# MAA 8 - Tilastot ja todennäköisyys 24.4.2023 (Pekka)

**VASTAA MONIVALINTAAN (tehtävä 1) JA KOLMEEN TEHTÄVÄÄN TEHTÄVISTÄ 2-6! YHTEENSÄ NELJÄÄN TEHTÄVÄÄN TULEE VASTATA! KOEAIKA 3h.**

**OHJELMISTOT OVAT HETI KÄYTÖSSÄ!**

**Aineistot-osiossa komentoja ohjelmistoihin!**

## Sisällys



## 1. Monivalintatehtävä 12 p.

**Vain yksi vaihtoehto on oikein!**

### 1.1 1 p.

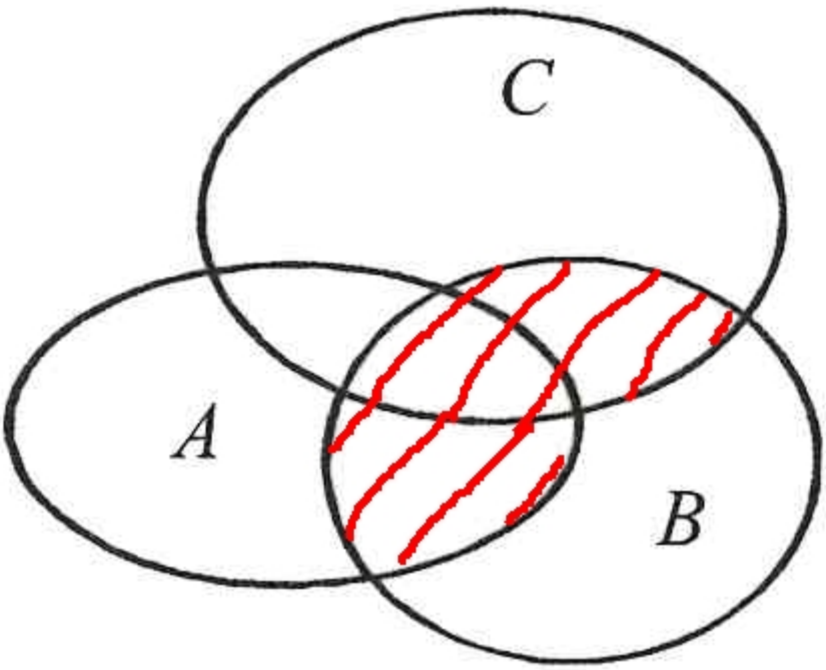
Moodilla tarkoitetaan...

- havaintoaineiston sitä arvoa, jonka frekvenssi on suurin.
- havaintoaineiston geometrista keskiarvoa.
- suuruusjärjestykseen asetetun havaintoaineiston keskimmäistä arvoa tai jos havaintoarvoja on parillinen määrä, niin kahden keskimmäisen arvon keskiarvoa.
- havaintoaineiston keskiarvon ja pienimmän arvon eroa.

***

### 1.2 1 p.

Kuvassa näkyvä viivoitettu alue voidaan joukko-opillisesti kirjoittaa muotoon




- $(A \cap B) \cup (B \cap C) \cup (A \cap C)$
- $(A \cap B) \cup C$
- $A \cap B \cap C$
- $(A \cap B) \setminus C$
- $(A \cap B) \cup (B \cap C)$

***

**1.3** 1 p.

Kolikkoa heitetään kuusi kertaa. Oletetaan, että kolikko jää aina joko kruuna- tai klaava-puoli ylöspäin, eli kolikko ei jää kyljelleen. Määritä tapahtumalle

