



Tutorium Allgemeines Lineares Modell

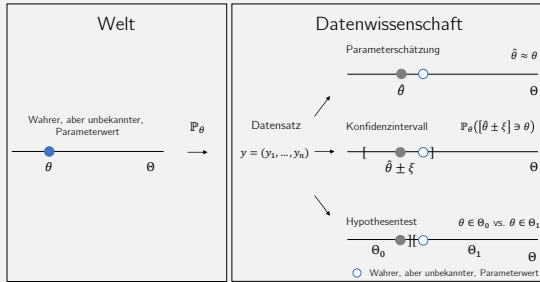
BSc Psychologie SoSe 2022

8. Termin (Teil 2): Modellevaluation_Teil_1

Belinda Fleischmann

Wiederholung - Frequentistisches Weltbild

Modell und Standardprobleme der Frequentistischen Inferenz



- Wir nehmen an, dass wahre, aber unbekannte Parameter existieren.
- Wir nehmen weiterhin an, dass probabilistische Prozesse existieren, die, gegeben dieser wahren, aber unbekannten Parameter, Datensätze generieren können.
- Für diese probabilistischen Prozesse gehen wir davon aus, dass ihnen bestimmte Verteilungen bzw. Wahrscheinlichkeitsdichten zugrundeliegen (z.B. Normalverteilung der Zufallsfehler)
- Wir verwenden erhobene Daten dafür, Parameterwerte zu schätzen (Wir berechnen eine "Einschätzung", was der wahre Wert sein könnte, den wir nicht beobachten können).
- Dabei bilden die angenommenen Verteilungen bzw. Wahrscheinlichkeitsdichten der probabilistischen Prozesse (die, wie wir annehmen, die Daten generiert haben), die Grundlage für Parameterschätzung und Modellevaluation.

Wiederholung - Standardannahmen frequentistischer Inferenz

- Gegeben sei ein statistisches Modell \mathbb{P}_θ , in dem probabilistische Prozesse definiert sind, die Datensätze generieren können.
- Es wird angenommen, dass ein vorliegender Datensatz eine der möglichen Realisierungen der Daten des Modells ist. Eine mögliche Realisierung wäre $y^{(1)}$, $y^{(2)}$ wäre eine andere, $y^{(3)}$ wäre nochmals eine andere.
- Aus frequentistischer Sicht kann man unendlich oft Datensätze basierend auf einem Modell generieren und zu jedem Datensatz Schätzer oder Statistiken auswerten, z.B. den Betaparameterschätzer (z.B. $\hat{\beta}^{(1)}$ für Datensatz $y^{(1)}$)

$$\text{Datensatz (1)} : y^{(1)} = \left(y_1^{(1)}, y_2^{(1)}, \dots, y_n^{(1)} \right)^T \text{ mit } \hat{\beta}^{(1)} = (X^T X)^{-1} X^T y^{(1)}$$

$$\text{Datensatz (2)} : y^{(2)} = \left(y_1^{(2)}, y_2^{(2)}, \dots, y_n^{(2)} \right)^T \text{ mit } \hat{\beta}^{(2)} = (X^T X)^{-1} X^T y^{(2)}$$

$$\text{Datensatz (3)} : y^{(3)} = \left(y_1^{(3)}, y_2^{(3)}, \dots, y_n^{(3)} \right)^T \text{ mit } \hat{\beta}^{(3)} = (X^T X)^{-1} X^T y^{(3)}$$

$$\text{Datensatz (4)} : y^{(4)} = \left(y_1^{(4)}, y_2^{(4)}, \dots, y_n^{(4)} \right)^T \text{ mit } \hat{\beta}^{(4)} = (X^T X)^{-1} X^T y^{(4)}$$

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

- Um die Qualität statistischer Methoden zu beurteilen betrachtet die frequentistische Statistik die Wahrscheinlichkeitsverteilungen von Schätzern und Statistiken unter Annahme der Datenverteilung (z.B. im ALM die Verteilung des Datenvektors $y = X\beta + \epsilon$ mit $\epsilon \sim N(0_n, \sigma^2 I_n)$, und damit $y \sim N(X\beta, \sigma^2 I_n)$, siehe Einheit (5) Theorem zu ALM Datenverteilung).
- Was ist zum Beispiel die Verteilung (möglicher) Betaparameterschätzer ($\hat{\beta}^{(1)}, \hat{\beta}^{(2)}, \hat{\beta}^{(3)}, \hat{\beta}^{(4)}, \dots$ also die Verteilung der Zufallsvariable $\hat{\beta} := (X^T X)^{-1} X^T y$?

