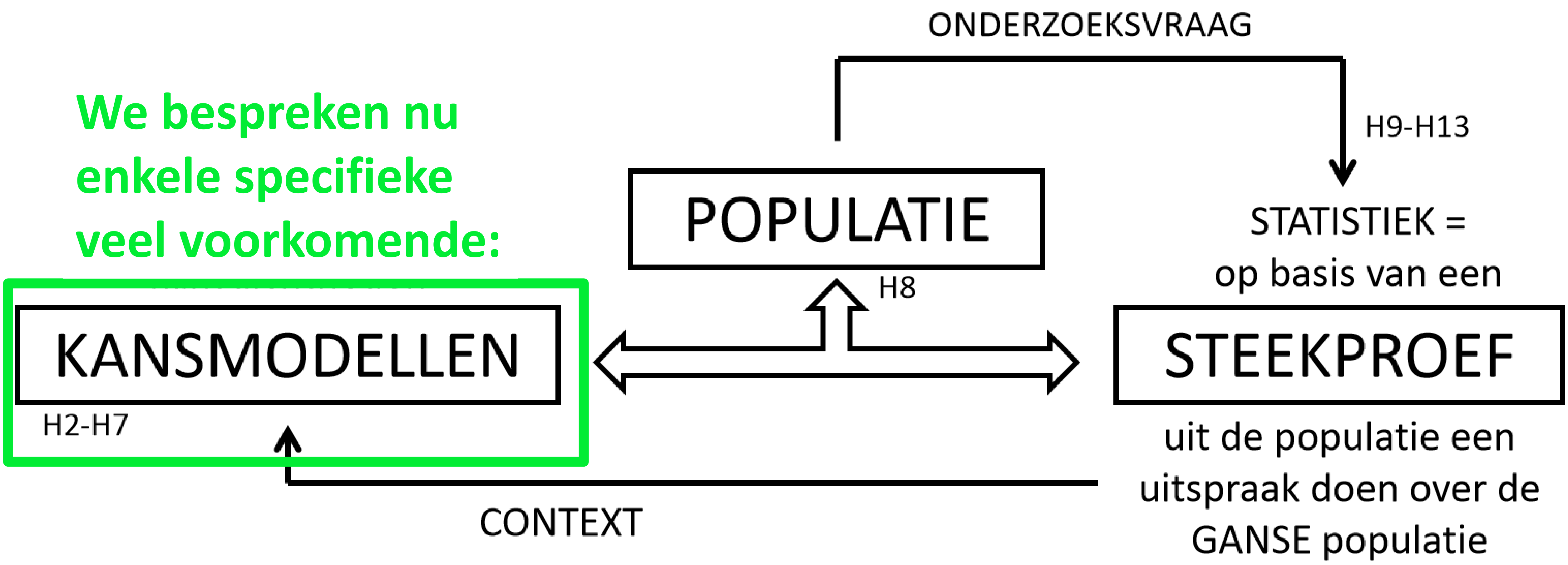
Even herhalen...

We bespreken nu enkele specifieke veel voorkomende:



Statistiek: les 3

a. Binomiale verdeling
b. Normale verdeling (zie ook week 2)
c. Lineaire combinaties (zie ook week 2) – het steekproefgemiddelde
d. Poisson verdeling
~~e. Exponentiële verdeling~~ (geen examenleerstof meer)
f. Benaderende verdelingen

Sabine Bertho sabine.bertho@kuleuven.be
Giovanni Vanroelen giovanni.vanroelen@uhasselt.be

Binomiaal kansexperiment: voorbeeld

Je gooit 5x met een dobbelsteen. Wat is de kans dat je hierbij 2x een zes gooit?

- P(eerst 2x zes, daarna 3x geen zes) $= \frac{1}{6} \cdot \frac{1}{6} \cdot \frac{5}{6} \cdot \frac{5}{6} \cdot \frac{5}{6} = \frac{125}{7776}$

Maar eerst 3x geen zes en daarna 2x zes is ook ok! Of alle andere combinaties:

Plaatskaartjes:

1	2	3	4	5

$$C_5^2 = C_5^3$$

Binomiaal kansexperiment: voorbeeld

Je gooit <u>5x</u> met een dobbelsteen. Wat is de kans dat je hierbij 2x een zes gooit?

- P(eerst 2x zes, daarna 3x geen zes) $= \frac{1}{6} \cdot \frac{1}{6} \cdot \frac{5}{6} \cdot \frac{5}{6} \cdot \frac{5}{6} = \frac{125}{7776}$

Maar eerst 3x geen zes en daarna 2x zes is ook ok! Of alle andere combinaties:

$$P(2x\text{ een zes}) = \underbrace{\frac{1}{6} \cdot \frac{1}{6}}_{\text{2x succes}} \cdot \underbrace{\frac{5}{6} \cdot \frac{5}{6} \cdot \frac{5}{6}}_{\text{3x mislukking}} \cdot \underbrace{C_5^2}_{\text{mogelijke combinaties}} = \frac{125}{7776} \cdot nCr(5,2) = \frac{625}{3888} = 0{,}160751$$

Binomiale kansexperiment: voorwaarden

Als een experiment aan de volgende voorwaarden voldoet, spreken we van een **binomiaal experiment**.

- Het experiment bestaat uit n deelexperimenten die onderling onafhankelijk zijn.
- Elk deelexperiment kan 2 waarden aannemen, die we noteren met 1 (*succes*) en 0 (*mislukking*).
- De kans op succes is voor elk deelexperiment gelijk aan dezelfde kans p. De kans op mislukking, aangegeven met $q = 1 - p$, verandert dus ook niet.

Notatie: $X \sim B(n, p)$

Binomiaal kansexperiment: voorbeelden

Binomiaal experiment	Parameters	Binomiaal verdeelde variabele X
Gokken op 3 meerkeuzevragen, elk met 4 mogelijke antwoorden.	$n = 3, p = 1/4$	X: aantal correcte antwoorden
Acht keer een muntstuk opgooien.	$n = 8, p = 1/2$	X: aantal keer kop
Van een lading van 200 radio's weet men dat er 40 defect zijn. Je neemt er 10 willekeurig uit (met teruglegging) en controleert of ze al of niet defect zijn.	$n = 10, p = 0.2$	X: aantal defecte radio's

Binomiale kansfunctie: formularium en rekentoestel

Wanneer een kansexperiment bestaat uit n *onafhankelijke* deelexperimenten met elk **dezelfde** kans op succes (p) spreken we van een **binomiaal kansexperiment**. De kansvariabele X = het aantal keer succes, volgt dan de binomiale kansverdeling: $X \sim B(n,p)$.

De bijhorende kansfunctie: $P(X=i) = \binom{n}{i} p^i (1-p)^{n-i}$ ← rekentoestel: binomPdf(n,p,i)

Voor een binomiale verdeling geldt: $\boxed{\mu_X = np}$ en $\sigma_X = \sqrt{np(1-p)}$.

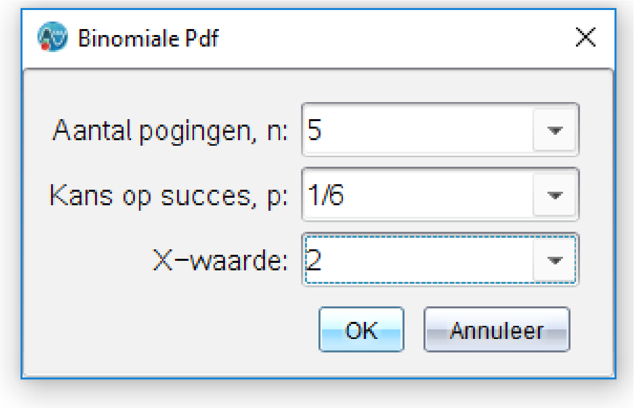
Stel dat in onze opleiding 10% meisjes zitten. Als je een groep van 50 studenten ziet, hoeveel meisjes verwacht je daar dan bij?

Binomiale kansfunctie: rekentoestel

TI-Nspire: menu – 5: Kansen – 5: Verdelingen... – A: Binomiale Pdf...

Voorbeeld: 2x zes gooien bij 5 worpen met een dobbelsteen:

$$\text{binomPdf}\left(5,\frac{1}{6},2\right)=0{,}160751$$

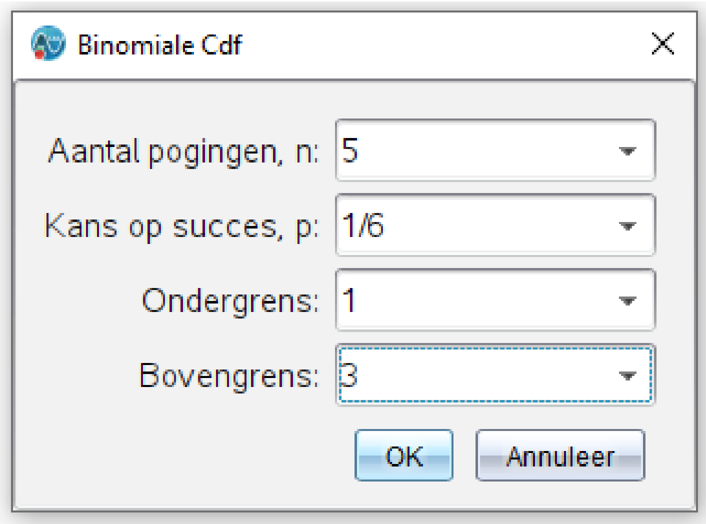


Of ook: menu – 5: Kansen – 5: Verdelingen... – B: Binomiale Cdf...

Voorbeeld: 1x <u>of</u> 2x <u>of</u> 3x een zes gooien bij 5 worpen met een dobbelsteen:

$$\text{binomCdf}\left(5,\frac{1}{6},1{,}3\right)=0{,}594779$$

interval



UHASSELT KU LEUVEN

Meerkeuzevraag

Bij Planet Yoghurt in Hasselt zijn aardbeien in de zomer een zeer populaire topping op frozen yoghurt (40% van de klanten kiest aardbeien). Op een dag zijn er te weinig aarbeien geleverd. In de namiddag ziet de uitbaatster dat er nog genoeg aardbeien zijn om 30 toppings te maken. Als er die dag nog 90 klanten komen, wat is dan de kans P dat elke liefhebber van aardbeien zijn favoriete topping kan kiezen?

A. $0{,}05 \leq P < 0{,}1$

B. $0{,}1 \leq P < 0{,}15$

C. $0{,}15 \leq P < 0{,}2$

D. $0{,}2 \leq P < 0{,}25$

Meerkeuzevraag

Bij Planet Yoghurt in Hasselt zijn aardbeien in de zomer een zeer populaire topping op frozen yoghurt (40% van de klanten kiest aardbeien). Op een dag zijn er te weinig aarbeien geleverd. In de namiddag ziet de uitbaatster dat er nog genoeg aardbeien zijn om 30 toppings te maken. Als er die dag nog 90 klanten komen, wat is dan de kans P dat elke liefhebber van aardbeien zijn favoriete topping kan kiezen?

Ga naar **wooclap.com** en gebruik de code **STAT2802**

Hoe definieer je de binomiaal verdeelde kansvariabele? X = ...

Meerkeuzevraag

Bij Planet Yoghurt in Hasselt zijn aardbeien in de zomer een zeer populaire topping op frozen yoghurt (40% van de klanten kiest aardbeien). Op een dag zijn er te weinig aarbeien geleverd. In de namiddag ziet de uitbaatster dat er nog genoeg aardbeien zijn om 30 toppings te maken. Als er die dag nog 90 klanten komen, wat is dan de kans P dat elke liefhebber van aardbeien zijn favoriete topping kan kiezen?

