Wahrscheinlichkeitstheorie und Frequentistische Inferenz

BSc Psychologie WiSe 2024/25

Prof. Dr. Dirk Ostwald

(8) Grundbegriffe Frequentistischer Inferenz

Frequentistische Inferenzmodelle

Statistiken und Schätzer

Standardproblemstellungen und Standardannahme

Selbstkontrollfragen

Frequentistische Inferenzmodelle

Statistiken und Schätzer

Standardproblemstellungen und Standardannahme

Selbstkontrollfragen

Definition (Frequentistisches Inferenzmodell)

Ein *Frequentistisches Inferenzmodell* ist ein Tripel

$$\mathcal{M} := (\mathcal{Y}, \mathcal{A}, \{\mathbb{P}_\theta | \theta \in \Theta\}) \tag{1}$$

bestehend aus einem *Datenraum* $\mathcal{Y}$, einer σ-Algebra $\mathcal{A}$ auf $\mathcal{Y}$ und einer mindestens zweielementigen Menge $\{\mathbb{P}_\theta | \theta \in \Theta\}$ von Wahrscheinlichkeitsmaßen auf $(\mathcal{Y}, \mathcal{A})$, die durch $\theta \in \Theta$ indiziert sind. Wenn $\Theta \subset \mathbb{R}^k$ ist, heißt ein Frequentistisches Inferenzmodell auch *parametrisches* Frequentistisches Inferenzmodell und Θ heißt *Parameterraum* des Frequentistischen Inferenzmodells.

Bemerkungen

- Für Erwartungswerte und (Ko)Varianzen bezüglich $\mathbb{P}_\theta$ schreiben wir $\mathbb{E}_\theta$, $\mathbb{V}_\theta$, $\mathbb{C}_\theta$.
- Ein Frequentistisches Inferenzmodell $\mathcal{M}$ heißt ein *diskretes Modell*, wenn $\mathcal{Y}$ endlich oder abzählbar ist und jedes $\mathbb{P}_\theta$ eine WMF p_θ besitzt, ein Frequentistisches Inferenzmodell $\mathcal{M}$ heißt ein *stetiges Modell*, wenn $\mathcal{Y} \subseteq \mathbb{R}^n$ ist und jedes $\mathbb{P}_\theta$ eine WDF p_θ besitzt.
- Für ein Frequentistisches Inferenzmodell $\mathcal{M}_0 := (\mathcal{Y}_0, \mathcal{A}_0, \{\mathbb{P}_\theta^0 | \theta \in \Theta\})$ heißt das Modell $\mathcal{M} := (\mathcal{Y}, \mathcal{A}, \{\mathbb{P}_\theta | \theta \in \Theta\})$, für das $\mathcal{Y}$ das n-fache kartesische Produkt von $\mathcal{Y}_0$ mit sich selbst, $\mathcal{A}$ die entsprechende Produkt-σ-Algebra ist, und $\{\mathbb{P}_\theta | \theta \in \Theta\}$ die entsprechende Menge an Produktmaßen ist, das zu $\mathcal{M}_0$ gehörige *Produktmodell*.
- Wenn für ein Produktmodell die Menge $\mathcal{Y}_0$ eindimensional ist, also z.B. $\mathcal{Y}_0 = \mathbb{R}$ gilt, spricht man von einem *univariaten Frequentistischen Inferenzmodell*. Wenn für ein Produktmodell die Menge $\mathcal{Y}_0$ mehrdimensional ist, also z.B. $\mathcal{Y}_0 = \mathbb{R}^m, m > 1$ ist, spricht man von einem *multivariaten Frequentistischen Inferenzmodell*.

