Andererseits: $A \subseteq [0,1) \Rightarrow \biguplus_{r \in \mathbb{Q} \cap [0,1)} (r+A) \subseteq [0,2)$

$$\Rightarrow 2 \overset{b)}{=} L([0,2)) \geq L\Big(\bigcup_{r \in \mathbb{Q} \cap [0,1)} (r+A)\Big) = \sum_{r \in \mathbb{Q} \cap [0,1)} L(r+A)$$

$$\overset{c)}{=} \sum_{r \in \mathbb{Q} \cap [0,1)} L(A) \Rightarrow L(A) = 0 \quad \text{ℇ}. \quad \square$$

5.2. Allg. W'räume.

In Abschnitt 5.1 haben wir gesehen, dass eine naive Verallg. von Def 1.7 auf nicht-diskrete Räume scheitert, da die Potenzmenge als Definitionsmenge für ein W'maß zu groß ist. Die Lsg besteht darin, sich auf eine Teilmenge, ~~eine~~ eine σ-Algebra "Ereignisse" zu beschränken (vgl Diskussion vor Def 1.6 für Interpretation):

Def 5.2 (σ-Algebra): Sei Ω eine Menge. Eine nichtleere Familie $\mathcal{F}$ von Teilmengen von Ω heißt σ-Algebra, falls

i) $\Omega \in \mathcal{F}$

ii) $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$

iii) Ist $(A_n)_n$ Folge in $\mathcal{F} \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$.

Bem 5.3. Zu jedem Mengensystem $C \subseteq \Omega$ gibt es eine kleinste σ-Algebra $\sigma(C)$, die C enthält.

Bsp 5.4: i) $\mathcal{P}(\Omega)$ ist eine σ-Algebra

ii) $\Omega = \mathbb{R}^n$, C sei die Familie der rechts halboffenen Intervalle, d.h. der Intervalle der Form

$]a,b] = \{x = (x_1, \dots, x_n) : a_i < x_i \leq b_i, \ 1 \leq i \leq n\}$

für $a, b \in \mathbb{R}^n$ mit $a_i < b_i$, $1 \leq i \leq n$

$\mathcal{B}_n = \sigma(C)$ heißt Borel'sche σ-Algebra in $\mathbb{R}^n$.

Elemente von $\mathcal{B}_n$ heißen Borelmengen. Insbesondere ist jede offene und abgeschlossene Menge Borel.

Def 5.5: Sei Ω eine Menge, $\mathcal{F}$ eine σ-Algebra von Teilmengen von Ω. Ein W'maß auf $(\Omega, \mathcal{F})$ ist eine Fkt $P: \mathcal{F} \to [0,1]$ s.d.

i) $P[\Omega] = 1$

ii) Für disjunkte $A_1, A_2, \ldots \in \mathcal{F}$ gilt
$$P[\bigcup_{n=1}^{\infty} A_n] = \sum_{n=1}^{\infty} P[A_n] \quad (\sigma\text{-Additivität}).$$

$(\Omega, \mathcal{F}, P)$ heißt W'raum.

Def 5.6: Sei $(\Omega, \mathcal{F}, P)$ ein W'raum. Eine Abb $X: \Omega \to \mathbb{R}$ heißt reelle ZV, wenn für alle $a \in \mathbb{R}$ gilt
$$X^{-1}((-\infty, a]) \in \mathcal{F}.$$

Bsp 5.7: Sei Ω diskret, $\mathcal{F} = \mathcal{P}(\Omega)$. Dann stimmt die Def 5.6 mit der bisherigen Def. von ZV überein.

~~Ukedes~~ Die andere fundamentale Klasse von W'räumen für diese Vorlesung beruht auf folgender Def:

Def 5.8: Bezeichne das Lebesgue-Integral (bzw Riemann-Integral) mit $\int \ldots dx$. Eine Lebesgue (Riemann)-integrierbare Fkt $f: \mathbb{R} \to [0, \infty)$ heißt Dichte / Wahrscheinlichkeitsdichte, wenn $\int_{-\infty}^{\infty} f(x)\, dx = 1$

~~Bsp~~ 5.9: Bem 5.3: Für $\Omega = \mathbb{R}$, $\mathcal{F} = \mathcal{B}_1$ definiert ~~d. Standard~~ $P[A] = \int_A f(x) dx$ eine W'maß auf $\mathbb{R}$.

