4.4 Niet-binomiaal

Inleiding

Het binomiale kansmodel is erg overzichtelijk: er zijn maar twee mogelijkheden ‘succes’ of ‘mislukking’ en het gaat om herhaling van steeds dezelfde kanssituatie. Maar natuurlijk bestaan er heel veel discrete stochasten waarbij er meer dan twee mogelijkheden zijn en/of er geen herhaling plaatsvindt. Denk bijvoorbeeld in het vaasmodel aan een trekking zonder teruglegging.
Je zult in dit onderdeel een paar discrete niet-binomiale stochasten tegenkomen.

Je leert in dit onderwerp

- werken met niet-binomiale discrete stochasten en de bijbehorende kansverdelingen opstellen;
- regels voor de verwachting en de standaardafwijking van niet-binomiale discrete stochasten;
- een binomiale stochast gebruiken bij een kleine steekproef uit een grote populatie.

Voorkennis

- het vaasmodel;
- een kansverdeling opstellen bij een vaasmodel;
- de regels voor de verwachting en de standaarddeviatie van de som van meerdere stochasten.

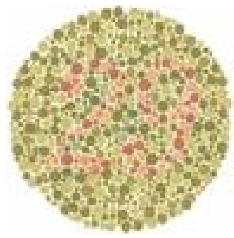
Verkennen

Opgave V1

In een groep van 25 mannen zijn 2 leden van die groep kleurenblind.
Je trekt aselect een steekproef van vier mannen uit deze groep.

a Hoe groot is de kans dat daarbij één kleurenblinde man zit?

b Hoe groot is de kans dat daarbij twee kleurenblinde mannen zitten?



Figuur 1

Uitleg 1

In een groep van dertig personen hebben tien mensen een bepaalde eigenschap en de rest niet. Uit die groep wordt aselect een steekproef van vijf getrokken.
Stochast M is het aantal mensen met deze eigenschap in deze steekproef.

Wil je nu een kansverdeling voor M opstellen, dan bedenk je dat het hier gaat om trekking zonder terugleggen. Dit betekent dat de kansen afhankelijk van elkaar zijn en dat een binomiaal kansmodel niet mogelijk is.
De kans op bijvoorbeeld $M = 2$ kun je zo berekenen:

$$\text{P}(M = 2) = \frac{10}{30} \cdot \frac{9}{29} \cdot \frac{20}{28} \cdot \frac{19}{27} \cdot \frac{18}{26} \cdot \binom{5}{2} \approx 0{,}3600$$

Dit kun je ook met behulp van combinaties berekenen.

$$\text{P}(M = 2) = \frac{\binom{10}{2} \cdot \binom{20}{3}}{\binom{30}{5}} \approx 0{,}3600$$

Ga na dat je de volgende kansverdeling krijgt.

m	0	1	2	3	4	5
$P(M=m)$	0,1088	0,3400	0,3600	0,1600	0,0295	0,0018

Tabel 1

Je kunt met behulp van de tabel de verwachting en de standaardafwijking berekenen.
Je vindt $E(M) \approx 1{,}667$ en $\sigma(M) \approx 0{,}979$.

Kennelijk gaat $E(M) = 5 \cdot \frac{10}{30} = 1\frac{2}{3}$ ook hier op, maar dit geldt niet voor de formule die bij de binomiale verdeling voor de standaardafwijking geldt.

Opgave 1

Bekijk in **Uitleg 1** de kansverdeling voor stochast M die het aantal mensen met een bepaalde eigenschap in een steekproef uit een kleine populatie van 30 personen weergeeft.

a Bereken $P(M=3)$.

b Bereken $E(M)$ en $\sigma(M)$ in vier decimalen nauwkeurig.

c Waarom is hier geen sprake van een binomiale kansverdeling?

Uitleg 2

In een groep van 30000 personen hebben 10000 mensen een bepaalde eigenschap en de rest niet.
Uit die groep wordt aselect een steekproef van 5 getrokken.
Stochast M is het aantal mensen met deze eigenschap in deze steekproef.

Wil je nu een kansverdeling voor M opstellen, dan bedenk je opnieuw dat het hier gaat om trekking zonder terugleggen. Dit betekent dat de kansen afhankelijk van elkaar zijn en dat een binomiaal kansmodel niet mogelijk is.
De kans op $M=2$ is:

$$P(M=2) = \frac{10000}{30000} \cdot \frac{9999}{29999} \cdot \frac{20000}{29998} \cdot \frac{19999}{29997} \cdot \frac{19998}{29996} \cdot \binom{5}{2} \approx 0{,}3292$$

Nu verschilt een breuk als $\frac{9999}{29999}$ vrijwel niet van $\frac{10000}{30000} = \frac{1}{3}$.

