Repetition
Probability Theory

Algorithm Theory
WS 2017/18

Fabian Kuhn

Randomized Algorithms

Randomized Algorithms

- An algorithm that uses (or can use) random coin flips in order to make decisions
- randomization can be a powerful tool to make algorithms faster or simpler

First: Short Repetition of Basic Probability Theory

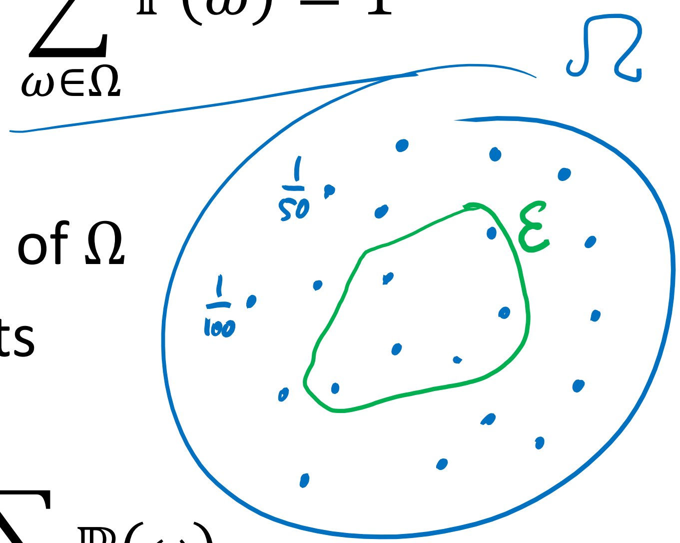
- We need: basic discrete probability theory
  - probability spaces, probability events, independence, random variables, expectation, linearity of expectation, Markov inequality
- Literature, for example
  - your old probability theory book / lecture notes / ...
  - Appendix C of book of Cormen, Rivest, Leiserson, Stein
  - http://www.ti.inf.ethz.ch/ew/courses/APC15/material/ra.pdf

Probability Space and Events

Definition: A (discrete) **probability space** is a pair $(\Omega, \mathbb{P})$, where

- Ω: (countable) set of elementary events
- $\mathbb{P}$: assigns a probability to each $\omega \in \Omega$

$$\mathbb{P} : \Omega \to \mathbb{R}_{\geq 0} \quad \text{s.t.} \quad \sum_{\omega \in \Omega} \mathbb{P}(\omega) = 1$$



Definition: An **event** $\mathcal{E}$ is a subset of Ω

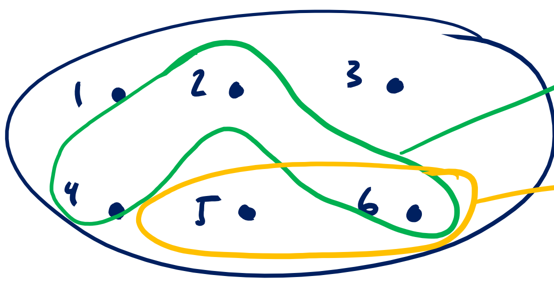
- Event $\mathcal{E} \subseteq \Omega$: set of basic events
- Probability of $\mathcal{E}$

$$\mathbb{P}(\mathcal{E}) := \sum_{\omega \in \mathcal{E}} \mathbb{P}(\omega)$$

Example: Probability Space, Events

roll a die

$\Omega = \{1,2,3,4,5,6\}$, $\mathbb{P}(1) = \mathbb{P}(2) = \ldots = \mathbb{P}(6) = 1/6$



$\mathcal{E}_{even}$ $\quad \mathbb{P}(\mathcal{E}_{even}) = \frac{1}{2}$

$\mathcal{E}_{5,6}$ $\quad \mathbb{P}(\mathcal{E}_{5,6}) = \frac{1}{3}$

roll 2 dice

$\Omega = \{(1,1), (1,2), \ldots, (1,6), (2,1), \ldots, (6,6)\}$ $\quad \mathbb{P}((i,j)) = \frac{1}{36}$

