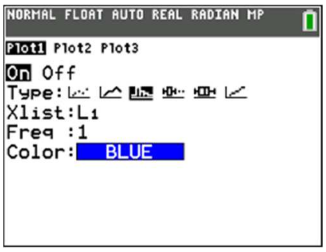
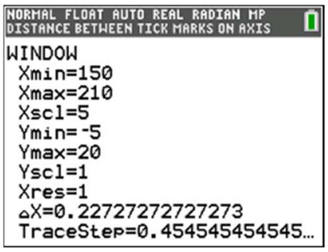
1. In de hoofdstad van Ivoorkust, Yamoussoukro, meet men de lengte van 100 mannen (in cm) :

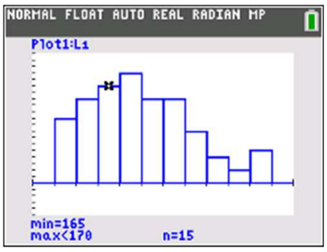
168,6	156,4	166,8	185,5	177,3	201,8	177,3	197,3	175,5	169,5
172,7	170,9	190,0	179,1	166,8	202,3	162,7	170,0	155,0	168,6
169,5	157,7	168,6	189,5	183,2	159,1	160,0	168,2	172,3	161,8
190,5	186,4	178,6	161,8	197,7	173,2	174,5	185,9	165,9	181,8
157,7	178,2	171,4	175,0	163,6	183,2	180,9	160,9	155,0	***204,9***
175,5	164,5	186,8	172,3	169,5	164,1	167,3	181,8	176,8	180,0
174,5	185,9	165,9	181,8	157,7	178,2	171,4	175,0	163,6	183,2
170,9	177,3	190,0	155,0	183,2	190,5	174,5	171,4	***155,0***	169,5
185,5	169,5	202,3	157,7	168,2	161,8	181,8	183,2	164,5	181,8
201,8	170,9	170,0	189,5	161,8	173,2	178,2	160,9	172,3	180,0

a) Stel een frequentietabel op met als eerste klasse $[150,155[$.

(Geef alle gegevens in als L_1 op je rekenmachine en plot het (gewone) histogram. Gebruik daarna de trace-functie om alle frequenties af te lezen.



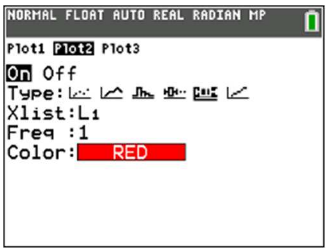


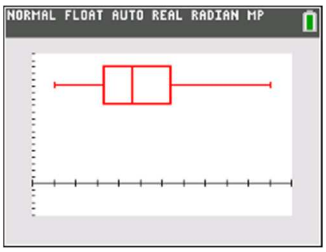


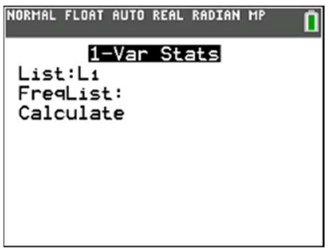
Klasse	x_i	n_i	cn_i	f_i	Genormaliseerde frequentie
$[150,155[$	152,5	0	0	0	0
$[155,160[$	157,5	10	10	0.1	0,02
$[160,165[$	162,5	13	23	0.13	0,026
$[165,170[$	167,5	15	38	0.15	0,03
$[170,175[$	172,5	17	55	0.17	0,034
$[175,180[$	177,5	13	68	0.13	0,026
$[180,185[$	182,5	13	81	0.13	0,026
$[185,190[$	187,5	8	89	0.08	0,016
$[190,195[$	192,5	4	93	0.04	0,008
$[195,200[$	197,5	2	95	0.02	0,004
$[200,205[$	202,5	5	100	0.05	0,01
$[205,210[$	207,5	0	0	0	0

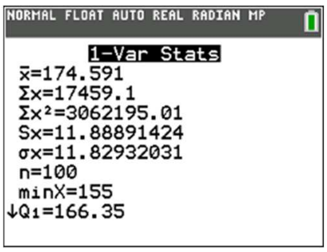
(voor de genormaliseerde frequentie deel je de relatieve frequentie door de klassenbreedte)

b) Teken een boxplot en duid alle relevante waarden aan.





