Mathe C Klausurzettel

Henri Heyden – stu240825

Analysis

Integrierbarkeit

RIEMANN SUMME
Sei $f : [a, b] \longrightarrow \mathbb{R}$, seien (x, ξ) Partition und Stützstellen aus $[a, b]$.
Dann nennen wir $R(f, x, \xi) := \sum_{i=1}^{n}(x_i - x_{i-1}) \cdot f(\xi_i)$ Riemann Summe.

INTEGRIERBARKEIT
Wir nennen f integrierbar, wenn
$\exists R_0 \in \mathbb{R} \forall \epsilon > 0 \exists \delta > 0 \forall (x, \xi) \in \mathrm{PS}(a, b, \delta) : |R(f, x, \xi) - R_0| < \epsilon$ gilt.
Das Integral ist eindeutig, schreibe $\int_a^b f$ oder $\int f$ hierfür.
Wir schreiben auch $\int f(x)\mathrm{d}x := \int f$
Ist f integrierbar, dann kann man das Integral mit einer Beliebigen Folge an $(x_n, \xi_n)_n$ finden wessen Feinheit den Limes 0 hat, sodass die Riemann-Summe konvergiert.
f ist genau dann integrierbar, wenn für alle 2 solcher Folgen ihre Differenz immer zu 0 konvergiert.

STETIG UND KOMPAKT
Eine Funktion f:
…ist stetig in $x \in \mathrm{dom}(f)$, wenn alle Funktionslimetes zu x gleich sind.
…ist beschränkt, wenn ihre Domain eine obere und untere Schranke hat.
…ist abgeschlossen, wenn das komplement ihrer Domain offen ist.
…ist kompakt, wenn sie beschränkt und abgeschlossen ist.
Ist eine Funktion kompakt stetig, dann ist sie gleichmäßig stetig und somit integrierbar.
Aus Lipschitzstetigkeit ($\exists L \in \mathbb{R} \colon \forall x_1, x_2 \in \mathrm{dom}(f) \colon |f(x_1) - f(x_2)| \leq L \cdot |x_1 - x_2|$) folgt auch gleichmäßige Stetigkeit.

ABSCHÄTZUNGEN
Für $f \leq g$ gilt: $\int f \leq \int g$.
Es gilt: $(b - a) \cdot \inf(f) \leq \int_a^b f \leq (b - a) \cdot \sup(f)$

Integrationstechniken

HAUPTSATZ DER DIFFERENZIALRECHNUNG
Schreibe $[\phi]_u^v := \phi(v) - \phi(u)$
Sei $f, F : \Omega \longrightarrow \mathbb{R}$ so, dass $F' = f$ gilt.
Dann gilt: $\int f = F(\sup(\Omega)) - F(\inf(\Omega)) = F(b) - F(a) = [F]_a^b$ für $\Omega = [a, b]$.

PARTIELLE INTEGRATION
Für $f, g : [a, b] \longrightarrow \mathbb{R}$ gilt: $\int_a^b fg' = [fg]_a^b - \int_a^b f'g$

SUBSTITUTION
Für f stetig reel und ϕ stetig differenzierbar reel mit $u, v \in \mathrm{dom}(\phi)$,
sodass $[u, v] \subseteq \mathrm{dom}(\phi)$ und $\phi^{\rightarrow}([u, v]) \subseteq \mathrm{dom}(f)$ ist, gilt:

$$\int_{\phi(u)}^{\phi(v)} f = \int_u^v (f \circ \phi) \cdot \phi'$$

STAMMFUNKTIONEN

Domain	$f(x)$	$F(x)$	args
$\mathbb{R}$	c	cx	$c \in \mathbb{R}$
$\mathbb{R}$	$\sum_{k=0}^{n} a_k x^k$	$\sum_{k=0}^{n} \frac{a_k}{k+1} x^{k+1}$	$a_0 \dots a_n \in \mathbb{R}$
$\mathbb{R}_{>0}$	x^{-1}	$\ln(x)$	
$\mathbb{R}$	b^x	$\frac{b^x}{\ln(b)}$	$b \in \mathbb{R}_{>0} \setminus \{1\}$
$\mathbb{R}_{>0}$	$\log_b(x)$	$\frac{x\ln(x) - x}{\ln(b)}$	$b \in \mathbb{R}_{>0} \setminus \{1\}$
]-1, 1[	$\frac{1}{\sqrt{1-x^2}}$	$\arcsin(x)$	
$\mathbb{R}$	$\frac{1}{1+x^2}$	$\arctan(x)$	

LOGARITHMUS UND EXPONENTIALFUNKTION

Es gilt $\frac{\mathrm{d}}{\mathrm{d}x}\log_a(x) = \frac{\mathrm{d}}{\mathrm{d}x}\frac{\ln(x)}{\ln(a)} = \frac{1}{x\cdot\ln(a)}$ und $\frac{\mathrm{d}}{\mathrm{d}x}a^x = e^{\ln(a)\cdot x} = \ln(a)\cdot e^{\ln(a)\cdot x} = \ln(a)\cdot a^x$

Uneigentliche Integrale

INTEGRAL ÜBER $x^{-\alpha}$

1) $\forall\alpha > 1 : \int_1^{+\infty} x^{-\alpha}\mathrm{d}x = \frac{1}{\alpha-1}$. Für $\alpha \in]0,1]$ divergiert das Integral.
2) $\forall\alpha \in]0,1[: \int_0^1 x^{-\alpha}\mathrm{d}x = \frac{1}{1-\alpha}$. Für $\alpha \geq 1$ divergiert das Integral.

VERGLEICHSKRITERIUM

Seien $f, g : [a, b[\longrightarrow \mathbb{R}$ integrierbar über alle Teilintervalle und $|f| \leq g$.
Konvergiert das uneigentliche Integral über g, dann konvergiert das uneigentliche Integral über f.

INTEGRALKRITERIUM

Warnung nicht im Skript, trotzdem leichtes Argument.
Sei $p \in \mathbb{Z}$ und $f : [p, \infty[\longrightarrow \mathbb{R}_{>0}$ monoton fallend.
Dann ist f integrierbar genau dann, wenn $\sum_{n=p}^{\infty} f(n)$ konvergiert.
Zur Motivation betrachte $\sum_{n=p}^{\infty} f(n)$ als obere Schranke und Annäherung des Integrals mittels Riemann Summe mit gleichmäßiger Feinheit 1.
Tipp für $\int_{-\infty}^{+\infty}$: Betrachte oberes und unteres Integral seperat und berechne $\lim_{b\to+\infty}\int_0^b$ bzw. andersrum.

