

# 806109P TILASTOTIETEEN PERUSMENETELMÄT I

## 2. välikoe 27.4.2012 (Jari Päckilä)

### VALITSE VIIDESTÄ TEHTÄVÄSTÄ NELJÄ JA VASTAA VAIN NIIHIN!

1. Valitse kohdissa A-F oikea (vain yksi) vaihtoehto. Perusteluja ei vaadita, mutta ne voi laittaa vastaukseen näkyviin. Oikeasta vastauksesta saat +1 pistettä, väärästä et menetä pisteitä.

A) Tapahtumat A ja B ovat toisistaan riippumattomat,  $P(A) = 0.2$  ja  $P(B) = 0.6$ . Tällöin todennäköisyys, että ainakin toinen tapahtumista A tai B sattuu, on

a1) 0.12    a2) 0.80    a3) 0.33    a4) 0.68    a5) 0.00    a6) ei voida määrätä.

B) Satunnaismuuttujan  $X$  todennäköisyysjakauma on:

$x_i$	0	1	5	7	10	Yht.
$p_i$	0.1	0.2	0.3	0.2	0.2	1

$X$ :n varianssi on

b1) 8.25    b2) 16.34    b3) 8.89    b4) 17.30    b5) 11.49    b6) 0.00.

C) Satunnaismuuttuja  $X \sim N(40, 4^2)$  ja  $Y \sim N(50, 5^2)$ . Lisäksi tiedetään, että  $X$  ja  $Y$  ovat toisistaan riippumattomia. Olkoon satunnaismuuttuja  $Z = 3X - Y + 5$ .  $Z$ :n jakauma on

c1)  $N(75, 23)$     c2)  $N(75, 28)$     c3)  $N(75, 73)$   
c4)  $N(75, 119)$     c5)  $N(75, 169)$     c6)  $N(75, 194)$

D) Normaalijakaumaa noudattavasta satunnaismuuttujasta  $X$  on saatu 30 kappaleen satunnaisotos.  $X$ :n jakauman odotusarvoon kohdistuvasta tilastollisesta analyysistä on käytössä alla esitetty R-ohjelman tulostus:

```
> t.test(x, mu=100, alternative='two.sided')

One Sample t-test

data:  x
t = -0.53, df = 29, p-value = x.xxxx
alternative hypothesis: true mean is not equal to 100
95 percent confidence interval:
 96.52537 102.04439
sample estimates:
mean of x
 99.28488
```

Tulostuksessa merkinnällä x.xxxx peitetylle p-arvolle pätee:

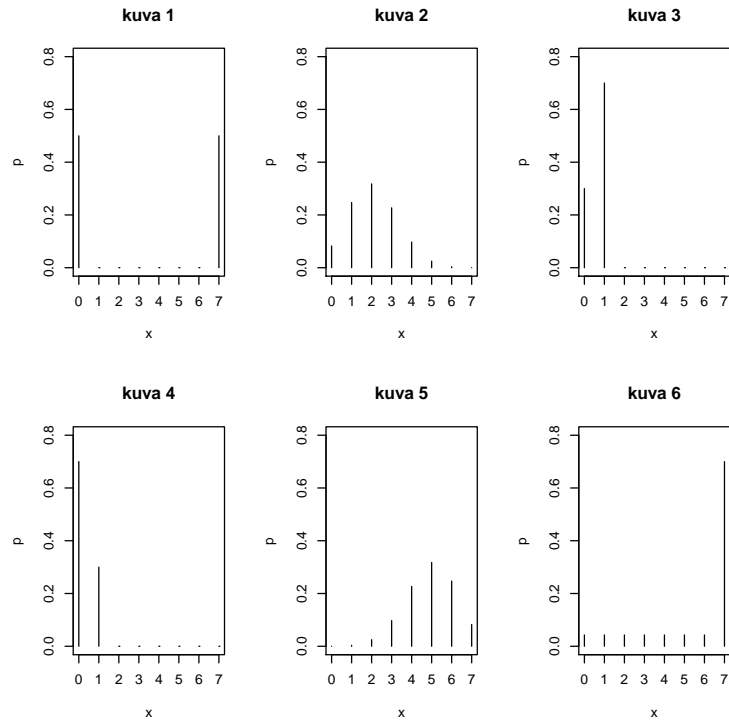
d1) 0.7000,    d2) 0.3000,    d3) 0.2981,    d4) 0.7019,    d5) 0.5962,    d6) 0.6000.

E) Satunnaismuuttuja  $X \sim \text{Exp}(0.25)$ . Todennäköisyys  $P(X \geq 5)$  on likimain

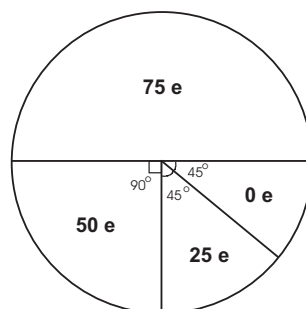
- e1) 0.0716    e2) 0.7135    e3) 0.7788    e4) 0.0067    e5) 0.2865    e6) 0.9284

F) Satunnaismuuttuja  $X \sim \text{Bern}(0.7)$ .  $X$ :n todennäköisyysjakauma on esitetty alla olevassa

- f1) kuvassa 1,                      f2) kuvassa 2,                      f3) kuvassa 3,  
 f4) kuvassa 4,                      f5) kuvassa 5,                      f6) kuvassa 6.



