Musterlösung

1. (11 Punkte)

a) Es handelt sich um einen gepaarten Test. Warum?
Jeder Temperatur aus dem Jahr 2012 kann eindeutig eine entsprechende Temperatur (diejenige aus demselben Monat) des Jahres 2013 zugeordnet werden. **(1 Punkt)**

b) Geben Sie die Null- und die Alternativhypothese an und begründen Sie kurz Ihre Wahl.

$$H_0 : \mu_D = \mu_0 = 2 \qquad \textbf{(0,5 Punkte)}$$
$$H_A : \mu_D > \mu_0 = 2 \qquad \textbf{(0,5 Punkte)}$$

wobei μ_D der Erwartungswert von $D_i = X_i - Y_i$ ist mit X_i: Temperatur im Monat i im Jahr 2012, Y_i: Temperatur im Monat i im Jahr 2013.
Man möchte im Sinne des Besitzers statistisch nachweisen, dass die Temperaturen im Jahr 2012 mehr als 2°C höher waren als im Jahr 2013, also dass die Differenz der Temperaturen im Jahr 2012 und 2013 grösser als 2 ist. Dabei handelt es sich um eine einseitige Alternativhypothese von der Form $H_A : \mu_D > 2$. **(1 Punkt)**

c) Führen Sie den geeigneten t-Test zum Niveau 0.05 durch: Bestimmen Sie den Wert der Teststatistik T, die Verteilung der Teststatistik T unter der Nullhypothese, den Verwerfungsbereich für T und den Testentscheid.

1. Modell: $D_i \sim \mathcal{N}(\mu_D, \sigma_D^2)$ iid
2. Nullhypothese: $H_0 : \mu_D = 2$
   Alternative: $H_A : \mu_D > 2$
3. Teststatistik:

$$T = \frac{\sqrt{n}(\bar{D} - \mu_0)}{\hat{\sigma}_D} \qquad \textbf{(0,5 Punkte)}$$
$$t = \frac{\sqrt{5} \cdot (2.5 - 2)}{0.64} \approx 1.75 \qquad \textbf{(0,5 Punkte)}$$

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

4. Signifikanzniveau: $\alpha = 0.05$
5. Verwerfungsbereich:

$$K = [t_{n-1;1-\alpha}; \infty) \qquad \textbf{(0,5 Punkte)}$$
$$= [t_{4;0.95}; \infty) = [2.132; \infty) \qquad \textbf{(0,5 Punkte)}$$

6. Testentscheid: $t \notin K \Rightarrow H_0$ wird nicht verworfen. **(0,5 Punkte, mit Folgefehler)**

d) Wie ist die Teststatistik des z-Tests unter der Nullhypothese verteilt? Geben Sie Verteilungsfamilie und Parameter an.
Die Teststatistik

$$Z = \frac{\sqrt{n}(\bar{D} - \mu_0)}{\sigma_D}$$

ist unter der Nullhypothese Standardnormalverteilt, d.h. $Z \sim \mathcal{N}(0, 1)$. **(1 Punkt)**
Alternativ: Die Teststatistik

$$Z = \bar{D}$$

ist unter der Nullhypothese $\mathcal{N}(\mu_0, \frac{\sigma_D^2}{n})$-verteilt, also $Z \sim \mathcal{N}(2, 0.082)$. **(1 Punkt)**

e) Bestimmen Sie ein einseitiges (passend zur Wahl der Alternativhypothese) 95%-Vertrauensintervall für μ_D. Verwirft der z-Test die Nullhypothese auf dem 5%-Niveau? Begründen Sie kurz. Da σ_D bekannt ist und die Alternativhypothese $H_A : \mu > 2$ nach oben zeigt hat das 95%-Vertrauensintervall die Form

$$VI = [\bar{D} - \Phi^{-1}(1-\alpha)\frac{\sigma_D}{\sqrt{n}}, \infty) \qquad \textbf{(0.5 Punkte)}$$
$$= [2.5 - 1.64 * \frac{0.64}{\sqrt{5}}, \infty) = [2.03, \infty) \qquad \textbf{(0.5 Punkte)}$$

