Lecture 9

Random Variables: A random variable is a real-valued function on the space of all events Ω.

eg: ① Sum of two rolls of a fair die

Sum → Function

two rolls → generate random numbers when we repeat the experiment.

② Outcome of a coin toss

If Heads, then $X=1$

If Tails, then $X=0$

Random variable X takes values: 0 or 1.

* Random Variables can take values in a discrete set or can take continuous values as well.

Discrete-type Random Variables

If X is a discrete-type random variable, then X takes values in a discrete set.

eg: ① Coin toss, $X \in \{0,1\}$

② Roll of a fair die, $X \in \{1,2,3,4,5,6\}$

③ Number of trials it may take for a coin to show heads $X \in \{1,2,3,4,\ldots\ldots\}$

* Probability Mass Function (pmf):

↳ defined for a discrete-type random variable

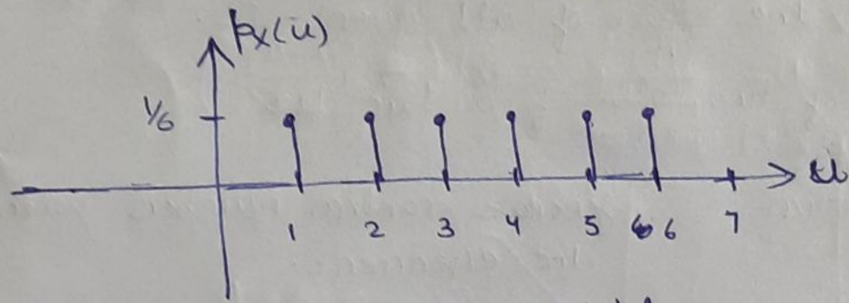
Let X be a discrete Random variable taking values: $u_1, u_2, \ldots, u_n$

A probability mass function (pmf) assigns to each $u \in \mathbb{R}$, the probability that $X=u$.

$P(X=u) \rightarrow$ Prob. that $X=u$

pmf for roll of a fair die

$P(X=1) = P(X=2) = \cdots = P(X=6) = \frac{1}{6}$ $\qquad p_X(u) = P(X=u)$



Mean and Variance of a random variable:

↳ Mean (or Average value) of a random variable

denoted by $\mathbb{E}[X]$, μ_X or $\overline{X}$

$$\boxed{\mathbb{E}[X] = \sum_i u_i P(X=u_i)}$$

↳ Variance of a random variable:

$$\begin{aligned} Var(X) &= \mathbb{E}[(X-\mu_X)^2] \\ &= \mathbb{E}[X^2 + \mu_X^2 - 2X\mu_X] \\ &= \mathbb{E}[X^2] + \mu_X^2 - 2\mu_X \underbrace{\mathbb{E}[X]}_{\mu_X} \end{aligned}$$

$$\boxed{Var(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2}$$

Ex: Suppose X is a random variable taking values in $\{-2,-1,0,1,2,3,4,5\}$ each with prob. 1/8.
Let $Y = X^2$. Find $\mathbb{E}[Y]$?

Y	P(Y=u)
0	1/8
1	1/4
4	1/4
9	1/8
16	1/8
25	1/8

$$\begin{aligned} \mathbb{E}[Y] &= 0\cdot\tfrac{1}{8} + 1\cdot\tfrac{1}{4} + 4\cdot\tfrac{1}{4} + 9\cdot\tfrac{1}{8} + 16\cdot\tfrac{1}{8} + 25\cdot\tfrac{1}{8} \\ &= \frac{2+8+9+16+25}{8} \\ &= 7.5 \end{aligned}$$

<u>Ex.</u> Let X be a random variable and consider the new random variable $Y = X^2 + 3X$.

$$\mathbb{E}[Y] = \mathbb{E}[X^2+3X] = \mathbb{E}[X^2] + 3\,\mathbb{E}[X].$$

Transformations of Random Variables:

- $\mathbb{E}[X+b] = \mathbb{E}[X] + b$
- $$\begin{aligned} Var(X+b) &= \mathbb{E}[(X+b-\mathbb{E}[X+b])^2] \\ &= \mathbb{E}[(X+b-\mathbb{E}[X]-b)^2] \\ &= \mathbb{E}[(X-\mathbb{E}[X])^2] = Var(X). \end{aligned}$$

- $\mathbb{E}[aX] = a\,\mathbb{E}[X]$
- $$\begin{aligned} Var(aX) &= \mathbb{E}[(aX-\mathbb{E}[aX])^2] \\ &= a^2\,\mathbb{E}[(X-\mathbb{E}[X])^2] = a^2\,Var(X) \end{aligned}$$

$\Rightarrow$ $\mathbb{E}[aX+b] = a\,\mathbb{E}[X] + b$

$Var(aX+b) = a^2\,Var(X)$

Data Standardization:

Let X be a random variable with mean μ_X and variance σ_X^2.

$$\Rightarrow \quad \mathbb{E}\left[\left(\frac{X-\mu_X}{\sigma_X}\right)\right] = \frac{1}{\sigma_X}(\mathbb{E}[X]-\mu_X) = 0$$

and, $$Var\left(\frac{X-\mu_X}{\sigma_X}\right) = \frac{1}{\sigma_X^2}\,Var(X) = 1$$

Thus $Y = \dfrac{X-\mu_X}{\sigma_X}$ is a random variable with mean 0 and variance 1.

- Independent Random Variables:

Defn Random Variables X and Y are independent if

$$P(X=i, Y=j) = P(X=i) \cdot P(Y=j)$$

- Common distributions of discrete random variables:

① A random variable X is said to have the Bernoulli distribution with parameter p, where $0 \leq p \leq 1$ if

$$P(X=1) = p \text{ and } P(X=0) = 1-p.$$

$$\boxed{X \sim Ber(p)}$$

$$E[X] = p$$

$$Var(X) = E[X^2] - E[X]^2 = p - p^2 = p(1-p)$$

② Binomial distribution: $X \sim Binomial(n, p)$

Suppose n independent Bernoulli trials are connected, each resulting in a 1 with prob. p and a 0 with prob $1-p$. Let X denote the total number of ones occurring in the n-trials. Any particular outcome with k ones and $(n-k)$ zeros has probability $p^k(1-p)^{n-k}$. There are ^nC_k such outcomes.

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

$$E[X] = np$$

$$Var(X) = np(1-p)$$

Ex: Suppose two teams A and B play a best-of-seven series of games. Assume that ties are not possible in each game, the team wins a game with prob. 1/2 and games are independent. The series ends once one of the teams has won four games. Let Y denote the total number of games played. Find pmf of Y.

A: Clearly Y is defined for $4 \le k \le 7$, i.e. $P(Y=k)=0$ for $k=1,2,3$.

For $\underline{k \ge 4}$:

$$P(Y=k) = P(Y=k \text{ and A wins series}) + P(Y=k \text{ and B wins series})$$
$$= 2P(Y=k \text{ and A wins series})$$

wins thrice in $(k-1)$ games and win the k^{th} game.

$$\Rightarrow P(Y=k \text{ and A wins series}) = \left[\binom{k-1}{3}\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^{k-1-3}\right]\frac{1}{2}$$

$$= \frac{1}{2}\left[\binom{k-1}{3}\left(\frac{1}{2}\right)^{k-1}\right]$$

$$\Rightarrow \boxed{P(Y=k) = \binom{k-1}{3}\left(\frac{1}{2}\right)^{k-1}}$$

③ Geometric distribution: $X \sim \text{Geometric}(p)$

X denotes the number of trials conducted until the outcome of the trial is 1.

$$P(X=k) = (1-p)^{k-1}p, \text{ for } k \ge 1$$

Clearly, $\sum_{k=1}^{\infty}(1-p)^{k-1}p = p + p(1-p) + p(1-p)^2 + \cdots$

$$= p \cdot \frac{1}{1-(1-p)} = 1$$

Can show that $\mathbb{E}[X] = \frac{1}{p}$

$$\text{Var}(X) = \frac{1-p}{p^2}$$

④ Poisson distribution: → used to model queues.

