CS70 – Spring 2024

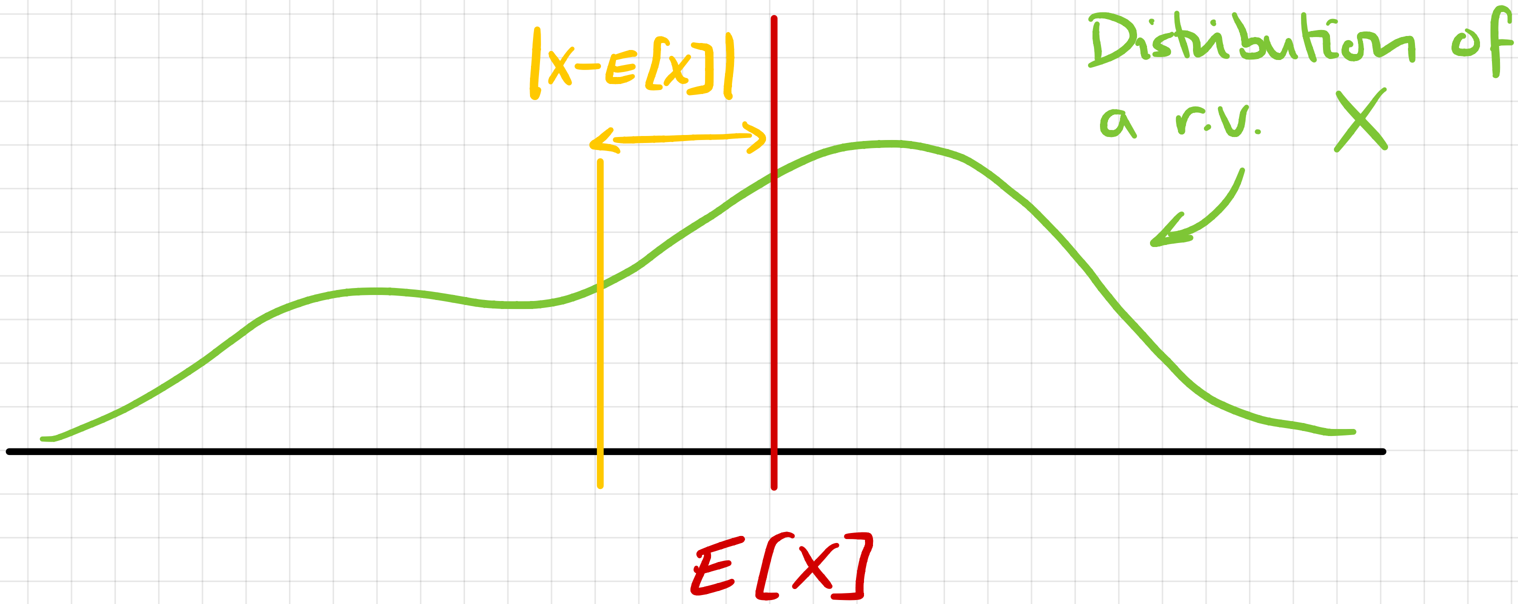
Lecture 20 – April 2

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

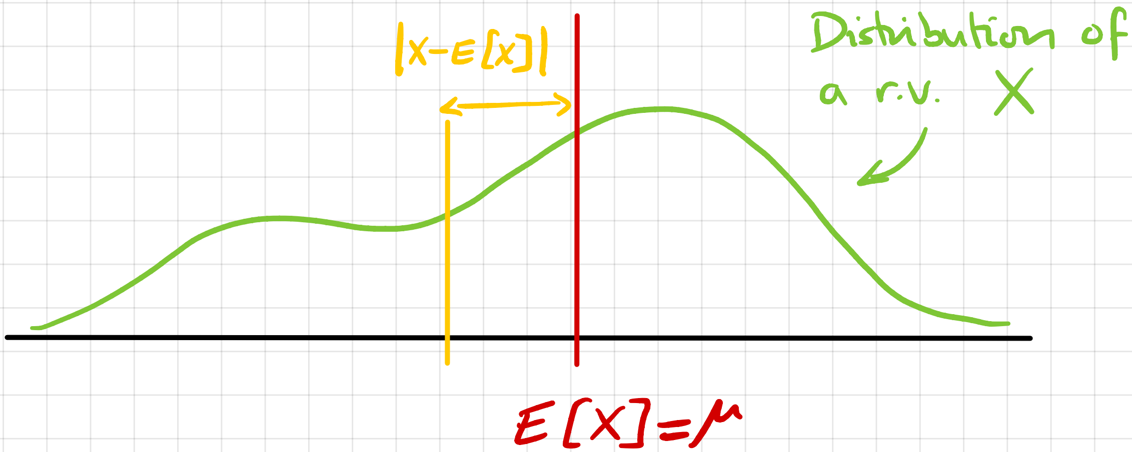
Variance

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

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

where $\mu = E[X]$.

