FORMELSAMMLUNG
für
Wahrscheinlichkeit und Statistik
FS 2021

1 Definitionen

1. **Wahrscheinlichkeitsraum**

   Ein Wahrscheinlichkeitsraum ist ein Tripel $(\Omega, \mathcal{A}, \mathbb{P})$, wobei

   (a) Ω ist eine beliebige nichtleere Menge.

   (b) $\mathcal{A}$ ist eine σ-Algebra auf Ω:

      i. $\Omega \in \mathcal{A}$

      ii. $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$

      iii. $A_1, A_2, \ldots \in \mathcal{A} \Rightarrow \cup_{n=1}^{\infty} A_n \in \mathcal{A}$

   (c) $\mathbb{P}\colon \mathcal{A} \to [0,1]$ ist ein Wahrscheinlichkeitsmass auf $(\Omega, \mathcal{A})$:

      i. $P(\Omega) = 1$

      ii. Für alle $A_1, A_2, \ldots \in \mathcal{A}$ mit $\forall i \neq j\colon A_i \cap A_j = \emptyset$ gilt $\mathbb{P}(\cup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mathbb{P}(A_n)$

2. **Zufallsvariable**

   Eine Zufallsvariable X ist eine $\mathcal{B}(\mathbb{R})/\mathcal{A}$-messbare Abbildung $X\colon \Omega \to \mathbb{R}$.

3. **Verteilung und Verteilungsfunktion**

   Die Verteilung einer Zufallsvariable X ist das Wahrscheinlichkeitsmass $\mu_X : \mathcal{B}(\mathbb{R}) \to [0,1]$ mit

   $$\forall\, B \in \mathcal{B}(\mathbb{R}): \quad \mu_X(B) = \mathbb{P}(X \in B).$$

   Die Verteilungsfunktion (VF) von X ist die Abbildung $F_X\colon \mathbb{R} \to [0,1]$ mit

   $$\forall\, x \in \mathbb{R}: \quad F_X(x) = \mathbb{P}(X \leq x) = \mu_X((-\infty, x]).$$

4. **Erwartungswert**

   Der Erwartungswert einer Zufallsvariable X, sofern er existiert, ist gegeben durch

   $$\begin{aligned} \mathbb{E}[X] &= \int_\Omega X(\omega)\mathbb{P}(d\omega) = \int_\mathbb{R} x\mu_X(dx) \\ &= \begin{cases} \sum_x x\mathbb{P}(X=x), & X \text{ diskret verteilt} \\ \int_\mathbb{R} x f_X(x)dx, & X \text{ absolut stetig verteilt} \end{cases}. \end{aligned}$$

5. **Varianz**

   Die Varianz einer Zufallsvariable X mit $\mathbb{E}[|X|] < \infty$ ist

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

6. **Kovarianz**

   Die Kovarianz von zwei Zufallsvariablen X, Y mit $\mathbb{E}[|X|] < \infty$, $\mathbb{E}[|Y|] < \infty$, $\mathbb{E}[|XY|] < \infty$ ist

   $$\begin{aligned} \mathrm{Cov}(X,Y) &= \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] \\ &= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]. \end{aligned}$$

7. **Korrelation**

   Die Korrelation von zwei Zufallsvariablen X, Y mit $0 < \mathrm{Var}(X) < \infty$, $0 < \mathrm{Var}(Y) < \infty$ ist

   $$\mathrm{corr}(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}.$$

8. **Momenterzeugende Funktion**

   Die momenterzeugende Funktion einer reellwertigen Zufallsvariablen X ist definiert durch

   $$\begin{aligned} M_X &: \mathbb{R} \to [0,\infty] \\ &t \mapsto M_X(t) = \mathbb{E}[e^{tX}]. \end{aligned}$$

9. **Charakteristische Funktion**

   Die charakteristische Funktion einer reellwertigen Zufallsvariable X ist definiert durch

   $$\begin{aligned} \varphi_X &: \mathbb{R} \to \mathbb{C} \\ &t \mapsto \varphi_X(t) = \mathbb{E}[e^{itX}]. \end{aligned}$$

2 Einige diskrete Verteilungen

1. **Bernoulli-Verteilung**

   $X \sim \mathrm{Ber}(p)$ für $p \in [0,1]$:

   $$\mathbb{P}(X = x) = \begin{cases} p, & x = 1 \\ 1-p, & x = 0 \end{cases}$$

   $\mathbb{E}[X] = p$, $\mathrm{Var}(X) = p(1-p)$.

