American Review of Mathematics and Statistics
June 2016, Vol. 4, No. 1, pp. 11-16
ISSN: 2374-2348 (Print), 2374-2356 (Online)

Published by American Research Institute for Policy Development
DOI: 10.15640/arms.v4n1a2
URL: https://doi.org/10.15640/arms.v4n1a2

An Alternate Test for Variances

Gerry La Bute[1]

Abstract

Current tests for variances have various limitations. Those such as the chi-square test for one variance or the F test for two variances require the assumption of normality. The Hartley test for multiple variances requires equal sample sizes as well as the assumption of normality. Levene's test has the limitation that if the variance of one group increases relative to the variance of the other groups, the variation within groups can increase at a greater rate than the variation between groups resulting in the ironic situation of a smaller test statistic. The purpose of this paper is to present a nonparametric test for variances that can be used for any number of populations without the restriction of normality or equal sample sizes.

Development of Test Statistic

For a value x from population i, Levene's Test [4] proposes $Z = |x - \bar{x}_i|$ and conducting one-way ANOVA on the subsequent Z values.

As a variation on the above transformation, I propose

$$(x - \bar{x}_i)^2$$

The sample mean of this quantity is

$$\frac{\sum(x - \bar{x}_i)^2}{n} = \frac{n-1}{n} s_i^2$$

If we subtract the sample mean from this quantity and divide by s_i, the result is

$$Q = \frac{(x - \bar{x}_i)^2 - \frac{n-1}{n} s_i^2}{s_i} = \frac{(x - \bar{x}_i)^2}{s_i} - \frac{n-1}{n} s_i$$

If a vector x is multiplied by a constant c, then $\bar{x}_i$ and s_i are also multiplied by c :

$$\frac{(cx - c\bar{x}_i)^2}{cs_i} - \frac{n-1}{n} cs_i$$

[1] The Statistician Ltd., 39 Falmere Way NE, Calgary, AB, Canada, T3J 2Z2.

$$= \frac{c^2(x-\bar{x}_i)^2}{cs_i} - \frac{n-1}{n}cs_i$$
$$= \frac{c(x-\bar{x}_i)^2}{s_i} - \frac{n-1}{n}cs_i$$
$$= c\left(\frac{(x-\bar{x}_i)^2}{s_i} - \frac{n-1}{n}s_i\right) = cQ$$

Thus, Q is also multiplied by c. Another property of Q is that

$$\lim_{s_i \to 0} |Q| = 0$$

and

$$\lim_{s_i \to \infty} |Q| = \infty$$

This property of Q allows us to use standard nonparametric ranking tests to test hypotheses of one or more variances:

One variance: Wilcoxon signed ranks test [6]

Two variances: Wilcoxon rank sum test

Three or more variances: Kruskal-Wallis test [3]

The requirement for two or more variances is that the samples are independent. The minimum scale for all tests is ordinal. Each sample size needs to be at least 3.

One Variance

Given the hypothesized standard deviationσ_o, we modify Q as

$$Q = \frac{(x-\bar{x})^2}{\sigma_o} - \frac{n-1}{n}\sigma_o$$

If $Q > 0$ then

$$\frac{(x-\bar{x})^2}{\sigma_o} > \frac{n-1}{n}\sigma_o$$
$$(x-\bar{x})^2 > \frac{n-1}{n}\sigma_o^2$$

From this

$$x < \bar{x} - \sqrt{\frac{n-1}{n}}\sigma_o$$

or

$$x > \bar{x} + \sqrt{\frac{n-1}{n}}\sigma_o$$

Given the alternative hypothesis $\sigma > \sigma_o$, as $\sigma_o \to \infty$, the range for x decreases, signifying an increase in T_- and a decrease in T_+, making it less likely to reject the null hypothesis. Conversely, as $\sigma_o \to 0$, the range for x increases, signifying a decrease in T_- and an increase in T_+, making it more likely to reject the null hypothesis.

