

SVR - Methoden en technieken 2010 / 1

Standaardisatie van een statistische waarneming voor één of meerdere kenmerken

Edwin Pelfrene

Studiedienst van de Vlaamse Regering

Vlaamse overheid



Standaardisatie van een statistische waarneming voor één of meerdere kenmerken

Edwin Pelfrene



Samenstelling

Diensten voor het Algemeen Regeringsbeleid
Studiedienst van de Vlaamse Regering (SVR)

Edwin Pelfrene

Leescomité

Marc Callens, wetenschappelijk attaché-expert SVR
Jan Pickery, adviseur SVR

Heidi Cloots, statisticus gezondheidsindicatoren, Vlaams agentschap Zorg en Gezondheid

Verantwoordelijke uitgever

Josée Lemaître
Administrateur-generaal
Boudewijnlaan 30 bus 23
1000 Brussel

Lay-out cover

Diensten voor het Algemeen Regeringsbeleid
Communicatie
Patricia Van Dichel

Druk

Agentschap voor Facilitair Management

Depotnummer

D/2010/3241/054

<http://www.vlaanderen.be/svr>

Korte samenvatting

Standaardisatie van een statistische waarneming voor één of meerdere kenmerken is een kunstgreep om *via weging te corrigeren* voor ongelijke verdelingen tussen groepen naar die kenmerken. De techniek is zeer gebruikelijk in vakgebieden als de demografie of de epidemiologie, bijvoorbeeld om de waargenomen sterfte in deelpopulaties goed te vergelijken. Vaak gaat het daarbij om standaardisatie voor leeftijd.

Bedoeling van dit rapport is om de techniek van standaardisatie van verdelingen toe te lichten aan de hand van enkele voorbeelden. Algemeen wordt onderscheid gemaakt tussen *directe* en *indirecte standaardisatie*. Bij de eerste techniek wordt uitgegaan van de waargenomen risico's op de uitkomst van belang (sterfte) in de deelgroepen (leeftijdsgroepen) van elk der onderscheiden vergelijkingsgroepen (stad/platteland), die vervolgens gewogen worden aan een goedgekozen standaard. De tweede techniek vertrekt niet van die waargenomen risico's, hetzij omdat informatie daarover ontbreekt of onbetrouwbaar is (onder meer als gevolg van een te klein aantal waarnemingen). In dat geval wordt een ratio berekend, met in de teller de waargenomen uitkomst van belang en in de noemer de verwachte uitkomst van belang in elk der vergelijkingsgroepen op basis van de risico's per deelgroep zoals waargenomen in de standaard.

De techniek van directe standaardisatie voor één kenmerk kan in het geval van een categorische uitkomst (wel/geen overlijden) vrij gemakkelijk uitgebreid worden naar meerdere kenmerken toe met behulp van logistische regressie. Hoe de vergelijking opstellen om gestandaardiseerde percentages te schatten wordt geïllustreerd. Aanvullend wordt nog ingegaan op de bepaling van het betrouwbaarheidsinterval rond die puntschatters.

Abstract

Standardisation as a technique *to control* for a potential confounder of the relationship between an explanatory variable and a single response variable is widely applied in different fields of research, particularly in demography and epidemiology. The technique seems most justified in cases of well known effects of the potential confounder on the relationship under study, e.g. the disturbing effect of the age-distribution on the relationship between a contagious agent and the disease.

Our aim is to present and clarify the technique of standardisation with the help of a few examples. In general, two types are distinguished: *direct* and *indirect standardisation*. The first technique re-weights observed risk factors within subgroups of the control variable with the chosen standard. The second technique does not rely on the risk factors within these subgroups, either for lack of information or for lack of reliability. Actually, a ratio is calculated, comparing the observed number having the characteristic of interest to what could be expected if the risk factors of the chosen standard would prevail in the study population. Most used in this respect are age-standardised mortality ratios (SMRs).

It is also shown that direct standardisation of proportions can easily be extended to a number of potentially confounding characteristics with the help of logistic regression. How to assess standardised log odds and how to translate these to standardised rates is demonstrated, together with some hints on how to calculate associated confidence intervals.

Dankwoord

Graag een bijzonder woord van dank voor de raad van twee mathematici, altijd onmisbare spitsen op dit veld, namelijk mevrouw Liesbeth Bruckers van het Centrum voor Statistiek van de Universiteit Hasselt, en Prof. Dr. Dirk De Bacquer, voorzitter van de Vakgroep Maatschappelijke Gezondheidskunde van de Universiteit Gent.

Tevens onze dank aan het Wetenschappelijk Instituut Volksgezondheid (WIV) van de FOD Volksgezondheid, Veiligheid van de voedselketen en Leefmilieu, voor het beschikbaar stellen van het bestand van de nationale gezondheidsenquête.

INHOUDSTAFEL

1.	Inleiding	5
2.	Directe standaardisatie	5
2.1.	Basismethode.....	5
2.1.1.	Voorbeeld 1: aanvragen voor een griepvaccin.....	5
2.1.2.	Voorbeeld 2: houding tegenover het dragen van de hoofddoek.....	7
2.2.	Uitbreiding via logistische regressie.....	11
2.2.1.	Werkwijze.....	11
2.2.2.	Voorbeeld 3: het gebruik van psychotrope geneesmiddelen.....	13
3.	Indirecte standaardisatie	18
3.1.	Werkwijze.....	18
3.2.	Voorbeelden.....	19
3.2.1.	Voorbeeld 4: kans op tenlasteneming door de Vlaamse zorgverzekering.....	19
3.2.2.	Voorbeeld 5: oversterfte in Vlaamse gemeenten.....	22
4.	Besluit	26
	LITERATUUR	27
	ADDENDUM	28
A1.	Bepaling ASE van $\text{logit}(p)$ via de variantie-covariantie matrix.....	28
A2.	Schatting van ASE van $\text{logit}(p)$ via transformaties van verklarende variabelen en herberekening van het model.....	30
A3.	Bepaling van het betrouwbaarheidsinterval voor SMR.....	32

1. Inleiding

Standaardisatie van een statistische waarneming voor één of meerdere kenmerken is een kunstgreep om via *weging* te corrigeren voor ongelijke verdelingen tussen groepen naar die kenmerken. De techniek is zeer gebruikelijk in vakgebieden als de demografie of de epidemiologie, bijvoorbeeld om de waargenomen sterfte in deelpopulaties goed te vergelijken. Vaak gaat het daarbij om standaardisatie voor leeftijd.

Anders dan bij de weging van een steekproef met het oog op een grotere representativiteit of een betere afspiegeling van de populatie, mikt de weging bij standaardisatie op het wegwerken van verschillen in de verdeling van kenmerken in subgroepen van de voorliggende dataset. Niet representativiteit maar causaliteit is de centrale zorg. Immers, zoals bij het experiment in een laboratorium wordt getracht de vergelijkingsgroepen slechts op één (oorzakelijke) factor te laten variëren, met gelijkschakeling van alle andere (belangrijk geachte) omstandigheden, cf. de zogenaamde *ceteris paribus clausule*.

