

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

Schreiben Sie für Aufgabe 2-4 stets alle Zwischenschritte und -rechnungen sowie Begründungen auf. Aufgabe 1 ist eine Multiple Choice Aufgabe (keine Begründungen notwendig).

Die für die Aufgaben benötigten Tabellen (Normalverteilung, Quantile der t-Verteilung, Verwerfungsbereiche für den Wilcoxon-Test) wurden mit der Prüfung ausgeteilt.

1. (10 Punkte)

Bei den folgenden 10 Fragen ist jeweils genau eine Antwort richtig. Es gibt pro richtig beantwortete Frage 1 Punkt und pro falsche Antwort 1/2 Punkt Abzug. Minimal erhält man für die gesamte Aufgabe 0 Punkte. **Bitte benützen Sie das beiliegende Antwortblatt.**

a) Sei $X \sim \text{Bernoulli}(p)$ mit $p \in (0, 1)$. Dann gilt

1. $E[X(1 - X)] = p^2$
2. $X^2 \sim \text{Bernoulli}(p)$
3. $E[X^3] = p^3$.

b) Sei $X \sim \text{Exp}(\lambda)$ mit $E[X] = 2$. Dann ist $\text{Var}(2X - 3)$ gleich

1. 16
2. 8
3. 13

Bitte wenden!

c) Sei f eine Dichtefunktion mit

$$f(x) = \begin{cases} cx^2, & 0 \leq x \leq 3 \\ 0, & \text{sonst.} \end{cases}$$

für ein $c > 0$. Dann gilt

1. $\mathbb{P}(X \leq 1) = \frac{1}{9}$.
2. Das 8/27-Quantil liegt in $[1.8, 2.2]$.
3. Weder 1. noch 2. ist richtig.

d) Es seien A und B Ereignisse mit $P(A \cup B) = \frac{2}{3}$ und $P(B) = \frac{1}{3}$. Dann

1. kann $P(A) = \frac{1}{4}$ gelten.
2. kann $P(A) = \frac{2}{3}$ gelten.
3. muss $P(A) = \frac{1}{3}$ gelten.

e) Seien X und Y zwei Zufallsvariablen welche nur die Werte 0, 1 und 2 annehmen können. Die Wahrscheinlichkeitsfunktion von (X, Y) ist veranschaulicht in folgender Tabelle (z.B. $P(X = 0, Y = 0) = p(0, 0) = 0.10$):

$X \setminus Y$	0	1	2
0	0.1	0.1	0.3
1	0.06	0.06	0.18
2	0.04	0.04	0.12

Sind X und Y unabhängig?

1. Ja.
2. Nein.
3. Ist aus der Tabelle nicht ersichtlich.

f) Eine Münze hat bei 6 Würfeln 6 mal Kopf ergeben. Wir möchten testen ob die Münze fair ist oder nicht und führen einen geeigneten einseitigen Binomialtest zum Niveau 5% mit Nullhypothese $H_0 : p = 0.5$ durch, wobei p die Wahrscheinlichkeit für 'Kopf' bezeichnet. Dies ergibt einen Verwerfungsbereich $K = \{6\}$. Dann gilt:

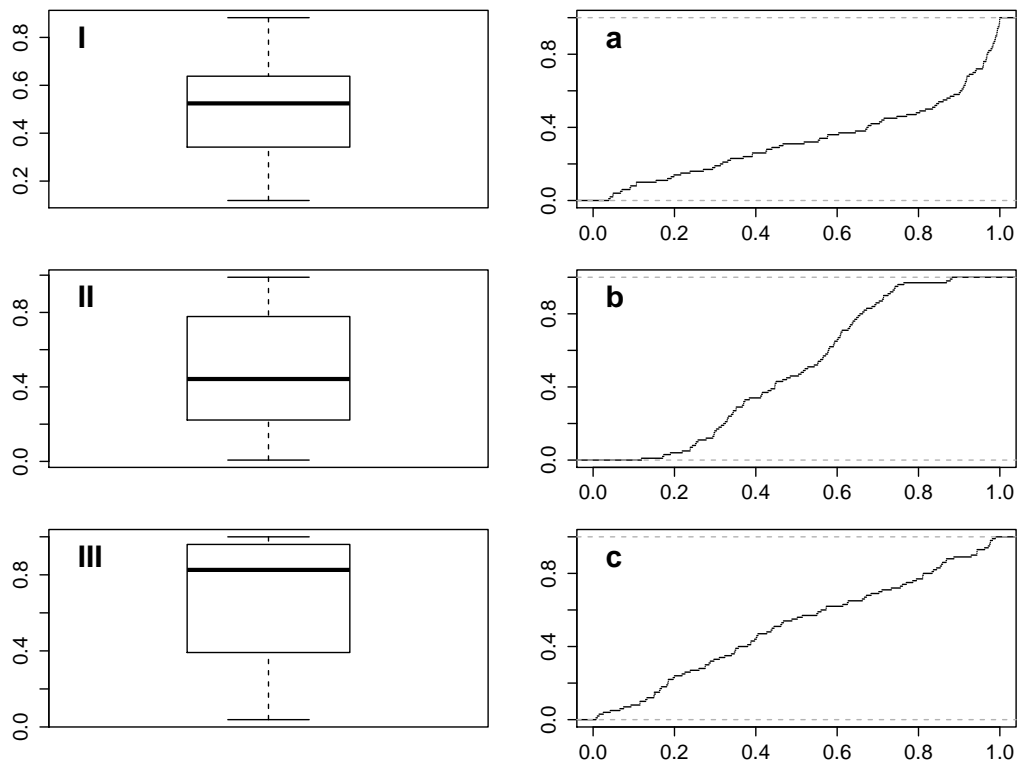
1. Die Wahrscheinlichkeit für einen Fehler 1. Art ist ungefähr 95%.
2. Die Wahrscheinlichkeit für einen Fehler 2. Art bei $p = 0.8$ ist ungefähr 26%.
3. Die Macht des Tests bei $p = 0.8$ ist ungefähr 26%.

g) Wird bei einem Z-Test anstelle eines 95%-Vertrauensintervalls ein 99%-Vertrauensintervall betrachtet (deren Längen wir mit $L_{0.95}$ und $L_{0.99}$ bezeichnen), so ist das Verhältnis der Länge der beiden Intervalle $\frac{L_{0.95}}{L_{0.99}}$ gegeben durch (wobei z_q das q-Quantil der Standardnormalverteilung ist):

1. $\frac{z_{0.975}}{z_{0.995}}$
2. $\frac{z_{0.95}}{z_{0.99}}$
3. Es sind zu wenige Informationen verfügbar um dies zu beurteilen.

Siehe nächstes Blatt!

- h) Es seien X_1, \dots, X_n unabhängig und identisch verteilt mit $X_i \sim \text{Pois}(\lambda)$. Dann ist $\frac{1}{n} \sum_{i=1}^n X_i$
1. Ein Schätzer für λ , aber nicht erwartungstreu.
 2. Ein erwartungstreuer Schätzer für λ .
 3. $\mathcal{N}(\lambda, \frac{\lambda}{n})$ -verteilt für $n = 10$.
- i) In einer medizinischen Fachzeitschrift wird ein zweiseitiger t-Test mit Nullhypothese $H_0 : \mu = 0$ angewendet um zwei Behandlungsmethoden zu vergleichen. Beide Methoden wurden an 100 Patienten erprobt. Die Wissenschaftler haben aus den Daten einen p-Wert von 3% errechnet. Dann gilt:
1. Ein zweiseitiger Test zum Niveau 1% wird die Nullhypothese verwerfen.
 2. Das entsprechende 99%-Vertrauensintervall enthält den Wert 0.
 3. Um Aussagen 1. und 2. beurteilen zu können benötigen Sie weitere Angaben aus dem Forschungsartikel.
- j) Für drei Stichproben vom Umfang $n = 100$ wurden je ein Boxplot und die empirische Verteilungsfunktion gezeichnet. Welche Zuordnung der Boxplots zu den entsprechenden empirischen Verteilungsfunktionen ist korrekt?



