

Einführung in die angewandte Stochastik

Fabian Meyer

5. August 2024

Inhaltsverzeichnis

1	Wahrscheinlichkeitsrechnung	3
1.1	Definitionen	3
1.2	Wahrscheinlichkeitsmaß, Wahrscheinlichkeitsverteilung, Zähldichte	3
1.3	Diskreter Wahrscheinlichkeitsraum	3
1.4	Laplace-Raum	3
1.5	Träger eines diskreten Wahrscheinlichkeitsraum	3
1.6	Diskrete Wahrscheinlichkeitsverteilungen	3
1.7	Nicht-diskrete Wahrscheinlichkeitsmaße mit Riemann-Dichten	3
1.8	σ -Algebra	4
1.9	Kolmogorov-Axiome	4
1.10	Borelsche σ -Algebra	4
1.11	Riemann-Dichte, Verteilungsfunktion	4
1.12	Träger einer Riemann-Dichtefunktion	4
1.13	Ereignisfolgen, limes superior, limes inferior	5
1.14	Siebformel von Sylvester-Poincaré	5
1.15	Bedingte Wahrscheinlichkeit	5
1.16	Stochastische Unabhängigkeit	5
1.17	Produktraum	5
2	Zufallsvariablen und Wahrscheinlichkeitsmaße	6
2.1	Indikatorfunktion	6
2.2	Stochastische Unabhängigkeit von Zufallsvariablen	6
2.3	Summe unabhängiger Zufallsvariablen, Faltung	6
2.4	Quantilfunktion	6
2.5	Multivariate/Mehrdimensionale Verteilungsfunktion	7
2.6	Randverteilung, Marginalverteilung, Randdichte	7
2.7	Erwartungswerte	7
2.8	Moment	8
2.9	k-tes Moment, Varianz, Kovarianz	8
2.10	Verschiebungssatz von Steiner	8

2.11	Unkorreliertheit, Korrelationskoeffizient	8
2.12	Ungleichungen mit Momenten	9
2.13	Erwartungswertvektor, Kovarianzmatrix	9
2.14	Erzeugende Funktion	9
2.15	Bedingte Verteilungen	10
2.16	Bedingter Erwartungswert, bedingte Varianz	10
2.17	Bedingte Erwartung	10
2.18	Eine Version des Schwachen Gesetzes großer Zahlen	10
2.19	1. Version des starken Gesetzes großer Zahlen	10
2.20	2. Version des starken Gesetzes großer Zahlen	10
2.21	Eine Version des zentralen Grenzwertsatzes	11

1 Wahrscheinlichkeitsrechnung

1.1 Definitionen

- Grundraum Ω (Grundmenge, Ergebnisraum)- Menge aller möglichen Ergebnisse eines Zufallsexperiments
- $\omega \in \Omega$ Ergebnis.
- Ereignis A (\mathcal{B}, \dots)- Menge von Ergebnissen. Ein Ereignis, das genau ein Element besitzt heißt Elementarereignis

1.2 Wahrscheinlichkeitsmaß, Wahrscheinlichkeitsverteilung, Zähldichte

Sei $\mathfrak{P} = \text{Pot}(\Omega)$ (Menge aller Ereignisse über Ω) und $p : \Omega \rightarrow [0, 1]$ Abbildung mit $\sum_{\omega \in \Omega} p(\omega) = 1$. Dann ist

$$P(A) = \sum_{\omega \in A} p(\omega), A \in \mathfrak{P}$$

Wahrscheinlichkeitsmaß/Wahrscheinlichkeitsverteilung auf \mathfrak{P} (oder Ω).

1.3 Diskreter Wahrscheinlichkeitsraum

Wenn $|\Omega|$ höchstens abzählbar unendlich dann ist (Ω, P) diskreter Wahrscheinlichkeitsraum.

1.4 Laplace-Raum

$$P(A) = \frac{|A|}{|\Omega|}$$

Ziehen aus Urne ohne Wiederholung ohne Reihenfolge kein Laplace-Raum.

