UNIVERSITY OF BIRMINGHAM

Current Topics in Data Science and AI

Denoising in Scientific Imaging

Background:

Random Variables

Discrete Probability Distributions

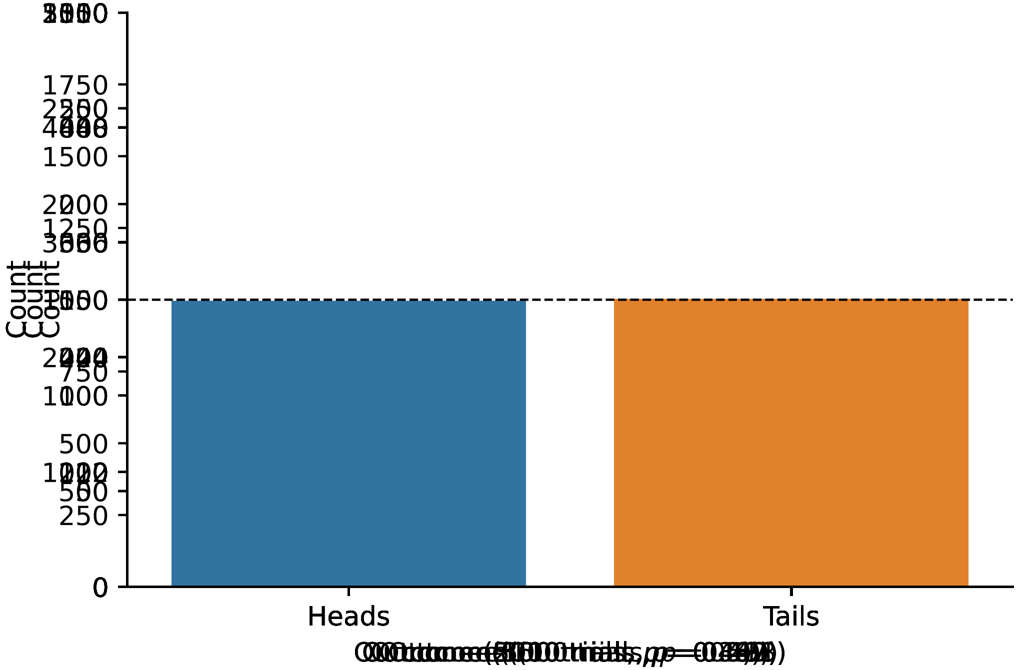
Coin flip experiment:

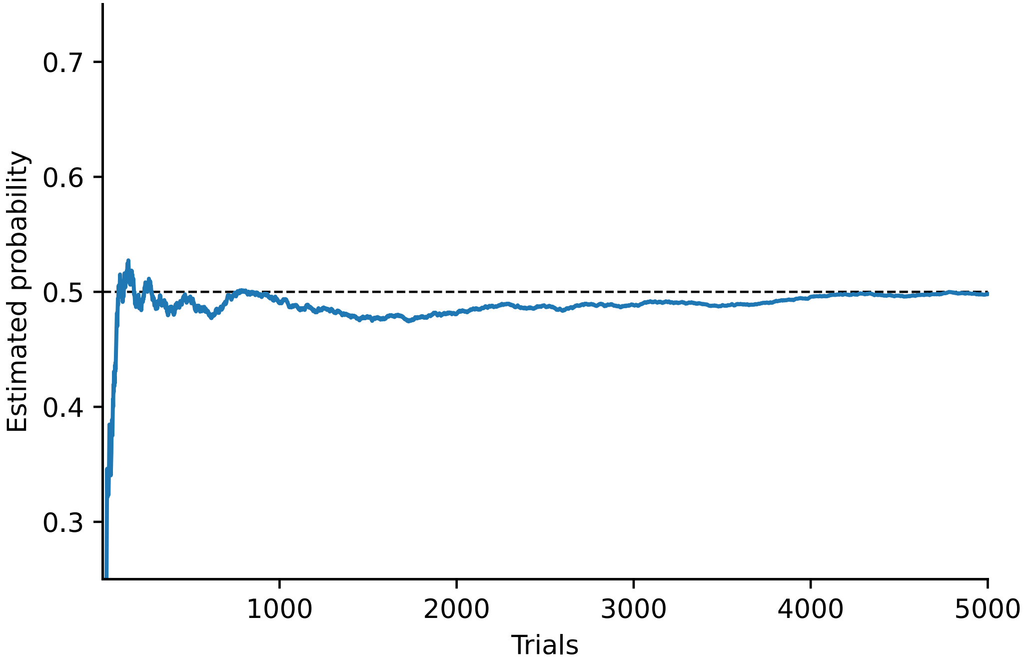
$X = 1$ (heads) $\qquad\qquad$ $X = 0$ (tails)

$$P(X = 1) = \lim_{K\to\infty} \frac{1}{K}\sum_{k=1}^{K} \mathbf{1}(X_k = 1)$$

Probability

Relative frequency





Discrete Probability Distributions

Coin flip experiment:

$X = 1$ (heads)

$X = 0$ (tails)

$$\underbrace{P(X=1)}_{\text{Probability}} = \lim_{K \to \infty} \underbrace{\frac{1}{K} \sum_{k=1}^{K} \mathbf{1}(X_k = 1)}_{\text{Relative frequency}}$$

Notation:

$P(X = x)$

often written $p(x)$

Variable

Value

Probability Mass Function (PMF):

$P(X = 1)$	$P(X = 0)$
0.5	0.5

$$\sum_x p(x) = 1$$

Expected Values

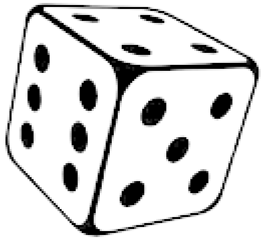
$$\mathbb{E}_{p(x)}[x] = \sum_{x} p(x)x$$

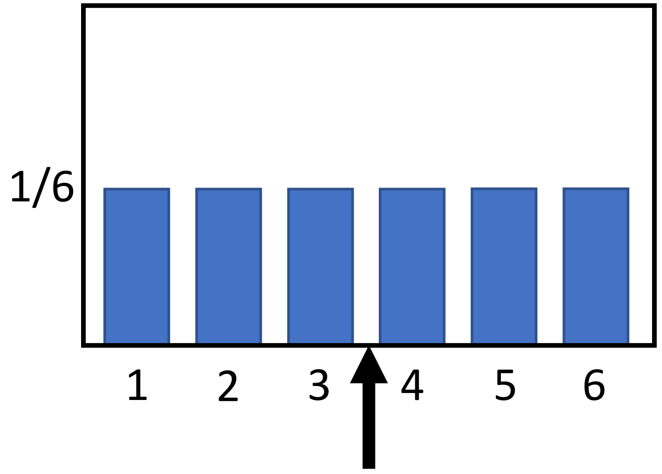
$$= \lim_{K\to\infty} \frac{1}{K} \sum_{k=1}^{K} x_k \quad | x_k \sim p(x)$$

- Center of mass of the distribution
- Expected value minimizes quadratic error

$(s\text{-}x)^2$

- Rolling dice:



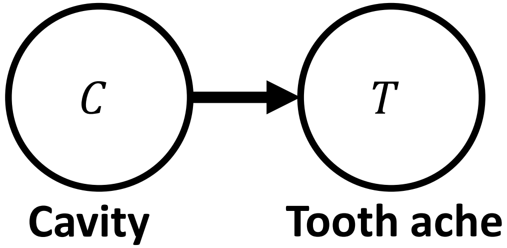


$$\mathbb{E}_{p(x)}[x] = \sum_{x=1}^{6} \frac{1}{6} x = 3.5$$

(1+2+3+4+5+6)/6=3.5

Joint Probability

- Consider two random variables:



Joint probability $p(c,t)$:

$p(c,t)$	$T=1$	$T=0$
$C=1$	0.1	0.02 p(c=1,t=0)=0.02
$C=0$	0.08	0.8

- Joint probabilities sum to 1:

$$\sum_c \sum_t p(c,t) = 1$$

Take Home Message – Probability:

Always true:

- Marginalisation: $p(t) = \sum_c p(t, c)$
  Probability of tooth ache
- Cond. prob.: $p(c|t) = \frac{p(c, t)}{p(t)}$
  Probability of cavity or tooth ache
- Product rule: $p(c, t) = p(c|t)\, p(t)$
  Probability of cavity and tooth ache
- Bayes rule: $p(c|t) = \frac{p(t|c)p(c)}{p(t)}$
  Probability of

Derive everything from

$p(t, c)$:

	$\boldsymbol{T = 1}$	$\boldsymbol{T = 0}$
$\boldsymbol{C = 1}$	0.1	0.02
$\boldsymbol{C = 0}$	0.08	0.8

Joint probability is complete model

Take Home Message – Probability (2):