$$A = \text{Klaavoja on enintään yksi tai vähintään neljä.}$$

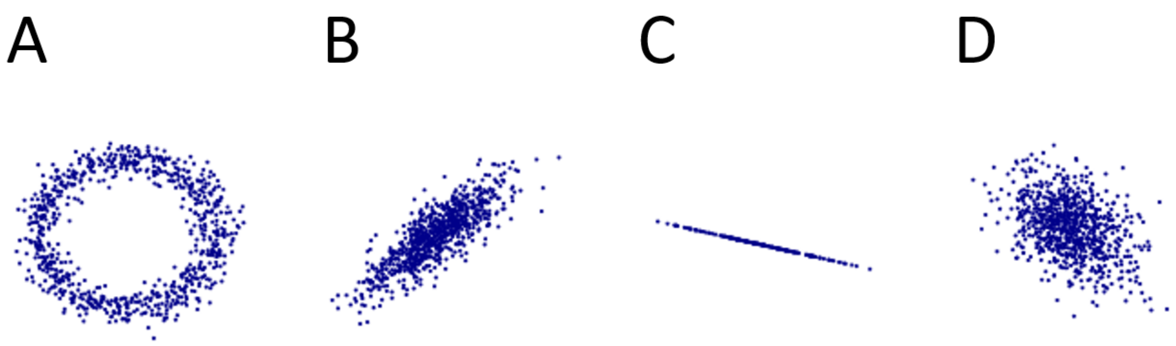
komplementtitapahtuma.

- $\overline{A} = \text{Saadaan nolla, viisi tai kuusi klaavaa.}$
- $\overline{A} = \text{Saadaan vähemmän kuin neljä klaavaa.}$
- $\overline{A} = \text{Saadaan enemmän kuin yksi klaava.}$
- $\overline{A} = \text{Saadaan nolla, kaksi, kolme, viisi tai kuusi klaavaa.}$
- $\overline{A} = \text{Saadaan kaksi tai kolme klaavaa.}$

***

**1.4** 1 p.

Mikä kuvioista vastaa korrelaatiokertoimen $|r|$ arvoa $0{,}4$.




- D
- A
- B
- C

***

**1.5** **1 p.**

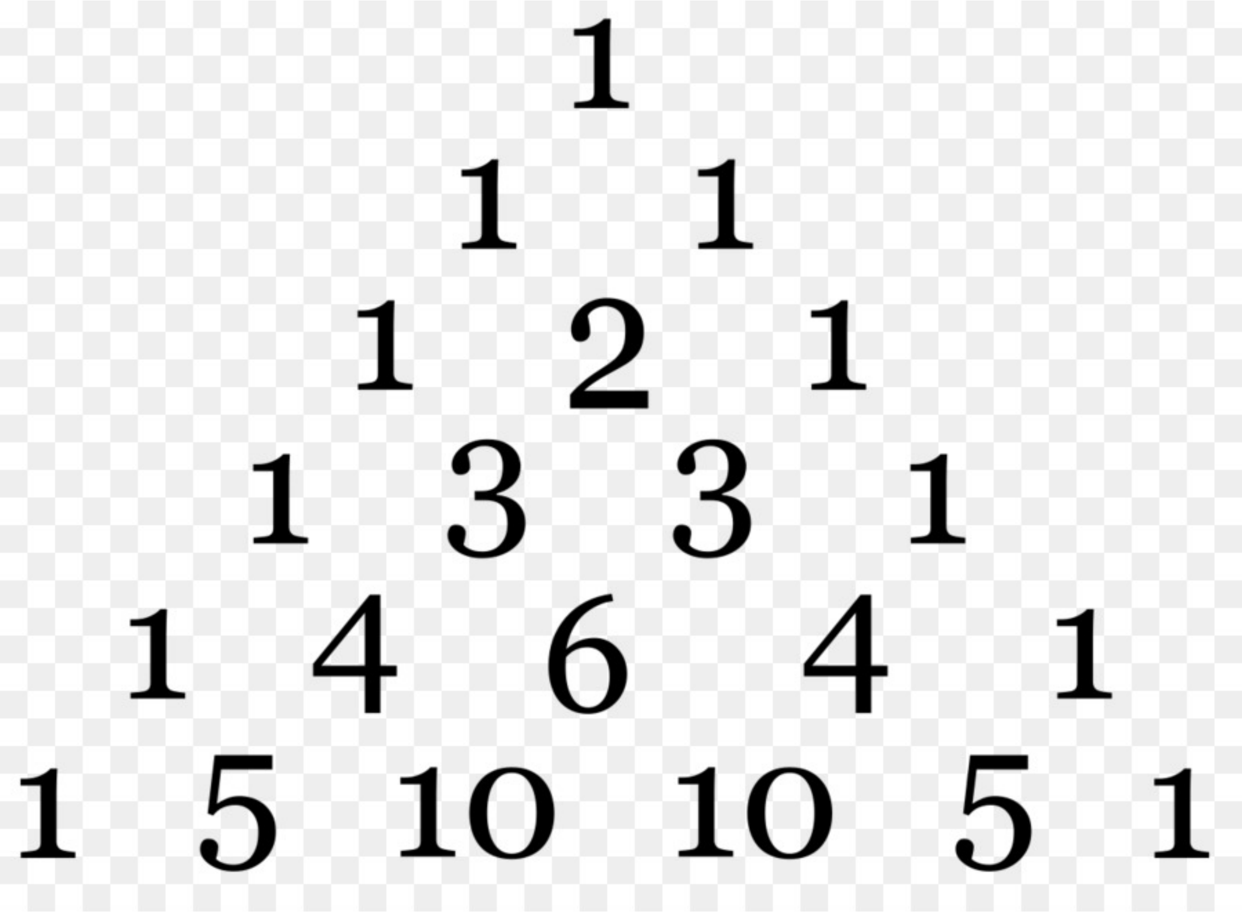
Mikä väittämä on ***väärin***?

