s^{(p)}(Y) \tag{3.2}$$

We assume here the convergence of (3.2). Multiplying both sides of (3.2) by $L_r^{(k)}(X)\, L_s^{(p)}(Y)$ and integrating over the domain ($0 \leqslant X \leqslant \infty$, $0 \leqslant Y \leqslant \infty$), we get

$$\beta_{rs} = E\left[\frac{L_r^{(k)}(X)\ L_s^{(p)}(Y)}{c_{r,0}^{(k)}\ c_{s,0}^{(p)}}\right] \tag{3.3}$$

where E denotes mathematical expectation, by virtue of orthogonality property (1.3). What we seek here is an approximation to $\phi(X, Y)$. We take

$$\phi(X, Y) \sim \sum_{r+s \leqslant 4}\sum [\beta_{rs} L_r^{(k)}(X)\, L_s^{(p)}(Y)]\, \psi_k(X)\, \psi_p(Y) \tag{3.4}$$

To evaluate β_{rs} for $r + s \leqslant 4$, we replace x by X, m by k and writing $u = X - k$, the polynomials given by (1.2) can be expressed in terms of u as

$$L_0^{(k)} = 1$$

$$L_1^{(k)} = -u$$

$$L_2^{(k)} = \tfrac{1}{2}u^2 - u - \tfrac{1}{2}k$$

$$L_3^{(k)} = -\tfrac{1}{6}u^3 + u^2 + (\tfrac{1}{2}k - 1)\,u - \tfrac{2}{3}k$$

$$L_4^{(k)} = \tfrac{1}{24}u^4 - \tfrac{1}{2}u^3 + \tfrac{1}{4}(6 - k)\,u^2 + (\tfrac{7}{6}k - 1)\,u + \tfrac{1}{8}k\,(k - 6) \tag{3.5}$$

We obtain similar polynomials $L_0^{(p)}, L_1^{(p)}, L_2^{(p)}, L_3^{(p)}$ and $L_4^{(p)}$ in V by writing $V = y - p$ and replacing m by p and x by y in (1.2). Using the relation

$$E(u^r V^s) = \left(\frac{1}{2\sigma^2}\right)^{r+s} E\left[(S_1 - E(S_1))^r\,(S_2 - E(S_2))^s\right] \tag{3.6}$$

and an ingenius process of converting product moments into product cumulants due to Kendall (1940) and a table due to David and Johnson (1951), we can very easily write the expressions for $E(u^r V^s)$ for $r + s \leqslant 4$; for example

$$E(u^2) = \frac{v_1^2}{4N}\,\Lambda_4 + \frac{v_1}{2}$$

$$E(uV) = \frac{v_1 v_2}{4N}\,\Lambda_4$$

$$E(V^2) = \frac{v_2^2}{4N}\,\Lambda_4 + \frac{v_2}{2}$$

$$E(u^3) = \frac{v_1^3}{8N^2}\,\Lambda_6 + \frac{3v_1^2}{2N}\,\Lambda_4 + \frac{v_1(v_1 - 1)}{2N}\,\Lambda_3^2 + v_1$$

$$E(uV^2) = \frac{v_1 v_2}{8N}\left(\frac{v_2}{N}\,\Lambda_6 + 4\Lambda_3^2 + 4\Lambda_4\right)$$

$$E(u^2 V) = \frac{v_1 v_2}{8N}\left(\frac{v_1}{N}\,\Lambda_6 + 4\Lambda_4\right)$$

$$E(V^3) = \frac{v_2^3}{8N^2}\,\Lambda_6 + \frac{3v_2^2}{2N}\,\Lambda_4 + \frac{v_2(v_2 - v_1 - 1)}{2N}\,\Lambda_3^2 + v_2 \tag{3.7}$$

and so on. From (3.7), it is very easy to obtain $E[L_r^{(k)} L_s^{(p)}]$ which when substituted in (3.3) give the following:

$$\beta_{00} = 1$$

$$\beta_{10} = 0$$

$$\beta_{20} = \frac{v_1}{N(v_1+2)} \Lambda_4$$

$$\beta_{11} = \frac{1}{N} \Lambda_4$$

$$\beta_{02} = \frac{v_2}{N(v_2+2)} \Lambda_4$$

$$\beta_{30} = -\frac{1}{N(v_1+2)(v_1+4)} \left(\frac{v_1^2}{N} \Lambda_6 + 4(v_1-1) \Lambda_3^2 \right)$$

$$\beta_{21} = -\frac{v_1}{N^2(v_1+2)} \Lambda_6$$

$$\beta_{12} = -\frac{1}{N(v_2+2)} \left(\frac{v_2}{N} \Lambda_6 + 4 \Lambda_3^2 \right)$$

$$\beta_{03} = -\frac{1}{N(v_2+2)(v_2+4)} \left(\frac{v_2^2}{N} \Lambda_6 + 4(v_2-v_1-1) \Lambda_3^2 \right)$$

$$\beta_{40} = \frac{1}{N^2(v_1+2)(v_1+4)(v_1+6)} \times \left(\frac{v_1^3}{N} \Lambda_8 + 32v_1(v_1-1)\Lambda_5\Lambda_3 + (3v_1^3+32v_1^2-8v_1+8)\Lambda_4^2 \right)$$

$$\beta_{31} = \frac{1}{N^2(v_1+2)(v_1+4)} \times \left(\frac{v_1^2}{N} \Lambda_8 + 8(v_1-1)\Lambda_5\Lambda_3 + 3v_1(v_1+4)\Lambda_4^2 \right)$$

$$\beta_{22} = \frac{1}{N^2(v_1+2)(v_2+2)} \times \left(\frac{v_1 v_2}{N} \Lambda_8 + 16v_1 \Lambda_5\Lambda_3 + (3v_1v_2+12v_1+4v_2)\Lambda_4^2 \right)$$

$$\beta_{13} = \frac{1}{N^2(v_2+2)(v_2+4)} \times \left(\frac{v_2^2}{N} \Lambda_8 + 8(4v_2-v_1-1)\Lambda_5\Lambda_3 + (3v_2^2+20v_2-8v_1-8)\Lambda_4^2 \right)$$

$$\beta_{04} = \frac{1}{N^2 (v_2 + 2)(v_2 + 4)(v_2 + 6)}$$

$$\times \left(\frac{(v_2{}^3}{N} \Lambda_8 + 32 v_2 (v_2 - v_1 - 1) \Lambda_5 \Lambda_3 + (v_2{}^2 (3v_2 + 32) \right.$$

$$\left. - 8 (v_1 + 1)(v_2 - v_1 - 1)) \Lambda_4{}^2 \right). \tag{3.8}$$

The coefficients (3.8) have the same interesting pattern as the result (2.11) of Roy and Tiku (1962); there is a misprint in the expression for their $a_3{}^{(m)}$; and the result (4.3) of author (1962). Given the values (3.8) we can write explicitly the approximation (3.4) for $\phi(X, Y)$. If we ignore Λ_r; $r > 4$ and $\Lambda_4{}^2$, $\phi(X, Y)$ is in perfect agreement with Gayen's formula (2.17). We note that the integral of $\phi(X, y)$ over the domain $(0 \leqslant X \leqslant \infty,\ 0 \leqslant Y \leqslant \infty)$ is obviously unity due to the following property

$$\int_0^\infty L_r{}^{(m)}(x)\, \psi_m(x)\, dx = \begin{cases} 1 \text{ if } r = 0 \\ 0 \text{ if } r > 0 \end{cases} \tag{3.9}$$

of Laguerre polynomials. Also by virtue of (3.9), the distribution of X or of Y can be very easily written from $\phi(X, Y)$, say for example the distribution of X is obtained as

