Matematiikan koe, Maa10 – Todennäköisyys ja tilastot – RATKAISUT **Sievin lukio**
Perjantai 31.5.2019
VASTAA YHTEENSÄ VIITEEN TEHTÄVÄÄN SITEN, ETTÄ OLET VASTANNUT TEHTÄVIIN 1 JA 2.

A-OSA

1. MONIVALINTA – Vain yksi vaihtoehto on oikein!

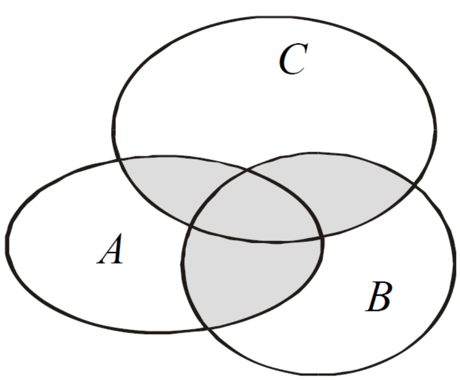
1.1 Moodilla tarkoitetaan...

- ○ suuruusjärjestykseen asetetun havaintoaineiston keskimmäistä arvoa tai jos havaintoarvoja on parillinen määrä, niin kahden keskimmäisen arvon keskiarvoa.
- ◉ havaintoaineiston sitä arvoa, jonka frekvenssi on suurin.
- ○ havaintoaineiston suurinta arvoa.
- ○ havaintoaineiston keskiarvon ja pienimmän arvon eroa.
- ○ havaintoaineiston sitä arvoa, jolle suhteellinen summafrekvenssi ylittää 50 %.

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

- ○ $k!$
- ○ $\frac{k!}{(n-k)!}$
- ○ $\frac{n!}{(n-k)!\,k!}$
- ○ $n!$
- ◉ $\frac{n!}{(n-k)!}$
- ○ $\frac{k!}{(n-k)!\,k!}$

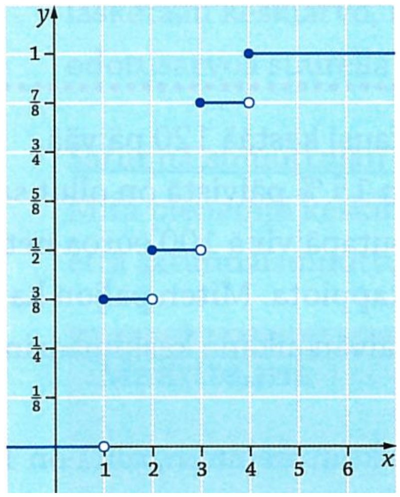
1.3 Kuvassa näkyvä alue voidaan joukko-opillisesti kirjoittaa muotoon



- ○ $A \cap B \cap C$
- ○ $(A \cap B) \cup C$
- ○ $(A \cap B) \cup (B \cap C)$
- ○ $A^c \cup (B \cap C)$
- ○ $(A \cap B) \setminus C$
- ◉ $(A \cap B) \cup (B \cap C) \cup (A \cap C)$

1.4 Kuvassa on satunnaismuuttujan X kertymäfunktion kuvaaja. Määritä pistetodennäköisyys

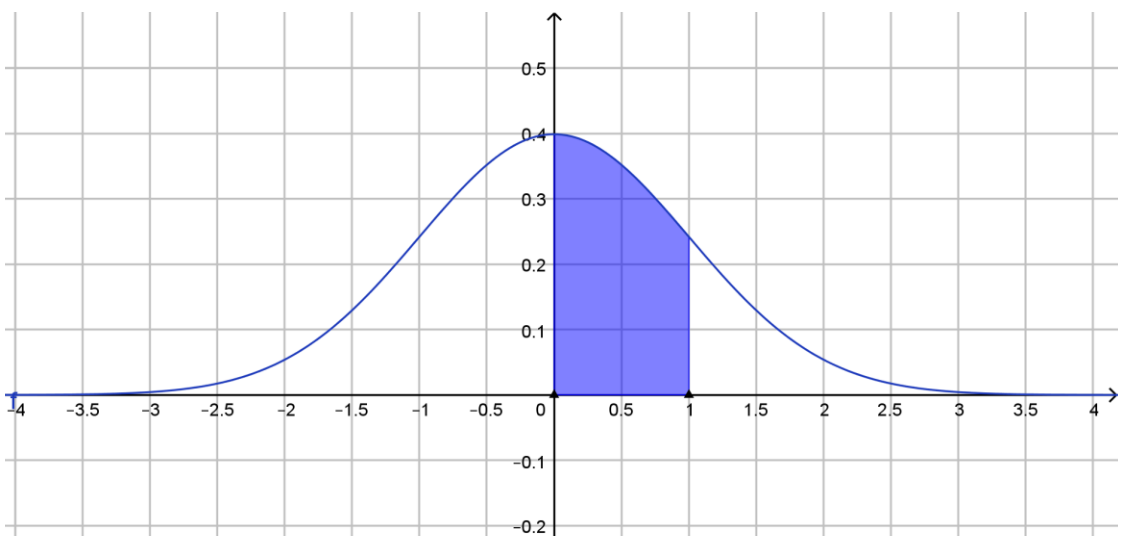
$$P(X = 3).$$



- $P(X = 3) = 3$
- $P(X = 3) = \frac{7}{8}$
- $P(X = 3) = \frac{2}{3}$
- $P(X = 3) = \frac{1}{2}$
- $P(X = 3) = \frac{3}{8}$

1.5 Määritä $P(0 \leq X \leq 1)$, kun $X \sim N(0,1)$ ja tiedetään, että

$$P(X \leq -1) = 0{,}1587.$$



- $P(0 \leq X \leq 1) = 0{,}6826$
- $P(0 \leq X \leq 1) = 0{,}3174$
- $P(0 \leq X \leq 1) = 0{,}3413$
- $P(0 \leq X \leq 1) = 0{,}1826$
- $P(0 \leq X \leq 1)$ ei voida määrittää annetuilla tiedoilla

1.6 Normaalia kuusitahkoista noppaa heitetään kolme kertaa. Määritä

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

- ○ $P(\text{"..."}) = \frac{1}{6}$
- ● $P(\text{"..."}) = \frac{1}{6^2}$
- ○ $P(\text{"..."}) = \frac{1}{6^3}$
- ○ $P(\text{"..."}) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6}$

1.7 Laatikossa on sinistä, valkoista ja keltaista samanlaista ja samankokoista palloa. Laatikosta nostetaan sokkona yksitellen 4 palloa siten, että nostettu pallo aina palautetaan laatikkoon. Mikä on todennäköisyys saada 4 kertaa **sininen pallo**?

- ○ $P(\text{"..."}) = 4 \cdot \frac{8}{18} = \frac{16}{9}$
- ○ $P(\text{"..."}) = \frac{\binom{8}{4}}{\binom{18}{4}} = \frac{70}{3060}$
- ● $P(\text{"..."}) = \left(\frac{8}{18}\right)^4 = \left(\frac{4}{9}\right)^4$
- ○ $P(\text{"..."}) = \frac{8}{18} \cdot \frac{7}{17} \cdot \frac{6}{16} \cdot \frac{5}{15} = \frac{1680}{73440}$

1.8 Laatikossa on sinistä, valkoista ja keltaista samanlaista ja samankokoista palloa. Laatikosta nostetaan sokkona yksitellen 4 palloa siten, että nostettua palloa ei palauteta takaisin laatikkoon. Mikä on todennäköisyys saada 4 kertaa **sininen pallo**?

- ○ $P(\text{"..."}) = 4 \cdot \frac{8}{18} = \frac{16}{9}$
- ○ $P(\text{"..."}) = \frac{\binom{8}{4}}{\binom{18}{4}} = \frac{70}{3060}$
- ○ $P(\text{"..."}) = \left(\frac{8}{18}\right)^4 = \left(\frac{4}{9}\right)^4$
- ○ $P(\text{"..."}) = \frac{8}{18} \cdot \frac{7}{17} \cdot \frac{6}{16} \cdot \frac{5}{15} = \frac{1680}{73440}$

1.9 Laatikossa on sinistä, valkoista ja keltaista samanlaista ja samankokoista palloa. Laatikosta nostetaan sokkona kerralla 4 palloa. Mikä on todennäköisyys, että kaikki pallot ovat **sinisiä**?

- ○ $P(\text{"..."}) = 4 \cdot \frac{8}{18} = \frac{16}{9}$
- ○ $P(\text{"..."}) = \left(\frac{8}{18}\right)^4 = \left(\frac{4}{9}\right)^4$

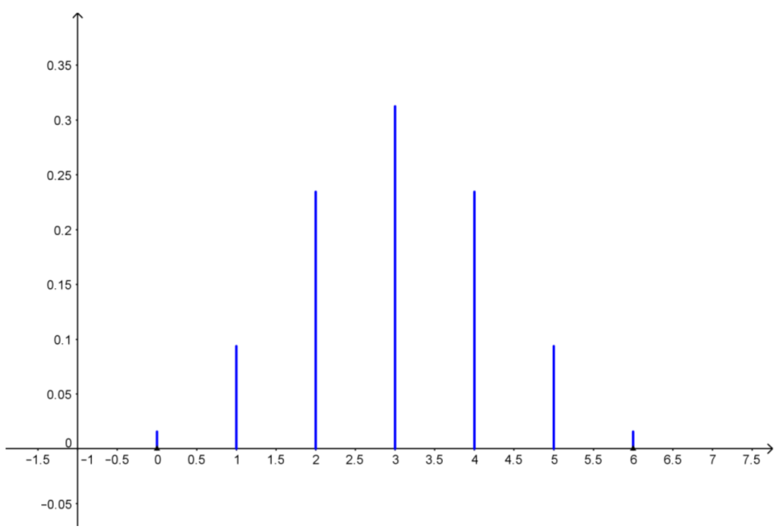
$$P(\text{"..."}) = \frac{\binom{8}{4}}{\binom{18}{4}} = \frac{70}{3060}$$

$$P(\text{"..."}) = \frac{8}{18} \cdot \frac{7}{17} \cdot \frac{6}{16} \cdot \frac{5}{15} = \frac{1680}{73440}$$

1.10 Kolikkoa heitetään viisi kertaa. (Oletetaan, että kolikko jää aina joko kruuna- tai klaava-puoli ylöspäin, eli kolikko ei jää kyljelleen.) Määritä tapahtumalle $A =$ "Klaavoja on enintään yksi tai vähintään neljä." komplementtitapahtuma.

