

Tilastotiede

- Regressiosuora (geogebra, 2. muuttujan tilasto)
- Keskiarvo, hajonta, maksimi, mediaani... (geogebra 1. muuttujan tilasto)
- Moodi (libreoffice taajuusfunktio)
- Normaalijakaumalla todennäköisyydet (geogebra todennäköisyyslaskuri)
- Mallien tulkitseminen
- Virhe ja luottamusväli (MAOL ja kertausdiat)

4. Frekvenssitaulukko 12 p.

Aineisto

4.A Taulukko: Koirien selän pituus

Taulukossa 4.A on esitetty aineisto beagle-koirien selän pituudesta.

1. Mikä on aineiston tyyppiarvo eli moodi? (2 p.)
2. Kuinka monta koiraa aineistoon sisältyy? (2 p.)
3. Mikä on aineiston koirien selän pituuden keskiarvo? (8 p.)

4.A Taulukko: Koirien selän pituus

Pituus (cm)	Lukumäärä
42	1
43	0
44	8
45	7
46	4
47	11
48	3

9. Noppatilastoja 12 p.

Aineisto

9.A Taulukko: Noppatilasto

Tämä tehtävä on tarkoitettu ohjelmistolla ratkaistavaksi. Vastaukset voi antaa likiarvoina, ja perusteluiksi riittävät kuvakaappaukset tai selitykset, joista ilmenee, mitä on tehty.

Opiskelijat yrittivät selvittää, minkälaisen jakauman viiden nopan silmäluvun summa toteuttaa. He heittivät 100 kertaa viittä noppaa ja kirjasivat tulokset, ks. taulukko 9.A. Laskemalla jokaisen heittokerran viiden nopan silmäluvut yhteen saadaan *summamuuttuja*.

1. Määritä tämän summamuuttujan minimi, maksimi, moodi, mediaani, keskiarvo ja keskihajonta. (6 p.)
2. Piirrä tämän summamuuttujan jakauman pylväskaavio. (3 p.)
3. Muistuttaako jakauma enemmän normaalijakaumaa vai tasaista jakaumaa? (3 p.)

12. Peruskoulujen lukumäärä Suomessa 12 p.

Aineisto

12.A Taulukko: Peruskoulujen lukumäärä

Peruskoulujen lukumäärä Suomessa vuosina 2005–2020 on esitetty taulukossa 12.A.

1. Piirrä diagrammi, joka esittää peruskoulujen lukumäärät vuosina 2005–2020. (2 p.)
2. Sovita aineistoon regressiosuora $y = a + bx$, kun vuosi on x -akselilla ja peruskoulujen määrä y -akselilla. Selitä sanallisesti kertoimien a ja b merkitys. (4 p.)
3. Sovita aineistoon regressiosuora käyttäen vain vuosien 2005–2008 peruskoulujen lukumääriä. (2 p.)
4. Arvioi kummankin mallin perusteella vuosi, jolloin peruskoulut häviävät Suomesta. (4 p.)

11. Kirja-arvioita 12 p.

Kirjaviisauteen kyllästynyt dosentti aikoo myydä kaikki kirjansa antikvariaattiin. Tarjouspyyntöä varten hän arvioi kirjojen lukumäärää.

1. Kolmella umpimähkäisellä mittauksella hän saa seuraavat tulokset:

Mittaus 1: 25 kirjaa yhteensä 65,0 cm

Mittaus 2: 28 kirjaa yhteensä 76,0 cm

Mittaus 3: 14 kirjaa yhteensä 30,0 cm

Missä näistä kolmesta mittauksesta on keskimäärin paksuimmat kirjat? (3 p.)

2. Kirjoja on yhteensä 38,07 hyllymetriä. Dosentti arvioi kirjojen kokonaismäärää kohdan 1 kolmen mittauksen perusteella. Hänellä on mielessään kaksi menetelmää:

Menetelmä 1: Jaan pituuden 38,07 kohdassa 1 laskettujen keskiarvojen keskiarvolla.

