D-MATL D-MAVT Neur. Syst. Comp.

Stochastik (401-0603-00L)

Lösungen

1. (9 Punkte) Die Freunde Alice, Bob, Charlie und Doris treffen sich zum Wichteln. Dafür bereitet Doris vier Kärtchen mit ihren Namen darauf vor und gibt sie zusammengefaltet in eine Schüssel. Reihum ziehen daraufhin alle ein Kärtchen aus der Schüssel ohne es zurückzulegen. Die Ziehung gelingt, wenn niemand das Kärtchen mit dem eigenen Namen zieht. Wir bezeichnen mit A_i das Ereignis, dass sich die i-te Person nicht selbst zieht. Aus Symmetriegründen hat man $\mathbb{P}(A_i) = \frac{3}{4}$ für $i = 1, \ldots, 4$.

a) (1 Punkt) Bestimmen Sie die Wahrscheinlichkeit, dass sich die ersten beiden Personen nicht selbst ziehen.

b) (2 Punkte) Bestimmen Sie die Wahrscheinlichkeit, dass sich die zweite Person selbst zieht, gegeben, dass sich die erste Person nicht selbst gezogen hat.

c) (2 Punkte) Sind A_1 und A_2 unabhängig? Begründen Sie Ihre Antwort.

Es gibt 9 Fälle, in denen niemand aus der Gruppe das Kärtchen mit dem eigenen Namen zieht.

d) (2 Punkte) Wie gross ist die Wahrscheinlichkeit, dass mindestens jemand aus der Gruppe das Kärtchen mit dem eigenen Namen zieht (d.h., dass die Ziehung nicht gelingt)? Begründen Sie Ihre Antwort.

Die vier Freunde wiederholen die Ziehungen so lange, bis eine gelingt. Sei X die Anzahl der Versuche, bis sich niemand selbst gezogen hat.

e) (1 Punkt) Wie hoch ist die erwartete Anzahl Versuche, die benötigt werden, bis eine Ziehung gelingt, d.h., bis sich niemand selbst gezogen hat?

f) (1 Punkt) Wie hoch ist die Wahrscheinlichkeit, dass unter 10 Versuchen genau zwei gültige Ziehungen (d.h., zwei Ziehungen in denen sich niemand selbst zieht) dabei sind?

Lösung:

a)

$$\mathbb{P}(A_1 \cap A_2) = \frac{2}{4}\frac{2}{3} + \frac{1}{4}\frac{3}{3} = \frac{7}{12}.$$

b)

$$\mathbb{P}(A_2^c \mid A_1) = 1 - \mathbb{P}(A_2 \mid A_1) = 1 - \frac{\mathbb{P}(A_1 \cap A_2)}{\mathbb{P}(A_1)} = 1 - \frac{\frac{7}{12}}{\frac{3}{4}} = 1 - \frac{7}{9} = \frac{2}{9}.$$

c) Nein.

$$\mathbb{P}(A_1 \cap A_2) = \frac{7}{12} \neq \frac{9}{16} = \mathbb{P}(A_1)\mathbb{P}(A_2).$$

d)

$$\mathbb{P}(\text{niemand zieht sich selbst}) = \frac{9}{4 \cdot 3 \cdot 2} = \frac{3}{8}.$$

Daher $\mathbb{P}(\text{mindestens eine Person zieht sich selbst}) = 1 - \frac{3}{8} = \frac{5}{8} = 0.625$.

e) X ist geometrisch verteilt mit Parameter $p = \frac{3}{8}$. Daher ist $\mathbb{E}[X] = \frac{1}{p} = \frac{8}{3} \approx 2.6667$.

f)

