Random variables, expectation, and variance

DSE 210

Random variables

Roll a die.

$$\text{Define } X = \begin{cases} 1 & \text{if die is} \geq 3 \\ 0 & \text{otherwise} \end{cases}$$

Here the sample space is $\Omega = \{1, 2, 3, 4, 5, 6\}$.

$$\omega = 1, 2 \Rightarrow X = 0$$

$$\omega = 3, 4, 5, 6 \Rightarrow X = 1$$

Random variables

Roll a die.

$$\text{Define } X = \begin{cases} 1 & \text{if die is} \geq 3 \\ 0 & \text{otherwise} \end{cases}$$

Here the sample space is $\Omega = \{1, 2, 3, 4, 5, 6\}$.

$$\omega = 1, 2 \Rightarrow X = 0$$
$$\omega = 3, 4, 5, 6 \Rightarrow X = 1$$

Roll n dice.

$$X = \# \text{ of 6's}$$
$$Y = \# \text{ of 1's before the first 6}$$

Both X and Y are defined on the same sample space, $\Omega = \{1, 2, 3, 4, 5, 6\}^n$. For instance,

$$\omega = (1, 1, 1, \ldots, 1, 6) \ \Rightarrow \ X = 1, Y = n - 1.$$

Random variables

Roll a die.

$$\text{Define } X = \begin{cases} 1 & \text{if die is } \geq 3 \\ 0 & \text{otherwise} \end{cases}$$

Here the sample space is $\Omega = \{1, 2, 3, 4, 5, 6\}$.

$$\omega = 1, 2 \Rightarrow X = 0$$
$$\omega = 3, 4, 5, 6 \Rightarrow X = 1$$

Roll n dice.

$$X = \# \text{ of 6's}$$
$$Y = \# \text{ of 1's before the first 6}$$

Both X and Y are defined on the same sample space, $\Omega = \{1, 2, 3, 4, 5, 6\}^n$. For instance,

$$\omega = (1, 1, 1, \ldots, 1, 6) \;\Rightarrow\; X = 1, Y = n - 1.$$

In general, a **random variable (r.v.)** is a defined on a probability space. It is a mapping from Ω to $\mathbb{R}$. We'll use capital letters for r.v.'s.

The distribution of a random variable

Roll a die. Define $X = 1$ if die is ≥ 3, otherwise $X = 0$.

- What is the set of values X can take?

$X \in \{0, 1\}$

- What is the probability of each of these values?

$\Pr(X = 0) = 1/3, \quad \Pr(X = 1) = 2/3$

Roll n dice. Define X = number of 6's.

- What is the set of values X can take?

$X \in \{0, 1, 2, \ldots, n\}$

- What is the probability of each of these values?

$$\Pr(X = k) = \underbrace{\binom{n}{k}}_{\text{\# ways to choose where the k sixes will be}} \underbrace{\left(\frac{1}{6}\right)^k \left(\frac{5}{6}\right)^{n-k}}_{\text{probability of that exact configuration}}$$

$\frac{5}{6} \quad \frac{1}{6} \quad \frac{5}{6} \quad \frac{5}{6} \quad \frac{1}{6} \quad \frac{5}{6}$

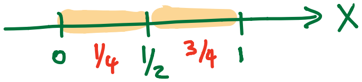
X X

1 2 3 n

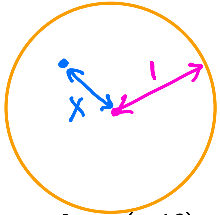
e.g. $k=2$

$$\left(\frac{5}{6}\right)^{n-k} \left(\frac{1}{6}\right)^k$$

Distribution of a continuous random variable



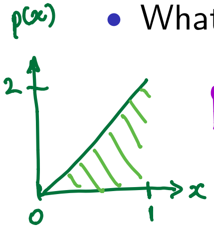
Throw a dart at a dartboard of radius 1. Let X be the distance to the center of the board.



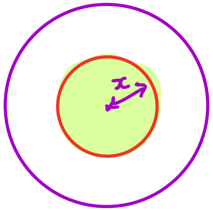
- What is the set of values X can take?

$$X \in [0,1] = \{0 \le x \le 1\}$$

- What is $F(x) = \Pr(X \le x)$? This is the **cumulative distribution function** (cdf) of X.

