Stochastik für Informatik TU Berlin SoSe 23

–Klausur vom 06.10.23 –

1. Bitte füllen Sie die folgenden Felder aus

Name:	(bitte Druckbuchstaben
Vornavme:	verwenden)
Matrikelnummer:	
Fach:	

2. Bitte nehmen Sie die folgenden Regeln für den Wissensteil zur Kenntnis

- Jede richtige Aufgabe ergibt 2 Punkte
- Wird die Aufgabe unvollständig oder in Teilen fehlerhaft gelöst, so erhalten Sie für die korrekt gelösten Teile anteilig Punkte
- Nicht gelöste Aufgaben geben weder einen Punkt noch einen Punktabzug
- Mit der Hälfte der erreichbaren Punkte haben Sie sicher bestanden. Wir behalten uns eine Senkung des Kriteriums vor.

3. Ihre Klausurnummer

Merken Sie sich Ihre Nummer

Aus Datenschutzgründen dürfen wir die Klausurergebnisse nicht in Verbindung mit Ihrer Matrikelnummer im Internet veröffentlichen. Bitte notieren Sie sich daher Ihre **Klausurnummer**, mit deren Hilfe Sie Ihr Klausurergebnis auf ISIS erfahren können.

Das Vorlesungsteam wünscht Ihnen eine erfolgreiche Klausur!

Platz für Notizen und Nebenrechnungen

A. Wahrscheinlichkeitsrechnung *(20 Punkte)*

A.a. Wissensteil 5×2 *=10 Punkte*

1. Drei Mal wird mit einer fairen Münze geworfen. Geben Sie die Ereignismenge Ω an und bestimmen Sie die Wahrscheinlichkeit, dass in den drei Versuchen insgesamt zweimal Zahl drankommt.

Betrachte

$$\Omega = \{(\omega_1, \omega_2, \omega_3) \mid \omega_i \in \{K, Z\} \text{ für alle } i \in [3]\}.$$

Dann gilt

$$\begin{aligned} P(\text{“zweimal Zahl”}) &= P((K, Z, Z)) + P((Z, K, Z)) + P((Z, Z, K)) \\ &= 3 \cdot 0.125 \\ &= 0.375. \end{aligned}$$

Ein Punkt für den richtigen W-Raum und ein Punkt für das Ergebnis.

2. Was bedeutet die **Additivität** und die **Normiertheit** bei einem Wahrscheinlichkeitsmaß P? Geben Sie jeweils die Definition an!

Sei (Ω, P) ein Wahrscheinlichkeitsraum.

Normiertheit: $P(\Omega) = 1$

Additivität: Für $A, B \subseteq \Omega$ mit $A \cap B = \emptyset$ gilt $P(A \cap B) = P(A) + P(B)$.
Ein Punkt pro richtig beschriebene Eigenschaft

3. Die Wahrscheinlichkeit für das Ereignis A beträgt 10%, die Wahrscheinlichkeit für B 40% und die Wahrscheinlichkeit für $A \cup B$ 47%. Wie groß ist die Wahrscheinlichkeit von $A \cap B$, also dass A und B gleichzeitig eintreten?

Es gilt

$$\begin{aligned} P(A \cap B) &= P(A) + P(B) - P(A \cup B) \\ &= 0.1 + 0.4 - 0.47 \\ &= 0.03. \end{aligned}$$

Ein Punkt für korrekte Rechenregeln mit W-keiten und ein Punkt für das richtige Ergebnis.

4. Ein neuronales Netzwerk, dass auf einem Bild-Datensatz Katzen und Hunde unterscheiden soll, erkennt Katzen mit einer Genauigkeit von 95% und Hunde mit einer Genauigkeit von 90%. Wie hoch ist die Genauigkeit auf dem gesamten Datensatz, wenn auf 80% der Bilder Katzen und nur auf 20% der Bilder Hunde zu sehen sind?

Sei R das Ereignis “richtig erkannt”, K das Ereignis “Bild enthält Katzen” und H das Ereignis “Bild enthält Hunde”. Dann gilt nach dem Satz von der Gesamtwahrscheinlichkeit

$$\begin{aligned} P(R) &= P(R|K)P(K) + P(R|H)P(H) \\ &= 0.95 \cdot 0.8 + 0.9 \cdot 0.2 \\ &= 0.94. \end{aligned}$$

Damit beträgt die Genauigkeit auf dem gesamten Datensatz 94 Prozent.
Ein Punkt für den Ansatz mit dem Gesetz der totalen Wahrscheinlichkeit ein halber Punkt für das korrekte Einsetzen und ein halber Punkt für das richtige Ergebnis.

