Um nun T, S und R zu vergleichen, kann man die Varianzen betrachten. Wünschenswert ist, dass ein asymptotisch erwartungstreuer Schätzer eine kleine Varianz hat (warum?).

Es gilt $\mathrm{Var}_\vartheta(T) = \frac{n\vartheta^2}{(n+1)^2(n+2)}$ für $T = \max_{1\le i\le n} X_i$

und $\mathrm{Var}_\vartheta(S) = \frac{2\vartheta^2}{(n+1)(n+2)}$ (ÜA).

Aber $\mathrm{Var}_\vartheta(R) = \frac{4}{n^2} \sum_{i=1}^{n} \underbrace{\mathrm{Var}_\vartheta(X_i)}_{\frac{\vartheta^2}{12}} = \frac{\vartheta^2}{3n}$

In Bezug auf die Varianz ist R also ein deutig schlechter als T und S.

Bem: T, S und R sind asymptotisch erwartungstreu. Man kann also die Qualität von Schätzern nach verschiedenen Kriterien beurteilen (ML-Schätzer, Konsistenz, Erwartungstreue, kleine Varianz), die wir nun etwas systematischer behandeln wollen.

6.3 Das Maximum-Likelihood Prinzip

Das ML-Prinzip besteht darin, denjenigen Parameter, bzw. dasjenige Modell zu wählen, das den beobachteten Daten die größte W'keit gibt.

Def 6.10. (ML-Schätzer):
Sei $(\mathcal{X}, \mathcal{F}, (\mathbb{P}_\vartheta;\ \vartheta\in\Theta))$ ein statistisches Modell.

a) Die Abb. $L: \mathcal{X}\times\Theta \to [0,\infty)$

$$L_\vartheta(x) = L(x,\vartheta) = \begin{cases} \mathbb{P}_\vartheta[\{x\}] & \text{, falls } \mathcal{X} \text{ diskret} \\ f_\vartheta(x) & \text{, falls } \mathcal{X}\subset\mathbb{R}^n,\ \mathbb{P}_\vartheta \text{ hat Dichte } f_\vartheta \end{cases}$$

heißt Likelihood-Funktion.

(78)

(ii) Ein Schätzer $T: \mathcal{X} \to \Theta$ für $\tau(\vartheta) = \vartheta$ heißt ML-Schätzer, falls

$$L(x, T(x)) = \max_{\vartheta \in \Theta} L(x, \vartheta) \qquad \forall x \in \mathcal{X}$$

Bsp 6.11: Ein auf den Boden geworfener Reisnagel fällt mit unbekannter W'keit $\vartheta \in [0,1]$ auf die Spitze, ansonsten auf den Rücken.
Wir wollen ϑ schätzen. Wir werfen n Mal und sehen, dass der Reisnagel x Mal auf der Spitze landet.
Wir nehmen an, dass die Experimente unabhängig sind und modellieren sie über ein Binomialmodell, d.h.

$\mathcal{X} = \{0, 1, \dots, n\}$

$\mathbb{P}_\vartheta = Bin(n, \vartheta)$, $\vartheta \in \Theta = [0,1]$

Die Likelihood-Fkt ist geg. durch $L(x, \vartheta) = \binom{n}{x} \vartheta^x (1-\vartheta)^{n-x}$

Die Maximumstellen von L ist identisch mit der von ~~log~~ $\log L$:

$$\frac{d}{d\vartheta} \log L(x, \vartheta) = \frac{x}{\vartheta} - \frac{n-x}{1-\vartheta}$$

und die rechte Seite ist fallend in ϑ mit

$$\frac{d}{d\vartheta} \log L(x, \vartheta) = 0 \iff \vartheta = \frac{x}{n} \qquad \text{(insb. ist } \tfrac{x}{n} \text{ Maximum!)}$$

$\Rightarrow \quad T(x) = \frac{x}{n}$ ist ML-Schätzer.

Bsp 6.9 (Fortsetzung, Gameshow Teil II).
Die Likelihoodfkt ist geg. durch

$$L(x, \vartheta) = \begin{cases} \frac{1}{\vartheta^n} & , x = (x_1, \dots, x_n) \in \{0, \dots, \vartheta\}^n \\ 0 & , \text{sonst} \end{cases}$$

$\Rightarrow \quad T(x) = \max_{1 \le i \le n} x_i$ ist ML-Schätzer.

Satz 6.12: Zu $n \in \mathbb{N}$ betrachte das Produktmodell $(\mathbb{R}^n, \mathcal{B}_n, (\mathcal{N}(\mu,\sigma^2))^n, (\mu,\sigma) \in \mathbb{R} \times \mathbb{R}_{>0} = \Theta))$.
Dann ist der ML-Schätzer für $c(\mu,\sigma^2) = (\mu,\sigma^2)$ geg. durch $T(x_1,\dots,x_n) = (\bar{x}_n, V_n)$, wobei $\bar{x}_n = \frac{1}{n}\sum_{i=1}^n x_i$ der empirische Mittelwert und $V_n = \frac{1}{n}\sum_{i=1}^n (x_i - \bar{x}_n)^2$ die empirische Varianz von $x_1,\dots,x_n$ bezeichnen.

L
VL8

Bew: Wir benutzen die Verschiebungsformel
$$\frac{1}{n}\sum_{i=1}^n (x_i - m)^2 = \frac{1}{n}\sum_{i=1}^n (x_i - \bar{x}_n)^2 + (\bar{x}_n - m) \quad (*) \quad \forall m \in \mathbb{R}.$$
Die Likelihood-Fkt ist geg durch
$$L(x,(\vartheta,\sigma^2)) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\Big(-\sum_{i=1}^n \frac{(x_i-\mu)^2}{2\sigma^2}\Big)$$
Wir wollen $L(x,(\vartheta,\sigma^2))$ über (ϑ,σ^2) maximieren, also $\sum_{i=1}^n \frac{(x_i-\mu)^2}{2\sigma^2}$ minimieren. Zunächst minimieren wir $\sum_{i=1}^n (x_i-\mu)^2$. Mit $(*)$ gilt
$$\frac{1}{n}\sum_{i=1}^n (x_i-\mu)^2 = \underbrace{\frac{1}{n}\sum_{i=1}^n (x_i-\bar{x}_n)^2}_{=(i)} + \underbrace{(\bar{x}_n-\mu)^2}_{=(ii)}$$

(i) hängt nicht von μ ab, (ii) wird minimal für $\mu = \bar{x}_n \Rightarrow \hat{\mu}(x_1,\dots,x_n) = \bar{x}_n$
Setzen wir dies in die Likelihoodfkt ein, erhalten wir:
$$\tilde{L}(x,(\hat{\mu},\sigma^2)) := \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\Big(-\sum_{i=1}^n \frac{(x_i-\bar{x}_n)^2}{2\sigma^2}\Big)$$

Nun müssen wir $v = \sigma^2$ so wählen, dass $L(x,(\hat{\mu},v))$ maximal wird. Wir gehen wieder zur log-Likelihoodfkt:
$$\log L(x,(\hat{\mu},v)) = -\frac{n}{2}\log(2\pi v) - \frac{1}{2v}\sum_{i=1}^n (x_i-\bar{x}_n)^2$$
$$\frac{d}{dv}\log L(x,(\hat{\mu},v)) = -\frac{n}{2v} + \frac{1}{2v^2}\sum_{i=1}^n (x_i-\bar{x}_n)^2 \overset{!}{=} 0$$
$$\Leftrightarrow \quad v = V_n = \frac{1}{n}\sum_{i=1}^n (x_i-x_n)^2$$
Und
$$\frac{d^2}{dv^2}\log L(x,(\hat{\mu},v))\Big|_{v=V_n} = \Big(\frac{n}{2v^2} - \frac{1}{v^3}\sum_{i=1}^n (x_i-\bar{x}_n)^2\Big)\Big|_{v=V_n}$$

(80)

$$= \frac{n}{v_n^2}\left(\frac{1}{2}-1\right) = -\frac{n}{2v_n^2} < 0 \Rightarrow$$

$\Rightarrow$ an der Stelle $v = v_n$ liegt ein Maximum vor.