Twee methodes voor standaardisatie staan ter beschikking: de directe en de indirecte standaardisatie.

2. Directe standaardisatie

2.1. Basismethode

2.1.1. Voorbeeld 1: aanvragen voor een griepvaccin

Vertrekken we van een voorbeeld, i.c. het 'aantal aanvragen voor een griepvaccin' (GV) in de studie- of *indexpopulaties* P1 en P2, resp. met omvang N1 en N2.

Tabel 1a Basisobservaties in de populaties P1 en P2

	P1				P2		
	N1	GV	Prop. GV*		N2	GV	Prop. GV*
Totaal	30	21	0,70	60	34	0,57	

*Prop. GV = proportie dat griepvaccin aanvraag

In P1 hebben 21 van de 30 leden het griepvaccin aangevraagd, in P2 is dat 34 van de 60 leden. Omgerekend naar (globale) kansen of risico's vinden we resp. 0,70 en 0,57. Of in percentages uitgedrukt: 70% versus 57%.

Een opsplitsing naar leeftijdsgroepen leert:

Tabel 1b Berekening van leeftijdsspecifieke proporties GV in P1 en P2

Leeftijdsgroep	P1				P2		
	N1	GV	Prop. GV		N2	GV	Prop. GV
Jong	4	2	0,50	40	20	0,50	
Middelbaar	6	4	0,67	12	8	0,67	
Oud	20	15	0,75	8	6	0,75	
Totaal	30	21	0,70	60	34	0,57	

Duidelijk is dat de kans op het aanvragen van een griepvaccin stijgt met de leeftijdsgroep, en wel op helemaal gelijke wijze in de beide studiepopulaties. Anders dan het *globale risico* verschillen de *leeftijdsspecifieke risico's* niet tussen P1 en P2.

Het verschil in het globale risico van de beide populaties wordt zodoende helemaal verklaard door hun verschil in leeftijdsverdeling: P1 vertoont een hoger globaal risico dan P2, precies omdat in P1 de *laagrisicogroep* ('jong') minder weegt en de *hoogrisicogroep* ('oud') zwaarder doorweegt dan in P2.

Welnu, '*standaardisatie voor een kenmerk*' is een techniek die erop gericht is om het *vertekenend effect* van de verdeling van dat kenmerk op het *globale risico* van een uitkomstmaat weg te werken. Voor ons voorbeeld betekent dat het neutraliseren van het effect van de leeftijdsverdeling in de onderscheiden groepen P1 en P2 op de globale (kans tot) aanvraag voor een griepvaccin.

Bij *directe standaardisatie* voor een kenmerk is de werkwijze algemeen als volgt:

- (1) Vertrokken wordt van de *specifieke risico's* op de uitkomst van belang zoals waargenomen per deelgroep van het betreffende kenmerk in elk der studiepopulaties. In ons voorbeeld dus de waargenomen kansen (of risico's) op het aanvragen van het griepvaccin per leeftijdsgroep, resp. in P1 en in P2.
- (2) Die set van risico's (of risicovector) per studiepopulatie wordt vervolgens toegepast op een goedgekozen populatie, te weten de *referentiepopulatie* of de *standaard*. In de praktijk betekent dat, apart voor elke studiepopulatie, de vermenigvuldiging van het specifieke risico met de overeenkomstige deelgroep van de *standaardpopulatie*. Dit levert dan het *verwachte aantal* met de gemeten uitkomst in de betreffende deelgroep, te weten *verwacht in de standaardpopulatie* volgens de waargenomen specifieke risico's van de *studiepopulatie*.
- (3) Opgeteld over alle deelgroepen van het kenmerk levert dit het *totale aantal te verwachten gevallen* met de uitkomst van belang in de standaardpopulatie, en dit apart berekend per studiepopulatie. In ons voorbeeld is dat het verwachte aantal aanvragen voor een griepvaccin in de standaardpopulatie, apart volgens de risicovector van P1 en van P2.

Eender welke standaardpopulatie kan gekozen worden. Wel wordt onderscheid gemaakt tussen:

- a) de *interne standaard*, met als standaardpopulatie de som van de diverse studiepopulaties, uiteraard met sommaties per deelgroep. Voor ons voorbeeld dus de som van N1 en N2, per leeftijdsgroep;
- b) een *externe standaard*, vaak een relevante populatie waarvan de verdeling van het te controleren kenmerk gemakkelijk beschikbaar is. Bijvoorbeeld de leeftijdsverdeling van de bevolking van het Vlaamse Gewest per 1/01/2010.

In ons voorbeeld is geopteerd voor de *interne standaard* (S):

Tabel 1c Interne standaardpopulatie

<i>Leeftijdsgroep</i>	S = N1+N2
Jong	44
Midden	18
Oud	28
Totaal	90

Toepassing van de waargenomen risicovectoren van P1 en P2 op de standaard, levert:

Tabel 1d Aanvragen voor een griepvaccin na standaardisatie voor leeftijd volgens de directe methode

Leeftijdsgroep	S	P1-risicogroep			P2-risicogroep		
		prop. GV	GV	Prop. GV	prop. GV	GV	Prop. GV
Jong	44	0,50	22		0,50	22	
Midden	18	0,67	12		0,67	12	
Oud	28	0,75	21		0,75	21	
Totaal	90		55	0,611		55	0,611

Met de risico's per leeftijdsgroep zoals waargenomen in P1 worden in de standaardpopulatie bij de 'jongeren' 22 aanvragen ($= 0,50 \times 44$) voor een griepvaccin verwacht, bij de leeftijdsgroep 'midden' zijn er dat 12 ($= 0,67 \times 18$) en bij de 'ouderen' 21 ($= 0,75 \times 28$). Samengeteld geeft dat een totaal van 55 aanvragen in de standaardpopulatie volgens de risicovector van P1. Dit is net zo voor P2, wat uiteraard voortvloeit uit het gegeven dat in beide indexpopulaties de risicovectoren, te weten de set van risico's per leeftijdsgroep, gelijk zijn. Omgerekend is de kans op 'aanvraag voor een griepvaccin' gelijk aan 0,611 ($= 61,1\%$), zowel in P1 als in P2.

De conclusie voor ons voorbeeld is dat na standaardisatie voor de leeftijdsgroep er niet langer een verschil is tussen P1 en P2 wat de globale proportie GV betreft. Het aanvankelijk waargenomen verschil (0,70 in P1 versus 0,57 in P2) wordt geheel verklaard door het verschil in de leeftijdsgroepverdeling van P1 en P2.

