UE Statistik 1, SS 2015, letztes Update am 27. April 2015

Übungsbeispiele

Beispiele mit Musterlösungen finden Sie auch in dem Buch Brannath, W., Futschik, A., Krall, C., (2010) *Statistik im Studium der Wirtschaftswissenschaften*. 3. Edition. Am Anfang der jeweiligen Kapitel finden Sie eine Tabelle mit Beispielen aus diesem Buch, die für Sie als Übungsbeispiele nützlich sind. Die zweite Zeile in der Tabelle ist für mögliche Korrekturen der Lösungen aus dem Buch.

1 Ereignisse und ihre Wahrscheinlichkeiten

Beispiel	3-37	3-38	3-41	3-45	3-46	3-50	3-54	3-60	3-61
Korrektur	-	-	-	-	-	-	-	-	-

Tabelle 1: Empfohlene Beispiele aus dem Buch.

1.1 Stichprobenräume und Mengenoperationen

1.1 Geben Sie den Stichprobenraum Ω[1] für folgende Zufallsexperimente an:

(a) Rating eines zufällig ausgewählten Landes beim Standard & Poor's (siehe Abbildung 19),
(b) zweifacher Würfelwurf,
(c) dreifacher Münzwurf,
(d) Lebensdauer (in Stunden) einer zufällig ausgewählten Glühbirne.

1.2 Welche Teilmengen des Stichprobenraums aus Beispiel 1b entsprechenden folgenden Ereignissen?

(a) Die Summe der Augenzahlen ist größer 4.
(b) Die Summe der Augenzahlen ist größer oder gleich 4.
(c) Die Summe der Augenzahlen ist größer 12.
(d) Zumindest einer der Würfel zeigt eine Augenzahl größer als 5.
(e) Keiner der Würfel zeigt eine Augenzahl größer als 3.

1.3 Bilden Sie die Mengen, die den Komplementärereignisen der Ereignisse aus Aufgabe 2 entsprechen.

1.4 Man bezeichne die in Aufgabe 2(a) gefundene Menge mit A, die in 2(b) gefundene Menge mit B usw. Bilden Sie folgende Mengen und beschreiben Sie die Ereigniße, die diesen Mengen entsprechen, in Worten.

(a) $A \cap B$
(b) $D \cup E$
(c) $D \cap E$
(d) $C \cup A$

[1] In allen Übungsbeispielen mit Würfeln/Münten wird angenommen, daßdie Würfel/Münzen gleichzeitig geworfen werden und man sie unterscheiden kann (z.B. durch verschiedene Farben). Überlegen Sie sich bitte, was sich ändern würde, falls wir die Würfel/Münzen nicht unterscheiden könnten.

(e) $C \cap A$

1.5 Zeigen Sie die Gültigkeit der folgenden Relationen mittels eines Venn-Diagramms.

(a) $\mathsf{A} \cap (\mathsf{B} \cup \mathsf{C}) = (\mathsf{A} \cap \mathsf{B}) \cup (\mathsf{A} \cap \mathsf{C})$
(b) $\mathsf{A} \cup (\mathsf{B} \cap \mathsf{C}) = (\mathsf{A} \cup \mathsf{B}) \cap (\mathsf{A} \cup \mathsf{C})$
(c) $(\mathsf{A} \cup \mathsf{B})^c = \mathsf{A}^c \cap \mathsf{B}^c$
(d) $(\mathsf{A} \cap \mathsf{B})^c = \mathsf{A}^c \cup \mathsf{B}^c$
(e) $\mathsf{B} = (\mathsf{B} \cap \mathsf{A}) \cup (\mathsf{B} \cap \mathsf{A}^c)$
(f) $\mathsf{A} \cup \mathsf{B} = \mathsf{A} \cup (\mathsf{B} \cap \mathsf{A}^c)$

1.2 Wahrscheinlichkeitsverteilung

1.6 Wir werfen gleichzeitig 4 Würfel. Geben Sie die Wahrscheinlichkeit an, dass

(a) alle Augenzahlen ungerade sind,
(b) die Summe aller Augenzahlen zusammen 6 ist,
(c) die Summe aller Augenzahlen zusammen > 5 ist.

1.7 Mit welcher Wahrscheinlichkeit ist **mindestens** eine Augenzahl gleich 6, falls wir 2 Würfel werfen?

1.8 Geben Sie die Wahrscheinlichkeiten der Ereignisse aus Aufgabe 2 an.

1.9 Aus den Zahlen 1 bis 49 werden beim „Lotto“ sechs verschiedene ausgewählt. Wie groß ist die Wahrscheinlichkeit, daß ein Spieler

(a) genau sechs Richtige,
(b) genau fünf Richtige,
(c) keine Richtige,
(d) höchstens zwei Richtige hat?

1.10 Bei einer Multiple-Choice-Prüfung sind sechs Fragen jeweils vier Antwortmöglichkeiten beigegeben, wovon jeweils nur eine richtig ist. Die Prüfung gilt als bestanden, wenn mindestens vier Fragen richtig beantwortet wurden. Ein Student kreuzt bei den sechs Fragen jeweils eine Antwort zufällig an. Mit welcher Wahrscheinlichkeit besteht er die Prüfung?

1.11 Sie haben in Ihrem Geldbeutel 4 Banknoten und zwar: zwei 5 € Banknoten, eine 10 € Banknote und eine 20 € Banknote. Ein Dieb nimmt zufällig zwei Banknoten (egal in welcher Reihenfolge) aus Ihrem Geldbeutel. (Er kann mit derselben Wahrscheinlichkeit jede Banknote nehmen). Es interessiert uns die Summe, die er uns gestohlen hat.

(a) Geben Sie den zugrundeliegenden Stichprobenraum Ω so wie die Mächtigkeit $|\Omega|$ für diesen Zufallsexperiment an.
(b) Berechnen Sie die Wahrscheinlichkeit für jeden Betrag der uns der Dieb stehlen kann.

1.12 Gegeben seien zwei Ereignisse A und B mit $\mathbb{P}(A) = \frac{1}{3}$ und $\mathbb{P}(B^c) = \frac{1}{4}$. Können die beiden Ereignisse disjunkt sein? Warum (nicht)?

1.13 Seien A und B unabhängige Ereignisse mit $\mathbb{P}(A) = 0.5$ und $\mathbb{P}(B \cup A) = 0.6$. Wie hoch ist $\mathbb{P}(B)$?

1.3 Bedingte Wahrscheinlichkeit und Satz von Bayes

1.14 Wir werfen gleichzeitig 2 unterscheidbare Würfel (z.B. einen grünen und einen roten). Sei A das Ereignis, dass eine der Augenzahlen gleich 6 ist und sei B das Ereignis, dass die Summe beider Augenzahlen 8 ist.

(a) Berechnen Sie $\mathbb{P}(A|B)$.

(b) Sind die Ereignisse A und B unabhängig?

1.15 Im Statistik 1 Kurs gibt es 70% Männer und 30% Frauen. Lange Haare tragen 10% der Männer und 80% der Frauen. Mit welcher Wahrscheinlichkeit rufe ich zur Tafel

(a) jemanden (egal ob Frau oder Mann), der lange Haare hat.

(b) eine Frau, unter der Bedingung, dass ich nur jemanden mit langen Haaren zur Tafel rufe.

1.16 Seien A und B zwei Ereignisse mit $\mathbb{P}(A) = 0.5$ und $\mathbb{P}(A \cup B) = 0.7$. Wie hoch ist $\mathbb{P}(B)$, wenn

(a) A und B unabhängig sind?

(b) A und B disjunkt sind?

(c) $\mathbb{P}(A|B) = 0.3$?

(d) $\mathbb{P}(A|B) = 0.5$?

