Probability review

Probability spaces, events, and conditioning

1. How to define the **probability space** for an experiment in which outcomes are random.
2. How to formulate an **event** of interest.
3. The probability that two events both occur.
4. The **conditional probability** that an event occurs, given that some other event has occurred.
5. **Bayes' rule**.

Probability spaces

You roll two dice.

What is the probability they add to 10?

The **probability space** has two components:

1. **Sample space** (space of outcomes).

2. **Probabilities of outcomes**, summing to 1.

Events

Probability space:

- Outcomes: $\Omega = \{\text{all possible pairs of dice rolls}\}$
- Every pair $z = (z_1, z_2) \in \Omega$ has probability $1/36$.

Event of interest: the two dice add up to 10.

Multiple events

You have ten coins. Nine are fair, but one is a bad coin that always comes up tails.

- You close your eyes and pick a coin at random.
- You toss it four times, and it comes up tails every time.

What is the probability you picked the bad coin?

- Ten coins: nine are fair, one is a bad coin that always comes up tails.
- You pick a coin at random, toss it four times, and it's tails every time.

Conditioning

For two events A, B, **conditional probability**

$\Pr(B|A)$ = probability that B occurs, given that A occurs

Conditioning formula: $\boxed{\Pr(A \cap B) = \Pr(A)\Pr(B|A)}$

In our example:

- A: the bad coin is chosen
- B: all four tosses are tails

Want $\Pr(A|B)$

- Ten coins: nine are fair, one is a bad coin that always comes up tails.
- You pick a coin at random, toss it four times, and it's tails every time.

Event A: the bad coin is chosen. Event B: all tails

Bayes' rule

Two events A, B

- We are interested in A
- We can observe B

If we find out B occurred, how does it alter the probability of A?

Bayes' rule: $\Pr(A|B) = \Pr(A) \times \dfrac{\Pr(B|A)}{\Pr(B)}$

Random variables, expectation, and variance

1. What is a random variable?
2. Expected value
3. Variance and standard deviation

Random variables

Roll two dice. Let X be their sum.

$$\begin{aligned} \text{outcome} = (1,1) &\quad \Rightarrow \quad X = 2 \\ \text{outcome} = (1,2) \text{ or } (2,1) &\quad \Rightarrow \quad X = 3 \end{aligned}$$

Probability space:

- Sample space: $\Omega = \{1,2,3,4,5,6\} \times \{1,2,3,4,5,6\}$.
- Each outcome equally likely.

Random variable X lies in $\{2,3,4,5,6,7,8,9,10,11,12\}$.

A **random variable (r.v.)** is a defined on a probability space.
It is a mapping from Ω (outcomes) to $\mathbb{R}$ (numbers).
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 $X = 0$.

Expected value, or mean

Expected value of a random variable X:

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

Roll a die. Let X be the number observed.
What is $\mathbb{E}(X)$?

Another example

A biased coin has heads probability p.
Let X be 1 if heads, 0 if tails. What is $\mathbb{E}(X)$?

A property of expected values

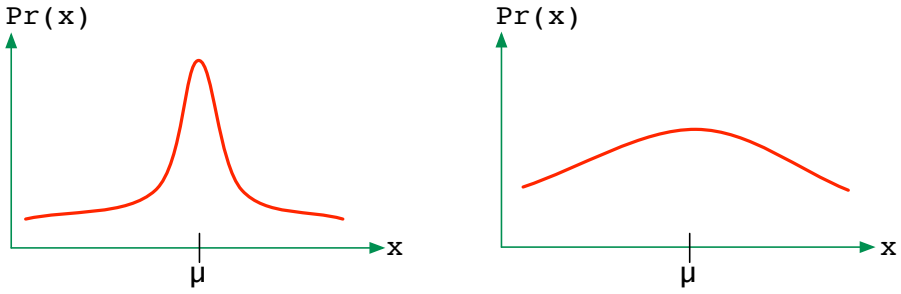
How is the average of a set of numbers affected if:

- You double the numbers?
- You increase each number by 1?

