

2 Wahrscheinlichkeitstheorie

In diesem Kapitel beschäftigen wir uns mit der Wahrscheinlichkeitsrechnung. Hierbei versuchen wir die wichtigsten grundlegenden Definitionen und Notationen möglichst elementar an einfach verständlichen Beispielen klar zu machen. Dies ist nicht so selbstverständlich, da die Wahrscheinlichkeitstheorie auf der Maßtheorie und auf der Funktionentheorie basiert und somit die meisten Bücher über Wahrscheinlichkeitstheorie mit einem stark theoretischen Ansatz beginnen. Für einen umfassenden Überblick sind an entsprechenden Stellen eine Literaturlauswahl zu diesem Thema aufgeführt, die jedoch einen grösseren mathematischen Background voraussetzen. Im Weiteren reichen jedoch die Vorkenntnisse der Mathematik für Informatiker Vorlesung (MfI) vollkommen aus.

2.1 Grundlagen und Definitionen

Es hat sich im Laufe der Jahre gezeigt, dass die Wahrscheinlichkeitstheorie in vielen Bereichen auf einer sehr allgemeinen theoretischen Grundlage entwickelt werden kann. Hieraus haben sich verschiedenen Ansätze und Betrachtungsweisen entwickelt, auf die wir im Folgenden nicht näher eingehen wollen. Erst im Jahre 1933 hat Kolmogorov die Wahrscheinlichkeitstheorie als ein Teilgebiet der Maßtheorie etabliert, in der Form in der wir sie heute am häufigsten zu Gesicht bekommen. Auf vereinfachte Art und Weise können wir wie folgt beginnen.

Definition 2.1. Sei S eine endliche Menge und sei p eine Abbildung von S in die positiven reellen Zahlen

$$p : S \rightarrow \mathbb{R}_{\geq 0}.$$

Für eine Teilmenge $A \subseteq S$ von S sei p definiert als

$$p(A) = \sum_{s \in A} p(s).$$

Dies entspricht nicht dem üblichen mathematischen Ansatz, ist jedoch für unsere Ziele die zweckmäßigste Vorgehensweise. Für einen Zugang über die Axiome von Kolmogorov sei der Leser an [?] verwiesen. Auf dieser Definition können wir einige neue Begriffe und Notationen aufbauen.

Definition 2.2.

1. (S, p) heißt **Wahrscheinlichkeitsraum**, wenn $p(S) = 1$.

2. Die endliche Menge S heißt der **Ereignisraum** (sample space).
3. Die Funktion p heißt die **Wahrscheinlichkeitsverteilung**.
4. Eine Teilmenge $A \subset S$ heißt eine **Ereignis** (event).
5. Enthält A genau ein Element der Menge S , so sprechen wir von einem **Elementarereignis**.
6. Hat jedes Elementarereignis die gleiche Wahrscheinlichkeit, so nennen wir das Zufallsexperiment **fair**.

Zur Verdeutlichung der doch etwas theoretisch anmutenden Konstrukte machen wir zuerst einmal einige kleine Beispiele, auf die wir uns auch im Weiteren immer wieder zurückziehen werden.

Beispiel 2.1. 1. Eines der einfachsten Zufallsexperimente, das schon jeder von uns einmal gemacht hat, ist das Werfen einer Münze und feststellen, ob sie nach der Landung Kopf (K) oder Zahl (Z) zeigt. Wir sprechen dann von einer Durchführung des Zufallsexperiments:

“Werfen einer Münze und Feststellen, welche Seite oben liegt.“

Die möglichen Ausgänge des Experiments lassen sich also in der Menge

$$S = \{K, Z\}$$

zusammenfassen. Falls es sich bei der Münze um eine faire Münze handelt, erscheinen Kopf und Zahl mit der gleichen Wahrscheinlichkeit. Definieren wir p als

$$p(K) = p(Z) = \frac{1}{2}$$

dann gilt

$$p(S) = p(K) + p(Z) = 1.$$

Somit ist (S, p) ein Wahrscheinlichkeitsraum mit Ereignisraum $S = \{K, Z\}$ und Wahrscheinlichkeitsverteilung p .

2. Ebenso einfach in der Durchführung jedoch mit mehreren möglichen Ergebnissen ist das Werfen eines Würfels und Feststellen der Augenzahl. Nach der Landung des Würfels sehen wir auf der Oberseite des Würfels 1, 2, 3, 4, 5 oder 6 Punkte, d.h. wir können die Ergebnisse in der Menge

$$S = \{1, 2, 3, 4, 5, 6\}$$

zusammenfassen. Falls es sich um einen fairen Würfel handelt, erscheint jede der sechs Zahlen mit der gleichen Wahrscheinlichkeit $\frac{1}{6}$. Definieren wir also p als

$$p(1) = p(2) = p(3) = p(4) = p(5) = p(6) = \frac{1}{6},$$

dann gilt

$$p(S) = \sum_{i=1}^6 p(i) = 6 \cdot \frac{1}{6} = 1.$$

Somit ist (S, p) ein Wahrscheinlichkeitsraum mit Ereignisraum $S = \{1, 2, 3, 4, 5, 6\}$ und Wahrscheinlichkeitsverteilung p .

2.1.1 Wahrscheinlichkeit von Ereignissen $A \subseteq S$

Betrachten wir eine Teilmenge $A \subseteq S$ von S , d.h. ein Ereignis, so sagt uns Definition 2.1 des Ereignisraums bzw. der Wahrscheinlichkeitsverteilung p , dass sich die Wahrscheinlichkeit von A wie folgt

$$p(A) = \sum_{s \in A} p(s)$$

berechnet. Die Wahrscheinlichkeit eines Ereignisses A ist also die Summe der Wahrscheinlichkeiten der in A enthaltenen Elementarereignisse. Dies kann wie folgt interpretiert werden.

