TODENNÄKÖISYYSLASKENNAN JA TILASTOTIETEEN PERUSKURSSI. 26.2.–16.4.2024. Aalto-yliopisto. Yliopistonlehtori Pekka Pere, tutkijatohtori Natalia Vesselinova.

Uusintakoe 7.6.2024

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

1. Potilas on sairastunut masennukseen. Potilas voi parantua lääkärin määräämällä pitkäaikaisella psykoanalyysilla tai lääkityksellä tai niiden yhdistelmällä.

a) Oletetaan, että potilas paranee masennuksesta psykoanalyysin avulla todennäköisyydellä 0.6, lääkityksen avulla todennäköisyydellä 0.3 ja että hoidot toimivat toisistaan riippumattomasti.[1] Lääkäri määrää potilaalle sekä psykoanalyysin että lääkityksen. Mikä on todennäköisyys, että potilas paranee masennuksesta?

b) Luovutaan oletuksista, että psykoanalyysi- ja lääkityshoidot toimisivat toisistaan riippumattomasti. Oletetaan, että potilas paranee

- pelkällä psykoanalyysillä todennäköisyydellä 0.6.
- pelkällä lääkityksellä todennäköisyydellä 0.3.
- yhtäaikaisella psykoanalyysilla ja lääkityksellä todennäköisyydellä 0.8.

Lääkäri määrää masennuspotilaalle sopivan hoitomuodon potilaasta riippuen huolellisen harkinnan ja tutkinnan jälkeen (onko potilaalla varaa maksaa psykoanalyysista, ovatko lääkityksen sivuvaikutukset potilaan mielestä kohtuulliset jne.). Lääkäri määrää hoidoksi

- pelkän psykoanalyysin todennäköisyydellä 0.05
- pelkän lääkityksen todennäköisyydellä 0.90
- sekä psykoanalyysin että lääkityksen todennäköisyydellä 0.05.

Mikä on todennäköisyys, että potilas paranee?

c) Lisätään mahdollisuus, että potilas paranee masennuksesta itsestään todennäköisyydellä 0.2 riippumatta siitä, saako hän hoitoa vai ei. Mikä on todennäköisyys, että potilas paranee a)-kohdan oletuksilla, jos lääkäri määrää potilaalle sekä psykoanalyysin että lääkityksen?

[1] Kiitän kehittämispäällikkö (Terveyden ja hyvinvoinnin laitos) Olavi Lindforsia masennustutkimusten osoittamisesta ja niiden tulosten selvittämisestä minulle 4.9.2012. Pelkkää psykoanalyysihoitoa saaneista parani 60 % tutkimuksessa P. Knekt, M.A. Laaksonen, T. Härkänen, T. Maljanen, E. Heinonen, E. Virtala ja O. Lindfors (2012): The Helsinki Psychotherapy Study: Effectiveness, Sufficiency, and Suitability of Short- and Long-Term Psychotherapy. Teoksessa R. Levy, J.S. Ablon ja H. Kächele (toim.): *Psychodynamic Psychotherapy Research.* Humana Press. S:t 71–94. Tehtävän muut todennäköisyydet ja riippumattomuusoletus eivät perustu tutkimustietoon. Lääkehoidon tehokkuudesta pitkällä aikavälillä ei ole luotettavaa tietoa.

2. Arpajaisia on järjestetty Euroopassa 1500- ja ehkä jopa 1400-luvulta lähtien. Arpajaiset yleistyivät 1700-luvulla ja erityisesti sen jälkipuoliskolla. Suomessa tiettävästi ensimmäinen arvonta järjestettiin 1745 Turun tuomiokirkon 1738 palaneen tornin korjauksen rahoittamiseksi.[2]

