

CS70 – Spring 2024

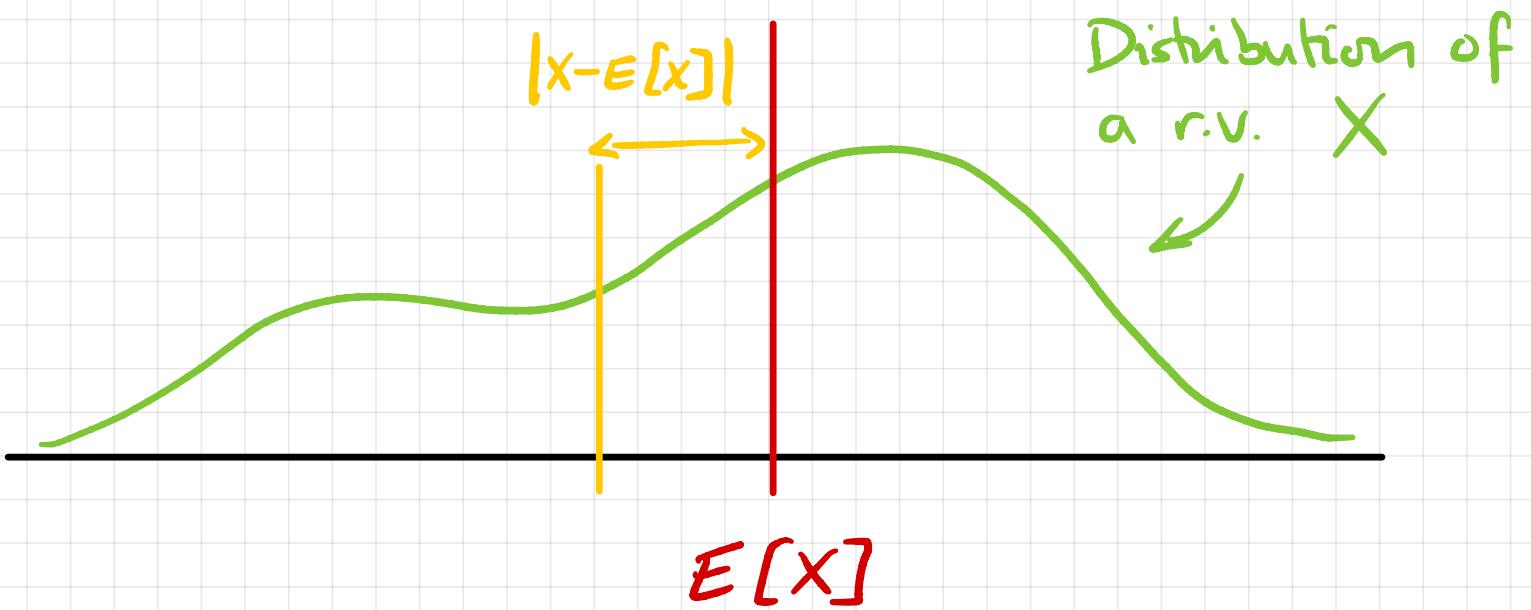
Lecture 20 – April 2

# Plan for Today

- Variance (& standard deviation) – a measure of “spread” of a random variable
- How to compute Variance
- Variance of Binomial, Geometric, Poisson
- Covariance & Correlation

