Tilastollinen päättely I, kevät 2017
Harjoitus 6 – palautuspäivämäärä 2. 5.

Tehtävät 1–5 liittyvät monisteen lukuun 6. Testien voiman käsitettä (jakso 6.3 ja osittain 6.5), testien ja luottamusjoukkojen duaalisuutta (jakso 6.6) ja binomijakauman testausta (jakso 6.8) emme kuitenkaan kurssilla käsittele, vaikkakin erityisesti jakso 6.6 on hyödyllistä luettavaksi itse. Tarvittava t-jakauman taulukko jaettiin harjoituksen 3 yhteydessä. Vaihtoehtoisesti käytä online-laskinta http://surfstat.anu.edu.au/surfstat-home/tables/t.php.

1. Professorilla on tavoitteena mitoittaa kurssinsa vaatima viikottainen työmäärä siten, että opiskelijat keskimäärin joutuisivat käyttämään kurssiin liittyvään opiskeluun enintään 10 tuntia. Eräällä viikolla tehdyssä otoksessa (kokoa 13) opiskelijat käyttivät aikaa opiskeluun keskimäärin 11.1 tuntia ja aikojen keskihajonta oli 1.8 tuntia. Testaa t-testillä ja 5 %:n merkitsevyystasolla, onko professorin syytä katsoa epäonnistuneen tavoitteessaan. Oletamme (paremman puutteessa), että vaihtelu opiskeluajoissa opiskelijoiden kesken on likimain normaalisti jakautunutta.

2. Jatkoa harjoituksen 4 tehtävään 2. Asetetaan hypoteesit $H_0\colon \mu = 3$ ja $H_1\colon \mu \neq 3$. Testaa näitä kaksisuuntaisella t-testillä ja merkitsevyystasolla 0.05 käyttäen a) kummankin tutkimusryhmän aineistoja erikseen, b) yhdistettyä aineistoa. Aineistoista lasketut tulokset (otoskoko, otoskeskiarvo ja otosvarianssi) olivat ryhmille erikseen ja yhdessä

$$\begin{aligned} n_A &= 10, \quad \bar{y}_A = 3.952, \quad s_A^2 = 1.981, \\ n_B &= 15, \quad \bar{y}_B = 3.411, \quad s_B^2 = 1.321 \\ n &= 25, \quad \bar{y} = 3.627, \quad s^2 = 1.587. \end{aligned}$$

3. Tutkitaan taas satunnaisotosta jakaumasta $N(\mu, \sigma^2)$ eli mallia $Y_1, \ldots, Y_n \sim N(\mu, \sigma^2) \perp\!\!\!\perp$, jossa sekä μ että σ^2 ovat tuntemattomia parametreja.

Tehtäväsi on johtaa testi keskihajonnan σ (tai yhtäpitävästi varianssin σ^2) testaamiseen. Tarkemmin sanoen tarkastellaan yksisuuntaista testausasetelmaa, jossa hypoteesit ovat $H_0\colon \sigma = \sigma_0$ ja $H_1\colon \sigma > \sigma_0$, kun $\sigma_0 > 0$ on tunnettu luku (H_0 voisi olla myös $\sigma \leq \sigma_0$). Katso muistiinpanojen kohtaa 5.6.4 (sivu 53) ja vastaa seuraaviin kysymyksiin:

a) Mikä olisi luonteva testisuure?

b) Millaiset testisuureen arvot (suuret vai pienet) ovat kriittisiä H_0:lle ja tukevat H_1:tä?

c) Mitä jakaumaa testisuure noudattaa, kun $\sigma = \sigma_0$ eli H_0 pätee?

d) Miten p-arvo lasketaan?

4. Palauta mieleen harjoituksen 1 tehtävän 4d kysymyksenasettelu ja siinä tehty lasku. Miten muotoilisit tämän täsmällisenä testausasetelmana? Mitkä ovat nollahypoteesi ja vastahypoteesi? Jos testisuureena käytetään suurinta havaintoa eli odotusaikaa $y_{(50)} = \max(y_1, \ldots, y_{50})$, millaiset testisuureen arvot (suuret vai pienet) ovat kriittisiä nollahypoteesin kannalta? Mikä on p-arvo eli havaittu merkitsevyystaso? Entä johtopäätökset?

5. Luennolla ja monisteessa (erityisesti jaksoissa 6.4 ja 6.9) on selostettu p-arvon ja siihen liittyvien johtopäätösten tekoa sekä oikeaa tulkintaa.

a) Eräs suomalainen oppikirja[1] kuvailee asiaa näin: *Hypoteesi ei koskaan jää voimaan absoluuttisesti, vaan se jää voimaan jollain todennäköisyydellä. P-arvot ovat yksinkertaisia todennäköisyyslukuja, jotka vaihtelevat välillä* $[0, 1]$ *kuten klassinen todennäköisyyskin. P-arvot ilmoittavat, kuinka suurella todennäköisyydellä vaihtoehtoinen hypoteesi on väärä.*

[1] L. Nummenmaa: *Käyttäytymistieteiden tilastolliset menetelmät*, Tammi, 2009. Sivut 148–149.

Mitä lähempänä p-arvo on ykköstä, sitä suuremmalla todennäköisyydellä nollahypoteesi on asetettu oikein. Jos p-arvo on taas lähellä nollaa, vaihtoehtoinen hypoteesi on erittäin todennäköisesti oikea. Koska p-arvot ovat todennäköisyyksiä, niitä voidaan ajatella myös prosentteina. Toisin sanoen $p = .5$ *tarkoittaa* 50 %*:n todennäköisyyttä.*

Millaisia virheitä ja/tai puutteita tässä kuvauksessa on?

b) Eräs englanninkielinen oppikirja[2] selostaa näin: *Most studies require very small p-values, such as* $p \leq 0.05$*, in order to reject* H_0*. In such cases, the results are said to be significant at the* 0.05 *level. -- Making a decision by rejecting or not rejecting a null hypothesis is an optional part of the significance test. -- the study should interpret the p-value in context. The smaller p is, the stronger the evidence against* H_0 *and in favor of* H_1*. -- Why do smaller p-values indicate stronger evidence against* H_0*? Because the data would then be more unusual if* H_0 *were true. -- In practise, it is sometimes necessary to decide whether the evidence against* H_0 *is strong enough to reject it. The decision is based on whether the p-value falls below a prespecified cutoff point. -- The* α*-level is a number such that we reject* H_0 *if the p-value is less than or equal to it. The* α*-level is called the significance level.*

Onko tämä kuvaus mielestäsi oikein ja yhtäpitävä kurssilla oppimasi kanssa?

6. a) Sukulaisesi, joka ei ole opiskellut lainkaan tilastotiedettä, pyytää sinua kertomaan yhdellä virkkeellä, mistä tilastollisessa päättelyssä on kysymys. Miten vastaisit hänelle tämän kurssin pohjalta?

b) Luettele kolme mielestäsi tärkeintä (tai ainakin mielenkiintoisinta) tällä kurssilla opittua tilastollisen päättelyn käsitettä tai menetelmää.

[2] A. Agresti ja B. Finlay: *Statistical Methods for the Social Sciences*, 4. laitos, Pearson, Lontoo, 2008.