Lösungsvorschlag zur 13. Übung

G40

Verwendetes Testverfahren: Vorzeichentest
Die Wertepaare $(x_1, y_1), \ldots, (x_{12}, y_{12})$ der mit den Verfahren A und B gemessenen Blutalkoholspiegel werden als Realisierungen der unabhängigen zweidimensionalen Zufallsvariablen $(X_1, Y_1), \ldots, (X_{12}, Y_{12})$ aufgefasst. Dann sind die Zufallsvariablen D_i, $i = 1, \ldots, 12$, mit

$$D_i := \begin{cases} 1 & \text{, falls } X_i \geq Y_i \\ 0 & \text{, sonst} \end{cases}$$

unabhängig.

Nullhypothese: $H_0 : P(D_i = 1) = P(D_i = 0) = \frac{1}{2} \quad, i = 1, \ldots, 12$

Testgröße: $V\left((X_1, Y_1), \ldots, (X_{12}, Y_{12})\right) = \sum_{i=1}^{12} D_i$
Unter H_0 ist die Testgröße $B(12, \frac{1}{2})$-verteilt.

Kritischer Bereich: $V < 3$ oder $V > 9$, da $P(V < 3 \text{ oder } V > 9) = 2P(V < 3) = 0.038 \leq 0.05$

Wert der Testgröße: Für $d_1, \ldots, d_{12}$ erhält man aus der gegebenen Stichprobe: 110100010010
$\Longrightarrow\ V\left((x_1, y_1), \ldots, (x_{12}, y_{12})\right) = 5$

Entscheidung: Wegen $3 \leq 5 \leq 9$ wird gegen H_0 nichts eingewendet.
Weder der Zweistichprobentest von Wilcoxon-Mann-Whitney noch der Run-Test von Wald und Wolfowitz können hier angewendet werden, da paarige Beobachtungsergebnisse vorliegen, d. h. für festes i können X_i und Y_i nicht als unabhängig angenommen werden.

G41

a) Ordnet man die Beobachtungserbebnisse der Größe nach, so erhält man:

$$X X X Y\ X X Y\ X X Y\ X Y\ X Y Y Y Y\ X Y Y$$

Nullhypothese: $H_0 : F = G \quad H_1 : F > G$ oder $F < G$

Testgröße: $U(X_1, \ldots, X_{10}, Y_1, \ldots, Y_{10}) = \sum_{i=1}^{10} \sum_{j=1}^{10} Z_{ij}$ mit

$$Z_{ij} = \begin{cases} 1 & \text{, falls } X_i > Y_j \\ 0 & \text{, falls } X_i \leq Y_j \end{cases}$$

d.h. die Anzahl der Inversionen in der sortierten Folge der Zufallsvariablen.

Wert der Testgröße: $U(x_1, \ldots, x_{10}, y_1, \ldots, y_{10}) = 1 \cdot 7 + 1 \cdot 5 + 1 \cdot 3 + 1 \cdot 2 + 4 \cdot 1 = 21$

Kritischer Bereich: Zunächst näherungsweise Bestimmung der benötigten Quantile:
U ist Näherungsweise $N\left(\frac{10 \cdot 10}{2}, \frac{1}{12} \cdot 10 \cdot 10 \cdot (10 + 10 + 1)\right)$-verteilt, d.h. $N(50, 175)$-verteilt.
Mit $u_{0.975} = 1.96$ folgt: H_0 ablehnen, falls $U < 50 - 1.96 \cdot \sqrt{175} = 24.07$ oder $U > 50 + 1.96 \cdot \sqrt{175} = 75.93$ ist.

Entscheidung: Wegen $U = 21 < 24.07$ wird H_0 abgelehnt.

b) Nullhypothese: $H_0 : F = G \quad H_1 : F \neq G$

Testgröße: Anzahl $R(X_1, \ldots, X_{10}, Y_1, \ldots, Y_{10})$ der Runs in der sortierten Folge der Zufallsvariablen.

Wert der Testgröße: $R(x_1, \ldots, x_{10}, y_1, \ldots, y_{10}) = 12$

Kritischer Bereich: Zunächst näherungsweise Bestimmung der benötigten Quantile:
R ist näherungsweise $N\left(1 + \frac{2 \cdot 10 \cdot 10}{10+10}, \frac{2 \cdot 10 \cdot 10 \cdot (2 \cdot 10 \cdot 10 - 10 - 10)}{(10+10)^2 (10+10-1)}\right)$-verteilt, d.h. $N(11, \frac{90}{19})$-verteilt.
Mit $u_{0.05} = -1.64$ folgt: H_0 ablehnen, falls $R < 11 - 1.64 \cdot \sqrt{\frac{90}{19}} \approx 7.4307$ gilt.