Ga naar **wooclap.com** en gebruik de code **STAT2802**

Welke binomiale verdeling volgt de kansvariabele?

wooclap

100 %

0

Meerkeuzevraag

Bij Planet Yoghurt in Hasselt zijn aardbeien in de zomer een zeer populaire topping op frozen yoghurt (40% van de klanten kiest aardbeien). Op een dag zijn er te weinig aarbeien geleverd. In de namiddag ziet de uitbaatster dat er nog genoeg aardbeien zijn om 30 toppings te maken. Als er die dag nog 90 klanten komen, wat is dan de kans P dat elke liefhebber van aardbeien zijn favoriete topping kan kiezen?

Ga naar **wooclap.com** en gebruik de code **STAT2802**

Schrijf de gezochte kans m.b.v. de kansvariabele X.

wooclap

100 %

0

Meerkeuzevraag

Bij Planet Yoghurt in Hasselt zijn aardbeien in de zomer een zeer populaire topping op frozen yoghurt (40% van de klanten kiest aardbeien). Op een dag zijn er te weinig aarbeien geleverd. In de namiddag ziet de uitbaatster dat er nog genoeg aardbeien zijn om 30 toppings te maken. Als er die dag nog 90 klanten komen, wat is dan de kans P dat elke liefhebber van aardbeien zijn favoriete topping kan kiezen?

U kunt niet langer stemmen

1. $0{,}05 \le P < 0{,}1$ — 0% — 0
2. $0{,}1 \le P < 0{,}15$ — 0% — 0 ✓
3. $0{,}15 \le P < 0{,}2$ — 0% — 0
4. $0{,}2 \le P < 0{,}25$ — 0% — 0

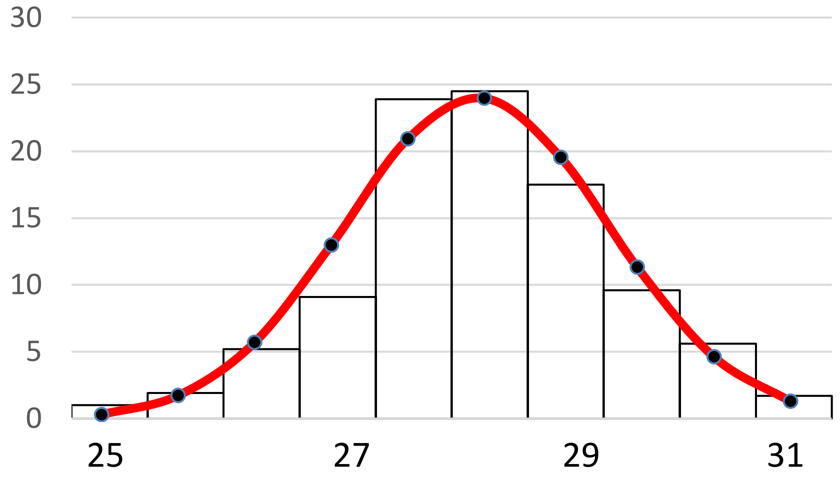
100 %

0% juist 0 / 0

Déjà vu…

<u>Grootte van de voet</u>

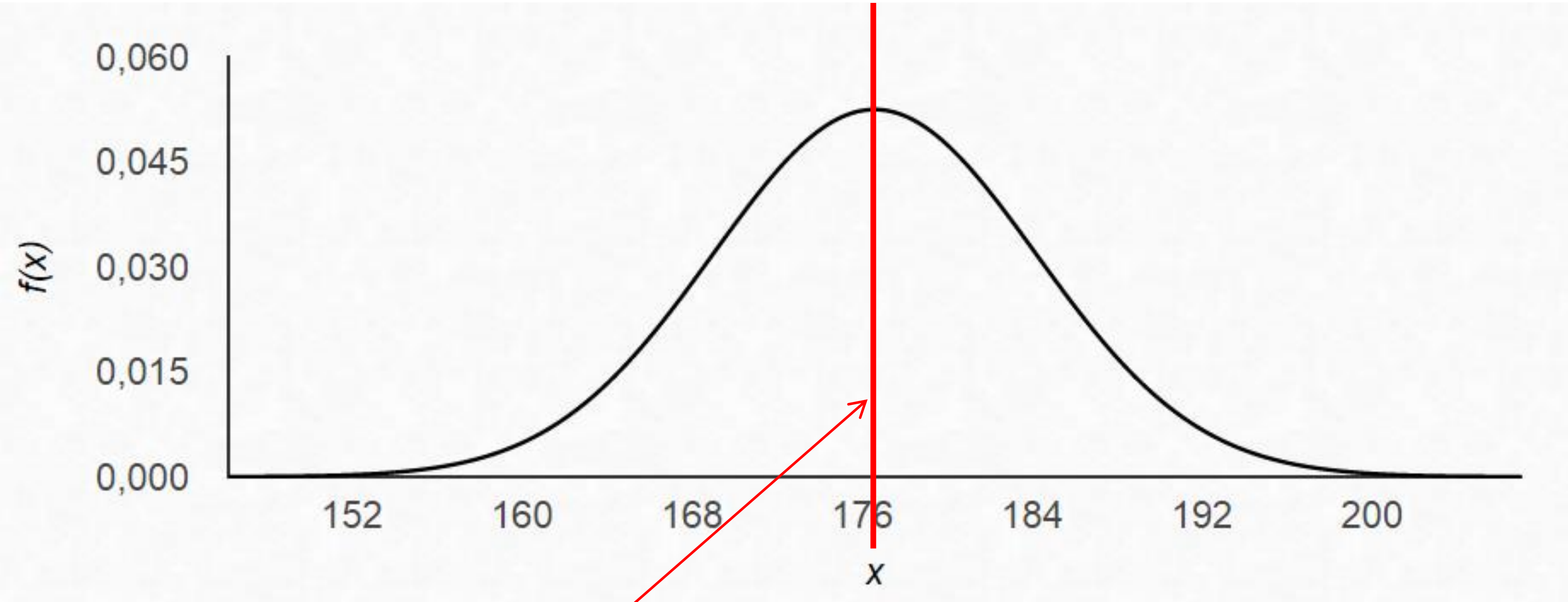
CONTINUE
KANSVARIABELE



CONTINUE
KANSVERDELING

Normale kansdichtheid

Voorbeeld: de lengte van Belgische mannen (2005) heeft volgende kansdichtheid:



$$X \sim N(\mu, \sigma)$$

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

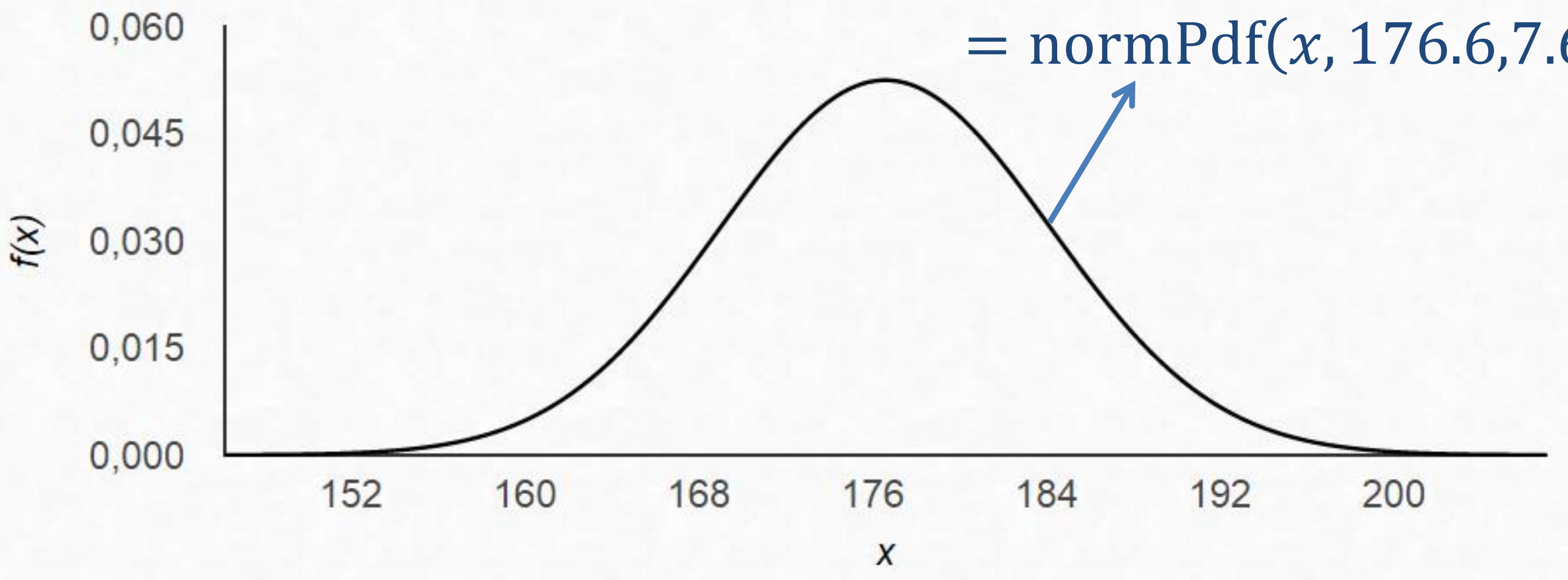
$\boxed{\mu = E(X) = 176.6}$ $\boxed{\sigma = SD(X) = 7.6}$ $\sigma^2 = Var(X) = 57.76$

Hoe kleiner σ, hoe kleiner de spreiding, hoe scherper de piek

Normale kansdichtheid

Voorbeeld: de lengte van Belgische mannen (2005) heeft volgende kansdichtheid:

Continue kansvariabele



$= \mathrm{normPdf}(x, 176.6, 7.6)$ = hoogte kansdichtheid

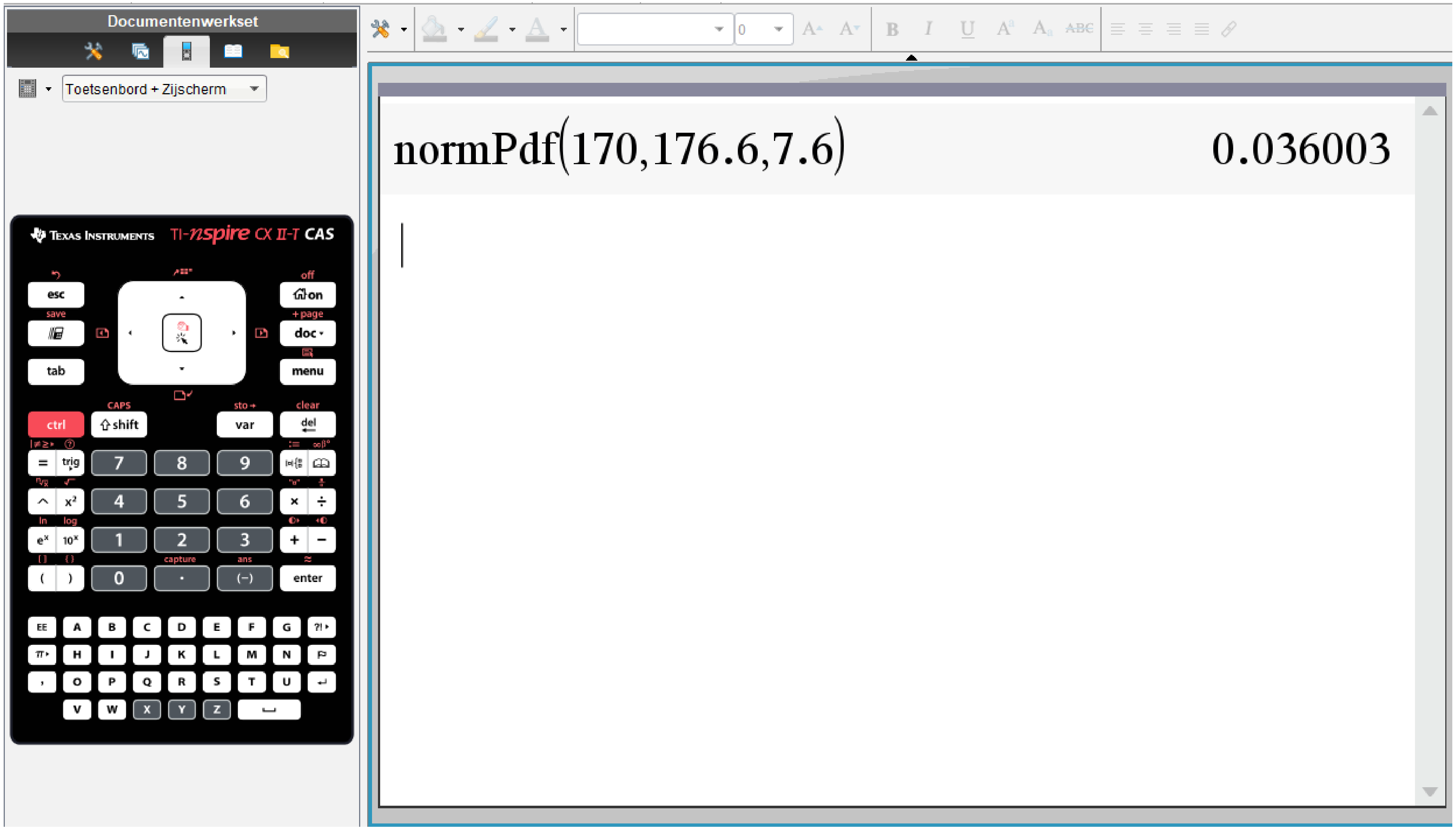
$\neq$ kans

vs. binomPdf = kans

Discrete kansvariabele

$\mu = E(X) = 176.6 \qquad \sigma = SD(X) = 7.6 \qquad \sigma^2 = Var(X) = 57.76$

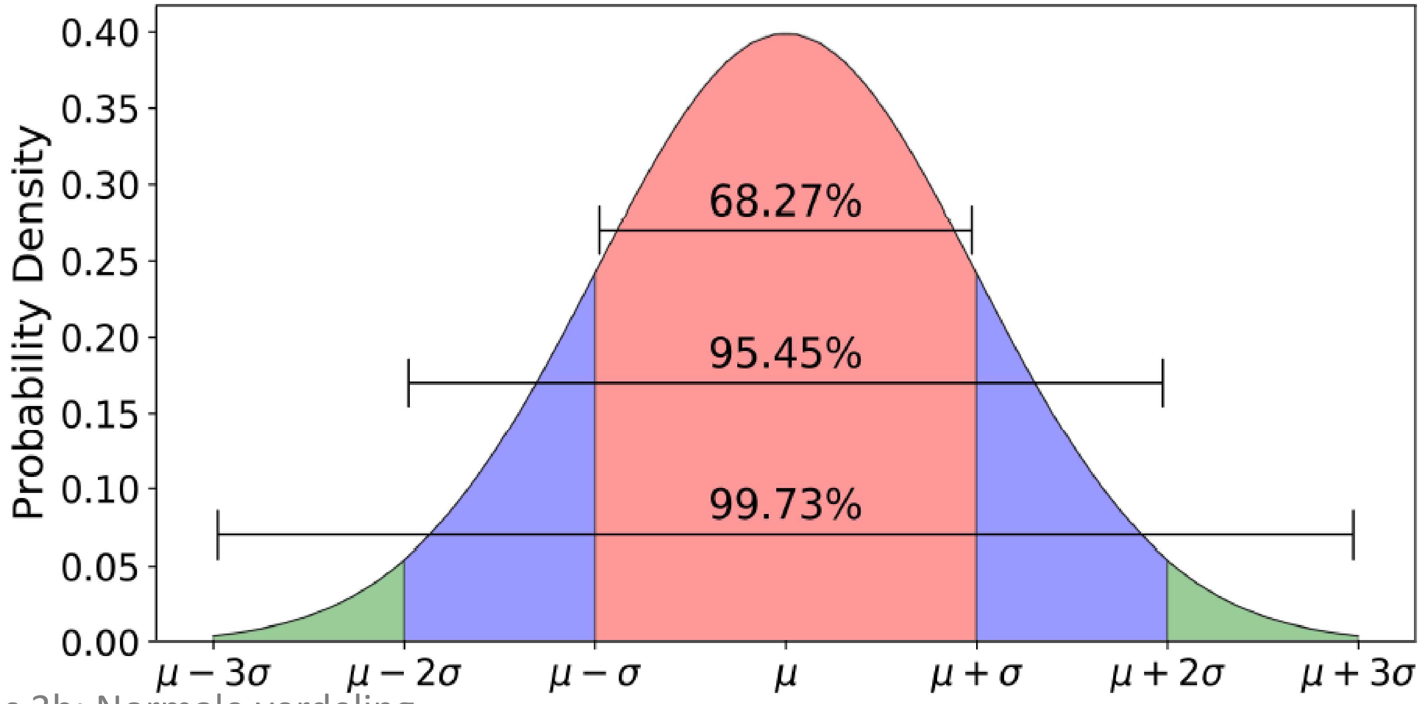
Normale kansdichtheid



Algemene normale kansdichtheid: eigenschappen

Eigenschap 6.6 *Een normaal verdeelde variabele neemt*

- *in* 68.3% *van de gevallen een waarde aan tussen* $\mu - \sigma$ *en* $\mu + \sigma$.
- *in* 95.5% *van de gevallen een waarde aan tussen* $\mu - 2\sigma$ *en* $\mu + 2\sigma$.
- *in* 99.7% *van de gevallen een waarde aan tussen* $\mu - 3\sigma$ *en* $\mu + 3\sigma$.



Reken zelf na!

normCdf(128,132,130,2) = …

normCdf(75,95,85,5) = …

normCdf(750,1650,1200,150) = …

Normale verdeling: rekentoestel

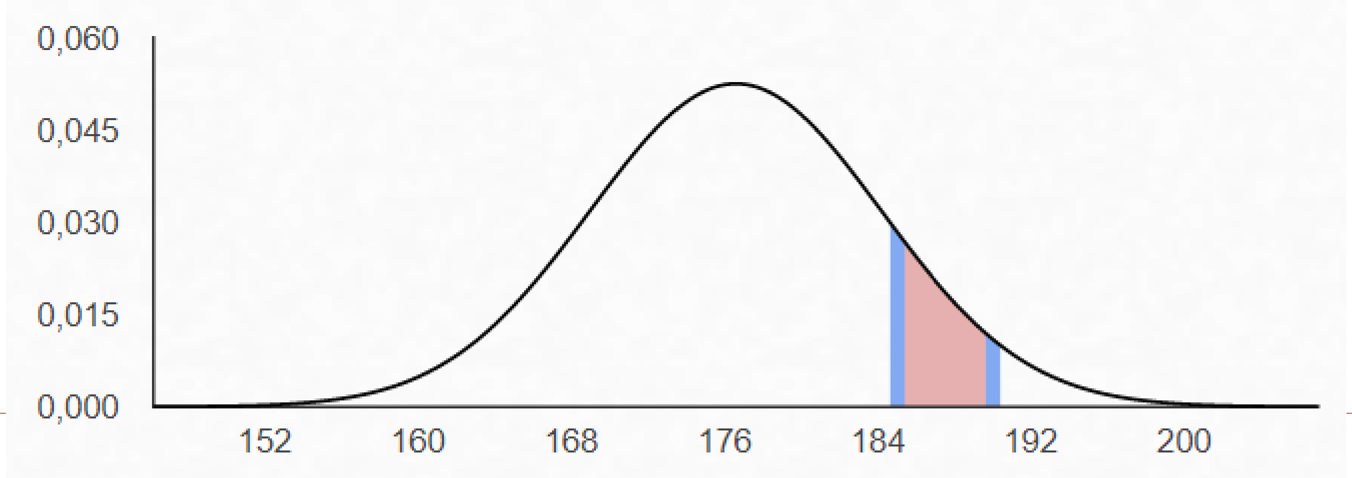
- menu – 6: Statistieken – 5: Verdelingen... – 1: Normale Pdf...

  Berekent de functiewaarde van de dichtheidsfunctie (zullen we bijna niet gebruiken)

- menu – 5: Kansen – 5: Verdelingen... – 2: Normale Cdf...

  Voorbeeld: Bereken de kans dat de lengte van een Belgische man tussen 185 cm en 190 cm ligt.

$$\mathrm{P}(185 < X < 190) = \mathrm{normCdf}(185, 190, 176.6, 7.6) = 0.095586$$

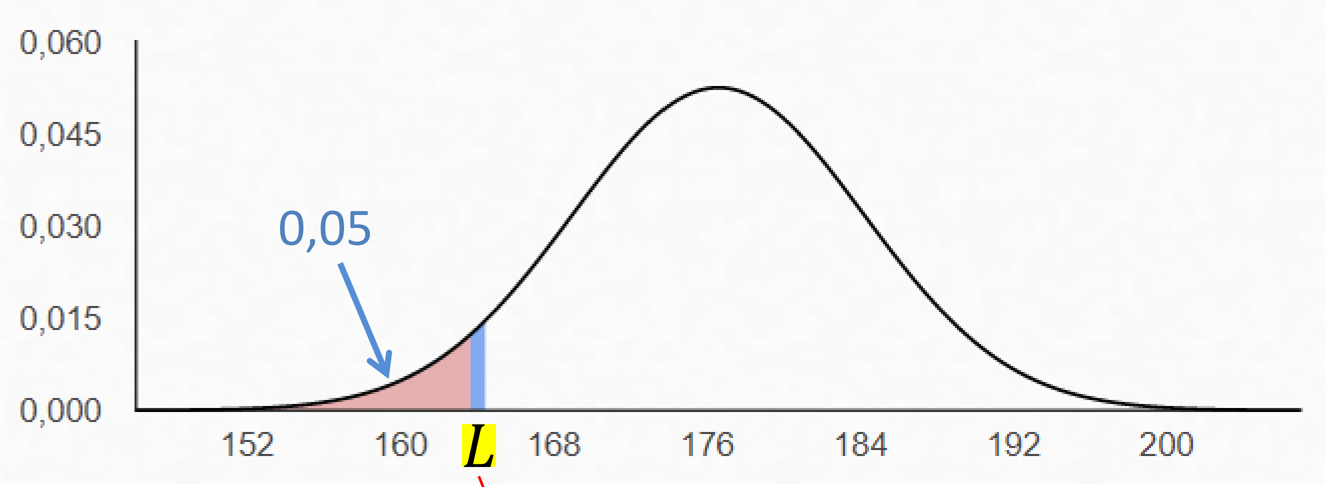


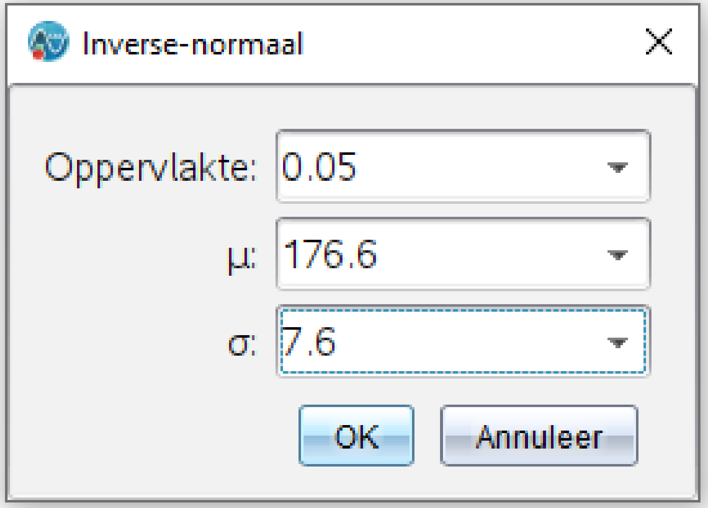
Normale verdeling: rekentoestel

- menu – 6: Statistieken – 5: Verdelingen… – 3: Inverse-normaal…

Voorbeeld: Bij welke lengte ligt het 5% percentiel (of het 0,05-kwantiel) van de Belgische mannen?