$A_= = \{(1,1), (2,2), \ldots, (6,6)\}$ $\quad \mathbb{P}(A_=) = 6/36 = \frac{1}{6}$

$A_{\neq} = \Omega \setminus A_= = \overline{A_=}$ $\quad \mathbb{P}(\overline{A_=}) = 1 - \mathbb{P}(A_=) = \frac{5}{6}$

Example: Probability Space, Events

flip (biased) coin $\{H, T\}$

prob. to get H is equal to p

experiment: flip coins until we get H

$$\Omega = \{H, TH, TTH, \dots, \underbrace{TTT\dots T}_{\infty}\}$$

$H \to e_0$, $TH \to e_1$, $TTH \to e_2$

$$\mathbb{P}(e_i) = (1-p)^i \cdot p$$

$$\sum_{i=0}^{\infty} \mathbb{P}(e_i) = p \cdot \underbrace{\sum_{i=0}^{\infty} (1-p)^i}_{\frac{1}{1-(1-p)}} = p \cdot \frac{1}{p} = 1$$

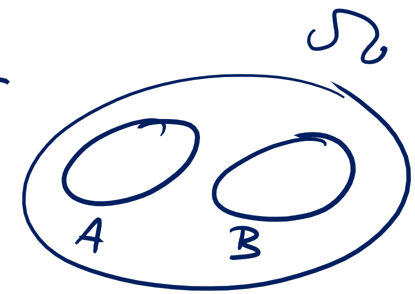
$$\mathcal{E} = \{e_i \mid i \text{ is even}\}$$

$$\mathbb{P}(\mathcal{E}) = \sum_{j=0}^{\infty} \mathbb{P}(e_{2j}) = p \cdot \sum_{j=0}^{\infty} \underbrace{(1-p)^{2j}}_{((1-p)^2)^j} = p \cdot \frac{1}{1-(1-p)^2} = p \frac{1}{2p - p^2} = \frac{1}{2-p}$$

Independent Events

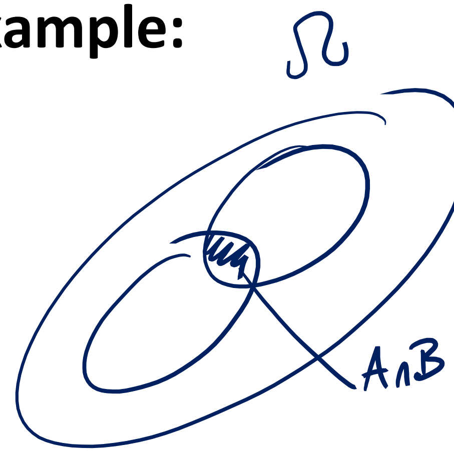
Definition: Events $\mathcal{A} \subseteq \Omega$ and $\mathcal{B} \subseteq \Omega$ are **independent** iff

$$\underline{\mathbb{P}(\mathcal{A} \cap \mathcal{B}) = \mathbb{P}(\mathcal{A}) \cdot \mathbb{P}(\mathcal{B})}$$



A and B disjoint

Example:



roll 2 dice

A: first die is even

B: second die is odd

$A \cap B = \{(2,1), (2,3), (2,5), (4,1), (4,3), (4,5), (6,1), (6,3), (6,5)\}$

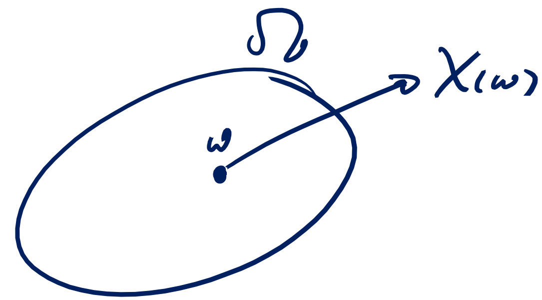
$|A \cap B| = 9 \qquad \mathbb{P}(A \cap B) = \frac{9}{36} = \frac{1}{4}$

