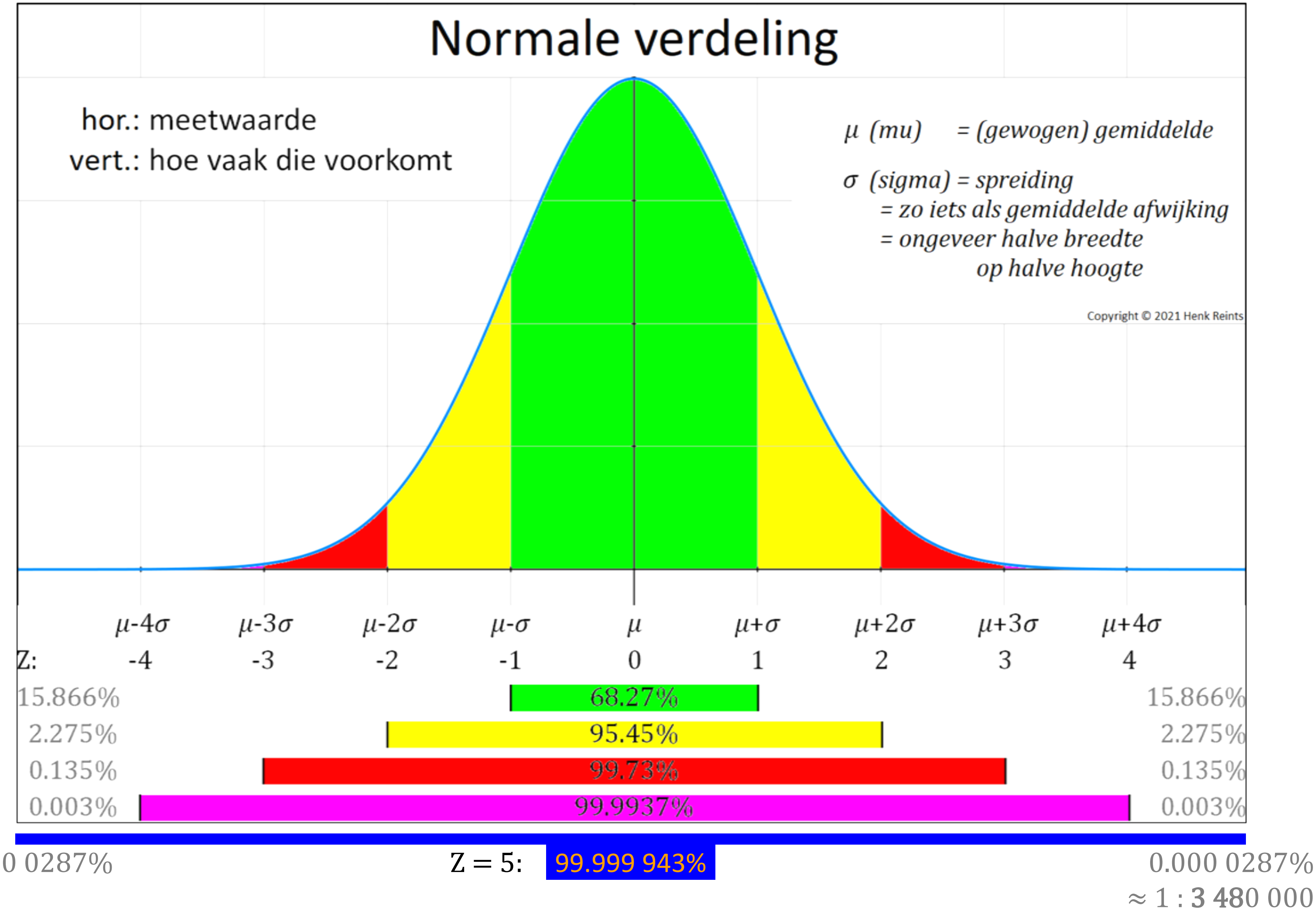
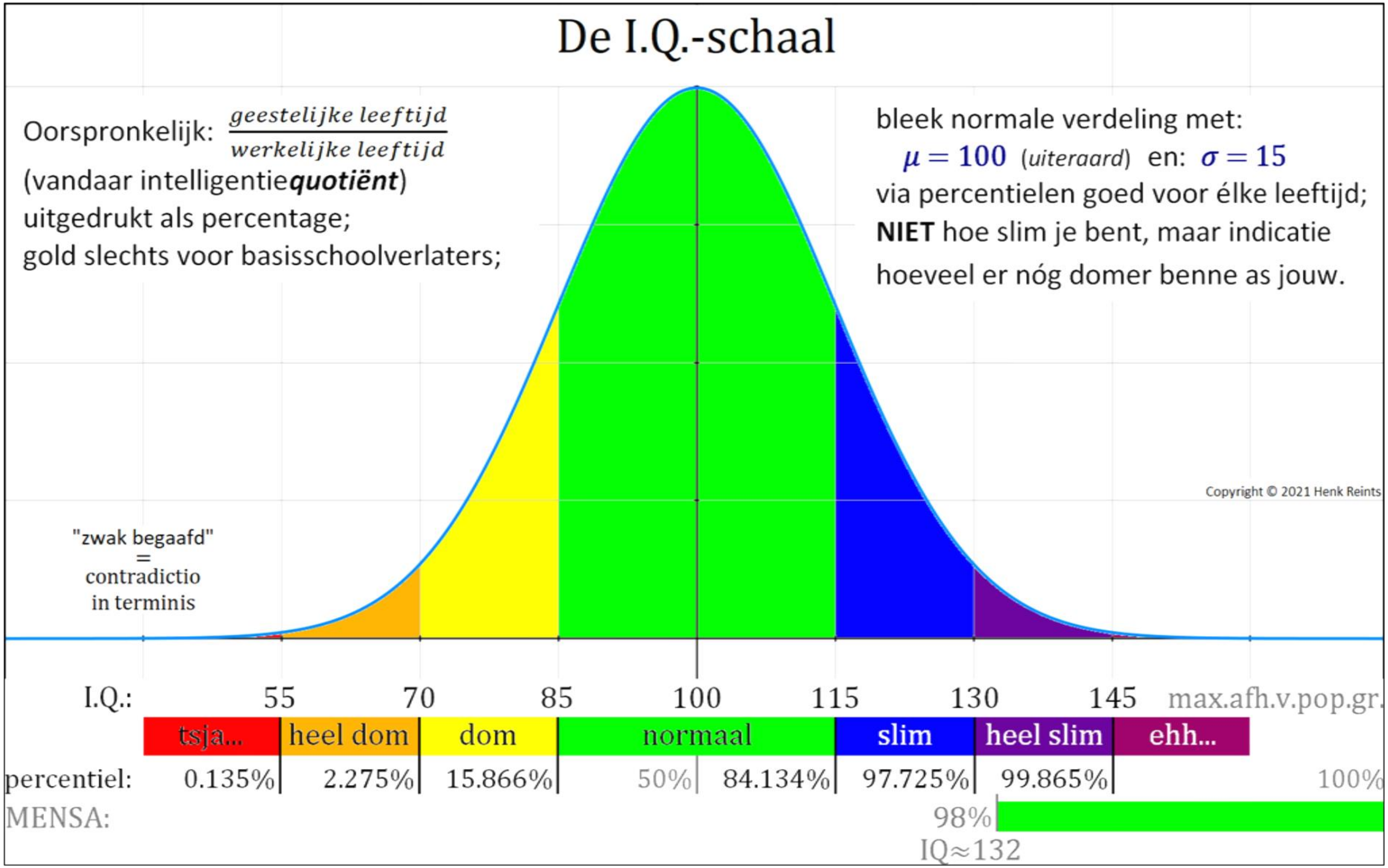
We gaan naar Australië! Konijntjes doodschieten!

