




---

# Context: een cruciale factor bij statistisch onderzoek

---

Prof. dr. Herman Callaert

Met voorbeelden tonen we in deze tekst dat context echt belangrijk is. Het eerste voorbeeld start heel eenvoudig met een grootte die iedereen denkt te kennen: het gemiddelde. De volgende voorbeelden komen uit het domein van de biostatistiek waar je de invloed van medicatie op de gezondheid bekijkt.

## 1 Een eerste voorbeeld: het gemiddelde.

In de klas van Emma zitten 22 leerlingen. Bij een toets Nederlands behaalde zij 7 punten op 10. Dit resultaat deelt de school als volgt mee:

- **niet:** Op Nederlands heeft Emma 7 op 10. De som van alle punten van de klas gedeeld door 22 is 6.
- **maar wel:** Op Nederlands heeft Emma 7 op 10. Het klasgemiddelde is 6.

Waarom gebruikt men het “statistische” woord gemiddelde terwijl “wiskundig” het gemiddelde niets anders is dan een synoniem voor “de som gedeeld door het aantal”? Waarom zeg je niet gewoon dat “de som gedeeld door het aantal” gelijk is aan 6?

Als je overstapt van “de som gedeeld door het aantal” op “gemiddelde”, dan geef je naast een wiskundige bewerking ook een **context** mee. “Het klasgemiddelde is 6” geeft voor veel mensen een gevoel van: “op deze toets is 6 een normale score”, dat is een resultaat dat je van een “gewone” leerling kan verwachten.

Het is deze **context** die niet zelden aanleiding geeft tot “een oordeel over de prestatie van Emma”.

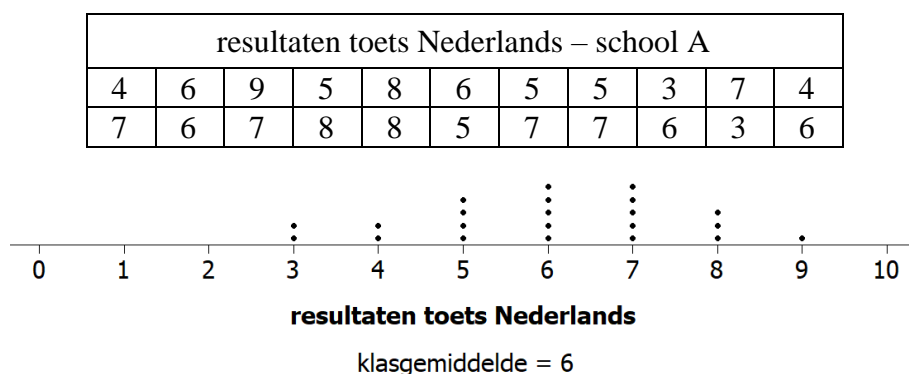
Het “gemiddelde” is een woord dat spontaan doet denken aan “de standaard”, “het typische”, “het globale”, “het centrum”, “de norm”... In veel gevallen is dit “spontane aanvoelen” een goede leidraad.

Er zijn ook heel wat situaties waar het gemiddelde een volledig verkeerd beeld geeft. Dat heeft niets te maken met “de berekening” van het gemiddelde die fout zou zijn, maar het is “de interpretatie” die helemaal de mist ingaat. Het gemiddelde heeft een context nodig. Zomaar een woord (gemiddelde) plakken op een getal (de som gedeeld door het aantal) zonder verdere context kan je ernstig op het verkeerde been zetten. Dat zie je hieronder.

Emma heeft 7 en het klasgemiddelde is 6. Emma heeft dus een behoorlijk resultaat. Is dat zo?

Je vergelijkt hier het resultaat van Emma met de resultaten van alle leerlingen in die klas. De punten van al die leerlingen vat je samen in 1 getal: het gemiddelde. Dat is handig en je kan dan gemakkelijk dat getal vergelijken met het resultaat van Emma. Maar is één getal wel een goede weergave van wat er in die klas gebeurd is?

