

Beispiel 62

Chebyshev-Ungleichung:

$$\Pr[X \geq 550] \leq \Pr[|X - 500| \geq 50] \leq \frac{250}{50^2} = 0,1.$$

Setze nun $n = 10000$ und betrachte wieder eine maximal 10%-ige Abweichung vom Erwartungswert:

$$\mathbb{E}[X] = 5000 \text{ und } \text{Var}[X] = 2500, \text{ und damit}$$
$$\Pr[X \geq 5500] \leq \Pr[|X - 5000| \geq 500] \leq \frac{2500}{500^2} = 0,01.$$

6.2 Gesetz der großen Zahlen

Wir haben diskutiert, wie Wahrscheinlichkeiten als Grenzwerte von relativen Häufigkeiten aufgefasst werden können.

Satz 63 (Gesetz der großen Zahlen)

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

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

$$\Pr[|Z - \mathbb{E}[X]| \geq \delta] \leq \varepsilon.$$

Beweis:

Für Z gilt

$$\mathbb{E}[Z] = \frac{1}{n} \cdot (\mathbb{E}[X_1] + \dots + \mathbb{E}[X_n]) = \frac{1}{n} \cdot n \cdot \mathbb{E}[X] = \mathbb{E}[X],$$

sowie

$$\text{Var}[Z] = \frac{1}{n^2} \cdot (\text{Var}[X_1] + \dots + \text{Var}[X_n]) = \frac{1}{n^2} \cdot n \cdot \text{Var}[X] = \frac{\text{Var}[X]}{n}.$$

Mit der Chebyshev-Ungleichung erhalten wir

$$\Pr[|Z - \mathbb{E}[X]| \geq \delta] = \Pr[|Z - \mathbb{E}[Z]| \geq \delta] \leq \frac{\text{Var}[Z]}{\delta^2} = \frac{\text{Var}[X]}{n\delta^2} \leq \varepsilon,$$

nach Wahl von n .



Wahrscheinlichkeit und relative Häufigkeit.

Sei X eine Indikatorvariable für ein Ereignis A , $\Pr[A] = p$. Somit ist X Bernoulli-verteilt mit $\mathbb{E}[X] = p$.

$Z = \frac{1}{n}(X_1 + \dots + X_n)$ gibt die relative Häufigkeit an, mit der A bei n Wiederholungen des Versuchs eintritt, denn

$$Z = \frac{\text{Anzahl der Versuche, bei denen } A \text{ eingetreten ist}}{\text{Anzahl aller Versuche}}.$$

Mit Hilfe des obigen Gesetzes der großen Zahlen folgt

$$\Pr[|Z - p| \geq \delta] \leq \varepsilon,$$

für genügend großes n . Also nähert sich die relative Häufigkeit von A bei hinreichend vielen Wiederholungen des Experiments mit beliebiger Sicherheit beliebig nahe an die „wahre“ Wahrscheinlichkeit p an.

Die obige Variante eines **Gesetzes der großen Zahlen** geht auf **Jakob Bernoulli** zurück, der den Satz in seinem Werk **ars conjectandi** zeigte.

Es soll betont werden, dass das Gesetz der großen Zahlen die

$$\text{relative Abweichung } \left| \frac{1}{n} \sum_i X_i - p \right|$$

und nicht die

$$\text{absolute Abweichung } \left| \sum_i X_i - np \right|$$

abschätzt!

6.3 Chernoff-Schranken

6.3.1 Chernoff-Schranken für Summen von 0–1–Zufallsvariablen

Die hier betrachtete Art von Schranken ist nach **Herman Chernoff** (*1923) benannt. Sie finden in der Komplexitätstheoretischen Analyse von Algorithmen eine sehr häufige Verwendung.

