Tutorium

Wahrscheinlichketstheorie und Frequentistische Inferenz

BSc Psychologie WiSe 2022/23

Belinda Fleischmann

(7) Ungleichungen und Grenzwerte

Selbstkontrollfragen

1. Geben Sie die Markov Ungleichung wieder.
2. Geben Sie die Chebyshev Ungleichung wieder.
3. Geben Sie die Cauchy-Schwarz Ungleichung wieder.
4. Geben Sie die Korrelationsungleichung wieder.
5. Definieren Sie den Begriff der Konvergenz in Wahrscheinlichkeit.
6. Definieren Sie den Begriff der Konvergenz in Verteilung.
7. Geben Sie das Schwache Gesetz der Großen Zahl wieder.
8. Erläutern Sie den Zentralen Grenzwertsatz nach Lindenberg und Lévy.
9. Erläutern Sie den Zentralen Grenzwertsatz nach Liapunov.
10. Warum sind die Zentralen Grenzwertsätze für die statistische Modellbildung wichtig?

1. Geben Sie die Markov Ungleichung wieder.

Theorem (Markov Ungleichung)

ξ sei eine Zufallsvariable mit $\mathbb{P}(\xi \geq 0) = 1$. Dann gilt für alle $x \in \mathbb{R}$, dass

$$\mathbb{P}(\xi \geq x) \leq \frac{\mathbb{E}(\xi)}{x}. \tag{1}$$

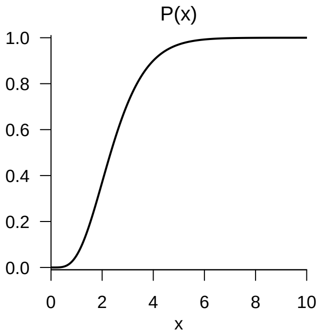
Bemerkungen

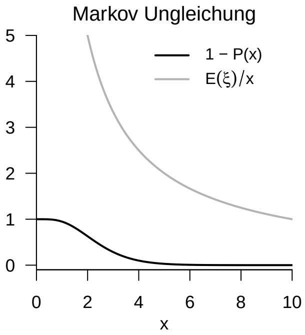
- Weil $\mathbb{P}(\xi \geq 0) = 1$ gilt, sagt man auch, dass ξ eine *nicht-negative* Zufallvariable ist.
- Die Ungleichung setzt Überschreitungswahrscheinlichkeiten und Erwartungswerte in Bezug.
- Gilt z.B. für eine nichtnegative Zufallsvariable ξ, dass $\mathbb{E}(\xi) = 1$, dann ist $\mathbb{P}(\xi \geq 100) \leq 0.01$.

Beispiel

Beispiel ($\xi \sim G(\alpha, \beta)$)

- Wir halten ohne Beweis fest, dass für $\xi \sim G(\alpha, \beta)$ gilt, dass $\mathbb{E}(\xi) = \alpha\beta$.
- Wir betrachten den Fall $\alpha := 5, \beta := 2$, so dass $G(x; 5, 2) = \chi^2(10)$





2. Geben Sie die Chebyshev Ungleichung wieder.

Theorem (Chebyshev Ungleichung)

Es sei ξ eine Zufallsvariable mit Varianz $\mathbb{V}(\xi)$. Dann gilt für alle $x \in \mathbb{R}$

$$\mathbb{P}(|\xi - \mathbb{E}(\xi)| \geq x) \leq \frac{\mathbb{V}(\xi)}{x^2}. \tag{2}$$

Bemerkungen

- Die Chebyshev Ungleichung setzt Abweichungen vom Erwartungswert in Bezug zur Varianz.
- Zum Beispiel gilt

$$\mathbb{P}\left(|\xi - \mathbb{E}(\xi)| \geq 3\sqrt{\mathbb{V}(\xi)}\right) \leq \frac{\mathbb{V}(\xi)}{\left(3\sqrt{\mathbb{V}(\xi)}\right)^2} = \frac{1}{9}. \tag{3}$$

3. Geben Sie die Cauchy-Schwarz Ungleichung wieder.

Theorem (Cauchy-Schwarz-Ungleichung)

ξ und υ seien zwei Zufallsvariablen und $\mathbb{E}(\xi\upsilon)$ sei endlich. Dann gilt

$$\mathbb{E}(\xi\upsilon)^2 \leq \mathbb{E}\left(\xi^2\right)\mathbb{E}\left(\upsilon^2\right). \tag{4}$$

Bemerkungen

- Analog gilt für Vektoren $x, y \in \mathbb{R}^n$, dass $\langle x, y\rangle^2 \leq \|x\| \cdot \|y\|$.
- Die Korrelationsungleichung ist eine direkte Konsequenz der Cauchy-Schwarz-Ungleichung.

