Statistical Inference

HW Author: Ali Mollahosseini, Pouya Gohari, Parham Sazdar
Instructor: Mohammadreza A. Dehaqani

Spring 2024

HW 3: Part I

- Since your midterm is coming up, we didn't want to keep you busy with computer assignments and topics that are not covered in your exam, so we broke this homework into two parts. Part I is lighter and midterm themed so prepared for a bigger Part II.
- Please note that this is the bare minimum number of questions you need to solve in preparation for your mideterm.
- If you have any questions about the homework, don't hesitate to use the class group to ask questions or drop an email to the HW Authors.
- Please consult the course page for important information on submission guidelines and delay policies to ensure your homework is turned in correctly and on time.
- This course aims to equip you with the skills to tackle all problems in this domain and encourages you to engage in independent research. Utilize your learnings to extend beyond the classroom teachings where necessary.

Problem 1: MoM

An urn contains B black balls and N - B white balls. A sample of n balls is to be selected without replacement. Let Y denote the number of black balls in the sample. Show that (N/n)Y is the method-of-moments (MOM) estimator of B.

Problem 2: Cramér-Rao

Let $X_1, \ldots, X_n$ be i.i.d. random variables from an exponential distribution with parameter $\lambda > 0$, which has the following probability density function:

$$f(x; \lambda) = \lambda e^{-\lambda x}, \text{ for } x > 0$$

Consider two unbiased estimators of λ:

- $\hat{\lambda}_1 = 1/\bar{X}$, where $\bar{X}$ is the sample mean.
- $\hat{\lambda}_2 = 2/(X_1 + X_n)$, which uses only the first and last observations.

1. Find the variances of $\hat{\lambda}_1$ and $\hat{\lambda}_2$.
2. Determine the Cramér-Rao lower bound for unbiased estimators of λ.
3. Which estimator is more efficient, $\hat{\lambda}_1$ or $\hat{\lambda}_2$? Justify your answer using the concept of efficiency and the Cramér-Rao inequality.

Problem 3: GLRT

Let $X_1, \ldots, X_n$ be i.i.d. from an exponential distribution with unknown mean θ. Consider testing $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$.

1. Derive the generalized likelihood ratio test for this setup.
2. Find the asymptotic distribution of $-2 \log \Lambda$ under H_0, where Λ is the likelihood ratio.

Problem 4: Playing around with the hypothesis

Suppose that a single observation X is taken from a uniform density on $[0, \theta]$, and consider testing $H_0 : \theta = 1$ versus $H_1 : \theta = 2$.

1. Find a test that has significance level $\alpha = 0$. What is its power?
2. For $0 < \alpha < 1$, consider the test that rejects when $X \in [0, \alpha]$. What is its significance level and power?
3. What is the significance level and power of the test that is rejected when $X \in [1 - \alpha, 1]$?
4. Find another test that has the same significance level and power as the previous one.
5. Does the likelihood ratio test determine a unique rejection region?
6. What happens if the null and alternative hypotheses are interchanged-$H_0 : \theta = 2$ versus $H_1 : \theta = 1$?

Problem 5: Largest Possible Power

Consider two probability density functions on $[0, 1]$: $f_0(x) = 1$, and $f_1(x) = 2x$. Among all tests of the null hypothesis $H_0 : X \sim f_0(x)$ versus the alternative $X \sim f_1(x)$, with significance level $\alpha = 0.10$, how large can the power possibly be?

Problem 6: Hypothesis about Distribution + more

1. For data $X_1, \ldots, X_n \in \mathbb{R}$ and two fixed and known values $\sigma_0^2 < \sigma_1^2$, consider the following testing problem:

$$H_0 : X_1, \ldots, X_n \sim N(0, \sigma_0^2)$$
$$H_1 : X_1, \ldots, X_n \overset{IID}{\sim} N(0, \sigma_1^2)$$

   What is the most powerful test for testing H_0 versus H_1 at level α? Letting $\chi_n^2(\alpha)$ denote the $1 - \alpha$ quantile of the χ_n^2 distribution, describe explicitly both the test statistic T and the rejection region for this test.

2. What is the distribution of this test statistic T under the alternative hypothesis H_1? Using this result, and letting F denote the CDF of the χ_n^2 distribution, provide a formula for the power of this test against H_1 in terms of $\chi_n^2(\alpha)$, σ_0^2, σ_1^2, and F. Keeping σ_0^2 fixed, what happens to the power of the test as σ_1^2 increases to ∞?