Veelal wordt de techniek toegepast in situaties waarin leeftijd een sterk bepalende factor is voor het optreden van de uitkomst (van belang). Bij uitstek is dat het geval in studies naar oorzaken van ziekte (morbiditeit) en sterfte (mortaliteit). In mortaliteitsstudies heeft men het aldus over 'het sterftecijfer gestandaardiseerd voor leeftijd (volgens de directe methode)'. In het gebruikelijke Engels is dat de 'age-standardised mortality rate (by direct standardisation)', kortweg de 'age standardised rate' (ASR).¹

2.1.2. Voorbeeld 2: houding tegenover het dragen van de hoofddoek

Elk jaar organiseert de Studiedienst van de Vlaamse Regering (SVR) in het voorjaar een enquête naar de 'sociaal-culturele verschuivingen in Vlaanderen', de SCV-survey. Het betreft een bevraging bij Nederlandssprekende Belgen tussen 18 en 85 jaar oud, woonachtig in het Vlaamse Gewest of in het Brussels Hoofdstedelijk Gewest. Uitgangspunt is de realisatie van ongeveer 1500 gevalideerde interviews.

In de SCV-survey van 2008 (SCV08) is volgende vraag opgenomen:

Ik ga u nu een aantal uitspraken voorlezen over migranten. Hieronder verstaan we vooral Turken en Marokkanen.

Vraag 113 (V113_12)

Kan u voor iedere uitspraak zeggen in welke mate u het ermee eens bent of niet?

Islamitische vrouwen en meisjes moeten altijd en overal een hoofddoek kunnen dragen

¹ Merk op dat de techniek veelal aangewend wordt voor het gecontroleerd vergelijken van verhoudingsmaten zoals sterftecijfers (taux, rates). Men kan de techniek ook toepassen op gemiddelden, waarbij waargenomen gemiddelden in deelgroepen van de onderscheiden studiepopulaties dezelfde weging meekrijgen. Dit levert dan telkens aan de standaard gewogen gemiddelden op.

Tabel 2a Antwoorden op V113_12

Antwoord	%	→	"Recode"	%
Volledig oneens	18,3		Oneens	61,0
Oneens	42,7			
Noch eens, noch oneens	19,2		Niet oneens	39,0
Eens	16,3			
Volledig eens	2,5			
Geen mening	1,0			
Totaal	100,0 (N = 1.464)		Totaal	100,0
Geen antwoord (missing)	(1)			

Bron: SCV-survey 2008

Een meerderheid (61%) onder de respondenten is het 'oneens' met de uitspraak, de minderheid (39%) is het 'niet oneens'. Men zou kunnen stellen dat die minderheid wat "toleranter" staat tegenover 'het altijd en overal kunnen dragen van de hoofddoek', zij het dat menigeen in die groep enige reserve behoudt (slechts 2,5% verklaart zich 'volledig eens').²

Verdere exploratie geeft aan dat er een verband is tussen de 'houding tegenover het altijd en overal kunnen dragen van de hoofddoek' en het 'al of niet wonen in een stedelijk gebied'. Die laatste variabele is hier een 1/0-dummy waarbij het 'stedelijke gebied' de ruimtelijke structuurgebieden 'grootstad', 'centrumstad' en 'stedelijke rand' omvat en het 'gebied erbuiten' de 'kleinere stad', het 'overgangsgebied' en het 'platteland'.

Tabel 2b Verband tussen houding tegenover het 'altijd en overal kunnen dragen van de hoofddoek' en verstedelijking, in %

Verstedelijking	Houding tegenover dragen hoofddoek		Totaal	(N)
	Oneens	Niet oneens		
Stedelijk gebied	56,7	43,3	100,0	(538)
Gebied erbuiten	63,5	36,5	100,0	(926)
$\chi^2 = 6,629$, $df = 1$, $p = 0,010$				

Het stedelijke gebied staat wat "toleranter" tegenover 'het altijd en overal kunnen dragen van de hoofddoek'.

Ook met het opleidingsniveau treedt een verband naar voren:

² Enig voorbehoud m.b.t. de indeling 'tolerant/niet tolerant' is gepast want niet uitgesloten moet worden dat iemand best aanvaardt dat islamitische vrouwen en meisjes de hoofddoek kunnen dragen op tal van publieke plaatsen, maar daarom niet "altijd en overal".

Tabel 2c Verband tussen houding tegenover ‘het altijd en overal kunnen dragen van de hoofddoek’ en opleidingsniveau, in %

Opleidingsniveau	Houding tegenover dragen hoofddoek		Totaal	(N)
	Oneens	Niet oneens		
Geen/LO	65,4	34,6	100,0	(272)
Secundair	62,4	37,6	100,0	(798)
Hoger_bachelor	57,7	42,3	100,0	(208)
Hoger_master	52,2	47,8	100,0	(186)
$\chi^2 = 9,997, df = 3, p = 0,019$				

Hoe hoger het opleidingsniveau, des te “toleranter”.

Vraag is of het hoger geschetste verband met verstedelijking standhoudt *na standaardisatie* voor opleidingsniveau (in doorsnee lager in het niet-stedelijke gebied)?

Bekijken we daartoe eerst de kans op een “tolerante houding” (TOL) naar opleidingsniveau binnen elke subgroep van ‘al of niet wonen in een stedelijk gebied’:³

Tabel 2d TOL per opleidingsgroep in P1 en P2

Opleidingsniveau	P1 (stedelijk gebied)			P2 (gebied erbuiten)		
	N	TOL	% TOL	N	TOL	% TOL
Geen/LO	83	36	43,4	189	58	30,7
Secundair	285	119	41,8	513	181	35,3
Hoger_bachelor	81	37	45,7	127	51	40,2
Hoger_master	89	41	46,1	97	48	49,5
$\chi^2 = 0,742, df = 3, p = 0,86$			$\chi^2 = 10,872, df = 3, p = 0,012$			

Merk op dat het positieve verband tussen een “tolerante houding” en het opleidingsniveau *enkel* nog in het ‘niet-stedelijke gebied’ wordt vastgesteld.

Nemen we vervolgens de globale verdeling naar opleidingsniveau als (interne) standaard:

³ Bij *controle voor opleidingsniveau* zouden we het verband tussen TOL en ‘al of niet wonen in een stedelijk gebied’ precies omgekeerd bekijken, namelijk binnen elke subgroep van opleidingsniveau. Voor de groep ‘geen/LO’ vinden we dan 43,4% TOL in P1 vs. 30,7% in P2 [$\chi^2 = 4,104, df = 1, p = 0,043$].

