More Independence; Variance

CSE 312 Spring 24
Lecture 11

More Independence

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

Mutual Independence

Events $A_1, A_2, \ldots, A_n$ are mutually independent if
$\mathbb{P}(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}) = \mathbb{P}(A_{i_1}) \cdot \mathbb{P}(A_{i_2}) \cdots \mathbb{P}(A_{i_k})$
for every subset $\{i_1, i_2, \ldots, i_k\}$ of $\{1, 2, \ldots, n\}$.

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"

$$\mathbb{P}(S=7, R=5) = \mathbb{P}(S=7)\,\mathbb{P}(R=5)$$

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

$$\mathbb{P}(S=2, R=5) = 0 \qquad \mathbb{P}(S=2)\,\mathbb{P}(R=5) = 1/216$$

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)

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 you?

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

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

What does XY mean? I rolled two dice, let X be the red die, Y the blue die. XY is the random variable that tells you the product of the two dice.

That's a function that takes in an outcome and gives you a number back…so a random variable!! (Same for $X + Y$).

Functions of a random variable

Let X, Y be random variables defined on the same sample space.

Functions of X and/or Y like

$X + Y$

$Z = X + Y$

X^2

$2X + 3$

Etc.

Are random variables! (Say what the outcome is, and these functions give you a number. They're functions from $\Omega \to \mathbb{R}$. **That's the definition of a random variable!**

Expectations of functions of random variables

Let's say we have a random variable X and a function g. What is $\mathbb{E}[g(X)]$?

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

Equivalently: $\mathbb{E}[g(X)] = \sum_{k\in\Omega_{g(X)}} k \cdot \mathbb{P}(g(X) = k)$

Notice that $\mathbb{E}[g(X)]$ might not be $g(\mathbb{E}[X])$.

Variance

Where are we?

A random variable is a way to summarize what outcome you saw.

The Expectation of a random variable is its average value.

A way to summarize a random variable

Variance

Another one number summary of a random variable.

But wait, we already have expectation, what's this for?

Consider these two games

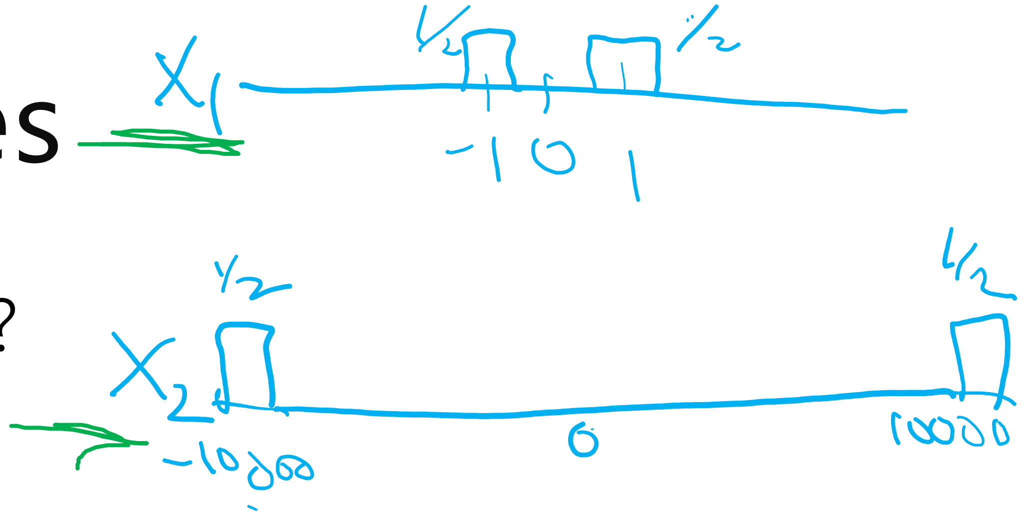
Would you be willing to play these games?

Game 1: I will flip a fair coin; if it's heads, I pay you \$1. If it's tails, you pay me \$1. Let X_1 be your profit if you play game 1

Game 2: I will flip a fair coin; if it's heads, I pay you \$10,000. If it's tails, you pay me \$10,000. Let X_2 be your profit if you play game 2.