Bsp 5.10: a) Die Standard-Normalvtlg ist definiert durch
$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad x \in \mathbb{R}$$

b) Die Normalvtlg mit Mittelwert $\mu \in \mathbb{R}$ und Varianz σ^2 ~~...~~ > 0 ist definiert durch

$$\varphi(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

c) Die Gleichverteilung auf $[a,b]$ mit $a<b$ ist definiert durch $f(x)=\frac{1}{b-a}\mathbb{1}_{[a,b]}$.

d) Die Exponentialvtlg zum Parameter $\lambda>0$ ist definiert durch

$$f(x) = \lambda e^{-\lambda x}\mathbb{1}_{x\geq 0}$$

e) Die Cauchy-Vtlg zum Parameter $c>0$ ist def. durch

$$f(x) = \frac{c}{\pi}\frac{1}{x^2+c^2}, \quad x\in\mathbb{R}$$

Def 5.11: Eine Fkt $F: \mathbb{R}\to[0,1]$ heißt Vtlgsfkt, falls

i) F mon. wachsend, d.h. $\forall s\leq t$, $F(s)\leq F(t)$

ii) F ist rechtsstetig, d.h. $\lim_{t\downarrow s} F(t) = F(s)$

iii) $\lim_{t\to\infty} F(t) = 1$, $\lim_{t\to-\infty} F(t) = 0$

Bsp 5.12: Für jede Dichte f ist $\int_{-\infty}^{t} f(x)\,dx$ eine stetige Vtlgsfkt.

Def 5.13: a) Sei $(\Omega,\mathcal{F},\mathbb{P})$ ein W'raum und X eine ZV, dann heißt $F_X(t) := \mathbb{P}[X\leq t]$ die Vtlgsfkt von X.

b) X heißt absolutstetig mit Dichte f, falls

$$F_X(t) = \int_{-\infty}^{t} f(x)\,dx \quad \forall t\in\mathbb{R}.$$

In diesem Fall nennt man auch F_X absolut stetig.

Hat eine ZV X eine Dichte f so gilt $\forall a<b$

$$\mathbb{P}[a<X\leq b] = \mathbb{P}[X\leq b] - \mathbb{P}[X\leq a] = \int_a^b f(x)\,dx$$

Insbesondere gilt $\mathbb{P}[X=x]=0$ und somit

$$\mathbb{P}[a\leq X\leq b] = \mathbb{P}[a<X<b].$$

Eine ZV mit Dichte heißt normalverteilt, uniformverteilt, gleichförmig verteilt bzw Cauchy verteilt, wenn sie eine Dichte gemäß Bsp 5.10 b), c), d), e) hat.

Def 5.13 Sei $(\Omega, \mathcal{F}, \mathbb{P})$ W'raum und X eine ZV mit Dichte f. Sei g eine messbare Fkt bzgl der Borel'schen σ-Algebra auf $\mathbb{R}$ (d.h. $g^{-1}(A) \in \mathcal{B}_1 \;\forall A \in \mathcal{B}_1$).

a) Ist $x \mapsto g(x) f(x)$ (Lebesgue/Riemann)-integrierbar, so existiert der Erwartungswert von $g(X)$, er ist def. durch

$$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f(x)\, dx$$

Ist $g(x) = x$, so sagen wir der Erwartungswert von X existiert.

b) existiert $\mathbb{E}X$ und ist $(x - \mathbb{E}X)^2 f(x)$ integrierbar, so ist die Varianz von X def. durch

$$Var(X) = \int_{-\infty}^{\infty} (x - \mathbb{E}X)^2 f(x)\, dx$$

Bem 5.14: ~~Im Falle diskreter Räume ist des "Referenzmaß" das Zählmaß und die Begriffe~~

Wie im diskreten Fall kann gezeigt werden, dass $\mathbb{E}$ linear ist und $Var(\lambda X + b) = \lambda^2 Var(X)$ für $\lambda, b \in \mathbb{R}$.

