MS-A0502 Todennäköisyyslaskennan ja tilastotieteen peruskurssi mallivastaukset

Date of exam: 5.12.2022 **Solved by**: mathwiz.ai

Disclaimer:
These problems are solved by mathwiz.ai, a service leveraging the latest generative AI technologies. This means the solutions come with inherent limitations:

1. **Understanding:** The AI doesn't understand mathematics like humans do; it predicts answers based on historical data patterns.
2. **Reliability and Context:** The AI's responses can be contextually inappropriate or inconsistent, and thus contain errors.
3. **Creativity:** AI applies known patterns, it does not invent new mathematical methods.

Use these solutions as guides, not definitive answers. Despite occasional errors, the solutions can be useful for providing context around the problems. For verifiable accurate solutions, seek guidance from a qualified educator. At the time of solving these problems, the mathematical capabilities of AIs were at the level of an advanced university mathematics student. If you find errors in the solutions, feel free to inform us. Contact information: mathwizcontact@gmail.com.

Problem 1

Herrat A, B ja C ovat metsästämässä ja ampuvat samanaikaisesti kohti samaa hirveä. Heidän todennäköisyytensä osua hirveen ovat $0.2, 0.3$ ja 0.4 , ja osumiset ovat riippumattomia. Olkoot A, B ja C indikaattorimuuttujat, jotka ilmaisevat, osuiko kyseinen herra hirveen vai ei.

(a) Luettele satunnaisvektorin (A, B, C) mahdolliset arvot ja niiden todennäköisyydet kolmella desimaalilla (yhdellä, jos käytät prosentteja). (2p)

(b) Olkoon X hirveen osuneiden luotien lukumäärä. Luettele $X : n$ mahdolliset arvot ja niiden todennäköisyydet. (2p)

(c) Jos hirveen osuu tasan yksi luoti, mikä on todennäköisyys, että se on herran C ampuma? (2p)

Problem 1 solution

(a) Satunnaisvektorin (A, B, C) mahdolliset arvot ja niiden todennäköisyydet ovat seuraavat:

(0, 0, 0): Ei kukaan osu hirveen. Todennäköisyys on $(1 - 0.2)(1 - 0.3)(1 - 0.4) = 0.336$.

(0, 0, 1): Vain herra C osuu hirveen. Todennäköisyys on $(1 - 0.2)(1 - 0.3)(0.4) = 0.224$.

(0, 1, 0): Vain herra B osuu hirveen. Todennäköisyys on $(1 - 0.2)(0.3)(1 - 0.4) = 0.144$.

(0, 1, 1): Herrat B ja C osuvat hirveen. Todennäköisyys on $(1 - 0.2)(0.3)(0.4) = 0.096$.

(1, 0, 0): Vain herra A osuu hirveen. Todennäköisyys on $(0.2)(1 - 0.3)(1 - 0.4) = 0.084$.

(1, 0, 1): Herrat A ja C osuvat hirveen. Todennäköisyys on $(0.2)(1 - 0.3)(0.4) = 0.056$.

(1, 1, 0): Herrat A ja B osuvat hirveen. Todennäköisyys on $(0.2)(0.3)(1 - 0.4) = 0.036$.

(1, 1, 1): Kaikki herrat osuvat hirveen. Todennäköisyys on $(0.2)(0.3)(0.4) = 0.024$.

Nyt todennäköisyyksien summa on 1.

(b) Satunnaisen muuttujan X mahdolliset arvot ja niiden todennäköisyydet ovat seuraavat:

$X = 0$: Ei yhtään osumaa. Todennäköisyys on 0.336.

$X = 1$: Yksi osuma. Todennäköisyys on $0.084 + 0.144 + 0.224 = 0.452$.

$X = 2$: Kaksi osumaa. Todennäköisyys on $0.096 + 0.056 + 0.036 = 0.188$.

$X = 3$: Kolme osumaa. Todennäköisyys on 0.024.

