Aufgabe 1

Die Zwillinge Daida und Iballa M. kämpfen um den Weltmeistertitel im Windsurfen. Vor einem Wettkampf in Torbole macht der berühmte Moderator Roberto H. einen Vorbericht. In seiner Analyse beziffert er die Wahrscheinlichkeit, dass Daida den Wettkampf gewinnt, mit 0.4 und die Wahrscheinlichkeit, dass Iballa gewinnt, mit 0.3. Er schließt jedoch die Möglichkeit aus, dass beide gewinnen, d.h., beide Ereignisse "Iballa gewinnt den Wettkampf" und "Daida gewinnt den Wettkampf" sind disjunkt.

1. Wie groß ist die Wahrscheinlichkeit, dass beide nicht gewinnen?
   Hinweis: Die Vierfeldertafel könnte hilfreich sein.

2. Wie groß ist die Wahrscheinlichkeit, dass Iballa gewinnt, wenn Daida bereits verletzungsbedingt ausgeschieden ist und daher nicht mehr gewinnen kann?

3. Erfahrungsgemäß wird in 30% aller Fälle der Sieger aus Torbole später auch Weltmeister, während in 50% aller Fälle der spätere Weltmeister nicht in Torbole gewonnen hat. Wie groß ist die Wahrscheinlichkeit, dass Daida Weltmeister wird?

Es sei bekannt, dass die Slalomzeit einer Runde von Iballa bei gegebenen Wetterkonditionen mit einem Mittelwert von 10 Minuten und einer Standardabweichung von einer Minute normalverteilt ist.

4. Wie groß ist die Wahrscheinlichkeit, dass Iballa für eine beliebige Runde nicht genau 12 Minuten braucht?

5. Weiterhin sei bekannt, dass die Slalomzeit einer Runde von Daida bei gegebenen Wetterkonditionen 1.5-mal so groß ist wie die ihrer Schwester. Wie und mit welchen Parametern ist die Slalomzeit einer Runde von Daida verteilt?

6. Wie groß ist die Wahrscheinlichkeit, dass Iballa für eine beliebige Runde höchstens 14 Minuten und Daida für eine beliebige Runde mehr als 12 Minuten braucht? Es wird dabei angenommen, dass die Slalomzeiten beider Schwestern stochastisch unabhängig sind.
   Hinweis: Falls Sie Teilaufgabe 5. nicht lösen konnten, nehmen Sie an, dass die Slalomzeit einer Runde von Daida normalverteilt ist mit Mittelwert 15 und Varianz 2.25.

Bei dem Surffestival waren ca. 4000 Zuschauer. Es wurde eine einfache Stichprobe von 145 Personen gezogen. Darunter waren 58 Frauen.

7. Bestimmen Sie das approximative 96%-Konfidenzintervall für den Anteil der Frauen bei den Zuschauern.

Lösen Sie folgende Aufgaben in **R**.

8. Zeichnen Sie mittels **R** die Dichte einer Normalverteilung mit Mittelwert 10 und Varianz 1 für den Wertebereich von -4 bis 4.

9. Berechnen Sie das 98%-Quantil einer Standardnormalverteilung mittels **R**.

Aufgabe 2

Eine Zufallsvariable X ist exponentialverteilt, d.h. die Dichte lautet:

$f_X(x) = \lambda e^{-\lambda x}$, für $x \geq 0$.

1. Zeigen Sie, dass für die Verteilungsfunktion $F_X(x)$ gilt:

   $F_X(x) = 1 - e^{-\lambda x}$ für $\quad x \geq 0$.

2. Zeigen Sie, dass der Median $x_{(0.5)} = \frac{ln2}{\lambda}$ lautet.

3. Ein Hersteller von Energiesparlampen wirbt damit, dass seine Lampen eine mittlere Brenndauer von 10000 Stunden aufweisen. Gehen Sie davon aus, dass die Brenndauer einer Energiesparlampe exponentialverteilt ist. Mit welchem Parameter ist dann die Zufallsvariable X: "Brenndauer einer Energiesparlampe" verteilt? Welche Brenndauer wird von einer Energiesparlampe in diesem Fall mit einer Wahrscheinlichkeit von 50% nicht unterschritten?

4. Mit welcher Wahrscheinlichkeit brennt eine Energiesparlampe mehr als 10000 Stunden?

5. Zeigen Sie, dass die Exponentialverteilung "gedächtnislos" ist, d.h., dass gilt:

   $$P(X \geq x + a | X \geq a) = P(X \geq x).$$

6. Mit welcher Wahrscheinlichkeit brennt eine Energiesparlampe noch 2000 Stunden, wenn Sie wissen, dass Sie bereits 8500 Stunden gebrannt hat?