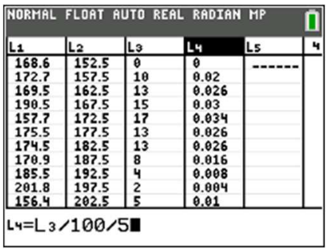


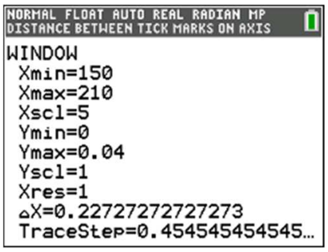


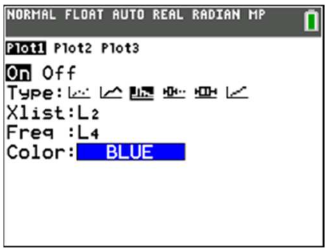
Je leest af dat $x_{\min} = 155$, $Q_1 = 166,35$, $Me = 172,95$, $Q_3 = 181,8$ en $x_{\max} = 204,9$.

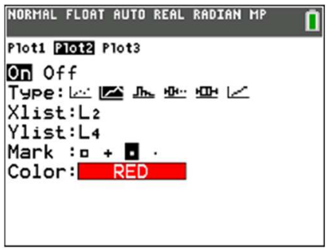
c) Teken het bijhorende histogram en bijhorend enkelvoudig frequentiepolygoon. Zorg ervoor dat je assen zo geijkt zijn dat het frequentiepolygoon kan opgevat worden als een dichtheidskromme.

Stop in L_2 *de klassemiddens, in* L_3 *de frequenties die je in vraag a vond en definieer* L_4 *als de genormaliseerde frequenties door te stellen dat* $L_4 = L_3/100/5$ *- de relatieve frequentie gedeeld door de klassebreedte).*







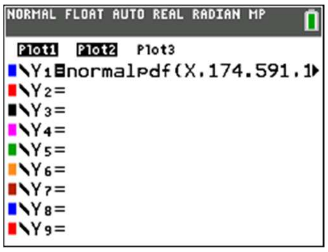


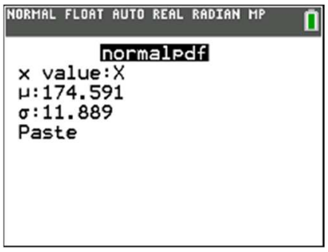
(voor het histogram en het enkelvoudig frequentiepolygoon: zie het antwoord op vraag e).

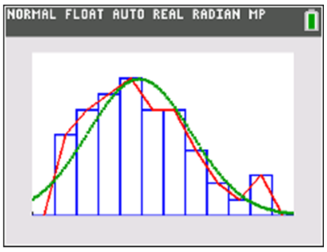
d) Bereken met behulp van je rekenmachine het gemiddelde en de standaardafwijking van deze steekproef.

$\overline{x} = 174,591$ en $s = 11,889$ (je neemt s ipv σ omdat het om een steekproef gaat).

e) Teken (op dezelfde grafiek als het histogram) de bijhorende normale verdelingskromme. Kan je besluiten dat de gegevens min of meer normaal verdeeld zijn?



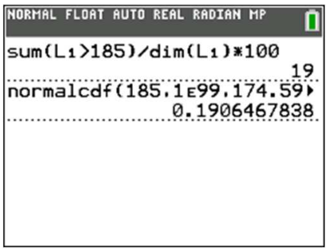




De normale verdelingskromme (de groene dichtheidskromme) is een vrij goede benadering van de genormaliseerde frequentiepolygoon. De gegevens zijn dus inderdaad vrij normaal verdeeld.

f) Hoeveel % van de mannen uit de steekproef meet meer dan 185 cm?