Da $\mu_0 = 2$ nicht im 95%-Vertrauensintervall enthalten ist, verwirft der z-Test die Nullhypothese auf dem Signifikanzniveau von 5%. **(1 Punkt)**

f) Wie gross muss c mindestens sein, damit sich der Testentscheid des t-Tests ändert? (*Hinweis: Überlegen Sie sich, wie sich $\bar{D}$ und $\hat{\sigma}_D$ ändern, wenn Sie bei jeder monatlichen Differenz den Wert c addieren. Benutzen Sie danach die Teststatistik des t-Tests.*)
Zu jeder monatlichen Differenz wird der Wert c addiert. Deshalb erhöht sich der Mittelwert ebenfalls um c, d.h.: $\bar{D} = 2.5 + c$. **(0.5 Punkte)**
Die Standardabweichung $\hat{\sigma}_D$ ändert sich nicht, also $\hat{\sigma}_D = 0.64$. **(0.5 Punkte)**
Die Teststatistik ist also gegeben durch:
$t = \frac{\sqrt{5}(2.5+c-2)}{0.64} = \frac{\sqrt{5}\cdot 0.5}{0.64} + \frac{\sqrt{5}\cdot c}{0.64} \approx 1.75 + \frac{\sqrt{5}\cdot c}{0.64}$. **(0.5 Punkte)**
Damit sich der Testentscheid ändert, muss t genau dem $t_{n-1,1-\alpha}$-Quantil ($= 2.132$) entsprechen, also $1.75 + \frac{\sqrt{5}\cdot c}{0.64} = 2.132$.
Auflösen nach c ergibt: $c \approx 0.11$. **(0.5 Punkte, aufpassen auf Folgefehler, falls Quantil in c) falsch bestimmt)**

2. (9 Punkte)

a) Wie ist X verteilt? Geben Sie die Verteilungsfamilie und die Parameter an.
X ist binomialverteilt **(0.5 Punkte)** mit $n = 8$ und $\pi = 0.5$. **(0.5 Punkte)**

b) Wie gross ist die Wahrscheinlichkeit, dass mindestens 7 Umfragen beantwortet werden?
$P(X \geq 7) = P(X = 7) + P(X = 8) = \binom{8}{7}0.5^7(1-0.5) + \binom{8}{8}0.5^8(1-0.5)^0 = 0.035$ **(1 Punkt)**

c) Wie gross ist die Wahrscheinlichkeit, dass genau ein Viertel der angeschriebenen Studenten die Umfrage beantwortet? Wie gross ist die Wahrscheinlichkeit, dass genau ein Viertel der Studenten die Umfrage *nicht* beantwortet?
Die Wahrscheinlichkeit, dass genau ein Viertel (also 2 der 8) Studenten antworten, ist gegeben durch:
$P(X = 2) = \binom{8}{2}0.5^2(1-0.5)^6 = 0.109$ **(0.5 Punkt)**
Die Wahrscheinlichkeit, dass genau ein Viertel nicht antwortet ist:
$P(X = 6) = \binom{8}{6}0.5^6(1-0.5)^2 = 0.109$ **(0.5 Punkt)**
Da die Wahrscheinlichkeit, dass ein Student antwortet ($\pi = 0.5$) gleich ist wie die Wahrscheinlichkeit, dass ein Student nicht antwortet ($1 - \pi = 0.5$), sind die beiden gesuchten Wahrscheinlichkeiten identisch.

d) Was sind die Null- und die Alternativhypothese?
$H_0 : \pi = 0.5$ **(0.5 Punkte)** und $H_A : \pi < 0.5$. **(0.5 Punkte)**

e) Führen Sie einen geeigneten exakten Test durch. Geben Sie den Verwerfungsbereich für X und den Testentscheid an.
(2 Punkte)
1. Modell: $X \sim \text{Bin}(8, \pi)$ ist die Anzahl der Studenten, die die Umfrage beantworten
2. Nullhypothese: $H_0 : \pi = 0.5$ Alternativhypothese: $H_A : \pi < 0.5$
3. Die Teststatistik ist: $T = X$, Verteilung von T unter H_0: $T \sim \text{Bin}(8, 0.5)$
4. Signifikanznivau: $\alpha = 0.05$
5. Verwerfungsbereich:
Da die Alternativhypothese nach unten zeigt hat der Verwerfungsbereich die Form $K = [0, c]$. Gesucht ist das grösste c, sodass $P[X \leq c] \leq 0.05$. Da
$P(X = 0) = \binom{8}{0}0.5^8 = 0.0039$,
$P(X = 1) = \binom{8}{1}0.5^8 = 0.0313$,
$P(X = 2) = \binom{8}{2}0.5^8 = 0.1094$,
gilt

