Statistiek voor MBW / KW
Uitwerkingen huiswerkopgaven

Dr. ir. D.A.M.P. Blom & Dr. J.B.M. Melissen

2025

Week 5: de normale verdeling en de centrale limietstelling

Hoofdstuk 5

Opdracht 5.m1: Voor een willekeurige normale verdeling geldt dat de volgende drie grootheden aan elkaar gelijk moeten zijn:

(a) gemiddelde, standaarddeviatie en mediaan.

(b) modus, de totale kans en de variantie

(c) mediaan, gemiddelde en modus

(d) standaarddeviatie, variantie en spreidingsbreedte

Uitwerking

Het juiste antwoord is (c).

Opdracht 5.m3: Voor de populatie van mannelijke economiestudenten is bekend dat het lichaamsgewicht kan worden beschreven door middel van een normaal verdeelde kansvariabele X met $\mu = 75$ kilogram en een standaarddeviatie van 6 kilogram. Hieruit volgt dat het percentage studenten van deze populatie dat een gewicht van meer dan 85 kilogram heeft ongeveer bedraagt:

(a) $95,2$

(b) 66

(c) 34

(d) $4,8$

Uitwerking

Om dit percentage te berekenen, maken we gebruik van de functie normalcdf op de grafische rekenmachine. In het bijzonder geldt namelijk dat de kans op een gewicht van meer dan 85 kilogram gelijk is aan:

$$P(X > 85) = \text{normalcdf}(a = 85; b = 10^{10}; \mu = 75; \sigma = 6) \approx 0,0478.$$

Met ongeveer $4,78\%$ kans heeft een economiestudent een gewicht van meer dan

85 kilogram. Het juiste antwoord is dus (d).

Opdracht 5.m4: Het aantal jaarlijkse reiskilometers aan woon-werkverkeer per auto is voor de werknemers van een groot concern nauwkeurig geregistreerd. Het bleek hierbij dat dit aantal goed kon worden beschreven door een normale verdeling met $\mu = 8000$ kilometer en $\sigma = 2500$. Op basis van deze verdeling kan worden vastgesteld dat de 10% werknemers die het grootste aantal woon-werkkilometers maken, per jaar per persoon meer reizen dan ongeveer ...

(a) 8000 km.

(b) 10500 km.

(c) 11200 km.

(d) 12100 km.

Uitwerking

In dit geval willen we een grenswaarde berekenen, dus moeten we gebruik maken van de functie invNorm op de grafische rekenmachine. Noteer met X de normaal verdeelde kansvariabele die het aantal woon-werkkilometers van een willekeurige werknemer beschrijft. In het bijzonder geldt dat de grenswaarde g van het aantal kilometers zodat 10% van de werknemers meer dan g km per persoon per jaar reizen gelijk is aan:

$$P(X > g) = 0,10 \rightarrow g = \text{invNorm}(opp = 1 - 0,10 = 0,90; \mu = 8000; \sigma = 2500) \\ \approx 11204\ km.$$

Het juiste antwoord is dus (c).

Opdracht 5.m5: Van een normaal verdeelde kansvariabele X is gegeven dat deze een standaarddeviatie heeft die gelijk is aan 9. Verder is gegeven dat $P(X < 15,5) = 0,3085$. De waarde van μ van deze variabele is dan ...

(a) 16

(b) 17

(c) 20

(d) 11

Uitwerking

In dit geval willen we bepalen wat de waarde van μ is zodanig dat geldt:

$$\text{normalcdf}(a = -10^{10}; b = 15,5; \mu =?; \sigma = 9) = 0,3085.$$

In de grafische rekenmachine kun je in het functiescherm het volgende invullen:

$$y_1 = \text{normalcdf}(-10^{10}; 15,5; x; 9)$$
$$y_2 = 0,3085$$

De optie intersect (2nd-TRACE-5) geeft in dat geval een snijpunt voor $x = 20,0010$. Dit betekent dat μ een waarde van ongeveer 20 heeft. Het juiste antwoord is dus (c).

Opdracht 5.2: Een vulmachine vult pakken rijst waarop staat dat daarin 1000 gram rijst zit. Omdat er altijd wat variatie zit in de afgeleverde hoeveelheden, staat de machine ingesteld op een gemiddeld vulgewicht van 1010 gram. We mogen ervan uitgaan dat de afgeleverde hoeveelheden beschreven worden door een normale verdeling met $\mu = 1010$ gram. De standaarddeviatie is echter onbekend.

