Random Variable

Defn: A random variable is a function from the sample space S to the set $\mathbb{R}$ of real numbers.

Example: Suppose $S = \{\text{rain}, \text{snow}, \text{clear}\}$.

Then, the function $X : S \longrightarrow \mathbb{R}$ defined by $X(\text{rain}) = 3$, $X(\text{snow}) = 6$ and $X(\text{clear}) = -2.7$ is a random variable.

Example: Suppose $S = \{5, 6, 7, 8, \ldots\}$, with $P(s) = \frac{1}{2^s}$, for $s \in S$. Then what is $P(\{5, 6, 7, 8, \ldots\})$.

Let $A_5 = \{5\}$, $A_6 = \{5, 6\}$, $\ldots$

$A_n = \{5, 6, \ldots, 7\}$.

Then $\{A_n\} \uparrow A$ and

$$P(A) = \lim_{n \to \infty} \left(\frac{1}{2^5} + \frac{1}{2^6} + \cdots + \frac{1}{2^n} \right)$$

$$= \lim_{n \to \infty} \left(\frac{2^{-5} - 2^{-n-1}}{1 - 2^{-1}} \right)$$

$$= \lim_{n\to\infty} \left(2^{-4} - 2^{-n}\right) = 2^{-4} = 1/16.$$

Example, For the case

$S = \{\text{rain, snow, clear}\}$, we can define another random variable $Y: S \to \mathbb{R}$ by

$$Y(\{\text{rain}\}) = 0$$

$$Y(\{\text{snow}\}) = -\frac{1}{2}$$

$$Y(\{\text{clear}\}) = 7/8.$$

Exam:

Suppose we flip three coins. Let X be the total number of heads showing. Clearly X is a random variable.

Let Y be the square of the number showing on the second die. Y is a random variable.

Let Z be the sum of the two numbers showing, and let W be the

Square of the sum of the two nos. showing. Clearly Z, W are random variables.

Ex: Constant as Random variables.

Every constant c is a random variable, by saying that $X(s) = c$, for all $c \in S$.

Ex: Indicator functions.

If A is any event, then the indicator function of A,

$$I_A(s) = \begin{cases} 1 & s \in A \\ 0 & s \notin A \end{cases}$$ is again a random variable.

Ex: Given random variables X and Y.

(i) $Z = X^2$ is a random variable.

(ii) $W = XY^3$ is a random variable.

(iii) $Z = X + Y$ is a random variable.

Ex: Consider rolling a fair six-sided die. Then $S = \{1, 2, 3, 4, 5, 6\}$.

Let X be the number showing and Y be three more than the number showing.

Then $Z = X^2 + Y$ is a random variable

$$Z(s) = (X(s))^2 + Y(s), \quad s \in S.$$

Defn:

Given random variables X, Y.

(a) We write $X = Y$ to mean $X(s) = Y(s)$ for all $s \in S$

(b) We write $X \leq Y$ to mean $X(s) \leq Y(s)$ for all $s \in S$

(c) We write $X \geq Y$ to mean

$X(s) \geq Y(s)$ for all $s \in S$.

Example: Let $S = \{1, 2, 3, 4, 5, 6\}$

Consider random variables,

$X(s) = s, \quad s \in S.$ and

Also, $P(X=17)=0$. Further,

$$P(X=x) = P(\phi) = 0 \quad \text{for all} \quad x \notin \{3, 6, -2.7\}$$

$$P(X \in \{3,6\}) = P(X=3) + P(X=6)$$

$$= 0.4 + 0.15 = 0.55$$

$$P(X \le 5) = P(X=3) + P(X=-2.7)$$

$$= 0.4 + 0.45 = 0.85.$$

Defn: If X is a random variable, then the distribution of X is the collection of probabilities $P(X \in B)$ for all subsets B of the real numbers.

Note: Strictly speaking B should be a Borel subset. Any subset that we could ever write down is a Borel subset.

Example: A very simple distribution:

Consider our above example,

$$P(X \in B) = 0.4\, I_B(3) + 0.15\, I_B(6) + 0.45\, I_B(-2.7).$$

Example:

Consider $S = \{\text{rain}, \text{snow}, \text{clear}\}$.

