IE 230 Seat # ________ Name ___ < KEY > ___

Closed book and notes. 120 minutes.

Cover page, five pages of exam.
No calculator. No need to simplify answers.

Recall: A true/false statement is true only if it is always true.

Score ___________________________

Closed book and notes. 120 minutes.

Consider an experiment that chooses a random sample $X_1, X_2, \ldots, X_n$ from a population with cdf F, mean $\mathrm{E}(X) = \mu$, variance $\mathrm{V}(X) = \sigma^2$ and median $x_{0.5} = F^{-1}(0.5)$. Let $\bar{X}$ denote the sample mean, $S^2 = [\sum_{i=1}^{n} (X_i - \bar{X})^2] / (n - 1)$ denote the sample variance, and $X_{(i)}$ the ith order statistic.

1. True or false. (2 points each)

(a) T ← F The standard error of a point estimator is its standard deviation.

(b) T F ← A statistic is an event computed from a population.

(c) T ← F A point estimator is a statistic that is used to estimate a population parameter.

(d) T ← F A sample is selected from a population.

(e) T ← F $X_{(1)} \le X_{(n)}$.

(f) T ← F $\mathrm{V}(X) = \mathrm{E}[(X - \mu_X)^2]$.

(g) T F ← $\bar{X} = \mathrm{E}(X)$.

(h) T F ← The sample size is $X_1 + X_2 + \ldots + X_n$.

(i) T F ← $X_{(1)} \le \mathrm{E}(X) \le X_{(n)}$.

(j) T ← F If $n = 1$, then $\bar{X} = X_1$.

(k) T ← F One advantage of a stem-and-leaf plot is that the observation values are not lost.

(l) T F ← One advantage of a box plot is that the observation values are not lost.

(m) T F ← One advantage of a histogram is that the observation values are not lost.

(n) T F ← The sample median is always less than or equal to the sample average.

(o) T ← F Computing the order statistics is necessary to create the empirical cdf.

2. (From Montgomery and Runger, fourth edition, Problem 7–30) Let X denote a random variable with probability density function $f_X(x) = (\theta+1)x^\theta$ for $0 \le x \le 1$ and zero elsewhere.

(a) (4 points) What values of the parameter θ are possible?

A pdf needs to satisfy two conditions:
A. $0 \le f_X(x)$ for every real number x
and
B. $\int_{-\infty}^{\infty} f_X(x)\,dx = 1$.
Condition A implies that $0 \le (\theta+1)x^\theta$, so $-1 \le \theta$ is required.
Condition B requires $\int_0^1 (\theta+1)x^\theta\,dx = 1$,
which is true for every θ value greater than -1.
Therefore, the set of possible values is all θ greater than -1. ←

(b) (4 points) Sketch the pdf for $\theta = 2$. Label and scale both axes.

Sketch two axes. Label the horizontal axis with a dummy variable, probably x.
Label the vertical axis with $f_X(x)$.
Scale the horizontal axis with at least two numbers, probably zero and one.
Scale the vertical axis with at least two numbers, probaby zero and $(2+1)(1)^2 = 9$.
Sketch the function $f_X(x) = 3x^2$ for $0 \le x \le 1$ and zero elsewhere.

(c) (4 points) On your sketch of Part (b), indicate the likelihood of $X = 0.75$.

The likelihood of $X = 0.75$ is $f_X(0.75) = 3(0.75)^2$
Mark the point $(0.75, 3(0.75)^2$ on the sketch.

(d) (4 points) On your sketch, indicate the value of $\mathrm{E}(X)$.

$\mathrm{E}(X) = \int_0^1 x\,[(\theta+1)\,x^\theta]\,dx = (\theta+1)\,/\,(\theta+2)$.
Therefore, the mean is $\mathrm{E}(X) = 3/4$. Indicate this point on the horizontal axis.

---

3. (2 points each) Choose one answer.

