CSE 312

Foundations of Computing II

Lecture 11: Variance and Independence of RVs

PAUL G. ALLEN SCHOOL OF COMPUTER SCIENCE & ENGINEERING

Rachel Lin, Hunter Schafer

Slide Credit: Based on Stefano Tessaro's slides for 312 19au
incorporating ideas from Alex Tsun's and Anna Karlin's slides for 312 20su and 20au

Recap Linearity of Expectation

Theorem. For any two random variables X and Y *(X, Y do not need to be independent)*

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

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

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

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

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

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

Recap Linearity is special!

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

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

- $\mathbb{E}(XY) \neq \mathbb{E}(X)\mathbb{E}(Y)$
- $\mathbb{E}(X/Y) \neq \mathbb{E}(X)/\mathbb{E}(Y)$
- $\mathbb{E}(X^2) \neq \mathbb{E}(X)^2$

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

Recap Expectation of $g(X)$

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

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

or equivalently

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

Example: Expectation of $g(X)$

Suppose we rolled a fair, 6-sided die in a game. You will win the square number rolled dollars, times 10. Let X be the result of the dice roll. What is your expected winnings?

$E[10X^2] =$

Agenda

- Variance
- Properties of Variance
- Independent Random Variables
- Properties of Independent Random Variables

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

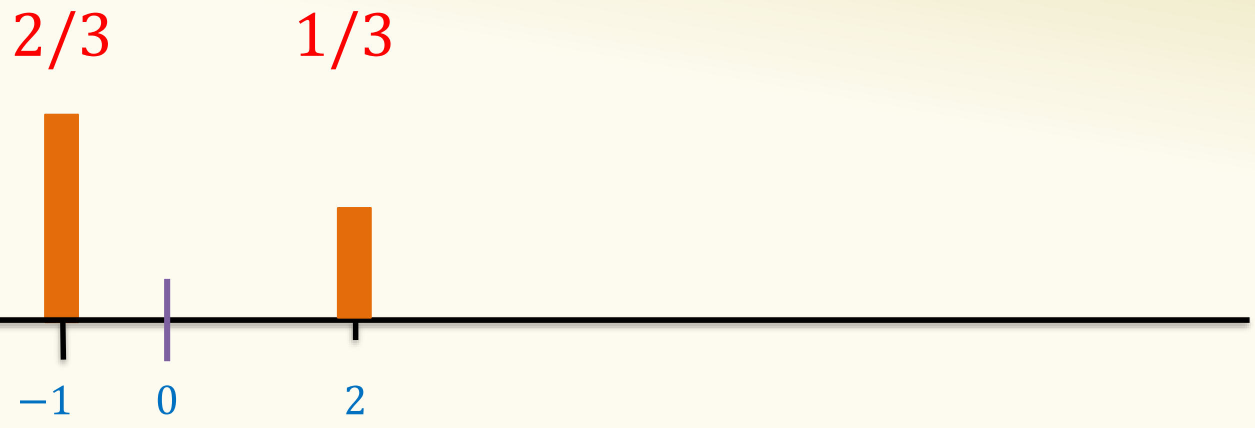
$\mathbb{P}(W_2 = 10) = \frac{1}{3}$, $\mathbb{P}(W_2 = -5) = \frac{2}{3}$

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

Which game would you rather play?

