Probability review II: Random variables and dependencies between them

Instructor: Taylor Berg-Kirkpatrick
Slides: Sanjoy Dasgupta

Random variables

Roll two dice. Let X be their sum.
E.g.,

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

Random variables

Roll two dice. Let X be their sum.
E.g.,

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

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 $\{2,3,4,5,6,7,8,9,10,11,12\}$.

Random variables

Roll two dice. Let X be their sum.
E.g.,

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

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 $\{2,3,4,5,6,7,8,9,10,11,12\}$.

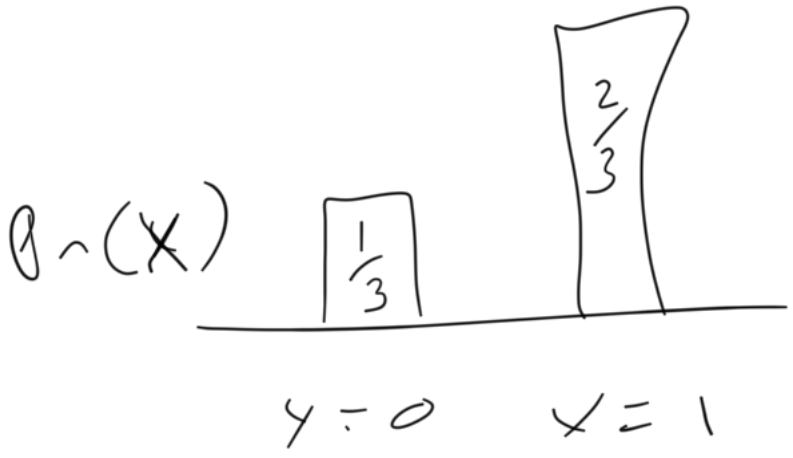
A **random variable (r.v.)** is a defined on a probability space.
It is a mapping from Ω (outcomes) to $\mathbb{R}$ (numbers).
We'll use capital letters for r.v.'s.

The distribution of a random variable

Roll a die. y

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

$X = \mathbb{1}[y \geq 3]$



$P(X = 1) = 2/3$

$P(X = 0) = 1/3$

Another example

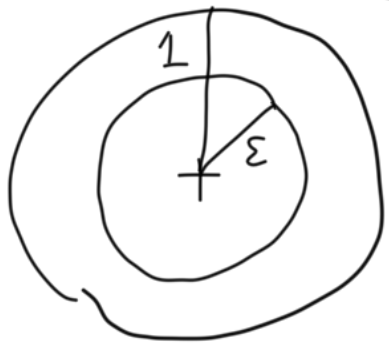
Throw a dart at a dartboard of radius 1.

Let X = distance to center of board.

$\Pr(X = \varepsilon)$?

$$\Pr(X \le \varepsilon) = \frac{\text{Area}(B_\varepsilon)}{\text{Area}(B_1)}$$

B_ε = ball of radius ε.



Expected value, or mean

Expected value of a random variable X:

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

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 $\mathbb{E}(X)$?

$$E[X] = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} - \cdots - 6 \cdot \frac{1}{6}$$

$$= 3.5$$

Another example

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

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

Linearity of expectation

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

- You double the numbers?

Linearity of expectation

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

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

Linearity of expectation

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

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

Summary: $\mathbb{E}(aX + b) = a\mathbb{E}(X) + b$
(any r.v. X, and constants a, b)

Linearity of expectation

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

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

Summary: $\mathbb{E}(aX + b) = a\mathbb{E}(X) + b$
(any r.v. X, and constants a, b)

Something more powerful is true:

- $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$ for any two random variables X, Y.

Linearity of expectation

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

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

Summary: $\mathbb{E}(aX + b) = a\mathbb{E}(X) + b$
(any r.v. X, and constants a, b)

Something more powerful is true:

- $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$ for any two random variables X, Y.
- Likewise: $\mathbb{E}(X + Y + Z) = \mathbb{E}(X) + \mathbb{E}(Y) + \mathbb{E}(Z)$, etc.