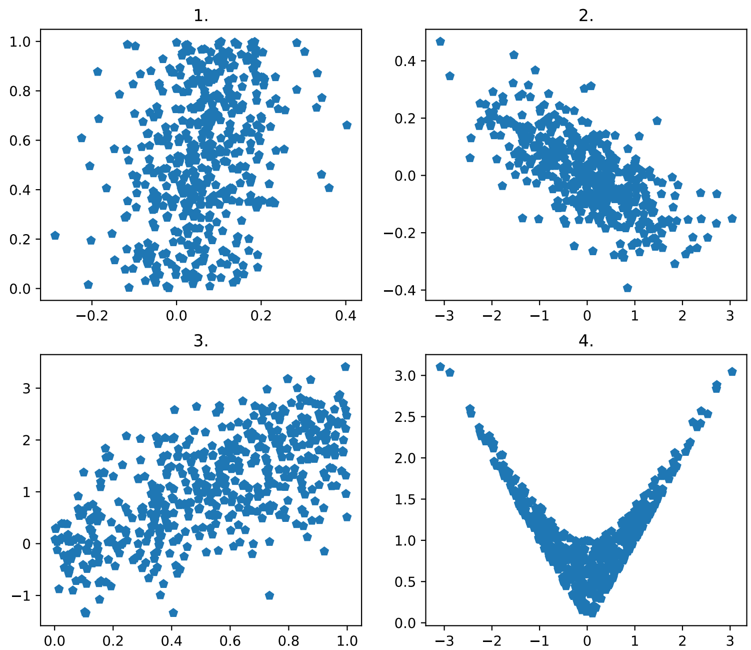
$$\binom{10}{2}(0.375)^2(1 - 0.375)^8 = \frac{10 \cdot 9}{2}(0.375)^2(0.625)^8 \approx 0.147.$$

2. (11 Punkte) Beurteilen Sie die folgenden 11 Aussagen auf ihre Richtigkeit. **Kreuzen Sie Ihre Antworten auf dem zusätzlich angehängten Antwortblatt an**. Nur das Antwortblatt wird ausgewertet. Pro korrekter Antwort gibt es einen Punkt. Es gibt **keinen** Punkteabzug für falsche Antworten.

Ordnen Sie die vier Scatterplots den empirischen Korrelationen

$$i = -0.69, \quad ii = 0.29, \quad iii = -0.07, \quad iv = 0.66$$

zu und bewerten Sie die folgenden Aussagen. (Bem.: Die Zuordnung der Scatterplots gibt keine zusätzlichen Punkte.)



a) (1 Punkt) Es gilt 2.i und 1.ii.

b) (1 Punkt) Es gilt 3.iv und 4.iii.

c) (1 Punkt) Die Ereignisse A_i, $i = 1, \ldots, 5$ sind unabhängig genau dann, wenn

$$\mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i)\mathbb{P}(A_j)$$

für alle $1 \leq i < j \leq 5$.

d) (1 Punkt) Zwei unkorrelierte Zufallsvariablen sind immer auch unabhängig.

e) (1 Punkt) Für zwei unabhängige Zufallsvariablen X, Y und zwei positive Konstanten $a, b > 0$ gilt

$$\text{Cov}(aX + bY) = a\text{Var}(X) + b\text{Var}(Y).$$

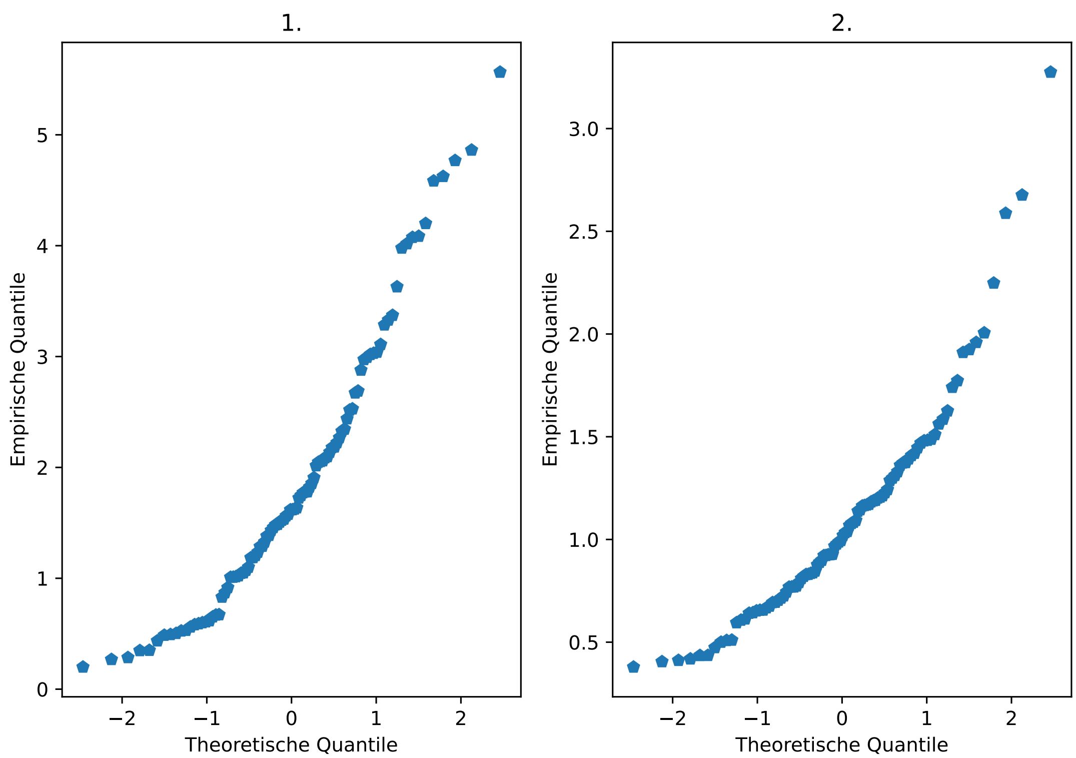
f) (1 Punkt) Für eine normalverteilte Zufallsvariable X und eine Konstante $c \in \mathbb{R}$ gilt

