6.2 Binomiale toetsen

Inleiding

Meestal ga je er van uit dat bij het verwekken van kind de kans op een jongen even groot is als die op een meisje: de kans op een meisje is 50% is je nulhypothese. Als bij een zekere Nederlandse gemeente in een bepaald jaar 60% van de geboren kinderen een meisje is dan denk je niet meteen dat de kans op een meisje nu wel 60% moet zijn geworden, maar je vraagt je wel af of de kans op een meisje in Nederland soms meer dan 50% is geworden. Zo'n hypothese kun je heel goed toetsen bijvoorbeeld door te kijken naar de geboren kinderen van het volgende jaar. Maar wanneer zeg je nu dat de kans op een meisje niet langer 50% is?

Je leert in dit onderwerp

- hypothesen toetsen met behulp van de binomiale kansverdeling;
- het begrip significantieniveau;
- bij een gegeven significantieniveau een binomiale toets uitvoeren.

Voorkennis

- werken met binomiale kansverdelingen;
- de begrippen nulhypothese, alternatieve hypothese en kritiek gebied.

Verkennen

Opgave V1

In 2016 was in de Nederlandse gemeente A 60% van de geboren kinderen een meisje. Je vraagt je af of de kans op een meisje in Nederland soms meer dan 50% is geworden. Je neemt in 2017 een steekproef van 650 in dat jaar geborenen door heel Nederland en vraagt of het een jongen dan wel een meisje betreft.

a Wat is de nulhypothese en wat de alternatieve hypothese in dit geval?

b Stel je voor dat er in je steekproef 348 meisjes voorkomen. Hoe groot is nu de kans dat de nulhypothese ten onrechte wordt verworpen?

c Stel je eens voor dat je de kans dat de nulhypothese ten onrechte wordt verworpen maximaal 1% wilt hebben. Wat wordt dan het kritieke gebied van de toets?

Uitleg

Omdat het uitvoeren van een hypothesetoets afhangt van een steekproef, bestaat er een kans op een foute conclusie. De kans op de foute conclusie
'H_0 is juist, maar H_1 wordt geaccepteerd'
wordt het significantieniveau genoemd. Deze significantie mag niet te groot zijn. In de praktijk wordt voor deze kans vaak 10% of 5% gebruikt. Bij belangrijk of nauwkeurig onderzoek wordt meestal 1% gebruikt.

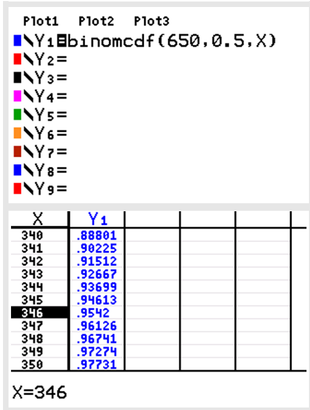
Als het significantieniveau (de foutkans) is gekozen, kun je daarmee het kritieke gebied berekenen. Bekijk de volgende toets:

M is het aantal meisjes in een steekproef met steekproefomvang n en p is het deel van de steekproef dat meisje is.
$H_0 : p = 0{,}5$ en $H_1 : p > 0{,}5$.
Neem aan dat M binomiaal verdeeld is.
De steekproefomvang $n = 650$.
Het significantieniveau (de kans dat H_0 juist is, maar H_1 wordt geaccepteerd) mag maximaal 5% zijn.
Als H_0 juist is, geldt: $p = 0{,}5$. Voor de grens g van het kritieke gebied moet dus gelden:

$\text{P}(M > g | p = 0{,}5 \text{ en } n = 650) \leq 0{,}05$

Ofwel: $\text{P}(M \leq g | p = 0{,}5 \text{ en } n = 650) \leq 0{,}95$.

De grafische rekenmachine geeft $g = 346$.



Figuur 1

Als er dus meer dan 346 meisjes in de steekproef worden geteld, mag je met een significantie van 5% besluiten dat het deel meisjes in de steekproef groter dan 0,5 is.

Opgave 1

Gebruik de gegevens uit de **Uitleg**.

a Het significantieniveau van de toets is 5%. Leg uit wat dit significantieniveau precies betekent.

b Waarom werk je hier met de waarde van p (het deel meisjes) die bij H_0 hoort?

c Reken het kritieke gebied van deze toets na.

