Basic statistics : Assignment

shala 2020

1 Distribution Functions

1. The probability density function for an exponential random variable is given by

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & , x \geq 0 \\ 0 & , x < 0 \end{cases}$$

   Find the mean and variance. Find the cumulative distribution function $F(X)$ and verify that it satisfies the properties of a CDF.

2. A game of 7up 7down has the following rules :
   You are given a pair of dice and have to estimate the sum you will obtain, before rolling the dice. Your estimate must lie in exactly one of the 3 categories : less than 7, equal to 7, and greater than 7.
   Find the probability of each category and the cumulative distribution function of the sum. Using the CDF obtained, find the smallest 'n' such that in a game of n up n down, the probability of n down (sum less than n) would be greater than 75% .

3. The probability mass function of a Poisson random variable is given below. Find its mean.

$$p(x) = \begin{cases} \frac{\lambda^n e^{-\lambda}}{n!} & , n \geq 0 \\ 0 & , n < 0 \end{cases}$$

4. A 3 sided die with faces 1, 2, 3 is rolled twice. Let random variables X, Y capture the sum and absolute difference of the outcomes. Find the joint mass function and the conditional distribution P(Y|X=4).

5. The joint density of random variables X and Y is given as

$$f(x,y) = \begin{cases} xe^{-(x+y)} & , x > 0,\ y > 0 \\ 0 & , \text{otherwise} \end{cases}$$

Find the marginal densities f(x) and f(y). Comment on the independence of X and Y.

2 MLE and Hypothesis Testing

1. Find the joint distribution for 'n' iid Bernoulli(p) RVs $X_1, X_2, ..., X_n$ using independence. Recall that a compact form of the pmf was :

$$prob(x) = p^x(1-p)^{1-x}$$

Differentiate the log likelihood function to obtain the MLE for p.

2. The uniform continuous iid RVs $X_1, X_2, ..., X_n$ have the pdf :

$$f(X) = \begin{cases} \frac{1}{N} & ;\ x \in [0, N] \\ 0 & ;\ \text{otherwise} \end{cases}$$

Find the joint distribution using iid assumption. Find the MLE for N (what value of N maximises the likelihood? Will differentiating the log likelihood help?)

3. Let $X_1, X_2, X_3, \ldots, X_n$ be iid $\mathcal{N}(\mu, 1)$ RVs, where μ is unknown. Let $H_0 : X_i \sim \mathcal{N}(\mu_0, 1)$, $H_1 : X_i \sim \mathcal{N}(\mu_1, 1)$ be the null and the alternate hypotheses, respectively, i.e. $\mu \in \{\mu_0, \mu_1\}$. Find a test $\varphi(\mathbf{X})$ that accepts the null hypothesis with significance α.
The significance level, is a measure of the strength of the evidence that must be present in your sample before you will reject the null hypothesis and conclude that the effect is statistically significant. Mathematically, The mapping φ is said to be a test of hypothesis $H_0 : \boldsymbol{\theta} \in \Theta_0$ against the alternatives $H_1 : \boldsymbol{\theta} \in \Theta_1$, with error probability α (also called level of significance or, simply, level) if

$$E_{\boldsymbol{\theta}}\varphi(\boldsymbol{X}) \leq \alpha \ \ \forall\ \boldsymbol{\theta} \in \Theta_0$$