Aufgabe 1

Folgende Daten geben die Anzahl der Buchpublikationen (X) und die Anzahl der Departments (Y) von 8 Universitäten im Jahr 2011 an.

Universität i	1	2	3	4	5	6	7	8
Anzahl der Buchpublikationen X	8	10	8	7	4	3	12	4
Anzahl der Departments Y	26	21	22	16	21	14	24	16

1. Berechnen Sie den Modus, den Median und das untere Quartil von X.

2. Berechnen Sie die Spannweite von X.

3. Berechnen und interpretieren Sie die Verteilungsfunktion von X an der Stelle 13.

4. Es wird vermutet, dass die Anzahl der Buchpublikationen (X) von der Anzahl der Departments (Y) abhängt. Berechnen Sie den Korrelationskoeffizienten, wenn bekannt ist, dass $\sum_{i=1}^{8} x_i = 56$, $\sum_{i=1}^{8} y_i = 160$, $s_Y^2 = \frac{1}{8}\sum_{i=1}^{8}(y_i - \overline{y})^2 = 15.75$ und $\frac{1}{8}\sum_{i=1}^{8} x_i y_i = 148$.

Eine Publikation soll auf Plagiate überprüft werden. Dazu sollen zufällig n Seiten ausgewählt und auf Plagiate durchsucht werden. Die Anzahl der Plagiate pro Seite (Z) sei normalverteilt mit unbekanntem Mittelwert μ und Varianz $\sigma^2 = 0.04$.

5. Bestimmen Sie den Stichprobenumfang n so, dass die mittlere Anzahl von Plagiaten in der Stichprobe mit einer Wahrscheinlichkeit von mindestens 95% um weniger als 0.1 von μ abweicht.

6. Es werden nun 20 zufällig ausgewählte Seiten geprüft und das Stichprobenmittel $\overline{Z_{20}}$ beträgt 2.2. Schätzen Sie die tatsächliche mittlere Anzahl der Plagiate pro Seite nach einer geeigneten Methode.

7. Geben Sie an, ob folgende Aussagen „wahr“ oder „falsch“ sind (ohne Begründung). Jede **richtige** Antwort wird mit **0.5**, jede **falsche** mit **-0.5** und **keine** Antwort mit **0** Punkten bewertet. **Insgesamt** können in dieser Teilaufgabe **mindestens 0** und **höchstens 2** Punkte erreicht werden.

(a) ________ Sei σ^2 die Varianz der Zufallsvariable X, dann hat die Zufallsvariable $Y = 2X$ die Varianz $4\sigma^2$.

(b) ________ X und Y sind zwei stochastisch unabhängige Zufallsvariablen. Sei $X \sim N(\mu, \sigma^2)$ und $Y \sim N(\mu, \sigma^2)$, dann hat die Zufallsvariable $Z = X + Y$ die Varianz $2\sigma^2$.

(c) ________ Die Kovarianz σ_{XY} von zwei Zufallsvariablen X und Y kann auch negative Werte annehmen, wenn ein negativer linearer Zusammenhang zwischen X und Y vorliegt.

(d) ________ Die Varianz des Stichprobenmittels steigt mit zunehmendem Stichprobenumfang.

Schmierpapier

Aufgabe 2

Peter möchte vor seiner Berufswahl wissen, ob unterschiedliche Wirtschaftszweige stärker als andere von Insolvenzen betroffen sind. Dazu hat er $n = 800$ zufällig ausgesuchte Unternehmen aus drei Wirtschaftszweigen über ein Jahr beobachtet. Betrachten Sie die folgende Kontingenztabelle der absoluten Häufigkeiten mit den beiden Merkmalen
A: „Wirtschaftszweig des Unternehmens“ und B: „Zahlungsfähigkeit des Unternehmens“:

$B\backslash A$	verarbeitendes Gewerbe	Baugewerbe	Handel	
solvent	$n_{11} = 278$	$n_{12} = 140$	$n_{13} = 297$	$n_{1\cdot} = 715$
insolvent	$n_{21} = 32$	$n_{22} = 24$	$n_{23} = 29$	$n_{2\cdot} = 85$
	$n_{\cdot 1} = 310$	$n_{\cdot 2} = 164$	$n_{\cdot 3} = 326$	$n = 800$