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

1.11 Kuvassa on annettu erään diskreetin satunnaismuuttujan X pistetodennäköisyysfunktio. Mikä seuraavista vaihtoehdoista on **väärin**?



- Satunnaismuuttuja on binomijakautunut.
- Satunnaismuuttuja ei ole normaalijakautunut.
- $P(X > 2) \approx 0{,}66$
- $P(X \leq 3) \approx 0{,}66$
- $P(X < 3) \approx 0{,}66$

1.12 Pieni kesälomatehtävä viimeiseksi :)

Tiedetään, että $X \sim \mathrm{Gamma}(\nu, \lambda)$. Tällöin X:n tiheysfunktio f on muotoa.

- ( ) $$f(x) = \frac{1}{\lambda\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{X-\nu}{\lambda}\right)^2}$$
- (•) $$f(x) = \frac{\lambda^\nu}{\Gamma(\nu)} \cdot x^{\nu-1} e^{-\lambda x}$$
- ( ) $$f(x) = \frac{\lambda^i}{i!} \cdot e^{-\lambda}$$
- ( ) $$f(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi} \cdot \Gamma\left(\frac{\nu}{2}\right)} \cdot \left(1 + \frac{x^2}{\lambda}\right)^{-\frac{\nu+1}{2}}$$

B-OSA

2. a) **Selitä lyhyesti käsite satunnaismuuttuja ja ilmaise matemaattisesti kuinka havaintoarvo x_i normitetaan, kun keskiarvo $\overline{x}$ ja keskihajonta s ovat tunnettuja. (4p)**

b) Avaa aineistot-osiosta *tehtava2.ods*-taulukkolaskentadokumentti ja vastaa sieltä saatavien tietojen perusteella seuraaviin kysymyksiin. Vastaukset voi tehdä myös esim. Geogebralla. (8p)

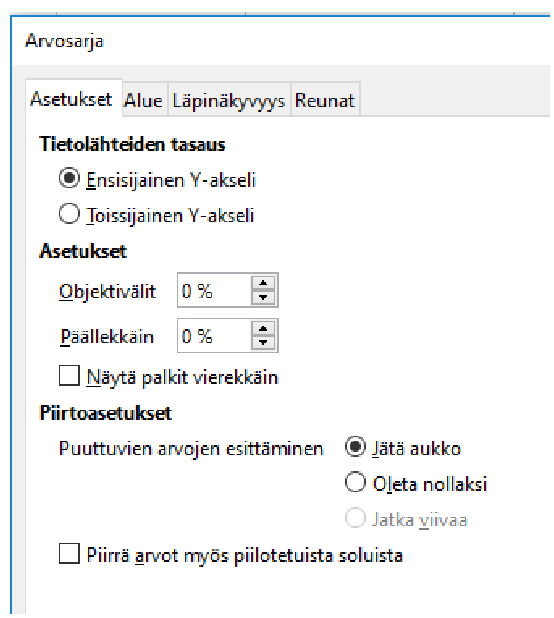
i) Määritä vuoden 2017 aineistosta vuoteiden lukumäärien keskiarvo ja mediaani.

ii) Jaa vuoden 2018 aineistosta huoneiden keskihinnat *neljään tasaväliseen luokkaan*, määritä luokkakeskukset ja muodosta histogrammi niin, että pylvään korkeus kuvaa ko. luokkaan kuuluvien huoneiden lukumäärää sekä pylvään keskikohtana on edellä määrittämäsi luokkakeskukset. Tulisi näyttää tältä:

luokka	luokkarajat	luokkakeskus	lukumäärät yht.
1	62,65-78,1555	70,4275	1746
2			
3			
4			

Huom! Muistapa lajitella aineisto.

Ohjelmisto-ohjeita: Hyödynnä Libre Calc:ssa pikavalintoja  **. Huomaa, että valitse vain haluamasi asia taulukosta jota haluat lajitella. Pylväät sait yhteen, kun painat pylvään kohdalla hiiren kakkospainiketta ja valitset "Muotoile arvosarjoja → Asetukset välilehti → Objektiväli 0%".**



Hyödynnä Geogebralla Taulukkolaskentaa ja Yhden muuttujan analyysia.

a) Satunnaismuuttuja

Satunnaismuuttuja on otosavaruudessa määritelty *reaaliarvoinen funktio* $\underline{x}: \boldsymbol{E} \rightarrow \mathbb{R}$.

Havaintoarvon x_i normitus: vähennetään havaintoarvosta keskiarvo ja jaetaan erotus keskihajonnalla. Siis,

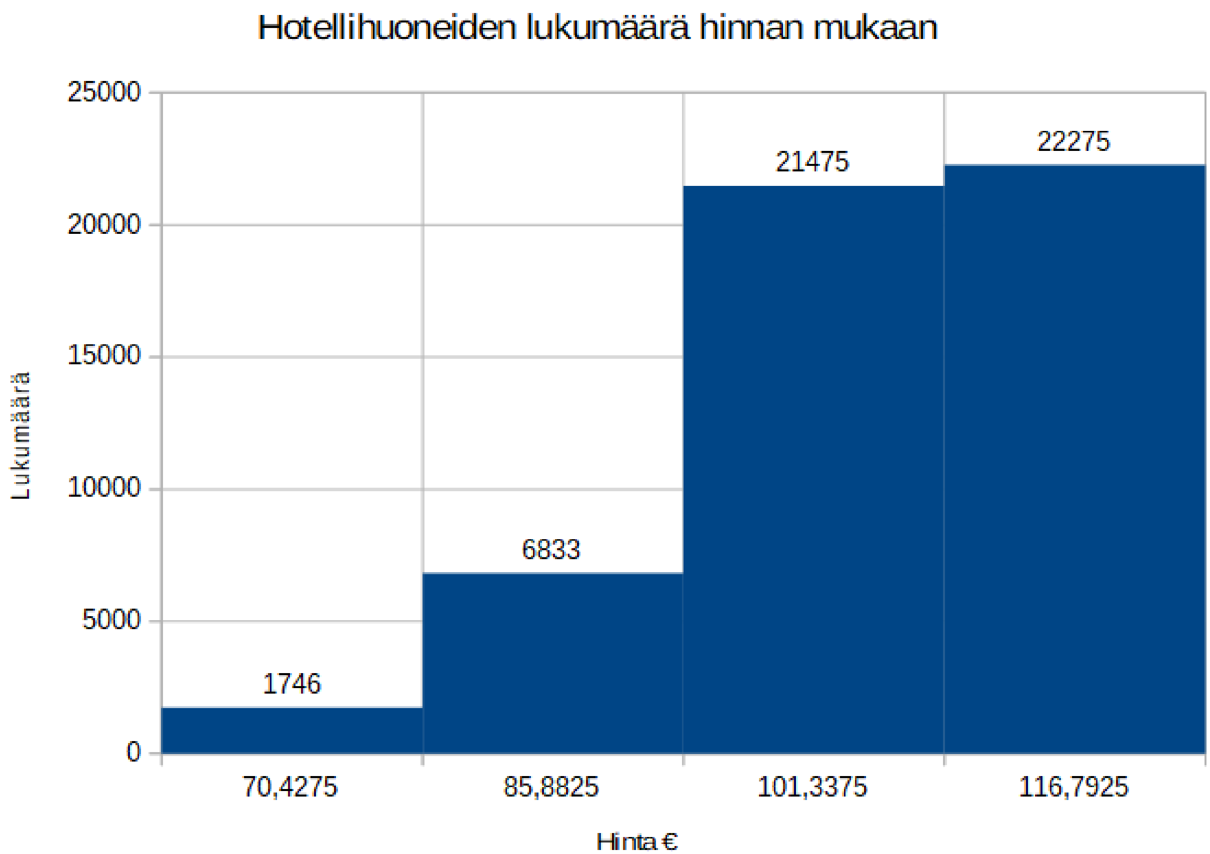
$$z_i = \frac{x_i - \overline{x}}{s}.$$

b) i) Käytetään Calc:n omia komentoja. Vuoteiden keskiarvoksi saadaan 6270 ja mediaaniksi 3871.

Hotellien keskimääräinen vuosikapasiteetti maakunnittain, lähde tilastokeskus (www)			
	2017		
MAAKUNTA	Huoneet, lkm	Huoneen keskihinta	Vuoteet, lkm
Uusimaa	15603	111,87	30261
Varsinais-Suomi	2870	93,73	5674
Satakunta	1117	89,89	2208
Kanta-Häme	1307	82,08	2669
Pirkanmaa	3916	100,37	7964
Päijät-Häme	1725	93,12	3547
Kymenlaakso	787	92,22	1534
Etelä-Karjala	1549	93,08	4195
Etelä-Savo	1631	85,62	3458
Pohjois-Savo	2334	93,57	5471
Pohjois-Karjala	1179	91,28	2787
Keski-Suomi	2703	92,19	7288
Etelä-Pohjanmaa	1502	81,78	3199
Pohjanmaa	1227	94,48	2358
Keski-Pohjanmaa	507	89,41	959
Pohjois-Pohjanmaa	3508	93,44	8042
Kainuu	1736	66,51	5556
Lappi	6234	112,33	15685
		keskiarvo=	6269,7222
		Mediaani =	3871

ii) Saadaan

luokka	luokkarajat	luokkakeskus	lukumäärät yht.
1	62,70-78,155	70,4275	1746
2	78,155-93,61	85,8825	6833
3	93,61-109,065	101,3375	21475
4	109,065-124,52	116,7925	22275
vaihteluväli=	61,82		
Luokkaväli =	15,455		

