Rosen

Math 214. Spring 2013. Final.

Show all work. For hypothesis testing, always state H_0, H_a and the test.

1) Find the mean, median, and standard deviation for the numbers: $2, 7, 6, 2, 1, 8, 11$. What are Q_1 and Q_3?

2) Sketch the frequency histogram for the following frequency distribution.

score	11-20	21-30	31-40	41-50
frequency	54	32	27	43

3) The drying time of a certain paint is normally distributed with mean 90 minutes and standard deviation of 6 minutes. What is the probability that a wall painted with this paint will need more than 100 minutes to dry?

4) A random sample of 200 egg cartons in a large supermarket found that 25 cartons had at least one broken egg. Find a 95% confidence interval for the proportion of egg cartons with at least one broken egg.

5) If we were designing the study of question 4 from scratch, how large a sample would be needed to have a margin of error $\leq .04$.

6) A random sample of 12 fisherman in Homser Lake, Oregon found a mean catch of 7.36 with standard deviation 4.03. At the 5% significance level test the claim that the catch differs from its historical average of 8.8.

7) In a study to evaluate the effectiveness of peer tutoring, the average score for 20 subjects in a control group on a vocabulary test was 349.2 with a standard deviation of 26.1. For a peer tutored group of 24 children the average score was 358.4 with a standard deviation of 19.5. At the 10% significance level, can we say that peer tutoring has any effect on test performance?

8) The following data is from 5 small cities, where y is the death rate per 1000 residents and x is the per capita income in thousands of dollars.

x (income)	8.6	9.3	10.1	8	8.3
y (death rate)	8.4	7.6	5.4	10.6	8.3

What is the correlation r? Find the equation of the least squares regression line for y as a function of x. Assuming that our data is from a random sample, find a 95% confidence interval for the slope of the regression line for the population.

9) What does your regression line predict would be the death rate in a city where the per capita income is $9,000?

10) The following table shows age distribution and location of a random sample of buffalo in Yellowstone Park.

Age	District I	District II	District III
Calf	64	15	17
Yearling	82	42	30
Adult	68	35	12

Use the χ^2 test to determine if age distribution and location are independent at the 5% significance level?

PLEASE TURN OVER

Some useful formulas

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}, \qquad s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$$

$$\sum x_i p_i, \qquad \sum_{i=1}^{n}(x_i - \mu_X)^2 p_i, \qquad \mu_{X^2} - (\mu_X)^2$$

$$np, \qquad \sqrt{np(1-p)}, \qquad z = \frac{x - \bar{x}}{s}$$

$$\bar{X} \pm z^* \sigma/\sqrt{n}, \qquad \bar{X} \pm t^* s/\sqrt{n}, \qquad \hat{p} \pm z^* \sqrt{\hat{p}(1-\hat{p})/n},$$

$$\frac{\bar{X} - \mu_0}{s/\sqrt{n}}, \quad \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}, \quad \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}}$$

$$\left(\frac{z^* \sigma}{m}\right)^2, \quad \frac{1}{4}\left(\frac{z^*}{m}\right)^2$$

$$\frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

$$\frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}, \qquad \bar{y} - b_1 \bar{x}, \qquad t = \frac{b_1}{SE_{b_1}}$$

$$\sum_{i=1}^{n}(x_i - \bar{x})^2 = s_x^2(n-1), \qquad \sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2} = s_x\sqrt{n-1}$$

$$SE_{b_1} = \frac{s}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}}, \qquad SE_{b_0} = s\sqrt{\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}}$$

$$SE_{\hat{\mu}} = s\sqrt{\frac{1}{n} + \frac{(x^* - \bar{x})^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}}, \qquad SE_{\hat{y}} = s\sqrt{1 + \frac{1}{n} + \frac{(x^* - \bar{x})^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}}$$

$$E_i = \frac{\text{row sum } \times \text{column sum}}{n}$$

$$\chi^2 = \sum (O_i - E_i)^2 / E_i, \qquad (r-1)(c-1)$$