

H7: confidentieniveaus

7.1 confidentieniveaus in beschrijvende statistiek

7.1.1 waarde

voorbeeld confidentie	<p>stel dat gewicht van zakken rijst Gaussisch verdeeld is met gemiddelde 520g en $\sigma=10\text{g}$</p> <ul style="list-style-type: none"> > 68% vd zakken ligt tss 510g en 530g > de <i>confidentie</i> dat een zak een gewicht tss 510g en 530g heeft is dus 68% <p>>> hoe groter de confidentie, hoe groter het interval van gewicht</p> <div style="display: flex; justify-content: space-around;"> kleiner kleiner </div> <p>> Typische waarden: 68% (1σ), 95.4% (2σ), 90% (1.64σ), 95% (1.96σ) en 99% (2.58σ).</p>
stelling: σ vs CL	Voor niet-Gaussische verdelingen gaat de relatie tss confidentieniveaus en aantal σ niet langer op

7.1.2 het belang van nul-resultaten

belang van nul-resultaten	vb: er wordt een nieuw deeltje gevonden met confidentieniveau 95% > echter wanneer dit 19 andere keren wordt getest wordt niets gevonden > dus: eerste keer had het geluk een meting te doen met een toevallige fluctuatie van 1 op 20 >> het is belangrijk te weten hoe vaak een experiment geen succes oplevert
---------------------------	---

7.1.3 interval

<p>confidentieniveau en grenzen van interval</p>	<p>Eenmaal een waarde voor confidentie C gekozen is moet je keuze maken voor interval x_- en x_+ voldoen voor een bepaalde C aan:</p> $\mathcal{P}(x_- \leq x \leq x_+) = \int_{x_-}^{x_+} P(x) dx = C$
--	---

def: tweezijdig interval

er zijn 3 manieren om de limieten rond het centrum te bepalen:

1. **Symmetrisch interval:** x_- en x_+ worden equidistant van de gemiddelde waarde μ gekozen:

$$x_+ - \mu = \mu - x_-$$
2. **Kortste interval:** De waarden worden zodanig gekozen dat $|x_+ - x_-|$ minimaal is.
3. **Centraal interval:** De waarschijnlijkheid dat een waarde beneden het interval bekomen wordt is even groot als de waarschijnlijkheid dat een waarde boven het interval gevonden wordt:

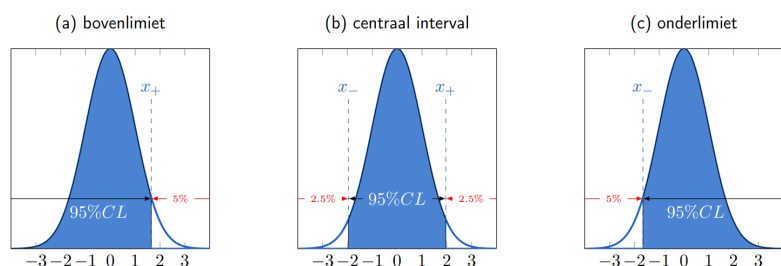
$$\int_{-\infty}^{x_-} P(x) dx = \int_{x_+}^{+\infty} P(x) dx = \frac{1 - C}{2}$$

st: Gauss en limiet rond centrum	Voor de Gauss en elke andere symmetrische verdeling zijn de 3 definities equivalent
----------------------------------	---

<p>def: bovenlimiet/onderlimiet</p>	<p>voor een bepaald confidentieniveau C ligt het gewicht vd zakken beneden de bovenlimiet:</p> $\mathcal{P}(x \leq x_+) = \int_{-\infty}^{x_+} P(x)dx = C$ <p>en het heeft geen belang wat hun gewicht is beneden deze limiet</p> <p>op zelfde manier liggen de gewichten voor bepaalde C boven de onderlimiet:</p> $\mathcal{P}(x \geq x_-) = \int_{x_-}^{+\infty} P(x)dx = C$ <p>en of ze daar nu veel of weinig boven liggen heeft geen belang</p>
-------------------------------------	---

>> oppassen op welke type interval dit is:

ie: er is nl een verschil in limiet voor bvb centraal en bovenlimiet interval:



7.1.4 discrete variabele

confidentie voor discrete vars

voor een discrete X zijn er geen x_+ zodanig dat de echte waarde van C te vinden is
> we kunnen echter wel stellen:

$$\sum_{x=x_-}^{x_+} P(x; X) \geq C$$

def: tweezijdig centraal interval

Voor een tweezijdig centraal interval zijn de boven- en onderlimiet:

$$x_- = \max \theta \quad \left| \quad \sum_{x=-\infty}^{\theta-1} P(x) \leq \frac{1-C}{2} \quad \text{en} \quad x_+ = \min \theta \quad \left| \quad \sum_{x=\theta+1}^{\infty} P(x) \leq \frac{1-C}{2} \right.$$

Dezelfde definities gelden ook voor eenzijdige intervallen, met C i.p.v. $\frac{1-C}{2}$

7.2 confidentieniveaus en schatten

7.2.1 Bayesiaanse aanpak (en zijn beperkingen)

Bayesiaanse aanpak

We willen de waarde van parameter θ bepalen

> met een apparaat met variantie $V(\theta)$ wordt een resultaat θ_m gehaald

> we willen dit omzetten in een confidentieniveau en zeggen:

θ ligt tss $\theta_m - \sigma$ en $\theta_m + \sigma$ met 68% confidentie

$\theta_m - 2\sigma$ en $\theta_m + 2\sigma$ met 95% confidentie

>> deze heeft verborgen Bayesiaanse aannames (zie vb)

vb: Bayesiaanse aanpak

je meet de massa van een e^- als $520 \text{ keV}/c^2$ met resolutie $10 \text{ keV}/c^2$

> je zou zeggen: 'massa van e^- ligt tss 510 en $530 \text{ keV}/c^2$ met confidentie van 68%'

Nu zegt theorema van Bayes: $p(m_e|m) = \frac{p(m|m_e)p(m_e)}{p(m)}$

voor experiment wist je niets over m_e , dus alle waarden zijn even waarschijnlijk

> je hebt meting m met resolutie σ dus:

$$p(m|m_e) \propto e^{-(m-m_e)^2/2\sigma^2}$$

pas Bayes toe om dit om te keren en vindt:

$$p(m_e|m) \propto e^{-(m_e-m)^2/2\sigma^2}$$

>> dit doe je onbewust bij Bayesiaanse aanpak

> maakt gebruik van **aanname van een initiële uniforme verdeling**

7.2.2 confidentie intervallen opbouwen

confidentiegebied

Stel voor een parameter θ met wss.verdeling voor de meting x : $P(x, \theta)$

> bij gewone meting met resolutie σ is dit een Gauss met gemm. $\theta = \mu$ en standaard afw. σ

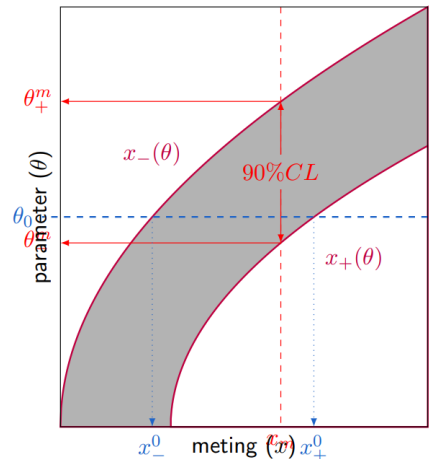
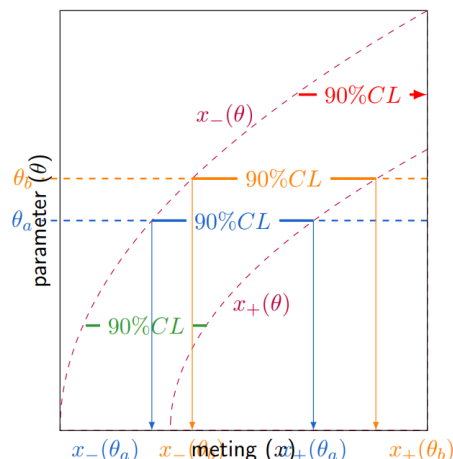
normaal gezien zal de verdeling een piek vertonen in de buurt van θ en dan afnemen

> hieruit kunnen we een confidentie-interval construeren

Voor een andere waarde θ_b krijg je andere grenzen

> x_- en x_+ kunnen als functies van θ beschouwd worden (zie figuur)