Tabel 2e TOL gestandaardiseerd voor opleidingsniveau voor P1 en P2

Opleidingsniveau	Standaard (N)	P1 (stedelijk gebied)			P2 (gebied erbuiten)		
		prop. TOL	TOL	% TOL	prop. TOL	TOL	% TOL
Geen/LO	272	0,434	118		0,307	84	
Secundair	798	0,418	334		0,353	282	
Hoger_bachelor	208	0,457	95		0,402	84	
Hoger_master	186	0,461	86		0,495	92	
Totaal	1.464		632	43,2		541	37,0

Met de risico's per opleidingsgroep zoals waargenomen in P1 worden in de standaardpopulatie bij de groep 'geen/LO' 118 leden (= $0,433 \times 272$) verwacht die zich "toleranter" opstellen, d.w.z. het 'niet oneens' zijn met de uitspraak van de SCV-survey. Aldus berekend voor alle opleidingsgroepen en samengeteld, geeft dat een totaal van 633 leden met een 'tolerante opstelling' in de standaardpopulatie. Vertrekken we anderzijds van de risico's per opleidingsgroep zoals waargenomen in P2, dan levert dat een totaal van 541 leden met een 'tolerante opstelling'. Omgerekend in percentages vinden we 43,2% voor P1 versus 37,0% voor P2.

Voor het toevoegen van een significantietoets, hier de *chikwadraattoets*, kan best opnieuw aangeknoopt worden bij de oorspronkelijke tabel die het verband tussen 'al of geen tolerante houding' en 'al of niet woonachtig in een stedelijk gebied' voor de ruwe data weergeeft (zie tabel 2b). Reden daartoe is dat de toetsmaat functie is van het effectief der waarnemingen.

Uitgaande van de ruwe data voor het stedelijke en niet-stedelijke gebied, en met toepassing van de berekende risico's voor elk der cellen na standaardisatie, komen we tot volgende tabel met verwachte frequenties:

Tabel 2f Aanmaak van een toetsbestand voor verband TOL - Verstedelijking⁴

Verstedelijking (URBAN)	(N)	Houding tegenover uitspraak hoofddoek (TOL)	Risico ná standaardisatie	Verwachte frequentie (W)
stedelijk gebied	(538)	'niet oneens'	0,432	232
		'oneens'	0,568	306
gebied erbuiten	(926)	'niet oneens'	0,370	342
		'oneens'	0,630	584
$\chi^2 = 5,470$, $df = 1$, $p = 0,019$				

Wat leren we? Het hoger *bruto verband* (43,3% TOL in stedelijk gebied vs. 36,5% in het gebied erbuiten; $p = 0,010$) blijft overeind *na standaardisatie* voor opleidingsniveau (43,2% versus 37,0%; $p = 0,019$). In het stedelijke gebied is er een 'meer tolerante houding'

⁴ In SPSS creëert men daartoe het bestand zoals weergegeven in de tabel, met specificatie van de variabele W als weegcoëfficiënt via de instructie: WEIGHT BY W.

tegenover het dragen van de hoofddoek (in de zin van de uitspraak), ook bij gelijke verdeling van het opleidingsniveau.

De standaardisatie mist hier, althans in zijn samenvattende weergave van het globale verband tussen studie- en uitkomstvariabele, wel de *pointe* van de analyse, namelijk de *specificatie* van het positieve verband tussen een “tolerante houding” en het opleidingsniveau: *niet* in het ‘stedelijke gebied’, *wel* in het ‘gebied erbuiten’. (Vooraf de groep met ‘geen/lager onderwijs’ in het ‘gebied erbuiten’ is het opvallend meer ‘oneens’ met de uitspraak, geheel in contrast met dezelfde opleidingsgroep in het stedelijke gebied).

Niettemin kunnen we lering trekken uit het voorbeeld: standaardisatie voor een kenmerk als controletechniek is vooral aangewezen als dit kenmerk kan verondersteld worden:

- (1) *een klassieke (gekende) versturende variabele* te zijn voor het verband tussen de studie- en de uitkomstvariabele (in een causaal model), omdat dan geen bijzondere analyse van één of andere specificatie van het verband vereist is;
- (2) *verschillend verdeeld* is binnen categorieën van de verklarende studievariabele (in een causaal model).

2.2. Uitbreiding via logistische regressie

Tot hiertoe werd in de voorbeelden maar voor één variabele gestandaardiseerd. Vaak is dat de leeftijdsverdeling als gekende potentieel versturende variabele. Men kan dit gemakkelijk uitbreiden naar twee variabelen, zeg leeftijd en geslacht, maar bij nog verdere uitbreiding wordt het hele proces een kluwen. Voor de hand liggend is dan meervoudige regressie, waarbij in voorkomend geval van een binaire uitkomstvariabele (al of niet aanwezig zijn van de uitkomst) meteen kan gedacht worden aan logistische regressie.

2.2.1. Werkwijze

Logistische regressie wordt aangewend voor het schatten van de effecten van één of meerdere verklarende variabelen op een categorische, vaak binaire, uitkomstvariabele. Centraal daarin staat het concept van de *odds ratio* die aanduidt met welke factor de *odds* van de uitkomst wijzigt bij een eenheidsstijging op een verklarende variabele, na controle voor de andere verklarende variabelen in het model (Agresti, 1996; Pickery, 2006 & 2008).⁵

De algemene modelvorm is:

$$p = \frac{e^z}{1 + e^z} \quad (1)$$

- ⇒ p = de kans op een gebeurtenis, de *toevalscomponent* van het model
- ⇒ e = het grondtal van de natuurlijke logaritme ($e \approx 2,71828 \dots$)
- ⇒ $z = \alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k$, de *systematische lineaire component*, de zgn. *lineaire voorspeller (linear predictor)*

De relatie volgt een *logistisch*, dus niet-lineair patroon.⁶ Via een transformatie van de uitkomstvariabele, i.c. met de ‘logit’ of de ‘log odds’ als *linkfunctie* tussen de toevals- en de systematische component, kan die herschreven worden als een lineair model:

⁵ *Odds* is een alternatieve uitdrukking voor een kans, te weten de kans op succes tegen de kans op falen [$(p/(1-p)) = p/q$], of het aantal successen tegen het aantal mislukkingen. De *odds ratio* is een ratio van 2 odds (OR = $\text{odds}_1/\text{odds}_2 = p_1 q_2/p_2 q_1$), en geeft zodoende aan met welke factor de ene van de andere verschilt.

⁶ De *logistische functie* werd voor het eerst ontwikkeld door P.-F. Verhulst (1804–1849), pupil van A. Quételet, als beschrijving van de verzadigingscurve van de bevolkingsgroei (Kint et al., 2006).

$$\text{logit}(p) = \log\left(\frac{p}{1-p}\right) = \alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k = \alpha + \sum_i \beta_i x_i = z \quad (2)$$

⇒ $\text{logit}(p)$ = de *log odds* van het optreden van de gebeurtenis⁷

Hieruit kunnen de *odds* (van p) afgeleid worden:

$$\text{Odds}(p) = \exp(\text{logit}(p)) = \frac{p}{1-p} = e^z \quad (3)$$

Bemerk dat de *odds* (van p) de ratio is van twee complementaire kansen, te weten de kans dat de gebeurtenis optreedt versus de kans dat ze niet optreedt.

