

# Klausuren BW24.1

Oliver Kirchkamp

28. Januar 2020

## Inhaltsverzeichnis

	12	Klausur BW24.1, Februar 2014	27
1	3	Klausur BW24.1, Mai 2019	29
2	6	Klausur BW24.1, Februar 2019	30
3	9	Klausur BW24.1, Mai 2018	32
4	12	Klausur BW24.1, Februar 2018	34
5	14	Klausur BW24.1, Mai 2017	36
6	16	Klausur BW24.1, Februar 2017	37
7	18	Klausur BW24.1, Mai 2016	39
8	20	Klausur BW24.1, Februar 2016	41
9	22	Klausur BW24.1, Mai 2015	43
10	24	Klausur BW24.1, Februar 2015	46
11	26	Klausur BW24.1, Mai 2014	
		12 Klausur BW24.1, Februar 2014	27
		13 Klausur BW24.1, Mai 2013	29
		14 Klausur BW24.1, Februar 2013	30
		15 Klausur BW24.1, Mai 2012	32
		16 Klausur BW24.1, Februar 2012	34
		17 Klausur BW24.1, Mai 2011	36
		18 Klausur BW24.1, Februar 2011	37
		19 Klausur BW24.1, Mai 2010	39
		20 Klausur BW24.1, Februar 2010	41
		21 Klausur Statistik II, Mai 2008	43
		22 Klausur Statistik II, Februar 2008	46

## Informationen zur Klausur BW 24.1

- In der Klausur erhalten Sie folgendes Material:
  - Lösungsbogen (Überprüfen Sie bitte, ob Ihr Name auf dem Bogen angegeben ist).
  - Aufgabenblatt (Wird erst ausgeteilt, wenn alle sitzen.)
  - Informationsblatt
  - Schmierpapier
- Geben Sie am Ende der Klausur *nur den Lösungsbogen* ab. Ihr Aufgabenblatt, Ihr Schmierpapier, und Ihr Informationsblatt nehmen Sie nach der Klausur mit nach Hause. Wenn Sie auf dem Aufgabenblatt Ihre Lösungen markiert haben, können Sie anhand der Musterlösung leicht Ihren Erfolg kontrollieren.

**Der Lösungsbogen wird in der gleichen Reihenfolge eingesammelt, in der auch das Auf-**

**gabenblatt ausgeteilt wurde. Damit die Bearbeitungszeit für alle Kandidaten etwa gleich ist, werden verspätet abgegebene Bögen nicht gewertet.**

Bleiben Sie *ruhig auf Ihrem Platz sitzen, bis auch die letzte Person ihre Klausur beendet hat*. Bitte nehmen Sie Rücksicht auf die anderen Kandidaten! Jeder sollte die Klausur in einer ruhigen Umgebung zuende schreiben.

- Es sind grundsätzlich *keine Hilfsmittel* zugelassen. Im Einzelfall notwendige Ausnahmen (Lese- oder Schreibhilfe, Wörterbücher) klären Sie bitte vorher ab. Taschen, Taschenrechner, Mobiltelefone, Fachliteratur, Formelsammlungen, Tafelwerk, Aufzeichnungen, etc., deponieren Sie bitte *vorne* im Hörsaal. Schalten Sie Ihre Mobiltelefone *vollständig aus*, bevor Sie sie vorne im Hörsaal deponieren. Es wird Ihnen und den anderen Kandidaten leichter fallen, sich auf die Klausur zu konzentrieren, wenn nicht immer irgendwo ein Telefon klingelt oder vibriert.

- Es gibt unterschiedliche Klausurversionen. Ihre Version ist auf Ihrem Aufgabenblatt oben links angegeben:

Version:

Übertragen Sie bitte Ihre Version, die Sie auf Ihrem Aufgabenblatt finden, in Ihren Lösungsbogen, indem Sie dort die entsprechenden Felder ausfüllen.

Bitte markieren Sie alle Antworten mit einem dunklen Stift.

- Bei einigen Aufgaben mit vorgegebenen Antworten ist nur eine Antwort richtig (diese Aufgaben sind nicht besonders gekennzeichnet).

Antwort	a	b	c	d	e
x.x	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

Wie beim Wettbewerb erhalten Sie bei diesen Aufgaben nur dann die volle Punktzahl, wenn Sie nur die richtige Lösung markieren. In allen anderen Fällen bekommen Sie keinen Punkt.

Bei anderen Aufgaben mit vorgegebenen Antworten sagen wir Ihnen nicht, wie viele Antworten richtig sind. Diese Aufgaben sind mit dem Vermerk: »mehrere Antworten möglich« gekennzeichnet. Wie im Wettbewerb bekommen Sie in diesem Fall für korrekt »ja« und »nein« markierte Antworten jeweils die Teilpunktzahl. Für falsch »ja« und »nein« markierte bekommen Sie jeweils keinen Punkt.

Beispiel: Es gibt fünf mögliche Lösungen: a, b, c, d, und e. Richtig sind c, d und e. Sie bekommen also für die richtige Antwort die volle Punktzahl:

Antwort	a	b	c	d	e
x.x: ja	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
x.x: nein	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Für die folgende Antwort erhalten Sie nur 4/5 der Punkte:

Antwort	a	b	c	d	e
x.x: ja	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
x.x: nein	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

(a,b,d,e ist richtig markiert, c ist falsch markiert). Für die folgende Antwort erhalten Sie 2/5 der Punkte:

Antwort	a	b	c	d	e
x.x: ja	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
x.x: nein	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>

(nur a und d sind richtig markiert). Für die folgende Antwort erhalten Sie keinen Punkt:

Antwort	a	b	c	d	e
x.x: ja	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
x.x: nein	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

(nichts ist richtig markiert). **Markieren Sie auf jeden Fall entweder »ja« oder »nein«.** Ansonsten verzögert sich die Bearbeitung für Sie und für alle anderen. Durch zufälliges Ankreuzen bei diesen Aufgaben erhalten Sie bereits 50% der Punktzahl. Zum Bestehen benötigen Sie mehr Punkte, als durch zufälliges Ankreuzen erreicht werden kann.

### Die letzten fünf Minuten:

- Übertragen Sie Ihre Antworten in den Lösungsbogen bitte **erst, wenn Sie sich Ihrer Antwort sicher sind.**

Wenn Sie, trotz aller Vorsicht, ein markiertes Feld wieder »löschen« wollen, kreuzen Sie das Feld bitte wie in diesem Beispiel durch. Im folgenden Beispiel ist nur Antwort (b) als richtig markiert:

Antwort	a	b	c	d	e
x.x:	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

- Vergessen Sie nicht, Ihren Lösungsbogen mit einem dokumentenechten Stift (Kugelschreiber) zu unterschreiben.
- Wir wünschen Ihnen viel Erfolg.

### Verteilungen in R:

Typ	Verteilung (F)	Quantil (Q)
Normalverteilung	pnorm	qnorm
t-Verteilung	pt	qt
$\chi^2$ -Verteilung	pchisq	qchisq
F-Verteilung	pf	qf

**Poisson Verteilung:**  $P_\lambda(X = k) = \lambda^k \cdot e^{-\lambda}/k!$ ;  $E[X] = \lambda$ ;  $\text{var}(X) = \lambda$

**Exponentialverteilung:**  $f_\lambda(X) = \begin{cases} \lambda e^{-\lambda X} & X \geq 0 \\ 0 & \text{sonst} \end{cases}$  ;

$$F_\lambda(X) = \begin{cases} 1 - e^{-\lambda X} & X \geq 0 \\ 0 & \text{sonst} \end{cases} ;$$

$$E[X] = 1/\lambda; \text{var}(X) = 1/\lambda^2$$

**Einige Stammfunktionen:**  $\int x dx = \frac{1}{2}x^2 + C$ ;

$$\int x^n dx = x^{n+1}/(n+1) + C; \int \frac{1}{x} dx = \log x + C;$$

$$\int a^x dx = a^x / \log a + C$$

### Ableitung der Log-Likelihood Funktion:

$$\frac{d}{d\theta} \log L(x_1, \dots, x_n | \theta) = \frac{f'(x_1 | \theta)}{f(x_1 | \theta)} + \dots + \frac{f'(x_n | \theta)}{f(x_n | \theta)}$$

**Erwartungswert:**  $E(c \cdot X) = c \cdot E(X)$ ;

$$E(X + Y) = E(X) + E(Y)$$

**Varianz:**  $\text{var}(c \cdot X) = c^2 \cdot \text{var}(X)$ ;

$$\text{var}(X + Y) = \text{var}(X) + \text{var}(Y) + 2 \cdot \text{cov}(X, Y)$$

**Varianz von  $\bar{X}$ :**  $\text{var}(\bar{X}) = \sigma_X^2/n$

**Standardabweichung von  $\bar{X}$ :**  $\sigma_{\bar{X}} = \sigma_X/\sqrt{n}$

**Schätzer für Erwartungswert:**  $\hat{\mu}_X = \bar{X} = \frac{1}{n} \sum_i X_i$

**Schätzer für Varianz:**  $\hat{\sigma}_X^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$

**Schätzer für Standardabweichung von X:**

$$\hat{\sigma}_X = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2}$$

**Schätzer für  $\sigma_{\bar{X}}$ :**  $\hat{\sigma}_{\bar{X}} = \hat{\sigma}_X / \sqrt{n}$

**Bias:**  $\text{Bias}(\hat{\theta}, \theta) = E(\hat{\theta}) - \theta$

**Konfidenzintervall für den Mittelwert:**

$$\left[ \bar{x} + \sigma_{\bar{x}} \cdot Q\left(\frac{\alpha}{2}\right); \bar{x} - \sigma_{\bar{x}} \cdot Q\left(\frac{\alpha}{2}\right) \right]$$

**Fehler 1. und 2. Art**

		tatsächliche Situation	
		$H_0$ falsch	$H_0$ wahr
Testergebnis	$H_0$ wird abgelehnt (positiv)	$1 - \beta$ , Power Sensitivität	$\alpha$ , Signifikanzniveau Fehler 1. Art
	$H_0$ wird angenommen (negativ)	$\beta$ Fehler 2. Art	$1 - \alpha$ Spezifität

**Vergleich von Mittelwerten (unverbundene Stichproben)**

$$\frac{\bar{x}_A - \bar{x}_B}{\sqrt{\frac{\hat{\sigma}_A^2}{n_A} + \frac{\hat{\sigma}_B^2}{n_B}}} \sim t_{n_A + n_B - 2}$$

**Vergleich von Mittelwerten (verbundene Stichproben)**

$$g = \frac{\bar{\Delta}}{\hat{\sigma}_{\Delta}} \sim t_{n-1} \text{ mit } \Delta_i = X_i - Y_i \text{ und}$$

$$\hat{\sigma}_{\Delta} = \frac{1}{\sqrt{n}} \sqrt{\frac{\sum_{i=1}^n (\Delta_i - \bar{\Delta})^2}{n-1}}$$

$$\chi^2\text{-Kontingenztest } e_{ij} = \frac{\sum_{j=1}^k X_{ij} \cdot \sum_{i=1}^n X_{ij}}{\sum_{i=1}^n \sum_{j=1}^k X_{ij}}$$

$$g = \sum_{i=1}^n \sum_{j=1}^k \frac{(X_{ij} - e_{ij})^2}{e_{ij}} \sim \chi^2_{(n-1) \cdot (k-1)}$$

$$\chi^2\text{-Anpassungstest: } g = \sum_{i=1}^k \frac{(X(a_i) - n \cdot P(a_i))^2}{n \cdot P(a_i)} \sim \chi^2_{k-1}$$

**Test von Mittelwerten:**  $g = \frac{\bar{x} - \mu_0}{\hat{\sigma}_{\bar{x}}}$ . Falls  $X \sim N$ :  $g \sim t_{n-1}$

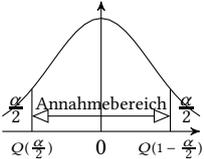
oder  $g^2 \sim F_{1, n-1}$ .

Falls  $n \rightarrow \infty$ :  $g \sim N(0, 1)$  oder  $g^2 \sim F_{1, \infty}$ .

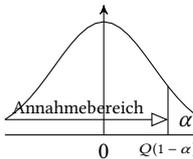
**AIC** =  $-2 \cdot L + 2 \cdot k$  (dabei ist  $L$  die Likelihood des Modells und  $k$  die Anzahl der Parameter).

**Signifikanztest:** Teststatistik:  $g = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$

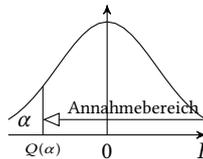
zweiseitig ( $H_1 : \mu \neq \mu_0$ )



einseitig ( $H_1 : \mu > \mu_0$ )

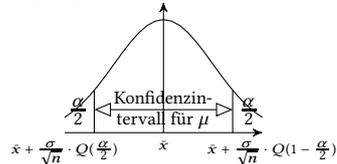


einseitig ( $H_1 : \mu < \mu_0$ )



$H_0$  wird verworfen, wenn  $g$  nicht im Annahmebereich ist.

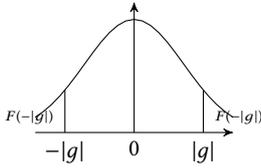
**Konfidenzintervall:**



$H_0 : \mu = \mu_0$  wird verworfen, wenn  $\mu_0$  nicht im Konfidenzintervall für  $\mu$  ist.

**p-Wert:** Teststatistik:  $g = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$

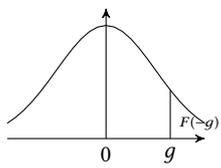
zweiseitig ( $H_1 : \mu \neq \mu_0$ )



$$p = 2 \cdot F(-|g|)$$

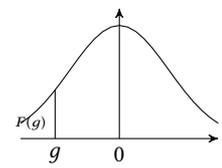
$H_0$  wird verworfen, wenn  $p < \alpha$ .

einseitig ( $H_1 : \mu > \mu_0$ )



$$p = F(-g)$$

einseitig ( $H_1 : \mu < \mu_0$ )



$$p = P(g)$$

	0.9	0.95	0.975	0.99	0.995	0.9975	0.999
qnorm(x)	1.28	1.64	1.96	2.33	2.58	2.81	3.09
qt(x,1)	3.08	6.31	12.71	31.82	63.66	127.32	318.31
qt(x,2)	1.89	2.92	4.30	6.96	9.92	14.09	22.33
qt(x,3)	1.64	2.35	3.18	4.54	5.84	7.45	10.21
qt(x,4)	1.53	2.13	2.78	3.75	4.60	5.60	7.17
qt(x,5)	1.48	2.02	2.57	3.36	4.03	4.77	5.89
qt(x,6)	1.44	1.94	2.45	3.14	3.71	4.32	5.21
qt(x,7)	1.41	1.89	2.36	3.00	3.50	4.03	4.79
qt(x,8)	1.40	1.86	2.31	2.90	3.36	3.83	4.50
qt(x,9)	1.38	1.83	2.26	2.82	3.25	3.69	4.30
qt(x,10)	1.37	1.81	2.23	2.76	3.17	3.58	4.14

	0.001	0.0025	0.005	0.01	0.025	0.05	0.1	0.9	0.95	0.975	0.99	0.995	0.9975	0.999
qchisq(x,1)	0.00	0.00	0.00	0.00	0.00	0.00	0.02	2.71	3.84	5.02	6.63	7.88	9.14	10.83
qchisq(x,2)	0.00	0.01	0.01	0.02	0.05	0.10	0.21	4.61	5.99	7.38	9.21	10.60	11.98	13.82
qchisq(x,3)	0.02	0.04	0.07	0.11	0.22	0.35	0.58	6.25	7.81	9.35	11.34	12.84	14.32	16.27
qchisq(x,4)	0.09	0.14	0.21	0.30	0.48	0.71	1.06	7.78	9.49	11.14	13.28	14.86	16.42	18.47
qchisq(x,5)	0.21	0.31	0.41	0.55	0.83	1.15	1.61	9.24	11.07	12.83	15.09	16.75	18.39	20.52
qchisq(x,6)	0.38	0.53	0.68	0.87	1.24	1.64	2.20	10.64	12.59	14.45	16.81	18.55	20.25	22.46
qchisq(x,7)	0.60	0.79	0.99	1.24	1.69	2.17	2.83	12.02	14.07	16.01	18.48	20.28	22.04	24.32
qchisq(x,8)	0.86	1.10	1.34	1.65	2.18	2.73	3.49	13.36	15.51	17.53	20.09	21.95	23.77	26.12
qchisq(x,9)	1.15	1.45	1.73	2.09	2.70	3.33	4.17	14.68	16.92	19.02	21.67	23.59	25.46	27.88
qchisq(x,10)	1.48	1.83	2.16	2.56	3.25	3.94	4.87	15.99	18.31	20.48	23.21	25.19	27.11	29.59

**1 Klausur BW24.1, Mai 2019**

Ihre Version ist . Bitte vergessen Sie nicht, die Version in den Lösungsbogen zu übertragen. Lösen Sie die Aufgaben zunächst hier (und ggf. auf dem Schmierpapier) und übertragen Sie die Lösungen zum Schluss in den Lösungsbogen. Geben Sie nur den Lösungsbogen ab, und behalten Sie dieses Aufgabenblatt. Wir wünschen Ihnen viel Erfolg!

**Aufgabe 1:** Ihre Stichprobe der Zufallsvariablen  $X$  enthält  $n$  unabhängige und normalverteilte Beobachtungen:  $X_1, \dots, X_n$ . Welche Schätzfunktionen für  $E(X)$  sind erwartungstreu?

1:	<sup>a</sup> $2X_4 - X_5$	<sup>b</sup> $X_1 + X_2 - X_5$	<sup>c</sup> $X_2 - X_5$	<sup>d</sup> $\sum_{i=1}^n X_i$	<sup>e</sup> $\frac{1}{3}(X_2 + X_3)$
----	---------------------------	--------------------------------	--------------------------	---------------------------------	---------------------------------------

**Aufgabe 2:** Betrachten Sie weiter die obige Stichprobe. Die Varianz von  $X$  sei  $\sigma_X^2$ . Wie groß ist die Varianz von  $2X_1 - 2X_2$ ?

2:	<sup>a</sup> anderer Wert	<sup>b</sup> 0	<sup>c</sup> $4\sigma_X^2$	<sup>d</sup> $8\sigma_X^2$	<sup>e</sup> $16\sigma_X^2$
----	---------------------------	----------------	----------------------------	----------------------------	-----------------------------

**Aufgabe 3:** Die Zufallsvariable  $X$  hat eine Varianz von  $\sigma_X^2 = 345$ . Sie haben vor, eine Stichprobe von  $n$  unabhängigen und identisch verteilten Beobachtungen zu ziehen. Berechnen Sie in R, wie groß Ihre Stichprobe sein muss, damit die Varianz des Stichprobenmittelwertes  $\sigma_{\bar{X}}^2 = 15$  ist?

3:	<sup>a</sup> anderer Wert	<sup>b</sup> 345/15	<sup>c</sup> $\sqrt{345/15}$	<sup>d</sup> $\sqrt{345}/15$	<sup>e</sup> $345/\sqrt{15}$
----	---------------------------	---------------------	------------------------------	------------------------------	------------------------------

**Aufgabe 4:** Eine Zufallsvariable  $X$  ist wie folgt verteilt:  $P(X = A) = \theta$ ,  $P(X = B) = 2\theta$ ,  $P(X = C) = 1 - 3\theta$  wobei  $\theta \in [0, 1/3]$ . Eine Stichprobe ergibt die folgenden Beobachtungen:  $\{A, B, A, A, C\}$ . Was ist der Maximum-Likelihood Schätzer für  $\theta$ ?

4:	<sup>a</sup> anderer Wert	<sup>b</sup> 0	<sup>c</sup> 4/15	<sup>d</sup> 1/3	<sup>e</sup> 1
----	---------------------------	----------------	-------------------	------------------	----------------

**Aufgabe 5:** Eine Zufallsvariable  $X$  folgt einer Verteilung  $\mathcal{X}_\theta$  mit Erwartungswert  $E(X) = 1/(\theta + 1)$  und Varianz  $1/(\theta^2 + 1)$  mit  $\theta > 0$ . Die Variable  $x$  enthält Ihre Stichprobe. Wie berechnen Sie mit R den Momentenschätzer für  $\theta$  auf Basis des ersten Moments?

5:	<sup>a</sup> anderer Wert	<sup>b</sup> $1/\text{mean}(x)-1$	<sup>c</sup> $1/(\text{mean}(x)-1)$	<sup>d</sup> $1/\text{mean}(x-1)$	<sup>e</sup> $\text{mean}(1/x-1)$
----	---------------------------	-----------------------------------	-------------------------------------	-----------------------------------	-----------------------------------

**Aufgabe 6:** Sie nehmen an, dass die Zufallsvariable  $Y$  einer Normalverteilung folgt:  $Y \sim N(\mu, \sigma^2)$ . Sie führen das Kommando

```
summary(MCMCregress(y ~ 1))
```

aus und erhalten die folgende Ausgabe:

```
1. Empirical mean and standard deviation for each variable,
plus standard error of the mean:
      Mean      SD Naive SE Time-series SE
(Intercept) -24.14 0.5257 0.005257      0.005408
sigma2      26.68 3.9133 0.039133      0.040175
2. Quantiles for each variable:
      2.5%    25%    50%    75%    97.5%
(Intercept) -25.17 -24.49 -24.14 -23.79 -23.10
sigma2      20.07 23.89 26.35 29.07 35.23
```

Welche Aussagen über die a-posteriori Verteilung von  $Y$  treffen zu?

- 6a: Mit Wahrscheinlichkeit 95% ist  $\sigma < 35.23$
- 6b: Mit Wahrscheinlichkeit 25% ist  $\sigma^2 < 29.07$
- 6c: Mit Wahrscheinlichkeit 75% ist  $\mu > -24.49$
- 6d: Mit Wahrscheinlichkeit 25% ist  $y < -23.79$
- 6e: Mit Wahrscheinlichkeit 2.5% ist  $y < -23.10$

**Aufgabe 7:** Sie schätzen das folgende Regressionsmodell:

$$Y \sim \beta_0 + \beta_1 X + \epsilon$$

Sie nehmen an, dass  $\epsilon$  einer Normalverteilung folgt:  $\epsilon \sim N(0, \sigma^2)$ . Sie führen das Kommando

```
summary(MCMCregress(y ~ x))
```

aus und erhalten folgenden Output:

```
1. Empirical mean and standard deviation for each variable,
plus standard error of the mean:
      Mean      SD Naive SE Time-series SE
(Intercept)  6.228 0.3104 0.003104      0.003104
x           -10.036 0.3372 0.003372      0.003372
sigma2      8.892 1.3597 0.013597      0.013992
2. Quantiles for each variable:
      2.5%    25%    50%    75%    97.5%
(Intercept)  5.623  6.022  6.231  6.434  6.836
x           -10.694 -10.263 -10.031 -9.811 -9.375
sigma2      6.622  7.936  8.753  9.712 11.888
```

Welche Aussagen über die a-posteriori Verteilung der Parameter Ihres Modells treffen zu?

- 7a: Mit Wahrscheinlichkeit 50% ist  $\beta_0 < 6.231$
- 7b: Mit Wahrscheinlichkeit 50% ist  $\beta_0 > 6.231$
- 7c: Mit Wahrscheinlichkeit 50% ist  $x < 6.231$
- 7d: Mit Wahrscheinlichkeit 95% ist  $\beta_1 \in [-10.694, -9.375]$
- 7e: Mit Wahrscheinlichkeit 50% ist  $\beta_1 \in [-10.263, -9.811]$

**Aufgabe 8:** Eine Zufallsvariable  $X$  ist normalverteilt mit unbekanntem Mittelwert und unbekannter Standardabweichung. Die Variable  $x$  enthält Ihre Stichprobe mit 18 Beobachtungen. Wie ermitteln Sie die Obergrenze des 99%-Konfidenzintervalls für den Mittelwert?

- 8a: Keine der folgenden Antworten ist richtig.
- 8b:  $\text{mean}(x) - \text{qt}(.99, \text{df}=17) * \text{sd}(x) / \sqrt{18}$
- 8c:  $\text{mean}(x) + \text{qt}(.99, \text{df}=17) * \text{sd}(x) / \sqrt{18}$
- 8d:  $\text{mean}(x) - \text{qt}(.98, \text{df}=17) * \text{sd}(x)$
- 8e:  $\text{mean}(x) + \text{qt}(.98, \text{df}=17) * \text{sd}(x)$

**Aufgabe 9:** Eine Zufallsvariable  $X$  ist normalverteilt mit unbekanntem Mittelwert und bekannter Standardabweichung  $s$ . Die Variable  $x$  enthält Ihre Stichprobe mit 18 Beobachtungen. Wie ermitteln Sie die Breite des 90%-Konfidenzintervalls für den Mittelwert?

- 9a: Keine der folgenden Antworten ist richtig.
- 9b:  $2 * \text{qnorm}(.90) * s / \sqrt{18}$
- 9c:  $2 * \text{qnorm}(.95) * s^2 / 18$
- 9d:  $2 * \text{qnorm}(.80) * s$
- 9e:  $2 * \text{qnorm}(.90) * s$

**Aufgabe 10:** Eine Zufallsvariable  $X$  ist normalverteilt mit unbekanntem Mittelwert  $\mu$  und unbekannter Varianz  $\sigma^2$ . Die Variable  $x$  enthält Ihre Stichprobe mit  $n$  Beobachtungen. Ihre Alternativhypothese ist  $\mu < 5$ . Wie berechnen Sie einen  $p$ -Wert zum Test der Nullhypothese?

- 10a: Keine der folgenden Antworten ist richtig.
- 10b:  $\text{pt}((\text{mean}(x) - 5) / \text{sd}(x) * \sqrt{n}), n-1)$
- 10c:  $1 - \text{pt}((\text{mean}(x) - 5) / \text{sd}(x) * \sqrt{n}), n-1)$
- 10d:  $\text{pt}((\text{mean}(x) - 5) / \text{sd}(x) * n), n-1)$
- 10e:  $1 - \text{pt}((\text{mean}(x) - 5) / \text{sd}(x) * n), n-1)$

**Aufgabe 11:** Betrachten Sie weiter die obige Stichprobe und die obige Hypothese. Mit welchem Kommando können Sie einen Test Ihrer Nullhypothese durchführen?

- 11a: Keine der folgenden Antworten ist richtig.
- 11b:  $\text{t.test}(x, \text{mu}=5, \text{alternative}="less")$
- 11c:  $\text{t.test}(x, \text{mu}=5, \text{alternative}="greater")$
- 11d:  $\text{t.test}(x, \text{mu}<5)$

11e: `t.test(x,mu>5)`

**Aufgabe 12:** Eine Zufallsvariable  $X$  ist normalverteilt mit unbekanntem Mittelwert  $\mu$  und unbekannter Varianz  $\sigma^2$ . Die Variable  $x$  enthält Ihre Stichprobe mit  $n$  Beobachtungen. Ihre Alternativhypothese ist  $\mu \neq 5$ . Wie berechnen Sie einen  $p$ -Wert zum Test der Nullhypothese?

- 12a: Keine der folgenden Antworten ist richtig.
- 12b: `2*pt(-abs(mean(x)-5)/sd(x)*sqrt(n),n-1)`
- 12c: `1-2*pt(-abs(mean(x)-5)/sd(x)*sqrt(n),n-1)`
- 12d: `2*pt(-abs(mean(x)-5)/sd(x)*n,n-1)`
- 12e: `1-2*pt(-abs(mean(x)-5)/sd(x)*n,n-1)`

**Aufgabe 13:** Die Zufallsvariable  $X$  folgt einer Normalverteilung mit unbekanntem Mittelwert  $\mu$  und unbekannter Varianz  $\sigma^2$ . Ihre Alternativhypothese ist, der Mittelwert einer Verteilung ist größer als ein nullhypothetischer Wert  $\mu_0$ . Ihre Stichprobe von 14 Beobachtungen ergibt den Mittelwert  $\bar{x}$ . Mit der Formel  $g = (\bar{x} - \mu_0) / \hat{\sigma}_{\bar{x}}$  bestimmen Sie Ihre Teststatistik  $g$ . Wann lehnen Sie Ihre Nullhypothese auf einem Signifikanzniveau von 5% ab?

- 13a: Keine der folgenden Antworten ist richtig.
- 13b: Wenn  $g > 1 - qt(.05, df=13)$
- 13c: Wenn  $g > qt(.05, df=13)$
- 13d: Wenn  $g < 1 - qt(.05, df=13)$
- 13e: Wenn  $g < qt(.05, df=13)$

**Aufgabe 14:** Sie vergleichen zwei Stichproben. In Ihrem Datensatz hat die Variable  $G$  den Wert 0, wenn eine Beobachtung aus der Population  $A$  gezogen wurde.  $G$  hat den Wert 1, wenn eine Beobachtung aus der Population  $B$  gezogen wurde.  $y$  ist die jeweilige Stichprobenbeobachtung. Sie führen das Kommando

```
summary(MCMCregress(y ~ G))
```

aus und erhalten folgenden Output:

```

1. Empirical mean and standard deviation for each variable,
plus standard error of the mean:
      Mean      SD Naive SE Time-series SE
(Intercept) -0.5938 0.7955 0.007955      0.007955
G            6.4845 1.1583 0.011583      0.011583
sigma2      17.7737 3.6570 0.036570      0.037929
2. Quantiles for each variable:
      2.5%   25%   50%   75%   97.5%
(Intercept) -2.157 -1.116 -0.5826 -0.06818 0.9508
G            4.198  5.720  6.4795  7.24578  8.8108
sigma2      12.036 15.198 17.2847 19.87401 26.2567

```

Welche der folgenden Aussagen treffen zu?

- 14a: Die Wahrscheinlichkeit, eine Stichprobe zu ziehen, die wenigstens so advers zur Nullhypothese ist, wie diese Stichprobe, ist 0.011583.
- 14b: Die Wahrscheinlichkeit, dass der Mittelwert für Beobachtungen aus  $B$  größer ist, als für solche aus  $A$ , ist größer als 97.5%.
- 14c: Die Wahrscheinlichkeit, dass der Mittelwert für Beobachtungen aus  $B$  um vier Einheiten größer ist, als für solche aus  $A$ , ist größer als 97.5%.
- 14d: Die Wahrscheinlichkeit, dass Beobachtungen aus  $B$  um vier Einheiten größer sind, als solche aus  $A$ , ist größer als 97.5%.
- 14e: Der Mittelwert der Beobachtungen aus  $A$  liegt mit Wahrscheinlichkeit 50% zwischen  $-1.116$  und  $-0.06818$ .

**Aufgabe 15:** Betrachten Sie das folgende Testergebnis.  $x$  und  $y$  sind jeweils Stichproben der Zufallsvariablen  $X$  bzw.  $Y$ . Verwenden Sie ein Signifikanzniveau von 5%.

```

Welch Two Sample t-test
data: x and y
t = 2.2034, df = 29.771, p-value = 0.9823
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval:
 -Inf 1.145146
sample estimates:
mean of x mean of y
0.09248883 -0.55431143

```

Was können Sie aus dem Testergebnis schließen?

- 15a: Es wurde ein paarweiser Test durchgeführt.
- 15b: Die Alternativhypothese in diesem Test war  $E(X) - E(Y) > 0$ .
- 15c: Die Nullhypothese  $E(X) - E(Y) \leq -1$  wird abgelehnt.
- 15d: Die Nullhypothese  $E(X) - E(Y) \geq 2$  wird abgelehnt.
- 15e: Die Nullhypothese  $E(X) - E(Y) \leq 1$  wird abgelehnt.

**Aufgabe 16:** Betrachten Sie das folgende Testergebnis. Verwenden Sie ein Signifikanzniveau von 5%.

```

Wilcoxon signed rank test
data: x and y
V = 105, p-value = 0.02884
alternative hypothesis: true location shift is greater than 0
Was können Sie aus dem Testergebnis schließen?

```

- 16a: Keine der folgenden Antworten ist richtig.
- 16b: Die Nullhypothese  $\text{median}(X - Y) \leq 0$  wird abgelehnt.
- 16c: Die Alternativhypothese war  $\text{median}(X - Y) > 0$ .
- 16d: Die Nullhypothese  $E(X) < E(Y)$  wird abgelehnt.
- 16e: Die Nullhypothese  $E(X) - E(Y) < 105$  wird abgelehnt.

**Aufgabe 17:** Das Merkmal  $X$  kann die Werte 1, 2 annehmen. Das Merkmal  $Y$  kann die Werte 1, 2, und 3 annehmen. Die Häufigkeiten in Ihrer Stichprobe sind durch die folgende Tabelle gegeben.

	$X = 1$	$X = 2$
$Y = 1$	3	18
$Y = 2$	12	16
$Y = 3$	18	13

Interpretieren Sie das folgende Testergebnis:

```

Pearson's Chi-squared test
data: Z
X-squared = 9.9468, df = 2, p-value = 0.00692

```

Was können Sie aus dem Testergebnis schließen?

- 17a: Für ein Signifikanzniveau von 1% findet man zwischen den beiden Merkmalen  $X$  und  $Y$  einen signifikanten Zusammenhang.
- 17b: Für ein Signifikanzniveau von 5% kann man die Nullhypothese,  $X$  und  $Y$  seien unabhängig, verwerfen.
- 17c: Das Testergebnis zeigt, dass  $Y$  und  $X$  positiv korreliert sind.
- 17d: Bei einem Signifikanzniveau von 1% kann man die Nullhypothese,  $X$  und  $Y$  seien voneinander abhängig, verwerfen.
- 17e: Die Nullhypothese,  $X$  und  $Y$  seien unabhängig, kann man nicht zum Signifikanzniveau von 0.1% verwerfen.

**Aufgabe 18:** In der folgenden Regression schätzen Sie den Effekt der Variablen  $a$ ,  $b$  und  $c$  auf die Variable  $Y$ . Sie erhalten den folgenden Output:

```
Call: lm(formula = y ~ a * b + c)
      Estimate Std. Error t value Pr(>|t|)
(Intercept) -1.00      0.79   -1.26  0.2091
a            1.00      1.00    1.00  0.3200
b           -2.00      1.00   -2.00  0.0485
c           -2.00      0.71   -2.83  0.0058
a:b          1.00      1.41    0.71  0.4813
Residual standard error: 3.464 on 91 degrees of freedom
Multiple R-squared:  0.1613, Adjusted R-squared:  0.1244
F-statistic: 4.375 on 4 and 91 DF, p-value: 0.002789
```

Welchen Wert für  $Y$  erwarten Sie, wenn  $a = b = c = 0$ ?

18: 

a	b	c	d	e
anderer Wert	-1	0	2	3

**Aufgabe 19:** Welchen Wert für  $Y$  erwarten Sie, wenn  $a = 1$  und  $b = c = 0$ ?

19: 

a	b	c	d	e
anderer Wert	-1	0	2	3

**Aufgabe 20:** Welchen Wert für  $Y$  erwarten Sie, wenn  $a = 1$ ,  $b = -1$  und  $c = -1$ ?

20: 

a	b	c	d	e
anderer Wert	-1	0	2	3

**Aufgabe 21:** Wie groß ist der marginale Effekt von  $b$  auf  $Y$  falls  $a = -1$ ?

21: 

a	b	c	d	e
anderer Wert	-3	-2	-1	0

**Aufgabe 22:** Wie groß ist der marginale Effekt von  $a$  auf  $Y$  falls  $b = -1$  und  $c = 3$ ?

22: 

a	b	c	d	e
anderer Wert	-2	0	1	4

**Aufgabe 23:** Wie bestimmen Sie für die obige Regression die Untergrenze des 90%-Konfidenzintervalls für den Koeffizienten von  $b$ ?

- 23a: Keine der folgenden Antworten ist richtig.  
 23b:  $-2 + qt(.05, df=91)$   
 23c:  $-2 - qt(.1, df=91)$   
 23d:  $-2 - qt(.9, df=91)$   
 23e:  $-2 - qt(.05, df=91)$

**Aufgabe 24:** Wie breit ist für die obige Regression das 90%-Konfidenzintervall für den Koeffizienten von  $a$ ?

- 24a: Keine der folgenden Antworten ist richtig.  
 24b:  $-2 * qt(.05, df=91)$   
 24c:  $qt(.95, df=91)$   
 24d:  $-2 * qt(.95, df=91)$

24e:  $qt(.05, df=91)$

**Aufgabe 25:** Betrachten Sie weiter die obige Regression. Gehen Sie von einem Signifikanzniveau von 5% aus. Nehmen Sie ferner an, dass die Residuen normalverteilt sind:

- 25a:  $b$  hat einen signifikanten Einfluss auf  $c$ .  
 25b: Falls  $a = 0$  hat  $b$  einen signifikanten Einfluss auf  $Y$ .  
 25c: Falls  $b = 0$  hat  $a$  einen signifikanten Einfluss auf  $Y$ .  
 25d: Unabhängig vom Wert von  $b$  hat  $c$  einen signifikanten Einfluss auf  $Y$ .  
 25e: Die Nullhypothese, der Koeffizient von  $a$  sei 2, wird nicht abgelehnt.

**Aufgabe 26:** Betrachten Sie weiter die obige Regression.  $b$  ist eine Dummy Variable und hat den Wert 1 im Fall A und den Wert 0 sonst. Nun codieren Sie  $b$  anders.  $b$  hat nun den Wert 0 im Fall A und den Wert 1 sonst. Was ist nun der marginale Effekt von  $b$  auf  $Y$  falls  $a = 0$ ?

26: 

a	b	c	d	e
anderer Wert	0	1	2	3

**Aufgabe 27:** Was ist (in der Regression mit der neuen Codierung von  $b$ ) der geschätzte Koeffizient von  $a$ ?

27: 

a	b	c	d	e
anderer Wert	0	1	2	3

**Aufgabe 28:** In der Gleichung

$\log Y = \beta_0 + \beta_1 X_1 + \beta_2 \log X_2 + u$  schätzen Sie  $\hat{\beta}_2 = 8.2$ . Was erwarten Sie?

- 28a: Keine der folgenden Antworten ist richtig.  
 28b: Wenn  $X_2$  um 1% steigt, steigt  $Y$  um 8.2%.  
 28c: Wenn  $X_2$  um 0.01 steigt, steigt  $Y$  um 8.2%.  
 28d: Wenn  $X_2$  um 1 steigt, steigt  $Y$  um 8.2%.  
 28e: Wenn  $X_2$  um 100% steigt, steigt  $Y$  um 8.2%.

**Aufgabe 29:** In der Gleichung  $Y = \beta_0 + \beta_1 X_1 + \beta_2 \log X_2 + u$  schätzen Sie  $\hat{\beta}_2 = 8.2$ . Was erwarten Sie?

- 29a: Keine der folgenden Antworten ist richtig.  
 29b: Wenn  $X_2$  um 1% steigt, steigt  $Y$  um 8.2%.  
 29c: Wenn  $X_2$  um 1% steigt, steigt  $Y$  um 0.082%.  
 29d: Wenn  $X_2$  um 1 steigt, steigt  $Y$  um 8.2%.  
 29e: Wenn  $X_2$  um 1 steigt, steigt  $Y$  um 820%.

**Aufgabe 30:** In der Gleichung  $\log Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + u$  schätzen Sie  $\hat{\beta}_2 = 8.2$ . Was erwarten Sie?

- 30a: Keine der folgenden Antworten ist richtig.  
 30b: Wenn  $X_2$  um 1% steigt, steigt  $Y$  um 8.2%.  
 30c: Wenn  $X_2$  um 1% steigt, steigt  $Y$  um 0.082%.  
 30d: Wenn  $X_2$  um 1 steigt, steigt  $Y$  um 8.2%.  
 30e: Wenn  $X_2$  um 0.01 steigt, steigt  $Y$  um 8.2%.

## 2 Klausur BW24.1, Februar 2019

Ihre Version ist . Bitte vergessen Sie nicht, die Version in den Lösungsbogen zu übertragen. Lösen Sie die Aufgaben zunächst hier (und ggf. auf dem Schmierpapier) und übertragen Sie die Lösungen zum Schluss in den Lösungsbogen. Geben Sie nur den Lösungsbogen ab, und behalten Sie dieses Aufgabenblatt. Wir wünschen Ihnen viel Erfolg!

**Aufgabe 1:** Ihre Stichprobe der Zufallsvariablen  $X$  enthält 6 unabhängige und normalverteilte Beobachtungen:  $X_1, \dots, X_6$ . Welche Schätzfunktionen für  $E(X)$  sind erwartungstreu?

