1 Konfidenzbereiche

Ziel: Einer Beobachtung $x \in \mathcal{X}$ soll nun ein Bereich $C(x) \subset \Theta$ zugeordnet werden, so dass der wahre Parameter ϑ mit großer Wahrscheinlichkeit in $C(x)$ enthalten ist.

Definition 1.1. Sei $\mathcal{M} = (\mathcal{X}, (P_\vartheta)_{\vartheta \in \Theta}$ ein statistisches Modell, Σ eine beliebige Menge und $\tau : \theta \to \Sigma$ die zu ermittelnde Kenngröße.
Eine Abbildung

$$C : \mathcal{X} \to \wp(\Sigma),$$

die jedem Beobachtungsergebnis $x \in \mathcal{X}$ eine Menge C(x) $\subset \Sigma$ zuordnet, heißt Konfidenzbereich für τ zum (Irrtums-)Niveau $\alpha \in (0,1)$, falls

$$\forall \vartheta \in \Theta : P_\vartheta(C(X) \ni \tau(\vartheta)) \geq 1 - \alpha \tag{1}$$

gilt. Insbesondere fordert man, dass $\{x \in \mathcal{X} : C(x) \ni \tau(\vartheta)\}$ ein Ereignis ist $(\forall \vartheta \in \theta)$.

Bemerkung 1.2.

i) Falls $\Sigma = \mathbb{R}$ und C(x) ein Intervall ist, dann sprechen wir von Konfidenzintervall.

ii) $C(x) = \Sigma, x \in \mathcal{X}$, erfüllt (1), ist aber völlig nutzlos. Gesucht wird ein möglichst kleiner Konfidenzbereich.

iii) Nicht ϑ variiert zufällig, sondern x und damit C(x).

Konstruktion von Konfidenzbereichen für $\tau = \mathrm{id}$

Die Angabe der Abbildung $C : \mathcal{X} \to \wp(\theta)$ ist äquivalent mit der Angabe der Menge

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

und damit äquivalent mit der Angabe aller „Schnitte“

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

Geeignete Mengen C_ϑ sind häufig leicht zu bestimmen und man gelangt von hier zu den Konfidenzbereichen mittels

$$C(x) = \{\vartheta \in \theta : x \in C_\vartheta\}.$$

D.h. C_ϑ ($\vartheta \in \Theta$) sind die Schnitte von C entlang der $\mathcal{X}$-Achse und $C(x)$ ($x \in \mathcal{X}$) sind die Schnitte von C entlang der Θ-Achse. Aufgrund der Äquivalenz

$$x \in C_\vartheta \Leftrightarrow \vartheta \in C(x)$$

gilt

$$P_\vartheta(C(X) \ni \vartheta) = P_\vartheta(X \in C_\vartheta)$$

und wir wählen für jedes $\vartheta \in \Theta$ eine möglichst kleine Menge C_ϑ mit

$$P_\vartheta(X \in C_\vartheta) \geq 1 - \alpha.$$

Bemerkung 1.3. Wir wählen häufig eine Menge der Form

$$C_\vartheta = \{x \in \mathcal{X} : p_\vartheta(x) \geq k_\vartheta\}$$

für eine geeignete Konstante $k_\vartheta > 0$. Hat man C_ϑ für alle ϑ spezifiziert, so gelangt man zur Abbildung $x \longmapsto C(x)$ mittels

$$C(x) = \{\vartheta \in \theta : x \in C_\vartheta\}.$$

Beispiel 1.4.

- Gegeben 10 Heizkraftwerke
- n = 4 zufällig ausgewählte Heizkraftwerke sollen auf schädliche Emissionen untersucht werden.

Ziel: Angabe eines Konfidenzbereichs für die Zahl ϑ der Kraftwerke mit schädlichen Emissionen zum Niveau $\alpha = 0.2$.

Statistisches Modell:

$$\mathcal{M} = (\mathcal{X}, (P_\vartheta)_{\vartheta\in\Theta} = \Big(\{0,1,2,3,4\}, \mathrm{Hyp}(4,\vartheta,10-\vartheta)_{\vartheta\in\{0,\cdots,10\}}\Big)$$

d.h. $p_\vartheta(x) = P_\vartheta(X = x) = \frac{\binom{\vartheta}{x}\binom{10-\vartheta}{4-x}}{\binom{10}{4}}$ $(\vartheta \in \{0,\cdots,10\}, x \in \{0,\cdots,4\})$

Nun konstruieren wir Mengen C_ϑ mit $P_\vartheta(X \in C_\vartheta) \geq 1-\alpha = 0.8 = \frac{168}{210}$ für alle ϑ. Dies liefert eine Menge C und die Θ-Schnitte definieren die Konfidenzbereiche:
Diagramm

$$\begin{aligned} C(0) &= \{0,1,2\} \\ C(1) &= \{1,2,3,4,5\} \\ C(2) &= \{3,\cdots,7\} \\ C(3) &= \{5,\cdots,9\} \\ C(4) &= \{8,9,10\} \end{aligned}$$

Obiger Zugang ist universell anwendbar. Häufig kann man aber durch geeignete Statistiken eleganter Konfidenzbereiche definieren!