1. Geben Sie das Theorem zur Frequentistischen Verteilung des Betaparameterschätzers wieder.
2. Geben Sie die Verteilung des Betaparameterschätzers im Szenario von n u.i.v. Zufallsvariablen wieder.
3. Geben Sie das Theorem zur Frequentistischen Verteilung des Varianzparameterschätzers wieder.
4. Geben Sie die Verteilung des skalierten Varianzparameterschätzers bei n u.i.v. Zufallsvariablen wieder.
5. Skizzieren Sie die WDFen von t -Zufallsvariablen mit 2, 10 und 30 Freiheitsgraden.
6. Skizzieren Sie die WDFen von nichtzentralen t -Zufallsvariablen mit Nichtzentralitätsparametern 0,5 und 15.
7. Geben Sie die Definition der T-Statistik wieder.
8. Erläutern Sie die Definition der T-Statistik.
9. Warum kann die T-Statistik als Signal-zu-Rauschen Verhältnis interpretiert werden.
10. Geben Sie die Form der T-Statistik im Szenario von n u.i.v. Zufallsvariablen wieder.
11. Erläutern Sie den Zusammenhang der T-Statistik und Cohen's d .
12. Geben Sie die Definition der T-Teststatistik wieder.

13. Erläutern Sie die Definition der T-Teststatistik.
14. Geben Sie das Theorem zur Verteilung der T-Teststatistik wieder.
15. Geben Sie die Definition eines vollständigen und eines reduzierten ALMs wieder.
16. Geben Sie die Definition des Likelihood-Quotienten eines vollständigen und eines reduzierten ALMs wieder.
17. Geben Sie das Theorem zu Likelihood-Quotient und Residualquadratsummendifferenz wieder.
18. Definieren Sie die F-Statistik.
19. Erläutern Sie den Zusammenhang zwischen F-Statistik und Likelihood-Quotient.
20. Definieren Sie und erläutern Sie den Zähler der F-Statistik.
21. Definieren Sie und erläutern Sie den Nenner der F-Statistik.
22. Erläutern Sie die F-Statistik.
23. Geben Sie das Theorem zur Verteilung der F-Statistik wieder.

1. Geben Sie das Theorem zur Frequentistischen Verteilung des Betaparameterschätzers wieder.

Theorem (Frequentistische Verteilung des Betaparameterschätzers)

Es sei

$$y = X\beta + \varepsilon \text{ mit } \varepsilon \sim N(0_n, \sigma^2 I_n) \quad (1)$$

das ALM in generativer Form. Weiterhin sei

$$\hat{\beta} = (X^T X)^{-1} X^T y \quad (2)$$

der Betaparameterschätzer. Dann gilt

$$\hat{\beta} \sim N(\beta, \sigma^2 (X^T X)^{-1}). \quad (3)$$

Anmerkungen:

- Da es sich bei $\hat{\beta}$ um eine linear-affine Transformation von y handelt, leitet sich die hier angegebene Verteilung für $\hat{\beta}$ aus dem Theorem zur linearen Transformation von multivariaten Normalverteilung aus Einheit (4) ab.
- Im Spezifischen besagt das Theorem in diesem Fall, dass wenn $y \sim N(X\beta, \sigma^2 I_n)$ und $\hat{\beta} = Ay + b$, dann gilt $\hat{\beta} \sim N(AX\beta + b, A\sigma^2 I_n A^T)$, wobei $A = (X^T X)^{-1} X^T$ der Faktor der linearen-affinen Transformation ist.

2. Geben Sie die Verteilung des Betaparameterschätzers im Szenario von n u.i.v. Zufallsvariablen wieder.

Das Szenario von n u.i.v. Zufallsvariablen ist in Matrixschreibweise gegeben durch

$$y \sim N(X\beta, \sigma^2 I_n) \text{ mit } X := \mathbf{1}_n \in \mathbb{R}^{n \times 1}, \beta := \mu \in \mathbb{R} \text{ und } \sigma^2 > 0,$$

wobei der Betaparameterschätzer gegeben ist durch $\hat{\beta} = \bar{y}$. Durch Einsetzen von $\beta := \mu$ und $X := \mathbf{1}_n$ in $\hat{\beta} \sim N(\beta, \sigma^2(X^T X)^{-1})$ (vgl. Theorem zur Frequentistischen Verteilung des Betaparameterschätzers), erhalten wir

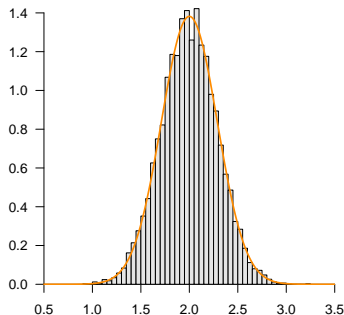
$$\bar{y} \sim N\left(\mu, \frac{\sigma^2}{n}\right).$$

Das Stichprobenmittel (\bar{y}) von n u.i.v. Zufallsvariablen mit Erwartungswertparameter μ und Varianzparameter σ^2 ist also normalverteilt mit Erwartungswertparameter μ und Varianzparameter σ^2/n .