5. Sie haben 3 weiße und 3 schwarze Kugeln, die Sie in **beliebiger Anzahl** in zwei Urnen verteilen können. In jeder Urne muss aber **mindestens eine Kugel** liegen. Wenn das abgeschlossen ist, entscheidet sich eine Kandidat*in erst blind für eine der beiden Urnen mit gleicher Wahrscheinlichkeit und zieht dann ebenfalls blind eine Kugel.

   Wie müssen Sie die Kugeln verteilen, sodass die Wahrscheinlichkeit, dass 'schwarz' gezogen wird, maximal wird. Wie hoch ist diese Wahrscheinlichkeit dann?

Die Kugeln müssen so verteilt werden, dass in einer Urne nur eine schwarze Kugel liegt und in der anderen Urne die übrigen Kugeln, d.h. zwei schwarze und drei weiße. Mit dieser Verteilung gilt

$$P(\text{“schwarze Kugel”}) = 0.5 \cdot 1 + 0.5 \cdot \frac{2}{5} = 0.7.$$

Ein Punkt für die richtige Idee, ein punkt für das richtige Ergebnis.

A.b.: Textaufgabe $4 \times 2.5 = 10$ *Punkte*

Ein*e Fahrschüler*in macht die Führerscheinprüfung. Gleich am Anfang der Strecke ist ein Stoppschild mit einer durchgezogenen Linie, vor der man anhalten muss. Hieran scheitern schon 10% der Fahrschüler*innen. Etwas später muss die Autobahnauffahrt genommen werden, wobei man den Blinker setzen muss. Den Blinker wird zu 20% vergessen. Danach fährt man durch ein Wohngebiet, wobei die Vorfahrtregel 'rechts vor links' zu beachten ist. In 50% der Fälle kommt im Wohngebiet kein Auto von rechts. In den anderen Fällen beachten die Fahrschüler*innen die Vorfahrtsregel zu 80% korrekt.

Geschieht während der Prüfung ein Fehler, so wird die Prüfung sofort abgebrochen und die/der Fahrschüler*in ist durchgefallen.

Bitte lösen Sie die folgenden Aufgaben unter der Annahme, dass die genannten Prozentzahlen für diese Prüfung gültig und die einzelnen Ereignisse unabhängig sind:

a) Zeichnen Sie den Wahrscheinlichkeitsbaum der Fahrprüfung mit den (bedingten) Wahrscheinlichkeiten an den 'Zweigen',

b) Mit welcher Wahrscheinlichkeit besteht der/die Fahrschüler*in?

c) Mit welcher Wahrscheinlichkeit besteht ein*e Fahrschüler*in, die/der an der Stopplinie korrekt gehalten hat.

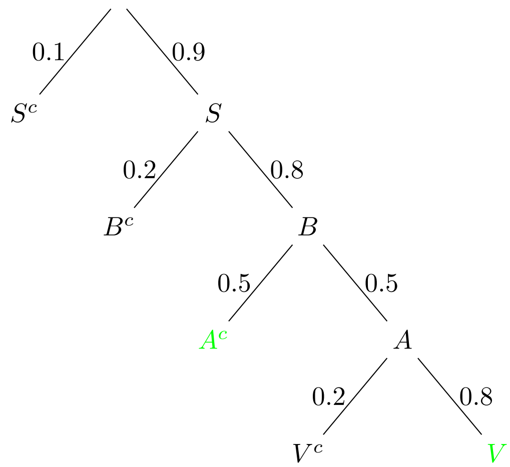
d) Wenn die/der Fahrschüler*in bestanden hat, mit welcher Wahrscheinlichkeit kam dann im Wohngebiet kein Auto von rechts?

Schreiben Sie Ihre Lösung auf die Seiten 5 bis 7.

Teil A–Aufgabe	1	2	3	4	5	Textaufgabe	$\sum$
Punkte							

Lösung zu Aufgabe A.b.1

a) Bezeichne S das Ereignis "am Stoppschild angehalten", B das Ereignis "Blinker gesetzt", A das Ereignis "Auto kommt von rechts" und V das Ereignis "Vorfahrtsregel beachtet". Die Szenarien, in denen der Fahrschüler bestanden hat, sind grün markiert.



b) Es gilt

$$P(\text{“bestanden”}) = 0.9 \cdot 0.8 \cdot 0.5 \cdot (1 + 0.8) = 0.648.$$

c) Es gilt

$$P(\text{“bestanden”}|S) = 0.8 \cdot 0.5 \cdot (1 + 0.8) = 0.72.$$

d) Wir setzen voraus, dass an dem Stoppschild angehalten und der Blinker gesetzt wurde und betrachten nur den untersten Bereich des Wahrscheinlichkeitsbaumes. Dann gilt nach dem Satz von Bayes

$$\begin{aligned} P(A^c|\text{“bestanden”}) &= \frac{P(\text{“bestanden”}|A^c)P(A^c)}{P(\text{“bestanden”}|A)P(A) + P(\text{“bestanden”}|A^c)P(A^c)} \\ &= \frac{0.5}{0.8 \cdot 0.5 + 0.5} \\ &= 0.556. \end{aligned}$$

a) 1,5 Punkte für den Baum und ein Punkt für die W-keiten.
b) Ein Punkt für die richtigen Zweige und 1,5 Punkte für das richtige Ergebnis
c) Ein Punkt für die richtigen Zweige und 1,5 Punkte für das richtige Ergebnis
d) Ein Punkt für die Bayes-Formal, ein halber Punkt für's richtige einsetzen und ein halber Punkt für das richtige Ergebnis.

