**Tilastollinen päättely I – kurssikoe 16. 5. 2016**
**Ratkaisuehdotukset ja pisteytyksistä**

Tehtävä 1: Joonas Nuutinen
Tehtävät 2,4: Topias Tolonen
Tehtävä 3: Aku Leivonen

**1.** Poimitaan satunnaisotos kokoa $n = 3$ (eli kolme riippumatonta havaintoa) eksponenttijakaumasta, jonka tiheysfunktio on

$$g(y; \lambda) = \lambda e^{-\lambda y}, \qquad y > 0,$$

ja jossa $\lambda > 0$ on tuntematon parametri. Kirjoita havaintoja $y_1 = 4.1$, $y_2 = 4.8$ ja $y_3 = 3.1$ vastaava uskottavuusfunktio ja johda derivaattatarkastelun avulla huolellisesti perustellen $\lambda$:n suurimman uskottavuuden estimaatti.

**Ratkaisu:**

Koska havainnot ovat riippumattomia, niin tilastollisen mallin lauseke eli yhteistiheysfunktio saadaan tulona

$$f(\mathbf{y}; \lambda) = \prod_{i=1}^{3} g(y_i; \lambda) = \lambda^3 e^{-\lambda \sum_{i=1}^{3} y_i} \quad \left(= \lambda^3 e^{-\lambda \cdot 3\bar{y}}\right).$$

Täten havaintoja $y_1 = 4.1$, $y_2 = 4.8$ ja $y_3 = 3.1$ vastaava uskottavuusfunktio on

$$L(\lambda) = \lambda^3 e^{-\lambda \sum_{i=1}^{3} y_i} = \lambda^3 e^{-12\lambda}, \qquad \lambda > 0.$$

Logaritmisen uskottavuusfunktion

$$l(\lambda) = \log L(\lambda) = 3 \log \lambda - 12\lambda$$

derivaatta on

$$l'(\lambda) = \frac{3}{\lambda} - 12 \quad \left(= \frac{3(1 - 4\lambda)}{\lambda}\right),$$

jolla on nollakohta $\frac{1}{4}$. Koska toinen derivaatta

$$l''(\lambda) = -\frac{3}{\lambda^2} < 0 \text{ kaikilla } \lambda > 0,$$

niin piste $\frac{1}{4}$ on log-uskottavuusfunktion globaali maksimikohta ja täten myös suurimman uskottavuuden estimaatti $\hat{\lambda} = \frac{1}{4} = 0.25$.

**Pisteytyksestä:**

Uskottavuusfunktion oikeasta muodostamisesta sai 3 pistettä ja su-estimaatin johtamisesta toiset 3 pistettä. Jos saatua derivaatan nollakohtaa ei oltu perusteltu maksimikohdaksi (esim. toisen derivaatan tai kulkukaavion avulla), niin siitä menetti yhden pisteen. Uskottavuusfunktion muodostamisesta jaettavat kolme pistettä menetti melko helposti, esimerkiksi jos muodosti kolme erillistä yksittäistä havaintoa vastaavaa uskottavuusfunktiota.

**2.** a) Erään kaupungin vuokrayksiöiden vuokratasoa selvitettiin satunnaisotoksella, jonka koko oli 25. Otoksessa keskimääräinen kuukausivuokra oli 450 euroa ja otoskeskihajonta 40 euroa. Muodosta 95 %:n luottamusväli kaupungin yksiöiden keskimääräiselle vuokratasolle, kun oletetaan, että vuokrien vaihtelut ovat (likimain) normaalisti jakautuneita.

b) Onko oikein sanoa, että noin 95 % koko kaupungin yksiöiden kuukausivuokrista sijoittuu ko. luottamusvälille? Perustele.

**Ratkaisu:**

a) Havaitaan $n = 25$, $\bar{y} = 450$, $s = 40$, sekä taulukosta näemme $t_{24}(0.025) \approx 2.064$.

Koska varianssiparametri on tuntematon, luottamusväliksi saadaan (monisteen kappale 5.6.)

$$\begin{aligned}
&[\bar{y} - t_{24}(0.025)\frac{s}{\sqrt{n}}, \bar{y} + t_{24}(0.025)\frac{s}{\sqrt{n}}] \\
=&[450 - 2.064 \cdot \frac{40}{\sqrt{25}}, [450 + 2.064 \cdot \frac{40}{\sqrt{25}}] \\
\approx&[433.488, 466.512].
\end{aligned}$$

b) Ei ole oikein sanoa. Väite on monella tapaa hyvin hullunkurinen: se ei ole yhtäpitävä kurssilla käytetyn luottamusvälin määritelmän kanssa, ja puhuu muun muassa yksittäisistä kuukausivuokrista eikä niinkään mallin parametrista.

