

Diskrete Wahrscheinlichkeitstheorie

Repetitorium

Jonas Hübötter

Outline

Zählen

Wahrscheinlichkeit

Bedingte Wahrscheinlichkeit

Diskrete Zufallsvariablen

Kontinuierliche Zufallsvariablen

Induktive Statistik

Markovketten

Plan I

Zählen

Ergebnismengen und Ereignisse
Abzählen von Mengen

Ergebnismengen und Ereignisse

Definition 1

Eine **Ergebnismenge** ist die Menge aller möglichen Elementarereignisse eines Experiments.

Ergebnismengen und Ereignisse

Definition 1

Eine **Ergebnismenge** ist die Menge aller möglichen Elementarereignisse eines Experiments.

Definition 2

Ein **Ereignis** ist eine Teilmenge der Ergebnismenge.

Ergebnismengen und Ereignisse

Definition 1

Eine **Ergebnismenge** ist die Menge aller möglichen Elementarereignisse eines Experiments.

Definition 2

Ein **Ereignis** ist eine Teilmenge der Ergebnismenge.

Naive Definition der Wahrscheinlichkeit eines Ereignisses A in der Ergebnismenge S :

Ergebnismengen und Ereignisse

Definition 1

Eine **Ergebnismenge** ist die Menge aller möglichen Elementarereignisse eines Experiments.

Definition 2

Ein **Ereignis** ist eine Teilmenge der Ergebnismenge.

Naive Definition der Wahrscheinlichkeit eines Ereignisses A in der Ergebnismenge S :

$$P(A) = \frac{\# \text{ günstige Ergebnisse}}{\# \text{ mögliche Ergebnisse}} = \frac{|A|}{|S|}$$

Ergebnismengen und Ereignisse

Definition 1

Eine **Ergebnismenge** ist die Menge aller möglichen Elementarereignisse eines Experiments.

Definition 2

Ein **Ereignis** ist eine Teilmenge der Ergebnismenge.

Naive Definition der Wahrscheinlichkeit eines Ereignisses A in der Ergebnismenge S :

$$P(A) = \frac{\# \text{ günstige Ergebnisse}}{\# \text{ mögliche Ergebnisse}} = \frac{|A|}{|S|}$$

Annahmen:

Ergebnismengen und Ereignisse

Definition 1

Eine **Ergebnismenge** ist die Menge aller möglichen Elementarereignisse eines Experiments.

Definition 2

Ein **Ereignis** ist eine Teilmenge der Ergebnismenge.

Naive Definition der Wahrscheinlichkeit eines Ereignisses A in der Ergebnismenge S :

$$P(A) = \frac{\# \text{ günstige Ergebnisse}}{\# \text{ mögliche Ergebnisse}} = \frac{|A|}{|S|}$$

Annahmen:

- alle Ergebnisse gleich wahrscheinlich

Ergebnismengen und Ereignisse

Definition 1

Eine **Ergebnismenge** ist die Menge aller möglichen Elementarereignisse eines Experiments.

Definition 2

Ein **Ereignis** ist eine Teilmenge der Ergebnismenge.

Naive Definition der Wahrscheinlichkeit eines Ereignisses A in der Ergebnismenge S :

$$P(A) = \frac{\# \text{ günstige Ergebnisse}}{\# \text{ mögliche Ergebnisse}} = \frac{|A|}{|S|}$$

Annahmen:

- alle Ergebnisse gleich wahrscheinlich
- endlicher Ergebnisraum

Abzählen von Mengen

Multiplikationsregel

Betrachte $i \in [m]$ Experimente mit n_i möglichen Ergebnissen. Dann ist die Gesamtanzahl an möglichen Ergebnissen

$$\prod_{i=1}^m n_i.$$

Kombinatorik-Tabelle

Gegeben n Objekte, wähle k Objekte.

	Reihenfolge	\neg Reihenfolge
Zurücklegen		
\neg Zurücklegen		

Kombinatorik-Tabelle

Gegeben n Objekte, wähle k Objekte.

	Reihenfolge	\neg Reihenfolge
Zurücklegen	n^k	
\neg Zurücklegen		

Kombinatorik-Tabelle

Gegeben n Objekte, wähle k Objekte.

	Reihenfolge	\neg Reihenfolge
Zurücklegen	n^k	
\neg Zurücklegen		$\binom{n}{k}$

Kombinatorik-Tabelle

Gegeben n Objekte, wähle k Objekte.

	Reihenfolge	\neg Reihenfolge
Zurücklegen	n^k	
\neg Zurücklegen	$\frac{n!}{(n-k)!}$	$\binom{n}{k}$

Kombinatorik-Tabelle

Gegeben n Objekte, wähle k Objekte.

	Reihenfolge	\neg Reihenfolge
Zurücklegen	n^k	$\binom{n+k-1}{k}$
\neg Zurücklegen	$\frac{n!}{(n-k)!}$	$\binom{n}{k}$

Plan I

Wahrscheinlichkeit

σ -Algebren

Wahrscheinlichkeitsräume

Multivariate- und Randwahrscheinlichkeiten

σ -Algebren

Definition 3

Gegeben die Ergebnismenge S . Die Menge $\mathcal{A} \subseteq \mathcal{P}(S)$ ist eine σ -Algebra über S wenn die folgenden Eigenschaften erfüllt sind:

σ -Algebren

Definition 3

Gegeben die Ergebnismenge S . Die Menge $\mathcal{A} \subseteq \mathcal{P}(S)$ ist eine σ -Algebra über S wenn die folgenden Eigenschaften erfüllt sind:

- $S \in \mathcal{A}$

σ -Algebren

Definition 3

Gegeben die Ergebnismenge S . Die Menge $\mathcal{A} \subseteq \mathcal{P}(S)$ ist eine σ -Algebra über S wenn die folgenden Eigenschaften erfüllt sind:

- $S \in \mathcal{A}$;
- falls $A \in \mathcal{A}$, dann $\bar{A} \in \mathcal{A}$

Definition 3

Gegeben die Ergebnismenge S . Die Menge $\mathcal{A} \subseteq \mathcal{P}(S)$ ist eine σ -Algebra über S wenn die folgenden Eigenschaften erfüllt sind:

- $S \in \mathcal{A}$;
- falls $A \in \mathcal{A}$, dann $\bar{A} \in \mathcal{A}$; und
- $\forall n \in \mathbb{N}. A_n \in \mathcal{A} \implies \bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$.

σ -Algebren

Definition 3

Gegeben die Ergebnismenge S . Die Menge $\mathcal{A} \subseteq \mathcal{P}(S)$ ist eine σ -Algebra über S wenn die folgenden Eigenschaften erfüllt sind:

- $S \in \mathcal{A}$;
- falls $A \in \mathcal{A}$, dann $\bar{A} \in \mathcal{A}$; und
- $\forall n \in \mathbb{N}. A_n \in \mathcal{A} \implies \bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$.

Warum benötigen wir σ -Algebren?

σ -Algebren

Definition 3

Gegeben die Ergebnismenge S . Die Menge $\mathcal{A} \subseteq \mathcal{P}(S)$ ist eine σ -Algebra über S wenn die folgenden Eigenschaften erfüllt sind:

- $S \in \mathcal{A}$;
- falls $A \in \mathcal{A}$, dann $\bar{A} \in \mathcal{A}$; und
- $\forall n \in \mathbb{N}. A_n \in \mathcal{A} \implies \bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$.

Warum benötigen wir σ -Algebren?

Um Ereignisse im Kontext eines Wahrscheinlichkeitsraumes beschreiben zu können.

Wahrscheinlichkeitsräume

Definition 4

Gegeben die Ergebnismenge S und die σ -Algebra \mathcal{A} über S .

Wahrscheinlichkeitsräume

Definition 4

Gegeben die Ergebnismenge S und die σ -Algebra \mathcal{A} über S . Die Funktion

$$P : \mathcal{A} \rightarrow [0, 1]$$

ist ein **Wahrscheinlichkeitsmaß** auf \mathcal{A} falls die **Kolmogorov Axiome** erfüllt sind:

Wahrscheinlichkeitsräume

Definition 4

Gegeben die Ergebnismenge S und die σ -Algebra \mathcal{A} über S . Die Funktion

$$P : \mathcal{A} \rightarrow [0, 1]$$

ist ein **Wahrscheinlichkeitsmaß** auf \mathcal{A} falls die **Kolmogorov Axiome** erfüllt sind:

- $P(S) = 1$;

Wahrscheinlichkeitsräume

Definition 4

Gegeben die Ergebnismenge S und die σ -Algebra \mathcal{A} über S . Die Funktion

$$P : \mathcal{A} \rightarrow [0, 1]$$

ist ein **Wahrscheinlichkeitsmaß** auf \mathcal{A} falls die **Kolmogorov Axiome** erfüllt sind:

- $P(S) = 1$;
- $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$ falls $\forall i \neq j. A_i \cap A_j = \emptyset$.

Definition 5

Für ein Ereignis $A \in \mathcal{A}$, $P(A)$ ist die **Wahrscheinlichkeit** von A .

Definition 5

Für ein Ereignis $A \in \mathcal{A}$, $P(A)$ ist die **Wahrscheinlichkeit** von A .

Definition 6

Ein **Wahrscheinlichkeitsraum** besteht aus

- einer Ergebnismenge S ;
- einer σ -Algebra \mathcal{A} über S ; und
- einem Wahrscheinlichkeitsmaß P auf \mathcal{A} .

Für einen Wahrscheinlichkeitsraum gelten die folgenden Eigenschaften:

Für einen Wahrscheinlichkeitsraum gelten die folgenden Eigenschaften:

- $P(\emptyset) = 0$

Für einen Wahrscheinlichkeitsraum gelten die folgenden Eigenschaften:

- $P(\emptyset) = 0$
- $P(S) = 1$

Für einen Wahrscheinlichkeitsraum gelten die folgenden Eigenschaften:

- $P(\emptyset) = 0$
- $P(S) = 1$
- $0 \leq P(A) \leq 1$ für alle $A \in \mathcal{A}$

Für einen Wahrscheinlichkeitsraum gelten die folgenden Eigenschaften:

- $P(\emptyset) = 0$
- $P(S) = 1$
- $0 \leq P(A) \leq 1$ für alle $A \in \mathcal{A}$
- $P(\bar{A}) = 1 - P(A)$ für alle $A \in \mathcal{A}$

Für einen Wahrscheinlichkeitsraum gelten die folgenden Eigenschaften:

- $P(\emptyset) = 0$
- $P(S) = 1$
- $0 \leq P(A) \leq 1$ für alle $A \in \mathcal{A}$
- $P(\bar{A}) = 1 - P(A)$ für alle $A \in \mathcal{A}$
- falls $A, B \in \mathcal{A}$ und $A \subseteq B$, dann $P(A) \leq P(B)$

Weiterhin gilt die **Siebformel**:

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{I \subseteq [n], I \neq \emptyset} (-1)^{|I|+1} \cdot P\left(\bigcap_{i \in I} A_i\right).$$

Weiterhin gilt die **Siebformel**:

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{I \subseteq [n], I \neq \emptyset} (-1)^{|I|+1} \cdot P\left(\bigcap_{i \in I} A_i\right).$$

Und die **Bool'sche Ungleichung**:

$$P\left(\bigcup_{i=1}^n A_i\right) \leq \sum_{i=1}^n P(A_i).$$

Multivariate- und Randwahrscheinlichkeiten

Eine **Randwahrscheinlichkeit** ist die Wahrscheinlichkeit eines einzelnen Ereignisses unabhängig von anderen Ereignissen.

Multivariate- und Randwahrscheinlichkeiten

Eine **Randwahrscheinlichkeit** ist die Wahrscheinlichkeit eines einzelnen Ereignisses unabhängig von anderen Ereignissen.

Eine **multivariate Wahrscheinlichkeit** ist die Wahrscheinlichkeit von zwei oder mehreren Ereignissen gleichzeitig aufzutreten:

$$P(A, B) = P(A \cap B).$$

Plan I

Bedingte Wahrscheinlichkeit

A-priori und a-posteriori

Unabhängigkeit

Konditionierung

A-priori und a-posteriori

Bedingte Wahrscheinlichkeit *aktualisiert* die Wahrscheinlichkeit eines Ereignisses A gegeben eine neue Information B .

A-priori und a-posteriori

Bedingte Wahrscheinlichkeit *aktualisiert* die Wahrscheinlichkeit eines Ereignisses A gegeben eine neue Information B .

$P(A)$ heißt **a-priori** und $P(A|B)$ **a-posteriori** Wahrscheinlichkeit.

A-priori und a-posteriori

Bedingte Wahrscheinlichkeit *aktualisiert* die Wahrscheinlichkeit eines Ereignisses A gegeben eine neue Information B .

$P(A)$ heißt **a-priori** und $P(A|B)$ **a-posteriori** Wahrscheinlichkeit.

$$P(A|B) = \frac{P(A, B)}{P(B)}.$$

Die a-posteriori Wahrscheinlichkeit ist die multivariate Wahrscheinlichkeit des Ereignisses A und der Information B relativ zu der Wahrscheinlichkeit der Information B .

Unabhängigkeit

Zwei Ereignisse sind **unabhängig** wenn das Auftreten des einen Ereignisses nicht die Wahrscheinlichkeit des anderen Ereignisses beeinflusst.

Unabhängigkeit

Zwei Ereignisse sind **unabhängig** wenn das Auftreten des einen Ereignisses nicht die Wahrscheinlichkeit des anderen Ereignisses beeinflusst.

Zwei Ereignisse A und B sind unabhängig

$$\iff P(A|B) = P(A) \text{ for } P(B) > 0$$

Unabhängigkeit

Zwei Ereignisse sind **unabhängig** wenn das Auftreten des einen Ereignisses nicht die Wahrscheinlichkeit des anderen Ereignisses beeinflusst.

Zwei Ereignisse A und B sind unabhängig

$$\iff P(A|B) = P(A) \text{ for } P(B) > 0$$

$$\iff P(B|A) = P(B) \text{ for } P(A) > 0$$

Unabhängigkeit

Zwei Ereignisse sind **unabhängig** wenn das Auftreten des einen Ereignisses nicht die Wahrscheinlichkeit des anderen Ereignisses beeinflusst.