Define
$$Y(\text{rain}) = 5$$
$$Y(\text{snow}) = 7$$
$$Y(\text{clear}) = 5.$$

What is the distribution of Y?

$$P(Y = 7) = P(\text{snow}) = 0.15$$
$$P(Y = 5) = P(\{\text{rain}, \text{clear}\}) = 0.4 + 0.45 = 0.85.$$

Therefore, if B is any subset of the real numbers, then

$$P(Y \in B) = 0.15\, I_B(7) + 0.85\, I_B(5).$$

Discrete Distributions:

Defn: A random variable X is discrete if there is a finite or countable sequence $x_1, x_2, \ldots$ of distinct real numbers, and a corresponding sequence $p_1, p_2, \ldots$ of nonnegative real numbers, such that $P(X = x_i) = p_i$ for all i, and $\sum_i p_i = 1$.

Defn:

For a discrete random variable X, its probability function is the function $p_X : \mathbb{R} \longrightarrow [0,1]$ defined by

$$p_X(x) = P(X = x).$$

If $x_1, x_2, \ldots$ are the distinct values such that $P(X = x_i) = p_i$ for all i with $\sum_i p_i = 1$.

Clearly, $P_X(x) = \begin{cases} p_i & x = x_i \text{ for some } i. \\ 0 & \text{otherwise} \end{cases}$

<u>Theorem</u>

(Law of total probability, discrete random variable version). Let X be a discrete random variable, and let A be some event. Then

$$P(A) = \sum_{x \in \mathbb{R}} P(X = x)\, P(A \mid X = x).$$

<u>Important discrete distributions:</u>

Example: Constant random variable:

For some $c \in \mathbb{R}$, $X \equiv c$ is a random variable.

Its probability function is

$$P_c(x) = \begin{cases} 1 & \text{if } x = c \\ 0 & \text{if } x \neq c. \end{cases}$$

Bernoulli random variable:

A random variable X is said to be a Bernoulli random variable with parameter p if X takes the values 0 and 1, and its probability mass function is given by

$$p(0) = P(X=0) = 1-p$$
$$p(1) = P(X=1) = p, \quad \text{where } 0 \leq p \leq 1.$$

Ex: Suppose that an experiment whose outcome is either success or failure

Then $S = \{\text{success, failure}\}$.

Define a function $X : S \longrightarrow \mathbb{R}$ by

$X(\text{success}) = 1, \quad X(\text{failure}) = 0.$

Then X is a Bernoulli random variable.

Why.

$$p_X(1) = P(\text{success}) = p$$

$$p_X(0) = P(\text{failure}) = 1-p, \text{ for some } 0 \leq p \leq 1$$

Binomial random variable:

A random variable X is said to be Binomial random variable with parameters (n, p) if X takes the values $0, 1, 2, \ldots, n$ and its probability mass function is given by

$$p_X(i) = \binom{n}{i} p^i (1-p)^{n-i}, \quad 0 \leq i \leq n.$$

Example: Consider flipping n fair coins. Let X be the total number of heads showing. Then, X is a binomial random variable with parameters $(n, \frac{1}{2})$.

Why?

