Bachelorprüfung: Statistik Musterlösung

1. **a)** **(1P)** Gepaart, da jeweils dieselbe Person einmal mit dem einen Wachs und einmal mit dem anderen Wachs fährt.

b) **(1P)** $H_0 : \mu_X = \mu_Y$, $H_A : \mu_X \neq \mu_Y$

c) **(1P)** $\frac{-0.33\sqrt{8}}{0.51} = -1.83$ und da $1.83 < t_{7,97.5\%} = 2.3646$, wird H_0 beibehalten.
Der Verwerfungsbereich ist

$$K = \left[-\infty, \frac{t_{7,97.5\%}0.51}{\sqrt{8}}\right] \cup \left[\frac{t_{7,97.5\%}0.51}{\sqrt{8}}, \infty\right] = [-\infty, -0.42637] \cup [0.42637, \infty].$$

Weil $-0.33 \notin K$, wird H_0 beibehalten.

d) **(2P)** $\left[-0.33 - \frac{2.3646 \cdot 0.51}{\sqrt{8}}, -0.33 + \frac{2.3646 \cdot 0.51}{\sqrt{8}}\right] = [-0.75637, 0.09637]$

e) **(2P)** Entweder: Mit zusätzlichem Wissen wird die Macht eines Testes grösser, daher wird der Verwerfungsbereich grösser.
Oder: Das 97.5%-Quantil von t_7 ist 2.3646 und das einer Standardnormalverteilung 1.96, vergleiche dann mit **c)**.

2. a) **(1P)** Bei zufälligen Antworten des Kandidaten hat dieser bei jeder Frage eine Erfolgswahrscheinlichkeit von 25%. Da die Richtigkeit seiner Antworten auf verschiedene Fragen zudem unabhängig ist, folgt die Anzahl X seiner richtigen Antworten einer Binomialverteilung **(1/2 P)** mit Parametern $n = 16$ (Stichprobengrösse) und $p = 0.25$ (Erfolgswahrscheinlichkeit) **(1/2 P)**.

b) **(1P)** Der Mittelwert beträgt $E[X] = np = 4$ **(1/2 P)**, die Varianz $Var[X] = np(1-p) = 3$ **(1/2 P)**.

c) **(1 1/2 P)** Die nützlichste stetige Verteilung hier ist die Normalverteilung **(1/2 P)**. Deren Parameter sollte man so wählen, dass sie denselben Mittelwert und dieselbe Varianz aufweist wie die zu approximierende Binomialverteilung. Also ist $\mu = E[X] = 50 \cdot 0.25 = 12.5$ **(1/2 P)** und $\sigma = \sqrt{Var[X]} = \sqrt{50 \cdot 0.25 \cdot 0.75} = 3.062$ **(1/2 P)**.

d) **(4 1/2 P)** Die Nullhypothese ist die Hypothese, deren Wahrheit uns am meisten interessiert:

$$H_0 : \mu = 12.5 \ (\mathbf{1/2P}) \ ,$$

und die Alternative dazu, die wir in Betracht ziehen, ist einseitig nach oben geöffnet, da alles Wissen vom Kandidaten zu einer höheren Erfolgswahrscheinlichkeit führen würde:

$$H_A : \mu > 12.5 \ (\mathbf{1/2P}) \ .$$

Unter der Nullhypothese ist die Standardabweichung von X gleich 3.062 (wie bereits gesehen); die Teststatistik ist der standardisisierte "Messwert":

$$Z = \frac{15 - \mu}{\sigma} = \frac{15 - 12.5}{3.062} = 0.816 \ (\mathbf{1P}) \ .$$

Der Verwerfungsbereich des Tests besteht aus jenem nach oben offenen Intervall, deren Wahrscheinlichkeit gerade noch 5% beträgt und in welchem daher die Erfolgsquoten liegen, die derart hoch sind, dass dadurch reines Raten unplausibel wird:

$$\begin{aligned} K &= [M, \infty) \text{ sodass } P(K) = 0.05 \text{ unter der Standard-Normalverteilung} \\ &= [\mu + \Phi^{-1}(0.95) * \sigma, \infty) = [17.54, \infty) \ (\mathbf{2P}) \ . \end{aligned}$$

Da der beobachtete Wert $X = 15$ nicht im Verwerfungsbereich liegt, wird hier die Nullhypothese *nicht verworfen* **(1/2 P)**.