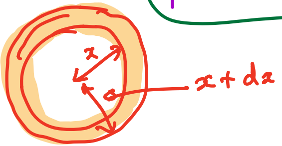


$$\Pr(X \le x) = \frac{\text{area of shaded region}}{\text{area of circle}} = \frac{\pi x^2}{\pi \cdot 1^2} = x^2$$



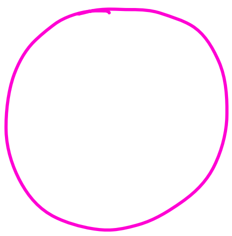
- The **density** of X is the derivative of the cdf. What is it?

$$p(x) = 2x$$



$$\frac{F(x+dx) - F(x)}{dx}$$

$$\therefore \Pr(X \in [a,b]) = \int_a^b p(x)\,dx = \int_a^b 2x\,dx = x^2\Big]_a^b = b^2 - a^2$$



X = distance to center

cdf $F(x) = \Pr(X \le x) = x^2$

What does the density $p(x)$ mean?

- $\Pr(X = x) = 0$
- $\underbrace{\Pr(X \in [x, x+dx])}_{\underbrace{\Pr(X \le x+dx) - \Pr(X \le x)}_{F(x+dx) - F(x)}} \approx p(x)\, dx$

$$\Rightarrow \quad p(x) \approx \frac{F(x+dx) - F(x)}{dx}$$

Expected value, or mean

The expected value of a discrete random variable X is

$$\mathbb{E}(X) = \sum_x x \Pr(X = x).$$

Expected value, or mean

The expected value of a discrete random variable X is

$$\mathbb{E}(X) = \sum_x x \Pr(X = x).$$

weighted average of the various values that X can take

Roll a die. Let X be the number observed.

- What is $\mathbb{E}(X)$?

$X \in \{1, 2, 3, 4, 5, 6\}$, all equally likely

$$\mathbb{E}[X] = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6}$$

$$= \frac{1+2+3+4+5+6}{6} = \frac{21}{6} = 3.5$$

Biased coin

<u>Biased coin.</u> A coin has heads probability p.

- Toss the coin once, and let X be 1 if heads, 0 if tails.
  What is $\mathbb{E}(X)$?

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

- Toss the coin repeatedly, until it comes up heads. Let T be the $\#$ of tosses.
  What is $\mathbb{E}(T)$? $1/p$

$$\begin{aligned}
\mathbb{E}(T) &= 1 \cdot \Pr(T=1) + 2 \cdot \Pr(T=2) + 3 \cdot \Pr(T=3) + \cdots \\
&= 1 \cdot p + 2 \cdot (1-p)p + 3 \cdot (1-p)^2 \cdot p + 4 \cdot (1-p)^3 \cdot p + \cdots \\
&= \cdots \\
&= 1/p
\end{aligned}$$

$$\mathbb{E}(T) = 1 + (1-p)\,\mathbb{E}(T)$$

$$\Rightarrow \quad p\,\mathbb{E}(T) = 1$$

$$\Rightarrow \quad \mathbb{E}(T) = 1/p$$

Pascal's wager

Pascal: I think there is some chance (say, probability $p > 0$) that God exists. Therefore I should act as if he exists.

Pascal's wager

Pascal: I think there is some chance (say, probability $p > 0$) that God exists. Therefore I should act as if he exists.

Let X = my level of suffering.

- Suppose I behave as if God exists (that is, I behave myself).
  Then X is some significant but finite amount, like 100.
  What is $\mathbb{E}(X)$?

  $$\mathbb{E}(X) = 100$$

- Suppose I behave as if God doesn't exists (I do whatever I want to).
  If indeed God doesn't exist: $X = 0$.
  But if God exists: $X = \infty$ (hell).
  What is $\mathbb{E}(X)$?

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

Expectation of a continuous random variable

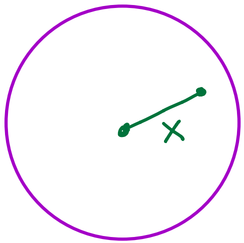
If continuous random variable X has density $p(x)$, its expected value is:

$$\mathbb{E}(X) = \int x \; p(x) \; dx.$$

Expectation of a continuous random variable

If continuous random variable X has density $p(x)$, its expected value is:

$$\mathbb{E}(X) = \int x \, p(x) \, dx.$$



Throw a dart at a dartboard of radius 1. Let X be the distance to the center of the board. What is $\mathbb{E}(X)$?