1. Welches Lagemaß ist für das Merkmal A sinnvoll?

2. Berechnen und interpretieren Sie $\frac{n_{11}}{n_{\cdot 1}}$ und $\frac{n_{1\cdot}}{n}$. Sind A und B unabhängig?

3. Berechnen Sie die relativen Randhäufigkeiten für das Merkmal A.

4. Berechnen und interpretieren Sie die normierte Entropie H_A^* für das Merkmal A.

5. Vervollständigen Sie folgende Indifferenztabelle für die Verteilung der Merkmale A und B:

$B \backslash A$	a_1	a_2	a_3	Σ
b_1	$n_{11}^* = 277.0625$	$n_{12}^* =$	$n_{13}^* =$	$n_{1\cdot} = 715$
b_2	$n_{21}^* = 32.9375$	$n_{22}^* =$	$n_{23}^* =$	$n_{2\cdot} = 85$
Σ	$n_{\cdot 1} = 310$	$n_{\cdot 2} = 164$	$n_{\cdot 3} = 326$	$n = 800$

6. Berechnen Sie die mittlere quadratische Kontingenz ϕ^2_{AB}. Folgende Zwischenergebnisse können verwendet werden:

$$\frac{\left(n_{11}-n^*_{11}\right)^2}{n^*_{11}} = 0.0032 \quad \frac{\left(n_{12}-n^*_{12}\right)^2}{n^*_{12}} = 0.2949 \quad \frac{\left(n_{13}-n^*_{13}\right)^2}{n^*_{13}} = 0.1091$$
$$\frac{\left(n_{22}-n^*_{22}\right)^2}{n^*_{22}} = 2.4810 \quad \frac{\left(n_{23}-n^*_{23}\right)^2}{n^*_{23}} = 0.9175$$

(**Hinweis:** Wenn Sie diese Aufgabe nicht lösen können, verwenden Sie im Folgenden $\phi^2_{AB} = 0.005$.)

7. Berechnen und interpretieren Sie das Cramérsche Kontingenzmaß V für die Merkmale A und B.

8. Führen Sie einen χ^2-Unabhängigkeitstest für die Daten aus Peters Erhebung durch. Testen Sie die Nullhypothese H_0 : „Zahlungsfähigkeit und Wirtschaftszweig sind unabhängig“ auf dem Signifikanzniveau $\alpha = 10\%$.
(**Hinweis:** Das Prüfmaß entspricht $n \cdot \phi^2_{AB}$ und ist unter Gültigkeit der Nullhypothese χ^2-verteilt mit 2 Freiheitsgraden.)

9. Geben Sie an, ob die folgenden Aussagen ”wahr” oder ”falsch” sind (ohne Begründung). Jede **richtige** Antwort wird mit **0.5**, jede **falsche** mit **-0.5** und keine Antwort wird mit **0** Punkten bewertet. **Insgesamt** können in dieser Teilaufgabe **mindestens 0** und **höchstens 2** Punkte erreicht werden.

(a) ________ Ist die Nullhypothese eines Tests wahr, dann stimmt die Wahrscheinlichkeit dafür, dass die Teststatistik in den Ablehnungsbereich fällt, genau mit dem Signifikanzniveau α überein.

(b) ________ Bei genau spezifizierter Alternativhypothese lässt sich die Wahscheinlichkeit für den Fehler zweiter Art nicht berechnen.

(c) ________ Hält man das Signifikanzniveau konstant, steigt die Güte eines Tests mit dem Stichprobenumfang.

(d) ________ Der p-Wert gibt die Güte des Tests an.

Schmierpapier

Schmierpapier

Aufgabe 3

Sie untersuchen im Auftrag eines Tierschutzverbandes, wieviele Hühner aus Massentierhaltung mit Antibiotika behandelt und anschließend in Supermärkten verkauft werden. Dazu führen Sie eine Zufallsstichprobe in zufällig ausgewählten Supermärkten durch. Die Zufallsvariable X: "Ein zufällig ausgewähltes Huhn wurde mit Antibiotika behandelt." sei bernoulliverteilt, d.h. X besitzt die Wahrscheinlichkeitsmassenfunktion

$$f_X(x) = p^x(1-p)^{1-x}$$

mit unbekanntem Parameter $p \in [0, 1]$. Dabei sei $X = 1$, falls das Huhn mit Antibiotika behandelt wurde und $X = 0$ falls nicht.

1. Geben Sie an, ob die folgenden Aussagen "wahr" oder "falsch" sind (ohne Begründung). Jede **richtige** Antwort wird mit **0.5**, jede **falsche** mit **-0.5** und keine Antwort wird mit **0** Punkten bewertet. **Insgesamt** können in dieser Teilaufgabe **mindestens 0** und **höchstens 2** Punkte erreicht werden.

