Wahrscheinlichkeitstheorie und Frequentistische Inferenz

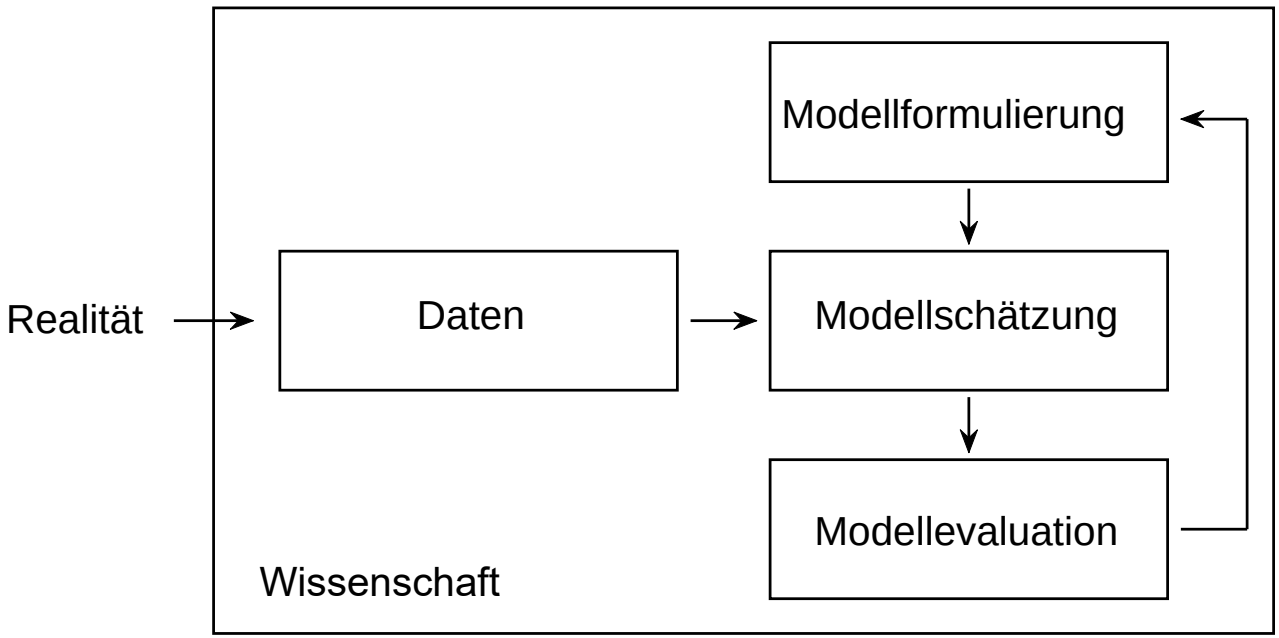
BSc Psychologie WiSe 2022/23

Prof. Dr. Dirk Ostwald

(9) Grundbegriffe Frequentistischer Inferenz

Datum	Einheit	Thema
13.10.2022	Einführung	(1) Einführung
20.10.2022	Wahrscheinlichkeitstheorie	(2) Wahrscheinlichkeitsräume
27.10.2022	Wahrscheinlichkeitstheorie	(3) Elementare Wahrscheinlichkeiten
03.11.2022	Wahrscheinlichkeitstheorie	(4) Zufallsvariablen I
10.11.2022	Wahrscheinlichkeitstheorie	(4) Zufallsvariablen II
17.11.2022	Wahrscheinlichkeitstheorie	(5) Multivariate Verteilungen
24.11.2022	Wahrscheinlichkeitstheorie	(6) Erwartungswert und Kovarianz
01.12.2022	Wahrscheinlichkeitstheorie	(7) Ungleichungen und Grenzwerte
08.12.2022	Wahrscheinlichkeitstheorie	(8) Transformationen der Normalverteilung
15.12.2022	Frequentistische Inferenz	(9) Grundbegriffe Frequentistischer Inferenz
	Weihnachtspause	
05.01.2023	Frequentistische Inferenz	(10) Parameterschätzung
12.01.2023	Frequentistische Inferenz	(11) Konfidenzintervalle
19.01.2023	Frequentistische Inferenz	(12) Hypothesentests I
26.01.2023	Frequentistische Inferenz	(12) Hypothesentests II
02.02.2023	Klausur	G44-H6, 16:00 - 17:00 Uhr
Jul 2023	Klausurwiederholungstermin	