B. Zufallsvariablen, Verteilungen und Grenzwertsätze *(20 Punkte)*

B.a. Wissensteil 5×2 *=10 Punkte*

1. Geben Sie je eine diskrete und eine kontinuierliche Verteilung an, mit denen man **Wartezeiten** modelliert und geben Sie die (Wahrscheinlichkeits-)Dichten an!

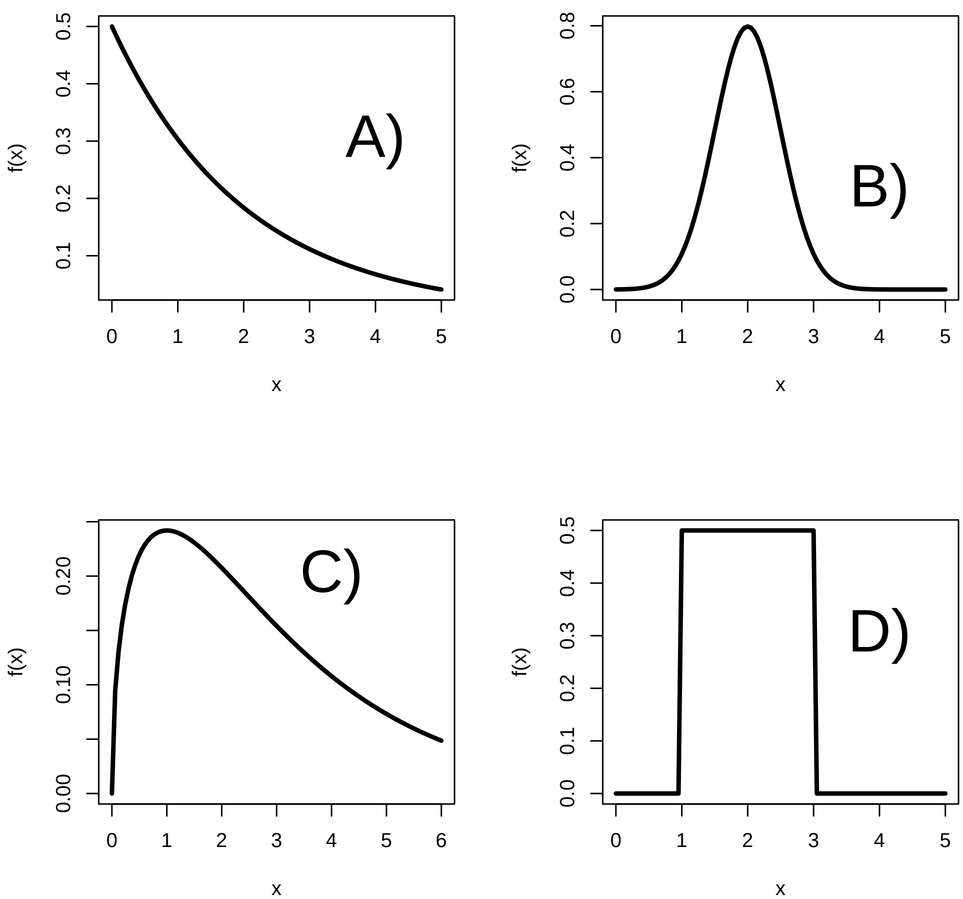
Diskret: Geometrische Verteilung mit Dichte $f(x) = p(1-p)^{x-1}$

Kontinuierlich: Exponentialverteilung mit Dichte

$$f(x) = \begin{cases} 0, & x < 0, \\ \lambda e^{-\lambda x}, & x \geq 0. \end{cases}$$

Je ein halber Punkt für je einen richigen Namen und ein halber Punkt für die richtige Formel für die Dichten

2. Die untenstehende Grafik zeigt vier Dichtefunktionen.



Ordnen Sie diese Dichtefunktionen den folgenden vier Verteilungen zu, indem Sie den entsprechenden Buchstaben hinter die richtige Verteilung schreiben.