Satz 64

Seien X_1, \dots, X_n unabhängige Bernoulli-verteilte Zufallsvariablen mit $\Pr[X_i = 1] = p_i$ und $\Pr[X_i = 0] = 1 - p_i$. Dann gilt für $X := \sum_{i=1}^n X_i$ und $\mu := \mathbb{E}[X] = \sum_{i=1}^n p_i$, sowie jedes $\delta > 0$, dass

$$\Pr[X \geq (1 + \delta)\mu] \leq \left(\frac{e^\delta}{(1 + \delta)^{1+\delta}} \right)^\mu.$$

Beweis:

Für $t > 0$ gilt

$$\Pr[X \geq (1 + \delta)\mu] = \Pr[e^{tX} \geq e^{t(1+\delta)\mu}].$$

Mit der Markov-Ungleichung folgt

$$\Pr[X \geq (1 + \delta)\mu] = \Pr[e^{tX} \geq e^{t(1+\delta)\mu}] \leq \frac{\mathbb{E}[e^{tX}]}{e^{t(1+\delta)\mu}}.$$

Wegen der Unabhängigkeit der Zufallsvariablen X_1, \dots, X_n gilt

$$\mathbb{E}[e^{tX}] = \mathbb{E}\left[\exp\left(\sum_{i=1}^n tX_i\right)\right] = \mathbb{E}\left[\prod_{i=1}^n e^{tX_i}\right] = \prod_{i=1}^n \mathbb{E}[e^{tX_i}].$$

Weiter ist für $i \in \{1, \dots, n\}$:

$$\mathbb{E}[e^{tX_i}] = e^{t \cdot 1} p_i + e^{t \cdot 0} (1 - p_i) = e^t p_i + 1 - p_i = 1 + p_i(e^t - 1),$$

Beweis (Forts.):

und damit

$$\begin{aligned}\Pr[X \geq (1 + \delta)\mu] &\leq \frac{\prod_{i=1}^n (1 + p_i(e^t - 1))}{e^{t(1+\delta)\mu}} \\ &\leq \frac{\prod_{i=1}^n \exp(p_i(e^t - 1))}{e^{t(1+\delta)\mu}} \\ &= \frac{\exp(\sum_{i=1}^n p_i(e^t - 1))}{e^{t(1+\delta)\mu}} = \frac{e^{(e^t-1)\mu}}{e^{t(1+\delta)\mu}} =: f(t).\end{aligned}$$

Wir wählen nun t so, dass $f(t)$ minimiert wird, nämlich

$$t = \ln(1 + \delta).$$

Damit wird

$$f(t) = \frac{e^{(e^t-1)\mu}}{e^{t(1+\delta)\mu}} = \frac{e^{\delta\mu}}{(1 + \delta)^{(1+\delta)\mu}}.$$

□

Beispiel 65

Wir betrachten wieder das Beispiel, dass wir eine faire Münze n -mal werfen und abschätzen wollen, mit welcher Wahrscheinlichkeit „Kopf“

$$\frac{n}{2}(1 + 10\%)$$

oder öfter fällt.

n	Chebyshev	Chernoff
1000	0,1	0,0889
10000	0,01	$0,308 \cdot 10^{-10}$
n	$\frac{\frac{1}{4}n}{(0,1 \cdot \frac{1}{2}n)^2}$	$\left(\frac{e^{0,1}}{(1+0,1)^{1+0,1}} \right)^{\frac{1}{2}n}$

Satz 66

Seien X_1, \dots, X_n unabhängige Bernoulli-verteilte Zufallsvariablen mit $\Pr[X_i = 1] = p_i$ und $\Pr[X_i = 0] = 1 - p_i$. Dann gilt für $X := \sum_{i=1}^n X_i$ und $\mu := \mathbb{E}[X] = \sum_{i=1}^n p_i$, sowie jedes $0 < \delta < 1$, dass

$$\Pr[X \leq (1 - \delta)\mu] \leq \left(\frac{e^{-\delta}}{(1 - \delta)^{1-\delta}} \right)^\mu.$$

Beweis:

Analog zum Beweis von Satz 64. □

Bemerkung: Abschätzungen, wie sie in Satz 64 und Satz 66 angegeben sind, nennt man auch **tail bounds**, da sie Schranken für die **tails**, also die vom Erwartungswert weit entfernten Bereiche angeben. Man spricht hierbei vom **upper tail** (vergleiche Satz 64) und vom **lower tail** (vergleiche Satz 66).

