

Ohje: Vastaa lyhyesti ja ytimekkäästi, mutta perustele ratkaisusi. Pelkkä lukuarvo vastauksena ei anna pisteitä. Kokeessa on 4 tehtävää, jokaisesta saa 0–6 pistettä. Merkitse jokaiseen vastauspaperiin:

- Kurssin nimi ja koodi
- SUKUNIMI ja ETUNIMET (tikkukirjaimin)
- Opiskelijanumero
- Koulutusohjelma ja vuosikurssi
- Päivämäärä ja nimikirjoitus

Sallitut apuvälineet: Mellinin tilastolliset taulukot, laskin ja a4-muistilappu (käsin kirjoitettu, tekstiä vain toisella puolella, oikeassa yläkulmassa oma nimi, ei tarvitse palauttaa)

T1 Tavallista 6-sivuista noppaa heitetään monta kertaa peräkkäin ja heittojen tuloksia merkitään X_1, X_2, \dots . Lisäksi merkitään $M_2 = \frac{1}{2}(X_1 + X_2)$ ja $M_{100} = \frac{1}{100}(X_1 + X_2 + \dots + X_{100})$. Määritä

- (a) satunnaismuuttujan X_1 odotusarvo ja keskihajonta, (2p)
- (b) satunnaismuuttujien X_1 ja M_2 yhteisjakauma, (1p)
- (c) satunnaismuuttujan M_2 jakauma, (1p)
- (d) satunnaismuuttujan M_{100} odotusarvo ja keskihajonta. (2p)

T2 Tenniksen kilpailutauon aikana Henri kokeili uutta harjoittelumuotoa. Sen tavoitteena oli parantaa ykkösyötön onnistumisprosenttia θ , joka viime kaudella oli 0.50. Uuden harjoittelumuodon vaikutusta testattiin tutkimalla 6 ykkösyöttöä uuden kilpailukauden alussa. Päätettiin testata nollahypoteesia $H_0 : \theta = 0.50$ suhteessa vastahypoteesiin $H_1 : \theta \neq 0.50$ valitsemalla testisuureksi $X = \text{”onnistuneiden ykkösyöttöjen lkm”}$. Testissä havaittiin tulos $x = 5$.

- (a) Mikä on onnistuneiden ykkösyöttöjen odotusarvo nollahypoteesin vallitessa? (1p)
- (b) Mikä on tapahtuman $X = 5$ todennäköisyys nollahypoteesin vallitessa? (1p)
- (c) Määritä havaitun testituloksen p-arvo ja tee sen pohjalta johtopäätös nollahypoteesin hylkäämisestä tai hyväksymisestä. (2p)
- (d) Määritä 5% merkitsevyystasoa vastaava testin hylkäysalue. (2p)

T3 Skeptiikan tentissä on kuusi kysymystä, joihin vastataan valitsemalla toinen kahdesta vastausvaihtoehdosta. Jokaisesta oikeasta vastauksesta saa yhden pisteen. Tenttiin osallistuu ahkeria ja laiskoja opiskelijoita. Ahkera opiskelija on valmistautunut tenttiin huolellisesti ja vastaa kuhunkin tehtävään oikein todennäköisyydellä 0.9, kun taas laiska opiskelija vastaa jokaiseen kysymykseen umpimähkään tasaisen satunnaisesti. Kurssipalautteen perusteella tentin järjestäjä arvioi, että ahkerien opiskelijoiden osuus tenttiin osallistujista on $2/3$.

- (a) Millä todennäköisyydellä ahkera opiskelija saa tentistä tasan 3 pistettä? **(2p)**
- (b) Millä todennäköisyydellä satunnainen opiskelija saa tentistä tasan 3 pistettä? **(2p)**
- (c) Millä todennäköisyydellä tentistä tasan 3 pistettä saanut opiskelija on laiska? **(2p)**

T4 Tuntematon datalähde tuottaa stokastisesti riippumattomia $\{0, 1\}$ -arvoisia satunnaislukuja tiheysfunktiona

$$f(x_i | \theta) = \begin{cases} 1 - \theta, & x_i = 0, \\ \theta, & x_i = 1, \end{cases}$$

missä $\theta \in (0, 1)$ on tuntematon parametri. Datalähteestä on havaittu luvut $x_1 = 1, x_2 = 0, x_3 = 1, x_4 = 1, x_5 = 1$. Auta Ronaldia, Karlia ja Thomasia estimoimaan parametrin θ arvo näiden havaintojen pohjalta.

- (a) Ronald päättää käyttää suurimman uskottavuuden estimaattia. Määritä parametrin uskottavuusfunktio ja laske sen avulla suurimman uskottavuuden estimaatti datajoukolle $x = (x_1, \dots, x_5)$. **(2p)**
- (b) Karlin käyttämän estimaatin arvo määräytyy ratkaisemalla θ yhtälöstä

$$E(X | \theta) = \frac{1}{n} \sum_{i=1}^n x_i,$$

jonka vasemmalla puolella on jakaumaa $f(x_i | \theta)$ noudattavan satunnaismuuttujan odotusarvo ja oikealla puolella havaintojen keskiarvo. Laske tämän estimaatin arvo datajoukolle x . **(2p)**

- (c) Thomas tulkitsee tuntemattoman parametrin satunnaismuuttujaksi ja valitsee priorijakaumaksi yksikkövälin tasajakauman tiheysfunktiona

$$p(\theta) = \begin{cases} 1, & \theta \in (0, 1), \\ 0, & \text{muuten.} \end{cases}$$

Thomas päättää käyttää estimaattina posteriorijakauman odotusarvoa. Laske tämän estimaatin arvo datajoukolle x . **(2p)**