6.3 Normale toetsen

Inleiding

Het toetsen van hypothesen kun je ook doen als een stochast normaal is verdeeld.
Een goed voorbeeld is de controle door de consumentenbond van het vulgewicht van kilopakken suiker. Ook daarbij speelt de significantie een grote rol. Maar bovendien wordt er vaak een steekproef getrokken met een bepaalde grootte n uit een normaal verdeelde populatie. En dan moet je met de wortel-n-wet rekening houden.

Figuur 1

Je leert in dit onderwerp

- het gemiddelde μ van een normaal verdeelde stochast toetsen;
- de wortel-n-wet gebruiken bij een steekproef van grootte n.

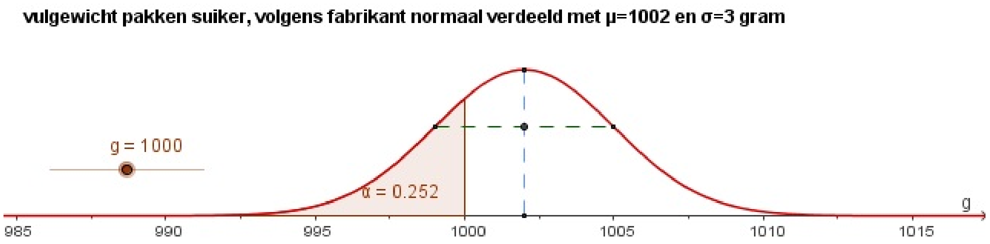
Voorkennis

- werken met binomiale toetsen;
- de begrippen nulhypothese, alternatieve hypothese en beslissingsvoorschrift en significantieniveau gebruiken.

Verkennen

Opgave V1

Bekijk de applet



Figuur 2

Volgens de fabrikant is het gewicht G (in gram) van zijn pakken suiker normaal verdeeld met $\mu(G) = 1002$ en $\sigma(G) = 3$.
Omdat de consumentenbond veel klachten heeft binnengekregen waarin wordt gemeld dat de pakken suiker van deze fabrikant te weinig suiker bevatten, wordt er door hen getwijfeld aan dit gemiddelde. De consumentenbond stelt dat $\mu(G) < 1002$.
Ze onderzoeken de bewering van de fabrikant door een pak suiker te wegen.
Ze vinden dat de fabrikant ongelijk heeft als dit pak suiker minder dan 998 gram weegt.

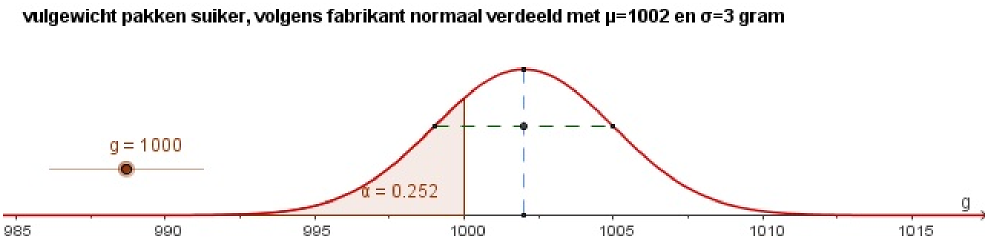
a Hoe groot is de kans dat men toevallig minder dan 998 gram suiker vindt?

b Wat zegt dit over de kans dat de consumentenbond een fout maakt?

c Hoe zou de consumentenbond dit resultaat kunnen verbeteren?

Uitleg 1

Bekijk de applet



Figuur 3

Een fabrikant beweert: het gewicht G (gram) van mijn pakken suiker is normaal verdeeld. Het gemiddelde gewicht $\mu(G) = 1002$ en de standaardafwijking $\sigma(G) = 3$.
De nulhypothese is dus:
H_0: $\mu(G) = 1002$
Een consumentenorganisatie twijfelt aan dit gemiddelde en komt met de alternatieve hypothese:
H_1: $\mu(G) < 1002$

De nulhypothese wordt getoetst door één pak suiker te wegen. Als beslissingsvoorschrift wordt genomen: H_0 wordt verworpen als dit pak suiker minder dan 998 gram weegt.

