MA 486-12 (Statistics), Dr Chernov Final exam Thu, June 3
Show your work. Every problem is 4 points. Full credit is given for 8 out of 9 problems

1. Let X_1, X_2, X_3, X_4 be four independent random variables that have normal distributions $N(\mu_i, \sigma^2)$. Test the hypothesis

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$$

at the level $\alpha = 5\%$. The observed data are given in the table below:

X_1 :	12	8	9	11
X_2 :	10	13	11	14
X_3 :	14	7	12	11
X_4 :	6	8	10	12

Compute $SS(E)$, $SS(T)$, and use the F-test.

Some answers:

$SS(E) = 66$, $SS(T) = 20$, $F = (20/3)/(66/12) = 1.21$. This is less than $F_{0.05}(3, 12) = 3.49$, so accept H_0.

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

	Gasoline				
Car	1	2	3	4	
1	18	16	19	19	
2	22	20	24	26	
3	17	18	14	15	

Compute $SS(E)$, $SS(A)$, $SS(B)$ and test the hypotheses

$$H_A : \alpha_1 = \alpha_2 = \alpha_3$$

and

$$H_B : \beta_1 = \beta_2 = \beta_3 = \beta_4$$

at the level $\alpha = 5\%$. Note: these are two different hypotheses, you need to do two tests!

Some answers:

$SS(E) = 30$, $SS(A) = 104$, $SS(B) = 6$.

H(A): $F = (104/2)/(30/6) = 10.4$ and $F_{0.05}(2, 6) = 5.14$. Reject.

H(B): $F = (6/3)/(30/6) = 0.4$ and $F_{0.05}(3, 6) = 4.76$. Accept.

3. Eight experimental data points are observed:

$$(-2,3),\ (-1,4),\ (0,1),\ (1,2),\ (2,1),\ (3,1),\ (4,0),\ (5,0)$$

Compute the Gaussian brackets $[X]$, $[Y]$, $[X^2]$, $[XY]$, $[Y^2]$. Estimate the parameters α and β of the regression line $y = \alpha + \beta x$. Draw a scatter plot, marking the data points and the regression line.

[Bonus] Compute the RSS (Residual Sum of Squares).

Some answers:

$[X] = 12$, $[Y] = 12$, $[X^2] = 60$, $[Y^2] = 32$, $[XY] = -3$.

$\alpha = 2.25$, $\beta = -0.5$, $RSS = 3.5$

4. In a regression problem, $n = 18$ data points are observed and the following values are found for the Gaussian brackets:

$$[X] = 6,\ [Y] = 12,\ [X^2] = 3,\ [XY] = 6,\ [Y^2] = 14$$

Find $\hat{\alpha}$, $\hat{\beta}$, and $\hat{\sigma^2}$. Find 95% confidence intervals for α, β, and σ^2.

[Bonus] Find a 95% prediction interval for Y when $x = 2/3$.

Some answers:

$\beta = 2$, $\alpha = 0$, $RSS = 2$, $\hat{\sigma^2} = 1/9$.

$t_{0.025}(16) = 2.12$

The CI for α is $[-0.177, 0.177]$

The CI for β is $[1.250, 2.750]$

The CI for σ^2 is $[0.0693, 0.289]$

The PI for Y is $[0.523, 2.143]$

5. (a) For 10 cars, the following measurements give the time in seconds for the car to go from 0 to 60 mph:

$$7.5,\ 6.8,\ 9.1,\ 5.4,\ 7.7,\ 6.1,\ 5.8,\ 7.0,\ 8.5,\ 5.1$$

Find a 97.86% confidence interval for the median, m. Use table 10.1.

(b) For a random sample of size $n = 300$ from an unknown distribution, find an approximate 90% confidence interval for the third quartile, $\pi_{0.75}$. Use normal approximation. Give the answer in the form $(y_{...}, y_{...})$.

Some answers:

$(y_2, y_9) = (5.4, 8.5)$

The CI is $[y_{213}, y_{238}]$

6. The median score in calculus tests is used to be 70 (out of 100). A new method of teaching calculus is proposed. It is used in a class of 16 students, and the scores in this class are

$$63, 79, 68, 82, 51, 94, 87, 60, 98, 74, 67, 85, 90, 59, 82, 96$$

Test the hypothesis $H_0 : m = 70$ against $H_1 : m > 70$.

(a) First, use the sign test, find the p-value. [You may use the exact binomial distribution or a normal approximation to binomial.]

(b) Then use the Wilcoxon test, again find the p-value.

Which hypothesis would you accept?

Some answers:

Sign test: $R = 6$, $P(b(16, 0.5) \leq 6) = 0.227$.

$W = 72$, $Z = 1.86$, p-value is 0.0314.

7. Test the hypothesis $H_0 : m_X = m_Y$ against $H_1 : m_X > m_Y$. The following data were observed:

$$X : \ 11,\ 3,\ 12,\ 8,\ 5,\ 15,\ 12,\ 9$$

$$Y : \ 10,\ 6,\ 2,\ 4,\ 10,\ 6,\ 4,\ 1$$

(a) Use the median test, compute the p-value (use the exact formula for the probabilities of the V-statistic).
(b) Use the Wilcoxon test (for two samples), again compute the p-value.
(c) Sketch a q-q plot.

Which hypothesis would you accept?

Some answers:

(a) $V = 2$, p-value is 0.066.

(b) $W = 50$, $Z = -1.89$, p-value is 0.0294

8. Test the hypothesis $H_0 : F_X = F_Y$ against $H_1 : F_X \neq F_Y$. The following data were observed:

$$X : \ 4,\ 12,\ 0,\ 3,\ 16,\ 15,\ 2,\ 17,\ 15,\ 4,\ 1,\ 6$$

$$Y : \ 7,\ 21,\ 5,\ 8,\ 9,\ 7,\ 26,\ 8,\ 11,\ 7,\ 22,\ 10$$

Use the run test, count the number of runs. Find ER and $\operatorname{Var} R$. Find the Z-value. Complete the test (chose a hypothesis) at the $\alpha = 5\%$ level.

[Bonus] Find the p-value of the test.

[Bonus] Compute the exact probability $P(R = 6)$.

Some answers:

$R = 6$, $ER = 13$, $\operatorname{Var} R = 5.74$

$Z = -2.92$, $z_\alpha = -1.645$, accept H_1

p-value is 0.0018.

$P(R = 6) = 0.002$

9. The following numbers were generated by a computer random number generator:

$$0.35,\ 0.21,\ 0.61,\ 0.03,\ 0.88,\ 0.49,\ 0.11,\ 0.69,\ 0.16,\ 0.85$$

Test the hypothesis that the computer generates a uniform random variable on the interval $(0, 1)$, i.e. $X = U(0, 1)$. Use the Kolmogorov-Smirnov test with Table VIII at $\alpha = 10\%$. Also, sketch an empirical distribution function.

[Bonus] Construct a 90% confidence band around the empirical distribution function.

Some answers:

$D = 0.19$, $d = 0.37$, accept H_0.