

Statistics 515 - Fall 2002 - Exam 3 (modified for practice in Spring 2003)

Part I: Answer seven of the following eight questions. If you complete more than eight, I will grade only the first eight. Five points each.

- 1) In performing a linear regression to predict y from x , what four assumptions must be satisfied?
- 2) Define what is meant by the p -value (or the *observed significance level*) of a test.
- 3) (Circle the correct answers) Power can be increased by decreasing / increasing the sample size or by decreasing / increasing the α -level.
- 4) Sketch an example of a power curve for testing the null hypothesis of $\mu=100$ against the alternate hypothesis of $\mu>100$ for $\alpha=0.10$.
- 5) (Circle the correct answers) A student achieved a score of 780 out of 800 on an aptitude test. If you knew nothing else about the student, regression to the mean would imply that they would score lower / about the same / higher if they retook the test. A student scoring a 500 out of 800 (near the average) would score lower / about the same / higher if they retook it.
- 6) Consider the following partial ANOVA table for a one-way analysis of variance. How many different treatment groups were there in this experiment?

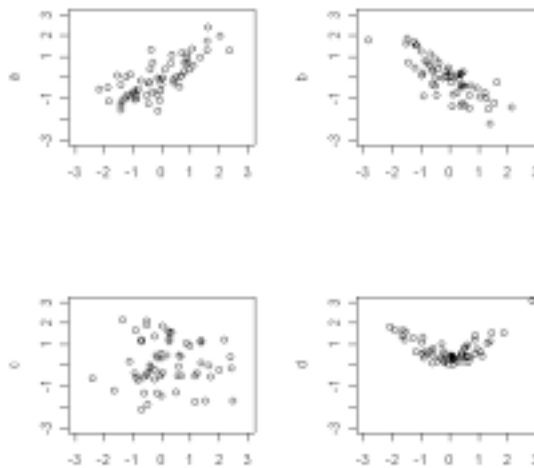
Source	SS	DF
Treatments	2794.39	3
Error	762.30	36
Total	3556.69	39

- 7) A one-way analysis of variance has five different treatment groups. The population means for these five groups are μ_1 , μ_2 , μ_3 , μ_4 , and μ_5 . Specify the null and alternate hypotheses that are tested by the p -value in the ANOVA table.

- 8) For each of the four graphs to the right (a-d), identify the correct correlation coefficient.

- $r = -1.00$ _____
 $r = -0.75$ _____
 $r = 0.0$ _____
 $r = 0.75$ _____
 $r = 1.00$ _____

You may use a value more than once, and obviously not all five values can be used for the four graphs.



Part II: Answer every part of the next two problems. Read each problem carefully, and show your work for full credit. Twenty points each.

1) A paper from 1857 by the Scottish physicist James D. Forbes attempts to measure the effect of barometric pressure (a measure of altitude) on the boiling point of water (in degrees Fahrenheit). The SAS input corresponding to this data is given on the attached sheets.

- Assuming the assumptions of the regression model are met, an increase in the pressure by 1 inch of mercury would be predicted to have what effect on the boiling point?
- Assuming the assumptions of the regression model are met, what is the p-value for testing the hypothesis that $\beta_1 = 0$? Do we accept or reject this null hypothesis at $\alpha = 0.05$?
- Assuming the assumptions of the regression model are met, what is our estimate for the standard deviation (σ) of the error terms in the regression equation?
- Assuming the assumptions of the regression model are met, what percent of the variation/error in the boiling point is explained by the barometric pressure?
- Which assumption(s) for performing a simple linear regression seems to be violated for this example, and how did you tell?

2) The following is the incomplete work for a linear regression problem.

$$\begin{aligned} SS_{xx} &= 1.90 & \text{average } x &= 1.49 \\ SS_{yy} &= 2330.21 & \text{average } y &= 36.94 \\ SS_{xy} &= 62.46 \end{aligned}$$

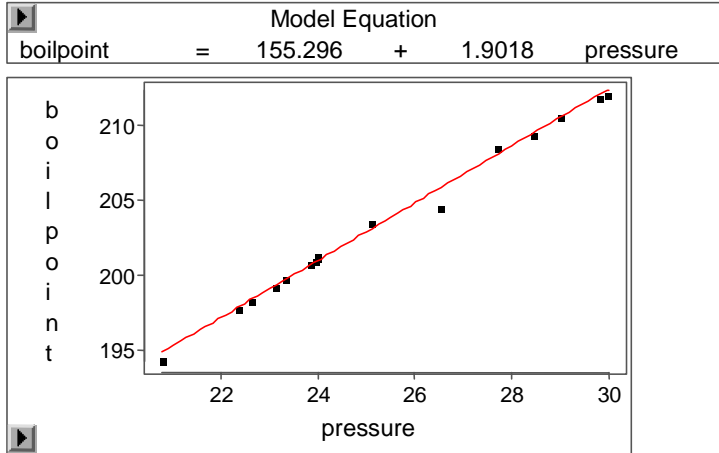
Source	SS	DF	MS	F	Prob>F
Regression	2053.64	1	_____	_____	<0.001
Error	_____	13	_____		
Total	2330.21	_____			

- Complete the table above by writing in the missing values.
- Determine the estimated regression equation.
- What was the original sample size?
- Determine a 90% confidence interval for the slope β_1 .

```

DATA forbes;
INPUT boilpoint pressure;
CARDS;
194.5 20.79
194.3 20.79
197.9 22.4
198.4 22.67
199.4 23.15
199.9 23.35
200.9 23.89
201.1 23.99
201.4 24.02
201.3 24.01
203.6 25.14
204.6 26.57
209.5 28.49
208.6 27.76
210.7 29.04
211.9 29.88
212.2 30.06
;
PROC INSIGHT;
OPEN forbes;
FIT boilpoint = pressure;
RUN;

```



Summary of Fit			
Mean of Response	202.9529	R-Square	0.9944
Root MSE	0.4440	Adj R-Sq	0.9941

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Stat	Pr > F
Model	1	527.8249	527.8249	2677.11	<.0001
Error	15	2.9574	0.1972		
C Total	16	530.7824			

