ECON 395 Fall 2011 Practice Exam SOLUTION

INSTRUCTIONS:
Show your formula and some steps in your calculations to receive points. You may use non-programmable calculators during the exam.

1. Three airlines serve a small city in Ontario. Airline A has 50% of all the scheduled flights; Airline B has 30% and Airline C the rest. 80% of Airline A's flights are at the scheduled time, while for Airline B it is 65% are at the scheduled time, and 40% for Airline C.

Departure and Arrival as per Schedule

		Not on time (0)	On Time (1)	f(X)
Airline	A	0.1	.4	.5
	B	.105	0.195	.3
	C	.12	0.08	0.2
	f(Y)	0.325	0.675	1

a) Fill in the missing elements **in the table**.
b) What is the probability that a plane coming in to land belongs to Airline B and it is not on time?

$P(B\ and\ Not\ on\ time) = 0.105$ From table above

The probability that a plane coming in to land belongs to Airline B and it is not on time is 0.105 or 10.5%.

c) A plane has just left, and it is on time. What is the probability the plane is from Airline A?

$P(A|On\ time) = \frac{P(A\ and\ On\ time)}{P(On\ time)} = \frac{0.4}{0.675} = 0.5926$

The probability that a plane that has just left on time is from Airline A is 0.5926 or 59.26%.

d) Is the Airline and whether the plane is "on time" independent? Prove your answer.

We can test for independence two ways,

1. If, for all Airlines and Schedule times, the conditional probability is equal to the marginal probability for the undetermined event. For example $P(A|On\ time) = ??P(A)$. In this case $P(A|On\ time) = 0.5926 \neq 0.50 = P(A)$. So, we can conclude that the Airline and the "On schedule" are not independent.
2. If, for all Airlines and Schedule times, the multiplication of the marginal probabilities is equal to the joint probability of the two event. In this case does $P(C\ and\ On\ time) = ???P(C) * P(On\ time)$ $P(C\ and\ On\ time) = 0.08 \neq 0.135 = 0.675 * .2 = P(C) * P(On\ time)$ So, we can conclude that the Airline and the "On schedule" are not independent.

2. Short answer to **TWO** of the following questions or statements. **[14]**

ECON 395 Fall 2011 Practice Exam SOLUTION

a) Explain the difference between an estimator and an estimate. Which one is a random variable? What is a "biased" estimator? What is a relatively efficient estimator?

While an estimator is a mathematical rule or formula that tells us how to combine data to get an estimate for some unknown population parameter, an estimate is just a number derived from using sample data in an estimator. The estimator is a random variable, but a number is the outcome of the random variable.

An estimator is "biased" if, on average, you do not get the true unknown parameter value when you use it. If you are centering on some other value we can measure the bias. For example: let $\hat{\theta}$ be our estimator for the population parameter θ. Then if $E(\hat{\theta}) = \theta + c$, where c is a constant, the bias would equal the c.

A relatively efficient estimator is the estimator with the smallest variance when compared to other unbiased estimators of the same unknown population parameter. We do not compare estimators for different population parameters to find the relatively efficient estimator.

b) What is a Type I error and a Type II error? Why do we get errors in our statistical results?

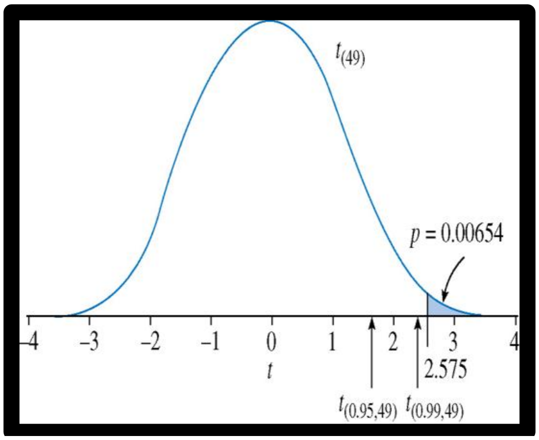
A Type I error is to reject the null hypothesis, H_0, when H_0 is true.
A Type II error is to not reject the null hypothesis, H_0, when H_0 is false.
Getting errors in our statistical results is unavoidable. When we reject, or do not reject, a null hypothesis there is a chance that we may be making a mistake.
We can control the probability of a Type I error by choosing the level of significance of a test.
The magnitude of a probability of a Type II error is not under our control and cannot be computed because it depends on the true value of the unknown parameter we are testing. If the true unknown value of the parameter is close to the value in the null hypothesis the probability of a Type II error increases. The probability of a Type II error varies inversely with the level of significance, which is the probability of a Type I error. So, there is a relationship between the probability we select for a Type I error and the probability of a Type II error.

