Algorithms and Probability

Week 8

G08 - mkilic

10.IV.2025

Overview

1. **Minitest**
2. **Mehrere Zufallsvariablen**
3. **Abschätzen von Wahrscheinlichkeiten**
4. **Randomisierte Algorithmen**
5. **Exercises**

Roadmap

1. Graphentheorie
   - Zusammenhang
   - Kreise
   - Matchings
   - Färbungen
2. W'keitstheorie
   - Bedingte W'keit
   - Unabhängigkeit
   - (mehrere) Zufallsvariablen
   - Diskrete Verteilungen
   - Abschätzen von W'keiten
   - Randomisierte Algorithmen
3. Algorithmen
   - Lange-Bunte Pfade
   - MaxFlow
   - MinCut
   - Kleinster umschliessender Kreis
   - Konvexe Hülle

Minitest

Passwort: variance

Gemeinsame Dichte und Randdichte

Gemeinsame Dichte

$$f_{X,Y}(x,y) = \Pr[X = x, Y = y]$$

"Randdichte von X" (= Dichte von X)

$$f_X(x) = \sum_{y \in \mathcal{W}_Y} f_{X,Y}(x,y) \quad \text{(nach dem Additionssatz)}$$

[0] *Engl. gemeinsame Dichte: Joint distribution, Randdichte: marginal distribution*

Gemeinsame Verteilung und Randverteilung

gemeinsame Verteilung

$$F_{X,Y}(x,y) := Pr[X \leq x, Y \leq y] = Pr[\omega \in \Omega | X(\omega) \leq x, Y(\omega) \leq y]$$
$$= \sum_{x' \leq x} \sum_{y' \leq y} f_{X,Y}(x', y')$$

Randverteilung von X

$$F_X(x) = \sum_{x' \leq x} f_X(x) = \sum_{x' \leq x} \sum_{y \in \mathcal{W}_y} f_{X,Y}(x', y)$$

[0] *Engl. gem. Verteilung: joint cumulative distr., Randverteilung: marginal cumulative distribution*

Bedingte Zufallsvariablen

Bedingte Zufallsvariable

Sei X Zufallsvariable auf W'raum Ω, $A \subseteq \Omega$ Ereignis mit $\Pr[A] > 0$. Die bedingte Zufallsvariable $X|A$ ist dieselbe Funktion wie X, aber der Definitionsbereich ist auf die Menge A eingeschränkt.

$$f_{X|A} : \mathbb{R} \to [0,1], \quad x \mapsto \Pr[X = x \mid A]$$

X ist unabhängig von A, falls $f_{X|A} = f_X$.

$$F_{X|A} : \mathbb{R} \to [0,1], \quad x \mapsto \Pr[X \leq x \mid A]$$

$$\mathbb{E}[X \mid A] := \sum_{x \in \mathcal{W}_X} x \cdot \Pr[X = x \mid A] = \frac{1}{\Pr[A]} \sum_{\omega \in A} X(\omega) \cdot \Pr[\omega]$$

Erinnerung: $X = x$ ist auch ein Ereignis, also $''X = x \subseteq \Omega''$

Unabhängigkeit von Zufallsvariablen

Unabhängigkeit von Zufallsvariablen

Die Ereignisse $X_1 = x_1, \ldots, X_n = x_n$ heißen unabhängig, wenn für alle $x_1, \ldots, x_n$ die Zufallsvariablen $X_1, \ldots, X_n$ unabhängig sind, dh.

$$\Pr[X_1 = x_1, \ldots, X_n = x_n] = \Pr[X_1 = x_1] \cdots \Pr[X_n = x_n].$$

Aber Viele der Gleichungen sind redundant. Siehe Beispiel in Vorlesungsslides.
Alternativ:

$$f_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = f_{X_1}(x_1) \cdots f_{X_n}(x_n)$$

für alle $(x_1, \ldots, x_n) \in \mathcal{W}_{X_1} \times \cdots \times \mathcal{W}_{X_n}$

Andere Definitionen und Korollare

Lemma 2.53.

