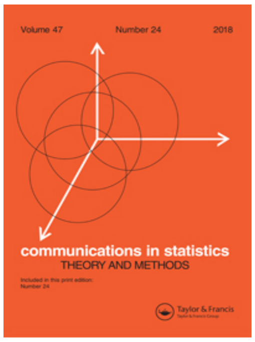

Taylor & Francis
Taylor & Francis Group

Communications in Statistics - Theory and Methods

ISSN: 0361-0926 (Print) 1532-415X (Online) Journal homepage: https://www.tandfonline.com/loi/lsta20

Approximate critical values with error bounds for bartlett's test of homogeneity of variances for unequal sample sizes

J.L. Cyr & E.B. Manoukian

To cite this article: J.L. Cyr & E.B. Manoukian (1982) Approximate critical values with error bounds for bartlett's test of homogeneity of variances for unequal sample sizes, Communications in Statistics - Theory and Methods, 11:15, 1671-1680, DOI: 10.1080/03610928208828340

To link to this article: https://doi.org/10.1080/03610928208828340

COMMUN. STATIST.-THEOR. METH., 11(15), 1671-1680 (1982)

APPROXIMATE CRITICAL VALUES WITH ERROR BOUNDS FOR BARTLETT'S TEST OF HOMOGENEITY OF VARIANCES FOR UNEQUAL SAMPLE SIZES.

J.L. Cyr and E.B. Manoukian

Department of National Defence, Royal Military College of Canada
Kingston, Ontario, Canada K7L 2W3

Key Words and Phrases: Bartlett's test of homogeneity of variances; tests of hypotheses.

ABSTRACT

Errors are given for the departure of Bartlett's exact distribution for the test of homogeneity of variances for unequal sample sizes from that of equal sample sizes. Let $\nu_j + 1$ denote the size of the j^{th} sample and define a departure number $\delta = \max\,(\nu_j - \underline{\nu})/\nu_j$, where $\underline{\nu} = \min\,\nu_j$. Then these errors may be expressed as functions of only three parameters: the number of populations k, the departure number δ and $\underline{\nu}$. Accurate critical values of the test, in the general case with unequal sample sizes, are given when these errors are small.

1. INTRODUCTION

There has been much interest recently (Layard, 1973; Glaser, 1976; Chao and Glaser, 1978; Dyer and Keating, 1980; Glaser, 1980; Nandi, 1980; Manoukian, 1980) in the classic Bartlett test (Bartlett, 1937) for the test of homogeneity of variances. The assumption of normality of the underlying populations is essential for this test as the latter is well known to be sensitive to departures from normality (i.e., non-robust) (Box, 1953; Box and Andersen, 1955; Gartside, 1972). The W test of Wilk and Shapiro (1968) for the joint assessment of normality of several independent samples may, for example, may be first carried out. If normality is in doubt robust procedures (Keselman et al., 1979) are, of course, to be used. The Bartlett test is unbiased (Pitman, 1939) and is consistent against all alternatives (Brown, 1939). The Bartlett test is general enough to allow samples of different sizes. It is well known that the chi-square distribution provides an approximation to the Bartlett distribution when the sample sizes are large (Bartlett, 1937). Improvements, however, to the chi-square as approximations to the Bartlett distribution have been suggested in the literature (e.g., Box, 1949; Hartley, 1940) and corresponding (and hence approximate) critical values of the test have been tabulated (Thompson and Merrington, 1946; Pearson and Hartley, 1966). Unfortunately, Bartlett's distribution involves too many parameters which prohibit tabulation of the critical values in the general case. The underlying mathematical analysis to this end, in the general case with unequal sample sizes, is quite complicated, but some interesting progress in this direction has been recently made (Glaser, 1980). For the equal sample size case the problem simplifies to some extent and the exact critical values of the test may be computed and have been tabulated (Harsaae, 1969; Glaser, 1976; Dyer and Keating, 1980). Recently (Manoukian, 1980) we have made a rigorous analysis in the comparison of the Bartlett distribution for unequal sample sizes with the corresponding one with equal sample sizes and we have obtained maximum error bounds to the approximation of the former by the latter. More precisely, let ν_j denote the degree of freedom of the j^{th} sample, $\underline{\nu} = \min \nu_j$ and define a departure number $\delta = \max (\nu_j - \underline{\nu})/\nu_j$; then we obtain an error estimate of the form : $\bar{\varepsilon}(\delta,\underline{\nu},k) \geq |P[M^1/C_1 \leq x] - P[M^2/C_2 \leq x]|$, uniformly in x, where k denotes the number of populations under consideration, M^1 and M^2 are Bartlett's variables with unequal and equal ($\nu_j = \underline{\nu}$) sample sizes, respectively, $C_1 = C_1(k,\{\nu_j\})$ and $C_2 = C_2(k,\underline{\nu})$ are suitable scaling factors with $C_2(k,\underline{\nu}) = C_1(k,\{\nu_j\})|_{\nu_j=\underline{\nu}}$. The right-hand-side of the above inequality clearly depends on k+2 parameters, whereas the error bound $\bar{\varepsilon}(\delta,\underline{\nu},k)$ depends only on the three parameters indicated.

