

APMA 1655 Honors Statistical Inference I

Homework 4

Name: *Tanish Makadia*

Due: *11 pm, March 10*

Collaborators: *Garv Gaur*

- You are strongly encouraged to work in groups, but solutions must be written independently.

1 Review

1.1 Definition of Discrete Random Variables

Let (Ω, \mathbb{P}) be a probability space. Suppose X is a random variable defined on Ω , and F_X is the CDF of X .

1. We say X is a **discrete random variable** if its CDF F_X is of the following form

$$F_X(x) = \sum_{k=0}^K p_k \cdot \mathbf{1}_{[x_k, +\infty)}(x), \quad (1)$$

where $p_k \geq 0$ for all $k = 1, 2, \dots, K$ and $\sum_{k=0}^K p_k = 1$; the K is allowed to be ∞ .

2. If X is a discrete random variable whose CDF is of the form in Eq. (1), we call the ordered sequence $\{p_k\}_{k=0}^K$ as the **probability mass function (PMF)**^{†1} of X .

1.2 Independence between Events

Let (Ω, \mathbb{P}) be a probability space. Suppose \tilde{A} and \tilde{B} are two events. We say \tilde{A} and \tilde{B} are **independent** if $\mathbb{P}(\tilde{A} \cap \tilde{B}) = \mathbb{P}(\tilde{A}) \cdot \mathbb{P}(\tilde{B})$.

1.3 Independence between Random Variables — Version I

Let Y and Z be two random variables defined on the probability space (Ω, \mathbb{P}) . We say that Y and Z are independent if they satisfy the following **for any** subsets $A \subset \mathbb{R}$ and $B \subset \mathbb{R}$

$$\mathbb{P}(\{\omega \in \Omega : Y(\omega) \in A \text{ and } Z(\omega) \in B\}) = \mathbb{P}(\{\omega \in \Omega : Y(\omega) \in A\}) \cdot \mathbb{P}(\{\omega \in \Omega : Z(\omega) \in B\}).$$

^{†1}: The ordered sequence $\{p_k\}_{k=0}^K$ is conventionally called as a function. You may view the map $k \mapsto p_k$ as a function. I think the reason $\{p_k\}_{k=0}^K$ is called a function is to make the names “PMF” and “PDF” look similar. In addition, if you are comfortable with the concept of vectors, you may view the ordered sequence $\{p_k\}_{k=0}^K$ as a vector (p_0, p_1, \dots, p_K) ; if $K = \infty$, this vector is infinitely long.

1.4 Independence between Random Variables — Version II

Let Y and Z be two random variables defined on the probability space (Ω, \mathbb{P}) . We say that Y and Z are independent if the following is true: **for any** subsets $A \subset \mathbb{R}$ and $B \subset \mathbb{R}$, the following two events are independent

$$\tilde{A} = \{\omega \in \Omega : Y(\omega) \in A\}, \quad \tilde{B} = \{\omega \in \Omega : Z(\omega) \in B\}.$$

2 Problem Set

- (2 points) Let n be a positive integer, and $\Omega \stackrel{\text{def}}{=} \{1, 2, \dots, n\}$. Suppose \mathbb{P} is a function of subsets of Ω defined as follows

$$\mathbb{P}(A) \stackrel{\text{def}}{=} \frac{\#A}{\#\Omega}, \quad \text{for all } A \subset \Omega.$$

We define a random variable X as follows

$$X(\omega) = \omega, \quad \text{for all } \omega \in \Omega = \{1, 2, \dots, n\}.$$

Suppose you have done the following

- You have proved that (Ω, \mathbb{P}) is a probability space (see HW 1).
- You have derived the CDF F_X of X (see HW 3).

Please represent the CDF F_X in the form in Eq. (1). Specifically, please show what the K , $\{p_k\}_{k=0}^K$, and $\{x_k\}_{k=0}^K$ in Eq. (1) should be.

$$\begin{aligned} F_X(x) &= \frac{1}{n} \mathbb{1}_{[1, +\infty)}(x) + \frac{1}{n} \mathbb{1}_{[2, +\infty)}(x) + \dots + \frac{1}{n} \mathbb{1}_{[n, +\infty)}(x) \\ &= \sum_{k=0}^{n-1} \frac{1}{n} \cdot \mathbb{1}_{[k+1, +\infty)}(x) \end{aligned}$$