Analytische Grundstrukturen

Metrische Räume

METRIK

Eine Metrik ist eine Funktion $d : M^2 \longrightarrow \mathbb{R}_{\geq 0}$ mit folgenden Eigentschaften:
1) $d(x, y) = 0 \iff x = y$ (positive Definitheit)
2) $d(x, y) = d(y, x)$ (Symmetrie)
3) $d(x, y) \leq d(x, z) + d(z, y)$ (Dreiecksungleichung)

EIGENSCHAFTEN DER METRIK

Für Metrische Räume gelten alle analytischen Gesetze und Sätze aus MatheB nur mit jeder Metrik nicht nur der Betragsmetrik. Somit sind Begriffe wir Stetigkeit, Limes, Kompaktheit, Funktionslimes etc. äquivalent.

Normierte Räume

NORM

Sei V Vektorraum, dann ist $||\cdot||$ Norm auf V, wenn folgende Gesetze gelten:
1) $||v|| = 0 \iff v = 0$
2) $||\lambda|| = |\lambda|\cdot||v||$
3) $||u + v|| \leq ||u|| + ||v||$

BEISPIELNORMEN ÜBER $\mathbb{R}^n$

Sei $k \in \mathbb{N}_{>0}$, dann ist $||\cdot||_k : \mathbb{R}^n \longrightarrow \mathbb{R}, v \mapsto \sqrt[k]{\sum_{i=1}^n |v_i|}$ Norm
Die Funktion $||\cdot||_\infty : \mathbb{R}^n \longrightarrow \mathbb{R}, v \mapsto \max_{i\in[n]}|v_i|$ ist Norm

EIGENSCHAFTEN DER NORM

Es gilt $||-v|| = ||v||$
Außerdem ist $V^2 \longrightarrow \mathbb{R}_{\geq 0}, (u, v) \mapsto ||u - v||$ Metrik. Somit lassen sich analytische Grundbegriffe für Vektorräume komposieren und alle Eigentschaften der Metrischen Räume für jene Metrik anwenden.
Normen sind stetig. Also $||\cdot||$ ist eine stetige Funktion.

ÄQUIVALENZ VON NORMEN

Seien $(V, ||\cdot||)$ und $(V, ||\cdot||')$ normierte Räume. Wir nennen $||\cdot||$ und $||\cdot||'$ äquivalent, wenn folgendes gilt: $\exists\alpha,\beta > 0 : \forall v \in V : \alpha||v|| \leq ||v||' \leq \beta||v||$
Umgebungen über äquivalente Normen sind äquivalent, genau wie Konvergenz mit den gleichen Limetes, Stetigkeiten und Kompaktheiten.
Auf endlich dimensionalen Vektorräumen sind alle Normen äquivalent.

Differentation im Mehrdimensionalen

Es seien V, W Vektorräume über $\mathbb{R}$, $\Omega \subseteq V$.

Differenzierbarkeit

<u>DEFINITON DER DIFFERENZIERBARKEIT</u>
Wir nennen $f : \Omega \longrightarrow W$ differenzierbar in v, wenn

$$\exists \phi \in \mathrm{L}(V,W) : \lim_{\tilde{v} \longrightarrow v} \frac{||f(\tilde{v}) - (f(v) + \phi(\tilde{v} - v))||_W}{||\tilde{v} - v||_V} = 0_{\mathbb{R}}$$

gilt. Dann nennen wir ϕ Ableitung von f und schreiben $D(f) := \phi$. Offenbar da ϕ linear stetig ist, existiert eine Matrix für ϕ. Folgende Definiton ist äquivalent:

$$\exists \phi \in \mathrm{L}(V,W) : \lim_{h \longrightarrow 0_V} \frac{||f(v+h) - (f(v) + \phi(h))||_W}{||h||_V} = 0_{\mathbb{R}}$$

Gilt $V = W = \mathbb{R}$, dann ist die Definiton äquivalent zur Differenzierbarkeit im bekannten Sinne.

Ableitungsregeln

Ist f differenzierbar, dann sind alle Komponentenfunktionen von f $(V, \mathbb{R})$-differenzierbar.

<u>KOMBINATIONSSÄTZE</u>
Für $\lambda : \Omega \longrightarrow \mathbb{R}$, $f, g : \Omega \longrightarrow V$ differenzierbar gelten folgende Regeln:
1) $D(f+g) = D(f) + D(g)$
2) $D(\lambda f) = \lambda D(f)$
3) $D\left(\frac{1}{\lambda}\right) = \frac{D(\lambda)}{\lambda^2}$
4) $D\left(\frac{f}{\lambda}\right) = \frac{\lambda D(f) - f D(\lambda)}{\lambda^2}$

Partielle Ableitungen

Seien $\Omega \subseteq \mathbb{R}^n$, $f : \Omega \longrightarrow \mathbb{R}^d$, $i \in [n]$, $j \in [d]$, $v \in \Omega$.

<u>PARTIELLE DIFFERENZIERBARKEIT</u>
Wir nennen f_j partiell differenzierbar in v nach der i-ten Variable, wenn $f_{j,v,i}$ in v_i differenzierbar ist.
Wir definieren $\partial_i f_j(v) := f'_{j,v,i}(v_i)$.

<u>BEISPIEL</u>
Sei $f : \mathbb{R} \times \mathbb{R}_{\geq 0} \longrightarrow \mathbb{R}^4, (x, y) \mapsto \left(x + y, xy, x^2 + 3y - 6, -2\sqrt{y}\right)$.
Dann gilt:
$\partial_1 f_2(x,y) = y, \quad \partial_1 f_3(x,y) = 2x, \quad \partial_1 f_4(x,y) = 0,$
$\partial_2 f_1(x,y) = 1, \quad \partial_2 f_3(x,y) = 3$

<u>GRADIENT UND JACOBI-MATRIX</u>
Wir definieren den Gradienten von f_j als $\mathrm{grad}_{f_j}(v) := (\partial_1 f_j(v), \ldots, \partial_n f_j(v))$.
Des Weiteren definieren wir die Jacobi-Matrix von f:

$$\mathbf{J}_f(v) := \begin{bmatrix} \partial_1 f_1(v) & \ldots & \partial_n f_1(v) \\ \vdots & & \vdots \\ \partial_1 f_d(v) & \ldots & \partial_n f_d(v) \end{bmatrix}$$

Somit gilt:

$$\mathbf{J}_f(v) = \begin{bmatrix} \mathrm{grad}_{f_1}(v) \\ \vdots \\ \mathrm{grad}_{f_d}(v) \end{bmatrix}$$

Ist f differenzierbar in v, dann ist $\mathbf{J}_f(v)$ Ableitung von f, aber Achtung: Vollständige partielle Differenzierbarkeit impliziert NICHT totale Differenzierbarkeit!