c) Explain the term "p-value" and how to use a p-value to determine the outcome of a hypothesis test; provide a sketch to illustrate.

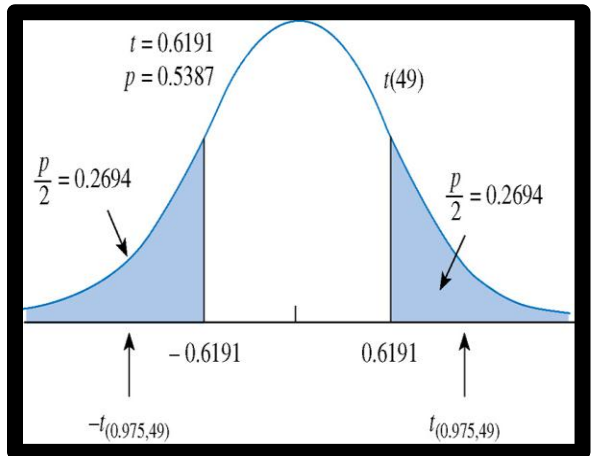
The term "p-value" refers to the probability of getting a test statistic value equal to or more extreme than the calculated value. For example $P(t_{0.95, 49} \geq 2.575) = 0.00654$.
***p*-value rule**: Reject the null hypothesis when the *p*-value is less than, or equal to, the level of significance, α. That is, if $p \leq \alpha$ then reject H_0. If $p > \alpha$ then do not reject H_0. A small p-value indicates that there is lots of evidence to support the alternative hypothesis when used in a hypothesis test.
How the p-value is computed depends on the alternative. If t is the calculated value [not the critical value t_c] of the t-statistic with N−k degrees of freedom, then:

- if $H_1: \mu > c$, p = probability to the right of t
- if $H_1: \mu < c$, p = probability to the left of t
- if $H_1: \mu \neq c$, p = <u>sum</u> of probabilities to the right of |t| <u>and</u> to the left of –|t|

A one-tail, right tail, test.



A two tail test.



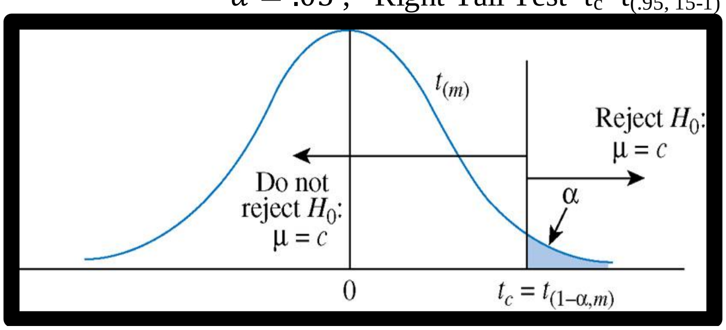
3. How much money do winners go home with from the television quiz show *Jeopardy?* To determine an answer, a random sample of winners was drawn and the amount of money each won was recorded. The mean of the sample of 15 winners was \$26,050 and the standard deviation was \$8,782.6.
Test the hypothesis, at a 5% significance level, that the mean winnings is greater than \$23,000.00

i. State the null and alternative hypothesis **[3]**

$$H_0: \mu = \$23{,}000$$
$$H_A: \mu > \$23{,}000$$

ii. Draw the distribution, identify and state the critical value on the graph and shade the rejection region, clearly labeling your drawing. **[5]**