(c) Jos hirveen osuu tasan yksi luoti, haluamme tietää todennäköisyyden, että se on herran C ampuma. Tätä varten voimme käyttää ehdollista todennäköisyyttä:

$\mathbb{P}(C = 1|X = 1) = \frac{\mathbb{P}(C=1 \cap X=1)}{\mathbb{P}(X=1)}$

Koska $X = 1$ ja $C = 1$, niin $A = 0$ ja $B = 0$. Tällöin:

$\mathbb{P}(C = 1|X = 1) = \frac{\mathbb{P}(A=0,B=0,C=1)}{\mathbb{P}(X=1)} = \frac{0.224}{0.452} \approx 0.496$

Vastaukset:

(a) Satunnaisvektorin (A, B, C) mahdolliset arvot ja niiden todennäköisyydet ovat esitetty yllä.

(b) Satunnaisen muuttujan X mahdolliset arvot ja niiden todennäköisyydet ovat esitetty yllä.

(c) Jos hirveen osuu tasan yksi luoti, todennäköisyys, että se on herran C ampuma, on noin 0.496.

Problem 2

Eräällä satunnaismuuttujalla X on jatkuva jakauma tiheysfunktiolla f, missä

$$f(x) = \begin{cases} x/2 - 5 & \text{jos } 10 \leq x \leq 12 \\ 0 & \text{muuten} \end{cases}$$

(a) Määritä X : n kertymäfunktio.
(b) Laske X : n odotusarvo.
(c) Laske X : n moodi (tiheysfunktion maksimikohta).
(d) Laske X :n mediaani, ts. sellainen luku m, että $\mathbb{P}(X \leq m) = 1/2$.
(e) Laske satunnaismuuttujan $Y = 12 - X$ odotusarvo.
(f) Laske $\mathbb{P}\left(X^2 < 121\right)$.
Kohdissa (b)-(e) ilmoita vastausten lukuarvot kolmella desimaalilla.

Problem 2 solution

(a) Määritetään X:n kertymäfunktio $F(x)$. Koska $f(x)$ on jatkuva tiheysfunktio, voimme integroida sen saadaksemme kertymäfunktion:

$$F(x) = \int f(x)dx = \begin{cases} 0 & \text{jos } x < 10 \\ \int_{10}^{x}(t/2 - 5)dt & \text{jos } 10 \leq x \leq 12 \\ 1 & \text{jos } x > 12 \end{cases}$$

Integroimalla saamme:

$$F(x) = \begin{cases} 0 & \text{jos } x < 10 \\ \frac{t^2}{4} - 5t\Big|_{10}^{x} & \text{jos } 10 \leq x \leq 12 \\ 1 & \text{jos } x > 12 \end{cases}$$

$$F(x) = \begin{cases} 0 & \text{jos } x < 10 \\ \frac{x^2}{4} - 5x - \left(\frac{100}{4} - 50\right) & \text{jos } 10 \leq x \leq 12 \\ 1 & \text{jos } x > 12 \end{cases}$$

$$F(x) = \begin{cases} 0 & \text{jos } x < 10 \\ \frac{x^2}{4} - 5x + 25 & \text{jos } 10 \leq x \leq 12 \\ 1 & \text{jos } x > 12 \end{cases}$$

(b) Laske X:n odotusarvo.

Odotusarvo on integraali $E[X] = \int_{-\infty}^{\infty} xf(x)dx$. Koska $f(x) = 0$ muualla kuin välillä $[10, 12]$, voimme laskea integraalin vain tältä väliltä:

$$E[X] = \int_{10}^{12} x(\frac{x}{2} - 5)dx = \int_{10}^{12} (\frac{x^2}{2} - 5x)dx$$

$$E[X] = \left[\frac{x^3}{6} - \frac{5x^2}{2}\right]_{10}^{12} = \left(\frac{12^3}{6} - \frac{5 \cdot 12^2}{2}\right) - \left(\frac{10^3}{6} - \frac{5 \cdot 10^2}{2}\right) = \frac{34}{3}$$

(c) Laske X:n moodi eli tiheysfunktion maksimikohta.

