CS 5100 02/03

* Probability ⇒ quantify uncertainty
* Same Representation for different problems

random variable X

X	0	1
P(X)	1/2	1/2

X	1	2	3	4	5	6
P(X)	1/6	1/6	1/6	1/6	1/6	1/6

* Look at an average value of the random variable

look at the probabilities to decide how much weight to put on every value of x

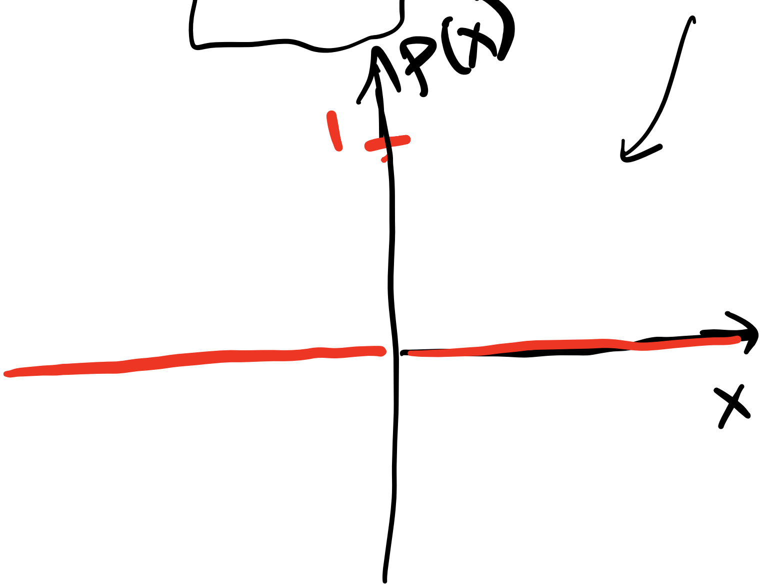
$$0 \cdot \tfrac{1}{2} + 1 \cdot \tfrac{1}{2} = \tfrac{1}{2}$$

$$1 \cdot \tfrac{1}{6} + 2 \cdot \tfrac{1}{6} + 3 \cdot \tfrac{1}{6} + 4 \cdot \tfrac{1}{6} + 5 \cdot \tfrac{1}{6} + 6 \cdot \tfrac{1}{6} = 3.5$$

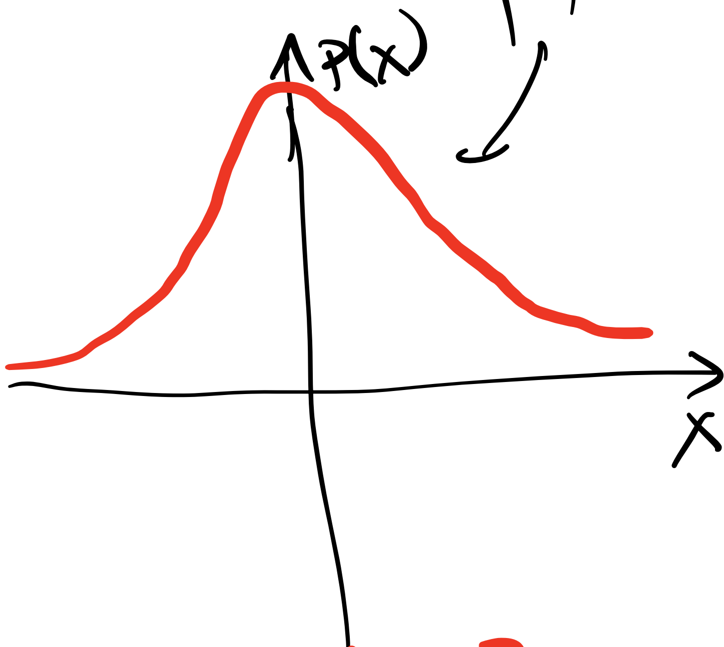
$$E(X) = \sum_a a\, p(X=a) = \text{average value of } X = \text{expected value of } X$$

X	2	3	4	5	6	7	8	9	10	11	12
$p(x)$	$1/36$	$2/36$	$3/36$	$4/36$	$5/36$	$6/36$	$5/36$	$4/36$	$3/36$	$2/36$	$1/36$