If $Q < 0$ then

$$\frac{(x-\bar{x}_i)^2}{\sigma_o} < \frac{n-1}{n}\sigma_o$$

$$(x-\bar{x}_i)^2 < \frac{n-1}{n}\sigma_o^2$$

From this

$$\bar{x} - \sqrt{\frac{n-1}{n}}\sigma_o < x < \bar{x} + \sqrt{\frac{n-1}{n}}\sigma_o$$

Given the alternative hypothesis $\sigma < \sigma_o$, as $\sigma_o \to 0$, the range for x decreases, signifying a decrease in T_- and a increase in T_+, making it less likely to reject the null hypothesis. Conversely, as $\sigma_o \to \infty$, the range for x increases, signifying an increase in T_- and a decrease in T_+, making it more likely to reject the null hypothesis.

Example 1

Given the data set $X = \{4,5,5,5,6,6,9,10\}$, $\bar{x} = 6.25$ and $s = 2.1213$.

Test the hypothesis σ = 1.5 versus σ > 1.5 at a 5% level of significance.

We reject the null hypothesis if $T_- \leq 6$

The following table summarizes the results:

X	Q	Rank Q
4	2.0625	6
5	-0.27083	2
5	-0.27083	2
5	-0.27083	2
6	-1.27083	4.5
6	-1.27083	4.5
9	3.729167	7
10	8.0625	8

$T_- = 15$. Do not reject the null hypothesis and conclude σ is not greater than 1.5 at a 5% level of significance.

A normality test using Lilliefors [5] indicates that assumptions of normality are not satisfied.

Test the hypothesis σ = 0.5 versus σ > 0.5 at a 5% level of significance.

The following table summarizes the results:

X	Q	Rank Q
4	9.6875	6
5	2.6875	4
5	2.6875	4
5	2.6875	4
6	-0.3125	1.5
6	-0.3125	1.5
9	14.6875	7
10	27.6875	8

$T_- = 3$. Reject the null hypothesis and conclude σ is greater than 0.5 at a 5% level of significance.

Two variances

We modify Q for each sample as:

$$Q = \left|\frac{(x - \bar{x}_i)^2}{s_i} - \frac{n-1}{n}s_i\right|$$

Example 2

Given these two data sets:

Group 1	102.7	99.9	104.4	98.7	99	99.5	98.6	99.4
Group 2	103.4	112.6	100.1	111.7	108	100.1	80	98.1

Their respective means are 100.275 and 101.75 and their respective variances are 4.4736 and 107.1914. A proposed hypothesis $\sigma_1 = \sigma_2$ versus $\sigma_1 < \sigma_2$ at a 5% level of significance.

We reject the null hypothesis if $T_1 \leq 52$ where T_1 represents the rank sum of Group 1. The following table summarizes the results:

Group 1	Group 2	Q1	Q2	RQ1	RQ2
102.7	103.4	0.929633	8.796204	4	14
99.9	112.6	1.78421	2.311334	8	9
104.4	100.1	6.194205	8.796204	11	14
98.7	111.7	0.67787	0.50322	3	1
99	108	1.082109	5.286222	5	10
99.5	100.1	1.566724	8.796204	7	14
98.6	80	0.524211	36.63266	2	16
99.4	98.1	1.488713	7.772379	6	12

The rank sum of Group 1 is 46. Reject the null hypothesis and conclude $\sigma_1 < \sigma_2$. The p-value for the test is 1.03%. Conducting the above test using Levene's test, the test statistic is 5.27 with a p-value of 3.8%. Conducting the Anderson-Darling normality test [1] on Group 1, the p-value is 1.6%.

Three variances

The value of Q is computed for each sample in the same manner as for two variances.

Example 3

Given these groups of data:

Group 1	102.7	99.9	104.4	98.7	99	99.5	98.6	99.4
Group 2	103.4	112.6	100.1	111.7	108	100.1	80	98.1
Group 3	103.4	112.6	100.1	101.7	108	100.1	90	98.1

Is there a significant difference in the variances of at least two groups?