Frequentistische Inferenzmodelle

Bemerkungen (fortgeführt)

- Der Vorgang der Datenbeobachtung wird durch einen Zufallsvektor y, der Werte in $\mathcal{Y}$ annimmt, beschrieben. Im Kontext von Inferenzmodellen nennt man diesen Zufallsvektor *Daten*, *Beobachtung*, *Messung* oder *Stichprobe*. Eine Realisierung von y, also konkret vorliegende Datenwerte, werden *Datensatz*, *Beobachtungswert*, *Messwert* oder *Stichprobenwert* genannt.
- Produktmodelle modellieren die n-fache unabhängige Wiederholung eines Zufallsvorgangs. Der entsprechende Zufallsvektor $y := (y_1, ..., y_n)$ entspricht dann einer Menge von n u.i.v. Zufallsvariablen.
- Im Gegensatz zum Wahrscheinlichkeitsraummodell betrachtet man bei Frequentistische Inferenzmodellen zwei oder mehr Wahrscheinlichkeitsmaße, die die Verteilung von y mutmaßlich bestimmen. Das jeweils zugrundeliegende Wahrscheinlichkeitsmaß ist mit $\theta \in \Theta$ indiziert,
- In einem konkreten Datenanalyseproblem nimmt man an, dass die beobachteten Werte von $y = (y_1, ..., y_n)$ durch θ^* generiert wurde, wobei θ^* hier den *wahren, aber unbekannten, Parameterwert* bezeichnet. Der wahre, aber unbekannten, Parameterwert θ^* bleibt auch nach der Datenanalyse unbekannt. In der mathematischen Analyse von Inferenzmethoden betrachtet man alle möglichen wahren, aber unbekannten, Parameterwerte, schreibt also einfach $\{\mathbb{P}_\theta | \theta \in \Theta\}$.

Definition (Normalverteilungsmodell)

Das univariate parametrische Produktmodell

$$\mathcal{M} := (\mathcal{Y}, \mathcal{A}, \{\mathbb{P}_\theta | \theta \in \Theta\}) \tag{2}$$

mit

$$\mathcal{Y} := \mathbb{R}^n, \mathcal{A} := \mathcal{B}(\mathbb{R}^n), \theta := (\mu, \sigma^2), \Theta := \mathbb{R} \times \mathbb{R}_{>0}, \tag{3}$$

also

$$\{\mathbb{P}_\theta | \theta \in \Theta\} := \left\{ \prod_{i=1}^{n} N(\mu, \sigma^2) | (\mu, \sigma^2) \in \mathbb{R} \times \mathbb{R}_{>0} \right\}, \tag{4}$$

und damit

$$y_1, ..., y_n \sim N(\mu, \sigma^2) \text{ mit } (\mu, \sigma^2) \in \mathbb{R} \times \mathbb{R}_{>0} \tag{5}$$

heißt *Normalverteilungsmodell*.

Bemerkungen

- Das Normalverteilungsmodell ist Grundlage vieler populärer Inferenzverfahren.
- Diese Verfahren werden im Allgemeinen Linearen Modell integrativ betrachtet.
- Das Modell ist grundlegend durch normalverteilte Fehlerterme motiviert.

Definition (Bernoullimodell)

Das univariate parametrische Produktmodell

$$\mathcal{M} := (\mathcal{Y}, \mathcal{A}, \{\mathbb{P}_\theta | \theta \in \Theta\}) \tag{6}$$

mit

$$\mathcal{Y} := \{0,1\}^n, \mathcal{A} := \mathcal{P}\left(\{0,1\}^n\right), \theta := \mu, \Theta :=]0,1[, \tag{7}$$

also

$$\{\mathbb{P}_\theta | \theta \in \Theta\} := \left\{ \prod_{i=1}^{n} \text{Bern}(\mu) | \mu \in]0,1[\right\}, \tag{8}$$

und damit

$$y_1, ..., y_n \sim \text{Bern}(\mu) \text{ mit } \mu \in]0,1[, \tag{9}$$

heißt *Bernoullimodell*.

Bemerkung

- Das Bernoullimodell spielt in der Anwendung eine eher untegeordnete Rolle.

Frequentistische Inferenzmodelle

Statistiken und Schätzer

Standardproblemstellungen und Standardannahme

Selbstkontrollfragen

Definition (Statistik)

$\mathcal{M}$ sei ein Frequentistisches Inferenzmodell und $(\Sigma, \mathcal{S})$ sei ein Messraum. Dann wird eine Zufallsvariable der Form

$$S : \mathcal{Y} \to \Sigma \tag{10}$$

Statistik genannt.