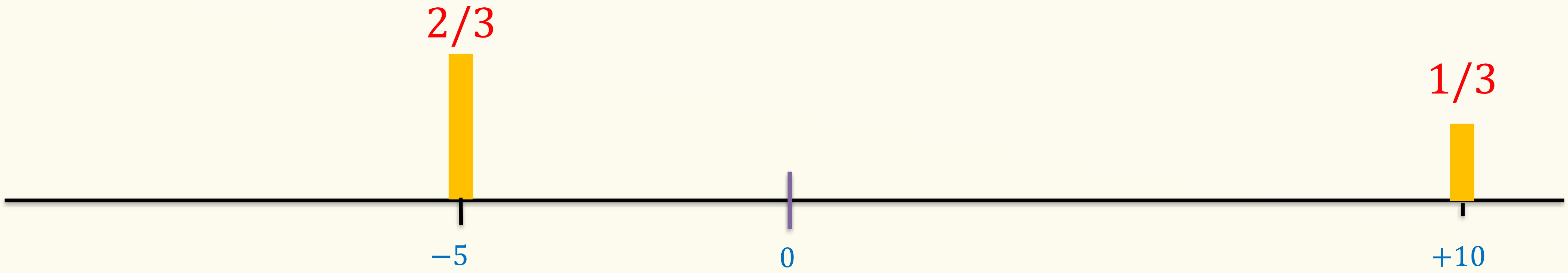
Somehow, Game 2 has higher volatility / exposure!

Two Games



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

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



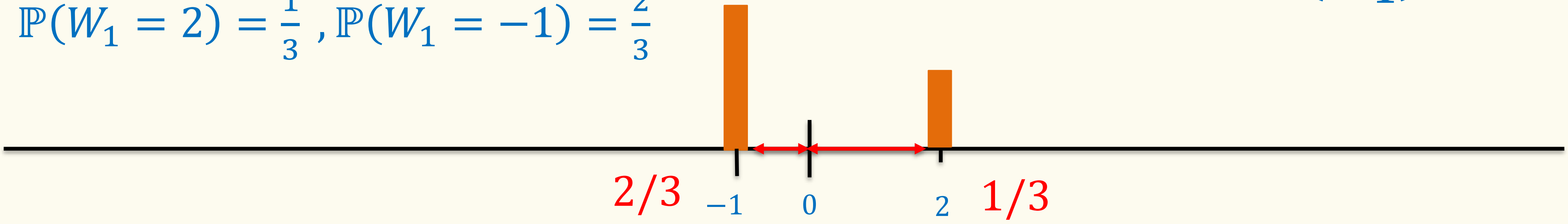
Same expectation, but clearly very different distribution.

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?

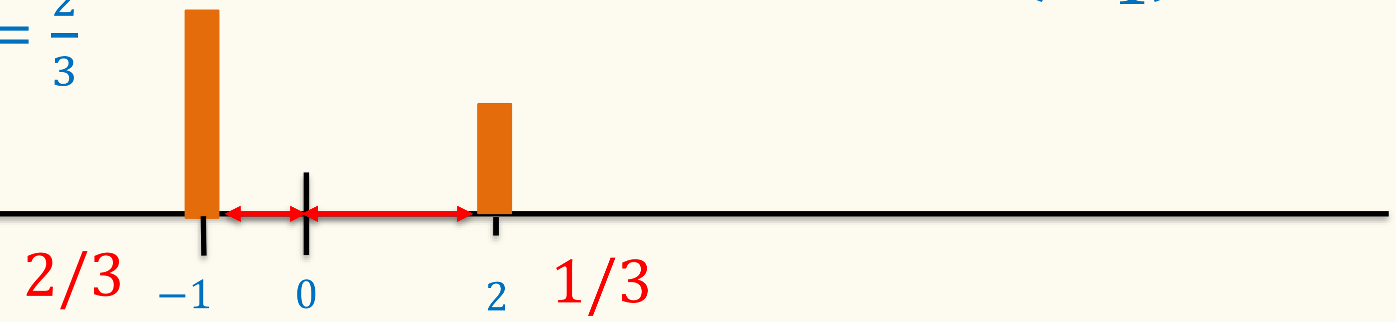
$\Delta(W_1) = W_1 - E[W_1]$

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

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?