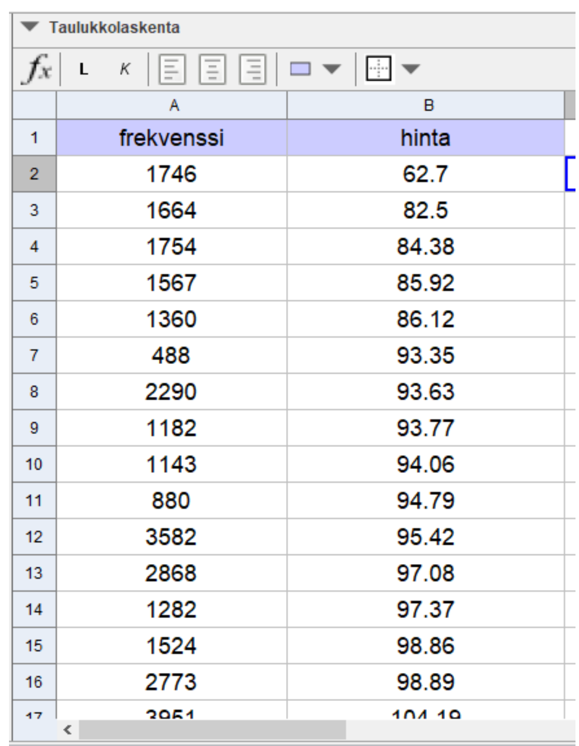


ii) TAI GEOEGBRALLA

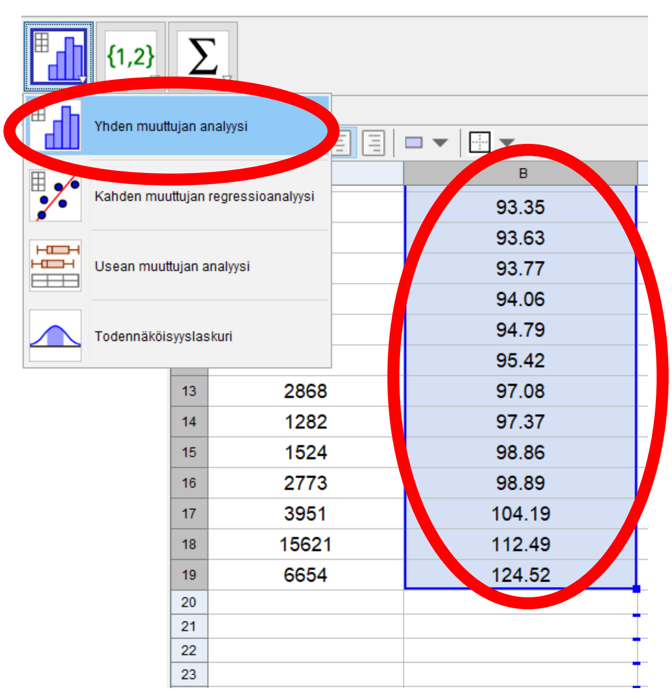
1) Kopioidaan ensin tiedot Geogebran taulukko-osaan.

→ Lisätään otsikoiksi "frekvenssi" ja "hinta"

2) Valitaan "Hinta"-sarakkeen tiedot ja painetaan "Yhden muuttujan analyysi" –painiketta

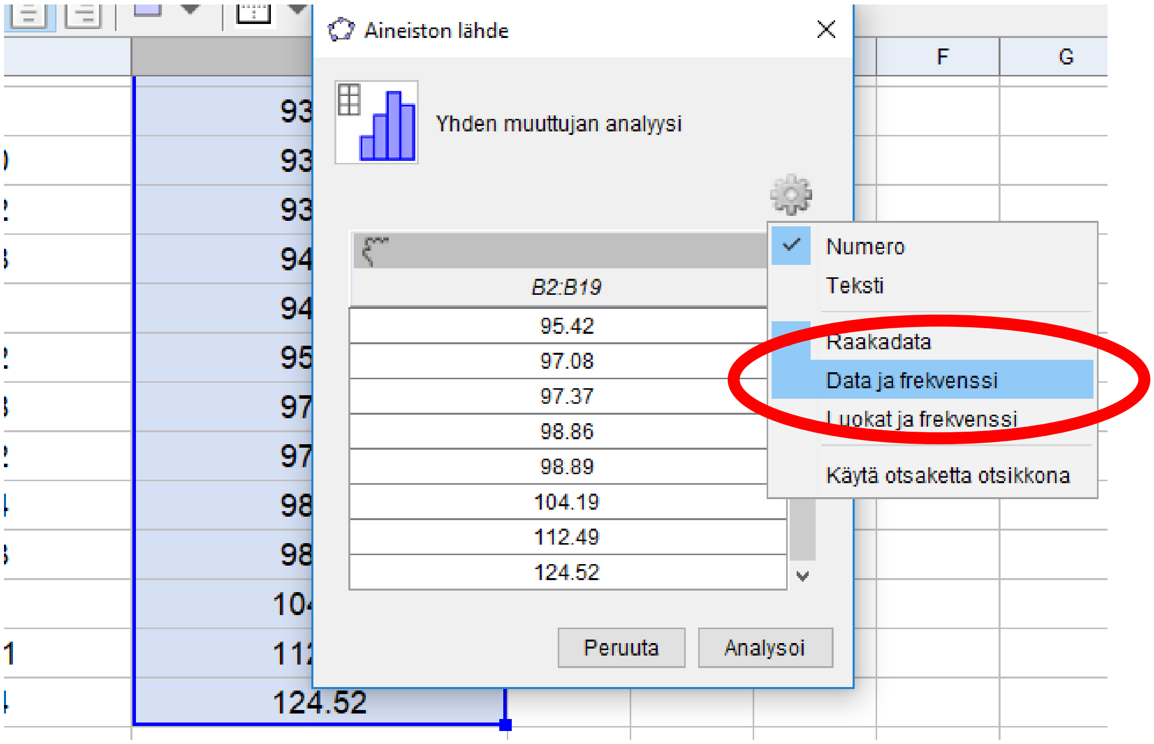

Taulukkolaskenta

	A	B
1	frekvenssi	hinta
2	1746	62.7
3	1664	82.5
4	1754	84.38
5	1567	85.92
6	1360	86.12
7	488	93.35
8	2290	93.63
9	1182	93.77
10	1143	94.06
11	880	94.79
12	3582	95.42
13	2868	97.08
14	1282	97.37
15	1524	98.86
16	2773	98.89

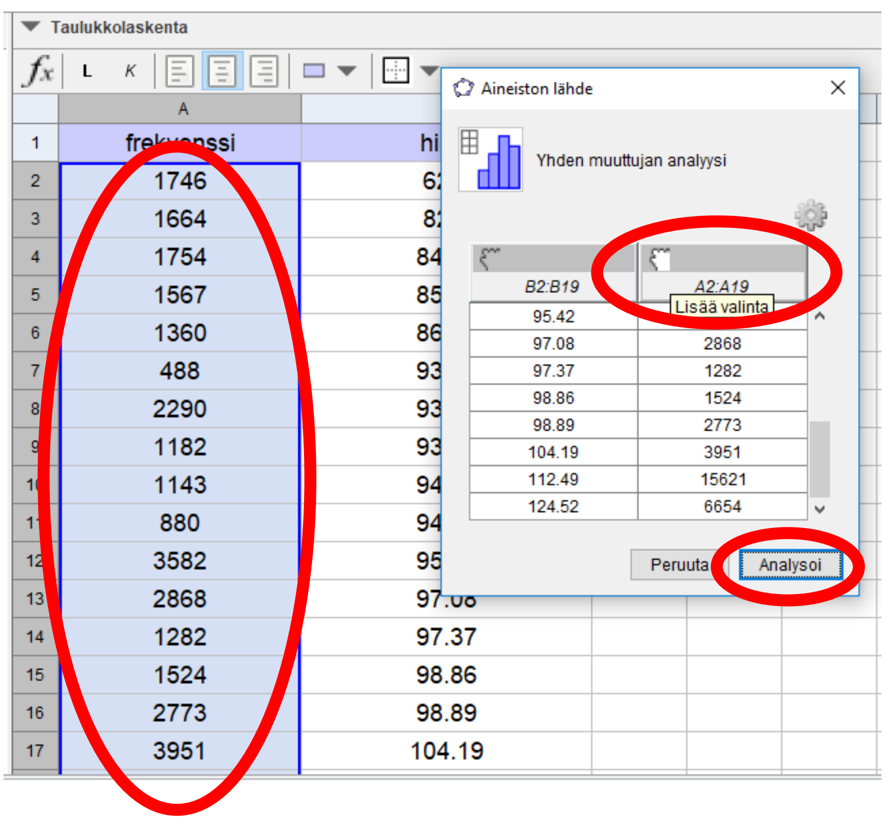


		B
		93.35
		93.63
		93.77
		94.06
		94.79
		95.42
13	2868	97.08
14	1282	97.37
15	1524	98.86
16	2773	98.89
17	3951	104.19
18	15621	112.49
19	6654	124.52
20		
21		
22		
23		

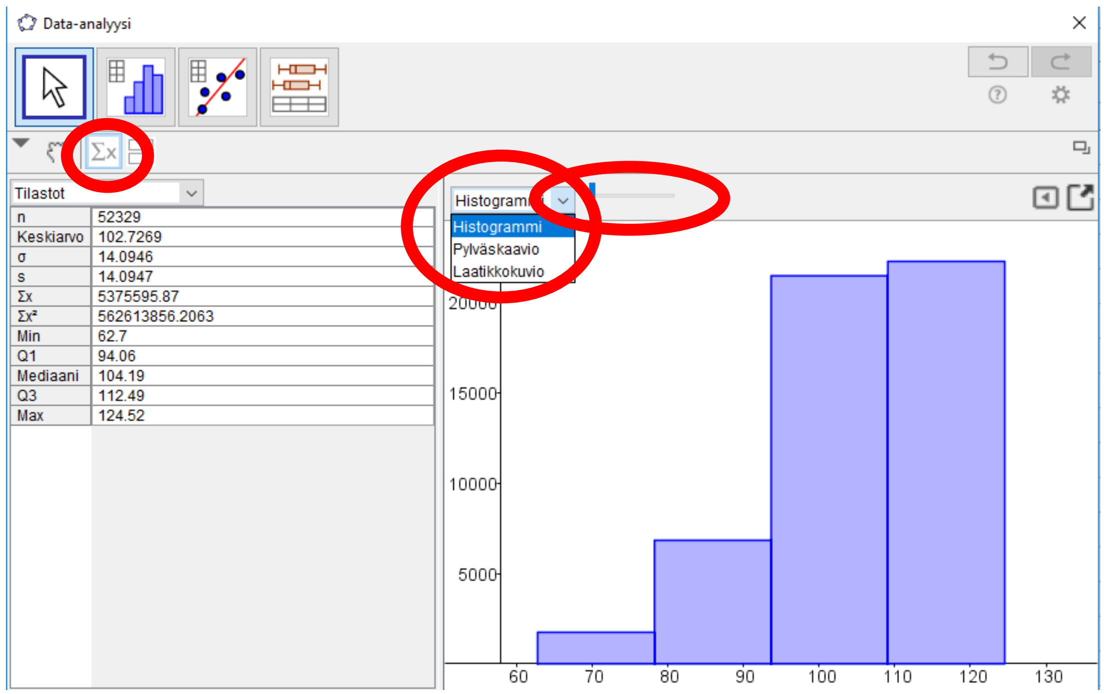
3) Aineiston lähde – ikkunasta valitaan Asetukset (ratas-symboli) – kohdasta Data ja frekvenssi. Tulee toinen sarake.