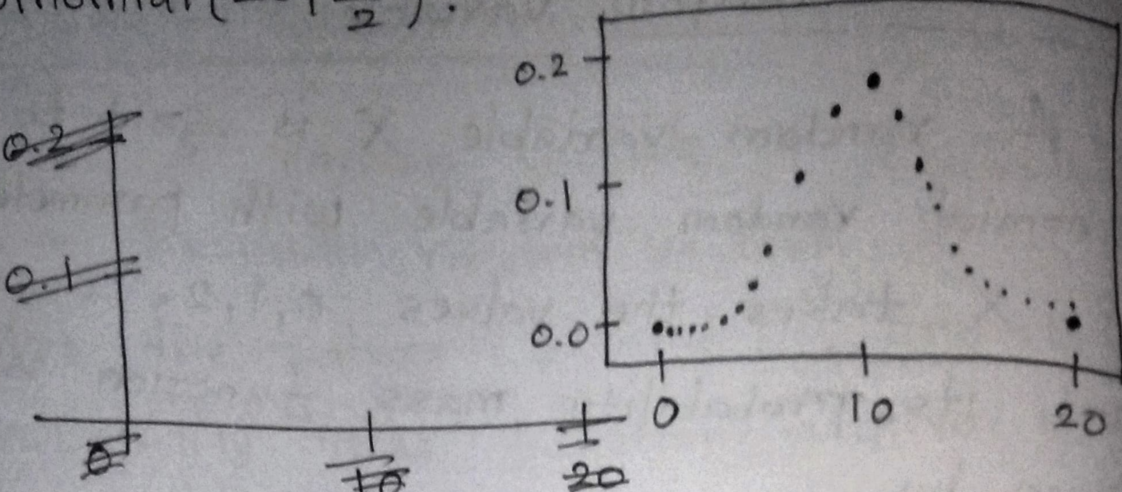
X takes the values $0, 1, 2, \ldots, n$

and

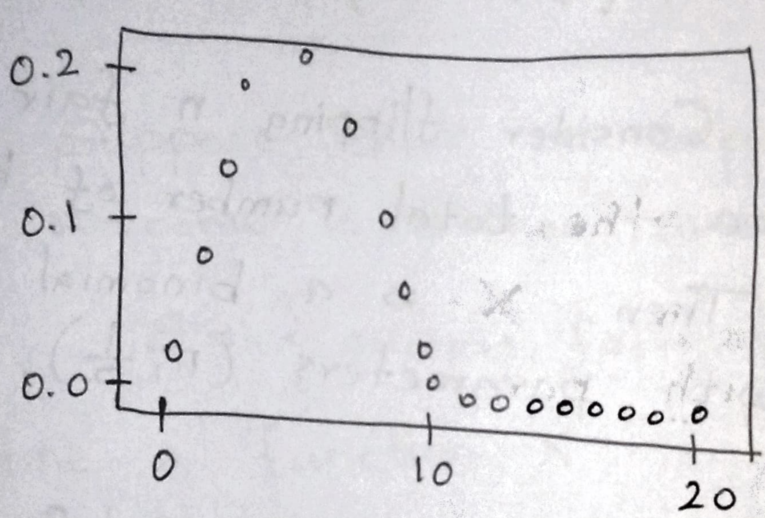
$$p_X(i) = \binom{n}{i} p^i (1-p)^i$$

We write this as $X \sim \text{Binomial}(n, p)$.

Binomial$(20, \frac{1}{2})$.



Binomial$(20, 1/5)$



Ans: If $X_1, X_2, \ldots, X_n$ has the Bernoulli(θ) distribution, then the function $Y = X_1 + X_2 + \cdots + X_n$ has the binomial distribution, Binomial(n, θ).

Example:

The poisson distribution:

A random variable X is said to be a Poisson random variable with parameter λ if takes one of the values 0, 1, 2, 3, ... and its probability distributive function is given by

$$P_X(i) = \frac{\lambda^i}{i!} e^{-\lambda}, \quad i = 0, 1, 2, \ldots$$

The Poisson random variable has a tremendous range of applications in diverse areas.

Consider a binomial random variable with parameters (n, p), with p is very small. Even further, n is larger and larger, Binomial (n, p) is an approximation to a Poisson distribution.

Let $\lambda = np$.

$$P(X = i) = \binom{n}{i} p^i (1-p)^{n-i}$$

$$= \frac{n!}{i!(n-i)!} p^i (1-p)^{n-i}.$$

$$= \frac{n(n-1)\cdots(n-i+1)}{i!}\left(\frac{\lambda}{n}\right)^{i}\left(1-\frac{\lambda}{n}\right)^{n-i}$$

$$= \frac{n(n-1)\cdots(n-i+1)}{n^{i}}\;\frac{\lambda^{i}}{i!}\;\frac{(1-\lambda/n)^{n}}{(1-\lambda/n)^{i}}$$

Since

$\frac{n(n-1)\cdots(n-i+1)}{n^{i}}$ Converges to 1 as $n \to \infty$

$(1-\lambda/n)^{i}$ cgs 1 as $\lim_{n\to\infty}\left(1+\frac{c}{n}\right)^{n} = e^{c}$

$(1-\lambda/n)^{n}$ cgs $e^{-\lambda}$ when n is large and λ moderate.

