

Zusammenfassung

Mihir Mahajan, Alfred Nguyen, Noah Kiefer Diaz

June 3, 2022

Contents

1	Diskrete Wahrscheinlichkeitsräume	2
1.1	Grundlagen	2
1.1.1	Definition 1	2
1.1.2	Lemma 8	3
1.1.3	Satz 9 Siebformel	3
1.1.4	Wahl der Wahrscheinlichkeiten	4
1.2	Bedingte Wahrscheinlichkeiten	4
1.2.1	Definition 12	4
1.2.2	Baba Beispiele	4
1.2.3	Satz 18 (Satz von der totalen Wahrscheinlichkeit) . . .	4
1.2.4	Satz 19 (Satz von Bayes)	4
2	Unabhängigkeit	5
3	Zufallsvariablen	5
3.1	Grundlagen	5
3.2	Erwartungswert und Varianz	5
3.2.1	Definition 29	5
3.2.2	Satz 32 Monotonie des Erwartungswerts	5
3.2.3	Rechenregeln für Erwartungswert	5
3.2.4	Satz 33 (Linearität des Erwartungswerts, einfache Ver- sion)	6
3.2.5	Satz 34	6
3.2.6	Satz 35	6
3.2.7	Satz 36	6
3.2.8	Varianz	6
3.3	Mehrere Zufallsvariablen	6
3.3.1	hypergeometrische Verteilung	7

3.3.2	Gemeinsame Dichte	7
3.3.3	Unabhängigkeit von Zufallsvariablen	7
3.3.4	Zusammengesetzte Zufallsvariablen	7
3.3.5	Momente zusammengesetzter Zufallsvariablen	7
4	Wichtige diskrete Verteilungen	8
4.1	Bernoulli Verteilung	8
4.2	Binomialverteilung	8
4.2.1	Definition 55	8
4.2.2	Satz 56	8
4.3	5.3 Geometrische Verteilung	8
4.3.1	Definition 57	9
4.3.2	Warten auf n-ten Erfolg	9
4.4	5.4 Poisson-Verteilung	9
4.4.1	5.4.1 Poisson-Verteilung als Grenzwert der Binomialverteilung	9
4.4.2	Satz 59	9
5	Abschätzen von Wahrscheinlichkeiten	10
5.1	Markov & Chebyshev Ungleichungen	10
5.2	Satz 60: Markov-Ungleichung	10
5.3	Satz 61: Chebyshev-Ungleichung	10
5.4	Satz 62: Gesetz der großen Zahlen	10
5.4.1	Wahrscheinlichkeit und relative Häufigkeit	10
5.5	Chernoff-Schranken	10

1 Diskrete Wahrscheinlichkeitsräume

1.1 Grundlagen

1.1.1 Definition 1

- Ein diskreter Wahrscheinlichkeitsraum ist durch eine **Ergebnismenge** $\Omega = \{\omega_1, \dots, \omega_n\}$ von Elementarereignissen gegeben
- Jedem Ereignis ω_i ist eine Wahrscheinlichkeit $0 \leq Pr[\omega_i] \leq 1$ zugeordnet

$$\sum_{\omega \in \Omega} Pr[\omega] = 1$$
- Die Menge $E \subseteq \Omega$ heißt Ereignis. $Pr[E] = \sum_{\omega \in E} Pr[\omega]$
- \bar{E} ist komplement zu E

Man kann standard Mengenoperationen auf Ereignisse machen, also bei Ereignissen A, B dann auch $A \cup B, A \cap B$

1.1.2 Lemma 8

Für Ereignisse $A, B, A_1, A_2, \dots, A_n$ gilt

- $Pr[\emptyset] = 0, Pr[\Omega] = 1$
- $0 \leq Pr[A] \leq 1$
- $Pr[\bar{A}] = 1 - Pr[A]$
- Wenn $A \subseteq B$ so folgt $Pr[A] \leq Pr[B]$
- Additionssatz: Bei **paarweise disjunkten** Ereignissen gilt:
 $Pr[\bigcup_{i=1}^n A_i] = \sum_{i=1}^n Pr[A_i]$
 Insbesondere gilt also:
 $Pr[A \cup B] = Pr[A] + Pr[B]$
 Und für unendliche Menge von **disjunkten** Ereignissen:
 $Pr[\bigcup_{i=1}^{\infty} A_i] = \sum_{i=1}^{\infty} Pr[A_i]$

1.1.3 Satz 9 Siebformel

Lemma 8, gilt nur für **disjunkte** Mengen. Das geht auch für nicht disjunkte!