Korollar 2.1. *Das Ereignis A tritt dann auf, wenn das Resultat des zufälligen Prozesses ein in A enthaltenes Elementarereignis ist.*

Beispiel 2.2. Betrachten wir unser Würfelbeispiel mit dem Ereignis $A = \{1, 6\}$, dann erhalten wir

$$p(A) = p(1) + p(6) = \frac{2}{6}.$$

Das Ereignis A tritt dann auf, wenn der Wurf eine 1 oder eine 6 liefert. Dies geschieht mit Wahrscheinlichkeit $\frac{2}{6}$.

Wir können unseren Versuchsaufbau verallgemeinern in dem wir diese Zufallsexperimente wiederholt durchführen.

Bemerkung 2.1. Werden n Zufallsexperimente nacheinander ausgeführt, so können wir dies als einmalige Durchführung eines n -stufigen Zufallsexperiments werten und somit als n -Tuple in unsere Ergebnismenge aufnehmen.

Dieses Vorgehen wird am folgenden Beispiel noch einmal verdeutlicht.

Beispiel 2.3. Wir erweitern unser Münzbeispiel und werfen eine faire Münze n mal. Als Resultat von diesen n Würfeln erhalten wir eine Folge der Länge n der Form

$$KZZKKKZKZK...ZZKZKK,$$

d.h wir haben nach jedem Wurf festgehalten, ob die Münze Kopf (K) oder Zahl (Z) gezeigt hat.

$$S = \{K, Z\}^n$$

Wir müssen uns jetzt die Frage stellen, wieviele verschiedene Elementarereignisse (Folgen von K 's und Z 's der Länge n) gibt es für unserer Zufallsexperiment. Die Antwort

auf diese Frage liefert die Kombinatorik oder ein einfaches Abzählprinzip. Auf den Bereich der Kombinatorik werden wir im Rahmen dieses Skripts nicht näher eingehen, wollen dem Leser am Ende des Kapitels jedoch eine kleine tabellarische Zusammenfassung einiger wichtiger kombinatorischer Ergebnisse liefern. Hieraus wird ersichtlich, dass wir

$$\#S = 2^n$$

als Mächtigkeit unseres Ereignisraums erhalten (Wir betrachten n Würfe, wobei jeder Wurf 2 Ergebnisse haben kann). Jedes Elementarereignis $e \in S$ tritt mit der gleichen Wahrscheinlichkeit

$$p(e) = \frac{1}{2^n}$$

auf. Es gilt dann

$$p(S) = \sum_{e \in S} p(e) = 2^n \cdot \frac{1}{2^n} = 1.$$

2.1.2 Bedingte Wahrscheinlichkeiten

Für die Wahrscheinlichkeit, dass ein Elementarereignis e_i eintritt unter der Bedingung, dass das Ereignis B bereits eingetreten ist, gilt offensichtlich 0, falls $e_i \notin B$ und $\frac{P(e_i)}{P(B)}$, falls $e_i \in A$. Verallgemeinern wir die Situation und betrachten anstelle des Elementarereignisses e_i ein beliebiges Ereignis A , dann ist die Wahrscheinlichkeit, mit der A eintritt, falls B bereits eingetreten ist

$$\frac{p(A \cap B)}{p(B)},$$

da nur die Ergebnisse zählen, welche in $A \cap B$ liegen. Dies führt uns zu folgender Definition.

Definition 2.3. Seien A und B Ereignisse mit $p(B) \neq 0$. Wir definieren die **bedingte Wahrscheinlichkeit** $p(A|B)$ von A unter der **Bedingung (Hypothese)** B als

$$p(A|B) = \frac{p(A \cap B)}{p(B)}.$$

Die bedingte Wahrscheinlichkeit $p(A|B)$ ist die Wahrscheinlichkeit, dass ein Elementarereignis $e \in A$ eintritt, unter der Annahme, dass die Hypothese B wahr ist, d.h., dass das eintretende Elementarereignis e mit Sicherheit ein Elementarereignis von B ist.

Wir erhalten aus dieser Definition direkt die Aussagen.

Folgerung 2.1. 1. Ist der Durchschnitt $A \cap B$ der Ereignisse A und B leer, so ist $p(A|B) = 0$.

2. Ist B in A enthalten, so ist $p(A|B) = 1$.

Beispiel 2.4. Greifen wir unser Zufallsexperiment „Werfen eines Würfels“ auf und betrachten die beiden Ereignisse $A = \{1, 2, 3\}$ und $B = \{3, 4, 5\}$ so erhalten wir für die bedingte Wahrscheinlichkeit

$$p(A|B) = \frac{p(3)}{p(\{3, 4, 5\})} = \frac{1}{3},$$

d.h. die Wahrscheinlichkeit das Ereignis A eintritt, wenn Hypothese B erfüllt ist, ist $\frac{1}{3}$. Dieses Ergebnis ist auch noch augenscheinlich klar, da die Schnittmenge von A und B nur aus einem einzigen Element besteht und die Wahrscheinlichkeit somit noch offensichtlich ist.

Um unsere Betrachtungen weiterführen zu können, benötigen wir noch eine einfache Notationsvereinbarung. Die jedoch teilweise beim Studium der diversen Fachliteratur für Verwirrung sorgt, da viele Autoren eine eigene Notation verwenden. Wir wollen uns im Folgenden an Definition 2.4 halten.

