5A Luottamusvälien määrittäminen

Normaalijakauman kertymäfunktion voi laskea R:ssä funktiolla `pnorm`, ja sen käänteisfunktion funktiolla `qnorm`, esimerkiksi `pnorm(1.96)` ≈ 0.975, ja vastaavasti `qnorm(0.975)` ≈ 1.96.

Matlabissa ja Octavessa vastaavat komennot ovat `normcdf` ja `norminv`.

Jos tietokonetta tai normaalijakauman osaavaa laskinta ei ole käytettävissä, kertymäfunktion arvoja voi etsiä taulukoista (ks. kurssisivu, kohta Materiaalit).

Tuntitehtävät

5A1 (Limuautomaatti) Limuautomaatti laskee mukiin juomaa määrän (ml), joka noudattaa likimain normaalijakaumaa odotusarvona μ ja keskihajontana $\sigma = 3$. Automaattia testatessa mitattiin mukeihin valutetuiksi juomamääriksi (ml): 304, 298, 301, 302, 301, 300, 305, 300, 306.

(a) Määritä 95% luottamustason väliestimaatti parametrille μ.

(b) Määritä 99% luottamustason väliestimaatti parametrille μ.

(c) Kun suoritetaan koe, jossa otetaan 9 mukia juomaa ja lasketaan luottamusväli a-kohdan mukaisesti, mikä on todennäköisyys, että luottamusväli (i) sisältää arvon μ, (ii) on kokonaan μ:n alapuolella, (iii) on kokonaan μ:n yläpuolella?

(d) Mitä enemmän dataa on saatavilla, sen kapeampi luottamusväli saadaan. Kuinka monta mittausta vaadittaisiin, että 95% luottamustason väliestimaatti saataisiin alle 1 ml levyiseksi (0.5 ml kumpaankin suuntaan)? Entä 0.1 ml levyiseksi?

5A2 (Mielipidemittaus) Helsingin Sanomien heinäkuussa 2016 raportoiman kyselytutkimuksen mukaan 89 prosenttia suomalaisista oli sitä mieltä, että presidentti Niinistö on suoriutunut tehtävästään erittäin tai melko hyvin. Kysely toteuttiin haastattelemalla puhelimitse 1002 suomalaista ikähaarukassa 15–79 vuotta ja virhemarginaalin kerrottiin olevan noin 3 prosenttiyksikköä suuntaansa. Oletetaan, että virhemarginaali on laskettu käyttämällä binaarimallin konservatiivista väliestimaattoria (luentomoniste, kaava (8.5)).

(a) Päättele annettujen tietojen perustella, mitä luottamustasoa kyselytutkimusten virhemarginaalin raportoinnissa käytettiin.

(b) Kuinka monta suomalaista olisi pitänyt haastatella, jos virhemarginaaliksi olisi samalla luottamustasolla haluttu noin 1 prosenttiyksikkö suuntaansa?

Kotitehtävät

5A3 (Monen puolueen kyselytutkimus) Eräästä suuresta populaatiosta otettiin $n = 100$ henkilön kokoinen satunnaisotos, jolta kysyttiin mitä neljästä puolueesta A,B,C,D vastaajat kannattavat. Kyseisten puolueiden kannattajia oli otoksessa 80, 18, 2 ja 0 henkilöä.

(a) Käsittele kunkin puolueen X kohdalla erikseen kysymystä "mikä osuus populaatiosta kannattaa puoluetta X" binaarisena kysymyksenä (ks. luentomonisteen luku 8.3) ja laske 95% luottamusväli kyseisen puoleen kannatukselle populaatiossa. Ilmoita välit alku- ja loppupisteen avulla kolmella desimaalilla esim. muodossa [0.500, 0.600].

(b) Toista edellisen kohdan laskut käyttäen konservatiivista väliestimaattoria.

(c) Ovatko edellisissä kahdessa kohdassa lasketut välit mielekkäitä? Jos eivät, millä tavalla eivät ja mitä arvelet syyksi? Pohdi mitä asialle voisi tehdä. Opastus: Pieni pohtiminen riittää. Tällaisiin tilanteisiin on olemassa kirjallisuudessa ratkaisuja, mutta ne ovat hiukan mutkikkaampia kuin tällä kurssilla esitetyt kaavat.

(d) Onko mahdollista, että puolueen kannatus populaatiossa on nollaa suurempi, mutta otoksessa sen osuus on nolla? (Järkeile, älä laske.)

(e) Onko mahdollista, että puolueen kannatus populaatiossa on tasan nolla, mutta otoksessa sen osuus on nollaa suurempi? (Järkeile, älä laske.)

5A4 (Alkeishiukkaset) Eräässä fysikaalisessa kokeessa havaitaan hiukkasten hajoamisia satunnaisin välein. Väliajat oletetaan riippumattomiksi ja eksponenttijakautuneiksi tuntemattomalla odotusarvolla μ, jolloin kyseisen jakauman taajuusparametri on $\lambda = 1/\mu$. Mitattiin 30 peräkkäistä hajoamisten väliaikaa. Havaittujen väliaikojen keskiarvo oli 13.15 sekuntia ja keskihajonta 12.18 sekuntia.

(a) Käyttäen yleistä odotusarvon luottamusväliä (monisteen luku 8.2 / luento 4B), määritä 90% luottamusväli odotusarvoparametrille μ. Ilmoita tulos kahdella desimaalilla. Huom. vaadittu luottamustaso.

(b) Yksittäiset hiukkashajoamiset eivät olleet normaalijakautuneita. Miksi tässä käytetty menetelmä on kuitenkin likimain oikea?

(c) Voisiko samaa menetelmää käyttää luotettavasti, jos havaintoja (hajoamisten väliaikoja) olisi vain $n = 3$ kappaletta? Miksi / miksi ei?