$\Omega = \Omega_1 \times \Omega_2$

Random Variables

Definition: A **random variable $\boldsymbol{X}$** is a real-valued function on the elementary events Ω

$$\boldsymbol{X} : \boldsymbol{\Omega} \to \mathbb{R}$$



- We usually write X instead of $X(\omega)$
- We also write

$$\mathbb{P}(X = x) = \mathbb{P}(\{\omega \in \Omega : X(\omega) = x\})$$

Examples:

- $\boldsymbol{X^{top}}$: $X^{top}(1) = 1, X^{top}(2) = 2, \dots, X^{top}(6) = 6$
- $\boldsymbol{X^{bot}}$: $X^{bot}(1) = 6, X^{bot}(2) = 5, \dots, X^{bot}(6) = 1$

- Note that for all $\omega \in \Omega$, $X^{top}(\omega) + X^{bot}(\omega) = 7$
- To denote this, we write $X^{top} + X^{bot} = 7$

Indicator Random Variables

A random variable with only takes values 0 and 1 is called a **Bernoulli random variable** or an **indicator random variable**.

roll a die, rand var. $Y = \begin{cases} 1 & \text{if even} \\ 0 & \text{otherwise} \end{cases}$

$Y(1) = 0,\ Y(2) = 1,\ Y(3) = 0,\ \dots$

$\mathbb{P}(Y = 1) = \frac{1}{2}$

Independent Random Variables

Definition: Two random variables X and Y are called **independent** if

$$\forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R} : \ \mathbb{P}(\boldsymbol{X} = \boldsymbol{x} \wedge \boldsymbol{Y} = \boldsymbol{y}) = \mathbb{P}(\boldsymbol{X} = \boldsymbol{x}) \cdot \mathbb{P}(\boldsymbol{Y} = \boldsymbol{y})$$

two coin flips (fair coin) Bernoulli rand. var X and Y

$X = 1 \iff$ 1st coin flip is H

$Y = 1 \iff$ exactly one coin flip is H

$\mathbb{P}(X=0 \wedge Y=0) = \mathbb{P}(\{TT\}) = \frac{1}{4}$

$\mathbb{P}(X=0 \wedge Y=1) = \mathbb{P}(\{TH\}) = \frac{1}{4}$

$\mathbb{P}(X=1 \wedge Y=0) = \mathbb{P}(\{HH\}) = \frac{1}{4}$

$\mathbb{P}(X=1 \wedge Y=1) = \mathbb{P}(\{HT\}) = \frac{1}{4}$

Independent Random Variables

$\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A) \cdot \mathbb{P}(B) \cdot \mathbb{P}(C)$

Definition: A collection of random variables $X_1, X_2, \dots, X_n$ on a probability space Ω is called **mutually independent** if

$$\forall k \geq 2, 1 \leq i_1 < \cdots < i_k \leq n, \forall x_{i_1}, \dots, x_{i_k} \in \mathbb{R}:$$
$$\mathbb{P}\left(X_{i_1} = x_{i_1} \wedge \cdots \wedge X_{i_k} = x_{i_k}\right) = \mathbb{P}\left(X_{i_1} = x_{i_1}\right) \cdot \ldots \cdot \mathbb{P}\left(X_{i_k} = x_{i_k}\right)$$

not the same as pairwise independence

example: 2 coin flips

$X_1 = 1 \iff$ 1st flip is H

$X_2 = 1 \iff$ 2nd flip is H

$X_3 = 1 \iff$ exactly one H

$\mathbb{P}(X_1 = 1 \wedge X_2 = 1 \wedge X_3 = 1) = 0$

Expectation

Definition: The **expectation** of a random variable X is defined as

$$\mathbb{E}[X] := \sum_{x \in X(\Omega)} x \cdot \mathbb{P}(X = x) = \sum_{\omega \in \Omega} X(\omega) \cdot \mathbb{P}(\omega)$$

Example:

- recall: X^{top} is outcome of rolling a die

Expectation: Examples

Sums and Products of Random Variables

Linearity of Expectation:

For random variables X and Y and any $c \in \mathbb{R}$, we have

$$\begin{aligned} \mathbb{E}[cX] &= c \cdot \mathbb{E}[X] \\ \mathbb{E}[X+Y] &= \mathbb{E}[X] + \mathbb{E}[Y] \end{aligned}$$

