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:

$$
\begin{aligned}
& i = 1 \\
& \textbf{While} \quad X_i = X_{i+1} = 1 \quad \textbf{do} \\
& \qquad\qquad i = i + 2 \\
& \textbf{endwhile} \\
& Z = X_i + 2 \cdot X_{i+1} \\
& \textbf{Return} \quad Z
\end{aligned}
$$

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

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

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

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

(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:

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

(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 kontruiert.

(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.

Lösung 6.1

(a) In der Lösung von Aufgabenteil 5.1 (a) haben wir bereits gezeigt, dass $\mathbb{P}[T \geq j] = (1/4)^j$ für $j \in \{0, 1, 2, \ldots\}$ gilt. Hieraus folgt

$$\mathbb{P}[T = j] = \mathbb{P}[T \geq j] - \mathbb{P}[T \geq j+1] = \left(\frac{1}{4}\right)^j - \left(\frac{1}{4}\right)^{j+1} = \frac{3}{4} \cdot \left(\frac{1}{4}\right)^j$$

und insbesondere ist $S := T + 1$ Geom(p)-verteilt mit Parameter $p = 3/4$.

(b) Wir betrachten eine deterministische Zufallsvariable X mit $\mathbb{P}[X = 1] = 1$. Man sieht leicht, dass $\mathbb{E}[X] = 1$ und die Varianz einer deterministischen Zufallsvariable ist immer gleich 0:

$$\sigma_X^2 = \mathbb{E}[(X-1)^2] = 0.$$

Da X deterministisch ist, sind T und X unabhängig und somit folgt aus Proposition 4.26

$$\sigma_S^2 = \sigma_T^2 + \sigma_X^2 = \sigma_T^2.$$

(c) Da S Geom(p)-verteilt ist mit Parameter $p = 3/4$, erhalten wir

$$\mathbb{E}[S^2] = \sum_{j=0}^{\infty} j^2 \mathbb{P}[S = j] = p \cdot \sum_{j=0}^{\infty} j^2 (1-p)^{j-1}.$$

Wir betrachten die Funktion $f : (0, 1] \to [0, \infty)$ definiert durch

$$f(p) := \sum_{j=0}^{\infty} (1-p)^{j+1} = \frac{1-p}{p}.$$

Als Potenzreihe ist die Funktion f auf ihrem Konvergenzbereich (hier das halboffene Interval $(0, 1]$)) beliebig oft stetig differenzierbar und es gilt

$$\begin{aligned}
-\frac{1}{p^2} &= f'(p) = \sum_{j=0}^{\infty} \frac{d}{dp}\,(1-p)^{j+1} = -\sum_{j=0}^{\infty} (j+1)(1-p)^j = -\sum_{j=0}^{\infty} j(1-p)^{j-1}, \\
\frac{2}{p^3} &= f''(p) = \sum_{j=0}^{\infty} \frac{d^2}{dp^2}\,(1-p)^{j+1} = \sum_{j=0}^{\infty} (j+1)j(1-p)^{j-1}.
\end{aligned}$$

Durch Umformen erhält man

$$\begin{aligned}
\mathbb{E}[S^2] = p \cdot \sum_{j=0}^{\infty} j^2 (1-p)^{j-1} &= p \cdot (f''(p) + f'(p)) \\
&= \frac{2}{p^2} - \frac{1}{p} = \frac{2-p}{p^2}
\end{aligned}$$

Mithilfe von Proposition 4.26 erhalten wir

$$\sigma_S^2 = \mathbb{E}[S^2] - \mathbb{E}[S]^2 = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2} = \frac{4}{9}.$$

(d) Nach Annahme sind T_1, T_2, T_3 unabhängig und somit folgt aus den vorherigen Aufgabenteilen und Proposition 4.26, dass

$$\sigma_{T'}^2 = \sigma_{T_1}^2 + \sigma_{T_2}^2 + \sigma_{T_3}^2 = 4/3.$$

Lösung 6.2

(a) Für $p = 1/2$ erhalten wir

$$\mathbb{P}[Y_n = 1] = \mathbb{P}\left[\sum_{i=1}^{n} X_i > \frac{n}{2}\right] \overset{\text{[Symmetrie]}}{=} \mathbb{P}\left[\sum_{i=1}^{n} X_i < \frac{n}{2}\right] = \mathbb{P}[Y_n = 0],$$

da die ZVen $(1 - X_i)_{i=1}^n$ ebenfalls unabhängig und Ber(1/2)-verteilt sind. Daraus folgt $\mathbb{P}[Y_n = 1] = 1/2$ und somit auch

$$\mathbb{E}[Y_n] = 1 \cdot \mathbb{P}[Y_n = 1] + 0 \cdot \mathbb{P}[Y_n = 0] = 1/2.$$

