CSE 312

Foundations of Computing II

Lecture 9: Linearity of Expectation, LOTUS, and Variance

PAUL G. ALLEN SCHOOL
OF COMPUTER SCIENCE & ENGINEERING

Aleks Jovcic

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

$n=5$
$k=2$

HHTTT $\leftarrow$ $p \cdot p \cdot (1-p) \cdot (1-p) \cdot (1-p) = p^2(1-p)^3$
THTTH $\leftarrow$ $(1-p)p(1-p)(1-p)p = p^2(1-p)^3$

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 ?

$\mathbb{P}(Z=k) = \{$ $k=0$, $k=1$, …

$$\mathbb{P}(Z=k) = \binom{n}{k} p^k (1-p)^{n-k}$$

Expectation of Random Variable

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

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

or equivalently

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

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 $\mathbb{E}(Z)$?

$$E[Z] = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k}$$

The brute force method

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

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

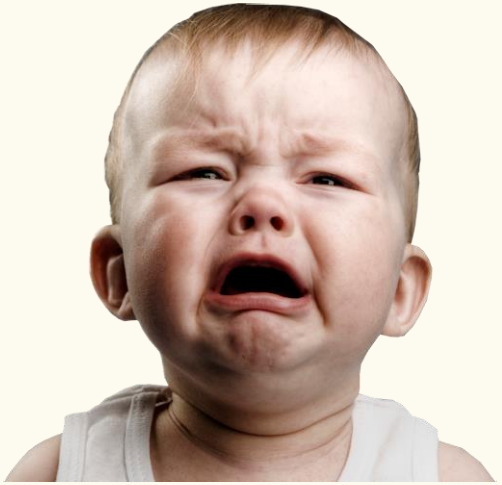
$$\mathbb{E}[Z] = \sum_{k=0}^{n} k \cdot P(Z=k) \quad = \sum_{k=0}^{n} k \cdot \binom{n}{k} p^k (1-p)^{n-k}$$

$$= \sum_{k=0}^{n} k \cdot \frac{n!}{k!\,(n-k)!}\, p^k (1-p)^{n-k} \quad = \sum_{k=1}^{n} \frac{n!}{(k-1)!\,(n-k)!}\, p^k (1-p)^{n-k}$$

$$= np \sum_{k=1}^{n} \frac{(n-1)!}{(k-1)!\,(n-k)!}\, p^{k-1} (1-p)^{n-k}$$

$$= np \sum_{k=0}^{n-1} \frac{(n-1)!}{k!\,(n-1-k)!}\, p^k (1-p)^{(n-1)-k}$$

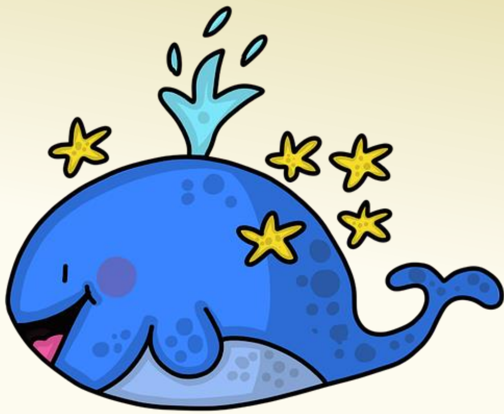
$$= np \sum_{k=0}^{n-1} \binom{n-1}{k} p^k (1-p)^{(n-1)-k} = np\big(p + (1-p)\big)^{n-1} = np \cdot 1 = \boxed{np}$$



$n=3$, $p=0.5$, $np=1.5$

$n=2$, $p=0.5$, $np=1$

Linearity of Expectation (Idea)



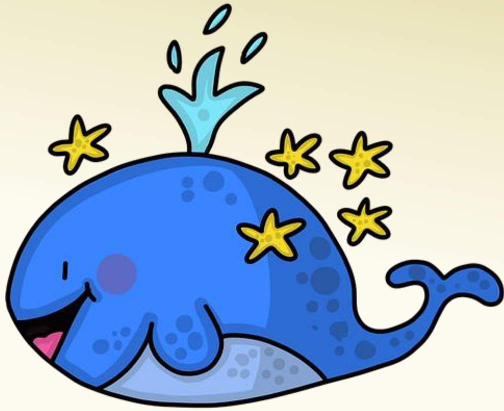
Let's say you and your friend sell fish for a living.