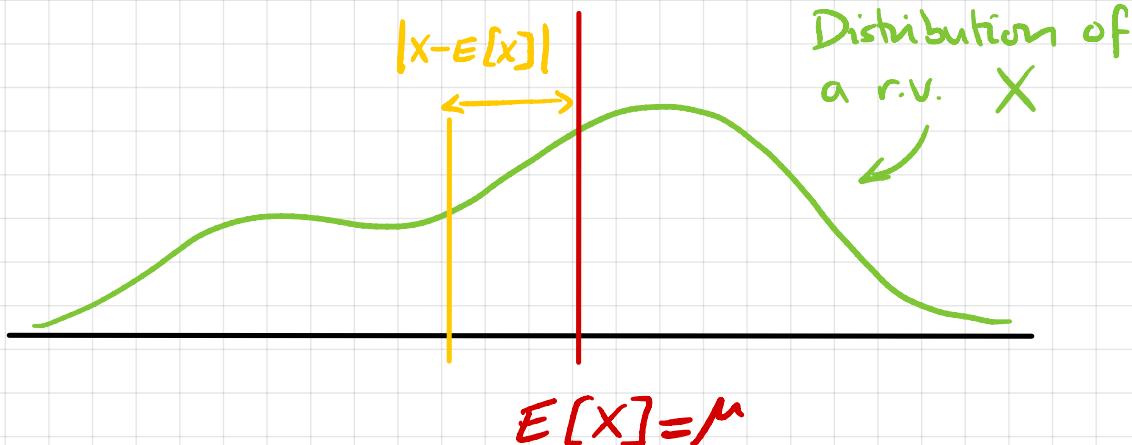
# Variance

Measures the typical distance of a r.v. from its expectation



Obvious measure of distance is  $|X - E[X]|$

More convenient to use  $(X - E[X])^2$



Definition : The variance of a random variable  $X$  is

$$\text{Var}(X) = E[(X - \mu)^2]$$

where  $\mu = E[X]$ .

The standard deviation of  $X$  is

$$\sigma(X) = \sqrt{\text{Var}(X)}$$

## A more convenient expression

Claim :  $\text{Var}(X) = E[(X-\mu)^2] = \boxed{E[X^2] - \mu^2}$

Proof :

$$\begin{aligned} \text{Var}(X) &= E[(X-\mu)^2] \\ &= E[X^2 - 2\mu X + \mu^2] \\ \text{LINEARITY} \rightarrow &= E[X^2] - 2\mu E[X] + E[\mu^2] \\ E[X] = \mu \rightarrow &= E[X^2] - 2\mu^2 + \mu^2 \\ &= E[X^2] - \mu^2 \end{aligned}$$

Q: How do we compute  $E[X^2]$

A:  $E[X^2] = \sum_a a^2 \times \Pr[X=a]$

$$E[X] = \sum_a a \times \Pr[X=a]$$

More generally, let  $y = f(x)$  be any function defined on the range of  $X$ . [E.g.  $y = x^2$ ]

Then  $Y$  is also a random variable:  $y(\omega) = f(X(\omega))$   
 $y(\omega) = X(\omega)^2$

Claim :  $E[f(X)] = \sum_a f(a) \times \Pr[X=a]$

Claim :  $E[f(X)] = \sum_a f(a) \times \Pr[X=a]$

Proof : Recall that, for any r.v.  $Y$ , we have

$$E[Y] = \sum_{\omega \in \Omega} Y(\omega) \times \Pr[\omega]$$



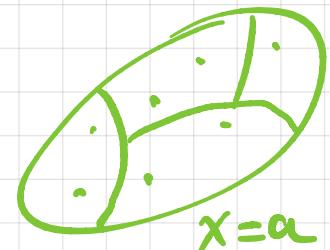
Apply this to the r.v.  $Y = f(X)$  :