Definition 2.4. Sei $A \subseteq S$ ein Ereignis. Mit A^c bezeichnen wir das **komplementäre Ereignis** mit

$$A^c = S \setminus A = \{e \in S | e \notin A\}.$$

Leider fehlen in vielen Fällen die entsprechenden Informationen über die Wahrscheinlichkeit des Durchschnitts von A und B . In diesen Fällen kann die folgende Regel von Bayes weiterhelfen.

Satz 2.1. (Bayes'sche Regel)

Seien A und B Ereignisse. Es gelte $p(b) \neq 0$, $p(A^c) \neq 0$ und $p(A) \neq 0$. Dann gilt

$$p(A|B) = \frac{p(A)p(B|A)}{p(A)p(B|A) + p(A^c)p(B|A^c)}.$$

Beweis.

$$\begin{aligned} p(A|B) &= \frac{p(B \cap A)}{p(B)} \\ &= \frac{p(A) \frac{p(B \cap A)}{p(A)}}{p(B)} \\ &= \frac{p(A)p(B|A)}{p(B)} \\ &= \frac{p(A)p(B|A)}{p(B \cap A) + p(B \cap A^c)} \\ &= \frac{p(A)p(B|A)}{p(A)p(B|A) + p(A^c)p(B|A^c)}. \end{aligned}$$

■

Die Formel aus 2.1 scheint den Sachverhalt nicht einvereinfacher zu machen, spiegelt jedoch in der Anwendung eine häufig auftretende Konstellation wieder, wie das folgende Beispiel veranschaulicht.

Beispiel 2.5. Gegeben eine faire Münze und eine gefälschte Münze, die immer Kopf zeigt. Wir wählen eine der beiden Münzen mit Wahrscheinlichkeit $\frac{1}{2}$ und werfen diese Münze zweimal. Betrachten wir nun auf der einen Seite das Ereignis

$$B = \{(KK)\}$$

und auf der anderen das Ereignis

$$A = \{\text{gefälschte Münze gewählt}\},$$

dann erhalten wir

$$\begin{aligned} p(B) &= p(A)p(B|A) + p(A^c)p(B|A^c) \\ &= \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{4} = \frac{5}{8}, \end{aligned}$$

da $p(A) = \frac{1}{2}$ nach Voraussetzung und somit $p(B|A) = 1$ und $p(B|A^c) = \frac{1}{4}$ gilt. Insgesamt erhalten wir für die bedingte Wahrscheinlichkeit

$$\begin{aligned} p(A|B) &= \frac{p(A)p(B|A)}{p(A)p(B|A) + p(A^c)p(B|A^c)} \\ &= \frac{\frac{1}{2} \cdot 1}{1 + \frac{1}{2} \cdot \frac{1}{4}} \\ &= \frac{4}{5}. \end{aligned}$$

2.2 Zufallsvariablen und ihre Wahrscheinlichkeitsverteilung

Definition 2.5. Gegeben ein Wahrscheinlichkeitsraum (S, p) . Eine Funktion

$$X : S \rightarrow \mathbb{R},$$

die jedem Elementarereignisse e eine reelle Zahl zuordnet, heißt **Zufallsvariable**.

Korollar 2.2. *Mittels Zufallsvariablen kann man zum Beispiel jedem Elementarereignis e Kosten zuordnen, die durch das zufällige “Auftreten” von e verursacht werden. Im Folgenden werden wir oft Zufallsvariablen betrachten, deren Kosten den “Laufzeiten” von Algorithmen (mit einer Zufallskomponente) entsprechen.*

Für ein vorgegebenes $x \in \mathbb{R}$ betrachten wir häufig die Menge aller Elementarereignisse, die durch die Zufallsvariable X auf x abgebildet werden

$$\{e \in S | X(e) = x\}.$$

Als Kurzschreibweise für diese Menge verwenden wir

$$\{X = x\}.$$

2.2.1 Erwartungswert

Definition 2.6. Gegeben ein Wahrscheinlichkeitsraum (S, p) und eine Zufallsvariable $X : S \rightarrow \mathbb{R}$. Der **Erwartungswert** der Zufallsvariablen X ist definiert als

$$E[X] = \sum_{e \in S} p(e) \cdot X(e).$$

Korollar 2.3. Repräsentiert die Zufallsvariable X eine Art **Kostenfunktion**, so kann der Erwartungswert als die (mittleren) zu erwartenden Kosten interpretiert werden.

Siehe folgendes Beispiel.

Beispiel 2.6. Wir werfen zweimal hintereinander mit einem Würfel. Die Elementarereignisse sind Paare (a, b) mit $a, b \in \{1, 2, 3, 4, 5, 6\}$. Die “Kosten“ $X(e)$ eines Elementarereignisses $e = (a, b)$ sind gleich der Summe $a + b$ der Augenzahlen, d.h.

$$X(e = (a, b)) = a + b.$$

Für die Wahrscheinlichkeiten der Elementarereignisse erhalten wir also

$$p(e = (a, b)) = \frac{1}{36}$$

und der Erwartungswert ergibt sich zu

$$\begin{aligned} E[X] &= \sum_{e \in S} p(e) X(e) = \sum_{e \in S} \frac{1}{36} X(e) = \frac{1}{36} \sum_{e \in S} X(e) \\ &= \frac{1}{36} \sum_{a=1}^6 \sum_{b=1}^6 (a + b) = \frac{1}{36} \sum_{a=1}^6 (6a + \sum_{b=1}^6 b) = \frac{1}{36} \sum_{a=1}^6 (6a + 21) \\ &= \frac{12 \cdot 21}{36} = \frac{12 \cdot 7}{12} = 7. \end{aligned}$$

