Prof. Dr. Jan Beyersmann
Matthias Weber

SS 2021
Übung: 23. Juli 2021, Abgabe: 22. Juli

Angewandte Stochastik 1 – Übungsblatt 7

Dieses Übungsblatt wird in der Übung am 23. Juli um 10:15 Uhr diskutiert. Bitte gebt eure Lösungen bis zum Donnerstag, 22. Juli um 12:00 Uhr über Moodle ab. Gebt eure Lösungen als exakt eine PDF-Datei ab, andere Dateiformate oder mehrere Dateien sind nicht erlaubt. Gebt die Lösungen nach Möglichkeit bitte in Gruppen von zwei oder drei Personen ab. Bitte beachtet auch die Hinweise zur Übung in Moodle.

Musterlösungen

Aufgabe 1 (2 + 2 + 2 + 3 = 9 Punkte)

Angenommen eine Münze ist 'gezinkt' und man möchte die Wahrscheinlichkeit p berechnen, dass diese nach einem Wurf 'Kopf' zeigt. Dazu wird die Münze n-mal geworfen und K_i bezeichne das Ereignis, im i-ten Wurf 'Kopf' zu werfen.

(a) Warum ist

$$\widehat{p} = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{K_i}$$

ein Näherungswert für p? Begründe mithilfe eines Gesetzes der großen Zahlen.

Lösung:
Es ist $\mathbb{1}_{K_i} \sim \mathrm{Bern}(p)$ iid., also folgt mit dem SGGZ: $\widehat{p} \to \mathbb{E}(\mathbb{1}_{K_i}) = p$.

Hinweis: Nutze den zentralen Grenzwertsatz, um die folgenden Aufgabenteile näherungsweise zu lösen.

Lösung:
Wir lösen als erstes Teil (d) und verwenden dessen Zwischenergebnisse für die anderen Teile.

(b) Angenommen, der wahre Wert von p ist 0.3 und die Münze wird 100 mal geworfen. Berechne $\mathbb{P}(0.25 \leq \widehat{p} \leq 0.35)$.

Lösung:
Wir nutzen A aus Teil (d) und sehen

$$\mathbb{P}(0.25 \leq \widehat{p} \leq 0.35) = A(n, p)$$

für $p = 0.3$ und $n = 100$. Also ist

$$\mathbb{P}(0.25 \leq \widehat{p} \leq 0.35) = 2\Phi\left(10\frac{0.05}{\sqrt{0.3 \cdot 0.7}}\right) - 1 = 2 \cdot 0.8621 - 1 = 0.7242\,.$$

(c) Wie oft muss die Münze geworfen werden, damit $\mathbb{P}(0.25 \leq \widehat{p} \leq 0.35) \geq 0.95$ für $p = 0.3$ gilt?

Lösung:
Mit A aus Teil (d) suchen wir n, sodass $A(n, 0.3) \geq 0.95$. Es ist

$$A(n, 0.3) = 2\Phi\left(\sqrt{n}\frac{0.05}{\sqrt{0.3 \cdot 0.7}}\right) - 1 \geq 0.95 \iff \sqrt{n}\frac{0.05}{\sqrt{0.3 \cdot 0.7}} \geq 1.96\,,$$

siehe (d). Also muss gelten

$$n \geq 322.7\,.$$

Da n ganzzahlig sein muss, folgt auch $n \geq 323$.

(d) Wie oft muss die Münze geworfen werden, damit $\mathbb{P}(p - 0.05 \leq \widehat{p} \leq p + 0.05) \geq 0.95$ für alle $p \in [0, 1]$ gilt?

Lösung:
Mit $S_n = n\widehat{p}$ und dem ZGWS gilt

$$\tilde{Y} := \frac{n\widehat{p} - np}{\sqrt{np(1-p)}} = \sqrt{n}\frac{\widehat{p} - p}{\sqrt{p(1-p)}} \to Y \sim N(0,1)\,.$$

Dann ist

$$\begin{aligned} A(n,p) := P(p - 0.05 \leq \widehat{p} \leq p + 0.05) &= P\left(\sqrt{n}\frac{-0.05}{\sqrt{p(1-p)}} \leq \tilde{Y} \leq \sqrt{n}\frac{0.05}{\sqrt{p(1-p)}}\right) \\ &\approx \Phi\left(\sqrt{n}\frac{0.05}{\sqrt{p(1-p)}}\right) - \Phi\left(-\sqrt{n}\frac{0.05}{\sqrt{p(1-p)}}\right) \\ &= 2\Phi\left(\sqrt{n}\frac{0.05}{\sqrt{p(1-p)}}\right) - 1 \end{aligned}$$

und

$$A(n,p) \geq 0.95 \iff \sqrt{n}\frac{0.05}{\sqrt{p(1-p)}} \geq \Phi^{-1}(0.975) = 1.96 \iff n \geq 1536p(1-p)\,.$$

Die Funktion $f(p) = p(1-p)$ wird maximal bei $p = 0.5$, da $\frac{d}{dp}f(p) = 1 - 2p = 0 \iff p = 0.5$ und $f(0) = f(1) = 0$ und $f(0.5) = 0.25$. Also gilt $A(p) \geq 0.95$ für alle $p \in (0,1)$ wenn $n \geq 384$.