Bemerkungen

- Daten und Statistiken werden durch Zufallsvariablen modelliert. Statistiken modellieren dabei von Datenwissenschaftler:innen konstruierte Funktionen, die bestenfalls datenbasierte Information liefern, aus der sich Schlüsse über die latenten datengenerierenden Prozesse ziehen lassen.

Beispiele (Statistiken)

$\mathcal{M}$ sei das Normalverteilungsmodell. Dann sind zum Beispiel folgende Zufallsvariablen Statistiken, was wir hier explizit durch die Notation deutlich machen wollen, was aber oft zur Vereinfachung der Notation implizit (aber trotzdem wichtig) bleibt:

- Das *Stichprobenmittel*

$$\bar{y} : \mathbb{R}^n \to \mathbb{R}, y \mapsto \bar{y}(y) := \frac{1}{n} \sum_{i=1}^{n} y_i, \tag{11}$$

- Die *Stichprobenvarianz*

$$s^2 : \mathbb{R}^n \to \mathbb{R}_{\geq 0}, y \mapsto s^2(y) := \frac{1}{n-1} \sum_{i=1}^{n} (y_i - \bar{y}(y))^2, \tag{12}$$

- Die *Stichprobenstandardabweichung*

$$s : \mathbb{R}^n \to \mathbb{R}_{\geq 0}, y \mapsto s(y) := \sqrt{s^2(y)}, \tag{13}$$

Definition (Schätzer)

$\mathcal{M}$ sei ein Frequentistisches Inferenzmodell, $(\Sigma, \mathcal{S})$ sei ein Messraum und $\tau : \Theta \to \Sigma$ sei eine Abbildung, die jedem $\theta \in \Theta$ eine Kenngröße $\tau(\theta) \in \Sigma$ zuordnet. Dann heißt eine Statistik

$$\hat{\tau} : \mathcal{Y} \to \Sigma \tag{14}$$

ein *Schätzer* für τ.

Bemerkungen

- Typische Beispiele für τ sind
  - $\tau(\theta) := \theta$ für die Schätzung von θ,
  - $\tau(\theta) := \theta_i$ mit $\theta \in \mathbb{R}^d, d > 1$ für die Schätzung einer Komponente von θ,
  - $\tau(\theta) := \mathbb{E}_\theta(y_1)$ für die Schätzung des Erwartungswert,
  - $\tau(\theta) := \mathbb{V}_\theta(y_1)$ für die Schätzung der Varianz.
- Für $\hat{\tau}$ bei $\tau(\theta) := \theta$ schreibt man üblicherweise $\hat{\theta}$.
- Schätzer nehmen Zahlwerte in Σ an und heißen deshalb auch *Punktschätzer*.
- Nicht jeder Schätzer ist ein guter Schätzer, man definiert deshalb *Schätzgütekriterien*.
- Gütekriterien für Schätzer sind der Inhalt von Vorlesungseinheit (9) Punktschätzung.

Beispiele (Schätzer)

$\mathcal{M}$ sei das Normalverteilungsmodell.

- Dann ist zum Beispiel das Stichprobenmittel $\bar{y} : \mathbb{R}^n \to \mathbb{R}$ ein Schätzer für

$$\tau : \mathbb{R} \times \mathbb{R}_{>0} \to \mathbb{R}, (\mu, \sigma^2) \mapsto \tau(\mu, \sigma^2) := \mu. \tag{15}$$

  Ebenso ist $\bar{y}$ ein Schätzer für

$$\tau : \mathbb{R} \times \mathbb{R}_{>0} \to \mathbb{R}, (\mu, \sigma_2) \mapsto \tau(\mu, \sigma^2) := \mathbb{E}_{\mu,\sigma^2}(y_1). \tag{16}$$

- Weiterhin ist die konstante Funktion

$$\hat{\tau} : \mathbb{R}^n \to \mathbb{R}, y \mapsto \hat{\tau}(y) := 42 \tag{17}$$

  ein Schätzer für

$$\tau : \mathbb{R} \times \mathbb{R}_{>0} \to \mathbb{R}_{>0}, (\mu, \sigma_2) \mapsto \tau(\mu, \sigma^2) := \sigma^2. \tag{18}$$

Dass eine Funktion $\hat{\tau} : \mathcal{Y} \to \Sigma$ ein Schätzer ist, heißt nicht, dass sie ein guter Schätzer ist!