2. **Binomialverteilung**

   $X \sim \mathrm{Bin}(n,p)$ für $n \in \mathbb{N}$, $p \in [0,1]$:

   $$\forall x \in \{0,1,\ldots,n\}: \quad \mathbb{P}(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$$

   $\mathbb{E}[X] = np$, $\mathrm{Var}(X) = np(1-p)$.

3. **Geometrische Verteilung**

   $X \sim \mathrm{Geo}(p)$ für $p \in (0,1)$:

   $$\forall x \in \mathbb{N}: \quad \mathbb{P}(X = x) = (1-p)^{x-1} p$$

   $\mathbb{E}[X] = \frac{1}{p}$, $\mathrm{Var}(X) = \frac{1-p}{p^2}$.

4. **Poisson-Verteilung**

   $X \sim \mathrm{Poi}(\lambda)$ für $\lambda > 0$:

   $$\forall x \in \mathbb{N}_0: \quad \mathbb{P}(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

   $\mathbb{E}[X] = \lambda$, $\mathrm{Var}(X) = \lambda$.

5. **Negative Binomialverteilung**

   $X \sim \mathrm{NB}(r,p)$ für $p \in (0,1), r \in \mathbb{N}$:

$$\forall x \in \{0,1,2,\dots\}: \quad \mathbb{P}(X=x) = \binom{x+r-1}{x} p^r (1-p)^x$$

   $\mathbb{E}[X] = \frac{r(1-p)}{p}$, $\mathrm{Var}(X) = \frac{r(1-p)}{p^2}$.

6. **Hypergeometrische Verteilung**

   $X \sim \mathrm{Hypergeom}(n,N,K)$ für $n,N,K \in \mathbb{N}$ mit $\max(n,K) \leq N$:

$$\forall x \in \{\max(0, n-N+K), ..., \min(n,K)\}: \quad \mathbb{P}(X=x) = \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}$$

   $\mathbb{E}[X] = n\frac{K}{N}$, $\mathrm{Var}(X) = n\frac{K}{N}\left(1-\frac{K}{N}\right)\frac{N-n}{N-1}$.

3 Einige stetige Verteilungen

1. **Gleichverteilung**

   $X \sim \mathcal{U}(a,b)$ für $a < b$:

$$\forall x \in \mathbb{R}: \quad f_X(x) = \frac{1}{b-a}\mathbb{1}_{[a,b]}(x)$$

   $\mathbb{E}[X] = \frac{a+b}{2}$, $\mathrm{Var}(X) = \frac{(b-a)^2}{12}$.

2. **Exponentialverteilung**

   $X \sim \mathrm{Exp}(\lambda)$ für $\lambda > 0$:

$$\forall x \in \mathbb{R}: \quad f_X(x) = \lambda e^{-\lambda x}\mathbb{1}_{(0,\infty)}(x)$$

   $\mathbb{E}[X] = \frac{1}{\lambda}$, $\mathrm{Var}(X) = \frac{1}{\lambda^2}$.

3. **Gammaverteilung**

   $X \sim \Gamma(\alpha,\beta)$ für $\alpha,\beta > 0$:

$$\forall x \in \mathbb{R}: \quad f_X(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}\mathbb{1}_{(0,\infty)}(x),$$

   $\mathbb{E}[X] = \frac{\alpha}{\beta}$, $\mathrm{Var}(X) = \frac{\alpha}{\beta^2}$.

4. **Normalverteilung**

   $X \sim \mathcal{N}(\mu,\sigma^2)$ für $\mu \in \mathbb{R}, \sigma > 0$:

$$\forall x \in \mathbb{R}: \quad f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

   $\mathbb{E}[X] = \mu$, $\mathrm{Var}(X) = \sigma^2$.

   Für $\mu = 0$ und $\sigma = 1$ nennen wir X standardnormalverteilt und bezeichnen die Verteilungsfunktion (VF) mit Φ.