For $\bar{\varepsilon}$ sufficiently small, accurate critical values of the test may be then determined to be $M_c^1 = (C_1/C_2)M_c^2$, where M_c^2 is the corresponding critical value for the equal sample size case and C_1, C_2 are readily calculable scaling factors. When $\bar{\varepsilon}$ is sufficiently small, the test criterion that we suggest in this paper is then to reject the hypothesis H_o of the equality of variances if $M^1 > M_c^1$.

2. ERROR ESTIMATES AND CRITICAL VALUES

Bartlett's variable may be written as

$$M^1 = -\sum_{j=1}^{k} \nu_j \ln(S_j^2/S^2) \ , \ S^2 = \sum_{j=1}^{k} \frac{\nu_j S_j^2}{\nu} \ , \ \nu = \sum_{j=1}^{k} \nu_j \ , \qquad (2.1)$$

where S_j^2 is the variance of the j^{th} sample of size ν_j+1. If the hypothesis of the equality of variances of the k populations is true, then M^1 may be written as

$$M^1 = -\sum_{j=1}^{k} \nu_j \ \ln\left(\frac{\chi_j^2/\nu_j}{\chi^2/\nu}\right) \quad , \qquad (2.2)$$

where χ_j^2 is a chi-square variable of ν_j degrees of freedom and χ^2 is of ν degrees of freedom. The characteristic function associated with the variable M^1/C_1, where C_1 is a suitable scaling factor, is well known (e.g., Bartlett, 1937 or Glaser, 1980) to be given by

$$\phi_1(t) = \frac{\prod_{j=1}^{k} \exp\left[i\frac{t}{C_1} \nu_j \ln \frac{\nu_j}{\nu}\right]\left[\Gamma(\frac{\nu_j}{2}(1 - \frac{2it}{C_1}))/\Gamma(\nu_j/2)\right]}{\left[\Gamma(\frac{\nu}{2}(1 - \frac{2it}{C_1}))/\ \Gamma(\nu/2)\right]} \qquad (2.3)$$

where $\Gamma(z)$ is the gamma function (Re $z > 0$). We may then write

$$|P[M^1/C_1 \leq x] - P[M^2/C_2 \leq x]|$$

$$= \frac{1}{2\pi}\left|\int_{-\infty}^{\infty}\left(\frac{e^{-itx} - 1}{it}\right)[\phi_1(t) - \phi_2(t)]\ dt\right| , \qquad (2.4)$$

where $M^2 = M^1\Big|_{\nu_j=\underline{\nu}}$, $\underline{\nu} = \min \nu_j$. We have estimated the magnitude of the error in (2.4) by breaking the t-integral into two parts: $0 < |t| \leq \lambda$ $\lambda < |t| < \infty$, where λ is an arbitrary positive parameter and is conveniently chosen to minimize our estimates. From the estimates (i) and (ii) in the following lemma we may then readily carry out the t-integral in (2.4) over a bound of $[\phi_1(t) - \phi_2(t)]$.

