Math 214. Fall 2011. Final.

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

1) Find the mean, median, sample variance and standard deviation for the numbers: $2, 5, 4, 1, 1, 7, 12$. What are Q_1 and Q_3?

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

score	10-10	11-20	21-30	31-40
frequency	16	12	27	13

3) The drying time of a certain paint is normally distributed with mean 90 minutes and standard deviation of 18 minutes. What is the probability that a wall painted with this paint will need more than 120 minutes to dry?

4) A random sample of 14 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.

5) In a study to evaluate the effectiveness of peer tutoring, the average score for 30 subjects in a control group on a vocabulary test was 349.2 with a standard deviation of 56.6. For a peer tutored group of 34 children the average score was 368.4 with a standard deviation of 39.5. At the 5% level of significance, test the claim that the peer tutored group performed better on the vocabulary test.

6) A random sample of 250 egg cartons in a large supermarket found that 40 cartons had at least one broken egg. Find a 95% confidence interval for the proportion of egg cartons with at least one broken egg.

7) If we were designing the study of question 6 from scratch, how large a sample would be needed to have a margin of error $\leq .05$.

8) The following data is from 6 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	8.7
y (death rate)	8.4	7.6	5.4	10.6	8.3	9.3

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 90% confidence interval for the slope of the regression line for the population. How many degrees of freedom are there?

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

10) A random sample of employees in three different branches of a life insurance company describes the sales volume per employee in hundreds of thousands of dollars.

Branch I	7.2	6.4	10.1	9.9	11	10.6
Branch II	8.8	9.8	10.7	11.1		
Branch III	6.9	8.7	10.5	11.4		

We wish to test the hypothesis that all branches have the same mean. State the null hypothesis and alternate hypothesis. What is the P-value? What would you conclude at the 1% significance level?

PLEASE TURN OVER

Some useful formulas

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

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

$$np, \qquad \sqrt{np(1-p)}$$

$$z = \frac{x - \bar{x}}{s}$$

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

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

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

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

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

$$\frac{r s_y}{s_x}, \qquad \bar{y} - b_1 \bar{x}, \qquad t = \frac{b_1}{SE_{b_1}}$$

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

$$s\sqrt{\frac{1}{n} + \frac{(x^* - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}}$$