

## 806109 TILASTOTIETEEN PERUSMENETELMÄT I

### Harjoitus 13, viikko 49, syksy 2010

1. Turvavöitä valmistava yritys on pitkäaikaisissa tutkimuksissaan todennut, että sen valmistamien vöiden murtumisluku on keskimäärin 600 kg. Yritys on nyt kehittänyt uuden valmistusmenetelmän ja väittää uusien vöiden murtumislukuun olevan aiempia vöitä korkeamman. Kymmenen satunnaisesti valitun uudella menetelmällä valmistetun vöiden murtumisluku olivat:

608, 650, 624, 598, 645, 612, 595, 672, 643, 640

Näyttääkö uusien vöiden keskimääräinen murtumisluku olevan vanhalla menetelmällä valmistettuja korkeampi? Analysoi aineistoa sopivalla merkitsevyydestillä. Murtumislukuun oletetaan noudattavan normaalijakaumaa.

2. Kuntasektorin eri toimialojen välisten palkkojen vertailemiseksi valittiin satunnaisesti 110 työntekijää toimialalta A ja 142 työntekijää toimialalta B. Näiden työntekijöiden kuukausipalkoista (euroa) lasketut keskiarvot ja keskihajonnat olivat:

	Toimiala A	Toimiala B
Keskiarvo	3405	3350
Keskihajonta	380	420

Tee tarvittavat oletukset ja tutki sopivan merkitsevyydestin avulla sitä, onko keskimääräinen kuukausipalkka toimialalla A suurempi kuin toimialalla B. Kuukausipalkkojen varianssien voidaan olettaa olevan yhtä suuret toimialoilla A ja B. Tee testaus merkitsevyydestä sisältyvien vaiheiden mukaisesti. Täydennä testin antamaa tulosta sopivalla 95%:n luottamusvälillä.

3. Halutaan tutkia, onko henkilöautojen hylkäämisprosentti sama katsastusasemilla A ja B. Satunnaisesti valituista 600:sta asemalla A käyneestä henkilöautosta hylättiin (ensimmäisellä käyntikerralla) 150, vastaava luku satunnaisesti valituista 800:sta asemalla B käyneestä henkilöautosta oli 224. Suorita tilanteeseen sopiva merkitsevyydestesti. Täydennä testin tulosta sopivalla 95%:n luottamusvälillä.
4. Yrityksen AB oman ilmoituksen mukaan AB:n palvelupuhelimeen soittavista vain 23% joutuu odottamaan yli minuutin. Väitettä tutkittiin soittamalla palveluun eri aikoina 50 kertaa, näistä 19 puhelua joutui jonottamaan yli minuutin. Tukeeko saatu tulos soittajien epäilyjä, että yli minuutin joutuu odottamaan enemmän kuin 23%. Analysoi aineistoa sopivalla merkitsevyydestillä.
5. Lääkäriaseman asiakkaista tiedetään aiempien vuosien kokemuksella, että 80 prosenttia saapuu vastaanotolle vähintään 10 minuuttia etuajassa, 10 prosenttia saapuu paikalle 0-9 minuuttia etuajassa, kuusi prosenttia myöhästyy korkeintaan 5 minuuttia ja neljä prosenttia myöhästyy yli 5 minuuttia. Lääkäriaseman henkilökunta on kuitenkin viime aikoina alkanut epäillä, että asiakkaiden saapumisaikataulussa on tapahtunut muutos ja asiaa päätetään selvittää satunnaisotoksen avulla. Neljästä sadasta satunnaisesti valitusta asiakkaasta vähintään 10 minuuttia etuajassa tuli 287, 0-9 minuuttia etuajassa oli 49, korkeintaan 5 minuuttia myöhästyi 30 ja loput myöhästyi yli 5 minuuttia. Onko saapumisaikan jakauma muuttunut? Analysoi aineistoa sopivalla merkitsevyydestillä.

