

## Viertes Übungsblatt - Transformation, Wahrscheinlichkeitsverteilungen

Abgabe spätestens am **19.11.2022, 16:00 Uhr MEZ**. Voraussetzungen: Die Vorlesungsinhalte zu Transformationen von Zufallsvariablen (VL05, 2.11.2022), Gleich-, Binomialverteilung (VL06, 7.11.2022), Normal-, Poisson-, Exponentialverteilung (VL07, 9.11.2022). In den mit \* gekennzeichneten Aufgaben kannst du gerne auch mit R (oder einer anderen Programmiersprache) arbeiten. Bitte gib dann jeweils als Teil der Lösung auch deinen wohldokumentierten Code mit ab.

### Aufgabe 4.1 - Transformation bekannter Verteilungen [(a-c) 4 + 1 BP, (d) 3 P]

Wende die Regeln zur Transformation aus der VL korrekt an. [Hinweis: dazu gehört auch die Überprüfung aller Voraussetzungen!]

- a) Welche Dichtefunktionen haben Zufallsvariablen (ZV), die durch die Transformationen (I)  $f(x) = -5x$ , (II)  $f(x) = x + 3$ , (III - Bonusaufgabe - 1 BP)  $f(x) = x^2$  einer ZV entstehen, deren Verteilung eine stetige Gleichverteilung auf dem Intervall  $[-3, -1] \subset \mathbb{R}$  aufweist?
- b) Welche Dichtefunktionen haben ZV, die durch die Transformationen (I)  $f(x) = -5x$  bzw. (II)  $f(x) = x + 3$  einer ZV mit Normalverteilung mit  $\mu = 2$ ,  $\sigma = 4$  entstehen?
- c) Welche Dichtefunktionen haben ZV, die durch die Transformationen (I)  $f(x) = 5x$  bzw. (II)  $f(x) = x + 3$  einer Zufallsvariablen mit Exponentialverteilung mit  $\lambda = 2$  entstehen?
- d\*) Skizziere die in a) - c) erhaltenen Dichtefunktionen (außer a) III) der transformierten ZV. Welche der transformierten ZV haben wieder aus der VL bekannte Verteilungen und welche (und mit welchen Parametern)?

### Aufgabe 4.2 - Modellierung in diskreten Wahrscheinlichkeitsräumen [Tutorium]

30% der Bewerbenden beim offenen Call for Talents eines IT-Konzerns verfügen über fortgeschrittene Programmierkenntnisse. Bewerbende führen unabhängig voneinander Vorstellungsgespräche und sind zufällig aus einer Bewerberliste gewählt. Die Zufallsvariable  $X$  sei so definiert, dass genau  $X = k$  Bewerbende in  $n$  Gesprächen fortgeschrittene Kenntnisse besitzen. Hinweis: Zur Modellierung darf angenommen werden, dass die Liste der Bewerbenden unendlich lang ist.

- a) Wie ist  $X$  verteilt? Begründe.
- b) Wie groß ist die Wahrscheinlichkeit, dass genau im dritten Gespräch das erste Mal eine Person mit fortgeschrittenen Programmierkenntnissen gefunden wurde?
- c\*) Wie groß ist die Wahrscheinlichkeit, dass bei  $n = 300$  Gesprächen insgesamt mindestens 80 Personen mit fortgeschrittenen Programmierkenntnissen gefunden wurden?

- d)  $p \in (0,1)$  beschreibe die Wahrscheinlichkeit, dass ein zufälliger Bewerbender über fortgeschrittene Programmierkenntnisse verfügt. Welches  $p \in (0,1)$  maximiert die Wahrscheinlichkeit, dass die Firma genau eine Person mit fortgeschrittenen Programmierkenntnissen in drei Vorstellungsgesprächen findet?

### Aufgabe 4.3 - Modellierung mit einer Normalverteilung [2.5 P + 1 BP]

Für die Quarantäne-Käfige einer Auffangstation für Löwen eines Nationalparks soll berücksichtigt werden, wie groß potentielle Bewohner der Käfige (d.h. Löwen) sind. Damit soll sichergestellt werden, dass sich nicht zu viele Löwen in den Käfigen und mit den zur Verfügung gestellten Spielzeugen unwohl fühlen, aber auch nicht zu viel Baumaterial benutzt werden muss. Hierbei wird auf große Datensammlungen zurückgegriffen, aus denen hervorgeht, dass bei erwachsenen Tieren die Schulterhöhe einer Löwin einer Normalverteilung mit  $\mu = 85\text{cm}$  und  $\sigma = 6\text{cm}$  folgt und die Schulterhöhe eines männlichen Löwen einer Normalverteilung mit  $\mu = 98\text{cm}$  und  $\sigma = 10\text{cm}$ .

- a) Analysiere, ob die Normalverteilungsannahme hier plausibel ist.
- b\*) Wie groß ist die Wahrscheinlichkeit, dass die Schulterhöhe einer Löwin größer als 95cm ist und wie groß ist die Wahrscheinlichkeit, dass die Schulterhöhe eines männlichen Löwen größer als 110cm ist?
- c\*) Die Käfige sollen so ausgelegt werden, dass sich Tiere mit einer Schulterhöhe zwischen 78cm und 108cm wohl fühlen. Ist die Wahrscheinlichkeit größer, dass ein weiblicher Löwe oder ein männlicher Löwe sich nicht in diesem Intervall befindet? (Mit Begründung!)
- d\*) [Bonusaufgabe - 1 BP] Die Auffangstation hat 30 Gehege für einzelne Löwen zur Quarantäne. Wie groß ist die Wahrscheinlichkeit, dass bei voller Besetzung der Käfige alle Löwen und Löwinnen eine Schulterhöhe zwischen 78cm und 108cm haben? Nimm dabei an, dass sich die Käfige genau gleich unter den Geschlechtern aufteilen (und die Schulterhöhen unabhängig voneinander sind, d.h. sich die WK für mehrere Käfige als Produkt der WK für die einzelnen Käfige ergibt). Interpretiere das Ergebnis.
- e\*) Bestimme eine Schulterhöhe, bei der die Dichtefunktion für Löwenmännchen und Löwinnen den gleichen Wert annimmt.

### Aufgabe 4.4 - Modellierung von Verteilungen [ $2 \times 1.5$ P + 1.5 BP]

Es liegen drei Datensätze vor:

1. ulfs\_cafe.csv: Anzahl der Studierenden, die an einem Wochentag zwischen 14 und 15 Uhr Ulf's Café betreten. Zusammenfassung der Beobachtungen von 875 Tagen.
  2. bewohner\_grossdorf.csv: Körpergröße der 5000 Bewohner von Grossdorf
  3. [Bonusdatensatz - 1.5 BP] s.bahn.csv: Wartezeit auf die S-Bahn, wenn diese nachts nur noch stündlich abfährt, Zusammenfassung für 100 Beobachtungen
- a\*) Veranschauliche jeden Datensatz graphisch, so dass die unterliegende Verteilung der im Text beschriebenen Messgröße sinnvoll dargestellt ist. Tipp: Generiere für Datensatz 2

und 3 eine ähnliche Zusammenfassung der Daten, wie für 1 schon gegeben ist, aber mit passenden Intervallen statt einzelner Werte!

- b) Überlege für jeden Datensatz, welche Verteilung die Daten am besten beschreibt und begründe jeweils deine Entscheidung.
- c\*) Visualisiere für jeden der Datensätze mindestens drei Parametrisierungen der in b) ausgewählten Verteilungen und beschreibe, welche der drei Parametrisierungen graphisch am besten geeignet erscheint.