Summary: Let X be any random variable.
If $V = aX + b$ (any constants a, b), then $\mathbb{E}(V) = a\mathbb{E}(X) + b$

Variance

Can summarize an r.v. X by its mean, μ. But this doesn't capture the **spread** of X:



A measure of spread: average distance from the mean, $\mathbb{E}(|X - \mu|)$?

- **Variance:** $\text{var}(X) = \mathbb{E}((X - \mu)^2)$, where $\mu = \mathbb{E}(X)$
- **Standard deviation** $\sqrt{\text{var}(X)}$:
  Roughly, the average amount by which X differs from its mean.

Variance: example

Choose X uniformly at random from $\{1, 2, 3, 4, 5\}$.

Variance: properties

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

- Variance is always ≥ 0
- How is the variance affected if:
  - You increase each number by 1?
  - You double each number?
- Summary: If $V = aX + b$ then $\text{var}(V) = a^2 \text{var}(X)$

Alternative formula for variance

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

Another way to write it: $\text{var}(X) = \mathbb{E}(X^2) - \mu^2$

Example: Choose X uniformly at random from $\{1, 2, 3, 4, 5\}$.

Measuring dependence

1. When are two random variables **independent**?
2. Qualitatively assessing dependence
3. Quantifying dependence: **covariance** and **correlation**

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 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.

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

Dependence

Example: Pick a person at random, and take

$$H = \text{height}$$
$$W = \text{weight}$$

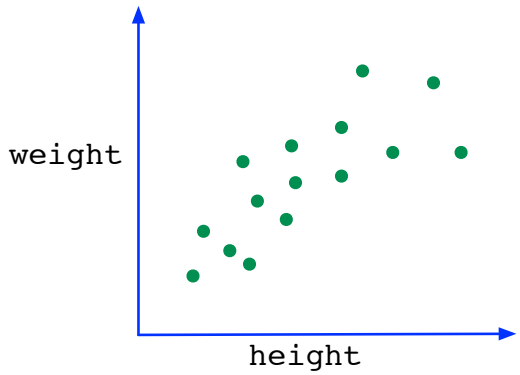
Independence would mean

$$\Pr(H = h, W = w) \;=\; \Pr(H = h)\Pr(W = w).$$

Not accurate: height and weight will be **positively correlated**.

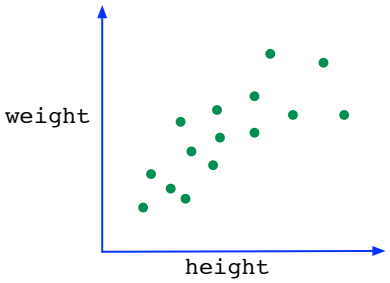
Positive correlation

H, W are **positively correlated**



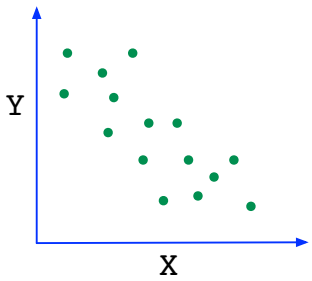
This also implies $\mathbb{E}[HW] > \mathbb{E}[H]\,\mathbb{E}[W]$.

