Mathematische Statistik

Konkrete Stichprobe und mathematische Stichprobe

·) Vorstellung: Stichprobe als Auswertung von Daten

Beispiel: Seien $n \in \mathbb{N}$ und $x_i \in \mathbb{R}$, $i \in \mathbb{N}$, $i \leq n$. (z.B. Lebensdauer von Bauelementen)
oder $(x_i, y_i) \in \mathbb{R}^2$, $i \in \mathbb{N}$, $i \leq n$ (z.B. (Spannung, Stromstärke))
usw.

Bemerkung: ·) Hier meist der eindimensionale Fall.

<u>Definition:</u> (konkrete Stichprobe)

Sei $n \in \mathbb{N}$ und $x = (x_1, \dots, x_n) \in \mathbb{R}^n$. Dann ist x eine konkrete Stichprobe vom Umfang n.

Notation: $x_1, \dots, x_n$ konkrete Stichprobe $:\Longleftrightarrow$ $(x_1, \dots, x_n) \in \mathbb{R}^n$ konkrete Stichprobe vom Umfang n

·) Ziel ist es nun Aussagen über den Vorgang zu treffen, der die Stichprobe erzeugt hat.

Stochastisches Modell:

Grundannahme: Werte einer konkreten Stichprobe lassen als Realisierung von Zufallsvariablen auffassen

<u>Definition:</u> (mathematische Stichprobe)

Sei $(\Omega, \mathcal{D}, P)$ ein Wahrscheinlichkeitsraum. Seien für $n \in \mathbb{N}$ $X_1, \dots, X_n$ Zufallsvariablen auf $(\Omega, \mathcal{D}, P)$. Dann ist $X = (X_1, \dots, X_n)$ (auch nur $X_1, \dots, X_n$) eine mathematische Stichprobe vom Umfang n. Dann ist $(\mathbb{R}^n, \mathcal{B}_n)$ als Bildraum von X der Stichprobenraum.

Allgemeine Voraussetzung elementarer math. Statistik: $X_1, \dots, X_n$ sind unabhängig und identisch verteilt ($X_1, \dots, X_n$ i.i.d.)

$\Longrightarrow$ im Unterschied zur Wahrscheinlichkeitstheorie ist hier P nicht bekannt

Der statistische Raum

Ansatz: $P_{X_i} \in \{ P_\vartheta \mid \vartheta \in \mathcal{M} \}$ mit $\mathcal{M} \subset \mathbb{R}^l$ mit $l \geq 1$

Es existiert $\vartheta^* \in \mathcal{M}$, sodass $P_{X_i} = P_{\vartheta^*}$

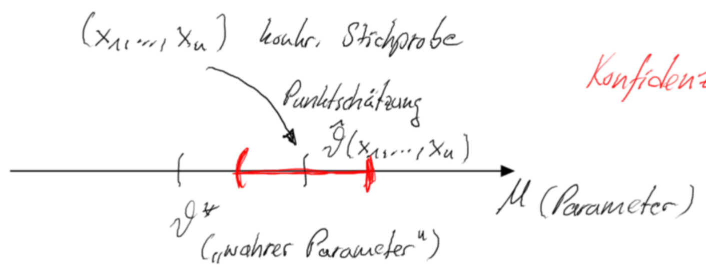
Definition: (Statistischer Raum)

Ein statistischer Raum (auch Experiment oder stat. Struktur) ist ein Tripel $(\mathbb{R}^n, \mathcal{R}_n, \{ P_\vartheta^{\otimes n} \mid \vartheta \in \mathcal{M} \})$, wobei $\mathcal{M} \subset \mathbb{R}^l$ mit $l \geq 1$ und P_ϑ ein Wahrscheinlichkeitsmaß auf $(\mathbb{R}, \mathcal{R})$ für alle $\vartheta \in \mathcal{M}$ ist.

Beispiel: ·) n-gliedriges Bernoulli-Schema

Klassifizierung parametrischer statistischer Verfahren: (nach Aussagen, die über wahren unbekannten Parameter gemacht werden)

- ·) Punktschätzungen
- ·) Konfidenzschätzung
- ·) Tests von Hypothesen



Konfidenzschätzung (Kugel mit wahrem Parameter)

Annahme: wahrer Parameter existiert

Statistiken / Stichprobenfunktion

Definition: (Statistik / Stichprobenfunktion)

Sei $(\mathcal{M}, \mathfrak{M})$ ein messbarer Raum mit $\mathcal{M} \neq \emptyset$. Dann ist eine $(\mathcal{R}_n, \mathfrak{M})$-messbare Funktion $T: \mathbb{R}^n \to \mathcal{M}$ eine Statistik.

