HARJOITUS 12

1. Merk. $X = \begin{cases} 1, & \text{kun kukansiemen itää} \\ 0, & \text{-"- -"- ei idä} \end{cases}$

$$X \sim \text{Bern}(\overset{\pi}{\overset{\shortparallel}{0.82}}) \quad \text{ja} \quad P(X=1) = 0.82$$

$(X_1, X_2, \ldots, X_{1000})$ on satunnaisotos $\text{Bern}(\overset{\pi}{\overset{\shortparallel}{0.82}})$-jakaumasta

$$P = \frac{\sum_{i=1}^{n} X_i}{n} \sim N(\pi, \tfrac{1}{n}\pi(1-\pi)) \text{ likimain}$$ (LUENTOMONISTE, SIVU 94)

Nyt siis

$$P = \frac{\sum_{i=1}^{1000} X_i}{1000} \sim N(0.82, \tfrac{1}{1000} \cdot 0.82 \cdot (1-0.82)) \text{ likimain}$$

$$\approx N(0.82, \sqrt{0.0001476}^2)$$

$$P(P > 0.80) = P(\underbrace{\frac{P - 0.82}{\sqrt{0.0001476}}}_{= Z \sim N(0,1)} > \frac{0.80 - 0.82}{\sqrt{0.0001476}})$$

$$= P(Z > -1.65)$$

$$= 1 - P(Z > 1.65) = 1 - 0.0495 = \underline{\underline{0.9505}}$$

2. (X_1, X_2) on satunnaisotos jakaumasta, jonka odotusarvo on μ ja varianssi on σ^2.

$\Rightarrow$ $E(X_1) = \mu$, $E(X_2) = \mu$, $D^2(X_1) = \sigma^2$, $D^2(X_2) = \sigma^2$ ja X_1 ja X_2 ovat riippumattomia toisistaan ($X_1 \perp\!\!\!\perp X_2$).

a) $$E(U) = E(\tfrac{1}{2}X_1 + \tfrac{1}{2}X_2) = E(\tfrac{1}{2}X_1) + E(\tfrac{1}{2}X_2)$$
$$= \tfrac{1}{2}E(X_1) + \tfrac{1}{2}E(X_2)$$
$$= \tfrac{1}{2}\cdot\mu + \tfrac{1}{2}\cdot\mu = \underline{\underline{\mu}}$$

$$E(V) = E(\tfrac{1}{4}X_1 + \tfrac{3}{4}X_2) = E(\tfrac{1}{4}X_1) + E(\tfrac{3}{4}X_2)$$
$$= \tfrac{1}{4}E(X_1) + \tfrac{3}{4}E(X_2)$$
$$= \tfrac{1}{4}\cdot\mu + \tfrac{3}{4}\cdot\mu = \underline{\underline{\mu}}$$

$$E(Y) = E(\tfrac{1}{3}X_1 + \tfrac{2}{3}X_2) = E(\tfrac{1}{3}X_1 + \tfrac{2}{3}X_2) = \tfrac{1}{3}E(X_1) + \tfrac{2}{3}E(X_2)$$
$$= \tfrac{1}{3}\cdot\mu + \tfrac{2}{3}\cdot\mu = \underline{\underline{\mu}}$$

Estimaattori on harhaton, jos sen odotusarvo = estimoitava parametri

Nyt siis kaikki kolme estimaattoria ovat harhattomia μ:n estimaattoreita.

b) $D^2(U) = D^2(\frac{1}{2}X_1 + \frac{1}{2}X_2) = D^2(\frac{1}{2}X_1) + D^2(\frac{1}{2}X_2)$

$= (\frac{1}{2})^2 D^2(X_1) + (\frac{1}{2})^2 \cdot D^2(X_2)$

$= \frac{1}{4}\cdot\sigma^2 + \frac{1}{4}\cdot\sigma^2 = \underline{\underline{\sigma^2/2}}$

$D^2(V) = D^2(\frac{1}{4}X_1 + \frac{3}{4}X_2) = D^2(\frac{1}{4}X_1) + D^2(\frac{3}{4}X_2)$

