STAT 461: Lab 9 - Two-Way ANOVA with No Replicate

1 Review: Factorial (or Crossed) Experimental Designs

Consider an experiment in which there are two factors (types of treatments), which we will call "Factor A" and "Factor B". If every possible combination of (1) the levels of Factor A and (2) the levels of Factor B are applied to experimental units, then we say that we have a **Factorial Treatment Design** or that **Factor A and Factor B are Crossed** with each other.

1.1 The Two-Way Complete ANOVA Model

The two-way complete ANOVA model for a factorial design is

$$Y_{ijt} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijt}, \quad \epsilon_{ijt} \overset{iid}{\sim} N(0, \sigma^2)$$

$$i = 1, 2, \ldots, a \qquad j = 1, 2, \ldots, b \qquad t = 1, 2, \ldots, r_{ij}$$

Where the first index i indexes the levels of Factor A, with

$$i = 1, 2, \ldots, a = \text{ the number of treatment levels of Factor A}$$

and the second index j indexes the levels of Factor B, with

$$j = 1, 2, \ldots b = \text{ the number of treatment levels of Factor B.}$$

The **main effects** are α_i, which encodes the effect of the i-th level of Factor A on the mean response, and β_j, which is the effect of the j-th level of Factor B on the mean response.

The **interaction effects** are $(\alpha\beta)_{ij}$ and encode the additional effect on the mean response of the joint combination of level i of Factor A and level j of Factor B.

2 Two-Factor ANOVA without Replication

For various reasons (cost, time, material limitations, for instance), it is sometimes the case that only one case is observed per treatment. If we only have a single observation in each "cell" ($r = 1$), we cannot do statistical inference anymore with a model including the interaction as we have no idea of the experimental error. Because for every treatment combination we only have one observation, error degrees of freedom is $ab(r-1) = 0$. No estimate of σ^2 will be available. ## 2.1 Assume Additivity One common way to handle this situation is to analyze the model as a two-factor ANOVA without an interaction term, where the error term is replaced by the interaction term. However, this additivity assumption is unverifiable. So that the model becomes:

$$Y_{ij} = \mu + \alpha_i + \beta_j + \epsilon_{ij}, \quad \epsilon_{ij} \overset{iid}{\sim} N(0, \sigma^2)$$

3 Example: Insurance Premium

A 3×2 factorial experiment with $r = 1$ per cell, where factor A is the size of city, and factor B is the region of the city. The following code reads the data into R and fits this two-way anova model. The ANOVA table for this analysis is

```
A = rep(c("small", "medium", "large"), 2)
B = c(rep("E", 3), rep("W",3))
value = c(140, 210, 220, 100, 180, 200)
df<-data.frame(A,B,value)

anova(lm(value ~ A+ B, data=df))
```

```
## Analysis of Variance Table
##
## Response: value
##           Df Sum Sq Mean Sq F value  Pr(>F)
## A          2   9300    4650      93 0.01064 *
## B          1   1350    1350      27 0.03510 *
## Residuals  2    100      50
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
summary(aov(value ~ A+ B, data=df))
```

```
##             Df Sum Sq Mean Sq F value Pr(>F)
## A            2   9300    4650      93 0.0106 *
## B            1   1350    1350      27 0.0351 *
## Residuals    2    100      50
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Homework Assignment

1.Brainstorming. A researcher investigated whether brainstorming is more effective for larger groups than for smaller ones by setting up four groups of agribusiness executives, the group sizes being two, three, four, and five, respectively. He also set up four groups of agribusiness scientists, the group sizes being the same as for the agribusiness executives. The researcher gave each group the same problem: "How can Canada increase the value of its agricultural exports?" Each group was allowed 30 minutes to generate ideas. The variable of interest was the number of different ideas proposed by the group. The results, classified by type of group (factor A) and size of group (factor B), were:

Table 1:

	B_1 (Two)	B_2 (Three)	B_3 (Four)	B_4 (Five)
A_1 (Agribusiness executives)	18	22	31	32
A_2 (Agribusiness scientists)	15	23	29	33

Assume that no-interaction ANOVA model is appropriate.

1. 20 points Plot the data in an interaction plot. Does it appear that interaction effects are present? Does it appear that factor A and factor B main effects are present? Discuss.
2. 40 points (By hand) Conduct separate tests for type of group and size of group main effects. In each test, use level of significance $\alpha = 0.01$ and state the alternatives, decision rule, and conclusion. What is the P-value for each test?
3. 40 points (By hand) Obtain confidence intervals for $D_1 = \mu_{.2} - \mu_{.1}$,$D_2 = \mu_{.3} - \mu_{.2}$, $D_3 = \mu_{.4} - \mu_{.3}$; use the Bonferroni procedure with a 95 percent family confidence coefficient. State your findings.