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

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

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

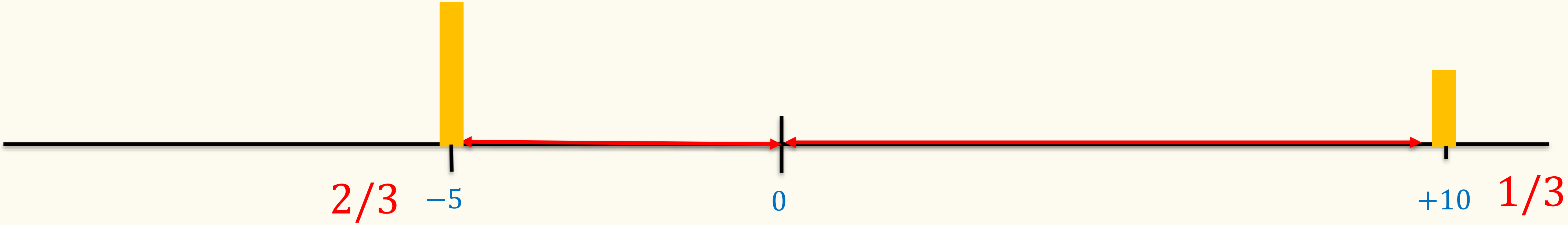
$$E[\Delta(W_1)] = E[(W_1 - E[W_1])^2]$$

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

$$= 2$$

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?

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

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

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

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

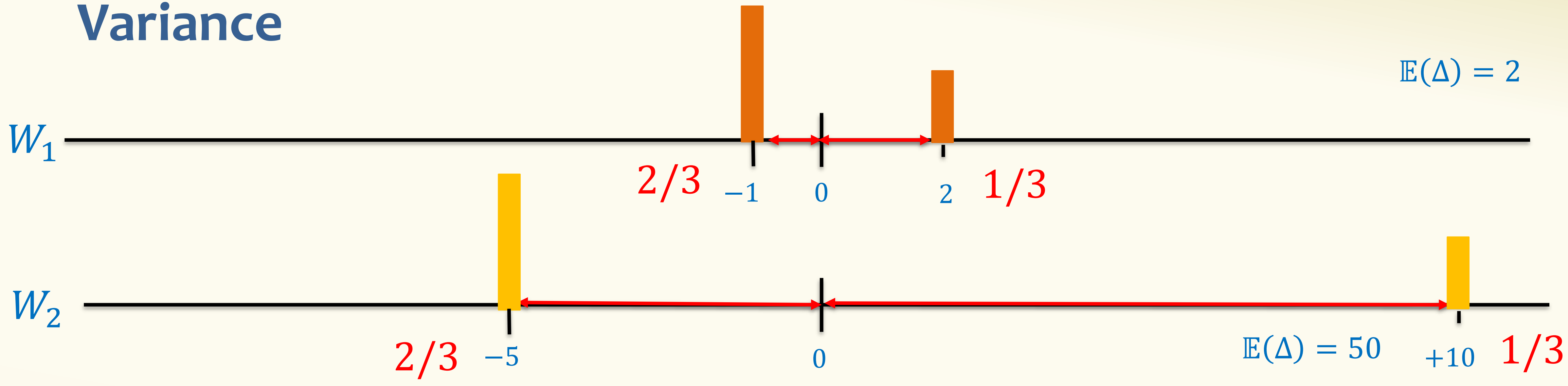
$$= \frac{2}{3} \cdot 25 + \frac{1}{3} \cdot 100$$

$$= 50$$

Poll:
pollev.com/hunter312

A. 0
B. 20/3
C. 50
D. 2500

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.

