Chapter 3-TWO WAY ANOVA PRACTICE

Due Wednesday, October 1st

Work in groups and turn in one paper for your group!

Gender Differences in Performance on Mathematics Achievement Tests Data set on 861 ACT Assessment Mathematics Usage Test scores from 1987. The test was given to a sample of high school seniors who met one of three profiles of high school mathematics course work: (a) Algebra I only; (b) two Algebra courses and Geometry; and (c) two Algebra courses, Geometry, Trigonometry, Advanced Mathematics and Beginning Calculus.
These data were generated from summary statistics for one particular form of the test as reported by Doolittle (1989).
Source: Ramsey, F.L. and Schafer, D.W. (2002). *The Statistical Sleuth: A Course in Methods of Data Analysis (2nd ed)*, Duxbury.

Summary statistics, side-by-side boxplots, and interaction plots are given for these data.

```
options(show.signif.stars = F)
require(Sleuth2)
require(mosaic)
math <- ex1320
names(math)

## [1] "Sex"        "Background" "Score"
```

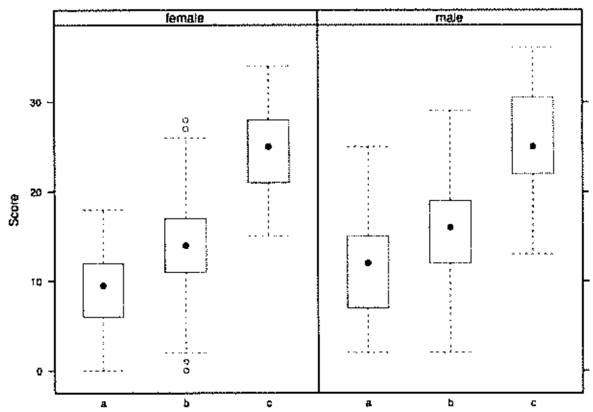
1. Is the design balanced? See the output below. No. Different 'n' for each group.

```
favstats(Score~Sex+Background, data = math)

##      .group min Q1 median   Q3 max   mean    sd   n missing
## 1 female.a    0  6    9.5 12.0  18  9.073 4.186  82       0
## 2   male.a    2  7   12.0 15.0  25 11.458 5.086  48       0
## 3 female.b    0 11   14.0 17.0  28 13.964 5.001 387       0
## 4   male.b    2 12   16.0 19.0  29 15.565 4.888 223       0
## 5 female.c   15 21   25.0 28.0  34 24.630 4.850  54       0
## 6   male.c   13 22   25.0 30.5  36 25.433 5.555  67       0
```

2. What do the side-by-side boxplots tell you about the effects of *Sex* and *Background*? In other words, describe the relationships you see in the boxplots.

```
bwplot(Score~Background|Sex, data = math)
```



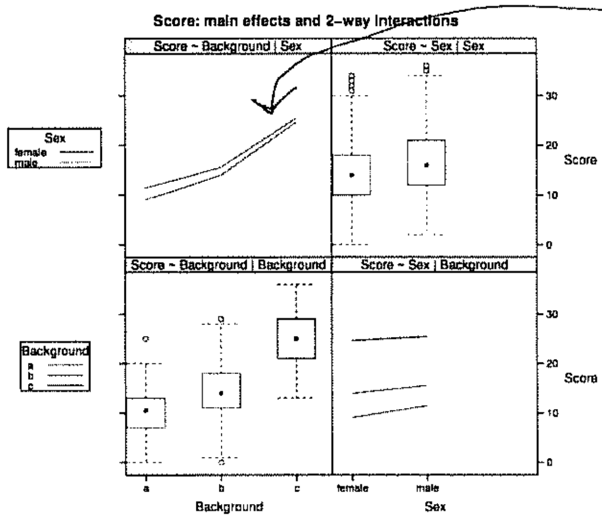
Group C has the highest ~~average~~ median scores + group A has the lowest ~~average~~ median scores for both sexes. In groups b + c, it appears that the ~~average~~ median score for males is higher than females

3. Interpret an interaction in this context.

How the difference in mean scores between males + females changes across levels of background.

4. Does the plot below suggest that there is an interaction between *Sex* and *Background*? Explain why or why not.

```
require(HH)
interaction2wt(Score~Background*Sex, data = math)
```



No - the lines are very close to parallel.