We've already seen that the density $p(x) = 2x$

$$\mathbb{E}(X) = \int_0^1 x \cdot (2x) \, dx = \int_0^1 2x^2 \, dx = \left. \frac{2x^3}{3} \right]_0^1 = \frac{2}{3}$$

Median

Two ways of summarizing a set of numbers by a single number.

- **The mean**
- **The median**: the number in the middle, if you sort them

Find the median of the following sets of numbers:

- $10, \cancel{-20}, \cancel{100}, 20, 50$

median = 20

- $50, 100, 60, 90, 20, 10$

10, 20, 50, 60, 90, 100 } median = anything in the range [50,60]

How can we define the median of a random variable X?

It is any value m for which

$\Pr(X \leq m) \geq \frac{1}{2}$ and $\Pr(X \geq m) \geq \frac{1}{2}$.

Mean vs median

The university wants to keep track of salaries of UCSD students ten years after graduation. They find that in a class of 100:

- 9 have a salary of 0
- 15 have a salary of 40K
- 25 have a salary of 100K
- 25 have a salary of 150K
- 25 have a salary of 200K
- 1 has become a billionaire

- What is the mean salary, roughly? $> \$10$ million

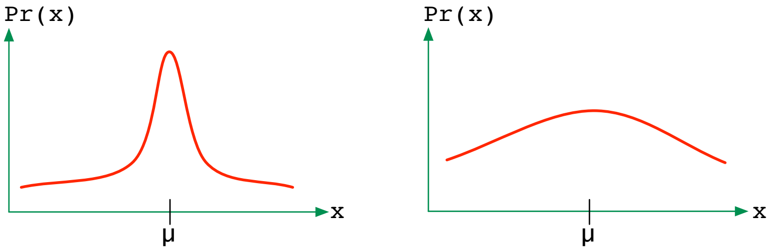
Which is the more honest assessment?

- What is the median salary? \$150K

Variance

We can summarize a random variable X by its mean μ (or median).

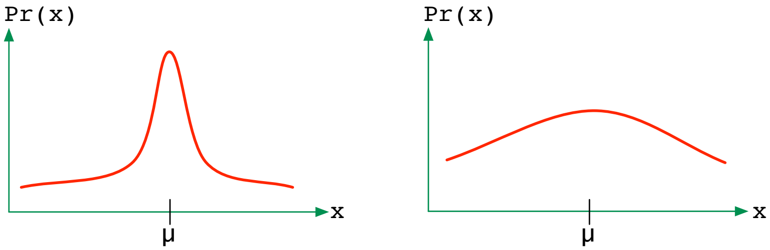
Problem: This doesn't capture the **spread** of X.



Variance

We can summarize a random variable X by its mean μ (or median).

Problem: This doesn't capture the **spread** of X.

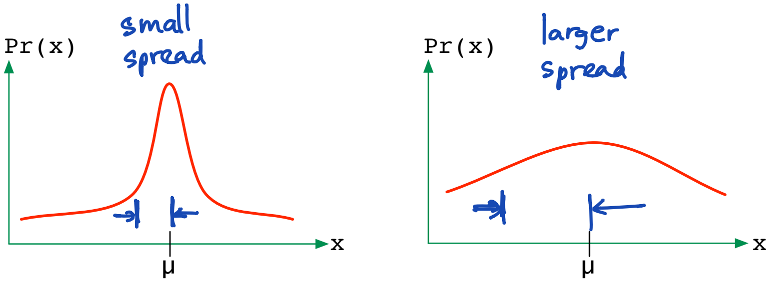


What is a good measure of spread? How about: average distance from the mean, $\mathbb{E}(|X - \mu|)$?

Variance

We can summarize a random variable X by its mean μ (or median).

Problem: This doesn't capture the **spread** of X.



What is a good measure of spread? How about: average distance from the mean, $\mathbb{E}(|X-\mu|)$?