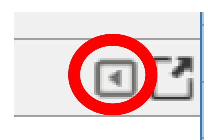
4) Valitaan frekvenssi – sarakkeen tiedot ja painetaan Lisää valinta (käsi-symboli) – kohdasta. Lopuksi painetaan Analysoi-painiketta.



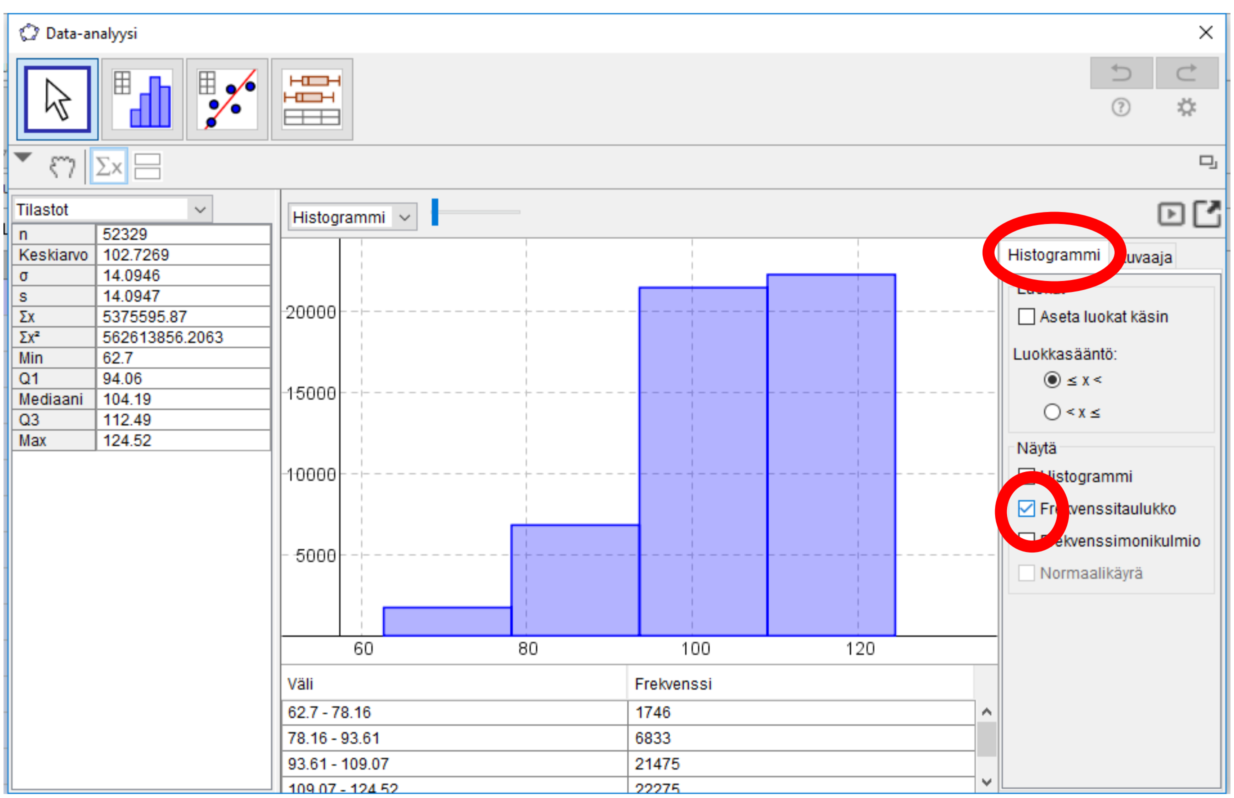
5) Aukeavasta Data analyysi – ikkunasta valitaan Histogrammi (ellei se jo ole valittuna) sekä tehtävänannon mukaisesti liukukytkimeltä 4 luokkaa. Histogrammi on automaattisesti tasavälinen. Halutessasi laita esiin Tilastot – tiedot painamalla symbolia $\sum x$.



6) Histogrammi on muuten valmis, mutta tehtävänannossa pyydetyt välit eli luokkarajat ja lukumäärät eli frekvenssit saat näkyviin painamalla symbolia →



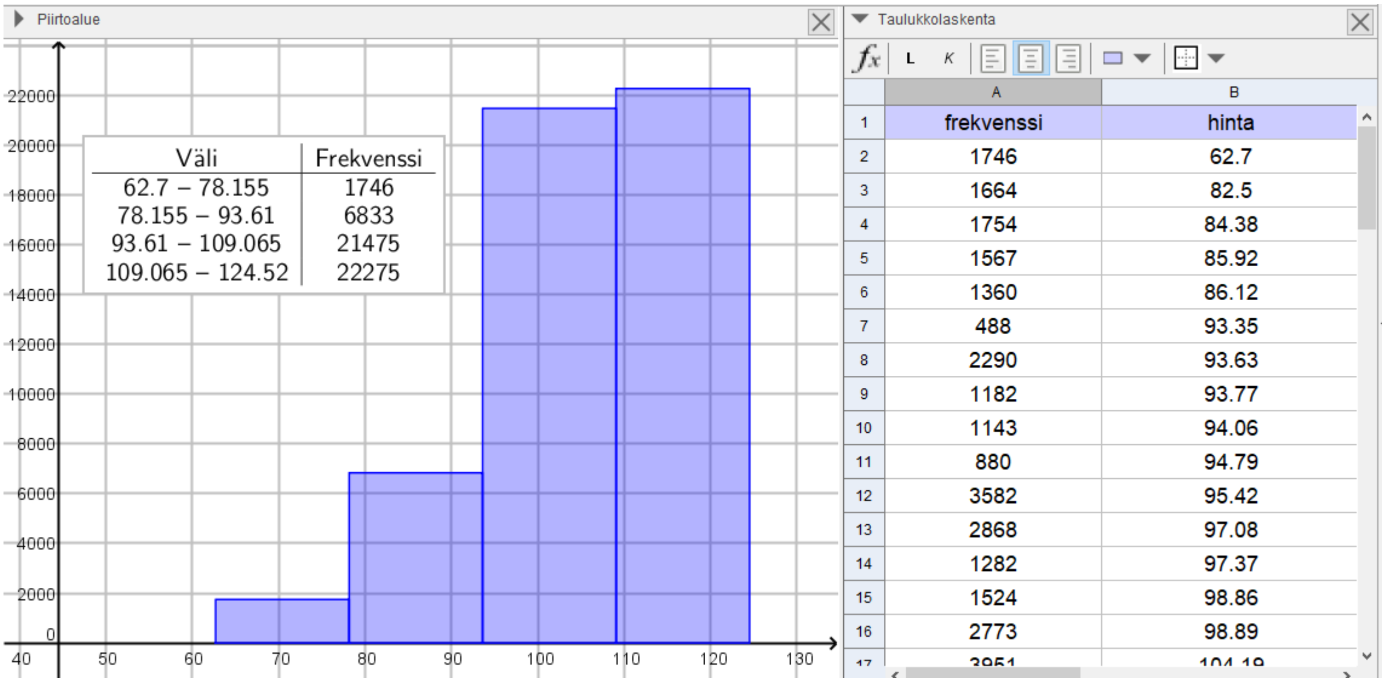
Aukeavasta valikosta valitaan Histogrammi – välilehti ja painetaan Näytä – kohdasta Frekvenssitaulukko



7) Lopuksi vie tiedot piirtoalueelle, sulje muut ikkunat kuin piirtoalue ja ota näyttökuva.

Luokkakeskukset tulee itse laskea.

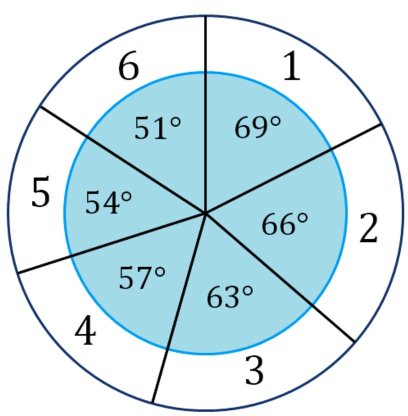
Muista lisätä desimaaleja, eli asetuset → pyöristystarkkuus esim. 4 desimaaleja.