$\boxed{E(X)} = 7$



$E(X) = 0$



$E(X) = 0$

$\Rightarrow$ the two distributions don't have the same spread and $E(X)$ is not able to capture such information

$\Rightarrow$ Need another metric

50 60 60 50 → 55

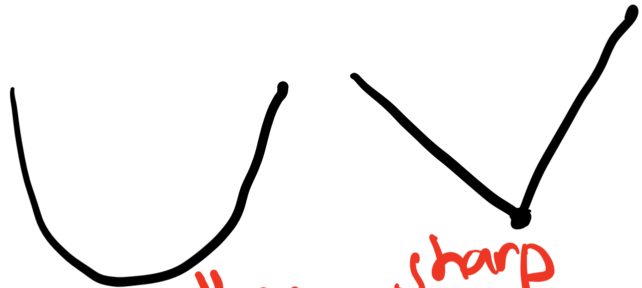
30 80 30 80 → 55

$$(50-55)+(60-55)+(60-55)+(50-55)=0$$

negative not good

square ✓

absolute value



smoother sharp

$$(50-55)^2+(60-55)^2+(60-55)^2+(50-55)^2=100$$

$$(30-55)^2+(80-55)^2+(30-55)^2+(80-55)^2=2500$$

give me intuition about the spread of the square values

variance

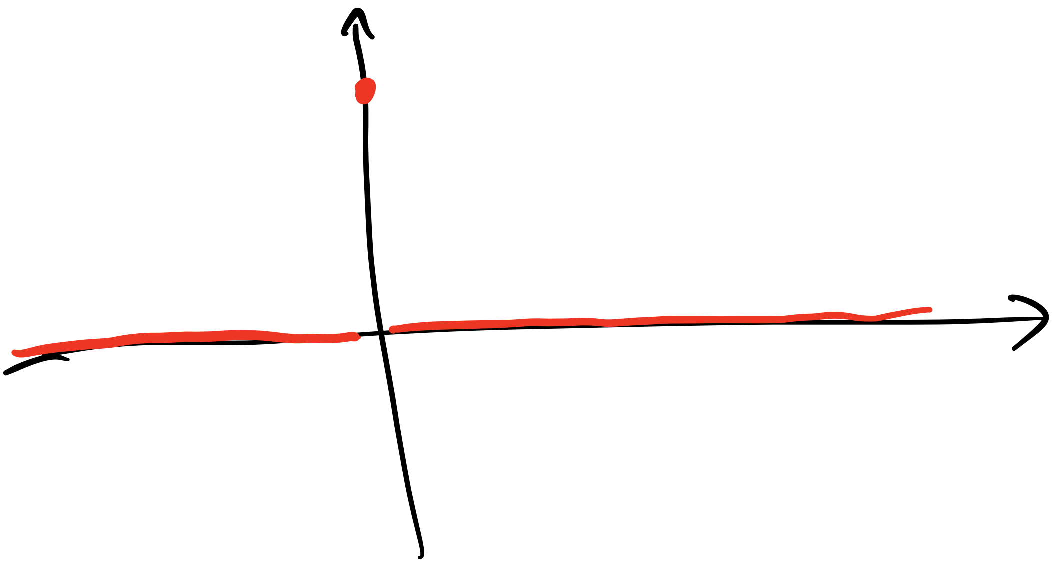
standard deviation

Let's take a square root

$$Var(X)=E\left[(X-E(X))^2\right] \checkmark$$

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

standard deviation



$E(X) = 0$

$E(X) = 1/2$

X	0	1
$P(X)$	1/2	1/2

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

$$= \boxed{E\left[X^2 - 2X\,E(X) + (E(X))^2\right]}$$

* 1 4 6 9 $E(X) = 5$

* 2 8 12 18 $E(2X) = 10 = 2E(X)$

$$E[aX) = a\,E(X)$$

more probability (trust me, will convince you later

* $E(X_1 + X_2) = E(X_1) + E(X_2)$ (pinkie promise)

two random variables X_1 and X_2

$$Var(X) = E\left[X^2 - 2XE(X) + (E(X))^2\right]$$

$$= E(X^2) + E[-2XE(X)] + E\left[\underbrace{(E(X))^2}_{}\right]$$

(think of it as a 4)

