11.0 Voorkennis

Let op:
Cumulatieve binomiale verdeling: $P(X \leq k)$ = binom**c**df(n,p,k)

Wanneer je met binomcdf werkt, werk je dus altijd met een kans van de vorm $P(X \leq k)$

Voorbeeld 1:
Binomiaal kansexperiment met n = 25 en p = 0,20 en X = aantal keer succes

$P(X \leq 7)$ = binomcdf(25, 0.20, 7) = 0,891

Voorbeeld 2:
$P(X \geq 8)$ = 1 – $P(X \leq 7)$ = 1 – binomcdf(25, 0.20, 7) = 1 – 0,891 = 0,109

Voorbeeld 3:
$P(X < 8)$ = $P(X \leq 7)$ = binomcdf(25, 0.20, 7) = 0,891

11.0 Voorkennis

Voorbeeld 4:
Gooien met muntstuk.
Bereken hoeveel keer je moet gooien zodat de kans op minstens drie keer munt groter dan 98% is.

X = aantal keer munt, p = 0,5, n = onbekend

Dus bij welke n geldt: $P(X \geq 3) = 1 - P(X \leq 2)$ is groter dan 0.98

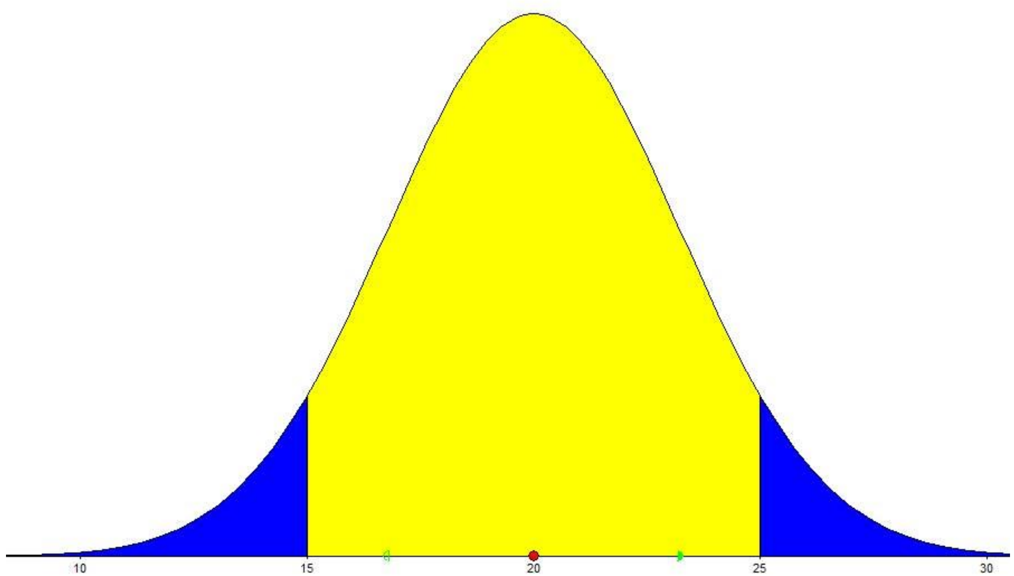
Invullen in de GR:
Y1 = 1 – binomcdf(X, 0.5, 2)
Y2 = 0.98

Aflezen uit TABLE geeft
$n = 11 \Rightarrow 1 - P(X \leq 2) \approx 0{,}967$
$n = 12 \Rightarrow 1 - P(X \leq 2) \approx 0{,}981$
Dus je moet minstens 12 keer gooien.

11.0 Voorkennis

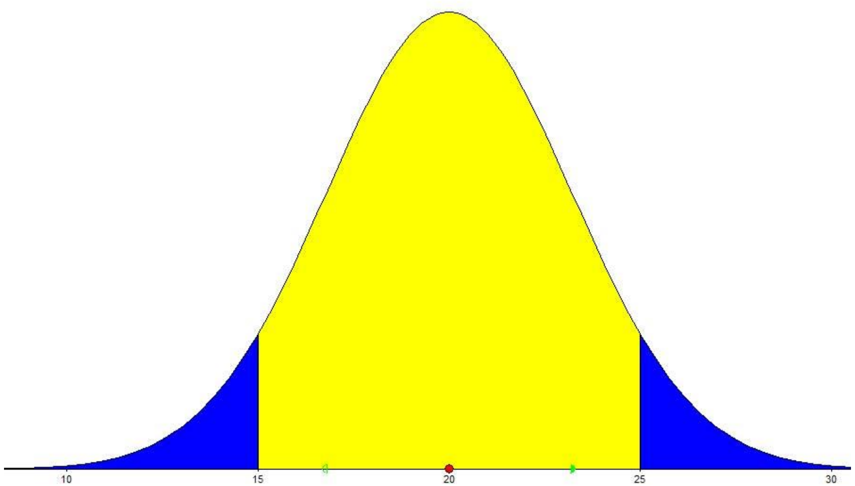
Wanneer van een normale verdeling, het gemiddelde (μ), de standaard-afwijking (σ) en een interval gegeven is, kun je de oppervlakte onder de normaalkromme voor dat interval berekenen.

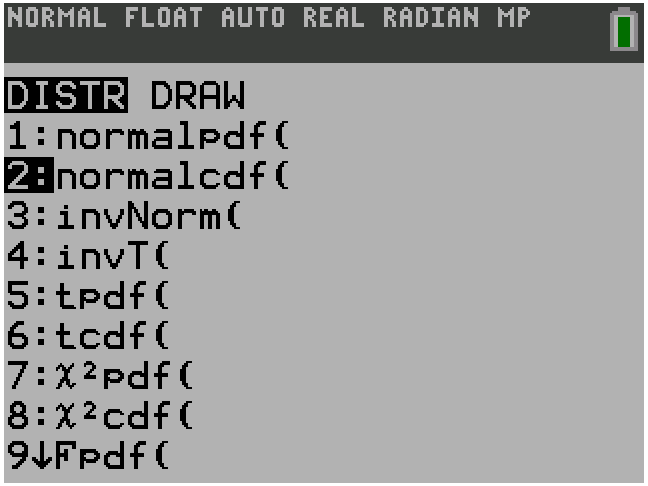
<u>Voorbeeld 5:</u> Normale verdeling met μ = 20 en σ = 3.2. Bepaal de oppervlakte onder de normaalkromme tussen 15 en 25.



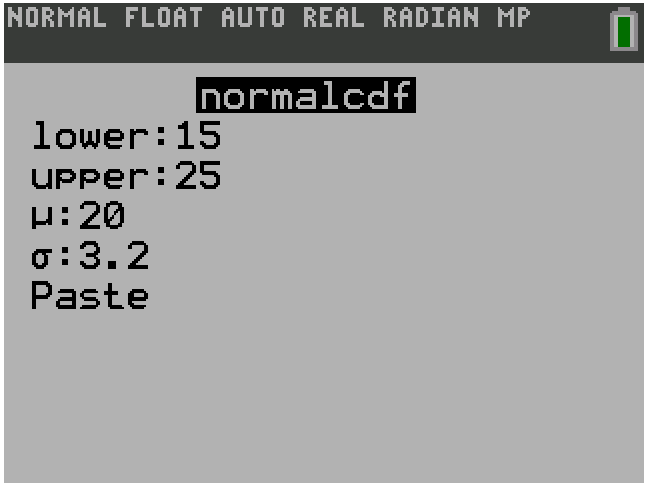
11.0 Voorkennis

<u>Voorbeeld 5:</u> Normale verdeling met μ = 20 en σ = 3.2. Bepaal de oppervlakte onder de normaalkromme tussen 15 en 25.





