OULUN YLIOPISTO
Matemaattisten tieteiden laitos

TENTTIVASTAUSPAPERI

Pvm 24.04.-13

Opiskelija

Kuulustelija

Opiskelijanro tai henkilötunnus

Tentittävä opintojakso TILASTOTIETEEN PERUSMENETELMÄT I, 2. välikoe

Koulutusohjelma/aloitusvuosi

Tentin arvostelu

1	2	3	4	5	6	Σ

Opintojakson tyyppi P A S

1.

A)
$$\left.\begin{aligned} P(B^c) &= 8/20 = 0.4 \\ P(A) &= 13/20 = 0.65 \\ P(A|B) &= 6/12 = 0.5 = P(A^c|B) \\ P(A\cap B) &= 6/20 = 0.3 \\ P(A^c \cup B^c) &= \frac{7+6+1}{20} = 14/20 = 0.7 \\ P(A^c \cap B^c) &= 1/20 = 0.05 \end{aligned}\right\} \Rightarrow \underline{\underline{a5}}$$

B) $P(A) = 0.65$ ja $P(A|B) = 0.5$ eli $P(A) \neq P(A|B)$

$\Rightarrow$ A ja B eivät ole toisistaan riippumattomia

$P(A\cap B) = 0.3 \neq 0 \Rightarrow$ A ja B eivät ole toisensa poissulkevia

$\Rightarrow \underline{\underline{b6}}$

C)
$$\begin{aligned} F(1) &= P(X \leq 1) \\ &= P(X=0) + P(X=1) \\ &= \frac{4^0}{0!}e^{-4} + \frac{4^1}{1!}e^{-4} \\ &= e^{-4} + 4e^{-4} \\ &\approx 0.0183 + 0.0733 \approx 0.092 \Rightarrow \underline{\underline{c3}} \end{aligned}$$

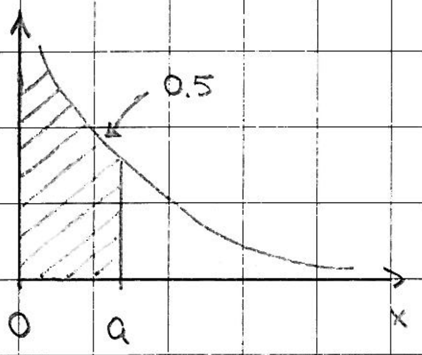
D) $X_i \sim \mathrm{Tas}[0,12]$ kaikilla $i = 1, \dots, 100$

$$\mu = E(X_i) = \frac{a+b}{2} = \frac{0+12}{2} = 6$$
$$\sigma^2 = D^2(X_i) = \frac{(b-a)^2}{12} = \frac{(12-0)^2}{12} = \frac{12^2}{12} = 12$$

$$\Rightarrow \bar{X} = \frac{1}{100}\sum_{i=1}^{100} X_i \overset{\text{likim.}}{\sim} N(\mu, \sigma^2/n) = N(6, 12/100)$$
$$= N(6, 0.12) \Rightarrow \underline{\underline{d6}}$$

E) $X \sim \mathrm{Exp}(\alpha) \Rightarrow P(X \le a) = 1 - e^{-\alpha x}$

Olkoon $\mathrm{Md}(X) = a \Rightarrow P(X \le a) = 0.5$



$$\Leftrightarrow 1 - e^{-2a} = 0.5 \quad \big| -1$$
$$\Leftrightarrow -e^{-2a} = -0.5 \quad \big| \cdot -1$$
$$\Leftrightarrow e^{-2a} = 0.5 \quad \big| \ln()$$
$$\Leftrightarrow -2a = \ln(0.5) \quad \big| : -2$$
$$\Leftrightarrow a = \frac{\ln(0.5)}{-2} \approx 0.347 \Rightarrow \underline{\underline{e2}}$$