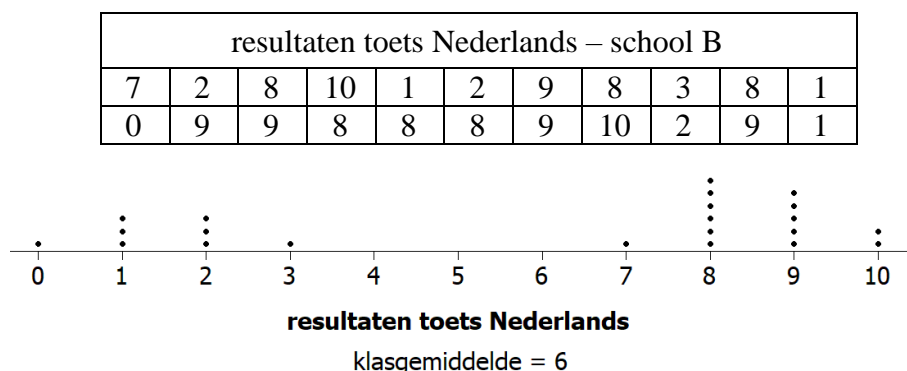
Hieronder zie je twee verschillende situaties. De resultaten zijn ook voorgesteld in een puntendiagram.



In school A haalden 2 leerlingen een 3, twee hadden een 4, er waren er 4 met een 5 enz. Het klasgemiddelde is 6 en de figuur toont dat het gemiddelde hier een goede weergave is van wat er globaal in die klas gebeurde. Meer dan de helft van de leerlingen (14 van de 22) hebben punten die niet ver van het gemiddelde afwijken (5 of 6 of 7 punten) en in de figuur zijn er geen eigenaardige patronen (zoals grote gaten of enorme extremen) te bespeuren.

Het resultaat van Emma is 1 punt meer dan het gemiddelde. Op basis hiervan (en bevestigd door het puntendiagram) zeg je dat Emma, in vergelijking met haar medeleerlingen, een goed resultaat heeft.

De resultaten in school B vertellen een ander verhaal, terwijl Emma ook hier 7 punten op 10 heeft behaald en het klasgemiddelde ook hier gelijk is aan 6.



Het puntendiagram toont dat er in deze klas iets speciaal aan de hand is. Acht leerlingen vormen een aparte groep met punten tussen 0 en 3. De overige 14 leerlingen vormen een andere groep met punten tussen 7 en 10. Het klasgemiddelde 6 is hier helemaal niet “typisch”, noch voor de ene groep (die heeft een gemiddelde van 1.5), noch voor de andere groep (met een eigen gemiddelde van 8.6). In vergelijking met deze laatste groep heeft Emma de laagste score, alle andere leerlingen hebben hier meer dan 7. Deze informatie haal je niet uit “Emma heeft 7 punten behaald en het klasgemiddelde is 6”.

*Nota.* Een startende klas met 8 nieuwe leerlingen die amper Nederlands kennen, samen met 14 leerlingen die Nederlands als moedertaal hebben, zou bij een eerste peiling een dergelijk resultaat kunnen opleveren.

*Nota.* Als je iets meer wil weten over criteria om punten in een klas te beoordelen, lees dan blz.1-4 van de tekst “Extra informatie: Standaardisatie en z-scores”. Je kan daarbij de stukjes over z-scores eventueel overslaan. Voor deze tekst ga je naar <https://www.uhasselt.be/lesmateriaal-statistiek>, klik dan op “Kansmodellen en de normale verdeling” en kijk onder “2A. Populatiemodellen”.

## 2 Voorbeelden uit de biostatistiek

### 2.1 Als je extra calcium aan je dieet toevoegt, daalt dan je bloeddruk?

De gegevens hieronder zijn een verkorte en aangepaste versie van een studie die in de USA plaats vond.

De data die je ziet, gaan over de bloeddruk (in mmHg) opgemeten bij 21 mensen. Daarbij is er een groep van 10 mensen die dagelijks een calciumsupplement namen. De andere 11 deden dat niet.