<u>TOTALE DIFFERENZIERBARKEIT AUS PARTIELLER DIFFERENZIERBARKEIT</u>
Ist f vollständig partiell differenzierbar und alle Ableitung sind stetig, dann ist f total differenzierbar.

C^1-FUNKTIONEN UND C^2-FUNKTIONEN
Wir nennen $f : \Omega \longrightarrow \mathbb{R}^d$ eine C^1-Funktion, wenn sie vollständig stetig partiell differenzierbar ist. Somit ist also eine C^1-Funktion total differenzierbar. f ist C^1-Funktion, wenn all ihre Komponentenfunktionen aus Kombinationen von Projektionen stetiger differenzierbaren eindimensionalen Funktionen bestehen.
Wir nennen $f : \Omega \longrightarrow \mathbb{R}$ eine C^2-Funktion, wenn sie eine C^1-Funktion ist, und $\forall i \in [n] : \partial_i f$ stetig partiell differenzierbar ist.
Dies ist äquivalent dazu, dass grad_f eine C^1-Funktion ist.
Die partielle Ableitung von $\partial_i f$ nach der j-ten Variable bezeichnen wir mit $\partial_j \partial_i f$.

HESSE-MATRIX
Sei $f : \Omega \longrightarrow \mathbb{R}$ eine C^2-Funktion, wir definieren die Hesse-Matrix von f als:

$$\mathbf{H}_f := \begin{bmatrix} \partial_1\partial_1 f(v) & \dots & \partial_1\partial_n f(v) \\ \vdots & & \vdots \\ \partial_n\partial_1 f(v) & \dots & \partial_n\partial_n f(v) \end{bmatrix}$$

offenbar gilt: $\mathbf{H}_f(v) = (\mathbf{J}_{\text{grad}_f}(v))^\top$, denn die Spalten der Hesse-Matrix sind die Gradienten der Komponentenfunktionen von grad_f. Die Transposition ist auch nicht nötig, da die Hesse-Matrix symmetrisch ist. Das heißt es gilt: $\forall i, j \in [n]\colon \partial_i\partial_j f = \partial_j\partial_i f$.

Lokale Extremstellen

DEFINITON
Wir nennen $v \in \Omega$ eine lokale Maximumstelle/Minimumstelle von f, wenn gilt:
$\exists U \in \mathcal{U}(v)\colon \forall u \in U\colon f(u) \leq f(v)$ bzw. $f(u) \geq f(v)$.
Wir bezeichnen die Menge der lokalen Maximumstellen $\text{LMAX}(f)$ und die Menge der lokalen Minimumstellen $\text{LMIN}(f)$. Die Vereinigung beider Mengen ist $\text{LEXT}(f)$.

DEFINITHEIT
Sei $A \in \mathbb{R}^{n\times n}$ symmetrisch, dann nennen wir $A \dots$
…positiv definit, wenn $\forall k \in [n] : \det(A_k) > 0$
…negativ definit, wenn $\forall k \in [n] : \det(A_k) < 0$
…indifinit, sonst.

DEFINITHEIT DER HESSE-MATRIX
Definiere die Menge der kritischen Stellen von f als $K(f) := \text{grad}_f^{\leftarrow}(0_{\mathbb{R}^n})$.
Es gilt für alle $v \in K(f)$:
1) $\mathbf{H}_f(v)$ positiv definit $\Longrightarrow v \in \text{LMIN}(f)$
2) $\mathbf{H}_f(v)$ negativ definit $\Longrightarrow v \in \text{LMAX}(f)$
3) $\mathbf{H}_f(v)$ indifinit $\Longrightarrow v \notin \text{LEXT}(f)$

DEFINITHEIT EINER 2×2-MATRIX
Definiere $A := \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ für $a, b, c \in \mathbb{R}$. Dann gilt:
1) $\det(A) = ac - b^2$
2) A positiv definit $\Longleftrightarrow a > 0 \land \det(A) > 0$.
3) A negativ definit $\Longleftrightarrow a < 0 \land \det(A) > 0$.
4) A indefinit $\Longleftrightarrow \det(A) < 0$

Stochastik

Mathematische Modellierung des Zufalls

σ-ALGEBRA, EVENTRAUM
Sei $\Omega \neq \emptyset$ Menge und $\mathcal{E} \subseteq \mathcal{P}(\Omega)$. Für $A \in \mathcal{P}(\Omega)$ definieren wir $A^c := \Omega \setminus A$.
Wir nennen $\mathcal{E}$ eine σ-Algebra über Ω, wenn gilt:
1) $\Omega \in \mathcal{E}$
2) $\forall A \in \mathcal{E}\colon A^c \in \mathcal{E}$
3) $\forall (A_i)_i \in \mathcal{S}(\mathcal{E})\colon \bigcup_i A_i \in \mathcal{E}$
Wir nennen $(\Omega, \mathcal{E})$ **Eventraum** und Ω **Ergebnisraum**.
Die Elemente von $\mathcal{E}$ nennen wir **Event** und die Elemente von Ω **Ergebnis**.
Einen E'Raum nennen wir diskret, wenn Ω abzählbar ist

BEISPIELE
Für Ω ist $\{\emptyset, \Omega\}$ σ-Algebra über Ω. Dann ist $(\Omega, \{\emptyset, \Omega\})$ Eventraum, genau wie $(\Omega, \mathcal{P}(\Omega))$

WAHRSCHEINLICHKEITSMASS
Sei $(\Omega, \mathcal{E})$ Eventraum. Sei eine Funktion $\mathbb{P} : \mathcal{E} \longrightarrow [0,1]$, dann nennen wir $\mathbb{P}$ Wahrscheinlichkeitsmaß auf $(\Omega, \mathcal{E})$, wenn gilt:
1) $\mathbb{P}(\Omega) = 1$
2) $\forall (A_i)_i \in \mathcal{S}(\mathcal{E}), \bigcap_i A_i = \emptyset \colon \mathbb{P}\left(\bigcup_i A_i\right) = \sum_i \mathbb{P}(A_i)$
Das Tripel $(\Omega, \mathcal{E}, \mathbb{P})$ nennen wir Wahrscheinlichkeitsraum.

