SI 140A-02 Probability & Statistics for EECS, Fall 2024
Homework 6

Name: **Wenye Xiong**
Student ID: 2023533141

Due on Nov. 19, 2024, 11:59 UTC+8

Read all the instructions below carefully before you start working on the assignment, and before you make a submission.

- You are required to write down all the major steps towards making your conclusions; otherwise you may obtain limited points of the problem.
- Write your homework in English; otherwise you will get no points of this homework.
- Any form of plagiarism will lead to 0 point of this homework.

Problem 1

Suppose a fair coin is tossed repeatedly, and we obtain a sequence of H and T(H denotes Head and T denotes Tail). Let N denote the number of tosses to observe the first occurrence of the pattern "HH". Find the PMF of N.

Solution

For $k \geq 3$ let N denote the need steps to find the first HH , $p_k = p(N = k)$, by using the first step method , if first coin is T , will find HH in the rest k-1 coins , the Probability is $\frac{1}{2}p_{k-1}$ for this condition because it is a fair coin .
If the first step is H , the second step must be T , so the Probability for this condition is $\frac{1}{4}p_{k-2}$, so we will have the equation $p_k = \frac{1}{2}p_{k-1} + \frac{1}{4}p_{k-2}$ solve this difference equation , by $p_1 = 0, p_2 = \frac{1}{4}, p_3 = \frac{1}{8}$, we will find

$$p(N = k) = \frac{5+\sqrt{5}}{40}(\frac{\sqrt{5}+1}{4})^{k-2} - \frac{\sqrt{5}}{10}(-\frac{\sqrt{5}-1}{4})^{k-1}$$

Problem 2

The Cauchy distribution has PDF

$$f(x) = \frac{1}{\pi\left(1 + x^2\right)}$$

for all x. Find the CDF of a random variable with the Cauchy PDF. Hint: Recall that the derivative of the inverse tangent function $\tan^{-1}(x)$ is $\frac{1}{1+x^2}$.

Solution

The CDF of a Cauchy is F given by

$$F(t) = \int_{-\infty}^{t} \frac{1}{\pi(1+x^2)} dx = \frac{1}{\pi} \int_{-\infty}^{t} \frac{1}{1+x^2} dx = \frac{1}{\pi} \tan^{-1}(x)\Big|_{-\infty}^{t} = \frac{1}{\pi} \tan^{-1}(t) + \frac{1}{2}$$

Problem 3

The Pareto distribution with parameter $a > 0$ has PDF

$$f(x) = \frac{a}{x^{a+1}}$$

for $x \geq 1$ (and 0 otherwise). This distribution is often used in statistical modeling. Find the CDF of a Pareto r.v. with parameter a; check that it is a valid CDF.

Solution

The CDF of a Pareto is F given by

$$F(t) = \int_1^t \frac{a}{x^{a+1}} dx = (-t^{-a})\Big|_1^t = 1 - \frac{1}{t^a}$$

For all $t \geq 1$, $F(t) = 1 - \frac{1}{t^a} \geq 0$. This is a valid CDF since it is increasing in y (this can be seen directly or from the fact that F 0 = f is nonnegative), right continuous (in fact it is continuous), and has limit 0 as x approaches negative infinity and limit 1 as x approaches positive infinity.

Problem 4

The *Beta distribution* with parameters $a = 3, b = 2$ has PDF

$$f(x) = 12x^2(1-x), \text{ for } 0 < x < 1$$

Let X have this distribution.

(a) Find the CDF of X.

(b) Find $P(0 < X < 1/2)$.

(c) Find the mean and variance of X (without quoting results about the Beta distribution).