Dit model is ook makkelijker te interpreteren, omdat het aangeeft met welke factor de *odds* (van p) veranderen bij een eenheidsstijging op één van de verklarende variabelen, terwijl de overige verklarende variabelen ongewijzigd blijven.

Via uitdrukking (1) kunnen vervolgens de kansen afgeleid worden.

Stapsgewijs kan nu als volgt te werk gegaan worden:

- 1° Preciseer de *studievariabele* en haar *cofactoren*. Immers, in een vraagstuk van standaardisatie voor kenmerken willen we concreet nagaan wat het effect is van de studievariabele op de uitkomstvariabele, na controle - i.c. standaardisatie - voor de cofactoren. Die laatste term duidt op de selectie van met de studie- en uitkomstvariabele verbonden variabelen die de onderlinge samenhang kunnen verstoren (vertekenen). Bepaal dus in het model conceptueel (vanuit een theorie) de uitkomst (p), de studievariabele (x) en de cofactoren (u_i).

In modelvorm laat zich dit schrijven als:

$$\text{logit}(p) = \log\left(\frac{p}{1-p}\right) = \alpha + \beta x + \sum_i \gamma_i u_i \quad (4)$$

- 2° *Schat de regressiecoëfficiënten* α , β en γ_i (via uitvoering van de regressie):

$$\text{logit}\left(\hat{p}\right) = \hat{\alpha} + \hat{\beta} x + \sum_i \hat{\gamma}_i u_i \quad (5)$$

→ $\hat{\alpha} = a$, $\hat{\beta} = b$, $\hat{\gamma}_i = c_i$, waaruit p als schatter voor π

- 3° Interpreteer dit model via inspectie van de *odds ratio's* (e^{β}) en hun betrouwbaarheidsintervallen, en *goodness-of-fit* maten (-2LL statistics).
- 4° Kies vervolgens de *standaard*, in dit geval vaste waarden voor alle u_i . Meest voor de hand liggend is om voor (elke dummy van) een categorische variabele de globale relatieve frequentie als standaardwaarde te nemen en voor een continue variabele haar globale gemiddelde. Merk op dat de enige verklarende variabele in dit model die dan nog effectief varieert, de studievariabele is; alle cofactoren zijn immers gelijkgesteld aan de gekozen standaard.
- 5° Vertaal de uitkomst naar waarden voor p (naargelang de gekozen waarden op de verklarende variabelen) via vergelijking (1).

⁷ De vermelding 'log odds' voor $\text{logit}(p)$ is gebruikelijk, zij het dat uit (1) volgt dat het de logaritme met grondtal e betreft, ofwel $\ln(\text{odds})$.

2.2.2. Voorbeeld 3: het gebruik van psychotrope geneesmiddelen

1) Gestandaardiseerde proporties naar opleidingsniveau

De *Gezondheidsenquête door middel van interview* die het Wetenschappelijk Instituut Volksgezondheid op regelmatige tijdstippen houdt (1997, 2001, 2004, 2008), verzamelt informatie over de gezondheidstoestand en de behoefte aan gezondheidszorgen van de bevolking. Het betreft een steekproefenquête waarbij initieel wordt gemikt op het realiseren van een steekproef van plusminus 10.000 personen woonachtig in België, waarvan 3.500 in het Vlaamse Gewest (Van der Heyden *e.a.*, 2003).⁸

In totaal werden in 2004 12.945 enquêtes gerealiseerd, waarvan 4.513 in het Vlaamse Gewest (Bayingana *e.a.*, 2006, p.36). Na weging komt dat uit op 13.831 enquêtes voor België en 4.777 voor het Vlaamse Gewest.

Hier presenteren we enige resultaten uit de enquête van 2004 over het gebruik van psychotrope geneesmiddelen (breed genomen slaap- en kalmeermiddelen of middelen tegen neerslachtigheid of angstgevoelens). Meer bepaald gaat het om het gebruik van psychotrope geneesmiddelen bij respondenten van 15 jaar of ouder (15+) die woonachtig zijn in het Vlaamse Gewest (N = 4.110, na weging).

Het gebruik van psychotrope geneesmiddelen is afgeleid uit de positieve antwoorden op volgende vragen:

Tabel 3a Gebruik van psychotrope geneesmiddelen bij de bevolking van het Vlaamse Gewest van 15 jaar of ouder, 2004, in % (N = 4.110)

Vraag		% pos.
<i>Hebt u de afgelopen 2 weken, die gisteren afliepen, X gebruikt die voorgeschreven waren door een arts?</i>		
1	X = slaapmiddelen	10,9
2	X = geneesmiddelen tegen depressie	4,6
3	X = kalmeermiddelen of andere geneesmiddelen voor de zenuwen	5,5
<i>Hebt u de afgelopen 2 weken, die gisteren afliepen, X gebruikt die niet voorgeschreven waren door een arts (vrij verkrijgbaar in de apotheek)?</i>		
4	X = slaapmiddelen	0,8
5	X = kalmeermiddelen of andere geneesmiddelen voor de zenuwen	0,4
=> Gebruik van psychotrope geneesmiddelen (via sommatie over alle X)		16,2

Bron: WIV - Gezondheidsenquête 2004 - mondelinge vragenlijst

1° Preciseer de studievariabele en haar cofactoren

Als studievariabele is gekozen voor het 'hoogste opleidingsniveau binnen het huishouden' van de respondent, een "indicator voor het socio-economisch niveau van het huishouden en van de leden die ze telt" (Bayingana, 2006, p. 40). Drie groepen zijn onderscheiden, elk van ongeveer gelijk gewicht: 'laag' (36,4%), 'midden' (31,1%) en 'hoog' (32,5%).

⁸ Zie ook: www.gezondheidsenquête.be, onder meer voor meer details over het steekproefplan (regionale stratificatie, leeftijdsstratificatie, provinciale oververtegenwoordiging, oververtegenwoordiging van ouderen, bevraging van maximaal 4 leden van hetzelfde huishouden, selectie van collectieve huishoudens, vervangingsprocedures, enzovoort).

Tabel 3b Gebruik van psychotrope geneesmiddelen naar het hoogste opleidingsniveau binnen het huishouden (HH), Vlaams Gewest, bevolking 15+, 2004, in % (N=3979)*

Variabele (x)	Waarden	Dummy	%
Opleidingsniveau HH	Laag (geen, LO of LSO)	x_1	25,2
	Midden (HSO)	x_2	12,3
	Hoog (HO)	<i>Referentie</i>	9,3

* 131 ontbrekende waarden (3,2%) zijn niet opgenomen

Bron: WIV - Gezondheidsenquête 2004. Bewerking: SVR

Exploratie van de dataset geeft aan dat zowel de uitkomstvariabele (psychotrope geneesmiddelen) als de studievvariabele (opleidingsniveau) samenhangen met volgende kenmerken: geslacht, leeftijdsgroep, al of niet alleenwonend.

