Levene's Test

In statistics, Levene's test is an inferential statistic used to assess the equality of variances in different samples. Some common statistical procedures assume that variances of the populations from which different samples are drawn are equal. Levene's test assesses this assumption. It tests the null hypothesis that the population variances are equal (called homogeneity of variance). If the resulting p-value of Levene's test is less than some critical value (typically 0.05), the obtained differences in sample variances are unlikely to have occurred based on random sampling. Thus, the null hypothesis of equal variances is rejected and it is concluded that there is a difference between the variances in the population. Procedures which typically assume homogeneity of variance include analysis of variance and t-tests. **One advantage of Levene's test is that it does not require normality of the underlying data. Levene's test is often used before a comparison of means.** When Levene's test is significant, modified procedures are used that do not assume equality of variance. Levene's test may also test a meaningful question in its own right if a researcher is interested in knowing whether population group variances are different. This test employs a transformation of the original data values to difference values around each mean.

Assumptions:

- Observed values $X_{1,1}$, $X_{1,2}$, $X_{1,3}$, ... $X_{1,n1}$ are a random sample from
- Observed values $X_{2,1}$, $X_{2,2}$, $X_{2,3}$, ... $X_{2,n2}$ are a random sample from
- Samples from the two samples are *independent.*

Hypotheses:

For a two sided case:

H_0: σ1^2 = σ2^2 … Meaning that there is no difference in variance between populations X1. & X2.
H_1: σ1^2≠σ2^2…Meaning that there is a difference in variance between populations X1 & X2

Decision Rule:

If $P<\alpha$, then reject H_0, otherwise, accept H_0
The p value can be ascertained from the output in R. The alpha value must be explicitly set, and is usually set to 0.05 or 0.01, depending on if the test is "minimally important" or "critically important," respectively.

The package that contains Levene's Test must be downloaded and installed before running the test.

1) Under the tab *packages*, select *install packages*
2) Select a CRAN mirror
3) Select package *car* and install it
4) **library(car)**
5) **?levene.test**

Example:

> Loblolly

	height	age	Seed
1	4.51	3	301
15	10.89	5	301
29	28.72	10	301
43	41.74	15	301
57	52.70	20	301
71	60.92	25	301
2	4.55	3	303
16	10.92	5	303
30	29.07	10	303
44	42.83	15	303
58	53.88	20	303
72	63.39	25	303
3	4.79	3	305
17	11.37	5	305
31	30.21	10	305
45	44.40	15	305
59	55.82	20	305
73	64.10	25	305
4	3.91	3	307
18	9.48	5	307
32	25.66	10	307
46	39.07	15	307
60	50.78	20	307
74	59.07	25	307
5	4.81	3	309
19	11.20	5	309
33	28.66	10	309
47	41.66	15	309
61	53.31	20	309
75	63.05	25	309
6	3.88	3	311
20	9.40	5	311
34	25.99	10	311
48	39.55	15	311
62	51.46	20	311
76	59.64	25	311
7	4.32	3	315
21	10.43	5	315
35	27.16	10	315
49	40.85	15	315
63	51.33	20	315
77	60.07	25	315
8	4.57	3	319
22	10.57	5	319
36	27.90	10	319
50	41.13	15	319
64	52.43	20	319
78	60.69	25	319
9	3.77	3	321
23	9.03	5	321
37	25.45	10	321
51	38.98	15	321
65	49.76	20	321
79	60.28	25	321
10	4.33	3	323
24	10.79	5	323
38	28.97	10	323
52	42.44	15	323
66	53.17	20	323
80	61.62	25	323
11	4.38	3	325
25	10.48	5	325
39	27.93	10	325
53	40.20	15	325
67	50.06	20	325
81	58.49	25	325
12	4.12	3	327
26	9.92	5	327
40	26.54	10	327
54	37.82	15	327
68	48.43	20	327
82	56.81	25	327
13	3.93	3	329
27	9.34	5	329
41	26.08	10	329
55	37.79	15	329
69	48.31	20	329
83	56.43	25	329
14	3.46	3	331
28	9.05	5	331
42	25.85	10	331
56	39.15	15	331
70	49.12	20	331
84	59.49	25	331

As an example, the dataset *Loblolly* (as presented) from R is used. The dataset uses samples from 84 Loblolly pine trees and lists their height, age and seed. The first column ranks each according to their height in increasing order.

To fit the format of the Levene Test, we must alter the data to fit what exactly we are trying to find, is there a variance between the age of the seed and the height it grows. To do this we categorize the ages of the seeds into *young* and *old*. Any seed ≤ 10 is deemed *young* and any seed > 10 is deemed *old.* Our information now looks like this:

height	age
4.51	Young
10.89	Young
28.72	Young
41.74	Old
52.7	Old
60.92	Old
4.55	Young
10.92	Young
29.07	Young
42.83	Old
53.88	Old
63.39	Old
4.79	Young
11.37	Young
30.21	Young
44.4	Old
55.82	Old
64.1	Old
3.91	Young
9.48	Young
25.66	Young
39.07	Old
50.78	Old
59.07	Old
4.81	Young
11.2	Young
28.66	Young
41.66	Old
53.31	Old
63.05	Old
3.88	Young
9.4	Young
25.99	Young
39.55	Old
51.46	Old
59.64	Old
4.32	Young
10.43	Young
27.16	Young
40.85	Old
51.33	Old
60.07	Old
4.57	Young
10.57	Young
27.9	Young
41.13	Old
52.43	Old
60.69	Old
3.77	Young
9.03	Young
25.45	Young
38.98	Old
49.76	Old
60.28	Old
4.33	Young
10.79	Young
28.97	Young
42.44	Old
53.17	Old
61.62	Old
4.38	Young
10.48	Young
27.93	Young
40.2	Old
50.06	Old
58.49	Old
4.12	Young
9.92	Young
26.54	Young
37.82	Old
48.43	Old
56.81	Old
3.93	Young
9.34	Young
26.08	Young
37.79	Old
48.31	Old
56.43	Old
3.46	Young
9.05	Young
25.85	Young
39.15	Old
49.12	Old
59.49	Old

It is now possible to run Levene's Test. To run the test in R, the program R must be able to read the data, so it is imperative to input the data into R.

```
> LT=read.table("c:/RData/LeveneTest.txt")
> LT
> attach(LT)
```

We can now run the test.

```
> leveneTest(height,age)
Levene's Test for Homogeneity of Variance (center = median)
      Df F value Pr(>F)
group  1  0.3892 0.5345
      82
```

We use the median as the center because it provides a more robust test. We have an alpha value of 0.05, and our p-value was found to be 0.5345. Based on decision rules, we can accept the Null Hypothesis that states that the population variances are equal.

-Ryan Dean Joe Crifo