7. Sie wissen aufgrund einer Stichprobe von 100 Energiesparlampen, dass die mittlere Brenndauer 9500 Stunden beträgt. Schätzen Sie den Parameter λ mit Hilfe der Methode der Momente.

8. Was bewirkt folgender **R**-Befehl?

   ```
   > fix(expo)
   ```

9. Schreiben Sie in **R** eine Funktion `expo`, die den Wert der Verteilungsfunktion an der Stelle `x` einer exponentialverteilten Zufallsvariablen zurückgibt. Hierzu steht Ihnen folgender **R**-Output zur Verfügung:

   ```
   > function (x, lambda)
   {
   ...
   }
   ```

10. Geben Sie den Befehl zur Berechnung des Wertes der Verteilungsfunktion einer exponentialverteilten Zufallsvarianblen an der Stelle `x=1, lambda=1` mit Hilfe der Funktion `expo` an.

Aufgabe 3

Sie sind Qualitätsmanager beim Hauptplatinenhersteller Abus. Täglich nehmen Sie eine Lieferung von Kondensatoren für die Produktion ab, von denen laut Hersteller A maximal 10% defekt sind. Sie führen regelmäßig Stichproben durch, um festzustellen, ob der Anteil der defekten Kondensatoren höchstens 12% ist. Ist er größer, so schicken Sie die Ware zurück.

1. Sie ziehen aus einer Lieferung von 50 Kondensatoren 10 Stück ohne Zurücklegen. Wie und mit welchen Parametern ist die Zufallsvariable X: "Anzahl der defekten Kondensatoren in der Stichprobe" verteilt?
   Hinweis: Falls Sie diese Aufgabe nicht lösen konnten, gehen Sie im Folgenden davon aus, dass $X \sim Poi(0.5)$ ist.

2. Wie groß ist die Wahrscheinlichkeit, dass bei 10 gezogenen Kondensatoren,
   (a) genau einer defekt ist,
   (b) mindestens 2 defekt sind.

3. Da das Unternehmen expandiert, erreichen nun täglich 10000 Kondensatoren das Werk. Sie ziehen 25 Stück mit Zurücklegen und untersuchen, ob der Kondensator defekt ist oder nicht.
   (a) Wie und mit welchen Parametern ist nun die Zufallsvariable Y: "Anzahl der defekten Kondensatoren in der Stichprobe" verteilt.
   Hinweis: Falls Sie diese Aufgabe nicht lösen konnten, gehen Sie im Folgenden davon aus, dass $Y \sim N(2.5, 2.25)$ ist.
   (b) Wie groß ist die Wahrscheinlichkeit, genau 2 defekte Kondensatoren zu ziehen?
   (c) Wie groß ist die Wahrscheinlichkeit, dass Sie die Ladung zurückschicken?
   (d) Wie groß ist die Wahrscheinlichkeit, die Ladung zu akzeptieren, obwohl 15% der Kondensatoren defekt sind?
   (e) Was hat man mit der Wahrscheinlichkeit aus Teilaufgabe 3d) im Rahmen eines Hypothesentests bestimmt, wenn die Nullhypothese "Anteil der defekten Kondensatoren ist höchstens 10%" gegen die Alternative "Anteil der defekten Kondensatoren beträgt 15%" getestet wird?

4. Sie haben einen zweiten Lieferanten (Hersteller B) hinzubekommen. Dieser verspricht, dass nur 5% seiner Kondensatoren defekt sind. Allerdings können Sie nur 20% der Kondensatoren von diesem Hersteller beziehen, während der Rest weiterhin von Hersteller A kommt. Sie greifen sich aus den gesamten, an einem Tag angelieferten Kondensatoren einen Defekten heraus. Mit welcher Wahrscheinlichkeit stammt dieser aus Werk B?