Bsp 5.15.

i) Sei X standard normalverteilt, $X \sim \mathcal{N}(0,1)$

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |x| e^{-\frac{x^2}{2}}\, dx = \frac{2}{\sqrt{2\pi}} \int_0^{\infty} x e^{-\frac{x^2}{2}}\, dx = \frac{2}{\sqrt{2\pi}} \left(-e^{-\frac{x^2}{2}}\right)\Big|_0^{\infty} = \sqrt{\frac{2}{\pi}} < \infty$$

$\Rightarrow \mathbb{E}X$ existiert und

$$\mathbb{E}X = \int_{-\infty}^{\infty} x e^{-\frac{x^2}{2}} \frac{1}{\sqrt{2\pi}}\, dx = 0, \text{ da } x e^{-\frac{x^2}{2}} \text{ eine}$$

ungerade Fkt ist

(10)

$$\mathrm{Var}(X) = \int_{-\infty}^{\infty} x^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \lim_{N\to\infty} \frac{1}{\sqrt{2\pi}} \int_{-N}^{N} x\left(x e^{-\frac{x^2}{2}}\right) dx$$

partielle Integr.

$$= \lim_{N\to\infty} \frac{1}{\sqrt{2\pi}} \left(-x e^{-\frac{x^2}{2}}\right)\Big|_{-N}^{N} + \underbrace{\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx}_{=1} = 0+1 = 1$$

ii) Sei X normalverteilt mit Mittelwert μ und Varianz $\sigma^2 > 0$, $X \sim \mathcal{N}(\mu, \sigma^2)$. Mit der Trafo $y = \frac{x-\mu}{\sigma}$ und Teil i)

folgt $\int_{-\infty}^{\infty} |x| \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \int |\mu + \sigma y| \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy$

$$\le |\mu| + |\sigma| \int |y| e^{-y^2/2} \frac{1}{\sqrt{2\pi}} dy < \infty.$$

$\Rightarrow$ Erwartungswert existiert und es gilt

$$\mathbb{E} X = \int x \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \int \frac{1}{\sqrt{2\pi}} (\mu + \sigma y) e^{-y^2/2} dy = \mu$$

und $\mathbb{E} X^2 = \ldots = \frac{\sigma^2}{\sqrt{2\pi}} \int y^2 e^{-y^2/2} dy = \sigma^2$.

$$\Rightarrow \quad X \sim \mathcal{N}(\mu, \sigma^2) \Leftrightarrow Y = \frac{X-\mu}{\sigma} \sim \mathcal{N}(0,1)$$

Falls $X \sim \mathcal{N}(\mu, \sigma^2)$, $a, b \in \mathbb{R}$, $a \neq 0$ $\Rightarrow$

$$Z = aX + b \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$$

Def 5.16: i) Eine uneigentl. Riemann-integrierbare ~~nichtnegative~~ Funktion $f: \mathbb{R}^n \to [0, \infty)$ heißt n-dimensionale Dichte, wenn $\int_{\mathbb{R}^n} f(x)\, dx = 1$

ii) Sei f eine n-dim. Dichte und $X_1, \ldots, X_n$ ZV. Man sagt, dass sie die gem. Dichte f haben, wenn

$$\mathbb{P}[X_1 \le a_1, X_2 \le a_2, \ldots, X_n \le a_n] = \int_{(-\infty, a_1] \times \ldots \times (-\infty, a_n]} f(x)\, dx \quad \forall a_1, \ldots, a_n \in \mathbb{R}.$$

Def 5.17. $X_1, \dots, X_n$ seien n ZV auf einem W'raum $(\Omega, \mathcal{F}, P)$. Sie heißen unabhängig, wenn für alle $a_1 \dots a_n \in \mathbb{R}$ gilt

$$P[X_1 \le a_1, \dots, X_n \le a_n] = \prod_{i=1}^{n} P[X_i \le a_i]$$

Satz 5.18 $X_1, \dots, X_n$ n ZV auf $(\Omega, \mathcal{F}, P)$. Jedes X_j habe Dichte f_j. Dann sind $X_1, \dots, X_n$ genau dann unabhängig, wenn eine gem. Dichte für $X_1, \dots, X_n$ durch $x = (x_1, \dots, x_n) \mapsto f_1(x_1) \cdots f_n(x_n)$ geg. ist.