Zwei Ereignisse A und B sind unabhängig

$$\iff P(A|B) = P(A) \text{ for } P(B) > 0$$

$$\iff P(B|A) = P(B) \text{ for } P(A) > 0$$

$$\iff P(A, B) = P(A)P(B).$$

Konditionierung

Einige Eigenschaften folgen direkt aus der Definition der bedingten Wahrscheinlichkeit:

Konditionierung

Einige Eigenschaften folgen direkt aus der Definition der bedingten Wahrscheinlichkeit:

- $P(A, B) = P(B)P(A|B)$

Konditionierung

Einige Eigenschaften folgen direkt aus der Definition der bedingten Wahrscheinlichkeit:

- $P(A, B) = P(B)P(A|B) = P(A)P(B|A)$,
da $A \cap B = B \cap A$

Konditionierung

Einige Eigenschaften folgen direkt aus der Definition der bedingten Wahrscheinlichkeit:

- $P(A, B) = P(B)P(A|B) = P(A)P(B|A)$,
da $A \cap B = B \cap A$
- $P(A_1, \dots, A_n) =$
 $P(A_1)P(A_2|A_1)P(A_3|A_1, A_2) \cdots P(A_n|A_1, \dots, A_{n-1})$
(Multiplikationssatz)

Konditionierung

Einige Eigenschaften folgen direkt aus der Definition der bedingten Wahrscheinlichkeit:

- $P(A, B) = P(B)P(A|B) = P(A)P(B|A)$,
da $A \cap B = B \cap A$
- $P(A_1, \dots, A_n) =$
 $P(A_1)P(A_2|A_1)P(A_3|A_1, A_2) \cdots P(A_n|A_1, \dots, A_{n-1})$
(Multiplikationssatz)
- $P(A|B) = \frac{P(B|A)P(A)}{P(B)}$ (Satz von Bayes)

Konditionierung

Einige Eigenschaften folgen direkt aus der Definition der bedingten Wahrscheinlichkeit:

- $P(A, B) = P(B)P(A|B) = P(A)P(B|A)$,
da $A \cap B = B \cap A$
- $P(A_1, \dots, A_n) =$
 $P(A_1)P(A_2|A_1)P(A_3|A_1, A_2) \cdots P(A_n|A_1, \dots, A_{n-1})$
(Multiplikationssatz)
- $P(A|B) = \frac{P(B|A)P(A)}{P(B)}$ (Satz von Bayes)
- $P(A) = P(A, B) + P(A, \bar{B}) = P(A|B)P(B) + P(A|\bar{B})P(\bar{B})$
(Gesetz der totalen Wahrscheinlichkeit)

Plan I

Diskrete Zufallsvariablen

Verteilungsfunktion

Diskrete Dichtefunktion

Unabhängigkeit

Bernoulli Verteilung

Erwartungswert

Indikatorvariablen

Binomialverteilung

Varianz

Geometrische Verteilung

Poisson Verteilung

Wahrscheinlichkeitserzeugende Funktionen

Momenterzeugende Funktionen

Multivariate Dichten

Bedingte Dichten

Plan II

Faltungen

Weitere Verteilungen

Ungleichungen

Diskrete Zufallsvariablen

Definition 7

Eine **Zufallsvariable** X ist eine Funktion

$$X : \mathcal{S} \rightarrow \mathbb{R}.$$

Diskrete Zufallsvariablen

Definition 7

Eine **Zufallsvariable** X ist eine Funktion

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

Eine Zufallsvariable heißt **diskret** wenn ihr Urbild S endlich oder abzählbar unendlich ist.

Diskrete Zufallsvariablen

Definition 7

Eine **Zufallsvariable** X ist eine Funktion

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

Eine Zufallsvariable heißt **diskret** wenn ihr Urbild S endlich oder abzählbar unendlich ist.

Der Wertebereich einer diskreten Zufallsvariable

$$X(S) = \{x \in \mathbb{R}. \exists A \in S. X(A) = x\}$$

ist ebenfalls abzählbar.

Verteilungsfunktion

$X \leq x$ ist ein Ereignis.

Verteilungsfunktion

$X \leq x$ ist ein Ereignis.

Definition 8

Die **Verteilungsfunktion** einer Zufallsvariable X ist definiert als $F_X(x) = P(X \leq x) \in [0, 1]$.

Verteilungsfunktion

$X \leq x$ ist ein Ereignis.

Definition 8

Die **Verteilungsfunktion** einer Zufallsvariable X ist definiert als $F_X(x) = P(X \leq x) \in [0, 1]$.

Eigenschaften von Verteilungsfunktionen:

Verteilungsfunktion

$X \leq x$ ist ein Ereignis.

Definition 8

Die **Verteilungsfunktion** einer Zufallsvariable X ist definiert als $F_X(x) = P(X \leq x) \in [0, 1]$.

Eigenschaften von Verteilungsfunktionen:

- monoton wachsend

Verteilungsfunktion

$X \leq x$ ist ein Ereignis.

Definition 8

Die **Verteilungsfunktion** einer Zufallsvariable X ist definiert als $F_X(x) = P(X \leq x) \in [0, 1]$.

Eigenschaften von Verteilungsfunktionen:

- monoton wachsend
- rechtsseitig stetig

Verteilungsfunktion

$X \leq x$ ist ein Ereignis.

Definition 8

Die **Verteilungsfunktion** einer Zufallsvariable X ist definiert als $F_X(x) = P(X \leq x) \in [0, 1]$.

Eigenschaften von Verteilungsfunktionen:

- monoton wachsend
- rechtsseitig stetig
- $F_X(x) \xrightarrow{x \rightarrow -\infty} 0$

Verteilungsfunktion

$X \leq x$ ist ein Ereignis.

Definition 8

Die **Verteilungsfunktion** einer Zufallsvariable X ist definiert als $F_X(x) = P(X \leq x) \in [0, 1]$.

Eigenschaften von Verteilungsfunktionen:

- monoton wachsend
- rechtsseitig stetig
- $F_X(x) \xrightarrow{x \rightarrow -\infty} 0$
- $F_X(x) \xrightarrow{x \rightarrow \infty} 1$

Verteilungsfunktion

$X \leq x$ ist ein Ereignis.

Definition 8

Die **Verteilungsfunktion** einer Zufallsvariable X ist definiert als $F_X(x) = P(X \leq x) \in [0, 1]$.

Eigenschaften von Verteilungsfunktionen:

- monoton wachsend
- rechtsseitig stetig
- $F_X(x) \xrightarrow{x \rightarrow -\infty} 0$
- $F_X(x) \xrightarrow{x \rightarrow \infty} 1$

Daher, $P(a < X \leq b) = F_X(b) - F_X(a)$.

Diskrete Dichtefunktion

Definition 9

Die **diskrete Dichtefunktion** einer diskreten Zufallsvariable X ist definiert als $f_X(x) = P(X = x) \in [0, 1]$ wobei

$$\sum_{x \in X(S)} f_X(x) = 1.$$

Die Verteilungsfunktion von X kann von der Dichtefunktion von X erhalten werden indem über die Dichtefunktion summiert wird

$$F_X(x) = \sum_{x' \leq x} f_X(x').$$

Die Verteilungsfunktion von X kann von der Dichtefunktion von X erhalten werden indem über die Dichtefunktion summiert wird

$$F_X(x) = \sum_{x' \leq x} f_X(x').$$

Die Dichtefunktion von X kann von der Verteilungsfunktion von X erhalten werden indem die *Sprünge* in der Verteilungsfunktion identifiziert werden

$$f_X(x) = F_X(x) - F_X(\text{prev}(x)).$$

Unabhängigkeit

Zwei Zufallsvariablen sind **unabhängig** wenn das Wissen des Wertes einer Zufallsvariable keine Auswirkungen auf die Verteilung der anderen Zufallsvariable hat.

Unabhängigkeit

Zwei Zufallsvariablen sind **unabhängig** wenn das Wissen des Wertes einer Zufallsvariable keine Auswirkungen auf die Verteilung der anderen Zufallsvariable hat.

Zwei diskrete Zufallsvariablen X und Y sind unabhängig
 \iff die Ereignisse $X = x$ und $Y = y$ sind unabhängig

Unabhängigkeit

Zwei Zufallsvariablen sind **unabhängig** wenn das Wissen des Wertes einer Zufallsvariable keine Auswirkungen auf die Verteilung der anderen Zufallsvariable hat.

Zwei diskrete Zufallsvariablen X und Y sind unabhängig

\iff die Ereignisse $X = x$ und $Y = y$ sind unabhängig

\iff die Ereignisse $X \leq x$ und $Y \leq y$ sind unabhängig.

Bernoulli Verteilung

Definition 10 ($X \sim \text{Bern}(p)$)

Eine diskrete Zufallsvariable X ist **Bernoulli**-verteilt mit Parameter p falls $X(S) = \{0, 1\}$ und $P(X = 1) = p$.

Bernoulli Verteilung

Definition 10 ($X \sim \text{Bern}(p)$)

Eine diskrete Zufallsvariable X ist **Bernoulli**-verteilt mit Parameter p falls $X(S) = \{0, 1\}$ und $P(X = 1) = p$.

Übersicht

- $E(X) = p$
- $\text{Var}(X) = p(1 - p)$
- $G_X(s) = 1 - p + ps$
- $M_X(s) = 1 - p + pe^s$

Erwartungswert

Definition 11

Der **Erwartungswert** $E(X)$ einer Zufallsvariable X ist das arithmetische Mittel einer großen Anzahl an Realisierungen von X .

Erwartungswert

Definition 11

Der **Erwartungswert** $E(X)$ einer Zufallsvariable X ist das arithmetische Mittel einer großen Anzahl an Realisierungen von X .

$$E(X) = \sum_{x \in X(S)} x \cdot P(X = x)$$

Erwartungswert

Definition 11

Der **Erwartungswert** $E(X)$ einer Zufallsvariable X ist das arithmetische Mittel einer großen Anzahl an Realisierungen von X .

$$\begin{aligned} E(X) &= \sum_{x \in X(S)} x \cdot P(X = x) \\ &= \sum_{A \in S} X(A) \cdot P(A). \end{aligned}$$

Erwartungswert

Definition 11

Der **Erwartungswert** $E(X)$ einer Zufallsvariable X ist das arithmetische Mittel einer großen Anzahl an Realisierungen von X .

$$\begin{aligned} E(X) &= \sum_{x \in X(S)} x \cdot P(X = x) \\ &= \sum_{A \in \mathcal{S}} X(A) \cdot P(A). \end{aligned}$$

Für unendlich große Wahrscheinlichkeitsräume ist **absolute Konvergenz** von $E(X)$ eine notwendige Bedingung für die Existenz von $E(X)$.

Eigenschaften des Erwartungswerts:

Eigenschaften des Erwartungswerts:

- falls $\forall A \in S. X(A) \leq Y(A)$, dann $E(X) \leq E(Y)$ (Monotonie)

Eigenschaften des Erwartungswerts:

- falls $\forall A \in S. X(A) \leq Y(A)$, dann $E(X) \leq E(Y)$ (Monotonie)
- $E(a \cdot X + b) = a \cdot E(X) + b$, $E(X + Y) = E(X) + E(Y)$
(Linearität)

Eigenschaften des Erwartungswerts:

- falls $\forall A \in S. X(A) \leq Y(A)$, dann $E(X) \leq E(Y)$ (Monotonie)
- $E(a \cdot X + b) = a \cdot E(X) + b$, $E(X + Y) = E(X) + E(Y)$
(Linearität)
- $E(\prod_{i=1}^n X_i) = \prod_{i=1}^n E(X_i)$ falls X_1, \dots, X_n unabhängig
(Multiplikativität).

Definition 12

$E(X^i)$ heißt *i -tes Moment* der Zufallsvariable X und
 $E((X - E(X))^i)$ heißt *i -tes zentrales Moment* von X .

Das sogenannte **law of the unconscious statistician (LOTUS)** kann verwendet werden, um den Erwartungswert von transformierten Zufallsvariablen zu finden.

$$E(g(X)) = \sum_{x \in X(S)} g(x) \cdot P(X = x).$$

Indikatorvariablen

Definition 13

Gegeben ein Ereignis A , die Zufallsvariable $I_A \sim \text{Bern}(P(A))$ ist die **Indikatorvariable** des Ereignisses A .

Indikatorvariablen

Definition 13

Gegeben ein Ereignis A , die Zufallsvariable $I_A \sim \text{Bern}(P(A))$ ist die **Indikatorvariable** des Ereignisses A .

Eigenschaften von Indikatorvariablen:

Indikatorvariablen

Definition 13

Gegeben ein Ereignis A , die Zufallsvariable $I_A \sim \text{Bern}(P(A))$ ist die **Indikatorvariable** des Ereignisses A .

Eigenschaften von Indikatorvariablen:

- $E(I_A) = P(A)$

Indikatorvariablen

Definition 13

Gegeben ein Ereignis A , die Zufallsvariable $I_A \sim \text{Bern}(P(A))$ ist die **Indikatorvariable** des Ereignisses A .

Eigenschaften von Indikatorvariablen:

- $E(I_A) = P(A)$
- $E(I_{A_1} \cdots I_{A_n}) = P(A_1 \cap \cdots \cap A_n)$.

Binomialverteilung

Definition 14 ($X \sim \text{Bin}(n, p)$)

Eine diskrete Zufallsvariable X ist **binomial**-verteilt mit Parametern n und p falls X die #Erfolge in n unabhängigen $\text{Bern}(p)$ Versuchen modelliert.

Binomialverteilung

Definition 14 ($X \sim \text{Bin}(n, p)$)

Eine diskrete Zufallsvariable X ist **binomial**-verteilt mit Parametern n und p falls X die #Erfolge in n unabhängigen $\text{Bern}(p)$ Versuchen modelliert.

$$f_X(k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

Binomialverteilung

Definition 14 ($X \sim \text{Bin}(n, p)$)

Eine diskrete Zufallsvariable X ist **binomial**-verteilt mit Parametern n und p falls X die #Erfolge in n unabhängigen $\text{Bern}(p)$ Versuchen modelliert.

$$f_X(k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

Übersicht

- $E(X) = np$

Binomialverteilung

Definition 14 ($X \sim \text{Bin}(n, p)$)

Eine diskrete Zufallsvariable X ist **binomial**-verteilt mit Parametern n und p falls X die #Erfolge in n unabhängigen $\text{Bern}(p)$ Versuchen modelliert.

$$f_X(k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

Übersicht

- $E(X) = np$
- $\text{Var}(X) = np(1-p)$
- $G_X(s) = (1-p+ps)^n$
- $M_X(s) = (1-p+pe^s)^n$

Varianz

Definition 15

Die **Varianz** $Var(X)$ einer Zufallsvariable X ist ein Maß der absoluten Abweichung einer Zufallsvariablen von ihrem Erwartungswert.

$$Var(X) = E((X - E(X))^2)$$

Varianz

Definition 15

Die **Varianz** $\text{Var}(X)$ einer Zufallsvariable X ist ein Maß der absoluten Abweichung einer Zufallsvariablen von ihrem Erwartungswert.

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

Varianz

Definition 15

Die **Varianz** $\text{Var}(X)$ einer Zufallsvariable X ist ein Maß der absoluten Abweichung einer Zufallsvariablen von ihrem Erwartungswert.

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

$SD(X) = \sqrt{\text{Var}(X)}$ heißt **Standardabweichung** von X .

Varianz

Definition 15

Die **Varianz** $Var(X)$ einer Zufallsvariable X ist ein Maß der absoluten Abweichung einer Zufallsvariablen von ihrem Erwartungswert.

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

$SD(X) = \sqrt{Var(X)}$ heißt **Standardabweichung** von X .

Eigenschaften der Varianz:

Varianz

Definition 15

Die **Varianz** $\text{Var}(X)$ einer Zufallsvariable X ist ein Maß der absoluten Abweichung einer Zufallsvariablen von ihrem Erwartungswert.

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

$SD(X) = \sqrt{\text{Var}(X)}$ heißt **Standardabweichung** von X .

Eigenschaften der Varianz:

- $\text{Var}(a \cdot X + b) = a^2 \cdot \text{Var}(X)$

Varianz

Definition 15

Die **Varianz** $\text{Var}(X)$ einer Zufallsvariable X ist ein Maß der absoluten Abweichung einer Zufallsvariablen von ihrem Erwartungswert.

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

$SD(X) = \sqrt{\text{Var}(X)}$ heißt **Standardabweichung** von X .

Eigenschaften der Varianz:

- $\text{Var}(a \cdot X + b) = a^2 \cdot \text{Var}(X)$
- $\text{Var}(\sum_{i=1}^n X_i) = \sum_{i=1}^n \text{Var}(X_i)$ falls X_1, \dots, X_n unabhängig.

Geometrische Verteilung

Definition 16 ($X \sim \text{Geom}(p)$)

Eine diskrete Zufallsvariable X ist **geometrisch** verteilt mit Parameter p falls X die #Versuche, die zu einem Erfolg führen, in unabhängigen $\text{Bern}(p)$ Versuchen modelliert.