Bemerkung: ·) Anwendung Statistik auf Stichprobe liefert meist Datenverdichtung, bei der teilweise Informationen verloren gehen
·) siehe suffiziente / erschöpfende Statistik

wichtige Statistiken: ·) Stichprobenmittel / empirischer Erwartungswert:

$$T: \mathbb{R}^n \to \mathbb{R}, \quad T(x_1, \ldots, x_n) = \frac{1}{n} \sum_{i=1}^n x_i \quad =: \overline{x}$$

·) r-tes Stichprobenmoment: $(x_1,\dots,x_n)$ konkrete Stichprobe mit Umfang n und $r \in \mathbb{N}$

$$T: \mathbb{R}^n \longrightarrow \mathbb{R}, \quad T(x_1,\dots,x_n) = \frac{1}{n} \sum_{i=1}^n x_i^r$$

·) Stichprobenstreuung / empirische Varianz:

$$T: \mathbb{R}^n \longrightarrow \mathbb{R}, \quad T(x_1,\dots,x_n) = \frac{1}{n} \sum_{i=1}^n (x_i - \overline{x})^2 =: \overline{\sigma}^2$$

·) korrigierte Stichprobenstreuung / korrigierte empirische Varianz:

$$T: \mathbb{R}^n \longrightarrow \mathbb{R}, \quad T(x_1,\dots,x_n) = \frac{1}{n-1} \sum_{i=1}^n (x_i - \overline{x})^2 = \frac{n}{n-1} \overline{\sigma}^2 =: \hat{\sigma}^2$$

·) konkrete geordnete Stichprobe / Variationsreihe:

$$T: \mathbb{R}^n \longrightarrow \mathbb{R}^n, \quad T(x_1,\dots,x_n) = (x_1', \dots, x_n') =: \sigma(x)$$

mit $x_1' \leq x_2' \leq \dots x_n'$ $\implies$ T ist Permutation, welche Stichprobe ordnet

·) i-te geordnete Statistik (Rangstatistik, Ranggröße): Sei $i \in \mathbb{N}$, $1 \leq i \leq n$

$$T: \mathbb{R}^n \longrightarrow \mathbb{R}, \quad T(x_1,\dots,x_2) = (\sigma(x_1,\dots,x_n))_i = x_i'$$

·) Spannweite der Stichprobe:

$$T: \mathbb{R}^n \longrightarrow \mathbb{R}, \quad T(x_1,\dots,x_n) = \sigma(x)_n - \sigma(x)_1 = x_n' - x_1'$$

·) Stichprobenmedian / empirische Zentralwert:

$$T: \mathbb{R}^n \longrightarrow \mathbb{R}, \quad T(x_1,\dots,x_n) := \tilde{x}_{\frac{1}{2}} := \begin{cases} \sigma(x)_{\frac{n+1}{2}} & : n \text{ ungerade} \\ \frac{1}{2}\left(\sigma(x)_{\frac{n}{2}} + \sigma(x)_{\frac{n}{2}+1}\right) & : \text{sonst} \end{cases}$$

·) Stichproben-α-Quantil / empirisches α-Quantil: $\alpha \in (0,1)$

$$T: \mathbb{R}^n \longrightarrow \mathbb{R}, \quad T(x_1,\dots,x_n) := \tilde{x}_\alpha := \begin{cases} \sigma(x)_{\lfloor n\alpha \rfloor + 1} & : n\alpha \notin \mathbb{Z} \\ \frac{1}{2}\left(\sigma(x)_{n\alpha} + \sigma(x)_{n\alpha+1}\right) & : \text{sonst} \end{cases}$$

·) empirische Verteilungsfunktion (relative Häufigkeit):

$$T: \mathbb{R}^n \longrightarrow \{F: \mathbb{R} \to \mathbb{R}\}, \quad T(x_1,\dots,x_n) = \frac{1}{n} \sum_{i=0}^n 1_{[x_i,\infty)} =: \hat{F}$$

·) Histogramm (der relativen Häufigkeiten): Sei $\Delta_1,\dots,\Delta_k$ eine Klasseneinteilung mit $k \in \mathbb{N}$ und Δ_i, Δ_j disjunkt für alle i,j, $i \neq j$

$$T: \mathbb{R}^n \longrightarrow \mathbb{R}^k, \quad T(x_1,\dots,x_n) = \left(\frac{1}{n} \sum_{i=1}^n 1_{\Delta_1}(x_i), \dots, \frac{1}{n} \sum_{i=1}^n 1_{\Delta_k}(x_i)\right)$$

$$\sum_{j=1}^k 1_{\Delta_j}(x_i) \overset{!}{=} 1$$

(Δ_i meist äquidistant)
Wahl von k nach Sturges: $k = 1 + 3{,}32 \lg n$

Simulation von Zufallsvariablen mit vorgegebener Verteilung

Sei zum Beispiel das Verteilungsgesetz der Zufallsvariable $T(X_1, \dots, X_n)$, $n \in \mathbb{N}$ gesucht.
Sei $m \in \mathbb{N}$ (sehr groß). Simulation von m Stichproben mit Umfang n

$$(x_1^{(1)}, \dots, x_n^{(1)}) \longrightarrow T(x_1^{(1)}, \dots, x_n^{(1)})$$
$$\vdots \qquad\qquad\qquad \vdots$$
$$(x_1^{(m)}, \dots, x_n^{(m)}) \longrightarrow T(x_1^{(m)}, \dots, x_n^{(m)})$$
$$\downarrow$$

empirische Verteilungsfunktion bzw. Histogramm

Notation:
Sei $T: \mathbb{R}^n \longrightarrow M$ die Statistik einer konkreten Stichprobe. Dann definiert man:

$$T(X_1, \dots, X_n): \Omega \to M$$
$$T(X_1, \dots, X_n)(\omega) = T(X_1(\omega), \dots, T(X_n(\omega)))$$