4 Unabhängigkeit und Satz von Bayes

1. **Unabhängigkeit**

   Sei I eine beliebige Menge und $(A_i)_{i\in I} \subseteq \mathcal{A}$ eine Kollektion von Ereignissen. Dann heisst $(A_i)_{i\in I}$ unabhängig, falls

$$\forall J \subseteq I \text{ endlich}: \quad \mathbb{P}(\cap_{j\in J} A_j) = \prod_{j\in J}\mathbb{P}(A_j).$$

2. **Bedingte Wahrscheinlichkeit**

   Seien $A, B \in \mathcal{A}$ mit $\mathbb{P}(B) > 0$. Die bedingte Wahrscheinlichkeit von A gegeben B ist

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

3. **Satz der totalen Wahrscheinlichkeit**

   Sei $(B_i)_{i\in I} \subseteq \mathcal{A}$ so, dass $\Omega = \cup_{i\in I} B_i$ und $\forall\, i \neq j$: $B_i \cap B_j = \emptyset$. Dann gilt

$$\forall A \in \mathcal{A}: \quad \mathbb{P}(A) = \sum_{i\in I:\, \mathbb{P}(B_i)>0} \mathbb{P}(A|B_i)\mathbb{P}(B_i).$$

4. **Satz von Bayes**

   Seien $A, B \in \mathcal{A}$ mit $\mathbb{P}(A) > 0$. Es gilt

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

5 Borel-Cantelli

Seien $A_1, A_2, \ldots \in \mathcal{A}$ Ereignisse und sei

$$A_\infty = \bigcap_{n=1}^\infty \bigcup_{m=n}^\infty A_m.$$

1. Aus $\sum_{n=1}^\infty \mathbb{P}(A_n) < \infty$ folgt $\mathbb{P}(A_\infty) = 0$.
2. Angenommen $(A_n)_{n=1}^\infty$ sind paarweise unabhängig und erfüllen $\sum_{n=1}^\infty \mathbb{P}(A_n) = \infty$. Dann ist $\mathbb{P}(A_\infty) = 1$.

6 Grenzwertsätze

1. **Konvergenz von Zufallsvariablen**

   (a) **Konvergenz in Wahrscheinlichkeit**

   Seien $X, X_1, X_2, \ldots$ Zufallsvariablen. Wir sagen, dass $(X_n)_{n=1}^\infty$ in Wahrscheinlichkeit gegen X konvergiert, falls

$$\forall\, \varepsilon > 0: \quad \lim_{n\to\infty} \mathbb{P}(|X_n - X| > \varepsilon) = 0.$$

   (b) **Fast sichere Konvergenz**

   Seien $X, X_1, X_2, \ldots$ Zufallsvariablen. Wir sagen, dass $(X_n)_{n=1}^\infty$ $\mathbb{P}$-fast sicher ($\mathbb{P}$-f.s.) gegen X konvergiert falls

$$\mathbb{P}(\{\omega \in \Omega\colon \lim_{n\to\infty} |X_n(\omega) - X(\omega)| = 0\}) = 1.$$

   (c) **Konvergenz in Verteilung**

   Seien $\mu, \mu_1, \mu_2, \ldots$ Verteilungen auf $\mathbb{R}$ und $F, F_1, F_2, \ldots$ die entsprechenden Verteilungsfunktionen. Wir sagen, dass $(\mu_n)_{n=1}^\infty$ schwach (in Verteilung) gegen μ konvergiert, falls eine der folgenden zwei äquivalenten Aussagen zutrifft:

   i. $\forall\, f \in C_b(\mathbb{R})$: $\lim_{n\to\infty} \int_{\mathbb{R}} f d\mu_n = \int_{\mathbb{R}} f d\mu$.
   ii. $\forall\, x \in \mathbb{R}, F$ stetig an x: $\lim_{n\to\infty} F_n(x) = F(x)$.

   (d) **Implikationen**

   Fast sichere Konvergenz $\Rightarrow$ Konvergenz in Wahrscheinlichkeit $\Rightarrow$ Konvergenz in Verteilung.

2. **Starkes Gesetz der grossen Zahlen**

   Seien $X_1, X_2, \ldots$ i.i.d. mit $\mathbb{E}[|X_1|^2] < \infty$. Sei $S_n = \sum_{i=1}^n X_i$ für alle $n \in \mathbb{N}$. Dann konvergiert S_n/n fast sicher gegen $\mathbb{E}[X_1]$.