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

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

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

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

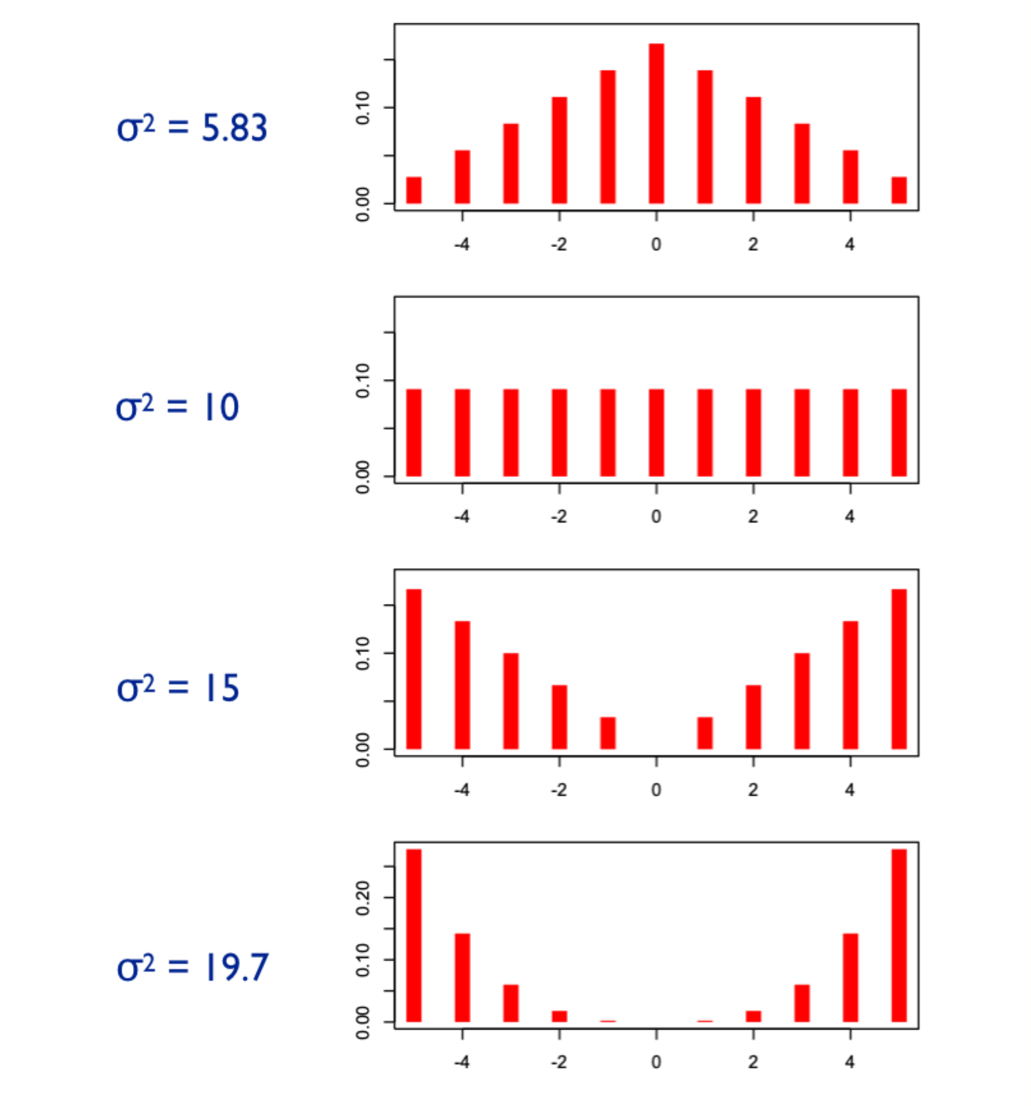
$$= \frac{1}{6}[(1 - 3.5)^2 + (2 - 3.5)^2 + (3 - 3.5)^2 + (4 - 3.5)^2 + (5 - 3.5)^2 + (6 - 3.5)^2]$$

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

Variance in Pictures

Captures how much "spread' there is in a pmf

All pmfs have same expectation



Agenda

- Variance
- Properties of Variance
- Independent Random Variables
- Properties of Independent Random Variables

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

Proof by counter-example:

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

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

Brain Break

Agenda

- Variance
- Properties of Variance
- Independent Random Variables
- Properties of Independent Random Variables

Random Variables and Independence

Definition. Two random variables X, Y are **(mutually) independent** if for all x, y,

$$\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x) \cdot \mathbb{P}(Y = y)$$

Intuition: Knowing X doesn't help you guess Y and vice versa

Definition. The random variables $X_1, \dots, X_n$ are **(mutually) independent** if for all $x_1, \dots, x_n$,

$$\mathbb{P}(X_1 = x_1, \dots, X_n = x_n) = \mathbb{P}(X_1 = x_1) \cdots \mathbb{P}(X_n = x_n)$$

Note: No need to check for all subsets, but need to check for all outcomes!