1. Zwei Mengen $Pr[A \cup B] = Pr[A] + Pr[B] - Pr[A \cap B]$
2. Drei Mengen $Pr[A_1 \cup A_2 \cup A_3] =$
 $Pr[A_1] + Pr[A_2] + Pr[A_3]$
 $- Pr[A_1 \cap A_2] - Pr[A_1 \cap A_3] - Pr[A_2 \cap A_3]$
 $+ Pr[A_1 \cap A_2 \cap A_3]$
3. n Mengen Veranschaulichen an Venn-Diagramm
 - (a) Alle aufaddieren
 - (b) Paarweise schnitte subtrahieren
 - (c) Dreifache schnitte dazuaddieren
 - (d) 4- fache schritte subtrahieren
 - (e) ...
4. für nerds: $Pr[\bigcup_{i=0}^n A_i] = \sum_{\emptyset \subset I \subseteq [n]} (-1)^{|I|+1} Pr[\bigcap_{i \in I} A_i]$

1.1.4 Wahl der Wahrscheinlichkeiten

Prinzip von Laplace (Pierre Simon Laplace (1749–1827)): Wenn nichts dagegen spricht, gehen wir davon aus, dass alle Elementarereignisse gleich wahrscheinlich sind. $Pr[E] = \frac{|E|}{|\Omega|}$

1.2 Bedingte Wahrscheinlichkeiten

1.2.1 Definition 12

A und B seien Ereignisse mit $Pr[B] > 0$. Die bedingte Wahrscheinlichkeit $Pr[A|B]$ von A gegeben B ist definiert als: $Pr[A|B] := \frac{Pr[A \cap B]}{Pr[B]}$

Umgangssprachlich: $Pr[A|B]$ beschreibt die Wahrscheinlichkeit, dass A eintritt wenn B eintritt.

Die bedingten Wahrscheinlichkeiten $Pr[\cdot|B]$ bilden für ein beliebiges Ereignis $B \subseteq \Omega$ mit $Pr[B] > 0$ einen neuen Wahrscheinlichkeitsraum über Ω .

1.2.2 Baba Beispiele

1. **TODO** Töchterproblem
2. **TODO** Ziegenproblem
3. **TODO** Geburtstagsproblem

1.2.3 Satz 18 (Satz von der totalen Wahrscheinlichkeit)

Die Ereignisse A_1, \dots, A_n seien paarweise disjunkt und es gelte $B \subseteq A_1 \cup \dots \cup A_n$.

$$Pr[B] = \sum_{i=1}^n Pr[B|A_i] * Pr[A_i]$$

analog für $n \rightarrow \infty$

1.2.4 Satz 19 (Satz von Bayes)

Es seien A_1, \dots, A_n paarweise disjunkt, mit $Pr[A_j] > 0$ für alle j . Außerdem sei $B \subseteq A_1 \cup \dots \cup A_n$ ein Ereignis mit $Pr[B] > 0$. Dann gilt für beliebiges $i \in [n]$

$$Pr[A_i|B] = \frac{Pr[A_i \cap B]}{Pr[B]} = \frac{Pr[B|A_i] * Pr[A_i]}{\sum_{j=1}^n Pr[B|A_j] * Pr[A_j]}$$

2 Unabhängigkeit

Wenn das auftreten von Ereignissen unabhängig ist.