1.5 Träger eines diskreten Wahrscheinlichkeitsraum

$$T = \{\omega \in \Omega \mid P(\omega) > 0\}$$

1.6 Diskrete Wahrscheinlichkeitsverteilungen

Zur Festlegung der Verteilung wird jedem Element von $T = \{x_1, \dots\}$ Wahrscheinlichkeit $p_k \in [0, 1)$ zugewiesen mit $\sum_k p_k = 1$

1.7 Nicht-diskrete Wahrscheinlichkeitsmaße mit Riemann-Dichten

- jedem Intervall wird eine Wahrscheinlichkeit zugewiesen
- statt Potenzmenge neues Mengensystem σ -Algebra

1.8 σ -Algebra

$\Omega \neq \emptyset$. $\mathfrak{P} \subset \text{Pot}(\Omega)$ heißt σ -Algebra von Ereignissen über Ω , falls:

1. $\Omega \in \mathfrak{P}$
2. $A \in \mathfrak{P} \Rightarrow A^c \in \mathfrak{P}$
3. Für jede Folge A_1, A_2, \dots in \mathfrak{P} gilt: $\cup_{n=1}^{\infty} A_n \in \mathfrak{P}$

1.9 Kolmogorov-Axiome

Sei $P : \mathfrak{P} \rightarrow [0, 1]$ mit

- $P(A) \geq 0$, für $A \in \mathfrak{P}$
- $P(\Omega) = 1$
- $P(\cup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n)$, für paarweise diskunkte Mengen

Dann heißt P Wahrscheinlichkeitsmaß/Wahrscheinlichkeitsverteilung auf Ω . $(\Omega, \mathfrak{P}, P)$ heißt Wahrscheinlichkeitsraum, (Ω, \mathfrak{P}) heißt messbarer Raum oder Messraum.

1.10 Borelsche σ -Algebra

Die kleinstmögliche σ -Algebra über einem Intervall (hier z.B. $[a, b]$), welche alle Teilmengen des Intervalls enthält heißt Borelsche σ -Algebra \mathcal{B}^1 . Für $\Omega \subset \mathbb{R}^n$ entsprechend \mathcal{B}^n .

1.11 Riemann-Dichte, Verteilungsfunktion

Sei $f : \mathbb{R} \rightarrow \mathbb{R}$ mit $f(x) \geq 0, x \in \mathbb{R}$ und $\int_{-\infty}^{\infty} f(x)dx = 1$ integrierbar, dann heißt f Riemann-Dichtefunktion. Wahrscheinlichkeitsmaß festgelegt über

$$F(x) = P((-\infty, x]) = \int_{-\infty}^x f(y)dy$$

F ist Verteilungsfunktion. Für Verteilungsfunktionen siehe Formelsammlung.

1.12 Träger einer Riemann-Dichtefunktion

Das größtmögliche Intervall I mit $f(x) > 0, x \in I$ heißt Träger der zugehörigen Verteilungsfunktion.

1.13 Ereignisfolgen, limes superior, limes inferior

$(A_n)_n$ in \mathfrak{P} heißt isoton (monoton wachsend), falls $A_n \subset A_{n+1}, \forall n \in \mathbb{N}$ oder antiton (monoton fallend), falls andersrum. Für eine Ereignisfolge ist:

$$\limsup_{n \rightarrow \infty} A_n = \lim_{n \rightarrow \infty} \left(\bigcup_{k=n}^{\infty} A_k \right) = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$$

$$\liminf_{n \rightarrow \infty} A_n = \lim_{n \rightarrow \infty} \left(\bigcap_{k=n}^{\infty} A_k \right) = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k$$

Es gilt:

$$\limsup_{n \rightarrow \infty} A_n = \{ \omega \in \Omega \mid \omega \text{ liegt in unendlich vielen der } A_i \}$$

$$\liminf_{n \rightarrow \infty} A_n = \{ \omega \in \Omega \mid \omega \text{ liegt in allen } A_i \text{ bis auf endlich viele} \}$$