- Uniforme Verteilung $U(a = 1, b = 3)$ D
- Exponentialverteilung $\text{Exp}(\lambda = 0.5)$ A
- Normalverteilung, $N(\mu = 2, \sigma^2 = 0.25)$ B
- Chi-Quadrat Verteilung $\chi^2(n = 3)$ C

Je ein halber Punkt pro richtige Zuordnung.

3. Betrachten Sie die folgende Funktion

$$f(x) = \begin{cases} 0 & x < 0 \\ Ce^{-x} & 0 \leq x \leq 2 \\ 0 & x > 2 \end{cases}$$

Welchen Wert muss die die Konstante $C > 0$ annehmen, sodass $f(x)$ eine kontinuierliche Dichtefunktion ist. Berechnen Sie C!

Es gilt

$$\int_{-\infty}^{\infty} f(x)\mathrm{d}x = \int_0^2 Ce^{-x}\mathrm{d}x = \left[-Ce^{-x}\right]_0^2 = C(1 - e^{-2}).$$

D.h.

$$\int_{-\infty}^{\infty} f(x)\mathrm{d}x = 1 \quad \Leftrightarrow \quad C = \frac{e^2}{e^2 - 1}.$$

Damit gilt auch $f(x) \geq 0$ für alle x.

Ein Punkt für den richtigen Ansatz, ein halber Punkt für das richtige Integral und ein halber punkt für das richtige Endergebnis.

4. Eine Zufallsvariable X kann die Werte $0, 1, 2$ annehmen, wobei jeder Wert mit gleicher Wahrscheinlichkeit $P(X = x) = \frac{1}{3}$, $x \in \{0, 1, 2\}$, angenommen wird.

Berechnen Sie den Erwartungswert und die Varianz von X.

Es gilt

$$\begin{aligned} E[X] &= \sum_{k \in \Omega(X)} k \cdot P(X = k) = 0 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} + 2 \cdot \frac{1}{3} = 1, \\ E[X^2] &= \sum_{k \in \Omega(X)} k^2 \cdot P(X = k) = 0^2 \cdot \frac{1}{3} + 1^2 \cdot \frac{1}{3} + 2^2 \cdot \frac{1}{3} = \frac{5}{3}, \\ \mathrm{Var}(X) &= E[X^2] - E[X]^2 = \frac{5}{3} - 1^2 = \frac{2}{3}. \end{aligned}$$

Ein halber Punkt für den richtigen Ansatz für die Summen und je ein halber Punkt für die richtige Auswertung der Summen und ein halber Punkt für das richtige Eendergebnis.

5. Sind die Zufallsvariablen X und Y mit der gemeinsamen Dichtefunktion unabhängig? Begründen Sie!

$$f_{X,Y}(x,y) = \begin{cases} 6\,e^{-3x-2y} & x,y>0 \\ 0 & \text{sonst} \end{cases}$$

Die Zufallsvariablen X und Y sind unabhängig, da die gemeinsame Dichtefunktion als Produkt der einzelnen Dichten geschrieben werden kann (betrachte $X \sim \text{Exp}(3)$ und $Y \sim \text{Exp}(2)$).
Ein Punkt für die richtige Antwort und ein Punkt für die richtige Begründung.

B.b.: Textaufgabe $4 \times 2,5 = 10$ *Punkte*
Eine Internetseite wird pro Tag von durchschnittlich 10 000 Kund*innen besucht. Dabei liegt die Standardabweichung der Besucherzahlen bei 1 500 Kund*innen.

Bearbeiten Sie die folgenden Fragestellungen unter der Annahme, dass die zufällige tägliche Anzahl X von Kund*innen normalverteilt ist (wir vernachlässigen, dass die Anzahl von Kund*innen immer eine natürliche Zahl sein muss). Geben Sie ihre Rechnungen an oder begründen Sie kurz!

a) Skizzieren Sie den Verlauf der **Dichtefunktion** der Besucherzahlen in der unten stehenden Grafik und zeichnen Sie die 1σ, 2σ und 3σ Umgebungen ein.

b) Mit welcher Wahrscheinlichkeit liegen die Besucherzahlen zwischen 8 500 und 11 500 Kund*innen?

c) Geben Sie eine Besucherzahl an, die zu 90% überschritten wird.

d) Mit welcher Wahrscheinlichkeit hat die Seite an einem Tag mehr als 12 000 Besucher*innen?

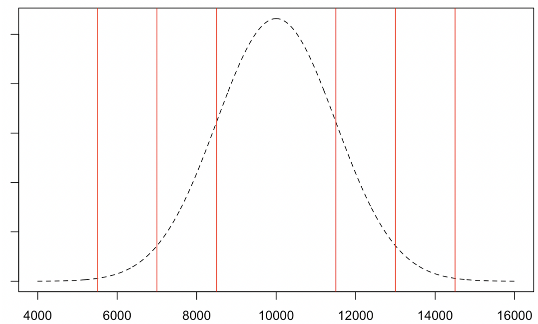
Hinweis: Die Werte der Verteilungsfunktion der Standardnormalverteilung sowie deren Quantile entnehmen Sie bitte den Tabellen am Ende der Klausur. Beachten Sie auch die Relation $z_\alpha = -z_{1-\alpha}$ für die Quantile der Standardnormalverteilung.
Schreiben Sie Ihre Lösung auf die Seiten 11–13.

