D-INFK
FS 2022

Prof. Vincent Tassion
Laurin Köhler-Schindler

Wahrscheinlichkeit & Statistik

Serie 6

Abgabe bis Mittwoch (06.04.2022) um 10:15 Uhr

Diese Serie beschäftigt sich mit dem Konzept der Varianz, wichtigen Ungleichungen und mit gemeinsamen Verteilungen diskreter Zufallsvariablen.

Weitere Informationen und Instruktionen zur Abgabe unter
https://metaphor.ethz.ch/x/2022/fs/401-0614-00L/

Aufgabe 6.1 [Varianz der Laufzeit eines Algorithmus]
In dieser Aufgabe betrachten wir erneut den Algorithmus aus den Aufgaben 3.1 und 5.1:

$i = 1$
While $X_i = X_{i+1} = 1$ **do**
  $i = i + 2$
endwhile
$Z = X_i + 2 \cdot X_{i+1}$
Return Z

Zur Erinnerung: $(X_i)_{i\geq 1}$ ist eine unendliche Folge von unabhängigen, Ber(1/2)-verteilten Zufallsvariablen. Sei T die Laufzeit des Algorithmus, also die Anzahl der Durchläufe der While-Schleife. Aus Aufgabe 5.1 wissen wir bereits, dass $\mathbb{E}[T] = 1/3$. Ziel dieser Aufgabe ist es, die Varianz σ_T^2 zu berechnen.

(a) Zeige, dass die Zufallsvariable $S := T + 1$ Geom(3/4)-verteilt ist.

(b) Zeige, dass $\sigma_S^2 = \sigma_T^2$.

(c) Berechne $\mathbb{E}[S^2]$. Nutze das Ergebnis, um die Varianz σ_S^2 zu bestimmen.

(d) Um drei unabhängige Zufallsvariablen zu erhalten, die gleichverteilt in $\{0, 1, 2\}$ sind, führen wir den Algorithmus dreimal hintereinander aus. Was ist die Varianz der Gesamtlaufzeit $T' := T_1 + T_2 + T_3$, wobei T_i die Laufzeit der i-ten Ausführung des Algorithmus bezeichnet?

Aufgabe 6.2 [Markow & Chebychev Ungleichung]
Sei $n \in \mathbb{N}$ ungerade. Wir definieren die Mehrheitsfunktion

$$\begin{aligned} \mathrm{MAJ}_n \colon \{0,1\}^n &\to \{0,1\} \\ (x_1, \dots, x_n) &\mapsto \begin{cases} 0, & \text{falls } \sum_{i=1}^n x_i < \frac{n}{2}, \\ 1, & \text{falls } \sum_{i=1}^n x_i > \frac{n}{2}. \end{cases} \end{aligned}$$

Sei $(X_i)_{i\geq 1}$ eine Folge von unabhängigen, Ber(p)-verteilten Zufallsvariablen mit Parameter $p \in [0, 1]$. Definiere

$$Y_n := \mathrm{MAJ}_n(X_1, \dots, X_n).$$

(a) Berechne $\mathbb{P}[Y_n = 1]$, $\mathbb{E}[Y_n]$ und $\sigma_{Y_n}^2$ für $p = 1/2$.

(b) Zeige, dass für alle $p \in [0,1]$,

$$\mathbb{E}[Y_n] \leq 2p \quad \text{und} \quad \sigma^2_{Y_n} \leq \frac{1}{4}.$$

Hinweis: Nutze die Markow Ungleichung, um $\mathbb{P}[\sum_{i=1}^n X_i > \frac{n}{2}]$ abzuschätzen.

(c) Zeige, dass

$$\mathbb{E}[Y_n] = \mathbb{P}[Y_n = 1] \xrightarrow{n\to\infty} \begin{cases} 0, & \text{falls } p < 1/2, \\ 1, & \text{falls } p > 1/2. \end{cases}$$

Hinweis: Nutze die Chebychev Ungleichung.

(d) Schlussfolgere aus (c), dass $\sigma^2_{Y_n} \xrightarrow{n\to\infty} 0$ für $p \neq 1/2$.

Aufgabe 6.3 [Gemeinsame Verteilung diskreter ZVen I]
Seien X und Y zwei diskrete Zufallsvariablen mit folgender gemeinsamer Gewichtsfunktion:

$$p(j,k) = \mathbb{P}\left[X = j, Y = k\right] = \begin{cases} C \cdot \left(\frac{1}{2}\right)^k & \text{für } k = 2,3,\dots \text{ und } j = 1,2,\dots,k-1 \\ 0 & \text{sonst.} \end{cases}$$

(a) Bestimme die Konstante C.

(b) Berechne die Gewichtsfunktionen p_X und p_Y der Randverteilungen von X und Y.

(c) Sind X und Y unabhängig?

Aufgabe 6.4 [Gemeinsame Verteilung diskreter ZVen II: Konstruktion]
Seien W_1 und W_2 endlich oder abzählbar und sei $p : W_1 \times W_2 \to [0,1]$ eine Funktion mit

$$\sum_{x_1 \in W_1, x_2 \in W_2} p(x_1, x_2) = 1.$$

Seien weiterhin U_1 und U_2 zwei unabhängige $\mathcal{U}([0,1])$-verteilter Zufallsvariablen. Ziel dieser Aufgabe ist es, mithilfe von U_1 und U_2 zwei diskrete Zufallsvariablen X_1 und X_2 zu konstruieren, sodass (X_1, X_2) die gemeinsame Gewichtsfunktion/Verteilung $p = (p(x_1,x_2))_{x_1 \in W_1, x_2 \in W_2}$ haben.

(a) Was ist die Gewichtsfunktion p_{X_1} der Randverteilung von X_1? Nutze U_1, um die Zufallsvariable X_1 zu konstruieren.
Erinnerung: In Aufgabe 3.2 haben wir bereits einige Zufallsvariablen aus einer $\mathcal{U}([0,1])$-verteilter Zufallsvariable konstruiert.

(b) Analog zu Aufgabenteil (a) könnte man nun auch U_2 nutzen, um die Zufallsvariable X_2 mit Gewichtsfunktion p_{X_2} (der Randverteilung von X_2) zu konstruieren. Zeige, dass (X_1, X_2) in diesem Fall die gemeinsame Gewichtsfunktion/Verteilung $q = (q(x_1,x_2))_{x_1 \in W_1, x_2 \in W_2}$ haben mit

$$q(x_1, x_2) = p_{X_1}(x_1) \cdot p_{X_2}(x_2).$$

Die Konstruktion aus Aufgabenteil (b) funktioniert also nur, falls $p(x_1,x_2) = p_{X_1}(x_1) \cdot p_{X_2}(x_2)$ für alle $x_1 \in W_1, x_2 \in W_2$ gilt, also wenn die ZVen X_1 und X_2 unabhängig sind.

(c) *[schwer]* Nutze U_2, um die Zufallsvariable X_2 so zu konstruieren, dass (X_1, X_2) die gemeinsame Gewichtsfunktion/Verteilung $p = (p(x_1,x_2))_{x_1 \in W_1, x_2 \in W_2}$ haben.