Ecn 102 - Analysis of Economic Data
University of California - Davis
Instructor: John Parman

January 28, 2010

Midterm 1

You have until 10:20am to complete this exam. Please remember to put your name, section and ID number on both your scantron sheet and the exam. Fill in test form A on the scantron sheet. Answer all multiple choice questions on your scantron sheet. Choose the single best answer for each multiple choice question. Answer the long answer questions directly on the exam. Keep your answers complete but concise. For the long answer questions, you must show your work where appropriate for full credit.

Name: **ID Number:** **Section:**

(POTENTIALLY) USEFUL FORMULAS

$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$

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

$CV = \frac{s}{\bar{x}}$

$skew = \frac{n}{(n-1)(n-2)}\sum_{i=1}^{n}(\frac{x_i-\bar{x}}{s})^3$

$kurt = \frac{n(n+1)}{(n-1)(n-2)(n-3)}\sum_{i=1}^{n}(\frac{x_i-\bar{x}}{s})^4 - \frac{3(n-1)^2}{(n-2)(n-3)}$

$\mu = E(X)$

$z^* = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}$

$t^* = \frac{\bar{x}-\mu}{\frac{s}{\sqrt{n}}}$

$Pr[T_{n-1} > t_{\alpha,n-1}] = \alpha$

$Pr[|T_{n-1}| > t_{\frac{\alpha}{2},n-1}] = \alpha$

$\sum_{i=1}^{n} a = na$

$\sum_{i=1}^{n}(ax_i) = a\sum_{i=1}^{n} x_i$

$\sum_{i=1}^{n}(x_i + y_i) = \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} y_i$

$s^2 = \bar{x}(1-\bar{x})$ for proportions data

$t_{\alpha,n-1} = TINV(2\alpha, n-1)$

$Pr(|T_{n-1}| \geq |t^*|) = TDIST(|t^*|, n-1, 2)$

$Pr(T_{n-1} \geq t^*) = TDIST(t^*, n-1, 1)$

(POTENTIALLY) USEFUL EXCEL OUPUT

TINV(.005,999)=2.81
TINV(.01,999)=2.58
TINV(.02,999)=2.33
TINV(.025,999)=2.24
TINV(.05,999)=1.96
TINV(.10,999)=1.65
TINV(.20,999)=1.28

TINV(.005,99)=2.87
TINV(.01,99)=2.63
TINV(.02,99)=2.36
TINV(.025,99)=2.28
TINV(.05,99)=1.98
TINV(.10,99)=1.66
TINV(.20,99)=1.29

SECTION I: MULTIPLE CHOICE (60 points)

1. Suppose we take a sample of 400 people and measure their heights. If we measure height in inches, the coefficient of variation will be:
   (a) Larger than if we measure height in meters.
   (b) Smaller than if we measure height in meters.
   (c) The same as when we measure height in meters.
   (d) (a), (b) or (c) could be true depending on the value of the sample mean.
2. Which of the following is not a random variable?
   (a) The sample mean.
   (b) The sample variance.
   (c) The population variance.
   (d) All of the above are random variables.
3. Suppose we can reject the null hypothesis that $\mu \geq 50$ at the 5% significance level. Which of the following statements is definitely true?
   (a) We can reject the null hypothesis that $\mu = 50$ at the 5% significance level.
   (b) We can reject the null hypothesis that $\mu \geq 50$ at the 10% significance level.
   (c) The value we obtained for t^* was positive.
   (d) The value we obtained for t^* was greater than the critical value for an upper one-tailed hypothesis test at the 5% significance level.
4. The sample variance of 100 observations of monthly unemployment rates will be ______ the sample variance of 100 observations of the unemployment rate covering the same time period for which each month's unemployment rate is an average of that month's rate and the previous two months' rates.
   (a) Less than or equal to.
   (b) Less than.
   (c) Equal to.
   (d) Greater than or equal to.
5. Which of the following is not a measure of central tendency?
   (a) The sample range.
   (b) The sample median.
   (c) The sample midrange.
   (d) The sample mean.
6. Suppose we have data on the interest rate, i_t. Which of the following could we use to get the percent change in the interest rate from year t to year $t+1$ (expressed as a decimal)?
   (a) $i_{t+1} - i_t$.
   (b) $ln(i_{t+1} - i_t)$.
   (c) $\frac{i_{t+1}}{i_t}$.
   (d) $ln(i_{t+1}) - ln(i_t)$.

