Tilastollinen päättely II, syksy 2015 – kevät 2016
Harjoitus 10 (9. ja 11. 2. 2016)

Kaikki tehtävät liittyvät testiteoriaan; keskeisimmin monisteen jaksoihin 5.2–5.4.

1. Olkoon $f_{\mathbf{Y}}(\mathbf{y};\theta)$ jokin tilastollinen malli ja $H_0\colon \theta \in \Omega_0$ siihen liittyvä hypoteesi, jota voidaan testata testisuureella $t = t(\mathbf{y})$ (suuret arvot H_0:lle kriittisiä). Olkoon g jokin aidosti kasvava funktio, ja määritellään $u = u(\mathbf{y}) = g(t(\mathbf{y}))$. Perustele, että testisuuretta u käyttämällä saadaan samat p-arvot ja samat kriittiset alueet kuin t:tä käyttämällä. Sanomme tällöin, että testisuureet t ja u ovat *ekvivalentit*.

2. Normaalijakauman testien kertaus (monisteen teht. 5.3). Kemiantehtaassa kone annostelee erästä kemikaalia kanistereihin. Oletetaan, että kerralla annostellun kemikaalin määrä (litroina) noudattaa normaalijakaumaa. Pyrkimyksenä on säätää kone siten, että keskimääräinen annos μ on 10 ja keskihajonta σ korkeintaan 0.2. Tutkittiin 20 kanisteria ja havaittiin, että niissä oli kemikaalia keskimäärin $\overline{y} = 9.86$ (litraa), keskihajonnan ollessa $s = 0.25$. Testaa kaksisuuntaisella t-testillä ja yksisuuntaisella χ^2-testillä, onko kone säädön tarpeessa. Käytä 5 %:n merkitsevyystasoa.

Ohje. Kertaa tarvittavat testit JTP-kurssilta. Voit myös katsoa monisteen jaksoa 5.4.

3. (Vrt. monisteen teht. 5.5.) Olkoot Y_1 ja Y_2 kaksi riippumatonta havaintoa Poissonin jakaumasta $P(\mu)$. Testataan hypoteesia $H_0\colon \mu = 2$ (tai $\mu \geq 2$) vastaan $H_1\colon \mu < 2$. Testisuureena on $T = Y_1 + Y_2 \sim P(2\mu)$.

a) Millaiset testisuureen arvot todistavat H_0:aa vastaan ja H_1:n puolesta: pienet vai suuret?

b) Laske testisuureen havaittuja arvoja $t = 0$, $t = 1$, $t = 2$ ja $t = 3$ vastaavat p-arvot.

4. Jatkoa edelliseen tehtävään.

c) Mitkä t:n arvot johtavat H_0:n hylkäämiseen ja H_1:n hyväksymiseen merkitsevyystasolla $\alpha = 0.1$? Mitkä havaintoparit (y_1, y_2) kuuluvat vastaavaan kriittiseen alueeseen?

d) Jos toimitaan merkitsevyystasolla $\alpha = 0.1$, niin missä tilanteessa tehdään hylkäämisvirhe ja missä hyväksymisvirhe? (Ts. mitkä μ:n arvot ja t:n arvot tai aineistot (y_1, y_2) johtavat näihin virheisiin?)

e) Onko olemassa mitään sellaista aineistoa (y_1, y_2), jonka havaitessasi voisit *varmuudella* sanoa, että H_0 pätee tai vastaavasti että se ei päde?

5. Pöydän kummallakin sivulla on kolme tuolia ja lisäksi kummassakin päässä yksi tuoli. Sosiologi tekee satunnaiskokeen, jossa hän antaa neljästä miehestä ja neljästä naisesta koostuvan ryhmän istuutua pöydän ympärille heidän haluamallaan tavalla.

a) Satunnaismuuttuja Y kertoo pöydän päissä istuvien miesten lukumäärän (sen mahdolliset arvot ovat siis 0, 1 ja 2). Johda Y:n pistetodennäköisyysfunktio ja laske sen odotusarvo sekä varianssi olettaen, että henkilöt valitsevat paikkansa täysin satunnaisesti.

b) Sosiologi toistaa satunnaiskokeen 100 kertaa toisistaan riippumattomasti valituille koehenkilöiden ryhmille (kussakin neljä miestä ja neljä naista). Satunnaismuuttuja Y_i kertoo pöydän päissä istuvien miesten lukumäärän i:nnessä kokeessa ja $T = Y_1 + \cdots + Y_{100}$. Oletetaan, että T:n arvoksi havaitaan $t = 120$. Testaa nollahypoteesia, jonka mukaan miesten ja naisten hakeutumisella päätypaikoille ei ole eroa.

Apu. Testisuureen T nollahypoteesijakaumaa voi approksimoida normaalijakaumalla KRL:n avulla (tarvittavat tiedot laskettu a-kohdassa).