IPVS – Universität Stuttgart
Abteilung Scientific Computing
Prof. Dirk Pflüger
Stefan Zimmer

WS 2021/22

Statistische & Stochastische Grundlagen der Informatik
Übungsblatt 3

Abgabe (Aufgaben 3.1, 3.2, 3.4, 3.6) bis Montag 29.11., 14:00
Von den 16 möglichen Punkten müssen Sie wenigstens 8 erreichen.
Besprechung in den Übungen 01.12.–10.12.

Aufgaben zur KW 46 (ab 15.11.)

3.1 Erwartungswert Münzwurfspiel (3 Punkte)

Die Spieler A und B werfen eine Münze, bis „Zahl“ erscheint. Solange „Wappen“ erscheint, bekommt im 1., 3., 5.,… Wurf A jeweils 2 € von B und im 2., 4., 6.,… Wurf B jeweils 3 € von A.
Als Grundraum nehmen wir $\Omega := \mathbb{N}$ (ohne Null), das Ergebnis ω gebe an, in welchem Wurf „Zahl“ erscheint.

a) Geben Sie die Wahrscheinlichkeitsverteilung $\mathbb{P}$ an und berechnen Sie die Wahrscheinlichkeit, dass im Zug k noch eine Auszahlung stattfindet (also noch nicht „Zahl“ geworfen wurde).

b) Nun beschreibe

- für ungerades i die Zufallsvariable X_i den Betrag, den B in diesem Zug an A zahlt (je nach Ergebnis ω kann ja in Zug i noch eine Auszahlung stattfinden oder nicht),
- die Zufallsvariable X die Summe aller X_i, also den Betrag, den B insgesamt an A zahlt.

Berechnen Sie $\mathbb{E}(X_i)$ und $\mathbb{E}(X)$.

Und was ist der Erwartungswert für den Betrag, den (in den geradzahligen Zügen) A an B zahlt?

c) Was wäre der faire Einsatz für das Spiel (ein fester Betrag, den vor dem Spiel ein Spieler an den anderen zahlt, um die unterschiedlichen Gewinnerwartungen auszugleichen)? Könnte man die Auszahlungen (2 € bzw. 3 €) anpassen, damit das Spiel ohne Einsatz fair ist?

Aufgabe 3.1

> A bekommt 2€ bei ungeraden Zahlen
> B bekommt 3€ bei geraden Zahlen
> ω: "Nummer des Zahl Wurfes", $\omega \in \mathbb{N}$
> $\Omega := \mathbb{N}$, $\omega \in \Omega$

a)

Wurf i	1	2	3	4	$\dots$	k
$p(i)$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$		$\frac{1}{2^k}$

$\Rightarrow \mathbb{P}(\{k\}) = \frac{1}{2^k}$

$\Rightarrow p(k) = \frac{1}{2^k}$ "im k-ten Zug Wappen"

b)

$$X_i(\omega) = \begin{cases} 2€, & \text{falls } i < \omega \\ 0€, & \text{falls } i = \omega \end{cases} \qquad (X_i : \Omega \to \{0, 2\})$$

$$X(k) = \begin{cases} -1€ \cdot \lfloor \frac{k}{2} \rfloor + 2€, & \text{falls } k\text{-ter Wurf Wappen} \\ -1€ \cdot \lfloor \frac{k}{2} \rfloor, & \text{falls } k\text{-ter Wurf Zahl} \end{cases}$$

Beachte $X_{\text{ungerade}} = \sum_i X_{i,\text{ungerade}}$, $X_{\text{gerade}} = \sum_i X_{i,\text{gerade}}$

$$\Rightarrow \mathbb{E}(X_{\text{ungerade}}) = \sum_{i=0}^{\infty} 2€ \cdot 2^{-(2i+1)}$$
$$= 2€ \cdot \sum_{i=0}^{\infty} 2^{-(2i+1)}$$
$$= 2€ \cdot \frac{1}{2} \cdot \sum_{i=0}^{\infty} \left(\frac{1}{2}\right)^{2i}$$
$$= 1€ \cdot \sum_{i=0}^{\infty} \left(\frac{1}{4}\right)^{i} \qquad \text{geometr. Reihe: } \sum_{k=0}^{\infty} q^k = \frac{1}{1-q}$$
$$= 1€ \cdot \frac{1}{1-\frac{1}{4}}$$
$$= 1€ \cdot \frac{4}{3}$$
$$= \frac{4}{3}€$$