4. Geben Sie die Korrelationsungleichung wieder.

Theorem (Korrelationsungleichung)

ξ und υ seien Zufallsvariablen mit $\mathbb{V}(\xi), \mathbb{V}(\upsilon) > 0$. Dann gilt

$$\rho(\xi, \upsilon)^2 = \frac{\mathbb{C}(\xi, \upsilon)^2}{\mathbb{V}(\xi)\mathbb{V}(\upsilon)} \leq 1. \tag{5}$$

Bemerkung

- Es gilt also

$$\rho(\xi, \upsilon)^2 \leq 1 \Leftrightarrow |\rho(\xi, \upsilon)| \leq 1 \Leftrightarrow \rho(\xi, \upsilon) \in [-1, 1]. \tag{6}$$

5. Definieren Sie den Begriff der Konvergenz in Wahrscheinlichkeit.

Definition (Konvergenz in Wahrscheinlichkeit)

Eine Folge von Zufallsvariable $\xi_1, \xi_2, \ldots$ *konvergiert gegen eine Zufallsvariable ξ in Wahrscheinlichkeit*, wenn für jedes noch so kleine $\epsilon > 0$ gilt, dass

$$\lim_{n\to\infty} \mathbb{P}(|\xi_n - \xi| < \epsilon) = 1 \Leftrightarrow \lim_{n\to\infty} \mathbb{P}(|\xi_n - \xi| \geq \epsilon) = 0 \tag{7}$$

Die Konvergenz von $\xi_1, \xi_2, \ldots.$ gegen ξ in Wahrscheinlichkeit wird geschrieben als

$$\xi_n \xrightarrow[n\to\infty]{\text{P}} \xi \tag{8}$$

Bemerkungen

- $\xi_n \xrightarrow[n\to\infty]{\text{P}} \xi$ heißt, dass sich die Wahrscheinlichkeit, dass ξ_n in dem zufälligen Intervall

  $$]\xi - \epsilon, \xi + \epsilon[\tag{9}$$

  liegt, unabhängig davon, wie klein dieses Intervall sein mag, 1 nähert, wenn n gegen Unendlich strebt.
- Intuitiv heißt das, dass sich für eine konstante Zufallsvariable $\xi := a$ die Verteilung von ξ_n mehr und mehr um a konzentriert, wenn n gegen Unendlich strebt.

6. Definieren Sie den Begriff der Konvergenz in Verteilung.

Definition (Konvergenz in Verteilung)

Eine Folge $\xi_1, \xi_2, \ldots$ von Zufallsvariablen *konvergiert in Verteilung gegen eine Zufallsvariable* ξ, wenn

$$\lim_{n \to \infty} P_{\xi_n}(x) = P_{\xi}(x). \tag{10}$$

für alle ξ an denen P_ξ stetig ist. Die Konvergenz in Verteilung von $\xi_1, \xi_2, \ldots$ gegen ξ wird geschrieben als

$$\xi_n \xrightarrow[n \to \infty]{\text{D}} \xi, \tag{11}$$

Gilt $\xi_n \xrightarrow[n \to \infty]{\text{D}} \xi$, dann heißt die Verteilung von ξ die *asymptotische Verteilung der Folge* $\xi_1, \xi_2, \ldots$.

Bemerkungen

- $\xi \xrightarrow[n \to \infty]{\text{D}} \xi$ ist eine Aussage über die Konvergenz von KVFs.
- Konvergenz in Wahrscheinlichkeit impliziert Konvergenz in Verteilung.