<u>Op de GR:</u> **2ND VARS | DISTR 2:normalcdf(| ENTER**

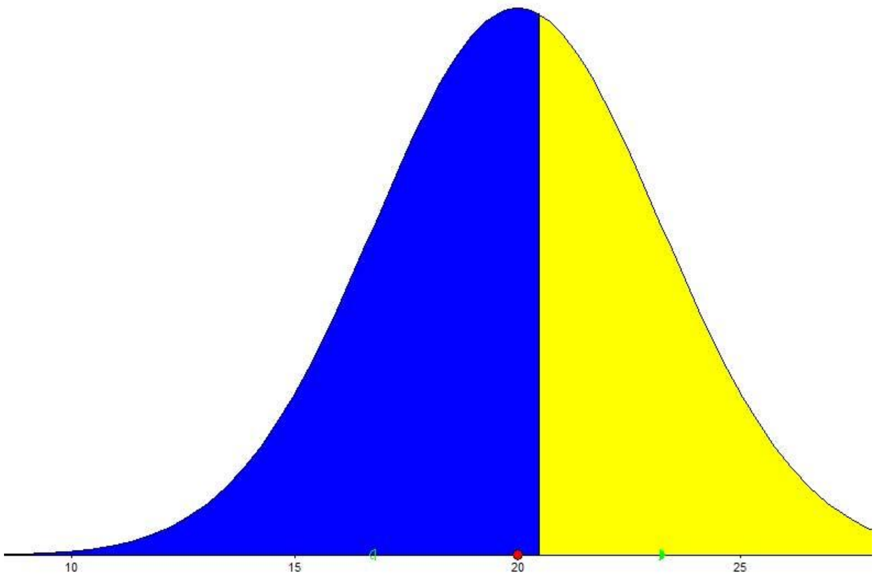


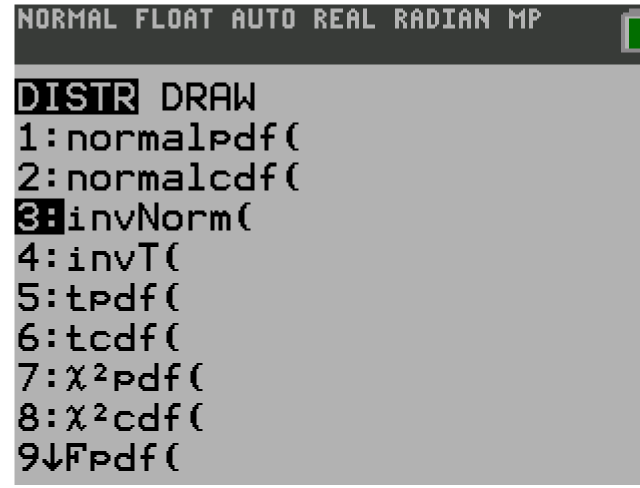
Vul bij “lower” de linkergrens in
Vul bij “upper” de rechtergrens in;
Vul bij “μ” het gemiddelde in;
Vul bij “σ” de standaardafwijking in.

Normalcdf(15, 25, 20, 3.2) ≈ 0,882

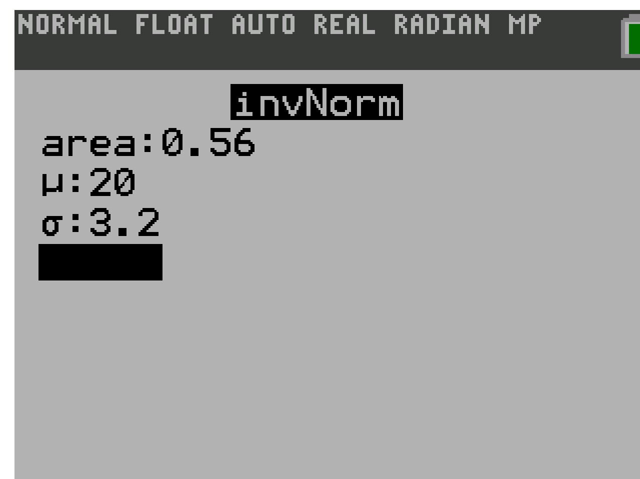
11.0 Voorkennis

Voorbeeld 6: Normale verdeling met $\mu = 20$ en $\sigma = 3.2$. De oppervlakte links van de grens a is 0,56. Bereken deze grens.





Op de GR: **2ND VARS | DISTR 3:invNorm(| ENTER**

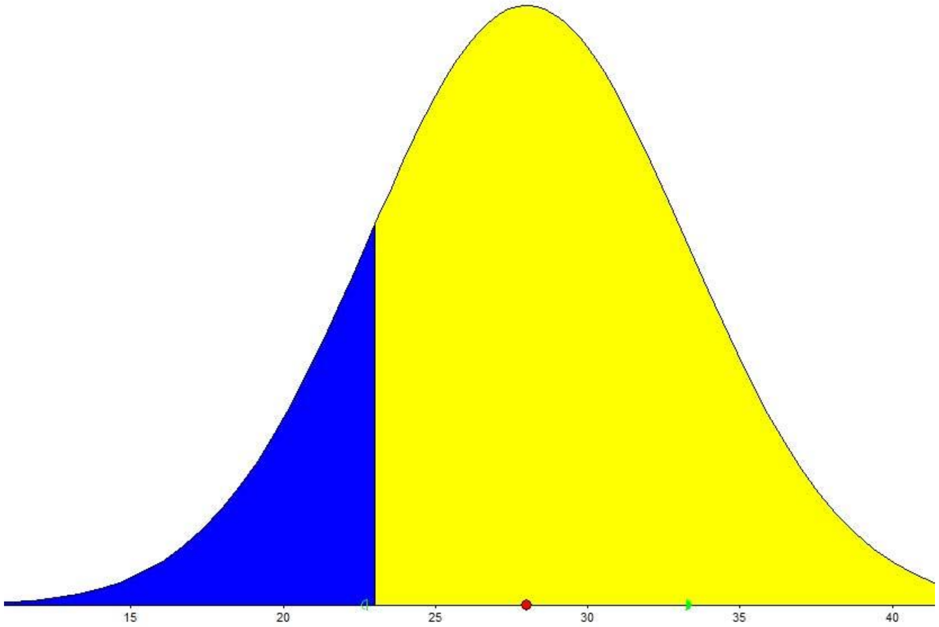


Vul bij “area” de oppervlakte links van de grens in;
Vul bij “μ” het gemiddelde in;
Vul bij “σ” de standaardafwijking in.

InvNorm(0.56, 20, 3.2) ≈ 20,48

11.0 Voorkennis

<u>Voorbeeld 7:</u> Normale verdeling met μ = 28 en σ = onbekend. De oppervlakte Rechts van 23 is 0,83. Bereken de standaardafwijking.



Er moet gelden normalcdf(23, 10^{99}, 28, σ) = 0,83

Met de GR: Y1 = normalcdf(23, 10^{99}, 28, σ)
Y2 = 0,83 en **INTERSECT**

[Let op grenzen van assen!!!]

11.1 De wortel-*n*-wet [1]

Voorbeeld 1: Een artikel wordt geproduceerd in twee fasen:
De productietijd X van fase I is normaal verdeeld met $\mu_x = 180$ en $\sigma_x = 2$
De productietijd Y van fase II is normaal verdeeld met $\mu_x = 23$ en $\sigma_y = 1$

Hoeveel procent van de artikelen heeft een totale productietijd Z ($X + Y$) van minder dan 200 seconden?

In dit voorbeeld zijn er twee **onafhankelijke** normaal verdeelde toevalsvariabelen waarvan je de som neemt. **Deze som *Z* is ook een normaal verdeelde toevalsvariabele** met:

$\mu_z = \mu_x + \mu_y$ en $\sigma_Z = \sqrt{\sigma_x^2 + \sigma_y^2}$

Wanneer je van twee onafhankelijke normaal verdeelde toevalsvariabelen het verschil neemt geldt:

$\mu_z = \mu_x - \mu_y$ en $\sigma_Z = \sqrt{\sigma_x^2 + \sigma_y^2}$

11.1 De wortel-*n*-wet [1]

<u>Voorbeeld 1:</u> Een artikel wordt geproduceerd in twee fasen:
De productietijd X van fase I is normaal verdeeld met $\mu_x = 180$ en $\sigma_x = 2$
De productietijd Y van fase II is normaal verdeeld met $\mu_y = 23$ en $\sigma_y = 1$

