PROBABILITY THEORY CLASS 3

- Agenda for today's class
  - Take home Q ✓
  - Random variables
  - PDF / PMF / CDF + Expected Value
  - Binomial distribution
  - Geometric distribution*]
  - Poisson dist.
  - Uniform dist.
  - Exponential dist.*]
  - Normal dist.

- Random Variables

  Formally, a RV is neither random, nor a variable. It is a function mapping events $\to \mathbb{R}$

$$\{ X : \omega \to \mathbb{R} , \ \forall \omega \in \Omega$$

Intuitively → Randomness: Experiment

Intuitively → Variable: storing events in a variable x

Ex 1: $\Omega = \{H, T\}$

$$X = \begin{cases} 1, & \text{if } H \\ 0, & \text{if } T \end{cases}$$

$x = 1$ is an event "H" occurs

Ex 2: 3 coin tosses, $X = \{\text{No: of } H\}$

$1/8$	$3/8$	$3/8$	$1/8$
$(X = 0)$	$(X = 1)$	$(X = 2)$	$X = 3$
TTT	TTH	THH	HHH
	THT	HTH	
	HTT	HHT	

$$P(2H) = P(X = 2) = \frac{3}{8}$$ Transition from events → X

PDF / PMF

Functions that summarize RVs

- → $\underbrace{\text{PDF}}$ / PMF
- → CDF
- → MGF
- → CGF

- Probability <u>Mass</u> Function (PMF)

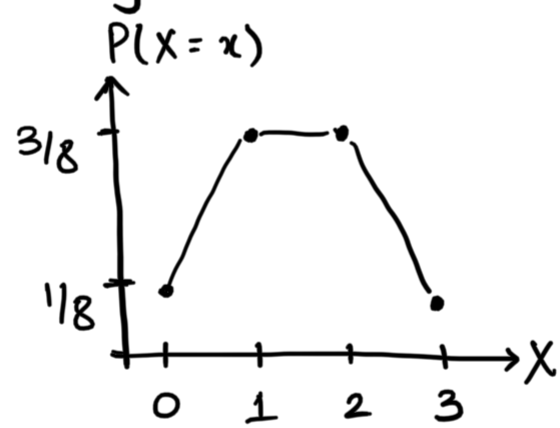
  Used for discrete values

  Say $\{a_1, a_2, \ldots, a_n\}$ are n-discrete points of an RV X, then

$$P(\underbrace{X = a_j}) = \underbrace{P_j} > 0$$

$$\sum_{j=1}^{n} P_j = 1 \quad \text{and} \quad \boxed{P_j \leq 1}$$

<u>Ex</u>: Tossing 3 coins, X = No: of H



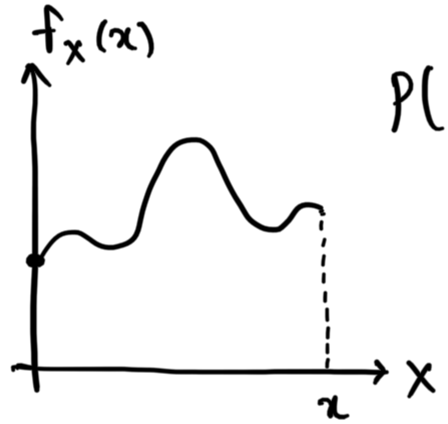
- Probability Density Function (PDF)

Used for continuous distributions.