F) $\underline{\underline{f1}}$

2. Merk. X = pesästä löytyvien munien lukumäärä

X:n tn-jakauma on:

x_i	0	1	2	3	4	Yht.
p_i	0.125	0.250	0.500	0.100	0.025	1

a) $P(X \leq 1) = P(X=0) + P(X=1)$
$= 0.125 + 0.250$
$= \underline{\underline{0.375}}$

b) Merk. Y = vähintään kolme munaa sisältävien pesien lukumäärä kahdeksan pesän joukossa

$Y \sim \text{Bin}(8, 0.125)$

$$P(Y \geq 3) = P(Y=3) + P(Y=4) + \dots + P(Y=8)$$
$$= 1 - P(Y \leq 2)$$
$$= 1 - (P(Y=0) + P(Y=1) + P(Y=2))$$
$$= 1 - \left(\binom{8}{0} \cdot 0.125^0 \cdot (1-0.125)^8 + \binom{8}{1} \cdot 0.125^1 \cdot (1-0.125)^7 + \binom{8}{2} \cdot 0.125^2 \cdot (1-0.125\right.$$
$$\approx 1 - (0.3436 + 0.3927 + 0.1963)$$
$$= \underline{\underline{0.0674}}$$

c) Merk. A = "maanantaina 1-2 munaa"
B = "tiistaina korkeintaan kahdessa vähintään 3 munaa"

$$P(A) = P(X=1) + P(X=2) = 0.250 + 0.500 = 0.750$$
$$P(B) = P(Y \leq 2) = 1 - P(Y \geq 3) \overset{a2)}{=} 1 - 0.0674 = 0.9326$$

$$P(A \cap B) = P(A)P(B)$$
$$= 0.750 \cdot 0.9326$$
$$\approx \underline{\underline{0.6995}}$$

b) $\sigma^2 = D^2(X) = \sum_{i=1}^{k} x_i^2 p_i - \mu^2$

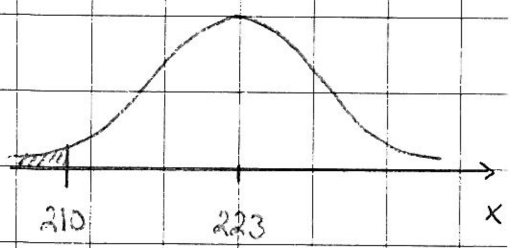
$$\mu = \sum_{i=1}^{k} x_i p_i = 0 \cdot 0.125 + 1 \cdot 0.250 + 2 \cdot 0.500 + 3 \cdot 0.100 + 4 \cdot 0.025$$
$$= 1.65$$

$$\sigma^2 = (0^2 \cdot 0.125 + 1^2 \cdot 0.250 + 2^2 \cdot 0.500 + 3^2 \cdot 0.100 + 4^2 \cdot 0.025) - 1.65^2$$
$$= 0.8275 \quad \Rightarrow \quad \sigma = \sqrt{0.8275} \approx \underline{\underline{0.9097}}$$

3. Merk. X = laulujoutsenen siipien kärkivälin pituus (cm)

$X \sim N(223, 6^2)$

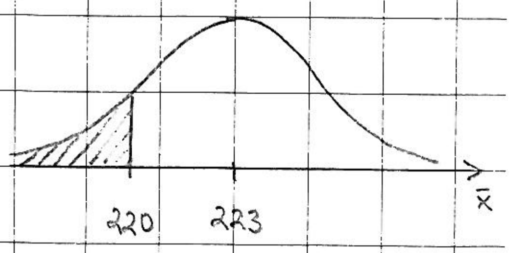
a) $$P(X \le 210) = P\Big(\underbrace{\frac{X-223}{6}}_{= Z \sim N(0,1)} \le \frac{210-223}{6}\Big)$$
$$= P(Z \le -2{,}17)$$
$$= P(Z \ge 2{,}17)$$
$$= \underline{\underline{0.0150}}$$