	$x = 0$	$x = 1$	$x = 2$
$P[X = x]$	0.0039	0.0313	0.1094
$P[X \leq x]$	0.0039	0.0352	0.1445

Der Verwerfungsbereich ist also gegeben durch $K = \{0, 1\}$. **(1 Punkt)**
6. Testentscheid: Da $2 \notin \{0, 1\}$ wird die Nullhypothese nicht verworfen **(1 Punkt, mit Folgefehler)**.

f) Berechnen Sie die Wahrscheinlichkeit für einen Fehler 2. Art, falls der Parameter der Binomialverteilung in Tat und Wahrheit $\pi = 0.25 \in H_A$ ist. Wie müsste das Signifikanzniveau α verändert werden, damit der Fehler 2. Art kleiner wird? Ein Fehler 2. Art liegt dann vor, wenn der Test die Nullhypothese beibehält, obwohl die Alternative ($\pi = 0.25$) wahr wäre.

$$\begin{aligned} P_{\pi=0.25}[X \notin K] &= P_{\pi=0.25}[X \geq 2] \qquad \textbf{(0.5 Punkte)} \\ &= 1 - P_{\pi=0.25}[X \leq 1] = 1 - \left(\binom{8}{0} 0.25^0 0.75^8 + \binom{8}{1} 0.25^1 0.75^7 \right) \\ &= 1 - 0.367 = 0.633. \qquad \textbf{(0.5 Punkte)} \end{aligned}$$

Damit der Fehler 2. Art kleiner wird, müsste das Signifikanzniveau α vergrössert werden. **(1 Punkt)**

g) Geben Sie ein approximatives zweiseitiges 95%-Vertrauensintervall für den unbekannten Parameter π an. Da n gross ist, kann man die Normalapproximation benutzen. Das 95%-Vertrauensintervall ist gegeben durch:
$I \approx \frac{x}{n} \pm 1.96\sqrt{\frac{x}{n}\left(1 - \frac{x}{n}\right)\frac{1}{n}} = \frac{211}{500} \pm 1.96\sqrt{\frac{211}{500}\frac{289}{500}\frac{1}{500}} = [0.379, 0.465]$. **(1 Punkt)**.

3. (9 Punkte)

a) Schätzen Sie $P[X = \text{Menu 1}]$, $P[X = \text{Vegi Menu}]$, $P[X = \text{Bio Menu}]$ und $P[X = \text{Menu spezial}]$ aus den Angaben in der Tabelle.
(1 Punkt)
Insgesamt wurden $420 + 160 + 40 + 380 = 1000$ Kunden beobachtet. Daher gilt:
$P[X = \text{Menu 1}] = \frac{420}{1000} = 0.42$. Analog findet man:
$P[X = \text{Vegi Menu}] = 0.16$, $P[X = \text{Bio Menu}] = 0.04$ und $P[X = \text{Menu spezial}] = 0.38$.

b) Wieviele CHF Einnahmen erwartet die Mensa pro Kunde? Mit welcher Varianz? (*Falls Sie Aufgabe a) nicht lösen konnten, verwenden Sie:* $P[X = \text{Menu 1}] = 0.35$, $P[X = \text{Vegi Menu}] = 0.23$, $P[X = \text{Bio Menu}] = 0.12$, und $P[X = \text{Menu spezial}] = 0.30$.)
Die erwarteten Einnahmen pro Kunde in CHF sind gegeben durch:
$\sum_{i=1}^{4} \text{Preis}(\text{Menu}_i) \cdot P[X = \text{Menu}_i] = 6.20 \cdot 0.42 + 6.20 \cdot 0.16 + 7.00 \cdot 0.04 + 7.90 \cdot 0.38 \approx 6.88.$
(1 Punkt) mit der Varianz
$(6.20 - 6.88)^2 \cdot 0.42 + (6.20 - 6.88)^2 \cdot 0.16 + (7.00 - 6.88)^2 \cdot 0.04 + (7.90 - 6.88)^2 \cdot 0.38 \approx 0.67.$
(1 Punkt)
Mit den alternativen Wahrscheinlichkeiten (falls man a) nicht gelöst hat) erhält man analog:
Erwarteter Gewinn in CHF: 6.81 , Varianz in CHF: 0.58 .

