



# Wahrscheinlichkeitstheorie und Frequentistische Inferenz

BSc Psychologie WiSe 2023/24

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

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

bestehend aus einem *Datenraum*  $\mathcal{Y}$ , einer  $\sigma$ -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  $\Theta$  heißt *Parameterraum* des Frequentistischen Inferenzmodells.

### Bemerkungen

- Für Erwartungswerte und (Ko)Varianzen bezüglich  $\mathbb{P}_\theta$  schreiben wir  $E_\theta, V_\theta, C_\theta$ .
- Ein Frequentistisches Inferenzmodell  $\mathcal{M}$  heißt ein *diskretes Modell*, wenn  $\mathcal{Y}$  endlich oder abzählbar ist und jedes  $\mathbb{P}_\theta$  eine WMF  $p_\theta$  besitzt, ein Frequentistisches Inferenzmodell  $\mathcal{M}$  heißt ein *stetiges Modell*, wenn  $\mathcal{Y} \subseteq \mathbb{R}^m$  ist und jedes  $\mathbb{P}_\theta$  eine WDF  $p_\theta$  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- $\sigma$ -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*.

## Bemerkungen (fortgeführt)

- Der Vorgang der Datenbeobachtung wird durch einen Zufallsvektor  $v$ , der Werte in  $\mathcal{Y}$  annimmt, beschrieben. Im Kontext von Inferenzmodellen nennt man diesen Zufallsvektor *Daten*, *Beobachtung*, *Messung* oder *Stichprobe*. Eine Realisierung von  $v$ , also konkret vorliegende Datenwerte  $y \in \mathcal{Y}$ , werden *Datensatz*, *Beobachtungswert*, *Messwert* oder *Stichprobenwert* genannt.
- Produktmodelle modellieren die  $n$ -fache unabhängige Wiederholung eines Zufallsvorgangs. Der entsprechende Zufallsvektor  $v := (v_1, \dots, v_n)$  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  $v$  mutmaßlich bestimmen. Das jeweils zugrundeliegende Wahrscheinlichkeitsmaß ist mit  $\theta \in \Theta$  indiziert,
- In einem konkreten Datenanalyseproblem nimmt man an, dass die beobachteten Werte  $y = (y_1, \dots, y_n)$  von  $v = (v_1, \dots, v_n)$  durch  $\theta^*$  generiert wurde, wobei  $\theta^*$  hier den *wahren, aber unbekanntem, Parameterwert* bezeichnet. Der wahre, aber unbekanntem, Parameterwert  $\theta^*$  bleibt auch nach der Datenanalyse unbekannt. In der mathematischen Analyse von Inferenzmethoden betrachtet man alle möglichen wahren, aber unbekanntem, Parameterwerte, schreibt also einfach  $\{\mathbb{P}_\theta | \theta \in \Theta\}$ .

## Definition (Normalverteilungsmodell)

Das univariate parametrische Produktmodell

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

mit

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

also

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

und damit

$$v_1, \dots, v_n \sim N(\mu, \sigma^2) \text{ mit } (\mu, \sigma^2) \in \mathbb{R} \times \mathbb{R}_{>0} \quad (5)$$

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

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

mit

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

also

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

und damit

$$v_1, \dots, v_n \sim \text{Bern}(\mu) \text{ mit } \mu \in ]0, 1[ \quad (9)$$

heißt *Bernoullimodell*.

Bemerkung

- Das Bernoullimodell spielt in der Anwendung eine eher untergeordnete 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

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

*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*

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

- Die *Stichprobenvarianz*

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

- Die *Stichprobenstandardabweichung*

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

## Definition (Schätzer)

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

$$\hat{\tau} : \mathcal{Y} \rightarrow \Sigma \quad (14)$$

ein *Schätzer* für  $\tau$ .

### Bemerkungen

- Typische Beispiele für  $\tau$  sind
  - $\tau(\theta) := \theta$  für die Schätzung von  $\theta$ ,
  - $\tau(\theta) := \theta_i$  mit  $\theta \in \mathbb{R}^d$ ,  $d > 1$  für die Schätzung einer Komponente von  $\theta$ ,
  - $\tau(\theta) := \mathbb{E}_\theta(y_1)$  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  $\Sigma$  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  $\bar{y} : \mathbb{R}^n \rightarrow \mathbb{R}$  ein Schätzer für

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

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

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

- Weiterhin ist die konstante Funktion

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

ein Schätzer für

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

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

---

Frequentistische Inferenzmodelle

Statistiken und Schätzer

**Standardproblemstellungen und Standardannahme**

Selbstkontrollfragen

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.