– ist dem Sichtpunkt geschuldet: A zahlt an B

$$\Rightarrow \mathbb{E}(X_{gerade}) = -\sum_{j=0}^{\infty} 3€ \cdot 2^{-(2j+2)}$$

$$= -3€ \cdot \frac{1}{4} \cdot \sum_{j=0}^{\infty} 2^{-2j}$$

geometr. Reihe: $\sum_{k=0}^{\infty} q^k = \frac{1}{1-q}$

$$= -\frac{3}{4}€ \cdot \underbrace{\frac{1}{1-\frac{1}{4}}}_{\frac{4}{3}}$$

$$= -1\,€$$

mit $X = \sum_i X_i$

$$= \sum_i X_{i,gerade} + X_{i,ungerade}$$

und $\mathbb{E}(X) = \mathbb{E}(\sum_i X_i)$

$$= \mathbb{E}\left(\sum_i X_{i,gerade} + X_{i,ungerade}\right)$$

Linearität d. Erwartungswertes:

$$= \mathbb{E}(X_{gerade}) + \mathbb{E}(X_{ungerade})$$

$$= -1€ + \frac{4}{3}€$$

$$= \frac{1}{3}€$$

c) Ein fairer Einsatz müsste $\mathbb{E}(X_{gerade})$ und $\mathbb{E}(X_{ungerade})$ ausgleichen.

$$\Rightarrow \mathbb{E}(X) \overset{!}{=} 0$$

Ein Beispiel wäre 2€ und 4€:

$\mathbb{E}(X_{ungerade}) = \frac{4}{3}€$ (s.o.)

$$\mathbb{E}(X_{gerade}) = -\sum_{j=0}^{\infty} 4€ \cdot 2^{-(2j+2)}$$

$$= -4€ \cdot \frac{1}{4} \cdot \sum_{j=0}^{\infty} 2^{-2j}$$

$$= -\frac{4}{4}€ \cdot \frac{1}{1-\frac{1}{4}}$$

$$= -\frac{4}{3}€$$

3.2 Geometrisch verteilte Zufallsvariablen (3 Punkte)

Seien $X_1 \sim \mathrm{G}(p_1)$ und $X_2 \sim \mathrm{G}(p_2)$ unabhängige Zufallsvariablen mit geometrischer Verteilung mit Parameter (Trefferwahrscheinlichkeit) p_1 bzw. p_2: $\mathbb{P}(X_i = k) = p_i q_i^k$ mit $q_i := 1 - p_i$.

a) Zeigen Sie, dass die Zufallsvariable $X_3 := \min(X_1, X_2)$ ebenfalls geometrisch verteilt ist. Wie lautet die zugehörige Trefferwahrscheinlichkeit p_3?

   Die folgenden Schritte sind dazu hilfreich (und es ist zweckmäßig, ganz in den Misserfolgswahrscheinlichkeiten q_i zu rechnen, das gibt einfachere Formeln als in den p_i):

   - Wie lauten die Verteilungsfunktionen von X_1 und X_2 bzw., wie groß sind $\mathbb{P}(X_i > k)$ $(k \in \mathbb{N}, i = 1, 2)$?
   - Und wie groß $\mathbb{P}(\min(X_1, X_2) > k)$?
   - Wie sieht die Verteilungsfunktion von $X_3 := \min(X_1, X_2)$ aus?

b) Begründen Sie, dass das Ergebnis plausibel ist!

c) Jemand möchte mit Ihnen folgendes Spiel spielen: Es wird eine Münze geworfen, bis „Zahl“ eintritt, die Zahl der benötigten Versuche wird notiert. Dann wird gewürfelt, bis eine Sechs gewürfelt wurde, die Zahl der benötigten Versuche wird ebenfalls notiert. Das Gesamtergebnis ist die kleinere der beiden Zahlen. Beispiel: Sie werfen „Wappen“, „Wappen“, „Zahl“ und würfeln 4,1,1,3,6. Das Ergebnis ist 3 als die kleinere Zahl von 3 und 5.

