Name: ______________________________

AP Statistics — Practice Exam

Instructions: Answer all questions in the spaces provided. Show your working — method marks are awarded even when the final answer is wrong. The answer key with worked solutions is available at examtex.com/free-exams.

Time Limit: 55 minutes

Multiple Choice (12 questions)

[2 points] 1. A distribution of household incomes is strongly skewed to the right. Which of the following is most likely true?

A. The mean is greater than the median
B. The mean is less than the median
C. The mean equals the median
D. The median is pulled toward the high incomes
E. The standard deviation is zero

[2 points] 2. A data set of 25 house prices has one extreme outlier added to it. Which pair of statistics is MOST resistant to the outlier?

A. Mean and standard deviation
B. Mean and range
C. Median and IQR
D. Median and standard deviation
E. Range and IQR

[2 points] 3. Scores on a test are approximately normal with mean 70 and standard deviation 10. A student scores 85. The z-score is:

A. 1.5, meaning the score is 1.5 standard deviations above the mean
B. 1.5, meaning the score is 1.5 points above the mean
C. 15, the distance from the mean
D. 0.85, the percentile of the score
E. −1.5, since most students scored lower

[2 points] 4. A least-squares regression of exam score on hours studied has correlation $r = 0.8$. Which statement is correct?

A. 80 percent of the variation in scores is explained by the regression
B. 64 percent of the variation in scores is explained by the linear relationship with hours studied
C. Each extra hour of study causes a 0.8-point score increase
D. The regression predicts scores correctly 80 percent of the time
E. The residuals have correlation 0.8 with the predictor

[2 points] 5. A principal wants a sample of 100 students that reflects the proportions of freshmen, sophomores, juniors, and seniors in the school. She randomly selects students separately from each grade, in proportion to grade size. This design is:

A. A simple random sample
B. A stratified random sample
C. A cluster sample
D. A systematic sample
E. A convenience sample

[2 points] 6. In a comparative experiment, the primary purpose of RANDOM ASSIGNMENT to treatments is to:

A. guarantee the sample represents the population
B. eliminate all variability between subjects
C. balance the effects of lurking variables across treatment groups
D. ensure the placebo effect does not occur
E. increase the sample size

[2 points] 7. At a school of 200 students, 120 are juniors and 80 are seniors. Forty-eight juniors and 40 seniors have part-time jobs. If a randomly selected student is a senior, the probability that they have a part-time job is:

A. 0.20
B. 0.44
C. 0.50
D. 0.45
E. 0.40

[2 points] 8. Events A and B have $P(A) = 0.4$, $P(B) = 0.5$, and $P(A \cap B) = 0.2$. Which is true?

A. A and B are mutually exclusive
B. A and B are independent, since $P(A)P(B) = P(A \cap B)$
C. A and B are dependent, since $P(A \cap B) \neq 0$
D. $P(A \cup B) = 0.9$
E. A and B cannot both occur

[2 points] 9. Which of the following scenarios describes a binomial random variable?

A. Counting made baskets in 20 independent free throws, each with success probability 0.7
B. Counting draws from a deck WITHOUT replacement until the first ace appears
C. Recording the time until a light bulb fails
D. Counting cars passing an intersection in an hour
E. Measuring the average height of 30 random students

[2 points] 10. A raffle sells 500 tickets at 2 dollars each, with a single prize worth 400 dollars. The expected net gain from buying one ticket is:

A. a loss of 1.20 dollars
B. a gain of 0.80 dollars
C. a loss of 2.00 dollars
D. a loss of 0.80 dollars
E. exactly zero

[2 points] 11. A population has standard deviation σ. If the sample size is increased from $n = 25$ to $n = 100$, the standard deviation of the sampling distribution of the sample mean:

A. is divided by 4
B. is divided by 2
C. is unchanged
D. is multiplied by 2
E. becomes zero

[2 points] 12. A 95 percent confidence interval for a population mean is computed as $(12.1,\ 15.3)$. Which interpretation is correct?

A. There is a 95 percent probability that the population mean is between 12.1 and 15.3
B. 95 percent of the data values lie between 12.1 and 15.3
C. If we repeated the sampling many times, about 95 percent of the intervals built this way would capture the population mean
D. 95 percent of sample means will fall between 12.1 and 15.3
E. The sample mean has a 95 percent chance of being in the interval

Calculations (2 questions)

[8 points] 13. Significance test. A city official claims that 60 percent of voters support a new transit measure. A polling firm takes a random sample of 600 voters from the city's 90,000 registered voters and finds 390 who support the measure. Does the sample provide convincing evidence, at the 5 percent significance level, that MORE than 60 percent of all the city's voters support the measure?

[2 points] (a) Define the parameter of interest and state the hypotheses.

[2 points] (b) Verify that the conditions for a one-proportion z-test are met.

[2 points] (c) Compute the test statistic and the p-value.

[2 points] (d) State a conclusion in the context of the problem.

[8 points] 14. Expected value. An insurance company sells a one-year policy for 150 dollars. Based on historical data, in a given year a policyholder files no claim with probability 0.95, a minor claim paying out 1,000 dollars with probability 0.04, and a major claim paying out 10,000 dollars with probability 0.01.

[2 points] (a) Let X be the payout on a randomly selected policy. Construct the probability distribution of X and verify it is valid.

