Repetition
Probability Theory

Algorithm Theory
WS 2016/17

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

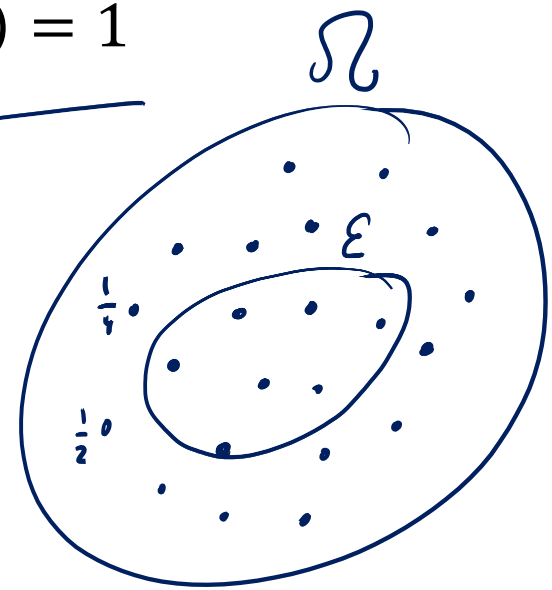
Probability Space and Events

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 $\omega \in \Omega$

$\mathbb{P}(\omega) \geq 0$

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



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

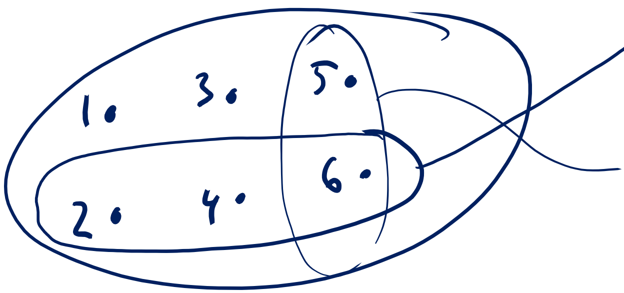
- Event $\mathcal{E} \subseteq \Omega$: set of basic events
- Probability of $\mathcal{E}$

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

Example: Probability Space, Events

roll a die

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



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

$\mathcal{E}_{\geq 5}$ $\quad \mathbb{P}(\mathcal{E}_{\geq 5}) = 1/3$

roll 2 dice

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

$A_= = \{(1,1), (2,2), (3,3), \ldots, (6,6)\}$ $\quad \mathbb{P}(A_=) = 6/36 = 1/6$

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

Example: Probability Space, Events

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

prob. to get H is p

experiment: flip coins until we get H

$$\Omega = \{\underbrace{H}_{e_0}, \underbrace{TH}_{e_1}, \underbrace{TTH}_{e_2}, \dots, \underbrace{\underbrace{TT\dots T}_{\infty}}_{e_\infty}\}$$

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

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

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

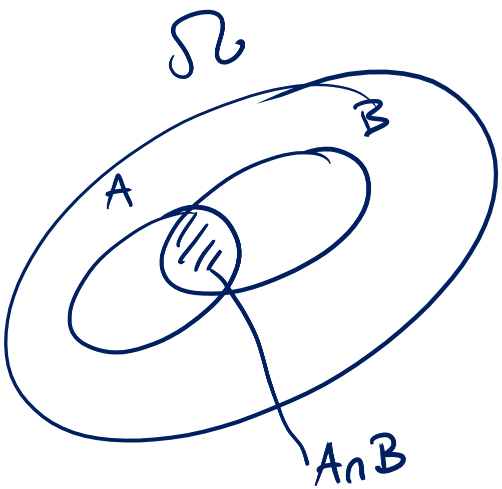
Independent Events

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

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

$1/4 \qquad\qquad 1/2 \qquad 1/2$

Example:



roll 2 dice

A: first die is even

B: second die is odd

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

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

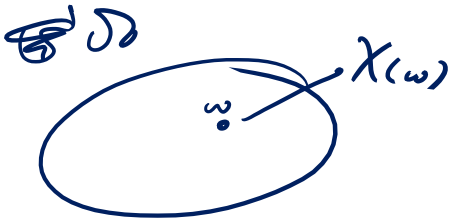
$\Omega = \Omega_1 \times \Omega_2$

1st die 2nd die

Random Variables

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

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



$\Omega = \{1, \dots, 6\}$

- 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. $Y = \begin{cases} 1 & \text{if even} \\ 0 & \text{if odd} \end{cases}$