- 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?

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

Linearity of Expectation (Idea)



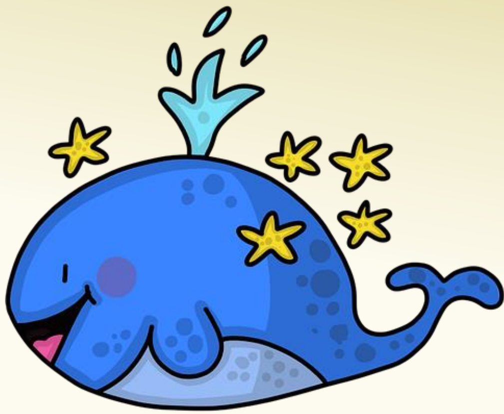
Let's say you and your friend sell fish for a living.

- 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

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?

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

Linearity of Expectation – Proof

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

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

$$\mathbb{E}(X+Y)=\sum_{\omega} P(\omega)(X(\omega)+Y(\omega))$$

$$=\sum_{\omega} P(\omega)X(\omega)+\sum_{\omega} P(\omega)Y(\omega)$$

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

Linearity of Expectation

$E(XY) \stackrel{?}{=} E(X)E(Y)$

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 $X_1, \dots, X_n$,

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

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

$$= \mathbb{E}(X_1 + \cdots + X_{n-1}) + \mathbb{E}(X_n) = \cdots$$

Coin flipping again

$C = A + B$

$n = 5$

Z_i = # of heads in i^{th} coin flip

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 $\mathbb{E}(Z)$?

$Z = Z_1 + Z_2 + \dots + Z_n$

$E[Z] = E[Z_1 + \dots + Z_n]$

$= E[Z_1] + E[Z_2] + \dots + E[Z_n]$

HTHTT

$Z = 2$

$Z_1 = 1$, $Z_2 = 0$, $Z_3 = 1$, $Z_4 = 0$, $Z_5 = 0$

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

$\Pr(Z_i = k) = \begin{cases} p & k = 1 \\ 1-p & k = 0 \end{cases}$

$p + p + \dots p = np$

Example – Coin Tosses

we flip n coins, each one heads with probability p

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

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

Fact. $Z = X_1 + \cdots + X_n$

Linearity of Expectation:

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

$\mathbb{P}(X_i = 1) = p$
$\mathbb{P}(X_i = 0) = 1 - p$

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

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 $Z = Z_1 + \cdots Z_n$
$$X = X_1 + \cdots + X_n$$
- 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

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

$\mathbb{P}(X = 1) = \mathbb{P}(\mathrm{A})$
$\mathbb{P}(X = 0) = 1 - \mathbb{P}(\mathrm{A})$

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

$Z_i = 1$ if the i^{th} coin flip is heads, 0 otherwise

$E[Z_i] = \mathbb{P}(\quad) = \mathbb{P}$

Example: Returning Homeworks

LOE + indicators

- Class with n students, randomly hand back homeworks. All permutations equally likely.

$X_i = 1$ if the i^{th} student gets their HW

- Let X be the number of students who get their own HW
- what is $\mathbb{E}(X)$?

Decompose: $X = X_1 + X_2 + \dots + X_n$

$n=3$ $E(X) = 1$

$\Pr(\omega)$	ω	$X(\omega)$
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

LOE: $E[X] = E[X_1 + \dots X_n] = E[X_1] + \dots E[X_n]$

Conquer: $E[X_i] = \mathbb{P}(X_i = 1) = \frac{1}{n}$

$E[X] = \frac{1}{n} \cdot n = \boxed{1}$

$\frac{|E|}{|\Omega|} = \frac{(n-1)!}{n!} = \frac{1}{n}$

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?