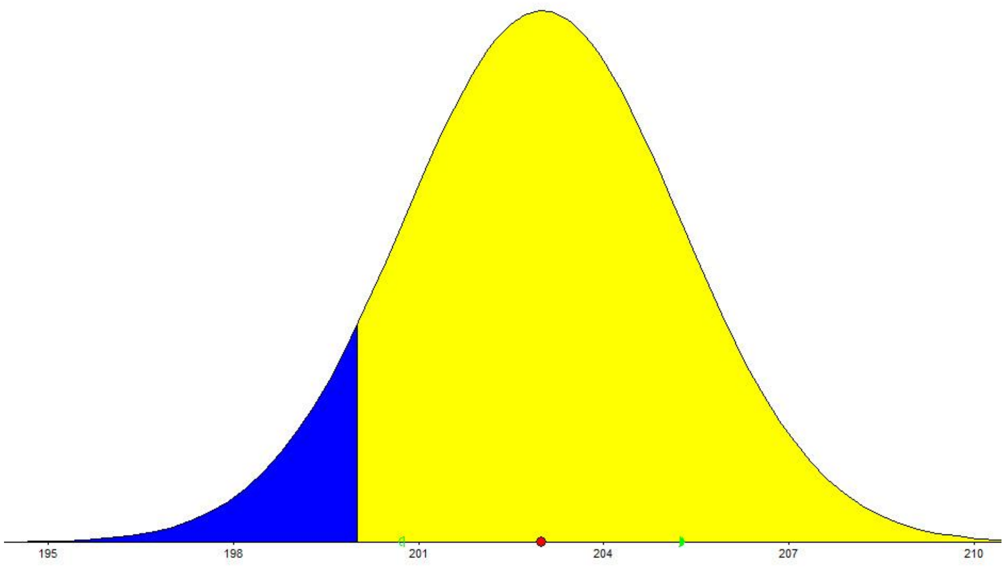
Hoeveel procent van de artikelen heeft een totale productietijd Z ($X + Y$) van minder dan 200 seconden?

Z is normaal verdeeld met gemiddelde μ_z en standaardafwijking σ_z:

$\mu_z = \mu_x + \mu_y = 180 + 23 = 203$ en $\sigma_Z = \sqrt{\sigma_x^2 + \sigma_y^2} = \sqrt{2^2 + 1^2} = \sqrt{5}$

11.1 De wortel-*n*-wet [1]

<u>Voorbeeld 1:</u> Hoeveel procent van de artikelen heeft een totale productietijd Z (X + Y) van minder dan 200 seconden?



Opp = normalcdf(-10^{99}, 200, 203, √5) ≈ 0,090

Dus 0,09 · 100% = 9,0% heeft een productietijd van minder dan 200 seconden.

11.1 De wortel-n-wet [2]

Voorbeeld 1: Een artikel wordt geproduceerd in drie fasen:
De productietijd X van fase I is normaal verdeeld met μ_x = 180 en σ_x = 2
De productietijd Y van fase II is normaal verdeeld met μ_y= 23 en σ_y = 1
De productietijd Z van fase III is normaal verdeeld met μ_z= 10 en σ_z = 0,5

Hoeveel procent van de artikelen heeft een totale productietijd T (X + Y + Z) van minder dan 210 seconden?

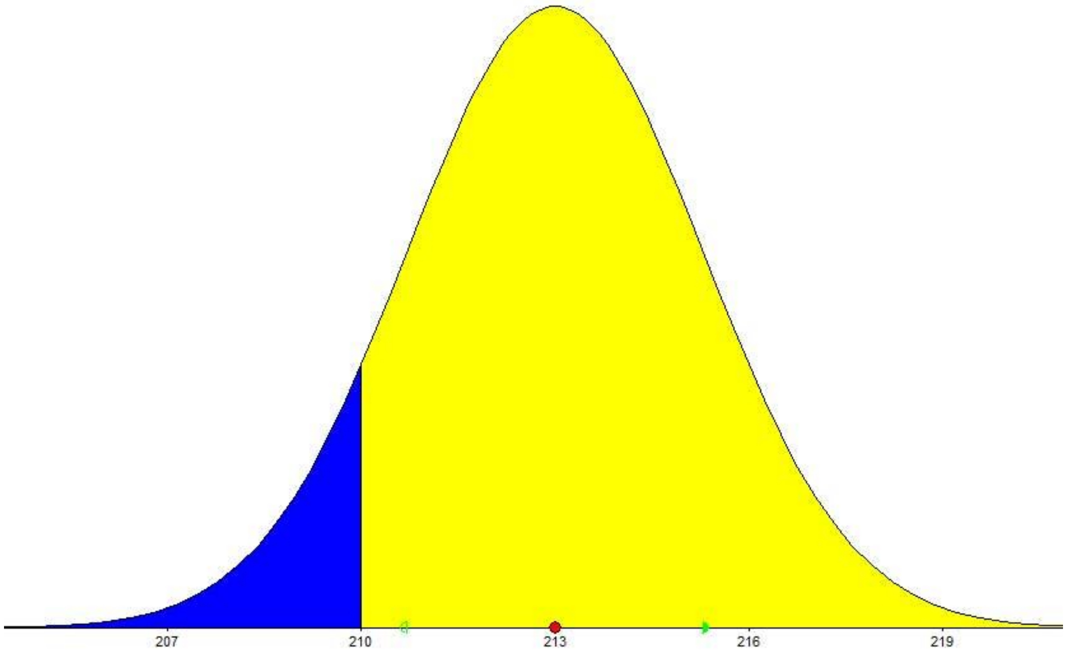
T is normaal verdeeld met gemiddelde μ_t en standaardafwijking σ_t:

$\mu_t = \mu_x + \mu_y + \mu_z = 180 + 23 + 10 = 213$ en

$$\sigma_t = \sqrt{\sigma_x^2 + \sigma_y^2 + \sigma_z^2} = \sqrt{2^2 + 1^2 + 0{,}5^2} = \sqrt{5{,}25}$$

11.1 De wortel-n-wet [2]

Voorbeeld 1: Hoeveel procent van de artikelen heeft een totale productietijd T $(X + Y + Z)$ van minder dan 210 seconden?



Opp = normalcdf(-10^{99}, 210, 213, $\sqrt{5{,}25}$) ≈ 0,095

Dus 0,095 · 100% = 9,5% heeft een productietijd van minder dan 210 seconden.

11.1 De wortel-*n*-wet [3]

<u>Voorbeeld 1:</u> Van een blik erwten uit een pallet is het gewicht X normaal verdeeld met $\mu_x = 500$ en $\sigma_x = 2$. Er wordt nu een steekproef van 10 blikken uit deze pallet genomen. Bereken de kans dat het gewicht van deze 10 blikken minder is dan 4985 gram.

Het totale gewicht van deze 10 blikken ($X_{\text{som}} = X + X + \dots + X$) is nu normaal verdeeld met:

$\mu_{X\text{som}} = \mu_x + \mu_x + \dots + \mu_x = 10 \cdot \mu_x = 10 \cdot 500 = 5000$

$$\sigma_{Xsom} = \sqrt{\sigma_x^2 + \sigma_x^2 + \dots + \sigma_x^2} = \sqrt{10 \cdot \sigma_x^2} = \sqrt{10 \cdot 2^2} = \sqrt{10} \cdot 2$$

Wanneer je een steekproef met een grootte van *n* neemt geldt:

De som ($X_{\text{som}} = X + X + \dots + X$) van deze steekproef is normaal verdeeld met:

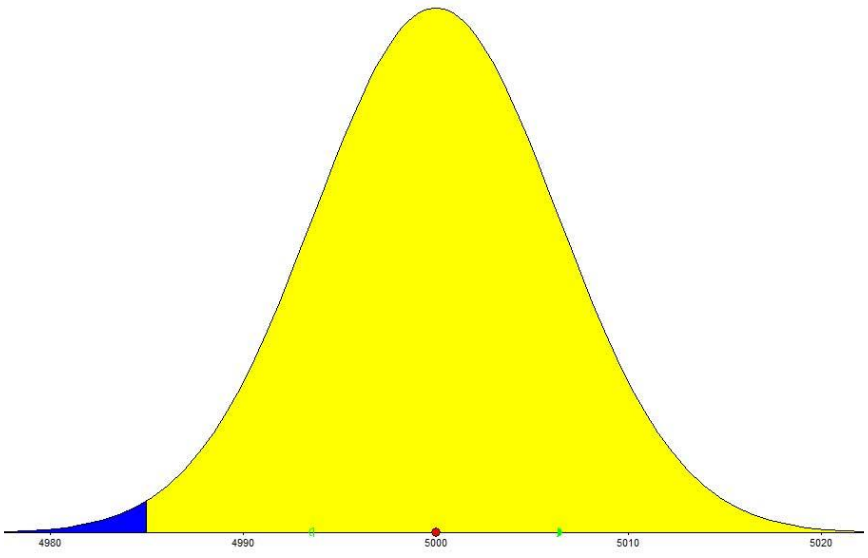
$\mu_{X\text{som}} = \mu_x + \mu_x + \dots + \mu_x = n \cdot \mu_x$ en

$$\sigma_{Xsom} = \sqrt{\sigma_x^2 + \sigma_x^2 + \dots + \sigma_x^2} = \sqrt{n \cdot \sigma_x^2} = \sqrt{n} \cdot \sigma_x$$

11.1 De wortel-*n*-wet [3]

Voorbeeld 2: Van een blik erwten uit een pallet is het gewicht X normaal verdeeld met $\mu_x = 500$ en $\sigma_x = 2$. Er wordt nu een steekproef van 10 blikken uit deze pallet genomen. Bereken de kans dat het gewicht van deze 10 blikken minder is dan 4985 gram.

