

Tilastollinen päättely I – erilliskoe 19.11.2015
(ent. Johdatus tilastolliseen päättelyyn)

Omia taulukoita, kaavakirjoja tai "lunttilappuja" ei saa käyttää. Tarvittava taulukko on kääntöpuolella.

1. Tarkastellaan seuraavanlaista koeasetelmaa: Tiettyä koetta toistetaan riippumattomasti samanlaisissa olosuhteissa n kertaa. Kukin koe joko *onnistuu* tai *epäonnistuu*, ja onnistumisen todennäköisyys on tuntematon luku θ . Satunnaismuuttuja X kertoo onnistumisten lukumäärän.
 - a) Minkä nimistä jakaumaa X noudattaa, ja mikä on sen pistetodennäköisyysfunktio?
 - b) Oletetaan, että toistoja tehtiin $n = 9$ ja niistä $X = 2$ onnistui. Muodosta tätä havaintoa vastaava uskottavuusfunktio ja määritä suurimman uskottavuuden estimaatti θ :lle.
2.
 - a) Tarkastellaan tilastollista mallia $f(y; \theta)$, jonka parametri θ on yksiulotteinen eli reaalinen. Oletetaan, että aineiston y perusteella on θ :lle muodostettu 95 %:n luottamusväli $[L, U]$, jossa L ja U ovat reaalitylukuja. Selosta, mitä tämä tarkoittaa ja erityisesti miten luku 95 % eli 0.95 on tulkinnallisesti ymmärrettävää.
 - b) Olkoon Y_1, \dots, Y_n satunnaisotos normaali-jakaumasta $N(\mu, 1)$. Kuinka suuri otoskoon n on oltava, jotta havaintojen y_1, \dots, y_n perusteella muodostettava tavanomainen 95 %:n luottamusväli μ :lle (ns. z -luottamusväli) olisi pituudeltaan (noin) 0.2 eli muotoa $[\bar{y} - 0.1, \bar{y} + 0.1]$, jossa $\bar{y} = (y_1 + \dots + y_n)/n$ on otoskeskiarvo?
3. Puutavaraerän laatuvaatimuksiin kuuluu, että puun keskimääräinen kosteusprosentti saa olla korkeintaan 20. Laadun varmistamiseksi tehtiin kuusi riippumatonta kosteusmittausta, joiden tulokset olivat 21, 18, 24, 23, 20 ja 26. Laske mittaustulosten keskiarvo ja keskihajonta sekä testaa merkitsevyystasolla 0.05, voidaanko päätellä, että puutavara ei täytä laatuvaatimusta. Muista lausua testattavat hypoteesit selvästi. Oletetaan, että kosteusprosentin vaihtelut ovat (likimain) normaalisti jakautuneita.
4. Laboratoriotestillä pyritään selvittämään, sisältääkö verenluovuttajan veri HIV:n vasta-aineita. Kuvataan vasta-aineiden olemassaoloa veressä muuttujalla θ siten, että $\theta = 0$, jos vasta-aineita ei ole, ja $\theta = 1$, jos niitä on. Testin herkkyys on 0.995, ts. testituloksella positiivinen tällä todennäköisyydellä, jos veressä todella on HIV:n vasta-aineita. Testi antaa todennäköisyydellä 0.010 (väärän) positiivisen tuloksen silloin, kun vasta-aineita ei oikeasti ole. Oletetaan, että HIV:n vasta-aineita on väestössä yhdellä tuhannesta ja että verenluovuttajat ovat satunnaisotos väestöstä.
 - a) Mikä on θ :n priorijakauma?
 - b) Oletetaan, että satunnaisesti (eli "umpimähkään") valittu luovuttaja on saanut positiivisen testituloksen. Mikä on θ :n posteriorijakauma?