1.17 Gegeben seien drei Ereignisse A_i, $i = 1{,}2{,}3$ und die Wahrscheinlichkeiten $\mathbb{P}(A_3|A_1 \cap A_2) = \frac{1}{5}$, $\mathbb{P}(A_1 \cap A_2) = \frac{1}{10}$ und $\mathbb{P}(A_3|A_1^c \cup A_2^c) = \frac{3}{5}$. Berechnen Sie $\mathbb{P}(A_1 \cap A_2|A_3)$.

1.18 Urne A enthält zwei weiße und zwei schwarze Kugeln. Urne B enthält drei weiße und zwei schwarze Kugeln. Eine Kugel wird von Urne A nach Urne B transferiert. Danach wird eine Kugel aus Urne B gezogen.

(a) Wie hoch ist die Wahrscheinlichkeit, dass die gezogene Kugel weiß ist?

(b) Gegeben die gezogene Kugel ist weiß. Wie hoch ist die Wahrscheinlichkeit, dass die transferierte Kugel auch weiß ist?

1.19 Sie haben drei Münzen, zwei davon sind normale Münzen und eine Münze hat zwei Kopf-Seiten. Sie wählen zufällig eine der Münzen und werfen Kopf. Wie hoch ist die Wahrscheinlichkeit eine normale Münze gezogen zu haben?

2 Diskrete Zufallsvariablen und Ihre Momente

Beispiel	4-3	4-4	4-5	4-6	4-8	4-35	4-37	4-40	4-41	4-63
Korrektur	-	-	-	-	-	-	-	-	-	-

Tabelle 2: Empfohlene Beispiele aus dem Buch.

Versuchen Sie beim Beispiel 4.35 zusätzlich den erwarteten Gewinn berechnen, falls Sie 6 Runden spielen und 1 € pro Runde für Ihre Teilnahme bezahlen. Versuchen Sie die Beispiele 4.40 und 4.63 mit binomischer Verteilung zu lösen.

2.1 Gegeben ist eine diskrete Zufallsvariable X mit der Verteilung:

$$X = \begin{cases} x_1 = -5, & \text{mit } \mathbb{P}(X = -5) = 0.3, \\ x_2 = 0, & \text{mit } \mathbb{P}(X = 0) = 0.45, \\ x_3 = 2, & \text{mit } \mathbb{P}(X = 2) = 0.25, \end{cases}$$

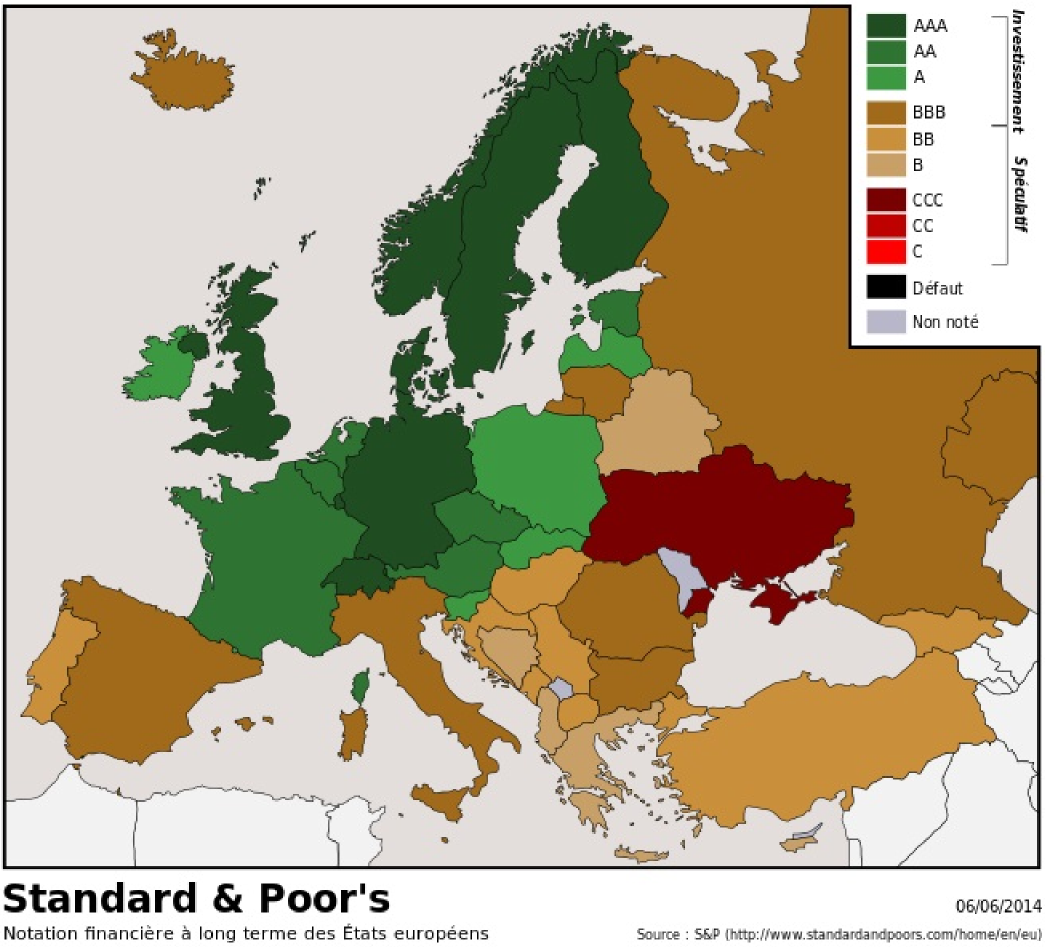

Abbildung 1: Die Ratings von Standard & Poor's für europäische Länder.

(a) Berechnen Sie die Standardabweichung von X

(b) Berechnen Sie die Verteilungsfunktion $F(x)$ der Zufallsvariablen X und zeichnen Sie sie.

(c) Berechnen Sie den Erwartungswert der Zufallsvariable Y, falls $Y = X^3$.

(d) Bestimmen Sie die Varianz der Zufallsvariable Z, falls $Z = -3X - 1$.

2.2 Sie haben in Ihrem Geldbeutel 4 Banknoten und zwar: zwei 5 € Banknoten, eine 10 € Banknote und eine 20 € Banknote. Ein Dieb nimmt zufällig zwei Banknoten (egal in welcher Reihenfolge) aus Ihrem Geldbeutel. (Er kann mit derselben Wahrscheinlichkeit jede Banknote nehmen). X ist eine Zufallsvariable, die sagt, wie viel Geld Ihnen genommen wird.

(a) Schreiben Sie die Verteilung der Zufallsvariable X auf.

(b) Der Dieb muss nachher 10 € Bestrafung für Falschparken bezahlen. Zuhause nimmt ihm seine Frau noch $\frac{4}{5}$ davon, was er gebracht hat. Y ist eine Zufallsvariable, die uns sagt, wie viel ihm am Ende des Tages übriggeblieben ist. Berechnen Sie die Wahrscheinlichkeit, dass er dann noch auf ein Bier mit seinen Freunden gehen kann d.h. $\mathbb{P}(Y \geq 2.5\,€) =?$ (Das Bier kostet nämlich nur 2.5 €).

(c) Skizzieren Sie die Verteilungsfunktion $F(x)$ der Zufallsvariablen X und Y.

(d) Berechnen Sie den erwarteten Verlust sowie var X.

(e) Berechnen Sie den Erwartungswert und die Varianz von Y.

2.3 Gegeben Sei eine diskrete Verteilung auf den Punkten 1,2,3 und 4. Die Wahrscheinlichkeiten der einzelnen Punkte sind $p_1 = 0.1$, $p_2 = 0.5$, $p_3 = 0.15$ und $p_4 =?$.

(a) Finden Sie p_4. Berechnen und zeichnen Sie die kumulative Verteilungsfunktion.

(b) Berechnen Sie Erwartungswert und Varianz.

