X Lecture 2, Sept. 15, 04

Random Variables

A random variable X is a function that assigns a real number $X(\omega)$ to each outcome of a random experiment $\omega \in S$.

S is the domain and $S_X = \{X(\omega) : \omega \in S\}$ is the range of the random variable X.

Example:

A binary communication system with

$S = \{0, 1\}$ and $X(\omega) \in \{-A, +A\}$

where $0 \to -A$ and $1 \to +A$, i.e.,

$S_X = \{-A, A\}$

CDF (Cumulative Distribution Function)

$F_X(x) = P[X \leq x] \qquad -\infty < x < \infty$

Properties of $F_X(x)$

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

2) $\lim_{x \to \infty} F_X(x) = 1$

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

4) $F_X(x)$ is a nondecreasing function of x, i.e.,

if $a < b$ then $F_X(a) \leq F_X(b)$

5) $F_X(x)$ is continuous from right, i.e.,

for $h > 0$ $\quad F_X(b) = \lim_{h \to 0} F_X(b+h) = F_X(b^+)$.

6)

$$P(a < X \leq b) = F_X(b) - F_X(a) \qquad \text{for } b \geq a$$

Proof:

$$\{X \leq a\} \cup \{a < X \leq b\} = \{X \leq b\}$$

and $\quad \{X \leq a\} \cap \{a < X \leq b\} = \phi$

So

$$P(X \leq a) + P(a < X \leq b) = P(X \leq b)$$

$$F_X(a) + P(a < X \leq b) = F_X(b)$$

$$\boxed{P(a < X \leq b) = F_X(b) - F_X(a)} \qquad \text{QED.}$$

7) $\quad P(X = b) = F_X(b) - F_X(b^-)$

Example: Consider throwing a dice and let

$$S = \{1, 2, \cdots 6\} = S_X$$

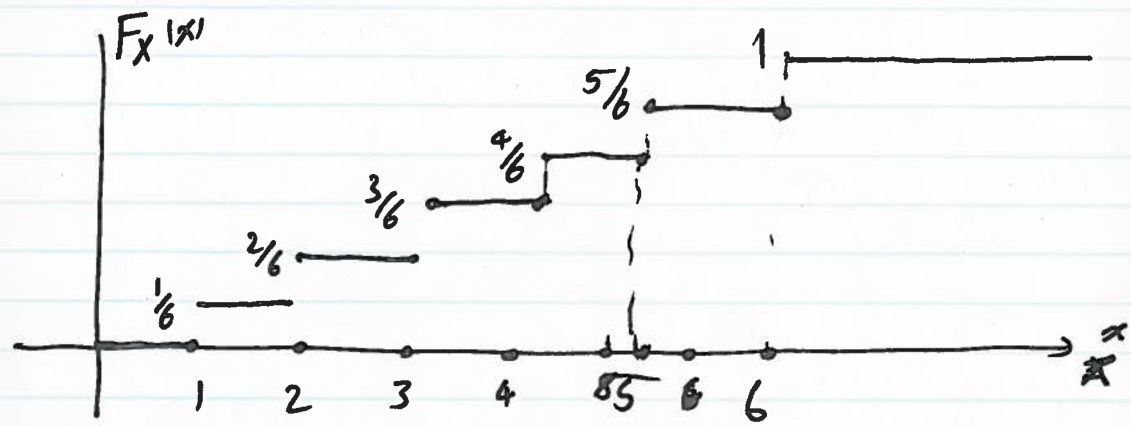
$$\cancel{F_X(x)} \quad F_X(x) = 0 \qquad x < 1$$

$$F_X(x) = \frac{1}{6} \qquad 1 \leq x < 2$$

$$F_X(x) = \frac{2}{6} \qquad 2 \leq x < 3$$

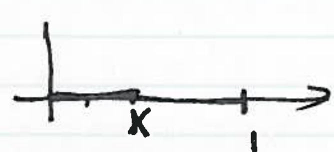
$$\vdots \qquad\qquad \vdots$$

$$F_X(x) = 1 \qquad x \geq 6$$

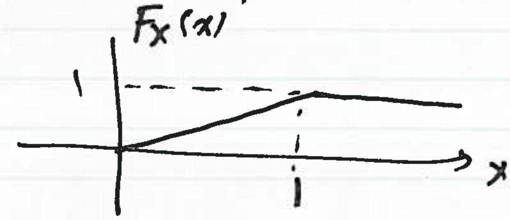


Example:

Picking a number between 0 and 1 uniformly (without prejudice).



$$F_X(x) = P(X \leq x) = \frac{x}{1} = x$$



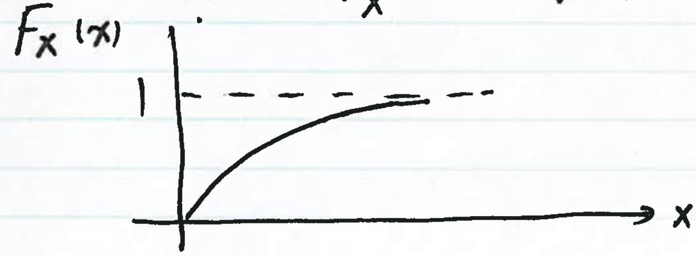
Example:

~~The transmission time of a message is distributed~~

The lifetime of a device, say, a light bulb, is distributed exponentially, i.e.,

$$P[X > t] = e^{-\lambda t}$$

So $F_X(x) = P(X \le x) = 1 - e^{-\lambda x}$



Discrete, Continuous and Mixed Random Variables

Discrete random variables take values from a finite or at most countably infinite set.

e.g., $\{1, 2, \cdots 6\}$, $\{0, 1\}$, $\{1, 2, \cdots, \cdots\}$

i.e.,

$$S_X = \{x_0, x_1, \cdots\}$$

The CDF of a discrete random variable is given as

$$F_X(x) = \sum_k P_X(x_k)\, u(x - x_k)$$

where $P_X(x_k)$ is the probability mass function of X given as

$$P_X(x_k) = P(X = x_k)$$

as an example for $S_X = \{1, 2, \cdots, 6\}$ with

$$P_X(x_k) = \frac{1}{6} \quad \forall k \in \{1, \cdots, 6\}$$

we have

$$F_X(x) = \frac{1}{6}u(x-1) + \frac{1}{6}u(x-2) + \cdots + \frac{1}{6}u(x-6)$$

A <u>Continuous Random Variable</u> has a CDF that is continuous everywhere.

A mixed R.V. has a cdf that is continuous over certain values but jumps at certain points. The CDF for a mixed R.V. can be written as

$$F_X(x) = pF_1(x) + (1-p)F_2(x)$$

where $0 < p < 1$ and $F_1(x)$, and $F_2(x)$ are CDF's that are discrete, and continuous, respectively.

The <u>probability density function</u>:

pdf is defined as

$$f_X(x) = \frac{d F_X(x)}{dx}$$

by definition

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\, dt$$

$$P[\, x < X \le x+h \,] = P[\, X < x+h \,] - P[\, X \le x \,]$$

$$= \cancel{F_X(x+h) - F_X(x)} = F_X(x+h) - F_X(x)$$

$$= \int_{-\infty}^{x+h} f_X(t)\, dt - \int_{-\infty}^{x} f_X(t)\, dt$$

$$= \int_{x}^{x+h} f_X(t)\, dx = f_X(x)\, h \quad \text{for small } h$$

So

$$P[\, x \le X \le x+h \,] = f_X(x)\, h$$

Thus $f_X(x)$ represents the "<u>density</u>" of the probability at point $X = x$.

Properties of pdf:

1) $f_X(x) \geq 0$

2) $P[a \leq X \leq b] = \int_a^b f_X(x)\,dx$

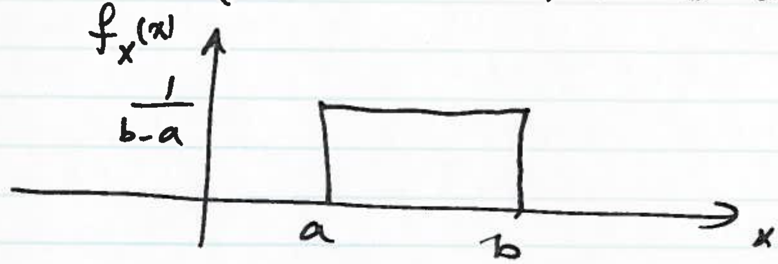
3) $F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt$