$\Pr(\omega)$	ω	$X(\omega)$
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 $\mathbb{E}(X)$?

$\Pr(\omega)$	ω	$X(\omega)$
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. ½

Pairs with same birthday

$(A, B) = (B, A)$ $\quad \frac{1}{2} m \cdot (m-1)$ $\quad \frac{365}{365 \cdot 365}$

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

$X_i = 1$ if pair shares a bday

$n = \#$ of pairs $= \binom{m}{2}$

Decompose: $X = X_1 + \dots X_n$

LOE: $E[X] = E[X_1] + \dots + E[X_n]$

$E[X] = E[X_i] \cdot n = \frac{\binom{m}{2}}{365}$

Conquer: $E[X_i] = \mathbb{P}(X_i = 1) = \frac{1}{365}$

Linearity of Expectation – Even stronger

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

$$\mathbb{E}(a_1 X_1 + \cdots + a_n X_n) = a_1 \mathbb{E}(X_1) + \cdots + a_n \mathbb{E}(X_n).$$

$E[aX] = aE[X]$ $\quad E[X+b] = E[X]+b$

$E[X] + E[b]$

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., $X = \begin{cases} 1 \;\; with\, prob \;\; 1/2 \\ -1 \;\; with\, prob\; 1/2 \end{cases}$

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

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

Expectation of $g(X)$

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

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

or equivalently

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

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 $Y = (X^2 + 4) \ mod \ 8.$
- what is $\mathbb{E}(Y)$?

$\Pr(\omega)$	ω	$X(\omega)$
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 $X_i \in \{1, \ldots, n\}$, how many rounds (in expectation) until we have one of each coupon?

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

Example, $n = 3$:

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

$$\mathbb{P}\big((1,2,3)\big) = \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{1}{3} \qquad \mathbb{P}\big((1,1,2,2,2,3)\big) = \left(\frac{1}{3}\right)^6 \qquad \ldots$$

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 ω]

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: $\mathbb{E}(T_n)$

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

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

⬇

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

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 $p = 1 - \frac{(i-1)}{n}$.

$\mathbb{p}_{Z_i}(1) = p \qquad \mathbb{p}_{Z_i}(2) = (1-p)p \qquad \cdots \qquad \mathbb{p}_{Z_i}(i) = (1-p)^{i-1}p$

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

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

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

n-th **harmonic number**

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

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

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

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

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

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

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

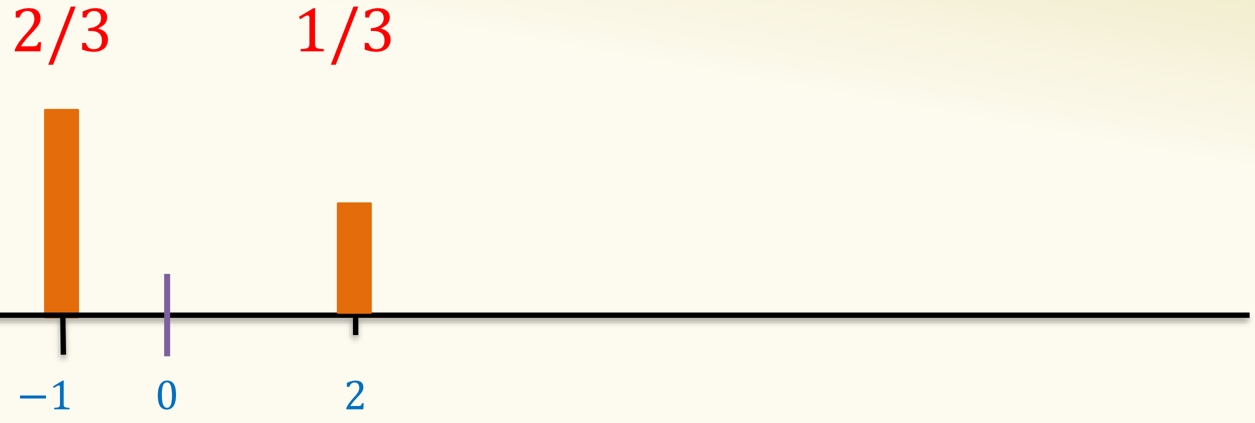
$\mathbb{E}(W_2) = 0$

Which game would you <u>rather</u> play?