$X \sim \text{Poisson}(\lambda)$ — Let arrival rate is λ, what is the prob. that queue length is k.

$$P(X=k) = \frac{e^{-\lambda}\lambda^k}{k!}, \quad k \geq 0$$

$$E[X] = \lambda$$
$$Var(X) = \lambda$$

—— x —— x —— x —— x ——

Continuous-type random variables.

"From peanuts to peanut butter"

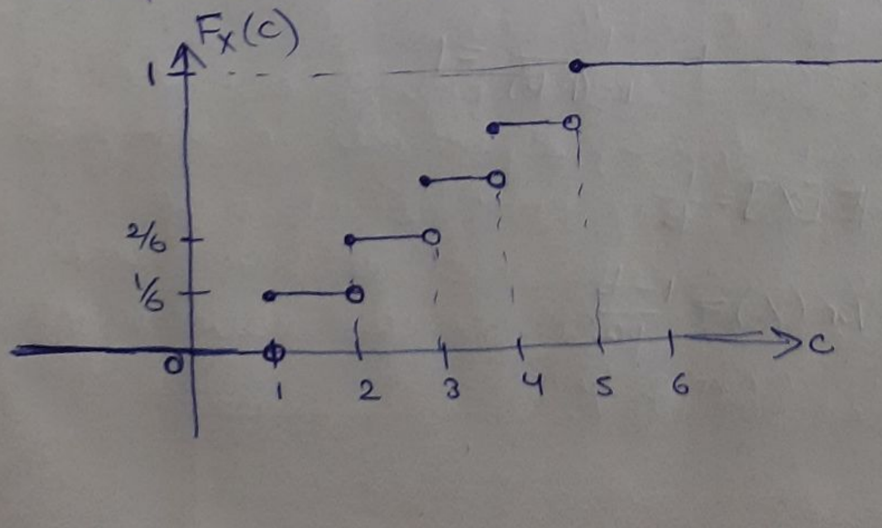
- peanuts ↓ Discrete random variable
- peanut butter: Mass of peanut butter — Continuous random variable

Cumulative ~~density~~ distribution function of a random variable
(CDF) — random variable: discrete or continuous

↓ $P(X \leq c) = F_X(c)$

Why CDF? Doesn't make sense to talk about $P(X=c)$ for continuous-type random variable. There are infinitely many choices. Prob. of picking exactly one specific number is zero.

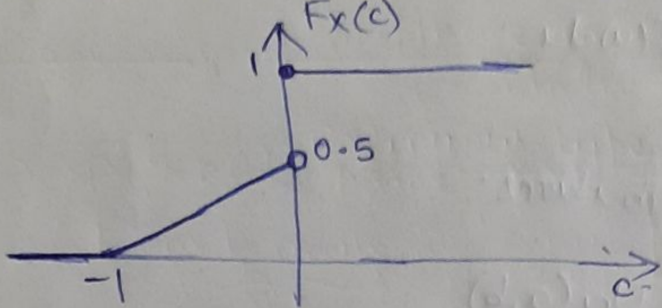
CDF for die roll experiment.



Properties of CDF: (F)

1. F is non-decreasing.
2. $\lim_{c \to +\infty} F(c) = 1$ and $\lim_{c \to -\infty} F(c) = 0$
3. F is right-continuous.

Ex: Let X have the CDF shown is fig.



1. Determine all values of u s.t. $P(X=u) > 0$.
2. Find $P(X \leq 0)$.
3. Find $P(X<0)$.

A: 1. $P(X=u) > 0$ only for $u=0$. At other places, probability of choosing a point is 0 (because of continuity).

