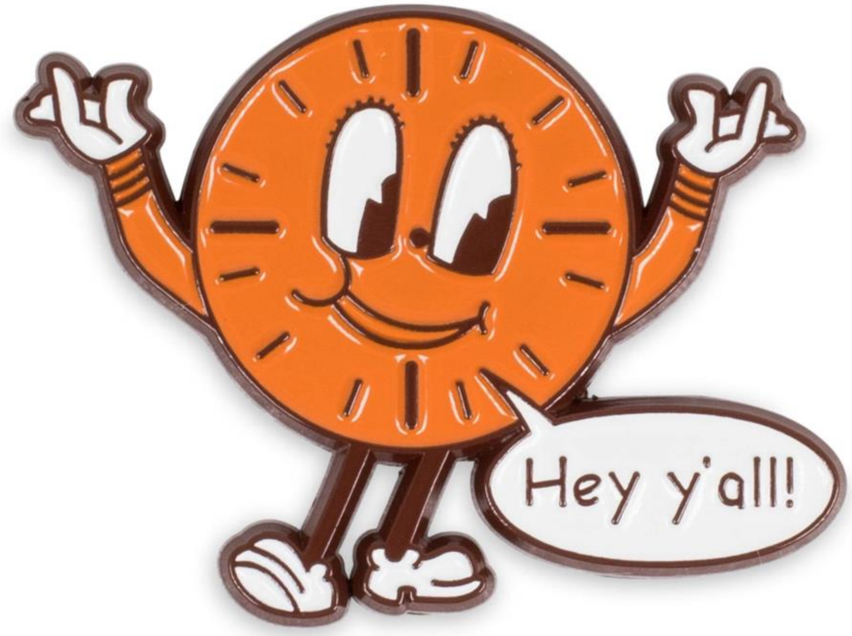
No activity slide today!



Variance & Independence of RVs

CSE 312 Summer 21
Lecture 11

Announcements

Annotations not saved in Wednesday's lecture 😢

Problem Set 4 and Review Summary 2 released.

Please check the course webpage / calendar for deadlines!

- Real World Mini-project 1 – next Wednesday
- Problem Set 4 – next Thursday
- Review Summary 2 – July 30th

Variance

A gamble to think about

Would you play either of these games?

I have a (biased) coin which flips heads with probability 1/3.

<u>Game 1:</u> I will flip the biased coin

- If a heads comes up, you win $2
- If a tails comes up, you pay me $1

<u>Game 2:</u> I will flip the biased coin

- If a heads comes up, you win $10
- If a tails comes up, you pay me $5

You are probably thinking about the average profit you get from these, or if you can get a higher payout than me

A gamble to think about

Would you play either of these games?