Groups 1 and 2 are the same groups used in Example 2. Group 3 is virtually identical to Group 2 with 80 being changed to 90 and 111.7 to 101.7 in order to create a data set with less variation than Group 2. We reject the null hypothesis if the test statistic is greater than 5.991. The following table summarizes the results:

Group 1	Group 2	Group 3	Q1	Q2	Q3	RQ1	RQ2	RQ3
102.7	103.4	103.4	0.929633	8.796204	5.477164	5	20	14
99.9	112.6	112.6	1.78421	2.311334	11.62968	9	10	22
104.4	100.1	100.1	6.194205	8.796204	5.477164	17	20	14
98.7	111.7	101.7	0.67787	0.50322	5.881778	4	2	16
99	108	108	1.082109	5.286222	0.071402	6	12	1
99.5	100.1	100.1	1.566724	8.796204	5.477164	8	20	14
98.6	80	90	0.524211	36.63266	14.65536	3	24	23
99.4	98.1	98.1	1.488713	7.772379	3.90036	7	18	11

The rank sums are 59, 126, and 115 respectively. The value of the test statistic is 6.455 which have a p-value of 3.97%. We reject the null hypothesis and conclude a significant difference in the variances of at least two groups. The Hartley test [2] produces a test statistic of 23.96. Based on an F distribution with 7 and 7 degrees of freedom, the p-value is 0.02%. However, given the p-value of 1.6% on Group 1 from the Anderson-Darling normality test, the results of the Hartley test are suspect.

As seen in Example 2, we know that the variance of Group 1 is significantly less than that of Group 2. Given the alternative hypothesis $\sigma_1 < \sigma_3$, here are the results:

Group 1	Group 3	Q1	Q3	RQ1	RQ3
102.7	103.4	0.929633	5.477164	4	11
99.9	112.6	1.78421	11.62968	8	15
104.4	100.1	6.194205	5.477164	14	11
98.7	101.7	0.67787	5.881778	3	13
99	108	1.082109	0.071402	5	1
99.5	100.1	1.566724	5.477164	7	11
98.6	90	0.524211	14.65536	2	16
99.4	98.1	1.488713	3.90036	6	9

We reject the null hypothesis if $T_1 \leq 52$ where T_1represents the rank sum of Group 1. The rank sum of Group 1 is 49. We reject the null hypothesis and conclude $\sigma_1 < \sigma_3$.

Given the alternative hypothesis $\sigma_2 > \sigma_3$, here are the results:

Group 2	Group 3	Q2	Q3	RQ2	RQ3
103.4	103.4	8.796204	5.443258	12	6
112.6	112.6	2.311334	11.85861	3	14
100.1	100.1	8.796204	5.516848	12	7.5
111.7	101.1	0.50322	5.836482	2	9
108	108	5.286222	0.061883	5	1
100.1	100.1	8.796204	5.516848	12	7.5
80	90	36.63266	14.37852	16	15
98.1	98.1	7.772379	3.985578	10	4

We reject the null hypothesis if $T_1 \geq 84$ where T_1 represents the rank sum of Group 2. The rank sum of Group 2 is 72. We do not reject the null hypothesis and conclude the variance of Group 2 is not significantly greater than that of Group 3.

Bibliography

Anderson, T.W., Darling, D.A. (1952). Asymptotic theory of certain "goodness-of-fit" criteria based on stochastic processes. *Annals of Mathematical Statistics, 23,* 193-212

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

Kruskal, W.H., and Wallis, W.A. (1952). Use of ranks on one-criterion variance analysis. *Journal of the American Statistical Association, 47,* 593-621

Levene, Howard (1960). Robust tests for equality of variances. In Ingram Olkin, Harold Hotelling et al. *Stanford University Press,* 278-292

Lilliefors, H.W. (1967). On the Kolmogorov-Smirnov test for normality with mean and variance unknown. *Journal of the American Statistical Association, 62,* 399-402

Wilcoxon, F (1945). Individual comparisons by ranking methods. *Biometrics, 1,* 80-83