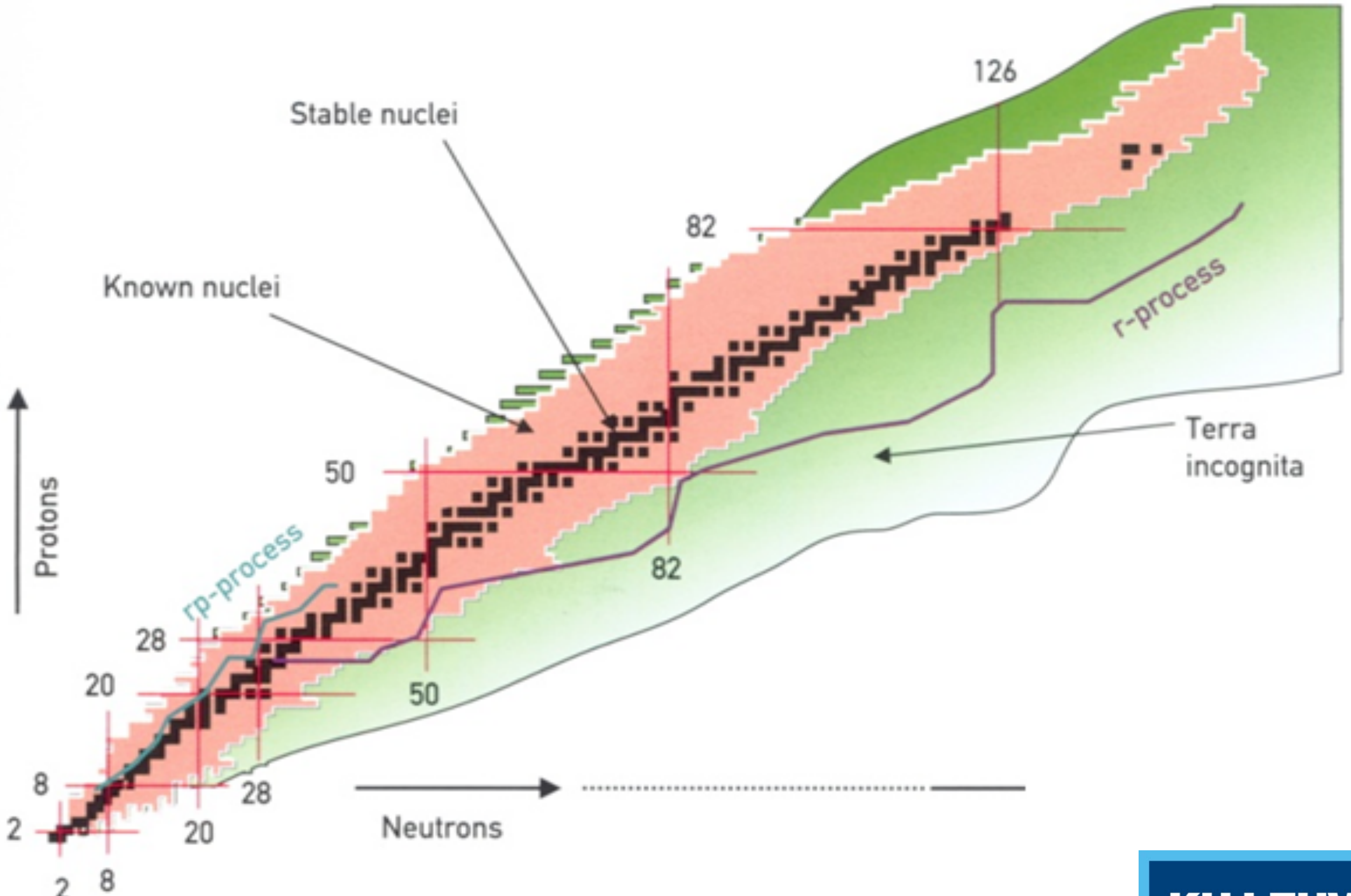
Meetfouten - Statistiek

- Meetfouten en hun schatting
- Testen van hypotheses
- Fitten

Future disciplinary self

Experimentele kernfysica

Studie van kernen ver van stabiliteit

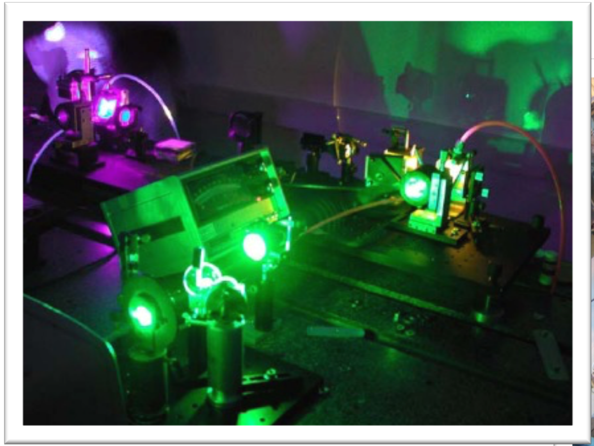


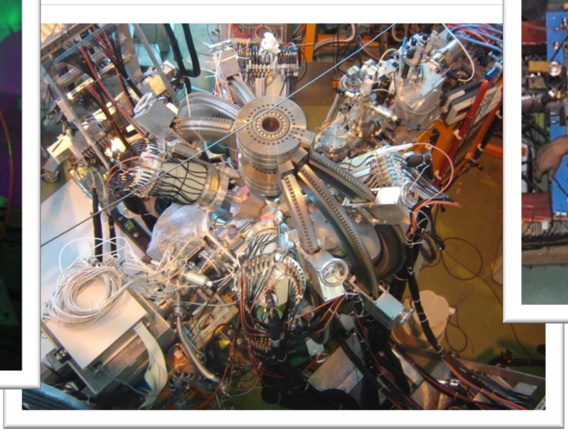
Future disciplinary self

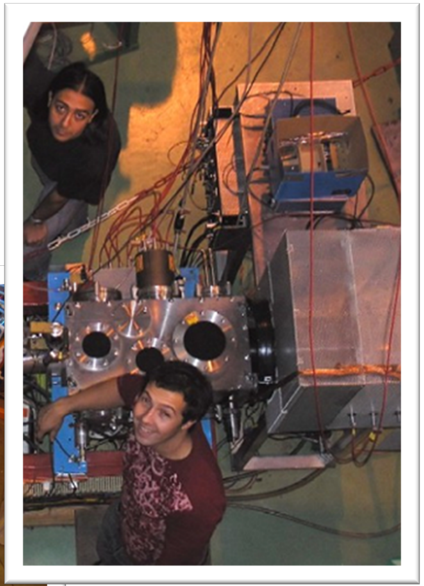
Experimentele kernfysica

Studie van kernen ver van stabiliteit

- Werken met schroevendraaiers
  en computers
- Technieken ontwikkelen
  Experimenten bedenken en realiseren
- Data analyseren en interpreteren







KU LEUVEN

Meetfouten - Statistiek

- **Meetfouten en hun schatting**
- Testen van hypotheses
- Fitten

Waarom foutentheorie

Exacte (=kwantitatieve) wetenschappen

Er kan een fout gegeven op de voorspellingen en resultaten

$$\text{echte waarde} = m \pm \varepsilon$$

Belang van foutenschatting

- Het meetresultaat is de echte waarde niet:
  hoe groot is het verschil?
  Hoe betrouwbaar is onze meting?
- Vergelijken van twee of meer meetresultaten:
  zijn ze compatibel?
- Verder gebruik van het resultaat:
  welke nauwkeurigheid is daarvoor nodig?
- Vergelijken met theorie:
  de gewenste fout bepaalt de soort experiment

Waarom foutentheorie

Exacte (=kwantitatieve) wetenschappen

Er kan een fout gegeven op de voorspellingen en resultaten

$$\text{echte waarde} = m \pm \varepsilon$$

Ons doel

- De fout ε bepalen
  met de betrouwbaarheidsniveaus (CL)
  CL = kans dat het interval de echte waarde dekt

  → belang van statistiek

Oorsprong onzekerheden

Grootheid

Constant of variabel

Meetproces

Statistische fouten

Systematische fouten

Meetinstrument

Resolutie (sensitiviteit)
Nauwkeurigheid

Voorbeelden

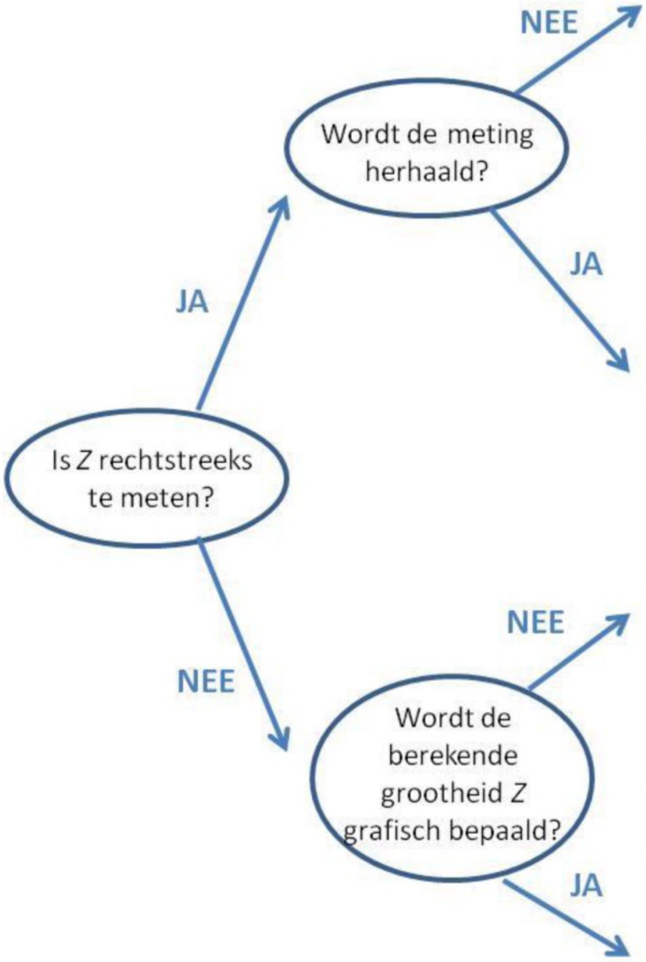
- Lading elektron, constante van Planck…
- Snelheid gasdeeltjes, aantal voorvallen…