$$\mathbb{P}(X \leq c) = \mathbb{P}(X < c).$$

g) **(1 Punkt)** Für eine standardnormalverteilte Zufallsvariable X gilt, dass $\sigma X - \mu$ normalverteilt ist mit Mittelwert μ und Standardabweichung σ.

h) **(1 Punkt)** Wenn F eine stetige Verteilungsfunktion und $U \sim \text{Unif}([0,1])$ eine uniformverteilte Zufallsvariable sind, dann gilt, dass $F(U)$ eine Zufallsvariable mit Verteilungsfunktion F ist.

Gegeben sind die folgenden Normal-QQ-Plots.



i) **(1 Punkt)** Der Median der 2. empirischen Verteilung ist grösser als jener der 1. empirischen Verteilung.

j) **(1 Punkt)** Für die empirische Verteilung aus dem 1. Plot gilt, dass es wahrscheinlicher ist Werte grösser als 2 zu beobachten als für die Standardnormalverteilung.

k) **(1 Punkt)** Das empirische 10%-Quantil der Verteilung im 2. Plot ist grösser als jenes der Standardnormalverteilung.

Lösung:

a) Wahr.

b) Wahr.

c) Nicht wahr. Paarweise Unabhängigkeit impliziert nicht Unabhängigkeit.

d) Nicht wahr. Unkorreliertheit impliziert nicht Unabhangigkeit.

e) Nicht wahr. Es gilt

$$\text{Cov}(aX + bY) = a^2\text{Var}(X) + b^2\text{Var}(Y).$$

f) Wahr. Für stetige Verteilungen gilt dies.

g) Nicht wahr. $\sigma X - \mu$ ist normalverteilt mit Mittelwert $-\mu$ und Standardabweichung σ.

h) Nicht wahr. Es gilt dann, dass $F^{-1}(U)$ eine Zufallsvariable mit Verteilungsfunktion F ist.

i) Nicht wahr. Der Median der 1. empirischen Verteilung ist grösser.

j) Wahr. Die empirische Verteilung im 1. Plot hat mehr Masse am oberen Ende.

k) Wahr.

3. (9 Punkte) Laut neuem Fahrplan soll in der Stosszeit alle 10 Minuten ein zusätzlicher Bus fahren. Allerdings ist nicht klar ersichtlich, wann die Stosszeit beginnt.

Die Zeit X, die Vergeht bis ein zusätzlicher Bus kommt, ist gemäss der Dichte

$$f(x) = \begin{cases} e^{-(x-\mu)}, & x > \mu \\ 0, & \text{sonst}, \end{cases}$$

verteilt, mit unbekanntem Parameter $\mu > 0$. Basierend auf einer Stichprobe von unabhängigen Wartedauern $X_1, \ldots, X_n$ möchten wir den Parameter μ schätzen.

a) (1 Punkt) Berechnen Sie die erwartete Wartedauer, bis ein zusätzlicher Bus kommt.

b) (1 Punkt) Wie lautet der Momentenschätzer $\hat{\mu}_{MoM}$ von μ?

c) (1 Punkt) Ist der Momentenschätzer $\hat{\mu}_{MoM}$ erwartungstreu? Begründen Sie Ihre Antwort.

d) (2 Punkte) Bestimmen Sie den Maximum-Likelihood-Schätzer $\hat{\mu}_{MLE}$ für μ.

e) (2 Punkt) Bestimmen Sie die Verteilungsfunktion von X.

f) (2 Punkte) Bestimmen Sie die Verteilungsfunktion von $\hat{\mu}_{MLE}$.