<u>Lemma 1</u>

(i) $|\phi_1(t) - \phi_2(t)| \leq e^{\delta k\theta(\underline{\nu}) + \delta k\theta(k\underline{\nu})}\, G(t),$

$$G(t) = \frac{2}{9}\,\delta\underline{\nu}^{-2}t^2(k+1) + \delta\underline{\nu}^{-3}\sum_{j=0}^{2} k[A_j(\underline{\nu},k) + A_j(k\underline{\nu},k)]\frac{|t|^{j+1}}{j+1},$$

$$\theta(\underline{\nu}) = \frac{1}{6\underline{\nu}}\left(1 + \frac{1}{2\underline{\nu}}\right) + \frac{1}{\underline{\nu}^3}\left(1 + \frac{1}{4\underline{\nu}}\right)\left[\frac{2}{15} + \frac{16}{315\,\underline{\nu}^2}\right],$$

$$A_0(\underline{\nu},k) = \frac{2}{5} + \frac{3}{5}\underline{\nu}^{-1} + \frac{16}{9}\underline{\nu}^{-2} + \frac{120}{63}\underline{\nu}^{-3} + \frac{16}{3}\underline{\nu}^{-4} + \frac{56}{15}\underline{\nu}^{-5} +$$

$$+ \frac{256}{33}\underline{\nu}^{-6} + \frac{128}{33}\underline{\nu}^{-7},$$

$$A_1(\underline{\nu},k) = \frac{32}{21}\underline{\nu}^{-2} + \frac{16}{9}\underline{\nu}^{-3} + \frac{128}{15}\underline{\nu}^{-4} + \frac{32}{5}\underline{\nu}^{-5} + \frac{512}{33}\underline{\nu}^{-6} + \frac{256}{33}\underline{\nu}^{-7},$$

$$A_2(\underline{\nu},k) = \frac{16}{5}\underline{\nu}^{-4} + \frac{8}{3}\underline{\nu}^{-5} + \frac{256}{33}\underline{\nu}^{-6} + \frac{128}{33}\underline{\nu}^{-7}.$$

(ii) $|\phi_1(t) - \phi_2(t)| \leq \left(\frac{C_2}{2|t|}\right)^{\frac{k-1}{2}} e^{\theta_1(\underline{\nu},k) + \delta k\theta(k\underline{\nu})} \times$

$$\times\left[\frac{\delta k}{6\underline{\nu}}\, e^{\delta k\theta(\underline{\nu})} + \delta\theta(\underline{\nu})\left(\frac{e^{\delta k\theta(\underline{\nu})} - 1}{e^{\delta\theta(\underline{\nu})} - 1}\right) e^{\delta\theta(\underline{\nu}) + \delta k\theta(k\underline{\nu})}\right],$$

$$\theta_1(\underline{\nu},k) = \frac{1}{6\underline{\nu}}\left(k + \frac{1}{k}\right) + \frac{2}{45\underline{\nu}^3}\left(k + \frac{1}{k^3}\right).$$

The estimates in (i) and (ii) are valid for all t, and the scaling factors C_1 and C_2 have been chosen to be

$$C_1 = 1 + \frac{1}{3(k-1)}\left(\sum_{j=1}^{k} \frac{1}{\nu_j} - \frac{1}{\nu}\right) \quad , \quad C_2 = 1 + \frac{1}{3\underline{\nu}}\left(1 + \frac{1}{k}\right) .$$

With the scaling factors as just chosen, the error $\bar{\varepsilon}$ may be shown (Manoukian, 1980) to be $\bar{\varepsilon} = O(\underline{\nu}^{-\eta})$, $\eta > 1$, for $\underline{\nu} \to \infty$. On the other hand if the choice $C_1 = C_2 = 1$ is made then $\bar{\varepsilon} = O(\underline{\nu}^{-1})$ for $\underline{\nu} \to \infty$.