Teil B–Aufgabe	1	2	3	4	5	Textaufgabe	$\sum$
Punkte							

Lösung zu Aufgabe B.b.1

Lösung zu Aufgabe B.b.1

a)



b) Mit $\mu = 10000$ und $\sigma = 1500$ gilt

$$\begin{aligned} P(8500 \leq X \leq 11500) &= P(X \leq 11500) - P(X \leq 8500) \\ &= P\left(\frac{X-\mu}{\sigma} \leq \frac{11500-\mu}{\sigma}\right) - P\left(\frac{X-\mu}{\sigma} \leq \frac{8500-\mu}{\sigma}\right) \\ &= \Phi\left(\frac{11500-\mu}{\sigma}\right) - \Phi\left(\frac{8500-\mu}{\sigma}\right) \\ &= \Phi(1) - \Phi(-1) \\ &= 0.8413 - 0.1587 \\ &= 0.6826. \end{aligned}$$

c) Gesucht ist x, sodass $P(X > x) = 0.9$. Es gilt

$$\begin{aligned} P(X \leq x) = \Phi\left(\frac{x-\mu}{\sigma}\right) = 0.1 \quad &\Leftrightarrow \quad \frac{\mu - x}{\sigma} = 1.282 \\ &\Leftrightarrow x = 8077.68. \end{aligned}$$

D.h. eine Besucherzahl von 8077 wird zu 90 Prozent überschritten.

d) Es gilt

$$\begin{aligned} P(X > 12000) &= 1 - P(X \leq 12000) \\ &= 1 - \Phi\left(\frac{12000-\mu}{\sigma}\right) \\ &= 1 - \Phi(1.333) \\ &= 1 - 0.9082 \\ &= 0.0918. \end{aligned}$$

a) Ein Punkt für die Kurvenform, win halber Punkt für richtige Lage (Erwartungswert) und ein Punkt für die richtigen 1,2,3-σ Umgebungen.

b) Ein halber Punkt für die richtige zu berechnende Wahrscheinlichkeit, ein Punkt für korrektes Standardisieren, ein halber Punkt für den korrekten Umgang mit der Verteilungsfunktion und ein halber Punkt für das richtige Endergebnis.

c) Ein Punkt für den richtigen Ansatz, ein Punkt für das richtige $z_{0.1}$-Quantil (0.5 neg. Vorzeichen, 0.5 Absolutwert) ein halber punkt für das richtige Endergebnis.

d) Ein halber Punkt für die richtige zu berechnende Wahrscheinlichkeit, ein Punkt für korrektes Standardisieren, ein halber Punkt für den korrekten Umgang mit der Verteilungsfunktion und ein halber Punkt für das richtige Endergebnis.

C. Statistik *(20 Punkte)*

C.a. Wissensteil 5×2 *=10 Punkte*

1. Beschreiben Sie kurz, wozu Parameterschätzer benötigt werden!

In der Regel werden Parameterschätzer benötigt, um mithilfe gegebener Daten Aussagen über eine unbekannte Grundgesamtheit herzuleiten. Dabei wird angenommen, dass die Daten Realisierungen von Zufallsvariablen sind, deren Verteilung aus einer Familie von Wahrscheinlichkeitsverteilungen stammen. Diese Wahrscheinlichkeitsverteilungen hängen von einem unbekannten Parameter ab, der geschätzt werden soll, um mittels der Stichprobe Aussagen über die zugrunde liegenden Zufallsvariablen bzw. über die gesamte Population treffen zu können.
Einen Punkt für die Erwähnung von parametrischen Verteilungsfamilien und ein Punkt für die Erwähnung von Daten.

2. Wie ist ein Konfidenzintervall $[\hat{\theta}_n^-, \hat{\theta}_n^+)$ für den Parameter θ mit Konfidenzniveau $1-\alpha$ definiert? Wir betrachten die Fragestellung für ein gegebenes statistisches Modell $\{P_\theta\}_{\theta\in\Theta}$ und den Datenprozess $X_1, \ldots, X_n$.

   i) Wie sind die Grenzen $\hat{\theta}_n^\pm$ des Konfidenzintervalls ganz allgemein definiert und wovon hängen sie ab?

   ii) Geben Sie die definierende Gleichung für das Konfidenzintervall zum Konfidenzniveau $1-\alpha$ an!

i) Die Grenzen $\hat{\theta}_n^\pm$ sind als Schätzer für θ definiert, sodass das Paar $\hat{\theta}_n^-, \hat{\theta}_n^+$ auch Intervallschätzer für θ genannt wird. Hierbei hängen die Schätzer von dem Datenprozess $X_1, \ldots, X_n$ ab.

ii) Die definierende Gleichung für das Konfidenzintervall zum Konfidenzniveau $1-\alpha$ ist

$$P_\theta(\theta \in [\hat{\theta}_n^-, \hat{\theta}_n^+)) = 1 - \alpha.$$

Je ein Punkt für die korrekte Antwort auf eine Teilaufgabe.

