2 Wahrscheinlichkeitstheorie

2.1 Kombinatorik, Satz der totalen Wahrscheinlichkeit, Satz von Bayes

Aufgabe 2.1.1 Souvenir ⊛ ⊛

Bradley bringt sich meist ein Tattoo als Souvenir aus dem Urlaub mit. Da er bei den meisten seiner Tätowierungen betrunken war, kann man davon ausgehen, dass er alle Tätowierer in seinem Umkreis mit gleicher Wahrscheinlichkeit aufsucht. Jetzt ist Bradley für zwei Wochen in Berlin. In seinem Umkreis gibt es die folgenden Tattoostudios:

- „Für immer und ehwig"
  Hier arbeitet Mike, der eine Lese-, Rechtschreibschwäche hat und mit einer Wahrscheinlichkeit von 45 % einen Rechtschreibfehler tätowiert.
- „Taotto"
  Der Laden gehört Analphabet Joe, welcher beim Abschreiben der Wörter in Tattoos häufig die Buchstaben verdreht. Ein Rechtschreibfehler passiert ihm sogar mit einer Wahrscheinlichkeit von 80 %.
- „Der Körper seines Geistes"
  Im dritten Tattoostudio lebt und arbeitet Jacky. Tattoos sind ihr Leben und sie macht nur selten Fehler. Ein Fehler in Orthographie oder Grammatik würde ihr nur mit einer Wahrscheinlichkeit von 5 % passieren.

(a) Bradley hat wieder über die Stränge geschlagen, geht nun zum Tätowierer und lässt sich den Namen seiner Freundin Lissy tätowieren. Mit welcher Wahrscheinlichkeit hat das fertige Tattoo keinen Rechtschreibfehler?
(b) Am nächsten Morgen sieht er das Tattoo. Er liest „Ssily". Mit welcher Wahrscheinlichkeit war er bei Joe?
(c) Angenommen die Reihenfolge der Buchstaben des Wortes „Lissy" entsteht bei Tätowierer Joe rein zufällig. Mit welcher Wahrscheinlichkeit schreibt er den Namen richtig?

P. Otto, A.-L. Lange, *Arbeitsbuch der Angewandten Statistik*,
DOI 10.1007/978-3-662-49212-3_2

Aufgabe 2.1.2 Lernen, lernen und nochmals lernen (W. I. Lenin) ⊛ ⊛

Marcel hat festgestellt, dass noch zwei Tage bis zum zweiten Termin der Statistik-Klausur verbleiben, und möchte sich nun auf diese vorbereiten und die beste Strategie zum Bestehen finden. Dazu befragt er seine Kommilitonen, die im ersten Termin geschrieben haben. Marcel weiß, dass dieser erste Termin mit einer Wahrscheinlichkeit von 70 % zu bestehen war.

Unter seinen Kommilitonen konnte er drei (sich gegenseitig ausschließende) Lernstrategien feststellen:

1. Regelmäßiges Lernen
2. Unterstützung durch Wiederholungskurse
3. Gänzlich ohne Lernen

Zunächst betrachtet er die Kommilitonen, die bestanden haben. Mit einer Wahrscheinlichkeit von 80 % gehört ein Student, der bestanden hat, der ersten Kategorie an. Mit einer Wahrscheinlichkeit von 15 % gehört ein Besteher der zweiten Kategorie an.

Marcel hat bei der Vorlesung zufällig gehört, dass man auch die Studenten, die nicht bestanden haben, betrachten sollte. Mit einer Wahrscheinlichkeit von 10 % gehört ein Nicht-Besteher zur ersten Kategorie und mit einer Wahrscheinlichkeit von 20 % gehört ein Nicht-Besteher zur zweiten Kategorie.

(a) Berechnen Sie die Wahrscheinlichkeit, dass ein Student in der Klausur sitzt und nicht gelernt hat!
(b) Marcel entscheidet sich für die zweite Lernstrategie. Wie groß ist die Wahrscheinlichkeit, dass er besteht?
(c) Wie groß ist die Wahrscheinlichkeit, dass sich unter den ersten zehn befragten Kommilitonen mindestens ein Student befindet, der nicht bestanden hat? (Gehen Sie davon aus, dass die Auswahl unabhängig und zufällig erfolgte und für alle Kommilitonen die obige Bestehenswahrscheinlichkeit gilt.)

Aufgabe 2.1.3 Verhaftung mit oder ohne Beute? ⊛ ⊛

Die Coups der Olsenbande, einer bekannten dänischen TV-Kriminalkömodie, glücken nur in den wenigstens aller Fälle. Dies liegt nicht unbedingt an den schlechten Plänen von Egon Olsen, sondern oftmals auch an der Vorbereitung der Coups durch Benny und Kield, welche mit einer Wahrscheinlichkeit von 80 % schlecht durchgeführt wird. Es gibt genau drei mögliche Ergebnisse des Coups:

- positiv – ohne Verhaftung und mit Beute
- neutral – ohne Verhaftung und ohne Beute
- negativ – mit Verhaftung und ohne Beute

Bei einer guten Vorbereitung endet der Coup mit einer Wahrscheinlichkeit von 20 % positiv. Bei einer schlechten Vorbereitung endet er nur in 5 % der Fälle positiv.

Mit einer Wahrscheinlichkeit von 30 % endet ein gut vorbereiteter Coup neutral und mit einer Wahrscheinlichkeit von 10 % endet ein schlecht vorbereiteter Coup neutral.

In allen anderen Fällen endet der Coup ohne Beute und mit einer Verhaftung von Egon Olsen.