Lösung:

a)

$$\begin{aligned}\mathbb{E}[X] &= \int_\mu^\infty x e^{-(x-\mu)}\,dx = -xe^{-(x-\mu)}\Big|_\mu^\infty + \int_\mu^\infty e^{-(x-\mu)}\,dx \\ &= \mu + (-1)e^{-(x-\mu)}\Big|_\mu^\infty = \mu + 1.\end{aligned}$$

b)

$$\mathbb{E}[X] \stackrel{!}{=} \frac{1}{n}\sum_{i=1}^n x_i \iff 1 + \mu = \frac{1}{n}\sum_{i=1}^n x_i \implies \hat{\mu}_{MoM} = \frac{1}{n}\sum_{i=1}^n X_i - 1$$

c) Ja.

$$\mathbb{E}[\hat{\mu}_{MoM}] = \mathbb{E}[X] - 1 = \mu.$$

d) Die Likelihoodfunktion

$$L(\mu|x_1, \ldots, x_n) = \begin{cases} \prod_{i=1}^n e^{-(x_i-\mu)}, & x_i \geq \mu\,\forall i \\ 0, & \text{sonst} \end{cases}$$

ist wachsend in μ, d.h. wir möchten μ so gross wie möglich wählen, ohne, dass einer der Faktoren null wird. Daher ist $\hat{\mu}_{MLE} = \min_{i=1,\ldots,n} X_i$.

e) Für $x \geq \mu$ ist

$$F_X(x) = \int_\mu^x e^{-(y-\mu)}\,dy = -e^{-(y-\mu)}\Big|_\mu^x = 1 - e^{-(x-\mu)},$$

und für $x < \mu$ ist $F_X(x) = 0$.

f)

$$
\begin{aligned}
F_{\hat{\mu}_{MLE}}(x) &= \mathbb{P}\left(\min_{i=1,\dots,n} X_i \leq x\right) = 1 - \mathbb{P}(X_i > x, \forall i = 1\dots,n) \\
&= 1 - \prod_{i=1}^{n} \mathbb{P}(X_i > x) = 1 - \mathbb{P}(X > x)^n = 1 - (1 - F_X(x))^n \\
&= 1 - e^{-n(x-\mu)},
\end{aligned}
$$

falls $x \geq \mu$ und für $x < \mu$ ist $F_{\hat{\mu}_{MLE}}(x) = 0$.

4. (11 Punkte) Wir haben zwei Münzen vor uns liegen und wissen, dass nur eine davon fair ist. Bezeichnen wir mit p die Wahrscheinlichkeit dass ein Münzwurf in *Kopf* resultiert. Wir wissen also, dass für eine der Münzen $p = 0.5$. Bei der anderen Münze wissen wir, dass die Wahrscheinlichkeit *Kopf* zu werfen höher ist.

Wir dürfen nun eine der Münzen nehmen und diese 10 Mal werfen. Wir vermuten, dass es die gezinkte Münze ist. Wie können wir unsere Vermutung statistisch belegen?

Wir entschliessen uns dazu einen Binomialtest durchzuführen um unsere Vermutung zu belegen.

a) (1 Punkt) Formulieren Sie geeignete Hypothesen.

b) (2 Punkte) Unser Test soll die Fehlerwahrscheinlichkeit 1. Art von 5% nicht überschreiten. Bestimmen Sie den maximalen Verwerfungsbereich.

c) (1 Punkt) Bei unseren 10 Münzwürfen bekommen wir 8 Mal *Kopf*. Führt das dazu, dass wir die Nullhypothese verwerfen? Begründen Sie Ihre Antwort.

d) (1 Punkt) Wie gross ist die Fehlerwahrscheinlichkeit 1. Art für einen einseitigen Binomialtest, der die Nullhypothese nach Beobachten von mindestens 8 Mal *Kopf* bei 10 Münzwürfen verwirft?

e) (1 Punkt) Angenommen, wir haben tatsächlich die gezinkte Münze genommen, und $p = 0.7$. Wie hoch ist bei dem Test aus Punkt **b)** die Wahrscheinlichkeit, dass wir die Nullhypothese nicht verwerfen, obwohl $p = 0.7$?