In deze situatie is de kans op onterecht verwerpen van H_0:
$P(G < 998 | \mu = 1002 \text{ en } \sigma = 3) \approx 0{,}091$

Er is dus iets meer dan 9% kans dat de consumentenorganisatie ten onrechte beweert dat de fabrikant ongelijk heeft. Het significantieniveau is dus 9%.

Opgave 1

Gebruik de gegevens uit **Uitleg 1**.

a Reken na dat de kans op de genoemde fout ongeveer 0,091 is.

b Stel dat de consumentenorganisatie vindt dat de fabrikant ongelijk heeft, als het gewogen pak suiker minder dan 997 gram weegt. Wat is nu in drie decimalen de kans op onterecht verwerpen van H_0?

c Welk bezwaar zit er aan deze toets?

Opgave 2

Bekijk **Uitleg 1**.

a Er is hier sprake van een ‘enkelzijdige normale toets op het gemiddelde’. Leg die naam uit.

b Waarom voert de consumentenorganisatie een enkelzijdige toets uit? Met wat voor soort toets zou de kwaliteitsafdeling van de fabrikant testen? Waarom?

c Wat moet de conclusie zijn als de consumentenorganisatie vooraf een significantieniveau van 5% wilde hanteren?

d Bij welk gewicht zou er bij een significantieniveau van 5% sprake zijn van een significant verschil?

Uitleg 2

Een fabrikant beweert: het gewicht G (gram) van mijn pakken suiker heeft een gemiddelde $\mu(G) = 1002$ en een standaardafwijking $\sigma(G) = 3$.
Merk op dat er NIET bij staat dat het gewicht G van de suiker normaal verdeeld is.

Toch geldt: het gemiddelde gewicht $\overline{G}$ van de pakken suiker in een steekproef van bijvoorbeeld 100 is WEL normaal verdeeld. Anders gezegd: als er veel steekproeven genomen worden, zijn de gemiddelde steekproefgewichten normaal verdeeld. Dit is de centrale limietstelling.

Om zijn bewering te kunnen toetsen moeten het gemiddelde van de steekproef $\overline{G}$ en de standaardafwijking van dit gemiddelde $\sigma\left(\overline{G}\right)$ bekend zijn. $\overline{G}$ wordt berekend uit de steekproefgegevens. $\sigma\left(\overline{G}\right)$ bereken je uit de standaardafwijking die de fabrikant opgeeft met gebruikmaking van de wortel-n-wet:

$\sigma\left(\overline{G}\right) = \frac{\sigma(G)}{\sqrt{n}}$

Als beslissingsvoorschrift wordt genomen: H_0 wordt verworpen als dit pak suiker minder dan 998 gram weegt. Hieruit volgt de overschrijdingskans:

$$\text{P}\left(\overline{G} < 998 | \mu(G) = 1002 \text{ en } \sigma\left(\overline{G}\right) = \frac{3}{\sqrt{100}}\right)$$

Met een grotere steekproef wordt de onbetrouwbaarheidsdrempel, het significantieniveau, kleiner. En de betrouwbaarheid van de toets dus groter.

Opgave 3

Bekijk **Uitleg 2**.

a Waarom kan de consumentenorganisatie met grote zekerheid H_0 verwerpen als het gemiddelde gewicht van 100 pakken suiker minder dan 998 gram is?

b Stel dat in de steekproef het gemiddelde gewicht minder dan 1000 gram is. Kan de consumentenorganisatie dan met een betrouwbaarheid van 99% de nulhypothese verwerpen?

Opgave 4

Bekijk nogmaals **Uitleg 2**. Neem nu aan de consumentenorganisatie een steekproef van 8 pakken hanteert.

a Geef de waarden van $\mu\left(\overline{G}\right)$ en $\sigma\left(\overline{G}\right)$.

b Stel dat in de steekproef het gemiddelde gewicht minder dan 1000 gram is. Wat is de conclusie als de consumentenorganisatie een significantieniveau van 1% hanteert?

Theorie en voorbeelden

Om te onthouden

Er wordt beweerd dat stochast X verdeeld is met gemiddelde $\mu(X) = \mu$.
Iemand anders vertrouwt dit gemiddelde niet en vermoedt $\mu(X) \neq \mu$.