Daarom kun je als je een kleine steekproef uit een heel grote populatie trekt, toch het binomiale kansmodel gebruiken. Hoewel het eigenlijk niet om onafhankelijke kansen gaat.

$$P(M=2) \approx \left(\frac{1}{3}\right)^2 \cdot \left(\frac{2}{3}\right)^3 \cdot \binom{5}{2} \approx 0{,}3292$$

Je ziet dat beide kansen bij benadering gelijk zijn aan elkaar. Daarom wordt in de praktijk bij een steekproef uit een veel grotere populatie waarbij het gaat om het wel of niet hebben van een bepaalde eigenschap, gewoon het binomiale kansmodel gebruikt.

Opgave 2

Bekijk in **Uitleg 2** de kansverdeling voor stochast M die het aantal mensen met een bepaalde eigenschap in een steekproef uit een grote populatie van 30000 personen weergeeft.

a Bereken $P(M=3)$ en $P(M=4)$. Benader deze kansen ook met behulp van het binomiale kansmodel. Rond in beide gevallen af op vier decimalen nauwkeurig.

b Bereken $E(M)$ en $\sigma(M)$ in vier decimalen nauwkeurig.

c Waarom kun je de kansverdeling van M heel goed benaderen door een binomiale kansverdeling?

Theorie en voorbeelden

Om te onthouden

Heel vaak is in een bepaalde kanssituatie helemaal geen sprake van een binomiale stochast. Dan is er geen sprake van een herhaling van onafhankelijke Bernoulli-experimenten (succes of mislukking).

Een belangrijk geval is de **hypergeometrische stochast**.
Daarbij gaat het om een **populatie** van N elementen waarvan er a een bepaalde eigenschap hebben. Je trekt daaruit zonder teruglegging een **steekproef** van n elementen. De hypergeometrische stochast X is dan het aantal elementen in de steekproef dat deze eigenschap heeft. De kans op $X = x$ is:

$$\text{P}(X = x) = \frac{a}{N} \cdot \frac{a-1}{N-1} \cdot \ldots \cdot \frac{N-a}{N-x} \cdot \frac{N-a-1}{N-x-1} \cdot \ldots \cdot \binom{n}{x}$$

Met behulp van combinaties kun je dit ook uitrekenen:

$$\text{P}(X = x) = \frac{\binom{a}{x} \cdot \binom{N-a}{n-x}}{\binom{N}{n}}$$

Voor de verwachtingswaarde geldt: $\text{E}(X) = n \cdot \frac{a}{N}$.
De standaardafwijking van X kun je nu alleen uit de kansverdeling zelf halen. Daarom bepaal je in de praktijk zowel $\text{E}(X)$ als $\sigma(X)$ met behulp van de grafische rekenmachine.

Bij een **kleine steekproef uit een heel grote populatie** kun je toch wel het binomiale kansmodel gebruiken, hoewel het eigenlijk niet om onafhankelijke kansen gaat. Dat komt omdat dan breuken als $\frac{a}{N}$ en $\frac{a-1}{N-1}$ vrijwel gelijk zijn.
In de praktijk wordt vaak bij een steekproef uit een veel grotere populatie waarbij het gaat om het wel of niet hebben van een bepaalde eigenschap, gewoon het binomiale kansmodel gebruikt.

Voorbeeld 1

In een klas zitten acht jongens en twaalf meisjes. Daaruit wordt een aselecte steekproef van vier personen getrokken. Stochast M is het aantal meisjes in de steekproef.
Stel een kansverdeling op voor M en bepaal de verwachting en de standaardafwijking van M.

Antwoord

Bij de steekproef gaat het om trekking zonder terugleggen van vier elementen uit een populatie van twintig. M is een hypergeometrische stochast.
De kans op bijvoorbeeld $M = 3$ is:

$$\text{P}(M = 3) = \frac{12}{20} \cdot \frac{11}{19} \cdot \frac{10}{18} \cdot \frac{8}{17} \cdot 4 \approx 0{,}3633$$

De complete kansverdeling wordt:

m	0	1	2	3	4
$\text{P}(M = m)$	0,0145	0,1387	0,3814	0,3633	0,1022

Tabel 2

Met de grafische rekenmachine vind je dan: $\text{E}(M) = 2{,}4$ en $\sigma(M) \approx 0{,}899$.