(a) Bij nauwkeurig nawegen van een groot aantal pakken bleek dat in 20% van de gevallen een pak minder dan 1000 gram rijst bevat. Bereken op basis van deze informatie de standaarddeviatie van de vulgewichten.

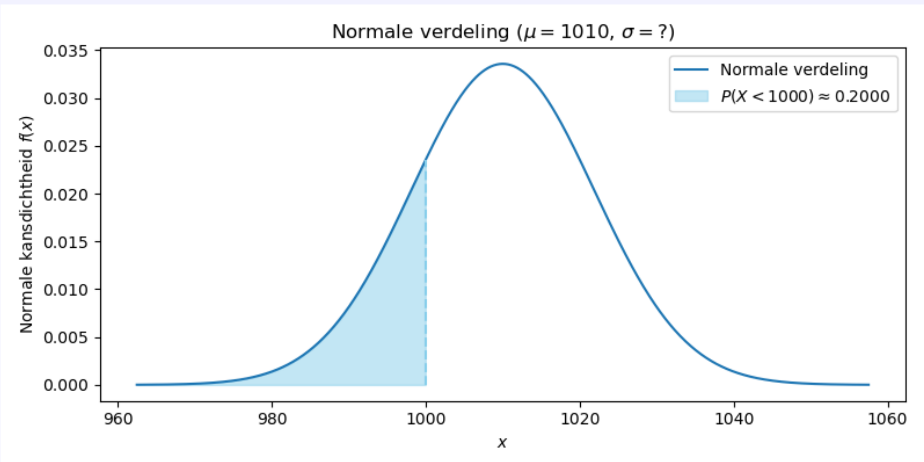
Uitwerking

Laat X de normaal verdeelde kansvariabele zijn met verwachtingswaarde $\mu = 1010$ en onbekende standaarddeviatie σ die het gewicht van een pak rijst beschrijft. In dat geval is de kansvariabele $Z = \frac{X-\mu}{\sigma}$ standaardnormaal verdeeld, oftewel $Z \sim N(0,1)$.
De z-waarde waarvoor geldt dat $P(Z < z) = 0,20$ is gelijk aan

$$z = \text{InvNorm}(opp = 0,20; \mu = 0; \sigma = 1) \approx -0,8416.$$

In andere woorden: de waarde 1000 g ligt 0,8416 standaarddeviaties onder het gemiddelde, oftewel:

$$\frac{1000 - 1010}{\sigma} = -0,8416 \to 1000 = 1010 - 0,8416 \cdot \sigma \to \sigma \approx 11,8818$$



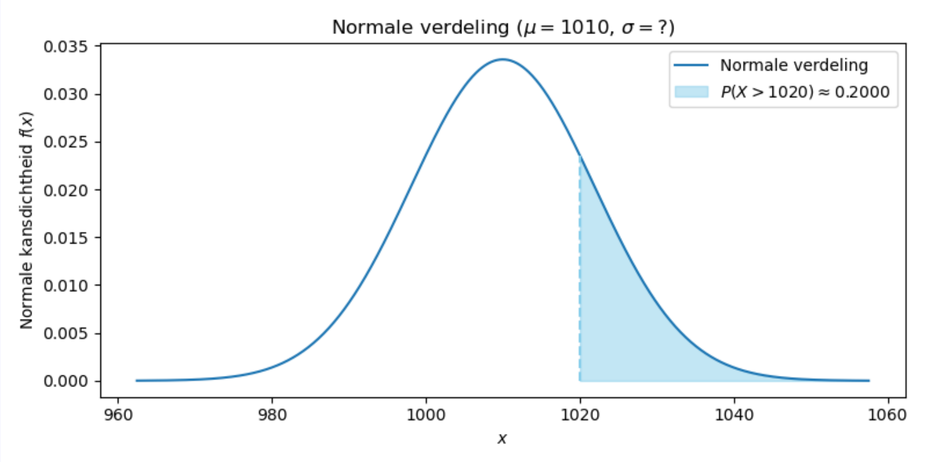
(b) Hoe groot is de kans dat een willekeurig pak rijst meer weegt dan 1020 gram?

