Chapter 11: Comparing several means

Oliver Twisted

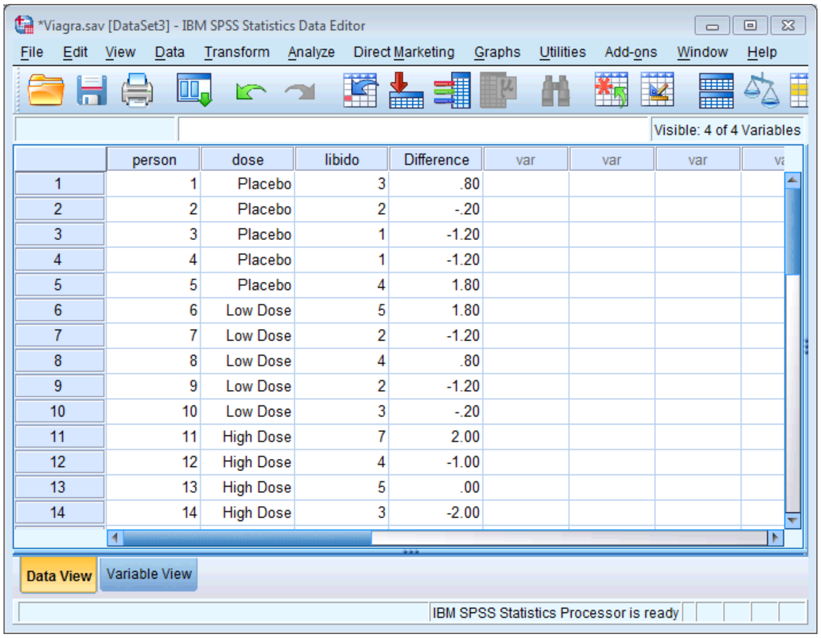
Please, Sir, can I have some more … Levene's test?

Levene's test is basically an ANOVA conducted on the absolute differences between the observed data and the mean from which the data came. To see what I mean, let's do a sort of manual Levene's test on the Viagra data. First we need to create a new variable called **difference** (short for 'Difference from group mean'), which is each score subtracted from the mean of the group to which that score belongs. Remember that means for the placebo, low-dose and high-dose groups were 2.2, 3.2 and 5 respectively, and the groups were coded 1, 2 and 3. We can compute this new variable using syntax:

```
IF  (dose = 1) Difference=libido − 2.2.

IF  (dose = 2) Difference=libido − 3.2.

IF  (dose = 3) Difference=libido − 5.

VARIABLE LABELS  Difference 'Difference from Group Mean'.

EXECUTE.
```

The first line just says that if dose = 1 (i.e., placebo) then the difference is the value of libido minus 2.2 (the mean of the placebo group). The next two lines do the same thing for the low- and high-dose group.

The resulting data look like this:



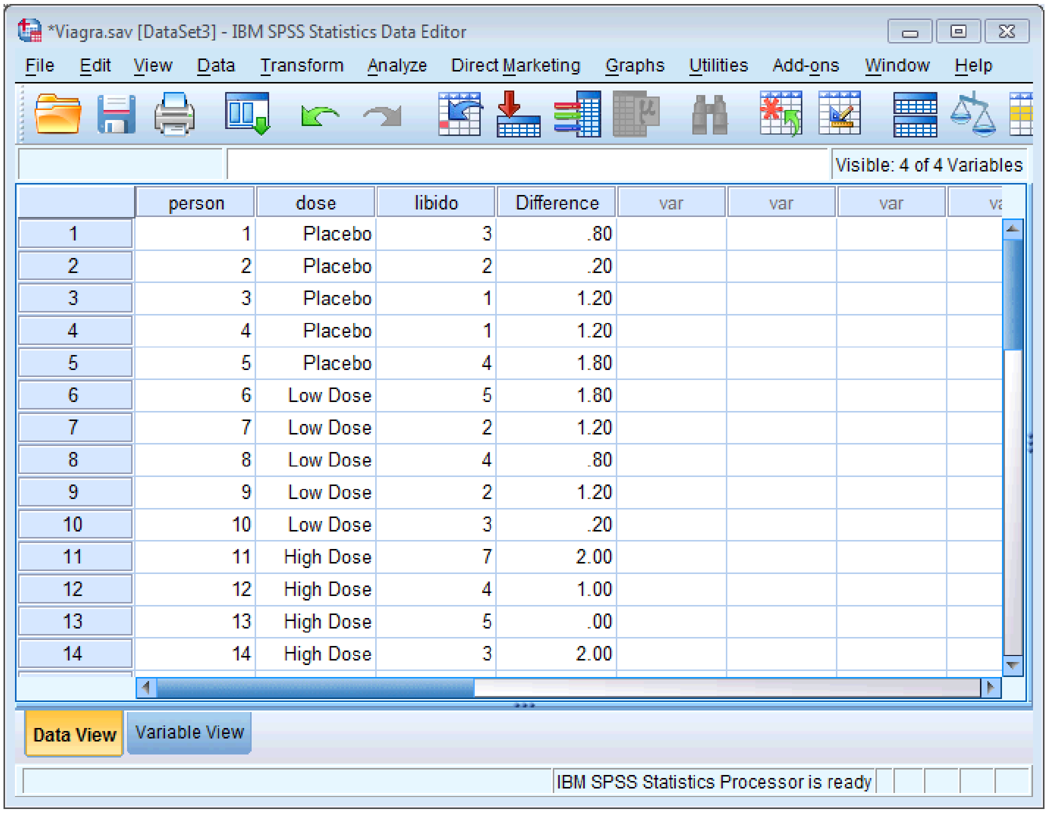
	person	dose	libido	Difference	var	var	var
1	1	Placebo	3	.80			
2	2	Placebo	2	-.20			
3	3	Placebo	1	-1.20			
4	4	Placebo	1	-1.20			
5	5	Placebo	4	1.80			
6	6	Low Dose	5	1.80			
7	7	Low Dose	2	-1.20			
8	8	Low Dose	4	.80			
9	9	Low Dose	2	-1.20			
10	10	Low Dose	3	-.20			
11	11	High Dose	7	2.00			
12	12	High Dose	4	-1.00			
13	13	High Dose	5	.00			
14	14	High Dose	3	-2.00			