Om dit te toetsen is een steekproef nodig met grootte n. De steekproefgemiddelden $\overline{X}$ van dergelijke steekproeven zijn altijd normaal verdeeld met $\mu\left(\overline{X}\right) = \overline{X}$ en $\sigma\left(\overline{X}\right) = \frac{\sigma(X)}{\sqrt{n}}$.

Je kunt nu vaststellen of de afwijking van $\mu\left(\overline{X}\right) = \overline{X}$ ten opzichte van $\mu(X)$ significant is.
Om deze berekening uit te kunnen voeren, moet de standaardafwijking bekend zijn.

De **toets van het (normaal verdeelde) gemiddelde** ziet er bijvoorbeeld zo uit:

- X is de gekozen variabele
- $\text{H}_0 : \mu(X) = \mu$
- $\text{H}_1 : \mu(X) \neq \mu$
- De steekproefomvang is n
- $\overline{X}$ is het berekende steekproefgemiddelde en $\sigma\left(\overline{X}\right) = \frac{\sqrt{n} \cdot \sigma}{n} = \frac{\sigma(X)}{\sqrt{n}}$ (wortel-n-wet)

Het kritieke gebied bereken je uit het gekozen **significantieniveau** α, bijvoorbeeld zo:

$$\text{P}\left(\overline{X} < g_1 \text{ of } \overline{X} > g_2 | \mu(X) = \mu \text{ en } \sigma\left(\overline{X}\right) = \frac{\sigma(X)}{\sqrt{n}}\right) = \alpha$$

Hierin zijn g_1 en g_2 de grenswaarden van het kritieke gebied.

In dit geval spreek je van een tweezijdige toets. Er is sprake van een linkszijdige toets bij het toetsen van $\mu(X) < \mu$ of van een rechtszijdige toets indien $\mu(X) > \mu$.

Het significantieniveau moet uiteraard vooraf worden gekozen.

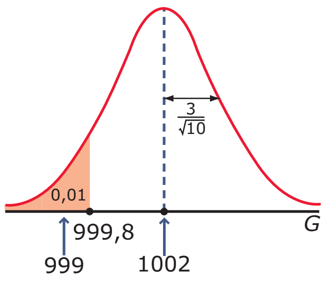
Voorbeeld 1

Bekijk de applet

Volgens de fabrikant is het gewicht G (in gram) van zijn pakken suiker normaal verdeeld met $\mu(G) = 1002$ en $\sigma(G) = 3$.

Omdat de consumentenorganisatie veel klachten heeft gekregen dat de pakken suiker van deze fabrikant te weinig suiker bevatten, wordt er door hen getwijfeld aan dit gemiddelde. De consumentenorganisatie stelt dat $\mu(G) < 1002$.

In een steekproef van 10 is het gemiddelde 999 gram. Is dit bij een significantieniveau van 1% voldoende reden om aan te nemen dat de fabrikant ongelijk heeft?



Figuur 4

Antwoord

G is het gewicht van pakken suiker in gram.

- $\text{H}_0 : \mu(G) = 1002$
- $\text{H}_1 : \mu(G) \leq 1002$
- $\overline{G} = 999$ en $\sigma\left(\overline{G}\right) = \frac{3}{\sqrt{10}} \approx 0{,}95$

$$\text{P}\left(\overline{G} < 999 | \mu = 1002 \text{ en } \sigma = \frac{3}{\sqrt{10}}\right) \approx 0{,}0008 < 0{,}01$$

Het in de steekproef gevonden gemiddelde geeft inderdaad aanleiding om de bewering van de fabrikant in twijfel te trekken bij een significantieniveau van 1%.

Je kunt dit ook bepalen door eerst het kritieke gebied te berekenen.

$$\text{P}\left(\overline{G} \leq g | \mu = 1002 \text{ en } \sigma = \frac{3}{\sqrt{10}}\right) = 0{,}01$$

Dit geeft $g \approx 999{,}79$.
Het kritieke gebied is $\overline{G} < 999{,}79$ en 999 valt binnen het kritieke gebied.