- Bereken dit exact met behulp van je rekenmachine.
- Bereken dit benaderend door ervan uit te gaan dat de verdeling normaal is.



Exact is dit 19% van de steekproef en als we de normale verdeling gebruiken komen we op 19,06% uit. Dit is nogmaals een bevestiging dat de normale verdeling een hele goede benadering is.

2. De tijdsduur van lokale telefoongesprekken is normaal verdeeld met een gemiddelde van 9,5 minuten en een standaardafwijking van 3 minuten.

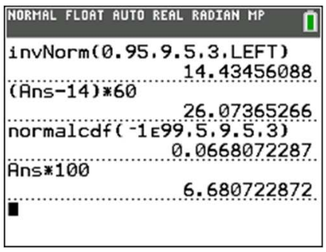
a) Welke tijdsduur wordt door 5% van de gesprekken overschreden?

$\boxed{X \sim N(9{,}5\,;3)}$. $P(X > a) = 0{,}05 \Leftrightarrow P(X < a) = 0{,}95 \Leftrightarrow a \approx 14{,}43456$

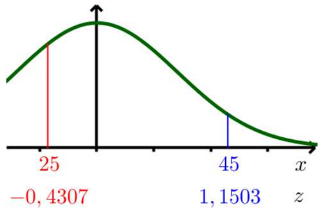
Dus 5% van de gesprekken duurt langer als 14'26".

b) Hoeveel procent van de gesprekken duurt minder dan 5 minuten?

$P(X < 5) \approx 6{,}68\%$



3. Als Mahamadou thuis om 8u00 vertrekt naar school dan is hij twee derde van de keren te laat voor de les die stipt om 8u25 begint. Vertrekt hij om 7u40 dan is hij slechts een achtste van de keren te laat. In de veronderstelling dat Mahamadou zijn reistijd naar school normaal verdeeld is, hoe laat moet hij dan thuis vertrekken om in niet meer dan 5% van de gevallen te laat te komen?

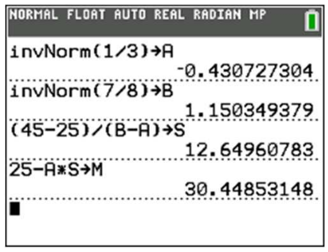


$2/3$ van de tijd doet hij er langer over dan 25 minuten

$P(Z > a) = 2/3 \Leftrightarrow P(Z < a) = 1/3 \Leftrightarrow a \approx -0,4307$, dus $-0,4307 \approx \dfrac{25-\mu}{\sigma}$.

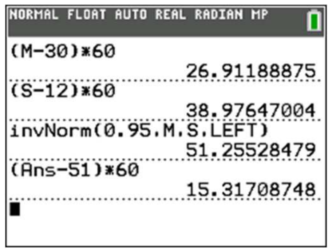
$1/8$ van de tijd doet hij er langer over dan 45 minuten

$P(Z > a) = 1/8 \Leftrightarrow P(Z < a) = 7/8 \Leftrightarrow a \approx 1,1503$, dus $1,1503 \approx \dfrac{45-\mu}{\sigma}$.



$$\begin{cases} -0,4307\sigma \approx 25-\mu \\ 1,1503\sigma \approx 45-\mu \end{cases} \Leftrightarrow \begin{cases} \mu - 0,4307\sigma \approx 25 \\ \mu + 1,1503\sigma \approx 45 \end{cases} \overset{*}{\Leftrightarrow} \begin{cases} \mu \approx 30,4485 \\ \sigma \approx 12,6496 \end{cases}$$

$$* : 25 + 0,4307\sigma \approx 45 - 1,1503\sigma \Leftrightarrow \sigma = \frac{45-25}{1,1503+0,4307}$$



Mahamadou doet dus gemiddeld $30'27"$ over de rit met een standaardafwijking van $12'39"$. De tijdsduur T voor een ritje naar school is dus ongeveer normaal verdeeld met $T \sim N(30,4485\,;\ 12,6496)$.