Opgave 2

Gebruik de gegevens uit de **Uitleg**.

a Voer deze toets nog eens uit, maar nu met een significantie van 1%. Geef in dit geval het kritieke gebied.

b Welke invloed heeft het verkleinen van de significantie op het kritieke gebied?

c Hoe verandert het kritieke gebied als de significantie lager wordt gemaakt? Wat is er aan de hand als de nulhypothese dan toch wordt verworpen?

Theorie en voorbeelden

Om te onthouden

Veronderstel dat voor een binomiale stochast X de gangbare opvatting is dat $p = p_0$ het deel is van de elementen van een steekproef met een bepaalde eigenschap. Iemand bestrijdt deze opvatting en beweert dat $p > p_0$.

Hier wordt dus $H_0 : p = p_0$ tegen $H_1 : p > p_0$ getoetst.

Omdat X een binomiale stochast is, heet dit een **binomiale toets** van de **proportie** (= deel).

Bij statistisch onderzoek wordt vaak geëist, dat de kans op de fout ‘H_0 wordt verworpen terwijl deze toch waar is’ klein is. De waarde van deze foutkans heet het **significantieniveau** of de **onbetrouwbaarheidsdrempel**. Dit wordt aangegeven met de Griekse letter α. De waarde van de significantie moet vooraf worden afgesproken, bijvoorbeeld: $\alpha = 0{,}05$.

Met deze waarde van de significantie kun je het kritieke gebied bij steekproefomvang N berekenen:
$P(H_0 \text{ verwerpen} \mid H_0 \text{ is waar}) = P(X \geq g | p = p_0 \text{ en } n = N) \leq \alpha$
De berekende g is dan de grens van het kritieke gebied.

Er wordt onderscheid gemaakt tussen drie verschillende soorten toetsen. Deze hangen af van de alternatieve hypothese:

- als $H_1 : p > p_0$ spreek je van een **rechtszijdige toets**;
- als $H_1 : p < p_0$ spreek je van een **linkszijdige toets**;
- als $H_1 : p \neq p_0$ spreek je van een **tweezijdige toets**.

Bij de tweezijdige toets bestaat het kritieke gebied (meestal) uit twee delen. De onbetrouwbaarheidsdrempel α_0 verdeel je dan in twee gelijke delen voor elk deel van het kritieke gebied.

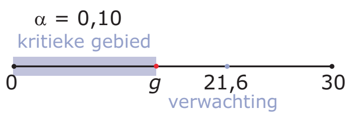
Voorbeeld 1

Voor een toets heeft tot nu toe 72% van de leerlingen een voldoende gescoord. Een klas van 30 leerlingen maakt de toets. Bij hoeveel voldoendes is het resultaat significant lager dan 72%, bij een significantieniveau van 10%?

Antwoord

‘Vertaal’ deze vraag naar een linkszijdige binomiale toets.

- Stochast X is het aantal van de 30 kandidaten dat een voldoende heeft en is binomiaal verdeeld.
- $H_0 : p = 0{,}72$
- $H_1 : p < 0{,}72$
- $n = 30$
- $\alpha = 0{,}1$



Figuur 2

Er geldt: $P(X \leq g | p = 0{,}72 \text{ en } n = 30) \leq 0{,}1$
Dit levert op: $g = 17$ en dus wordt het kritieke gebied $X \leq 17$.

Opgave 3

Gebruik de gegevens uit **Voorbeeld 1**.

a Reken na, dat het kritieke gebied gelijk is aan $X \leq 17$.

b Bepaal nogmaals het kritieke gebied, maar nu met een betrouwbaarheid van 95%.

Opgave 4

Je toetst $H_0 : p = 0{,}35$ tegen $H_1 : p > 0{,}35$ met een significantieniveau van 5%.

a Bepaal het kritieke gebied bij een steekproef met grootte 100.

b Doe hetzelfde bij een steekproef met grootte 1000.

c Welke invloed heeft de grootte van de steekproef op de grens van het kritieke gebied?

d Waarom neemt men niet altijd een zo groot mogelijke steekproef?

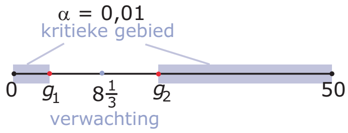
Voorbeeld 2

Om te bepalen of een dobbelsteen zuiver is, kun je bijvoorbeeld 50 keer met deze dobbelsteen werpen en het aantal keren ‘zes ogen’ tellen. Bij hoeveel keer ‘zes ogen’ mag je dan besluiten dat hij niet zuiver is? Neem een significantieniveau van 1%.