> gebied tss de twee curven is het confidentiegebied



7.2.3 interpretatie van confidentie diagram

interpretatie confidentie diagram

eenmaal je een meting x hebt wordt ze verticaal afgelezen

> stel je meet een waarde x_m

> x_- curve geeft bovenlimiet θ_+^m

> betekenis: als de echte waarde θ groter of gelijk is aan θ_+^m dan is de waarschijnlijkheid om een meting als x of kleiner te vinden minder dan 5% is

op zelfde manier voor x_+ curve vind je een onderlimiet θ_-

> 90% confidentie-interval voor echte waarde θ tss θ_- en θ_+

>> dit is een stelling over θ_- en θ_+

nl: stel dat θ_0 de echte waarde is van θ is en je deze meerdere malen meet

> metingen zullen 90% binnen x_-^0 en x_+^0 bepaald door θ_0

> punten in dit confidentiegebied vallen in horizontale en verticale grenzen

ie: vallen binnen $[x_-^0, x_+^0]$ en $[\theta_-^m, \theta_+^m]$

> 90% punten in $[x_-^0, x_+^0]$ hebben ook θ tussen $[\theta_-^m, \theta_+^m]$

7.2.4 confidentieniveaus voor een Gauss

confidentieniveau Gauss

Voor een meting x vh gemiddelde μ met een gekende σ worden waarden van μ_- en μ_+ voor een confidentie-interval C voor μ bepaald door:

$$\int_{x_m}^{+\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu_-)^2}{2\sigma^2}} dx = \frac{1-C}{2} = \int_{-\infty}^{x_m} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu_+)^2}{2\sigma^2}} dx$$

de vgl voor μ_- vereist dat x een bepaald aantal standaardafwijkingen boven μ_- moet liggen

> ie: μ_- ligt hetzelfde aantal σ beneden de gemeten waarde x_m :

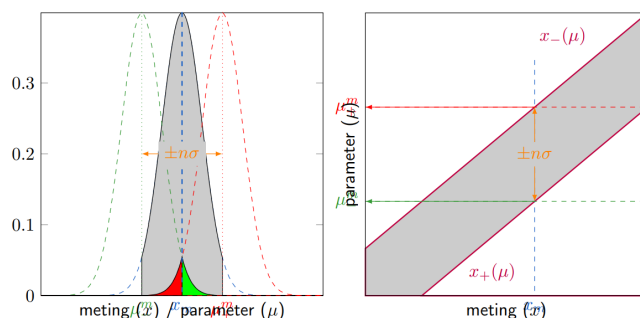
$$\int_{-\infty}^{\mu_-} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu_-)^2}{2\sigma^2}} dx = \frac{1-C}{2}$$

>> confidentielimieten kunnen voor Gaussische schatters gevonden worden uit standaard tabellen voor de Gauss integraal

> curven worden rechte lijnen met helling:

- verticaal afgelezen: $\mu_{\pm} = x_m \pm n\sigma$

- horizontaal afgelezen: $x_{\pm} = \mu \pm n\sigma$



7.2.5 Poisson confidentiegebieden

confidentieniveau Poisson

wanneer n evenementen geobserveerd worden voor Poisson met onbekend gemm. N

> dan is de 90% bovenlimiet bvb de waarde N_+ waarvoor:

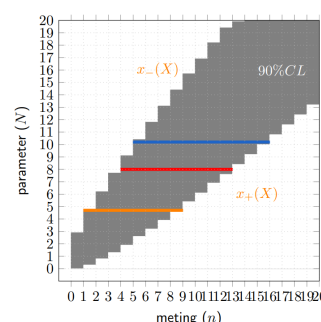
$$\sum_{r=n+1}^{\infty} P(r; N_+) = 0.90 \quad \text{of nog} \quad \sum_{r=0}^n P(r; N_+) = 0.10$$

ie: als de ware waarde N echt N_+ is, dan is de wssheid om een resultaat n te vinden dat zo klein is maar 10% en voor N groter dan N_+ is ze nog kleiner

> we zeggen dat we 90% zeker zijn dat N niet groter is dan N_+

Voor de 90% onderlimiet zoek je dan N_- waarvoor geldt:

$$\sum_{r=0}^{n-1} P(r; N_-) = 0.90$$



Figuur 7.5: 90% confidentiegebied voor Poisson.

