Wahrscheinlichkeitstheorie und Frequentistische Inferenz

BSc Psychologie WiSe 2021/22

Prof. Dr. Dirk Ostwald

(7) Ungleichungen und Grenzwerte

Wahrscheinlichkeitsungleichungen

Erwartungswertungleichungen

Gesetze der Großen Zahl

Zentrale Grenzwertsätze

Selbstkontrollfragen

Wahrscheinlichkeitsungleichungen

Erwartungswertungleichungen

Gesetze der Großen Zahl

Zentrale Grenzwertsätze

Selbstkontrollfragen

Theorem (Markov Ungleichung)

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

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

Bemerkungen

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

Wahrscheinlichkeitsungleichungen

Beweis

Wir betrachten den Fall einer kontinuierlichen Zufallsvariable X mit WDF p. Wir halten zunächst fest, dass

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} \xi\, p(\xi)\, d\xi = \int_{0}^{\infty} \xi\, p(\xi)\, d\xi = \int_{0}^{x} \xi\, p(\xi)\, d\xi + \int_{x}^{\infty} \xi\, p(\xi)\, d\xi, \tag{2}$$

weil X nicht-negativ ist. Es folgt dann

$$\mathbb{E}(X) \ge \int_{x}^{\infty} \xi\, p(\xi)\, d\xi \ge \int_{x}^{\infty} x\, p(\xi)\, d\xi = x \int_{x}^{\infty} p(\xi)\, d\xi = x\, \mathbb{P}(X \ge x). \tag{3}$$

Dabei gilt die erste Ungleichung weil

$$\int_{-\infty}^{x} \xi\, p(\xi)\, d\xi = 0 \text{ für } x \le 0 \text{ und } \int_{-\infty}^{x} \xi\, p(\xi)\, d\xi > 0 \text{ für } x > 0. \tag{4}$$

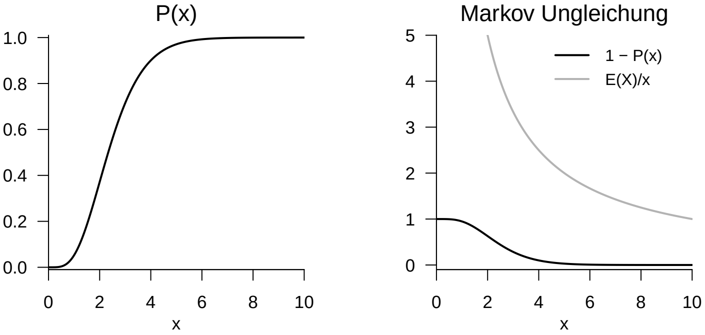
und die zweite Ungleichung gilt, weil $x \le \xi$ für $\xi \in [x, \infty[$. Es folgt also, dass

$$\mathbb{E}(X) \ge x\, \mathbb{P}(X \ge x) \Leftrightarrow \mathbb{P}(X \ge x) \le \frac{\mathbb{E}(X)}{x}. \tag{5}$$

$\square$

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

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



Theorem (Chebyshev Ungleichung)

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

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq x) \leq \frac{\mathbb{V}(X)}{x^2}. \quad (6)$$

Bemerkungen

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

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

Wahrscheinlichkeitsungleichungen

Beweis

Wir halten zunächst fest, dass für $a, b \in \mathbb{R}$ gilt

$$a^2 \geq b^2 \Leftrightarrow |a| \geq b. \tag{7}$$

Dazu betrachten wir die folgenden vier möglichen Fälle

$$\begin{array}{lclclr}
|a| \geq b \text{ für } a > 0, b > 0 & \Leftrightarrow & a \geq b & \Leftrightarrow & & a^2 \geq b^2 \\
|a| \geq b \text{ für } a > 0, b < 0 & \Leftrightarrow & a \geq b & \Leftrightarrow & & a^2 \geq b^2 \\
|a| \geq b \text{ für } a < 0, b > 0 & \Leftrightarrow & -a \geq b & \Leftrightarrow & (-a)^2 \geq b^2 & = a^2 \geq b^2 \\
|a| \geq b \text{ für } a < 0, b < 0 & \Leftrightarrow & -a \geq b & \Leftrightarrow & (-a)^2 \geq b^2 & = a^2 \geq b^2
\end{array}$$

