

CSE 312

# Foundations of Computing II


## Lecture 9: Linearity of Expectation, LOTUS, and Variance



**Aleks Jovicic**

Slide Credit: Based on Stefano Tessaro's slides for 312 19au incorporating ideas from Anna Karlin, Alex Tsun, Rachel Lin, Hunter Schafer & myself 😊

# Agenda

- Linearity of Expectation 
- Indicator Random Variables
- LOTUS
- Variance

# Coin flipping again

$$n=5$$
$$k=2$$

HTTT  
+

Suppose we flip a coin independently  $n$  times with probability  $p$  of coming up Heads each time. Let the r.v.  $Z$  be the number of Heads in the  $n$  coin flips. What is the p.m.f. of  $Z$ ?

$$\underline{P(Z=k)} = \begin{cases} k=0 \\ k=1 \\ \dots \end{cases} \quad P(Z=k) =$$

# Expectation of Random Variable

**Definition.** Given a discrete RV  $X: \Omega \rightarrow \mathbb{R}$ , the **expectation or expected value** of  $X$  is

$$E[X] = \sum_{\omega \in \Omega} X(\omega) \cdot \Pr(\omega)$$

or equivalently

$$E[X] = \sum_{x \in \Omega_X} x \cdot \Pr(X = x)$$

Intuition: “Weighted average” of the possible outcomes (weighted by probability)

## Coin flipping again

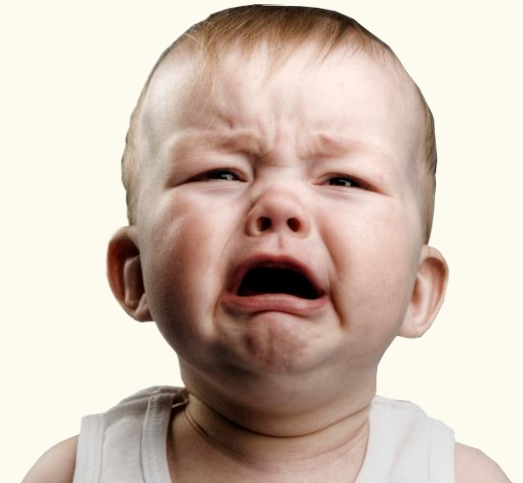
Suppose we flip a coin independently  $n$  times with probability  $p$  of coming up Heads each time. Let the r.v.  $Z$  be the number of Heads in the  $n$  coin flips. What is the  $\mathbb{E}(Z)$  ?

# The brute force method

we flip  $n$  coins, each one heads with probability  $p$ ,

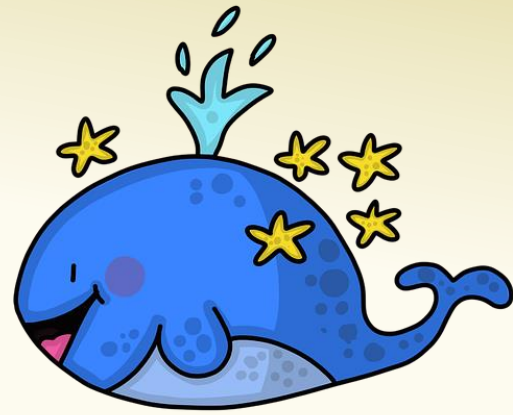
$Z$  is the number of heads, what is  $\mathbb{E}(Z)$ ?

$$\begin{aligned}\mathbb{E}[Z] &= \sum_{k=0}^n k \cdot P(Z = k) = \sum_{k=0}^n k \cdot \binom{n}{k} p^k (1-p)^{n-k} \\ &= \sum_{k=0}^n k \cdot \frac{n!}{k! (n-k)!} p^k (1-p)^{n-k} = \sum_{k=1}^n \frac{n!}{(k-1)! (n-k)!} p^k (1-p)^{n-k} \\ &= np \sum_{k=1}^n \frac{(n-1)!}{(k-1)! (n-k)!} p^{k-1} (1-p)^{n-k} \\ &= np \sum_{k=0}^{n-1} \frac{(n-1)!}{k! (n-1-k)!} p^k (1-p)^{(n-1)-k} \\ &= np \sum_{k=0}^{n-1} \binom{n-1}{k} p^k (1-p)^{(n-1)-k} = np(p + (1-p))^{n-1} = np \cdot 1 = np\end{aligned}$$



[This Photo](#) by Unknown Author is licensed under [CC BY-NC](#)

## Linearity of Expectation (Idea)

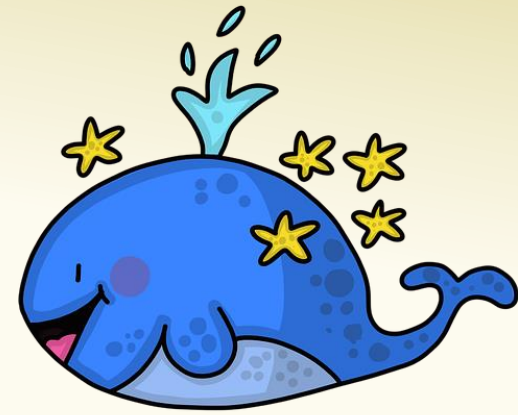


Let's say you and your friend sell fish for a living.

