OULUN YLIOPISTO
Matemaattisten tieteiden laitos

TENTTIVASTAUSPAPERI

Pvm ______

Opiskelija ______

Kuulustelija ______

Opiskelijanro tai henkilötunnus ______

Tentittävä opintojakso 806109P TILASTOTIETEEN PERUSMENETELMÄT I

2. välikoe

Koulutusohjelma/aloitusvuosi ______

Opintojakson tyyppi P A S

Tentin arvostelu	1	2	3	4	5	6	Σ

1. A) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$= P(A) + P(B) - P(A)P(B)$

$= 0.2 + 0.6 - 0.2 \cdot 0.6$

$= 0.68 \Rightarrow \underline{\underline{a4}}$

B) $\mu = E(X) = \sum_{i=1}^{4} x_i p_i$

$= 0 \cdot 0.1 + 1 \cdot 0.2 + 5 \cdot 0.3 + 7 \cdot 0.2 + 10 \cdot 0.2$

$= 5.1$

$\sigma^2 = D^2(X) = \sum_{i=1}^{4} x_i^2 p_i - \mu^2$

$= (0^2 \cdot 0.1 + 1^2 \cdot 0.2 + 5^2 \cdot 0.3 + 7^2 \cdot 0.2 + 10^2 \cdot 0.2) - 5.1^2$

$= 37.5 - 5.1^2 = 11.49 \Rightarrow \underline{\underline{b5}}$

C) $X \sim N(40, 4^2)$ ja $Y \sim N(50, 5^2)$

$Z = 3X - Y + 5$

$E(Z) = E(3X - Y + 5) = 3E(X) - E(Y) + 5$

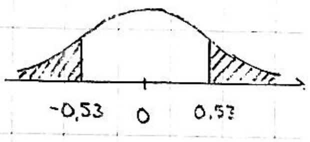
$= 3 \cdot 40 - 50 + 5 = 75$

$D^2(Z) = D^2(3X - Y + 5) = 3^2 D^2(X) + D^2(Y)$

$= 9 \cdot 4^2 + 5^2 = 9 \cdot 16 + 25 = 169 \Rightarrow \underline{\underline{c5}}$

D) Käytetyn testin testisuure $T = \dfrac{\bar{X} - \mu_0}{S_x/\sqrt{n}} \sim t(n-1)$, kun H_0 on tosi

Nyt $T \sim t(30-1) = t(29)$-jakaumaa ja



P-arvo $= P(|T| \geq 0.53 \mid H_0)$

$= P(T \leq -0.53$ tai $T \geq 0.53 \mid H_0)$

$= 2 \cdot 0.3 \leftarrow$ t-jakauman taulukosta

$= 0.6 \Rightarrow \underline{\underline{d6}}$

E) Kun $X \sim Exp(\alpha)$, niin $F(x) = P(X \leq x) = 1 - e^{-\alpha x}$

$P(X \geq 5) = 1 - \overbrace{P(X \leq 5)}^{=F(5)}$

$= 1 - (1 - e^{-0.25 \cdot 5})$

$= e^{-1.25} \approx 0.2865 \Rightarrow \underline{\underline{e5}}$

F) $X \sim Bern(0.7) \Rightarrow P(X = 1) = 0.7$ ja $P(X = 0) = 1 - 0.7 = 0.3$

$\Rightarrow \underline{\underline{f3}}$

2. a) X:n mahdolliset arvot ovat 0, 25, 50 ja 75.

$P(X = 0) = 1/8 = 0.125$

$P(X = 25) = 1/8 = 0.125$

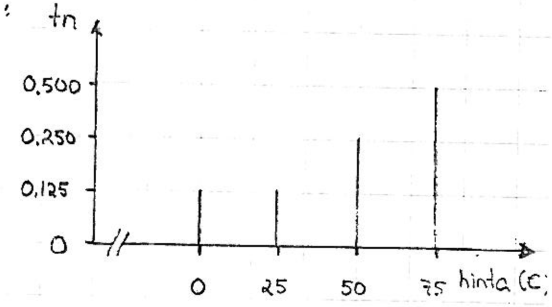
$P(X = 50) = 1/4 = 0.250$

$P(X = 75) = 1/2 = 0.500$

⇒ X:n todennäköisyysjakauma on:

x_i	0€	25€	50€	75€	Yht.
p_i	0.125	0.125	0.250	0.500	1

tn-jakauma graafisesti:



b) 100 puhelimen hankkiminen on maksanut kauppiaalle
$100 + 40 \cdot 100 = 4100$ euroa

Yhden puhelimen myyntihinnan odotusarvo on
$E(X) = \sum_{i=1}^{4} x_i p_i$
$= 0 \cdot 0.125 + 25 \cdot 0.125 + 50 \cdot 0.250 + 75 \cdot 0.500$
$= 53.125$ euroa