Geometrische Verteilung

Definition 16 ($X \sim \text{Geom}(p)$)

Eine diskrete Zufallsvariable X ist **geometrisch** verteilt mit Parameter p falls X die #Versuche, die zu einem Erfolg führen, in unabhängigen $\text{Bern}(p)$ Versuchen modelliert.

$$f_X(k) = p(1 - p)^{k-1}, k \in \mathbb{N}.$$

Geometrische Verteilung

Definition 16 ($X \sim \text{Geom}(p)$)

Eine diskrete Zufallsvariable X ist **geometrisch** verteilt mit Parameter p falls X die #Versuche, die zu einem Erfolg führen, in unabhängigen $\text{Bern}(p)$ Versuchen modelliert.

$$f_X(k) = p(1-p)^{k-1}, k \in \mathbb{N}.$$

$$F_X(k) = 1 - (1-p)^{\lfloor k \rfloor}.$$

Geometrische Verteilung

Definition 16 ($X \sim \text{Geom}(p)$)

Eine diskrete Zufallsvariable X ist **geometrisch** verteilt mit Parameter p falls X die #Versuche, die zu einem Erfolg führen, in unabhängigen $\text{Bern}(p)$ Versuchen modelliert.

$$f_X(k) = p(1-p)^{k-1}, k \in \mathbb{N}.$$

$$F_X(k) = 1 - (1-p)^{\lfloor k \rfloor}.$$

Übersicht

- $E(X) = \frac{1}{p}$

Geometrische Verteilung

Definition 16 ($X \sim \text{Geom}(p)$)

Eine diskrete Zufallsvariable X ist **geometrisch** verteilt mit Parameter p falls X die #Versuche, die zu einem Erfolg führen, in unabhängigen $\text{Bern}(p)$ Versuchen modelliert.

$$f_X(k) = p(1-p)^{k-1}, k \in \mathbb{N}.$$

$$F_X(k) = 1 - (1-p)^{\lfloor k \rfloor}.$$

Übersicht

- $E(X) = \frac{1}{p}$
- $\text{Var}(X) = \frac{1-p}{p^2}$

Geometrische Verteilung

Definition 16 ($X \sim \text{Geom}(p)$)

Eine diskrete Zufallsvariable X ist **geometrisch** verteilt mit Parameter p falls X die #Versuche, die zu einem Erfolg führen, in unabhängigen $\text{Bern}(p)$ Versuchen modelliert.

$$f_X(k) = p(1-p)^{k-1}, k \in \mathbb{N}.$$

$$F_X(k) = 1 - (1-p)^{\lfloor k \rfloor}.$$

Übersicht

- $E(X) = \frac{1}{p}$
- $\text{Var}(X) = \frac{1-p}{p^2}$
- $G_X(s) = \frac{ps}{1-(1-p)s}$

Gedächtnislosigkeit

Durchführen von x Versuchen, von denen keiner zum Erfolg führt, verändert nicht die Wahrscheinlichkeit dafür, dass die nächsten y Versuche einen Erfolg beinhalten.

Gedächtnislosigkeit

Durchführen von x Versuchen, von denen keiner zum Erfolg führt, verändert nicht die Wahrscheinlichkeit dafür, dass die nächsten y Versuche einen Erfolg beinhalten.

Diese Eigenschaft kann wie folgt formalisiert werden:

$$P(X > y + x | X > x) = P(X > y).$$

Gedächtnislosigkeit

Durchführen von x Versuchen, von denen keiner zum Erfolg führt, verändert nicht die Wahrscheinlichkeit dafür, dass die nächsten y Versuche einen Erfolg beinhalten.

Diese Eigenschaft kann wie folgt formalisiert werden:

$$P(X > y + x | X > x) = P(X > y).$$

Die geometrische Verteilung ist die **einzig**e gedächtnislose diskrete Verteilung.

Poisson Verteilung

Definition 17 ($X \sim Po(\lambda)$)

Eine diskrete Zufallsvariable X ist **Poisson**-verteilt mit Parameter λ falls X die #Ereignisse in einem festen Intervall mit Rate λ modelliert, wobei die Ereignisse unabhängig von der Zeit seit dem letzten Ereignis auftreten.

Poisson Verteilung

Definition 17 ($X \sim Po(\lambda)$)

Eine diskrete Zufallsvariable X ist **Poisson**-verteilt mit Parameter λ falls X die #Ereignisse in einem festen Intervall mit Rate λ modelliert, wobei die Ereignisse unabhängig von der Zeit seit dem letzten Ereignis auftreten.

$$f_X(k) = \frac{e^{-\lambda} \cdot \lambda^k}{k!}, k \in \mathbb{N}_0.$$

Poisson Verteilung

Definition 17 ($X \sim Po(\lambda)$)

Eine diskrete Zufallsvariable X ist **Poisson**-verteilt mit Parameter λ falls X die #Ereignisse in einem festen Intervall mit Rate λ modelliert, wobei die Ereignisse unabhängig von der Zeit seit dem letzten Ereignis auftreten.

$$f_X(k) = \frac{e^{-\lambda} \cdot \lambda^k}{k!}, k \in \mathbb{N}_0.$$

$$F_X(k) = e^{-\lambda} \cdot \sum_{i=0}^{\lfloor k \rfloor} \frac{\lambda^i}{i!}.$$

Poisson Verteilung

Definition 17 ($X \sim Po(\lambda)$)

Eine diskrete Zufallsvariable X ist **Poisson**-verteilt mit Parameter λ falls X die #Ereignisse in einem festen Intervall mit Rate λ modelliert, wobei die Ereignisse unabhängig von der Zeit seit dem letzten Ereignis auftreten.

$$f_X(k) = \frac{e^{-\lambda} \cdot \lambda^k}{k!}, k \in \mathbb{N}_0.$$

$$F_X(k) = e^{-\lambda} \cdot \sum_{i=0}^{[k]} \frac{\lambda^i}{i!}.$$

Übersicht

- $E(X) = \lambda$

Poisson Verteilung

Definition 17 ($X \sim Po(\lambda)$)

Eine diskrete Zufallsvariable X ist **Poisson**-verteilt mit Parameter λ falls X die #Ereignisse in einem festen Intervall mit Rate λ modelliert, wobei die Ereignisse unabhängig von der Zeit seit dem letzten Ereignis auftreten.

$$f_X(k) = \frac{e^{-\lambda} \cdot \lambda^k}{k!}, k \in \mathbb{N}_0.$$

$$F_X(k) = e^{-\lambda} \cdot \sum_{i=0}^{[k]} \frac{\lambda^i}{i!}.$$

Übersicht

- $E(X) = \lambda$
- $Var(X) = \lambda$

Poisson Verteilung

Definition 17 ($X \sim Po(\lambda)$)

Eine diskrete Zufallsvariable X ist **Poisson**-verteilt mit Parameter λ falls X die #Ereignisse in einem festen Intervall mit Rate λ modelliert, wobei die Ereignisse unabhängig von der Zeit seit dem letzten Ereignis auftreten.

$$f_X(k) = \frac{e^{-\lambda} \cdot \lambda^k}{k!}, k \in \mathbb{N}_0.$$

$$F_X(k) = e^{-\lambda} \cdot \sum_{i=0}^{[k]} \frac{\lambda^i}{i!}.$$

Übersicht

- $E(X) = \lambda$
- $Var(X) = \lambda$
- $G_X(s) = \exp(\lambda(s - 1))$
- $M_X(s) = \exp(\lambda(e^s - 1))$

Poisson-Approximation der Binomialverteilung

Sei $X \sim \text{Bin}(n, \lambda/n)$.

Poisson-Approximation der Binomialverteilung

Sei $X \sim \text{Bin}(n, \lambda/n)$.

Dann konvergiert die Verteilung von X zu $Po(\lambda)$ mit $n \rightarrow \infty$

Poisson-Approximation der Binomialverteilung

Sei $X \sim \text{Bin}(n, \lambda/n)$.

Dann konvergiert die Verteilung von X zu $Po(\lambda)$ mit $n \rightarrow \infty$
(d.h. für kleine λ/n).

Wahrscheinlichkeitserzeugende Funktionen

Definition 18

Gegeben eine diskrete Zufallsvariable X mit $X(S) \subseteq \mathbb{N}_0$ ist die **wahrscheinlichkeitserzeugende Funktion** definiert als

$$G_X(s) = \sum_{x \in X(S)} s^x \cdot P(X = x)$$

Wahrscheinlichkeitserzeugende Funktionen

Definition 18

Gegeben eine diskrete Zufallsvariable X mit $X(S) \subseteq \mathbb{N}_0$ ist die **wahrscheinlichkeitserzeugende Funktion** definiert als

$$\begin{aligned} G_X(s) &= \sum_{x \in X(S)} s^x \cdot P(X = x) \\ &= E(s^X). \end{aligned}$$

Wahrscheinlichkeitserzeugende Funktionen

Definition 18

Gegeben eine diskrete Zufallsvariable X mit $X(S) \subseteq \mathbb{N}_0$ ist die **wahrscheinlichkeitserzeugende Funktion** definiert als

$$\begin{aligned} G_X(s) &= \sum_{x \in X(S)} s^x \cdot P(X = x) \\ &= E(s^X). \end{aligned}$$

Die wahrscheinlichkeitserzeugende Funktion einer Zufallsvariable X erzeugt die Dichtefunktion von X :

$$P(X = i) = \frac{G_X^{(i)}(0)}{i!}.$$

Eigenschaften von wahrscheinlichkeitserzeugenden Funktionen:

Eigenschaften von wahrscheinlichkeitserzeugenden Funktionen:

- $E(X) = G'_X(1)$

Eigenschaften von wahrscheinlichkeitserzeugenden Funktionen:

- $E(X) = G'_X(1)$
- $Var(X) = G''_X(1) + G'_X(1) - (G'_X(1))^2$

Eigenschaften von wahrscheinlichkeitserzeugenden Funktionen:

- $E(X) = G'_X(1)$
- $Var(X) = G''_X(1) + G'_X(1) - (G'_X(1))^2$
- $G_{X+t}(s) = s^t \cdot G_X(s), t \in \mathbb{N}_0$

Eigenschaften von wahrscheinlichkeitserzeugenden Funktionen:

- $E(X) = G'_X(1)$
- $Var(X) = G''_X(1) + G'_X(1) - (G'_X(1))^2$
- $G_{X+t}(s) = s^t \cdot G_X(s), t \in \mathbb{N}_0$
- $G_{X+Y}(s) = G_X(s) \cdot G_Y(s)$ falls X, Y unabhängig

Eigenschaften von wahrscheinlichkeitserzeugenden Funktionen:

- $E(X) = G'_X(1)$
- $Var(X) = G''_X(1) + G'_X(1) - (G'_X(1))^2$
- $G_{X+t}(s) = s^t \cdot G_X(s), t \in \mathbb{N}_0$
- $G_{X+Y}(s) = G_X(s) \cdot G_Y(s)$ falls X, Y unabhängig
- $G_Z(s) = G_N(G_X(s))$ für $Z = X_1 + \dots + X_N, X_i$ unabhängig und identisch verteilt mit wahrscheinlichkeitserzeugender Funktion G_X und N unabhängig.

Momenterzeugende Funktionen

Definition 19

Gegeben eine Zufallsvariable X ist die **momenterzeugende Funktion** definiert als

$$M_X(s) = \sum_{x \in X(S)} e^{sx} \cdot P(X = x)$$

Momenterzeugende Funktionen

Definition 19

Gegeben eine Zufallsvariable X ist die **momenterzeugende Funktion** definiert als

$$\begin{aligned}M_X(s) &= \sum_{x \in X(S)} e^{sx} \cdot P(X = x) \\ &= E(e^{sX})\end{aligned}$$

Momenterzeugende Funktionen

Definition 19

Gegeben eine Zufallsvariable X ist die **momenterzeugende Funktion** definiert als

$$\begin{aligned}M_X(s) &= \sum_{x \in X(S)} e^{sx} \cdot P(X = x) \\&= E(e^{sX}) \\&= \sum_{i=0}^{\infty} \frac{E(X^i)}{i!} \cdot s^i.\end{aligned}$$

Momenterzeugende Funktionen

Definition 19

Gegeben eine Zufallsvariable X ist die **momenterzeugende Funktion** definiert als

$$\begin{aligned}M_X(s) &= \sum_{x \in X(S)} e^{sx} \cdot P(X = x) \\ &= E(e^{sX}) \\ &= \sum_{i=0}^{\infty} \frac{E(X^i)}{i!} \cdot s^i.\end{aligned}$$

Die momenterzeugende Funktion einer Zufallsvariable X erzeugt das i -te Moment von X :

$$E(X^i) = M_X^{(i)}(0).$$

Eigenschaften von momenterzeugenden Funktionen:

Eigenschaften von momenterzeugenden Funktionen:

- $M_X(s) = G_X(e^s)$ if $X(S) \subseteq \mathbb{N}_0$

Eigenschaften von momenterzeugenden Funktionen:

- $M_X(s) = G_X(e^s)$ if $X(S) \subseteq \mathbb{N}_0$
- $M_{X+Y}(s) = M_X(s) \cdot M_Y(s)$ falls X, Y unabhängig.

Multivariate Dichten

Definition 20

Eine **multivariate Dichte** ist die Dichte von zwei oder mehr Zufallsvariablen.

$$f_{X,Y}(x,y) = P(X = x, Y = y).$$

Multivariate Dichten

Definition 20

Eine **multivariate Dichte** ist die Dichte von zwei oder mehr Zufallsvariablen.

$$f_{X,Y}(x,y) = P(X = x, Y = y).$$

Die **Randdichte** einer Zufallsvariablen kann aus einer multivariaten Dichte gewonnen werden indem über alle anderen Zufallsvariablen summiert wird:

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

Bedingte Dichten

Definition 21

Gegeben eine multivariate Dichte von zwei Zufallsvariablen X und Y ist die **bedingte Dichte** von X gegeben Y die Dichte von X wenn der konkrete Wert von Y bekannt ist.

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{f_{Y|X}(y|x) \cdot f_X(x)}{f_Y(y)}.$$

Bedingte Dichten

Definition 21

Gegeben eine multivariate Dichte von zwei Zufallsvariablen X und Y ist die **bedingte Dichte** von X gegeben Y die Dichte von X wenn der konkrete Wert von Y bekannt ist.

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{f_{Y|X}(y|x) \cdot f_X(x)}{f_Y(y)}.$$

Der **bedingte Erwartungswert** der Zufallsvariablen $X|Y = y$ ist der Erwartungswert der Dichte $f_{X|Y=y}$:

$$E(X|Y = y) = \sum_{x \in X(S)} x \cdot f_{X|Y}(x|y).$$

Faltungen

Definition 22

Seien die Zufallsvariablen X und Y unabhängig und $Z = X + Y$.
Dann gilt

$$f_Z(z) = \sum_{x \in X(S)} f_X(x) \cdot f_Y(z - x).$$

Faltungen

Definition 22

Seien die Zufallsvariablen X und Y unabhängig und $Z = X + Y$.
Dann gilt

$$f_Z(z) = \sum_{x \in X(S)} f_X(x) \cdot f_Y(z - x).$$

Die Herleitung der Verteilung einer Summe von Zufallsvariablen gegeben deren Randverteilungen bezeichnet man auch als **Faltung** oder **Konvolution**.

Weitere Verteilungen

Definition 23 ($X \sim \text{HypGeom}(r, a, b)$)

Eine diskrete Zufallsvariable X ist **hypergeometrisch** verteilt mit Parametern r , a und b falls X die # von gezogenen Objekten, die eine spezifische Eigenschaft haben, in r Ziehungen ohne Zurücklegen aus $a + b$ Objekten modelliert wobei b Objekte die Eigenschaft aufweisen.