Menetelmä 2: Lasken kolmen mittauksen kirjamäärät ja yhteispaksuuden ja saan näin arvion keskimääräisen kirjan paksuudesta. Jaan pituuden 38,07 tällä luvulla.

Laske kummankin menetelmän mukainen arvio kirjojen kokonaismäärälle ja arvioi, kumpi menetelmä on parempi. (9 p.)

12. Maailman väestö 12 p.

Aineisto

12.A Taulukko: Väestömäärä YK:n arvioiden mukaan

Taulukossa 12.A on esitetty maailman väestömäärä vuodesta 1950 vuoteen 2020.

1. Mallinna väestömäärää sovittamalla aineistoon regressiosuora $y = a + bx$, kun x on vuosi ja y väestömäärä. (4 p.)
2. Väestönkasvua voi mallintaa myös lukujonolla $y_n = c q^n$, missä n on vuosi ja y_n on väestömäärä. Määritä parametrit c ja q käyttämällä vuosien 2010 ja 2020 väestömääriä. Vaihtoehtoisesti voit käyttää mallia $y_n = c e^{kn}$, joka on käytössä tietyissä ohjelmistoissa. (4 p.)
3. Määritä ennusteet vuoden 2040 väestömäärästä osatehtävien 12.1 ja 12.2 mallien perusteella. (2 p.)
4. Arvioi osatehtävien 12.1 ja 12.2 mallien hyviä tai huonoja puolia. Mainitse kummastakin mallista yksi seikka. (2 p.)

9. Suomalaisten miesten ja naisten pituudet 12 p.

Suomalaisten naisten keskipituus on 167,5 cm ja pituuden keskihajonta 5,40 cm. Suomalaisten miesten keskipituus on 181,0 cm ja pituuden keskihajonta 6,06 cm. Oletetaan suomalaisten miesten ja naisten pituudet normaalijakautuneiksi.

1. Laske todennäköisyys, että kun valitaan satunnaisesti yksi suomalainen mies ja yksi suomalainen nainen, niin molemmat ovat lyhyempiä kuin 160 cm. (6 p.)
2. Annelin pituus on 171 cm. Hän laskee todennäköisyyden, että satunnaisesti valittu suomalainen nainen on häntä itseään pidempi mutta ei kuitenkaan enempää kuin x cm, ja saa tulokseksi 0,1. Mikä on Annelin käyttämä luku x ? (6 p.)

9. Älypuhelinien käyttöikä 12 p.

Aineisto

9.A Taulukko: Älypuhelinien käyttöikä

Älypuhelinien keskimääräiset käyttöiät eräissä maissa on esitetty taulukossa 9.A.

1. Laske taulukossa esitettyjen keskimääräisten käyttöikäiden keskiarvo $\bar{x}$ ja keskihajonta s . (4 p.)
2. Miksi osatehtävässä 9.1 laskettu keskiarvo $\bar{x}$ ei tuota kyseisten maiden älypuhelinien käyttöikäiden keskiarvoa? (4 p.)
3. Oletetaan, että älypuhelinien käyttöikä koko maailmassa noudattaa normaalijakaumaa, jonka odotusarvo on 23 kuukautta ja keskihajonta 2,55 kuukautta. Millä todennäköisyydellä satunnaisesti valitun älypuhelimien käyttöikä on vähintään kolme vuotta? (4 p.)

9. Luottamusväli 12 p.

Pakkauskone on säädetty täyttämään puolen kilon muropaketteja. Laadunvalvonnassa punnitaan 100 pakettia, joiden painon keskiarvoksi saadaan 509 grammaa ja keskihajonnaksi 6,0 grammaa. Oletetaan, että paketin paino noudattaa normaalijakaumaa $N(509; 6,0)$.

1. Millä todennäköisyydellä paketin paino on korkeintaan 500 grammaa? **(4 p.)**
2. Määritä paketin painon keskiarvon keskivirhe ja 95 %:n luottamusväli, eli luottamusväli 95 %:n luottamustasolla. **(8 p.)**