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;
- werken met de normale verdeling en bijbehorende kansen berekenen;
- de wortel-n-wet gebruiken;
- de begrippen nulhypothese, alternatieve hypothese en beslissingsvoorschrift en significantieniveau gebruiken.

Verkennen

Opgave V1

Bekijk de applet: normale verdeling gewicht kilopakken suiker

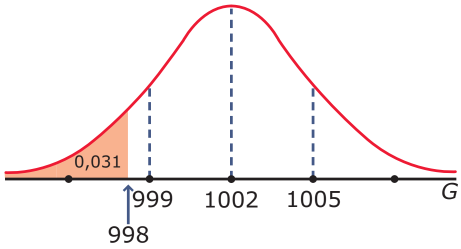
Volgens de fabrikant is het gewicht G (in gram) van zijn pakken suiker normaal verdeeld met $\mu(G) = 1002$ en $\sigma(G) = 3$.
Omdat de consumentenbond veel klachten heeft binnengekregen waarin wordt gemeld dat de pakken suiker van deze fabrikant te weinig suiker bevatten, wordt er door hen getwijfeld 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? Hoe zou de consumentenbond dit resultaat kunnen verbeteren?

Uitleg 1

Bekijk de applet: normale verdeling gewicht kilopakken suiker



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. Dit is het beslissingsvoorschrift.

Dan bestaat de kans dat dit pak toevalligerwijs minder dan 998 gram weegt, terwijl toch de fabrikant gelijk heeft. De consumentenbond zou dan moeten beweren dat de fabrikant ongelijk heeft, terwijl dat niet waar is.
Ga na, dat die kans is: $\text{P}(G < 998 | \mu = 1002 \text{ en } \sigma = 3) \approx 0{,}091$.
Er is dus meer dan 9% kans dat de consumentenbond ten onrechte beweert dat de fabrikant ongelijk heeft.

De bewering van de fabrikant is de nulhypothese H_0: $\mu(G) = 1002$.
De consumentenbond stelt als alternatieve hypothese H_1: $\mu(G) < 1002$.

De mogelijkheid dat dit pak toevalligerwijs minder dan 998 gram weegt, terwijl de fabrikant toch gelijk heeft, is een fout van de eerste soort. De kans op zo'n fout heet het significantieniveau van de toets. Die kans is hier $\alpha \approx 0{,}091$.
Meestal wordt geprobeerd om α kleiner te krijgen, want de consumentenbond doet niet graag foute uitspraken. Ze stellen dan voorafgaande aan de toets vast welke waarde van α nog acceptabel is en bepalen het bijbehorende beslissingsvoorschrift.
En natuurlijk nemen ze een grotere steekproef. Want dan geldt de $\sqrt{n}$-wet...

Opgave 1

In **Uitleg 1** zie je hoe de consumentenbond het gewicht van kilopakken suiker controleert.

a Er is hier sprake van een enkelzijdige normale toets. Kun je die naam verklaren?

b Waarom voert de consumentenbond een enkelzijdige toets uit? Hoe zou dat met de fabrikant zelf zijn?

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

d Bij welk gewicht zou er dan sprake zijn van een significant verschil?

Opgave 2

Gebruik de gegevens van **Uitleg 1**.

a 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?

b Welk bezwaar zit er aan deze toets?

Uitleg 2

Je gaat nu het gemiddelde gewicht van de pakken suiker toetsen door een steekproef van 100 pakken.

De fabrikant beweert dat het gewicht G (gram) van zijn pakken suiker een gemiddelde $\mu(G) = 1002$ en een standaardafwijking $\sigma(G) = 3$ hebben.
Het gewicht G van die pakken suiker hoeft niet normaal verdeeld te zijn.

Toch is het gemiddelde gewicht $\overline{G}$ van de pakken suiker uit deze steekproef automatisch normaal verdeeld. Want uit de centrale limietstelling volgt: als er veel steekproeven genomen worden, zijn de gemiddelde steekproefgewichten normaal verdeeld. Dat geldt ook voor de gemiddelden van steekproeven van 100.

Om te kunnen toetsen moeten het gemiddelde van de steekproef $\mu(\overline{G})$ en de standaardafwijking van dit gemiddelde $\sigma(\overline{G})$ bekend zijn.

