Chapter 2

Logan Rosentreter

Random Variable and Distribution

Intro

Def : A R.V. Y is a function where domain is sample space and range is the real number i.e. $Y : S \to \mathbb{R}$ e.g. toss coin 3 times $S = \{(HHH), \dots, (TTT)\}$

$$Y = \#\text{ of H}$$

$$Y(HHH) = 3 \quad Y(HTH) = 2 \quad Y(HTT) = 1 \quad Y(TTT) = 0$$

Ex : covid test of 30 subjects

$$S = \{(0, 0, \dots, 0), (0, 0, \dots, 1), \dots\}$$

$$N_S = 2^{30}$$

$$Let X = \#\text{ of positive cases}$$

$$X((0, 0, \dots, 0)) = 0$$

$$X((0, 0, \dots, 1)) = 1$$

$$X : (0 : 30) = 0, 1, 2, \dots, 30$$

Def : The support of a random variable Y is the set of all possible values it can assume

Discrete R.V.

Def : A R/V/ X is called "discrete if its support is countable $X : S \to \mathbb{R}$

Def : Probability Mass Function (PMF) : pmf of a discrete R.V> X is defined as $f(x) = f_X(x) = P(X = x)$

Capital letter : name of R.V.

Lowercase letter : value of R.V.

$$f(3) = P(X = 3) = f_X(3)$$
$$Recall : Y = \# \text{ of H}$$
$$pmf : f(0) = P(Y = 0) = 1/8$$
$$f(1) = P(Y = 1) = 3/8$$
$$f(2) = P(Y = 2) = 3/8$$
$$f(3) = P(Y = 3) = 1/8$$
$$f(10) = P(Y = 10) = 0$$
$$f(x) = 0 \text{ when } x \neq 0, 1, 2, 3$$

If x is not in support of $X, f(x) = 0$

Property of pmf (Thm)

A function $f(x)$ is a valid discrete pmf iff

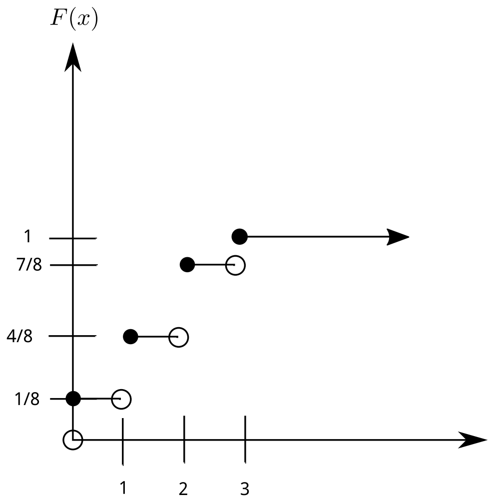
1. $0 \leq f(x) \leq 1$ for all x
2. $\sum_X f(x) = 1$

Def : Cumulative Distributive Fucntion (CDF)

$$P(X \leq x) = F(x) = F_X(x)$$

Taking pmf values from previous example

x	0	1	2	3
$f(x)$	1/8	3/8	3/8	1/8
$F(x)$	1/8	4/8	7/8	1



$$F(-0.5) = P(X \leq -0.5) = 0$$

$$F(1.5) = P(X \leq 1.5) = P(X \leq 1) = 1/2$$

Remark

- $P(X \leq x) = \sum_{k \leq x} f(k)$
- $F(x)$ is a nondecreasing funciton and always right continuous
- $0 \leq F(x) \leq 1$

Theorem 0.1

$\lim_{x \to -\infty} F(x) = 0$ $\lim_{x \to \infty} F(x) = 1$

Notate : pmf and CDF defines the distribution of a R.V.

$$X \sim f(x)$$

$$X \sim F(x)$$

where $\sim$ means "has distribution defined by"