- Tapahtuman $A$ geometrinen todennäköisyys on
$$\mathrm{P}(A) = \frac{\text{Tapahtumaa } A \text{ vastaavan kuvion osan geometrinen mitta.}}{\text{Koko kuvion geometrinen mitta.}}\ .$$
- Riippumattomien tapahtumien kertolaskusääntö on
$$\mathrm{P}(A \cap B) = \mathrm{P}(A) \cdot \mathrm{P}(B)\ .$$
- Ehdollinen todennäköisyys on
$$\mathrm{P}(A|B) = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)}\ .$$
- Yleinen yhteenlaskusääntö on
$$\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B)\ .$$

***

**1.6** **1 p.**

Pascalin kolmio liittyy (valitse ***oikea*** vaihtoehto)




- harmonisen keskiarvon juureen
- kovarianssiin
- permutaatioiden lukumääriin
- binomikertoimiin

***

**1.7** 1 p.

Mikä seuraavista vaihtoehdoista tarkoittaa $n$ -alkioisen perusjoukon $k$ -variaatiota? (Huom! $0 < k < n$ .)

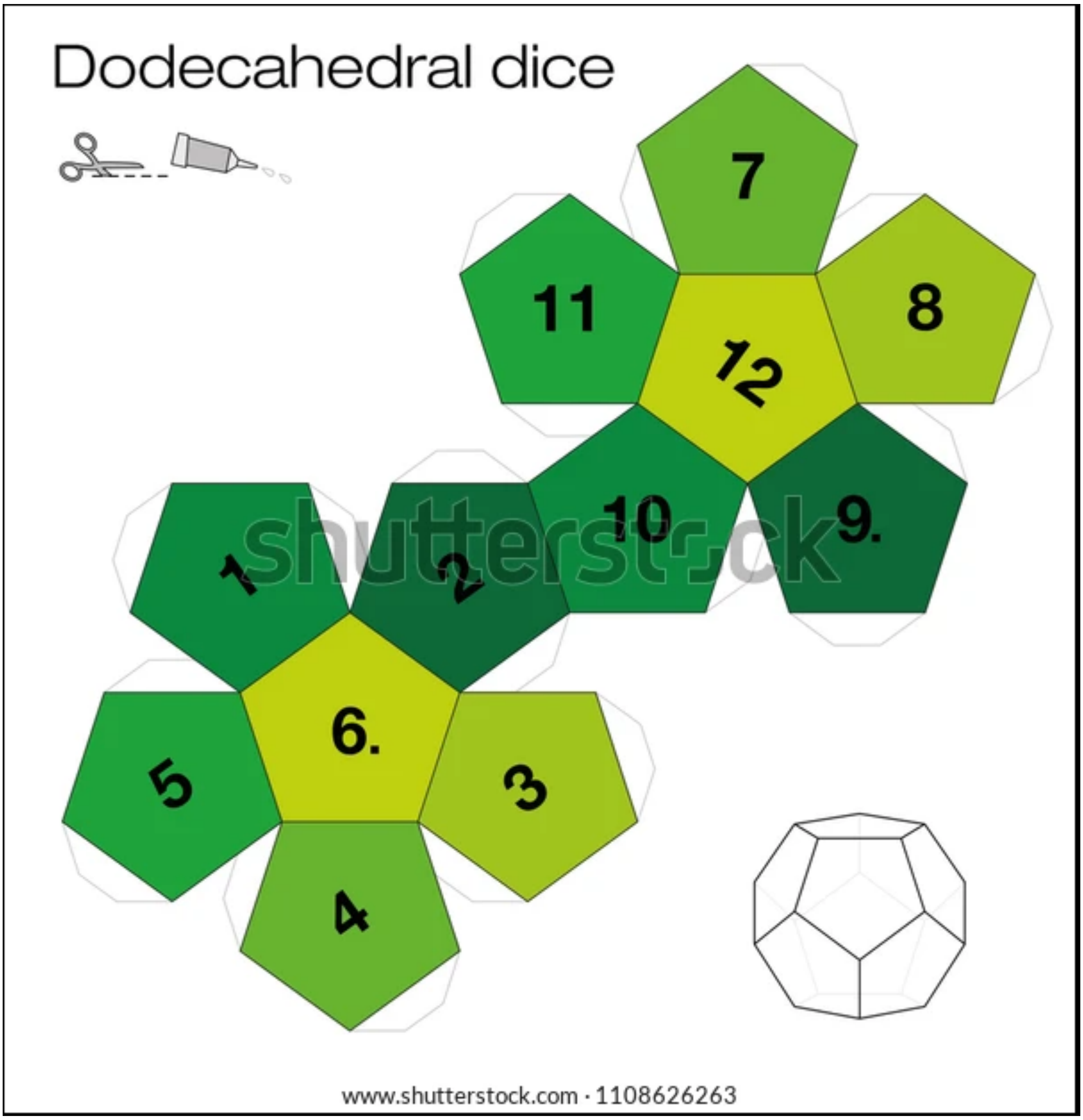
- $\frac{k!}{(n-k)!}$
- $\frac{n!}{(n-k)!k!}$
- $k!$
- $n!$
- $\frac{n!}{(n-k)!}$

***

**1.8** 1 p.

Kaksitoistatahkoista, dodekaedrin muotoista, noppaa heitetään kolme kertaa (kaikkien tahkojen todennäköisyys tulla päällimmäiseksi on sama!). Määritä

$P(\text{Saadaan sama silmäluku kaikilla heitoilla.})$




- $\frac{1}{12}+\frac{1}{12}+\frac{1}{12}$
- $\frac{1}{12}\cdot\frac{1}{12}\cdot\frac{1}{12}$
- $\frac{1}{12^2}$
- $\frac{1}{12}$

***

**1.9** **1 p.**

Laatikossa on $8$ sinistä, $11$ valkoista ja $5$ keltaista samanlaista ja samankokoista palloa. Laatikosta nostetaan sokkona ***yksitellen*** 5 palloa siten, että nostettu pallo ***aina palautetaan*** laatikkoon. Mikä on todennäköisyys saada viisi kertaa **sininen pallo** ?