Modellbasierte Datenwissenschaft



Frequentistische Inferenz

Modellformulierung	$\Rightarrow$ Statistische Modelle
Modellschätzung	$\Rightarrow$ Parameterschätzung und Konfidenzintervalle
Modellevaluation	$\Rightarrow$ Hypothesentests (cf. Allgemeines Lineares Modell)

Statistische Modelle

Statistiken und Schätzer

Standardprobleme Frequentistischer Inferenz

Selbstkontrollfragen

Statistische Modelle

Statistiken und Schätzer

Standardprobleme Frequentistischer Inferenz

Selbstkontrollfragen

Definition (Statistisches Modell)

Ein *statistisches Modell* 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 statistisches Modell auch *parametrisches* statistisches Modell und Θ heißt *Parameterraum* des statistischen Modells.

Bemerkungen

- Für Erwartungswerte und (Ko)Varianzen bezüglich $\mathbb{P}_\theta$ schreiben wir $\mathbb{E}_\theta$, $\mathbb{V}_\theta$, $\mathbb{C}_\theta$.
- Ein statistisches Modell $\mathcal{M}$ heißt ein *diskretes Modell*, wenn $\mathcal{Y}$ diskret ist und jedes $\mathbb{P}_\theta$ eine WMF p_θ besitzt, ein statistisches Modell $\mathcal{M}$ heißt ein *stetiges Modell*, wenn $\mathcal{Y} \subset \mathbb{R}^n$ ist und jedes $\mathbb{P}_\theta$ eine WDF p_θ besitzt.
- Für ein statistisches Modell $\mathcal{M}_0 := (\mathcal{Y}_0, \mathcal{A}_0, \{\mathbb{P}_\theta^0 | \theta \in \Theta\})$ heißt das statistische 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 statistischen Modell*. 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 statistischen Modell*.

Statistische Modelle

Bemerkungen (fortgeführt)

- Der Vorgang der Datenbeobachtung wird durch einen Zufallsvektor y, der Werte in $\mathcal{Y}$ annimmt, beschrieben. Im Kontext statistischer Modelle nennt man diesen Zufallsvektor *Daten*, *Beobachtung*, *Messung* oder *Stichprobe*. Eine Realisierung von y, also konkret vorliegende Datenwerte $\tilde{y} \in \mathcal{Y}$, werden *Datensatz*, *Beobachtungswert*, *Messwert* oder *Stichprobenwert* genannt.
- Produktmodelle modellieren die n-fache unabhängige Wiederholung eines Zufallsvorgangs. Der entsprechende Zufallsvektor $y := (y_1, \ldots, y_n)$ entspricht dann einer Menge von n unabhängigen Zufallsvariablen/vektoren.
- Im Gegensatz zum Wahrscheinlichkeitsraummodell betrachtet man bei statistische Modellen 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 $\tilde{y} = (\tilde{y}_1, \ldots, \tilde{y}_n)$ von $y = (y_1, \ldots, y_n)$ durch θ^* generiert wurde, wobei θ^* hier den *wahren, aber unbekannten, Parameterwert* bezeichnet. Der wahre, aber unbekannten, Parameterwert θ^* bleibt auch nach der statistischen Analyse 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

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

heißt *Normalverteilungsmodell*.

Bemerkungen

- Das Normalverteilungsmodell ist Grundlage vieler populärer statistischen Verfahren.
- 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

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

heißt *Bernoullimodell*.

Bemerkung

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

Statistische Modelle

Statistiken und Schätzer

Standardprobleme Frequentistischer Inferenz

Selbstkontrollfragen

Definition (Statistik)

$\mathcal{M}$ sei ein statistisches Modell 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*

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

- Die *Stichprobenvarianz*

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

- Die *Stichprobenstandardabweichung*

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

Definition (Schätzer)

$\mathcal{M}$ sei ein statistisches Modell, $(\Sigma, \mathcal{S})$ 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 (10) Parameterschätzung.

Beispiele (Schätzer)

$\mathcal{M}$ sei das Normalverteilungsmodell.

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

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

  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!