$\underbrace{f_X(x)}$ is a PDF for an RV X iff

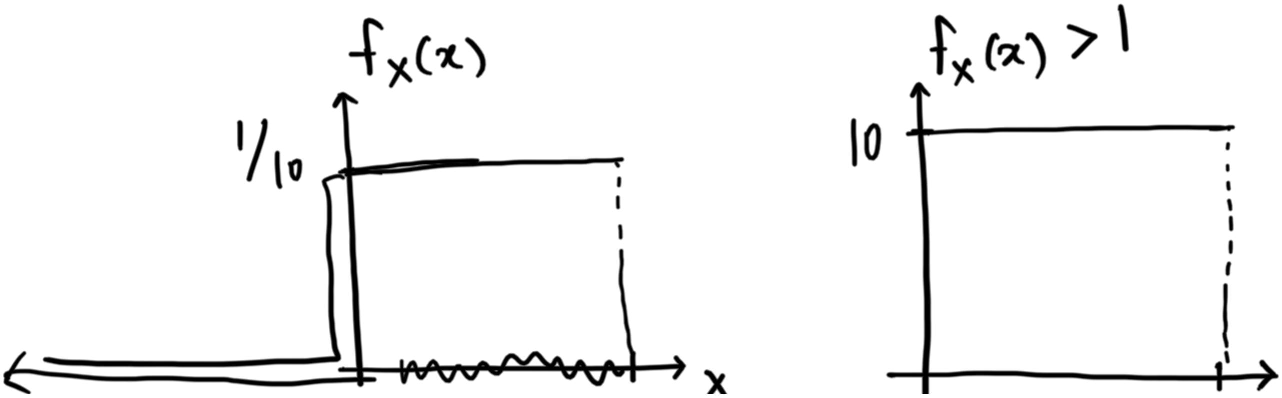
$$\begin{cases} \int_{-\infty}^{\infty} f_X(x)\,dx = 1, \quad \underbrace{f_X(x) \geq 0} \\ P(a < X < b) = \int_a^b f_X(x)\,dx \end{cases}$$

Ex 1:



$P(X = x) \neq \overbrace{f_X(x)}$

Ex 2: (WARNING!!)



$-\infty$ 10 $1/10$

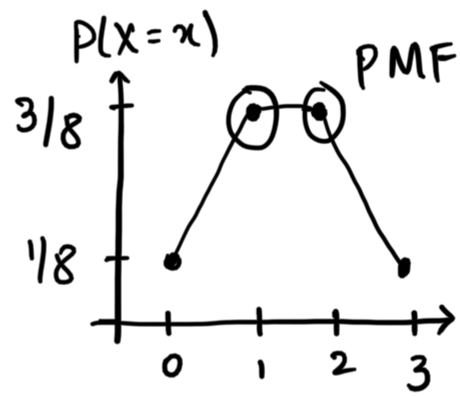
- Cumulative Distribution Function (CDF)

Instead of $\underbrace{X = x}$, $\underbrace{X \le x}$ can be of interest.

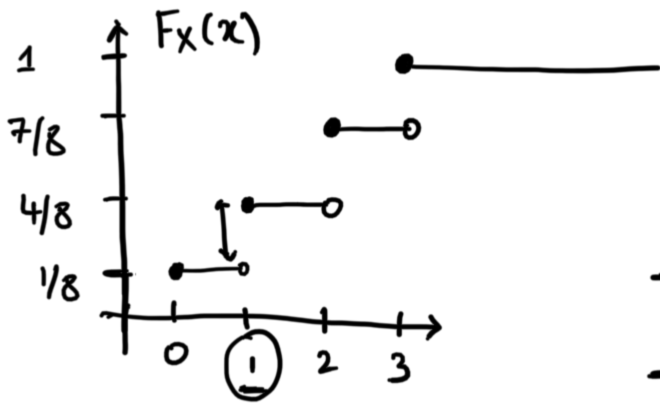
[Discrete] : $F_X(x) = P(X \le x)$.

[Cont.] : $F_X(x) = \int_{-\infty}^{x} f_X(x)\,dx$

Ex: 3 Tosses, X = No: of H



$$F_X(x) = \begin{cases} 1/8, & x \le 0 \\ 4/8, & x \le 1 \\ 7/8, & x \le 2 \\ 1, & x \le 3 \end{cases}$$



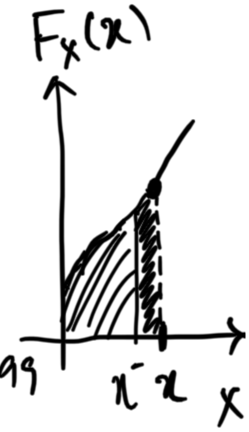
- Non-decreasing
- Right-continuous
- $F(-\infty) = 0$