The standard deviation of X is

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

A more convenient expression

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

Proof :

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

$\boxed{Q}$: How do we compute $E[X^2]$

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

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: $Y(\omega) = f(X(\omega))$

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

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 = score on roll of a fair die

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

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

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

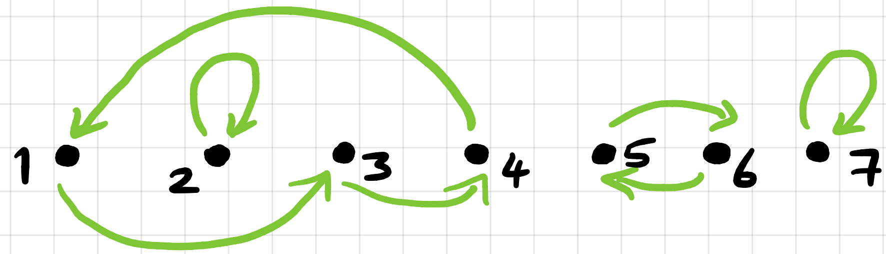
$$= \frac{91}{6} - \frac{49}{4}$$

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

2. X = # fixed points in a random permutation

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

E.g.:



$X = 2$

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

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

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

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

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

$$E[X^2] = E[(X_1 + X_2 + \cdots + X_n)^2]$$

$$= \sum_{i=1}^{n} E[X_i^2] + 2\sum_{i<j} E[X_i X_j]$$

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

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

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

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

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 : $\sum_{k=1}^{\infty} \underbrace{k^2(1-p)^{k-1}}_{= \frac{1}{p}E[X^2]\,!} = \frac{2-p}{p^3}$

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

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

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

$$E[X] = \lambda$$

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

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

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

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

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

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

5. $X \sim$ Binomial (n, p)

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

$E[X] = np$

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

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

$$E[X^2] = E[(X_1 + \cdots + X_n)^2]$$

$$= \sum_{i=1}^{n} E[X_i^2] + 2 \sum_{i<j} E[X_i X_j]$$

$$= \cdots$$

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:

$$\begin{aligned} Var(X+Y) &= E[(X+Y)^2] - (E[X+Y])^2 \\ &= E[X^2+Y^2+2XY] - (E[X]+E[Y])^2 \\ &= E[X^2]+E[Y^2]+2E[XY] - E[X]^2 - E[Y]^2 \\ &\quad - 2E[X]E[Y] \\ &= \boxed{Var(X) + Var(Y) + 2(\underbrace{E[XY]-E[X]E[Y]}_{\substack{\text{Covariance} \\ Cov(X,Y)}})} \end{aligned}$$

Claim: If X, Y are *independent* then $Cov(X,Y) = 0$

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

Claim: If X, Y are <u>independent</u> then $Cov(X,Y) = 0$

Proof: Recall that X, Y independent means that

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

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

INDEPENDENCE $\longrightarrow$ $$= \sum_a \sum_b ab \times \Pr[X=a] \times \Pr[Y=b]$$

$$= \left(\sum_a a \Pr[X=a]\right)\left(\sum_b b \Pr[Y=b]\right)$$

$$= E[X]\, E[Y]$$

Hence $Cov(X,Y) = E[XY] - E[X]E[Y] = 0$

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

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

$$E[X] = np$$

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

Since the X_i are independent:

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

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

$$E[X] = np \qquad \text{Var}(X) = np(1-p)$$

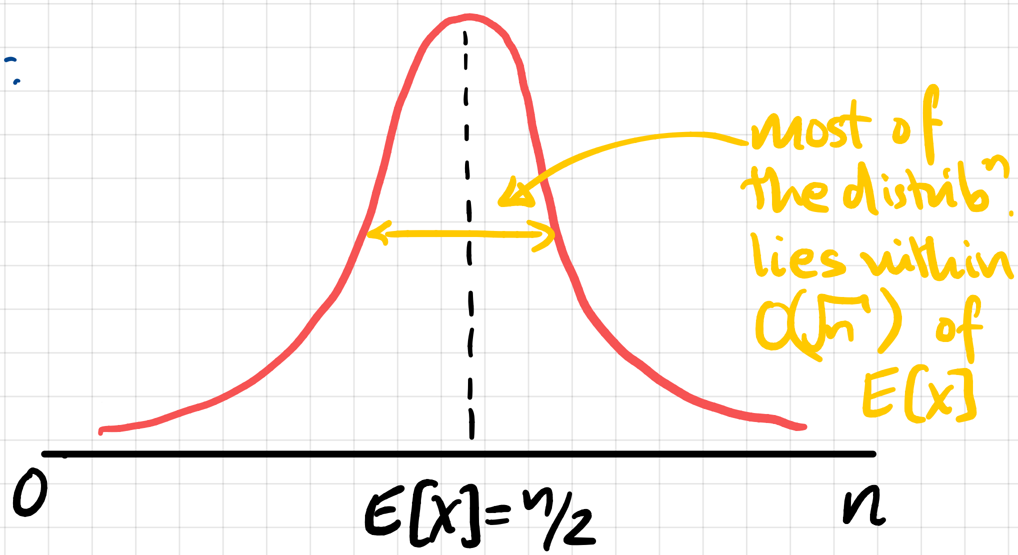
E.g. $p = 1/2$ (fair coin):

$$E[X] = \frac{n}{2} \qquad \text{Var}(X) = \frac{n}{4}$$

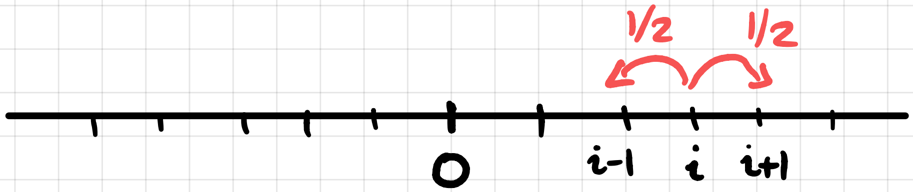
$$\boxed{\sigma(X) = \sqrt{\text{Var}(X)} = \frac{\sqrt{n}}{2}}$$

<u>Intuitive Interpretation</u>:

We will make this idea precise in the next lecture



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



- Start at 0
- Each second, move
$$\begin{cases} 1 \text{ step right} & \text{w. prob. } 1/2 \\ 1 \text{ step left} & \text{— " —} \end{cases}$$

S_n = position after n seconds

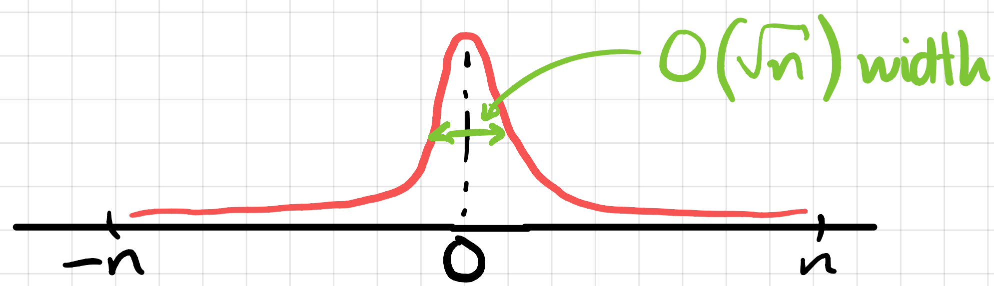
$$S_n = \sum_{i=1}^{n} X_i \quad \text{where} \quad X_i = \begin{cases} 1 & \text{if ith step to right} \\ -1 & \text{— " — left} \end{cases}$$

$$\mathrm{Var}(X_i) = E[X_i^2] - E[X_i]^2 = 1 - 0 = 1$$

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

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

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



Covariance & Correlation

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

Equivalently: $\mathrm{Cov}(X,Y) = E[(X-\mu_x)(Y-\mu_y)]$

$$= E[XY] - \mu_x E[Y] - \mu_y E[X] + \mu_x\mu_y$$

$$= E[XY] - \mu_x\mu_y - \mu_x\mu_y + \mu_x\mu_y$$

$$= E[XY] - E[X]E[Y]$$

Properties

1. X, Y independent $\Rightarrow \mathrm{Cov}(X,Y) = 0$ [Note: This is NOT $\Leftrightarrow$]
2. $\mathrm{Cov}(X,X) = \mathrm{Var}(X)$
3. $\mathrm{Cov}(X,Y) = \frac{1}{2}\left[\mathrm{Var}(X+Y) - \mathrm{Var}(X) - \mathrm{Var}(Y)\right]$

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

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

Better measure :

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

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

Properties

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

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

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

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