Die Chernoff-Schranken hängen **exponentiell** von μ ab!

Lemma 67

Für $0 \leq \delta < 1$ gilt

$$(1 - \delta)^{1-\delta} \geq e^{-\delta+\delta^2/2} \quad \text{und} \quad (1 + \delta)^{1+\delta} \geq e^{\delta+\delta^2/3}.$$

Beweis:

Wir betrachten

$$f(x) = (1 - x) \ln(1 - x) \quad \text{und} \quad g(x) = -x + \frac{1}{2}x^2.$$

Es gilt für $0 \leq x < 1$:

$$g'(x) = x - 1 \leq -\ln(1 - x) - 1 = f'(x)$$

sowie

$$f(0) = 0 = g(0),$$

also im angegebenen Intervall $f(x) \geq g(x)$.

Die Herleitung der zweiten Ungleichung erfolgt analog. □

Korollar 68

Seien X_1, \dots, X_n unabhängige Bernoulli-verteilte Zufallsvariablen mit $\Pr[X_i = 1] = p_i$ und $\Pr[X_i = 0] = 1 - p_i$. Dann gelten folgende Ungleichungen für $X := \sum_{i=1}^n X_i$ und $\mu := \mathbb{E}[X] = \sum_{i=1}^n p_i$:

- 1 $\Pr[X \geq (1 + \delta)\mu] \leq e^{-\mu\delta^2/3}$ für alle $0 < \delta \leq 1$,
- 2 $\Pr[X \leq (1 - \delta)\mu] \leq e^{-\mu\delta^2/2}$ für alle $0 < \delta \leq 1$,
- 3 $\Pr[|X - \mu| \geq \delta\mu] \leq 2e^{-\mu\delta^2/3}$ für alle $0 < \delta \leq 1$,
- 4 $\Pr[X \geq (1 + \delta)\mu] \leq \left(\frac{e}{1+\delta}\right)^{(1+\delta)\mu}$ und
- 5 $\Pr[X \geq t] \leq 2^{-t}$ für $t \geq 2e\mu$.

Beweis:

1 und 2 folgen direkt aus Satz 64 bzw. 66 und Lemma 67.

Aus 1 und 2 zusammen folgt 3.

Die Abschätzung 4 erhalten wir direkt aus Satz 64, da für den Zähler gilt

$$e^\delta \leq e^{(1+\delta)}.$$

5 folgt aus 4, indem man $t = (1 + \delta)\mu$ setzt, $t \geq 2e\mu$:

$$\left(\frac{e}{1+\delta}\right)^{(1+\delta)\mu} \leq \left(\frac{e}{t/\mu}\right)^t \leq \left(\frac{1}{2}\right)^t.$$

□

Beispiel 69

Wir betrachten wieder **balls into bins** und werfen n Bälle unabhängig und gleichverteilt in n Körbe. Sei

$$X_i := \text{Anzahl der Bälle im } i\text{-ten Korb}$$

für $i = 1, \dots, n$, sowie $X := \max_{1 \leq i \leq n} X_i$.

Für die Analyse von X_i ($i \in \{1, \dots, n\}$ beliebig) verwenden wir Aussage 5 von Korollar 68, mit $p_1 = \dots = p_n = \frac{1}{n}$, $\mu = 1$ und $t = 2 \log n$. Es folgt

$$\Pr[X_i \geq 2 \log n] \leq 1/n^2.$$

Daraus ergibt sich

$$\Pr[X \geq 2 \log n] = \Pr[X_1 \geq 2 \log n \vee \dots \vee X_n \geq 2 \log n] \leq n \cdot \frac{1}{n^2} = \frac{1}{n}.$$

Es gilt also mit Wahrscheinlichkeit $\geq 1 - 1/n$, dass $X < 2 \log n$ ist.

