

SM230
Final Exam – Spring 2004

Directions: This exam is "closed book and closed notes" except that you may use your own (8 ½ by 11 both sides) resource sheet of paper with your handwritten notes and formulas. You may use your graphing calculator; you are not to borrow a class mate's calculator. You are to give answers as simple fractions or decimal numbers. Please show all work and answers in you blue books or on your solution sheets. Box your answers. Show all work, including which calculator programs or commands you used to solve the problem.

1. The amount of time it takes a restricted MID to clean the head has a normal distribution with a mean of 18 minutes and a standard deviation of 2 minutes.
- a) What is the probability that it will take a restricted MID longer than 21.5 minutes to clean the head?

$$\begin{aligned} \text{Normal CDF: } LV &= 21.5 \\ UV &= \infty \\ \mu &= 18 \\ \sigma &= 2 \end{aligned}$$

0.040059

- b) 90% of all restricted MIDs can clean the head in less than how many minutes?

$$\begin{aligned} \text{Inverse Normal: Area} &= 0.90 \\ \mu &= 18 \\ \sigma &= 2 \end{aligned}$$

20.5631

- c) In one company we take a random sample of 8 restricted MIDs and find that the average time for all eight of these restricted MIDS to individually clean the head is 16.5 minutes with a sample standard deviation of 2.25 minutes. Determine a 95% confidence interval (CI) for the mean time to clean the head in this company.

$$\begin{aligned} \text{t-interval} \\ 16.5 \pm (2.3646)2.25/\sqrt{8} \\ \mathbf{(14.62, 18.38)} \end{aligned}$$

- d) Use the information in part c) to perform a hypothesis test of $H_0: \mu = 18$ versus $H_a: \mu \neq 18$.

$$\begin{aligned} \text{t-test} \\ t = \frac{16.5 - 18}{2.25/\sqrt{8}} = -1.8856 \end{aligned}$$

t-CDF
 LV = $-\infty$
 UV = -1.8856
 DF = 7

p-value = $2 * 0.05065 = 0.1013$

Do not reject H_0 , there isn't enough evidence to conclude that the mean is different from 18.

2. A study in a major city was done in order to investigate the incomes of business professionals who belong to Rotary International. Random samples of size 30 were taken from the membership roles. The sample mean income of those belonging to Rotary International was \$122,000 with a sample standard deviation of \$5,000.

- a) Construct a 90% CI for the overall (population) mean income of those belonging to Rotary International.

t-interval
 $122000 \pm 1.69913 (5000 / \sqrt{30})$

(120449, 123551)

- b) The president of the club tells you that the mean income is actually \$125,000. Based on the random sample of 30 members what do you think of the \$125,000 figure?

The \$125,000 figure is not in your 90% CI so I suspect either the president is either stretching the truth, or we just observed a rare event. In fact, the probability of seeing a sample mean of \$122,000 or less under the assumption that the mean is \$125,000 is 0.0013. Not likely!!

3. The following data is for 12 perch caught in a lake in Finland:

<i>Weight (grams)</i>	<i>Length (cm)</i>	<i>Width (cm)</i>	<i>Weight (grams)</i>	<i>Length (cm)</i>	<i>Width (cm)</i>
5.9	8.8	1.4	300	28.7	5.1
100	19.2	3.3	500	30.1	4.6
110	22.5	3.6	685	39.0	6.9
120	23.5	3.5	650	41.4	6.0
150	24.0	3.6	820	42.5	6.6
145	25.5	3.8	1000	46.6	7.6

The data is in your calculator as follows:

ex1044a – Observation #
 ex1044b – Weight
 ex1044c – Length
 ex1044d – Width

- a) Perform a multiple regression to predict Weight from Length and Width. Your answer should include estimates of the model parameters, R^2 and your ANOVA table.

$$\hat{\text{Weight}} = -466.8 + 13.33 \text{ Length} + 94.61 \text{ Width}$$

$$R^2 = 0.91$$

ANOVA Table

Source	Df	SS	MS	F	p-value
Model	2	1,114,340	557,170	46.94	<0.0001
Error	9	106,836	11,870.7		
Total	11	1,221,180	569,041		

- b) What (if anything) does the intercept represent?

The weight of a fish with length = 0 and width = 0. Meaningless.

- c) Does your model in part a) accurately predict Weight? Explain.

Yes, $R^2 = 0.91$ is quite high, and the ANOVA p-value < 0.0001.

- d) Another fisherman catches a perch that has a Length of 21.2cm and a Width of 3.4cm. Use your model to construct a 95% CI for the predicted Weight of his perch.

Multi-regression 95% prediction interval.
(-125, 399.8)

- e) Test the hypothesis $H_0 : \beta_2 = 0$ vs. $H_a : \beta_2 \neq 0$ where β_2 is the parameter for Width. What can you conclude from your test?

t-statistic = 1.0417
p-value = 0.3247

The parameter for width is insignificant in predicting weight; however, when using length and weight, the model works quite well. (See part c above)

4. A study was done to see if there was a link between exercise and healthy bones.

The data is in your calculator as follows:

ex1223a – Observation #
ex1223b – Group (0 – Control, 1 – Low Jump, 2 – High Jump)
ex1223c – Not needed
ex1223d – Bone density (mg/cm^3)

Note: You can assume the population std. deviations are equal for all 3 groups.