Example

Let X be the number of heads in n *independent* coin flips of the same coin. Let $Y = \mathrm{X} \bmod 2$ be the parity (even/odd) of X.

Are X and Y independent?

Poll:

pollev.com/hunter312

A. Yes
B. No

Example

Make $2n$ *independent* coin flips of the same coin. Let X be the number of heads in the first n flips and Y be the number of heads in the last n flips.

Are X and Y independent?

Poll:
pollev.com/hunter312

A. Yes
B. No

Agenda

- Variance
- Properties of Variance
- Independent Random Variables
- Properties of Independent Random Variables

Important Facts about Independent Random Variables

Theorem. If X, Y independent, $\mathbb{E}(X \cdot Y) = \mathbb{E}(X) \cdot \mathbb{E}(Y)$

Theorem. If X, Y independent, $\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$

Corollary. If $X_1, X_2, \ldots, X_n$ mutually independent,

$$\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i}^{n} \mathrm{Var}(X_i)$$

(Not Covered) Proof of $\mathbb{E}(X \cdot Y) = \mathbb{E}(X) \cdot \mathbb{E}(Y)$

Theorem. If X, Y independent, $\mathbb{E}(X \cdot Y) = \mathbb{E}(X) \cdot \mathbb{E}(Y)$

Proof

Let $x_i, y_i, i = 1, 2, \ldots$ be the possible values of X, Y.

$$\begin{aligned} E[X \cdot Y] &= \sum_i \sum_j x_i \cdot y_j \cdot P(X = x_i \wedge Y = y_j) \\ &= \sum_i \sum_j x_i \cdot y_j \cdot P(X = x_i) \cdot P(Y = y_j) \\ &= \sum_i x_i \cdot P(X = x_i) \cdot \left(\sum_j y_j \cdot P(Y = y_j) \right) \\ &= E[X] \cdot E[Y] \end{aligned}$$

independence

Note: *NOT* true in general; see earlier example $E[X^2] \neq E[X]^2$

(Not Covered) Proof of $\text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y)$

Theorem. If X, Y independent, $\text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y)$

Proof

$$Var[X+Y]$$

$$= E[(X+Y)^2] - (E[X+Y])^2$$

$$= E[X^2 + 2XY + Y^2] - (E[X] + E[Y])^2$$

$$= E[X^2] + 2E[XY] + E[Y^2] - ((E[X])^2 + 2E[X]E[Y] + (E[Y])^2)$$

$$= E[X^2] - (E[X])^2 + E[Y^2] - (E[Y])^2 + 2(E[XY] - E[X]E[Y])$$

$$= Var[X] + Var[Y] + 2(E[X]E[Y] - E[X]E[Y])$$

$$= Var[X] + Var[Y]$$

Example – Coin Tosses

We flip n *independent* coins, each one heads with probability p

- $X_i = \begin{cases} 1, & i\text{–th outcome is heads} \\ 0, & i\text{–th outcome is tails.} \end{cases}$
- $Z =$ number of heads

Fact. $Z = \sum_{i=1}^{n} X_i$

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

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

What is $\mathrm{E}[Z]$? What is $\mathrm{Var}(Z)$?

Note: $X_1, \dots, X_n$ are mutually independent! [Verify it formally!]

$$\mathrm{Var}(Z) = \sum_{i=1}^{n} \mathrm{Var}(X_i) = n \cdot p(1-p)$$

Note $\mathrm{Var}(X_i) = p(1-p)$