

2.8 Tschebyscheffsche Ungleichung

Wir haben bisher Verfahren kennen gelernt, Wahrscheinlichkeiten auszurechnen, wenn eine konkrete Verteilung vorliegt. Einige dieser Verteilungen lassen sich „aus der Natur heraus“ erklären – wie die Binomial, die Hypergeometrische, die geometrische, die Multinomial-, die diskrete und die stetige Gleichverteilung – andere müssen dem Benutzer gegeben sein, so wie die Normal- die Poisson- und die Exponentialverteilung.

Wenn nun die Verteilung einer Zufallsvariablen nicht bekannt ist, so ist es nicht möglich, Wahrscheinlichkeiten exakt zu berechnen. Trotzdem hat man in manchen Fällen Glück und kann gewisse Wahrscheinlichkeiten zumindest abschätzen. Hierzu dient die Tschebyscheffsche Ungleichung. Wichtig ist, daß man die Voraussetzungen überprüft. Das Vorgehen erläutert das nachfolgende

LAMBERT-KOCHREZEPT TSCHEBYSCHEFFSCHE UNGLEICHUNG

Voraussetzungen

- die konkrete Verteilung ist unbekannt
- Erwartungswert μ und Varianz σ^2 hingegen sind bekannt.
- es ist gefragt nach einer Mindest- oder nach einer Höchstwahrscheinlichkeit, nicht danach, eine Wahrscheinlichkeit exakt zu berechnen.
- der fragliche Bereich, über den eine Wahrscheinlichkeitsaussage gemacht werden soll, liegt *symmetrisch* um den Erwartungswert μ herum. Rechts und links um den Erwartungswert herum beträgt die Länge des fraglichen Bereichs genau ε bzw. $k \cdot \sigma$.

Anwendung der Ungleichung

- 1. Fall

- verbal: dann ist die Wahrscheinlichkeit, daß die Zufallsvariable X vom eigenen Erwartungswert μ um

höchstens ε abweicht, mindestens $1 - \frac{\sigma^2}{\varepsilon^2}$

- in Formeln:

$$P(|X - \mu| < \varepsilon) \geq 1 - \frac{\sigma^2}{\varepsilon^2}$$

- graphisch: die Wahrscheinlichkeit also, daß X in den schraffierten Bereich fällt, ist mindestens $1 - \frac{\sigma^2}{\varepsilon^2}$

- 2. Fall

- verbal: die Wahrscheinlichkeit, daß die Zufallsvariable X vom eigenen Erwartungswert μ um mindestens ε abweicht, beträgt

höchstens $\frac{\sigma^2}{\varepsilon^2}$

- in Formeln: $P(|X - \mu| > \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}$

die Wahrscheinlichkeit, dass X in den schraffierten Bereich fällt,

ist höchstens $\frac{\sigma^2}{\varepsilon^2}$

MERKE:

- im Fall 1 ist es entscheidend, daß die Wahrscheinlichkeiten

nicht ungefähr $1 - \frac{\sigma^2}{\varepsilon^2}$ ist, sondern irgendwo rechts von dieser

Zahl liegt (und klarerweise kleiner oder gleich 1 ist). Gleiches

gilt für Fall 2 mit der Zahl $\frac{\sigma^2}{\varepsilon^2}$. Hier ist die Wahrscheinlichkeit

nicht ungefähr gleich diese Zahl, sondern liegt links hiervon (und ist größer oder gleich null)

- Die Tschebyscheffsche Ungleichung ist also eine *Abschätzung*, keine Schätzung!

Aufgabe 33

Läßt sich etwas über die Wahrscheinlichkeit aussagen, daß eine Zufallsvariable, deren Verteilung unbekannt ist, die aber den Erwartungswert von 5 und die Streuung von 1 hat, zwischen 3 und 7 liegt?

Gefragt ist also nach einer Aussage über $P(3 \leq X \leq 7)$. Dies lässt sich umrechnen in $P(3 \leq X \leq 7) = P(3 - 5 \leq X - 5 \leq 7 - 5)$

$$= P(-2 \leq X - 5 \leq 2) = P(|X - 5| \leq 2).$$

Damit ist genau die Gestalt der Tschebyscheffschen Ungleichung erreicht. Der Erwartungswert μ ist 5, der Abstand ε zwischen den Grenzen 3 und 7 zu diesem Wert ist jeweils genau $\varepsilon = 2$ (und also liegt 5 in der Mitte des fraglichen Bereichs!) und die Streuung lautet

$$\sigma = 1. \text{ Also } P(|X - 5| < 2) \geq 1 - \frac{1^2}{2^2} = 1 - \frac{1}{4} = 0,75$$

Damit liegt die gefragte Wahrscheinlichkeit zwischen 0,75 und 1.