Als nächstes definieren wir $Y := (X - \mathbb{E}(X))^2$. Dann folgt aus der Markov Ungleichung

$$\begin{aligned}
\mathbb{P}\left(Y \geq x^2\right) &\leq \frac{\mathbb{E}(Y)}{x^2} \\
\Leftrightarrow \mathbb{P}\left((X - \mathbb{E}(X))^2 \geq x^2\right) &\leq \frac{\mathbb{E}\left((X - \mathbb{E}(X))^2\right)}{x^2} \\
\Leftrightarrow \mathbb{P}(|X - \mathbb{E}(X)| \geq x) &\leq \frac{\mathbb{V}(X)}{x^2}.
\end{aligned} \tag{8}$$

$\square$

Wahrscheinlichkeitsungleichungen

Erwartungswertungleichungen

Gesetze der Großen Zahl

Zentrale Grenzwertsätze

Selbstkontrollfragen

Theorem (Cauchy-Schwarz-Ungleichung)

Es seien X, Y zwei Zufallsvariablen und $\mathbb{E}(XY)$ endlich. Dann gilt

$$\mathbb{E}(XY)^2 \le \mathbb{E}\left(X^2\right)\mathbb{E}\left(Y^2\right) \tag{9}$$

Bemerkungen

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

Erwartungswertungleichungen

Beweis

Wir betrachten den Fall, dass $0 < \mathbb{E}(X^2) < \infty$ und $0 < \mathbb{E}(Y^2) < \infty$. Für alle $a, b \in \mathbb{R}$ gilt dann

$$0 \le \mathbb{E}\left((aX + bY)^2\right) \text{ und } 0 \le \mathbb{E}\left((aX - bY)^2\right). \tag{10}$$

Also folgt

$$0 \le a^2\mathbb{E}(X^2) + b^2\mathbb{E}(Y^2) + 2ab\mathbb{E}(XY) \text{ und } 0 \le a^2\mathbb{E}(X^2) + b^2\mathbb{E}(Y^2) - 2ab\mathbb{E}(XY). \tag{11}$$

Definition von $a := \sqrt{\mathbb{E}(Y^2)}$ und $b := \sqrt{\mathbb{E}(X^2)}$ ergibt dann

$$\begin{aligned}
0 &\le \mathbb{E}(Y^2)\mathbb{E}(X^2) + \mathbb{E}(X^2)\mathbb{E}(Y^2) + 2\sqrt{\mathbb{E}(Y^2)}\sqrt{\mathbb{E}(X^2)}\mathbb{E}(XY) \\
\Leftrightarrow -2\sqrt{\mathbb{E}(Y^2)}\sqrt{\mathbb{E}(X^2)}\mathbb{E}(XY) &\le 2\mathbb{E}(X^2)\mathbb{E}(Y^2) \\
\Leftrightarrow -\sqrt{\mathbb{E}(Y^2)\mathbb{E}(X^2)}\mathbb{E}(XY) &\le \sqrt{\mathbb{E}(Y^2)\mathbb{E}(X^2)}\sqrt{\mathbb{E}(Y^2)\mathbb{E}(X^2)} \\
\Leftrightarrow -\mathbb{E}(XY) &\le \sqrt{\mathbb{E}(Y^2)\mathbb{E}(X^2)},
\end{aligned} \tag{12}$$