⇒ 100 puhelimen myynnistä saatavan rahasumman odotusarvo $E(100X) = 100\,E(X)$
$= 100 \cdot 53.125 = 5312.50$ euroa

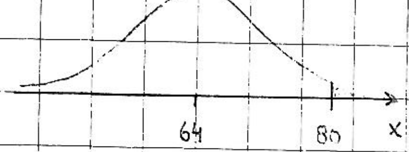
⇒ kauppiaan voitto on $5312.50 - 4100 = \underline{\underline{1212.50 \text{ euroa}}}$

c) Merk. Y = nolla euroa maksavien puhelimien lkm neljän puhelimen joukossa
$Y \sim Bin(4, 0.125)$

$P(Y \geq 2) = P(Y=2) + P(Y=3) + P(Y=4)$
$= \binom{4}{2} \cdot 0.125^2 (1-0.125)^2 + \binom{4}{3} \cdot 0.125^3 (1-0.125)^1 + \binom{4}{4} \cdot 0.125^4 (1-0.125)^0$
$= 0.0718 + 0.0068 + 0.0002$
$= \underline{\underline{0.0788}}$

(3.) Merk. X = kananmunan paino grammoina
$X \sim N(64, 7^2)$

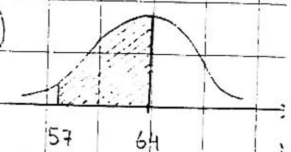
a) $P(X > 80) = P(\underbrace{\frac{X-64}{7}}_{=Z \sim N(0,1)} > \frac{80-64}{7})$
$= P(Z > 2.29)$
$= \underline{\underline{0.0110}}$



b) $\bar{X} = \sum_{i=1}^{7} X_i$

b1) $\bar{X} \sim N(\mu, \sigma^2/n) = N(64, 7^2/7) = N(64, 7) = \underline{\underline{N(64, \sqrt{7}^2)}}$

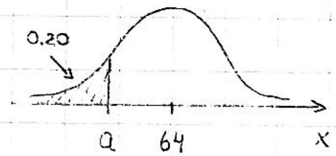
b2) $P(57 \leq \bar{X} \leq 64) = P(\frac{57-64}{\sqrt{7}} \leq \underbrace{\frac{\bar{X}-64}{\sqrt{7}}}_{=Z \sim N(0,1)} \leq \frac{64-64}{\sqrt{7}})$
$= P(-2.65 \leq Z \leq 0)$
$= P(Z \geq 0) - P(Z \geq 2.65)$



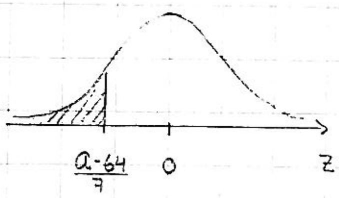
$= 0.5000 - 0.0040$

$= \underline{\underline{0.4960}}$

c) $P(X \leq a) = 0.20 \iff P(\underbrace{\frac{X-64}{7}}_{=Z\sim N(0,1)} \leq \frac{a-64}{7}) = 0.20$



$\iff P(Z \leq \frac{a-64}{7}) = 0.20$



$N(0,1)$-jakauman taulukko:

$P(Z \geq 0.84) \approx 0.20$

$\Rightarrow P(Z \leq -0.84) \approx 0.20$

$\Rightarrow \frac{a-64}{7} = -0.84$

$\Rightarrow a = -0.84 \cdot 7 + 64$

$= \underline{\underline{58.12 \text{ grammaa}}}$

(4.) a) Merk. X = tuotteen toimitusaika vuorokausina

a1) Tulostukessa on laskettu 99 % luottamusväli parametrille μ, kun oletetaan, että $X \sim N(\mu, \sigma^2)$, missä μ ja σ^2 ovat tuntemattomia. Ko. luottamusväli on (12.50696 vrk, 14.42637 vrk). Koska viikko eli 14 vrk kuuluu lasketun luottamusvälin sisään, jonka ainoastaan perusteella on uskottavaa, että tuotteet toimitetaan keskimäärin kahdessa viikossa.

a2) Merk. X = toimitusnopeus, missä

$$X = \begin{cases} 1, & \text{kun toimitusaika on yli kaksi viikkoa} \\ 0, & \text{— " — ei ole — " —} \end{cases}$$

$X \sim \text{Bern}(\pi)$. Lasketaan 95 % luottamusväli parametrille π.