Antwoord

Je kunt deze vraag ‘vertalen’ naar een tweezijdige binomiale toets:

- X is het aantal keren ‘zes’ gegooid met een dobbelsteen, X is binomiaal verdeeld.
- (De verwachte waarde van X is $50 \cdot \frac{1}{6} = 8\frac{1}{3}$.)
- $\text{H}_0 : p = \frac{1}{6}$
- $\text{H}_1 : p \neq \frac{1}{6}$
- steekproefgrootte 50
- $\alpha = 0{,}01$



Figuur 3

Nu moet $\text{P}\left(X \leq g_1 \text{ of } X > g_2 | p = \frac{1}{6} \text{ en } n = 50\right) \leq 0{,}01$.

Je bepaalt de twee grenzen daarom uit:

- $\text{P}\left(X \leq g_1 | p = \frac{1}{6} \text{ en } n = 50\right) \leq 0{,}005$
- $\text{P}\left(X \geq g_2 | p = \frac{1}{6} \text{ en } n = 50\right) \leq 0{,}005$

Hieruit volgt $g_1 = 1$ en $g_2 = 16$. De kritieke gebieden zijn dus $0 \leq X \leq 1$ en $16 \leq X \leq 50$.

Opgave 5

Gebruik de gegevens uit **Voorbeeld 2**.

a Waarom is deze toets tweezijdig?

b Hoe wordt de onbetrouwbaarheidsdrempel α hier verwerkt?

c Bereken het kritieke gebied bij een significantieniveau van 5%.

Opgave 6

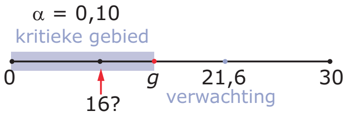
Bij een spel moet er met een achtvlaksdobbelsteen gegooid worden. Manon vermoedt dat deze dobbelsteen niet zuiver is. Ze gooit er 100 keer mee en telt hoe vaak ze acht ogen gooit. Bij hoeveel keer ‘acht ogen’ mag Manon besluiten dat de dobbelsteen niet zuiver is? Neem een significantieniveau van 2,5%.

Voorbeeld 3

Gegeven is een toets die door 72% van de leerlingen voldoende werd gemaakt.
De toets is nog een keer gemaakt door een groep van 30 leerlingen. 16 leerlingen hebben de toets voldoende gemaakt. Is deze toets bij een significantie van 10% slechter gemaakt dan tot nu toe?

Antwoord

Bekijk de afbeelding.



Figuur 4

Voor de grens g van het kritieke gebied geldt:
de kans dat er g of minder leerlingen een voldoende hebben gehaald, is kleiner dan de significantie 0,1. En er geldt: de kans dat er meer dan g leerlingen een voldoende hebben gehaald, is groter dan 0,1.

Dus ook geldt: als 16 in het kritieke gebied ligt, ligt 16 links van g en is de kans op 16 of minder leerlingen met een voldoende kleiner dan 0,1.

Door te controleren of geldt dat $P(X \leq 16) < 0{,}1$ is direct duidelijk of 16 in het kritieke gebied ligt. Je hoeft het kritieke gebied niet meer te bepalen. Deze kans wordt overschrijdingskans genoemd.

Hier geldt $P(X \leq 16 | p = 0{,}72 \text{ en } n = 30) = 0{,}0225$. Omdat dit kleiner is dan de significantie (0,1), is de conclusie: deze toets is met een significantie van 10% slechter gemaakt dan tot nu toe.

Opmerking: bij een tweezijdige toets kan de proportie van de steekproef links of rechts van de proportie van de populatie liggen. In dat geval kies je de bijbehorende overschrijdingskans (links of rechts) en vergelijk je met $\frac{1}{2}\alpha$.

Opgave 7

Bekijk **Voorbeeld 3**.

a Waarom is het in dit geval niet nodig om het kritieke gebied vast te stellen?

b Wat betekent het als de overschrijdingskans kleiner is dan de onbetrouwbaarheidsdrempel?

c Wat betekent het als de overschrijdingskans groter is dan de onbetrouwbaarheidsdrempel?