2.4 Gegeben sei eine Zufallsvariable X, die einer Gleichverteilung auf den Punkten $-10, 0, 1$ folgt. Berechnen Sie die Wahrscheinlichkeitsfunktion. Berechnen Sie den Erwartungswert und die Varianz.

2.5 Betrachten Sie das Experiment: Würfeln mit zwei sechsseitigen Würfeln.

(a) Berechnen Sie die Wahrscheinlichkeitsfunktion und die Verteilungsfunktion der Zufallsvariablen Y, die die absoluten Differenz der beiden Augenzahlen beschreibt.

(b) Berechnen Sie den Erwartungswert und die Varianz der Zufallsvariablen Y.

2.6 Für eine binomialverteilte Zufallsvariable X mit $\mathbb{E}(X) = 5$ und $Var(X) = 4$ finden Sie die Parameter n und p.

2.7 Betrachten Sie die Zufallsvariable X mit $\mathbb{E}(X) = 10$ und $Var(X) = 25$. Geben Sie die positiven Zahlen a und b an, so dass die Zufallsvariable $Y = aX - b$ Erwartungswert 0 und Varianz 1 hat.

2.8 Gegeben sei eine Zufallsvariable X mit $\mathbb{E}(X) = \mu$ und $Var(X) = \sigma^2$. Berechnen Sie Erwartungswert und Varianz der Zufallsvariablen

$$Y = \frac{X - \mu}{\sigma}.$$

2.9 Ein Kontrollor weiß, dass jeder 10. Fahrgast ohne Fahrschein unterwegs ist. Er kontrolliert 20 Fahrgäste. Wie groß ist die Wahrscheinlichkeit, dass er

(a) keinen Schwarzfahrer,

(b) einen Schwarzfahrer,

(c) mindestens 2 Schwarzfahrer erwischt?

2.10 Ein Würfel wird so oft geworfen bis er „6“ zeigt. Es sei T die Anzahl der Fehlversuche bis zur ersten „6“. Berechnen Sie die Wahrscheinlichkeit, dass mehr als 3 bzw. mehr als 5 Fehlversuche auftreten

2.11 Betrachten Sie das vorige Beispiel. Berechnen Sie die Wahrscheinlichkeit, dass genau drei Fehlversuche vorliegen bis der Würfel zum ersten Mal „6“ zeigt. Berechnen Sie die Wahrscheinlichkeit, dass 6 Fehlversuche vorliegen, gegeben dass schon mindestens drei Fehlversuche vorliegen. Was fällt Ihnen auf?

2.12 Gegeben sei eine nichtleere Menge Ω. Weiters seien $A \in \Omega$ und $B \in \Omega$. Argumentieren Sie, ob folgende Mengensysteme σ-Algebren auf Ω sind:

(a) $\mathcal{F}_1 = \{\emptyset, \Omega\}$

(b) $\mathcal{F}_2 = \{\emptyset, A, B^c, \Omega\}$

(c) $\mathcal{F}_3 = \{\emptyset, A, A^c, \Omega\}$

(d) $\mathcal{F}_4 = \{\emptyset, A, B, A^c, B^c, A \cup B, \Omega\}$

3 Asymptotik und stetige Zufallsvariablen

Beispiel	4.42	4.43b)	4.44	4.45	4.47
Korrektur	-	-	-	✗	-

Tabelle 3: Empfohlene Beispiele aus dem Buch:

Für das Beispiel 4.45 gibt es mindestens 3 mögliche Lösungswege. Probieren sie alle drei und enscheiden Sie sich, welche war die einfachste. Lösungen: Bi: 0.908, Poi: 0.907, N: 0.921.

3.1 Die Zufallsvariable X sei standardnormalverteilt. Mit welcher Wahrscheinlichkeit:

(a) Ist X kleiner als 1?

(b) Liegt X zwischen 1 und 2?

(c) Ist X entweder größer als 2 oder kleiner als -2?

(d) Wie verändert sich das Ergebnis aus (c) falls $X \sim N(-1.5,2)$.

3.2 Die Zufallsvariablen $X_1, \ldots, X_n$ sind i.i.d, wobei $X_i \sim N(\mu,\sigma^2)$ für jede $i = 1, \ldots ,n$.

(a) Mit welcher Wahrscheinlichkeit ist $\bar{X}_n$ kleiner als die Konstante μ?

(b) Wie groß ist n, so dass $\mathbb{P}(\bar{X}_n > 0) = 99\%$, angenommen dass $\mu = 4$ und $\sigma^2 = 1$?

3.3 Es seien $X_1, \ldots, X_n$ i.i.d. $N(\mu, \sigma^2)$ mit $\mu \in \mathbb{R}$ und $0 < \sigma^2 < \infty$.

(a) Bestimmen Sie die Verteilung von $X_1 + \ldots + X_n$.

(b) Bestimmen Sie die Verteilung von $X_1^2 + \ldots + X_n^2$, wenn $\mu = 0$ und $\sigma^2 = 1$.

(c) Bestimmen Sie die Verteilung von $\frac{X_1}{\sqrt{(X_2^2+\ldots+X_n^2)/(n-1)}}$, wenn $\mu = 0$ und $\sigma^2 = 1$.

(d) Bestimmen Sie $\mathbb{E}[\bar{X}_n]$ und $Var(\bar{X}_n)$, wobei $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$. Brauchen Sie hierfür die Annahme, dass die Zufallsvariablen normalverteilt sind?

3.4 Es seien $X_1, \ldots X_n$ i.i.d. mit Erwartungswert 0 und Varianz $0 < \sigma^2 < \infty$. Gilt $\mathbb{E}[(\bar{X}_n - 1)^2] \to 0$ für $n \to \infty$, wobei $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$?

3.5 Es sei X gleichverteilt auf dem Intervall $[-1{,}3]$. Berechnen Sie

(a) $P(X \leq -1)$,
(b) $P(X \leq 2)$,
(c) $P(X \leq 3)$,
(d) $P(1 \leq X \leq 2)$.

3.6 Es sei X gleichverteilt auf dem Intervall $[0, \theta]$, wobei $\theta > 0$. Berechnen Sie den Erwartungswert und die Varianz von X.

3.7 Es seien $X_1, \ldots$ i.i.d. exponentialverteilt mit Parameter $\tau = 2$. Welche Aussage trifft zu? Korrigieren Sie den/die Fehler.

(a) $\sqrt{n}\frac{\bar{X}_n - 2}{2} \xrightarrow{d} N(0{,}1)$
(b) $\sqrt{n}\frac{\bar{X}_n - 2\sqrt{n-1}/\sqrt{n}}{2} \xrightarrow{d} N(1{,}1)$
(c) $\sqrt{n}\frac{\bar{X}_n - 2}{4} \xrightarrow{d} N(0{,}1)$

3.8 Berechnen Sie den Erwartungswert und die Varianz der Zufallsvariable X, wobei X eine Verteilung mit folgender Verteilungsfunktion hat:

$$F(x) = \begin{cases} 0 & x < 0 \\ x^3 & 0 \leq x < 1 \\ 1 & x \geq 1. \end{cases}$$

3.9 Sei X die Lebensdauer von Glühbirnen (in Stunden). Die Dichte von X sei durch

$$f(x) = \begin{cases} 2x & 0 \leq x < \frac{1}{2} \\ \frac{3}{4} & 2 < x < 3 \\ 0 & \text{sonst} \end{cases}$$

gegeben.

(a) Welcher Prozentsatz an Glühbirnen überlebt länger als 15 Minuten?
(b) Berechnen Sie $\mathbb{E}[X]$ und $Var(X)$.
(c) Berechnen Sie $\mathbb{P}(0.25 < X \leq 2.2 | X > 1)$.

3.10 Es sei X eine Zufallsvariable mit der Dichte

$$f(x) = \begin{cases} \frac{2}{x^3}, & x \geq 1 \\ 0, & x < 1 \end{cases}.$$

Berechnen Sie $\mathbb{E}[X]$ und $Var(X)$.