1: 

a	b	c	d	e
$\frac{X_5}{2} + \frac{X_6}{2}$	$X_4$	$X_1 - X_2 - X_3$	$\frac{1}{6} \sum_{i=1}^6 X_i$	$\sum_{i=1}^6 X_i$

**Aufgabe 2:** Welche Schätzfunktionen für  $E(X)$  sind weniger effizient als  $\frac{X_5}{2} + \frac{X_6}{2}$ ?

2:	$\frac{X_5}{2} - \frac{X_6}{2}$	$X_4$	$X_1 - X_2 - X_3$	$\frac{1}{5} \sum_{i=1}^5 X_i$	$\sum_{i=1}^5 X_i$
----	---------------------------------	-------	-------------------	--------------------------------	--------------------

**Aufgabe 3:** Die Zufallsvariable  $X$  hat eine Varianz von  $\sigma_X^2$ . Sie haben vor, eine Stichprobe von  $n$  unabhängigen und identisch verteilten Beobachtungen zu ziehen. Wie groß muß  $n$  sein, damit die Varianz des Stichprobenmittelwerts  $\sigma_{\bar{X}}^2 = 12$  ist?

3:	a anderer Wert	b $\sigma_X^2/12$	c $\sigma_X/\sqrt{12}$	d $\sigma_X^2/12^2$	e $12 \cdot \sigma_X^2$
----	----------------	-------------------	------------------------	---------------------	-------------------------

**Aufgabe 4:**

Die Dichtefunktion von  $X$  ist

$$f(X|\theta) = \begin{cases} \frac{2X-6}{(\theta-6)\theta} & \text{falls } X \in [6, \theta] \\ 0 & \text{sonst} \end{cases}$$

Was ist der ML-Schätzer für  $\theta$ , wenn die Stichprobe die Beobachtungen: {18, 19, 27, 31} enthält?

4:	a anderer Wert	b 6	c 18	d 31	e $\infty$
----	----------------	-----	------	------	------------

**Aufgabe 5:** Eine Zufallsvariable  $X$  hat den Erwartungswert  $E(X) = 1/\theta$  und die Varianz  $1/\theta^2$ . Die Variable  $x$  enthält Ihre Stichprobe. Wie berechnen Sie mit R den Momentenschätzer für  $\theta$  auf Basis des ersten Moments?

5:	a anderer Wert	b $1/\text{mean}(x)$	c $1/\text{var}(x)$	d $\text{var}(x)/\text{mean}(x)$	e $2*\text{sd}(x)$
----	----------------	----------------------	---------------------	----------------------------------	--------------------

**Aufgabe 6:** Sie nehmen an, dass die Zufallsvariable  $Y$  einer Normalverteilung folgt:  $Y \sim N(\mu, \sigma^2)$ . Sie führen das Kommando

```
summary(MCMCregress(y ~ 1))
```

aus und erhalten die folgende Ausgabe:

```
1. Empirical mean and standard deviation for each variable,
plus standard error of the mean:
      Mean      SD Naive SE Time-series SE
(Intercept)  9.393  0.4065  0.004065      0.004127
sigma2      11.273  1.9670  0.019670      0.020123
2. Quantiles for each variable:
      2.5%  25%   50%   75%  97.5%
(Intercept)  8.596  9.122  9.394  9.661 10.20
sigma2      8.031  9.845 11.070 12.452 15.71
```

Welche Aussagen über die a-posteriori Verteilung von  $Y$  treffen zu?

- 6a: Mit Wahrscheinlichkeit 95% ist  $\sigma < 11.070$
- 6b: Mit Wahrscheinlichkeit 25% ist  $\sigma^2 < 9.845$
- 6c: Mit Wahrscheinlichkeit 75% ist  $\mu > 9.122$
- 6d: Mit Wahrscheinlichkeit 25% ist  $\mu < 9.661$
- 6e: Mit Wahrscheinlichkeit 2.5% ist  $y < 8.596$

**Aufgabe 7:** Sie schätzen das folgende Regressionsmodell:

$$Y \sim \beta_0 + \beta_1 X + \epsilon$$

Sie nehmen an, dass  $\epsilon$  einer Normalverteilung folgt:  $\epsilon \sim N(0, \sigma^2)$ . Sie führen das Kommando

```
summary(MCMCregress(y ~ x))
```

aus und erhalten folgenden Output:

1. Empirical mean and standard deviation for each variable, plus standard error of the mean:

```
      Mean      SD Naive SE Time-series SE
(Intercept)  4.852  0.3350  0.003350      0.003343
x            3.166  0.3307  0.003307      0.003345
sigma2      7.798  1.3886  0.013886      0.014308
```

2. Quantiles for each variable:

```
      2.5%  25%   50%   75%  97.5%
(Intercept)  4.197  4.631  4.855  5.075  5.504
x           2.517  2.942  3.172  3.386  3.814
sigma2     5.542  6.816  7.630  8.617 10.966
```

Welche Aussagen über die a-posteriori Verteilung der Parameter treffen zu?

- 7a: Mit Wahrscheinlichkeit 50% ist  $\beta_0 < 4.855$
- 7b: Mit Wahrscheinlichkeit 50% ist  $\beta_0 > 4.855$
- 7c: Mit Wahrscheinlichkeit 50% ist  $x < 3.172$
- 7d: Mit Wahrscheinlichkeit 75% ist  $\beta_1 > 2.942$
- 7e: Mit Wahrscheinlichkeit 25% ist  $\beta_1 > 3.386$

**Aufgabe 8:** Eine Zufallsvariable  $X$  ist normalverteilt mit unbekanntem Mittelwert und unbekannter Standardabweichung. Die Variable  $x$  enthält Ihre Stichprobe mit  $n$  Beobachtungen. Wie ermitteln Sie die Untergrenze des 95%-Konfidenzintervalls für den Mittelwert?

- 8a: Keine der folgenden Antworten ist richtig.
- 8b:  $\text{mean}(x) - \text{qt}(.05, \text{df}=n-1) * \text{var}(x) / \text{sqrt}(n)$
- 8c:  $\text{mean}(x) - \text{qt}(.025, \text{df}=n-1) * \text{sqrt}(\text{var}(x)/n)$
- 8d:  $\text{mean}(x) - \text{qt}(.025, \text{df}=n-1) * \text{sd}(x)$
- 8e:  $\text{mean}(x) + \text{qt}(.025, \text{df}=n-1) * \text{sd}(x)$

**Aufgabe 9:** Eine Zufallsvariable  $X$  ist normalverteilt mit unbekanntem Mittelwert und unbekannter Standardabweichung. Ihre Stichprobe enthält 16 Beobachtungen. Der Mittelwert Ihrer Stichprobe ist 12, die Varianz in Ihrer Stichprobe ist 9. Wie breit ist das 99%-Konfidenzintervall für den Mittelwert?

- 9a: Keine der folgenden Antworten ist richtig.
- 9b:  $3/4 * \text{pt}(.995, \text{df}=15)$
- 9c:  $3/4 * \text{qt}(.995, \text{df}=15)$
- 9d:  $3/2 * \text{qt}(.995, \text{df}=15)$
- 9e:  $4 * \text{qt}(.995, \text{df}=15) / 3$

**Aufgabe 10:** Ihre Stichprobe  $x$  enthält 9 Beobachtungen einer normalverteilten Zufallsvariable mit unbekanntem Mittelwert  $\mu$  und unbekannter Varianz  $\sigma^2$ . Sie finden einen Stichprobenmittelwert von 2 und eine Stichprobenvarianz von 36. Ihre Alternativhypothese ist  $\mu < 3$ . Wie berechnen Sie einen  $p$ -Wert zum Test der Nullhypothese?

- 10a: Keine der folgenden Antworten ist richtig.
- 10b:  $\text{pt}(1/2, 8)$
- 10c:  $\text{pt}(-1/2, 8)$
- 10d:  $\text{pt}(-1/36, 8)$
- 10e:  $\text{pt}(-1/6, 8)$

**Aufgabe 11:** Eine Zufallsvariable  $X$  ist normalverteilt mit unbekanntem Mittelwert  $\mu$  und unbekannter Varianz  $\sigma^2$ . Ihre Alternativhypothese ist  $\mu \neq 4$ . Wie können Sie einen Test Ihrer Nullhypothese durchführen?

- 11a: Keine der folgenden Antworten ist richtig.
- 11b:  $\text{t.test}(x, \text{mu}=4, \text{alternative}='less')$
- 11c:  $\text{t.test}(x, \text{mu}=4, \text{alternative}='greater')$

11d: `t.test(x,mu=4,alternative='equal')`  
 11e: `t.test(x,mu=4,alternative='two.sided')`

**Aufgabe 12:** Sie vergleichen zwei Stichproben,  $x$  und  $y$ . Ihre Nullhypothese ist, dass beide Stichproben aus der gleichen Population stammen. Ihre Alternativhypothese ist, dass der Mittelwert der Population, aus der Sie  $x$  gezogen haben, kleiner als der von  $y$  ist. Wie können Sie einen Test Ihrer Nullhypothese durchführen?

- 12a: Keine der folgenden Antworten ist richtig.  
 12b: `t.test(x,y,alternative='two.sided')`  
 12c: `t.test(x,y,alternative='less')`  
 12d: `t.test(x,y,alternative='left.sided')`  
 12e: `t.test(x,y,alternative='more')`

**Aufgabe 13:** Ihr Signifikanzniveau ist 5%. Bei welchen Werten für den  $p$ -Wert aus dem obigen Test können Sie Ihre Nullhypothese ablehnen?

13:	<sup>a</sup> 0.1	<sup>b</sup> 0.99	<sup>c</sup> 0.01	<sup>d</sup> 0.03	<sup>e</sup> 0.97
-----	------------------	-------------------	-------------------	-------------------	-------------------

**Aufgabe 14:** Nehmen Sie an, dass Sie im obigen (einseitigen) Test einen  $p$ -Wert von 0.04 erhalten haben. Nun wollen Sie eine andere Alternativhypothese testen, und zwar dass der Mittelwert der Population, aus der Sie  $x$  gezogen haben, größer oder kleiner als der von  $y$  ist. Sie wollen also einen zweiseitigen Test durchführen. Was ist der  $p$ -Wert für diesen zweiseitigen Test?

14:	<sup>a</sup> anderer Wert	<sup>b</sup> 0.02	<sup>c</sup> 0.04	<sup>d</sup> 0.08	<sup>e</sup> 0.96
-----	---------------------------	-------------------	-------------------	-------------------	-------------------

**Aufgabe 15:** Welche Wahrscheinlichkeit gibt der  $p$ -Wert an?

- 15a: Keine der folgenden Antworten ist richtig.  
 15b: Die Wahrscheinlichkeit, eine Stichprobe zu ziehen, die mindestens so advers zur Nullhypothese ist, wie die gezogene Stichprobe, falls die Nullhypothese falsch ist.  
 15c: Die Wahrscheinlichkeit, eine Stichprobe zu ziehen, die mindestens so advers zur Nullhypothese ist, wie die gezogene Stichprobe.  
 15d: Die Wahrscheinlichkeit eines Fehlers erster Art.  
 15e: Die Wahrscheinlichkeit eines Fehlers zweiter Art.

**Aufgabe 16:** Betrachten Sie das folgende Testergebnis.  $x$  und  $y$  sind jeweils Stichproben der Zufallsvariablen  $X$  bzw.  $Y$ . Verwenden Sie ein Signifikanzniveau von 5%.

```
Welch Two Sample t-test
data: x and y
t = 1.5696, df = 20.394, p-value = 0.06594
alternative hypothesis: true difference in means
is greater than 0
95 percent confidence interval:
-0.07426527 Inf
sample estimates:
mean of x mean of y
0.2684472 -0.4910772
```

Was können Sie aus dem Testergebnis schließen?

- 16a: Es wurde ein paarweiser Test durchgeführt.  
 16b: Die Alternativhypothese in diesem Test war  $E(X) - E(Y) > 0$ .  
 16c: Die Nullhypothese  $E(X) - E(Y) \leq -0.1$  wird abgelehnt.  
 16d: Die Nullhypothese  $E(X) - E(Y) \leq 1$  wird abgelehnt.  
 16e: Die Nullhypothese  $E(X) - E(Y) \geq 2$  wird abgelehnt.

**Aufgabe 17:** Mit dem Kommando `wilcox.test(x ~ i)` erhalten Sie folgendes Testergebnis. Verwenden Sie ein Signifikanzniveau von 5%.

```
Wilcoxon rank sum test
data: x by i
W = 6, p-value = 0.07323
alternative hypothesis: true location shift is not equal to 0
Was können Sie aus dem Testergebnis schließen?
17a: Keine der folgenden Antworten ist richtig.
17b: Die getestete Nullhypothese wird nicht abgelehnt
17c: Die Alternativhypothese in diesem Test war  $E(X) < E(Y)$ .
17d: Die Nullhypothese  $E(X) \geq E(Y)$  wird abgelehnt.
17e: Die Nullhypothese  $E(X) - E(Y) \geq 6$  wird abgelehnt.
```

**Aufgabe 18:** Sie untersuchen den Einfluss eines Treatments  $t$  auf ein Messergebnis  $x$ . Die Variable  $t$  hat entweder den Wert 0 oder 1. Ihre Nullhypothese ist, dass die Messergebnisse in beiden Fällen im Mittel gleich ausfallen. Wie testen Sie Ihre Hypothese?

- 18a: Keine der folgenden Antworten ist richtig.  
 18b: `t.test(x ~ t)`  
 18c: `t.test(x ~ t,alternative='greater')`  
 18d: `t.test(x ~ t,alternative='smaller')`  
 18e: `t.test(x ~ t,alternative='equal')`

**Aufgabe 19:** Nun wollen Sie für die gleichen Daten ein Bayesianisches Verfahren anwenden. Sie verwenden das Kommando:

```
summary(MCMCregress(x ~ t))
```

und erhalten folgenden Output:

```
1. Empirical mean and standard deviation for each variable,
plus standard error of the mean:
      Mean   SD Naive SE Time-series SE
(Intercept) 0.53576 0.2123 0.002123    0.002118
t            0.07352 0.2767 0.002767    0.002767
sigma2       1.26449 0.2252 0.002252    0.002320
2. Quantiles for each variable:
      2.5%   25%   50%   75%  97.5%
(Intercept) 0.1206 0.3956 0.53753 0.6768 0.9485
t           -0.4718 -0.1088 0.07145 0.2544 0.6224
sigma2       0.8987  1.1053 1.23730 1.3973 1.7783
```

- 19a: Keine der folgenden Antworten ist richtig.  
 19b: Für  $t = 1$  sind die Messwerte im Mittel um 0.07352 größer als für  $t = 0$ .  
 19c: Der Messwert für  $t = 0$  ist im Mittel 0.1206.  
 19d: Für  $t = 1$  sind die Messwerte mit Wahrscheinlichkeit 0.2252 signifikant.  
 19e: Die Wahrscheinlichkeit, dass die Messwerte für  $t = 1$  im Mittel größer sind als für  $t = 0$ , ist größer als 50%.

**Aufgabe 20:**

Zwei Merkmale,  $X$  und  $Y$ , können jeweils die Werte 1, 2, 3 annehmen. Die Häufigkeiten in Ihrer Stichprobe sind durch die Tabelle  $Z$  gegeben:

	$X = 1$	$X = 2$	$X = 3$
$Y = 1$	28	20	30
$Y = 2$	24	16	29
$Y = 3$	13	6	26

Interpretieren Sie das folgende Testergebnis:

Pearson's Chi-squared test

data: Z

X-squared = 5.0299, df = 4, p-value = 0.2842

Was können Sie aus dem Testergebnis schließen?

20a: Für ein Signifikanzniveau von 10% findet man zwischen den beiden Merkmalen X und Y einen signifikanten Zusammenhang.

20b: Für ein Signifikanzniveau von 5% kann man die Nullhypothese, X und Y seien unabhängig, verwerfen.

20c: Bei einem Signifikanzniveau von 5% kann man die Nullhypothese, X und Y seien voneinander abhängig, verwerfen.

20d: Um zu testen, ob X und Y voneinander unabhängig sind, sollte man besser einen Rangsummentest verwenden.

20e: Es wurde ein einseitiger Test durchgeführt.

**Aufgabe 21:** In der folgenden Regression schätzen Sie den Effekt der Variablen a, b und c auf die Variable Y. Sie erhalten den folgenden Output:

```
Call: lm(formula = y ~ a * b + c)
      Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.00      0.79    2.53  0.01516 *
a            2.00      1.00    2.00  0.05184 .
b            3.00      1.00    3.00  0.00448 **
c            2.00      0.71    2.83  0.00708 **
a:b          2.00      1.41    1.41  0.16450

Residual standard error: 2.449 on 43 degrees of freedom
Multiple R-squared:  0.5825, Adjusted R-squared:  0.5437
F-statistic: 15 on 4 and 43 DF, p-value: 9.434e-08
```

Welchen Wert für Y erwarten Sie, wenn  $a = b = c = 0$ ?

21: 

a	anderer Wert	b	2.53	c	0	d	2	e	3
---	--------------	---	------	---	---	---	---	---	---

**Aufgabe 22:** Welchen Wert für Y erwarten Sie, wenn  $a = 1$  und  $b = c = 0$ ?

22: 

a	anderer Wert	b	0	c	2	d	3	e	4
---	--------------	---	---	---	---	---	---	---	---

**Aufgabe 23:** Welchen Wert für Y erwarten Sie, wenn  $a = 1$ ,  $b = -1$  und  $c = 1$ ?

23: 

a	anderer Wert	b	-1	c	0	d	1	e	3
---	--------------	---	----	---	---	---	---	---	---

**Aufgabe 24:** Wie groß ist der marginale Effekt von b auf Y falls  $a = -1$ ?

24: 

a	anderer Wert	b	-1	c	0	d	1	e	5
---	--------------	---	----	---	---	---	---	---	---

**Aufgabe 25:** Wie groß ist der marginale Effekt von a auf Y falls  $b = -1$  und  $c = 2$ ?

25: 

a	anderer Wert	b	-1	c	0	d	1	e	5
---	--------------	---	----	---	---	---	---	---	---

**Aufgabe 26:** Wie bestimmen Sie für die obige Regression die Obergrenze des 99%-Konfidenzintervalls für den Koeffizienten von a?

26a: Keine der folgenden Antworten ist richtig.

26b:  $2 - qt(.005, df=43)$

26c:  $2 + qt(.005, df=43)$

26d:  $2 + pt(.05, df=43)$

26e:  $2 + qt(.05, df=43)$

**Aufgabe 27:** Betrachten Sie weiter die obige Regression. Gehen Sie von einem Signifikanzniveau von 5% aus. Nehmen Sie ferner an, dass die Residuen normalverteilt sind:

27a: b hat einen signifikanten Einfluss auf c.

27b: Falls  $a = 0$  hat b einen signifikanten Einfluss auf Y.

27c: Falls  $b = 0$  hat a einen signifikanten Einfluss auf Y.

27d: Unabhängig vom Wert von b hat c einen signifikanten Einfluss auf Y.

27e: Die Nullhypothese, der Koeffizient von a sei 1, wird nicht abgelehnt.

**Aufgabe 28:** Betrachten Sie weiter die obige Regression. c ist eine Dummy Variable und hat den Wert 1 im Fall A und den Wert 0 sonst. Nun codieren Sie c anders. c hat nun den Wert 0 im Fall A und den Wert 1 sonst. Was ist nun der marginale Effekt von c auf Y?

28: 

a	anderer Wert	b	-2	c	-1	d	0	e	2
---	--------------	---	----	---	----	---	---	---	---

**Aufgabe 29:**

In der Gleichung  $Y = \beta_0 + \beta_1 X_1 + \beta_2 \log X_2 + u$  schätzen Sie  $\hat{\beta}_2 = 0.52$ . Was erwarten Sie?

29a: Keine der folgenden Antworten ist richtig.

29b: Wenn  $X_2$  um 1% steigt, steigt Y um 0.52.

29c: Wenn  $X_2$  um 1% steigt, steigt Y um 0.0052.

29d: Wenn  $X_2$  um 1 steigt, steigt Y um 0.52%.

29e: Wenn  $X_2$  um 1 steigt, steigt Y um 52%.

**Aufgabe 30:**

In der Gleichung  $\log Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + u$  schätzen Sie  $\hat{\beta}_2 = 1.6$ . Was erwarten Sie?

30a: Keine der folgenden Antworten ist richtig.

30b: Wenn  $X_2$  um 1% steigt, steigt Y um 1.6.

30c: Wenn  $X_2$  um 1% steigt, steigt Y um 0.016.

30d: Wenn  $X_2$  um 1 steigt, steigt Y um 1.6%.

30e: Wenn  $X_2$  um 0.01 steigt, steigt Y um 1.6%.

**Aufgabe 31:**

In der Gleichung  $\log Y = \beta_0 + \beta_1 X_1 + \beta_2 \log X_2 + u$  schätzen Sie  $\hat{\beta}_2 = 0.18$ . Was erwarten Sie?

31a: Keine der folgenden Antworten ist richtig.

31b: Wenn  $X_2$  um 1% steigt, steigt Y um 0.18%.

31c: Wenn  $X_2$  um 0.01 steigt, steigt Y um 0.18%.

31d: Wenn  $X_2$  um 1 steigt, steigt Y um 0.18%.

31e: Wenn  $X_2$  um 100% steigt, steigt Y um 0.18%.

### 3 Klausur BW24.1, Mai 2018

Ihre Version ist . Bitte vergessen Sie nicht, die Version in den Lösungsbogen zu übertragen. Lösen Sie die Aufgaben zunächst hier (und ggf. auf dem Schmierpapier) und übertragen Sie die Lösungen zum Schluss in den Lösungsbogen. Geben Sie nur den Lösungsbogen ab, und behalten Sie dieses Aufgabenblatt. Wir wünschen Ihnen viel Erfolg!

**Aufgabe 1:** Ihre Stichprobe der Zufallsvariablen  $X$  enthält  $n$  unabhängige und normalverteilte Beobachtungen:  $X_1, \dots, X_n$ . Welche Schätzfunktionen für  $E(X)$  sind erwartungstreu?

1:	$2X_4 - 2X_5$	$X_4 + X_5 - X_1$	$X_2$	$\sum_{i=1}^n X_i$	$\frac{1}{n} \sum_{i=1}^n X_i$
----	---------------	-------------------	-------	--------------------	--------------------------------

**Aufgabe 2:** Welche Schätzfunktionen für  $E(X)$  sind weniger effizient als  $(X_1 + X_2)/2$ ?

2:	$2X_4 - 2X_5$	$X_4 + X_5 - X_1$	$X_2$	$\sum_{i=1}^n X_i$	$\frac{1}{n} \sum_{i=1}^n X_i$
----	---------------	-------------------	-------	--------------------	--------------------------------

**Aufgabe 3:** Die Zufallsvariable  $X$  hat eine Varianz von 12. Sie haben vor, eine Stichprobe von  $n$  unabhängigen und identisch verteilten Beobachtungen zu ziehen. Wie groß muß  $n$  sein, damit die Standardabweichung des Stichprobenmittelwerts 10 ist?

3:	anderer Wert	12/10	10/12	$10/\sqrt{12}$	12/100
----	--------------	-------	-------	----------------	--------

**Aufgabe 4:**

Eine Zufallsvariable nimmt mit Wahrscheinlichkeit  $\theta$  den Wert  $A$  und mit Wahrscheinlichkeit  $1 - \theta$  den Wert  $B$  an. In Ihrer Stichprobe beobachten Sie 10 mal ein  $A$  und 1 mal ein  $B$ . Was ist der Maximum-Likelihood Schätzer für  $\theta$ ?

4:	anderer Wert	11/10	1/11	10/11	90/110
----	--------------	-------	------	-------	--------

**Aufgabe 5:** Eine Zufallsvariable  $X$  folgt einer Verteilung  $\mathcal{N}(\theta, \theta)$  mit Erwartungswert  $E(X) = \sqrt{\theta}$ . Die Variable  $x$  enthält Ihre Stichprobe. Wie berechnen Sie mit  $R$  den Momentenschätzer für  $\theta$  auf Basis des ersten Moments?

5:	anderer Wert	$1 - \text{mean}(x)^2$	$\text{sqrt}(x)$	$x^2$	$\text{mean}(x)^2$
----	--------------	------------------------	------------------	-------	--------------------

**Aufgabe 6:** Eine Zufallsvariable  $X$  ist normalverteilt mit unbekanntem Mittelwert und unbekannter Standardabweichung. Die Variable  $x$  enthält Ihre Stichprobe mit  $n$  Beobachtungen. Wie ermitteln Sie die Obergrenze des 90%-Konfidenzintervalls für den Mittelwert?

- 6a: Keine der folgenden Antworten ist richtig.  
 6b:  $\text{mean}(x) - \text{qt}(.05, \text{df}=(n-1)) * \text{var}(x) / \text{sqrt}(n)$   
 6c:  $\text{mean}(x) + \text{qt}(.05, \text{df}=(n-1)) * \text{var}(x) / \text{sqrt}(n)$   
 6d:  $\text{mean}(x) - \text{qt}(.025, \text{df}=(n-1)) * \text{sd}(x)$   
 6e:  $\text{mean}(x) + \text{qt}(.025, \text{df}=(n-1)) * \text{sd}(x)$

**Aufgabe 7:** Eine Zufallsvariable  $X$  ist normalverteilt mit unbekanntem Mittelwert und unbekannter Standardabweichung. Ihre Stichprobe enthält 28 Beobachtungen. Der Mittelwert Ihrer Stichprobe ist 18, die Varianz in Ihrer Stichprobe ist 16. Wie breit ist das 99%-Konfidenzintervall für den Mittelwert?

- 7a: Keine der folgenden Antworten ist richtig.  
 7b:  $8/28 * \text{qt}(.995, \text{df}=27)$   
 7c:  $4/28 * \text{qt}(.995, \text{df}=27)$   
 7d:  $8/\text{sqrt}(28) * \text{qt}(.995, \text{df}=27)$   
 7e:  $4/\text{sqrt}(28) * \text{qt}(.995, \text{df}=27)$

**Aufgabe 8:** Eine Zufallsvariable  $X$  ist normalverteilt mit unbekanntem Mittelwert  $\mu$  und unbekannter Varianz  $\sigma^2$ . Ihre Alternativhypothese ist  $\mu > 4$ . Wie berechnen Sie einen  $p$ -Wert zum Test der Nullhypothese?

- 8a: Keine der folgenden Antworten ist richtig.

- 8b:  $\text{pt}((\text{mean}(x)-4)/\text{sd}(x)*\text{sqrt}(n), n-1)$   
 8c:  $\text{pt}((4-\text{mean}(x))/\text{sd}(x)*\text{sqrt}(n), n-1)$   
 8d:  $1-\text{pt}((\text{mean}(x)-4)/\text{sd}(x)*n, n-1)$   
 8e:  $\text{pt}((4-\text{mean}(x))/\text{sd}(x), n-1)$

**Aufgabe 9:** Eine Zufallsvariable  $X$  ist normalverteilt mit unbekanntem Mittelwert  $\mu$  und unbekannter Varianz  $\sigma^2$ . Ihre Alternativhypothese ist  $\mu \neq 4$ . Wie berechnen Sie einen  $p$ -Wert zum Test der Nullhypothese?

- 9a: Keine der folgenden Antworten ist richtig.  
 9b:  $1-2*\text{pt}((\text{mean}(x)-4)/\text{sd}(x)*\text{sqrt}(n), n-1)$   
 9c:  $2*\text{pt}((-abs(\text{mean}(x)-4))/\text{sd}(x)*\text{sqrt}(n), n-1)$   
 9d:  $2*\text{pt}((4-\text{mean}(x))/\text{sd}(x), n-1)$   
 9e:  $1-2*\text{pt}((4-\text{mean}(x))/\text{sd}(x), n-1)$

**Aufgabe 10:** Welche Wahrscheinlichkeit gibt der  $p$ -Wert an?

- 10a: Keine der folgenden Antworten ist richtig.  
 10b: Die Wahrscheinlichkeit, eine Stichprobe zu ziehen, die mindestens so advers zur Nullhypothese ist, wie die gezogene Stichprobe, falls die Nullhypothese wahr ist.  
 10c: Die Wahrscheinlichkeit, eine Stichprobe zu ziehen, die mindestens so advers zur Nullhypothese ist, wie die gezogene Stichprobe, falls die Nullhypothese falsch ist.  
 10d: Die Wahrscheinlichkeit, eine Stichprobe zu ziehen, die mindestens so advers zur Nullhypothese ist, wie die gezogene Stichprobe.  
 10e: Die Wahrscheinlichkeit, die Nullhypothese abzulehnen, wenn sie wahr ist.

**Aufgabe 11:** Betrachten Sie das folgende Testergebnis.  $x$  und  $y$  sind jeweils Stichproben von  $X$  bzw.  $Y$ . Verwenden Sie ein Signifikanzniveau von 5%.

```
Welch Two Sample t-test
data: x and y
t = 4.4222, df = 19.575, p-value = 0.0001372
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 0.885849 Inf
sample estimates:
mean of x mean of y
0.1447171 -1.3084985
Was können Sie aus dem Testergebnis schließen?
11a: Es wurde ein paarweiser Test durchgeführt.
11b: Die Alternativhypothese in diesem Test war  $E(X) - E(Y) > 0$ .
11c: Die Nullhypothese  $E(X) - E(Y) \leq -1$  wird abgelehnt.
11d: Die Nullhypothese  $E(X) - E(Y) \leq 1$  wird abgelehnt.
11e: Die Nullhypothese  $E(X) - E(Y) \geq 2$  wird abgelehnt.
```

**Aufgabe 12:** Betrachten Sie das folgende Testergebnis. Verwenden Sie ein Signifikanzniveau von 5%.

```
Wilcoxon signed rank test
data: x and y
V = 72, p-value = 0.9976
alternative hypothesis: true location shift is less than 0
Was können Sie aus dem Testergebnis schließen?
12a: Keine der folgenden Antworten ist richtig.
12b: Die getestete Nullhypothese wird nicht abgelehnt.
12c: Die Alternativhypothese in diesem Test war  $E(X) < E(Y)$ .
12d: Die Nullhypothese  $E(X) \geq E(Y)$  wird abgelehnt.
```

12e: Die Nullhypothese  $E(X) - E(Y) \geq 72$  wird abgelehnt.

**Aufgabe 13:**

Zwei Merkmale,  $X$  und  $Y$ , können jeweils die Werte 1, 2, 3 annehmen. Die Häufigkeiten in Ihrer Stichprobe sind durch die folgende Tabelle gegeben.

	$X = 1$	$X = 2$	$X = 3$
$Y = 1$	21	28	16
$Y = 2$	23	19	2
$Y = 3$	24	22	16

Interpretieren Sie das folgende Testergebnis:

Pearson's Chi-squared test

data: Z

X-squared = 10.26, df = 4, p-value = 0.03626

Was können Sie aus dem Testergebnis schließen?

- 13a: Für ein Signifikanzniveau von 10% findet man zwischen den beiden Merkmalen  $X$  und  $Y$  einen signifikanten Zusammenhang.
- 13b: Für ein Signifikanzniveau von 5% kann man die Nullhypothese,  $X$  und  $Y$  seien unabhängig, verwerfen.
- 13c: Das Testergebnis zeigt, dass  $Y$  und  $X$  positiv korreliert sind.
- 13d: Bei einem Signifikanzniveau von 5% kann man die Nullhypothese,  $X$  und  $Y$  seien voneinander abhängig, verwerfen.
- 13e: Es wurde ein einseitiger Test durchgeführt.

**Aufgabe 14:** In der folgenden Regression schätzen Sie den Effekt der Variablen  $a$ ,  $b$  und  $c$  auf die Variable  $Y$ . Sie erhalten den folgenden Output:

```
Call: lm(formula = y ~ a * b + c)
      Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.00      0.79    2.53  0.0131
a            1.00      1.00    1.00  0.3200
b            1.00      1.00    1.00  0.3200
c            3.00      0.71    4.24  0.0001
a:b          3.00      1.41    2.12  0.0366

Residual standard error: 3.464 on 91 degrees of freedom
Multiple R-squared:  0.343, Adjusted R-squared:  0.3141
F-statistic: 11.87 on 4 and 91 DF, p-value: 8.33e-08
```

Welchen Wert für  $Y$  erwarten Sie, wenn  $a = b = c = 0$ ?

14: 

a	anderer Wert	b	-1	c	0	d	2	e	3
---	--------------	---	----	---	---	---	---	---	---

**Aufgabe 15:** Welchen Wert für  $Y$  erwarten Sie, wenn  $a = 1$  und  $b = c = 0$ ?

15: 

a	anderer Wert	b	-1	c	0	d	2	e	3
---	--------------	---	----	---	---	---	---	---	---

**Aufgabe 16:** Welchen Wert für  $Y$  erwarten Sie, wenn  $a = 1$ ,  $b = -1$  und  $c = 1$ ?

16: 

a	anderer Wert	b	-1	c	0	d	2	e	3
---	--------------	---	----	---	---	---	---	---	---

**Aufgabe 17:** Wie groß ist der marginale Effekt von  $b$  auf  $Y$  falls  $a = -1$ ?

17: 

a	anderer Wert	b	-2	c	0	d	1	e	4
---	--------------	---	----	---	---	---	---	---	---

**Aufgabe 18:** Wie groß ist der marginale Effekt von  $a$  auf  $Y$  falls  $b = -1$  und  $c = 2$ ?

18: 

a	anderer Wert	b	-2	c	0	d	1	e	4
---	--------------	---	----	---	---	---	---	---	---

**Aufgabe 19:** Wie bestimmen Sie für die obige Regression die Untergrenze des 90%-Konfidenzintervalls für den Koeffizienten von  $b$ ?

- 19a: Keine der folgenden Antworten ist richtig.
- 19b:  $1 + qt(.05, df=91)$
- 19c:  $1 - qt(.1, df=91)$
- 19d:  $1 - qt(.9, df=91)$
- 19e:  $1 - qt(.05, df=91)$

**Aufgabe 20:** Wie breit ist für die obige Regression das 90%-Konfidenzintervalls für den Koeffizienten von  $a$ ?

- 20a: Keine der folgenden Antworten ist richtig.
- 20b:  $-2 * qt(.05, df=91)$
- 20c:  $qt(.95, df=91)$
- 20d:  $-2 * qt(.95, df=91)$
- 20e:  $qt(.05, df=91)$

**Aufgabe 21:** Betrachten Sie weiter die obige Regression. Gehen Sie von einem Signifikanzniveau von 5% aus. Nehmen Sie ferner an, dass die Residuen normalverteilt sind:

- 21a:  $b$  hat einen signifikanten Einfluss auf  $c$ .
- 21b: Falls  $a = 0$  hat  $b$  einen signifikanten Einfluss auf  $Y$ .
- 21c: Falls  $b = 0$  hat  $a$  einen signifikanten Einfluss auf  $Y$ .
- 21d: Unabhängig vom Wert von  $b$  hat  $c$  einen signifikanten Einfluss auf  $Y$ .
- 21e: Die Nullhypothese, der Koeffizient von  $a$  sei 1, wird nicht abgelehnt.

**Aufgabe 22:** Betrachten Sie weiter die obige Regression.  $b$  ist eine Dummy Variable und hat den Wert 1 im Fall A und den Wert 0 sonst. Nun codieren Sie  $b$  anders.  $b$  hat nun den Wert 0 im Fall A und den Wert 1 sonst. Was ist nun der marginale Effekt von  $b$  auf  $Y$  falls  $a = 0$ ?

22: 

a	anderer Wert	b	-3	c	-1	d	0	e	1
---	--------------	---	----	---	----	---	---	---	---

**Aufgabe 23:** Was ist (in der Regression mit der neuen Codierung von  $b$ ) der geschätzte Koeffizient von  $a$ ?

23: 

a	anderer Wert	b	-1	c	1	d	3	e	4
---	--------------	---	----	---	---	---	---	---	---

**Aufgabe 24:**

In der Gleichung  $Y = \beta_0 + \beta_1 X_1 + \beta_2 \log X_2 + u$  schätzen Sie  $\hat{\beta}_2 = 0.31$ . Was erwarten Sie?

- 24a: Keine der folgenden Antworten ist richtig.
- 24b: Wenn  $X_2$  um 1% steigt, steigt  $Y$  um 0.31.
- 24c: Wenn  $X_2$  um 1% steigt, steigt  $Y$  um 0.0031.
- 24d: Wenn  $X_2$  um 1 steigt, steigt  $Y$  um 0.31%.
- 24e: Wenn  $X_2$  um 1 steigt, steigt  $Y$  um 31%.

**Aufgabe 25:**

In der Gleichung  $\log Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + u$  schätzen Sie  $\hat{\beta}_2 = 4.3$ . Was erwarten Sie?

- 25a: Keine der folgenden Antworten ist richtig.
- 25b: Wenn  $X_2$  um 1% steigt, steigt  $Y$  um 4.3.
- 25c: Wenn  $X_2$  um 1% steigt, steigt  $Y$  um 0.043.

- 25d: Wenn  $X_2$  um 1 steigt, steigt  $Y$  um 4.3%.  
 25e: Wenn  $X_2$  um 0.01 steigt, steigt  $Y$  um 4.3%.

**Aufgabe 26:**

In der Gleichung  $\log Y = \beta_0 + \beta_1 X_1 + \beta_2 \log X_2 + u$  schätzen

Sie  $\hat{\beta}_2 = 7.2$ . Was erwarten Sie?

- 26a: Keine der folgenden Antworten ist richtig.  
 26b: Wenn  $X_2$  um 1% steigt, steigt  $Y$  um 7.2%.  
 26c: Wenn  $X_2$  um 0.01 steigt, steigt  $Y$  um 7.2%.  
 26d: Wenn  $X_2$  um 1 steigt, steigt  $Y$  um 7.2%.  
 26e: Wenn  $X_2$  um 100% steigt, steigt  $Y$  um 7.2%.

## 4 Klausur BW24.1, Februar 2018

Ihre Version ist . Bitte vergessen Sie nicht, die Version in den Lösungsbogen zu übertragen. Lösen Sie die Aufgaben zunächst hier (und ggf. auf dem Schmierpapier) und übertragen Sie die Lösungen zum Schluss in den Lösungsbogen. Geben Sie nur den Lösungsbogen ab, und behalten Sie dieses Aufgabenblatt. Wir wünschen Ihnen viel Erfolg!

**Aufgabe 1:** Ihre Stichprobe der Zufallsvariablen  $X$  enthält  $n$  unabhängige und normalverteilte Beobachtungen:  $X_1, \dots, X_n$ . Welche Schätzfunktionen für  $E(X)$  sind erwartungstreu?

- 1: 

<sup>a</sup> $X_1 + X_n$	<sup>b</sup> $X_1 - X_2 + X_3$	<sup>c</sup> $\frac{1}{n} \sum_{i=1}^n X_i$	<sup>d</sup> $(X_1 + X_n)/2$	<sup>e</sup> $\frac{X_1 + X_n}{n}$
--------------------------	--------------------------------	---	------------------------------	------------------------------------

**Aufgabe 2:** Welche Schätzfunktionen für  $E(X)$  sind effizienter als  $X_2$ ?

- 2: 

<sup>a</sup> $X_1 + X_n$	<sup>b</sup> $(X_1 + X_n)/2$	<sup>c</sup> $\frac{X_1 + X_n}{n}$	<sup>d</sup> $X_1 - X_2 + X_3$	<sup>e</sup> $\frac{1}{n} \sum_{i=1}^n X_i$
--------------------------	------------------------------	------------------------------------	--------------------------------	---

**Aufgabe 3:** Die Zufallsvariable  $X$  hat eine Varianz von 25. Sie haben vor, eine Stichprobe von  $n$  unabhängigen und identisch verteilten Beobachtungen zu ziehen. Wie groß muß  $n$  sein, damit die Standardabweichung des Stichprobenmittelwerts  $\sigma_{\bar{X}}$  ist?

- 3: 

<sup>a</sup> anderer Wert	<sup>b</sup> $25/\sigma_{\bar{X}}^2$	<sup>c</sup> $\sigma_{\bar{X}}^2/5$	<sup>d</sup> $5/\sigma_{\bar{X}}$	<sup>e</sup> $\sigma_{\bar{X}}^2/25$
---------------------------	--------------------------------------	-------------------------------------	-----------------------------------	--------------------------------------

**Aufgabe 4:**

Eine Zufallsvariable  $X$  ist Poissonverteilt:  $P_\lambda(X = k) = \lambda^k \cdot e^{-\lambda}/k!$ . Ihre Stichprobe enthält drei Beobachtungen:  $\{1, 2, 3\}$ .

Was ist der Maximum-Likelihood Schätzer für  $\lambda$ ?

- 4: 

<sup>a</sup> anderer Wert	<sup>b</sup> 1	<sup>c</sup> 1/2	<sup>d</sup> 4/3	<sup>e</sup> 2
---------------------------	----------------	------------------	------------------	----------------

**Aufgabe 5:** Eine Zufallsvariable  $X$  folgt einer Verteilung  $\mathcal{X}_\theta$  mit Erwartungswert  $E(X) = 1 - \theta$ . Die Variable  $x$  enthält Ihre Stichprobe. Wie berechnen Sie mit R den Momentenschätzer für  $\theta$  auf Basis des ersten Moments?