und analog

$$\begin{aligned}
0 &\le \mathbb{E}(Y^2)\mathbb{E}(X^2) + \mathbb{E}(X^2)\mathbb{E}(Y^2) - 2\sqrt{\mathbb{E}(Y^2)}\sqrt{\mathbb{E}(X^2)}\mathbb{E}(XY) \\
\Leftrightarrow 2\sqrt{\mathbb{E}(Y^2)}\sqrt{\mathbb{E}(X^2)}\mathbb{E}(XY) &\le 2\mathbb{E}(X^2)\mathbb{E}(Y^2) \\
\Leftrightarrow \sqrt{\mathbb{E}(Y^2)\mathbb{E}(X^2)}\mathbb{E}(XY) &\le \sqrt{\mathbb{E}(Y^2)\mathbb{E}(X^2)}\sqrt{\mathbb{E}(Y^2)\mathbb{E}(X^2)} \\
\Leftrightarrow \mathbb{E}(XY) &\le \sqrt{\mathbb{E}(Y^2)\mathbb{E}(X^2)}.
\end{aligned} \tag{13}$$

Zusammengenommen ergibt sich für den betrachteten Fall also die Cauchy-Schwarz-Ungleichung. Ein vollständiger Beweis findet sich bei DeGroot and Schervish (2012), Theorem 4.6.2.

□

Theorem (Korrelationsungleichung)

X und Y seien Zufallsvariablen mit $\mathbb{V}(X), \mathbb{V}(Y) > 0$. Dann gilt

$$\rho(X,Y)^2 = \frac{\mathbb{C}(X,Y)^2}{\mathbb{V}(X)\mathbb{V}(Y)} \le 1. \tag{14}$$

Bemerkung

- Es gilt also $\rho(X,Y)^2 \le 1 \Leftrightarrow |\rho(X,Y)| \le 1 \Leftrightarrow \rho(X,Y) \in [-1,1]$.

Beweis

Mit der Cauchy-Schwarz-Ungleichung für zwei Zufallsvariablen U und V gilt, dass

$$(\mathbb{E}(UV))^2 \le \mathbb{E}(U^2)\mathbb{E}(V^2). \tag{15}$$

Wir definieren nun $U := X - \mathbb{E}(X)$ und $V := Y - \mathbb{E}(Y)$.

Dann besagt die Cauchy-Schwarz Ungleichung, dass

$$\mathbb{E}\left((X-\mathbb{E}(X))(Y-\mathbb{E}(Y)\right)^2 \le \mathbb{E}\left((X-\mathbb{E}(X))^2\right)\mathbb{E}\left((Y-\mathbb{E}(Y))^2\right). \tag{16}$$

Also gilt

$$\mathbb{C}(X,Y)^2 \le \mathbb{V}(X)\mathbb{V}(Y) \Leftrightarrow \frac{\mathbb{C}(X,Y)^2}{\mathbb{V}(X)\mathbb{V}(Y)} \le 1. \tag{17}$$

$\square$

Theorem (Jensensche Ungleichung)

X sei eine Zufallsvariable und $g : \mathbb{R} \to \mathbb{R}$ eine konvexe Funktion, d.h.

$$g(\lambda x_1 + (1-\lambda)x_2) \le \lambda g(x_1) + (1-\lambda)g(x_2) \tag{18}$$

für alle $x_1, x_2 \in \mathbb{R}$, $\lambda \in [0,1]$. Dann gilt

$$\mathbb{E}(g(X)) \ge g(\mathbb{E}(X)). \tag{19}$$

Analog sei $g : \mathbb{R} \to \mathbb{R}$ eine konkave Funktion, d.h.

$$g(\lambda x_1 + (1-\lambda)x_2) \ge \lambda g(x_1) + (1-\lambda)g(x_2) \tag{20}$$

für alle $x_1, x_2 \in \mathbb{R}$, $\lambda \in [0,1]$. Dann gilt

$$\mathbb{E}(g(X)) \le g(\mathbb{E}(X)). \tag{21}$$

Bemerkungen

- Bei konvexem g liegt der Funktionsgraph unter der Geraden von $g(x_1)$ zu $g(x_2)$.
- Bei konkavem g liegt der Funktionsgraph über der Geraden von $g(x_1)$ zu $g(x_2)$.
- Der Logarithmus ist eine konkave Funktion, also gilt $\mathbb{E}(\ln X) \le \ln \mathbb{E}(X)$.