$$E[Y] = \sum_{\omega \in \Omega} Y(\omega) \times \Pr[\omega]$$

$$= \sum_{\omega \in \Omega} f(X(\omega)) \times \Pr[\omega]$$

$$= \sum_a \sum_{\omega: X(\omega)=a} f(a) \times \Pr[\omega]$$

$$= \sum_a f(a) \times \Pr[X=a]$$



## Variance: Examples

1.  $X = \text{score on roll of a fair die}$

$$E[X] = \frac{1}{6} (1+2+3+4+5+6) = \boxed{\frac{7}{2}}$$

$$E[X^2] = \frac{1}{6} (1+4+9+16+25+36) = \boxed{\frac{91}{6}}$$

$$\text{So } \text{Var}(X) = E[X^2] - E[X]^2$$

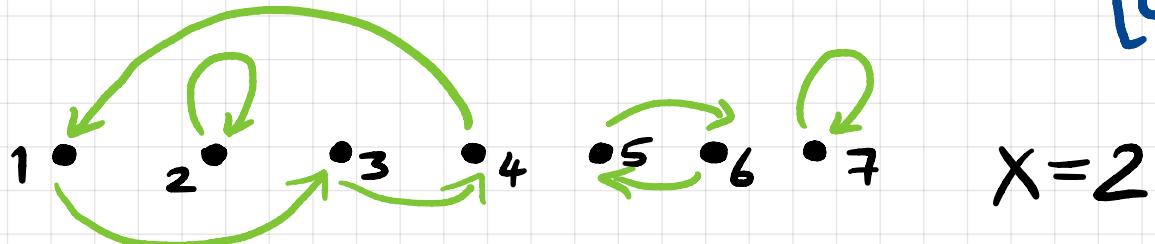
$$= \frac{91}{6} - \frac{49}{4}$$

$$= \boxed{\frac{35}{12}}$$

## 2. $X = \#$ fixed points in a random permutation

Recall:  $X = X_1 + X_2 + \dots + X_n$  where  $X_i = \begin{cases} 1 & \text{if } i \text{ is a fixed pt.} \\ 0 & \text{otherwise} \end{cases}$

E.g. :



$$E[X] = \sum_{i=1}^n E[X_i] = \sum_{i=1}^n \frac{1}{n} = 1$$

$$\begin{aligned} E[X_i] &= \Pr[X_i = 1] \\ &= \Pr[i \text{ a fixed pt.}] \\ &= \frac{1}{n} \end{aligned}$$

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

So we just need to compute  $E[X^2]$  . . .

So we just need to compute  $E[X^2]$  . . .

$$\begin{aligned} E[X^2] &= E[(X_1 + X_2 + \dots + X_n)^2] && (X_1 + X_2 + X_3)^2 \\ &= \sum_{i=1}^n E[X_i^2] + 2 \sum_{i < j} E[X_i X_j] && + 2(X_1 X_2 + X_1 X_3 + X_2 X_3) \end{aligned}$$

Note that: •  $X_i^2 = X_i$ , so  $E[X_i^2] = E[X_i] = \frac{1}{n}$

•  $X_i X_j = \begin{cases} 1 & \text{if } i, j \text{ both fixed points} \\ 0 & \text{otherwise} \end{cases}$

So  $E[X_i X_j] = \Pr[i, j \text{ both fixed points}] = \frac{1}{n(n-1)}$

Thus  $E[X^2] = \left(n \times \frac{1}{n}\right) + 2 \binom{n}{2} \frac{1}{n(n-1)} = 1 + 1 = \boxed{2}$

And so  $\text{Var}(X) = E[X^2] - E[X]^2 = 2 - 1 = \boxed{1}$

### 3. $X \sim \text{Geometric}(p)$

Recall :  $\Pr[X = k] = (1-p)^{k-1} p$   $k = 1, 2, 3, \dots$

$$E[X] = \frac{1}{p}$$

To compute  $E[X^2] = \sum_{k=1}^{\infty} k^2 (1-p)^{k-1} p$  :

- Start from  $\sum_{k=0}^{\infty} (1-p)^k = \frac{1}{p}$
  - Differentiate w.r.t.  $p$  :  $\sum_{k=1}^{\infty} k (1-p)^{k-1} = \frac{1}{p^2}$
  - Multiply by  $1-p$  :  $\sum_{k=1}^{\infty} k (1-p)^k = \frac{1-p}{p^2}$
  - Differentiate again :  $\sum_{k=1}^{\infty} k^2 (1-p)^{k-1} = \frac{2-p}{p^3}$
- $\leftarrow = \frac{1}{p} E[X^2]!$

Hence  $E[X^2] = \frac{2-p}{p^2} \Rightarrow \text{Var}(X) = \frac{2-p}{p^2} - \frac{1}{p^2} = \boxed{\frac{1-p}{p^2}}$