$= (\frac{1}{4})^2 D^2(X_1) + (\frac{3}{4})^2 \cdot D^2(X_2)$

$= \frac{1}{16}\cdot\sigma^2 + \frac{9}{16}\cdot\sigma^2 = \frac{10}{16}\cdot\sigma^2 = \underline{\underline{\frac{5}{8}\sigma^2}}$

$D^2(Y) = D^2(\frac{1}{3}X_1 + \frac{2}{3}X_2) = D^2(\frac{1}{3}X_1) + D^2(\frac{2}{3}X_2)$

$= (\frac{1}{3})^2 D^2(X_1) + (\frac{2}{3})^2 D^2(X_2)$

$= \frac{1}{9}\sigma^2 + \frac{4}{9}\sigma^2 = \underline{\underline{\frac{5}{9}\sigma^2}}$

$\Rightarrow$ tehokkain estimaattori on $\underline{\underline{U}}$.

c) a) & b) -kohtien perusteella μ:n estimoinnissa kannattaa käyttää estimaattoria U:

$$\hat{\mu} = \frac{1}{2}\cdot 5 + \frac{1}{2}\cdot 10 = \underline{\underline{7.5}}$$

3. X = viipymisaika minuuteissa, $X \sim N(\mu, \sigma^2)$ ← Huom. σ^2 tunnetaan!

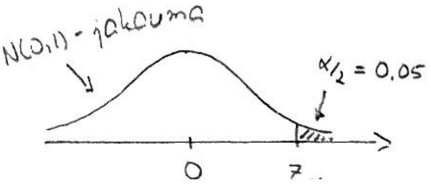
$\bar{X} = \frac{1}{16}\sum_{i=1}^{16} X_i = 25$

a) $\hat{\mu} = \bar{X} = \underline{\underline{25}}$ min

b) b1) $100(1-\alpha) = 90 \Leftrightarrow 100 - 100\alpha = 90 \Leftrightarrow 100\alpha = 10 \Rightarrow \alpha = 0.10$ ja $\alpha/2 = 0.05$

$\Rightarrow z_{0.05} = 1.64$ (tai 1.65 tai 1.645)

μ:n 90 %:n lv. on $(\bar{X} - z_{0.05}\cdot\sigma/\sqrt{n},\ \bar{X} + z_{0.05}\cdot\sigma/\sqrt{n})$



$\Leftrightarrow (25 - 1.64\cdot 6/\sqrt{16},\ 25 + 1.64\cdot 6/\sqrt{16})$

$\Leftrightarrow \underline{\underline{(22.54, 27.46)}}$

TULKINTA: Tuntematon parametri μ (= kaikkien asiakkaiden keskimääräinen viipymisaika liikkeessä) kuuluu 90 % varmuudella välille (22.54 min, 27.46 min).

b2) $100(1-\alpha) = 95 \Rightarrow \alpha = 0.05 \Rightarrow \alpha/2 = 0.025 \Rightarrow z_{\alpha/2} = z_{0.025} = 1.96$

kysytty μ:n lv.: $(\bar{X} - z_{0.025} \cdot \sigma/\sqrt{n}\,,\ \bar{X} + z_{0.025} \cdot \sigma/\sqrt{n})$

$\Leftrightarrow (25 - 1.96 \cdot 6/\sqrt{16}\,,\ 25 + 1.96 \cdot 6/\sqrt{16})$

$\Leftrightarrow \underline{\underline{(22.06, 27.94)}}$

TULKINTA: μ kuuluu 95 %:n varmuudella välille (22.06 min, 27.94 min.)

c) virhemarginaali $d \geq z_{\alpha/2} \cdot \sigma/\sqrt{n} \quad | \cdot \sqrt{n}$

$\Leftrightarrow \sqrt{n}\, d \geq z_{\alpha/2} \cdot \sigma \quad | : d$

$\Leftrightarrow \sqrt{n} \geq \dfrac{z_{\alpha/2} \cdot \sigma}{d} \quad | ()^2$