7. Geben Sie das Schwache Gesetz der Großen Zahl wieder.

Theorem (Schwaches Gesetz der Großen Zahl)

$\xi_1, \ldots, \xi_n$ seien unabhängig und gleichverteilte Zufallsvariablen mit $\mathbb{E}(\xi_i) = \mu$ für alle $i = 1, \ldots, n$. Weiterhin bezeichne

$$\bar{\xi}_n := \frac{1}{n} \sum_{i=1}^{n} \xi_i \tag{12}$$

das Stichprobenmittel der ξ_i, $i = 1, \ldots, n$. Dann konvergiert $\bar{\xi}_n$ in Wahrscheinlichkeit gegen μ,

$$\bar{\xi}_n \xrightarrow[n \to \infty]{\mathsf{P}} \mu. \tag{13}$$

Bemerkungen

- $\bar{\xi}_n \xrightarrow[n \to \infty]{\mathsf{P}} \mu$ heißt, dass die Wahrscheinlichkeit, dass das Stichprobenmittel nahe dem Erwartungswert der zugrundeliegenden Verteilung liegt, sich 1 nähert, wenn $n \to \infty$.

8. Erläutern Sie den Zentralen Grenzwertsatz nach Lindenberg und Lévy.

Erläuterung:

Der zentrale Grenzwertsatz nach Lindenberg und Lévy besagt, dass wenn wir n unabhängig und identisch verteilten ZVen mit gleichem Erwartungswert $\mathbb{E}(\xi_i) = \mu$ und Varianz $\mathbb{V}(\xi_i) := \sigma^2 > 0$ haben, dann entspricht die KVF $P_{\zeta_n}(z)$ einer ZV ζ_n welche definiert ist als

$$\zeta_n := \sqrt{n}\left(\frac{\bar{\xi}_n - \mu}{\sigma}\right)$$

bei $n \to \infty$ der KVF der Standardnormalverteilung, welche wir mit Φ bezeichnen.

Theorem (Zentraler Grenzwertsatz nach Lindenberg und Lévy)

$\xi_1, \ldots, \xi_n$ seien unabhängig und identisch verteilte Zufallsvariablen mit

$$\mathbb{E}(\xi_i) := \mu \text{ und } \mathbb{V}(\xi_i) := \sigma^2 > 0 \text{ für alle } i = 1, \ldots, n. \tag{14}$$

Weiterhin sei ζ_n die Zufallsvariable definiert als

$$\zeta_n := \sqrt{n}\left(\frac{\bar{\xi}_n - \mu}{\sigma}\right). \tag{15}$$

Dann gilt für alle $z \in \mathbb{R}$

$$\lim_{n \to \infty} P_{\zeta_n}(z) = \Phi(z), \tag{16}$$

wobei Φ die kumulative Verteilungsfunktion der Standardnormalverteilung bezeichnet.

Bemerkung

- Wir zeigen später, dass damit für $n \to \infty$ asymptotisch auch gilt, dass

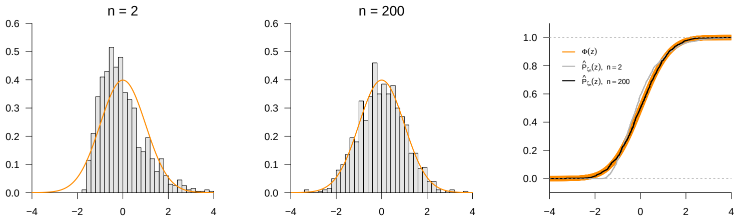
$$\sum_{i=1}^{n} \xi_i \sim N(n\mu, n\sigma^2) \text{ und } \bar{\xi}_n \sim N\left(\mu, \frac{\sigma^2}{n}\right). \tag{17}$$

Beispiel ($\xi_1, \ldots, \xi_n \sim \chi^2(k)$)

- Wir halten ohne Beweis fest, dass $\mathbb{E}(\xi_i) = k$ und $\mathbb{V}(\xi_i) = 2k$.
- Wir betrachten das Szenario $\xi_i \sim \chi(3)$ für $i = 1, \ldots, n$.
- Die linken Abbildungen zeigen Histogrammschätzer der Wahrscheinlichkeitsdichte von

$$\zeta_n := \sqrt{n}\left(\frac{\bar{\xi}_n - \mu}{\sigma}\right) \tag{18}$$

  basierend auf 1000 Realisationen von ζ_n für $n = 2$ und $n = 200$, sowie die WDF von $N(0, 1)$.
- Die rechte Abbildung zeigt die entsprechenden (empirischen) kumulativen Verteilungsfunktionen.