[2 points] (b) Compute the expected payout $E(X)$ and interpret it in context.

[2 points] (c) Compute the standard deviation of X.

[2 points] (d) What is the company's expected profit per policy, and why can the company rely on this number even though individual policies vary so much?

Answer Key

1. **A**

2. **C**

3. **A**

4. **B**

5. **B**

6. **C**

7. **C**

8. **B**

9. **A**

10. **A**

11. **B**

12. **C**

13. (a) Let p = the true proportion of all the city's registered voters who support the transit measure. $H_0 : p = 0.60$ versus $H_a : p > 0.60$. (Hypotheses are about the parameter p, never about $\hat{p}$.)

(b) Random: stated random sample. 10 percent condition: $600 \leq 0.10 \times 90{,}000 = 9{,}000$, so independence is reasonable. Large counts (using p_0): $np_0 = 600(0.6) = 360 \geq 10$ and $n(1 - p_0) = 600(0.4) = 240 \geq 10$, so the sampling distribution of $\hat{p}$ is approximately normal.

(c) $\hat{p} = \frac{390}{600} = 0.65$. $z = \frac{0.65-0.60}{\sqrt{\frac{0.60\times0.40}{600}}} = \frac{0.05}{\sqrt{0.0004}} = \frac{0.05}{0.02} = 2.5$. p-value $= P(Z \geq 2.5) \approx 0.0062$.

(d) Because the p-value $0.0062 < \alpha = 0.05$, we reject H_0. There is convincing statistical evidence that more than 60 percent of the city's registered voters support the transit measure. (Both the comparison to α AND the context sentence are required for full credit.)

14. (a) X takes values 0, 1,000, and 10,000 dollars with probabilities 0.95, 0.04, and 0.01. Valid because each probability is between 0 and 1 and $0.95 + 0.04 + 0.01 = 1$.

(b) $E(X) = 0(0.95) + 1000(0.04) + 10000(0.01) = 40 + 100 = 140$ dollars. Interpretation: over MANY policies, the company pays out an average of about 140 dollars per policy per year — not a prediction for any single policyholder.

(c) $E(X^2) = 0^2(0.95) + 1000^2(0.04) + 10000^2(0.01) = 40{,}000 + 1{,}000{,}000 = 1{,}040{,}000$. $\text{Var}(X) = 1{,}040{,}000 - 140^2 = 1{,}020{,}400$, so $\sigma_X = \sqrt{1{,}020{,}400} \approx 1{,}010$ dollars.

(d) Expected profit $= 150 - E(X) = 150 - 140 = 10$ dollars per policy. Although one policy's outcome is highly variable ($\sigma \approx 1{,}010$ dollars), the law of large numbers means the AVERAGE payout across thousands of independent policies will be very close to 140 dollars — so the average profit per policy stabilizes near 10 dollars.

Worked Solutions

Question 13. *(a)* Let p = the true proportion of all the city's registered voters who support the transit measure. $H_0 : p = 0.60$ versus $H_a : p > 0.60$. (Hypotheses are about the parameter p, never about $\hat{p}$.)

(b) Random: stated random sample. 10 percent condition: $600 \leq 0.10 \times 90{,}000 = 9{,}000$, so independence is reasonable. Large counts (using p_0): $np_0 = 600(0.6) = 360 \geq 10$ and $n(1 - p_0) = 600(0.4) = 240 \geq 10$, so the sampling distribution of $\hat{p}$ is approximately normal.

(c) $\hat{p} = \frac{390}{600} = 0.65$. $z = \frac{0.65-0.60}{\sqrt{\frac{0.60\times0.40}{600}}} = \frac{0.05}{\sqrt{0.0004}} = \frac{0.05}{0.02} = 2.5$. p-value $= P(Z \geq 2.5) \approx 0.0062$.

(d) Because the p-value $0.0062 < \alpha = 0.05$, we reject H_0. There is convincing statistical evidence that more than 60 percent of the city's registered voters support the transit measure. (Both the comparison to α AND the context sentence are required for full credit.)

Question 14. *(a)* X takes values 0, 1,000, and 10,000 dollars with probabilities 0.95, 0.04, and 0.01. Valid because each probability is between 0 and 1 and $0.95 + 0.04 + 0.01 = 1$.

(b) $E(X) = 0(0.95) + 1000(0.04) + 10000(0.01) = 40 + 100 = 140$ dollars. Interpretation: over MANY policies, the company pays out an average of about 140 dollars per policy per year — not a prediction for any single policyholder.

(c) $E(X^2) = 0^2(0.95) + 1000^2(0.04) + 10000^2(0.01) = 40{,}000 + 1{,}000{,}000 = 1{,}040{,}000$. $\text{Var}(X) = 1{,}040{,}000 - 140^2 = 1{,}020{,}400$, so $\sigma_X = \sqrt{1{,}020{,}400} \approx 1{,}010$ dollars.

(d) Expected profit $= 150 - E(X) = 150 - 140 = 10$ dollars per policy. Although one policy's outcome is highly variable ($\sigma \approx 1{,}010$ dollars), the law of large numbers means the AVERAGE payout across thousands of independent policies will be very close to 140 dollars — so the average profit per policy stabilizes near 10 dollars.