Opgave 3

Bekijk **Voorbeeld 1**. Stochast J is het aantal jongens in de steekproef.

a Waarom is J geen binomiale stochast?

b Bereken zelf de kansen in de kansverdeling J. Rond af op vier decimalen.

c Bereken de verwachtingswaarde en de standaardafwijking van J in vier decimalen nauwkeurig.

d Bereken de kans dat er minstens drie jongens in de steekproef voorkomen. Rond af op drie decimalen.

Opgave 4

In een vaas zitten twee witte en drie rode balletjes. Uit deze vaas worden zonder teruglegging balletjes getrokken, totdat er een wit balletje wordt getrokken.
Wat is de verwachting en de variantie van het aantal benodigde trekkingen?

Voorbeeld 2

Op een scholengemeenschap zitten 800 jongens en 1200 meisjes. Daaruit wordt een aselecte steekproef van 4 personen getrokken. Stochast M is het aantal meisjes in de steekproef. Stel een kansverdeling op voor M en bepaal de verwachting en de standaardafwijking van M. Laat zien dat je kansen vrijwel hetzelfde zijn als je een binomiaal kansmodel gebruikt.

Antwoord

Bij de steekproef gaat het om trekking zonder terugleggen van 4 elementen uit een populatie van 2000. M is een hypergeometrische stochast.
De kans op bijvoorbeeld $M = 3$ is:

$$\text{P}(M = 3) = \frac{1200}{2000} \cdot \frac{1199}{1999} \cdot \frac{1198}{1998} \cdot \frac{800}{1997} \cdot 4 \approx 0{,}3458$$

Als je M benadert als binomiale stochast is deze kans gelijk aan: $\text{P}(M = 3) = \left(\frac{1200}{2000}\right)^3 \cdot \frac{800}{2000} \cdot 4 \approx 0{,}3456$.

Je ziet dat je de kansen met een binomiaal kansmodel goed kunt benaderen:

m	0	1	2	3	4
$\text{P}(M = m)$	0,0256	0,1536	0,3456	0,3456	0,1296

Tabel 3

Nu vind je: $\text{E}(M) = 4 \cdot \frac{1200}{2000} = 2{,}4$ en $\sigma(M) = \sqrt{4 \cdot \frac{1200}{2000} \cdot \frac{800}{2000}} \approx 0{,}980$.

Opgave 5

In **Voorbeeld 2** gaat het om een steekproef van 4 uit een populatie van 2000 personen. M is het aantal meisjes in de steekproef.

a Waarom is M nog steeds geen binomiale stochast? Maar waarom kun je M nu wel goed benaderen met een binomiale stochast?

b Bereken zelf de kansen in de kansverdeling M.

c Reken de verwachtingswaarde en de standaardafwijking van M na.

d Bereken de kans dat er minstens 3 meisjes in de steekproef voorkomen.

Opgave 6

Een gezelschap bestaat uit drie mannen, vier vrouwen en vijf kinderen. Op een buurtfeest moet, om aan een spel deel te nemen, op aselecte wijze een team van vier personen uit de groep worden samengesteld.

a Welk kansmodel moet je gebruiken om de kans te berekenen dat in de groep twee kinderen zitten? Waarom?

b Hoe groot is de kans bedoeld in a? Rond af op vier decimalen.

c Hoe groot is de kans dat in de groep van vier minstens twee vrouwen zitten? Rond af op vier decimalen.

d Hoe groot is de kans dat de groep alleen uit vrouwen en kinderen bestaat? Rond af op vier decimalen.

e Hoeveel kinderen mag je in de groep verwachten?

Voorbeeld 3

Je hebt ergens gelezen dat op dit moment 23% van alle Nederlandse meisjes van 12 tot en met 18 jaar rookt. Je weet dat deze groep meisjes uit ongeveer 450000 personen bestaat. Je vraagt 50 voor jou onbekende Nederlandse meisjes uit die leeftijdscategorie of ze roken.
Hoe groot is de kans dat minstens 15 daarvan roken?

Antwoord

Hier is sprake van een steekproef uit een veel grotere populatie. Hoewel er in feite sprake is van een hypergeometrische stochast, kun je het aantal rokende meisjes M in de steekproef opvatten als binomiale stochast.