(a) Mit welcher Wahrscheinlichkeit endet der Coup ohne Beute und mit einer Verhaftung für Egon Olsen?
(b) Zeigen Sie, dass eine gute und schlechte Vorbereitung gleichwahrscheinlich ist, wenn der Coup positiv endet!
(c) Sind die Ereignisse „Coup endet positiv“ und „Vorbereitung ist gut“ stochastisch unabhängig?

Aufgabe 2.1.4 Autohauseröffnung I ⊛ ⊛
Zur Eröffnung eines Autohauses plant der Inhaber eine große Rabattaktion in Form eines Sofortrabattes bei der Anschaffung eines neuen Fahrzeugs. Am Tag der Eröffnung sollen genau 10 Personen ausgelost werden, die den Rabatt erhalten können. Jeder Besucher kann dabei jeweils nur einmal gewinnen. Es kommen leider nur 30 Personen, unter diesen befinden sich zu allem Übel 12 Personen, die bereits einen Wagen dieser Marke fahren und keine Neuanschaffung planen.

(a) Mit welcher Wahrscheinlichkeit befindet sich unter den ausgelosten Gewinnern mindestens ein Kunde, der bereits ein Fahrzeug der Marke fährt?
(b) Das Autohaus hat 15 verschiedene Fahrzeuge auf Lager. Wie viele Möglichkeiten gibt es, die Fahrzeuge auf die 10 Gewinner zu verteilen? Jeder erhält nur ein Fahrzeug.

Aufgabe 2.1.5 Autohauseröffnung II ⊛
Zur Eröffnung eines Autohauses plant der Inhaber eine große Rabattaktion in Form eines Sofortrabattes bei der Anschaffung eines neuen Fahrzeugs. Am Tag der Eröffnung sollen genau 5 Personen ausgelost werden, die den Rabatt erhalten können. Jeder Besucher kann dabei jeweils nur einmal gewinnen. Es kommen leider nur 30 Personen, unter diesen befinden sich zu allem Übel 12 Personen, die bereits einen Wagen dieser Marke fahren und keine Neuanschaffung planen.

(a) Mit welcher Wahrscheinlichkeit befindet sich unter den ausgelosten Gewinnern mindestens ein Kunde, der bereits ein Fahrzeug der Marke fährt?
(b) Leider ist die Wahrscheinlichkeit gering, dass man gewinnt und der Händler das persönliche Lieblingsfahrzeug auf Lager hat. Sie beträgt nur 10 %. Wie groß ist die Wahrscheinlichkeit, dass das persönliche Lieblingsfahrzeug auf Lager ist, wenn man bei dem Gewinnspiel gewinnt?

Aufgabe 2.1.6 Iskander ⊛ ⊛ ⊛
Bereits in altägyptischen Gräbern von 3500 v. Chr. wurden manipulierte Würfel gefunden. Der alte Ägypter Iskander war sehr gerissen, dessen Würfel besaß die folgende Wahrscheinlichkeitsverteilung:

$$P(\{1\}) = 0{,}1, \quad P(\{2\}) = P(\{3\}) = P(\{4\}) = P(\{5\}) = 0{,}15, \quad P(\{6\}) = 0{,}3$$

Definiert seien die beiden Ereignisse A und B:

A – Augenzahl ist gerade
B – Augenzahl ist ein Vielfaches von 3

(a) Bestimmen Sie die Wahrscheinlichkeit für das Auftreten einer geraden Augenzahl!
(b) Lässt sich die Wahrscheinlichkeit für das Auftreten des Ereignisses A als Laplace-Wahrscheinlichkeit ermitteln? Begründen Sie Ihre Aussage in einem Satz!
(c) Bestimmen Sie $P(A \cup B)$ und $P(A \cap B)$!
(d) Bestimmen Sie $P(A|B)$!
(e) Unter welcher Bedingung gilt $P(A|B) = P(A)$? Begründen Sie in einem Satz!
(f) Überprüfen Sie, ob A und B stochastisch unabhängig sind!
(g) Welcher Zusammenhang besteht zwischen der relativen Häufigkeit und der Wahrscheinlichkeit? Erläutern Sie kurz mit eigenen Worten!
(h) Es gelten die Axiome von Kolmogorov. Zeigen Sie, dass für $C \subseteq A$ gilt, dass $P(C) \leq P(A)$!

Aufgabe 2.1.7 Hütchenspieler Alejandro ⊛ ⊛
An der Strandpromenade spielt der Hütchenspieler Alejandro mit dem Passanten Horst-Günther folgendes Spiel. Er hat 5 Hütchen und 2 rote, 2 blaue und eine gelbe Kugel. Diese Kugeln versteckt er unter den Hütchen, dreht die Hütchen und ordnet sie unabhängig in einer Reihe von links nach rechts.

(a) Wie viele mögliche Anordnungen der Kugeln gibt es, wenn alle Kugeln verschiedene Farben hätten?
(b) Wie viele mögliche Anordnungen der Kugeln mit den oben genannten Farben gibt es?
(c) Horst-Günther gewinnt das Spiel, wenn er die gelbe Kugel aufdeckt, wobei er drei Hütchen wählen darf. Mit welcher Wahrscheinlichkeit gewinnt er das Spiel, wenn er die drei Hütchen zufällig auswählt?
(d) Mit welcher Wahrscheinlichkeit ist die erste gezogene Kugel blau?
(e) Es gibt insgesamt 18 Kombinationen, bei denen Horst-Günther gewinnt, bei 6 dieser Kombinationen ist dabei die erste Kugel blau.
Horst-Günther ist außerdem ein strategischer Spieler und stellt sich die folgende Frage: „Wie groß ist die Wahrscheinlichkeit, dass ich gewinne, wenn ich als erstes eine blaue Kugel aufdecke?" Beantworten Sie ihm die Frage!
(f) Gehen Sie davon aus, dass alle Spiele unabhängig erfolgen. Mit welcher Wahrscheinlichkeit gewinnt Horst-Günther erst beim dritten Spiel?