- Every day you catch  $X$  fish, with  $E[X] = 3$ .
- Every day your friend catches  $Y$  fish, with  $E[Y] = 7$ .

How many fish do the two of you bring in ( $Z = X + Y$ ) on an average day?

## Linearity of Expectation (Idea)



Let's say you and your friend sell fish for a living.

- Every day you catch  $X$  fish, with  $E[X] = 3$ .
- Every day your friend catches  $Y$  fish, with  $E[Y] = 7$ .

How many fish do the two of you bring in ( $Z = X + Y$ ) on an average day?

$$E[Z] = E[X + Y] = E[X] + E[Y] = 3 + 7 = 10$$



## Linearity of Expectation (Idea)



Let's say you and your friend sell fish for a living.

- Every day you catch  $X$  fish, with  $E[X] = 3$ .
- Every day your friend catches  $Y$  fish, with  $E[Y] = 7$ .

How many fish do the two of you bring in ( $Z = X + Y$ ) on an average day?

$$E[Z] = E[X + Y] = E[X] + E[Y] = 3 + 7 = 10$$

You can sell each fish for \$5 at a store, but you need to pay \$20 in rent. How much profit do you expect to make?

$$E[5Z - 20] = 5E[Z] - 20 = 5 \times 10 - 20 = 30$$

# Linearity of Expectation – Proof

**Theorem.** For **any** two random variables  $X$  and  $Y$

$$\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y).$$

$$\begin{aligned}\mathbb{E}(X + Y) &= \sum_{\omega} P(\omega)(X(\omega) + Y(\omega)) \\ &= \sum_{\omega} P(\omega)X(\omega) + \sum_{\omega} P(\omega)Y(\omega) \\ &= \mathbb{E}(X) + \mathbb{E}(Y)\end{aligned}$$

# Linearity of Expectation

**Theorem.** For **any** two random variables  $X$  and  $Y$

$$\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y).$$

Or, more generally: For any random variables  $X_1, \dots, X_n$ ,

$$\mathbb{E}(X_1 + \dots + X_n) = \mathbb{E}(X_1) + \dots + \mathbb{E}(X_n).$$

**Because:**  $\mathbb{E}(X_1 + \dots + X_n) = \mathbb{E}((X_1 + \dots + X_{n-1}) + X_n)$   
 $= \mathbb{E}(X_1 + \dots + X_{n-1}) + \mathbb{E}(X_n) = \dots$

## Coin flipping again

Suppose we flip a coin independently  $n$  times with probability  $p$  of coming up Heads each time. Let the r.v.  $Z$  be the number of Heads in the  $n$  coin flips. What is the  $\mathbb{E}(Z)$  ?

## Example – Coin Tosses

we flip  $n$  coins, each one heads with probability  $p$

$Z$  is the number of heads, what is  $\mathbb{E}(Z)$  ?

$$- X_i = \begin{cases} 1, & i\text{-th coin-flip is heads} \\ 0, & i\text{-th coin-flip is tails.} \end{cases}$$

$$\text{Fact. } Z = X_1 + \cdots + X_n$$

### Linearity of Expectation:

$$\mathbb{E}(Z) = \mathbb{E}(X_1 + \cdots + X_n) = \mathbb{E}(X_1) + \cdots + \mathbb{E}(X_n) = n \cdot p$$

$$\begin{aligned} \mathbb{P}(X_i = 1) &= p \\ \mathbb{P}(X_i = 0) &= 1 - p \end{aligned}$$

$$\mathbb{E}(X_i) = p \cdot 1 + (1 - p) \cdot 0 = p$$

# Computing complicated expectations

Often boils down to the following three steps

- Decompose: Finding the right way to decompose the random variable into sum of simple random variables

$$X = X_1 + \cdots + X_n$$

- LOE: Observe linearity of expectation.

$$\mathbb{E}(X) = \mathbb{E}(X_1) + \cdots + \mathbb{E}(X_n).$$

- Conquer: Compute the expectation of each  $X_i$

Often,  $X_i$  are **indicator** (0/1) random variables.

# Agenda

- Linearity of Expectation
- **Indicator Random Variables** ◀
- LOTUS
- Variance

## Indicator random variable

For any event  $A$ , can define the indicator random variable  $X$  for  $A$

$$X = \begin{cases} 1 & \text{if event } A \text{ occurs} \\ 0 & \text{if event } A \text{ does not occur} \end{cases}$$

$$\begin{aligned} \mathbb{P}(X = 1) &= \mathbb{P}(A) \\ \mathbb{P}(X = 0) &= 1 - \mathbb{P}(A) \end{aligned}$$

$$E[X] = \mathbb{P}(A) \cdot 1 + (1 - \mathbb{P}(A)) \cdot 0 = \mathbb{P}(A)$$



## Example: Returning Homeworks

- Class with  $n$  students, randomly hand back homeworks. All permutations equally likely.
- Let  $X$  be the number of students who get their own HW
- what is  $\mathbb{E}(X)$ ?

