Repetition
Probability Theory

Algorithm Theory
WS 2018/19

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 $\mathcal{E}$

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

Example: Probability Space, Events

Example: Probability Space, Events

Independent Events

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

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

Example:

Random Variables

Definition: A **random variable X** is a real-valued function on the elementary events Ω

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

- 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**.

Independent Random Variables

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

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

Independent Random Variables

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

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

Expectation

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

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

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: $C_1, C_2, \ldots \in \{H, T\}$

- Stop as soon as the first H turns up

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

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

Markov's Inequality

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

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

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

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

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

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

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

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

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

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