Beweis

Wir zeigen die Ungleichung für den Fall einer konkaven Funktion g.

- Es sei f eine Tangente am Punkt $g(\mathbb{E}(X))$.
- f sei also eine linear-affine Funktion der Form

$$f(X) := aX + b \text{ für } a, b \in \mathbb{R} \text{ mit } f(\mathbb{E}(X)) = g(\mathbb{E}(X)). \tag{22}$$

- Weil g konkav ist, gilt $g(x) \le ax + b$ für alle $x \in \mathbb{R}$, also $g(X) \le aX + b$.
- Also gilt

$$\mathbb{E}(g(X)) \le \mathbb{E}(aX + b) = a\mathbb{E}(X) + b = f(\mathbb{E}(X)) = g(\mathbb{E}(X)). \tag{23}$$

□

Wahrscheinlichkeitsungleichungen

Erwartungswertungleichungen

Gesetze der Großen Zahl

Zentrale Grenzwertsätze

Selbstkontrollfragen

Überblick

- Es gibt ein *Schwaches Gesetz der Großen Zahl* und ein *Starkes Gesetz der Großen Zahl*.
- Intuitiv besagen beide Gesetze, dass sich das Stichprobenmittel von unabhängigen und identisch verteilten Zufallsvariablen für eine große Anzahl an Zufallsvariablen dem Erwartungswert der zugrundeliegenden Verteilung nähert.
- Das Schwache und das Starke Gesetz der Großen Zahl unterscheiden sich in Hinblick auf die zu ihrer Formulierung benutzen Formen der *Konvergenz von Zufallsvariablen*.
  - Das Schwache Gesetz basiert auf der *Konvergenz in Wahrscheinlichkeit*.
  - Das Starke Gesetz basiert auf der *fast sicheren Konvergenz*.
- Wir begnügen uns mit Konvergenz in Wahrscheinlichkeit und dem Schwachen Gesetz.

Definition (Konvergenz in Wahrscheinlichkeit)

Eine Folge von Zufallsvariable $X_1, X_2, ...$ *konvergiert gegen eine Zufallsvariable X in Wahrscheinlichkeit*, wenn für jedes noch so kleine $\epsilon > 0$ gilt, dass

$$\lim_{n\to\infty} \mathbb{P}(|X_n - X| < \epsilon) = 1 \Leftrightarrow \lim_{n\to\infty} \mathbb{P}(|X_n - X| \geq \epsilon) = 0 \tag{24}$$

Die Konvergenz von $X_1, X_2, ...$ gegen X in Wahrscheinlichkeit wird geschrieben als

$$X_n \xrightarrow[n\to\infty]{\mathrm{P}} X \tag{25}$$

Bemerkungen

- $X_n \xrightarrow[n\to\infty]{\mathrm{P}} X$ heißt, dass sich die Wahrscheinlichkeit, dass X_n in dem zufälligen Intervall $]X - \epsilon, X + \epsilon[$ 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 $X := a$ die Verteilung von X_n mehr und mehr um a konzentriert, wenn n gegen Unendlich strebt.

Theorem (Schwaches Gesetz der Großen Zahl)

$X_1, ..., X_n$ seien unabhängig und gleichverteilte Zufallsvariablen mit $\mathbb{E}(X_i) = \mu$ für alle $i = 1, ..., n$. Weiterhin bezeichne

$$\bar{X}_n := \frac{1}{n}\sum_{i=1}^{n} X_i \tag{26}$$

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

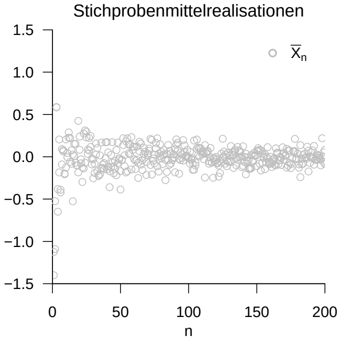
$$\bar{X}_n \xrightarrow[n\to\infty]{\mathrm{P}} \mu. \tag{27}$$

