ÜBUNGSKLAUSUR

02.10.2020

Rechnen Sie jedes Beispiel auf einem eigenen Blatt. Schreiben Sie auf jedes Blatt ihren Namen und Ihre Matrikelnummer.

1) Wahrscheinlichkeitsverteilungen

a) (3 P) Gegeben sei die Wahrscheinlichkeitsdichte

$$p(x) = \begin{cases} \frac{ax^2+bx+c}{Z} & \text{für } x \in [0,1] \\ 0 & \text{sonst} \end{cases} .$$

Berechnen Sie die Normierung, den Mittelwert und die Varianz dieser Wahrscheinlichkeitsdichte. Die Parameter a,b und c entnehmen Sie auf folgende Weise Ihrer Matrikelnummer: a ist die letzte Ziffer, b die vorletzte Ziffer und c die vorvorletzte Ziffer.

b) (1 P) Wir ändern nun das Vorzeichen des zweiten Terms ($+bx \to -bx$) und erhalten

$$p(x) = \begin{cases} \frac{ax^2-bx+c}{Z} & \text{für } x \in [0,1] \\ 0 & \text{sonst} \end{cases} .$$

Zu welchem Problem könnte diese harmlos scheinende Änderung nun führen?

Hinweis: Wann kann eine Funktion als Wahrscheinlichkeitsdichte verwendet werden?

c) (1 P) Im Folgenden finden Sie eine Tabelle mit Paaren verschiedener Eigenschaften. Bei welchen Eigenschaften würden Sie eine Kovarianz von *0* erwarten? Bei welchen eine Kovarianz größer bzw. kleiner als *0*? Begründen Sie Ihre Antworten in ein bis zwei kurzen Sätzen.

Eigenschaft A	Eigenschaft B
Körpergröße	Körpergewicht
Körpergröße	Schuhgröße
Körpergewicht	Augenfarbe

ZB (2 P) Berechnen Sie die charakteristische Funktion der Gleichverteilung $p_G(x)$ (siehe unten). Zeigen Sie außerdem unter Verwendung der charakteristischen Funktion, dass $Z = b - a$.

$$p_G(x) = \begin{cases} \frac{1}{Z} & \text{für } x \in [a,b] \\ 0 & \text{sonst} \end{cases} .$$

Hinweis: Ausdrücke der Form $C = \frac{0}{0}$ können unter Zuhilfenahme der Regel von de l'Hospital berechnet werden

$$C = \lim_{x\to 0} \frac{f(x)}{g(x)} = \lim_{x \to 0} \frac{f'(x)}{g'(x)}.$$

2) Parameterschätzen

a) (1 P) Mithilfe des MaxEnt-Prinzips soll eine Wahrscheinlichkeitsdichte bestimmt werden, die folgende Bedingungen erfüllen soll:

- Die Normierung ist *1*
- Der Erwartungswert ist *5*
- Die Varianz ist *3*

Stellen Sie die Lagrangefunktion für dieses Problem auf. Die Wahrscheinlichkeitsdichte **muss nicht** ausgerechnet werden.

b) (1 P) Gegeben sei folgende Likelihood

$$p(x|\alpha) = \begin{cases} \alpha e^{-\alpha x} & \text{für x} \geq 0 \\ 0 & \text{sonst} \end{cases} .$$

Bestimmen Sie den Maximum-Likelihood-Schätzwert für eine Messreihe $\vec{x} = \{x_1, x_2, x_3, ..., x_N\}$.

Hinweis: Nutzen Sie folgende Relation $\overline{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$.

c) (2 P) Berechnen Sie als Nächstes den Maximum-a-posterior-Schätzwert. Die Likelihood sei jene aus **(b)**. Der Prior sei

$$p(\alpha) = \frac{1}{\alpha^m}.$$

Welchen Einfluss hat der Prior auf den Schätzwert? Welchen Einfluss hat m auf den Schätzwert? Wie würden sich die Schätzwerte für z.B. $m = 1$ und $m = 10$ unterscheiden? Gehen Sie einfachheitshalber davon aus, dass $N > m$.

Hinweis: Lassen Sie sich nicht vom Schätzwert im **Zusatzbeispiel** beirren.

d) (1 P) Nun wollen wir uns mit der (Un)verzerrtheit von Schätzwerten beschäftigen. Beantworten Sie hierzu folgende Fragen:

- Was unterscheidet einen verzerrten von einem unverzerrten Schätzwert?
- Welchen Schätzwert sollte man verwenden? Einen Verzerrten oder einen Unverzerrten? Wieso?
- Worauf könnte ein verzerrter Schätzwert schließen?

ZB (2 P) Gegegen sei der Maximum-Likelihood-Schätzwert $\alpha_{ML} = \overline{x}$. Bestimmen Sie durch explizites Berechnen, ob dieser Schätzwert verzerrt ist. Verwenden Sie hierzu die Likelihood aus **(b)**. Gehen Sie einfachheitshalber von einem einzigen Messpunkt aus, i.e. $N = 1$.

Ist dieser Schätzwert verzerrt oder nicht? Worauf lässt Ihr Ergebnis schließen?

3) Hypothesentest

Wir haben drei experimentelle Setups zur Verfügung: ein faires und zwei manipulierte Setup. Im fairen Setup seien alle Ausgänge gleich wahrscheinlich, während das manipulierte Setup die jeweiligen Ausgänge mit folgenden Wahrscheinlichkeiten liefert:

Ausgang	Wahrscheinlichkeit
1	1/20
2	2/20
3	3/20
4	4/20
5	4/20
6	3/20
7	2/20
8	1/20

Gehen Sie für die folgenden Berechnungen davon aus, dass die Experimente durch ein Urnenmodell *Ziehen mit Zurücklegen* beschrieben werden können.

a) (1 P) Schreiben Sie zunächst das Theorem von Bayes auf und erklären Sie die Bedeutung der auftretenden Terme anhand unseres Beispiels.

b) (2 P) Es wird nun zufällig ein Setup gewählt. Bei 50 durchgeführten Experimenten tritt der Ausgang 1 insgesamt 10mal ein. Berechnen Sie die Posterior-Wahrscheinlichkeit, dass es sich bei dem Setup, um ein manipuliertes Setup handelt.

c) (1 P) Überprüfen Sie die Hypothesen "faires Setup" und "manipuliertes Setup" mithilfe des ODDS-Ratio. Wie ist das Ergebnis des Hypothesentests zu verstehen? Haben wir nun absolute Gewissheit, ob unser Setup fair oder manipuliert ist?

d) (1 P) Wie würden sich Mittelwert und Standardabweichung eines fairen Setups und eines manipulierten Setups voneinander unterscheiden? Wie könnten Sie Mittelwert und Standardabweichung nutzen, um einem Setup die Eigenschaft "fair" oder "manipuliert" zuzuweisen?

Hinweis: Mittelwert und Standardabweichung müssen **nicht** explizit berechnet werden. Wer möchte kann dies aber gerne tun.

ZB (2 P) Wir wiederholen das Experiment mit dem gleichen Setup und erhalten bei 50 Durchgängen 7 mal den Ausgang 4. Berechnen Sie mithilfe dieser Informationen und jenen aus Aufgabenteil (b) das ODDS-Ratio (Stichwort: Mehrfachtests). Erklären Sie wie der Prior-ODDS das Ergebnis beeinflusst. Worin liegt der Unterschied zwischen Bayes-Faktor und ODDS-Ratio?