Linearity: examples

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

Linearity: examples

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
$\Pr(Z = z)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Linearity: examples

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
$\Pr(Z = z)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

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

Linearity: examples

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
$\Pr(Z = z)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

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

Method 2

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 $Z = X_1 + X_2$,

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

Another example

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

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

$$\mathbb{E}(X) = \mathbb{E}(X_1) + \cdots + \mathbb{E}(X_n)$$

$$= \underbrace{p + \cdots + p}_{n}$$

$$= np$$

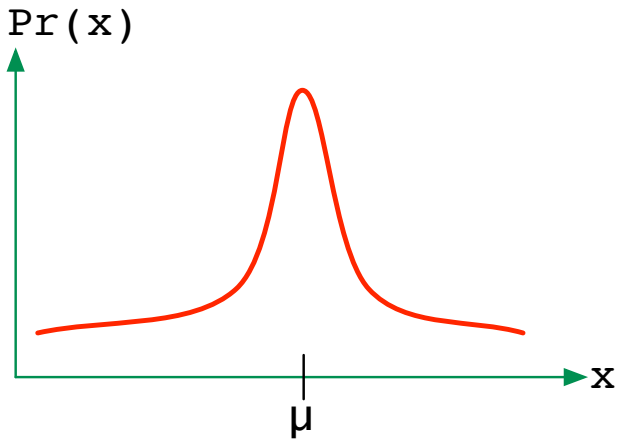
Variance

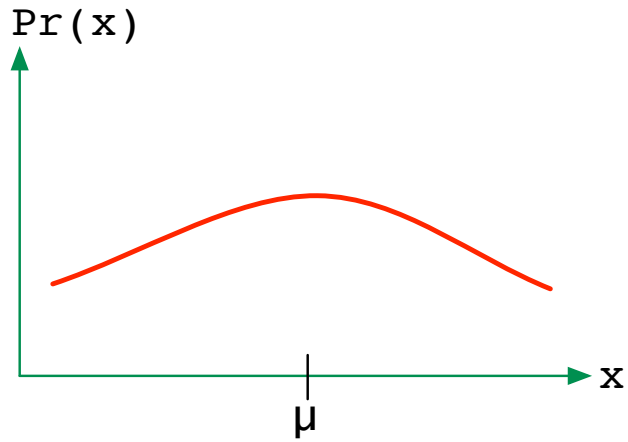
Can summarize an r.v. X by the mean/median, μ.

Variance

Can summarize an r.v. X by the mean/median, μ.

But these don't capture the **spread** of X:

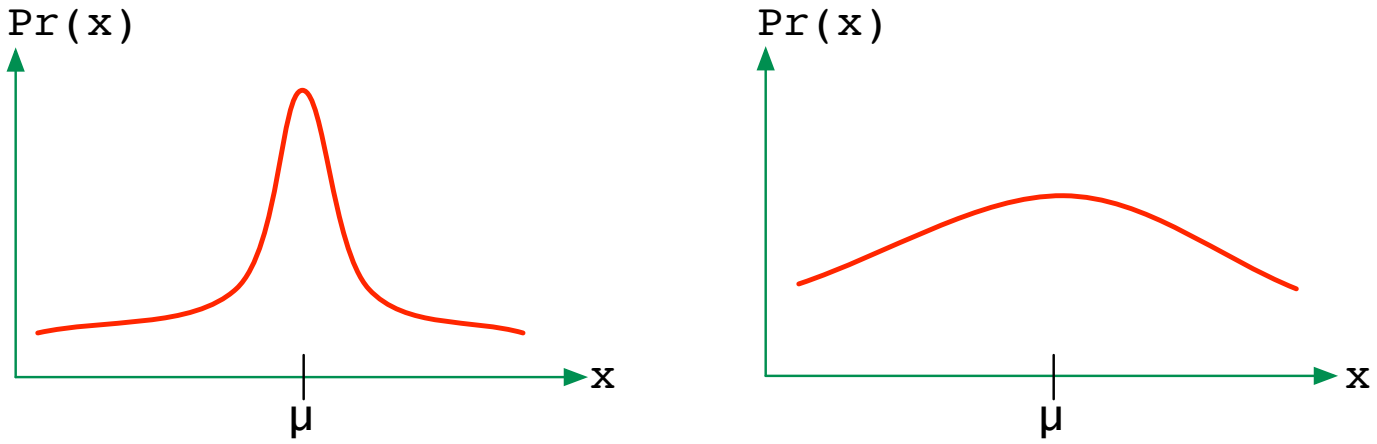




Variance

Can summarize an r.v. X by the mean/median, μ.

But these don't capture the **spread** of X:

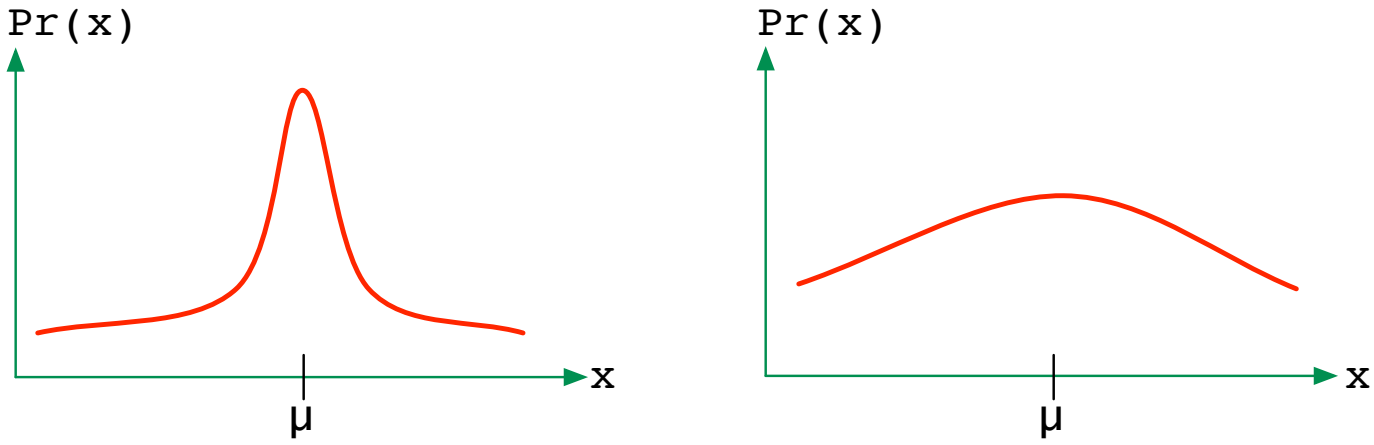


A measure of spread: average distance from the mean, $\mathbb{E}(|X - \mu|)$?

Variance

Can summarize an r.v. X by the mean/median, μ.

But these don'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: properties

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

- Variance is always ≥ 0
- Another way to write it: $\text{var}(X) = \mathbb{E}(X^2) - \mu^2$

Variance: properties

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

- Variance is always ≥ 0
- Another way to write it: $\text{var}(X) = \mathbb{E}(X^2) - \mu^2$