Frequentistische Inferenzmodelle

Statistiken und Schätzer

Standardproblemstellungen und Standardannahme

Selbstkontrollfragen

Standardproblemstellungen und Standardannahme

Mithilfe von Modellen behandelt die Frequentistische Inferenz folgende Standardproblemstellungen:

(1) Punktschätzung

Ziel der Punktschätzung ist es, einen möglichst guten Tipp für wahre, aber unbekannte, Parameterwerte oder Funktionen dieser abzugeben, typischerweise mithilfe der Daten.

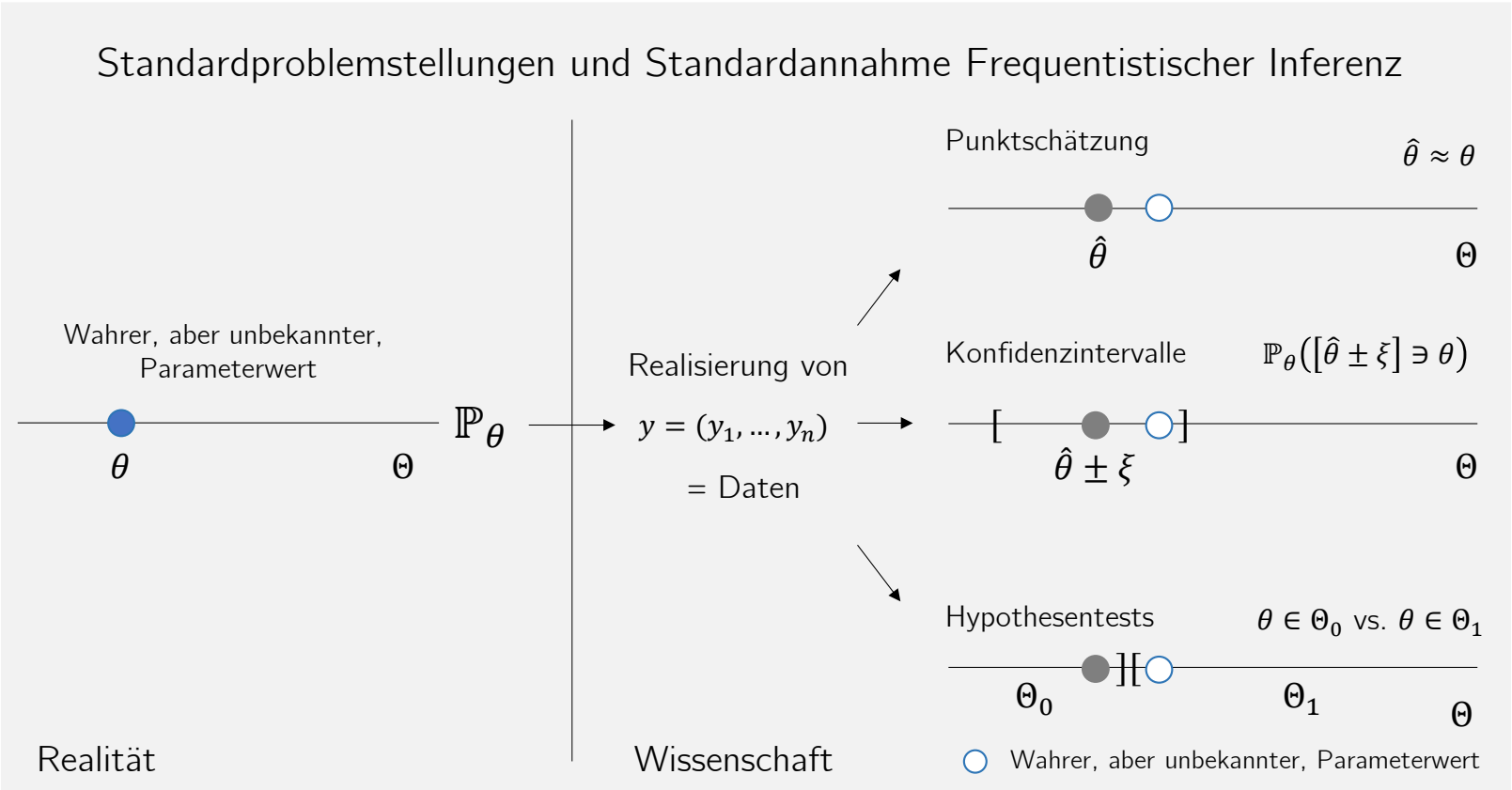
(2) Konfidenzintervalle

Ziel der Konfidenzintervallbestimmung ist es, basierend auf der angenommenen Datenverteilung und den beobachteten Daten durch eine Intervallschätzung einen möglichst sicheren, wenn auch oft unpräzisen, Tipp für wahre, aber unbekannte, Parameterwerte abzugeben.

(3) Hypothesentests

Ziel des Hypothesentestens ist es, basierend auf der angenommenen Verteilung der Daten in einer möglichst zuverlässigen Form zu entscheiden, ob ein wahrer, aber unbekannter, Parameterwert in einer von zwei sich gegenseitig ausschließenden Untermengen des Parameterraumes liegt.

Standardproblemstellungen und Standardannahme



Standardannahme der Frequentistischen Inferenz

$\mathcal{M}$ sei ein Frequentistisches Inferenzmodell mit Zufallsvektor y.

Es wird angenommen, dass ein vorliegender Datensatz eine der möglichen Realisierungen von y ist.

