

MS-A0503 Todennäköisyyslaskennan ja tilastotieteen peruskurssi

Tentti ja välikokeet 12.3.2014/ Aro

Kirjoita selvästi *jokaiseen koepaperiin* seuraavat tiedot:

- Kurssikoodi, päiväys, ja mitä vaihtoehtoa olet suorittamassa (ks. alla)
- opiskelijanumero + kirjain
- TEKSTATEN sukunimi ja kaikki etunimet
- koulutusohjelma ja vuosikurssi
- nimikirjoitus

Lue *tarkasti* seuraavat ohjeet:

- (1) Jos olet suorittamassa kurssia tentillä, vastaa kysymyksiin 1-5.
- (2) Jos olet suorittamassa 1. välikoetta erikoisluvalla, vastaa kysymyksiin 1, 2, 6 ja 7.
- (3) Jos olet suorittamassa 2. välikoetta erikoisluvalla, vastaa kysymyksiin 3-5 ja 8.

Tentin kesto on 3 tuntia. 1. välikokeen kesto on 2 tuntia ja 2. välikokeen 3 tuntia.

Merkitse vastauspapereihisi selvästi mitä vaihtoehtoa olet suorittamassa!

Sallitut apuvälineet: *Laskin* ja Mellinin kaava- ja taulukkokokoelmat.

Vastausohjeet: Vastaa *lyhyesti* ja *ytimekkäästi*, mutta *perustele ratkaisusi*.
Esimerkiksi pelkkä lukuarvo vastauksena *ei anna* pisteitä.

1. Alla oleva taulukko esittää diskreettien satunnaismuuttujien X ja Y yhteisjakauman pistetodennäköisyysfunktiota, $f_{XY}(x, y) = \Pr(X=x \text{ ja } Y=y)$.

$f_{XY}(x,y)$		x		
		-1	1	2
y	3	0	0	1/4
	1	1/4	1/4	0
	-2	1/4	0	0

Määrää:

- (a) Satunnaismuuttujien X ja Y reunajakaumat.
 - (b) Satunnaismuuttujien X ja Y odotusarvot, varianssit ja standardipoikkeamat
2. Tehdas valmistaa tuotetta, jolla on erittäin korkeat laatuksiteerit. Keskimäärin vain 60 % tuotteista täyttää kriteerit. Poimitaan tuotteita tarkastettavaksi satunnaisesti yksi kerrallaan.
Tehtävä 1: Mikä on todennäköisyys sille, että joudumme tarkastamaan vähintään 4 tuotetta *ensimmäisen* viallisen tuotteen löytämiseksi?

Tehtävä 2: Mikä on odotettavissa oleva lukumäärä tuotteille, jotka joudumme tarkastamaan *ensimmäisen* viallisen tuotteen löytämiseksi?

Tehtävä 3: Mikä on todennäköisyys, että joudumme tarkastamaan ainakin 5 tuotetta *kolmannen* viallisen tuotteen löytämiseksi?

3. Sparks Electric Institute on julkaissut lukuja, jotka kertovat kuinka monta kilowattituntia eri kodin sähkölaitteet kuluttavat vuosittain. Se väittää, että pölynimuri kuluttaa keskimäärin 46 kilowattituntia vuodessa.

Valitaan 12 pölynimurin suuruinen yksinkertainen satunnaisotos, ja kirjataan imurien vuosittainen kilowattituntien kulutus. Aritmeettiseksi keskiarvoksi saadaan $\bar{x} = 42$ ja otosvarianssiksi $s^2 = 141.61$. Löytyykö tästä riittävästi todisteita 5% merkitsevyystasolla, että pölynimurit kuluttavat keskimäärin vähemmän kuin 46 kilowattituntia vuosittain? Oleta, että pölynimurin vuosittain kuluttamien kilowattituntien määrä on normaalijakautunut.

4. Kyselytutkimuksessa haluttiin verrata Vasemmistoliiton (Vas), Sosiaalidemokraattisen puolueen (SDP) ja Vihreän liiton (Vih) kannatuksen jakautumista äänioikeutettujen joukossa kolmessa kunnassa A, B ja C.