Sind $X_1, \ldots, X_n$ unabhängige Zufallsvariablen und sind $S_1, \ldots, S_n \subseteq \mathbb{R}$ beliebige Mengen, dann gilt

$$Pr[X_1 \in S_1, \ldots, X_n \in S_n] = Pr[X_1 \in S_1] \cdot \cdots \cdot Pr[X_n \in S_n]$$

Korollar 2.54

Sind $X_1, \ldots X_n$ unabhängige Zufallsvariablen und ist $I = \{i_1, \ldots, i_k\} \subseteq [n]$, dann sind $X_{i_1}, \ldots, X_{i_k}$ ebenfalls unabhängig.

Funktionen von Zufallsvariablen

Anwendung einer Funktion f auf eine Zufallsvariable X liefert wieder eine Zufallsvariable $f(X)$. Falls X eine Zufallsvariable ist dann ist z.Bsp. $\sqrt{X^2 + X + 1}$ auch eine Zufallsvariable.

Funktionen und Unabhängigkeit (Satz 2.55.)

Seien $f_1, \ldots, f_n$ reellwertige Funktionen ($f_i \colon \mathbb{R} \to \mathbb{R}$ für $i = 1, \ldots, n$). Wenn die Zufallsvariablen $X_1, \ldots, X_n$ unabhängig sind, dann gilt dies auch für $f_1(X_1), \ldots, f_n(X_n)$.

Zusammengesetzte Zufallsvariablen

1. Aus Zufallsvariablen $X_1, \ldots, X_n \hookrightarrow$ neue Zufallsvariable $Y := g(X_1, \ldots, X_n)$
2. Die W'keit $''Y = y''$ berechnet man durch:

$$Pr[Y = y] = Pr[\omega \in \Omega \mid Y(\omega) = y] = Pr[\omega \mid g(X_1(\omega), \ldots, X_n(\omega)) = y]$$

Beispiel zsm.gesetzer Zufallsvariablen

X und Y bezeichnen die Augenzahl im ersten und zweiten Wurf eines Würfels. Lass uns die Zufallsvariable $Z := X + Y$ betrachten, die selber die Summe der Augenzahlen bezeichnet. Es gilt:

$$Pr[Z = 1] = Pr[\emptyset] = 0, Pr[Z = 3] = Pr[(1,2),(2,1)] = \frac{2}{36} \text{ usw.}$$

Summe von Zufallsvariablen

Summe von Zufallsvariablen (Satz 2.58.)

Für zwei unabhängige Zufallsvariablen X und Y sei $Z := X + Y$ Es gilt

$$f_Z(z) = \sum_{x \in \mathcal{W}_x} f_X(x) \cdot f_Y(z - x)$$

Summe von Zufallsvariablen

Summe von Zufallsvariablen (Satz 2.58.)

Für zwei unabhängige Zufallsvariablen X und Y sei $Z := X + Y$ Es gilt

$$f_Z(z) = \sum_{x \in \mathcal{W}_x} f_X(x) \cdot f_Y(z - x)$$

Beweis:

$$f_Z(z) = \Pr[Z = z] = \sum_{x \in \mathcal{W}_X} \Pr[X + Y = z \mid X = x] \cdot \Pr[X = x]$$

$$= \sum_{x \in \mathcal{W}_X} \Pr[Y = z - x] \cdot \Pr[X = x] = \sum_{x \in \mathcal{W}_X} f_X(x) \cdot f_Y(z - x).$$

$\rightarrow$ $Bin(n, p) + Bin(m, p) = Bin(n + m, p)$
$\rightarrow$ $Poisson(\lambda_1) + Poisson(\lambda_2) = Poisson(\lambda_1 + \lambda_2)$

Waldshce Identität

Waldsche Identität

N und X seien zwei unabhängige Zufallsvariablen, wobei für den Wertebereich von N gilt: $W_N \subseteq \mathbb{N}$. Weiter sei

$$Z := \sum_{i=1}^{N} X_i,$$

wobei $X_1, X_2, \ldots$ unabhängige Kopien von X seien. Dann gilt:

$$\mathbb{E}[Z] = \mathbb{E}[N] \cdot \mathbb{E}[X].$$

$\rightarrow$ Bei $Z := X + Y$ ist die Anzahl summanden konstant. Nun ist sie auch eine Zufallsvariable.