Na afloop van de jacht maken we een tentoonstelling door onze trofeeën gesorteerd op gewicht op te stapelen. De lichtste en zwaarste komen uiteraard helemaal aan de linker- en rechterkant en het gemiddelde zal ergens in het midden terechtkomen.

In het midden blijkt onze uitstalling ook nog eens het hoogst te worden. De meeste konijntjes wijken kennelijk niet zo heel veel af van het gemiddelde. Hoe verder van het gemiddelde, hoe minder konijntjes, dus hoe lager de stapel aldaar.

Naarmate we meer konijntjes aan de stapel toevoegen gaat onze expositie er qua vorm steeds meer uitzien als een zogeheten *normale verdeling*, zie de volgende pagina. Bestudeer die maar eens even goed.


Normale verdeling
hor.: meetwaarde
vert.: hoe vaak die voorkomt
μ (mu) = (gewogen) gemiddelde
σ (sigma) = spreiding
= zo iets als gemiddelde afwijking
= ongeveer halve breedte
op halve hoogte
Copyright © 2021 Henk Reints
μ-4σ μ-3σ μ-2σ μ-σ μ μ+σ μ+2σ μ+3σ μ+4σ
Z: -4 -3 -2 -1 0 1 2 3 4
15.866% 68.27% 15.866%
2.275% 95.45% 2.275%
0.135% 99.73% 0.135%
0.003% 99.9937% 0.003%
0.000 0287% Z = 5: 99.999 943% 0.000 0287%
≈ 1 : 3 480 000



Ja, onze intelligentie is ook normaal verdeeld. Intelligente mensen gebruiken de I.Q.-schaal echter niet; zij werken met percentielscores.

Het I.Q. geeft aan in hoeverre je afwijkt van het gemiddelde (dat is per definitie 100), in termen van de spreiding (die is 15 en zegt hoever men gemiddeld afwijkt van het gemiddelde) en als jij dit niet snapt zit je ruim onder de 100. Het I.Q. verloopt qua percentielscores (percentage dommer dan of even slim als de proband = *beproefde*) helemaal niet lineair en met name de rechtlijnig denkende leek komt dan gemakkelijk op kromme gedachten. Aan de percentielen kun je zien dat het nogal tamelijk vrij behoorlijk redelijk zinloos is om een vermeend I.Q. van boven de $145 = \mu + 3\sigma \triangleq 99.865\%$ serieus te nemen.

Het maximaal haalbare I.Q. hangt ook nog eens af van de bevolkingsomvang! Bij de huidige wereldbevolking (7.89 miljard d.d. 6 sep. 2021) zou de allerslimste ter wereld (percentiel = 100*(N-1)/N) een I.Q. van 195 (194.87) hebben.

Dat geldt echter slechts als die bij een groot aantal tests telkens weer als enige de allerslimste van de wereld blijkt. Zhij moet wel écht de nummer één zijn.

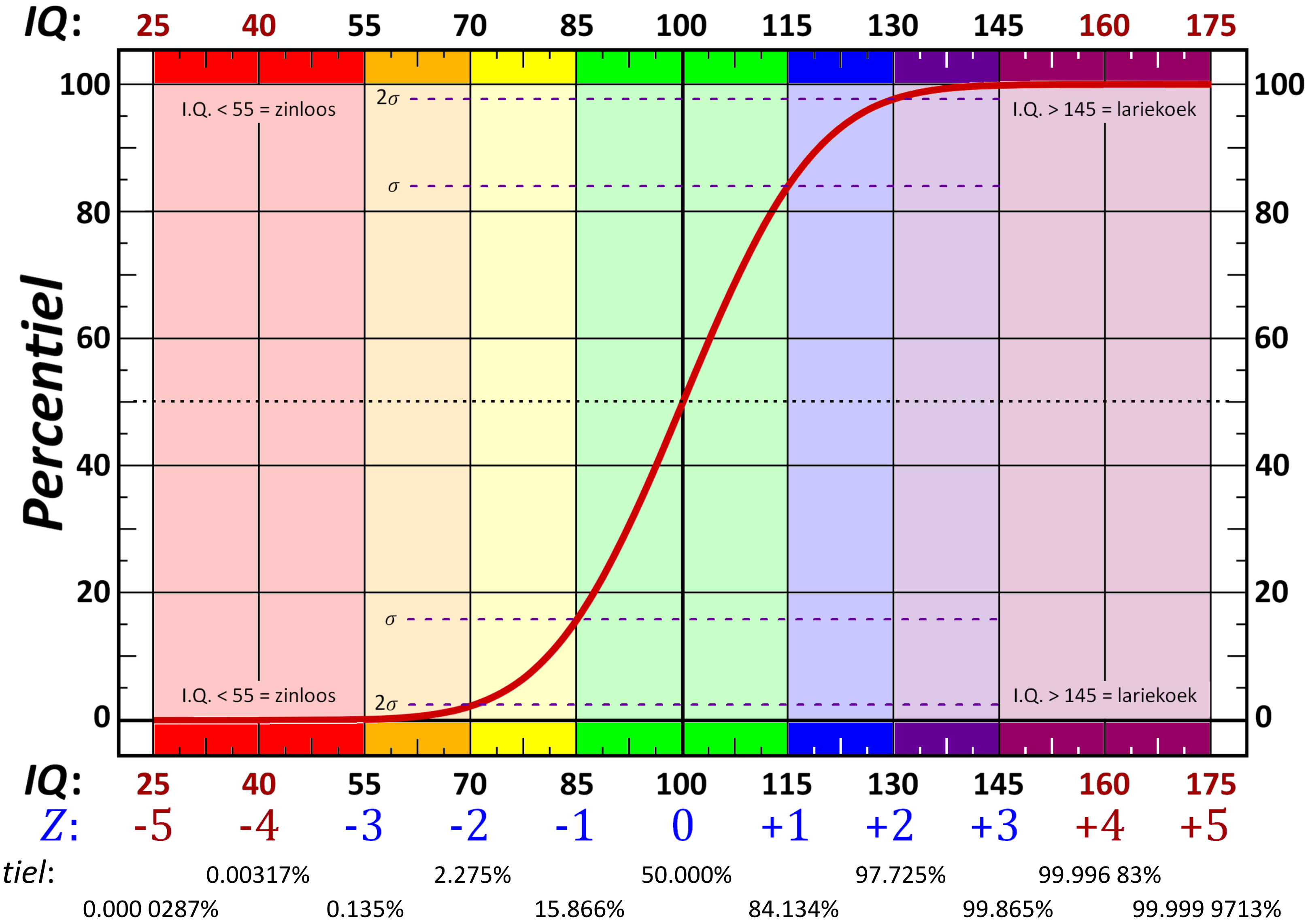
Dat *kleiner-of-gelijk* in de definitie van percentielscore impliceert dat je in geval van een ex aequo de proband laat prevaleren. Alsof die toch slimmer is. Zhíj is immers degene die getest wordt en alleen degenen die nóg slimmer zijn tellen dus niet mee.

