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!

Es gibt unterschiedliche Versionen — Sie finden die richtigen Antworten jeweils unter verschiedenen Buchstaben (a,b,c,d,e), die Antworten sind aber in allen Versionen die gleichen.

Aufgabe: Sie interessieren sich für den Erwartungswert von X. Ihre Stichprobe enthält 9 unabhängige Beobachtungen: $x_1, \ldots, x_9$. Welche Schätzfunktionen für $E(X)$ sind erwartungstreu? (mehrere Antworten möglich, 15 Punkte)

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: Welche Schätzfunktionen für $E(X)$ sind effizient?
(mehrere Antworten möglich, 5 Punkte)

2a: x_9
2b: $2x_1 + 2x_7 - 3x_9$
2c: $\frac{1}{9}\sum_i^9 x_i$
2d: $\sum_i^9 x_i$
2e: $x_1 + x_7 - x_9$

Aufgabe: Welche Schätzfunktionen für $E(X)$ sind effizienter?
(mehrere Antworten möglich, 15 Punkte)

3a: x_9 ist effizienter als $x_1 + x_7 - x_9$
3b: x_1 ist effizienter als $(x_2 + x_3)/2$
3c: $\frac{1}{9}\sum_i^9 x_i$ ist effizienter als $\frac{1}{7}\sum_i^7 x_i$
3d: x_3 ist effizienter als $\frac{1}{7}\sum_i^7 x_i$
3e: $x_1 + x_7 - x_9$ ist effizienter als $2 \cdot x_9 - x_8$

Aufgabe: Eine Zufallsvariable ist wie folgt verteilt: $P(X=1)=\theta$, $P(X=3)=\theta$, $P(X=5)=1-2\theta$. Es gilt $0 \le \theta \le \frac{1}{2}$. Ihre Stichprobe enthält zwei Beobachtungen: 1 und 3. Was ist der Maximum-Likelihood-Schätzer für θ?
(12 Punkte)

4:

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

Die Likelihoodfunktion ist θ^2. Wenn man über den Definitionsbereich $0 \le \theta \le \frac{1}{2}$ maximiert, liegt das Maximum bei $\theta = \frac{1}{2}$

Aufgabe: Ihre Stichprobe enthält nun drei Beobachtungen: 1, 3 und 5. Was ist der Momentenschätzer auf Basis des ersten Moments für θ?
(8 Punkte)

5:

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

Das erste Moment (der Mittelwert) ist 3. Das theoretische Moment ist $\theta \cdot 1 + \theta \cdot 3 + (1 - 2 \cdot \theta) \cdot 5 = -6\theta + 5$ Wir lösen also $-6\theta + 5 = 3 \to \theta = 1/3$

Aufgabe: Sie wollen mit einem Normal Q-Q Plot überprüfen, ob eine Variable normalverteilt ist. Letzteres ist der Fall, wenn... (5 Punkte)

6a: alle Punkte im Plot etwa auf einer horizontalen Linie liegen
6b: alle Punkte im Plot etwa auf einer Linie liegen
6c: alle Punkte im Plot etwa auf einer Linie mit Steigung von ca. 1.96 liegen
6d: alle Punkte im Plot gleichmäßig um eine Horizontale streuen
6e: fast alle Punkte im Plot innerhalb einer Ellipse um den Mittelwert liegen

Aufgabe: 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?
(mehrere Antworten möglich, 10 Punkte)

7a: `200-qnorm(0.975)`
7b: `200+qnorm(0.025)`
7c: `200+100*qnorm(0.025)`
7d: `200+10*qt(0.025,df=9)`
7e: `200-10*qnorm(0.025)`

Aufgabe: Ihre Nullhypothese sei H_0, die Alternativhypothese sei H_1. Welche der folgenden Aussagen trifft zu?
(mehrere Antworten möglich, 10 Punkte)

8a: Der Fehler 1. Art gibt an, wie häufig H_0 angenommen wird, obwohl sie falsch ist.
8b: Der Fehler 2. Art gibt an, wie häufig H_0 angenommen wird, obwohl sie falsch ist.
8c: Der Fehler 1. Art gibt an, wie häufig H_1 abgelehnt wird, obwohl sie wahr ist.
8d: Der Fehler 2. Art gibt an, wie häufig H_1 angenommen wird, obwohl sie falsch ist.
8e: Der Fehler 1. Art gibt an, wie häufig H_0 abgelehnt wird, obwohl sie wahr ist.

Aufgabe: 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
(mehrere Antworten möglich, 10 Punkte)

9:

a $g > Q_{n-1}^t(0.95)$	b $g > Q_{n-1}^t(0.975)$	c $g > Q_n^t(0.975)$	d $g < Q^N(0.025)$	e $g < Q_{n-1}^t(0.975)$

Aufgabe: 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?
(mehrere Antworten möglich, 10 Punkte)

10a: Die Verfahren führen zu einem signifikant verschiedenen Output.
10b: Die gemessenen Unterschiede sind klein und deuten nicht auf einen signifikanten Unterschied hin.
10c: Mit den Ihnen vorliegenden Daten sollte man besser einen paarweisen Test durchführen.
10d: Bei einem kleineren Signifikanzniveau würde man einen signifikanten Unterschied finden.
10e: Bei einer so kleinen Stichprobe sollte man besser einen t-Test durchführen.