  $F(\infty) = 1$

Properties of CDF function:

(a) $$\frac{\partial}{\partial x} F_X(x) = f_X(x)$$

(b) $$F_X(x) = \int_{-\infty}^{x} f_X(x)\, dx$$



Continuous distributions

(c) $$\underbrace{P(X = x)}_{\text{cont. dist.}} = \underbrace{\overset{0.5}{F(x)}} - \underbrace{\overset{0.4999}{F(x^-)}}$$

(d) $$P(x < X \leq y) = F(y) - F(x)$$

(e) $$\underbrace{P(X > x)} = \underbrace{1 - \overbrace{P(X \leq x)}} = 1 - F_X(x)$$

- Averages / Expectation

Say $X = \{2, 2, 2, 3, 3, 4\}$

$$E[X] = \frac{2+2+2+3+3+4}{6} = \frac{16}{6}$$

$$E[X] = \underbrace{\frac{3}{6}} \times 2 + \underbrace{\frac{2}{6}} \times 3 + \underbrace{\frac{1}{6}} \times 4$$

x

$$P(X = x)$$

[Discrete]: $E[X] = \sum_{j=1}^{n} P(X = x_i) \cdot x_i \}$

[Cont^n]: $E[X] = \int_{-\infty}^{\infty} x \cdot f_X(x)\, dx$

Properties of expectation:

(a) $E[X+Y] = E[X] + E[Y] \Big\}$ indep. / dep. X, Y

(b) $E[c \cdot X] = c \cdot E[X]$

(c) $\sum E[X] = E[\sum X]$

- Independence of RVs

Joint Dist.

$$\overline{P(\underbrace{X = x}_{A}, \underbrace{Y = y}_{B})} = P(X = x) \cdot P(Y = y)$$

$$f_{XY}(x, y) = f_X(x) \cdot f_Y(y) \}$$

Similarly $P(X \le x, Y \le y) = P(X \le x) \cdot P(Y \le y)$

$$\left\{ F_{XY}(x,y) = F_X(x) \cdot F_Y(y) \right.$$

- Bernoulli Distribution

$$\underbrace{X \sim \text{Bern}(p)}, \quad \underline{X} = \begin{cases} 1, & \text{with prob. } \underline{p} \\ 0, & \underline{1-p} \end{cases}$$

Only 2 outcomes $\{0, 1\}$

$$E[X] = 0 \cdot (1-p) + 1 \cdot p = p$$

$$\underbrace{E[X] = p} = \mathbb{1}_{\underbrace{\{x=1\}}} \Bigg\} \text{Indicator Function}$$

- Binomial Distribution

$X \sim \text{Binom}(n, p)$,

n – independent identical bernoulli trials

- $\exists$ 2 outcomes
- p is same for all trials

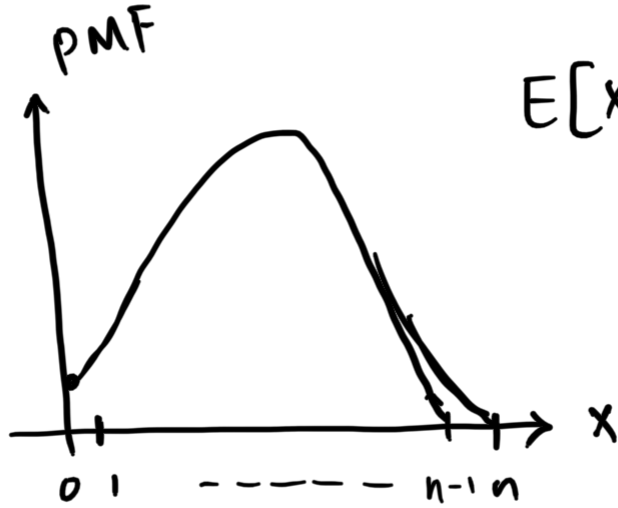
- n is fixed ✓
- iid }

$n=10$, $k=4$

$$\underbrace{P(X=k)} = \underbrace{{}^{n}C_{k}} \cdot \underbrace{\boxed{p^{k}}} \cdot \underbrace{(1-p)}^{\,n-k}$$