(a) ________ Die Punktwahrscheinlichkeit bei einer Bernoulliverteilung ist immer gleich 0.

(b) ________ Die Dichtefunktion bei stetigen Verteilungsmodellen kann Werte größer 1 annehmen.

(c) ________ Sind die Ereignisse A und B unabhängig, so gilt stets: $P(A \cup B) = P(A) + P(B)$.

(d) ________ Sind die Ereignisse A und B unabhängig, so gilt stets: $P(A \cap B) = P(A) \cdot P(B)$.

2. Zeigen Sie, dass die Loglikelihoodfunktion $\ln L(p)$ einer Bernoulliverteilung folgende Gestalt hat:

$$\ln L(p) = \sum_{i=1}^{n} (x_i \ln(p) + (1 - x_i) \ln(1 - p)).$$

3. Zeigen Sie, dass der ML-Schätzer $\hat{p}_{ML}$ für p durch $\hat{p}_{ML} = \frac{1}{n} \sum_{i=1}^{n} x_i$ gegeben ist.

Die Lebensmittelindustrie wirbt damit, dass in deutschen Supermärkten der Anteil der Hühner, die vorher mit Antibiotika behandelt wurden, maximal 20% betrage. Aufgrund unabhängiger Recherche vermuten Tierschutzverbände einen Anteil von mehr als 60%. Um diese Aussagen zu überprüfen, formulieren Sie einen Hypothesentest der Form:

$$H_0 : \ p = 0.2, \quad H_1 : \ p = 0.6$$

Anschließend ziehen Sie eine Stichprobe der Länge $n = 20$ und untersuchen, wie groß der Anteil der mit Antibiotika behandelten Hühner in der Stichprobe ist. Hierfür ziehen Sie die Zufallsvariable T: "Anzahl der Hühner, die mit Antibiotika behandelt wurden" als Prüfgröße heran. Sie wissen, dass T binomialverteilt mit Parameter p ist. Sie geben sich eine Irrtumswahrscheinlichkeit von 3.21% vor.

4. Berechnen Sie die Wahrscheinlichkeit, unter Gültigkeit von H_0 höchstens 5 mit Antibiotika behandelte Hühner zu beobachten.

5. Legen Sie die kritische Schranke k^* fest, unter der H_0 gerade noch beibehalten wird. (**Hinweis**: Wenn Sie k^* nicht bestimmen konnten, dann verwenden Sie im Folgenden $k^* = 7$.)

6. Aufgrund der Stichprobe wissen Sie, dass $T = 10$. Treffen Sie eine Testentscheidung.

7. Berechnen Sie den p-Wert für $T = 10$.

8. Berechnen Sie die Wahrscheinlichkeit des Fehlers 2. Art des obigen Tests, d.h. $\beta = P(T \leq k^*|p = 0.6)$. Verwenden Sie hierfür, falls notwendig, dass $F_{Binom}(x; n, p) = 1 - F_{Binom}(n - x - 1; n, 1 - p)$ ist.

Schmierpapier

Schmierpapier

Aufgabe 4

Der R-Data Frame `NW` enthält die folgenden Merkmale von 50 zufällig ausgewählten sozialen Netzwerken, die derzeit im Internet vertreten sind.

Merkmal A:	Anzahl registrierter Nutzer in Mio.	Spalte `Nutzer`
Merkmal B:	durchschnittl. tägl. Nutzungsdauer in Min.	Spalte `Dauer`
Merkmal C:	Thema	Spalte `Thema`
Merkmal D:	Verbreitung	Spalte `Verbreitung`

Beim Merkmal D handelt es sich um ein qualitatives Merkmal mit den Ausprägungen d_1: national und d_2: international.

Betrachten Sie nun den zum Data Frame `NW` gehörenden R-Output.

```
> table(NW$Verbreitung,NW$Thema)

                Business Freizeit Spezial
  international       13       12       7
  national             8       10       0
> cor(NW$Nutzer,NW$Dauer)
[1] 0.8411377
> max(NW$Nutzer)-min(NW$Nutzer)
[1] 132.96
> D_i=NW[NW$Verbreitung=="international","Dauer"]
> D_n=NW[NW$Verbreitung=="national","Dauer"]
> quantile(D_n,1,type=1)
100%
19.5
> max(D_i)
[1] 48
> N_i=NW[NW$Verbreitung=="international","Nutzer"]
> N_n=NW[NW$Verbreitung=="national","Nutzer"]
> mean(N_i)
[1] 17.44369
> mean(N_n)
[1] 6.078453
> quantile(N_i,0.75,type=1)
75%
 20
```