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

Independent random variables

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

$$\frac{P(X, Y)}{P(Y)} = P(X)$$

$$P(X \mid Y) = P(X)$$

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

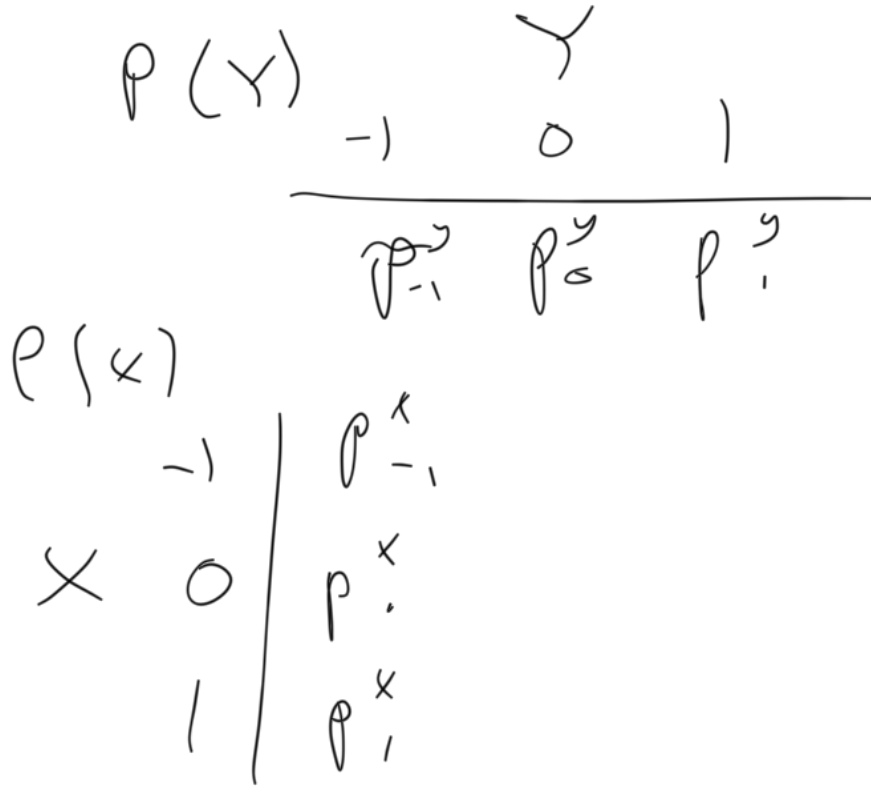
Flip a fair coin 10 times.
$X = \#$ heads and $Y =$ last toss.

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
X	-1	0.4	0.16	0.24
	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

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

$$P(H \mid W) = P(H)$$

Dependence

Example: Pick a person at random, and take

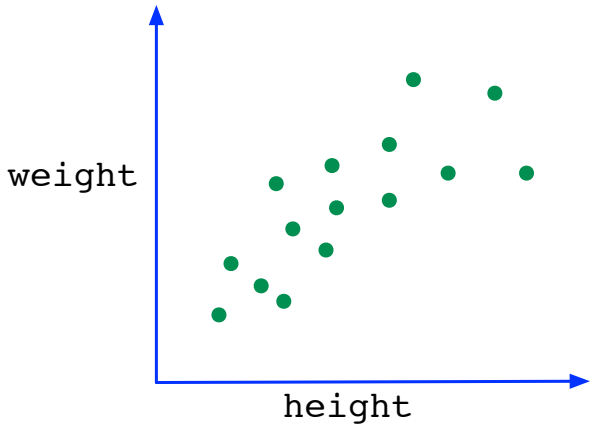
$$H = \text{height}$$
$$W = \text{weight}$$