Man beachte:

$$E[X] = \sum_{e \in S} p(e) X(e) = \sum_{k \in X(S)} p(\{X = k\}) \cdot k$$

Beispiel 2.7. Wir werfen eine faire Münze n mal hintereinander. Die Zufallsvariable X ordnet einem Elementarereignis $e \in S = \{K, Z\}^n$ die Anzahl der Würfe mit Resultat Kopf in e zu. Wie groß ist die Wahrscheinlichkeit, dass man k Köpfe in einer Folge von n Würfeln erhält, und welchen Wert hat der Erwartungswert $E[X]$ von X ?

Für die Wahrscheinlichkeit erhalten wir mit Hilfe der Kombinatorik

$$p(\{X = k\}) = \binom{n}{k} \cdot 2^{-n}$$

und somit ergibt sich für den Erwartungswert

$$\begin{aligned}
 E[X] &= \sum_{k=0}^n \binom{n}{k} \cdot 2^{-n} \cdot k \\
 &= \sum_{k=1}^n \binom{n-1}{k-1} \cdot n \cdot 2^{-n} \\
 &= \sum_{k=0}^n \binom{n-1}{k} \cdot n \cdot 2^{-n} \\
 &= n \cdot 2^{-n} \cdot \sum_{k=0}^n \binom{n-1}{k} \\
 &= \frac{n}{2}.
 \end{aligned}$$

Satz 2.2. Seien X und Y Zufallsvariablen über dem gleichen Ereignisraum S , dann gilt

$$E[X + Y] = E[X] + E[Y].$$

Beweis.

$$\begin{aligned}
 E[X + Y] &= \sum_{e \in S} p(e)(X + Y)(e) \\
 &= \sum_{e \in S} p(e)(X(e) + Y(e)) \\
 &= \sum_{e \in S} p(e)X(e) + \sum_{e \in S} p(e)Y(e) \\
 &= E(X) + E(Y)
 \end{aligned}$$

■

Beispiel 2.8. Im Folgenden betrachten wir das zweimalige Werfen eines Würfels und betrachten jeweils die folgenden verschiedenen Zufallsvariablen.

$X : S \rightarrow \mathbb{R}$ bildet ein Elementarereignis auf die Augensumme ab.

$X_1 : S \rightarrow \mathbb{R}$ bildet ein Elementarereignis auf die Augenzahl des ersten Wurfes ab

$X_2 : S \rightarrow \mathbb{R}$ bildet ein Elementarereignis auf die Augenzahl des zweiten Wurfes ab

Es gilt nun offensichtlich $X = X_1 + X_2$ und wir erhalten mit Satz 2.2 für den Erwartungswert

$$E[X] = E[X_1 + X_2] = E[X_1] + E[X_2] = 3.5 + 3.5 = 7.$$

Beispiel 2.9. Betrachten wir nun n Würfe mit einer fairen Münze, d.h. $S = \{K, Z\}^n$. Wir nehmen jetzt folgende Zufallsvariablen unter die Lupe

$$X_i = \begin{cases} 1, & \text{falls } i\text{-ter Wurf Kopf ist} \\ 0, & \text{falls } i\text{-ter Wurf Zahl ist} \end{cases} \quad i \in \{1, \dots, n\}.$$

Für jede dieser Zufallsvariablen erhalten wir den Erwartungswert

$$E[X_i] = \frac{1}{2}, \quad i \in \{1, \dots, n\}.$$

Wir definieren die Zufallsvariable X als $X = \sum_{i=1}^n X_i$ und erhalten mit Satz 2.2 einen Erwartungswert von

$$E[X] = \sum_{i=1}^n E[X_i] = \sum_{i=1}^n \frac{1}{2} = \frac{n}{2}.$$

Bemerkung 2.2. An dieser Stelle sollte der Leser nicht in die Hoffnung verfallen, dass sich mathematische Operationen immer auf so einfache Art und Weise auf Zufallsvariablen und ihre Erwartungswerte übertragen lassen.

Im Allgemeinen gilt für zwei Zufallsvariablen X und Y

$$E[XY] \neq E[X] \cdot E[Y].$$

Es gibt jedoch Ausnahmen die wir im Folgenden näher beschreiben werden.

Definition 2.7. Zwei Zufallsvariablen X und Y über dem gleichen Ereignisraum heißen **unabhängig**, wenn für alle $x, y \in \mathbb{R}$ gilt

$$p(\{X = x\} \cap \{Y = y\}) = p(\{X = x\}) \cdot p(\{Y = y\}).$$

Satz 2.3. Wenn X und Y unabhängig sind, dann gilt

$$E[XY] = E[X] \cdot E[Y].$$

Beweis.

$$\begin{aligned} E[XY] &= \sum_{e \in S} p(e)(XY)(e) \\ &= \sum_{z \in \mathbb{R}} p(\{XY = z\})z \\ &= \sum_{z \in \mathbb{R}} \sum_{x, y; xy=z} p(\{X = x\} \cap \{Y = y\})xy \\ &= \sum_{x, y} p(\{X = x\} \cap \{Y = y\})xy \\ &= \sum_{x, y} p(\{X = x\}) \cdot p(\{Y = y\})xy \\ &= \sum_x p(\{X = x\})x \sum_y p(\{Y = y\})y \\ &= E[X] \cdot E[Y]. \end{aligned}$$

2.2.2 Varianz

Häufigkeitsverteilungen sind durch ihren Erwartungswert nicht vollständig charakterisiert. Betrachten wir beispielsweise die beiden folgenden Verteilungen so erkennen wir direkt, dass sie beide den gleichen Erwartungswert besitzen, jedoch unterschiedlich um diesen streuen. Um diesen Sachverhalt betrachten zu können benötigen wir den Begriff der Varianz einer Zufallsvariablen.