1. Ia, IIc, IIIb.
2. Ic, IIb, IIIa.
3. Ib, IIc, IIIa.

Bitte wenden!

- 2. (Total 8 Punkte)** Alice und Bob arbeiten in zwei verschiedenen Labors an der Herstellung von Nanopartikeln. Um aus der Ausgangsprobe ein einzelnes Partikel herzustellen ist eine Abfolge von Schritten nötig, die einen ganzen Tag in Anspruch nehmen. An jedem Werktag wird also ein Partikel hergestellt. In diesem Prozess gibt es ein paar schwierige Schritte, die beide noch nicht ganz beherrschen. Wird die Vorbereitung von einem Laboranten übernommen, so gelingt Alice die Herstellung in 80% der Fälle, Bob immerhin noch in 70% der Fälle. Manchmal ist jedoch kein Laborant zur Stelle, dann misslingt Alice der Herstellungsprozess in 70% der Fälle und Bob in 60% der Fälle. Erfahrungsgemäss arbeitet der Laborant in Alices Labor an 75% der Werktage, in Bobs Labor sind es sogar 90%. Wir nehmen an, dass die Herstellung in den beiden Labors unabhängig voneinander ist.
- a) **(2 Punkte)** Alices Betreuerin besucht das Labor an einem zufällig ausgewählten Tag. Wie gross ist die Wahrscheinlichkeit, dass Alice an diesem Tag erfolgreich ein Nanopartikel herstellt?
- b) **(2 Punkte)** An einem zufällig ausgewählten Tag besuchen beide Betreuerinnen das Labor ihrer Studenten Alice bzw. Bob. Danach wollen Sie sich zu viert auf ein Bier treffen, dies aber nur, wenn die Herstellung in beiden Labors gelungen ist. Berechnen Sie die Wahrscheinlichkeit, dass die vier auf das Bier verzichten müssen.
- c) **(2 Punkte)** Angenommen man besucht Alices Labor am Ende eines zufällig gewählten Tages und ihr ist die Herstellung gelungen. Berechnen Sie die Wahrscheinlichkeit, dass ihr der Laborant an diesem Tag geholfen hat.
- d) **(2 Punkte)** Da die Schweiz Steuern senken möchte, sollen Forschungsgelder eingespart werden. Daher dürfen der Laborant in Alices Labor und der Laborant in Bobs Labor zusammen nur noch ein 100%-Pensum an Anstellung haben. Die Universitätsleitung fragt sich nun wie sie dieses optimal aufteilen soll. Wir betrachten also dasselbe Setup wie oben, nur dass nun $p \in [0, 1]$ die Wahrscheinlichkeit ist, dass in Alices Labor ein Laborant anwesend ist und $1 - p$ die Wahrscheinlichkeit, dass in Bobs Labor ein Laborant anwesend ist. Bestimmen Sie p so, dass die Wahrscheinlichkeit, dass an einem zufällig gewählten Tag beide Labors erfolgreich ein Nanopartikel herstellen, maximiert wird.

Siehe nächstes Blatt!

3. (Total 9 Punkte) Der Fabrikant F. Engels stellt in Wädenswil Schokoladeneis her. Die Anlage ist leider ziemlich veraltet und produziert täglich etwas zuviel oder zuwenig. Der angepeilte Wert ist 50 Tonnen, die Abweichung U (in Tonnen) davon ist $\text{Unif}(-\vartheta, \vartheta)$ -verteilt, wobei $\vartheta = 3$.