$a = -2E(X)$

$$= E(X^2) + \underbrace{E[-2XE(X)]}_{} + (E(X))^2$$

$$= E(X^2) + (\underbrace{-2E(X)}_{a})E(X) + (E(X))^2$$

$$= E(X^2) - 2(E(X))^2 + (E(X))^2$$

$$= E(X^2) - (E(X))^2$$

Coin example

$E(X) = \frac{1}{2}$

X	0	1
$P(X)$	$\frac{1}{2}$	$\frac{1}{2}$

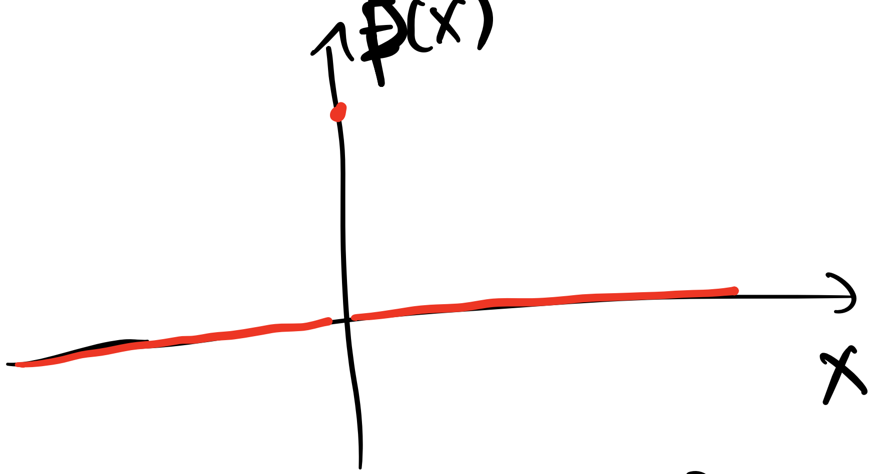
X^2	0^2	1^2
$P(X)$	$\frac{1}{2}$	$\frac{1}{2}$

$$E(X^2) = 0^2 \cdot \tfrac{1}{2} + 1^2 \cdot \tfrac{1}{2} = \tfrac{1}{2}$$

$$Var(X) = \tfrac{1}{2} - \left(\tfrac{1}{2}\right)^2 = \tfrac{1}{4}$$

$$\sigma(X) = \sqrt{Var X} = \tfrac{1}{2}$$

Yay makes sense



$E(X)=0$

$$P(x) = \begin{cases} 1 & \text{if } x=0 \\ 0 & \text{if } x\neq 0 \end{cases}$$

$$E(X^2) = 0^2 \cdot 1 = 0$$

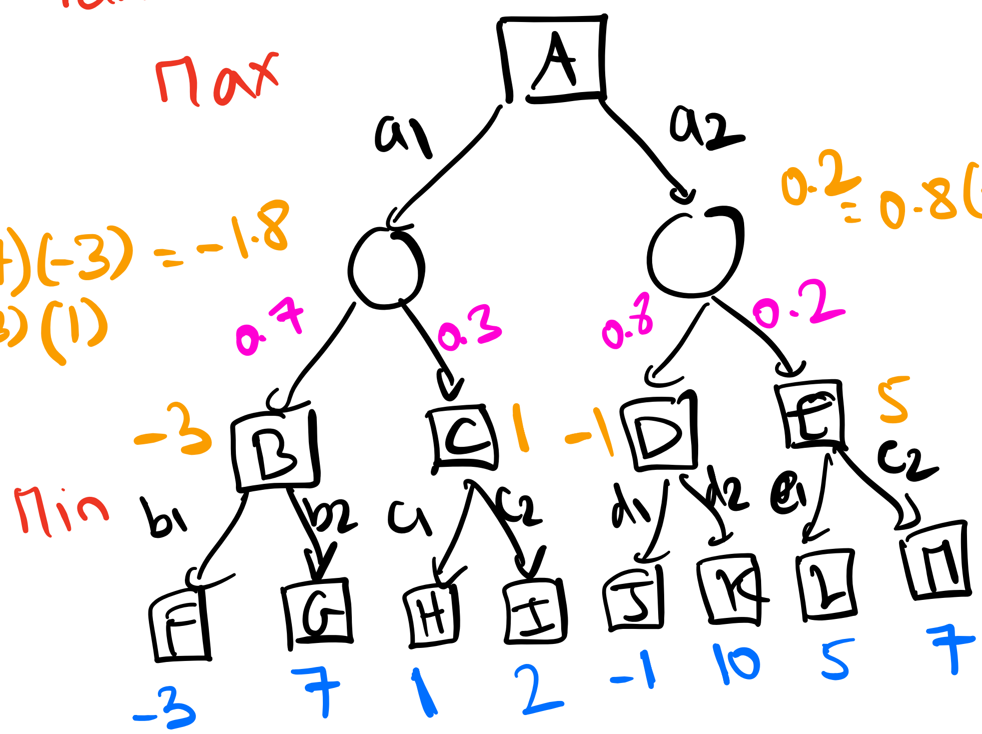
$$Var(X) = E(X^2) - E(X)^2 = 0 - 0^2 = 0.$$