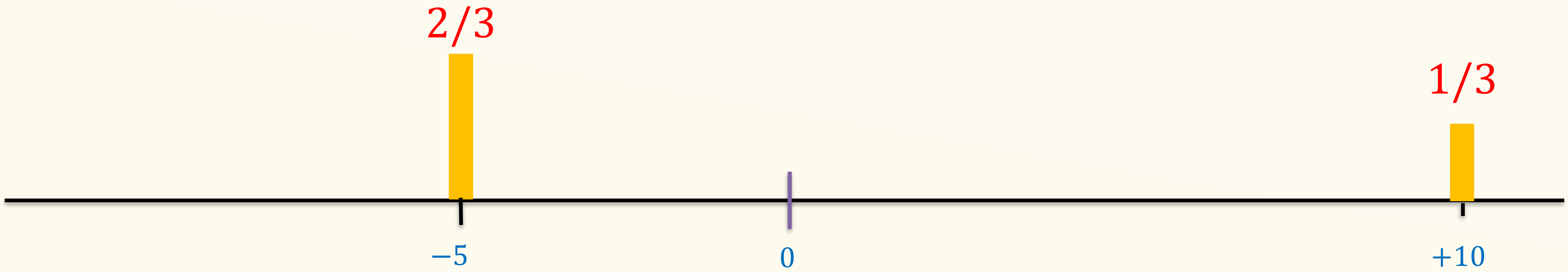
Somehow, Game 2 has higher volatility!

Two Games

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



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



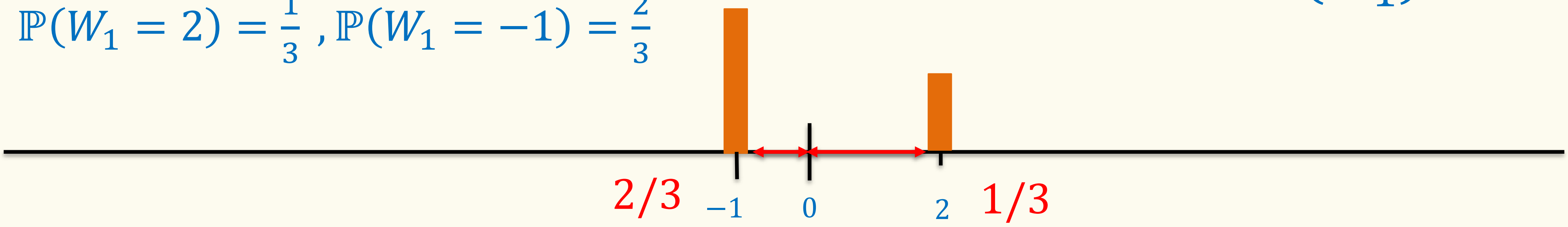
Same expectation, but clearly very different distribution.

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

Variance (Intuition, First Try)

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

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



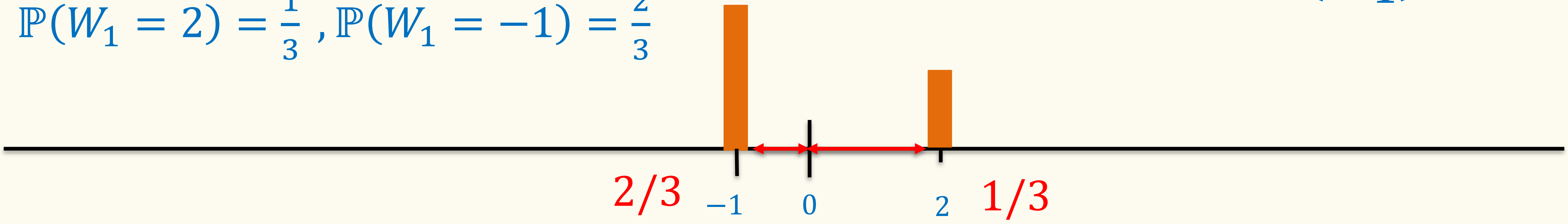
New quantity (random variable): How far from the expectation?