Entscheidung: Wegen $7.4307 < 12 = R$ wird H_0 nicht abgelehnt.

G42

Verwendetes Testverfahren: Kruskal-Wallis-Test

Voraussetzungen: $X_1, \ldots, X_4$, $Y_1, \ldots, Y_4$, $Z_1, \ldots, Z_4$ sind unabhängig und $X_1, \ldots, X_4$ sowie $Y_1, \ldots, Y_4$ und $Z_1, \ldots, Z_4$ sind jeweils identisch verteilt.

Nullhypothese: $H_0: \quad F_X = F_Y = F_Z$

Wert der Testgröße: Mittlerer Rang von $x_1, \ldots, x_4$: $r_x = \frac{1}{4}(1+3+4+5) = \frac{13}{4}$,
mittlerer Rang von $y_1, \ldots, y_4$: $r_y = \frac{1}{4}(2+6+8+10) = \frac{26}{4}$,
mittlerer Rang von $z_1, \ldots, z_4$: $r_z = \frac{1}{4}(7+9+11+12) = \frac{39}{4}$.

$$\begin{aligned} K(x_1, \ldots, x_4, y_1, \ldots, y_4, z_1, \ldots, z_4) &= \frac{12}{12 \cdot 13}\left(4(3.25-6.5)^2 + 4(6.5-6.5)^2 + 4(9.75-6.5)^2\right) \\ &= \frac{4}{13}(3.25^2 + 3.25^2) = 6.5 \end{aligned}$$

Entscheidung: Wegen $K = 6.5 > 4.60 = \chi^2_{2;0.9}$ wird die Nullhypothese auf dem Niveau 0.1 verworfen.

G43

a) Zu den Meßwerten erhält man:

$$\sum_{i=1}^{9} x_i = 48.5 \qquad \sum_{i=1}^{9} x_i^2 = 565.85 \qquad \sum_{i=1}^{9} x_i \cdot y_i = 4187.6$$

$$\sum_{i=1}^{9} y_i = 1424 \qquad \sum_{i=1}^{9} y_i^2 = 273588$$

Daraus folgt: $\bar{x} = 5.39 \qquad \bar{y} = 158.22$

$$s_x^2 = \frac{1}{8}\left(\sum_{i=1}^{9} x_i^2 - 9\bar{x}^2\right) = 38.06$$

$$s_y^2 = \frac{1}{8}\left(\sum_{i=1}^{9} y_i^2 - 9\bar{y}^2\right) = 6034.94$$

$$r_{xy} = \frac{\frac{1}{8}\left(\sum_{i=1}^{9} x_i y_i - 9\bar{x}\bar{y}\right)}{\sqrt{s_x^2\, s_y^2}} = -0.9096$$

b) Für die Koeffizienten der Regressionsgeraden $y = ax + b$ ergeben sich folgende Schätzer:

$$\hat{a} = \frac{sxy}{ssx} = \frac{\sum_{i=1}^{9} x_i y_i - 9\bar{x}\bar{y}}{\sum_{i=1}^{9} x_i^2 - n\bar{x}^2} = \frac{-3487.65}{304.38} = -11.46$$

$$\hat{b} = \bar{y} - \hat{a}\bar{x} = 219.99$$

Regressionsgerade: $y = -11.46x + 219.99$

c) Wir gehen wie in Beispiel 3.64 (4) auf Seite 195 vor. Zunächst berechnen wir $y(-3.9) = 264.68$,

$$t_* = t_{7;0.975} = 2.365$$

$$s_* = \sqrt{1 + \frac{1}{9} + \frac{(-3.9-5.39)^2}{304.38}} = 1.181$$

$$ssr = ssy(1 - r_{xy}^2) = 8335.49$$

$$t_* \cdot s_* \sqrt{\frac{ssr}{7}} = 2.365 \cdot 1.181 \cdot 34.51 = 96.39$$

Wir erhalten somit $[264.68 - 96.39; 264.68 + 96.39] = [168.29; 361.07]$ als Prognoseintervall.
Der durchschnittliche Verbrauch von $246dm^3$ liegt also unter dem alten Erwartungswert von $264.68dm^3$. Da er jedoch noch deutlich im Prognoseintervall liegt, ist der Unterschied nicht signifikant.

