Random Variables

Saravanan Vijayakumaran

Department of Electrical Engineering
Indian Institute of Technology Bombay

January 22, 2025

Measurements in Experiments

- In many experiments, we are interested in some real-valued measurement
- Example
  - A coin is tossed twice. We want to count the number of heads which appear.
  - $\Omega = \{HH, HT, TH, TT\}$
  - Let $X(\omega)$ be the number of heads for $\omega \in \Omega$.
  - $X(HH) = 2, X(HT) = 1, X(TH) = 1, X(TT) = 0$
- We are also interested in knowing which measurements are more likely and which are less likely
- The distribution function $F : \mathbb{R} \to [0, 1]$ captures this information where

$$\begin{aligned} F(x) &= \text{Probability that } X(\omega) \text{ is less than or equal to } x \\ &= P\left(\{\omega \in \Omega : X(\omega) \leq x\}\right) \end{aligned}$$

- Is $\{\omega \in \Omega : X(\omega) \leq x\}$ always an event? Does it always belong to the σ-field $\mathcal{F}$ of the experiment?

Random Variables

Definition (Random Variable)

A random variable is a function $X : \Omega \to \mathbb{R}$ with the property that $\{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{F}$ for each $x \in \mathbb{R}$.

Definition (Distribution Function)

The distribution function of a random variable X is the function $F : \mathbb{R} \to [0, 1]$ given by $F(x) = P(X \leq x)$

Examples

- Counting heads in two tosses of a coin.
- Constant random variable

$$X(\omega) = c \ \text{ for all } \ \omega \in \Omega$$

Properties of the Distribution Function

- $P(X > x) = 1 - F(x)$
- $P(x < X \leq y) = F(y) - F(x)$
- If $x < y$, then $F(x) \leq F(y)$
- $\lim_{x \to -\infty} F(x) = 0$
- $\lim_{x \to \infty} F(x) = 1$
- F is right continuous, $F(x + h) \to F(x)$ as $h \downarrow 0$
- $P(X = x) = F(x) - \lim_{y \uparrow x} F(y)$

Discrete Random Variables

Discrete Random Variables

Definition

A random variable is called discrete if it takes values only in some countable subset $\{x_1, x_2, x_3, \ldots\}$ of $\mathbb{R}$.

Definition

A discrete random variable X has a probability mass function $f : \mathbb{R} \to [0, 1]$ given by $f(x) = P[X = x]$

Example

- Bernoulli random variable

  $\Omega = \{0, 1\}$

$$P[X = x] = \begin{cases} p & \text{if } x = 1 \\ 1 - p & \text{if } x = 0 \end{cases}$$

  where $0 \leq p \leq 1$

Properties of the Probability Mass Function

Let F be the distribution function and f be the mass function of a random variable

- $F(x) = \sum_{i:x_i \leq x} f(x_i)$
- $\sum_{i=1}^{\infty} f(x_i) = 1$
- $f(x) = F(x) - \lim_{y \uparrow x} F(y)$

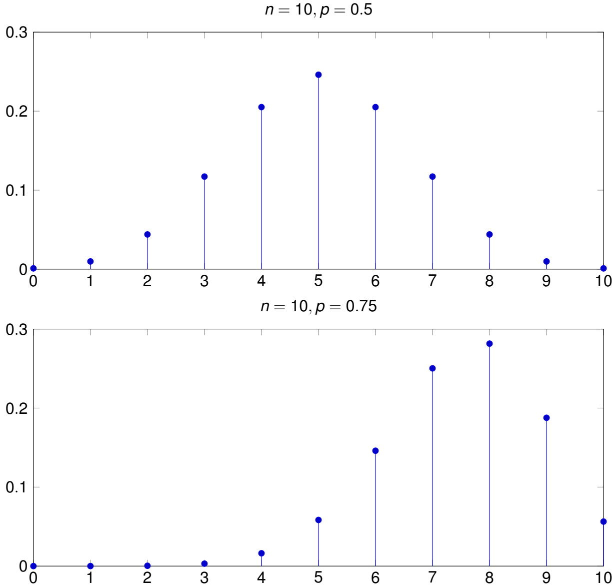
Binomial Random Variable

- An experiment is conducted n times and it succeeds each time with probability p and fails each time with probability $1 - p$
- The sample space is $\Omega = \{0, 1\}^n$ where 1 denotes success and 0 denotes failure
- Let X denote the total number of successes
- $X \in \{0, 1, 2, \ldots, n\}$
- The probability mass function of X is

$$P[X = k] = \binom{n}{k} p^k (1 - p)^{n-k} \quad \text{if } 0 \leq k \leq n$$

- X is said to have the binomial distribution with parameters n and p
- X is the sum of n independent and identically distributed Bernoulli random variables $Y_1 + Y_2 + \cdots + Y_n$