1.1 Konstruktion von Konfidenzbereichen mittels geeigneter Statistiken

Ziel: Konstruktion eines Konfidenzbereichs für eine Kenngröße $\tau : \Theta \to \Sigma$ in einem statistischen Modell $\mathcal{M} = (\mathcal{X}, (P_\vartheta)_{\vartheta\in\Theta})$.

Definition 1.5. Eine Familie von Statistiken $S_z : \mathcal{X} \to \Sigma'$ $(z \in \Sigma)$ (in einen beliebigen messbaren Raum Σ') heißt *pivotall* für τ, wenn für alle $\vartheta \in \Theta$ die Verteilung von $S_{\tau(\vartheta)}$ unter P_ϑ die gleiche ist, sagen wir Q.

Mittels einer universellen Familie von Statistiken $(S_z)_{z\in\Sigma}$ für τ lassen sich einfach Konfidenzbereiche für τ generieren: Für ein gegebenes Konfidenzniveau $\alpha \in (0,1)$ wählt man $U \subset \Sigma'$ mit $Q(U) \geq 1-\alpha$. Dann definiert

$$C(x) := \{\vartheta \in \Theta : S_{\tau(\vartheta)}(x) \in U\} \qquad (x \in \mathcal{X}) \tag{2}$$

ein Konfidenzbereich zum Niveau α für τ, da für $\vartheta \in \Theta$

$$P_\vartheta(C(X) \ni \vartheta) = P_\vartheta(S_{\tau(\vartheta)}(X) \in U) = Q(U) \geq 1-\alpha.$$

D.h. man erhält Konfidenzbereiche durch folgendes Vorgehen:

i) Konstruktion einer universellen Familie (S_z)

ii) Wahl von $U \subset \Sigma'$ mit $Q(U) \geq 1-\alpha$

iii) Berechnung von C(x) mittels (2)

Wir nutzen diese Rangehensweise zur Konstruktion von Konfidenzbereichen im Gaußmodell.

Konfidenbzereiche für den Erwartungswert im Gaußmodell mit bekannter Varianz

Statistisches Modell

$$\begin{aligned} \mathcal{M} &= (\mathcal{X}, (P_\vartheta)_{\vartheta\in\Theta}) \\ &= \left(\mathbb{R}^n, (\mathcal{N}(\mu,v)^{\otimes n})_{\mu\in\mathbb{R}}\right) \end{aligned}$$

wobei die Varianz $v > 0$ fest vorgegeben ist.

Ziel: Konfidenzbereich für den Erwartungswert, d.h. $\tau = \mathrm{id}$.
Zunächst müssen wir in Schritt (i) universelle Statistiken angeben.

Lemma 1.6. *Für $\mu \in \mathbb{R}$ ist*

$$S_\mu := \frac{1}{\sqrt{nv}} \sum_{j=1}^{n} (X_j - \mu) \qquad (\mu \in \mathbb{R})$$

unter P_μ standardnormalverteilt, d.h. (S_μ) is eine universelle standardnormalverteilte Familie für $\tau = \mathrm{id}$.

Beweis. Seien $Y_1, \cdots, Y_n$ unabhängig $\mathcal{N}(0,1)$-verteilte ZV. Dann gilt

$$S_\mu \begin{pmatrix} \sqrt{v}Y_1 + \mu \\ \cdots \\ \sqrt{v}Y_n + \mu \end{pmatrix} = \frac{1}{\sqrt{n}} \underbrace{\sum_{j=1}^{n} Y_j}_{\mathcal{N}(0,n\cdot 1)} \sim \mathcal{N}(0,1)$$

Da $\sqrt{v}Y_1 + \mu, \ldots, \sqrt{v}Y_n + \mu$ unabhängig $\mathcal{N}(\mu, v)$-verteilt sind, folgt

$$P_\mu(S_\mu(X) \leq t) = P\left(\frac{1}{\sqrt{n}} \sum_{j=1}^n Y_j \leq t)\right) = \Phi(t)$$

und S_μ (bzw. $S_\mu(X)$) ist unter P_μ $\mathcal{N}(0,1)$-verteilt. □

Definition 1.7. Für $\alpha \in (0,1)$ und eine Verteilung Q auf $\mathbb{R}$ heißt $q \in \mathbb{R}$ α-*Quantil*, wenn

- $Q((-\infty, t]) \geq \alpha$ und
- $Q((-\infty, t)) \leq \alpha$

Bemerkung 1.8. :

- Jede reelle Verteilung Q besitzt für jedes $\alpha \in (0,1)$ ein α-Quantil.
- Hat Q keine Atome, d.h. $\forall t \in \mathbb{R}$ gilt $Q(\{t\}) = 0$, so ist $q \in \mathbb{R}$ ein α-Quantil genau dann wenn
$$F_Q(q) := Q((-\infty, q]) = \alpha$$
- Ist die Verteilungsfunktion strikt wachsend, so gibt es zu jedem $\alpha \in (0,1)$ genau ein α-Quantil.