Aufgabe 2.1.8 Du zahlst ⊛ ⊛ ⊛

Bei der Bezahlung des Urlaubes ziehen das Ehepaar Anton und Beate gern eine Münze zu Rate. Sie werfen die Münze immer zweimal hintereinander. Die beiden Würfe sind unabhängig. Anton wird dabei das Ereignis $A = \{(K, K), (K, Z)\}$ (d. h. das Ereignis, dass beim ersten Wurf Kopf kommt) und Beate das Ereignis $B = \{(K, K), (Z, K)\}$ (d. h. das Ereignis, dass beim zweiten Wurf Kopf auftritt) zugeordnet.

(a) Bestimmen Sie die Grundgesamtheit Ω für den zweimaligen Münzwurf!

(b) Ergänzen Sie die nachfolgende Tabelle mit den entsprechenden Erläuterungen oder Mengendarstellungen!

Mengendarstellung	Ereignis, dass
	mindestens einmal Kopf auftritt
	zweimal Kopf auftritt
$\bar{A} \cup \bar{B}$	mindestens einmal Zahl auftritt
$\bar{A} \cap \bar{B}$	
$(A \cap \bar{B}) \cup (\bar{A} \cap B)$	
$(A \cap B) \cup (\bar{A} \cap \bar{B})$	

(c) Bestimmen Sie die Wahrscheinlichkeit für das Ereignis A!

(d) Sei $P(\cdot)$ ein Wahrscheinlichkeitsmaß und es gelten die Axiome von Kolmogorov. Zeigen Sie, dass $P(\bar{A}) = 1 - P(A)$ gilt!

(e) Gilt $P(A|B) = P(A)$? Interpretieren Sie das Ergebnis bezüglich der Unabhängigkeit von A und B!

(f) Anton und Beate wiederholen den Wurf zweier Münzen zehnmal. Anton interessiert dabei, wie oft er gewinnen wird. Sei X die Anzahl, wie oft Anton beim n-maligen Wurf gewinnt. Wie ist X verteilt? Geben Sie wenn nötig auch die Parameter der Verteilung an!

(g) Von welcher Verteilung spricht man in (f), wenn $n = 1$ ist?

(h) Spricht man bei der obigen Verteilung aus (f) von einer diskreten oder stetigen Verteilung? Begründen Sie!

(i) Die Verteilungsfunktion einer Zufallsvariablen X sei gegeben durch

$$F_X(x) = P\left(\{\omega \in \Omega : X(\omega) \leq x\}\right), \quad x \in \mathbb{R}.$$

Was verstehen Sie unter einer Verteilungsfunktion? Erläutern Sie in einem Satz!

(j) Bestimmen Sie den Funktionswert der Verteilungsfunktion einer Normalverteilung mit den Parametern $\mu = 2$ und $\sigma^2 = 4$ an der Stelle 3!

(k) Warum ist die folgende Funktion G keine Verteilungsfunktion?

$$G(x) = \begin{cases} 1 - \frac{1}{x} & \text{für} \quad x \in (0, \infty) \\ 0 & \text{für} \quad x \in (-\infty, 0] \end{cases}$$

(l) Nennen Sie einen sinnvollen Schätzer für die Verteilungsfunktion!

Aufgabe 2.1.9 Klausurendoping ⊛ ⊛ ⊛ ⊛

In den letzten Jahren hat sich unter Studenten, die eine Klausur gerade noch bestehen wollen, eine neue leistungssteigernde Droge verbreitet.

Leider wird sie häufig nur von Studenten genommen, bei denen das Bestehen bereits fast aussichtslos ist. Unter diesen mit einer Wahrscheinlichkeit von 90 %. Zur Vereinfachung sei anzunehmen, dass die Wahrscheinlichkeit, dass ein guter Student das Doping vornimmt, vernachlässigbar gering ist.

Mit einer Wahrscheinlichkeit von 10 % gehört ein Student zur Gruppe der Aussichtslosen.

Die Wirksamkeit des Doping ist sehr hoch. Sie steigert die Bestehenswahrscheinlichkeit um das 15-fache. Folgende Bestehenswahrscheinlichkeiten sollen angenommen werden:

- Wenn das Bestehen bereits fast aussichtslos ist und er keine Drogen nimmt: 0,01.
- Wenn der Student kein aussichtsloser Fall ist: 0,8.

(a) Mit welcher Wahrscheinlichkeit besteht man die Klausur?

(b) Mit welcher Wahrscheinlichkeit ist der Student „gedopt", wenn er die Klausur nicht besteht?

2.2 Univariate Zufallsgrößen

Aufgabe 2.2.1 Noch ein Bier ⊛ ⊛ ⊛

Im All-Inclusive-Urlaub nutzen die meisten Gäste am Abend nur noch maximal N Worte. Der Barkeeper, ein studierter Linguist, ist von der Sprache fasziniert und versucht sich der quantitativen Linguistik. Er erinnert sich von seinem Studium an das Zipf'sche Gesetz. Sei $X \in \{1, 2, \ldots, N\}$ eines der N Wörter. Gegeben sei weiterhin die Wahrscheinlichkeitsfunktion f, welche das Auftreten des k-häufigsten Wortes beschreibt und vom Parameter $s \in \mathbb{N}$ abhängt. Sie sei

$$f_s(k) = P(X = k) = \frac{k^{-1}}{\sum_{i=1}^{N} i^{-s}}, \quad k \in \{1, 2, \ldots, N\}.$$

(a) Bestimmen Sie den Parameter s so, dass die Funktion f eine Wahrscheinlichkeitsfunktion sein kann!