Literatur:

-  Torben Hagerup, Christine Rüb:
A guided tour of Chernoff bounds
Inf. Process. Lett. **33**, pp. 305–308 (1990)

7. Erzeugende Funktionen

7.1 Einführung

Definition 70

Für eine Zufallsvariable X mit $W_X \subseteq \mathbb{N}_0$ ist die (wahrscheinlichkeits-)erzeugende Funktion definiert durch

$$G_X(s) := \sum_{k=0}^{\infty} \Pr[X = k] \cdot s^k = \mathbb{E}[s^X].$$

Die obige Definition gilt für allgemeine $s \in \mathbb{R}$, wir werden uns aber auf $s \in [-1, 1]$ konzentrieren.

Eine wahrscheinlichkeitserzeugende Funktion ist also die (gewöhnliche) erzeugende Funktion der Folge $(f_i)_{i \in \mathbb{N}_0}$ mit $f_i := \Pr[X = i]$.

Bei wahrscheinlichkeitserzeugenden Funktionen haben wir kein Problem mit der **Konvergenz**, da für $|s| < 1$ gilt

$$\begin{aligned} |G_X(s)| &= \left| \sum_{k=0}^{\infty} \Pr[X = k] \cdot s^k \right| \\ &\leq \sum_{k=0}^{\infty} \Pr[X = k] \cdot |s^k| \leq \sum_{k=0}^{\infty} \Pr[X = k] = 1. \end{aligned}$$

Beobachtung:

Sei $Y := X + t$ mit $t \in \mathbb{N}_0$. Dann gilt

$$G_Y(s) = \mathbb{E}[s^Y] = \mathbb{E}[s^{X+t}] = \mathbb{E}[s^t \cdot s^X] = s^t \cdot \mathbb{E}[s^X] = s^t \cdot G_X(s).$$

Ebenso lässt sich leicht nachrechnen, dass

$$G'_X(s) = \sum_{k=1}^{\infty} k \cdot \Pr[X = k] \cdot s^{k-1}, \text{ also}$$

$$G'_X(0) = \Pr[X = 1], \text{ sowie}$$

$$G_X^{(i)}(0) = \Pr[X = i] \cdot i!, \text{ also}$$

$$G_X^{(i)}(0)/i! = \Pr[X = i].$$

Satz 71 (Eindeutigkeit der w.e. Funktion)

Die Dichte und die Verteilung einer Zufallsvariablen X mit $W_X \subseteq \mathbb{N}$ sind durch ihre Wahrscheinlichkeitserzeugende Funktion eindeutig bestimmt.

Beweis:

Folgt aus der Eindeutigkeit der Potenzreihendarstellung. □

Bernoulli-Verteilung

Sei X eine Bernoulli-verteilte Zufallsvariable mit $\Pr[X = 0] = 1 - p$ und $\Pr[X = 1] = p$. Dann gilt

$$G_X(s) = \mathbb{E}[s^X] = (1 - p) \cdot s^0 + p \cdot s^1 = 1 - p + ps.$$

Gleichverteilung auf $\{0, \dots, n\}$

Sei X auf $\{0, \dots, n\}$ gleichverteilt, d.h. für $0 \leq k \leq n$ ist $\Pr[X = k] = 1/(n + 1)$.
Dann gilt

$$G_X(s) = \mathbb{E}[s^X] = \sum_{k=0}^n \frac{1}{n + 1} \cdot s^k = \frac{s^{n+1} - 1}{(n + 1)(s - 1)}.$$

Binomialverteilung

Für $X \sim \text{Bin}(n, p)$ gilt nach der binomischen Formel

$$G_X(s) = \mathbb{E}[s^X] = \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} \cdot s^k = (1-p+ps)^n.$$

Geometrische Verteilung

Sei X eine geometrisch verteilte Zufallsvariable mit Erfolgswahrscheinlichkeit p . Dann gilt

$$\begin{aligned} G_X(s) = \mathbb{E}[s^X] &= \sum_{k=1}^{\infty} p(1-p)^{k-1} \cdot s^k \\ &= ps \cdot \sum_{k=1}^{\infty} ((1-p)s)^{k-1} = \frac{ps}{1-(1-p)s}. \end{aligned}$$

