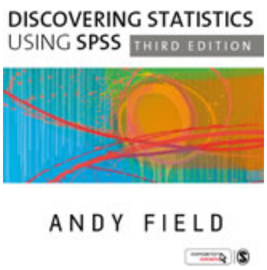
Discovering Statistics Using SPSS (3rd Edition)

Additional material

Professor Andy Field



This file contains additional material for Field, A. P. (2009). *Discovering statistics using SPSS (and sex and drugs and rock 'n' roll)*, 3rd edition. London: Sage.

Welch's F

Oliver Twisted: Please Sir, Can I have Some More … Welch's *F*?

'You're only telling us about the Brown-Forsythe F because you don't understand Welch's F,' taunts Oliver, 'Andy, Andy, brains all sandy ….' Whatever, Oliver. Like the Brown-Forsythe F, Welch's F adjusts F and the residual degrees of freedom to combat problems arising from violations of the homogeneity of variance assumption. There is a lengthy explanation about Welch's F in the additional material available on the companion website. Oh, and Oliver, microchips are made of sand.

The Welch (1951) *F*-ratio is somewhat more complicated (and 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 times it by its weight, w_k, do this for each group and add them up, then divide this total by the sum of weights:

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

© Professor Andy Field, 2005

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:

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

Think back, to equation 8.5, the model sum of squares was:

$$SS_M = \sum n_k \left(\bar{x}_k - \bar{x}_{grand}\right)^2$$

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

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

And to create a mean squares we divide by the degrees of freedom, *k* – 1:

$$\begin{aligned} MS_M^{Welch} &= \frac{\sum w_k \left(\bar{x}_k - \bar{x}_{grand}^{Welch}\right)^2}{K-1} \\ &= \frac{2.941(2.2-3.284)^2 + 2.941(3.2-3.284)^2 + 2(5-3.284)^2}{2} \\ &= 4.683 \end{aligned}$$

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 equation 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:

© Professor Andy Field, 2005

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

The *F* ratio is then given by:

$$F_W = \frac{MS_M^{Welch}}{1+\frac{2\Lambda(K-2)}{3}}$$

Where, *K*, is the total number of groups. So, for the Viagra data we get:

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

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/Λ (in this case, 1/0.126 = 7.94).

© Professor Andy Field, 2005