In de praktijk zal/zullen een aantal mensen aan de top als even slim moeten worden beschouwd. Binnen de top 1‰ (1 op 1000) kun je nauwelijks onderscheid maken en dat is ook zinloos.

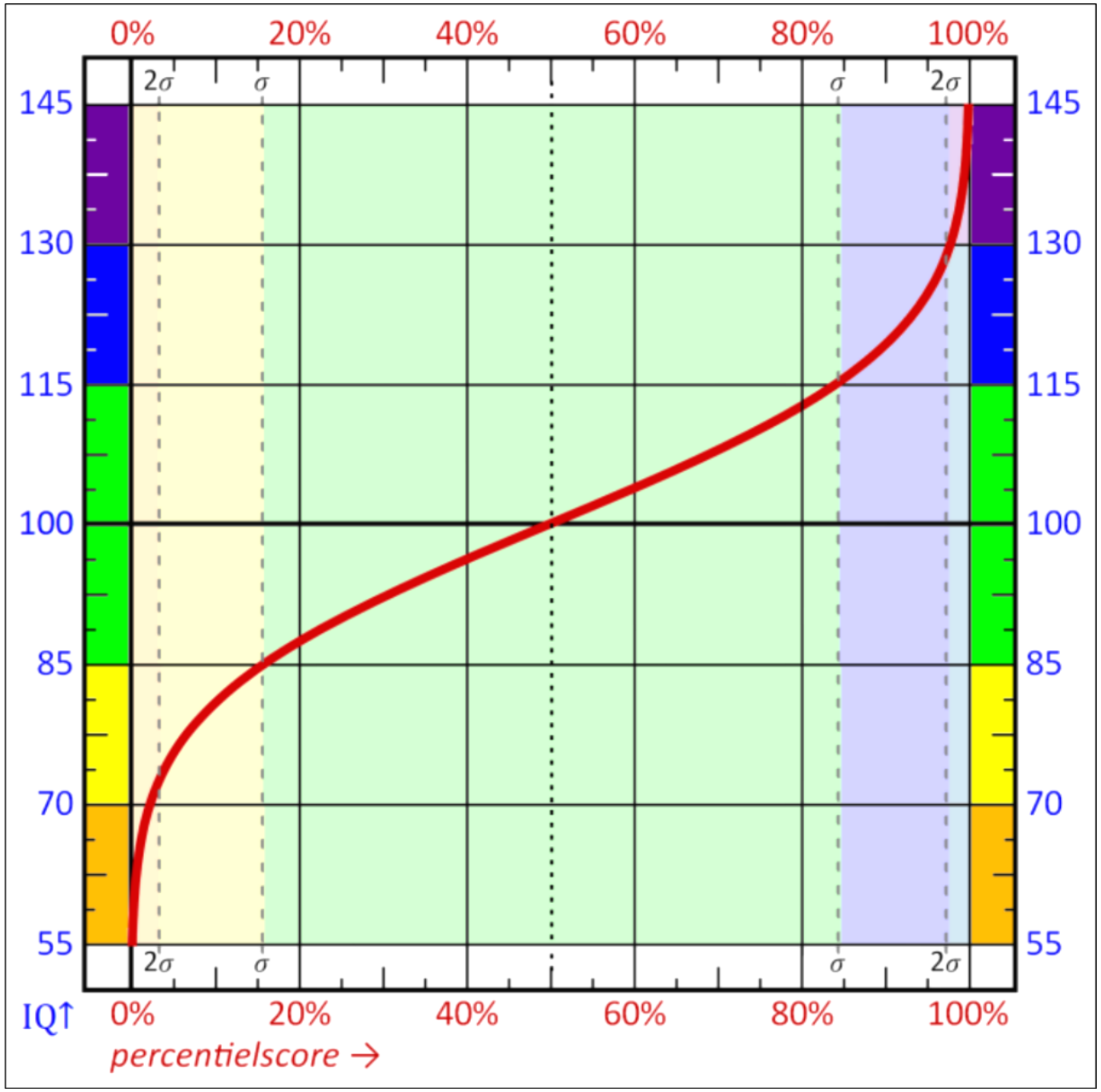
Geen enkele psycholoog kan een test ontwerpen die zijn eigen intelligentie wezenlijk te boven gaat, dus alleen die top 1‰ zélf kan de verschillen ontdekken, maar wij houden ons liever bezig met dingen die wél zinvol en verstandig zijn.

