Stochastik - Musterlösung
(BSc D-MAVT / BSc D-MATH / BSc D-MATL)

1. **a)** 2. **b)** 1. **c)** 2. **d)** 2. **e)** 1. **f)** 3. **g)** 1. **h)** 2. **i)** 2. **j)** 3.

In detail:

a) $X(1-X)$ ist immer 0 (man könnte aber denken, dass es $E[X] - \text{Var}(X) = p^2$ ist). $X^2 = X$, daher stimmt 2. $E[X^3] \neq E[X]^3$.

b) Aus der Formel für die Momente einer Exponentialverteilung erhält man $\lambda = \frac{1}{2}$, damit ist $\text{Var}(2X-3) = 2^2\text{Var}(X) = 2^2 \cdot 2^2 = 16$. Die anderen Antworten erhält man, wenn man nicht quadriert oder denkt, dass die Varianz additiv sei.

c) Zuerst berechnen wir $F(x) = \frac{c}{3}x^3$, $x \in [0,3]$. Aus $F(3) = 1$ folgt $c = \frac{1}{9}$ (Man muss also realisieren, dass c eindeutig bestimmt ist). Somit ist $F(1) = \frac{1}{27}$ und 1. falsch. Das 8/27-Quantil q erhalten wir als $8/27 = \frac{1}{27}q^3$, also $q = 2$.

d) Aus

$$\frac{2}{3} = P(A \cup B) = P(A) + \frac{1}{3} - P(A \cap B)$$

und $P(A \cap B) \geq 0$ erhalten wir $P(A) \geq \frac{1}{3}$ womit 1. falsch ist. Für $B \subset A$ erhält man $P(A) = P(A \cup B) = \frac{2}{3}$ womit 3. falsch und 2. richtig ist.

e) Summieren gibt $P(X=0) = 0.5, P(X=1) = 0.3, P(X=2) = 0.2$ sowie $P(Y=0) = 0.2$,$P(Y=1) = 0.2$ und $P(Y=2) = 0.6$. Nachprüfen zeigt nun, dass tatsächlich $P(X=x, Y=y) = P(X=x)P(Y=y)$.

f) Für $X \sim Bin(6, \frac{1}{2})$ haben wir $P(X=6) = 0.016$, $P(X=5) = 0.094$. Ein Test zum Niveau 5% wird also tatsächlich nur bei 6 verwerfen. 1. ist falsch, die Wahrscheinlichkeit für einen Fehler 1. Art ist kleiner als 5% (genauer: 1.6%). Andererseits ist die Wahrscheinlichkeit für einen Fehler 2. Art bei $p = 0.8$ gegeben durch $P_{0.8}(X < 6) = 1 - P_{0.8}(X=6) = 1 - 0.8^6 = 0.738$ bzw. die Macht durch 0.262.

g) Die Länge des $1-\alpha$-Vertrauensintervalls ist $2\frac{\sigma}{\sqrt{n}}z_{1-\frac{\alpha}{2}}$. Also ist das Verhältnis gegeben als $z_{0.975}/z_{0.995}$ was wir aus der Tabelle als ungefähr $1.96/2.58 \approx 0.76$ (oder $1.96/2.57 \approx 0.76$) identifizieren, also ist 1. richtig. Die Zahlen in 2. erhält man aus $z_{0.95}/z_{0.99}$.

h) Es gilt $E[\frac{1}{n}\sum X_i] = \lambda$, darum ist 2. richtig. 3. ist falsch, da der zentrale Grenzwertsatz nur *genähert* gilt.

Bitte wenden!

i) 1. ist falsch, da der p-Wert grösser als das Niveau ist. Also wird die Nullhypothese nicht verworfen und der dazugehörige Parameter, 0 ist im Vertrauensintervall enthalten, also stimmt 2.

j) III. ist rechtsschief, genau wie a (e.g.:Median betrachten). II. und I. sind beide symmetrisch, aber II. hat eine grössere Streuung und gehört daher zu c (e.g.: oberes Quartil betrachten).

2. Sei G_A das Ereignis, dass Alice die Herstellung gelingt und L_A das Ereignis, dass die Vorbereitung von einem Laboranten übernommen wird (und analog für Bob G_B und L_B). Dann ist $P(G_A|L_A) = 0.8$, $P(G_B|L_B) = 0.7$, $P(G_A^c|L_A^c) = 0.7$, $P(G_B^c|L_B^c) = 0.6$, $P(L_A) = 0.75$ und $P(L_B) = 0.9$ aus der Aufgabenstellung.

a) Mit dem Satz der totalen W'keit und der Rechenregel für Komplenentärereignisse erhalten wir

$$\begin{aligned} P(G_A) &= P(G_A|L_A)P(L_A) + P(G_A|L_A^c)P(L_A^c) \\ &= 0.8 \cdot 0.75 + (1 - 0.7) \cdot (1 - 0.75) = 0.675. \end{aligned}$$

b) Aus der Aufgabenstellung wissen wir, dass G_A und G_B unabhängig sind. Also gilt

$$\begin{aligned} P(\text{'die vier müssen auf das Bier verzichten'}) &= 1 - P(G_A \cap G_B) = 1 - P(G_A)P(G_B) \\ &= 1 - 0.675P(G_B) = 1 - 0.675 \cdot 0.67 \\ &= 0.54775 \end{aligned}$$

wobei wir $P(G_B) = 0.67$ wie in a) berechnet haben.