Tabel 3c Verdeling van achtergrondkenmerken waarvoor gestandaardiseerd wordt, Vlaams Gewest, bevolking 15+, 2004, in % (N=4110)

Variabele (z_i)	Waarden	Dummy	%
Geslacht	Man	<i>Referentie</i>	47,5
	Vrouw	u_1	52,5
Leeftijdsgroep	15-24 j	<i>Referentie</i>	12,1
	25-44 j	u_{21}	29,1
	45-64 j	u_{22}	28,0
	65-74 j	u_{23}	14,6
	75+	u_{24}	16,1
Alleenwonend	Neen of geen antwoord	<i>Referentie</i>	77,3
	ja	u_3	22,7

Bron: WIV - Gezondheidsenquête 2004. Bewerking: SVR

Vraag is of de %-verschillen tussen de opleidingsniveaus standhouden na controle voor die achtergrondkenmerken, in casu na standaardisatie voor die kenmerken?

2° Schat de parameters van het regressiemodel

Inpassen van de verklarende variabelen met *dummycoding* zoals aangeduid in de tabellen 3b en 3c in vergelijking (5) levert:

$$\text{logit}(p) = a + b_1x_1 + b_2x_2 + c_1u_1 + c_{21}u_{21} + c_{22}u_{22} + c_{23}u_{23} + c_{24}u_{24} + c_3u_3 = z \quad (6)$$

Tabel 3d Parameters van het logistisch regressiemodel voor schatting van het gebruik van psychotrope geneesmiddelen, Vlaams Gewest, bevolking 15+, 2004 (N = 3979)

	Cte	Studievar.		Cofactoren					
		Opleiding HH		Gesl.	Leeftijdsgroep				Alleen
Dummy		x ₁	x ₂	u ₁	u ₂₁	u ₂₂	u ₂₃	u ₂₄	u ₃
Parameter	a	b ₁	b ₂	c ₁	c ₂₁	c ₂₂	c ₂₃	c ₂₄	c ₃
Schatting	-3,886	0,334	0,101	0,662	0,562	1,437	2,091	2,664	0,359
p	<0,001	<0,01	0,46	<0,001	0,043	<0,001	<0,001	<0,001	<0,001
OR*	0,021	1,40	1,11	1,94	1,75	4,21	8,10	14,36	1,43
Model chi-square = 535,953, df = 8, p < 0,001 Hosmer-Lemeshow Goodness-of-Fit Test = 3,901, df = 8, p = 0,87 Nagelkerke R ² = 0,215									

Bron: WIV - Gezondheidsenquête 2004. Bewerking: SVR

3° Interpretatie van het model

Het model is voorspellend voor de uitkomstvariabele.

De meeste parameterwaarden zijn significant op het 0,01-niveau. Uitzonderingen zijn 1) het *middelste opleidingsniveau*, dat qua waarde voor gebruik van psychotrope geneesmiddelen niet significant afwijkt van de referentiecategorie in casu het hoogste opleidingsniveau (van het huishouden); 2) zo ook de *leeftijdsgroep 25-44-jarigen* ten opzichte van de referentiegroep 15-24-jarigen. De waarde voor de groep met het laagste opleidingsniveau (van het huishouden) wijkt wel significant af van die voor de referentiecategorie. Die groep telt naar verhouding beduidend meer gebruikers van psychotrope geneesmiddelen, en dit bij gelijkblijvende waarden op de andere variabelen in het model.

4° Kies de standaard

De globale verdeling van de achtergrondkenmerken (cofactoren), in casu de percentages zoals aangeduid in tabel 3c. Vertaal dit naar proporties zoals onderlijnd in de vergelijking hieronder.

5° Bepaal logit(p)

Invullen van vergelijking (6) levert:

$$\text{logit}(p) = -3,886 + 0,334(x_1) + 0,101(x_2) + 0,662(\underline{0,525}) + 0,562(\underline{0,291}) \\ + 1,437(\underline{0,280}) + 2,091(\underline{0,146}) + 2,664(\underline{0,161}) + 0,359(\underline{0,227}) = z$$

-> x₁ en x₂ nemen de waarden 0 en/of 1 aan.

6° Bereken de gestandaardiseerde proportie

Uit vergelijking (1) volgen dan de gestandaardiseerde proporties per opleidingsniveau.

Tabel 3e Bepaling van de logits, odds en afgeleide percentages (%*) voor het gebruik van psychotrope geneesmiddelen voor de verschillende categorieën van het hoogste opleidingsniveau in het huishouden volgens het schattingsmodel met gestandaardiseerde cofactoren, en vergelijking met de geobserveerde percentages (%), Vlaams Gewest, bevolking 15+, 2004 (N=3979)

Opleidingsniveau huishouden	Dummy		Logit	Odds	%*	%
	X ₁	X ₂				
Laag	1	0	-1,823	0,161	13,9	25,2
Midden	0	1	-2,056	0,128	11,3	12,3
Hoog (ref.)	0	0	-2,157	0,116	10,4	9,3

Bron: WIV -Gezondheidsenquête 2004. Bewerking: SVR

Vastgesteld wordt dat de nettoverschillen tussen de opleidingsniveaus qua proportie gebruik van psychofarmaca (%*), na standaardisatie voor de gekozen kenmerken, veel kleiner zijn dan de brutowaarden zoals geobserveerd in de kruistabel. Die correctie is opvallend voor het laagste opleidingsniveau. Met andere woorden, voor een belangrijk deel zijn de oorspronkelijke (ruwe) percentageverschillen toe te schrijven aan het verschil in samenstelling van de onderscheiden opleidingsgroepen naar geslacht, leeftjidsverdeling en het aandeel in de bevolking 15+ die aangeeft alleen te wonen.

De rangorde tussen de opleidingsniveaus blijft wel overeind. Te onthouden valt dat de groep met het laagste opleidingsniveau (voor het huishouden) toch nog netto een licht maar niettemin beduidend hoger gebruik van psychotrope geneesmiddelen rapporteert dan de andere opleidingsniveaus.

2) Bepaling van het betrouwbaarheidsinterval rond de gestandaardiseerde proporties

Het betrouwbaarheidsinterval (BI) rond een waargenomen proportie p bij een gegeven betrouwbaarheid gelijk aan $(1-\alpha)$, kan voor grote steekproeven ($n > 30$) en geen al te zeldzame proporties vrij eenvoudig bepaald worden:

$$(1-\alpha)\% \text{BI voor } p = \hat{p} \pm z_{1-\alpha/2} (SE) = \hat{p} \pm z_{1-\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \quad (7)$$

-> Bij $\alpha=0,05$ is $z_{1-\alpha/2} = 1,96$

De bepaling van het betrouwbaarheidsinterval op de uit een logistische regressievergelijking geschatte proportie p vraagt meer rekenwerk. Ze vertrekt van het betrouwbaarheidsinterval rond de logit:

$$(1-\alpha)\% \text{BI voor } \text{logit}(p) = \text{logit} \left(\hat{p} \right) \pm z_{1-\alpha/2} (ASE) \quad (8)$$

ASE staat voor de *asymptotische standaardfout* van $\text{logit}(p)$.⁹ De waarde wordt niet in alle statistische pakketten berekend. Bij ontstentenis kan de waarde berekend worden, via de variantie-covariantie matrix van de modelparameters (zie bijlage A1), of via een transformatie van de verklarende variabelen annex herberekening van het model (zie bijlage A2).