- $Pr[A \cap B] = Pr[A] * Pr[B]$
- $Pr[A|B] = Pr[A]$

3 Zufallsvariablen

3.1 Grundlagen

Anstatt der Ereignisse selbst sind wir oft an "Auswirkungen" oder "Merkmalen" der (Elementarereignisse) interessiert

Sei ein Wahrscheinlichkeitsraum auf der Ergebnismenge Ω gegeben. Eine Abbildung $X : \Omega \rightarrow R$ heißt (numerische) Zufallsvariable. Eine Zufallsvariable X über einer endlichen oder abzählbar unendlichen Ergebnismenge heißt **diskret**

3.2 Erwartungswert und Varianz

3.2.1 Definition 29

Zu einer Zufallsvariablen X definieren wir den **Erwartungswert** $E[X]$ durch $E[X] := \sum_{x \in W_X} x * Pr[X = x] = \sum x * f_X(x)$ sofern $\sum_{x \in W_X} |x| * Pr[X = x]$ konvergiert

3.2.2 Satz 32 Monotonie des Erwartungswerts

Seien X und Y Zufallsvariablen über dem Wahrscheinlichkeitsraum ω mit $X(\omega) \leq Y(\omega)$ für alle $\omega \in \Omega$. Dann gilt $E[X] \leq E[Y]$ $E[X] = \sum_{\omega \in \Omega} X(\omega) * Pr[\omega] \leq \sum_{\omega \in \Omega} Y(\omega) * Pr[\omega] = E[Y]$

3.2.3 Rechenregeln für Erwartungswert

Oft betrachtet man eine Zufallsvariable X nicht direkt, sondern wendet noch eine Funktion darauf an: $Y := f(X) = f \circ X$, wobei $f : D \rightarrow R$ eine beliebige Funktion sei mit $W_X \subseteq D \subseteq R$. Beobachtung: $f(X)$ ist wieder eine Zufallsvariable.

3.2.4 Satz 33 (Linearität des Erwartungswerts, einfache Version)

Für eine beliebige Zufallsvariable X und $a, b \in \mathbb{R}$ gilt

$$\mathbb{E}[a * X + b] = a * \mathbb{E}[X] + b$$

3.2.5 Satz 34

Sei X eine Zufallsvariable mit $W_x \subseteq \mathbb{N}_0$. Dann gilt

$$\mathbb{E}[X] = \sum_{i=1}^{\infty} \Pr[X \geq i]$$

3.2.6 Satz 35

Sei X eine Zufallsvariable und A ein Ereignis mit $\Pr[A] > 0$. Die **bedingte Zufallsvariable** $X|A$ besitzt die Dichte:

$$f_{X|A}(x) := \Pr[X = x|A] = \frac{\Pr[\"X=x\" \cap A]}{\Pr[A]}$$

Die Definition ist zulässig, da

$$\sum_{x \in W_x} f_{X|A}(x) = \sum_{x \in W_x} \frac{\Pr[\"X=x\" \cap A]}{\Pr[A]} = 1$$

Somit ist $\mathbb{E}[X|A] = \sum_{x \in W_x} x * f_{X|A}(x)$

3.2.7 Satz 36

TODO

3.2.8 Varianz

1. Definition 38 Für eine Zufallsvariable X mit $\mu = E[X]$ definieren wir die Varianz $Var[X]$ durch
$$Var[X] := E[(X - \mu)^2] = \sum_{x \in W_X} (x - \mu)^2 * \Pr[X = x]$$
Die Größe $\sigma := \sqrt{Var[X]}$
 $Var[X]$ heißt Standardabweichung von X .
2. Satz 39 Für eine beliebige Zufallsvariable X gilt
$$Var[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$
3. Satz 41 Für eine beliebige Zufallsvariable X und $a, b \in \mathbb{R}$ gilt:
$$Var[a * X + b] = a^2 * Var[X]$$

3.3 Mehrere Zufallsvariablen

Wie kann man mit mehreren Zufallsvariablen über demselben Wahrscheinlichkeitsraum rechnen, auch wenn sie, wie im obigen Beispiel, sehr voneinander abhängig sind? Wir untersuchen Wahrscheinlichkeiten der Art
$$\Pr[X = x, Y = y] = \Pr[\omega; X(\omega) = x, Y(\omega) = y]$$

3.3.1 hypergeometrische Verteilung

Allgemein nennt man Zufallsvariablen mit der Dichte $Pr[X = x] = \frac{\binom{b}{x} * \binom{a}{r-x}}{\binom{a+b}{r}}$

hypergeometrisch verteilt. Durch diese Dichte wird ein Experiment modelliert, bei dem r Elemente ohne Zurücklegen aus einer Grundmenge der Mächtigkeit a + b mit b besonders ausgezeichneten Elementen gezogen werden.

