Todennäköisyyslaskennan ja tilastotieteen peruskurssi — Tentti
Matematiikan ja systeemianalyysin laitos — 15.2.2017
Aalto-yliopisto — L Leskelä

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 Uusi doping-testi paljastaa 95% erään kielletyn hormonivalmisteen käyttäjistä, mutta tuottaa (virheellisen) positiivisen testituloksen myös 2%:lle urheilijoista, jotka eivät käytä kyseistä valmistetta. Hiihtomaajoukkueesta, jonka urheilijoista 1% käyttää kyseistä valmistetta, poimitaan satunnaisotannalla yksi urheilija testattavaksi. Määritä todennäköisyydet tapahtumille:

(a) "Testitulos on positiivinen", **(2p)**

(b) "Positiivisen testituloksen saanut urheilija ei käytä kyseistä hormonivalmistetta", **(2p)**

(c) "Negatiivisen testituloksen saanut urheilija ei käytä kyseistä hormonivalmistetta". **(2p)**

T2 Tenniksen kilpailutauon aikana Henri kokeili uutta harjoittelumuotoa. Sen tavoitteena oli parantaa ykkössyötön onnistumisprosenttia θ, joka viime kaudella oli 0.50. Uuden harjoittelumuodon vaikutusta testattiin tutkimalla 6 ykkössyöttöä uuden kilpailukauden alussa. Päätettiin testata nollahypoteesia $H_0 : \theta = 0.50$ suhteessa vastahypoteesiin $H_1 : \theta \neq 0.50$ valitsemalla testisuureeksi X = "onnistuneiden ykkössyöttöjen lkm". Testissä havaittiin tulos $x = 5$.

(a) Mikä on onnistuneiden ykkössyöttöjen 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 Tieverkon sulana pitämiseen on varastoitu suolaa 200 cm lumimäärän varalle. Yksittäisen talvipäivän aikana lunta sataa keskimäärin 4.5 cm, keskihajonnan ollessa 2.5 cm. Tulevien talvipäivien lumimäärät oletetaan toisistaan riippumattomiksi.

(a) Laske 50 talvipäivän aikana sataneen lumimäärän odotusarvo ja keskihajonta. **(2p)**

(b) Onko yo. tietojen pohjalta mahdollista laskea, millä todennäköisyydellä yksittäisenä talvipäivänä sataa lunta yli 12 cm? Jos on, ilmoita vastaus. Jos ei, selitä miksi ei. **(1p)**

(c) Laske normaaliapproksimaation avulla arvio todennäköisyydelle, että suola riittää 50 talvipäiväksi. **(1p)**

(d) Arvioi normaaliapproksimaatiota käyttäen, kuinka paljon enemmän suolaa pitäisi varata, jotta sitä riittäisi 50 talvipäiväksi vähintään 99% todennäköisyydellä? **(1p)**

(e) Mitkä ylläolevista laskelmista pitävät paikkansa myös silloin, kun talvipäivien riippumattomuusoletuksesta luovutaan? **(1p)**

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

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

Neljänä eri matkustuspäivänä on havaittu odotusajat $x_1 = 1$, $x_2 = 6$, $x_3 = 5$, $x_4 = 4$. Auta Ronaldia, Karlia ja Thomasia estimoimaan parametrin θ arvo näiden havaintojen pohjalta.

(a) Ronald päättää käyttää suurimman uskottavuuden estimaattia. Laske tämän estimaatin arvo datajoukolle $x = (x_1, \ldots, x_4)$. **(2p)**

(b) Karlin käyttämän estimaatin arvo määräytyy ratkaisemalla θ yhtälöstä

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

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

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

$$f_0(\theta) \;=\; \begin{cases} 3\theta^{-4}, & \theta \ge 1, \\ 0, & \text{muuten.} \end{cases}$$

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