Not to be cited without prior reference to the authors

Working Paper No 8/86

BARTLETT'S TEST

by

Aileen M Shanks and Julia C Hutton

June 1986

Department of Agriculture and Fisheries for Scotland
Marine Laboratory
Victoria Road
Aberdeen

This working paper is one of a series on statistical procedures which will be issued from time to time. Most, though not all, will contain details of associated computer programs implemented on the Marine Laboratory's PDP 11/70 computer and will include instructions to users on how to access and run the programs and interpret the output.

Earlier papers in the series are:-

Robust Estimators of Location. Smith G.L., Michie C. and Pope J.A. Working Paper No 77/17.

The Poisson and Negative Binomial Distribution. Michie C. Working Paper No 78/3.

Robust Regression Methods. Smith G.L., Michie C. and Pope J.A. Working Paper No 79/1.

Multivariate Analysis of Variance. Shanks A.M. Working Paper No 80/22.

HIST Histogram Computer Program. Emslie D.C. Working Paper No 6/84.

Summary Statistics. Shanks A.M. Working Paper No 10/84.

Principle Component Analysis. Pope J.A. and Shanks A.M. Working Paper No 2/85.

Stem-and-Leaf Displays, Box-and-Whisker Plots and Rootograms. Pope J. A. and Shanks A.M. Working Paper No 4/85.

Contents

BARTLETT'S TEST

by

Aileen M Shanks and Julia C Hutton

Introduction

There are many instances when it is desirable to test the homogeneity of a set of variance estimates. Many statistical techniques rely upon an underlying assumption of equal variance. If variability within groups is not uniform then simple comparisons of the groups in other respects, say means, may be misleading. It is, therefore, important to be able to test whether or not the variance in different groups or classes is constant.

If $x_1, x_2, \ldots, x_n$ is a random sample of n independent observations from a population with true variance σ^2, the latter quantity may be estimated using the estimator

$$s^2 = \sum_{i=1}^{n} (x_i-\bar{x})^2 / (n-1)$$

where $\bar{x}$ is the ordinary sample mean. This estimator is unbiased whatever form the distribution of the individual observations takes. If this distribution is Normal, that of the values of s^2 in repeated samples of size n is completely known. This distribution is a chi-squared distribution with (n-1) degrees of freedom.

If $x_1, x_2, \ldots, x_n$ and $y_1, y_2, \ldots, y_m$ are random samples of size n and m respectively taken from two Normal distributions with variances σ_x^2 and σ_y^2, an exact test of the hypothesis $\sigma_x^2 = \sigma_y^2$ is provided by computing the variance ratio s_x^2/s_y^2. When $\sigma_x^2 = \sigma_y^2 = \sigma^2$, say, the probability distribution of the ratio of the two sample estimates of σ^2, namely s_x^2 and s_y^2, does not depend on σ^2. It does, however, depend on the degrees of freedom (n-1) and (m-1). This distribution is commonly referred to as the F-distribution. Tables of percentiles of the F-distribution for various combinations of (n-1) and (m-1) are widely available.

As an example, suppose two samples give estimates of 6.53 based on 10 degrees of freedom and 1.97 based on 15 degrees of freedom. The hypothesis that these samples come from (Normal) distributions with the same variance is to be tested at (say) the 5% level. The ratio of the sample estimates of variance is first computed. However, two ratios may be calculated, namely 6.53/1.97 = 3.31 and 1.97/6.53 = 0.30. As far as the hypothesis being tested is concerned, we shall want to reject it if an observed variance ratio is "significantly large" or "significantly small". If the former ratio, 3.31 is considered, it follows an F-distribution with 10 and 15 degrees of freedom, the first of these always referring to the degrees of freedom associated with the numerator of the ratio. The upper $2\frac{1}{2}$% value of such a distribution is 3.06

which the observed ratio of 3.31 exceeds. The ratio 1.97/6.53 follows an F-distribution with 15 and 10 degrees of freedom and we require to compare it with the lower 2½% point of such a distribution. Now it may be shown that the percentile points of this F-distribution are related in a very simple way to those of the F-distribution with 10 and 15 degrees of freedom. In particular, the lower 2½% point of the F-distribution with 15 and 10 degrees of freedom is equal to the inverse of the upper 2½% point of the F-distribution with 10 and 15 degrees of freedom. Hence, the hypothesis of equality of variances will be rejected at the 2½% level if the observed ratio 1.97/6.53 is less than the inverse of the upper 2½% point of the F-distribution with 10 and 15 degrees of freedom, ie if $1.97/6.53 < 1/3.06$. This is the same as rejecting the hypothesis if $6.53/1.97 > 3.06$. Hence, the hypothesis of equality may be tested at the 5% level by always taking whichever of the estimates is the larger as the numerator and comparing the observed ratio with the upper 2½% point of the appropriate F-distribution. In the present example, the required ratio is $6.53/1.97 = 3.31$ which exceeds the upper 2½% point (3.06) of the F-distribution with 10 and 15 degrees of freedom. Hence the hypothesis of equality of variances is rejected at the 5% level.