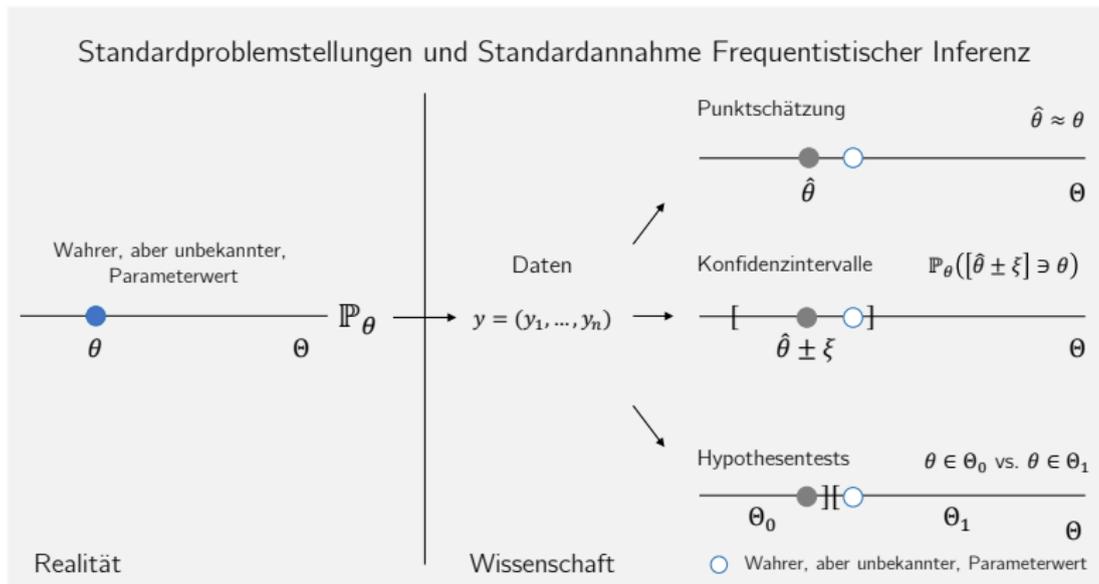
## (2) Konfidenzintervalle

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



# Standardproblemstellungen und Standardannahme

## Standardannahme Frequentistischer Inferenz

$\mathcal{M}$  sei ein Frequentistisches Inferenzmodell mit Zufallsvektor  $v$ . Es wird angenommen, dass ein Datensatz  $y \in \mathbb{R}^n$  eine der möglichen Realisierungen von  $v$  ist.

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

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

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

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

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

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

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  $\bar{y}^{(1)}, \bar{y}^{(2)}, \bar{y}^{(3)}, \bar{y}^{(4)}, \dots$  also die Verteilung der Zufallsvariable  $\bar{v}_n$ ?

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.

# Standardproblemstellungen und Standardannahme

## Beispiel | Evidenzbasierte Evaluation von Psychotherapie bei Depression

Pre-BDI



n = 12

Post-BDI

⇒ Pre-Post BDI Score Reduktion

BDI-II		Fragebogen	
Name		Alter	Erkrankt
		von 18	ja/nein
<b>Aufgabe:</b> Dieser Fragebogen enthält 21 Gruppen von Aussagen. Bitte lesen Sie jede dieser Aussagen von Anfang bis Ende und wählen 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 für Sie am besten passt. Es ist kein Ziel, alle Aussagen richtig zu beantworten, sondern für die Aussage, mit der Sie sich am besten identifizieren können, diejenige auszuwählen, die am besten beschreibt, wie Sie sich in den letzten zwei Wochen gefühlt haben. Bitte beachten Sie, dass es sich nicht um Fragen zu Verhaltensweisen, die Sie planen zu tun, sondern um Aussagen über das, was Sie in der Vergangenheit getan haben.			
<b>1.) Traurigkeit</b>		<b>6.) Berufungsgefühle</b>	
1 Ich bin nicht traurig.	2 Ich bin oft traurig.	3 Ich bin ständig traurig.	4 Ich bin so traurig oder unglücklich, dass ich es nicht aushalte.
<b>3.) Passivitäten</b>		<b>7.) Selbstabwertung</b>	
1 Ich sehe nicht mehr in die Zukunft als sonst.	2 Ich sehe nicht mehr in die Zukunft als sonst.	3 Ich bin müde und erwarte nicht, dass Dinge glücklicher werden.	4 Ich glaube, dass meine Zukunft hoffnungslos ist und nur noch schlechter wird.
<b>3.) Verweigerungsgefühle</b>		<b>8.) Selbstverurteilung</b>	
1 Ich fühle mich nicht als Verweigerer.	2 Ich habe häufiger Verweigerungsgefühle.	3 Wenn ich zurückblicke, sehe ich eine Menge Missetzungen.	4 Ich fühle das Gefühl, ein Mensch zu sein, der viele Missetzungen begangen hat.
<b>4.) Verlust von Freude</b>		<b>9.) Selbstmordgedanken</b>	
1 Ich kann die Dinge genauso gut genießen wie früher.	2 Ich kann die Dinge nicht mehr so genießen wie früher.	3 Ich denke, ich würde mich selbst verletzen, aber ich würde es nicht tun.	4 Ich möchte mich selbst verletzen, aber ich bin zu ängstlich, dies zu tun.
<b>5.) Schuldgefühle</b>		<b>10.) Werten</b>	
1 Ich habe keine besonderen Schuldgefühle.	2 Ich habe oft Schuldgefühle wegen Dingen, die ich getan habe oder habe tun sollen.	3 Ich habe die meiste Zeit Schuldgefühle.	4 Ich habe ständig Schuldgefühle.
1 Ich habe mich nicht über ein Gefühl beklagt.	2 Ich habe mich manchmal um Selbstmord, aber ich würde es nicht tun.	3 Ich würde mich verletzen, wenn ich die Gelegenheit dazu hätte.	4 Ich würde mich selbst verletzen, wenn ich könnte.

## 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

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

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  $\varepsilon_i$  erklärt.

Wie unten gezeigt ist dieses Modell äquivalent zum oben eingeführten Normalverteilungsmodell

$$v_1, \dots, v_n \sim N(\mu, \sigma^2). \quad (20)$$

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

- (1) Was sind sinnvolle Tipps für die wahren, aber unbekanntenen, Parameterwerte  $\mu$  und  $\sigma^2$ ?
- (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. mit  $i = 1, \dots, n$ , dass

$$v_i = f(\varepsilon_i) \text{ mit } f: \mathbb{R} \rightarrow \mathbb{R}, e_i \mapsto f(e_i) := e_i + \mu. \quad (21)$$

Dann gilt

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

also  $v_i \sim N(\mu, \sigma^2)$  u.i.v. und damit  $v_1, \dots, v_n \sim N(\mu, \sigma^2)$ .

---

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.

1. Siehe Definition (Frequentistischen Inferenzmodell)
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.