$\Pr(\omega)$	$\omega$	$X(\omega)$
1/6	1, 2, 3	3
1/6	1, 3, 2	1
1/6	2, 1, 3	1
1/6	2, 3, 1	0
1/6	3, 1, 2	0
1/6	3, 2, 1	1

## Example: Returning Homeworks

- Class with  $n$  students, randomly hand back homeworks. All permutations equally likely.
- Let  $X$  be the number of students who get their own HW
- what is  $\mathbb{E}(X)$ ?
- Use Linearity of Expectation Decompose: What is  $X_i$ ?

$\Pr(\omega)$	$\omega$	$X(\omega)$
1/6	1, 2, 3	3
1/6	1, 3, 2	1
1/6	2, 1, 3	1
1/6	2, 3, 1	0
1/6	3, 1, 2	0
1/6	3, 2, 1	1

## Example: Returning Homeworks

- Class with  $n$  students, randomly hand back homeworks. All permutations equally likely.
- Let  $X$  be the number of students who get their own HW
- what is  $\mathbb{E}(X)$ ?

$\Pr(\omega)$	$\omega$	$X(\omega)$
1/6	1, 2, 3	3
1/6	1, 3, 2	1
1/6	2, 1, 3	1
1/6	2, 3, 1	0
1/6	3, 1, 2	0
1/6	3, 2, 1	1

Decompose:  $X_i$  indicates if student  $i$  got their own HW back

LOE:

Conquer: What is  $\mathbb{E}(X_i)$ ?

A.  $\frac{1}{n}$  B.  $\frac{1}{n-1}$  C.  $1/2$

## Pairs with same birthday

- In a class of  $m$  students, on average how many pairs of people have the same birthday?

Decompose:

LOE:

Conquer:


# Linearity of Expectation – Even stronger

**Theorem.** For any random variables  $X_1, \dots, X_n$ , and real numbers  $a_1, \dots, a_n \in \mathbb{R}$ ,

$$\mathbb{E}(a_1X_1 + \dots + a_nX_n) = a_1\mathbb{E}(X_1) + \dots + a_n\mathbb{E}(X_n).$$

Very important: In general, we do not have  $\mathbb{E}(X \cdot Y) = \mathbb{E}(X) \cdot \mathbb{E}(Y)$

# Agenda

- Linearity of Expectation
- Indicator Random Variables
- **LOTUS** 
- Variance

## Linearity is special!

In general  $\mathbb{E}(g(X)) \neq g(\mathbb{E}(X))$

E.g.,  $X = \begin{cases} 1 & \text{with prob } 1/2 \\ -1 & \text{with prob } 1/2 \end{cases}$

$$\mathbb{E}(X^2) \neq \mathbb{E}(X)^2$$

How DO we compute  $\mathbb{E}(g(X))$ ?

# Expectation of $g(X)$

**Definition.** Given a discrete RV  $X: \Omega \rightarrow \mathbb{R}$ , the **expectation or expected value** of  $g(X)$  is

$$E[g(X)] = \sum_{\omega \in \Omega} g(X(\omega)) \cdot \Pr(\omega)$$

or equivalently

$$E[g(X)] = \sum_{x \in X(\Omega)} g(x) \cdot \Pr(X = x)$$



## Example: Returning Homeworks

- Class with 3 students, randomly hand back homeworks. All permutations equally likely.
- Let  $X$  be the number of students who get their own HW
- Let  $Y = (X^2 + 4) \bmod 8$ .
- what is  $\mathbb{E}(Y)$ ?

$\Pr(\omega)$	$\omega$	$X(\omega)$
1/6	1, 2, 3	3
1/6	1, 3, 2	1
1/6	2, 1, 3	1
1/6	2, 3, 1	0
1/6	3, 1, 2	0
1/6	3, 2, 1	1

## Rotating the table

$n$  people are sitting around a circular table. There is a name tag in each place. Nobody is sitting in front of their own name tag.

Rotate the table by a random number  $k$  of positions between 1 and  $n-1$  (equally likely).

$X$  is the number of people that end up front of their own name tag.

What is  $E(X)$ ?

Decompose:

LOE:

Conquer:

## Take Home FUN Example – Coupon Collector Problem

Say each round we get a random coupon  $X_i \in \{1, \dots, n\}$ , how many rounds (in expectation) until we have one of each coupon?

Formally: Outcomes in  $\Omega$  are sequences of integers in  $\{1, \dots, n\}$  where each integer appears at least once (+ cannot be shortened).

Example,  $n = 3$ :

$$\Omega = \{(1,2,3), (1,1,2,3), (1,2,2,3), (1,2,3), (1,1,1,3,3,3,3,3,2), \dots\}$$

$$\mathbb{P}((1,2,3)) = \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{1}{3} \quad \mathbb{P}((1,1,2,2,2,3)) = \left(\frac{1}{3}\right)^6 \quad \dots$$

## Example – Coupon Collector Problem

Say each round we get a random coupon  $X_i \in \{1, \dots, n\}$ , how many rounds (in expectation) until we have one of each coupon?