Da Y_n Werte in $\{0, 1\}$ annimmt, gilt insbesondere $(Y_n)^2 = Y_n$ und somit folgt

$$\sigma_{Y_n}^2 = \mathbb{E}[(Y_n)^2] - \mathbb{E}[Y_n]^2 = \mathbb{E}[Y_n](1 - \mathbb{E}[Y_n]) = 1/4.$$

(b) In Aufgabenteil (a) haben wir gezeigt, dass

$$\sigma_{Y_n}^2 = \mathbb{E}[Y_n](1 - \mathbb{E}[Y_n]).$$

Da die Funktion $g : [0, 1] \to \mathbb{R}$ mit $g(x) = x(1 - x)$ ihr Maximum an der Stelle $x = 1/2$ annimmt (mit $g(1/2) = 1/4$) und $\mathbb{E}[Y_n] \in [0, 1]$ gilt, da $Y_n \in \{0, 1\}$, erhalten wir

$$\sigma_{Y_n}^2 \leq 1/4$$

für alle $p \in [0, 1]$. Weiterhin gilt

$$\mathbb{E}[Y_n] = \mathbb{P}[Y_n = 1] = \mathbb{P}\left[\sum_{i=1}^{n} X_i > \frac{n}{2}\right] \overset{\text{[Markov]}}{\leq} \frac{\mathbb{E}\left[\sum_{i=1}^{n} X_i\right]}{n/2},$$

wobei wir die Markov Ungleichung für die nicht-negative Zufallsvariable $\sum_{i=1}^{n} X_i$ mit $a = n/2$ angewendet haben. Aufgrund der Linearität des Erwartungswerts erhalten wir

$$\mathbb{E}[Y_n] \leq \frac{\sum_{i=1}^{n} \mathbb{E}[X_i]}{n/2} = \frac{np}{n/2} = 2p.$$

(c) Wir wollen die Chebychev Ungleichung anwenden, um

$$\mathbb{E}[Y_n] = \mathbb{P}[Y_n = 1] = \mathbb{P}\left[\sum_{i=1}^{n} X_i > \frac{n}{2}\right]$$

für $p \neq 1/2$ anzuschätzen. Wir definieren $S_n := \sum_{i=1}^{n} X_i$ und stellen wie in Aufgabenteil (b) fest, dass

$$\mathbb{E}[S_n] = \mathbb{E}\left[\sum_{i=1}^{n} X_i\right] = np.$$

Ausserdem gilt aufgrund der Unabhängigkeit

$$\sigma_{S_n}^2 = \sum_{i=1}^{n} \sigma_{X_i}^2 = np(1 - p).$$

Sei $p < 1/2$. Dann gilt

$$\begin{aligned}\mathbb{E}[Y_n] &= \mathbb{P}\left[S_n > \frac{n}{2}\right] = \mathbb{P}\left[S_n - np > \frac{n}{2} - np\right] \\ &\le \mathbb{P}\left[|S_n - np| > n(1/2 - p)\right] \overset{\text{[Chebychev]}}{\le} \frac{\sigma_{S_n}^2}{(n(1/2-p))^2} \\ &= \frac{np(1-p)}{(n(1/2-p))^2} = \frac{1}{n} \cdot \underbrace{\frac{p(1-p)}{(1/2-p)^2}}_{<\infty} \xrightarrow{n\to\infty} 0,\end{aligned}$$

wobei wir die Chebychev Ungleichung für die Zufallsvariable S_n mit $a = n(1/2-p)$ angewendet haben. Sei nun $p > 1/2$. Dann gilt

$$\begin{aligned}\mathbb{P}\left[S_n < \frac{n}{2}\right] &= \mathbb{P}\left[np - S_n > np - \frac{n}{2}\right] \le \mathbb{P}\left[|S_n - np| > n(p - 1/2)\right] \\ &\overset{\text{[Chebychev]}}{\le} \frac{\sigma_{S_n}^2}{(n(p-1/2))^2} = \frac{1}{n} \cdot \underbrace{\frac{p(1-p)}{(p-1/2)^2}}_{<\infty} \xrightarrow{n\to\infty} 0,\end{aligned}$$

und somit folgt, dass

$$\mathbb{E}[Y_n] = \mathbb{P}\left[S_n > \frac{n}{2}\right] = 1 - \mathbb{P}\left[S_n < \frac{n}{2}\right] \xrightarrow{n\to\infty} 1.$$

(d) Wie in Aufgabenteil (a) stellen wir fest, dass für $p \in [0,1]$

$$\sigma_{Y_n}^2 = \mathbb{E}[Y_n](1 - \mathbb{E}[Y_n]).$$

Da für $p \neq 1/2$, $\mathbb{E}[Y_n]$ gegen 0 oder 1 konverigert für $n \to \infty$ (siehe Aufgabenteil (c)), folgt direkt, dass $\sigma_{Y_n}^2 \to 0$ für $n \to \infty$.