2. $P(X \leq 0) = 1$.
3. $P(X<0) =$ 0.5.

Continuous Random Variable: A random variable X is a continuous-type random variable if the CDF is the integral of a function,

$$F_X(c) = \int_{-\infty}^{c} f_X(u)\, du$$

$f_X(u)$ is defined as the probability density function (pdf)

Note: $f_X(u)$ is not the probability of choosing u.

- $P(a \le X \le b) = F_X(b) - F_X(a) = \int_a^b f_X(u)\,du$

- Mean of continuous-type random variable,

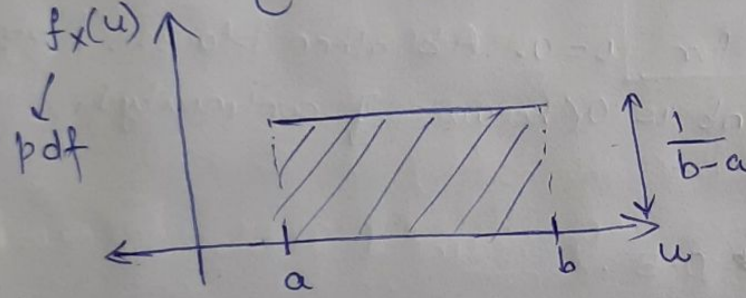
$$\mathbb{E}[X] = \int_{-\infty}^{\infty} u\, f_X(u)\,du$$

$$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(u)\, f(u)\,du$$

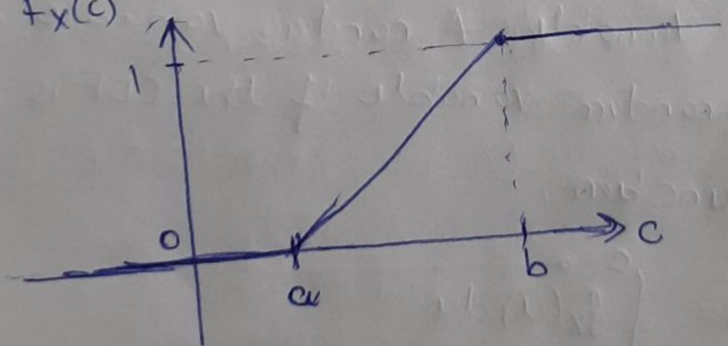
* Common continuous-distributions:

① Uniform distribution: $X \sim Uni(a,b)$

$$f_X(u) = \begin{cases} \frac{1}{b-a}, & a \le u \le b \\ 0, & \text{else} \end{cases}$$



CDF:

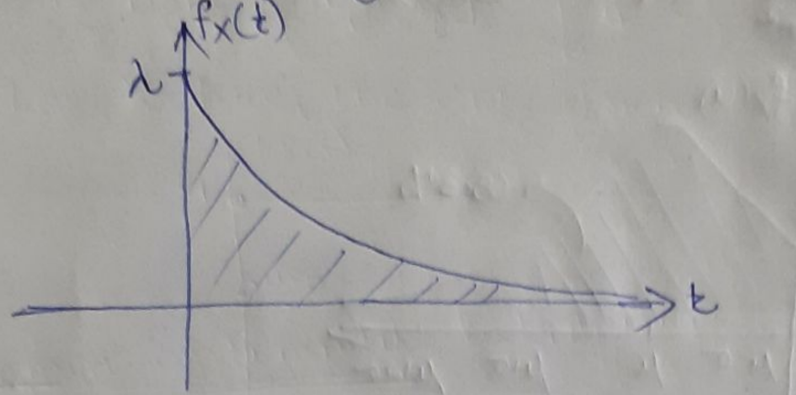


$$\mathbb{E}[X] = \frac{a+b}{2}$$

$$Var(X) = \frac{(a-b)^2}{12}$$

② Exponential distribution: $X \sim \exp(\lambda)$

pdf, $f_X(t) = \begin{cases} \lambda e^{-\lambda t} & t \geq 0 \\ 0 & \text{else} \end{cases}$



