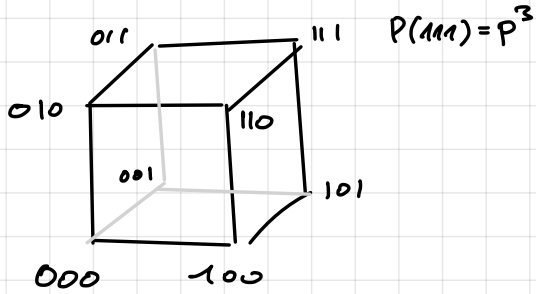
Example 3.4 $n=3$, $q=2$ $\{0,1\}^3$



$\Omega = \{0,1\}^3$ noise that was observed

$\omega = (\omega_1, \omega_2, \omega_3) \in \Omega$ with

$\omega_i = 1$ if there was noise in the i-th trial, $\omega_i = 0$ if there was no noise.

If $x \in \{0,1\}^3$ is the word we send,

$x = (x_1, x_2, x_3)$, then

$$x \oplus \omega = (x_1 \oplus \omega_1, x_2 \oplus \omega_2, x_3 \oplus \omega_3)$$

$$(x, \omega) \longmapsto \underbrace{x \oplus \omega}_{\text{randomly disturbed } x.}$$

$\oplus$ is another notation for XOR, also called addition modulo two.

The probability of the noise occuring in the ith symbol is set to be some number

$p \in (0,1)$.

$010 \rightsquigarrow 110$ $\quad p = \frac{1}{4}$

$3 \cdot \frac{1}{4} \cdot \left(\frac{3}{4}\right)^2$ the probability of exactly error.

$$P(1 \text{ error}) = 3 \cdot p\,(1-p)^2$$

P (exactly one error, which is in the first position)

$= p \cdot (1-p)^2$

Def 3.5 We introduce the probability space that models n independent coin tosses with the fair probability equal to $p \in (0,1)$.

$\Omega = \{0,1\}^n$ with

$P(\omega_1 \ldots \omega_n) = p^{wt(\omega)} \cdot (1-p)^{n-wt(\omega)}$, where

$\omega_1 \in \{0,1\}$, $\omega_n \in \{0,1\}$

$\omega = \omega_1 \ldots \omega_n$ and

$wt(\omega) = \omega_1 + \ldots + \omega_n$.

The probability space is the pair (Ω, P).

Def 3.6 Events $A, B \subseteq \Omega$ are called independent if $P(A \cap B) = P(A)\,P(B)$

A system of events $A_1, \ldots, A_m \subseteq \Omega$ is called independent if

$$P(A_{i_1} \cap \ldots \cap A_{i_k}) = P(A_{i_1}) \cdot \ldots \cdot P(A_{i_k})$$

for all $1 \le i_1 < \ldots < i_k \le m$.

Def 3.7 If S is a set, then $X: \Omega \to S$ is called a random variable with values in S.

Notation: $X \in_R S$. The function

$$s \in S \longmapsto P(X=s) = P(\{\omega \in \Omega : X(\omega) = s\})$$

is called the distribution

Ex 3.8 $\Omega = \{0,1\}^3 = \{000, 001, 010, 011, 100, 101, 110, 111\}$

$P(\omega) = \frac{1}{8}$ for every $\omega \in \Omega$.

(3 independent tosses of a fair coin).

$X(\omega) = X(\omega_1, \omega_2, \omega_3) = \omega_1 + \omega_2 + \omega_3$.

X is the number of tails tossed

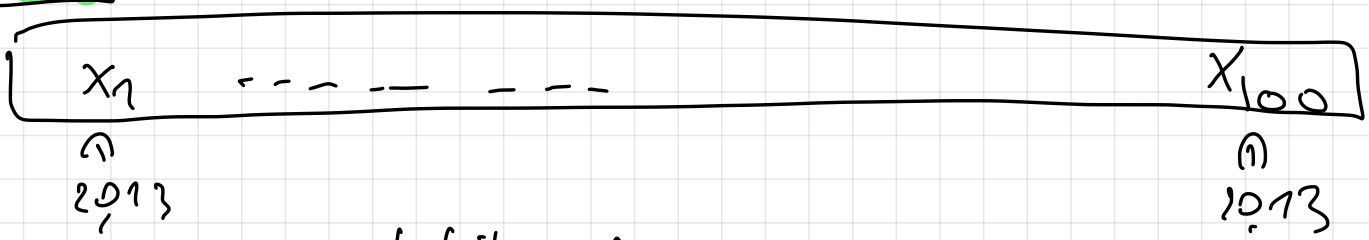
$P(X=0) = P(\{000\}) = \frac{1}{8}$

$P(X=1) = P(\{100, 010, 001\}) = \frac{3}{8}$

$P(X=2) = P(\{011, 101, 011\}) = \frac{3}{8}$

$P(X=3) = P(\{111\}) = \frac{1}{8}$

Ex 3.9



probability of error in any given position is 5%.