3.11 Eine Zufallsvariable X hat die Verteilung mit folgender Dichtefunktion:

$$f(x) = \begin{cases} 2c + 2x & \text{für } 0 \leq x \leq 1, \\ 0 & \text{sonst.} \end{cases}$$

(a) Zeigen Sie, dass $f(x)$ eine Dichtefunktion ist nur wenn $c = 0$. (Hinweis: Benutzen Sie diese Eigenschaften der Dichtefunktion: $\int_0^1 f(x)dx = 1$)
(b) Berechnen Sie $\mathbb{E}[X]$,

(c) Zeichnen Sie $f(x)$ im $[x,f(x)]$ Koordinatensystem, wobei $0 \le x \le 1$. Wo liegt $\mathbb{E}[X]$ auf Ihrem Bild?

(d) Berechnen Sie den Median der Verteilung von X ohne die Verteilungsfunktion zu berechnen. Wo liegt der Median auf Ihrem Bild? (Hinweis: Median ist so eine reale Zahl, dass $\mathbb{P}(\text{median} \le X) = 1/2$).

3.12 Eine Zufallsvariable X hat die Verteilung mit folgender Dichtefunktion:

$$f(x) = \begin{cases} c * e^{-x/5} & \text{falls } x > 0, \\ 0 & \text{sonst.} \end{cases}$$

(a) Berechnen Sie die Konstante c, so dass $f(x)$ eine Dichtefunktion ist.

(b) Berechnen Sie die Verteilungsfunktion dieser Zufallsvariablen und versuchen Sie den Graphen dieser Funktion zu skizzieren.

(c) Berechnen Sie EX und var(X).

(d) Berechnen Sie die Quantilsfunktion und den Median. Versuchen Sie den Graphen der Quantilsfunktion zu skizzieren.

3.13 Es sei $5X - 2$ exponentialverteilt mit Parameter $\tau = 3$. Bestimmen Sie $\mathbb{E}[X]$ und $Var(X)$.

3.14 Nehmen Sie an, dass die Dauer einer Reparatur (in Stunden) exponentialverteilt mit Parameter $\tau = 2$ ist. Wie groß ist die Wahrscheinlichkeit, dass die Reparatur länger als 4 Stunden dauert. Wenn eine Reparatur schon länger als 4 Stunden dauert, wie groß ist die Wahrscheinlichkeit, dass die Reparatur länger als 8 Stunden dauert.

3.15 Die Zeit zwischen dem Eintreffen zweier Kunden (in Minuten) sei exponentialverteilt mit Parameter $\tau = 3$. Man bestimme die Wahrscheinlichkeit, dass zwischen den Ankünften zweier Kunden mehr als 3 Minuten vergehen.

3.16 Es sei $3 - \frac{5}{2}X$ standardnormalverteilt. Bestimmen Sie $E(X)$ und $V(X)$.

3.17 Ein Speditionsunternehmen weiß , dass die Fahrtdauer (in Minuten) zwischen zwei Städten normalverteilt mit $\mu = 230$ und $\sigma = 15$ ist.

(a) Wie groß ist die Wahrscheinlichkeit, dass eine Fahrt länger als 4 Stunden dauert?

(b) Wöchentlich gibt es 5 Fahrten zwischen diesen Städten. Wie ist die durchschnittliche Fahrtdauer verteilt?

(c) Wie groß ist die Wahrscheinlichkeit, dass die durchschnittliche Fahrtdauer über 4 Stunden liegt?

3.18 In einer Bevölkerungsgruppe sei der Intelligenzquotient IQ normalverteilt mit $IQ \sim N(105, 100)$.

(a) Mit welcher Wahrscheinlichkeit hat eine zufällig ausgewählte Person einen IQ über 110?

(b) Mit welcher Wahrscheinlichkeit ist der durchschnittliche IQ von 4 zufällig ausgewählten Personen über 110?

(c) Mit welcher Wahrscheinlichkeit ist der durchschnittliche IQ von 25 zufällig ausgewählten Personen über 110?

3.19 Um sich für die Polizeischule zu qualifizieren, müssen die Kandidaten eine Eignungsprüfung machen. Es wird angenommen, dass die Punkte, die bei der Prüfung erreicht werden, normalverteilt mit $\mu = 200$ und $\sigma = 20$ sind. Um sich zu qualifizieren, müssen die Kandidaten im Punktebereich der oberen 10% liegen. Wie viele Punkte muss ein Kandidat mindestens erreichen?

In den folgenden Aufgaben sollten Sie den Zentralen Grenzwertsatz benutzen. Überlegen Sie sich, bei jedem Beispiel, ob die Annahmen des Satzes erfühlt sind.

3.20 Eine Kiste enthält 5000 Schrauben, von denen 250 ein defektes Gewinde besitzen. Aus der Kiste werden 100 Schrauben **mit** Zurücklegen entnommen.

(a) Welche exakte Verteilung besitzt die Zufallsvariable X, die Anzahl der defekten Schrauben unter den 100 gezogenen ergibt? Durch welche Verteilung lässt sich diese exakte Verteilung approximieren?

(b) Berechnen Sie approximativ die Wahrscheinlichkeit, dass maximal 3 Schrauben unter den gezogenen defekt sind.

3.21 Ein Beamter verlässt an den 225 Arbeitstagen eines Jahres sein Büro immer erst kurz nach Dienstschluss. Die Dauern der täglichen zusätzlichen Arbeitszeiten lassen sich jeweils durch exponentialverteilte Zufallsvariablen $X_1, \ldots, X_{225}$ mit einem Erwartungswert von 5 Minuten angemessen beschreiben und sind unabhängig.

(a) Verwenden sie die Zentrale Grenzwertsatz damit Sie die Parameter der Normalverteilung von $\sum_{i=1}^{225} X_i$ berechnen.

(b) Berechnen Sie (approximativ[2]) die Wahrscheinlichkeit, dass unser Beamter in einem Jahr mehr als 16 Stunden zusätzlich arbeitet.

3.22 Eine Firma hat 100 Kunden. Jeder Kunde bezahlt der Firma einen Betrag für das nächste Jahr. Die Zufallsvariable X_i entspricht der Zahlung des i-ten Kunden. Wir nehmen an, dass die Zufallsvariablen $X_1, \ldots, X_{100}$ unabhängig sind mit $\mathrm{E}X_i = \mu = 170€$ und $\mathrm{var}X_i = \sigma^2 = 2500€^2$.

(a) Bezeichnen wir als G die Gesamteinnahmen der Firma. Wie kann man mit Hilfe der Zufallsvariablen $X_1,...,X_{100}$ die Gesamteinnahmen G der Firma ausdrücken?

(b) Bezeichnen wir mit A die Gesamtausgaben der Firma. Wie groß können diese sein, sodass sie mit einer Wahrscheinlichkeit von 90% nicht die Gesamteinnahmen der Firma übersteigen? Mit anderen Worten, berechnen Sie A, so dass $\mathbb{P}(G \geq A) = 90\%$.

3.23 Sie wollen Ihre eigene Firma mit 100 Angestellten gründen. Jeder von Ihren Angestellten würde einen Gehalt ungefähr in der Höhe von $\mu = 1400$ € (netto) monatlich **erwarten**. Die Standardabweichung σ des Gehaltswunsches eines Angestellten ist 300 €. Mit wenigstens wie viel Geld sollten Sie beim Budgetentwurf für die Lohnausgaben rechnen, so dass Sie diese Ausgaben mit der Wahrscheinlichkeit von 99% nicht unterschätzen?

4 Schätzen und Testen

Beispiel	6.19	6.20	6.22	6.23	6.25	6.26
Korrektur	-	-	-	-	-	-

Tabelle 4: Empfohlene Beispiele aus dem Buch.