1.14 Siebformel von Sylvester-Poincaré

$$P\left(\bigcup_{k=1}^n A_k\right) = \sum_{k=1}^n P(A_k) - \sum_{1 \leq i_1 < i_2 \leq n} P(A_{i_1} \cap A_{i_2}) + \sum_{1 \leq i_1 < i_2 < i_3 \leq n} P(A_{i_1} \cap A_{i_2} \cap A_{i_3}) - \dots$$

1.15 Bedingte Wahrscheinlichkeit

$P(A|B) = \frac{P(A \cap B)}{P(B)}$ heißt elementar bedingte Wahrscheinlichkeit von A und B . $P(A|B)$ bildet wiederum eine Wahrscheinlichkeitsverteilung und $(\Omega, \mathfrak{P}, P(\cdot|B))$ ist ein Wahrscheinlichkeitsraum. Es gilt:

- $P(A|B) = P(B|A) \cdot \frac{P(A)}{P(B)}$
- $P\left(\bigcap_{i=1}^n A_i\right) = P(A_1) \cdot P(A_2 | A_1) \cdot P(A_3 | A_2 \cap A_1)$

1.16 Stochastische Unabhängigkeit

A_1 und A_2 heißen paarweise stochastisch unabhängig, falls $P(A_i \cup A_j) = P(A_1) \cdot P(A_2)$. A_1, A_2, \dots heißen (gemeinsam) stochastisch unabhängig, falls für jede endliche Auswahl von Indizes $\{i_1, \dots, i_s\}$ gilt: $P(A_{i_1} \cap \dots \cap A_{i_s}) = P(A_{i_1}) \cdot \dots \cdot P(A_{i_s})$.

1.17 Produktraum

Für diskrete Wahrscheinlichkeitsräume $(\Omega_i, \mathfrak{P}_i, P_i)$ heißt $(\Omega, \mathfrak{P}, P)$ mit $\Omega = \{(\omega_1, \dots, \omega_n \mid \omega_i \in \Omega_i)\}$, $\mathfrak{P} = \text{Pot}(\Omega)$ und $P(\{\omega\}) = \prod_{i=1}^n P_i(\{\omega_i\})$, $\omega = (\omega_1, \dots, \omega_n) \in \Omega$ heißt Produktraum.

$$(\Omega, \mathfrak{P}, P) = (\Omega_1, \mathfrak{P}_1, P_1) \times \dots \times (\Omega_n, \mathfrak{P}_n, P_n)$$

2 Zufallsvariablen und Wahrscheinlichkeitsmaße

Zufallsvorgänge hier wieder beschrieben durch den Wahrscheinlichkeitsraum $(\Omega, \mathfrak{P}, P)$, wobei der Ausgang des Vorgangs $\omega \in \Omega$ ist. Dabei ist häufig nicht ω von Interesse sondern ein Funktionswert $X(\omega)$, wobei X eine Abbildung auf $X : \omega \rightarrow \mathbb{R}^n$ ist.

$$P(X = k) = P(\{\omega \in \Omega \mid X(\omega) = k\}) = P^X(\{k\})$$

und

$$P^X : \mathcal{B} \rightarrow [0, 1]$$

X heißt Zufallsvariable (falls $n = 1$), sonst Zufallsvektor. Eine Zufallsvariable erzeugt im Wertebereich der Funktion eine neue Wahrscheinlichkeitsverteilung.

2.1 Indikatorfunktion

$\mathfrak{I}_A : \Omega \rightarrow \mathbb{R} = 1$, falls $\omega \in A$, sonst 0, heißt Indikatorfunktion von A . Und es gilt \mathfrak{I}_A ist $bin(1, p)$ verteilt, $p = P(A)$.