6. Satunnaisesti valituilta 1772 yli 18-vuotiaalta henkilöltä USA:ssa kysyttiin alkoholin kulutusta ja siviilisäätyä ja saatiin seuraava taulukko:

Siviilisääty	Alkoholiannoksia kk:ssa			yht.
	Ei yhtään	1-60	yli 60	
Naimaton	67	213	74	354
Naimisissa	411	633	129	1173
Leski	85	51	7	143
Eronnut	27	60	15	102
Yht.	590	957	225	1772

Näyttääkö siviilisäädyn ja alkoholin käytön välillä olevan riippuvuutta USA:ssa? Analysoi aineistoa sopivalla merkitsevyydestillä.

7. Erään pankin asiakasrekisteristä poimittiin 500 henkilöasiakkaan satunnaisotos. Otokseen mukaan tulleille asiakkaille lähetettiin kyselylomake, joka sisälsi mm. seuraavat kaksi kysymystä: "Uskotteko oman taloutenne kohenevan seuraavan vuoden aikana?" ja "Kuinka paljon käytätte aikaa (minuuteissa) pankkimme verkkopalveluihin yhden kuukauden aikana?".
- a) Kyselyyn vastasi 240 miestä ja 260 naista. Oman taloutensa kohenemiseen seuraavan vuoden aikana uskoi 72 miestä ja 91 naista. Tutki sopivan 95% luottamustavan välin avulla väitteen "Miehet ja naiset uskovat yhtä vahvasti oman taloutensa kohenemiseen seuraavan vuoden aikana" uskottavuutta. Merkitse vastaukseesi näkyviin myös laskelmissa tarvittavat mallioletukset ja tulkitse saamasi tulos lyhyesti.
- b) Ennen kyselyn tekemistä pankissa arvioitiin verkkopalveluita käyttävän asiakkaan viipyvän pankin verkkopalveluissa keskimäärin 45 minuuttia kuukaudessa. Seuraavassa R-ohjelman tulostuksessa on analysoitu kyselyä saatua aineistoa.

```
> # muuttuja verkko aika = asiakkaan viettämä aika pankin verkko-
> # palveluissa yhden kuukauden aikana
>
> t.test(verkko aika,mu=45,conf.level=0.9)
```

One Sample t-test

```
data: verkko aika
t = 10.9918, df = 450, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 45
90 percent confidence interval:
 xx.xxxx-xx.xxxx
sample estimates:
mean of verkko aika
50.1700
```

b1) Lomakkeen kysymykseen ”Kuinka paljon käytätte aikaa (minuuteissa) pankkimme verkkopalveluihin yhden kuukauden aikana?” pyydettiin vastaus ainoastaan niiltä asiakkailta, joilla oli ko. pankin verkkotunnukset. Kuinka moni vastasi ao. kysymykseen?

b2) Laske tulostuksessa esiintyvän luottamusvälin ala- ja yläraja (tulostuksessa rajat peitetty).

b3) Näyttäisikö pankin ennakoarvio keskimääräisestä 45 minuutin verkkopalveluiden käyttöajasta olevan realistinen? Perustele vastauksesi lyhyesti käyttämällä perustelussasi sekä b2)-kohdassa laskemaasi luottamusväliä että R:n tulostuksen merkitsevyyystestausta.

**HUOM. 1** Viikolla 49 harjoituksia on seuraavasti: ti 7.12. klo 10-14 salissa FY1120, ke 8.12. klo 12-18 salissa IT112 ja pe 10.12. klo 8-13 salissa IT112. Viimeinen varsinainen luento on ke 8.12. klo 10-12 salissa L6, ma 13.12. on ylimääräinen kertaustuento klo 12-14 salissa L6.

**HUOM. 2** 2. välikoe on tiistaina 14.12. klo 10-14 salissa L1.

**HUOM. 3** Palautetta kurssista voi antaa matemaattisten tieteiden laitoksen nettisivuilta (<http://math oulu.fi/>) löytyvällä palautelomakkeella.