1. Welche Ausprägungen des Merkmals Thema kommen bei den 50 Netzwerken vor?

2. Geben Sie den (die) R-Befehl(e) an, um die relativen Häufigkeiten des Merkmals Thema geeignet graphisch darzustellen.

3. Geben Sie den R-Befehl an, um zu überprüfen, ob die Spalte `Dauer` den passenden Datentyp besitzt.

4. Geben Sie an, ob folgende Aussagen „wahr“ oder „falsch“ sind (ohne Begründung). Jede **richtige** Antwort wird mit **0.5**, jede **falsche** mit **-0.5** und **keine** Antwort mit **0** Punkten bewertet. **Insgesamt** können in dieser Teilaufgabe **mindestens 0** und **höchstens 2** Punkte erreicht werden.

(a) ________ Insgesamt sind 18 nationale Netzwerke im Datensatz vertreten.

(b) ________ Der Quartilsabstand von `NW$Nutzer` beträgt maximal 132.96.

(c) ________ In allen Netzwerken wurde eine durchschnittl. tägl. Nutzungsdauer von 19.5 Min. nicht überschritten.

(d) ________ Es besteht ein positiver linearer Zusammenhang zwischen den Merkmalen A und B.

5. Vervollständigen Sie die folgenden Aussagen:

   (a) Ein nationales Netzwerk hat durchschnittlich __________ Mio. Nutzer.

   (b) _____% der internationalen Netzwerke verzeichneten maximal 20 Mio. Nutzer.

6. Vervollständigen Sie die folgende R-Funktion so, dass damit die Wahrscheinlichkeitsfunktion einer Binomialverteilung mit Stichprobenumfang `n` und Erfolgswahrscheinlichkeit `pr` gezeichnet wird.

```
Bin_plot=function(n,pr){
x=seq(0,n,1)
fx=
plot(x,fx,type="p") }
```

Es wird angenommen, dass es sich bei den internationalen und den nationalen Netzwerken im Datensatz jeweils um einfache Zufallsstichproben handelt und dass für die Zufallsvariablen X: "durchschnittl. tägl. Nutzungsdauer in einem internationalen Netzwerk" und Y: "durchschnittl. tägl. Nutzungsdauer in einem nationalen Netzwerk" gelte:

$X \sim N(\mu_X, \sigma_X^2),\ Y \sim N(\mu_Y, \sigma_Y^2)$ und $\sigma_X^2 = \sigma_Y^2$

7. Bei einer Irrtumswahrscheinlichkeit von 1% soll die Vermutung getestet werden, dass $\mu_X - \mu_Y > 2$.
   Vervollständigen Sie entsprechend den folgenden Befehl (sodass im Output auch das korrekte Konfidenzintervall ausgegeben wird).

```
> t.test(D_i,D_n,                                       ,
        paired=FALSE,var.equal=TRUE,                   )
```

8. Zusätzlich sei bekannt, dass $\sigma_X^2 = 73$. Es soll $H_0: \ \mu_X \leq 14$ gegen $H_1: \ \mu_X > 14$ bei $\alpha = 0.1$ getestet werden.
   Treffen Sie eine Testentscheidung mit Hilfe des folgenden Outputs (mit <u>kurzer</u> Begründung).

```
> n=length(D_i)
> t=(mean(D_i)-14)/(sqrt(73/n))
> t
[1] 1.515553
> pnorm(t,mean=0,sd=1)
[1] 0.9351838
> 1-pnorm(mean(D_i),mean=14,sd=sqrt(73/n))
[1] 0.06481617
```

9. Warum sind die oben getroffenen Verteilungsannahmen unter Berücksichtigung des Trägers von X bzw. Y zu kritisieren?

Aufgabe 1

1.

i	x_i	n_i	f_i	F_i
1	3	1	0.125	0.125
2	4	2	0.25	0.375
3	7	1	0.125	0.5
4	8	2	0.25	0.75
5	10	1	0.125	0.875
6	12	1	0.125	1

Modi: 4 und 8, Median: $x_{(0.5)} = 7$, Unteres Quartil: $x_{(0.25)} = 4$.

2. Spannweite: $12 - 3 = 9$

3. $F(13) = 1$
   Alle betrachtete Universitäten haben im Jahr 2011 maximal 13 Buchpublikationen.

