Placement Exam for MA117

September 5, 2025

Central Limit Theorem

10 points for scoring at least 2 out of 3.

[A] Suppose that the number of times that individual Americans have been to France is highly skewed, with a mean of 1.4 and a standard deviation of 2.5. Which of the following statements is true?

(i) In a random sample of 100 Americans, the number of times that these individuals have been to France will follow an approximately normal distribution with mean 1.4 and standard deviation 2.5.

(ii) In a random sample of 100 Americans, the number of times that these individuals have been to France will follow an approximately normal distribution with mean 1.4 and standard deviation 0.25.

(iii) In a random sample of 100 Americans, the average number of times that these individuals have been to France will come from an approximately normal distribution with mean 1.4 and standard deviation 0.25.

(iv) In a random sample of 100 Americans, the number of times that these individuals have been to France will follow an approximately normal distribution with mean 1.4 and standard deviation 0.025.

(v) In a random sample of 100 Americans, the average number of times that these individuals have been to France will come from an approximately normal distribution with mean 1.4 and standard deviation 2.5.

[B] Which of the following is a consequence of the Central Limit Theorem?

(i) In any sufficiently large sample, the data points will approximately follow a normal distribution.

(ii) In a large sample of independent identically distributed random variables, the values or the random variables will approximately follow a normal distribution.

(iii) The average from a large sample of independent measurements of the same variable comes from an approximately normal distribution of possible values.

(iv) The sample mean for any set of independent identically distributed random variables comes from an approximately normal distribution of possible values.

[C] Which of the following statements best captures why the Central Limit Theorem is important in statistics?

(i) It explains why we can use the normal distribution and related distributions (T distribution, chi-squared distribution) to analyze many data sets even when we don't know the shape of the population distribution.

(ii) It allows us to avoid Type 1 and Type 2 errors in hypothesis testing.

(iii) It allows us to conclude that a population distribution is normal, based on a sample.

(iv) It allows us to extrapolate beyond the range of observed data when doing linear regression.

Confidence Intervals

20 points for scoring at least 3 out of 4.

[A] Suppose that researchers measure 200 adult barn owls and conclude that the 95% confidence interval for the average weight of an adult barn owl is 310g +/- 15g. Which of the following is the correct interpretation?

(i) 95% of adult barn owls weigh between 295g and 325g.

(ii) If another sample of 200 adult barn owls were measured, there is a 95% chance that their average weight will be between 295g and 325g.

(iii) The researchers are 95% confident that the average weight of all adult barn owls is between 295g and 325g.

(iv) The researchers are 95% confident that the average weight of the owls in their sample is between 295g and 325g.

[B] Suppose that a sample size of 100 produces a 95% confidence interval for the population mean: 13.5 +/- 1.6 . Which of the following statements is true?

(i) The 90% confidence interval from the same data will have a larger margin of error.

(ii) A sample of size of 200 would produce a 95% confidence interval: 13.5 +/- 0.8 .

(iii) A sample size of 400 would produce a 95% confidence interval: 13.5 +/- 0.8 .

(iv) A sample of size of 200 would produce a 95% confidence interval: 27.0 +/- 3.2 .

[C] A survey of 80 randomly chosen high school juniors finds that 32 of them are planning to attend college. Which of the following is the approximate 95% confidence interval for the proportion of all high school juniors who plan to attend college?

(i) 0.32 +/- 0.11

(ii) 0.40 +/- 0.11

(iii) 0.40 +/- 0.96

(iv) 0.32 +/- 0.05

[D] A random sample of 20 adult mice finds that their average weight is 18.5 g, with a standard deviation of 3.5 g. Which of the following is the approximate 95% confidence interval for the average weight of all adult mice?

(i) (15.0, 22.0)

(ii) (17.0, 20.0)

(iii) (16.9, 20.1)

(iv) (11.2, 25.8)

Hypothesis Testing

20 points for scoring at least 4 out of 5.

[A] Which of the following is an accurate interpretation of what the p-value from a hypothesis test means?

(i) The probability that the null hypothesis is true.

(ii) The probability of obtaining data as "extreme" as what was observed, if we assume the null hypothesis is true.

(iii) The probability of obtaining data as "extreme" as what was observed, if we assume the null hypothesis is false.

(iv) The probability that the null hypothesis is false, give the data that was observed.

(v) The probability that our conclusion is wrong.

[B] Suppose that an appropriately conducted hypothesis test based on a sample of size 20 yields a p-value of 0.12. What is the most reasonable conclusion?

(i) We can reject the null hypothesis.

(ii) We cannot reject the null hypothesis, but there is some evidence that this issue may be worth studying with more data.

(iii) We can confidently accept the alternative hypothesis.

(iv) We cannot reject the alternative hypothesis without more data.

[C] Researchers want to know whether treatment "A" or treatment "B" leads to better outcomes for patients. Of 100 patients who have treatment "A", 68 recover. Of 100 patients who have treatment "B", 56 recover. What are the most appropriate null and alternative hypotheses for a hypothesis test?