Tämä voidaan löytää derivoimalla tiheysfunktio ja etsimällä sen nollakohdat:

$$f'(x) = \frac{d}{dx}(\frac{x^2}{2} - 5x) = x - 5$$

Asetetaan $f'(x) = 0$:

$$x - 5 = 0 \Rightarrow x = 5$$

Koska $x = 5$ ei ole tiheysfunktion määrittelyjoukossa $[10, 12]$, moodi on joko 10 tai 12. Vertaamalla tiheysfunktion arvoja näissä pisteissä:

$$f(10) = \frac{10}{2} - 5 = 0$$
$$f(12) = \frac{12}{2} - 5 = 1$$

Moodi on $x = 12$, koska siinä tiheysfunktio saavuttaa suurimman arvon.

.

(d) Laske X:n mediaani eli sellainen luku m, että $\mathbb{P}(X \leq m) = \frac{1}{2}$.

Käytetään kertymäfunktiota:

$$F(m) = \frac{1}{2} \Rightarrow \frac{m^2}{4} - 5m + 25 = \frac{1}{2}$$

$$m^2 - 20m + 100 = 2$$

$$m^2 - 20m + 98 = 0$$

Ratkaisemalla toisen asteen yhtälön saamme $m = 10 - \sqrt{2}$ tai $m = 10 + \sqrt{2}$. Koska X on määritelty välillä $[10, 12]$, mediaani on $m = 10 + \sqrt{2}$.

(e) Laske satunnaismuuttujan $Y = 12 - X$ odotusarvo.

Käytetään odotusarvon ominaisuuksia:

$$E[Y] = E[12 - X] = 12 - E[X] = 12 - \frac{34}{3} = \frac{2}{3}$$

(f) Laske $\mathbb{P}(X^2 < 121)$.

Tämä voidaan ilmaista kertymäfunktion avulla:

$$\mathbb{P}(X^2 < 121) = \mathbb{P}(-11 < X < 11)$$

Koska X on määritelty välillä $[10, 12]$, voimme rajoittaa tapauksen:

$$\mathbb{P}(10 < X < 11) = F(11) - F(10) = \left(\frac{11^2}{4} - 5 \cdot 11 + 25\right) - \left(\frac{10^2}{4} - 5 \cdot 10 + 25\right) = \frac{1}{4}$$

Vastaukset:

(a) Kertymäfunktio $F(x)$ on esitetty aiemmassa vastauksessa.

(b) X:n odotusarvo on $\frac{34}{3}$.

(c) X:n moodi on 12.

(d) X:n mediaani on $10 + \sqrt{2}$.

(e) Satunnaismuuttujan $Y = 12 - X$ odotusarvo on $\frac{2}{3}$.

(f) $\mathbb{P}(X^2 < 121) = \frac{1}{4}$.

Problem 3

Luennoijalla on neljä 6-sivuista noppaa, joista kolme on tavallisia (tulokset $1, \ldots, 6$ ovat yhtä todennäköiset) ja neljäs on painotettu siten, että tuloksen i todennäköisyys on $i/21$, kun $i = 1, \ldots, 6$. Luennoija on poiminut yhden nopista satunnaisesti ja heittänyt sitä seitsemän kertaa saaden tulosjonon $(6, 6, 2, 5, 6, 6, 5)$. Merkitään $\Theta = 1$ jos poimittu noppa on tavallinen ja $\Theta = 2$ jos se on painotettu.
(a) Laske todennäköisyys saada juuri tämä tulosjono, jos noppa oli tavallinen.

(b) Laske todennäköisyys saada juuri tämä tulosjono, jos noppa oli painotettu.

(c) Määritä parametrin Θ posteriorijakauma.