Für gegebenes Signifikanzniveau $\alpha \in (0,1)$ wählen wir nun

$$(I) \quad U = [q_{\alpha/2}, q_{1-\alpha/2}] = [-q_{1-\alpha/2}, q_{1-\alpha/2}]\,,$$

wobei allgemein q_β für $\beta \in (0,1)$ das β-Quantil von $\mathcal{N}(0,1)$ bezeichne, d.h. $q_\beta = \Phi^{-1}(\beta)$. Dann gilt

$$Q(U) = \Phi(q_{1-\frac{\alpha}{2}}) - \Phi(q_{\frac{\alpha}{2}}) = 1 - \frac{\alpha}{2} - \frac{\alpha}{2} = 1 - \alpha$$

Alternative Möglichkeiten für die Wahl von U:

$$(II) \;\; U = (-\infty, q_{1-\alpha}] \quad \text{oder} \quad (III) \;\; U = [q_\alpha, \infty).$$

Wir führen nun die Bestimmung des zugehörigen Konfidenzintervalls durch (Schritt (iii)): Für die Wahl (I) erhalten wir

$$\begin{aligned}
C(x) &= \{\mu \in \mathbb{R} : S_\mu(x) \in U\} \\
&= \{\mu \in \mathbb{R} : -q_{1-\frac{\alpha}{2}} \leq \frac{1}{\sqrt{n\,v}} \sum_{j=1}^n (x_j - \mu) \leq q_{1-\frac{\alpha}{2}}\} \\
&= \{\mu \in \mathbb{R} : -\sqrt{\frac{v}{n}} q_{1-\frac{\alpha}{2}} \leq \underbrace{\frac{1}{n} \sum_{j=1}^n x_j}_{=M(x)} - \mu \leq \sqrt{\frac{v}{n}} q_{1-\frac{\alpha}{2}}\} \\
&= \left[M(x) - \sqrt{\frac{v}{n}} q_{1-\frac{\alpha}{2}}\,, M(x) + \sqrt{\frac{v}{n}} q_{1-\frac{\alpha}{2}}\right]
\end{aligned}$$

(symmetrisches Konfidenzintervall)

In den Fällen (II) und (III) erhält man die Konfidenzintervalle

$$\text{(II) } C(x) = \left[M(x) - \sqrt{\frac{v}{n}} q_{1-\alpha}, \infty\right) \text{ bzw.} C(x) = \left(-\infty, M(x) + \sqrt{\frac{v}{n}} q_{1-\alpha}\right]$$

(einseitige Konfidenzintervalle)

Satz 1.9. *Im Gaußmodell mit bekannter Varianz $v > 0$ und Stichprobengröße $n \in \mathbb{N}$ erhält man als Konfidenzintervalle zum Niveau $\alpha \in (0,1)$ das symmetrische Intervall*

$$\left[M - \sqrt{\frac{v}{n}} q_{1-\frac{\alpha}{2}}, M + \sqrt{\frac{v}{n}} q_{1-\frac{\alpha}{2}}\right]$$

und die einseitigen Intervalle

$$\left(-\infty, M + \sqrt{\frac{v}{n}} q_{1-\frac{\alpha}{2}}\right] \quad \textit{und} \quad \left[M - \sqrt{\frac{v}{n}} q_{1-\frac{\alpha}{2}}, \infty\right),$$

wobei $q_\beta = \phi^{-1}(\beta)$ das β-Quantil von $\mathcal{N}(0,1)$ bezeichne.

Konfidenzbereiche für den Erwartungswert und die Varianz im allgemeinen Gaußmodell

Statistisches Modell:

$$\mathcal{M} = (\mathcal{X}, (P_\vartheta)_{\vartheta \in \theta}) = \left(\mathbb{R}^n, (\mathcal{N}(\mu, v)^{\otimes n})_{\mu \in \mathbb{R}, v > 0}\right) \qquad (n \geq 2)$$

Konfidenzintervalle für den Erwartungswert:
Hierzu nutzt man, dass die Statistiken

$$t_\mu = \frac{1}{\sqrt{nV^*}} \sum_{j=1}^{n} (X_j - \mu)$$

unter $P_{\mu,v}$ Student-t-verteilt mit $n-1$ Freiheitsgraden sind. Hierbei bezeichne

$$M = \frac{1}{n} \sum_{j=1}^{n} X_j \text{ und } V^* = \frac{1}{n-1} \sum_{j=1}^{n} (X_j - M)^2.$$

t_μ wird auch t-Statistik genannt. Wir zeigen lediglich, dass die Verteilung nicht von der Wahl von μ und v abhängt.