- $P(\text{Saadaan viisi kertaa sininen pallo.}) = \frac{8}{24}\cdot\frac{8}{23}\cdot\frac{8}{22}\cdot\frac{8}{21}\cdot\frac{8}{20} = \frac{4096}{5313}$
- $P(\text{Saadaan viisi kertaa sininen pallo.}) = 5\cdot\frac{8}{24} = \frac{5}{3}$

- $P(\text{Saadaan viisi kertaa sininen pallo.}) = \frac{\binom{8}{5}}{\binom{24}{5}} = \frac{1}{759}$
- $P(\text{Saadaan viisi kertaa sininen pallo.}) = \left(\frac{1}{3}\right)^5 = \frac{1}{243}$

***

**1.10** 1 p.

Laatikossa on $8$ sinistä, $11$ valkoista ja $5$ keltaista samanlaista ja samankokoista palloa. Laatikosta nostetaan sokkona ***yksitellen*** 5 palloa siten, että nostettua palloa ***ei palauteta*** laatikkoon. Mikä on todennäköisyys saada viisi kertaa **sininen pallo** ?

- $P(\text{Saadaan viisi kertaa sininen pallo.}) = \frac{\binom{24}{5}}{\binom{8}{5}} = \frac{759}{1}$
- $P(\text{Saadaan viisi kertaa sininen pallo.}) = \left(\frac{1}{3}\right)^5 = \frac{1}{243}$
- $P(\text{Saadaan viisi kertaa sininen pallo.}) = \frac{8}{24} \cdot \frac{7}{23} \cdot \frac{6}{22} \cdot \frac{5}{21} \cdot \frac{4}{20} = \frac{1}{759}$
- $P(\text{Saadaan viisi kertaa sininen pallo.}) = 5 \cdot \frac{8}{24} = \frac{5}{3}$

***

**1.11** 1 p.

Laatikossa on $8$ sinistä, $11$ valkoista ja $5$ keltaista samanlaista ja samankokoista palloa. Laatikosta nostetaan sokkona ***kerralla*** 5 palloa. Mikä on todennäköisyys, että kaikki pallot ovat **sinisiä** ?

- $P(\text{Kaikki pallot ovat sinisiä.}) = \frac{8}{24} \cdot \frac{8}{23} \cdot \frac{8}{22} \cdot \frac{8}{21} \cdot \frac{8}{20} = \frac{4096}{5313}$
- $P(\text{Kaikki pallot ovat sinisiä.}) = \frac{1}{3}$
- $P(\text{Kaikki pallot ovat sinisiä.}) = \frac{\binom{8}{5}}{\binom{24}{5}} = \frac{1}{759}$
- $P(\text{Kaikki pallot ovat sinisiä.}) = \left(\frac{1}{3}\right)^5 = \frac{1}{243}$

***

**1.12** 1 p.

Mikä väittämä on **väärin**.

- Tilastomuuttujien $X$ ja $Y$ kovarianssi $s_{XY}$ eli yhteisvaihtelu kuvaa havaintopisteiden $(x_i, y_i)$ sijoittumista keskiarvoihin $\overline{x}$ ja $\overline{y}$ nähden.
- $n$ -alkioisen perusjoukon $k$ -alkioiset $(0 \leq k \leq n)$ kombinaatiot kuvaavat osajoukkojen määrää.
- Tilastomuuttujien $X$ ja $Y$ korrelaatiokerroin saadaan jakamalla kovarianssi $s_{XY}$ tilastomuuttujien $X$ ja $Y$ keskihajontojen $s_X$ ja $s_Y$ tulolla.
- Muuttujan arvoa $x$ vastaa normitettu arvo $z = \frac{x - s_n}{\overline{x}}$ , missä $\overline{x}$ on keskiarvo ja $s_n$ keskihajonta.