Weitere Verteilungen

Definition 23 ($X \sim \text{HypGeom}(r, a, b)$)

Eine diskrete Zufallsvariable X ist **hypergeometrisch** verteilt mit Parametern r , a und b falls X die # von gezogenen Objekten, die eine spezifische Eigenschaft haben, in r Ziehungen ohne Zurücklegen aus $a + b$ Objekten modelliert wobei b Objekte die Eigenschaft aufweisen.

$$f_X(x) = \frac{\binom{b}{x} \binom{a}{r-x}}{\binom{a+b}{r}}.$$

Weitere Verteilungen

Definition 23 ($X \sim \text{HypGeom}(r, a, b)$)

Eine diskrete Zufallsvariable X ist **hypergeometrisch** verteilt mit Parametern r , a und b falls X die $\#$ von gezogenen Objekten, die eine spezifische Eigenschaft haben, in r Ziehungen ohne Zurücklegen aus $a + b$ Objekten modelliert wobei b Objekte die Eigenschaft aufweisen.

$$f_X(x) = \frac{\binom{b}{x} \binom{a}{r-x}}{\binom{a+b}{r}}.$$

Übersicht

- $E(X) = r \cdot \frac{b}{a+b}$

Definition 24 ($Z \sim \text{NegBin}(n, p)$)

Eine diskrete Zufallsvariable Z ist **negativ binomialverteilt** mit Parametern n und p falls Z die # von unabhängigen $\text{Bern}(p)$ Versuchen bevor dem n -ten Erfolg modelliert.

Definition 24 ($Z \sim \text{NegBin}(n, p)$)

Eine diskrete Zufallsvariable Z ist **negativ binomialverteilt** mit Parametern n und p falls Z die # von unabhängigen $\text{Bern}(p)$ Versuchen bevor dem n -ten Erfolg modelliert.

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

Definition 24 ($Z \sim \text{NegBin}(n, p)$)

Eine diskrete Zufallsvariable Z ist **negativ binomialverteilt** mit Parametern n und p falls Z die # von unabhängigen $\text{Bern}(p)$ Versuchen bevor dem n -ten Erfolg modelliert.

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

Example 25

Seien $X_1, \dots, X_n \sim \text{Geom}(p)$ unabhängig und gleichverteilt.
Dann gilt $Z = X_1 + \dots + X_n \sim \text{NegBin}(n, p)$.

Ungleichungen

Ungleichungen vs Approximationen

Approximationen erlauben uns komplexere Probleme zu modellieren, doch ist oft nicht klar wie genau die Approximation ist.

Ungleichungen

Ungleichungen vs Approximationen

Approximationen erlauben uns komplexere Probleme zu modellieren, doch ist oft nicht klar wie genau die Approximation ist. *Ungleichungen* erlauben uns definitive Aussagen (Schranken) bezüglich der Wahrscheinlichkeit von Ereignissen zu treffen.

Definition 26 (Markov)

Gegeben eine Zufallsvariable $X \geq 0$ und $t > 0$

$$P(X \geq t) \leq \frac{E(X)}{t}.$$

Definition 26 (Markov)

Gegeben eine Zufallsvariable $X \geq 0$ und $t > 0$

$$P(X \geq t) \leq \frac{E(X)}{t}.$$

Definition 27 (Chebyshev)

Gegeben eine Zufallsvariable X und $t > 0$

$$P(|X - E(X)| \geq t) \leq \frac{\text{Var}(X)}{t^2}.$$

Definition 28 (Chernoff)

Seien X_1, \dots, X_n unabhängige, Bernoulli-verteilte Zufallsvariablen mit $X_i \sim \text{Bern}(p_i)$. Dann gelten die folgenden Ungleichungen für $X = \sum_{i=1}^n X_i$ und $\mu = E(X) = \sum_{i=1}^n p_i$.

Definition 28 (Chernoff)

Seien X_1, \dots, X_n unabhängige, Bernoulli-verteilte Zufallsvariablen mit $X_i \sim \text{Bern}(p_i)$. Dann gelten die folgenden Ungleichungen für $X = \sum_{i=1}^n X_i$ und $\mu = E(X) = \sum_{i=1}^n p_i$.

- $P(X \geq (1 + \delta)\mu) \leq \left(\frac{e^\delta}{(1+\delta)^{1+\delta}}\right)^\mu$ für alle $\delta > 0$;
- $P(X \leq (1 - \delta)\mu) \leq \left(\frac{e^{-\delta}}{(1-\delta)^{1-\delta}}\right)^\mu$ für alle $0 < \delta < 1$

Definition 28 (Chernoff)

Seien X_1, \dots, X_n unabhängige, Bernoulli-verteilte Zufallsvariablen mit $X_i \sim \text{Bern}(p_i)$. Dann gelten die folgenden Ungleichungen für $X = \sum_{i=1}^n X_i$ und $\mu = E(X) = \sum_{i=1}^n p_i$.

- $P(X \geq (1 + \delta)\mu) \leq \left(\frac{e^\delta}{(1+\delta)^{1+\delta}}\right)^\mu$ für alle $\delta > 0$;
- $P(X \leq (1 - \delta)\mu) \leq \left(\frac{e^{-\delta}}{(1-\delta)^{1-\delta}}\right)^\mu$ für alle $0 < \delta < 1$;
- $P(X \geq (1 + \delta)\mu) \leq e^{-\mu\delta^2/3}$ für alle $0 < \delta \leq 1$;
- $P(X \leq (1 - \delta)\mu) \leq e^{-\mu\delta^2/2}$ für alle $0 < \delta \leq 1$;
- $P(|X - \mu| \geq \delta\mu) \leq 2e^{-\mu\delta^2/3}$ für alle $0 < \delta \leq 1$;
- $P(X \geq (1 + \delta)\mu) \leq \left(\frac{e}{1+\delta}\right)^{(1+\delta)\mu}$; and
- $P(X \geq t) \leq 2^{-t}$ für alle $t \geq 2e\mu$.

Plan I

Kontinuierliche Zufallsvariablen

Messbarkeitstheorie

Kontinuierliche Wahrscheinlichkeitsräume

Gleichverteilung

Normalverteilung

γ -Quantil

Exponentialverteilung

Multivariate Verteilungen

Weitere Verteilungen

Kontinuierliche Zufallsvariablen

Definition 29

Eine **kontinuierliche Zufallsvariable** X ist eine Funktion

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

wobei $X(S)$ überabzählbar ist.

Kontinuierliche Zufallsvariablen

Definition 29

Eine **kontinuierliche Zufallsvariable** X ist eine Funktion

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

wobei $X(S)$ überabzählbar ist.

Die Verteilung von X ist definiert durch die **kontinuierliche Dichtefunktion** $f_X : \mathbb{R} \rightarrow \mathbb{R}_0^+$ mit der Eigenschaft

$$\int_{-\infty}^{+\infty} f_X(x) dx = 1.$$

Messbarkeitstheorie

Definition 30

- Eine **Borel'sche Menge** für \mathbb{R} ist eine Teilmenge $A \subseteq \mathbb{R}$, die als durch abzählbare Vereinigungen und Schnitte von Intervallen (offen, halboffen, oder geschlossen) auf \mathbb{R} dargestellt werden kann.

Messbarkeitstheorie

Definition 30

- Eine **Borel'sche Menge** für \mathbb{R} ist eine Teilmenge $A \subseteq \mathbb{R}$, die als durch abzählbare Vereinigungen und Schnitte von Intervallen (offen, halboffen, oder geschlossen) auf \mathbb{R} dargestellt werden kann.
- Eine Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ ist **(Borel-)messbar**, falls das Urbild jeder Borel'schen Menge ebenfalls eine Borel'sche Menge ist.

Definition 30

- Eine **Borel'sche Menge** für \mathbb{R} ist eine Teilmenge $A \subseteq \mathbb{R}$, die als durch abzählbare Vereinigungen und Schnitte von Intervallen (offen, halboffen, oder geschlossen) auf \mathbb{R} dargestellt werden kann.
- Eine Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ ist **(Borel-)messbar**, falls das Urbild jeder Borel'schen Menge ebenfalls eine Borel'sche Menge ist.
- Für eine messbar Funktion f schreiben wir das **Lebesgue-Integral** als $\int f \, d\lambda$.

Messbarkeitstheorie

Definition 30

- Eine **Borel'sche Menge** für \mathbb{R} ist eine Teilmenge $A \subseteq \mathbb{R}$, die als durch abzählbare Vereinigungen und Schnitte von Intervallen (offen, halboffen, oder geschlossen) auf \mathbb{R} dargestellt werden kann.
- Eine Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ ist **(Borel-)messbar**, falls das Urbild jeder Borel'schen Menge ebenfalls eine Borel'sche Menge ist.
- Für eine messbar Funktion f schreiben wir das **Lebesgue-Integral** als $\int f \, d\lambda$.

Example 31 (Beispiele messbarer Funktionen)

- die charakteristische Funktion 1_A der Menge A ,
- stetige Funktionen, und
- Summen und Produkte messbarer Funktionen.

Wahrscheinlichkeitsräume über Borel'schen Mengen

Die Menge von Borel'schen Mengen \mathcal{A} ist eine σ -Algebra über \mathbb{R} .

Wahrscheinlichkeitsräume über Borel'schen Mengen

Die Menge von Borel'schen Mengen \mathcal{A} ist eine σ -Algebra über \mathbb{R} .

Eine Borel-messbare Funktion f mit den Eigenschaften einer kontinuierlichen Dichtefunktion definiert den Wahrscheinlichkeitsraum $(\mathbb{R}, \mathcal{A}, P)$

Wahrscheinlichkeitsräume über Borel'schen Mengen

Die Menge von Borel'schen Mengen \mathcal{A} ist eine σ -Algebra über \mathbb{R} .

Eine Borel-messbare Funktion f mit den Eigenschaften einer kontinuierlichen Dichtefunktion definiert den Wahrscheinlichkeitsraum $(\mathbb{R}, \mathcal{A}, P)$ mit

$$P : A \mapsto \int f \cdot 1_A d\lambda.$$

Wahrscheinlichkeitsräume über Borel'schen Mengen

Die Menge von Borel'schen Mengen \mathcal{A} ist eine σ -Algebra über \mathbb{R} .

Eine Borel-messbare Funktion f mit den Eigenschaften einer kontinuierlichen Dichtefunktion definiert den Wahrscheinlichkeitsraum $(\mathbb{R}, \mathcal{A}, P)$ mit

$$P : A \mapsto \int f \cdot 1_A d\lambda.$$

Insbesondere erfüllt P die Kolmogorov Axiome.

Kontinuierliche Wahrscheinlichkeitsräume

Definition 32

Ein Ereignis ist eine Menge $A = \bigcup_k I_k \subseteq \mathbb{R}$, die durch eine Vereinigung abzählbar vieler paarweise disjunkter Intervalle repräsentiert werden kann. Die Wahrscheinlichkeit von A ist gegeben als

$$P(A) = \int_A f_X(x) dx = \sum_k \int_{I_k} f_X(x) dx.$$

Kontinuierliche Wahrscheinlichkeitsräume

Definition 32

Ein Ereignis ist eine Menge $A = \bigcup_k I_k \subseteq \mathbb{R}$, die durch eine Vereinigung abzählbar vieler paarweise disjunkter Intervalle repräsentiert werden kann. Die Wahrscheinlichkeit von A ist gegeben als

$$P(A) = \int_A f_X(x) dx = \sum_k \int_{I_k} f_X(x) dx.$$

Die Wahrscheinlichkeit des Ereignisses $A = \{x\}, x \in \mathbb{R}$ ist immer 0.

Verteilungsfunktion

Die Verteilungsfunktion einer kontinuierlichen Zufallsvariable X ist gegeben als

$$F_X(x) = P(X \leq x)$$

Verteilungsfunktion

Die Verteilungsfunktion einer kontinuierlichen Zufallsvariable X ist gegeben als

$$F_X(x) = P(X \leq x) = P(X < x)$$

Verteilungsfunktion

Die Verteilungsfunktion einer kontinuierlichen Zufallsvariable X ist gegeben als

$$\begin{aligned} F_X(x) &= P(X \leq x) = P(X < x) \\ &= \int_{-\infty}^x f_X(t) dt. \end{aligned}$$

Verteilungsfunktion

Die Verteilungsfunktion einer kontinuierlichen Zufallsvariable X ist gegeben als

$$\begin{aligned} F_X(x) &= P(X \leq x) = P(X < x) \\ &= \int_{-\infty}^x f_X(t) dt. \end{aligned}$$

Die Dichtefunktion von X kann durch die Verteilungsfunktion von X erhalten werden indem ihre Ableitung bezüglich x gefunden wird:

$$f_X(x) = \frac{dF_X}{dx}.$$

Intervalle

Nach dem Hauptsatz der Differential- und Integralrechnung, ist die Wahrscheinlichkeit von $X \in [a, b]$ gegeben als

$$P(a \leq X \leq b) = F_X(b) - F_X(a) = \int_a^b f_X(x) dx.$$

Erwartungswerte

Der Erwartungswert einer kontinuierlichen Zufallsvariable X ist gegeben als

$$E(X) = \int_{-\infty}^{+\infty} x \cdot f_X(x) dx.$$

Erwartungswerte

Der Erwartungswert einer kontinuierlichen Zufallsvariable X ist gegeben als

$$E(X) = \int_{-\infty}^{+\infty} x \cdot f_X(x) dx.$$

LOTUS gilt auch im kontinuierlichen Fall:

$$E(g(X)) = \int_{-\infty}^{+\infty} g(x) \cdot f_X(x) dx.$$

Gleichverteilung

Definition 33 ($X \sim Unif(a, b)$)

Eine kontinuierliche Zufallsvariable X ist **gleichverteilt** mit Parametern a und b falls X den Ausgang eines Experimentes modelliert, wo alle Ergebnisse, die in dem Intervall $[a, b]$ liegen, gleichwahrscheinlich sind.

Gleichverteilung

Definition 33 ($X \sim Unif(a, b)$)

Eine kontinuierliche Zufallsvariable X ist **gleichverteilt** mit Parametern a und b falls X den Ausgang eines Experimentes modelliert, wo alle Ergebnisse, die in dem Intervall $[a, b]$ liegen, gleichwahrscheinlich sind.

$$f_X(x) = \begin{cases} \frac{1}{b-a} & \text{für } x \in [a, b] \\ 0 & \text{sonst} \end{cases} .$$

Gleichverteilung

Definition 33 ($X \sim Unif(a, b)$)

Eine kontinuierliche Zufallsvariable X ist **gleichverteilt** mit Parametern a und b falls X den Ausgang eines Experimentes modelliert, wo alle Ergebnisse, die in dem Intervall $[a, b]$ liegen, gleichwahrscheinlich sind.

$$f_X(x) = \begin{cases} \frac{1}{b-a} & \text{für } x \in [a, b] \\ 0 & \text{sonst} \end{cases} . \quad F_X(x) = \begin{cases} 0 & \text{für } x < a \\ \frac{x-a}{b-a} & \text{für } x \in [a, b] \\ 1 & \text{für } x > b \end{cases} .$$

Gleichverteilung

Definition 33 ($X \sim Unif(a, b)$)

Eine kontinuierliche Zufallsvariable X ist **gleichverteilt** mit Parametern a und b falls X den Ausgang eines Experimentes modelliert, wo alle Ergebnisse, die in dem Intervall $[a, b]$ liegen, gleichwahrscheinlich sind.

$$f_X(x) = \begin{cases} \frac{1}{b-a} & \text{für } x \in [a, b] \\ 0 & \text{sonst} \end{cases} . \quad F_X(x) = \begin{cases} 0 & \text{für } x < a \\ \frac{x-a}{b-a} & \text{für } x \in [a, b] \\ 1 & \text{für } x > b \end{cases} .$$

Übersicht

- $E(X) = \frac{a+b}{2}$

Gleichverteilung

Definition 33 ($X \sim Unif(a, b)$)

Eine kontinuierliche Zufallsvariable X ist **gleichverteilt** mit Parametern a und b falls X den Ausgang eines Experimentes modelliert, wo alle Ergebnisse, die in dem Intervall $[a, b]$ liegen, gleichwahrscheinlich sind.

$$f_X(x) = \begin{cases} \frac{1}{b-a} & \text{für } x \in [a, b] \\ 0 & \text{sonst} \end{cases} . \quad F_X(x) = \begin{cases} 0 & \text{für } x < a \\ \frac{x-a}{b-a} & \text{für } x \in [a, b] \\ 1 & \text{für } x > b \end{cases} .$$

Übersicht

- $E(X) = \frac{a+b}{2}$
- $Var(X) = \frac{(a-b)^2}{12}$

Universalität der Gleichverteilung

Sei $X \sim F$. Dann gilt $F(X) \sim Unif(0, 1)$.