Uitwerking

Zodra we de verwachtingswaarde μ en de standaarddeviatie σ van een normale verdeling weten, is het niet meer moeilijk om een kans te bepalen: De kans dat een willekeurig pak rijst meer dan 1020 gram weegt is gelijk aan:

$$\begin{aligned} P(X > 1020) &= \text{normalcdf}(a = 1020; b = 10^{10}; \mu = 1010; \sigma = 11,8818) \\ &\approx 0,20. \end{aligned}$$

Merk op dat deze kans hetzelfde is als de kans $P(X < 1000)$, aangezien de kansverdeling van X symmetrisch is rond $\mu = 1010$.



(c) Stel we kiezen willekeurig 16 pakken rijst. Hoe groot is de kans dat deze gemiddeld minder dan 1000 gram rijst bevatten?

Uitwerking

Laat $X_1, X_2, \ldots, X_16$ de gewichten zijn van 16 willekeurig gekozen pakken rijst. Volgens de centrale limietstelling geldt dat het gemiddelde $\overline{X} = \frac{X_1+X_2+\ldots+X_{16}}{16}$ van deze gewichten opnieuw normaal verdeeld is met verwachtingswaarde $\mu = 1010$ en standaardafwijking $\frac{\sigma}{\sqrt{16}} = \frac{\sigma}{4} \approx 2,9705$. De kans dat de 16 willekeurig gekozen pakken rijst gemiddeld minder dan 1000 gram rijst bevatten is gelijk aan:

$$\begin{aligned} P(\overline{X} < 1000) &= \text{normalcdf}(a = -10^{10}; b = 1000; \mu = 1010; \sigma = 2,9705) \\ &\approx 3,8072 \cdot 10^{-4}. \end{aligned}$$

Opdracht 5.4: Bepaal voor de standaardnormaal verdeelde variabele Z:

- $P(Z < -0,35)$
- $P(-0,16 < Z < 0,16)$
- $P(Z > -1,15)$

Uitwerking

Opdracht 5.5: Een garagebedrijf bestudeert de tijdsduur waarin de jaarlijkse servicebeurt van een auto kan worden uitgevoerd. Deze kan worden beschouwd als een normaal verdeelde variabele X met $\mu = 120$ minuten en $\sigma = 20$ minuten.

(a) Hoe groot is de kans dat een servicebeurt meer dan 150 minuten vergt?

Uitwerking

(b) Een klant komt zijn auto brengen voor service. De werkzaamheden gaan onmiddellijk beginnen. De garagehouder zegt tegen de klant: “Komt u over X minuten maar terug.”. Hoe groot moet X worden gekozen als we eisen dat de klant minder dan 5% kans wil hebben dat hij nog moet wachten als hij zich na X minuten weer meldt bij de garage? En hoe zit dat bij 1% kans?

Uitwerking

Opdracht 5.6: De gewichten van appels uit een grote partij blijken normaal verdeeld te zijn met $\mu = 100$ g en $\sigma = 20$ g. We willen deze appels in vijf gewichtsklassen verdelen die allemaal evenveel appels bevatten.

(a) Wat is de klassengrens van de 20% appels met het geringste gewicht?

Uitwerking

(b) Bepaal ook de andere klassengrenzen?

Uitwerking

Opdracht 5.10: In een kliniek wordt van een groep 50-plussers bloed afgenomen voor de bepaling van het cholesterolgehalte. Het is een bekend gegeven dat de hoeveelheid cholesterol per buisje bloed kan worden weergegeven door een normaal verdeelde variabele X met $\mu = 5,20$ mg en $\sigma = 0,9$ mg. Het 98-ste percentiel van een verdeling is het punt op de x-as waarboven zich 2% van de waarnemingen bevindt, dus waaronder 98% van de waarnemingen ligt.

(a) Welke x-waarde geeft hier het 98-ste percentiel aan?

Uitwerking

(b) Hoeveel standaarddeviaties ligt dat punt verwijderd van het gemiddelde van $5,20$ mg.

Uitwerking

(c) In een gemeente wordt een bevolkingsonderzoek gehouden naar het cholesterolgehalte. Hierbij waren 8500 personen betrokken. Mensen met een cholesterolwaarde boven het 98-ste percentiel worden voor nader onderzoek opgeroepen. Hoeveel personen krijgt naar verwachting een oproep?

Uitwerking

(d) Voor een willekeurig persoon uit de groep 50-plussers bepalen we het cholesterolgehalte. Noem deze uitkomst y. Wat is het twee-sigmagebied voor Y? en het drie-sigmagebied?

Uitwerking