(b) Bestimmen Sie die folgenden Wahrscheinlichkeiten! Gehen Sie davon aus, dass $s = 1$ und $N = 6$ ist.
   i. $P(X \leq 3)$
   ii. $P(1 \leq X < 4)$
   iii. $P(X = 1)$

(c) Bestimmen Sie den Erwartungswert und die Varianz von X für den Parameter $s = 1$ und beliebige $N \in \mathbb{N}$! Hinweis: $\sum_{i=1}^{N} i = \frac{1}{2}N(N+1)$

Aufgabe 2.2.2 Beta-Verteilung ⊛ ⊛ ⊛
Gegeben sei die BETA-Verteilung

$$f(x) = cx^{\alpha-1}(1-x)^{\beta-1}, \quad x \in [0,1],$$

wobei die beiden Parameter $\alpha, \beta > 0$ und $c \in \mathbb{R}$.

(a) Bestimmen Sie den Paramater c so, dass die Funktion f eine Dichtefunktion sein kann! Wählen Sie dabei die Parameter $\alpha = 1$ und $\beta = 3$.
(b) Bestimmen Sie für $\alpha = \beta = c = 1$ die folgenden Wahrscheinlichkeiten!
   i. $P(X \geq 0{,}6)$
   ii. $P(X = 0{,}5)$
   iii. $P(X < 1{,}8)$
(c) Bestimmen Sie den Erwartungswert von X für beliebige $c \in \mathbb{R}$ und $\alpha > 0$! Wählen Sie dabei den Parameter $\beta = 3$.

Aufgabe 2.2.3 Bis die Maschine ausfällt ⊛ ⊛
In der Qualitätskontrolle wird zur Beurteilung der Zuverlässigkeit von Maschinen die mittlere Betriebsdauer X zwischen zwei Betriebsausfällen in Tagen (mean-time-between-failures (MTBF)) herangezogen. Diese sei mit folgender Dichte exponentialverteilt.

$$f(x) = \lambda e^{-\lambda x}, \quad x \geq 0, \lambda > 0$$

(a) Zeigen Sie, dass f unabhängig von der Parameterwahl λ eine Dichtefunktion von X ist.
(b) Im folgenden sei die MTBF der Maschine $X \sim \text{Exp}(\lambda = 0{,}2)$.
   Berechnen Sie die Wahrscheinlichkeit, dass die mittlere Betriebsdauer:
   i. mindestens 10 Tage beträgt.
   ii. weniger als 8 Tage beträgt.
   iii. zwischen 5 und 15 Tagen liegt.
(c) Für welche Parameter $\lambda > 0$ existiert der Erwartungswert von e^X?

Aufgabe 2.2.4 Rudolph und Gisela ⊛ ⊛ ⊛ ⊛
Rentier Rudolph und Henne Gisela streiten sich wie jedes Jahr, wer mehr Geschenke im Auftrag ihrer Chefs verteilen muss. Um die Arbeit vergleichbar zu machen, einigen sie sich die Menge aller Geschenke, die auf einen Schlitten oder in ein Eier-Körbchen passen, als eine Zufallsgröße X in kg zu bezeichnen.

(a) Beide wissen zunächst nicht, wie die Geschenkemenge X verteilt ist. Sie wissen lediglich, dass man eine Geschenkemenge von 25 kg bei einer Varianz von 5 (kg)^2 erwarten kann. Wie groß ist die Wahrscheinlichkeit mindestens, dass X im Intervall $(21, 29)$ liegt?

(b) Rudolph musste in diesem Jahr insgesamt 5.000 von verschiedenen Wichteln unabhängig bepackte Schlitten ziehen. Gehen Sie davon aus, dass die Geschenkemenge auf jedem Schlitten der gleichen Verteilung mit dem Erwartungswert und der Varianz aus Aufgabenteil (a) folgt. Wie groß ist die Wahrscheinlichkeit näherungsweise, dass er mehr als 125,2 t Geschenke ausgeliefert hat? Hinweis: 1 t = 1.000 kg

(c) Gehen Sie im Folgenden davon aus, dass X mit identischem Erwartungswert und gleicher Varianz normalverteilt ist. Bestimmen Sie die Wahrscheinlichkeit aus Aufgabenteil (a) unter dieser Zusatzannahme!

(d) Gisela zweifelt an Rudolphs Rechnungen und hat selber das Gewicht der letzten 100 Schlitten gemessen. Die Geschenkemenge dieser 100 Schlitten betrug insgesamt 2.550 kg. Sie möchte überprüfen, ob der Erwartungswert 25 kg signifikant überschritten wird. Führen Sie einen geeigneten Test zum Signifikanzniveau $\alpha = 0{,}05$ durch! Gehen Sie davon aus, dass die Varianz von 5 $(\text{kg})^2$ bekannt ist. *vgl.* Abschn. 3.3.1.

(e) Rudolphs Chef Claus weiß, dass der wahre Erwartungswert der Geschenkemenge X exakt 24,78 kg beträgt. Bestimmen und interpretieren Sie für das Testproblem aus (d) die Wahrscheinlichkeit H_0 abzulehnen! *vgl.* Abschn. 3.3.1.