b) $(X_1, X_2, \ldots, X_8)$ on satunnaisotos $N(223, 6^2)$ - jakaumasta

$$\Rightarrow \bar{X} = \frac{1}{8}\sum_{i=1}^{8} X_i \sim N(223, 6^2/8) = N(223, \sqrt{4.5}^2)$$

$$P(\bar{X} \le 220) = P\Big(\underbrace{\frac{\bar{X}-223}{\sqrt{4.5}}}_{= Z \sim N(0,1)} \le \frac{220-223}{\sqrt{4.5}}\Big)$$
$$= P(Z \le -1{,}41)$$
$$= P(Z \ge 1{,}41) = \underline{\underline{0.0793}}$$



c) Merk. Y = kurjen siipien kärkivälin pituus (cm)

$Y \sim N(201, 7)$

$S = X - Y$

$$E(S) = E(X - Y) = E(X) - E(Y)$$
$$= 223 - 201$$
$$= 22$$

$$D^2(S) = D^2(X - Y) = D^2(X) + D^2(Y)$$
$$= 6^2 + 7^2$$
$$= 85$$

$$\Rightarrow S \sim N(22, \sqrt{85}^2)$$

$$P(S > 0) = P\Big(\underbrace{\frac{S - 22}{\sqrt{85}}}_{= Z \sim N(0,1)} > \frac{0 - 22}{\sqrt{85}}\Big)$$
$$= P(Z > -2{,}39)$$
$$= 1 - P(Z \leq -2.39)$$
$$= 1 - P(Z \geq 2{,}39)$$
$$= 1 - 0.0084$$
$$= \underline{\underline{0.9916}}$$

4. Merk. X = hapenottokyky (ml/kg/min, Perun alkuperäisväestö)

$(X_1, X_2, \ldots, X_{20})$ on satunnaisotos $N(\mu, \sigma^2)$-jakaumasta, missä μ ja σ^2 ovat tuntemattomia.

a) 95% luottamusväli μ:lle: $100(1-\alpha) = 95 \Rightarrow \alpha = 0.05$ ja $\alpha/2 = 0.025$

$$\left(\bar{X} - t_{\alpha/2} \, S/\sqrt{n},\ \bar{X} + t_{\alpha/2} \, S/\sqrt{n}\right)$$

$t(n-1) = t(20-1) = t(19)$-jakauman taulukko: $P(T \geq 2.093) = 0.025$

$$\Rightarrow t_{\alpha/2} = t_{0.025} = 2.093$$

$\Rightarrow$ Kysytty lv. on: $\left(46.3 - 2.093 \cdot 5.0/\sqrt{20},\ 46.3 + 2.093 \cdot 5.0/\sqrt{20}\right)$

$$\Leftrightarrow \underline{\underline{(43.96, 48.64)}}$$

b) 98% luottamusväli σ^2:lle: $100(1-\alpha) = 98 \Rightarrow \alpha = 0.02$

$$\Rightarrow \alpha/2 = 0.01 \text{ ja } 1-\alpha/2 = 0.99$$

$$\left(\frac{(n-1)S^2}{\chi^2_{\alpha/2}},\ \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}}\right)$$

$\chi^2(n-1) = \chi^2(20-1) = \chi^2(19)$-jakauman taulukko:

$$P(\chi^2 \geq 36.19) = 0.01 \Rightarrow \chi^2_{\alpha/2} = \chi^2_{0.01} = 36.19$$
$$P(\chi^2 \geq 7.63) = 0.99 \Rightarrow \chi^2_{1-\alpha/2} = \chi^2_{0.99} = 7.63$$

$\Rightarrow \sigma^2$:n 98% lv. on: $\left(\frac{(20-1)\cdot 5.0^2}{36.19},\ \frac{(20-1)\cdot 5.0^2}{7.63}\right) \Leftrightarrow (13.125, 62.254)$

