

## Tilastotieteen johdantokurssin tentti 10.3.2115

Tentissä saa olla mukana kirjoitusvälineet ja taskulaskin.

1. Seuraavassa taulukossa on eräiden maiden kahvin ja teen kulutuksia. Tutki kahvin ja teen kulutuksen riippuvuutta sopivalla kuviolla ja tunnusluvulla ja tulkitse kuvailu.

| Maa            | Tee    | Kahvi   |
|----------------|--------|---------|
| Austria        | 0.3 kg | 6.1 kg  |
| Canada         | 0.2 kg | 6.5 kg  |
| Denmark        | 0.2 kg | 8.7 kg  |
| Egypt          | 1.1 kg | 0.1 kg  |
| Finland        | 0.3 kg | 12.0 kg |
| India          | 0.8 kg | 0.1 kg  |
| Ireland        | 2.0 kg | 3.5 kg  |
| Italy          | 0.1 kg | 5.9 kg  |
| Japan          | 0.9 kg | 3.3 kg  |
| Netherlands    | 0.8 kg | 8.4 kg  |
| Norway         | 0.4 kg | 9.9 kg  |
| Sweden         | 0.4 kg | 8.2 kg  |
| Switzerland    | 0.4 kg | 7.9 kg  |
| Turkey         | 2.1 kg | 0.4 kg  |
| United Kingdom | 2.2 kg | 2.8 kg  |
| United States  | 0.2 kg | 4.2 kg  |

2. Kuvaile kahvin kulutuksen jakauma sopivan taulukon, kuvion ja tunnuslukujen avulla. Tulkitse kuvailu.

3. Mitta-asteikot. Kerro kaikki mitta-asteikot, mitä ominaisuuksia niillä on ja anna pari esimerkkiä jokaisesta mitta-asteikosta. Mitä kuvailumenetelmiä voit käyttää kunkin mitta-asteikon muuttujaa kuvaillessasi.

4. Tutkimuksessa on kysytty 'Mikä on mielestänne perheen ihanteellinen lapsilukumäärä?'. Sukupuolittain saatiin seuraavat tulokset:

| Sukupuoli | n   | keskiarvo | keskihajonta |
|-----------|-----|-----------|--------------|
| Mies      | 445 | 2.42      | 1.15         |
| Nainen    | 470 | 2,69      | 1.13         |

Laske ryhmittäisille keskiarvoille 95%:n luottamusvälit, piirrä näistä kuvio ja tulkitse tulos.

5. Korttipakassa on 52 korttia, jotka ovat jaettu 4 maahan (pata, ruutu, risti ja hertta) sekä 13 numeroon (ässistä (=1) kuninkaaseen (=13)). Korttipakasta vedetään kaksi korttia.

Millä todennäköisyydellä

- molemmat ovat patoja
- ainakin toinen on pata
- molemmat ovat ässiä.

Kaavoja:

$$\bar{x} = \frac{1}{n} \sum x_i, s = \sqrt{\frac{1}{n-1} \left[ \sum x_i^2 - \frac{1}{n} (\sum x_i)^2 \right]}$$

Pearsonin korrelaatiokerroin

$$r = \frac{n \sum x y - \sum x \sum y}{\sqrt{(n \sum x^2 - (\sum x)^2)(n \sum y^2 - (\sum y)^2)}}$$

$$\chi^2 \text{-testi: } \chi^2 = \sum \sum \frac{(f_{ij} - e_{ij})^2}{e_{ij}} \sim \chi^2(df), \text{ missä } df = (r-1)(s-1) \text{ ja } e_{ij} = \frac{f_{i\cdot} \cdot f_{\cdot j}}{n}$$

Keskiarvon luottamusväli  $P\left(\bar{x} - z \cdot \frac{s}{\sqrt{n}} < \mu < \bar{x} + z \cdot \frac{s}{\sqrt{n}}\right) = 0.95$  missä  $z$  t-jakaumasta  
df= n-1.

$$\text{Spearmanin järjestyskorrelaatiokerroin: } r_s = 1 - \frac{6 \sum d^2}{n^3 - n}$$