7.3 confidentieniveaus met bijkomende voorwaarden

voorbeeld: foute conclusies

bekijk de massa ν voorwerp, deze moet altijd positief zijn

vb: neem $\theta_0 = 0,1\text{g}$ met resolutie $0,2\text{g}$

> neem als illustratie 2σ (95,4%) centrale intervallen

> meting x geeft een confidentie-interval:

$$\theta_{\pm}(x) = x \pm 0.4 \text{ g op de parameter } \theta.$$

in praktijk is deze stelling over confidentie-interval te eenvoudig

> verschillende waarden ν d metingen kunnen tot verschillende conclusies leiden:

- (a) Er is een 2.3% waarschijnlijkheid dat de meting x groter zal zijn dan $\theta_0 + 2\sigma = 0.5 \text{ g}$. Uit dergelijke meting zullen we limieten afleiden die verkeerd zijn, maar de 2.3% waarschijnlijkheid is er nu eenmaal. Het is een deel van het spel.
- (b) Als de meting x tussen 0.4 en 0.5 ligt gaan we opnieuw limieten θ_{\pm} afleiden die in dit geval wel nog net de echte waarde van 0.1 bevatten.
- (c) Als de meting x een beetje onder 0.4 ligt, bv. bij 0.3, dan worden de limieten $\theta_{-}(0.3) = -0.1$ en $\theta_{+}(0.3) = 0.7$. De onderlimiet θ_{-} kunnen we eventueel aanpassen tot $\theta_{-}^{*} = 0.0$ op basis van ons 'gezond verstand'. Deze limieten zijn opnieuw correct. En zo gaat het door tot meetwaarden van -0.3.
- (d) Beneden -0.3 ($x = \theta_0 - 2\sigma$) komen we opnieuw tot conclusies die verkeerd zijn. Deze keer omdat de bovenlimiet θ_{+} kleiner uitvalt dan $\theta_0 = 0.1$. Vermits dit gebeurt met een kleine waarschijnlijkheid van 2.3% is het aanvaardbaar.
- (e) Maar eenmaal we met een meting x van -0.5 te maken krijgen moeten we in principe een gebied van -0.9 tot -0.1 opgeven. Dit is overduidelijk absurd. Door een 95.4% niveau te gebruiken weten we dat 4.6% van onze conclusies verkeerd zullen zijn en we aanvaarden de consequenties daarvan.

>> onze aanpak van confidentieniveaus gebaseerd op frequentiedistributie leer ons niets

7.3.1 Bayesiaanse statistiek

Bayesiaanse methode

Stel geen confidentiediagram op

> bekijk de wet van Bayes:

$$p(\theta|x) = \frac{p(x|\theta)p(\theta)}{p(x)}$$

waarbij de waarde van $p(x)$ er niet toedoet en meegenomen wordt in de normering

conditionele wssheid $p(x|\theta)$ is de Gauss verdeling voor een meting x afkomstig van een echte θ met resolutie σ :

$$p(x|\theta) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\theta)^2}{2\sigma^2}}$$

$p(\theta)$ is de wssheidsverdeling van θ

> stel dat deze uniform en constant is $p(\theta)=U$

> dan krijgen we:

$$p(\theta|x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\theta)^2}{2\sigma^2}} \frac{U}{N} = p(x|\theta)$$

Nu kunnen we de kennis dat θ positief moet zijn eenvoudig meenemen

nl: de initiële $p(\theta)$ is een stapfunctie die 0 is voor $\theta < 0$ en cte voor $\theta > 0$

> dan wordt via de wet van Bayes met correcte normering:

$$p(\theta|x) = \frac{e^{-(x-\theta)^2/2\sigma^2}}{\int_0^{\infty} e^{-(x-\mu)^2/2\sigma^2} d\mu} (x \geq 0)$$

Hieruit kunnen weer confidentieniveaus afgeleid worden

> gebruik makend ν d tabellen met de Gauss integraal

>> distributie is nu asymmetrisch, zodat er een keuze moet gemaakt worden voor een symmetrisch, kortst of centraal interval

7.4 confidentieniveaus in de praktijk

7.4.1 de normale verdeling gebruiken

gevolg centraal limiet theorema: steekproefgemiddelde is benaderd normaal verdeeld ongeacht de wssheidsverd. vd gegevens
> voorwaarde dat aantal gegevens N vd steekproef groot genoeg is

a) interval voor gemiddelde met bekende variantie

st: confidentie-interval	<p>Elke waarneming heeft een gemiddelde μ en bekende resolutie σ^2 > gemm. van N steekproefgegevens X_1, \dots, X_N is Gauss met variantie σ^2/N zodat:</p> $\bar{X} \propto \mathcal{N}\left(\mu, \frac{\sigma}{\sqrt{N}}\right)$ <p>Het confidentie-interval over het gemiddelde μ ve brede steekproef van N waarnemingen met bekende resolutie σ is gegeven door:</p> $\left[\bar{X} - z_{\alpha/2} \frac{\sigma}{\sqrt{N}}, \bar{X} + z_{\alpha/2} \frac{\sigma}{\sqrt{N}} \right]$
st: voor kleine steekproeven	dit interval geldt ook voor kleine steekproeven, zolang de gegevens normaal verdeeld zijn