Independence would mean

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

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

Types of correlation

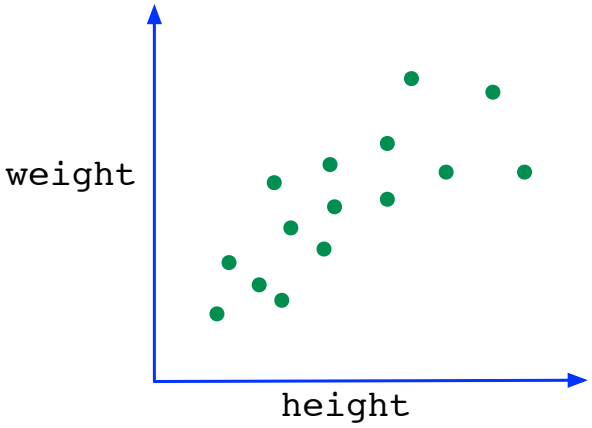


H, W **positively correlated**

This also implies

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

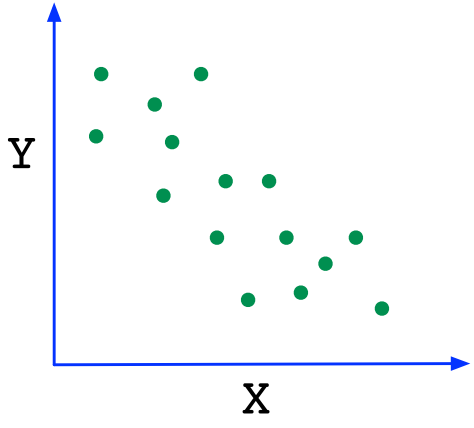
Types of correlation



H, W **positively correlated**

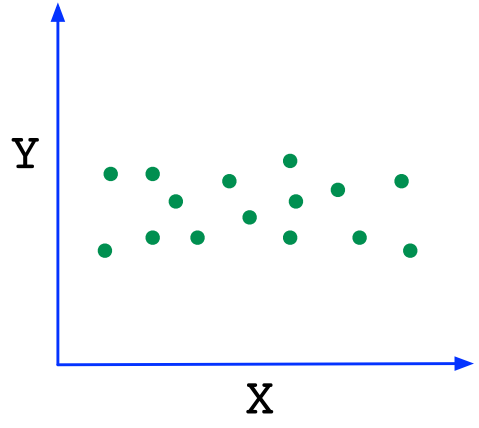
This also implies

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



X, Y **negatively correlated**

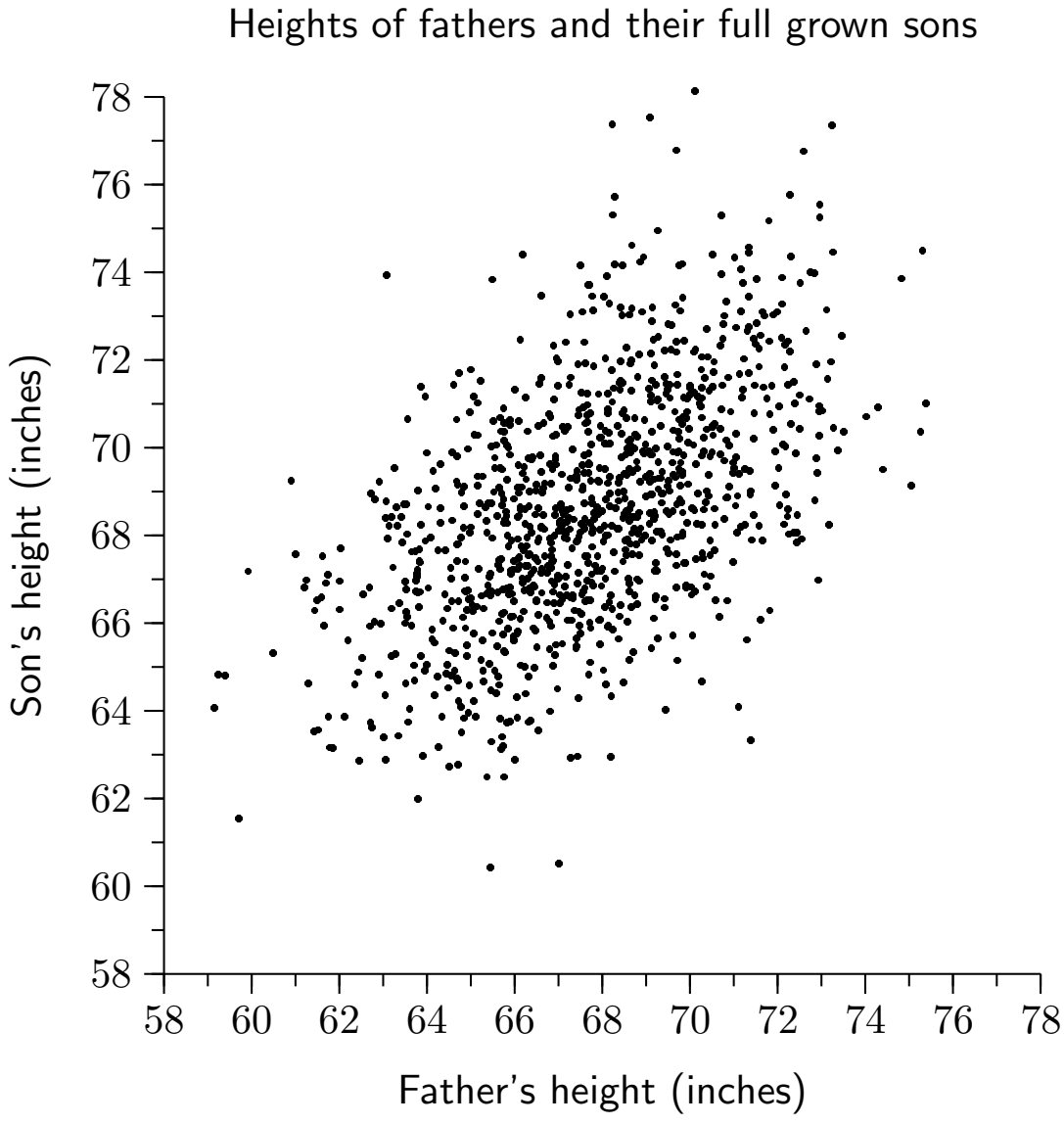
$$\mathbb{E}[XY] < \mathbb{E}[X]\,\mathbb{E}[Y]$$



