Introduction to Econometrics (Econ 107)
Drake University, Spring 2010
William M. Boal

Signature: ______________________________

Printed name: ______________________________

MIDTERM EXAMINATION #1 VERSION A
"Introduction and Statistics Review"
February 12, 2010

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. Use margins for scratch work. [2 pts each–18 pts total]

(1) $\frac{d}{d\alpha}\sum_{i=1}^{n}(y_i - 3\alpha)^2 =$

a. $\sum_{i=1}^{n} -6\alpha x_i$.

b. $-6\sum_{i=1}^{n}(y_i - 3\alpha)$.

c. $2\sum_{i=1}^{n} y_i$.

d. $-2\sum_{i=1}^{n} 3$.

e. $\sum_{i=1}^{n}(y_i - 3\alpha)^2$.

(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. $f(\beta_1, \beta_2) = \sum(\beta_1 + \beta_2 x_i)^2$.

b. $f(\beta_1, \beta_2) = \sum(y_i - \beta_1 - \beta_2 x_i)$.

c. $f(\beta_1, \beta_2) = \sum\left(y_i^2 - (\beta_1 + \beta_2 x_i)^2\right)$.

d. $f(\beta_1, \beta_2) = \sum(y_i - \beta_1 - \beta_2 x_i)^2$.

e. $f(\beta_1, \beta_2) = \sum|y_i - \beta_1 - \beta_2 x_i|$.

(3) Suppose we wish to estimate the mean income of the population of all families in Iowa. Our estimator will be the sample mean, from a random sample of 100 families. So-called "classical statistics" treats the population mean as __________ and the estimator as __________.

a. a fixed constant … a random variable.
b. a random variable … a fixed constant.
c. a random variable … also a random variable.
d. a fixed constant … also a fixed constant.

(4) Suppose $\hat{\theta}$ is an estimator of an unknown population parameter θ. If $E(\hat{\theta}) = \theta$, then the estimator $\hat{\theta}$ is said to be

a. unbiased.
b. asymptotically unbiased.
c. best (or minimum-variance) unbiased.
d. consistent.

(5) Suppose $\hat{\theta}$ is an estimator of an unknown population parameter θ. If $\lim_{n\to\infty} \text{Prob}\left(\left|\hat{\theta}-\theta\right| > \delta\right) = 0$, then the estimator $\hat{\theta}$ is said to be

a. unbiased.
b. asymptotically unbiased.
c. best (or minimum-variance) unbiased.
d. consistent.

(6) The principle for finding an estimator that uses the joint density function of the sample is called the

a. method of maximum likelihood.
b. method of steepest descent.
c. method of scoring.
d. method of moments.

(7) A symmetric 90 percent confidence interval for an unknown parameter is necessarily ________________ a symmetric 95 percent confidence interval for the same parameter, computed from the same data.

a. narrower than
b. wider than
c. identical to
d. wider or narrower, depending on the true value of the parameter, than

(8) The probability that a test will mistakenly reject the null hypothesis when it is true 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.

(9) Suppose the p-value for a test statistic is 0.022. If the size of the test is 1 percent, 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. SHORT ANSWER: Please write your answers in the boxes on this question sheet. Use margins for scratch work.

(1) [Summation operator: 8 pts] Let $\bar{x} = \left(\frac{1}{n}\right)\sum_{i=1}^{n} x_i$ and consider the following equations. Write "TRUE" for equations that are necessarily true for any values of x_i and y_i. Write "FALSE" for equations that are not necessarily true.

a. $\sum(\alpha x_i) = \alpha \sum x_i$.	
b. $\sum(x_i - y_i)^2 = \sum x_i^2 - \sum y_i^2$.	
c. $\sum(x_i - \bar{x}) = 0$.	
d. $\sum x_i^2 = \left(\sum x_i\right)^2$	

(2) [Bernoulli random variable: 4 pts] Suppose X is a Bernoulli random variable, with Prob{X=1} = 0.2 and Prob{X=0} = 0.8.

a. Compute the mean of X, that is, E(X).

b. Compute the variance of X, that is, Var(X).

(3) [Conditional mean: 6 pts] Consider the following table, which shows the joint probabilities of discrete random variables X and Y .

		X = 1	X = 2	X = 3
Y	1	Prob = 0.1	Prob = 0.4	Prob = 0.05
	2	Prob = 0.3	Prob = 0.1	Prob = 0.05

a. Compute the conditional mean E(Y|X=1).

b. Compute the conditional mean E(Y|X=2).

c .Compute the unconditional mean E(Y).

(4) [Mean and variance of linear function: 4 pts] Suppose X is a random variable with mean E(X) = 2 and variance Var(X) = 3 . Suppose Y = 2 + 4 X .

a. Compute the mean of Y, that is, E(Y).

b. Compute the variance of Y, that is, Var(Y).