- **Resolutie**: de kleinste nog te onderscheiden grootte van de meetwaarde
- **Nauwkeurigheid**: hoe dicht de meetwaarden de echte waarde benaderen

Meestal worden de twee opgesomd

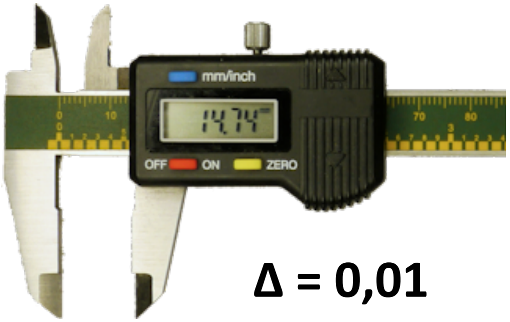
KU LEUVEN

Overzicht


Is Z rechtstreeks te meten?
JA
Wordt de meting herhaald?
NEE
JA
NEE
Wordt de berekende grootheid Z grafisch bepaald?
NEE
JA

Fouten van het instrument

- **Resolutie Δ**: de kleinste nog te onderscheiden grootte van de meetwaarde
  → **aflezing**
  - **digitaal**: de laatste digit
    Kansverdeling: uniform, CL 100%
  - **analoog**: de kleinste verdeling
    Kansverdeling: uniform of driehoek



Δ = 0,01

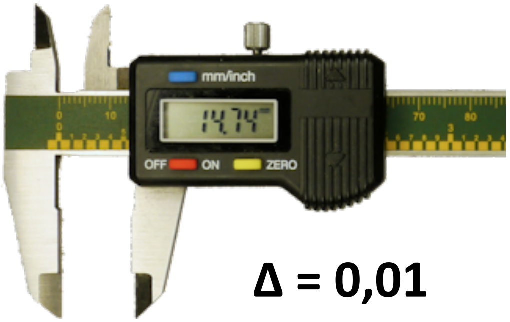
Δ = 0,2

- **Nauwkeurigheid δ**: hoe dicht de meetwaarden de echte waarde benaderen
  Wordt door de producent gegeven
  Goed instrument: δ < Δ → aflezing bepaalt de fout

Fouten van het instrument

- **Meestal:**
  resolutie en nauwkeurigheid opsommen
  in een gegeven "systematische fout"
  van het instrument (ook $\mathbf{\Delta_{sys}}$)
  Kan niet worden verwijderd

- **Voorbeeld**: We meten een spanning
  met een digitale multimeter:
  30,4 mV → Δ = 0,1 mV
  De nauwkeurigheid is 1% → δ = 0,3 mV
  → Δ_{sys} = 0,4 mV



Δ = 0,01

Δ = 0,2

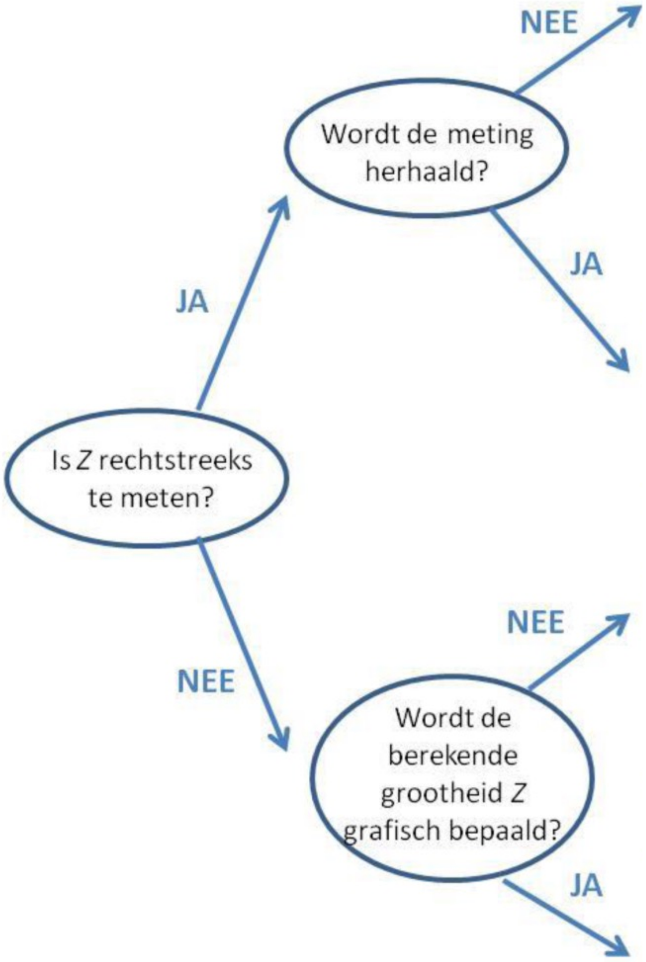
Het meetresultaat wordt gegeven als:

$$m \pm \Delta_{sys}/2$$

Kansverdeling: uniform

CL 100%

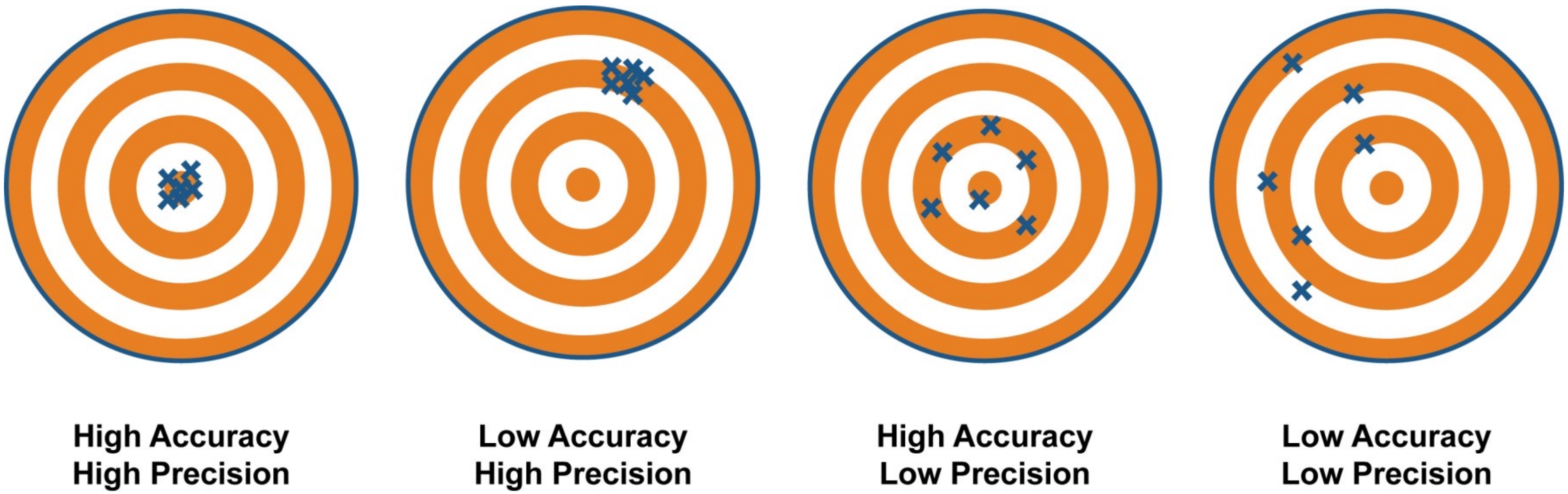
Overzicht



$$\varepsilon = \frac{\Delta_{sys}}{2}, \text{ uniform, CL } 100\%$$

Nauwkeurigheid en precisie van een meetresultaat

- Precisie (spreiding) ↔ statistische fouten
  Willekeurige fluctuaties door onafhanekelijke invloeden
  Gemiddelde van de verdeling = echte waarde (hypothese!)
- Nauwkeurigheid ↔ systematische fouten
  Door foute procedures of modellen
  Afwijking niet stochastisch
  Moeten kunnen worden verwijderd (groot probleem...!)



Resultaat bij statistische fouten: schatters

- De meetresultaten vormen een steekproef
  uit een bepaalde verdeling (normale, poisson, exponentiële, binomiaal...)
- Gemiddelde van de verdeling = echte waarde (hypothese!)
  Standaarddeviatie = meetfout (onzekerheid van één meetresultaat)
- We moeten die parameters (van de onderliggende verdeling) schatten
  Goede schatters: gemiddelde en standaarddeviatie van de steekproef:

$$\boxed{m = \frac{1}{N}\sum_{i=1}^{N} x_i}$$

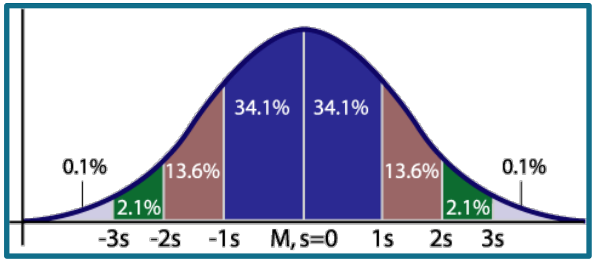
$$\boxed{s = \sqrt{\frac{\sum_{i=1}^{N}(x_i - m)^2}{N-1}}}$$

- Ons eindresultaat is het gemiddelde
  De variantie van m is σ^2/N, benaderd door s^2/N
  → de fout op het gemiddelde is $s/\sqrt{N}$
- Om de CL te bepalen, moeten we de verdeling van het gemiddelde kennen

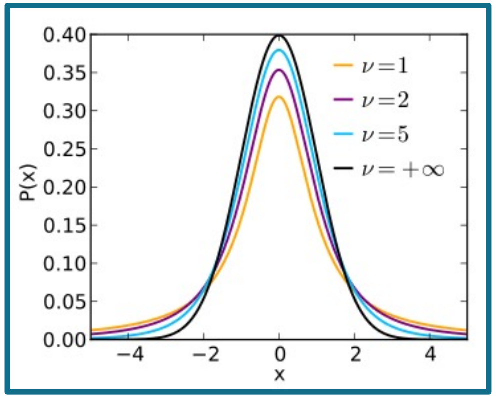
KU LEUVEN

Betrouwbaarheidsniveaus van de fout op het gemiddelde

- Grote steekproeven $N \gg 1$ (werkt voor $N > 20{,}30$):
  Centrale limietstelling
  → intervallen zijn normaal-verdeeld
  Resultaat wordt gegeven als $m \pm a\frac{s}{\sqrt{N}}$
  met 68% CL voor $a = 1$, 95% voor $a = 2$, 99.7% voor $a = 3$