Waldshce Identität

Waldsche Identität

N und X seien zwei unabhängige Zufallsvariablen, wobei für den Wertebereich von N gilt: $W_N \subseteq \mathbb{N}$. Weiter sei

$$Z := \sum_{i=1}^{N} X_i,$$

wobei $X_1, X_2, \ldots$ unabhängige Kopien von X seien. Dann gilt:

$$\mathbb{E}[Z] = \mathbb{E}[N] \cdot \mathbb{E}[X].$$

→ Bei $Z := X + Y$ ist die Anzahl summanden konstant. Nun ist sie auch eine Zufallsvariable.

→ *"Wirf eine Münze bis zum ersten Kopf N-mal. Dann wirf die Münze noch N-mal und notiere mit Z die Anzahl Kopf bei diesen zweiten N-Versuchen.* $\mathbb{E}[Z] =$*?"*

Rechenregeln für Momente

Multiplikativität des Erwartungswerts

Für unabhängige Zufallsvariablen $X_1, \dots X_n$ gilt

$$\mathbb{E}[X_1 \cdot \dots \cdot X_n] = \mathbb{E}[X_1] \cdot \dots \cdot \mathbb{E}[X_n]$$

Rechenregeln für Momente

Multiplikativität des Erwartungswerts

Für unabhängige Zufallsvariablen $X_1, \dots X_n$ gilt

$$\mathbb{E}[X_1 \cdot \dots \cdot X_n] = \mathbb{E}[X_1] \cdot \dots \cdot \mathbb{E}[X_n]$$

Beweis:

$$\begin{aligned}
\mathbb{E}[X \cdot Y] &= \sum_{x \in \mathcal{W}_X} \sum_{y \in \mathcal{W}_Y} xy \cdot \Pr[X = x, Y = y] \\
&\overset{\text{Unabh.}}{=} \sum_{x \in \mathcal{W}_X} \sum_{y \in \mathcal{W}_Y} xy \cdot \Pr[X = x] \cdot \Pr[Y = y] \\
&= \sum_{x \in \mathcal{W}_X} x \cdot \Pr[X = x] \sum_{y \in \mathcal{W}_Y} y \cdot \Pr[Y = y] = \mathbb{E}[X] \cdot \mathbb{E}[Y]. \square
\end{aligned}$$

Rechenregeln für Momente

Additivität von der Varianz

Für unabhängige Zufallsvariablen $X_1, \ldots, X_n$ und $X := X_1 + \cdots + X_n$ gilt

$$\mathrm{Var}[X] = \mathrm{Var}[X_1] + \cdots + \mathrm{Var}[X_n].$$

Rechenregeln für Momente

Additivität von der Varianz

Für unabhängige Zufallsvariablen $X_1, \ldots, X_n$ und $X := X_1 + \cdots + X_n$ gilt

$$\mathrm{Var}[X] = \mathrm{Var}[X_1] + \cdots + \mathrm{Var}[X_n].$$

Beweis:

$$\mathbb{E}[(X+Y)^2] = \mathbb{E}[X^2 + 2XY + Y^2] = \mathbb{E}[X^2] + 2\mathbb{E}[X]\mathbb{E}[Y] + \mathbb{E}[Y^2]$$
$$\mathbb{E}[X+Y]^2 = (\mathbb{E}[X] + \mathbb{E}[Y])^2 = \mathbb{E}[X]^2 + 2\mathbb{E}[X]\mathbb{E}[Y] + \mathbb{E}[Y]^2,$$