5. Interpretieren Sie folgenden **R**-Code und geben Sie an, was berechnet wird. Verwenden Sie hierbei, dass `choose(a,b)=`$\binom{a}{b}$.

```
>x=c(0,1,2)
>zaehler=choose(5,x)*choose(45,10-x)
>nenner=choose(50,10)
>dhyp=zahler/nenner
>phyp=sum(dhyp)
>phyp
>0.9517397
```

Aufgabe 4

Die Assistenten F.T. und D.S. haben acht Gutscheine für eine Fast-Food-Filiale in Nürnberg über je einen Burger erhalten. Das Gewicht eines Burger (B) in Gramm sei normalverteilt mit $\mu_B = 211$ und unbekannter Varianz σ_B^2. Mit einer Waage ausgestattet gehen die beiden Assistenten zum Mittagessen in das Schnellrestaurant und wiegen die Burger vor dem Verzehr. Dabei haben sich folgende Gewichte ergeben:

Gutschein i	1	2	3	4	5	6	7	8
Gewicht in Gramm	209	205	206	211	212	209	210	204

1. Berechnen Sie die Stichprobenvarianz $S^2_{B;n}$.
2. Wie und mit welchen Parametern ist das Stichprobenmittel $\overline{B}_n$ verteilt?
3. Wie groß ist die Wahrscheinlichkeit, dass das Stichprobenmittel $\overline{B}_n$ kleiner als 211 Gramm ist? (Mit Begründung!)
4. Bestimmen Sie ein zweiseitiges 95%-Konfdenzintervall für die Varianz von B.
5. Nehmen Sie Stellung zu folgendenden Aussagen:
   (a) Eine Erhöhung der Irrtumswahrscheinlichkeit α führt bei einem Konfidenzintervall für den Mittelwert bei bekannter Varianz dazu, dass die Grenzen des Konfidenzintervalles näher beieinander liegen.
   (b) Ein Fehler 2. Art liegt vor, wenn die Nullhypothese abgelehnt wird, obwohl sie richtig ist.
6. Testen Sie bei einer Irrtumswahrscheinlichkeit von 5% die Nullhypothese $\mu_B = 211$.
7. Gegeben ist folgender Output in **R**:

```
t.test(X,Y,alternative='two.sided',mu=0,paired=FALSE,var.equal=TRUE,conf.level=0.975)

        Two Sample t-test

data:  X and Y
t = -4.0996, df = 18, p-value = 0.0006729
alternative hypothesis: true difference in means is not equal to 0
97.5 percent confidence interval:
 -30.464722  -7.701945
sample estimates:
mean of x mean of y
 208.2500  227.3333
```

   (a) Welcher Test wurde hier angewendet?
   (b) Was besagt der p-Wert?
   (c) Treffen Sie die Testentscheidung anhand des p-Wertes.

Lösung 1

1. D: "Daida gewinnt den Wettkampf", I:"Iballa gewinnt den Wettkampf",
   gegeben: $P(D) = 0.4$, $P(I) = 0.3$, $P(D \cap I) = 0$. Vierfeldertafel:

	I	$\bar{I}$	$\sum$
D	0	0.4	0.4
$\bar{D}$	0.3	0.3	0.6
$\sum$	0.3	0.7	1

   $P\left(\bar{D} \cap \bar{I}\right) = 0.3$

2. $P\left(I|\bar{D}\right) = \dfrac{P\left(I \cap \bar{D}\right)}{P\left(\bar{D}\right)} = \dfrac{0.3}{0.6} = 0.5$

3. W:"Daida wird Weltmeister",
   $P(W) = P(W|D)P(D) + P\left(W|\bar{D}\right) P\left(\bar{D}\right) = 0.3 \cdot 0.4 + 0.5 \cdot 0.6 = 0.42$

4. S_I:"Slalomzeit einer Runde von Iballa", gegeben: $S_I \sim N(10, 1)$

   $$P(S_I \neq 12) = 1 - P(S_I = 12) = 1 - 0 = 1$$

5. S_D:"Slalomzeit einer Runde von Daida", $S_D \sim N\left(1.5 \cdot 10 = 15, 1.5^2 \cdot 1 = 2.25\right)$
   Reproduktivität der Normalverteilung

6.

   $$\begin{aligned} P\left(S_I \leq 14, S_D \geq 12\right) &= P\left(S_I \leq 14\right) \cdot P\left(S_D \geq 12\right) = \Phi\left(\frac{14-10}{1}\right) \cdot \left(1 - \Phi\left(\frac{12-15}{\sqrt{2.25}}\right)\right) \\ &= \Phi(4)\left(1 - \Phi(-2)\right) = \Phi(4)\Phi(2) = 0.9772 \end{aligned}$$