- Kleine steekproeven $N < 30$:
  - Als x_i normaal-verdeeld dan
    intervallen volgen de Student t-verdeling met $\nu = N - 1$
    Resultaat wordt gegeven als $m \pm a\frac{s}{\sqrt{N}}$
    met CL die afhangen van N
    (worden normaal verdeeld voor $N > 10$)



  - Als x_i niet normaal-verdeeld: CL onbekend

Herhaalde metingen: gevallen

- Is de standaarddeviatie (veel) kleiner dan de fout van het instrument: dan schrijven we het resultaat als:

$$m \pm \frac{\Delta_{\text{sys}}}{2} \quad \text{100\% CL, uniform kansverdeling}$$

- Is de standaarddeviatie (veel) groter dan de fout van het instrument: dan gebruiken we de fout op het gemiddelde:

$$m \pm \frac{s}{\sqrt{N}} \quad \text{normale of Student kansverdeling}$$

- Zijn Δ_{sys} en s vergelijkbaar: meestal worden de twee apart gegeven

$$m \pm \frac{s}{\sqrt{N}} \pm \frac{\Delta_{\text{sys}}}{2}$$

- Zijn de fouten ε_i op elke meting verschillend: gebruiken we het **gewogen gemiddelde**

  De gewichten zijn: $g_i = \frac{1}{\epsilon_i^2}$

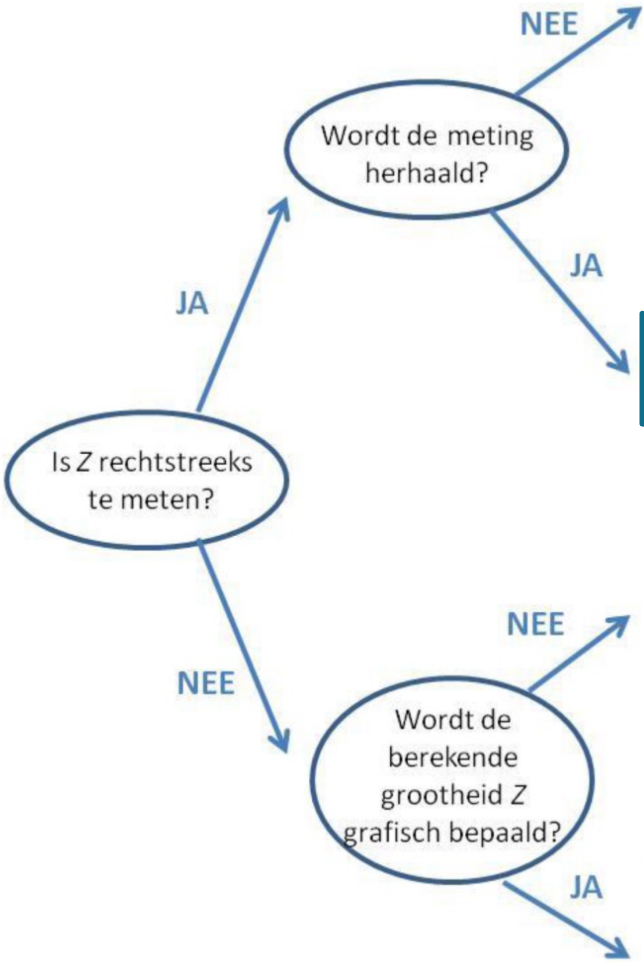
  Het resultaat is $m_g \pm \epsilon$

$$\boxed{m_g = \frac{\sum_{i=1}^{N} g_i x_i}{\sum_{i=1}^{N} g_i}}$$

$$\boxed{\epsilon = \frac{1}{\sqrt{\sum_{i=1}^{N} g_i}}}$$

KU LEUVEN

Overzicht



$\varepsilon = \frac{\Delta_{sys}}{2}$, uniform, CL 100%

$m \pm \frac{s}{\sqrt{N}} \pm \frac{\Delta_{\mathrm{sys}}}{2}$ of gewogen gemiddelde

Foutenpropagatie

$z = f(x,y)$

We meten $x \pm s_x$ en $y \pm s_y$. Wat is de fout op z?

- Algemeen: functies van stochastische variabelen:

$$s_z = \sqrt{\left(\frac{df}{dx}\right)^2 s_x^2 + \left(\frac{df}{dy}\right)^2 s_y^2}$$

Oppassen!!
- s_x^2, s_y^2 zijn de ***varianties***!
- geldt als s_x^2, s_y^2 klein (lineaire benadering)
  en x en y onafhankelijk van elkaar
- welke verdeling voor z? **Normale verdeling**!

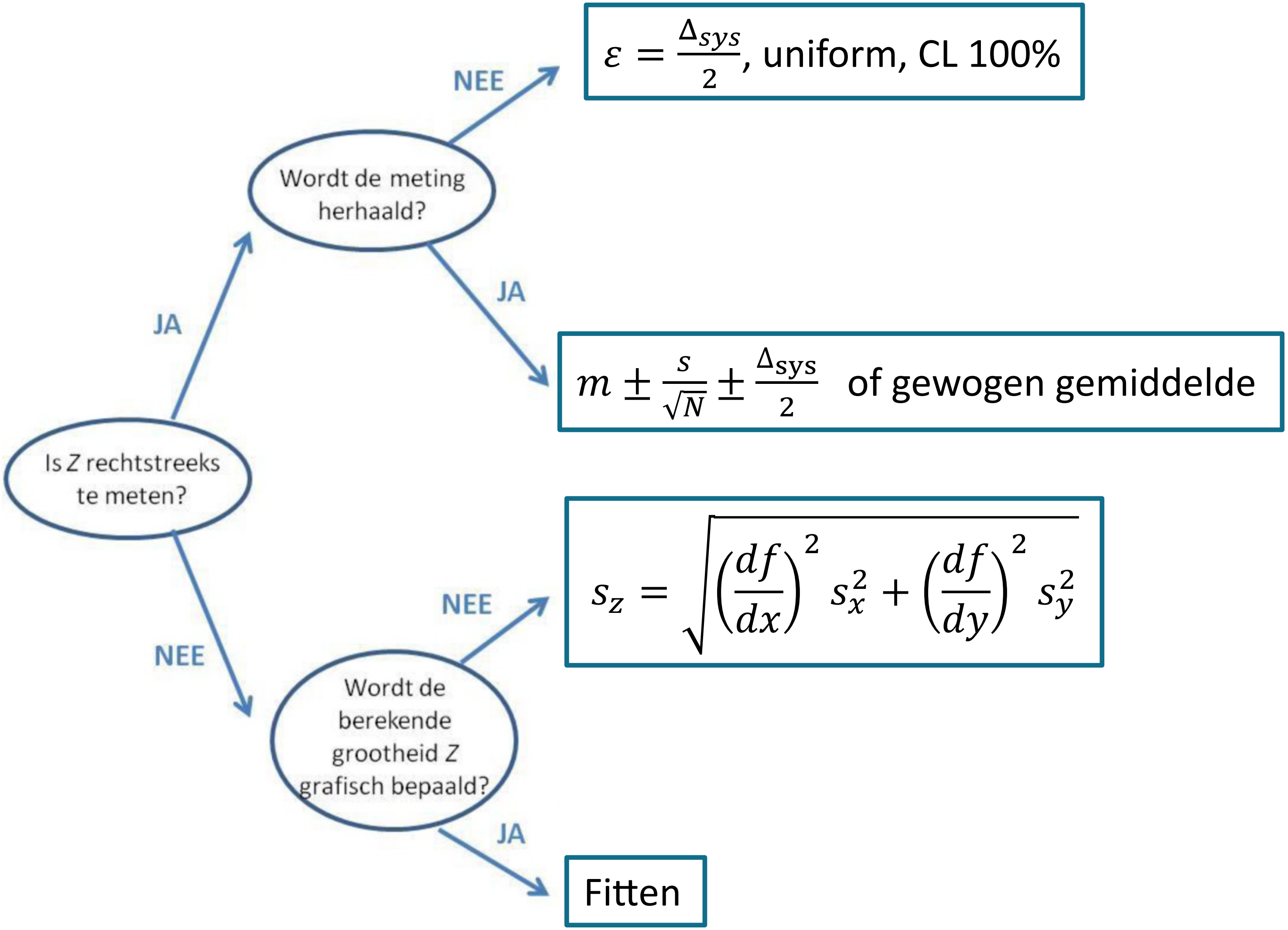
- Rekenregels (equivalent aan de algemene formule):

$$z = ax + by \qquad s_z^2 = a^2 s_x^2 + b^2 s_y^2$$

$$z = x^p y^q \qquad \left(\frac{s_z}{z}\right)^2 = p^2\left(\frac{s_x}{x}\right)^2 + q^2\left(\frac{s_y}{y}\right)^2$$

KU LEUVEN

Overzicht


Is Z rechtstreeks te meten?
JA
Wordt de meting herhaald?
NEE
$\varepsilon = \frac{\Delta_{sys}}{2}$, uniform, CL 100%
JA
$m \pm \frac{s}{\sqrt{N}} \pm \frac{\Delta_{sys}}{2}$ of gewogen gemiddelde
NEE
Wordt de berekende grootheid Z grafisch bepaald?
NEE
$s_z = \sqrt{\left(\frac{df}{dx}\right)^2 s_x^2 + \left(\frac{df}{dy}\right)^2 s_y^2}$
JA
Fitten

Meetfouten - Statistiek

- Meetfouten en hun schatting
- **Testen van hypotheses**
- Fitten

Statistiek

We moeten informatie inwinnen rond de verdeling van onze gemeten grootheid

Die komt uit onze <u>steekproef</u>

(Kans)Verdeling
alle mogelijke gebeurtenissen

Statistiek

Steekproef
een aantal gebeurtenissen

Experiment
- metingen
- beschrijvende statistiek

Onderliggende verdeling

Welke mogelijke onderliggende verdeling?