Aufgabe: Sie schätzen ein lineares Regressionsmodell um den Einfluss von drei Variablen, `x1`, `x2`,und `x3` auf `y` zu ermitteln. Sie erhalten folgendes Ergebnis:

```
lm(formula = y ~ x1 + x2 + x3)
Coefficients:
              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 `x2` 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? (10 Punkte)

11:

a 1.5	b 0.05	c anderer Wert	d 0.15	e 0.5

$|t| = |1.5 - 20/10| = 0.5$

Aufgabe: Welchen p-Wert erhalten Sie für den Test der obigen Hypothese? (6 Punkte)

12:

a 0.137	b anderer Wert	c 0.618	d 0.842	e 0.320

Hier braucht man den p-Wert der zu einer t-Statistik von 0.5 passt. Den finden wir in der Zeile von `x1`*: $p = 0.618$*

Aufgabe: Sie untersuchen die Wirksamkeit von zwei Marketingmaßnahmen mit Hilfe einer linearen Regression. Ihre abhängige Variable `y` ist der Umsatz. Die unabhängigen Variablen `x1` und `x2` sind jeweils 1 wenn Maßnahme 1 bzw. 2 eingesetzt wurde und 0 sonst. Sie erhalten folgendes Ergebnis:

```
lm(formula = y ~ x1 * x2)
Coefficients:
            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? (7 Punkte)

13:	a	b	c	d	e
	-3.2	1.8	anderer Wert	3.5	0.3

1.5+3.0+0.5-3.2=1.8

Aufgabe: 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:

	$\hat\beta$	$\hat\sigma$
β_0	50	10
β_P	5	1
β_X	2	2

Wie groß ist die geschätzte Preiselastizität? (10 Punkte)

14:	a	b	c	d	e
	2	1/5	2/5	anderer Wert	5

Da $Q = e^{\beta_0}\cdot P^{\beta_P}\cdot X^{\beta_X}$ *ist* $\frac{\partial Q}{\partial P} = e^{\beta_0}\cdot\beta_P P^{\beta_P-1}\cdot X^{\beta_X}$ *und mithin* $\frac{\partial Q}{\partial P}\frac{P}{Q} = \beta_P = 5$

Aufgabe: 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 Punkte)

15:	a	b	c	d	e
	0	1	100	anderer Wert	1/2

Der geeignete Test ist hier ein χ^2*-Kontingenztest. Die Tabelle der* x_{ij} *ist*

	A	*B*	*C*	
Männer	*8*	*8*	*24*	*40*
Frauen	*12*	*12*	*36*	*60*
	20	*20*	*60*	*100*

Wenn man die e_{ij} *ausrechnet, erhält man die gleiche Tabelle,* $\sum_{i=1}^n\sum_{j=1}^k\frac{(x_{ij}-e_{ij})^2}{e_{ij}} \sim \chi^2_{(n-1)\cdot(k-1)}$ *ist also Null.*

Aufgabe: Nehmen Sie an, Sie hätten in der obigen Aufgabe eine Teststatistik von 10 berechnet. Wie berechnen Sie den p-Wert für Ihren Test? (mehrere Antworten möglich, 5 Punkte)

16a: `pchisq(10,2)`
16b: `1-pchisq(10,6)`
16c: `pt(10,5,lower=FALSE)`
16d: `pnorm(-10)`
16e: `1-pchisq(10,2)`

Aufgabe: 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 β_1 ist (5 Punkte)

17:	a	b	c	d	e
	0	$\frac{1}{2}$	2	anderer Wert	$\frac{1}{5}$

$Y = \frac{1}{2}$

Ihr R^2 ist (3 Punkte)

18:	a	b	c	d	e
	0	$\frac{1}{2}$	1	anderer Wert	$\frac{1}{5}$

Der Fit von $Y = \frac{1}{2}$ *ist perfekt.*

Was passiert, wenn Z in das Modell aufgenommen wird? (mehrere Antworten möglich, 10 Punkte)

19a: Die Schätzung für β_1 verändert sich. Vorher war sie verzerrt, weil das Modell unterspezifiziert war.
19b: AIC sinkt
19c: R^2 steigt
19d: R^2 sinkt
19e: AIC steigt

Y wird durch X bereits perfekt erklärt, Z bringt keine Verbesserung, sondern nur eine Verschlechterung. Deshalb steigt das AIC.

---

maximal erreichbare Punktzahl: 171
davon durch Randomisieren erreichbar: 61.2
hinreichend: 100.7