Aufgabe 2.2.5 Plausibilitätsprüfungen ⊛ ⊛

Finanzämter überprüfen teilweise die Plausibilität von Bilanzdaten anhand der Verteilung der ersten Ziffer aller Zahlen in einer Bilanz. Sei X die erste Ziffer einer beliebigen Zahl. Gegeben sei für $d \in \{1, 2, \ldots, 9\}$ die folgende Wahrscheinlichkeitsfunktion f.

$$f(d) = P(X = d) = \log_B \left(\frac{d+1}{d} \right)$$

(a) Zeigen Sie, dass für die Basis $B = 10$ die oben gegebene Funktion f eine Wahrscheinlichkeitsfunktion ist!

(b) Berechnen Sie die Wahrscheinlichkeit, für die folgenden Ereignisse:
   i. Die erste Ziffer ist 4.
   ii. Die erste Ziffer ist größer als 2, aber kleiner als 5.
   iii. Die erste Ziffer ist größer 2.

(c) Berechnen Sie den Erwartungswert und die Varianz von X!

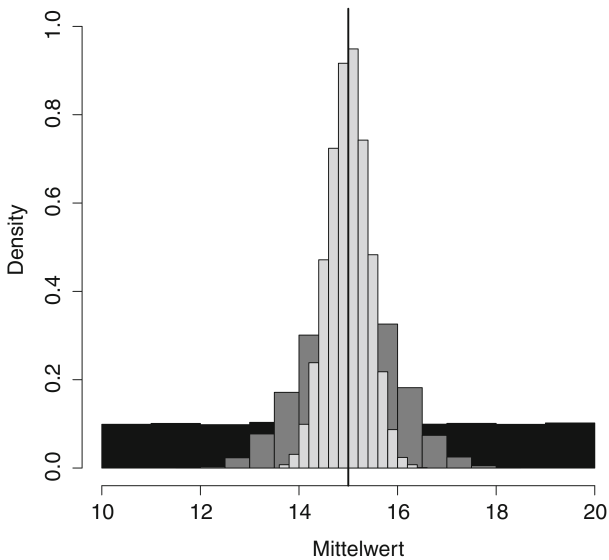
Aufgabe 2.2.6 Simulationsstudie ⊛ ⊛ ⊛ ⊛

In einer Simulationsstudie soll das schwache Gesetz der großen Zahlen veranschaulicht werden.

Es seien $X_1, X_2, \ldots$ unabhängig und identisch verteilte Zufallsvariablen mit $E(X_i) = \mu$. Dann gilt für alle $\varepsilon > 0$

$$\lim_{n \to \infty} P\left(|\bar{X} - \mu| > \varepsilon\right) = 0\,.$$

Hierzu wurde eine Simulationsstudie durchgeführt, welche $n = (1, 10, 50)$ auf dem Intervall $[a, b]$ gleichverteilte Zufallszahlen simuliert und das Ergebnis in einem Histogramm darstellt.




```
> m <- 10^4
> a <- 10
> b <- 20
> mu <- 0.5 * (a + b)
> # Definition der Vektoren x_bar
> x_bar_1 <- numeric(m)
> x_bar_10 <- numeric(m)
> x_bar_50 <- numeric(m)
> # Simulationsstudie
> for (i in 1:m)
+ {
+   x_bar_1[i] <- mean(runif(1, min = a, max = b))
+   x_bar_10[i] <- mean(runif(10, min = a, max = b))
+   x_bar_50[i] <- mean(runif(50, min = a, max = b))
+ }
> # Abbildung Histogramme
> hist(x_bar_1, prob = TRUE, breaks = 12, col = gray(0.2), ylim = c(0,1)
   , xlim = c(a,b), main = "", xlab = "Mittelwert")
> hist(x_bar_10, prob = TRUE, breaks = 12, col =  gray(0.5), add = TRUE)
> hist(x_bar_50, prob = TRUE, breaks = 12, col = gray(0.8), add = TRUE)
> abline(v = mu, lwd = 2)
> # Signifikanztests
> library("TeachingDemos")
> sigma.test(x_bar_1, sigmasq = 8, alternative = "two.sided")

        One sample Chi-squared test for variance

data:  x_bar_1
X-squared = 10353.85, df = 9999, p-value = 0.01296
alternative hypothesis: true variance is not equal to 8
95 percent confidence interval:
 8.058997 8.518419
```

```
32 sample estimates:
33 var of x_bar_1
34       8.283911
35
36 > var.test(x_bar_10, x_bar_50, alternative = "greater")
37
38         F test to compare two variances
39
40 data:  x_bar_10 and x_bar_50
41 F = 4.8726, num df = 9999, denom df = 9999, p-value < 2.2e-16
42 alternative hypothesis: true ratio of variances is greater than 1
43 95 percent confidence interval:
44  4.714935      Inf
45 sample estimates:
46 ratio of variances
47           4.872638
```

(a) Für wie viele Simulationswiederholungen wurde die Studie durchgeführt? Nennen Sie die entsprechende Zeile im `R`-Code!
(b) Wie viele Werte hat der Vektor `x_bar_50`? Begründen Sie Ihre Antwort!
(c) Ergänzen Sie in der obigen Abbildung bei jedem der drei Histogramme die Zeilennummer des Codes, mit welchem es erzeugt wurde!
(d) Stellen Sie die Hypothesen des in Zeile 23 durchgeführten Tests auf! Treffen Sie die Entscheidung zu einem Signifikanzniveau von $\alpha = 0{,}01$!
(e) Welcher Test wird in Zeile 36 durchgeführt? Bestimmen Sie die Hypothesen und erläutern Sie in eigenen Worten, was untersucht werden soll!
(f) Zu welcher Entscheidung kommt der in Zeile 36 durchgeführte Test ($\alpha = 0{,}01$)?
(g) Sie möchten die Ergebnisse reproduzierbar machen und wollen einen Startwert für die Simulation setzen. Was müssen Sie am gegebenen Code ändern/ergänzen? Geben Sie auch an, an welcher Stelle die Änderung(en) zu erfolgen hat(haben).