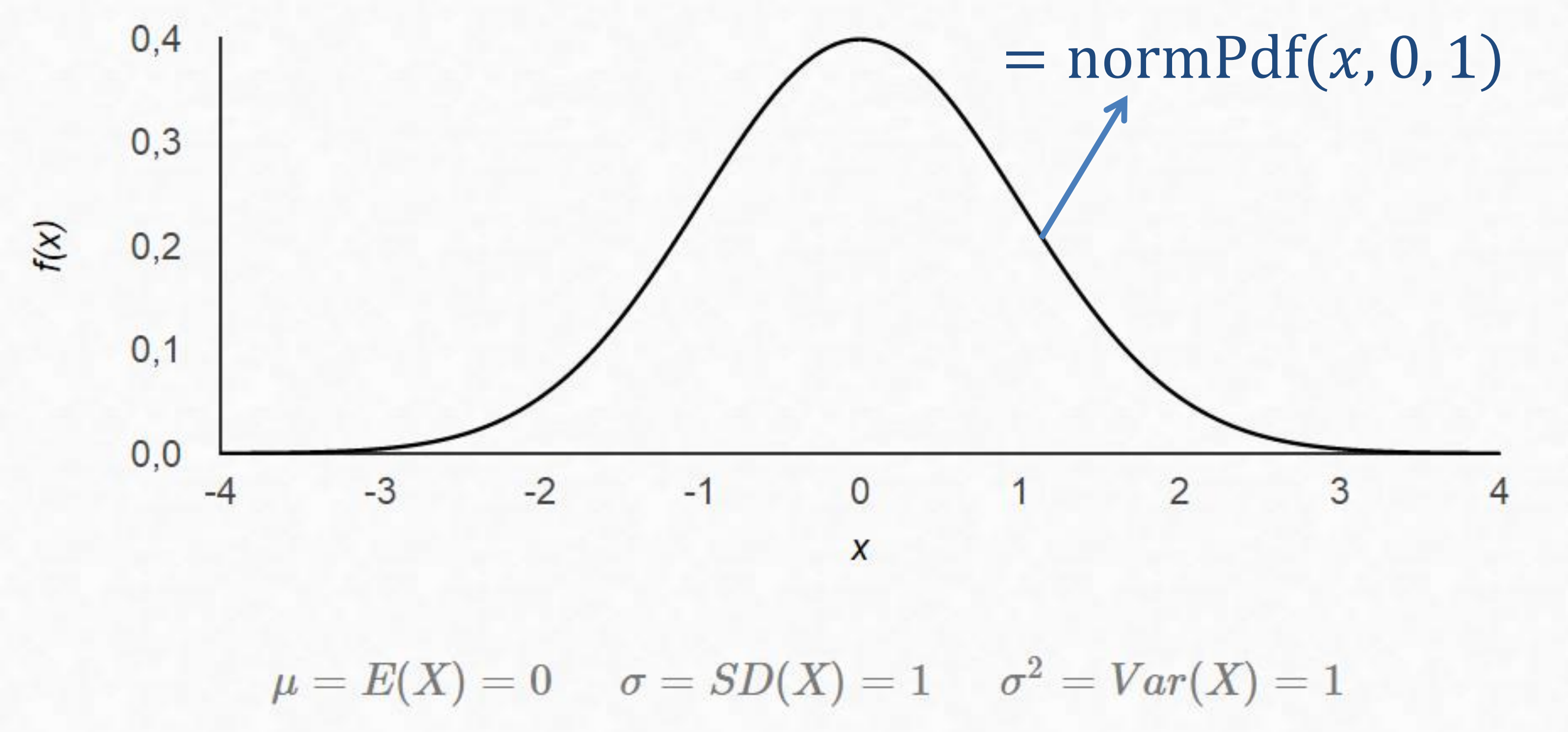
$$P(X < L) = 0.05 \Rightarrow L = \text{invNorm}(0.05, 176.6, 7.6) = 164.099 \text{ cm}$$





waarde op horizontale as zoeken

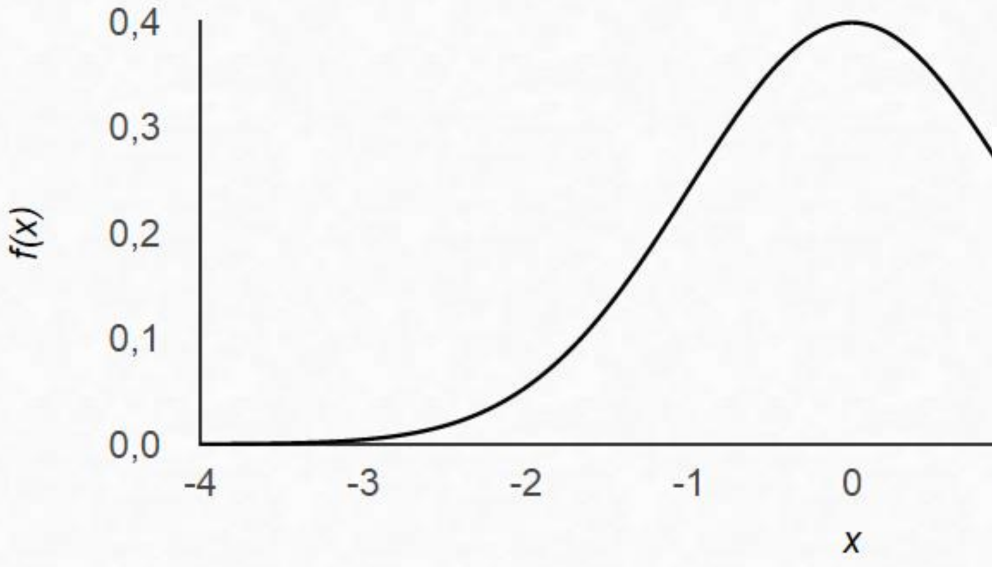
De standaard normale kansdichtheid



$$Z \sim N(0, 1)$$

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

De standaard normale kansdichtheid



$= \mathrm{normPdf}(x, 0, 1)$

$\mu = E(X) = 0 \qquad \sigma = SD(X) = 1$

$Z \sim N(0, 1)$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.

Gestandaardiseerde kansvariabele

Eigenschap 6.4 *Als $X \sim N(\mu, \sigma)$, dan heeft de gestandaardiseerde kansvariabele*

$$Z = \frac{X - \mu}{\sigma}$$

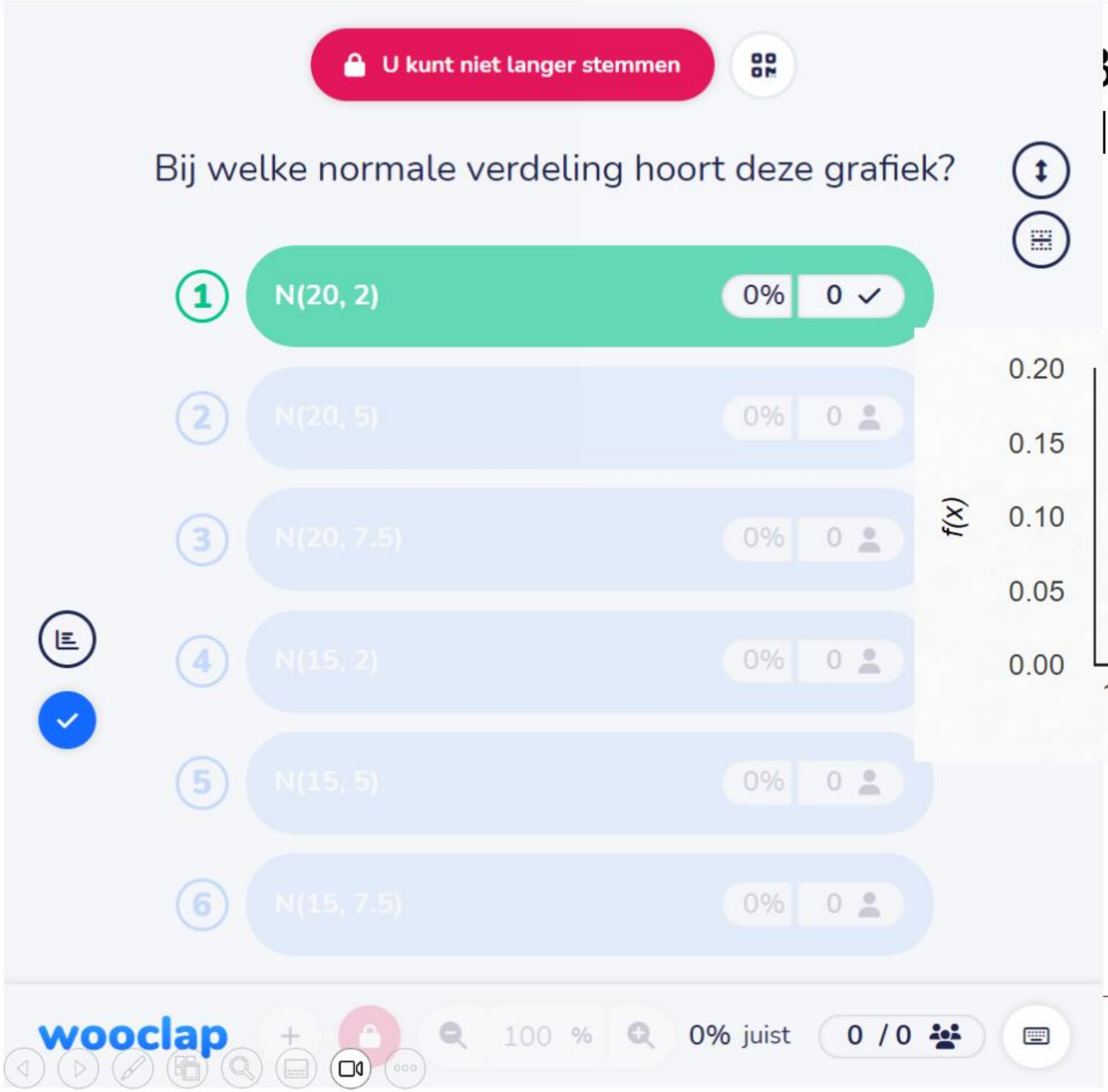
een standaard normale kansdichtheid $N(0,1)$ met verwachting 0 en standaardafwijking 1.