$P(T \geq t) = 5\% \Leftrightarrow P(T < t) = 95\% \Leftrightarrow t \approx 51,2553$.

Mahamadou zal er dus in 5% van de ritten langer dan $51'15"$ op rijden. Hij moet om zeker te zijn voor 7u34 vertrekken.

4. Een munstuk wordt net zo vaak opgeworpen tot er kop verschijnt, of tot er drie keer na elkaar munt verschijnt. Bepaal de verwachtingswaarde en de variantie van het aantal worpen.

uitkomst	x_i	p_i
K	1	$1/2$
MK	2	$(1/2)^2 = 1/4$
MMK of MMM	3	$2.(1/2)^3 = 1/4$

Verwachtingswaarde:

$$\mu = E(X) = \sum p_i x_i = \frac{1}{2}\cdot 1 + \frac{1}{4}\cdot 2 + \frac{1}{4}\cdot 3 = \frac{7}{4}$$

Variantie:

$$\sigma^2 = Var(X) = \sum p_i x_i^2 - \mu^2 = \frac{1}{2}\cdot 1^2 + \frac{1}{4}\cdot 2^2 + \frac{1}{4}\cdot 3^2 - \frac{49}{16} = \frac{11}{16}$$

(Dus de standaardafwijking is $\sigma = \dfrac{\sqrt{11}}{4} \approx 0,82916$)

5. Bij het kaartspel Whist, in de volksmond 'wiezen' genoemd, krijgt iedereen bij aanvang 13 kaarten uitgedeeld. Noem de stochastische variabele X het aantal azen van één van de spelers.

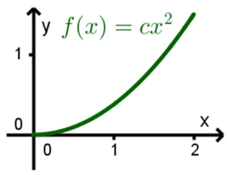
Bereken de verwachtingswaarde en standaardafwijking μ_X en σ_X.

uitkomst	x_i	p_i	verklaring
0 azen	0	$0,30382$	$P(0\ azen)=\dfrac{C_4^0\cdot C_{48}^{13}}{C_{52}^{13}}\approx 0,30382$
1 aas	1	$0,43885$	$P(1\ aas)=\dfrac{C_4^1\cdot C_{48}^{12}}{C_{52}^{13}}\approx 0,43885$
2 azen	2	$0,21349$	$P(2\ azen)=\dfrac{C_4^2\cdot C_{48}^{11}}{C_{52}^{13}}\approx 0,21349$
3 azen	3	$0,04120$	$P(3\ azen)=\dfrac{C_4^0\cdot C_{48}^{13}}{C_{52}^{13}}\approx 0,04120$
4 azen	4	$0,00264$	$P(4\ azen)=\dfrac{C_4^4\cdot C_{48}^{9}}{C_{52}^{13}}\approx 0,00264$

Dus $E(X)\approx 0.0,30382+1.0,43885+2.0,21349+3.0,04120+4.0,00264\approx 1$ (kon je dit voorspellen?)

En $Var(X)\approx 0.0,30382+1.0,43885+4.0,21349+9.0,04120+16.0,00264-1=0,70585\Rightarrow \sigma\approx 0,84$

6. In Spanje mogen de werknemers van firma Tammenont 's middags een siësta houden van maximaal 2 uur.



De tijdsduur van die siësta kan beschreven worden door de stochast X met als dichtheidsfunctie $f(x)=\begin{cases} cx^2 & ,x\in[0,2] \\ 0 & ,x\notin[0,2] \end{cases}$, met $c\in\mathbb{R}$ een constante.

a) Bepaal de waarde van de reële constante c opdat f wel degelijk een dichtheidsfunctie is.