X, Y **uncorrelated**

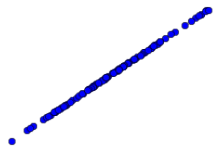
$$\mathbb{E}[XY] = \mathbb{E}[X]\,\mathbb{E}[Y]$$

Pearson (1903): fathers and sons

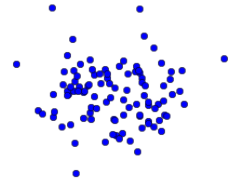


Correlation coefficient: pictures

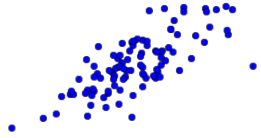
$r = 1$



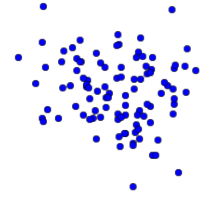
$r = 0$



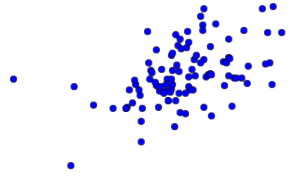
$r = 0.75$



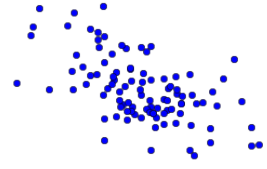
$r = -0.25$



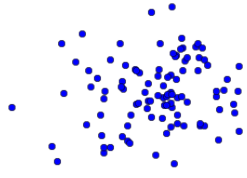
$r = 0.5$



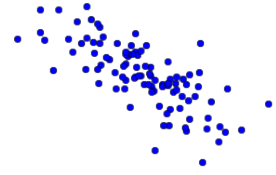
$r = -0.5$



$r = 0.25$



$r = -0.75$



Covariance and correlation

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

- Covariance

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

$\mathrm{cov}(X, X) = \mathrm{Var}(X)$

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

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 $[-1, 1]$.

Covariance and correlation: example 1

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

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

Covariance and correlation: example 2

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

x	y	$\Pr(x, y)$
-1	-10	$1/6$
-1	10	$1/3$
1	-10	$1/3$
1	10	$1/6$