

WEB-DISTANCE ST 370 Quiz 1 FALL 2007 ver. B

NAME _____

ID # _____

I will neither give nor receive help from other students during this quiz

Sign _____

PROBLEM 1: If the number 3 is added to every member of a sample of observations (as might happen if an ohm meter was not zeroed before doing measurements) THEN

- a) the sample standard deviation of the sample does not change
- b) the sample variance is increased by 9 times
- c) the sample variance is increased by 3 times
- d) the sample mean does not change
- e) the sample median does not change

PROBLEMS 2 and 3: A sample of 5 cars parked along the Hillsborough Street across from the library revealed the following data on the NUMBER OF M&M's on the floor under the passenger-side seat

33, 7, 3, 54, 12

2) The value of the (Sample Mean MINUS the Sample median) for this dataset is :

- a) 5.2 b) 9.8 c) 18.5 d) 12.3 e) none of these

3) What is the 37th percentile for this dataset

- a) 9.6 b) 8.4 c) 8.1 d) 7.4 e) none of these

4) How is the correlation coefficient r related to the least squares slope b

- a) $r = b s_X / s_Y$
- b) $r = s_X / (b s_Y)$
- c) $r = b s_Y / s_X$
- d) $r = s_Y / (b s_X)$
- e) none of these

Problem 5: The following lm output on response of time from 0 to 60 mph for different tire pressure and weight in trunk of car indicate which terms are statistically significant

- a) intercept, weight, pressure
- b) only intercept and weight
- c) only weight* pressure
- d) no terms are significant
- e) all are significant

PROBLEM 6: for a weight of 100 lbs and a pressure of 20 psi, the predicted time is to 3 decimal places

- a) 7.498 b) 7.574 c) 7.238 d) 7.322 e) none of these

```

III. lm>> model time=weight+pressure+weight*pressure

Sequential Sums of Squares ANOVA Table

Source      df      SS      MS      F      P-val
Intercept    1  4645.16810  4645.168100  304629.00890  0.000000000
weight       1   71.82400   71.824000  4710.20070  0.000000000
pressure     1    0.46650    0.466500   30.59270  8.6862e-007
weight*pressure 1    0.00720    0.007200    0.47217  0.494820000
Error       56    0.85392    0.015249

R-square 0.98833

Standard Error 0.12349

Parameter Estimates

Source      Parameter Estimate  Std. Error    t    p-val
Intercept    6.868200  0.148700000  46.18960  0.000000000
weight       0.003545  0.000287950  12.31120  0.000000000
pressure     0.018173  0.004509000   4.03040  0.00016977
weight*pressure -6e-006.0  8.7317e-006  -0.68715  0.49482000

```

PROBLEM 7: A data set for precipitation has 86 observations and the stem and leaf plot is given below.

The 25th percentile is

- a) 2.0 b) 2.075 c) 2.135 d) 1.95 e) 1.8

LEAF (LAST DIGIT)	
	0. 3 3 4 5 6 7 7 8
	1. 0 1 2 2 2 5 5 6 8 8 8
Stem	2. 0 0 1 2 2 2 3 4 4 6 6 6 7 7 8 9 9 9 9 9 9 9
	3. 0 1 1 1 2 2 2 3 3 3 4 4 4 4 5 6 6 6 6 6 7 7 7 8 8 9 9
	4. 0 0 0 0 0 1 2 4 5 8
	5. 0 4 4 8 9

PROBLEM 8: The _____ refers to all possible units of interest or to a measurement on all units of interest

- a) sample
- b) population
- c) dependent variable
- d) independent variable
- e) none of the above

PROBLEM 9: A very small p value means

- a) sample sizes were too small.
- b) an important scientific finding.
- c) strong evidence against the null hypothesis.
- d) that roundoff error is likely.
- e) none of these.