Ziehe die zweite Gleichung von der Ersten ab:

$$\mathbb{E}[(X+Y)^2] - \mathbb{E}[X+Y]^2 = \mathbb{E}[X^2] - \mathbb{E}[X]^2 + \mathbb{E}[Y^2] - \mathbb{E}[Y]^2. \square$$

Rechenregeln für Momente

$$
\begin{aligned}
\mathbb{E}[X+Y] &= \mathbb{E}[X]+\mathbb{E}[Y] && \forall X, Y \\
\mathbb{E}[X\cdot Y] &= \mathbb{E}[X]\cdot\mathbb{E}[Y] && \forall X, Y \text{ unabhängig} \\
\mathrm{Var}[X+Y] &= \mathrm{Var}[X]+\mathrm{Var}[Y] && \forall X, Y \text{ unabhängig} \\
\mathrm{Var}[X\cdot Y] &\neq \mathrm{Var}[X]\cdot\mathrm{Var}[Y] && \text{i.A. (auch für unabhängige ZV)}
\end{aligned}
$$

Die Ungleichung von Markov

Markov's Ungleichung

Sei X eine Zufallsvariable, die nur nicht-negative Werte annimmt. Dann gilt für alle $t \in \mathbb{R}$ mit $t > 0$, dass

$$\Pr[X \geq t] \leq \frac{\mathbb{E}[X]}{t}.$$

Oder äquivalent dazu

$$\Pr[X \geq t \cdot \mathbb{E}[X]] \leq \frac{1}{t}.$$

Die Ungleichung von Markov

Markov's Ungleichung

Sei X eine Zufallsvariable, die nur nicht-negative Werte annimmt. Dann gilt für alle $t \in \mathbb{R}$ mit $t > 0$, dass

$$\Pr[X \geq t] \leq \frac{\mathbb{E}[X]}{t}.$$

Oder äquivalent dazu

$$\Pr[X \geq t \cdot \mathbb{E}[X]] \leq \frac{1}{t}.$$

Beweis: *Weglassen von Summanden in der Definition von* $\mathbb{E}[X]$

$$\mathbb{E}[X] = \sum_{x \in \mathcal{W}_X} x \cdot Pr[X = x] \geq \sum_{x \in \mathcal{W}_X, x \geq t} x \cdot Pr[X = x]$$

$$\geq t \cdot \sum_{x \in \mathcal{W}_X, x \geq t} Pr[X = x] = t \cdot Pr[x \geq t]$$

Die Ungleichung von Chebyshev

unleichung von Chebyshev

Sei X eine Zufallsvariable und $t \in R$ mit $t > 0$. Dann gilt

$$Pr[|X - \mathbb{E}[X]| \geq t] \leq \frac{Var[X]}{t^2}$$

oder äquivalent dazu $Pr[|X - \mathbb{E}[X]| \geq t\sqrt{Var[X]}] \leq \frac{1}{t^2}$

Die Ungleichung von Chebyshev

unleichung von Chebyshev

Sei X eine Zufallsvariable und $t \in R$ mit $t > 0$. Dann gilt

$$Pr[|X - \mathbb{E}[X]| \geq t] \leq \frac{Var[X]}{t^2}$$

oder äquivalent dazu $Pr[|X - \mathbb{E}[X]| \geq t\sqrt{Var[X]}] \leq \frac{1}{t^2}$

Beweis. Es gilt $\Pr[|X - \mathbb{E}[X]| \geq t] = \Pr[(X - \mathbb{E}[X])^2 \geq t^2]$

Die Zufallsvariable $Y := (X - \mathbb{E}[X])^2$ ist nicht-negativ und hat nach Definition der Varianz den Erwartungswert $\mathbb{E}[Y] = \mathrm{Var}[X]$. Damit folgt die Behauptung durch Anwendung der Markov-Ungleichung:

$$\Pr[|X - \mathbb{E}[X]| \geq t] = \Pr[Y \geq t^2] \leq \frac{\mathbb{E}[Y]}{t^2} = \frac{\mathrm{Var}[X]}{t^2}.$$

