Math 484   Final Fall 2016   Name___________

```
Current terms: (Ht9 Lg18 Lg9 Soma St9 Wt18 Wt2 Wt9)
                df    RSS          |   k    C_I
Delete: Wt2     62     538.56      |   8    5.663
Delete: Lg9     62    538.791      |   8    5.688
...
Delete: Ht9     62    974.015      |   8   53.903

Current terms: (Ht9 Lg18 Lg9 Soma St9 Wt18 Wt9)
                df    RSS          |   k    C_I
Delete: Lg9     63    545.972      |   7    4.484
Delete: St9     63    552.069      |   7    5.159
...
Delete: Ht9     63    993.575      |   7   54.070

Current terms: (Ht9 Lg18 Soma St9 Wt18 Wt9)
                df    RSS          |   k    C_I
Delete: St9     64     557.89      |   6    3.804   Imin
Delete: Lg18    64    564.086      |   6    4.490
...
Delete: Ht9     64    1000.18      |   6   52.802

Current terms: (Ht9 Lg18 Soma Wt18 Wt9)
                df    RSS          |   k    C_I
Delete: Wt9     65    582.362      |   5    4.515  ← I_I   4.515 < 4.804 =
Delete: Lg18    65    583.148      |   5    4.602                Cp(Imin) + 1
...
Delete: Ht9     65    1065.78      |   5   58.069

Current terms: (Ht9 Lg18 Soma Wt18)
                df    RSS          |   k    C_I
Delete: Lg18    66    624.386      |   4    7.171  ←   7.171 < Cp(Imin) + 4
Delete: Wt18    66    763.596      |   4   22.593              = 7.804
Delete: Soma    66    843.467      |   4   31.441
Delete: Ht9     66    1249.22      |   4   76.391
```

Q8d22

1) The *ARC* data set BGSgirls predicted $Y = HT18 = height\ at\ age\ 18$ for ~~boys~~ girls from various predictors such as weights at age 2, 9, and 18; heights at age 2 and 9; leg circumference at age 8 and 19, stregth at age 9 and 19.

a) What terms are in model I_I? constant, HT9, LG18, Soma, WT18 -or -7

18

b) How many other models should be looked at?

$\boxed{1}$

(constant, HT9, Soma, WT18)

17

See 484 Q7d22

35

or -10

Label	Estimate	Std. Error	t-value	p-value
Constant	-2.93739	1.42523	-2.061	0.0422
mlife	1.12359	0.0229362	48.988	0.0000

R Squared: 0.96424 Sigma hat: 2.11667
Number of cases: 91 Degrees of freedom: 89

Summary Analysis of Variance Table

Source	df	SS	MS	F	p-value
Regression	1	10751.8	10751.8	2399.80	0.0000
Residual	89	398.746	4.48029		

2) The output above is for predicting Y = *female life expectancy* from X_2 = *male life expectancy*.

a) Predict Y if $X_2 = 67$.

$\hat{Y} = -2.93739 + 1.12359\,(67) = 72.3431$

b) Find a 95% confidence interval for β_2. $df = n-2 = 89$ so $t^* = z^* = 1.96$ or -2

$\hat{\beta}_2 \pm t^* SE(\hat{\beta}_2) = 1.12359 \pm 1.96(0.02294)$

$= 1.12359 \pm 0.04495 =$

$[1.0786\ ,\ 1.1685]$

c) Test whether X_2 is useful for predicting Y with an ANOVA F test.

t test −4

$H_0\ \beta_2 = 0 \quad H_A\ \beta_2 \neq 0$

$F_0 = 2399.8$

Pval $= 0$

reject H_0 there is an MLR relationship between the response female life expectancy and the predictor male life expectancy.

