

Übungsblatt 7 Statistik

Aufgabe 7.1 Beschreibende Statistik

Gegeben sei die folgende Stichprobe von (bereits sortierten) Messwerten z

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

- Ermitteln und zeichnen Sie Histogramm und Häufigkeitsverteilungsfunktion für die Klasseneinteilungen
 $[0,8]$, $[8,16]$, $[16,24]$, $[24,32]$
- 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! Wie ändern sich Box und Whisker?

Aufgabe 7.2

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?

Wie groß sind Erwartungswert und Varianz der Anzahl X der fehlerhaften Bauteile in der Stichprobe?

Aufgabe 7.3

Bei der geheimen Wahl zum Dekan treten drei Kandidaten D, E, F an, die von 9 Stimmen gewählt werden. Wieviele Wahlausgänge gibt es insgesamt? Wieviele mit der absoluten Mehrheit für Kandidat E? Wieviele mit einem "Patt"?

["Patt" = kein Kandidat kann sich durchsetzen]

Aufgabe 7.4

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 Hundennamen stehen zu haben?

Aufgabe 7.5 Geburtstag

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?"

Hinweis: Betrachte komplementäres Ereignis "keine 2 Personen haben am gleichen Tag Geburtstag".

OPTIONAL {in MAPLE oder Excel/VBA oder Java}: Schreiben Sie eine Prozedur, mit der Sie zahlreiche Gruppen variabler Größe simulieren können und überprüfen Sie so empirisch Ihre Ergebnisse.

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 Spam-Filter

In Ihrer Mailbox sind 25% der neu eintreffenden Nachrichten SPAM.

Sie stellen aufgrund der Inhalte Ihrer Mailbox folgende Tabelle für das Vorkommen bestimmter Wörter in Ihren Spam- oder Nicht-Spam-Mails (SPAM oder HAM) auf:

	„Gewinn“	„Million“
SPAM	19%	25%
HAM	1%	10%

Diese Tabelle ist wie folgt zu lesen: Die Wahrscheinlichkeit, in einer SPAM-Nachricht das Wort W_1 ="Gewinn" zu finden, beträgt $P(W_1|S)=19\%$.

- Eine neu eintreffende Nachricht enthält das Wort „Gewinn“. Mit welcher Wahrscheinlichkeit ist sie SPAM?
- Eine neu eintreffende Nachricht enthält das Wort „Million“. Mit welcher Wahrscheinlichkeit ist sie SPAM?
- Eine neu eintreffende Nachricht enthält beide Worte „Gewinn“ UND „Million“. Mit welcher Wahrscheinlichkeit ist sie SPAM?

Hinweise: 1) Benutzen Sie die Bayes-Formel. 2) Machen Sie bei (c) die (nicht unbedingt realistische, aber das Leben stark vereinfachende) Annahme, dass das Auftreten der beiden Wörter W_1 und W_2 in beiden Mail-Kategorien (SPAM oder HAM) statistisch unabhängig sei, also $P(W_1 \cap W_2) = P(W_1) \cdot P(W_2)$.

Aufgabe 7.8 Mammographie

Die Wahrscheinlichkeit für Frauen, an Brustkrebs zu erkranken, beträgt 1%. Ein Mammographie-Test mit Ergebnis „pos“ oder „neg“ hat folgende Zuverlässigkeitsquoten:

	krank	gesund
pos	95%	10%
neg	5%	90%

also fällt der Test für kranke Frauen zu 95% positiv und zu 5% negativ aus., d.h.

$P(\text{pos}|\text{krank}) = 95\%$.

Wenn für eine zufällig ausgewählte Frau der Test positiv ausfällt, mit welcher Wahrscheinlichkeit ist sie dann krank?

Aufgabe 7.9 Normalverteilung 1

Es sei bekannt, dass in einem Rechnernetz pro 1000 übertragener Pakete die Anzahl X der fehlerhaft übertragenen Pakete normalverteilt sei mit Mittelwert $m=80$ und Standardabweichung $\sigma=15$. Berechnen Sie die Wahrscheinlichkeit dafür,

(a) dass weniger als 50 aus 1000 Paketen

(b) dass mehr als 100 aus 1000 Paketen

fehlerhaft übertragen werden.

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

Aufgabe 7.10 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 7.11 (+) Miraculus der Magier

Miraculus der Magier späht vor seinem Auftritt durch den Vorhang. Der Saal mit seinen 2000 Plätzen ist zu über 95% besetzt, dabei sind die Frauen in der Mehrheit. Wie immer wird er zwei zufällig aus dem Publikum ausgewählte Personen als Assistenten auf die Bühne bitten. Am liebsten wären ihm eine Frau und ein Mann. Tatsächlich wird dieses Ereignis mit einer Wahrscheinlichkeit von genau 50% eintreten.

Aus wie vielen Frauen und Männern setzt sich das Publikum zusammen?

Tipp: Versuchen Sie, die Gleichung $2F - X = \sqrt{X}$ zu zeigen (X : Größe Publikum, F : Anzahl Frauen), und argumentieren Sie von da aus weiter.