For convenience, take the square instead of the absolute value.

$$\textbf{Variance:} \qquad \mathrm{var}(X) = \mathbb{E}\left[(X-\mu)^2\right] = \mathbb{E}(X^2) - \mu^2,$$

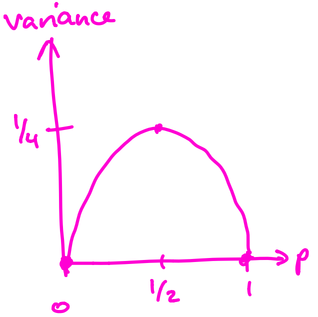
where $\mu = \mathbb{E}(X)$. The $\underbrace{\text{variance is always } \geq 0.}_{\text{why?}}$

Example: Variance of a biased coin

$(X-\mu)^2 \in \{(1-p)^2, p^2\}$

Recall: $\mathrm{var}(X) = \mathbb{E}(X-\mu)^2 = \mathbb{E}(X^2) - \mu^2$, where $\mu = \mathbb{E}(X)$.

Toss a coin of bias p. Let $X \in \{0, 1\}$ be the outcome. Compute the following:



- $\mathbb{E}(X) =$ p
- $\mathbb{E}(X^2) =$ p since $X^2 = X$
- $\mathbb{E}[(X-\mu)^2] =$ $(1-p)^2 \cdot \Pr(X=1) + p^2 \cdot \Pr(X=0) = (1-p)^2 p + p^2(1-p) = p(1-p)$
- $\mathbb{E}(X^2) - \mu^2 =$ $p - p^2$

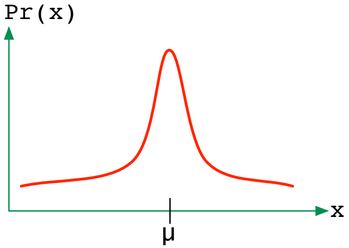
equal

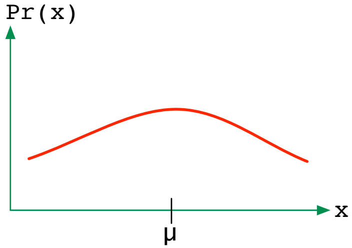
For what value of p does the coin have highest variance?

$p = \frac{1}{2}$

Standard deviation

Recall: $\text{var}(X) = \mathbb{E}(X - \mu)^2$, where $\mu = \mathbb{E}(X)$.





The **standard deviation** of X is $\text{std}(X) = \sqrt{\text{var}(X)}$.
It is, *roughly*, the average amount by which X differs from its mean.

Question: How does $\text{std}(X)$ relate to $\mathbb{E}(|X - \mu|)$? Are they equal?

Not exactly the same, but closely related. In general

$$\mathbb{E}\left[|X-\mu|\right] \leq \text{std}(X)$$

Worksheet 4

1. A die is thrown twice. Outcomes: X_1, X_2.
Define X to be the minimum of X_1 and X_2.
What is the distribution of X?

(a) What values can X take?

$X \in \{1, 2, 3, 4, 5, 6\}$

(b) What is the probability of each of these values?

$Pr(X = 6) = Pr(X_1 = 6, X_2 = 6) = 1/36$

$Pr(X = 5) = Pr((5,5), (5,6), (6,5)) = 3/36 = 1/12$

$Pr(X = 4) = Pr((6,4), (5,4), (4,4), (4,5), (4,6)) = 5/36$

X_2 \ X_1 ←	1	2	3	4	5	6
1						
2						
3			3	3	3	3
4			3	4	4	4
5			3	4	5	5
6			3	4	5	6

$Pr(X = 3) = ?$

$Pr(X = 2) = ?$

$Pr(X = 1) = ?$

2. A fair die is rolled repeatedly until a six is seen.
What is the expected # of rolls?

This is like a coin of bias $1/6$. ← p

$E[\# \text{ rolls}] = 1/p = 6$

$\boxed{3.}$ A die has the following probabilities:

$$Pr(1) = Pr(2) = Pr(3) = Pr(4) = \frac{1}{8}$$

$$Pr(5) = Pr(6) = \frac{1}{4}$$

Roll the die; let Z be the outcome.

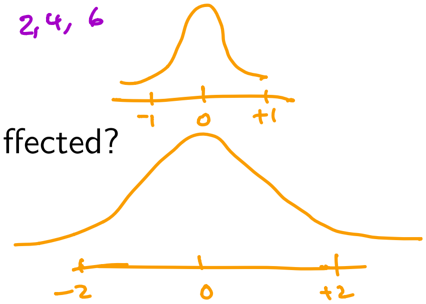
(a) What is $\mathbb{E}[Z]$?

$$\mathbb{E}(Z) = 1 \cdot \frac{1}{8} + 2 \cdot \frac{1}{8} + 3 \cdot \frac{1}{8} + 4 \cdot \frac{1}{8} + 5 \cdot \frac{1}{4} + 6 \cdot \frac{1}{4}$$

$$= \frac{10}{8} + \frac{11}{4} = \frac{32}{8} = 4.$$

Linear functions of a single random variable

1, 2, 3 ←
2, 4, 6



- If you double a set of numbers, how are their mean and variance affected?