(d) Käyttäen saatua posteriorijakaumaa, määritä todennäköisyys saada kolme kuutosta, jos poimittua noppaa heitetään vielä kolme kertaa.

Problem 3 solution

(a)
Laske todennäköisyys saada juuri tämä tulosjono, jos noppa oli tavallinen. Tavallisessa nopassa jokaisen tuloksen todennäköisyys on $\frac{1}{6}$. Koska heitot ovat riippumattomia, voimme laskea todennäköisyyden kertomalla yksittäisten heittojen todennäköisyydet:

$$P((6,6,2,5,6,6,5)|\Theta = 1) = \left(\frac{1}{6}\right)^7 \approx 0.00000357$$

(b) Laske todennäköisyys saada juuri tämä tulosjono, jos noppa oli painotettu. Painotetussa nopassa tuloksen i todennäköisyys on $\frac{i}{21}$. Lasketaan todennäköisyys samalla tavalla kuin kohdassa (a):

$$P((6,6,2,5,6,6,5)|\Theta = 2) = \frac{6}{21} \cdot \frac{6}{21} \cdot \frac{2}{21} \cdot \frac{5}{21} \cdot \frac{6}{21} \cdot \frac{6}{21} \cdot \frac{5}{21} \approx 0.00003598$$

Oletetaan, että ennen heittoja todennäköisyys valita tavallinen noppa oli $\frac{3}{4}$ ja painotettu noppa $\frac{1}{4}$. Lasketaan posteriorijakauma molemmille arvoille:

$$P(\Theta = 1|(6,6,2,5,6,6,5)) = \frac{P((6,6,2,5,6,6,5)|\Theta = 1)P(\Theta = 1)}{P((6,6,2,5,6,6,5))}$$

$$P(\Theta = 2|(6,6,2,5,6,6,5)) = \frac{P((6,6,2,5,6,6,5)|\Theta = 2)P(\Theta = 2)}{P((6,6,2,5,6,6,5))}$$

$$P((6,6,2,5,6,6,5)|\Theta = 1)P(\Theta = 1) \approx 0.00000357 \cdot \frac{3}{4} \approx 0.0000026775$$

$$P((6,6,2,5,6,6,5)|\Theta = 2)P(\Theta = 2) \approx 0.00003598 \cdot \frac{1}{4} \approx 0.000008995$$

$$P(\Theta = 1|(6,6,2,5,6,6,5)) = \frac{0.0000026775}{0.0000026775 + 0.000008995} \approx 0.229$$

$$P(\Theta = 2|(6,6,2,5,6,6,5)) = \frac{0.000008995}{0.0000026775 + 0.000008995} \approx 0.771$$

Posteriorijakauma on $P(\Theta = 1|(6,6,2,5,6,6,5)) \approx 0.229$ ja $P(\Theta = 2|(6,6,2,5,6,6,5)) \approx 0.771$.

Problem 4

Eräästä datalähteestä saadaan riippumattomia satunnaislukuja, joista kukin saa arvon 1 todennäköisyydellä θ, ja arvon 0 todennäköisyydellä $1-\theta$. Parametrin $\theta \in [0,1]$ arvo on tuntematon. Lähteestä on saatu havainnot $x_1 = 0, x_2 = 0, x_3 = 0$ ja $x_4 = 1$. Auta Ronaldia, Karlia ja Thomasia estimoimaan $\theta :$ n arvoa havaintojen perusteella. Kaikissa kohdissa vastaus on perusteltava laskutoimituksin (vetoaminen jakauman tunnistamiseen ja sen tunnettuihin ominaisuuksiin ei riitä).

(a) Ronald käyttää suurimman uskottavuuden estimaattia. Määritä parametrin uskottavuusfunktio havaintojen perusteella ja sen perusteella parametrin su-estimaatti.

(b) Karl käyttää momenttiestimaattia, eli hän etsii sellaisen $\theta :$ n arvon, että datalähteen odotusarvo vastaa havaintojen keskiarvoa. Määritä Karlin estimaatti.

