

## 16.11: Summary

- Factorial ANOVA with balanced designs, without interactions (Section 16.1) and with interactions included (Section 16.2)
- Effect size, estimated means, and confidence intervals in a factorial ANOVA (Section 16.3)
- Understanding the linear model underling ANOVA (Sections 16.5, 16.6 and 16.7)
- Post hoc testing using Tukey's HSD (Section 16.8), and a brief commentary on planned comparisons (Section 16.9)
- Factorial ANOVA with unbalanced designs (Section 16.10)

### References

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

230. The R command we need is `xtabs(~ drug+gender, clin.trial)` .
231. The nice thing about the subscript notation is that generalises nicely: if our experiment had involved a third factor, then we could just add a third subscript. In principle, the notation extends to as many factors as you might care to include, but in this book we'll rarely consider analyses involving more than two factors, and never more than three.
232. Technically, marginalising isn't quite identical to a regular mean: it's a weighted average, where you take into account the frequency of the different events that you're averaging over. However, in a balanced design, all of our cell frequencies are equal by definition, so the two are equivalent. We'll discuss unbalanced designs later, and when we do so you'll see that all of our calculations become a real headache. But let's ignore this for now.
233. English translation: "least tedious".
234. This chapter seems to be setting a new record for the number of different things that the letter R can stand for: so far we have R referring to the software package, the number of rows in our table of means, the residuals in the model, and now the correlation coefficient in a regression. Sorry: we clearly don't have enough letters in the alphabet. However, I've tried pretty hard to be clear on which thing R is referring to in each case.
235. Implausibly large, I would think: the artificiality of this data set is really starting to show!
236. In fact, there's a function `Effect()` within the `effects` package that has slightly different arguments, but computes the same things, and won't give you this warning message.
237. Due to the way that the `leveneTest()` function is implemented, however, if you use a formula like `mood.gain ~\ drug + therapy + drug:therapy` , or input an ANOVA object based on a formula like this, you actually get the error message. That shouldn't happen, because this actually is a fully crossed model. However, there's a quirky shortcut in the way that the `leveneTest()` function checks whether your model is fully crossed that means that it doesn't recognise this as a fully crossed model. Essentially what the function is doing is checking that you used `*` (which ensures that the model is fully crossed), and not `+` or `:` in your model formula. So if you've manually typed out all of the relevant terms for a fully crossed model, the `leveneTest()` function doesn't detect it. I think this is a bug.
238. There could be all sorts of reasons for doing this, I would imagine.
239. This is cheating in some respects: because ANOVA and regression are provably the same thing, R is lazy: if you read the help documentation closely, you'll notice that the `aov()` function is actually just the `lm()` function in disguise! But we shan't let such things get in the way of our story, shall we?
240. In the example given above, I've typed `summary( regression.model )` to get the hypothesis tests. However, the `summary()` function does produce a lot of output, which is why I've used the `BLAH BLAH BLAH` text to hide the unnecessary parts of the output. But in fact, you can use the `coef()` function to do the same job. If you the command `coef( summary( regression.model ) )` you'll get exactly the same output that I've shown above (minus the `BLAH BLAH BLAH` ). Compare and contrast this to the output of `coef( regression.model )` .
241. Advanced users may want to look into the `model.matrix()` function, which produces similar output. Alternatively, you can use a command like `contr.treatment(3)[clin.trial\$drug, ]` . I'll talk about the `contr.treatment()` function later.
242. Future versions of this book will try to be a bit more consistent with the naming scheme for variables. One of the many problems with having to write a lengthy text very quickly to meet a teaching deadline is that you lose some internal consistency.
243. The `lsr` package contains a more general function called `permuteLevels()` that can shuffle them in any way you like.
244. Technically, this list actually stores the functions themselves. R allows lists to contain functions, which is really neat for advanced purposes, but not something that matters for this book.

245. If, for instance, you actually would find yourself interested to know if Group A is significantly different from the mean of Group B and Group C, then you need to use a different tool (e.g., Scheffe's method, which is more conservative, and beyond the scope this book). However, in most cases you probably are interested in pairwise group differences, so Tukey's HSD is a pretty useful thing to know about.
246. This discrepancy in standard deviations might (and should) make you wonder if we have a violation of the homogeneity of variance assumption. I'll leave it as an exercise for the reader to check this using the `leveneTest()` function.
247. Actually, this is a bit of a lie. ANOVAs can vary in other ways besides the ones I've discussed in this book. For instance, I've completely ignored the difference between fixed-effect models, in which the levels of a factor are "fixed" by the experimenter or the world, and random-effect models, in which the levels are random samples from a larger population of possible levels (this book only covers fixed-effect models). Don't make the mistake of thinking that this book – or any other one – will tell you "everything you need to know" about statistics, any more than a single book could possibly tell you everything you need to know about psychology, physics or philosophy. Life is too complicated for that to *ever* be true. This isn't a cause for despair, though. Most researchers get by with a basic working knowledge of ANOVA that doesn't go any further than this book does. I just want you to keep in mind that this book is only the beginning of a very long story, not the whole story.
248. The one thing that might seem a little opaque to some people is why the residual degrees of freedom in this output look different from one another (i.e., ranging from 12 to 17) whereas in the original one the residual degrees of freedom is fixed at 12. It's actually the case that R uses a residual df of 12 in all cases (that's why the p values are the same in the two outputs, and it's enough to verify that `pf(6.7495, 2, 12, lower.tail=FALSE)` gives the correct answer of  $p=.010863$ , for instance, whereas `pf(6.7495, 2, 15, lower.tail=FALSE)` would have given a p-value of about .00812. It's the residual degrees of freedom in the full model (i.e., the last one) that matters here.
249. Or, at the very least, rarely of interest.
250. Yes, I'm actually a big enough nerd that I've written my own functions implementing Type II tests and Type III tests. I only did it to convince myself that I knew how the different Types of test worked, but it did turn out to be a handy exercise: the `etaSquared()` function in the `lsr` package relies on it. There's actually even an argument in the `etaSquared()` function called `anova`. By default, `anova=FALSE` and the function just prints out the effect sizes. However, if you set `anova=TRUE` it will spit out the full ANOVA table as well. This works for Types I, II and III. Just set the `types` argument to select which type of test you want.
251. Note, of course, that this does depend on the model that the user specified. If original ANOVA model doesn't contain an interaction term for `B:C`, then obviously it won't appear in either the null or the alternative. But that's true for Types I, II and III. They never include any terms that you *didn't* include, but they make different choices about how to construct tests for the ones that you did include.
252. I find it amusing to note that the default in R is Type I and the default in SPSS is Type III (with Helmert contrasts). Neither of these appeals to me all that much. Relatedly, I find it depressing that almost nobody in the psychological literature ever bothers to report which Type of tests they ran, much less the order of variables (for Type I) or the contrasts used (for Type III). Often they don't report what software they used either. The only way I can ever make any sense of what people typically report is to try to guess from auxiliary cues which software they were using, and to assume that they never changed the default settings. Please don't do this... now that you know about these issues, make sure you indicate what software you used, and if you're reporting ANOVA results for unbalanced data, then specify what Type of tests you ran, specify order information if you've done Type I tests and specify contrasts if you've done Type III tests. Or, even better, do hypotheses tests that correspond to things you really care about, and then report those!

---

This page titled [16.11: Summary](#) 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.