Geben Sie für die folgenden Problemstellungen die jeweilige Funktion in `R` an!

(h) Sie möchten für `x_bar_10` und `x_bar_50` testen, ob die Erwartungswerte gleich sind.
(i) Sie wollen für die Funktion `hist()` die Hilfe aufrufen.
(j) Wie müssen Sie die Zeilen 12–14 ändern, damit $X_1, X_2, \ldots$ unabhängig und mit $\mu = 10$ und $\sigma = 5$ identisch normalverteilt ist.

Aufgabe 2.2.7 In der Weihnachtsbäckerei ⊛ ⊛ ⊛ ⊛

In der Weihnachtsbäckerei für Kinder werden die Mengen von Mehl und Zucker oft nur abgeschätzt. Der Inhaber der „Bäckerei Gutbrodt" beschäftigt sich in seiner Freizeit gern mit Statistik und bezeichnet mit einer Zufallsvariable X die Menge Mehl in g, welche in eine Kinderhand passt.

(a) Herr Gutbrodt weiß, dass eine Handvoll Mehl, welche in eine Kinderhand passt, immer zwischen 0 g und 30 g liegt. Angenommen, die Menge Mehl sei auf diesem Intervall gleichverteilt, bestimmen Sie den Erwartungswert und die Varianz von X!

(b) Wie groß ist die Wahrscheinlichkeit mindestens, dass die Menge Mehl um weniger als 12 g nach oben bzw. unten vom Erwartungswert abweicht, wenn man von einer beliebigen Verteilung ausgeht, aber Erwartungswert und Varianz unverändert sind?
(c) Bestimmen Sie die approximative Wahrscheinlichkeit, dass 100 Handvoll Mehl weniger als 1.400 g sind! Gehen Sie davon aus, dass jedes Kind einmal eine Handvoll Mehl greift und X für jedes Kind unabhängig der gleichen Verteilung folgt.
(d) Um die Mehlmenge in einer Kinderhand statistisch zu untersuchen, hat Herr Gutbrodt 50 Kinder jeweils eine Handvoll Mehl in eine Schüssel füllen lassen. Insgesamt haben die Kinder dadurch 800 g Mehl in die Schüssel gefüllt. Überprüfen Sie, ob die erwartete Menge einer Handvoll Mehl über 14 g liegt. Gehen Sie davon aus, dass X mit dem Erwartungswert $\mu = 15\,g$ und der Varianz $\sigma^2 = 25\,\text{g}^2$ normalverteilt ist. Führen Sie einen geeigneten Test zum Signifikanzniveau $\alpha = 0{,}05$ durch! *vgl.* Abschn. 3.3.1.
(e) Bestimmen und interpretieren Sie für dieses Testproblem die Wahrscheinlichkeit H_0 abzulehnen! Hinweis: Der wahre Erwartungswert ist aus Aufgabenteil (a) bekannt. *vgl.* Abschn. 3.3.1.

Aufgabe 2.2.8 Schwaches Gesetz der großen Zahlen ⊛ ⊛ ⊛ ⊛ ⊛

Simulieren Sie $n = (1, 10, 50)$ Zufallsvariablen, die auf dem Intervall $[0, 1]$ gleichverteilt sind. Veranschaulichen Sie das schwache Gesetz der großen Zahlen mittels einer Simulationsstudie. Bestimmen Sie für 500 Wiederholungen $\bar{X}$ und analysieren Sie die Verteilung. Erzeugen Sie eine Grafik, welche die Histogramme von $\bar{X}$ für alle n in verschiedenen Farben zeigt. (`set.seed(1)`)

Wiederholen Sie die Aufgabe für exponentialverteilte Zufallsvariablen ($\lambda = 5$). Wählen Sie hierbei allerdings $n = (100, 1.000, 10.000)$. *vgl.* Abschn. 1.2.

Aufgabe 2.2.9 Zentraler Grenzwertsatz ⊛ ⊛ ⊛ ⊛ ⊛

Veranschaulichen Sie den Zentralen Grenzwertsatz, indem Sie je 25 Zufallsvariablen 1.000-mal simulieren und die (normierten) Mittelwerte der simulierten Werte in einem Histogramm darstellen. Vergleichen Sie die Histogramme mit der Dichtefunktion der Normalverteilung. Führen Sie die Simulationstudie für gleich-, binomial- und χ^2-verteilte Zufallsvariablen durch. Wählen Sie die nötigen Parameter beliebig. (`set.seed(65535)`) *vgl.* Abschn. 1.2.

Aufgabe 2.2.10 Erlang-Verteilung ⊛ ⊛ ⊛ ⊛ ⊛ ⊛

Gegeben seien n unabhängig und identisch ERLANG-verteilte Zufallsvariablen.

$$X_1, \ldots, X_n \overset{\text{i.i.d}}{\sim} \mathrm{Erl}(\lambda, k), \quad k \in \mathbb{N} \setminus \{0\}, \lambda > 0$$

Veranschaulichen Sie in einer Simulationsstudie, dass der Schätzer $\hat{\lambda} = \frac{k}{\bar{x}}$ ein erwartungstreuer und schwach konsistenter Schätzer ist!