2. Kauppias haluaa myydä matkapuhelinvarastonsa tyhjäksi. Myyntiä vauhdittaakseen hän aloittaa kampanjan, jossa matkapuhelimen hinta määräytyy tikanheiton perusteella. Tikkataulu on jaettu hintasektoreihin seuraavalla tavalla:



Tikkaa heitettäessä taulu pyörii keskipisteensä ympäri ja matkapuhelimen hinta on ensimmäisen tauluun osuneen tikan osoittaman euromäärän suuruinen.

a) Määrää matkapuhelimen hinnan ( $= X$ ) todennäköisyysjakauma ja esitä se graafisesti.

(2 p)

- b) Kauppiaalla on kampanjan alkaessa 100 matkapuhelinta varastossaan. Puhelimia hankkiessaan kauppias on maksanut perusmaksuna 100 euron toimitusmaksun, minkä lisäksi hän on maksanut jokaisesta puhelimesta 40 euroa. Paljonko kauppias jää keskimäärin voitolle/tappiolle, kun hän on myynyt em. kampanjassa kaikki puhelimet? (2 p)
- c) Nelihenkinen Niemisen perhe päättää ostaa kaikille uudet matkapuhelimet (4 kpl). Millä todennäköisyydellä he saavat ainakin kaksi matkapuhelinta ilmaiseksi? (Jokaisen matkapuhelimen hinta määräytyy erillisellä tikanheitolla). (2 p)

**3.** Eräessä kanalassa munittujen kananmunien paino (grammoina) noudattaa normaalijakamaa odotusarvolla 64 grammaa ja keskihajonnalla 7 grammaa.

- a) Millä todennäköisyydellä edellä kuvatussa kanalasta satunnaisesti valittu kananmuna painaa yli 80 grammaa? (2 p)
- b) Kanalan munista on poimittu satunnaisesti seitsemän munaa.
- b1) Mitä jakaumaa tässä tapauksessa kananmunien painon otoskeskiarvo (grammoina) noudattaa ja millä parametrien arvoilla? (1 p)
- b2) Millä todennäköisyydellä satunnaisotoksen kananmunien painon keskiarvo on vähintään 57 grammaa, mutta korkeintaan 64 grammaa? (1 p)
- c) Määrää sellainen kananmunan paino, jota kevyempiä on 20 % kanalassa munituista kananmunista. (2 p)

**4.** a) Verkkokaupparyitys ilmoittaa toimittavansa tilatut tuotteet kahden viikon kuluessa tilauksesta. Toimintansa kehittämiseksi yritys teki omille asiakkailleen tyytyväisyyskyselyyn, jossa tilattujen tuotteiden toimitusajat (vuorokausina) olivat seuraavat.

```
> toimitusaika <- c(12, 15, 13, 16, 16, 14, 12, 16, 12, 15,
+                  14, 14, 16, 14, 14, 15, 13, 12, 14, 12,
+                  14, 14, 11, 10, 15, 12, 17, 11, 11, 10)
```

- a1) Toimittaako yritys tilatut tuotteet keskimäärin kahden viikon kuluessa? Perustele vastauksesi lyhyesti alla olevan R-tulostuksen luottamusvälin perusteella. (1 p)

```
> t.test(toimitusaika, alternative='two.sided', mu=14, conf.level=.99)
```

```
One Sample t-test
```

```
data: toimitusaika
t = -1.5318, df = 29, p-value = 0.1364
alternative hypothesis: true mean is not equal to 14
99 percent confidence interval:
 12.50696 14.42637
sample estimates:
mean of x
 13.46667
```

- a2) Edellä olevista R-tulostuksista huomataan, että yhdeksän tuotteen toimitusaika oli yli kaksi viikkoa. Kuinka suuri on niiden verkkokaupparytityksen toimittamien tuotteiden osuus, joiden toimitusaika on yli kaksi viikkoa? Laske tilanteeseen sopiva 95 %:n luottamusväli. (2 p)
- b) Erään mielipidekyselyn perusteella on tarkoitus arvioida kokoomuksen kannatusosuutta.
- b1) Kuinka monta haastattelua kyselyssä täytyy vähintään tehdä, kun kokoomuksen kannatusosuuden 95 %:n luottamusvälin pituuden halutaan olevan korkeintaan neljä prosenttiyksikköä (eli virhemarginaali on korkeintaan kaksi prosenttiyksikköä suuntaansa)? Oletetaan, että minkään puolueen kannatus ei ole tutkimushetkellä suurempi kuin 26.0 prosenttia. (2 p)
- b2) Kuinka paljon enemmän haastatteluja olisi pitänyt tehdä, jotta kyseisessä tutkimuksessa ilmoitetun luottamusvälin pituus olisi ollut korkeintaan kolme prosenttiyksikköä? (1 p)

**5.** Eräässä ravintolassa on tarjolla neljä ruokavaihtoehtoa. Maaliskuussa ravintolan kaikista tilauksista 48 % oli hampurilaisaterioita, 27 % grillilautasia, 13 % päivän wokkeja ja loput 12 % salaatteja. Maaliskuun jälkeen ravintola uudisti ruokavaihtoehtojensa sisältöä: hampurilaisaterian hampurilaiseen otettiin suurempi pihvi, grillilautaseen erilaisia ranskalaisia jne. Uudistuksen jälkeen tehdyssä tutkimuksessa 1047 satunnaisesti valituista tilauksista hampurilaisateria oli tilattu 380 kertaa, grillilautanen 313 kertaa, päivän wokki 192 kertaa ja salaatti 162 kertaa. Näyttääkö aineiston perusteella sille, että ravintolan eri ruokavaihtoehtojen suosiossa olisi tapahtunut uudistuksen yhteydessä muutos? Tee tilanteeseen sopiva merkitsevyytestaus. Merkitse vastauksessasi näkyviin kaikki testauksen vaiheet. (6 p)