Lemma 1.10. *Die Verteilung von t_μ unter $P_{\mu,v}$ hängt nicht von der Wahl von $\mu \in \mathbb{R}$ und $v > 0$ ab.*

Beweis. Wie zuvor: Seien $Y_1, \cdots, Y_n$ unabhängig $\mathcal{N}(0,1)$-verteilte ZV. Dann gilt

$$t_\mu \begin{pmatrix} \sqrt{v}Y_1 + \mu \\ \cdots \\ \sqrt{v}Y_n + \mu \end{pmatrix} = t_0 \begin{pmatrix} Y_1 \\ \cdots \\ Y_n \end{pmatrix}$$

Da $\sqrt{v}Y_1 + \mu, \ldots, \sqrt{v}Y_n + \mu$ unabhängig $\mathcal{N}(\mu, v)$-verteilt sind, folgt

$$P_{\mu,v}(t_\mu(X) \leq t) = P\left(t_0(Y) \leq t\right).$$

Dies beweist die Aussage, da die rechte Seite nicht von μ und v abhängt. □

Obige Argumentation führt nun zu den folgenden Konfidenzintervallen für das allgemeine Gaußmodell:

Satz 1.11. *Im allgemeinen Gaußmodell mit Stichprobengröße $n \geq 2$ erhält man als Konfidenzintervalle für den Erwartungswert μ zum Niveau $\alpha \in (0,1)$ das symmetrische Intervall*

$$\left[M - \sqrt{\frac{V^*}{n}} q_{1-\frac{\alpha}{2}}, M + \sqrt{\frac{V^*}{n}} q_{1-\frac{\alpha}{2}}\right]$$

und die einseitigen Intervalle

$$\left(-\infty, M + \sqrt{\frac{V^*}{n}} q_{1-\frac{\alpha}{2}}\right] \quad \textit{und} \quad \left[M - \sqrt{\frac{V^*}{n}} q_{1-\frac{\alpha}{2}}, \infty\right),$$

wobei q_β das β-Quantil der Student-t-Verteilung mit $n-1$ Freiheitsgraden bezeichne (kurz t_{n-1}-Verteilung).

Konfidenzintervalle für die Varianz

Zur Bestimmung von Konfidenzintervallen für die Varianz nutzt man die Statistiken

$$S_v = \frac{\sum_{j=1}^n (X_j - M)^2}{v} = \frac{(n-1)V^*}{v}$$

Analog zu obiger Rechnung zeigt man, dass deren Verteilung unter $P_{\mu,v}$ nicht von der Wahl von μ und v abhängt. Die entsprechende Verteilung Q ist in diesem Falle die $\mathcal{X}^2$-Verteilung mit $n-1$ Freiheitsgraden. Es ist die Verteilung der Summe von $n-1$ unabhängigen standardnormalverteilten Zufallsvariablen. Das gleiche vorgehen von zuvor liefert nun folgenden Satz:

Satz 1.12. *Im allgemeinen Gaußmodell mit Stichprobengröße $n \geq 2$ erhält man als Konfidenzintervalle für die Varianz v zum Niveau $\alpha \in (0,1)$ das beidseitige Intervall*

$$C(x) = \left[\frac{n-1}{q_{1-\frac{\alpha}{2}}} V^*, \frac{n-1}{q_{\frac{\alpha}{2}}} V^*\right]$$

und die einseitigen Intervalle

$$C(x) = \left(0, \frac{n-1}{q_\alpha} V^*\right] \quad \textit{und} \quad C(x) = \left[\frac{n-1}{q_{1-\alpha}} V^*, \infty\right).$$

Hierbei bezeichne allgemein q_β das β-Quantil der $\mathcal{X}^2$-Verteilung mit $n-1$ Freiheitsgraden.

2 Hypothesentests

2.1 Einführung

Wir betrachten ein statistisches Modell $\mathcal{M} = (\mathcal{X}, (P_\vartheta)_{\vartheta\in\Theta})$ und teilen den Parameterbereich Θ in zwei Hälften

- die Hypothese $H \subset \Theta$ und
- die Alternative $A = H^c \subset \Theta$

Ziel ist es nun aufgrund der Beobachtung zu entscheiden, ob der wahre Parameter in H oder A liegt. In einem Hypothesentest teilen wir $\mathcal{X}$ in zwei Hälften

- den *Akzeptanzbereich* $\mathcal{H} \subset \mathcal{X}$ und
- den *Verwerfungsbereich* $\mathcal{A} = \mathcal{H}^c$

und wir

- *akzeptieren* die Hypothese H, wenn wir eine Realisierung in $\mathcal{H}$ beobachten und
- *verwerfen* die Hypothese H, wenn wir eine Realisierung in $\mathcal{A}$ beobachten.