Die Ungleichung von Chernoff

Chernoff-Schranken

Seien $X_1, \ldots, X_n$ unabhängige Bernoulli-verteilte Zufallsvariablen mit $\Pr[X_i = 1] = p_i$ und $\Pr[X_i = 0] = 1 - p_i$. Dann gilt für $X := \sum_{i=1}^{n} X_i$:

(i) $\Pr[X \geq (1+\delta)\mathbb{E}[X]] \leq e^{-\frac{1}{3}\delta^2 \mathbb{E}[X]}$ für alle $0 < \delta \leq 1$,

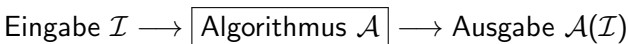
(ii) $\Pr[X \leq (1-\delta)\mathbb{E}[X]] \leq e^{-\frac{1}{2}\delta^2 \mathbb{E}[X]}$ für alle $0 < \delta \leq 1$,

(iii) $\Pr[X \geq t] \leq 2^{-t}$ für $t \geq 2e\mathbb{E}[X]$.

Präziser als Markov und Chebyshev. Beweis für (iii) auf der Tafel.

Randomisierte Algorithmen

→ **Deterministisch:**



Randomisierte Algorithmen

→ **Deterministisch:**

$$\text{Eingabe } \mathcal{I} \longrightarrow \boxed{\text{Algorithmus } \mathcal{A}} \longrightarrow \text{Ausgabe } \mathcal{A}(\mathcal{I})$$

→ **Randomisiert:**

$$\text{Eingabe } \mathcal{I} + \boxed{\text{Zufallsquelle: } \mathcal{R}} \longrightarrow \boxed{\text{Algorithmus } \mathcal{A}} \longrightarrow \text{Ausgabe } \mathcal{A}(\mathcal{I}, \mathcal{R})$$

Monte Carlo & Las Vegas

Monte Carlo Algorithmen

$\rightarrow$ immer schnell

$\rightarrow$ manchmal flasch
(Korrektheit ist Zufallsvariable)

Beispiele:

- ?

Las Vegas Algorithmen

$\rightarrow$ immer korrekt

$\rightarrow$ manchmal langsam
(Laufzeit ist Zufallsvariable)

Beispiele:

- ?

Monte Carlo & Las Vegas

Monte Carlo Algorithmen

→ immer schnell

→ manchmal flasch
(Korrektheit ist Zufallsvariable)

Beispiele:

Immer dieselbe Antwort geben

Las Vegas Algorithmen

→ immer korrekt

→ manchmal langsam
(Laufzeit ist Zufallsvariable)

Beispiele:

Randomisiertes QuickSort

Monte Carlo & Las Vegas

Monte Carlo Algorithmen

→ immer schnell

→ manchmal flasch
(Korrektheit ist Zufallsvariable)

Beispiele:

Immer dieselbe Antwort geben

Primzahltest durch 10 Teiler ausprobieren

Las Vegas Algorithmen

→ immer korrekt

→ manchmal langsam
(Laufzeit ist Zufallsvariable)

Beispiele:

Randomisiertes QuickSort

Primzahltest durch Teiler ausprobieren mit obere Schranke

Las Vegas mit oberer Schranke

Ein Las Vegas Algorithmus läuft bis er die richtige Antwort findet. Wir können die Laufzeit eines Las Vegas Algorithmus beschränken sodass der Algorithmus nur für eine beschränkte Zeit läuft, und falls am Ende die Antowrt noch nicht bekannt ist, ('???') zurückgibt.

```
1        Las_Vegas_Algo(params):
2
3            while(runtime < LIMIT):
4                do_work(params)
5
6            return ('???')
7
```

Beachte, dass diese Version **keine** Instanz von Monte-Carlo-Algorithmus ist. Bei einer Monte Carlo Algorithmus wird immer eine Antwort zurückgegeben, während hier der Algorithmus manchmal "Keine Ahnung!" sagt.