Lösung 6.3

(a) Da p die gemeinsame Gewichtsfunktion/Verteilung von X und Y ist, muss gelten

$$\begin{aligned}1 &\overset{!}{=} \sum_{k=2}^{\infty}\sum_{j=1}^{k-1} \mathbb{P}[X = j, Y = k] = C\sum_{k=2}^{\infty}\sum_{j=1}^{k-1}\left(\frac{1}{2}\right)^k = C\sum_{j=1}^{\infty}\sum_{k=j+1}^{\infty}\left(\frac{1}{2}\right)^k \\ &= C\sum_{j=1}^{\infty}\left(\frac{1}{2}\right)^{j+1}\sum_{k=0}^{\infty}\left(\frac{1}{2}\right)^k = C\sum_{j=1}^{\infty}\frac{\left(\frac{1}{2}\right)^{j+1}}{1-\frac{1}{2}} = C\sum_{j=1}^{\infty}\left(\frac{1}{2}\right)^j = C\frac{\frac{1}{2}}{1-\frac{1}{2}} = C.\end{aligned}$$

Also ist $C = 1$.

(b) Wir berechnen für $j \ge 1$,

$$p_X(j) = \sum_{k=2}^{\infty} p(j,k) = \sum_{k=j+1}^{\infty}\left(\frac{1}{2}\right)^k = \left(\frac{1}{2}\right)^{j+1}\sum_{k=0}^{\infty}\left(\frac{1}{2}\right)^k = \left(\frac{1}{2}\right)^j$$

und für $k \ge 2$

$$p_Y(k) = \sum_{j=1}^{k-1} p(j,k) = (k-1)\left(\frac{1}{2}\right)^k.$$

(c) Die Zufallsvariablen X und Y sind genau dann unabhängig, wenn für alle $k \geq 2$ und für alle $j \geq 1$ gilt

$$p(j,k) = p_X(j) \cdot p_Y(k).$$

Dies ist nicht der Fall, z.B. gilt für $j = 1$ und $k = 2$

$$p(1,2) = \frac{1}{4} \neq \frac{1}{2} \cdot \frac{1}{4} = p_X(1) \cdot p_Y(2).$$

Lösung 6.4

(a) Mittels Proposition 5.4 muss für die Gewichtsfunktion der Randverteilung von X_1 gelten, dass für $x_1 \in W_1$,

$$p_{X_1}(x_1) = \sum_{x_2 \in W_2} p(x_1, x_2).$$

Nach Annahme gilt weiterhin $\sum_{x_1 \in W_1} p_{X_1}(x_1) = 1$. Da W_1 endlich oder abzählbar ist, können wir ohne Beschränkung der Allgemeinheit annehmen, dass $W_1 = \{w_1, w_2, \ldots\}$. Wir definieren

$$q_0 := 0 \quad \text{und für } j \geq 1, \quad q_j := \sum_{i=1}^{j} p_{X_1}(w_i)$$

und

$$X_1(\omega) := \sum_{j \geq 1} w_j \ \mathbb{1}_{U_1(\omega) \in (q_{j-1}, q_j]}.$$

Um zu überprüfen, dass X_1 die gewünschte Gewichtsfunktion/Verteilung hat, berechnen wir

$$\mathbb{P}[X_1 = x_1] = \sum_{j \geq 1} \underbrace{\mathbb{P}\big[U_1(\omega) \in (q_{j-1}, q_j]\big]}_{= q_j - q_{j-1} = p_{X_1}(w_j)} \mathbb{1}_{x_1 = w_j} = p_{X_1}(x_1),$$

was zu zeigen war.