4. $s_{XY} = \overline{xy} - \overline{x} \cdot \overline{y} = 148 - \frac{56}{8} \cdot \frac{160}{8} = 8$
   $s_X^2 = \frac{1}{8} \sum_{i=1}^{8} (x_i - \overline{x})^2 = \frac{1}{8}(1 + 9 + 1 + 9 + 16 + 25 + 9) = 8.75$
   $r_{XY} = \frac{s_{XY}}{s_X s_Y} = \frac{8}{\sqrt{8.75}\sqrt{15.75}} = 0.6815$
   oder
   $s_X^2 = \frac{1}{8-1} \sum_{i=1}^{8} (x_i - \overline{x_8})^2 = \frac{1}{7}(1 + 9 + 1 + 9 + 16 + 25 + 9) = 10$
   $s_Y^2 = 15.75 \cdot \frac{8}{7} = 18$
   $r_{XY} = \frac{S_{XY}}{S_X S_Y} = \frac{8}{\sqrt{10}\sqrt{18}} = 0.5963$

5. $P(|\overline{Z_n} - \mu| < 0.1) = 0.95$
   $P(-0.1 < \overline{Z_n} - \mu < 0.1) = 0.95$
   $= P(\frac{-0.1}{\sigma/\sqrt{n}} < \frac{\overline{Z_n} - \mu}{\sigma/\sqrt{n}} < \frac{0.1}{\sigma/\sqrt{n}}) = \Phi(\frac{0.1}{\sigma/\sqrt{n}}) - \Phi(\frac{-0.1}{\sigma/\sqrt{n}}) = 2\Phi(\frac{0.1}{\sigma/\sqrt{n}}) - 1 = 0.95$
   $\frac{0.1}{0.2/\sqrt{n}} = 1.96 \quad \Rightarrow \sqrt{n} = 3.92 \quad \Rightarrow n = 15.3664 \quad \Rightarrow n \geq 16$

6. Methode der Momente. Das erste Potenzmoment der Stichprobe (Stichprobenmittel) wird dem Erwartungswert der Grundgesamtheit gleichgesetzt.
   $\widehat{\mu} = \overline{Z_{20}} = 2.2$

7. W, W, W, F

Aufgabe 2

1. Modus.

2. Anteil der solventen Unternehmen innerhalb des verarbeitenden Gewerbes: 89.6774%

   Anteil der solventen Unternehmen innerhalb der beobachteten Unternehmen: 89.375%

   Nein.

3. $f_{\cdot 1} = 0.3875$, $f_{\cdot 2} = 0.205$ und $f_{\cdot 3} = 0.4075$

4. $H_A^* = \frac{-\sum_{i=1}^{3} f_{\cdot i} \log_2 f_{\cdot i}}{\log_2 3} = 0.9631$

   Die Verteilung von Merkmal A hat eine hohe Streuung.

5. $n_{22}^* = \frac{85 \cdot 164}{800} = 17.425$

   $n_{23}^* = 85 - 32.9375 - 17.425 = 34.6375$

   $n_{12}^* = 164 - 17.425 = 146.575$

   $n_{13}^* = 326 - 34.6375 = 291.3625$

6. $\frac{\left(n_{21} - n_{21}^*\right)^2}{n_{21}^*} = 0.0267$

   $\phi_{AB}^2 = \frac{1}{n} \sum_{i=1}^{2} \sum_{j=1}^{3} \frac{\left(n_{ij} - n_{ij}^*\right)^2}{n_{ij}^*} = 0.0048$

7. $V = \sqrt{\frac{\phi_{AB}^2}{\min\{3-1, 2-1\}}} = 0.0693$. Der Grad der Abhängigkeit der Merkmale A und B ist sehr gering.

8. $800 \cdot 0.0048 = 3.84 < 4.61 = q_{\chi^2(2)}^{0.9}$

   Nullhypothese kann auf 10% Signifikanziveau nicht abgelehnt werden.