- holds also if the random variables are not independent

Product of Random Variables:

For two **independent** random variables X and Y, we have

$$\mathbb{E}[X \cdot Y] = \mathbb{E}[X] \cdot \mathbb{E}[Y]$$

Sums and Products of Random Variables

Linearity of Expectation:

For random variables X and Y and any $c \in \mathbb{R}$, we have

$$\mathbb{E}[\boldsymbol{cX}] = \boldsymbol{c} \cdot \mathbb{E}[\boldsymbol{X}], \qquad \mathbb{E}[\boldsymbol{X} + \boldsymbol{Y}] = \mathbb{E}[\boldsymbol{X}] + \mathbb{E}[\boldsymbol{Y}]$$

Sums and Products of Random Variables

Product of Random Variables:

For two **independent** random variables X and Y, we have

$$\mathbb{E}[\boldsymbol{X} \cdot \boldsymbol{Y}] = \mathbb{E}[\boldsymbol{X}] \cdot \mathbb{E}[\boldsymbol{Y}]$$

Linearity of Expectation: Example

Sequence of coin flips: $C_1, C_2, \ldots \in \{H, T\}$

$\mathbb{P}(H) = p$

e_i

- Stop as soon as the first H turns up

Random variable $\boldsymbol{X}$: number of T before first H

$\mathbb{P}(X=i) = p(1-p)^i$

$\mathbb{E}[X]$

Indicator random variable $\boldsymbol{X_i}$ ($\boldsymbol{i \geq 1}$):

- $X_i = 1$: i^{th} coin flip happens and its outcome is T
  $X_i = 0$: otherwise

$X = X_1 + X_2 + \ldots + X_\infty$

$\mathbb{P}(X_i = 1) = (1-p)^i \qquad \mathbb{E}[X_i] = (1-p)^i$

$\mathbb{E}[X] = \mathbb{E}[X_1 + X_2 + \ldots + X_\infty]$

$= \sum_{i=1}^{\infty} \mathbb{E}[X_i] = \sum_{i=1}^{\infty} (1-p)^i = (1-p) \cdot \frac{1}{1-(1-p)}$

lin. of exp.

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

$\mathbb{E}[X] = \sum_{i=0}^{\infty} i \cdot p \cdot (1-p)^i$

$= \ldots$

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

Markov's Inequality

Lemma: Let X be a nonnegative random variable.
Then for all $c > 0$

$$\mathbb{P}(X \geq c \cdot \mathbb{E}[X]) \leq \frac{1}{c}$$

$$\mathrm{Var}(X) := \mathbb{E}\big[\underbrace{(X - \mathbb{E}[X])^2}_{Z \geq 0}\big] \qquad \mathbb{P}(Z \geq c^2 \cdot \underbrace{\mathbb{E}[Z]}_{\mathrm{Var}(X)}) \leq \frac{1}{c^2}$$

$$\mathbb{P}\big((X - \mathbb{E}[X])^2 \geq c^2 \cdot \mathrm{Var}(X)\big) \leq \frac{1}{c^2}$$

$$\underbrace{\mathbb{P}\big(|X - \mathbb{E}[X]| \geq c \cdot \sigma(X)\big) \leq \frac{1}{c^2}}_{\text{Chebyshev's ineq.}}$$

$$\sigma(X) := \sqrt{\mathrm{Var}(X)}$$

Conditional Probabilities

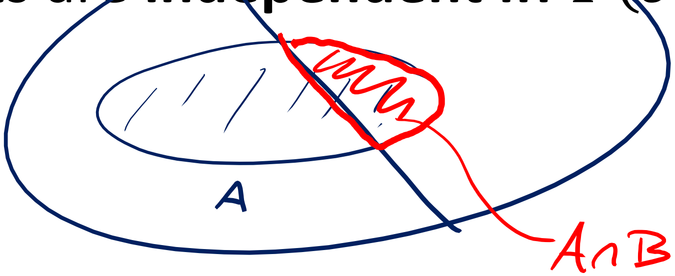
For events $\mathcal{A} \subseteq \Omega$ and $\mathcal{B} \subseteq \Omega$, the **conditional probability** of $\mathcal{A}$ given $\mathcal{B}$ is defined as

$$\underline{\mathbb{P}(\mathcal{A}|\mathcal{B})} := \frac{\mathbb{P}(\mathcal{A} \cap \mathcal{B})}{\mathbb{P}(\mathcal{B})}$$