Opgave 5

Je ziet in **Voorbeeld 1** hoe de consumentenorganisatie met een steekproef van 10 pakken het gewicht van kilopakken suiker controleert.

a Reken na dat $\text{P}\left(\overline{G} < 999 | \mu = 1002 \text{ en } \sigma = \frac{3}{\sqrt{10}}\right) \approx 0{,}0008$.

b Reken na dat $g \approx 999{,}79$.

c Voer de toets nog eens uit, maar nu met een betrouwbaarheid van 99,5%. Is er nog steeds sprake van een significante afwijking?

d In plaats van een steekproef van 10 pakken wordt een steekproef van 50 pakken suiker genomen. Bij welke gewichten krijgt de consumentenorganisatie nu met 99% betrouwbaarheid gelijk?

Opgave 6

In een fabriek heeft men het vermoeden dat het koolstofgehalte van een bepaalde staalsoort groter is dan 0,200%. Uit een steekproef van 80 metingen wordt een gemiddelde gevonden van 0,213%. De standaardafwijking van het koolstofgehalte is bekend en bedraagt $\alpha = 0{,}041\%$.

a Formuleer een geschikte nulhypothese en een alternatieve hypothese.

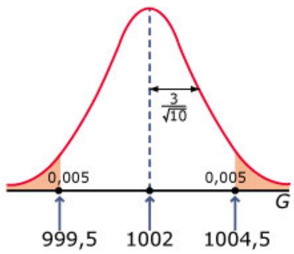
b Toets de hypothese met een onbetrouwbaarheidsdrempel van 0,01. Wat is je conclusie?

Voorbeeld 2

Bekijk de applet

Volgens de fabrikant is zijn vulmachine zo ingesteld dat het gewicht G (gram) van zijn pakken suiker normaal verdeeld is met $\mu(G) = 1002$ en $\sigma(G) = 3$.

De fabrikant test zijn vulmachine door van een steekproef van 10 pakken suiker het gemiddelde gewicht te berekenen. Hij doet (uiteraard) een dubbelzijdige toets.
Wat is het beslissingsvoorschrift bij een significantieniveau van 1%?



Figuur 5

Antwoord

- $H_0 : \mu(G) = 1002$
- $H_1 : \mu(G) \neq 1002$
- $\overline{G}$ is onbekend en $\sigma(\overline{G}) = \frac{3}{\sqrt{10}} \approx 0{,}95$

$\mathrm{P}\left(\overline{G} \leq g_1 \text{ of } \overline{G} \geq g_2 | \mu = 1002 \text{ en } \sigma = \frac{3}{\sqrt{10}}\right) = 0{,}01$ geeft:

$g_1 \approx 999{,}56$ en $g_2 \approx 1004{,}44$.
Het kritieke gebied wordt daarom $\overline{G} < 999{,}6$ of $\overline{G} > 1004{,}4$.

Opgave 7

In **Voorbeeld 2** zie je hoe je bij een tweezijdige toets te werk kunt gaan.

a Waarom is dit een tweezijdige toets? Wat gebeurt er met de onbetrouwbaarheidsdrempel α?

b Voer de beschreven toets zelf uit, maar nu met een significantieniveau van 5%.

Opgave 8

In een medisch laboratorium worden voortdurend cholesterolgehaltes in bloedmonsters bepaald. De gebruikte apparatuur wordt elk uur gecontroleerd met behulp van een ijkmonster. Hiervan is bekend dat het gemiddelde 175 mg per 100 mL zou moeten zijn. De controlemetingen aan het ijkmonster leveren op: 168, 170, 188, 170, 174, 190, 188, 171.
Is er met een significantie van $\alpha = 0{,}01$ reden om aan te nemen dat de meetapparatuur niet goed meer werkt?

Gebruik de standaardafwijking van de controlemetingen als schatting voor de standaardafwijking van de populatie.

Verwerken

Opgave 9

Een consumentenbond toetst of verpakkingen van maaltijdmixen goed zijn gevuld.

a Leg uit dat de toets links-, rechts- of tweezijdig is.

b Het gewicht G in gram van een maaltijdmix is ingesteld als $\mu(G) = 50$ en $\sigma(G) = 2$. De steekproefgrootte is 100. Wat is het kritieke gebied bij een significantieniveau van 1%?

