

Tilastomenetelmien peruskurssi (TILP150)

Syksy 2006

5. harjoitukset/viikko 41.

Näissä harjoituksissa käsitellään Lukua 8, erityisesti 8.2 (1.-5. tehtävät) ja Lukua 8.3 (6.-7. tehtävät).

1. ”Teoriatehtävä”: Yksi syy, että normaalijakauma on tärkeä, on se, että se kuvaa, miten keskiarvo vaihtelee otoksesta toiseen siirryttäessä.

Isommat eläimet näyttävät kantavan pidemmän aikaa jälkeläisiään. Tarkastellaan seuraavaksi hevosen raskauden pituutta, joka noudattaa likimain normaalijakaumaa odotusarvolla $\mu = 336$ päivää ja keskihajonnalla $\sigma = 3$ päivää (Moore & McCabe, 2006, teht. 1.87). Merkitään $X \simeq N(336, 3^2)$. Tällöin keskiarvo $\bar{X} \simeq N(336, 3^2/n)$.

- a) Millä välillä keskiarvoista on 95 %, kun otoskoko on $n = 625$?
- b) Millä välillä keskiarvoista on 95 %, kun otoskoko $n = 81$?

2. ”Teoriatehtävä”: Haluat vuokrata kalustamattoman yksión ensi kevääksi. Merkitään X = ”vuokra kuukaudessa (e)”. Sanomalehti-ilmoituksista tehdyn satunnaisotoksen ($n = 30$) perusteella vuokrien keskiarvo on 270 euroa eräällä alueella.

Oletetaan, että vuokran varianssi on tunnettu, ts. $\sigma^2 = 45^2 = 2025$ euroa. Oletetaan $\bar{X}$:n noudattavan likimain normaalijakaumaa $N(\mu, 45^2/30)$, koska otoskoko on ”suuri” eli $n \geq 30$. Arvioi odotusarvoa μ ko. alueella

- a) keskiarvolla eli piste-estimaatilla.
- b) 95 %:n luottamusvälillä.
- c) 99 %:n luottamusvälillä.
- d) Miten valittu riski vaikuttaa luottamusväliin? Vertaile kohtia b) ja c).

3. Tarkastellaan systolista verenpainetta (mmHg) seuraavan aineiston avulla ($n=5$): 125, 134, 138, 141, 145 (Zar, 1999, mukailtu esim. 6.5).

- a) Laske keskiarvo ja keskiarvon keskivirheen estimaatti tai lyhyesti keskiarvon keskivirhe.
- b) Mieti, mitä edellä lasketulla keskiarvon keskivirheellä arvioidaan?

4. Jatkoa edelliseen tehtävään. Oletetaan, että systolinen verenpaine X noudattaa normaalijakaumaa $N(\mu, \sigma^2)$. Varianssi σ^2 on tuntematon. Arvioi edellisen tehtävän aineiston avulla systolisen verenpaineen odotusarvoa μ

a) keskiarvolla eli piste-estimaatilla.

b) 95 %:n luottamusvälillä.

Vinkki: Kohdassa b) luottamusvälin laskemiseksi tarvitset lukua (fraktiilia) $t_{0.05/2}(4) = 2.776$. Se saadaan t -jakaumataulukosta: $f = 4$ (neljäs rivi) ja kolmas sarake (luvun 2.776 oikealla puolella oleva pinta-ala on 0.025 eli 2.5 %).

5. Teoriatehtävä: Tarkastellaan $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$.

Ratkaise a seuraavasta

$$\mathbb{P}(a \leq \bar{X}) = 0.975.$$

Vinkki: Tee muunnos $Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$ ja käytä tietoa, että $\mathbb{P}(-1.96 \leq Z) = 0.975$.

6. ”Teoriatehtävä”: Toinen syy, että normaalijakauma on tärkeä, on se, että se kuvaa, miten esim. puolueen kannatusosuus mielipidekyselyssä vaihtelee, jos mielipidekysely tehtäisiin useita kertoja.

Noin 40 % amerikkalaisista sanoo pelkäävänsä ulosmenoa iltaisin mahdollisen rikoksen takia. Ts. oletetaan, että $\pi = 0.40$ (pelkäävien osuus). Arvotaan amerikkalaisista useita otoksia, joissa kussakin on 1050 henkeä. Pelkäävien osuuden estimaatit noudattavat likimain normaalijakaumaa, jonka odotusarvo on 0.40 ja varianssi 0.015^2 . Ts.

$$\mathcal{P} \simeq N(0.40, 0.015^2).$$

Totea, että ehdot $n\pi > 5$ ja $n(1 - \pi) > 5$ ovat voimassa. Millä välillä suhteellisen osuuden estimaateista on 95 % (käytä luentomonisteen tarkempaa sääntöä)? (Lähde: Moore & McCabe, 2006, teht. 1.88).

7. Tutkitaan erään siemenlajikkeen itävyysprosenttia.

a) Otoksessa on 1500 ($= n$) siementä, joista iti 930. Laske 95 %:n luottamusväli suhteelliselle osuudelle π .

b) Otoksessa on 3000 ($= n$) siementä, joista iti 1860. Laske 95 %:n luottamusväli suhteelliselle osuudelle π .

c) Miten otoskoko vaikuttaa luottamusväliin? Vertaile kohtia a) ja b).

Demohyvytys: 6 tehtävää=0.75 pistettä, 3 tehtävää=0.35 pistettä