b) interval voor een proportie

benadering via proportie	<p>Het is ook mogelijk om een confidentieniveau te maken via proportie π ipv. gemm. μ nl: steekproefproportie \hat{P} is bij benadering normaal verdeeld > met verwachte waarde π en variantie $\frac{\pi(1-\pi)}{N}$</p>
st: Wilson score interval	<p>Het confidentie-interval over de populatieproportie π ve steekproef van N waarnemingen met gemeten proportie \hat{P} is gegeven door:</p> $\left[\frac{(2N\hat{P} + z_{\alpha/2}^2) - z_{\alpha/2} \sqrt{z_{\alpha/2}^2 + 4N\hat{P}(1-\hat{P})}}{2(N + z_{\alpha/2}^2)} ; \frac{(2N\hat{P} + z_{\alpha/2}^2) + z_{\alpha/2} \sqrt{z_{\alpha/2}^2 + 4N\hat{P}(1-\hat{P})}}{2(N + z_{\alpha/2}^2)} \right]$
toepassing Wilson score	<p>Werkt goed voor zowel grote als kleine waarnemingen en kansen > indien de ondergrens negatief is wordt ze vervangen door 0 bovengrens groter dan 1 is wordt ze vervangen door 1</p>
st: vereenvoudiging Wilson score	<p>In sommige praktische toepassingen geldt dat:</p> <ul style="list-style-type: none"> - N veel groter is dan $z_{\alpha/2}^2$ - $2N\hat{P}$ veel groter is dan $z_{\alpha/2}^2$ - $4N\hat{P}(1-\hat{P})$ veel groter is dan $z_{\alpha/2}^2$ <p>> alle contributies van $z_{\alpha/2}^2$ mogen verwaarloosd worden. > confidentie-interval vereenvoudigd tot:</p> $\left[\hat{P} - z_{\alpha/2} \sqrt{\frac{\hat{P}(1-\hat{P})}{N}}, \hat{P} + z_{\alpha/2} \sqrt{\frac{\hat{P}(1-\hat{P})}{N}} \right]$

7.4.2 de Student verdeling gebruiken

situatie student verdeling	<p>Als je een meting doet met een apparaat met bekende resolutie > dan is dit een confidentie-interval voor een Gaussische verdeling</p> <p>Vaak is de resolutie van de apparatuur echter onbekend > confidentie-interval uit vorige paragraaf niet bruikbaar</p>
----------------------------	---

a) de student verdeling

student verdeling	<p>onderzoeker kent de variantie σ^2 niet > variantie kan geschat worden nl: voor een gekende μ:</p> $\hat{\sigma} = \sqrt{(X - \mu)^2}$ <p>Als μ onbekend is gebruiken we (vgl.5.6):</p> $\hat{\sigma} = s = \sqrt{\frac{N}{N-1} (X - \bar{X})^2}$ <p>Het tweede geval komt vaker voor, maar toch niet altijd.</p> <p>ipv de variabele $Z=(X-\mu)/\sigma$ die verdeeld is volgens een eenheidsgauss > hier hebben we te maken met de variabele</p> $t = \frac{(X - \mu)}{\hat{\sigma}}$
st: eigenschappen $\hat{\sigma}$	<p>1: t is niet normaal verdeeld met variantie 1, zoals dat zou zijn als $\hat{\sigma}$ gelijk ware aan σ 2: significantie ve afwijking tss X en μ is kleiner indien $\hat{\sigma}$ gebruikt wordt ipv σ 3: voor kleine N is $\hat{\sigma}$ een slechte schatting voor σ</p>
def: t-verdeelde toevallige variabele T	$T = \frac{(X - \mu)/\sigma}{\hat{\sigma}/\sigma} = \frac{Z}{\sqrt{\chi^2/n}}$ <p>met - eenheidsgauss verdeelde variabele Z - χ^2-verdeelde variabele - aantal vrijheidsgraden n van de χ^2-verdeelde variabele > 2 verdelingen dienen onafh. te zijn van elkaar</p>
st: Student wssheidsdichtheid	<p>De waarschijnlijkheidsdichtheid van de Student verdeelde toevallige variabele t met één parameter n wordt gegeven door:</p> $f(t; n) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi} \Gamma(\frac{n}{2}) (1 + \frac{t^2}{n})^{\frac{n+1}{2}}}$
b) interval voor gemiddelde met onbekende variantie	
gemiddelde ve t-verdeling	<p>voor een t-verdeling met steekproefvariantie s^2 > het steekproefgemiddelde is benaderd normaal verdeeld met geschatte variantie:</p> $\bar{X} \propto \mathcal{N}\left(\mu, \frac{S^2}{n}\right)$ <p>met n = N of N-1 naargelang de schatter voor de spreiding S zonder (σ) of met (s) Bessel correctie gebruikt werd</p>
st: confidentie-interval Student	<p>Het confidentie-interval over het gemiddelde μ van een brede steekproef N waarnemingen met onbekende spreiding S is gegeven door:</p> $\left[\bar{X} - t_{\alpha/2; (n-1)} \frac{S}{\sqrt{n}}, \bar{X} + t_{\alpha/2; (n-1)} \frac{S}{\sqrt{n}} \right]$ <p>met n = N of N-1 naargelang de schatter voor de spreiding S zonder (σ) of met (s) Bessel correctie gebruikt werd</p>