Aufgabe 34

Was kann man über die Wahrscheinlichkeit aussagen, daß eine Zufallsvariable X , deren Verteilung unbekannt ist, deren Erwartungswert μ und deren Varianz σ^2 hingegen bekannt sind, um

mehr als das Doppelte ihrer eigenen Streuung vom Erwartungswert abweicht?

Hier ist die Abweichung vom Erwartungswert μ genau $\varepsilon = 2 \cdot \sigma$. Damit rechnet man

$$P(|X - \mu| > 2 \cdot \sigma) \leq \frac{\sigma^2}{(2 \cdot \sigma)^2} = \frac{\sigma^2}{4 \cdot \sigma^2} = \frac{1}{4} = 0,25$$

An dieser Stelle lässt sich wegen $\varepsilon = k \cdot \sigma$ eine leicht anders aussehende Darstellung der Tschebyscheffschen Ungleichung angeben.

Fall 1

- verbal: Die Wahrscheinlichkeit, daß die Zufallsvariable X vom eigenen Erwartungswert um höchstens das k -fache der eigenen Streuung σ abweicht, ist mindestens $1 - \frac{1}{k^2}$.
- in Formeln: $P(|X - \mu| \leq k \cdot \sigma) \geq 1 - \frac{1}{k^2}$

Fall 2

Verbal: Genauso ist die Wahrscheinlichkeit, daß die Zufallsvariable X vom eigenen Erwartungswert μ um mindestens das k -fache der eigenen Streuung σ abweicht, höchstens $\frac{1}{k^2}$.

In Formeln: $P(|X - \mu| \geq k \cdot \sigma) \leq \frac{1}{k^2}$

Interessant ist nun zu beobachten, was passiert, wenn eine konkrete Verteilung bekannt ist. Dazu eine Erweiterung der obigen Aufgabe:

Aufgabe 35

Im Beispiel 1 sei nachträglich bekannt geworden, daß die Zufallsvariable X normalverteilt ist mit dem Erwartungswert 5 und der Streuung 1, daß also $X \sim N(5; 1)$ gilt.

Dann kann man die Wahrscheinlichkeit für das fragliche Ereignis genau ausrechnen:

$$\begin{aligned} P(|X - 5| \leq 2) &= P(3 \leq X \leq 7) = P\left(\frac{3-5}{1} \leq X_{St} \leq \frac{7-5}{1}\right) \\ &= P(-2 \leq X_{St} \leq 2) = \phi(2) - \phi(-2) = \phi(2) - [1 - \phi(2)] = 2 \cdot \phi(2) - 1 \\ &= 2 \cdot 0,9772 - 1 = 0,9544 \end{aligned}$$

MERKE:

1. die gute Nachricht damit zuerst: die Aussage der Tschebyscheffschen Ungleichung, daß die gesuchte Wahrscheinlichkeit größer als 0,75 ist, war richtig.
2. die schlechte Nachricht: die Wahrscheinlichkeit kann dann nochmal sehr weit weg liegen von 0,75. Hieran wird am Beispiel die obige Aussage klar, dass die Tschebyscheffsche Ungleichung eine *Abschätzung*, hingegen *keine Schätzung* ist.

Aufgabe 35

Eine faire Münze werde zehnmal geworfen. Die Zufallsvariable X sei definiert als die Anzahl der gefallen Köpfe.

- a) Wie ist die Zufallsvariable X verteilt und was sind ihr Erwartungswert und ihre Varianz?
- b) Berechne die Wahrscheinlichkeit, dass mehr als zwei und weniger als achtmal Kopf fällt?
- c) Berechne die unter b) gesuchte Wahrscheinlichkeit unter der Annahme, dass die Verteilung von X mit ausreichender Genauigkeit durch eine Normalverteilung mit den unter a) berechneten Parametern angenähert werden kann.

- d) Gib eine Abschätzung der unter b) gesuchten Wahrscheinlichkeit mit Hilfe der Tschebyscheffschen Ungleichung an.
- e) Vergleiche die Ergebnisse b)-d) und erläutere die Unterschiede.

Richtig-Falsch-Fragen Tschebyscheffsche Ungleichung

Die folgenden Aussagen sind richtig oder falsch. Entscheide.

1) Die Tschebyscheffsche ist eine Schätzung und nicht lediglich eine Abschätzung, wenn die Zufallsvariable normalverteilt ist.

2) Wenn man die Tschebyscheffsche Ungleichung rechnet und hinterher als Zusatzinformation erfährt, dass der Verteilungstyp doch bekannt ist und die gesuchte Wahrscheinlichkeit mit dieser Zusatzinformation nochmals ausrechnet, dann kann die Aussage der Tschebyscheffschen Ungleichung manchmal als verkehrt angesehen werden.

3) Die Tschebyscheffsche Ungleichung kann auch solche Ergebnisse

liefern wie $P(|X - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2} = 1,6$, d.h. eine Tautologie, denn dass

eine Wahrscheinlichkeit kleiner oder gleich 1,6 ist, ist sowieso immer richtig, denn sie liegt ja zwischen 0 und 1.