Statistische Modelle

Statistiken und Schätzer

Standardprobleme Frequentistischer Inferenz

Selbstkontrollfragen

Standardprobleme Frequentistischer Inferenz

Mithilfe statistischer Modelle behandelt die Frequentistische Inferenz folgende Standardprobleme:

(1) Parameterschätzung

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

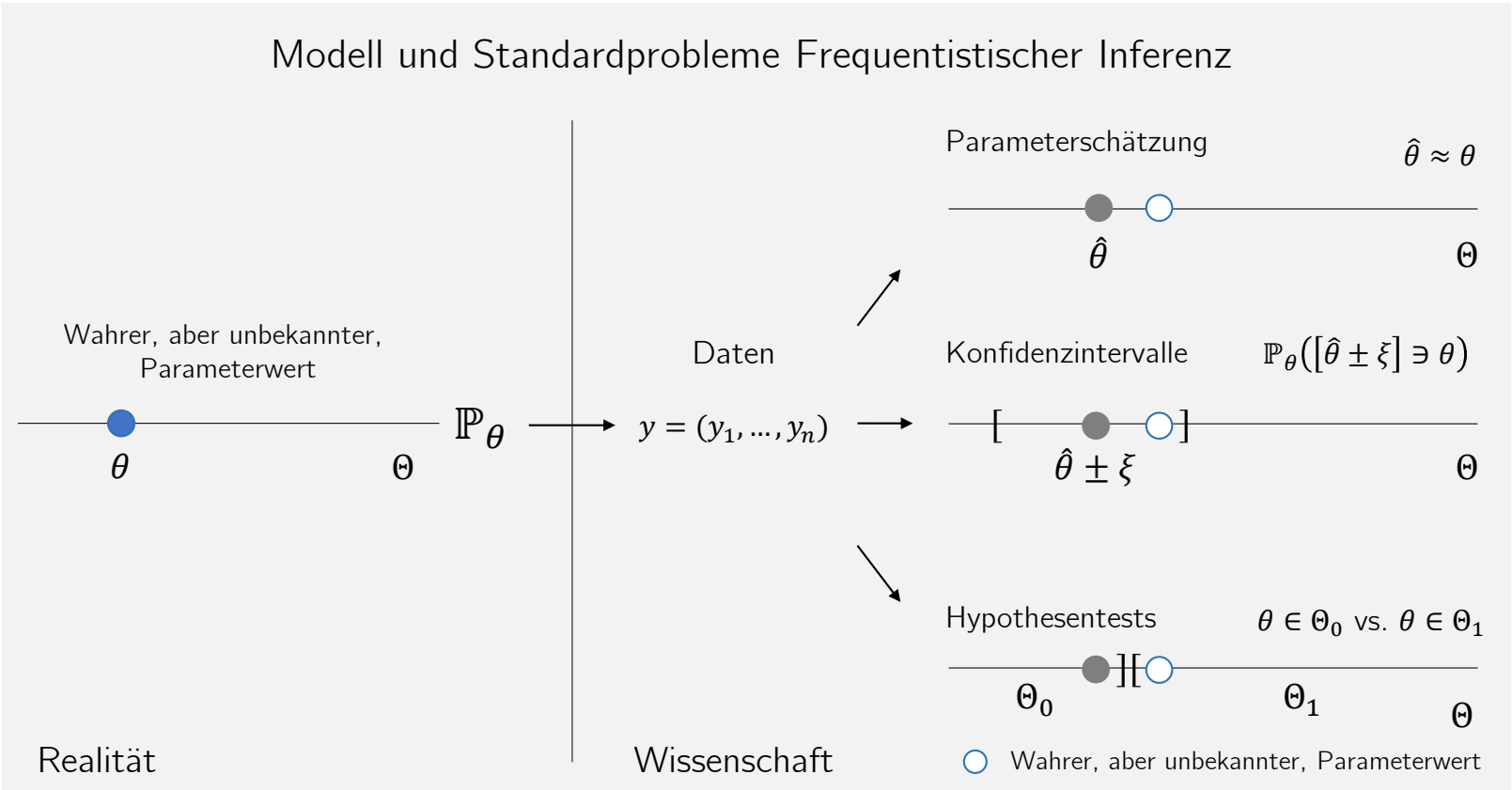
(2) Konfidenzintervalle

Ziel der Bestimmung von Konfidenzintervallen ist es, basierend auf der angenommenen Verteilung der Daten eine quantitative Aussage über die mit Schätzwerten assoziierte Unsicherheit zu treffen.