$\Leftrightarrow n \geq \left(\dfrac{z_{\alpha/2} \cdot \sigma}{d}\right)^2$

c1) $n \geq \left(\dfrac{1.64 \cdot 6}{2}\right)^2 \Leftrightarrow n \geq 24.2 \Rightarrow$ otoskoko vähintään $\underline{\underline{25}}$

c2) $n \geq \left(\dfrac{1.96 \cdot 6}{2}\right)^2 \Leftrightarrow n \geq 34.6 \Rightarrow$ otoskoko vähintään $\underline{\underline{35}}$

4. $X \sim N(\mu, 17^2)$ ja $\bar{X} = 123$

a) 90 %:n lv. μ:lle: $100(1-\alpha) = 90 \Rightarrow \alpha = 0.10$ ja $\alpha/2 = 0.05 \Rightarrow z_{\alpha/2} = z_{0.05} = 1.64$

$\Rightarrow (\bar{X} - z_{0.05} \cdot \sigma/\sqrt{n}\,,\ \bar{X} + z_{0.05} \cdot \sigma/\sqrt{n})$

a1) $(123 - 1.64 \cdot 17/\sqrt{9}\,,\ 123 + 1.64 \cdot 17/\sqrt{9}) \Leftrightarrow \underline{\underline{(113.71, 132.29)}}$

a2) $(123 - 1.64 \cdot 17/\sqrt{36}\,,\ 123 + 1.64 \cdot 17/\sqrt{36}) \Leftrightarrow \underline{\underline{(118.35, 127.65)}}$

a3) $(123 - 1.64 \cdot 17/\sqrt{100}\,,\ 123 + 1.64 \cdot 17/\sqrt{100}) \Leftrightarrow \underline{\underline{(120.21, 125.79)}}$

b) Mitä suurempi otoskoko sitä pienempi virhemarginaali (ja kapeampi

c) $100(1-\alpha) = 95 \Rightarrow \alpha = 0.05$ ja $\alpha/2 = 0.025 \Rightarrow z_{\alpha/2} = z_{0.025} = 1.96$

$$\Rightarrow \left(\bar{X} - z_{0.025} \cdot \sigma/\sqrt{n}\,,\ \bar{X} + z_{0.025} \cdot \sigma/\sqrt{n}\right)$$

$$\Leftrightarrow \left(123 - 1.96 \cdot 17/\sqrt{9}\,,\ 123 + 1.96 \cdot 17/\sqrt{9}\right) \Leftrightarrow \underline{\underline{(111.89,\ 134.11)}}$$

Mitä suurempi luottamustaso sitä suurempi virhemarginaali (ja leveämpi luottamusväli).

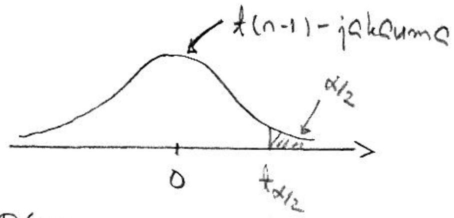
(5.) Merk. X = ajoaika paikasta A paikkaan B (minuuteissa)

$X \sim N(\mu, \sigma^2)$, missä μ ja σ^2 ovat tuntemattomia

a) $\hat{\mu} = \bar{X} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{10}(49 + 52 + \ldots + 51) = \underline{\underline{51.1}}$ min.

b) μ:n $100(1-\alpha)$ %:n luottamusväli on $\left(\bar{X} - t_{\alpha/2} \cdot S/\sqrt{n}\,,\ \bar{X} + t_{\alpha/2} \cdot S/\sqrt{n}\right)$

laskimesta: $S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2} = \ldots \approx 3.1429$



b1) $100(1-\alpha) = 95 \Rightarrow \alpha = 0.05$ ja $\alpha/2 = 0.025$

$t(n-1) = t(10-1) = t(9)$-jakauman taulukosta: $P(T \geq 2.262) = 0.025$

$$\Rightarrow t_{\alpha/2} = t_{0.025} = 2.262$$

$\Rightarrow$ kysytty lv. on: $\left(51.1 - 2.262 \cdot 3.1429/\sqrt{10}\,,\ 51.1 + 2.262 \cdot 3.1429/\sqrt{10}\right)$

$$\Leftrightarrow \underline{\underline{(48.85,\ 53.35)}}$$

lv:n tulkinta: tuntematon parametri μ eli keskimääräinen ajoaika paikasta A paikkaan B kuuluu 95 % varmuudella välille (48.85 min, 53.35 min).