De gevraagde kans is $\text{P}(M \geq 15) = 1 - \text{P}(M \leq 14) \approx 0{,}1565$.

Opgave 7

In **Voorbeeld 3** gaat het om het berekenen van kansen dat een bepaald aantal meisjes in een steekproef van 50 uit een populatie van 450000 meisjes rookt.

a Geef aan hoe je $\text{P}(M = 15)$ zou berekenen.

b Waarom kun je in dit geval heel goed met een binomiaal kansmodel werken?

c Benader M nu als binomiale stochast en bereken $\text{P}(M = 15)$ in vier decimalen nauwkeurig.

d Controleer dat $\text{P}(M \geq 15) \approx 0{,}1565$.

Opgave 8

Van alle leerlingen uit het basisonderwijs is bekend dat 90% rechtshandig is. Hoe groot is de kans dat je in een willekeurig gekozen groep van 20 kinderen minder dan 16 rechtshandigen aantreft? Rond af op vier decimalen nauwkeurig.

Verwerken

Opgave 9

In een klas zitten 8 jongens en 12 meisjes. Daaruit wordt een aselecte steekproef van 3 personen getrokken. Stochast J is het aantal jongens in de steekproef.

a Stel de kansverdeling voor J op. Rond de kansen af op vier decimalen.

b Bereken de verwachtingswaarde en de standaardafwijking van J. Rond af op één decimaal.

Opgave 10

Het bestuur van een politieke partij bestaat uit 20 personen, waarvan 40% jonger is dan 28 jaar. Door het lot worden 4 personen aangewezen om deel te nemen aan een buitenlandse reis.

a Hoeveel personen van de groep van 4 zijn naar verwachting jonger dan 28 jaar?

b Bepaal in vier decimalen nauwkeurig de kans dat 3 van de 4 personen jonger zijn dan 28 jaar.

c Benader deze kans ook met behulp van een binomiaal kansmodel. Hoe groot is de afwijking met de juiste kans?

Tijdens een regionale bijeenkomst van diezelfde partij zijn 100 leden aanwezig. Van deze leden is 40% jonger dan 35 jaar. Door het lot worden 4 personen aangewezen om deze regionale groepering op het landelijk congres van de partij te vertegenwoordigen.

d Bepaal in vier decimalen nauwkeurig de kans dat 3 van de 4 afgevaardigden jonger zijn dan 35 jaar.

e Benader ook deze kans binomiaal. Vind je nu een groot verschil? Licht je antwoord toe.

Opgave 11

In een doos zitten 30 uiterlijk allemaal dezelfde bonbons. Vijf bonbons hebben echter een roomvulling, de andere een caramelvulling. Uit de doos worden vier bonbons genomen.

a Hoe groot is de kans dat er precies één bonbon met een roomvulling uit de doos wordt gehaald? Rond af op vier decimalen nauwkeurig.

b Hoe groot is de kans dat er twee of meer bonbons met roomvulling uit de doos worden gehaald? Rond af op vier decimalen nauwkeurig.

c Hoe groot is de kans dat van de vier uitgenomen bonbons op één na allemaal een roomvulling hebben? Rond af op vier decimalen nauwkeurig.

Opgave 12

Een partij van 1000 blikken met groente heeft lange tijd in een magazijn gelegen. Je mag aannemen dat van 10% van de blikken de uiterste verkoopdatum verstreken is. Je kiest aselect 8 blikken uit de partij en controleert de verkoopdatum. Je vraagt je af hoe groot de kans is dat je in die steekproef 3 blikken aantreft die te oud zijn.

a Hoe groot is de genoemde kans? Rond af op vier decimalen nauwkeurig.

b Bereken deze kans ook met het binomiale kansmodel in vier decimalen nauwkeurig. Hoe groot is het verschil tussen beide berekeningen?

c Bereken in drie decimalen nauwkeurig de kans dat je maximaal 3 blikken gekozen hebt waarvan de uiterste verkoopdatum verstreken is.

Opgave 13

Een grote partij wijnflessen wordt gekeurd door uit de partij een aselecte steekproef van 20 flessen te nemen. Elke fles wordt nauwkeurig onderzocht op gebreken. Wordt er in de steekproef meer dan één fles met gebreken gevonden, dan wordt de gehele partij afgekeurd. Als er maximaal één fles wordt gevonden die niet voldoet, dan wordt de gehele partij goedgekeurd.

a Hoe groot is de kans dat de partij wordt goedgekeurd als 5% van de gehele partij flessen gebreken vertoont? Rond af op vier decimalen nauwkeurig.

b Hoe groot is de kans dat de partij wordt goedgekeurd als $\frac{1}{5}$ van de totale partij gebreken heeft? Rond af op vier decimalen nauwkeurig.

