

Vorlesung: Statistik II für Wirtschaftswissenschaft

Prof. Dr. Helmut Küchenhoff

Institut für Statistik, LMU München

Sommersemester 2017



- Einführung
- 1 Wahrscheinlichkeit: Definition und Interpretation
- 2 Elementare Wahrscheinlichkeitsrechnung
- 3 Zufallsgrößen
- 4 Spezielle Zufallsgrößen
- 5 Mehrdimensionale Zufallsvariablen
- 6 Genzwertsätze
- 7 Statistische Inferenz: Punktschätzer
- 8 Statistische Inferenz: Konfidenzintervalle
- 9 Statistische Inferenz: Statistische Tests
- 10 Spezielle statistische Tests
- 11 Lineare Regression
- 12 Bayes-Statistik

Subjektiver Wahrscheinlichkeitsbegriff

Laplace, Ramsey, de Finetti:

„Die Wahrscheinlichkeit eines Ereignisses ist der Grad der Überzeugung, mit der ein Beobachter aufgrund eines bestimmten Informationsstandes an das Eintreten eines Ereignisses glaubt“

$P(A)$ ist der Wetteinsatz in Euro, den eine Person höchstens einzugehen bereit ist, falls diese bei Eintreten von A einen Euro gewinnt.

Beispiel: Trifft Müller beim Elfmeter?

3 Theorien :

UH immer $p = 1.0$

AR $p = 0.8$

JL $p = 0.7$

Ansatz: Theorien habe gleiche Wahrscheinlichkeit

$$P(UH) = P(AR) = P(JL) = 1/3$$

Daten : Ein Treffer $X_1 = 1$

Berechnung der Wahrscheinlichkeit der Beobachtung unter den verschiedenen Theorien (Likelihood)

$$P(X_1 = 1|UH) = 1$$

$$P(X_1 = 1|AR) = 0.8$$

$$P(X_1 = 1|JL) = 0.7$$

ML-Prinzip: Theorie UH stimmt

Satz von Bayes

Bilden die A_j eine vollständige Zerlegung von Ω und ist B irgendein Ereignis, so gilt:

$$P(A_j|B) = \frac{P(B|A_j) \cdot P(A_j)}{\sum_{i=1}^n P(B|A_i) \cdot P(A_i)}.$$

In unserem Fall entsprechen die A_j den Theorien $A_1 = UH$ etc. und das Ereignis B entspricht den Daten $X_1 = 1$

$$P(UH) = P(AR) = P(JL) = 1/3$$

$$P(X_1 = 1|UH) = 1$$

$$P(X_1 = 1|AR) = 0.8$$

$$P(X_1 = 1|JL) = 0.7$$

Posteriori- Wahrscheinlichkeiten

Wir berechnen jetzt die Posteriori-Wahrscheinlichkeit der Theorien gegeben die Daten:

$$P(UH|X_1 = 1) = \frac{P(UH) \cdot P(X_1 = 1|UH)}{P(UH) \cdot P(X_1 = 1|UH) + P(AR) \cdot P(X_1 = 1|AR) + P(JL) \cdot P(X_1 = 1|JL)}$$

$$\begin{aligned}P(UH|X = 1) &= \frac{\frac{1}{3} \cdot 1}{\frac{1}{3} \cdot 1 + \frac{1}{3} \cdot 0.8 + \frac{1}{3} \cdot 0.7} \\ &= \frac{1}{1 + 0.8 + 0.7} = 0.4 \\ P(AR|X = 1) &= \frac{0.8}{1 + 0.8 + 0.7} = 0.32 \\ P(JL|X = 1) &= \frac{0.7}{1 + 0.8 + 0.7} = 0.28\end{aligned}$$

Leichte Verschiebung

Verwendung weiterer Daten

13 von 15 (TM)

Berechnung der Likelihood mit Hilfe der Binomialverteilung(i.i.d)

$$P(X_2 = 13|UH) = 0$$

$$P(X_2 = 13|AR) = 0.23$$

$$P(X_2 = 13|JL) = 0.09$$

Wir berechnen jetzt die Posteriori-Wahrscheinlichkeit der Theorien gegeben die Daten:

$$P(UH|X_2 = 13) = \frac{0}{0 + 0.23 + 0.09} = 0$$

$$P(AR|X_2 = 13) = \frac{0.23}{0 + 0.23 + 0.09} = 0.72$$

$$P(JL|X_2 = 13) = \frac{0.09}{0 + 0.23 + 0.09} = 0.28$$

Es spricht Einiges für AR



50 von 62 (GM)