Games with randomness

(Stochastic games)

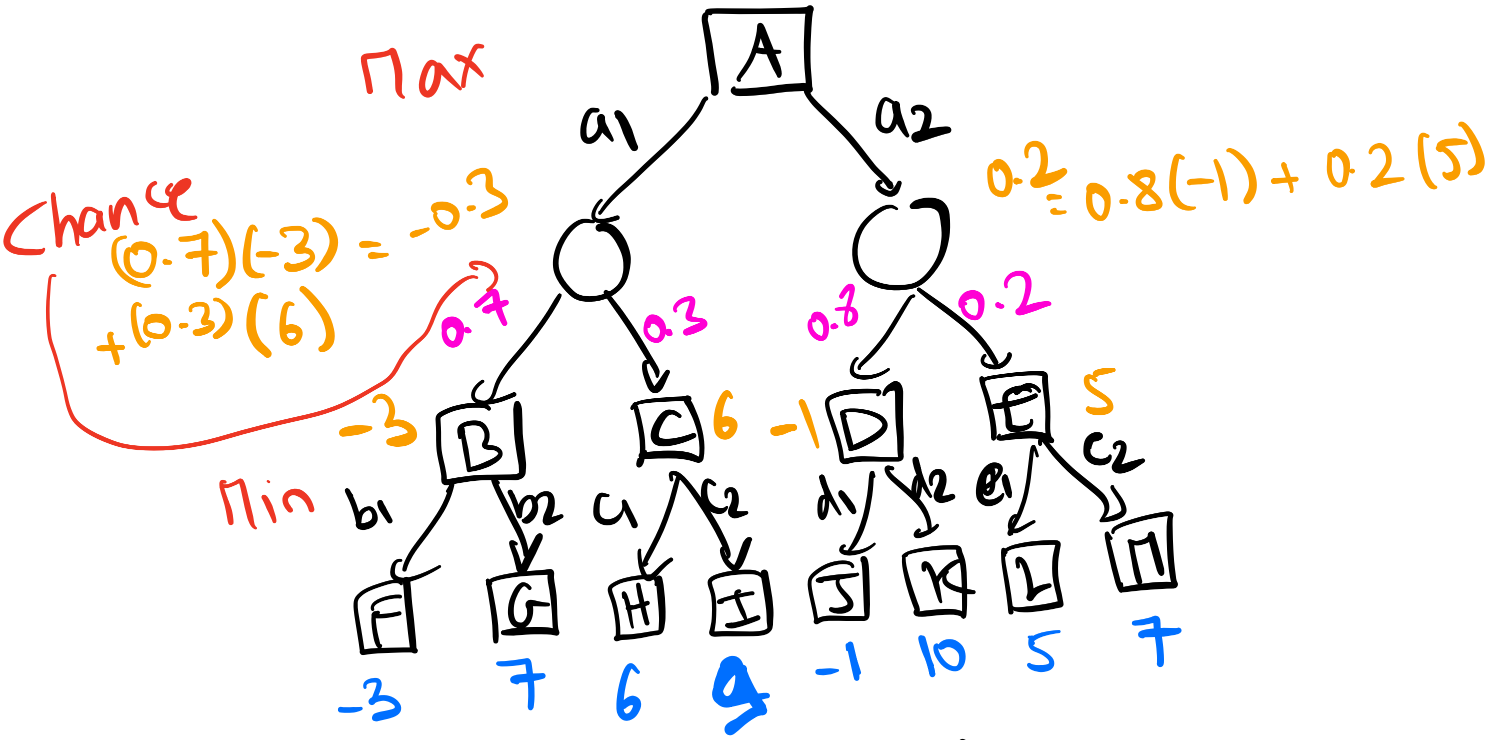
$\parallel$

random



$(0.7)(-3) + (0.3)(1) = -1.8$

$0.2 = 0.8(-1) + 0.2(5)$



$$
\text{Expectiminimax}(s) = \begin{cases} \text{Utility}(s) & \text{if } s \text{ is a terminal state} \\ \max\limits_{\substack{a \in \text{actions} \\ s' \in \text{succ}(s)}} \text{Expectiminimax}(s') & \text{if Player} = \text{Max} \\ \min\limits_{\substack{a \in \text{actions} \\ s' \in \text{succ}(s)}} \text{Expectiminimax}(s') & \text{if Player} = \text{Min} \end{cases}
$$

$\sum_{s'} P(s \to s') \cdot \text{Expectiminimax}(s')$; if Player = Chance

[more for coding
chance is not really a player]

$P(B)\,\text{Expectiminimax}(B)$ s_1'
$+\,P(C)$ " (C) s_2'

$$\sum_{s'} P(s \to s')\ \text{Expectiminimax}(s')$$

Algorithm (similar to what we did last time
we will need function
max, min and chance which
call each other)

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

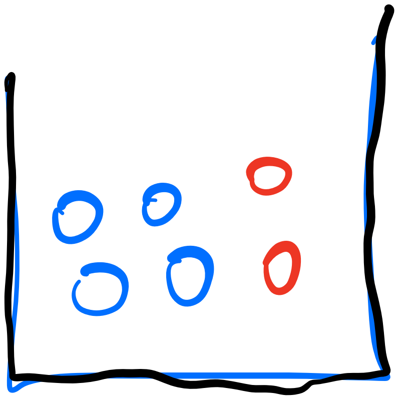
Exercise