  Mean ×2

  Variance ×4 (standard deviation ×2)

- If you increase a set of numbers by 1, how much do their mean and variance change?

  Mean +1

  Variance unchanged

- Let X be any random variable.
  For some constants a, b, define a new random variable $V = aX + b$.
  Express $\mathbb{E}(V)$ and $\text{var}(V)$ in terms of $\mathbb{E}(X)$ and $\text{var}(X)$.

  $\mathbb{E}(V) = a\,\mathbb{E}(X) + b$

  $\text{var}(V) = a^2\,\text{var}(X)$

Linearity of expectation

A powerful and extremely useful property:

Linearity of expectation: For any random variables $X_1, \ldots, X_m$,

$$\mathbb{E}(X_1 + X_2 + \cdots + X_m) = \mathbb{E}(X_1) + \mathbb{E}(X_2) + \cdots + \mathbb{E}(X_m).$$

Mean of the sum = sum of the means

Linearity: example

Roll 2 dice and let Z denote the sum. What is $\mathbb{E}(Z)$?

Linearity: example

Roll 2 dice and let Z denote the sum. What is $\mathbb{E}(Z)$?

(5,6)
(6,5) (6,6)

Method 1

Distribution of Z:

z	2	3	4	5	6	7	8	9	10	11	12
$\Pr(Z = z)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Now use formula for expected value:

$$\mathbb{E}(Z) = 2 \cdot \frac{1}{36} + 3 \cdot \frac{2}{36} + 4 \cdot \frac{3}{36} + \cdots = 7.$$

Linearity: example

Roll 2 dice and let Z denote the sum. What is $\mathbb{E}(Z)$?

Method 1

Distribution of Z:

z	2	3	4	5	6	7	8	9	10	11	12
$\Pr(Z = z)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

somewhat painful

Now use formula for expected value:

$$\mathbb{E}(Z) = 2 \cdot \frac{1}{36} + 3 \cdot \frac{2}{36} + 4 \cdot \frac{3}{36} + \cdots = 7.$$

Method 2 ← Much easier

Let X_1 be the first die and X_2 the second die. Each of them is a single die and thus (as we saw earlier) has expected value 3.5. Since $Z = X_1 + X_2$,

$$\mathbb{E}(Z) = \mathbb{E}(X_1) + \mathbb{E}(X_2) = 3.5 + 3.5 = 7.$$

Using linearity of expectation

The typical scenario:

- You want to compute the expectation of a random variable X.
- But it is complicated, or messy.
- So instead you notice that X can be rewritten in the form

$$X = X_1 + X_2 + \cdots + X_m,$$

  where $X_1, X_2, \ldots, X_m$ are much simpler random variables whose expected values are easy to compute.
- By linearity, $\mathbb{E}(X) = \mathbb{E}(X_1) + \cdots + \mathbb{E}(X_m)$.

$X = \#\text{heads} \in \{0, 1, 2, \ldots, n\}$

Toss n coins of bias p, and let X be the number of heads. What is $\mathbb{E}(X)$?

- Write out the formula for $\mathbb{E}(X)$. Does it look easy to figure out?

$$\mathbb{E}(X) = \sum_{h=0}^{n} h \cdot \Pr(X = h) = \sum_{h=0}^{n} h \binom{n}{h} p^h (1-p)^{n-h} = ?? \quad \text{tricky}$$

- Can you write X as the sum of simpler random variables?

Define $X_i = \begin{cases} 1 & \text{if } i^{th} \text{ coin is heads} \\ 0 & \text{if tails} \end{cases}$

Then $X = X_1 + X_2 + \cdots + X_n$

- Compute $\mathbb{E}(X)$ using this alternative representation.

$\mathbb{E}(X_i) = p$ (since X_i is a coin of bias p)

$\mathbb{E}(X) = np$

Coupon collector, again

$X_1 = 1$, X_2 is most likely to be 1

Each cereal box has one of k action figures. What is the expected number of boxes you need to buy in order to collect all the figures?