PROBLEM 10: What type of data is the following:

area codes used in long distance dialing

- a) nominal b) interval or ratio c) ordinal d) temperamental e) none of these

The next set of problems concern a project from last semester. BG is an avid basketball fan so they looked at factors affecting the number of shots made (through the goal) out of 10 shots. Two different ball types were used : a standard basketball and a tennis ball. The balls were shot from distances of 5, 10, 15 and 20 feet. The EIGHT possible combinations of ball type and distance were tried on Saturday in a random order. All EIGHT combinations were repeated on the next day Sunday in a random order.

PROBLEM 11: Tennis ball is _____

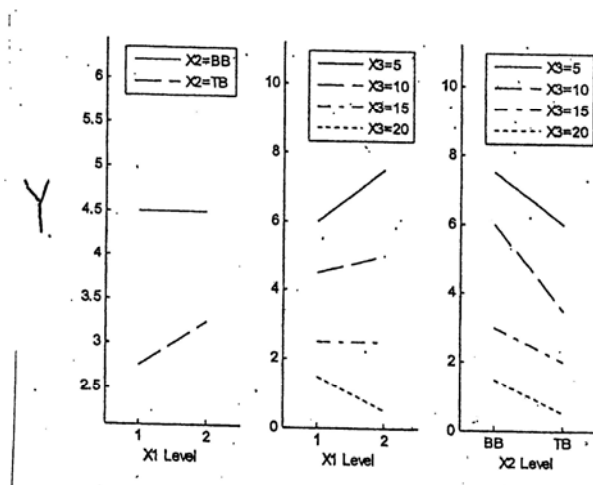
- a) a factor
- b) a level of a factor
- c) a confounding factor
- d) a treatment
- e) a blocking factor

PROBLEM 12: This project is what kind of study

- a. An observational study.
- b. A completely randomized design.
- c. A randomized complete block design.
- d. An incomplete block design.
- e. none of the above.

The following graph gives the M-plots (means plots) for this study using the notation Note the vertical scale is NOT the same for all 3 plots.

X1 = DAY X2 = BALL(type) X3 = DISTANCE



PROBLEM 13: The means plot suggests

- a) an interaction between balltype and distance
- b) an interaction between day and distance
- c) all three interactions suggested
- d) more replication is needed
- e) an interaction between day and balltype

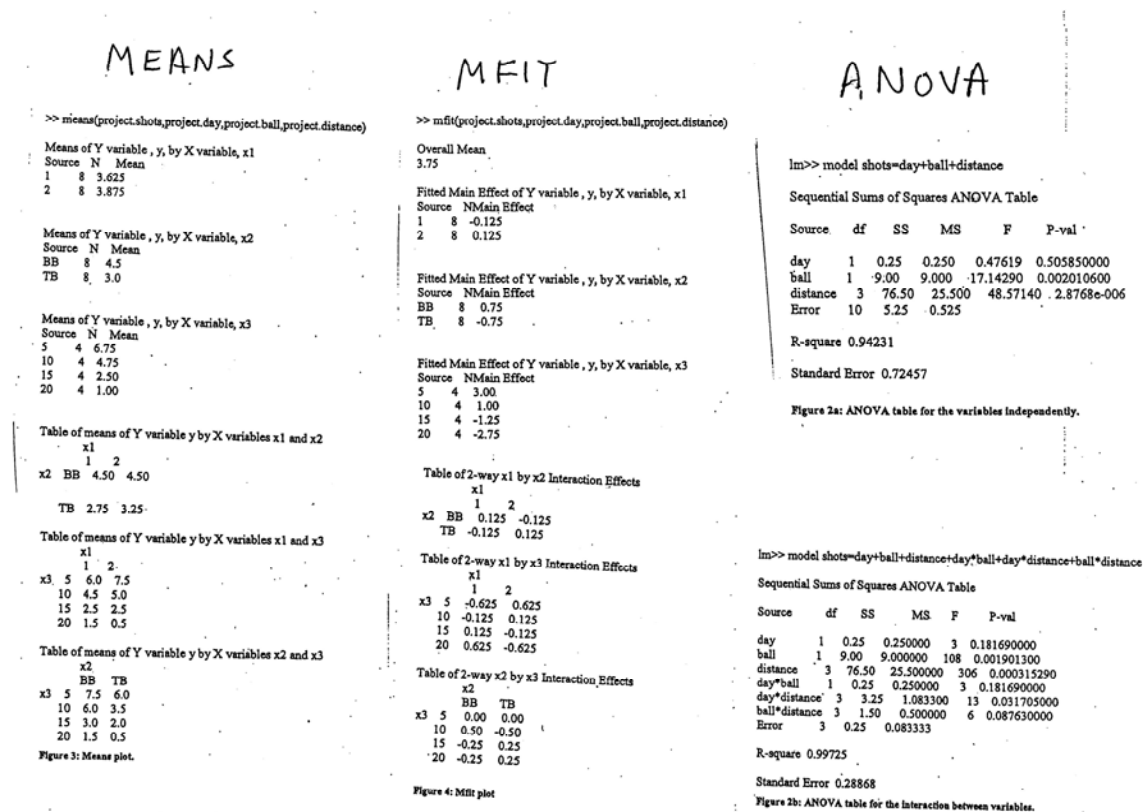
the means and mfit output and two ANOVA outputs are attached below