$$\sum_{r=0}^{4} \beta_{ro} L_r{}^{(k)}(X)\, \psi_k(X). \tag{3.10}$$

4. *Distribution of F*

Submitting $\phi(X, y)$ to the transformation

$$\frac{v_1}{v_2} F = \frac{X}{Y} \tag{4.1}$$

and integrating over Y from zero to infinity noting that a typical term $X^r y^s \psi_k(X) \psi_p(Y)$ integrates to

$$\frac{\Gamma\left(\frac{v_1}{2} + r\right) \Gamma\left(\frac{v_2}{2} + s\right)}{\Gamma\left(\frac{v_1}{2}\right) \Gamma\left(\frac{v_2}{2}\right)}\ \beta\left(\frac{v_1}{2} + r,\ \frac{v_2}{2} + s;\ F\right) \tag{4.2}$$

where

$$\beta\left(\frac{v_1}{2}+r,\ \frac{v_2}{2}+s;\ F\right)$$

$$=\frac{\left(\dfrac{v_1}{v_2}\right)^{(v_1/2)+r}}{\beta\left(\dfrac{v_1}{2}+r, \dfrac{v_2}{2}+s\right)}\cdot\frac{F^{(v_1/2)+r-1}}{\left(1+\dfrac{v_1}{x_2}F\right)^{(v_1+v_2/2)+r+s}}.$$

The probability density function $p(F)$ of F is thus obtained as

$$p(F) \sim F_0 + \Lambda_4 F_{\Lambda_4} + \Lambda_3^2 F_{\Lambda_3^2} + \Lambda_6 F_{\Lambda_6} + \Lambda_4^2 F_{\Lambda_4^2}$$
$$+ \Lambda_5\Lambda_3 F_{\Lambda_5\Lambda_3} + \Lambda_8 F_{\Lambda_8} \tag{4.3}$$

where

$$F_0 = \beta\left(\frac{v_1}{2}, \frac{v_2}{2};\ F\right)$$

and the other terms may be called corrective terms due to finite population cumulants. F_{Λ_4} and $F_{\Lambda_3^2}$ are given by Gayen. For brevity we reproduce here F_{Λ_6} and F_{Λ_8} only; the other terms can be very easily written from $\phi(X, y)$ using (4.2) and (4.1) together with (3.8) or can be had from the author.

$$F_{\Lambda_6} = \frac{1}{48N^2}\Bigg[\left\{v_1^3\,\beta\left(\frac{v_1}{2}+3, \frac{v_2}{2};\ F\right) - 3v_1^2(v_1+v_2)\,\beta\left(\frac{v_1}{2}+2, \frac{v_2}{2};\ F\right)\right.$$
$$\left.+ 3v_1(v_1+v_2)^2\,\beta\left(\frac{v_1}{2}+1, \frac{v_2}{2};\ F\right) - (v_1+v_2)^3\,\beta\left(\frac{v_1}{2}, \frac{v_2}{2};\ F\right)\right\}$$
$$+3\left\{v_1^2v_2\,\beta\left(\frac{v_1}{2}+2, \frac{v_2}{2}+1;\ F\right) - 2v_1v_2(v_1+v_2)\,\beta\left(\frac{v_1}{2}+1, \frac{v_2}{2}+1;\ F\right)\right.$$
$$\left.+v_2(v_1+v_2)^2\,\beta\left(\frac{v_1}{2}, \frac{v_2}{2}+1;\ F\right)\right\} + 3\left\{v_1v_2^2\,\beta\left(\frac{v_1}{2}+1, \frac{v_2}{2}+2;\ F\right)\right.$$
$$\left.- v_2^2(v_1+v_2)\,\beta\left(\frac{v_1}{2}, \frac{v_2}{2}+2;\ F\right)\right\} + v_2^3\,\beta\left(\frac{v_1}{2}, \frac{v_2}{2}+3;\ F\right)\Bigg] \tag{4.4}$$