Als we deze 1‰ dus niet meenemen in de percentielberekening heeft elk toplid een percentielscore van 99.9% en daarbij hoort een I.Q. van 146. Eén en slechts één punt boven de 3σ. Een vermeend hoger I.Q. is zinloze lariekoek van een domoor met een beduidend lager I.Q.

De I.Q.-percentielomrekening gaat bij extreme I.Q.'s helemaal plat liggen en gezien vanaf de μ = 50%-stippellijn ligt de 1σ-markering vér voorbij de helft van 2σ. Niet getekend is 3σ, dat met deze lijndikte samenvalt met onder- of bovenrand, vrijwel direct volgend op 2σ. En 4σ? 5σ? Doe 's ff normaal!



Onderstaand plaatje is eigenlijk hetzelfde:



De grootste percentielverschillen zitten nu net in het gebied van $\mu \pm \sigma$ dat als *normaal begaafd* wordt aangemerkt en dat is meer dan 2/3 van de hele meute. Overigens zijn er meerdere klassificatie-systemen in de psychologie. Vaag vak.

Het moge duidelijk zijn dat iemand met I.Q. 110 nauwelijks een diepgaand gesprek kan hebben met iemand die een I.Q. van 90 heeft, of het moet over diepgaande seks gaan.

De normale verdeling is b.v. zinvol om meetnauwkeurigheden te bepalen of om te toetsen hoe groot de kans is dat iets door toeval is opgetreden.

Bij intelligentie zijn we echter vaak geïnteresseerd in de uitersten en daar is de normale verdeling dus verre van lineair, zelfs vrijwel verticaal en daar kun je niets mee.

De zinvolle I.Q.'s lopen dus van 55 t/m 145. Vooruit, van 50 t/m 150. Dat is een omvang van precies 100. Hmm, dat hebben percentielen ook, maar dan gewoon van 0 t/m 100. Dus wat maakt 't nou uit?

Wel, met I.Q.'s leveren "gewone" rekensommetjes gegarandeerd een nietszeggende verkeerde uitkomst, maar percentielen lopen keurig linear. En geen mens komt op het idee om daarmee boven de 100 te gaan. Maar zoek eens op internet naar een I.Q.-top-10. Je kunt er diverse vinden, alle met volslagen idiote waardes.

Piet Stoteles, Arie Tagoras, Euclides, Newton, Mozart, Beethoven en Ikzeidegek schitteren in al die lijsten door afwezigheid en alle lijsten gaan over I.Q.'s van 160 of veel hoger, met als absolute topscorer ene William James Sidis met een I.Q. van 250 tot 300. Ahum.

Het universum telt $2{\times}10^{11}$ (200 miljard) sterrenstelsels met gemiddeld 10^{11} sterren. Stel nu eens dat om 1 op de 100 000 sterren een planeet draait met een intelligente beschaving van gemiddeld 1 miljard individuen. Dat is heus niet zo onwaarschijnlijk, hoor! Het betekent 10^{15} sapiensen per sterrenstelsel en $2{\times}10^{26}$ in de hele kosmos.