$$\int_0^2 cx^2dx=1\Leftrightarrow\left[\frac{cx^3}{3}\right]_0^2=1\Leftrightarrow c\cdot\frac{8}{3}=1\Leftrightarrow c=\frac{3}{8}.$$

b) Bereken hoeveel procent van de siësta's langer dan 1,5u duren.

$$P(X>1,5)=\int_{1,5}^2\frac{3}{8}x^2dx=\frac{1}{8}\left[x^3\right]_{1,5}^2=\frac{1}{8}\left(2^3-1,5^3\right)=\frac{37}{64}=57,8125\%$$

c) Bereken de verwachtingswaarde μ en de standaardafwijking σ van de continue stochast X.

$$\mu=\int_{-\infty}^{+\infty}x f(x)dx=\int_0^2\frac{3}{8}x^3dx=\frac{3}{2}=1,5 \text{ en } \sigma^2=\int_{-\infty}^{+\infty}x^2 f(x)dx-\mu^2=\int_0^2\frac{3}{8}x^4dx-\frac{9}{4}=\frac{3}{20}\Rightarrow\sigma=\frac{\sqrt{15}}{10}$$

7. Een koperen ring moet passen om een ijzeren staaf. Dus de binnendiameter van de ring moet groter zijn dan de diameter van de staaf. Bij de productie van de ringen hangt de precieze waarde van de binnendiameter X van het toeval af, net als de precieze diameter van de staven Y. Zowel X als Y zijn dus stochastische variabelen die uitgedrukt worden in millimeter.

Stel nu dat $X \sim N(45\,;0,3)$ en $Y \sim N(44\,;0,4)$ (X en Y zijn onafhankelijk van elkaar).

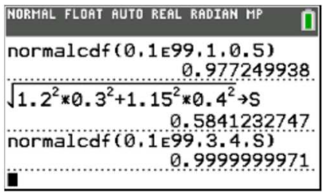
a) Bereken de kans dat een willekeurige ring om een willekeurige staaf past in dit geval.

Ook de stochast $V = X - Y$ is normaal verdeeld, met $\mu = \mu_X - \mu_Y = 45 - 44 = 1$ en $\sigma_V = \sqrt{\sigma_X^2 + \sigma_Y^2} = 0,5$.

Dus $P(X > Y) = P(X - Y > 0) \approx 0,97725 = 97,725\%$

b) Men beslist om de ringen met factor $1,2$ en de staven met factor $1,15$ te vergroten. Bereken in dat geval opnieuw de kans dat een willekeurige ring om een willekeurige staaf past.

Ook de stochast $V' = 1,2X - 1,15Y$ is normaal verdeeld, met $\mu = 1,2\mu_X - 1,15\mu_Y = 1,2.45 - 1,15.44 = 3,4$



en $\sigma_V = \sqrt{1,2^2.\sigma_X^2 + 1,15^2.\sigma_Y^2} \approx 0,5841$.

Dus $P(1,2X > 1,15Y) = P(1,2X - 1,15Y > 0) \approx 1 \approx 100\%$

8. De inhoud van een blikje cola is normaal verdeeld met $\mu = 33\,cl$ en $\sigma = 0,5\,cl$.

a) Je giet drie blikjes cola uit in een glas van exact $1\,l$. Wat is de kans dat het glas overloopt?

De stochast $3C$ die de inhoud van drie blikjes cola weergeeft is ook normaal verdeeld: $3C \sim N\left(99; 0,5\sqrt{3}\right)$

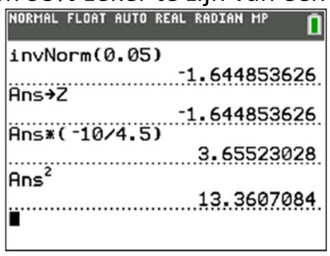
Dus $P(3C > 100) \approx 0,1241 = 12,41\%$

b) Je koopt een verpakking van 6 blikjes cola. Wat is de kans dat de gemiddelde inhoud kleiner is dan $32,6\,cl$?

