MA 486-4A (Statistics), Dr Chernov — Final exam — Thu, Apr 28, 2005

Student's name ____________________

1. (15 pts) Given random sample

24.2, 35.7, 22.1, 11.5, −23.2, 40.6, 28.7, 39.1, 28.5, 20.4, 16.2, 33.7, 12.6, 29.9, 96.1

(a) find a point estimate for the median, m;

Answer: 28.5

(b) find point estimates for the quartiles, $\pi_{0.25}$ and $\pi_{0.75}$;

Answers: 16.2 and 35.7

(c) find an approximate 97% confidence interval for the median, m. Give the exact confidence level;

Answers: $(16.2, 35.7)$ at 96.48%

(d) Find $P(12.6 < \pi_{0.6} \leq 39.1)$;

Answer: $P(3 \leq b(15, 0.6) \leq 12) = 0.9716$.

2. (10 pts) Estimate the main effect, the three two-factor interactions, and the three-factor interaction in a 2^3 factorial design experiment. The data, in the canonical order (see Section 20.6), are: $x_1 = 6$, $x_2 = 6.5$, $x_3 = 4.3$, $x_4 = 3$, $x_5 = 4$, $x_6 = 5$, $x_7 = 2$, and $x_8 = 3$. Construct an approximate $q - q$ plot to see if any of these effects seem to be significantly larger than the others.

Partial answers: $[A] = 0.15$, $[B] = -1.15$, $[C] = -0.725$, $[AB] = -0.225$, $[AC] = 0.35$, $[BC] = 0.15$, $[ABC] = 0.225$.

3. (15 pts) In a linear regression problem $y = \alpha + \beta(x - \bar{x})$, 50 data points are observed and the following accumulated values are obtained:

$$\sum x_i = 50,\ \sum y_i = 100,\ \sum x_i^2 = 60,\ \sum y_i^2 = 240,\ \sum x_i y_i = 118$$

(a) Find $\hat{\alpha}$, $\hat{\beta}$, and $\hat{\sigma^2}$.

Answers: $al\hat{p}ha = 2$, $\hat{\beta} = 1.8$, $\hat{\sigma^2} = 0.15$.

(b) Find 98% confidence intervals for α, β, and σ^2.

(c) Test the hypothesis H_0: $\beta = 0$ against H_1: $\beta > 0$ at the level 95%. What is the p-value?

Answers: test statistic is $T \approx 14$, p-value is practically zero.

(Bonus) Find a 90% prediction interval for y_{exp} at the point $x = 2$.

4. (15 pts) Each of three cars is driven each of four different brands of gasoline. The number of miles per gallon driven for each of $3 \times 4 = 12$ combinations is recorded in the table below.

	Gasoline			
Car	1	2	3	4
1	22	21	23	22
2	22	17	21	16
3	19	16	22	19

Apply the two factor analysis of variance. Compute $SS(E)$, $SS(A)$, $SS(B)$. Test the hypotheses that the first factor (the car) affects the gas mileage. Test the hypotheses that the second factor (the gasoline) affects the gas mileage. Use the level $\alpha = 5\%$.

Answers: $SS(A) = 24$, $SS(B) = 30$, $SS(E) = 16$.

$F_A = 4.49 > F_{0.05}(2, 6) = 5.14$, accept H_A.

$F_B = 3.75 > F_{0.05}(3, 6) = 4.76$, accept H_B.

5. (15 pts) The following numbers were generated by a computer program:

$$2.8,\ 1.3,\ -0.6,\ 1.8,\ 0.5.$$

(a) Sketch an empirical distribution function. Show the boundaries of the 90% confidence band.

(b) Test the hypothesis that the program generates values of the normal random variable $N(1,4)$; use the Kolmogorov-Smirnov test at the level $\alpha = 10\%$.

Answer: $D = 0.2119 < d = 0.51$, accept the null hypothesis.

(c) Give the formula for the p-value. Compute it approximately by using the first two terms in the series.

6. (10 pts) Two samples from random variables X and Y were recorded:

$$X: \ 67, \ 24, \ 95, \ 89, \ 49, \ 57$$

$$Y: \ 27, \ 65, \ 33, \ 71, \ 15, \ 40, \ 28, \ 51, \ 24$$

Use the Wilcoxon test (for two samples) to test the hypothesis $H_0 : m_X = m_Y$ against $H_1 : m_X > m_Y$. Compute the p-value. Which hypothesis would you accept? What would change if the alternative was $H_1 : m_X \neq m_Y$?

Answers: $W = 58.5$, $\mu = 72$, $\sigma^2 = 72$, $Z = -1.59$, p-value is 0.0559. If $H_1 : m_X \neq m_Y$, then the p-value would be 0.1118.

7. (10 pts) Two samples from random variables X and Y were observed:

$$X: \ 8, \ 10, \ 5, \ 11, \ 9,$$

$$Y: \ 3, \ 7, \ 1, \ 2.$$

Use the run test to test the hypothesis $H_0: F_X = F_Y$ against $H_1: F_X \neq F_Y$. Compute R and the (exact) p-value. Also, compute the p-value by normal approximation. Which hypothesis would you accept?

[Bonus] Test the same hypothesis by the Kolmogorov-Smirnov test for two samples at the level $\alpha = 0.2$.

Answers: $R = 4$, the exact p-value is 0.262, by normal approximation the p-value is $\Phi(-0.75) = 0.227$.

8. (10 pts) It is claimed that the median price of houses in Birmingham area is \$180,000. A random sample of houses on the market gives the following prices (in thousands of dollars):

$$223,\ 145,\ 98,\ 195,\ 137,\ 319,\ 268,\ 480,\ 162,\ 206,\ 149,\ 352,\ 177$$

(a) Test the hypothesis $H_0 : m = 180$ against $H_1 : m \neq 180$ at the level 10%. Compute the p-value. Use the Wilcoxon test. What would change if the alternative was $H_1 : m > 180$?

(b) Explain the formula for σ^2 that you are using.

Answers: $W = 28$, $\sigma = 28.26$, $Z \approx 1$. The p-value is 0.32. If $H_1 : m > 180$, then the p-value would be 0.16.