$$F_{\Lambda_8} = \frac{1}{384N^3}\Bigg[\left\{v_1^4\,\beta\left(\frac{v_1}{2}+4, \frac{v_2}{2};\ F\right) - 4v_1^3(v_1+v_2)\,\beta\left(\frac{v_1}{2}+3, \frac{v_2}{2};\ F\right)\right.$$
$$+ 6v_1^2(v_1+v_2)^2\,\beta\left(\frac{v_1}{2}+2, \frac{v_2}{2};\ F\right) - 4v_1(v_1+v_2)^3\,\beta\left(\frac{v_1}{2}+1, \frac{v_2}{2};\ F\right)$$
$$\left.+ (v_1+v_2)^4\,\beta\left(\frac{v_1}{2}, \frac{v_2}{2};\ F\right)\right\} + 4v_2\left\{v_1^3\,\beta\left(\frac{v_1}{2}+3, \frac{v_2}{2}+1;\ F\right)\right.$$

$$-3v_1^2(v_1+v_2)\,\beta\left(\frac{v_1}{2}+2,\frac{v_2}{2}+1;F\right)+3v_1(v_1+v_2)^2\beta\left(\frac{v_1}{2}+1,\frac{v_2}{2}+1;F\right)$$

$$-(v_1+v_2)^3\,\beta\left(\frac{v_1}{2},\frac{v_2}{2}+1;F\right)\Big\}+6v_2^2\Big\{v_1^2\,\beta\left(\frac{v_1}{2}+2,\frac{v_2}{2}+2;F\right)$$

$$-2v_1(v_1+v_2)\,\beta\left(\frac{v_1}{2}+1,\frac{v_2}{2}+2;F\right)+(v_1+v_2)^2\,\beta\left(\frac{v_1}{2},\frac{v_2}{2}+2;F\right)\Big\}$$

$$+4v_2^3\Big\{v_1\,\beta\left(\frac{v_1}{2}+1,\frac{v_2}{2}+3;F\right)-(v_1+v_2)\,\beta\left(\frac{v_1}{2},\frac{v_2}{2}+3;F\right)$$

$$+v_2^4\,\beta\left(\frac{v_1}{2},\frac{v_2}{2}+4;F\right)\Big\}\Big] \qquad (4.5)$$

Summary

The first few terms of the Laguerre product series expansion of the distribution of variance components of one-way classification are worked out in terms of population cumulants, of up to the eighth order. The distribution of variance ratio is given explicitly.

Remark

Similar corrective terms to the distribution of the ratio of variances of two random samples from two different populations have been obtained by the same technique and can be had from the author.

Further Work

The order of approximation of (4.3) is under investigation.

Acknowledgement

I am highly grateful to Dr. J. Roy for his kind guidance and encouragements and a referee for valuable suggestions.

References

David, F. N. and Johnson, N. L. .. "The effect of non-normality on the power function of the F-test in the analysis of variance," *Biometrika*, 1951, **38**, 43–57.

David, F. N. .. "The moments of *Z*- and *F*-distribution," *ibid.*, 1949, **36**, 394–403.

Fisher, R. A. .. "Moments and product moments of sampling distributions," *Proc. Lond. Math. Soc.*, 1928, **30 B**, 199–238.

Gayen, A. K. .. "The distribution of the variance ratio in random samples of any size drawn from non-normal universes," *Biometrika*, 1950, **37**, 236-51.

Geary, R. C. .. "Tests of Normality, *ibid.*, 1947, **34**, 209-42.

Kendall, M. G. .. "The derivation of multivariate sampling formula by symbolic operation," *Annals of Eugenics*, 1940, **10**, 392–402.

Roy, J. and Tiku, M. L. .. "A Laguerre series approximation to the sampling distribution of the variance," *Sankhya*, *Series A*, 1962, **24** (2), 181-84.

Sansone, G. .. *Orthogonal Functions*, Interscience publishers, Inc., New York, 1959.

Szego, G. .. "Orthogonal polynomials," *Amer. Math. Soc. Colloquium Publication*, 1939, **23**.

Tiku, M. L. .. "Distribution of Student's *t* in terms of Hermite and Laguerre polynomilas," 1962 (to be published).