3. Geben Sie die Definition des Maximum Likelihood Prinzips an! Beschreiben Sie das Prinzip kurz in ihren eigenen Worten (nicht eins zu eins aus dem Buch abschreiben).

Den Ausgangspunkt für die Maximum-Likelihood-Methode bilden beobachtete Daten $x_1, \ldots, x_n$. Es wird angenommen, dass dies Realisierungen von u.i.v. Zufallsvariablen sind, deren Dichte P_θ von einem unbekannten Parameter θ abhängt. Ziel der Maximum-Likelihood-Methode ist die Bestimmung des Parameters, dessen zugehörige Verteilung die Daten mit der höchsten Wahrscheinlichkeit erklärt. Im Allgemeinen ist hierzu das Optimierungsproblem

$$\hat{\theta}_{ML} \in \underset{\theta \in \Theta}{\arg\max}\, P_\theta^{(n)}(\{(x_1, \ldots, x_n)\})$$

zu lösen.

Ein Punkt für den Bezug auf ein parametriertes statistisches Modell, ein Punkt für die maximierung der W.-keiten unter besagtem Modell.

4. Betrachten Sie die Exponentialverteilung parametrisiert über den Ratenparameter $\eta = \lambda^{-1}$, also

$$f_\eta(x) = \frac{1}{\eta} e^{-x/\eta}, \quad x > 0.$$

Stellen Sie die Maximum-Likelihood-Gleichungen für den Skalenparameter $\eta > 0$ für die Daten $x_1, \ldots, x_n$, $x_j > 0$, $j = 1, \ldots, n$, auf und bestimmen Sie den Maximum-Likelihood-Schätzer $\hat{\eta}_{\mathrm{ML}}$.

Für $x > 0$ ist

$$L(\eta|x) = \prod_{i=1}^{n} f(x_i|\eta) = \prod_{i=1}^{n} \eta^{-1} e^{-\frac{x_i}{\eta}}.$$

Damit ist die Log-Likelihood-Funktion

$$l(\eta|x) = \sum_{i=1}^{n} \left[\ln\left(\frac{1}{\eta}\right) - \frac{x_i}{\eta} \right].$$

Es gilt

$$\frac{\partial l(\eta|x)}{\partial \eta} = -\frac{n}{\eta} + \frac{1}{\eta^2} \sum_{i=1}^{n} x_i = 0 \quad \Leftrightarrow \quad \hat{\eta}_{\mathrm{ML}} = \frac{1}{n} \sum_{i=1}^{n} x_i.$$

Ein halber Punkt für die korrekte Likelihood, ein Punkt für die ML-Gleichungen, ein halber Punkt für die korrekte berechnung des ML-Schätzers.

5. In einem statistischen Test wird ein p - Wert von 0.0073 angegeben. Beantworten Sie die untenstehenden Fragen und **begründen** Sie kurz!

   i) Wird die Gegenhypothese H_1 bei einem gegebenen Signifikanzniveau 1% angenommen?

   ii) Wann ist eine statistische Aussage H_1 eigentlich stärker durch Daten belegt - wenn der p-Wert größer wird oder wenn er kleiner wird?

i) Da $0.0073 < 0.01$, wird die Nullhypothese abgelehnt.

ii) Je kleiner der p - Wert, desto unwahrscheinlicher ist es, dass die beobachteten Daten unter der Nullhypothese auftreten. D.h. H_1 ist stärker durch Daten belegt, wenn der p - Wert kleiner wird.

Je ein Punkt pro Teilaufgabe.

C.b.: Textaufgabe *($4 \times 2.5 = 10$ Punkte)*
Eine Firma, die Spülmaschinen herstellt, misst den Energieverbrauch ihrer Maschine pro Spülgang. Hierbei wird davon ausgegangen, dass die einzelnen Messungen eine normalverteilte Streuung aufweisen und unabhängig voneinander sind.