c) Begründen Sie kurz, weshalb die Poissonverteilung ein geeignetes Modell für Y ist. Wie lautet eine geeignete Schätzung für λ?
Da wir eine Anzahl modellieren möchten, kommen nur diskrete Verteilungen in Frage. Da der Wertebereich nicht beschränkt ist, bietet sich die Poissonverteilung an. **(1 Punkt, sinngemäss)**
Eine geeignete Schätzung für λ ist gerade die Beobachtung, also $\hat{\lambda} = 420$. **(1 Punkt)**

d) Wie lautet die *exakte* Verteilung für die Anzahl Kunden pro Arbeitswoche (5 Tage), die sich für das Menu 1 entscheiden? Geben Sie die Verteilungsfamilie und Parameter an. Sie dürfen annehmen, dass die Anzahl Kunden an verschiedenen Arbeitstagen unabhängig voneinander sind.
Es handelt sich hierbei um eine Summe von unabhängigen, Poisson-verteilten Zufallsvariablen. Sei Y_i: Anzahl Kunden am i-ten Arbeitstag. Es gilt: $Y_i \sim \text{Poisson}(420)$.
Die Anzahl Kunden, die sich pro Arbeitswoche (5 Tage) für das "Menu 1" entscheidet ist gegeben durch: $S = Y_1 + Y_2 + ... + Y_5 \sim \text{Poisson}(\lambda_S)$ **(0.5 Punkte)** mit $\lambda_S = 420 + 420 + ... + 420 = 2100$ **(0.5 Punkte)**.

e) Sei nun die Zufallsvariable D die Gesamtanzahl Kunden der Mensa pro Tag. Nehmen Sie an, $D \sim \text{Poisson}(1000)$. Berechnen Sie die Wahrscheinlichkeit, dass an einem Tag zwischen 990 und 1020 Kunden in der Mensa essen. Benutzen Sie eine geeignete Normalapproximation.

Da $\mathbf{E}[D] = \text{Var}(D) = 1000$ ist die Gesamtanzahl der Kunden der Mensa pro Tag approximativ $\mathcal{N}(1000, 1000)$-verteilt **(1 Punkt)**. Die gesuchte Wahrscheinlichkeit $P[990 \leq D \leq 1020]$ kann durch Standardisieren berechnet werden. Es gilt:
$P[990 \leq D \leq 1020] = P[\frac{990-1000}{\sqrt{1000}} \leq \frac{D-\mathbf{E}[D]}{\sqrt{\text{Var}(D)}} \leq \frac{1020-1000}{\sqrt{1000}}] = P[-0.32 \leq Z \leq 0.63]$
wobei $Z := \frac{D-\mathbf{E}[D]}{\sqrt{\text{Var}(D)}} \sim \mathcal{N}(0,1)$. **(1 Punkt)**
Mit der Tabelle der kumulativen Standardnormalverteilung erhält man:

$$
\begin{aligned}
P[-0.32 \leq Z \leq 0.63] &= \phi(0.63) - \phi(-0.32) = \phi(0.63) - (1 - \phi(0.32)) \\
&= 0.7357 - (1 - 0.6255) = 0.3612 \approx 0.36. \qquad \textbf{(1 Punkt)}
\end{aligned}
$$

4. (7 Punkte) Oskar und Max werfen nacheinander einen Ball von unterschiedlich hohen Positionen (hoehe) auf den Boden und messen die Zeit (zeit) bis der Ball aufschlägt. Folgendes Modell wurde an die Daten angepasst:

$$
\text{hoehe}_i = \beta_0 + \beta_1 \cdot \text{zeit}_i + \beta_2 \cdot \text{zeit}_i^2 + \varepsilon_i, \qquad \varepsilon_i \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2).
$$