   Das ist Ihnen zu kompliziert – Erfinden Sie einen einfacheren Mechanismus, der Ergebnisse mit der gleichen Verteilung wie beim vorliegenden Spiel produziert.

3.3 Poisson-Verteilung, bedingte Wahrscheinlichkeiten (keine Abgabe)

An einem Studienprojekt mit Programmierung in `C++` arbeiten zwei Studenten, Ano und Nym. Da sie zur Speicherverwaltung viel mit `new` und `delete` arbeiten, letzteres aber oft vergessen, produzieren sie eine Menge Memory-Leaks. Andere Fehler machen sie aber nicht. Die Arbeit ist so aufgeteilt, dass jedes Modul von einem der beiden ausprogrammiert wird.
Die Anzahl der produzierten Memory-Leaks pro Modul sei bei Ano Poisson-verteilt mit Parameter $\lambda_1 = 1$, bei Nym Poisson-verteilt mit Parameter $\lambda_2 = 10$.
Allerdings macht Ano nur deshalb im Mittel weniger Fehler, da er viel langsamer programmiert. Daher stammen $\frac{3}{4}$ von den entwickelten Modulen von Nym, nur $\frac{1}{4}$ von Ano.

a) Ein Betreuer schaut in den Code eines Moduls. Was ist die Wahrscheinlichkeit, dass das Modul genau ein Memory-Leak enthält?

b) Entzückt findet der Betreuer endlich ein völlig fehlerfreies Modul. Was ist jetzt die Wahrscheinlichkeit, dass Nym das Modul geschrieben hat?

c) Welchen Anteil $q \in [0, 1]$ am Code müsste Nym geschrieben haben, sodass fehlerlose Module mit gleicher Wahrscheinlichkeit von Ano oder Nym geschrieben wurden?

d) Wie viele Fehler pro Modul darf der Betreuer im Mittel erwarten?

Aufgabe 3.2

$$\mathbb{P}(X_1 = k) = p_1 \cdot (1-p_1)^k = p_1 q_1^k \qquad (1)$$

$$\mathbb{P}(X_2 = k) = p_2 \cdot (1-p_2)^k = p_2 q_2^k \qquad (2)$$

"W'keit k Misserfolge vor dem ersten Erfolg zu haben"

a) $$F^{X_1}(k) = \mathbb{P}(X_1 \leq k) = \sum_{j=0}^{k} \mathbb{P}(X_1 = j)$$
$$= p_1 \sum_{j=0}^{k} q_1^j$$
$$= p_1 \frac{1-q_1^{k+1}}{1-q_1}$$
$$= 1 - q_1^{k+1}$$

$$\Rightarrow F^{X_2}(k) = \mathbb{P}(X_2 \leq k) = 1 - q_2^{k+1}$$

$$\Rightarrow \mathbb{P}(X_1 > k) = 1 - \mathbb{P}(X_1 \leq k)$$
$$= q_1^{k+1}$$

$$\mathbb{P}(X_2 > k) = q_2^{k+1}$$

vgl

$$\Rightarrow \mathbb{P}(\min(X_1, X_2) > k) = [\min(q_1, q_2)]^{k+1} = (1 - \max(p_1, p_2))^{k+1}$$

$$F^{X_3}(k) = \mathbb{P}(X_3 \leq k)$$
$$= \mathbb{P}(\min(X_1, X_2) \leq k)$$
$$= 1 - \mathbb{P}(\min(X_1, X_2) > k)$$
$$= 1 - (\min(q_1, q_2))^{k+1}$$

$$\mathbb{P}(X_3 \leq k) = 1 - (\min(q_1, q_2))^{k+1}$$
$$= 1 - q_3^{k+1} \quad , \text{mit } p_3 = \max(p_1, p_2)$$
$$= p_3 \frac{1-q_3^{k+1}}{1-q_3}$$
$$= p_3 \sum_{j=0}^{k} q_3^j$$

$$= \sum_{j=0}^{k} p_3 q_3^j$$

$$= \sum_{j=0}^{k} \mathbb{P}(X_3 = j)$$

$$\Rightarrow \quad \mathbb{P}(X_3 = k) = p_3 q_3^k$$

$$\Leftrightarrow \quad \mathbb{P}(\min(X_1, X_2) = k) = \max(p_1, p_2) \cdot (\min(q_1, q_2))^k$$

