Factorial ANOVA (Between Groups)

A factorial design is when there are two or more independent variables, and every level of each variable is tested at every level of the other variables. That is, there are data from every possible combination of levels of the IVs.

Let's begin by looking at a problem we've already had reason to look at.

> ls()  # make sure your workspace is clear
character(0)
>附件("datasets.RData")
>names(scarify)  # Zach Rizutto, Psyc 497, Spring 2013
[1] "Age" "Sex" "Tattoos" "Num.Tattoos"
> my.scar=scarify[,c(2,9,11)]
> summary(my.scar)
Sex       Cut.Scar        RSES
F:80  No :124   Min. : 7.00
M:60  Yes: 16   1st Qu.:19.00
     Median :23.00
     Mean :22.62
     3rd Qu.:27.00
     Max. :30.00
> attach(my.scar)
>table(Sex,Cut.Scar)  # design and group sizes (crosstabs of IVs)
Sex Cut.Scar
F 65 15
M 59 1

Descriptive Statistics

> tapply(RSES,list(Sex,Cut.Scar),mean)  # tapply comes into its own!
  No      Yes                          # Question: are these weighted or
  F 22.75385 18.93333                          #    unweighted means?
  M 23.32203 28.00000
> tapply(RSES,list(Sex,Cut.Scar),sd)        # why is one NA?
  No      Yes
  F 4.670488 4.978047
  M 5.018726 NA

A Graph of the Results

> interaction.plot(Sex,Cut.Scar,RSES,type="b",pch=1:2,legend=T)    # next page

What effects are we seeing (in the sample)?
a) main effects
b) simple effects (not terribly interesting unless there is an interaction)
c) interaction effects (a significant interaction trumps all other effects!)

What problem do we have here with putting a lot of faith in this result?
This problem is not a good candidate for ANOVA for the following reasons:
1) the design is grotesquely unbalanced
2) one of the cells is very small (n = 1) - a reliability/influence issue

The ANOVA

For the record, this is how the ANOVA can (but not necessarily WILL) be done.

> aov.out = aov(RSES ~ Sex * Cut.Scar)     # unbalanced designs make interpretation
> summary(aov.out)                         # very complex (we'll come back to it)

<table>
<thead>
<tr>
<th>Df</th>
<th>Sum Sq</th>
<th>Mean Sq</th>
<th>F value</th>
<th>Pr(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td>1</td>
<td>64</td>
<td>63.65</td>
<td>2.702</td>
</tr>
<tr>
<td>Cut.Scar</td>
<td>1</td>
<td>134</td>
<td>133.69</td>
<td>5.675</td>
</tr>
<tr>
<td>Sex:Cut.Scar</td>
<td>1</td>
<td>66</td>
<td>65.72</td>
<td>2.790</td>
</tr>
<tr>
<td>Residuals</td>
<td>136</td>
<td>3204</td>
<td>23.56</td>
<td></td>
</tr>
</tbody>
</table>

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

The check for normality can be done like this.

> plot(aov.out,2)     # qnorm plot of the residuals (resid = score - cell mean)
Warning message:
not plotting observations with leverage one:
  19
> plot(aov.out,4)     # influential cases (we know we have at least one!)
Let's look at a better behaved problem.

```r
> detach(my.scar)
> rm(my.scar,aov.out)
> names(brainsize)
[1] "gender" "FSIQ" "wgt" "hgt" "count" "IQ"
> my.brain=brainsize[,c(1,6,5)]
> summary(my.brain)

gender      IQ         count
  Male  :20   High:20   Min.   : 790619
  Female:20   Low :20   1st Qu.: 855918
        Median : 905399
        Mean   : 908755
        3rd Qu.: 950078
        Max.   :1079549
```

The researchers were interested in whether or not brain size is related to IQ.

# Reference: In vivo brain size and intelligence. L. Willerman, R. Schultz, # J.N. Rutledge, E.D. Bigler (1991). Intelligence (Norwood), 15(22), 223-228, # Elsevier. The study was done at a large university in the southwestern U.S. # 40 students in an introductory psychology class (all right handed, all # "Anglo", with no history of alcoholism, unconsciousness, brain damage, heart # disease, or epilepsy) served as subjects. The bimodal distribution of FSIQ # was intentional as subjects were selected only if they had IQ scores above # 132 or below 104.

```r
> attach(my.brain)
> table(gender,IQ)
   IQ
 gender High Low
  Male 10 10
 Female 10 10
> tapply(count,list(gender,IQ),mean)
  IQ    count
  High 975346.6 934364.2
  Low  876549.5 848759.7
> tapply(count,list(gender,IQ),sd)
  IQ    count
  High 49404.70 56794.28
  Low  65073.99 43950.38
```

Any concerns?

```r
> interaction.plot(gender,IQ,count,type="b",pch=1:2,legend=T)     # next page
```

What (if any) effects are we seeing in the sample? Describe the following effects.

1) main effect of gender

2) main effect of IQ

3) simple effect of IQ at gender = female (at gender = male)

4) gender x IQ interaction
Effect Size

The usual effect size measure applies: $R^2 = \eta^2 = \frac{SS(\text{effect})}{SS(\text{total})}$

Sometimes partial $\eta^2$ is used: $\frac{SS(\text{effect})}{SS(\text{effect}) + SS(\text{error})}$
Interpreting Effects - Some Examples from the Literature

[Graphs showing the relationship between different dietary habits and systolic blood pressure for males and females.]

[Graphs showing the relationship between genetic risk score and the likelihood of developing major depressive disorder for different levels of childhood trauma for all MDD and severe MDD.]

[Graph showing the relationship between work-family conflict and employee loyalty for different levels of positive thinking.]

5
Note: MS = mortality-primed group
Performance on Two Tests Taken "On Holiday" or "At Work"

Performance on Stroop Task

Adults

Children

Mean RTs (ms)

Interval Between Stimulus Presentations (ms)
**Fig. 1.** Interaction effect between the culture orientation and the relationship of perceived CSR on the job satisfaction.

**Fig. 2.** Interaction effect between the power distance interaction and the relationship of perceived CSR on the job satisfaction.

CSR = corporate social responsibility
DV is score on a job satisfaction measure