Opgave 8

Je toetst $H_0 : p = 0{,}75$ tegen $H_1 : p \neq 0{,}75$ met een onbetrouwbaarheidsdrempel van $\alpha = 0{,}05$.

a Bepaal het kritieke gebied als je een representatieve steekproef met omvang 100 gebruikt.

b Stel dat je vooraf hebt bepaald dat in de steekproef 80 elementen de betreffende eigenschap hebben. Laat zien hoe je in zo'n geval sneller te werk kunt gaan.

Verwerken

Opgave 9

Op een school slaagt elk jaar ongeveer 96% van de eindexamenkandidaten. Het afgelopen jaar viel het resultaat behoorlijk tegen. Slechts 92 van de 107 kandidaten haalden het eindexamen. Op andere scholen in de buurt waren er geen grote veranderingen ten opzichte van de slagingspercentages van de voorafgaande jaren. Er wordt geroepen: "De kwaliteit van de school holt achteruit". Een van de geslaagden is het daarmee niet eens: "Nee, dat is niet waar. Dat kan eens een jaar voorkomen. Die kans bestaat nou eenmaal".

a Toets de uitspraak van deze geslaagde leerling. Hoe luiden de nulhypothese en de alternatieve hypothese?

b Bereken de kans op 92 of minder geslaagden als de nulhypothese waar is.

c Krijgt de geslaagde leerling gelijk als het significantieniveau 0,01 is?

Opgave 10

In 2015 zijn er in Nederland 170510 baby's geboren. 83083 meisjes en 87427 jongens. Toets met deze getallen of de kans op een meisje kleiner is dan de kans op een jongen.
De nulhypothese is dan weer: de kans op een meisje is 0,5.

a Neem aan dat het aantal meisjes binomiaal is verdeeld. Welke parameters heeft deze verdeling?

b Voer de toets uit. Moet de nulhypothese worden verworpen of geaccepteerd als de onbetrouwbaarheidsdrempel 0,1% is?

Opgave 11

Iemand wil de zuiverheid van een dobbelsteen controleren en besluit er 600 keer mee te werpen. Hij let op de uitkomsten 1 en 2. Bij welke aantallen zal hij besluiten dat de dobbelsteen onzuiver is als hij een betrouwbaarheid van 99,9% hanteert?

Opgave 12

Het hoofd van de personeelsadministratie van een bedrijf weet dat het aantal ziekteverzuimdagen op alle dagen van de week ongeveer even groot was. Maar nu er nieuwe werktijden zijn ingevoerd, wil zij onderzoeken of het aantal ziekteverzuimdagen op maandag veel groter is geworden dan op de andere dagen van de week. Om dit te onderzoeken vraagt zij het aantal ziekteverzuimdagen per werkdag op in een bepaalde maand. In deze tabel zie je de resultaten.

dag	ma	di	wo	do	vr
aantal ziekteverzuimdagen	95	61	58	63	11

Tabel 1

a Hoeveel ziekteverzuimdagen zou je op maandag mogen verwachten als het ziekteverzuim onafhankelijk is van de weekdag?

b Het vermoeden van het hoofd van de personeelsadministratie kun je met deze gegevens toetsen. Formuleer de nulhypothese en de alternatieve hypothese. Je mag uitgaan van een binomiale verdeling.

c Krijgt het hoofd van de personeelsadministratie gelijk? Neem een significantieniveau van $\alpha = 0{,}05$.

Opgave 13

Bij een loterij die elke week plaatsvindt, moet je op een formulier met daarop de getallen 1 tot en met 19, drie getallen aankruisen. Nadat de inlevertijd is verstreken, worden aselect drie getallen getrokken. Een deelnemer die op zijn lot ten minste twee van de drie getrokken getallen heeft aangekruist, krijgt een prijs.

a Toon aan dat, afgerond op twee decimalen, de kans op een prijs gelijk is aan 0,05.

b Een deelnemer beweert dat zijn kans op een prijs groter is dan 0,05. Deze bewering wordt getoetst. Er wordt een significantieniveau van 1% gekozen. Gedurende 100 weken wordt het aantal prijzen van deze persoon bijgehouden. Hij wint 11 prijzen. Voer een toets uit om te beoordelen of de deelnemer gelijk heeft.