<u>Game 1:</u> Win \$2, lose \$1

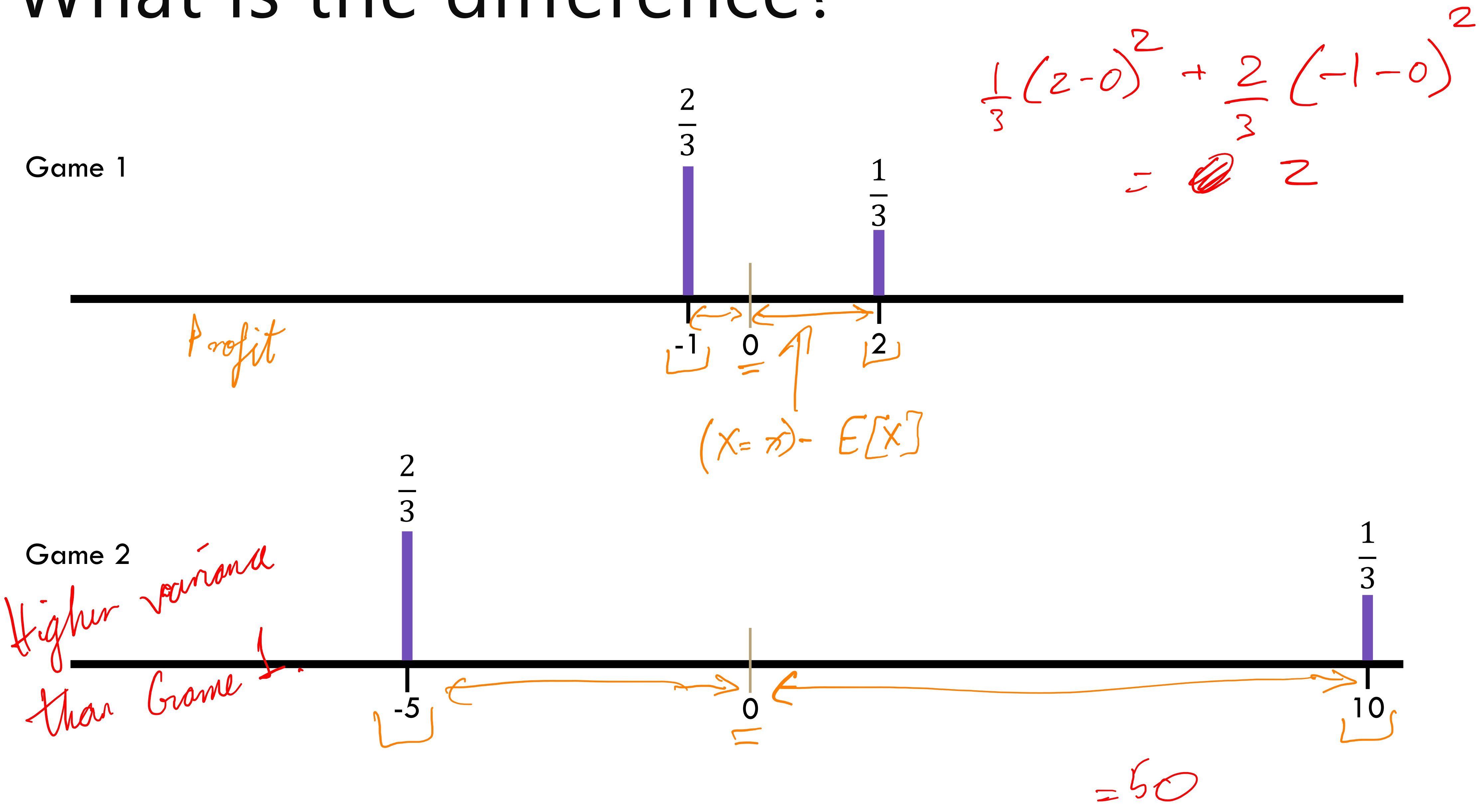
<u>Game 2:</u> Win \$10, lose \$5

Use random variables X_1 and X_2 to represent your profit from playing game 1 and game 2 respectively and then find the expected profit.

$$\mathbb{E}[X_1] = \mathbb{E}[X_2] = 0$$

Both games have an average profit of \$0. But, somehow one of the two is more enticing to some of you. Even with a \$0 average profit you believe that one of the games is more volatile.

What is the difference?



What is the difference?

Expectation tells you what the average will be…

But it doesn’t tell you how “extreme” your results could be.

Nor how likely those extreme results are.

Game 2 has more extreme results when compared to Game 1.

In expectation they “cancel out” but if you can only play once…

…it would be nice to measure that.

Designing a Measure – Try 1

Well let's measure how far all the events are away from the center, and how likely they are

$$\sum_\omega \mathbb{P}(\omega) \cdot (X(\omega) - \mathbb{E}[X])$$

What happens with Game 1?

$$\frac{1}{3} \cdot (2 \ - 0) + \frac{2}{3} \cdot (-1 - 0)$$

$$\frac{2}{3} - \frac{2}{3} = 0$$

What happens with Game 2?

$$\frac{1}{3} \cdot (10 \ - 0) + \frac{2}{3} \cdot (-5 - 0)$$

$$\frac{10}{3} - \frac{10}{3} = 0$$

Designing a Measure – Try 2

How do we prevent cancelling? Squaring makes everything positive.

$$\sum_\omega \mathbb{P}(\omega) \cdot (X(\omega) - \mathbb{E}[X])^2$$

What happens with Game 1?

$$\frac{1}{3} \cdot (2 \; - 0)^2 + \frac{2}{3} \cdot (-1 - 0)^2$$

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

What happens with Game 2?

$$\frac{1}{3} \cdot (10 \; - 0)^2 + \frac{2}{3} \cdot (-5 - 0)^2$$

$$\frac{100}{3} + \frac{50}{3} = 50$$

We say that X_2 has "higher variance" than X_1.

