

## FINAL EXAM

**Problem 1:** (part a = 5%; part b = 10%; TOTAL = 15%)

Data on 15 adult men were collected to see if there was a significant relationship between a person's height and his weight. The following EXCEL results summarize the regression

$$\text{Weight} = \beta_0 + \beta_1 \text{Height} \quad \text{where } \mathbf{Weight} \text{ is measured in pounds and } \mathbf{Height} \text{ in inches.}$$

Regression Statistics	
Multiple R	0.879424
R Square	0.773386
Adjusted R Square	0.755955
Standard Error	10.16656
Observations	15

ANOVA

	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	1	4585.666	4585.666	44.36638	1.56E-05
Residual	13	1343.667	103.359		
Total	14	5929.333			

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>
Intercept	-314.886	70.6437	-4.45738	0.000646
Height	6.854508	1.02908	6.660809	1.56E-05

- a. Based on this output is there a significant relationship between a person's height and his weight?  
In other words, if you test

$$H_0: \beta_1 = 0$$

$$H_1: \beta_1 \neq 0$$

at any reasonable level of significance (e.g.,  $\alpha=0.1, 0.05, 0.01$ ), do you reject the null hypothesis?

**Justify your answer briefly.**

- b. Previous research has suggested that on average men weigh 7 pounds more for every inch of height. Can you reject the null hypothesis that  $\beta_1 = 7$  in favor of the alternative hypothesis that  $\beta_1 \neq 7$  at  $\alpha = 0.05$ ? **Again show your work and justify your answer.**

**Problem 2:** (part a = 2% for each of the 6 numbers; part b = 5%; **TOTAL = 17%**)

For the 15 men used in the regression for problem 1, the raw data yielded the following sums where  $h_i$  is the height of the  $i^{\text{th}}$  person and  $w_i$  is the weight of the  $i^{\text{th}}$  person.

$$\sum_{i=1}^{15} h_i = 1,029$$

$$\sum_{i=1}^{15} w_i = 2,330$$

$$\sum_{i=1}^{15} h_i^2 = 70,687$$

$$\sum_{i=1}^{15} w_i^2 = 367,856$$

$$\sum_{i=1}^{15} h_i w_i = 160,507$$

- a. Complete the following table:

	Sample Average	Sample Variance	Sample standard deviation
Height			
Weight			

**SHOW YOUR WORK and the EQUATION S you use ON THE FOLLOWING BLANK PAGE**

- b. What is the sample covariance?

Place your name on ALL pages NOW.  
THANKS!

Name \_\_\_\_\_

**This is a work page for problem 2**

**Problem 3:** (part a = 6%; part b = 8%; part c = 8%; TOTAL = 20%)

A sample of 120 people in a large professional organization found that 88 people were in favor of changing the bylaws to allow a new membership category called “Fellows” to be established within the professional organization to recognize those people who had demonstrated significant achievements in the profession. Clearly, the other 32 did not favor this change.

a. Find the sample proportion who favor the change?

b. Consider the following hypothesis:

$$H_0: p \leq 2/3$$

$$H_1: p > 2/3$$

$$\alpha = 0.01$$

Can you reject the null hypothesis at the indicated level of significance? **Clearly show your work.**

c. If a 2/3 majority of the membership is needed to institute this change, find the probability that the change will be adopted when the resolution is put to the entire membership. **Clearly show your work.**

**Problem 4:** (part a = 2% for each of the 8 numbers; part b = 6%; TOTAL = 22%)

A tire manufacturer collects data (from a tire store) on the number of miles that its tires have been driven since they were last replaced. The manufacturer believes that the data are Normally distributed. A sample of 100 sets of tires (a set is 4 tires, but this is not relevant to the problem) is collected. The sample mean for the 100 sets of tires is 41,599 miles, while the sample standard deviation is 10,322 miles.

- a. Using this information, complete the following table. In other words, find the values for the numbers that should go in the 8 highlighted cells. *This table will be used in part (b) to test the Normality of the data.*

Show your work clearly in the space below, but put your answers in the indicated cells.