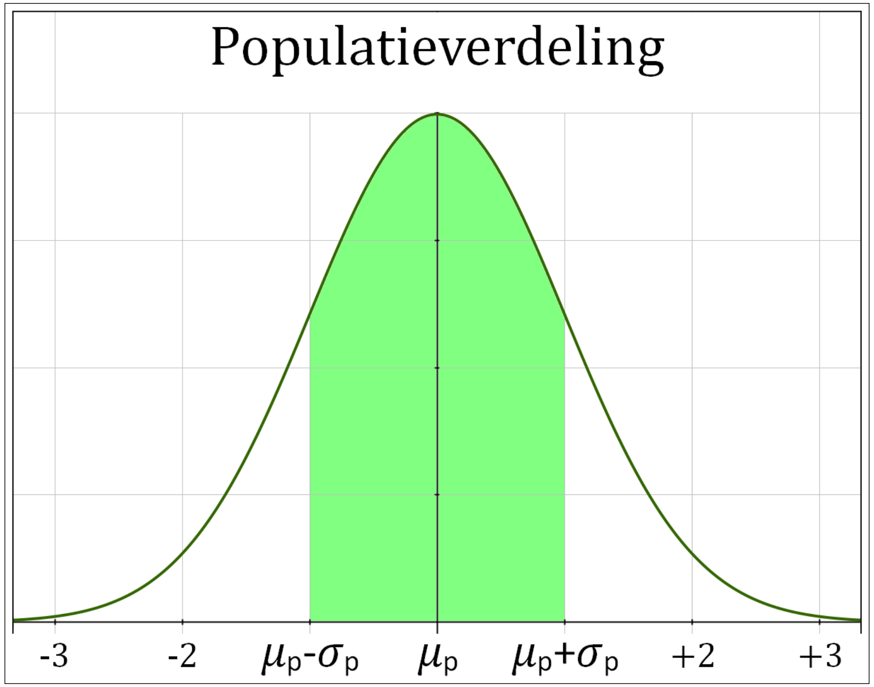
Voor elk sterrenstelsel is het maximale I.Q. dan 219 en voor het totale heelal 261.

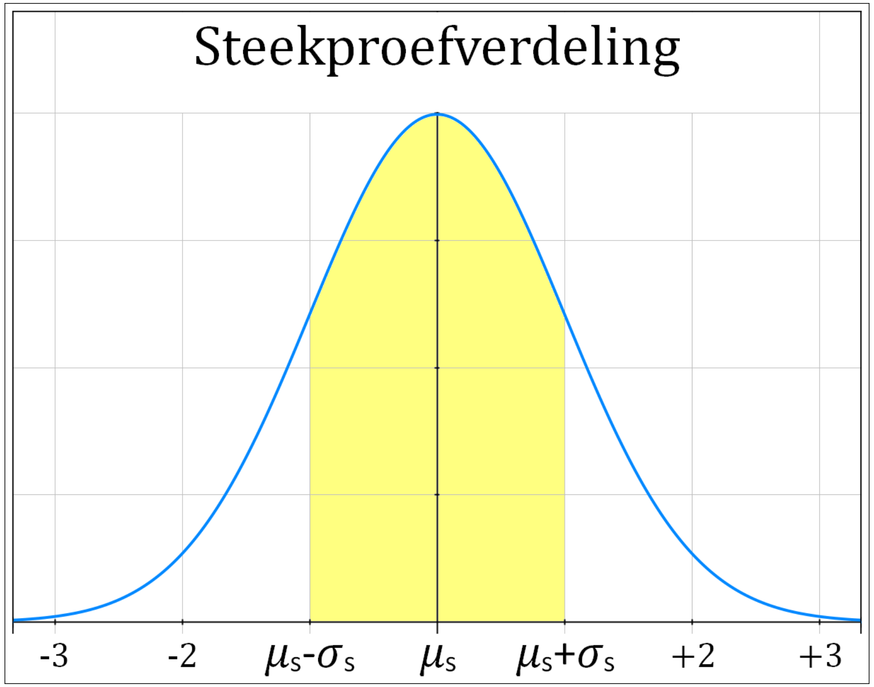
Voor een I.Q. van 250 is dan een minimumpopulatie nodig van 131 miljoen sterrenstelsels, voor I.Q. 275 reeds 53000 heelallen en voor I.Q. 300 zelfs 67.5 **b**iljoen (67 500 000 000 000) universums!