Std dev(X) = $\sqrt{Var(X)}$

Why Squaring?

Why not absolute value? Or Fourth power?

Squaring is nicer algebraically.

Our goal with variance was to talk about the spread of results. Squaring makes extreme results even more extreme.

Fourth power over-emphasizes the extreme results (for our purposes).

Variance

Variance

The variance of a random variable X is

$$\mathrm{Var}(X) = \sum_{\omega} \mathbb{P}(\omega) \cdot (X(\omega) - \mathbb{E}[X])^2 = \mathbb{E}[(X - \mathbb{E}[X])^2]$$

$\mathbb{E}[g(X)]$

$g(X)$

The first forms are the definition.

We can simplify the formula with an algebra trick.

$X^2 \to$ Do experiment, get X, square each ~~outcome~~ value

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

X^2 is a rv

Variance and Standard Deviation

Standard Deviation

The standard deviation of a random variable X is

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

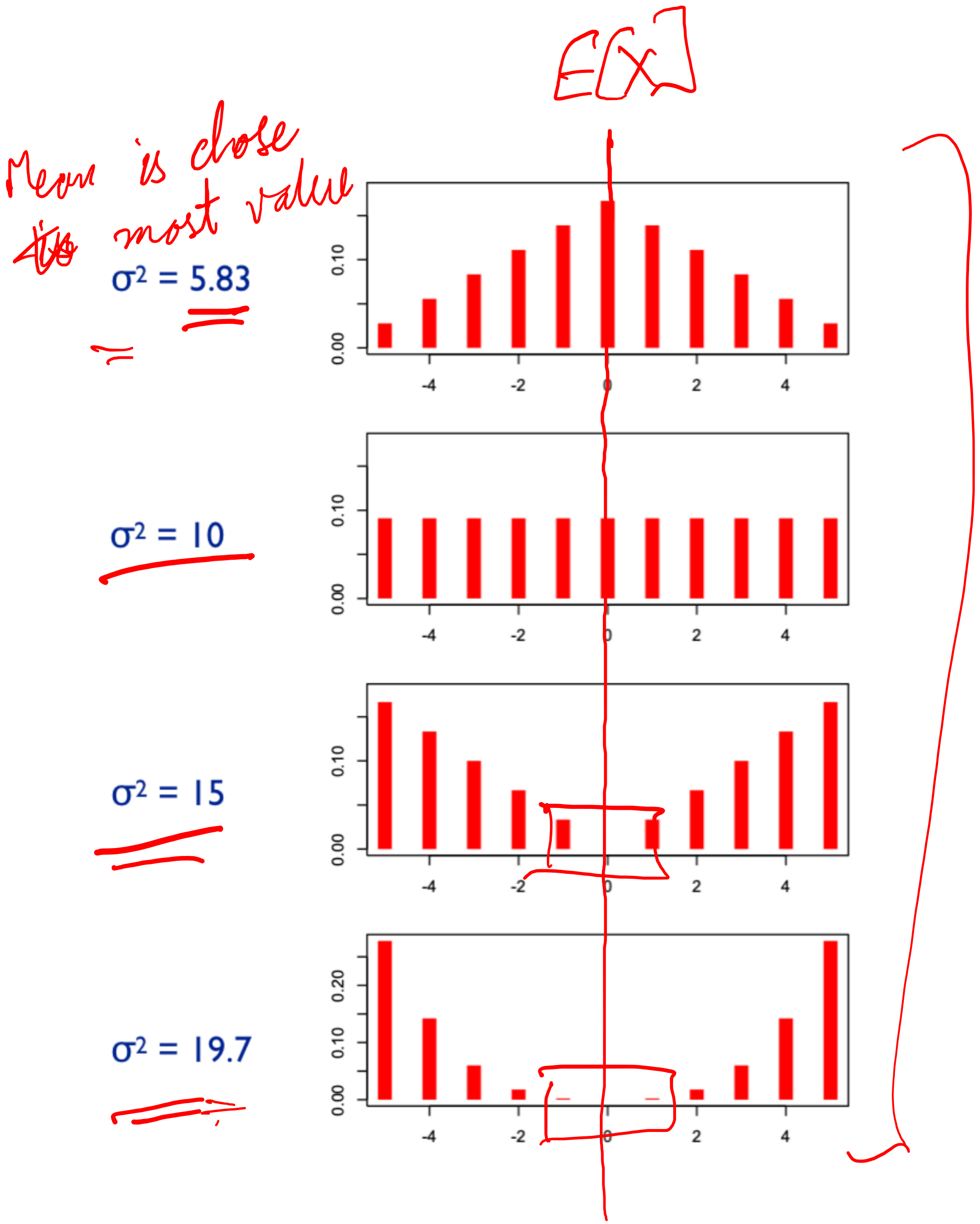
Variance (or standard deviation) is a quantity that measures, in expectation, how “far” the random variable is from its expectation.