Berechnung der Likelihood mit Hilfe der Binomialverteilungsannahme (i.i.d)

$$P(X_3 = 50|UH) = 0$$

$$P(X_3 = 50|AR) = 0.126$$

$$P(X_3 = 50|JL) = 0.021$$

Wir berechnen jetzt die Posteriori-Wahrscheinlichkeit der Theorien gegeben die Daten:

$$P(UH|X_2 = 13) = \frac{0}{0 + 0.126 + 0.021} = 0$$

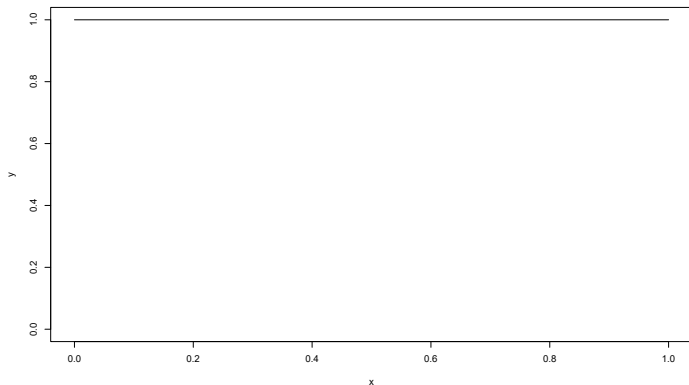
$$P(AR|X_2 = 13) = \frac{0.126}{0 + 0.126 + 0.021} = 0.86$$

$$P(JL|X_2 = 13) = \frac{0.021}{0 + 0.126 + 0.021} = 0.14$$

Es spricht Einiges für AR

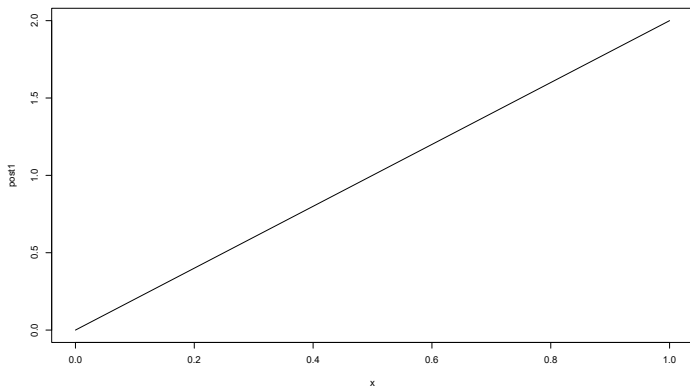
Inferenz über unbekanntem Parameter

Parameter θ unbekannt
Apriori Gleichverteilung



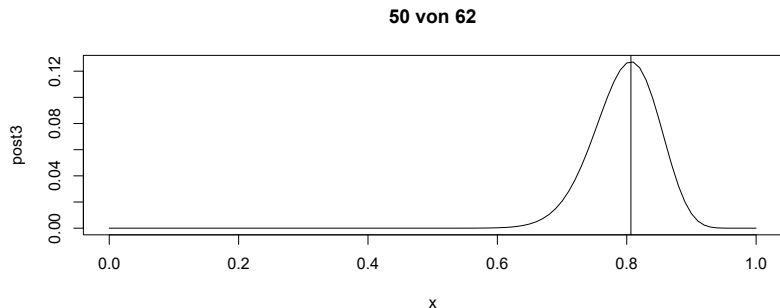
Berechnung der Posteriori-Verteilung für einen Treffer

Inferenz über unbekanntem Parameter



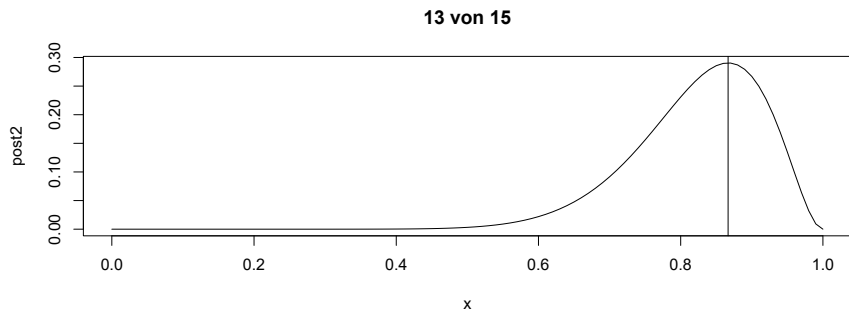
50 von 62 (GM)