Beispiel: Primzahltest

```
Monte_Carlo_Prime_Test(n, k):

  repeat k times:

    pick random a in [2, n-2]

    if a^(n-1) mod n != 1:
      return 'Composite'

  return 'Probably Prime'

```

```
Las_Vegas_Prime_Test(n):

  factors = shuffle([sqrt(n)])

  for d in factors:

    if n mod d == 0:
      return 'Composite'

  return 'Prime'

```

Reduktion der Fehlerwahrscheinlichkeit

Reduktion der Fehlerw'keit - Las Vegas

Sei $\mathcal{A}$ ein randomisierter Algorithmus, der nie eine falsche Antwort gibt, aber zuweilen '???' ausgibt, wobei

$$\Pr[\mathcal{A}(\mathcal{I}) \text{ korrekt}] \geq \varepsilon.$$

Dann gilt für alle $\delta > 0$:
Sei $\mathcal{A}_\delta$ der Algorithmus, der $\mathcal{A}$ solange aufruft, bis entweder

- ein Wert verschieden von '???' ausgegeben wird (und diesen Wert dann ebenfalls ausgibt) oder bis
- $N = \lceil \varepsilon^{-1} \ln \delta^{-1} \rceil$ mal '???' ausgegeben wurde (und dann '???' ausgibt),

so gilt

$$\Pr[\mathcal{A}_\delta(\mathcal{I}) \text{ korrekt}] \geq 1 - \delta.$$

Beweis: Auf der Tafel (Skript s. 147)

Reduktion der Fehlerwahrscheinlichkeit

Reduktion der Fehlerw'keit - Monte Carlo - Einseitiger Feher

Sei $\mathcal{A}$ ein randomisierter Algorithmus, der immer entweder `JA` oder `NEIN` ausgibt, wobei

$$\Pr[\mathcal{A}(\mathcal{I}) = \texttt{JA}] = 1, \quad \text{falls } \mathcal{I} \text{ eine JA-Instanz ist, und}$$

$$\Pr[\mathcal{A}(\mathcal{I}) = \texttt{NEIN}] \geq \varepsilon, \quad \text{falls } \mathcal{I} \text{ eine NEIN-Instanz ist.}$$

Dann gilt für alle $\delta > 0$:
Sei $\mathcal{A}_\delta$ der Algorithmus, der $\mathcal{A}$ solange aufruft, bis entweder

- der Wert `NEIN` ausgegeben wird (und dann selbst `NEIN` ausgibt) oder bis
- $N = \lceil \varepsilon^{-1} \ln \delta^{-1} \rceil$ mal `JA` ausgegeben wurde (und dann selbst '`JA`' ausgibt),

so gilt

$$\Pr[\mathcal{A}_\delta(\mathcal{I}) \text{ korrekt}] \geq 1 - \delta.$$

Reduktion der Fehlerwahrscheinlichkeit

Reduktion der Fehlerw'keit - Monte Carlo - Zweiseitiger Feher

Sei $\varepsilon > 0$ und $\mathcal{A}$ ein randomisierter Algorithmus, der immer eine der beiden Antworten `JA` oder `NEIN` ausgibt, wobei

$$\Pr[\mathcal{A}(\mathcal{I}) \text{ korrekt}] \geq \tfrac{1}{2} + \varepsilon.$$

Dann gilt für alle $\delta > 0$: bezeichnet man mit $\mathcal{A}_\delta$ den Algorithmus, der $N = 4\varepsilon^{-2} \ln \delta^{-1}$ unabhängige Aufrufe von $\mathcal{A}$ macht und dann die Mehrheit der erhaltenen Antworten ausgibt, so gilt für den Algorithmus $\mathcal{A}_\delta$, dass

$$\Pr[\mathcal{A}_\delta(\mathcal{I}) \text{ korrekt}] \geq 1 - \delta.$$

Beweis: Auf der Tafel (Skript s. 149)

Exercises

Quiz auf der Webseite!

The End

♠