(Pseudo –) Zufallsgeneratoren:
- deterministischer Algorithmus, der bei Festlegung eines Startwertes ein eindeutiges n-Tupel erzeugt
- erzeugen eines n-Tupels mit „wesentlichen" Eigenschaften einer Folge von Realisierungen einer Zufallsvariable hat

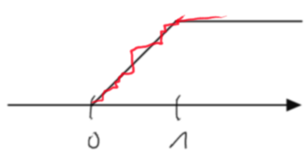
Kongruenzgeneratoren für stetige Gleichverteilung auf $(0,1)$

- Bestimmungsstücke:
  - Model: $M \in \mathbb{N}$
  - Mutiplikator: $a \in \mathbb{N}$
  - Verschiebung: $c \in \mathbb{N}_0$
  - Startwert: $x_0 \in \mathbb{N}_0$, $x_0 < M$
  - rekursive Rechenvorschrift: $\boxed{x_{n+1} = (a x_n + c) \bmod M}$ für alle $n \in \mathbb{N}$
  - Ermittlung der Gleichverteilung $(u_n)_{n \in \mathbb{N}}$: $\boxed{u_n = \frac{x_n}{M}}$ für alle $n \in \mathbb{N}$

$\Longrightarrow$ periodische Folge (u_n) (Periodenlänge höchstens M)

Beispiel: SAS (alt): $M = 2^{31} - 1$ (Primzahl)
$a = 397\,204\,094$
$c = 1$

Überprüfung der Eigenschaften:
- $\mathbb{E}u \overset{!}{=} \frac{1}{2}$
- $\operatorname{var} U \overset{!}{=} \frac{1}{12}$
- Anpassungstests



·) Schwachstellen können aufgrund der deterministischen Berechnung immer wieder auftreten (z.B. durch Transformationen)

$\Longrightarrow$ bei Verwendung von Zufallsgeneratoren: protokollieren, welcher Generator und wie (Startwert möglichst selbst festlegen)

Erzeugung von (Pseudo-) Zufallszahlen anderer Verteilungstypen

·) Ausgangspunkt: Generator erzeugt $(u_n)_{n\in\mathbb{N}}$ in $[0,1)$ mit Gleichverteilung (0 kann vorkommen)

·) Ziel: Erzeugung von $(y_n)_{n\in\mathbb{N}}$ in $\mathbb{R}$

a) Inversionsmethode:

Definition: (wirksame und beste Schätzungen)

Seien $(\mathbb{R}^n, \mathcal{B}_n, \{P_\vartheta^{\otimes n} \mid \vartheta \in \Theta\})$ ein statistischer Raum, $\gamma: \Theta \to \mathbb{R}$ und T^*, T_1, T_2 erwartungstreue Schätzungen für γ mit $\mathrm{var}_\vartheta T_i(x_1, \ldots, x_n) < \infty$, $i = 1, 2$.

a) T_1 ist wirksamer als T_2 $:\Longleftrightarrow$ $\mathrm{var}_\vartheta T_1(x_1, \ldots, x_n) \leq \mathrm{var}_\vartheta T_2(x_1, \ldots, x_n)$ für alle $\vartheta \in \Theta$

b) T^* heißt beste erwartungstreue Schätzung für γ, wenn $\mathrm{var}_\vartheta T^*(x_1, \ldots, x_n) < \infty$ und

$$\mathrm{var}_\vartheta T^*(x_1, \ldots, x_n) \leq \mathrm{var}_\vartheta T(x_1, \ldots, x_n)$$

für alle $\vartheta \in \Theta$ und erwartungstreuen Schätzungen T von γ

Bemerkung: ·) Alternatives Kriterium: Minimierung $\mathbb{E}_\vartheta\left(T(x_1, \ldots, x_n) - \gamma(\vartheta)\right)^2$ für alle $\vartheta \in \Theta$

$$(\mathbb{E}_\vartheta(T - \gamma(\vartheta)))^2 = \mathbb{E}_\vartheta(T - \mathbb{E}_\vartheta T + \mathbb{E}_\vartheta T - \gamma(\vartheta))^2$$

$$= \underbrace{(\mathbb{E}_\vartheta T - \gamma(\vartheta))^2}_{=:\, \mathrm{bias}_\vartheta^2} + \underbrace{\mathbb{E}_\vartheta(T - \mathbb{E}_\vartheta T)^2}_{\mathrm{var}_\vartheta T} + \underbrace{2\, \mathbb{E}_\vartheta(T - \mathbb{E}_\vartheta T)(\mathbb{E}_\vartheta T - \gamma(\vartheta))}_{= 0}$$

$$= \mathrm{bias}_\vartheta^2 + \mathrm{var}_\vartheta T$$

·) man kann z.B. zeigen, dass $\hat{\mu} = \bar{x}$ die beste erwartungstreue Schätzung für Normalverteilung, Exponentialverteilung, Poissonverteilung, Binomialverteilung

