November 11, 2022

Tutorial 8 - Probability and Statistics

Problem 1

The following data is from a random sample: 5, 1, 3, 3, 8.
Compute the sample mean and sample standard deviation.

Problem 2

Let X_1, X_2, X_3 be a random sample of size $n = 3$ from a distribution with the geometric probability mass function:

$$f(x) = \left(\frac{3}{4}\right)\left(\frac{1}{4}\right)^{x-1}$$

for $x = 1, 2, 3, \ldots$. What is $P(\max X_i \leq 2)$?

Problem 3

If X_i is a Bernoulli random variable with parameter p, then:

$$\hat{p} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

is the maximum likelihood estimator (MLE) of p. Is the MLE of p an unbiased estimator of p?

Problem 4

If X_i are normally distributed random variables with mean μ and variance σ^2, then:

$$\hat{\mu} = \frac{\sum X_i}{n} = \bar{X} \text{ and } \hat{\sigma}^2 = \frac{\sum(X_i - \bar{X})^2}{n}$$

are the maximum likelihood estimators of μ and σ^2, respectively. Are the MLEs unbiased for their respective parameters?

Tutorial 8 - Probability and Statistics

Problem 5

Let $Y_1 < Y_2 < Y_3 < Y_4 < Y_5 < Y_6$ be the order statistics associated with n = 6 independent observations each from the distribution with probability density function:

$$f(x) = \frac{1}{2}x \quad ; \text{for } 0<x<2$$

What is the probability that the next-to-largest order statistic, that is, Y_5, is less than 1?

Problem 6

Example 2: Suppose $X_1, X_2, \cdots, X_n$ are i.i.d. random variables with density function $f(x|\sigma) = \frac{1}{2\sigma} \exp\left(-\frac{|x|}{\sigma}\right)$, please find the maximum likelihood estimate of σ.

Problem 7

Example 3. Light bulbs
Suppose that the lifetime of *Badger* brand light bulbs is modeled by an exponential distribution with (unknown) parameter λ. We test 5 bulbs and find they have lifetimes of 2, 3, 1, 3, and 4 years, respectively. What is the MLE for λ?

Problem 8

Suppose our data $x_1, \ldots x_n$ are independently drawn from a uniform distribution $U(a, b)$. Find the MLE estimate for a and b.

November 11, 2022

Tutorial 8 - Probability and Statistics

Problem 9

Example 7. Hardy-Weinberg. Suppose that a particular gene occurs as one of two alleles (A and a), where allele A has frequency θ in the population. That is, a random copy of the gene is A with probability θ and a with probability $1-\theta$. Since a diploid genotype consists of two genes, the probability of each genotype is given by:

genotype	AA	Aa	aa
probability	θ^2	$2\theta(1-\theta)$	$(1-\theta)^2$

Suppose we test a random sample of people and find that k_1 are AA, k_2 are Aa, and k_3 are aa. Find the MLE of θ.

Problem 10

Example 2.2.2 (Weibull with known α) $\{Y_i\}$ *are iid random variables, which follow a Weibull distribution, which has the density*

$$\frac{\alpha y^{\alpha-1}}{\theta^\alpha}\exp(-(y/\theta)^\alpha) \qquad \theta,\alpha>0.$$

Suppose that α is <u>known</u>, but θ is <u>unknown</u>. Our aim is to fine the MLE of θ.

Tutorial 8 - Probability and Statistics

Problem 11

Let X_1, X_2, X_3, ..., X_n be a random sample from the following distribution:

$$f_X(x) = \begin{cases} \theta\left(x - \frac{1}{2}\right) + 1 & \text{for } 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

where $\theta \in [-2, 2]$ is an unknown parameter. We define the estimator $\hat{\Theta}_n$ as

$$\hat{\Theta}_n = 12\overline{X} - 6$$

to estimate θ.

(a) Is $\hat{\Theta}_n$ an unbiased estimator of θ?

(b) Is $\hat{\Theta}_n$ a consistent estimator of θ?

(c) Find the mean squared error (MSE) of $\hat{\Theta}_n$.