$\Rightarrow$ $(\hat{\mu}, v_n) = (\bar{x}_n, v_n)$ ist der (kein eindeutige!) ML-Schätzer für (μ, σ^2). Δ

Bem.: $X_1, X_2, \dots, X_n$ uiv mit Dichte $f_\vartheta(x) = \frac{1}{2} e^{-|x-\vartheta|}, x \in \mathbb{R}$

Dann ist ein ML-Schätzer für ϑ geg. durch

$T(x_1, \dots, x_n) = M(x_1, \dots, x_n) =$ Median von $x_1, \dots, x_n$ (ÜA)

6.4. Erwartungstreue und quadr. Fehler

Def 6.13: Sei $(\mathcal{X}, \mathcal{F}, (P_\vartheta, \vartheta \in \Theta))$ ein stat. Modell und

$\tau: \Theta \to \mathbb{R}$ eine reelle Kenngröße. Ein Schätzer $T: \mathcal{X} \to \mathbb{R}$

für τ heißt erwartungstreu, falls

$$\mathbb{E}_\vartheta[T] = \tau(\vartheta) \qquad \forall \vartheta \in \Theta$$

Die Differenz $B_\vartheta(T) = \mathbb{E}_\vartheta[T] - \tau(\vartheta)$ heißt der Bias oder systematische Fehler von T.

In Bsp 6.8 und 6.9 haben wir gesehen, dass ML-Schätzer nicht immer erwartungstreu sind. Auch die empirische Varianz ist i.A. kein erwartungstreuer Schätzer.

Satz 6.14: Es sei $(\mathbb{R}^n, \mathcal{B}_n, (P_\vartheta^n, \vartheta \in \Theta))$ ein n-faches Produktmodell. Für jedes $\vartheta \in \Theta$ sei der Erwartungswert $m(\vartheta)$ und die Varianz $V(\vartheta)$ der Vtlg P_ϑ definiert und endlich (d.h. falls eine ZV X die Vtlg P_ϑ hat gilt

$$\mathbb{E}_\vartheta[X] = m(\vartheta) < \infty, \ \mathrm{Var}_\vartheta[X] = V(\vartheta) < \infty).$$

Dann sind der empirische Mittelwert $\bar{x}_n = \frac{1}{n}\sum_{i=1}^n x_i$ und die korrigierte Stichprobenvarianz $v_n^* = \frac{1}{n-1}\sum_{i=1}^n (x_i - \bar{x}_n)^2 = \frac{n}{n-1} v_n$

erwartungstreuen Schätzer für $m(\vartheta)$ und $V(\vartheta)$.

Bew.: 1. $\mathbb{E}_\vartheta[\bar X_n] = \frac{1}{n} \sum_{i=1}^n \mathbb{E}_\vartheta[X_i] = \frac{1}{n} \cdot n \cdot m(\vartheta) = m(\vartheta)$ $\quad\forall \vartheta\in\Theta$

2. Da $\mathbb{E}_\vartheta[X_i - \bar X_n] = 0$ folgt

$$\mathbb{E}_\vartheta[V_n^*] = \frac{1}{n-1}\sum_{i=1}^n \mathbb{E}_\vartheta[(X_i-\bar X_n)^2]$$

$$= \frac{1}{n-1}\sum_{i=1}^n Var_\vartheta(X_i-\bar X_n) = \frac{n}{n-1} Var_\vartheta(X_1-\bar X_n)$$

$$= \frac{n}{n-1} Var_\vartheta\left(\frac{n-1}{n}X_1 - \frac{1}{n}\sum_{i=2}^n X_i\right)$$

$$\overset{X_1 \text{ unab. von } (X_2,\dots,X_n)}{=} \frac{n}{n-1}\left(Var_\vartheta\left(\frac{n-1}{n}X_1\right) + Var_\vartheta\left(\frac{1}{n}\sum_{i=2}^n X_i\right)\right)$$

$$= \frac{n}{n-1}\left(\left(\frac{n-1}{n}\right)^2 V(\vartheta) + \frac{n-1}{n^2}V(\vartheta)\right)$$

$$= \frac{n}{n-1}\cdot\frac{n-1}{n} V(\vartheta) = V(\vartheta). \quad \square$$

Intuitive Erklärung: V_n unterschätzt $V(\vartheta)$ systematisch, da die Streuung bereits durch die Wahl von $\bar X_n$ (für $m(\vartheta)$) klein gemacht wird.

Ein Maß für die Qualität von Schätzern ist der mittlere quadratische Fehler:

Def 6.15: Sei $(\mathcal{X}, \mathcal{F}, (P_\vartheta, \vartheta\in\Theta))$ ein stat. Modell und $\tau: \Theta\to\mathbb{R}$ eine reelle Kenngröße. Der mittlere quadr. Fehler eines Schätzers $T:\mathcal{X}\to\mathbb{R}$ für τ ist geg. durch

$$R_\vartheta(T) := \mathbb{E}_\vartheta[(T-\tau(\vartheta))^2]$$

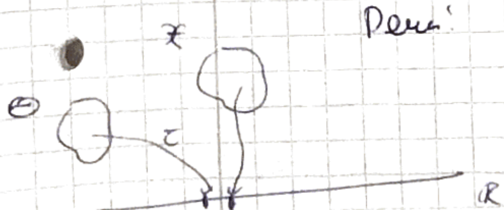
Man sieht leicht, dass

$$R_\vartheta(T) = Var_\vartheta(T) + B_\vartheta(T)^2$$

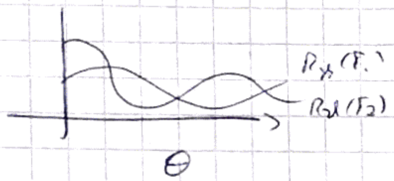
Denn: $R_\vartheta(T) = \mathbb{E}_\vartheta[T^2] - 2\mathbb{E}_\vartheta[T]\tau(\vartheta) + \tau(\vartheta)^2$

$= Var_\vartheta(T) + \mathbb{E}_\vartheta[T]^2 - 2\mathbb{E}_\vartheta[T]\tau(\vartheta) + \tau(\vartheta)^2$

$= Var_\vartheta(T) + B_\vartheta(T)^2$



Bem 6.16: Das Kriterium, dass $R_\vartheta(T)$ möglichst klein sein soll, liefert i.A. <u>keine</u> Totalordnung für die Schätzer von τ.