⁹ Asymptotische standaardfouten worden onder meer in aanmerking genomen als de regressieparameters geschat zijn via de *maximum likelihood* methode, wat het geval is binnen *Generalised Linear Models* (Hutcheson & Sofroniou, 1999, p. 121).

Voor ons voorbeeld is ASE van $\text{logit}(p)$ gelijk aan 0,102 voor de groep met een 'hoog' opleidingsniveau, 0,084 voor de groep met een 'midden' opleidingsniveau, en 0,095 voor de groep met een 'laag' opleidingsniveau.

Bepalen we het 95%-betrouwbaarheidsinterval (95%BI) voor de groep met een 'hoog' opleidingsniveau (cf. de vergelijkingen (8) en (7) resp.):

$$95\%BI \text{ voor } \text{logit}(p) = [-2,157-(1,96)(0,102); -2,157+(1,96)(0,102)] = (-2,357; -1,957)$$

$$95\%BI \text{ voor } p = (e^{-2,357}/1 + e^{-2,357}; e^{-1,957}/1 + e^{-1,957}) = (0,087; 0,124) = (8,7\%; 12,4\%)$$

De berekening werd analoog uitgevoerd voor de andere opleidingsgroepen.

Tabel 3f Waargenomen (%) en geschatte percentages (%*) voor gebruik van psychotrope geneesmiddelen naar opleidingsniveau (van het huishouden), met hun 95%-betrouwbaarheidsinterval, Vlaams Gewest, bevolking 15+, 2004 (N=3979)

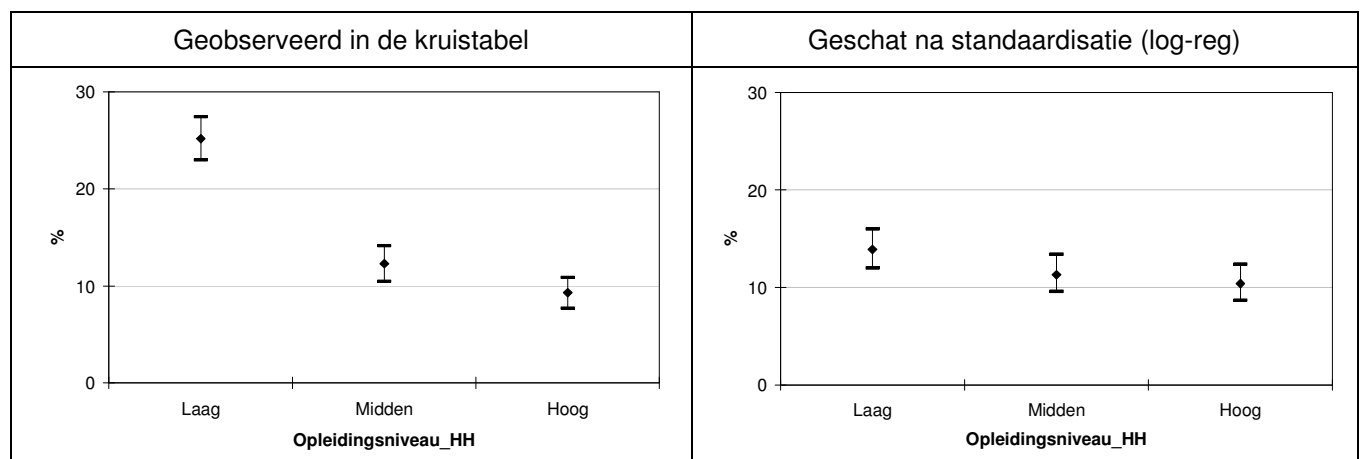
Opleidingsniveau	Geobserveerd in de kruistabel			Geschat na standaardisatie		
	%	OG	BG	%*	OG	BG
Laag	25,2	23,0	27,4	13,9	12,0	16,0
Midden	12,3	10,5	14,1	11,3	9,6	13,4
Hoog	9,3	7,7	10,9	10,4	8,7	12,4

-> OG = ondergrens; BG = bovengrens

Bron: WIV - Gezondheidsenquête 2004. Bewerking: SVR

Dit visualiseert zich als volgt:

Figuur 1 Verband tussen gebruik van psychotrope geneesmiddelen en opleidingsniveau (van het huishouden) zoals geobserveerd in de kruistabel en volgens de schatting na standaardisatie, met aanduiding van het 95%-betrouwbaarheidsinterval



Bron: WIV - Gezondheidsenquête 2004. Bewerking: SVR

Goed zichtbaar is hoe de uitkomstwaarden na standaardisatie dichter bij elkaar komen te liggen.¹⁰

¹⁰ Voor het al of niet aanduiden van significante verschillen kan men zich best verlaten op de resultaten van de logistische regressie, cf. de rij met p -waarden in tabel 3d wat gekoppeld is aan verschiltoetsen.

3. Indirecte standaardisatie

3.1. Werkwijze

Hier gaat men uit van de specifieke risico's op de uitkomst van belang in deelgroepen van de *standaard* en past deze, - net andersom dan bij directe standaardisatie -, toe op de *studiepopulaties*. Vaak betreft het ook hier *standaardisatie voor leeftijd* als 'gekende versturende variabele'. Voor de eenvoud van presentatie houden we dit ook zo aan in de navolgende bespreking.

Indirecte standaardisatie wordt aangewend als informatie beschikbaar is over:

- 1) het *globale risico* op een uitkomst (sterfte, hospitalisatie, geboorte, ...) in elk der studiepopulaties;
- 2) de *leeftijdverdeling* in elk der studiepopulaties;
- 3) de *leeftijdsspecifieke risico's* op een uitkomst in de *standaardpopulatie*.

Wat hier ontbreekt, of althans niet de aandacht heeft, zijn de *leeftijdsspecifieke risico's* op de uitkomst van belang in elk der *studiepopulaties*. Bij ontstentenis daarvan zou men de methode van indirecte standaardisatie dan ook kunnen zien als een goed "tweede alternatief".

Het Agentschap Zorg en Gezondheid van de Vlaamse overheid wendt de techniek nochtans ook veelvuldig aan in gevallen waarin de risicoprofielen naar leeftijd in de studiepopulaties (bijvoorbeeld gemeenten) wel gekend zijn. Dit heeft ermee te maken dat de techniek van indirecte standaardisatie zich goed leent voor een naar leeftijd gecontroleerde analyse van een zeldzaam voorkomende uitkomst van belang (bv. sterfte voor kleinere doodsoorzaken, met in sommige leeftijdsgroepen weinig of zelfs helemaal geen sterfgevallen). In zulke gevallen wordt de techniek van directe standaardisatie immers geheel onbetrouwbaar.

