Math 484 Exam 3 Fall 2016 Name_____

One Way Analysis of Variance Table

Source	df	SS	MS	F	p-value
Treatments	2	100.647	33.549	12.08	0.0006
Error	28	33.328	2.777		
Total	15	133.974			

1) Nematodes are microscopic worms. A botanist desires to learn how the presence of the nematodes affects tomato growth. She uses 16 pots each with a tomato seedling. Four pots get 0 nematodes, four get 1000, four get 5000, and four get 10000. These four groups are denoted by "none," "n1000," "n5000" and "n10000" respectively. The seedling growths were all recorded and the table above is the one way ANOVA results.

a) What is $\hat{\mu}_1 = \hat{\mu}_{none}$? See c) below.

10.65

b) Do a four step test for whether the four mean growths are equal.

$\mu_{none} = \mu_{N1000} = \mu_{N5000} = \mu_{N10000}$

Ho $\mu_1 = \mu_2 = \mu_3 = \mu_4$ H_A not Ho

$F_o = 12.08$

pval = 0.0006

reject Ho the mean growths depends on the number of nematodes (worms)

c) Below is a Bonferroni comparison of means. Which groups of means are not significantly different?

VARIABLE	MEAN	HOMOGENEOUS GROUPS	
NONE	10.650	I	
N1000	10.425	I	
N5000	5.600		I
N10000	5.450		I

μ_{none} and μ_{1000} (1 and 2) are not significantly different and μ_{N5000} and μ_{N10000} (3 and 4)

1 2

first group of symbols

21

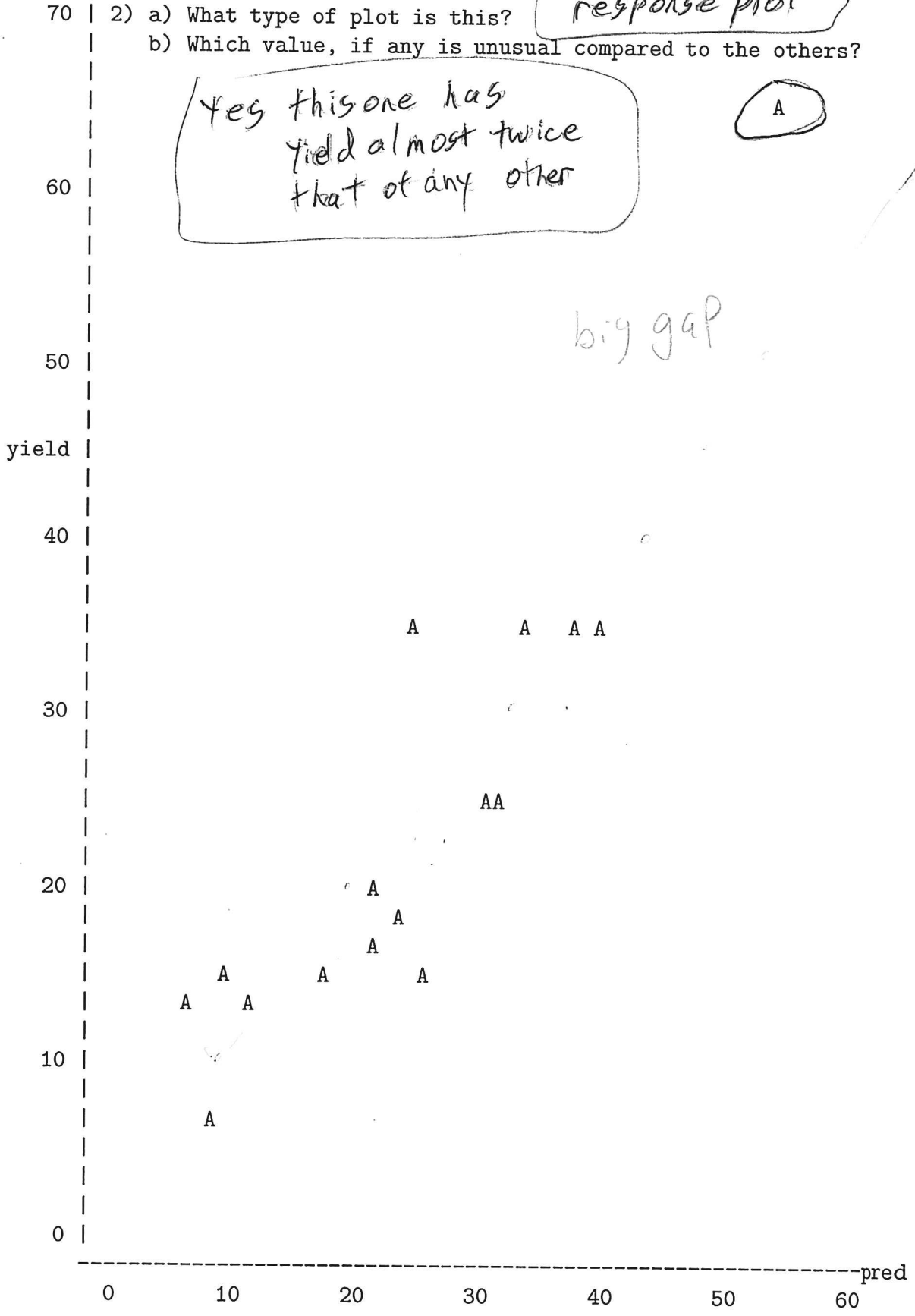
2) a) What type of plot is this? response plot

b) Which value, if any is unusual compared to the others?