Opgave 10

Een firma die accu's levert voor rekenmachines, beweert dat die accu's geschikt zijn om gemiddeld zo’n apparaat 8,5 weken te laten werken. De firma gaat ervan uit dat die levensduur normaal is verdeeld met een standaarddeviatie van 1,5 weken.
In een aselect gekozen groep van 60 rekenmachines doe je de accu's van deze firma. De gemiddelde levensduur van de accu's blijkt 8,2 weken te zijn. Kun je op grond van dit resultaat met een betrouwbaarheid van 97% de bewering van de firma verwerpen?

Opgave 11

Volgens een wetenschappelijk tijdschrift is het gewicht van zeventienjarigen normaal verdeeld met een gemiddelde van 54,2 kg en een standaarddeviatie van 4,7 kg. Om deze bewering te toetsen wordt door een kritische lezer door middel van een steekproef van 200 aselect gekozen zeventienjarigen het gewicht bepaald.

a Als het gemiddelde gewicht in de steekproef 53,3 kg is, heeft het tijdschrift dan met een significantie van 2,5% gelijk?

b Bij welk significantieniveau verwerp je de mening van het tijdschrift?

c Bij welk significantieniveau had je de mening van het tijdschrift verworpen als je in een veel kleinere steekproef van 10 zeventienjarigen hetzelfde gemiddelde gewicht had aangetroffen? Geef een verklaring voor het verschil met het antwoord bij b.

Opgave 12

Vacuüm verpakte vleeswaren mogen maximaal 0,022% natriumnitriet bevatten. Bij de Nederlandse Voedsel- en Warenautoriteit controleren ze dit percentage. Natriumnitriet remt de groei van de bacterie die botulisme veroorzaakt en zorgt voor behoud van de kleur van de vleeswaren. Maar een te hoog gehalte is giftig.

a Formuleer de hypothese die getoetst wordt. Neem aan dat het natriumnitrietpercentage normaal verdeeld is.

b Is de toets eenzijdig of tweezijdig? Formuleer ook de alternatieve hypothese.

Hieronder zie je 25 meetresultaten:

0,0219	0,0226	0,0225	0,0225	0,0216
0,0219	0,0220	0,0216	0,0229	0,0226
0,0214	0,0219	0,0226	0,0220	0,0212
0,0225	0,0223	0,0215	0,0221	0,0223
0,0224	0,0215	0,0228	0,0223	0,0223

Tabel 1

c Toets met behulp van deze steekproef of er reden is tot bezorgdheid. Neem een significantieniveau van 5%.

Opgave 13

Op een pak melk staat: "Het natuurlijke vetgehalte van melk - zoals die van de koe komt - varieert van 3,7% tot 4,3%".
Volle melk wordt in de fabriek altijd afgeroomd tot 3,5%. Een consumentenorganisatie besluit na te gaan of volle melk 3,5% vet bevat. In een aselecte steekproef van 20 pakken volle melk vindt ze de percentages die je in de tabel ziet. Men veronderstelt dat het vetgehalte van pakken melk normaal verdeeld is.

a Schat met behulp van deze steekproef het gemiddelde en de parameter σ van de normale verdeling.

b Toets met significantieniveau 0,05 of de consumentenorganisatie op grond van de steekproef de bewering op het pak kan verwerpen.

percentage	frequentie
3,445– < 3,455	1
3,455– < 3,465	0
3,465– < 3,475	4
3,475– < 3,485	3
3,485– < 3,495	3
3,495– < 3,505	4
3,505– < 3,515	2
3,515– < 3,525	2
3,525– < 3,535	1

Tabel 2

Toepassen

Opgave 14: Vulmachine instellen

Een fabrikant produceert een vloeistof in vaten van ruim 100 liter. Door middel van steekproeven voert de kwaliteitsdienst controles op de inhoud uit. Het nemen van steekproeven kost geld, dus worden er zo weinig mogelijk steekproeven genomen.

a De vulmachine staat ingesteld op vullen met 100,5 liter per vat. De standaardafwijking van dit vulproces is 1 liter. De kwaliteitsdienst wil zo weinig mogelijk steekproeven nemen. Maar de dienst wil ook met een significantie van hoogstens 1% weten of er gemiddeld minstens 100 liter in de vaten zit. Hoeveel vaten moet de dienst testen?