T_1 kann für gewisse Werte von ϑ besser sein als T_2, für andere schlechter.

Es gibt zwei Ansätze, T_1 und T_2 bzgl $R_\vartheta(\cdot)$ zu vgl.

1. Bayes-Ansatz:
Geg. eine sogenannte a-priori-VtG $Q(d\vartheta)$ auf Θ, minimiere den <u>erwarteten</u> mittleren quadr. Schätzfehler bzgl ϑ, d.h. minimiere $\int R_\vartheta(T)\, Q(d\vartheta)$

2. Minimax-Ansatz (aus der Spieltheorie)
minimiere: $\max_{\vartheta \in \Theta} R_\vartheta(T)$ („worst case analysis")

6.5. Varianzminimierende Schätzer

Wir betrachten Schätzer, die erwartungstreu sind und unter allen erwartungstreuen Schätzern optimal sind in dem Sinne, dass sie am wenigsten streuen.

Def. 6.17 Sei $(\mathcal{X}, \mathcal{F}, (P_\vartheta, \vartheta \in \Theta))$ ein stat. Modell, $\tau: \Theta \to \mathbb{R}$ eine Kenngröße. Ein erwartungstreuer Schätzer T für $\tau(\vartheta)$ heißt <u>Varianz-minimierend</u> oder <u>glm. bester Schätzer</u>, falls für jeden erwartungstreuen Schätzer S gilt: $Var_\vartheta(T) \le Var_\vartheta(S) \; \forall \vartheta \in \Theta$.

Bem 6.18: Da T und S erwartungstreu, gilt

$$Var_\vartheta(T) = R_\vartheta(T), \quad Var_\vartheta(S) = R_\vartheta(S).$$

Es kann nicht erwartungstreue Schätzer $\tilde{T}$ geben mit
$Var_\vartheta(\tilde{T}) < Var_\vartheta(T)$

83

Def 6.19: Ein einparametrisches stat. Modell $(\mathcal{X}, \mathcal{F}, (P_\vartheta, \vartheta \in \Theta))$ heißt regulär, falls Θ ein offenes Intervall in $\mathbb{R}$ ist ($\Theta = \mathbb{R}$ ist erlaubt) und folgende Bedingungen erfüllt sind:

i) Die Likelihoodfkt L ist auf $\mathcal{X} \times \Theta$ strikt positiv und stetig nach ϑ diffbar. Insbesondere existiert die Scorefkt U_ϑ geg. durch

$$U_\vartheta(x) = \frac{d}{d\vartheta} \log L(x, \vartheta) = \frac{\frac{d}{d\vartheta} L(x, \vartheta)}{L(x, \vartheta)}$$

ii) Für jedes $\vartheta \in \Theta$ existiert die Varianz $I(\vartheta) = Var_\vartheta(U_\vartheta)$ und ist strikt positiv und es gibt die Vertauschungsrelation

$$\int_\mathcal{X} \frac{d}{d\vartheta} L(x, \vartheta)\, dx = \frac{d}{d\vartheta} \int_\mathcal{X} L(x, \vartheta)\, dx$$

(falls $\mathcal{X}$ diskret: $\sum_{x \in \mathcal{X}} \frac{d}{d\vartheta} L(x, \vartheta) = \frac{d}{d\vartheta} \sum_{x \in \mathcal{X}} L(x, \vartheta)$)

Die Fkt $I : \Theta \to (0, \infty)$ heißt die Fisher-Information des Modells.

Bem 6.20: Falls (ii) in Def 6.19 erfüllt ist, gilt:

$$\mathbb{E}_\vartheta(U_\vartheta) = \int_\mathcal{X} \frac{\frac{d}{d\vartheta} L(x, \vartheta)}{L(x, \vartheta)} \underbrace{f_\vartheta(x)}_{= L(x, \vartheta)} dx = \int_\mathcal{X} \frac{d}{d\vartheta} L(x, \vartheta) dx$$

$$\overset{ii)}{=} \frac{d}{d\vartheta} \underbrace{\int L(x, \vartheta)\, dx}_{=1} = 0.$$

$$\hookrightarrow I(\vartheta) = Var(U_\vartheta) = \mathbb{E}_\vartheta(U_\vartheta^2) = \int_\mathcal{X} \left(\frac{\frac{\partial}{\partial\vartheta} L(x, \vartheta)}{L(x, \vartheta)} \right)^2 L(x, \vartheta)\, dx$$

$$\overset{f_\vartheta(x) = L(x, \vartheta)}{=} \int_\mathcal{X} \frac{\left(\frac{\partial}{\partial\vartheta} f_\vartheta(x)\right)^2}{f_\vartheta(x)} dx$$

(analog für den diskreten Fall).

Bsp 6.21: a) $\mathcal{X} = \{0, 1, \dots\}$, $\Theta = (0, \infty)$, $P_\vartheta \sim Poi(\vartheta)$

Fisher-Information?

(Bsp)

$$L(x,\vartheta) = e^{-\vartheta} \frac{\vartheta^x}{x!}$$

$$\log L(x,\vartheta) = -\vartheta + x \log \vartheta - \log(x!)$$

$$U_\vartheta(x) = \frac{d}{d\vartheta} \log L(x,\vartheta) = -1 + \frac{x}{\vartheta}$$

$$I(\vartheta) = Var_\vartheta(U_\vartheta) = \frac{1}{\vartheta^2} Var_\vartheta(X) = \frac{1}{\vartheta^2} \cdot \vartheta = \frac{1}{\vartheta}$$

b) $\mathcal{X} = \mathbb{R}$, $\Theta = \mathbb{R}$, $\sigma > 0$ fix, $P_\vartheta \sim \mathcal{N}(\mu, \sigma^2)$

$\hookrightarrow$ $L(x,\vartheta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\vartheta)^2}{2\sigma^2}\right)$

$$U_\vartheta(x) = -\frac{x-\vartheta}{\sigma^2}, \quad I(\vartheta) = \frac{1}{\sigma^2} \quad \forall \vartheta \in \Theta.$$

Lemma 6.22: (Additivität der Fisher-Information).
Sei $(\mathcal{X}, \mathcal{F}, (P_\vartheta, \vartheta \in \Theta))$ ein reguläres stat. Modell mit Fisher Info I. Dann hat das zugehörige n-fache Produktmodell Fisher Information $I_n = n \cdot I$