- 5: 

<sup>a</sup> anderer Wert	<sup>b</sup> $1/(1-\text{mean}(x))$	<sup>c</sup> $1-\text{mean}(x)$	<sup>d</sup> $1-\text{var}(x)$	<sup>e</sup> $(1-\text{mean}(x))^2$
---------------------------	-------------------------------------	---------------------------------	--------------------------------	-------------------------------------

**Aufgabe 6:** Eine Zufallsvariable  $X$  ist normalverteilt mit unbekanntem Mittelwert und unbekannter Standardabweichung. Die Variable  $x$  enthält Ihre Stichprobe mit 12 Beobachtungen. Wie ermitteln Sie die Obergrenze des 95%-Konfidenzintervalls für den Mittelwert?

- 6a: Keine der folgenden Antworten ist richtig.  
 6b:  $\text{mean}(x) + \text{qt}(.95, \text{df}=11) * \text{sd}(x) / \text{sqrt}(12)$   
 6c:  $\text{mean}(x) - \text{qt}(.975, \text{df}=11) * \text{var}(x)$   
 6d:  $\text{mean}(x) - \text{qt}(.025, \text{df}=11) * \text{sd}(x) / \text{sqrt}(12)$   
 6e:  $\text{mean}(x) - \text{qt}(.975, \text{df}=11) * \text{sd}(x) * \text{sqrt}(12)$

**Aufgabe 7:** Eine Zufallsvariable  $X$  ist normalverteilt mit unbekanntem Mittelwert und unbekannter Standardabweichung. Ihre Stichprobe enthält 16 Beobachtungen. Der Mittelwert Ihrer Stichprobe ist 10, die Varianz in Ihrer Stichprobe ist 36. Wie breit ist das 95%-Konfidenzintervall für den Mittelwert?

- 7a: Keine der folgenden Antworten ist richtig.  
 7b:  $3 * \text{qt}(.975, \text{df}=15)$   
 7c:  $3/2 * \text{qt}(.975, \text{df}=15)$   
 7d:  $10 - 3/2 * \text{qt}(0.025, \text{df}=15)$   
 7e:  $-\text{qt}(.025, \text{df}=15)/3$

**Aufgabe 8:** Eine Zufallsvariable  $X$  ist normalverteilt mit unbekanntem Mittelwert  $\mu$  und unbekannter Varianz. Eine Stichprobe von 9 Beobachtungen ergibt einen Stichprobenmittelwert  $\bar{x} = 3$  und eine Stichprobenvarianz  $\text{var}(x) = 81$ . Ihre Alternativhypothese ist  $\mu < 5$ . Wie berechnen Sie einen  $p$ -Wert zum Test der Nullhypothese?

- 8: 

<sup>a</sup> anderer Wert	<sup>b</sup> $\text{pt}(2/3, \text{df}=8)$	<sup>c</sup> $\text{pt}(-2/3, \text{df}=8)$	<sup>d</sup> $\text{pt}(-2/9, \text{df}=8)$	<sup>e</sup> $\text{pt}(-2/9, \text{df}=8)$
---------------------------	--	---	---	---

**Aufgabe 9:** Eine Zufallsvariable  $X$  ist normalverteilt mit unbekanntem Mittelwert  $\mu$  und bekannter Varianz  $\sigma^2 = 9$ . Eine Stichprobe von 9 Beobachtungen ergibt einen Stichprobenmittelwert  $\bar{x} = 2$ . Ihre Alternativhypothese ist  $\mu \neq -1$ . Wie berechnen Sie einen  $p$ -Wert zum Test der Nullhypothese?

- 9: 

<sup>a</sup> anderer Wert	<sup>b</sup> $\text{pnorm}(3) * 2$	<sup>c</sup> $\text{pnorm}(3)$	<sup>d</sup> $(1 - \text{pnorm}(3)) * 2$	<sup>e</sup> $1 - \text{pnorm}(-3)$
---------------------------	------------------------------------	--------------------------------	--	-------------------------------------

**Aufgabe 10:** Welche Wahrscheinlichkeit gibt der  $p$ -Wert an?

- 10a: Keine der folgenden Antworten ist richtig.  
 10b: Die Wahrscheinlichkeit, eine Stichprobe zu ziehen, die mindestens so advers zur Nullhypothese ist, wie die gezogene Stichprobe, falls die Nullhypothese wahr ist.  
 10c: Die Wahrscheinlichkeit, eine Stichprobe zu ziehen, die mindestens so advers zur Nullhypothese ist, wie die gezogene Stichprobe, falls die Nullhypothese falsch ist.  
 10d: Die Wahrscheinlichkeit eines Fehlers erster Art.  
 10e: Die Wahrscheinlichkeit, die Alternativhypothese abzulehnen, wenn sie falsch ist.

**Aufgabe 11:** Betrachten Sie das folgende Testergebnis.  $x$  und  $y$  sind jeweils Stichproben von  $X$  bzw.  $Y$ . Verwenden Sie ein Signifikanzniveau von 5%.

Welch Two Sample t-test  
 data: x and y  
 t = 5.2674, df = 31.884, p-value = 9.219e-06  
 alternative hypothesis: true difference in means is not equal to 0  
 95 percent confidence interval:  
 0.9924276 2.2442334  
 sample estimates:  
 mean of x mean of y  
 0.252889 -1.365442

Was können Sie aus dem Testergebnis schließen?

- 11a: Es wurde ein paarweiser Test durchgeführt.
- 11b: Die Alternativhypothese in diesem Test war  $E(X) - E(Y) > 0$ .
- 11c: Die Nullhypothese  $E(X) - E(Y) = 0.5$  wird abgelehnt.
- 11d: Die Nullhypothese  $E(X) - E(Y) = 1$  wird abgelehnt.
- 11e: Die Nullhypothese  $E(X) - E(Y) = 2$  wird nicht abgelehnt.

**Aufgabe 12:** Betrachten Sie das folgende Testergebnis. Verwenden Sie ein Signifikanzniveau von 5%.

Wilcoxon signed rank test  
 data: x and y  
 V = 161, p-value = 0.000164  
 alternative hypothesis: true location shift is greater than 0

Was können Sie aus dem Testergebnis schließen?

- 12a: Keine der folgenden Antworten ist richtig.
- 12b: Die Alternativhypothese in diesem Test war  $E(X) > E(Y)$ .
- 12c: Die Alternativhypothese in diesem Test war  $E(X) \neq E(Y)$ .
- 12d: Die Nullhypothese  $E(X) = E(Y)$  wird abgelehnt.
- 12e: Die Nullhypothese  $E(X) - E(Y) = 1$  wird abgelehnt.

**Aufgabe 13:**

Zwei Merkmale, X und Y, können jeweils die Werte 1, 2, 3, 4, bzw. 1 und 2 annehmen. Die Häufigkeiten in Ihrer Stichprobe sind durch die folgende Tabelle gegeben.

	X = 1	X = 2	X = 3	X = 4
Y = 1	13	14	15	14
Y = 2	20	4	11	15

Interpretieren Sie das folgende Testergebnis:

Pearson's Chi-squared test  
 data: Z  
 X-squared = 7.3743, df = 3, p-value = 0.06088

Was können Sie aus dem Testergebnis schließen?

- 13a: Für ein Signifikanzniveau von 10% findet man zwischen den beiden Merkmalen X und Y keinen signifikanten Zusammenhang.
- 13b: Für ein Signifikanzniveau von 10% kann man die Nullhypothese, X und Y seien unabhängig, verwerfen.
- 13c: Das Testergebnis zeigt, dass mit großen Werten von Y auch große Werte von X wahrscheinlicher sind.
- 13d: Bei einem Signifikanzniveau von 5% kann man die Nullhypothese, X und Y seien voneinander abhängig, verwerfen.
- 13e: Es wurde ein zweiseitiger Test durchgeführt.

**Aufgabe 14:** In der folgenden Regression schätzen Sie den Effekt der Variablen a, b und c auf die Variable Y. Sie erhalten den folgenden Output:

Call: lm(formula = y ~ a \* b + c)  
 Estimate Std. Error t value Pr(>|t|)  
 (Intercept) -1.00 0.79 -1.26 0.2098  
 a 2.00 1.00 2.00 0.0491  
 b 3.00 1.00 3.00 0.0037  
 c 3.00 0.71 4.24 0.0001  
 a:b 2.00 1.41 1.41 0.1614  
 Residual standard error: 1.033 on 75 degrees of freedom  
 Multiple R-squared: 0.8974, Adjusted R-squared: 0.892  
 F-statistic: 164.1 on 4 and 75 DF, p-value: < 2.2e-16

Welchen Wert für Y erwarten Sie, wenn  $a = b = c = 0$ ?

14:	a anderer Wert	b -4	c -1	d 0	e 1
-----	----------------	------	------	-----	-----

**Aufgabe 15:** Welchen Wert für Y erwarten Sie, wenn  $a = 1$  und  $b = c = 0$ ?

15:	a anderer Wert	b -4	c -1	d 0	e 1
-----	----------------	------	------	-----	-----

**Aufgabe 16:** Welchen Wert für Y erwarten Sie, wenn  $a = 1$ ,  $b = -1$  und  $c = 0$ ?

16:	a anderer Wert	b -4	c -1	d 0	e 1
-----	----------------	------	------	-----	-----

**Aufgabe 17:** Wie groß ist der marginale Effekt von b auf Y falls  $a = 2$ ?

17:	a anderer Wert	b 2	c 3 + c	d 5	e 7
-----	----------------	-----	---------	-----	-----

**Aufgabe 18:** Wie groß ist der marginale Effekt von a auf Y falls  $b = -1$ ?

18:	a anderer Wert	b -1	c 0	d 1	e 2 + c
-----	----------------	------	-----	-----	---------

**Aufgabe 19:** Wie bestimmen Sie für die obige Regression die Obergrenze des 95%-Konfidenzintervalls für den Koeffizienten von b?

- 19a: Keine der folgenden Antworten ist richtig.
- 19b:  $3 - qt(.025, df=75)$
- 19c:  $3 - qt(.05, df=75)$
- 19d:  $3 - qt(.95, df=75)$
- 19e:  $3 + qt(.025, df=75)$

**Aufgabe 20:** Wie breit ist für die obige Regression das 95%-Konfidenzintervall für den Koeffizienten von a?

- 20a: Keine der folgenden Antworten ist richtig.
- 20b:  $2 * qt(.975, df=75)$
- 20c:  $qt(.95, df=75)$
- 20d:  $-2 * qt(.975, df=75)$
- 20e:  $qt(.025, df=75)$

**Aufgabe 21:** Betrachten Sie weiter die obige Regression. Gehen Sie von einem Signifikanzniveau von 5% aus. Nehmen Sie ferner an, dass die Residuen normalverteilt sind:

- 21a: b hat einen signifikanten Einfluss auf c.
- 21b: Falls  $a = 0$  hat b einen signifikanten Einfluss auf Y.
- 21c: Falls  $b = 0$  hat a einen signifikanten Einfluss auf Y.
- 21d: Die Nullhypothese, der Koeffizient von a sei -1, wird abgelehnt.
- 21e: Das Modell erklärt 89.74% der Varianz von Y.

**Aufgabe 22:** Betrachten Sie weiter die obige Regression.  $b$  ist eine Dummy Variable und hat den Wert 1 im Fall A und den Wert 0 sonst. Nun codieren Sie  $b$  anders.  $b$  hat nun den Wert 0 im Fall A und den Wert 1 sonst. Was ist nun der marginale Effekt von  $b$  auf  $Y$  falls  $a = 0$ ?

22:	<sup>a</sup> anderer Wert	<sup>b</sup> -1	<sup>c</sup> 5	<sup>d</sup> -3	<sup>e</sup> 3
-----	---------------------------	-----------------	----------------	-----------------	----------------

**Aufgabe 23:** Was ist (in der Regression mit der neuen Codierung von  $b$ ) der geschätzte Koeffizient von  $a$ ?

23:	<sup>a</sup> anderer Wert	<sup>b</sup> -2	<sup>c</sup> -1	<sup>d</sup> 2	<sup>e</sup> 4
-----	---------------------------	-----------------	-----------------	----------------	----------------

**Aufgabe 24:**

In der Gleichung  $Y = \beta_0 + \beta_1 \log X + u$  schätzen Sie  $\hat{\beta}_1 = 2.4$ . Was erwarten Sie?

- 24a: Keine der folgenden Antworten ist richtig.
- 24b: Wenn  $X$  um 1% steigt, steigt  $Y$  um 2.4.
- 24c: Wenn  $X$  um 1% steigt, steigt  $Y$  um 0.024.
- 24d: Wenn  $X$  um 1 steigt, steigt  $Y$  um 2.4%.

24e: Wenn  $X$  um 1 steigt, steigt  $Y$  um 240%.

**Aufgabe 25:**

In der Gleichung  $\log Y = \beta_0 + \beta_1 X + u$  schätzen Sie  $\hat{\beta}_1 = 0.12$ . Was erwarten Sie?

- 25a: Keine der folgenden Antworten ist richtig.
- 25b: Wenn  $X$  um 1% steigt, steigt  $Y$  um 0.12.
- 25c: Wenn  $X$  um 1% steigt, steigt  $Y$  um 0.0012.
- 25d: Wenn  $X$  um 1 steigt, steigt  $Y$  um 0.12%.
- 25e: Wenn  $X$  um 0.01 steigt, steigt  $Y$  um 0.12%.

**Aufgabe 26:**

In der Gleichung  $\log Y = \beta_0 + \beta_1 \log X + u$  schätzen Sie  $\hat{\beta}_1 = 21.4$ . Was erwarten Sie?

- 26a: Keine der folgenden Antworten ist richtig.
- 26b: Wenn  $X$  um 1% steigt, steigt  $Y$  um 21.4%.
- 26c: Wenn  $X$  um 0.01 steigt, steigt  $Y$  um 21.4%.
- 26d: Wenn  $X$  um 1 steigt, steigt  $Y$  um 21.4%.
- 26e: Wenn  $X$  um 100% steigt, steigt  $Y$  um 21.4%.

## 5 Klausur BW24.1, Mai 2017

Ihre Version ist . Bitte vergessen Sie nicht, die Version in den Lösungsbogen zu übertragen. Lösen Sie die Aufgaben zunächst hier (und ggf. auf dem Schmierpapier) und übertragen Sie die Lösungen zum Schluss in den Lösungsbogen. Geben Sie nur den Lösungsbogen ab, und behalten Sie dieses Aufgabenblatt. Wir wünschen Ihnen viel Erfolg!

**Aufgabe 1:** Ihre Stichprobe der Zufallsvariablen  $X$  enthält 7 unabhängige und normalverteilte Beobachtungen:  $X_1, \dots, X_7$ . Welche Schätzfunktionen für  $E(X)$  sind erwartungstreu?

1:	<sup>a</sup> $\frac{X_7 + X_6}{2}$	<sup>b</sup> $\frac{X_7 + X_5}{3}$	<sup>c</sup> $\frac{2X_7 + X_3}{3}$	<sup>d</sup> $\frac{X_7 + X_6 + X_5}{X_7 + \frac{X_6 - X_5}{2}}$
----	------------------------------------	------------------------------------	-------------------------------------	--

**Aufgabe 2:** Betrachten Sie weiter die obige Stichprobe. Die Varianz von  $X$  sei 3. Wie groß ist die Varianz von  $X_1 + X_2 - 2X_3$ ?

2:	<sup>a</sup> anderer Wert	<sup>b</sup> 0	<sup>c</sup> 12	<sup>d</sup> 18	<sup>e</sup> 24
----	---------------------------	----------------	-----------------	-----------------	-----------------

**Aufgabe 3:** Welche Schätzfunktionen für  $E(X)$  sind weniger effizient als  $X_1 + X_2 - X_3$ ?

3:	<sup>a</sup> $X_1 + X_2$	<sup>b</sup> $3X_1$	<sup>c</sup> $3X_1 + 3X_2 - 5X_3$	<sup>d</sup> $\frac{1}{7} \sum_{i=1}^7 X_i$	<sup>e</sup> $3X_1 - 2X_2$
----	--------------------------	---------------------	-----------------------------------	---	----------------------------

**Aufgabe 4:** Die Zufallsvariable  $X$  hat eine Standardabweichung von 18. Sie haben vor, eine Stichprobe von  $n$  unabhängigen und identisch verteilten Beobachtungen zu ziehen. Wie groß muß  $n$  sein, damit die Standardabweichung des Stichprobenmittelwerts  $\sigma_{\bar{X}} = 2$  ist?

4:	<sup>a</sup> anderer Wert	<sup>b</sup> 3	<sup>c</sup> 9	<sup>d</sup> 36	<sup>e</sup> 81
----	---------------------------	----------------	----------------	-----------------	-----------------

**Aufgabe 5:**

Die Dichtefunktion von  $X$  ist

$$f(X|\theta) = \begin{cases} \frac{20-2X}{(\theta-10)^2} & \text{falls } X \in [\theta, 10] \\ 0 & \text{sonst} \end{cases}$$

wobei  $\theta < 10$ . Eine Stichprobe ergibt zwei Beobachtungen:  $\{3, 2\}$ .

Was ist der Maximum-Likelihood Schätzer für  $\theta$ ?

5:	<sup>a</sup> anderer Wert	<sup>b</sup> 0	<sup>c</sup> 2	<sup>d</sup> 3	<sup>e</sup> 10
----	---------------------------	----------------	----------------	----------------	-----------------

**Aufgabe 6:** Die Zufallsvariable  $X$  folgt einer Verteilung  $\mathcal{X}_\theta$  mit Erwartungswert  $E(X) = \theta^2$  und Varianz  $1/\theta^2$  mit  $\theta > 0$ .

Die Variable  $x$  enthält Ihre Stichprobe. Wie berechnen Sie mit  $R$  den Momentenschätzer für  $\theta$  auf Basis des zweiten Moments?

6:	<sup>a</sup> anderer Wert	<sup>b</sup> $1/\text{sd}(x)$	<sup>c</sup> $1/\text{var}(x)$	<sup>d</sup> $\text{mean}(x)$	<sup>e</sup> $\text{mean}(x)^{-2}$
----	---------------------------	-------------------------------	--------------------------------	-------------------------------	------------------------------------

**Aufgabe 7:** Der Vektor  $x$  enthält Ihre Stichprobe mit 16 Beobachtungen,  $s$  ist die geschätzte Standardabweichung von  $x$ . Sie gehen davon aus, dass  $x$  einer Normalverteilung folgt. Wie berechnen Sie die Breite des 90%-Konfidenzintervalls für den Mittelwert von  $x$ ?

- 7a: Keine der folgenden Antworten ist richtig.
- 7b:  $s/2 * \text{qt}(.90, \text{df}=15)$
- 7c:  $s/4 * \text{qt}(.90, \text{df}=15)$
- 7d:  $s/2 * \text{qt}(.95, \text{df}=15)$
- 7e:  $s/4 * \text{qt}(.95, \text{df}=15)$

**Aufgabe 8:** Sie betrachten weiter die obige Stichprobe. Ihre Nullhypothese ist  $E(x) \leq 8$ , Ihre Alternativhypothese ist  $E(x) > 8$ . Ihre Teststatistik haben Sie mit dem Kommando

$$t \leftarrow (\text{mean}(x) - 8) / (s/4)$$

berechnet, wobei  $\text{mean}(x)$  den Mittelwert Ihrer Stichprobe enthält. In welchem Bereich lehnen Sie die Nullhypothese bei einem Signifikanzniveau von 5% ab?

- 8a: Keine der folgenden Antworten ist richtig.
- 8b:  $t < \text{qt}(.05, \text{df}=15)$
- 8c:  $t < -\text{qt}(.05, \text{df}=15)$
- 8d:  $t > \text{qt}(.025, \text{df}=15)$
- 8e:  $t > -\text{qt}(.25, \text{df}=15)$

**Aufgabe 9:** Sie testen weiter die Nullhypothese aus der obigen Aufgabe. Wie bestimmen Sie den  $p$ -Wert?

- 9a: Keine der folgenden Antworten ist richtig.
- 9b:  $\text{pt}(t, \text{df}=15)$
- 9c:  $\text{pt}(-t, \text{df}=15)$
- 9d:  $2 * \text{pt}(t, \text{df}=15)$
- 9e:  $2 * \text{pt}(-t, \text{df}=15)$

**Aufgabe 10:** Welche Wahrscheinlichkeit gibt der  $p$ -Wert an?

- 10a: Keine der folgenden Antworten ist richtig.
- 10b: Die Wahrscheinlichkeit, eine Stichprobe zu ziehen, die mindestens so advers zur Nullhypothese ist, wie die gezogene Stichprobe, falls die Nullhypothese wahr ist.
- 10c: Die Wahrscheinlichkeit, eine Stichprobe zu ziehen, die mindestens so advers zur Nullhypothese ist, wie die gezogene Stichprobe, falls die Nullhypothese falsch ist.
- 10d: Die Wahrscheinlichkeit, eine Stichprobe zu ziehen, die mindestens so advers zur Nullhypothese ist, wie die gezogene Stichprobe.
- 10e: Die Wahrscheinlichkeit, die Nullhypothese abzulehnen, wenn sie wahr ist.

**Aufgabe 11:** Betrachten Sie das folgende Testergebnis.  $x$  und  $y$  sind jeweils Stichproben von  $X$  bzw.  $Y$ . Verwenden Sie ein Signifikanzniveau von 5%.

```
Paired t-test
data: x and y
t = -5.5664, df = 11, p-value = 0.0001686
alternative hypothesis: true difference in means
is not equal to 0

95 percent confidence interval:
-2.502550 -1.084296
sample estimates:
mean of the differences
-1.793423
```

Was können Sie aus dem Testergebnis schließen?

- 11a: Es wurde ein paarweiser Test durchgeführt.
- 11b: Die Alternativhypothese in diesem Test war  $E(X) - E(Y) > 0$ .
- 11c: Die Alternativhypothese in diesem Test war  $E(X) \neq E(Y)$ .
- 11d: Die Nullhypothese  $E(X - Y) = -2$  kann abgelehnt werden.
- 11e: Die Nullhypothese  $E(X - Y) = -4$  kann abgelehnt werden.

**Aufgabe 12:** Betrachten Sie das folgende Testergebnis. Verwenden Sie ein Signifikanzniveau von 5%.

```
Wilcoxon signed rank test
data: x and y
V = 1, p-value = 0.0009766
alternative hypothesis: true location shift is not equal to 0
```

Was können Sie aus dem Testergebnis schließen?

- 12a: Es wurde ein paarweiser Test durchgeführt.
- 12b: Die Alternativhypothese in diesem Test war  $E(X) > E(Y)$ .
- 12c: Die Alternativhypothese in diesem Test war  $E(X) \neq E(Y)$ .
- 12d: Die Nullhypothese  $E(X) = E(Y)$  kann abgelehnt werden.
- 12e: Die Nullhypothese  $E(X) - E(Y) = 1$  kann abgelehnt werden.

**Aufgabe 13:**

Zwei Merkmale,  $X$  und  $Y$ , können jeweils die Werte  $A, B, C$  bzw.  $D$  und  $E$  annehmen. Die Häufigkeiten sind durch die folgende Tabelle gegeben.

	$X = A$	$X = B$	$X = C$
$Y = D$	23	11	22
$Y = E$	12	13	23

Ihr Signifikanzniveau ist 5%. Interpretieren Sie das folgende Testergebnis:

Pearson's Chi-squared test

```
data: Z
X-squared = 3.0487, df = 2, p-value = 0.2178
```

Was können Sie aus dem Testergebnis schließen?

- 13a: Zwischen den beiden Merkmalen  $X$  und  $Y$  finden Sie keinen signifikanten Zusammenhang.
- 13b: Die Nullhypothese,  $X$  und  $Y$  seien unabhängig, kann man zum gegebenen Signifikanzniveau verwerfen.
- 13c: Das nicht-signifikante Testergebnis zeigt, dass mit großen Werten von  $Y$  kleine Werte von  $X$  wahrscheinlicher sind.
- 13d: Die Nullhypothese,  $X$  und  $Y$  seien voneinander abhängig, kann man zum gegebenen Signifikanzniveau verwerfen.
- 13e: Um zu testen, ob  $X$  und  $Y$  voneinander unabhängig sind, sollte man besser einen Rangsummentest verwenden.

**Aufgabe 14:** In der folgenden Regression schätzen Sie den Effekt der Variablen  $a$  und  $b$  auf die Variable  $Y$ . Sie erhalten das folgende Ergebnis:

```
Call: lm(formula = y ~ a * b)
      Estimate Std. Error t value Pr(>|t|)
(Intercept) -3.00      0.71   -4.24  0.0028
a            -3.00      1.00   -3.00  0.0171
b             2.00      1.00    2.00  0.0805
a:b           1.00      1.41    0.71  0.4996

Residual standard error: 1.225 on 8 degrees of freedom
Multiple R-squared: 0.7612, Adjusted R-squared: 0.6716
F-statistic: 8.5 on 3 and 8 DF, p-value: 0.007196
```

Welchen Wert für  $Y$  erwarten Sie, wenn  $a = 0$  und  $b = 0$ ?

14:	<sup>a</sup> anderer Wert	<sup>b</sup> -9	<sup>c</sup> -6	<sup>d</sup> -3	<sup>e</sup> 0
-----	---------------------------	-----------------	-----------------	-----------------	----------------

**Aufgabe 15:** Welchen Wert für  $Y$  erwarten Sie, wenn  $a = 1$  und  $b = 0$ ?

15:	<sup>a</sup> anderer Wert	<sup>b</sup> -9	<sup>c</sup> -6	<sup>d</sup> -3	<sup>e</sup> 0
-----	---------------------------	-----------------	-----------------	-----------------	----------------

**Aufgabe 16:** Welchen Wert für  $Y$  erwarten Sie, wenn  $a = 1$  und  $b = -1$ ?

16:	<sup>a</sup> anderer Wert	<sup>b</sup> -9	<sup>c</sup> -6	<sup>d</sup> -3	<sup>e</sup> 0
-----	---------------------------	-----------------	-----------------	-----------------	----------------

**Aufgabe 17:** Wie groß ist der geschätzte marginale Effekt von  $b$  auf  $Y$  falls  $a = -1$ ?

17:	<sup>a</sup> anderer Wert	<sup>b</sup> 1	<sup>c</sup> 2	<sup>d</sup> 3	<sup>e</sup> 4
-----	---------------------------	----------------	----------------	----------------	----------------

**Aufgabe 18:** Wie groß ist der geschätzte marginale Effekt von  $a$  auf  $Y$  falls  $b = -1$ ?

18:	<sup>a</sup> anderer Wert	<sup>b</sup> -4	<sup>c</sup> -3	<sup>d</sup> -2	<sup>e</sup> -1
-----	---------------------------	-----------------	-----------------	-----------------	-----------------

**Aufgabe 19:** Wie bestimmen Sie für die obige Regression die Obergrenze des 95%-Konfidenzintervalls für den Koeffizienten von  $a$ ?

- 19a: Keine der folgenden Antworten ist richtig.  
 19b:  $-3 + qt(.975, df=8)$   
 19c:  $-3 + qt(.95, df=8)$   
 19d:  $-3 + qt(.95, df=8)$   
 19e:  $-3 + qt(.975, df=8)$

**Aufgabe 20:** Sie verwenden das folgende quadratische Modell, um den Zusammenhang zwischen der erklärenden Variablen  $X$  und der abhängigen Variablen  $Y$  zu schätzen:

$$Y = \beta_0 + \beta_1 X + \beta_2 X^2 + u$$

Sie erwarten folgendes: Für  $X = 1$  ist der marginale Effekt von  $X$  auf  $Y$  gleich 0. Für  $X = 2$  ist der marginale Effekt von  $X$  auf  $Y$  gleich 2.

Welchen Wert erwarten Sie für  $\beta_1$ ?

20:	<sup>a</sup> anderer Wert	<sup>b</sup> -2	<sup>c</sup> 0	<sup>d</sup> 2	<sup>e</sup> 4
-----	---------------------------	-----------------	----------------	----------------	----------------

**Aufgabe 21:** Welchen Wert erwarten Sie für  $\beta_2$ ?

21:	<sup>a</sup> anderer Wert	<sup>b</sup> -1	<sup>c</sup> 0	<sup>d</sup> 1	<sup>e</sup> 2
-----	---------------------------	-----------------	----------------	----------------	----------------

**Aufgabe 22:** Sie schätzen den Zusammenhang zwischen der erklärenden Variablen  $X$  und der abhängigen Variablen  $Y$  mit einem Polynom  $Y = \sum_{k=0}^r \beta_k X^k + u$ . Für verschiedene Werte von  $r$  erhalten Sie die folgenden Werte für das AIC:

	1	2	3	4	5	6
AIC	104.98	105.04	107.01	107.37	105.82	107.77

Welchen Wert von  $r$  sollten Sie auf Basis des AIC wählen?

22:	<sup>a</sup> anderer Wert	<sup>b</sup> 1	<sup>c</sup> 2	<sup>d</sup> 3	<sup>e</sup> 5
-----	---------------------------	----------------	----------------	----------------	----------------

**Aufgabe 23:** Sie schätzen den Zusammenhang zwischen  $X$  und  $Y$  mit Hilfe unterschiedlicher Modelle und erhalten die folgenden Ergebnisse:

	$\hat{\beta}_0$	$\hat{\beta}_1$
$Y = \beta_0 + \beta_1 \cdot X + u$	6.20	2.10
$Y = \beta_0 + \beta_1 \cdot \log X + u$	7.73	0.43
$\log Y = \beta_0 + \beta_1 \cdot X + u$	1.83	0.29
$\log Y = \beta_0 + \beta_1 \cdot \log X + u$	2.04	0.06

Wie groß schätzen Sie die marginale Änderung von  $Y$  bei einer Änderung von  $X$  um 1 Prozent?

23:	<sup>a</sup> anderer Wert	<sup>b</sup> 2.10	<sup>c</sup> 0.43	<sup>d</sup> 0.29	<sup>e</sup> 0.06
-----	---------------------------	-------------------	-------------------	-------------------	-------------------

**Aufgabe 24:** Wie groß schätzen Sie die prozentuale Änderung von  $Y$  bei einer Änderung von  $X$  um 1 Prozent?

24:	<sup>a</sup> anderer Wert	<sup>b</sup> 2.10%	<sup>c</sup> 0.43%	<sup>d</sup> 0.29%	<sup>e</sup> 0.06%
-----	---------------------------	--------------------	--------------------	--------------------	--------------------

**Aufgabe 25:** Wie groß schätzen Sie die prozentuale Änderung von  $Y$  bei einer Änderung von  $X$  um eine Einheit?

25:	<sup>a</sup> anderer Wert	<sup>b</sup> 210%	<sup>c</sup> 43%	<sup>d</sup> 29%	<sup>e</sup> 6%
-----	---------------------------	-------------------	------------------	------------------	-----------------

**Aufgabe 26:** Wie groß schätzen Sie die marginale Änderung von  $Y$  bei einer Änderung von  $X$  um eine Einheit?

26:	<sup>a</sup> anderer Wert	<sup>b</sup> 2.10	<sup>c</sup> 0.43	<sup>d</sup> 0.29	<sup>e</sup> 0.06
-----	---------------------------	-------------------	-------------------	-------------------	-------------------

## 6 Klausur BW24.1, Februar 2017

Ihre Version ist . Bitte vergessen Sie nicht, die Version in den Lösungsbogen zu übertragen. Lösen Sie die Aufgaben zunächst hier (und ggf. auf dem Schmierpapier) und übertragen Sie die Lösungen zum Schluss in den Lösungsbogen. Geben Sie nur den Lösungsbogen ab, und behalten Sie dieses Aufgabenblatt. Wir wünschen Ihnen viel Erfolg!

**Aufgabe 1:** Ihre Stichprobe der Zufallsvariablen  $X$  enthält 5 unabhängige und normalverteilte Beobachtungen:  $X_1, \dots, X_5$ . Welche Schätzfunktionen für  $E(X)$  sind erwartungstreu?

1:	<sup>a</sup> $X_3 - X_1$	<sup>b</sup> $(X_2 + X_1)/2$	<sup>c</sup> $2X_4 - X_5$	<sup>d</sup> $X_4 - X_2 + X_1$	<sup>e</sup> $X_3$
----	--------------------------	------------------------------	---------------------------	--------------------------------	--------------------

**Aufgabe 2:** Welche Schätzfunktionen für  $E(X)$  sind weniger effizient als  $X_5$ ?

2:	<sup>a</sup> $X_3 - X_1$	<sup>b</sup> $(X_2 + X_1)/2$	<sup>c</sup> $X_4 - X_2 + X_1$	<sup>d</sup> $2X_4 - X_5$	<sup>e</sup> $(X_3 + X_2 + X_1)/3$
----	--------------------------	------------------------------	--------------------------------	---------------------------	------------------------------------

**Aufgabe 3:** Die Zufallsvariable  $X$  hat eine Standardabweichung von  $\sigma_X$ . Sie haben vor, eine Stichprobe von  $n$  unabhängigen und identisch verteilten Beobachtungen zu ziehen. Wie groß muß  $n$  sein, damit die Standardabweichung des Stichprobenmittelwerts  $\sigma_{\bar{X}} = 4$  ist?

3:	<sup>a</sup> anderer Wert	<sup>b</sup> $\sigma_X^2/16$	<sup>c</sup> $16/\sigma_X^2$	<sup>d</sup> $\sigma_X^2/4$	<sup>e</sup> $4 \cdot \sigma_X$
----	---------------------------	------------------------------	------------------------------	-----------------------------	---------------------------------

**Aufgabe 4:**

Eine Zufallsvariable  $X$  ist wie folgt verteilt:  $P(X = 1) = 2\theta$ ,  $P(X = 2) = \theta$ ,  $P(X = 3) = 1 - 3\theta$  wobei  $\theta \in [0, 1/3]$ . Eine Stichprobe ergibt zwei Beobachtungen:  $\{3, 2\}$ .

Was ist der Maximum-Likelihood Schätzer für  $\theta$ ?

4:	<sup>a</sup> anderer Wert	<sup>b</sup> 0	<sup>c</sup> 1/6	<sup>d</sup> 1/3	<sup>e</sup> 1
----	---------------------------	----------------	------------------	------------------	----------------

**Aufgabe 5:** Die Zufallsvariable  $X$  folgt einer Verteilung  $\mathcal{X}_\theta$  mit Erwartungswert  $E(X) = 2 + \theta$  und Varianz  $\theta$  mit  $\theta > 0$ . Die Variable  $x$  enthält Ihre Stichprobe. Wie berechnen Sie mit  $R$  den Momentenschätzer für  $\theta$  auf Basis des ersten Moments?

5:	<sup>a</sup> anderer Wert	<sup>b</sup> $\text{var}(x)$	<sup>c</sup> $1/\text{var}(x)$	<sup>d</sup> $\text{var}(x) - 2$	<sup>e</sup> $\text{mean}(x) - 2$
----	---------------------------	------------------------------	--------------------------------	----------------------------------	-----------------------------------

**Aufgabe 6:** Eine Zufallsvariable  $X$  ist normalverteilt mit unbekanntem Mittelwert und unbekannter Standardabweichung. Die Variable  $x$  enthält Ihre Stichprobe mit 25 Beobachtungen. Wie ermitteln Sie die Breite des 95%-Konfidenzintervalls für den Mittelwert?

- 6a: Keine der folgenden Antworten ist richtig.  
 6b:  $2 * \text{sd}(x) / 5 * qt(.95, df=24)$   
 6c:  $2 * \text{sd}(x) / 5 * qt(.975, df=24)$   
 6d:  $\text{mean}(x) + \text{sd}(x) * qt(.95, df=24)$   
 6e:  $\text{mean}(x) - \text{sd}(x) / 5 * qt(.95, df=24)$

**Aufgabe 7:** Die Zufallsvariable  $X$  ist normalverteilt mit unbekanntem Mittelwert  $\mu$  und Varianz  $\sigma^2 = 36$ . Eine Stichprobe von 9 Beobachtungen ergibt einen Stichprobenmittelwert  $\bar{x} = 2$ . Ihre Alternativhypothese ist  $\mu \neq 5$ . Was ist der Absolutwert der Teststatistik?

7:	a	b	c	d	e
	anderer Wert	1/4	3/4	3/2	9/2

**Aufgabe 8:** Die Zufallsvariable  $X$  ist normalverteilt mit unbekanntem Mittelwert  $\mu$  und Varianz  $\sigma^2 = 25$ . Eine Stichprobe von 16 Beobachtungen ergibt einen Stichprobenmittelwert  $\bar{x} = 11$ . Ihre Alternativhypothese ist  $\mu < 10$ . Wie bestimmen Sie den  $p$ -Wert für Ihren Hypothesentest?

- 8a: Keine der folgenden Antworten ist richtig.  
 8b: `pt(-.64, df=15)`  
 8c: `pt(-.16, df=15)`  
 8d: `pt(.8, df=15)`  
 8e: `qt(.8, df=15)`

**Aufgabe 9:** Welche Wahrscheinlichkeit gibt der  $p$ -Wert an?

- 9a: Keine der folgenden Antworten ist richtig.  
 9b: Die Wahrscheinlichkeit, eine Stichprobe zu ziehen, die mindestens so advers zur Nullhypothese ist, wie die gezogene Stichprobe, falls die Nullhypothese falsch ist.  
 9c: Die Wahrscheinlichkeit, eine Stichprobe zu ziehen, die mindestens so advers zur Nullhypothese ist, wie die gezogene Stichprobe.  
 9d: Die Wahrscheinlichkeit eines Fehlers erster Art.  
 9e: Die Wahrscheinlichkeit, die Alternativhypothese abzulehnen, wenn sie falsch ist.

**Aufgabe 10:** Betrachten Sie das folgende Testergebnis.  $x$  und  $y$  sind jeweils Stichproben von  $X$  bzw.  $Y$ . Verwenden Sie ein Signifikanzniveau von 5%.

```
Welch Two Sample t-test
data: x and y
t = 3.2625, df = 43.074, p-value = 0.001083
alternative hypothesis: true difference in means
is greater than 0
95 percent confidence interval:
0.5239417 Inf
sample estimates:
mean of x mean of y
0.1010021 -0.9798666
```

Was können Sie aus dem Testergebnis schließen?

- 10a: Es wurde ein paarweiser Test durchgeführt.  
 10b: Die Alternativhypothese in diesem Test war  $E(X) - E(Y) > 0$ .  
 10c: Die Alternativhypothese in diesem Test war  $E(X) \neq E(Y)$ .  
 10d: Die Nullhypothese  $E(X) \leq E(Y)$  wird abgelehnt.  
 10e: Die Nullhypothese  $E(X) - E(Y) \leq 1$  wird nicht abgelehnt.

**Aufgabe 11:** Betrachten Sie das folgende Testergebnis. Verwenden Sie ein Signifikanzniveau von 5%.

```
Wilcoxon rank sum test
data: x and y
W = 441, p-value = 0.000636
alternative hypothesis: true location shift is greater than 0
```

- Was können Sie aus dem Testergebnis schließen?  
 11a: Es wurde kein paarweiser Test durchgeführt.

- 11b: Die Alternativhypothese in diesem Test war  $E(X) > E(Y)$ .  
 11c: Die Alternativhypothese in diesem Test war  $E(X) \neq E(Y)$ .  
 11d: Die Nullhypothese  $E(X) = E(Y)$  wird abgelehnt.  
 11e: Die Nullhypothese  $E(X) - E(Y) = 1$  wird abgelehnt.

**Aufgabe 12:**

Zwei Merkmale,  $X$  und  $Y$ , können jeweils die Werte 1, 2, 3 bzw. 1 und 2 annehmen. Die Häufigkeiten sind durch die folgende Tabelle gegeben.

	$X = 1$	$X = 2$	$X = 3$
$Y = 1$	20	4	21
$Y = 2$	18	14	14

Ihr Signifikanzniveau ist 5%. Interpretieren Sie das folgende Testergebnis:

```
Pearson's Chi-squared test
data: Z
X-squared = 7.0507, df = 2, p-value = 0.02944
```

Was können Sie aus dem Testergebnis schließen?

- 12a: Zwischen den beiden Merkmalen  $X$  und  $Y$  finden Sie keinen signifikanten Zusammenhang.  
 12b: Die Nullhypothese,  $X$  und  $Y$  seien unabhängig, kann man zum gegebenen Signifikanzniveau verwerfen.  
 12c: Das signifikante Testergebnis zeigt, dass mit großen Werten von  $Y$  auch große Werte von  $X$  wahrscheinlicher sind.  
 12d: Die Nullhypothese,  $X$  und  $Y$  seien voneinander abhängig, kann man zum gegebenen Signifikanzniveau verwerfen.  
 12e: Um zu testen, ob  $X$  und  $Y$  voneinander unabhängig sind, sollte man besser einen Rangsummentest verwenden.