- Keep buying boxes until you get all the figures. Let X be the number of boxes bought. We want $\mathbb{E}(X)$.
- Let X_i be the number of additional boxes needed to collect the ith action figure, after the first $i-1$ distinct action figures have been obtained. Thus $X = X_1 + X_2 + \cdots + X_k$.
- Can you show that $\mathbb{E}(X_i) = \frac{k}{k-i+1}$?

Already have i-1 action figures.

$$\left.\Pr(\text{random box contains a new figure}) = \frac{\#\text{remaining figures}}{k} = \frac{k-i+1}{k}\right\} \therefore \mathbb{E}[\#\text{boxes before a new figure appears}] = \frac{k}{k-i+1} \quad \left(= \frac{1}{p}\right)$$

- What is $\mathbb{E}(X)$?

$$\mathbb{E}(X) = \mathbb{E}(X_1) + \mathbb{E}(X_2) + \mathbb{E}(X_3) + \cdots + \mathbb{E}(X_k)$$

$$= \frac{k}{k} + \frac{k}{k-1} + \frac{k}{k-2} + \frac{k}{k-3} + \cdots + \frac{k}{1}$$

first figure: 1

last figure: k boxes

$$= k\Big(\underbrace{1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots + \frac{1}{k}}_{\text{harmonic series}}\Big) \approx k \ln k$$

Linearity of variance

We've seen that $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$.
Is this also true of variance, i.e., is $\text{var}(X + Y) = \text{var}(X) + \text{var}(Y)$?

- In general, **no**. Give a counterexample.

E.g. $Y = X$

$\text{var}(X+Y) = \text{var}(2X) = 4\,\text{var}(X) \neq \text{var}(X) + \text{var}(X)$

- But it is true if X and Y are **independent**...

Independent random variables

Random variables X, Y are **independent** if $\Pr(X = x, Y = y) = \Pr(X = x)\Pr(Y = y)$.

Independent random variables

Random variables X, Y are **independent** if $\Pr(X = x, Y = y) = \Pr(X = x)\Pr(Y = y)$.

Pick a card out of a standard deck.
X = suit and Y = number. ← independent

$\Pr(X = \heartsuit, Y = 10) = 1/52$

$\Pr(X = \heartsuit) = 1/4$

$\Pr(Y = 10) = 1/13$

$\Pr(X = \heartsuit, Y = 10)$

$= \Pr(X = \heartsuit) \Pr(Y = 10)$

Works for any choice of <u>suit</u> and <u>number</u>

Independent random variables

Random variables X, Y are **independent** if $\Pr(X = x, Y = y) = \Pr(X = x)\Pr(Y = y)$.

Flip a fair coin 10 times.
$X = \#$ heads and $Y =$ last toss.

not independent

$\Pr(X = 0, Y = \text{heads}) = 0$

$\Pr(X = 0) = \frac{1}{2^{10}}$

$\Pr(Y = \text{heads}) = \frac{1}{2}$

$\Pr(X = 0, Y = \text{heads}) \neq$

$\Pr(X = 0) \cdot \Pr(Y = \text{heads})$

Independent random variables

Random variables X, Y are **independent** if $\Pr(X = x, Y = y) = \Pr(X = x)\Pr(Y = y)$.

$X, Y \in \{-1, 0, 1\}$, with these probabilities:

		Y		
		-1	0	1
	-1	0.4	0.16	0.24
X	0	0.05	0.02	0.03
	1	0.05	0.02	0.03

this means $Pr(X=-1, Y=0) = 0.16$

check: $0.03 = 0.1 \times 0.3$? ✓

check: $0.4 = 0.8 \times 0.5$ ✓

What are the marginal distributions of X and Y?

x	Pr
-1	0.8
0	0.1
1	0.1

y	Pr
-1	0.5
0	0.2
1	0.3

For all x, y, $Pr(X=x, Y=y) = Pr(X=x)\, Pr(Y=y)$.

$\therefore$ Indpt.

