

## 12.8: Checking the Homogeneity of Variance Assumption

There's more than one way to skin a cat, as the saying goes, and more than one way to test the homogeneity of variance assumption, too (though for some reason no-one made a saying out of that). The most commonly used test for this that I've seen in the literature is the **Levene test** (Levene 1960), and the closely related **Brown-Forsythe test** (Brown and Forsythe 1974), both of which I'll describe here. Alternatively, you could use the Bartlett test, which is implemented in R via the `bartlett.test()` function, but I'll leave it as an exercise for the reader to go check that one out if you're interested.

Levene's test is shockingly simple. Suppose we have our outcome variable  $Y_{ik}$ . All we do is define a new variable, which I'll call  $Z_{ik}$ , corresponding to the absolute deviation from the group mean:

$$Z_{ik} = |Y_{ik} - \bar{Y}_k|$$

Okay, what good does this do us? Well, let's take a moment to think about what  $Z_{ik}$  actually is, and what we're trying to test. The value of  $Z_{ik}$  is a measure of how the  $i$ -th observation in the  $k$ -th group deviates from its group mean. And our null hypothesis is that all groups have the same variance; that is, the same overall deviations from the group means! So, the null hypothesis in a Levene's test is that the population means of  $Z$  are identical for all groups. Hm. So what we need now is a statistical test of the null hypothesis that all group means are identical. Where have we seen that before? Oh right, that's what ANOVA is... and so all that the Levene's test does is run an ANOVA on the new variable  $Z_{ik}$ .

What about the Brown-Forsythe test? Does that do anything particularly different? Nope. The only change from the Levene's test is that it constructs the transformed variable  $Z$  in a slightly different way, using deviations from the group *medians* rather than deviations from the group *means*. That is, for the Brown-Forsythe test,

$$Z_{ik} = |Y_{ik} - \text{median}_k(Y)|$$

where  $\text{median}_k(Y)$  is the median for group  $k$ . Regardless of whether you're doing the standard Levene test or the Brown-Forsythe test, the test statistic – which is sometimes denoted  $F$ , but sometimes written as  $W$  – is calculated in exactly the same way that the  $F$ -statistic for the regular ANOVA is calculated, just using a  $Z_{ik}$  rather than  $Y_{ik}$ . With that in mind, let's just move on and look at how to run the test in R.

### 12.8.1 Running the Levene's test in R

Okay, so how do we run the Levene test? Obviously, since the Levene test is just an ANOVA, it would be easy enough to manually create the transformed variable  $Z_{ik}$  and then use the `aov()` function to run an ANOVA on that. However, that's the tedious way to do it. A better way to do run your Levene's test is to use the `leveneTest()` function, which is in the `car` package. As usual, we first load the package

```
library( car )
```

```
## Loading required package: carData
```

and now that we have, we can run our Levene test. The main argument that you need to specify is `y`, but you can do this in lots of different ways. Probably the simplest way to do it is actually input the original `aov` object. Since I've got the `my.anova` variable stored from my original ANOVA, I can just do this:

```
leveneTest( my.anova )
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##           Df F value Pr(>F)
## group    2  1.4672 0.2618
##           15
```

If we look at the output, we see that the test is non-significant ( $F_{2,15}=1.47, p=.26$ ), so it looks like the homogeneity of variance assumption is fine. Remember, although R reports the test statistic as an  $F$ -value, it could equally be called  $W$ , in which case you'd

just write  $W_{2,15}=1.47$ . Also, note the part of the output that says `center = median`. That's telling you that, by default, the `leveneTest()` function actually does the Brown-Forsythe test. If you want to use the mean instead, then you need to explicitly set the `center` argument, like this:

```
leveneTest( y = my.anova, center = mean )
```

```
## Levene's Test for Homogeneity of Variance (center = mean)
##      Df F value Pr(>F)
## group  2  1.4497 0.2657
##      15
```

That being said, in most cases it's probably best to stick to the default value, since the Brown-Forsythe test is a bit more robust than the original Levene test.

### 12.8.2 Additional comments

Two more quick comments before I move onto a different topic. Firstly, as mentioned above, there are other ways of calling the `leveneTest()` function. Although the vast majority of situations that call for a Levene test involve checking the assumptions of an ANOVA (in which case you probably have a variable like `my.anova` lying around), sometimes you might find yourself wanting to specify the variables directly. Two different ways that you can do this are shown below:

```
leveneTest(y = mood.gain ~ drug, data = clin.trial) # y is a formula in this case
leveneTest(y = clin.trial$mood.gain, group = clin.trial$drug) # y is the outcome
```

Secondly, I did mention that it's possible to run a Levene test just using the `aov()` function. I don't want to waste a lot of space on this, but just in case some readers are interested in seeing how this is done, here's the code that creates the new variables and runs an ANOVA. If you are interested, feel free to run this to verify that it produces the same answers as the Levene test (i.e., with `center = mean`):

```
Y <- clin.trial $ mood.gain # the original outcome variable, Y
G <- clin.trial $ drug      # the grouping variable, G
gp.mean <- tapply(Y, G, mean) # calculate group means
Ybar <- gp.mean[G]          # group mean associated with each obs
Z <- abs(Y - Ybar)          # the transformed variable, Z
summary( aov(Z ~ G) )      # run the ANOVA
```

```
##      Df Sum Sq Mean Sq F value Pr(>F)
## G      2  0.0616  0.03080    1.45  0.266
## Residuals 15  0.3187  0.02125
```

That said, I don't imagine that many people will care about this. Nevertheless, it's nice to know that you could do it this way if you wanted to. And for those of you who do try it, I think it helps to demystify the test a little bit when you can see – with your own eyes – the way in which Levene's test relates to ANOVA.

This page titled [12.8: Checking the Homogeneity of Variance Assumption](#) is shared under a [CC BY-SA 4.0](#) license and was authored, remixed, and/or curated by [Danielle Navarro](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.

- [14.8: Checking the Homogeneity of Variance Assumption](#) by [Danielle Navarro](#) is licensed [CC BY-SA 4.0](#). Original source: <https://bookdown.org/ekothe/navarro26/>.