

UNIVERSITY OF TEXAS AT AUSTIN, DEPARTMENT OF MATHEMATICS
M358K - Applied Statistics

MOCK IN-TERM III

Problem 1.1. (5 points)

Lord Clarence Emsworth keeps meticulous records of the feeding patterns of his prized pet hog *the Empress of Blandings*. He observes that the average weight of slops consumed per diem during the month of November equals 19 lbs with a sample standard deviation of 3lbs. As the incessantly perused tome “The Care of the Pig” states that the daily intake of slops should amount to 20 lbs, Lord Emsworth gets worried.

Luckily, his secretary, the efficient Baxter, knows a bit of statistics and can test whether the Empress is underfed. Which p -value will the efficient Baxter report?

- a. Between 0 and 0.005.
- b. Between 0.005 and 0.01.
- c. Between 0.01 and 0.025.
- d. Between 0.025 and 0.05.
- e. None of the above.

Problem 1.2. (5 points)

Bertie and Tuppy are playing a game. Bertie simulates 5 draws from a normal distribution without telling the parameter values to Tuppy. Then, Tuppy calculates the 95% confidence interval for the mean parameter.

Bertie simulated the values which resulted in a sample average of 5.88 and the sample standard deviation of 1.96. Which confidence interval should Tuppy get?

- a. 5.88 ± 2.4333
- b. 5.88 ± 2.7209
- c. 5.88 ± 2.7764
- d. 5.88 ± 2.2532
- e. None of the above.

Problem 1.3. (5 points) *Source: "Mathematical Statistics with Applications in R" by Ramachandran and Tsokos.*

A cross is hypothesized to result in a 3 : 1 phenotypic ratio of red-flowered to white-flowered plants. You set up a hypothesis test to test this claim. Suppose your cross actually produces 170 red- and 30 white-flowered plants. What is the p -value you obtain?

- a. Less than 0.005.
- b. Between 0.005 and 0.01.
- c. Between 0.01 and 0.025.
- d. Between 0.025 and 0.05.
- e. None of the above.

Problem 1.4. (5 points) A die is rolled 60 times and the face values are recorded. The results are as follows:

Up face	1	2	3	4	5	6
Number of occurrences	8	11	5	12	15	9

You test the hypothesis that the die is fair. What can you say about the p -value?

- (a) Less than 0.05.
- (b) Between 0.05 and 0.10.
- (c) Between 0.10 and 0.20.

- (d) Between 0.20 and 0.30.
- (e) It's greater than 0.30.

Problem 1.5. The final exam in a particular course has 100 multiple-choice questions: for each question there are five offered answers exactly one of which is correct. Out of the 100 questions, 36 questions come from a public problem bank. A student diligently memorizes the correct answers to all of those questions. However, since the student learned by rote, they are not able to do any work on the remaining questions. So, in the exam, they are able to answer exactly 36 questions correctly. For the remaining questions, the student guesses completely at random and independently between problems. Approximately, what is the probability that the student achieves a passing score of 65?

- (a) Closest to zero.
- (b) Closest to 0.001.
- (c) Closest to 0.01
- (d) Closest to 0.1.
- (e) Closest to 0.5.

Problem 1.6. (5 points) An improvement in a process for manufacturing is being considered. Samples are taken both from parts manufactured using the existing method and the new method. It is found that 50 out of a sample of 1000 items produced using the existing method are defective. It is also found that 40 out of a sample of 1600 items produced using the new method are defective. The two samples are independent.

Find the 80%-confidence interval for the true difference in the proportions of defectives produced using the existing and the new method. *Note: Round your point estimate and the margin of error to four places after the decimal point.*

- (a) (0.0149, 0.0351)
- (b) (0.0171, 0.0329)
- (c) (0.0120, 0.0380)
- (d) (0.0095, 0.0405)
- (e) None of the above.

