Harjoitus 7

1. $X_i = i$:nnen henkilön massa, $\mu = 78, \sigma = 20$.

n:n henkilön kokonaismassa $S = \sum_{i=1}^{n} X_i \sim N\left(n\mu, n\sigma^2\right) \quad \Rightarrow Z = \frac{S-n\mu}{\sqrt{n}\,\sigma} \sim N(0,1)$

a) $n = 100$

$P(S > 10000) = 1 - P\left(\frac{S-n\mu}{\sqrt{n}\,\sigma} \leq \frac{10000-n\mu}{\sqrt{n}\,\sigma}\right)$

$= 1 - P\left(Z \leq \frac{10000-n\mu}{\sqrt{n}\,\sigma}\right) = 1 - \Phi\left(\frac{10000-n\mu}{\sqrt{n}\,\sigma}\right) = 1 - \Phi(11) \approx 0.$

b) $P(S > 10000) =$ VERTAA KOHTA a) $= 1 - \Phi\left(\frac{10000-n\mu}{\sqrt{n}\,\sigma}\right) \leq 0,01$

$\Rightarrow \Phi\left(\frac{10000-n\mu}{\sqrt{n}\,\sigma}\right) \geq 0,99 \overset{\text{taulukosta 1}}{=} \Phi(2,33)$

Kertymäfunktio $\Phi(x)$ on kasvava funktio $\Rightarrow \frac{10000-n\mu}{\sqrt{n}\sigma} \geq 2,33 \quad \Rightarrow$

$n \cdot \mu + 2,33 \cdot \sigma \cdot \sqrt{n} - 10000 \leq 0$

$\sqrt{n} = \begin{cases} 11,02 \\ (-11,62) \end{cases} \quad \Rightarrow n = 121,62 \ \Rightarrow \ n \leq 121.$

2. Suhteellisen osuuden (tapahtumatodennäköisyyden/prosenttiluvun) p luottamusväli.

$p = P$(rintasyöpädiagnoosin saanut kuolee).

$X =$ rintasyöpään kuolleiden lkm n:n diagnoosin saaneen joukossa; $X \sim Bin(n,p)$.

$\frac{X}{n} \overset{\text{likim.}}{\sim} N(p, \frac{p(1-p)}{n})$; $\quad p$:n 95%:n luottamusväli ($\alpha = 0,05$):

$P(-z_{\frac{\alpha}{2}} \leq \frac{\frac{X}{n}-p}{\sqrt{\frac{p(1-p)}{n}}} \leq z_{\frac{\alpha}{2}}) = 1 - \alpha \quad \Rightarrow \quad \hat{p} - z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \leq p \leq \hat{p} + z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

Sijoitetaan lukuarvot $\hat{p} = \frac{X}{n} = \frac{6675}{28028}$; $\quad z_{\frac{\alpha}{2}} = 1,96$ (arvo taulukosta 1)

$\Rightarrow \quad 0,233 \leq p \leq 0,243.$

3. Odotusarvon μ luottamusväli

$X_i =$ Näytteen i lämpölaajenemiskerroin; $X_i \sim N(\mu, \sigma^2), (\sigma$ tuntematon$)$

Teoria: $\frac{\overline{X}-\mu}{\frac{s}{\sqrt{n}}} \sim t, n-1$

Havainnoista saadaan $\overline{X} = 1,0076383333$ ja $s = 1,459 \cdot 10^{-3}$.

μ:n 95%:n luottamusväli:

$P(-t_p \leq \frac{\overline{X}-\mu}{\frac{s}{\sqrt{n}}} \leq t_p) = 0,95.$

$t_p = 2,201$ (arvo taulukosta 2; vapausasteet $f = n - 1 = 11$).

Joten $-2,201 \leq \frac{\overline{X}-\mu}{\frac{s}{\sqrt{n}}} \leq 2,201$

$\Leftrightarrow \overline{X} - 2,201\frac{s}{\sqrt{n}} \leq \mu \leq \overline{X} + 2,201\frac{s}{\sqrt{n}}$.

Sijoitetaan havainnoista saadut arvot $\Rightarrow \mu \in [1,00671; 1,00857]$.

4. $X_i = i$:s kuivumisajan mittaustulos; $X_i \sim N(\mu, \sigma^2)$, ($\sigma$ tuntematon)

Teoria: $\frac{\overline{X}-\mu}{\frac{s}{\sqrt{n}}} \sim t, n-1$

Havainnoista saadaan $\overline{X} = 3,79 \quad s = 0,9709$ ja $n = 15$.

a) Todellisen kuivumisajan μ 98%:n luottamusväli saadaan ehdosta

$P(-t_p \leq \frac{\overline{X}-\mu}{\frac{s}{\sqrt{n}}} \leq t_p) = 0,98$.

Taulukosta 2 saadaan $t_p = 2,624$ (vapausasteet $f = 15-1 = 14$),

joten $-2,624 \leq \frac{\overline{X}-\mu}{\frac{s}{\sqrt{n}}} \leq 2,624$

$\Leftrightarrow \overline{X} - 2,624\frac{s}{\sqrt{n}} \leq \mu \leq \overline{X} + 2,624\frac{s}{\sqrt{n}}$.

Sijoitetaan havainnoista saadut arvot $\Rightarrow \mu \in [3,13; 4,44]$.

b) Virhe on itseisarvoltaan korkeintaan $2,624\frac{s}{\sqrt{n}} = 0,66$ 98 %:n varmuudella.

5. $H(1) = 100e^{A+\sigma W(1)} = 100e^A \cdot e^{\sigma Z}$,

missä $A = b - \frac{1}{2}\sigma^2; \quad b = 0,1; \quad \sigma = 0,3$ ja $Z = W(1) \sim N(0,1)$.

a) $E(H(1)) = 100e^A E(e^{\sigma Z}) = 100e^A \int_{-\infty}^{\infty} e^{\sigma x}\varphi(x)\mathrm{d}x = 100e^A \int_{-\infty}^{\infty} e^{\sigma x}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\mathrm{d}x$

$= 100e^A \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(x^2-2\sigma x)}\mathrm{d}x = 100e^A \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(x^2-2\sigma x+\sigma^2-\sigma^2)}\mathrm{d}x$

$= 100e^A \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(x-\sigma)^2+\frac{1}{2}\sigma^2}\mathrm{d}x = 100e^{A+\frac{1}{2}\sigma^2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(x-\sigma)^2}\mathrm{d}x$

sijoitus $t = x - \sigma; \quad \mathrm{d}x = \mathrm{d}t$ (rajat säilyvät samoina)

$$= 100e^{A+\frac{1}{2}\sigma^2} \underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}t^2}\mathrm{d}t}_{=1} = 100e^b \approx 110,52.$$

b) $P(a_1 \leq 100e^A \cdot e^{\sigma Z} \leq a_2) = P(\frac{a_1}{100e^A} \leq e^{\sigma Z} \leq \frac{a_2}{100e^A})$

$= P\left(\frac{1}{\sigma}\ln\frac{a_1}{100e^A} \leq Z \leq \frac{1}{\sigma}\ln\frac{a_2}{100e^A}\right) = P(-r_0 \leq Z \leq r_0) = 0,95$

Taulukosta 1 saadaan $r_0 = 1,96$

$\Rightarrow a_1 = 100e^A e^{-\sigma r_0} \approx 58,68; \quad a_2 = 100e^A e^{\sigma r_0} \approx 190,22$