## 2. Aineiston tutkimista ja päättelyn tekoa 12 p.

Aineistot-osiossa on 189 eri valtion bruttokansantuotteet asukasta kohti ja lasten koulunkäyntivuodet.
**a)** Kumpi tilastomuuttuja mielestäsi selittää paremmin toista? Perustele lyhyesti. (1p)
**b)** Muodosta hajontakuviot (2 kpl), jossa koulunkäyntivuodet selittää bruttokansantuotetta ja sovita pisteisiin *lineaarisen* mallin lisäksi *eksponentiaalinen* malli. Muodosta siis kaksi kertaa sama pistekuvio ja molempiin eri sovitus selkeyden vuoksi. ***Mallien yhtälöt tulee olla näkyvissä***. Kumpi malli vaikuttaisi paremmalta? (4p)
OHJE: Kannattanee käyttää CALC:a, erityisesti "Lisää kaavio" -ohjattua toimintoa pistejoukon mallien luomisessa. Katso myös aineistot-osio!
**c)** Määritä vielä erikseen lineaarisen mallin $y = bx + a$ kulmakertoimessa $b = \frac{s_{XY}}{(s_X)^2}$ ja vakiotermissä $a = \overline{y} - b\overline{x}$ esiintyvät termit. Käytä jälleen CALC:n suoria komentoja. Vertaa arvoja **b)**-kohdassa sovittamaasi lineaarisen mallin arvoihin. (Saitko samat?) (3p)
**d)** Kirjoita/Laske selitysasteet $R^2$ näkyviin ja tulkitse niitä? (2p)
**e)** Miten eksponentiaalinen malli ennustaa valtion bruttokansantuotteen muuttuvan, jos koulunkäyntivuodet kasvaa yhdeksästä vuodesta (ns. peruskoulu) kahteentoista vuoteen (ns. toinen aste)? (2p)

**Aineisto**

2.A BKT ja koulunkäyntivuodet

## 3. Todennäköisyyksien määrittämistä 12 p.

**Hyödynnä konetta kaikissa alakohdissa!**

**a)** Todennäköisyys, että erään tulppaanilajikkeen sipuli itää on $0{,}6$. Kuinka monta sipulia on vähintään istutettava, jotta niistä ainakin kolme itäisi yli $99\ \%$ todennäköisyydellä? Muista, kone laskee, mutta laskukaava tulee olla vastauksessa. (6p)

**b)** Pekka, Lasse ja Juho pelaavat peliä. Pekan voiton todennäköisyys on $0{,}3$, lassen $0{,}35$ ja Juhon $0{,}2$. Kun peliä on pelattu neljä kierrosta, Pekalla ja Lassella on yksi voitto ja Juholla kaksi voittoa. Kuinka suurella todennäköisyydellä Pekka tällöin saavuttaa ensimmäisenä kolme voittoa? (4p)

**c)** Uimahallissa aaltokone luo viiden minuutin ajan aaltoja aina puolelta ja tasatunnein. Altaaseen, jossa aaltokone on, saapuu uimari satunnaiseen aikaan kelloon katsomatta. Kuinka suurella todennäköisyydellä uimari **i)** sattuu altaaseen aaltojen aikaan **ii)** joutuu odottamaan aaltojen saapumista yli $15$ minuuttia. (2p)

## 4. Todennäköisyyden laskusääntöjä sekä kombinatoriikkaa 12 p.

**Hyödynnä konetta kaikissa alakohdissa!**

**a)** Kaksi opiskelijaa, Seppo-Gustaf ja Siiri-Irmeli, suunnittelevat koulun kevätjuhlaan ohjelmarunkoa. Aikaisemmista tilanteista tiedetään, että Seppo-Gustaf on onnistunut vastaavissa tilanteissa todennäköisyydellä $5/9$, Siiri-Irmeli todennäköisyydellä $5/13$ ja ainakin toinen todennäköisyydellä $2/3$.

- Millä todennäköisyydellä molemmat opiskelijat onnistuvat? (1p)

- Entä millä todennäköisyydellä vain toinen opiskelija onnistuu? (1,5p)

- Jos vain toinen opiskelija onnistuu, millä todennäköisyydellä se on Seppo-Gustaf? (1,5p)

Ohje: Venn-diagrammilla kannattaa lähteä liikkeelle. Voit halutessasi liittää sen ratkaisuusi.

