Quick Calculation of Main Effects in a Fixed-Effects
Completely Randomized Factorial ANOVA with Equal *n*

Psychology 311
Spring 2013

Recall that, in the Psychology 310 handout introducing 1-Way ANOVA, we pointed out that a 1-Way, equal *n* ANOVA may be calculated as

$$F = \frac{nS_{\bar{X}}^2}{\hat{\sigma}^2} \tag{1.1}$$

where $S_{\bar{X}}^2$ is the variance of the cell means, and $\hat{\sigma}^2$ is the mean of the cell variances of the individual cells.

So, in our little (highly artificial) sample data set, in which we had data as shown in the table below, the *F* statistic is the variance of the 3 means (2,5,8) times 3, divided by 1.

Factor A		
Mean = 2	Mean = 5	Mean = 8
sd = 1	sd = 1	sd = 1
n = 3	*n* = 3	*n* = 3

Since the three means are evenly spaced with a spacing of 3, they have a standard deviation of 3, and a variance of 9, and the *F*-statistic is

$$F = \frac{nS_{\bar{X}}^2}{\hat{\sigma}^2} = \frac{(3)(9)}{1} = 27 \tag{1.2}$$

Exactly the same principle holds when evaluating main effects in a fixed-effects, completely randomized factorial design. Here is what you do in a 2-Way ANOVA: (1) reduce the problem to a 1-Way ANOVA by computing row (or column) means, and the *effective n* on which the effect was based, (2) compute $\hat{\sigma}^2$ as the *average of the individual cell variances from all of the individual cells,* (3) compute $S_{\bar{X}_{effect}}^2$ as the variance of the *marginal means* used to calculate the effect, and then compute the *F* statistic as the analog of the equation above, i.e.,

$$F_{effect} = \frac{n_{effect}S_{\bar{X}_{effect}}^2}{\hat{\sigma}^2} \tag{1.3}$$

Let's look at an example.

Here is a table giving cell means, standard deviations, and sample sizes. I've computed the marginal means, along with the effective sample sizes used to compute them.

		Factor.B		Marginal
Factor.A	M=100 sd=10 n=3	M=105 sd=11 n=3	M=110 sd=12 n=3	Xbar=105 $n_{effect}=9$
	M=80 sd=11 n=3	M=90 sd=10 n=3	M=100 sd=12 n=3	Xbar=90 $n_{effect}=9$
Marginal	Xbar=90 $n_{effect}=6$	Xbar=97.5 $n_{effect}=6$	Xbar=105 $n_{effect}=6$	

The general formula for equal n multi-way main effects calculation in a fixed effects completely randomized factorial ANOVA is this.

$$F_{effect} = \frac{n_{effect} S^2_{\bar{X}_{effect}}}{\hat{\sigma}^2} \tag{1.4}$$

where n_{effect} is the total n involved in each mean used to calculate the effect, $S^2_{\bar{X}_{effect}}$ is the variance of the *marginal means* used to calculate the effect, and $\hat{\sigma}^2$ is the mean of the variances in all of the cells. So, for example, $S^2_{\bar{X}_A}$ would be the variance of the numbers 105 and 90, and $n_A = 9$, since each row mean in the A effect was based on an n of 3+3+3=9.

Remember, $\hat{\sigma}^2$ is the mean of the *variances*, not the mean of the s.d.'s.

$$\hat{\sigma}^2 = \frac{S^2_{11} + S^2_{12} + S^2_{13} + S^2_{21} + S^2_{22} + S^2_{23}}{6} = \frac{10^2 + 11^2 + 12^2 + 11^2 + 10^2 + 12^2}{6} = 121.6667$$

In that case, you would have gotten an F value for your A effect of

```
> means<-c(105,90)
> variance.of.means <- var(means)
> sds <- c(10,11,12,11,10,12)
> sigma.hat.squared<- sum(sds^2)/length(sds)
> n.effect<-9
> F<-n.effect*variance.of.means/(sigma.hat.squared)
> F
[1] 8.321918
> sigma.hat.squared
[1] 121.6667
> variance.of.means
[1] 112.5
> n.effect
[1] 9
```

Note that the mean squares for the effect and residual are the numerator and denominator.

```
> mean.square.effect <- n.effect * variance.of.means
> mean.square.residuals <- sigma.hat.squared
> mean.square.A
[1] 1012.5
> mean.square.residuals
[1] 121.6667
```

