Probability Distributions of Random Variables

Let $S \rightsquigarrow$ sample space of an experiment.

A random variable $X \rightsquigarrow$ function $X : S \longrightarrow \mathbb{R}$

$\Downarrow$

the values X of X define a new sample space for the original experiment

$\Downarrow$

X (or $\mathcal{X}$) extracts specific information about an outcome in S.

Examples:

Experiment	Random variable	Sample values
① Toss a coin five times	count the number of heads	$X(HHTTH) = 3 = X(HHHTT)$ $X(HTTTT) = 1$
	proportion of heads in the outcome	$X(HHTTH) = X(HHHT) = \frac{3}{5}$ $X(HTTTT) = X(THTTT) = \frac{1}{5}$
② Roll a six sided dice 4-times	X = sum of the numbers facing up	$X(1,2,4,4) = 11$ $X(1,1,1,6) = 9$ $X(6,6,6,2) = 20$
	X = average of numbers in the sample	$X(1,2,4,4) = \frac{11}{4}$ $X(1,1,1,6) = 9/4$ $X(6,6,6,2) = 20/4$
③ Choose a box of milk cartons	X = weight of the box.	$X(\text{box } 1) = 5.1$ $X(\text{box } 2) = 5.2$ $X(\text{box } 3) = 5.0$

Note Given an experiment with sample space S, and an random variable $X : S \longrightarrow \mathbb{R}$

$\Downarrow$

two sample spaces: S and $\mathcal{X}$ = values of X

For simplicity, assume $|S|$ is finite

$\Downarrow$

$S = \{s_1, s_2, s_3, \dots, s_n\}$

If $X : S \longrightarrow \mathbb{R}$ is a r.v

$\Downarrow$

Let $\mathcal{X}$ = values of X $= \{x_1, x_2, \dots, x_k\}$

Two sample spaces for original experiment

- S $\Downarrow$ Simple outcomes are $\{s_j\}$ $j = 1,2,3,\dots,n$ $\Downarrow$ event $E \subseteq S$ described interms of $\{s_j\}$
- $\mathcal{X}$ $\Downarrow$ Simple out comes are $\{x_i\}$ $i = 1,2,\dots,k$ $\Downarrow$ $\{X = x_i\} = \{s \in S : X(s) = x_i\} \subseteq S$ $\Downarrow$ events $E \subset \mathcal{X}$ are described interms of $\{X = x_i\}$

Let P be a probabolity function for S.

$\Downarrow$

use P to "induce" a probability function P_X on $\mathcal{X}$

$\Downarrow$

Step 1) $$\boxed{P_X(\{x_i\}) := P(X = x_i) = \sum_{\omega \in \{X = x_i\}} P(\omega)}$$

outcomes in $\mathcal{X}$ — Event in S — outcome in S.

Step 2) Suppose $E \subseteq \mathcal{X}$, $E = \bigcup_{i=1}^{r} \{X = x_i\}$

$$\boxed{P_X(E) := P\left(\bigcup_{i=1}^{r} \{X = x_i\}\right) = \sum_{i=1}^{r} P(X = x_i)}$$

Note:

① The random variable will be denoted by upper case letters $X, Y, \dots$ etc

② The values of the random variable will be denoted by lower case letters: $x, y, \dots$ etc.

③ when there is no ambiguity, we will use the notation $P(X = x_i)$ instead of $P_X(X = x_i)$.

Distributions of Random Variables

Let $X \rightsquigarrow$ random variable

$\Downarrow$

P_X : probability function for X.

The cummulative distribution function of X

$\Downarrow$

$$\boxed{F_X(x) := P(X \le x) \quad \text{for all } x}$$

$F_X : \mathbb{R} \longrightarrow \mathbb{R}$

$x \longmapsto F_X(x) = P_X(X \le x)$

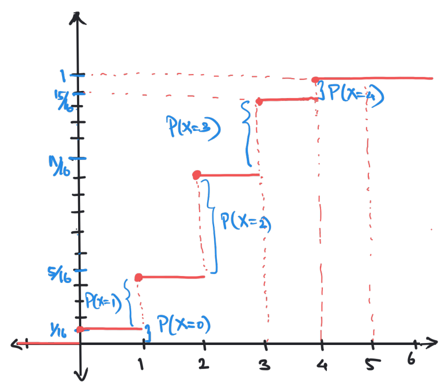
Example: Suppose we toss a fair coin 4 times. X = # heads in 4 tosses.

We can calculate

$P(X=0) = \frac{1}{16}$, $P(X=1) = \frac{4}{16}$, $P(X=2) = \frac{6}{16}$, $P(X=3) = \frac{4}{16}$, $P(X=4) = \frac{1}{16}$

We can calculate $F_X(x)$ as

$$F_X(x) = \begin{cases} 0 & \text{if } -\infty < x < 0 \\ 1/16 & \text{if } 0 \le x < 1 \\ 5/16 & \text{if } 1 \le x < 2 \\ 11/16 & \text{if } 2 \le x < 3 \\ 15/16 & \text{if } 3 \le x < 4 \\ 16/16 = 1 & \text{if } 4 \le x < \infty \end{cases}$$



Note:

① $F_X(x)$ has jumps at $x = 0, 1, 2, 3, 4$ and the jump equals $P(X = x_i)$.

② $F_X(x)$ is a non-decreasing function

③ $\lim_{x \to -\infty} F(x) = 0$ and $\lim_{x \to \infty} F(x) = 1$

④ $F(x) \ge 0$ $\forall x \in \mathbb{R}$

Note: Every cdf satisfies properties ②, ③ and ④.

Thm: A function $F(x)$ is a cdf if and only if the following three conditions are satisfied

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

b) $F(x)$ is a non-decreasing function

c) $F(x)$ is right continuous; ie $\lim_{x \to x_0^+} F(x) = F(x_0)$, $\forall x_0 \in \mathbb{R}$

$X \rightsquigarrow$ a random variable

- X is Discrete if $F_X(x)$ is a step function $\Downarrow$ The probability mass function pmf is $$\boxed{f_X(x) = P(X = x) \quad \forall x}$$
- X is Continuous if $F_X(x)$ is a continuous function $\Downarrow$ The probability density function (pdf) $f_X(x)$ satisfies $$\boxed{F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt \quad \forall x.}$$

Thm: $f_X(x)$ is a pdf (or pmf) of a random variable if and only if

a) $f(x) \ge 0$ for all x

b) $\sum_x f_X(x) = 1$ or $\int_{-\infty}^{\infty} f(x)\,dx = 1$