LINEAIRE COMBINATIES: Als $X_1, X_2, \ldots, X_n$ onderling **onafhankelijke** normaal verdeelde toevalsvariabelen zijn met verwachting μ_i en standaardafwijking σ_i voor *i* = 1, 2, …, *n* dan is $Y = a_1X_1 + a_2X_2 + \cdots + a_nX_n + b$ ook normaal verdeeld met verwachting $\mu = a_1\mu_1 + a_2\mu_2 + \cdots + a_n\mu_n + b$ en standaardafwijking $\sigma = \sqrt{a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + \cdots + a_n^2\sigma_n^2}$. *(opmerking: hierbij mogen de constanten a_i en b negatief zijn !!)*

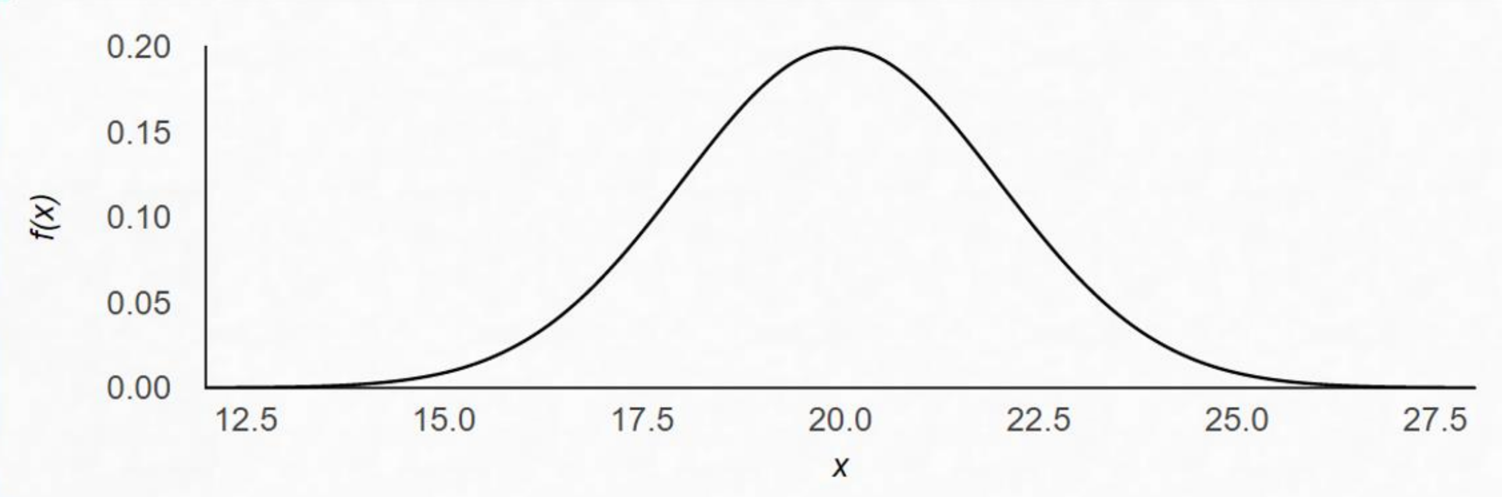
Oef 39 p17 – zie werkzitting

De levensduur van een machine is normaal verdeeld met een verwachtingswaarde van 3000 uren. Uit ervaring is gebleken dat 50% van deze machines minder dan 2632 uren of meer dan 3368 uren werken. Bereken de standaardafwijking van de levensduur van deze machines.

Meerkeuzevraag



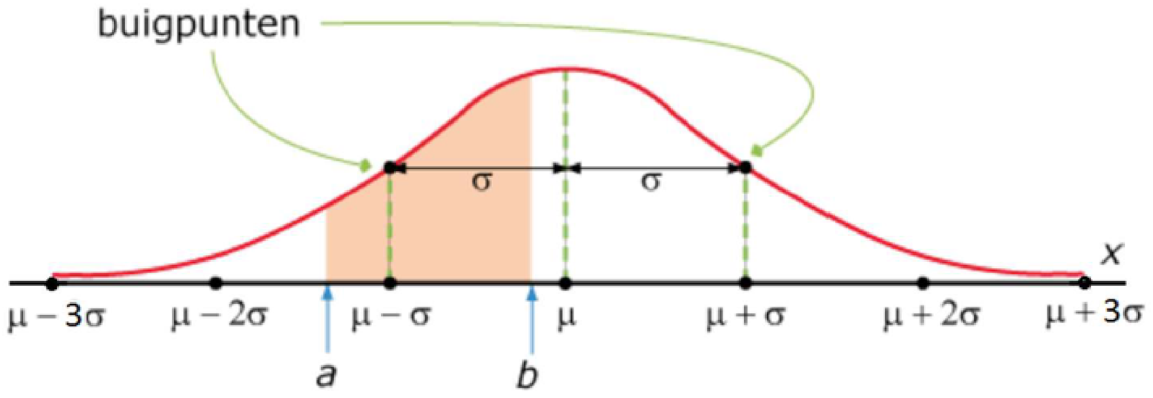
Bij welke normale verdeling hoort deze grafiek?



Formularium

KANSDICHTHEIDSFUNCTIE: $f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ ← rekentoestel: normPdf(*x, μ, σ*)

NOTATIE: $X \sim N(\mu, \sigma)$ → E(*X*) = μ, Var(*X*) = σ²



KANSVERDELINGSFUNCTIE: $P(X \le x) = \int_{-\infty}^{x} f_X(t)dt$ ← rekentoestel: normCdf(-∞, x, μ, σ)

KANS MET ONDER- en BOVENGRENS: $P(a \le X \le b)$ = normCdf(*a, b, μ, σ*)

BELANGRIJKSTE σ-NIVEAUS: Voor elke normale kansverdeling $X \sim N(\mu, \sigma)$ geldt

$P(\mu - \sigma \le X \le \mu + \sigma) = 68{,}27\%$ $\quad$ $P(\mu - 1.645\sigma \le X \le \mu + 1.645\sigma) \approx 90\%$

$P(\mu - 2\sigma \le X \le \mu + 2\sigma) = 95{,}45\%$ $\quad$ $P(\mu - 1.96\sigma \le X \le \mu + 1.96\sigma) \approx 95\%$

$P(\mu - 3\sigma \le X \le \mu + 3\sigma) = 99{,}73\%$ $\quad$ $P(\mu - 2.576\sigma \le X \le \mu + 2.576\sigma) \approx 99\%$

p-KWANTIELEN (0 ≤ *p* ≤ 1): zoek *c* zodat $P(X \le c) = p$ ← rekentoestel: invNorm(*p*, μ, σ)

LINK met **STANDAARDNORMALE VERDELING:** Als $X \sim N(\mu, \sigma)$, dan heeft de gestandaardiseerde kansvariabele $Z = \frac{X-\mu}{\sigma}$ (=*Z*-score) een normale verdeling *N*(0,1) met verwachting 0 en standaardafwijking 1.

De Poissonverdeling – voorbeeld

In de kerkstraat in Diepenbeek worden regelmatig snelheidscontroles gedaan. Gemiddeld worden er 28 auto’s per uur geflitst. Wat is de kans dat er bij de volgende controle 35 auto’s per uur geflitst worden?

De Poissonverdeling – voorbeeld

In de kerkstraat in Diepenbeek worden regelmatig snelheidscontroles gedaan. Het aantal auto’s dat geflitst wordt is Poisson verdeeld met een gemiddelde van 28 auto’s per uur. Wat is de kans dat er bij de volgende controle 35 auto’s per uur geflitst worden?

Poisson: voorwaarden

1. X is het aantal gebeurtenissen in een continu interval van een bepaalde vaste grootte.
2. Het verwachte aantal gebeurtenissen in 2 intervallen van dezelfde lengte is gelijk en is evenredig met de lengte.
3. Het aantal gebeurtenissen in een interval is onafhankelijk van het aantal gebeurtenissen in een niet-overlappend interval.

Dit is niet zo gemakkelijk als het lijkt!

Voorbeeld: bakker maakt krentenbrood met gemiddeld 8 krenten per snee brood.

5	4	9	5	10	13	10	8

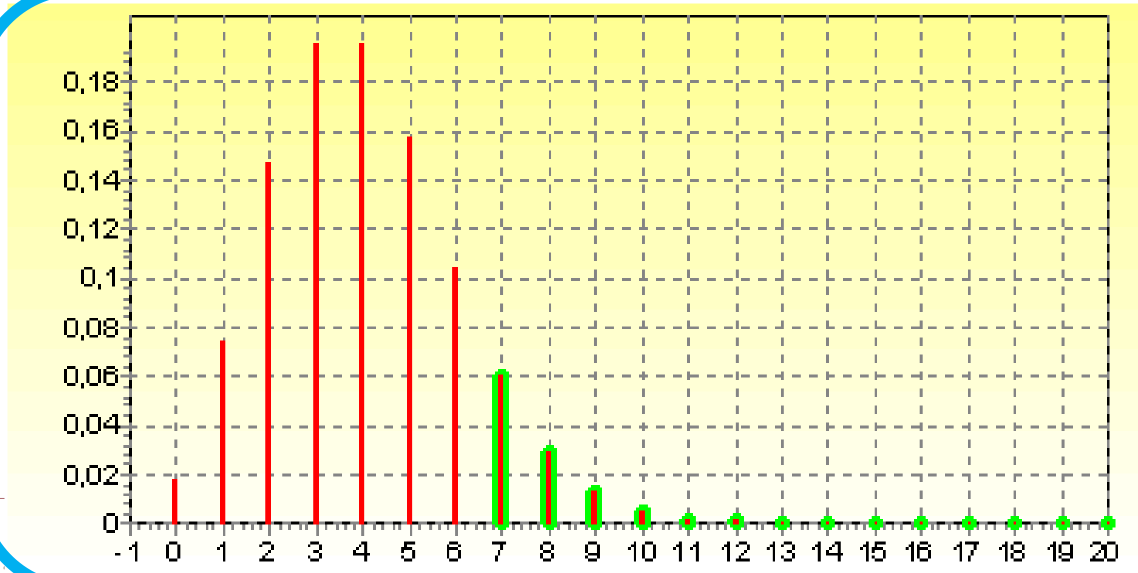
Poisson: Formularium

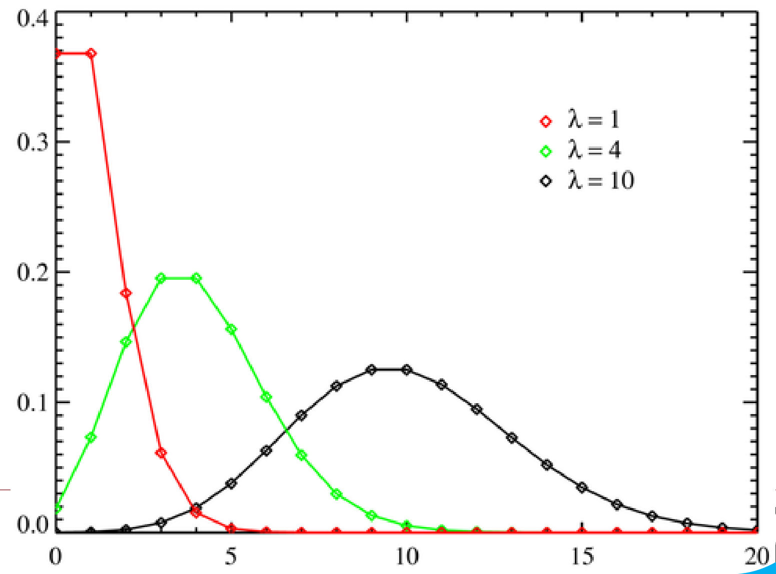
Noem X het aantal gebeurtenissen in een continu interval (bvb. een bepaalde afstand of een tijdsinterval).