(c) Thomas pitää parametria satunnaismuuttujana Θ, jonka priorijakauman tiheys on $f_\Theta(\theta) = \theta(1-\theta)$, kun $0 \leq \theta \leq 1$, ja nolla muualla. Määritä parametrin posteriorijakauman tiheysfunktio (voit jättää normalisointivakion ratkaisematta) sekä sen perusteella parametrin posteriorimoodi eli $\theta :$ n arvo, jossa posterioritiheys on suurimmillaan.

Problem 4 solution

(a)

Ronald käyttää suurimman uskottavuuden estimaattia. Määritetään parametrin uskottavuusfunktio havaintojen perusteella ja sen perusteella parametrin su-estimaatti.

Uskottavuusfunktio on:

$$L(\theta) = P(x_1 = 0, x_2 = 0, x_3 = 0, x_4 = 1|\theta) = (1-\theta)^3\theta$$

Suurimman uskottavuuden estimaatti saadaan maksimoimalla uskottavuusfunktio. Derivoimalla uskottavuusfunktio ja etsimällä sen nollakohdat:

$$\frac{dL(\theta)}{d\theta} = -3(1-\theta)^2\theta + (1-\theta)^3 = 0$$

Tämä yhtälö ei ratkea analyyttisesti, mutta voimme tarkastella sen käyttäytymistä. Huomaa, että $\frac{dL(\theta)}{d\theta} > 0$, kun $0 < \theta < \frac{1}{2}$ ja $\frac{dL(\theta)}{d\theta} < 0$, kun $\frac{1}{2} < \theta < 1$. Tämä tarkoittaa, että uskottavuusfunktio saavuttaa maksiminsa, kun $\theta = \frac{1}{2}$. Suurimman uskottavuuden estimaatti on siis $\hat{\theta} = \frac{1}{2}$.

(b)

Karl käyttää momenttiestimaattia, eli hän etsii sellaisen θ:n arvon, että datalähteen odotusarvo vastaa havaintojen keskiarvoa. Määritetään Karlin estimaatti.

Datalähteen odotusarvo on $E[X] = \theta$. Havaintojen keskiarvo on $\bar{x} = \frac{0+0+0+1}{4} = \frac{1}{4}$. Momenttiestimaatti saadaan asettamalla odotusarvo yhtä suureksi kuin havaintojen keskiarvo:

$$\hat{\theta} = \bar{x} = \frac{1}{4}$$

(c)

Thomas pitää parametria satunnaismuuttujana Θ, jonka priorijakauman tiheys on $f_\Theta(\theta) = \theta(1-\theta)$, kun $0 \leq \theta \leq 1$, ja nolla muualla. Määritetään parametrin posteriorijakauman tiheysfunktio (voit jättää normalisointivakion ratkaisematta) sekä sen perusteella parametrin posteriorimoodi eli θ:n arvo, jossa posterioritiheys on suurimmillaan.

Posteriorijakauman tiheysfunktio saadaan Bayesin kaavalla:

$$f_{\Theta|X}(\theta|x) \propto L(\theta) f_\Theta(\theta) = (1-\theta)^3\theta \cdot \theta(1-\theta) = \theta^2(1-\theta)^4$$

Maksimoimalla posterioritiheys löydämme posteriorimoodin. Derivoimalla ja etsimällä nollakohdat:

$$\frac{df_{\Theta|X}(\theta|x)}{d\theta} = 2\theta(1-\theta)^4 - 4\theta^2(1-\theta)^3 = 0$$

$$\theta(1-\theta)^3(2(1-\theta) - 4\theta) = 0$$

Tästä saamme kolme ratkaisua: $\theta = 0$, $\theta = 1$, ja $\theta = \frac{1}{3}$. Koska $\theta = 0$ ja $\theta = 1$ ovat ääripäät, posteriorimoodi on $\theta = \frac{1}{3}$.