

Oplossingen hoofdstuk 9

1. Bestaat er een verband tussen het geslacht en het voorkomen van dyslexie?

Uit een aselechte steekproef van 200 leerlingen (waarvan 50% jongens en 50% meisjes) uit het basisonderwijs (allen 11 à 12 jaar) komen volgende gegevens naar voren:

Jongens: 10 deelnemers vertonen dyslexie.

Meisjes: 2 deelnemers vertonen dyslexie.

Stel een kruistabel op.

	Dyslexie	niet dyslexie	
Meisjes	2	98	100
Jongens	10	90	100
	12	188	200

Bestaat er een betekenisvol verband tussen het geslacht en het voorkomen van dyslexie?

Gebruik de Chi-kwadraattoets voor kruistabellen.

H₀ er bestaat geen verband tussen beide variabelen

H_A er bestaat een verband tussen beide variabelen

Test:

Fe:	6	94
	6	94

Berekening Chi-kwadraat: $16/6 + 16/6 + 16/94 + 16/94 = 5,68$

Vergelijk dit met de kritische waarde bij 1 vrijheidsgraad op 5% niveau. Deze kritische waarde is 3,84. Dus de gevonden Chi-kwadraat is groter dan deze 3,84.

Vandaar dat de H₀ verworpen wordt. Er bestaat een significante samenhang tussen de twee variabelen.

Indien ja, hoe sterk is dit verband?

Gebruik de Cramer's V. en de Phi-coëfficiënt. Wat blijkt?

$V = \text{wortel uit } 5,68/200 = 0,17$

$\text{Phi} = \text{wortel uit } 5,68/200 = 0,17$

De Cramer's V en de Phi-coefficient zijn aan elkaar gelijk.

Dit is enkel het geval in een viervelden tabel.

Formuleer uw besluit hierover.

Er bestaat een significante samenhang. Maar deze samenhang is eerder matig.

Werk deze opdracht handmatig uit en controleer vervolgens via SPSS of de gevonden oplossing klopt.

gesl * status Crosstabulation

Count		status		Total
		dyslexie	niet dyslexie	
gesl	meisje	2	98	100
	jongen	10	90	100
Total		12	188	200

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	5,674 ^b	1	,017		
Continuity Correction ^a	4,344	1	,037		
Likelihood Ratio	6,163	1	,013		
Fisher's Exact Test				,033	,017
Linear-by-Linear Association	5,645	1	,018		
N of Valid Cases	200				

a. Computed only for a 2x2 table

b. 0 cells (,0%) have expected count less than 5. The minimum expected count is 6,00.

Symmetric Measures

		Value	Approx. Sig.
Nominal by Nominal	Phi	-,168	,017
	Cramer's V	,168	,017
	Contingency Coefficient	,166	,017
N of Valid Cases		200	

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

We kunnen vaststellen dat SPSS gelijklopende uitkomsten bekomt. Via deze output lezen we eveneens dat de Chi-kwadraat geïnterpreteerd mag worden. We stellen eveneens vast dat de Phi waarde een negatief teken bevat. Eerder hebben we gezegd dat dit teken geen betekenis heeft en afhankelijk is van de wijze waarop de kruistabel opgesteld wordt.

2. Herneem de gegevens over de hart- en vaatziekten bij rokers en niet rokers (tabel 9.1). Voer deze gegevens in SPSS via de Weight Cases en onderzoek de samenhang. Bespreek de bekomen resultaten in termen van samenhang en sterkte van deze samenhang.

roker * hartprobl Crosstabulation

Count		hartprobl		Total
		hart en vaatziekten	geen hart en vaatziekten	
roker	1	5	15	20
	2	1	99	100
Total		6	114	120

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	20,211 ^b	1	,000		
Continuity Correction ^a	15,474	1	,000		
Likelihood Ratio	13,950	1	,000		
Fisher's Exact Test				,000	,000
Linear-by-Linear Association	20,042	1	,000		
N of Valid Cases	120				

a. Computed only for a 2x2 table

b. 1 cells (25,0%) have expected count less than 5. The minimum expected count is 1,00.

Symmetric Measures

		Value	Approx. Sig.
Nominal by Nominal	Phi	,410	,000
	Cramer's V	,410	,000
N of Valid Cases		120	

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

Uit de output blijkt dat we de Chi-kwadraat niet kunnen interpreteren. Dus op grond van deze gegevens kunnen we niet aantonen dat er een verband bestaat tussen het al dan niet roken en het voorkomen van hart en vaatziekten.

Het gering aantal respondenten in deze steekproef speelt een rol in het feit dat deze Chi-kwadraat niet kan geïnterpreteerd worden.

3. Van 120 studenten wordt bij het begin van het eerste jaar bachelor een motivatieproef afgenomen. Van deze studenten blijken er 80 werkelijk gemotiveerd te zijn. Op het einde van het academiejaar gaat men voor elke student het studieresultaat na. Van de gemotiveerde studenten zijn er 10 niet geslaagd, terwijl dit aantal 30 bedraagt bij de niet-gemotiveerde collega's.