b2) $100(1-\alpha) = 99 \Rightarrow \alpha = 0.01$ ja $\alpha/2 = 0.005$

$t(9)$-jakauman taulukko: $P(T \geq 3.250) = 0.005 \Rightarrow t_{\alpha/2} = t_{0.005} = 3.250$

$\Rightarrow$ kysytty lv. on: $\left(51.1 - 3.250 \cdot 3.1429/\sqrt{10}\,,\ 51.1 + 3.250 \cdot 3.1429/\sqrt{10}\right)$

$$\Leftrightarrow \underline{\underline{(47.87,\ 54.33)}}$$

Tulkinta: keskimääräinen ajoaika paikasta A paikkaan B kuuluu 99 % varmuudella välille (47.87 min, 54.33 min).

Koska 50 min kuuluu laskettujen luottamusvälien sisään, on yrityksen ilmoitus aineiston perusteella uskottava.

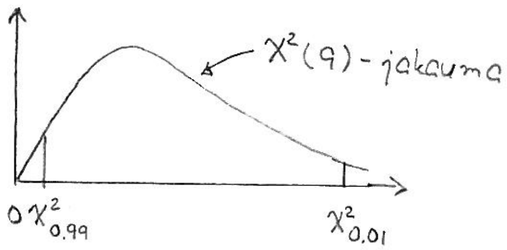
c) c1) $\hat{\sigma} = s \approx \underline{\underline{3,143}}$ min

c2) $100(1-\alpha) = 98 \Rightarrow \alpha = 0.02 \Rightarrow \alpha/2 = 0.01$ ja $1-\alpha/2 = 0,99$

$\chi^2(n-1) = \chi^2(10-1) = \chi^2(9)$ - jakauman taulukosta:

$P(\chi^2 \geq 21.67) = 0.01 \Rightarrow \chi^2_{\alpha/2} = \chi^2_{0.01} = 21.67$

$P(\chi^2 \geq 2.09) = 0,99 \Rightarrow \chi^2_{1-\alpha/2} = \chi^2_{0.99} = 2.09$



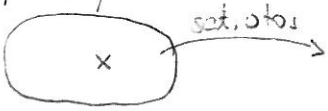
σ^2:n 98 %:n lv. on $\left(\dfrac{(n-1)s^2}{\chi^2_{0.01}}, \dfrac{(n-1)s^2}{\chi^2_{0.99}} \right)$

$\Leftrightarrow \left(\dfrac{(10-1)\cdot 3.143^2}{21.67}, \dfrac{(10-1)\cdot 3.143^2}{2.09} \right) \Leftrightarrow (4.1027, 42.5388)$

$\Rightarrow \sigma$:n 98 %:n lv. on $(\sqrt{4.1027}, \sqrt{42.5388}) \Leftrightarrow \underline{\underline{(2.025, 6.522)}}$

TULKINTA: tuntematon parametri σ (= ajoaikojen keskihajonta) kuuluu 98 %:n varmuudella välille (2.03 min, 6.52 min).

6. a) perusjoukko



$(X_1, \ldots, X_{800})$

$X_i = \begin{cases} 1, & \text{kun matkustajalla } i \text{ on opiskelijalippu} \\ 0, & \text{—"— } i \text{ ei ole opiskelijalippua} \end{cases}$

$X_i \sim \text{Bern}(\pi)$ ja $P(X_i = 1) = \pi \quad \forall\, i = 1, \ldots, 800$