Both games are "fair" ($\mathbb{E}[X_1] = \mathbb{E}[X_2] = 0$)

Consider these two games



Would you be willing to play these games?

Game 1: I will flip a fair coin; if it's heads, I pay you \$1. If it's tails, you pay me \$1. Let X_1 be your profit if you play game 1

Game 2: I will flip a fair coin; if it's heads, I pay you \$10,000. If it's tails, you pay me \$10,000. Let X_2 be your profit if you play game 2.

Both games are "fair" ($\mathbb{E}[X_1] = \mathbb{E}[X_2] = 0$)

What’s 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 many (well, only) very extreme results.

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?

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

$$\frac{1}{2} - \frac{1}{2} = 0$$

What happens with Game 2?

$$\frac{1}{2} \cdot (100000 - 0) + \frac{1}{2} \cdot (-100000 - 0)$$

$$5000 - 5000 = 0$$

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?

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

What happens with Game 2?

$$\frac{1}{2} \cdot (10{,}000 - 0)^2 + \frac{1}{2} \cdot (-10{,}000 - 0)^2$$
$$50{,}000{,}000 + 50{,}000{,}000 = 10^8$$

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

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

The first two forms are the definition. The last one is an algebra trick.

Variance of a die

Let X be the result of rolling a fair die.

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

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

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

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

$E[g(X)] \neq g(E(X))$

$E[XY] = E[X]E[Y]$

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'll see next time $\mathbb{E}[X] = np$.

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

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

Are the X_i independent? Yes!

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

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

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

Plugging In

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

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

$p(1-p)$.

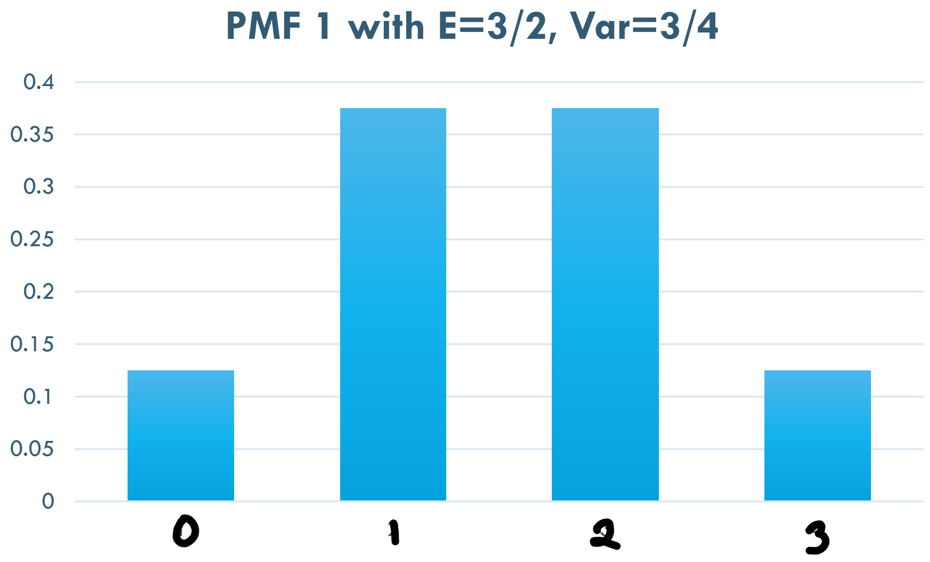
$\text{Var}(X) = \sum_{i=1}^{n} p(1-p) = np(1-p)$.