De stochast $\overline{X}$ die de gemiddelde inhoud weergeeft is ook normaal verdeeld: $\overline{X} \sim N\left(33; 0,5/\sqrt{6}\right)$

Dus $P(\overline{X} < 32,6) \approx 0,0250 = 2,5\%$

9. Leerlingen van het Oscar Romerocollege scoren gemiddeld 84 op de eerste ronde van de wiskunde olympiade, met een standaardafwijking van 10. Hoeveel leerlingen moeten we minstens afvaardigen om 95% zeker te zijn van een gemiddelde van minstens 80?



De stochast $\overline{X}$ die het gemiddelde van n leerlingen geeft is verdeeld als $\overline{X} \sim N\left(84, \dfrac{10}{\sqrt{n}}\right)$.

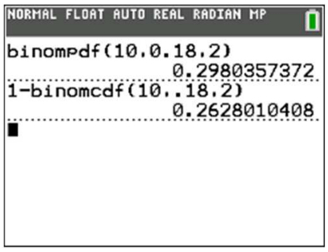
Rekening houdend met de continuïteitscorrectie worden punten vanaf 79,5 afgerond tot 80.

We weten dat $P(\overline{X} > 79,5) = P(Z > z) = 95\% \Leftrightarrow P(Z < z) = 5\% \Leftrightarrow z = -1,6449$.

Dus moet er gelden dat $-1,6449 > \dfrac{79,5 - 84}{10/\sqrt{n}} \Leftrightarrow \sqrt{n} > \dfrac{1,6449.10}{4,5} \Leftrightarrow n > 13,36$.

Vanaf 14 leerlingen zal het gemiddelde dus met 95% waarschijnlijkheid meer zijn dan 80.

10. Bij het spelen van Guitar Hero blijkt dat 18% van de spelers last heeft van hoofdpijn achteraf. Op een vrijgezellenweekend speelt een groepje van 10 vrienden 's avonds Guitar Hero.



a) Wat is de kans dat achteraf precies 2 personen last hebben van hoofdpijn?

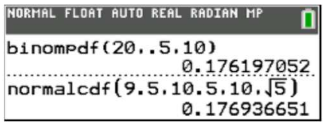
Dit aantal X is binomiaal verdeeld $X \sim B(10;\ 0{,}18)$.

Dus $P(X=2) \approx 0{,}2980$ (ongeveer 30%)

b) Wat is de kans dat minstens 3 vrienden achteraf last heeft van hoofdpijn?

$P(X \geq 3) = 1 - P(X \leq 2) \approx 0{,}2628$ (ongeveer 26%)

11. Iemand gooit een zuiver muntstuk een even aantal keer en wil berekenen hoe groot de kans is dat precies de helft van de keren kop verschijnt.



a) Bereken deze kans als hij 20 keer gooit exact. Rond af op 4 decimalen nauwkeurig.

$P(B(20,1/2)=10) \approx 0{,}1762$

b) Bereken deze kans als hij 20 keer gooit met een normale benadering. Rond af op 4 decimalen nauwkeurig.

$B(20,1/2) \sim N(10,\sqrt{5})$, want $\mu = np = 20 \cdot \frac{1}{2} = 10$ en $\sigma = \sqrt{npq} = \sqrt{20 \cdot \frac{1}{2} \cdot \frac{1}{2}} = \sqrt{5}$.

De continuïteitscorrectie leert ons: $P(B(20,1/2)=10) \approx P(9{,}5 \leq N(10,\sqrt{5}) < 10{,}5) \approx 0{,}1769$

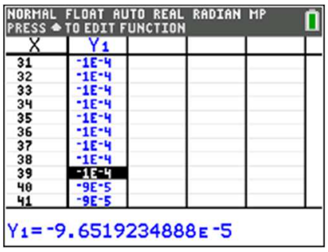
Het verschil tussen de antwoorden op de vorige vragen is niet zo groot.

