

Konfidenzintervalle

Annahme: $X_1, \dots, X_n \stackrel{iid}{\sim} F_\theta$.

Gesucht: $U = U(X_1, \dots, X_n)$, $O = O(X_1, \dots, X_n)$, sodass für das wahre θ gilt

$$P_\theta(U \leq \theta \leq O) = 1 - \alpha, \quad \alpha \in (0, 1).$$

Das Intervall $[U, O]$ ist ein Konfidenzintervall (KIV) für θ zum Niveau $1 - \alpha$.

Intuitive Bedeutung: KIV ist ein Intervall, das mit Wahrscheinlichkeit $1 - \alpha$ (groß) den unbekanntem Parameter θ überdeckt.

Aber: Sei (x_1, \dots, x_n) eine Realisation, dann enthält $[u, o]$ den wahren Parameter θ oder eben nicht.

Zu sagen, dass $\theta \in [u, o]$ mit W! $1 - \alpha$ ist somit unsinnig!

Und trotzdem: Sei $[U_r, O_r]$, $r = 1, \dots, R$, eine Folge iid KIVs für θ zum Niveau $1 - \alpha$, dann resultiert mit dem Starken Gesetz der großen Zahlen (SLLN)

$$\frac{1}{R} \sum_{r=1}^R I_{[U_r, O_r]}(\theta) \xrightarrow{f.s.} 1 - \alpha.$$

Hierbei gilt für die Indikatoren $I_{[U_r, O_r]}(\theta) \stackrel{iid}{\sim} \text{Bernoulli}(1 - \alpha)$, also

$$\mathbb{E}\left(I_{[U_r, O_r]}(\theta)\right) = 1 - \alpha.$$

Das heißt ?

Konfidenzintervalle bei Normalverteilung: $X_i \stackrel{iid}{\sim} N(\mu, \sigma^2)$

1. Für μ (σ bekannt):

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \Rightarrow Z := \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$$

$$\begin{aligned} P(z_{\alpha/2} \leq Z \leq z_{1-\alpha/2}) &= P\left(\bar{X} - \frac{\sigma}{\sqrt{n}} z_{1-\alpha/2} \leq \mu \leq \bar{X} + \frac{\sigma}{\sqrt{n}} z_{1-\alpha/2}\right) \\ &= 1 - \alpha. \end{aligned}$$

2. Für μ (σ unbekannt):

$$T := \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t_{n-1}$$

$$P(t_{n-1;\alpha/2} \leq T \leq t_{n-1;1-\alpha/2}) =$$

$$P\left(\bar{X} - \frac{S}{\sqrt{n}} t_{n-1;1-\alpha/2} \leq \mu \leq \bar{X} + \frac{S}{\sqrt{n}} t_{n-1;1-\alpha/2}\right) = 1 - \alpha.$$

3. Für σ^2 (μ unbekannt):

$$Y := \frac{(n-1)S^2}{\sigma^2} \sim \chi_{n-1}^2$$

$$\begin{aligned} P(\chi_{n-1;\alpha/2}^2 \leq Y \leq \chi_{n-1;1-\alpha/2}^2) &= P\left(\frac{(n-1)S^2}{\chi_{n-1;1-\alpha/2}^2} \leq \sigma^2 \leq \frac{(n-1)S^2}{\chi_{n-1;\alpha/2}^2}\right) \\ &= 1 - \alpha. \end{aligned}$$

Für eine beliebige Verteilung F mit $E(X_i) = \mu$ und $\text{var}(X_i) = \sigma^2$ folgt mit ZGWS

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \stackrel{as}{\approx} N(0, 1).$$

KIVs wie zuvor, aber mit asymptotischer Überdeckungswahrscheinlichkeit $1 - \alpha$.

Hypothesentests

Zweck: Aussagen oder **Hypothesen** über Verteilung einer ZV'en Y anhand der Stichprobe X_1, \dots, X_n zu untermauern.

Ein Hypothesentest beinhaltet:

- **Testproblem:** Nullhypothese H_0 und Alternativhypothese H_1 .
- **Teststatistik:** $T = T(X_1, \dots, X_n)$.
- **Entscheidungsregel:** Jeder Realisation von T wird Entscheidung für oder gegen die vorliegende Hypothese zugeordnet:
 - (a) Entscheidung für H_0 , falls T nicht in C realisiert.
 - (b) Entscheidung für H_1 , falls T in C realisiert.