GRUNDLEGENDE EIGENSCHAFTEN DER WAHRSCHEINLICHKEIT
1) $\mathbb{P}(\emptyset) = 0$
2) $\forall A, B \in \mathcal{E}, A \cap B = \emptyset \colon \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$
3) $\forall A, B \in \mathcal{E}, A \subseteq B \colon \mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A)$
4) $\forall A \in \mathcal{E} : \mathbb{P}(A^c) = 1 - \mathbb{P}(A)$
5) $\forall A, B \in \mathcal{E} \colon \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$

UNIFORMES W'MASS
Nun beachte, dass wenn $\mathbb{P}(A)$ und $\mathbb{P}(B)$ bekannt sind für $A \cap B = \emptyset$, dass dann auch $\mathbb{P}(A \cup B)$ berechnet werden kann. Somit lassen sich Wahrscheinlichkeitsmaße konstruieren, wenn $\forall \omega \in \Omega : \mathbb{P}(\omega)$ bekannt ist.
Eines solcher W'Maße ist das Uniforme mit $\mathbb{P}(\omega) := \frac{1}{|\Omega|}$. Dann gilt: $\mathbb{P}(A) = \frac{|A|}{|\Omega|}$.

KONDITIONIERUNG
Seien $A, B \in \mathcal{E}$ so, dass $\mathbb{P}(B) > 0$. Wir definieren die W'keit von A konditioniert auf B, oder W'keit von A gegeben B als $\mathbb{P}(A|B) := \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$.

BAYES
Für $A, B \in \mathcal{E}$ mit $\mathbb{P}(A), \mathbb{P}(B) > 0$ gilt: $\mathbb{P}(A|B) = \frac{\mathbb{P}(A)}{\mathbb{P}(B)} \cdot \mathbb{P}(B|A)$

ODDS
Wir definieren für $A, B \in \mathcal{E}$ mit $\mathbb{P}(B) > 0$ die Odds von A konditioniert auf B, bzw die Odds von A als:

$$\mathrm{Od}(A|B) := \begin{cases} \frac{\mathbb{P}(A|B)}{\mathbb{P}(A^c|B)} = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A^c \cap B)} & \mathbb{P}(A|B < 1) \\ +\infty & \mathbb{P}(A|B = 1) \end{cases}$$

$$\mathrm{Od}(A) := \begin{cases} \frac{\mathbb{P}(A)}{\mathbb{P}(A^c)} & \mathbb{P}(A < 1) \\ +\infty & \mathbb{P}(A = 1) \end{cases}$$

Folgende Umrechnungsregeln gelten:
$o := \frac{p}{1-p}$ sind die odds für p Wahrscheinlichkeit.
$p := \frac{o}{o+1}$ ist die Wahrscheinlichkeit für o odds.

UNABHÄNGIGKEIT
Sei $(A_i)_{i \in I}$ eine endliche Familie von Events. Wir nennen diese unabhängig, wenn gilt:
$\forall I' \subseteq I \colon \mathbb{P}\left(\bigcap_{i \in I'} A_i\right) = \prod_{i \in I'} \mathbb{P}(A_i)$

Zufallsvariablen über diskreten Wahrscheinlichkeitsräumen

ZUFALLSVARIABLEN
Sei $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ diskreter W'Raum.
Wir nennen alle Funktionen $X \colon \Omega \longrightarrow \mathbb{R}$ Zufallsvariablen.

EVENTS ÜBER ZUFALLSVARIABLEN
Sei ϕ ein logischer Ausdruck der Form: „$X \in U / X \leq u / X = u$“, dann bezeichnen wir damit das Event, für das für $X(\omega)$ der Ausdruck gilt für alle $\omega \in \Omega$. Somit lassen sich Events mit ZV elegant konstruieren, und es lassen sich Wahrscheinlichkeiten wie $\mathbb{P}(X = u) = \mathbb{P}(\{\omega \in \Omega \mid X(\omega) = u\})$ berechnen.

GESETZE DER NEUEN EVENTS
1) $\mathbb{P}(\bot \mid \beta) = 0 \wedge \mathbb{P}(\top|\beta) = 1$
2) $\mathbb{P}(\neg\phi \mid \beta) = 1 - \mathbb{P}(\phi|\beta)$
3) $\mathbb{P}(\phi \vee \psi \mid \beta) = \mathbb{P}(\phi|\beta) + \mathbb{P}(\psi|\beta) - \mathbb{P}(\phi \wedge \psi|\beta)$

Verteilungen über Zufallsvariablen

VERTEILUNG
Man sieht, dass die Funktion $\mathbb{P}_X : \mathcal{P}(\mathbb{R}) \longrightarrow [0,1], U \mapsto \mathbb{P}(X{\in}U)$ ein W'maß für den Eventraum $(\mathbb{R}, \mathcal{P}(\mathbb{R}))$ ergibt gegeben der ZV X und dem diskretem Eventraum $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$.
Diese Funktion nennen wir Verteilung von X.

TRÄGER
Wir definieren den **Träger** von X als: $\mathrm{supp}(X) := \{u \in \mathbb{R} \mid \mathbb{P}(X = u) > 0\}$.
Offenbar ist $X^{\rightarrow}(\Omega) = \mathrm{img}(X)$ abzählbar (da Ω abzählbar ist), und es gilt: $\mathrm{supp}(X) \subseteq \mathrm{img}(X)$, also ist der Träger einer ZV über diskreten W'räumen abzählbar.
Somit schreibt man auch $\mathbb{P}(X{\in}U) = \mathbb{P}(X{\in}U \cap \mathrm{supp}(X)) = \sum_{u \in U \cap \mathrm{supp}(X)} \mathbb{P}(X{=}u) =: \sum_{u \in U} \mathbb{P}(X{=}u)$.

PROBABILITY MASS FUNCTION (PMF)
Wir nennen $\rho_X : \mathbb{R} \longrightarrow [0,1],\ u \mapsto \mathbb{P}(X = u)$ die PMF von X.

