Statistiek deel 2

Week 2

Schatters (2)

Betrouwbaarheidsintervallen

Steekproefgrootteberekeningen

Schatter voor gemiddelde μ en standaardvariatie σ

We hebben gezien dat je de verwachtingswaarde μ **van een willekeurige kansverdeling** kunt schatten met het **steekproefgemiddelde** van een aantal waarden uit die kansverdeling $x_1, \ldots, x_n$:

$$\hat{\mu} = \bar{x} = \frac{1}{n}(x_1 + \cdots + x_n)$$

Dit is een zuivere schatter.

Voor de standaarddeviatie **van een willekeurige kansverdeling** is de **steekproefstandaarddeviatie** $\hat{\sigma}$ een zuivere schatter voor de standaarddeviatie σ:

$$\hat{\sigma} = s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \hat{\mu})^2}$$

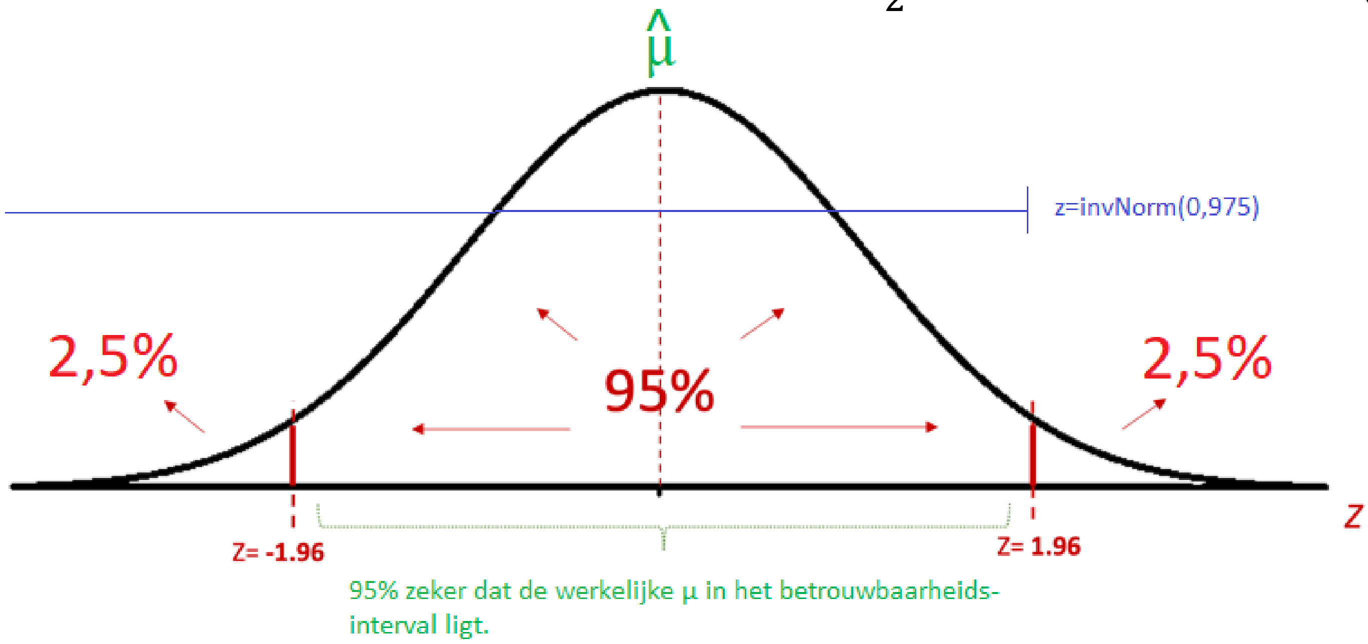
Betrouwbaarheidsinterval voor μ voor normale verdeling bij bekende σ

We zagen ook hoe je voor een normaal verdeelde variabele een betrouwbaarheidsinterval berekent voor de onbekende verwachtingswaarde μ, als daarbij de (werkelijke) standaarddeviatie σ gegeven is. Je neemt een steekproef $x_1, \dots, x_n$ en schat μ met behulp van het steekproefgemiddelde:

$$\hat{\mu} = \bar{x} = \frac{1}{n}(x_1 + \cdots + x_n).$$

Voor het betrouwbaarheidsinterval kies je eerst een onbetrouwbaarheid α (bv. 0,05, ofwel 5%), of een betrouwbaarheid $1 - \alpha$ (bv. 95%).