- $\mu(\overline{G})$ volgt uit de steekproefgegevens: $\mu(\overline{G}) = \mu(G)$
- $\sigma(\overline{G})$ volgt uit de wortel-n-wet: $\sigma(\overline{G}) = \frac{\sigma(G)}{\sqrt{n}}$

Bij gebruik van het kritieke gebied uit het eerdere voorbeeld volgt nu de overschrijdingskans:

$$\text{P}\left(\overline{G} < 998 | \mu(\overline{G}) = 1002 \text{ en } \sigma(\overline{G}) = \frac{3}{\sqrt{100}}\right) \approx 0$$

De kans dat er een foute conclusie wordt getrokken is dus veel kleiner geworden.

Opgave 3

Gebruik de gegevens in **Uitleg 1**. Nu neemt de consumentenbond een steekproef van 100 pakken.

a Wat betekent dit voor de toets?

b Als in die steekproef het gemiddelde gewicht 999 gram zou zijn, zou de consumentenbond dan met een betrouwbaarheid van 99% gelijk krijgen?

Opgave 4

Gebruik de gegevens uit **Uitleg 2**. Neem nu aan de consumentenorganisatie een steekproef van 8 pakken hanteert.

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

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

Iemand beweert: stochast X is normaal verdeeld met $\mu(X) = \mu$ en $\sigma(X) = \sigma$.
Iemand anders vertrouwt het gemiddelde niet en vermoedt (bijvoorbeeld): $\mu(X) \neq \mu$.
Dit wordt getoetst met een steekproef van grootte n. Je bepaalt dan het gemiddelde in de steekproef en kijkt of de afwijking van μ significant is.

Bij zo'n **normale toets van het gemiddelde** is de **nulhypothese** $H_0 : \mu(X) = \mu$.
Daarnaast staat een **alternatieve hypothese** $H_1 : \mu(X) \neq \mu$.

Het gemiddelde $\overline{X}$ in de steekproef is ook normaal verdeeld met $\mu\left(\overline{X}\right) = \frac{n \cdot \mu}{n} = \mu$ en $\sigma\left(\overline{X}\right) = \frac{\sqrt{n} \cdot \sigma}{n} = \frac{\sigma}{\sqrt{n}}$ als de nulhypothese inderdaad waar is.
Bij de alternatieve hypothese hoort een **kritiek gebied** dat aangeeft waar de afwijking van μ zo groot is dat je de nulhypothese verwerpt. Dat kritieke gebied bepaal je op grond van een vooraf vastgesteld **significantieniveau** α, bijvoorbeeld:

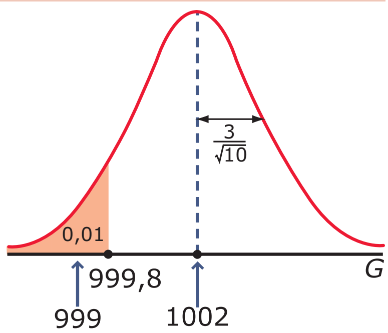
$$\mathrm{P}\left(\overline{X} \leq g_1 \vee \overline{X} > g_2 \mid \mu\left(\overline{X}\right) = \mu \text{ en } \sigma\left(\overline{X}\right) = \frac{\sigma}{\sqrt{n}}\right) \leq \alpha$$

Hierin zijn g_1 en g_2 de grenswaarden van het kritieke gebied $\overline{X} \leq g_1$ of g_2 de grenswaarden van het kritieke gebied $\overline{X} > g_2$.
Het significantieniveau kies je voordat je de toets uitvoert, bijvoorbeeld 10%, 5% of 1%. De keuze geeft informatie over de significantie van de toets.

Voorbeeld 1

Bekijk de applet: normale verdeling gewicht kilopakken suiker

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 $\mu(G) < 1002$.
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 3

Antwoord

- $H_0 : \mu\left(\overline{G}\right) = 1002$ en $\sigma\left(\overline{G}\right) = \frac{3}{\sqrt{10}} \approx 0{,}95$.
- $H_1 : \mu\left(\overline{G}\right) < 1002$

Het significantieniveau is $\alpha = 0{,}01$.
$\mathrm{P}\left(\overline{G} < g \mid \mu = 1002 \text{ en } \sigma = 0{,}95\right) \leq 0{,}01$ geeft: $g \approx 999{,}8$.
Het kritieke gebied wordt daarom: $\overline{G} \leq 999{,}8$.
Het in de steekproef gevonden gemiddelde geeft daarom inderdaad aanleiding om de bewering van de fabrikant in twijfel te trekken bij een significantieniveau van 1%.