Universalität der Gleichverteilung

Sei $X \sim F$. Dann gilt $F(X) \sim Unif(0, 1)$.

Realisierungen einer Zufallsvariablen mit Verteilung F und inverser Verteilungsfunktion F^{-1} können mittels Realisierungen einer gleichverteilten Zufallsvariablen Y simuliert werden: $F^{-1}(Y) \sim F$.

Normalverteilung

Definition 34 ($X \sim \mathcal{N}(\mu, \sigma^2)$)

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) =: \varphi(x; \mu, \sigma).$$

$$F_X(x) =: \Phi(x; \mu, \sigma).$$

Normalverteilung

Definition 34 ($X \sim \mathcal{N}(\mu, \sigma^2)$)

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) =: \varphi(x; \mu, \sigma).$$

$$F_X(x) =: \Phi(x; \mu, \sigma).$$

Übersicht

- $E(X) = \mu$

Normalverteilung

Definition 34 ($X \sim \mathcal{N}(\mu, \sigma^2)$)

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) =: \varphi(x; \mu, \sigma).$$

$$F_X(x) =: \Phi(x; \mu, \sigma).$$

Übersicht

- $E(X) = \mu$
- $\text{Var}(X) = \sigma^2$

Normalverteilung

Definition 34 ($X \sim \mathcal{N}(\mu, \sigma^2)$)

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) =: \varphi(x; \mu, \sigma).$$

$$F_X(x) =: \Phi(x; \mu, \sigma).$$

Übersicht

- $E(X) = \mu$
- $\text{Var}(X) = \sigma^2$
- $M_Z(s) = \exp(\mu s + \frac{(\sigma s)^2}{2})$

$\mathcal{N}(0, 1)$ heißt Standardnormalverteilung.

$\mathcal{N}(0, 1)$ heißt **Standardnormalverteilung**.

Lineare Transformation

Sei $X \sim \mathcal{N}(\mu, \sigma^2)$. Dann ist für alle $a \in \mathbb{R} \setminus \{0\}$ und $b \in \mathbb{R}$ die Zufallsvariable

$$Y = aX + b$$

normalverteilt mit Erwartungswert $a\mu + b$ und Varianz $a^2\sigma^2$.

$\mathcal{N}(0, 1)$ heißt **Standardnormalverteilung**.

Lineare Transformation

Sei $X \sim \mathcal{N}(\mu, \sigma^2)$. Dann ist für alle $a \in \mathbb{R} \setminus \{0\}$ und $b \in \mathbb{R}$ die Zufallsvariable

$$Y = aX + b$$

normalverteilt mit Erwartungswert $a\mu + b$ und Varianz $a^2\sigma^2$.

Normierung

Sei $X \sim \mathcal{N}(\mu, \sigma^2)$ und $Y = \frac{X - \mu}{\sigma}$. Dann gilt $Y \sim \mathcal{N}(0, 1)$.

$\mathcal{N}(0, 1)$ heißt **Standardnormalverteilung**.

Lineare Transformation

Sei $X \sim \mathcal{N}(\mu, \sigma^2)$. Dann ist für alle $a \in \mathbb{R} \setminus \{0\}$ und $b \in \mathbb{R}$ die Zufallsvariable

$$Y = aX + b$$

normalverteilt mit Erwartungswert $a\mu + b$ und Varianz $a^2\sigma^2$.

Normierung

Sei $X \sim \mathcal{N}(\mu, \sigma^2)$ und $Y = \frac{X - \mu}{\sigma}$. Dann gilt $Y \sim \mathcal{N}(0, 1)$.
Die Zufallsvariable Y heißt auch **normiert**.

Additivität

Seien X_1, \dots, X_n unabhängig und normalverteilt mit Parametern μ_i, σ_i^2 . Dann ist die Zufallsvariable

$$Z = a_1 X_1 + \dots + a_n X_n$$

normalverteilt mit Erwartungswert $a_1 \mu_1 + \dots + a_n \mu_n$ und Varianz $a_1^2 \sigma_1^2 + \dots + a_n^2 \sigma_n^2$.

Additivität

Seien X_1, \dots, X_n unabhängig und normalverteilt mit Parametern μ_i, σ_i^2 . Dann ist die Zufallsvariable

$$Z = a_1 X_1 + \dots + a_n X_n$$

normalverteilt mit Erwartungswert $a_1 \mu_1 + \dots + a_n \mu_n$ und Varianz $a_1^2 \sigma_1^2 + \dots + a_n^2 \sigma_n^2$.

Normal-Approximation der Binomialverteilung

Sei $X \sim \text{Bin}(n, p)$ mit Verteilungsfunktion $F_n(t)$. Dann kann

$$F_n(t) \approx \Phi \left(\frac{t - np}{\sqrt{p(1-p)n}} \right)$$

als Approximation verwendet werden falls $np \geq 5$ und $n(1-p) \geq 5$.

γ -Quantil

Definition 35

Sei X eine kontinuierliche Zufallsvariable mit Verteilung F_X . Eine Zahl x_γ mit

$$F_X(x_\gamma) = \gamma$$

heißt γ -Quantil von X bzw. der Verteilung F_X .

γ -Quantil

Definition 35

Sei X eine kontinuierliche Zufallsvariable mit Verteilung F_X . Eine Zahl x_γ mit

$$F_X(x_\gamma) = \gamma$$

heißt γ -Quantil von X bzw. der Verteilung F_X .

Definition 36

Für die Standardnormalverteilung bezeichnet z_γ das γ -Quantil.

Exponentialverteilung

Definition 37 ($X \sim \text{Exp}(\lambda)$)

Eine kontinuierliche Zufallsvariable X ist **exponentialverteilt** mit Parameter λ falls X die Zeit zwischen Ereignissen eines Poisson-Prozesses modelliert.

Exponentialverteilung

Definition 37 ($X \sim \text{Exp}(\lambda)$)

Eine kontinuierliche Zufallsvariable X ist **exponentialverteilt** mit Parameter λ falls X die Zeit zwischen Ereignissen eines Poisson-Prozesses modelliert.

$$f_X(x) = \lambda e^{-\lambda x}.$$

Exponentialverteilung

Definition 37 ($X \sim \text{Exp}(\lambda)$)

Eine kontinuierliche Zufallsvariable X ist **exponentialverteilt** mit Parameter λ falls X die Zeit zwischen Ereignissen eines Poisson-Prozesses modelliert.

$$f_X(x) = \lambda e^{-\lambda x}.$$

$$F_X(x) = 1 - e^{-\lambda x}.$$

Exponentialverteilung

Definition 37 ($X \sim \text{Exp}(\lambda)$)

Eine kontinuierliche Zufallsvariable X ist **exponentialverteilt** mit Parameter λ falls X die Zeit zwischen Ereignissen eines Poisson-Prozesses modelliert.

$$f_X(x) = \lambda e^{-\lambda x}.$$

$$F_X(x) = 1 - e^{-\lambda x}.$$

Übersicht

- $E(X) = \frac{1}{\lambda}$

Exponentialverteilung

Definition 37 ($X \sim \text{Exp}(\lambda)$)

Eine kontinuierliche Zufallsvariable X ist **exponentialverteilt** mit Parameter λ falls X die Zeit zwischen Ereignissen eines Poisson-Prozesses modelliert.

$$f_X(x) = \lambda e^{-\lambda x}.$$

$$F_X(x) = 1 - e^{-\lambda x}.$$

Übersicht

- $E(X) = \frac{1}{\lambda}$
- $\text{Var}(X) = \frac{1}{\lambda^2}$

Exponentialverteilung

Definition 37 ($X \sim \text{Exp}(\lambda)$)

Eine kontinuierliche Zufallsvariable X ist **exponentialverteilt** mit Parameter λ falls X die Zeit zwischen Ereignissen eines Poisson-Prozesses modelliert.

$$f_X(x) = \lambda e^{-\lambda x}.$$

$$F_X(x) = 1 - e^{-\lambda x}.$$

Übersicht

- $E(X) = \frac{1}{\lambda}$
- $\text{Var}(X) = \frac{1}{\lambda^2}$
- $M_X(s) = \frac{\lambda}{\lambda - s}, s < \lambda$

Skalierung

Sei $X \sim \text{Exp}(\lambda)$. Falls $a > 0$, dann ist $Y = aX$ exponentialverteilt mit dem Parameter λ/a .

Skalierung

Sei $X \sim \text{Exp}(\lambda)$. Falls $a > 0$, dann ist $Y = aX$ exponentialverteilt mit dem Parameter λ/a .

Gedächtnislosigkeit

Die Exponentialverteilung ist die **einzig**e gedächtnislose kontinuierliche Verteilung. Daher ist jede kontinuierliche Zufallsvariable X für die

$$P(X > y + x | X > x) = P(X > y)$$

für all $x, y > 0$ gilt, exponentialverteilt.

Warten auf mehrere Ereignisse

Seien X_1, \dots, X_n unabhängige, exponentialverteilte Zufallsvariablen mit Parametern $\lambda_1, \dots, \lambda_n$. Dann ist $X = \min\{X_1, \dots, X_n\}$ exponentialverteilt mit Parameter $\lambda_1 + \dots + \lambda_n$.

Warten auf mehrere Ereignisse

Seien X_1, \dots, X_n unabhängige, exponentialverteilte Zufallsvariablen mit Parametern $\lambda_1, \dots, \lambda_n$. Dann ist $X = \min\{X_1, \dots, X_n\}$ exponentialverteilt mit Parameter $\lambda_1 + \dots + \lambda_n$.

Exponential-Approximation der geometrischen Verteilung

Sei $X_n \sim \text{Geom}(\lambda/n)$. Die Verteilung der skalierten geometrisch verteilten Zufallsvariablen $Y_n = \frac{1}{n}X_n$ konvergiert mit $n \rightarrow \infty$ zu einer Exponentialverteilung mit Parameter λ .

Poisson-Prozess

Seien $T_1, T_2, \dots \sim \text{Exp}(\lambda)$ unabhängige und gleichverteilte Zufallsvariablen, die die Zeit zwischen dem $(i - 1)$ -ten und dem i -ten Ereignis modellieren.

Poisson-Prozess

Seien $T_1, T_2, \dots \sim \text{Exp}(\lambda)$ unabhängige und gleichverteilte Zufallsvariablen, die die Zeit zwischen dem $(i - 1)$ -ten und dem i -ten Ereignis modellieren.

Für $t > 0$ definieren wir

$$X(t) = \max\{n \in \mathbb{N} \mid T_1 + \dots + T_n \leq t\},$$

das die Anzahl der Ereignisse repräsentiert, die bis zum Zeitpunkt t aufgetreten sind.

Poisson-Prozess

Seien $T_1, T_2, \dots \sim \text{Exp}(\lambda)$ unabhängige und gleichverteilte Zufallsvariablen, die die Zeit zwischen dem $(i - 1)$ -ten und dem i -ten Ereignis modellieren.

Für $t > 0$ definieren wir

$$X(t) = \max\{n \in \mathbb{N} \mid T_1 + \dots + T_n \leq t\},$$

das die Anzahl der Ereignisse repräsentiert, die bis zum Zeitpunkt t aufgetreten sind.

Dann ist $X(t)$ Poisson-verteilt mit Parameter $t\lambda$.

Multivariate Verteilungen

Randverteilungen finden

Gegeben eine multivariate Verteilung $f_{X,Y}$ kann die Randverteilung f_X wie folgt gefunden werden:

$$f_X(x) = \int_{-\infty}^{+\infty} f_{X,Y}(x, y) dy.$$

Multivariate Verteilungen

Randverteilungen finden

Gegeben eine multivariate Verteilung $f_{X,Y}$ kann die Randverteilung f_X wie folgt gefunden werden:

$$f_X(x) = \int_{-\infty}^{+\infty} f_{X,Y}(x, y) dy.$$

Wahrscheinlichkeiten berechnen

Gegeben ein Ereignis $A \in \mathbb{R}^2$. Die Wahrscheinlichkeit von A ist die Fläche unter der Dichtefunktion von X :

$$P(A) = \iint_A f_{X,Y}(x, y) dx dy.$$

Dichtefunktionen finden

Gegeben eine Verteilungsfunktion $F_{X,Y}$ kann die Dichtefunktion $f_{X,Y}$ wie folgt gefunden werden:

$$f_{X,Y}(x,y) = \frac{\partial^2 F_{X,Y}}{\partial x \partial y}(x,y).$$

Dichtefunktionen finden

Gegeben eine Verteilungsfunktion $F_{X,Y}$ kann die Dichtefunktion $f_{X,Y}$ wie folgt gefunden werden:

$$f_{X,Y}(x,y) = \frac{\partial^2 F_{X,Y}}{\partial x \partial y}(x,y).$$

Verteilungsfunktionen finden

Gegeben eine Dichtefunktion $f_{X,Y}$ kann die Verteilungsfunktion $F_{X,Y}$ wie folgt gefunden werden:

$$F_{X,Y}(x,y) = \int_{-\infty}^y \int_{-\infty}^x f_{X,Y}(u,v) du dv.$$

Weitere Verteilungen

Definition 38 ($X \sim \text{Lognormal}(\mu, \sigma^2)$)

Eine kontinuierliche Zufallsvariable X ist **logarithmisch normalverteilt** mit Parametern μ und σ^2 falls $Y = \ln(X) \sim \mathcal{N}(\mu, \sigma^2)$.

Weitere Verteilungen

Definition 38 ($X \sim \text{Lognormal}(\mu, \sigma^2)$)

Eine kontinuierliche Zufallsvariable X ist **logarithmisch normalverteilt** mit Parametern μ und σ^2 falls $Y = \ln(X) \sim \mathcal{N}(\mu, \sigma^2)$.

$$f_X(x) = \frac{1}{x\sigma\sqrt{2\pi}} \cdot \exp\left(-\frac{(\ln(x) - \mu)^2}{2\sigma^2}\right)$$

für $x > 0$.

Plan I

Induktive Statistik

Schätzer

Maximum-Likelihood-Schätzer

Gesetz der großen Zahlen

Zentraler Grenzwertsatz

Konfidenzintervalle

Hypothesentests

Statistische Tests

Induktive Statistik

Induktive Statistik versucht mittels gemessener Größen auf zugrundeliegende Gesetzmäßigkeiten zu schließen.

Induktive Statistik

Induktive Statistik versucht mittels gemessener Größen auf zugrundeliegende Gesetzmäßigkeiten zu schließen.

Um Daten zu generieren, werden n unabhängige Kopien eines identischen Experimentes durchgeführt, das durch die Zufallsvariable X modelliert wird. Eine Messung, die aus einem dieser Experimente resultiert, heißt **Stichprobe**.

Induktive Statistik

Induktive Statistik versucht mittels gemessener Größen auf zugrundeliegende Gesetzmäßigkeiten zu schließen.

Um Daten zu generieren, werden n unabhängige Kopien eines identischen Experimentes durchgeführt, das durch die Zufallsvariable X modelliert wird. Eine Messung, die aus einem dieser Experimente resultiert, heißt **Stichprobe**.

Jede Stichprobe wird durch eine separate Zufallsvariable X_i repräsentiert, die als **Stichprobenvariable** bezeichnet wird.

Schätzer

Definition 39

Ein **Schätzer** für Parameter θ ist eine Zufallsvariable, die mehrere Stichprobenvariablen kombiniert und verwendet wird, um θ abzuschätzen.

Schätzer

Definition 39

Ein **Schätzer** für Parameter θ ist eine Zufallsvariable, die mehrere Stichprobenvariablen kombiniert und verwendet wird, um θ abzuschätzen.

Der **Bias** eines Schätzers U ist gegeben als $E(U - \theta)$.

Definition 39

Ein **Schätzer** für Parameter θ ist eine Zufallsvariable, die mehrere Stichprobenvariablen kombiniert und verwendet wird, um θ abzuschätzen.