Hinweis: Benötigte Funktionswerte der Verteilungsfunktion der Standardnormalverteilung bzw. der Quantilfunktion sind: $\Phi(1.09) = 0.8621$, $\Phi^{-1}(0.975) = 1.96$.

Aufgabe 2 (3 + 3 = 6 Punkte)

Es soll der durchschnittliche Wasserverbrauch pro Person und Tag (in Liter l) einer Stadt bestimmt werden. Wir nehmen an, dass die unabhängigen und identisch verteilten Zufallsvariablen $Y_1, \ldots, Y_m$ den Wasserverbrauch von m Personen beschreiben. Es gelte $\mathbb{E}\, Y_j = 50l$ und $\operatorname{Var} Y_j = 87l^2$ für alle $j = 1, \ldots, m$.

(a) Verwende die Tschebyscheff-Ungleichung, um festzustellen, von wie vielen Personen der Wasserverbrauch mindestens untersucht werden muss, damit die Wahrscheinlichkeit, dass $\overline{Y}_m = \frac{1}{m} \sum_{j=1}^{m} Y_j$ um nicht mehr als $8l$ vom Erwartungswert $\mathbb{E}\, Y_j = 50l$ abweicht, mindestens 0.99 beträgt.

Lösung:
Mit der Tschebyscheff-Ungleichung gilt

$$P(|\overline{Y}_m - \mathbb{E}\, \overline{Y}_m| \geq 8) \leq \frac{\operatorname{Var} \overline{Y}_m}{8^2}\,.$$

Außerdem ist $\mathbb{E}\, \overline{Y}_m = EY_j = 50$ *und* $\operatorname{Var} \overline{Y}_m = \frac{1}{m} \operatorname{Var} Y_1 = 87/m$. *Also ist*

$$P(|\overline{Y}_m - 50| \geq 8) \leq 0.01\,,$$

wenn

$$\frac{87}{m8^2} \leq 0.01$$

oder

$$m \geq 135.9\,.$$

Da m ganzzahlig sein muss, muss also gelten $m \geq 136$.

(b) Verwende nun statt der Tschebyscheff-Ungleichung den zentralen Grenzwertsatz, um die Mindestanzahl an Personen zu bestimmen, sodass die Bedingung aus (a) erfüllt ist.

Lösung:
Mit dem ZGWS ist für $Y \sim N(0, 1)$

$$P(42 \leq \overline{Y}_m \leq 58) \approx P\left(\sqrt{m}\frac{-8}{\sqrt{87}} \leq Y \leq \sqrt{m}\frac{8}{\sqrt{87}}\right) = 2\Phi(\sqrt{m} \cdot 0.858) - 1$$

und

$$2\Phi(\sqrt{m} \cdot 0.858) - 1 \geq 0.99 \iff m \geq 9.05\,.$$

Da m ganzzahlig sein muss, muss also gelten $m \geq 10$.

Hinweis: Benötigter Wert der Quantilfunktion der Standardnormalverteilung: $\Phi^{-1}(0.995) = 2.58$.

Aufgabe 3 (3 + 2 = 5 Bonuspunkte)

In den folgenden Aufgaben sollen Monte-Carlo Methoden (siehe Kapitel 5.1.2) genutzt werden. Implementiere die Lösungen in einer Programmiersprache deiner Wahl und gib den Code und den dazugehörigen Output ab.

(a) Gegeben seien zwei Kreise mit Mittelpunkten $c_1 = (0, 0.5)$ und $c_2 = (1, 0.5)$ und identischem Radius $r = \frac{1}{\sqrt{2}}$. Berechne die Fläche des Schnitts der beiden Kreise approximativ basierend auf 10^4 zufällig gewählten Punkten in einem Rechteck $[a, b] \times [c, d]$. Wähle die Werte a, b, c, d, so, dass der Schnitt der beiden Kreise vollständig in diesem Rechteck enthalten und die Fläche des Rechtecks minimal ist.