Poisson-Verteilung

Für $X \sim \text{Po}(\lambda)$ gilt

$$G_X(s) = \mathbb{E}[s^X] = \sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} \cdot s^k = e^{-\lambda + \lambda s} = e^{\lambda(s-1)}.$$

Beispiel 72

Sei X binomialverteilt mit $X \sim \text{Bin}(n, \lambda/n)$, Für $n \rightarrow \infty$ folgt

$$G_X(s) = \left(1 - \frac{\lambda}{n} + \frac{\lambda s}{n}\right)^n = \left(1 + \frac{\lambda(s-1)}{n}\right)^n \rightarrow e^{\lambda(s-1)}.$$

Man kann beweisen, dass aus der Konvergenz der wahrscheinlichkeitserzeugenden Funktion die Konvergenz der Verteilung folgt.

7.1.1 Zusammenhang zwischen der w.e. Funktion und den Momenten

Da

$$G_X(s) := \sum_{k=0}^{\infty} \Pr[X = k] \cdot s^k = \mathbb{E}[s^X],$$

gilt

$$G'_X(1) = \sum_{k=1}^{\infty} k \cdot \Pr[X = k] = \mathbb{E}[X].$$

Beispiel 73

Sei X binomialverteilt mit $X \sim \text{Bin}(n, p)$, also

$$G_X(s) = (1 - p + ps)^n.$$

Dann gilt

$$G'_X(s) = n \cdot (1 - p + ps)^{n-1} \cdot p$$

und somit

$$\mathbb{E}[X] = G'_X(1) = np.$$

Beispiel 73

Ebenso ergibt sich

$$\mathbb{E}[X(X-1)\dots(X-k+1)] = G_X^{(k)}(1),$$

also etwa

$$\begin{aligned}\text{Var}[X] &= \mathbb{E}[X(X-1)] + \mathbb{E}[X] - \mathbb{E}[X]^2 \\ &= G_X''(1) + G_X'(1) - (G_X'(1))^2.\end{aligned}$$

Andere Momente von X kann man auf ähnliche Art und Weise berechnen.

Momenterzeugende Funktionen

Definition 74

Zu einer Zufallsvariablen X ist die **momenterzeugende Funktion** gemäß

$$M_X(s) := \mathbb{E}[e^{Xs}]$$

definiert.

Es gilt

$$M_X(s) = \mathbb{E}[e^{Xs}] = \mathbb{E}\left[\sum_{i=0}^{\infty} \frac{(Xs)^i}{i!}\right] = \sum_{i=0}^{\infty} \frac{\mathbb{E}[X^i]}{i!} \cdot s^i$$

und

$$M_X(s) = \mathbb{E}[e^{Xs}] = \mathbb{E}[(e^s)^X] = G_X(e^s).$$

7.2 Summen von Zufallsvariablen

Satz 75 (Erzeugende Funktion einer Summe)

Für unabhängige Zufallsvariablen X_1, \dots, X_n und die Zufallsvariable $Z := X_1 + \dots + X_n$ gilt

$$G_Z(s) = G_{X_1}(s) \cdot \dots \cdot G_{X_n}(s).$$

Ebenso gilt

$$M_Z(s) = M_{X_1}(s) \cdot \dots \cdot M_{X_n}(s).$$

Beweis:

Wegen der Unabhängigkeit von X_1, \dots, X_n gilt

$$G_Z(s) = \mathbb{E}[s^{X_1 + \dots + X_n}] = \mathbb{E}[s^{X_1}] \cdot \dots \cdot \mathbb{E}[s^{X_n}] = G_{X_1}(s) \cdot \dots \cdot G_{X_n}(s).$$



Beispiel 76

Seien X_1, \dots, X_k mit $X_i \sim \text{Bin}(n_i, p)$ unabhängige Zufallsvariable und $Z := X_1 + \dots + X_k$. Dann gilt

$$G_Z(s) = \prod_{i=1}^k (1 - p + ps)^{n_i} = (1 - p + ps)^{\sum_{i=1}^k n_i}$$

und somit

$$Z \sim \text{Bin}\left(\sum_{i=1}^k n_i, p\right)$$

(vgl. Satz 56).

