Musterlösung

1. (**9 Punkte**) Lukas und Markus haben bisher immer "Feinste Mini-Brezeln 100g" des Herstellers Gammelbrot und Söhne zum Znüni gegessen. Vom ständigen Hungerklagen von Markus genervt schlägt Lukas nun vor auf "Rustikale Brez'n Bayrischer Art 100g" des Herstellers Wolpertinger Backwaren zu wechseln. Lukas vermutet, dass - obwohl beide Hersteller das selbe Gewicht angeben - die Brezeln der Firma Gammelbrot und Söhne leichter als das Konkurenzprodukt sind. Um dies zu untersuchen bringt Markus seine Küchenwaage mit und wiegt einige Brezeln beider Marken. Die Ergebnisse (alles in g) sind:

Brezeln von Gammelbrot	97	97	97	98	93
Brezeln von Wolpertinger	100	101	99	99	98

Nun wollen die beiden testen ob die Vermutung von Lukas richtig ist. Dazu nehmen sie an, dass die Gewichts-Differenzen normalverteilt mit Erwartungswert μ und Varianz σ^2 sind (mit Gewichtsdifferenz ist hier und im Folgenden gemeint: Gewicht einer Brezel von Gammelbrot - Gewicht einer Brezel von Wolpertinger).

Anmerkung: Die Aufgabe war unklar gestellt. Folgender Zusatz hat gefehlt: "Es werden von jedem Brezelhersteller 5 Brezeln unterschiedlicher Sorten gekauft (mit Mohn, Sonnenblumenkernen, Kürbiskernen, Lauge, und gemischt). In der Tabelle entspricht jede Spalte einer Sorte."

a) Handelt es sich um einen gepaarten oder einen ungepaarten Test?
Gepaart. (1 Punkt)

b) Geben Sie die Null-und die Alternativhypothese an.
$H_0 : \mu_D = 0$ (0,5 Punkte)
$H_A : \mu_D < 0$ (0,5 Punkte)
mit $D_i = X_i - Y_i$, $\mu_D = \mu_X - \mu_Y$,
X_i: Gewichte der Brezeln von Gammelbrot,
Y_i: Gewichte der Brezeln von Wolpertinger.

c) Geben Sie eine Schätzung für die Varianz σ^2 der Differenz an.

$$\bar{D} = \frac{1}{n}\sum_{i=1}^{n} D_i = \frac{1}{5}(-3-4-2-1-5) = -3$$

$$\begin{aligned}\hat{\sigma}_D^2 &= \frac{1}{n-1}\sum_{i=1}^{n}(D_i - \bar{D})^2 \qquad (0,5 \text{ Punkte})\\ &= \frac{1}{4}(-3-(-3))^2 + (-4-(-3))^2 + (-2-(-3))^2 + (-1-(-3))^2 + (-5-(-3))^2)\\ &= 2,5 \qquad (0,5 \text{ Punkte})\end{aligned}$$

d) Führen Sie den geeigneten t-Test durch: Bestimmen Sie den Wert der Teststatistik T, den Verwerfungsbereich für T und den Testentscheid. (Wenn sie obige Aufgabe nicht lösen konnten, benutzen Sie im Folgenden als Ersatzwert $\sigma^2 = 2.5$.)
Teststatistik:

$$\begin{aligned}T &= \frac{\sqrt{n}\bar{D}}{\hat{\sigma}_D} \qquad (0,5 \text{ Punkte})\\ &= \frac{\sqrt{5}\cdot(-3)}{\sqrt{2,5}} = -4,24 \qquad (0,5 \text{ Punkte})\end{aligned}$$

Verteilung von T unter H_0: $T \sim t_{n-1} = t_4$. (0,5 Punkte)
Verwerfungsbereich:

$$\begin{aligned}K &= (-\infty; t_{n-1;1-\alpha}] \qquad (0,5 \text{ Punkte})\\ &= (-\infty; t_{4;0.95}] = (-\infty; -2,132] \qquad (0,5 \text{ Punkte})\end{aligned}$$

Testentscheid: $T \in K \Rightarrow H_0$ wird verworfen. (0,5 Punkte)

e) Bestimmen Sie ein einseitiges 95%-Vertrauensintervall für μ.

$$\mathrm{VI} = \left(-\infty; \bar{D} + t_{n-1;1-\alpha}\frac{\hat{\sigma}_D}{\sqrt{n}}\right] \qquad (0,5 \text{ Punkte})$$
$$= (-\infty; -1,49] \qquad (0,5 \text{ Punkte})$$

f) Angenommen, der Wert σ^2 wäre nicht aus den Daten geschätzt, sondern bekannt: Wie lautet dann das einseitige 95%-Vertrauensintervall? Geben Sie eine kurze Erklärung für den Unterschied zu e)!