In de fysica komen deze meestal voor:

- Tel-experiment: Poisson
- Systematisch fout domineert (of slechts een waarde): uniforme verdeling
- Statistische fout domineert: normale verdeling
- Bij foutenschatting/statistiek: Student t-verdeling en χ^2

Schatters

Welke zijn de waarden van de parameters van de verdeling, en hoe nauwkeurig zijn ze bepaald?

We gebruiken schatters, berekend uit de steekproef

- Wens: de verwachtingswaarde van de schatter T moet gelijk zijn aan de te schatten grootheid: $\mathrm{E}[T] = \theta$
- Ook: $\mathrm{Var}[T]$ moet zo klein mogelijk zijn

Schatters

uitdrukkingen, op basis van de waarden van de steekproef, die een benadering geven van de parameters van de verdeling

- Notatie:
  t berekende waarde voor een bepaalde steekproef
  T stochastische variabele, uit de herhaalde berekening van t
  θ echte waarde van de te schatten parameter

Uitdrukkingen voor enkele schatters

Parameter van de verdeling	Schatter
Kans	Frequentie $f = \frac{n_k}{N}$ $\mathrm{Var}[F] = \frac{f(1-f)}{N}$
Gemiddelde	Steekproefgemiddelde $m = \frac{1}{N}\sum_{i=1}^{N} x_i$ $\mathrm{Var}[M] = \frac{\sigma^2}{N} \simeq \frac{s^2}{N}$
Variantie	Steekproefvariantie $s^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2$ of $s^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - m)^2$

Betrouwbaarheidsintervallen voor schatters

	X_i normaal verdeeld		X_i algemeen	
	Interval	CL	Interval	CL
kans				
$Nf < 30$	—	—	$p \in \dfrac{f + \frac{b^2}{2N}}{1 + \frac{b^2}{N}} \pm \dfrac{b\sqrt{\frac{b^2}{4N^2} + \frac{f(1-f)}{N}}}{1 + \frac{b^2}{N}}$	binomiaal
$Nf > 30$ $N(1-f) > 30$	—	—	$p \in f \pm z\sqrt{f(1-f)/N}$	normaal
gemiddelde				
$N < 30$	$\mu \in m \pm t\dfrac{s}{\sqrt{N}}$	student $\nu = N - 1$	$\mu \simeq m \pm v\dfrac{s}{\sqrt{N}}$	?
$N > 30$	$\mu \in m \pm z\dfrac{s}{\sqrt{N}}$	normaal	$\mu \in m \pm z\dfrac{s}{\sqrt{N}}$	normaal
variantie				
$N < 100$	$\dfrac{s^2}{\chi^2_{R1}} \leq \sigma^2 \leq \dfrac{s^2}{\chi^2_{R2}}$	$\chi^2_R(N-1)$	$\sigma^2 \simeq s^2 \pm v\sqrt{\dfrac{D_4 - s^4}{N-1}}$	?
$N > 100$	$\sigma^2 \in \dfrac{s^2}{1 \pm z\sqrt{2/(N-1)}}$	normaal	$\sigma^2 \in s^2 \pm z\sqrt{\dfrac{D_4 - s^4}{N-1}}$	normaal
standaard deviatie				
$N < 100$	$\sqrt{\dfrac{s^2}{\chi^2_{R1}}} \leq \sigma \leq \sqrt{\dfrac{s^2}{\chi^2_{R2}}}$	maxwell	$\sigma \simeq s \pm v\sqrt{\dfrac{D_4 - s^4}{4s^2(N-1)}}$	?
$N > 100$	$\sigma \in \dfrac{s}{1 \pm z\sqrt{1/[2(N-1)]}}$	normaal	$\sigma \in s \pm z\sqrt{\dfrac{D_4 - s^4}{4s^2(N-1)}}$	normaal

KU LEUVEN

We stappen verder

Met dezelfde ideeën kunnen we ook complexere vragen beantwoorden

- Is een theorie/model fout?
- Zijn de resultaten van twee metingen compatibel?
- Wat is eigenlijk de vorm van de onderliggende verdeling?
- Hoe kunnen we de betrouwbaarheid van een fit checken?

Algemene idee: een grootheid opbouwen, die een maat is voor die toetsen

Testen van hypothesen

= een waarde vergelijken met een voorspelling

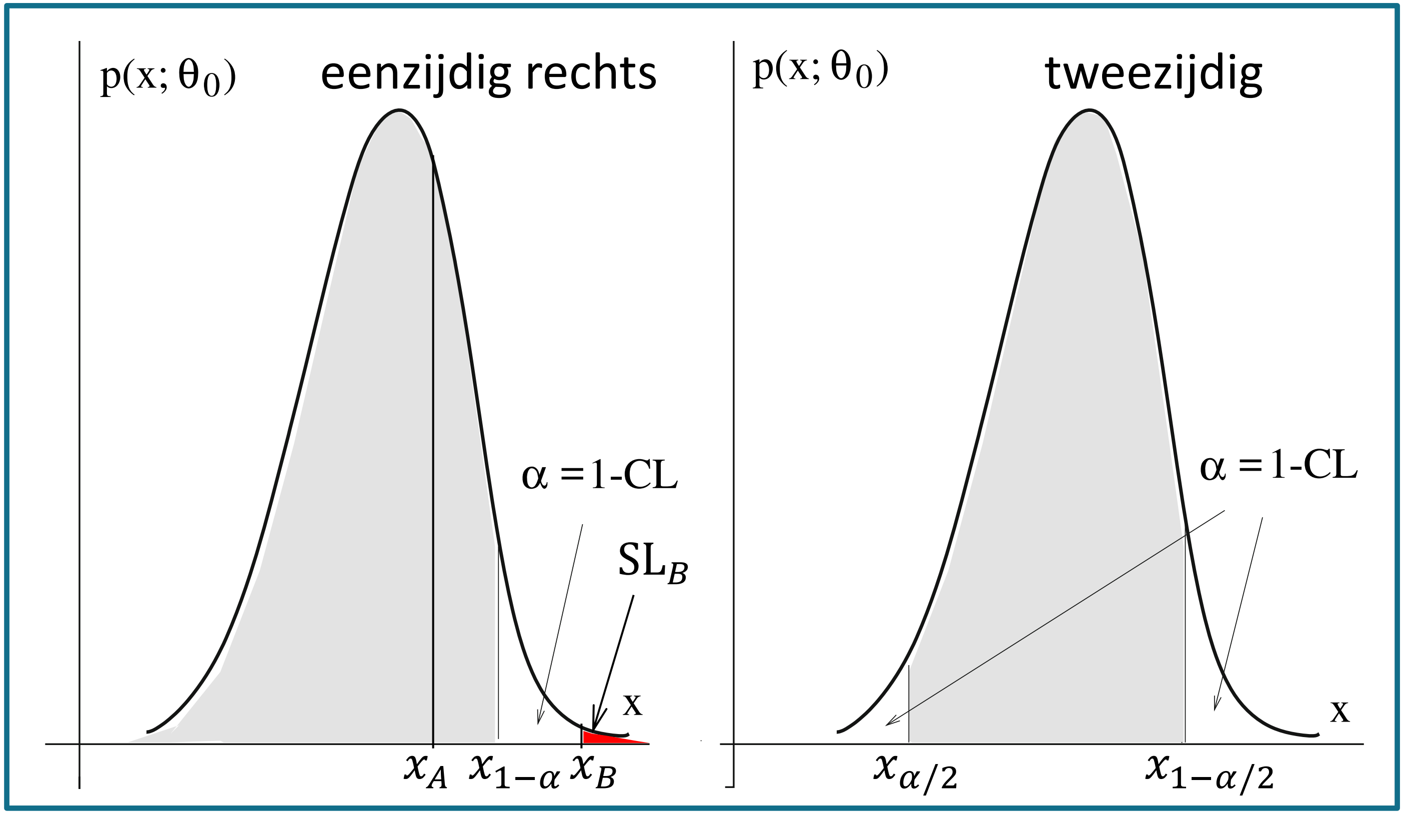
Stappen:

- Een (nul-)hypothese/model/theorie geeft een waarde w_0 (soms met een fout erop)
  We moeten onze resultaat w (met zijn fout) vergelijken met die waarde
- We bouwen de grootheid op: $x = |w - w_0|$ of, beter: $x = |w - w_0|/s_{\text{tot}}$
  Die is een "standaard" variabele; $s_{\text{tot}}^2 = s_{w_0}^2 + s_w^2$
  (de verdeling ervan moeten we kennen)
- We kiezen een <u>niveau</u> α van de test (onbetrouwbaarheidsdrempel)
  $\alpha = 1 - \text{CL}$ α moet klein zijn: 1%-5%
- We bereken de <u>p-waarde</u> (of <u>significantieniveau</u> SL) van onze meting:
  kans om x te behalen of overschrijden = <u>integraal</u> van de verdeling verder dan x
  $p(X < x)$ eenzijdig links
  $p(X > x)$ eenzijdig rechts
  $p(|X| > |x|)$ tweezijdig (meest voorkomende voor normale verdelingen)

$$\boxed{\text{(Nul-)Hypothese verwerpen als } p\text{-waarde} < \alpha}$$

KU LEUVEN

Testen van hypothesen



- Geval A: we meten x_A → p-waarde > α → hypothese niet verwerpen
- Geval B: we meten x_B → p-waarde < α → hypothese verwerpen

> Als we de hypothese verwerpen,
> is de kans om een fout te maken kleiner dan α

Oefening 1

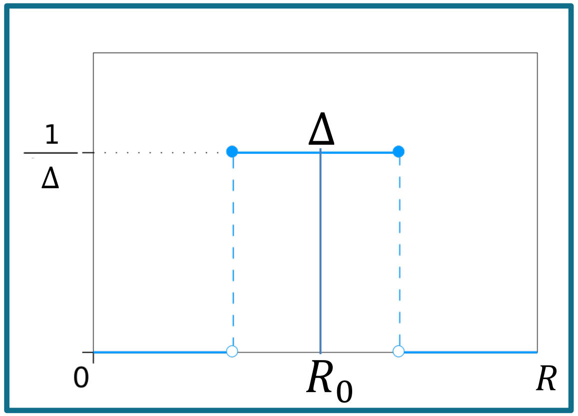
De gegeven waaarde van een weerstand (kleurcodes) is $R_0 = (4{,}7 \pm 0{,}2)\ \mathrm{k\Omega}$.
De gemeten waarde is $R = (4{,}3 \pm 0{,}1)\ \mathrm{k\Omega}$ (1σ-fout).
Mogen we de gegeven waarde aanvaarden, op een niveau van 1%?