60

Math 484 Fall 2016 Final

Analysis of Variance for Full Model

Source	DF	SS	MS	F	P
Regression	3	4133.6	1377.9	13.01	0.000
Residual Error	19	2011.6	105.9		
Total	22	6145.2			

SSE(F)   MSE(F) df_F

Analysis of Variance for Reduced Model

Source	DF	SS	MS	F	P
Regression	2	4081.2	2040.6	19.77	0.000
Residual Error	20	2064.0	103.2		
Total	22	6145.2			

SSE(R) df_R

3) A hospital administrator wished to investigate the relationship between *patient satisfaction* (Y), and the predictors *age* (X_2), severity of illness (X_3), and *anxiety level* (X_4). Test whether X_4 can be dropped from the MLR model given that X_2 and X_3 are in the model. Use the output above.

i) H_0 reduced model is good H_A use the full model

ii) $$F_0 = \frac{\left[\frac{SSE(R) - SSE(F)}{df_R - df_F}\right]}{MSE(F)} = \frac{\left[\frac{2064 - 2011.6}{1}\right]}{105.9} = \underline{0.4948}$$

iii) $pval = P(F_{1,19} > 0.4948)$

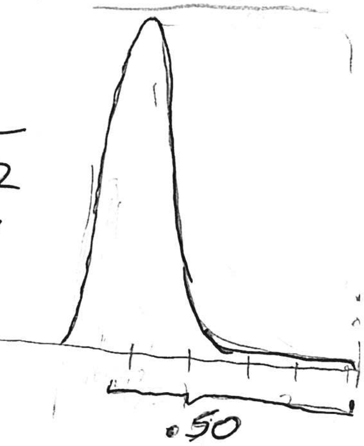
.05

		1
20	.5	.472
	.9	2.97

$0.5 = 1 - 0.5$

$0.1 = 1 - .9$

$0.1 < pval < 0.5$



iv) Fail to reject H_0, the reduced model is good

20

A	B	C
9.5	8.5	7.7
3.2	9.0	11.3
4.7	7.9	9.7
7.5	5.0	11.5
8.3	3.2	12.4

Analysis of Variance for Time

Source	DF	SS	MS	F	P
Design	2	49.168	24.584	4.4625	0.0356
Error	12	66.108	5.509		

4) Ledolter and Swersey (2007, p. 49) describe a one way Anova design used to study the effectiveness of 3 product displays (A, B and C). Fifteen stores were used and each display was randomly assigned to 5 stores. The response Y was the sales volume for the week during which the display was present compared to the base sales for that store.

E3d22

a) Find $\hat{\mu}_2 = \hat{\mu}_B$.

$$\frac{8.5+9+7.9+5+3.2}{5} = \frac{33.6}{5} = 6.72$$

17

b) Perform a 4 step Anova F test.

19

random effects

~ -7 -8

~ or

Ho $\mu_1 = \mu_2 = \mu_3$ HA not Ho

$F_o = 4.4625$

pval = 0.0356

reject Ho, the mean sales volume depends on the display

or -3

35

OIC M408T HW4 #5

Source	df	SS	MS	F	P
brand	5	854.53	170.906	238.71	0.0000
error	42	30.07	0.716		

5) A researcher is interested in the amount of sodium in beer. She selects 6 brands of beer at random from 127 brands and the response is the average sodium content measured from 8 cans of each brand.

a) State whether this is a random or fixed effects one way Anova. Explain briefly.

random effects, the beer brands are selected at random

E3022

17

b) Using the output above, perform the appropriate 4 step Anova F test.