Väitteen voi siis esimerkiksi kumota huomauttamalla sen eroavaisuudet frekventistisen päättelyn pointteihin luottamusvälistä. Alla myös kaksi näkökulmaa monisteesta, jotka kumoavat väiteen:
1) Aineistosta laskettu luottamusväli joko sisältää todellisen parametrinarvon tai tai ei sisällä sitä. Emme voi pelkästään aineistoa tarkastelemalla sanoa mitään sen enempää, vaan tätä varten pitäisi tuntea todellinen parametrinarvo.
2) Frekventistisessä tilastotieteessä parametri on tuntematon, mutta kiinteä (siis ei-satunnainen). Tämän lähestymistavan puitteissa väite $\mu \in [433.488, 466.512]$ on joko tosi tai epätosi. Tälläisen väitteen todennäköisyys ei taatusti ole 0.95.

**Pisteytyksistä:**

a)-kohdasta yhteensä neljä pistettä. Nyrkkisääntönä pisteytyksessä oli, että luottamusvälin oikeasta johtamisesta tai muistamisesta sai kaksi pistettä, ja järkevän oloisen vastauksen saamisesta toiset kaksi pistettä.

Jos t-jakauman yläkvantiilifunktion sijaan käytti normaalijakauman yläkvantiilifunktiota $z(\alpha/2)$, vähensin kaksi pistettä. Tämä on mielestäni kohtuu tiukka linja, mutta tehtävässä ei ollut juuri muuta arvosteltavaa, niin tällä vastauksiin tulee hieman hajontaa. Lisäksi tässä onnistuminen näyttää, että tiedostaa tuntemattoman ja tunnetun varianssin eron väliestimoinnissa. Jos $z$-yläkvantiilin perusteli kohtuu suurella otoskoolla, hyväksyin vastauksen sellaisenaan. 90% - luottamusvälin laskemisesta vähensin yhden pisteen, jos vastaus oli muuten oikein. Luottamusvälin johtamisen alku antoi aina yhden pisteen.

b)-kohdassa "Ei oikein-vastaus antoi yhden pisteen, ja kohtuu lepsusti annoin "järkevistä perusteluista"toisen pisteen, esimerkiksi ylläolevien argumenttien mukaisesti. Kuitenkin asiavirheistä - esimerkiksi LV:n idean frekventistisestä näkökulmasta väärin muistamisen - rokotin helposti yhden pisteen.

Tehtävä oli osattu yleisellä tasolla hyvin. a)-kohdassa hyvin suuri osa osasi muodostaa jonkinlaisen luottamusvälin, joksikin joukossa oli useita $z$-yläkvantiilifunktiota käyttäneitä. b)-kohdassa lähes kaikki osasivat todeta väitteen vääräksi (hyvä te!), mutta useat muistivat frekventistisen lähestymistavan luottamusväleihin virheellisesti..

**3.** (”teoriatehtävä”) Havainnot $y_1, \ldots, y_n$ ovat peräisin mallista, jossa vastaavat satunnaismuuttujat $Y_1, \ldots, Y_n$ ovat satunnaisotos normaalijakaumasta $N(\mu, \sigma^2)$. Testataan nollahypoteesia $H_0\colon \mu \leq \mu_0$, kun vastahypoteesi on $H_1\colon \mu > \mu_0$ ja $\mu_0$ on tunnettu reaaliluku. Kurssilla on opittu, että testaus perustuu testisuureeseen

$$t = \frac{\bar{y} - \mu_0}{s/\sqrt{n}},$$

jossa $\bar{y}$ on havainnoista laskettu otoskeskiarvo ja $s$ niiden otoskeskihajonta.

a) Millä kaavalla $s$ lasketaan havainnoista?

b) Miten määritellään ja lasketaan testin $p$-arvo (esim. sopivaa taulukkoa tai tietokoneohjelmaa käyttäen)?

c) Miten menetellään, jos halutaan tehdä päätös $H_0$:n hyväksymisestä tai sen hylkäämisestä ja $H_1$:n hyväksymisestä merkitsevyystasolla $\alpha = 0.01$? Millaiset virhemahdollisuudet päätöksentekoon liittyvät ja mitä nimityksiä niistä käytetään?

**Ratkaisu:**

a)

$$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

b) Testisuuretta vastaava satunnaismuuttuja

$$T = \frac{\bar{Y} - \mu_0}{S/\sqrt{n}},$$

noudattaa $t$-jakaumaa $n-1$:llä vapausasteella, kun $\mu = \mu_0$. Koska kyseessä on yksisuuntainen testi ja suuret testisuureen arvot ovat nollahypoteesin kannalta kriittisiä, niin testin $p$-arvo määritellään ja lasketaan kaavalla

$$p = P_{\mu=\mu_0}(T \geq t) = 1 - P_{\mu=\mu_0}(T < t) = 1 - F_{n-1}(t),$$

missä $F_{n-1}$ on $n-1$ vapausasteisen t-jakauman kertymäfunktio.

c) $H_0$ hyväksytään, mikäli $p \geq 0.01$. Vastaavasti $H_0$ hylätään ja $H_1$ hyväksytään, mikäli $p < 0.01$.