7. Theoretisches approximatives 96%-KI für den Anteil der Frauen

   $$P\left(\bar{X}_n - \lambda_{1-\alpha/2} \cdot \sqrt{\frac{\bar{X}_n \cdot (1-\bar{X}_n)}{n}} \leq p \leq \bar{X}_n + \lambda_{1-\alpha/2} \cdot \sqrt{\frac{\bar{X}_n \cdot (1-\bar{X}_n)}{n}}\right) \approx 1 - \alpha$$

   Realisiertes approximatives 96%-Konfidenzintervall für den Anteil der Frauen

   $$\begin{aligned} KI &= \left[0.4 - 2.0537 \cdot \sqrt{\frac{0.4 \cdot 0.6}{145}}; 0.4 + 2.0537 \cdot \sqrt{\frac{0.4 \cdot 0.6}{145}}\right] \\ &= [0.3164; 0.4836], \end{aligned}$$

   $\bar{X}_n = 58/145 = 0.4$, $1 - \alpha = 0.96 \Rightarrow \alpha = 0.04$, $\lambda_{1-\alpha/2} = \lambda_{0.98} = 2.0537$

8.
   ```
   x=seq(-4,4,,100)
   y=dnorm(x,10,1)
   plot(x,y)
   ```

9. `qnorm(0.98)`

Lösung zu Aufgabe 2

1. Zu zeigen ist: $F_X(x) = \int_{-\infty}^{x} f(t)dt = 1 - e^{-\lambda x}$

$$\begin{aligned}\int_{-\infty}^{x} \lambda e^{-\lambda t} dt &= -e^{-\lambda t}\Big|_0^x \\ &= -e^{-\lambda x} + e^{-\lambda \cdot 0} = 1 - e^{-\lambda x}\end{aligned}$$

2. Der Median ist definiert als 50%-Quantil, d.h. $F(x_{(0.5)}) = 0.5$. Es gilt:

$$\begin{aligned}1 - e^{-\lambda x_{(0.5)}} &= 0.5 \text{Auflösen nach } x_{(0.5)} \\ e^{-\lambda x_{(0.5)}} &= 0.5 \\ -\lambda x_{(0.5)} &= \ln(0.5) \\ x_{(0.5)} &= \frac{-\ln(0.5)}{\lambda} = \frac{\ln(2)}{\lambda}\end{aligned}$$

3.
   - $X \sim Exp(\lambda = 1/10000)$.
   - $x_{(0.5)} = \ln(2)/1/10000 = 6931.47$.

4. $P(X > 10000) = 1 - F(10000) = 1 - (1 - e^{-1}) = 0.3678794$.

5.

$$\begin{aligned}P(X > x + a | X > a) &= \frac{1 - (1 - \exp(-\lambda(x + a)))}{1 - (1 - \exp(-\lambda \dot{a}))} \\ &= \frac{\exp(-\lambda(x + a))}{\exp(-\lambda a)} \\ &= \exp(-\lambda x - \lambda a + \lambda a) = \exp(-\lambda x)\end{aligned}$$

6. $P(X > 2000 + 8500 | X > 8500) = \exp(-2000/10000) = 0.8187308$.

7. Methode der Momente: Gleichsetzen des Stichprobenmomentes mit seinem Erwartungswert. Also hier: $\bar{x} = E[X] = \frac{1}{\lambda}$. Da $\bar{x} = 9500$ folgt daraus $\hat{\lambda} = 1/9500$.

8. Der Editor zum Schreiben der Funktion `expo` öffnet sich.

9.
```
> y=1-exp(-x*lambda)
> return(y)
```

10.
```
> expo(1,1)
```

Lösung zu Aufgabe 3

1. $X \sim Hyp(N = 50, n = 10, M = 5)$
2. (a)
$$P(X = 1) = \frac{\binom{5}{1}\binom{45}{9}}{\binom{50}{10}} = 0.4313372$$
   (b)
$$\begin{aligned} P(X \geq 2) &= 1 - P(X \leq 1) = 1 - (P(X = 0) + P(X = 1)) \\ &= 1 - (\frac{\binom{5}{0}\binom{45}{10}}{\binom{50}{10}} + 0.4313372) = 0.2581 \end{aligned}$$
3. $Y \sim Bin(n = 25, p = 0.1)$
4. $P(Y = 2) = f_{Bin(25,0.1)}(2) = 0.2658881$
5. Die Ladung wird zurückgeschickt, wenn mehr als 12% defekte Kondensatoren bei der Stichprobe dabei sind, d.h. hier: $0.12 \cdot 25 = 3$. Also insgesamt:
$$P(Y > 3) = 1 - P(Y \leq 3) = 1 - F_{Bin(25,0.1)}(3) = 1 - 0.7635914 = 0.2364086.$$
6. Man akzeptiert die Ladung, wenn man höchstens 3 defekte Kondensatoren zieht, d.h. dann:
$$P(Y \leq 3) = F_{Bin(25,0.15)} = 0.4711213.$$
7. Die Wahrscheinlichkeit aus Teilaufgabe 3d) ist der Fehler 2. Art oder β-Fehler im Hypothesentest.
8. Gegeben sind folgende Ereignisse:
   - D=Kondesnator defekt
   - B=Kondensator stammt von Hersteller B