Conditioning on event $\mathcal{B}$ defines a new probability space $(\cancel{\Omega \setminus} \mathcal{B}, \mathbb{P}')$

$$\forall \omega \in \Omega \setminus B : \ \mathbb{P}'(\omega) = \frac{\mathbb{P}(\omega)}{\mathbb{P}(\mathcal{B})}.$$

Two events are **independent iff** $\mathbb{P}(\mathcal{A}|\mathcal{B}) = \mathbb{P}(\mathcal{A})$



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

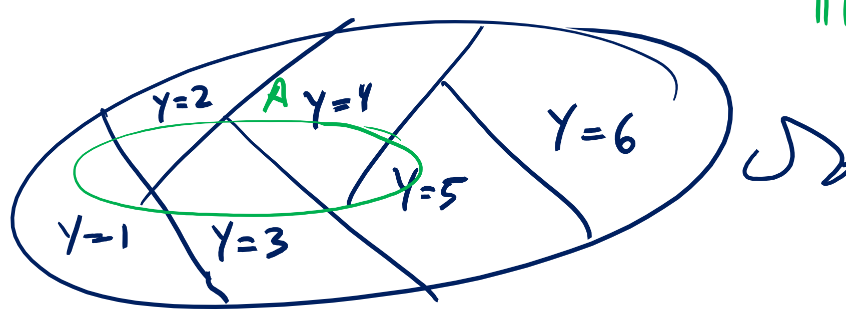
$\iff A, B$ indep.

Law of Total Probability / Expectation

Lemma: Let X and Y be two random variables on the same probability space $(\Omega, \mathbb{P})$. We then have

$$\forall \boldsymbol{x} \in \mathbb{R} : \mathbb{P}(\boldsymbol{X} = \boldsymbol{x}) = \sum_{\boldsymbol{y} \in \boldsymbol{Y}(\boldsymbol{\Omega})} \mathbb{P}(\boldsymbol{X} = \boldsymbol{x} \mid \boldsymbol{Y} = \boldsymbol{y}) \cdot \mathbb{P}(\boldsymbol{Y} = \boldsymbol{y}).$$

$$\mathbb{E}[X] = \sum_{y \in \mathrm{Y}(\Omega)} \underbrace{\mathbb{E}[X \mid Y = y]}_{:= \sum_x x \cdot \mathbb{P}(X=x \mid Y=y)} \cdot \mathbb{P}(Y = y)$$



$$\mathbb{P}(A) = \mathbb{P}(A \mid Y=1) \cdot \mathbb{P}(Y=1)$$
$$+ \mathbb{P}(A \mid Y=2) \cdot \mathbb{P}(Y=2)$$
$$+ \ldots$$

Important Discrete Prob. Distributions

Bernoulli Random Variable $\boldsymbol{X : \Omega \to \{0,1\}}$

$$\mathbb{P}(X = 1) = p, \mathbb{P}(X = 0) = 1 - p, \qquad \mathbb{E}[X] = p$$

Binomial Random Variable $\boldsymbol{X \sim \mathrm{Bin}(n,p)}$

$$\forall k \in \{0, \dots, n\} : \ \mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \qquad \mathbb{E}[X] = np$$

- measures number of ones in n independent biased coin flip

Geometric Random Variables $\boldsymbol{X \sim \mathrm{Geom}(p)}$

$$\forall k \geq 1 : \ \mathbb{P}(X = k) = p(1-p)^{k-1}, \qquad \mathbb{E}[X] = \frac{1}{p}$$

- measures number independent biased coin flips are necessary to get one “heads”