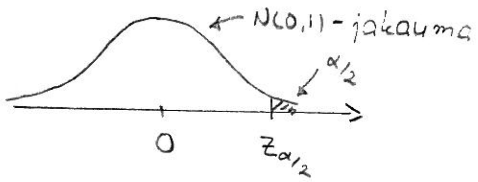
Merk. P = opiskelijalipulla matkustavien suhteellinen osuus otoksessa

a1) $\hat{\pi} = P = 440/800 = \underline{\underline{0,55}}$

a2) π:n likimääräinen $100(1-\alpha)$ %:n lv. on $\left(P - z_{\alpha/2}\sqrt{P(1-P)/n}, P + z_{\alpha/2}\sqrt{P(1-P)/n} \right)$

Nyt $100(1-\alpha) = 99 \Rightarrow \alpha = 0,01$ ja $\alpha/2 = 0,005$

$\Rightarrow z_{\alpha/2} = z_{0,005} = 2,58$



$\Rightarrow$ kysytty lv. on $\left(0.55 - 2.58 \cdot \sqrt{\dfrac{0.55\cdot(1-0,55)}{800}}, 0.55 + 2.58 \cdot \sqrt{\dfrac{0.55\cdot(1-0.55)}{800}} \right)$

$\Leftrightarrow \underline{\underline{(0,505, 0.595)}}$

TULKINTA: Tuntematon parametri π (= opiskelijalipulla matkustavien suhteellinen osuus kaikkien matkustajien joukossa linjalla XX) on 99 % varmuudella välillä (0.505, 0.595)

b) $100(1-\alpha) = 95 \Rightarrow \alpha = 0.05$ ja $\alpha/2 = 0.025 \Rightarrow z_{\alpha/2} = z_{0.025} = 1.96$

lv:n pituus: $2 z_{\alpha/2} \cdot \sqrt{\frac{p(1-p)}{n}} \leq 0.05 \quad | : 2 z_{\alpha/2}$

$$\Leftrightarrow \sqrt{\frac{p(1-p)}{n}} \leq \frac{0.05}{2 z_{\alpha/2}} \quad | ()^2$$

$$\Leftrightarrow \frac{p(1-p)}{n} \leq \left(\frac{0.05}{2 z_{\alpha/2}}\right)^2 \quad | \cdot n$$

$$\Leftrightarrow p(1-p) \leq \left(\frac{0.05}{2 z_{\alpha/2}}\right)^2 \cdot n \quad | : \left(\frac{0.05}{2 z_{\alpha/2}}\right)^2$$

$$\Rightarrow n \geq \frac{p(1-p)}{\left(\frac{0.05}{2 z_{\alpha/2}}\right)^2} = \frac{0.5 \cdot (1-0.5)}{\left(\frac{0.05}{2 \cdot 1.96}\right)^2} = 1536.64$$

$\Rightarrow$ otoskoon tulisi olla vähintään $\underline{\underline{1537}}$

7. a) $X \sim \underline{\underline{Bin(5, 0.5)}}$

b) $P(X<1) = P(X=0) = \binom{5}{0} \cdot 0.5^0 \cdot (1-0.5)^{5-0} = 0.5^5 = \underline{\underline{0.03125}}$

c) Merkitsevyystestaus:

1) Oletetaan, että $X \sim Bin(5, p)$

2) Hypoteesit: $\begin{cases} H_0: p = 0.5 & \text{"Jarkko arvaa"} \\ H_1: p > 0.5 & \text{"Jarkko tietää"} \end{cases}$

3) Testisuure $X \sim Bin(5, 0.5)$, kun H_0 on tosi

4) Testisuureen havaittu arvo $x = 4$

5) p-arvo $= P(X \geq 4 \mid H_0 \text{ tosi})$

$$= P(X=4 \mid H_0) + P(X=5 \mid H_0)$$

$$= \binom{5}{4} \cdot 0.5^4 \cdot (1-0.5)^{5-4} + \binom{5}{5} \cdot 0.5^5 \cdot (1-0.5)^{5-5}$$

$$= 0.15625 + 0.03125 = \underline{\underline{0.1875}}$$

($\Rightarrow$ 6) johtopäätös: Aineisto on sopusoinnussa H_0:n kanssa, ts. Jarkko näyttää arvanneen. Mutta: Oikeammin tulkittuna aineiston pienuuden takia johtopäätöksiä ei juurikaan kannata tehdä!)