(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.

Standardprobleme Frequentistischer Inferenz



Standardannahmen frequentistischer Inferenz

$\mathcal{M}$ sei ein statistisches Modell mit $y_1, \ldots, y_n \sim p_\theta$.

Es wird angenommen, dass ein konkreter Datensatz eine der möglichen Realisierungen von $y_1, \ldots, y_n \sim p_\theta$ ist.

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

$$\text{Datensatz (1)} : \tilde{y}^{(1)} = \left(\tilde{y}_1^{(1)}, \tilde{y}_2^{(1)}, \ldots, \tilde{y}_n^{(1)}\right) \text{ mit } \bar{y}_n^{(1)} = \frac{1}{n}\sum_{i=1}^n \tilde{y}_i^{(1)}$$

$$\text{Datensatz (2)} : \tilde{y}^{(2)} = \left(\tilde{y}_1^{(2)}, \tilde{y}_2^{(2)}, \ldots, \tilde{y}_n^{(2)}\right) \text{ mit } \bar{y}_n^{(2)} = \frac{1}{n}\sum_{i=1}^n \tilde{y}_i^{(2)}$$

$$\text{Datensatz (3)} : \tilde{y}^{(3)} = \left(\tilde{y}_1^{(3)}, \tilde{y}_2^{(3)}, \ldots, \tilde{y}_n^{(3)}\right) \text{ mit } \bar{y}_n^{(3)} = \frac{1}{n}\sum_{i=1}^n \tilde{y}_i^{(3)}$$

$$\text{Datensatz (4)} : \tilde{y}^{(4)} = \left(\tilde{y}_1^{(4)}, \tilde{y}_2^{(4)}, \ldots, \tilde{y}_n^{(4)}\right) \text{ mit } \bar{y}_n^{(4)} = \frac{1}{n}\sum_{i=1}^n \tilde{y}_i^{(4)}$$

$$\text{Datensatz (5)} : \tilde{y}^{(5)} = \ldots$$

Um die Qualität statistischer Methoden zu beurteilen betrachtet die Frequentistische Statistik deshalb die Wahrscheinlichkeitsverteilungen von Schätzern und Statistiken unter Annahme von $y_1, \ldots, y_n \sim p_\theta$. Was zum Beispiel ist die Verteilung von $\bar{y}_n^{(1)}$, $\bar{y}_n^{(2)}$, $\bar{y}_n^{(3)}$, $\bar{y}_n^{(4)}$, $\ldots$ also die Verteilung der Zufallsvariable $\bar{y}_n$?

Wenn eine statistische Methode im Sinne der Frequentistischen Standardannahmen "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

```
fname = file.path(getwd(), "9_Grundbegriffe_Frequentistischer_Inferenz.csv")
D     = read.table(fname, sep = ",", header = T)
```

i	BDI.Reduktion
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 Standardprobleme 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 hoch ist im frequentistischen Sinn die mit diesen Tipps assoziierte Unsicherheit?
(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., dass

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

Dann gilt

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

also $y_i \sim N(\mu, \sigma^2)$ u.i.v. und damit $y_1, \ldots, y_n \sim N(\mu, \sigma^2)$.

Statistische Modelle

Statistiken und Schätzer

Standardprobleme Frequentistischer Inferenz

Selbstkontrollfragen

Selbstkontrollfragen

1. Definieren und erläutern Sie den Begriff des parametrischen statistischen Modells.
2. Definieren und erläutern Sie den Begriff eine parametrischen statistischen Produktmodells.
3. Erläutern Sie den Unterschied zwischen univariaten und multivariaten statistischen Modellen.
4. Formulieren und erläutern Sie das Normalverteilungsmodell.
5. Formulieren und erläutern Sie das Bernoullimodell.
6. Definieren und erläutern Sie den Begriff der Statistik.
7. Definieren und erläutern Sie den Begriff des Schätzers.
8. Nennen und erläutern Sie die Standardprobleme der frequentistischen Inferenz.
9. Erläutern Sie die Standardannahmen der frequentistischen Inferenz.