Problem 1.7. (5 points) Let the random sample X_1, \dots, X_{10} be drawn from a normal distribution with mean 2 and variance 1. Define

$$Y = \sum_{i=1}^{10} (X_i - 2)^2.$$

Find the constant q such that

$$\mathbb{P}[Y \leq q] = 0.975.$$

- (a) 18.31
- (b) 19.02
- (c) 20.48
- (d) 21.92
- (e) None of the above.

Problem 1.8. (5 points) In a simple random sample of 1000 Austinites owning televisions, it is found that 480 do not have cable (but do have Netflix or some such or just game on the big screen). Find an 92% confidence interval for the true proportion of Austinites with television who do not have cable. *Note: Round the margin of error to four places after the decimal point.*

- (a) 0.48 ± 0.0277
- (b) 0.48 ± 0.0260
- (c) 0.48 ± 0.0310
- (d) 0.48 ± 0.0158
- (e) None of the above.

Problem 1.9. Let p_m and p_f be the population proportions of male and female warblers who return to their hatching site. You want to test whether the two proportions are different. The observed number of males who returned is 135 out of 900, while the observed number of females who returned is 84 out of 700. What is your decision for this hypothesis test? *Note: Keep four significant places after the decimal point for all your point estimates.*

- (a) Reject at the 0.01 significance level.
- (b) Fail to reject at the 0.01 significance level; reject at the 0.025 significance level.
- (c) Fail to reject at the 0.025 significance level; reject at the 0.05 significance level.
- (d) Fail to reject at the 0.05 significance level; reject at the 0.10 significance level.
- (e) Fail to reject at the 0.10 significance level.

Problem 1.10. (5 points) *Source: "Probability and Statistical Inference" by Hogg, Tanis, and Zimmerman.*

A study was conducted to determine whether there is an association between the choice of the most credible media source for reporting news and the education level. The results are displayed in the following table:

	Newspaper	Television	Radio	Total
Grade School	45	22	6	73
High School	94	115	30	239
College	49	52	13	114
Total	188	189	49	426

Your goal is to test whether the choice of the most credible medium is independent from the education level. The observed value of the relevant test statistic is 11.399. What is your decision?

- (a) Reject at the 0.01 significance level.
- (b) Fail to reject at the 0.01 significance level; reject at the 0.025 significance level.
- (c) Fail to reject at the 0.025 significance level; reject at the 0.05 significance level.
- (d) Fail to reject at the 0.05 significance level; reject at the 0.10 significance level.
- (e) Fail to reject at the 0.10 significance level.

Problem 1.11. An experiment was designed to test whether people's mean reaction times to an auditory stimulus are different depending on whether the room temperature is 68F or 85F. Upon hearing a sound, the subjects would hit a button and the reaction time (in seconds) would be recorded. The reaction time (in seconds) of 25 subjects was recorded. The average reaction time for the cool setting was 0.1825 seconds, and the average reaction time for the warm setting was 0.1485 seconds. The sample standard deviation of the differences between reaction times was 0.0465. What is the 80%-confidence interval for the difference in mean reaction times? Assume the normal model for the reaction times.

- (a) 0.034 ± 0.0163
- (b) 0.034 ± 0.0465
- (c) 0.034 ± 0.0093
- (d) 0.034 ± 0.0128
- (e) None of the above.

Problem 1.12. Sixteen bright middle schoolers were put on an intense sudoku regimen in the hope of improving their scores in the *Canvas Algebra Test (CAT)*. Assume that the distribution of the differences in the scores is normal. Let μ_d be the mean difference between the "pre-test" (before the regimen) and "post-test" (after the regimen). We want to test

$$H_0 : \mu_d = 0 \quad \text{vs.} \quad H_a : \mu_d > 0.$$

The observed average difference in the scores was $\bar{x}_d = 1$ while the sample standard deviation was $s_d = 3$. What can you say about the p -value for this hypothesis test?

- (a) It's below 0.05.
- (b) It's between 0.05 and 0.10.
- (c) It's between 0.10 and 0.15.
- (d) It's between 0.15 and 0.20.
- (e) It's over 0.20.