[Need to check all 9 entries]

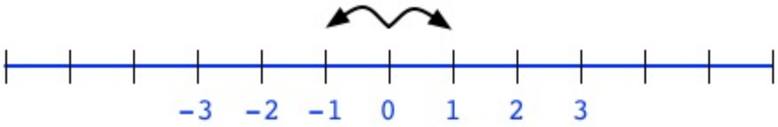
Variance of a sum

$\mathrm{var}(X_1 + \cdots + X_k) = \mathrm{var}(X_1) + \cdots + \mathrm{var}(X_k)$ if the X_i are independent.

Variance of a sum

$\text{var}(X_1 + \cdots + X_k) = \text{var}(X_1) + \cdots + \text{var}(X_k)$ if the X_i are independent.

Symmetric random walk. A drunken man leaves a bar. At each time step, he either moves one step to the right or one step to the left, with equal probabilities. Where is he after n steps?



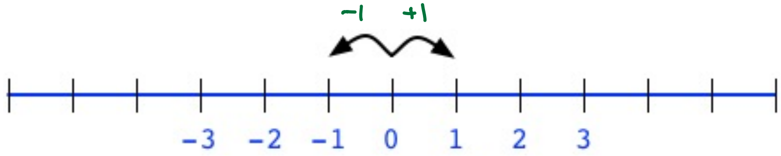
Variance of a sum

$X \in \{-n, .., 0, .., n\}$

X_i -1 $+1$

$\text{var}(X_1 + \cdots + X_k) = \text{var}(X_1) + \cdots + \text{var}(X_k)$ if the X_i are independent.

Symmetric random walk. A drunken man leaves a bar. At each time step, he either moves one step to the right or one step to the left, with equal probabilities. Where is he after n steps?



Let $X_i \in \{-1, 1\}$ be his ith step. His position after n steps is $X = X_1 + \cdots + X_n$.

- What are $\mathbb{E}(X_i)$ and $\text{var}(X_i)$?

$\mathbb{E}(X_i) = (+1)\cdot\frac{1}{2} + (-1)\cdot\frac{1}{2} = 0$; $X_i^2 = 1$; $\mathbb{E}(X_i^2) = 1$; $\text{var}(X_i) = \mathbb{E}(X_i^2) - (\mathbb{E}(X_i))^2 = 1$

- What are $\mathbb{E}(X)$, $\text{var}(X)$, and $\text{std}(X)$?

$\mathbb{E}(X) = 0$, $\text{var}(X) = n$, $\text{std}(X) = \sqrt{n}$

- Roughly how far from the starting point would we expect him to be after n steps?

Roughly $\sqrt{n}$ steps: doesn't get far! E.g. $n = 10{,}000 \Rightarrow \sqrt{n} = 100$

Worksheet # 4

→ 1, 2, 3, 8, 9, 11, 12, 18, 20

9 (a) n people.

Each person chooses a number in $\{1, .., 10\}$ at random.

$$\Pr(\text{exactly one person chooses the number } i) = \binom{n}{1} \frac{1}{10} \cdot \left(\frac{9}{10}\right)^{n-1}$$

$$= n \cdot \frac{1}{10} \cdot \left(\frac{9}{10}\right)^{n-1}$$

$$\left(\underset{\substack{1 \\ \frac{9}{10}}}{\_\_}, \underset{\substack{2 \\ \frac{9}{10}}}{\_\_}, \underset{\substack{3 \\ \frac{1}{10}}}{\_\_}, \ldots , \underset{\substack{n \\ \frac{9}{10}}}{\_\_} \right)$$

$$\Pr(\text{exactly } k \text{ people choose the number } i) = \Pr(k \text{ of the coins are heads})$$

- Each person is a coin toss

  Heads $\equiv$ choosing i

- $\Pr(\text{heads}) = p = \frac{1}{10}$

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

$$= \binom{n}{k} \left(\frac{1}{10}\right)^k \left(\frac{9}{10}\right)^{n-k}$$

(b) $X_i = \begin{cases} 1 & \text{if exactly one person chooses floor } i \\ 0 & \text{o.w.} \end{cases}$

$\mathbb{E}(X_i) = \frac{n}{10} \left(\frac{9}{10}\right)^{n-1}$ ← from previous problem

$X = X_1 + X_2 + \ldots + X_{10} \Rightarrow \mathbb{E}(X) = n \left(\frac{9}{10}\right)^{n-1}$

$$\left\{ 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \ldots + \frac{1}{n} \right\} \sim \ln n$$

$$\int_1^n \frac{1}{x}\, dx = \ln x \Big]_1^n = \ln n$$

equal to the shaded area