e) **(2P)** Mit der neuen Erfolgswahrscheinlichkeit $p = 0.4$ ändern sich auch die Parameter der Normalapproximation: $\mu = 50 \cdot 0.4 = 20$ **(1/2 P)** sowie $\sigma = \sqrt{50 \cdot 0.4 \cdot 0.6} = 3.464$ **(1/2 P)**. Die Wahrscheinlichkeit, dass trotzdem die Nullhypothese beibehalten wird, berechnet sich aus der neuen Wahrscheinlichkeit des alten Verwerfungsbereichs und beträgt

$$\begin{aligned} P_{20,3.464}(X < 17.54) &= P_{0,1}\left(X < \frac{17.54 - 20}{3.464}\right) \\ = P_{0,1}(X < -0.71) &= 1 - \Phi(0.71) = 0.24 \ (\mathbf{1P}) \ . \end{aligned}$$

3. Die Wahrscheinlichkeit dass ein Wort richtig übermittelt wird sei $p = 3/4$. Die Wahrscheinlichkeit dass ein Wort nicht übermittelt wird ist $1 - p = 1/4$. Der Satz besteht aus 9 Wörtern.

a) (2 Punkte)
$\mathrm{P}[\{\text{Satz wird richtig übermittelt}\}] = p^9 = (3/4)^9 = 0.075$ (1 Punkt)

$\mathrm{P}[\{\text{Nur das vierte Wort im Satz wird nicht übermittelt}\}] = p^8(1-p) = (3/4)^8(1/4) = 0.025$ (1 Punkt)

b) (2 Punkte)

$$\begin{aligned}\mathrm{P}[\{\text{mindestens 2 Wörter richtig}\}] &\overbrace{=}^{1P} 1 - \mathrm{P}[\{\text{kein Wort richtig}\} \cup \{\text{1 Wort richtig}\}]\\ &\overbrace{=}^{1P} 1 - ((1/4)^9 + 9\frac{3}{4}(\frac{1}{4})^8) = 0.9998932\end{aligned}$$

c) (2 Punkte)
Markus spricht ein Wort mehreremale aus. Z sei die Anzahl solcher Ausprechungen desselben Wortes bis zur ersten Übermittelung.

$\mathrm{P}[Z = k] = (\frac{1}{4})^{(k-1)}\frac{3}{4}$ (1 Punkt)

$\mathrm{P}[Z > 3] \overbrace{=}^{0.5P} 1 - \mathrm{P}[Z \leq 3] \overbrace{=}^{0.5P} 1 - (\frac{3}{4} + \frac{1}{4}\frac{3}{4} + (\frac{1}{4})^2\frac{3}{4}) = 0.015625$ (1 Punkt)

d) (2 Punkte)

$$E[Z] \overbrace{=}^{1P.} \sum_{k=1}^{\infty} k\,\mathrm{P}[Z = k] \overbrace{=}^{0.5P} \sum_{k=1}^{\infty} k(1-p)^{k-1}p = p\sum_{k=1}^{\infty} k(1-p)^{k-1} \overbrace{=}^{0.5P} p\frac{1}{p^2} = \frac{1}{p} = 4/3$$

e) (2 Punkte)
Betrachte zuerst die Übermittelung eines Wortes durch Wiederholung:

$\mathrm{P}[\{\text{mind. eine der 4 Wortwiederholungen wird richtig übermittelt}\}]$
$\overbrace{=}^{0.5P} 1-\mathrm{P}[\{\text{alle 4 Wiederholungen werden nicht übermittelt}\}] \overbrace{=}^{0.5P} 1-(\frac{1}{4})^4 = 0.9960938$

Der ganze Satz besteht aus 9 Wörtern und jedes dieser Wörter spricht Markus 4 mal aus (0.5 P):

$\mathrm{P}[\{\text{Corinne versteht den Satz richtig}\}] \overbrace{=}^{0.5P} (1 - (\frac{1}{4})^4)^9 = 0.965388$

4. 1) b)

2) a)

3) a), der F-Test hat einen sehr kleinen p-Wert

4) b), der P-Wert für das Alter ist grösser als 5%

5) c), Trotz signifikantem F-Test ist es bei korrelierten erklärenden Variablen möglich, dass beide t-Tests nicht signifikant sind

6) d), $\hat{\beta}_1 \pm \widehat{s.e.}(\hat{\beta}_1) \cdot t_{30;0.975}$

7) c), keine Aussage möglich ohne Angabe des Altersunterschieds

8) c)

5. 1) e)
2) c)
3) b)
4) d)
5) c)
6) b)
7) d)
8) c)