Binomial Random Variable PMF



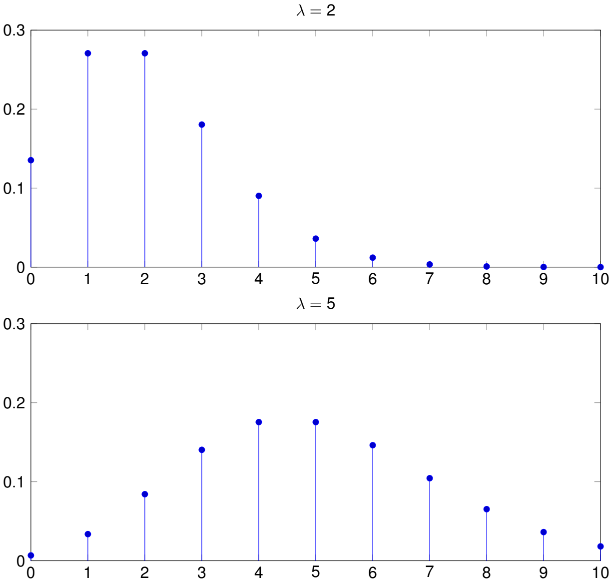
Poisson Random Variable

- The sample space of a Poisson random variable is $\Omega = \{0, 1, 2, 3, \ldots\}$
- The probability mass function is

$$P[X = k] = \frac{\lambda^k}{k!} e^{-\lambda} \quad k = 0, 1, 2, \ldots$$

where $\lambda > 0$

Poisson Random Variable PMF



Independence

- Discrete random variables X and Y are independent if the events $\{X = x\}$ and $\{Y = y\}$ are independent for all x and y
- Examples
  - Two fair dice are rolled. Let X be the number shown by the first die and Y be the sum of the numbers shown by both dice. Are X and Y independent?
  - Binary symmetric channel with crossover probability p. If the input is equally likely to be 0 or 1, are the input and output independent?
- A family of discrete random variables $\{X_i : i \in I\}$ is an independent family if

$$P\left(\bigcap_{i \in J} \{X_i = x_i\}\right) = \prod_{i \in J} P(X_i = x_i)$$

  for all sets $\{x_i : i \in I\}$ and for all finite subsets $J \in I$
- Example

  Let X and Y be independent random variables, each taking values -1 or 1 with equal probability $\frac{1}{2}$. Let $Z = XY$.

  Are X, Y, and Z independent?

Consequences of Independence

- If X and Y are independent, then the events $\{X \in A\}$ and $\{Y \in B\}$ are independent for any subsets A and B of $\mathbb{R}$
- If X and Y are independent, then for any functions $g, h : \mathbb{R} \to \mathbb{R}$ the random variables $g(X)$ and $h(Y)$ are independent
- Exercise
  - Let X and Y be independent discrete random variables taking values in the positive integers
  - Both of them have the same probability mass function given by

$$P[X = k] = P[Y = k] = \frac{1}{2^k} \quad \text{for } k = 1, 2, 3, \ldots$$

  - Find the following
    - $P(\min\{X, Y\} \leq x)$
    - $P[X = Y]$
    - $P[X > Y]$
    - $P[X \geq nY]$ for a given positive integer n
    - $P[X \text{ divides } Y]$