The estimate in (i) will be used when integrating over the region $0 < |t| \leq \lambda$, and the one in (ii) over the region $\lambda < |t| < \infty$. The proof of the lemma is tedious and quite long and makes use, in the process, of properties of the gamma function $\Gamma(z)$. A complete proof of the lemma is given in (Manoukian, 1980). From the estimates (i), (ii) and $|(e^{-itx} - 1)/t| \leq 2|t|^{-1}$, we may then readily carry out the t-integral to obtain the following maximum error bound:

$$|P[M^1/C_1 \leq x] - P[M^2/C_2 \leq x]| \leq \bar{\varepsilon}(\delta,\underline{\nu},k) \; ,$$

$$\bar{\varepsilon}(\delta,\nu,k) = e^{\delta k\theta(\underline{\nu})+\delta k\theta(k\underline{\nu})}\,\bar{\varepsilon}_1 + \left(\frac{C_2}{2\lambda}\right)^{\frac{k-1}{2}} e^{\theta_1(\underline{\nu},k)+\delta k\theta(k\underline{\nu})}\,\bar{\varepsilon}_2 ,$$

$$\bar{\varepsilon}_1 = \frac{2}{\pi}\left[\frac{\delta}{9}\,\underline{\nu}^{-2}\lambda^2(k+1) + \delta\underline{\nu}^{-3}\sum_{j=0}^{2} k[A_j(\underline{\nu},k) + A_j(k\underline{\nu},k)]\frac{\lambda^{j+1}}{(j+1)^2}\right],$$

$$\bar{\varepsilon}_2 = \frac{4}{\pi(k-1)}\left[\frac{\delta k}{6}\,\underline{\nu}^{-1}\,e^{\delta k\theta(\underline{\nu})} + \delta\theta(\underline{\nu})\left(\frac{e^{\delta k\theta(\underline{\nu})}-1}{e^{\delta\theta(\underline{\nu})}-1}\right)e^{\delta\theta(\underline{\nu})} + \delta k\theta(k\underline{\nu})\right] ,$$

$$\delta = \max\,(\nu_j-\underline{\nu})/\nu_j \; , \qquad \underline{\nu} = \min\,\nu_j \; .$$

The parameter λ is an arbitrary positive number and may be so adjusted as to minimize the error $\bar{\varepsilon}(\delta,\underline{\nu},k)$. For each triplet $(\delta,\underline{\nu},k)$, in general, a different minimizing λ is found. Once such values for λ are determined the errors are then computed numerica-

NUMERICAL ERROR ESTIMATES

	k=3			k=4			k=5			k=6			k=7		
$\nu \backslash \delta$	.16	.21	.26	.16	.21	.26	.16	.21	.26	.16	.21	.26	.16	.21	.26
4	--	.0379	.0473	--	.0332	.0415	--	.0319	.0399	--	.0319	.0400	--	.0326	.0410
5	.0185	.0244	.0304	.0153	.0202	.0252	.0141	.0186	.0232	.0136	.0181	.0226	.0136	.0180	.0225
6	.0133	.0175	.0217	.0106	.0139	.0173	.0094	.0124	.0155	.0089	.0118	.0147	.0087	.0115	.0144
7	.0102	.0134	.0167	.0079	.0103	.0129	.0068	.0090	.0112	.0064	.0084	.0104	.0061	.0081	.0101
$\nu \backslash \delta$	.31	.36	.41	.31	.36	.41	.31	.36	.41	.31	.36	.41	.31	.36	.41
4	.0568	.0664	.0762	.0499	.0584	.0671	.0480	.0566	.0649	.0482	.0567	.0654	.0496	.0584	.0675
5	.0364	.0425	.0486	.0302	.0353	.0404	.0279	.0326	.0375	.0271	.0319	.0365	.0271	.0319	.0367
6	.0260	.0303	.0347	.0208	.0242	.0277	.0186	.0217	.0249	.0176	.0206	.0236	.0173	.0203	.0232
7	.0199	.0232	.0265	.0154	.0179	.0205	.0134	.0157	.0179	.0125	.0146	.0167	.0121	.0141	.0161
$\nu \backslash \delta$	.46	.51	.66	.46	.51	.66	.46	.51	.66	.46	.51	.66	.46	.51	.66
4	.0860	.0961	.1269	.0759	.0849	.1128	.0736	.0824	.1101	.0743	.0835	.1122	.0768	.0864	.1169
5	.0548	.0611	.0802	.0457	.0509	.0672	.0423	.0473	.0626	.0414	.0463	.0615	.0416	.0466	.0622
6	.0391	.0435	.0570	.0312	.0348	.0457	.0280	.0313	.0412	.0267	.0298	.0394	.0263	.0294	.0389
7	.0298	.0332	.0434	.0231	.0257	.0337	.0202	.0225	.0296	.0188	.0210	.0277	.0182	.0204	.0269

