6.2 Binomiale toetsen

Inleiding

Meestal ga je er van uit dat bij het verwekken van een 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

In 2011 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 2018 een steekproef van 650 in dat jaar geborenen door heel Nederland en vraagt of het een jongen dan wel een meisje betreft.
Je wilt nu vooraf vaststellen bij welk aantal meisjes M in deze steekproef je kunt zeggen dat de kans op een meisje inderdaad meer dan 50% is. M is een binomiale stochast.

Je toetst H_0: $p = 0{,}5$ tegen H_1: $p > 0{,}5$.
Je gebruikt een steekproef van grootte $n = 650$.

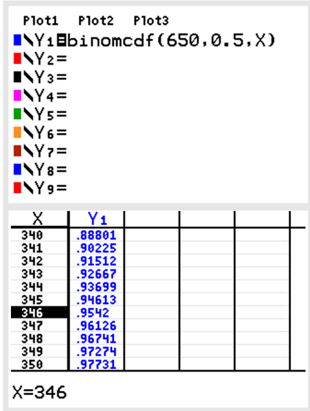
Bij zo'n toets neem je als uitgangspunt het vermijden van een fout van de eerste soort: de kans dat je H_0 ten onrechte verwerpt moet kleiner zijn dan (bijvoorbeeld) 5%. Deze 5% heet het significantieniveau of de onbetrouwbaarheidsdrempel van de toets en wordt aangegeven met α. Je spreekt dus VOORAF af dat (bijvoorbeeld) $\alpha = 0{,}05$.

Je kritieke gebied is $M > g$. Dan moet:

$\text{P}(\text{H}_0 \text{ verwerpen} \mid \text{H}_0 \text{ is waar}) =$
$= \text{P}(M > g \mid p = 0{,}5 \wedge n = 650) \leq 0{,}05$

Dus moet $\text{P}(M \leq g \mid p = 0{,}5 \wedge n = 650) \leq 0{,}95$.
Je grafische rekenmachine geeft: $g = 346$.

Als je dus meer dan 346 meisjes in je steekproef aantreft mag je met een significantieniveau van 5% besluiten dat de kans op een meisje meer dan 50% is.



Figuur 1

Opgave 1

In de **Uitleg** gaat het weer om de hypothese dat de kans op de geboorte van een meisje 50% is. Het significantieniveau van de toets is 5%.

a Leg uit wat dit significantieniveau precies betekent.

b Er is hier sprake van een enkelzijdig binomiale toets. Kun je die naam verklaren?

c Reken zelf het kritieke gebied van deze toets na.

d Waarom kun je ook zeggen dat de beschreven toets een betrouwbaarheid heeft van 95%?

Opgave 2

Bekijk nog eens de toets in de **Uitleg**.

a Voer deze toets nog eens uit, maar nu met een significantieniveau van 1%.

b Welke invloed heeft het verkleinen van het significantieniveau op het kritieke gebied?

c Waarom wordt het significantieniveau niet nog veel kleiner genomen, zeg 0,1%?

Theorie en voorbeelden

Om te onthouden

Stel je voor dat voor een binomiale stochast X de gangbare opvatting is dat $p = p_0$ de kans is dat een element van de steekproef een bepaalde eigenschap heeft. Iemand bestrijdt deze opvatting en beweert dat $p > p_0$.

Je toetst $\text{H}_0 \colon p = p_0$ tegen $\text{H}_1 \colon p > p_0$ met behulp van een steekproef van grootte N.

Het **kritieke gebied** (de waarden in de steekproef waarbij je H_0 verwerpt) van de toets stel je vast door te eisen dat de kans, dat je H_0 verwerpt terwijl hij toch waar is, kleiner blijft dan een bepaalde waarde die je vooraf kiest. Die waarde noem je het **significantieniveau** of de **onbetrouwbaarheidsdrempel** van de toets en hij wordt aangegeven met α. Je spreekt dus VOORAF af dat (bijvoorbeeld) $\alpha = 0{,}05$. Dit betekent:

$\text{P}(\text{H}_0 \text{ verwerpen} \mid \text{H}_0 \text{ is waar}) = \text{P}(X > g \mid p = p_0 \wedge n = N) \leq 0{,}05$

Op grond hiervan bereken je g en zo stel je het kritieke gebied van de toets vast.