Bew: Die Likelihood-Fkt $L^{(n)}$ in Produktmodell ist geg. durch

$$L^{(n)}(x,\vartheta) = \prod_{i=1}^n L(x_i,\vartheta) \quad \text{für } x = (x_1, \dots, x_n)$$

$$\hookrightarrow U_\vartheta^{(n)} = \frac{d}{d\vartheta} \log L^{(n)}(x,\vartheta) = \sum_{i=1}^n U_\vartheta(x_i)$$

$$\hookrightarrow I_\vartheta^{(n)} = Var(U_\vartheta^{(n)}) \underset{\text{wg Produktmodell}}{\overset{\text{unabh.}}{=}} \sum_{i=1}^n Var_\vartheta(U_\vartheta(X_i))$$

$$= n \, Var_\vartheta(U_\vartheta) = n \, I(\vartheta).$$

Def 6.23. Ein erwartungstreuer Schätzer T für $\tau(\vartheta)$ heißt regulär, falls für jedes $\vartheta \in \Theta$ gilt:

$$\int_{\mathcal{X}} T(x) \frac{d}{d\vartheta} L(x,\vartheta)\, dx = \frac{d}{d\vartheta} \int T(x)\, L(x,\vartheta)\, dx$$

Satz 6.24. (Informationsungl. von Rao-Cramér)
Sei $(\mathcal{X}, \mathcal{F}, (P_\vartheta, \vartheta \in \Theta))$ ein reguläres stat. Modell, $\tau : \Theta \to \mathbb{R}$ sei stet. diff.bar Fkt mit $\tau' \neq 0$

und T sei regulärer erwartungstreuer Schätzer für $\tau(\vartheta)$. Dann gilt:

$$\mathrm{Var}_\vartheta(T) \geq \frac{(\tau'(\vartheta))^2}{I(\vartheta)} \quad \forall \vartheta \in \Theta \qquad (*)$$

„=" gilt für ϑ $\iff$ $T - \tau(\vartheta) = \frac{\tau'(\vartheta)}{I(\vartheta)} U_\vartheta \quad \forall \vartheta \in \Theta$,

d.h. wenn die Likelihood-Fkt die Form

$$L(x,\vartheta) = e^{a(\vartheta)T(x) - b(\vartheta)} h(x) \qquad (**)$$

hat, mit $a: \Theta \to \mathbb{R}$, a Stammfkt von $\frac{I}{\tau'}$, $h: \mathcal{X} \to (0,\infty)$ und $b(\vartheta) = \log \int_{\mathcal{X}} \exp(a(\vartheta)T(x)) \cdot h(x)\, dx$.

Wir behandeln den stetigen Fall.

Bew: Zuerst:

$$\mathrm{Cov}_\vartheta(T, U_\vartheta) = \mathbb{E}_\vartheta[T \cdot U_\vartheta] - \mathbb{E}_\vartheta[T]\underbrace{\mathbb{E}_\vartheta[U_\vartheta]}_{=0 \text{ (Bem 6.20)}}$$

$$= \int_{\mathcal{X}} T(x)\left(\frac{d}{d\vartheta}\log L(x,\vartheta)\right)\underbrace{f_\vartheta(x)}_{=L(x,\vartheta)}dx$$

$$= \int_{\mathcal{X}} T(x)\frac{d}{d\vartheta}L(x,\vartheta)\,dx \overset{T\text{ regulär}}{=} \frac{d}{d\vartheta}\underbrace{\int_{\mathcal{X}} T(x)L(x,\vartheta)\,dx}_{=\mathbb{E}_\vartheta[T]}$$

$$\overset{T\text{ erwartungstreu}}{=} \frac{d}{d\vartheta}\tau(\vartheta) = \tau'(\vartheta)$$

Setze nun $c(\vartheta) = \frac{\tau'(\vartheta)}{I(\vartheta)}$ ~~und~~ ~~erhalte~~ Dann t

$$0 \leq \mathrm{Var}_\vartheta(T - c(\vartheta)U_\vartheta)$$

$$= \mathrm{Var}_\vartheta(T) + c(\vartheta)^2\mathrm{Var}_\vartheta(U_\vartheta) - 2c(\vartheta)\mathrm{Cov}_\vartheta(T, U_\vartheta)$$

$$= \mathrm{Var}_\vartheta(T) + c(\vartheta)^2 I(\vartheta) - 2c(\vartheta)\tau'(\vartheta)$$

$$= \mathrm{Var}_\vartheta(T) + \left(\frac{\tau'(\vartheta)}{I(\vartheta)}\right)^2 I(\vartheta) - 2\frac{(\tau'(\vartheta))^2}{I(\vartheta)}$$

$$= \mathrm{Var}_\vartheta(T) - \frac{(\tau'(\vartheta))^2}{I(\vartheta)} \implies (*).$$

(86). Zusätzlich gilt "=" $\Leftrightarrow$ $T - c(\vartheta) U_\vartheta$ bzgl P_ϑ konstant ist und daher ~~konstant~~ insbesondere mit ihrem Erwartungswert übereinstimmt, der wegen $E_\vartheta(U_\vartheta) = 0$ mit $\tau(\vartheta)$ identisch ist. Da das stat Modell regulär ist folgt $L(x,\vartheta) > 0$ und

$$T(x) - \tau(\vartheta) = c(\vartheta) \underbrace{U_\vartheta(x)}_{\frac{d}{d\vartheta}\log L(x,\vartheta)} \quad \forall x \in \mathcal{X}, \forall \vartheta \in \Theta.$$

Das ist equivalent zu

$$\frac{d \log}{d\vartheta} L(x,\vartheta) = \frac{I(\vartheta)}{\tau'(\vartheta)} (T(x) - \tau(\vartheta)).$$

Mit Integration folgt (**):

$$b(\vartheta) = -\int_\Theta \tau(\vartheta) \frac{I(\vartheta)}{\tau'(\vartheta)} d\vartheta, \quad \log h(x) \text{ ist}$$

Integrationskonstante, $a(\cdot)$ ist Stammfkt von $\frac{I}{\tau'}$

Außerdem gilt: $b(\vartheta) = \log \int_\mathcal{X} e^{a(\vartheta) T(x)} h(x)\, dx$

Die Umkehrung folgt direkt: Falls $L(x,\vartheta)$ die Form (*) hat, so gilt

$$U_\vartheta = a'(\vartheta) T(x) - b'(\vartheta) \overset{(**)}{=} \frac{I(\vartheta)}{\tau'(\vartheta)} T(x) - \tau(\vartheta) \frac{I(\vartheta)}{\tau'(\vartheta)}$$

also $T(x) - c(\vartheta) U_\vartheta = \tau(\vartheta) \quad \forall x \in \mathcal{X}, \vartheta \in \Theta$.

(**) gilt, da $b(\vartheta) = \log \int e^{a(\vartheta) T(x)} h(x)\, dx$.

Damit:

$$b'(\vartheta) = \frac{\int_\mathcal{X} a'(\vartheta) T(x) e^{a(\vartheta) T(x)} h(x)\, dx}{\int_\mathcal{X} e^{a(\vartheta) T(x)} h(x)\, dx}$$

$$= a'(\vartheta) \frac{\int_\mathcal{X} T(x) e^{a(\vartheta) T(x) - b(\vartheta)} h(x) dx}{\int_\mathcal{X} \underbrace{e^{a(\vartheta) T(x) - b(\vartheta)} h(x)}_{L(x,\vartheta)} dx} = a'(\vartheta) \tau(\vartheta) \quad \square$$

Folgerungen aus Satz 6.24:

1. Bei n-facher unabh. Wdh. eines Experimentes, das durch ein reguläres stat. Modell beschrieben wird, gilt für jeden reg. erwartungstreuen Schätzer T weg. Add. der Fisher Info

$$\mathrm{Var}_\vartheta(T) \geq \frac{1}{n} g(\vartheta)$$

mit einer geeigneten Fkt $g: \Theta \to (0,\infty)$. In Bsp 6.8 ($\max x_i$) und 6.9 (Gauss-Strom) hatten wir Schätzer gesehen, deren Varianz die Größenordnung $\frac{1}{n^2}$ hat. Diese Modelle sind nicht regulär. (Dichte war nicht überall positiv!)