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

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?

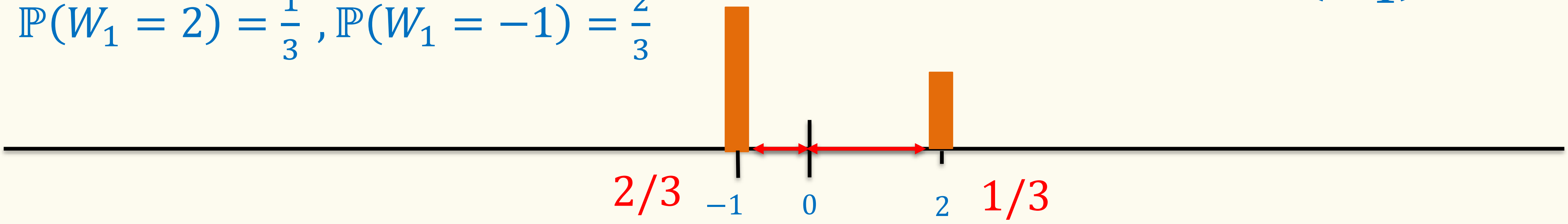
$\Delta(W_1) = W_1 - E[W_1]$

$$E[\Delta(W_1)] = E[W_1 - E[W_1]]$$
$$= E[W_1] - E[E[W_1]]$$
$$= E[W_1] - E[W_1]$$
$$= 0$$

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?

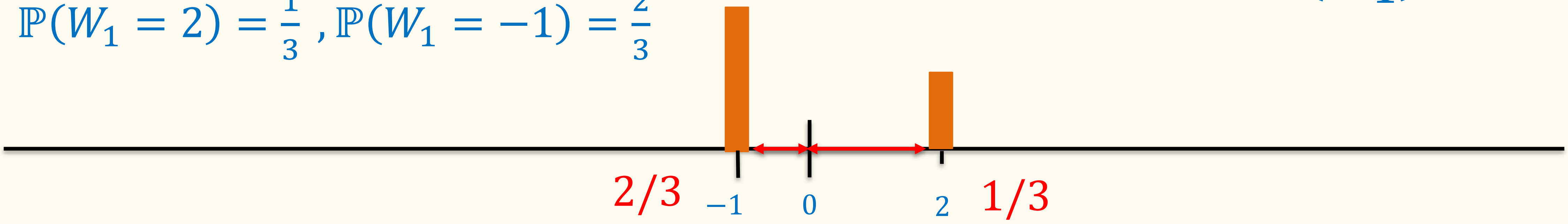
$\Delta(W_1) = (W_1 - E[W_1])^2$

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

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?

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

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

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

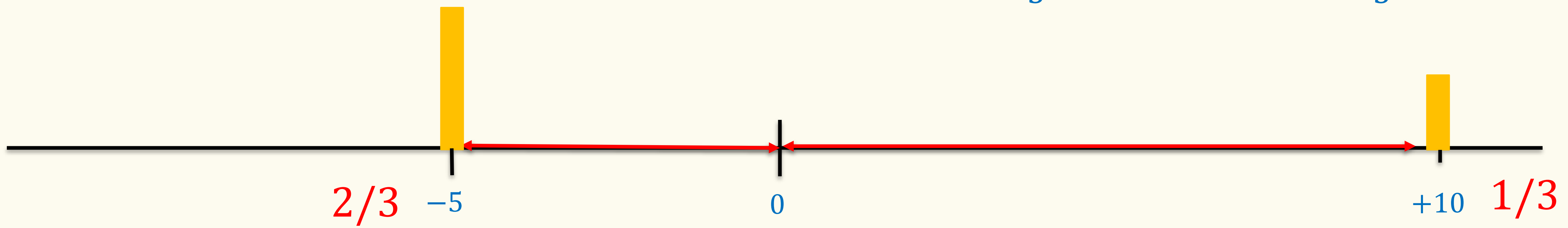
$$E[\Delta(W_1)] = E[(W_1 - E[W_1])^2]$$