X_1 , X_2 random variables

$$P(X_1 = a \mid X_2 = b) = \frac{P(X_1 = a \text{ and } X_2 = b)}{P(X_2 = b)}$$

$$\boxed{P(X_2 = b \mid X_1 = a) = \frac{P(X_2 = b \text{ and } X_1 = a)}{P(X_1 = a)}}$$

$$\boxed{P(X_1=a \text{ and } X_2=b) = P(X_1=a)\,P(X_2=b \mid X_1=a)}$$



$P(X_1 = \bigcirc)$

$P(X_2 = \bigcirc)$ $\leftarrow$ depends on what I got for X_1

$$P(X_1 = a \text{ and } X_2 = b \text{ and } X_3 = c)$$

$$= \underbrace{P(\overbrace{X_1 = a,\ X_2 = b}^{B},\ \overbrace{X_3 = c}^{A})}_{\text{joint distribution}} = P(X_1=a)\cdot P(X_2=b \mid X_1=a) \cdot P(X_3=c \mid X_1=a, X_2=b)$$

$$P(A \cap B) = P(X_3=c \mid X_1=a, X_2=b) \cdot P(X_1=a, X_2=b)$$

$$P(A \mid B) = \frac{P(A\cap B)}{P(B)} \Rightarrow P(A\cap B) = P(A\mid B)\cdot P(B)$$

$$P(X_1=a, X_2=b, X_3=c) = P(X_3=c \mid X_1=a, X_2=b)\, \underbrace{P(\overbrace{X_1=a}^{B},\ \overbrace{X_2=b}^{A})}_{\substack{= P(A\cap B) \\ = P(A\mid B)\cdot P(B)}}$$

$$= P(X_3 = c \mid X_1 = a, X_2 = b) \cdot P(X_2 = b \mid X_1 = a) \cdot P(X_1 = a)$$