4.  $X \sim \text{Poisson}(\lambda)$

Recall:  $\Pr[X = k] = e^{-\lambda} \frac{\lambda^k}{k!}$   $k=0,1,2, \dots$

$$E[X] = \lambda$$

$$E[X^2] = \sum_{k=0}^{\infty} k^2 e^{-\lambda} \frac{\lambda^k}{k!}$$

$$= \lambda e^{-\lambda} \sum_{k=1}^{\infty} k \frac{\lambda^{k-1}}{(k-1)!}$$

$$k = (k-1) + 1$$

$$= \lambda e^{-\lambda} \sum_{k=1}^{\infty} \left[ (k-1) \frac{\lambda^{k-1}}{(k-1)!} + \frac{\lambda^{k-1}}{(k-1)!} \right]$$

$$= \lambda e^{-\lambda} \left[ \sum_{k=2}^{\infty} \frac{\lambda^{k-1}}{(k-2)!} + \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} \right]$$

$$= \lambda e^{-\lambda} [\lambda e^{\lambda} + e^{\lambda}] = \boxed{\lambda^2 + \lambda}$$

Hence  $\text{Var}(X) = (\lambda^2 + \lambda) - \lambda^2 = \boxed{\lambda}$

## 5. $X \sim \text{Binomial}(n, p)$

Recall :  $\Pr[X=k] = \binom{n}{k} p^k (1-p)^{n-k}$

$$E[X] = np$$

Also :  $X = X_1 + X_2 + \dots + X_n$  where  $X_i = \begin{cases} 1 & \text{if } i\text{th toss is Heads} \\ 0 & \text{otherwise} \end{cases}$

To compute  $E[X^2]$ , we could use :

$$\begin{aligned} E[X^2] &= E[(X_1 + \dots + X_n)^2] \\ &= \sum_{i=1}^n E[X_i^2] + 2 \sum_{i < j} E[X_i X_j] \\ &= \dots \end{aligned}$$

But it's much easier to use the fact that the  $X_i$  are independent

## Variance of a Sum

For any two random variables  $X, Y$ , we have :

$$\begin{aligned}\text{Var}(X+Y) &= E[(X+Y)^2] - (E[X+Y])^2 \\ &= E[X^2 + Y^2 + 2XY] - (E[X] + E[Y])^2 \\ &= E[X^2] + E[Y^2] + 2E[XY] - E[X]^2 - E[Y]^2 \\ &\quad - 2E[X]E[Y]\end{aligned}$$

$$= \boxed{\text{Var}(X) + \text{Var}(Y) + 2\underbrace{(E[XY] - E[X]E[Y])}_{\text{Cov}(X,Y)}}$$

$$\text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X,Y)$$

Covariance  
 $\text{Cov}(X,Y)$

Claim : If  $X, Y$  are independent then  $\text{Cov}(X,Y) = 0$

Corollary :  $X, Y$  independent  $\Rightarrow \text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y)$

Claim : If  $X, Y$  are independent then  $\text{Cov}(X, Y) = 0$

Proof : Recall that  $X, Y$  independent means that

$$\Pr[X=a, Y=b] = \Pr[X=a] \times \Pr[Y=b] \quad \forall a, b$$

$$E[XY] = \sum_a \sum_b ab \times \Pr[X=a, Y=b]$$

INDEPENDENCE  $\rightarrow$

$$\begin{aligned} &= \sum_a \sum_b ab \times \Pr[X=a] \times \Pr[Y=b] \\ &= \left( \sum_a a \Pr[X=a] \right) \left( \sum_b b \Pr[Y=b] \right) \\ &= E[X] E[Y] \end{aligned}$$

Hence  $\text{Cov}(X, Y) = E[XY] - E[X]E[Y] = 0$

## 5. $X \sim \text{Binomial}(n, p)$

Recall :  $\Pr[X=k] = \binom{n}{k} p^k (1-p)^{n-k}$

$$E[X] = np$$

Also :  $X = X_1 + X_2 + \dots + X_n$  where  $X_i = \begin{cases} 1 & \text{if } i\text{th toss is Heads} \\ 0 & \text{otherwise} \end{cases}$