$\Rightarrow$ X_3 ist geometrisch verteilt.

b)

Da die Minimumfunktion entweder X_1 oder X_2 wählt und X_1 wie auch X_2 geometrisch verteilt sind, muss also auch X_3 geometrisch verteilt sein.

c)

$X :=$ "Münze werfen bis Zahl"

$Y :=$ "Würfel werfen bis 6"

$$\mathbb{P}(X = k) = p_x \cdot q_x^k = \left(\frac{1}{2}\right)^{k+1}$$

$$\mathbb{P}(Y = k) = p_y \cdot q_y^k = \frac{1}{6} \cdot \left(\frac{5}{6}\right)^k = \frac{5^k}{6^{k+1}}$$

mit $Z = \min(X, Y)$ folgt aus a)

$$\mathbb{P}(Z = k) = \max(p_x, p_y) \cdot (\min(q_x, q_y))^k$$
$$= \frac{1}{2} \cdot \left(\frac{1}{2}\right)^k$$
$$= \left(\frac{1}{2}\right)^{k+1}$$

$\Rightarrow$ Nur X, also das Münzwurfexperiment ist relevant.

3.4 Würfelsumme (6 Punkte)

Im Kapitel über das Gesetz der großen Zahlen hatten wir gesehen, dass die Tschebyscheff-Ungleichung für den Mittelwert M_n aus n Würfen eines fairen Würfels besagt, dass für $0 < p \leq 1$ gilt $\mathbb{P}(|M_n - \mu| \geq \epsilon) \leq p$ mit $\epsilon := \sqrt{\sigma^2/(np)}$, $\mu = 7/2$ und $\sigma^2 = 35/12$. Wir setzen ab jetzt wieder $p = 1/10$, dann entsprechen die Grenzen $\mu \pm \epsilon$ gerade den schwarzen Punkten des Diagramms im Skript. In dieser Aufgabe wollen wir ϵ durch eine schärfere (d.h., kleinere) Schranke ersetzen. Dazu ist es praktisch, nicht im Mittelwert, sondern in der Summe W_n der Augenzahlen zu rechnen – es ist also einfach $W_n = n \cdot M_n$ und das Ereignis $\{|M_n - \mu| \geq \epsilon\}$ lässt sich sehr leicht in ein Ereignis in W_n umschreiben. $W_n = n \cdot M_n \quad , M_n = \frac{1}{n} \sum_{j=1}^{n} X_j$

a) Berechnen Sie für $n = 1, \ldots, 40$ die Verteilungen von W_n, also die $\mathbb{P}(W_n = k)$ für alle Werte k, die W_n annehmen kann. min : 40 max : 240

Halten Sie sich dabei nicht damit auf, eine geschlossene Formel dafür zu finden, sondern lassen Sie Ihren Rechner arbeiten: Der Fall $n = 1$ ist einfach und für $n > 1$ gilt

W'keit für Augensumme k

$$\mathbb{P}(W_n = k) = \frac{1}{6} \sum_{j=1}^{6} \mathbb{P}(W_{n-1} = k - j)$$

$n = 5 \; , \Rightarrow k = \{5, \ldots, 30\}$

Speichern Sie die $\mathbb{P}(W_n = k)$ z.B. in einer Hashtabelle mit n und k als Schlüsseln.

Abzugeben sind von dieser Teilaufgabe zwei Bilder: Verteilung und Verteilungsfunktion von W_5.

b) Erweitern Sie Ihr Programm so, dass es für $n = 2, \ldots, 40$ das größte $k^*(n) \in \mathbb{Z}$ bestimmt mit $\mathbb{P}(W_n \leq k^*(n)) \leq p/2$.