Der (unvollständige) Regressionsoutput sieht wie folgt aus:

```
Residuals:
   Min     1Q Median     3Q    Max
-2.359 -0.692 -0.185  0.926  2.522

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)     1.987      2.572     ???      ???
zeit              ???      2.326   -0.82     0.42
zeit.quadrat    5.340      0.492   10.86  2.9e-12
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.31 on 32 degrees of freedom
Multiple R-squared:  0.994,       Adjusted R-squared:  0.994
F-statistic: 2.81e+03 on 2 and 32 DF,  p-value: <2e-16
```

1) Wie viele Messungen haben Oskar und Max gemacht?
   a) 30
   b) 32
   c) 34
   d) 35

   d (35)

2) Wird die Nullhypothese $H_0 : \beta_0 = 0$ auf dem 5%-Niveau verworfen?
   a) Ja.
   b) Nein.
   c) Keine Angabe möglich.

   b (Der Verwerfungsbereich ist $K = (-\infty, -t_{32,0.975}] \cup [t_{32,0.975}, \infty) = (-\infty, -2.037] \cup [2.037, \infty)$ und $t = \frac{1.987}{2.572} = 0.773$ ist nicht in K.)

3) Wie gross ist die Schätzung von $\hat{\beta}_1$?
   a) -2.84
   b) -1.91

c) -0.35
d) -0.34

b (se(b)· t-value(b) $= 2.326 \cdot (-0.82) = -1.91$)

4) Berechnen Sie das exakte zweiseitige 99%-Konfidenzintervall für β_2.

a) $[3.99, 6.69]$
b) $[4.34, 6.34]$
c) $[4.36, 6.32]$
d) $[4.59, 5.99]$

a ($[\hat{\beta}_2 - t_{32,0.995} \cdot se(\hat{\beta}_2), \hat{\beta}_2 + t_{32,0.995} \cdot se(\hat{\beta}_2)] = [5.340 - 2.738 \cdot 0.492, 5.340 + 2.738 \cdot 0.492] = [3.99, 6.69]$)

5) Oskar und Max werfen den Ball aus 20 Metern Höhe herunter und messen dabei 2 Sekunden bis zum Aufprall. Wie gross ist der Fehler gemäss dem Modell ($\hat{=}$ Residuum)?

   a) 1.4
   b) 0.5
   c) -1.4
   d) -0.5

   b (Erwartet: $1.987 - 1.91 \cdot 2 + 5.340 \cdot 4 = 19.5$, Residuum: $20 - 19.5 = 0.5$)

6) Handelt es sich bei dem obigen Modell um ein lineares Modell?

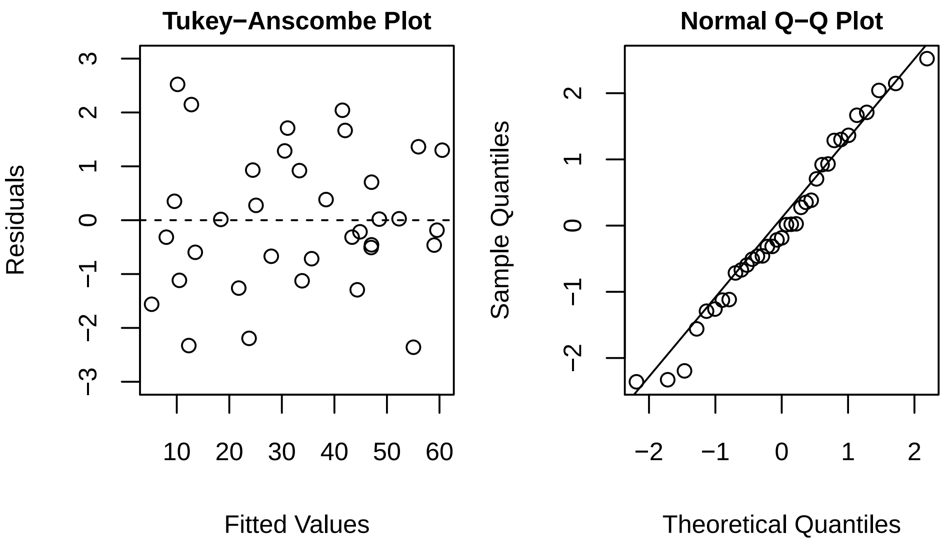
   a) Nein, weil das Modell einen quadratischen Regressionsterm (`zeit.quadrat`) enthält.
   b) Nein, weil die Variable `zeit` zweifach als erklärende Variable im Modell vorkommt.
   c) Ja, weil das Modell linear in den Koeffizienten ist.
   d) Ja, weil die Zielvariable (`hoehe`) nicht transformiert ist.