TUPEL UND FAMILIEN VON ZV
Wir definieren eine Familie von ZV als: $I \subseteq \mathbb{N}, X := \bigtimes_{i \in I} X_i$, dann ist X Familie über alle ZV $X_i \in M_i \in \mathcal{P}(M)^I$ mit M Menge an ZV. Ein Tupel einer ZV ist ein Spezialfall dessen, und ZVar, dass
$\exists n \in \mathbb{N} : I = [n]$ gilt. Wir schreiben dann $X := (X_1 \cdot X_n)$ für Tupel und $(X_i)_{i \in I}$ für die Familie.
Wir definieren somit folgende Events: $X \in U :\Longleftrightarrow \forall i \in I : X_i \in U_i$ und $X = u :\Longleftrightarrow \forall i \in I : X_i = u_i$. Die Verteilung, der Träger und die PMF werden gleich definiert.

UNIFORME VERTEILUNG
Für eine Familie von ZV oder einer ZV X sagen wir, dass X uniform verteilt ist,
wenn $\forall u \in M : \mathbb{P}(X = u) = \frac{1}{|M|}$ gilt. Folgendes ist die PMF einer Uniformen Verteilung:
$\rho_{\mathrm{Unif}(M)} : \mathbb{R}^I \longrightarrow [0,1], u \mapsto \begin{cases} \frac{1}{|M|} & u \in M \\ 0 & \text{sonst} \end{cases}$

GEOMETRISCHE VERTEILUNG
Sei $p \in]0,1[$, wir nennen eine ZV X geometrisch verteilt, wenn gilt: $\mathbb{P}(X{=}k) = (1-p)^k p$

POISSON-VERTEILUNG
Sei $\lambda > 0$, wir nennen eine ZV X poisson-verteilt mit EW λ, wenn gilt: $\mathbb{P}(X{=}k) = \frac{\lambda^k}{k!} e^{-\lambda}$

MARGINALISIERUNG BEISPIEL
Sei $(X_1, X_2) \sim \mathrm{Unif}(\{(0,1),(0,2),(2,1)\})$. Dann gilt:
$\mathbb{P}(X_1 = 0)$
$= \mathbb{P}((X_1 = 0) \wedge (X_2 = 1 \vee X_2 = 2))$
$= \mathbb{P}((X_1, X_2) = (0,1) \vee (X_1, X_2) = (0,2))$
$= \mathbb{P}((X_1, X_2) = (0,1)) + \mathbb{P}((X_1, X_2) = (0,2))$
$= \frac{2}{3}$

BERNOULLIVERTEILUNG
Sei $p \in [0,1]$ für eine ZV X sagen wir, dass sie Bernoulli-verteilt ist, wenn gilt:
$\mathbb{P}(X{=}1) = p$ und $\mathbb{P}(X{=}0) = 1{-}p$, dann ist die PMF:

$$\rho_{\mathrm{Ber}(p)} : \mathbb{R} \longrightarrow [0,1], u \mapsto \begin{cases} p & u = 1 \\ 1-p & u = 0 \\ 0 & \text{sonst} \end{cases}$$

UNABHÄNGIGKEIT
Wir nennen eine endliche Familie von ZV $X = (X_i)_{i \in I}$ unabhängig, wenn gilt:

$$\forall (U_i)_{i \in I} \in \mathcal{P}(\mathbb{R})^I : \mathbb{P}\left(\bigwedge_{i \in I} (X_i \in U_i)\right) = \prod_{i \in I} \mathbb{P}(X_i \in U_i)$$

Wir sagen $(X_i)_{i \in I} \sim \mathcal{D}$ i.i.d. für eine Verteilung $\mathcal{D}$, wenn wir meinen, dass alle X_i unabhängig aber gleich verteilt sind.

Summe unabhängiger Bernoulli-ZV

Sei $p \in [0,1]$ und $X = (X_1, \dots, X_n) \sim \text{Ber}(p)$ i.i.d., und definiere $S := \sum_{i=1}^n X_i$, dann gilt für $k \in [n]$:

$$\mathbb{P}(S{=}k) = \binom{n}{k} p^k (1-p)^{n-k}$$

Binomialverteilung

Sei $p \in [0,1]$ und $X = (X_1, \dots, X_n) \sim \text{Ber}(p)$ i.i.d., dann nennen wir X binomialverteilt und schreiben:

$$(X_1, \dots, X_n) \sim \text{Ber}(p) \text{ i.i.d.} \Longrightarrow \sum_{i=1}^n X_i \sim \text{Bin}(n,p)$$

Erwartungswert und Varianz von ZV über diskreten W'räumen

Erwartungswert

Sei X ein ZV. X hat einen Erwartungswert, genau dann, wenn $\mathbb{E} := \sum_{u \in \mathbb{R}} u \cdot \mathbb{P}(X{=}u)$ definiert ist bzw. summierbar ist für $|u|$ statt u. Offenbar gilt: $\mathbb{E}(X) = \sum_{\omega \in \Omega} X(\omega) \cdot \mathbb{P}(\omega)$.

Erwartungswerte von Verteilungen

Es gilt: $\mathbb{E}(\text{Ber}(p)) = p$, $\mathbb{E}(\text{Bin}(n,p)) = n \cdot p$, $\mathbb{E}(\text{Pois}(\lambda)) = \lambda$ und $\mathbb{E}(\text{Geo}(p)) = \frac{1-p}{p}$. Des Weiteren gilt: $|\mathbb{E}(X)| \le \mathbb{E}(|X|)$

Law of the unconscious statistician (LOTUS)

Sei X ZV oder endliche Familie davon und sei $f\colon \text{supp}(X) \longrightarrow \mathbb{R}$. Dann hat die ZV $f \circ X$ einen EW genau dann, wenn $\sum_{u \in R^I} |f(u)| \cdot \mathbb{P}(X = u) < +\infty$ gilt.
Des Weiteren lässt sich der EW von $f \circ X$ so berechnen: $\mathbb{E}(f \circ X) = \sum_{u \in R^I} f(u) \cdot \mathbb{P}(X = u)$.