$\Rightarrow \sigma$:n 98 % lv on: $(\sqrt{13.125}, \sqrt{62.254}) \Leftrightarrow \underline{\underline{(3.62, 7.89)}}$

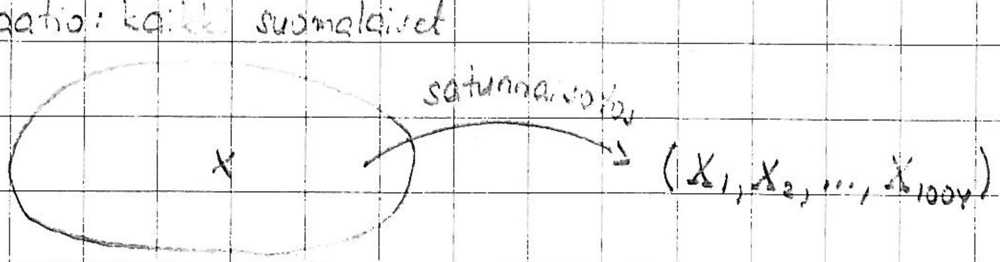
c) R:n tulostuksella on tutkittu sitä onko Perun alkuperäisväestön ja tasankoalueilla syntyneiden, mutta suuriin korkeuksiin sopeutuneiden amerikkalaisten keskimääräiset hapenottokyvyt yhtä suuret vai ei.

Nollahypoteesi $H_0: \mu_x = \mu_y$ (tai $\mu_x - \mu_y = 0$) eli ei eroa ryhmien välillä. (Testi on tehty kaksisuuntaisena eli vastahypoteesi on muotoa $H_1: \mu_x \neq \mu_y$).

Testin p-arvo = 0.0006772 eli aineisto on selvässä ristiriidassa H_0:in kanssa, H_1 on uskottavampi. Vertailtavien ryhmien välillä näyttäisi siis olevan eroa keskimääräisissä hapenottokyvyissä.

Tulostuksesta löytyvä 95 % luottamusväli on laskettu odotusarvojen erotukselle $\mu_x - \mu_y$ ja se on (3.62, 11.98). Nolla ei kuulu lasketun luottamusvälin sisään, joten H_0 ei ole uskottava. Perun alkuperäisväestöllä näyttäisi olevan keskimäärin korkeampi hapenottokyky.

5. 1) populaatio: kaikki suomalaiset



X = suhtautuminen nelosoluen myynnin vapauttamiseen, missä

$$X = \begin{cases} 1, & \text{kun henkilö kannattaa myynnin vapauttamista} \\ 0, & \text{–"– –"– ei kannata –"– –"–} \end{cases}$$

Oletetaan, että $X \sim Bern(\pi)$, missä $P(X=1) = \pi$.

2) Hypoteesit: $\begin{cases} H_0: \pi = 0.5 \\ H_1: \pi \neq 0.5 \end{cases}$

3) Testisuure $Z = \dfrac{P - \pi_0}{\sqrt{\dfrac{\pi_0(1-\pi_0)}{n}}} \sim N(0,1)$ likimain, kun H_0 on tosi

4) Testisuureen havaittu arvo:

$$z = \frac{462/1004 - 0.5}{\sqrt{\dfrac{0.5 \cdot (1-0.5)}{1004}}} \approx -2.52$$

5) P-arvo $= P(Z \leq -2.52 \text{ tai } Z \geq 2.52 \mid H_0)$
$= 2P(Z \geq 2.52 \mid H_0)$

$= 2 \cdot 0.0059$
$= 0.0118$

6) Johtopäätökset: Havaintoaineisto on selvässä ristiriidassa H_0:n, H_1 uskottavampi. Aineiston perusteella näyttäisi siis siltä, että nelosoluen myyntiä ruokakaupoissa kannattavien osuus kaikkien suomalaisten keskuudessa ei ole 50%.