

16.3: Effect Size, Estimated Means, and Confidence Intervals

In this section I'll discuss a few additional quantities that you might find yourself wanting to calculate for a factorial ANOVA. The main thing you will probably want to calculate is the effect size for each term in your model, but you may also want to R to give you some estimates for the group means and associated confidence intervals.

16.3.1 Effect sizes

The effect size calculations for a factorial ANOVA is pretty similar to those used in one way ANOVA (see Section 14.4). Specifically, we can use η^2 (eta-squared) as simple way to measure how big the overall effect is for any particular term. As before, η^2 is defined by dividing the sum of squares associated with that term by the total sum of squares. For instance, to determine the size of the main effect of Factor A, we would use the following formula

$$\eta_A^2 = \frac{SS_A}{SS_T}$$

As before, this can be interpreted in much the same way as R^2 in regression.²³⁴ It tells you the proportion of variance in the outcome variable that can be accounted for by the main effect of Factor A. It is therefore a number that ranges from 0 (no effect at all) to 1 (accounts for *all* of the variability in the outcome). Moreover, the sum of all the η^2 values, taken across all the terms in the model, will sum to the total R^2 for the ANOVA model. If, for instance, the ANOVA model fits perfectly (i.e., there is no within-groups variability at all!), the η^2 values will sum to 1. Of course, that rarely if ever happens in real life.

However, when doing a factorial ANOVA, there is a second measure of effect size that people like to report, known as partial η^2 . The idea behind partial η^2 (which is sometimes denoted $p\eta^2$ or η_p^2) is that, when measuring the effect size for a particular term (say, the main effect of Factor A), you want to deliberately ignore the other effects in the model (e.g., the main effect of Factor B). That is, you would pretend that the effect of all these other terms is zero, and then calculate what the η^2 value would have been. This is actually pretty easy to calculate. All you have to do is remove the sum of squares associated with the other terms from the denominator. In other words, if you want the partial η^2 for the main effect of Factor A, the denominator is just the sum of the SS values for Factor A and the residuals:

$$\text{partial } \eta_A^2 = \frac{SS_A}{SS_A + SS_R}$$

This will always give you a larger number than η^2 , which the cynic in me suspects accounts for the popularity of partial η^2 . And once again you get a number between 0 and 1, where 0 represents no effect. However, it's slightly trickier to interpret what a large partial η^2 value means. In particular, you can't actually compare the partial η^2 values across terms! Suppose, for instance, there is no within-groups variability at all: if so, $SS_R=0$. What that means is that *every* term has a partial η^2 value of 1. But that doesn't mean that all terms in your model are equally important, or indeed that they are equally large. All it mean is that all terms in your model have effect sizes that are large *relative to the residual variation*. It is not comparable across terms.

To see what I mean by this, it's useful to see a concrete example. Once again, we'll use the `etaSquared()` function from the `lsr` package. As before, we input the `aoV` object for which we want the η^2 calculations performed, and R outputs a matrix showing the effect sizes for each term in the model. First, let's have a look at the effect sizes for the original ANOVA without the interaction term:

```
etaSquared( model.2 )
```

```
##           eta.sq eta.sq.part
## drug      0.7127623  0.7888325
## therapy  0.0964339  0.3357285
```

Looking at the η^2 values first, we see that `drug` accounts for 71.3% of the variance (i.e. $\eta^2=0.713$) in `mood.gain`, whereas `therapy` only accounts for 9.6%. This leaves a total of 19.1% of the variation unaccounted for (i.e., the residuals constitute 19.1% of the variation in the outcome). Overall, this implies that we have a very large effect²³⁵ of `drug` and a modest effect of `therapy`.

Now let's look at the partial η^2 values. Because the effect of `therapy` isn't all that large, controlling for it doesn't make much of a difference, so the partial η^2 for `drug` doesn't increase very much, and we obtain a value of $p\eta^2=0.789$). In contrast, because the effect of `drug` was very large, controlling for it makes a big difference, and so when we calculate the partial η^2 for `therapy` you can see that it rises to $p\eta^2=0.336$. The question that we have to ask ourselves is, what does these partial η^2 values actually *mean*? The way I generally interpret the partial η^2 for the main effect of Factor A is to interpret it as a statement about a hypothetical experiment in which *only* Factor A was being varied. So, even though in *this* experiment we varied both A and B, we can easily imagine an experiment in which only Factor A was varied: the partial η^2 statistic tells you how much of the variance in the outcome variable you would expect to see accounted for in that experiment. However, it should be noted that this interpretation – like many things associated with main effects – doesn't make a lot of sense when there is a large and significant interaction effect. Speaking of interaction effects, here's what we get when we calculate the effect sizes for the model that includes the interaction term. As you can see, the η^2 values for the main effects don't change, but the partial η^2 values do:

```
etaSquared( model.3 )
```

```
##                eta.sq eta.sq.part
## drug           0.71276230  0.8409091
## therapy        0.09643390  0.4169559
## drug:therapy    0.05595689  0.2932692
```