3.3.2 Gemeinsame Dichte

Die Funktion

$f_{X,Y}(x, y) := Pr[X = x, Y = y]$ heißt gemeinsame Dichte der Zufallsvariablen X und Y.

Aus der gemeinsamen Dichte $f_{X,Y}$ kann man ableiten

$$f_X(x) = \sum_{y \in W_Y} f_{X,Y}(x, y)$$

$$f_Y(y) = \sum_{x \in W_X} f_{X,Y}(x, y)$$

Die Funktionen f_X und f_Y nennt man Randdichten.

3.3.3 Unabhängigkeit von Zufallsvariablen

1. Definition 45 Die Zufallsvariablen X_1, \dots, X_n heißen unabhängig, wenn für alle $(x_1, \dots, x_n) \in W_{X_1} \times \dots \times W_{X_n}$ gilt:

$$Pr[X_1 = x_1, \dots, X_n = x_n] = Pr[X_1 = x_1] * \dots * Pr[X_n = x_n]$$

Analog: Gesamte Dichte ist Produkt aus einzelnen Dichten. Analog: Gesamte Verteilung ist Produkt aus einzelnen Verteilungen.

3.3.4 Zusammengesetzte Zufallsvariablen

1. Satz 49 Für zwei unabhängige Zufallsvariablen X und Y sei $Z := X + Y$. Es gilt $f_Z(z) = \sum_{x \in W_X} f_X(x) * f_Y(z - x)$

3.3.5 Momente zusammengesetzter Zufallsvariablen

1. Satz 50 (Linearität des Erwartungswerts) Für Zufallsvariablen X_1, \dots, X_n und $X := a_1 X_1 + \dots + a_n X_n$ mit $a_1, \dots, a_n \in R$ gilt $\mathbb{E}[X] = a_1 \mathbb{E}[X_1] + \dots + a_n \mathbb{E}[X_n]$

2. Satz 52 (Multiplikativität des Erwartungswerts) Für unabhängige Zufallsvariablen X_1, \dots, X_n gilt $E[X_1 * \dots * X_n] = E[X_1] * \dots * E[X_n]$
3. Definition 53 Zu einem Ereignis A heißt die Zufallsvariable

$$I_A = \begin{cases} 1 & \text{falls A eintritt} \\ 0 & \text{sonst} \end{cases}$$
Indikatorvariable des Ereigniss A

4 Wichtige diskrete Verteilungen

4.1 Bernoulli Verteilung

Eine Zufallsvariable X mit $W_X = 0, 1$ und der Dichte $f_X(x) = \begin{cases} p & x = 1 \\ 1 - p & x = 0 \end{cases}$ heißt Bernoulli-verteilt. Den Parameter p nennen wir Erfolgswahrscheinlichkeit. Eine solche Verteilung erhält man z.B. bei einer einzelnen Indikatorvariablen. Es gilt mit $q := 1 - p$ $E[X] = p$ und $Var[X] = pq$, wegen $E[X^2] = p$ und $Var[X] = E[X^2]E[X]^2 = pp^2$.

4.2 Binomialverteilung

Eine Bernoulli-verteilte Zufallsvariable entspricht der Verteilung einer Indikatorvariablen. Häufig betrachtet man jedoch Summen von Indikatorvariablen.

4.2.1 Definition 55

Sei $X := X_1 + \dots + X_n$ als Summe von n unabhängigen, Bernoulli-verteilten Zufallsvariablen mit gleicher Erfolgswahrscheinlichkeit p definiert. Dann heißt X binomialverteilt mit den Parametern n und p . In Zeichen schreiben wir $X \sim Bin(n, p)$

4.2.2 Satz 56

Wenn $X \sim Bin(n_x, p)$ und $Y \sim Bin(n_y, p)$ unabhängig sind, dann gilt für $Z := X + Y$, dass $Z \sim Bin(n_x + n_y, p)$.