Definition 2.8. Sei X eine Zufallsvariable. Dann heißt

$$\text{Var}[X] := E[(X - E[X])^2]$$

die **Varianz** von X . Die positive Wurzel der Varianz heißt **Standardabweichung** und wird mit σ bezeichnet.

Beispiel 2.10. Simulieren wir wieder n Würfe mit einer fairen Münze und betrachten wir die Zufallsvariablen aus Beispiel 2.9

$$X_i = \begin{cases} 1, & \text{falls } i\text{-ter Wurf Kopf ist} \\ 0, & \text{falls } i\text{-ter Wurf Zahl ist} \end{cases} \quad i \in \{1, \dots, n\},$$

so erhalten wir für die Varianz

$$\text{Var}[X_i] = \frac{1}{2} \cdot \left(1 - \frac{1}{2}\right)^2 + \frac{1}{2} \cdot \left(0 - \frac{1}{2}\right)^2 = \frac{1}{4}$$

für $i \in \{1, \dots, n\}$.

Lemma 2.1. Sei X eine Zufallsvariable über einem Wahrscheinlichkeitsraum, dann gilt für die Varianz von X

$$\text{Var}[X] = E[X^2] - (E[X])^2.$$

Beweis.

$$\begin{aligned} \text{Var}[X] &= E[(X - E[X])^2] \\ &= E[X^2 - 2E[X]X + (E[X])^2] \\ &= E[X^2] - 2E[X]E[X] + (E[X])^2 \\ &= E[X^2] - (E[X])^2. \end{aligned}$$

■

Satz 2.4. Wenn X und Y unabhängig sind, dann gilt

$$\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y].$$

Beweis.

$$\begin{aligned}
 \text{Var}[X + Y] &= E[(X + Y)^2] - (E[X + Y])^2 \\
 &= E[X^2 + 2XY + Y^2] - (E[X] + E[Y])^2 \\
 &= E[X^2] + 2E[XY] + E[Y^2] - (E[X])^2 - 2E[X]E[Y] - (E[Y])^2 \\
 &= E[X^2] - (E[X])^2 + E[Y^2] - (E[Y])^2 + 2E[X]E[Y] - 2E[X]E[Y] \\
 &= \text{Var}[X] + \text{Var}[Y].
 \end{aligned}$$

■

Beispiel 2.11. Durchführung des Experiments n Würfe mit einer fairen Münzen. Sei

$$X = X_1 + X_2 + \dots + X_n$$

die Zahl der Köpfe. Die Zufallsvariablen X_i sind unabhängig. Dann erhalten wir mit Satz 2.4

$$\text{Var}[X] = \sum_{i=1}^n \text{Var}[X_i] = \sum_{i=1}^n \frac{1}{4} = \frac{n}{4}.$$

Satz 2.5. (Markow Ungleichung)

Sei X eine nicht-negative Zufallsvariable. Dann gilt für alle Konstanten $t \geq 1$

$$p(\{X \geq t \cdot E[X]\}) \leq \frac{1}{t}.$$

Beweis.

$$\begin{aligned}
 E[X] &= \sum_{e \in S} p(e) \cdot X(e) \geq \sum_{e \in S, X(e) \geq t \cdot E[X]} p(e) \cdot X(e) \\
 &\geq \sum_{e \in S, X(e) \geq t \cdot E[X]} p(e) \cdot t \cdot E[X] \\
 &= t \cdot E[X] \cdot \sum_{e \in S, X(e) \geq t \cdot E[X]} p(e) \\
 &= t \cdot E[X] \cdot p(\{X \geq t \cdot E[X]\}).
 \end{aligned}$$

Division durch $t \cdot E[X]$ liefert die Behauptung. ■

Bemerkung 2.3. Die obige Abschätzung ist sehr schwach. Wenn man etwas über die Varianz weiß, dann kann man bessere Abschätzungen erhalten. Doch leider reichen die vorhandenen Informationen nicht für eine bessere Abschätzung aus.

Satz 2.6. (Chebycheff Ungleichung)

Sei X eine nicht-negative Zufallsvariable. Dann gilt für alle Konstanten $t \geq 1$

$$p(\{|X - E[X]| \geq t\}) \leq \frac{\text{Var}[X]}{t^2}$$

bzw.

$$p(\{|X - E[X]| \geq t \cdot \sqrt{\text{Var}[X]}\}) \leq \frac{1}{t^2}.$$

Beweis.

$$\begin{aligned}
 \text{Var}[X] &= E[(X - E[X])^2] \\
 &= \sum_{e \in S} p(e)(X - E[X])^2(e) \\
 &= \sum_{e \in S} p(e)(X(e) - E[X])^2 \\
 &\geq \sum_{e \in S, |X(e) - E[X]| \geq t} p(e) \cdot t^2 \\
 &= t^2 \cdot \sum_{e \in S, |X(e) - E[X]| \geq t} p(e) \\
 &= p(\{|X - E[X]| \geq t\}).
 \end{aligned}$$

■

Beispiel 2.12. Schauen wir wieder auf das Beispiel n Würfe einer fairen Münzen. Für die Anzahl X der Köpfe gilt $E[X] = n/2$ und $\text{Var}[X] = n/4$. Damit ist $\sigma = \sqrt{n/4}$ und daher gilt für $t = 10$

$$p(\{|X - \frac{n}{2}| \geq 5 \cdot \sqrt{n}\}) \leq \frac{1}{100}.$$

Daher ist X mit Wahrscheinlichkeit $1 - (1/100)$ in dem folgenden Intervall

$$x \in \left[\frac{n}{2} - 5 \cdot \sqrt{n}, \frac{n}{2} + 5 \cdot \sqrt{n} \right]$$

enthalten.