(a) μ_X	constant ←	random variable	event	meaningless
(b) $\bar{X}$	constant	random variable ←	event	meaningless
(c) $\mathrm{E}(X)$	constant ←	random variable	event	meaningless
(d) $X_{(1)}$	constant	random variable ←	event	meaningless

4. Result: If $X_1, X_2, \ldots, X_n$ is a random sample from a population with cdf F_X and mean μ_X and variance σ_X^2, then the sample mean $\bar{X} = \sum_{i=1}^{n} X_i \,/\, n$, is an unbiased estimator of μ_X.

   For each line of the proof, choose from these reasons. (3 points each)

   (i) Axiom 3: If events A and B are mutually exclusive, then $\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B)$.
   (ii) Independence
   (iii) Algebra (no probability required)
   (iv Definition of expected value
   (v) Definition of sample mean
   (vi) $\mathrm{V}(X) = \mathrm{E}(X^2) - \mu_X^2$
   (vii) For a random sample, $\mathrm{V}(\bar{X}) = \sigma_X^2 \,/\, n$
   (viii) Same event
   (ix) Identically distributed data
   (x) Expected value of a linear combination
   (xi) Definition of mean squared error
   (xii) Definition of sample variance

$$\mathrm{E}(\bar{X}) = \mathrm{E}\left[\frac{\sum_{i=1}^{n} X_i}{n}\right]$$ (a) ___ < (v) > ___

$$= \mathrm{E}\left[\sum_{i=1}^{n} (1/n) X_i\right]$$ (b) ___ < (iii) > ___

$$= \sum_{i=1}^{n} (1/n)\mathrm{E}(X_i)$$ (c) ___ < (x) > ___

$$= \sum_{i=1}^{n} (1/n)\mu_X$$ (d) ___ < (ix) > ___

$$= \mu_X$$ (e) ___ < (iii) > ___

5. (2 points each) We discussed the analogy of point estimators being like target shooting. Error in target shooting is caused by either aiming the wrong direction or by not holding the gun steady. Error in point estimation is caused by bias and standard error.

   (a) T F ← Mean squared error is the sum of bias and standard error.

   (b) T ← F Bias is analogous to aiming in the wrong direction.

6. A computer software package was used to calculate some numerical summaries of a random sample of data from a population with mean μ and variance σ^2. The results are in the following table.

Variable	n	mean	ste mean	stdev	variance
X	10	3.841	?	?	22.891

(a) (5 points) (first missing value) What is the value of the estimated standard error of the sample mean?

$$\hat{\sigma}_{\overline{X}} = \sqrt{s^2 / n} = \sqrt{22.891 / 10} \leftarrow$$

(b) (5 points) (second missing value) What is the value of the sample standard deviation?

$$\sqrt{s^2} = \sqrt{22.891} \leftarrow$$

(c) (5 points) What is the value of the point estimate of σ^2?

$$s^2 = 22.891 \leftarrow$$

7. Consider the exponential distribution with rate λ. The mean and standard deviation are then $1/\lambda$, the pdf is $f(x) = \lambda e^{-\lambda x}$ for $0 \le x$ and zero elsewhere, and cdf $F(x) = 1 - e^{-\lambda x}$ for $0 \le x$ and zero elsewhere. A sample of size $n = 3$ is taken, with observations $x_1 = 3.5$, $x_2 = 7.2$, $x_3 = 11.0$.

(a) (5 points) State the method-of-moments estimator of λ.

The method-of-moments point esimtator is obtain by setting the population mean equal to the sample average.

Therefore,

$$1/\hat{\lambda} = \bar{X}$$

which implies

$$\hat{\lambda} = 1/\bar{X} \leftarrow$$

(b) (2 points) T F $\leftarrow$ Because of the memoryless property of the exponential distribution, there is no maximum-likelihood estimator of λ.

(c) (5 points) Using the given data, compute the observed value of the sample variance, $S^2 = [\sum_{i=1}^{n} X_i^2 - n\bar{X}^2]/(n-1)$. (Be clear, using parentheses and numbers. No need to simplify.)

$$S^2 = \frac{[(3.5)^2 + (7.2)^2 + (11.0)^2] - 3[(3.5 + 7.2 + 11.0)/3]^2}{2} \leftarrow$$
