Todennäköisyyslaskennan ja tilastotieteen peruskurssi
Matematiikan ja systeemianalyysin laitos
Aalto-yliopisto

Tentti
03.04.2018
J. Virta/C. Webb

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: 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 – ajatellaan, että ykkössyötön onnistuminen on riippumatonta muiden ykkössyöttöjen onnistumisesta. 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 Satunnaismuuttuja X saa arvoja joukosta $\{-1, 0, 1\}$ ja satunnaismuuttuja Y saa arvoja joukosta $\{0, 1\}$. Näiden yhteisjakauma on annettu allaolevassa taulukossa.

	Y	
X	0	1
-1	$\frac{1}{3}$	0
0	0	$\frac{1}{3}$
1	$\frac{1}{3}$	0

Toisin sanoen, $\mathbb{P}(Y = 0 \text{ ja } X = \pm 1) = \frac{1}{3} = \mathbb{P}(Y = 1 \text{ ja } X = 0)$ ja $\mathbb{P}(Y = 1 \text{ ja } X = \pm 1) = 0 = \mathbb{P}(Y = 0 \text{ ja } X = 0)$.

(a) Laske yhteisjakauman reunajakaumien tiheysfunktiot $f_X(x)$ ja $f_Y(y)$. **(1p)**

(b) Laske satunnaismuuttujien X ja Y odotusarvot $\mathbb{E}(X)$ ja $\mathbb{E}(Y)$. **(1p)**

(c) Laske satunnaismuuttujien X ja Y keskihajonnat $\mathrm{SD}(X)$ ja $\mathrm{SD}(Y)$. **(1p)**

(d) Laske satunnaismuuttujien X ja Y kovarianssi $\mathrm{Cov}(X, Y)$. **(1p)**

(e) Laske satunnaismuuttujien X ja Y korrelaatio $\mathrm{Cor}(X, Y)$. **(1p)**

(f) Ovatko satunnaismuuttujat X ja Y riippumattomat? Perustele vastauksesi. **(1p)**

T4 Kokeellinen fyysikko uskoo löytäneensä uuden alkeishiukkasen. Hänellä on laite, joka tuottaa näitä hiukkasia ja antaa mitata tuotetun hiukkasen elinajan. Fyysikko olettaa, että tuotettujen hiukkasten elinajat ovat satunnaisia, toisistaan riippumattomia ja noudattavat eksponenttijakaumaa, jonka tiheysfunktio on $f(t|\lambda) = \lambda e^{-\lambda t}$, $t > 0$ jollakin tuntemattomalla parametrilla $\lambda > 0$. Hän suorittaa mittauksen viisi kertaa ja havaitsee elinajat $t_1 = 3.5$ s, $t_2 = 2.7$ s, $t_3 = 8.5$ s, $t_4 = 4.2$ s, ja $t_5 = 4.7$ s.

(a) Auta kokeellista fyysikkoa muodostamaan parametrin λ uskottavuusfunktio havaitun datajoukon $t_1, ..., t_5$ suhteen. **(1p)**

(b) Auta kokeellista fyysikkoa muodostamaan parametrin λ suurimman uskottavuuden estimaatti havaitun datajoukon $t_1, ..., t_5$ suhteen. **(2p)**

(c) Paikalle saapuu tunnettu teorettiinen fyysikko Huu Haa ja julistaa: "Tämä on selvästi ennustamani Ω-hiukkanen, jonka elinajan uskon olevan noin 4 s". Huu Haa haluaisi laskea Bayes-estimaatin parametrin λ arvolle. Hänen mielestään uskomusta neljän sekunnin elinajasta kuvaisi hyvin priorijakauma, jonka tiheysfunktio on $p_0(\lambda) = 4e^{-4\lambda}$, $\lambda > 0$.

Auta Huu Haata muodostamaan Bayes-estimaatti parametrin λ arvolle laskemalla piste, jossa posteriorijakauman tiheysfunktio saa suurimman arvonsa, kun posteriorijakauma laskataan käyttämällä priorijakaumaa p_0, datapisteitä $t_1, \ldots, t_5$ ja samaa stokastista mallia kuin a-kohdassa.

(**Vihje:** Sinun ei tarvitse laskea posteriorijakauman normitusvakiota vaan pystyt ratkaista tehtävän pelkän normittamattoman posteriorijakauman avulla.) **(3p)**