Since the  $X_i$  are independent :

$$\text{Var}(X) = \sum_{i=1}^n \text{Var}(X_i)$$

$$\text{But } \text{Var}(X_i) = E[X_i^2] - E[X_i]^2 = p - p^2 = p(1-p)$$

Hence  $\text{Var}(X) = \boxed{np(1-p)}$

For  $X \sim \text{Binomial}(n, p)$  :

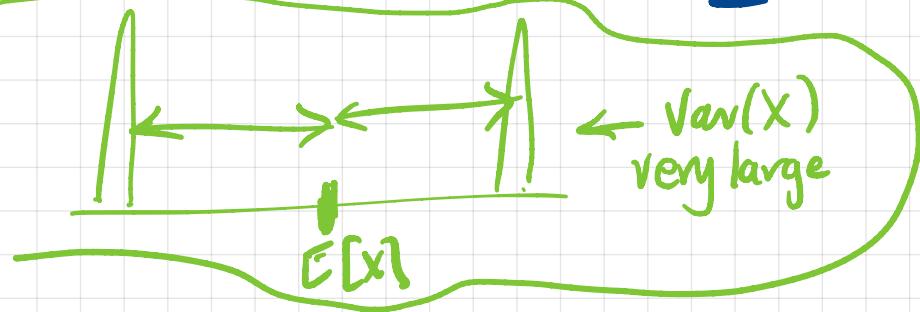
$$E[X] = np$$

$$\text{Var}(X) = np(1-p)$$

E.g.  $p = 1/2$  (fair coin) :

$$E[X] = \frac{n}{2}$$

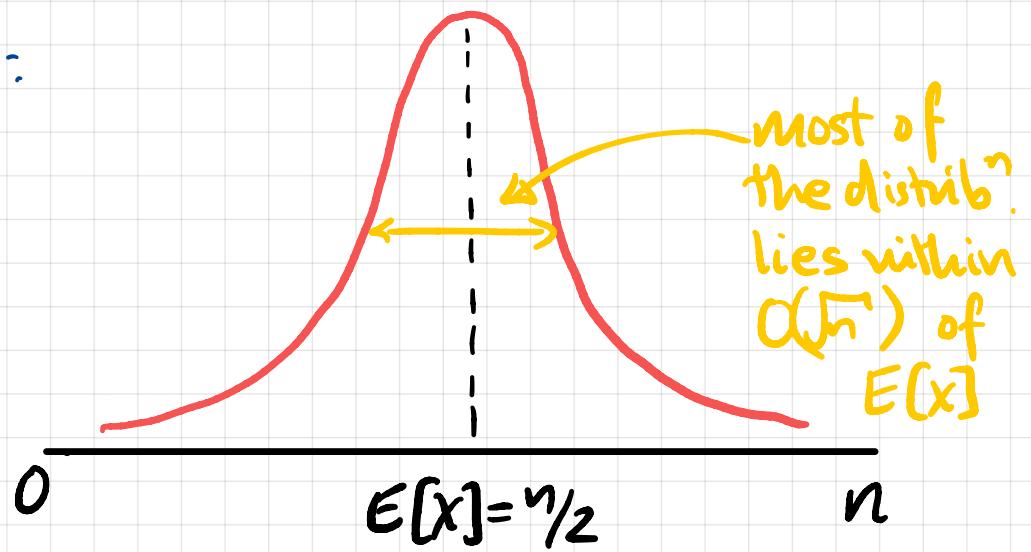
$$\text{Var}(X) = \frac{n}{4}$$



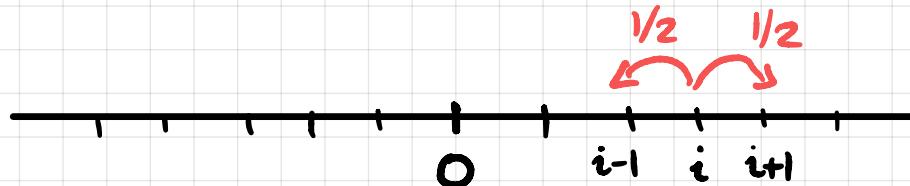
$$\sigma(X) = \sqrt{\text{Var}(X)} = \frac{\sqrt{n}}{2}$$