Daarbij hoort een z-waarde met linkeroverschrijdingskans $1 - \frac{\alpha}{2}$, dus $z = \mathrm{invNorm}\left(1 - \frac{\alpha}{2}\right)$ $(= 1{,}96)$:



Betrouwbaarheidsinterval voor μ voor normale verdeling bij bekende σ

Dit levert een (95%) **betrouwbaarheidsinterval voor $\boldsymbol{\mu}$ bij bekende $\boldsymbol{\sigma}$**:

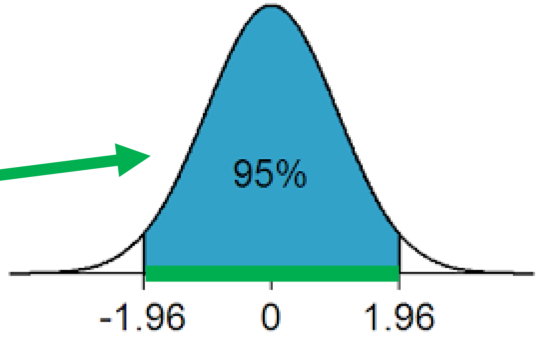
$$\left[\bar{x} - z\frac{\sigma}{\sqrt{n}}, \bar{x} + z\frac{\sigma}{\sqrt{n}}\right]$$

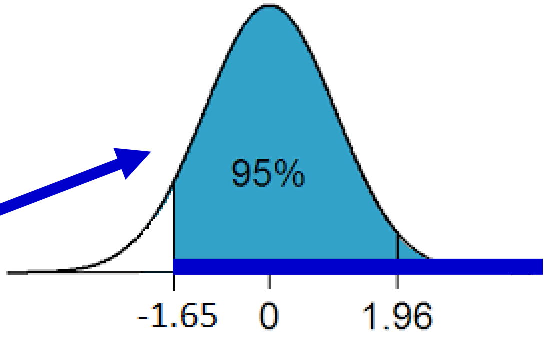
Voorspellings-/betrouwbaarheidsinterval symmetrisch of niet?

Normaal wordt een voorspellings-/betrouwbaarheidsinterval symmetrisch gekozen, bv. 95% betrouwbaarheid (in het interval), met 2,5% onbetrouwbaarheid links en 2,5% onbetrouwbaarheid rechts van het interval.

$$\bar{x} - z\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + z\frac{\sigma}{\sqrt{n}} \quad \text{met } z = \text{invNorm}(0{,}975)$$

% links	z links	% in interval	z rechts	% rechts	lengte
2,5	**-1,960**	**95**	**1,960**	**2,5**	**3,92**
3,0	-1,880	95	2,054	2,0	3,93
3,5	-1,812	95	2,170	1,5	3,98
4,0	-1,751	95	2,326	1,0	4,08
4,5	-1,695	95	2,576	0,5	4,27
4,9	-1,655	95	3,090	0,1	4,75
5	**-1,645**	**95**	∞	**0,0**	∞





In de tabel staat op elke regel een 95% betrouwbaarheidsinterval voor μ waarbij het linker interval groeit van 2,5 naar 5%.

Voor de normale verdeling is het symmetrische betrouwbaarheidsinterval het kortst qua lengte, dus het meest precies.

Het 95%-interval in de onderste regel is zoveel mogelijk naar rechts geschoven. Van boven naar beneden wordt het interval steeds groter (qua lengte) en dus onnauwkeuriger.

Voorspellings-/betrouwbaarheidsinterval symmetrisch of niet?

Voorbeeld: Een cosmeticabedrijf heeft een product ontwikkeld dat met een mascaraborsteltje op de wimpers kan worden aangebracht. Na drogen levert dit verdikte en verlengde wimperharen op.
Onbehandelde wimperhaartjes zijn gemiddeld 10 mm lang en er zitten er 100 in een wimper. Het bedrijf heeft door een gecertificeerd onderzoeksbureau vast laten stellen dat de **verlenging** in een steekproef van 10 wimperhaartjes gemiddeld $\bar{x} = 4{,}0$ mm bedraagt. De standaarddeviatie is $\sigma = 2{,}0$ mm.

Het 95% betrouwbaarheidsinterval voor de lengte is:

$$\left[4{,}0 - 1{,}96 \cdot \frac{2{,}0}{\sqrt{10}}\, , 4{,}0 + 1{,}96 \cdot \frac{2{,}0}{\sqrt{10}}\right] = [2{,}76\, , 5{,}24]$$

Oftewel, voor de verlenging geldt een percentage van

$$\left[\frac{2{,}76}{10} \cdot 100\%\, , \frac{5{,}25}{10} \cdot 100\%\right] = [27{,}6\%\, , 52{,}5\%]$$

Met 95% betrouwbaarheid ligt de verlenging van de wimperhaartjes tussen 27,6% en 52,4% ($z = 1,96$).
Het 95% betrouwbaarheidsinterval voor de ondergrens van de lengte is:

$$\left[4{,}0 - 1{,}645 \cdot \frac{2{,}0}{\sqrt{10}}\, , \infty\right] = [2{,}96\, , \infty]$$

Met 95% betrouwbaarheid is de verlenging van de wimperhaartjes minstens 29,6% ($z = 1,645$).

Voorspellings-/betrouwbaarheidsinterval symmetrisch of niet?

Op grond van deze meetresultaten kun je met betrouwbaarheidsintervallen het volgende constateren:
Met 95% betrouwbaarheid ligt de verlenging van de wimperhaartjes tussen 27,6% en 52,4% ($\boldsymbol{z = 1,96}$).
Met 95% betrouwbaarheid is de verlenging van de wimperhaartjes minstens 29,6% ($\boldsymbol{z = 1,645}$).