De genoteerde getallen tonen de “*daling*” van de systolische bloeddruk (de bovendruk) na 12 weken. Het resultaat 7 geeft aan dat bij die persoon de bloeddruk na 12 weken met 7 mmHg gedaald is. Als je -4 ziet staan dan is dat “*een daling van -4 mmHg*”, wat erop neerkomt dat bij die persoon de bloeddruk met 4 mmHg gestegen is.

Groep 1 (nam extra calcium): 7, -4, 18, 17, -3, -5, 6, 10, 11, -2  $\bar{x}_1 = 5.5$

Groep 2 (nam geen extra calcium): -1, 12, -1, -3, 3, -5, 5, 2, -11, -1, -3  $\bar{x}_2 = -0.3$

De bloeddruk van de eerste groep laat een gemiddelde daling van 5.5 mmHg zien. Bij de tweede groep die geen extra calcium nam is er gemiddeld zo goed als geen verschil over die periode van 12 weken (gemiddeld is er een zeer lichte stijging van 0.3 mmHg).

Om te weten of het verschil tussen de twee groepen al dan niet aan het toeval te wijten is, kan je een hypothesetoets uitvoeren. Een klassieke t-toets bijvoorbeeld leert ons dat het verschil “*statistisch significant*” is.

Wat weet je nu?

De vraag is niet of de data een significant verschil weergeven. De vraag is: helpt calcium om de bloeddruk te verlagen?

Op deze vraag kan je niet antwoorden als je alleen maar de data hebt. Je moet weten hoe die data tot stand zijn gekomen. **Je hebt context nodig.**

Enkele bedenkingen die je hierbij kan maken zien er als volgt uit:

- hoe zijn de deelnemers in deze studie terechtgekomen?
  - o zij hebben zich spontaan aangeboden na een advertentie
  - o zij namen aan de studie deel op advies van hun huisarts
  - o zij zijn lukraak gekozen uit een populatie van zwarte Amerikaanse mannen
  - o ...
- hoe zijn die 21 deelnemers in 2 groepen verdeeld?
  - o dat hebben zij zelf mogen kiezen
  - o zij zijn na een grondig cardiologisch onderzoek aan een groep toegewezen
  - o door het opgooien van een muntstuk zijn ze lukraak in een groep terechtgekomen
  - o ...
- hoe verliep de studie?
  - o sommige deelnemers die hun groep zelf hadden gekozen, veranderden ook hun levensstijl
  - o de deelnemers wisten niet welk supplement (calcium of placebo) zij kregen
  - o ...

Als je een statistisch significant verschil in bloeddruk ziet, is calcium dan **de oorzaak?**

## 2.2 Hormoontherapie en menopauze

Na de menopauze is er een sterke daling van het vrouwelijke hormoon (oestrogeen). Deze daling kan aanleiding geven tot klachten. Een behandeling kan hier HST (=hormonale substitutietherapie) zijn waarbij vrouwen extra hormonen krijgen toegediend.

Midden de jaren '90 las je in de media:

- Hormonen helpen na de menopauze (HET VOLK van 05/04/1994)
- Hormoontherapie: het levenselixir voor vrouwen? (STEM DER VROUW nr 05 van 1998)

In 2002 schreven kranten iets helemaal anders:

- Hormoontherapie na menopauze doet meer kwaad dan goed (DE STANDAARD van 10/07/2002)
- Paniek in VS over hormoontherapie na menopauze (DE MORGEN van 12/07/2002).

Wat was er gebeurd?

Het gaat hier telkens over een klinische studie waarbij (naast veel andere gezondheidsrisico's) gekeken werd naar het risico op hartinfarct bij vrouwen die na de menopauze hormoontherapie kregen. De studies staan bekend als NHS (Nurses' Health Study) en WHI (Women's Health Initiative).

### 1. The Nurses' Health Study

De universiteit van Harvard startte in 1976 een studie met 121 700 verpleegsters. Om de twee jaar vulden die verpleegsters een enquêteformulier in. Het was de bedoeling om verbanden te vinden tussen dieet, roken en lichaamsbeweging enerzijds en het risico op hartaandoening, longziekte en beroerte anderzijds.