   Die gesuchte Wahrscheinlichkeit ergibt sich mit der Formel von Bayes:
$$P(B|D) = \frac{P(D|B \cdot P(B)}{P(D)},$$
   wobei $P(D)$ mit Hilfe der Formel der totalen Wahrscheinlichkeit berechnet wird:
$$P(D) = P(B) \cdot P(D|B) + P(A) \cdot P(D|A) = 0.2 \cdot 0.05 + 0.8 \cdot 0.1 = 0.09$$
$$\begin{aligned} P(B|D) &= (P(B) \cdot P(D|B))/P(D) \\ &= 0.2 \cdot 0.05/0.09 = 0.01/0.09 = 0.111. \end{aligned}$$
   Antwort: Die Wahrscheinlichkeit, dass ein defekter Kondensator aus Werk B stammt beträgt 11,1%.
9. Erste Zeile: Ein Vektor mit den Werten 0,1,2 wird definiert.
   Zweite und dritte Zeile: Der Zähler und der Nenner einer hypergeometrisch verteilten Zufallsvariable mit $N = 50$,$M = 5$ und $n = 10$ wird berechnet.
   Vierte Zeile: Der Wert der Wahrscheinlichkeitsmassenfunktion an den Stellen 0,1 und 2 einer hypergeometrisch verteilten Zufallsvariable wird berechnet.
   Fünfte Zeile: Der Wert der Verteilungsfunktion an der Stelle 2 von dieser Zufallsvariable wird berechnet.

Lösung 4

1. $\overline{B}_n = \frac{1}{8}(209 + 205 + \cdots + 204) = 208.25$

   $S^2_{B;n} = \frac{1}{8-1}((209 - 208.25)^2 + \cdots + (204 - 208.25)^2) = 8.5$

2. $\overline{B}_n \sim N(\mu, \sigma^2/n)$.

3. $P(\overline{B}_n \leq 211) = 0.5$ wegen Symmetrie der NV.

4. Die Varianz bei bekanntem Mittelwert:
   $\sum(B_i - \mu_B)^2 = (209 - 211)^2 + \cdots + (204 - 211)^2 = 120$

   $\chi^2_{0.975;8} = 17.53$ und $\chi^2_{0.025;8} = 2.18$

   KI$=\left(\frac{120}{17.53}; \frac{120}{2.18}\right) = [6.8454; 55.0459]$

5. (a) Ja, richtig, weil sich die jeweiligen Quantile wertmäßig gegeneinander annähern.
   (b) Nein, falsch, Nullhypothese wird nicht abgelehnt obwohl sie falsch ist.

6. Mittelwerttest bei unbekannter Varianz:

   $H_0 : \mu_B = 211$, $H_A : \mu_B \neq 211$

   Kritischer Bereich: $\mid \sqrt{n}\frac{\overline{B}_n - \mu_0}{S_{B;n}} \mid > t_{1-\alpha/2;n-1}$;

   Kritische Schranke: $t_{1-\alpha/2;n-1} = t_{0.975,7} = 2.365$

   $T = \mid \sqrt{8}\frac{208.25-211}{\sqrt{8.5}} \mid = \mid -2.66789 \mid > 2.36$

   Da Teststatistik im kritischen Bereich liegt, muss H_0 mit einer IW $\alpha = 0.05$ abgelehnt werden.
   Bemerkung: Eigentlich war der wahre Mittelwert laut Angabe bekannt - ein Test daher unnötig. Diese Alternative wäre auch Lösung der Aufgabe.

7. Mittelwertdifferenzentest bei unverbundenen Stichproben. Der p-Wert gibt die Wahrscheinlichkeit an, mit der ein so extremes Stichprobenergebnis oder ein noch extremeres auftreten kann. Der p-Wert ist kleiner als die Irrtumswahrscheinlichkeit, daher wird die Nullhypothese abgelehnt.