Wanneer het **verwachte** aantal gebeurtenissen λ **evenredig** is met de lengte van het interval en bovendien het aantal gebeurtenissen die optreden in 2 niet-overlappende (deel)intervallen *onafhankelijk* zijn van elkaar, dan is de kansvariabele X **Poisson** verdeeld met gemiddelde λ: $X \sim Poi(\lambda)$.

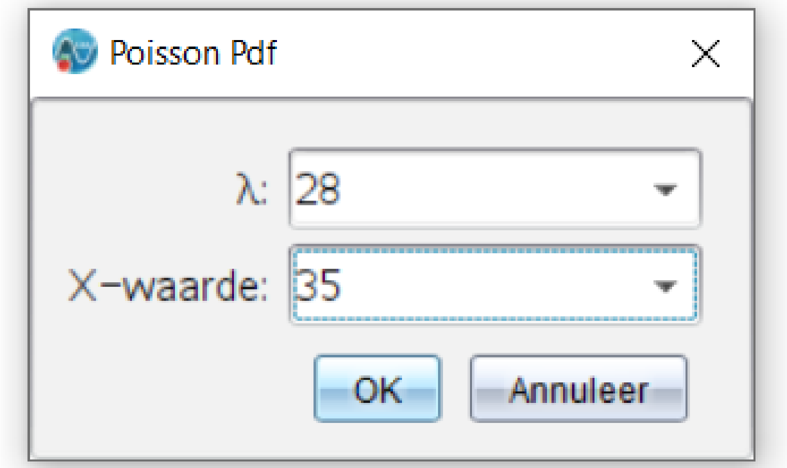
De bijhorende kansfunctie: $P(X = k) = \dfrac{e^{-\lambda}\lambda^k}{k!}$ ← rekentoestel: poissPdf(λ,k)

Voor een Poisson verdeling geldt: $\mu_X = \lambda$ en $\sigma_X = \sqrt{\lambda}$.

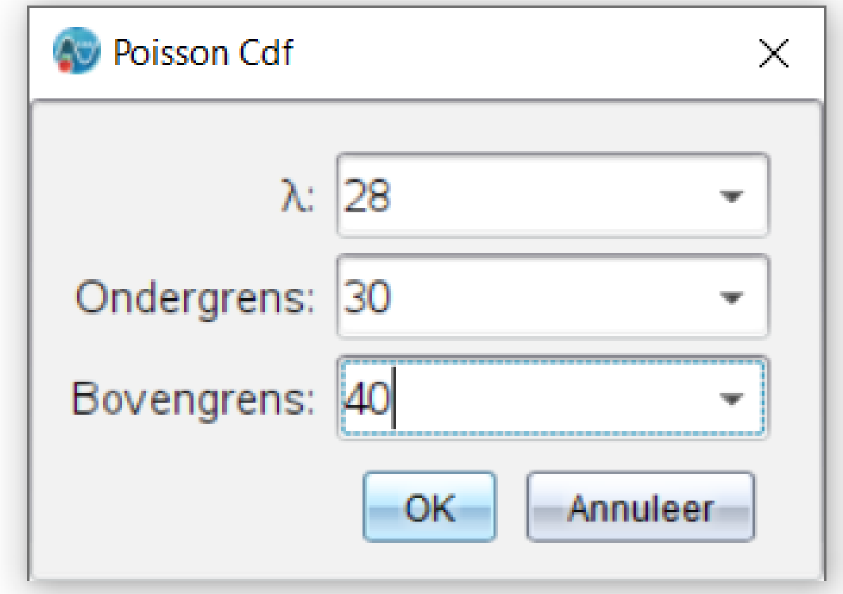




Poisson: rekentoestel



- Notatie: $X \sim Poi(\lambda)$
- TI-Nspire: menu – 6: Statistieken – 5: Verdelingen… – H: Poisson Pdf…
  of I: Poisson Cdf…



Voorbeeld: snelheidscontrole

In de kerkstraat in Diepenbeek worden regelmatig snelheidscontroles gedaan. Het aantal auto's dat geflitst wordt is Poisson verdeeld met een gemiddelde van 28 auto's per uur. Wat is de kans dat er bij de volgende controle 35 auto's per uur geflitst worden?

$$\mathrm{P}(35\ \mathrm{auto's}) = \mathrm{poissPdf}(28{,}35) = 0.029926$$

Wat is de kans dat er tussen de 30 en de 40 auto's per uur geflitst worden?

$$\mathrm{P}(30\ \mathrm{tot}\ 40\ \mathrm{auto's}) = \mathrm{poissCdf}(28{,}30{,}40) = 0.364925$$

Meerkeuzevraag

Het aantal bezoekers per minuut op een webpagina is Poisson verdeeld. Als het gemiddeld aantal bezoekers per minuut gelijk is aan 4, wat is dan de kans P dat er meer dan 25 bezoekers zijn in 5 minuten?

A. $0.05 \leq P \leq 0.10$

B. $0.10 \leq P \leq 0.15$

C. $0.15 \leq P \leq 0.20$

D. $0.20 \leq P \leq 0.25$

Meerkeuzevraag

Het aantal bezoekers per minuut op een webpagina is Poisson verdeeld. Als het gemiddeld aantal bezoekers per minuut gelijk is aan 4, wat is dan de kans P dat er meer dan 25 bezoekers zijn in 5 minuten?

Ga naar **wooclap.com** en gebruik de code **STAT2802**

Hoe definieer je de kansvariabele bij deze vraag? X = ...

wooclap

100 %

0

Meerkeuzevraag

Het aantal bezoekers per minuut op een webpagina is Poisson verdeeld. Als het gemiddeld aantal bezoekers per minuut gelijk is aan 4, wat is dan de kans P dat er meer dan 25 bezoekers zijn in 5 minuten?

Ga naar **wooclap.com** en gebruik de code **STAT2802**

Welke kansverdeling volgt de kansvariabele?

wooclap

100 %

0

Meerkeuzevraag

Het aantal bezoekers per minuut op een webpagina is Poisson verdeeld. Als het gemiddeld aantal bezoekers per minuut gelijk is aan 4, wat is dan de kans P dat er meer dan 25 bezoekers zijn in 5 minuten?

Ga naar **wooclap.com** en gebruik de code **STAT2802**

Schrijf de gezochte kans m.b.v. de kansvariabele X.

wooclap

100 %

0

Meerkeuzevraag

Het aantal bezoekers per minuut op een webpagina is Poisson verdeeld. Als het gemiddeld aantal bezoekers per minuut gelijk is aan 4, wat is dan de kans P dat er meer dan 25 bezoekers zijn in 5 minuten?

U kunt niet langer stemmen

En nu nog berekenen! In welk interval ligt de gevraagde kans P?

1. $0.05 \leq P \leq 0.10$ — 0% — 0
2. $0.10 \leq P \leq 0.15$ — 0% — 0 ✓
3. $0.15 \leq P \leq 0.20$ — 0% — 0
4. $0.20 \leq P \leq 0.25$ — 0% — 0

wooclap

100 %

0% juist

0 / 0

Lineaire combinaties van normaal verdeelde kansvariabelen

Déjà vu: formularium

LINEAIRE COMBINATIES: Als $X_1, X_2, \ldots, X_n$ onderling **onafhankelijke** ~~normaal verdeelde~~ toevalsvariabelen zijn met verwachting μ_i en standaardafwijking σ_i voor *i* = 1, 2, ..., *n* dan is $Y = a_1X_1 + a_2X_2 + \cdots + a_nX_n + b$ ~~ook normaal verdeeld~~ met verwachting $\mu = a_1\mu_1 + a_2\mu_2 + \cdots + a_n\mu_n + b$ en standaardafwijking $\sigma = \sqrt{a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + \cdots + a_n^2\sigma_n^2}$. *(opmerking: hierbij mogen de constanten a_i en b negatief zijn !!)*

Lineaire combinaties: (steekproef)gemiddelde

Eerste schroef:

$X_1 \sim N(\mu, \sigma)$

Tweede schroef:

$X_2 \sim N(\mu, \sigma)$

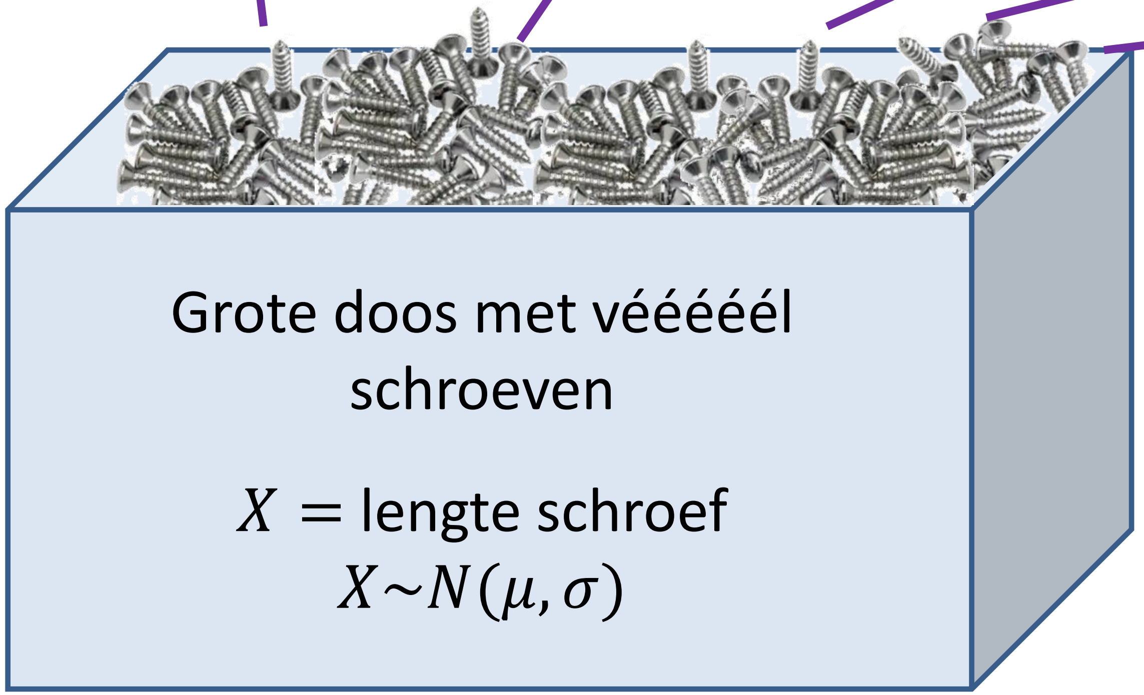
Derde schroef:

$X_3 \sim N(\mu, \sigma)$

…

n-de schroef:

$X_n \sim N(\mu, \sigma)$



Steekproefgemiddelde:

$$\bar{X} = \frac{1}{n} \cdot (X_1 + X_2 + \cdots + X_n)$$

Lineaire combinaties: (steekproef)gemiddelde

Gevolg 6.9 *Als $X_i \sim N(\mu, \sigma)$, $i = 1, \ldots, n$ onderling onafhankelijk, dan is*

$$\bar{X} = \frac{1}{n}(X_1 + \ldots + X_n)$$

ook normaal verdeeld met gemiddelde μ en standaardafwijking $\frac{\sigma}{\sqrt{n}}$.

$$\mu_{\bar{X}} = \frac{1}{n} \cdot \mu_{X_1} + \frac{1}{n} \cdot \mu_{X_2} + \cdots + \frac{1}{n} \cdot \mu_{X_n} = \frac{1}{n} \cdot n \cdot \mu = \mu$$

LINEAIRE COMBINATIES: Als $X_1, X_2, \ldots, X_n$ onderling **onafhankelijke** normaal verdeelde toevalsvariabelen zijn met verwachting μ_i en standaardafwijking σ_i voor *i* = 1, 2, …, *n* dan is $Y = a_1X_1 + a_2X_2 + \cdots + a_nX_n + b$ ook normaal verdeeld met verwachting $\mu = a_1\mu_1 + a_2\mu_2 + \cdots + a_n\mu_n + b$ en standaardafwijking $\sigma = \sqrt{a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + \cdots + a_n^2\sigma_n^2}$. *(opmerking: hierbij mogen de constanten a_i en b negatief zijn !!)*

Lineaire combinaties: (steekproef)gemiddelde

Gevolg 6.9 *Als* $X_i \sim N(\mu, \sigma)$, $i = 1, \ldots, n$ *onderling onafhankelijk, dan is*

$$\bar{X} = \frac{1}{n}(X_1 + \ldots + X_n)$$

ook normaal verdeeld met gemiddelde μ *en standaardafwijking* $\dfrac{\sigma}{\sqrt{n}}$.

$$\mu_{\bar{X}} = \frac{1}{n} \cdot \mu_{X_1} + \frac{1}{n} \cdot \mu_{X_2} + \cdots + \frac{1}{n} \cdot \mu_{X_n} = \frac{1}{n} \cdot n \cdot \mu = \mu$$

$$\mu_{\bar{X}} = \mu$$

Het gemiddelde van het steekproefgemiddelde is gelijk aan het populatiegemiddelde

Lineaire combinaties: (steekproef)gemiddelde

Gevolg 6.9 *Als* $X_i \sim N(\mu, \sigma)$, $i = 1, \ldots, n$ *onderling onafhankelijk, dan is*

$$\bar{X} = \frac{1}{n}(X_1 + \ldots + X_n)$$

ook normaal verdeeld met gemiddelde μ *en standaardafwijking* $\frac{\sigma}{\sqrt{n}}$.

$$\mu_{\bar{X}} = \frac{1}{n} \cdot \mu_{X_1} + \frac{1}{n} \cdot \mu_{X_2} + \cdots + \frac{1}{n} \cdot \mu_{X_n} = \frac{1}{n} \cdot n \cdot \mu = \mu$$

$$\sigma_{\bar{X}} = \sqrt{\frac{1}{n^2} \cdot \sigma_{X_1}^2 + \frac{1}{n^2} \cdot \sigma_{X_2}^2 + \cdots + \frac{1}{n^2} \cdot \sigma_{X_n}^2} = \sqrt{\frac{1}{n^2} \cdot n \cdot \sigma^2} = \sqrt{\frac{\sigma^2}{n}}$$

LINEAIRE COMBINATIES: Als $X_1, X_2, \ldots, X_n$ onderling **onafhankelijke** normaal verdeelde toevalsvariabelen zijn met verwachting μ_i en standaardafwijking σ_i voor *i* = 1, 2, …, *n* dan is $Y = a_1X_1 + a_2X_2 + \cdots + a_nX_n + b$ ook normaal verdeeld met verwachting $\mu = a_1\mu_1 + a_2\mu_2 + \cdots + a_n\mu_n + b$ en standaardafwijking $\sigma = \sqrt{a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + \cdots + a_n^2\sigma_n^2}$. *(opmerking: hierbij mogen de constanten a_i en b negatief zijn !!)*

Lineaire combinaties: (steekproef)gemiddelde

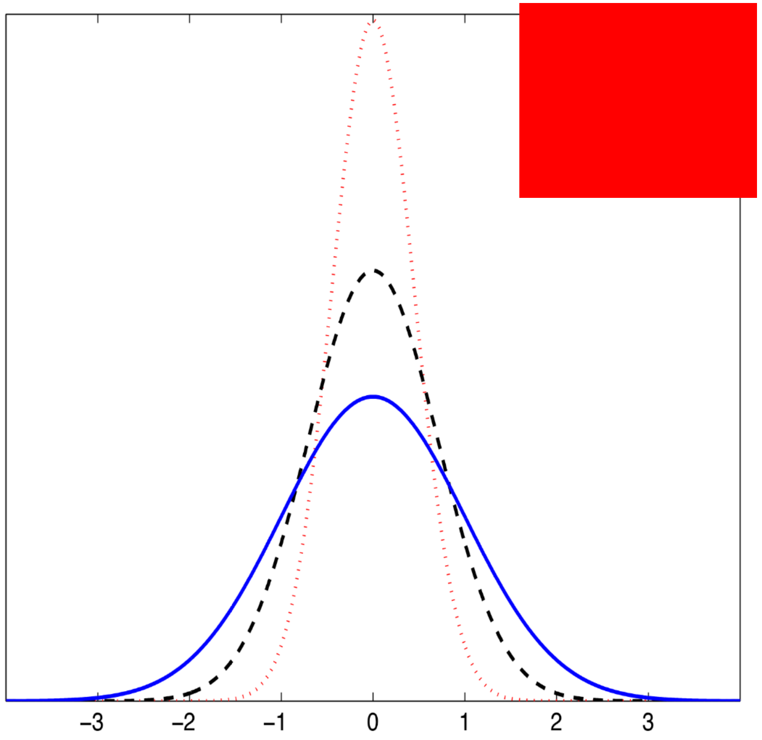
Een lineaire combinatie die veelvuldig voorkomt is het gemiddelde van n onderling onafhankelijke normaal verdeelde variabelen met dezelfde verwachting μ en standaardafwijking σ. Dan is het gemiddelde ook normaal verdeeld met verwachting μ en kleinere standaardafwijking $\sigma/\sqrt{n}$. Dus

Gevolg 6.9 *Als $X_i \sim N(\mu, \sigma)$, $i = 1, \ldots, n$ onderling onafhankelijk, dan is*

$$\bar{X} = \frac{1}{n}(X_1 + \ldots + X_n)$$

ook normaal verdeeld met gemiddelde μ en standaardafwijking $\dfrac{\sigma}{\sqrt{n}}$.

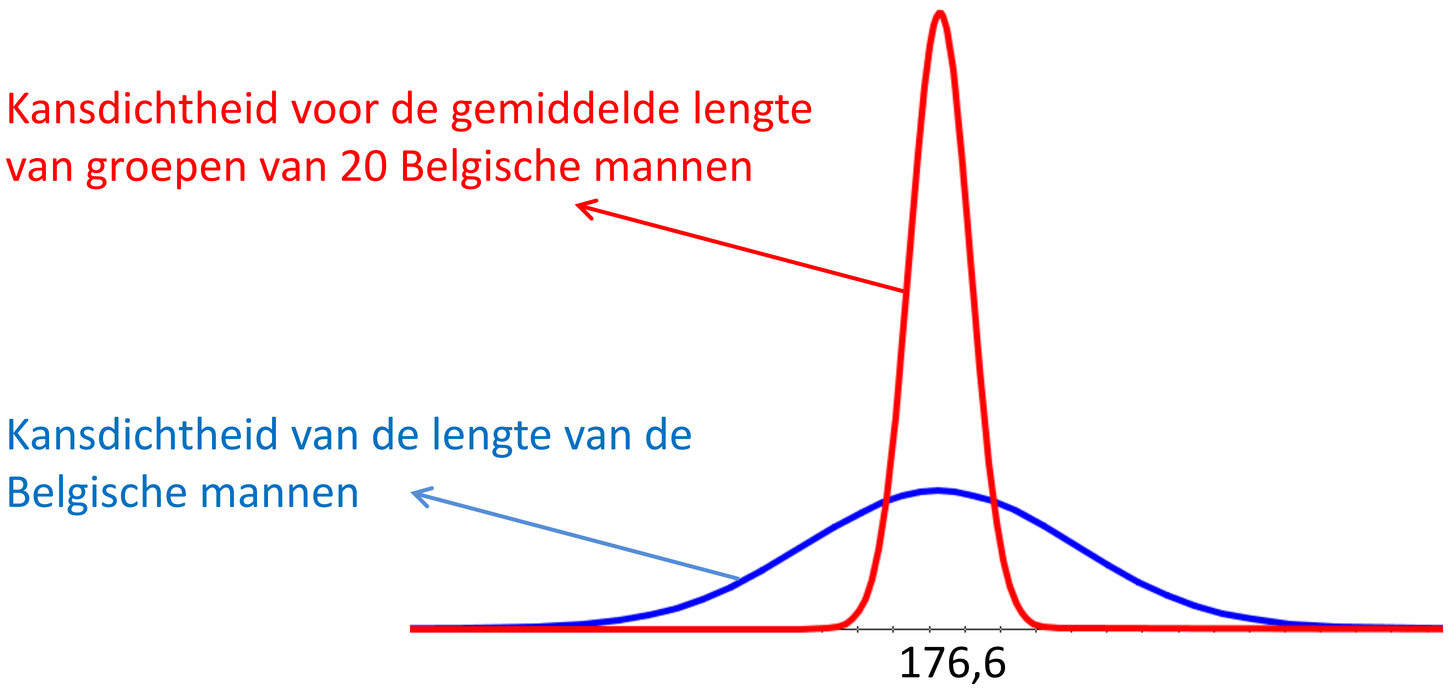
Hiernaast zie je de kansverdeling van 3 steekproeven.
Welke verdeling hoort bij de grootste steekproef?



Lineaire combinaties: (steekproef)gemiddelde

Voorbeeld: de gemiddelde lengte $\bar{X}$ van een groep van 20 Belgische mannen

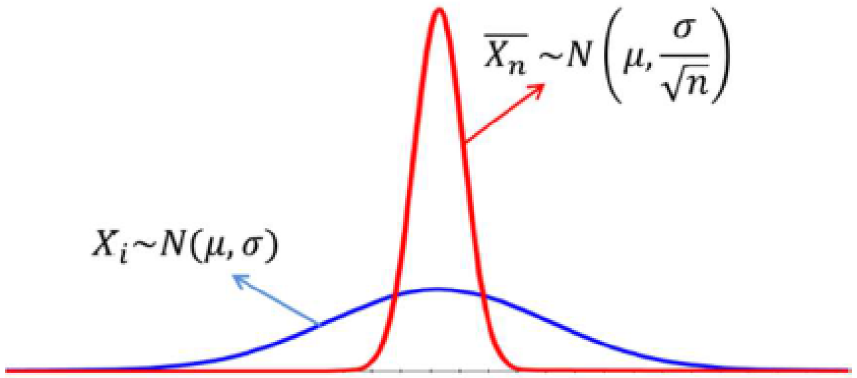
$$\bar{X} \sim N(176.6, \frac{7.6}{\sqrt{20}})$$



Formularium

LINEAIRE COMBINATIES: Als $X_1, X_2, \ldots, X_n$ onderling **onafhankelijke** normaal verdeelde toevalsvariabelen zijn met verwachting μ_i en standaardafwijking σ_i voor *i* = 1, 2, …, *n* dan is $Y = a_1X_1 + a_2X_2 + \cdots + a_nX_n + b$ ook normaal verdeeld met verwachting $\mu = a_1\mu_1 + a_2\mu_2 + \cdots + a_n\mu_n + b$ en standaardafwijking $\sigma = \sqrt{a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + \cdots + a_n^2\sigma_n^2}$. *(opmerking: hierbij mogen de constanten a_i en b negatief zijn !!)*

STEEKPROEFGEMIDDELDE: Als $X_1, X_2, \ldots, X_n$ onderling **onafhankelijke** normaal verdeelde toevalsvariabelen zijn met verwachting μ en standaardafwijking σ voor *i* = 1, 2, …, *n* dan is $\overline{X_n} \sim N(\mu, \frac{\sigma}{\sqrt{n}})$.