**Aufgabe 13:** In der folgenden Regression schätzen Sie den Effekt der Variablen  $a$  und  $b$  auf die Variable  $Y$ . Sie erhalten den folgenden Output:

```
Call: lm(formula = y ~ a * b)
      Estimate Std. Error t value Pr(>|t|)
(Intercept)  3.00      0.71    4.24  0.0004
a             3.00      1.00    3.00  0.0071
b             2.00      1.00    2.00  0.0593
a:b           2.00      1.41    1.41  0.1727
Residual standard error: 1.732 on 20 degrees of freedom
Multiple R-squared:  0.7222, Adjusted R-squared:  0.6806
F-statistic: 17.33 on 3 and 20 DF, p-value: 8.749e-06
```

Welchen Wert für  $Y$  erwarten Sie, wenn  $a = 0$  und  $b = 0$ ?

13:	a	b	c	d	e
	anderer Wert	0	2	3	4

**Aufgabe 14:** Welchen Wert für  $Y$  erwarten Sie, wenn  $a = 1$  und  $b = 0$ ?

14:	a	b	c	d	e
	anderer Wert	0	2	4	6

**Aufgabe 15:** Welchen Wert für  $Y$  erwarten Sie, wenn  $a = 1$  und  $b = -1$ ?

15:	a	b	c	d	e
	anderer Wert	0	2	4	6

**Aufgabe 16:** Wie groß ist der marginale Effekt von  $b$  auf  $Y$  falls  $a = 2$ ?

16:	a	b	c	d	e
	anderer Wert	2	3	5	6

**Aufgabe 17:** Wie groß ist der marginale Effekt von  $a$  auf  $Y$  falls  $b = -1$ ?

17:	<sup>a</sup> anderer Wert	<sup>b</sup> -1	<sup>c</sup> 0	<sup>d</sup> 1	<sup>e</sup> 3
-----	---------------------------	-----------------	----------------	----------------	----------------

**Aufgabe 18:** Wie bestimmen Sie für die obige Regression die Untergrenze des 95%-Konfidenzintervalls für den Koeffizienten von  $a$ ?

- 18a: Keine der folgenden Antworten ist richtig.
- 18b:  $3 - qt(.025, df=20)$
- 18c:  $3 - qt(.05, df=20)$
- 18d:  $3 - qt(.95, df=20)$
- 18e:  $3 + qt(.025, df=20)$

**Aufgabe 19:** Betrachten Sie weiter die obige Regression. Gehen Sie von einem Signifikanzniveau von 5% aus. Nehmen Sie ferner an, dass die Residuen normalverteilt sind:

- 19a:  $b$  hat einen signifikanten Einfluss auf  $a$ .
- 19b:  $b$  hat einen signifikanten Einfluss auf  $Y$ .
- 19c:  $a$  hat einen signifikanten Einfluss auf  $Y$ .
- 19d: Die Nullhypothese, der Koeffizient von  $a$  sei  $-1$ , wird abgelehnt.
- 19e: Das Modell erklärt 17.33% der Varianz von  $Y$ .

**Aufgabe 20:**

In der Gleichung  $Y = \beta_0 + \beta_1 \cdot \log X + u$ ? schätzen Sie  $\hat{\beta}_1 = 8$ .

Was erwarten Sie?

- 20a: Keine der folgenden Antworten ist richtig.
- 20b: Wenn  $X$  um 1% steigt, steigt  $Y$  um 8.
- 20c: Wenn  $X$  um 1% steigt, steigt  $Y$  um 0.08.
- 20d: Wenn  $X$  um 1 steigt, steigt  $Y$  um 8%.
- 20e: Wenn  $X$  um 1 steigt, steigt  $Y$  um 80%.

**Aufgabe 21:**

In der Gleichung  $\log Y = \beta_0 + \beta_1 X + u$ ? schätzen Sie  $\hat{\beta}_1 = 0.63$ . Was erwarten Sie?

- 21a: Keine der folgenden Antworten ist richtig.
- 21b: Wenn  $X$  um 1% steigt, steigt  $Y$  um 0.63.
- 21c: Wenn  $X$  um 1% steigt, steigt  $Y$  um 0.0063.
- 21d: Wenn  $X$  um 1 steigt, steigt  $Y$  um 0.63%.
- 21e: Wenn  $X$  um 1 steigt, steigt  $Y$  um 63%.

**Aufgabe 22:**

In der Gleichung  $\log Y = \beta_0 + \beta_1 \log X + u$ ? schätzen Sie  $\hat{\beta}_1 = 9.2$ . Was erwarten Sie?

- 22a: Keine der folgenden Antworten ist richtig.
- 22b: Wenn  $X$  um 1% steigt, steigt  $Y$  um 920%.
- 22c: Wenn  $X$  um 1% steigt, steigt  $Y$  um 9.2%.
- 22d: Wenn  $X$  um 1 steigt, steigt  $Y$  um 920%.
- 22e: Wenn  $X$  um 92% steigt, steigt  $Y$  um 1.

## 7 Klausur BW24.1, Mai 2016

Ihre Version ist . Bitte vergessen Sie nicht, die Version in den Lösungsbogen zu übertragen. Lösen Sie die Aufgaben zunächst hier (und ggf. auf dem Schmierpapier) und übertragen Sie die Lösungen zum Schluss in den Lösungsbogen. Geben Sie nur den Lösungsbogen ab, und behalten Sie dieses Aufgabenblatt. Wir wünschen Ihnen viel Erfolg!

**Aufgabe 1:** Ihre Stichprobe der Zufallsvariablen  $X$  enthält 5 unabhängige und normalverteilte Beobachtungen:  $X_1, \dots, X_5$ . Welche Schätzfunktionen für  $E(X)$  sind erwartungstreu?

1:	<sup>a</sup> $2X_1 - X_2$	<sup>b</sup> $\frac{X_1 + X_5}{2}$	<sup>c</sup> $\frac{X_1 + X_5}{3}$	<sup>d</sup> $X_4 + X_5 + X_1$	<sup>e</sup> $X_4 - X_5 + X_1$
----	---------------------------	------------------------------------	------------------------------------	--------------------------------	--------------------------------

**Aufgabe 2:** Welche Schätzfunktionen für  $E(X)$  sind effizienter als  $(X_3 + X_4)/2$ ?

2:	<sup>a</sup> $\sum_{i=1}^5 X_i$	<sup>b</sup> $\frac{X_1 + X_2 + X_3}{3}$	<sup>c</sup> $X_1 - X_2 + X_3$	<sup>d</sup> $\frac{1}{5} \sum_{i=1}^5 X_i$	<sup>e</sup> $X_3$
----	---------------------------------	--	--------------------------------	---	--------------------

**Aufgabe 3:** Die Zufallsvariable  $X$  hat eine Varianz von  $\sigma_X^2$ . Sie haben vor, eine Stichprobe von 100 unabhängigen und identisch verteilten Beobachtungen zu ziehen. Wie groß muß  $\sigma_X^2$  sein, damit die Varianz des Stichprobenmittelwerts  $\sigma_{\bar{X}}^2 = 1$  ist?

3:	<sup>a</sup> anderer Wert	<sup>b</sup> 1/100	<sup>c</sup> 1	<sup>d</sup> 10	<sup>e</sup> 100
----	---------------------------	--------------------	----------------	-----------------	------------------

**Aufgabe 4:**

Eine Zufallsvariable  $X$  folgt einer Exponentialverteilung mit Parameter  $\lambda$ . Eine Stichprobe ergibt drei Beobachtungen:  $\{2, 1, 3\}$ .

Was ist der Maximum-Likelihood Schätzer für  $\lambda$ ?

4:	<sup>a</sup> anderer Wert	<sup>b</sup> 0	<sup>c</sup> 1/3	<sup>d</sup> 1/2	<sup>e</sup> 1
----	---------------------------	----------------	------------------	------------------	----------------

**Aufgabe 5:** Was ist der Momentenschätzer für  $\lambda$  auf Basis des ersten Moments?

5:	<sup>a</sup> anderer Wert	<sup>b</sup> 0	<sup>c</sup> 1/3	<sup>d</sup> 1/2	<sup>e</sup> 1
----	---------------------------	----------------	------------------	------------------	----------------

**Aufgabe 6:** Eine Zufallsvariable  $X$  ist normalverteilt mit unbekanntem Mittelwert und unbekannter Varianz. Die Variable  $x$  enthält Ihre Stichprobe mit 9 Beobachtungen. Wie ermitteln Sie die Breite des 90%-Konfidenzintervalls für den Mittelwert?

- 6a: Keine der folgenden Antworten ist richtig.
- 6b:  $2 * sd(x) * qt(.95, df=8)$
- 6c:  $2/3 * sd(x) * qt(.95, df=8)$
- 6d:  $2 * var(x) * qt(.95, df=8)$
- 6e:  $2/9 * var(x) * qt(.95, df=8)$

**Aufgabe 7:** Die Zufallsvariable  $X$  ist normalverteilt mit unbekanntem Mittelwert  $\mu$  und Standardabweichung  $\sigma_X = 4$ . Eine Stichprobe von 16 Beobachtungen ergibt einen Stichprobenmittelwert  $\bar{x} = 4$ . Ihre Alternativhypothese ist  $\mu \neq 5$ . Was ist der Absolutwert der Teststatistik?

7:	<sup>a</sup> anderer Wert	<sup>b</sup> 1/4	<sup>c</sup> 1	<sup>d</sup> 5/4	<sup>e</sup> 4
----	---------------------------	------------------	----------------	------------------	----------------

**Aufgabe 8:** Die Zufallsvariable  $X$  ist normalverteilt mit unbekanntem Mittelwert  $\mu$  und unbekannter Varianz. Eine Stichprobe von 16 Beobachtungen ergibt einen Stichprobenmittelwert  $\bar{x} = 6$  mit einer Standardabweichung  $\sigma_X = 4$ . Ihre Alternativhypothese ist  $\mu < z$ . Für welche Werte

von  $z$  können Sie auf einem Signifikanzniveau von 1% die Nullhypothese ablehnen?

- 8a: Keine der folgenden Antworten ist richtig.
- 8b: `qt(.01, 15)+6 > z`
- 8c: `z > qt(.99, 15)+6`
- 8d: `z > qt(.01, 15)+6`
- 8e: `qt(.99, 15)/4+6 > z`

**Aufgabe 9:** Die Zufallsvariable  $X$  ist normalverteilt mit unbekanntem Mittelwert  $\mu$  und unbekannter Varianz. Eine Stichprobe von 12 Beobachtungen ergibt einen Stichprobenmittelwert  $\bar{x} = 3$  mit einer Varianz  $\sigma_X^2 = 48$ . Ihre Alternativhypothese ist  $\mu \neq 5$ . Wie bestimmen Sie den  $p$ -Wert zu Ihrem Hypothesentest?

- 9a: Keine der folgenden Antworten ist richtig.
- 9b: `2*pt((3-5)/2, 11)`
- 9c: `pt((3-5)/2, 11)`
- 9d: `2*pt((5-3)/2, 11)`
- 9e: `pt((5-3)/2, 11)`

**Aufgabe 10:** Betrachten Sie das folgende Testergebnis.  $x$  und  $y$  sind jeweils Stichproben von  $X$  bzw.  $Y$ . Verwenden Sie ein Signifikanzniveau von 5%.

```
Welch Two Sample t-test
data: x and y
t = -1.6006, df = 17.465, p-value = 0.1274
alternative hypothesis: true difference in means
is not equal to 0
95 percent confidence interval:
-4.7912324 0.6527738
sample estimates:
mean of x mean of y
2.010298 4.079527
```

Was können Sie aus dem Testergebnis schließen?

- 10a: Es wurde ein paarweiser Test durchgeführt.
- 10b: Die Alternativhypothese in diesem Test war  $E(X) \neq E(Y)$ .
- 10c: Die Nullhypothese  $E(X) = E(Y)$  wird abgelehnt.
- 10d: Die Nullhypothese  $E(X) - E(Y) = 1$  wird abgelehnt.
- 10e: Die Nullhypothese  $E(X) - E(Y) = -1$  wird abgelehnt.

**Aufgabe 11:** Betrachten Sie das folgende Testergebnis. Verwenden Sie ein Signifikanzniveau von 5%.

```
Wilcoxon signed rank test
data: x and y
V = 16, p-value = 0.07715
alternative hypothesis: true location shift
is not equal to 0
```

Was können Sie aus dem Testergebnis schließen?

- 11a: Es wurde ein paarweiser Test durchgeführt.
- 11b: Die Alternativhypothese in diesem Test war  $E(X) > E(Y)$ .
- 11c: Die Nullhypothese  $E(X) = E(Y)$  wird abgelehnt.
- 11d: Die Nullhypothese  $E(X) - E(Y) = 1$  wird abgelehnt.
- 11e: Die Nullhypothese  $E(X) - E(Y) = 2$  wird abgelehnt.

**Aufgabe 12:**

Zwei Merkmale,  $X$  und  $Y$ , können jeweils die Werte  $A, B, C$  bzw.  $D$  und  $E$  annehmen. Die Häufigkeiten sind durch die

folgende Tabelle gegeben.

	$X = A$	$X = B$	$X = C$
$Y = D$	3	17	18
$Y = E$	18	17	19

Ihr Signifikanzniveau ist 5%. Interpretieren Sie das folgende Testergebnis:

```
Pearson's Chi-squared test
data: Z
X-squared = 8.2069, df = 2, p-value = 0.01652
```

Was können Sie aus dem Testergebnis schließen?

- 12a: Zwischen den beiden Merkmalen  $X$  und  $Y$  finden Sie keinen signifikanten Zusammenhang.
- 12b: Die Nullhypothese,  $X$  und  $Y$  seien unabhängig, kann man zum gegebenen Signifikanzniveau verwerfen.
- 12c: Die Nullhypothese,  $X$  und  $Y$  seien voneinander abhängig, kann man zum gegebenen Signifikanzniveau verwerfen.
- 12d: Die Nullhypothese,  $X$  und  $Y$  seien unabhängig, kann man nicht zum gegebenen Signifikanzniveau verwerfen.
- 12e: Um zu testen, ob  $X$  und  $Y$  voneinander unabhängig sind, sollte man besser den Korrelationskoeffizienten betrachten.

**Aufgabe 13:** Welche Wahrscheinlichkeit gibt der  $p$ -Wert an?

- 13a: Keine der folgenden Antworten ist richtig.
- 13b: Die Wahrscheinlichkeit, eine Stichprobe zu ziehen, die mindestens so advers zur Nullhypothese ist, wie die gezogene Stichprobe, falls die Nullhypothese falsch ist.
- 13c: Die Wahrscheinlichkeit eines Fehlers erster Art.
- 13d: Die Wahrscheinlichkeit eines Fehlers zweiter Art.
- 13e: Die Wahrscheinlichkeit, die Alternativhypothese abzulehnen, wenn sie falsch ist.

**Aufgabe 14:** In der folgenden Regression schätzen Sie den Effekt der Variablen  $a$  und  $b$  auf die Variable  $Y$ . Die Variablen  $a$  und  $b$  können jeweils die Werte 0 oder 1 annehmen, Sie erhalten den folgenden Output:

```
Call: lm(formula = y ~ a * b)
      Estimate Std. Error t value Pr(>|t|)
(Intercept)  3.00      0.71    4.24    0.00
a             3.00      1.00    3.00    0.01
b             1.00      1.00    1.00    0.33
a:b          2.00      1.41    1.41    0.17
Residual standard error: 1.732 on 20 degrees of freedom
Multiple R-squared: 0.6774, Adjusted R-squared: 0.629
F-statistic: 14 on 3 and 20 DF, p-value: 3.795e-05
```

Welchen Wert für  $Y$  erwarten Sie, wenn  $a = 0$  und  $b = 0$ ?

14: 

<sup>a</sup> anderer Wert	<sup>b</sup> 0	<sup>c</sup> 3	<sup>d</sup> 4	<sup>e</sup> 6
---------------------------	----------------	----------------	----------------	----------------

**Aufgabe 15:** Welchen Wert für  $Y$  erwarten Sie, wenn  $a = 1$  und  $b = 0$ ?

15: 

<sup>a</sup> anderer Wert	<sup>b</sup> 0	<sup>c</sup> 3	<sup>d</sup> 4	<sup>e</sup> 6
---------------------------	----------------	----------------	----------------	----------------

**Aufgabe 16:** Welchen Wert für  $Y$  erwarten Sie, wenn  $a = 1$  und  $b = 1$ ?

16: 

<sup>a</sup> anderer Wert	<sup>b</sup> 4	<sup>c</sup> 6	<sup>d</sup> 7	<sup>e</sup> 9
---------------------------	----------------	----------------	----------------	----------------

**Aufgabe 17:** Wie groß ist der marginale Effekt von  $b$  auf  $Y$  falls  $a = 1$ ?

17:	<sup>a</sup> anderer Wert	<sup>b</sup> 1	<sup>c</sup> 2	<sup>d</sup> 3	<sup>e</sup> 4
-----	---------------------------	----------------	----------------	----------------	----------------

**Aufgabe 18:** Wie groß ist der marginale Effekt von  $a$  auf  $Y$  falls  $b = 0$ ?

18:	<sup>a</sup> anderer Wert	<sup>b</sup> 1	<sup>c</sup> 2	<sup>d</sup> 3	<sup>e</sup> 4
-----	---------------------------	----------------	----------------	----------------	----------------

**Aufgabe 19:** Wie bestimmen Sie für die obige Regression die Untergrenze des 95%-Konfidenzintervalls für den Koeffizienten von  $a$ ?

- 19a: Keine der folgenden Antworten ist richtig.  
 19b:  $3+qt(.025, df=20)$   
 19c:  $3+qt(.975, df=20)$   
 19d:  $3+qt(.975, df=3)$   
 19e:  $3-qt(.975, df=3)$

**Aufgabe 20:** Betrachten Sie weiter die obige Regression. Gehen Sie von einem Signifikanzniveau von 5% aus. Nehmen Sie ferner an, dass die Residuen normalverteilt sind. Welche Aussagen werden durch den obigen Output gestützt?

- 20a:  $b$  hat einen signifikanten Einfluss auf  $a$ .  
 20b:  $b$  hat einen signifikanten Einfluss auf  $Y$  falls  $a = 0$ .  
 20c:  $a$  hat einen signifikanten Einfluss auf  $Y$  falls  $b = 0$ .  
 20d: Die Hypothese, der Koeffizient von  $b$  sei  $-2$ , wird abgelehnt.

20e: Die Hypothese, der Koeffizient von  $b$  sei 1, wird abgelehnt.

**Aufgabe 21:** Über den Zusammenhang zwischen  $X$  und  $Y$  wissen Sie folgendes:

- Wenn  $X$  um 1 steigt, dann steigt  $Y$  um  $-0.35$ .  
 Wenn  $X$  um 1% steigt, dann steigt  $Y$  um  $-0.01$ .  
 Wenn  $X$  um 1 steigt, dann steigt  $Y$  um  $-8.3\%$ .  
 Wenn  $X$  um 1% steigt, dann steigt  $Y$  um  $-0.28\%$ .

Welchen Wert für  $\beta_1$  schätzen Sie in  $Y = \beta_0 + \beta_1 \cdot X + u$ ?

21:	<sup>a</sup> anderer Wert	<sup>b</sup> -35	<sup>c</sup> -10	<sup>d</sup> -8.3	<sup>e</sup> -0.35
-----	---------------------------	------------------	------------------	-------------------	--------------------

**Aufgabe 22:** Welchen Wert für  $\beta_1$  schätzen Sie in  $Y = \beta_0 + \beta_1 \cdot \log X + u$ ?

22:	<sup>a</sup> anderer Wert	<sup>b</sup> -830	<sup>c</sup> -1	<sup>d</sup> -0.83	<sup>e</sup> -0.01
-----	---------------------------	-------------------	-----------------	--------------------	--------------------

**Aufgabe 23:** Welchen Wert für  $\beta_1$  schätzen Sie in  $\log Y = \beta_0 + \beta_1 \cdot X + u$ ?

23:	<sup>a</sup> anderer Wert	<sup>b</sup> -830	<sup>c</sup> -1	<sup>d</sup> -0.083	<sup>e</sup> -0.01
-----	---------------------------	-------------------	-----------------	---------------------	--------------------

**Aufgabe 24:** Welchen Wert für  $\beta_1$  schätzen Sie in  $\log Y = \beta_0 + \beta_1 \cdot \log X + u$ ?

24:	<sup>a</sup> anderer Wert	<sup>b</sup> -28	<sup>c</sup> -8.3	<sup>d</sup> -1	<sup>e</sup> -0.28
-----	---------------------------	------------------	-------------------	-----------------	--------------------

## 8 Klausur BW24.1, Februar 2016

Ihre Version ist . Bitte vergessen Sie nicht, die Version in den Lösungsbogen zu übertragen. Lösen Sie die Aufgaben zunächst hier (und ggf. auf dem Schmierpapier) und übertragen Sie die Lösungen zum Schluss in den Lösungsbogen. Geben Sie nur den Lösungsbogen ab, und behalten Sie dieses Aufgabenblatt. Wir wünschen Ihnen viel Erfolg!

**Aufgabe 1:** Ihre Stichprobe der Zufallsvariablen  $X$  enthält 6 unabhängige und normalverteilte Beobachtungen:  $X_1, \dots, X_6$ . Welche Schätzfunktionen für  $E(X)$  sind erwartungstreu?

1:	<sup>a</sup> $\bar{X}_1 + X_2 + X_3$	<sup>b</sup> $\frac{1}{3}X_2 + \frac{2}{3}X_3$	<sup>c</sup> $\bar{X}_1 - X_2 + X_3$	<sup>d</sup> $X_3^2/X_3$	<sup>e</sup> $\bar{X}_1 - X_2 - X_3$
----	--------------------------------------	--	--------------------------------------	--------------------------	--------------------------------------

**Aufgabe 2:** Welche Schätzfunktionen für  $E(X)$  sind effizienter als  $X_3$ ?

2:	<sup>a</sup> $\frac{1}{6} \sum_{i=1}^6 X_i$	<sup>b</sup> $\sum_{i=1}^6 X_i$	<sup>c</sup> $\bar{X}_1 + X_2 + X_3$	<sup>d</sup> $\bar{X}_1 - X_2 + X_3$	<sup>e</sup> $\bar{X}_1 - X_2 - X_3$
----	---	---------------------------------	--------------------------------------	--------------------------------------	--------------------------------------

**Aufgabe 3:** Die Zufallsvariable  $X$  hat eine Varianz von  $\sigma_X^2$ . Sie haben vor, eine Stichprobe von  $n$  unabhängigen und identisch verteilten Beobachtungen zu ziehen. Wie groß muß  $n$  sein, damit die Varianz des Stichprobenmittelwerts  $\sigma_{\bar{X}}^2 = 10$  ist?

3:	<sup>a</sup> anderer Wert	<sup>b</sup> $\sigma_X^2/10$	<sup>c</sup> $\sigma_X/\sqrt{10}$	<sup>d</sup> $\sigma_X^2/100$	<sup>e</sup> $10 \cdot \sigma_X^2$
----	---------------------------	------------------------------	-----------------------------------	-------------------------------	------------------------------------

**Aufgabe 4:**

Eine Zufallsvariable  $X$  folgt einer Exponentialverteilung mit Parameter  $\lambda$ . Eine Stichprobe ergibt zwei Beobachtungen:  $\{1, 1\}$ .

Was ist der Maximum-Likelihood Schätzer für  $\lambda$ ?

4:	<sup>a</sup> anderer Wert	<sup>b</sup> 0	<sup>c</sup> 1	<sup>d</sup> 2	<sup>e</sup> 3
----	---------------------------	----------------	----------------	----------------	----------------

**Aufgabe 5:** Was ist der Momentenschätzer für  $\lambda$  auf Basis des ersten Moments?

5:	<sup>a</sup> anderer Wert	<sup>b</sup> 0	<sup>c</sup> 1	<sup>d</sup> 2	<sup>e</sup> 3
----	---------------------------	----------------	----------------	----------------	----------------

**Aufgabe 6:** Eine Zufallsvariable  $X$  ist normalverteilt mit unbekanntem Mittelwert und Varianz  $\sigma_X^2 = 9$ . Die Variable  $x$  enthält Ihre Stichprobe mit 16 Beobachtungen. Wie ermitteln Sie die Untergrenze des 99%-Konfidenzintervalls für den Mittelwert?

- 6a: Keine der folgenden Antworten ist richtig.  
 6b:  $\text{mean}(x) - 3 * \text{sd}(x) / 4 * \text{qt}(.995, df=7)$   
 6c:  $\text{mean}(x) + 3 * \text{qnorm}(.005)$   
 6d:  $\text{mean}(x) + 3 / 4 * \text{qt}(.995, df=7)$   
 6e:  $\text{mean}(x) - 3 / 4 * \text{qnorm}(.995)$

**Aufgabe 7:** Die Zufallsvariable  $X$  ist normalverteilt mit unbekanntem Mittelwert  $\mu$  und Varianz  $\sigma^2 = 9$ . Eine Stichprobe von 25 Beobachtungen ergibt einen Stichprobenmittelwert  $\bar{x} = -3$ . Ihre Alternativhypothese ist  $\mu < -1$ . Was ist der Absolutwert der Teststatistik?

7:	<sup>a</sup> anderer Wert	<sup>b</sup> 2/5	<sup>c</sup> 2/3	<sup>d</sup> 10/3	<sup>e</sup> 50/9
----	---------------------------	------------------	------------------	-------------------	-------------------

**Aufgabe 8:** Die Zufallsvariable  $X$  ist normalverteilt mit unbekanntem Mittelwert  $\mu$  und unbekannter Varianz. Eine Stichprobe von 9 Beobachtungen ergibt einen

Stichprobenmittelwert  $\bar{x} = 2$  mit einer Standardabweichung  $\sigma_X = 6$ . Ihre Alternativhypothese ist  $\mu > z$ . Für welche Werte von  $z$  können Sie auf einem Signifikanzniveau von 5% die Nullhypothese ablehnen?

- 8a: Keine der folgenden Antworten ist richtig.
- 8b:  $z > qt(.95, 8) * 6 + 2$
- 8c:  $z > qt(.95, 8) * 2 + 2$
- 8d:  $z < qt(.05, 8) * 2 + 2$
- 8e:  $z < qt(.05, 8) * 6 + 2$

**Aufgabe 9:** Die Zufallsvariable  $X$  ist normalverteilt mit unbekanntem Mittelwert  $\mu$  und unbekannter Varianz. Eine Stichprobe von 9 Beobachtungen ergibt einen Stichprobenmittelwert  $\bar{x} = 12$  mit einer Standardabweichung  $\sigma_X = 12$ . Ihre Alternativhypothese ist  $\mu \neq 2$ . Wie bestimmen Sie den  $p$ -Wert zu Ihrem Hypothesentest?

- 9a: Keine der folgenden Antworten ist richtig.
- 9b:  $2 * pt((12-2)/(12/3), 8)$
- 9c:  $2 * pt(-(12-2)/12, 8)$
- 9d:  $pt(-(12-2)/(12/3), 8)$
- 9e:  $2 * pt(-(12-2)/(12/3), 8)$

**Aufgabe 10:** Betrachten Sie das folgende Testergebnis.  $x$  und  $y$  sind jeweils Stichproben von  $X$  bzw.  $Y$ . Verwenden Sie ein Signifikanzniveau von 5%.

```
Welch Two Sample t-test
data: x and y
t = -2.2134, df = 10.974, p-value = 0.04898
alternative hypothesis: true difference in means
is not equal to 0
95 percent confidence interval:
-1.835183279 -0.004881935
sample estimates:
mean of x mean of y
-0.1644569 0.7555757
```

Was können Sie aus dem Testergebnis schließen?

- 10a: Es wurde ein paarweiser Test durchgeführt.
- 10b: Die Alternativhypothese in diesem Test war  $E(X) \neq E(Y)$ .
- 10c: Die Nullhypothese  $E(X) = E(Y)$  wird abgelehnt.
- 10d: Die Nullhypothese  $E(X) - E(Y) = 1$  wird abgelehnt.
- 10e: Die Nullhypothese  $E(X) - E(Y) = -1$  wird abgelehnt.

**Aufgabe 11:** Betrachten Sie das folgende Testergebnis. Verwenden Sie ein Signifikanzniveau von 5%.

```
Wilcoxon rank sum test
data: x and y
W = 15, p-value = 0.08298
alternative hypothesis: true location shift
is not equal to 0
```

Was können Sie aus dem Testergebnis schließen?

- 11a: Es wurde kein paarweiser Test durchgeführt.
- 11b: Die Alternativhypothese in diesem Test war  $E(X) > E(Y)$ .
- 11c: Die Nullhypothese  $E(X) = E(Y)$  wird abgelehnt.
- 11d: Die Nullhypothese  $E(X) - E(Y) = 1$  wird abgelehnt.
- 11e: Die Nullhypothese  $E(X) - E(Y) = 2$  wird abgelehnt.

**Aufgabe 12:**

Zwei Merkmale,  $X$  und  $Y$ , können jeweils die Werte  $A, B, C$  bzw.  $D$  und  $E$  annehmen. Die Häufigkeiten sind durch die

folgende Tabelle gegeben.

	$X = A$	$X = B$	$X = C$
$Y = D$	11	26	13
$Y = E$	14	12	14

Ihr Signifikanzniveau ist 10%. Interpretieren Sie das folgende Testergebnis:

```
Pearson's Chi-squared test
data: Z
X-squared = 4.4994, df = 2, p-value = 0.1054
```

Was können Sie aus dem Testergebnis schließen?

- 12a: Zwischen den beiden Merkmalen  $X$  und  $Y$  finden Sie keinen signifikanten Zusammenhang.
- 12b: Die Nullhypothese,  $X$  und  $Y$  seien unabhängig, kann man zum gegebenen Signifikanzniveau verwerfen.
- 12c: Die Nullhypothese,  $X$  und  $Y$  seien voneinander abhängig, kann man zum gegebenen Signifikanzniveau verwerfen.
- 12d: Die Nullhypothese,  $X$  und  $Y$  seien unabhängig, kann man nicht zum gegebenen Signifikanzniveau verwerfen.
- 12e: Um zu testen, ob  $X$  und  $Y$  voneinander unabhängig sind, sollte man besser den Korrelationskoeffizienten betrachten.

**Aufgabe 13:** Welche Wahrscheinlichkeit gibt der  $p$ -Wert an?

- 13a: Keine der folgenden Antworten ist richtig.
- 13b: Die Wahrscheinlichkeit, eine Stichprobe zu ziehen, die mindestens so advers zur Nullhypothese ist, wie die gezogene Stichprobe, falls die Nullhypothese wahr ist.
- 13c: Die Wahrscheinlichkeit, eine Stichprobe zu ziehen, die mindestens so advers zur Nullhypothese ist, wie die gezogene Stichprobe.
- 13d: Die Wahrscheinlichkeit, die Alternativhypothese abzulehnen, wenn sie falsch ist.
- 13e: Die Wahrscheinlichkeit, die Nullhypothese abzulehnen, wenn sie wahr ist.

**Aufgabe 14:** In der folgenden Regression schätzen Sie den Effekt der Variablen  $a$  und  $b$  auf die Variable  $Y$ . Die Variablen  $a$  und  $b$  können jeweils die Werte 0 oder 1 annehmen, Sie erhalten den folgenden Output:

```
Call: lm(formula = y ~ a * b)
      Estimate Std. Error t value Pr(>|t|)
(Intercept) -2.00      0.71    -2.83  0.01
a             -2.00      1.00    -2.00  0.06
b              3.00      1.00     3.00  0.01
a:b           1.00      1.41     0.71  0.49
Residual standard error: 1.732 on 20 degrees of freedom
Multiple R-squared: 0.596, Adjusted R-squared: 0.5354
F-statistic: 9.833 on 3 and 20 DF, p-value: 0.0003411
```

Welchen Wert für  $Y$  erwarten Sie, wenn  $a = 0$  und  $b = 0$ ?

14:	<sup>a</sup> anderer Wert	<sup>b</sup> -2	<sup>c</sup> 1	<sup>d</sup> 2	<sup>e</sup> 3
-----	---------------------------	-----------------	----------------	----------------	----------------

**Aufgabe 15:** Welchen Wert für  $Y$  erwarten Sie, wenn  $a = 1$  und  $b = 0$ ?

15:	<sup>a</sup> anderer Wert	<sup>b</sup> -4	<sup>c</sup> 0	<sup>d</sup> 2	<sup>e</sup> 4
-----	---------------------------	-----------------	----------------	----------------	----------------

**Aufgabe 16:** Welchen Wert für  $Y$  erwarten Sie, wenn  $a = 1$  und  $b = 1$ ?

16:	<sup>a</sup> anderer Wert	<sup>b</sup> -4	<sup>c</sup> -1	<sup>d</sup> 0	<sup>e</sup> 1
-----	---------------------------	-----------------	-----------------	----------------	----------------

**Aufgabe 17:** Wie groß ist der marginale Effekt von  $b$  auf  $Y$  falls  $a = 1$ ?

17:	<sup>a</sup> anderer Wert	<sup>b</sup> 1	<sup>c</sup> 2	<sup>d</sup> 3	<sup>e</sup> 4
-----	---------------------------	----------------	----------------	----------------	----------------

**Aufgabe 18:** Wie groß ist der marginale Effekt von  $a$  auf  $Y$  falls  $b = 0$ ?

18:	<sup>a</sup> anderer Wert	<sup>b</sup> -4	<sup>c</sup> -3	<sup>d</sup> -2	<sup>e</sup> 0
-----	---------------------------	-----------------	-----------------	-----------------	----------------

**Aufgabe 19:** Wie bestimmen Sie für die obige Regression die Obergrenze des 95%-Konfidenzintervalls für den Koeffizienten von  $a$ ?

- 19a: Keine der folgenden Antworten ist richtig.  
 19b:  $-2 + qt(.95, df=20)$   
 19c:  $3 - qt(.975, df=20)$   
 19d:  $3 + qt(.975, df=20)$   
 19e:  $-2 + qt(.975, df=20)$

**Aufgabe 20:** Betrachten Sie weiter die obige Regression. Gehen Sie von einem Signifikanzniveau von 5% aus. Nehmen Sie ferner an, dass die Residuen normalverteilt sind:

- 20a:  $b$  hat einen signifikanten Einfluss auf  $a$ .  
 20b:  $b$  hat einen signifikanten Einfluss auf  $Y$ .  
 20c: Die Hypothese, der Koeffizient von  $b$  sei  $-1$ , wird abgelehnt.  
 20d: Die Hypothese, der Koeffizient von  $a$  sei  $-1$ , wird abgelehnt.

20e: Das Modell erklärt 59.6% der Varianz von  $Y$ .

**Aufgabe 21:** Über den Zusammenhang zwischen  $X$  und  $Y$  wissen Sie folgendes:

- Wenn  $X$  um 1 steigt, dann steigt  $Y$  um  $-2.9$ .  
 Wenn  $X$  um 1% steigt, dann steigt  $Y$  um  $-0.0021$ .  
 Wenn  $X$  um 1 steigt, dann steigt  $Y$  um  $-540\%$ .  
 Wenn  $X$  um 1% steigt, dann steigt  $Y$  um  $-0.48\%$ .

Welchen Wert für  $\beta_1$  schätzen Sie in  $Y = \beta_0 + \beta_1 \cdot X + u$ ?

21:	<sup>a</sup> anderer Wert	<sup>b</sup> -0.48	<sup>c</sup> -0.29	<sup>d</sup> -2.9	<sup>e</sup> -5.4
-----	---------------------------	--------------------	--------------------	-------------------	-------------------

**Aufgabe 22:** Welchen Wert für  $\beta_1$  schätzen Sie in  $Y = \beta_0 + \beta_1 \cdot \log X + u$ ?

22:	<sup>a</sup> anderer Wert	<sup>b</sup> -0.48	<sup>c</sup> -0.21	<sup>d</sup> -540	<sup>e</sup> -5.4
-----	---------------------------	--------------------	--------------------	-------------------	-------------------

**Aufgabe 23:** Welchen Wert für  $\beta_1$  schätzen Sie in  $\log Y = \beta_0 + \beta_1 \cdot X + u$ ?

23:	<sup>a</sup> anderer Wert	<sup>b</sup> -0.48	<sup>c</sup> -0.21	<sup>d</sup> -0.0021	<sup>e</sup> -5.4
-----	---------------------------	--------------------	--------------------	----------------------	-------------------

**Aufgabe 24:** Welchen Wert für  $\beta_1$  schätzen Sie in  $\log Y = \beta_0 + \beta_1 \cdot \log X + u$ ?

24:	<sup>a</sup> anderer Wert	<sup>b</sup> -0.48	<sup>c</sup> -0.0048	<sup>d</sup> -48	<sup>e</sup> -540
-----	---------------------------	--------------------	----------------------	------------------	-------------------

## 9 Klausur BW24.1, Mai 2015

Ihre Version ist . Bitte vergessen Sie nicht, die Version in den Lösungsbogen zu übertragen. Lösen Sie die Aufgaben zunächst hier (und ggf. auf dem Schmierpapier) und übertragen Sie die Lösungen zum Schluss in den Lösungsbogen. Geben Sie nur den Lösungsbogen ab, und behalten Sie dieses Aufgabenblatt. Wir wünschen Ihnen viel Erfolg!

**Aufgabe 1:** Ihre Stichprobe der Zufallsvariablen  $X$  enthält 5 unabhängige und identisch verteilte Beobachtungen:  $X_1, \dots, X_5$ . Welche Schätzfunktionen für  $E(X)$  sind erwartungstreu?

1:	<sup>a</sup> $X_1 - X_2$	<sup>b</sup> $(X_3 + X_4)/2$	<sup>c</sup> $X_3 - X_2 + X_4$	<sup>d</sup> $X_5$	<sup>e</sup> $X_1 + X_2$
----	--------------------------	------------------------------	--------------------------------	--------------------	--------------------------

**Aufgabe 2:** Welche Schätzfunktionen für  $E(X)$  sind weniger effizient als  $X_1$ ?

2:	<sup>a</sup> $\frac{1}{5} \sum_{i=1}^5 X_i$	<sup>b</sup> $X_2 + X_3 - X_5$	<sup>c</sup> $2X_1 - X_2$	<sup>d</sup> $X_1 - X_2$	<sup>e</sup> $X_1/X_2$
----	---	--------------------------------	---------------------------	--------------------------	------------------------

**Aufgabe 3:** Die Zufallsvariable  $X$  hat eine Varianz von  $\sigma^2$ . Sie haben vor, eine Stichprobe unabhängiger und identisch verteilter Beobachtungen  $X_1, \dots, X_5$  zu ziehen. Wie berechnen Sie die Varianz von  $(3X_3 - X_4)$ ?

3:	<sup>a</sup> anderer Wert	<sup>b</sup> $\sigma$	<sup>c</sup> $2\sigma^2$	<sup>d</sup> $4\sigma^2$	<sup>e</sup> $10\sigma^2$
----	---------------------------	-----------------------	--------------------------	--------------------------	---------------------------

**Aufgabe 4:** Eine Zufallsvariable  $X$  folgt einer Exponentialverteilung mit Parameter  $\lambda$ . Eine Stichprobe ergibt drei Beobachtungen:

$\{0, 1, 0\}$ .

Was ist der Maximum-Likelihood Schätzer für  $\lambda$ ?

4:	<sup>a</sup> anderer Wert	<sup>b</sup> 0	<sup>c</sup> 1	<sup>d</sup> 2	<sup>e</sup> 3
----	---------------------------	----------------	----------------	----------------	----------------

**Aufgabe 5:** Was ist der Momentenschätzer für  $\lambda$  auf Basis des ersten Moments?

5:	<sup>a</sup> anderer Wert	<sup>b</sup> 0	<sup>c</sup> 1	<sup>d</sup> 2	<sup>e</sup> 3
----	---------------------------	----------------	----------------	----------------	----------------

**Aufgabe 6:** Eine Zufallsvariable  $X$  ist normalverteilt mit unbekanntem Mittelwert und unbekannter Standardabweichung. Die Variable  $x$  enthält Ihre Stichprobe mit 8 Beobachtungen. Wie ermitteln Sie die Untergrenze des 99%-Konfidenzintervalls für den Mittelwert?

- 6a: Keine der folgenden Antworten ist richtig.  
 6b:  $\text{mean}(x) - \text{sd}(x) / \sqrt{8} * qt(.995, df=7)$   
 6c:  $\text{mean}(x) - \text{sd}(x) / 8 * qt(.995, df=7)$   
 6d:  $\text{mean}(x) - \text{sd}(x) / 8 * qnorm(.995)$   
 6e:  $\text{mean}(x) + \text{sd}(x) / \sqrt{8} * qt(.995, df=7)$

**Aufgabe 7:** Die Zufallsvariable  $X$  ist normalverteilt mit unbekanntem Mittelwert  $\mu$  und unbekannter Standardabweichung  $\sigma$ . Eine Stichprobe von 25 Beobachtungen ergibt einen Stichprobenmittelwert  $\bar{x} = -8$  und eine Standardabweichung der Stichprobenbeobachtungen von  $\hat{\sigma}_X = 35$ . Ihre Alternativhypothese ist  $\mu \leq -1$ . Was ist der Absolutwert der Teststatistik?