Het aantal vaten dat moet worden getest bij een significantie van 1% hangt af van de instelling van het vullen van het vat. Dit gaat met stapjes van 0,1 liter. Hoe hoger de instelling, hoe minder vaten er hoeven te worden getest, maar hoe meer het vullen kost. Het testen van een vat kost € 10,00 euro. 1 liter vloeistof kost € 1,50. De kwaliteitsdienst controleert steeds een partij van 1000 vaten door daar een aantal van te testen.

b Op welke instelling moet de fabrikant de vulmachine zetten, zodat de kosten voor steekproef nemen en vloeistof samen zo laag mogelijk zijn?

Opgave 15: Student's t-toets

Bij een normale toets van het gemiddelde gebruik je de standaardafwijking van de steekproef.
William Sealy Gosset (1876—1937) was een Engelse statisticus, die publiceerde onder het pseudoniem 'Student'. Hij heeft onderzoek gedaan naar de fout die hierdoor ontstaat, immers ook die standaardafwijkingen zullen per steekproef iets verschillen.
Hij ontdekte dat er in dat geval beter een iets andere verdeling dan de normale verdeling gebruikt kan worden. Deze verdeling wordt de Student- of t-verdeling (onder die letter t vind je hem vaak op de grafische rekenmachine) genoemd.
Voor de Studentverdeling moet de steekproefomvang bekend zijn. Als deze n is, is het aantal vrijheidsgraden $n - 1$.

a Teken met behulp van bijvoorbeeld de grafische rekenmachine de normale verdeling en de t-verdeling. Gebruik bij de normale verdeling $\mu = 0$ en $\sigma = 1$. Bij de Studentverdeling is dit vaak automatisch het geval. Neem een steekproefomvang 2.

b Leg aan de hand van deze grafieken uit dat het steekproefgemiddelde bij de Studentverdeling een grotere afwijking van het gemiddelde moet hebben om de nulhypothese te kunnen verwerpen, dan bij de normale verdeling.

c Vanaf welke steekproefomvang is te zeggen dat de Studentverdeling en de normale verdeling op elkaar lijken?

Testen

Opgave 16

In een melkfabriek worden flessen machinaal gevuld. De gedoseerde hoeveelheid per fles is normaal verdeeld. Bij juiste instelling is de verwachte hoeveelheid 250 g per fles. Een kwaliteitsinspecteur neemt een steekproef van 9 flessen en vindt voor het gemiddelde 252 g. De standaardafwijking is 2 g.
Toets, met significantieniveau $\alpha = 0{,}05$, $H_0 : \mu = 250$ g tegen $H_1 : \mu \neq 250$ g.

Opgave 17

Een partij kobaltchloride wordt bij levering door de ontvanger gekeurd op het gehalte kobalt. Volgens de leverancier is het gehalte 16,4%. Hieronder zie je de resultaten van de metingen.

16,2	15,8	16,1	15,8	15,9
15,9	16,2	16,1	16,2	16,0
15,8	15,9	16,1	15,8	16,0
16,0	16,0	15,9	16,2	16,2
16,0	16,1	16,0	15,9	16,3

Tabel 3

Toets met behulp van deze resultaten of de leverancier gelijk heeft. Je mag aannemen dat het gehalte kobalt normaal verdeeld is. Neem als significantieniveau 0,02.

Deze paragraaf is een onderdeel van het Math4All wiskundemateriaal.

Math4All stelt het op prijs als onvolkomenheden in het materiaal worden gemeld en ideeën voor verbeteringen in de content of dienstverlening kenbaar worden gemaakt. Klik op in de marge bij de betreffende opgave. Uw mailprogramma wordt dan geopend waarbij het emailadres en onderwerp al zijn ingevuld. U hoeft alleen uw opmerkingen nog maar in te voeren.

Email: `f.spijkers@math4all.nl`

Met de Math4All Foliostraat kunnen complete readers worden samengesteld en toetsen worden gegenereerd. Docenten kunnen bij `a.f.otten@math4all.nl` een gratis inlog voor de maatwerkdienst aanvragen.