Ho $\sigma_\mu^2 = 0$ H_A $\sigma_\mu^2 > 0$

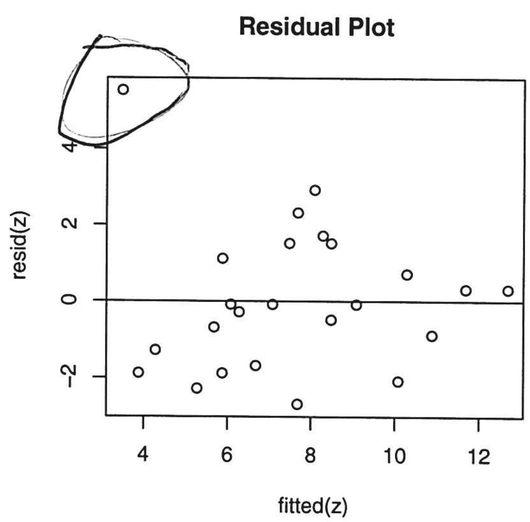
$F_0 = 238.71$

pval = 0

reject Ho, $\sigma_\mu^2 > 0$, the mean amount of sodium depends on the beer brand

19

35



E2022

Figure 1:

6) The residual plot shown above is from a one way block design. Circle the residual that is a potential outlier.

15

Source	DF	SS	MS	F	P
treat	3	0.9212	0.3071	13.8056	0.000
ptype	2	1.0330	0.5165	23.2217	0.000
Interaction	6	0.2501	0.0417	1.8743	0.1123
Error	36	0.8007	0.0222		

A

7) The Box et al. (2005, p. 318) poison data has 4 types of treatments (1,2,3,4) and 3 types of poisons (1,2,3). Each animal is given a poison and a treatment, and the response is survival in hours.

B swap A

B

a) Give a four step test for the "A*B" interaction.

i) Ho there is no interaction HA there is an interaction

ii) $F_{AB} = 1.8743$

iii) pval $= 0.1123$

iv) fail to reject Ho, there is no interaction between treatment and poison

AB -2

b) Give a four step test for the A main effects.

i) Ho $\mu_{10} = \cdots = \mu_{40}$ HA not Ho

4

ii) $F_A = 13.8056$

iii) pval $= 0$

iv) reject Ho, the mean survival time depends on treatment

c) Give a four step test for the B main effects.

i) Ho $\mu_{01} = \mu_{02} = \mu_{03}$ 3 HA not Ho

ii) $F_B = 23.2217$

iii) pval $= 0$

iv) reject Ho the mean survival time depends on the poison

60

8) Suppose that the regression model is $Y_i = 7 + \beta X_i + \epsilon_i$ for $i = 1, ..., n$ where the ϵ_i are iid $N(0, \sigma^2)$ random variables. The least squares criterion is $Q(\eta) = \sum_{i=1}^{n} (Y_i - 7 - \eta X_i)^2$.

a) What is $E(Y_i)$?

$$\boxed{7 + \beta X_i}$$

10

b) Find the least squares estimator $\hat{\beta}$ of β by setting the first derivative $\frac{d}{d\eta} Q(\eta)$ equal to zero.

E2022

$$\frac{dQ}{d\eta} = \sum 2 (Y_i - 7 - \eta X_i)(-X_i)$$

$$\text{or } -2 \sum (Y_i - 7 - \eta X_i) X_i \overset{set}{=} 0$$

$$\text{or } \sum (Y_i - 7) X_i = \eta \sum X_i^2$$

$$\text{or } \hat{\eta} = \boxed{\hat{\beta} = \frac{\sum_{i=1}^{n} (Y_i - 7) X_i}{\sum_{i=1}^{n} X_i^2}}$$

15

c) Show that your $\hat{\beta}$ is the global minimizer of the least squares criterion Q by showing that the second derivative $\frac{d^2}{d\eta^2} Q(\eta) > 0$ for all values of η.

$$\frac{d^2 Q(\eta)}{d\eta^2} = \frac{d}{d\eta} \left[-2 \sum (Y_i - 7) X_i + 2\eta \sum X_i^2 \right]$$

$$= \boxed{2 \sum_{i=1}^{n} X_i^2 > 0}$$

15

$$\left(= \frac{d}{d\eta} \; -2 \sum X_i Y_i + 14 \sum X_i + 2\eta \sum X_i^2 \right)$$

40