- a) **(1 Punkt)** Wie gross ist die Wahrscheinlichkeit, dass die Menge an Eis, die an einem bestimmten Tag produziert wird, mehr als eine Tonne vom angepeilten Wert abweicht?
- b) Da es in der Fabrik einen grossen Kühlraum gibt, spielt es keine Rolle, wenn zuviel Eis produziert wird. Heute morgen hat F. Engels keinerlei Eis auf Lager. Er erhält einen grossen Auftrag: Er hat 30 Tage Zeit um 1515 Tonnen Eis zu produzieren.
- (i) **(1 Punkt)** Wie gross muss die durchschnittliche tägliche Abweichung vom Sollwert von 50 Tonnen mindestens sein, damit er dieses Ziel erreicht?
- (ii) **(1 Punkt)** Unter der Annahme, dass die Abweichungen vom angepeilten Wert für jeden der 30 Tage unabhängig sind, durch welche Verteilung kann man die durchschnittliche tägliche Abweichung approximieren? (Geben Sie sowohl die Art der Verteilung als auch den/die zugehörigen Parameter an).
- (iii) **(2 Punkte)** Wie gross ist die Wahrscheinlichkeit, dass er die verlangte Menge an Eis liefern kann? Benutzen Sie zur Berechnung die Approximation aus (ii).

F. Engels ist sich nun etwas unsicher über den Wert von ϑ . Er weiss also, dass die Abweichung an einem bestimmten Tag $\text{Unif}(-\vartheta, \vartheta)$ -verteilt ist, allerdings ist ϑ unbekannt. Er überwacht die Eisproduktion an 5 Tagen und beobachtet folgende Abweichungen vom Sollwert:

Tag Nr.	1	2	3	4	5
Abweichung (in Tonnen)	1.9	-2.3	1.6	0.5	-1.2

- c) **(2 Punkte)** Leiten Sie einen Momentenschätzer für ϑ her und werten Sie diesen aus für die angegebene Stichprobe. **Tipp:** Sie können das 2. Moment verwenden.
- d) **(2 Punkte)** Leiten Sie den Maximum-Likelihood Schätzer für ϑ her und werten Sie diesen aus für die angegebene Stichprobe.

Bitte wenden!

4. **(Total 9 Punkte)** Im Alltag von Frau Fuchs und Herrn Hürzeler bleibt es nur wenig Zeit für Hausarbeit. Sie möchten sich daher einen Saugroboter kaufen, der selbständig ihre riesige gemeinsame Wohnung mit 99 Zimmern staubsaugt. Leider können sie sich nicht auf ein Modell einigen und entscheiden daher, sowohl den „Neato Signature“ als auch den „Ecovacs Deboot D76“ zu kaufen und selbst zu testen, welcher von beiden schneller saubermacht. Die Hersteller geben an, dass die Roboter gleich leistungsfähig sind. Frau Fuchs glaubt jedoch, dass „Ecovacs“ schneller ist und möchte dies mit einem statistischen Test zum Niveau 5% überprüfen.