2.2 Stochastische Unabhängigkeit von Zufallsvariablen

Die Zufallsvariablen $X_i : (\Omega, \mathfrak{P}, P) \rightarrow (\Omega_i, \mathfrak{P}_i, P^{X_i})$ heißen stochastisch unabhängig, falls

$$P\left(\bigcap_i \{X_i \in A_i\}\right) = \prod_i P(X_i \in A_i), \forall A_i \in \mathfrak{P}_i$$

2.3 Summe unabhängiger Zufallsvariablen, Faltung

Für stochastisch unabhängige Zufallsvariablen X, Y auf \mathbb{Z} mit Zähldichten f, g gilt, die Zähldichte von $X + Y = h$:

$$h(k) = \sum_{j \in \mathbb{Z}} f(j) \cdot g(k - j) = \sum_{j \in \mathbb{Z}} f(k - j) \cdot g(j) = P(X + Y = k)$$

$h = f * g$ ist Faltung der Dichten f und g . In \mathbb{R} ergibt sich die Zähldichte als Integral statt als Summe über obige Funktion.

2.4 Quantilfunktion

Die Umkehrfunktion $F^{-1}(y) = \inf\{x \in \mathbb{R} \mid F(x) \geq y\}$ der Verteilungsfunktion F heißt Quantilfunktion oder Pseudoinverse von F (existiert nur, falls F bijektiv, also streng monoton (wachsend)).

2.5 Multivariate/Mehrdimensionale Verteilungsfunktion

Sei $X = (X_1, \dots, X_n)$ ein n -dimensionaler Zufallsvektor. Dann ist die multivariate/mehrdimensionale Verteilungsfunktion:

$$F^X(x) = P(X_1 \in (-\infty, x_1], \dots, X_n \in (-\infty, x_n]) = P(X_1 \leq x_1, \dots, X_n \leq x_n)$$

mit

$$x = (x_1, \dots, x_n) \in \mathbb{R}^n$$

2.6 Randverteilung, Marginalverteilung, Randdichte

Sei $X = (X_1, \dots, X_n)$ ein Zufallsvektor und $m < n$. Dann heißt $(X_{i_1}, \dots, X_{i_m})$ Rand- oder Marginalverteilung zu (i_1, \dots, i_m) . Die Randverteilung wird bestimmt, indem man in die nicht benötigten Komponenten \mathbb{R} einsetzt. Die i -te Randdichte wird wie folgt bestimmt:

$$f^{X_i}(t) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f^X(x_1, \dots, x_{i-1}, t, x_{i+1}, \dots, x_n) dx_1 \dots dx_{i-1} dx_{i+1} \dots dx_n$$

X_1, \dots, X_n sind genau dann stochastisch unabhängige Zufallsvariablen, wenn

$$f^{(X_1, \dots, X_n)}(x_1, \dots, x_n) = \prod_{i=1}^n f^{X_i}(x_i)$$

Zudem sind (X_1, X_2) genau dann stochastisch unabhängig, wenn sie normalverteilt sind, mit Parameter $\rho = 0$.

2.7 Erwartungswerte

Sei X eine Zufallsvariable mit Zähldichte p oder Riemann-Dichte f . Dann gilt für den Erwartungswert EX :

1. Sei $X(\Omega) \subset [0, \infty)$ oder $X(\Omega) \subset (-\infty, 0]$.

a) $EX = E(X) = \sum_{x \in X(\Omega)} xp(x)$, oder

b) $EX = E(X) = \int_{-\infty}^{\infty} xf(x)dx$

2. Ist $E(\max(X, 0)) < \infty$, oder $E(\min(X, 0)) > -\infty$, dann heißt EX wie in 1 Erwartungswert von X (unter P)

Für Zufallsvariablen mit Werten in \mathbb{N}_0 gilt auch:

$$EX = \sum_{n=1}^{\infty} P(X \geq n)$$

2.8 Moment

Als (allgemeines Moment) bezeichnet man den Erwartungswert einer Funktion $g(x) = E(g(x))$. Um ihn zu berechnen, berechnet man die Wahrscheinlichkeit des Auftretens jedes Ereignisses und multipliziert dies mit dem Wert von g an dieser Stelle also:

$$E(g(x)) = \sum_{(t_1, \dots, t_k) \in \text{supp}(P^X)} P^X((t_1, \dots, t_k)) \cdot g(t_1, \dots, t_k)$$

bzw.