Het bedrijf vindt 29,6% verlenging niet zo overtuigend klinken en heeft het volgende bedacht:

Ons product maakt uw wimpers tot 77% langer.

Is deze uitspraak verantwoord bij een betrouwbaarheid van 95%?

De kans dat een haartje minstens 7,7 mm langer wordt (77% verlenging) is

$$\text{normalcdf}(7.7\,,10^{10},4.0\,,2.0) = 0{,}0322$$

De kans dat minstens één van de 100 wimperhaartjes in een wimper met minstens 77% wordt verlengd is:

$$P(\underline{k} \geq 1) = 1 - P(\underline{k} = 0) = 1 - \text{binompdf}(100\,,0.0322\,,0) = 0{,}96$$

De uitspraak van de producent is dus waar met een betrouwbaarheid van 96%! Het zegt alleen niet zoveel.

Steekproefomvang voor de normale verdeling bij gegeven σ

Het $1-\alpha$ betrouwbaarheidsinterval is

$$\left[\bar{x} - z\frac{\sigma}{\sqrt{n}}, \bar{x} + z\frac{\sigma}{\sqrt{n}}\right].$$

De gezochte waarde van μ zit (gegeven de betrouwbaarheid) in dit interval. De **onnauwkeurigheid** van het steekproefgemiddelde als schatter $\hat{\mu} = \bar{x}$ is de maximale afstand tussen $\bar{x}$ en μ: $\quad \mathrm{a} = z\frac{\sigma}{\sqrt{n}}$.

Je kunt de onnauwkeurigheid verkleinen door de steekproefgrootte te vergroten (bij gelijke betrouwbaarheid).

Bereken de steekproefgrootte n uit

$$n = \frac{z^2\sigma^2}{a^2}$$

σ is gegeven, z volgt uit de gekozen betrouwbaarheid, de onnauwkeurigheid a heb je gekozen, en daarmee kun je n berekenen.

Voorbeeld: $\sigma = 3$, betrouwbaarheid 98%, dan is $z = \mathrm{invNorm}\left(0{,}98 + \frac{1-{,}098}{2}\right) = 2{,}3263$, $a = 1{,}4$

$$n = \frac{z^2\sigma^2}{a^2} = \frac{2{,}3263^2 3^2}{1{,}4^2} = 24{,}8495, \qquad \text{dus} \quad n = 25$$

Schatters voor de normale verdeling bij onbekende μ, σ (Boek/reader 8.7)

Het komt vaker voor dat bij onbekende μ ook de σ onbekend is. Wat doe je dan?

Voor μ kun je nog steeds het steekproefgemiddelde als zuivere schatter gebruiken:

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Voor de standaarddeviatie hebben we eerder gezien dat er een zuivere schatter is:

$$\hat{\sigma} = s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \hat{\mu})^2}\,.$$

Om de berekening hiervan te vereenvoudigen kun je eventueel gebruik maken van de rekenregel

$$\sum_{i=1}^{n}(x_i - \hat{\mu})^2 = \sum_{i=1}^{n} {x_i}^2 - n\hat{\mu}^2.$$

De schatters $\hat{\mu}$ en $\hat{\sigma}$ kun je vervolgens gebruiken om de onbekende μ, σ te schatten.

Schatters voor de normale verdeling bij onbekende μ, σ Voorbeeld 8.8 uit Buijs

Een beleggingsmaatschappij weet van acht willekeurige beleggers hoeveel zij willen beleggen in een nieuw op te richten fonds (in €1000): 6,7 8,5 11,8 9,4 7,6 9,2 10,0 en 8,8. Het steekproefgemiddelde is dan

$$\hat{\mu} = \frac{1}{8}(6{,}7 + 8{,}5 + 11{,}8 + 9{,}4 + 7{,}6 + 9{,}2 + 10{,}0 + 8{,}8) = 9{,}0$$

De steekproefstandaarddeviatie is:

$$s = \hat{\sigma} = \sqrt{\frac{1}{7}\left((6{,}7 - 9{,}0)^2 + (8{,}5 - 9{,}0)^2 + (11{,}8 - 9{,}0)^2 + \cdots + (10{,}0 - 9{,}0)^2 + (8{,}8 - 9{,}0)^2\right)} = 1{,}5390$$ Of met de rekenregel:

$$s = \hat{\sigma} = \sqrt{\frac{1}{7}(6{,}7^2 + 8{,}5^2 + 11{,}8^2 + 9{,}4^2 + 7{,}6^2 + 9{,}2^2 + 10{,}0^2 + 8{,}8^2 - 8 \cdot 9{,}0^2)} = 1{,}5390$$

Betrouwbaarheidsinterval voor de normale verdeling bij onbekende μ, σ

We willen een betrouwbaarheidsinterval voor μ hebben, zoals

$$\left[\hat{\mu} - z\frac{\sigma}{\sqrt{n}}, \hat{\mu} + z\frac{\sigma}{\sqrt{n}}\right],$$

waarbij de betrouwbaarheid $1-\alpha$ is en de z-waarde hoort bij linkeroverschrijdingskans $1-\frac{\alpha}{2}$, dus $z = \mathrm{invNorm}\left(1-\frac{\alpha}{2}\right)$.