Wählen Sie für jeden der Simulationsdurchgänge den Parameter k zufällig aus dem Intervall $[1, 10]$ und wählen Sie $\lambda = 3$. Simulieren Sie die Zufallszahlen $N = 10^4$-mal

(Simulationswiederholungen). Erhöhen Sie den Stichprobenumfang n in vier Schritten, $n = (20, 50, 100, 500)$. Stellen Sie für alle n die ermittelten Schätzungen mithilfe eines Histogrammes dar! Zeichnen Sie alle Darstellungen in eine Grafik! Setzen Sie den Startwert auf `set.seed(1020)`. *vgl.* Abschn. 1.2.

Hilfestellung: Nutzen Sie bei der Generierung der Zufallszahlen den Zusammenhang, dass die ERLANG-Verteilung eine GAMMA-Verteilung mit Parametern $r = k$ und $\lambda_G = 1/\lambda_E$ ist.

```
rgamma(n, shape = k, scale = 1 / lambda)
```

Aufgabe 2.2.11 Pareto-Verteilung ⊛ ⊛ ⊛

Im Risikocontrolling von Versicherungsgesellschaften spielt die Betrachtung von Versicherungsfällen mit sehr hohen Schadenssummen eine große Rolle.

Die Zufallsvariable X gebe die Schadenshöhe in Mio. € an, die eine bestimmte Mindesthöhe m überschreitet. Die dazugehörige Dichte f einer PARETO-Verteilung für die Parameter $k > 0$, $m > 0$ sei gegeben durch

$$f(x) = \begin{cases} \frac{k}{m}\left(\frac{m}{x}\right)^{k+1} & x \geq m \\ 0 & x < m \end{cases}$$

(a) Zeigen Sie, dass die dazugehörige Verteilungsfunktion

$$F(x) = 1 - \left(\frac{m}{x}\right)^k, \quad x \geq m$$

ist!

(b) Das Versicherungsunternehmen V interessiert sich für Versicherungsfälle, die 1 Mio. € überschreiten, somit ist $X \sim \text{Par}(k = 1, m = 1)$. Bestimmen Sie die folgenden Wahrscheinlichkeiten
   i. $P(X \geq 1{,}5)$
   ii. $P(1 \leq X \leq 1{,}5)$
   iii. $P(1{,}5 < X < 4)$

(c) Bestimmen Sie den Erwartungswert von X für die Parameter $k > 1$ und $m > 0$!

Aufgabe 2.2.12 Laplace-Verteilung ⊛ ⊛

Eine mit $E(X) = 0$ LAPLACE-verteilte Zufallsvariable X hat die Dichtefunktion:

$$f(x) = \frac{1}{2\lambda}e^{-\frac{|x|}{\lambda}}, \quad x \in \mathbb{R}, \lambda > 0$$

(a) Zeigen Sie, dass f unabhängig von λ eine Dichtefunktion ist!

(b) Zeigen Sie, dass die dazugehörige Verteilungsfunktion

$$F(x) = \begin{cases} \frac{1}{2}e^{\frac{x}{\lambda}} & x \leq 0 \\ 1 - \frac{1}{2}e^{-\frac{x}{\lambda}} & x > 0 \end{cases}$$

ist!

2.3 Bivariate Zufallsgrößen

Aufgabe 2.3.1 Unabhängigkeit ⊛ ⊛

Sind die Zufallsvariablen X_1 und X_2 stochastisch unabhängig, wenn ihre Verteilungsfunktion

$$F(x_1, x_2) = \frac{1}{1 + e^{-x_1} + e^{-x_2} + e^{-x_1 - x_2}}$$

ist?

Aufgabe 2.3.2 Portfolio ⊛ ⊛ ⊛

Im Folgenden werde ein Portfolio bestehend aus zwei Aktien mit den relativen Anteilen ω und $(1 - \omega)$ betrachtet. Die Renditen R_1 und R_2 seien zweidimensional normalverteilt. Sei $E(R_1) = \mu_1$, $E(R_2) = \mu_2$, $\mathrm{Var}(R_1) = \sigma_1^2$, $\mathrm{Var}(R_2) = \sigma_2^2$ und $\mathrm{Corr}(R_1, R_2) = \rho$.

(a) Bestimmen Sie die erwartete Portfoliorendite R_P!
(b) Wie ist die Portfoliorendite R_P verteilt? Geben Sie auch die Parameter der Verteilung an.
(c) Bestimmen Sie für $\omega = 0{,}7$, $\mu_1 = 0{,}05$, $\mu_2 = 0{,}08$, $\sigma_1 = 0{,}06$ und $\sigma_2 = 0{,}1$ ein Intervall für die Portfoliostandardabweichung $\sqrt{\mathrm{Var}(R_P)}$!
(d) Welche der folgenden Aussage(n) ist(sind) bezüglich der Abb. 2.1 wahr?
- ☐ Es ist die Verteilungsfunktion einer zweidimensionalen Zufallsvariable dargestellt.
- ☐ Es ist die Verteilungsfunktion einer dreidimensionalen Zufallsvariable dargestellt.
- ☐ Es ist die Dichtefunktion einer zweidimensionalen Zufallsvariable dargestellt.
- ☐ Es ist die Dichtefunktion einer dreidimensionalen Zufallsvariable dargestellt.
- ☐ Es ist weder eine Verteilungs- noch eine Dichtefunktion dargestellt.