Dit is een **rechtszijdige toets** omdat $\text{H}_1 \colon p > p_0$.
Zo bestaat er ook een **linkszijdige toets** met $\text{H}_1 \colon p < p_0$.
En een **tweezijdige toets** met $\text{H}_1 \colon p \neq p_0$.
In dit laatste geval trek je aan beide zijden van het op grond van H_0 verwachte aantal in de steekproef een grens. De onbetrouwbaarheidsdrempel α verdeel je dan in twee gelijke delen voor elk deel van het kritieke gebied.

Voorbeeld 1

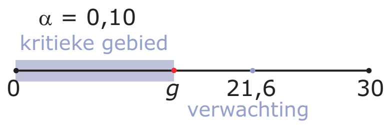
Voor een bepaalde toets scoort gemiddeld 72% van de kandidaten een voldoende. Deze keer hebben 16 van de 30 kandidaten een voldoende gehaald, duidelijk minder dan 72%. Wijkt dit resultaat significant af van het verwachte resultaat als je een significantieniveau van 10% hanteert?

Antwoord

Je kunt deze vraag vertalen naar een linkszijdige binomiale toets:

- $\text{H}_0 : p = 0{,}72$
- $\text{H}_1 : p < 0{,}72$

De steekproefgrootte is 30 en stochast X is het aantal kandidaten met een voldoende in deze steekproef. Het significantieniveau is $\alpha = 0{,}10$.



Figuur 2

Je bepaalt het kritieke gebied uit: $\text{P}(X \leq g | n = 30 \wedge p = 0{,}72) \leq 0{,}10$.
Dit levert op: $g = 17$ en dus wordt het kritieke gebied $X \leq 17$.

Het resultaat van 16 voldoendes ligt binnen het kritieke gebied, dus inderdaad wijkt het resultaat significant af van het verwachte resultaat.

Opgave 3

Bestudeer **Voorbeeld 1**.

a Voer de beschreven toets zelf uit.

b Voer de toets nog eens uit, maar nu met een betrouwbaarheid van 95%. Is er nog steeds sprake van een significante afwijking?

Voorbeeld 2

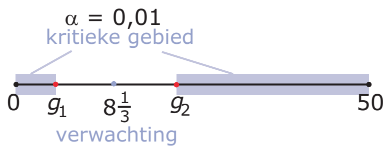
Om te bepalen of een dobbelsteen eerlijk is, kun je bijvoorbeeld 50 keer met deze dobbelsteen werpen en het aantal keren 'zes ogen' tellen. Bij hoeveel keer 'zes ogen' mag je dan besluiten dat hij niet eerlijk is? Neem een significantieniveau van 1%.

Antwoord

Je kunt deze vraag vertalen naar een tweezijdige binomiale toets:

- $\text{H}_0 : p = \frac{1}{6}$
- $\text{H}_1 : p \neq \frac{1}{6}$

De steekproefgrootte is 50 en de stochast X is het aantal keren 'zes ogen' in deze steekproef. Het significantieniveau is $\alpha = 0.01$.



Figuur 3

Nu moet $P\left(X \leq g_1 \vee X > g_2 | n = 50 \wedge p = \frac{1}{6}\right) \leq 0{,}01$.
Je bepaalt de twee grenzen daarom uit:

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

Ga na, dat het kritieke gebied wordt: $X \leq 1 \vee X \geq 16$.
Bij deze aantallen zessen mag je aannemen dat de dobbelsteen niet eerlijk is.

Opgave 4

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%.

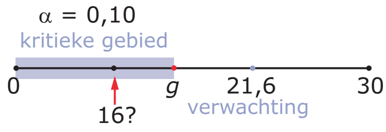
Voorbeeld 3

Voor een bepaalde toets scoort gemiddeld 72% van de kandidaten een voldoende. Deze keer hebben 16 van de 30 kandidaten een voldoende gehaald, duidelijk minder dan 72%. Wijkt dit resultaat significant af van het verwachte resultaat als je een significantieniveau van 10% hanteert?

Antwoord

Dit is hetzelfde probleem als dat in **Voorbeeld 1**. Je kunt het echter ook op een andere manier oplossen. Daarbij maak je meteen gebruik van het testresultaat: 16 van de 30 kandidaten haalden een voldoende.