Het probleem is alleen dat we σ niet kennen.

We hebben wel een steekproefstandaarddeviatie s berekend, kunnen we die niet in plaats van σ gebruiken?

Betrouwbaarheidsinterval voor μ voor normale verdeling bij bekende σ

$$\left[\bar{x} - z\frac{\sigma}{\sqrt{n}}, \bar{x} + z\frac{\sigma}{\sqrt{n}}\right] \qquad \left[\bar{x} - z\frac{s}{\sqrt{n}}, \bar{x} + z\frac{s}{\sqrt{n}}\right]$$

=NORM.INV.N(ASELECT();10;3)

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
Zes randomwaarden uit N(10 , 3)							sigma bekend			sigma onbekend gebruik s				
	1	2	3	4	5	6	gem	a	b	nu in [a,b]	s	a	b	nu in [a,b]?
1	10,22187	9,013294	14,21821	13,09941	8,744016	11,93728	11,20568	8,805179	13,60618	binnen	2,237454	9,415344	12,99602	binnen
2	7,751074	15,71936	12,33761	7,871149	8,217147	9,126518	10,17048	7,769976	12,57098	binnen	3,210369	7,601646	12,73931	binnen
3	11,088	6,703456	11,7249	5,166345	7,106865	4,969895	7,793243	5,392743	10,19374	binnen	2,92725	5,450955	10,13553	binnen
4	8,030991	7,64967	13,87062	6,209669	7,629507	10,09976	8,915037	6,514537	11,31554	binnen	2,731876	6,729082	11,10099	binnen
5	8,179286	12,52791	9,233288	4,903749	9,229958	16,37362	10,07463	7,674135	12,47513	binnen	3,935671	6,925442	13,22383	binnen
6	7,498492	15,9013	13,72508	11,22609	7,371192	11,21221	11,15573	8,755228	13,55623	binnen	3,371297	8,458128	13,85333	binnen
7	11,48249	6,846795	14,72336	8,490522	10,72353	7,992577	10,04321	7,642713	12,44371	binnen	2,87302	7,744319	12,34211	binnen
8	12,19754	13,7509	13,71343	13,3449	14,10297	8,017845	12,52126	10,12076	14,92176	buiten	2,302129	10,67917	14,36335	buiten
9	9,567964	11,22646	8,36358	16,2863	7,008039	10,56229	10,50244	8,101937	12,90294	binnen	3,213708	7,930935	13,07394	binnen
10	6,794982	7,240796	7,066758	5,764489	14,65401	3,593539	7,519096	5,118596	9,919595	buiten	3,747053	4,520829	10,51736	binnen
11	9,470847	11,27528	14,66435	12,42258	10,83662	11,42184	11,68192	9,281419	14,08242	binnen	1,748131	10,28312	13,08072	buiten
998	16,59292	11,80693	8,011349	11,41841	9,855745	6,985459	10,77847	8,377969	13,17897	binnen	3,410876	8,0492	13,50774	binnen
999	11,1663	10,65615	10,42129	10,01449	13,5776	10,78059	11,10274	8,702236	13,50324	binnen	1,271286	10,0855	12,11998	buiten
1000	13,96284	10,11844	14,13862	9,248242	12,44243	7,315389	11,20433	8,803826	13,60483	binnen	2,752512	9,001858	13,40679	binnen
									# buiten	51			# buiten	119
									frac buiten	0,051			frac binner	0,119

Betrouwbaarheidsinterval voor de normale verdeling bij onbekende μ, σ

We willen een betrouwbaarheidsinterval voor μ hebben, zoals

$$\left[\hat{\mu} - z\frac{\sigma}{\sqrt{n}}, \hat{\mu} + z\frac{\sigma}{\sqrt{n}}\right],$$

waarbij de betrouwbaarheid $1 - \alpha$ is en de z-waarde hoort bij linkeroverschrijdingskans $1 - \frac{\alpha}{2}$, dus
$z = \mathrm{invNorm}\left(1 - \frac{\alpha}{2}\right)$.

Het probleem is alleen dat we σ niet kennen.
We hebben wel een steekproefstandaarddeviatie s gevonden, kunnen we die niet in plaats van σ gebruiken?

Het antwoord is: Nee, tenzij n groot genoeg is.