C heißt **kritischer Bereich**.

Parametrische Statistik: Hypothesen über Werte eines Parameters θ .

Bezeichne Ω_0 die unter H_0 zulässigen Parameterwerte und Ω_1 die unter H_1 zulässigen. Je nach Gestalt dieser Mengen unterscheiden wir

- (a) einfache
- (b) zusammengesetzte Testprobleme.

Bei univariaten Testproblemen gibt es im wesentlichen

- (a) einseitige und
- (b) zweiseitige Fragestellungen.

Die Entscheidung für H_0 oder H_1 kann richtig oder falsch sein:

H_0	Entscheidung	
	annehmen	ablehnen
richtig	$1 - \alpha$	α
falsch	$1 - \beta$	β

Ziel: α und $1 - \beta$ möglichst klein halten. (Widersprüchliche Forderung!!)

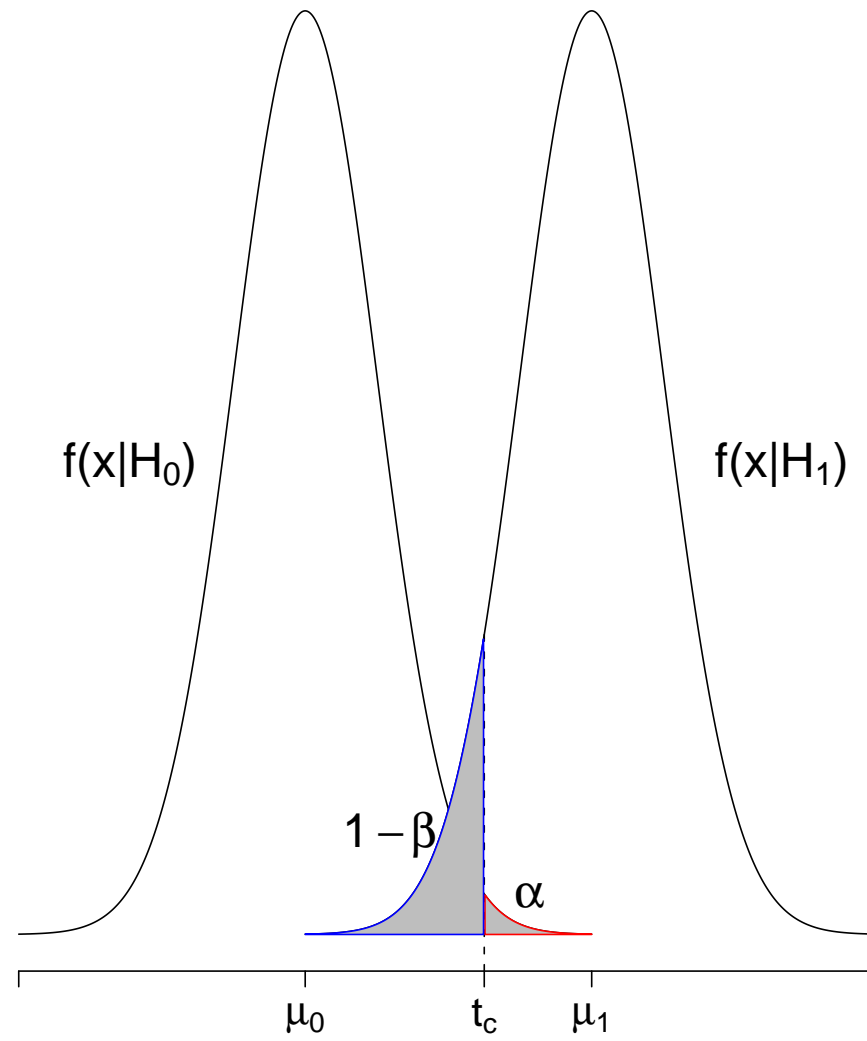
Üblich: α vorgegeben, Annahmebereich bestimmen und Fehler $1 - \beta$ berechnen.

$1 - \beta$ kann groß werden. Da der wahre Wert des Parameters unbekannt, kann man über Fehler 2. Art keine genaue Auskunft geben. Nur Fehler 1. Art ist unter Kontrolle und damit nur die damit verbundene Entscheidung: " H_0 verwerfen".

Richtige Entscheidungen:

(a) $P(T \notin C | H_0 \text{ richtig}) = 1 - \alpha$ (durch Niveau bestimmt!!)

