9.8 Chi-kwadraat toetsen*

In paragraaf 9.7 is de Chi-kwadraat toets voor onafhankelijkheid bij kruistabellen behandeld. Hier behandelen we de Chi-kwadraat toets voor het vergelijken van percentages uit verschillende steekproeven, wat uiteindelijk ook weer het toetsen van onafhankelijkheid bij een kruistabel zal blijken te zijn. Vervolgens behandelen we de Chi-kwadraat toets voor het toetsen van een (kans)verdeling.

* Nummering sluit aan bij het leerboek.

9.8.1 Chi-kwadraattoets voor het vergelijken van percentages uit verschillende steekproeven

We gaan de uitkomsten van steekproeven vergelijken.

Voorbeeld 9.11

Van een steekproef van 145 bedrijven in Zeeland maakte 19% verlies. Van een steekproef van 609 bedrijven in Noord-Holland maakte 16% verlies. Kunnen we nu stellen dat er verschil is in prestaties tussen de provincies?

Op het eerste gezicht lijkt het niet om een kruistabel te gaan maar die kunnen we wel maken. Bijvoorbeeld: 19% van 145 is 28 bedrijven enzovoort. We krijgen dan de volgende kruistabel.

Tabel 9.6 **Kruistabel**

	Winst	Verlies	Totaal
Zeeland	117	28	145
Noord-Holland	512	97	609
Totaal	629	125	754

Nu we zien dat het weer een kruistabel betreft, kunnen we op dezelfde wijze als in voorbeeld 9.10 (op p. 246 leerboek, echter verkeerd aangeduid met 9.7) toetsen op onafhankelijkheid (een associatietoets).

We toetsen de hypothese:

Er is geen verschil tussen de twee provincies wat betreft het verlies maken of winst boeken van bedrijven.

Naast de waargenomen frequenties in tabel 9.6 maken we weer een tabel met de verwachte frequenties.

Tabel 9.7 **Verwachte frequenties**

	Winst	Verlies	Totaal
Zeeland	121	24	145
Noord-Holland	508	101	609
Totaal	629	125	754

Toelichting op de tabel met verwachte frequenties:

$\frac{629}{754}$ = 83,4% deel van alle bedrijven maakt winst.

Als de hypothese juist is, is er *geen* verschil tussen de provincies, en geldt dit percentage voor zowel Zeeland als Noord-Holland. We kunnen dan voor Zeeland en Noord-Holland het verwachte aantal winstgevende bedrijven berekenen:

Voor Zeeland: $\frac{629}{754} \times 145 = 121$ bedrijven met winst.

Voor Noord-Holland: $\frac{629}{754} \times 609 = 508$ bedrijven met winst.

We berekenen $X^2 = \frac{(117 - 121^2}{121} + \frac{(512 - 508)^2}{508} + \frac{(28 - 24)^2}{24} + \frac{(97 - 101)^2}{101} = 0{,}99$

Het aantal vrijheidsgraden is n = (2 – 1) × (2 – 1) = 1 × 1 = 1.

Kiezen we de betrouwbaarheid 95%, dan vinden we in de tabel de grenswaarde 3,84.

Tabel 9.8 **Enkele kritieke grenzen voor de *X*²-verdeling** (aangegeven zijn: grenswaarden g_2 met $P(X^2 < g_2) = \alpha$,*n* is het aantal vrijheidsgraden)

n	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
$g_{0,01}$	–	0,02	0,11	0,30	0,55	0,87	1,24	1,65	2,09	2,56
$g_{0,025}$	–	0,05	0,22	0,48	0,83	1,24	1,69	2,18	2,70	3,25
$g_{0,05}$	–	0,10	0,35	0,71	1,15	1,64	2,17	2,73	3,33	3,94
$g_{0,95}$	3,84	5,99	7,81	9,49	11,07	12,59	14,07	15,51	16,92	18,31
$g_{0,975}$	5,02	7,38	9,35	11,41	12,83	14,45	16,01	17,53	19,02	20,48
$g_{0,99}$	6,63	9,21	11,34	13,28	15,09	16,81	18,48	20,09	21,67	23,21

n	**12**	**14**	**16**	**18**	**20**	**25**	**30**	**50**	**100**
$g_{0,01}$	3,57	4,66	5,81	7,01	8,26	11,52	14,95	29,71	70,06
$g_{0,025}$	4,40	5,63	6,91	8,23	9,59	13,12	16,79	32,36	74,22
$g_{0,05}$	5,23	6,57	7,96	9,39	10,85	14,61	18,49	34,76	77,93
$g_{0,95}$	21,03	23,68	26,30	28,87	31,41	37,65	43,77	67,50	124,34
$g_{0,975}$	23,34	26,12	28,85	31,53	34,17	40,65	46,98	71,42	129,56
$g_{0,99}$	26,22	29,14	32,00	34,81	37,57	44,31	50,89	76,15	135,81

De berekende waarde van $X^2 = 0,99$, ligt onder de grenswaarde (0,99 < 3,84). Conclusie: naar aanleiding van de steekproef is er geen reden om aan te nemen dat er verschil in prestatie is tussen beide provincies (met een betrouwbaarheid van 95%).