Ausnahme: stetige Gleichverteilung auf unbekanntem Intervall (hier ist max und min beste Schätzung)

Definition: (schwache Konsistenz)

Sei $((\mathbb{R}^n, \mathcal{B}_n, \{P_\vartheta^{\otimes n} \mid \vartheta \in \Theta\}))_{n \in \mathbb{N}}$ eine Folge von statistischen Räumen. Seien $\gamma: \Theta \to \mathbb{R}$, $(T_n: \mathbb{R}^n \to \mathbb{R})_{n \in \mathbb{N}}$ eine Folge von Statistiken.
Die Folge (T_n) heißt (schwach) konsistent für γ, wenn für alle $\vartheta \in \Theta$ und alle $\varepsilon > 0$

$$\lim_{n \to \infty} P_\vartheta^{\otimes n}\left(|T_n(x_1, \ldots, x_n) - \gamma(\vartheta)| > \varepsilon\right) = 0$$

Bemerkung: ·) auch Konvergenz in Wahrscheinlichkeitsraum
·) starke Konsistenz $\Longleftrightarrow$ fast-sicher konvergent

Lemma: (hinreichendes Kriterium)

Hinreichend für schwache Konsistenz einer Folge $(T_n)_{n\in\mathbb{N}}$ von Statistiken ist:

- $\lim_{n\to\infty} \mathbb{E}_\vartheta T_n(x_1,\dots,x_n) = \gamma(\vartheta)$ (asymptotische Erwartungstreue)
- $\lim_{n\to\infty} \operatorname{var}_\vartheta T_n(x_1,\dots,x_n) = 0$

für alle $\vartheta \in \Theta$

Beispiel:
- $\mu_n = \frac{1}{n} \sum_{i=1}^n x_i$ ist schwach konsistent für $\mu(\vartheta)$ falls $\operatorname{var} X_i < \infty$ für alle $\vartheta \in \Theta$
- $\hat{\sigma}_n = \frac{1}{n-1} \sum_{i=1}^n (x_i - \overline{x}_{(n)})^2$ ist schwach konsistent für $\sigma^2(\vartheta)$ falls $\mathbb{E}(x_i^4) < \infty$ für alle $\vartheta \in \Theta$

Maximum-Likelihood-Methode

heuristische Idee: Sei $(x_1,\dots,x_n)$ eine konkrete Stichprobe. Man sucht nun $\vartheta \in \Theta$ für den diese Stichprobe

- die größte Wahrscheinlichkeit besitzt, falls diskret verteilt
- den Maximalwert der Dichte liefert, falls stetig verteilt

Definition: (Likelihood-Funktion)

Sei $(\mathbb{R}^n, \mathcal{B}_n, \{P_\vartheta^{\otimes n} \mid \vartheta \in \Theta\})$ für $n \in \mathbb{N}$ ein statistischer Raum. Dann ist die Likelihood-Funktion definiert durch:

$$L : \Theta \times \mathbb{R}^n \longrightarrow [0, \infty)$$

a) $L(\vartheta, x_1,\dots,x_n) := \prod_{i=1}^n f_\vartheta(x_i)$ falls für alle $\vartheta \in \Theta$ die Funktion f_ϑ eine Dichte von P_ϑ ist

b) $L(\vartheta, x_1,\dots,x_n) := \prod_{i=1}^n P_\vartheta(X_i = x_i)$ falls für alle $\vartheta \in \Theta$ P_ϑ diskret

Bemerkung: Produkt kommt zustande, da $x_1,\dots,x_n$ unabhängig entstanden sind

Definition:

Der Schätzwert $\hat{\vartheta}^*(x_1,\dots,x_n) \in \Theta$ heißt Maximum-Likelihood-Schätzwert zur konkreten Stichprobe $x_1,\dots,x_n$, wenn für alle $\vartheta \in \Theta$

$$L(\hat{\vartheta}^*(x_1,\dots,x_n), x_1,\dots,x_n) \geq L(\vartheta, x_1,\dots,x_n)$$

Bemerkung: ·) Zur Berechnung einer ML-Schätzung betrachtet man oft $\ln L$, sodass Summen entstehen.

·) ML-Schätzung sagt nichts über Güteeigenschaften aus

Definition: (Wichtige Verteilungen)

Die Zufallsvariable Y besitzt eine χ^2-Verteilung mit r Freiheitsgraden, wobei $r \in \mathbb{N}$, wenn sie folgende Verteilungsdichte besitzt:

$$f_Y(x) = \left(2^{\frac{r}{2}}\,\Gamma\left(\tfrac{r}{2}\right)\right)^{-1} x^{\frac{r}{2}-1} e^{-\frac{x}{2}} \, 1_{[0,\infty)}(x)$$

Man schreibt dann auch $Y \sim \chi_r^2$.