Seien $X_1, \dots, X_k \sim \text{Po}(\lambda)$ unabhängige Zufallsvariablen. Dann folgt für $Z := X_1 + \dots + X_k$

$$G_Z(s) = \prod_{i=1}^k e^{\lambda(s-1)} = e^{k\lambda(s-1)}$$

und somit $Z \sim \text{Po}(k\lambda)$ (vgl. Satz 59).

7.2.1 Zufällige Summen

Wir betrachten die Situation, dass $Z := X_1 + \dots + X_N$, wobei N ebenfalls eine Zufallsvariable ist.

Satz 77

Seien X_1, X_2, \dots unabhängige und identisch verteilte Zufallsvariablen mit der Wahrscheinlichkeitserzeugenden Funktion $G_X(s)$. N sei ebenfalls eine unabhängige Zufallsvariable mit der Wahrscheinlichkeitserzeugenden Funktion $G_N(s)$. Dann besitzt die Zufallsvariable $Z := X_1 + \dots + X_N$ die Wahrscheinlichkeitserzeugende Funktion $G_Z(s) = G_N(G_X(s))$.

Beweis:

Nach Voraussetzung ist $W_N \subseteq \mathbb{N}_0$. Deshalb folgt mit Satz 36

$$\begin{aligned} G_Z(s) &= \sum_{n=0}^{\infty} \mathbb{E}[s^Z \mid N = n] \cdot \Pr[N = n] \\ &= \sum_{n=0}^{\infty} \mathbb{E}[s^{X_1 + \dots + X_n}] \cdot \Pr[N = n] \\ &= \sum_{n=0}^{\infty} \mathbb{E}[s^{X_1}] \cdot \dots \cdot \mathbb{E}[s^{X_n}] \cdot \Pr[N = n] \\ &= \sum_{n=0}^{\infty} (G_X(s))^n \cdot \Pr[N = n] \\ &= \mathbb{E}[(G_X(s))^N] \\ &= G_N(G_X(s)). \end{aligned}$$

□

8. Formelsammlung

8.1 Gesetze zum Rechnen mit Ereignissen

Im Folgenden seien A und B , sowie A_1, \dots, A_n Ereignisse. Die Notation $A \uplus B$ steht für $A \cup B$ und zugleich $A \cap B = \emptyset$ (disjunkte Vereinigung). $A_1 \uplus \dots \uplus A_n = \Omega$ bedeutet also, dass die Ereignisse A_1, \dots, A_n eine Partition der Ergebnismenge Ω bilden.

$$\Pr[\emptyset] = 0$$

$$0 \leq \Pr[A] \leq 1$$

$$\Pr[\bar{A}] = 1 - \Pr[A]$$

$$A \subseteq B \implies \Pr[A] \leq \Pr[B]$$

$\forall i \neq j : A_i \cap A_j = \emptyset \implies$ $\Pr \left[\bigcup_{i=1}^n A_i \right] = \sum_{i=1}^n \Pr[A_i]$	Additionssatz
$\Pr[A \cup B] = \Pr[A] + \Pr[B] - \Pr[A \cap B]$ allgemeine Form: siehe Satz 9	Inklusion/Exklusion, Siebformel
$\Pr \left[\bigcup_{i=1}^n A_i \right] \leq \sum_{i=1}^n \Pr[A_i]$	Boolesche Ungleichung
$\Pr[A B] = \frac{\Pr[A \cap B]}{\Pr[B]}$ für $\Pr[B] > 0$	Def. bedingte Ws.