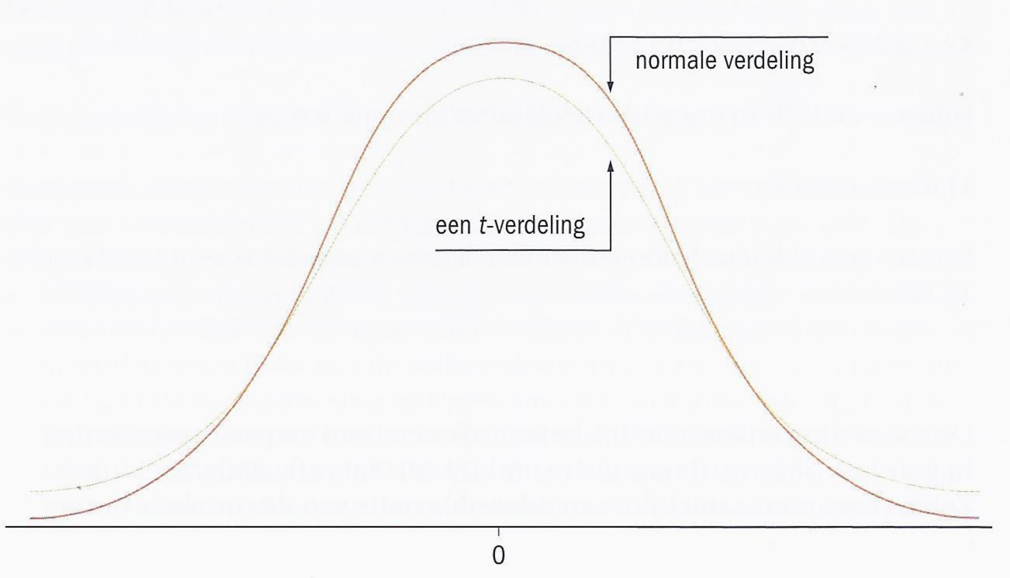
Voor onbekende σ blijkt dat als je de (geschatte) steekproefstandaarddeviatie s gebruikt i.p.v. de werkelijke σ, dat dan de geschatte $\hat{\mu}$ niet voldoet aan de normale verdeling maar aan een zogenaamde t-verdeling.

Student t-verdeling

Om dit probleem op te lossen is door William Gossett de zogenaamde **Student *t*-verdeling** bedacht. Dit is een standaardverdeling voor de ***t*-waarde**. De t-waarde is vergelijkbaar met de z-waarde van de standaard normale verdeling:

$$\underline{t} = \frac{\hat{\mu} - \mu}{\left(\frac{s}{\sqrt{n}}\right)} \quad \text{met } s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \hat{\mu})^2}{n-1}}$$

FIGUUR 8.4 Normale verdeling en een *t*-verdeling



Naast t heeft de verdeling nog een extra parameter, namelijk $v = n - 1$, het **aantal vrijheidsgraden** (DF = degrees of freedom) van de verdeling. De t-verdelingen lijken erg op de normale verdeling maar corrigeren voor het feit dat je niet de exacte standaarddeviatie kent maar de steekproefstandaarddeviatie als benadering daarvan gebruikt. De t-verdelingen zijn daardoor wat breder maar de verschillen met de standaard normale verdeling worden steeds kleiner naarmate n groter wordt.

Praktisch gezien pas je de t-verdeling toe als $n \leq 30$. Daarboven mag je ook gewoon de normale verdeling toepassen en neem je aan dat $\sigma = s = \hat{\sigma}$. Maar de t-verdeling werkt altijd.

Betrouwbaarheidsinterval voor de normale verdeling bij onbekende μ, σ

We kunnen nu het betrouwbaarheidsinterval voor μ opschrijven:

$$\left[\bar{x} - t\frac{s}{\sqrt{n}}, \bar{x} + t\frac{s}{\sqrt{n}}\right]$$

Voor de berekening van t wordt de t-verdeling gebruikt.
Voor voorbeeld 8.8 kunnen we hiermee een 95% betrouwbaarheidsinterval berekenen:
We hebben $n = 8$ steekproefwaarden. Het aantal vrijheidsgraden is $8 - 1 = 7$.
We hebben al berekend dat $\bar{x} = 9{,}0$ en $s = 1{,}5390$.

De t-waarde moet horen bij betrouwbaarheid van 95%, dus een linkeroverschrijdingskans van 0,975:

$$\mathrm{P}(\underline{t} < t) = \mathrm{tcdf}(-10^{10}, t?\ , 7) = 0{,}975$$

Los dit op met de GR solver
of met $t = \mathrm{invT}(0.975\ , 7) = 2{,}3646$.
Deze waarde is groter (onnauwkeuriger) dan de z-waarde van de normale verdeling (z = 1,96).

Dit levert als 95% betrouwbaarheidsinterval voor μ: **[7,7134 , 10,2866].**

Steekproefomvang voor de normale verdeling bij onbekende σ

Het berekenen van de steekproefomvang is wel een probleem, want we kunnen niet de formule

$$n = \frac{z^2\sigma^2}{a^2}$$

toepassen omdat σ onbekend is en voor z eigenlijk de t-waarde gebruikt moet worden, maar die heeft n nodig die nog onbekend is.