Die Zufallsvariable Z besitzt eine t-Verteilung mit r Freiheitsgraden, wenn sie folgende Verteilungsdichte besitzt

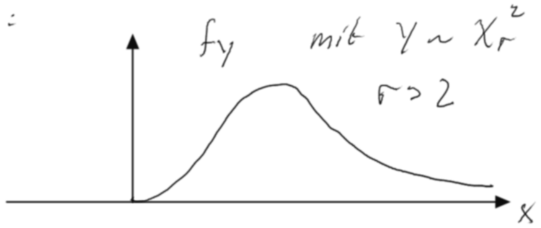
$$f_Z(x) = \frac{\Gamma\left(\frac{r+1}{2}\right)}{\Gamma\left(\frac{r}{2}\right)\sqrt{\pi r}} \left(1 + \frac{x^2}{r}\right)^{-\frac{r+1}{2}}$$

Man schreibt dann auch $Z \sim t_r$.

Die Zufallsvariable Y besitzt eine F-Verteilung mit $(s,r) \in \mathbb{N}^2$ Freiheitsgraden (auch $Y \sim F_{s,r}$), wenn sie folgende Verteilungsdichte besitzt:

$$f_Y(x) = \left(\frac{s}{r}\right)^{\frac{s}{2}} x^{\frac{s}{2}-1} \left(1 + \frac{s}{r} x\right) 1_{[0,\infty)}(x)$$

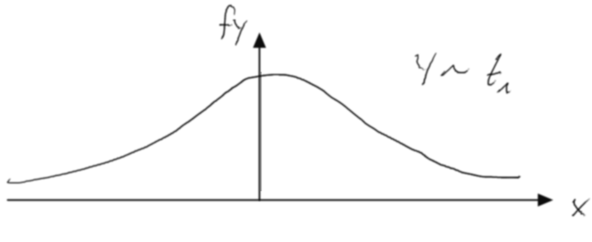
Bemerkung:



$\chi_2^2 \sim Exp\left(\frac{1}{2}\right)$, $\chi_r^2 = \Gamma_{\frac{r}{2},\frac{1}{2}}$

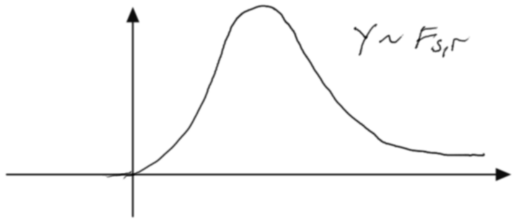
$Y_1 \sim \chi_{r_1}^2$, $Y_2 \sim \chi_{r_2}^2$, Y_1, Y_2 unabhängig

$\Longrightarrow \quad Y_1 + Y_2 \sim \chi^2_{r_1+r_2}$



·) Erwartungswert von t_1 existiert nicht

·) $f_Y(x) \longrightarrow (2\pi)^{-\frac{1}{2}} e^{-x^2}$, $r \to \infty$



Proposition:

Seien $r, s \in \mathbb{N}$ und $X_0, \dots, X_{r+s}$ unabhängig und identisch standard-normalverteilt. Dann gilt:

$$\boxed{\begin{aligned} &\cdot)\quad \sum_{i=1}^{r} X_i^2 \sim \chi_r^2 \\ &\cdot)\quad \frac{X_0}{\left(\frac{1}{r}\sum_{i=1}^{r} X_i^2\right)^{\frac{1}{2}}} \sim t_r^2 \\ &\cdot)\quad \left(\frac{1}{s}\sum_{i=r+1}^{s+r} X_i^2\right)\left(\frac{1}{r}\sum_{i=1}^{r} X_i^2\right)^{-1} \sim F_{s,r} \end{aligned}}$$

Beweis: hier nicht $\square$

Folgerung: $Y \sim F_{s,r} \implies Y^{-1} \sim F_{r,s}$ (wichtig für Tabellenbenutzung)

Lemma:

Seien $X_1, \dots, X_n$, $n \in \mathbb{N}$ iid Zufallsgrößen mit $X_i \sim N_{\mu,\sigma^2}$, $i \in \mathbb{N}, i \leq n$. Dann sind

$$\bar{X} := \frac{1}{n}\sum_{i=1}^{n} X_i$$

und

$$\tilde{X} := \sum_{i=1}^{n} (X_i - \bar{X})^2$$

unabhängig.