9. Erläutern Sie den Zentralen Grenzwertsatz nach Liapunov.

Erläuterung:

Der zentrale Grenzwertsatz nach Liapunov besagt, dass wenn wir n unabhängig aber nicht notwendigerweise identisch verteilten ZVen mit jeweiligen Erwartungswerten $\mathbb{E}(\xi_i) = \mu_i$ und Varianzen $\mathbb{V}(\xi_i) = \sigma_i^2$ haben, dann entspricht die KVF $P_{\zeta_n}(z)$ einer ZV ζ_n welche definiert ist als

$$\zeta_n := \frac{\sum_{i=1}^n \xi_i - \sum_{i=1}^n \mu_i}{\sqrt{\sum_{i=1}^n \sigma_i^2}}$$

bei $n \to \infty$ der KVF der Standardnormalverteilung Φ, wenn für $\xi_1, \ldots, \xi_n$ die folgenden zwei Eigenschaften gelten:

$$\mathbb{E}(|\xi_i - \mu_i|^3) < \infty \text{ und } \lim_{n\to\infty} \frac{\sum_{i=1}^n \mathbb{E}\left(|\xi_i - \mu_i|^3\right)}{(\sum_{i=1}^n \sigma_i^2)^{3/2}} = 0.$$

Theorem (Zentraler Grenzwertsatz nach Liapounov)

$\xi_1, \ldots, \xi_n$ seien unabhängige aber nicht notwendigerweise identisch verteilten Zufallsvariablen mit

$$\mathbb{E}(\xi_i) := \mu_i \text{ und } \mathbb{V}(\xi_i) := \sigma_i^2 > 0 \text{ für alle } i = 1, \ldots., n. \tag{19}$$

Weiterhin sollen für $\xi_1, \ldots, \xi_n$ folgend Eigenschaften gelten:

$$\mathbb{E}(|\xi_i - \mu_i|^3) < \infty \text{ und } \lim_{n \to \infty} \frac{\sum_{i=1}^n \mathbb{E}\left(|\xi_i - \mu_i|^3\right)}{(\sum_{i=1}^n \sigma_i^2)^{3/2}} = 0. \tag{20}$$

Dann gilt für die Zufallsvariable ζ_n definiert als

$$\zeta_n := \frac{\sum_{i=1}^n \xi_i - \sum_{i=1}^n \mu_i}{\sqrt{\sum_{i=1}^n \sigma_i^2}}, \tag{21}$$

für alle $z \in \mathbb{R}$, dass

$$\lim_{n \to \infty} P_{\zeta_n}(z) = \Phi(z), \tag{22}$$

wobei Φ KVF der Standardnormalverteilung bezeichnet.

Bemerkungen

- Wir zeigen später, dass dann auch gilt, dass $\sum_{i=1}^n \xi_i \sim N\left(\sum_{i=1}^n \mu_i, \sum_{i=1}^n \sigma_i^2\right)$.

10. Warum sind die Zentralen Grenzwertsätze für die statistische Modellbildung wichtig?

- Die Zentralen Grenzwertsätze besagen, dass die Summe von unabhängigen Zufallsvariablen mit Erwartungswert 0 asymptotisch, d.h. für unendlich viele Zufallsvariablen, normalverteilt mit Erwartungswertparameter 0 ist.
- Modelliert man eine Messgröße y also als Summe eines deterministischen Einflusses μ und der Summe

$$\varepsilon := \sum_{i=1}^{n} \xi_i \tag{23}$$

einer Vielzahl von unabhängigen Zufallsvariablen $\xi_i, i = 1, \ldots, n$, welche unbekannte Störeinflüsse beschreiben, so ist für großes n die Annahme

$$y = \mu + \varepsilon \text{ mit } \varepsilon \sim N(0, \sigma^2) \tag{24}$$

also mathematisch gerechtfertigt. Wie wir später sehen werden, liegt die Annahme in Gleichung (24) vielen statischen Modellen zugrunde.