Aus Frequentistischer Sicht kann man eine Studie unendlich oft wiederholen und zu jedem Datensatz Schätzer oder Statistiken auswerten, z.B. das Stichprobenmittel:

Datensatz (1) : $y^{(1)} = \left(y_1^{(1)}, y_2^{(1)}, ..., y_n^{(1)}\right)$ mit $\bar{y}^{(1)} = \frac{1}{n}\sum_{i=1}^{n} y_i^{(1)}$

Datensatz (2) : $y^{(2)} = \left(y_1^{(2)}, y_2^{(2)}, ..., y_n^{(2)}\right)$ mit $\bar{y}^{(2)} = \frac{1}{n}\sum_{i=1}^{n} y_i^{(2)}$

Datensatz (3) : $y^{(3)} = \left(y_1^{(3)}, y_2^{(3)}, ..., y_n^{(3)}\right)$ mit $\bar{y}^{(3)} = \frac{1}{n}\sum_{i=1}^{n} y_i^{(3)}$

Datensatz (4) : $y^{(4)} = \left(y_1^{(4)}, y_2^{(4)}, ..., y_n^{(4)}\right)$ mit $\bar{y}^{(4)} = \frac{1}{n}\sum_{i=1}^{n} y_i^{(4)}$

Datensatz (5) : $y^{(5)} = ...$

Um die Qualität ihrer Methoden zu beurteilen betrachtet die Frequentistische Inferenz deshalb die Wahrscheinlichkeitsverteilungen von Schätzern und Statistiken. Was zum Beispiel ist die Verteilung der $\bar{y}^{(1)}$, $\bar{y}^{(2)}$, $\bar{y}^{(3)}$, $\bar{y}^{(4)}$, ... also die Verteilung der Zufallsvariable $\bar{y}_n$?

Wenn eine Methode im Sinne der Frequentistischen Standardannahme "gut" ist, dann heißt das also, dass sie bei häufiger Anwendung "im Mittel gut" ist. Im Einzelfall, also im Normalfall nur eines vorliegenden Datensatzes, kann sie auch "schlecht" sein.