Toinen tapa ottaa kantaa nollahypoteesin hyväksymiseen tai sen hylkäämiseen ja vastahypoteesin hyväksymiseen on ratkaista valitun merkitsevyystason $\alpha = 0.01$ perusteella kriittinen piste, joka määrää testin hylkäysalueen. Koska kyseessä on yksisuuntainen testi ja nollahypoteesin kannalta kriittisiä ovat suuret testisuureen arvot, niin nollahypoteesi hylätään ja vastahypoteesi hyväksytään, mikäli

$$t > t_{n-1}(0.01),$$

missä $t_{n-1}(0.01)$ on $n-1$ vapausasteisen t-jakauman 0.01-yläkvantiili. Vastaavasti nollahypoteesi hyväksytään, mikäli $t \leq t_{n-1}(0.01)$.

Jos $H_0$ pitää paikkansa, mutta testi hylkää sen, tapahtuu hylkäysvirhe (käytetään myös nimitystä tyypin I virhe). Jos $H_1$ pitää paikkansa, mutta testi hyväksyy $H_0$:n, tapahtuu hyväksymisvirhe (käytetään myös nimitystä tyypin II virhe).

**Pisteytyksistä:**

a-kohdasta yksi piste, b-kohdasta kaksi pistettä ja c-kohdasta kolme pistettä.

Koska a-kohdassa oli vain yksi piste jaossa, siitä sai pisteen vain, jos kaava oli täysin oikein. Esimerkiksi neliöjuuren puuttuminen tai $n$:n käyttäminen jakajana $n-1$:n sijasta ei tuonut pisteitä. Moni muisti kyllä tämän kaavan hienosti ilman lunttilappuakin.

Huomatkaa, että b-kohdassa ei kysytty $p$-arvon sanallista tulkintaa. Moni saattoi menettää pisteitä sen vuoksi, että itse kysymykseen: "Miten määritellään ja lasketaan testin $p$-arvo...?" ei oikeastaan edes vastattu. Tyypillinen virhe oli myös sekoittaa $t$-testisuure $z$-testisuureeseen ja käyttää $p$-arvon laskemiseen normaalijakauman kertymäfunktiota $t$-jakauman kertymäfunktion sijasta. Toinen yleinen virhe oli käyttää kaksisuuntaista $t$-testiä ja laskea $p$-arvo kaavalla $2(1 - F_{n-1}(|t|))$. Näistä virheistä vähennettiin yksi piste.

Lähtökohtaisesti b-kohdassa yhden pisteen sai $p$-arvon oikeasta määrittelystä ja toisen pisteen $p$-arvon laskemisesta $t$-jakauman kertymäfunktion avulla. Myös täysin sanallisista vastauksista saattoi saada pisteitä, jos niistä kävi selvästi ilmi, miten $p$-arvo määritellään ja lasketaan tässä tapauksessa.

c-kohdassa oli tarjolla kolme pistettä ja erityisesti seuraavat virheet johtivat pisteiden vähentämiseen:

- Otettiin kantaa vain nollahypoteesin hylkäämiseen, mutta hyväksymisestä ei puhuttu mitään tai se ei käynyt ilmi vastauksesta.
- Johtopäätös väärin päin (Esim. ”Nollahypoteesi hylätään, mikäli $p > 0.01$”)
- Virhetyyppien nimien puuttuminen (Vaadittiin joko hylkäysvirhe/hyväksymisvirhe tai tyypin I/II virhe tai vastaavat englanninkieliset nimitykset).
- Hylkäys- ja hyväksymisvirheet selitettyinä väärin päin. Pisteitä ei kuitenkaan vähennetty, jos puhuttiin tyypin I ja II virheistä, mutta ei sattunut muistamaan kumpi tarkoitti hylkäys- ja kumpi hyväksymisvirhettä.

Yhtäsuuruuksien käyttämistä epäyhtälöissä toisin päin ei pidetty virheenä. Esim. oli yhtä oikein sanoa, että nollahypoteesi hylätään, mikäli $p \leq 0.01$ ja hyväksytään, mikäli $p > 0.01$.

Lähtökohtaisesti täydet pisteet c-kohdasta sai, jos oli menetelty oikein päätöksenteon kanssa nollahypoteeesin hylkäämisestä ja hyväksymisestä (kummalla tahansa tavalla) sekä todettu mahdolliset virheet ja näiden nimitykset.