4.3 5.3 Geometrische Verteilung

Man betrachte ein Experiment, das so lange wiederholt wird, bis Erfolg eintritt. Gelingt ein einzelner Versuch mit Wahrscheinlichkeit p , so ist die Anzahl der Versuche bis zum Erfolg geometrisch verteilt.

4.3.1 Definition 57

Eine geometrisch verteilte Zufallsvariable X mit Parameter (Erfolgswahrscheinlichkeit) $p \in (0, 1]$ und $q := 1 - p$ hat die Dichte $f_X(i) = pq^{i-1}$ für $i \in \mathbb{N}$. Für Erwartungswert und Varianz geometrisch verteilter Zufallsvariablen gilt:

$$E[X] = \frac{1}{p} \text{ und } Var[X] = \frac{q}{p^2}$$

4.3.2 Warten auf n-ten Erfolg

$$f_Z(z) = \binom{z-1}{n-1} * p^n (1-p)^{z-n}$$

4.4 5.4 Poisson-Verteilung

Die Poisson-Verteilung kann verwendet werden, um die Anzahl von Ereignissen zu modellieren, welche mit konstanter Rate und unabhängig voneinander in einem Zeitintervall auftreten. Eine Poisson-verteilte Zufallsvariable X mit dem Parameter $\lambda \geq 0$ hat den Wertebereich $W_X = \mathbb{N}_0$ und besitzt die Dichte $f_X(i) = \frac{e^{-\lambda} \lambda^i}{i!}$ für $i \in \mathbb{N}_0$.

Als Erwartungswert ergibt sich:

$$E[X] = \lambda$$

Und für die Varianz:

$$Var[X] = \lambda$$

4.4.1 5.4.1 Poisson-Verteilung als Grenzwert der Binomialverteilung

Wir betrachten eine Folge von binomialverteilten Zufallsvariablen X_n mit $X_n \sim Bin(n, p_n)$, wobei $p_n = \lambda/n$. Für ein beliebiges k mit $0 \leq k \leq n$ ist die Wahrscheinlichkeit, dass X_n den Wert k annimmt gleich:

$$b(k; n, p_n) = \frac{\lambda^k}{k!} * \frac{n^k}{n^k} (1 - \frac{\lambda}{n})^{n-k}$$

Damit folgt:

$$\lim_{n \rightarrow \infty} b(k; n, p_n) = e^{-\lambda} * \frac{\lambda^k}{k!}$$

4.4.2 Satz 59

Sind X und Y unabhängige Zufallsvariablen mit $X \sim Po(\lambda)$ und $Y \sim Po(\mu)$, dann gilt

$$Z := X + Y \sim Po(\lambda + \mu)$$

5 Abschätzen von Wahrscheinlichkeiten

5.1 Markov & Chebyshev Ungleichungen

5.2 Satz 60: Markov-Ungleichung

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

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

Äquivalent dazu: $\Pr[X > t \bullet \mathbb{E}[X]] \leq 1/t$

5.3 Satz 61: Chebyshev-Ungleichung

Sei X eine Zufallsvariable, und sei $t \in \mathbb{R}$ mit $t > 0$. Dann gilt

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

Äquivalent dazu: $\Pr[|X - \mathbb{E}[X]| \geq t\sqrt{\text{Var}[X]}] \leq 1/t^2$

5.4 Satz 62: Gesetz der großen Zahlen

Gegeben sei eine Zufallsvariable X & seien $\epsilon, \delta > 0$ beliebig aber fest. Dann gilt für alle $n \geq \frac{\text{Var}[X]}{\epsilon\delta^2}$

Sind X_1, \dots, X_n unabhängige Zufallsvariablen mit derselben Verteilung wie X und setzt man $Z := \frac{X_1 + \dots + X_n}{n}$ so gilt $\Pr[|Z - \mathbb{E}[X]| \geq \delta] \leq \epsilon$

5.4.1 Wahrscheinlichkeit und relative Häufigkeit

TODO: slide 161

5.5 Chernoff-Schranken