

TODENNÄKÖISYYSLASKENNAN JA TILASTOTIETEEN PERUSKURSSI (MS-A0501/2/3/4). 26.2.–16.4.2024. Aalto-yliopisto. Pekka Pere, Natalia Vesselinova ja Jukka Kohonen.

## Koe 16.4.2024

Kokeessa saa käyttää funktiolaskinta (laskee potenssi-, eksponentti-, logaritmi- ja trigonometristen funktioiden arvoja sekä binomikertoimia). Kiellettyjä ovat ohjelmoitavat, symboliset ja graafiset laskimet. Laskimet, joilla saa yhteyden Internettiin (esim. kännykän laskin) ovat myös kiellettyjä. Voit tuoda kokeeseen käsin kirjoittamasi kaksipuolisen A4-kokoisen muistilapun. **Esitä kaikkien laskujesi välivaiheet, ja perustele kaikki vastauksesi yksityiskohtaisesti. Pelkästä oikeasta vastauksesta ei saa pisteitä.** Kaikki tehtävät ovat kuuden pisteen arvoisia. Vastaa rauhassa. Tarkista vastauksesi! Ja vielä toisen kerran! Parhainta koemenestystä!

1. Kohta a) 4 pistettä. Kohta b) 2 pistettä.

a) Oletetaan, että  $P(A) = 0.5$ ,  $P(B) = 0.3$  ja  $P(A \text{ ja } B) = 0.2$ . Määrää seuraavat todennäköisyydet:

- joko A tai B tai molemmat tapahtuvat
- vain A tapahtuu (B ei tapahdu)
- ainakin A tapahtuu
- korkeintaan A tapahtuu (ainakaan B ei tapahdu).

b) Aku on kolmekymmentäyksivuotias ja naimaton. Häntä pidettiin nörttinä jo lukiossa. Opiskeluaikanaan Aku kiinnostui sellaisista arvoista ja asioista kuin sananvapaus, yksilönsuoja, viestintäsalaisuus, yksilön oikeusturva, kansalaisten yhteiskunnalliseen päätöksentekoon osallistumisen kehittäminen, yhteiskunnallisen datan julkistaminen ja avoin data ylipäänsä. Aku ei valmistunut yliopistosta, koska hän käytti valtaosan ajastaan aktivismiin edellä lueteltujen asioiden puolesta. Kumpi on todennäköisempää?: i) Aku on verovirkailija. ii) Aku on verovirkailija ja aktiivinen Piraattipuolueen jäsen. (Vihje: Mitenkään ei välttämätöntä, mutta saatat haluta piirtää Venn-diagrammin.)<sup>1</sup>

2.

a) *Tyrannosaurus rexin* keskimääräiseksi ekologiseksi painoksi (*ecological body mass*) on arvioitu 5 200 kg ja ekologisen painon jakauman 0.025. ja 0.975. kvantiileiksi 3 700 kg ja 6 900 kg. Voidaanko tiedoista päätellä, että *tyrannosaurus rexin* ekologisen painon jakauma on symmetrinen?

b) Keskimääräisellä ekologisella painolla tarkoitettiin edellä eri ikäisten *Tyrannosaurus rexien* keskimääräistä painoa. Täysikasvuisen *Tyrannosaurus rexin* painon on arvioitu olevan keskimäärin 7 100 kg ja painon jakauman 0.025. ja 0.975. kvantiilien olevan 5 100 kg ja 9 100 kg. Voidaanko tiedoista päätellä, että täysikasvuisen *Tyrannosaurus rexin* painon jakauma on symmetrinen?

c) Täysikasvuisen *Tyrannosaurus rexin* painon on arvioitu olevan normaali-jakautunut kohdassa b) todetuin parametrein. Mitkä ovat täysikasvuisen *Tyrannosaurus rexin* painon keskihajonta ja varianssi?<sup>2</sup> (Vihje: Standardinormaalijakauman 0.025. ja 0.975. kvantiili ovat  $-1.959964$  ja  $1.959964$ .)

<sup>1</sup>Tehtävä on mukaelma psykologi Daniel Kahnemanin kirjan (2011) *Thinking, Fast and Slow* luvun 15 tehtävästä.