3. **Zentraler Grenzwertsatz**

   Seien $X_1, X_2, \dots$ i.i.d. mit $\mathbb{E}[|X_1|^2] < \infty$. Seien $\mu = \mathbb{E}[X_1]$ und $\sigma^2 = \mathrm{Var}(X_1)$, und sei $S_n^* = \frac{\sum_{i=1}^n X_i - n\mu}{\sigma\sqrt{n}}$ für alle $n \in \mathbb{N}$. Dann konvergiert die Verteilung von S_n^* schwach gegen die Standardnormalverteilung $\mathcal{N}(0,1)$:

   $$\forall x \in \mathbb{R}: \quad \lim_{n\to\infty} \mathbb{P}(S_n^* \le z) = \Phi(z).$$

7 Einige Ungleichungen

1. **Markov-Ungleichung**

   Seien X eine Zufallsvariable und $g\colon \mathbb{R} \to [0,\infty)$ monoton wachsend. Für jedes c mit $g(c) > 0$ gilt dann

   $$\mathbb{P}(X \ge c) \le \frac{\mathbb{E}[g(X)]}{g(c)}.$$

2. **Chebyshev-Ungleichung**

   Sei X eine Zufallsvariable mit $\mathbb{E}[|X|^2] < \infty$. Für alle $c > 0$ gilt

   $$\mathbb{P}(|X - \mathbb{E}[X]| \ge c) \le \frac{\mathrm{Var}(X)}{c^2}.$$

3. **Cauchy–Schwarz-Ungleichung**

   Seien X, Y Zufallsvariablen mit $\mathbb{E}[X^2] < \infty$, $\mathbb{E}[Y^2] < \infty$. Dann gilt

   $$|\mathbb{E}[XY]| \le \sqrt{\mathbb{E}[X^2]\mathbb{E}[Y^2]}.$$

4. **Jensen-Ungleichung**

   Seien X eine Zufallsvariable mit $\mathbb{E}[|X|] < \infty$ und $g\colon \mathbb{R} \to \mathbb{R}$ eine konvexe Funktion. Dann gilt

   $$\mathbb{E}[g(X)] \ge g(\mathbb{E}[X]).$$

   Ist g strikt konvex, so gilt Gleichheit genau dann, wenn $\mathbb{P}(X = \mathbb{E}[X]) = 1$.

8 Punktschätzer

Sei $X = (X_1, X_2, \dots, X_n)$ ein $\mathbb{R}^n$-wertiger Zufallsvektor. Sei Θ eine Menge und sei $(\mu_\theta)_{\theta\in\Theta}$ Familie von möglichen Verteilungen von X. Sei $g\colon \Theta \to \mathbb{R}^k$ eine Abbildung. Wir sind daran interessiert $\eta = g(\theta)$ anhand von X zu schätzen.

1. **Definition**

   Ein Punktschätzer für $g(\theta)$ ist eine Abbildung

   $$T\colon \mathbb{R}^n \to \mathbb{R}^k.$$

2. **Erwartungstreue**

   Der Schätzer T heisst erwartungstreu für $g(\theta)$ falls

   $$\forall\, \theta \in \Theta\colon \mathbb{E}_\theta[T(X)] = g(\theta).$$

3. **Konsistenz**

   Wir nehmen nun n als Variabel an und betrachten eine Folge $T_n\colon \mathbb{R}^n \to \mathbb{R}$, $n \in \mathbb{N}$, von Punktschätzern für $g(\theta) \in \mathbb{R}$. Die Folge $(T_n)_{n\in\mathbb{N}}$ heisst konsistent für $g(\theta)$ falls

   $$\forall\, \theta \in \Theta, \varepsilon > 0\colon \limsup_{n\to\infty} \mathbb{P}_\theta(|T_n(X_1,\dots,X_n) - g(\theta)| > \varepsilon) = 0.$$

9 Statistische Tests

Wir betrachten eine Nullhypothese $\Theta_0 \subseteq \Theta$ und eine Alternative $\Theta_A \subseteq \Theta_0^c$.

1. **Definition**

   Ein statistischer Test ist eine Abbildung $\varphi\colon \mathbb{R}^n \to \{0,1\}$. Für $x \in \mathbb{R}^n$ intepretiert man:

   - $\varphi(x) = 0 \Rightarrow$ "Die Nullhypothese wird beibehalten"
   - $\varphi(x) = 1 \Rightarrow$ "Die Nullhypothese wird abgelehnt"

   Ein randomisierter Test ist eine Abbildung $\varphi\colon \mathbb{R}^n \to [0,1]$. Für $x \in \mathbb{R}^n$ intepretiert man $\varphi(x)$ als die Wahrscheinlichkeit die Nullhypothese abzulehnen und $(1 - \varphi(x))$ als die Wahrscheinlichkeit die Nullhypothese beizubehalten.