We vinden de steekproefgrootte door met de solver van de GR de volgende vergelijking op te lossen voor n (gebruik niet de methode uit het boek):

$$\mathrm{tcdf}\left(\frac{a}{s}\sqrt{n?}\ , 10^{10}, n?-1\right) = \frac{\alpha}{2}$$

Voorbeeld 8.8. De onnauwkeurigheid van het 95% betrouwbaarheidsinterval is $a = \frac{10{,}2866-7{,}7134}{2} = 1{,}2866$. Stel, je wilt een maximale afwijking van $a = 0{,}75$. Hoe groot moet dan je steekproefgrootte n zijn? Doe opnieuw, los op $\mathrm{tcdf}\left(\frac{0{,}75}{1{,}539} * \sqrt{n}\ , 10^{10}, n-1\right) = 0{,}025$. Dit levert $n = 18{,}64$, ofwel, naar boven afronden, de gewenste steekproefgrootte is $n = 19$ i.p.v. de oorspronkelijke $n = 8$ beleggers.

Schatters voor de Bernoulli/binomiale verdeling (Boek/reader 8.3: Fractie)

De Bernouilliverdeling hoort bij een experiment met twee mogelijke uitkomsten:

1 (= succes) met kans p

0 (= falen) met kans $1 - p$

Hoe kun je nu met een steekproef de waarde van de parameter p (ofwel de fractie) schatten?
Dat is eenvoudig, **je deelt het aantal successen door het totaal aantal waarnemingen van de steekproef**.
Anders geformuleerd: Voor een steekproef k_1, …., k_n van n waarnemingen (telkens 0 of 1) is de steekproefverwachting een schatting voor de succeskans p:

$$\hat{p} = \bar{k} = \frac{1}{n}\sum_{i=1}^{n} k_i$$

De som van alle waarden $\sum_{i=1}^{n} k_i$ is namelijk precies het aantal successen en bij middelden deel je door n, het totale aantal waarnemingen.

We hebben eerder al gezien dat het steekproefgemiddelde $\bar{k}$ een zuivere schatting is voor de verwachtingswaarde μ van een willekeurige verdeling. Voor de Bernoulliverdeling is dat p, dus
$\hat{p}$ is een zuivere schatter voor de succeskans p in de binomiale verdeling.

Betrouwbaarheidsinterval voor de Bernoulli/binomiale verdeling

Als je met behulp van een steekproef een succeskans hebt geschat, bv. $p = 0{,}2857$, dat kan het zijn dat je in een steekproef van $n = 7$ twee succesgevallen hebt gevonden, of 86 succesgevallen in een steekproef van $n = 301$, maar dat maakt nogal uit voor de betrouwbaarheid van je (punt)schatting.
Het is dus beter om ook een betrouwbaarheidsinterval te bepalen.
Voor de Bernoulliverdeling is dat lastig en er bestaan tientallen varianten. In het boek wordt een combinatie van de formule van Wilson (voor n tussen 20 en 200) en van Walt (voor n boven 200) behandeld.

In plaats daarvan gaan we een andere methode gebruiken (**Clopper-Pearson**) die niet met een expliciete formule werkt, maar die altijd werkt en die je met je GR kunt uitrekenen. **Dit staat niet in het boek!**

Je hebt een steekproef gedaan van n metingen met k successen. Kies het betrouwbaarheidsniveau $1 - \alpha$.

Stap 1: Bepaal de $p = p_2$ waarvoor $P(\underline{k} \leq k) = \frac{\alpha}{2}$ ofwel los op met GR: $\text{binomcdf}(n, p?, k) = \frac{\alpha}{2}$

Stap 2: Bepaal de $p = p_1$ waarvoor $P(\underline{k} \geq k) = \frac{\alpha}{2}$ ofwel, los op: $\text{binomcdf}(n, p?, k - 1) = 1 - \frac{\alpha}{2}$

Dit kun je doen met SOLVER of INTERSECT op de GR.
Het $100(1 - \alpha)\%$ betrouwbaarheidsinterval is dan $[p_1, p_2]$.

