MATH 145 - Elementary Statistics

Instructions: Please include all relevant work to get full credit. Encircle your final answers.

- 1. An article in *Professional Geographer* (Feb. 2000) investigated whether there has been a significant suburban-to-urban shift in the location of major sport facilities. In 1985, 40% of all major sport facilities were located downtown, 30% in central city, and 30% in suburban areas. To check if these proportions remain at this levels, 113 major sports franchises that existed in 1997 were randomly selected and of the 113, 58 were built downtown, 26 in a central city, and 29 in a suburban area.
 - a. Formulate the appropriate null and alternative hypotheses for the test to determine whether the proportions of major sports facilities in downtown, central city, and suburban areas in 1997 are the same as in 1985.
 - **b.** Using $\alpha = 0.05$, define your rejection rule.
 - **c.** If the null hypothesis, in part (a), is true, how many of the 113 sports facilities in 1997 would you expected to be located in downtown, central city, and suburban areas, respectively? [4]
 - **d.** Compute the observed value of the appropriate test statistic.

e. Do you reject the null hypothesis? Write a practical conclusion.

[3]

[3]

[5]

2. Clinical Test of Lipitor. The cholesterol-reducing drug Lipitor consists of atorvastatin calcium, and results summarizing headaches as an adverse reaction in clinical tests are given in the table (based on data from Parke-Davis). Using a 0.01 level of significance, test the claim that getting a headache is independent of the amount of atorvastatin used as a treatment.

	Placebo	10 mg	20/40 mg	80 mg	Total
Headache	19	47	8	6	
No Headache	251	816	107	88	
Headache					
Total					

a. Formulate the appropriate null and alternative hypotheses for this problem.

b. Using $\alpha = 0.01$, define your rejection rule.

c. Compute the observed value of the appropriate test statistic.

d. Do you reject the null hypothesis? Write a practical conclusion.

3. The USGA wants to compare the mean distances associated with four different brands of golf balls when struck with a driver. A completely randomized design is employed, with Iron Byron, the USGA's robotic golfer, using a driver to hit a random sample of 11 balls of each brand in a random sequence. The distances were recorded and a summary of the results is given below.

Brand	n_i	Average (\bar{x}_i)	Sample Variance (s_i^2)
A	11	250.78	22.42
В	11	261.03	14.95
С	11	269.95	20.26
D	11	249.32	27.07

a. Formulate the most appropriate null and alternative hypotheses for this problem. Define clearly the parameters you used in the null hypothesis. [4]

[3]

[2]

[10]

[3]

c. Using $\alpha = 0.01$, define your rejection rule.

- d. Do you reject the null hypothesis? Write a practical conclusion.
- e. What are the assumptions for the ANOVA procedure?
- 4. In the state of Florida, elementary school performance is based on the average score obtained by students on a standardized exam, called the Florida Comprehensive Assessment Test (FCAT). An analysis of the link between FCAT scores and socio-demographic factors was published in the *Journal of Educational and Behavioral Statistics* (Spring 2004). Part of the data on average math and reading FCAT scores of third-graders, as well as the percentage of students below the poverty level, for a sample of 22 Florida elementary schools are listed in the table.

Elementary School	1	2	3		21	22
FCAT-Math (y)	166.4	159.6	159.1	• • •	182.8	186.1
FCAT-Reading	165.0	157.2	164.4	• • •	181.6	183.8
% Below Poverty (x)	91.7	90.2	86.0		26.5	13.8

$$\sum x = 1292.7 \qquad \sum x^2 = 88,668.43 \qquad s_x = 24.6$$
$$\sum y = 3781.1 \qquad \sum y^2 = 651,612.45 \qquad s_y = 9.16$$
$$n = 22 \qquad \sum xy = 218,291.63$$

[3]

[3]

[3]

a. Compute the value of S_{xx} , S_{yy} , and S_{xy} .

- **b.** Determine the correlation coefficient r.
- c. Would you recommend using simple linear regression to model the relationship FCAT-Math and the percentage of students below the poverty level? Explain your answer. [3]
- ${\bf d.}$ Determine the regression line.
- e. If another Florida elementary school has 50% of its students below the poverty level, predict the average FCAT-math score of this school. [3]
- **f.** Determine the coefficient of determination r^2 . Explain the meaning of this quantity in the context of this problem. [3]
- **g.** Using $\alpha = 0.01$, test $H_0: b = 0$ vs. $H_1: b \neq 0$

h. Construct and interpret a 99% confidence interval for b.

[4]

[7]

[9]