(b) $P(T \in C | H_1 \text{ richtig}) = \beta$ (Macht des Tests).



Parametrische Tests bei Normalverteilung

1. Test auf μ bei σ bekannt (Gaußtest):

H_0	H_1	Entscheidung gegen H_0 , falls	kritische Werte
$\mu = \mu_0$	$\mu \neq \mu_0$	$\bar{X} < c_3$ oder $\bar{X} > c_4$	$c_3 = \mu_0 - z_{1-\alpha/2} \sigma / \sqrt{n}$ $c_4 = \mu_0 + z_{1-\alpha/2} \sigma / \sqrt{n}$
$\mu \leq \mu_0$	$\mu > \mu_0$	$\bar{X} > c_1$	$c_1 = \mu_0 + z_{1-\alpha} \sigma / \sqrt{n}$
$\mu \geq \mu_0$	$\mu < \mu_0$	$\bar{X} < c_2$	$c_2 = \mu_0 - z_{1-\alpha} \sigma / \sqrt{n}$

2. Test auf μ bei σ unbekannt (t -Test):

H_0	H_1	Entscheidung gegen H_0 , falls	kritische Werte
$\mu = \mu_0$	$\mu \neq \mu_0$	$\bar{X} < c_3$ oder $\bar{X} > c_4$	$c_3 = \mu_0 - t_{n-1;1-\alpha/2} S/\sqrt{n}$ $c_4 = \mu_0 + t_{n-1;1-\alpha/2} S/\sqrt{n}$
$\mu \leq \mu_0$	$\mu > \mu_0$	$\bar{X} > c_1$	$c_1 = \mu_0 + t_{n-1;1-\alpha} S/\sqrt{n}$
$\mu \geq \mu_0$	$\mu < \mu_0$	$\bar{X} < c_2$	$c_2 = \mu_0 - t_{n-1;1-\alpha} S/\sqrt{n}$

mit $S^2 = (n - 1)^{-1} \sum_{i=1}^n (X_i - \bar{X})^2$.

3. Test auf σ^2 bei μ bekannt (χ^2 -Test):

H_0	H_1	Entscheidung gegen H_0 , falls	kritische Werte
$\sigma^2 = \sigma_0^2$	$\sigma^2 \neq \sigma_0^2$	$T < c_3$ oder $T > c_4$	$c_3 = \sigma_0^2 \chi_{n;\alpha/2}^2$ $c_4 = \sigma_0^2 \chi_{n;1-\alpha/2}^2$
$\sigma^2 \leq \sigma_0^2$	$\sigma^2 > \sigma_0^2$	$T > c_1$	$c_1 = \sigma_0^2 \chi_{n;1-\alpha}^2$
$\sigma^2 \geq \sigma_0^2$	$\sigma^2 < \sigma_0^2$	$T < c_2$	$c_2 = \sigma_0^2 \chi_{n;\alpha}^2$

mit $T = \sum_{i=1}^n (X_i - \mu)^2$.

4. Test auf σ^2 bei μ unbekannt (χ^2 -Test):

H_0	H_1	Entscheidung gegen H_0 , falls	kritische Werte
$\sigma^2 = \sigma_0^2$	$\sigma^2 \neq \sigma_0^2$	$T < c_3$ oder $T > c_4$	$c_3 = \sigma_0^2 \chi_{n-1; \alpha/2}^2$ $c_4 = \sigma_0^2 \chi_{n-1; 1-\alpha/2}^2$
$\sigma^2 \leq \sigma_0^2$	$\sigma^2 > \sigma_0^2$	$T > c_1$	$c_1 = \sigma_0^2 \chi_{n-1; 1-\alpha}^2$
$\sigma^2 \geq \sigma_0^2$	$\sigma^2 < \sigma_0^2$	$T < c_2$	$c_2 = \sigma_0^2 \chi_{n-1; \alpha}^2$

mit $T = \sum_{i=1}^n (X_i - \bar{X})^2$.

p-Wert

Für Tests liefern Computerprogramme keine logische Entscheidung sondern den p-Wert. Dieser ist die anhand der Stichprobe beobachtete Type I Error Rate.