Linearität des EW

Für X ZV oder Familie davon mit EW und $\lambda \in \mathbb{R}$ oder gleichlange Familie davon gilt: $\mathbb{E}(\lambda \cdot X) = \lambda \cdot \mathbb{E}(X)$ es gilt auch: $\mathbb{E}(X + \lambda) = \mathbb{E}(X) + \lambda$

Varianz

Wir definieren die Varianz einer ZV X als: $\text{Var}(X) := \mathbb{E}((X - \mathbb{E}(X))^2)$. Es gilt: $\text{Var}(\lambda X + a) = \lambda^2 \cdot \text{Var}(X)$.
Die Varianz kann auch so berechnet werden: $\text{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$
Es gilt: $\text{Var}(\text{Ber}(p)) = p(1-p)$, $\text{Var}(\text{Bin}(n,p)) = np(1-p)$, $\text{Var}(\text{Pois}(\lambda)) = \lambda$ und $\text{Var}(\text{Geo}(p)) = \frac{1-p}{p^2}$
Für $(X_i)_{i \in I}$ eindliche Familie von ZV, die Paarweise unkorreliert sind und Varianz haben, dann gilt: $\text{Var}(\sum_{i \in I} X_i) = \sum_{i \in I} \text{Var}(X_i)$.

Kovarianz und Unkorreliertheit

Seien X, Y ZV mit Varianz. Definiere die Kovarianz $\text{Cov}(X,Y) := \mathbb{E}(X{\cdot}Y) - \mathbb{E}(X){\cdot}\mathbb{E}(Y)$.
Wir nennen solche ZV unkorreliert, wenn $\text{Cov}(X,Y) = 0$ gilt. Offenbar folgt dies aus (X,Y) unabhängig.

Abschätzungen

Markow-Ungleichung: Sei X ZV mit EW, sei $\mathbb{P}(X \ge 0) = 1$ und $a > 0$. Dann gilt $\mathbb{P}(X \ge a) \le \frac{\mathbb{E}(X)}{a}$.
Tschebyschow-Ungleichung: Sei X ZV mit Varianz und $a > 0$, dann gilt:
$\mathbb{P}(|X - \mathbb{E}(X)| \ge a) \le \frac{\text{Var}(X)}{a^2}$ und $\mathbb{P}(X - \mathbb{E}(X) \ge a) \le \frac{\text{Var}(X)}{\text{Var}(X)+a^2}$
Sei $\mu := \mathbb{E}(X), \sigma := \sigma(X), k > 0$, dann gilt nach Tschebyschow:
$\mathbb{P}(\mu - k\sigma < X < \mu + k\sigma) = \mathbb{P}(|X - \mu| < k\sigma) \ge 1 - \frac{1}{k^2}$

Verteilungen mit Dichte

Sei $(\Omega, \mathcal{E}, \mathbb{P})$ W'Raum.

Allgemeine ZV

Wir nennen $X\colon \Omega \longrightarrow \mathbb{R}$ eine ZV, wenn gilt: $\forall u \in \mathbb{R} : [X \le u] \in \mathcal{E}$

Borellmengen

Um sinnvolle Events für eine allgemeine ZV zu konstruieren, definieren wir eine σ-Algebra, die nicht die Potenzmenge von $\mathbb{R}$ ist.
Für $\mathcal{G} := \{]-\infty, u[\mid u \in \mathbb{R}\}$, was selbst schon eine gültige σ-Algebra ist, definieren wir $\mathcal{B} := \bigcap_{\mathcal{G} \subseteq \mathcal{E}' \subseteq \mathcal{P}(\mathbb{R}) \wedge \mathcal{E} \text{ ist } \sigma\text{-Algebra über } \mathbb{R}} \mathcal{E}'$.
$\mathcal{B}$ ist σ-Algebra.
Folgendes sind Borellmengen: Intervalle, abzählbare Mengen, Offene Mengen, Vereinigungen abzählbar vieler Borellmengen,

Schnitte abzählbar vieler Borellmengen, Komplemente von Borellmengen.
Außerdem ist das kartesische Produkt über Borellmengen wieder Borellmenge, somit übertragen sich auch noch äquivalente Definitionen und Regeln für Folgen von ZV, welche über Folgen von Borellmengen agieren.

KOMB'SATZ FÜR ALLGEMEINE ZV
Seien X, Y ZV, $\lambda \in \mathbb{R}$, $f\colon D \longrightarrow \mathbb{R}$ stetig für $\operatorname{img}(X) \subseteq D \in \mathcal{B}$.
Dann sind folgendes auch ZV: $X + \lambda$, λX, $X + Y$, XY, $f \circ X$

VERTEILUNG EINER ALLGEMEINEN ZV
Folgendes W'maß nennen wir die Verteilung einer allgemeinen ZV X:
$\mathbb{P}_X\colon \mathcal{B} \longrightarrow [0,1], U \mapsto \mathbb{P}(X \in U)$

KUMULATIVE VERTELUNGSFUNKTION
Sei X ZV, dann nennen wir $F_X\colon \mathbb{R} \longrightarrow [0,1], u \mapsto \mathbb{P}(X{\le}u)$, die kumulative Vertelungsfunktion oder auch CDF von X. Wir nennen X stetige ZV, wenn F_X stetig ist. Für die CDF gelten folgende Eigenschaften: F_X ist monoton steigend, F_X ist rechtsstetig (also stetig für rechte Grenze), $\lim_{\xi\to+\infty} F_X = 1 \wedge \lim_{\xi\to-\infty} F_X = 0$
Es gilt außerdem wenn F_X stetig ist: $\mathbb{P}(X{=}u) = 0$ für alle $u \in \mathbb{R}$ und $\mathbb{P}(X{\in}I) = \mathbb{P}(X{\in}I^o)$ für I^o das offene Intervall von I.
Man kann eine CDF auch aus der PMF definieren mit: $\mathbb{R} \longrightarrow [0,1], u \mapsto \sum_{\xi\le u} \rho_X(\xi)$, offenbar sind ZV, die solch eine PMF bzw. CDF haben dann immer diskret, wie wenn die Bildmenge von X abzählbar ist.

UNABHÄNGIGKEIT ALLGEMEINER ZV
Der Begriff der Unabhängigkeit überträgt sich für allgemeine ZV, nur, dass wir nicht mehr alle Teilmengen sondern nur noch alle Borellmengen betrachten und unendliche ZV-Familien sind erlaubt.