Im Beispiel 6.20 brauchen Sie die Fragestellung nicht aus dem Konfidenzintervall ermitteln sondern einen statistischen Test verwenden.

4.1 Mehrdimensionale Zufallsvariablen

4.1 Es wird mit zwei vierseitigen Würfeln gewürfelt. Berechnen Sie die gemeinsame Wahrscheinlichkeitsfunktion der Zufallsvariable X, die die maximale Augenzahl beschreibt, und der Zufallsvariablen Y, die die Summe der beiden Augenzahlen beschreibt. Sind die beiden Variablen unabhängig?

[2] Da die Anzahl der Beobachtungen nicht zum ∞ konvergiert, sondern nur beim 225 bleibt, ist unsere Ergebnis nicht exakt, sondern nur approximativ.

4.2 Gegeben sei folgende gemeinsame Wahrscheinlichkeitsfunktion von X und Y.

$Y \backslash X$	1	2	3	4	5	6
0	0	0	$\frac{1}{32}$	$\frac{2}{32}$	$\frac{2}{32}$	$\frac{3}{32}$
1	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$
2	$\frac{1}{32}$	$\frac{1}{32}$	$\frac{1}{64}$	$\frac{1}{64}$	0	$\frac{2}{64}$

Berechnen Sie

(a) die marginale Wahrscheinlichkeitsfunktion von X und von Y,

(b) die Kovarianz und die Korrelation zwischen X und Y und

(c) die Wahrscheinlichkeiten $\mathbb{P}(X \leq 1, Y = 2)$, $\mathbb{P}(X \leq 1)$, $\mathbb{P}(Y \leq 3)$ und $\mathbb{P}(X < 3, Y \leq 4)$.

(d) Angenommen Sie haben folgende Stichprobe gegeben: $X_1 = 1$, $X_2 = 4$, $X_3 = 4$, $X_4 = 2$, $X_5 = 6$, $Y_1 = 0$, $Y_2 = 2$, $Y_3 = 2$, $Y_4 = 0$, $Y_5 = 1$. Berechnen Sie die Stichprobenkovarianz.

4.3 Gegeben sei folgende gemeinsame Wahrscheinlichkeitsfunktion von X und Y.

$Y \backslash X$	1	2	3
0	$\frac{1}{6}$	$\frac{1}{18}$	$\frac{1}{36}$
3	$\frac{1}{2}$	$\frac{1}{6}$	$\frac{1}{12}$

(a) Berechnen Sie die Kovarianz und die Korrelation zwischen X und Y.

(b) Sind die beiden Zufallsvariablen unabhängig?

4.4 Gegeben sei folgende gemeinsame Wahrscheinlichkeitsfunktion von X und Y.

$Y \backslash X$	-2	-1	1	2
-2	$\frac{1}{10}$	0	0	0
-1	0	0	$\frac{2}{5}$	0
1	0	$\frac{2}{5}$	0	0
2	0	0	0	$\frac{1}{10}$

(a) Berechnen Sie die Kovarianz und die Korrelation zwischen X und Y.

(b) Sind die beiden Zufallsvariablen unabhängig?

(c) Berechnen Sie den Median der Zufallsvariablen X.

4.5 Gegeben sei ein Zufallsvektor $X = (X_1, X_2)$, der die Werte $\{(0,0), (0,1), (1,0), (1,1)\}$ mit jeweils gleichen Wahrscheinlichkeiten annimmt.

(a) Berechnen Sie die Erwartungswerte und die Varianzen von X_1 und X_2.

(b) Berechnen Sie die Kovarianz und die Korrelation zwischen X_1 und X_2. Sind sie unabhängig?

(c) Ändern Sie die Verteilung so, dass $\mathbb{P}(X_1 = 1, X_2 = 1) = \frac{3}{4}$ und die übrigen Werte die Wahrscheinlichkeit $\frac{1}{12}$ haben und wiederholen Sie die Analyse aus (a), (b).

4.2 Schätzen

4.6 Es sei $X_1, \dots X_n$ i.i.d. mit $\mathbb{E}[X_i] = \mu$ und $\mathrm{Var}(X_i) = \sigma^2$.

(a) Ist der Schätzer $\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i$ unverzerrt für den Parameter μ?

(b) Angenommen Sie kennen den Parameter μ. Ist der Schätzer $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2$ unverzerrt für den Parameter σ^2?

(c) Angenommen Sie kennen den Parameter μ NICHT und müssen auch diesen Parameter schätzen. Hierfür verwenden wir den Schätzer $\hat{\mu}$ aus (a). Ist der Schätzer $\breve{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \hat{\mu})^2$ unverzerrt für den Parameter σ^2?

4.7 Gegeben sei eine i.i.d. Zufallsstichprobe $Y_1, \dots, Y_n$ aus einer Exponentialverteilung mit Parameter $\tau > 0$. Die Dichte ist folgendermaßen definiert:

$$f(y) = \begin{cases} \frac{1}{\tau}\,\mathrm{e}^{-\frac{y}{\tau}}, & y > 0 \\ 0, & y \leq 0. \end{cases}$$

(a) Ist der Schätzer

$$\hat{\tau} = \frac{1}{n}\sum_{i=1}^{n} Y_i$$

unverzerrt für τ?

(b) Berechnen Sie den Median dieser Verteilung.

(c) Geben Sie einen unverzerrten Schätzer für den Median an.

4.8 Gegeben sei eine i.i.d. Zufallsstichprobe $X_1, \dots, X_n$ mit $\mathbb{E}[X_1] = \mu$ und $\mathrm{Var}(X_1) = \sigma^2$. Für den Erwartungswert μ werden folgende Schätzfunktionen vorgeschlagen:

$$T_1 = \bar{X},\; T_2 = X_n,\; T_3 = \frac{X_1 + X_n}{2},\; T_4 = \frac{1}{n-1}\sum_{i=1}^{n} X_i,\; T_5 = \frac{1}{n^2}\sum_{i=1}^{n} X_i.$$

(a) Welche der Schätzfunktionen sind linear?

(b) Welche der Schätzfunktionen sind unverzerrt?

(c) Welche der unverzerrten Schätzfunktionen hat die kleinste Varianz?

(d) Berechnen Sie den MSE (=mean-squared error) für alle Schätzfunktionen.

4.9 Gegeben seien zwei unabhängige Schätzer T_1, T_2 für einen unbekannten statistischen Parameter θ. Für die Schätzfunktionen gilt: $\mathbb{E}[T_1] = 3\theta$, $\mathrm{Var}(T_1) = 1$ und $\mathbb{E}[T_2] = 2\theta$, $\mathrm{Var}(T_2) = 9$. Betrachten Sie Linearkombinationen $T = a_1T_1 + a_2T_2$ der beiden Schätzer, wobei $a_1 \in \mathbb{R}$ und $a_2 \in \mathbb{R}$.

(a) Welche Bedingungen müssen die a_i erfüllen, sodass T ein unverzerrter Schätzer für θ ist.

(b) Bestimmen Sie a_i, sodass T varianzminimal und unverzerrt ist.

4.10 Eine Stichprobe der Größe 2 wird aus folgender Verteilung gezogen

$$f(y) = \begin{cases} 2y\theta^2, & 0 < y < 1/\theta \\ 0, & \text{sonst.} \end{cases}$$

Betrachten Sie den Schätzer $T = c(Y_1 + 2Y_2)$, wobei $c \in \mathbb{R}$. Für welchen Wert c ist T ein unverzerrter Schätzer für $1/\theta$?

4.11 Gegeben sind zwei normalverteilte Zufallsvariablen $X_1 \sim N(-1, 9)$ und $X_2 \sim N(1, 1)$, die unkorreliert sind. Welche der Abbildungen A, B, C, D (siehe 11) könnte den zweidimensionalen Daten $[x_1, x_2]$ entsprechen, wobei x_1 Realisationen von X_1 und x_2 Realisationen von X_2 sind?