Variance in Pictures

Captures how much "spread" there is in a pmf

All the PMFs graphed on the right have the same expectation of 0.

Expectation does not give the complete picture.



Variance of a die

Let X be the result of rolling a fair die.

$\mathrm{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[(X - 3.5)^2]$

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

$= \frac{35}{12} \approx 2.92.$

Or $\mathbb{E}[X^2] - (E[X])^2 = \sum_{k=1}^{6} \frac{1}{6} \cdot k^2 - 3.5^2 = \frac{91}{6} - 3.5^2 \approx 2.92$

Variance of n Coin Flips

Flip a coin n times, where it comes up heads with probability p each time (independently). Let X be the total number of heads.

We saw last time $\mathbb{E}[X] = np$.

$$X_i = \begin{cases} 1 & \text{if flip } i \text{ is heads} \\ 0 & \text{otherwise} \end{cases}$$

$$\mathbb{E}[X] = \mathbb{E}[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} \mathbb{E}[X_i] = \sum_{i=1}^{n} p = np.$$

Variance of n Coin Flips

Flip a coin n times, where it comes up heads with probability p each time (independently). Let X be the total number of heads.

What about $\text{Var}(X)$?

$$\mathbb{E}\left[(X - \mathbb{E}[X])^2\right] = \sum_{\omega} \mathbb{P}(\omega)(X(\omega) - np)^2$$

$$= \sum_{k=0}^{n} \binom{n}{k} \cdot p^k (1-p)^{n-k} \cdot (k - np)^2$$

Algebra time?

Variance

If X and Y are independent, then
$$\mathbf{Var}(X + Y) = \mathbf{Var}(X) + \mathbf{Var}(Y)$$

We'll talk about what it means for random variables to be independent soon…

For now, in this problem X_i is independent of X_j for $i \neq j$ where

$$X_i = \begin{cases} 1 & \text{if flip } i \text{ was heads} \\ 0 & \text{otherwise} \end{cases}$$

Variance of n Coin Flips

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

What's the $\mathrm{Var}(X_i)$?

$\mathbb{E}\left[(X_i - \mathbb{E}[X_i])^2\right]$

$= \mathbb{E}[(X_i - p)^2]$

$= p(1-p)^2 + (1-p)(0-p)^2$

$= p(1-p)[(1-p)+p] = p(1-p).$

OR

$Var(X_i) = \mathbb{E}\left[X_i^2\right] - \mathbb{E}[X_i]^2 = \mathbb{E}\left[X_i^2\right] - p^2 = p - p^2 = p(1-p)$.

$X_i = \begin{cases} 1 & \text{if } i\text{th flip is heads} \\ 0 & \text{o/w} \end{cases}$

$E[X_i] = P(X=1)$

$X_i^2 = \begin{cases} 1^2 = 1 & \text{if } i\text{th flip is heads} \\ 0^2 = 0 & \text{o/w} \end{cases}$

$E[X_i^2] = E[X_i] = p$

Plugging In

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

because each flip is independent

What's the $\mathrm{Var}(X_i)$?

$p(1-p)$.

$\mathrm{Var}(X) = \sum_{i=1}^{n} p(1-p) = np(1-p)$.

Variance simplification algebra trick

$\mathbb{E}\left[(X-\mathbb{E}[X])^2\right] = \mathbb{E}\left[X^2 - 2X\mathbb{E}[X] + (\mathbb{E}[X])^2\right]$ expanding the square

$= \mathbb{E}\left[X^2\right] - \mathbb{E}\left[2X\mathbb{E}[X]\right] + \mathbb{E}[(\mathbb{E}[X])^2]$ linearity of expectation.

