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)

T1 Neljältä Mac-käyttäjältä ja kahdelta PC-käyttäjältä kysyttiin, kuinka monta kertaa he sammuttivat tietokoneensa viime viikolla. Mac-käyttäjät vastasivat: 1, 2, 2 ja 6 kertaa. PC-käyttäjät vastasivat: 2 ja 5 kertaa. Kyselyyn osallistuineiden opiskelijoiden keskuudesta valitaan satunnaisesti yksi, ja merkitään

$$
\begin{aligned}
T &= \text{satunnaisesti valitun opiskelijan tietokoneen tyyppi (0=Mac, 1=PC)},\\
X &= \text{satunnaisesti valitun opiskelijan tietokoneen sammutuskertojen lukumäärä.}
\end{aligned}
$$

(a) Kirjoita satunnaismuuttujien T ja X yhteisjakaumaa kuvaava taulukko. **(2p)**

(b) Määritä T:n ja X:n odotusarvot. **(1p)**

(c) Ovatko T ja X riippuvat vai riippumattomat? Perustele. **(1p)**

Valitaan sitten toisistaan riippumattomasti yksi satunnainen Mac-käyttäjä ja yksi satunnainen PC-käyttäjä, ja merkitään heidän tietokoneidensa sammutuskertojen lukumääriä X_0 ja X_1.

(d) Laske todennäköisyys $\mathrm{P}(X_0 > X_1)$. **(1p)**

(e) Määritä X_0:n ja X_1:n korrelaatio. **(1p)**

T2 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)**

T3 Datalähteen tiedetään tuottavan riippumattomia normaalijakautuneita satunnaislukuja odotusarvona μ (tuntematon) ja keskihajontana $\sigma = 2$. Datalähteestä on kerätty $n = 100$ havaintoa. Testaa 5% merkitsevyystasolla nollahypoteesia $H_0 : \mu = 3$ suhteessa vastahypoteesiin $H_1 : \mu \neq 3$. Käytä testisuuretta

$$t(x) = \frac{m(x) - 3}{\sigma/\sqrt{n}},$$

missä $m(x) = \frac{1}{n}\sum_{i=1}^n x_i$ on havaitun datajoukon keskiarvo.

(a) Määritä testisuureen p-arvo datajoukolle, jonka keskiarvo on 3.6. **(2p)**

(b) Mikä on testin johtopäätös (a)-kohdan havainnolle? **(1p)**

(c) Määritä nollahypoteesin hylkäämiseen johtavien testisuureen arvojen joukko. **(1p)**

(d) Määritä testin hyväksymisvirheen todennäköisyys, kun tuntemattoman parametrin todellisen arvon tiedetään olevan $\mu = 3.2$. **(2p)**

T4 Eri matkustuspäivien odotusajat (min) eräällä bussipysäkillä ovat toisistaan riippumattomia ja noudattavat jatkuvan välin $[0, \theta]$ tasajakaumaa tiheysfunktiona

$$f(t \mid \theta) = \begin{cases} \frac{1}{\theta}, & 0 < t < \theta, \\ 0, & \text{muuten.} \end{cases}$$

Viiden matkustuspäivän aikana on havaittu odotusajat $x = (1, 6, 3, 4, 8)$. Auta Ronaldia, Karlia ja Thomasia estimoimaan parametrin θ arvo näiden havaintojen pohjalta.

(a) Ronald päättää käyttää suurimman uskottavuuden estimaattia. Laske tämän estimaatin arvo datajoukolle x. **(2p)**

(b) Karlin käyttämän estimaatin arvo määräytyy ratkaisemalla θ yhtälöstä

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

jossa vasemmalla on jakaumaa $f(t \mid \theta)$ noudattavan satunnaismuuttujan odotusarvo ja oikealla havaintojen keskiarvo. Laske tämän estimaatin arvo datajoukolle x. **(1p)**

(c) Selitä mikä on harhaton estimaattori. Onko Karlin estimaattori harhaton? **(1p)**

(d) Thomas tulkitsee tuntemattoman parametrin satunnaismuuttujaksi Θ ja valitsee priorijakauman tiheysfunktioksi

$$f_0(\theta) = \begin{cases} 2\theta^{-3}, & \theta \geq 1, \\ 0, & \text{muuten.} \end{cases}$$

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