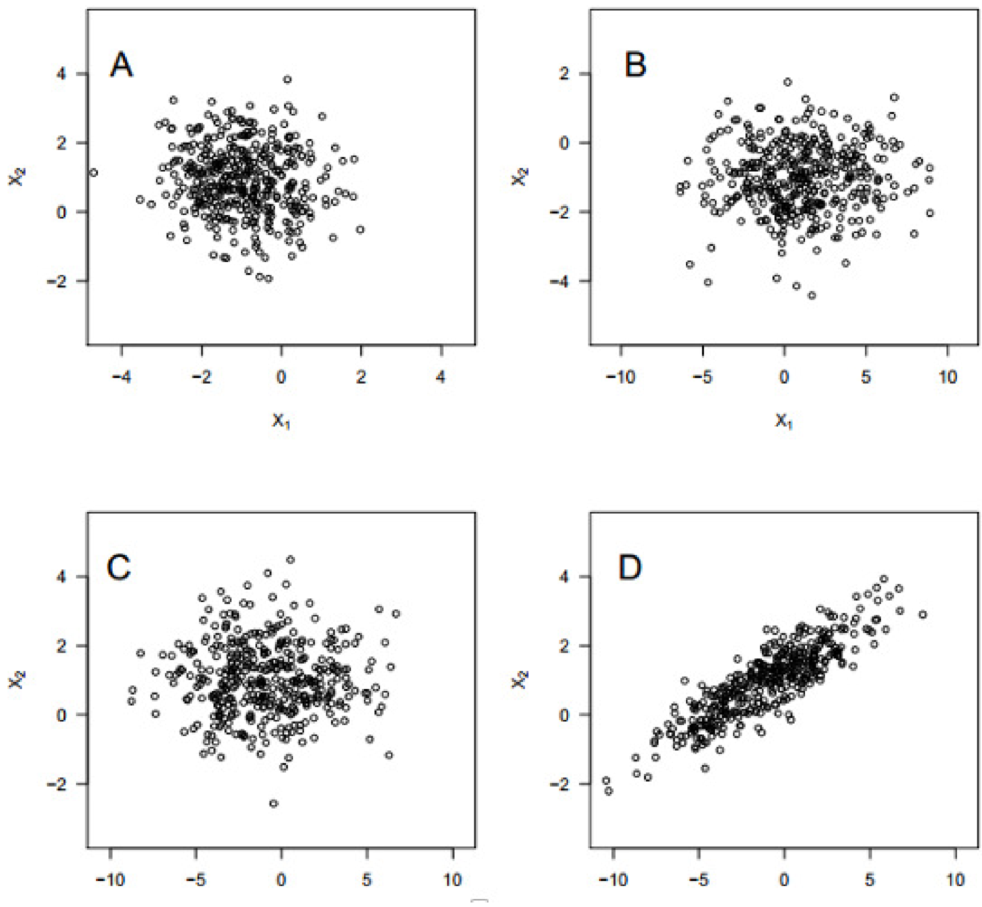

Abbildung 2: Streudiagramm der Daten $[x_1, x_2]$.

4.3 Testen

*In einigen der folgenden Aufgaben haben Sie einen Teil des Outputs der statistischen Software **R** zur Verfügung.*

4.12 (Fortsetzung des Beispiels 3.23) Es ist schon ein paar Jahre her, dass Sie Ihre Firma gegründet haben. Der Firmenvorstand will nun die Gehaltspolitik der Firma beurteilen. Dafür haben Sie eine Umfrage bei Ihren Angestellten durchgeführt. Die Ergebnisse dieser Umfrage werden Ihnen später helfen Ihre Gehaltspolitik und anschließend auch die Personalpolitik zu optimieren. Die Datenbank (*zufriedenheit.txt*) mit den Ergebnissen der Umfrage beinhaltet folgende Variablen:

Id	die Identifikationsnummer des Angestellten (1 bis 100)
Geschlecht	das Geschlecht des Angestellten (Frau, Mann)
Dauer	die Dauer, die der Angestellte in Ihren Firma besuchäftigt ist (in Jahren)
Zufriedenheit	die Zufriedenheit des Angestellten mit der Firma (zufrieden, eher zufrieden, eher unzufrieden, unzufrieden)
Ausbildung	die erreichte Ausbildung (Grundschule, Matura, Diplom)
Gehalt	monatliches Gehalt (in €)

(a) Schauen Sie sich die Daten gut an. Welche Variable(n) können wir als nicht-stetige Variable(n) behandeln?

(b) Der folgende Output zeigt die geschätzten Quantile der Variable **Gehalt**.

```
Quantil 0%     5%     10%     25%     50%     75%     90%     95%     100%
Gehalt 685.9 948.6 1038.8 1189.1 1371.8 1611.0 1875.0 1933.8 2162.1
```

Wie viel muss ein Angestellter monatlich verdienen, so dass er innerhalb der 5% der Top-Verdiener ist?

(c) In Abbildung 3 sind zwei Boxplots, getrennt für Frauen und Männer, dargestellt. In einem Boxplot kann man die Quartile ablesen.[3] Überlegen Sie sich, ob wir aus der Graphik schließen können, dass das Gehalt vom Geschlecht des Angestellten abhängt. Warum ja/nein? (Hinweis: die Mediane vergleichen)

(d) In Abbildung 4 finden Sie das Histogramm der Variable Gehalt. Überlegen Sie sich, ob Sie daraus schließen können, dass die Variable normalverteilt ist. Warum ja/nein?

4.13 In Tabelle 5 haben wir als repräsentative Stichprobe, 10 unserer Angestellten ausgewählt. X_i ist eine Zufallsvariable, die das Gehalt des i-ten Angestellten repräsentiert. Wir nehmen an, dass $X_i \sim N(\mu, \sigma^2)$ und dass X_i voneinander unabhängig sind.

(a) Benutzen Sie Tabelle 5, um einen Punktschätzer für μ und σ^2 zu bestimmen. Vergleichen Sie ihre Werte mit den Werten, die die Software **R** aus der gesamten Datenbank berechnet hat, d.h. $\bar{X}_n = 1408.45$ und $s = 308.84$.

(b) Nehmen Sie an, Sie kennen σ^2. Standardisieren Sie Ihren Punktschätzer für μ und bestimmen Sie die Verteilung der standardisierten Größe.

(c) Im Allgemeinen ist σ leider unbekannt. Deshalb müssen wir es durch eine bekannte Größe (= einen Schätzer für σ) ersetzen. Machen Sie dies für die obige standardisierte Größe und bestimmen Sie nun die Verteilung.

ID	3	4	5	6	7	8	9	10	11	12
Gehalt	1400.67	949.47	1299.6	1659.47	1191.93	1255.8	1466.87	1368.67	1168.27	1761.2

Tabelle 5: Das Gehalt von 10 ausgewählten Angestellten.

4.14 Im vorigen Beispiel wollen wir testen, ob das Gehalt im Mittel 1500 € beträgt, i.d., wir wollen die Nullhypothese $H_0 : \mu = 1500$ testen. Unter den Annahmen des vorherigen Beispiels und dass σ^2 unbekannt ist, berechnen Sie den konkreten Wert der Teststatistik und den Ablehnungsbereich für den t-Test mit Signifikanzniveau $\alpha = 10\%$ (Hinweis: siehe Seite 173 in Brannath/Futschik/Krall).

(a) Finden Sie eine passende Alternativhypothese.

(b) Entscheiden Sie, ob Sie H_0 auf dem 10% Signifikanzniveau verwerfen oder nicht.

(c) Den p-Wert des Tests finden wir im Software Output. Er beträgt 0.0815. Benutzen Sie nun den p-Wert, um zu entscheiden, ob wir H_0 verwerfen oder nicht. (Das Signifikanzniveau bleibt $\alpha = 10\%$)[4]

(d) Führen Sie nun die Berechnungen in (b) und (c) nochmals mit dem Signifikanzniveau $\alpha = 0.05$ durch. Würde sich Ihre Entscheidung ändern?