$$\sum_{k=0}^{n} P(X=k) = 1$$

$$\sum_{k=0}^{n} {}^{n}C_{k}\, \underline{p}^{k}\, (1-\underline{p})^{n-k} = (p+1-p)^{n} = 1$$



$$E[X] = n \cdot p$$

$n=5$, $k=2$, $p=0.8$

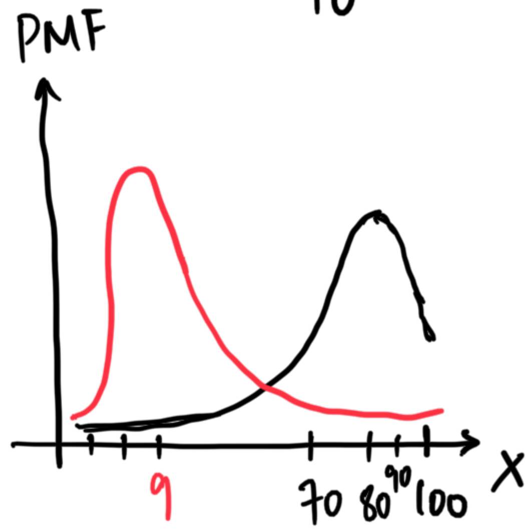
$$P(X=2) = {}^{5}C_{2} \cdot (0.8)^{2} \cdot (0.2)^{3}$$

$$= \frac{\overset{2}{\cancel{4}}\times 5}{\cancel{5!}_{0}} \cdot \frac{64}{\ldots} \times \frac{8}{1000}$$

$$\underbrace{\cancel{3\%} : \cancel{2\%}}_{10} \qquad 100 \qquad 1000$$

$n = 100$

$p = 0.09$



- Poisson Dist.

"Discrete"

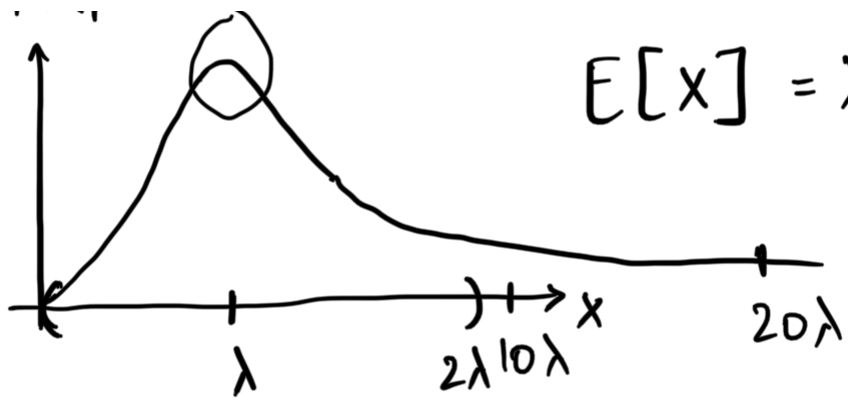
Describes no: of events in a fixed time

Ex: → No: of clicks
→ No: of buses

$$P_X(x) = \underbrace{e^{-\lambda}} \cdot \frac{\boxed{\lambda^x}}{x!}, \quad x = 0, 1, \ldots$$

$X \sim Pois(\lambda)$ → intensity const.

PMF



$E[X] = \lambda$ (const.)