Ko. lv. on: $\left(p - z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}},\ p + z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}\right)$

95 % lv.: $100(1-\alpha) = 0.95 \Rightarrow \alpha = 0.05$ ja $\alpha/2 = 0.025$

$N(0,1)$-jakauman taulukko: $P(Z \geq 1.96) = 0.025$

$\Rightarrow z_{\alpha/2} = z_{0.025} = 1.96$

$p = \frac{6}{20} = 0.3$

$\Rightarrow$ kysytty lv. on $\left(0.3 - 1.96 \cdot \sqrt{\frac{0.3\cdot(1-0.3)}{20}},\ 0.3 + 1.96 \cdot \sqrt{\frac{0.3\cdot(1-0.3)}{20}}\right)$

$\iff \underline{(0.136, 0.464)}$

95 %:n varmuudella kysytty suhteellinen osuus on välillä (0.136, 0.464).

b) 95 %:n lv. π:lle (= kokoomuksen kannatusosuus) lasketaan samalla kaavalla kuin kohdassa a2. Nyt

$z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}} \leq 0.02 \quad | : z_{\alpha/2}$

$\sqrt{\frac{p(1-p)}{n}} \leq \frac{0.02}{z_{\alpha/2}} \quad | ()^2$

$$\frac{p(1-p)}{n} \leq \left(\frac{0.02}{z_{\alpha/2}}\right)^2 \qquad \Big| \cdot n$$

$$\Leftrightarrow p(1-p) \leq \left(\frac{0.02}{z_{\alpha/2}}\right)^2 n \qquad \Big| : \left(\frac{0.2}{z_{\alpha/2}}\right)^2$$

$$\Leftrightarrow n \geq \frac{p(1-p)}{\left(\frac{0.02}{z_{\alpha/2}}\right)^2}$$

95 % lv.: $z_{\alpha/2} = 1.96$ ja $p = 0.26$

$$\Rightarrow n \geq \frac{0.26 \cdot (1-0.26)}{\left(\frac{0.02}{1.96}\right)^2} \approx 1847.8 \Rightarrow \underline{\text{vähintään } 1848}$$

b2) samoin kuin b1, mutta 0,02 muuttuu luvuksi 0,015

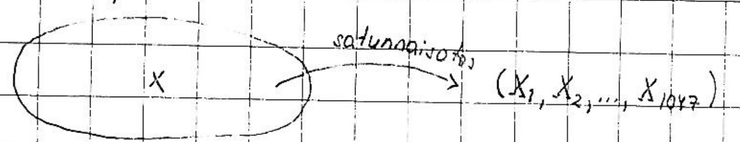
$$\Rightarrow n \geq \frac{0.26 \cdot (1-0.26)}{\left(\frac{0.015}{1.96}\right)^2} \approx 3285$$

Tarvitaan siis 3285 - 1848 = 1437 haastateltavaa enemmän

(5.) χ^2-yhteensopivuustesti

1) Populaatio: ravintolassa tehdyt tilaukset maaliskuun jälkeen

X = tilausvaihtoehdot



X luokiteltu

2) Hypoteesit: H_0: X:n todennäköisyysjakauma on

X_i	hampurilaisateria	grillilautanen	päivän wokki	salaatti	Yht.
p_i^0	0.48	0.27	0.13	0.12	1

3) Testisuure $\chi^2 = \sum_{i=1}^{k} \frac{(F_i - e_i)^2}{e_i} \sim \chi^2(k-m-1)$ likimain, kun H_0 tos

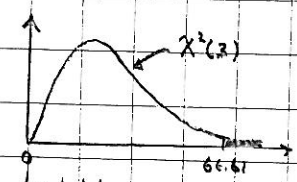
4) Testisuureen havaittu arvo:

tilausvaihtoehdot	F_i	p_i^0	$e_i = np_i^0$
hampurilaisateria	380	0.48	$1047 \cdot 0.48 \approx 502.56$
grillilautanen	313	0.27	$1047 \cdot 0.27 \approx 282.69$
päivän wokki	192	0.13	$1047 \cdot 0.13 \approx 136.11$
salaatti	162	0.12	$1047 \cdot 0.12 \approx 125.64$
Yhteensä	1047	1	1047

$$\Rightarrow \chi^2 = \sum_{i=1}^{4} \frac{(F_i - e_i)^2}{e_i}$$

$$= \frac{(380-502.56)^2}{502.56} + \frac{(313-282.69)^2}{282.69} + \frac{(192-136.11)^2}{136.11} + \frac{(162-125.64)^2}{125.64}$$

$$\approx 66.61$$



5) P-arvo = $P(\chi^2 \geq 66.61 \mid H_0)$

$\chi^2(k-m-1) = \chi^2(4-0-1) = \chi^2(3)$-jakauman taulukko: p-arvo < 0.001

6) Johtopäätökset: havaintoaineisto on selvässä ristiriidassa H_0:n kanssa, H_1 uskottavampi. Aineiston perusteella näyttää siis siltä, että ruokavaihtoehtojen suosiossa on tapahtunut muutos uudistuksen yhteydessä