G44

a) Berechnung der M-L-Schätzwerte:

$$\hat{a} = \frac{sxy}{ssx} = \frac{\sum_{i=1}^{20} x_i y_i - 20 \cdot \bar{x} \cdot \bar{y}}{\sum_{i=1}^{20} x_i^2 - 20\bar{x}^2} = \frac{240}{300} = 0.8$$

$$\hat{b} = \bar{y} - \hat{a} \cdot \bar{x} = 61.625 - 0.8 \cdot 168.25 = -72.975$$

$$\hat{\sigma}^2 = \frac{ssr}{n} = \frac{ssy \cdot \left(1 - \frac{sxy^2}{ssx \cdot ssy}\right)}{n} = \frac{88}{20} = 4.4$$

b) Nullhypothese: $H_0 : a = 1 \quad H_1 : a \neq 1$
Vorgehensweise wie im Beispiel 3.64 (3) auf Seite 194:
H_0 ist zu verwerfen, falls

$$(\hat{a} - 1)^2 > t_*^2 \cdot \frac{ssr}{(n-2) \cdot ssx} \text{ mit } t_* = t_{n-2;1-\frac{\alpha}{2}}$$

$$t_* = t_{18;0.95} = 1.734$$

$$(\hat{a} - 1)^2 = (0.8 - 1)^2 = 0.04$$

$$t_*^2 \cdot \frac{ssr}{18 \cdot ssx} = 1.734^2 \cdot \frac{88}{18 \cdot 300} = 0.049$$

Da $0.04 < 0.049$ ist, wird H_0 nicht verworfen.

c)

$$L(b; (x_1, y_1), \ldots, (x_n, y_n)) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(y_i - x_i - b)^2}{2\sigma^2}} = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \cdot e^{\frac{-\sum_{i=1}^{n}(y_i - x_i - b)^2}{2\sigma^2}}$$

Die M-L-Funktion wird (für beliebiges σ^2) maximal, wenn $\sum_{i=1}^{n}(y_i - x_i - b)^2$ minimal ist.

$$\frac{\partial}{\partial b} \sum_{i=i}^{n} (y_i - x_i - b)^2 = -2 \cdot \sum_{i=1}^{n} (y_i - x_i - b)$$

$$\frac{\partial^2}{\partial b^2} \sum_{i=1}^{n} (y_i - x_i - b)^2 = 2n > 0$$

Berechnung des Maximums:

$$-2\sum_{i=1}^{n}(y_i - x_i - b) \stackrel{!}{=} 0 \implies \sum_{i=1}^{n}(y_i - x_i) = \sum_{i=1}^{n} b = n \cdot b \implies \tilde{b} = \frac{1}{n}\sum_{i=1}^{n}(y_i - x_i) = \frac{1}{n}\sum_{i=1}^{n} y_i - \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{y} - \bar{x}$$

ist der M-L-Schätzwert.
hier: $\tilde{b} = 61.625 - 168.25 = -106.625$

G45

Zusätzliche Annahme:
$X_1, \ldots, X_4 \sim N(\mu_x, \sigma^2)$, $Y_1, \ldots, Y_4 \sim N(\mu_y, \sigma^2)$ und $Z_1, \ldots, Z_4 \sim N(\mu_z, \sigma^2)$

Nullhypothese: $H_0 : \mu_x = \mu_y = \mu_z$

$$\bar{x} = \frac{1}{4}\sum_{i=1}^{4} x_i = 10 \qquad \bar{y} = 15 \qquad \bar{z} = 20 \qquad \bar{g} = \frac{\bar{x} + \bar{y} + \bar{z}}{3} = 15$$

$$sst = n_1 \cdot (\bar{x} - \bar{g})^2 + n_2 \cdot (\bar{y} - \bar{g})^2 + n_3 \cdot (\bar{z} - \bar{g})^2 = 4 \cdot [(10-15)^2 + (15-15)^2 + (20-15)^2] = 200$$

$$\begin{aligned} sse &= (4-10)^2 + (11-10)^2 + (12-10)^2 + (13-10)^2 + (10-15)^2 + (14-15)^2 \\ &+ (17-15)^2 + (19-15)^2 + (16-20)^2 + (18-20)^2 + (20-20)^2 + (26-20)^2 \\ &= 36 + 1 + 4 + 9 + 25 + 1 + 4 + 16 + 16 + 4 + 0 + 36 \\ &= 152 \end{aligned}$$

$$T(x_1, \ldots, x_4, y_1, \ldots, y_4, z_1, \ldots, z_4) = \frac{\frac{1}{2}sst}{\frac{1}{9}sse} = \frac{9 \cdot 200}{2 \cdot 152} = 5.921$$

Da $T = 5.921 > 3.01 = F_{2;9;0.9}$ ist, wird H_0 auf dem Niveau 0.1 verworfen.