Opgave 5

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

a Voer de beschreven toets zelf uit.

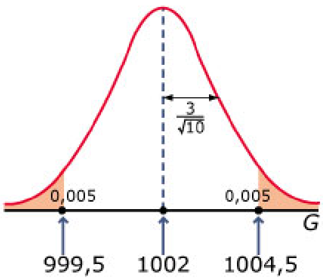
b Voer de toets nog eens uit, maar nu met een betrouwbaarheid van 99,5%. Is er nog steeds sprake van een significante afwijking?

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

Voorbeeld 2

Bekijk de applet: normale verdeling gewicht kilopakken suiker

Volgens de fabrikant is het gewicht G (in gram) van zijn pakken suiker normaal verdeeld met $\mu(G) = 1002$ en $\sigma(G) = 3$.
De fabrikant test zijn vulmachine door van een steekproef van 10 pakken het gemiddelde gewicht te berekenen. Hij doet uiteraard een dubbelzijdige toets.
Wat is het beslissingsvoorschrift bij een significantieniveau van 1%?



Figuur 4

Antwoord

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

Het significantieniveau is $\alpha = 0{,}01$.

$P\left(\overline{G} \leq g_1 \vee \overline{G} > g_2 | \mu = 1002 \text{ en } \sigma = 0{,}95\right) \leq 0{,}01$ geeft: $g_1 \approx 999{,}5$ en $g_2 \approx 1004{,}5$.
Het kritieke gebied wordt daarom: $\overline{G} \leq 999{,}5$ of $\overline{G} \geq 1004{,}5$.

Opgave 6

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 7

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?

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?

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 batterijen levert voor rekenmachines, beweert dat die batterijtjes geschikt zijn om gemiddeld zo'n apparaat 3600 uur te laten werken. Ze gaan er van uit dat die levensduur normaal is verdeeld met een standaarddeviatie van 600 uur.
In een aselect gekozen groep van 60 rekenmachines stop je de batterijen van deze firma. De gemiddelde levensduur blijkt 3300 uur 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 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? Kun je een verklaring geven voor het verschil met het antwoord bij c?

Opgave 12

Vacuüm verpakte vleeswaren mogen maximaal 0,022% natriumnitriet bevatten. Bij de keuringsdienst van waren controleren ze dit percentage.

a Formuleer de hypothese die getoetst wordt. Je mag aannemen dat het natriumnitrietpercentage normaal verdeeld is.

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

Bij de nulhypothese zijn er nog meer mogelijkheden voor μ. Omdat zelfs 0,022 toegestaan is, wordt er uitgegaan van $H_0 : \mu = 0{,}022$. 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 natuurlijk 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.

percentage	freq.
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

a Onderzoek met normaal waarschijnlijkheidspapier of de resultaten van deze proef reden geven om deze veronderstelling in twijfel te trekken.

b Schat met behulp van deze steekproef de parameter σ van de normale verdeling.

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

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.

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.

a 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. 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 normale toets van het gemiddelde kan de standaardafwijking van de steekproef worden gebruikt. In de theorie van dit onderwerp wordt opgemerkt dat er dan rekening moet worden gehouden met een grotere fout.
William Sealy Gosset (1876—1937) was een Engelse statisticus, die publiceerde onder het pseudoniem ‘Student’. Deze wiskundige heeft onderzoek gedaan naar deze fout.
Hij ontdekte dat er in dat geval van een iets andere verdeling dan de normale verdeling gebruikgemaakt kan worden. Deze verdeling wordt de Student- of t-verdeling genoemd.
Voor de Studentverdeling moet de steekproefomvang bekend zijn. Als deze n is, moet bij vrijheidsgraden $n - 1$ worden ingevuld.

a Teken met behulp van bijvoorbeeld de grafische rekenmachine de normale verdeling en de Studentverdeling. Gebruik bij de normale verdeling $\mu = 0$ en $\sigma = 1$. Bij de Studentverdeling is dit vaak automatisch het geval. De Studentverdeling is in de meestal beschikbaar als tPdf-functie. 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.