Bew. ÜA.

Satz 5.19. Seien X, Y unabh. ZV mit Dichten f und g. Dann hat die ZV $X+Y$ die Dichte h geg. durch die Faltung von f und g: $f * g\ (x) := h(x) := \int_{-\infty}^{\infty} f(x-y)\, g(y)\, dy$

Bew: Nach Satz 5.18 ist die gem. Dichte von (X,Y) geg. durch $f(x)\, g(y)$. Mit $C_a = \{(x,y) \in \mathbb{R}^2 : x+y \le a\}$ folgt

$$P[X+Y \le a] = P[(X,Y) \in C_a]$$

$$= \int_{C_a} f(x)\, g(y)\, dx\, dy$$

$$\overset{u=x+y,\ v=y}{=} \int_{-\infty}^{a} \int_{-\infty}^{\infty} f(u-v)\, g(v)\, dv\, du \qquad \square$$

Satz 5.20: Es seien $(X_i)_{1 \le i \le n}$ unabhängige normalverteilte ZV mit Erwartungswerten μ_i und Varianzen σ_i^2. Dann ist

$$\sum_{i=1}^{n} X_i \sim \mathcal{N}\left(\sum_{i=1}^{n} \mu_i, \sum_{i=1}^{n} \sigma_i^2\right)$$

Beweis: Induktion über n (Ann. impliziert $X_1 + \dots + X_{n-1}$ unabh. von X_n (warum?)). Fall $n=2$ Rechnung mittels Satz 5.19

$\square$

(72)

Bem 5.21: Der zentrale Grenzwertsatz kann als Konvergenz gegen einen Fixpunkt interpretiert werden. Für unabh. ZV mit derselben Dichte f mit $\mathbb{E}X_i = \mu$, $Var(X_i) = \sigma^2$ gilt

$$\mathbb{P}\left[\frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n\sigma^2}} \leq t\right] \xrightarrow[n\to\infty]{} \Phi(t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{t} e^{-\frac{x^2}{2}}\,dx$$

Für $X_i \sim \mathcal{N}(\mu, \sigma^2)$ gilt "="

IV Statistik

6 Schätztheorie

6.1 Grundbegriffe

Bsp 6.1: Sprödigkeit von Röhren

Datensätze: Messungen $x_1, \ldots, x_n$, Deutung: n Wdh. eines unabhängigen, identisch ~~gleich~~verteilten Experimentes, d.h. $x_1, \ldots, x_n$ sind Realisierung von ZV $X_1, \ldots, X_n$ u.i.v.

z.B. $X_i \sim \mathcal{N}(\mu, \sigma^2)$, $\mu \hat{=}$ Sprödigkeit des Rohres, μ unbekannt

$\sigma^2 \hat{=}$ Ungenauigkeit des Messgeräts, $\sigma^2 > 0$ bekannt

μ soll geschätzt werden, das bedeutet, zu finden ist eine Vorschrift $\hat{\mu}$, die zu jedem Datensatz $x_1, \ldots, x_n$ einen Wert $\hat{\mu} = \hat{\mu}(x_1, \ldots, x_n)$ liefert, den Schätzwert.