9.8.2 Toetsen van een (kans)verdeling

In paragraaf 9.4 is gezegd dat er geen echt verschil is tussen het toetsen van een populatieproportie en het toetsen van een binomiale kans. In die paragraaf is er geen expliciet voorbeeld gegeven van het laatste.
In deze subparagraaf gaan we niet één kans toetsen, maar meerdere kansen tegelijk, een (kans)verdeling. Hier geven we zowel een voorbeeld met een theoretische kansverdeling als met een empirische (kans)verdeling. We zullen zien dat dit voor het toetsen verder geen verschil maakt.

In voorbeeld 9.7 op p. 242 leerboek hebben we een slagingskans getoetst, $p = 0{,}7$. Hier toetsen we niet één kans, maar een hele (kans)verdeling. Deze toets wordt ook wel goodness-of-fit toets genoemd. De toets onderzoekt of de uitkomsten van een steekproef goed passen bij de veronderstelde (kans)verdeling.

Voorbeeld 9.12 Theoretische kansverdeling

We gaan toetsen of een dobbelsteen zuiver is.

We kunnen de te toetsen hypothese als volgt opschrijven:
nulhypothese: de dobbelsteen is zuiver.
(De alternatieve hypothese is dat de dobbelsteen niet zuiver is.)
Maar we kunnen de hypothese ook in de vorm van een kansverdeling opschrijven:

nulhypothese: $p_1 = \frac{1}{6}$, $p_2 = \frac{1}{6}$, $p_3 = \frac{1}{6}$, $p_4 = \frac{1}{6}$, $p_5 = \frac{1}{6}$ en $p_6 = \frac{1}{6}$.

Om de hypothese te toetsen, hebben we 60 keer gegooid, een steekproef met grootte 60. De uitkomst staat in tabel 9.9.

Tabel 9.9 **Waargenomen frequenties**

Aantal ogen	1	2	3	4	5	6
Waargenomen frequentie	7	10	12	9	9	13

Als de hypothese ‘de dobbelsteen is zuiver’ of anders opgeschreven

$p_1 = \frac{1}{6}$, $p_2 = \frac{1}{6}$, $p_3 = \frac{1}{6}$, $p_4 = \frac{1}{6}$, $p_5 = \frac{1}{6}$ en $p_6 = \frac{1}{6}$ waar is, dan kunnen we ook een tabel met verwachte frequenties opstellen.

Bij 60 keer gooien verwachten we dat elk aantal ogen (gemiddeld) $1/6 \times 60 = 10$ keer gegooid wordt. Dus 10 keer een 1, 10 keer een 2 enzovoort.
Ook dit zetten we in een tabel.

Tabel 9.10 **Verwachte frequenties**

Aantal ogen	1	2	3	4	5	6
Verwachte frequentie	10	10	10	10	10	10

We gaan nu een zogenoemde toetsingsgrootheid X^2 berekenen.

De formule hiervoor is weer $X^2 = \Sigma \frac{(W - V)^2}{V}$

waarin:

W = waargenomen frequentie

V = de verwachte frequentie.

Vullen we de formule in, dan krijgen we:

$$X^2 = \frac{(7-10)^2}{10} + \frac{(10-10)^2}{10} + \frac{(12-10)^2}{10} + \frac{(9-10)^2}{10} + \frac{(9-10)^2}{10} + \frac{(13-10)^2}{10}$$

X^2 = 0,9 + 0 + 0,4 + 0,1 + 0,1 + 0,9

= 2,4.

We gaan deze waarde weer vergelijken met de grenswaarde uit de tabel.

Voordat we de grenswaarde in de tabel kunnen opzoeken, hebben we nog twee gegevens nodig: het aantal vrijheidsgraden en de betrouwbaarheid.

Het *aantal vrijheidsgraden* bij deze toets bepalen we als volgt. Als we zoals hier een kansverdeling met 6 kansen hebben, dan zou na 5 vrij gekozen kansen de 6e vastliggen, de som is immers 1. Dit aantal van 5 heet het aantal vrijheidsgraden (n). Dus het aantal vrijheidsgraden is n = k -/- 1, k is het aantal kansen of klassen (het aantal ogen kun je als een klasse zien).