$$= \frac{2}{3} \cdot 1 + \frac{1}{3} \cdot 4$$

$$= 2$$

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?

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

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

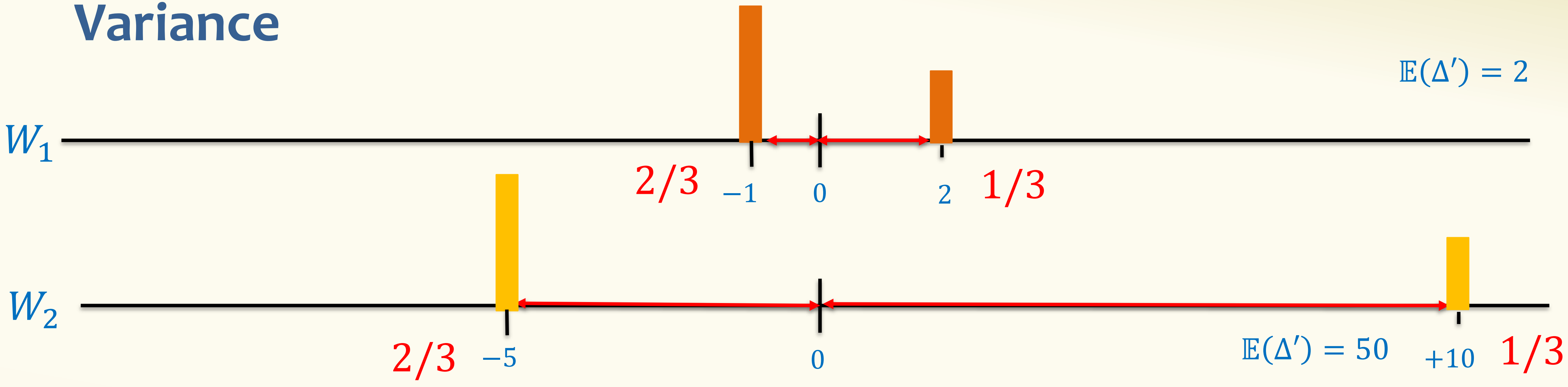
$$\mathbb{P}(\Delta'(W_2) = 100) = \frac{1}{3}$$

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

$$= \frac{2}{3} \cdot 25 + \frac{1}{3} \cdot 100$$

$$= 50$$

Variance



We say that W_2 has **"higher variance"** than W_1.

Variance

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

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

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

<u>Intuition:</u> 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

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

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

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

<u>Intuition:</u> 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

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

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

Variance – Example 1

X fair die

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

$$\mathrm{Var}(\mathrm{X}) = \sum_x \mathbb{P}(X = x) \cdot \big(x - \mathbb{E}(X)\big)^2$$