Tests for comparing more than two estimates of variance have been proposed by Bartlett (1937), Cochran (1941) and Hartley (1950). Bartlett's test is one of homogeneity of a set of variances and is the most commonly required test. It is described more fully below. Cochran proposed a test of the significance of the amount by which the largest of a set of variances exceeds the others. Tables of significance points for this test are not widely available, probably because of the specific nature of the hypothesis being tested. The test proposed by Hartley is in the nature of a quick, easy to compute, approximate test. It consists of selecting the largest and the smallest estimates, $s^2(\max)$ and $s^2(\min)$, and referring the ratio $s^2(\max)/s^2(\min)$ to specially constructed tables of significance values. Again, tables of these values are not widely available and the prime motivation for using the test, namely to ease the burden of calculation involved in computing Barlett's test, no longer holds with the widespread use of computers.

Bartlett's Test

M S Bartlett (1937) proposed a test of the homogeneity of a set of variance estimates as follows. Let $s_1^2, s_2^2, \ldots, s_k^2$ be k variance estimates distributed independently as $\chi^2 \sigma^2/n_i$ for $n_1, n_2, \ldots, n_k$ degrees of freedom, respectively, and s^2 be the pooled variance of the k samples, namely

$$s^2 = \Sigma n_i s_i^2/n$$

where $n = \Sigma n_i$

Bartlett's test calls for the evaluation of B where

$$B = (n \ln s^2 - \Sigma n_i \ln s_i^2)/C$$

where

$$C = 1 + (\Sigma n_i^{-1} - n^{-1})/3(k - 1)$$

B is approximately distributed as χ^2 with k-1 degrees of freedom. The following examples illustrate the computation of Bartlett's test.

Example 1

The following variance estimates were computed for the amount of thyroxine found in samples of ten smolting juvenile salmon on nine occasions in spring:-

11.58
15.47
13.81
10.22
9.41
15.56
11.75
8.00
13.72

For the above variances χ^2 = 1.76 with 8 degrees of freedom which is not significant. The nine variance estimates may be accepted as being homogeneous therefore. Exhibits 1a and 1c show the input file constructed for these data and the output file produced from a computer run of these data.

Example 2

In an intercalibration exercise 14 laboratories submitted results of their determinations of the copper content of a fish flour. The variance estimates and their corresponding degrees of freedom are as follows:-

Laboratory	Variance Estimates	Degrees of Freedom
1	15.5	5
2	20.8	5
3	6.6	5
4	59.8	4
5	18.3	5
6	25.9	7
7	12.3	5
8	19.8	5
9	37.8	11
10	31.4	5
11	223.9	5
12	14.2	9
13	17.1	5
14	27.7	5

These data give a χ^2 value of 30.28 with 13 degrees of freedom which exceeds the 1% value (27.688) of χ^2 with 13 degrees of freedom, hence χ^2 is significant and the variance estimates are not homogeneous. It is clear that the variance for laboratory 11 is very large, being four times that of laboratory 4 which has the next highest variance. When the test is recalculated excluding laboratory 11, it is found that the χ^2 value for the remaining laboratories (9.07) does not exceed the 5% value (21.026) of χ^2 with 12 degrees of freedom. Hence there is no evidence to reject the hypothesis that the variance estimates for the remaining laboratories are homogeneous. Exhibits 2a and 2c show the input file for these data and the output file of the results from computing Bartlett's test.

It has been established that Bartlett's test is not robust to departures from Normality in the distribution being compared. The Type I error depends very much on the population kurtosis, the difference between the true percentage and the nominal value of 5% increasing markedly with the number of samples being compared.

Computer Program

A computer program BRTLT has been written in FORTRAN-77 to run on the Marine Laboratory's PDP 11/70 to produce a Bartlett's test, as described in the previous section of this working paper.

The user must prepare and verify an input data file prior to running the program as there are no data entry or checking facilities in BRTLT. The user has the option of recalculating the test excluding one or more variance estimates. The program caters for up to 50 variances.

Data Input Instructions

An input file should be created as follows:-

Number of variances
1st variance, degrees of freedom
2nd variance, degrees of freedom
"
"
"
Last variance, degrees of freedom

Exhibits 1a and 2a show input data files suitable for BRTLT.

Program Running Instructions

To run the program type RUN SP:BRTLT. The user is asked to give the name of the input file and the name of a file for results. (Note the file for results must not already exist or an error message will occur.) The user is then requested to supply a heading for the results file. A successful run of the program terminates with BRTLT -- STOP being printed on the terminal before $ is returned, otherwise an error indicator is printed at the terminal. If an error code is printed, the user should consult the Statistics Section.