- a) Is there a difference between the average bone density between the control group and the Low Jump group? (Your answer should include the p-value for your test and your conclusion.)

$$H_0 : \mu_{control} = \mu_{LowJump}$$

$$H_a : \mu_{control} \neq \mu_{LowJump}$$

2-sample t-test

$$t = -1.07606$$

$$p\text{-value} = 0.2977$$

There is not enough evidence to conclude that there is a difference between the groups.

- b) Construct a 95% CI for the average bone density difference (High Jump group – Control Group) between the High Jump group and the control group.

2 – sample t interval

$$(16.01, 59.19)$$

- c) Is there a difference between the average bone densities of all three groups? (Your answer should include the p-value for your test and your conclusion.)

$$H_0 : \mu_{control} = \mu_{LowJump} = \mu_{HighJump}$$

$$H_a : \text{Not all the means are equal}$$

$$F = 10.88$$

$$p\text{-value} = 0.0003$$

There is enough evidence to suggest that the mean bone densities are different.

5. Two groups were sampled to see who they might vote for in the November election (George W. Bush or John F. Kerry). The results are listed in the following table:

n1 - # of people in group 1

n2 - # of people in group 2

x1 - # of people who said they would vote for George W. Bush in group 1

x2 - # if people who said they would vote for George W. Bush in group 2

n1	124
x1	68
n2	133
x2	70

- a) Test the hypothesis: $H_0: p_1 = 0.5$ vs. $H_a: p_1 > 0.5$. (Your answer should include the p-value for your test and your conclusion.)

1-proportion Z test

$$z = 1.07763$$

$$p\text{-value} = 0.1406$$

There is enough evidence to conclude that the percentage of bush supporters in group 1 exceeds 50%.

- b) Construct a 95% CI for p_1 .

1- Proportion Z interval

$$\hat{p} = 0.5484$$

$$(0.46, 0.64)$$

- c) Construct a 90% CI for difference in percentages ($p_2 - p_1$) of Bush supporters.

2 – proportion Z interval

$$\hat{p}_{\text{diff}} = 0.0221$$

$$(-0.08, 0.12)$$

- d) Test to see if there is a difference between the percentages of Bush supporters. (use a 10% level of significance)

2 – proportion Z interval

$$H_0 : p_2 - p_1 = 0$$

$$H_a : p_2 - p_1 \neq 0$$

$$p\text{-value} = 0.7229$$

There is not enough evidence to conclude that the difference between the percentage of Bush supporters in the groups is different from 0.

- e) Compare your answers to parts c) and d). What do you conclude?

Both answers back each other up. The confidence interval in part c) includes some negative values as well as some positive values indicating that sometimes p_1 is bigger than p_2 and vice versa. Part d) tells us we can't reject the hypothesis that the difference is not zero.

6. One of your classmates has just completed a survey using a random sample from the brigade. They asked each midshipman whether they thought their summer training was preparing them adequately to be Naval officers. They were allowed to answer the question by stating that whether the training was excellent, adequate or inadequate. The tabulated results separated the answers by class (plebe through 1st class).

- a) You want to see if there is a relationship between response (excellent, adequate or inadequate) and class (plebe, youngster, etc.). Which statistical test should you perform?

Chi-squared test

- b) The value of the appropriate test statistic was calculated to be 13.25. What are your conclusions?

H_0 : There is not a relationships between class of midshipmen and response (excellent, adequate, inadequate). (i.e. the two are independent)

H_a : There is a relationship

$$\chi^2 = 13.25$$

$$Df = (\text{rows} - 1)(\text{columns} - 1) = (4 - 1)(3 - 1) = 3(2) = 6$$

Chi-2 CDF:

$$LV = 13.25$$

$$UV = \infty$$

$$Df = 6$$

$$p\text{-value} = 0.0392$$

The small p-value indicates that there is a relationship between the response and midshipmen class.

7. One of your classmates was trying to determine if there was a difference between family incomes among different classes in the brigade (plebe through 1st class). To do this, they took random samples of each class; the results are summarized in the following table.

Class	Number Sampled	Average Family Income
Plebe	25	\$103,000
Youngster	20	\$112,000
2 nd Class	28	\$96,000
1 st Class	22	\$85,000

- a) To see if there is a difference between average family income among the classes, which statistical test should you perform?

One-way ANOVA.

- b) What additional information do you need?

The sample standard deviations.

- c) The additional information was provided, and the appropriate test statistic was calculated to be 2.66. What are your conclusions?

$$H_0 : \mu_{plebe} = \mu_{youngster} = \mu_{2ndClass} = \mu_{1stClass}$$

H_a : Not all the means are equal.

$$F = 2.66$$

$$Df \text{ Numerator} = \text{groups} - 1 = 3 - 1 = 2$$

$$Df \text{ Denominator} = \text{sample points} - \text{groups} = 95 - 3 = 92$$

$$p\text{-value} = 0.0753$$

The small p-value indicates that the group means are not all equal.