Beweis: Seien $Y_1, \dots, Y_n$ iid mit $Y_i \sim N_{0,1}$ und $A \in \mathbb{R}^{n \times n}$ eine orthonormale Matrix. Dann sei

$$\vec{Z} := A\vec{Y} = (Z_1, \dots, Z_n)$$

$$\overset{(\text{Satz})}{\implies} Z_1, \dots, Z_n \text{ iid mit } Z_i \sim N_{0,1}$$

Wähle $a_{1j} = \frac{1}{\sqrt{n}}$ für alle j

$$\implies \bar{Y} = \frac{1}{\sqrt{n}} Z_1$$

$$\implies \tilde{Y} = \sum_{i=2}^{n} Z_i^2 \implies \bar{Y} \text{ unabh. von } \tilde{Y}$$

$\square$

Bemerkung: Man kann zeigen: Sind $\bar{X}, \tilde{X}$ unabhängig, so folgt $X_1, \dots, X_n \sim N_{\mu,\sigma^2}$.

Satz:

Seien $X_1, \dots, X_n$, $n \in \mathbb{N}$ iid mit $X_i \sim N_{\mu, \sigma^2}$. Dann gilt:

$$\boxed{\begin{aligned} &\cdot)\ \frac{1}{\sigma} \sum_{i=1}^{n} (X_i - \bar{X})^2 \sim \chi^2_{n-1} \\ &\cdot)\ \frac{\bar{X} - \mu}{\left(\frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2\right)^{\frac{1}{2}}} \sim t_{n-1} \end{aligned}}$$

Konfidenzintervalle

Ziel: Aufgrund Vorkenntnisse und Stichprobe Konstruktion eines (zufälligen) Bereiches, der den wahren unbekannten Parameter ϑ bezüglich $\gamma(\vartheta)$ mit einer vorgegebenen Mindestwahrscheinlichkeit überdeckt.

Definition: (Konfidenzintervallschätzung)

Seien $(\mathbb{R}^n, \mathcal{B}_n, \{P_\vartheta^{\otimes n} \mid \vartheta \in \Theta\})$ ein statistischer Raum, $\gamma : \Theta \to \mathbb{R}$ und $(1-\alpha) \in (0,1)$. Eine Statistik $C : \mathbb{R}^n \longrightarrow \{[a,b] \mid a,b \in \mathbb{R}, a < b\}$ heißt $(1-\alpha)$-Konfidenzintervallschätzung für $\gamma(\vartheta)$, wenn

$$P_\vartheta^{\otimes n}(\gamma(\vartheta) \in C(X_1, \dots, X_n)) \geq 1 - \alpha \quad \text{für alle } \vartheta \in \Theta$$

Bemerkung: $\cdot)$ häufige $1-\alpha$: $0{,}9$; $0{,}95$; $0{,}99$

Beispiel: (normalverteilte Zufallsgröße mit bekannter Varianz)

Sei $\sigma_0^2 \in (0, \infty)$. (genutzter stat. Raum $(\mathbb{R}^n, \mathcal{B}_n, \{N_{\mu,\sigma_0^2}^{\otimes n} \mid \mu \in \mathbb{R}\})$)

Gesucht: $(1-\alpha)$-Konfidenzintervall für μ

$\Rightarrow$ Ansatz: such KI $[\bar{x} - a, \bar{x} + a]$ für $a > 0$, sodass

$$P_\mu^{\otimes n}(\bar{X} - a \leq \mu \leq \bar{X} + a) \geq 1 - \alpha$$

$$\Rightarrow P_\mu^{\otimes n}\Big(-\frac{\sqrt{n}\,a}{\sigma_0} \leq \underbrace{\sqrt{n}\,\frac{\bar{X} - \mu}{\sigma_0}}_{= N_{0,1}} \leq \frac{\sqrt{n}\,a}{\sigma_0}\Big) \geq 1 - \alpha$$

$$\Rightarrow \phi\Big(\frac{\sqrt{n}\,a}{\sigma_0}\Big) \geq 1 - \frac{\alpha}{2} \quad \Rightarrow \quad a \geq \frac{\sigma_0}{\sqrt{n}}\, \phi^{-1}\Big(1 - \frac{\alpha}{2}\Big)$$

$$\Longrightarrow \boxed{C(x_1,\dots,x_n) = \left[\bar{X} - \frac{\sigma_0}{\sqrt{n}}\phi^{-1}\left(1-\frac{\alpha}{2}\right),\ \bar{X} + \frac{\sigma_0}{\sqrt{n}}\phi^{-1}\left(1-\frac{\alpha}{2}\right)\right]}$$

ist eine $1-\alpha$ - Konfidenzintervallschätzung

Bemerkung: Länge des Intervalls hängt nicht von Stichprobe ab, sondern nur vom Stichprobenumfang n

Beispiel: (beliebige normalverteilte Zufallsgröße)

$\Longrightarrow$ genutzter stat. Raum $(\mathbb{R}^n, \mathcal{B}_n, \{P_{\mu,\sigma^2}^{\otimes n} \mid \mu \in \mathbb{R}, \sigma^2 \in (0,\infty)\})$

(gleicher Ansatz)

$$\Longrightarrow N_{\mu,\sigma^2}^{\otimes n}\left(\frac{-\sqrt{n}a}{\sqrt{\hat{\sigma}^2}} \leq \underbrace{\sqrt{n}\,\frac{\bar{X}-\mu}{\sqrt{\hat{\sigma}^2}}}_{\sim\, t_{n-1}} \leq \frac{\sqrt{n}a}{\sqrt{\hat{\sigma}^2}}\right) \geq 1-\alpha$$