$$\mathrm{VI} = \left(-\infty; \bar{D} + z_{1-\alpha}\frac{\sigma_D}{\sqrt{n}}\right] \qquad (0,5 \text{ Punkte})$$
$$= (-\infty; -1,83] \qquad (0,5 \text{ Punkte})$$

Das Vertrauensintervall wird kleiner da die Unsicherheit aus der Schätzung von σ wegfällt. (1 Punkt)

2. (9 Punkte) Das Pharmaunternehmen Life Co. hat ein neues Medikament zur Bekämpfung von ADHS entwickelt. Um die Wirksamkeit festzustellen wurde das Medikament mit $n = 10$ Patienten getestet. Die derzeitige Standardmethode zeigt bei 30% der behandelten Patienten eine Wirkung.

a) Angenommen das neue Medikament ist genauso wirksam wie die Standardmethode, wie gross ist die Wahrscheinlichkeit, dass die Behandlung bei genau 2 Patienten eine Wirkung zeigt? Wie gross ist die Wahrscheinlichkeit, dass sie bei höchstens 2 Patienten eine Wirkung zeigt?
(1 Punkt)

$$P(X = 2) = \binom{10}{2} 0.3^2 0.7^8 = 0.23$$
$$P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2) = 0.7^{10} + \binom{10}{1} 0.3^1 0.7^9 + \binom{10}{2} 0.3^2 0.7^8 = 0.38$$

b) Die Behandlung mit dem neuen Medikament war bei 4 Patienten erfolgreich. Führen Sie einen einseitigen Hypothesentest durch um festzustellen ob das neue Medikament wirksamer ist als die Standardmethode (bei einem Signifikanzniveau von 5%). Geben Sie explizit alle Schritte an.
(3 Punkte)

1. Modell: X ist die Anzahl erfolgreich behandelter Patienten, $X \sim \mathrm{Bin}(10, \pi)$.
2. Die Nullhypothese ist $H_0 : \pi = 0.3$, die Alternative ist $H_A : \pi > 0.3$.
3. Die Teststatistik ist X: $P(X = x|H_0) = \binom{10}{x} 0.3^x 0.7^{10-x}$.
4. Das Signifikanzniveau ist $\alpha = 0.05$.
5. Verwerfungsbereich:

	$x = 5$	$x = 6$	$x = 7$	$x = 8$	$x = 9$	$x = 10$
$P(X \geq x)$	0.1503	0.0473	0.0106	0.0016	0.0001	5.9×10^{-6}

   Daher ist der Verwerfungsbereich $K = \{6, 7, 8, 9, 10\}$.
6. Testentscheid: Da $4 \notin K$ wird H_0 nicht verworfen. Eine erhöhte Wirksamkeit des neuen Medikaments kann nicht nachgewiesen werden.

c) Wie ist die *Macht* eines Hypothesentests definiert? Geben Sie die Macht an für den Test H_0 : $\pi = 0.3$ vs. $H_A : \pi = 0.6$ (π ist die Wirksamkeit).
(1 Punkt) Die Macht eines Tests ist die Wahrscheinlichkeit, dass die Nullhypothese verworfen wird, wenn die Alternative stimmt: $P(X \in K|H_A)$. (Alternativ: $\text{Macht} = 1 - P(\text{Fehler 2. Art}) = 1 - P(X \notin K|H_A)$)
Im konkreten Fall:

$$\text{Macht} = \binom{10}{6} 0.6^6 0.4^4 + \binom{10}{7} 0.6^7 0.4^3 + \binom{10}{8} 0.6^8 0.4^2 + \binom{10}{9} 0.6^9 0.4 + 0.6^1 0 = 0.6331$$

d) An einem zweiten Test nehmen 200 Patienten teil. Diesmal zeigt das Medikament bei 80 Patienten eine Wirkung. Führen Sie erneut einen einseitigen Hypothesentest durch und berechnen Sie anschliessend das (einseitige!) 95%-Intervall für π.

(4 Punkte) Hypothesentest (da n gross ist verwenden wir die Normalapproximation):