Hierbei können zwei Fehler auftreten:

- Bei wahrem Parameter $\vartheta \in H$ beobachten wir eine Realisierung in $\mathcal{A}$ und verwerfen die Hypothese (Fehler erster Art).
- Bei wahrem Parameter $\vartheta \in A$ beobachten wir eine Realisierung in $\mathcal{H}$ und akzeptieren die Hypothese (Fehler zweiter Art).

Wir bezeichnen mit

$$\beta(\vartheta) := P_\vartheta(X \in \mathcal{A}) \qquad (\vartheta \in \Theta)$$

die Verwerfungswahrscheinlichkeit und bemerken, dass bei wahrem Parameter $\vartheta \in H$ ein Fehler erster Art mit Wahrscheinlichkeit $\beta(\vartheta)$ eintritt. Ferner tritt bei wahrem Parameter $\vartheta \in A$ ein Fehler zweiter Art mit Wahrscheinlichkeit $1 - \beta(\vartheta)$ ein. In der Praxis ist es meist nicht möglich Fehler 1. und 2. Art auszuschließen und beim Hypothesentest einigt man sich auf die Konvention den Fehler <u>1. Art</u> klein zu halten und sich im Zweifel für die Hypothese zu entscheiden.
Wir nennen den Test der Hypothese H mit Akzeptanzbereich $\mathcal{H}$ vom *Niveau* $\alpha \in (0,1)$, wenn

$$\sup_{\vartheta\in H} \beta(\vartheta) \leq \alpha$$

gilt. Hiermit wird sichergestellt, dass für jede Wahl des Parameters $\vartheta \in H$ ein Fehler erster Art mit Wahrscheinlichkeit kleiner gleich α eintritt. Ferner heißt

$$\inf_{\vartheta\in A} \beta(\vartheta)$$

Güte oder *Power* des Tests.

Beispiel 2.1. Es soll überprüft werden, ob ein neues Medikament besser als das alte (mit einem Behandlungserfolg in 60% der Fälle) wirkt. In einer Studie wird die Wirksamkeit des Mediaments an n Personen getestet und die Zahl der Behandlungserfolge ermittelt.

Statistisches Modell: $(\{0,\ldots,n\},(\mathrm{Bin}(n,\vartheta))_{\vartheta\in[0,1]})$

Es ist ein Hypothesentest der Hypothese $H = [0, 0.6]$ durchzuführen. Wir machen den Ansatz

$$\mathcal{H} = \{0,\ldots,m-1\} \text{ und } \mathcal{A} = \{m,\ldots,n\},$$

wobei m noch in Abhängigkeit vom spezifizierten Signifikanzniveau $\alpha \in (0,1)$ geeignet zu wählen ist. Wir wählen das minimale m mit

$$\forall\vartheta \in H : \beta(\vartheta) = P_\vartheta(X \geq m) \leq \alpha.$$

Wegen der Monotonie der Wahrscheinlichkeit in ϑ, ist dies das minimale m mit $P_{0.6}(X \geq m) \leq \alpha$. Wir nutzen den zentralen Grenzwertsatz um m zuapproximieren:

$$P_{0.6}(X \geq m) = P_{0.6}(X \geq m - \frac{1}{2}) = P\Big(\frac{X-\frac{3}{5}n}{\sqrt{n\,0.6\,0.4}} \geq \frac{m-\frac{1}{2}-\frac{3}{5}n}{\sqrt{\frac{6}{25}n}}\Big) \approx 1-\Phi\Big(\frac{m-\frac{1}{2}-\frac{3}{5}n}{\sqrt{\frac{6}{25}n}}\Big).$$

Approximativ kann nun m als die minimale ganze Zahl mit

$$\frac{m-\frac{1}{2}-\frac{3}{5}n}{\sqrt{\frac{6}{25}n}} \geq \Phi^{-1}(1-\alpha)$$

gewählt werden.

2.2 Hypothesentests und Konfidenzbereiche

Sei $\tau : \Theta \to \Sigma$ die Kenngröße eines statistischen Modells und betrachte die Hypothese

$$H = \{\vartheta \in \Theta : \tau(\vartheta) \in H'\}$$

für eine Menge $H' \subset \Sigma$.

Satz 2.2. *Ist $C : \mathcal{X} \to \mathcal{P}(\Sigma)$ ein Konfidenzbereich für τ zum Niveau α, so definiert*

$$\mathcal{A} = \{x \in \mathcal{X} : C(x) \cap H' = \emptyset\}$$

den Verwerfungsbereich eines Hypothesentests für obiges H zum Niveau α.