(kann nicht mehr normalverteilt sein, da $\hat{\sigma}^2$ ein Schätzwert ist)

mit $\hat{\sigma}^2$ als korrigierte emp. Varianz

$$\Longrightarrow N_{\mu,\sigma^2}^{\otimes n}\left(-t_{n-1,1-\frac{\alpha}{2}} \leq \sqrt{n}\,\frac{\bar{X}-\mu}{\sqrt{\hat{\sigma}^2}} \leq t_{n-1,1-\frac{\alpha}{2}}\right) \geq 1-\alpha$$

$$\Longrightarrow \boxed{C(x_1,\dots,x_n) = \left[\bar{X} - \frac{\sqrt{\hat{\sigma}^2}}{\sqrt{n}}t_{n-1,1-\frac{\alpha}{2}},\ \bar{X} + \frac{\sqrt{\hat{\sigma}^2}}{\sqrt{n}}t_{n-1,1-\frac{\alpha}{2}}\right]}$$

ist eine $(1-\alpha)$-KIS für μ

Bemerkung: $t_{n-1,1-\frac{\alpha}{2}}$ ist $1-\frac{\alpha}{2}$ - Quantil der t_{n-1} Verteilung

Weiterhin gilt: $(n-1)\frac{\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{n-1}$ (wichtig: Verteilung hängt nicht von σ^2 ab, aber σ^2 kommt vor)

$$\Longrightarrow N_{\mu,\sigma^2}^{\otimes n}\left(\chi_{n-1,\frac{\alpha}{2}} \leq \frac{(n-1)\hat{\sigma}^2}{\sigma^2} \leq \chi_{n-1,1-\frac{\alpha}{2}}\right) \geq 1-\alpha$$

$$\Longrightarrow N_{\mu,\sigma^2}^{\otimes n}\left(\frac{(n-1)\hat{\sigma}^2}{\chi_{n-1,1-\frac{\alpha}{2}}} \leq \sigma^2 \leq \frac{(n-1)\hat{\sigma}^2}{\chi_{n-1,\frac{\alpha}{2}}}\right) \geq 1-\alpha$$

$$\Longrightarrow \boxed{C(x_1,\dots,x_n) = \left[\frac{(n-1)\hat{\sigma}^2}{\chi^2_{n-1,1-\frac{\alpha}{2}}},\ \frac{(n-1)\hat{\sigma}^2}{\chi^2_{n-1,\frac{\alpha}{2}}}\right]}$$ ist eine $(1-\alpha)$-KIS für σ^2

Bemerkung: Gemeinsamer $(1-\alpha)$ - Konfidenzintervall $f(\mu,\sigma^2) \in \mathbb{R}\times(0,\infty)$ (ergibt sich nicht als Kreuzprodukt der Intervalle)

es gelte: $P_{\mu,\sigma^2}^{\otimes n}(\mu \in C_1) \geq 1-\frac{\alpha}{2}$, $P_{\mu,\sigma^2}^{\otimes n}(\sigma^2 \in C_2) \geq 1-\frac{\alpha}{2}$

$$\Longrightarrow P_{\mu,\sigma^2}^{\otimes n}(\mu \in C_1, \sigma^2 \in C_2) \overset{(*)}{\geq} P(\mu \in C_1) + P(\sigma^2 \in C_2) - 1$$

$$\geq 1-\frac{\alpha}{2} + 1-\frac{\alpha}{2} - 1 = 1-\alpha$$

$(*)$ $P(A\cap B) = P(A) + P(B) - P(A\cup B) \geq P(A) + P(B) - 1$

Prinzipien bei Konstruktion von Konfidenzintervallen:

- ·) gute Punktschätzung für gesuchten Parameter $\mu(\vartheta)$ suchen
- ·) aus gegebenen Informationen Zufallsvariable konstruieren, in der nur der zu schätzende Parameter vorkommt
- ·) Verteilung der Zufallsvariable darf nicht von unbekanntem Parameter abhängen
- ·) Mit Hilfe von Quantilen ein Intervall angeben, in dem gesuchter Parameter mit Wahrscheinlichkeit größer gleich $1-\alpha$ liegt
- ·) Ungleichung nach $\mu(\vartheta)$ umstellen und Konfidenzintervall aufzustellen

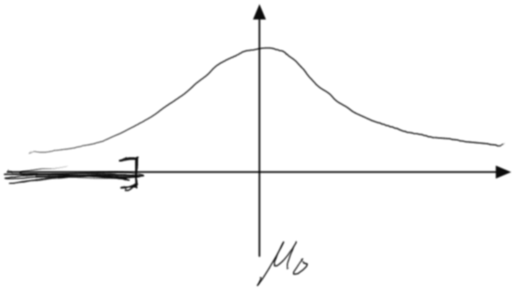
b) einseitige Alternativhypothese:

(i) $H_0: \mu \geq \mu_0$ (ii) $H_0^{'}: \mu \leq \mu_0$

$H_1: \mu < \mu_0$ $H_1^{'}: \mu > \mu_0$

·) Testgröße wie bei zweiseitiger Alternativhypothese

·) Unterschied bei Festlegung des kritischen Bereiches



$$T(x_1, \dots, x_n) = \frac{\sqrt{n}\,(\bar{x} - \mu_0)}{\sqrt{\sigma_0^2}}$$

$$\Longrightarrow K = (-\infty, -\Phi^{-1}(1-\alpha)]$$

$$K^{'} = [\Phi^{-1}(1-\alpha), \infty)$$

(Einstichproben)-t-Test

(zur Prüfung des Erwartungswerts bei unbekannter Varianz)

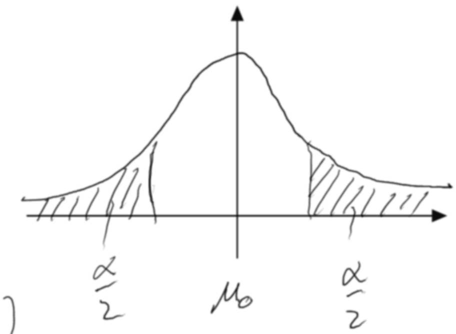
$\Longrightarrow$ stat. Raum: $(\mathbb{R}^n, \mathcal{B}_n, \{ N_{\mu,\sigma^2}^{\otimes n} \mid \mu \in \mathbb{R}, \sigma^2 \in (0,\infty) \})$

a) zweiseitige Alternativhypothese: $\mu_0 \in \mathbb{R}$ vorgegeben

$H_0: \mu = \mu_0$

$H_1: \mu \neq \mu_0$

Testgröße: $T(x_1, \dots, x_n) = \frac{\sqrt{n}}{\sqrt{\hat{\sigma}^2}} (\bar{x} - \mu_0)$



$\Longrightarrow T(X_1, \dots, X_n) \sim t_{n-1}$ falls $\mu = \mu_0$

$\Longrightarrow K = (-\infty, -t_{n-1,1-\frac{\alpha}{2}}] \cup [t_{n-1,1-\frac{\alpha}{2}}, \infty)$

Bemerkung: (andere Variante mit gleicher Testgröße)

·) jetzt nicht Bestimmung des krit. Bereiches, sondern Überschreitungswahrsch.:

$$p_ü := N_{\mu_0,\sigma^2}^{\otimes n}\Big(\underbrace{|T(X_1, \dots, X_n)|}_{\sim t_{n-1}} > \underbrace{|T(x_1, \dots, x_n)|}_{\text{konkrete Stichpr.}} \Big)$$

Entscheidung: wenn $p_ü(x_1, \dots, x_n) \leq \alpha$, dann H_0 ablehnen, sonst H_0 annehmen

·) weitere äquivalente Möglichkeiten zur Durchführung des t-Tests

Bestimmung aller μ_0 für die H_0 angenommen wird

$\Longrightarrow \quad T(x_1, \dots, x_n) \notin K \iff \mu_0 \in \left(\bar{x} - \frac{\sqrt{\hat{\sigma}^2}}{\sqrt{n}} t_{n-1, 1-\frac{\alpha}{2}}, \bar{x} + \frac{\sqrt{\hat{\sigma}^2}}{\sqrt{n}} t_{n-1, 1-\frac{\alpha}{2}} \right)$

$\iff \mu_0$ liegt im $(1-\alpha)$-KI für μ

Ein Prinzip zur Konstruktion von Tests

(auch Maximum-Likelihood-Quotienten-Test)

Betrachte Testproben: $(\mathbb{R}^n, \mathcal{B}_n, \{P_\vartheta^{\otimes n} \mid \vartheta \in \Theta\})$

$H_0: \vartheta \in \Theta_0$

$H_1: \vartheta \in \Theta_1$

Beschreiben jetzt randomisierten Test als

$\varphi: \mathbb{R}^n \longrightarrow [0,1]$, wobei $\varphi(x)$ für $x \in \mathbb{R}^n$ die Wahrscheinlichkeit für die Annahme von H_1 angibt, wenn eine konkrete Stichprobe vorliegt

(damit auch nicht-random. Test: $\tilde{\varphi}: \mathbb{R} \to \{0,1\}$)

Gütefunktion eines randomisierten Tests:

$$\beta_\varphi : \Theta \longrightarrow [0,1] \text{ mit } \beta_\varphi(\vartheta) = \mathbb{E}_\vartheta \varphi(X_1, \dots, X_n)$$

Betrachte für $x \in \mathbb{R}^n$ den Likelihood-Quotient:

$$R(x) = \begin{cases} \dfrac{\sup\{L(\vartheta, x) \mid \vartheta \in \Theta_1\}}{\sup\{L(\vartheta, x) \mid \vartheta \in \Theta_0\}} & : \text{falls Nenner} > 0 \\ \infty & : \text{sonst} \end{cases}$$