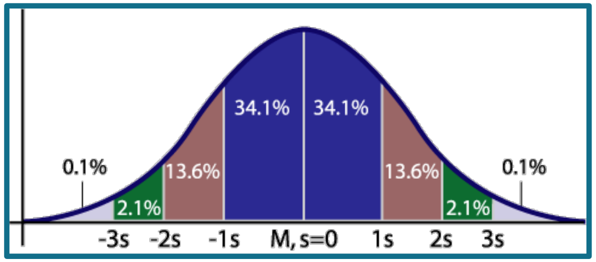
KU LEUVEN

Oefening 1

De gegeven waarde van een weerstand (kleurcodes) is $R_0 = (4{,}7 \pm 0{,}2)\ \mathrm{k\Omega}$.
De gemeten waarde is $R = (4{,}3 \pm 0{,}1)\ \mathrm{k\Omega}$ (1σ-fout).
Mogen we de gegeven waarde aanvaarden, op een niveau van 1%?

- We bouwen de standaard variabele op:
  $x = \frac{|R - R_0|}{s_{tot}}$
  Het probleem is symmetrisch → tweezijdige toets
- Wat is s_{tot}?
  $s_{tot}^2 = s_{R_0}^2 + s_R^2$ (we gebruiken de gemeten waarden)
  - De verdeling van R_0 is <u>uniform</u> en de Δ is 0,4
    $\rightarrow\ s_{R_0}^2 = \frac{\Delta^2}{12} = \frac{0{,}4^2}{12} = 0{,}013$
  - De verdeling van R is <u>normaal</u> (afgeleide grootheid)
    $\rightarrow\ s_R^2 = 0{,}01$
  Dus is $s_{tot} = \sqrt{0{,}013 + 0{,}01} = 0{,}15$
- De waarde van x voor deze toets is
  $x = \frac{|4{,}3 - 4{,}7|}{0{,}15} = 2{,}62$



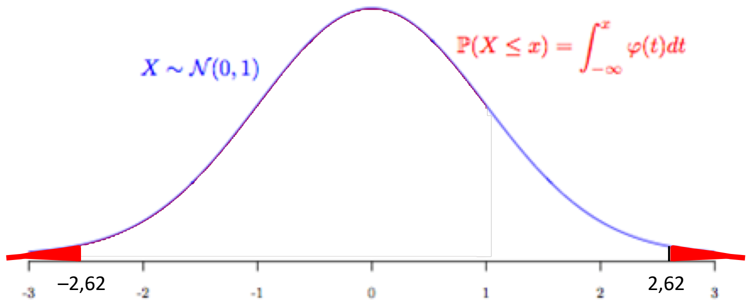


Oefening 1

De gegeven waarde van een weerstand (kleurcodes) is $R_0 = (4{,}7 \pm 0{,}2)\ \mathrm{k\Omega}$.
De gemeten waarde is $R = (4{,}3 \pm 0{,}1)\ \mathrm{k\Omega}$ (1σ-fout).
Mogen we de gegeven waarde aanvaarden, op een niveau van 1%?

$x = 2{,}62$; welke is de p-waarde?

- Welke verdeling voor X?
  Samenstelling van uniforme en normaal
  → normale verdeling
- $p(|X| > 2{,}62) = 2{\times}p(X > 2{,}62) =$
  $= 2{\times}(1 - 0{,}9956) = 0{,}88\%$
- p-waarde < α
  → we moeten de hypothese <u>verwerpen</u>!
- De kans om een fout te maken door de hypothese te verwerpen is kleiner dan 1%



Belangrijk: een hypothese niet verwerpen ("aanvaarden") betekent niet dat de hypothese correct is.
Het betekent: we kunnen de hypothese niet verwerpen op basis van onze resultaten alleen.

Vergelijking tussen twee meetresultaten

- We hebben twee metingen w_1 en w_2 en we vragen ons af of ze uit dezelfde populatie komen
- We bouwen de standaard variabele op: $x = \frac{|w_1 - w_2|}{\sqrt{s_1^2 + s_2^2}}$
- Als w_1 en w_2 normaal verdeeld zijn, dan is x ook normaal verdeeld en zijn de betrouwbaarheidsniveaus die van een normale verdeling
- Het probleem is symmetrisch, we gebruiken een tweezijdig toets

Oefening 2

Twee verschillende experimenten hebben dezelfde grootheid gemeten, met de resultaten: $x_1 = 12{,}3 \pm 0{,}5$ en $x_2 = 13{,}5 \pm 0{,}8$, waar de betrouwbaarheidsintervallen normaal verdeeld zijn (en die zijn 1σ-fouten).

Zijn de twee resultaten compatibel, op een niveau van 5%?

Oefening 2

Twee verschillende experimenten hebben dezelfde grootheid gemeten, met de resultaten: $x_1 = 12{,}3 \pm 0{,}5$ en $x_2 = 13{,}5 \pm 0{,}8$, waar de betrouwbaarheidsintervallen normaal verdeeld zijn (en die zijn 1σ-fouten).

Zijn de twee resultaten compatibel, op een niveau van 5%?

- We bouwen de standaard variabele op

$$x = \frac{|12{,}3 - 13{,}5|}{\sqrt{(0{,}5)^2 + (0{,}8)^2}} = 1{,}27$$

- p-waarde: $p(|X| > 1{,}27) = 2\times(1 - 0{,}898) = 20{,}4\%$

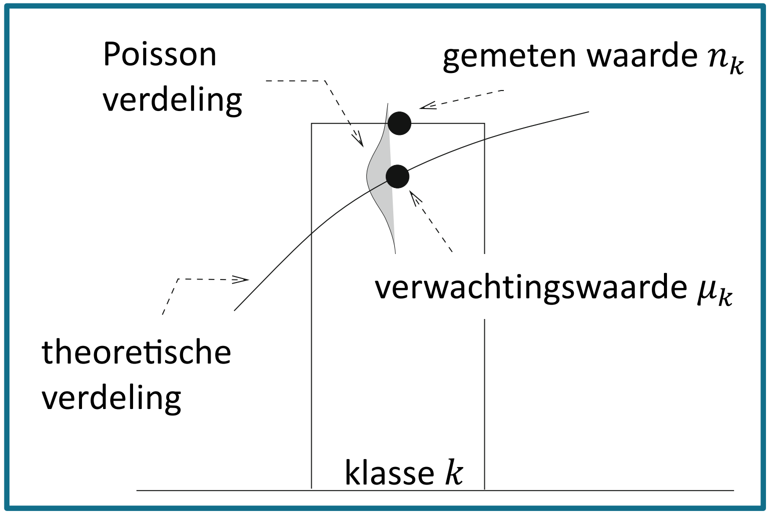
- p-waarde > α → hypothese niet verwerpen!!
  (we kunnen niet zeggen dat de twee metingen niet compatibel zijn)

Toets voor de vorm van een verdeling: χ^2-toets

- We gebruiken dezelfde principe:
  een variabele opbouwen uit het verschil tussen de verwachte waarde en de gemeten waarde

- We doen dat voor <u>elke klasse k</u> van de steekproef:
  μ_k verwachte waarde in de klasse k
  n_k gemeten waarde in de klasse k
  variabele: $t_k = \frac{|n_k - \mu_k|}{\sigma_k}$

- De waarde μ_k zijn <u>gegevens</u> uit ons model:
  <u>alle</u> verdelingen zijn mogelijk.
  De waarde σ_k hangt af van de verdeling van de tellingen n_k binnen de klasse:
  het is een tel-experiment → <u>Poisson</u> → $\sigma_k = \sqrt{\mu_k}$

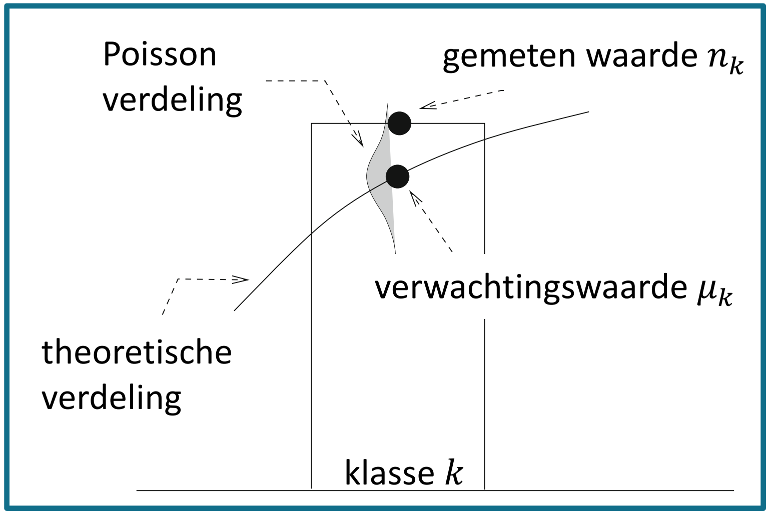


verder →

KU LEUVEN

Toets voor de vorm van een verdeling: χ²-toets

- We groeperen de klassen met minder dan 5 tellingen
  → de statistiek van elke klasse wordt normaal



- We tellen de variabelen, eigenlijk hun kwadraten, op:

$$\boxed{\chi^2_{K-1} = \sum_{k=1}^{K} \frac{(n_k - \mu_k)^2}{\mu_k}}$$

  (de som van de kwadraten van normale variabelen is en χ^2-variabele)

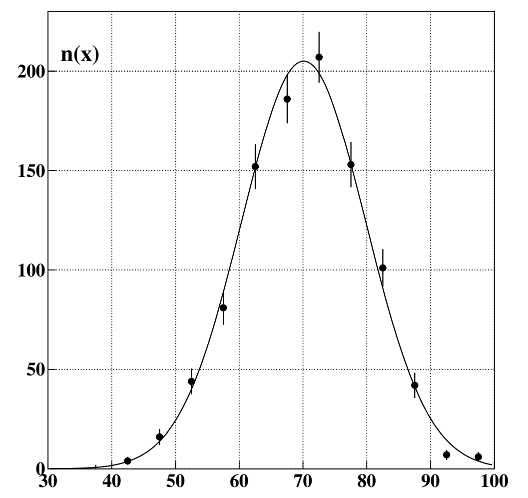
- We gebruiken de kwantielen van de χ^2-verdeling om de toets te maken, meestal eenzijdig rechts