2. Falls die Likelihoodfkt von der Form $(**)$ in Satz 6.24 ist, so ist T, falls T erwartungstreu für $\tau(\vartheta)$, Varianz minimierend (zumindest in der Klasse der regulären Schätzer für τ). Unter weiteren (milden) Regularitätskriterien kann man zeigen, dass T in dieser Situation varianzminimierend ist ($\to$ Vorlesung "Statistik").

3. Sei $\mathcal{M} = (\mathcal{X}, \mathcal{F}, (P_\vartheta; \vartheta \in \Theta))$ reg. stat. Modell mit Fisher Information I, τ sei wie in Satz 6.24 und $\tilde{T}$ sei reg. Schätzer erwartungstreu für τ, der die Cramér-Rao Schranke erreicht, d.h.

$$\mathrm{Var}_\vartheta(\tilde{T}) = \frac{\tau'(\vartheta)^2}{I(\vartheta)} \quad \forall \vartheta \in \Theta$$

Sei $\mathcal{M}^n = (\mathcal{X}^n, \mathcal{F}^n, (P_\vartheta^n, \vartheta \in \Theta))$ das zugehörige Produktmodell. Dann ist

$$T(x) = \frac{1}{n}\sum_{i=1}^n \tilde{T}(x_i), \quad x = (x_1, \dots, x_n)$$

regulärer erwartungstreuer Schätzer für τ in $\mathcal{M}^n$, der die Cramér-Rao Schranke erreicht, d.h.

(88) $\quad \operatorname{Var}_\vartheta(T) = \frac{\tau'(\vartheta)^2}{n I(\vartheta)} = \frac{\tau'(\vartheta)^2}{I^{(n)}(\vartheta)}$

Bew. Nach Satz 6.24 folgt

$$L(x_i, \vartheta) = e^{a(\vartheta) T(x_i) - b(\vartheta)} h(x_i)$$

Damit

$$L^{(n)}(x, \vartheta) = e^{n a(\vartheta) T(x) - n b(\vartheta)} \prod_{i=1}^{n} h(x_i)$$

und $n a(\vartheta)$ ist Stammfkt von $\frac{n I(\vartheta)}{\tau'(\vartheta)} = \frac{I^{(n)}(\vartheta)}{\tau'(\vartheta)}$

und damit gilt "=" für T in (*) in Satz 6.24. □

Bsp 6.25: Poisson Modell:

Sei $\mathcal{X} = \{0,1,2,\dots\}$, $\Theta = (0,\infty)$, $P_\vartheta \sim Poi(\vartheta)$

$\hookrightarrow L(x, \vartheta) = e^{-\vartheta} \frac{\vartheta^x}{x!} = e^{x \log \vartheta - \vartheta} \frac{1}{x!}$

Damit hat $L(x, \vartheta)$ die erwünschte Form (**) in Satz 6.24

mit $T(x) = x$, $a(\vartheta) = \log(\vartheta)$. Für $\tau(\vartheta) = \vartheta$

gilt $\tau'(\vartheta) = 1$ und ~~damit~~ da $I(\vartheta) = \frac{1}{\vartheta}$ (Bsp 6.21 b))

folgt $a'(\vartheta) = \frac{1}{\vartheta} = \frac{I(\vartheta)}{\tau'(\vartheta)}$

$\Rightarrow T(x) = x$ ist erwartungstreuer Schätzer für $\tau(\vartheta) = \vartheta$, der die Rao-Cramér Schranke erreicht, insbesondere ist T Varianz minimierend.

Bsp 6.26:

$\mathcal{X} = \mathbb{R}$, $\Theta = \mathbb{R}$, $\sigma > 0$ fest, $P_\vartheta = \mathcal{N}(\vartheta, \sigma^2)$

Dann $L(x, \vartheta) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\vartheta)^2}{2\sigma^2}}$

$$= \exp\left(-\frac{1}{2} \log(2\pi\sigma^2) - \frac{\vartheta^2}{2\sigma^2} + \frac{\vartheta}{\sigma^2} x - \frac{x^2}{2\sigma^2}\right)$$

Also von der Form (**) in Satz 6.24 mit $T(x) = x$,

$a(\vartheta) = \frac{\vartheta}{\sigma^2}$, $b(\vartheta) = \frac{\vartheta^2}{2\sigma^2} - \frac{1}{2} \log(2\pi\sigma^2)$, $h(x) = e^{-\frac{x^2}{2\sigma^2}}$

Für $\tau(\vartheta) = \vartheta$ gilt, $a'(\vartheta) = \frac{1}{\sigma^2} = \frac{I(\vartheta)}{\tau'(\vartheta)}$ (vgl Bsp 6.21 iii))

Also ist $T(x) = x$ erwartungstreuer Schätzer für $\tau(\vartheta) = \vartheta$ und Varianz-minimierend:

$$Var_\vartheta(T) = \frac{1}{\sigma^2}.$$

6.6. Konsistenz:

Def 6.27: (Konsistente Folge von Schätzern)

Für jedes $n \in \mathbb{N}$ sei $(\mathcal{X}_n, \mathcal{F}_n, (P_{\vartheta,n}, \vartheta \in \Theta))$ ein stat. Modell mit Parametermenge Θ und T_n ein Schätzer für ~~die~~ ein reelle Kenngröße τ im n-ten Modell.

Die Folge von Schätzern $(T_n)_n$ heißt konsistent, falls für jedes $\vartheta \in \Theta$ gilt:

$$\lim_{n\to\infty} P_{\vartheta,n}[\,|T_n - \tau(\vartheta)| \geq \varepsilon\,] = 0 \qquad \forall \varepsilon > 0.$$

Bsp: $(\mathcal{X}_n, \mathcal{F}_n, (P_{\vartheta,n}, \vartheta \in \Theta)) = (\mathcal{X}^n, \mathcal{F}^n, (P_\vartheta^n, \vartheta \in \Theta))$

Bsp 6.28: Raten des Bereichs von Zufallszahlen (vgl Bsp 6.9)

n-faches Produktmodell mit $\mathcal{X} = [0, \infty)$, $\mathcal{F} = \mathcal{B}_1$, $P_\vartheta = \mathrm{Unif}([0, \vartheta])$. Haben gesehen, dass $T_n(x) = \max_{1 \leq i \leq n} x_i$ und $R_n = \frac{2}{n} \sum_{i=1}^n x_i$ konsistente Folgen von Schätzern sind.

Allgemeiner gilt:

Ist T_n der empirische Mittelwert von n unabh. Beob., so folgt die Kvgz von T_n aus dem schwachen Gesetz der großen Zahlen. Dies gilt ebenfalls für die empirische Varianz.