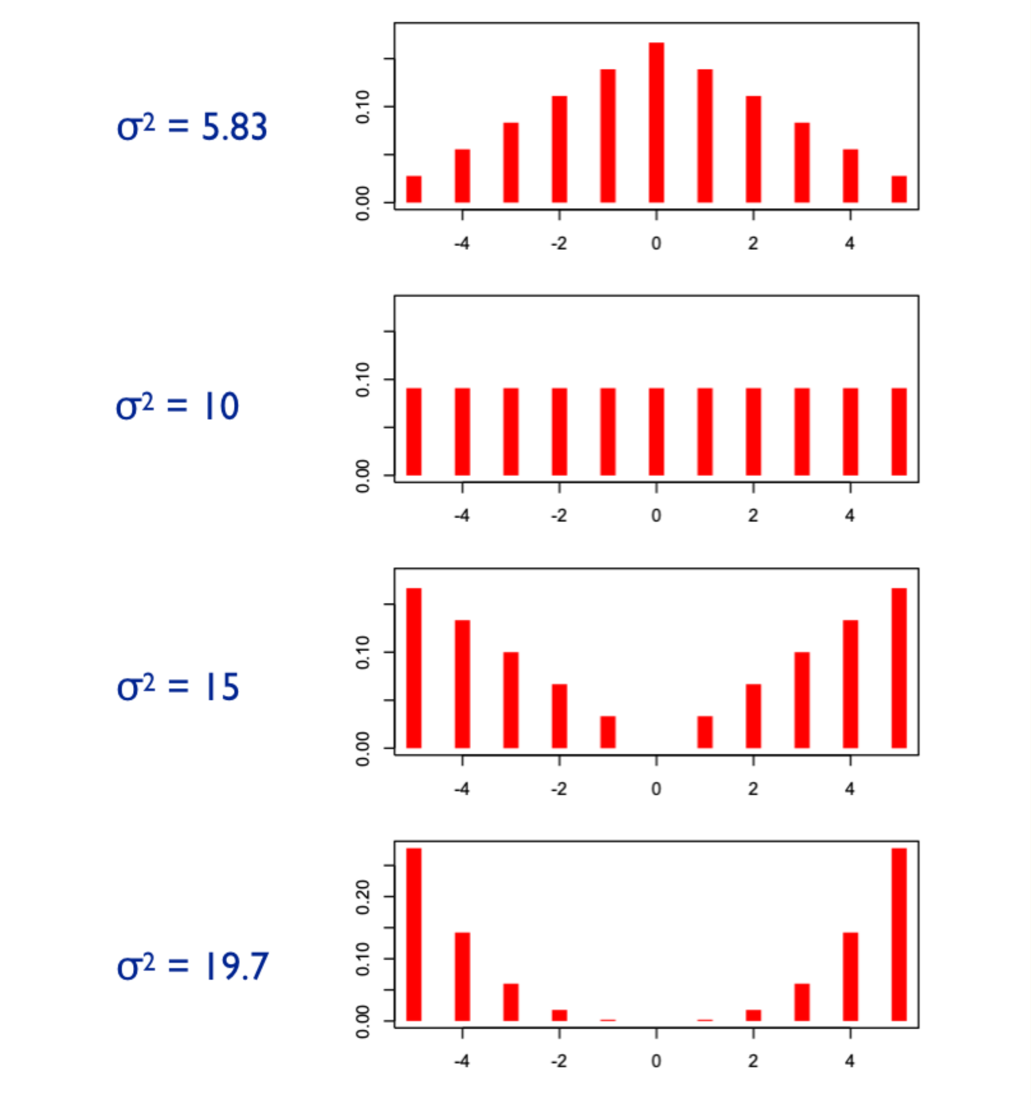
$$= \frac{1}{6}[(1-3.5)^2 + (2-3.5)^2 + (3-3.5)^2 + (4-3.5)^2 + (5-3.5)^2 + (6-3.5)^2]$$

$$= \frac{2}{6}[2.5^2 + 1.5^2 + 0.5^2] = \frac{2}{6}\left[\frac{25}{4} + \frac{9}{4} + \frac{1}{4}\right] = \frac{35}{12} \approx 2.91677\ldots$$

Variance in Pictures

Captures how much "spread' there is in a pmf

All pmfs in picture
have same expectation



Variance – Properties

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

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

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

(Proof: Exercise!)

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

Variance

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

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

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

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

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

(linearity of expectation!)

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

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

Variance – Example 1

X fair die

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

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

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

Example to show this:

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

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

Example to show this:

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

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

Example to show this:

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

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