"Memoryless"

$$P(\overline{X > t+s} \mid \overline{X > t}) = P(X > s)$$

$t = 5,\ s = 5$

$$X \sim \text{Pois}(\lambda),\quad P(X = x) = \frac{e^{-\lambda}\lambda^{x}}{x!}$$

$$X \sim \text{Binom}(n, p),\quad P(X = k) = {}^{n}C_{k}\, p^{k} (1-p)^{n-k}$$

$$\lambda = np,\quad p = \lambda/n$$

$$= \frac{\overset{n(n-1)\cdots(n-k+1)}{\cancel{n!}}}{k!\,\cancel{(n-k)!}} \cdot \left(\frac{\lambda}{n}\right)^{k} \cdot \left(1 - \frac{\lambda}{n}\right)^{n-k}$$

$$\lambda^{k} \quad n(n-1)(n-2)\cdots(n-k+1) \left(1-\lambda\right)^{n} \left(1-\lambda\right)^{-k}$$

$$= \frac{\lambda}{k!} \cdot \underbrace{\frac{n(n-1)(n-2)\cdots}{n^k}}_{\checkmark} \cdot \left(\quad \frac{}{n} \right) \cdot \underbrace{\left(\quad \frac{}{n} \right)}_{\checkmark}$$

$$\lim_{n\to\infty} \frac{n(n-1)\cdots(n-k+1)}{n^k} = 1$$

$$\lim_{n\to\infty} \Big(1 - \underbrace{\frac{\lambda}{n}}_{\to 0}\Big)^{-k} = 1$$

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

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

$n \to \infty$ or $p \to 0$ $\quad \text{Binom}(n,p) \to \text{Pois}(\lambda)$

$\lambda = np$

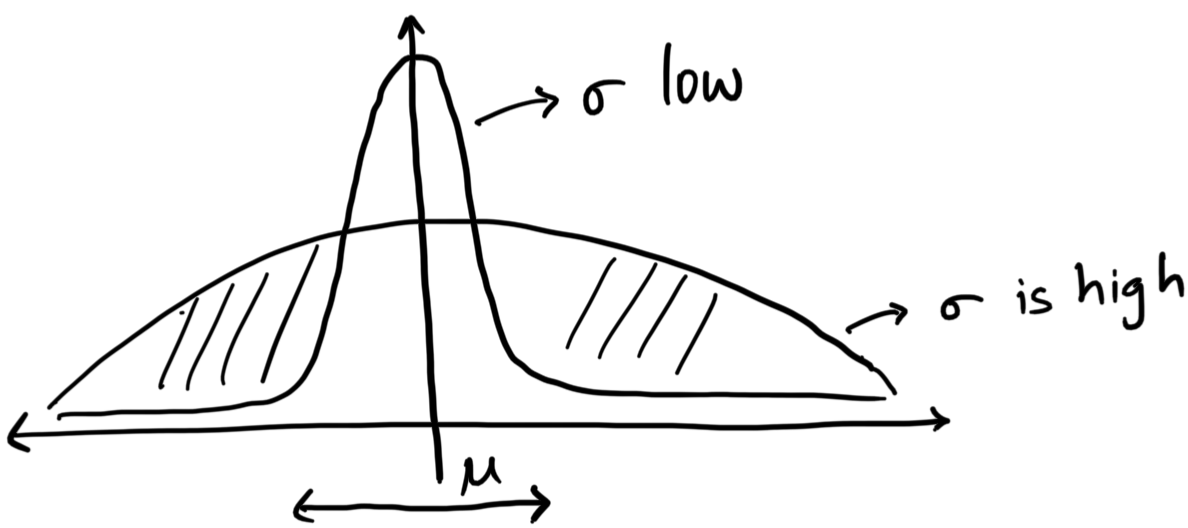
$$P(X=k) = e^{-\lambda} \cdot \underbrace{\frac{\lambda^k}{k!}}$$

- Normal Dist:

$$f_X(x) = \underbrace{\frac{1}{\sigma\sqrt{2\pi}}}_{\text{Normalizing}} \cdot \underbrace{\exp\left(-\frac{(\overbrace{x-\mu}^{z})^2}{2\sigma^2}\right)}_{\underline{\text{Kernel}}}$$

Kernel is big when x is close to μ //

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



Rate at which $P \to 0$ is at $e^{-(x-\mu)^2}$