Satz 6.29:

Sei $(\mathcal{X}, \mathcal{B}, (P_\vartheta, \vartheta \in \Theta))$ stat. Modell und für jedes ϑ existiere der Erwartungswert $m(\vartheta)$ und die Varianz $V(\vartheta)$ und sei endlich (vgl. Satz 6.14). Wir nehmen an, dass für eine ZV X_ϑ mit Vtlg P_ϑ das vierte Moment endlich ist: $\mathbb{E}_\vartheta X_\vartheta^4 < \infty \quad \forall \vartheta \in \Theta$.
Betrachte im n-fachen Produktmodell $(\mathcal{X}^n, \mathcal{B}^n, (P_\vartheta^n, \vartheta \in \Theta))$ die Schätzer $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ für $m(\vartheta)$ und $V_n = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$ für $V(\vartheta)$. Dann sind die Folgen $(\bar{X}_n)$ und (V_n) konsistent.

Bew.: 1. Das schwache Gesetz der großen Zahlen impliziert für alle $\vartheta \in \Theta$, alle $\varepsilon > 0$, dass

$$P_\vartheta^n[\,|\bar{X}_n - m(\vartheta)| \geq \varepsilon\,] \xrightarrow[n\to\infty]{} 0$$

2. Setze $\tilde{V}_n = \frac{1}{n}\sum_{i=1}^{n}(X_i - m(\vartheta))^2$. Dann gilt

$$V_n = \tilde{V}_n - (\bar{X}_n - m(\vartheta))^2$$ (vgl. Verschiebungsformel Satz 6.12)

Schwaches Gesetz der großen Zahlen impliziert $\forall \vartheta\ \forall \varepsilon > 0$

$$P_\vartheta^n[\,|\tilde{V}_n - V(\vartheta)| \geq \varepsilon\,] \xrightarrow[n\to\infty]{} 0$$

(mit $Y_i = (X_i - m(\vartheta))^2$, beachte $\mathbb{E}_\vartheta[Y_i^2] < \infty$ da das vierte Moment endlich ist).

Mit 1. folgt mit

$$|V_n - V(\vartheta)| \leq |\tilde{V}_n - V(\vartheta)| + (\bar{X}_n - m(\vartheta))^2,$$

dass

$$P_\vartheta^n[\,|V_n - V(\vartheta)| \geq \varepsilon\,] \leq P_\vartheta^n[\,|\tilde{V}_n - V(\vartheta)| \geq \tfrac{\varepsilon}{2}\,] + P_\vartheta^n[\,|\bar{X}_n - m(\vartheta)|^2 \geq \tfrac{\varepsilon}{2}\,] \xrightarrow[n\to\infty]{} 0$$

$$(x + y \geq \varepsilon \Rightarrow x \geq \varepsilon/2 \text{ oder } y \geq \varepsilon/2)$$ □

Bem.: $V_n^* = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$ ist ebenfalls konsistent für $V(\vartheta)$.

Bsp 6.30: Sei $(\mathcal{X}^n, \mathcal{F}^n, (P_\vartheta^n, \vartheta \in \Theta))$ das n-fache Produktmodell mit $\mathcal{X} = \{0, 1, 2, \dots\}$, $P_\vartheta = Poi(\vartheta)$, $\Theta = (0, \infty)$. Dann gilt $m_4(\vartheta) = E_\vartheta[X^4] = \sum_{k=0}^{\infty} k^4 e^{-\vartheta} \frac{\vartheta^k}{k!} < \infty$ für alle $\vartheta \in \Theta$. Damit ist die Folge $T_n = \overline{X}_n$ mit $\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$ eine konsistente Folge von Schätzern für $\tau(\vartheta) = \vartheta$.

Bsp 6.31: Wir wollen die mittlere Lebensdauer von Glühbirnen schätzen. Gegeben sind n Messungen $x_1, \dots, x_n$, wobei x_i = Lebensdauer der i-ten Glühbirne.

Sei $(\mathcal{X}^n, \mathcal{F}^n, (P_\vartheta, \vartheta \in \Theta))$ das n-fache Produktmodell mit $\mathcal{X} = (0, \infty)$, $\mathcal{F} = \mathcal{B}_1$ (Borel σ-Algebra), P_ϑ sei die Exponentialvtlg mit Parameter ϑ, d.h. P_ϑ hat Dichte

$$f_\vartheta(x) = \begin{cases} \vartheta e^{-\vartheta x} & , x \geq 0 \\ 0 & x < 0 \end{cases}$$

Der Erwartungswert ist $E_\vartheta X = \frac{1}{\vartheta}$. Also folgt aus dem schwachen Gesetz der großen Zahlen $\forall \varepsilon > 0$

$$P_\vartheta^n\left[\left|\frac{1}{n} \sum_{i=1}^{n} X_i - \frac{1}{\vartheta}\right| \geq \varepsilon\right] \xrightarrow[n \to \infty]{} 0 \quad \forall \vartheta \in \Theta.$$

Damit ist $\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$ eine konsistente Folge für $\tau(\vartheta) = \frac{1}{\vartheta}$. Die Stetigkeit der Funktion $z \mapsto \frac{1}{z}$ für $z \in (0, \infty)$ impliziert dann, dass für alle $\varepsilon > 0$ ein $\delta > 0$ existiert, so dass

$$\left|\frac{1}{\frac{1}{n}\sum_{i=1}^{n} X_i} - \frac{1}{\frac{1}{\vartheta}}\right| \leq \delta \quad \Rightarrow \quad \left|\frac{1}{n} \sum_{i=1}^{n} X_i - \frac{1}{\vartheta}\right| \leq \varepsilon$$

$\Rightarrow \forall \delta > 0$ folgt

$$P_\vartheta^n\left[\left|\frac{1}{\frac{1}{n} \sum_{i=1}^{n} X_i} - \frac{1}{\frac{1}{\vartheta}}\right| \geq \delta\right]$$

$$= P_\vartheta^n\left[\left|\frac{n}{\sum_{i=1}^{n} X_i} - \vartheta\right| \geq \delta\right] \xrightarrow[n \to \infty]{} 0 \quad \forall \vartheta \in \Theta.$$

(92)

Also ist $(T_n)_n(X) = \left(\frac{n}{\sum_{i=1}^{n} X_i} \right)_n$ eine konstante Folge für $\tau(\vartheta) = \vartheta$.

7. Konfidenzbereiche

Ein Schätzer für eine unbekannte Kenngröße liefert einen ersten Anhaltspunkt für den "wahren Wert", macht aber keine Aussage über die Zuverlässigkeit, d.h. über mögliche Abweichungen vom wahren Wert.

Betrachten wir nochmals Bsp 6.11: Ein Reißnagel fällt mit W'keit $\vartheta \in [0,1]$ auf die Spitze, $P_\vartheta \sim Bin(n,\vartheta)$ und wir haben gesehen, dass $T_n = \frac{x}{n}$ ein ML-Schätzer ist. Falls jedoch ϑ irrational ist, so gilt $T_n \neq \vartheta$ $\forall n, \forall x$.

Besser ist es also einen (von den Beob. abhängigen) Bereich anzugeben, in dem die Kenngröße mit vorgeg. W'keit $1-\alpha$ enthalten ist.