Plausible Schätzer:

1. $\hat{\mu}(x_1, \ldots, x_n) = \frac{1}{n}\sum_{i=1}^{n} x_i$ Stichprobenmittel
(wenig robust gegen „Ausreißer")

2. Stichprobenmedian

$x_{(1)} \leq x_{(2)} \leq \ldots \leq x_{(n)}$ Anordnung der Stichprobe

$$\hat{\mu}(x_1, \ldots, x_n) = \begin{cases} x_{(k)} & n = 2k-1 \\ \frac{1}{2}(x_{(k)} + x_{(k+1)}) & n = 2k \end{cases}$$

3. Getrimmte Mittel

Fixiere $0 < \gamma < \frac{1}{2}$. $\exists! \, k \in \mathbb{N}$ mit $\frac{k}{n} \le \gamma < \frac{k+1}{n}$

Eliminiere die k kleinsten und die k größten Werte und bilde das arithmetische Mittel der verbleibenden Werte.

Def 6.2: $\mathcal{X}$ Menge, $\mathcal{F}$ σ-Alg. auf $\mathcal{X}$, $(\mathbb{P}_\vartheta, \vartheta \in \Theta)$ Familie von W'maßen auf $(\mathcal{X}, \mathcal{F})$. Dann heißt

- $(\mathcal{X}, \mathcal{F}, (\mathbb{P}_\vartheta, \vartheta \in \Theta))$ statistisches Modell
- $\mathcal{X}$ Stichprobenraum
- Θ Parametermenge; Falls $\Theta \subseteq \mathbb{R}^d$ für $d \in \mathbb{N}$, heißt das Modell parametrisch. Falls $\Theta \subseteq \mathbb{R}$, heißt das Modell einparametrisch. Falls $\Theta \not\subseteq \mathbb{R}^d$ heißt das Modell nicht-parametrisch.

Bem 6.3: Für den Erwartungswert und Varianz bzgl $\mathbb{P}_\vartheta$ schreiben wir $\mathbb{E}_\vartheta(\cdot)$, $\mathrm{Var}_\vartheta(\cdot)$

In dieser Vorlesung nur zwei Arten von Modellen:

i) $\mathcal{X}$ abzählbar, $\mathcal{F} = \mathcal{P}(\mathcal{X})$

ii) $\mathcal{X} = \mathbb{R}^k$, $\mathcal{F} = \mathcal{B}_k$ (Borel'sche σ-Algebra), $\mathbb{P}_\vartheta$ hat Dichte f_ϑ

Bsp 6.4: i) $\mathcal{X} = \mathbb{N}_0$, $\mathbb{P}_\vartheta = \mathrm{Poi}(\vartheta)$, $\vartheta > 0$ $(\Theta = (0, \infty))$

ii) $\mathcal{X} = \mathbb{R}$, $\mathcal{F} = \mathcal{B}_1$, $\mathbb{P}_\vartheta = \mathcal{N}(\mu, \sigma^2)$, $\vartheta = \mu \in \Theta = \mathbb{R}$, σ fix

Def 6.5: Sei $(\mathcal{X}, \mathcal{F}, (\mathbb{P}_\vartheta, \vartheta \in \Theta))$ ein stat. Modell, wir setzen $\mathcal{X}^n = \mathcal{X} \times \mathcal{X} \times \dots \times \mathcal{X}$, $\mathcal{F}^n = \mathcal{F} \otimes \dots \otimes \mathcal{F}$ (~~je n-Mal~~), $\mathbb{P}^n_\vartheta = \mathbb{P}_\vartheta \otimes \dots \otimes \mathbb{P}_\vartheta$ (je n-Mal). Dann heißt $(\mathcal{X}^n, \mathcal{F}^n, (\mathbb{P}^n_\vartheta, \vartheta \in \Theta))$ n-faches Produktmodell.

(74)

(Erinnerung)
Wir schreiben $X_i : \mathcal{X}^n \to \mathcal{X}$, $X_i(x_1 \dots x_n) = x_i$
und $X_1, \dots, X_n$ uiv, X_i hat Vtlg P_ϑ
(vgl Bsp 6.1. $(\mathbb{R}^n, \mathcal{B}_n, (\mathcal{N}(\mu,\sigma^2)^n, \vartheta \in \Theta = \mathbb{R}))$.)