   c (Ja, weil das Modell linear in den Koeffizienten ist.)

7) Betrachten Sie die nachfolgenden Plots. Welche der folgenden Aussagen trifft zu?

   a) Alle Modellannahmen sind erfüllt.
   b) Die Fehlervarianz ist nicht konstant, aber die Normalverteilungsannahme ist plausibel.
   c) Die Fehlervarianz ist konstant, aber die Normalverteilungsannahme trifft nicht zu.
   d) Sowohl konstante Fehlervarianz als auch Normalverteilungsannahme treffen nicht zu.

   a (Alle Modellannahmen sind erfüllt.)



5. (9 Punkte) Die folgenden Aufgaben sind zufällig angeordnet und insbesondere nicht nach Schwierigkeitsgrad sortiert.

1) Seien $X \sim \mathcal{N}(2, 16)$ und $Y \sim$ Poisson(3) zwei unabhängige Zufallsvariablen. Berechnen Sie $\text{Var}(X + 3Y - 5)$.

   a) 20
   b) 25
   c) 30
   d) 38
   e) 43

   e), $\text{Var}(X) + 9 \cdot \text{Var}(Y) = 16 + 27 = 43$.

2) Sei $X \sim \text{Uniform}([-2; 2])$, dann sind X und $3 - X$

   a) Unabhängig und unkorreliert.
   b) Abhängig aber unkorreliert.
   c) Unabhängig aber korreliert.

d) Abhängig und korreliert.

d), X und $3 - X$ sind korreliert mit Korrelation -1. Korrelation impliziert Abhängigkeit. Alternativ berechnet man die Korrelation formal wie folgt:
$Cov(X, 3 - X) = \mathbf{E}[X(3 - X)] - \mathbf{E}[X]\mathbf{E}[3 - X] = 3\mathbf{E}[X] - \mathbf{E}\left[X^2\right] - 3\mathbf{E}[X] + \mathbf{E}[X]^2 = -\operatorname{Var}(X) < 0$.
$\rho(X, 3 - X) = \frac{Cov(X,3-X)}{\sqrt{\operatorname{Var}(X)}\sqrt{\operatorname{Var}(3-X)}} = \frac{-\operatorname{Var}(X)}{\operatorname{Var}(X)} = -1$.

3) Sei $X \sim \text{Exp}(1/3)$ eine Zufallsvariable. Berechnen Sie $P[X \geq 5]$.

a) 0
b) 0.19
c) 0.76
d) 0.81
e) 0.94

b), $1 - (1 - e^{-5/3}) = 0.19$.

4) Bei welcher der folgenden Funktionen handelt es sich um eine Dichtefunktion?

a)
$$f(x) = 0$$

b)
$$f(x) = \begin{cases} 0 & \text{für} \quad x \leq 0 \\ 2x & \text{für} \quad x \in (0;1) \\ 1 & \text{für} \quad x \geq 1 \end{cases}$$

c)
$$f(x) = \begin{cases} -4x^3 & \text{für} \quad x \in [-1;0] \\ 0 & \text{sonst} \end{cases}$$

d)
$$f(x) = \begin{cases} -1 & \text{für} \quad x \in [0;1] \\ 1 & \text{für} \quad x \in (1;3] \\ 0 & \text{sonst} \end{cases}$$

c), es muss gelten $f(x) \geq 0$ und $\int_{\mathbb{R}} f(x)dx = 1$.