Maar als ie op die lijst staat is 't kennelijk gemeten... En die éne vent uit 53000 heelallen heeft op ónze aardkloot gewoond! Broeva! Haro!

Maar goed, we hadden dus konijntjes doodgeschoten en die arme beestjes vormen een steekproef (dat woord hebben we te danken aan de keurmeester op de kaasmarkt) van de totale Australische konijnenpopulatie.

Nu hebben uiteraard zowel de steekproef als de gehele populatie elk hun eigen gemiddelde en spreiding. Daarmee hebben we dus het *populatiegemiddelde* en de *populatiespreiding* alsmede het *steekproefgemiddelde* en de *steekproefspreiding*.

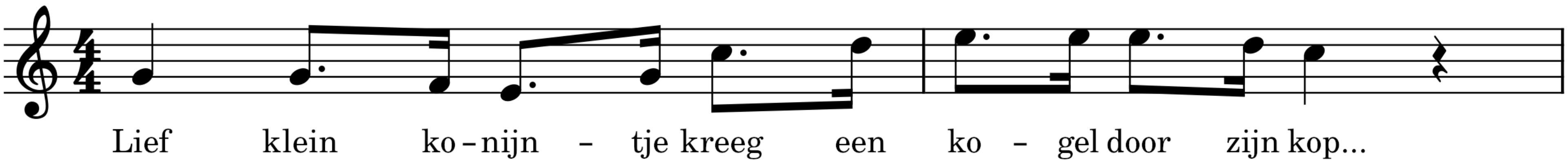




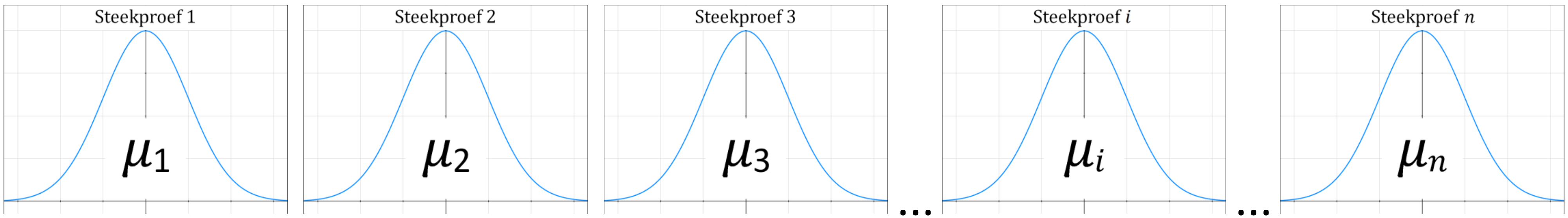
Lees eventjes goed welke symbolen eronder staan.

Nu zijn – mits we voldoende konijntjes hebben geschoten – μ_s en σ_s een goede indicatie (in de wiskunde heet dat een *schatter*) van μ_p resp. σ_p. We kunnen dus op basis van de steekproef iets zeggen over de hele populatatie.

En nu gaan we de hele vakantie onder luid gezang elke dag weer op jacht.



We schieten er telkens even veel. Aan het eind van onze vakantie hebben we dan een heleboel jachttrofeeëntentoonstellingsstapelopbouwevenementen en dan hebben we dus allemaal verschillende maar even grote steekproeven met elk hun eigen steekproefgemiddelde (en -spreiding).



Nu moet je op een hoger abstractieniveau kunnen denken.

Al die *steekproefgemiddelden* hebben zélf namelijk óók weer een gemiddelde en een spreiding. Dat zijn dus het *gemiddelde steekproefgemiddelde* en de *steekproefgemiddeldenspreiding*.

Hallo! Tuut tuut! Ben je d'r nog? Is het nog te volgen?

Laten we ze μ_μ en σ_μ noemen en de steekproefgrootte N.

Dus dan hebben we:

μ_p : het *populatiegemiddelde*;
$\boldsymbol{\sigma_p}$: de *populatiespreiding* daaromheen;
N : de *steekproefgrootte*;
n : het aantal *steekproeven* van die grootte;
$\mu_1, \mu_2, \mu_3, \ldots \mu_i, \ldots \mu_n$: de diverse *steekproefgemiddelden*;
μ_μ : het *gemiddelde steekproefgemiddelde* (som van μ_1 t/m μ_n gedeeld door n);
$\boldsymbol{\sigma_\mu}$: de *spreiding* in al die μ_i's rondom μ_μ.