Satz 2.7. (Chernoff Schranke)

Sei X die Anzahl der Köpfe in n Würfen einer fairen Münze und sei $\epsilon \geq 0$. Dann gilt

$$p(\{|X - \frac{n}{2}| \geq \epsilon \cdot \frac{n}{2}\}) \leq 2 \cdot e^{-\epsilon^2 \cdot \frac{n}{4}}.$$

Beweis. Buch von Motwani/Ragharan über Randomized Algorithms. ■

Beispiel 2.13. Schätzen wir nun wieder die Wahrscheinlichkeit für Beispiel 2.12 ab. So erhalten wir für $\epsilon = \frac{10}{\sqrt{n}}$

$$p(\{|X - \frac{n}{2}| \geq 5 \cdot \sqrt{n}\}) \leq 2 \cdot e^{-25},$$

und sehen, dass diese Abschätzung erheblich stärker als die Chebycheff Schranke ist.

2.3 Zusammenfassung einiger kombinatorischer Formeln

In der Kombinatorik wird untersucht, wie eine gegebene Anzahl von Elementen unterschiedlich angeordnet und zu Gruppen zusammengefasst werden kann. Die Wahrscheinlichkeit für das Eintreten des Ereignisses A wird wie folgt definiert

$$p(A) = \frac{|A|}{|S|} = \frac{\text{Anzahl der für } A \text{ günstigen Elementarereignisse}}{\text{Anzahl aller möglichen Elementarereignisse}}$$

wobei unendliche Anzahlen ausgeschlossen werden und davon ausgegangen wird, dass jedes Elementarereignis die gleiche Eintrittschance (Wahrscheinlichkeit) hat. Um jetzt die Anzahl der für A günstigen Elementarereignisse, bzw. die Anzahl aller möglichen Elementarereignisse zu bestimmen benötigen wir einige Ergebnisse der Kombinatorik.

Im Wesentlichen beschäftigen wir uns nun mit den vier typischen Kombinationen von Auswahlverfahren.

1. Anzahl an Kombinationen von k aus n Elementen mit Zurücklegen und mit Berücksichtigung der Reihenfolge
2. Anzahl an Kombinationen von k aus n Elementen ohne Zurücklegen und mit Berücksichtigung der Reihenfolge
3. Anzahl an Kombinationen von k aus n Elementen ohne Zurücklegen und ohne Berücksichtigung der Reihenfolge
4. Anzahl an Kombinationen von k aus n Elementen mit Zurücklegen und ohne Berücksichtigung der Reihenfolge

Kombinatorische Probleme dieser Art werden häufig anhand von Urnen- oder Schubladenmodellen beschrieben. Dort werden die Kombinationen dann als das Ziehen von Kugeln aus einer Urne oder als Ablage von Kugeln in Schubladen betrachtet. Näheres zu diesen Vorstellungsmodellen findet der Leser in [?]. Wir werden im Folgenden für die oben erwähnten Fälle die allgemeingültigen Formeln an Beispielen herleiten, eine allgemeine Herleitung kann in jedem Analysisbuch oder Kombinatorikbuch nachgelesen werden.

Ihr Studium hat Sie hoffentlich schon zum stolzen Besitzer von fünf Mathematik und Informatikbüchern gemacht (mindestens!!). Die wollen Sie nun auf Ihrem neuen Regal eines schwedischen Möbeldiscounters unterbringen. Wieviele Möglichkeiten (im Sinne von Reihenfolgen) gibt es, die 5 Bücher nebeneinander auf das Regal zu stellen?

Für den ersten Platz haben Sie die Auswahl unter fünf Büchern, denn noch haben Sie alle Bücher in der Hand und das Regal ist leer. Wieviel Möglichkeiten bleiben Ihnen für den zweiten Platz auf dem Regal? Vier, da Sie nur noch vier Bücher auf der Hand haben. Führen wir diese Vorgehen fort, so erhalten wir folgendes erstes Ergebnis für die Anzahl der möglichen Anordnungen (Permutationen) Ihrer Bücher.

Anzahl der Permutationen von 5 Elementen ist $5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 120$.

Allgemein erhalten wir mit dieser Logik für die Anzahl der Permutationen von n Elementen

Anzahl der Permutationen von n Elementen ist $n \cdot (n - 1) \cdot \dots \cdot 2 \cdot 1 = n!$.

Neben diesem Permutationsbegriff werden wir im Folgenden auch häufig von Kombinationen sprechen. Während eine Permutation eine mögliche Anordnung aller Elemente einer gegebenen Menge n darstellt, ist eine Kombination eine Anordnung einer Teilmenge der Grösse k aus der Menge n .