6.2 Bsp für Schätzer

Def 6.6: Sei $(\mathcal{X}, \mathcal{F}, (P_\vartheta, \vartheta \in \Theta))$ ein stat. Modell, $\Gamma \subset \mathbb{R}^d$ für ein $d \in \mathbb{N}$.
i) Eine Abb. $S: \mathcal{X} \to \Gamma$ heißt Statistik, falls $S^{-1}(A) \in \mathcal{F}$ für alle $A \in \mathcal{B}_d$ (falls $d = 1$ so ist eine Statistik eine reelle ZV).
ii) Sei $\tau : \Theta \to \Gamma$; dann heißt jede Statistik S Schätzer für τ.

Bem: Die offensichtliche mehrfache Benennung desselben Objektes soll die jeweilige Interpretation in den Vordergrund schieben.

Bsp 6.7. $\mathcal{X} = \mathbb{R}^n$, $\mathcal{F} = \mathcal{B}_n$, $P_\vartheta = \mathcal{N}(\mu, \sigma^2)$, $\vartheta = \mu \in \mathbb{R} = \Theta$,
$\tau(\vartheta) = \vartheta$, $\rightsquigarrow \Gamma = \Theta = \mathbb{R}$.
$x = (x_1, \dots, x_n) \mapsto T(x_1, \dots, x_n) = \frac{1}{n} \sum_{i=1}^{n} x_i$ ist Schätzer für τ.

Ziel: Gütekriterien, Vergleichskriterien für Schätzer
Herleitung über Bsp.

Bsp 6.8. (Taxiproblem).
Schätzung der Zahl der Taxen in einer Stadt: Taxen sind nummeriert $1, 2, \dots, \vartheta$; $\vartheta \in \mathbb{N}$ unbekannt, soll geschätzt werden.
Die Nummern der ersten n vorbeikommenden Taxen werden notiert $\rightsquigarrow$ Datensatz: $x = (x_1, \dots, x_n)$, $1 \le x_i \le \vartheta$

Modell: n-faches Produktmodell mit

$\mathcal{X} = \mathbb{N}$, $\mathcal{F} = \mathcal{P}(\mathcal{X})$, $P_\vartheta = \mathrm{Unif}(\{1, \dots, \vartheta\})$, $\Theta = \mathbb{N}$.

Betrachte Schätzer: $T(x) = T(x_1, \dots, x_n) = \max\limits_{1 \le i \le n} x_i$

Eigenschaften von T:

1. T ist sogenannter Maximum-Likelihood-Schätzer für ϑ, d.h.

$$\mathbb{P}^n_{T(x)}[X_1 = x_1, \dots, X_n = x_n] = \max_{\vartheta \in \mathbb{N}} \mathbb{P}^n_\vartheta[X_1 = x_1, \dots, X_n = x_n]$$

<u>für alle</u> $x = (x_1, \dots, x_n) \in \mathbb{N}^n$

Denn: $\mathbb{P}^n_\vartheta[X_1 = x_1, \dots, X_n = x_n] = \begin{cases} \frac{1}{\vartheta^n}, & \max\limits_{1\le i\le n} x_i \le \vartheta \\ 0, & \text{sonst.} \end{cases}$

2. Es wäre wünschenswert, dass man zumindest im Mittel richtig schätzt, d.h.

$$\mathbb{E}_\vartheta[T] \overset{!}{=} \vartheta \quad \forall \vartheta \in \Theta$$

Diese Eigenschaft heißt <u>erwartungstreu</u>. $T = \max\limits_{1\le i\le n}(X_i)$ ist nicht erwartungstreu, denn

$$\mathbb{P}^n_\vartheta[T \le \vartheta] = 1, \ \mathbb{P}^n_\vartheta[T < \vartheta] > 0 \Rightarrow \mathbb{E}_\vartheta[T] < \vartheta,$$

d.h. T unterschätzt ϑ im Mittel. $\forall \vartheta \in \mathbb{N}$.