Asetuksessa 1772 kuvataan arvonta Turun hiippakunnan sairaalan (Lazaretti) hyväksi. Lottoarvat (Lottisedelit) olivat myyneet huonosti, joten voittoja korotettiin. Kussakin arvonnassa oli 3 200 arpaa, jotka kukin maksoivat 1.5 taaleria (1 taaleria ja 16 äyriä). Päävoitto oli 900 taaleria. Muita voittavia arpoja oli 400, joilla kullakin voitti 8.31 taaleria (aiemmin 6 taaleria). Sairaala sai 12 % (= avustus) arpojen myyntituloista (vähentämättä menoja palkinnoista). Oletetaan, että kaikki lottoarvat myydään.

a) Mikä on yhden arvan voiton odotusarvo? (Voitto = arvan tuottama raha huomioimatta arvan hintaa.)

b) Mikä on yhden arvan voiton varianssi? Entä keskihajonta?

c) Mikä on yhden arvan tuoton odotusarvo? (Tuotto = voitto − arvan hinta.)

d) Mikä on yhden arvan tuoton varianssi? Entä keskihajonta?

e) Mikä on sairaalan saaman avustuksen odotusarvo?

f) Mikä on sairaalan saaman avustuksen varianssi? Entä keskihajonta?

3. Eurooppalaisessa sosiaalitutkimuksessa (*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.[3]

Haastattelut ovat riippumattomia. Testaa, onko kyselyssä havaittu äänestysosuus 0.7711765 satunnaisvaihtelun puitteissa sopusoinnussa todellisen äänestysosuuden 0.705 kanssa.

a) Selitä huolellisesti, miksi kyselyssä havaittu äänestysosuus, sopivasti standardoituna, noudattaa likimain standardinormaalijakaumaa. (Vihje: Keskeinen raja-arvolause.)

b) Muodosta a)-kohdassa muodostetun standardoidun äänestysosuuden avulla testisuure nollahypoteesille “äänestysosuus on 0.705”, ja laske sen arvo.

c) Mitä päättelet? Käytä merkitsevyystasoa 0.01 (kaksisuuntainen testaus). (Vihje: Standardinormaalijakauman 0.95., 0.975., 0.99., 0.995., 0.999. ja 0.9995 kvantiilit ovat 1.645, 1.960, 2.326, 2.576, 3.090 ja 3.291.)

[2] S.M. Stigler (2022): *Casanova's Lottery*. University of Chicago Press. Luvut 5, 14. O.E.A. Hjelt (1893): *Svenska och finska medicinalverkets historia 1663–1812, osa 3*. Helsingfors central-tryckeri. S. 89. Kotimaisten kielten keskus: `https://kaino.kotus.fi/korpus/vks/meta/lait/as1700_rdf.xml` (haettu 24.5.2024). Kiitän Katariina Lehtoa Suomen ensimmäisen arvonnan ja Hjeltin kirjan osoittamisesta. Asetuksen tulkinta on tehtävänlaatijan.

[3] `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 äänestysoikeutetut 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.

4. Potilaat saapuvat sairaalan päivystysyksikköön tuntemattomalla intensiteetillä λ. Saapumisten oletetaan olevan toisistaan riippumattomia ja niiden lukumäärän noudattavan Poisson-jakaumaa tuntemattomalla parametrilla $\lambda > 0$.

Poisson-jakautuneen satunnaismuuttujan X pistetodennäköisyysfunktio on

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

Yllä $x = 0, 1, \ldots$ on tapahtumien lukumäärä ajanjakson aikana.

Tuntemattoman parametrin λ priorijakauma on eksponenttijakauma

$$p(\lambda) = \begin{cases} e^{-\lambda} & \lambda > 0, \\ 0 & \text{muuten.} \end{cases}$$

a) Potilaiden määrä yhden työviikon tietyn tunnin aikana on $x_1 = 100$, $x_2 = 110$, $x_3 = 90$, $x_4 = 50$ ja $x_5 = 150$. Määritä λ:n (vakiolla normeeraamaton) posteriorijakauma.

b) Viikonloppuna saadaan uusia aiemmista havainnoista $(x_1, \ldots, x_5)$ riippumattomia havaintoja: $x_6 = 250$ ja $x_7 = 200$. Päivitä (vakiolla normeeraamaton) posteriorijakauma.