WAHRSCHEINLICHKEITSDICHTEFUNKTIONEN
Sei $\rho\colon \mathbb{R} \longrightarrow \mathbb{R}_{\ge 0}$ integrierbar über alle kompakten Intervalle, sodass $\int_{-\infty}^{+\infty} \rho = 1$ gilt, dann nennen wir ρ PDF oder auch Wahrscheinlichkeitsdichtefunktion.
Ist ρ PDF, dann ist $F\colon \mathbb{R} \longrightarrow [0,1], u \mapsto \int_{-\infty}^{u} \rho$ eine stetige CDF.
Hat die ZV X eine PDF ρ, dann gilt: $\mathbb{P}(a \le X \le b) = \int_a^b \rho$. Beachte hier, dass da die Punktevents keine Wahrscheinlichkeit haben, deswegen es egal ist, ob man $<$ statt $\le$ verwendet etc.

ERWARTUNGSWERT VIA PDF, LOTUS FÜR ALLGEMEINE ZV
Sei X allgemeine ZV mit PDF ρ, dann hat X genau dann einen EW, wenn $\int_{-\infty}^{+\infty} |t|{\cdot}\rho(t)\,\mathrm{d}t < +\infty$ gilt, dann gilt:

$$\mathbb{E}(X) = \int_{-\infty}^{+\infty} t \cdot \rho(t)\,\mathrm{d}t$$

Sei $f\colon D \longrightarrow \mathbb{R}$ stetig mit $\operatorname{img}(X) \subseteq D \subseteq \mathbb{R}$, dann hat $f \circ X$ gena dann einen EW, wenn $\int_{-\infty}^{+\infty} |f|{\cdot}\rho < +\infty$ gilt, dann gilt:

$$\mathbb{E}(f \circ X) = \int_{-\infty}^{+\infty} f \cdot \rho$$

DIVERSE VERTEILUNGEN
Uniforme Verteilung
Seien $a < b \in \mathbb{R}$, wir sagen eine ZV X hat uniforme Verteilung zwischen a und b und schreiben $X \sim \operatorname{Unif}(a,b)$, wenn X folgende PDF hat: $\rho\colon \mathbb{R} \longrightarrow \mathbb{R}_{\ge 0}, t \mapsto \frac{1}{b-a}$, wenn $t \in]a,b[$ sonst 0. Es gilt: $\mathbb{E}(\operatorname{Unif}(a,b)) = \frac{a+b}{2}$ und $\operatorname{Var}(\operatorname{Unif}(a,b)) = \frac{(b-a)^2}{12}$.
CDF: $t \mapsto \frac{t-a}{b-a}$ im Falle $t \in]a,b[$ sonst links 0 und rechts 1.
Exponentialverteilung
Sei $\lambda > 0$, wir sagen eine ZV X hat Exponentialverteilung mit Rate λ, wenn folgende Funktion die PDF von X ist:
$\rho\colon \mathbb{R} \longrightarrow \mathbb{R}_{\ge 0}, t \mapsto \lambda e^{-\lambda t}$, wenn $t \ge 0$ und sonst 0. Es gilt: $\mathbb{E}(\operatorname{Exp}(\lambda)) = \lambda^{-1}$ und $\operatorname{Var}(\operatorname{Exp}(\lambda)) = \lambda^{-2}$. CDF: $t \mapsto 1 - e^{-\lambda t}$
Pareto-Verteilung
Sei $a > 0$, wir sagen eine ZV X hat Pareto-Verteilung, wenn X folgende PDF hat: $\rho\colon \mathbb{R} \longrightarrow \mathbb{R}_{\ge 0}, t \mapsto \frac{a}{t^{a+1}}$, wenn $t \ge 1$ und sonst 0. Es gilt: $\mathbb{E}(\operatorname{Par}(a)) = \frac{a}{a-1}$, wenn $a > 1$ und wenn $a > 2$ gilt: $\operatorname{Var}(\operatorname{Par}(a)) = \frac{a}{(a-1)^2(a-2)}$. CDF: $t \mapsto 1 - t^{-a}$
Cauchy-Verteilung
Sei $t_0 \in \mathbb{R}, \gamma > 0$, wir sagen X hat Cauchy-Verteilung mit Lageparameter t_0 und Skalenparameter γ, wenn X folgende PDF hat:
$\rho\colon \mathbb{R} \longrightarrow \mathbb{R}_{\ge 0}, t \mapsto \left(\pi\gamma\left(1 + \left(\frac{t-t_0}{\gamma}\right)^2\right)\right)^{-1}$. Die Cauchy-Verteilung hat keinen EW oder Varianz. CDF: $\frac{1}{2} + \frac{1}{\pi} \cdot \arctan\left(\frac{t-t_0}{\gamma}\right)$
Normalverteilung
Sei $\mu \in \mathbb{R}$ und $\sigma > 0$. Wir sagen X hat Normalverteilung mit EW μ und Varianz σ^2,
und schreiben $X \sim \mathcal{N}(\mu, \sigma^2)$, wenn X folgende PDF hat: $\rho\colon \mathbb{R} \longrightarrow \mathbb{R}_{>0}, t \mapsto \frac{e^{-\frac{1}{2}\cdot\left(\frac{t-\mu}{\sigma}\right)^2}}{\sqrt{2\pi}\cdot\sigma}$

Eigentschaften der Normalverteilung

Sei $X \sim \mathcal{N}(\mu, \sigma^2)$ und $\delta \in \mathbb{R}$, $\lambda \neq 0$. Dann gilt: $\lambda X + \delta \sim \mathcal{N}(\lambda\mu+\delta, \lambda^2\sigma^2)$, insbesondere $\frac{X-\mu}{\sigma} \sim \mathcal{N}(0,1)$. Es gilt auch: $\mathbb{E}(X) = \mu$ und $\mathrm{Var}(X) = \sigma^2$. Für $(X_i)_{i\in I}$ endliche Familie an ZV mit $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$ i.d.d. gilt: $\sum_{i\in I} X_i \sim \mathcal{N}\left(\sum_{i\in I}\mu_i, \sum_{i\in I}\sigma_i^2\right)$.

Gedächtnislosigkeit

Sei X eine verteilte ZV. Wir nennen die Verteilung gedächtnislos, wenn gilt: $\forall n,k > 0\colon \mathbb{P}(X \geq n+k \mid X \geq k) = \mathbb{P}(X \geq k)$
Beweis für $X \sim \mathrm{Geo}(p)$: Es gilt: $\mathbb{P}(X \geq n+k \mid X \geq n) = \frac{\mathbb{P}(X\geq n+k \wedge X \geq n)}{\mathbb{P}(X\geq n)} = \frac{\mathbb{P}(X\geq n+k)}{\mathbb{P}(X\geq n)} = \frac{(1-p)^{n+k-1}}{(1-p)^{n-1}} = (1-p)^k = \mathbb{P}(X \geq k)$.
Außerdem ist die Geometrische Verteilung gedächtnislos – ich bin das auch.