Thus $\cancel{P(X=i) = e^{-\lambda}\frac{\lambda^{i}}{i!}}$

$$\lim_{n\to\infty} P(X=i) = e^{-\lambda}\frac{\lambda^{i}}{i!}.$$

Intuitively, if we flip a very large number of coins, and each coin has a very small probability p of coming up heads, then the probability that the total number of heads will be i is approximately $\frac{e^{-\lambda} \lambda^i}{i!}$.

Task:

Plot Binomial $(100, 1/10)$ and Poisson(10) for the values $0, 1, \ldots 20$.

Other examples are

(1) The number of misprints on a page of a book.

(2) The number of people in a community who survive to age 100.

(3) The number of wrong telephone nos. that are dialed in a day

(4) The number of packages of dog biscuits sold in a particular ~~day~~ store each day.

(5) The number of customers entering a post office on a given day.

(6) The number of vacancies occuring during a year in the Indian Judicial system.

Example:

Suppose that the probability that an item produced by a certain machine will be defective is 0.1. Find the probability that a sample of 10 items will contain at most 1 defective item.

Let X be the number of defective items. Then X is a binomial random variable on S with parameter (10, 0.1).

The required probability is

$$P(X \leq 1) = P(X=0) + P(X=1)$$

$$= \binom{10}{0}(0.1)^0(0.9)^{10} + \binom{10}{1}(0.1)^1(0.9)^9$$

$$= 0.7361.$$

Ex: Suppose that the number of typographical errors on a single page in a book has a poisson distribution with parameter $\lambda = \frac{1}{2}$. Calculate the probability

Cumulative distribution functions

Defn: Given a random variable X, its cumulative distribution function (or distribution function, or cdf for short) is the function

$F_X : \mathbb{R} \longrightarrow [0,1]$, defined by

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

Suppose that $B = (a, b]$, a left-open interval.

$$\begin{aligned} P(X \in B) &= P(a < X \leq b) \\ &= P(X \leq b) - P(X \leq a) \\ &= F_X(b) - F_X(a) \end{aligned}$$

Properties of Distribution Functions:

Theorem: Let F_X be the cumulative distribution function of a random variable X. Then

(a) $0 \leq F_X(x) \leq 1$ for all x.

(b) $F_X(x) \leq F_X(y)$ whenever $x \leq y$.

(c) $\lim_{x \to +\infty} F_X(x) = 1$.

(d) $\lim_{x \to -\infty} F_X(x) = 0$.

CDFs of Discrete distributions:

Thm: Let X be a discrete random variable with probability function P_X. Then its cumulative distribution function F_X satisfies $F_X(x) = \sum_{y \leq x} P_X(y)$.

Proof:

Let $x_1, x_2, \ldots$ be the possible values of X. Then

$$F_X(x) = P(X \leq x) = \sum_{x_i \leq x} P(X = x_i)$$

$$= \sum_{y \le x} P(X = y) = \sum_{y \le x} p_X(y)$$

Example:

Consider rolling one fair six-sided die, so that $S = \{1, 2, 3, 4, 5, 6\}$, with $P(s) = 1/6$.

Let X be the number showing on the die divided by 6, so that $X(s) = s/6$, for $s \in S$.

Then

$$F_X(x) = \begin{cases} 0 & x < 1/6 \\ 1/6 & 1/6 \le x < 2/6 \\ 2/6 & 2/6 \le x < 3/6 \\ 3/6 & 3/6 \le x < 4/6 \\ 4/6 & 4/6 \le x < 5/6 \\ 5/6 & 5/6 \le x < 1 \\ 1 & 1 \le x \end{cases}$$