**b)** Sienikurssilla opetettiin tunnistamaan $78$ erilaista sientä, joista kurssilainen oppi kuitenkin vain $49$. Kuinka suurella todennäköisyydellä hän tunnisti oikein satunnaisesti esitetyt kuusi erilaista kurssilla opetettua sientä? [YO s1996/7] (4p)

**c)** Pekka heittää viidestä vapaaheitosta koriin yhtä todennäköisesti kolme kuin kahdeksasta vapaaheitosta kuusi. Millä todennäköisyydellä Pekka saa yksittäisen vapaaheiton koriin? (4p)

## 5. Pieniä erillisiä tehtäviä 12 p.

**a)** Ville-Ensio ja Mauno-Taavetti osallistuivat pituushyppykilpailuun ja saivat oheisen taulukon mukaiset tulokset.

| | | | | |
|---|---|---|---|---|
| **Ville-Ensio:** | 5,44 | 5,85 | 6,05 | 5,20 |
| **Mauno-Taavetti:** | 5,62 | 5,56 | 5,70 | 6,02 |

Kumman paras hyppy poikkesi suhteessa enemmän keskiarvosta? (3p)

**b)** Euron kolikko (halkaisija $23{,}25$ mm) tipahtaa shakkilaudalle, jonka ruutujen sivujen pituus on $50$ mm. Millä todennäköisyydellä kolikko **i)** on kokonaan jonkin ruudun sisällä, **ii)** peittää jonkin ruudun nurkan? Perustele. (4p) Vihje: Kannattaa piirtää kuva tai kuvia.

**c)** Tapahtumille $A$ ja $B$ on $P(A) = \frac{1}{5}$, $P(B) = \frac{1}{2}$ ja $P(A \cup B) = \frac{7}{10}$. Laske

$$\textbf{i)} \quad P(A|B) \qquad\qquad \textbf{ii)} \quad P(B|\overline{A})\ .$$

(5p)

## 6. Koulumatkojen tarkastelua 12 p.

Aineistossa on erään suuren koulun oppilaiden koulumatkat (pyöristettynä 100 m tarkkuudella).

**a)** Määritä keskiarvo, moodi, mediaani sekä populaatiokeskihajonta käyttäen taulukkolaskenta-ohjelmaa (saa valita joko Libre Calc tai Geogebra). Liitä vastaukseen mukaan kuva käytetystä komennosta ohjelman syöttöriviltä.) Onko koulumatka diskreetti vai jatkuva muuttuja? (2p)

**b)** Luokittele aineisto 2000 m luokkiin (100 m tarkkuudella eli ensimmäinen luokka on 0 – 1900), määritä luokkakeskukset, luokkien frekvenssit, summafrekvenssit, suhteelliset frekvenssit ja suhteelliset summafrekvenssit. Havainnollista tilannetta pystypylväskaaviolla, jossa pylväs on luokkakeskuksen kohdalla ja korkeus on kyseisen luokan frekvenssin arvo. (5p)
Muista lajitella aineisto ensin, käytä frekvenssien määrittämiseksi seuraavan kaltaista komentoa:
=LASKE.JOS(tietoalue;ehto)-LASKE.JOS(tietoalue;ehto)
esimerkiksi =LASKE.JOS(A2:A1025;">0")-LASKE.JOS(A2:A1025;">1000")

**c)** Piirrä kertymäkuvaaja ja arvioi sen perusteella prosentuaalinen määrä niistä oppilaista, joiden koulumatkat on enintään 6000 m. Arvioi kertymäkuvaajaa käyttäen kuinka pitkä koulumatka (100 m tarkkuudella) vähintään on niillä oppilailla, jotka kuuluvat pisimmän 15 % joukkoon. (5p)

**Aineisto**

6.A Koulumatkat

*Kokeen tehtävät loppuvat tähän.*

Siirry tarkastelemaan vastauksiasi

Tarkastelun jälkeen voit vielä palata muokkaamaan vastauksia, tai päättää kokeen.