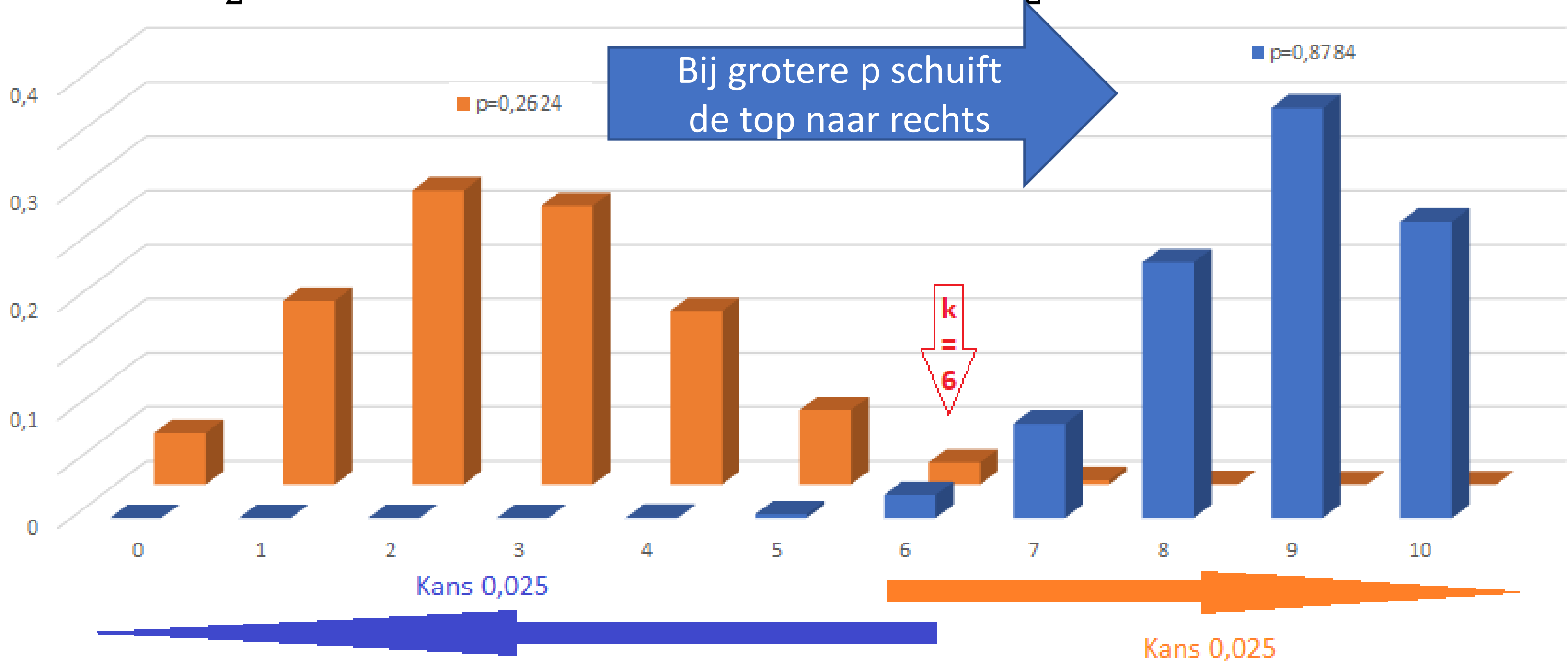
Je gebruikt de binomiale verdeling, want die geeft het aantal successen van een steekproef uit Bernoulli.

Betrouwbaarheidsinterval voor de Bernoulli/binomiaal Voorbeeld

Stap 1: Bepaal de $p = p_2$ waarvoor $P(\underline{k} \leq k) = \frac{\alpha}{2}$ ofwel: $\text{binomcdf}(n, p, k) = \frac{\alpha}{2}$

Stap 2: Bepaal de $p = p_1$ waarvoor $P(\underline{k} \geq k) = \frac{\alpha}{2}$ ofwel: $\text{binomcdf}(n, p, k - 1) = 1 - \frac{\alpha}{2}$

Voorbeeld: Je hebt een dobbelsteen die vals lijkt, want van de $n = 10$ keer gooi je $k = 6$ keer een zes. Bereken een 95% betrouwbaarheidsinterval.



1. Los $\text{binomcdf}(10, p?, 6) = \frac{0,05}{2}$ op:

$$p_2 = 0,8784$$

2. Los $\text{binomcdf}(10, p, 6 - 1) = 1 - \frac{0,05}{2}$ op: $p_1 = 0,2624$

Het 95% betrouwbaarheidsinterval is dus $[\mathbf{0,2624}\ ,\mathbf{0,8784}]$

Let op: Als je de formules uit het boek gebruikt kunnen de antwoorden enigszins afwijken van deze methode.

Het betrouwbaarheidsinterval ligt niet precies symmetrisch om de geschatte waarde van $p = k/n$.

Betrouwbaarheidsinterval voor de Bernoulli/binomiaal **Voorbeeld**

Het 95% betrouwbaarheidsinterval is dus $[0{,}2624\ ,0{,}8784]$.

Kun je met 95% zekerheid zeggen dat de dobbelsteen vals is?

Een eerlijke dobbelsteen heeft slaagkans (slagen is 6 gooien) 1/6 = 0,1667. Dit getal zit **niet** in het 95% betrouwbaarheidsinterval. De kans dat de werkelijke slaagkans buiten het 95% betrouwbaarheidsinterval ligt is 5%. Dus: Ja, met 95% zekerheid vals.

Hij is zelfs nog met 99,5% zekerheid vals.
Het 99,5% betrouwbaarheidsinterval is $[\mathbf{0,1675,0,9365}]$ (reken na).
Daar ligt 1/6 = 0,16666 net niet in.

Extra, niet voor het tentamen: Je kunt ook een enkelzijdig betrouwbaarheidsinterval $[p_1, 1]$ uitrekenen bij een betrouwbaarheid van 99,75%, ofwel $\alpha = 0{,}0003$:
Stap 2: Bepaal de $p = p_1$ waarvoor $P(\underline{k} \geq k) = \alpha$ ofwel, los op: $\text{binomcdf}(n, p?\,, k - 1) = 1 - \alpha$
Dan vind je $[\mathbf{0,1675\ ,1}]$. Ook daar ligt 1/6 = 0,16666 net niet in. De dobbelsteen is dus zelfs vals met een betrouwbaarheid van 99,75%.