$B \subseteq A_1 \uplus \dots \uplus A_n \implies$ $\Pr[B] = \sum_{i=1}^n \Pr[B A_i] \cdot \Pr[A_i]$	Satz von der totalen Wahrscheinlichkeit
$\Pr[B] > 0, B \subseteq A_1 \uplus \dots \uplus A_n \implies$ $\Pr[A_i B] = \frac{\Pr[B A_i] \cdot \Pr[A_i]}{\sum_{i=1}^n \Pr[B A_i] \cdot \Pr[A_i]}$	Satz von Bayes
$\Pr[A_1 \cap \dots \cap A_n] = \Pr[A_1] \cdot \Pr[A_2 A_1] \cdot$ $\dots \cdot \Pr[A_n A_1 \cap \dots \cap A_{n-1}]$	Multiplikationssatz
$A \text{ und } B \text{ unabhängig} \iff$ $\Pr[A \cap B] = \Pr[A] \cdot \Pr[B]$	Definition Unabhängigkeit

8.2 Erwartungswert und Varianz diskreter Zufallsvariablen

Sei X eine diskrete Zufallsvariable. Für Erwartungswert und Varianz gelten die folgenden Formeln (sofern $\mathbb{E}[X]$ und $\text{Var}[X]$ existieren).

$$\begin{aligned}\mathbb{E}[X] &= \sum_{x \in W_X} x \cdot \Pr[X = x] \\ &= \sum_{\omega \in \Omega} X(\omega) \cdot \Pr[\omega] && \text{Erwartungswert} \\ \left(&= \sum_{i=1}^{\infty} \Pr[X \geq i], \quad \text{falls } W_X \subseteq \mathbb{N}_0 \right)\end{aligned}$$

$$\begin{aligned}\text{Var}[X] &= \mathbb{E}[(X - \mathbb{E}[X])^2] \\ &= \sum_{x \in W_X} \Pr[X = x] \cdot (x - \mathbb{E}[X])^2 && \text{Varianz}\end{aligned}$$

8.3 Gesetze zum Rechnen mit Zufallsvariablen

Seien $a, b, a_1, \dots, a_n \in \mathbb{R}$, $f_1, \dots, f_n : \mathbb{R} \rightarrow \mathbb{R}$.

$$\begin{aligned} X_1, \dots, X_n \text{ unabhängig} &\iff \text{für alle } (a_1, \dots, a_n): \\ &\Pr[X_1 = a_1, \dots, X_n = a_n] \\ &= \Pr[X_1 = a_1] \cdot \dots \cdot \Pr[X_n = a_n] \end{aligned}$$

$$X_1, \dots, X_n \text{ unabhängig} \implies f_1(X_1), \dots, f_n(X_n) \text{ unabhängig}$$

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

$$X(\omega) \leq Y(\omega) \text{ für alle } \omega \in \Omega \implies \\ \mathbb{E}[X] \leq \mathbb{E}[Y]$$

Monotonie des
Erwartungswerts

$$\mathbb{E}[X] = \sum_{i=1}^n \mathbb{E}[X|A_i] \cdot \Pr[A_i]$$

$$\text{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

$$\text{Var}[a \cdot X + b] = a^2 \cdot \text{Var}[X]$$

$\mathbb{E}[a_1 X_1 + \dots + a_n X_n]$ $= a_1 \mathbb{E}[X_1] + \dots + a_n \mathbb{E}[X_n]$	<p>Linearität des Erwartungswerts</p>
$X_1, \dots, X_n \text{ unabhängig} \implies$ $\mathbb{E}[X_1 \cdot \dots \cdot X_n] = \mathbb{E}[X_1] \cdot \dots \cdot \mathbb{E}[X_n]$	<p>Multiplikatitivität des Erwartungswerts</p>
$X_1, \dots, X_n \text{ unabhängig} \implies$ $\text{Var}[X_1 + \dots + X_n] = \text{Var}[X_1] + \dots + \text{Var}[X_n]$	<p>Varianz einer Summe</p>

$X \geq 0 \implies$ $\Pr[X \geq t] \leq \mathbb{E}[X]/t$ für $t > 0$	Markov
$\Pr[X - \mathbb{E}[X] \geq t]$ $\leq \text{Var}[X]/t^2$ für $t > 0$	Chebyshev
siehe Satz 63	Gesetz der großen Zahlen