$$E(g(x)) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f^X(t_1, \dots, t_k) \cdot g(t_1, \dots, t_k) dt_1 \dots dt_k$$

Für stochastisch unabhängige Zufallsvariablen mit (endlichen) Erwartungswerten gilt:

$$E\left(\prod_{i=1}^n X_i\right) = \prod_{i=1}^n E(X_i)$$

2.9 k-tes Moment, Varianz, Kovarianz

1. $m_k(c) = E((X-c)^k)$ heißt k -tes Moment von X um c (unter P). (zentrales Moment, falls $c = 0$)
2. $\text{Var} X = E((X - EX)^2)$ ist die Varianz (Streuung) von X
3. $\text{Kov}(X, Y) = E((X - EX)(Y - EY))$ ist die Kovarianz von X und Y

Zudem ist:

$$\text{Var}\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \text{Var} X_i + 2 * \sum_{1 \leq i < j \leq n} \text{Kov}(X_i, X_j)$$

und $\text{Kov}(X, Y) = 0$, falls X, Y stochastisch unabhängig.

2.10 Verschiebungssatz von Steiner

$$E((X - a)^2) = \text{Var} X + (EX - a)^2$$

2.11 Unkorreliertheit, Korrelationskoeffizient

1. X, Y heißen unkorreliert, falls $\text{Kov}(X, Y) = 0$
2. Der Korrelationskoeffizient ist $\text{Korr}(X, Y) = \frac{\text{Kov}(X, Y)}{\sqrt{\text{Var} X} \cdot \sqrt{\text{Var} Y}} \in [-1, 1]$

Für unkorrelierte Zufallsvariablen gilt:

$$\text{Var}\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \text{Var}(X_i)$$

Stochastisch unabhängige Variablen sind unkorreliert. (aber nicht andersrum (außer bei Normalverteilungen)).

2.12 Ungleichungen mit Momenten

1. Ungleichung von Jensen: Sei $h : \mathbb{R} \rightarrow \mathbb{R}$ eine konvexe Funktion (linksgekrümmt), und $E(h(X))$ und EX existieren und sind endlich. Dann ist $E(h(X)) \geq h(EX)$ (bzw. andersrum im konkaven Fall)
2. Ungleichung von Markov: Sei $g : [0, \infty) \rightarrow [0, \infty)$ monoton wachsend. Dann ist:

$$P(|X| > \epsilon) \leq P(|X| \geq \epsilon) \leq \frac{1}{g(\epsilon)} E(g(|X|)), \epsilon > 0$$

3. Ungleichung von Tschebyscheff:

$$P(|X - EX| \geq \epsilon) \leq \frac{\text{Var} X}{\epsilon^2}$$

2.13 Erwartungswertvektor, Kovarianzmatrix

Hier X nicht Zufallsvariable, sondern Zufallsvektor.

1. $E(X) = (EX_1, \dots, EX_n)$ ist der Erwartungsvektor von X
2. $Kov(X)$ Kovarianzmatrix von X mit $Kov(X)_{i,j} := Kov(X_i, X_j)$

2.14 Erzeugende Funktion

$g(t) = Et^X$ (für alle t , für die der Erwartungswert endlich existiert) heißt (wahrscheinlichkeits)erzeugende Funktion von X (bzw. von P^X).

