Todennäköisyyslaskennan ja tilastotieteen peruskurssi
Matematiikan ja systeemianalyysin laitos
Aalto-yliopisto

Tentti
10.12.2018
J. Virta

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 Heitetään kahta kuusisivuista noppaa. Olkoon satunnaismuuttuja X_1 ensimmäisen nopan silmäluku ja X_2 toisen nopan silmäluku (voidaan olettaa, että X_1 ja X_2 ovat riippumattomia). Merkitään lisäksi $N = \min(X_1, X_2)$ (silmäluvuista pienempi) ja $M = \max(X_1, X_2)$ (silmäluvuista suurempi).

(a) Etsi satunnaismuuttujan N jakauma ja odotusarvo **(2p)**

(b) Etsi satunnaismuuttujien N ja M yhteisjakauma. Ovatko satunnaismuuttujat N ja M riippumattomat (perustele vastauksesi)? **(3p)**

(c) Etsi satunnaismuuttujan N ehdollinen jakauma ehdolla $M = 4$. **(1p)**

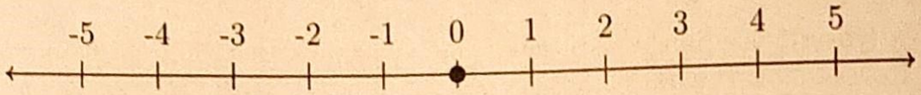
T2 Satunnaiskulkija lähtee liikkeelle origosta ja liikkuu sekunnin välein joko yhden yksikön vasemmalle, yhden yksikön oikealle tai pysyy paikallaan (ks. kuva seuraavalla sivulla). Kaikki kolme vaihtoehtoa ovat yhtä todennäköiset, ja se mitä kulkija tekee kullakin ajanhetkellä (liikkuu vasemmalle/liikkuu oikealle/pysyy paikallaan) on riippumaton kaikesta sen aikaisemmasta toiminnasta.

Olkoon X_i satunnaismuuttuja, joka saa arvon -1, jos kulkija ottaa i:nnen sekunnin aikana askeleen vasemmalle, arvon 0, jos kulkija pysyy i:nnen sekunnin aikana paikallaan ja arvon 1, jos kulkija ottaa i:nnen sekunnin aikana askeleen oikealle. Tällöin summa $S_n = \sum_{i=1}^{n} X_i$ kertoo satunnaiskulkijan sijainnin n sekunnin jälkeen.

(a) Määritä satunnaismuuttujan X_1 odotusarvo ja keskihajonta. **(2p)**

(b) Määritä satunnaismuuttujan S_{100} odotusarvo ja keskihajonta. **(2p)**

(c) Arvioi normaaliapproksimaatiota käyttäen likimääräinen todennäköisyys sille, että satunnaiskulkijan etäisyys origosta on 100 sekunnin jälkeen yli 10 yksikköä. **(2p)**



T3 Signaalinkäsittelijät ovat kehittäneet kohinaisen tiedonsiirtokanavan, jonka viesteihin aiheuttamien virheiden tiedetään olevan normaalijakautuneita satunnaislukuja odotusarvona μ (tuntematon) ja keskihajontana $\sigma = 3$ (voidaan olettaa, että kanavan eri viesteihin aiheuttamat virheet ovat riippumattomia). He lähettivät kanavan läpi $n = 100$ viestiä ja mittasivat näihin aiheutuneiden virheiden suuruudet $x_1, \ldots, x_n$.

Auta signaalinkäsittelijöitä testaamaan 5% merkitsevyystasolla nollahypoteesia $H_0 : \mu = 0$ suhteessa vastahypoteesiin $H_1 : \mu \neq 0$. Käytä testisuuretta

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

missä $m(x) = \frac{1}{n}\sum_{i=1}^{n} x_i$ on havaitun datajoukon keskiarvo. *Huom. riippumattomien normaalijakautuneiden satunnaismuuttujien keskiarvo on myös normaalijakautunut.*

(a) Määritä testisuureen p-arvo datajoukolle, jonka keskiarvo on $m(x) = 0.9$. **(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 $\mu = 0.3$. **(2p)**

T4 Tietyn lajin perhosilla on joko keltaiset, oranssit tai punaiset siivet. Aloitteleva entomologi Aino ei tiedä eri värien osuuksia, mutta on päätellyt, että kelta- ja punasiipisiä perhosia on kumpiakin koko populaatiosta osuus θ ja oranssisiipisiä osuus $1 - 2\theta$, missä $0 < \theta < 1/2$. Selvittääkseen tuntemattoman parametrin θ arvon, Aino kerää viisi kyseisen lajin perhosta ja havaitsee niiden siipien väreiksi: *punainen, keltainen, punainen, oranssi, keltainen.*

Olkoon X_i satunnaismuuttuja, joka saa arvon 0, jos i:nnes perhonen on keltasiipinen, arvon 1, jos i:nnes perhonen on oranssisiipinen ja arvon 2, jos i:nnes perhonen on punasiipinen. Tällöin satunnaismuuttujan X_i tiheysfunktio on,

$$f(x_i\,|\,\theta) = \begin{cases} \theta, & x_i = 0, \\ 1 - 2\theta, & x_i = 1, \\ \theta, & x_i = 2, \end{cases}$$

ja Aino havaitsee siis arvot $x_1 = 2$, $x_2 = 0$, $x_3 = 2$, $x_4 = 1$, $x_5 = 0$. Perhospopulaation suuren koon vuoksi voidaan olettaa, että satunnaismuuttujat $X_1, \ldots, X_5$ ovat riippumattomia toisistaan.

(a) Auta Ainoa muodostamaan parametrin θ uskottavuusfunktio hänen havaitsemansa datajoukon $x_1, \ldots, x_5$ suhteen. **(1p)**

(b) Auta Ainoa muodostamaan parametrin θ suurimman uskottavuuden estimaatti hänen havaitsemansa datajoukon $x_1, \ldots, x_5$ suhteen. **(2p)**

(c) Ainon kollega Uolevi epäilee, että vain pieni osa kyseisen lajin perhosista on oranssisiipisiä. Tätä uskomusta kuvastaa hänen mielestään hyvin priorijakauma, jonka tiheysfunktio on $p_0(\theta) = 8\theta$, $0 < \theta < 1/2$.

Auta Ainoa muodostamaan Bayes-estimaatti parametrin θ arvolle laskemalla piste, jossa posteriorijakauman tiheysfunktio saa suurimman arvonsa, kun posteriorijakauma lasketaan käyttämällä priorijakaumaa p_0, datapisteitä $x_1, \ldots, x_5$ ja samaa stokastista mallia kuin (a)-kohdassa.

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