7:	<sup>a</sup> anderer Wert	<sup>b</sup> 1/5	<sup>c</sup> 1	<sup>d</sup> 2	<sup>e</sup> 5
----	---------------------------	------------------	----------------	----------------	----------------

**Aufgabe 8:** Nehmen Sie nun an, dass Ihre Alternativhypothese im obigen Problem  $\mu < x$  ist. Für welche Werte von  $x$  können Sie auf einem Signifikanzniveau von 5% die Nullhypothese ablehnen?

- 8a: Keine der folgenden Antworten ist richtig.  
 8b:  $x > -qnorm(.975)*7-8$   
 8c:  $x < qt(.025, df=24)*1.4-8$   
 8d:  $x < qt(.05, df=24)*7-8$   
 8e:  $x > qnorm(.05)*35-8$

**Aufgabe 9:** Betrachten Sie das folgende Testergebnis.  $x$  und  $y$  sind jeweils Stichproben von  $X$  bzw.  $Y$ . Verwenden Sie ein Signifikanzniveau von 5%.

```
Welch Two Sample t-test, data: x and y
t = -1.9814, df = 21.85, p-value = 0.06026
alternative hypothesis: true difference
in means is not equal to 0
95 percent confidence interval:
-2.0671217 0.0475229
```

```
sample estimates:
mean of x mean of y
0.1737619 1.1835613
```

- 9a: Es wurde ein paarweiser Test durchgeführt.  
 9b: Die Alternativhypothese in diesem Test war  $E(X) \neq E(Y)$ .  
 9c: Die Nullhypothese  $E(X) = E(Y)$  wird nicht abgelehnt.  
 9d: Die Nullhypothese  $E(X) - E(Y) = 1$  wird abgelehnt.  
 9e: Die Nullhypothese  $E(X) - E(Y) = -2$  wird abgelehnt.

**Aufgabe 10:** Betrachten Sie das folgende Testergebnis. Verwenden Sie ein Signifikanzniveau von 5%.

```
Wilcoxon rank sum test
data: x and y
W = 38, p-value = 0.05186
alternative hypothesis: true location shift
is not equal to 0
```

- 10a: Es wurde kein paarweiser Test durchgeführt.  
 10b: Die Alternativhypothese in diesem Test war  $E(X) > E(Y)$ .  
 10c: Die Nullhypothese  $E(X) = E(Y)$  wird abgelehnt.  
 10d: Die Nullhypothese  $E(X) - E(Y) = 1$  wird abgelehnt.  
 10e: Die Nullhypothese  $E(X) - E(Y) = 2$  wird abgelehnt.

**Aufgabe 11:** Sie untersuchen den Zusammenhang zwischen zwei Merkmalen,  $X$  und  $Y$ . Das Merkmal  $X$  kann die beiden Werte  $A$  und  $B$  annehmen.  $Y$  hat entweder den Wert  $C$ ,  $D$  oder  $E$ . Die Häufigkeit der Merkmale ist durch die folgende Tabelle gegeben:

	$Y = C$	$Y = D$	$Y = E$
$X = A$	28	12	16
$X = B$	14	19	14

Ihr Signifikanzniveau ist 10%. Interpretieren Sie das folgende Testergebnis:

```
Pearson's Chi-squared test
X-squared = 5.6373, df = 2, p-value = 0.05969
```

- 11a: Zwischen den beiden Merkmalen  $X$  und  $Y$  gibt es keinen signifikanten Zusammenhang.  
 11b: Die Nullhypothese,  $X$  und  $Y$  seien unabhängig, kann man zum gegebenen Signifikanzniveau verwerfen.  
 11c: Die Nullhypothese,  $X$  und  $Y$  seien voneinander abhängig, kann man zum gegebenen Signifikanzniveau verwerfen.

- 11d: Die Nullhypothese,  $X$  und  $Y$  seien unabhängig, kann man nicht zum gegebenen Signifikanzniveau verwerfen.  
 11e: Um zu testen, ob  $X$  und  $Y$  voneinander unabhängig sind, sollte man besser den Korrelationskoeffizienten betrachten.

**Aufgabe 12:** Welche Wahrscheinlichkeit gibt der  $p$ -Wert an?

- 12a: Keine der folgenden Antworten ist richtig.  
 12b: Die Wahrscheinlichkeit, eine Stichprobe zu ziehen, die mindestens so advers zur Nullhypothese ist, wie die gezogene Stichprobe, falls die Nullhypothese wahr ist.  
 12c: Die Wahrscheinlichkeit, eine Stichprobe zu ziehen, die mindestens so advers zur Nullhypothese ist, wie die gezogene Stichprobe, falls die Nullhypothese falsch ist.  
 12d: Die Wahrscheinlichkeit, eine Stichprobe zu ziehen, die mindestens so advers zur Nullhypothese ist, wie die gezogene Stichprobe.  
 12e: Die Wahrscheinlichkeit eines Fehlers erster Art.

**Aufgabe 13:** In der folgenden Regression schätzen Sie den Effekt der Variablen  $a$  und  $b$  auf die Variable  $Y$ . Die Variablen  $a$  und  $b$  können jeweils die Werte 0 oder 1 annehmen, Sie erhalten den folgenden Output:

```
Call: lm(formula = y ~ a * b)
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 4.0000    0.7071    5.657 1.55e-05
a             4.0000    1.0000    4.000 0.000704
b            -4.0000    1.0000   -4.000 0.000704
a:b           -2.0000    1.4142   -1.414 0.172678
Residual standard error: 1.732 on 20 degrees of freedom
Multiple R-squared: 0.7778
F-statistic: 23.3 on 3 and 20 DF, p-value: 9.7e-07
```

Welchen Wert für  $Y$  erwarten Sie, wenn  $a = 0$  und  $b = 0$ ?

13:	<sup>a</sup> anderer Wert	<sup>b</sup> 0	<sup>c</sup> 2	<sup>d</sup> 4	<sup>e</sup> 5
-----	---------------------------	----------------	----------------	----------------	----------------

**Aufgabe 14:** Welchen Wert für  $Y$  erwarten Sie, wenn  $a = 1$  und  $b = 0$ ?

14:	<sup>a</sup> anderer Wert	<sup>b</sup> -2	<sup>c</sup> 0	<sup>d</sup> 4	<sup>e</sup> 8
-----	---------------------------	-----------------	----------------	----------------	----------------

**Aufgabe 15:** Welchen Wert für  $Y$  erwarten Sie, wenn  $a = 1$  und  $b = 1$ ?

15:	<sup>a</sup> anderer Wert	<sup>b</sup> 0	<sup>c</sup> 2	<sup>d</sup> 4	<sup>e</sup> 8
-----	---------------------------	----------------	----------------	----------------	----------------

**Aufgabe 16:** Wie groß ist der marginale Effekt von  $b$  auf  $Y$  falls  $a = 1$ ?

16:	<sup>a</sup> anderer Wert	<sup>b</sup> -8	<sup>c</sup> -6	<sup>d</sup> -4	<sup>e</sup> 4
-----	---------------------------	-----------------	-----------------	-----------------	----------------

**Aufgabe 17:** Wie groß ist der marginale Effekt von  $a$  auf  $Y$  falls  $b = 0$ ?

17:	<sup>a</sup> anderer Wert	<sup>b</sup> -4	<sup>c</sup> -2	<sup>d</sup> 4	<sup>e</sup> 8
-----	---------------------------	-----------------	-----------------	----------------	----------------

**Aufgabe 18:** Wie bestimmen Sie für die obige Regression die Untergrenze des 95%-Konfidenzintervalls für den Koeffizienten von  $b$ ?

- 18a: Keine der folgenden Antworten ist richtig.  
 18b:  $4-2*qt(.975, df=20)$   
 18c:  $4-qt(.975, df=20)$

18d:  $-4 - qt(.975, df=20)$

18e:  $4 + qt(.975, df=20)$

**Aufgabe 19:** Betrachten Sie weiter die obige Regression. Gehen Sie von einem Signifikanzniveau von 5% aus. Nehmen Sie ferner an, dass die Residuen normalverteilt sind:

19a:  $a$  hat einen signifikanten Einfluss auf  $b$ .

19b:  $b$  hat einen signifikanten Einfluss auf  $Y$ .

19c: Die Hypothese, der Koeffizient von  $b$  sei 5, wird abgelehnt.

19d: Die Hypothese, der Koeffizient von  $b$  sei  $-1$ , wird abgelehnt.

19e: Das Modell erklärt 23.3% der Varianz von  $Y$ .

**Aufgabe 20:** Sie verwenden das folgende Modell, um den Zusammenhang zwischen den erklärenden Variablen  $X$  und  $Z$  und der abhängigen Variablen  $Y$  zu schätzen:

$$\log Y = \beta_0 + \beta_1 \log X + \beta_2 Z^2 + u$$

Ihre Schätzung ergibt  $\beta_0 = 3, \beta_1 = -1, \beta_2 = 2$ .

20a: Keine der folgenden Antworten ist richtig.

20b: Wenn  $X$  um 1% steigt, verringert sich  $Y$  um 1%.

20c: Wenn  $X$  um 0.01 steigt, verringert sich  $Y$  um 1%.

20d: Wenn  $X$  um 1% steigt, verringert sich  $Y$  um 0.01.

20e: Wenn  $X$  um 0.01 sinkt, verringert sich  $Y$  um 0.01.

## 10 Klausur BW24.1, Februar 2015

Ihre Version ist . Bitte vergessen Sie nicht, die Version in den Lösungsbogen zu übertragen. Lösen Sie die Aufgaben zunächst hier (und ggf. auf dem Schmierpapier) und übertragen Sie die Lösungen zum Schluss in den Lösungsbogen. Geben Sie nur den Lösungsbogen ab, und behalten Sie dieses Aufgabenblatt. Wir wünschen Ihnen viel Erfolg!

**Aufgabe 1:** Ihre Stichprobe der Zufallsvariablen  $X$  enthält 5 unabhängige und identisch verteilte Beobachtungen:

$X_1, \dots, X_5$ . Welche Schätzfunktionen für  $E(X)$  sind erwartungstreu?

1:	<sup>a</sup> $X_1 + X_2$	<sup>b</sup> $(X_1 + X_2)/2$	<sup>c</sup> $X_1 - X_2$	<sup>d</sup> $X_1$	<sup>e</sup> $X_1/X_2$
----	--------------------------	------------------------------	--------------------------	--------------------	------------------------

**Aufgabe 2:** Welche Schätzfunktionen für  $E(X)$  sind weniger effizient als  $(X_1 + X_2)/2$ ?

2:	<sup>a</sup> $\frac{1}{5} \sum_{i=1}^5 X_i$	<sup>b</sup> $X_2$	<sup>c</sup> $X_1 + X_2 - X_3$	<sup>d</sup> $X_1$	<sup>e</sup> $X_1/X_2$
----	---	--------------------	--------------------------------	--------------------	------------------------

**Aufgabe 3:** Die Zufallsvariable  $X$  hat eine Varianz von  $\sigma^2$ . Sie haben vor, eine Stichprobe unabhängiger und identisch verteilter Beobachtungen  $X_1, \dots, X_5$  zu ziehen. Wie berechnen Sie die Varianz von  $\frac{1}{4}(X_1 + X_2 + X_3 + X_4)$ ?

3:	<sup>a</sup> $4\sigma^2$	<sup>b</sup> $\sigma^2$	<sup>c</sup> $\sigma^2/2$	<sup>d</sup> $\sigma^2/4$	<sup>e</sup> $\sigma^2/16$
----	--------------------------	-------------------------	---------------------------	---------------------------	----------------------------

**Aufgabe 4:** Eine Zufallsvariable  $X$  ist wie folgt verteilt:  $P(X = A) = \theta, P(X = B) = 2\theta, P(X = C) = 1 - 3\theta$  wobei  $\theta \in [0, 1/3]$ . Eine Stichprobe ergibt zwei Beobachtungen:  $\{B, C\}$ .

Was ist der Maximum-Likelihood Schätzer für  $\theta$ ?

4:	<sup>a</sup> anderer Wert	<sup>b</sup> 0	<sup>c</sup> 1/6	<sup>d</sup> 1/3	<sup>e</sup> 1/2
----	---------------------------	----------------	------------------	------------------	------------------

**Aufgabe 5:** Eine Zufallsvariable  $X$  hat den Erwartungswert  $E(X) = 2\theta$  und die Varianz  $1/\theta$ . Es gilt  $\theta > 0$ . Die Variable  $x$  enthält Ihre Stichprobe. Wie berechnen Sie mit  $R$  den Momentenschätzer für  $\theta$  auf Basis des zweiten Moments?

5:	<sup>a</sup> anderer Wert	<sup>b</sup> $1/\text{mean}(x)$	<sup>c</sup> $1/\text{var}(x)$	<sup>d</sup> $\text{mean}(x)/2$	<sup>e</sup> $\text{sd}(x)$
----	---------------------------	---------------------------------	--------------------------------	---------------------------------	-----------------------------

**Aufgabe 6:** Eine Zufallsvariable  $X$  ist normalverteilt mit unbekanntem Mittelwert und unbekannter Standardabweichung. Die Variable  $x$  enthält Ihre Stichprobe

mit 12 Beobachtungen. Wie ermitteln Sie die Obergrenze des 95%-Konfidenzintervalls für den Mittelwert?

6a: Keine der folgenden Antworten ist richtig.

6b:  $\text{mean}(x) + \text{sd}(x) / \sqrt{12} * qt(.975, df=11)$

6c:  $\text{mean}(x) + \text{sd}(x) / 12 * qt(.975, df=11)$

6d:  $\text{mean}(x) + \text{sd}(x) / 12 * qnorm(.975)$

6e:  $\text{mean}(x) - \text{sd}(x) / \sqrt{12} * qt(.975, df=11)$

**Aufgabe 7:** Die Zufallsvariable  $X$  ist normalverteilt mit unbekanntem Mittelwert  $\mu$  und unbekannter Standardabweichung  $\sigma$ . Eine Stichprobe von 9 Beobachtungen ergibt einen Stichprobenmittelwert  $\bar{x} = 12$  und eine Standardabweichung der Stichprobenbeobachtungen von  $\hat{\sigma}_X = 15$ . Ihre Nullhypothese ist  $\mu = 2$ . Was ist der Absolutwert der Teststatistik?

7:	<sup>a</sup> anderer Wert	<sup>b</sup> 2	<sup>c</sup> 6	<sup>d</sup> 12	<sup>e</sup> 15
----	---------------------------	----------------	----------------	-----------------	-----------------

**Aufgabe 8:** Sie lehnen die Nullhypothese ab, wenn die Teststatistik kleiner als  $-g^*$  oder größer als  $+g^*$  ist. Wie groß ist  $g^*$  bei einem Signifikanzniveau von 5%?

8:	<sup>a</sup> anderer Wert	<sup>b</sup> 1.86	<sup>c</sup> 1.96	<sup>d</sup> 2.26	<sup>e</sup> 2.31
----	---------------------------	-------------------	-------------------	-------------------	-------------------

**Aufgabe 9:** Betrachten Sie das folgende Testergebnis. Verwenden Sie ein Signifikanzniveau von 5%.

One Sample t-test

data: x

t = 0.6982, df = 14, p-value = 0.4965

alternative hypothesis: true mean is not equal to 0  
95 percent confidence interval:

-0.3157490 0.6205177

sample estimates: mean of x 0.1523843

9a: Es wurde ein paarweiser Test durchgeführt.

9b: Die Alternativhypothese in diesem Test war  $E(X) \neq E(Y)$ .

9c: Die Nullhypothese  $E(X) = 0$  wird nicht abgelehnt.

9d: Die Nullhypothese  $E(X) = 1$  wird abgelehnt.

9e: Die Nullhypothese  $E(X) = 2$  wird abgelehnt.

**Aufgabe 10:** Betrachten Sie das folgende Testergebnis. Verwenden Sie ein Signifikanzniveau von 5%.

Wilcoxon signed rank test  
 data: x  
 V = 66, p-value = 0.7615  
 alternative hypothesis: true location shift  
 is not equal to 0

- 10a: Es wurde kein paarweiser Test durchgeführt.
- 10b: Die Alternativhypothese in diesem Test war  $E(X) \neq E(Y)$ .
- 10c: Die Nullhypothese  $E(X) = 0$  wird nicht abgelehnt.
- 10d: Die Nullhypothese  $E(X) = 1$  wird abgelehnt.
- 10e: Die Nullhypothese  $E(X) = 2$  wird abgelehnt.

**Aufgabe 11:** Sie untersuchen den Zusammenhang zwischen zwei Merkmalen,  $X$  und  $Y$ . Das Merkmal  $X$  kann die beiden Werte  $A$  und  $B$  annehmen.  $Y$  hat entweder den Wert  $C$ ,  $D$  oder  $E$ . Die Häufigkeit der Merkmale ist durch die folgende Tabelle gegeben:

	$Y = C$	$Y = D$	$Y = E$
$X = A$	12	23	14
$X = B$	18	16	13

Ihr Signifikanzniveau ist 5%. Interpretieren Sie das folgende Testergebnis:

Pearson's Chi-squared test

X-squared = 2.4528, df = 2, p-value = 0.2933

- 11a: Zwischen den beiden Merkmalen  $X$  und  $Y$  gibt es keinen signifikanten Zusammenhang.
- 11b: Die Nullhypothese,  $X$  und  $Y$  seien unabhängig, kann man zum gegebenen Signifikanzniveau verwerfen.
- 11c: Die Nullhypothese,  $X$  und  $Y$  seien voneinander abhängig, kann man nicht zum gegebenen Signifikanzniveau verwerfen.
- 11d: Die Nullhypothese,  $X$  und  $Y$  seien unabhängig, kann man nicht zum gegebenen Signifikanzniveau verwerfen.
- 11e: Um zu testen, ob  $X$  und  $Y$  voneinander unabhängig sind, sollte man besser den Korrelationskoeffizienten betrachten.

**Aufgabe 12:** Welche Wahrscheinlichkeit gibt der  $p$ -Wert an?

- 12a: Keine der folgenden Antworten ist richtig.
- 12b: Die Wahrscheinlichkeit eines Fehlers erster Art.
- 12c: Die Wahrscheinlichkeit eines Fehlers zweiter Art.
- 12d: Die Wahrscheinlichkeit die Alternativhypothese abzulehnen, wenn sie falsch ist.
- 12e: Die Wahrscheinlichkeit die Nullhypothese abzulehnen, wenn sie falsch ist.

**Aufgabe 13:** Sie verwenden eine lineare Regression um den folgenden Zusammenhang zu schätzen:

$$Y = 3 - 3X - 2D + u$$

$D$  ist eine Dummy-Variable und hat den Wert  $D = 1$  für treue Kunden und den Wert  $D = 0$  für andere Kunden. Jetzt kodieren Sie  $D$  anders: Treue Kunden haben den Wert  $D = 0$ , andere Kunden haben  $D = 1$ . Was ist nun der marginale Effekt von  $D$ ?

13:	a anderer Wert	b 0	c -3	d -2	e 2
-----	----------------	-----	------	------	-----

**Aufgabe 14:** In der Schätzung oben bezeichnet  $X$  das Einkommen der Kunden in €. Jetzt kodieren Sie  $X$  anders:  $X$

bezeichnet nun das Einkommen in Tausend €. Was ist jetzt der marginale Effekt von  $X$ ?

14:	a anderer Wert	b 1/3000	c -3	d -3000	e 3/1000
-----	----------------	----------	------	---------	----------

**Aufgabe 15:** Sie untersuchen den Effekt der Variablen  $X$  und  $Z$  auf die Variable  $Y$ . Die Variable  $X$  kann die Werte  $a$  und  $b$  annehmen, die Variable  $Z$  kann die Werte  $d$  und  $e$  annehmen. In Ihrer Regression codieren Sie  $a$  und  $d$  jeweils als Null und  $b$  und  $e$  jeweils als Eins. Die Erwartungswerte für  $Y$  sind in der folgenden Tabelle angegeben:

$X$	$Z$	$Y$
$a$	$d$	2
$a$	$e$	12
$b$	$d$	5
$b$	$e$	3

Sie schätzen die Gleichung

$$Y = \beta_0 + \beta_1 X + \beta_2 Z + \beta_3 XZ + u$$

Welchen Wert erhalten Sie für  $\beta_0$ ?

15:	a anderer Wert	b 2	c 3	d 5	e 10
-----	----------------	-----	-----	-----	------

Welchen Wert erhalten Sie für  $\beta_1$ ?

16:	a anderer Wert	b 2	c 3	d 5	e 10
-----	----------------	-----	-----	-----	------

Welchen Wert erhalten Sie für  $\beta_2$ ?

17:	a anderer Wert	b 2	c 3	d 5	e 10
-----	----------------	-----	-----	-----	------

Welchen Wert erhalten Sie für  $\beta_3$ ?

18:	a anderer Wert	b -2	c -3	d 3	e 5
-----	----------------	------	------	-----	-----

**Aufgabe 19:** Betrachten Sie das folgende Ergebnis einer Regression. Gehen Sie weiter von einem Signifikanzniveau von 5% aus. Nehmen Sie ferner an, dass die Residuen normalverteilt sind:

Call: lm(formula = y ~ x1 + x2)

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	1.596	1.711	0.932	0.3711
x1	3.627	2.045	1.774	0.1038
x2	3.893	1.748	2.227	0.0477 *

Residual standard error: 6.371 on 11 degrees of freedom

Multiple R-squared: 0.42, Adjusted R-squared: 0.3145

F-statistic: 3.983 on 2 and 11 DF, p-value: 0.04999

- 19a:  $x_1$  hat einen signifikanten Einfluss auf  $x_2$ .
- 19b:  $x_2$  hat einen signifikanten Einfluss auf  $y$ .
- 19c: Die Hypothese, der marginale Effekt von  $x_2$  auf  $y$  sei 8, wird abgelehnt.
- 19d: Die Hypothese, der marginale Effekt von  $x_1$  auf  $y$  sei 7, wird nicht abgelehnt.
- 19e: Das Modell erklärt 42% der Varianz von  $Y$ .

**Aufgabe 20:** Sie verwenden das folgende quadratische Modell, um den Zusammenhang zwischen der erklärenden Variablen  $X$  und der abhängigen Variablen  $Y$  zu schätzen:

$$Y = \beta_0 + \beta_1 X + \beta_2 X^2 + u$$

Sie erwarten folgendes: Für  $X = 1$  ist der marginale Effekt von  $X$  auf  $Y$  gleich 3. Für  $X = 2$  ist der marginale Effekt von  $X$  auf  $Y$  gleich 1.

Welchen Wert erwarten Sie für  $\beta_1$ ?

20:	<sup>a</sup> anderer Wert	<sup>b</sup> -1/2	<sup>c</sup> 1	<sup>d</sup> 2	<sup>e</sup> 5
-----	---------------------------	-------------------	----------------	----------------	----------------

Welchen Wert erwarten Sie für  $\beta_2$ ?

21:	<sup>a</sup> anderer Wert	<sup>b</sup> -1	<sup>c</sup> 1	<sup>d</sup> 2	<sup>e</sup> 3
-----	---------------------------	-----------------	----------------	----------------	----------------

## 11 Klausur BW24.1, Mai 2014

Ihre Version ist . Bitte vergessen Sie nicht, die Version in den Lösungsbogen zu übertragen. Lösen Sie die Aufgaben zunächst hier (und ggf. auf dem Schmierpapier) und übertragen Sie die Lösungen zum Schluss in den Lösungsbogen. Geben Sie nur den Lösungsbogen ab, und behalten Sie dieses Aufgabenblatt. Wir wünschen Ihnen viel Erfolg!

**Aufgabe 1:** Ihre Stichprobe enthält 5 unabhängige und identisch verteilte Beobachtungen:  $X_1, \dots, X_5$ . Welche Schätzfunktionen für  $E(X)$  sind erwartungstreu?

1:	<sup>a</sup> $\frac{X_2}{2} + \frac{X_3}{2}$	<sup>b</sup> $X_3$	<sup>c</sup> $X_1 + X_2 - X_3$	<sup>d</sup> $\frac{1}{5} \sum_{i=1}^5 X_i$	<sup>e</sup> $\sum_{i=1}^5 X_i$
----	--	--------------------	--------------------------------	---	---------------------------------

**Aufgabe 2:** Betrachten Sie weiter die obige Stichprobe. Die Varianz von  $X$  sei 90. Wie groß ist die Varianz von  $\frac{1}{3}(X_1 + X_5)$ ?

2:	<sup>a</sup> anderer Wert	<sup>b</sup> 20	<sup>c</sup> 30	<sup>d</sup> 60	<sup>e</sup> 90
----	---------------------------	-----------------	-----------------	-----------------	-----------------

**Aufgabe 3:** Welcher Schätzer für  $E(X)$  ist effizienter?

- 3a:  $\frac{X_2}{2} + \frac{X_3}{2}$  ist effizienter als  $X_3$   
 3b:  $X_3$  ist effizienter als  $X_1 + X_2 - X_3$   
 3c:  $\frac{1}{5} \sum_{i=1}^5 X_i$  ist effizienter als  $\sum_{i=1}^5 X_i$   
 3d:  $\frac{1}{5} \sum_{i=1}^5 X_i$  ist effizienter als  $X_3$   
 3e:  $\frac{X_2}{2} + \frac{X_3}{2}$  ist effizienter als  $X_3 + X_4 - X_5$

**Aufgabe 4:** Die Zufallsvariable  $X$  folgt einer Verteilung  $\mathcal{X}_\theta$  mit Erwartungswert  $E(X) = 1/\theta$  und Varianz  $\theta^2$ . Es gilt  $\theta > 0$ . Die Variable  $x$  enthält Ihre Stichprobe. Wie berechnen Sie mit  $R$  den Momentenschätzer für  $\theta$  auf Basis des zweiten Moments?

4:	<sup>a</sup> anderer Wert	<sup>b</sup> $1/\text{mean}(x)$	<sup>c</sup> $1/\text{sd}(x)$	<sup>d</sup> $\text{sd}(x)$	<sup>e</sup> $\text{sd}(x)/2$
----	---------------------------	---------------------------------	-------------------------------	-----------------------------	-------------------------------

**Aufgabe 5:** Eine Zufallsvariable ist wie folgt verteilt:  $P(X = 1) = \theta$ ,  $P(X = 2) = \theta$ ,  $P(X = 3) = 1 - 2\theta$  wobei  $\theta \in [0, 1/2]$ . Eine Stichprobe ergibt die Beobachtungen  $\{1, 2, 1, 1, 1, 1\}$ .

Was ist der Maximum-Likelihood Schätzer für  $\theta$ ?

5:	<sup>a</sup> anderer Wert	<sup>b</sup> 0	<sup>c</sup> 1/4	<sup>d</sup> 1/2	<sup>e</sup> 1
----	---------------------------	----------------	------------------	------------------	----------------

**Aufgabe 6:** Der Vektor  $x$  enthält Ihre Stichprobe mit 9 Beobachtungen,  $s$  ist die geschätzte Standardabweichung von  $x$ . Sie gehen davon aus, dass  $x$  einer Normalverteilung folgt. Wie berechnen Sie die Breite des 99%-Konfidenzintervalls für den Mittelwert von  $x$ ?

- 6a: Keine der folgenden Antworten ist richtig.  
 6b:  $2*s*qt(.995, df=8)/3$   
 6c:  $2*s*qt(.99, df=8)/3$   
 6d:  $2*s*qnorm(.995)/3$   
 6e:  $2*s*qnorm(.99)/3$

**Aufgabe 7:** Sie betrachten weiter die obige Stichprobe. Ihre Nullhypothese ist  $E(x) \geq 7$ , Ihre Alternativhypothese ist  $E(x) < 7$ . Ihre Teststatistik haben Sie mit dem Kommando  $t <- (\text{mean}(x)-7)/(s/3)$

berechnet, wobei  $\text{mean}(x)$  den Mittelwert Ihrer Stichprobe enthält. In welchem Bereich lehnen Sie die Nullhypothese bei einem Signifikanzniveau von 1% ab?

7:	<sup>a</sup> anderer Wert	<sup>b</sup> $t < -3.36$	<sup>c</sup> $t < -2.9$	<sup>d</sup> $t < 2.9$	<sup>e</sup> $t < 3.36$
----	---------------------------	--------------------------	-------------------------	------------------------	-------------------------

**Aufgabe 8:** Sie testen weiter die Nullhypothese aus der obigen Aufgabe. Wie bestimmen Sie den  $p$ -Wert?

- 8a: Keine der folgenden Antworten ist richtig.  
 8b:  $2*pt(t, df=8)$   
 8c:  $1-pt(t, df=8)$   
 8d:  $pt(-t, df=8)$   
 8e:  $pt(t, df=8)$

**Aufgabe 9:** Sie wollen die Alternativhypothese testen, die Varianz  $\hat{\sigma}^2$  Ihrer Stichprobe sei signifikant größer als ein nullhypothetischer Wert  $\sigma_0^2$ . Mit der Formel  $g = (n-1)\hat{\sigma}^2/\sigma_0^2$  bestimmen Sie Ihre Teststatistik  $g$ . Sie wissen, dass Ihre Teststatistik einer  $\chi^2$  Verteilung mit  $n-1$  Freiheitsgraden folgt. In Ihrem Fall ist  $n = 6$ . Wie bestimmen Sie den  $p$ -Wert?

- 9a: Keine der folgenden Antworten ist richtig.  
 9b:  $1-qchisq(g, df=5)$   
 9c:  $qchisq(g, df=5)$   
 9d:  $1-pchisq(g, df=5)$   
 9e:  $pchisq(g, df=5)$

**Aufgabe 10:** Betrachten Sie das folgende Testergebnis. Verwenden Sie ein Signifikanzniveau von 5%.

```
Welch Two Sample t-test
data: x and y
t = -0.2718, df = 14.731, p-value = 0.7896
alternative hypothesis: true difference in means
is not equal to 0
95 percent confidence interval:
-20.37493 15.77302
sample estimates:
mean of x mean of y
28.42173 30.72268
```

- 10a: Es wurde ein paarweiser Test durchgeführt.  
 10b: Die Alternativhypothese in diesem Test war  $E(X) - E(Y) \neq 0$ .  
 10c: Die Nullhypothese  $E(X) - E(Y) \geq 16$  wird abgelehnt.  
 10d: Die Nullhypothese  $E(X) - E(Y) \leq 16$  wird abgelehnt.  
 10e: Die Nullhypothese  $E(X) - E(Y) \geq -16$  wird abgelehnt.

**Aufgabe 11:** Betrachten Sie das folgende Testergebnis. Verwenden Sie ein Signifikanzniveau von 5%.

```
Wilcoxon rank sum test
data: x and y
W = 87, p-value = 0.01839
alternative hypothesis: true location shift
is not equal to 0
```

- 11a: Es wurde ein zweiseitiger Test durchgeführt.

- 11b: Die Hypothese,  $x$  und  $y$  stammen aus der gleichen Verteilung, wird nicht abgelehnt.  
 11c: Die Hypothese  $E(X) - E(Y) \leq 16$  kann verworfen werden.  
 11d: Die Hypothese  $E(X) - E(Y) \geq 16$  kann verworfen werden.  
 11e: Der Test setzt nicht voraus, dass  $X$  und  $Y$  einer Normalverteilung folgen.

**Aufgabe 12:** Die Besucher einer Vorlesung lassen sich wie folgt verschiedenen Studienrichtungen zuordnen:

	BWL	VWL	WiPäd
Anzahl	20	40	20

Sie erwarten, dass von allen Studierenden 25% BWL, 25% VWL, und der Rest WiPäd studieren. Welchen Test benutzen Sie, um die von Ihnen beobachteten Häufigkeiten mit den von Ihnen erwarteten Häufigkeiten zu vergleichen?

- 12a: Keine der folgenden Antworten ist richtig.  
 12b: Einen paarweisen  $t$ -Test  
 12c: Einen  $t$ -Test für unverbundene Stichproben  
 12d: Einen  $\chi^2$  Anpassungstest  
 12e: Einen Rangsummentest

**Aufgabe 13:** In der folgenden Regression schätzen Sie den Effekt der Variablen  $a$  und  $b$  auf die Variable  $Y$ . Die Variablen  $a$  und  $b$  können jeweils die Werte 0 oder 1 annehmen, Sie erhalten den folgenden Output:

```
Call: lm(formula = y ~ a * b)
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  4.000      1.414    3.036  0.0162 *
a             -2.000      2.000   -1.000  0.3466
b              1.000      2.000    0.500  0.6305
a:b           1.000      2.828    0.354  0.7328
Residual standard error: 2.449 on 8 degrees of freedom
Multiple R-squared:  0.2289
F-statistic: 0.7917 on 3 and 8 DF, p-value: 0.5319
```

Welchen Wert für  $Y$  erwarten Sie, wenn  $a = 0$  und  $b = 0$ ?

13: 

a anderer Wert	b	c	d	e
	-2	2	3	4

Welchen Wert für  $Y$  erwarten Sie, wenn  $a = 1$  und  $b = 0$ ?

14: 

a anderer Wert	b	c	d	e
	-2	2	3	4

Welchen Wert für  $Y$  erwarten Sie, wenn  $a = 1$  und  $b = 1$ ?

15: 

a anderer Wert	b	c	d	e
	-2	2	3	4

Wie groß ist der marginale Effekt von  $b$  auf  $Y$  falls  $a = 1$ ?

16: 

a anderer Wert	b	c	d	e
	0	1	2	4

Wie groß ist der marginale Effekt von  $a$  auf  $Y$  falls  $b = 0$ ?

17: 

a anderer Wert	b	c	d	e
	-2	-1	0	1

**Aufgabe 18:** Wie bestimmen Sie die Untergrenze des 95%-Konfidenzintervalls für den Koeffizienten von  $a$ ?

- 18a: Keine der folgenden Antworten ist richtig.  
 18b:  $2 - 2 * qt(.975, df=8)$   
 18c:  $-2 * qt(.975, df=8)$   
 18d:  $-2 - 2 * qt(.975, df=8)$   
 18e:  $-2 + 2 * qt(.975, df=8)$

**Aufgabe 19:** Sie schätzen den Zusammenhang zwischen der erklärenden Variablen  $X$  und der abhängigen Variablen  $Y$  mit einem Polynom  $Y = \sum_{k=0}^r \beta_k X^k + u$ . Für verschiedene Werte von  $r$  erhalten Sie die folgenden Werte für das AIC:

$r$	1	2	3	4	5	6
AIC	124.78	125.44	126.94	127.18	129.10	131.04

Welchen Wert von  $r$  sollten Sie auf Basis des AIC wählen?

19: 

a anderer Wert	b	c	d	e
	1	2	3	6

**Aufgabe 20:** Sie verwenden das folgende quadratische Modell, um den Zusammenhang zwischen der erklärenden Variablen  $X$  und der abhängigen Variablen  $Y$  zu schätzen:

$$Y = \beta_0 + \beta_1 X + \beta_2 X^2 + u$$

Sie erwarten folgendes: Für  $X = 0$  ist der marginale Effekt von  $X$  auf  $Y$  gleich  $-3$ . Für  $X = -1$  ist der marginale Effekt von  $X$  auf  $Y$  gleich 1.

Welchen Wert erwarten Sie für  $\beta_1$ ?

20: 

a anderer Wert	b	c	d	e
	-3	-1	0	1

Welchen Wert erwarten Sie für  $\beta_2$ ?

21: 

a anderer Wert	b	c	d	e
	-3	-2	-1	0

## 12 Klausur BW24.1, Februar 2014

TEST-13

Ihre Version ist . Bitte vergessen Sie nicht, die Version in den Lösungsbogen zu übertragen. Lösen Sie die Aufgaben zunächst hier (und ggf. auf dem Schmierpapier) und übertragen Sie die Lösungen zum Schluss in den Lösungsbogen. Geben Sie den Lösungsbogen ab, und behalten Sie dieses Aufgabenblatt. Wir wünschen Ihnen viel Erfolg!

**Aufgabe 1:** Ihre Stichprobe enthält 4 unabhängige und identisch verteilte Beobachtungen:  $X_1, \dots, X_4$ . Welche Schätzfunktionen für  $E(X)$  sind erwartungstreu?

- 1a:  $\frac{1}{4} \sum_{i=1}^4 \sqrt{X_i^2}$   
 1b:  $\frac{1}{4} \sum_{i=1}^4 \sqrt[3]{X_i^3}$   
 1c:  $\frac{1}{3} X_1 + \frac{2}{3} X_2$

- 1d:  $\frac{1}{3} X_1 + \frac{1}{3} X_3$   
 1e:  $X_1 - X_2$

**Aufgabe 2:** Betrachten Sie weiter die obige Stichprobe. Die Varianz von  $X$  sei 10. Wie groß ist die Varianz von  $X_1 + 2X_2 - X_4$ ?

2: 

a anderer Wert	b	c	d	e
	10	40	50	60

**Aufgabe 3:** Die Zufallsvariable  $X$  folgt einer Verteilung  $\mathcal{Z}_\theta$  mit Erwartungswert  $E(X) = 2\theta$  und Varianz  $\theta^2$ . Die Variable  $x$  enthält Ihre Stichprobe. Wie berechnen Sie mit R den Momentenschätzer für  $\theta$  auf Basis des ersten Moments?

3: 

a anderer Wert	b	c	d	e
	$2/\text{mean}(x)$	$\text{mean}(x)/2$	$1/\text{mean}(x)$	$\text{var}(x)/2$

**Aufgabe 4:** Eine Zufallsvariable ist wie folgt verteilt:  
 $P(X = -1) = \theta$ ,  $P(X = 1) = \theta$ ,  $P(X = 0) = 1 - 2\theta$ . Eine Stichprobe ergibt die Beobachtungen  $\{-1, 0\}$ :  
 Was ist der Maximum-Likelihood Schätzer für  $\theta$ ?

4:	a	b	c	d	e
	anderer Wert	0	-1/4	1/4	1/2

**Aufgabe 5:** Der Vektor  $x$  enthält Ihre Stichprobe mit 16 Beobachtungen,  $s$  ist die bekannte Standardabweichung von  $x$ . Sie gehen davon aus, dass  $x$  einer Normalverteilung folgt. Wie berechnen Sie die Breite des 95%-Konfidenzintervalls für den Mittelwert von  $x$ .

- 5a: Keine der folgenden Antworten ist richtig.
- 5b:  $2*s*qt(.975)/16$
- 5c:  $2*s*qt(.975)/4$
- 5d:  $2*s*pnorm(.975)/16$
- 5e:  $2*s*pnorm(.975)/4$

**Aufgabe 6:** Sie wollen testen, ob der Mittelwert  $\bar{x}$  Ihrer Stichprobe mit 10 normalverteilten Beobachtungen signifikant kleiner als ein nullhypothetischer Wert  $\mu_0$  ist. Mit der Formel  $g = (\bar{x} - \mu_0) / \hat{\sigma}_{\bar{x}}$  bestimmen Sie Ihre Teststatistik  $g$ . Wie bestimmen Sie den  $p$ -Wert?

- 6a: Keine der folgenden Antworten ist richtig.
- 6b:  $1-qt(g, df=9)$
- 6c:  $qt(g, df=9)$
- 6d:  $1-pt(g, df=9)$
- 6e:  $pt(g, df=9)$

**Aufgabe 7:** Sie wollen testen, ob die Varianz  $\hat{\sigma}$  Ihrer Stichprobe signifikant größer als ein nullhypothetischer Wert  $\sigma_0^2$  ist. Mit der Formel  $g = (n - 1) \hat{\sigma}^2 / \sigma_0^2$  bestimmen Sie Ihre Teststatistik  $g$ . Sie wissen, dass Ihre Teststatistik einer  $\chi^2$  Verteilung mit  $n - 1$  Freiheitsgraden folgt. In Ihrem Fall ist  $n = 10$ . Wie bestimmen Sie den  $p$ -Wert?

- 7a: Keine der folgenden Antworten ist richtig.
- 7b:  $1-qchisq(g, df=9)$
- 7c:  $qchisq(g, df=9)$
- 7d:  $1-pchisq(g, df=9)$
- 7e:  $pchisq(g, df=9)$

**Aufgabe 8:**

Betrachten Sie das folgende Testergebnis. Verwenden Sie ein Signifikanzniveau von 5%.

```
Welch Two Sample t-test
data: x and y
t = 4.4379, df = 15.892, p-value = 0.0004198
alternative hypothesis: true difference in means
is not equal to 0
95 percent confidence interval:
14.25951 40.36870
sample estimates:
mean of x mean of y
57.34573 30.03163
```

- 8a: Es wurde ein paarweiser Test durchgeführt.
- 8b: Die Alternativhypothese in diesem Test war  $E(X) - E(Y) > 0$ .
- 8c: Die Nullhypothese  $E(X) - E(Y) \leq 7$  wird abgelehnt.
- 8d: Die Nullhypothese  $E(X) - E(Y) \geq 7$  wird abgelehnt.
- 8e: Die Nullhypothese  $E(X) - E(Y) \geq -7$  wird abgelehnt.