Het blijkt dat, als het aantal keer gooien (n) groter wordt, dan het verschil tussen de antwoorden kleiner wordt.

c) Leg uit waarom dat zo is

Omdat de normale benadering beter wordt als n groter wordt wegens de centrale limietstelling.

d) Bepaal met je GR vanaf welke n de normale benadering minder dan 0,0001 verschilt met de exacte berekening.

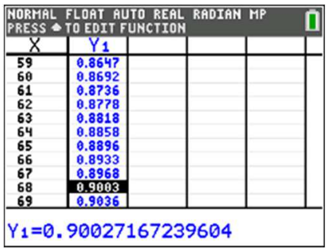


Stel $n = 2x$, dan is $\mu = np = 2x \cdot \frac{1}{2} = x$ en $\sigma = \sqrt{npq} = \sqrt{2x \cdot \frac{1}{2} \cdot \frac{1}{2}} = \sqrt{x/2}$.

Geef bij Y_1 in je rekenmachine de functie in die het verschil aangeeft tussen de exacte waarde $\text{binompdf}(2X,1/2,X)$ en de benadering met de normale verdeling $\text{normalcdf}\left(X-1/2, X+1/2, X, \sqrt{X/2}\right)$.

We kunnen dan eenvoudig aflezen dat het verschil kleiner wordt dan $0{,}0001 = 10^{-4}$ vanaf $n = 2.39 = 78$.

12. Als je om 1u 's nachts langs de N41 naar huis rijdt heb je 1 kans op 30 om een alcoholcontrole te moeten ondergaan.



Hoeveel keer moet je daar passeren om 1u 's nachts om 90% zeker te zijn dat je (minstens) een alcoholcontrole tegenkomt?

$P(B(n,1/30) \geq 1) = 1 - P(B(n,1/30)=0) > 90\% \Leftrightarrow n \geq 68$

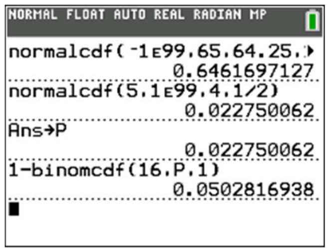
Dit kan echter ook zonder GRM: $1 - P(B(n,1/30)=0) > 90\% \Leftrightarrow \left(\frac{29}{30}\right)^n < 0{,}1 \Leftrightarrow n > \frac{\log 0{,}1}{\log(29/30)} \approx 67{,}92$

13. Voor de lengte L van een liedje uit een hele verzameling liedjes geldt $L \sim N(\mu_L, \sigma_L)$, met $\mu_L = 4 \text{ min}$ en $\sigma_L = 30 \text{ sec}$.

a) DJ Didier stelt een playlist samen van 16 willekeurige liedjes. Hoe is de totale lengte van deze playlist verdeeld? Bereken het gemiddelde en de standaardafwijking.

Ook de totale lengte Pl is normaal verdeeld met $\mu_{Pl} = 16.4 = 64$ en $\sigma_{Pl} = \sqrt{16} \cdot \frac{1}{2} = 2$

b) Op een CD is er plaats voor 65 minuten muziek. Bereken de kans dat Didier zijn playlist op één CD past als je weet dat de brandsoftware tussen elk liedje een pauze van 1 seconde plaatst.



Het gemiddelde neemt dan toe met 15 seconden, dus $\mu_{Pl'} = 64,25$, maar de standaardafwijking blijft dezelfde.

$P(Pl' \leq 65) \approx 0,6462 \Rightarrow$ Er is een vrij grote kans dat de playlist op de CD past: zo’n 64,62%

c) Bereken de kans dat er bij zo’n playlist minstens twee liedjes zijn die langer dan 5 minuten duren.

De kans dat één liedje langer dan 5 minuten duurt is $P(L > 5) \approx 0,02275$

De kans dat er van die 16 liedjes minstens twee langer duren is dan binomiaal te berekenen. We krijgen:

$P(B(16, p) \geq 2) = 1 - P(B(16, p) \leq 1) \approx 0,0503$, dus ongeveer 5% kans.