We hebben ook nog de *betrouwbaarheid* nodig. De betrouwbaarheid is hier niet gegeven. We kiezen weer de veel genomen betrouwbaarheid van 95%.

We kunnen nu in tabel 9.11 bij n = 5 en $g_{0,95}$ de grenswaarde voor X^2 vinden, namelijk 11,07.

Tabel 9.11 **Enkele kritieke grenzen voor de X^2-verdeling** (aangegeven zijn: grenswaarden g_2 met $P(X^2 < g_2) = \alpha$,*n* is het aantal vrijheidsgraden)

n	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
$g_{0,01}$	–	0,02	0,11	0,30	0,55	0,87	1,24	1,65	2,09	2,56
$g_{0,025}$	–	0,05	0,22	0,48	0,83	1,24	1,69	2,18	2,70	3,25
$g_{0,05}$	–	0,10	0,35	0,71	1,15	1,64	2,17	2,73	3,33	3,94
$g_{0,95}$	3,84	5,99	7,81	9,49	11,07	12,59	14,07	15,51	16,92	18,31
$g_{0,975}$	5,02	7,38	9,35	11,41	12,83	14,45	16,01	17,53	19,02	20,48
$g_{0,99}$	6,63	9,21	11,34	13,28	15,09	16,81	18,48	20,09	21,67	23,21

n	**12**	**14**	**16**	**18**	**20**	**25**	**30**	**50**	**100**
$g_{0,01}$	3,57	4,66	5,81	7,01	8,26	11,52	14,95	29,71	70,06
$g_{0,025}$	4,40	5,63	6,91	8,23	9,59	13,12	16,79	32,36	74,22
$g_{0,05}$	5,23	6,57	7,96	9,39	10,85	14,61	18,49	34,76	77,93
$g_{0,95}$	21,03	23,68	26,30	28,87	31,41	37,65	43,77	67,50	124,34
$g_{0,975}$	23,34	26,12	28,85	31,53	34,17	40,65	46,98	71,42	129,56
$g_{0,99}$	26,22	29,14	32,00	34,81	37,57	44,31	50,89	76,15	135,81

De berekende waarde $X^2 = 2{,}4$ ligt onder deze grens (2,4 < 11,07).
De statistische conclusie is: we aanvaarden de nulhypothese.
Conclusie: naar aanleiding van de steekproef geloven we dat de dobbelsteen zuiver is met een betrouwbaarheid van 95%.

NB. Let er bij een Chi-kwadraat toets op dat de verwachte frequenties niet kleiner dan 5 zijn.

In voorbeeld 9.12 hebben we een theoretische kansverdeling getoetst. In voorbeeld 9.13 toetsen we een empirische (kans)verdeling. Voor de manier van toetsen maakt dit op zich geen verschil.

Voorbeeld 9.13 Empirische (kans)verdeling

Uit het verleden is bekend dat van de afnemers van een groot bedrijf 15% binnen 15 dagen betaalt, 75% betaalt vanaf de 15e dag tot aan de 30e dag en 10% doet er 30 dagen of meer over. We willen toetsen of deze percentages nog steeds gelden.

Het resultaat van een steekproef van 200 afnemers staat in tabel 9.12.

Tabel 9.12

Klasse	< 15 dagen	15-< 30 dagen	≥ 30 dagen
Waargenomen frequentie	27	141	32

Als de hypothese ‘de percentages zijn nog hetzelfde’ of
$p_1 = 0{,}15$, $p_2 = 0{,}75$ en $p_3 = 0{,}10$ waar is, dan kunnen we weer een tabel met verwachte frequenties opstellen.

Tabel 9.13

Klasse	< 15 dagen	15-< 30 dagen	≥ 30 dagen
Verwachte frequentie	30	150	20

We berekenen weer:

$$X^2 = \frac{(27-30)^2}{30} + \frac{(141-150)^2}{150} + \frac{(32-20)^2}{20}$$

$= 0{,}3 + 0{,}54 + 7{,}2$

$= 8{,}04$.

Ook nu moeten we weer het aantal vrijheidsgraden en de betrouwbaarheid bepalen. We hebben hier te maken met n = 3 -/- 1 = 2 vrijheidsgraden. Voor de betrouwbaarheid kiezen we 95%. Dan vinden we bij n = 2 en $g_{0,95}$ de grenswaarde voor X^2 in tabel 9.5: 5,99.

De berekende waarde van X^2, 8,04, overschrijdt deze waarde (8,04 > 5,99). Conclusie: voor het betaalgedrag van de afnemers gelden de percentages uit het verleden niet meer. (De conclusie wordt getrokken met een betrouwbaarheid van 95%.)