Beispiel | Evidenzbasierte Evaluation von Psychotherapie bei Depression

Pre-BDI

n = 12

BDI-II Fragebogen

Name | Alter | Geschlecht m / w | Datum

Anleitung: Dieser Fragebogen enthält 21 Gruppen von Aussagen. Bitte lesen Sie jede dieser Gruppen von Aussagen sorgfältig durch und suchen Sie sich dann in jeder Gruppe **eine Aussage** heraus, die am besten beschreibt, wie Sie sich **in den letzten zwei Wochen, einschließlich heute, gefühlt haben**. Kreuzen Sie die Zahl neben der Aussage an, die Sie sich herausgesucht haben (0, 1, 2 oder 3). Falls in einer Gruppe mehrere Aussagen gleichermaßen auf Sie zutreffen, kreuzen Sie die Aussage mit der höheren Zahl an. Achten Sie bitte darauf, dass Sie in jeder Gruppe nicht mehr als eine Aussage ankreuzen, das gilt auch für Gruppe 16 (Veränderungen der Schlafgewohnheiten) oder Gruppe 18 (Veränderungen des Appetits).

1.) Traurigkeit
0 Ich bin nicht traurig.
1 Ich bin oft traurig.
2 Ich bin ständig traurig.
3 Ich bin so traurig oder unglücklich, dass ich es nicht aushalte.

2.) Pessimismus
0 Ich sehe nicht mutlos in die Zukunft.
1 Ich sehe mutloser in die Zukunft als sonst.
2 Ich bin mutlos und erwarte nicht, dass meine Situation besser wird.
3 Ich glaube, dass meine Zukunft hoffnungslos ist und nur noch schlechter wird.

3.) Versagensgefühle
0 Ich fühle mich nicht als Versager.
1 Ich habe häufiger Versagensgefühle.
2 Wenn ich zurückblicke, sehe ich eine Menge Fehlschläge.
3 Ich habe das Gefühl, als Mensch ein völliger Versager zu sein.

4.) Verlust von Freude
0 Ich kann die Dinge genauso gut genießen wie früher.
1 Ich kann die Dinge nicht mehr so genießen wie früher.
2 Dinge, die mir früher Freude gemacht haben, kann ich kaum mehr genießen.
3 Dinge, die mir früher Freude gemacht haben, kann ich überhaupt nicht mehr genießen.

5.) Schuldgefühle
0 Ich habe keine besonderen Schuldgefühle.
1 Ich habe oft Schuldgefühle wegen Dingen, die ich getan habe oder hätte tun sollen.
2 Ich habe die meiste Zeit Schuldgefühle.
3 Ich habe ständig Schuldgefühle.

6.) Bestrafungsgefühle
0 Ich habe nicht das Gefühl, für etwas bestraft zu sein.
1 Ich habe das Gefühl, vielleicht bestraft zu werden.
2 Ich erwarte, bestraft zu werden.
3 Ich habe das Gefühl, bestraft zu sein.

7.) Selbstablehnung
0 Ich halte von mir genauso viel wie immer.
1 Ich habe Vertrauen in mich verloren.
2 Ich bin von mir enttäuscht.
3 Ich lehne mich völlig ab.

8.) Selbstvorwürfe
0 Ich kritisiere oder tadle mich nicht mehr als sonst.
1 Ich bin mir gegenüber kritischer als sonst.
2 Ich kritisiere mich für all meine Mängel.
3 Ich gebe mir die Schuld für alles Schlimme, was passiert.

9.) Selbstmordgedanken
0 Ich denke nicht daran, mir etwas anzutun.
1 Ich denke manchmal an Selbstmord, aber ich würde es nicht tun.
2 Ich möchte mich am liebsten umbringen.
3 Ich würde mich umbringen, wenn ich die Gelegenheit dazu hätte.

10.) Weinen
0 Ich weine nicht öfter als früher.
1 Ich weine jetzt mehr als früher.
2 Ich weine beim geringsten Anlass.
3 Ich möchte gern weinen, aber ich kann nicht.