D.h. wir machen die Aussage <u>ungenauer</u>, aber dafür <u>zuverlässiger</u>.

Def 7.1 (Konfidenzbereiche):

Sei $(\mathcal{X}, \mathcal{F}, (P_\vartheta, \vartheta \in \Theta))$ ein stat. Modell, $\tau: \Theta \to \mathbb{R}$ eine reelle Kenngröße und $\alpha \in (0,1)$ eine Fehlerschranke. Eine Familie $(C(x))_{x \in \mathcal{X}}$ von Mengen $C(x) \subset \mathbb{R}$ heißt ein Konfidenzbereich für τ zum Irrtumsniveau α (bzw. zum Sicherheitsniveau $1-\alpha$), falls für jedes $\vartheta \in \Theta$ gilt, dass

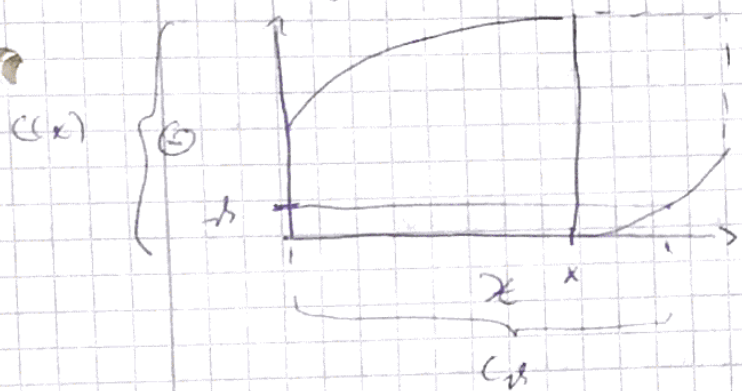
$P_\vartheta[\tau(\vartheta) \in C(X)] \geq 1-\alpha$, $(P_\vartheta[\tau(\vartheta) \in C(X)] = P_\vartheta[\{x \in \mathcal{X} : \tau(\vartheta) \in C(x)\}])$

Falls jedes $C(x)$ ein Intervall ist, so spricht man auch von einem Konfidenzintervall.

Bem 7.2: Konfidenzbereiche sind nur nützlich, wenn sie klein sind. Genauer, seien $(C(x))_{x\in\mathcal{X}}$ und $(C'(x))_{x\in\mathcal{X}}$ zwei Konfidenzbereiche zum Irrtumsniveau α mit $C(x) \subseteq C'(x)$ für alle x. So ist $(C(x))_{x\in\mathcal{X}}$ zu bevorzugen.

§2 Konstruktion möglichst kleiner Konfidenzbereiche.

Wir betrachten nur $\tau(\vartheta) = \vartheta$. Betrachte folgende zweidimensionale Menge: $C = \{(x,\vartheta) \in \mathcal{X}\times\Theta : \vartheta \in C(x)\}$



Für jedes $\vartheta \in \Theta$ ist

$C_\vartheta = \{x \in \mathcal{X} : \vartheta \in C(x)\}$

$= \{x \in \mathcal{X} : (x,\vartheta) \in C\}$

der horizontale Schnitt durch C

Die Bed. i Def 7.1

$$P_\vartheta[\{x \in \mathcal{X} : \tau(\vartheta) \in C(x)\}] \geq 1-\alpha \qquad \forall \vartheta \in \Theta$$

$$\iff P_\vartheta[C_\vartheta] \geq 1-\alpha \qquad \forall \vartheta \in \Theta.$$

Wir definieren also für jedes $\vartheta \in \Theta$ eine (möglichst kleine) Menge C_ϑ, so dass $P_\vartheta[C_\vartheta] \geq 1-\alpha$ $\forall \vartheta \in \Theta$. Auf diese Art konstruieren wir eine Familie $(C_\vartheta)_{\vartheta\in\Theta}$ von horizontalen Schnitten, und damit eine Menge C. Dann sind die Mengen $C(x)$ = vertikale Schnitte von C unsere Konfidenzbereiche. Ein Verfahren ist das ~~möglichste~~ folgende:

Konstruktion eines Konfidenzbereiches:

Sei $\alpha > 0$ und die Likelihoodfkt $L(x,\vartheta)$ gegeben.

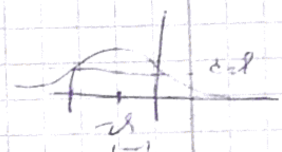
1. Für jedes $\vartheta \in \Theta$ bestimme eine Menge der Form $C_\vartheta = \{x \in \mathcal{X} : L(x,\vartheta) \geq d_\vartheta\}$, wobei $d_\vartheta > 0$

möglichst groß gewählt wird, so dass immer noch
$P_\vartheta[C_\vartheta] \geq 1-\alpha$ gilt.

2. Setze $C = \{(x,\vartheta) : x \in C_\vartheta\}$, dann bilden die x-Schnitte von C, $C(x) = \{\vartheta \in \Theta : (x,\vartheta) \in C\}$
$= \{\vartheta \in \Theta : x \in C_\vartheta\}$
einen Konfidenzbereich für $\tau(\vartheta) = \vartheta$ zum Irrtumsniveau α.

Bsp 7.3: Sei $\sigma > 0$, $(\mathcal{X}, \mathcal{F}, (P_\vartheta, \vartheta \in \Theta)) = (\mathbb{R}, \mathcal{B}_1, (\mathcal{N}(\vartheta, \sigma^2))_{\vartheta \in \Theta})$
mit $\Theta = \mathbb{R}$. Wir suchen Konfidenzbereich für $\tau(\vartheta) = \vartheta$.

$C_\vartheta = \{x \in \mathbb{R} : f_\vartheta(x) = L(x,\vartheta) \geq d_\vartheta\}$



$= [\vartheta - \gamma, \vartheta + \gamma]$

Damit $C(x) = \{\vartheta \in \Theta : x \in C_\vartheta\} = \{\vartheta \in \mathbb{R} : \vartheta - \gamma \leq x \leq \vartheta + \gamma\}$
$= [x - \gamma, x + \gamma]$

Also müssen wir γ so wählen, dass $P_\vartheta[\vartheta \in C(x)]$
$= P_\vartheta[C_\vartheta] \geq 1 - \alpha$

Also $P_\vartheta[[\vartheta - \gamma, \vartheta + \gamma]] \geq 1 - \alpha$

Aber $P_\vartheta[[\vartheta - \gamma, \vartheta + \gamma]] = P_0[[-\gamma, \gamma]]$

$$= \int_{-\gamma}^{\gamma} \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{x^2}{2\sigma^2}} dx \overset{!}{=} 1 - \alpha \quad \text{(Damit } C(x) \text{ klein wird)}$$

$$\Leftrightarrow \int_{-\infty}^{-\gamma} \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{x^2}{2\sigma^2}} dx = \frac{\alpha}{2} \quad (*)$$

Wähle γ so, (*) gilt. Dann ist $(C(x))_{x \in \mathbb{R}} = ([x - \gamma, x + \gamma])_{x \in \mathbb{R}}$
der kleinste Konfidenzbereich für $\tau(\vartheta) = \vartheta$ zum Irrtumsniveau α.