$\mu_{X\text{som}} = 5000$ en $\sigma_{X\text{som}} = \sqrt{10} \cdot 2$



Opp = normalcdf(-10^{99}, 4985, 5000, $\sqrt{10} \cdot 2$) = 0,00886

11.1 De wortel-*n*-wet [4]

Voorbeeld 1: Van een blik erwten uit een pallet is het gewicht X normaal verdeeld met $\mu_x = 500$ en $\sigma_x = 2$. Er wordt nu een steekproef van 10 blikken uit deze pallet genomen.

Het totale gewicht van deze 10 blikken ($X_{\text{som}} = X + X + \ldots + X$) is nu normaal verdeeld met:

$\mu_{X\text{som}} = \mu_x + \mu_x + \ldots + \mu_x = 10 \cdot \mu_x = 10 \cdot 500 = 5000$

$$\sigma_{Xsom} = \sqrt{\sigma_x^2 + \sigma_x^2 + \ldots + \sigma_x^2} = \sqrt{10 \cdot \sigma_x^2} = \sqrt{10 \cdot 2^2} = \sqrt{10} \cdot 2$$

Het gemiddelde gewicht van deze 10 blikken ($\bar{X}$ = **steekproefgemiddelde**) is ook normaal verdeeld met:

$$\mu_{\bar{X}} = \frac{\mu_{X-som}}{10} = \frac{10 \cdot \mu_X}{10} = \mu_X$$

$$\sigma_{\bar{X}} = \frac{\sigma_{X-som}}{10} = \frac{\sqrt{10} \cdot \sigma_X}{10} = \frac{\sigma_X}{\sqrt{10}}$$

Algemeen:

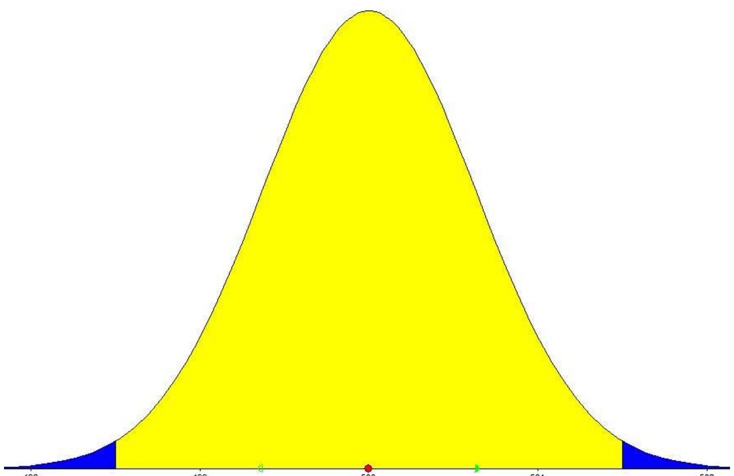
Bij een steekproef van grootte n geldt: $\bar{X}$ normaal verdeeld met $\mu_{\bar{X}} = \mu_X$ en $\sigma_{\bar{X}} = \frac{\sigma_X}{\sqrt{n}}$

11.1 De wortel-*n*-wet [4]

<u>Voorbeeld 1:</u> Van een blik erwten uit een pallet is het gewicht X normaal verdeeld met μ_x = 500 en σ_x = 2. Er wordt nu een steekproef van 10 blikken uit deze pallet genomen.

Bereken de kans dat het steekproefgemiddelde ($\bar{X}$) minder dan 1.5 van μ_x afwijkt

$\bar{X}$ is normaal verdeeld met $\mu_{\bar{X}} = \mu_X = 500 \; en \; \sigma_{\bar{X}} = \frac{\sigma_X}{\sqrt{n}} = \frac{2}{\sqrt{10}}$



P(498.5 < $\bar{X}$ < 501.5) = normalcdf(498.5, 501.5, 500, 2/√10) ≈ 0,9821

11.2 Discrete en continue verdelingen [1]

Voorbeeld:

Een machine vult pakken koffie met een gemiddelde van 510 gram en een standaardafwijking van 5 gram.

Bereken de kans dat in een steekproef van 20 pakken minstens drie pakken minder dan 505 gram bevatten.

X = aantal pakken koffie dat minder dan 505 gram bevat

$n = 20$

p = kans dat één pak koffie minder dan 505 gram bevat

p = normalcdf(-10^99, 505, 510, 5) = 0.159…

$P(X \geq 3) = 1 - P(X \leq 2) = 1 -$ binomcdf(20, 0.159…, 2) ≈ 0.637

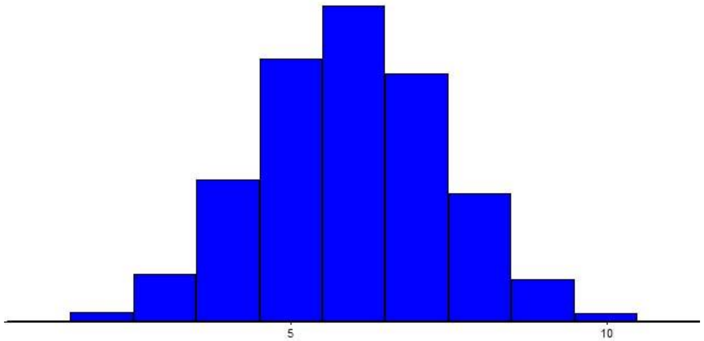
11.2 Discrete en continue verdelingen [2]

Continue toevalsvariabele *Y*:

- Alle waarden zijn mogelijk
- Kansverdeling is een vloeiende kromme;
- Bv.: Lengte van mannen, Gewicht van vrouwen, alles wat normaal verdeeld is;
- $P(Y < 5) = P(Y \leq 5)$.

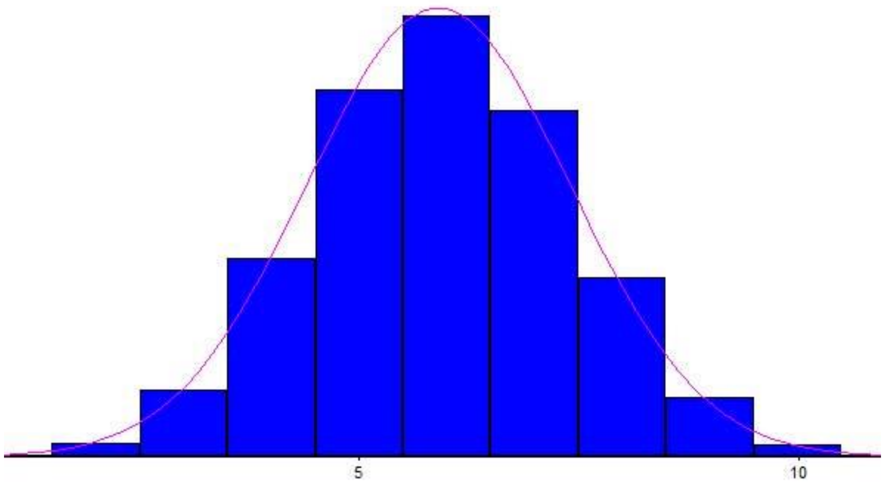
Discrete toevalsvariabele *X*:

- Alleen een aantal losse waarden zijn mogelijk;
- Kansverdeling is een histogram;
- Bv.: Aantal auto's op een weg per minuut, De schoenmaat van volwassenen;



- $P(X < 5) = P(X \leq 4)$.