Spülgang	1	2	3	4	5
Energieverbrauch in kWh	0.8044	0.8064	0.8478	0.8056	0.7335

Bearbeiten Sie die folgenden Aufgaben unter der Annahme, dass bei der Messung des Energieverbrauchs eine normalverteilte statistische Streuung auftritt und dass die einzelnen Messungen statistisch unabhängig voneinander sind.

a) Berechnen Sie das arithmetische Mittel und die Standardabweichung für die Werte.

b) Bestimmen Sie ein zweiseitiges 95%-Konfidenzintervall für den wahren Wert μ des Energieverbrauches pro Spülgang.

c) Ein Premium-Modell der Konkurrenz verbraucht 0.78kWh pro Spülgang. Ist es bei einem Konfidenzniveau von 90% sicher, dass die Konkurrenzmaschine sparsamer ist?

d) Für die Einstufung der Maschine in die Verbrauchsklasse mit dem niedrigsten Verbrauch 'A+++' darf die Maschine höchstens 0.8464 kWh pro Spülgang verbrauchen. Können sich die Entwickler*innen zu 99% sicher sein, dass ihre Maschine dieser Effizienzklasse entspricht?

Hinweis: Die Quantile der Normalverteilung und der t-Verteilung finden Sie im Anhang dieser Klausur.
Schreiben Sie die Lösung der Aufgaben auf die Seiten 16-18

Lösung zu Aufgabe C.b.1

a) Es gilt

$$\begin{aligned}\hat{\mu}_n &= \frac{1}{n}\sum_{i=1}^{n} x_i = 0.79954,\\ \hat{\sigma}_n^2 &= \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \hat{\mu}_n)^2 = 0.001699,\\ \hat{\sigma}_n &= 0.041219.\end{aligned}$$

b) Mit $t_{4,0.975} = 2.776$ ist das Konfidenzintervall

$$\left[\hat{\mu}_n - \frac{\hat{\sigma}_n}{\sqrt{n}} t_{n-1,1-\frac{\alpha}{2}}, \hat{\mu}_n + \frac{\hat{\sigma}_n}{\sqrt{n}} t_{n-1,1-\frac{\alpha}{2}}\right] = [0.748368, 0.850712].$$

c) Gesucht ist ein rechtsoffenes Konfidenzintervall zum Niveau 90 %. Mit $t_{4,0.9} = 1.533$ gilt für die untere Schranke

$$\hat{\mu}_n - \frac{\hat{\sigma}_n}{\sqrt{n}} t_{n-1,1-\alpha} = 0.77128 < 0.78.$$

D.h. es ist nicht zu 90 % sicher, dass die Konkurrenzmaschine sparsamer ist.

d) Gesucht ist ein linksoffenes Konfidenzintervall zum Niveau 99 %. Mit $t_{4,0.99} = 3.747$ gilt für die obere Schranke

$$\hat{\mu}_n + \frac{\hat{\sigma}_n}{\sqrt{n}} t_{n-1,1-\alpha} = 0.868611 > 0.8464.$$

D.h. die Entwickler können sich nicht zu 99% sicher sein, dass ihre Maschine der Effizienzklasse entspricht.

a) Je ein halber Punkt für Formel und Ergebnis (arithm. Mittel und emp. Standardabweichung. Halber Pkt Abzug wenn nur Varianz berechnet wurde.
b) Ein Punkt für korrekte Formel, halber punkt für korrektes Quantil und halber Punkt für korrektes Ergebnis.
c) Ein Punkt für korrekte Formel, halber punkt für korrektes Quantil und halber Punkt für korrektes Ergebnis.
d) Ein Punkt für korrekte Formel, halber punkt für korrektes Quantil und halber Punkt für korrektes Ergebnis.

Lösung zu Aufgabe C.b.1

Teil	A	B	C	$\sum$	Note
Punkte					

Verteilungsfunktion der Standardnormalverteilung $F_X(x) = \int\limits_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}\, dz$

	0,00	**0,01**	**0,02**	**0,03**	**0,04**	**0,05**	**0,06**	**0,07**	**0,08**	**0,09**
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0, 6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0, 7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0, 9990
3,1	0,9990	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993
3,2	0,9993	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995	0,9995
3,3	0,9995	0,9995	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997
3,4	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9998