Konvergenz in Verteilung und Zentraler Grenzwertsatz

Konvergenz in Verteilung: Sei $(X_n)_n$ beliebige Folge von ZV, wir sagen $(X_n)_n$ konvergiert in Verteilung gegen $\mathcal{D}$, wenn gilt: $\forall u \in \mathbb{R}$, sodass $F_\mathcal{D}$ stetig in $u\colon \lim_n F_{X_n}(u) = F_\mathcal{D}(u)$
Zentraler Grenzwertsatz: Seien $\mu \in \mathbb{R}, v > 0, (X_i)_{i\geq 1}$ Folge von i.i.d ZV über dem selben W'raum mit $\mathbb{E}(X_i) = \mu$ und $\mathrm{Var}(X_i) = v$ für alle $i \geq 1$. Dann konvergiert $\left(\frac{1}{\sqrt{n}}\sum_{i=1}^n \frac{X_i-\mu}{\sqrt{v}}\right)_{n\geq 1}$ in Verteilung gegen $\mathcal{N}(0,1)$.

Konfidenzintervalle in Normalverteilungsmodellen

Statistisches Modell

Sei $n, d \in \mathbb{N}_1, \Theta \subseteq \mathbb{R}^d$. Ein statistisches Modell ist eine Familie $(\mathcal{D}_i^\vartheta)_{(\vartheta,i)\in\Theta\times[n]}$ von Verteilungen für eine ZV.
Nach Existenzsatz existiert für jeden Parameter $\vartheta \in \Theta$ eine Familie $(X_i^\vartheta)_{i\in[n]}$ von ZV mit folgenden Eigenschaften:
X^ϑ ist unabhängig und $\forall i \in [n]\colon X_i^\vartheta \sim \mathcal{D}_i^\vartheta$. Um nicht immer den Existenzsatz anzuwenden schreiben wir $(X \sim \mathcal{D}_i^\vartheta)_{(\vartheta,i)\in\Theta\times[n]}$ um direkt sowohl die Familie der Verteilungen als auch die existierende Familie an unabhängigen ZV zu definieren.
Bei Modellen mit endlichen oder kleiner Anzahl an Verteilungen pro Paramenter ϑ schreiben wir auch gerne Beispielweise $(X = (X_1 \dots X_n) \sim \mathcal{D})_{\vartheta\in\Theta}$ oder $\begin{pmatrix} X=(X_1,\dots,X_n)\sim\mathcal{D}_1^\vartheta \\ Y=(Y_1,\dots,Y_n)\sim\mathcal{D}_2^\vartheta \end{pmatrix}_{\vartheta\in\Theta}$ für die Verteilungen $\mathcal{D}^\vartheta$ bzw. $\mathcal{D}_1^\vartheta, \mathcal{D}_2^\vartheta$.

Reele Kenngrösse

Sei $(X \sim \mathcal{D}_i^\vartheta)_{(\vartheta,i)\in\Theta\times[n]}$ ein statistisches Modell und $\tau\colon \Theta \longrightarrow \mathbb{R}$. Wir nennen τ eine reelle Kenngröße, wenn wir mittels τ Parameter auf eine bestimmte Größe abbilden wollen. Zum Beispiel wäre $\vartheta \mapsto \frac{b-a}{2}$ eine reele Kenngröße für uniform verteilte Modelle, wenn $(a,b) = \vartheta$ gilt.

Motivation

Sei $u \in \mathbb{R}^n$ eine Stichprobe, dann wollen wir ein statistisches Modell $(X \sim \mathcal{D}_i^\vartheta)_{(\vartheta,i)\in\Theta\times[n]}$ so wählen, dass es ein ϑ^* gibt, den sogenannten „wahren Parameter", sodass u durch einen Prozess entstanden ist, der gut modelliert wird durch X^{ϑ^*}. Man würde gerne ϑ^* und dann auch $\tau(\vartheta^*)$ approximieren.

Punktschätzung

Punktschätzer

Sei $(X \sim \mathcal{D}_i^\vartheta)_{(\vartheta,i)\in\Theta\times[n]}$ statistisches Modell und $\tau\colon \Theta \longrightarrow \mathbb{R}$ eine reelle Kenngröße. Ein Punktschätzer für τ ist schlicht eine Funktion $T\colon \mathbb{R}^n \longrightarrow \mathbb{R}$. Um möglichst effiziente Punktschätzer zu konstruieren, werden wir einige Eigenschaften definieren.

Erwartungstreue

Für alle $\vartheta \in \Theta$ definieren wir den Bias von T für τ als $\mathbb{B}^{\vartheta,\tau}(T(X)) := \mathbb{E}^\vartheta(T(X)) - \tau(\vartheta)$, wenn der EW existiert. Wir nennen T einen erwartungstreuen oder unbiased Punktschätzer für τ, wenn der Bias stets 0 ist, also gilt: $\forall\vartheta \in \Theta\colon \mathbb{E}^\vartheta(T(X)) = \tau(\vartheta)$

Erwartungstreue Schätzer für EW und Varianz

Sei $n \geq 2$ und seien $\mu\colon \Theta \longrightarrow \mathbb{R}$ und $v\colon \Theta \longrightarrow \mathbb{R}_{\geq 0}$ reele Kenngrößen, sodass $\mu(\vartheta) = \mathbb{E}(\mathcal{D}^\vartheta)$ und $v(\vartheta) = \mathrm{Var}(\mathcal{D}^\vartheta)$ für alle $\vartheta \in \Theta$
Dann sind mit $\langle\cdot\rangle\colon \mathbb{R}^n \longrightarrow \mathbb{R}, u \mapsto \frac{1}{n}\sum_{i=1}^n u_i$ die Funktionen $\langle\cdot\rangle$ und $s^2\colon \mathbb{R}^n \longrightarrow \mathbb{R}_{\geq 0}, u \mapsto \frac{1}{n-1}\sum_{i=1}^n (u_i - \langle u\rangle)^2$ jeweils erwartungstreue Schätzer für μ (also $\langle\cdot\rangle$) bzw. v (also s^2).
Sei $u \in \mathbb{R}^n$ Stichprobe, dann nennen wir $\langle u\rangle$ den empirischen Erwartungswert von u und $s^2(u)$ die empirische Varianz von u.