Ondanks de controle blijkt uiteindelijk 2% van de goedgekeurde wijnflessen alsnog gebreken te hebben. Dit komt helaas pas aan het licht als het te laat is.
Een groot restaurant heeft 250 wijnflessen op voorraad. Op een dag wordt een bedrijfsdiner in het restaurant geserveerd. Hiervoor worden acht flessen wijn opzij gelegd.

c Bereken de kans dat één van deze flessen een gebrek heeft, in vier decimalen nauwkeurig.

Niels gaat over de boekhouding bij het restaurant. Omdat het sneller is, gebruikt Niels als het kan liever een binomiaal verdeelde benadering dan een hypergeometrische verdeling. De baas van Niels is echter niet zo optimistisch over de situatie en laat een andere boekhouder de getallen nog eens nakijken. Als de kans van de benadering binnen 0,5% van de eigenlijke kans komt, is de baas tevreden.

d Onderzoek of dit bij de situatie bij c het geval is.

Toepassen

Opgave 14: Steekproefgrootte berekenen

Van een grote populatie is bekend dat 35% een bepaalde eigenschap bezit. Uit deze populatie wordt een willekeurige groep van 100 mensen gekozen.

a De kans dat in deze steekproef minder mensen aangetroffen worden met die eigenschap is 15%. Bepaal het maximale aantal mensen in de steekproef met die eigenschap.

Van een andere populatie is bekend dat $\frac{1}{6}$ een bepaalde eigenschap bezit. Uit deze populatie wordt een steekproef getrokken. De kans dat in deze steekproef hoogstens drie elementen worden aangetroffen met die eigenschap is 0,75.

b Bepaal de grootte van de steekproef.

Opgave 15: Steaks

Van een lading van meer dan 20 steaks is bekend dat er twee bedorven zijn. Er worden willekeurig vijf steaks uit de lading gekozen. Twee ervan blijken bedorven. Wat is de minimale grootte van de lading steaks zodanig dat de kans op deze hypergeometrisch verdeelde gebeurtenis hooguit 0,25% verschilt met de binomiale benadering?

Testen

Opgave 16

In een vaas zitten vijf balletjes genummerd 2, 4, 6, 8 en 10. Er worden zonder teruglegging twee balletjes uit de vaas getrokken. Stochast V is het verschil van de nummers van de twee balletjes.

a Stel de kansverdeling voor V op.

b Bereken zonder grafische rekenmachine de verwachtingswaarde, de variantie en de standaardafwijking.

Opgave 17

Bij een experiment heb je de beschikking over vijf vrouwelijke en vijf mannelijke proefpersonen. Je verdeelt ze willekeurig in twee groepen A en B van elk vijf personen.

a Hoe groot is de kans dat in groep A minstens vier vrouwen terechtkomen? Bereken deze kans met een hypergeometrisch kansmodel en benader de kans daarna met een binomiaal kansmodel. Rond indien nodig af op vier decimalen nauwkeurig.

b Welk van beide antwoorden op de vraag in a is juist? Licht je antwoord toe.

Opgave 18

Op zaterdagavond zit Jos, die elke week in de Lotto meespeelt, gespannen voor de tv om de trekking van de zes getallen mee te maken. (Het zogenaamde reservegetal wordt buiten beschouwing gelaten.) Er zitten 41 balletjes met daarop de getallen 1 tot en met 41 in een ronddraaiende trommel waaruit telkens één balletje wordt getrokken. Als de zes getrokken getallen overeenkomen met de getallen op zijn lot, ongeacht de volgorde, heeft Jos de hoofdprijs gewonnen.

a Hoe groot is de kans dat er zes even nummers worden getrokken? Rond af op vier decimalen nauwkeurig.

b Als er twee even nummers zijn getrokken, hoe groot is dan nog de kans dat de volgende vier balletjes ook een even nummer hebben? Rond af op vier decimalen nauwkeurig.

c Hoe groot is de kans, dat elk van de zes getrokken getallen kleiner is dan 15? Rond af op vier decimalen nauwkeurig.

Jos heeft de nummers 5, 10, 15, 20, 25 en 30 op zijn lot aangekruist.

d Hoe groot is de kans dat hij de hoofdprijs wint?

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.