11.2 Discrete en continue verdelingen [2]



Als we een discrete toevalsvariabele X benaderen door een continue toevalsvariabele Y geldt:

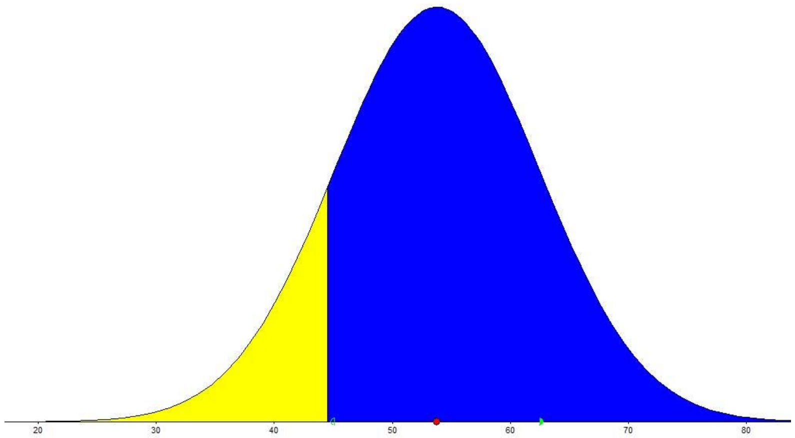
$P(X < 5) = P(X \leq 4) = P(Y \leq 4{,}5)$

Algemeen:
$P(X \leq k) = P(Y \leq k + 0{,}5)$

11.2 Discrete en continue verdelingen [3]

Voorbeeld:
Het aantal auto's X per uur op een weg is te benaderen door een normaal verdeelde toevalsvariabele Y met $\mu_Y = 53{,}8$ en $\sigma_Y = 8{,}7$.

Gedurende een uur wordt het aantal auto's op de weg geteld.
Bereken in hoeveel procent van de gevallen er minder dan 45 auto's per uur worden geteld.



$P(X < 45) = P(X \leq 44) = P(Y \leq 44{,}5) =$
normalcdf(-10^{99}, 44.5, 53.8, 8.7) $\approx 0{,}143$

11.3 Beslissen op grond van een steekproef [1]

Voorbeeld 1:
Een vulmachine vult flessen met een inhoud van X ml.
X is normaal verdeeld met $\mu_x = 400$ en $\sigma_x = 4$

Er wordt een steekproef genomen van 40 flessen.
Uit de $\sqrt{n}$-wet volgt nu: $\mu_{\bar{x}} = \mu_x = 400$ en $\sigma_{\bar{x}} = \dfrac{\sigma_x}{\sqrt{40}} = \dfrac{4}{\sqrt{40}}$

- De vraag is nu of de machine opnieuw ingesteld moet worden als bv. gemeten wordt dat de gemiddelde inhoud van de flessen 399 of 398 ml is;
- Vanwege meetafwijkingen wil dit nog niet meteen zeggen dat de machine ook niet meer goed bijvult;
- Ten onrechte bijvullen leidt tot het onnodig stil leggen van de productie;
- Niet tijdig bijvullen leidt tot een slechtere kwaliteit van het product.

11.3 Beslissen op grond van een steekproef [1]

Voorbeeld 1:
Een vulmachine vult flessen met een inhoud van X ml.
X is normaal verdeeld met $\mu_x = 400$ en $\sigma_x = 4$. De steekproefgrootte is 40.

Uit de √n-wet volgt nu: $\mu_{\bar{x}} = \mu_x = 400$ en $\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{40}} = \frac{4}{\sqrt{40}}$

Er wordt afgesproken dat de machine wordt bijgesteld als: $\bar{X} \leq 398{,}5 \textit{ of } \bar{X} \geq 401{,}5$

Situatie 1: De machine wordt ten onrechte bijgesteld:
- Als de machine ten onrechte wordt bijgesteld betekent dit dat voor het steekproefresultaat geldt: $\bar{X} \leq 398{,}5 \textit{ of } \bar{X} \geq 401{,}5$
- Als de machine ten onrechte wordt bijgesteld geldt nog steeds: $\mu = 400$

P(ten onrechte bijstellen) = P($\bar{X} \leq 398{,}5 \textit{ of } \bar{X} \geq 401{,}5$) =

1 – Normalcdf(398.5, 401.5, 400, $\frac{4}{\sqrt{40}}$) ≈ 0,018

De kans dat ten onrechte wordt bijgesteld is dus 0,018

11.3 Beslissen op grond van een steekproef [1]

Voorbeeld 1:
Een vulmachine vult flessen met een inhoud van X ml.
X is normaal verdeeld met μ_x = 400 en σ_x = 4. De steekproefgrootte is 40.

Uit de √n-wet volgt nu: $\mu_{\bar{x}} = \mu_x = 400$ en $\sigma_{\bar{x}} = \dfrac{\sigma_x}{\sqrt{40}} = \dfrac{4}{\sqrt{40}}$

Er wordt afgesproken dat de machine wordt bijgesteld als: $\bar{X} \leq 398{,}5 \; of \; \bar{X} \geq 401{,}5$

Situatie 2: Machine werkt met gemiddelde van 398 en wordt niet bijgesteld:
- Het steekproefresultaat ligt dus tussen de 398,5 en de 401,5:
- Als gemiddelde wordt nu 398 gebruikt.

P(niet bijstellen) = P($398{,}5 < \bar{X} < 401{,}5$) =

1 – Normalcdf(398.5, 401.5, 398, $\dfrac{4}{\sqrt{40}}$) ≈ 0,215

De kans dat dus ten onrechte niet wordt bijgesteld is 0,215

11.3 Beslissen op grond van een steekproef [2]

Voorbeeld 1:

Een vulmachine vult flessen met een inhoud van X ml.
X is normaal verdeeld met $\mu_x = 400$ en $\sigma_x = 4$. De steekproefgrootte is 25.

Uit de √n-wet volgt nu: $\mu_{\bar{x}} = \mu_x = 400$ en $\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{25}} = \frac{4}{\sqrt{25}} = 0{,}8$

Er worden twee veronderstellingen geformuleerd:

1. De inhoud van een fles is normaal verdeeld met $\mu = 400$.
   Dit is de **nulhypothese H_0**;
2. De inhoud van een fles is normaal verdeeld met $\mu \neq 400$.
   Dit is de **alternatieve hypothese H_1**.

Bij het beslissingsvoorschrift "verwerp H_0 als $\bar{X} \leq 399$ *of* $\bar{X} \geq 401$ geldt:

P(onterecht verwerpen) = P($\bar{X} \leq 399$ *of* $\bar{X} \geq 401$) =

1 – Normalcdf(399, 401, 400, 0.8) ≈ 0,211

De kans dat de nulhypothese ten onrechte verworpen wordt is 0,211.

11.3 Beslissen op grond van een steekproef [2]

Afhankelijk van het beslissingsvoorschrift kan het dus voorkomen dat de nulhypothese ten onrechte verworpen wordt

- Ten onrechte bijvullen leidt tot het onnodig stil leggen van de productie;
- Niet tijdig bijvullen leidt tot een slechtere kwaliteit van het product.

In de praktijk wordt daarom een maximale kans gekozen dat de nulhypothese ten onrechte verworpen wordt:

Significantieniveau (α) = de maximale kans dat de nulhypothese ten onrechte verworpen wordt.

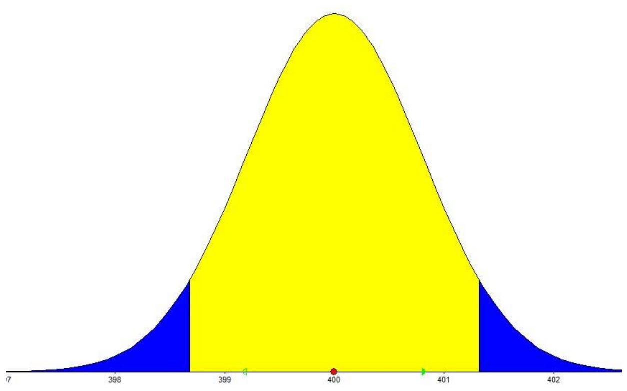
11.3 Beslissen op grond van een steekproef [2]

Voorbeeld 2:

Een vulmachine vult flessen met een inhoud van X ml.
X is normaal verdeeld met $\mu_x = 400$ en $\sigma_x = 4$. De steekproefgrootte is 25.