2. **Fehler 1. und 2. Art**

   - Fehler 1. Art: Die Nullhypothese wird abgelehnt, obwohl sie richtig ist.
   - Fehler 2. Art: Die Nullhypothese wird beibehalten, obwohl sie falsch ist.

3. **Niveau und Macht eines Tests**

   Sei $\alpha \in [0,1]$. Ein Test $\varphi\colon \mathbb{R}^n \to \{0,1\}$ hat Niveau α für die Nullhypothese Θ_0, falls

   $$\sup_{\theta\in\Theta_0} \mathbb{E}_\theta[\varphi(X)] \le \alpha.$$

   Für $\theta \in \Theta_0^c$ nennt man $\mathbb{E}_\theta[\varphi(X)]$ die Macht des Tests an der Stelle θ.

4. **Neyman-Pearson-Lemma**

   Sei $\Theta = \{0,1\}$, $\Theta_0 = \{0\}$, $\Theta_A = \{1\}$, sei $p_i\colon \mathbb{R}^n \to [0,\infty)$ die Dichte von μ_i bezüglich $\mu_0 + \mu_1$ für $i \in \{0,1\}$, und sei $\alpha \in [0,1]$. Dann

   (a) existiert ein randomisierter Test φ und ein $c \in [0,\infty]$, so dass

   $$\mathbb{E}_0[\varphi] = \alpha, \tag{1}$$

   $$\forall\, x \in \mathbb{R}^n: \quad \varphi(x) = \begin{cases} 1 & \text{falls } p_1(x) > cp_0(x) \\ 0 & \text{falls } p_1(x) < cp_0(x) \end{cases}, \tag{2}$$

   (b) jeder Test, der (1) und (2) erfüllt, ist ein mächtigster Test zum Niveau α, und

   (c) jeder mächtigste Test zum Niveau α erfüllt (2) $(\mu_0+\mu_1)$-f.ü. (fast überall). Er erfüllt auch (1), ausser wenn es einen Test φ' gibt mit $\mathbb{E}_1[\varphi'] = 1$ und $\mathbb{E}_0[\varphi'] < \alpha$.

   Bemerkung:

   Sind μ_0 und μ_1 absolut stetige Verteilungen mit Dichten f_0 resp. f_1, so kann man im Punkt (a) des Lemmas p_i jeweils durch f_i ersetzten.

10 Einige Teststatistiken

Seien $X_1, \ldots, X_n$ i.i.d. Zufallsvariablen mit einer geeigneten Verteilung (die vom gewünschten Test abhängt). Seien $Y_1, \ldots, Y_m$ i.i.d. Zufallsvariablen, die unabhängig von $X_1, \ldots, X_n$ sind und mit einer geeigneten Verteilung (die vom gewünschten Test abhängt). Sei

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i, \quad S_X^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2,$$
$$\overline{Y}_m = \frac{1}{m}\sum_{k=1}^{m} Y_k, \quad S_Y^2 = \frac{1}{m-1}\sum_{k=1}^{m}(Y_k - \overline{Y}_m)^2.$$

1. **Ein-Stichproben-t-Test**

$$T = \frac{\overline{X}_n - \mu_0}{S_X/\sqrt{n}}.$$

2. **z-Test**

$$T = \frac{\overline{X}_n - \mu_0}{\sigma/\sqrt{n}}.$$

3. **Zwei-Stichproben-t-Test**

$$T = \frac{\overline{X}_n - \overline{Y}_m}{\sqrt{\frac{1}{n} + \frac{1}{m}}\sqrt{\frac{1}{n+m-2}((n-1)S_X^2 + (m-1)S_Y^2}}.$$

4. **Vorzeichentest**

$$T = \sum_{i=1}^{n} \mathbb{1}_{\{X_i > \mu_0\}}.$$

5. **Zwei-Stichproben-Wilcoxon-Test**

$$T = \sum_{i=1}^{n}\sum_{k=1}^{m} \mathbb{1}_{\{X_i < Y_k\}}.$$

6. **χ^2-Anpassungstest**

$$T = \sum_{i=1}^{k} \frac{(N_i - n\theta_{0,i})^2}{n\theta_{0,i}}.$$

In allen Fällen kennt man die Verteilung der Teststatistik T unter der Nullhypothese (exakt oder approximativ) und findet die relevanten Quantile in den entsprechenden Tabellen.