$= \mathbb{E}\left[X^2\right] - 2\mathbb{E}[X]\mathbb{E}[X] + \mathbb{E}[(\mathbb{E}[X])^2]$ linearity of expectation.

$= \mathbb{E}\left[X^2\right] - 2\mathbb{E}[X]\mathbb{E}[X] + (\mathbb{E}[X])^2$ expectation of a constant is the constant

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

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

So, $\mathrm{Var}(X) = \mathbb{E}\left[X^2\right] - (\mathbb{E}[X])^2$.

$$P(A \cap B) = P(A) \cdot P(B)$$

A & B are independence

$$P(A|B) = P(A)$$

More Independence

or

$$P(B|A) = P(B)$$

Independence of events

Recall the definition of independence of **events**:

Independence

Two events A, B are independent if

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$$

Independence for 3 or more events

For three or more events, we need two kinds of independence

Pairwise Independence

Events $A_1, A_2, \ldots, A_n$ are pairwise independent if

$$\mathbb{P}(\boldsymbol{A_i} \cap \boldsymbol{A_j}) = \mathbb{P}(\boldsymbol{A_i}) \cdot \mathbb{P}(\boldsymbol{A_j}) \text{ for all } \boldsymbol{i, j}$$

Mutual Independence

A_1, A_2, A_3

(A_1, A_2) (A_1, A_3) (A_2, A_3) & (A_1, A_2, A_3)

Events $A_1, A_2, \ldots, A_n$ are mutually independent if

$$\mathbb{P}(\boldsymbol{A_{i_1}} \cap \boldsymbol{A_{i_2}} \cap \cdots \cap \boldsymbol{A_{i_k}}) = \mathbb{P}(\boldsymbol{A_{i_1}}) \cdot \mathbb{P}(\boldsymbol{A_{i_2}}) \cdots \mathbb{P}(\boldsymbol{A_{i_k}})$$

for every subset $\{\boldsymbol{i_1}, \boldsymbol{i_2}, \ldots, \boldsymbol{i_k}\}$ of $\{\mathbf{1}, \mathbf{2}, \ldots, \boldsymbol{n}\}$.

Pairwise Independence vs. Mutual Independence

Roll two fair dice (one red one blue) independently

R ="red die is 3"

B ="blue die is 5"

S ="sum is 7"

How should we describe these events?

Pairwise Independence

$\mathbb{P}(R \cap B) ?= \mathbb{P}(R)\mathbb{P}(B)$

$\frac{1}{36} = \frac{1}{6} \cdot \frac{1}{6}$ Yes! (These are also independent by the problem statement)

$\mathbb{P}(R \cap S) ?= \mathbb{P}(R)\mathbb{P}(S)$

$\frac{1}{36} ?= \frac{1}{6} \cdot \frac{1}{6}$ Yes!

$\mathbb{P}(B \cap S) ?= \mathbb{P}(B)\mathbb{P}(S)$

$\frac{1}{36} ?= \frac{1}{6} \cdot \frac{1}{6}$ Yes!

R, B, S are pairwise independent

Since all three pairs are independent, we say the random variables are pairwise independent.

Mutual Independence

$\mathbb{P}(R \cap B \cap S) = 0$

if the red die is 3, and blue die is 5 then the sum is 8 (so it can't be 7)

$\mathbb{P}(R)\mathbb{P}(B)\mathbb{P}(S) = \left(\frac{1}{6}\right)^3 = \frac{1}{216} \neq 0$

R, B, S are not mutually independent.

Checking Mutual Independence

It's not enough to check just $\mathbb{P}(A \cap B \cap C)$ either.

Roll a fair 8-sided die.

Let A be $\{1,2,3,4\}$

B be $\{2,4,6,8\}$

C be $\{1,2,3,7\}$

$$\mathbb{P}(A \cap B \cap C) = \mathbb{P}(\{2\}) = \frac{1}{8}$$

$$\mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}$$

$P(A \cap B) = P(A) \cdot P(B)$

$P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)$

Not sufficient

Checking Mutual Independence

It's not enough to check just $\mathbb{P}(A \cap B \cap C)$ either.

Roll a fair 8-sided die.