Satz 1. [Probability Integral Transformation] *Habe X stetige Verteilungsfunktion $F_X(x)$ und sei $Y = F_X(X)$. Dann ist Y gleichverteilt auf $(0, 1)$, d.h.*

$$P(Y \leq y) = y, \quad 0 < y < 1.$$

Beweis:

$$\begin{aligned} P(Y \leq y) &= P(F_X(X) \leq y) = P(F_X^{-1}(F_X(X)) \leq F_X^{-1}(y)) \\ &= P(X \leq F_X^{-1}(y)) = F_X(F_X^{-1}(y)) = y. \end{aligned}$$

Bemerkung: Ist X diskret, so gilt: $P(Y \leq y) \leq y$, für $0 \leq y \leq 1$.

Definition 1. F_X ist stochastisch größer als F_Y , falls $F_X(t) \leq F_Y(t)$ für alle t gilt. Für $X \sim F_X$ und $Y \sim F_Y$ folgt $P(X \leq t) = F_X(t) \leq F_Y(t) = P(Y \leq t)$ und für alle t gilt

$$P(X > t) \geq P(Y > t).$$

Nach dem Test wird Ergebnis mitgeteilt. Eine Möglichkeit ist es α und damit die Entscheidung bzgl. H_0 zu berichten. Alternativ kann p-Wert übermittelt werden.

Definition 2. Der p-Wert $p(X)$ ist eine Teststatistik mit $0 \leq p(x) \leq 1$. Kleine Werte von $p(X)$ weisen auf die Richtigkeit von H_1 hin. Ein p-Wert ist gültig, falls für jedes $\theta \in \Theta_0$ und jedes $0 \leq \alpha \leq 1$ gilt

$$P_\theta(p(X) \leq \alpha) \leq \alpha.$$

Ist $p(X)$ gültig, kann damit ein Level α Test konstruiert werden. Der Test, der H_0 genau dann verwirft wenn $p(X) \leq \alpha$ ist ein Level α Test.

Wie kann nun ein gültiger p-Wert definiert werden?

Satz 2. Sei $W(X)$ eine Teststatistik. Große Werte von W sprechen gegen H_0 .
Definiere für einen beliebigen Stichprobenpunkt x

$$p(x) = \sup_{\theta \in \Theta_0} P_{\theta}(W(X) \geq W(x)).$$

Damit ist $p(X)$ ein gültiger p -Wert.

Beweis: Fixiere ein $\theta \in \Theta_0$. Sei dafür $F_{\theta}(w)$ die cdf von $-W(X)$. Definiere dafür

$$p_{\theta}(x) = P_{\theta}(W(X) \geq W(x)) = P_{\theta}(-W(X) \leq -W(x)) = F_{\theta}(-W(x)).$$

Für dieses θ entspricht die ZV'e $p_{\theta}(X)$ dem $F_{\theta}(-W(X))$. Mit Satz 1 folgt, dass die Verteilung von $p_{\theta}(X)$ stochastisch größer oder gleich einer Uniform(0, 1) ist. D.h. für jedes $0 \leq \alpha \leq 1$ gilt $P_{\theta}(p_{\theta}(X) \leq \alpha) \leq \alpha$.

Nun ist der p-Wert definiert über alle $\theta \in \Theta_0$, und es gilt dafür für jedes x

$$p(x) = \sup_{\theta' \in \Theta_0} p_{\theta'}(x) \geq p_{\theta}(x),$$

da der größte p-Wert für alle Elemente in Θ_0 zumindest so groß ist als für unseren Wert θ . Somit gilt auch für jedes $\theta \in \Theta_0$ und jedes $0 \leq \alpha \leq 1$

$$P_{\theta}(p(X) \leq \alpha) \leq P_{\theta}(p_{\theta}(X) \leq \alpha) \leq \alpha$$

und $p(X)$ ist daher ein gültiger p-Wert.

Beispiel: Sei X_1, \dots, X_n Zufallsstichprobe aus $N(\mu, \sigma^2)$ und teste $H_0: \mu = \mu_0$ gegen $H_1: \mu \neq \mu_0$.

LRT verwirft H_0 für große Werte von $W(X) = |\bar{X} - \mu_0|/(S/\sqrt{n})$.

Für $\mu = \mu_0$ folgt $(\bar{X} - \mu_0)/(S/\sqrt{n})$ einer t_{n-1} -Verteilung, unabhängig von σ .
Deshalb gilt hierfür

$$p(x) = P_{\theta_0}(W(X) \geq W(x)) = 2P\left(T_{n-1} \geq (\bar{x} - \mu_0)/(s/\sqrt{n})\right).$$