PEARSON PsychCorp

Summe Seite 1:

Post-BDI

⇒ Pre-Post BDI Score Reduktion

Beispiel | Evidenzbasierte Evaluation von Psychotherapie bei Depression

i	dBDI
1	-1
2	3
3	-2
4	9
5	3
6	-2
7	4
8	5
9	5
10	1
11	9
12	4

Beispiel | Evidenzbasierte Evaluation von Psychotherapie bei Depression

Für die Pre-Post BDI Score Reduktion y_i der $i$ten von n Patient:innen legen wir das Modell

$$y_i = \mu + \varepsilon_i \text{ mit } \varepsilon_i \sim N(0, \sigma^2) \text{ u.i.v. für } i = 1, ..., n \tag{19}$$

zugrunde. Dabei wird die Pre-Post BDI Reduktion y_i der $i$ten Patient:in also mithilfe einer über die Gruppe von Patient:innen identischen Pre-Post BDI Score Reduktion $\mu \in \mathbb{R}$ und einer Patient:innen-spezifischen normalverteilten Pre-Post BDI Score Reduktionsabweichung ε_i erklärt.

Wie unten gezeigt ist dieses Modell äquivalent zum oben eingeführten Normalverteilungsmodell

$$y_1, ..., y_n \sim N(\mu, \sigma^2). \tag{20}$$

Die Standardproblemstellungen der Frequentistischen Inferenz führen in diesem Szenario auf folgende Fragen:

(1) Was sind sinnvolle Tipps für die wahren, aber unbekannten, Parameterwerte μ und σ^2?
(2) Wie gelingt im Sinne einer Intervallschätzung eine möglichst sichere Schätzung von μ und σ^2?
(3) Entscheiden wir uns sinnvollerweise für die Hypothese, dass gilt $\mu \neq 0$?

Beispiel | Evidenzbasierte Evaluation von Psychotherapie bei Depression

Die Äquivalenz beider Modellformen folgt direkt aus der Transformation normalverteilter Zufallsvariablen durch linear-affine Funktionen. Speziell gilt im vorliegenden Fall für $\varepsilon_i \sim N(0, \sigma^2)$ u.i.v. mit $i = 1, ..., n$, dass

$$y_i = f(\varepsilon_i) \text{ mit } f : \mathbb{R} \to \mathbb{R}, e_i \mapsto f(e_i) := e_i + \mu. \tag{21}$$

Dann gilt

$$\begin{aligned} p(y_i) &= \frac{1}{|1|} p_{\varepsilon_i} \left(\frac{y_i - \mu}{1} \right) \\ &= N\left(y_i - \mu; 0, \sigma^2\right) \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2} (y_i - \mu - 0)^2 \right) \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2} (y_i - \mu)^2 \right) \\ &= N(y_i; \mu, \sigma^2), \end{aligned} \tag{22}$$

also $y_i \sim N(\mu, \sigma^2)$ u.i.v. und damit $y_1, ..., y_n \sim N(\mu, \sigma^2)$, wobei wir notationell nicht zwischen der Zufallsvariable y_i und ihren Werten unterschieden haben.

Frequentistische Inferenzmodelle

Statistiken und Schätzer

Standardproblemstellungen und Standardannahme

Selbstkontrollfragen

Selbstkontrollfragen

1. Geben Sie die Definition eines parametrischen Frequentistischen Inferenzmodells wieder.

2. Was ist der Unterschied zwischen univariaten und multivariaten Produktmodell?

3. Geben Sie die Definition des Normalverteilungsmodells wieder.

4. Geben Sie die Definition des Begriffs der Statistik wieder.

5. Geben Sie die Definition des Begriffs des Schätzers wieder.

6. Erläutern Sie die Standardproblemstellungen der Frequentistischen Inferenz.

7. Erläutern Sie die Standardannahme der Frequentistischen Inferenz.

Selbstkontrollfragen - Lösungen

1. Siehe Definition (Frequentistischen Inferenmodell)
2. Bei einem univariaten Produktmodell ist der Datenraum eindimensional und ein einzelner Datenpunkt ist ein Skalar. Bei einem multivariaten Produktmodell ist der Datenraum mehrdimensional und ein einzelner Datenpunkt ist ein (mehrdimensionaler) Vektor.
3. Siehe Definition (Normalverteilungsmodell).
4. Siehe Definition (Statistik).
5. Siehe Definition (Schätzer).
6. Siehe Folien 15 und 16.
7. Siehe Folie 17.