yes this one has yield almost twice that of any other

A

big gap



12

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Treatment	3	0.88837	0.29612	44.853	0.000
Error	12	0.07923	0.00660		

3) Following Snedecor and Cochran (1967, p. 281), the response variable is the *calcium concentration* in turnip leaves. Four measurements were made on each of 4 leaves randomly selected from a population of leaves. The four measurements can be treated as a random sample from all possible measurements that could be made on the leaf. The researchers want to know if the calcium concentration depends on the leaf.

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

random effects, the leaves were randomly selected from a pop of leaves

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

$H_0: \sigma_u^2 = 0 \quad H_A: \sigma_u^2 > 0$

$F_0 = 44.853$

pval = 0

reject H_0, $\sigma_u^2 > 0$, the mean calcium concentration depends on the leaf

14e 4) Use the output from the command below

```
> sample(9)
 [1] 7 9 8 1 6 3 2 4 5
```

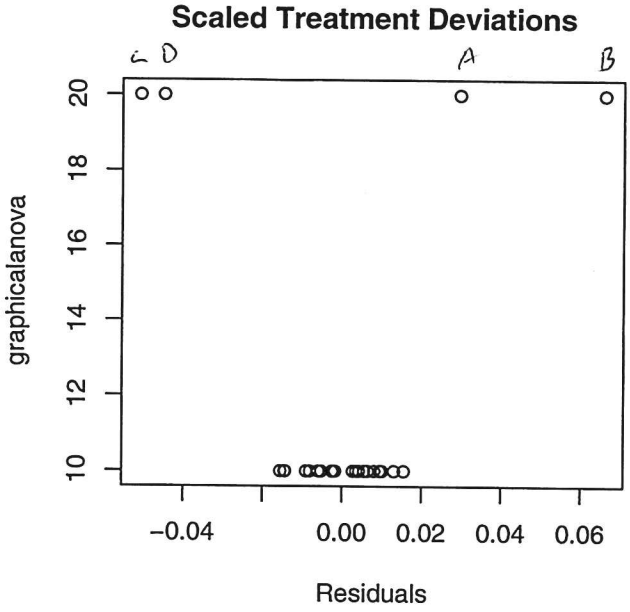
to assign the following 9 people to three groups.

Carroll (1), Collin (2), Crawford (3), Halverson (4), Lawes (5), Stach (6), Wayman (7), Wenslow (8), Xumong (9)

group 1: wayman xumong wenslow

group 2: carroll stach crawford

group 3: collin halverson lawes




```
ganova(x,y)
smn       0.02955502  0.06611268 -0.05080048 -0.04486722
Treatments "A"        "B"         "C"          "D"
```

5) Cobb (1998, p. 160) describes a one way Anova design used to study the amount of calcium in the blood. For many animals, the body's ability to use calcium depends on the level of certain hormones in the blood. The response was 1/(level of plasma calcium). The four groups were A: Female controls, B: Male controls, C: Females given hormone and D: Males given hormone. There were 10 birds of each gender, and five from each gender were given the hormone.

a) Which effects (from A, B, C and D) yielded similar mean response values?

C and D

b) Which effects (from A, B, C and D) appear to be significant?

all effects A, B, C, D are Significant

(out side the range of the residuals)

12

Math 484 Fall 2016 Exam 3

\begin{verbatim}

Math 583 Quiz 3 Fall 2008 Name_____

fat1	fat2	fat3	fat4
64	78	75	55
72	91	93	66
68	97	78	49
77	82	71	64
56	85	63	70
95	77	76	68

One way Anova for Fat1 Fat2 Fat3 Fat4

Source	DF	SS	MS	F	P
treatment	3	1636.5	545.5	5.41	0.0069
error	20	2018.0	100.9		

6) The output above represents grams of fat absorbed by doughnuts using 4 types of fat. Let μ_i denote the mean amount of fat$i$ absorbed by doughnuts, $i = 1, 2, 3$ and 4. Find $\hat{\mu}_3$.

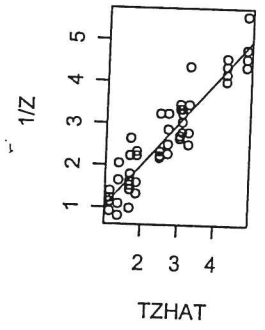
$= \frac{75 + 93 + 78 + 71 + 63 + 76}{6} = \frac{456}{6} = 76$

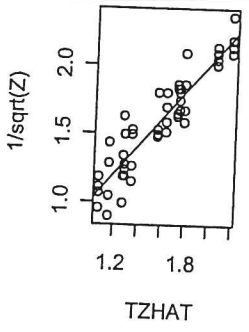
7

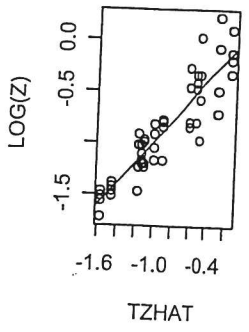
7) Using the transformation plots shown below, which two response transformations are the best?