7. Increasing the sample size used for a two-tailed hypothesis test will tend to:
   (a) Decrease the probability of a Type I error.
   (b) Decrease the probability of a Type II error.
   (c) Increase the probability of a Type I error.
   (d) Increase the probability of a Type II error.

8. Suppose that we have data on coin flips. The variable X is equal to one if the coin flip is heads and zero if the coin flip is tails. In a sample of 100 coin flips, the sample mean of X turns out to be .6. The value of the skewness for the sample will be:
   (a) Positive.
   (b) Negative.
   (c) Zero.
   (d) Not enough information.

9. Which of the following statements is not true?
   (a) Two random variables can have the same mean but different medians.
   (b) Two random variables can have the same mean but different variances.
   (c) The mean of the sum of two random variables is always equal to the sum of their means.
   (d) The variance of the sum of two random variables is always equal to the sum of their variances.

10. If the distribution of X is symmetric and reaches its highest point at the exact middle of the distribution:
    (a) The mode of X will be equal to the median of X.
    (b) The mean of X will be equal to the median of X.
    (c) The median of X will be at the exact middle of the distribution.
    (d) All of the above are true.

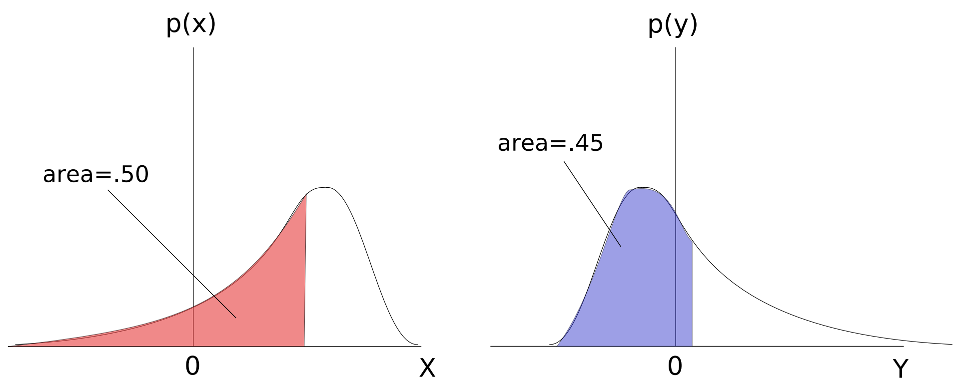
11. Which of the following would increase the probability of a Type I eror?
    (a) Decreasing the sample size.
    (b) Increasing the sample size.
    (c) Decreasing the significance level.
    (d) Increasing the significance level.

12. Suppose that the 95% confidence interval for the population mean hours of studying per week is $(6.5, 6.9)$. Which of the following statements is true?
    (a) You would reject the null hypothesis that $\mu \geq 6.5$ at the 10% significance level.
    (b) You would reject the null hypothesis that $\mu = 7.1$ at the 5% significance level.
    (c) You would reject the null hypothesis that $\mu \leq 6.9$ at the 5% significance level.
    (d) All of the above are true.

13. Suppose you have a dataset of the unemployement rate for 100 different cities for the month of December, with one observation per city. These are:
    (a) Cross-sectional data.
    (b) Panel data.
    (c) Time-series data.
    (d) Both (a) and (c).

14. Suppose that rather than the sample mean, we use another statistic that we'll call $\tilde{X}$ to estimate the population mean. If the distribution of $\tilde{X}$ is symmetric, what must be true for $\tilde{X}$ to be an unbiased estimator for the population mean?