Problem 12

Let X be one observation from a $N(0, \sigma^2)$ distribution.

(a) Find an unbiased estimator of σ^2.

(b) Find the log likelihood, $\log(L(x; \sigma^2))$, using

$$f_X(x; \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} exp\left\{-\frac{x^2}{2\sigma^2}\right\}$$

as the PDF.

(c) Find the Maximum Likelihood Estimate (MLE) for the standard deviation σ, $\hat{\sigma}_{ML}$.

Tutorial 8 - Probability and Statistics

Problem 13

15. Let X_1, X_2, X_3, X_4, X_5 be a random sample from a $N(\mu, 1)$ distribution, where μ is unknown. Suppose that we have observed the following values

$$5.45, \quad 4.23, \quad 7.22, \quad 6.94, \quad 5.98$$

We would like to decide between

H_0: $\mu = \mu_0 = 5$,

H_1: $\mu \neq 5$.

(a) Define a test statistic to test the hypotheses and draw a conclusion assuming $\alpha = 0.05$.

(b) Find a 95% confidence interval around $\overline{X}$. Is μ_0 included in the interval? How does the exclusion of μ_0 in the interval relate to the hypotheses we are testing?

Tutorial 8 - Probability and Statistics

Problem 14

Let X_1, X_2 ,..., X_{150} be a random sample from an unknown distribution. After observing this sample, the sample mean and the sample variance are calculated to be as follows:

$$\overline{X} = 52.28, \qquad S^2 = 30.9$$

Design a level 0.05 test to choose between

H_0: $\mu = 50$,

H_1: $\mu > 50$.

Do you accept or reject H_0?

Problem 15

Let X_1, X_2 ,..., X_{121} be a random sample from an unknown distribution. After observing this sample, the sample mean and the sample variance are calculated to be as follows:

$$\overline{X} = 29.25, \qquad S^2 = 20.7$$

Design a test to decide between

H_0: $\mu = 30$,

H_1: $\mu < 30$,

and calculate the P-value for the observed data.

November 11, 2022

Tutorial 8 - Probability and Statistics

Problem 16

Example 2.9 Let $X_1, X_2, \ldots, X_n$ be a random sample from $f(x, \alpha, \beta) = \beta e^{-\beta(x-\alpha)}$; $\alpha \leq x < \infty$ and $\beta > 0$.

Find MLE's of α, β.

Problem 17

Example 2.10 Let $X_1, X_2, \ldots, X_n$ be a random sample from $f(x, \alpha, \beta) =$

$$\begin{cases} \frac{1}{\beta-\alpha}, & \alpha \leq x \leq \beta \\ 0, & \text{Otherwise} \end{cases}$$

(a) Show that the MLE of (α, β) is (Min X_i, Max X_i).

(b) Also find the estimators of α and β by the method of moments.

Problem 18

Example 5.2 Let $X_1, X_2, \ldots, X_n$ be a random sample from $N(\mu, \sigma^2)$. Find $(1 - \alpha)$ level confidence interval for σ^2 when (i) μ is known and (ii) μ is unknown.

Problem 19

Example 5.3 Let $X_1, X_2, \ldots, X_n$ be a random sample from density function $f(x|\theta) = \left(\frac{1}{\theta}\right)$, $0 < x < \theta$. Find $100(1 - \alpha)\%$ confidence interval of θ.

Tutorial 8 - Probability and Statistics

Problem 20

There are two candidates in a presidential election: Candidate A and Candidate B. Let θ be the portion of people who plan to vote for Candidate A. Our goal is to find a confidence interval for θ. Specifically, we choose a random sample (with replacement) of n voters and ask them if they plan to vote for Candidate A. Our goal is to estimate the θ such that the margin of error is 3 percentage points. Assume a 95% confidence level. That is, we would like to choose n such that

$$P\left(\overline{X} - 0.03 \leq \theta \leq \overline{X} + 0.03\right) \geq 0.95,$$

where $\overline{X}$ is the portion of people in our random sample that say they plan to vote for Candidate A. How large does n need to be?