Beweis. Für $\vartheta \in H$ gilt

$$\alpha \geq P_\vartheta(C(X) \not\ni \tau(\vartheta)) \geq P_\vartheta(C(X) \cap H' = \emptyset) = P_\vartheta(\mathcal{A}) = \beta(\vartheta).$$

□

Wir benutzen nun exemplarisch ein Konfidenzintervall im Gaußmodell zur Konstruktion eines Hypothesentests. Wir betrachten das

Statistisches Modell: $(\mathbb{R}^n, (\mathcal{N}(\mu, v)^{\otimes n})_{\mu \in \mathbb{R}, v>0})$ für ein $n \geq 2$

Ziel: Konstruktion eines Hypothesentests für

H: $\mu \leq \mu_0$, d.h. $H = (-\infty, \mu_0] \times (0, \infty)$, bzw. $\tau(\mu, v) = \mu$ und $H' = (-\infty, \mu_0]$

für einen fest spezifizierten Wert $\mu_0 \in \mathbb{R}$.

Erinnerung: Sei $q_{1-\alpha}$ das $1-\alpha$-Quantil der t_{n-1}-Verteilung. Die Menge

$$C(x) = [M(x) - \sqrt{\frac{V^*(x)}{n}} q_{1-\alpha}, \infty)$$

definiert ein Konfidenzintervall für τ zum Niveau α.

Der ensprechende Verwerfungsbereich ist nun

$$\mathcal{A} = \{x \in \mathcal{X} : C(x) \cap (-\infty, \mu_0] = \emptyset\} = \{x \in \mathcal{X} : \sqrt{\frac{n}{V^*(x)}}(M(x) - \mu_0) > q_{1-\alpha}\}$$

Alle Konfidenzintervalle von zuvor liefern Hypothesentests:

Satz 2.3. *Im Gaußmodell erhält man für ein gegebenes Niveau $\alpha \in (0,1)$ folgende Verwerfungsbereiche:*

i) Für die Hypothese $H : \mu \leq \mu_0$ erhält man

$$\mathcal{A} = \{x \in \mathcal{X} : \sqrt{\frac{n}{V^*(x)}}(M(x) - \mu_0) > q_{1-\alpha}\}$$

und für die Hypothese $H : \mu = \mu_0$

$$\mathcal{A} = \{x \in \mathcal{X} : \sqrt{\frac{n}{V^*(x)}}|M(x) - \mu_0| > q_{1-\frac{\alpha}{2}}\},$$

wobei die Quantile q von der t_{n-1}-Verteilung stammen.

ii) Für die Hypothese $H : v \leq v_0$ erhält man

$$\mathcal{A} = \{x \in \mathcal{X} : \frac{n-1}{q_{1-\alpha}} V^*(x) > v_0\},$$

und für $H : v = v_0$

$$\mathcal{A} = \{x \in \mathcal{X} : [\frac{n-1}{q_{1-\frac{\alpha}{2}}} V^*(x), \frac{n-1}{q_{\frac{\alpha}{2}}} V^*(x)] \not\ni v_0\},$$

wobei die Quantile q von der $\mathcal{X}^2_{n-1}$-Verteilung stammen.

2.3 Der Neyman-Pearson-Test (NP-Test)

Wir betrachten nun den Spezialfall einer einelementigen Hypothese und Alternative. Sei $\mathcal{M} = (\mathcal{X}, (P_\vartheta)_{\vartheta \in \{\theta_0, \theta_1\}})$ ein statistisches Standardmodell und betrachte

- die Hypothese $H = \{\theta_0\}$ und
- die Alternative $A = \{\theta_1\}$.

Wir betrachten eine ganze Familie von Tests, die sogennaten *Neyman-Pearson-Tests* (indexiert durch $c > 0$), die durch die Verwerfungsbereiche

$$\mathcal{A}_c = \{x \in \mathcal{X} : c\, p_{\theta_0}(x) \leq p_{\theta_1}(x)\}$$

beschrieben werden. Die Signifikanzniveaus der Tests sind

$$\alpha(c) := \beta_c(\theta_0) = \begin{cases} \int_{\mathcal{A}_c} p_{\theta_0}(x)\, dx & \text{wenn } \mathcal{M} \text{ ein stetiges Modell ist} \\ \sum_{x \in \mathcal{A}_c} p_{\theta_0}(x) & \text{sonst.} \end{cases}$$

Satz 2.4. *Für jedes $c > 0$ ist der Neyman-Pearson-Test mit Verwerfungsbereich $\mathcal{A}_c$, der Test mit maximaler Güte unter allen Tests zum Niveau $\alpha(c)$.*