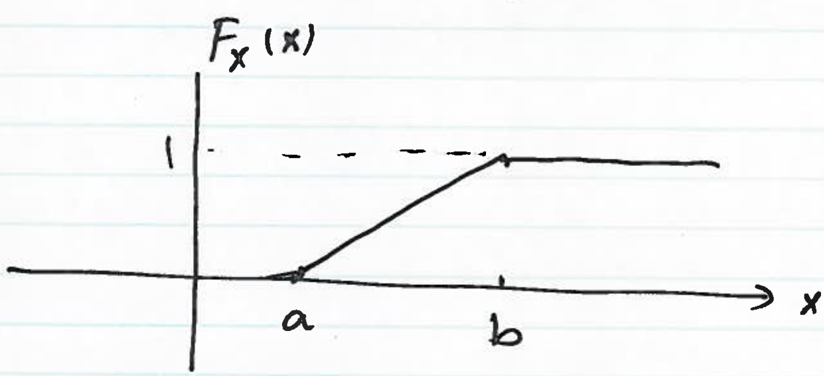
4) $\int_{-\infty}^{\infty} f_X(x)\,dx = 1$

- Examples of some random variables and their pdf

1) Uniform random variable

$$f_X(x) = \begin{cases} \dfrac{1}{b-a} & a \leq x \leq b \\ 0 & x < a \text{ and } x > b \end{cases}$$

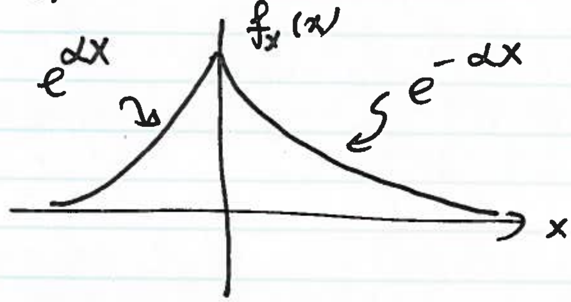




$$F_X(x) = \begin{cases} 0 & x < a \\ \dfrac{x-a}{b-a} & a \leq x \leq b \\ 1 & x > b \end{cases}$$

The pdf of the speech samples ~~are~~ is given by:

$$f_X(x) = c\, e^{-\alpha|x|} \qquad -\infty < x < \infty$$



a) Find the constant c.

b) Find $P[|X| < v]$

$$\int_{-\infty}^{\infty} f_X(x) = 2\int_0^{\infty} c e^{-\alpha x}\, dx = \frac{2c}{\alpha} = 1$$

$$c = \frac{\alpha}{2}$$

$$P[|X| < v] = \frac{\alpha}{2}\int_{-v}^{v} e^{-\alpha|x|}\, dx = \frac{\cancel{2}\alpha}{\cancel{2}} \int_0^v e^{-\alpha x}\, dx$$

$$= 1 - e^{-\alpha v}$$

For a discrete random variable, the CDF is not continuous at some points $\Rightarrow$ the derivative does not exist. To overcome this problem, we define the <u>delta function</u> $\delta(t)$

which is the derivative of $u(x)$

$$u(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases}$$

$$u(x) = \int_{-\infty}^{\infty} \delta(t)\, dt$$

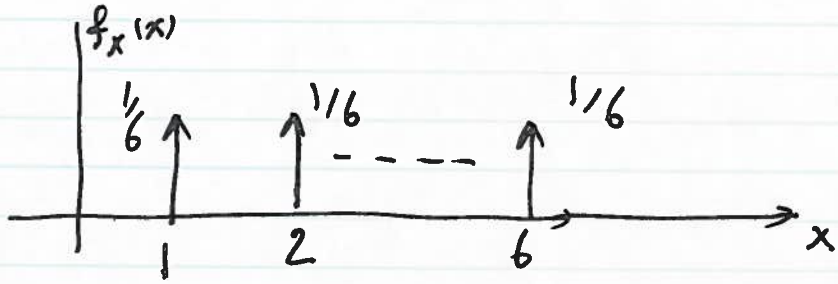
For a discrete r.v., we have:

$$F_X(x) = \sum_k P_X(x_k)\, u(x - x_k)$$

So

$$f_X(x) = \sum_k P_X(x_k)\, \delta(x - x_k)$$

e.g.



example: X = # of heads in three coin tosses.

