A6M33SSL: Domácí úloha: DU3

Varianta: mariklu1

1. **[stat_mm_1m2c, 0.200 b.]** Mějme náhodnou veličinu X, která má rozdělení $R(a, b)$, tj. jakési teoretické rozdělení se dvěma parametry, a a b. O tomto rozdělení víme, že střední hodnota $EX = a \cdot b$ a rozptyl $DX = a \cdot b^2$. Z realizace $\mathbf{x}$ náhodného výběru $\mathbf{X}$ známe realizaci prvního obecného momentu $m_X = m$ a realizaci druhého centrálního momentu $m^*_{X^2} = c$. S využitím metody momentů odvoďte odhady $\hat{a}, \hat{b}$ parametrů a, b jako funkce m a c a tyto odhady vyčíslete. (Výsledky zaokrouhlete na 4 platná místa.)

   Parametry: $m = 8.0$, $c = 3.9$

   Požadované výsledky: $\hat{a}$, $\hat{b}$

2. **[stat_mm_2unif, 0.200 b.]** Mějme náhodnou veličinu X s hustotou pravděpodobnosti

$$f_X(x) = \begin{cases} 0 & \text{pro } x \in (-\infty, 0), \\ 2c & \text{pro } x \in \langle 0, ka), \\ c & \text{pro } x \in \langle ka, a), \\ 0 & \text{pro } x \in \langle a, \infty), \end{cases}$$

   kde $k \in (0, 1)$ je známý parametr tohoto rozdělení, ale a neznáme. (Rozdělení je vlastně vytvořeno ze 2 "stejně vysokých" rovnoměrných rozdělení, jedno je na (kratším) intervalu $\langle 0, ka\rangle$, druhé na (delším) intervalu $\langle 0, a\rangle$.)

   a) Určete konstantu c jako funkci neznámého parametru a a známého parametru k tak, aby výše uvedená funkce byla skutečně hustotou pravděpodobnosti. (Nápověda: Samozřejmě lze počítat integrál od $-\infty$ do $+\infty$ a položit jej roven 1. Ale lze postupovat i geometricky, tj. vyjádřit si obsahy jednotlivých obdélníků a jejich součet položit roven 1.) Pro účely tohoto podúkolu nyní předpokládejme, že hodnotu a známe a že $a = a^*$ (viz parametry úlohy). Vyčíslete konstantu c a výsledek zaokrouhlete na 4 platné číslice.

   b) Vyjádřete střední hodnotu EX tohoto rozdělení jako funkci neznámého parametru a. (Nápověda: Opět lze pro výpočet EX využít definiční vztah s integrálem, nebo lze EX vyjádřit jako vážený průměr středních hodnot 2 rovnoměrných rozdělení.) Rozmyslete si, jaká by měla být střední hodnota rozdělení pro $k = 0$ a pro $k = 1$. Zkontrolujte, zda vámi odvozený vztah pro EX tyto vlastnosti splňuje. Pro účely tohoto podúkolu nyní předpokládejme, že hodnotu a známe a že $a = a^*$ (viz parametry úlohy). Vyčíslete střední hodnotu EX_0 pro $k = 0$ a střední hodnotu EX_1 pro $k = 1$. (V dalších podúkolech je opět a neznámé.)

   c) Získali jsme realizaci náhodného výběru z tohoto rozdělení. Vypočetli jsme výběrový průměr $\bar{\mathbf{x}} = m$. Metodou momentů odvoďte odhad $\hat{a}$ parametru rozdělení a. (Nápověda: vztah pro EX položte roven $\bar{\mathbf{x}}$ a vyřešte rovnici.) Odhad vyčíslete a zaokrouhlete na 4 platné číslice. Pokud by hodnota vámi odvozeného odhadu vyšla nesmyslně, nahraďte ji nejbližší hodnotou, kterou považujete za rozumnou/přípustnou.

   Parametry: $k = 0.8$, $a^* = 10$, $m = 6.1$

   Požadované výsledky: c, EX_0, EX_1, $\hat{a}$

3. **[stat_mle_exam1, 0.200 b.]** Ústní zkoušku ze statistiky se kromě vás v daný termín chystá složit n dalších studentů. Zkoušející je znám tím, že u ústní zkoušky používá pouze stupně A, B, nebo C. Z minulých let známe pravděpodobnosti p_A, p_B a p_C jednotlivých klas. stupňů. Vy půjdete na řadu až poslední, takže můžete sledovat, kolik kterých známek už zkoušející udělil, tj. budete znát sekvenci známek $\mathbf{x}$ délky n obsahující a, b, resp. c udělených známek A, B, resp. C (přičemž $a + b + c = n$).