Note that for person 1, the difference score is 3 – 2.2 = 0.8, for person 2 it is 2 – 2.2 = −0.20. As we move into the low-dose group we subtract the mean of that group, so for person 6 the difference score is 5 – 3.2 = 1.8, for person 7 it is 2 – 3.2 = −1.20. In the high-dose group, the group mean is 5, so for person 11 we get a difference of 7 – 5 = 2, and so on. Think about what these differences are; they are deviations from the mean, the same deviations that we calculate when we compute the sums of squares, variance and standard deviation. They represent variation from the mean. When we compute the variance we square the values to get rid of the plus and minus signs (otherwise the positive and negative deviations will cancel out). Levene's test doesn't do this (because we don't want to change the units of measurement by squaring the values), but instead simply takes the absolute values; that is, it pretends that all of the deviations are positive.

To get the absolute values of these differences (i.e. we need to make them all positive values), again we can do this with syntax:

```
Compute Difference = abs(Difference).

VARIABLE LABELS  Difference 'Absolute Difference from Group Mean'.

EXECUTE.
```

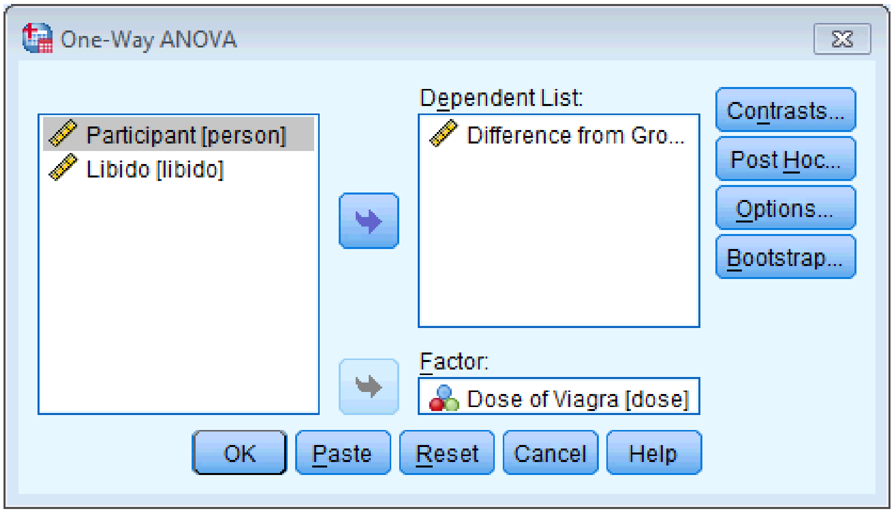
The first line just changes the variable **Difference** to be the absolute value of itself. The second line renames the variable to reflect the fact that it now contains absolute values. The data now look like this:



	person	dose	libido	Difference	var	var	var
1	1	Placebo	3	.80			
2	2	Placebo	2	.20			
3	3	Placebo	1	1.20			
4	4	Placebo	1	1.20			
5	5	Placebo	4	1.80			
6	6	Low Dose	5	1.80			
7	7	Low Dose	2	1.20			
8	8	Low Dose	4	.80			
9	9	Low Dose	2	1.20			
10	10	Low Dose	3	.20			
11	11	High Dose	7	2.00			
12	12	High Dose	4	1.00			
13	13	High Dose	5	.00			
14	14	High Dose	3	2.00			

Note that the difference scores are the same magnitude, it's just that the minus signs have gone. These values still represent deviations from the mean, or variance, but we now don't have the problem of positive and negative deviations cancelling each other out.