**Aufgabe 9:** Betrachten Sie das folgende Testergebnis. Verwenden Sie ein Signifikanzniveau von 5%.

```
wilcox.test(x,y)
Wilcoxon rank sum test
data: x and y
W = 22, p-value = 0.1135
alternative hypothesis: true location shift
is not equal to 0
```

- 9a: Es wurde ein zweiseitiger Test durchgeführt.
- 9b: Die Hypothese,  $x$  und  $y$  folgen der gleichen Verteilung, wird nicht abgelehnt.
- 9c: Die Hypothese  $E(X) - E(Y) \leq 7$  kann verworfen werden.
- 9d: Die Hypothese  $E(X) - E(Y) \geq 7$  kann verworfen werden.
- 9e: Der Test setzt nicht voraus, dass  $X$  und  $Y$  einer Normalverteilung folgen.

**Aufgabe 10:** Sie verkaufen die Produkte A und B an Kunden aus verschiedenen Städten. Die Anzahl der verkauften Produkte ist wie folgt:

	Erfurt	Gera	Jena
A	17	12	8
B	27	12	4

Welchen Test benutzen Sie, um zu testen, ob die Vorlieben Ihrer Kunden für A oder B von der Stadt abhängen?

- 10a: Keine der folgenden Antworten ist richtig.
- 10b: Einen paarweisen  $t$ -Test
- 10c: Einen  $t$ -Test für unverbundene Stichproben
- 10d: Einen  $\chi^2$  Kontingenztest
- 10e: Einen Rangsummentest

**Aufgabe 11:**

In der folgenden Regression schätzen Sie den Effekt der Variablen  $a$  und  $b$  auf die Variable  $Y$ . Die Variablen  $a$  und  $b$  können jeweils die Werte 0 oder 1 annehmen, Sie erhalten den folgenden Output:

```
Call: lm(formula = y ~ a * b)
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 3.0000    0.7071    4.243 0.000148 ***
a            2.0000    1.0000    2.000 0.053088 .
b            1.0000    1.0000    1.000 0.323985
a:b          2.0000    1.4142    1.414 0.165890
Residual standard error: 2.236 on 36 degrees of freedom
Multiple R-squared: 0.4375
F-statistic: 9.333 on 3 and 36 DF, p-value: 0.0001061
```

Welchen Wert für  $Y$  erwarten Sie, wenn  $a = 0$  und  $b = 0$ ?

11:	a	b	c	d	e
	anderer Wert	1	2	3	5

Welchen Wert für  $Y$  erwarten Sie, wenn  $a = 1$  und  $b = 0$ ?

12:	a	b	c	d	e
	anderer Wert	1	2	3	5

Welchen Wert für  $Y$  erwarten Sie, wenn  $a = 1$  und  $b = 1$ ?

13:	a	b	c	d	e
	anderer Wert	2	3	5	8

Wie groß ist der marginale Effekt von  $b$  auf  $Y$  falls  $a = 1$ ?

14:	a	b	c	d	e
	anderer Wert	1	2	3	5

Wie groß ist der marginale Effekt von  $a$  auf  $Y$  falls  $b = 0$ ?

15:	a	b	c	d	e
	anderer Wert	1	2	3	5

**Aufgabe 16:**

Wie bestimmen Sie die Untergrenze des 95%-Konfidenzintervalls für den Koeffizienten von a?

- 16a: Keine der folgenden Antworten ist richtig.  
 16b:  $2 - 2 * qt(.975, df=36)$   
 16c:  $1 - qt(.975, df=36)$   
 16d:  $2 - qt(.975, df=36)$   
 16e:  $2 + qt(.975, df=36)$

**Aufgabe 17:** Sie schätzen den Zusammenhang zwischen der erklärenden Variablen  $X$  und der abhängigen Variablen  $Y$  mit einem Polynom  $Y = \sum_{k=0}^r \beta_k X^k + u$ . Für verschiedene Werte von  $r$  erhalten Sie die folgenden Werte für das AIC:

$r$	1	2	3	4	5	6
AIC	8.52	0.41	-7.70	-5.84	-3.84	-2.10

Welchen Wert von  $r$  sollten Sie auf Basis des AIC wählen?

17: 

a) anderer Wert	b) 1	c) 2	d) 3	e) 6
-----------------	------	------	------	------

**Aufgabe 18:**

Sie verwenden das folgende quadratisches Modell um den Zusammenhang zwischen der erklärenden Variablen  $X$  und der abhängigen Variablen  $Y$  zu schätzen:

$$Y = \beta_0 + \beta_1 X + \beta_2 X^2 + u$$

Sie erwarten folgendes: Für  $X = 0$  ist der marginale Effekt von  $X$  auf  $Y$  gleich 5. Für  $X = 1$  ist der marginale Effekt von  $X$  auf  $Y$  gleich 7.

Welchen Wert erwarten Sie für  $\beta_1$ ?

18: 

a) anderer Wert	b) -5	c) 2	d) 5	e) 7
-----------------	-------	------	------	------

Welchen Wert erwarten Sie für  $\beta_2$ ?

19: 

a) anderer Wert	b) 1/2	c) 1	d) 2	e) 9
-----------------	--------	------	------	------

### 13 Klausur BW24.1, Mai 2013

**Aufgabe 1:**

Ihre Stichprobe enthält 4 unabhängige und identisch verteilte Beobachtungen:  $X_1, \dots, X_4$ . Welche Schätzfunktionen für  $E(X)$  sind erwartungstreu?

- 1a:  $(X_1 + X_2 + X_3 + X_4)/2$ .  
 1b:  $(X_1 + X_2 + X_3 + X_4)/4$ .  
 1c:  $(X_3 + X_4)/2$ .  
 1d:  $(X_4 - X_3)/2$ .  
 1e:  $X_1 - X_2 + X_3$ .

**Aufgabe 2:** Welche Schätzfunktionen für  $E(X)$  sind effizienter?

- 2a:  $(X_1 - X_4)/2$  ist effizienter als  $X_2$ .  
 2b:  $\sum_{i=1}^4 X_i/4$  ist effizienter als  $(X_1 + X_2)/2$ .  
 2c:  $X_3$  ist effizienter als  $X_2 + X_3 - X_4$ .  
 2d:  $2X_3 - X_2$  ist effizienter als  $X_2$ .  
 2e:  $X_1 + X_2 + 2X_3 - 3X_4$  ist effizienter als  $2X_3 - X_2$

**Aufgabe 3:** Ein Taxifahrer untersucht, ob seine Kundschaft eher aus Männern oder aus Frauen besteht. Er codiert Männer als 0 und Frauen als 1 und nennt diese Zufallsvariable  $X$ . Er zieht nun eine Stichprobe der Größe 4 (d.h. er betrachtet vier Fahrten) und stellt fest, dass er in allen vier Fällen eine Frau gefahren hat. Was ist die beste erwartungstreue Schätzung für  $\mu_X$ ?

3: 

a) anderer Wert	b) 0	c) 1/2	d) 1	e) 4
-----------------	------	--------	------	------

**Aufgabe 4:** Was ist die Standardabweichung Ihres Schätzers für  $\mu_X$ ?

4: 

a) anderer Wert	b) 0	c) 1/4	d) 1/2	e) 1
-----------------	------	--------	--------	------

**Aufgabe 5:** Der Taxifahrer geht nun davon aus, dass seine Zufallsvariable  $X$  mit Wahrscheinlichkeit  $\theta$  den Wert 0 und mit Wahrscheinlichkeit  $1 - \theta$  den Wert 1 hat. Was ist der Maximum-Likelihood Schätzer für  $\theta$ ?

5: 

a) anderer Wert	b) 0	c) 1/2	d) $\sqrt{3/4}$	e) 1
-----------------	------	--------	-----------------	------

**Aufgabe 6:** Für die Zufallsvariable  $X$  gilt  $E[X|\theta] = 1 - \theta$ . Was ist der Momentenschätzer für  $\theta$  auf Basis des ersten Moments?

6: 

a) anderer Wert	b) 0	c) 1/2	d) $\sqrt{3/4}$	e) 1
-----------------	------	--------	-----------------	------

**Aufgabe 7:** Ein anderer Taxifahrer nimmt ebenfalls eine Stichprobe von 4 Beobachtungen und ermittelt einen Mittelwert der Stichprobe von  $\bar{x} = 1/2$  mit einer Standardabweichung von  $\hat{\sigma}_x = 1/2$ . Wie kann er die Breite des 95%-Konfidenzintervalls für  $\bar{x}$  bestimmen?

- 7a: Keine der folgenden Antworten ist richtig.  
 7b: Weil  $X$  nicht normalverteilt und die Stichprobe klein ist, kann man hier weder mit qt noch mit qnorm arbeiten.  
 7c: Weil  $X$  nicht normalverteilt ist, kann man hier nicht mit qt sondern nur mit qnorm arbeiten.  
 7d: Weil  $X$  nicht normalverteilt ist, kann man hier nicht mit qnorm sondern nur mit qt arbeiten.  
 7e:  $2 * qt(.975, df=3)/4$

**Aufgabe 8:** Betrachten Sie das folgende Testergebnis. Verwenden Sie ein Signifikanzniveau von 5%.

```
Welch Two Sample t-test
data: x and y
t = -0.7614, df = 15.892, p-value = 0.2288
alternative hypothesis: true difference in
means is less than 0
95 percent confidence interval:
-Inf 6.063979
sample estimates:
mean of x mean of y
25.34573 30.03163
```

- 8a: Es wurde ein paarweiser Test durchgeführt.  
 8b: Die Alternativhypothese in diesem Test war  $E(X) - E(Y) > 0$ .  
 8c: Die Nullhypothese  $E(X) - E(Y) \leq 7$  wird abgelehnt.  
 8d: Die Nullhypothese  $E(X) - E(Y) \geq 7$  wird abgelehnt.  
 8e: Die Nullhypothese  $E(X) - E(Y) \geq -7$  wird abgelehnt.

**Aufgabe 9:** Betrachten Sie das folgende Testergebnis. Verwenden Sie ein Signifikanzniveau von 5%.

```
wilcox.test(x,y,paired=TRUE)
Wilcoxon signed rank test
data: x and y
V = 16, p-value = 0.4961
alternative hypothesis: true location shift
is not equal to 0
```

- 9a: Es wurde ein zweiseitiger Test durchgeführt.
- 9b: Die Hypothese, x und y folgen der gleichen Verteilung, wird nicht abgelehnt.
- 9c: Die Hypothese  $E(X) - E(Y) \leq 7$  kann verworfen werden.
- 9d: Der Test setzt voraus, dass X und Y einer Normalverteilung folgen.
- 9e: Für x und y liegen gleich viele Beobachtungen vor.

**Aufgabe 10:** In einem Versuch untersuchen Sie den Effekt der Variablen  $X_1$  und  $X_2$  auf die Variable Y. Die Variable  $X_1$  kann die Werte A und B annehmen,  $X_2$  kann die Werte C und D annehmen. In Ihrer Regression codieren Sie jeweils A und C als Null und B bzw. D als Eins. Sie erhalten den folgenden Output:

```
lm(formula = y ~ x1 * x2)
      Estimate Std. Error t value Pr(>|t|)
(Intercept)  3.0000    0.7071   4.243 6.17e-05 ***
x1           -2.0000    1.0000  -2.000 0.04907 *
x2           -3.0000    1.0000  -3.000 0.00365 **
x1:x2         4.0000    1.4142   2.828 0.00598 **
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 3.162 on 76 degrees of freedom
Multiple R-squared:  0.1163
F-statistic: 3.333 on 3 and 76 DF, p-value: 0.02381
Welchen Wert für Y erwarten Sie, wenn  $X_1 = A$  und  $X_2 = C$ ?

```

10:	a	b	c	d	e
	anderer Wert	-3	-2	3	4

**Aufgabe 11:** Welchen Wert für Y erwarten Sie, wenn  $X_1 = A$  und  $X_2 = D$ ?

11:	a	b	c	d	e
	anderer Wert	-3	-2	0	3

**Aufgabe 12:** Welchen Wert für Y erwarten Sie, wenn  $X_1 = B$  und  $X_2 = D$ ?

12:	a	b	c	d	e
	anderer Wert	-3	-2	2	4

**Aufgabe 13:** Wie groß ist der marginale Effekt von  $X_1$  auf Y falls  $X_2 = C$ ?

13:	a	b	c	d	e
	anderer Wert	-3	-2	2	4

**Aufgabe 14:** Wie groß ist der marginale Effekt von  $X_1$  auf Y falls  $X_2 = D$ ?

14:	a	b	c	d	e
	anderer Wert	-3	-2	2	4

**Aufgabe 15:** Wie bestimmen Sie die Untergrenze des 95%-Konfidenzintervalls für den Koeffizienten von  $x_2$ ?

- 15a: Keine der folgenden Antworten ist richtig.
- 15b:  $-3 - qt(.05, df=76)$
- 15c:  $-3 + qt(.975, df=76)$
- 15d:  $-3 - qt(.95, df=76)$
- 15e:  $-3 - qt(.975, df=76)$

**Aufgabe 16:** Betrachten Sie weiter den Output der Regression aus Aufgabe 10. Gehen Sie weiter von einem Signifikanzniveau von 5% aus:

- 16a:  $X_1$  hat einen signifikanten Einfluss auf Y wenn  $X_2 = C$ .
- 16b:  $X_2$  hat einen signifikanten Einfluss auf Y wenn  $X_1 = A$ .
- 16c: Die Hypothese, der marginale Effekt von  $X_1$  auf Y sei  $-3$  wenn  $X_2 = C$ , wird abgelehnt.
- 16d: Die Hypothese, der marginale Effekt von  $X_1$  auf Y sei  $-4$  wenn  $X_2 = C$ , wird abgelehnt.
- 16e: Ihre Stichprobe enthält 77 Beobachtungen

**Aufgabe 17:** Über den Zusammenhang zwischen X und Y wissen Sie folgendes:

- 17a: Wenn X um 1 steigt, dann steigt Y um  $-3.4$ .
- 17b: Wenn X um 1% steigt, dann steigt Y um  $-0.6$ .
- 17c: Wenn X um 1 steigt, dann steigt Y um  $-8.3\%$ .
- 17d: Wenn X um 1% steigt, dann steigt Y um  $-1.4\%$ .

Welchen Wert für  $\beta_1$  schätzen Sie in  $Y = \beta_0 + \beta_1 \cdot X + u$ ?

17:	a	b	c	d	e
	anderer Wert	-3.4	-0.6	-8.3	-1.4

**Aufgabe 18:** Welchen Wert für  $\beta_1$  schätzen Sie in  $Y = \beta_0 + \beta_1 \cdot \log X + u$ ?

18:	a	b	c	d	e
	anderer Wert	-60	8.3	-0.6	-0.083

**Aufgabe 19:** Welchen Wert für  $\beta_1$  schätzen Sie in  $\log Y = \beta_0 + \beta_1 \cdot X + u$ ?

19:	a	b	c	d	e
	anderer Wert	-60	8.3	-0.6	-0.083

**Aufgabe 20:** Welchen Wert für  $\beta_1$  schätzen Sie in  $\log Y = \beta_0 + \beta_1 \cdot \log X + u$ ?

20:	a	b	c	d	e
	anderer Wert	-140	-3.4	-1.4	-0.014

## 14 Klausur BW24.1, Februar 2013

Ihre Version ist . Bitte vergessen Sie nicht, die Version in den Lösungsbogen zu übertragen. Lösen Sie die Aufgaben zunächst hier (und ggf. auf dem Schmierpapier) und übertragen Sie die Lösungen zum Schluss in den Lösungsbogen. Geben Sie den Lösungsbogen ab, und behalten Sie dieses Aufgabenblatt. Wir wünschen Ihnen viel Erfolg!

**Aufgabe 1:** Ihre Stichprobe enthält 5 unabhängige und identisch verteilte Beobachtungen:  $X_1, \dots, X_5$ . Welche Schätzfunktionen für  $E(X)$  sind erwartungstreu?

- 1a:  $(X_1 + X_2 + 2X_5 - X_3 + X_4)/2$ .
- 1b:  $(X_1 + X_5)/2$ .
- 1c:  $X_2$ .
- 1d:  $2X_1 - X_2 + X_3 - X_4$ .
- 1e:  $3X_1 - X_2 - X_3$ .

**Aufgabe 2:** Welche Schätzfunktionen für  $E(X)$  sind effizienter?

- 2a:  $(X_1 + X_5)/2$  ist effizienter als  $X_2$ .  
 2b:  $\sum_{i=1}^5 X_i/5$  ist effizienter als  $(X_1 + X_2 + 2X_5 - X_3 + X_4)/2$ .  
 2c:  $\sum_{i=1}^5 X_i/5$  ist effizienter als  $(X_1 + X_5)/2$ .  
 2d:  $X_3$  ist effizienter als  $\sum_{i=1}^5 X_i/5$ .  
 2e:  $2X_1 - X_2 + X_3 - X_4$  ist effizienter als  $X_2$ .

**Aufgabe 3:** Die Zufallsvariable  $X$  hat Mittelwert  $\mu$  und Varianz 100. Sie ziehen eine Stichprobe der Größe 25 und stellen fest, dass  $\sum_{i=1}^{25} X_i = 2500$ .

Was ist die beste erwartungstreue Schätzung für  $\mu$ ?

3:	a anderer Wert	b 10	c 100	d 500	e 2500
----	----------------	------	-------	-------	--------

**Aufgabe 4:** Was ist die Standardabweichung Ihres Schätzers für  $\mu$ ?

4:	a anderer Wert	b 2	c 4	d 10	e 20
----	----------------	-----	-----	------	------

**Aufgabe 5:** Die Zufallsvariable  $X$  nimmt mit Wahrscheinlichkeit  $\theta$  den Wert  $A$  und mit Wahrscheinlichkeit  $1 - \theta$  den Wert  $B$  an. In Ihrer Stichprobe beobachten Sie  $\{A, B, B\}$ . Was ist der Maximum-Likelihood Schätzer für  $\theta$ ?

5:	a 0	b 1/3	c 1/2	d 2/3	e 1
----	-----	-------	-------	-------	-----

**Aufgabe 6:** Für die Zufallsvariable  $Y$  gilt  $E[Y|\theta] = 4 - 3\theta$ . In Ihrer Stichprobe beobachten Sie für  $Y$  die Werte  $\{1, 2, 3\}$ . Was ist der Momentenschätzer für  $\theta$  auf Basis des ersten Moments?

6:	a -2/3	b 1/3	c 2/3	d 1	e 4/3
----	--------	-------	-------	-----	-------

**Aufgabe 7:** Aus einer Population normalverteilter Beobachtungen mit unbekanntem Mittelwert  $\mu$  und unbekannter Standardabweichung  $\sigma$  ziehen Sie eine Stichprobe mit 9 Beobachtungen. Der Mittelwert dieser Beobachtungen ist  $\bar{x} = 3$ . Die Standardabweichung der Beobachtungen ist  $\hat{\sigma}_x = 4$ . Mit welchen Kommandos bestimmen Sie die Breite Ihres 95%-Konfidenzintervalls?

- 7a: Keine der folgenden Antworten ist richtig.  
 7b:  $2 * qt(.975, df=8) * 4/9$   
 7c:  $2 * qt(.975, df=8) * 4/3$   
 7d:  $(-2) * qnorm(.025) * 3/4$   
 7e:  $2 * qnorm(.95) * 2/3$

**Aufgabe 8:** Betrachten Sie das folgende Testergebnis. Verwenden Sie ein Signifikanzniveau von 5%.

```
Welch Two Sample t-test
data: x and y
t = -0.5058, df = 15.583, p-value = 0.69
alternative hypothesis: true difference in
means is greater than 0
95 percent confidence interval:
-11.00789 Inf
sample estimates:
mean of x mean of y
28.2340 30.7037
```

- 8a: Es wurde ein paarweiser Test durchgeführt.  
 8b: Die Alternativhypothese in diesem Test war  $E(X) - E(Y) > 0$ .  
 8c: Die Nullhypothese  $E(X) - E(Y) \geq 12$  wird abgelehnt.  
 8d: Die Nullhypothese  $E(X) - E(Y) \leq 12$  wird abgelehnt.

8e: Die Nullhypothese  $E(X) - E(Y) \leq -12$  wird abgelehnt.

**Aufgabe 9:** Betrachten Sie das folgende Testergebnis. Verwenden Sie ein Signifikanzniveau von 5%.

```
Wilcoxon rank sum test
data: x and y
W = 3, p-value = 0.0002879
alternative hypothesis: true location shift
is not equal to 0
```

- 9a: Es wurde ein zweiseitiger Test durchgeführt.  
 9b: Die Hypothese,  $x$  und  $y$  folgen der gleichen Verteilung, wird nicht abgelehnt.  
 9c: Die Hypothese,  $E(X) - E(Y) \leq 12$ , kann verworfen werden.  
 9d: Der Test setzt nicht voraus, dass  $X$  und  $Y$  einer Normalverteilung folgen.  
 9e: Wenn  $X$  und  $Y$  nicht normalverteilt sind, kann der Test nur bei einer sehr großen Anzahl an Beobachtungen angewendet werden.

**Aufgabe 10:** In einem Versuch untersuchen Sie den Effekt der Variablen  $X_1$  und  $X_2$  auf die Variable  $Y$ . Die Variable  $X_1$  kann die Werte  $A$  und  $B$  annehmen,  $X_2$  kann die Werte  $C$  und  $D$  annehmen. In Ihrer Regression codieren Sie jeweils  $A$  und  $C$  als Null und  $B$  bzw.  $D$  als Eins.

$X_1$	$X_2$	$Y$
$A$	$C$	5
$A$	$D$	10
$B$	$C$	12
$B$	$D$	7

Sie schätzen die Gleichung

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 X_2 + u$$

Welchen Wert erhalten Sie für  $\beta_0$ ?

10:	a anderer Wert	b 0	c 2	d 3	e 5
-----	----------------	-----	-----	-----	-----

**Aufgabe 11:**

Welchen Wert erhalten Sie für  $\beta_1$ ?

11:	a anderer Wert	b 3	c 5	d 7	e 12
-----	----------------	-----	-----	-----	------

**Aufgabe 12:**

Welchen Wert erhalten Sie für  $\beta_2$ ?

12:	a anderer Wert	b 0	c 5	d 10	e 12
-----	----------------	-----	-----	------	------

**Aufgabe 13:**

Welchen Wert erhalten Sie für  $\beta_3$ ?

13:	a anderer Wert	b -10	c 0	d 7	e 10
-----	----------------	-------	-----	-----	------

**Aufgabe 14:** Sie schätzen den Zusammenhang zwischen  $X$  und  $Y$  mit Hilfe unterschiedlicher Modelle und erhalten die folgenden Ergebnisse:

	$\hat{\beta}_0$	$\hat{\beta}_1$
$Y = \beta_0 + \beta_1 \cdot X + u$	12.30	16.00
$Y = \beta_0 + \beta_1 \cdot \log X + u$	24.92	4.30
$\log Y = \beta_0 + \beta_1 \cdot X + u$	2.57	0.82
$\log Y = \beta_0 + \beta_1 \cdot \log X + u$	3.23	0.23

Wie groß schätzen Sie die marginale Änderung von Y bei einer Änderung von X um 1 Prozent?

14:	<sup>a</sup> anderer Wert	<sup>b</sup> 0.23	<sup>c</sup> 0.82	<sup>d</sup> 4.30	<sup>e</sup> 16.00
-----	---------------------------	-------------------	-------------------	-------------------	--------------------

**Aufgabe 15:**

Wie groß schätzen Sie die prozentuale Änderung von Y bei einer Änderung von X um 1 Prozent?

15:	<sup>a</sup> anderer Wert	<sup>b</sup> 0.23%	<sup>c</sup> 0.82%	<sup>d</sup> 4.30%	<sup>e</sup> 16.00%
-----	---------------------------	--------------------	--------------------	--------------------	---------------------

**Aufgabe 16:**

Wie groß schätzen Sie die prozentuale Änderung von Y bei einer Änderung von X um eine Einheit?

16:	<sup>a</sup> anderer Wert	<sup>b</sup> 0.23%	<sup>c</sup> 0.82%	<sup>d</sup> 4.30%	<sup>e</sup> 16.00%
-----	---------------------------	--------------------	--------------------	--------------------	---------------------

**Aufgabe 17:**

Wie groß schätzen Sie die marginale Änderung von Y bei einer Änderung von X um eine Einheit?

17:	<sup>a</sup> anderer Wert	<sup>b</sup> 0.23	<sup>c</sup> 0.82	<sup>d</sup> 4.30	<sup>e</sup> 16.00
-----	---------------------------	-------------------	-------------------	-------------------	--------------------

**Aufgabe 18:** Betrachten Sie das folgende Ergebnis einer Regression. Gehen Sie weiter von einem Signifikanzniveau

von 5% aus. Nehmen Sie ferner an, dass die Residuen normalverteilt sind:

```
Call: lm(formula = y ~ x1 + x2)
      Estimate Std. Error t value Pr(>|t|)
(Intercept) -1.880      2.049   -0.917 0.39436
x1           -3.080      2.956   -1.042 0.33757
x2           14.303      2.507    5.706 0.00125 **
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 5.209 on 6 degrees of freedom
Multiple R-squared: 0.8444, Adjusted R-squared: 0.7926
F-statistic: 16.28 on 2 and 6 DF, p-value: 0.003766
18a: x1 hat einen signifikanten Einfluss auf y.
18b: x2 hat einen signifikanten Einfluss auf y.
18c: Die Hypothese, der marginale Effekt von x1 auf y sei -6, wird abgelehnt.
18d: Die Hypothese, der marginale Effekt von x1 auf y sei -7, wird abgelehnt.
18e: Das Modell erklärt 84.44% der Varianz von Y.
```

**Aufgabe 19:** Wie bestimmen Sie die Obergrenze des 95%-Konfidenzintervalls für den Koeffizienten von x1?

- 19a: Keine der folgenden Antworten ist richtig.  
 19b:  $-3.08 + 2.956 * 2.57$   
 19c:  $-3.08 + 2.956 * 1.96$   
 19d:  $-3.08 + 2.956 * 2.45$   
 19e:  $-3.08 + 2.956/3 * 2.45$

## 15 Klausur BW24.1, Mai 2012

Ihre Version ist . Bitte vergessen Sie nicht, die Version in den Lösungsbogen zu übertragen. Lösen Sie die Aufgaben zunächst hier (und ggf. auf dem Schmierpapier) und übertragen Sie die Lösungen zum Schluss in den Lösungsbogen. Geben Sie den Lösungsbogen ab, und behalten Sie dieses Aufgabenblatt. Wir wünschen Ihnen viel Erfolg!

**Aufgabe 1:** Die Zufallsvariable X ist normalverteilt mit Mittelwert  $\mu$  und unbekannter Standardabweichung. Eine Stichprobe von 16 Beobachtungen ergibt einen Stichprobenmittelwert von  $\bar{x} = 12$  und eine Standardabweichung der Stichprobenbeobachtungen von 8. Was ist die beste erwartungstreue Schätzung für  $\mu$ ?

1:	<sup>a</sup> anderer Wert	<sup>b</sup> 1	<sup>c</sup> 3	<sup>d</sup> 4	<sup>e</sup> 12
----	---------------------------	----------------	----------------	----------------	-----------------

**Aufgabe 2:** Was ist die Standardabweichung Ihres Schätzers für  $\mu$  aus Aufgabe 1?

2:	<sup>a</sup> anderer Wert	<sup>b</sup> 1/2	<sup>c</sup> 2	<sup>d</sup> 4	<sup>e</sup> 8
----	---------------------------	------------------	----------------	----------------	----------------

**Aufgabe 3:** Gehen Sie nun davon aus, dass die Standardabweichung Ihres Schätzers  $\hat{\mu}$  den Wert 5 hat. Nehmen Sie ferner an, dass Sie einen Wert von  $\hat{\mu} = 10$  erhalten haben. Ihre Nullhypothese ist  $\mu = 20$ . Was ist der Absolutwert der Teststatistik?

3:	<sup>a</sup> anderer Wert	<sup>b</sup> 1/2	<sup>c</sup> 2	<sup>d</sup> 4	<sup>e</sup> 8
----	---------------------------	------------------	----------------	----------------	----------------

**Aufgabe 4:** Wir bleiben bei der oben beschriebenen Situation, gehen aber jetzt davon aus, dass Ihre Stichprobe 8 Beobachtungen enthält, und Ihre Teststatistik den Wert  $-2$  hat. Ihre Nullhypothese ist nach wie vor  $\mu = 20$ , die Alternativhypothese ist  $\mu \neq 20$ . Ihr Signifikanzniveau ist 5%.

- 4a: Zum Testen sollte man die Normalverteilung verwenden.  
 4b: Die Nullhypothese wird abgelehnt.  
 4c: Die Alternativhypothese wird abgelehnt.  
 4d: Zum Testen sollte man die t-Verteilung mit 5 Freiheitsgraden verwenden.  
 4e: Der Annahmehereich für die Teststatistik reicht von  $-2.36$  bis  $+2.36$ .

**Aufgabe 5:** Die Zufallsvariable X ist Poissonverteilt,  $P_\lambda(X=k) = \lambda^k \cdot e^{-\lambda}/k!$ . Sie haben nur eine Beobachtung:  $X_1 = 1$ . Was ist der Maximum Likelihood Schätzer für den Parameter  $\lambda$ ?

5:	<sup>a</sup> anderer Wert	<sup>b</sup> 0	<sup>c</sup> 1/2	<sup>d</sup> 1	<sup>e</sup> 2
----	---------------------------	----------------	------------------	----------------	----------------

**Aufgabe 6:** In Ihrer Stichprobe mit 25 Beobachtungen einer normalverteilten Zufallsvariablen mit bekannter Varianz 25 bestimmen Sie einen Mittelwert von 10. Ihr Konfidenzniveau ist 95%. Wie bestimmen Sie die Untergrenze des Konfidenzintervalls für den Mittelwert?

- 6a: Keine der folgenden Antworten ist richtig.  
 6b:  $10 + \text{qnorm}(0.05)$   
 6c:  $10 + 5 * \text{qnorm}(0.05)$   
 6d:  $10 + 5 * \text{qnorm}(0.025)$   
 6e:  $10 - \text{qnorm}(0.975)$

**Aufgabe 7:** Welche Wahrscheinlichkeit gibt der p-Wert an?  
 7a: Keine der folgenden Antworten ist richtig.

- 7b: Die Wahrscheinlichkeit eines Fehlers erster Art.
- 7c: Die Wahrscheinlichkeit eines Fehlers zweiter Art.
- 7d: Die Wahrscheinlichkeit die Alternativhypothese abzulehnen, wenn sie falsch ist.
- 7e: Die Wahrscheinlichkeit die Nullhypothese abzulehnen, wenn sie falsch ist.

**Aufgabe 8:** Ihre Stichprobe enthält 7 unabhängige und identisch verteilte Beobachtungen:  $X_1, \dots, X_7$ . Welche Schätzfunktionen für  $E(X)$  sind erwartungstreu?

- 8a:  $X_3 - X_5$
- 8b:  $X_7$
- 8c:  $X_3 + X_4 - X_5$
- 8d:  $2X_3 + X_4 - 2X_5$
- 8e:  $2X_3 + 2X_4 - 2X_5$

**Aufgabe 9:** Welche Schätzfunktionen für  $E(X)$  sind effizienter?

- 9a:  $2X_3 + X_4 - 2X_5$  ist effizienter als  $X_3 + X_4 - X_5$ .
- 9b:  $X_7$  ist effizienter als  $X_3 + X_4 - X_5$ .
- 9c:  $X_3 + X_4 - X_5$  ist effizienter als  $2X_3 + X_4 - 2X_5$ .
- 9d:  $2X_3$  ist effizienter als  $2X_3 + X_4 - 2X_5$ .
- 9e:  $2X_3 + X_4 - 2X_5$  ist effizienter als  $X_3$ .

**Aufgabe 10:** Sie betrachten zwei Stichproben,  $x_1$  und  $x_2$ . Sie stellen fest, dass im Mittel  $x_2$  kleiner als  $x_1$  ist. Interpretieren Sie das folgende Testergebnis bei einem Signifikanzniveau von 5%:

```

Paired t-test
data: x1 and x2
t = 4.9339, df = 9, p-value = 0.000809
alternative hypothesis: true difference
in means is not equal to 0
95 percent confidence interval:
 2.395042 6.450793
sample estimates:
mean of the differences
4.422917

```

- 10a: Die Mittelwerte von  $x_1$  und  $x_2$  sind signifikant voneinander verschieden.
- 10b: Die Hypothese, die Differenz zwischen  $x_1$  und  $x_2$  sei im Erwartungswert 8, kann verworfen werden.
- 10c: Die Hypothese,  $x_1$  sei im Erwartungswert um 8 Einheiten größer als  $x_2$ , wird nicht abgelehnt.
- 10d: Bei einer kleinen Stichprobe setzt dieser Test voraus, dass  $x_1$  und  $x_2$  einer Normalverteilung folgen.
- 10e: Es wurde ein paarweiser Test durchgeführt.

**Aufgabe 11:** Betrachten Sie das folgende Testergebnis. Verwenden Sie ein Signifikanzniveau von 5%.

```

Wilcoxon signed rank test
data: x1 and x2
V = 55, p-value = 0.001953
alternative hypothesis: true location
shift is not equal to 0

```

- 11a: Es wurde ein einseitiger Test durchgeführt.
- 11b: Die Hypothese,  $x_1$  und  $x_2$  folgen der gleichen Verteilung, wird abgelehnt.
- 11c: Die Hypothese, die Differenz zwischen  $x_1$  und  $x_2$  sei im Mittel 1, kann verworfen werden.
- 11d: Der Test setzt voraus, dass  $x_1$  und  $x_2$  einer Normalverteilung folgen.

11e: Bei einer kleinen Anzahl an Beobachtungen sollte man besser einen  $t$ -Test durchführen.

**Aufgabe 12:** Sie schätzen ein lineares Regressionsmodell um den Einfluss von drei Variablen,  $x_1$ ,  $x_2$  und  $x_3$  auf  $y$  zu bestimmen. Verwenden Sie ein Signifikanzniveau von 5% um das folgende Ergebnis zu interpretieren:

```

Call: lm(formula = y ~ x1 + x2 + x3)
      Estimate Std. Error t value Pr(>|t|)
(Intercept)  -1          2     -0.5  0.6182
x1            6          1      6.0  3.48e-08
x2           -1          1     -1.0  0.3198
x3            2          1      2.0  0.0483

```

- 12a:  $x_1$  hat einen signifikanten Einfluss auf  $y$ .
- 12b:  $x_2$  hat einen signifikanten Einfluss auf  $y$ .
- 12c:  $x_3$  hat einen signifikanten Einfluss auf  $y$ .
- 12d: Wenn  $x_1$  um 10 Einheiten größer wird, erwarten Sie, dass  $y$  um etwa 6 Einheiten größer wird.
- 12e: Wenn  $x_2$  um 10 Einheiten größer wird, erwarten Sie, dass  $y$  um etwa 10 Einheiten kleiner wird.

**Aufgabe 13:**

Ihre Nullhypothese ist, der marginale Effekt von  $x_3$  sei 6. Ihre Alternativhypothese ist, der marginale Effekt sei ungleich 6. Wie groß ist der Absolutwert der Teststatistik?

13:	a anderer Wert	b 0	c 2	d 4	e 6
-----	----------------	-----	-----	-----	-----

**Aufgabe 14:** Nehmen Sie an, die Teststatistik hätte den Wert 1. Wie groß ist der  $p$ -Wert für den Test der Nullhypothese?

14:	a anderer Wert	b 0.0483	c 0.3198	d 0.6182	e 1
-----	----------------	----------	----------	----------	-----

**Aufgabe 15:** Mit den gleichen Daten schätzen Sie ein neues Regressionsmodell und erhalten das folgende Ergebnis:

```

Call: lm(formula = y ~ log(x1) + x2 + x3)
      Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.0          2.0      1.0  0.319962
log(x1)     25.0          2.0     12.5 < 2e-16
x2          -2.0          0.5     -4.0  0.000129
x3           2.5          0.5      5.0  2.77e-06

```

**Aufgabe 15:** Wie groß ist der marginale Effekt von  $x_1$  an der Stelle  $x_1 = 50$ ?

15:	a anderer Wert	b -1	c 1/2	d 1	e 25
-----	----------------	------	-------	-----	------

**Aufgabe 16:**

Wie groß ist der marginale Effekt von  $x_2$  an der Stelle  $x_2 = 50$ ?

16:	a anderer Wert	b -2	c -1/25	d 1/2	e 2
-----	----------------	------	---------	-------	-----

**Aufgabe 17:** Betrachten Sie das folgende Ergebnis einer Regression:

```

Call: lm(formula = y ~ x1 * x2)
      Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.0          2.0      0.5  0.61822
x1           7.5          0.5     15.0 < 2e-16
x2          -4.0          4.0     -1.0  0.31982
x1:x2        3.0          1.0      3.0  0.00344

```

Die Variable  $x_2$  ist eine Dummy-Variable. Sie hat den Wert 0 für Männer und den Wert 1 für Frauen. Wie groß ist der marginale Effekt von  $x_1$  für Männer?

17:	a) anderer Wert	b) -4	c) 1	d) 3.5	e) 7.5
-----	-----------------	-------	------	--------	--------

**Aufgabe 18:**

Wie groß ist der marginale Effekt von  $x_1$  für Frauen?

18:	a) anderer Wert	b) -4	c) 1	d) 3.5	e) 7.5
-----	-----------------	-------	------	--------	--------

**Aufgabe 19:** Ein Kollege rechnet mit den gleichen Daten, allerdings codiert er  $x_2$  anders. Bei ihm hat diese Variable den Wert 1 für Männer und 2 für Frauen. Wie groß ist in diesem Fall der marginale Effekt von  $x_1$  für Frauen?

19:	a) anderer Wert	b) -4	c) 1	d) 3.5	e) 7.5
-----	-----------------	-------	------	--------	--------

## 16 Klausur BW24.1, Februar 2012

Ihre Version ist . Bitte vergessen Sie nicht, die Version in den Lösungsbogen zu übertragen. Lösen Sie die Aufgaben zunächst hier (und ggf. auf dem Schmierpapier) und übertragen Sie die Lösungen zum Schluss in den Lösungsbogen. Geben Sie den Lösungsbogen ab, und behalten Sie dieses Aufgabenblatt. Wir wünschen Ihnen viel Erfolg!

**Aufgabe 1:** Die Zufallsvariable  $X$  hat Mittelwert  $\mu$  und Varianz 25. Sie ziehen eine Stichprobe der Größe 100 und stellen fest, dass  $\sum_{i=1}^{100} X_i = 500$ . Was ist die beste erwartungstreue Schätzung für  $\mu$ ? Übertragen Sie die Lösung in Feld 1 des Lösungsbogens.

**Aufgabe 2:** Was ist die Standardabweichung Ihres Schätzers für  $\mu$  aus Aufgabe 1? Übertragen Sie die Lösung in Feld 2 des Lösungsbogens.

**Aufgabe 3:** Gehen Sie nun davon aus, dass die Standardabweichung Ihres Schätzers  $\hat{\mu}$  den Wert 1 hat. Nehmen Sie ferner an, dass Sie einen Wert von  $\hat{\mu} = 8$  erhalten haben. Ihre Nullhypothese ist  $\mu = 4$ . Was ist der Absolutwert der Teststatistik? Übertragen Sie die Lösung in Feld 3 des Lösungsbogens.

**Aufgabe 4:** Die Zufallsvariable  $X$  ist Poissonverteilt,  $P_\lambda(X=k) = \lambda^k \cdot e^{-\lambda}/k!$ . Sie haben zwei Beobachtungen:  $X_1 = 1$ ,  $X_2 = 0$ . Was ist der Maximum Likelihood Schätzer für den Parameter  $\lambda$ ? Übertragen Sie die Lösung in Feld 4 des Lösungsbogens.

**Aufgabe 5:** In Ihrer Stichprobe mit 4 Beobachtungen einer normalverteilten Zufallsvariablen mit unbekannter Varianz ermitteln Sie einen Mittelwert von 10 und eine Varianz von 100. Ihr Konfidenzniveau ist 95%. Wie bestimmen Sie die Breite des Konfidenzintervalls für den Mittelwert?

- 5a: Keine der folgenden Antworten ist richtig.
- 5b: `qt(10, 4) * 0.95`
- 5c: `qnorm(10) * 5`
- 5d: `qt(.95, 3) * 10`
- 5e: `qnorm(.95) * 10`

**Aufgabe 6:** Welche der folgenden Aussagen trifft zu?

- 6a: Jeder Schätzer für  $\mu$  ist eine Stichprobenfunktion.
- 6b: Jede Stichprobenfunktion ist ein Schätzer für  $\mu$ .
- 6c: Nur konsistente Stichprobenfunktionen sind Schätzer.

6d: Der Mittelwert der Stichprobe ist ein Schätzer für den Mittelwert der Population.

6e: Die Verteilung eines Schätzers  $\hat{\theta}$  über verschiedene mögliche Stichproben nennen wir Stichprobenverteilung von  $\hat{\theta}$ .