Intuitive Interpretation :

We will make this idea precise in the next lecture



## 6. Random Walk ("Drunkard's Walk")



- Start at 0
- Each second, move
  - { 1 step right w. prob. 1/2
  - { 1 step left —..—

$S_n$  = position after  $n$  seconds

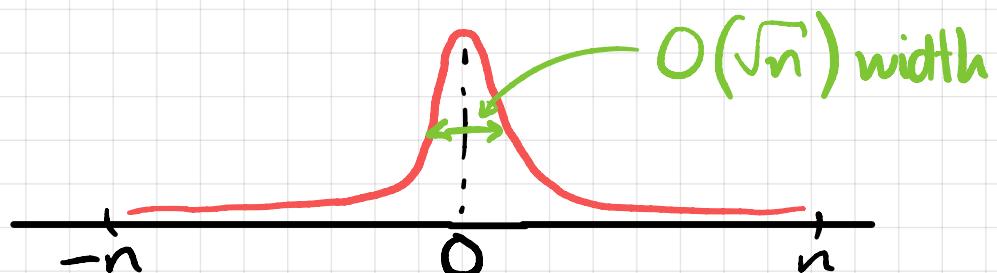
$$S_n = \sum_{i=1}^n X_i \quad \text{where} \quad X_i = \begin{cases} 1 & \text{if } i\text{th step to right} \\ -1 & \text{—..— left} \end{cases}$$

$$E[S_n] = \sum_{i=1}^n E[X_i] = 0$$

$$\text{Var}(S_n) = \sum_{i=1}^n \text{Var}(X_i) = n \times 1 = n$$

$$\text{So } \sigma(S_n) = \sqrt{n}$$

$$\begin{aligned} \text{Var}(X_i) &= \\ E[X_i^2] - E[X_i]^2 &= \\ 1 - 0 &= 1 \end{aligned}$$



## Covariance & Correlation

Recall :  $\text{Cov}(X, Y) = E[XY] - E[X]E[Y]$

Equivalently :  $\text{Cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)]$

$$\begin{aligned}
 \text{Var}(X) &= E[(X - E[X])^2] \geq 0 \\
 &= E[XY] - \mu_X E[Y] - \mu_Y E[X] + \mu_X \mu_Y \\
 &= E[XY] - \mu_X \mu_Y - \mu_X \mu_Y + \mu_X \mu_Y \\
 &= E[XY] - E[X]E[Y]
 \end{aligned}$$

c.f.  $\text{Var}(X) = E[X^2] - E[X]^2$

c.f.  $\text{Var}(X) = E[(X - \mu_X)^2]$

## Properties

1.  $X, Y$  independent  $\Rightarrow \text{Cov}(X, Y) = 0$

[Note : This is  
NOT  $\Leftrightarrow$ ]

2.  $\text{Cov}(X, X) = \text{Var}(X)$

3.  $\text{Cov}(X, Y) = \frac{1}{2} [\text{Var}(X+Y) - \text{Var}(X) - \text{Var}(Y)]$

Recall :  $\text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y) - 2\text{Cov}(X, Y)$

$\text{Cov}(X, Y) > 0 \Rightarrow X, Y$  positively correlated

$\text{Cov}(X, Y) < 0 \Rightarrow X, Y$  negatively correlated

Better measure :

Defn: The correlation  $\text{Corr}(X, Y)$  is defined as

$$\text{Corr}(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma(X)\sigma(Y)}$$

### Properties

1.  $-1 \leq \text{Corr}(X, Y) \leq +1$  [Proof: see notes]

2.  $X, Y$  independent  $\Rightarrow \text{Corr}(X, Y) = 0$

3.  $\text{Corr}(X, Y) = +1 \Rightarrow Y = aX + b$  for constants  $a > 0, b$

$\text{Corr}(X, Y) = -1 \Rightarrow Y = aX + b$   $\dots \quad a < 0, b$