3. a) Oletetaan, että $X \sim N(32, 5)$. Määritä sellainen luku $a > 0$, että

$$P(|X - 32| \leq a) = 0,40\,.\ (5\text{p})$$

b) Mikä on todennäköisyys saada alla olevan kuvion onnenpyörällä pisteluku 2 tai 6?(2p)



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

i) $P(A|B)$, ii) $P(B|\overline{A})$ (5p)

a) Koska $X \sim N(32{,}5)$, niin normitus antaa

$$X = \frac{X - 32}{5} \qquad \Longrightarrow \qquad Z \sim \mathrm{N}(0,1)$$

Näin ollen

$$P(|X - 32| \leq a) = 0{,}40 \qquad \Longrightarrow \qquad P(-a \leq X - 32 \leq a) = 0{,}40$$

$$\Longrightarrow \qquad P(-a + 32 \leq X \leq a + 32) = 0{,}40$$

$$\Longrightarrow \qquad P\left(\frac{-a + 32 - 32}{5} \leq Z \leq \frac{a + 32 - 32}{5}\right) = 0{,}40$$

$$\Longrightarrow \qquad P\left(\frac{-a}{5} \leq Z \leq \frac{a}{5}\right) = 0{,}40$$

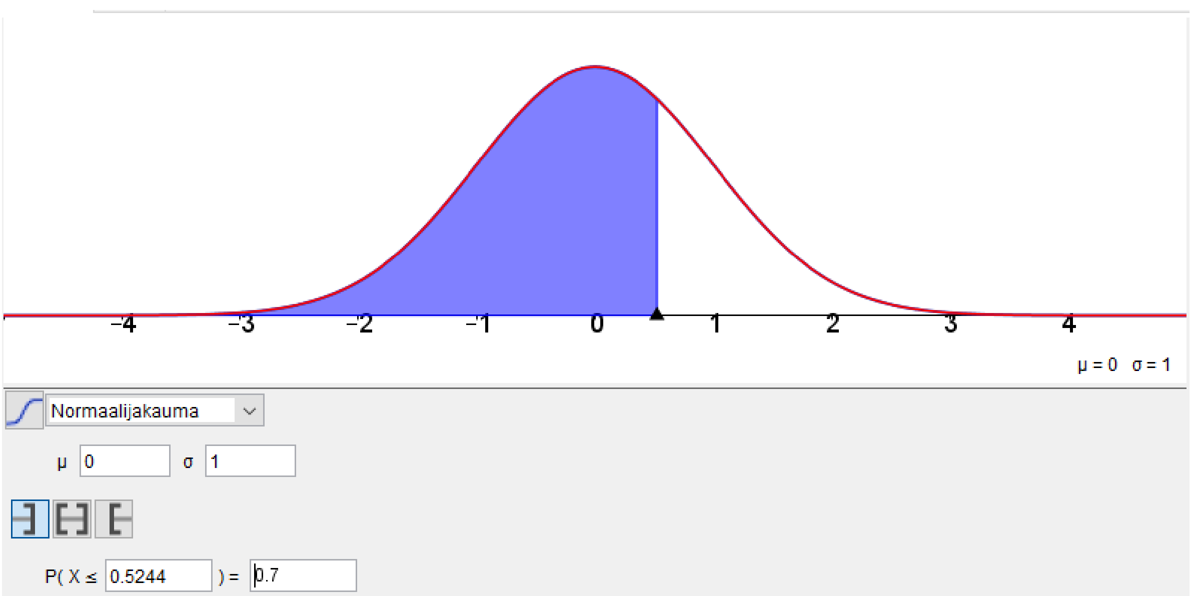
Koska $a > 0$ niin

$$\Longrightarrow \quad \Phi\left(\frac{a}{5}\right) - \left(1 - \Phi\left(\frac{a}{5}\right)\right) = 2 \cdot \Phi\left(\frac{a}{5}\right) - 1 = 0{,}40$$

$$\Longrightarrow \quad \frac{a}{5} = 0{,}5244\ldots \qquad \Longrightarrow \qquad a = 2{,}622\ldots \approx 2{,}6$$

TAI koneella

$$\mathrm{solve}\left(2 \cdot \mathrm{normCdf}\left(-\infty, \frac{a}{5}, 0, 1\right) - 1 = 0.4, a\right) \qquad a = 2.622$$



b) Olkoon tapahtuma A:"Saadaan pisteluku 2 tai 6". Tällöin tapahtuman A todennäköisyys on

$$P(A) = \frac{66° + 51°}{360°} = \frac{117°}{360°} = 0{,}325.$$

c) Nyt siis

$$P(A)=\frac{1}{5},\qquad P(B)=\frac{1}{2}\qquad \text{ja}\qquad P(A\cup B)=\frac{7}{10}$$

joten yhteenlaskusäännöstä

$$P(A\cup B)=P(A)+P(B)-P(A\cap B)$$

saadaan

$$P(A\cap B)=P(A)+P(B)-P(A\cup B)=\frac{1}{5}+\frac{1}{2}-\frac{7}{10}=0.$$

Eli joukot A ja B ovat erilliset. Toisin sanoen niillä ei ole yhteisiä alkioita.

Nyt voidaan ehdollisen todennäköisyyden laskukaavaa $P(A|B)=\frac{P(A\cap B)}{P(B)}$ käyttää ja saadaan **i)**-kohtaan

$$P(A|B)=\frac{0}{1/2}=0\,.$$

Mikä tietysti on totta, eli A ei riipu B:stä millään tavoin.

Vastaavasti koska $P(\overline{A})=1-P(A)=4/5$, ja tapahtumat A ja B ovat erilliset, eli $A\cap B=\emptyset$ niin

$$\overline{A}\cap B=B\quad\Rightarrow\quad P(\overline{A}\cap B)=P(B)\,.$$

Nyt ehdollisen todennäköisyyden laskukaavaa käyttäen saadaan **ii)**-kohtaan

$$P(B|\overline{A})=\frac{P(B\cap\overline{A})}{P(\overline{A})}=\frac{P(B)}{P(\overline{A})}=\frac{1/2}{4/5}=5/8\,.$$

4. a) Kuinka monta korttia pakasta (52 korttia) pitää nostaa, jotta todennäköisyys saada ainakin yksi ässä on suurempi kuin $0{,}9$? Nostettuja kortteja ei palauteta takaisin pakkaan. Hyödynnä laskinohjelmistoja, mutta idea ja jokin matemaattinen kaava, yhtälö, epäyhtälö tms. tulee olla vastauksessa mukana.

b) Valitaan satunnaisesti kaksi henkilöä. Millä todennäköisyydellä syntymäaikojen $[0,24[$ välillä on ainakin 10 tuntia? Perustele huolella.

c) Millä todennäköisyydellä kolmilapsisen perheen lapsista kaksi on poikia, kun tiedetään, että perheen lapsista ainakin yksi on poika? Syntyvistä lapsista $51{,}2$ % on poikia.

a) Koska sana *ainakin* niin hyödynnetään komplementtisääntöä.

Aluksi $P(\text{"Saadaan ässä."}) = 4/52$ josta $P(\text{"Ei saada ässää."}) = 48/52$, joten, kun nostettuja kortteja ei palauteta pakkaan (jonoajatus), niin

$$P(\text{"Ainakin yksi ässä } x\text{:llä nostolla."}) = 1 - P(\text{"Ei yhtään ässää } x\text{:llä nostolla."})$$

$$= 1 - \prod_{i=1}^{x} \frac{48-(i-1)}{52-(i-1)} > 0{,}9$$

$$\Rightarrow \quad \prod_{i=1}^{x} \frac{49-i}{53-i} < 0{,}1$$

$$\Rightarrow \quad \frac{48}{52}\cdot\frac{47}{51}\cdot\dots\cdot\frac{49-x}{53-x} = \frac{\overbrace{48\cdot 47\cdot\dots\cdot(49-x)}^{\mathrm{nPr}(48,x)}}{\underbrace{52\cdot 51\cdot\dots\cdot(53-x)}_{\mathrm{nPr}(53,x)}} < 0{,}1$$

Laskimella kokeilua

$$\frac{\mathrm{nPr}(48,\{10,20,30\})}{\mathrm{nPr}(52,\{10,20,30\})} \qquad \{0.413,0.133,0.027\}$$

$$\frac{\mathrm{nPr}(48,\{21,22,23,24,25\})}{\mathrm{nPr}(52,\{21,22,23,24,25\})} \qquad \{0.116,0.101,0.088,0.076,0.065\}$$

joten voidaan todeta, että tarvitaan vähintään $x = 23$ nostoa.

b) Olkoon $x =$ ”Henkilön A syntymäaika $[0{,}24[$.” ja $y =$ ” Henkilön B syntymäaika $[0{,}24[$.”. Tällöin tarkasteltavan tapahtuman alkeistapaus on järjestetty pari, merkitään (x, y), missä $0 \le x < 24$ ja samoin $0 \le y < 24$, arvot tunteja. Perusjoukoksi E saadaan (huomaa puoliavoimet välit.)

$$E = \{(x,y) \in \mathbb{R}^2 \mid 0 \le x < 24, 0 \le y < 24\}.$$

Tapahtumalle $A =$”*syntymäaikojen* $[0{,}24[$ *välillä on ainakin 10 tuntia.*” suotuisat alkeistapaukset toteuttavat ehdon, kun siis henkilön A syntymäaika on x (vaaka-akseli) ja henkilön B syntymäaika y (pystyakseli)

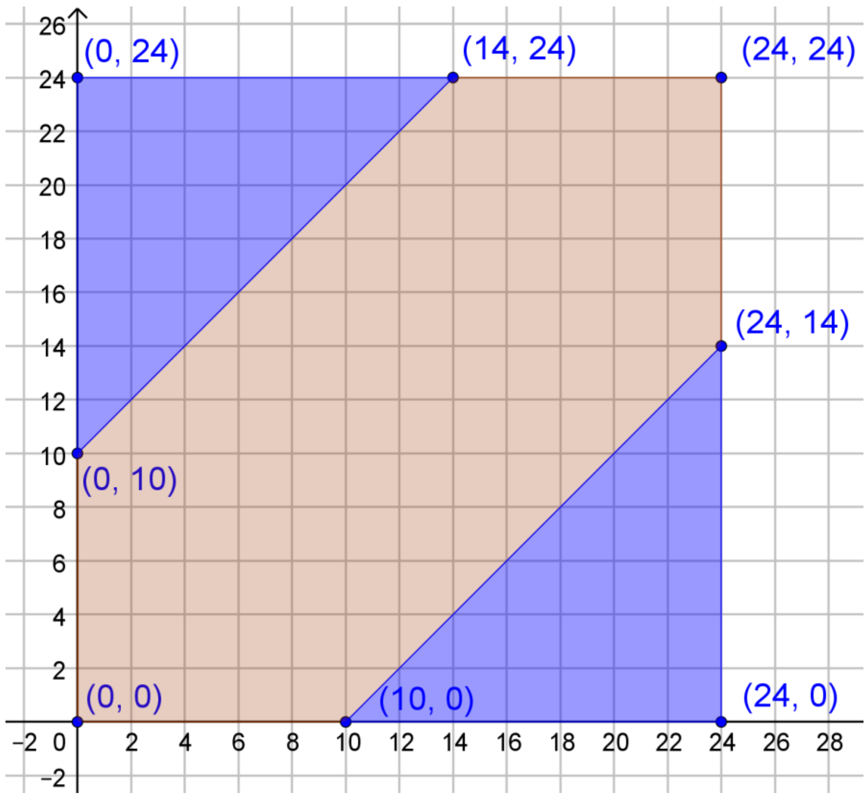
$$|y - x| \ge 10 \quad \text{(tai yhtä hyvin } |x - y| \ge 10)$$