**Aufgabe 7:** Für Ihre Schätzung verwenden Sie die Regressionsgleichung

$$Y = \beta_0 + \beta_1 X + u.$$

Welche Aussage trifft zu:

- 7a: Wenn  $\beta_1$  signifikant von 0 verschieden ist, dann beeinflusst die Variable  $X$  die Variable  $Y$  kausal.
- 7b: Wenn Sie feststellen, dass  $\beta_1$  signifikant von 0 verschieden ist, dann gibt es einen linearen Zusammenhang zwischen  $X$  und  $Y$ .
- 7c: Wenn 0 im Konfidenzintervall für  $\beta_1$  liegt, dann ist  $\beta_1$  nicht signifikant von 0 verschieden
- 7d: Wenn 0 nicht im Konfidenzintervall für  $\beta_1$  liegt, dann ist  $\beta_1$  signifikant von 0 verschieden
- 7e: Die Schätzung für  $\beta_1$  wird um so genauer, je stärker  $X$  und  $u$  korreliert sind.

**Aufgabe 8:** Ihre Stichprobe enthält 5 unabhängige und identisch verteilte Beobachtungen:  $X_1, \dots, X_5$ . Welche Schätzfunktionen für  $E(X)$  sind erwartungstreu?

- 8a:  $(X_1 - X_2 + 2X_5 - X_3 + X_4)/2$ .
- 8b:  $(X_1 - X_5)/2$ .
- 8c:  $X_3$ .
- 8d:  $2X_1 - X_2 + X_3$ .
- 8e:  $3X_1 - X_2 - X_3$ .

**Aufgabe 9:** Welche Schätzfunktionen für  $E(X)$  sind effizienter?

- 9a:  $(X_1 - X_5)/2$  ist effizienter als  $\sum_{i=1}^5 X_i/5$ .
- 9b:  $\sum_{i=1}^5 X_i/5$  ist effizienter als  $(X_1 - X_5)/2$ .
- 9c:  $X_3$  ist effizienter als  $\sum_{i=1}^5 X_i/5$ .
- 9d:  $\sum_{i=1}^5 X_i/5$  ist effizienter als  $X_3$ .
- 9e:  $X_3$  ist effizienter als  $3X_1 - X_2 - X_3$ .

**Aufgabe 10:** Betrachten Sie das folgende Testergebnis. Verwenden Sie ein Signifikanzniveau von 5%.

```

Paired t-test
data: x1 and x2
t = -5.7127, df = 9, p-value = 0.0002896
alternative hypothesis: true difference in means
is not equal to 0
95 percent confidence interval:
-7.853645 -3.398087
sample estimates:
mean of the differences
-5.625866

```

- 10a: Es wurde ein paarweiser Test durchgeführt.  
10b: Die Mittelwerte von  $x_1$  und  $x_2$  sind nicht signifikant voneinander verschieden.  
10c: Die Hypothese, der Betrag der Differenz zwischen  $x_1$  und  $x_2$  sei im Erwartungswert 7, kann nicht verworfen werden.  
10d: Die Hypothese, der Betrag der Differenz zwischen  $x_1$  und  $x_2$  sei im Erwartungswert 7, wird abgelehnt.  
10e: Bei einer kleinen Stichprobe setzt dieser Test nicht voraus, dass  $x_1$  und  $x_2$  einer Normalverteilung folgen.

**Aufgabe 11:** Betrachten Sie das folgende Testergebnis. Verwenden Sie ein Signifikanzniveau von 5%.

```

Wilcoxon signed rank test
data: x1 and x2
V = 1, p-value = 0.003906
alternative hypothesis: true location shift
is not equal to 0

```

- 11a: Es wurde ein zweiseitiger Test durchgeführt.  
11b: Die Hypothese,  $x_1$  und  $x_2$  folgen der gleichen Verteilung, wird nicht abgelehnt.  
11c: Die Hypothese, die Differenz zwischen  $x_1$  und  $x_2$  sei 1, kann verworfen werden.  
11d: Der Test setzt nicht voraus, dass  $x_1$  und  $x_2$  einer Normalverteilung folgen.  
11e: Der Test kann auch bei einer kleinen Anzahl an Beobachtungen angewendet werden.

**Aufgabe 12:** Betrachten Sie das folgende Ergebnis einer Regression. Gehen Sie weiter von einem Signifikanzniveau von 5% aus:

```

Call: lm(formula = y ~ x1 + x2)
      Estimate Std. Error t value Pr(>|t|)
(Intercept)  6.967      8.329   0.836  0.4305
x1           8.222      2.697   3.049  0.0186 *
x2          -3.182      1.453  -2.189  0.0648 .
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 11.09 on 7 degrees of freedom
Multiple R-squared:  0.5962, Adjusted R-squared:  0.4808
F-statistic:5.168 on 2 nd 7 DF, p-value:.04184

```

- 12a:  $x_1$  hat einen signifikanten Einfluss auf  $y$ .  
12b:  $x_2$  hat einen signifikanten Einfluss auf  $y$ .  
12c: Die Hypothese, der marginale Effekt von  $x_1$  auf  $y$  sei  $-5$ , wird abgelehnt.  
12d: Die Hypothese, der marginale Effekt von  $x_2$  auf  $y$  sei  $-5$ , wird abgelehnt.  
12e: Die Hypothese, der Zusammenhang von  $x_1$  auf  $y$  sei nicht linear, kann abgelehnt werden.

**Aufgabe 13:** Nun erweitern Sie die obige Regression um die Variable  $x_3$ . Sie erhalten folgendes Ergebnis:

```

Call: lm(formula = y ~ x1 + x2 + x3)
      Estimate Std. Error t value Pr(>|t|)
(Intercept)  6.967      8.329   0.836  0.4305
x1           8.222      2.697   3.049  0.0186 *
x2          -3.182      1.453  -2.189  0.0648 .
x3              NA         NA      NA      NA
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 11.09 on 7 degrees of freedom
Multiple R-squared:  0.5962, Adjusted R-squared:  0.4808
F-statistic: 5.168 on 2 and 7 DF, p-value: 0.04184

```

13a: Der Einfluss von  $x_3$  auf  $y$  ist nicht signifikant.  
13b: Der marginale Effekt von  $x_3$  ist Null.  
13c:  $x_3$  ist eine lineare Funktion von  $x_1$  und  $x_2$ .  
13d:  $x_3$  ist eine nicht-lineare Funktion von  $x_1$  und  $x_2$ .  
13e: Das Modell erklärt 59.62% der Varianz von  $y$ .

**Aufgabe 14:** Betrachten Sie das folgende Regressionmodell:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 X_2 + u$$

$X_1$  und  $X_2$  sind jeweils Dummyvariablen die entweder den Wert 0 oder 1 haben. Sie messen die folgenden Mittelwerte für  $Y$ :

	$X_1$	$X_2$	$\bar{Y}$
	0	0	3
	1	0	4
	0	1	5
	1	1	6

Welchen Wert hat  $\hat{\beta}_0$ ?

14:	a anderer Wert	b 0	c 1	d 2	e 3
-----	----------------	-----	-----	-----	-----

**Aufgabe 15:** Welchen Wert hat  $\hat{\beta}_1$ ?

15:	a anderer Wert	b 0	c 1	d 2	e 3
-----	----------------	-----	-----	-----	-----

**Aufgabe 16:** Welchen Wert hat  $\hat{\beta}_2$ ?

16:	a anderer Wert	b 0	c 1	d 2	e 3
-----	----------------	-----	-----	-----	-----

**Aufgabe 17:** Welchen Wert hat  $\hat{\beta}_3$ ?

17:	a anderer Wert	b -2	c -1	d 0	e 1
-----	----------------	------	------	-----	-----

**Aufgabe 18:** Betrachten Sie das folgende Regressionsmodell:

$$Y = \beta_0 + \beta_1 \ln X_1 + \beta_2 \ln X_2 + u$$

Ihre Nullhypothese ist, dass bei einem Wert von  $X_1 = 10$  der marginale Effekt von  $X_1$  den Wert  $-2$  hat. Welchen Wert erwarten Sie also für  $\hat{\beta}_1$ ?

18:	a anderer Wert	b -20	c -2	d -1/20	e 10
-----	----------------	-------	------	---------	------

**Aufgabe 19:** Für  $\beta_2$  schätzen Sie in diesem Modell einen Wert von 12. Was ist der marginale Effekt von  $X_2$  auf  $Y$  wenn  $X_2$  den Wert 3 annimmt?

19:	a anderer Wert	b 3	c 4	d 12	e 36
-----	----------------	-----	-----	------	------

# 17 Klausur BW24.1, Mai 2011

Ihre Version ist . Bitte vergessen Sie nicht, die Version in den Lösungsbogen zu übertragen. Lösen Sie die Aufgaben zunächst hier (und ggf. auf dem Schmierpapier) und übertragen Sie die Lösungen zum Schluss in den Lösungsbogen. Geben Sie den Lösungsbogen ab, und behalten Sie dieses Aufgabenblatt. Wir wünschen Ihnen viel Erfolg!

**Aufgabe 1:** Die Varianz von  $X$  sei 2. Was ist die Varianz von  $3X_2 - 2X_2 - 2X_3 + 2X_4$ ? Übertragen Sie die Lösung in Feld 1 des Lösungsbogens.

**Aufgabe 2:** Das Gewicht von Widgets ist gleichverteilt (rechteckverteilt) mit unbekannter Standardabweichung. Eine Stichprobe von 25 Widgets ergibt ein mittleres Gewicht von 50 und eine Standardabweichung von 20. Was ist die Standardabweichung des Mittelwertes,  $\hat{\sigma}_{\bar{x}}$ ? Übertragen Sie die Lösung in Feld 2 des Lösungsbogens.

**Aufgabe 3:** Gehen Sie nun davon aus, dass die geschätzte Standardabweichung des Mittelwertes 1 ist. Alle anderen Angaben sind wie in der obigen Aufgabe. Bestimmen Sie mit Hilfe der Tabellen im beiliegenden Informationsblatt die Obergrenze des 90% Konfidenzintervalls für den Mittelwert. Falls Sie im Informationsblatt keine exakt passende Verteilung finden, gehen Sie davon aus, dass die Stichprobe hinreichend groß ist, um eine Approximation verwenden zu können. Übertragen Sie die Lösung in Feld 3 des Lösungsbogens.

**Aufgabe 4:** Die Verteilung  $Q$  hängt von einem Parameter  $\theta$  ab. Das erste Moment dieser Verteilung ist  $\mu_1 = 3 - \theta/2$ . Unsere Stichprobe enthält drei Beobachtungen: -1, 0, 1. Was ist der Momentenschätzer auf Basis des ersten Moments für  $\theta$ ? Übertragen Sie die Lösung in Feld 4 des Lösungsbogens.

**Aufgabe 5:** Die Zufallsvariable  $X$  ist normalverteilt. Unsere Stichprobe enthält drei Beobachtungen: 10, 20, und 30. Geben Sie den unverzerrten Schätzer für die Varianz von  $X$  an. Übertragen Sie die Lösung in Feld 5 des Lösungsbogens.

**Aufgabe 6:**  $X$  ist entsprechend der folgenden Wahrscheinlichkeitsfunktion verteilt:

$$P(X = x|\theta) = \begin{cases} 1/4 & \text{falls } x = \theta - 4 \\ 1/2 & \text{falls } x = \theta - 2 \\ 1/4 & \text{falls } x = \theta \end{cases}$$

Was ist der ML-Schätzer für  $\theta$ , wenn die Stichprobe die drei Beobachtungen: 5, 5, 7 enthält? Übertragen Sie die Lösung in Feld 6 des Lösungsbogens.

**Aufgabe 7:** Ihre Stichprobe enthält 7 unabhängige und identisch verteilte Beobachtungen:  $X_1, \dots, X_7$ . Welche Schätzfunktionen für  $E(X)$  sind erwartungstreu?

- 7a:  $(X_2 + X_3)/2$
- 7b:  $(2X_2 - X_3 + X_4)/2$
- 7c:  $X_4$
- 7d:  $2X_2 - X_3 + X_4$
- 7e:  $2X_2 - 3X_2 + X_4$

**Aufgabe 8:** Die Länge von Widgets ist normalverteilt mit bekannter Standardabweichung. Eine Stichprobe von 25 Widgets ergibt eine mittleres Länge von 250 und eine Standardabweichung von 25. Wie bestimmen Sie die Obergrenze des 99% Konfidenzintervalls für den Mittelwert?

- 8a:  $250+5*qt(0.005, df=24)$
- 8b:  $250+5*qt(0.99, df=24)$
- 8c:  $250-25*qt(0.005, df=24)$
- 8d:  $250-25*qnorm(0.005)$
- 8e:  $250-5*qnorm(0.005)$

**Aufgabe 9:**  $X$  ist normalverteilt mit Mittelwert 70 und Varianz 100. Wie wahrscheinlich ist es, dass der Mittelwert einer Stichprobe von 25 Einheiten größer als 71 ist?

- 9a:  $1-pt(1/4, df=24)$
- 9b:  $1-pnorm(1/2)$
- 9c:  $pnorm(1/2)$
- 9d:  $pt(2, df=24)$
- 9e:  $pnorm(2)$

**Aufgabe 10:** Sie untersuchen den Zusammenhang zwischen zwei Merkmalen,  $X$  und  $Y$ . Das Merkmal  $X$  kann die beiden Werte 1 und 2 annehmen,  $Y$  hat entweder den Wert  $A$  oder  $B$ . Die Häufigkeit der Merkmale ist durch die folgende Tabelle gegeben:

	$Y = A$	$Y = B$
$X = 1$	22	13
$X = 2$	18	16

Ihr Signifikanzniveau ist 5%. Interpretieren Sie das folgende Testergebnis:

Pearson's Chi-squared test with Yates' continuity correction  
 $X\text{-squared} = 0.3485$ ,  $df = 1$ ,  $p\text{-value} = 0.555$

- 10a: Zwischen den beiden Merkmalen  $X$  und  $Y$  gibt es einen signifikanten Zusammenhang.
- 10b: Die Nullhypothese,  $X$  und  $Y$  seien unabhängig, kann man zum gegebenen Signifikanzniveau verwerfen.
- 10c: Die Nullhypothese,  $X$  und  $Y$  seien voneinander abhängig, kann man zum gegebenen Signifikanzniveau verwerfen.
- 10d: Die Nullhypothese,  $X$  und  $Y$  seien unabhängig, kann man nicht zum gegebenen Signifikanzniveau verwerfen.
- 10e: Um zu testen, ob  $X$  und  $Y$  voneinander unabhängig sind, sollte man besser den Korrelationskoeffizienten betrachten.

**Aufgabe 11:** Ihr Signifikanzniveau ist 5%. Interpretieren Sie das folgende Testergebnis:

Wilcoxon rank sum test with continuity correction  
 data: x1 and x2  
 $W = 4182$ ,  $p\text{-value} = 0.04577$   
 alternative hypothesis: true location shift  
is not equal to 0

- 11a: Es wurde ein einseitiger Test durchgeführt.
- 11b: Es wurde ein zweiseitiger Test durchgeführt.
- 11c: Dieser Test gibt nur dann ein sinnvolles Ergebnis, wenn  $x_1$  und  $x_2$  nicht normalverteilt sind.
- 11d: Die Nullhypothese kann bei einem Signifikanzniveau von 5% abgelehnt werden.
- 11e:  $x_1$  und  $x_2$  sind nicht signifikant verschieden.

**Aufgabe 12:** Das Gewicht von Widgets ist normalverteilt mit unbekannter Varianz. Ihre Nullhypothese ist, das mittlere Gewicht sei kleiner oder gleich 5. Ihre Alternativhypothese ist, das mittlere Gewicht sei größer als 12. In Ihrer Stichprobe von 5 Beobachtungen messen Sie einen Mittelwert von  $\bar{x}$ . Für  $\bar{x}$  schätzen Sie eine Standardabweichung von  $\hat{\sigma}_{\bar{x}}$ . Ihr Signifikanzniveau ist 5%. Sie berechnen eine Teststatistik

$$g = \frac{\bar{x} - 5}{\hat{\sigma}_{\bar{x}}}$$

Bestimmen Sie mit Hilfe der Tabellen im beiliegenden Informationsblatt die Obergrenze des Annahmebereichs für  $g$ .

12:	a	b	c	d	e
	anderer Wert	1.64	2.02	2.13	2.78

**Aufgabe 13:** In einem linear-log Modell wird  $Y$  als abhängig von  $X$  und  $C$  erklärt:

$$Y = \beta_0 + \beta_1 \ln X + \beta_2 \ln C + u$$

- 13a:  $\beta_1$  gibt die Elastizität von  $Y$  auf  $X$  an  
 13b:  $\beta_1$  gibt an, um wieviel Prozent sich  $Y$  ändert, wenn  $X$  um 0.01 wächst  
 13c:  $1/\beta_1$  gibt an, um welchen Betrag sich  $Y$  ändert, wenn  $X$  um 1% wächst  
 13d:  $\beta_1/100$  gibt an, um welchen Betrag sich  $Y$  ändert, wenn  $X$  um 1% wächst  
 13e: keine der Antworten ist richtig

**Aufgabe 14:** Die Variable  $Y$  hängt von den beiden binären Variablen  $A$  und  $B$  ab. Sie beobachten folgende Mittelwerte für  $Y$  für unterschiedliche Kombinationen von  $A$  und  $B$ :

	B	kein B
A	3	5
kein A	7	11

Mit den gleichen Daten schätzen Sie auch eine lineare Regression:

$$Y = \beta_0 + \beta_A d_A + \beta_B d_B + \beta_{AB} d_A d_B + u$$

Dabei sind  $d_A$  und  $d_B$  jeweils Dummyvariablen, die 1 sind, wenn  $A$  bzw.  $B$  eintritt und 0 sonst. Welchen Wert schätzen Sie für  $\beta_B$ ?

14:	a	b	c	d	e
	anderer Wert	-4	-6	2	11

**Aufgabe 15:** Welchen Wert schätzen Sie für  $\beta_{AB}$ ?

15:	a	b	c	d	e
	anderer Wert	-4	-6	2	11

**Aufgabe 16:** Sie schätzen den Zusammenhang  $Y = \beta_0 + \beta_1 \cdot X_1 + \beta_2 \cdot X_2 + \beta_3 \cdot X_1 \cdot X_2 + u$  und erhalten das folgende Ergebnis:

```
Call: lm(formula = Y ~ X1 * X2)
      Estimate Std. Error t value Pr(>|t|)
(Intercept)  -5          5      -1  0.187
X1            -1          1      -1  0.187
X2            6          2       3  0.020 *
X1:X2        -3          1      -3  0.020 *
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 30.89 on 78 degrees of freedom
Multiple R-squared:  0.4331, Adjusted R-squared:  0.4113
F-statistic: 19.87 on 3 and 78 DF, p-value: 1.157e-09
```

**Aufgabe 16:** Was ist der Erwartungswert von  $Y$ , wenn  $X_1 = 2$  und  $X_2 = 1$  ist?

16:	a	b	c	d	e
	anderer Wert	-6	-1	0	1

**Aufgabe 17:** Was ist der marginale Effekt von  $X_2$ , wenn  $X_1 = 0$  ist?

17:	a	b	c	d	e
	anderer Wert	0	2	3	6

**Aufgabe 18:** Verwenden Sie ein Signifikanzniveau von 5%.

- 18a: Die Hypothese, die Interaktion zwischen  $X_1$  und  $X_2$  sei Null, wird abgelehnt.  
 18b: 19.87% der Varianz von  $Y$  werden durch das Modell erklärt.  
 18c: 43.31% der Varianz von  $Y$  werden durch das Modell erklärt.  
 18d: Der marginale Effekt von  $X_1$  ist signifikant abhängig von  $X_2$ .  
 18e: Die Hypothese  $\beta_2 = 0$  wird abgelehnt.

**Aufgabe 19:** Ihre Nullhypothese ist nun, dass  $\beta_3 = 1$ . Wie groß ist der Absolutbetrag Ihrer Teststatistik?

19:	a	b	c	d	e
	anderer Wert	1	2	3	4

## 18 Klausur BW24.1, Februar 2011

Übertragen Sie die Lösung in Feld 1 des Lösungsbogens.

Ihre Version ist . Bitte vergessen Sie nicht, die Version in den Lösungsbogen zu übertragen. Lösen Sie die Aufgaben zunächst hier (und ggf. auf dem Schmierpapier) und übertragen Sie die Lösungen zum Schluss in den Lösungsbogen. Geben Sie den Lösungsbogen ab, und behalten Sie dieses Aufgabenblatt. Wir wünschen Ihnen viel Erfolg!

**Aufgabe 2:** Die Dichtefunktion von  $X$  ist

$$f(X|\theta) = \begin{cases} \frac{X-5}{2\theta-10} & \text{falls } X \in [5, \theta] \\ 0 & \text{sonst} \end{cases}$$

Was ist der ML-Schätzer für  $\theta$  wenn die Stichprobe die Beobachtungen: 6, 7, 7 und 11 enthält? Übertragen Sie die Lösung in Feld 2 des Lösungsbogens.

**Aufgabe 1:** Die Verteilung  $Q$  hängt von einem Parameter  $\theta$  ab. Das erste Moment dieser Verteilung ist  $\mu_1 = \theta - 1$ . Unsere Stichprobe enthält 4 Beobachtungen: 10, 20, 30 und 40. Was ist der Momentenschätzer auf Basis des ersten Moments für  $\theta$ ?

**Aufgabe 3:** Die Länge von Knorz ist normalverteilt mit unbekannter Standardabweichung. Eine Stichprobe von 9 Stück Knorz ergibt eine mittlere Länge von 100 und eine Standardabweichung von 27. Was ist die Standardabweichung

des Mittelwertes,  $\hat{\sigma}_{\bar{x}}$ ? Übertragen Sie die Lösung in Feld 3 des Lösungsbogens.

**Aufgabe 4:** Gehen Sie nun davon aus, dass die geschätzte Standardabweichung des Mittelwertes 10 ist. Alle anderen Angaben sind wie in der obigen Aufgabe. Bestimmen Sie mit Hilfe der Tabellen im beiliegenden Informationsblatt die Obergrenze des 90% Konfidenzintervalls für den Mittelwert. Übertragen Sie die Lösung in Feld 4 des Lösungsbogens.

**Aufgabe 5:** Die Varianz von  $X$  sei 10. Was ist die Varianz von  $2X_2 - 3X_3 + 2X_4$ ? Übertragen Sie die Lösung in Feld 5 des Lösungsbogens.

**Aufgabe 6:** Ihre Stichprobe enthält 5 unabhängige und identisch verteilte Beobachtungen:  $X_1, \dots, X_5$ . Welche Schätzfunktionen für  $E(X)$  sind erwartungstreu?

- 6a:  $X_2$
- 6b:  $2X_2 - X_3$
- 6c:  $2X_2 - X_3 + X_4$
- 6d:  $2X_2 - X_3 + 2X_4$
- 6e:  $2X_2 - 3X_3 + 2X_4$

**Aufgabe 7:** Die Länge von Knipps ist normalverteilt mit unbekannter Standardabweichung. Eine Stichprobe von 25 Stück Knipps ergibt eine mittlere Länge von 2000 und eine Standardabweichung von 100. Wie bestimmen Sie die Obergrenze des 99% Konfidenzintervalls für den Mittelwert?

- 7a:  $2000 - 20 * qt(0.005, df=24)$
- 7b:  $2000 + 20 * qt(0.005, df=24)$
- 7c:  $2000 - 100 * qt(0.005, df=24)$
- 7d:  $2000 - 20 * qnorm(0.005)$
- 7e:  $2000 - 100 * qnorm(0.005)$

**Aufgabe 8:** Sie testen Ihre Nullhypothese  $H_0$  mit einem gegebenen Datensatz. Welche der folgenden Aussagen trifft normalerweise zu: Je kleiner das Signifikanzniveau ist, um so...

- 8a: ... öfter wird  $H_0$  abgelehnt, obwohl sie wahr ist.
- 8b: ... seltener wird  $H_0$  abgelehnt, obwohl sie wahr ist.
- 8c: ... öfter wird  $H_0$  angenommen, obwohl sie falsch ist.
- 8d: ... seltener wird  $H_0$  angenommen, falls sie wahr ist.
- 8e: ... öfter wird  $H_0$  angenommen, falls sie wahr ist.

**Aufgabe 9:** Ihr Signifikanzniveau ist 5%. Interpretieren Sie das folgende Testergebnis:

```
Paired t-test
data: x1 and x2
t = -6.3706, df = 36, p-value = 2.228e-07
alternative hypothesis: true difference in means is
not equal to 0
```

95 percent confidence interval:  
-0.13222183 -0.06836484

sample estimates:  
mean of the differences  
-0.1002933

- 9a:  $x_1$  ist signifikant von  $x_2$  verschieden
- 9b:  $x_1$  ist nicht signifikant von  $x_2$  verschieden
- 9c: Das Konfidenzintervall für die Differenz zwischen  $x_1$  und  $x_2$  enthält die Null.
- 9d: Die Nullhypothese in diesem Test war, dass die Mittelwerte von  $x_1$  und  $x_2$  gleich sind.

9e: Weil die  $t$ -Statistik so klein ist, kann die Nullhypothese nicht abgelehnt werden.

**Aufgabe 10:** Interpretieren Sie das folgende Testergebnis:

```
Wilcoxon signed rank test
data: x1 and x2
V = 46, p-value = 1
alternative hypothesis: true location shift
is greater than 0
```

- 10a: Es wurde ein einseitiger Test durchgeführt.
- 10b: Es wurde ein zweiseitiger Test durchgeführt.
- 10c: Eine notwendige Annahme für diesen Test ist, dass  $x_1$  und  $x_2$  normalverteilt sind.
- 10d: Die Nullhypothese kann bei einem Signifikanzniveau von 5% abgelehnt werden.
- 10e: Das Testergebnis zeigt, dass  $x_1$  nicht signifikant von  $x_2$  verschieden ist.

**Aufgabe 11:** Die Länge von Knorz ist normalverteilt. Ihre Nullhypothese ist, die mittlere Länge von Knorz sei größer oder gleich 12. Ihre Alternativhypothese ist, die mittlere Länge von Knorz sei kleiner 12. In Ihrer Stichprobe von 5 Beobachtungen messen Sie einen Mittelwert von  $\bar{x}$ . Für  $\bar{x}$  schätzen Sie eine Standardabweichung von  $\hat{\sigma}_{\bar{x}}$ . Ihr Signifikanzniveau ist 5%. Sie berechnen eine Teststatistik

$$g = \frac{12 - \bar{x}}{\hat{\sigma}_{\bar{x}}}$$

Bestimmen Sie mit Hilfe der Tabellen im beiliegenden Informationsblatt die Obergrenze des Annahmebereichs für  $g$ .

11:	a	b	c	d	e
	anderer Wert	1.64	2.02	2.13	2.78

**Aufgabe 12:** In einem log-log Modell wird  $Y$  als abhängig von  $X$  und  $C$  erklärt:

$$\ln Y = \beta_0 + \beta_1 \ln X + \beta_2 \ln C + u$$

- 12a:  $\beta_1$  gibt die Elastizität von  $Y$  auf  $X$  an
- 12b:  $\beta_1$  gibt die Elastizität von  $X$  auf  $Y$  an
- 12c:  $\beta_1$  gibt an, um wieviel Prozent sich  $Y$  ändert, wenn  $X$  um 0.01 wächst
- 12d:  $\beta_1/100$  gibt an, um welchen Betrag sich  $Y$  ändert, wenn  $X$  um 1% wächst
- 12e: keine der Antworten ist richtig

**Aufgabe 13:** Die Variable  $Y$  hängt von den beiden binären Variablen  $A$  und  $B$  ab. Sie beobachten folgende Mittelwerte für  $Y$  für unterschiedliche Kombinationen von  $A$  und  $B$ :

	A	kein A
B	12	5
kein B	3	6

Mit den gleichen Daten schätzen Sie auch eine lineare Regression

$$Y = \beta_0 + \beta_A d_A + \beta_B d_B + \beta_{AB} d_A d_B + u$$

Dabei sind  $d_A$  und  $d_B$  jeweils Dummyvariablen, die 1 sind, wenn  $A$  bzw.  $B$  eintritt und 0 sonst.

Welchen Wert schätzen Sie für  $\beta_0$ ?

13:	a	b	c	d	e
	anderer Wert	3	5	6	12

**Aufgabe 14:** Welchen Wert schätzen Sie für  $\beta_{AB}$ ?

14:	a	b	c	d	e
	anderer Wert	7	9	10	12

**Aufgabe 15:** Sie schätzen den Zusammenhang  $Y = \beta_0 + \beta_1 \cdot X + \beta_2 \cdot Z + \beta_3 \cdot X \cdot Z + u$  und erhalten das folgende Ergebnis:

```
Call: lm(formula = Y ~ X * Z)
            Estimate Std. Error t value Pr(>|t|)
(Intercept)    -2.00         4.00   -0.50 0.321665
X                 8.00         1.00   8.00 0.000662 ***
Z                -2.00         1.00  -2.00 0.058058 .
X:Z               3.00         1.00   3.00 0.019971 *
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.0 ' '
Residual standard error: 25.81 on 61 degrees of freedom
Multiple R-squared:  0.7723, Adjusted R-squared:  0.7611
F-statistic: 68.95 on 3 and 61 DF, p-value: < 2.2e-16
Was ist der marginale Effekt von X, wenn Z = 1 ist?
```

15:	a	b	c	d	e
	anderer Wert	3	5	7	8

**Aufgabe 16:** Was ist der Erwartungswert von Y, wenn  $X = 1$  und  $Z = -1$  ist?

16:	a	b	c	d	e
	anderer Wert	5	7	9	11

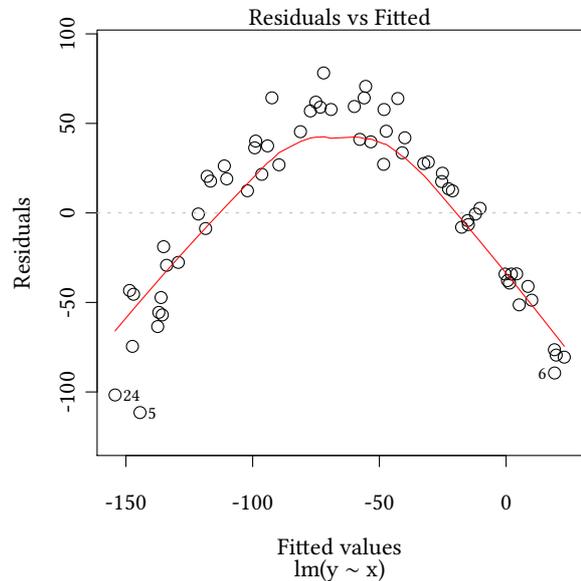
**Aufgabe 17:** Verwenden Sie ein Signifikanzniveau von 5%.

- 17a: Die Hypothese, die Interaktion zwischen X und Z sei Null, wird abgelehnt.  
 17b: 25.81% der Varianz von Y werden durch das Modell erklärt.  
 17c: 77.23% der Varianz von Y werden durch das Modell erklärt.  
 17d: Der marginale Effekt von X ist abhängig von Z.  
 17e: Die Hypothese  $\beta_2 = 0$  wird abgelehnt.

**Aufgabe 18:** Ihre Nullhypothese ist nun, dass  $\beta_3 = 5$ . Wie groß ist der Absolutbetrag Ihrer Teststatistik?

18:	a	b	c	d	e
	anderer Wert	1	2	3	4

**Aufgabe 19:** Sie schätzen den Zusammenhang  $Y = \beta_0 + \beta_1 X + u$ . Betrachten Sie den folgenden diagnostischen Plot einer linearen Regression:



- 19a: Die Regression ist korrekt spezifiziert.  
 19b: Die Annahme identisch verteilter  $(X_i, Y_i)$  ist verletzt.  
 19c: Die Annahme  $E(u|X) = 0$  ist möglicherweise verletzt.  
 19d: Es wäre besser, ein quadratisches Modell zu schätzen.  
 19e: X und Y sind offensichtlich nicht miteinander korreliert.

## 19 Klausur BW24.1, Mai 2010

Ihre Version ist . Bitte vergessen Sie nicht, die Version in den Lösungsbogen zu übertragen. Lösen Sie die Aufgaben zunächst hier (und ggf. auf dem Schmierpapier) und übertragen Sie die Lösungen zum Schluss in den Lösungsbogen. Geben Sie den Lösungsbogen ab, und behalten Sie dieses Aufgabenblatt. Wir wünschen Ihnen viel Erfolg!

**Aufgabe 1:** X ist gleichverteilt (rechteckverteilt) über dem Intervall  $[a, b]$ . Wir interessieren uns für die Obergrenze dieses Intervalls, b. Das erste Moment der Gleichverteilung ist  $E(X) = (a + b)/2$ . Ihre Stichprobe enthält 9 unabhängige Beobachtungen:  $x_1, \dots, x_9$ . Welche Schätzfunktionen für E(b) sind erwartungstreu?

- 1a:  $x_9$   
 1b:  $x_1 + x_7 - a$   
 1c:  $2x_1 + 2x_7 - 2x_9 - a$   
 1d:  $x_1 - x_7 - x_9 + 2 \cdot a$   
 1e:  $a + x_1 + x_7 - x_9$

**Aufgabe 2:** Welche Schätzfunktionen für E(b) sind effizienter?

- 2a:  $x_2 - a + x_3$  ist effizienter als  $\frac{1}{2}(4x_1 - 2a)$

- 2b:  $x_2 - a + x_3$  ist effizienter als  $2 \cdot x_1 + x_7 - x_6 - a$   
 2c:  $\frac{1}{2}(4x_1 - 2a)$  ist effizienter als  $2 \cdot x_1 + x_7 - x_6 - a$   
 2d:  $2 \cdot x_1 - a$  ist effizienter als  $2 \cdot x_1 + x_7 - x_6 - a$   
 2e:  $-a + \frac{2}{9} \sum_{i=1}^9 x_i$  ist effizienter als  $2 \cdot x_1 + x_7 - x_6 - a$

**Aufgabe 3:** Eine Zufallsvariable ist gleichverteilt (rechteckverteilt) über dem Intervall  $[a, b]$ . Ihre Stichprobe enthält drei Beobachtungen: 1, 2 und 3. Welche Aussagen über den ML (Maximum-Likelihood) Schätzer treffen zu?

- 3a: Der ML Schätzer für a ist 0  
 3b: Der ML Schätzer für a ist 1  
 3c: Der ML Schätzer für a ist 2  
 3d: Der ML Schätzer für b ist 3  
 3e: Der ML Schätzer für b ist 4

**Aufgabe 4:** Ihre Stichprobe enthält wieder drei Beobachtungen: 1, 3 und 5. Außerdem wissen Sie, dass  $b - a = 10$  ist. Was ist der Momentenschätzer für a auf Basis des ersten Moments?

4:	a	b	c	d	e
	anderer Wert	-5	-2	-3	5

**Aufgabe 5:** Die Länge von Knipps ist normalverteilt mit Standardabweichung 100 und unbekanntem Mittelwert. Eine

Stichprobe von 25 Stück Knipps ergibt eine mittlere Länge von 2000. Wie bestimmen Sie die Obergrenze des 99% Konfidenzintervalls für den Mittelwert?

- 5a:  $2000 - qnorm(0.005)$
- 5b:  $2000 + 20 * qnorm(0.005)$
- 5c:  $2000 - 20 * qnorm(0.005)$
- 5d:  $2000 + 4 * qnorm(0.99)$
- 5e:  $2000 - 100 * qnorm(0.005)$

**Aufgabe 6:** Wie bestimmen Sie die Untergrenze des 90% Konfidenzintervalls für den Mittelwert?

- 6a:  $2000 - 20 * qnorm(0.05)$
- 6b:  $2000 - qnorm(0.005)$
- 6c:  $2000 - 20 * qnorm(0.95)$
- 6d:  $2000 + 4 * qnorm(0.1)$
- 6e:  $2000 - 100 * qnorm(0.05)$

**Aufgabe 7:** In der Vorlesung haben wir das Verfahren der Varianzanalyse (AOV) behandelt. Welche der folgenden Aussagen trifft für dieses Verfahren zu?

- 7a: Es vergleicht die Mediane mehrerer Stichproben
- 7b: Es ist ein nichtparametrisches Verfahren.
- 7c: Es lässt sich nur anwenden, wenn die abhängige Variable  $F$ -verteilt ist
- 7d: Es betrachtet den Anteil der erklärten Varianz an der nicht erklärten Varianz
- 7e: Es lässt sich nur anwenden, wenn die erklärende Variable diskret ist

**Aufgabe 8:** Die Länge von Knorz ist normalverteilt. Ihre Nullhypothese ist, die mittlere Länge von Knorz sei kleiner oder gleich 12. Ihre Alternativhypothese ist, die mittlere Länge von Knorz sei größer 12. In Ihrer (eher kleinen) Stichprobe der Größe  $n$  messen Sie einen Mittelwert von  $\bar{x}$ . Für  $\bar{x}$  schätzen sie eine Standardabweichung von  $\hat{\sigma}_{\bar{x}}$ . Ihr Signifikanzniveau ist 10%. Sie berechnen eine Teststatistik

$$g = \frac{12 - \bar{x}}{\hat{\sigma}_{\bar{x}}}$$

Sei  $Q_n^t$  die Quantilsfunktion der  $t$ -Verteilung mit  $n$  Freiheitsgraden und  $Q^N$  die Quantilsfunktion der Normalverteilung. Sie lehnen die Nullhypothese ab, falls

- 8a:  $g > Q_{n-1}^t(0.9)$
- 8b:  $g < Q_{n-1}^t(0.95)$
- 8c:  $g > Q_{n-1}^t(0.05)$
- 8d:  $g < Q_{n-1}^t(0.1)$
- 8e:  $g > Q_{n-1}^N(0.9)$

**Aufgabe 9:** Jetzt sei Ihr Signifikanzniveau 1%. Sie lehnen die Nullhypothese ab, falls

- 9a:  $g < -Q_{n-1}^t(0.99)$
- 9b:  $g < Q_{n-1}^t(0.01)$
- 9c:  $g < Q^N(0.01)$
- 9d:  $g < Q^N(0.005)$
- 9e:  $g < -Q^N(0.005)$

**Aufgabe 10:** Ihr Signifikanzniveau ist immer noch 1%, allerdings ist Ihre Nullhypothese nun, die mittlere Länge von Knorz sei gleich 12. Ihre Alternativhypothese ist nun, die

mittlere Länge von Knorz sei ungleich 12. Sie lehnen die Nullhypothese ab, falls

- 10a:  $|g| > Q_{n-1}^t(0.995)$
- 10b:  $|g| < Q_{n-1}^t(0.995)$
- 10c:  $-|g| < Q_{n-1}^t(0.005)$
- 10d:  $-|g| < Q^N(0.005)$
- 10e:  $-|g| < Q^N(0.01)$

**Aufgabe 11:** Um den Beratungserfolg zu messen, vergleiche eine Unternehmensberatung bei 12 Unternehmen den Gewinn jeweils im Jahr vor und nach der Umstrukturierung der von ihr beratenen Unternehmen. Der Gewinn nach der Beratung steht in der Variablen  $x$ . Der Gewinn vor der Beratung steht in  $y$ . Nun führen Sie einen einseitigen Wilcoxon Test durch. Ihr Signifikanzniveau  $\alpha$  ist 1%. Sie erhalten folgendes Ergebnis:

```
Wilcoxon signed rank test
data: x and y
V = 72, p-value = 0.003418
alternative hypothesis: true location shift
is greater than 0
```

Was ist Ihre Schlussfolgerung?

- 11a: Die Gewinne nach Beratung sind signifikant größer.
- 11b: Bei einer so kleinen Stichprobe ist es besser, einen  $F$ -Test durchzuführen.
- 11c: Die gemessenen Unterschiede sind klein und deuten nicht auf einen signifikanten Unterschied hin.
- 11d: Wenn die Alternativhypothese ist, der Gewinn sei nach Beratung größer als vorher, ist es besser, einen zweiseitigen Test durchzuführen.
- 11e: Bei einem größeren Signifikanzniveau ( $\alpha$  wird also größer, das muss nicht besser sein!) würde man keinen signifikanten Unterschied mehr finden.

**Aufgabe 12:** Die Unternehmensberatung misst die Steigerung des Gewinns nach Beratung der 12 Unternehmen als  $S$ . Sie betrachten zwei mögliche Faktoren, die den Erfolg der Beratung vielleicht beeinflussen: Dauer der Beratung  $L$  in Stunden und Sektor des Unternehmens. Der Sektor ist codiert als  $D=1$  für Dienstleistung und  $D=0$  sonst. Sie erhalten das folgende Ergebnis:

```
Call: lm(formula = S ~ L * D)
      Estimate Std. Error t value Pr(>|t|)
(Intercept)  10          5        2   0.081 .
L             20          5        4   0.004 ***
D            -10         10       -1   0.347
L:D           -5          1       -5   0.001 ***
```

Ein Unternehmen aus dem Sektor „Dienstleistung“ wird 100 Stunden beraten. Welche Steigerung des Gewinns erwarten Sie?