Solution

(a):
The CDF of X is

$$F(X) = \int_0^x 12t^2(1-t)dt = 4t^3 - 3t^4\Big|_0^x = 4x^3 - 3x^4 \quad \text{for } 0 < x < 1$$

(b):
According to the CDF of X, we have

$$P(0 < X < 1/2) = F(1/2) = \tfrac{5}{16}$$

(c):
According to PDF of X, the mean of X is

$$E(X) = \int_0^1 xf(x)dx = \int_0^1 12x^3(1-x)dx = \int_0^1 12x^3dx - \int_0^1 12x^4dx = \tfrac{3}{5}$$

We also have:

$$E(X^2) = \int_0^1 x^2f(x)dx = \int_0^1 12x^4(1-x)dx = \int_0^1 12x^4dx - \int_0^1 12x^5dx = \tfrac{2}{5}$$

Thus, the variance of X is

$$Var(X) = E(X^2) - E(X)^2 = \tfrac{2}{5} - (\tfrac{3}{5})^2 = \tfrac{1}{25}$$

Problem 5

The Exponential is the analog of the Geometric in continuous time. This problem explores the connection between Exponential and Geometric in more detail, asking what happens to a Geometric in a limit where the Bernoulli trials are performed faster and faster but with smaller and smaller success probabilities.
Suppose that Bernoulli trials are being performed in continuous time; rather than only thinking about first trial, second trial, etc., imagine that the trials take place at points on a timeline. Assume that the trials are at regularly spaced times $0, \Delta t, 2\Delta t, \ldots$, where Δt is a small positive number. Let the probability of success of each trial be $\lambda \Delta t$, where λ is a positive constant. Let G be the number of failures before the first success (in discrete time), and T be the time of the first success (in continuous time).

(a) Find a simple equation relating G to T. Hint: Draw a timeline and try out a simple example.

(b) Find the CDF of T. Hint: First find $P(T > t)$.

(c) Show that as $\Delta t \to 0$, the CDF of T converges to the Expo(λ) CDF, evaluating all the CDFs at a fixed $t \geq 0$.

Solution

(a):
$T = G\Delta t$
(b):
For $t > 0$, $P(T > t)$ is the probability that no success in the first $\lfloor \frac{t}{\Delta t} \rfloor$ trials,

$$P(T \leq t) = 1 - P(G > \tfrac{t}{\Delta t}) = 1 - (1 - \lambda \Delta t)^{\lfloor \frac{t}{\Delta t} \rfloor + 1}$$

(c):
As $\Delta t \to 0$, we have

$$\lim_{\Delta t \to 0} P(T \leq t) = \lim_{\Delta t \to 0} 1 - (1 - \lambda \Delta t)^{\lfloor \frac{t}{\Delta t} \rfloor + 1} = 1 - e^{-\lambda t}$$

Thus, for all $t \geq 0$, the CDF of T converges to the Expo(λ) CDF as $\Delta t \to 0$.

Problem 6

The *Gumbel distribution* is the distribution of $-\log X$ with $X \sim \text{Expo}(1)$.

(a) Find the CDF of the Gumbel distribution.

(b) Let $X_1, X_2, \ldots$ be i.i.d. Expo(1) and let $M_n = \max(X_1, \ldots, X_n)$. Show that as $n \to \infty$, the CDF of $M_n - \log n$ converges to the Gumbel CDF.

Solution

(a):
Let G be the Gumbel distribution, and $X \sim Expo(1)$. The CDF of G is

$$F(t) = P(G \leq t) = P(-\log X \leq t) = P(X \geq e^{-t}) = e^{-e^{-t}}$$

(b):
The CDF of $M_n - \log n$ is

$$P(M_n - \log n \leq t) = P(X_1 \leq t + \log n, X_2 \leq t + \log n, \ldots, X_n \leq t + \log n) = (1 - e^{-(t+\log n)})^n = P(X \leq t + \log n)^n = (1 - e^{-(t+\log n)})^n = (1 - \frac{1}{n}e^{-t})^n$$

As $n \to \infty$, we have

$$\lim_{n\to\infty}(1 - \frac{1}{n}e^{-t})^n = e^{-e^{-t}}$$

Problem 7

(**Optional Challenging Problem**)
Let $X \sim \mathcal{N}(0,1)$, its corresponding CDF is denoted as Φ and the corresponding PDF is denoted as φ.

(a) If $x > 0$, show the following inequality holds:

$$\frac{x}{x^2+1}\varphi(x) \leq 1 - \Phi(x) \leq \frac{1}{x}\varphi(x)$$

(b) Define the function $g(x)$ as follows:

$$g(x) = \frac{2}{\sqrt{\pi}} \int_x^{\infty} e^{-t^2} dt, \forall x \geq 0$$

Show the following inequality holds:

$$g(x) \leq e^{-x^2}, \forall x \geq 0$$