Tehtävä oli osattu kohtuullisen hyvin, vaikka ihan täysiin pisteisiin ylsi vain muutama. Kompastuskiveksi monelle osoittautui b-kohta, jossa tehtävänanto saatettiin ohittaa jopa täysin. Toisaalta taas harmillisen monella saattoi olla otoskeskihajonnan kaavassa jokin pieni virhe, vaikka tehtävä olisi muuten osattu täydellisesti.

**4.** Esillä on kolme identtistä kulhoa, joista kussakin on viisi palloa:

- kulhossa yksi on viisi mustaa palloa.
- kulhossa kaksi on kaksi valkoista ja kolme mustaa palloa
- kulhossa kolme on neljä valkoista ja yksi musta pallo

Yksi kulhoista valitaan umpimähkään, mutta valitun kulhon numeroa $\theta$ (joka on 1, 2 tai 3) ei paljasteta, vaan se on tuntematon parametri. Valitusta kulhosta nostetaan umpimähkään kaksi palloa siten, että ensiksi nostettu pallo palautetaan kulhoon ennen toista nostoa. Molemmat nostetut pallot ovat valkoisia. Esitä parametrin $\theta$ priorijakauma, uskottavuusfunktio ja posteriorijakauma (niiden arvot luettelemalla).

**Ratkaisu:** //

Koska kaikkien kulhojen valinta on yhtä todennäköistä, priorijakauma $p(\theta) = \frac{1}{3}$, kun $\theta = 1, 2, 3$. Kokeessa havaittiin kaksi valkoista palloa kahdesta nostosta, merkitään $A =$ ”$kaksivalkoistahavaittu$”. Tätä havaintoa vastaava uskottavuusfunktio on

$$f(A|\theta) = p(A|\theta) = \begin{cases} 0^2 = 0, \text{ kun } \theta = 1, \\ (\frac{2}{5})^2 = \frac{4}{25}, \text{ kun } \theta = 2, \\ (\frac{4}{5})^2 = \frac{16}{25}, \text{ kun } \theta = 3. \end{cases}$$

Kokonaistodennäköisyydeksi saadaan

$$f(A) = p(1)f(A|1) + p(2)f(A|2) + p(3)f(A|3) = \frac{1}{3}(0 + \frac{4}{25} + \frac{16}{25}),$$

ja lopulta saamme posteriorijakaumaksi

$$p(\theta|A) = \frac{p(\theta)f(A\theta)}{f(A)} = \frac{f(A|\theta)}{\frac{20}{25}} = \begin{cases} 0, \text{ kun } \theta = 1, \\ \frac{4}{20} = \frac{1}{5}, \text{ kun } \theta = 2, \\ \frac{16}{20} = \frac{4}{5}, \text{ kun } \theta = 3. \end{cases}$$

**Pisteytyksistä:**

Karkeasti arvioituna tehtävästä sai priorin osaamisella yhden pisteen, uskottavuusfunktion osaamisella kaksi pistettä, ja posteriorin laskemisella kolme pistettä. Jos posteriorin kanssa oli ongelmia, annoin pisteitä seuraavasti: kokonaistodennäköisyys yksi piste, Bayesin kaava yksi piste, viimeinen piste järkevästä vastauksesta, tai jotain näiden järkeviä yhdistelmiä.

Jos priorin tai uskottavuusfunktion laskemisessa oli ongelmia, mutta posteriorijakauma osattiin näiden perusteella laskea, annoin posteriorin laskusta kolme pistettä (mikäli posteriori muuten järkevä).

Jos laskuissa ei ollut mainittu yhtäkään välivaihetta esimerkiksi posteriorissa, otin kaksi pistettä pois. "Oikean"vastauksen voi saada yllättävän vahingossa. Jos priori- ja uskottavuusfunktioita ei ollut erikseen tunnistettu tai ne oli tunnistettu virheellisesti, otin pisteitä pois virheen suuruuden mukaisesti.

Lisäksi virhetilanteissa otin surutta lisäpisteitä pois jos virheen seurauksena esimerkiksi posteriorijakauma sai arvoksi jotain suurempaa kuin yksi, tai ei summautunut ykköseksi.

Tehtävä oli osattu todella hyvin! Papereista jäi hyvä fiilis, mutta muutamaa virhettä esiintyi. Osa ei muistanut Bayesin kaavaa, yllättävän moni laski uskottavuusfunktion viidellä nostolla (kaksi valkoista, kolme mustaa) ja murto-osalla oli priorin ja uskottavuusfunktion käsitteet sekaisin. Näistä huolimatta (luultavasti) valtaosa sai täydet pisteet.