2.3.1 Geordnete Stichprobe ohne Zurücklegen

Beim Aufstellen der Bücher bemerken Sie, dass leider nur drei der fünf Bücher auf Ihrem neuen Regal, das zwar gut aussieht, aber wohl doch etwas klein geraten ist, platz haben. Sie stehen also nun vor dem Entscheidungsproblem aus den fünf Büchern drei auszuwählen und diese Teilmenge dann auf dem Regal anzuordnen. Aus Ihrer noch vorhandenen Unwissenheit heraus bleibt Ihnen nichts anderes übrig als alle Möglichkeiten für sinnvoll zu erachten. Auf ein Urnenmodell übertragen handelt es sich also um ein dreimaliges Ziehen aus einer Urne mit fünf Kugeln die mit den Nummern von 1 bis 5 beschriftet sind. Dies geschieht ohne Zurücklegen und mit Berücksichtigung der Reihenfolge, da sie jedes Buch nur einmal auf das Regal stellen können und Ihnen die Reihenfolge wichtig ist, auch wenn Sie noch nicht so ganz wissen warum.

Wieviele Kombinationen der Anordnung gibt es in diesem Fall?

Betrachten wir die Situation wieder schrittweise. Für den ersten Platz haben Sie 5 Bücher zur Wahl, für den zweiten noch 4 und für den dritten noch 3, d.h.

Anzahl der Kombinationen von 3 aus 5 Elementen ohne Zurücklegen mit Berücksichtigung der Reihenfolge ist $5 \cdot 4 \cdot 3 = 60$.

Wir haben also keine vollständige Permutation, wie in unserer ersten Betrachtung sondern nur eine Kombination, allgemein heisst das

Anzahl der Kombinationen von k aus n Elementen ohne Zurücklegen mit Berücksichtigung der Reihenfolge ist $\frac{n!}{(n-k)!}$.

2.3.2 Ungeordnete Stichprobe ohne Zurücklegen

Da diese Anzahl an Möglichkeiten Sie in Ihrer momentanen Situation überfordert und Sie durch das Lösen der diversen Übungsblätter keine Entscheidungskraft mehr aufbringen können, entscheiden Sie sich doch dazu, dass Ihnen die Reihenfolge der Bücher auf Ihrem Regal egal ist, da Ihnen ausser Ihren Übungen und Vorlesungen auch alles andere in Ihrem Leben egal geworden ist. Ihr Entscheidungsproblem reduziert sich also auf die Frage welche drei Bücher Sie auf Ihr neues Regal stellen. Wir sind also jetzt im Fall ohne Zurücklegen und ohne Berücksichtigung der Reihenfolge.

Wieviele Kombinationen der Anordnung gibt es in diesem Fall noch?

Wir haben schon gelernt, dass die Anzahl der Permutationen von k Elementen $k!$ ist. Für jede dreier Auswahl gäbe es also $3! = 6$ Anordnungsmöglichkeiten, die wir nun nicht mehr betrachten wollen. Gehen wir nun von dem geordneten Fall mit seinen 60 Kombinationen aus, so reduziert sich ohne die Reihenfolge die Anzahl der Kombinationen auf ein Sechstel, da alle Permutationen der drei gewählten Bücher als identisch betrachtet werden. Wir erhalten also in unserem Beispiel

Anzahl der Kombinationen von 3 aus 5 Elementen ohne Zurücklegen ohne Berücksichtigung der Reihenfolge ist $\frac{60}{6} = 10$.

und im Allgemeinen

Anzahl der Kombinationen von k aus n Elementen ohne Zurücklegen ohne Berücksichtigung der Reihenfolge ist $\frac{n!}{(n-k)! \cdot k!} = \binom{n}{k}$.

2.3.3 Geordnete Stichprobe mit Zurücklegen

Die ersten Wochen des Semesters sind vorbei und nun sollte jedem klar sein, dass das benötigte Wissen sich nicht in fünf Büchern zusammenfassen lässt. Vielmehr scheint der Quellfluss an Informationen unendlich zu sein und nie zu stoppen, da Sie mittlerweile die Fähigkeit erlangt haben Skripte und Bücher aus der Bibliothek und dem Netz zu ergattern. Betrachten wir in dieser Situation unser Buchproblem von neuem, so stellt sich nun nicht mehr die Frage, wie wir unsere spärlichen fünf eigenen Bücher anordnen, sondern welche und wieviele Bücher wir uns aus den verschiedenen Bereichen auf unserem Regal anordnen. Betrachten wir in einer Kernausswahl Bücher aus folgenden fünf Fachgebieten

Analysis, Algebra, Wahrscheinlichkeitstheorie, Maßtheorie und C++,

die wir in beliebiger Fülle in der Bibliothek finden oder wenigstens finden sollten. Wir sind jetzt also in dem Fall mit Zurücklegen und mit Berücksichtigung der Reihenfolge, da wir immer wieder nach einem Buch der gleichen Kategorie greifen und in unserem Regal platzieren können.

Wieviele Kombinationen der Anordnung gibt es in diesem Fall?

Für den ersten Platz auf unserem Regal haben wir ein Buch der 5 Gebiete zur Auswahl, für den zweiten Platz haben wir wieder ein Buch der fünf Gebiete zur Auswahl, da unserer Bibliothek die Bücher ja nie ausgehen. Führen wir diese Gedankenmodell fort so erhalten wir

Anzahl der Kombinationen von 3 aus 5 Elementen mit Zurücklegen mit Berücksichtigung der Reihenfolge ist $5 \cdot 5 \cdot 5 = 125$.

für unser Beispiel und allgemein die Formel

Anzahl der Kombinationen von k aus n Elementen mit Zurücklegen mit Berücksichtigung der Reihenfolge ist n^k .