c) Mit dem Satz von Bayes' und dem Ergebnis aus a) berechnen wir die gesuchte Wahrscheinlichkeit als

$$P(L_A|G_A) = \frac{P(G_A|L_A)P(L_A)}{P(G_A)} = \frac{0.8 \cdot 0.75}{0.675} = \frac{8}{9}.$$

d) Wir möchten $P(G_A \cap G_B)$ maximieren. Wegen der Unabhängigkeit gilt $P(G_A \cap G_B) = P(G_A)P(G_B)$. Wie in a) berechnen wir also

$$P(G_A) = P(G_A|L_A)P(L_A) + P(G_A|L_A^c)P(L_A^c) = 0.8p + 0.3(1-p)$$

und analog $P(G_B)$ und erhalten

$$\begin{aligned} P(G_A \cap G_B) = P(G_A)P(G_B) &= [0.8p + 0.3(1-p)][0.7(1-p) + 0.4p] \\ &= (0.5p + 0.3)(0.7 - 0.3p) = 0.26p - 0.15p^2 + 0.21. \end{aligned}$$

Diese quadratische Funktion in p ist maximal für $p = \frac{26}{30}$.

3. a) Wir berechnen $P(|U| > 1) = P(\{U < -1\} \cup \{U > 1\}) = \frac{-1-(-3)}{6} + 1 - \frac{1-(-3)}{6} = 2/3$.

Siehe nächstes Blatt!

b) (i) Pro Tag müssen im Schnitt $1515/30 = 50.5$ Tonnen Eis produziert werden, also 0.5 mehr als der Sollwert.

(ii) Es sei X_i die Abweichung vom angepeilten Wert am i-ten Tag. Nach Annahme sind die X_i i.i.d. und jedes X_i hat dieselbe Verteilung wie U. Mit $S_{30} = X_1 + \cdots + X_{30}$ ist dann nach dem Zentralen Grenzwertsatz die durchschnittliche tälgiche Abweichung $S_{30}/30$ ungefähr $\mathcal{N}(0, 3/30)$-verteilt, weil $\mathrm{Var}(U) = 3$.

(iii) Aus (i) folgt, dass die verlangte Menge an Eis geliefert werden kann, wenn $S_{30}/30 \geq 0.5$. Die Wahrscheinlichkeit dafür ist also mit der Approximation aus (ii) gegeben durch $P(S_{30}/30 \geq 0.5) \approx 1 - \Phi(0.5/\sqrt{1/10}) = 1 - 0.943 = 0.057$

c) $E[U] = 0$, also verwenden wir (siehe Hinweis) z.B. die Identität $E[U^2] = \frac{1}{12} \cdot (2\vartheta)^2 = \frac{1}{3}\vartheta^2$ und ersetzen $E[U^2]$ durch $\frac{1}{5}\sum_{i=1}^5 X_i^2$. Auflösen nach ϑ gibt

$$\hat{\vartheta} = \left(\frac{3}{5}\sum_{i=1}^{5} X_i^2\right)^{\frac{1}{2}}$$

als Momentenschätzer. Auswerten an der gegebenen Stichprobe gibt ungefähr 2.8.

Alternativ könnte man auch die Varianz oder ein höheres Moment benutzen, dies gibt auch volle Punkte.

d) Jedes X_i hat Dichte $\frac{1}{2\vartheta}$ auf $[-\vartheta, \vartheta]$ und 0 sonst. Die Likelihoodfunktion zur Stichprobe $x_1, \cdots, x_5$ ist also gegeben als

$$L(\vartheta) = \begin{cases} \left(\frac{1}{2\vartheta}\right)^5, & x_1, \cdots, x_5 \in [-\vartheta, \vartheta] \\ 0, & \text{sonst.} \end{cases}$$

Maximieren (nicht ableiten, nur überlegen) gibt $\hat{\vartheta} = \max_i |X_i|$, d.h. hier 2.3.

4. **a)** Die Stichprobe ist gepaart.

b) Die für den Test relevanten Daten sind die Stichprobe $(d_1, \ldots, d_{10})$. Plot 1 sind falsche Datenpunkte. In Plot 2 sind die Datenpunkte nicht richtig eingezeichnet (nicht geordnet). Plot 3 ist richtig. In Plot 4 sind auf der x-Achse nicht die Normalquantile aufgetragen (z.B. sieht man, dass die aufgetragenen x-Werte nicht symmetrisch sind um 0).

c) Aufgrund der Werte in den 'tails' scheint keine Normalverteilung vorzuliegen. Die Verteilung ist aber symmetrisch und wir entscheiden uns daher für den Wilcoxon-Test (Vorzeichentest würde die Symmetrie nicht ausnutzen).

d)
- Wir wollen herausfinden ob der Mittelwert μ der Differenzen ($d_i = a_i - b_i$ sind Realisierungen davon, analog zum Skript 8.2) grösser ist als 0, das heißt wir werden einen einseitigen Test zum Niveau $\alpha = 5\%$ mit den Hypothesen
  $H_0 : \mu = 0$
  $H_A : \mu > 0$
  durchführen.

Bitte wenden!

- Wir berechnen die Differenzen, Ränge und Vorzeichen als

Zimmer Nr (i)	1	2	3	4	5	6	7	8	9	10
(d_i)	-3	4	-1	2	18	8	4	31	6	-25
Rang($\|d_i\|$)	3	4.5	1	2	8	7	4.5	10	6	9
V_i	0	1	0	1	1	1	1	1	1	0

  Damit ist die Teststatistik $W = 42$.
- Aus der Tabelle sehen wir, dass der Verwerfungsbereich durch $[45, 55]$ gegeben ist, d.h., dass
- die Nullhypothese hier nicht verworfen wird. Es kann also nicht nachgewiesen werden, dass Ecovacs schneller ist.