16.3.2 Estimated group means

In many situations you will find yourself wanting to report estimates of all the group means based on the results of your ANOVA, as well as confidence intervals associated with them. You can use the `effect()` function in the `effects` package to do this (don't forget to install the package if you don't have it already!). If the ANOVA that you have run is a **saturated model** (i.e., contains all possible main effects and all possible interaction effects) then the estimates of the group means are actually identical to the sample means, though the confidence intervals will use a pooled estimate of the standard errors, rather than use a separate one for each group. To illustrate this, let's apply the `effect()` function to our saturated model (i.e., `model.3`) for the clinical trial data. The `effect()` function contains two arguments we care about: the `term` argument specifies what terms in the model we want the means to be calculated for, and the `mod` argument specifies the model:

```
library(effects)
eff <- effect( term = "drug*therapy", mod = model.3 )
eff
```

```
##
## drug*therapy effect
##           therapy
## drug      no.therapy      CBT
## placebo    0.300000 0.600000
## anxifree    0.400000 1.033333
## joyzepam    1.466667 1.500000
```

Notice that these are actually the same numbers we got when computing the sample means earlier (i.e., the `group.means` variable that we computed using `aggregate()`). One useful thing that we can do using the effect variable `eff`, however, is extract the confidence intervals using the `summary()` function:

```
summary(eff)
```

```
##
## drug*therapy effect
##      therapy
## drug      no.therapy      CBT
## placebo    0.300000 0.600000
## anxifree   0.400000 1.033333
## joyzepam   1.466667 1.500000
##
## Lower 95 Percent Confidence Limits
##      therapy
## drug      no.therapy      CBT
## placebo   0.006481093 0.3064811
## anxifree  0.106481093 0.7398144
## joyzepam  1.173147759 1.2064811
##
## Upper 95 Percent Confidence Limits
##      therapy
## drug      no.therapy      CBT
## placebo   0.5935189 0.8935189
## anxifree  0.6935189 1.3268522
## joyzepam  1.7601856 1.7935189
```

In this output, we see that the estimated mean mood gain for the placebo group with no therapy was 0.300, with a 95% confidence interval from 0.006 to 0.594. Note that these are not the same confidence intervals that you would get if you calculated them separately for each group, because of the fact that the ANOVA model assumes homogeneity of variance and therefore uses a pooled estimate of the standard deviation.

When the model doesn't contain the interaction term, then the estimated group means will be different from the sample means. Instead of reporting the sample mean, the `effect()` function will calculate the value of the group means that would be expected on the basis of the marginal means (i.e., assuming no interaction). Using the notation we developed earlier, the estimate reported for μ_{rc} , the mean for level r on the (row) Factor A and level c on the (column) Factor B would be $\mu_{..} + \alpha_r + \beta_c$. If there are genuinely no interactions between the two factors, this is actually a better estimate of the population mean than the raw sample mean would be. The command to obtain these estimates is actually identical to the last one, except that we use `model.2`. When you do this, R will give you a warning message:

```
eff <- effect( "drug*therapy", model.2 )
```

```
## NOTE: drug:therapy does not appear in the model
```

but this isn't anything to worry about. This is R being polite, and letting you know that the estimates it is constructing are based on the assumption that no interactions exist. It kind of makes sense that it would do this: when we use `"drug*therapy"` as our input, we're telling R that we want it to output the estimated group means (rather than marginal means), but the actual input `"drug*therapy"` might mean that you want interactions included or you might not. There's no *actual* ambiguity here, because the model itself either does or doesn't have interactions, but the authors of the function thought it sensible to include a warning just to make sure that you've specified the actual model you care about. But, assuming that we genuinely don't believe that there are any interactions, `model.2` is the right model to use, so we can ignore this warning.²³⁶ In any case, when we inspect the output, we get the following table of estimated group means:

```
eff
```

```
##
## drug*therapy effect
##      therapy
## drug      no.therapy      CBT
## placebo    0.2888889 0.6111111
## anxifree    0.5555556 0.8777778
## joyzepam    1.3222222 1.6444444
```

As before, we can obtain confidence intervals using the following command:

```
summary( eff )
```

```
##
## drug*therapy effect
##      therapy
## drug      no.therapy      CBT
## placebo    0.2888889 0.6111111
## anxifree    0.5555556 0.8777778
## joyzepam    1.3222222 1.6444444
##
## Lower 95 Percent Confidence Limits
##      therapy
## drug      no.therapy      CBT
## placebo    0.02907986 0.3513021
## anxifree    0.29574653 0.6179687
## joyzepam    1.06241319 1.3846354
##
## Upper 95 Percent Confidence Limits
##      therapy
## drug      no.therapy      CBT
## placebo    0.5486979 0.8709201
## anxifree    0.8153646 1.1375868
## joyzepam    1.5820313 1.9042535
```

but the output looks pretty much the same as last time, and this book is already way too long, so I won't include it here.

This page titled [16.3: Effect Size, Estimated Means, and Confidence Intervals](#) 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.