Bewerten Sie folgende Aussagen und begründen Sie Ihre Antwort.

f) (1 Punkt) Wenn wir das Niveau des Tests erhöhen, wird die Fehlerwahrscheinlichkeit 2. Art geringer.

g) (1 Punkt) Wir bezeichnen mit X die Anzahl der Würfe, die in *Kopf* resultieren. Für 10 Münzwürfe ist X approximativ normalverteilt mit Mittelwert p und Varianz $\frac{p(1-p)}{10}$.

Wir führen zum Vergleich auch noch einen einseitigen Z-Test durch.

h) (1 Punkt) Ein solcher Test zum Niveau 0.05 verwirft die Nullhypothese bei 8 Mal *Kopf* unter 10 Münzwürfen.

i) (2 Punkte) Wie viele Münzwürfe dürfen wir höchstens machen, damit die Nullhypothese auf dem 5% Niveau verworfen wird, wenn 7 Mal *Kopf* beobachtet wird?

Lösung:

a)

$$H_0 : p = p_0 = 0.5 \qquad H_A : p > p_0$$

b) Die Anzahl X der Würfe, die unter 10 Würfen in *Kopf* resultieren, ist unter der Nullhypothese Binomialverteilt mit $n = 10$ und $p = p_0 = 0.5$. Wir möchten das kleinstmögliche c finden, sodass

$$0.05 \geq \mathbb{P}(X \geq c \mid H_0).$$

Es gilt

$$\mathbb{P}(X = 10) = 0.5^{10}, \quad \mathbb{P}(X = 9) = 0.5^{10} \cdot 10, \quad \mathbb{P}(X = 8) = 0.5^{10}45,$$

also $\mathbb{P}(X \geq 9) = 0.5^{10}11 = 0.0107 < 0.05 < 0.05468 = 0.5^{10}56 = \mathbb{P}(X \geq 8)$. Der maximale Verwerfungsbereich zum Niveau 0.05 ist also $K = \{9, 10\}$.

c) Nein. Da $8 \notin K$ verwerfen wir die Nullhypothese nicht.

d) $\mathbb{P}(X \geq 8) = 56 \cdot 0.5^{10} = 0.05468$. (p-Wert)

e)

$$\begin{aligned}\mathbb{P}(X \notin K \mid p = 0.7) &= 1 - \mathbb{P}(X \in K \mid p = 0.7) \\ &= 1 - (0.7^{10} + 10 \cdot 0.7^9 0.3) \\ &= 1 - 0.1493 = 0.8507\end{aligned}$$

f) Richtig. Wenn wir das Niveau erhöhen, vergrössern wir den Verwerfungsbereich. Daher wird die Wahrscheinlichkeit (unter der Alternativhypothese) nicht zu verwerfen geringer.

g) Falsch. $X \overset{\text{appr.}}{\sim} \mathcal{N}(10p, 10p(1-p))$.

h) Richtig. Für $Z := \frac{X - 10p_0}{\sqrt{10p_0(1-p_0)}} \overset{\text{appr.}}{\sim}_{H_0} \mathcal{N}(0,1)$ ergibt sich der Verwerfungsbereich zum 5% Niveau als $K = [5 + 1.645\sqrt{\frac{10}{4}}, \infty) = [7.601, \infty)$. Da $8 \in K$ verwirft der Test die H_0.

i) Bei Stichprobengrösse n erhalten wir den Verwerfungsbereich $[0.5n + 1.645\sqrt{n}0.5, \infty)$. Wir möchten also, dass

$$\begin{aligned} & 7 \geq 0.5n + 1.645\sqrt{n}0.5 \\ \iff & 0 \geq n + 1.645\sqrt{n} - 14 \\ \implies & \sqrt{n} \leq -0.5 \cdot 1.645 + \sqrt{1.645^2 \cdot 0.25 + 14} \approx 3.008. \end{aligned}$$

Also $n \leq 9.05$, und wir dürfen höchstens 9 Münzwürfe machen.