Wegen der Symmetrie der Verteilung und $W_n = n \cdot M_n$ gilt nun $\mathbb{P}(|M_n - \mu| \geq \tilde{\epsilon}) \leq p$ mit $\tilde{\epsilon} := \mu - k^*(n)/n$.

Geben Sie ein Diagramm ab, das für $n = 2, \ldots, 40$ das $\tilde{\epsilon}$ und das ϵ aus der Tschebyscheff-Ungleichung darstellt.

3.5 Binomialverteilung, zentraler Grenzwertsatz (keine Abgabe)

Die Sprossen einer in beiden Richtungen unendlichen Leiter seien mit den ganzen Zahlen nummeriert. Ein Frosch sitze anfangs auf Sprosse 0. Nun steigt er

- an allen Tagen, an denen er gut gelaunt ist, eine Sprosse hoch,
- an den anderen Tagen eine Sprosse hinab (am Tag 1 bewegt er sich bereits).

Die Laune des Froschs am Tag i werde beschrieben durch unabhängige Zufallsvariablen

$$X_i \sim \text{Bin}(1, p)$$

(mit festem Parameter $0 < p < 1$): $X_i = 1$ bedeute „gute Laune".

a) Geben Sie an, wie sich aus den X_i die folgenden Zufallsvariablen ergeben und berechnen Sie jeweils Erwartungswert und Varianz (alles in Abhängigkeit von p):

(i) Y_n mit Werten ± 1 gebe die Höhenveränderung im Tag n an

(ii) Z_n mit Werten aus $\mathbb{Z}$ gebe die Sprosse an, die er am Tag n erreicht.

b) Bestimmen Sie $\alpha_n > 0$ und β_n so, dass $U_n := \alpha_n(Z_n - \beta_n)$ Erwartungswert 0 und Varianz 1 hat.

$X_j := 1\!\!1_A \qquad A \subset \Omega$

$\Omega = \{1, 2, 3, 4, 5, 6\}$

$\mu = \mathbb{P}(A)$

$\mu = \frac{7}{2}$

$\Rightarrow 1\!\!1_A(\omega) = \begin{cases} 1 & , \omega \in A \\ 0 & \text{sonst} \end{cases}$

$W_n = n \cdot M_n$

$M_n = \frac{1}{n} \sum_{j=1}^{n} X_j$

$\Leftrightarrow W_n = \sum_{j=1}^{n} X_j$

$W_1 :=$ "1x Würfeln"

$= \frac{1}{6} \quad \frac{1}{6}$

$1 - 6$

Aufgabe zur KW 47 (ab 22.11.)

3.6 Stetige Verteilung (4 Punkte)

a) Für eine stetige Zufallsvariable X sei die Dichte

$$f^X : \mathbb{R} \to \mathbb{R}, \quad x \mapsto \begin{cases} \frac{x}{2} - \frac{1}{2} & 1 \leq x \leq 3, \\ 0 & \text{sonst.} \end{cases}$$

Geben Sie die Verteilungsfunktion F^X von X an und berechnen Sie Erwartungswert $\mathbb{E}(X)$ und Varianz $\mathbb{V}(X)$.

b) Es sei $X \sim \mathcal{U}(0,1)$ eine Zufallsvariable mit Grundraum Ω und $n \in \mathbb{N}$, $n > 0$. Geben Sie ein $g : \mathbb{R} \to \mathbb{R}$ an, so dass mit der Zufallsvariable $Y := g(X)$ (zur Notation vgl. Skript, Folgerung 3.10) gilt, dass $(Y(\Omega), \mathbb{P}^Y)$ ein Laplace-Experiment mit n Ergebnissen ist ($\mathbb{P}^Y$ nach Skript, Gl. (3.15)).