   a) Určete vztah pro výpočet pravděpodobnosti $p = p(\mathbf{x}|p_A, p_B, p_C)$, s jakou bychom měli sekvenci $\mathbf{x}$ pozorovat, a vyčíslete ji pro známá a, b, c.

   b) Určete věrohodnost $L = L(p_A, p_B, p_C)$ hodnot parametrů rozdělení vzhledem k pozorovaným datům.

   c) Kterou ze známek A, B, nebo C by vám měl zkoušející dát, kdyby chtěl maximalizovat pravděpodobnost výsledků dnešního termínu jako celku (tj. kdyby nebral ohled na vaše znalosti)? Pokud existuje více než jedna možnost, vyberte tu, která je první v abecedě. (Do JSON souboru uveďte buď `"z": "A"`, nebo `"z": "B"`, nebo `"z": "C"`.)

   Výsledky zaokrouhlete na 4 platné číslice.

   Parametry: $p_A = 0.1$, $p_B = 0.45$, $p_C = 0.45$, $a = 1$, $b = 2$, $c = 5$

   Požadované výsledky: p, L, z

4. **[stat_mle_ae-ax, 0.200 b.]** Mějme náhodnou veličinu X s rozdělením popsaným hustotou $f_X(x) = a\mathrm{e}^{-ax}$. Z realizace náhodného výběru z tohoto rozdělení jsme spočetli výběrový průměr $\bar{\mathbf{x}} = m$. Metodou maximální věrohodnosti odvoďte odhad $\hat{a}$ parametru a a vyčíslete jej. Výsledek zaokrouhlete na 4 platné číslice.

   Parametry: $m = 0.2$

   Požadované výsledky: $\hat{a}$

5. **[stat_mle_mm_exam2, 0.200 b.]** Ústní zkoušku ze statistiky se kromě vás v daný termín chystá složit n dalších studentů. Zkoušející je znám tím, že u ústní zkoušky používá pouze stupně A, B, nebo C. Vy půjdete na řadu až poslední, takže můžete sledovat, kolik kterých známek už zkoušející udělil (a Áček, b Béček a c Céček). Narozdíl od minulého příkladu pravděpodobnosti p_A, p_B a p_C neznáme. Jediné, co vám zkoušející sdělí navíc, je to, že dlouhodobě platí, že dává stejně Béček jako Áček a Céček dohromady, tedy že $p_B = p_A + p_C$. Zkoušející po vás při vaší ústní zkoušce chce, abyste neznámé pravděpodobnosti odhadli z dostupných dat (a, b, c), tj. abyste odvodili vztahy pro výpočet odhadů $\hat{p_A}$, $\hat{p_B}$ a $\hat{p_C}$ jako funkce pozorovaných četností a, b, c při respektování podmínky na p_B.

   a) Odvoďte vztahy pro $r_A = \hat{p_A}$, $r_B = \hat{p_B}$ a $r_c = \hat{p_C}$ metodou maximální věrohodnosti a vyčíslete je. (Nápověda: Máme 3 neznámé parametry a 2 rovnice, které pro ně musí být splněny - rovnice pro součet všech 3 pstí, a rovnice pro p_B. Vyjádřete tedy 2 neznámé parametry pomocí toho zbývajícího. Zkonstruujte funkci věrohodnosti jako funkci jediného neznámého parametru. Zlogaritmujte, zderivujte, najděte maximum.)

   b) Odvoďte vztahy pro $q_A = \hat{p_A}$, $q_B = \hat{p_B}$ a $q_C = \hat{p_C}$ metodou momentů a vyčíslete je. Abyste metodu momentů mohli aplikovat, nahraďte písmenné označení klasifikačních stupňů A, B, C číselnými známkami 1, 2, 3. (Nápověda: Vyjádřete střední hodnotu rozdělení jako funkci parametrů p_A, p_B a p_C, resp. jako funkci jediného z nich, protože zbylé 2 parametry lze dopočítat stejně, jako u metody maximální věrohodnosti. Vyjádřete výběrový průměr jako funkci a, b, c. Sestavte rovnici a vyřešte ji pro neznámý parametr, čímž dostanete vztah pro jeho odhad.)

   Vztahy vyčíslete pro známá a, b, c. Výsledky zaokrouhlete na 4 platné číslice. Pokud by hodnota vámi odvozeného odhadu vyšla nesmyslně, nahraďte ji nejbližší hodnotou, kterou považujete za rozumnou/přípustnou.

   Parametry: $a = 2$, $b = 5$, $c = 4$

   Požadované výsledky: r_A, r_B, r_C, q_A, q_B, q_C