12:	a 1000	b -490	c anderer Wert	d 2000	e 2010
-----	--------	--------	----------------	--------	--------

**Aufgabe 13:** Ein Unternehmen aus einem anderen Sektor (nicht „Dienstleistung“) wird ebenfalls beraten. Wie groß ist die marginale Steigerung des Gewinns pro Stunde?

13:	a -5	b 30	c 10	d 15	e 20
-----	------	------	------	------	------

**Aufgabe 14:** Ihr Signifikanzniveau ist 5%. Welche der folgenden Aussagen treffen zu?

- 14a: Die durch Beratung erzielte Steigerung des Gewinns hängt nicht signifikant vom Sektor ab.

- 14b: Die Dauer der Beratung hat einen signifikanten Einfluss auf den Gewinn.  
 14c: Die  $t$ -Statistik folgt immer einer  $t$ -Verteilung, egal welcher Verteilung die Störterme folgen.  
 14d: Die obigen Schätzergebnisse zeigen, dass man den Interaktionsterm zwischen L und D besser weglassen sollte.  
 14e: Wenn man L aus der Schätzgleichung weglässt, wird das  $R^2$  voraussichtlich sinken.

**Aufgabe 15:** Sie erheben Daten zu einer abhängigen Variablen  $Y$  und einer unabhängigen Variablen  $X$ . Die Daten sind durch folgende Tabelle gegeben:

$X$	0	0	2	2
$Y$	0	4	0	4

Nun schätzen Sie ein lineares Regressionsmodell ohne Konstante:

$$Y_i = \beta_1 X_i + u_i$$

Wie groß ist der OLS Schätzer  $\hat{\beta}_1$ ?

15:	<sup>a</sup> -1	<sup>b</sup> anderer Wert	<sup>c</sup> 1	<sup>d</sup> 2	<sup>e</sup> 4
-----	-----------------	---------------------------	----------------	----------------	----------------

**Aufgabe 16:** Nun erweitern Sie ihr Modell um einen konstanten Term:

$$Y_i = \beta_1 X_i + \beta_0 + u_i$$

Wie groß ist jetzt Ihr OLS Schätzer  $\hat{\beta}_1$ ?

16:	<sup>a</sup> -1	<sup>b</sup> 0	<sup>c</sup> 1	<sup>d</sup> 2	<sup>e</sup> anderer Wert
-----	-----------------	----------------	----------------	----------------	---------------------------

**Aufgabe 17:** Welche Aussagen treffen zu:

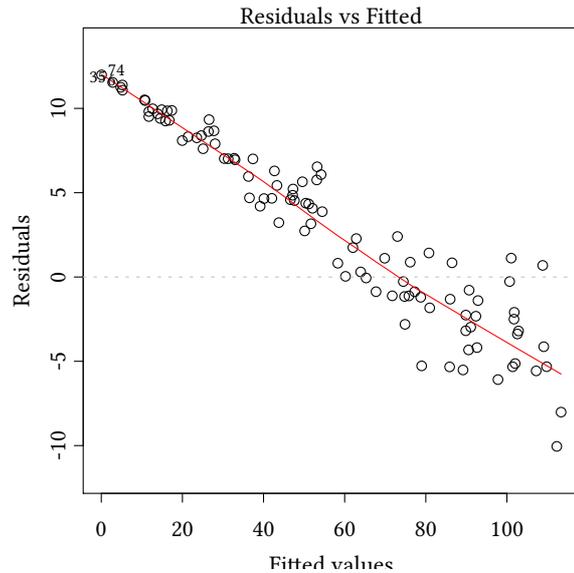
- 17a: Das zweite Modell hat ein kleineres  $R^2$  als das erste.  
 17b: Das zweite Modell hat ein größeres  $R^2$ .  
 17c: Das  $R^2$  in beiden Modellen ist gleich.  
 17d: Die Summe der Fehlerquadrate im zweiten Modell sind größer als im ersten Modell.  
 17e: Die Summe der Fehlerquadrate im zweite Modell sind kleiner als im ersten Modell.

**Aufgabe 18:**

Sie schätzen ein lineares Regressionsmodell ohne Konstante

$$Y_i = \beta_1 X_i + u_i$$

und erhalten den folgenden diagnostischen Plot ihrer Regression. Welche Schlussfolgerungen ziehen Sie?



- 18a: Zwischen unabhängiger und abhängiger Variablen besteht ein fallender Zusammenhang  
 18b: Wenn Sie das Modell um eine Konstante erweitern, wird sich der Fit der Regression verbessern  
 18c: Die Annahme  $E(u_i|X_i = x) = 0$  ist nicht erfüllt  
 18d: Die Annahme,  $\text{var}(u_i|X_i = x)$  ist konstant, ist nicht erfüllt.  
 18e: Die Annahme, große Ausreißer in  $X$  und  $Y$  seien selten, ist nicht erfüllt.

**Aufgabe 19:** Eine Gruppe von Sportlern bereitet sich auf einen Wettkampf vor. Über die Sportler liegen folgende Informationen vor: Alter ( $A$ ), Geschlecht ( $G$ ; 1 falls weiblich, 0 sonst), tägliches Training ( $T$ , 1 falls ja, 0 sonst), gesunde Ernährung ( $E$ ; 1 falls ja, 0 sonst) und Ranglistenpunkte ( $R$ ). Alter und Geschlecht sind mit den anderen Variablen nicht korreliert. Sie vermuten, dass Sportler nur dann besonders viele Ranglistenpunkte haben, wenn sie täglich trainieren und sich gesund ernähren; ein tägliches Training wirkt nur zusammen mit gesunder Ernährung. Was wären mögliche Spezifikationen des Modells um Ihren Verdacht zu überprüfen. (Hier ist *nicht* nach der „besten“ Spezifikation gefragt)

- 19a:  $R = \beta_0 + \beta_1 \cdot T + \beta_2 \cdot E + u$   
 19b:  $T \cdot E = \beta_0 + \beta_1 \cdot R + u$   
 19c:  $R = \beta_0 + \beta_1 \cdot A + \beta_2 \cdot G + \beta_3 \cdot T + \beta_4 \cdot E + u$   
 19d:  $R = \beta_0 + \beta_1 \cdot A + \beta_2 \cdot G + \beta_3 \cdot T + \beta_4 \cdot E + \beta_5 \cdot T \cdot E + u$   
 19e:  $R = \beta_0 + \beta_1 \cdot T + \beta_2 \cdot E + \beta_3 \cdot T \cdot E + u$

## 20 Klausur BW24.1, Februar 2010

Ihre Version ist . Bitte vergessen Sie nicht, die Version in den Lösungsbogen zu übertragen. Lösen Sie die Aufgaben zunächst hier (und ggf. auf dem Schmierpapier) und übertragen Sie die Lösungen zum Schluss in den Lösungsbogen. Geben Sie den Lösungsbogen ab, und behalten Sie dieses Aufgabenblatt. Wir wünschen Ihnen viel Erfolg!

**Aufgabe 1:** Sie interessieren sich für den Erwartungswert von  $X$ . Ihre Stichprobe enthält 9 unabhängige Beobachtungen:

$x_1, \dots, x_9$ . Welche Schätzfunktionen für  $E(X)$  sind erwartungstreu?

- 1a:  $x_9$   
 1b:  $x_1 + x_7 - x_9$   
 1c:  $2x_1 + 2x_7 - 3x_9$   
 1d:  $x_1 - x_7 - x_9$   
 1e:  $2x_1 + 3x_7 - 2x_9$

**Aufgabe 2:** Welche Schätzfunktionen für  $E(X)$  sind effizient?

- 2a:  $x_9$   
 2b:  $x_1 + x_7 - x_9$

- 2c:  $2x_1 + 2x_7 - 3x_9$   
 2d:  $\frac{1}{9} \sum_i^9 x_i$   
 2e:  $\sum_i^9 x_i$

**Aufgabe 3:** Welche Schätzfunktionen für  $E(X)$  sind effizienter?

- 3a:  $x_9$  ist effizienter als  $x_1 + x_7 - x_9$   
 3b:  $x_1 + x_7 - x_9$  ist effizienter als  $2 \cdot x_9 - x_8$   
 3c:  $x_1$  ist effizienter als  $(x_2 + x_3)/2$   
 3d:  $\frac{1}{9} \sum_i^9 x_i$  ist effizienter als  $\frac{1}{7} \sum_i^7 x_i$   
 3e:  $x_3$  ist effizienter als  $\frac{1}{7} \sum_i^7 x_i$

**Aufgabe 4:** Eine Zufallsvariable ist wie folgt verteilt:  
 $P(X = 1) = \theta, P(X = 3) = \theta, P(X = 5) = 1 - 2\theta$ . Es gilt  
 $0 \leq \theta \leq \frac{1}{2}$ . Ihre Stichprobe enthält zwei Beobachtungen: 1 und 3. Was ist der Maximum-Likelihood-Schätzer für  $\theta$ ?

4:	a	b	c	d	e
	0	1/4	1/2	anderer Wert	1

**Aufgabe 5:** Ihre Stichprobe enthält nun drei Beobachtungen: 1, 3 und 5. Was ist der Momentenschätzer auf Basis des ersten Moments für  $\theta$ ?

5:	a	b	c	d	e
	anderer Wert	1/4	1/3	1/2	1

**Aufgabe 6:** Sie wollen mit einem Normal Q-Q Plot überprüfen, ob eine Variable normalverteilt ist. Letzteres ist der Fall, wenn...

- 6a: alle Punkte im Plot etwa auf einer horizontalen Linie liegen  
 6b: fast alle Punkte im Plot innerhalb einer Ellipse um den Mittelwert liegen  
 6c: alle Punkte im Plot etwa auf einer Linie liegen  
 6d: alle Punkte im Plot etwa auf einer Linie mit Steigung von ca. 1.96 liegen  
 6e: alle Punkte im Plot gleichmäßig um eine Horizontale streuen

Die Länge von Knorz ist normalverteilt mit Varianz 100 und unbekanntem Mittelwert. Eine Stichprobe von 100 Stück Knorz ergibt eine mittlere Länge von 200. Wie bestimmen Sie die Untergrenze des 95% Konfidenzintervalls für den Mittelwert?

- 7a:  $200 - \text{qnorm}(0.975)$   
 7b:  $200 - 10 * \text{qnorm}(0.025)$   
 7c:  $200 + \text{qnorm}(0.025)$   
 7d:  $200 + 10 * \text{qnorm}(0.025)$   
 7e:  $200 + 10 * \text{qt}(0.025, \text{df}=9)$

Ihre Nullhypothese sei  $H_0$ , die Alternativhypothese sei  $H_1$ . Welche der folgenden Aussagen trifft zu?

- 8a: Der Fehler 1. Art gibt an, wie häufig  $H_0$  angenommen wird, obwohl sie falsch ist.  
 8b: Der Fehler 1. Art gibt an, wie häufig  $H_0$  abgelehnt wird, obwohl sie wahr ist.  
 8c: Der Fehler 2. Art gibt an, wie häufig  $H_0$  angenommen wird, obwohl sie falsch ist.  
 8d: Der Fehler 1. Art gibt an, wie häufig  $H_1$  abgelehnt wird, obwohl sie wahr ist.  
 8e: Der Fehler 2. Art gibt an, wie häufig  $H_1$  angenommen wird, obwohl sie falsch ist.

**Aufgabe 9:** Die Länge von Knipps ist normalverteilt. Ihre Nullhypothese ist, die mittlere Länge von Knipps sei 100. Ihre Alternativhypothese ist, die mittlere Länge von Knipps sei ungleich 100. In Ihrer Stichprobe der Größe  $n$  messen Sie einen Mittelwert von  $\bar{x}$  und eine Standardabweichung von  $\hat{\sigma}_x$ . Ihr Signifikanzniveau ist 5%. Sie berechnen eine Teststatistik

$$g = \sqrt{n} \cdot \frac{\bar{x} - 100}{\hat{\sigma}_x}$$

Sei  $Q_n^t$  die Quantilsfunktion der  $t$ -Verteilung mit  $n$  Freiheitsgraden und  $Q^N$  die Quantilsfunktion der Normalverteilung. Sie lehnen die Nullhypothese ab, falls

$$9: \begin{matrix} \text{a} & \text{b} & \text{c} & \text{d} & \text{e} \\ g > Q_{n-1}^t(0.95) & g < Q_{n-1}^t(0.975) & g > Q_{n-1}^t(0.975) & g > Q_n^t(0.975) & g < Q^N(0.025) \end{matrix}$$

**Aufgabe 10:** Knorz kann auf zwei verschiedene Arten hergestellt werden. Sie haben an 9 Tagen den Output des ersten Verfahrens gemessen. Das Ergebnis (die 9 Zahlen) schreiben Sie in die Variable x. Sie haben an 11 weiteren Tagen des Output des zweiten Verfahrens gemessen. Das Ergebnis (die 11 Zahlen) schreiben Sie in die Variable y. Nun führen Sie einen Mann-Whitney U Test durch. Ihr Signifikanzniveau ist 10%. Sie erhalten folgendes Ergebnis:

```
Wilcoxon rank sum test
data: x and y
W = 13, p-value = 0.9734
alternative hypothesis: true location shift
is not equal to 0
```

Was ist Ihre Schlussfolgerung?

- 10a: Die Verfahren führen zu einem signifikant verschiedenen Output.  
 10b: Bei einer so kleinen Stichprobe sollte man besser einen  $t$ -Test durchführen.  
 10c: Die gemessenen Unterschiede sind klein und deuten nicht auf einen signifikanten Unterschied hin.  
 10d: Mit den Ihnen vorliegenden Daten sollte man besser einen paarweisen Test durchführen.  
 10e: Bei einem kleineren Signifikanzniveau würde man einen signifikanten Unterschied finden.

**Aufgabe 11:** Sie schätzen ein lineares Regressionsmodell um den Einfluss von drei Variablen,  $x_1$ ,  $x_2$ , und  $x_3$  auf  $y$  zu ermitteln. Sie erhalten folgendes Ergebnis:

```
Call: lm(formula = y ~ x1 + x2 + x3)
      Estimate Std. Error t value Pr(>|t|)
(Intercept)  -0.2         1.0      -0.2   0.842
x1             0.5         1.0       0.5   0.618
x2             1.5         1.0       1.5   0.137
x3            -1.5         1.5      -1.0   0.320
```

Ihre Nullhypothese ist, dass bei einer Steigerung von  $x_2$  um 10 Einheiten die Variable  $y$  um 20 Einheiten wächst. Wie groß ist bei einem zweiseitigen Test der absolute Betrag Ihrer Teststatistik, wenn Ihr Datensatz 100 Beobachtungen enthält?

11:	a	b	c	d	e
	1.5	0.5	0.05	anderer Wert	0.15

**Aufgabe 12:** Welchen  $p$ -Wert erhalten Sie für den Test der obigen Hypothese?

12:	a	b	c	d	e
	0.137	0.320	anderer Wert	0.618	0.842

**Aufgabe 13:** Sie untersuchen die Wirksamkeit von zwei Marketingmaßnahmen mit Hilfe einer linearen Regression. Ihre abhängige Variable  $y$  ist der Umsatz. Die unabhängigen Variablen  $x_1$  und  $x_2$  sind jeweils 1 wenn Maßnahme 1 bzw. 2 eingesetzt wurde und 0 sonst. Sie erhalten folgendes Ergebnis:

```
Call: lm(formula = y ~ x1 + x2)
      Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.5         0.6      2.5  0.014 *
x1           3.0         0.8     3.75 0.000 ***
x2           0.5         0.8     0.625 0.533
x1:x2       -3.2         1.0    -3.2  0.002 ***
```

Welchen Umsatz erwarten Sie, wenn Sie gleichzeitig Maßnahme 1 und Maßnahme 2 einsetzen?

13:	a	b	c	d	e
	-3.2	0.3	1.8	anderer Wert	3.5

**Aufgabe 14:** Sei  $P$  der Preis und  $Q$  die nachgefragte Menge,  $X$  die Sonnenscheindauer an einem Verkaufstag, und  $u$  der Störterm der Regression. Wir nennen  $\frac{\partial Q}{\partial P} \frac{P}{Q}$  die Preiselastizität der Nachfrage. Sie schätzen die Gleichung

$$\log Q = \beta_0 + \beta_P \log P + \beta_X X + u$$

und erhalten folgendes Ergebnis:

	$\hat{\beta}$	$\hat{\sigma}$
$\beta_0$	50	10
$\beta_P$	5	1
$\beta_X$	2	2

Wie groß ist die geschätzte Preiselastizität?

14:	a	b	c	d	e
	2	5	1/5	2/5	anderer Wert

**Aufgabe 15:** Sie beobachten, dass von 40 Männern 8 Produkt A kaufen, weitere 8 kaufen Produkt B, und 24 kaufen Produkt C. Von 60 Frauen kaufen 12 Produkt A, 12 kaufen Produkt B, und 36 kaufen Produkt C. Ihre Nullhypothese  $H_0$  ist, dass

Präferenzen für A, B, und C unabhängig vom Geschlecht sind. Wie groß ist Ihre Teststatistik?

15:	a	b	c	d	e
	0	1/2	1	100	anderer Wert

**Aufgabe 16:** Nehmen Sie an, Sie hätten in der obigen Aufgabe eine Teststatistik von 10 berechnet. Wie berechnen Sie den  $p$ -Wert für Ihren Test?

- 16a: `pchisq(10,2)`  
 16b: `1-pchisq(10,2)`  
 16c: `1-pchisq(10,6)`  
 16d: `pt(10,5,lower=FALSE)`  
 16e: `pnorm(-10)`

**Aufgabe 17:** Betrachten Sie folgenden Datensatz:

	Y	X	Z
1	2	-1	
3	6	3	
5	10	-5	
7	14	7	

Sie schätzen  $Y_i = \beta_0 + \beta_1 X_i + u_i$  mithilfe eines OLS-Schätzers. Ihr Schätzer für  $\beta_1$  ist

17:	a	b	c	d	e
	0	1/5	1/2	2	anderer Wert

Ihr  $R^2$  ist

18:	a	b	c	d	e
	0	1/5	1/2	1	anderer Wert

Was passiert, wenn  $Z$  in das Modell aufgenommen wird?

- 19a: Die Schätzung für  $\beta_1$  verändert sich. Vorher war sie verzerrt, weil das Modell unterspezifiziert war.  
 19b: AIC steigt  
 19c: AIC sinkt  
 19d:  $R^2$  steigt  
 19e:  $R^2$  sinkt

## 21 Klausur Statistik II, Mai 2008

- Bearbeiten Sie die Klausur bitte in zwei Stunden und ohne Hilfsmittel.

Viel Erfolg!

Sei  $X$  eine binomialverteilte Zufallsvariable mit Stichprobengröße 6 und Erfolgswahrscheinlichkeit 0,85. Die folgende Tabelle gibt die Verteilungsfunktion  $P(X = x)$  und die kumulierte Verteilungsfunktion  $F(x) = P(X < x)$  an:

$x$	0	1	2	3	4	5	6
$P(X = x)$	0.00001	0.00039	0.00549	0.04145	0.17618	0.39933	0.37715
$F(x) = P(X < x)$	0	0.00001	0.0004	0.00589	0.04734	0.22352	0.62285

- (a) Bestimmen Sie das 25% Quantil von  $X$ .  
 (b) Wie wahrscheinlich ist es, dass  $X$  kleiner als 3 ist?  
 (c) Wie wahrscheinlich ist es, dass  $X$  zwischen 3,5 und 7,5 liegt?

Sei  $X$  eine auf dem Intervall  $[0, 200]$  gleichverteilte Zufallsvariable.

- (a) Bestimmen Sie den Median von  $X$ .  
 (b) Bestimmen Sie den Mittelwert von  $X$ .  
 (c) Wie groß muß  $X^*$  sein, damit 95% aller Werte größer als  $X^*$  sind?  
 (d) Zeichnen Sie die Dichtefunktion (vergessen Sie nicht, die Achsen zu beschriften).  
 (e) Zeichnen Sie die Verteilungsfunktion.

Betrachten Sie die zweidimensionale Zufallsvariable  $(X, Y)$  mit der Dichtefunktion

$$f(x, y) = \begin{cases} \frac{1}{2}(3x + y) & \text{falls } x \in [0, 1] \text{ und } y \in [0, 1] \\ 0 & \text{sonst} \end{cases}$$

- Zeigen Sie, dass  $f(x, y)$  tatsächlich eine Dichtefunktion ist.
- Bestimmen Sie  $P\left(X < \frac{1}{2}, Y < 2\right)$
- Bestimmen Sie  $P\left(X > \frac{1}{3}, Y > -1\right)$
- Bestimmen Sie  $P(X > 1)$

Über das Konsumverhalten in der Mensa wissen wir folgendes:

- 50% der Gäste essen keine Vorspeise.
- 30% der Gäste essen keinen Nachtisch.
- 10% der Gäste essen weder Vorspeise noch Nachtisch.

- Wie groß ist die Wahrscheinlichkeit, dass ein Gast, der keinen Nachtisch isst, auch keine Vorspeise isst?
- Wie groß ist die Wahrscheinlichkeit, dass ein Gast, der eine Vorspeise gewählt hat, auch noch einen Nachtisch nimmt?

Ihre Firma stellt Antriebsketten her. Die Länge der einzelnen Kettenglieder ist unabhängig voneinander normalverteilt. Die mittlere Länge eines Kettengliedes ist 10. Insgesamt haben 95% aller Kettenglieder eine Länge zwischen 9 und 11. Jede Kette ist so lang wie die Summe der einzelnen Glieder. Sie interessieren sich für Ketten, die aus 100 Kettengliedern bestehen.

- Wie wahrscheinlich ist es, dass eine solche Kette eine Gesamtlänge von über 1000 hat (verwenden Sie, falls notwendig, einen R-Ausdruck, um die Länge zu bestimmen)?
- Wie wahrscheinlich ist es, dass eine solche Kette eine Gesamtlänge zwischen 900 und 1100 hat (verwenden Sie, falls notwendig, einen R-Ausdruck, um die Länge zu bestimmen)?
- Wie wahrscheinlich ist es, dass eine solche Kette eine Gesamtlänge zwischen 990 und 1010 hat (verwenden Sie, falls notwendig, einen R-Ausdruck, um die Länge zu bestimmen)?
- Wie wahrscheinlich ist es, dass eine solche Kette eine Gesamtlänge zwischen 999 und 1001 hat (verwenden Sie, falls notwendig, einen R-Ausdruck, um die Länge zu bestimmen)?

Für die Verteilungsfunktion einer diskreten Zufallsvariablen  $X$  gilt

$$P(X = x) = \begin{cases} \frac{1}{4} & \text{falls } x = \theta \\ \frac{1}{2} & \text{falls } x = \theta + 1 \\ \frac{1}{4} & \text{falls } x = \theta + 2 \\ 0 & \text{sonst} \end{cases}$$

Der Parameter  $\theta$  ist unbekannt und soll geschätzt werden.

- Bestimmen Sie den Maximum Likelihood Schätzer für  $\theta$ , wenn Ihre Stichprobe die drei Werte  $\{2, 2, 3\}$  enthält.
- Bestimmen Sie den Momentenschätzer für  $\theta$  (auf Basis des ersten Moments)

Sie testen eine Abfüllmaschine für Joghurt. Aus einer Stichprobe von 10 Joghurtbechern bestimmen Sie eine mittlere Füllmenge von 150 Gramm bei einer Varianz von 17. Sie nehmen an, dass die Füllmenge normalverteilt ist.

- Der Hersteller der Maschine hat Ihnen versprochen, die Varianz der Füllmenge sei langfristig 9. Sie sind beunruhigt, weil die gemessene Varianz bei Ihrer Maschine größer ist und rufen empört beim Hersteller an. Dessen Telefonhotline versichert Ihnen, dass die von Ihnen gemessene Abweichung zufällig sei, und langfristig auch Ihre Maschine eine Varianz von nicht größer als 9 haben würde. Testen Sie diese Hypothese mit einem einseitigen Test bei einem Signifikanzniveau von 5%.
- Nehmen Sie an, die Varianz der Füllmenge Ihrer Maschine sei langfristig 10. Bestimmen Sie ein 95%-Konfidenzintervall für die mittlere Füllmenge.
- Gehen Sie nun wieder von einer langfristigen Varianz von 9 aus. Wie groß muß Ihr Stichprobenumfang mindestens sein, damit das 95%-Konfidenzintervall höchstens eine Breite von 0,392 Gramm hat
- Eine andere Joghurtmaschine befüllt Becher mit im Mittel 50 Gramm Joghurt. Die Standardabweichung beträgt 10 Gramm. Sie nehmen eine Stichprobe von 100 Bechern. Wie groß ist die Wahrscheinlichkeit, dass die Gesamtmenge Joghurt in diesen 100 Bechern kleiner als 4 800 Gramm ist?
- Eine weitere Joghurtmaschine produziert nach Herstellerspezifikation bei 80 Joghurtbechern im Mittel 4 Becher Ausschuss. Sie nehmen eine Stichprobe von 80 Bechern, und finden darunter 6 Becher Ausschuss. Ist die Qualität dieser Maschine signifikant verschieden von der Herstellerspezifikation?

- Schreiben Sie eine Nullhypothese auf und testen Sie bei einem Signifikanzniveau von 5%.
- Falls Sie „krumme“ Zahlen erhalten, verwenden Sie eine möglichst genaue Abschätzung.
- Gehen Sie davon aus, dass Sie die Binomialverteilung durch eine Normalverteilung approximieren können, ohne dass eine Stetigkeitskorrektur notwendig ist.

Sie vergleichen zwei Düngemittel für Salatköpfe: X und Y. In Ihrer Entwicklungsabteilung sind jeweils 30 Salatköpfe mit X und 30 Salatköpfe mit Y behandelt worden. Das Gewicht der Salatköpfe hat Ihr Assistent bereits in den Variablen x und y eingetragen. Ihr Assistent hat ferner zwei Tests mit R durchgeführt und präsentiert Ihnen den folgenden Output:

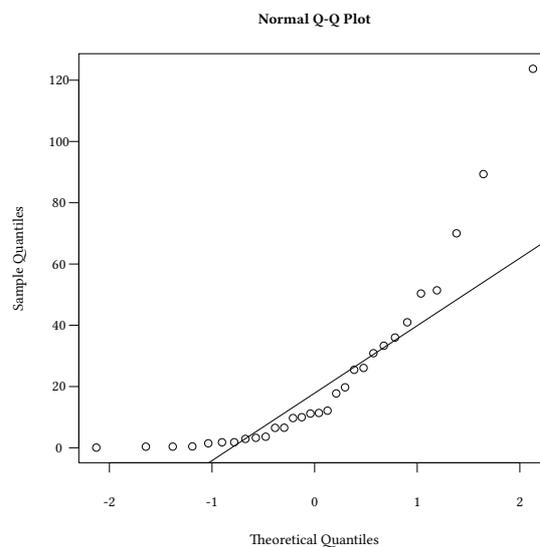
Test 1:

```
t.test(x,y,paired=TRUE)
  Paired t-test
data: x and y
t = -2.0558, df = 29, p-value = 0.0489
alternative hypothesis: true difference in
                      means is not equal to 0
95 percent confidence interval:
-22.02793284 -0.05680469
sample estimates:
mean of the differences
-11.04237
```

Test 2:

```
t.test(x,y)
  Welch Two Sample t-test
data: x and y
t = -2.0203, df = 31.547, p-value = 0.05191
alternative hypothesis: true difference in
                      means is not equal to 0
95 percent confidence interval:
-22.18182418 0.09708665
sample estimates:
mean of x mean of y
23.28710 34.32947
```

- Welche Annahmen werden von beiden obigen Tests vorausgesetzt?
- Nehmen Sie an, dass diese Annahmen erfüllt sind. Welcher der beiden Tests ist in diesem Fall angemessen? Warum?
- Ihre Nullhypothese ist, dass die beiden Düngemittel das Gewicht nicht in unterschiedlicher Weise beeinflussen. Ihr Signifikanzniveau ist 5%. Können Sie mit dem Test, den Sie in Teilaufgabe (b) ausgewählt haben, Ihre Nullhypothese ablehnen? Warum?
- Ihr fleißiger Assistent präsentiert Ihnen die folgende Graphik die er mit dem Kommando `qqnorm(x)`; `qqline(x)`; erzeugt hat. Was sagt Ihnen diese Graphik über die Voraussetzungen der obigen Tests?



- Ihr Assistent präsentiert nun zwei weitere Tests. Sind diese Tests grundsätzlich besser oder weniger gut geeignet als Test 1 oder 2 aus dieser Aufgabe, um Ihre Hypothese zu überprüfen?

Test 3

```
wilcox.test(x,y)
  Wilcoxon rank sum test
data: x and y
W = 220, p-value = 0.0005109
alternative hypothesis: true location shift
is not equal to 0
```

Test 4:

```
wilcox.test(x,y,paired=TRUE)
  Wilcoxon signed rank test
data: x and y
V = 111, p-value = 0.01130
alternative hypothesis: true location shift
is not equal to 0
```

- (f) Welcher der beiden Tests ist hier angemessen?
- (g) Ihre Nullhypothese ist, dass die beiden Düngemittel das Gewicht nicht in unterschiedlicher Weise beeinflussen. Ihr Signifikanzniveau ist 5%. Können Sie mit dem Test, den Sie in Teilaufgabe (f) ausgewählt haben, Ihre Nullhypothese ablehnen? Warum?

## 22 Klausur Statistik II, Februar 2008

- Bearbeiten Sie die Klausur bitte in zwei Stunden und ohne Hilfsmittel.

Viel Erfolg!

$X$  ist eine auf dem Intervall  $[2, 10]$  gleichverteilte (rechteckverteilte) Zufallsvariable.

- Bestimmen Sie das 25% Quantil von  $X$
- Wie wahrscheinlich ist es, dass  $X$  kleiner als 3 ist?
- Wie wahrscheinlich ist es, dass  $X$  zwischen 3 und 7 liegt?

Betrachten Sie die zweidimensionale Zufallsvariable  $(X, Y)$  mit der Dichtefunktion

$$f(x, y) = \begin{cases} \frac{1}{16}(x + 3y) & \text{falls } x \in [0, 2] \text{ und } y \in [0, 2] \\ 0 & \text{sonst} \end{cases}$$

- Zeigen Sie, dass  $f(x, y)$  tatsächlich eine Dichtefunktion ist.
- Bestimmen Sie  $P(X > \frac{1}{2}, Y > 3)$
- Bestimmen Sie  $P(Y < 1)$
- Bestimmen Sie  $P(X < 1, Y < 1)$

Um unnötigen Klausuraufwand zu vermeiden, soll ein Test zeigen, ob Studenten „fit fürs Examen“ sind. Der Test zeigt bei 90% der Studenten, die tatsächlich „fit“ sind ein positives Resultat. Außerdem werden 10% der Studenten die „nicht fit“ sind positiv diagnostiziert. In unserer Stichprobe ist nur jeder vierte Student „fit“. Wie wahrscheinlich ist es, dass ein zufällig ausgewählter Student aus dieser Stichprobe, der in unserem Test ein positives Resultat erreicht, tatsächlich „fit“ ist?

Die Dichtefunktion der Zufallsvariablen  $X$  ist

$$f(x) = \begin{cases} 1 - \frac{x - \theta}{2} & \text{falls } x \in [\theta, \theta + 2] \\ 0 & \text{sonst} \end{cases}$$

Der Erwartungswert  $E(X) = \frac{2}{3} + \theta$ . Der Parameter  $\theta$  ist unbekannt und soll geschätzt werden.

- Bestimmen Sie den Maximum Likelihood Schätzer für  $\theta$ , wenn Ihre Stichprobe die beiden Werte  $\{2, 3\}$  enthält.
- Bestimmen Sie den Momentenschätzer für  $\theta$  (auf Basis des ersten Moments), wenn Ihre Stichprobe die beiden Werte  $\{2, 3\frac{1}{3}\}$  enthält.

Die durchschnittliche Füllmenge eines Sacks Kartoffeln ist 50 kg, die Standardabweichung ist 2 kg. Nehmen Sie an, dass die Füllmenge normalverteilt ist. Sie kaufen 4 Sack Kartoffeln und bestimmen die durchschnittliche Füllmenge  $x$  dieser 4 Säcke. Schreiben Sie jeweils ein R Kommando auf, das die Antworten auf die folgenden Fragen berechnet.

- Wie wahrscheinlich ist es, dass  $x > 50$  kg?
- Wie wahrscheinlich ist es, dass  $x > 52$  kg?
- Benutzen Sie nun die Tabelle der Quantile, die Sie in der Anlage zur Klausur erhalten haben, und bestimmen Sie (ohne R) das Gewicht  $x^*$ , so dass in 99% aller Fälle der Wert  $x$  über diesem Gewicht liegt.

Sie entwickeln einen Energiedrink zur Verbesserung des Muskelaufbaus. Den Muskelaufbau pro Person messen wir als  $X$ . Sie testen diesen Drink an Versuchspersonen in vier (identischen) Fitness-Studios. Die folgende Tabelle zeigt die Anzahl der Versuchspersonen und den durchschnittlichen Muskelaufbau pro Versuchsperson in diesen vier Studios:

	Versuchspersonen	durchschnittlicher Muskelaufbau
Studio 1	10	$\bar{X}_1$
Studio 2	20	$\bar{X}_2$
Studio 3	30	$\bar{X}_3$
Studio 4	30	$\bar{X}_4$

Nun wollen Sie den Erwartungswert des Muskelaufbaus  $E(X)$  bestimmen (Die  $\bar{X}_1, \bar{X}_2, \bar{X}_3, \bar{X}_4$  sind Mittelwerte von jeweils unabhängig und identisch verteilten  $X$ ).

- Ist  $\bar{X}_1$  ein erwartungstreuer Schätzer für den Erwartungswert des Muskelaufbaus  $E(X)$ ?
- Hängt Ihre Antwort auf Frage (a) von der Verteilung von  $X$  ab?
- Geben Sie einen effizienten (wirksamsten) erwartungstreuen Schätzer für den Erwartungswert des Muskelaufbaus pro Person  $E(X)$  an (mit Formel).

Sie stellen Tachometer her, die die Geschwindigkeit von Kraftfahrzeugen messen. Ein Automobilclub interessiert sich für die Geschwindigkeit, die Ihre Tachometer im Durchschnitt bei einer tatsächlichen Geschwindigkeit von 50 km/h anzeigen. Sie wissen, dass in diesem Bereich die Standardabweichung der Anzeige Ihrer Tachometer 5 km/h beträgt. Außerdem gehen Sie davon aus, dass der Fehler der Anzeige Ihrer Tachometer normalverteilt ist. Sie nehmen 100 Tachometer aus der laufenden Produktion, und stellen fest, dass diese Tachometer bei einer tatsächlichen Geschwindigkeit von 50 km/h im Mittel 55 km/h anzeigen.

- Geben Sie die Unter- und Obergrenzen des 99% Konfidenzintervalls für den Erwartungswert der durchschnittlich angezeigten Geschwindigkeit an. Schreiben Sie Ihr Ergebnis als R Ausdruck auf und verwenden Sie dabei, dass `qnorm(0.99)` das 99% Quantil der Normalverteilung, `qnorm(0.98)` das 98% Quantil, sowie `qnorm(0.995)` das 99.5% Quantil der Normalverteilung ergibt. Verwenden Sie diese Ausdrücke in Ihrer Lösung.
- In Teilaufgabe (a) haben Sie angenommen, dass die Standardabweichung bekannt ist. Wie ändert sich Ihr Ergebnis, wenn Sie mit der empirischen Standardabweichung rechnen. R gibt Ihnen die empirische Standardabweichung als `sd(x)` an.
- Eigentlich gehen Sie von einem Erwartungswert für die angezeigte Geschwindigkeit von 56 km/h aus. Stellen Sie eine Nullhypothese auf, berechnen Sie eine Teststatistik für den zweiseitigen Test (als R-Kommando und mit der empirischen Standardabweichung), und rechnen Sie einen  $p$ -Wert aus (ebenfalls als R-Kommando).

Sie führen Ihre Berechnungen in R durch. Der Vektor  $x$  enthält die möglichen Realisierungen der Zufallsvariablen  $X$ , der Vektor  $w$  ist genauso lang wie  $x$  und enthält die Wahrscheinlichkeiten der jeweiligen Realisierungen. Nehmen Sie an, dass die Variablen ERW und VAR den korrekten Erwartungswert, bzw. die Varianz von  $X$  enthalten. Berechnen Sie Erwartungswert und Varianz der transformierten Zufallsvariable  $Y = 5 \cdot X - 2$ . Kreuzen Sie die richtige Lösung an (Mehrere richtige Antworten sind möglich):

- $E(Y) = (x*5-2)*w$
- $E(Y) = \text{mean}(x)*5-2$
- $E(Y) = \text{mean}(x)*5/\text{VAR}-2$
- $E(Y) = \text{ERW}*5-2$
- $E(Y) = \text{mean}(x)*5/\text{VAR}^2-2$
- $\text{var}(Y) = \text{VAR}*5-2$
- $\text{var}(Y) = \text{VAR}^2*5-2$
- $\text{var}(Y) = \text{VAR}*5^2$
- $\text{var}(Y) = \text{sqrt}(\text{VAR})*5^2$
- $\text{var}(Y) = \text{VAR}*sqrt(5)$

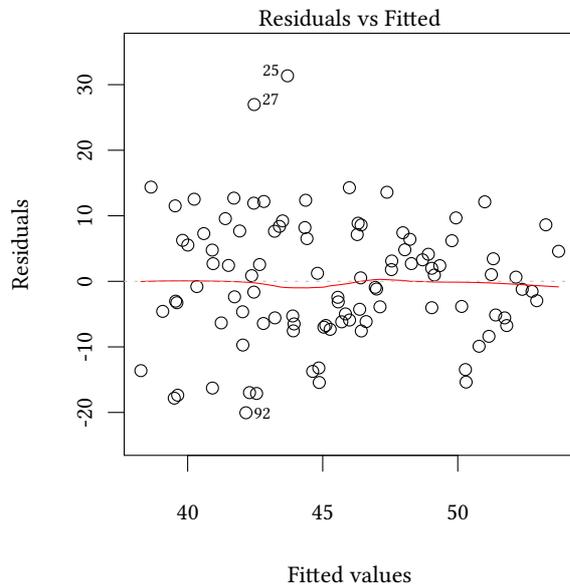
Sie betreiben eine Baumschule und versuchen, das Wachstum Ihrer Bäume mit unterschiedlichen Düngemitteln  $A, B$ , und  $C$  zu steigern. Sie düngen 100 Bäume mit unterschiedlichen Kombinationen der drei Düngemittel und messen die Größe nach einem Jahr. In Ihrem Datensatz ist  $y$  die Größe des Baums, und  $a, b$  und  $c$  sind die Mengen der Düngemittel. Sie berechnen in R eine lineare Regression der abhängigen Variablen  $y$  auf die unabhängigen  $a, b$ , und  $c$ . Den Output sehen Sie links, einen diagnostischen Plot der Residuen über die "fitted values" rechts:

```

Call: lm(formula = y ~ a + b + c)

            Estimate Std. Error t value Pr(>|t|)
(Intercept)  52.815     3.033   17.412 < 2e-16 ***
a              1.513     3.360    0.450  0.653
b             -1.062     3.372   -0.315  0.754
c            -14.899     3.633   -4.101 8.62e-05 ***
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 9.571 on 96 degrees of freedom
Multiple R-Squared:  0.1533, Adjusted R-squared:  0.1269
F-statistic: 5.796 on 3 and 96 DF, p-value: 0.001101

```



- Betrachten Sie den diagnostischen Plot. Hilft Ihnen dieser Plot zu entscheiden, ob Annahmen des klassischen Regressionsmodells erfüllt oder verletzt sind? Falls ja, welche Annahmen sind erfüllt bzw. verletzt bzw. können anhand der Graphik nicht beurteilt werden?
- Gehen Sie nun davon aus, dass alle Annahmen des Regressionsmodells erfüllt sind. Um welchen Betrag steigt die Größe Ihrer Bäume etwa, wenn Sie den Anteil des Düngemittels C um eine Einheit erhöhen?
- Ein befreundeter Gärtner behauptet, die von Ihnen eingesetzten Düngemittel hätten überhaupt keine Wirkung. Falls Sie eine Wirkung beobachtet hätten, sei das purer Zufall. Können Sie dieser Ansicht etwas entgegenhalten? Wenn ja, für welche Düngemittel? Begründen Sie Ihre Antwort.
- Durch ein Versehen wurden 100 neugepflanzte Bäume mit jeweils 10 Einheiten des Düngemittels B behandelt. Erwarteten Sie, dass diese Bäume größer oder kleiner als unbehandelte Bäume werden? Wenn ja, um wieviel werden sie im Mittel größer oder kleiner?