14.1: Summary

There's a fair bit covered in this chapter, but there's still a lot missing. Most obviously, I haven't yet discussed any analog of the paired samples t-test for more than two groups. There is a way of doing this, known as *repeated measures ANOVA*, which will appear in a later version of this book. I also haven't discussed how to run an ANOVA when you are interested in more than one grouping variable, but that will be discussed in a lot of detail in Chapter 16. In terms of what we have discussed, the key topics were:

- The basic logic behind how ANOVA works (Section 14.2) and how to run one in R (Section 14.3).
- How to compute an effect size for an ANOVA (Section 14.4)
- Post hoc analysis and corrections for multiple testing (Section 14.5).
- The assumptions made by ANOVA (Section 14.6).
- How to check the homogeneity of variance assumption (Section 14.7) and what to do if it is violated (Section 14.8).
- How to check the normality assumption (Section 14.9 and what to do if it is violated (Section 14.10).

As with all of the chapters in this book, there are quite a few different sources that I've relied upon, but the one stand-out text that I've been most heavily influenced by is Sahai and Ageel (2000). It's not a good book for beginners, but it's an excellent book for more advanced readers who are interested in understanding the mathematics behind ANOVA.

References

Hays, W. L. 1994. *Statistics*. 5th ed. Fort Worth, TX: Harcourt Brace.

Shaffer, J. P. 1995. "Multiple Hypothesis Testing." *Annual Review of Psychology* 46: 561–84.

Hsu, J. C. 1996. *Multiple Comparisons: Theory and Methods*. London, UK: Chapman; Hall.

Dunn, O.J. 1961. "Multiple Comparisons Among Means." *Journal of the American Statistical Association* 56: 52–64.

Holm, S. 1979. "A Simple Sequentially Rejective Multiple Test Procedure." *Scandinavian Journal of Statistics* 6: 65–70.

Levene, H. 1960. "Robust Tests for Equality of Variances." In *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling*, edited by I. Olkin et al, 278–92. Palo Alto, CA: Stanford University Press.

Brown, M. B., and A. B. Forsythe. 1974. "Robust Tests for Equality of Variances." *Journal of the American Statistical Association* 69: 364–67.

Welch, B. 1951. "On the Comparison of Several Mean Values: An Alternative Approach." *Biometrika* 38: 330–36.

Kruskal, W. H., and W. A. Wallis. 1952. "Use of Ranks in One-Criterion Variance Analysis." *Journal of the American Statistical Association* 47: 583–621.

Sahai, H., and M. I. Ageel. 2000. *The Analysis of Variance: Fixed, Random and Mixed Models*. Boston: Birkhauser.

---

201. When all groups have the same number of observations, the experimental design is said to be "balanced". Balance isn't such a big deal for one-way ANOVA, which is the topic of this chapter. It becomes more important when you start doing more complicated ANOVAs.
202. In a later versions I'm intending to expand on this. But because I'm writing in a rush, and am already over my deadlines, I'll just briefly note that if you read ahead to Chapter 16 and look at how the "treatment effect" at level k of a factor is defined in terms of the αk values (see Section 16.2), it turns out that Q refers to a weighted mean of the squared treatment effects, $Q=\left(\sum_{k=1}^{G} N_k\alpha_k^2\right)/(G-1)$
203. we want to be sticklers for accuracy, $1+\frac{2}{df_2-2}$
204. o be precise, party like "it's 1899 and we've got no friends and nothing better to do with our time than do some calculations that wouldn't have made any sense in 1899 because ANOVA didn't exist until about the 1920s".
205. Actually, it *also* provides a function called `anova()` , but that works a bit differently, so let's just ignore it for now.
206. It's worth noting that you can get the same result by using the command `anova( my.anova )` .
207. A potentially important footnote – I wrote the `etaSquared()` function for the `lsr` package as a teaching exercise, but like all the other functions in the `lsr` package it hasn't been exhaustively tested. As of this writing – `lsr` package version 0.5 – there is at least one known bug in the code. In some cases at least, it doesn't work (and can give very silly answers) when

you set the `weights` on the observations to something other than uniform. That doesn't matter at all for this book, since those kinds of analyses are well beyond the scope, but I haven't had a chance to revisit the package in a long time; it would probably be wise to be very cautious with the use of this function in any context other than very simple introductory analyses. Thanks to Emil Kirkegaard for finding the bug! (Oh, and while I'm here, there's an interesting blog post by Daniel Lakens suggesting that eta-squared itself is perhaps not the best measure of effect size in real world data analysis: http://daniellakens.blogspot.com.au/2015/06/why-you-should-use-omega-squared.html

208. I should point out that there are other functions in R for running multiple comparisons, and at least one of them works this way: the `TukeyHSD()` function takes an `aov` object as its input, and outputs Tukey's "honestly significant difference" tests. I talk about Tukey's HSD in Chapter 16.
209. If you *do* have some theoretical basis for wanting to investigate some comparisons but not others, it's a different story. In those circumstances you're not really running "post hoc" analyses at all: you're making "planned comparisons". I do talk about this situation later in the book (Section 16.9), but for now I want to keep things simple.
210. It's worth noting in passing that not all adjustment methods try to do this. What I've described here is an approach for controlling "family wise Type I error rate". However, there are other post hoc tests seek to control the "false discovery rate", which is a somewhat different thing.
211. There's also a function called `p.adjust()` in which you can input a vector of raw p-values, and it will output a vector of adjusted p-values. This can be handy sometimes. I should also note that more advanced users may wish to consider using some of the tools provided by the `multcomp` package.
212. Note that neither of these figures has been tidied up at all: if you want to create nicer looking graphs it's always a good idea to use the tools from Chapter 6 to help you draw cleaner looking images.
213. A technical term.