5. Write out the two-way anova interaction model in terms of y_{ijk}'s, τ_j's, γ_k's and ω_{jk}'s. Define all of the parameters and don't forget the errors!

$y_{ijk} = \mu + \tau_j + \gamma_k + \omega_{jk} + \varepsilon_{ijk}$

y_{ijk} = i^{th} response in $(j,k)^{th}$ treatment group

ε_{ijk} = error for i^{th} response in $(j,k)^{th}$ treatment group

μ = overall mean

τ_j = Sex $j = \{M, F\}$ γ_k = Background $k = \{A, B, C\}$

6. Below is output for the two-way ANOVA model with an interaction.

```
fit.Math <- lm(Score~Background*Sex, data = math)
anova(fit.Math)


## Analysis of Variance Table
##
## Response: Score
##                 Df Sum Sq Mean Sq F value  Pr(>F)
## Background       2  15619    7809  319.83 < 2e-16
## Sex              1    517     517   21.16 4.9e-06
## Background:Sex   2     38      19    0.77    0.46
## Residuals      855  20877      24
```

(a) Write the null and alternative hypothesis for the test you should look at first.

H_0: All ω_{jk}'s $= 0$

H_A: at least one $\omega_{jk} \neq 0$

(b) What is the distribution of the F statistic under the null hypothesis for this test?

F distribution with 2 and 855 df

(c) What is the value of the F statistic?

0.77

(d) What is the p-value?

0.46

(e) What would you conclude about the interaction effect?

There is no evidence that the difference in mean ACT scores between males and females changes across levels of background (p-value = 0.46 from F-stat = 0.77 on 2 + 855 df).

(f) Would you use an additive model or an interaction model for these data?

additive model

7. Which anova function should you use to fit an additive model? (circle the correct choice) **anova** or **Anova**

8. The type II sums of squares ANOVA table is given below:

```
fit.Math.add <- lm(Score~Background + Sex, data = math)
Anova(fit.Math.add)


## Anova Table (Type II tests)
##
## Response: Score
##              Sum Sq  Df F value  Pr(>F)
## Background  14705   2   301.3 < 2e-16
## Sex           517   1    21.2 4.8e-06
## Residuals   20914 857
```

For *Background:*

(a) Write the null and alternative hypothesis for testing the effect of background.

H_0: All γ_k's $= 0$ OR $\gamma_A = \gamma_B = \gamma_C = 0$

H_A: at least one $\gamma_k \neq 0$

(b) What is the distribution of the F statistic under the null hypothesis for this test?

F distribution on 2 and 857 df

(c) What is the value of the F statistic?

~~21.2~~ 301.3

(d) What is the p-value?

~~4.8×10^-6~~ 2×10^{-16}

(e) What would you conclude about the effect of *Background*?

There is strong evidence of at least one difference in mean ACT scores across backgrounds, after controlling for sex (p-value < 0.0001 from F-stat = 301.3 on 2 & 857 df).

For *Sex:*

(a) Write the null and alternative hypothesis for testing the effect of sex.

H_0: All τ_j's $= 0$ or $\tau_F = \tau_M = 0$

H_A: at least one $\tau_j \neq 0$

(b) What is the distribution of the F statistic under the null hypothesis for this test?

F distribution on 1 + 857 df

(c) What is the value of the F statistic?

21.2

(d) What is the p-value?

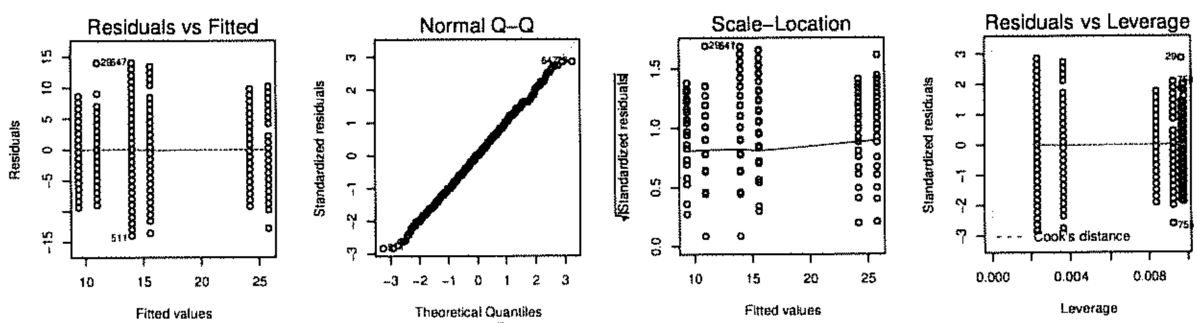
4.8×10^{-6}

(e) What would you conclude about the effect of *Sex*?