- **!! Aantal vrijheidsgraden = aantal klassen – 1**

KU LEUVEN

Oefening 3

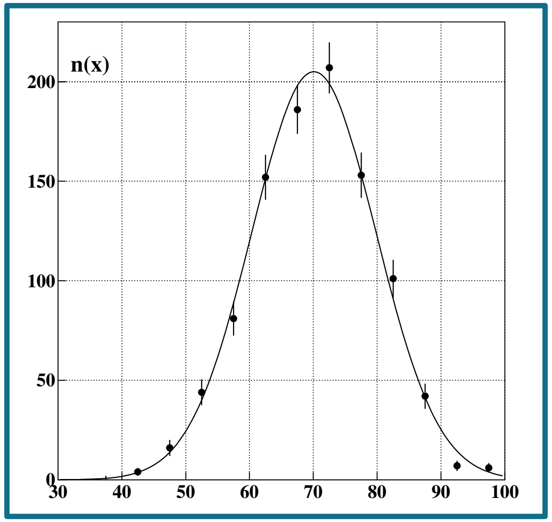
Beschouw het histogram in de figuur. De inhoud van elke klasse is in de tabel weergegeven. Bereken of de waarden afkomstig zijn van een normale verdeling met met $\mu = 71$, $\sigma = 10$, op een niveau van 5%.



1	2
x_i	n_i
37.5	1
42.5	4
47.5	16
52.5	44
57.5	81
62.5	152
67.5	186
72.5	207
77.5	153
82.5	101
87.5	42
92.5	7
97.5	6

Oefening 3

Beschouw het histogram in de figuur. De inhoud van elke klasse is in de tabel weergegeven. Bereken of de waarden afkomstig zijn van een normale verdeling met met $\mu = 71$, $\sigma = 10$, op een niveau van 5%.



1	2
x_i	n_i
37.5	1
42.5	4
47.5	16
52.5	44
57.5	81
62.5	152
67.5	186
72.5	207
77.5	153
82.5	101
87.5	42
92.5	7
97.5	6

- We bouwen de standaard variabelen op:
  $t_k = \frac{|n_k - \mu_k|}{\sqrt{\mu_k}}$
  waar de μ_k zijn berekend uit de normale verdeling met de gegeven μ en σ, voor elke waarde van x_k:
  $$\mu_k = N \times \Delta x \times \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x_k - \mu)^2}{2\sigma^2}\right)$$
- We sommen de eerste twee klassen op: $t_{12} = \frac{|n_1 - \mu_1 + n_2 - \mu_2|}{\sqrt{\mu_1 + \mu_2}}$
- We berekenen de som van de kwadraten van de t_i
  $\chi^2 = \sum_{k=1}^{12} \frac{(n_k - \mu_k)^2}{\mu_k} = 14{,}9$
- p-waarde (11 vrijheidsgraden, eenzijdig rechts):
  $p(Q_{11}^2 > 14{,}9) = 18{,}6\% > 5\%$ dus ok (hypothese niet verwerpen)

Meetfouten - Statistiek

- Meetfouten en hun schatting
- Testen van hypotheses
- **Fitten**

Fitten

- Soms moeten we een model toetsen, die geen simpel waarde is
  b.v. een verband tussen twee (of meer) grootheden
  Het model voorspelt de vorm van de functie en de parameters

- We meten de grootheden
  Hoe kunnen we de parameters schatten?
  Welke is de onzekerheid op die geschatte waarden?
  → ML-methode

- Hoe kunnen we de betrouwbaarheid van een fit checken?
  → LS-methode

Aannemelijkheidsfunctie (Maximum Likelihood ML)

Het probleem in het algemeen, nog een keer

- We hebben N waarden waargenomen van de variabele X:
  $x_1, x_2, \dots, x_N$
  De variabele X is verdeeld volgens $p(x, \boldsymbol{\theta})$,
  waar $\boldsymbol{\theta} = \theta_1, \theta_2, \dots, \theta_P$ zijn de echte waarden van de parameters
  (meest voorkomende) voorbeelden van die $\boldsymbol{\theta}$: het gemiddelde en de variantie

- We moeten de "beste" waarden van de parameters $\boldsymbol{\theta}$ vinden
  door het gebruik van **schatters** $\boldsymbol{T}$
  we gebruiken het concept: "beste" = meest aannemelijk

- Wat is de kans om precies <u>die waarden</u> $x_1, x_2, \dots, x_N$ te meten?
  Het is het product van de kans om elke x_i te meten:

$$L(\boldsymbol{\theta}, \boldsymbol{x}) = p(x_1, \boldsymbol{\theta})p(x_2, \boldsymbol{\theta}) \dots p(x_N, \boldsymbol{\theta}) = \prod_{i=1}^{N} p(x_i, \boldsymbol{\theta})$$

> ML methode: de schatters $\boldsymbol{T}$ berekenen
> als waarden van $\boldsymbol{\theta}$ waarvoor $L(\boldsymbol{\theta}, \boldsymbol{x})$ maximaal is

als functie van de $\boldsymbol{\theta}$; terwijl x_i zijn constanten

KU LEUVEN

Aannemelijkheidsfunctie (Maximum Likelihood ML)

- Heel vaak gebruiken we

$$\mathcal{L}(\boldsymbol{\theta}, \boldsymbol{x}) = -\ln L(\boldsymbol{\theta}, \boldsymbol{x}) = -\sum_{i=1}^{N} \ln p(x_i, \boldsymbol{\theta})$$

en we vinden het <u>minimum</u> ervan

- Dat kunnen we, door de afgeleiden te berekenen:

$$\frac{\partial \mathcal{L}(\boldsymbol{T}, \boldsymbol{x})}{\partial \theta_k} = -\sum_{i=1}^{N} \left[\frac{\partial \ln p(x_i, \boldsymbol{T})}{\partial \theta_k}\right] = 0$$

voor het P aantal parameters θ_1, θ_2,…, θ_P.
Deze zijn P vergelijkingen in de onbekende T_k; die T_k zijn onze <u>schatters</u>
Ze zijn stochastische variabelen, die van de waarden x_i afhangen

- De vorm van de verdelingen $p(x, \boldsymbol{\theta})$ wordt meestal aangenomen en op het einde getoetst (b.v. met een χ^2-toets)

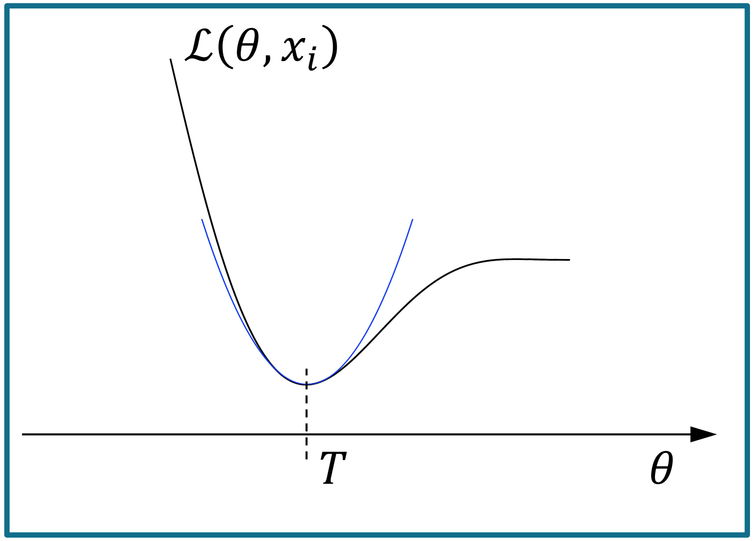
KU LEUVEN

Fout op de ML schatters

- Beschouw de tweede afgeleide van $\mathcal{L}$: $\frac{\partial^2 \mathcal{L}(\theta,\boldsymbol{x})}{\partial \theta^2}$
- De verwachtingswaarde (of echte gemiddelde): $\mathrm{E}\left[\frac{\partial^2 \mathcal{L}(\theta,\boldsymbol{X})}{\partial \theta^2}\right]$ is de informatie functie van ons model (functie van θ)
- We kunnen zien dat de informatie een maat is van hoe "scherp" het minimum is. Hoe groter de informatie, hoe beter is het minimum bepaald
- Formeel resultaat:

$$\boxed{\mathrm{Var}[T] = \left(\mathrm{E}\left[\frac{\partial^2 \mathcal{L}(\theta,\boldsymbol{X})}{\partial \theta^2}\right]\right)^{-1}}$$

Die uitdrukking bevat soms θ (of andere parameters). Als het zo is, dan moeten we in principe de "echte" waarde voor θ gebruiken. Maar die is meestal onbekend, dus gebruiken we de geschatte waarde T



Eigenschappen van de ML schatters

$$\frac{\partial \mathcal{L}(\boldsymbol{T}, \boldsymbol{x})}{\partial \theta_k} = -\sum_{i=1}^{N} \left[\frac{\partial \ln p(x_i, \boldsymbol{T})}{\partial \theta_k} \right] = 0$$

Voordelen van de ML-methode (hier niet bewezen):

- (Consistentie)
  De asymptotische* (voor grote N) verwachtingswaarde $\mathrm{E}[T_k]$ van de schatters T_k zijn de echte waarden van de parameters $\boldsymbol{\theta_k}$ van de onderliggende verdeling
- (Efficiëntie)
  De variantie van T_k is de kleinste mogelijk van alle schatters
- De schatters T_k zijn (asymptotisch) normaal verdeeld
  We gebruiken normale betrouwbaarheidsniveaus al voor $N > 10$

*Idealiter, zou het moeten zijn: $\mathrm{E}[(T_k)_N] = \theta_k$ voor *alle* waarden van N (zuiver schatter) maar het is zo alleen voor $N \to \infty$

Oefening 1

In en tel-experiment meten we N waarden x_i van onze grootheid.
Bereken de ML waarde van het gemiddelde en zijn standaarddeviatie.

Oefening 1

In en tel-experiment meten we N waarden x_i van onze grootheid.
Bereken de ML waarde van het gemiddelde en zijn standaarddeviatie.