Therefore,

$$K = n - 1; \quad \{p_k\}_{k=0}^K = \left\{\frac{1}{n}\right\}_{k=0}^{n-1}; \quad \{x_k\}_{k=0}^{n-1} = \{k + 1\}_{k=0}^{n-1}$$

- (2 points) Let Y and Z be random variables defined on the probability space (Ω, \mathbb{P}) ; the distribution of the random variable X defined as follows

$$\begin{aligned} X(\omega) &\stackrel{\text{def}}{=} Y(\omega) + (1 - Y(\omega)) \cdot Z(\omega), \quad \text{for all } \omega \in \Omega, \quad \text{where} \\ Y &\sim \text{Bernoulli}\left(\frac{1}{2}\right), \quad Z \sim N(0, 1), \\ Y \text{ and } Z &\text{ are independent.} \end{aligned} \tag{2}$$

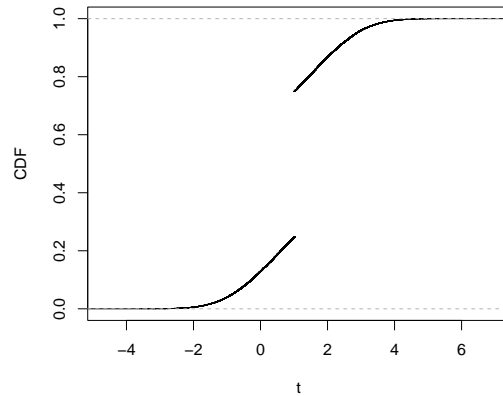


Figure 1: The CDF of the distribution of X defined in Eq. (2). This function is neither continuous nor piecewise constant/step-like.

Then, we claim that the CDF of the random variable X defined in Eq. (2) is the following

$$\begin{aligned} F_X(x) &= \frac{1}{2} \cdot \mathbf{1}_{[1,+\infty)}(x) + \frac{1}{2} \cdot F_Z(x) \\ &= \frac{1}{2} \cdot \mathbf{1}_{[1,+\infty)}(x) + \frac{1}{2} \cdot \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{t^2}{2}} dt, \end{aligned} \quad (3)$$

where F_X denotes the CDF of X , and F_Z denotes the CDF of Z (i.e., the CDF of $N(0, 1)$). The graph of the CDF in Eq. (3) is presented in Figure 1.

Please prove the formula in Eq. (3).

3. (2 points) Let Y , Z , and W be random variables defined on the probability space (Ω, \mathbb{P}) . Suppose

- $Y \sim \text{Bernoulli}(p)$;
- the CDFs of Z and W are F_Z and F_W , respectively;
- Y , Z , and W are mutually independent, i.e., Y and Z are independent, Y and W are independent, Z and W are independent.

We define a new random variable X by $X(\omega) \stackrel{\text{def}}{=} Y(\omega) \cdot Z(\omega) + (1 - Y(\omega)) \cdot W(\omega)$ for all $\omega \in \Omega$. **Please prove that the CDF of X is the following**

$$F_X(x) = p \cdot F_Z(x) + (1 - p) \cdot F_W(x).$$

4. (2 points) Let Y , Z , and W be random variables defined on the probability space (Ω, \mathbb{P}) . Suppose

- $Y \sim \text{Bernoulli}(1/3)$;
- $Z \sim \text{Pois}(1)$;
- $W \sim N(0, 1)$;
- Y , Z , and W are mutually independent, i.e., Y and Z are independent, Y and W are independent, Z and W are independent.

We define a new random variable X by $X(\omega) \stackrel{\text{def}}{=} Y(\omega) \cdot Z(\omega) + (1 - Y(\omega)) \cdot W(\omega)$ for all $\omega \in \Omega$. Let F_X denote the CDF of X . **Please draw the graph of $F_X(x)$ for $-1 \leq x \leq 5.5$, i.e.,**

$$\{(x, F_X(x)) : -1 \leq x \leq 5.5\}.$$

5. (2 points) Let $p_k = \frac{\lambda^k e^{-\lambda}}{k!}$ for all $k = 0, 1, 2, \dots$, where $k!$ denotes the factorial of k ; conventionally, $0! = 1$ (see Wikipedia). **Please prove the following identity**

$$\sum_{k=0}^{\infty} k \cdot p_k = \lambda. \tag{4}$$

Remark: Eq. (4) shows that the “expected value” of $\text{Pois}(\lambda)$. We will discuss the concept of expected values in Chapter 3 of my lecture notes.