Als 16 in het kritieke gebied van de toets ligt, dan moet dit gebied minstens $X \leq 16$ zijn (wellicht groter). En dan moet $P(X \leq 16 | n = 30 \wedge p = 0{,}72) \leq 0{,}10$.
Deze **overschrijdingskans** is 0,0225 en dus inderdaad kleiner dan $\alpha = 0{,}10$.



Figuur 4

Je weet nu zeker dat 16 in het kritieke gebied van de toets ligt, ook al heb je de grens ervan niet vastgesteld. De nulhypothese wordt verworpen, het resultaat wijkt significant af.

Opgave 5

In **Voorbeeld 3** wordt de toets van **Voorbeeld 1** opnieuw bekeken.

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 6

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 van grootte 1000.

c En nog eens bij een steekproef met grootte 10.000.

d Welke invloed heeft de grootte van de steekproef op de grens van het kritieke gebied? Wat heeft dit te maken met de standaardafwijking van de gebruikte binomiale verdeling?

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

Opgave 7

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 van 100 gebruikt.

b Stel je voor dat je vooraf hebt bepaald dat in de steekproef 80 elementen de betreffende eigenschap hebben. Leg uit waarom in dat geval het berekenen van het kritieke gebied niet nodig is. Laat zien hoe je in zo'n geval sneller te werk kunt gaan.

Verwerken

Opgave 8

Op een grote 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 hoor; dat is helemaal niet waar. Dat kan gewoon eens een jaar voorkomen. Die kans bestaat nou eenmaal".

a De uitspraak van deze geslaagde kun je toetsen. 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 9

Het hoofd van de personeelsadministratie van een groot bedrijf heeft het idee dat het ziekteverzuim op maandag veel groter is dan op de andere dagen in de week. Om dit te onderzoeken vraagt hij het aantal ziekteverzuimdagen per werkdag op in een bepaalde maand. In deze tabel zie je de resultaten.

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

Tabel 1

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

Het vermoeden van de bedrijfsleider kun je met deze gegevens toetsen. Als nulhypothese kun je nemen: $H_0 : p = 0{,}2$.

b Formuleer de alternatieve hypothese.

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

Opgave 10

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 tenminste 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.

Iemand beweert dat zijn kans op een prijs groter is dan 0,05. Deze bewering wordt getoetst door gedurende honderd weken het aantal prijzen van deze persoon te noteren. Er wordt een significantieniveau van 1% gekozen.

b Hoeveel prijzen moet deze persoon in die honderd weken tenminste hebben gewonnen om hem gelijk te kunnen geven?

Opgave 11

Iemand wil de zuiverheid van een geldstuk controleren en besluit er 500 keer mee te werpen.

Bij welke aantallen ‘kop’ zal hij besluiten dat het geldstuk onzuiver is als hij een betrouwbaarheid van 90% hanteert?

Opgave 12

In 1998 zijn er in Nederland 179.568 baby’s geboren, te weten 90.096 meisjes en 89.472 jongens. Met deze getallen kun je toetsen of toentertijd de kans op een meisje gelijk was aan 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 Moet de nulhypothese verworpen of geaccepteerd worden als de onbetrouwbaarheidsdrempel 5% is?

c En als de toets een betrouwbaarheid van 99% moet hebben?

Toepassen

Opgave 13: Onderdelen produceren

In een fabriek wordt een cilindervormig onderdeeltje voor machines gemaakt. Voor de diameter van dit onderdeel, D (in millimeter), geldt: D is normaal verdeeld met een gemiddelde van 10 millimeter en een standaardafwijking van 0,2 millimeter. 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 millimeter. 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 14: 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 tweede 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 a en b nog een keer uit, maar nu met een steekproefomvang 1000. Vergelijk de beide uitkomsten.

Testen

Opgave 15

Op het instituut voor toepast 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 16

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?

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.

Email: `f.spijkers@math4all.nl`

Met de Math4All maatwerkdienst kunnen complete readers worden samengesteld en toetsen worden gegenereerd. Docenten kunnen bij `a.f.otten@xs4all.nl` een gratis inlog voor de maatwerkdienst aanvragen.