Ist $(0, 1 + \epsilon) \subset K$ für ein $\epsilon > 0$, so existieren alle Momente $EX^k, k \in \mathbb{N}$ und es gilt:

$$g^{(k)}(1) = E\left(\prod_{i=0}^{k-1} (X - i)\right), k \in \mathbb{N}$$

$$g'(1) = EX$$

K ist hier der Konvergenzbereich von $\sum_{k=0}^{\infty} t^k p_k = Et^X$. Und X eine Zufallsvariable mit diskreter Wahrscheinlichkeitsverteilung auf \mathbb{N}_0 .

Falls X, Y stochastisch unabhängig mit diskreten Wahrscheinlichkeitsverteilungen auf \mathbb{N}_0 sind gilt $Et^{X+Y} = Et^X \cdot Et^Y$. h mit:

$$h(t) = Ee^{tX}$$

heißt momenterzeugende Funktion von X . Existiere $h(t)$ für $t \in (-\epsilon, \epsilon), \epsilon > 0$. Dann gilt:

1. h bestimmt die zugrundeliegende Wahrscheinlichkeitsverteilung eindeutig.
2. Es existieren alle absoluten Momente $E(|X|^k)$, $k \in \mathbb{N}$ endlich.
3. h ist im Nullpunkt beliebig oft differenzierbar, und es gilt: $h^{(k)}(0) = EX^k$, $k \in \mathbb{N}$

2.15 Bedingte Verteilungen

Sei (X, Y) diskret verteilter Zufallsvektor. Dann heißt

$$p^{Y|X}(y | x) = p^{Y|X=x}(y) = P(Y = y | X = x) = \begin{cases} \frac{p^{(X,Y)}(x,y)}{p^X(x)}, & p^X(x) > 0 \\ p^Y(y), & p^X(x) = 0 \end{cases}$$

bedingte Wahrscheinlichkeitsverteilung von Y unter (der Hypothese) $X = x$ und $p^{Y|X}$ heißt bedingte Zähldichte von Y unter X . Für Dichte ist die Definition analog.

2.16 Bedingter Erwartungswert, bedingte Varianz

Erwartungswerte definiert wie im nicht bedingten Fall, als Wahrscheinlichkeit wird jedoch $p^{Y|X=x}(y)$ bzw. $f^{Y|X=x}(y)$ betrachtet.

2.17 Bedingte Erwartung

Analog wird die Bedingte Erwartung von Y unter X $E(Y | X)$ definiert.

2.18 Eine Version des Schwachen Gesetzes großer Zahlen

Seien X_1, \dots paarweise unkorrelierte Zufallsvariablen mit $EX_i = \mu$ und $Var X_i \leq M < \infty$ für Konstante $M > 0$. Dann ist:

$$P \left(\left| \frac{1}{n} \sum_{i=1}^n X_i - \mu \right| \geq \epsilon \right) \leq \frac{M}{n\epsilon^2} \xrightarrow{n \rightarrow \infty} 0, \epsilon > 0$$

2.19 1. Version des starken Gesetzes großer Zahlen

Seien X_1, \dots stochastisch unabhängige Zufallsvariablen auf einem Wahrscheinlichkeitsraum mit endlichen Varianzen und es gelte $\sum_{n=1}^{\infty} \frac{Var X_n}{n^2} < \infty$. Dann ist:

$$P \left(\left\{ \omega \in \Omega \mid \frac{1}{n} \sum_{i=1}^n X_i - \frac{1}{n} \sum_{i=1}^n EX_i \xrightarrow{n \rightarrow \infty} 0 \right\} \right) = 1$$

2.20 2. Version des starken Gesetzes großer Zahlen

Seien X_1, \dots stochastisch unabhängig und identisch verteilt, mit $EX_1 = \mu$. Dann ist:

$$\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{n \rightarrow \infty} \mu$$

2.21 Eine Version des zentralen Grenzwertsatzes

Scheint mir nicht in der Klausur vorzukommen, da Terme zu lang. Außerdem habe ich dann eine Ausrede um das mir nicht angucken zu müssen.