Toepassen

Opgave 14: Onderdelen produceren

In een fabriek wordt een cilindervormig onderdeeltje voor machines gemaakt. Voor de diameter van dit onderdeel, D (μm, micrometer), geldt: D is normaal verdeeld met een gemiddelde van 10 μm en een standaardafwijking van 0,2 μm. De geproduceerde onderdelen worden als volgt gecontroleerd: Er wordt een steekproef genomen van de geproduceerde onderdelen met een omvang van 100. Deze onderdelen worden gepast in een rond gat met $D = 10{,}3$ μm. Het aantal niet passende onderdelen wordt geteld. Ga uit van een significantie van 5%.

a Hoe groot is de kans dat een onderdeel niet in het gat past?

b Voer een hypothesetoets uit om te bepalen hoeveel niet passende onderdelen er in de steekproef mogen zitten vóór men het productieproces gaat bijstellen.

Opgave 15: Fout van de tweede soort

Bij het toetsen van hypothesen kan er nog een andere fout optreden, namelijk dat de nulhypothese niet wordt verworpen, terwijl deze toch niet klopt. Deze fout heet ‘fout van de 2^e soort’. De kans op deze fout is alleen te berekenen als de waarde van p in de populatie bekend is. Maar dat is bijna nooit zo. Deze fout kan wel worden onderzocht.

Bekijk de volgende toets:
M is het aantal keren munt bij het werpen met een geldstuk, M is binomiaal verdeeld.
De indruk bestaat dat het geldstuk niet zuiver is. Er wordt een toets uitgevoerd met steekproefomvang 100, om te bepalen of het geldstuk zuiver is.

Er wordt een significantie van 5% gebruikt.

a Bepaal het kritieke gebied.

b Neem nu voor het kritieke gebied $M \geq 59$. Onderzoek de kans dat H_0 niet verworpen wordt, terwijl H_0 niet juist is, dus als $p > 0{,}5$. Doe dit door een tabel te maken waarin p loopt van 0,50 tot 0,60 met stapjes van 0,02 en bij deze waarden de gevraagde kans te berekenen.

c Hoe kan deze foutkans worden verkleind?

d Voor de opdrachten a en b nog een keer uit, maar nu met een steekproefomvang 1000. Vergelijk de beide uitkomsten.

Testen

Opgave 16

Op het instituut voor toegepast psychologisch onderzoek onderzoekt men helderziendheid. Mensen die beweren helderziend te zijn, worden uitgenodigd voor het volgende experiment. De helderziende wordt opgesloten in een kamer. Hij krijgt een serie van drie zeer verschillende plaatjes. Op hetzelfde ogenblik ontvangt een ander persoon in dezelfde ruimte dezelfde drie plaatjes. Deze persoon krijgt de opdracht om zich vijf minuten lang op één van de drie plaatjes te concentreren. Na die vijf minuten moet de helderziende dan aangeven op welk plaatje de ander zich geconcentreerd heeft. De twee deelnemers kunnen elkaar niet zien. Ze hebben een koptelefoon op waardoor ze steeds een nieuwe opdracht krijgen. Dit proces wordt in totaal 40 keer herhaald met steeds nieuwe series plaatjes. De nulhypothese die men wil toetsen luidt: de helderziende is niet helderziend. Dat kan men toetsen met behulp van het aantal keren dat de helderziende hetzelfde plaatje aangeeft als de ander. Als de nulhypothese waar is dan is het aantal keren ‘hetzelfde plaatje’ binomiaal verdeeld.

a Welke parameters heeft deze binomiale verdeling?

b Helderziende X geeft 17 keer het juiste plaatje aan. Bereken de kans dat dit gebeurt als de nulhypothese juist is.

c Bereken de kans dat de nulhypothese ten onrechte verworpen wordt als er afgesproken wordt dat er sprake is van helderziendheid bij meer dan 20 juiste plaatjes.

d Er wordt afgesproken dat het significantieniveau 0,05 is. Wordt helderziende Y helderziend verklaard als hij 18 plaatjes goed heeft?

Opgave 17

Een kweker wil onderzoeken of zijn kruisingsmethode als resultaat heeft dat 25% van de bessenstruiken gevoelig is voor meeldauw. Hij constateert dat van de 100 bessenstruiken er 33 last hebben van meeldauw.
Mag hij nu met een significantieniveau van 5% aannemen dat toch maar 25% gevoelig is voor meeldauw?