Let A be $\{1,2,3,4\}$

B be $\{2,4,6,8\}$

C be $\{1,2,3,7\}$

$$\mathbb{P}(A \cap B \cap C) = \mathbb{P}(\{2\}) = \frac{1}{8}$$

$$\mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}$$

But A and C aren't independent (nor are B, C; though A and B are independent). Because there's a subset that's not independent, A, B, C are not mutually independent.

$(A \cap B) = \{2,4\}$

$P(A \cap B) = \frac{2}{8} = \frac{1}{4}$

$P(A) \cdot P(B) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$

$P(A \cap C) = \frac{3}{8}$

$P(A) \cdot P(C) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$

$P(B \cap C) = \frac{1}{8}$

$P(B) \cdot P(C) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$

Checking Mutual Independence

To check mutual independence of events:

Check **every** subset.

To check pairwise independence of events:

Check **every** subset of size two.

Independence of Random Variables

That’s for events…what about random variables?

Independence (of random variables)

X and Y are independent if for all k, ℓ

$$\mathbb{P}(X = k, Y = \ell) = \mathbb{P}(X = k)\mathbb{P}(Y = \ell)$$

We’ll often use commas instead of ∩ symbol.

Independence of Random Variables

The "for all values" is important.

We say that the event "the sum is 7" is independent of "the red die is 5"

What about S ="the sum of two dice" and R ="the value of the red die"

Independence of Random Variables

The "for all values" is important.

We say that the event "the sum is 7" is independent of "the red die is 5"

What about S ="the sum of two dice" and R ="the value of the red die"

NOT independent.

$\mathbb{P}(S = 2, R = 5) \neq \mathbb{P}(S = 2)\mathbb{P}(R = 5)$ (for example)

|| 0 || Non-0

Independence of Random Variables

Flip a coin independently $2n$ times.

Let X be "the number of heads in the first n flips."

Let Y be "the number of heads in the last n flips."

X and Y are independent.

Mutual Independence for RVs

A little simpler to write down than for events

Mutual Independence (of random variables)

$X_1, X_2, \dots, X_n$ **are mutually independent if for all** $x_1, x_2, \dots, x_n$

$$\mathbb{P}(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) = \mathbb{P}(X_1 = x_1)\mathbb{P}(X_2 = x_2) \cdots \mathbb{P}(X_n = x_n)$$

DON'T need to check all subsets for random variables…

But you do need to check all values (all possible x_i) still.

What Does Independence Give Us?

If X and Y are independent random variables, then

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$$

$$\mathbb{E}[X \cdot Y] = \mathbb{E}[X]\mathbb{E}[Y]$$

$$\mathbb{E}\left[\frac{X}{Y}\right] = \mathbb{E}[X] \cdot \mathbb{E}\left[\frac{1}{Y}\right]$$

$$\neq \frac{\mathbb{E}[X]}{\mathbb{E}[Y]}$$

Facts about Variance

Shifting the variance

We know that

$\mathbb{E}[aX + c] = a\mathbb{E}[X] + c$

What happens with variance?

i.e., What is $Var(aX + c)$? What is $Var(aX)$?

Facts about Variance

$\text{Var}(X + c) = \text{Var}(X)$

Proof:

$$\text{Var}(X + c) = \mathbb{E}\left[(X + c)^2\right] - \mathbb{E}[X + c]^2$$

$$= \mathbb{E}\left[X^2\right] + \mathbb{E}[2Xc] + \mathbb{E}\left[c^2\right] - (\mathbb{E}[X] + c)^2$$

$$= \mathbb{E}\left[X^2\right] + 2c\mathbb{E}[X] + c^2 - \mathbb{E}[X]^2 - 2c\mathbb{E}[X] - c^2$$

$$= \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2$$

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

Facts about Variance

$\mathrm{Var}(aX) = a^2\mathrm{Var}(X)$

Proof:

$Var(aX) = \mathbb{E}\left[(aX)^2\right] - (\mathbb{E}[aX])^2$

$= a^2\mathbb{E}\left[X^2\right] - (a\mathbb{E}[X])^2$

$= a^2\mathbb{E}\left[X^2\right] - a^2\mathbb{E}[X]^2$

$= a^2\left(\mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2\right)$

$= a^2 Var(X)$