Jointly Distributed Discrete Random Variables

Jointly Distributed Discrete Random Variables

Definition

The joint probability distribution function of discrete RVs X and Y is given by

$$F_{X,Y}(x,y) = P\left(X \leq x \bigcap Y \leq y\right).$$

The joint probability mass function is given by

$$f_{X,Y}(x,y) = P\left(X = x \bigcap Y = y\right).$$

Definition

Given the joint pmf, the marginal pmfs are given by

$$f_X(x) = P(X = x) = \sum_y f_{X,Y}(x,y)$$

$$f_Y(y) = P(Y = y) = \sum_x f_{X,Y}(x,y)$$

Properties of the Joint PMF

- $\sum_x \sum_y f_{X,Y}(x, y) = 1$
- X and Y are independent if and only if

$$f_{X,Y}(x, y) = f_X(x) f_Y(y) \qquad \text{for all } x, y \in \mathbb{R}$$

Exercises

- The joint probability mass function of two discrete random variables X and Y is given by $f(x, y) = c(2x + y)$ where x and y take integer values such that $0 \leq x \leq 2$, $0 \leq y \leq 3$, and $f(x, y) = 0$ otherwise. Find the value of c.
- Given independent random variables $X_1, X_2, \ldots, X_n$ with probability mass functions $f_1, f_2, \ldots, f_n$ respectively, find the probability mass functions of the following
  - $\max(X_1, X_2, \ldots, X_n)$
  - $\min(X_1, X_2, \ldots, X_n)$

Conditional Distribution

Definition

The conditional probability distribution function of Y given $X = x$ is defined as

$$F_{Y|X}(y|x) = P(Y \leq y|X = x)$$

for any x such that $P(X = x) > 0$.
The conditional probability mass function of Y given $X = x$ is defined as

$$f_{Y|X}(y|x) = P(Y = y|X = x)$$

Properties

- $\sum_y f_{Y|X}(y|x) = 1$
- $\sum_x f_{Y|X}(y|x) f_X(x) = f_Y(y)$

Example

Using the clues given below, find the missing entries in the joint probability mass function of X and Y.

Y/X	1	2	3
1	?	?	?
2	?	0	?
3	0	?	0

Table: Joint probability mass function $f_{X,Y}(x, y)$

For $k = 1, 2, 3$,

- $P(Y = 1|X = k) = \frac{2}{3}$
- $P(X = k|Y = 1) = \frac{k}{6}$

Sum of Discrete Random Variables

Theorem

For discrete random variables X and Y with joint pmf $f(x, y)$, the pmf of $X + Y$ is given by

$$P(X + Y = z) = \sum_x f(x, z - x) = \sum_y f(z - y, y)$$

If X and Y are independent, the pmf of $X + Y$ is the convolution of the pmfs of X and Y.

$$P(X + Y = z) = \sum_x f_X(x) f_Y(z - x) = \sum_y f_X(z - y) f_Y(y)$$

Exercise

X is the sum of n independent and identically distributed Bernoulli random variables $Y_1 + Y_2 + \cdots + Y_n$. Can we derive the pmf of X using the above identity?

Continuous Random Variables

Continuous Random Variables

Definition

A random variable is called continuous if its distribution function can be expressed as

$$F(x) = \int_{-\infty}^{x} f(u)\, du \quad \text{for all} \quad x \in \mathbb{R}$$

for some integrable function $f : \mathbb{R} \to [0, \infty)$ called the probability density function of X. If F is differentiable at u, then $f(u) = F'(u)$.

Example

Uniform random variable on $[0, 1]$
$\Omega = [0, 1]$, $X(\omega) = \omega$, $X \sim U[0, 1]$

$$f(x) = \begin{cases} 1 & \text{for } 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

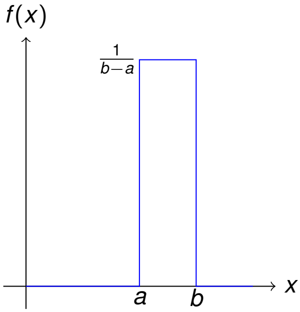
$$F(x) = \begin{cases} 0 & x < 0 \\ x & 0 \leq x \leq 1 \\ 1 & x > 1 \end{cases}$$

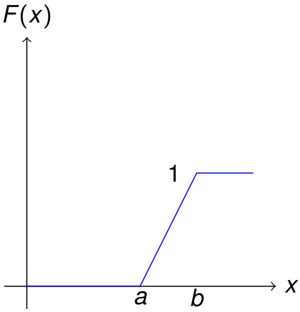
Uniform Random Variable on $[a, b]$

Example

$X \sim U[a, b]$
$\Omega = [a, b]$, $a < b$, $X(\omega) = \omega$,

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{for } a \le x \le b \\ 0 & \text{otherwise} \end{cases} \qquad F(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \le x \le b \\ 1 & x > b \end{cases}$$





Properties of the Probability Density Function

- The numerical value $f(x)$ is not a probability. It can be larger than 1.
- $f(x)dx$ can be intepreted as the probability $P(x < X \leq x + dx)$ since

$$P(x < X \leq x + dx) = F(x + dx) - F(x) \approx f(x)\ dx$$

- $P(a \leq X \leq b) = \int_a^b f(x)\ dx$
- $\int_{-\infty}^{\infty} f(x)\ dx = 1$
- $P(X = x) = 0$ for all $x \in \mathbb{R}$

Independence

- Continuous random variables X and Y are independent if the events $\{X \leq x\}$ and $\{Y \leq y\}$ are independent for all x and y in $\mathbb{R}$
- If X and Y are independent, then the random variables $g(X)$ and $h(Y)$ are independent
- Exercise
  - Let X and Y be independent continuous random variables with common distribution function F and density function f. Find the density functions of $\max(X, Y)$ and $\min(X, Y)$.