$T_i = \#$  of rounds until we have accumulated  $i$  distinct coupons

[Aka: length of the sampled  $\omega$ ]

**Wanted:**  $\mathbb{E}(T_n)$

Hard to think about  $T_n$  directly,  
Can we decompose  $T_n$  as a sum of simpler random variables?

$$Z_i = T_i - T_{i-1}$$

# of rounds needed to go from  $i - 1$  to  $i$  coupons

## Example – Coupon Collector Problem

$T_i = \#$  of rounds until we have accumulated  $i$  distinct coupons

**Wanted:**  $\mathbb{E}(T_n)$

$$Z_i = T_i - T_{i-1}$$

$$T_n = T_1 + (T_2 - T_1) + (T_3 - T_2) + \cdots + (T_n - T_{n-1}) = T_1 + Z_2 + \cdots + Z_n$$



$$\begin{aligned}\mathbb{E}(T_n) &= \mathbb{E}(T_1) + \mathbb{E}(Z_2) + \mathbb{E}(Z_3) + \cdots + \mathbb{E}(Z_n) \\ &= \underline{1 + \mathbb{E}(Z_2) + \mathbb{E}(Z_3) + \cdots + \mathbb{E}(Z_n)}\end{aligned}$$

**Wanted:**  $\mathbb{E}(Z_i)$

## Example – Coupon Collector Problem

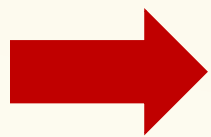
$T_i = \#$  of rounds until we have accumulated  $i$  distinct coupons

$$Z_i = T_i - T_{i-1}$$

**Wanted:**  $\mathbb{E}(Z_i)$

If we have accumulated  $i - 1$  coupons, the number  $Z_i$  of attempts needed to get the  $i$ -th coupon is **geometric** with parameter  $p = 1 - \frac{(i-1)}{n}$ .

$$\mathbb{P}_{Z_i}(1) = p \quad \mathbb{P}_{Z_i}(2) = (1 - p)p \quad \dots \quad \mathbb{P}_{Z_i}(i) = (1 - p)^{i-1}p$$



$$\underline{\mathbb{E}[Z_i] = \frac{1}{p} = \frac{n}{n - i + 1}}$$

Expectation of geometric distribution shown in last lecture, for the example #coin tosses to see first head

## Example – Coupon Collector Problem

$T_i$  = # of rounds until we have accumulated  $i$  distinct coupons

$$Z_i = T_i - T_{i-1} \quad \mathbb{E}(Z_i) = \frac{1}{p} = \frac{n}{n-i+1}$$

$$\mathbb{E}(T_n) = 1 + \mathbb{E}(Z_2) + \mathbb{E}(Z_3) + \cdots + \mathbb{E}(Z_n)$$

$$= 1 + \frac{n}{n-1} + \frac{n}{n-2} + \cdots + \frac{n}{1}$$

$$= n \cdot \left( \frac{1}{n} + \frac{1}{n-1} + \cdots + \frac{1}{2} + 1 \right) = n \cdot H_n \approx n \cdot \ln(n)$$

$n$ -th **harmonic number**

$$H_n = \sum_{i=1}^n \frac{1}{i}$$

$$\ln(n) \leq H_n \leq \ln(n) + 1$$

# Agenda

- Linearity of Expectation
- Indicator Random Variables
- LOTUS
- Variance





## Two Games

**Game 1:** In every round, you win \$2 with probability  $1/3$ , lose \$1 with probability  $2/3$ .

$W_1$  = payoff in a round of Game 1

$$\mathbb{P}(W_1 = 2) = \frac{1}{3}, \mathbb{P}(W_1 = -1) = \frac{2}{3}$$

## Two Games

**Game 1:** In every round, you win \$2 with probability  $1/3$ , lose \$1 with probability  $2/3$ .

$W_1$  = payoff in a round of Game 1

$$\mathbb{P}(W_1 = 2) = \frac{1}{3}, \mathbb{P}(W_1 = -1) = \frac{2}{3}$$

**Game 2:** In every round, you win \$10 with probability  $1/3$ , lose \$5 with probability  $2/3$ .

$W_2$  = payoff in a round of Game 2

$$\mathbb{P}(W_2 = 10) = \frac{1}{3}, \mathbb{P}(W_2 = -5) = \frac{2}{3}$$

Which game would you rather play?

## Two Games

**Game 1:** In every round, you win \$2 with probability  $1/3$ , lose \$1 with probability  $2/3$ .

$W_1$  = payoff in a round of Game 1

$$\mathbb{P}(W_1 = 2) = \frac{1}{3}, \mathbb{P}(W_1 = -1) = \frac{2}{3}$$

$$\mathbb{E}(W_1) = 0$$

**Game 2:** In every round, you win \$10 with probability  $1/3$ , lose \$5 with probability  $2/3$ .

$W_2$  = payoff in a round of Game 2

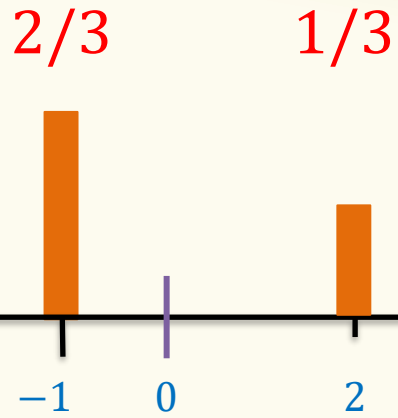
$$\mathbb{P}(W_2 = 10) = \frac{1}{3}, \mathbb{P}(W_2 = -5) = \frac{2}{3}$$