$\alpha = .05$, Right Tail Test $t_c = t_{(.95,\, 15-1)} = 1.761$



iii. Calculate the test statistic. **[4]**

$$t = \frac{\bar{X} - \mu}{s/\sqrt{n}} = \frac{26{,}050 - 23{,}000}{8782.6/\sqrt{15}} = 1.345$$

iv. What is your decision rule, and your decision? (State your conclusion, in terms of the problem, in words)

Two Solutions:

1. Decision Rule: Reject H_0 if $t > t_c$
   $1.345 < 1.761$ Do not reject H_0

There is insufficient evidence in this sample to reject the hypothesis that the mean winnings in the television quiz show *Jeopardy* are $23,000.

2. Decision Rule: Reject H_0 if *p-value* < $\alpha = .05$
$P(t_{14} > 1.345) = 0.1$ 0.1> 0.05 Do not reject H_0
There is insufficient evidence in this sample to reject the hypothesis that the mean winnings in the television quiz show *Jeopardy* are $23,000.

4. Many people assume that a person with more experience should receive a higher wage rate. The question of interest then would be, "How much does experience affect wage rates?" A simple specification for this model is given by

$$WAGE_i = \beta_1 + \beta_2 EXPER_i + \varepsilon_i, \qquad (1)$$

where $WAGE_i$ is in dollars per hour and $EXPER_i$ is measured in years (rounded to nearest year). (Assume the error term is i.i.d. $N(0,\sigma^2)$.)
A sample of data of 15 people gives you the following information:

$$\sum_{i=1}^{15} WAGE_i = 150.14 \qquad \sum_{i=1}^{15} EXPER_i = 282$$

$$\sum_{i=1}^{15}(EXPER_i - \overline{EXPER})(WAGE_i - \overline{WAGE}) = 597.778$$

$$\sum_{i=1}^{15}(WAGE_i - \overline{WAGE})^2 = 280.9349; \qquad \sum_{i=1}^{15}(EXPER_i - \overline{EXPER})^2 = 2360.4$$

$$\sum_{i=1}^{15}(WAGE_i - \hat{WAGE}_i)^2 = 129.5459$$

a. Determine the ordinary least squares regression line.

$$b_2 = \frac{\sum_{i=1}^{N}(X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{N}(X_i - \bar{X})^2} = \frac{597.778}{2360.4} = 0.2533$$

$$b_1 = \bar{Y} - b_2\bar{X} = \left(\frac{150.14}{15}\right) - \left(0.2533 * \left(\frac{282}{15}\right)\right) = 10.0093 - (4.7620) = 5.2473$$

The equation to the regression line is $\hat{WAGE}_i = 5.2473 + 0.2533 EXPER_i$

b. Interpret the coefficients of the regression line.

If a person has no experience their starting wage rate, on average, would be $5.25. Since we do not know the range of the values for experience in this dataset we should be very cautious using this interpretation.
As experience increases by one year the wage rate increases by $0.25, on average.

c. Using the equation you calculated in part (a), what is the average wage of a person with 19 years of experience?

$$\hat{WAGE}_i = 5.2473 + 0.2533 EXPER_i = 5.2473 + (0.2533 * 19) = \$10.06$$

The average wage for a person with 19 years of experience is $10.06.

Since you are curious about your results you get a much larger sample, and you use Stata to get the following output:

regress WAGE EXPER

Source	**SS**	**df**	**MS**
Model	**????????**	1	868.6881
Residual	38112.809	998	38.1892
Total	38981.4974	999	39.0205

Number of obs	=	1000
F(1, 998)	=	22.7470
Prob > F	=	0.0000
R-squared	=	?????
Adj R-squared	=	0.0213
Root MSE	=	6.1797

WAGE	**Coef.**	**Std. Err.**	**t**	**P>t**	**[95% Conf.**	**Interval]**
EXPER	0.0824	**???????**	4.7694	0.000	0.0485	0.1163
_cons	8.6658	0.3787	22.8822	0.000	7.9227	9.4090

d. Using the information from your Stata regression, answer the question, “If a person has one more year of experience, how much does their wage rise?”

If a person has one more year of experience, their wage rises by $0.08, on average.