

Übungsblatt 7 Statistik

Aufgabe 7.1 Kennzahlen von Stichproben

Gegeben sei die folgende Stichprobe Z von (bereits sortierten) Messwerten:

z	0.1	0.15	0.2	0.2	0.25	0.3	0.31	0.35
	0.7	1.0	1.0	1.1	1.5	2.5	2.5	2.7
	2.8	2.95	3.1	3.1	3.15	3.15	3.2	4.0
	4.2	5.2	7.5	10.5	10.7	20.0	25.5	31.5

-
- Berechnen Sie Mittelwert, Median und Standardabweichung für diese Stichprobe.
- Wie ändern sich Mittelwert, Median und Standardabweichung, wenn statt $z=0.25$ ein Wert $z=100.25$ gemessen wird?
- Können Sie eine Kennzahl angeben, die wie die Standardabweichung die Breite einer Verteilung misst, aber durch den einzelnen Ausreißer 100.25 kaum gestört wird?
- Zeichnen Sie den Boxplot für Z (Geben Sie dabei die notwendigen Berechnungsschritte an)
- Zeichnen Sie den Boxplot für die gemäß (c) abgeänderte Stichprobe Z' ! Wie ändern sich Box und Whisker?

Aufgabe 7.2 Kombinatorik I

- Wie viele Wörter über dem Alphabet $\{v,w,x,y,z\}$ mit 10 Buchstaben gibt es, die genau 8 z's enthalten? Wie viele Wörter enthalten mehr als 7 z's?
- 5 Würfel werden gleichzeitig geworfen. Wie viele Wurfresultate gibt es?
- Wie groß ist die Wahrscheinlichkeit, mit 3 Würfeln mindestens 2 Fünfer zu erzielen?
- Ein Würfel wird fünfmal geworfen. Wie viele Resultate gibt es? Wie wahrscheinlich ist es, mit einer Eins zu beginnen und einer Fünf zu enden?

Aufgabe 7.3 Kombinatorik II

- Bei der geheimen Wahl zum Dekan treten drei Kandidaten D, E, F an, die von 9 Stimmen gewählt werden. Wie viele Wahlausgänge gibt es insgesamt? Wie viele mit der absoluten Mehrheit für Kandidat E? Wie viele mit einem "Patt"?
- Bei einem Hunderennen mit 9 Hunden sind auf dem Wettschein Platz 1., 2., 3. und 4. zu notieren. Wie groß ist die Wahrscheinlichkeit, durch zufälliges Raten die richtigen Namen (in richtiger oder falscher Reihenfolge) auf dem Wettschein zu haben UND auf Platz 2. auch genau den richtigen Hundenamen stehen zu haben?

[Zu (a): "Patt" = kein Kandidat kann sich durchsetzen]

Aufgabe 7.4 (A) Anzahl Teilmengen

Erklären (beweisen) Sie, dass die Anzahl der Teilmengen einer n -elementigen Ereignismenge genau 2^n beträgt.

[Hinweis: Dies ist auch die Anzahl der möglichen Rucksäcke, die man packen kann, wenn n unterschiedliche Gegenstände verfügbar sind.]

Aufgabe 7.4 (B) Augensumme Würfel

Sei X die Augensumme beim Wurf mit zwei Würfeln. Berechnen Sie über **Laplacesche Wahrscheinlichkeiten** (Elementarereignisse geeignet wählen!):

- (a) $P(X=2)$ (b) $P(X=7)$ (c) $P(6 \leq X \leq 8)$ (d) $P(3 \leq X \leq 11)$

Aufgabe 7.5 Geburtstag + Hashtabelle

- a) Wie groß ist die Wahrscheinlichkeit, dass von 10 Personen mindestens 2 am gleichen Tag Geburtstag haben? Erst "aus dem Bauch heraus" schätzen, dann ausrechnen! Wie lautet die Formel allgemein für k Personen? Ab welcher Gruppengröße lohnt sich im Schnitt eine Wette: "Wetten, dass in dieser Gruppe mindestens 2 Personen am gleichen Tag Geburtstag haben?"
- b) Beim **Hashverfahren** in der Informatik passiert im Grunde dasselbe: Daten d aus einem beliebigen Datenraum D werden über eine **Hashfunktion** $h(d) : D \rightarrow \{0, \dots, n-1\}$ Schlüssel-Zahlen zugeordnet (in Analogie zu a): Menschen $\rightarrow \{1, \dots, 365\}$. Eine **Kollision** liegt vor, wenn zwei verschiedene Daten d_1 und d_2 die gleichen Schlüssel $h(d_1)$ und $h(d_2)$ besitzen. Überlegen Sie sich eine Formel für die Kollisionswahrscheinlichkeit und schreiben Sie eine kleine Prozedur {in Maple oder Excel oder Java oder ...}, mit der Sie die Formel für verschiedene Werte ausrechnen können. Beantworten Sie dann folgende Fragen:
- Wie wahrscheinlich ist bei $n=10000$ und $a=40$ Aufrufen der Hashfunktion mindestens eine Kollision?
 - Bis zu welcher Zahl a ist bei $n=10000$ die Kollisionswahrscheinlichkeit unter 25%?