Uit de $\sqrt{n}$-wet volgt nu: $\mu_{\bar{x}} = \mu_x = 400$ en $\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{25}} = \frac{4}{\sqrt{25}} = 0{,}8$

- Het significantieniveau (α) wordt bepaald op 0,10.
- Het beslissingsvoorschrift wordt: "verwerp H_0 als $\bar{X} \leq g_l$ *of* $\bar{X} \geq g_r$ ".
- Hieruit volgt: P(verwerpen H_0) = P($\bar{X} \leq g_l$ *of* $\bar{X} \geq g_r$) = 0,10



11.3 Beslissen op grond van een steekproef [2]

Voorbeeld 2:

Een vulmachine vult flessen met een inhoud van X ml.
X is normaal verdeeld met μ_x = 400 en σ_x = 4. De steekproefgrootte is 25.

Uit de $\sqrt{n}$-wet volgt nu: $\mu_{\bar{x}} = \mu_x = 400$ en $\sigma_{\bar{x}} = \dfrac{\sigma_x}{\sqrt{25}} = \dfrac{4}{\sqrt{25}} = 0{,}8$

- Het significantieniveau (α) wordt bepaald op 0,10.
- Het beslissingsvoorschrift wordt: "verwerp H_0 als $\bar{X} \leq g_l$ *of* $\bar{X} \geq g_r$ ".
- Hieruit volgt: P(verwerpen H_0) = P($\bar{X} \leq g_l$ *of* $\bar{X} \geq g_r$) = 0,10
- Vanwege de symmetrie van de normale verdeling geldt:
  - $P(\bar{X} \leq g_l)$ = 0,05 => g_l = invNorm(0.05, 400, 0.8) ≈ 398,68
  - $P(\bar{X} \geq g_r)$ = 0,05 => g_r = invNorm(1-0.05, 400, 0.8) ≈ 401,32

Wanneer er nu een steekproef wordt uitgevoerd zal de nulhypothese niet verworpen worden als het steekproefgemiddelde ligt tussen 398,68 en 401,32.

11.3 Beslissen op grond van een steekproef [3]

Voorbeeld 1:
X = aantal kilometers dat autoband meegaat.
X is normaal verdeeld met $\mu_x = 35.000$ en $\sigma_x = 4.000$.
H_0: $\mu = 35.000$, H_1: $\mu \neq 35.000$ en $\alpha = 0{,}05$.
Dit beslissingsvoorschrift betekent dat de 2,5% kleinste waarnemingen en de 2,5% grootste waarnemingen leiden tot een verwerping van de nulhypothese.

Een steekproef met een grootte van 64 geeft een gemiddelde van 33.844 kilometer.

Moet de nulhypothese nu verworpen worden?

$$P(\overline{X} \leq 33.844) = normalcdf(-10^{99}, 33.844, 35.000, \frac{4000}{\sqrt{64}}) \approx 0{,}01$$

33.844 behoort bij de 2,5% kleinste waarnemingen, dus het steekproefresultaat wijkt significant af van 35.000. De Nulhypothese wordt verworpen.

De **overschrijdingskans** is nu 0,01. Merk op dat de overschrijdingskans kleiner is dan de helft van het significantieniveau.

11.3 Beslissen op grond van een steekproef [3]

Voorbeeld 2:

X = aantal kilometers dat autoband meegaat.

X is normaal verdeeld met μ_x = 35.000 en σ_x = 4.000.

H_0: μ = 35.000, H_1: $\mu \neq$ 35.000 en α = 0,05.

Dit beslissingsvoorschrift betekent dat de 2,5% kleinste waarnemingen en de 2,5% grootste waarnemingen leiden tot een verwerping van de nulhypothese.

Een steekproef met een grootte van 64 geeft een gemiddelde van 34.831 kilometer.

Moet de nulhypothese nu verworpen worden?

$$P(\bar{X} \leq 34.831) = normalcdf(-10^{99}, 34.831, 35.000, \frac{4.000}{\sqrt{64}}) \approx 0,368$$

34.831 behoort niet bij de 2,5% kleinste waarnemingen en ook niet bij de 2,5% grootste waarnemingen. Het steekproefresultaat wijkt dus niet significant af van 35.000. De nulhypothese wordt niet verworpen.

De **overschrijdingskans** is nu 0,368. Merk op dat de overschrijdingskans groter is dan de helft van het significantieniveau.

11.3 Beslissen op grond van een steekproef [3]

Voorbeeld 3:
X = aantal kilometers dat autoband meegaat.
X is normaal verdeeld met μ_x = 35.000 en σ_x = 4.000.
H_0: μ = 35.000, H_1: $\mu \neq$ 35.000 en α = 0,05.
Dit beslissingsvoorschrift betekent dat de 2,5% kleinste waarnemingen en de 2,5% grootste waarnemingen leiden tot een verwerping van de nulhypothese.

Een steekproef met een grootte van 64 geeft een gemiddelde van 35.682 kilometer.

Moet de nulhypothese nu verworpen worden?

$$P(\overline{X} \geq 35.682) = normalcdf(35.682, 10^{99}, 35.000, \frac{4.000}{\sqrt{64}}) \approx 0{,}086$$

35.682 behoort niet bij de 2,5% grootste waarnemingen. Het steekproefresultaat wijkt dus niet significant af van 35.000. De nulhypothese wordt dus niet verworpen.

De **overschrijdingskans** is nu 0,086. Merk op dat de overschrijdingskans groter is dan de helft van het significantieniveau.

11.3 Beslissen op grond van een steekproef [3]

Algemeen:
X = aantal kilometers dat autoband meegaat.
X is normaal verdeeld met $\mu_x = \mu_0$ en $\sigma_x = \sigma$.
H_0: $\mu = \mu_0$, H_1: $\mu \neq \mu_0$ en significantieniveau α.

Een steekproef met een grootte van n geeft een gemiddelde van k kilometer.

Voor de overschrijdingskans van k geldt nu:

$P(\bar{X} \leq k)$ als $k < \mu_0$. (Het steekproefresultaat is hier kleiner dan μ_0)
$P(\bar{X} \geq k)$ als $k > \mu_0$. (Het steekproefresultaat is hier groter dan μ_0)

Als de overschrijdingskans van k kleiner dan of gelijk aan $0{,}5\alpha$ is, dan wordt de nulhypothese verworpen.

11.4 Eenzijdig en tweezijdig toetsen [1]

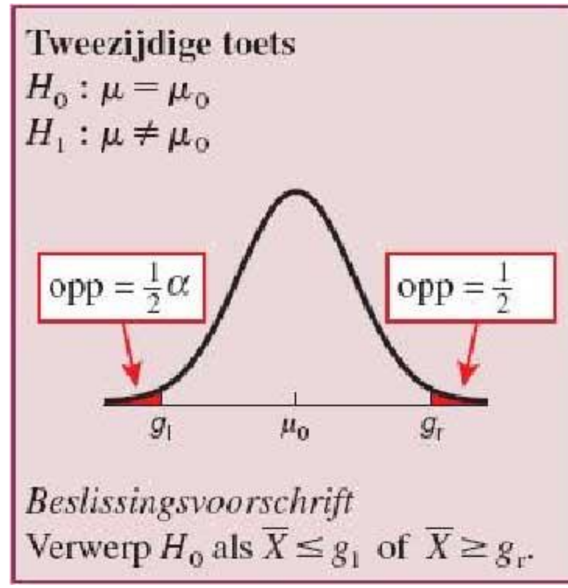
Tweezijdige toets:

X = aantal kilometers dat autoband meegaat.
X is normaal verdeeld met $\mu_x = \mu_0$ en $\sigma_x = \sigma$.
H_0: $\mu = \mu_0$, H_1: $\mu \neq \mu_0$ en significantieniveau α.

Een steekproef met een grootte van n geeft een gemiddelde van k kilometer.