$$\mathbb{E}(W_2) = 0$$

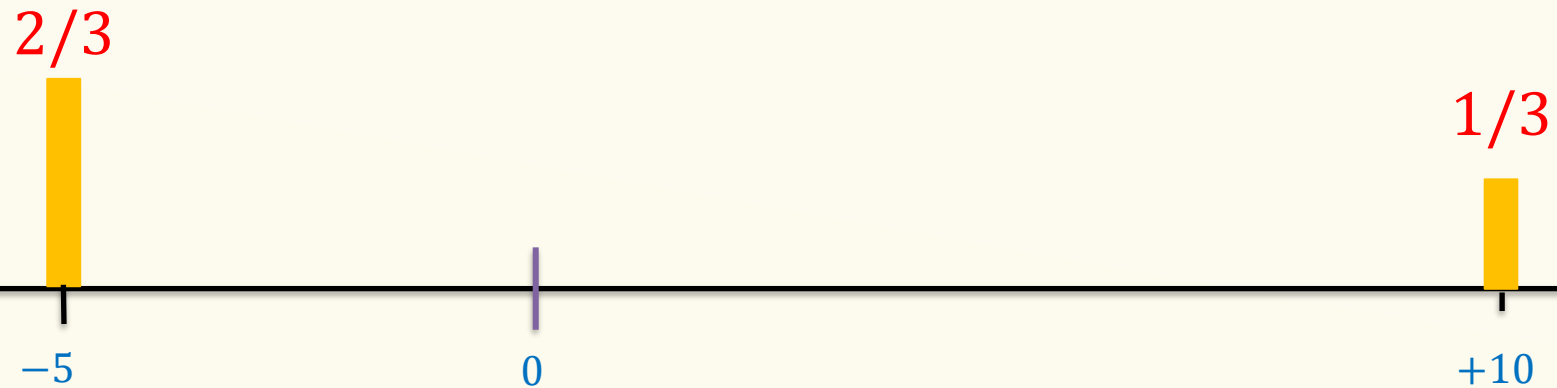
*Which game would you rather play?* Somehow, Game 2 has higher volatility!

## Two Games

$$\mathbb{P}(W_1 = 2) = \frac{1}{3}, \mathbb{P}(W_1 = -1) = \frac{2}{3}$$



$$\mathbb{P}(W_2 = 10) = \frac{1}{3}, \mathbb{P}(W_2 = -5) = \frac{2}{3}$$



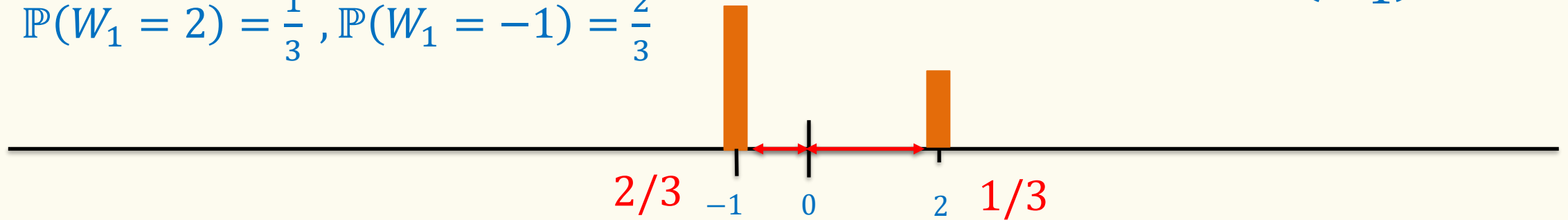
Same expectation, but clearly very different distribution.

We want to capture the difference – **New concept: Variance**

## Variance (Intuition, First Try)

$$\mathbb{E}(W_1) = 0$$

$$\mathbb{P}(W_1 = 2) = \frac{1}{3}, \mathbb{P}(W_1 = -1) = \frac{2}{3}$$



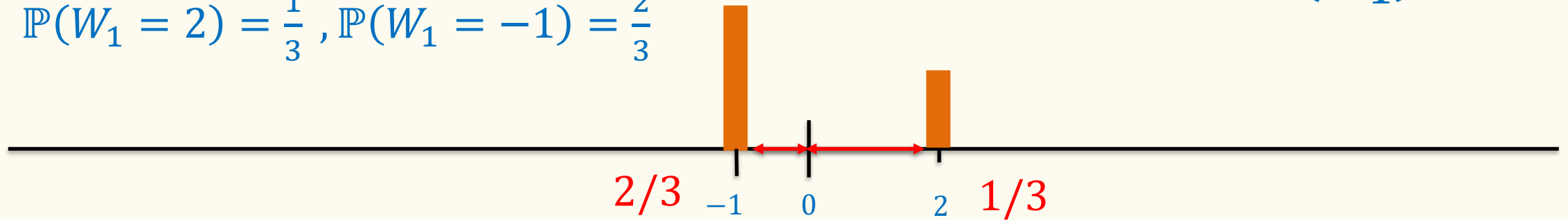
New quantity (random variable): How far from the expectation?

$$\Delta(W_1) = W_1 - E[W_1]$$

## Variance (Intuition, First Try)

$$\mathbb{E}(W_1) = 0$$

$$\mathbb{P}(W_1 = 2) = \frac{1}{3}, \mathbb{P}(W_1 = -1) = \frac{2}{3}$$



New quantity (random variable): How far from the expectation?