There is strong evidence of at least one difference in mean ACT scores across genders, after controlling for background (p-value <0.0001 from F-stat = 21.2 on 1 + 857 df).

9. Use the following plots to assess the assumptions and conditions required for the ANOVA.

```
par(mfrow = c(1,4))
plot(fit.Math.add)
```



equal spread across fitted values indicates that the equal variance assumption is <u>not</u> violated

points lie close to diagonal line. Normality assumption does not appear to be violated.

10. Use the output below to find the fitted values. In other words, I want you to tell me the estimated means for each treatment combination. *HINT: There are six, and here is the first one: the estimated mean test score for females with background a is* 9.363.

```
summary(fit.Math.add)

##
## Call:
## lm(formula = Score ~ Background + Sex, data = math)
##
## Residuals:
##      Min       1Q  Median      3Q     Max
## -13.964  -3.363   0.036   3.435  14.037
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)    9.363      0.452   20.72  < 2e-16
## Backgroundb    4.601      0.477    9.64  < 2e-16
## Backgroundc   14.825      0.627   23.63  < 2e-16
## Sexmale        1.601      0.348    4.60  4.8e-06
##
## Residual standard error: 4.94 on 857 degrees of freedom
## Multiple R-squared:  0.436,Adjusted R-squared:  0.434
## F-statistic:  220 on 3 and 857 DF,  p-value: <2e-16
```

Fem a $= 9.363$

Male a $= 9.363 + 1.601 = 10.964$

Fem b $= 9.363 + 4.601 = 13.964$

Fem c $= 9.363 + 14.825 = 24.188$

Male b $= 9.363 + 4.601 + 1.601 = 15.565$

Male c $= 9.363 + 14.825 + 1.601 = 25.789$

11. Which combination of *Background* and *Sex* had the highest mean score?

Males with background c

Two Questions about Multiple Comparisons

12. When should you use a multiple comparison procedure? Circle all that apply.

    A. You have multiple groups in the study, and you want to look at all the possible pairwise comparisons. EX: You have three groups, A, B, and C. You want to compare the means of A and B, B and C, and A and B.

    B. You have multiple groups in your study, you see the data, get a little curious, and then you decide you want to compare the means of two of the groups.

    C. You only have two groups in your study, and you are doing a two-sample t-test to compare the means.

    D. You only have two groups in your study, and you want to make a confidence interval for the true difference in means.

    E. You have multiple groups in your study, and you want to make confidence intervals for all the pairwise differences.

13. The "warpbreak" data(manipulated below) gives the number of warp breaks per loom, where a loom corresponds to a fixed length of yarn. The tension refers to the tension of the loom, set at low, medium, or high.

```
require(multcomp)
tension <- factor(warpbreaks$tension, levels=c("H", "M", "L"))
amod <- aov(warpbreaks$breaks ~ tension)
ps <- glht(amod, linfct = mcp(tension = "Tukey"))
confint(ps)
```

```
##
##   Simultaneous Confidence Intervals
##
## Multiple Comparisons of Means: Tukey Contrasts
##
##
## Fit: aov(formula = warpbreaks$breaks ~ tension)
##
## Quantile = 2.414
## 95% family-wise confidence level
##
##
## Linear Hypotheses:
##            Estimate lwr    upr
## M - H == 0  4.722   -4.837 14.281
## L - H == 0 14.722    5.163 24.281
## L - M == 0 10.000    0.441 19.559
```

Choose the correct interpretation of the confidence interval(s) above.

(a) We are 95% confident that the true difference in mean breaks between those yarns at a low tension and those yarns at a high tension is between 24.282 and 5.163 breaks.

(b) We are 95% confident that the high group has more breaks than the low group.

(c) We are 95% confident that all of the intervals above contain their respective true differences in mean breaks.

(d) We are more than 95% confident that all of the intervals above contain their respective true differences in mean breaks.