Berechnung der Likelihood und der Posteriori-Verteilung mit Hilfe der Binomialverteilungsannahme (i.i.d)



13 von 15 (TM)

Berechnung der Likelihood und der Posteriori-Verteilung mit Hilfe der Binomialverteilungsannahme (i.i.d)



Berechnung der Posteriori-Verteilung

Inferenz für Parameter θ bei Beobachtung x

- $f(\theta)$ Priori - Verteilung von θ
- $f(x|\theta)$ Wahrscheinlichkeitsfunktion/Dichte von x bei Parameter θ
- $f(x)$ a priori Randverteilung von x
- $f(\theta|x)$ Posteriori - Verteilung von θ gegeben die Beobachtung x

Berechnung der Posteriori-Dichte

$$f(\theta|x) = \frac{f(x|\theta)f(\theta)}{f(x)} = \frac{f(x|\theta)f(\theta)}{\int f(x|\theta)f(\theta)d\theta}$$

Allgemeine Form

Bei Beobachtungen x_1, \dots, x_n wird die gemeinsame Dichte betrachtet. Man erhält für unabhängige Beobachtungen:

$$f(x_1, \dots, x_n | \theta) = f(x_1 | \theta) \cdot f(x_2 | \theta) \cdots f(x_n | \theta) = L(\theta)$$

$L(\theta)$ ist die Likelihoodfunktion



Bayes-Inferenz

Die Wahrscheinlichkeits- oder Dichtefunktion von X gegeben der Parameter θ ist gegeben durch

$$f(x|\theta)$$

Die Likelihood ist

$$L(\theta) = f(x_1, \dots, x_n|\theta)$$

Für den unbekanntem Parameter θ ist die Priori-Dichte gegeben

$$f(\theta)$$

Dann gilt für die Posteriori- Dichte von θ

$$f(\theta|x_1, \dots, x_n) = \frac{f(x_1, \dots, x_n|\theta)f(\theta)}{f(x)} = \frac{L(\theta)f(\theta)}{\int L(\theta)f(\theta)d\theta}$$

- Die Posteriori-Verteilung $f(\theta|x_1, \dots, x_n)$ enthält die gesamte Information der Daten über den Parameter θ
- Die Posteriori-Verteilung hat folgende Darstellung :

$$f(\theta|x_1, \dots, x_n) = cL(\theta)f(\theta)$$

Dabei kann c als von θ unabhängiger Normierungsfaktor angesehen werden

- Ein zentrales Problem bei der Bayes-Inferenz ist die Wahl der Priori-Verteilung. Man wählt häufig sog. nicht-informativen Priori-Verteilungen

$$X \sim \text{Bin}(n, \theta)$$

$$f(x|\theta) = \binom{n}{x} \theta^x \cdot (1 - \theta)^{n-x}$$

Priori - Dichte

$$f(\theta) = 1 \text{ für } 0 \leq \theta \leq 1$$

Posteriori

$$f(\theta|x) = c \binom{n}{x} \theta^x \cdot (1 - \theta)^{n-x}$$

Die Inferenz erfolgt mit der Posteriori-Verteilung f_{POST} Punktschätzung von θ durch

- Posteriori - Modus, d.h. Maximum von f_{POST}
- Posteriori - Erwartungswert, d.h. Erwartungswert von θ unter $f_{POST}(\theta)$

Strategie: Finde Intervall, in dem der Parameter mit Wahrscheinlichkeit γ liegt

$$P(\theta \in [\theta_u, \theta_o]) = \gamma$$

Bezeichnung: Kreditabilitätsintervalle

Beispiel: Normalverteilung (Fahrmeir et al.)

Beispiel übernommen aus:

L. Fahrmeir, Ch. Heumann, R. Künstler, I. Pigeot und G. Tutz: Statistik - Der Weg zur Datenanalyse, (8. Auflage), Springer-Verlag, 2016.