PROBLEM 14: The mfit output suggests

- distance is more important than balltype
- day and first two interactions are equally important since they all have at least one value of 0.125
- an R^2 value close to 1
- The overall mean is too small to make any comparisons
- day is more important than balltype

PROBLEM 15: The ANOVA output indicates

- only distance is statistically significant
- standard error is too small to make conclusions
- day is the most important factor
- balltype, distance and the interaction of day and distance are significant
- day and day*ball are most important terms



PROBLEMS 16, 17, 18:

The next three questions relate to the following printout from a MATLAB least squares of Y versus x :

R-square 0.85231

Standard Error 0.43321

Parameter Estimates

Source	Parameter Estimate	Std. Error	t	p-val
Intercept	4.51	0.71183	6.3358	0.0079511
x	-0.57	0.13699	-4.1608	0.0252500

PROBLEM 16: The least squares equation is

- a. $Y = 6.3358 - 4.1608x$
- b. $Y = 0.71183Z + 0.13699x$
- c. $Y = 4.51 - 0.57x$
- d. $Y = -0.57 + 0.13699x$
- e. None of the above.

PROBLEM 17: The "the proportion of variation in Y explained by the least squares line" (to two decimal places) is

- a. 0.92
- b. 0.43
- c. 0.85
- d. 0.66
- e. None of the above.

PROBLEM 18: The correlation coefficient r (to two decimal places) is :

- a) 0.92
- b) -0.92
- c) 0.43
- d) 0.85
- e) -0.85

PROBLEM 19: The following sample data represent the gasoline mileages (in miles per gallon) determined for cars in a particular weight class:

25.1 29.0 34.5 35.7 37.9
34.9 24.3 26.6 27.3 32.0
30.0 34.5 35.3 33.5 36.6
34.8 16.2 13.1 24.5 33.6
28.0 33.9 30.7 32.0 37.7
21.1 31.2 35.6 34.4 25.2
35.9 18.3 29.4 29.5 34.8
29.4 26.4 38.8 36.0 28.7
23.4 35.3 33.7 38.1 28.6
34.2 34.8 39.2 39.9 36.8

Using 10 mpg as the lower limit of the first class interval, construct a histogram with intervals of width 2 mpg. What is the midpoint of the cell with the largest number of observations (the mode)?

- a) 27.5 b) 30 c) 34.2 d) 32.5 e) 35

PROBLEM 20:

For the histogram in problem 19, is the histogram

- a) symmetric
- b) skewed to the right
- c) skewed to the left
- d) none of these

EXTRA CREDIT:

What topic in the course are you having the most difficulty with ?

Any suggestions for improving the course ?