2. Geben Sie die Verteilung des Betaparameterschätzers im Szenario von n u.i.v. Zufallsvariablen wieder.

Visualisierung von $N\left(\hat{\beta}; \beta, \frac{\sigma^2}{n}\right)$ für $y \sim N(X\beta, \sigma^2 I_n)$ mit $\beta = \mu = 2$ und $\sigma^2 = 1$. Mit anderen Worten, wie sieht die Verteilung der Zufallsvariable $\hat{\beta}$ aus, wenn wir aus der von uns angenommenen "wahren" Verteilung des Zufallsvektors $y \sim N(X\beta, \sigma^2 I_n)$ ganz viele Realisierungen generieren, (z.B. 10000 Realisierungen $y^{(1)}, \dots, y^{(10000)}$), und für jede dieser Datensätze einen Betaparameter schätzen ($\hat{\beta}^{(1)}, \dots, \hat{\beta}^{(10000)}$).

$$N\left(\hat{\beta}; \beta, \frac{\sigma^2}{n}\right)$$



3. Geben Sie das Theorem zur Frequentistischen Verteilung des Varianzparameterschätzers wieder.

Theorem (Frequentistische Verteilung des Varianzparameterschätzers)

Es sei

$$y = X\beta + \varepsilon \text{ mit } \varepsilon \sim N(0_n, \sigma^2 I_n) \quad (4)$$

das ALM in generative Form. Weiterhin sei

$$\hat{\sigma}^2 = \frac{(y - X\hat{\beta})^T (y - X\hat{\beta})}{n - p} \quad (5)$$

der Varianzparameterschätzer. Dann gilt

$$\frac{n - p}{\sigma^2} \hat{\sigma}^2 \sim \chi^2(n - p) \quad (6)$$

4. Geben Sie die Verteilung des skalierten Varianzparameterschätzers bei n u.i.v. Zufallsvariablen wieder.

Wie in Aufg. 2, haben wir im Szenario von n u.i.v. Zufallsvariablen

$$y \sim N(X\beta, \sigma^2 I_n) \text{ mit } X := 1_n \in \mathbb{R}^{n \times 1}, \beta := \mu \in \mathbb{R} \text{ und } \sigma^2 > 0,$$

wobei der Betaparameterschätzer gegeben ist durch $\hat{\beta} = \bar{y}$ und der Varianzparameterschätzer durch $\hat{\sigma}^2 = s^2$. Durch Einsetzen von $p = 1$ in die Verteilung des skalierten Varianzparameterschätzers ($\frac{n-p}{\sigma^2} \hat{\sigma}^2$), wie im Theorem zur Frequentistischen Verteilung des Varianzparameterschätzers gegeben, erhalten wir

$$\frac{n-1}{\sigma^2} \hat{\sigma}^2 \sim \chi^2(n-1)$$

Das ist identisch mit der in Einheit (11) Konfidenzintervalle von Wahrscheinlichkeitstheorie und Frequentistische Inferenz gelernten U-Statistik. Wdhl.: Für den Fall von n unabhängigen und identisch normalverteilten Zufallsvariablen ist die U-Statistik definiert als

$$U := \frac{n-1}{\sigma^2} s^2,$$

wobei

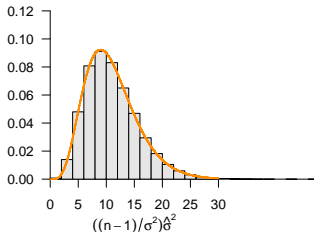
$$U \sim \chi^2(n-1).$$

Offenbar ist U für $p = 1$ mit der im obigen Theorem betrachteten Zufallsvariable $\frac{n-p}{\sigma^2} \hat{\sigma}^2$ identisch.