5) Aus einer Stichprobe von 2 Werten berechnet man den Mittelwert ($\bar{x} = 5.5$) und die empirische Varianz ($\hat{\sigma}^2 = 4.5$). Der kleinere der beiden Werte ist dann:

a) -1
b) 1
c) 3.25
d) 4
e) Es gibt nicht genügend Informationen, um die Frage zu beantworten.

d), Die 2 Werte haben den gleichen Abstand d zum Mittelwert. $\hat{\sigma}^2 = 2d^2$ und folglich $d = 1.5$. $x_{1,2} = \bar{x} \pm 1.5$

6) Ueli untersucht, ob das Geschlecht einer Person einen Einfluss auf die Orientierungsfähigkeit hat. Er testet 50 Frauen und 50 Männer und stellt dann die folgenden Hypothesen auf:
H_0: Es gibt keinen Unterschied in der Orientierungsfähigkeit von Mann und Frau.
H_A: Es gibt einen Unterschied in der Orientierungsfähigkeit von Mann und Frau.
Diese Hypothesen möchte Ueli nun anhand eines ungepaarten, 2-seitigen t-Tests auf dem 5%-Niveau analysieren. Der p-Wert des Testergebnisses beträgt 0.99. Ueli schliesst daraus, dass es mit hoher Wahrscheinlichkeit keinen Unterschied in der Orientierungsfähigkeit von Mann und Frau gibt. Welchen Fehler hat Ueli gemacht?

a) Der Test hätte gepaart durchgeführt werden sollen.
b) Die Alternativhypothese hätte einseitig gewählt werden müssen.
c) Das Testergebnis (p-Wert$= 99\%$) ist signifikant auf dem Niveau $\alpha = 5\%$, folglich hat das Geschlecht einen Einfluss.
d) Wenn die Nullhypothese akzeptiert wird, kann Ueli nicht mit hoher Wahrscheinlichkeit schliessen, dass sie richtig ist.

d)

7) Sei $X \sim \mathcal{N}(3, 8)$ eine Zufallsvariable. Berechnen Sie $P[X = 5]$.

a) 0
b) 0.24
c) 0.76
d) 0.88
e) 0.999

a), für stetige Verteilungen gilt $P(X = x) = 0$ für jedes x.

8) Für zwei Ereignisse A und B hat man $P(A) = 0.8$ und $P(B) = 0.9$. Welche Werte kann $P(A \cap B)$ ausschliesslich annehmen?

a) $0 \leq P(A \cap B) \leq 1$
b) $0.5 \leq P(A \cap B) \leq 0.9$
c) $0.7 \leq P(A \cap B) \leq 0.9$
d) $0.7 \leq P(A \cap B) \leq 0.8$
e) $0.8 \leq P(A \cap B) \leq 0.9$

d), die obere Schranke findet man mithilfe $P(A \cap B) \leq P(A) = 0.8$, es kommt also nur d) in Frage. (Untere Schranke: $P(A \cap B) = 1 - P((A \cap B)^c) = 1 - P(A^c \cup B^c) \geq 1 - (P(A^c) + P(B^c)) = 1 - (1 - P(A) + 1 - P(B)) = P(A) + P(B) - 1 = 0.7$)

9) Eine Stichprobe besteht aus n Beobachtungen. Der Mittelwert der Stichprobe beträgt 7. Die Beobachtung $x_1 = 31$ wird jetzt gestrichen. Danach ist der Mittelwert der restlichen Beobachtungen $x_2, ..., x_n$ gleich 5. Wie viele Beobachtungen beinhaltete die ursprüngliche Stichprobe?

a) $n = 5$
b) $n = 12$
c) $n = 13$
d) $n = 24$
e) $n = 57$

c), da $\frac{1}{n} \sum_{i=1}^{n} x_i = 7$ und $\frac{1}{n-1} \sum_{i=2}^{n} x_i = 5$ gilt $\sum_{i=1}^{n} x_i = 7n$ und $\sum_{i=2}^{n} x_i = 5(n-1)$. Folglich $x_1 = \sum_1^n x_i - \sum_2^n x_i = 7n - 5(n-1)$. Einsetzen von $x_1 = 31$ ergibt $n = 13$.