Aufgabe 3.6

a) Für eine stetige Zufallsvariable X sei die Dichte

$$f^X : \mathbb{R} \to \mathbb{R}, \quad x \mapsto \begin{cases} \frac{x}{2} - \frac{1}{2} & 1 \le x \le 3, \\ 0 & \text{sonst.} \end{cases}$$

Geben Sie die Verteilungsfunktion F^X von X an und berechnen Sie Erwartungswert $\mathbb{E}(X)$ und Varianz $\mathbb{V}(X)$.

a)

$$F^X = \mathbb{P}(X \le x)$$

$$\Rightarrow P(X \le x) = \begin{cases} 0 & , x < 1 \\ \int_{-\infty}^{x} f^X(t)\, dt & , 1 \le x \le 3 \\ 1 & , x > 3 \end{cases}$$

Betrachte $1 \le x \le 3$:

$$\Rightarrow \int_{-\infty}^{x} f^X(x)\, dx = \int_{-\infty}^{x} \left(\frac{x}{2} - \frac{1}{2} \right) dx$$

$$= \left[\frac{x^2}{4} - \frac{x}{2} + c \right]$$

Dabei muss für die Verteilungsfunktion gelten, dass

$$\lim_{x \to -\infty} F^X = 0 \quad \text{und} \quad \lim_{x \to \infty} F^X = 1$$

$$\Rightarrow \frac{x^2}{4} - \frac{x}{2} + c \overset{!}{=} 1 \quad \text{für } x = 3$$

$$\Rightarrow \frac{9}{4} - \frac{3}{2} + c = 1$$

$$\Leftrightarrow \frac{3}{4} + c = 1$$

$$\Leftrightarrow c = \frac{1}{4}$$

Prüfe ob das an der linken Grenze ebenfalls gilt:

$$\Rightarrow \frac{x^2}{4} - \frac{x}{2} + \frac{1}{4} \overset{!}{=} 0 \quad \text{für } x = 1$$

$$\Rightarrow \frac{1}{4} - \frac{1}{2} + \frac{1}{4} = 0$$

$$0 = 0 \quad \checkmark$$

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x \cdot f(x)\, dx$$

$$\Rightarrow \int_1^3 x \cdot \left(\frac{x}{2} - \frac{1}{2}\right) dx$$

$$= \int_1^3 \left(\frac{x^2}{2} - \frac{x}{2}\right) dx$$

$$= \left[\frac{x^3}{6} - \frac{x^2}{4}\right]_1^3$$

$$= \left(\frac{27}{6} - \frac{9}{4}\right) - \left(\frac{1}{6} - \frac{1}{4}\right)$$

$$= \frac{26}{6} - \frac{8}{4} = \frac{26}{6} - 2 = \frac{14}{6} = \frac{7}{3}$$

$$\Rightarrow \mathbb{E}(X) = \begin{cases} \frac{7}{3} & , 1 \leq x \leq 3 \\ 0 & , \text{sonst} \end{cases}$$

$$\mathbb{V}(X) = \int_{-\infty}^{\infty} (x - \mathbb{E}(x))^2 \cdot f(x)\, dx$$

$$= \int_1^3 \left(x - \frac{7}{3}\right)^2 \cdot \left(\frac{x}{2} - \frac{1}{2}\right) dx$$

$$= \int_1^3 \left(x^2 - \frac{14}{3}x + \frac{49}{9}\right) \cdot \left(\frac{x}{2} - \frac{1}{2}\right) dx$$

$$= \int_1^3 \left(\frac{x^3}{2} - \frac{7}{3}x^2 + \frac{49}{18}x - \frac{x^2}{2} + \frac{7}{3}x - \frac{49}{18}\right) dx$$

$$= \int_1^3 \left(\frac{x^3}{2} - \frac{17}{6}x^2 + \frac{91}{18}x - \frac{49}{18}\right) dx$$

$$= \left[\frac{x^4}{8} - \frac{17}{18}x^3 + \frac{91}{36}x^2 - \frac{49}{18}x\right]_1^3$$

$$= -\frac{19}{24} - \left(-\frac{73}{72}\right) = \frac{2}{9}$$

b)

$$\mathbb{P}^Y : \mathcal{P}(Y(\Omega)) \to \mathbb{R}$$

$$\mathbb{P}^Y(B)$$