Na een twintigtal jaren beschikte men over een massa opmetingen. Hieruit bleek dat vrouwen die na de menopauze hormoontherapie kregen een verkleind risico hadden op hartinfarct.

### 2. The Women's Health Initiative

Een andere studie, die gepland was om 8 jaar te duren, werd in juli 2002 voortijdig gestopt omdat er toen al aanwijzingen waren dat vrouwen die hormoontherapie kregen 30 % meer hartinfarcten kregen dan vrouwen die geen hormoontherapie kregen.

De studie werkte met 16 608 postmenopauzale vrouwen tussen 50 en 79 jaar. Zij werden lukraak toegewezen aan een behandeling: ofwel HST (hormonale substitutietherapie) ofwel placebo. Zij wisten niet welke behandeling zij kregen en ook hun arts was daar niet van op de hoogte.

Bij klinische studies (zoals bij elk ander statistisch onderzoek) is **de context cruciaal**. Die context was grondig verschillend bij de 2 studies hierboven.

Bij de NHS-studie werkt men met een grote groep verpleegsters die jarenlang gevolgd worden en waarbij men gewoon noteert wat er gebeurt. Verder legt het studieprotocol niets op aan de deelnemers, het enige wat men doet is observeren. Hoe weet je dan wat de echte oorzaak is van een lager percent hartinfarcten? Is het te wijten aan de hormoontherapie? Of kan het ook met andere factoren te maken hebben? Het kan bijvoorbeeld ook de levensstijl zijn. Een verpleegster kan grondig op haar gezondheid letten, gezond eten, veel bewegen, en na overleg met haar arts kiezen voor HST. Een andere verpleegster leeft erop los als een “bon vivant”. Zolang alles goed gaat mag er uitbundig gegeten en gedronken worden en aan medische ondersteuning voor nu en dan wat klachten wordt zelfs niet gedacht. Heel wat factoren (zoals de levensstijl) kunnen een invloed hebben op hartinfarcten. Je weet niet welke rol hormoontherapie hierbij speelt. Je ziet enkel een samenhang: “minder hartinfarcten bij verpleegsters die ook hormoontherapie krijgen”. Dat bewijst nog geen oorzakelijk verband.

De context van de WHI-studie is totaal anders. Het studieprotocol grijpt hier wel in en zegt op welke manier de deelnemers in groepen worden verdeeld en wat die groepen moeten doen. Door een lukraak kansmechanisme te gebruiken (denk aan het gooien van een muntstuk) creëer je twee groepen die “globaal evenwaardig” zijn. Een muntstuk weet niet wat je levensstijl is, aan welke stoffen je allergisch bent of welke bloedgroep je hebt, enz. Als er verder niets gebeurt, dan verwacht je dat bij die twee “evenwaardige” groepen het percent hartinfarcten ongeveer hetzelfde zal zijn. Die “evenwaardige” groepen worden nu op één factor verschillend gemaakt. Het studieprotocol zegt dat de ene groep hormoontherapie krijgt en de andere helemaal niet. Als er daarna bij die groepen een significant verschil is in het percent hartinfarcten, dan kan dat alleen maar te wijten zijn aan die ene factor waarop die groepen verschillen: de hormoontherapie. Zo toon je aan dat hormoontherapie de oorzaak is.

Nota. De WHI-studie werkte met vrouwen tussen 50 en 79 jaar. Latere studies tonen aan dat de leeftijd waarop hormoontherapie gestart wordt belangrijk is. Een studie met vrouwen tussen 45 en 58 jaar toont dat bij hen hormoontherapie het risico op hartinfarcten niet vergroot, integendeel.

Nota. Als je iets meer wil weten over een basisschema voor studies naar samenhang, lees dan blz.45-47 van de tekst “Studies naar samenhang: basisbegrippen”. Voor deze tekst ga je naar <https://www.uhasselt.be/lesmateriaal-statistiek>, klik dan op “Methoden en technieken bij een statistisch onderzoek” en ga naar “Studies naar samenhang”.

Statistiek = de wetenschap van het “leren uit data”  
in aanwezigheid van variabiliteit en toeval  
en **waarbij de context van een authentiek onderzoek cruciaal is.**