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

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)

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

$\{Y=1\} = \{HT, TH\}$

$Y = 1 \iff$ exactly one of the coin flips is H

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

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

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

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 andom 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 indep.

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

$$\mathbb{E}[X^{top}] = \sum_{i=1}^{6} i \cdot \frac{1}{6} = \frac{21}{6} = 3{,}5$$

$$X^2(\omega) = (X(\omega))^2$$

$$\mathbb{E}[X^2] \neq \mathbb{E}[X]^2$$

$$\mathbb{E}[X^{top^2}] = \sum_{i=1}^{6} i^2 \cdot \frac{1}{6} = \frac{1+4+9+\ldots+36}{6} = \frac{91}{6} = 15.16\ldots$$

$$\mathbb{E}[X^{top} X^{bot}] = \frac{1\cdot 6 + 2\cdot 5 + 3\cdot 4 + 4\cdot 3 + 5\cdot 2 + 6\cdot 1}{6} = \frac{28}{6} = 9.33\ldots$$

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

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

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

- 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

$\mathbb{P}(Y=1) = p$

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

Sequence of coin flips: $C_1, C_2, \dots \in \{H, T\}$ $\quad e_0, e_1$

- Stop as soon as the first H turns up

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

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

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

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

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

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

$X = X_1 + X_2 + X_3 + \dots + X_\infty$

$\mathbb{E}[X] = \mathbb{E}[X_1 + X_2 + \dots] = \sum_{i=1}^{\infty} \mathbb{E}[X_i] = \sum_{i=1}^{\infty} (1-p)^i = \frac{1-p}{1-(1-p)} = \frac{1-p}{p}$

Markov's Inequality

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

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

$$\mathrm{Var}(X) := \mathbb{E}\big[\underbrace{(X - \mathbb{E}(X))^2}_{Z \geq 0}\big]$$

$$\mathbb{P}(Z \geq c \cdot \underbrace{\mathbb{E}[Z]}_{\mathrm{Var}(X)}) \leq \frac{1}{c}$$

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

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

Conditional Probabilities

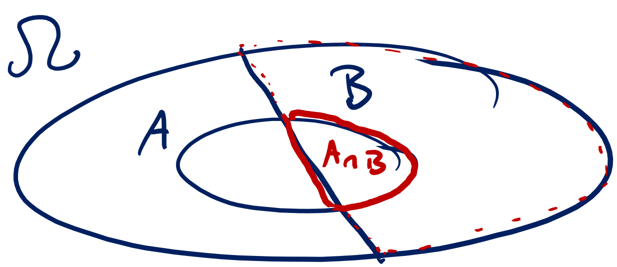
For events $\mathcal{A} \subseteq \Omega$ and $\mathcal{B} \subseteq \Omega$, the **conditional probability** of $\mathcal{A}$ given $\mathcal{B}$ is defined as

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

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

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

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



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

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

$$\mathbb{P}(A) = \mathbb{P}(\Omega_1) \cdot \mathbb{P}(A \mid \Omega_1) + \mathbb{P}(\Omega_2) \cdot \mathbb{P}(A \mid \Omega_2) + \ldots$$

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

Important Discrete Prob. Distributions

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

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

$X = X_1 + \cdots + X_n$

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

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

- measures number of ones in n independent biased coin flip

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

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

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