Expectation and Variance aren't everything

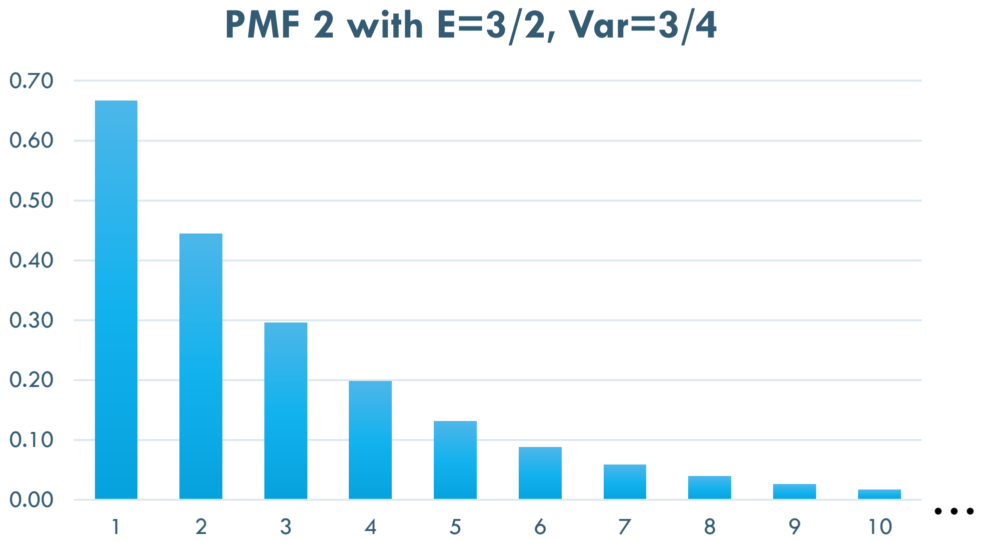
Alright, so expectation and variance is everything right?

No!

Flip a fair coin 3 times indep. Count heads.



Flip a biased coin (prob heads=2/3) until heads. Count flips.



A PMF or CDF *does* fully describe a random variable.

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

Useful Facts

Make a prediction

How should $\mathrm{Var}(X + c)$ relate to $Var(X)$ if c is a constant?

How should $\mathrm{Var}(\mathrm{aX})$ relate to $\mathrm{Var}(X)$ is a is a constant?

Make a prediction

How should $\mathrm{Var}(X + c)$ relate to $Var(X)$ if c is a constant?

How should $\mathrm{Var(aX)}$ relate to $\mathrm{Var}(X)$ is a is a constant?

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

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

Facts About Variance

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

Proof:

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

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

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

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

$= \mathrm{Var}(X)$

Facts about Variance

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

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

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

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

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

Extra Practice

More Practice

Suppose you flip a coin until you see a heads for the first time.

Let X be the number of trials (including the heads)

What is the pmf of X?

The cdf of X?

$\mathbb{E}[X]$?

More Practice

Suppose you flip a coin until you see a heads for the first time.

Let X be the number of trials (including the heads)

What is the pmf of X? $f_X(x) = 1/2^x$ for $x \in \mathbb{Z}^+$, 0 otherwise

The cdf of X? $F_X(x) = 1 - 1/2^{\lfloor x \rfloor}$ for $x \geq 0$, 0 for $x < 0$.

$\mathbb{E}[X]$? $\sum_{i=1}^{\infty} \frac{i}{2^i} = 2$

More Random Variable Practice

Roll a fair die n times. Let X be the number of rolls that are $5s$ or $6s$.

What is the pmf?

Don’t try to write the CDF…it’s a mess…

Or try for a few minutes to realize it isn’t nice.

What is the expectation?

More Random Variable Practice

Roll a fair die n times. Let Z be the number of rolls that are $5s$ or $6s$.

What's the probability of getting exactly k 5's/6's? Well we need to know which k of the n rolls are 5's/6's. And then multiply by the probability of getting exactly that outcome

$$f_Z(z) = \begin{cases} \binom{n}{z} \cdot \left(\frac{1}{3}\right)^z \left(\frac{2}{3}\right)^{n-z} & \text{if } z \in Z, 0 \leq z \leq n \\ 0 & \text{otherwise} \end{cases}$$

Expectation formula is a mess. If you plug it into a calculator you'll get a nice, clean simplification: $n/3$.