2.3.4 Ungeordnete Stichprobe mit Zurücklegen

Wieder taucht das alte Problem auf, die Studenten sind überfordert, die Übungsblätter viel zu schwer und Ordnung herrscht in Ihrem Leben schon lange nicht mehr. Also lassen wir die Ordnung mal wieder wie überall in Ihrem Zimmer auch auf Ihrem Regal unter den Tisch fallen. Es stellt sich also nur noch die Frage welche Bücher auf Ihrem Regal landen, die Reihenfolge ist vollkommen egal. Betrachten wir hierzu erst einmal den uns schon bekannten Fall ungeordnet, ohne Zurücklegen. Dort konnten wir jedes der 5 Gebiete nur einmal auswählen, wir hatten also die Wahl 3 aus 5. Im Falle mit Zurücklegen erhöht sich jedoch unsere Auswahlmöglichkeit, da wir immer wieder aus allen Gebieten ein Buch wählen können. Wir wählen also nicht k aus n sondern k aus $n + k - 1$. Was uns in unserem Beispiel auf

Anzahl der Kombinationen von 3 aus 5 Elementen mit Zurücklegen ohne Berücksichtigung der Reihenfolge ist $\binom{7}{3} = 35$.

führt und im Allgemeinen

Anzahl der Kombinationen von k aus n Elementen mit Zurücklegen ohne Berücksichtigung der Reihenfolge ist $\binom{n+k-1}{k}$.

Ein alternativer Zugang zu dieser Formel erschließt sich, wenn wir unser Vorgehen etwas abändern. Nehmen wir an Sie konnten an der Kasse des schwedischen Möbeldiscounters nicht widerstehen und haben sich zu Ihrem Regal noch vier unnötige Buchstützen gekauft. Aber vielleicht waren diese ja nicht vollkommen sinnlos. Stellen wir uns jetzt vor sie wollten drei Bücher aus den fünf Fachgebieten auf Ihrem Regal platzieren und die Buchstützen trennen die verschiedenen Fachgebiete voneinander ab. So beschreibt jetzt die Anzahl an Anordnungen der drei Bücher und der vier Buchstützen alle verschiedenen Möglichkeiten diese auf die Fachgebiete zu verteilen. Wie wir bei Permutationen schon gelernt haben gibt es $7!$ Möglichkeiten die sieben Gegenstände hintereinander zu ordnen. Dies setzt jedoch voraus, dass wir alle sieben Gegenstände unterscheiden können. Da dies nicht der Fall ist, müssen wir die zu viel gezählten Permutationen wieder eliminieren. Wir befinden uns nun in einem klassischen Problem der Kombinatorik, das unter dem Namen Mississippi-Problem in die Literatur eingegangen ist. Wir müssen also die unterschiedlichen Reihenfolgen der identischen Gegenstände wieder aus unserer Gesamtanzahl heraus dividieren, d.h. $3!$ für die Bücher und $4!$ für die Buchstützen. Insgesamt erhalten wir somit

Anzahl der Kombinationen von 3 aus 5 Elementen mit Zurücklegen ohne Berücksichtigung der Reihenfolge ist $\frac{7!}{3!4!} = \binom{7}{3} = 35$.

und auch allgemein kommen wir auf unsere angegebene Formel

Anzahl der Kombinationen von k aus n Elementen mit Zurücklegen ohne Berücksichtigung der Reihenfolge ist $\frac{(n+k-1)!}{k!(n-k)!} = \binom{n+k-1}{k}$.

2.3.5 Zusammenfassung

Nach der Lektüre dieses Kapitels sollte Ihnen klar geworden sein, dass Sie sich mindestens ein Wahrscheinlichkeitsbuch anschaffen müssen und dieses immer an erster Stelle in Ihrem Regal platzieren sollten. Für alle die dies immer noch nicht glauben wollen: „Pech“ und wir hoffen Sie kommen mit einer Zusammenfassung der von uns hergeleiteten Formeln für die Anzahl an Kombinationen von k aus n aus der Kombinatorik über die Runden.

$X(e) = a + b = k$	2	3	4	5	6	7	8	9	10	11	12
$p(\{X = k\})$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Tabelle 2.1: Wahrscheinlichkeitsverteilung

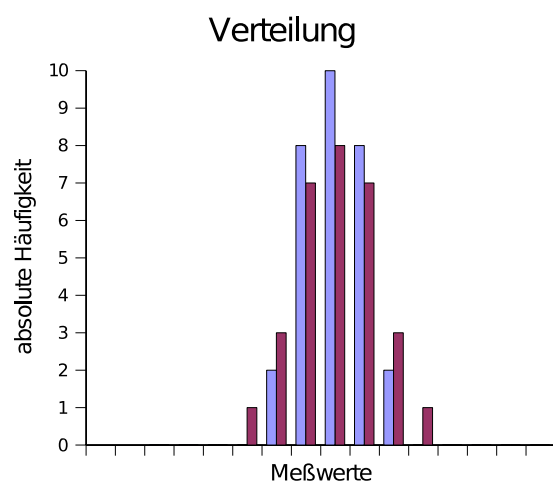


Abbildung 2.1: Verteilungen mit unterschiedlicher Streuung

Stichprobe k aus n	mit Zurücklegen	ohne Zurücklegen
Geordnet	n^k	$\frac{n!}{(n-k)!}$
Ungeordnet	$\binom{n+k-1}{k}$	$\binom{n}{k}$

Tabelle 2.2: Formeln der Kombinatorik