Hinweis zu a): Betrachte komplementäres Ereignis "Es gibt **keine** 2 Personen, die am gleichen Tag Geburtstag haben".

Hinweis zu b): Ein Beispiel für eine solche Hashfunktion ist $h(d) = d \bmod n$ (wenn d beliebig ganzzahlig ist).

Aufgabe 7.6 Hellseher

Ein Hellseher wirbt mit folgender Anzeige:

**Sage werdenden Eltern das Geschlecht ihres Kindes voraus.
Bei Nichteintreffen Geld zurück!**

Die Vorhersage kostet 100€, der Hellseher wirft eine Münze. Die Wahrscheinlichkeit einer Mädchengeburt sei 0.465. Berechnen Sie das mittlere Jahreseinkommen des Hellsehers, wenn 100 Anfragen je Monat eintreffen. Mit welcher Strategie kann er sein Jahreseinkommen (bei gleichem Tarif) verbessern?

[Hinweis: Stellen Sie eine Zufallsvariable "Gewinn" auf und berechnen Sie ihren Erwartungswert.]

Aufgabe 7.7 Zufallsvariable I

Eine diskrete Zufallsvariable X habe Ausprägungen x_i mit Wahrscheinlichkeiten p_i gemäß

x_i	1	5	10	20	100	
p_i	0.55	0.25	0.1	0.05	0.05	

- Berechnen Sie Erwartungswert $E(X)$ und Varianz $\text{Var}(x)$.
- Berechnen Sie Erwartungswert $E(2X+5)$ und Varianz $\text{Var}(2X+5)$.

Aufgabe 7.8 Zufallsvariable II: Bitfolgen

- a) Es werden Bitfolgen der Länge 3 zufällig generiert, wobei 0 und 1 gleichwahrscheinlich sind. Ermitteln Sie die Wahrscheinlichkeiten $p_k = P(X=k)$ der Zufallsvariablen $X = \text{Anzahl der Einsen in der Bitfolge}$ ferner den Erwartungswert, die Varianz und die Standardabweichung von X . Was ist $P(X \geq 1)$?
- b) Wiederholen Sie alle Berechnungen aus a) für Bitfolgen der Länge 4! Erkennen Sie eine Systematik bei den Wahrscheinlichkeiten? Können Sie raten oder beweisen, wie $P(X=k)$ für beliebige Länge n aussieht?
- c) Wiederholen Sie alle Berechnungen aus a) für Bitfolgen der Länge 2, wobei nun aber bei der Generierung $p(0)=0.2$ und $p(1)=0.8$ gilt. Können Sie den Rechenweg durch einen Baum visualisieren?

Aufgabe 7.9 Verteilungen

Eine Serienproduktion von Transistoren hat einen gleichbleibenden Ausschuss-Anteil von $p=30\%$. Wie groß ist die Wahrscheinlichkeit, dass in einer Stichprobe von $n=10$ entnommenen Transistoren

- (a) keine, (b) genau neun, (c) mehr als acht

fehlerhafte Bauteile sind? Überlegen Sie zunächst, welche Verteilung Sie anwenden!

- (d) Wie groß sind Erwartungswert und Varianz der Anzahl X der fehlerhaften Bauteile in der Stichprobe?

Aufgabe 7.10 Punkte verteilen I

Sie haben $a = 5$ Punkte, die Sie auf $n = 2$ Kategorien verteilen dürfen. (Denken Sie z.B. an Likes für einen Social-Media-Kanal, wobei Sie – anders als sonst – Mehrfach-Likes im Rahmen Ihres Budgets $a = 5$ vergeben können.)