3. Eine weitere wünschenswerte und wichtige Eigenschaft ist, dass $T = T_n$ für $n \to \infty$ gegen ϑ konvergiert, genauer:

$$\forall \varepsilon > 0 \quad \mathbb{P}^n_\vartheta[|T_n - \vartheta| \ge \varepsilon] \xrightarrow[n\to\infty]{} 0 \quad \forall \vartheta \in \Theta$$

(stochastische Konvergenz, vgl. Bem. 3.8)

Diese Eigenschaft heißt <u>Konsistenz</u>. T_n ist konsistent:

Für $0 < \varepsilon < 1$: $\mathbb{P}^n_\vartheta[|T_n - \vartheta| \ge \varepsilon] \underset{T_n \le \vartheta}{=} \mathbb{P}^n_\vartheta[T_n \le \vartheta - \varepsilon]$

$$\overset{\varepsilon < 1}{=} \mathbb{P}^n_\vartheta[T_n \le \vartheta - 1] = \left(\frac{\vartheta - 1}{\vartheta}\right)^n \xrightarrow[n\to\infty]{} 0.$$

Ein anderer Schätzer für ϑ ist z.B. $S(x) = S(x_1, \dots, x_n)$

$$= \max_{1\le i\le n} x_i + \min_{1\le i\le n} x_i - 1$$

Eig.: i) $S(x)$ ist kein ML Schätzer (klar wg. oben)

ii) $S(x)$ ist erwartungstreu (ÜA)

iii) $S(x)$ ist konsistent. (ÜA).

Ein stetiges Analogon zu Bsp 6.8 ist:

Bsp 6.3: Erraten [illegible] von Zufallszahlen / Gameshow.

Der Moderator einer Fernsehshow führt einen Apparat vor, der Zufallszahlen mit Gleichverteilung auf $[0,\vartheta]$ erzeugt, wobei der Parameter ϑ vom Moderator frei eingestellt ~~wurde~~ ~~kann~~ wurde. Zwei Spieler dürfen den Apparat k n-mal bedienen und sollen den Wert von ϑ möglichst gut schätzen.

Modell: Produktmodell: $\mathcal{X}^n = (0,\infty)^n$, $\mathcal{F} = \mathcal{B}^n$,

$(\mathbb{P}_\vartheta^n)_{\vartheta\in\Theta}$, $\Theta = (0,\infty)$, $P_\vartheta = \mathrm{Unif}([0,\vartheta])$

Wie beim Taxiproblem können wir $T(x) = \max_{1\le i\le n} x_i$ oder $S(x) = \max_{1\le i\le n} x_i + \min_{1\le i\le n} x_i$ ~~...~~ wählen.

Wie oben ist T nicht erwartungstreu aber asymptotisch erwartungstreu (ÜA)
$\vartheta > \mathbb{E}_\vartheta[T] = \frac{n}{n+1}\vartheta \xrightarrow[n\to\infty]{} \vartheta \quad \forall\vartheta\in\Theta$

Hingegen ist S erwartungstreu:

$\mathbb{E}_\vartheta[S] = \frac{n}{n+1}\vartheta + \frac{1}{n+1}\vartheta = \vartheta$ (ÜA)

T und S sind konsistent (ÜA)

Eine weitere Möglichkeit wäre $R_n = \frac{2}{n}\sum_{i=1}^n X_i$. Denn

R_n ist erwartungstreu und konsistent

Bew: 1. $\mathbb{E}_\vartheta[R_n] = \frac{2}{n}\sum_{i=1}^n \underbrace{\mathbb{E}_\vartheta[X_i]}_{=\vartheta/2} = \frac{2}{n}\cdot n\cdot\frac{\vartheta}{2} = \vartheta \quad \forall\vartheta\in\Theta$.

2. $P_\vartheta^n[|R_n - \vartheta| \ge \varepsilon] = P_\vartheta^n[|\frac{1}{n}\sum_{i=1}^n X_i - \frac{\vartheta}{2}| \ge \frac{\varepsilon}{2}] \to 0$

nach dem schwachen Gesetz der großen Zahlen, Satz 3.7.