Jointly Distributed Continuous Random Variables

Jointly Distributed Continuous Random Variables

Definition

The joint probability distribution function of RVs X and Y is given by

$$F_{X,Y}(x,y) = P\left(X \leq x \bigcap Y \leq y\right) = P(X \leq x, Y \leq y).$$

X and Y are said to be jointly continuous random variables with joint pdf $f_{X,Y}(x,y)$ if

$$F(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(u,v)\, du\, dv$$

for all x, y in $\mathbb{R}$

Definition

Given the joint pdf, the marginal pdfs are given by

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\, dy$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\, dx$$

Properties of the Joint PDF

- $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x, y)\, dx\, dy = 1$
- X and Y are independent if and only if

$$f_{X,Y}(x, y) = f_X(x) f_Y(y) \qquad \text{for all } x, y \in \mathbb{R}$$

Exercise

- The joint probability density function of two continuous random variables X and Y is given by $f(x, y) = c(2x + y)$ where x and y take real values such that $0 \leq x \leq 2$, $0 \leq y \leq 3$, and $f(x, y) = 0$ otherwise. Find the value of c.
- Given independent random variables $X_1, X_2, \ldots, X_n$ with probability density functions $f_1, f_2, \ldots, f_n$ respectively, find the probability density functions of the following
  - $\max(X_1, X_2, \ldots, X_n)$
  - $\min(X_1, X_2, \ldots, X_n)$

Conditional Distribution Function

- For discrete RVs, the conditional distribution was defined as $F_{Y|X}(y|x) = P(Y \leq y|X = x)$ for any x such that $P(X = x) > 0$
- For continuous RVs, $P(X = x) = 0$ for all x
- But considering an interval around x such that $f_X(x) > 0$, we have

$$\begin{aligned} P(Y \leq y|x \leq X \leq x + dx) &= \frac{P(Y \leq y, x \leq X \leq x + dx)}{P(x \leq X \leq x + dx)} \\ &\approx \frac{\int_{v=-\infty}^{y} f(x, v)\, dx\, dv}{f_X(x)\, dx} \\ &= \int_{v=-\infty}^{y} \frac{f(x, v)}{f_X(x)}\, dv \end{aligned}$$

Definition

The conditional distribution function of Y given $X = x$ is the function $F_{Y|X}(\cdot|x)$ given by

$$F_{Y|X}(y|x) = \int_{v=-\infty}^{y} \frac{f(x, v)}{f_X(x)}\, dv$$

for any x such that $f_X(x) > 0$. It is sometimes denoted by $P(Y \leq y|X = x)$.

Conditional Density Function

Definition

The conditional density function of Y given $X = x$ is given by

$$f_{Y|X}(y|x) = \frac{f(x, y)}{f_X(x)}$$

for any x such that $f_X(x) > 0$.

Properties

- $\int_{-\infty}^{\infty} f_{Y|X}(y|x)\, dy = 1$
- $\int_{-\infty}^{\infty} f_{Y|X}(y|x) f_X(x)\, dx = f_Y(y)$

Sum of Continuous Random Variables

Theorem

If X and Y have a joint density function f, then $X + Y$ has density function

$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f(x, z - x)\, dx.$$

If X and Y are independent, then

$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z - x)\, dx = \int_{-\infty}^{\infty} f_X(z - y) f_Y(y)\, dy.$$

The density function of the sum is the convolution of the marginal density functions.

Example (Sum of Uniform RVs)

Let $X \sim U[0, 1]$ and $Y \sim U[0, 1]$ be independent. What is the density function of $X + Y$?

Reading Assignment

- Sections 2.1, 2.3, 2.4, 2.5, 3.1, 3.2, 4.1, 4.2 from *Probability and Random Processes*, G. Grimmett and D. R. Stirzaker, 2001 (3rd Edition)