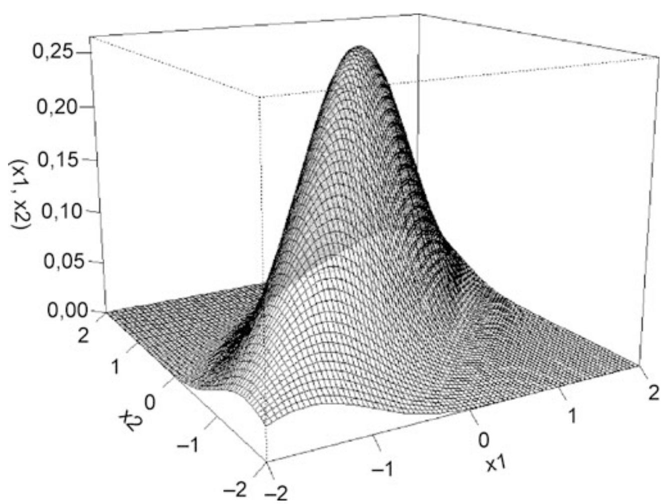


Abb. 2.1 Grafik zur Aufgabe 2.3.2 (d)

- ☐ Der Erwartungswert μ beträgt in etwa 0,25.
- ☐ Der Erwartungswert $\mu = (\mu_1, \mu_2)$ ist $(0, 0)$.
- ☐ Der Erwartungswert $\mu = (\mu_1, \mu_2)$ ist $(2, 0)$.
- ☐ Der Erwartungswert $\mu = (\mu_1, \mu_2)$ ist $(0, 2)$.
- ☐ Der Erwartungswert $\mu = (\mu_1, \mu_2)$ ist $(2, 2)$.
- ☐ Der Erwartungswert ist in diesem Fall nicht definiert.
- ☐ Die Zufallsvariable ist diskret.
- ☐ Die Darstellung ist nicht sinnvoll, da sie nicht punktsymmetrisch um den Punkt $(0, 0)$ ist.

(e) Sind die Zufallsvariablen in Abb. 2.1 korreliert? Wenn ja, treffen Sie die Aussage, ob es sich um eine positive oder negative Korrelation handelt. Begründen Sie!

(f) Erläutern Sie mit je einem Satz die Funktionen `plot` und `lines`!

(g) Die Verteilungsfunktion von (R_1, R_2) nimmt an der Stelle $(-0{,}01, -0{,}02)$ den Wert 0,159 an. Was sagt dieser Wert aus?

(h) Im Folgenden soll vereinfacht angenommen werden, dass R_1 und R_2 standardnormalverteilt sind. Gegeben sind die Dichtefunktion von (R_1, R_2) sowie die Randdichtefunktion von R_1 und R_2.

$$\begin{aligned} f_{R_1,R_2}(r_1, r_2) &= \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{r_1^2 - 2\rho r_1 r_2 + r_2^2}{2\cdot(1-\rho^2)}\right) \\ f_{R_1}(r_1) &= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{r_1^2}{2}\right) \\ f_{R_2}(r_2) &= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{r_2^2}{2}\right) \end{aligned}$$

Zeigen Sie, dass die beiden Komponenten des Zufallsvektors $R = (R_1, R_2)$ nicht unabhängig sind!

(i) Bestimmen Sie die bedingte Dichte $f_{R_1|R_2}(r_1|r_2)$ für den Fall, dass $\rho = 0$ ist!

(j) Wie groß ist der Erwartungswert und die Varianz von R_1 aus Aufgabenteil (h)? *vgl.* Abschn. 2.2.

Aufgabe 2.3.3 Dichte einer zweidimensionalen normalverteilten Zufallsgröße
⊛⊛⊛⊛⊛⊛

Sei der zweidimensionale Zufallsvektor (X_1, X_2) normalverteilt mit

$$f(x_1, x_2) = \frac{1}{2\,\pi\,\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)}\left\{\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 - 2\rho\frac{(x_1-\mu_1)}{\sigma_1}\frac{(x_2-\mu_2)}{\sigma_2} + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right\}\right]$$

Somit ist die marginalen Dichte (Randdichte) von X_1 und X_2 die Dichte der Normalverteilung mit den Parametern (μ_1, σ_1) für X_1 und (μ_2, σ_2) für X_2.

(a) Gehen Sie davon aus, dass $X_1, X_2 \sim N(0, 1)$ und $\rho = 0$. Zeichnen Sie die zweidimensionale Dichtefunktion mithilfe der Funktion (`persp`)! Stellen Sie die Dichte ebenfalls zweidimensional durch die Verwendung von Höhenlinien dar (`image` oder `contour`)!

(b) Ändern Sie die Parameter wie folgt: $X_1, X_2 \sim N(0, 1)$, $\rho = -0{,}8$ und $X_1 \sim N(0, 1)$, $X_2 \sim N(1, 3)$, $\rho = 0$. Analysieren Sie den Einfluss der Parameter auf die Lage und Streuung!

Hilfestellung: Definieren Sie sich zunächst eine Funktion, der Sie den Namen `normdens2d` zuweisen. Diese Funktion soll von den benötigten Parametern sowie dem Definitionsbereich von X_1 und X_2 abhängen (`normdens2d <- function( ... ) { ... }`). Die Parameter der Funktion sind in den runden Klammern zu definieren; der Code, den die Funktion ausführen soll, wird innerhalb der geschweiften Klammern geschrieben.

Die Funktion soll eine Matrix erzeugen, die für alle möglichen Kombinationen von X_1 und X_2 den entsprechenden Wert der zweidimensionalen Dichte der Normalverteilung enthält.

http://www.springer.com/978-3-662-49211-6

Arbeitsbuch der Angewandten Statistik
Mit Aufgaben zur Software R und detaillierten Lösungen
Otto, P.; Lange, A.-L.
2017, VIII, 158 S. 4 Abb., Softcover
ISBN: 978-3-662-49211-6