Benaderende verdelingen

Binomale verdeling:

→ Poisson benadering

→ Normale benadering

Poisson verdeling:

→ Normale benadering

Dit wordt nuttig als we hypothese-toetsen gaan uitvoeren

Poisson benadering van de binomiale verdeling

7.1.4 Poisson benadering van binomiale verdeling

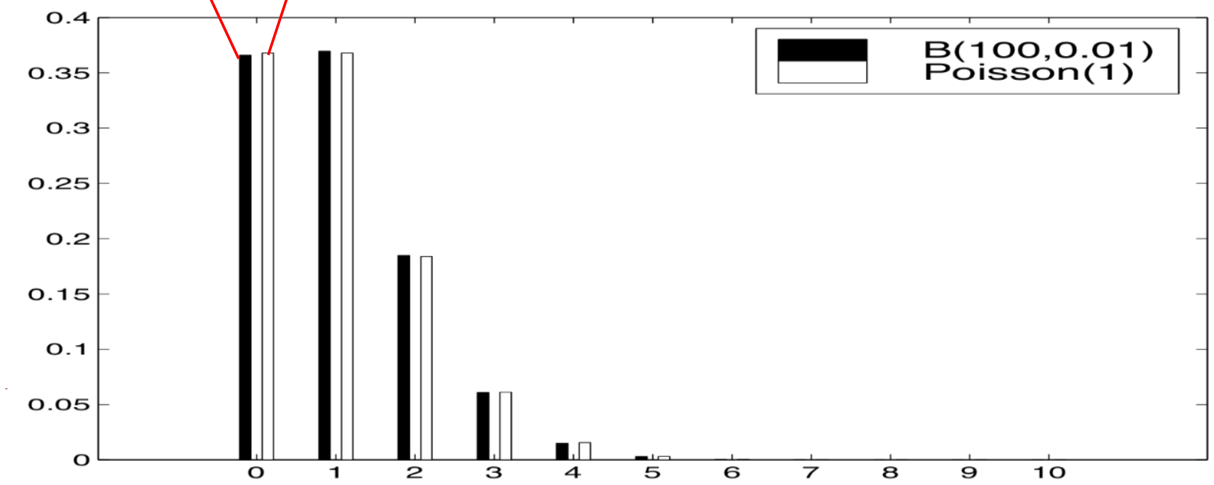
Stelling 7.4 *Als $X \sim B(n, p)$ met het aantal deelexperimenten n groot en de kans op succes p klein, dan is X bij benadering Poisson verdeeld met parameter $\lambda = np$.* → Gemiddelde $B(n,p)$

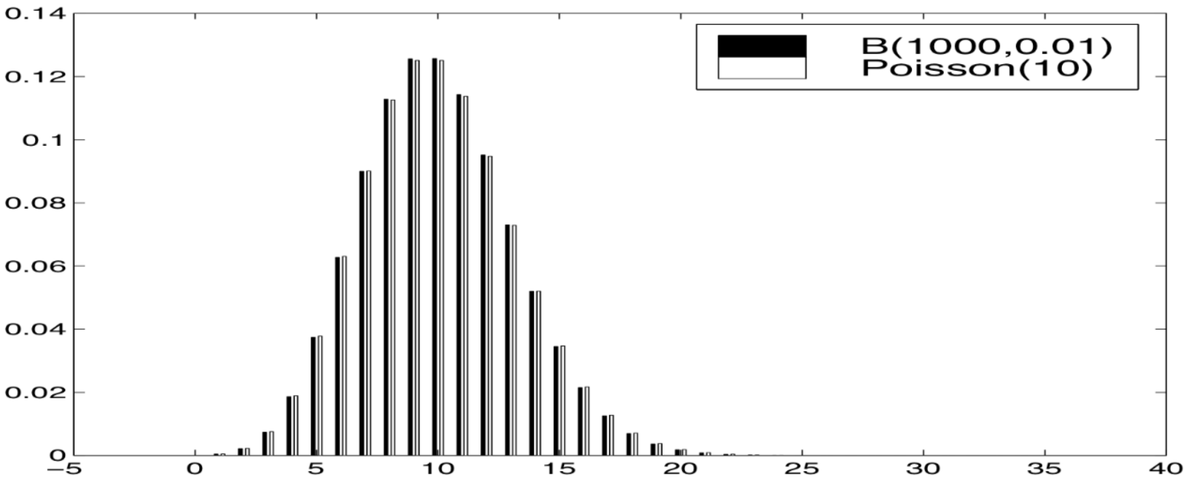
We zeggen dat n groot is en p klein als $n \geq 50$ en $p \leq 0.1$. Dan is de benadering voldoende goed. Zie figuur 7.1 voor een illustratie.

binomPdf(100,0.01,0)

poissPdf(1,0)

Figuur 7.1: Illustratie van de Poisson benadering van de binomiale verdeling.





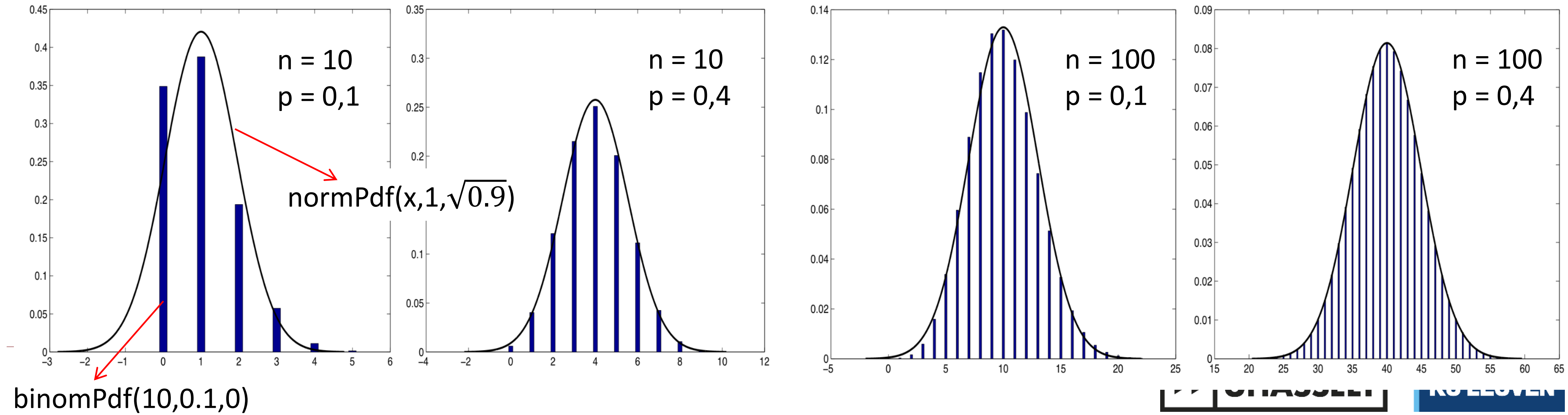
Normale benadering van de binomiale verdeling

Eigenschap 6.10 *Als $X \sim B(n,p)$ dan is X bij benadering*

$$X \sim N\left(np, \sqrt{np(1-p)}\right)$$

verdeeld voor n groot.

Voor vaste n is deze benadering het beste voor p dicht bij 0.5. In praktijk is het een goede benadering als $np \geq 5$ en $n(1-p) \geq 5$. Zie figuur 6.3 voor een illustratie.



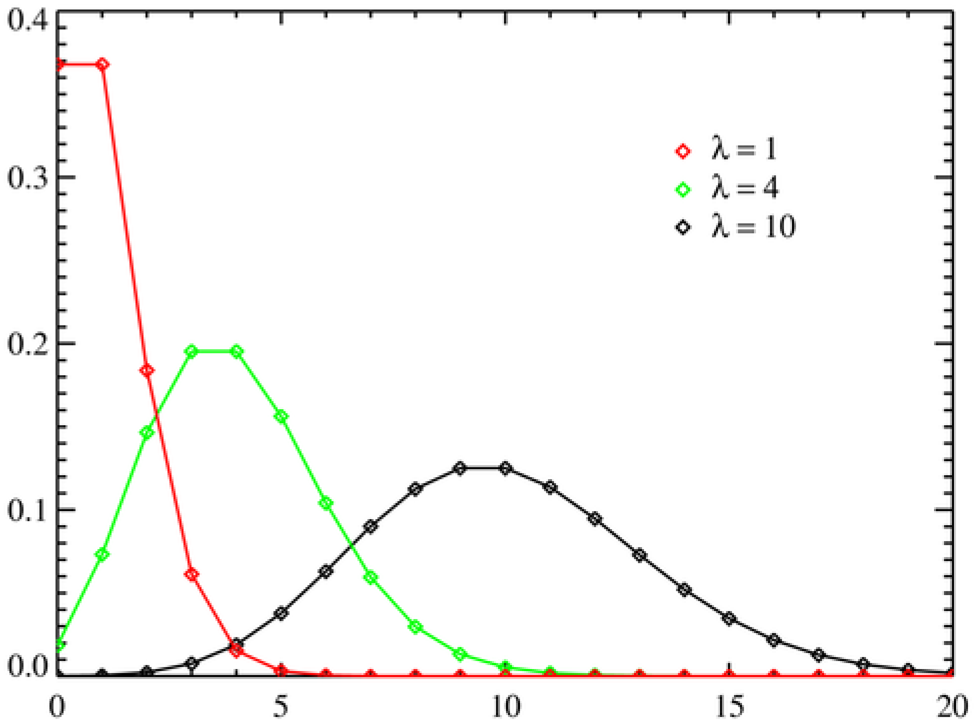
Normale benadering van de poisson verdeling

Eigenschap 7.5 *Als X Poisson verdeeld is met parameter λ en λ is groot, dan is X bij benadering normaal verdeeld:*

$$X \sim N\left(\lambda, \sqrt{\lambda}\right)$$

μ σ

In de praktijk is dit een goede benadering als $\lambda \geq 10$.



Formularium

NORMALE BENADERING: Als $X \sim \mathrm{B}(n, p)$, dan is X bij benadering $N(np, \sqrt{np(1-p)})$ als $np > 5$ en $n(1-p) > 5$.

Als $X \sim \mathrm{Poi}(\lambda)$, dan is X bij benadering $N(\lambda, \sqrt{\lambda})$ als $\lambda > 10$.

TODO vóór de werkzitting

Oefeningen 32 + 34 bekijken (oplossing staat op Toledo).

Dit duurt ongeveer 10 minuten per oefening.