(i) $H_0 : p_A > p_B$ and $H_A : p_A < p_B$

(ii) $H_0 : p_A - p_B = 0.12$ and $H_A : p_A - p_B \neq 0.12$

(iii) $H_0 : p_A = p_B$ and $H_A : p_A \neq p_B$

(iv) $H_0 : p_A = p_B$ and $H_A : p_A - p_B = 0.12$

(v) $H_0 : p_A > 0.68$ and $H_A : p_B > 0.56$

[D] Suppose that we conduct a hypothesis test for the population proportion, with the null hypothesis $H_0 : p = 0.3$ against the alternative $H_A : p \neq 0.3$. We have a sample size of 100, of which 38 are "successes". What are the approximate test statistic and p value for this test?

(i) Z = 1.75, p = 0.08

(ii) Z = 3.50, p ¡ 0.0002

(iii) Z = 1.65, p = 0.099

(iv) T = 1.75, 0.05 ¡ p ¡ 0.10

(v) T = 3.5, p ¡ 0.001

(vi) T = 1.65, 0.10 ¡ p ¡ 0.20

[E] Suppose that we conduct a hypothesis test for the population mean, with the null hypothesis $H_0 : \mu = 20$ against the alternative $H_A : \mu \neq 20$. We have a sample size of 30, with a sample mean of 17 and sample standard deviation of 6. What are the approximate test statistic and p value for this test?

(i) Z = -2.74, p = 0.0062

(ii) Z = -5.48, p ¡ 0.0002

(iii) Z = -2.74, p = 0.0031

(iv) T = -0.50, p ¿ 0.25

(v) T = -2.74, 0.01¡ p ¡ 0.02

(vi) T = -5.48, p ¡ 0.0005

Normal and Binomial

10 points for scoring at least 2 out of 2.

[A] Suppose that the length of squirrel tails is approximately normally distributed, with a mean of 24 cm and a standard deviation of 4 cm. What is the approximate probability that a squirrel has a tail that is between 18 cm and 30 cm?

(i) 0.87

(ii) 0.93

(iii) 0.13

(iv) 0.44

[B] Let X be the number of times needed to roll a fair die until it lands on 6 for the first time. What is the name of the distribution of X?

(i) Normal

(ii) Binomial

(iii) Neither of these

Random Variables

15 points for scoring at least 2 out of 3.

[A] Suppose that half of the households in a small town have no dogs, a quarter of the households have one dog, and the rest of the households have two dogs. If we pick a random household and let X be the number of dogs, what is the expected value of X?

(i) 0.25

(ii) 0.75

(iii) 1

(iv) 0.33

[B] Suppose that X and Y are random variables. X has mean 3 and standard deviation 1. Y has mean 4 and standard deviation 3. What are the mean and standard deviation of X+Y?

(i) Mean is 7/2 and standard deviation is 4

(ii) Mean is 7 and standard deviation is $\sqrt{10}$

(iii) Mean is 7/2 and standard deviation is unknown (not enough information)

(iv) Mean is 7 and standard deviation is unknown (not enough information)

[C] Suppose that X is a random variable with mean 10 and standard deviation 2. What are the mean and standard deviation of the quantity 3X+5?

(i) Mean is 35 and standard deviation is 11

(ii) Mean is 30 and standard deviation is 10

(iii) Mean is 35 and standard deviation is 6

(iv) Mean is 30 and standard deviation is 11

Probability

15 points for scoring at least 2 out of 3.

[A] Suppose that events E and F are independent of each other. Which of the following statement(s) is/are true?

(i) P(E—F) = P(E)

(ii) P(E and F) = P(E)P(F)

(iii) P(E or F) = P(E) + P(F)

(iv) If F occurs, E cannot occur.

(v) The first two statements.

(vi) The second and third statement.

(vii) The third and four statement.

[B] Suppose that we pick a random person from the population. Let E be the event that they are left-handed, and let F be the event that they like spicy food. What is the most reasonable assumption about these two events?

(i) P(E and F) = P(E)P(F)

(ii) P(E or F) = P(E) + P(F)

(iii) Both of these.

(iv) Neither of these.

[C] Suppose that 12% of people are left-handed. In a random sample of 5 people, what is the approximate probability that at least one of them is left-handed?

(i) 0.00025

(ii) 0.025

(iii) 0.528

(iv) 0.6

(v) 0.472

Regression

10 points for scoring at least 2 out of 2.

[A] Suppose that researchers use linear regression to study the relationship between the weight (in kg) and blood pressure (in mmHg) of cats. There were 100 cats in the study, and they weighed between 3 and 7 kg. The researchers report "We found a significant increase in blood pressure with weight: y = 110 + 8 x; p = 0.03, $R^2 = 0.85$." Which of the following can you most reasonably conclude from this statement?

(i) A 20 kg cat should have blood pressure of approximately 270 mmHg.

(ii) If Dexter weighs 2 kg more than Fluffy, then Dexter's blood pressure should be about 16 mmHg higher than Fluffy's.

(iii) The average blood pressure of a 3 kg cat is about 110 mmHg.

(iv) The relationship between cat's weights and blood pressure is not linear.

[B] Researchers report the following result: "We used linear regression to test for the presence of a relationship between x and y. The hypothesis test yielded p = 0.001, with $R^2 = 0.1$." Which of the following is the most reasonable interpretation of these results?

(i) There is a strong and statistically significant relationship between x and y.

(ii) There is no apparent relationship between x and y.

(iii) The relationship between x and y is statistically significant but weak. This tends to happen when the sample size is small.

(iv) The relationship between x and y is statistically significant but weak. This tends to happen when the sample size is very large.

(v) The relationship between x and y may be strong but this result is not statistically significant. More data is needed.