- L-functie: product van N keer de waarde van de Poisson verdeling voor de gemeten waarden $x_1, x_2, \dots x_N$:

$$L(\mu, x_i) = \prod_{i=1}^{N} \frac{\mu^{x_i} e^{-\mu}}{x_i!}$$

- De (-)logaritme ervan is:

$$\mathcal{L}(\mu, x_i) = -\sum_{i=1}^{N} (x_i \ln \mu - \mu - \ln x_i!) = -\ln \mu \sum_{i=1}^{N} x_i + N\mu + \sum_{i=1}^{N} \ln x_i!$$

- De afgeleide t.o.v. de parameter “gemiddelde”:

$$\frac{\partial \mathcal{L}}{\partial \mu} = -\frac{\sum_{i=1}^{N} x_i}{\mu} + N \ ; \ -\frac{\sum_{i=1}^{N} x_i}{M} + N = 0 \rightarrow M = \sum_{i=1}^{N} \frac{x_i}{N}$$

Oefening 1

In en tel-experiment meten we N waarden x_i van onze grootheid.
Bereken de ML waarde van het gemiddelde en zijn standaarddeviatie.

- Voor de standaarddeviatie berekenen we de variantie:

$$\mathrm{Var}[M] = \left(\mathrm{E}\left[\frac{\partial^2 \mathcal{L}}{\partial \mu^2}\right]\right)^{-1}$$

$$\frac{\partial^2 \mathcal{L}}{\partial \mu^2} = +\frac{\sum_{i=1}^{N} x_i}{\mu^2} \rightarrow \mathrm{E}\left[\frac{\partial^2 \mathcal{L}}{\partial \mu^2}\right] = \frac{N\mu}{\mu^2} = \frac{N}{\mu} \rightarrow \mathrm{Var}[M] = \frac{\mu}{N} \backsimeq \frac{M}{N}$$

- De ML-schatting van het gemiddelde van een Poisson verdeling is dus

$$\mu = M \pm \frac{\sqrt{\mu}}{\sqrt{N}} \simeq \sum_{i=1}^{N} \frac{x_i}{N} \pm \frac{\sqrt{M}}{\sqrt{N}}$$

met normale betrouwbaarheidsniveaus voor $N > 10$.

Oefening 2

We meten N waarden x_i van een normaal-verdeelde grootheid.
Bereken de ML waarde van het gemiddelde en zijn standaarddeviatie.

KU LEUVEN

Oefening 2

We meten N waarden x_i van een normaal-verdeelde grootheid.
Bereken de ML waarde van het gemiddelde en zijn standaarddeviatie.

- L-functie: product van N keer de waarde van de normaal verdeling voor de gemeten waarden $x_1, x_2, \dots x_N$:

$$L(\mu,\sigma,x_i) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x_i-\mu)^2}{2\sigma^2}\right] = \frac{1}{\left(\sqrt{2\pi\sigma^2}\right)^N} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{N}(x_i-\mu)^2\right]$$

- De (-)logaritme ervan is:

$$\mathcal{L}(\mu,\sigma,x_i) = +\frac{N}{2}\ln(2\pi\sigma^2) + \frac{1}{2\sigma^2}\sum_{i=1}^{N}(x_i-\mu)^2$$

- De afgeleide t.o.v. de parameter "gemiddelde":

$$\frac{\partial\mathcal{L}}{\partial\mu} = -\frac{1}{\sigma^2}\sum_{i=1}^{N}(x_i-\mu) \quad ; \quad -\frac{1}{\sigma^2}\sum_{i=1}^{N}(x_i-M) = 0 \rightarrow M = \sum_{i=1}^{N}\frac{x_i}{N}$$

Oefening 2

We meten N waarden x_i van een normaal-verdeelde grootheid.
Bereken de ML waarde van het gemiddelde en zijn standaarddeviatie.

- De tweede afgeleide is:
$$\frac{\partial^2 \mathcal{L}}{\partial \mu^2} = \frac{\partial}{\partial \mu}\left(-\frac{1}{\sigma^2}\sum_{i=1}^{N}(x_i - \mu)\right) = \cancel{\frac{\partial}{\partial \mu}\left(-\frac{\sum_{i=1}^{N} x_i}{\sigma^2}\right)} + \frac{\partial}{\partial \mu}\left(\frac{N\mu}{\sigma^2}\right) = \frac{N}{\sigma^2}$$
- En dus de variantie is:
$$\mathrm{Var}[M] = \left(\mathrm{E}\left[\frac{\partial^2 \mathcal{L}}{\partial \mu^2}\right]\right)^{-1} = \frac{\sigma^2}{N} \backsimeq \frac{S^2}{N}$$
- De ML-schatting van het gemiddelde van een normale verdeling is:
$$\mu = M \pm \frac{\sqrt{\sigma^2}}{\sqrt{N}} \simeq \sum_{i=1}^{N} \frac{x_i}{N} \pm \frac{\sqrt{S^2}}{\sqrt{N}}$$
met normale betrouwbaarheidsniveaus voor $N > 10$.

Oefening 3

We meten N waarden x_i van een normaal-verdeelde grootheid.
Bereken de ML waarde van de variantie.

Oefening 3

We meten N waarden x_i van een normaal-verdeelde grootheid.
Bereken de ML waarde van de variantie.

- Als in oefening 2, is de logaritme van de L-functie:

$$\mathcal{L}(\mu, \sigma, x_i) = +\frac{N}{2}\ln(2\pi\sigma^2) + \frac{1}{2\sigma^2}\sum_{i=1}^{N}(x_i - \mu)^2$$

- De afgeleide t.o.v. de parameter "variantie" σ^2 is:

$$\frac{\partial\mathcal{L}}{\partial(\sigma^2)} = \frac{N}{2\sigma^2} - \frac{1}{2(\sigma^2)^2}\sum_{i=1}^{N}(x_i - \mu)^2$$

- En dus:

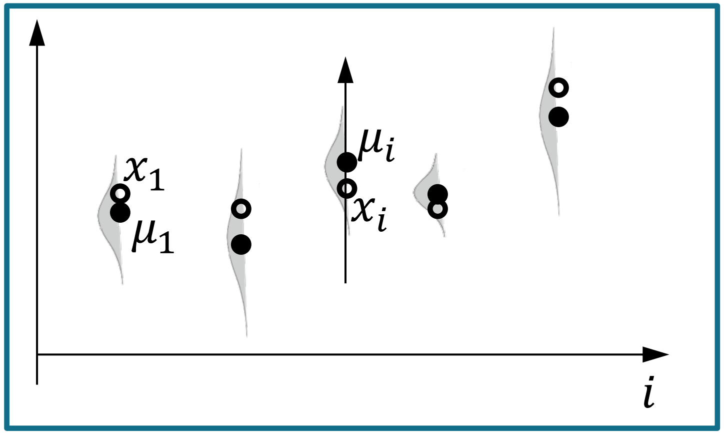
$$\frac{N}{2S^2} - \frac{1}{2(S^2)^2}\sum_{i=1}^{N}(x_i - \mu)^2 = 0 \;\rightarrow\; S^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2$$

Voor μ moeten we de andere vergelijking $\partial\mathcal{L}/\partial\mu = 0$ gebruiken, en dus de uitdrukking voor de schatter $M = \sum_{i=1}^{N}\frac{x_i}{N}$.

De ML-schatter van σ^2 is niet zuiver!

Kleinste-kwadratenmethode (Least-Squares LS)

- We willen nu een P aantal parameters $\boldsymbol{\theta}$ vinden die een reeks N variabelen x_i uit <u>verschillende</u> normale verdelingen beschrijven
  [De verdelingen zijn niet helemaal onafhankelijke van elkaar, door de parameters $\boldsymbol{\theta}$]



- ML methode:

$$L(\boldsymbol{\theta}, \boldsymbol{x}) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma_i^2(\boldsymbol{\theta})}} \exp\left[-\frac{\left(x_i - \mu_i(\boldsymbol{\theta})\right)^2}{2\sigma_i^2(\boldsymbol{\theta})}\right]$$

$$\mathcal{L}(\boldsymbol{\theta}, \boldsymbol{x}) = \frac{1}{2}\sum_{i=1}^{N} \ln 2\pi\sigma_i^2(\boldsymbol{\theta}) + \frac{1}{2}\sum_{i=1}^{N} \frac{\left(x_i - \mu_i(\boldsymbol{\theta})\right)^2}{\sigma_i^2(\boldsymbol{\theta})}$$

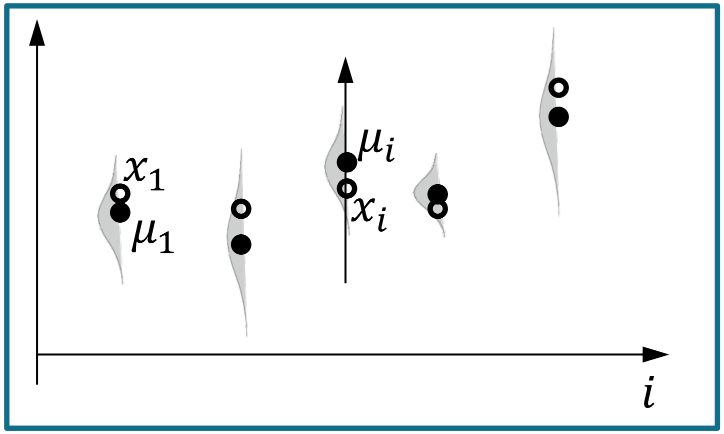
- We moeten een minimum van $\mathcal{L}$ vinden
  Als de σ_i gekend zijn, of als we de s_i gebruiken: dan hangen de σ_i niet af van $\boldsymbol{\theta}$
  Dan moeten we een minimum vinden van

$$\boxed{\mathrm{LS}(\boldsymbol{\theta}) \equiv \sum_{i=1}^{N} \frac{\left(x_i - \mu_i(\boldsymbol{\theta})\right)^2}{s_i^2}}$$