The file of results which is suitably formatted for output on A4 paper may then be printed.

Exhibits 1b and 2b show examples of the running of the program, ie all information printed at the user's terminal. Exhibits 1c and 2c show what is written to the file of results. The user must compare the computed χ^2 value with the appropriate tabulated value to test the significance of the computed value.

REFERENCES

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

Cochran, W.G. (1941) The distribution of the largest of a set of estimated variances as a fraction of their total. Ann. Eugen., 11, 47-52.

Hartley, H.O. (1950) The maximum F-ratio as a short-cut test for heterogeneity of variances. Biometrika, 37, 308-312.

EXHIBIT 1a

An input data file (EXAMl.BRT) suitable for BRTLT (the data are those given in the text as example 1)

```
9
11.58,9
15.47,9
13.81,9
10.22,9
9.41,9
15.56,9
11.75,9
8.00,9
13.72,9
```

EXHIBIT 1b

A run of BRTLT using the input data file EXAM.1.BRT

```
RUN SP:BRTLT

Enter name of input data file.
EXAM1.BRT
Enter name of file for results.
EXAM1.OUT
Enter heading for output, up to 50 characters.
Start with a space.
 THYROXINE LEVELS IN SMOLTING JUVENILE SALMON
      DATA SET    VARIANCE    DEGREES OF FREEDOM

          1         11.5800              9
          2         15.4700              9
          3         13.8100              9
          4         10.2200              9
          5          9.4100              9
          6         15.5600              9
          7         11.7500              9
          8          8.0000              9
          9         13.7200              9

B = 1.8306    C =     1.0412
CHISQUARED(B/C) =  1.76    D.F. =  8


Do you wish to exclude any data set(s)  (Y or N)?
N

BRTLT -- STOP
```

EXHIBIT 1c

Print-up of output file EXAM1.OUT created from a run of BRTLT shown in Exhibit 1b

PROGRAM BRTLT   Date of run 09-APR-86

THYROXINE LEVELS IN SMOLTING JUVENILE SALMON

Data from input file EXAM1.BRT

DATA SET	VARIANCE	DEGREES OF FREEDOM
1	11.5800	9
2	15.4700	9
3	13.8100	9
4	10.2200	9
5	9.4100	9
6	15.5600	9
7	11.7500	9
8	8.0000	9
9	13.7200	9

B = 1.8306   C =   1.0412
CHISQUARED(B/C) = 1.76   D.F. = 8

EXHIBIT 2a

An input data file (EXAM2.BRT) suitable for BRTLT (the data are those given in the text as example 2)

```
14
15.5,5
20.8,5
6.6,5
59.8,4
18.3,5
25.9,7
12.3,5
19.8,5
37.8,11
31.4,5
223.9,5
14.2,9
17.1,5
27.7,5
```

EXHIBIT 2b

A run of BRTLT using the input data file EXAM2.BRT

```
RUN SP:BRTLT


Enter name of input data file.
EXAM2.BRT
Enter name of file for results.
EXAM2.OUT
Enter heading for output, up to 50 characters.
Start with a space.
 COPPER CONTENT OF FISH FLOUR
```

DATA SET	VARIANCE	DEGREES OF FREEDOM
1	15.5000	5
2	20.8000	5
3	6.6000	5
4	59.8000	4
5	18.3000	5
6	25.9000	7
7	12.3000	5
8	19.8000	5
9	37.8000	11
10	31.4000	5
11	223.9000	5
12	14.2000	9
13	17.1000	5
14	27.7000	5

```
B =  32.2873   C =   1.0662
CHISQUARED(B/C) = 30.28   D.F. = 13

Do you wish to exclude any data set(s)  (Y or N)?
Y
How many data sets do you wish to exclude?
1
Enter 1st data set number.
11


B =  9.6735   C =  1.0662
CHISQUARED(B/C) =  9.07   D.F. = 12


Do you wish to exclude any data set(s)  (Y or N)?
N
BRTLT -- STOP
```

EXHIBIT 2c

Print-up of output file EXAM2.OUT created from a run of BRTLT shown in Exhibit 2b

PROGRAM BRTLT   Date of run 09-APR-86

COPPER CONTENT OF FISH FLOUR

Data from input file EXAM2.BRT

DATA SET	VARIANCE	DEGREES OF FREEDOM
1	15.5000	5
2	20.8000	5
3	6.6000	5
4	59.8000	4
5	18.3000	5
6	25.9000	7
7	12.3000	5
8	19.8000	5
9	37.8000	11
10	31.4000	5
11	223.9000	5
12	14.2000	9
13	17.1000	5
14	27.7000	5

B = 32.2873   C = 1.0662
CHISQUARED(B/C) = 30.28   D.F. = 13

Bartlett's test recalculated.
The following data sets are excluded:

11

B = 9.6735   C = 1.0662
CHISQUARED(B/C) = 9.07   D.F. = 12