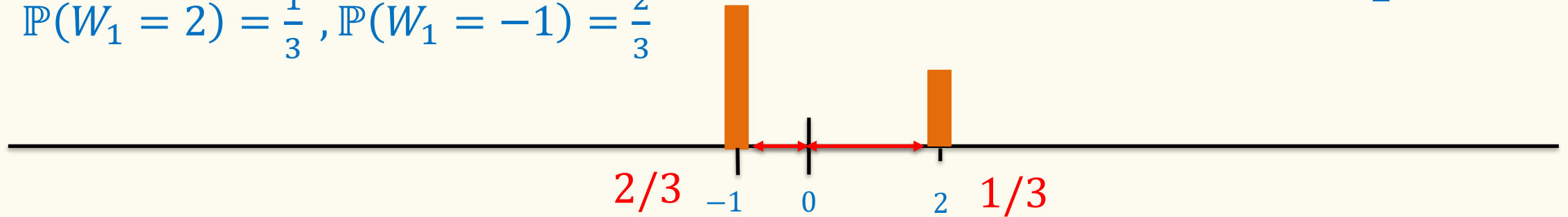
$$\Delta(W_1) = W_1 - E[W_1]$$

$$\begin{aligned} E[\Delta(W_1)] &= E[W_1 - E[W_1]] \\ &= E[W_1] - E[E[W_1]] \\ &= E[W_1] - E[W_1] \\ &= 0 \end{aligned}$$

## Variance (Intuition, Better Try)

$$\mathbb{E}(W_1) = 0$$

$$\mathbb{P}(W_1 = 2) = \frac{1}{3}, \mathbb{P}(W_1 = -1) = \frac{2}{3}$$



A better quantity (random variable): How far from the expectation?

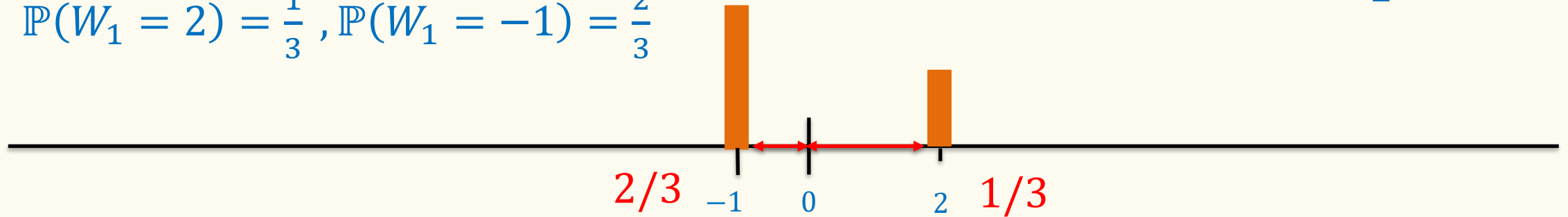
$$\Delta(W_1) = (W_1 - E[W_1])^2$$

$$E[\Delta(W_1)] = E[(W_1 - E[W_1])^2]$$

# Variance (Intuition, Better Try)

$$\mathbb{E}(W_1) = 0$$

$$\mathbb{P}(W_1 = 2) = \frac{1}{3}, \mathbb{P}(W_1 = -1) = \frac{2}{3}$$



A better quantity (random variable): How far from the expectation?

$$\Delta(W_1) = (W_1 - E[W_1])^2$$

$$\mathbb{P}(\Delta(W_1) = 1) = \frac{2}{3}$$

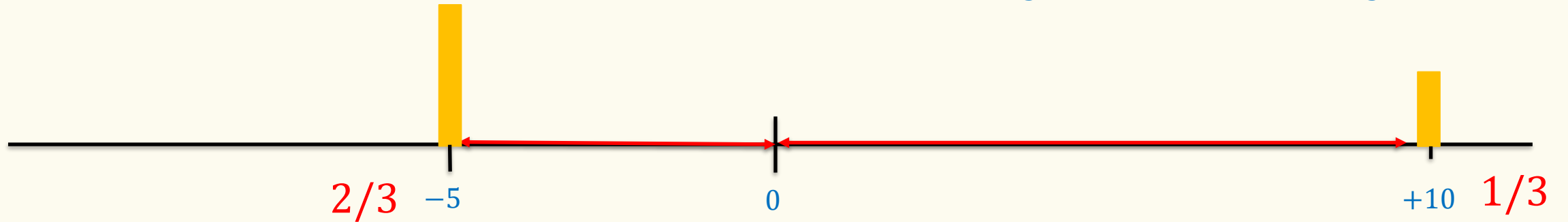
$$\mathbb{P}(\Delta(W_1) = 4) = \frac{1}{3}$$

$$\begin{aligned} E[\Delta(W_1)] &= E[(W_1 - E[W_1])^2] \\ &= \frac{2}{3} \cdot 1 + \frac{1}{3} \cdot 4 \\ &= 2 \end{aligned}$$



# Variance (Intuition, Better Try)

$$\mathbb{P}(W_2 = 10) = \frac{1}{3}, \mathbb{P}(W_2 = -5) = \frac{2}{3}$$



A better quantity (random variable): How far from the expectation?