Toisin sanoen (aukaistaan itseisarvot)

$$\begin{cases} y - x \leq -10 \\ y - x \geq 10 \end{cases} \Leftrightarrow \begin{cases} y \leq x - 10 \\ y \geq x + 10 \end{cases}.$$

Kaksoisepäyhtälöstä saadaan siis kaksi suoraa $y = x + 10$ ja $y = x - 10$, joiden rajaamat kaksi kolmiota neliöstä vastaavat kysyttyä tapahtumaa A. Geometrista todennäköisyyttä käyttäen saadaan

$$P(A) = \frac{A_{\text{kolmiot}}}{A_{\text{neliö}}} = \frac{2 \cdot \left(\frac{1}{2} \cdot 14 \cdot 14\right)}{24^2} = \frac{196}{576} \approx 0{,}340\,277\ldots.$$



c) Merkitään $A =$"Poika.", jolloin $P(A) = 0{,}512$. Lisäksi

$$P(\text{"Kolmilapsisen perheen lapsista ainakin yksi lapsista on poika."})$$

$$= 1 - P(\text{"Kolmilapsisen perheen lapsista kukaan ei ole poika."})$$

$$= 1 - 0{,}488^3 = 0{,}883\,785\ldots$$

Näin ollen ehdollista todennäköisyyttä käyttäen saadaan

$$P(\text{"Lapsista 2 on poikia.} \mid \text{Ainakin yksi lapsista on poika."})$$

$$= \frac{P(\text{"Lapsista 2 on poikia ja ainakin yksi poika."})}{P(\text{"Ainakin yksi lapsista on poika."})} = \frac{P(\text{"PPT tai PTP tai TPP"})}{0{,}883\,785\ldots}$$

$$\Longrightarrow \quad = \frac{3 \cdot 0{,}512^2 \cdot 0{,}488}{0{,}883\,785\ldots} = \frac{0{,}383\,778\ldots}{0{,}883\,785\ldots} = 0{,}434\,244\ldots \approx 0{,}43$$

5. a) Mitkä kolme ehtoa on oltava voimassa jatkuvasti jakautuneen satunnaismuuttujan tiheysfunktiolle? (2p)

b) Määritä sellainen vakio $\boldsymbol{a}$, että funktio

$$\boldsymbol{f}: \boldsymbol{f}(\boldsymbol{x}) = \begin{cases} \boldsymbol{0}, & \textbf{kun } \boldsymbol{x} \leq \boldsymbol{0} \\ \boldsymbol{ax}(\boldsymbol{x}^{\boldsymbol{2}} - \boldsymbol{5}), & \textbf{kun } \boldsymbol{0} < \boldsymbol{x} \leq \boldsymbol{2} \\ -\boldsymbol{ax}, & \textbf{kun } \boldsymbol{2} < \boldsymbol{x} \leq \boldsymbol{4} \\ \boldsymbol{0}, & \textbf{kun } \boldsymbol{4} < \boldsymbol{x} \end{cases}$$

on erään jatkuvan satunnaismuuttujan $\boldsymbol{X}$ tiheysfunktio. Hyödynnä aineistot-osion *tehtava5.ggb*-tiedostoa ja etsi ensin likiarvo kertoimelle $\boldsymbol{a}$ kahden desimaalin tarkkuudella (liitä näyttökuva vastaukseesi) ja sitten laskinohjelmistoja hyödyntäen tarkka arvo viiden desimaalin tarkkuudella. Vastauksesta tulee löytyä laskukaava, jolla tarkka arvo on saatu. (Vihje: Mietipä a)-kohdan ehtoja samalla.) (4p)

c) Määritä satunnaismuuttujan $\boldsymbol{X}$ kertymäfunktio. Eli muodosta kertymäfunktion lauseke. (4p)

d) Määritä kahden desimaalin tarkkuudella

$$\boldsymbol{P}(\boldsymbol{2} \leq \boldsymbol{X} < \boldsymbol{3})$$

ja keskihajonta

$$\mathbb{D}(\boldsymbol{X})\,.$$

Hyödynnä laskinohjelmistoja, idea/laskukaava tulee olla näkyvissä.(2p)

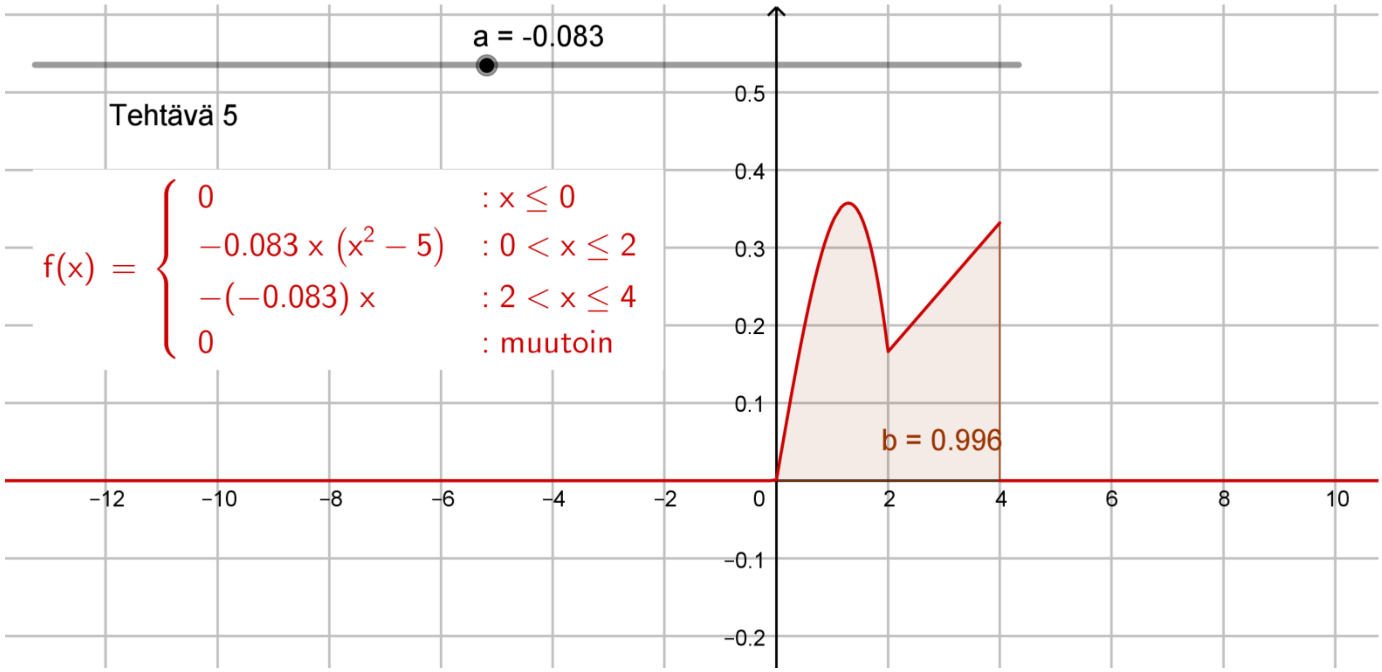
a) Ehdot ovat:

1. Tiheysfunktion f tulee olla ei-negatiivista kaikilla $x \in \mathbb{R}$.
2. Tiheysfunktion f kuvaajan ja x-akselin väliin jäävän alueen pinta-alan on oltava tasan 1, joten saadaan

$$\int_{-\infty}^{\infty} f(x)\,dx \overset{\text{ehto}}{\triangleq} 1\,.$$

3. Tiheysfunktiolla f on oltava äärellinen määrä epäjatkuvuuskohtia.

b) Likiarvo *tehtava5.ggb*-tiedostoa käyttäen käytännössä tarkoittaa ehdon 2 määrittämistä. Samalla toki saadaan ehdot 1 ja 3 voimaan. Kertoimen a likiarvoksi kahden desimaalin tarkkuudella saadaan $-0{,}08$.



Ei-negatiivisuusehdosta saadaan, että $a < 0$. Epäjva-kohtia on korkeintaan kolme (itse asiassa vain yksi kun $a \neq 0$ ja ei yhtään kun $a = 0$), joten 3.ehto myös OK. Vielä pitäisi määrittää integraali, saadaan

$$\int_{-\infty}^{\infty} f(x)\,dx = \underbrace{\int_{-\infty}^{0} 0\,dx}_{=0} + \int_{0}^{2} ax(x^2-5)\,dx + \int_{4}^{\infty} -ax\,dx + \underbrace{\int_{4}^{\infty} 0\,dx}_{=0} = 1$$

$$\Longrightarrow \qquad \underbrace{\int_{0}^{2} ax(x^2-5)\,dx + \int_{2}^{4} -ax\,dx}_{=-12a} = 1$$

Voi tehdä käsin mutta laskinohjelmistot siis riittää, $a = -\frac{1}{12} \approx -0{,}08333$ viiden desimaalin tarkkuudella.

$$\int_0^2 \left(a\cdot x\cdot\left(x^2-5\right)\right)dx+\int_2^4\left(-a\cdot x\right)dx \qquad -12\cdot a$$

$$\text{solve}\left(\int_0^2 \left(a\cdot x\cdot\left(x^2-5\right)\right)dx+\int_2^4\left(-a\cdot x\right)dx=1,a\right) \qquad a=\frac{-1}{12}$$

$$\text{solve}\left(\int_0^2 \left(a\cdot x\cdot\left(x^2-5\right)\right)dx+\int_2^4\left(-a\cdot x\right)dx=1,a\right) \qquad a=-0.08333333333$$