KU LEUVEN

LS methode: eigenschappen

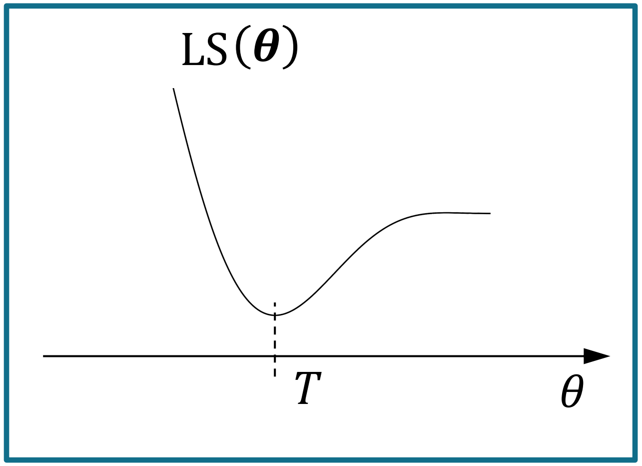
$$\mathrm{LS}(\boldsymbol{\theta}) \equiv \sum_{i=1}^{N} \frac{\left(x_i - \mu_i(\boldsymbol{\theta})\right)^2}{s_i^2}$$



- We hebben dat afgeleid voor x_i normaal verdeeld en met gekende σ_i
- Maar (eigenschap van LS, niet bewezen):
  De LS schatters zijn altijd zuiver ($\mathrm{E}[T] = \theta$),
  ook als de variabelen x_i niet normaal verdeeld zijn!
  En we hoeven die verdelingen niet te kennen! (alleen de verwachtingswaarden)
- Als de x_i toch normaal verdeeld zijn, dan is de LS functie verdeeld als een χ^2 met $N - P$ vrijheidsgraden (P aantal parameters)
  Dat is meestal het geval, vooral voor fysische meetresultaten;
  daarom gebruiken we soms het χ^2 symbool voor de LS functie
- Als de $\mu_i(\boldsymbol{\theta})$ lineair zijn in de parameters $\boldsymbol{\theta}$,
  dan heeft de LS-methode alle eigenschappen van de ML-methode
  (de variantie van de LS-schatters is de kleinste mogelijk)

LS methode: oplossing

$$\mathrm{LS}(\boldsymbol{\theta}) \equiv \sum_{i=1}^{N} \frac{\left(x_i - \mu_i(\boldsymbol{\theta})\right)^2}{s_i^2}$$



- Het minimum vinden we op de waarden $\boldsymbol{T}$ waar de afgeleiden van $LS(\boldsymbol{\theta})$ nul zijn

$$\boxed{\frac{\partial(\mathrm{LS})}{\partial\theta_k} = 0}$$

- Als de $\mu_i(\boldsymbol{\theta})$ lineair zijn in de parameters $\boldsymbol{\theta}$
  dan is die een stelsel van lineaire vergelijkingen
  → <u>kan analytisch worden opgelost</u>

- Als de $\mu_i(\boldsymbol{\theta})$ <u>niet</u> lineair zijn in de parameters $\boldsymbol{\theta}$
  dan moeten we
  - eerst een schatting maken van de initiële waarden
  - dan een nul van de afgeleide vinden
    m.b.v. iteratieve methoden (minimalisatie-programma's)
    (b.v. met de methode van Newton)

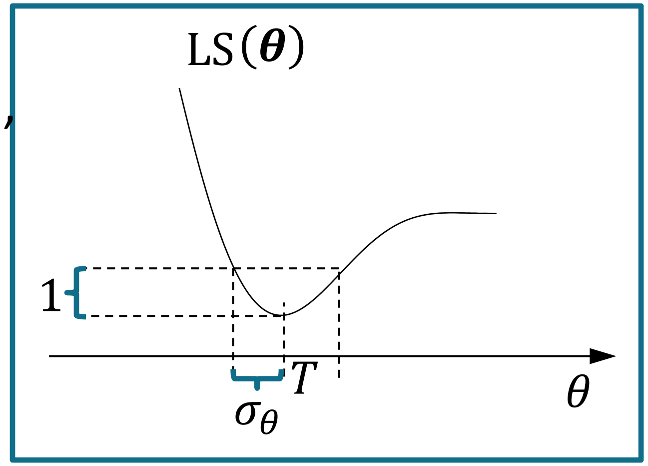
LS methode: fout van de schatters

- De variantie van de geschatte parameters

  is dezelfde als de ML-schatters: omdat $\frac{\partial^2 \mathcal{L}}{\partial \theta^2} = \frac{1}{2}\frac{\partial^2 (\text{LS})}{\partial \theta^2}$,

  $$\rightarrow \boxed{\text{Var}[T] = 2\left(\frac{\partial^2 (\text{LS})}{\partial \theta^2}\right)^{-1}_{\theta=T}}$$

  en de betrouwbaarheidsniveaus zijn normaal



- Hoe verandert $\text{LS}(\boldsymbol{\theta})$ rond het minimum?
  Taylorontwikkeling:

$$\text{LS}(\theta) = \text{LS}(T) + \cancel{\left(\frac{\partial (\text{LS})}{\partial \theta}\right)_{\theta=T}}(\theta - T) + \frac{1}{2}\left(\frac{\partial^2 (\text{LS})}{\partial \theta^2}\right)_{\theta=T}(\theta - T)^2 + \cdots$$

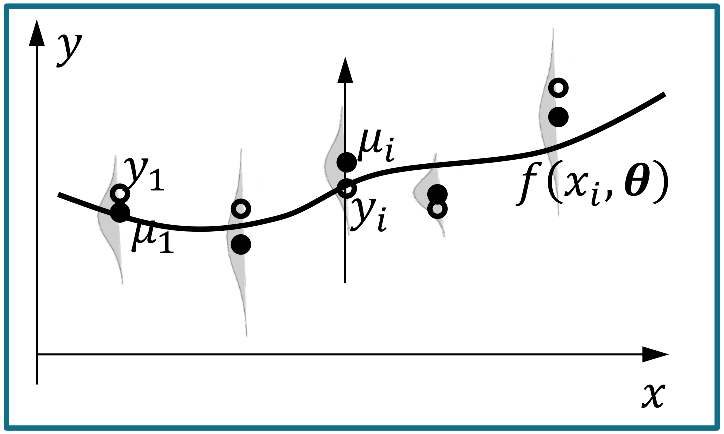
Als we één standaarddeviatie σ_θ nemen voor $\theta - T$, dan is het

$$\text{LS}(T \pm \sigma_\theta) - \text{LS}(T) = \frac{1}{2}\left(\frac{\partial^2 (\text{LS})}{\partial \theta^2}\right)_{\theta=T} \times \text{Var}[T] = 1$$

De verandering van de LS-functie op afstand σ_θ is 1
We kunnen dat gebruiken om σ_θ te berekenen

Toepassing LS methode I: fit van een model

- De μ_i zijn vaak functies van een onafhankelijke (niet-stochastische) variabele
  Dan schrijven we:



gemeten waarden

gekozen functie (model)

waarden van de onafhankelijke variabele

$$\mathrm{LS}(\boldsymbol{\theta}) \equiv \sum_{i=1}^{N} \frac{\left(y_i - f(x_i, \boldsymbol{\theta})\right)^2}{s_i^2}$$

Oppassen: de x_i van de vorige slides zijn nu y_i geworden
de verwachten waarden zijn nu $f(x_i, \boldsymbol{\theta})$
de nieuwe x_i zijn de waarden van de onafhankelijke variabele

- Voorbeeld: lineaire regressie $f(x, \boldsymbol{\theta}) = ax + b$
  zie ook formularium

- Checken of de fit goed is (alleen als de y_i normaal verdeeld zijn):
  - $\mathrm{LS} \equiv \chi^2 \approx N - p$ of $\chi^2/(N-p) \approx 1$ (p is het aantal parameters)
  - Meer quantitatief: χ^2-toets

Toepassing LS methode II: gewogen gemiddelde

- Ander probleem: stel nu voor dat we meerdere (K) steekproeven hebben van dezelfde grootheid X.
  De steekproef i heeft N_i elementen $x_{1i}, x_{2i}, \dots, x_{ji}, \dots, x_{N_i i}$
  en een (steekproef)variantie s_i^2

  Wat is het gemiddelde van X? We gebruiken de LS-schatter

- De LS-functie is: $\chi^2(\boldsymbol{\theta}) \equiv \sum_{i=1}^{K} \sum_{j=1}^{N_i} \frac{(x_{ji}-\mu)^2}{s_i^2}$

- De schatter voor μ is
  $$\frac{\partial \chi^2}{\partial \mu} = -2 \sum_{i=1}^{K} \sum_{j=1}^{N_i} \frac{x_{ji}-\mu}{s_i^2} = -2 \sum_{i=1}^{K} \frac{N_i m_i}{s_i^2} + 2 \sum_{i=1}^{K} \frac{N_i \mu}{s_i^2}$$
  $$\frac{\partial \chi^2}{\partial \mu} = 0 \quad \rightarrow \quad M = \frac{\sum_{i=1}^{K} N_i m_i / s_i^2}{\sum_{i=1}^{K} N_i / s_i^2}$$

- We stellen nu $g_i = N_i / s_i^2$: dat is het gewogen gemiddelde
  De fout erop vinden we met de tweede afgeleide.

Notatie

Afspraak voor beduidende cijfers van meetresultaten en onzekerheden

- Schrijf de onnauwkeurigheid met één (of twee) beduidend(e) cijfer(s)
  Als het beduidend cijfer een "1" is → gebruik steeds 2 beduidende cijfers
- Pas het maatgetal aan aan de onnauwkeurigheid:
  het meetresultaat moet met eenzelfde aantal cijfers na de komma geschreven worden als de onnauwkeurigheid aangeeft (waarbij maatgetal en onnauwkeurigheid uitgedrukt zijn in dezelfde eenheden)

Voorbeelden

Niet $(1{,}9342 \pm 0{,}03)$ cm	maar	$(1{,}93 \pm 0{,}03)$ cm
Niet $(1{,}3 \pm 0{,}02)$ kg	maar	$(1{,}30 \pm 0{,}02)$ kg
Niet $(3{,}213 \pm 0{,}490)$ ms^{-1}	maar	$(3{,}2 \pm 0{,}5)$ ms^{-1}
Niet (2203 ± 200) s	maar	$(22 \pm 2)\cdot 10^2$ s

$v = (3.610 \pm 0.013)$ ms^{-1} = $(3610 \pm 13)\cdot 10^{-3}$ ms^{-1}