(5) [Mean and variance of sample mean: 4 pts] Suppose $X_1, \ldots, X_{10}$ are random variables with identical mean 4 and variance 5 . Assume $Cov(X_i,X_j) = 0$ for all $i \neq j$. Let $\bar{X} = \frac{1}{10}\sum_{i=1}^{10} X_i$.

a. Compute the mean of $\bar{X}$, that is, E($\bar{X}$).

b. Compute the variance of $\bar{X}$, that is, Var($\bar{X}$).

(6) [Important distributions: 8 pts] Consider the following probability distributions. Some of them have density functions which are symmetric about zero. Write "SYMMETRIC" for these distributions. Write "ASYMMETRIC" for the others.

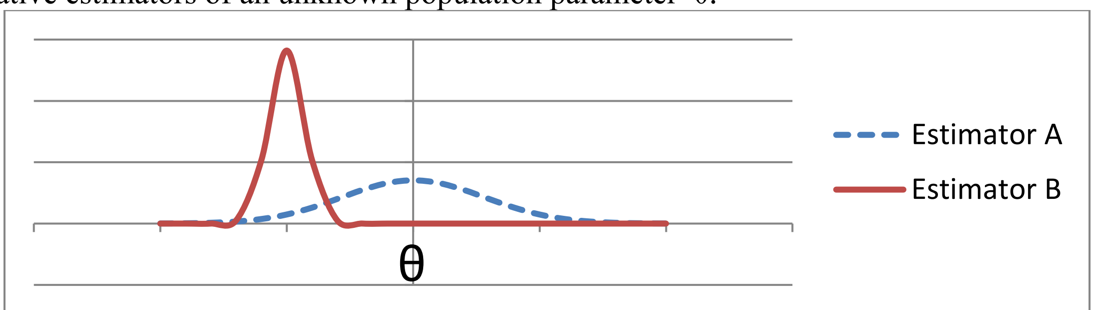
a. Chi-square (χ^2) distribution with 6 degrees of freedom.

b. Standard normal distribution.

c . *t* distribution with 9 degrees of freedom.

d. *F* distribution with 2 degrees of freedom in the numerator and 12 degrees of freedom in the denominator.

(7) [Properties of estimators: 4 pts] The graph below shows the density functions for two alternative estimators of an unknown population parameter θ.

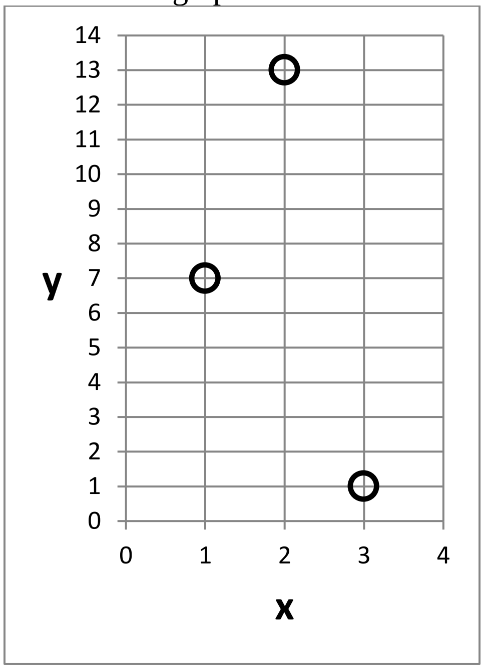


a. Which estimator has greater **bias**? Answer "A," "B," or "EQUAL."

b. Which estimator has greater **variance**? Answer "A," "B," or "EQUAL."

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

(1) [Least-squares calculation: 14 pts] Suppose we have three observations on x_i and y_i as shown in the graph below.



a. [3 pts] Compute $\hat{\beta}_2$, the least-squares estimate of the slope of the line $y = \beta_1 + \beta_2 x$.

b. [3 pts] Compute $\hat{\beta}_1$, the least-squares estimate of the y-intercept of the same line.

c. [3 pts] Compute the three fitted values $\hat{y}_i$ of this least-squares estimated regression line.

d. [3 pts] Compute the three residuals $\hat{\varepsilon}_i$ of this estimated least-squares regression line.

e. [2 pts] Sketch the least-squares estimated line in the graph above.

(2) [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 $MSE(\hat{\mu})$.

(3) [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? Explain your answer.

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

c. [3 pts] Compute the standard error of your estimate of μ.

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

e. [6 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 a table, and your conclusion (whether you can reject null hypothesis).

Value of test statistic = ________________. Critical point(s) = __________________.

Can you reject null hypothesis? ______________________.

[end of exam]