Calculating Main Effect of Factor B proceeds in much the same way, except that now we process the 3 column means, and must remember that n_{effect} has changed to 6.

```
> means<-c(90,97.5,105)
> variance.of.means <- var(means)
> n.effect<-6
> sigma.hat.squared<-
+ F<- n.effect * variance.of.means/sigma.hat.squared
> F
[1] 2.773973
> mean.square.B <- n.effect * variance.of.means
> mean.square.B
[1] 337.5
> mean.square.residuals
[1] 121.6667
```

Remember, you almost always have the ability to quickly check out any calculation method by creating a simple example with real numbers. So suppose you simply change the cell means to 3 in your example, and for simplicity make 10,11,12 the standard deviations, not the variances. Remember the 3 evenly-spaced number rule!

Then the actual raw data corresponding to our cell means and standard deviations could be

	Factor.B			
	110	116	122	Xbar=105
	100	105	110	
Factor.A	90	94	98	
	91	100	112	Xbar=90
	80	90	100	
	69	80	88	
	Xbar=90	Xbar=97.5	Xbar=105	

Now, we know that everything remains the same except that, $n = 3$ and $n_{effect} = 9$ f

Let's verify with R. Run the R code below and see what I mean.

```
> score <-
c(110,100,90,116,105,94,122,110,98,91,80,69,100,90,80,112,100,88
)
> factor.A <- gl(2,9,18,labels = c("A1","A2"))
> factor.B <- gl(3,3,18,labels = c("B1","B2","B3"))
> data <- data.frame(score,factor.A,factor.B)
> data
   score factor.A factor.B
1    110       A1       B1
2    100       A1       B1
3     90       A1       B1
4    116       A1       B2
5    105       A1       B2
6     94       A1       B2
7    122       A1       B3
8    110       A1       B3
9     98       A1       B3
10    91       A2       B1
11    80       A2       B1
12    69       A2       B1
13   100       A2       B2
14    90       A2       B2
15    80       A2       B2
16   112       A2       B3
17   100       A2       B3
18    88       A2       B3
> summary(aov(score~factor.A * factor.B,data=data))
                 Df Sum Sq Mean Sq F value Pr(>F)
factor.A          1   1012  1012.5   8.322 0.0137 *
factor.B          2    675   337.5   2.774 0.1023
factor.A:factor.B 2     75    37.5   0.308 0.7404
Residuals        12   1460   121.7
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
>
```

Now this is easy if $n = 3$ per cell, but what about more complex data? In that case, you can download *MakeExactData.txt* from the R Support Routines section of the website. The *MakeExactData* function in those routines will create data exactly according to your specifications.

It is a routine that can not only create independent sample data, but also correlated data. But for now we won't need to use its advanced capabilities. Note, you should read the entire file into R if you intend to use the function.

Suppose you need 10 numbers with a mean of 10.45 and a sd of 2.97.

The MakeExactData function requires a mean, variance, and sample size as input. Here are some examples.

```
> x <- MakeExactData(10.45,2.97,10)
> mean(x)
[1] 10.45
> var(x)
     [,1]
[1,] 2.97
> x
             [,1]
 [1,] 11.971767
 [2,] 12.234273
 [3,] 10.500073
 [4,]  9.771050
 [5,] 12.014888
 [6,]  6.881049
 [7,] 12.066168
 [8,] 10.146610
 [9,] 10.129717
[10,]  8.784407
```

Here, I show how to create 3 columns of data that have means of 90,100,110 and standard deviations of 10,11,12 (note I square them before inputting them).

```
> xx <- MakeExactData(c(90,100,100),c(10^2,11^2,12^2),10)
> xx
            [,1]       [,2]       [,3]
 [1,]  75.37103 111.51990 107.63941
 [2,] 103.72597 109.68827 110.05878
 [3,]  85.71843  94.83606 115.36738
 [4,]  98.71229  81.61405 104.89867
 [5,]  81.32397  86.12281  98.06540
 [6,]  92.66404 115.88383  97.22963
 [7,]  75.58967  98.61341  95.06226
 [8,]  92.18418 103.49959  74.36122
 [9,]  96.93206 103.41786 108.13391
[10,]  97.77835  94.80422  89.18333
> apply(xx,2,sd)
[1] 10 11 12
> apply(xx,2,mean)
[1]  90 100 100
```