Homework 3

ESE 402/542

October 2, 2024

Problem 1. Suppose that $X_1, \ldots, X_n$ are an i.i.d. random sample of size n with sample mean $\bar{X} = 12$ and sample variance $s^2 = 5$.

(a) Let $n = 5$ and suppose the samples are drawn from a Normal distribution with unknown mean μ and known variance $\sigma^2 = 9$. Let the null hypothesis be $H_0 : \mu = 10$ and the alternative hypothesis be $H_a : \mu \neq 10$. Calculate the relevant test statistic value and p-value. Determine the decision rule for $\alpha = 0.05$.

(b) Using the acceptance region of this test, construct a 95% confidence interval. (Hint: Think about how we can reverse our hypothesis test to construct the equivalent 95% confidence interval)

Problem 2. Federal investigators identified a strong association between chemicals in drywall and electrical problems, and there is also strong evidence of respiratory difficulties due to emission of hydrogen sulfide gas. An extensive examination of 51 homes found that 41 had such problems. Suppose that 51 were randomly sampled from the population of all homes having drywall.

(a) Does the data provide strong evidence for concluding that more than 50% of all homes with drywall have electrical/environmental problems? Carry out a test of hypotheses using $\alpha = 0.01$.

(b) Calculate a lower bound of a 99% confidence interval for the percentage of all such homes that have electrical/environmental problems.

Problem 3. Consider a random sample of size $n = 100$ with sample proportion $\hat{p} = 0.2$ from a population with a true unknown proportion p.

(a) For the test $H_0 : p = 0.25$ versus $H_a : p < 0.25$, calculate the relevant test statistic value and p-value. Determine the decision rule for $\alpha = 0.05$ and $\alpha = 0.01$.

(b) For the test $H_0 : p = 0.25$ versus $H_a : p \neq 0.25$, calculate the relevant test statistic value and p-value. Determine the decision rule for $\alpha = 0.05$ and $\alpha = 0.01$.

Problem 4. Recall that α represents the probability of a type I error. On the other hand, β represents the probability of a type II error. The power of a test is the probability that the null hypothesis is rejected when it is false, and is therefore defined as $1 - \beta$. For this problem, you will explore how power depends on multiple factors.

(a) Suppose that $X_1, \ldots, X_n$ are an i.i.d. random sample of size $n = 15$ drawn from a Normal distribution with unknown mean μ and known variance $\sigma^2 = 4$. Using Python, plot the power of the test $H_0 : \mu = 0$ versus $H_a : \mu \neq 0$ as the true mean μ varies at level $\alpha = 0.05$. Make sure to display your graph clearly.

(b) Now using the same information from part (a) except for fixing μ at $\mu = 3$, plot the power of the test as α varies.

(c) Fix α back to $\alpha = 0.05$ and plot the power of the test as n varies.

(d) Fix n back to $n = 15$ and plot the power of the test as σ^2 varies.

(e) Compare and interpret your results.

Problem 5. Suppose that $X_1, \ldots, X_n$ are an i.i.d. random sample of size n drawn from a Poisson distribution with unknown parameter λ. Find the likelihood ratio for testing $H_0 : \lambda = \lambda_0$ versus $H_a : \lambda = \lambda_1$ where $\lambda_1 > \lambda_0$. Use the fact that the sum of independent Poisson random variables follows a Poisson distribution to explain how to determine a rejection region for a test at significance level α_0.

Problem 6. Let $X_1, \ldots, X_n$ be a random sample from an exponential distribution with the density function $f(x|\theta) = \theta \exp(-\theta x)$. Derive a likelihood ratio test of $H_0 : \theta = \theta_0$ verses $H_A : \theta \neq \theta_0$, and show that the rejection region is of the form $\bar{X} \exp(-\theta_0 \bar{X}) \leq c$.

Problem 7. We have access to a data set $X_1, X_2, \cdots, X_n$ where X_i 's are generated i.i.d, according to a distribution with the following pdf:

$$f(x \mid \sigma) = \frac{1}{2\sigma} e^{-\frac{|x|}{\sigma}}$$

For this distribution, we know that $\mathbb{E}[X_i] = 0, \mathbb{E}[|X_i|] = \sigma$, and $\operatorname{Var}(X_i) = 2\sigma^2$. We consider a hypothesis testing problem with $H_0 : \sigma = \sigma_0$ and $H_a : \sigma = \sigma_1$. You may assume that $\sigma_0 < \sigma_1$.

(a) Write down the likelihood-ratio test-statistic for this setting. (You answer should be a function $T(X_1, \cdots, X_n)$).

(b) Given a significance level α, specify the acceptance/rejection regions for the test-statistic obtained in part (a). [Hint: (i) By simplifying the likelihood ratio, show that the acceptance/rejection regions can be written in terms of the quantity $\frac{1}{n} \sum_{i=1}^{n} |X_i|$. (ii) Use the central limit theorem to argue that $\frac{1}{n} \sum_{i=1}^{n} |X_i|$ has a Gaussian distribution. (iii) Use this Gaussian distribution to derive the acceptance/rejection regions based on the significance level α.]

(c) Consider the likelihood-ratio test in part (a) with the acceptance/rejection regions derived in part (b). Let t_{data} denote the realized value of the test on the sample data; i.e., if the sample data is $x_1, x_2, \cdots, x_n$, then $t_{\text{data}} = T(x_1, x_2, \cdots, x_n)$. Calculate the p -value for t_{data} and simplify your answer as much as you can.