Der **Bias** eines Schätzers U ist gegeben als $E(U - \theta)$.

Ein Schätzer U ist **erwartungstreu** bezüglich dem Parameter θ falls $E(U) = \theta$
(d.h. der Bias des Schätzers ist Null).

Definition 40

Das **Stichprobenmittel** \bar{X} ist ein erwartungstreuer Schätzer für $E(X)$.

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i.$$

Definition 40

Das **Stichprobenmittel** \bar{X} ist ein erwartungstreuer Schätzer für $E(X)$.

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i.$$

Definition 41

Die **Stichprobenvarianz** S^2 ist ein erwartungstreuer Schätzer für $\text{Var}(X)$.

$$S = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2}.$$

Definition 42

Der **mean squared error** ist ein Gütemaß für einen Schätzer U .

$$MSE(U) = E((U - \theta)^2).$$

Definition 42

Der **mean squared error** ist ein Gütemaß für einen Schätzer U .

$$MSE(U) = E((U - \theta)^2).$$

Falls U erwartungstreu ist, so gilt $MSE(U) = Var(U)$.

Definition 42

Der **mean squared error** ist ein Gütemaß für einen Schätzer U .

$$MSE(U) = E((U - \theta)^2).$$

Falls U erwartungstreu ist, so gilt $MSE(U) = Var(U)$.

Ein Schätzer A ist **effizienter** als ein anderer Schätzer B falls $MSE(A) < MSE(B)$.

Definition 42

Der **mean squared error** ist ein Gütemaß für einen Schätzer U .

$$MSE(U) = E((U - \theta)^2).$$

Falls U erwartungstreu ist, so gilt $MSE(U) = Var(U)$.

Ein Schätzer A ist **effizienter** als ein anderer Schätzer B falls $MSE(A) < MSE(B)$.

Ein Schätzer U ist **konsistent im quadratischen Mittel** falls $MSE(U) \xrightarrow{n \rightarrow \infty} 0$.

Maximum-Likelihood-Schätzer

Die **Maximum-Likelihood-Konstruktion** ist ein Verfahren zur Konstruktion eines Schätzers für Parameter von einer gegebenen Verteilung. Wir finden also den Parameter, für den die gemessenen Werte am wahrscheinlichsten sind. In anderen Worten, wir finden die Funktion, die am wahrscheinlichsten die Stichproben erklärt.

Maximum-Likelihood-Schätzer

Die **Maximum-Likelihood-Konstruktion** ist ein Verfahren zur Konstruktion eines Schätzers für Parameter von einer gegebenen Verteilung. Wir finden also den Parameter, für den die gemessenen Werte am wahrscheinlichsten sind. In anderen Worten, wir finden die Funktion, die am wahrscheinlichsten die Stichproben erklärt.

Gegeben Stichprobenvariablen $\vec{X} = (X_1, \dots, X_n)$ und Stichproben $\vec{x} = (x_1, \dots, x_n)$, finde einen Maximum-Likelihood-Schätzer für X mit Parameter θ .

1. konstruiere $L(\vec{x}; \theta) = f_{\vec{X}}(\vec{x}; \theta) = \prod_{i=1}^n f_{X_i}(x_i; \theta)$, das die Wahrscheinlichkeit modelliert, dass die Stichproben \vec{x} durch θ beschrieben werden
2. finde θ , das L maximiert, oder äquivalent $\ln L(\vec{x}; \theta) = \sum_{i=1}^n \ln f_{X_i}(x_i; \theta)$
3. der Wert für θ , der L maximiert, ist ein Maximum-Likelihood-Schätzer für θ

Gesetz der großen Zahlen

Das Gesetz der großen Zahlen besagt, dass das Stichprobenmittel aus unabhängigen und gleichverteilten Stichprobenvariablen \bar{X} mit Wahrscheinlichkeit 1 gegen den Erwartungswert $E(X)$ konvergiert während sich die Stichprobengröße n Unendlich annähert.

Gesetz der großen Zahlen

Das Gesetz der großen Zahlen besagt, dass das Stichprobenmittel aus unabhängigen und gleichverteilten Stichprobenvariablen \bar{X} mit Wahrscheinlichkeit 1 gegen den Erwartungswert $E(X)$ konvergiert während sich die Stichprobengröße n Unendlich annähert.

$$P(|\bar{X} - E(X)| \geq \delta) \leq \epsilon$$

für $\delta, \epsilon > 0$ und $n \geq \frac{\text{Var}(X)}{\epsilon\delta^2}$.

Zentraler Grenzwertsatz

Der zentrale Grenzwertsatz besagt, dass die normierte Summe von Stichprobenvariablen sich einer Standardnormalverteilung annähert während sich die Stichprobengröße n Unendlich annähert selbst wenn die zugrundeliegende Verteilung nicht die Normalverteilung ist.

Zentraler Grenzwertsatz

Der zentrale Grenzwertsatz besagt, dass die normierte Summe von Stichprobenvariablen sich einer Standardnormalverteilung annähert während sich die Stichprobengröße n Unendlich annähert selbst wenn die zugrundeliegende Verteilung nicht die Normalverteilung ist.

$$\frac{\sum_{i=1}^n X_i - n\mu}{\sigma\sqrt{n}} \xrightarrow{n \rightarrow \infty} \mathcal{N}(0, 1) \text{ in der Verteilung}$$

für X_i unabhängig und identisch verteilt.

Zentraler Grenzwertsatz

Der zentrale Grenzwertsatz besagt, dass die normierte Summe von Stichprobenvariablen sich einer Standardnormalverteilung annähert während sich die Stichprobengröße n Unendlich annähert selbst wenn die zugrundeliegende Verteilung nicht die Normalverteilung ist.

$$\frac{\sum_{i=1}^n X_i - n\mu}{\sigma\sqrt{n}} \xrightarrow{n \rightarrow \infty} \mathcal{N}(0, 1) \text{ in der Verteilung}$$

für X_i unabhängig und identisch verteilt.

Equivalent:

$$\sqrt{n} \left(\frac{\bar{X} - \mu}{\sigma} \right) \xrightarrow{n \rightarrow \infty} \mathcal{N}(0, 1) \text{ in der Verteilung.}$$

Grenzwertsatz nach de Moivre

Der Grenzwertsatz nach de Moivre ist ein Spezialfall des zentralen Grenzwertsatzes und sagt aus, dass die Normalverteilung als Approximation für die Binomialverteilung verwendet werden kann.

Grenzwertsatz nach de Moivre

Der Grenzwertsatz nach de Moivre ist ein Spezialfall des zentralen Grenzwertsatzes und sagt aus, dass die Normalverteilung als Approximation für die Binomialverteilung verwendet werden kann.

Seien $X_1, \dots, X_n \sim \text{Bern}(p)$ unabhängig und gleichverteilt sowie $H_n = X_1 + \dots + X_n$. Dann gilt

$$H_n^* = \frac{H_n - np}{\sqrt{np(1-p)}} \xrightarrow{n \rightarrow \infty} \mathcal{N}(0, 1) \text{ in der Verteilung.}$$

Konfidenzintervalle

Oft werden zwei Schätzer verwendet, um die abzuschätzende Größe aus beiden Richtungen abzuschätzen.

Konfidenzintervalle

Oft werden zwei Schätzer verwendet, um die abzuschätzende Größe aus beiden Richtungen abzuschätzen.

Die beiden Schätzer U_1 und U_2 werden so gewählt, dass

$$P(U_1 \leq \theta \leq U_2) \geq 1 - \alpha.$$

Konfidenzintervalle

Oft werden zwei Schätzer verwendet, um die abzuschätzende Größe aus beiden Richtungen abzuschätzen.

Die beiden Schätzer U_1 und U_2 werden so gewählt, dass

$$P(U_1 \leq \theta \leq U_2) \geq 1 - \alpha.$$

Die Wahrscheinlichkeit $1 - \alpha$ heißt **Konfidenzniveau**.

Konfidenzintervalle

Oft werden zwei Schätzer verwendet, um die abzuschätzende Größe aus beiden Richtungen abzuschätzen.

Die beiden Schätzer U_1 und U_2 werden so gewählt, dass

$$P(U_1 \leq \theta \leq U_2) \geq 1 - \alpha.$$

Die Wahrscheinlichkeit $1 - \alpha$ heißt **Konfidenzniveau**.

Berechnen wir für konkrete Stichproben die Schätzer U_1 und U_2 und erwarten $\theta \in [U_1, U_2]$, dann liegt die Fehlerwahrscheinlichkeit bei α . $[U_1, U_2]$ ist ein **Konfidenzintervall**.

Konfidenzintervalle

Oft werden zwei Schätzer verwendet, um die abzuschätzende Größe aus beiden Richtungen abzuschätzen.

Die beiden Schätzer U_1 und U_2 werden so gewählt, dass

$$P(U_1 \leq \theta \leq U_2) \geq 1 - \alpha.$$

Die Wahrscheinlichkeit $1 - \alpha$ heißt **Konfidenzniveau**.

Berechnen wir für konkrete Stichproben die Schätzer U_1 und U_2 und erwarten $\theta \in [U_1, U_2]$, dann liegt die Fehlerwahrscheinlichkeit bei α . $[U_1, U_2]$ ist ein **Konfidenzintervall**.

Oft wird ein einziger Schätzer U verwendet, um das symmetrische Konfidenzintervall $[U - \delta, U + \delta]$ zu definieren.

Hypothesentests

Gegeben Stichprobenvariablen $\vec{X} = (X_1, \dots, X_n)$ und
Stichprobenwerte $\vec{x} = (x_1, \dots, x_n)$ entscheide, ob eine Hypothese
akzeptiert oder abgelehnt werden soll.

Hypothesentests

Gegeben Stichprobenvariablen $\vec{X} = (X_1, \dots, X_n)$ und Stichprobenwerte $\vec{x} = (x_1, \dots, x_n)$ entscheide, ob eine Hypothese akzeptiert oder abgelehnt werden soll.

$K = \{\vec{x} \in \mathbb{R}^n \mid \vec{x} \text{ resultiert in Ablehnung der Hypothese}\}$ heißt **kritischer Bereich** (oder **Ablehnungsbereich**) eines Tests.

Hypothesentests

Gegeben Stichprobenvariablen $\vec{X} = (X_1, \dots, X_n)$ und Stichprobenwerte $\vec{x} = (x_1, \dots, x_n)$ entscheide, ob eine Hypothese akzeptiert oder abgelehnt werden soll.

$K = \{\vec{x} \in \mathbb{R}^n \mid \vec{x} \text{ resultiert in Ablehnung der Hypothese}\}$ heißt **kritischer Bereich** (oder **Ablehnungsbereich**) eines Tests.

K wird basierend auf den konkreten Werten der **Testvariablen** T gewählt, die sich aus den Stichprobenvariablen zusammensetzt.

Hypothesentests

Gegeben Stichprobenvariablen $\vec{X} = (X_1, \dots, X_n)$ und Stichprobenwerte $\vec{x} = (x_1, \dots, x_n)$ entscheide, ob eine Hypothese akzeptiert oder abgelehnt werden soll.

$K = \{\vec{x} \in \mathbb{R}^n \mid \vec{x} \text{ resultiert in Ablehnung der Hypothese}\}$ heißt **kritischer Bereich** (oder **Ablehnungsbereich**) eines Tests.

K wird basierend auf den konkreten Werten der **Testvariablen** T gewählt, die sich aus den Stichprobenvariablen zusammensetzt.

Ein Test heißt **einseitig** falls K ein halboffenes Intervall in $T(S)$ ist und **beidseitig** falls K ein geschlossenes Intervall in $T(S)$ ist.

H_0 ist die Hypothese auf die getestet wird, auch als **Nullhypothese** bezeichnet.

H_1 ist die **Alternative**. H_1 ist **trivial** falls es die einfache Negation von H_0 ist.

H_0 ist die Hypothese auf die getestet wird, auch als **Nullhypothese** bezeichnet.

H_1 ist die **Alternative**. H_1 ist **trivial** falls es die einfache Negation von H_0 ist.

Fehler

- **Fehler 1. Art** oder **α -Fehler** oder **Signifikanzniveau**
 H_0 gilt, aber $\vec{x} \in K$

$$\alpha = \sup_{p \in H_0} P_p(T \in K).$$

H_0 ist die Hypothese auf die getestet wird, auch als **Nullhypothese** bezeichnet.

H_1 ist die **Alternative**. H_1 ist **trivial** falls es die einfache Negation von H_0 ist.

Fehler

- **Fehler 1. Art** oder **α -Fehler** oder **Signifikanzniveau**
 H_0 gilt, aber $\vec{x} \in K$

$$\alpha = \sup_{p \in H_0} P_p(T \in K).$$

- **Fehler 2. Art** oder **β -Fehler**
 H_1 gilt, aber $\vec{x} \notin K$

$$\beta = \sup_{p \in H_1} P_p(T \notin K).$$

Die **Gütefunktion** g beschreibt die Wahrscheinlichkeit, dass ein Test die Nullhypothese ablehnt.

$$g(p) = P_p(T \in K).$$

Statistische Tests

Eigenschaften

Statistische Tests können anhand einiger Merkmale unterschieden werden:

- **Anzahl beteiligter Zufallsgrößen**

Statistische Tests

Eigenschaften

Statistische Tests können anhand einiger Merkmale unterschieden werden:

- **Anzahl beteiligter Zufallsgrößen**
Vergleich zweier Zufallsvariablen mit potenziell unterschiedlichen Verteilungen (**Zwei-Stichproben-Test**), oder Untersuchung einer Zufallsvariable (**Ein-Stichproben-Test**)?

Statistische Tests

Eigenschaften

Statistische Tests können anhand einiger Merkmale unterschieden werden:

- **Anzahl beteiligter Zufallsgrößen**
Vergleich zweier Zufallsvariablen mit potenziell unterschiedlichen Verteilungen (**Zwei-Stichproben-Test**), oder Untersuchung einer Zufallsvariable (**Ein-Stichproben-Test**)?
 - Unabhängigkeit beteiligter Zufallsvariablen
Bei dem Vergleich mehrerer Zufallsvariablen wird unterschieden, ob **unabhängige Messungen** (Unabhängigkeit) oder **verbundene Messungen** (Abhängigkeit) vorgenommen werden.

Statistische Tests

Eigenschaften

Statistische Tests können anhand einiger Merkmale unterschieden werden:

- **Anzahl beteiligter Zufallsgrößen**

Vergleich zweier Zufallsvariablen mit potenziell unterschiedlichen Verteilungen (**Zwei-Stichproben-Test**), oder Untersuchung einer Zufallsvariable (**Ein-Stichproben-Test**)?

- Unabhängigkeit beteiligter Zufallsvariablen

Bei dem Vergleich mehrerer Zufallsvariablen wird unterschieden, ob **unabhängige Messungen** (Unabhängigkeit) oder **verbundene Messungen** (Abhängigkeit) vorgenommen werden.

- Betrachtung des Zusammenhangs mehrerer Zufallsvariablen
Wird der funktionale Zusammenhang mehrerer Zufallsvariablen untersucht spricht man von einer **Regressionsanalyse**. Werden die Zufallsvariablen auf Unabhängigkeit untersucht spricht man von einer **Zusammenhangsanalyse**.

- Formulierung der Nullhypothese

- **Formulierung der Nullhypothese**

Über welche **Lageparameter** macht der Test eine Aussage (z.B. Erwartungswert oder Varianz), oder wird auf eine vorgegebene Verteilung getestet?

- **Formulierung der Nullhypothese**

Über welche **Lageparameter** macht der Test eine Aussage (z.B. Erwartungswert oder Varianz), oder wird auf eine vorgegebene Verteilung getestet?

- **Annahmen über die Zufallsgrößen**

- **Formulierung der Nullhypothese**

Über welche **Lageparameter** macht der Test eine Aussage (z.B. Erwartungswert oder Varianz), oder wird auf eine vorgegebene Verteilung getestet?