Def : expected value

$$E(x) = \sum_X x f(x)$$

Ex :

x	0	1	2	3
$f(x)$	1/8	3/8	3/8	1/8

$$\begin{aligned} E(X) &= 0 \cdot 1/8 + 1 \cdot 3/8 + 2 \cdot 3/8 + 3 \cdot 1/8 \\ &= 1.5 \end{aligned}$$

The expected Value is a "weighted average" or the "center" or the distribution or a long run average

Notate : $\mu = E(X) = \mu_X =$ mean

Ex : Y is a R.V. w/ pmf defined

$$f(y) = \begin{cases} 1/N & 1, 2, \ldots, N \text{ where } N \geq 1 \text{ and } N \in \mathbb{Z} \\ 0 & otherwise \end{cases}$$

1. support of $Y : 1, 2, \ldots, N$

2. f(y) is a valid pmf?

$$0 \leq f(y) \leq 1 \implies 0 \leq 1/N \leq 1 \checkmark$$

$$\sum_Y 1/N = 1/N + 1/N + \ldots + 1/N = N(1/N) = 1 \checkmark$$

Find CDF

$$F(y) = P(Y \leq y) = \sum_{k \leq y} 1/N$$

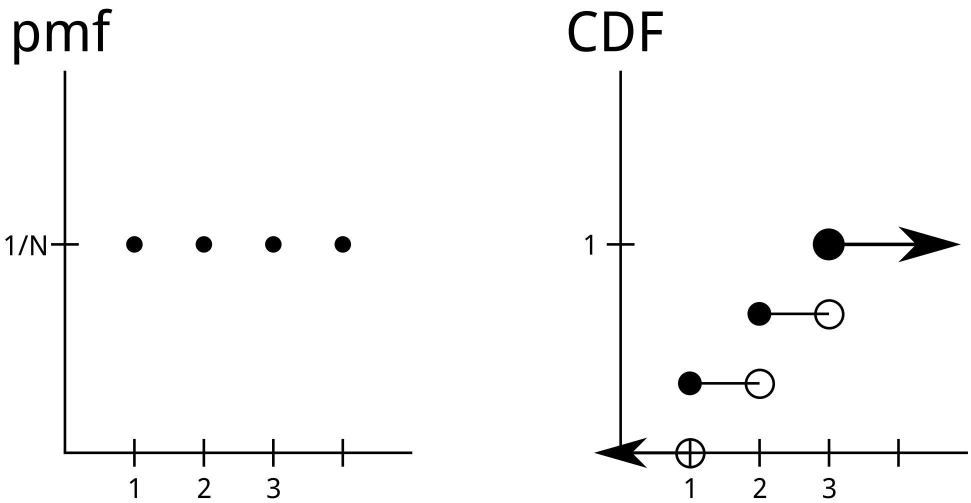
$$F(1) = 1/N$$

$$F(2) = 2/N$$

$$F(2.5) = 2/N$$

$$F(y) = \lfloor y \rfloor / N \text{ where } \lfloor y \rfloor = \lfloor ofywheny \geq 0 \rfloor$$

$$F(y) = 0 \text{ when } y < 0$$



$$E(Y) = 0f(0) + 1f(1) + \ldots + Nf(N)$$
$$= 0 \cdot 0/N + 1 \cdot 1/N + \ldots + N/N$$
$$= 1/N(1 + 2 + \ldots + N)$$
$$= 1/N(\frac{N(N+1)}{2}) = \frac{N+1}{2}$$

Continuous R.V.

Recall CDF for discrete R.V. is a step function Def : A R.V. X is called a continuous R.V. if the CDF is a continuous function There is a function $f(x)$ s.t.

$$F(x) = \int_{-\infty}^{x} f(t)dt \qquad F'(x) = f(x)$$

We don't have pmf

$$P(X = x)0 \text{ for continuous}$$

instead we have pdf (probability density function)

Remark : for continuous R.V.

$$\begin{aligned} P(a \leq X \leq b) = P(X \leq b) - P(X \leq a) = F(b) - F(a) = P(a < X < b) \\ = \int_a^b f(x)dx \end{aligned}$$