(b) Wir konstruieren zunächst X_2 aus U_2 analog zu Aufgabenteil (a). Für die Gewichtsfunktion der Randverteilung von X_2 muss gelten, dass für $x_2 \in W_2$,

$$p_{X_2}(x_2) = \sum_{x_1 \in W_1} p(x_1, x_2).$$

Ohne Beschränkung der Allgemeinheit können wir annehmen, dass $W_2 = \{w'_1, w'_2, \ldots\}$. Wir definieren

$$q'_0 := 0 \quad \text{und für } j \geq 1, \quad q'_j := \sum_{i=1}^{j} p_{X_2}(w'_i)$$

und

$$X_2(\omega) := \sum_{j \geq 1} w'_j \ \mathbb{1}_{U_2(\omega) \in (q'_{j-1}, q'_j]}.$$

X_2 hat die gewünschte Gewichtsfunktion/Verteilung, da

$$\mathbb{P}[X_2 = x_2] = \sum_{j \geq 1} \underbrace{\mathbb{P}\big[U_2(\omega) \in (q'_{j-1}, q'_j]\big]}_{= q'_j - q'_{j-1} = p_{X_2}(w'_j)} \mathbb{1}_{x_2 = w'_j} = p_{X_2}(x_2).$$

Da die Zufallsvariablen U_1 und U_2 unabhängig sind, folgt dass die Zufallsvariablen $X_1(= \phi_1(U_1))$ und $X_2(= \phi_2(U_2))$ unabhängig sind (siehe Abschnitt 2.3 im Skript). Dies ist gemäss Proposition 5.6 äquivalent zu der Aussage, dass für die gemeinsame Gewichtsfunktion/Verteilung gilt

$$q(x_1, x_2) = p_{X_1}(x_1) \cdot p_{X_2}(x_2), \quad \forall x_1 \in W_1,\ x_2 \in W_2.$$

(c) Sei X_1 konstruiert wie in Aufgabenteil (a) beschrieben. Insbesondere gilt also für alle $x_1 \in W_1$,

$$\mathbb{P}[X_1 = x_1] = p_{X_1}(x_1)$$

und X_1 ist unabhängig von U_2.

Für die Konstruktion von X_2 nutzen wir die folgende Idee: Wir konstruieren zuerst X_1. Abhängig von dem Wert von X_1, sagen wir $X_1 = x_1$, konstruieren wir X_2 dann mithilfe der 'bedingten' Gewichtsfunktion/Verteilung von X_2, also $\frac{p(x_1,\cdot)}{p_{X_1}(x_1)}$.

Wie zuvor nehmen wir ohne Beschränkung der Allgemeinheit an, dass $W_2 = \{w'_1, w'_2, \dots\}$ und definieren für jedes $x_1 \in W_1$,

$$q_0^{(x_1)} := 0 \quad \text{und für } j \geq 1, \quad q_j^{(x_1)} := \sum_{i=1}^{j} \frac{p(x_1, w'_i)}{p_{X_1}(x_1)}.$$

Basierend auf U_2 und X_1 definieren wir X_2 nun wie folgt

$$X_2(\omega) := \sum_{x_1 \in W_1} \mathbb{1}_{X_1(\omega)=x_1} \sum_{j \geq 1} w'_j \, \mathbb{1}_{U_2(\omega) \in \left(q_{j-1}^{(x_1)}, q_j^{(x_1)}\right]}.$$

Wir zeigen nun, dass (X_1, X_2) die gewünschte gemeinsame Gewichtsfunktion/Verteilung hat: Für alle $x_1 \in W_1, x_2 \in W_2$ gilt

$$\begin{aligned}
\mathbb{P}[X_1 = x_1, X_2 = x_2] &= \mathbb{P}\left[\sum_{x_1 \in W_1} \mathbb{1}_{X_1=x_1} \sum_{j \geq 1} w'_j \, \mathbb{1}_{U_2 \in \left(q_{j-1}^{(x_1)}, q_j^{(x_1)}\right]} = x_2,\ X_1 = x_1\right] \\
&= \mathbb{P}\left[\sum_{j \geq 1} w'_j \, \mathbb{1}_{U_2 \in \left(q_{j-1}^{(x_1)}, q_j^{(x_1)}\right]} = x_2,\ X_1 = x_1\right] \\
&\overset{[X_1, U_2 \text{unabhängig}]}{=} \mathbb{P}\left[\sum_{j \geq 1} w'_j \, \mathbb{1}_{U_2 \in \left(q_{j-1}^{(x_1)}, q_j^{(x_1)}\right]} = x_2\right] \cdot \mathbb{P}[X_1 = x_1] \\
&= \left(\sum_{j \geq 1} \underbrace{\mathbb{P}\left[q_{j-1}^{(x_1)} < U_2 \leq q_j^{(x_1)}\right]}_{= q_j^{(x_1)} - q_{j-1}^{(x_1)} = \frac{p(x_1, w'_j)}{p_{X_1}(x_1)}} \mathbb{1}_{w'_j = x_2}\right) \cdot p_{X_1}(x_1) \\
&= \frac{p(x_1, x_2)}{p_{X_1}(x_1)} \cdot p_{X_1}(x_1) = p(x_1, x_2).
\end{aligned}$$