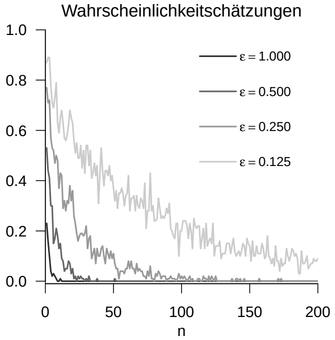
Bemerkungen

- Für einen Beweis siehe zum Beispiel Georgii (2009), Abschnitt 5.1.
- $\bar{X}_n \xrightarrow[n\to\infty]{P} \mu$ heißt, dass die Wahrscheinlichkeit, dass das Stichprobenmittel nahe dem Erwartungswert der zugrundeliegenden Verteilung liegt, sich 1 nähert, wenn $n \to \infty$.

Beispiel $(X_1, ..., X_n \sim N(0,1))$

- Die linke Abbildung zeigt Realisationen von $\bar{X}_n$ als Funktion von n.
- Die rechte Abbildung zeigt Schätzungen von $\mathbb{P}(|\bar{X}_n - \mu| \geq \epsilon)$ als Funktionen von n und ϵ.





Wahrscheinlichkeitsungleichungen

Erwartungswertungleichungen

Gesetze der Großen Zahl

Zentrale Grenzwertsätze

Selbstkontrollfragen

Überblick

- Die Zentralen Grenzwertsätze besagen grob, dass die Summe von unabhängigen Zufallsvariablen asymptotisch, d.h. für unendlich viele Zufallsvariablen, normalverteilt ist.

- Modelliert man eine Messgröße Y also als Summe eines deterministischen Einflusses μ und der Summe $\varepsilon := \sum_{i=1}^n X_i$ einer Vielzahl von unabhängigen Zufallsvariablen $X_i, i = 1, ..., n$, welche unbekannte Störeinflüsse beschreiben, so ist für großes n die Annahme

$$Y = \mu + \varepsilon \text{ mit } \varepsilon \sim N(m, s^2) \tag{28}$$

  also mathematisch gerechtfertigt. Wie wir später sehen werden, liegt die Annahme in Gleichung (28) vielen statischen Modellen zugrunde.

- In der "Lindenberg und Lévy" Form des Zentralen Grenzwertsatzes werden unabhängig und identische Zufallsvariablen vorausgesetzt. In der "Liapunov" Form werden nur unabhängige Zufallsvariablen voraussetzt. Der Beweis der "Lindenberg und Lévy" Form ist einfacher als der Beweis der "Liapunov" Form. Wir verzichten hier aberauf die Angabe von Beweisen.

- In beiden Formulierungen des Zentralen Grenzwertsatzes die betrachtete Konvergenz von Zufallsvariablen die *Konvergenz in Verteilung*, welche wir zunächst einführen.

Definition (Konvergenz in Verteilung)

Eine Folge $X_1, X_2, ...$ von Zufallsvariablen *konvergiert in Verteilung gegen eine Zufallsvariable* X, wenn

$$\lim_{n \to \infty} P_{X_n}(x) = P_X(x). \tag{29}$$

für alle x an denen P_X stetig ist.
Die Konvergenz in Verteilung von $X_1, X_2, ...$ gegen X wird geschrieben als

$$X_n \xrightarrow[n \to \infty]{\text{D}} X, \tag{30}$$

Gilt $X_n \xrightarrow[n \to \infty]{\text{D}} X$, dann heißt die Verteilung von X die *asymptotische Verteilung der Folge* $X_1, X_2, ...$

Bemerkungen

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

Theorem (Zentraler Grenzwertsatz nach Lindenberg und Lévy)

$X_1, ..., X_n$ seien unabhängig und identisch verteilte Zufallsvariablen mit Erwartungswert $\mu := \mathbb{E}(X_i)$ und Varianz $\sigma^2 := \mathbb{V}(X_i)$, $0 < \sigma^2 < \infty$ für alle $i = 1, ..., n$.
Dann gilt für jedes feste x, dass

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \le x\right) = \Phi(x), \tag{31}$$

wobei Φ KVF der Standardnormalverteilung bezeichnet.