7.4.3 de χ^2 -verdeling gebruiken

confidentie bij meerdere schatters

Als er 2+ variabelen tegelijkertijd geschat worden
 > dan is het bepalen van confidentielimieten op beide heel lastig
 > vaak: bepaal een confidentiegebied in de parameterruimte waarbinnen de echte parameters liggen met een bepaald confidentieniveau
 > komt vaak voor bij maximum likelihood methode

> voor één parameter

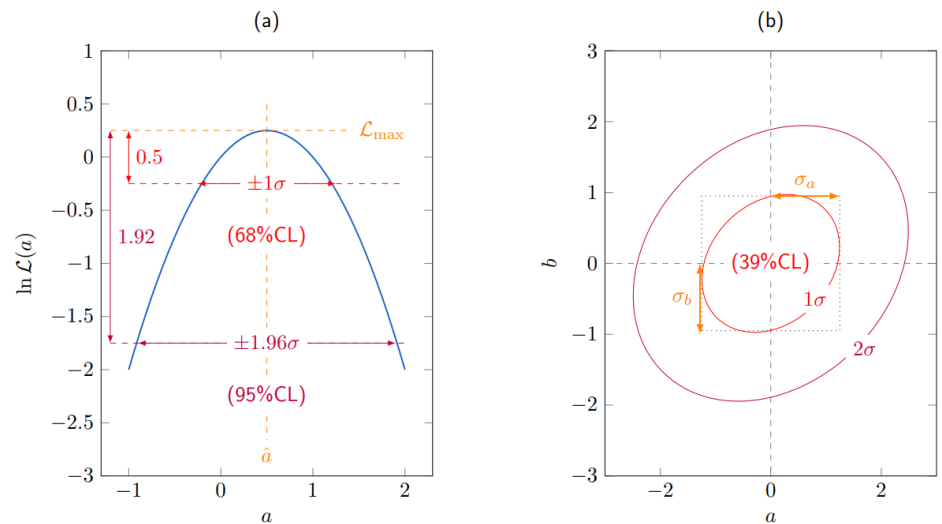
beschouw een MLLH voor 1 parameter
 > limiet voor grote N: '1 σ ' limieten voor de schatter = waarden vd parameter a waarbij de log likelihood functie in waarde 0.5 afgenomen is tov de piekwaarde
 > kunnen als Gaussische meetwaarden genomen worden

 gebied voor a waarbinnen log likelihood functie $\ln \mathcal{L}$ binnen 0.5 van zijn piekwaarde ligt is het 68% confidentie-interval
 > kan uitgebreid worden voor kleine N

> voor meerdere parameters

likelihood voor grote N:
 - 2 parameters: oppervlakte van gelijke waarschijnlijkheid zijn ellipsen
 - 3+ parameters: oppervlakte van gelijke waarschijnlijkheid zijn hyperellipsen

 men kan de (hyper)ellips waarbij $\ln \mathcal{L}=0.5$ kleiner geworden is tonen en stellen dat de echte parameterwaarden er met een zeker confidentieniveau binnen liggen
 > dit niveau is niet meer 68%
 nl: we zijn overgegaan naar een multidimensionale Gauss
 > bekijk tabel p175



Figuur 7.9: Confidentie gebied uit de maximum likelihood methode voor (a) één parameter en (b) twee parameters in de limiet van grote N.