Low	High	Probability	Expected	Observed	$(O_i - E_i)^2 / E_i$
0	25,000			6	
25,001	29,000	0.0572	5.7215	6	0.0136
29,001	33,000	0.0913	9.1285	7	0.4963
33,001	37,000	0.1256	12.5565	13	0.0157
37,001	41,000	0.1489	14.8909	17	0.2987
41,001	45,000	0.1523	15.2251	12	0.6832
45,001	49,000	0.1342	13.4211	17	0.9544
49,001	53,000	0.1020	10.2000	7	1.0039
53,001	57,000	0.0668	6.6833	10	1.6459
57,001	infinite				
Total		1.0000	100.0000	100	

Place your name on ALL pages NOW.  
THANKS!

Name \_\_\_\_\_

- b. Test the hypothesis that the data come from a Normal distribution at a significance level of  $\alpha=0.05$ . Can you reject the null hypothesis that tire life is Normally distributed?  
**Clearly justify your answer and show your work below.**

<b>Note: If you did not get part (a), use the partial information shown in the table above to do this test.</b>
---

**Problem 5:** (part a = 2% for each of the 10 numbers; part b = 6%; TOTAL = 26%)

Data are collected on the number of calls for ambulance services in each of 12 suburban areas over a 10-day period. The data, along with the town populations in 1000s of people, are shown below:

Town	Population (1000s)	Calls over 10 days
Buffalo Grove	42.9	30
Deerfield	18.4	24
Evanston	74.2	57
Highland Park	31.4	32
Kennilworth	2.5	11
Lake Forest	20.1	11
Lincolnwood	12.4	5
Morton Grove	22.5	3
Niles	30.1	31
Skokie	63.3	57
Wilmette	27.7	17
Winnetka	12.4	24
Total	357.9	302

For these data, we can compute the following summary statistics:

$$\bar{X} = \frac{\sum_{i=1}^{12} x_i}{n} = \frac{357.9}{12} = 29.825$$

$$\bar{Y} = \frac{\sum_{i=1}^{12} y_i}{n} = \frac{302}{12} = 25.1667$$

$$S_{xx} = \sum_{i=1}^n (x_i - \bar{X})^2 = \sum_{i=1}^n x_i^2 - \frac{1}{n} \left( \sum_{i=1}^n x_i \right)^2 = 15574.79 - \frac{1}{12} 357.9^2 = 4900.4225$$

$$S_{yy} = \sum_{i=1}^n (y_i - \bar{Y})^2 = \sum_{i=1}^n y_i^2 - \frac{1}{n} \left( \sum_{i=1}^n y_i \right)^2 = 11100 - \frac{1}{12} 302^2 = 3499.6667$$

$$S_{xy} = \sum_{i=1}^n (x_i - \bar{X})(y_i - \bar{Y}) = \sum_{i=1}^n x_i y_i - \frac{1}{n} \left( \sum_{i=1}^n x_i \right) \left( \sum_{i=1}^n y_i \right) = 12650.6 - \frac{1}{12} 357.9 \times 302 = 3643.45$$

Place your name on ALL pages NOW.  
THANKS!

Name \_\_\_\_\_

- a. Complete the following regression table from EXCEL. **Fill in all highlighted areas.** If a requested P-value is smaller than any value you can find in the text (e.g., less than 0.01) simply note that result (e.g., <0.01).

<i>Regression Statistics</i>	
Multiple R	0.8798
R Square	
Adjusted R Square	Ignore this cell
Standard Error	
Observations	12

ANOVA

	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	1		2,708.8946		
Residual		790.7721			
Total	11	3,499.6667			

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>
Intercept		4.5765	0.6539	0.5280
Population (1000s)		0.1270		

Show your work below. For the Significance F and the requested P-value, be sure to show the appropriate number of degrees of freedom in your work below and on the next page if necessary.



**Place your name on ALL pages NOW.  
THANKS!**

**Name** \_\_\_\_\_

**Work area for problem 5, part (a) continued.**

- b. A general rule of thumb is that the number of calls **per day** from a region is equal to 1 call per 10,000 people in the region. ***Recall that the observations are for 10 days and that the population is measured in 1,000s of people.*** Based on the results above, test the following hypothesis:

$H_0$ : Calls per day are 1 per day per 10,000 people

$H_1$ : Calls per day are NOT 1 per day per 10,000 people

$\alpha = 0.05$

**Note that you will have to translate the null hypothesis and the alternate hypothesis into appropriate hypotheses about either  $\beta_1$  or  $\beta_0$  and then perform the test.**