$\nu \backslash \delta$	.06	.11	.16	.06	.11	.16	.06	.11	.16	.06	.11	.16	.06	.11	.16
9	--	.0046	.0068	--	.0034	.0050	--	.0029	.0042	--	.0026	.0038	--	.0024	.0036
14	.0013	.0024	.0034	.0009	.0016	.0024	.0007	.0013	.0019	.0006	.0011	.0016	.0006	.0010	.0015
24	.0006	.0011	.0016	.0004	.0007	.0010	.0003	.0005	.0008	.0002	.0004	.0006	.0002	.0004	.0005
29	.0005	.0008	.0012	.0003	.0005	.0008	.0002	.0004	.0006	.0002	.0003	.0005	.0001	.0003	.0004
$\nu \backslash \delta$	.21	.36	.51	.21	.36	.51	.21	.36	.51	.21	.36	.51	.21	.36	.51
9	.0089	.0154	.0219	.0065	.0113	.0162	.0055	.0095	.0136	.0050	.0086	.0124	.0047	.0081	.0117
14	.0045	.0078	.0111	.0031	.0053	.0076	.0025	.0043	.0060	.0021	.0037	.0052	.0020	.0033	.0048
24	.0021	.0036	.0055	.0013	.0023	.0032	.0010	.0017	.0024	.0008	.0014	.0020	.0007	.0012	.0016
29	.0016	.0027	.0039	.0010	.0017	.0024	.0007	.0013	.0018	.0006	.0010	.0015	.0005	.0010	.0013
$\nu \backslash \delta$	.66	.76	.91	.66	.76	.91	.66	.76	.91	.66	.76	.91	.66	.76	.91
9	.0285	.0330	.0398	.0211	.0244	.0295	.0176	.0207	.0250	.0161	.0188	.0227	.0153	.0178	.0216
14	.0144	.0167	.0200	.0099	.0114	.0138	.0079	.0091	.0110	.0068	.0079	.0095	.0062	.0072	.0087
24	.0066	.0076	.0091	.0042	.0049	.0058	.0031	.0036	.0043	.0026	.0030	.0036	.0023	.0026	.0032
29	.0050	.0059	.0070	.0031	.0036	.0044	.0023	.0027	.0032	.0019	.0022	.0026	.0016	.0019	.0023

lly (Cyr, 1981) by standard methods. These errors are tabulated below. We note that since $\delta = \max (\nu_j - \underline{\nu})/\nu_j = (\bar{\nu}-\underline{\nu})/\bar{\nu}$, where $\bar{\nu} = \max \nu_j$, and $\bar{\nu} \geq \underline{\nu}+1$, we must have $\delta \geq (1+\underline{\nu})^{-1}$ in practice.

3. DISCUSSION

The usefulness of the above results is immediate. Suppose we have k=6 populations and sample sizes 25,26,28,29,30 and 31. Then $\underline{\nu} = 24$, $\delta = .2$, $C_1 = 1.01441$, $C_2 = 1.01620$ and according to the table $\bar{\varepsilon}=.0008$. Hence with a nominal type I error probability of .05, the critical value of the test is $M_c^1 = (C_1/C_2)M_c^2 = .9982 \times 11.2420 = 11.2218$, where in the notation of Glaser (1976) $M_c^2 = - k\, \underline{\nu} \ln A^*$, and the true size of the test lies in the interval .05 ± .0008. It is interesting to compare this critical value with other approximation methods: M_c(Bartlett) = 11.2295 = M_c(Hartley).

In general, our nominally size α procedure rejects the hypothesis of equality of variances provided $M^1 > M_c^1$, where $M_c^1 = (C_1/C_2)M_c^2$ and M_c^2 is the tabulated (Harsae, 1969; Glaser, 1976 and Dyer and Keating, 1980) critical values of the comparable equal sample sizes test based on each $\nu_j = \nu$ (of size α). The true size of our test lies in the interval $\alpha \pm \bar{\varepsilon}$; thus our procedure should yield a satisfactory approximation provided $\bar{\varepsilon}$ is sufficiently small.