Types of correlation

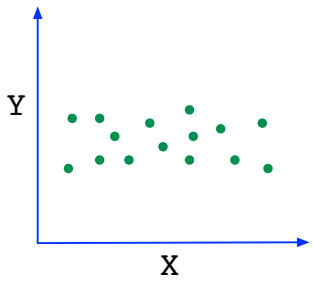


H, W **positively correlated**
This also implies

$$\mathbb{E}[HW] > \mathbb{E}[H]\,\mathbb{E}[W]$$

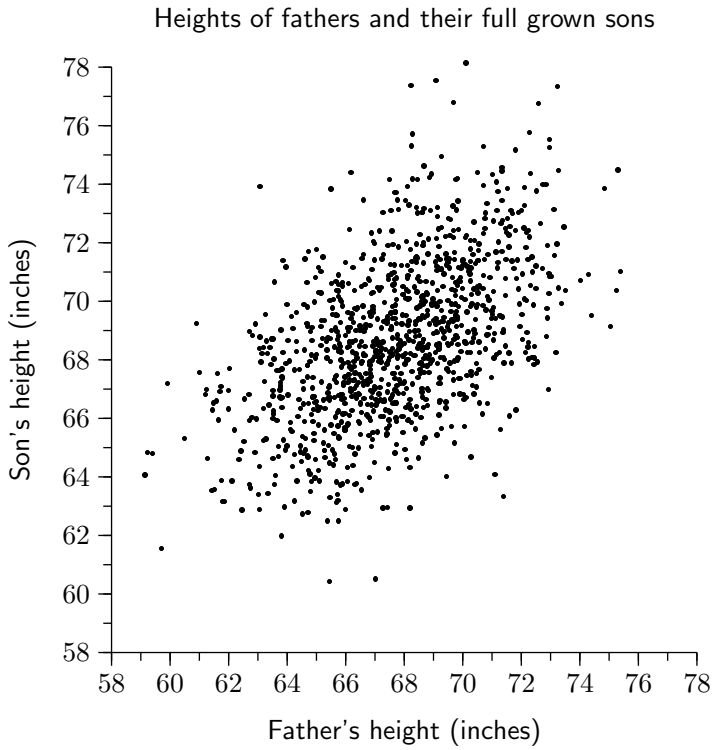


X, Y **negatively correlated**
$\mathbb{E}[XY] < \mathbb{E}[X]\,\mathbb{E}[Y]$



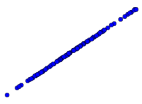
X, Y **uncorrelated**
$\mathbb{E}[XY] = \mathbb{E}[X]\,\mathbb{E}[Y]$

Pearson (1903): fathers and sons

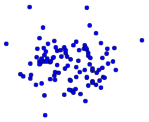


Correlation coefficient: pictures

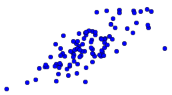
$r = 1$



$r = 0$

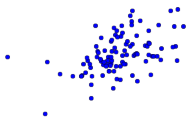


$r = 0.75$

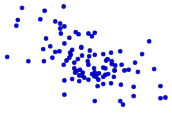


$r = -0.25$

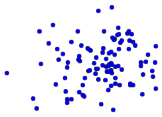
$r = 0.5$



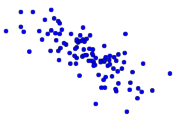
$r = -0.5$



$r = 0.25$



$r = -0.75$



Covariance and correlation

- Covariance

$$\begin{aligned}\mathrm{cov}(X, Y) &= \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] \\ &= \mathbb{E}[XY] - \mathbb{E}[X]\,\mathbb{E}[Y]\end{aligned}$$

  Maximized when $X = Y$, in which case it is $\mathrm{var}(X)$.
  In general, it is at most $\mathrm{std}(X)\mathrm{std}(Y)$.

- Correlation

$$\mathrm{corr}(X, Y) = \frac{\mathrm{cov}(X, Y)}{\mathrm{std}(X)\mathrm{std}(Y)}$$

  This is always in the range $[-1, 1]$.

If X, Y independent then $\mathrm{cov}(X, Y) = 0$.
But the converse need not be true.

Covariance and correlation: example

Find $\mathrm{cov}(X, Y)$ and $\mathrm{corr}(X, Y)$

x	y	$\Pr(x, y)$
-1	-3	$1/6$
-1	3	$1/3$
1	-3	$1/3$
1	3	$1/6$