Näin ollen tiheysfunktio f on muotoa

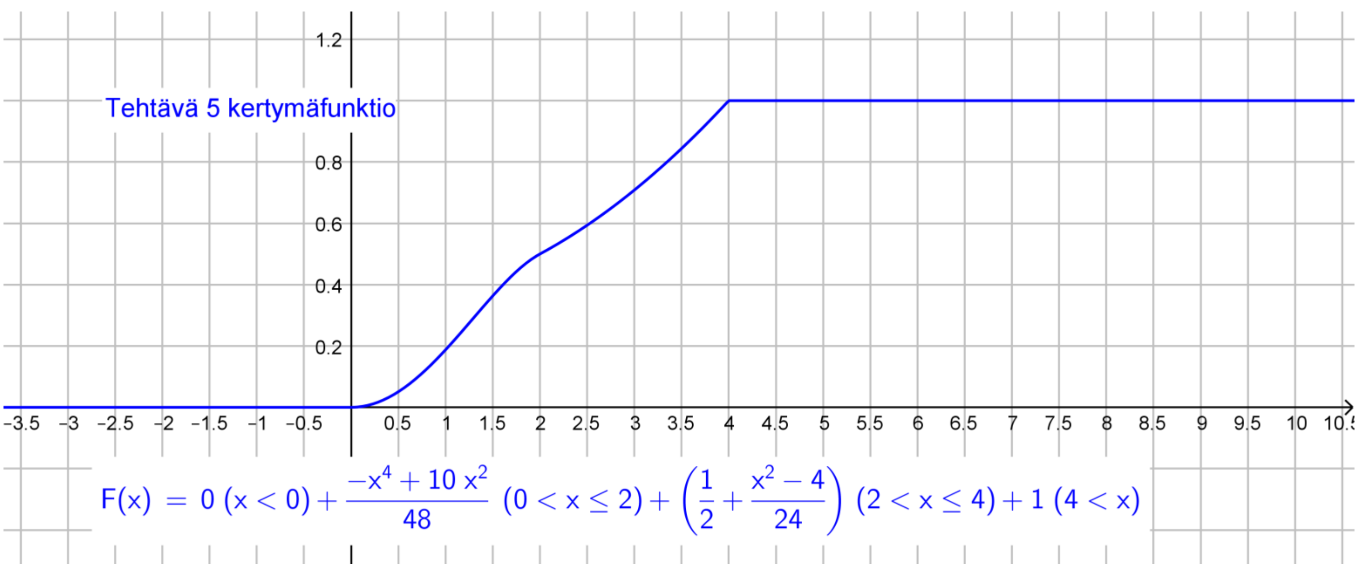
$$f: f(x) = \begin{cases} 0, & \text{kun } x \leq 0 \\ -\dfrac{1}{12}x(x^2-5), & \text{kun } 0 < x \leq 2 \\ \dfrac{1}{12}x, & \text{kun } 2 < x \leq 4 \\ 0, & \text{kun } 4 < x \end{cases} = \begin{cases} 0, & \text{kun } x \leq 0 \\ -\dfrac{1}{12}x^3 + \dfrac{5x}{12}, & \text{kun } 0 < x \leq 2 \\ \dfrac{1}{12}x, & \text{kun } 2 < x \leq 4 \\ 0, & \text{kun } 4 < x \end{cases}$$

c) Muodostetaan kertymäfunktio F.

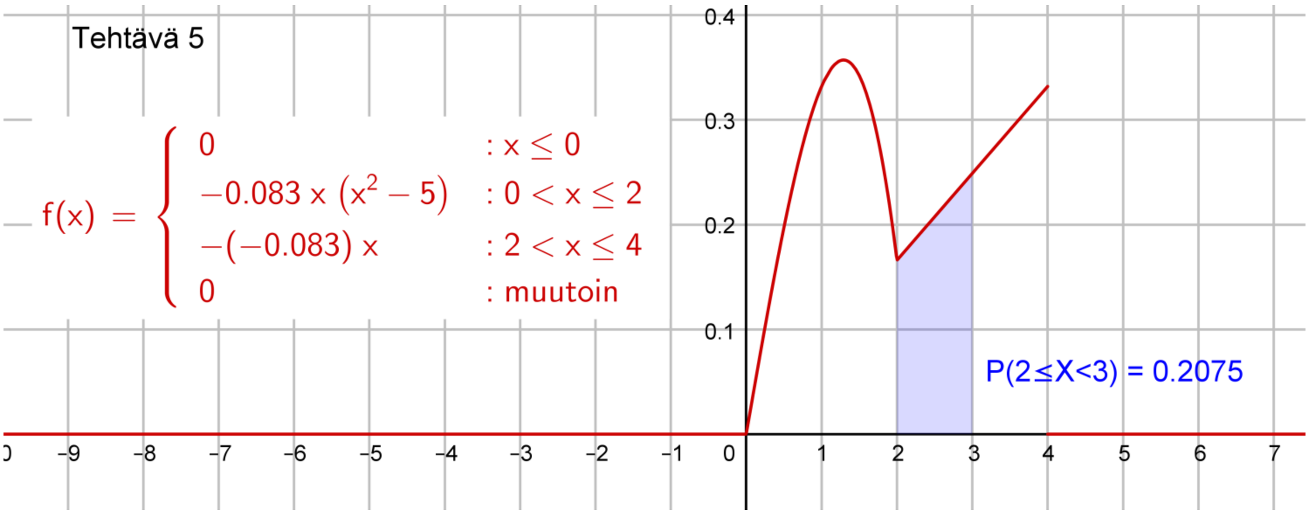
$$F:\ F(x) = \begin{cases} \displaystyle\int_{-\infty}^{x} 0\, dt = 0, & \text{kun } x \leq 0 \\ \displaystyle F(0) + \int_0^x \left(-\frac{1}{12}t^3 + \frac{5t}{12}\right) dt = 0 + \frac{-x^4+10x^2}{48}, & \text{kun } 0 < x \leq 2 \\ \displaystyle F(2) + \int_2^x \frac{1}{12}t\, dt = \frac{1}{2} + \frac{x^2-4}{24}, & \text{kun } 2 < x \leq 4 \\ \displaystyle F(2) + \int_4^{\infty} 0\, dt = 1 + 0 = 1, & \text{kun } 4 < x \end{cases}$$

Eli, (kuvaajaa ei vaadittu!)

$$F:\ F(x) = \begin{cases} 0, & \text{kun } x \leq 0 \\ \dfrac{-x^4+10x^2}{48}, & \text{kun } 0 < x \leq 2 \\ \dfrac{1}{2} + \dfrac{x^2-4}{24}, & \text{kun } 2 < x \leq 4 \\ 1, & \text{kun } 4 < x \end{cases}$$



d) Todennäköisyyttä $P(2 \leq X < 3)$ varten hyödynnetään esim. Geogebraa. Huomaa, että välin päätekohdan kuuluminen tai kuulumattominen eivät jatkuvan satunnaismuuttujan tapauksessa vaikuta todennäköisyyteen.



ja laskien

$$P(2 \leq X < 3) = F(3) - F(2) = \underbrace{\left(\frac{1}{2} + \frac{3^2 - 4}{24}\right)}_{=\frac{17}{24}} - \underbrace{\left(\frac{-2^4 + 10 \cdot 2^2}{48}\right)}_{=\frac{1}{2}} = \frac{5}{24} = 0{,}208\overline{3} \approx 0{,}21.$$

TAI TI-ohjelmistolla

$$f(x) := \begin{cases} 0, & x \leq 0 \\ \frac{-1}{12} \cdot x^3 + \frac{5 \cdot x}{12}, & 0 < x \leq 2 \\ \frac{1}{12} \cdot x, & 2 < x \leq 4 \\ 0, & 4 < x \end{cases} \qquad \textit{Valmis}$$

$$\int_2^3 f(x)\, dx \qquad \frac{5}{24}$$

$$\int_2^3 f(x)\, dx \qquad 0.2083333$$

Ero johtuu siitä, että Geogebrassa on kertoimen a arvona likiarvo $a \approx -0{,}083$.

Lopuksi keskihajonta (ensin odotusarvo)

$$\mu = \mathbb{E}(X) = \int_{-\infty}^{\infty} x \cdot f(x)\, dx = \ \dots\ = \int_0^2 x \cdot \left(-\frac{1}{12}x^3 + \frac{5x}{12}\right) dx + \int_2^4 x \cdot \frac{x}{12}\, dx = \frac{32}{15}$$

$$\sigma = \mathbb{D}(X) = \sqrt{\int_{-\infty}^{\infty} (x - \mu)^2 \cdot f(x)\, dx} = \ \dots\ = \sqrt{\int_0^2 \left(x - \frac{32}{15}\right)^2 \cdot \left(-\frac{1}{12}x^3 + \frac{5x}{12}\right) dx + \int_2^4 \left(x - \frac{32}{15}\right)^2 \cdot \frac{x}{12}\, dx} = \frac{2\sqrt{69}}{15}$$

6. Arpajaisissa on jäljellä 16 arpaa, joista viidessä on voitto. Määritä satunnaismuuttujan X ="voittoarpojen lukumäärä" jakauma sekä laske odotusarvo ja keskihajonta yhden desimaalin tarkkuudella, kun ostetaan 4 arpaa. Liitä jakaumasta kuva vastaukseesi. Käytä esim. Geogebrassa komentoa *Pylväskaavio(<datalista>, <frekvenssilista>, <palkin leveys>)*, eli esim. *Pylväskaavio({4, 5, 6, 7}, {0.3, 0.4, 0.2, 0.1}, 0.01)* ja voit sitten *ominaisuuksista -> objektin tyyli -> viivan paksuus* laittaa viivan paksuutta tarvittaessa lisää.

Tulkitse selkosuomeksi merkinnät $X \geq 2$ ja $P(X < 2)$.