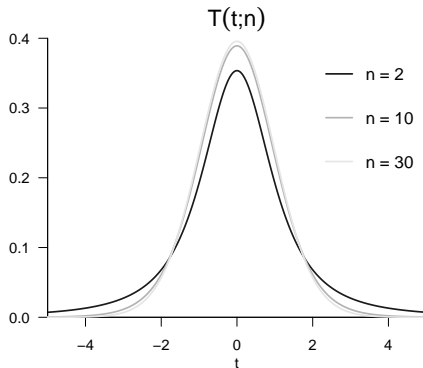
4. Geben Sie die Verteilung des skalierten Varianzparameterschätzers bei n u.i.v. Zufallsvariablen wieder.

Visualisierung von $\chi^2(\frac{n-1}{\sigma^2}\hat{\sigma}^2; n-1)$ für $y \sim N(X\beta, \sigma^2 I_n)$ mit $\beta = \mu = 2$ und $\sigma^2 = 1$. Mit anderen Worten, wie sieht die Verteilung der skalierten Zufallsvariable $\hat{\sigma}^2$, also $\frac{n-p}{\sigma^2}\hat{\sigma}^2$ aus, wenn wir aus der von uns angenommen "wahren" Verteilung des Zufallsvektors $y \sim N(X\beta, \sigma^2 I_n)$ ganz viele Realisierungen generieren, (z.B. 10000 Realisierungen $y^{(1)}, \dots, y^{(10000)}$), und für jede dieser Datensätze einen Varianzparameter schätzen ($\hat{\sigma}^{2(1)}, \dots, \hat{\sigma}^{2(10000)}$).

$$\chi^2(\frac{n-p}{\sigma^2}\hat{\sigma}^2; n-1)$$



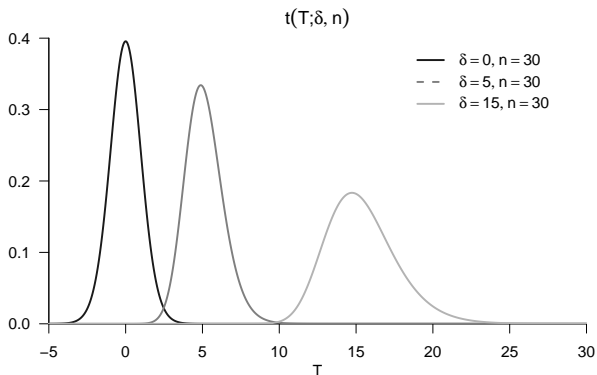
5. Skizzieren Sie die WDFen von t -Zufallsvariablen mit 2, 10 und 30 Freiheitsgraden.



Anmerkungen:

- Die Verteilung ist um 0 symmetrisch
- Steigendes n verschiebt Wahrscheinlichkeitsmasse aus den Ausläufen zum Zentrum

6. Skizzieren Sie die WDFen von nichtzentralen t -Zufallsvariablen mit Nichtzentralitätsparametern 0,5 und 15.



7. Geben Sie die Definition der T-Statistik wieder.

Definition (T-Statistik)

Es sei

$$y = X\beta + \varepsilon \text{ mit } \varepsilon \sim N(0_n, \sigma^2 I_n) \quad (7)$$

das ALM in generativer Form. Weiterhin seien

$$\hat{\beta} := (X^T X)^{-1} X^T y \text{ und } \hat{\sigma}^2 := \frac{(y - X\hat{\beta})^T (y - X\hat{\beta})}{n - p} \quad (8)$$

die Betaparameter- und Varianzparameterschätzer, respektive. Dann ist für einen *Kontrastgewichtsvektor* $c \in \mathbb{R}^p$ die *T-Statistik* definiert als

$$T := \frac{c^T \hat{\beta}}{\sqrt{\hat{\sigma}^2 c^T (X^T X)^{-1} c}}. \quad (9)$$

8. Erläutern Sie die Definition der T-Statistik.

- Die T-Statistik hängt durch die Parameterschätzer $\hat{\beta}$ und $\hat{\sigma}^2$ von den Daten y ab.
- Der Kontrastgewichtsvektor $c \in \mathbb{R}^p$ projiziert $\hat{\beta}$ auf einen Skalar $c^T \hat{\beta}$.
- Intuitiv quantifiziert die T-Statistik das Verhältnis von geschätzte Effektstärke zur geschätzten stichprobenumfangskalierten Datenvariabilität

$$T = \frac{\text{Geschätzte Effektstärke}}{\text{Geschätzte stichprobenumfangskalierte Datenvariabilität}}$$

und repräsentiert damit ein Signal-zu-Rauschen Verhältnis (signal-to-noise ratio).