    (a) The distribution of $\tilde{X}$ must get narrower as the sample size increases.
    (b) The variance of $\tilde{X}$ should not depend on the sample size.
    (c) The distribution of $\tilde{X}$ must be centered at the population mean.
    (d) The value of $\tilde{X}$ should approach the population mean as the sample size goes to infinity.



Use the figure above to answer questions 15 through 17. The graph on the left shows the distribution of random variable X and the graph on the right shows the distribution of random variable Y.

15. The median of X is ______ and the median of Y is ______.

    (a) Positive, positive.
    (b) Positive, negative.
    (c) Negative, negative.
    (d) Not enough information.

16. Which of the following statements is true?

    (a) $skewness_x > skewness_y$.
    (b) $skewness_y > skewness_x$.
    (c) $skewness_x < 0$ and $skewness_y < 0$.
    (d) $skewness_x = 0$.

17. The distribution of the sample mean of Y will be:

    (a) Right skewed.
    (b) Left skewed.
    (c) Symmetric.
    (d) Centered at 0.

18. Which of the following would narrow the confidence interval for the population mean?
    (a) Decreasing the sample size.
    (b) Decreasing the significance level.
    (c) Both (a) and (b) would narrow the confidence interval.
    (d) Neither (a) nor (b) would narrow the confidence interval.
19. The distribution of the sample mean of X (using a sample size of 10) has a:
    (a) Larger variance than the distribution of X.
    (b) Smaller variance than the distriubution of X.
    (c) Variance equal to the variance of X.
    (d) None of the above.
20. The probability of a Type I error when doing a two-tailed hypothesis test at a 5% significance level is:
    (a) Equal to the probability of a Type I error when doing a one-tailed hypothesis test at a 2.5% significance level.
    (b) Equal to the probability of a Type I error when doing a one-tailed hypothesis test at a 5% significance level.
    (c) Equal to the probability of a Type I error when doing a one-tailed hypothesis test at a 10% significance level.
    (d) Equal to the probability of a Type I error when doing a one-tailed hypothesis test at a 20% significance level.

SECTION II: SHORT ANSWER (40 points)

1. (14 points) A survey is given to 1000 college students asking them how many years they think it will take them to graduate from college. Responses ranged from three years to seven years. The distribution of students by their responses is given in the table below:

Years to graduate	Number of students
3	50
4	400
5	300
6	200
7	50

   (a) Calculate a 90% confidence interval for the proportion of students in the population who believe they will graduate in exactly four years.

   (b) Suppose a researcher wants to use these data to create a 95% confidence interval for the average number of years it takes students to graduate from college. At what value would this confidence interval be centered?

   (c) Why might the researcher reach an incorrect conclusion? Fully explain your answer including whether the researcher would be likely to overestimate or underestimate the average number of years it takes students to graduate.

2. (14 points) The true population mean of a random variable X is 50. Suppose that you draw a sample of 100 observations of X that has a sample standard deviation of 10 and you use this to test the following set of hypotheses using a 5% significance level:

$$H_o\colon\ \mu = 55$$

$$H_a\colon\ \mu \neq 55$$

   (a) Write down a formula for the test statistic you would use for the test. The only variable in your expression should be the sample mean, $\bar{x}$ (you should plug in the appropriate numerical values for any other variables or constants).
   (b) For what range of values of $\bar{x}$ would you end up committing a Type II error?

3. (12 points) For each scenario below, write down the null and alternative hypotheses you would use, the formula for the test statistic you would use, the critical value you would use, and the basis on which you would reject the null hypothesis. For the test statistic and the critical value, include specific numbers whenever possible.

   (a) You have a sample of 100 books. You want to test whether or not the average number of pages in a book is greater than 200 pages using a 5% significance level.

   (b) You take the temperatures of 100 people to test whether the average body temperature is equal to 98.6 degrees. You want to use a 5% significance level.

   (c) You have a sample of 100 people's wages. You want to use a lower one-tailed test whether the average wage is below $20 an hour using a 10% significance level.