CDF: $F_X(t) = \int_{-\infty}^{t} f_X(s)\,ds = 1 - e^{-\lambda t}$

$$F_X(t) = \begin{cases} 1 - e^{-\lambda t} & , t \geq 0 \\ 0 & , t < 0 \end{cases}$$

$\mathbb{E}[X] = \frac{1}{\lambda}$, $Var(X) = \frac{1}{\lambda^2}$

Ex. Let X be exponentially distributed random variable with parameter $\lambda = \ln 2$ ① Find the simplest possible expression for $P(X \geq t)$ as a function of $t \geq 0$.

$$P(X \geq t) = 1 - \underbrace{P(X < t)}_{F_X(t)}$$

$$= e^{-\lambda t} = e^{-(\ln 2) \cdot t} = 2^{-t}$$

② Find $P(X \leq 1 \mid X \leq 2)$.

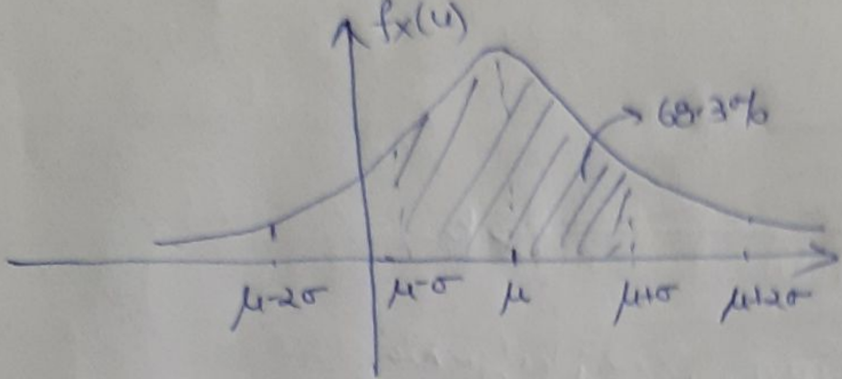
$$P(X \leq 1 \mid X \leq 2) = \frac{P(X \leq 1, X \leq 2)}{P(X \leq 2)} = \frac{P(X \leq 1)}{P(X \leq 2)}$$

$$= \frac{1 - 2^{-1}}{1 - 2^{-2}} = \frac{2}{3}$$

③ The Gaussian (Normal) distribution:

$X \sim N(\mu, \sigma^2)$

$$f(u) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$



CDF of X: $\Phi(c) = \int_{-\infty}^{c} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$

$X \sim N(0,1)$

Complementary CDF: $Q(c) = 1 - \Phi(c) = \underbrace{\Phi(-c)}_{\text{due to symmetry}}$

Ex: Let $X \sim N(10, 16)$. Find the numerical values of the following probabilities:

$P(X \geq 15)$, $P(X \leq 5)$, $P(X^2 \geq 400)$ and $P(X = 2)$.

A: Clearly $P(X=2) = 0$. (Prob. of choosing a point is 0).

$X \sim N(10, 16)$ define $Y = \frac{X-10}{4}$, then $Y \sim N(0,1)$

$\Rightarrow P(X \geq 15) = P\left(\frac{X-10}{4} \geq \frac{15-10}{4}\right) = P(Y \geq 1.25) = \Phi(-1.25)$

And, $P(X \leq 5) = P\left(\frac{X-10}{4} \leq \frac{5-10}{4}\right) = P(Y \leq -1.25) = \Phi(-1.25)$

And, $P(X^2 \geq 400) = P(X \geq 20) + P(X \leq -20)$

$= P\left(\frac{X-10}{4} \geq \frac{10}{4}\right) + P\left(\frac{X-10}{4} \leq -\frac{30}{4}\right)$

$= \Phi(-2.5) + \Phi(-7.5)$ —— Ans.