9. W, F, W, F

Aufgabe 3

1. F, W, F, W

2. Bei dieser Aufgabe handelt es sich um Stichproben aus bernoulliverteilten Grundgesamtheiten, d.h. $X \sim Ber(p)$ wobei $f_X(x) = p^x(1-p)^{1-x}$ und $x \in \{0,1\}$. Likelihoodfunktion aufstellen:

$$\begin{aligned} L(p) &= \prod_{i=1}^{n} p_i^x (1-p)^{1-x_i} \\ &= p^{\sum_{i=1}^n x_i}(1-p)^{\sum_{i=1}^n 1-x_i}. \end{aligned}$$

   Übergang zur Loglikelihoodfunktion:

$$\begin{aligned} LL(p) &= \ln L(p) = \ln\left(p^{\sum_{i=1}^n x_i}(1-p)^{\sum_{i=1}^n 1-x_i}\right) \\ &= \sum_{i=1}^{n} x_i \ln(p) + (\sum_{i=1}^{n} 1 - x_i)\ln(1-p). \end{aligned}$$

3. Zur Bestimmung des ML-Schätzer sind 3 Schritte nötig:

   (a) Ableiten der Loglikelihood nach p.

   (b) Ableitung gleich Null setzen.

   (c) Nach p auflösen.

   Also:

$$\frac{\partial LL(p)}{\partial p} = \sum_{i=1}^{n} x_i \frac{1}{p} + (\sum_{i=1}^{n} 1 - x_i)\frac{1}{1-p}(-1).$$

   Ableitung gleich 0 setzen und nach p auflösen:

$$\begin{aligned} \frac{\partial LL(p)}{\partial p} &\overset{!}{=} 0 \\ \sum_{i=1}^{n} x_i \frac{1}{p} - (\sum_{i=1}^{n} 1 - x_i)\frac{1}{1-p} &\overset{!}{=} 0 \\ \sum_{i=1}^{n} x_i \frac{1}{p} &= (n - \sum_{i=1}^{n} x_i)\frac{1}{1-p} \\ (1-p)\sum_{i=1}^{n} x_i &= (n - \sum_{i=1}^{n} x_i)p \\ \Longrightarrow \hat{p}_{ML} &= 1/n \sum_{i=1}^{n} x_i. \end{aligned}$$

4. $P(T \leq 5) = F_{Binom(n=20,p=0.2)}(5) = 0.8042.$

5. Der Ansatz für den kritischen Bereich und gegebener Prüfgröße T lautet:

$$P(H_0 \text{ ablehnen}|H_0 \text{ gilt }) = P(T > k^*|p = 0.2) = \alpha = 0.0321.$$

   Also insgesamt:

$$\begin{aligned} P(T > k^*|p = 0.2) &= 1 - P(T \leq k^*|p = 0.2) = 0.0321 \\ &\iff 1 - F_{Binom(n=20,p=0.2)}(k^*) = 0.0321 \\ &\iff F_{Binom(n=20,p=0.2)}(k^*) = 0.9679 \end{aligned}$$

   Tabelle Binomialverteilung für $p = 0.2, n = 20$ nachschauen: $k^* = 7$.

6. Da $T > k^*$, die Prüfgröße somit im kritischen Bereich liegt, lehne H_0 auf dem 3.21%-Niveau ab.

7. $P(T \geq 10|p = 0.2) = 1 - P(T \leq 9|p = 0.2) = 1 - F_{Binom(n=20,p=0.2)}(9) = 1 - 0.9974 = 0.0026.$

8.

$$\begin{aligned} \beta &= P(H_0 \text{ nicht ablehnen}|H_1 \text{ ist richtig}) = P(T \leq 7|p = 0.6) \\ &= 1 - F_{Binom(n=20,p=0.4)}(20 - 7 - 1) \text{ siehe Hinweis!} \\ &= 1 - F_{Binom(n=20,p=0.4)}(12) = 1 - 0.9790 = 0.0210. \end{aligned}$$

Aufgabe 4

1. Business, Freizeit, Spezial
2. (a) `barplot(table(NW$Thema)/length(NW$Thema))`

   (b) `is.nunmeric(NW$Dauer)` oder `str(NW)`, `class(NW$Dauer)`
3. W, W, F, W
4. (a) 6.078453

   (b) 75
5. ```
   Bin_plot=function(n,pr){
   x=seq(0,n,1)
   fx= dbinom(x,size=n,prob=pr)
   plot(x,fx,type="p") }
   ```
6. ```
   t.test(D_i,D_n,mu=2,paired=FALSE,var.equal=TRUE,alternative="greater",
   conf.level=0.99)
   ```
7. H_0 ist abzulehnen, da p-Wert des Tests mit 0.06481617 kleiner ist als vorgegebene Irrtumswahrscheinlichkeit $von 0.1$.
8. Träger der Normalverteilung ist $\mathbb{R}$, d.h. damit wären auch Realisationen < 0 möglich, negative Werte für das Merkmal Nutzungsdauer jedoch nicht möglich.