Def 7.4: (Quantil)
Sei Q ein W'maß auf $(\mathbb{R}, \mathcal{B})$ und $\alpha \in (0,1)$.
Jede Zahl $q \in \mathbb{R}$ mit $Q((-\infty, q]) \geq \alpha$ und
$Q([q, \infty)) \geq 1 - \alpha$ heißt ein <u>α-Quantil</u> von Q.

Ein $\frac{1}{2}$-Quantil von Q heißt Median von Q.

Quantile sind i.A. nicht eindeutig bestimmt. Falls Q eine Dichte f hat und $\{x : f(x) > 0\}$ ein Intervall ist, so ist das α-Quantil von Q die (eindeutig bestimmte) Zahl q mit $\int_{-\infty}^{q} f(x)\,dx = \alpha$.

Bsp. 7.5 Emissionskontrolle

Von $N = 10$ Kraftwerken überprüfen wir $n = 4$ zufällig ausgewählte auf ihre Energiewerte. Gesucht ist ein Konfidenzbereich für die Anzahl ϑ der Kraftwerke mit zu hohen Emissionswerten. X sei die Anzahl der Kraftwerke mit zu hohen Emissionswerten in der Stichprobe.

Also: Stichprobe der Größe 4 ohne Zurücklegen aus einer Urne mit zehn Kugeln, von denen eine unbekannte Anzahl ϑ schwarz ist. X = # schwarzen gezogenen Kugeln

$\mathcal{X} = \{0, 1, 2, 3, 4\}$, $\Theta = \{0, 1, \ldots, 10\}$

P_ϑ = Hypergeom. Vert. mit Parametern $N = 10$, $n = 4$, ϑ

d.h. $L(x, \vartheta) = \dfrac{\binom{\vartheta}{x}\binom{10-\vartheta}{4-x}}{\binom{10}{4}}$ für $\vartheta = x, x+1, \ldots, x+6$

Tabelle für $\binom{10}{4} L(x, \vartheta)$:

	$x=0$	$x=1$	$x=2$	$x=3$	$x=4$
$\vartheta=0$	210	0	0	0	0
$\vartheta=1$	126	84	0	0	0
$\vartheta=2$	70	112	28	0	0
$\vartheta=3$	35	105	63	7	0
$\vartheta=4$	15	80	90	24	1
$\vartheta=5$	5	50	100	50	5

(96)

Für $\vartheta > 5$ mit Symm. $L(x,\vartheta) = L(4-x, 10-\vartheta)$

Sei $\alpha = \frac{1}{5}$ vorgegeben. Wir brauchen für jedes ϑ so viele Werte x, bis $\sum_{x \in C_\vartheta} L(x,\vartheta) \geq \frac{4}{5} = 1-\alpha$,

also muss die Summe in einer Zeile (oben) ≥ 168 sein, da $168/\binom{10}{4} = \frac{168}{210} = \frac{4}{5}$.

Damit sind C_ϑ, $\vartheta \in \Theta$ und C bestimmt.

$C(x)$ ergibt sich indem man in der x-ten Spalte diejenigen ϑ betrachtet, deren Eintrag unterstrichen sind, d.h.

$$C(0) = \{0,1,2\}$$
$$C(1) = \{1,2,3,4,5\}$$
$$C(2) = \{3,4,5,6,7\}$$
$$C(3) = \{5,6,7,8,9\}$$
$$C(4) = \{8,9,10\} \quad \dots$$

Bsp 7.6: (Binomialmodell, vgl Bsp 6.11)

$\mathcal{X} = \{0, 1, \dots, 4\}$

$P_\vartheta = Bin(4, \vartheta)$

Gesucht ist ein Konfidenzbereich für $\tau(\vartheta) = \vartheta$ zu geg. $\alpha \in (0,1)$. Wir wissen, dass $T = \frac{x}{4}$ varianzminimierender Schätzer ist und machen daher den Ansatz:

$C(x) = (\frac{x}{4} - \varepsilon, \frac{x}{4} + \varepsilon)$ wobei wir ε so wählen, dass

$$\underbrace{P_\vartheta [\{ x : |\tfrac{x}{4} - \vartheta| \geq \varepsilon \}]}_{} \overset{!}{\leq} \alpha \qquad (*)$$

$$= \sum_{x : |\frac{x}{4} - \vartheta| \geq \varepsilon} \binom{4}{x} \vartheta^x (1-\vartheta)^{4-x} \overset{!}{\leq} \alpha$$

Wir suchen also das $1-\alpha$ Quantil für die

Verteilung von $\left|\frac{X}{n} - \vartheta\right|$ wobei $X \sim Bin(n, \vartheta)$.

Zwei Möglichkeiten zur Näherung:

1) Tschebyschev Ungleichung liefert:

$$\mathbb{P}_\vartheta\left[\left|\frac{X}{n} - \vartheta\right| \geq \varepsilon\right] \leq \frac{Var_\vartheta[X]}{n^2 \varepsilon^2}$$

$$= \frac{\vartheta(1-\vartheta)}{n \varepsilon^2} \leq \frac{1}{4n\varepsilon^2} \qquad \forall \vartheta \in \Theta$$

Damit ~~gilt~~ ist $(*)$ erfüllt, falls $\frac{1}{4n\varepsilon^2} < \alpha$, d.h.

$$\varepsilon > \frac{1}{\sqrt{4n\alpha}}$$

Für $n = 1000$ und $\alpha = 0{,}025$ genügt $\varepsilon > 0{,}1$

Da wir keine Eig. der Binomialvert. benutzt haben, ist ε i.A. zu groß.

2) Zentraler Grenzwertsatz liefert

$$\mathbb{P}_\vartheta\left[\left|\frac{X}{n} - \vartheta\right| < \varepsilon\right] = \mathbb{P}_\vartheta\left[\left|\frac{X - n\vartheta}{\sqrt{n\vartheta(1-\vartheta)}}\right| \leq \varepsilon \frac{\sqrt{n}}{\sqrt{\vartheta(1-\vartheta)}}\right]$$

$$\underset{\text{Moivre Laplace}}{\overset{\text{ZGWS}}{\approx}} 2\,\Phi\left(\varepsilon \sqrt{\frac{n}{\vartheta(1-\vartheta)}}\right) - 1 \overset{!}{\geq} 1 - \alpha$$

(wegen $\Phi(x) = 1 - \Phi(-x)$). Also wähle ε so, dass

$$\Phi\left(\varepsilon \sqrt{\frac{n}{\vartheta(1-\vartheta)}}\right) \geq 1 - \frac{\alpha}{2}$$

Da wir ϑ nicht kennen, wähle ε so, dass

$$\Phi(2\varepsilon\sqrt{n}) \geq 1 - \frac{\alpha}{2}.$$

Für $n = 1000$ und $\alpha = 0{,}025$ ergibt sich

$$\Phi(2\varepsilon\sqrt{n}) \geq 0{,}9875$$

$\rightsquigarrow$ $2\varepsilon\sqrt{n} \geq 2{,}24$ (Tabelle)

$\rightsquigarrow$ $\varepsilon > 0{,}04$ reicht!