Stochastik-Praktikum **WS 2023-2024**

Prof. Dr. Matthias Birkner

Jan Lukas Igelbrink

Blatt 9

Aufgabe 1 (3+5 = 8 Punkte)

X sei eine Zufallsvariable mit Verteilungsfunktion

$$F(x) = 1 - x^{-\frac{3}{2}} \quad (\text{für } x \geq 1, \text{ und } F(x) = 0 \text{ für } x < 1) \tag{1}$$

a) Implementieren Sie eine Funktion, die als Argument eine Anzahl n nimmt, und n Samples einer Zufallsvariable mit Verteilungsfunktion F zurückgibt.

b) Untersuchen Sie, ob die Verteilung von X dem zentralen Grenzwertsatz genügt: Es seien $X_1, \dots, X_n$ unabhängige Kopien von X. Betrachten Sie k Samples von (es gilt $\mathbb{E}[X] = 3$)

$$S_n := \frac{1}{\sqrt{n}} \sum_{\ell=1}^{n} \big(X_\ell - \mathbb{E}[X]\big) \tag{2}$$

und illustrieren Sie die empirische Verteilung von S_n. Können Sie die Verteilung mit einer Normalverteilung vergleichen? Wählen Sie verschiedene k und n, beschriften Sie die Plots sinnvoll, und wählen Sie selbst mindestens drei sinnvolle Paare k und n.

Aufgabe 2 (3/2 + 3/2 + 3/2 + 3/2 = 6 Punkte)

In dieser Aufgabe wollen wir die Grenzen der *Gesetze großer Zahlen* anhand der Beispiele eines kleinen Artikels von Nancy Geller aus den 70er Jahren, siehe *Some Examples of the Weak and Strong Laws of Large Numbers for Averages of Mutually Independent Random Variables, The American Statistician, Vol. 32, 1978* (online zugänglich unter `https://www.jstor.org/stable/2683474`) kennenlernen. (Beachten Sie dazu, dass Theorem 3 in der zitierten Arbeit aus Fellers bekanntem Werk *An introduction to probability theory and its applications, Vol. II* falsch zitiert wurde, siehe dazu `https://www.stochastik.mathematik.uni-mainz.de/files/2023/12/notes_lln.pdf`. Dies soll Sie hier allerdings nicht weiter beschäftigen, außer Ihr Interesse sollte damit geweckt sein.) Veranschaulichen Sie, ob oder ob nicht ein *Gesetz der großen Zahlen* für die Beispiele

a) Beispiel (1a) für $\delta = \frac{1}{2}$,

b) Beispiel (1b) für $\delta = \frac{3}{2}$,

c) Beispiel (2a) für $\delta = \frac{1}{4}$,

d) Beispiel (2b) für $\delta = \frac{1}{5}$

aus der zitierten Arbeit von Nancy Geller gilt.

In `https://www.stochastik.mathematik.uni-mainz.de/files/2023/12/notes_lln.pdf` finden Sie Aufgabenteil **b)** bereits beispielhaft gelöst, siehe Abbildung 1.

Aufgabe 3 (1 + 1 + 2 = 4 Punkte)
Laden Sie die aktuelle Excel-Tabelle zur Weinmosternte von 1950-2022 in RLP unter
`https://www.statistik.rlp.de/de/wirtschaftsbereiche/landwirtschaft/zeitreihen-land/tabelle-10/`
herunter.

a) Berechnen Sie den empirischen Mittelwert und die empirische Standardabweichung der drei Zeitreihen zum Hektarertrag.

b) Generieren Sie Boxplots zu den Zeitreihen von Weiß- und Rotmost in einer Graphik. Erstellen Sie eine Graphik mit Boxplots für die Hektarerträge von Weiß- und Rotmost. Interpretieren Sie die Graphik knapp in Hinblick auf den Mittelwert und die Streuung der Erträge. Welcher Typ Information geht in Boxplots verloren?

c) Generieren Sie einen Scatterplot, in dem Sie Weiß- gegen Rotmostertrag für die verschiedenen beobachteten Jahre plotten.

d) Veranschaulichen Sie die Zeitreihen zum Hektarertrag für Weißmost, Rotmost und deren Summe aussagekräftig, zum Beispiels mittels eines Flächendiagrammes. (Ein schönes Beispiel sehen Sie unter `https://coinmarketcap.com/charts/` mit dem Titel *Bitcoin dominance*.)
Hinweis: Sie finden diesen Typ Chart auch als `Stacked Area Graph`. Das Paket `ggplot2` kann hilfreich sein.

Denken Sie daran, all Ihre Plots sinnvoll und aussagekräftig zu beschriften.