1. Modell: X_k ist 1, falls der k-te Patient erfolgreich behandelt wurde, und sonst 0. $X \sim \text{Bernoulli}(\pi)$.
2. Die Nullhypothese ist $H_0 : \pi = 0.3$, die Alternative ist $H_A : \pi > 0.3$.
3. Die Teststatistik ist $\bar{X}_n = \frac{1}{n}\sum_{k=1}^{n} X_k$. Verteilung der Teststatistik unter H_0:
$$\bar{X}_n \sim \mathcal{N}(0.3, \sigma_X^2/n) \quad \text{(Normalapproximation)}, \quad \sigma_X^2 = 0.3 \cdot 0.7 = 0.21$$
4. Das Signifikanzniveau ist $\alpha = 0.05$.
5. Verwerfungsbereich: Wir transformieren zuerst $Z = \frac{\bar{X}_n - 0.3}{\sigma_X/\sqrt{n}}$, sodass $Z \sim \mathcal{N}(0,1)$. Dann:
$$P(\bar{X}_n \geq c) = P\left(Z \geq \frac{c - 0.3}{\sigma_X/\sqrt{n}}\right) = 1 - \Phi\left(\frac{c-0.3}{\sigma_X/\sqrt{n}}\right) < \alpha.$$
Somit ist der Verwerfungsbereich:
$$K = \left[\frac{\sigma_X}{\sqrt{n}}\Phi^{-1}(1-\alpha) + 0.3, \infty\right) = \left[\sqrt{\frac{0.21}{200}} \cdot 1.645 + 0.3, \infty\right) = [0.35, \infty)\,.$$
6. Testentscheid: Da $\bar{x}_n = \frac{80}{200} = 0.4 \in K$ wird H_0 verworfen. Die Wirksamkeit des neuen Medikaments ist signifikant höher als die der Standardmethode.

Vertrauensintervall (mit Normalapproximation):
$$I \approx \left[\bar{x}_n - 1.645\sqrt{\bar{x}_n\,(1-\bar{x}_n)\,\frac{1}{n}}, \infty\right) = [0.34, \infty)$$

3. (7 Punkte) Ein Mass für die technische Entwicklung (bzw. Rückständigkeit) eines Landes ist der Anteil der arbeitenden Bevölkerung im Agrarsektor. Es soll nun der Einfluss auf das Pro-Kopf-Einkommen untersucht werden. Dafür liegen Daten aus 20 Ländern aus dem Jahr 1960 vor. Folgendes Modell wurde angepasst:
$$\text{pcinc}_i = \beta_0 + \beta_1 \cdot \text{agr}_i + \varepsilon_i, \qquad \varepsilon_i \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2),$$
wobei pcinc_i das durchschnittliche Pro-Kopf-Einkommen (in USD) und agr_i der Anteil der arbeitenden Bevölkerung im Agrarsektor (in $\% \cdot 100$) sind. D.h. wenn z.B. 30% der arbeitenden Bevölkerung im Land i im Agrarsektor tätig sind, dann ist $\text{agr}_i = 30$. Der (unvollständige) Regressionsoutput sieht wie folgt aus:

```
Residuals:
    Min      1Q  Median      3Q     Max
-351.60 -211.36   -3.75  105.93  590.10

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 1317.905        ???  12.406 2.95e-10
agr          -18.858        ???  -5.554 2.84e-05
---

Residual standard error: 273.1 on  ? degrees of freedom
Multiple R-squared: 0.6315,       Adjusted R-squared: 0.611
F-statistic: 30.85 on 1 and  ? DF,  p-value: 2.845e-05
```

1) Was ist $\hat{\beta}_1$?

a) 1317.905

b) 12.406
c) -18.858
d) -5.554

c (-18.858)

2) Was ist der Standardfehler von $\hat{\beta_0}$?
   a) 0.009
   b) 0.295
   c) 3.395
   d) 106.230

   d (106.230)

3) Mit wievielen Freiheitsgraden wurde der "residual standard error" berechnet?
   a) 18
   b) 22
   c) 2
   d) 20

   a (18)

4) Berechnen Sie das 99%-Konfidenzintervall für β_1 (ohne Normalapproximation).
   a) $[-22.25, -15.46]$
   b) $[-25.99, -11.73]$
   c) $[-27.52, -10.19]$
   d) $[-28.63, -9.09]$

   d ($[-28.63, -9.09]$)

5) Kann die Nullhypothese $H_0 : \beta_0 = 0$ auf dem 5% Signifikanzniveau verworfen werden?
   a) Ja.
   b) Nein.
   c) Keine Angabe möglich.

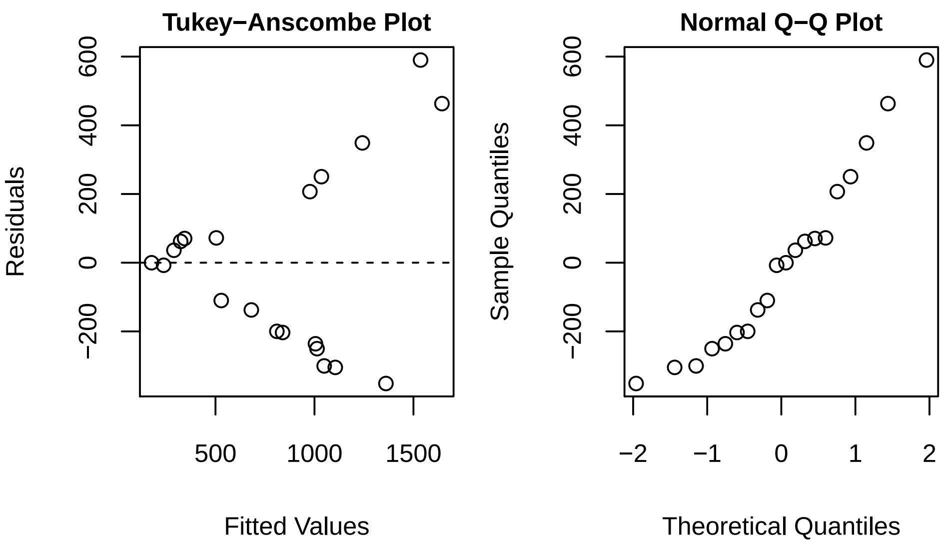
   a (Ja)