Voor de overschrijdingskans van k geldt nu:

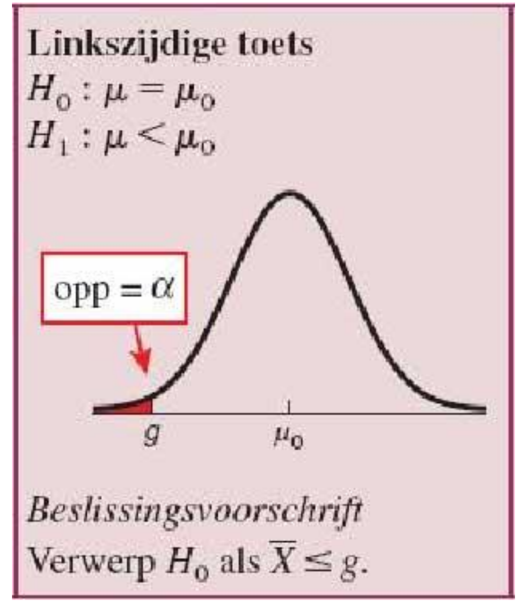
$P(\bar{X} \leq k)$ als $k < \mu_0$.
(Het steekproefresultaat is hier kleiner dan μ_0)
$P(\bar{X} \geq k)$ als $k > \mu_0$.
(Het steekproefresultaat is hier groter dan μ_0)



Als de overschrijdingskans van k [$P(\bar{X} \leq k)$ of $P(\bar{X} \geq k)$] kleiner dan of gelijk aan $0{,}5\alpha$ is, dan wordt de nulhypothese verworpen.

11.4 Eenzijdig en tweezijdig toetsen [1]

Linkszijdige toets:
X = aantal kilometers dat autoband meegaat.
X is normaal verdeeld met $\mu_x = \mu_0$ en $\sigma_x = \sigma$.
H_0: $\mu = \mu_0$, H_1: $\mu < \mu_0$ en significantieniveau α.



Een steekproef met een grootte van n geeft een gemiddelde van k kilometer.

Voor de overschrijdingskans van k geldt nu:

$P(\overline{X} \leq k)$ als $k < \mu_0$.
(Het steekproefresultaat is hier kleiner dan μ_0)

Als de overschrijdingskans van k $[P(\overline{X} \leq k)]$ kleiner dan of gelijk aan α is, dan wordt de nulhypothese verworpen.

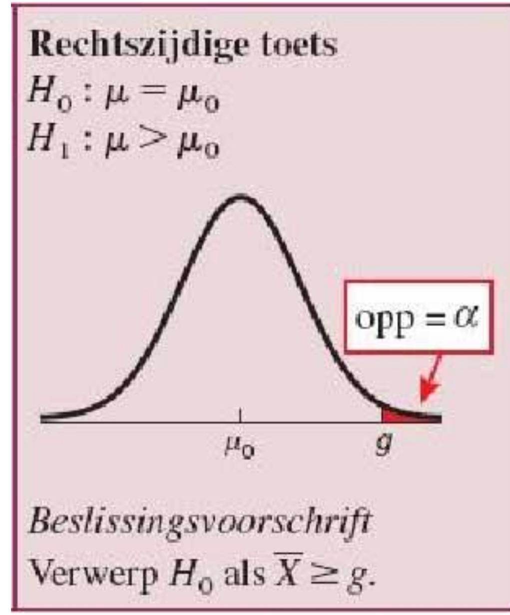
11.4 Eenzijdig en tweezijdig toetsen [1]

Rechtszijdige toets:
X = aantal kilometers dat autoband meegaat.
X is normaal verdeeld met $\mu_x = \mu_0$ en $\sigma_x = \sigma$.
H_0: $\mu = \mu_0$, H_1: $\mu > \mu_0$ en significantieniveau α.

Een steekproef met een grootte van n geeft een gemiddelde van k kilometer.

Voor de overschrijdingskans van k geldt nu:

$P(\bar{X} \geq k)$ als $k > \mu_0$.
(Het steekproefresultaat is hier kleiner dan μ_0)



Als de overschrijdingskans van k [$P(\bar{X} \geq k)$] kleiner dan of gelijk aan α is, dan wordt de nulhypothese verworpen.

11.4 Eenzijdig en tweezijdig toetsen [1]

Voorbeeld:
X = levensduur in uren van een nieuw soort batterij.
X is normaal verdeeld met $\mu_x = 800$ en $\sigma_x = 40$. Neem $\alpha = 0{,}025$
Een consumentenorganisatie beweert dat de gemiddelde levensduur minder dan 800 uur is.

Een aselecte steekproef van 100 batterijen geeft een levensduur van 793,8 uur.

H_0: $\mu = 800$, H_1: $\mu < 800$ en $\alpha = 0{,}025$.

$P(\bar{X} \leq 793{,}8)$ = normalcdf(-10^{99}, 793.8, 800, 40/$\sqrt{100}$) $\approx 0{,}061 > \alpha$

De nulhypothese wordt niet verworpen, dus de bewering van de consumenten-organisatie klopt niet en de bewering van de fabrikant over de levensduur hoeft niet in twijfel getrokken te worden.

11.4 Eenzijdig en tweezijdig toetsen [2]

Let bij het opstellen van hypotheses op de volgende zaken:

- De nulhypothese is de hypothese die als uitgangspunt dient en die door middel van de alternatieve hypothese in twijfel wordt getrokken;
- De nulhypothese dient een **enkelvoudige hypothese** van de vorm H_0: $\mu = \mu_0$ te zijn.

11.5 Binomiale toetsen [1]

Voorbeeld:

X = aantal personen dat frisdrank A het lekkerst vindt:
p = kans dat iemand frisdrank A het lekkerst vindt (p = 0.4)
n = aantal keer dat aan één persoon gevraagd wordt welke frisdrank hij het lekkerst vindt (n = 100)
α = significantieniveau (α = 0.05)

Een bedrijf dat frisdrank B produceert zegt dat minder dan 40% van de mensen frisdrank A het lekkerst vindt. Klopt deze uitspraak?

De nulhypothese wordt nu: $H_0 : p = 0.4$
De alternatieve hypothese wordt nu: $H_1 : p < 0.4$

De nulhypothese wordt verworpen als het aantal personen in de steekproef dat frisdrank A het lekkerste vindt, klein is. Dus **Verwerp $\boldsymbol{H_0}$ als $\boldsymbol{X \leq g}$**.

We hebben nu een binomiale toets opgesteld want X is een binomiaal verdeelde toevalsvariabele.

11.5 Binomiale toetsen [1]

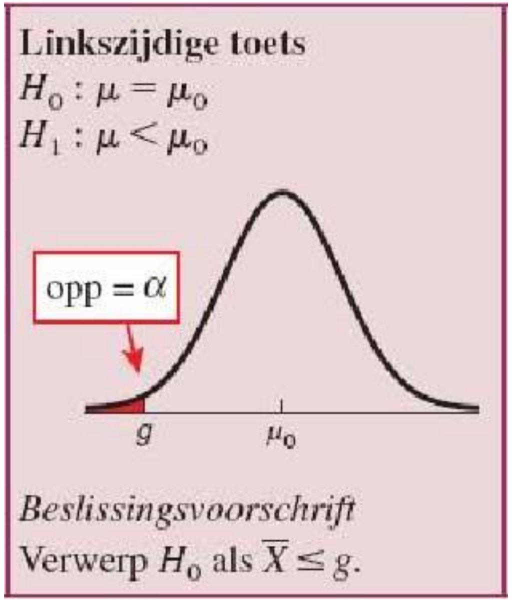
Voorbeeld:
X = aantal personen dat frisdrank A het lekkerst vindt:
p = kans dat iemand frisdrank A het lekkerst vindt ($p = 0.4$)
n = aantal keer dat aan één persoon gevraagd wordt welke frisdrank hij het lekkerst vindt ($n = 100$)
α = significantieniveau ($\alpha = 0.05$)
$H_0 : p = 0.4$ versus $H_1 : p < 0.4$

Bereken de overschrijdingskans van 28

De overschrijdingskans van 28 is $P(X \leq 28)$ = binomcdf(100, 0.4, 28) ≈ 0.0084
$P(X \leq 28) \leq \alpha$, dus de nulhypothese wordt verworpen

Als uit een steekproef blijkt dat maar 28 personen frisdrank A het lekkerst vinden klopt de nulhypothese dus niet en heeft het bedrijf dat frisdrank B produceert gelijk.

