

MIDTERM EXAMINATION #1 VERSION C
“Introduction and Statistics Review”
February 14, 2008

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. $\sum x_i^2 = (\sum x_i)^2$.
- b. $\sum (x_i - \bar{x}) = 0$.
- c. $\sum x_i = n\bar{x}$.
- d. $\sum (\alpha x_i) = \alpha \sum x_i$.
- e. $\sum (x_i - \bar{x})^2 = (\sum x_i^2) - n\bar{x}^2$.

(2) $\frac{\partial}{\partial \alpha} \sum_{i=1}^n (\alpha x_i + 7)^2 =$

- a. $\sum_{i=1}^n (\alpha x_i + 7)^2$.
- b. $14 \sum_{i=1}^n (\alpha x_i + 7)$.
- c. $2 \sum_{i=1}^n (\alpha x_i + 7)x_i$.
- d. $\sum_{i=1}^n x_i$.
- e. $\sum_{i=1}^n (x_i + 7)^2$.

(3) 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. $f(\beta_1, \beta_2) = \sum (y_i - \beta_1 - \beta_2 x_i)^2$.
- b. $f(\beta_1, \beta_2) = \sum (\beta_1 + \beta_2 x_i)^2$.
- c. $f(\beta_1, \beta_2) = \sum (y_i - \beta_1 - \beta_2 x_i)$.
- d. $f(\beta_1, \beta_2) = \sum |y_i - \beta_1 - \beta_2 x_i|$.
- e. $f(\beta_1, \beta_2) = \sum (y_i^2 - (\beta_1 + \beta_2 x_i)^2)$.

The next two questions assume the following. Suppose X is a Bernoulli random variable, with $\text{Prob}\{X=1\} = 0.9$ and $\text{Prob}\{X=0\} = 0.1$.

(4) The mean or expected value of X is

- a. zero.
- b. 0.09.
- c. 0.1.
- d. 0.9.
- e. one.

- (5) The variance of \bar{X} is
- zero.
 - 0.09 .
 - 0.1 .
 - 0.9 .
 - one.

- (6) Which of the following distributions has a symmetric bell-shaped density function?
- normal distribution.
 - t distribution.
 - all of the above have symmetric bell-shaped density functions.
 - none of the above have symmetric 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) $\text{Var}(\bar{X}) =$
- σ^2 .
 - $\sigma^2 / (n-1)$.
 - σ^2 / n .
 - σ^2 / n^2 .
 - zero.
 - one.

- (8) $E(\bar{X}) =$
- μ .
 - $\mu / (n-1)$.
 - μ / n .
 - μ / n^2 .
 - zero .
 - one.

- (9) An estimator $\hat{\theta}$ of an unknown population parameter θ is said to be asymptotically unbiased if

- $E(\hat{\theta}) = 0$.
- $E(\hat{\theta}) = \theta$.
- $\lim_{n \rightarrow \infty} E(\hat{\theta}) = 0$.
- $\lim_{n \rightarrow \infty} E(\hat{\theta}) = \theta$.
- $\lim_{n \rightarrow \infty} \text{Prob}(|\hat{\theta} - \theta| > \delta) = 0$, for all $\delta > 0$.

- (10) An estimator $\hat{\theta}$ of an unknown population parameter θ is said to be unbiased if

- $E(\hat{\theta}) = 0$.
- $E(\hat{\theta}) = \theta$.
- $\lim_{n \rightarrow \infty} \text{MSE}(\hat{\theta}) = 0$.
- $\lim_{n \rightarrow \infty} E(\hat{\theta}) = \theta$.
- $\lim_{n \rightarrow \infty} \text{Prob}(|\hat{\theta} - \theta| > \delta) = 0$, for all $\delta > 0$.

- (11) The principle for finding an estimator that proceeds by setting the sample mean equal to the population mean, the sample variance equal to the population variance, etc., is called

- the scientific method.
- Newton's method.
- the method of maximum likelihood.
- the method of moments.

- (12) A 99 percent confidence interval is necessarily

- narrower than a 95 percent confidence interval.
- wider than a 95 percent confidence interval.
- identical to a 95 percent confidence interval.
- answer cannot be determined from the information given.

(13) The probability that a test will correctly reject the null hypothesis when it is false is called the

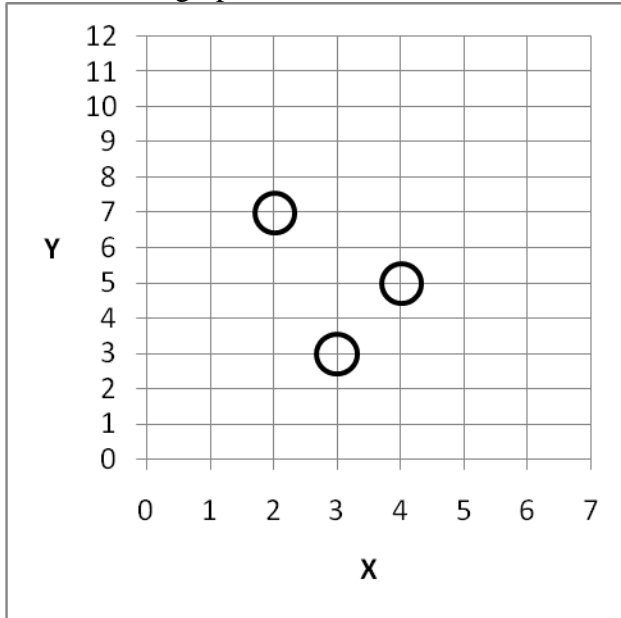
- a. test statistic.
- b. standard error.
- c. critical point of the test.
- d. power of the test.
- e. size or significance of the test.

(14) Suppose the p-value for a test statistic is 0.067. If the size of the test is 5 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. PROBLEMS: Please write your answers in the boxes on this question sheet. Show your work and circle your final answers.

(1) [Least-squares calculation: 15 pts] Suppose we have three observations on x_i and y_i as shown in the graph below.



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.

e. Sketch the least-squares estimated line in the graph above.

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

$$E(X_1) = 4$$

$$\text{Var}(X_1) = 3$$

$$\text{Cov}(X_1, X_2) = 0.25$$

$$E(X_2) = 7$$

$$\text{Var}(X_2) = 12$$

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

a. Compute $E(Y)$.

b. Compute $\text{Var}(Y)$.

c. Compute $\text{SD}(Y)$.

d. Compute $\text{Corr}(X_1, X_2)$.

(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.

$$\hat{\mu} = 5 + \frac{1}{20} \sum_{i=1}^{10} x_i$$

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 $\text{Var}(X_i) = 120$. Circle your final answers.

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

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

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

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

(4) [Inference for arbitrary distribution, large sample: 18 pts] Suppose we wish to analyze the distribution of the number of cars per household in a population. Let μ denote the unknown true population mean number of cars per household. Observations X_i have been collected on 500 households, with the following summary values. Here, \bar{X} is the sample mean.

$$\sum_{i=1}^{500} X_i = 950 \qquad \sum_{i=1}^{500} (X_i - \bar{X})^2 = 3025$$

- 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 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) = _____.

Reject null hypothesis? _____.

[end of exam]