Random Variables

Important Note: Review Calculus

Previously we have discussed the results of random experiments in terms of outcomes and events. Although the result of a random experiment may not be a number, we are generally interested in some numerical measurement of the outcome.

A random variable (r.v.) is a function that assigns a real number to each outcome in the sample space of a random experiment. In other words, each outcome is represented by a numerical quantity.

A r.v. is discrete if it takes on values from a finite or countable set.

A r.v. is continuous if it takes on values from an uncountable set.

Ex (i) Roll a die

Let r.v. X = face value

X is a discrete r.v.

(ii) Measure temperature of sample

Let r.v. X = temperature in C°

X is a continuous r.v.

(iii) Flip a coin

Let $X=0$ if tails ; $X=1$ if heads

X is a discrete r.v.

Notation: R.V. are usually denoted by capital letters (X, Y, Z)

When discussing r.v. s probabilities are assigned to a subset of the real line.

For a discrete r.v., we assign a probability to each value taken by the r.v.

For a continuous r.v., we assign a probability to each range of values. (intervals)

Note: For a continuous r.v., the probability that the r.v. takes on a particular value is 0.

Probability Mass Function

The probability mass function (pmf) of a discrete r.v. X is defined as:

$$P_X(x_i) = \Pr(X = x_i) \quad , \text{ where } x_i \in S_X = \begin{cases} \{x_1, \dots, x_n\} \\ \{x_1, x_2, \dots\} \end{cases}$$

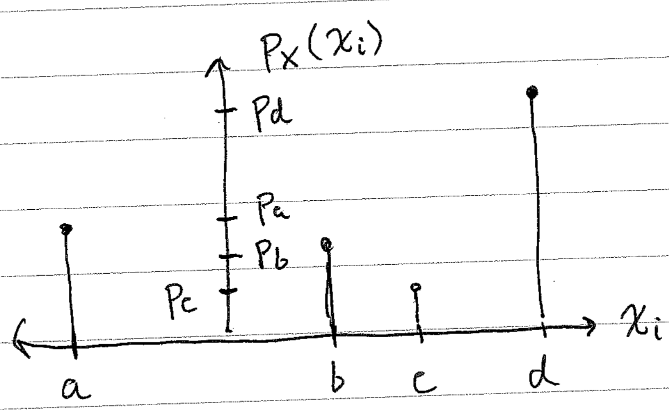
The pmf gives probabilities as its output and therefore satisfies the axioms of probability.

Properties of pmf

1) $0 \leq P_X(x_i) \leq 1$

2) $\sum_{x_i} P_X(x_i) = 1$

3) $\Pr(X \in A) = \sum_{x_i \in A} P_X(x_i)$



Ex: Let X be the number of heads in three independent coin flips with $Pr(\{H\}) = p$ on each flip. Find the pmf of X.

$$S_X = \{0, 1, 2, 3\}$$

$$p_X(0) = Pr(\{TTT\}) = (1-p)^3$$

$$p_X(1) = Pr(\{HTT\}) + Pr(\{THT\}) + Pr(\{TTH\}) = 3(1-p)^2 p$$

$$p_X(2) = 3(1-p)p^2$$

$$p_X(3) = Pr(\{HHH\}) = p^3$$

$$\sum_{x_i} p_X(x_i) = 1$$

Ex Let X be a discrete r.v. with pmf:

$$p_X(x_i) = \begin{cases} 1/3, & x_i = 0 \\ 1/6, & x_i = 1 \\ 1/4, & x_i = 2, 3 \end{cases}$$

Evaluate the following probabilities:

a) $Pr(X \geq 2)$

X takes on values $x_i = 0, 1, 2, 3$ with $Pr(X = x_i) = p_X(x_i)$

$$Pr(X \in A) = \sum_{x_i \in A} p_X(x_i)$$

$Pr(X \geq 2) = Pr(X \in \{2,3\})$

$= \sum_{x_i \in \{2,3\}} P_X(x_i)$

$= P_X(2) + P_X(3)$

$= 1/4 + 1/4 = 1/2$

b) $Pr(X < 3) = Pr(X \in \{1,2,0\})$

$= P_X(1) + P_X(2) + P_X(0)$

$= 1/4 + 1/6 + 1/3 = 3/4$

c) $Pr(1 < X \leq 3) = Pr(X \in \{2,3\})$

$= P_X(2) + P_X(3)$

$= 1/4 + 1/4 = 1/2$

d) $Pr(X \leq 2 \text{ or } X = 3) = Pr(X \leq 2) + Pr(X = 3)$

$= Pr(X < 3) + Pr(X = 3)$

$= 3/4 + 1/4$

$= 1$

Cumulative Distribution Function

The cumulative distribution function (cdf) of a r.v. X is defined as:

$$F_X(x) = \Pr(X \le x) , \quad -\infty < x < \infty$$

Properties of cdf:

1) $0 \le F_X(x) \le 1$

2) $F_X(-\infty) = 0 , \quad F_X(+\infty) = 1$

$$\lim_{x \to -\infty} F_X(x) = 0 , \quad \lim_{x \to \infty} F_X(x) = 1$$

3) $F_X(x_1) \le F_X(x_2)$ for $x_1 \le x_2$

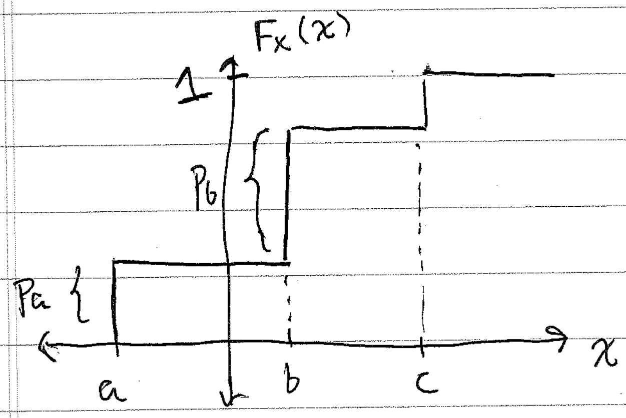
4) $\Pr(X > x) = 1 - F_X(x)$

5) $\Pr(x_1 < X \le x_2) = F_X(x_2) - F_X(x_1)$ for $x_1 \le x_2$

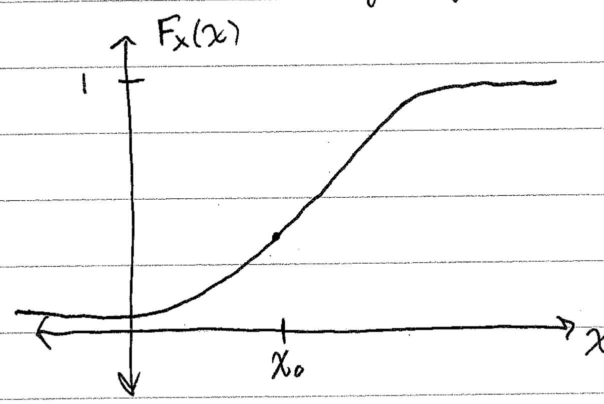
Can define discrete and continuous r.v.s in terms of cdf:

A <u>discrete</u> r.v. X has a piecewise constant cdf ($F_X(x)$ only has jumps)

A <u>continuous</u> r.v. X has a continuous cdf (has no jumps)



Discrete



Continuous

At any point $x_0 \in \mathbb{R}$ we have that

$$\Pr(X = x_0) = \lim_{\varepsilon \to 0^+} F_X(x_0 + \varepsilon) - \lim_{\cancel{x}\,\varepsilon \to 0^+} F_X(x_0 - \varepsilon)$$

$$= \lim_{\varepsilon \to 0^+} \Pr(x_0 - \varepsilon < X \leq x_0 + \varepsilon)$$

$$= \text{height of jump in } F_X(x) \text{ at } x = x_0$$

For continuous r.v. $Pr(X = x_0) = 0$

" discrete r.v. $Pr(X = a) = p_a$, $Pr(X = b) = p_b \ldots$

<u>Note</u>: The cdf of a discrete r.v. can be written as

$$F_X(x) = \sum_{x_i} Pr(X = x_i)\, u(x - x_i)$$

$$= \sum_{x_i} p_X(x_i)\, u(x - x_i)$$

where $p_X(x_i)$ is the pmf of X and $u(x)$ is the unit step function