$P_X(0) = \frac{1}{8}$ TTT $P_X(1) = \frac{3}{8}$, $P_X(2) = \frac{3}{8}$, $P_X(3) =$

then:

$$f_X(x) = \frac{1}{8}\delta(x) + \frac{3}{8}\delta(x-1) + \frac{3}{8}\delta(x-2) + \frac{1}{8}\delta(x-3)$$

Since:

$$F_X(x) = \frac{1}{3}u(x) + \frac{3}{8}u(x-1) + \frac{3}{8}u(x-2) + \frac{1}{8}u(x-3)$$

$$P[1 < X \leq 2] = \int_{1^+}^{2^+} f_X(x)\, dx = \frac{3}{8}$$

Conditional CDF and pdf

$$F_X(x|A) = P[X \le x|A] = \frac{P[\{X \le x\} \cap A]}{P[A]} \quad \text{if } P[A] > 0$$

$$f_X(x) = \frac{d}{dx} F_X(x|A)$$

example:

The life time of a device (a machine, a bulb) has CDF $F_X(x)$. Find the condition CDF and pdf given the even $A = \{X > t\}$, i.e., given that the device has servived until $X = t$ assess ~~find~~ its future survivability.

$$F_X(x|X>t) = P[X \le x | X > t]$$

$$= \frac{P[\{X \le x\} \cap \{X > t\}]}{P[\{X > t\}]}$$

a) if $x < t \Rightarrow \{X \le x\} \cap \{X > t\} = \Phi \Rightarrow F_X(x|X>t) = 0$

b) if $x \ge t \Rightarrow \{X \le x\} \cap \{X > t\} = \{t < X \le x\}$

$$\Rightarrow F_X(x) = \frac{F_X(x) - F_X(t)}{1 - F_X(t)}$$

So:

$$F_X(x|X>t) = \begin{cases} 0 & x \le t \\ \dfrac{F_X(x) - F_X(t)}{1 - F_X(t)} & x > t \end{cases}$$

and

$$f_X(x \mid X \geq t) = \frac{f_X(x)}{1 - F_X(t)} \qquad x \geq t$$

Well-known Random variables

1) Bernouli R.V.

$S_X = \{0, 1\}$

$P_X(0) = 1 - p$ and $P_X(1) = p$

it can model any binary experiment, e.g., tossing of a coin, value of a random bit, state of a device (working/broken).

let S be a sample space and $A \subset S$ be an event. Define

$$I_A(\omega) = \begin{cases} 0 & \text{if } \omega \notin A \\ 1 & \text{if } \omega \in A \end{cases}$$

I_A is called the indicator of event A.

I_A is ~~at~~ a random variable, since it assigns a number to each outcome of S.

$P(A) = P(\omega \in A) = P(I = 1) = p$

~~P~~ So $p_I(1) = p$ $\quad p_I(0) = 1 - p$

~~many~~

2) Binomial Distribution:

Assume we repeat a random experiment n times and Let X be the number of times the event A occurs.

So,

$$S_X = \{0, 1, 2, \cdots n\}$$

Let I_j denote the indicator function of the event A in j-th trial $0 \leq j \leq n$. Then:

$$X = I_1 + I_2 + \cdots + I_n$$

is ~~a~~ the sum of Bernouli trials.

$$P[X=k] = \binom{n}{k} p^k (1-p)^{n-k} \qquad k = 0, 1, \cdots, n.$$

X is called ~~the~~ the Binomial random variable.

3) Geometrical R.V.

example:

In an ARQ system, one sends a packet, waits for an Ack and resends if a NAK is received or in the case of time-out. Let the probability of Succesful transmission (no error in the packet) be p. Then the probability of ~~that the message is sent~~ number ~~of attempts~~ of transmission be k is

$$P[X=k] = (1-p)^{k-1} p \qquad k = 1, 2, \cdots$$

This is called a Geometric R.V.

~~Lecture~~ 3: Sept. 22, 2004

4) Poisson Random Variable

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

This models, e.g., the number of phone calls in a certain time interval or the number of packets needed to be transmitted in certain interval. α is the average number of occurences of that event in that time interval.

The Poisson probabilities can be derived from Binomial Distribution if we assume $\alpha = np$ and keep α fixed while letting $n \to \infty$, i.e.

$$p_k = \binom{n}{k} p^k (1-p)^{n-k} \longrightarrow \frac{\alpha^k}{k!} e^{-\alpha} \quad k = 0, 1, \ldots$$