En nou hebben slimmeriken het volgende uitgevogeld:

$$\sigma_\mu = \frac{\sigma_p}{\sqrt{N}}$$

de *spreiding in de steekproefgemiddelden* is gelijk aan de *populatiespreiding* gedeeld door de wortel uit de *steekproefgrootte*.

En als we nu een *steekproefgrootte* hebben van $N = 9$ vinden we:

$$\sigma_\mu = \frac{\sigma_p}{\sqrt{9}} = \frac{\sigma_p}{3}, \quad \text{ofwel:} \quad \sigma_p = 3\sigma_\mu$$

Bij *steekproefgrootte negen* past *drie* keer de *spreiding in de steekproefgemiddelden* precies in de *populatiespreiding*.

En kijk nu nog eens naar de afbeelding op blz. 2. Daar staat dat 99.73% binnen 3σ valt. Da's dus praktisch de hele zwik.

Dat betekent dat bij *steekproefgrootte negen* iedere steekproef (met een waarschijnlijkheid van 99.73%) binnen de *populatiespreiding* valt.

Je kunt *negen* dus zien als de minimale *steekproefgrootte* die nodig is om enigszins betrouwbare uitspraken te kunnen doen over een populatie.

Overigens moet een steekproef niet alleen maar groot genoeg zijn, maar hij moet ook nog voldoen aan de volgende voorwaarden:

1. hij moet **aselect** zijn, d.w.z. naar willekeur, dus zonder enige voorkeur uit de populatie geplukt; de groep mensen die jij (ja, jij) aanduidt als "iedereen" is wellicht een goed voorbeeld van een niet zo heel erg aselecte steekproef;
2. hij moet **representatief** zijn, d.w.z. de in de populatie voorkomende deelgroepen moeten met dezelfde percentages in de steekproef voorkomen; bij mensen gaat het dan o.a. om beroepsgroepen, opleidingsniveaus, religie en wat al dies meer zij; deze eis van representativiteit kan grote invloed hebben op de benodigde steekproefgrootte.

Een deugdelijke steekproef is bijvoorbeeld die van Stichting KijkOnderzoek.

https://kijkonderzoek.nl/onderzoek/methodologie/kijkonderzoek-in-het-kort

(screenshot door mij bewerkt om de afbeelding wat kleiner te maken)

$$\sigma_\mu = \frac{\sigma_p}{\sqrt{N}} = \frac{\sigma_p}{\sqrt{2800}} = \frac{\sigma_p}{52.92}$$

dus: $\boldsymbol{3\sigma_\mu} = \frac{\sigma_p}{17.64} < \boldsymbol{\frac{\sigma_p}{17}}$ en: $\boldsymbol{5\sigma_\mu} = \frac{\sigma_p}{10.58} < \boldsymbol{\frac{\sigma_p}{10}}$

Kans dat kijkcijfers er méér dan $\frac{1}{10}\sigma_p$ naast zitten:

(100 - 99.999 943 = 0.000 057)% = 1 op 1¾ miljoen.

Nog even terug naar de I.Q.-schaal:

schooljaar	aantal vwo-gediplomeerden	aantal NL-ers in loop van examenjaar 18 geworden	percentage vwo-dipl.	
2019/20	38 455	211 000	18.225%	uitschieter door corona?
2018/19	35 094	218 000	16.098%	
2017/18	35 221	228 000	15.448%	
2016/17	34 478	224 000	15.392%	
2015/16	32 677	223 000	14.653%	
2014/15	33 182	218 000	15.221%	
2013/14	31 816	218 000	14.594%	
2012/13	32 842	219 000	14.996%	
2011/12	32 089	226 000	14.199%	
2010/11	32 641	227 000	14.379%	
totaal	338 495	2 212 000	15.303%	

https://opendata.cbs.nl/statline/#/CBS/nl/dataset/80119NED/table?fromstatweb
https://www.cbs.nl/nl-nl/visualisaties/dashboard-bevolking/bevolkingspiramide

Dit zijn 10 ongeveer even grote steekproeven plus een meerjarige steekproef ruim groter dan 2 miljoen waaruit moge blijken dat 15.3% van de Nederlanders een vwo-diploma op zak heeft. Als we dat "afronden" naar 15.866% heeft 84.134% dus geen vwo-diploma en laat dat nou precies $\mu+1\sigma$ zijn, zie de afbeelding op blz. 3. Het is het percentiel voor een I.Q. van 115. Dat is dus het laagste I.Q. onder de gediplomeerde vwo-ers. Het hoogst haalbare is 145.