<sup>2</sup>C.R. Marshall ym. (2021): Absolute Abundance and Preservation Rate of *Tyrannosaurus Rex*. *Science*, 372, 284–287. Erityisesti artikkelin taulukko 1 ja kuvion 2 osio C.

3. Eurooppalaisessa yhteiskunnallisessa kyselyssä (*European Social Survey ESS*) haastateltiin vuonna 2012 noin 1 700 täysi-ikäistä suomalaista. Haastattelut tehtiin noin tunnin kestäneissä tapaamisissa kasvokkain haastattelijan kanssa. Haastatelluista 1 311 eli 77.11765 prosenttia kertoi äänestäneensä vuoden 2011 eduskuntavaaleissa. Äänestysprosentti oli vaaleissa 70.5.<sup>3</sup>

Laske ESS:n kyselyn perusteella kaksisuuntainen 95%:n luottamusväli äänestysprosentille vuoden 2011 eduskuntavaaleissa. Kattaako luottamusväli todellisen äänestysosuuden 70.5 %? (Vihje: Standardinormaalijakauman 0.95., 0.975., 0.99. ja 0.995. kvantiilit ovat 1.645, 1.960, 2.326 ja 2.576.)

4. Mobiiliverkkoon saapuvien puhelinpalvelupyyntöjen määrä oletetaan toisistaan riippumattomiksi, ja niiden keskimääräinen määrä kiinteää aikaväliä kohti oletetaan vakioksi  $\lambda$ . Tällöin puhelujen määrää voidaan mallintaa Poisson-jakaumalla, jonka parametri on  $\lambda > 0$ . Poisson-satunnaismuuttujan  $X$  pistodennäköisyysfunktio (tiheysfunktio) on

$$p_X(x) = P(X = x) = e^{-\lambda} \frac{\lambda^x}{x!}.$$

Tässä  $x = 0, 1, \dots$  on välin aikana vastaanotettujen soittopyyntöjen määrää ja  $\lambda > 0$ .

Puhelinoperaattori rekisteröi tietyssä mobiiliverkon sijaintipaikassa  $x_1 = 10$ ,  $x_2 = 5$  ja  $x_4 = 7$  puhelut kolmen samanpituisen, toisistaan riippumattoman ajanjakson aikana. Kun tiedetään, että soittopyyntöjen määrä noudattaa Poisson-jakaumaa, operaattori haluaa arvioida niiden keskimääräisen määrän  $\lambda$  parantaakseen verkon resurssien hallintaa ja puheluiden laatua.

a) Mikä on  $\lambda$ :n suurimman uskottavuuden estimaatti?

b) Puhelinoperaattorin näkemystä parametrin  $\lambda$  mahdollisista arvoista kuvaa eksponenttijakauma

$$e^{-\lambda}.$$

Johda annettuihin tietoihin perustuva normittamaton posteritiheysfunktio (normitusvakiota ei tarvitse laskea). Etsi  $\lambda$ :n arvo, joka maksimoi normittamattoman posterioritiheysfunktion (ns. suurimman posterioritiheyden estimaatti eli MAP-estimaatti).

c) Puhelinoperaattori rekisteröi uuden edellisiä vastaavan riippumattoman lisähavainnon  $x_4 = 20$ . Päivitä normittamaton posterioritiheysfunktio, ja laske päivitetty MAP-estimaatti  $\lambda$ :lle.

---

<sup>3</sup>[http://tilastokeskus.fi/til/evaa/2011/evaa\\_2011\\_2011-04-29\\_tie\\_001\\_fi.html](http://tilastokeskus.fi/til/evaa/2011/evaa_2011_2011-04-29_tie_001_fi.html) (haettu 26.11.2020). Suomessa asuvien suomalaisten äänestysprosentti on sivun mukaan 70.5. Sivulla annetuista tiedoista voidaan laskea, että jos huomioidaan ulkomailla asuvat äänestysaikeutetut suomalaiset (joiden äänestysprosentti oli 10.6), että kaikkien äänioikeutettujen äänestysprosentti oli 67.4. Suomessa ESS on toteutettu tasavälipoiminnalla, mutta voit olettaa tehtävää ratkaistaessasi, että otantamenetelmä olisi ollut yksinkertainen satunnaisotanta. Kiitän tilastotieteen professori emeritus Seppo Laaksosta, joka osoitti minulle tehtävän ilmiön ja tiedot siitä.