Introduction to Econometrics (Econ 107)
Drake University, Spring 2006
William M. Boal

Signature: ____________________

Printed name: ____________________

MIDTERM EXAMINATION #1 VERSION B
"Introduction and Statistics Review"
February 7, 2006

INSTRUCTIONS: This exam is closed-book, closed-notes. You may use a calculator on this exam, but not a graphing calculator or a calculator with alphabetical keys. Point values for each question are noted in brackets. A table of the t-distribution is attached.

I. MULTIPLE CHOICE: Circle the one best answer to each question. Feel free to use margins for scratch work [3 pts each—42 pts total]

(1) Which of the following is *not* necessarily true?
a. $\sum(\alpha x_i) = \alpha \sum x_i$.
b. $\sum(x_i - \bar{x})^2 = \left(\sum x_i^2\right) - n\bar{x}^2$.
c. $\sum(x_i - \bar{x}) = 0$.
d. $\sum x_i = n\bar{x}$.
e. $\sum(x_i / y_i) = \left(\sum x_i\right) / \left(\sum y_i\right)$.

(2) Suppose we wish to fit the equation $y = \beta_1 + \beta_2 x$ to data by the method of least squares. This method minimizes which function of the data?
a. $f(\beta_1, \beta_2) = \sum(y_i - \beta_1 - \beta_2 x)$.
b. $f(\beta_1, \beta_2) = \sum|y_i - \beta_1 - \beta_2 x|$.
c. $f(\beta_1, \beta_2) = \sum(\beta_1 + \beta_2 x)^2$.
d. $f(\beta_1, \beta_2) = \sum(y_i - \beta_1 - \beta_2 x)^2$.
e. $f(\beta_1, \beta_2) = \sum\left(y_i^2 - (\beta_1 - \beta_2 x)^2\right)$.

The next two questions assume the following. Suppose X is a Bernoulli random variable, with Prob{X=1} = 0.8 and Prob{X=0} = 0.2.

(3) The mean or expected value of X is
a. one.
b. 0.8 .
c. 0.2 .
d. 0.16 .
e. zero.

(4) The variance of X is
a. one.
b. 0.8 .
c. 0.2 .
d. 0.16 .
e. zero.

(5) The correlation of any two independent random variables is
a. negative.
b. zero.
c. between negative one and one.
d. exactly one.
e. infinite.

(6) Which of the following distributions has a symmetric bell-shaped density function?
a. t distribution.
b. chi-square distribution.
c. F distribution.
d. all of the above have bell-shaped density functions.
e. none of the above have bell-shaped density functions.

The next two questions assume the following. Suppose a random sample of size n is drawn from some population. The population has mean μ and variance σ^2. Consider the sample mean, defined as

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \ .$$

(7) The variance of $\bar{X}$ is
a. zero.
b. σ^2 .
c. σ^2 / n .
d. σ^2 / n^2 .
e. σ^2 / (n-1) .

(8) The expected value of $\bar{X}$ is
a. zero .
b. μ .
c. μ / n .
d. μ / n^2 .
e. μ / (n-1) .

(9) Let $\hat{\theta}$ be an estimator of an unknown population parameter θ, with the property that the probability of $\hat{\theta}$ being more than any given distance from θ shrinks to zero, as the sample size increases without bound. The estimator $\hat{\theta}$ is said to be
a. unbiased
b. asymptotically unbiased.
c. consistent.
d. linear.
e. efficient.

(10) If an estimator is unbiased, then its mean square error (MSE)
a. must be greater than its variance.
b. must be equal to its variance.
c. must be less than its variance.
d. can be greater or less than its variance.
e. must be identically zero.

(11) Which method(s) for finding an estimator does *not* require us to know the formula for the underlying density function of the population?
a. method of maximum likelihood.
b. method of moments.
c. both of the above require us to know the formula for the density function.
d. neither of the above require us to know the formula for the density function.

(12) A wider confidence interval is obtained by
a. increasing the confidence level.
b. decreasing the confidence level.
c. increasing the sample size.
d. both (b) and (c).
e. none of the above.

(13) The probability that a test will correctly reject the null hypothesis when it is false is called the
a. critical point of the test.
b. power of the test.
c. size or significance of the test.
d. test statistic.
e. standard error.

(14) If the computed p-value for a test statistic is greater than the size of the test, we
a. can reject the null hypothesis.
b. cannot reject the null hypothesis.
c. cannot compute the test statistic.
d. answer cannot be determined from the information given.

II. PROBLEMS: Please write your answers in the boxes on this question sheet. Show your work and circle your final answers.

(1) [Least-squares calculation: 12 pts] Suppose the following three observations on x_i and y_i are given.

Observation (i)	x_i	y_i
1	4	3
2	8	5
3	6	7

a. Compute $\hat{\beta}_2$, the least-squares estimate of the slope of the line $y = \beta_1 + \beta_2 x$.

b. Compute $\hat{\beta}_1$, the least-squares estimate of the y-intercept of the same line.

c. Compute the three fitted values $\hat{y}_i$ of this least-squares estimated regression line.

d. Compute the three residuals $\hat{\varepsilon}_i$ of this estimated least-squares regression line.

(2) [Moments: 12 pts] Suppose X_1 and X_2 are random variables with the following moments.

$E(X_1) = 6$	$Var(X_1) = 9$	$Cov(X_1,X_2) = 0.75$
$E(X_2) = 5$	$Var(X_2) = 4$	

Now let $Y = 2X_1 + 3X_2$. Compute the following and circle your final answers.

a. Compute E(Y).

b. Compute Var(Y).

c. Compute SD(Y).

d. Compute $Corr(X_1,X_2)$.

(3) [Estimation: 12 pts] Suppose we wish to estimate the mean of a population using the following (peculiar) estimator applied to a random sample of 10 observations.

$$\hat{\mu} = -2 + \frac{1}{6}\sum_{i=1}^{10} x_i$$

Compute the following properties of the estimator under the assumption that the true population mean is $E(X_i) = 6$ and the true population variance is $Var(X_i) = 18$. Circle your final answers.

a. Compute $E(\hat{\mu})$.

b. Compute $Bias(\hat{\mu})$.

c. Compute $Var(\hat{\mu})$.

d. Compute $MSE(\hat{\mu})$.

(4) [Inference for arbitrary distribution, large sample: 18 pts] Suppose we wish to analyze the distribution of the number of children per family in a population. Let μ denote the unknown true population mean number of children per family. Observations X_i have been collected on 600 families, with the following summary values. Here, $\overline{X}$ is the sample mean.

$$\sum_{i=1}^{600} X_i = 1260 \qquad \sum_{i=1}^{600} \left(X_i - \overline{X}\right)^2 = 144$$

a. [3 pts] Is the population distribution discrete or continuous? Justify your answer.

b. [3 pts] Compute an unbiased estimate of μ.

c. [3 pts] Compute a 95% asymptotic confidence interval for μ.

d. [9 pts] Test the null hypothesis that $\mu = 2$ against the one-sided alternative hypothesis that $\mu > 2$, at 5% significance using an asymptotic test. Give
- the *value* of the test statistic
- the *critical point* from the appropriate table
- your conclusion: whether you reject the null hypothesis at 5% significance.

III. CRITICAL THINKING [4 pts] Suppose a researcher proposes to apply the following estimator for the population mean to a sample of n observations.

$$\hat{\mu} = \frac{1}{n-5}\sum_{i=1}^{n} x_i$$

What is wrong with this estimator? Is it biased? Is it asymptotically biased? Is it inconsistent? Justify your answer.

[end of exam]