- **Annahmen über die Zufallsgrößen**

Welche Annahmen macht der Test über die Zufallsvariablen wie zum Beispiel die Art der Verteilung, Erwartungswert, oder Varianz?

Wichtige statistische Testverfahren

Wichtige statistische Testverfahren

- Approximativer Binomialtest

Wichtige statistische Testverfahren

- Approximativer Binomialtest
- Gaußtest

Wichtige statistische Testverfahren

- Approximativer Binomialtest
- Gaußtest
- t -Test

Wichtige statistische Testverfahren

- Approximativer Binomialtest
- Gaußtest
- t -Test
- Zwei-Stichproben- t -Test

Wichtige statistische Testverfahren

- Approximativer Binomialtest
- Gaußtest
- t -Test
- Zwei-Stichproben- t -Test
- χ^2 -Anpassungstest

Plan I

Markovketten

Stochastische Prozesse

Markov-Bedingung

Repräsentationen

Wahrscheinlichkeiten

Übergangszeiten

Stationäre Verteilung

Exkurs: Diagonalisierung

Konvergenz

Eigenschaften

Stochastische Prozesse

Definition 43

Ein **stochastischer Prozess** ist eine Folge von Zufallsvariablen $(X_t)_{t \in \mathcal{T}}$, die das Verhalten eines Systems zu verschiedenen Zeitpunkten t angeben.

Stochastische Prozesse

Definition 43

Ein **stochastischer Prozess** ist eine Folge von Zufallsvariablen $(X_t)_{t \in T}$, die das Verhalten eines Systems zu verschiedenen Zeitpunkten t angeben.

Gilt $T = \mathbb{N}_0$, so spricht man von einem stochastischen Prozess mit **diskreter Zeit**.

Stochastische Prozesse

Definition 43

Ein **stochastischer Prozess** ist eine Folge von Zufallsvariablen $(X_t)_{t \in T}$, die das Verhalten eines Systems zu verschiedenen Zeitpunkten t angeben.

Gilt $T = \mathbb{N}_0$, so spricht man von einem stochastischen Prozess mit **diskreter Zeit**. Gilt hingegen $T = \mathbb{R}_0^+$, so spricht man von einem stochastischen Prozess mit **kontinuierlicher Zeit**.

Falls X_t diskret ist, spricht man auch von unterschiedlichen **Zuständen** die ein System zum Zeitpunkt t annimmt.

Markov-Bedingung

Definition 44

Ein stochastischer Prozess erfüllt die **Markov-Bedingung**, falls die Wahrscheinlichkeitsverteilung der Zustände zum Zeitpunkt $t + 1$ nur von der Wahrscheinlichkeitsverteilung der Zustände zum Zeitpunkt t abhängt, nicht aber von Zuständen zum Zeitpunkt $< t$.

Markov-Bedingung

Definition 44

Ein stochastischer Prozess erfüllt die **Markov-Bedingung**, falls die Wahrscheinlichkeitsverteilung der Zustände zum Zeitpunkt $t + 1$ nur von der Wahrscheinlichkeitsverteilung der Zustände zum Zeitpunkt t abhängt, nicht aber von Zuständen zum Zeitpunkt $< t$.

Diese Bedingung kann wie folgt formalisiert werden:

$$P(X_{t+1} = j | X_t = i_t, \dots, X_0 = i_0) = P(X_{t+1} = j | X_t = i_t)$$

Markov-Bedingung

Definition 44

Ein stochastischer Prozess erfüllt die **Markov-Bedingung**, falls die Wahrscheinlichkeitsverteilung der Zustände zum Zeitpunkt $t + 1$ nur von der Wahrscheinlichkeitsverteilung der Zustände zum Zeitpunkt t abhängt, nicht aber von Zuständen zum Zeitpunkt $< t$.

Diese Bedingung kann wie folgt formalisiert werden:

$$P(X_{t+1} = j | X_t = i_t, \dots, X_0 = i_0) = P(X_{t+1} = j | X_t = i_t) =: p_{i_t j}^t.$$

Definition 45

Eine (endliche) Markov-Kette (mit diskreter Zeit) über Zustandsmenge $S = \{0, \dots, n - 1\}$ besteht aus einer unendlichen Folge von Zufallsvariablen $(X_t)_{t \in \mathbb{N}_0}$ mit Wertebereich S

Definition 45

Eine (endliche) Markov-Kette (mit diskreter Zeit) über Zustandsmenge $S = \{0, \dots, n - 1\}$ besteht aus einer unendlichen Folge von Zufallsvariablen $(X_t)_{t \in \mathbb{N}_0}$ mit Wertebereich S und Startverteilung q_0 mit $q_0^T \in \mathbb{R}^n$.

Definition 45

Eine (endliche) Markov-Kette (mit diskreter Zeit) über Zustandsmenge $S = \{0, \dots, n - 1\}$ besteht aus einer unendlichen Folge von Zufallsvariablen $(X_t)_{t \in \mathbb{N}_0}$ mit Wertebereich S und Startverteilung q_0 mit $q_0^T \in \mathbb{R}^n$. q_0 muss als Zeilenvektor eine valide diskrete Dichtefunktion auf der Zustandsmenge S beschreiben.

Definition 45

Eine (endliche) Markov-Kette (mit diskreter Zeit) über Zustandsmenge $S = \{0, \dots, n - 1\}$ besteht aus einer unendlichen Folge von Zufallsvariablen $(X_t)_{t \in \mathbb{N}_0}$ mit Wertebereich S und Startverteilung q_0 mit $q_0^T \in \mathbb{R}^n$. q_0 muss als Zeilenvektor eine valide diskrete Dichtefunktion auf der Zustandsmenge S beschreiben. Weiterhin muss die Markov-Bedingung erfüllt sein.

Repräsentationen

Definition 46

Sind die Übergangswahrscheinlichkeiten $p_{ij} = P(X_{t+1} = j | X_t = i)$ konstant mit der Zeit t , so spricht man von einer (zeit-)homogenen Markov-Kette.

Repräsentationen

Definition 46

Sind die Übergangswahrscheinlichkeiten $p_{ij} = P(X_{t+1} = j | X_t = i)$ konstant mit der Zeit t , so spricht man von einer (zeit-)homogenen Markov-Kette.

In diesem Fall definiert man die Übergangsmatrix durch

$$P = (p_{ij})_{0 \leq i, j < n}.$$

Repräsentationen

Definition 46

Sind die Übergangswahrscheinlichkeiten $p_{ij} = P(X_{t+1} = j | X_t = i)$ konstant mit der Zeit t , so spricht man von einer (zeit-)homogenen Markov-Kette.

In diesem Fall definiert man die Übergangsmatrix durch

$$P = (p_{ij})_{0 \leq i, j < n}.$$

Das Übergangsdiagramm ist ein Graph bestehend aus den Zuständen S wobei die gewichteten Kanten durch P repräsentiert werden.

Repräsentationen

Definition 46

Sind die Übergangswahrscheinlichkeiten $p_{ij} = P(X_{t+1} = j | X_t = i)$ konstant mit der Zeit t , so spricht man von einer (zeit-)homogenen Markov-Kette.

In diesem Fall definiert man die Übergangsmatrix durch

$$P = (p_{ij})_{0 \leq i, j < n}.$$

Das Übergangsdiagramm ist ein Graph bestehend aus den Zuständen S wobei die gewichteten Kanten durch P repräsentiert werden.

Einen konkreten Ablauf des Systems kann man sich auch als Random Walk auf dem Übergangsdiagramm vorstellen.

Wahrscheinlichkeiten

Die Verteilung einer Markov-Kette kann iterativ für immer größere Zeitpunkte t bestimmt werden:

$$q_{t+1} = q_t \cdot P$$

Wahrscheinlichkeiten

Die Verteilung einer Markov-Kette kann iterativ für immer größere Zeitpunkte t bestimmt werden:

$$q_{t+1} = q_t \cdot P$$

$$q_t = q_0 \cdot P^t$$

Wahrscheinlichkeiten

Die Verteilung einer Markov-Kette kann iterativ für immer größere Zeitpunkte t bestimmt werden:

$$q_{t+1} = q_t \cdot P$$

$$q_t = q_0 \cdot P^t$$

$$q_{t+k} = q_t \cdot P^k.$$

Wahrscheinlichkeiten

Die Verteilung einer Markov-Kette kann iterativ für immer größere Zeitpunkte t bestimmt werden:

$$q_{t+1} = q_t \cdot P$$

$$q_t = q_0 \cdot P^t$$

$$q_{t+k} = q_t \cdot P^k.$$

Definition 47

q_t wird auch als **Zustandsvektor** der Markov-Kette zum Zeitpunkt t bezeichnet.

Wahrscheinlichkeiten

Die Verteilung einer Markov-Kette kann iterativ für immer größere Zeitpunkte t bestimmt werden:

$$q_{t+1} = q_t \cdot P$$

$$q_t = q_0 \cdot P^t$$

$$q_{t+k} = q_t \cdot P^k.$$

Definition 47

q_t wird auch als **Zustandsvektor** der Markov-Kette zum Zeitpunkt t bezeichnet.

Die Einträge von P^k geben an, mit welcher Wahrscheinlichkeit ein Übergang von Zustand i in Zustand j in genau k Schritten erfolgt:

$$p_{ij}^{(k)} = P(X_{t+k} = j | X_t = i) = (P^k)_{ij}.$$

Übergangszeiten

Definition 48

Die **Übergangszeit** von Zustand i in Zustand j wird durch die folgende Zufallsvariable modelliert:

$$T_{ij} = \min\{n \geq 1 \mid X_n = j, \text{ wenn } X_0 = i\}.$$

Übergangszeiten

Definition 48

Die **Übergangszeit** von Zustand i in Zustand j wird durch die folgende Zufallsvariable modelliert:

$$T_{ij} = \min\{n \geq 1 \mid X_n = j, \text{ wenn } X_0 = i\}.$$

Die **erwartete Übergangszeit** ist gegeben durch

$$h_{ij} = E(T_{ij})$$

Übergangszeiten

Definition 48

Die **Übergangszeit** von Zustand i in Zustand j wird durch die folgende Zufallsvariable modelliert:

$$T_{ij} = \min\{n \geq 1 \mid X_n = j, \text{ wenn } X_0 = i\}.$$

Die **erwartete Übergangszeit** ist gegeben durch

$$\begin{aligned} h_{ij} &= E(T_{ij}) \\ &= 1 + \sum_{k \neq j} p_{ik} h_{kj}. \end{aligned}$$

Die Wahrscheinlichkeit vom Zustand i in beliebig vielen Schritten in den Zustand j zu gelangen, heißt **Ankunftswahrscheinlichkeit** f_{ij} :

$$f_{ij} = P(T_{ij} < \infty)$$

Die Wahrscheinlichkeit vom Zustand i in beliebig vielen Schritten in den Zustand j zu gelangen, heißt **Ankunftswahrscheinlichkeit** f_{ij} :

$$\begin{aligned} f_{ij} &= P(T_{ij} < \infty) \\ &= p_{ij} + \sum_{k \neq j} p_{ik} f_{kj}. \end{aligned}$$

Die Wahrscheinlichkeit vom Zustand i in beliebig vielen Schritten in den Zustand j zu gelangen, heißt **Ankunftswahrscheinlichkeit** f_{ij} :

$$\begin{aligned} f_{ij} &= P(T_{ij} < \infty) \\ &= p_{ij} + \sum_{k \neq j} p_{ik} f_{kj}. \end{aligned}$$

Definition 49

Die Zufallsvariable $T_i = T_{ii}$ gibt die **Rückkehrzeit** von Zustand i zu Zustand i an.

Die Wahrscheinlichkeit vom Zustand i in beliebig vielen Schritten in den Zustand j zu gelangen, heißt **Ankunftswahrscheinlichkeit** f_{ij} :

$$\begin{aligned} f_{ij} &= P(T_{ij} < \infty) \\ &= p_{ij} + \sum_{k \neq j} p_{ik} f_{kj}. \end{aligned}$$

Definition 49

Die Zufallsvariable $T_i = T_{ii}$ gibt die **Rückkehrzeit** von Zustand i zu Zustand i an.

Die **erwartete Rückkehrzeit** $h_i = h_{ii}$ und die **Rückkehrwahrscheinlichkeit** $f_i = f_{ii}$ sind analog zu der erwarteten Übergangszeit und der Ankunftswahrscheinlichkeit definiert.

Stationäre Verteilung

Definition 50

Ein Zustandsvektor π mit $\pi = \pi \cdot P$ heißt **stationäre Verteilung** einer Markov-Kette.

Stationäre Verteilung

Definition 50

Ein Zustandsvektor π mit $\pi = \pi \cdot P$ heißt **stationäre Verteilung** einer Markov-Kette.

Eine Markovkette konvergiert nicht notwendigerweise in eine ihrer stationären Verteilungen. Konvergenz hängt von den Eigenschaften der Markovkette und der Startverteilung ab.

Exkurs: Diagonalisierung

Für Eigenvektoren x_i und zugehörige Eigenwerte λ_i einer Matrix A gilt $A \cdot x_i = \lambda_i \cdot x_i$.

Exkurs: Diagonalisierung

Für Eigenvektoren x_i und zugehörige Eigenwerte λ_i einer Matrix A gilt $A \cdot x_i = \lambda_i \cdot x_i$.

Damit gilt für eine quadratische Matrix A mit Eigenvektoren x_1, \dots, x_n und zugehörigen Eigenwerten $\lambda_1, \dots, \lambda_n$

Exkurs: Diagonalisierung

Für Eigenvektoren x_i und zugehörige Eigenwerte λ_i einer Matrix A gilt $A \cdot x_i = \lambda_i \cdot x_i$.

Damit gilt für eine quadratische Matrix A mit Eigenvektoren x_1, \dots, x_n und zugehörigen Eigenwerten $\lambda_1, \dots, \lambda_n$, dass

$$A \cdot [x_1 \quad \cdots \quad x_n]$$

Exkurs: Diagonalisierung

Für Eigenvektoren x_i und zugehörige Eigenwerte λ_i einer Matrix A gilt $A \cdot x_i = \lambda_i \cdot x_i$.

Damit gilt für eine quadratische Matrix A mit Eigenvektoren x_1, \dots, x_n und zugehörigen Eigenwerten $\lambda_1, \dots, \lambda_n$, dass

$$A \cdot \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix} = \begin{bmatrix} \lambda_1 x_1 & \cdots & \lambda_n x_n \end{bmatrix}$$

Exkurs: Diagonalisierung

Für Eigenvektoren x_i und zugehörige Eigenwerte λ_i einer Matrix A gilt $A \cdot x_i = \lambda_i \cdot x_i$.

Damit gilt für eine quadratische Matrix A mit Eigenvektoren x_1, \dots, x_n und zugehörigen Eigenwerten $\lambda_1, \dots, \lambda_n$, dass

$$\begin{aligned} A \cdot [x_1 \quad \cdots \quad x_n] &= [\lambda_1 x_1 \quad \cdots \quad \lambda_n x_n] \\ &= [x_1 \quad \cdots \quad x_n] \cdot \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{bmatrix}. \end{aligned}$$

Exkurs: Diagonalisierung

Für Eigenvektoren x_i und zugehörige Eigenwerte λ_i einer Matrix A gilt $A \cdot x_i = \lambda_i \cdot x_i$.

Damit gilt für eine quadratische Matrix A mit Eigenvektoren x_1, \dots, x_n und zugehörigen Eigenwerten $\lambda_1, \dots, \lambda_n$, dass

$$\begin{aligned} A \cdot [x_1 \quad \cdots \quad x_n] &= [\lambda_1 x_1 \quad \cdots \quad \lambda_n x_n] \\ &= [x_1 \quad \cdots \quad x_n] \cdot \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{bmatrix}. \end{aligned}$$

Sei V die Matrix bestehend aus den Eigenvektoren von A als Spaltenvektoren und sei Λ die Diagonalmatrix bestehend aus den Eigenwerten von A .