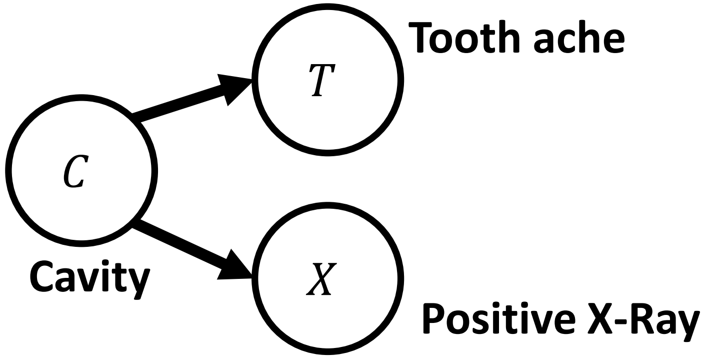
Everything holds when conditioned on additional variable:

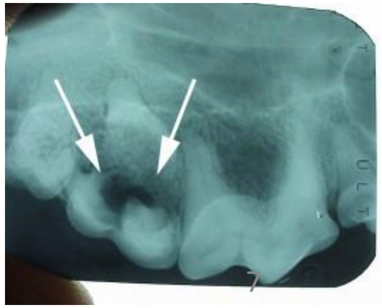
- Marginalisation: $p(t|x) = \sum_c p(t, c|x)$
- Cond. Prob.: $p(c|t, x) = \dfrac{p(c, t|x)}{p(t|x)}$
- Product Rule: $p(c, t|x) = p(c|t, x)\, p(t|x)$
- Bayes Rule: $p(c|t, x) = \dfrac{p(t|c, x)p(c|x)}{p(t|x)}$

Derive everything from

$p(t, c|X = 1)$:

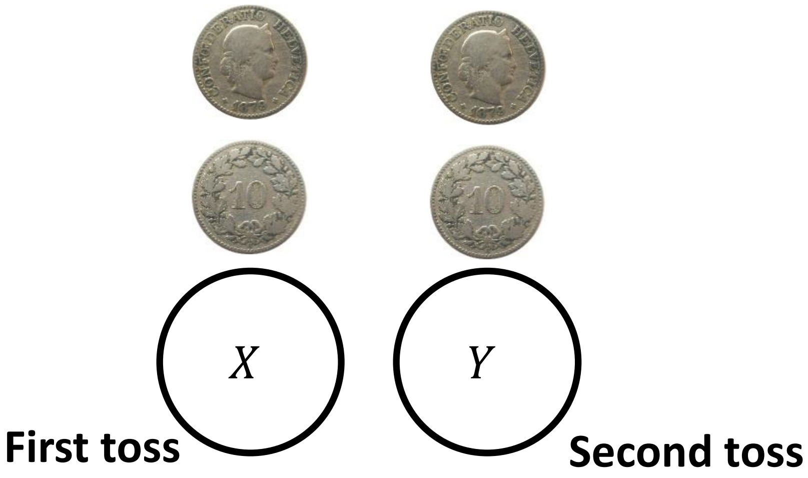
	$\boldsymbol{T = 1}$	$\boldsymbol{T = 0}$
$\boldsymbol{C = 1}$	0.1	0.02
$\boldsymbol{C = 0}$	0.08	0.8





Independence

- Example: Toss a coin twice



$p(x)$:

$p(X=1)$	$p(X=0)$
0.5	0.5

$p(y)$:

$p(Y=1)$	$p(Y=0)$
0.5	0.5

Can calculate

$p(x,y) = p(x)p(y)$

	$X=1$	$X=0$
$Y=1$	0.25	0.25
$Y=0$	0.25	0.25

- Two random variables X and Y are **independent** iff:

  $p(x|y) = p(x)$, $\quad p(y|x) = p(y)$, $\quad p(x,y) = p(x)\,p(y)$ for all x and y.

  - Observing one does not give information about the other

Written:
$X \perp\!\!\!\perp Y$

Conditional Independence

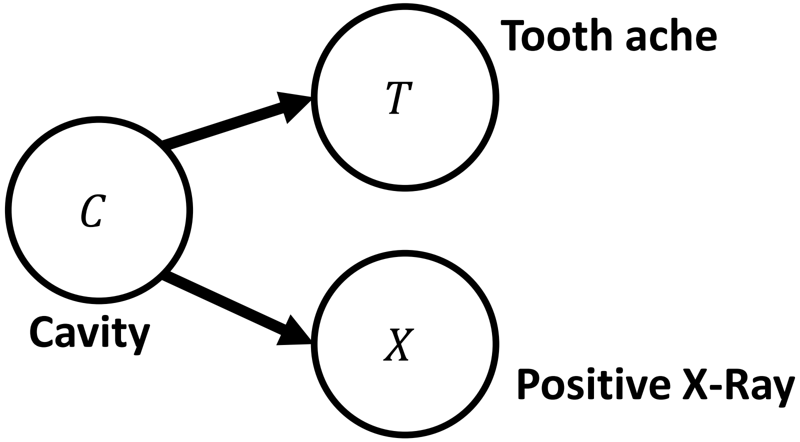
- X and T are **conditionally independent** given C iff:

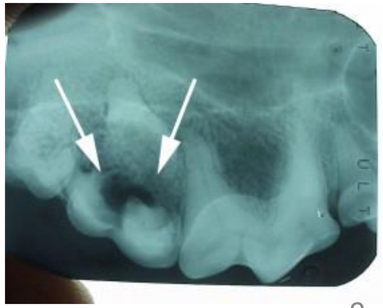
Written:
$X \perp\!\!\!\perp T \mid C$

$$p(x|t,c) = p(x|c), \qquad p(t|x,c) = p(t|c), \qquad p(x,t|c) = p(x|c)\, p(t|c) \quad \text{for all } x, t \text{ and } c.$$

- Observing one does not give information about the other, provided C is observed.

- Example: Dentist





Continuous Probability Distributions

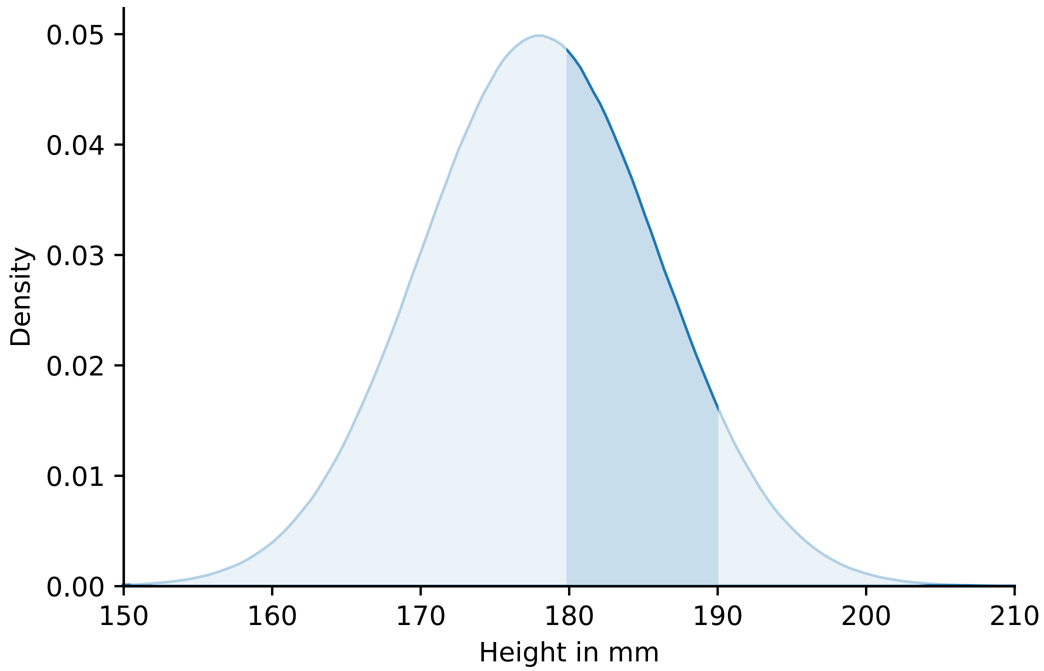
Σ

- $\underline{\int_{-\infty}^{\infty}} f(x)\, dx = 1$

Calculus

- Large values -> large sample density
- Completely describes distribution

- $\int_a^b f(x)\, dx$ is probability $a \leq x_i \leq b$



Probability Density Function (PDF)

Take Home Message – Continuous Probability:

Always true:

These are PDFs

- Marginalisation: $p(t) = \int_{-\infty}^{\infty} p(t,c)\, dc$
- Cond. prob.: $p(c|t) = \frac{p(c,t)}{p(t)}$
- Product rule: $p(c,t) = p(c|t)\, p(t)$
- Bayes rule: $p(c|t) = \frac{p(t|c)p(c)}{p(t)}$

Derive everything from

$p(t,c)$:

Expected value:

$$\mathbb{E}_{p(x)}[x] = \int_{-\infty}^{\infty} p(x)x = \lim_{K\to\infty} \frac{1}{K} \sum_{k=1}^{K} x_k \; | x_k \sim p(x)$$

(Cond.) Independence

Like in discrete case