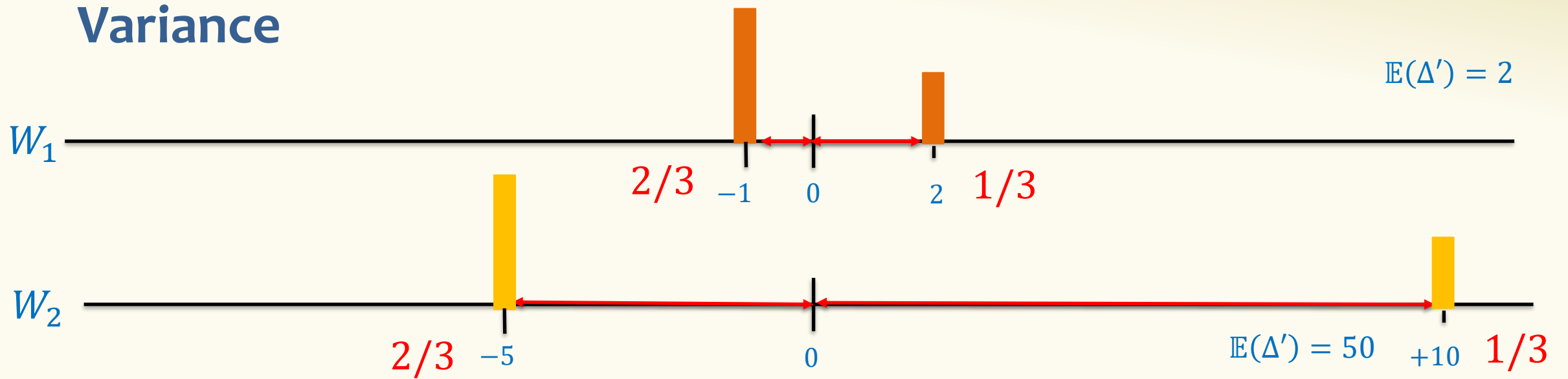
$$\Delta'(W_2) = (W_2 - E[W_2])^2$$

$$\mathbb{P}(\Delta'(W_2) = 25) = \frac{2}{3}$$

$$\mathbb{P}(\Delta'(W_2) = 100) = \frac{1}{3}$$

$$\begin{aligned} E[\Delta'(W_2)] &= E[(W_2 - E[W_2])^2] \\ &= \frac{2}{3} \cdot 25 + \frac{1}{3} \cdot 100 \\ &= 50 \end{aligned}$$

# Variance



We say that  $W_2$  has “**higher variance**” than  $W_1$ .

# Variance

**Definition.** The **variance** of a (discrete) RV  $X$  is

$$\text{Var}(X) = \mathbb{E} \left[ (X - \mathbb{E}(X))^2 \right] = \sum_x \mathbb{P}_X(x) \cdot (x - \mathbb{E}(X))^2$$

Recall  $\mathbb{E}(X)$  is a **constant**, not a random variable itself.

Intuition: Variance is a quantity that measures, in expectation, how “far” the random variable is from its expectation.

# Variance

**Definition.** The **variance** of a (discrete) RV  $X$  is

$$\text{Var}(X) = \mathbb{E} \left[ (X - \mathbb{E}(X))^2 \right] = \sum_{\mathbf{x}} \mathbb{P}_X(\mathbf{x}) \cdot (\mathbf{x} - \mathbb{E}(X))^2$$

**Standard deviation:**  $\sigma(X) = \sqrt{\text{Var}(X)}$

Recall  $\mathbb{E}(X)$  is a **constant**, not a random variable itself.

Intuition: Variance (or standard deviation) is a quantity that measures, in expectation, how “far” the random variable is from its expectation.

# Variance – Example 1

$X$  fair die

- $\mathbb{P}(X = 1) = \dots = \mathbb{P}(X = 6) = 1/6$
- $\mathbb{E}(X) = 3.5$

$\text{Var}(X) = ?$

# Variance – Example 1

$X$  fair die

- $\mathbb{P}(X = 1) = \dots = \mathbb{P}(X = 6) = 1/6$
- $\mathbb{E}(X) = 3.5$

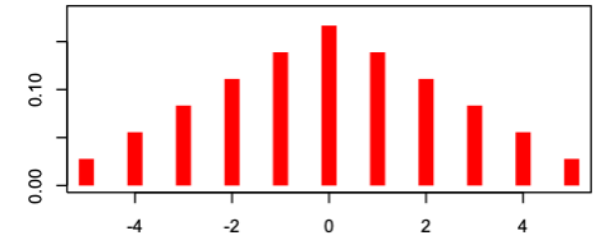
$$\begin{aligned}\text{Var}(X) &= \sum_{\mathbf{x}} \mathbb{P}(X = \mathbf{x}) \cdot (\mathbf{x} - \mathbb{E}(X))^2 \\ &= \frac{1}{6} [(1 - 3.5)^2 + (2 - 3.5)^2 + (3 - 3.5)^2 + (4 - 3.5)^2 + (5 - 3.5)^2 + (6 - 3.5)^2] \\ &= \frac{2}{6} [2.5^2 + 1.5^2 + 0.5^2] = \frac{2}{6} \left[ \frac{25}{4} + \frac{9}{4} + \frac{1}{4} \right] = \frac{35}{12} \approx 2.91677 \dots\end{aligned}$$