Koska kyseessä on diskreetti satunnaismuuttuja (ostettujen voittavien arpojen lukumäärä on jokin kokonaisluku väliltä $0 - 4$), niin määritetään pistetodennäköisyydet:

$$P(X = x_i) = \frac{\binom{5}{x_i}\binom{11}{4-x_i}}{\binom{16}{4}} = \frac{\binom{5}{x_i}\binom{11}{4-x_i}}{1820},$$

missä $x_i \in \{0,1,2,3,4\}$. Eli

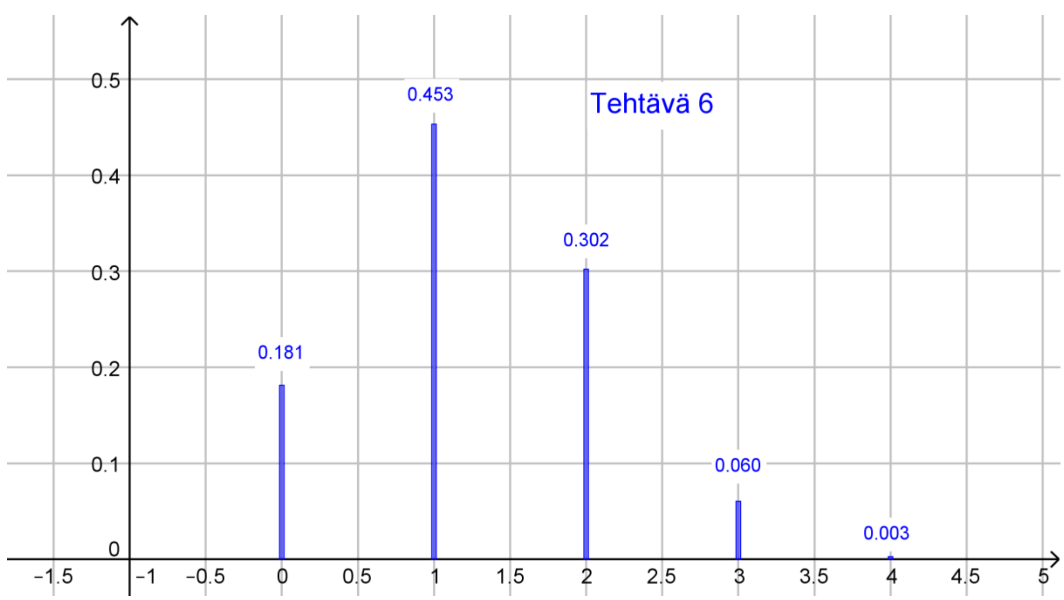
$$P(X = 0) = \frac{\binom{5}{0}\binom{11}{4-0}}{\binom{16}{4}} = \frac{1 \cdot 330}{1820} = \frac{33}{182} \approx 0{,}181$$

$$P(X = 1) = \frac{\binom{5}{1}\binom{11}{4-1}}{\binom{16}{4}} = \frac{5 \cdot 165}{1820} = \frac{165}{364} \approx 0{,}453$$

$$P(X = 2) = \frac{\binom{5}{2}\binom{11}{4-2}}{\binom{16}{4}} = \frac{10 \cdot 55}{1820} = \frac{55}{182} \approx 0{,}302$$

$$P(X = 3) = \frac{\binom{5}{3}\binom{11}{4-3}}{\binom{16}{4}} = \frac{10 \cdot 11}{1820} = \frac{11}{182} \approx 0{,}060$$

$$P(X = 4) = \frac{\binom{5}{4}\binom{11}{4-4}}{\binom{16}{4}} = \frac{5 \cdot 1}{1820} = \frac{1}{364} \approx 0{,}003$$



Odotusarvoksi saadaan

$$\mu = \mathbb{E}(X) = 0 \cdot \frac{33}{182} + 1 \cdot \frac{165}{364} + 2 \cdot \frac{55}{182} + 3 \cdot \frac{11}{182} + 4 \cdot \frac{1}{364} = \frac{5}{4} = 1{,}25\,.$$

ja keskihajonnaksi

$$\sigma = \mathbb{D}(X) = \sqrt{\left(0 - \frac{5}{4}\right)^2 + \left(1 - \frac{5}{4}\right)^2 + \left(2 - \frac{5}{4}\right)^2 + \left(3 - \frac{5}{4}\right)^2 + \left(4 - \frac{5}{4}\right)^2}$$

$$= \frac{\sqrt{205}}{4} = 3{,}579\,455\,2\,\ldots \approx 3{,}6$$

Merkintä $X \geq 2$ tarkoittaa selkosuomeksi “Voittoarpojen lukumäärä on vähintään 2.” ja merkintä $P(X < 2)$ tarkoittaa “Todennäköisyys, että voittoarpoja on alle yksi.” tai “Todennäköisyys, että voittoarpoja on alle yksi.”

7. **a) Koulussa on 300 lamppua, joiden kestoikä noudattaa normaalijakaumaa keskiarvon ollessa 1200 h ja keskihajonnan 200 h. Lamput palavat 55 h viikossa. Kaikki lamput vaihdetaan kerralla. Kuinka usein tämä on tehtävä, jotta vain 50 lamppua ehtii sammua. [lyhyt K1993/10] Liitä vastaukseesi kuva normaalijakaumasta, joka vastaa tehtävänannon tilannetta**. **Vastauksesta tulee löytyä idea ja laskut, joilla kysytty todennäköisyys saadaan, mutta esim. geogebran todennäköisyyslaskuria tai TI:tä saa käyttää tukena. (9p)**

   b) Juhlissa on 115 vierasta. Kuinka monta kättelyä on suoritettava, jotta jokainen vieras on kätellyt jokaista toista vierasta täsmälleen kerran? (3p)

a) Merkitään satunnaismuuttujalla X lampun kestoikää, eli $X =$”Lampun paloaika/kestoikä.” Tiedetään, että

$$X \sim N(1200{,}200) \qquad \Longrightarrow \qquad Z = \frac{X - 1200}{200} \sim N(0{,}1)$$

Vain 50 lamppua sammuu, eli $\frac{50}{300} = \frac{1}{6}$, joten jos merkitään $a =$ ”lamppujen paloaika(kestoikä) kunnes 50 lamppua on sammunut.”, niin

$$P(X < a) = \frac{1}{6} \qquad \Longrightarrow \qquad P\left(Z < \frac{a - 1200}{200}\right) = \frac{1}{6} < \frac{1}{2}.$$

Siis $a < 1200$!

Taulukkotietoja käyttäen

$$\Rightarrow \quad P\left(Z < \frac{1200 - a}{200}\right) = \frac{1}{6} \qquad \text{TAI} \quad P\left(Z < -\frac{a - 1200}{200}\right) = \frac{5}{6}$$

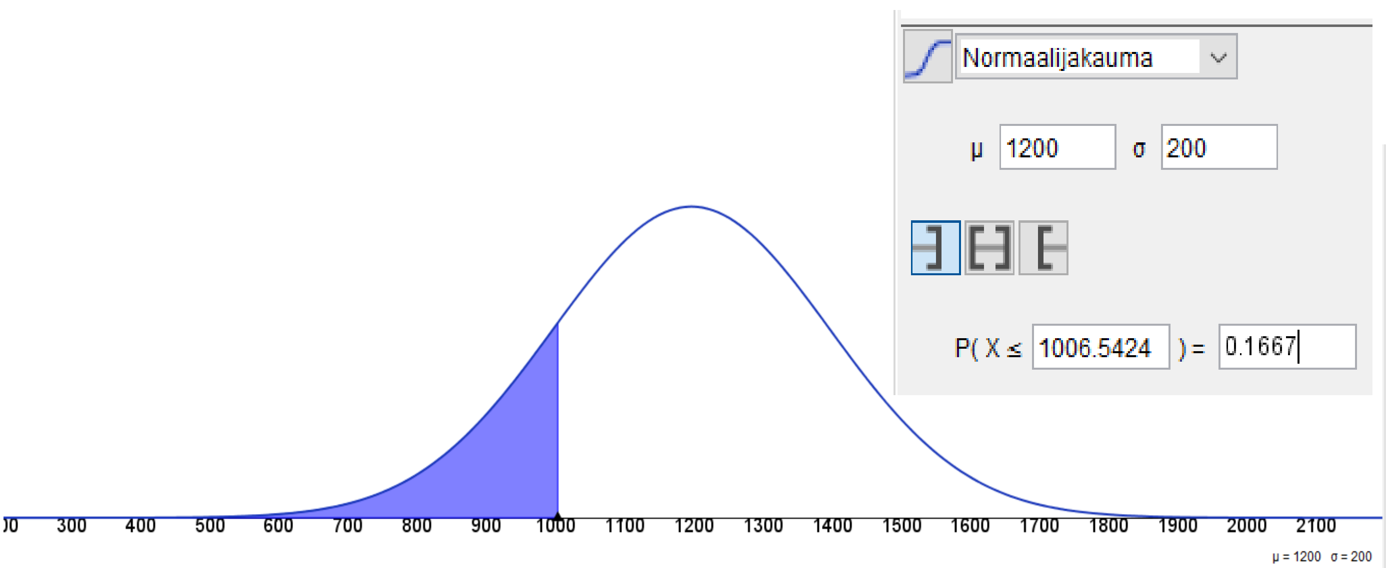
$$\Rightarrow \quad \frac{a - 1200}{-200} = 0{,}967\,421\ldots \quad \Rightarrow \quad a = 1006{,}515\ldots \approx 1007\ \text{h}\,.$$

Viikkoina 1007 tuntia on noin 18,3.

TAI TI:llä

$$\text{solve}\left(\text{normCdf}\left(-\infty, a, 1200, 200\right) = \frac{1}{6}, a\right) \qquad a = 1006.5156926$$

TAI Geogebran todennäköisyyslaskuriin laittaa puoliavoimen välin ja todennäköisyyden arvoksi 1/6, jolloin ylärajaksi tulee tuo 1006,5 ...katso kuva alla. Huom, nyt arvo heittää hieman koska todennäköisyys 1/6 pyöristyy arvoksi 0,1667!



Siis lamput tulee vaihtaa noin 18 viikon välein.

b) Käytetään ensin osajoukkoideaa. Yksi kättely vastaa kahden vieraan osajoukkoa, eli kättelyitä = kahden vieraan osajoukkoa kaikista vieraista on

$$\binom{115}{2} = 6\,555$$

kappaletta. Toisaalta voidaan hyödyntää jonoajattelua:

$$114 + 113 + 112 + \cdots + 2 + 1 \overset{\substack{\text{aritm.}\\ \text{summa}}}{\widehat{=}} \frac{1 + 114}{2} \cdot 114 = 6\,555$$

115. vieras kättelee muut ja poistuu.

114. vieras kättelee muut ja poistuu.

3. vieras kättelee muut ja poistuu.

2. vieras kättelee muut ja poistuu.

$\mathrm{nCr}(115,2)$	6555
$\sum_{i=1}^{114}(i)$	6555