Vertailua varten kuntien A, B ja C äänioikeutettujen joukosta poimittiin toisistaan riippumattomat yksinkertaiset satunnaisotokset, joiden koot olivat 300 (kunta A), 180 (kunta B) ja 340 (kunta C) ja otokseen poimituilta kysyttiin, mitä puoluetta he äänestäisivät seuraavissa vaaleissa. Tulokset kyselyistä on annettu alla olevassa taulukossa.

Testaa nollahypoteesia, että puolueiden kannatus jakautuu samalla tavalla kunnissa A, B ja C, kun vaihtoehtoisena hypoteesina on, että kannatus ei jakaudu samalla tavalla. Käytä testissä 5 %:n merkitsevyystasoa.

	Puolue	Vas	Sdp	Vih	Otoskoko
Kunta	A	60	160	80	300
	B	40	80	60	180
	C	60	160	120	340

5. Kokeessa tutkittiin seitsemän kuorma-auton polttoainetaloudellisuuden (muuttuja y , yksikkönä *mi/gal*, mailia per gallona) riippuvuutta ajoneuvon painosta (muuttuja x , yksikkönä *ton*, tonni). Kokeesta saadut tiedot on annettu alla olevassa taulukossa.

i	1	2	3	4	5	6	7
x_i	8.00	24.50	27.00	14.50	28.50	12.75	21.25

y_i	7.69	4.97	4.56	6.49	4.34	6.24	4.45
-------	------	------	------	------	------	------	------

Aineistoa kuvaavat otostunnusluvut ovat:

$$\bar{x} = 19.5 \qquad \bar{y} = 5.5343$$

$$s_x^2 = 61.354 \qquad s_y^2 = 1.655$$

$$s_x = 7.833 \qquad s_y = 1.286$$

$$s_{xy} = -9.518 \qquad r_{xy} = -0.945$$

- (a) Määrää tavanomaisen yhden selittäjän lineaarisen regressiomallin

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \varepsilon_i : N(0, \sigma^2), i = 1, 2, \dots, K, n$$

regressiokertoimien β_0 ja β_1 pienimmän neliösumman (PNS-) estimaatit.

- (b) Jännösvarianssin σ^2 harhattomaksi estimaatiksi tutkijat laskivat arvon $s^2 = 0.214$. Ennusta muuttujan y keskimääräinen arvo, kun $x = 40$. Määrää myös 95%:n luottamusväli ennusteelle.

6. Satunnaismuuttujan X tiheysfunktio on muotoa

$$f(x) = \begin{cases} ax^2 - 2ax, & \text{kun } 0 \leq x \leq 1 \\ 0, & \text{muulloin} \end{cases}$$

- (a) Määrää vakion $a < 0$ arvo.
 (b) Määrää satunnaismuuttujan X kertymäfunktio.
 (c) Määrää tapahtuman $X = 0.5$ todennäköisyys.
 (d) Määrää tapahtuman $0 \leq X \leq 0.25$ todennäköisyys tiheysfunktion avulla.
 (e) Määrää tapahtuman $0 \leq X \leq 0.25$ todennäköisyys kertymäfunktion avulla.

7. Heität virheetöntä noppaa 24000 kertaa. Mikä on todennäköisyys, että kuutosten lukumäärä on suljetulla välillä [3920, 4160]? Käytä normaalijakauma-aproksimaatiota.

8. Tehdas väittää, että normaalitilanteessa sen valmistamista tuotteista 2.5 % on viallisia. Asiakas poimii tuotteiden joukosta yksinkertaisen satunnaisotoksen, jonka koko on 100, ja löytää 9 viallista tuotetta. Voidaanko tehtaan väitettä viallisten tuotteiden suhteellisesta osuudesta pitää oikeutettuna?

Vastaa kysymykseen määräämällä otoksesta 95 %:n ja 99 %:n luottamusvälit viallisten tuotteiden todelliselle suhteelliselle osuudelle ja tekemällä johtopäätökset tehtaan väitteen oikeutuksesta niiden perusteella