- (a) Wie viele Möglichkeiten der Punkteverteilung gibt es, wenn Sie alle Punkte verteilen?
- (b) Wie viele Möglichkeiten gibt es, wenn Sie auch beliebig viele von den a Punkten zurückhalten dürfen?
- (c) Wie lautet die Antwort zu (a) für $a = 8$ und $n = 5$? Geben Sie eine Formel für beliebige a und n an!
- (d) Wie lautet die Antwort zu (b) für $a = 8$ und $n = 5$? Geben Sie eine Formel für beliebige a und n an!

Hinweise:

- 1) Bei (a) und (b) kann man sich wegen $n = 2$ einfach ein Bild von der Situation machen und abzählen.
- 2) Bei (c) und (d) geht das wegen $n = 5$ nicht mehr so einfach. Versuchen Sie eine Analogie zu einem der 4 Kombinatorik-Fälle der Vorlesung herzustellen.

Aufgabe 7.11 Punkte verteilen II

In Fortsetzung von Aufgabe 7.10, Fall $a = 8$ und $n = 5$: Nun dürfen Sie für die erste Kategorie K_1 der 5 Kategorien auch Likes und Dis-Likes vergeben. D.h. wenn Sie Punkte für K_1 vergeben, dürfen Sie zusätzlich entscheiden, ob diese positiv oder negativ gewertet werden. (Bei den anderen Kategorien $K_2, \dots, K_5$ weiterhin nur Likes.)

- (a) Wie viele Möglichkeiten der Punkteverteilung gibt es, wenn Sie alle Punkte verteilen?
- (b) Wie viele Möglichkeiten gibt es, wenn Sie auch beliebig viele von den a Punkten zurückhalten dürfen?

Können Sie für beide Fälle (a) und (b) eine Formel für beliebige a und n angeben?

Hinweise:

- 1) Vielleicht machen Sie sich auch hier erst für den einfachen Fall $a = 5$ und $n = 2$ und für beide Teilaufgaben (a) und (b) je ein Bild und zählen ab. Überlegen Sie dann, welche allgemeinere Formel zum gleichen Ergebnis führt.
- 2) Sie müssen die Anzahl der Fälle mit „0 Punkte für K1“ ermitteln und geeignet abziehen.

Aufgabe 7.12 Normalverteilung 1

Es sei bekannt, dass in einem Rechnernetz pro 1000 übertragener Pakete die Anzahl X der fehlerhaft übertragenen Pakete binomialverteilt sind mit Erwartungswert $\mu=80$.

- (a) Zeigen Sie, dass die auf 2 Nachkommastellen gerundete Standardabweichung $\sigma=8.58$ betragen muss.
- (b) Überprüfen Sie die Voraussetzungen für den Satz von Moivre-Laplace und berechnen Sie dann:
- (c) die Wahrscheinlichkeit, dass weniger als 70 aus 1000 Paketen und
- (d) die Wahrscheinlichkeit, dass mehr als 90 aus 1000 Paketen

fehlerhaft übertragen werden.

Hinweis: Tabelle der Funktion $\Phi(x)$ aus Formelsammlung benutzen.

Aufgabe 7.13 Normalverteilung 2

Eine Maschine produziere Werkstücke mit einem Sollmaß von 800 mm. Die Produktion sei **normalverteilt** mit $\mu= 800$ mm und $\sigma= 5$ mm. Es sollen Werte zwischen $(800-t)$ und $(800+t)$ mm eingehalten werden. Wie groß muss $t>0$ mindestens sein, damit der mittlere Ausschussanteil höchstens 5% ist?

Aufgabe 7.14 Normalverteilung 3 (aus einer Klausur)

Bei der Fertigung von SSD-Chips treten 20% ‚Bad Blocks‘ (defekte Datenblöcke) auf.

- a) Berechnen Sie die Wahrscheinlichkeit, dass unter 100 Blöcken $X=25$ oder mehr ‚Bad Blocks‘ sind.
- b) Bestimmen Sie ein zu 20 symmetrisches Intervall mit der Eigenschaft, dass die Anzahl der ‚Bad Blocks‘ mit einer Wahrscheinlichkeit von 95% in diesem Intervall liegt.

Hinweis zu b): Verwenden Sie den Grenzwertsatz von Moivre-Laplace. Überprüfen Sie zunächst, ob Sie ihn anwenden dürfen. Entnehmen Sie die Werte der Standardnormalverteilung einer geeigneten Tabelle.