Exkurs: Diagonalisierung

Für Eigenvektoren x_i und zugehörige Eigenwerte λ_i einer Matrix A gilt $A \cdot x_i = \lambda_i \cdot x_i$.

Damit gilt für eine quadratische Matrix A mit Eigenvektoren x_1, \dots, x_n und zugehörigen Eigenwerten $\lambda_1, \dots, \lambda_n$, dass

$$\begin{aligned} A \cdot [x_1 \quad \cdots \quad x_n] &= [\lambda_1 x_1 \quad \cdots \quad \lambda_n x_n] \\ &= [x_1 \quad \cdots \quad x_n] \cdot \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{bmatrix}. \end{aligned}$$

Sei V die Matrix bestehend aus den Eigenvektoren von A als Spaltenvektoren und sei Λ die Diagonalmatrix bestehend aus den Eigenwerten von A .

Dann heißt $V^{-1} \cdot A \cdot V = \Lambda$ **Diagonalisierung** von A .

Exkurs: Diagonalisierung

Für Eigenvektoren x_i und zugehörige Eigenwerte λ_i einer Matrix A gilt $A \cdot x_i = \lambda_i \cdot x_i$.

Damit gilt für eine quadratische Matrix A mit Eigenvektoren x_1, \dots, x_n und zugehörigen Eigenwerten $\lambda_1, \dots, \lambda_n$, dass

$$\begin{aligned} A \cdot [x_1 \quad \cdots \quad x_n] &= [\lambda_1 x_1 \quad \cdots \quad \lambda_n x_n] \\ &= [x_1 \quad \cdots \quad x_n] \cdot \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{bmatrix}. \end{aligned}$$

Sei V die Matrix bestehend aus den Eigenvektoren von A als Spaltenvektoren und sei Λ die Diagonalmatrix bestehend aus den Eigenwerten von A .

Dann heißt $V^{-1} \cdot A \cdot V = \Lambda$ **Diagonalisierung** von A .
Andererseits gilt auch, dass $A = V \cdot \Lambda \cdot V^{-1}$.

Konvergenz

Aus der Diagonalisierung der Übergangsmatrix folgt direkt, dass

$$P^t = V \cdot \Lambda^t \cdot V^{-1}.$$

Konvergenz

Aus der Diagonalisierung der Übergangsmatrix folgt direkt, dass

$$P^t = V \cdot \Lambda^t \cdot V^{-1}.$$

Damit kann das Verhalten der Markovkette für $t \rightarrow \infty$ beschrieben werden:

$$\lim_{t \rightarrow \infty} q_t$$

Konvergenz

Aus der Diagonalisierung der Übergangsmatrix folgt direkt, dass

$$P^t = V \cdot \Lambda^t \cdot V^{-1}.$$

Damit kann das Verhalten der Markovkette für $t \rightarrow \infty$ beschrieben werden:

$$\lim_{t \rightarrow \infty} q_t = \lim_{t \rightarrow \infty} q_0 \cdot P^t.$$

Konvergenz

Aus der Diagonalisierung der Übergangsmatrix folgt direkt, dass

$$P^t = V \cdot \Lambda^t \cdot V^{-1}.$$

Damit kann das Verhalten der Markovkette für $t \rightarrow \infty$ beschrieben werden:

$$\lim_{t \rightarrow \infty} q_t = \lim_{t \rightarrow \infty} q_0 \cdot P^t.$$
$$\lim_{t \rightarrow \infty} P(X_t = j \mid X_0 = i)$$

Konvergenz

Aus der Diagonalisierung der Übergangsmatrix folgt direkt, dass

$$P^t = V \cdot \Lambda^t \cdot V^{-1}.$$

Damit kann das Verhalten der Markovkette für $t \rightarrow \infty$ beschrieben werden:

$$\begin{aligned}\lim_{t \rightarrow \infty} q_t &= \lim_{t \rightarrow \infty} q_0 \cdot P^t. \\ \lim_{t \rightarrow \infty} P(X_t = j \mid X_0 = i) &= \lim_{t \rightarrow \infty} P^t(i, j).\end{aligned}$$

Eigenschaften

Weißt eine Markovkette bestimmte Eigenschaften auf, können Rückschlüsse auf ihre stationäre Verteilung gezogen werden.

Eigenschaften

Weißt eine Markovkette bestimmte Eigenschaften auf, können Rückschlüsse auf ihre stationäre Verteilung gezogen werden.

Definition 51

Ein Zustand i heißt **absorbierend**, falls $p_{ii} = 1$, d.h. aus ihm führen keine Übergänge heraus.

Eigenschaften

Weißt eine Markovkette bestimmte Eigenschaften auf, können Rückschlüsse auf ihre stationäre Verteilung gezogen werden.

Definition 51

Ein Zustand i heißt **absorbierend**, falls $p_{ii} = 1$, d.h. aus ihm führen keine Übergänge heraus.

Ein Zustand i heißt **rekurrent**, falls $f_i = 1$, d.h. mit Wahrscheinlichkeit 1 kehrt die Markovkette in den Zustand i zurück.

Eigenschaften

Weißt eine Markovkette bestimmte Eigenschaften auf, können Rückschlüsse auf ihre stationäre Verteilung gezogen werden.

Definition 51

Ein Zustand i heißt **absorbierend**, falls $p_{ii} = 1$, d.h. aus ihm führen keine Übergänge heraus.

Ein Zustand i heißt **rekurrent**, falls $f_i = 1$, d.h. mit Wahrscheinlichkeit 1 kehrt die Markovkette in den Zustand i zurück.

Falls hingegen $f_i < 1$ gilt, heißt der Zustand i **transient**.

Definition 52

Eine Markov-Kette heißt **irreduzibel** wenn man von jedem Zustand jeden anderen Zustand mit positiver Wahrscheinlichkeit erreicht, wenn man nur genügend Schritte ausführt.

Definition 52

Eine Markov-Kette heißt **irreduzibel** wenn man von jedem Zustand jeden anderen Zustand mit positiver Wahrscheinlichkeit erreicht, wenn man nur genügend Schritte ausführt. Das heißt formal:

$$\forall i, j \in S. \exists n \in \mathbb{N}. p_{ij}^{(n)} > 0.$$

Definition 52

Eine Markov-Kette heißt **irreduzibel** wenn man von jedem Zustand jeden anderen Zustand mit positiver Wahrscheinlichkeit erreicht, wenn man nur genügend Schritte ausführt. Das heißt formal:

$$\forall i, j \in S. \exists n \in \mathbb{N}. p_{ij}^{(n)} > 0.$$

Eine endliche Markov-Kette ist genau dann irreduzibel wenn ihr Übergangsdiagramm stark zusammenhängend ist.

Definition 52

Eine Markov-Kette heißt **irreduzibel** wenn man von jedem Zustand jeden anderen Zustand mit positiver Wahrscheinlichkeit erreicht, wenn man nur genügend Schritte ausführt. Das heißt formal:

$$\forall i, j \in S. \exists n \in \mathbb{N}. p_{ij}^{(n)} > 0.$$

Eine endliche Markov-Kette ist genau dann irreduzibel wenn ihr Übergangsdiagramm stark zusammenhängend ist.

Wenn eine endliche Markov-Kette irreduzibel ist gilt zudem:

Definition 52

Eine Markov-Kette heißt **irreduzibel** wenn man von jedem Zustand jeden anderen Zustand mit positiver Wahrscheinlichkeit erreicht, wenn man nur genügend Schritte ausführt. Das heißt formal:

$$\forall i, j \in S. \exists n \in \mathbb{N}. p_{ij}^{(n)} > 0.$$

Eine endliche Markov-Kette ist genau dann irreduzibel wenn ihr Übergangendiagramm stark zusammenhängend ist.

Wenn eine endliche Markov-Kette irreduzibel ist gilt zudem:

- $f_{ij} = 1, \forall i, j \in S$

Definition 52

Eine Markov-Kette heißt **irreduzibel** wenn man von jedem Zustand jeden anderen Zustand mit positiver Wahrscheinlichkeit erreicht, wenn man nur genügend Schritte ausführt. Das heißt formal:

$$\forall i, j \in S. \exists n \in \mathbb{N}. p_{ij}^{(n)} > 0.$$

Eine endliche Markov-Kette ist genau dann irreduzibel wenn ihr Übergangendiagramm stark zusammenhängend ist.

Wenn eine endliche Markov-Kette irreduzibel ist gilt zudem:

- $f_{ij} = 1, \forall i, j \in S$;
- h_{ij} existiert, $\forall i, j \in S$

Definition 52

Eine Markov-Kette heißt **irreduzibel** wenn man von jedem Zustand jeden anderen Zustand mit positiver Wahrscheinlichkeit erreicht, wenn man nur genügend Schritte ausführt. Das heißt formal:

$$\forall i, j \in S. \exists n \in \mathbb{N}. p_{ij}^{(n)} > 0.$$

Eine endliche Markov-Kette ist genau dann irreduzibel wenn ihr Übergangendiagramm stark zusammenhängend ist.

Wenn eine endliche Markov-Kette irreduzibel ist gilt zudem:

- $f_{ij} = 1, \forall i, j \in S$;
- h_{ij} existiert, $\forall i, j \in S$; und
- es existiert eine eindeutige stationäre Verteilung π mit $\pi(j) = \frac{1}{h_j}, \forall j \in S$.

Definition 52

Eine Markov-Kette heißt **irreduzibel** wenn man von jedem Zustand jeden anderen Zustand mit positiver Wahrscheinlichkeit erreicht, wenn man nur genügend Schritte ausführt. Das heißt formal:

$$\forall i, j \in S. \exists n \in \mathbb{N}. p_{ij}^{(n)} > 0.$$

Eine endliche Markov-Kette ist genau dann irreduzibel wenn ihr Übergangendiagramm stark zusammenhängend ist.

Wenn eine endliche Markov-Kette irreduzibel ist gilt zudem:

- $f_{ij} = 1, \forall i, j \in S$;
- h_{ij} existiert, $\forall i, j \in S$; und
- es existiert eine eindeutige stationäre Verteilung π mit $\pi(j) = \frac{1}{h_j}, \forall j \in S$.

Die Markov-Kette muss nicht zwingend in eine stationäre Verteilung konvergieren (Periodizität!).

Nun wollen wir die Periodizität von Zuständen betrachten.

Nun wollen wir die Periodizität von Zuständen betrachten.

Definition 53

Für einen Zustand i definieren wir

$$T(i) = \{n \geq 1 \mid P^n(i, i) > 0\}.$$

Nun wollen wir die Periodizität von Zuständen betrachten.

Definition 53

Für einen Zustand i definieren wir

$$T(i) = \{n \geq 1 \mid P^n(i, i) > 0\}.$$

Dann ist die **Periode** des Zustandes i definiert als $d_i = \gcd(T(i))$.

Nun wollen wir die Periodizität von Zuständen betrachten.

Definition 53

Für einen Zustand i definieren wir

$$T(i) = \{n \geq 1 \mid P^n(i, i) > 0\}.$$

Dann ist die **Periode** des Zustandes i definiert als $d_i = \gcd(T(i))$.

Falls eine Markovkette irreduzibel ist, ist die Periode aller Zustände gleich. Diese Periode wird dann als Periode der Markovkette bezeichnet.

Definition 54

Ein Zustand i heißt **aperiodisch**, falls $d_i = 1$

Definition 54

Ein Zustand i heißt **aperiodisch**, falls $d_i = 1$, oder equivalent, falls $\exists n_0 \in \mathbb{N}. \forall n \geq n_0. p_{ii}^{(n)} > 0$.

Definition 54

Ein Zustand i heißt **aperiodisch**, falls $d_i = 1$, oder equivalent, falls $\exists n_0 \in \mathbb{N}. \forall n \geq n_0. p_{ii}^{(n)} > 0$.

Somit ist ein Zustand i aperiodisch genau dann wenn es im Übergansdiagramm für alle $n \in \mathbb{N}$ ab einem $n_0 \in \mathbb{N}$ einen geschlossenen Weg der Länge n von i nach i gibt.

Definition 54

Ein Zustand i heißt **aperiodisch**, falls $d_i = 1$, oder equivalent, falls $\exists n_0 \in \mathbb{N}. \forall n \geq n_0. p_{ii}^{(n)} > 0$.

Somit ist ein Zustand i aperiodisch genau dann wenn es im Übergangsdiagramm für alle $n \in \mathbb{N}$ ab einem $n_0 \in \mathbb{N}$ einen geschlossenen Weg der Länge n von i nach i gibt.

Das heißt Zustand i ist sicherlich dann aperiodisch wenn er im Übergangsdiagramm

- eine Schleife besitzt ($p_{ii} > 0$)

Definition 54

Ein Zustand i heißt **aperiodisch**, falls $d_i = 1$, oder equivalent, falls $\exists n_0 \in \mathbb{N}. \forall n \geq n_0. p_{ii}^{(n)} > 0$.

Somit ist ein Zustand i aperiodisch genau dann wenn es im Übergangsdiagramm für alle $n \in \mathbb{N}$ ab einem $n_0 \in \mathbb{N}$ einen geschlossenen Weg der Länge n von i nach i gibt.

Das heißt Zustand i ist sicherlich dann aperiodisch wenn er im Übergangsdiagramm

- eine Schleife besitzt ($p_{ii} > 0$) oder
- auf mindestens zwei geschlossenen Pfaden P_1 und P_2 liegt, deren Längen teilerfremd sind.

Definition 54

Ein Zustand i heißt **aperiodisch**, falls $d_i = 1$, oder equivalent, falls $\exists n_0 \in \mathbb{N}. \forall n \geq n_0. p_{ii}^{(n)} > 0$.

Somit ist ein Zustand i aperiodisch genau dann wenn es im Übergangsdiagramm für alle $n \in \mathbb{N}$ ab einem $n_0 \in \mathbb{N}$ einen geschlossenen Weg der Länge n von i nach i gibt.

Das heißt Zustand i ist sicherlich dann aperiodisch wenn er im Übergangsdiagramm

- eine Schleife besitzt ($p_{ii} > 0$) oder
- auf mindestens zwei geschlossenen Pfaden P_1 und P_2 liegt, deren Längen teilerfremd sind.

Eine Markov-Kette heißt **aperiodisch** falls alle ihre Zustände aperiodisch sind.

Definition 55

Eine irreduzible und aperiodische Markov-Kette heißt **ergodisch**.

Definition 55

Eine irreduzible und aperiodische Markov-Kette heißt **ergodisch**.

Für jede endliche ergodische Markov-Kette gilt unabhängig der Startverteilung q_0 , dass

$$\lim_{t \rightarrow \infty} q_t = \pi$$

wobei π die eindeutige stationäre Verteilung bezeichnet.

Definition 56

Eine quadratische Matrix A heißt **stochastisch**, falls sich alle Zeilen zu Eins aufsummieren.

Definition 56

Eine quadratische Matrix A heißt **stochastisch**, falls sich alle Zeilen zu Eins aufsummieren.

Jede Übergangsmatrix P ist stochastisch.

Definition 56

Eine quadratische Matrix A heißt **stochastisch**, falls sich alle Zeilen zu Eins aufsummieren.

Jede Übergangsmatrix P ist stochastisch.

A heißt zusätzlich **doppeltstochastisch**, falls sich auch alle Spalten zu Eins aufsummieren.

Definition 56

Eine quadratische Matrix A heißt **stochastisch**, falls sich alle Zeilen zu Eins aufsummieren.

Jede Übergangsmatrix P ist stochastisch.

A heißt zusätzlich **doppeltstochastisch**, falls sich auch alle Spalten zu Eins aufsummieren.

Für jede endliche ergodische Markov-Kette mit doppelstochastischer Übergangsmatrix gilt, dass die stationäre Verteilung jedem Zustand die gleiche Wahrscheinlichkeit zuweist:

$$\pi \equiv \frac{1}{|S|}.$$