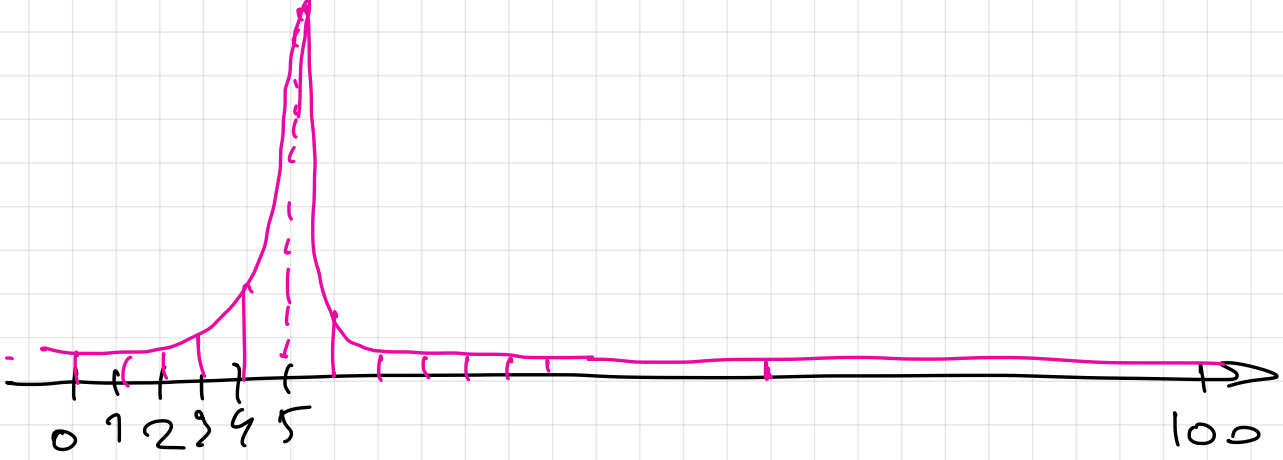
n tosses of a biased coin with tails prob. equal to $p = 0,05$

$n = 100$.

$X(\omega_1, \ldots \omega_n) = \omega_1 + \ldots + \omega_n$

The number of tails thrown.

How does the distribution of X look like?



Def 3.10 A random variable X is called uniformly distributed in S if

$$P(X=s) = \frac{1}{|S|} \quad (S \text{ is a finite set})$$

holds for all $s \in S$.

Def 3.11 If $A, B \subseteq \Omega$ are events with $P(B) > 0$, then

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

is called the conditional probability, probability of A given B.

Remark 3.12 If $A, B \subseteq \Omega$ and $P(B) > 0$, then the independence of A and B is equivalent to $P(A \mid B) = P(A)$.

Theorem 3.13 (Bayes Formula)

For events $A, B \subseteq \Omega$ with $P(A), P(B) > 0$ one has

$$P(A|B) = \frac{P(B|A)}{P(B)} P(A)$$

Proof:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A \cap B)}{P(A) \cdot P(B)} \cdot P(A)$$

$$= \frac{P(B|A)}{P(B)} \cdot P(A) \qquad \square$$

Remark 3.14 It's important not to confuse $P(A|B)$ and $P(B|A)$, because these values are in general not the same.

$P(\text{NBA players} \mid \text{higher than 2m})$ vs.

$P(\text{higher than 2m} \mid \text{NBA players})$

If you know $P(\text{higher than 2m})$ and $P(\text{NBA player})$, then you would

connect these numbers above by the Bayes formula.

$P(\text{spam} \mid \text{message contains "you won!"})$

$P(\text{message contains "you won!"} \mid \text{spam})$

$P(y \text{ was received} \mid x \text{ was sent})$

$P(x \text{ was sent} \mid y \text{ was received})$

↑ We want to know this one.

Prop 3.15 If Ω is a disjoint union of events $B_1, \dots, B_m$ and $A \subseteq \Omega$ is an event, then

$$P(A) = \sum_{i=1}^{m} P(A \mid B_i) \cdot P(B_i),$$

where the terms with $P(B_i) = 0$ are interpreted as zero.

Proof:

$$\sum_{i=1}^{m} P(A \mid B_i) \cdot P(B_i) = \sum_{i=1}^{m} \frac{P(A \cap B_i)}{P(B_i)} \cdot P(B_i)$$

$$= \sum_{i=1}^{m} P(A \cap B_i) = P\big((A \cap B_1) \cup \dots \cup (A \cap B_m)\big)$$

$A \cap B_i$ are disjoint, because $B_1, \dots, B_m$ are disjoint

$$= P(A \cap (B_1 \cup \dots \cup B_m))$$
$$= P(A \cap \Omega) = P(A). \qquad \square$$

Def 3.16. Random variables $X_1, \dots, X_m$ distributed in sets $S_1, \dots, S_m$ respectively are called independent if

$$P(X_1 \in T_1, X_2 \in T_2, \dots, X_m \in T_m) = \prod_{i=1}^{m} P(X_i \in T_i)$$

for all $T_1 \subseteq S_1, \dots, T_m \subseteq S_m$.

Def 3.17 For a random variable $X: \Omega \to \mathbb{R}$, the expectation of X is

$$E(X) = \sum_{\omega \in \Omega} X(\omega) \cdot P(\omega).$$

and the variance is

$$V(X) = E\left((X - E(X))^2\right)$$

$\sqrt{V(X)}$ is called the standard deviation.

LQ7 Recall the probability theory