6) In der Schweiz betrug 1960 das durchschnittliche Pro-Kopf-Einkommen 1361 USD, der Agrarsektor hatte einen Anteil von 11%. Wie hoch ist das Residuum für diesen Datenpunkt in unserem Modell?
   a) 45.17
   b) 250.53
   c) 1110.47
   d) 1315.83

   b (250.53)

7) Betrachten Sie die nachfolgenden Plots. Welche der folgenden Aussagen trifft zu?
   a) Alle Modellannahmen sind erfüllt.
   b) Die Fehlervarianz scheint nicht konstant zu sein, aber die Normalverteilungsannahme ist plausibel.
   c) Die Fehlervarianz scheint konstant zu sein, aber die Normalverteilungsannahme scheint nicht zuzutreffen.
   d) Sowohl konstante Fehlervarianz als auch Normalverteilungsannahme treffen nicht zu.

   b (Die Fehlervarianz scheint nicht konstant zu sein, aber die Normalverteilungsannahme ist plausibel.)



4. Je einen Punkt für:

(8 Punkte)

1) c (Median $= 6$, Mean $= 43.4$)
2) c ($\text{odds}(A) = 1/2$)

$$\text{odds}(A) = \frac{P(A)}{1 - P(A)} = \frac{1/3}{1 - 1/3} = \frac{1}{2}$$

3) a (Ja)
Das 95%-Vertrauensintervall beinhaltet alle Werte μ_0, bei denen die Nullhypothese $H_0 : \mu = \mu_0$ auf dem 5% Signifikanzniveau nicht verworfen wird. Da 500 nicht im gegebenen Vertrauensintervall liegt, wird die Nullhypothese $\mu = 500$ verworfen.
4) d (1080)
Die Formel für das zweiseitige 95%-Vertrauensintervall mit Normalapproximation ist:

$$I \approx \hat{p} \pm 1.96\sqrt{\frac{\hat{p}}{n}\,(1 - \hat{p})}$$

Die Unsicherheit des Vertrauensintervall ist somit proportional zu $1/\sqrt{n}$. Um die Unsicherheit auf ein Drittel zu reduzieren, werden daher 9-mal so viele Lose benötigt.
5) a (Auf dem 10%-Niveau verwerfen und auf dem 5%-Niveau nicht verwerfen.)
Die Nullhypothese wird verworfen, wenn $p < \alpha$ (α ist Signifikanzniveau).
6) b (1.8%)

$$\begin{aligned} P(K|\text{pos. Test}) &= \frac{P(\text{pos. Test}|K)P(K)}{P(\text{pos. Test})} = \frac{P(\text{pos. Test}|K)P(K)}{P(\text{pos. Test|K})P(K) + P(\text{pos. Test|K}^c)P(K^c)} \\ &= \frac{0.9 \cdot 0.001}{0.9 \cdot 0.001 + 0.05 \cdot 0.999} = 0.018 \end{aligned}$$

7) c ($1/k$)
Die Likelihood L ist gegeben durch:

$$\begin{aligned} &L(\pi|k) = (1 - \pi)^{k-1}\pi \\ &\frac{\partial \log L}{\partial \pi}(\hat{\pi}) = 0 \quad \Rightarrow \quad \hat{\pi} = \frac{1}{k} \end{aligned}$$

8) f (2.3%)

Die Verteilung der Überstunden $S = \sum X_i$ kann mit dem zentralen Grenzwertsatz approximiert werden (da $n = 256 \gg 1$):

$$S \sim \mathcal{N}(n\mu_X, n\sigma_X^2) = \mathcal{N}(128, 256) \quad \Rightarrow \quad Z = \frac{S - 128}{16} \approx \mathcal{N}(0,1)$$

$$P(S > 160) = P\left(Z > \frac{160 - 128}{16}\right) = P(Z > 2) = 1 - P(Z \leq 2) = 1 - \Phi(2) = 0.0228$$