Lösung:

Die beiden Kreise schneiden sich in den Punkten $(0.5, 0)$ *und* $(0.5, 1)$. *Die maximale Breite des Schnitts ist bei* $y = 0.5$ *von* $x = 1 - r \approx 0.29$ *(abgerundet) bis* $x = 0 + r = 0.71$ *(aufgerundet). Also ist* $[a, b] \times [c, d] = [0.29, 0.71] \times [0, 1]$.

In R implementiert:

```r
n = 10^4
x = runif(n, min = 0.29, max = 0.71)
y = runif(n)

# Abstand zu Mittelpunkt (0, 0.5) von Kreis 1
d1 = sqrt((x - 0)^2 + (y - 0.5)^2)

# Abstand zu Mittelpunkt (1, 0.5) von Kreis 2
d2 = sqrt((x -1)^2 + (y - 0.5)^2)

# Punkte im Schnitt (d1 <= 1/sqrt(2) und d2 <= 1/sqrt(2))
 hits = d1 <= 1/sqrt(2) & d2 <= 1/sqrt(2)

# Wahrscheinlichkeit, den Kreisschnitt zu treffen
prob = mean(hits)

# Die Fläche der "Grundmenge" [0.29, 0.71] x [0, 1] ist
a = (0.71 - 0.29) * (1 - 0)

# Die geschätzte Fläche des Kreisschnitts ist
prob * a
```

Damit schätzen wir die Fläche des Schnitts auf 0.285138. *Zum Vergleich: Der wahre Wert ist ungefähr* 0.28540

(b) Berechne das folgende Integral approximativ mittels Monte-Carlo Integration (siehe Kapitel 5.1.2) basierend auf 10^4 zufällig gewählten Punkten in $[0,1]^3$:

$$\int_{[0,1]^3} \cos(x_1 x_2) \cdot \sin(\cos(x_3 x_1))\, dx_1 dx_2 dx_3$$

Lösung:
In R implementiert:

```
n = 10^4
x = runif(n)
y = runif(n)
z = runif(n)

f = function(x, y, z) {
        cos(x * y) * sin(cos(z * x))
}

werte = f(x, y, z)

# Da x, y, z aus [0,1]^3 gezogen wurden (und dessen Volumen = 1 ist ), ist das
     Integral ungefähr
mean(werte)
```

Der geschätzte Wert des Integrals ist 0.7663139. *Zum Vergleich: Der wahre Wert ist ungefähr* 0.76653.

Multiple Choice Aufgaben

Bei den folgenden Multiple Choice Aufgaben sind die richtigen Aussagen anzugeben. Dabei kann genau eine Aussage zutreffen, mehrere Aussagen können zutreffen oder aber keine der Aussagen trifft zu.

Für die Bewertung des Übungsblatts gibt jede richtige Antwort 1 Punkt, jede falsche Antwort gibt 1 Punkt Abzug. Wie in der Prüfung auch, habt ihr hier die Möglichkeit, für einen Teil keine Antwort anzugeben und damit keine positiven oder negativen Punkte dafür zu erhalten. Je Aufgabe können minimal 0 Punkte erreicht werden.

Zur Abgabe könnt ihr diese Aufgabe gerne ausdrucken und ankreuzen oder die Antworten (richtig/falsch/keine Antwort) mit einer klaren Nummerierung (1-5 oder a-e) versehen.

Aufgabe 4 (5 Punkte)

Betrachte eine Folge von Zufallsvariablen $X_1, X_2, X_3, \ldots$ mit gemeinsamen Erwartungswert $\mu = E(X_1) = E(X_2) = E(X_3) = \ldots$. Die Zufallsvariablen können im Allgemeinen abhängig oder unabhängig sein. Es sei $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$.

Richtig	Falsch	
$\otimes$	$\bigcirc$	$X_1, X_2, X_3, \ldots$ seien abhängig. Dann gilt $E(\bar{X}) = \mu$.
$\bigcirc$	$\otimes$	Es gelte $E(\bar{X}) = \mu$. Dann muss auch $\sigma^2 = Var(X_1) = Var(X_2) = Var(X_3) = \ldots$ gelten.
$\bigcirc$	$\otimes$	$X_1, X_2, X_3, \ldots$ seien identisch verteilt. Dann gilt das starke Gesetz der großen Zahlen.
$\bigcirc$	$\otimes$	$X_1, X_2, X_3, \ldots$ seien unabhängig identisch verteilt mit Varianz $\sigma^2 \in (0, \infty)$. Dann gilt $\sqrt{n}(\bar{X} - \mu) \xrightarrow{P} 0$.
$\otimes$	$\bigcirc$	$X_1, X_2, X_3, \ldots$ seien unabhängig identisch Bernoulli-verteilt mit Erfolgswahrscheinlichkeit $p \in (0, 1)$. Die Anzahl der Erfolge ist dann ungefähr normalverteilt.