For handling situations when δ is not small, it may be interesting to carry out an analysis, as the one we have carried out, where the median of $\nu_1, \nu_2, \ldots, \nu_k$ replaces $\underline{\nu}$ in our estimates. This more difficult problem will be studied in a future work.

ACKNOWLEDGEMENTS

This work was supported by the Department of National Defence Award under CRAD No: 3610-637. One of the authors (E.B.M.) wishes to thank the referee for many valuable comments and suggestions.

BIBLIOGRAPHY

Bartlett, M.S. (1937). Properties of Sufficiency and Statistical Tests. Proc. Roy. Soc. London, Ser. A., 160, 268-282.

Box, G.E.P. (1949). A General Distribution Theory for a Class of Likelihood Criteria. Biometrika, 36, 317-345.

Box, G.E.P. (1953). Nonnormality and Tests on Variances. Biometrika, 40, 318-335.

Box, G.E.P. and Andersen, S.L. (1955). Permutation Theory in Derivation of Robust Criteria and the Study of the Departures from Assumption. Journ. Roy. Statist. Soc. Ser.B, 17, 1.26.

Brown, G.W. (1939). On the Power of the L_1 Test for Equality of Several Variances. Ann. Math. Statist., 10, 119-128.

Chao, M.T. and Glaser, R.E. (1978). The Exact Distribution of Bartlett's Test Statistic for Homogeneity of Variances with Unequal Sample Sizes. Journ. Amer. Statist. Assoc., 73, 422-426.

Cyr, J.L. (1981). Valeurs Critiques pour le Test d'Homogénéité des Variances pour des Echantillons de Différentes Tailles. Rapport d'ingenieur soumis pendant l'année finale du cours de Génie Gestion au R.M.C.

Dyer, D.D. and Keating, J.P. (1980). On the Determination of Critical Values for Bartlett's Test. Journ. Amer. Statist. Assoc., 75,313-319.

Gartside, P.S. (1972). A Study of Methods for Comparing Several Variances. Journ. Amer. Statist. Assoc., 67, 342-346.

Glaser, R.E. (1976). Exact Critical Values for Bartlett's Test for Homogeneity of Variances. Journ. Amer. Statist. Assoc., 71, 488-490.

Glaser, R.E. (1980). A Characterization of Bartlett's Statistic Involving Incomplete Beta Functions. Biometrika, 67, 53-58.

Harsaae, E. (1969). On the Computation and Use of a Table of Percentage Points of Bartlett's M. Biometrika, 56, 273-281.

Hartley, H.O. (1940). Testing the Homogeneity of a Set of Variances. Biometrika, 31, 249-255.

Keselman, J.H., Games, P.A. and Clinch, J.J. (1979). Tests for Homogeneity of Variances. Commun. Statist. B,8, 113-129

Layard, M.W. (1973). Robust Large-Sample Tests for Homogeneity of Variances. Journ. Amer. Statist. Assoc., 68, 195-198.

Manoukian, E.B. (1980). Departure of Bartlett's Distribution for the Homogeneity of Variances for Unequal Sample Sizes from that of Equal Sample Sizes. To be published in Metrika.

Nandi, S.B. (1980). On the Exact Distribution of a Normalized Ratio of the Weighted Geometric Mean to the Unweighted Arithmetic Mean in Samples from Gamma Distribution. Journ. Amer. Statist. Assoc., 75, 217-220.

Pearson, E.S. and Hartley, H.O. (1966). Biometrika Tables for Statisticians, Vol. 1, 3rd Edition, Cambridge University Press.

Pitman, E.J.G. (1939). Tests of Hypotheses Concerning Location and Scale Parameters. Biometrika, 31, 200-215.

Thompson, C.M. and Merrington, M. (1946). Tables for Testing the Homogeneity of a Set of Estimated Variances. Biometrika, 33, 296-301.

Wilk, M.B. and Shapiro, S.S. (1968). The Joint Assessment of Normality of Several Independent Samples. Technometrics, 10, 825-839.

Received by Editorial Board member September, 1981; Revised December, 1981 and January, 1982.

Recommended by Samuel Kotz, University of Maryland, College Park, MD

Refereed by R. E. Glaser, University of California at Livermore, CA