Beweis. Wir betrachten nur den Fall in dem $\mathcal{M}$ ein diskretes Standardmodell ist. Wir halten $c > 0$ fest und setzen $\mathcal{A} = \mathcal{A}_c$. Sei $\tilde{\mathcal{A}}$ der Verwerfungsbereich eines weiteren Tests zum Niveau $\alpha(c)$. Dann gilt

$$\alpha(c) = \sum_{x \in \mathcal{A}} p_{\theta_0}(x) = \sum_{x \in \mathcal{A} \cap \tilde{\mathcal{A}}} p_{\theta_0}(x) + \sum_{x \in \mathcal{A} \setminus \tilde{\mathcal{A}}} p_{\theta_0}(x) \quad \text{und}$$
$$\alpha(c) \geq \sum_{x \in \tilde{\mathcal{A}}} p_{\theta_0}(x) = \sum_{x \in \mathcal{A} \cap \tilde{\mathcal{A}}} p_{\theta_0}(x) + \sum_{x \in \tilde{\mathcal{A}} \setminus \mathcal{A}} p_{\theta_0}(x)$$

sodass $\sum_{x \in \tilde{\mathcal{A}} \setminus \mathcal{A}} p_{\theta_0}(x) \leq \sum_{x \in \mathcal{A} \setminus \tilde{\mathcal{A}}} p_{\theta_0}(x)$. Ferner gilt für die Güte des zweiten Tests

$$\tilde{\beta}(\theta_1) = \sum_{x \in \tilde{\mathcal{A}}} p_{\theta_1}(x) = \sum_{x \in \mathcal{A} \cap \tilde{\mathcal{A}}} p_{\theta_1}(x) + \sum_{x \in \tilde{\mathcal{A}} \setminus \mathcal{A}} p_{\theta_1}(x)$$

Wir bemerken, dass in der letzten Summe für alle Summanden $p_{\theta_1}(x) < c p_{\theta_0}(x)$ gilt und wir folgern

$$\sum_{x \in \tilde{\mathcal{A}} \setminus \mathcal{A}} p_{\theta_1}(x) \leq c \sum_{x \in \tilde{\mathcal{A}} \setminus \mathcal{A}} p_{\theta_0}(x) \leq c \sum_{x \in \mathcal{A} \setminus \tilde{\mathcal{A}}} p_{\theta_0}(x) \leq \sum_{x \in \mathcal{A} \setminus \tilde{\mathcal{A}}} p_{\theta_1}(x).$$

Insgesamt gilt also

$$\tilde{\beta}(\theta_1) \leq \sum_{x \in \mathcal{A} \cap \tilde{\mathcal{A}}} p_{\theta_1}(x) + \sum_{x \in \mathcal{A} \setminus \tilde{\mathcal{A}}} p_{\theta_1}(x) = \beta(\theta_1).$$

□

Beispiel 2.5. Wir betrachten das statistische Modell mit $\mathcal{X} = [0, \infty)$, $P_{\theta_0} = \text{Exp}(1)$ und $P_{\theta_1} = \text{Exp}(2)$. Dann gilt für $x \geq 0$

$$p_{\theta_0}(x) = e^{-x} \text{ und } p_{\theta_1}(x) = 2e^{-2x}.$$

Die Neyman-Pearson Verwerfungsbereiche sind nun von der Form

$$\mathcal{A}_\delta = [0, \delta] \qquad (\delta > 0).$$

Für gegebenes Signifikanzniveau α ist δ so zu wählen, dass

$$\alpha \stackrel{!}{=} \beta_\delta(\theta_0) = \int_0^\delta e^{-x}\, dx = 1 - e^{-\delta}$$

gilt. Dies liefert die Güte

$$\beta_\delta(\theta_1) = \int_0^\delta 2e^{-2x}\, dx = 1 - e^{-2\delta} = 1 - (1-\alpha)^2.$$

2.4 Der p-Wert

Wir haben in den letzten zwei Unterkapiteln explizite Hypothesentests kennengelernt. De Facto handelt es sich hierbei um Familien von Hypothesentests: bei den Neyman-Pearson-Tests sind die Tests durch den Paramter c und bei den Hypothesentests im Gaußmodell durch den Parameter α indiziert. Man kann leicht nachprüfen, dass die Tests hierarchisch angeordnet sind: Die Akzeptanz der Hypothese im NP-Test mit Parameter c impliziert die Akzeptanz für alle größeren Parameter. Ferner impliziert die Akzeptanz der Hypothese im Gaußmodell zum Parameter α die Akzeptanz für alle kleineren Parameter. Der p-Wert einer Realisierung ist nun das Signifikanzniveaus des kritischen Tests, der gerade noch zur Akzeptanz bzw. Verwerfung der Hypothese führt.

