STA 100A SSI 2023 – Practice Final

July 27, 2023

Name: _____

Student ID: _____

- Testing time: 90 minutes.
- Format: Closed-book, calculator allowed, 5 written answer questions, each with multiple parts.
- Scoring: Each question is worth 20 points.
- Scratch work: One page provided for scratch work.
- Required information: Print your **name** and fill in your student **ID** number (nine digits).
- Space limitation: Written answers and supporting work must **fit in the provided spaces**.
- Supporting work: Answers without supporting work may not receive credit.
- Submission: **Remove scratch paper, formula sheet, and statistical tables** before submission.

1. For a certain group of people, the probability they have a specific disease allele is 0.50. If they have this gene, the probability of developing kidney cancer is 0.30. If they do not have this gene, the probability of developing kidney cancer is 0.01. Assume a person was randomly selected from this group.

   (a) Find the probability that he/she develops kidney cancer.

   (b) Given that the person develops kidney cancer, what is the probability that he/she has the gene?

(c) Find the probability that he/she either develops kidney cancer, has the gene, or both.

(d) If he/she does not develop kidney cancer, find the probability that he/she has the gene.

(e) Using a normal approximation with continuity correction, what is the probability of having at least 50 people with the gene if we randomly sample 100 individuals?

2. Two treatment costs for a particular medical issue are being compared. The average cost of a sample of 50 randomly selected subjects for Treatment A was \$490, with a standard deviation of \$32. The average cost for a sample of 70 randomly selected subjects for Treatment B was \$500, with standard deviation \$48. A medical researcher claims that Treatment A (μ_1) is cheaper than Treatment B (μ_2) on average. Test the researcher's claim at $\alpha = 0.05$. Assume the degrees of freedom $\nu = 100$.

   (a) State the null and alternative hypotheses.

   (b) Calculate the test statistic.

(c) Find the range of p-value.

(d) Do you reject or fail to reject the null?

(e) Would the upper one-sided 95% confidence interval for $\mu_1 - \mu_2$ include zero? Please explain your answer without calculating the confidence interval.

3. The following data summarizes the incidence of Coronary Heart Disease (CHD) for people who smoked regularly and for people who did not smoke regularly:

	CHD	No CHD
Smoked	84	87
Did not smoke	296	491

Assume we want to test if the incidence of CHD was independent of smoking with $\alpha = 0.05$.

(a) State the null and alternative hypotheses.

(b) Calculate the test statistic.

(c) Find the critical value. Do you reject or fail to reject the null?

(d) Find a 95% confidence interval for the difference in proportion of CHD for smokers versus non-smokers.

(e) Is your confidence interval consistent with the conclusion in (c)? Explain.

4. Different methods of treating cloth to repeal mosquitoes were being tested, with treatments T1 = Chemical I, T2 = Chemical II, and C = Combination of chemicals I and II. The number of bites suffered by the subjects wearing the cloth in one day was measured, and summary statistics follow:

Treatment	T1	T2	C
$\bar{Y}_i$	8	8	5
s_i	3	3	3
n_i	21	21	21

Test whether there is a significant difference among the three different treatments in terms of the mean number of bites suffered using $\alpha = 0.01$. Denote the population mean number of bites with Treatment T1, T2, and C as μ_1, μ_2, and μ_3, separately. Assume the number of bites with any of the three treatments is normally distributed with equal standard deviations.

(a) Calculate MSB and MSW.

(b) State the null and alternative hypotheses.

(c) Calculate the test statistic.

(d) Find the range of p-value.

(e) Do you reject or fail to reject the null.

5. 60 guinea pigs were measured, and the amount of vitamin C (in mg) and the length of their adult teeth (in mm) was measured. We believe that vitamin C may increase the length of their teeth. Summary statistics are:

	Vitamin C (mg)	Tooth length (mm)
Sample mean	1.17	18.81
Sample standard deviation	0.63	7.65

In addition, $r = 0.803$, and SSE $= 1227.905$.

(a) Find the fitted regression line.

(b) If a guinea pig that received 2 mg of vitamin C has a tooth length of 29.4 mm, find the prediction error based on your fitted regression line. Did the line over estimate or under estimate the actual value?

(c) Find and interpret the residual standard deviation s_e.

(d) Find and interpret the coefficient of determination r^2.

(e) Find the 95% confidence interval for the slope, and interpret it in terms of the problem.

SCRATCH PAPER
Remove this page

SCRATCH PAPER
Remove this page