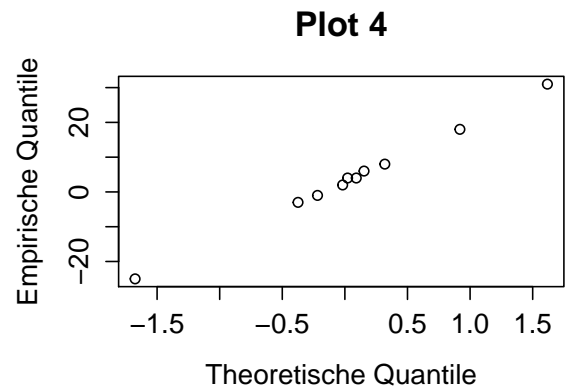
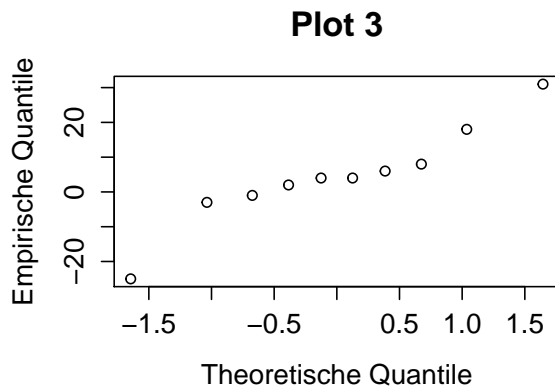
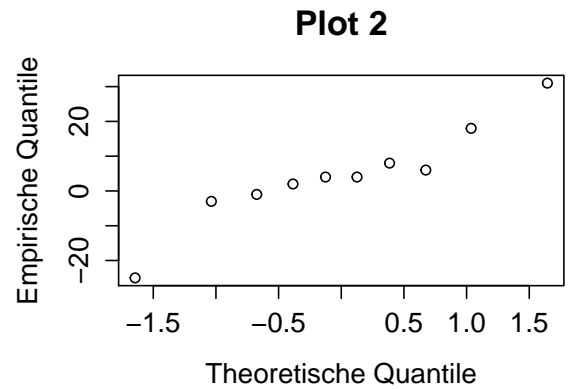
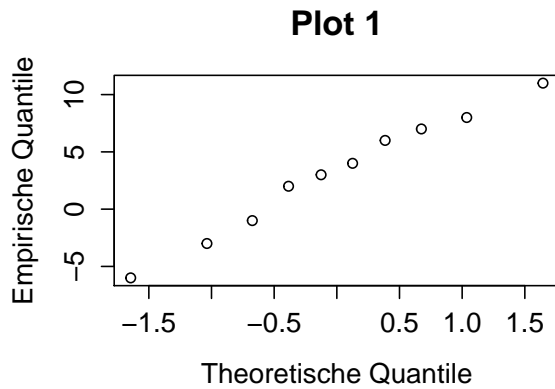
Dazu misst sie in einer Auswahl von Zimmern ihrer Wohnung die Saugzeit und erhält folgende Zahlen (in Minuten):

Zimmer Nr (i)	1	2	3	4	5	6	7	8	9	10
Neato (a_i)	12	10	8	7	40	20	15	50	13	31
Ecovacs (b_i)	15	6	9	5	22	12	11	19	7	56
$d_i = a_i - b_i$	-3	4	-1	2	18	8	4	31	6	-25

Es ergeben sich daraus die folgenden empirischen Kennwerte (in Minuten): Mittelwert: $\bar{a}_{10} = 20.6$, $\bar{b}_{10} = 16.2$, $\bar{d}_{10} = 4.4$. Streuung: $s_a = 14.8$, $s_b = 15$, $s_d = 14.4$. Gepoolte Streuung von a und b : $s_{pool} = 14.9$.

- a) **(1 Punkt)** Ist die Stichprobe gepaart oder ungepaart?
- b) **(2 Punkte)** Welche der folgenden Graphiken ist der korrekte Normal-QQ-Plot der Differenzen (d_1, \dots, d_{10}) (in Minuten)? Begründen Sie Ihre Antwort, d.h. geben Sie für jeden der anderen Plots eine Begründung warum dieser nicht der geeignete QQ-Plot ist.

Siehe nächstes Blatt!



c) (2 Punkte) Ist die Annahme der Normalverteilung für die für den Test relevanten Daten gerechtfertigt? Geben Sie auch hier eine Begründung an. Welcher Test ist daher (unter den üblichen i.i.d. Annahmen) am besten geeignet um die gegebenen Informationen bestmöglich auszunutzen?

d) (4 Punkte) Führen Sie den Test durch, für den Sie sich in b) entschieden haben, d.h.

- formulieren Sie die Null- und Alternativhypothese
- berechnen Sie den realisierten Wert der Teststatistik
- geben Sie den Verwerfungsbereich an
- bestimmen Sie den Testentscheid. Kann nachgewiesen werden, dass „Ecovacs“ schneller ist?