11.5 Binomiale toetsen [1]



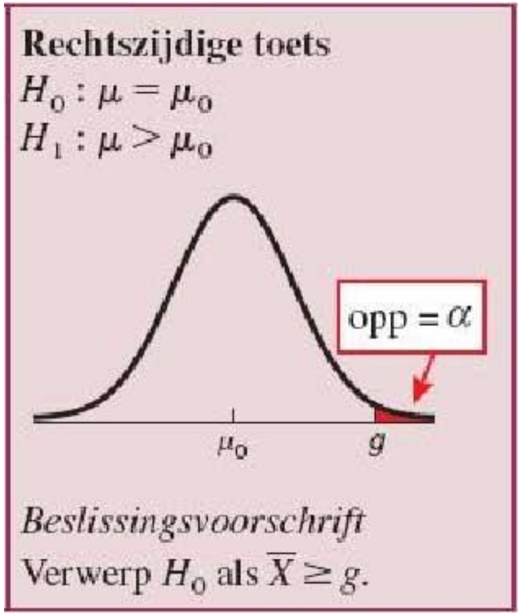
Normaal verdeelde toevalsvariabele (Linkszijdige toets):

- H_0: $\mu = \mu_0$, H_1: $\mu < \mu_0$ en significantieniveau α.
- Voor de overschrijdingskans van k geldt nu: $P(\overline{X} \leq k)$
- Als de overschrijdingskans van k $[P(\overline{X} \leq k)]$ kleiner dan of gelijk aan α is, dan wordt de nulhypothese verworpen.

Binomiaal verdeelde toevalsvariabele (Linkszijdige toets):

- H_0: $p = p_0$ en H_1 : $p < p_0$ en significantieniveau α.
- Voor de overschrijdingskans van k geldt nu: is $P(X \leq k)$
- Als de overschrijdingskans van k $[P(X \leq k)]$ kleiner dan of gelijk aan α is, dan wordt de nulhypothese verworpen.

11.5 Binomiale toetsen [1]



Normaal verdeelde toevalsvariabele (Rechtszijdige toets):

- H_0: $\mu = \mu_0$, H_1: $\mu > \mu_0$ en significantieniveau α.
- Voor de overschrijdingskans van k geldt nu: $P(\overline{X} \geq k)$
- Als de overschrijdingskans van k [$P(\overline{X} \geq k)$] kleiner dan of gelijk aan α is, dan wordt de nulhypothese verworpen.

Binomiaal verdeelde toevalsvariabele (Rechtszijdige toets):

- H_0: $p = p_0$ en H_1 : $p > p_0$ en significantieniveau α.
- Voor de overschrijdingskans van k geldt nu: is $P(X \leq k)$
- Als de overschrijdingskans van k [$P(X \leq k)$] kleiner dan of gelijk aan α is, dan wordt de nulhypothese verworpen.

11.5 Binomiale toetsen [2]

Voorbeeld:
X is binomiaal verdeeld met n = 100, α = 0,05 en
$H_0 : p = 0{,}42$ versus $H_1 : p > 0{,}42$

Algemeen:
- Het beslissingsvoorschrift wordt: "verwerp H_0 als $X \geq g$";
- Hieruit volgt: P(verwerpen H_0) = $P(X \geq g) \leq \alpha$ (= 0,05);
- We moeten dus nu de waarde van g vinden waarvoor dit geldt.

$P(X \geq g)$ = 1 - [P($X \leq g - 1$) ≤ 0,05

1 – P($X \leq g - 1$) ≤ α
1 – binomcdf(100, 0.42, g-1) ≤ 0.05

Vul in Y1 = 1 – binomcdf(100, 0.42, X-1)
Maak een tabel en lees af:
Voor X = 50 krijg je Y1 ≈ 0,065
Voor X = 51 krijg je Y1 ≈ 0,043
Dus verwerp H_0 als X ≥ 51

11.5 Binomiale toetsen [3]

<u>Voorbeeld:</u>
X = aantal huishoudens dat breedbeeld tv bezit
p = kans dat iemand een breedbeeld tv bezit (p = 0.38)
n = aantal keer dat aan één huishouden gevraagd wordt of het een breedbeeld tv bezit (n = 120)
α = significantieniveau (α = 0.05)

Wanneer klopt de bewering dat 38% van de huishoudens een breedbeeld tv bezit ?

De nulhypothese wordt nu: $H_0 : p = 0.38$
De alternatieve hypothese wordt nu: $H_1 : p \neq 0.38$

De nulhypothese wordt verworpen als het aantal personen dat een breedbeeld-tv bezit te veel afwijkt van 38%. Dus: **Verwerp $\boldsymbol{H_0}$ als $\boldsymbol{X \leq g_l}$ of $\boldsymbol{X \geq g_r}$.**

$P(\text{verwerpen } H_0) = P(X \geq g_r) \leq 0{,}5\alpha$ (= 0,025);
$P(\text{verwerpen } H_0) = P(X \leq g_l) \leq 0{,}5\alpha$ (= 0,025);

11.5 Binomiale toetsen [3]

Voorbeeld:

X = aantal huishoudens dat breedbeeld tv bezit

p = kans dat iemand een breedbeeld tv bezit (p = 0.38)

n = aantal keer dat aan één huishouden gevraagd wordt of het een breedbeeld tv bezit (n = 120)

α = significantieniveau (α = 0.05)

$H_0 : p = 0.38$ versus $H_1 : p \neq 0.38$

$\boldsymbol{P(\text{verwerpen } H_0) = P(X \geq g_r) \leq 0{,}5\alpha \ (= 0{,}025);}$

$P(X \geq g_r) = 1 - P(X \leq g_r - 1) \leq 0.025$ (= $0{,}5\alpha$)

1 – binomcdf(120, 0.38, g_r -1) ≤ 0.025

g_r = 56 geeft 0,032 en g_r = 57 geeft 0,021, dus $X \geq 57$

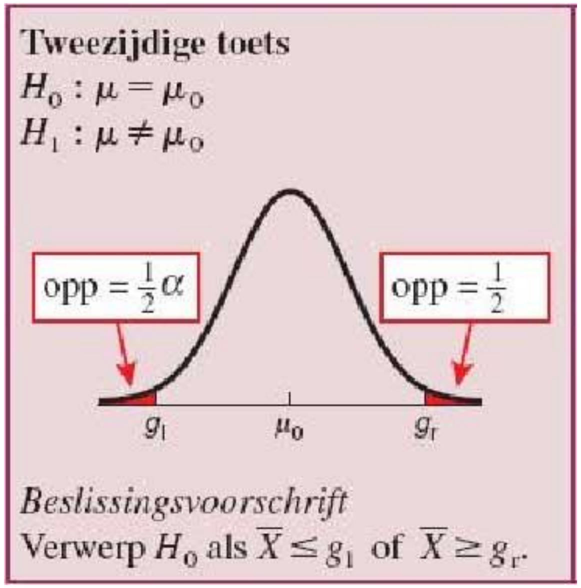
$\boldsymbol{P(\text{verwerpen } H_0) = P(X \leq g_l) \leq 0{,}5\alpha \ (= 0{,}025);}$

$P(X \leq g_l) \leq 0.025$ (= $0{,}5\alpha$)

binomcdf(120, 0.38, g_l) ≤ 0.025

g_l = 34 geeft 0,017 en g_l = 35 geeft 0,027, dus $X \leq 34$.

11.5 Binomiale toetsen [3]



Normaal verdeelde toevalsvariabele (Tweezijdige toets):

- H_0: $\mu = \mu_0$, H_1: $\mu \neq \mu_0$ en significantieniveau α.
- Voor de overschrijdingskans van k geldt nu:
  $P(\overline{X} \leq k)$ als $k < \mu_0$.
  $P(\overline{X} \geq k)$ als $k > \mu_0$.
- Als de overschrijdingskans van k [$P(\overline{X} \leq k)$ of $P(\overline{X} \geq k)$] kleiner dan of gelijk aan 0,5α is, dan wordt de nulhypothese verworpen.

Binomiaal verdeelde toevalsvariabele (Tweezijdige toets):

- H_0: $p = p_0$ en H_1 : $p \neq p_0$ en significantieniveau α.
- Als de overschrijdingskans van k [$P(X \leq k)$ of $P(X \geq k)$] kleiner dan of gelijk aan 0,5α is, dan wordt de nulhypothese verworpen.
- Je kunt de linker- en rechtergrens nu bepalen met:
  P(verwerpen H_0) = $P(X \geq g_r) \leq 0{,}5\alpha$;
  P(verwerpen H_0) = $P(X \leq g_l) \leq 0{,}5\alpha$.