Now, using what you learnt in the book, conduct a one-way ANOVA on these difference scores: **dose** is the independent variable and **Difference** is the dependent variable (don't select any special options, just run a basic analysis). The main dialog box should look like this:



You'll find that the *F*-ratio for this analysis is 0.092, which is significant at $p = 0.913$; that is, the same values as Levene's test in the book!

ANOVA

Absolute Difference from Group Mean

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.085	2	.043	.092	.913
Within Groups	5.584	12	.465		
Total	5.669	14			

Levene's test is, therefore, testing whether the 'average' absolute deviation from the mean is the same in the three groups. Clever, eh?

Please, Sir, can I have some more … Welch's *F*?

The Welch (1951) *F*-ratio is somewhat more complicated (hence why it's stuck on the website). First we have to work out a weight that is based on the sample size, n_k, and variance, s_k^2, for a particular group:

$$w_k = \frac{n_k}{s_k^2}$$

We also need to use a grand mean based on a weighted mean for each group. So we take the mean of each group, $\bar{x}_k$, and multiply it by its weight, w_k, do this for each group and add them up, then divide this total by the sum of weights:

$$\bar{x}_{\text{grand}}^{\text{Welch}} = \frac{\sum w_k \bar{x}_k}{\sum w_k}$$

The easiest way to do this is in table form:

Group	**Variance** s^2	**Sample size,** n_k	**Weight** w_k	**Mean** $\bar{x}_k$	$w_k \times \bar{x}_k$
Placebo	1.70	5	2.941	2.2	6.4702
Low Dose	1.70	5	2.941	3.2	9.4112
High Dose	2.50	5	2.000	5.0	10.000
			Σ = 7.882		Σ = 25.8814

So we get:

$$\bar{x}_{\text{grand}}^{\text{Welch}} = \frac{25.8814}{7.882} = 3.284$$

Think back to equation (11.4), the model sum of squares was:

$$SS_{\text{M}} = \sum_{n=1}^{k} n_k \left(\bar{x}_k - \bar{x}_{\text{grand}}\right)^2$$

In Welch's *F* this is adjusted to incorporate the weighting and the adjusted grand mean:

$$\text{SS}_{\text{M}}^{\text{Welch}} = \sum w_k \left(\bar{x}_k - \bar{x}_{\text{grand}}^{\text{Welch}}\right)^2$$

And to create a mean square we divide by the degrees of freedom, $k - 1$:

$$\text{MS}_{\text{M}}^{\text{Welch}} = \frac{\sum w_k \left(\bar{x}_k - \bar{x}_{\text{grand}}^{\text{Welch}}\right)^2}{k - 1}$$

We obtain

$$\text{MS}_{\text{M}}^{\text{Welch}} = \frac{2.941(2.2 - 3.284)^2 + 2.941(3.2 - 3.284)^2 + 2.000(5.0 - 3.284)^2}{2} = 4.683$$

We now have to work out a term called lambda, which is based again on the weights:

$$\Lambda = \frac{3\sum \frac{\left(1 - \frac{w_k}{\sum w_k}\right)^2}{n_k - 1}}{k^2 - 1}$$

This looks horrendous, but is just based on the sample size in each group, the weight for each group (and the sum of all weights), and the total number of groups, *k*. For the Viagra data this gives us:

$$\begin{aligned}\Lambda &= \frac{3\left(\frac{\left(1-\frac{2.941}{7.882}\right)^2}{5-1}+\frac{\left(1-\frac{2.941}{7.882}\right)^2}{5-1}+\frac{\left(1-\frac{2}{7.882}\right)^2}{5-1}\right)}{3^2-1}\\ &= \frac{3(0.098+0.098+0.139)}{8}\\ &= 0.126\end{aligned}$$

The F ratio is then given by:

$$F_W = \frac{\mathrm{MS}_{\mathrm{M}}^{\;(W)}}{1+\frac{2\Lambda(k-2)}{3}}$$

where k is the total number of groups. So, for the Viagra data we get:

$$\begin{aligned}F_W &= \frac{4.683}{1+\frac{(2\times 0.126)(3-2)}{3}}\\ &= \frac{9.336}{1.084}\\ &= 4.32\end{aligned}$$

As with the Brown–Forsythe F, the model degrees of freedom stay the same at $k - 1$ (in this case 2), but the residual degrees of freedom, df_R, are $1/\Lambda$ (in this case, $1/0.126 = 7.94$).