Berekening van de steekproefgrootte bij Bernoulli/binomiaal (Boek 8.6)

Bij een bepaalde grootte van de steekproef en een bepaald betrouwbaarheidsniveau kunnen we nu een betrouwbaarheidsinterval uitrekenen. Dit interval heeft een bepaalde breedte. Het idee is dat de puntschatter in het midden van het interval ligt en dat het betrouwbaarheidsinterval een maximale toegestane **afwijking a** naar links en naar rechts specificeert. De breedte van het interval is dus $2a$.

Als de steekproefgrootte toeneemt dan neemt de toegestane afwijking af. Bij een vooraf gespecificeerde afwijking a en betrouwbaarheid $1-\alpha$ kun je uitrekenen hoe groot de steekproef n moet zijn om binnen die afwijking te blijven.

Omdat dat lastig is doen we deze berekening alleen voor $n \geq 200$. In dat geval is de binomiale verdeling voor het steekproefgemiddelde goed te benaderen met een normale verdeling (centrale limietstelling) met

$$\mu = \frac{np}{n} = p \quad \text{en} \quad \sigma = \frac{\sqrt{np(1-p)}}{n} = \sqrt{\frac{p(1-p)}{n}}.$$

Het betrouwbaarheidsinterval is dus

$$\left[p - z\sqrt{\frac{p(1-p)}{n}}, p + z\sqrt{\frac{p(1-p)}{n}}\right]$$

Berekening van de steekproefgrootte bij Bernoulli/binomiaal

$$\left[p - z\sqrt{\frac{p(1-p)}{n}}, p + z\sqrt{\frac{p(1-p)}{n}}\right],$$

De onnauwkeurigheid is de halve breedte van het betrouwbaarheidsinterval:

$$a = z\sqrt{\frac{p(1-p)}{n}}.$$

Deze formule hebben we ook al gezien in het geval van de normale verdeling met bekende σ. Daar was

$$a = z\sqrt{\frac{\sigma}{n}}$$

en kon je bij gegeven σ, betrouwbaarheid (z) en afwijking (a) de benodigde n uitrekenen.

Nu is dat lastiger. Je krijgt als formule voor n:

$$n = \frac{p(1-p)z^2}{a^2}.$$

Probleem: p ken je niet!

Berekening van de steekproefgrootte bij Bernoulli/binomiaal

Het probleem is alleen dat je p niet kent, want die wil je juist schatten en je kunt de schatting nog niet gebruiken, want daarvoor heb je de steekproefgrootte nodig die je hier juist wilt uitrekenen.
Je kunt dat op drie manieren oplossen:

1. Als je uit eerdere ervaring al een idee hebt hoe groot p ongeveer is (zeg $\tilde{p}$), gebruik die waarde dan: $n = \frac{\tilde{p}(1-\tilde{p})z^2}{a^2}$
2. $p(1-p)$ is een bergparabool met maximale waarde $\frac{1}{4}$ in $p = \frac{1}{2}$, dus als je $p(1-p) = \frac{1}{4}$ gebruikt krijg je een veilige waarde voor n, die nooit te klein is, maar vrijwel altijd te groot:

$$n = \frac{z^2}{4a^2}$$

3. Je kunt eventueel de gemiddelde waarde gebruiken:

$$\frac{1}{1}\int_0^1 p(1-p)dp = \left[\frac{1}{2}p^2 - \frac{1}{3}p^3\right]_{p=0}^{1} = \frac{1}{6}$$

$$n = \frac{z^2}{6a^2}$$

Deze waarde is kleiner, niet gegarandeerd goed maar minder conservatief dan 2 en werkt vaak wel.

Berekening van de steekproefgrootte bij Bernoulli/binomiaal Voorbeeld

1. Gebruik een schatting voor $\tilde{p}$ en bereken $\quad n = \frac{\tilde{p}(1-\tilde{p})z^2}{a^2}$
2. Veilige waarde voor n $\quad n = \frac{z^2}{4a^2}$
3. Gemiddelde waarde $\quad n = \frac{z^2}{6a^2}$

De waarde van n altijd naar boven afronden op een geheel getal!

Voorbeeld 8.6. Stel je wilt met 98% betrouwbaarheid en 5% onnauwkeurigheid het percentage aanhangers van de monarchie in Nederland schatten. Hoeveel personen moet je daarvoor interviewen?

We nemen onnauwkeurigheid $a = 0{,}05$, onbetrouwbaarheid $\alpha = 1 - 0{,}98 = 0{,}02$
dus $z = \mathrm{invNorm}(1 - 0{,}02/2) = 2{,}3263$.

1. Uit eerdere schattingen bleek $\tilde{p} = 0{,}65$. Je krijgt dan $n = 493$ (naar boven afronden).
2. De veilige schatting is $n = 542$.
3. De gemiddelde schatting levert $n = 361$.

Als schatting 1 realistisch is, zal schatting 3 misschien wat aan de krappe kant zijn, en schatting 2 is erg ruim. Bedenk steeds, dat dit benaderingen zijn die werken als $n \geq 200$.