Definition 2.6. Wir nennen eine Familie von Hypothesentests *hierarchisch*, wenn

- die Tests durch eine monoton fallende Familie von Verwerfungsbereichen $\mathcal{A}_c$ ($c \in I$) beschrieben werden, wobei $I \subset \mathbb{R}$ eine abgeschlossene Menge bezeichne, d.h.
$$\forall c_1, c_2 \in I \text{ mit } c_1 < c_2 : \mathcal{A}_{c_1} \supset \mathcal{A}_{c_2},$$
- die Familie der Verwerfungsbereiche rechtsstetig ist d.h. wenn für jede fallende konvergente Folge $(c_n)_{n \in I}$ mit $c_n \downarrow c$
$$\mathcal{A}_c = \bigcup_{n \in \mathbb{N}} \mathcal{A}_{c_n}$$
gilt und
- $\bigcap_{c \in I} \mathcal{A}_c = \emptyset$.

Seien im folgenden $(\mathcal{A}_c)_{c\in I}$ die Verwerfungsbereiche einer hierarchischen Familie von Hypothesentests. Wir bezeichnen mit

$$\alpha(c) := \sup_{\theta\in H} \beta_c(\theta) \qquad (c \in I)$$

die entsprechenden Signifikanzniveaus. Die Abbildung $\alpha : I \to [0,1]$ ist monoton fallend und wegen der Rechtsstetigkeit der Familie $(\mathcal{A}_c)$ auch rechtsseitig stetig (Übungsaufgabe).

Satz 2.7. *Für $x \in \mathcal{X}$ gibt es einen minimalen Index $c(x) \in I$ sodass für $c \in I$*

$$x \in \mathcal{H}_c := \mathcal{A}_c^c \;\Leftrightarrow\; c(x) \le c.$$

D.h. $c(x)$ ist der minimale Index, der für die Realisierung x zur Akzeptanz der Hypothese führt. Das entsprechende Signifikanzniveau

$$p(x) = \alpha(c(x))$$

*heißt p-*Wert *von x bzgl. der Familie $(\mathcal{A}_c)$. Der p-Wert hängt nich von der Parametrisierung der Familie ab.*

Viele Tests, die wir bisher betrachtet haben, haben Verwerfungsbereiche der Form

$$\mathcal{A}_c := \{x \in \mathcal{X} : T(x) > c\} \qquad (c \in \mathbb{R}),$$

wobei T eine Statistik ist, die unter der Hypothese H immer die gleiche Verteilung hat, sagen wir mit Verteilungsfunktion F. Man sieht leicht, dass die Mengen $(\mathcal{A}_c)_{c\in\mathbb{R}}$ eine hierarchische Familie von Tests beschreiben und es stellt sich die Frage nach den p-Werten.

Für gegebene Realisierung $x \in \mathcal{X}$ ist $c(x) = T(x)$ der minimale Index der zur Akzeptanz der Hyposthese führt und wir erhalten als p-Wert

$$p(x) = \alpha(c(x)) = \alpha(T(x)) = P_\vartheta(T(X) > T(x)) = 1 - F(T(x)).$$

Beispiel 2.8. Sei $\mu_0 \in \mathbb{R}$. Wählt man zum Beispiel im Gaußmodell

$$t(x) = \sqrt{\frac{n}{V^*(x)}}(M(x) - \mu_0)$$

und $T(x) := |t(x)|$ so erhält man die Verwerfungsbereiche aus Satz 2.3 für die Hypothese $H : \mu = \mu_0$. Der entsprechende p-Wert einer Realisierung $x \in \mathcal{X}$ ist

$$p(x) = 1 - F_T(T(x)) = P_{\mu_0,v}(T(X) > T(x)) = P_{\mu_0,v}(t(X) \notin [-T(x), T(x)]).$$

Die Statistik t ist unter $P_{\mu_0,v}$ t_{n-1}-verteilt und der p-Wert einer Realisierung x kann nun aus der Verteilungstabelle abgelesen werden.

Lemma 2.9. *Ist F die* stetige *Verteilungsfunktion einer Statistik $T : \mathcal{X} \to \mathbb{R}$ unter der Hypothese H, so ist die Abildung*

$$p : \mathcal{X} \to [0,1], x \mapsto 1 - F(T(x))$$

unter der Hypothese auf $[0,1]$ uniform verteilt. Ferner ist der Verwerfungsbereich der zugehörigen Familie von Hypothesentests zum Niveau $\alpha \in (0,1)$

$$\{x \in \mathcal{X} : p(x) < \alpha\}.$$

Bemerkung 2.10. Die meisten gängigen Hypothesentests sind in Form von hierarschischen Tests darstellbar und Statistikprogramme geben die entsprecheden p-Werte aus. Der Wert erlaubt einen Rückschluss darauf wie „unwahrscheinlich“ eine Hypothese ist und zu welchem Signifikanzniveau sie verworfen werden kann.