$\tilde{x}_p$-Quantile der Student-t_n-Verteilung							
	p						
n	**0,90**	**0,95**	**0,975**	**0,99**	**0,995**	**0,999**	**0,9995**
1	3,078	6,314	12,706	31,821	63,656	318,289	636,578
2	1,886	2,920	4,303	6,965	9,925	22,328	31,600
3	1,638	2,353	3,182	4,541	5,841	10,214	12,924
4	1,533	2,132	2,776	3,747	4,604	7,173	8,610
5	1,476	2,015	2,571	3,365	4,032	5,894	6,869
6	1,440	1,943	2,447	3,143	3,707	5,208	5,959
7	1,415	1,895	2,365	2,998	3,499	4,785	5,408
8	1,397	1,860	2,306	2,896	3,355	4,501	5,041
9	1,383	1,833	2,262	2,821	3,250	4,297	4,781
10	1,372	1,812	2,228	2,764	3,169	4,144	4,587
11	1,363	1,796	2,201	2,718	3,106	4,025	4,437
12	1,356	1,782	2,179	2,681	3,055	3,930	4,318
13	1,350	1,771	2,160	2,650	3,012	3,852	4,221
14	1,345	1,761	2,145	2,624	2,977	3,787	4,140
15	1,341	1,753	2,131	2,602	2,947	3,733	4,073
16	1,337	1,746	2,120	2,583	2,921	3,686	4,015
17	1,333	1,740	2,110	2,567	2,898	3,646	3,965
18	1,330	1,734	2,101	2,552	2,878	3,610	3,922
19	1,328	1,729	2,093	2,539	2,861	3,579	3,883
20	1,325	1,725	2,086	2,528	2,845	3,552	3,850
21	1,323	1,721	2,080	2,518	2,831	3,527	3,819
22	1,321	1,717	2,074	2,508	2,819	3,505	3,792
23	1,319	1,714	2,069	2,500	2,807	3,485	3,768
24	1,318	1,711	2,064	2,492	2,797	3,467	3,745
25	1,316	1,708	2,060	2,485	2,787	3,450	3,725
26	1,315	1,706	2,056	2,479	2,779	3,435	3,707
27	1,314	1,703	2,052	2,473	2,771	3,421	3,689
28	1,313	1,701	2,048	2,467	2,763	3,408	3,674
29	1,311	1,699	2,045	2,462	2,756	3,396	3,660
∞	1,282	1,645	1,960	2,327	2,577	3,092	3,293
Die Letzte Zeile "∞ënthält die Quantile der Standard-Normalverteilung und gilt in guter Näherung für die t_n-Verteilung mit $n \geq 30$							

Beispiel: Das $p = 99\%$-Quantil der t-Verteilung mit $n = 5$ Freiheitsgraden beträgt

$$t_{0.99}(5) = 3.365$$

$\tilde{x}_p$-**Quantile der** χ^2_n-**Verteilung**									
n	p								
	0,005	0,01	0,025	0,05	0,9	0,95	0,975	0,99	0,995
1	0,000	0,000	0,001	0,004	2,706	3,841	5,024	6,635	7,879
2	0,010	0,020	0,051	0,103	4,605	5,991	7,378	9,210	10,597
3	0,072	0,115	0,216	0,352	6,251	7,815	9,348	11,345	12,838
4	0,207	0,297	0,484	0,711	7,779	9,488	11,143	13,277	14,860
5	0,412	0,554	0,831	1,145	9,236	11,070	12,832	15,086	16,750
6	0,676	0,872	1,237	1,635	10,645	12,592	14,449	16,812	18,548
7	0,989	1,239	1,690	2,167	12,017	14,067	16,013	18,475	20,278
8	1,344	1,647	2,180	2,733	13,362	15,507	17,535	20,090	21,955
9	1,735	2,088	2,700	3,325	14,684	16,919	19,023	21,666	23,589
10	2,156	2,558	3,247	3,940	15,987	18,307	20,483	23,209	25,188
11	2,603	3,054	3,816	4,575	17,275	19,675	21,920	24,725	26,757
12	3,074	3,571	4,404	5,226	18,549	21,026	23,337	26,217	28,300
13	3,565	4,107	5,009	5,892	19,812	22,362	24,736	27,688	29,819
14	4,075	4,660	5,629	6,571	21,064	23,685	26,119	29,141	31,319
15	4,601	5,229	6,262	7,261	22,307	24,996	27,488	30,578	32,801
16	5,142	5,812	6,908	7,962	23,542	26,296	28,845	32,000	34,267
17	5,697	6,408	7,564	8,672	24,769	27,587	30,191	33,409	35,718
18	6,265	7,015	8,231	9,390	25,989	28,869	31,526	34,805	37,156
19	6,844	7,633	8,907	10,117	27,204	30,144	32,852	36,191	38,582
20	7,434	8,260	9,591	10,851	28,412	31,410	34,170	37,566	39,997
21	8,034	8,897	10,283	11,591	29,615	32,671	35,479	38,932	41,401
22	8,643	9,542	10,982	12,338	30,813	33,924	36,781	40,289	42,796
23	9,260	10,196	11,689	13,091	32,007	35,172	38,076	41,638	44,181
24	9,886	10,856	12,401	13,848	33,196	36,415	39,364	42,980	45,558

Hinweis: n steht hier in der Tabelle für den Wert "(Zeilen-1)"×"(Spalten-1)ünd nicht für den Stichprobenumfang!

Teil	A	B	C	$\sum$	Note
Punkte					