4.15 Zur Beurteilung eines Trainingsprogramms zur friedlichen Lösung sozialer Konflikte wird an 5 Personen die jeweilige Gewaltneigung vor und nach dem Programm miteinander verglichen. Zur Messung wird ein normalverteilter Index verwendet. Dabei ergaben sich folgende Werte (siehe die

[3]Falls Sie den Begriff Quartil nicht aus der Vorlesung kennen, finden Sie die Definition in dem Buch von Brannath/Futschik/Krall Seite 121 oder bei Wikipedia. Dort finden Sie auch eine Beschreibung, wie man einen Boxplot interpretiert.

[4]Ein statistischer Test liefert uns nur eine Entscheidung über H_0 egal ob wir den p-Wert verwenden oder überprrufen, ob die Teststatistik im Ablehnungsbereich liegt.

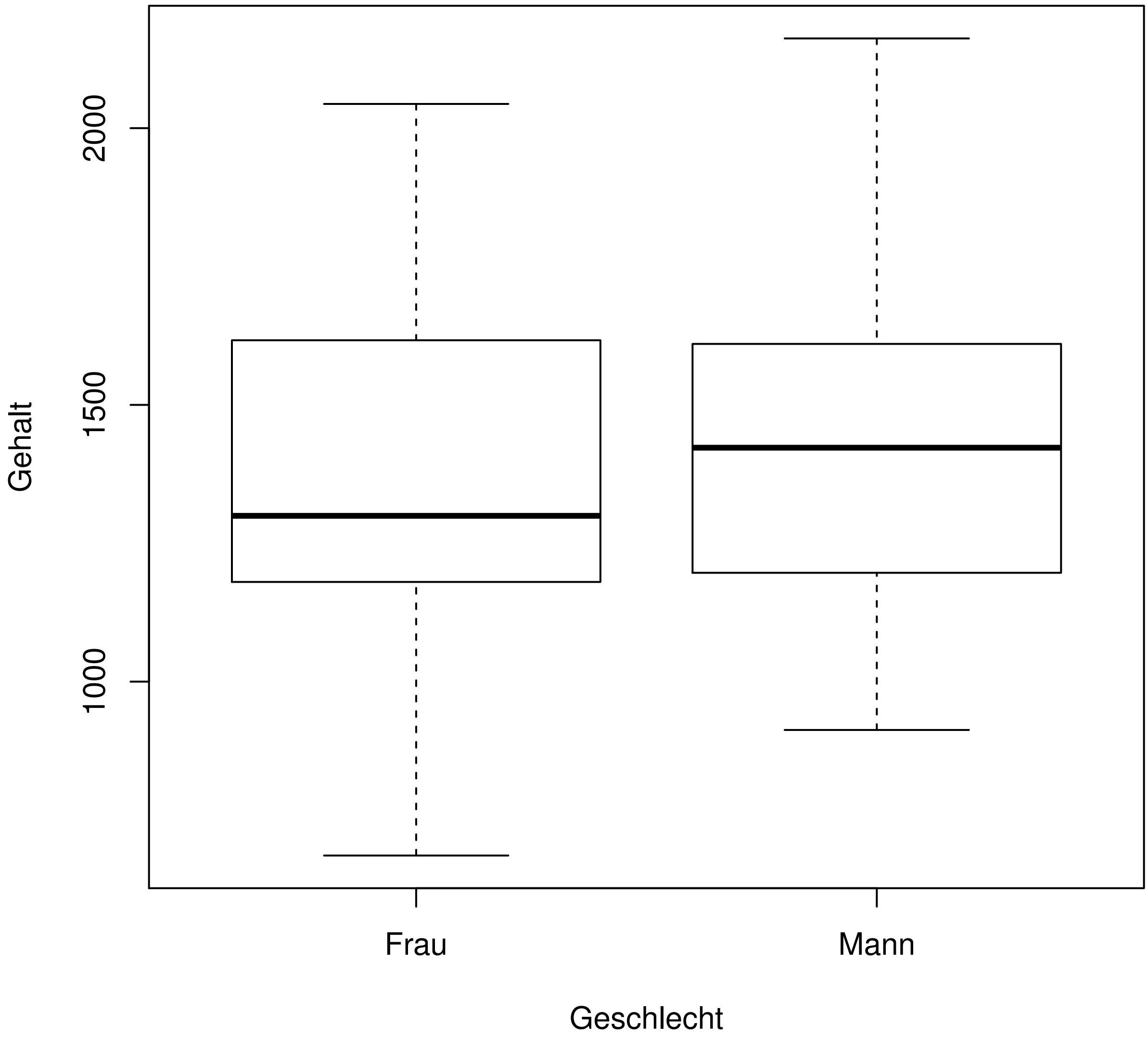

Abbildung 3: Boxplot der Variable Gehalt, getrennt für Männer und Frauen

Tabelle 6). Lässt sich zum Signifikanzniveau von 5% bestätigen, dass diese Maßnahme erfolgreich war? (Hinweis: Bilden Sie die Differenzen zwischen vorher und nachher und führen Sie einen einseitigen t-Test durch.)

ID	1	2	3	4	5
vorher	10	15	5	5	20
nachher	0	10	0	5	25

Tabelle 6: Die Messung des Gewaltindex an 5 Personen vor und nach dem Trainingsprogramm.

4.16 Wir wollen testen, ob eine Münze fair ist, i.e., $H_0 : p = 1/2$ gegen $H_1 : p \neq 1/2$. Dafür werfen wir die Münze vier mal und entscheiden uns für die Alternativhypothese, wenn bei allen