# Variance in Pictures

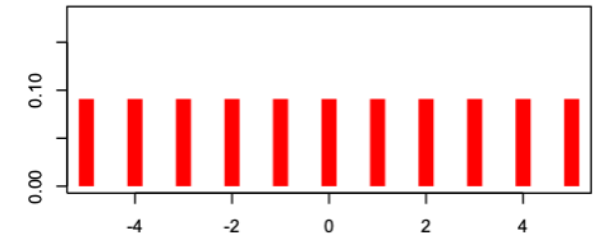
Captures how much  
“spread” there is in a pmf

All pmfs in picture  
have same expectation

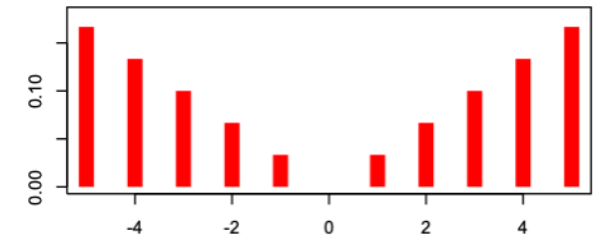
$$\sigma^2 = 5.83$$



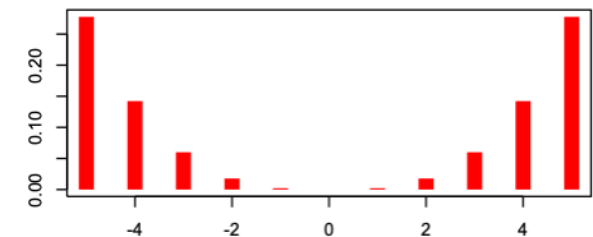
$$\sigma^2 = 10$$



$$\sigma^2 = 15$$



$$\sigma^2 = 19.7$$



# Variance – Properties

**Definition.** The **variance** of a (discrete) RV  $X$  is

$$\text{Var}(X) = \mathbb{E} \left[ (X - \mathbb{E}(X))^2 \right] = \sum_x \mathbb{P}_X(x) \cdot (x - \mathbb{E}(X))^2$$

**Theorem.** For any  $a, b \in \mathbb{R}$ ,  $\text{Var}(a \cdot X + b) = a^2 \cdot \text{Var}(X)$

(Proof: Exercise!)

**Theorem.**  $\text{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$



# Variance

**Theorem.**  $\text{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$

**Proof:**  $\text{Var}(X) = \mathbb{E} \left[ (X - \mathbb{E}(X))^2 \right]$

Recall  $\mathbb{E}(X)$  is a **constant**

$$= \mathbb{E}[X^2 - 2\mathbb{E}(X) \cdot X + \mathbb{E}(X)^2]$$

$$= \mathbb{E}(X^2) - 2\mathbb{E}(X)\mathbb{E}(X) + \mathbb{E}(X)^2$$

$$= \mathbb{E}(X^2) - \mathbb{E}(X)^2$$

(linearity of expectation!)

$\mathbb{E}(X^2)$  and  $\mathbb{E}(X)^2$   
are different !

# Variance – Example 1

$X$  fair die

- $\mathbb{P}(X = 1) = \dots = \mathbb{P}(X = 6) = 1/6$
- $\mathbb{E}(X) = \frac{21}{6}$
- $\mathbb{E}(X^2) = \frac{91}{6}$

$$\text{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \frac{91}{6} - \left(\frac{21}{6}\right)^2 = \frac{105}{36} \approx 2.91677$$

**In General,  $\text{Var}(X + Y) \neq \text{Var}(X) + \text{Var}(Y)$**

Example to show this:

- Let  $X$  be a r.v. with pmf  $\mathbb{P}(X = 1) = \mathbb{P}(X = -1) = 1/2$ 
  - What is  $\mathbb{E}[X]$  and  $\text{Var}(X)$ ?

**In General,  $\text{Var}(X + Y) \neq \text{Var}(X) + \text{Var}(Y)$**

Example to show this:

- Let  $X$  be a r.v. with pmf  $\mathbb{P}(X = 1) = \mathbb{P}(X = -1) = 1/2$ 
  - $E[X] = 0$  and  $\text{Var}(X) = 1$
- Let  $Y = -X$ 
  - What is  $E[Y]$  and  $\text{Var}(Y)$ ?

**In General,  $\text{Var}(X + Y) \neq \text{Var}(X) + \text{Var}(Y)$**

Example to show this:

- Let  $X$  be a r.v. with pmf  $\mathbb{P}(X = 1) = \mathbb{P}(X = -1) = 1/2$ 
  - $E[X] = 0$  and  $\text{Var}(X) = 1$
- Let  $Y = -X$ 
  - $E[Y] = 0$  and  $\text{Var}(Y) = 1$

What is  $\text{Var}(X + Y)$ ?