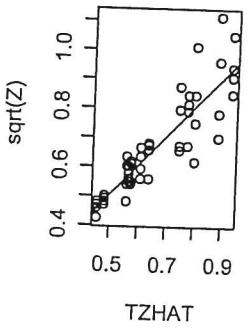
$Y = \frac{1}{Z}$ and $Y = \frac{1}{\sqrt{Z}}$

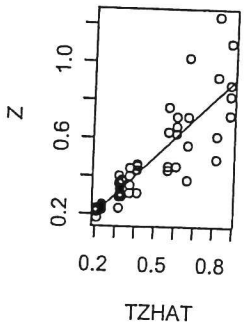
the others have spread that ↑ with TZHAT











7

```
Base terms: Intercept
           df    RSS          |   k    C_I
Add: X4    11    883.867      |   2  138.731
Add: X2    11    906.336      |   2  142.486
Add: X1    11    1265.69      |   2  202.549
Add: X3    11     1939.4      |   2  315.154

Base terms: (X4)
           df    RSS          |   k    C_I
Add: X1    10    74.7621      |   3    5.496
Add: X3    10    175.738      |   3   22.373
Add: X2    10     868.88      |   3  138.226

Base terms: (X4 X1)
           df    RSS          |   k    C_I
Add: X2     9    47.9727      |   4    3.018
Add: X3     9    50.8361      |   4    3.497

Base terms: (X4 X1 X2)
           df    RSS          |   k    C_I
Add: X3     8    47.8636      |   5    5.000

Current terms: (X1 X2 X3 X4)
              df    RSS          |   k    C_I
Delete: X3     9    47.9727      |   4    3.018
Delete: X4     9    48.1106      |   4    3.041
Delete: X2     9    50.8361      |   4    3.497
Delete: X1     9    73.8146      |   4    7.337

Current terms: (X1 X2 X4)
              df    RSS          |   k    C_I
Delete: X4    10    57.9045      |   3    2.678
Delete: X2    10    74.7621      |   3    5.496
Delete: X1    10     868.88      |   3  138.226

Current terms: (X1 X2)
              df    RSS          |   k    C_I
Delete: X1    11    906.336      |   2  142.486
Delete: X2    11    1265.69      |   2  202.549
```

too big

not as good as I_I

too many predictors

too many predictors

$I_{min} = I_I$

too big

8) From the output above, which terms, including the intercept, are in the "best" submodel (X1 is not the intercept, X1, X2, X3, and X4 are the four nontrivial predictors)?

constant, X1, X2

6

Math 484 Fall 2016 Exam 3

Base terms: (HT9 LG9)

	df	RSS	k	C_I
Add: WT18	62	562.396	4	5.726
Add: ST18	62	571.22	4	6.726
Add: WT2	62	571.515	4	6.760
...				
Add: Soma	62	582.183	4	7.968

← constant HT9 LG9 WT18

look at these 2 models

Base terms: (HT9 LG9 WT18)

	df	RSS	k	C_I
Add: WT2	61	545.381	5	5.798
Add: ST9	61	547.521	5	6.041
...				
Add: WT9	61	560.304	5	7.489

← constant HT9 LG9 WT18 WT2

Base terms: (HT9 LG9 WT18 WT2)

	df	RSS	k	C_I
Add: HT2	60	517.76	6	4.669
Add: ST9	60	532.773	6	6.370
Add: ST18	60	542.006	6	7.416
Add: WT9	60	542.249	6	7.443
Add: LG18	60	542.69	6	7.493
Add: Soma	60	544.485	6	7.697

$\leftarrow I_I$

two model with NO more predictors than I_{min} with $Cp(I) \leq Cp(I_{min}) + 1$

Base terms: (HT9 LG9 WT18 WT2 HT2)

	df	RSS	k	C_I
Add: ST9	59	497.552	7	4.379
Add: ST18	59	513.111	7	6.142
Add: WT9	59	514.527	7	6.302
Add: LG18	59	514.68	7	6.320
Add: Soma	59	515.498	7	6.412

$\leftarrow I_{min}$

Base terms: (HT9 LG9 WT18 WT2 HT2 ST9)

	df	RSS	k	C_I
Add: LG18	58	495.259	8	6.119
Add: WT9	58	495.294	8	6.123
Add: ST18	58	496.95	8	6.311
Add: Soma	58	497.514	8	6.374

9) The *ARC* data set BGSboys predicted $Y = HT18 = height\ at\ age\ 18$ for boys 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, LG9, WT18, WT2, HT2

b) How many other models should be looked at?

2

14 fewer predictors than I_I with $Cp(I) < Cp(I_{min}) + 4$!