Voer de gegevens in SPSS via het commando Weigth Cases.

Bestaat er een samenhang tussen de motivatie en het slagen/mislukken?

Als er een samenhang is bereken de sterkte van de samenhang.

In de Viewer ontstaat volgende output

gemotiv * slagen Crosstabulation

Count		slagen		Total
		geslaagd	niet geslaagd	
gemotiv	gemot	70	10	80
	niet gemot	10	30	40
Total		80	40	120

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	46,875 ^b	1	,000		
Continuity Correction ^a	44,105	1	,000		
Likelihood Ratio	47,493	1	,000		
Fisher's Exact Test				,000	,000
Linear-by-Linear Association	46,484	1	,000		
N of Valid Cases	120				

a. Computed only for a 2x2 table

b. 0 cells (,0%) have expected count less than 5. The minimum expected count is 13,33.

Symmetric Measures

		Value	Approx. Sig.
Nominal by	Phi	,625	,000
Nominal	Cramer's V	,625	,000
N of Valid Cases		120	

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

We stellen vast dat er een significant verband bestaat tussen de beide variabelen. De associatiematen – Phi en Cramer's V – tonen aan dat er een redelijke samenhang bestaat tussen de twee variabelen.

4. We onderzoeken of de steekproef voor een onderzoek naar leerproblemen een weerspiegeling is van de werkelijke samenstelling van de populatie. In de steekproef hebben we 100 algemeen secundair onderwijs (ASO) leerlingen, 50 technisch secundair onderwijs (TSO) en 50 uit het beroepssecundair onderwijs (BSO). Uit de gegevens van het ministerie van onderwijs blijkt dat er in Vlaanderen volgende % voorkomen:

ASO 35%, TSO 30% en BSO 35%.

Kunnen we stellen dat deze steekproef een goede weerspiegeling is van de populatie? Voer de gegevens via Weight Cases in en onderzoek de vraagstelling.

Na het invoeren van de data, en het betreffende commando ontstaat in de Viewer volgende output:

studie			
	Observed N	Expected N	Residual
ASO	100	70.0	30.0
TSO	50	60.0	-10.0
BSO	50	70.0	-20.0
Total	200		

Test Statistics	
	studie
Chi-Square ^a	20.238
df	2
Asymp. Sig.	.000

a. 0 cells (.0%) have expected frequencies less than 5. The minimum expected cell frequency is 60.0.

H0 stelt dat de steekproef een goede weerspiegeling is van de samenstelling van de populatie. HA stelt dat dit niet zo is.

Uit de output blijkt dat de Chi-kwadraat geïnterpreteerd kan worden, en dat deze significant van niveau is. We kunnen dus stellen dat deze steekproef geen goede afspiegeling is van de populatie.

5. Men gaat bij 120 studenten na of er een voorkeur bestaat voor het wonen:

A in een grote stad

B in een kleine stad

C op het platteland

Uit de gegevens van het onderzoek blijkt:

52 studenten hebben een voorkeur voor A

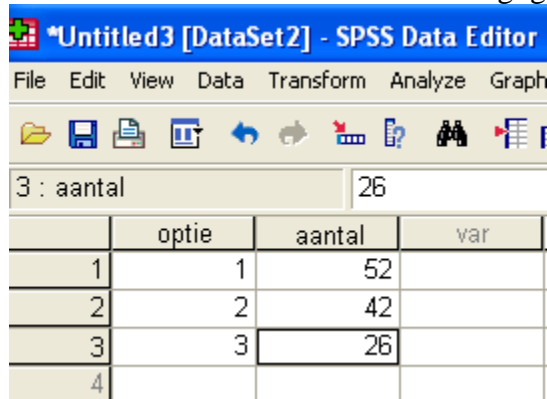
42 studenten hebben een voorkeur voor B

26 studenten hebben een voorkeur voor C

Bestaat er een voorkeur voor één van deze drie opties bij de studenten?

Voer de gegevens in via Weight Cases en onderzoek de vraagstelling.

De invoer in de data editor kan als volgt gebeuren:



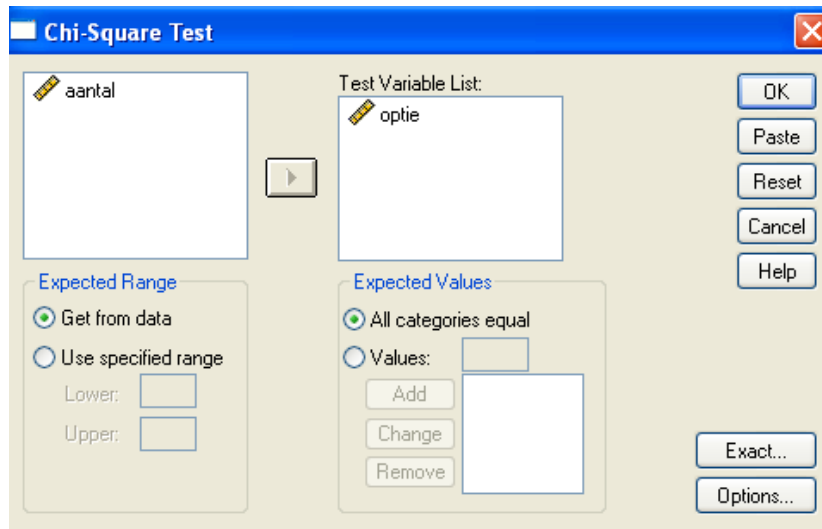
The screenshot shows the SPSS Data Editor interface. At the top, the title bar reads '*Untitled3 [DataSet2] - SPSS Data Editor'. Below it is a menu bar with 'File', 'Edit', 'View', 'Data', 'Transform', 'Analyze', and 'Graph'. A toolbar contains various icons for file operations and data manipulation. A data entry window is open, showing '3 : aantal' with the value '26' entered. Below this is a data table with the following structure:

	optie	aantal	var
1	1	52	
2	2	42	
3	3	26	
4			

Nadien kunnen via Data – Weight Cases de drie regels gewogen worden in functie van het aantal.

Voor het opvragen van de Chi-kwadraat kiezen we voor Analyze – Nonparametric Tests – Chi-Square.

In het dialoogvenster Chi-Square Test kiezen we voor de variabele ter studie en we duiden vervolgens aan 'All categories equal', overeenkomstige de nulhypothese dat alle opties even vaak voorkomen. De volgende figuur toont hoe het moet.



Dit levert volgende output op in de viewer:

optie

	Observed N	Expected N	Residual
1	52	40,0	12,0
2	42	40,0	2,0
3	26	40,0	-14,0
Total	120		

Test Statistics

	optie
Chi-Square ^a	8,600
df	2
Asymp. Sig.	,014

a. 0 cells (.0%) have expected frequencies less than 5. The minimum expected cell frequency is 40,0.

We kunnen deze Chi-kwadraat interpreteren. Deze blijkt significant te zijn (omdat .014 kleiner is dan .05). We kunnen de nulhypothese verwerpen. We zeggen dat de studenten wel degelijk een voorkeur hebben in deze kwestie.

6. Wie is er delinquent?

We stellen de kruistabel op

	delinquent	niet delinquent	totaal
Jongens	5	185	190
Meisjes	1	209	210

Via Weight Cases worden de data ingevoerd in SPSS

Eerst wordt de Variable View ingevuld met drie variabelen: geslacht, status en aantal

Nadien wordt de Data View ingevuld als volgt

*Untitled1 [DataSet0] - IBM SPSS Statistics Data Editor

	geslacht	status	aantal	va
1	jongen	delinquent	5	
2	jongen	niet delinq...	185	
3	meisje	delinquent	1	
4	meisje	niet delinq...	209	
5				

Nadien kun je via Data – Weight Cases - Weight cases by the aantal de kruistabel formeren. Nadien vragen we de chikwadraat op

Chi-Square Tests

	Value	df	Asymptotic Significance (2- sided)	Exact Sig. (2- sided)	Exact Sig. (1- sided)
Pearson Chi-Square	3,136 ^a	1	,077		
Continuity Correction ^b	1,847	1	,174		
Likelihood Ratio	3,373	1	,066		
Fisher's Exact Test				,106	,086
Linear-by-Linear Association	3,129	1	,077		
N of Valid Cases	400				

a. 2 cells (50,0%) have expected count less than 5. The minimum expected count is 2,85.

b. Computed only for a 2x2 table

Wat blijkt uit deze resultaten?

50% van de cellen van expected count zijn kleiner dan 5, vandaar dat deze chikwadraat niet kan geïnterpreteerd worden.

Dus we kunnen niets besluiten over een mogelijke samenhang in de populatie.

Het berekenen van de associatiematen heeft dan geen zin.

7. Bestaat er een samenhang tussen geslacht en slagen en mislukken?

Deze chikwadraat mag geïnterpreteerd worden, gezien de voetnoot onder de cijfers.

Wat blijkt uit het resultaat?

Het resultaat blijkt niet significant te zijn (.764). Dat wil zeggen dat er niet aangetoond werd dat er een verband bestaat tussen geslacht en slagen en mislukken. Jongens en meisjes hebben evenveel kans om te slagen.

8. Bestaat er een verband tussen geslacht en dyspraxie?

Deze chikwadraat mag geïnterpreteerd worden.

In dit geval bestaat er een significant verband (.014) tussen geslacht en dyspraxie.

In dit geval wordt de nulhypothese (er is geen verband) verworpen. Er bestaat in de populatie een verband tussen geslacht en dyspraxie.

9. Het aantal vrijheidsgraden is in deze situatie gelijk aan 2. Dit resultaat wordt bereikt door $(3-1)*(2-1) = 1$.