Seien X_1, \dots, X_n unabhängige Wiederholungen von $X \sim N(\mu, \sigma^2)$, wobei μ zu schätzen ist, aber nun σ^2 als bekannt angenommen wird. Als a priori Dichte für μ wählen wir eine $N(\mu_0, \sigma_0^2)$ -Verteilung, also

$$f(\mu) = \frac{1}{(2\pi\sigma_0^2)^{1/2}} \exp\left\{-\frac{(\mu - \mu_0)^2}{2\sigma_0^2}\right\}.$$

Die a posteriori Dichte ergibt sich also aus:

$$f(\mu \mid x_1, \dots, x_n) = \frac{L(\mu, \sigma)f(\mu)}{\int L(\mu, \sigma)f(\mu) d\mu}$$

Beispiel: Normalverteilung (Fortsetzung)

Es ergibt sich eine Normalverteilung

$$\mu \mid x_1, \dots, x_n \sim N(\tilde{\mu}, \tilde{\sigma}^2)$$

mit a posteriori Erwartungswert

$$\tilde{\mu} = \frac{n\sigma_0^2}{n\sigma_0^2 + \sigma^2} \bar{x} + \frac{\sigma^2}{n\sigma_0^2 + \sigma^2} \mu_0$$

und a posteriori Varianz

$$\tilde{\sigma}^2 = \frac{\sigma^2}{n + \sigma^2/\sigma_0^2}.$$

Beispiel: Normalverteilung (Zusammenfassung)

- Für $\sigma_0^2 \rightarrow 0$ (“exaktes Vorwissen”) gilt $\tilde{\mu} \rightarrow \mu_0$
- Für $\sigma_0^2 \rightarrow \infty$ (“kein Vorwissen”) ergibt sich $\tilde{\mu} \rightarrow \bar{x}$, also die Maximum Likelihood-Schätzung $\hat{\mu} = \bar{x}$ aus der Stichprobe
- Analog für die Varianz

Also:

$$\mu \mid x_1, \dots, x_n \rightarrow N(\bar{x}, \sigma^2/n) \quad \text{für } \sigma_0^2 \rightarrow \infty$$

bei nichtvorhandenem Vorwissen über μ , und

$$\mu \mid x_1, \dots, x_n \rightarrow N(\mu_0, 0)$$

bei sicherem Vorwissen $\mu = \mu_0$.

Der “Hyperparameter” σ_0^2 steuert also den Kompromiss zwischen Stichprobeninformation und subjektiver a priori Information.

Vergleich: Frequentistisch vs. Bayes–Inferenz

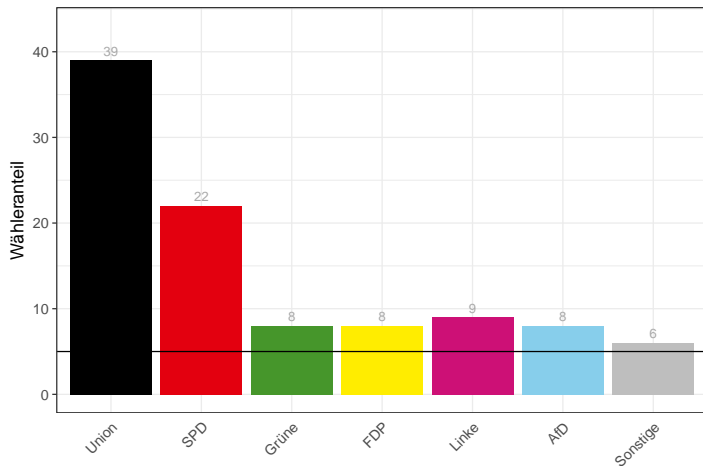
- Da die Likelihood wesentlich in die Berechnung eingeht gibt es teilweise ähnliche Ergebnisse, z.B. Kreditabilitätsintervalle mit nicht informativer Prior sind in bestimmten Fällen identisch mit Konfidenzintervallen
- Unterschiedliche Interpretation
- Laufende wissenschaftliche Diskussion



Auswertung von Umfragen: Wahlistik

Verwende Umfrageergebnisse, z.B.

Aktuelle forsa-Umfrage



Bayes Schätzung von Wahrscheinlichkeiten

Unionsgeführte Große Koalition

>99%

Union-Grüne-FDP

58.6%

Union-FDP

33.2%

Union-Grüne

23.1%

SPD-geführte Große Koalition

<1%

SPD-Grüne-Linke

<1%

SPD-Grüne-FDP

<1%

SPD-Grüne

<1%

SPD-Linke

<1%

SPD-FDP

<1%