Bemerkung

- Wir beweisen später noch folgende Implikationen des Zentralen Grenzwertsatzes:
  - Für $n \to \infty$ gilt asymptotisch $\sum_{i=1}^n X_i \sim N(n\mu, n\sigma^2)$.
  - Für $n \to \infty$ gilt asymptotisch $\bar{X}_n \sim N\left(\mu, \frac{\sigma^2}{n}\right)$.

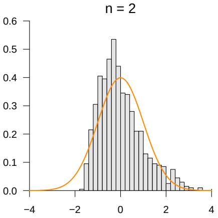
Beispiel ($X_1, ..., X_n \sim \chi^2(k)$)

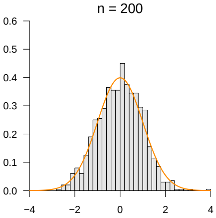
- Wir halten ohne Beweis fest, dass $\mathbb{E}(X_i) = k$ und $\mathbb{V}(X_i) = 2k$.
- Wir betrachten das Szenario $X_i \sim \chi(3)$ für $i = 1, ..., n$.
- Die linken Abbildungen zeigen Histogrammschätzer der Wahrscheinlichkeitsdichte von

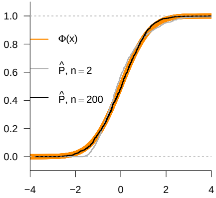
$$Y_n := \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \tag{32}$$

basierend auf 1000 Realisationen von Y_n für $n = 2$ und $n = 200$, sowie die WDF von $N(0, 1)$.

- Die rechte Abbildung zeigt die entsprechenden (empirischen) kumulativen Verteilungsfunktionen.







Theorem (Zentraler Grenzwertsatz nach Liapounov)

$X_1, ..., X_n$ seien unabhängige aber nicht notwendigerweise identisch verteilten Zufallsvariablen mit den Eigenschaften

$$\mathbb{E}(|X_i - \mu_i|^3) < \infty \text{ und } \lim_{n\to\infty} \frac{\sum_{i=1}^n \mathbb{E}\left(|X_i - \mu_i|^3\right)}{(\sum_{i=1}^n \sigma_i^2)^{3/2}} = 0. \tag{33}$$

Es sei weiterhin $\mu_i := \mathbb{E}(X_i)$ und $\sigma_i^2 = \mathbb{V}(X_i)$ für $i = 1, ..., n$. Dann gilt für jedes feste x, dass

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{\sum_{i=1}^n X_i - \sum_{i=1}^n \mu_i}{\sqrt{\sum_{i=1}^n \sigma_i^2}} \le x\right) = \Phi(x), \tag{34}$$

wobei Φ KVF der Standardnormalverteilung bezeichnet.

Bemerkungen

- Für $n \to \infty$ gilt also asymptotisch $\sum_{i=1}^n X_i \sim N\left(\sum_{i=1}^n \mu_i, \sum_{i=1}^n \sigma_i^2\right)$.
- Die Summe vieler unabhängiger Zufallsvariablen ist asymptotisch normalverteilt.

Wahrscheinlichkeitsungleichungen

Erwartungswertungleichungen

Gesetze der Großen Zahl

Zentrale Grenzwertsätze

Selbstkontrollfragen

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?

References

DeGroot, Morris H., and Mark J. Schervish. 2012. *Probability and Statistics*. 4th ed. Boston: Addison-Wesley.

Georgii, Hans-Otto. 2009. *Stochastik: Einführung in die Wahrscheinlichkeitstheorie und Statistik*. 4., überarb. und erw. Aufl. De-Gruyter-Lehrbuch. Berlin: de Gruyter.