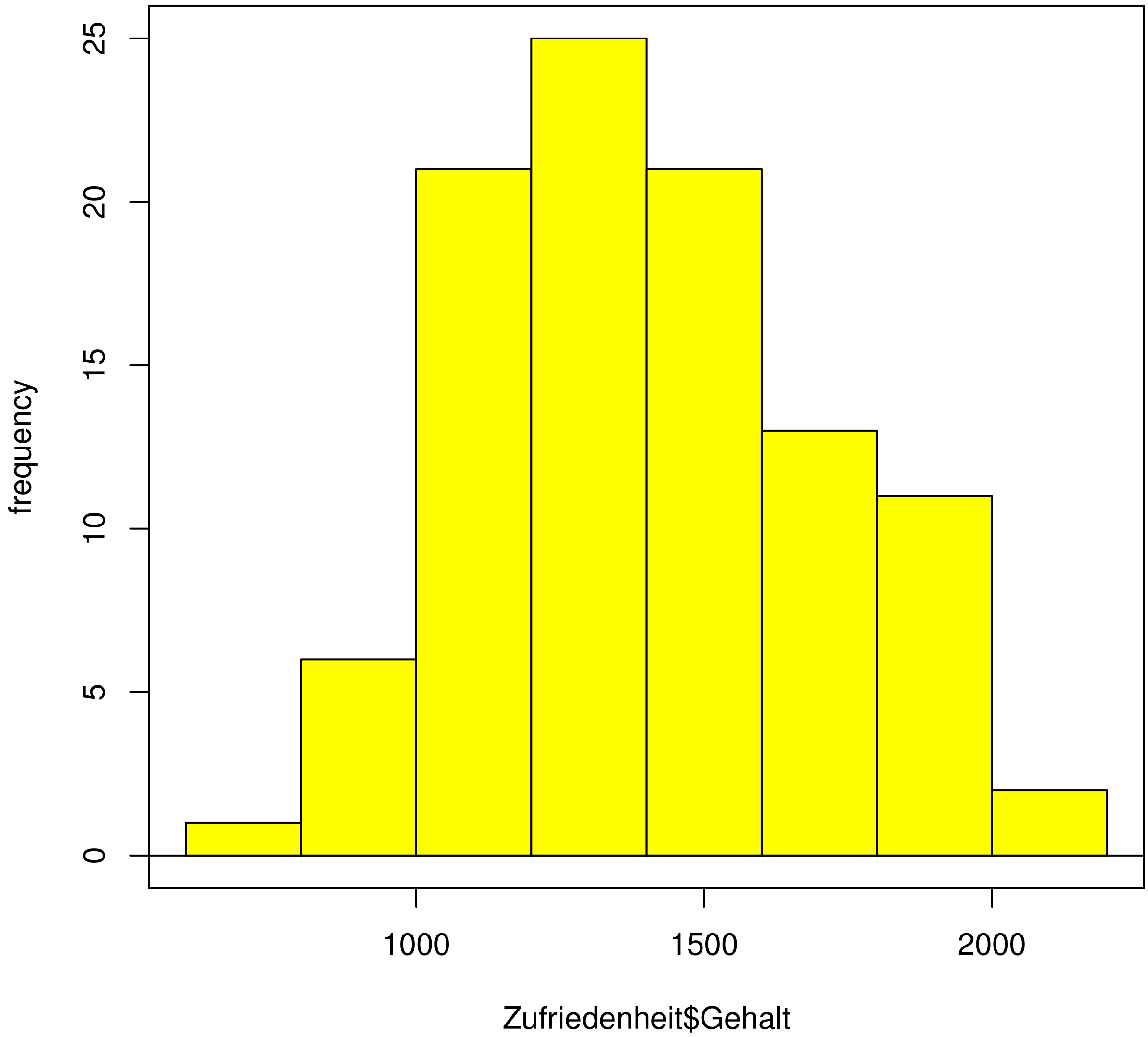

Abbildung 4: Histogram der Variable Gehalt

Münzen Kopf erscheint. In allen anderen Fällen entscheidet er sich für die Hypothese. Geben Sie die Wahrscheinlichkeit für den Fehler 1. Art an.

4.17 Der Soziologe Max Weber führt in dem Aufsatz "Zur Psychophysik der industriellen Arbeit" folgende Verteilung der Arbeitsunfälle männlicher Arbeiter in Kopen hagen 1898-1907 dar:

Wochentag	Mo	Di	Mi	Do	Fr	Sa
Anzahl	50	46	34	34	33	43

(a) Die hohen Frequenzen am Samstag und Montag hält Weber für eine Folge des Alkohols (Freitag ist Löhnungstag) bzw. für eine Folge größerer gesundheitlicher Strapazen am Wochenende. Testen Sie auf eine Gleichverteilung ($\alpha = 0.01$).

(b) Vorausgesetzt, die Stichprobe wäre 10-mal so groß wie oben und die Verteilung sähe wie folgt aus:

Wochentag	Mo	Di	Mi	Do	Fr	Sa
Anzahl	500	460	340	340	330	430

Testen Sie wie unter (a).

(c) Vergleichen Sie die gefundenen Ergebnisse miteinander und kommentieren Sie diese.

5 Lineare Regression

Beispiel						
Korrektur	-	-	-	-	-	-

Tabelle 7: Empfohlene Beispiele aus dem Buch.

In den folgenden Beispielen wird angenommen, dass die Fehler $U_1, \ldots, U_n$ i.i.d. mit Erwartungswert 0 und Varianz $\sigma^2 > 0$ sind. Weiters nehmen wir an, dass X_t nicht zufällig sind, dass $X_t \neq 0$ für ein $t = 1, \ldots, n$ und dass $X_t \neq X_s$ für mindestens ein Paar $t \neq s$ gilt.

5.1 Betrachten Sie das lineare Regressionsmodell $Y_t = a + U_t$, $t = 1, \ldots, n$.

(a) Bestimmen Sie den Kleinst-Quadrate Schätzer für a.

(b) Ist der Kleinst-Quadrate Schätzer unverzerrt?

5.2 Betrachten Sie das homogene lineare Regressionsmodell $Y_t = bX_t + U_t$, $t = 1, \ldots, n$.

(a) Bestimmen Sie den Kleinst-Quadrate Schätzer für b.

(b) Ist der Kleinst-Quadrate Schätzer unverzerrt?

5.3 Betrachten Sie das lineare Regressionsmodell $Y_t = a + bX_t + U_t$, $t = 1, \ldots, 10$. Gegeben seien die folgenden Daten

$x:$	5	6	7	8	9	10	11	12	13	40
$y:$	1	2	4	2	3	5	7	6	9	10

(a) Berechnen Sie die Kleinst-Quadrate-Schätzer $\hat{a}$ und $\hat{b}$.

(b) Veranschaulichen Sie die Einpassung der Regressionsgeraden in den Punktschwarm durch eine Zeichnung.

(c) Ermitteln und interpretieren Sie die Stichprobenkorrelation $r_{x,y}$.

(d) Führen Sie die Berechnungen in (a), (b) und (c) nochmals für folgende Daten durch:

$x:$	5	6	7	8	9	10	11	12	13	14
$y:$	1	2	4	2	3	5	7	6	9	10

Diskutieren Sie die Unterschiede!

5.4 Bestimmen Sie die unbekannten Größen c bzw. d.

(a) Betrachten Sie das lineare Regressionsmodell $Y_t = a + bX_t + U_t$, $t = 1, \ldots, 5$. Gegeben sind $\hat{U}_1 = -\frac{1}{2}$, $\hat{U}_2 = 1$, $\hat{U}_3 = 2$, $\hat{U}_4 = -3$, $\hat{U}_5 = c$ sowie $X_1 = 1$, $X_2 = 2$, $X_3 = d$, $X_4 = 0$ und $X_5 = 1$.

(b) Betrachten Sie das lineare Regressionsmodell $Y_t = a + U_t$, $t = 1, \ldots, 5$. Gegeben ist der Kleinst-Quadrate Schätzer für a, i.e., $\hat{a} = -6$ sowie $Y_1 = -5$, $Y_2 = -10$, $Y_3 = -12$, $Y_4 = -d$ und $Y_5 = 1$.

(c) Betrachten Sie das homogene lineare Regressionsmodell $Y_t = bX_t + U_t$, $t = 1, \ldots, 5$. Gegeben sind $\hat{U}_1 = 1$, $\hat{U}_2 = -\frac{1}{2}$, $\hat{U}_3 = 4$, $\hat{U}_4 = \frac{3}{2}$, $\hat{U}_5 = c$ sowie $X_1 = d$, $X_2 = 1$, $X_3 = \frac{1}{2}$, $X_4 = -1$ und $X_5 = 2$.

5.5 (a) Betrachten Sie das lineare Regressionsmodell $Y_t = a + bX_t + U_t$, $t = 1, \ldots, n$ und die Schätzer

$$\tilde{b} = \frac{\sum_{t=1}^{n} X_t Y_t}{\sum_{t=1}^{n} X_t^2} \quad \text{und} \quad \hat{b} = \frac{\sum_{t=1}^{n}(X_t - \bar{X})(Y_t - \bar{Y})}{\sum_{t=1}^{n}(X_t - \bar{X})^2}.$$

Welcher Schätzer hat eine kleinere Varianz? Warum?

(b) Betrachten Sie das homogene lineare Regressionsmodell $Y_t = bX_t + U_t$, $t = 1, \ldots, n$ und die Schätzer

$$\tilde{b} = \frac{\sum_{t=1}^{n} X_t Y_t}{\sum_{t=1}^{n} X_t^2} \quad \text{und} \quad \hat{b} = \frac{\sum_{t=1}^{n}(X_t - \bar{X})(Y_t - \bar{Y})}{\sum_{t=1}^{n}(X_t - \bar{X})^2}.$$

Welcher Schätzer hat eine kleinere Varianz? Warum?