Het algoritme is als volgt.

- 1) Bereken het aantal gevallen met de uitkomst van belang (sterfte, hospitalisaties, geboorten, ...) dat per leeftijdsgroep in de studiepopulatie is te verwachten met behulp van de *leeftijdsspecifieke risico's van de standaard(populatie)*, via weging van die risico's met de overeenkomstige leeftijdsgroepen van de studiepopulaties.
- 2) Tel samen. Dit geeft het *verwachte aantal* gevallen met de uitkomst van belang onder voorwaarde dat de studiepopulatie dezelfde leeftijdsspecifieke risico's kent als de standaard.
- 3) Deel vervolgens het *waargenomen (totale) aantal* gevallen met de uitkomst van belang door het berekende *verwachte (totale) aantal* gevallen.

Dit levert, naargelang de uitkomst van belang, de "*indirect gestandaardiseerde mortaliteitsratio*", de "*indirect gestandaardiseerde hospitalisatiesratio*", enzovoort.

Meest gekend is de “*indirect gestandaardiseerde mortaliteitsratio*”, in gebruikelijk Engels de “*Standardized Mortality Ratio*” (SMR) genoemd:¹¹

$$SMR = \frac{d}{e} \quad (9)$$

Waarin $d = \textit{diagnosed}$ = waargenomen aantal sterfgevallen

$$e = \textit{expected}$$
 = verwacht aantal sterfgevallen = $\sum_{i=1}^k n_i R_i$

-> n_i = het aantal personen(-tijd) in het *i*-de stratum van de studiebevolking
 R_i de (referentie)kans voor het *i*-de stratum van de standaardbevolking

SMR gelijk aan 1 betekent dat het waargenomen aantal gelijk is aan het verwachte aantal, een ratio kleiner dan 1 betekent dat het waargenomen aantal kleiner is dan het verwachte aantal en een ratio groter dan 1 dat het waargenomen aantal groter is dan het verwachte aantal.

Vaak wordt de ratio vermenigvuldigd met 100. SMR = 120 wil dan zeggen dat in de studiebevolking de waargenomen sterfte 20% hoger ligt dan verwacht op basis van de leeftijds(groep)specifieke sterfterisico's in de standaardbevolking. SMR = 80 betekent dat de waargenomen sterfte 20% lager ligt dan verwacht.

3.2. Voorbeelden

3.2.1. Voorbeeld 4: kans op tenlasteneming door de Vlaamse zorgverzekering

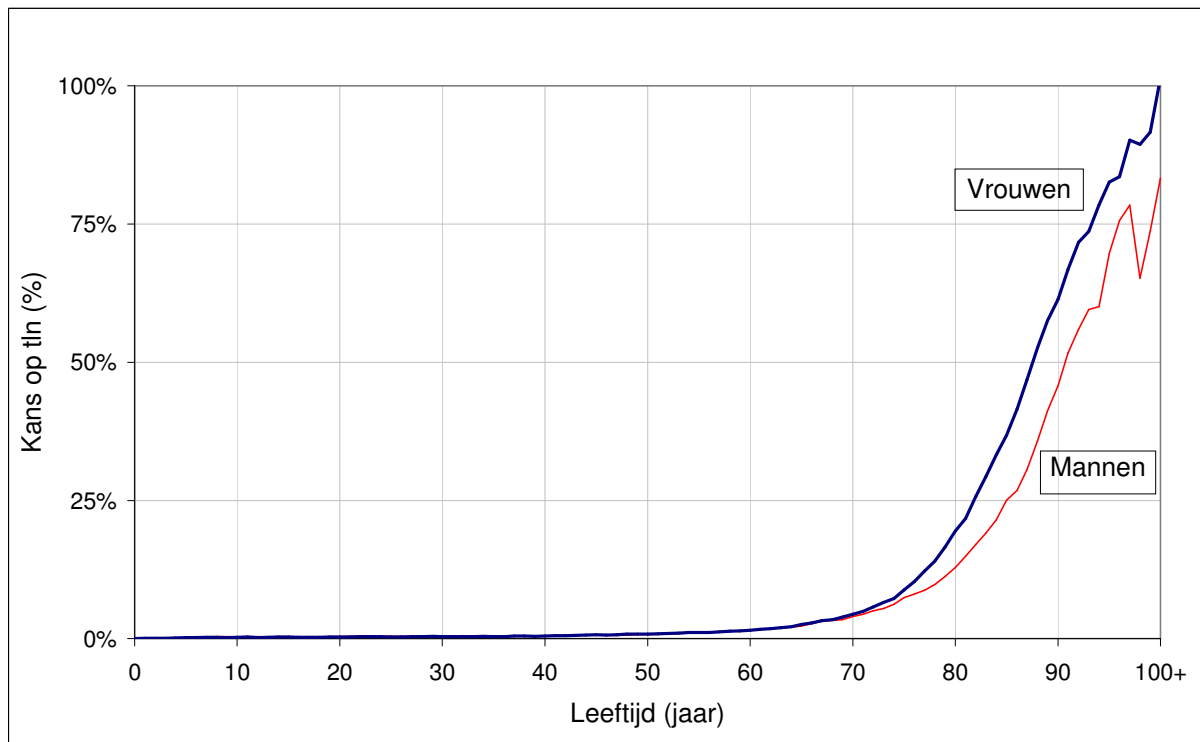
De afdeling Vlaamse Zorgverzekering telde eind 2007 in totaal 180.321 “tenlastenemingen” door de Vlaamse zorgverzekering.¹² Concreet betekent dit dat voor die groep personen een reguliere tegemoetkoming vanwege de Vlaamse overheid toegekend is voor niet-medische zorgkosten. In overgrote meerderheid gaat het om residenten van het Vlaamse Gewest (177.123 personen, waarvan 2/3 vrouwen en 1/3 mannen).

Op de hele bevolking woonachtig in het Vlaamse Gewest is de kans op een tenlasteneming, hetzij in het kader van mantel- of thuiszorg of van residentiële zorg, gelijk aan 0,029. Bij vrouwen (0,038) is dat vrijwel dubbel zo hoog als bij mannen (0,020). Merk op dat in beide geslachtsgroepen die kans flink gaat stijgen vanaf leeftijd 70 jaar.

¹¹ Analooq spreekt men over de “*indirect gestandaardiseerde hospitalisatieratio*” (SHR) en de “*indirect gestandaardiseerde geboorteteratio*” (SBR) naargelang het hospitalisaties of geboorten betreft. SMR is wel de meer generieke term.

¹² De tellingen worden verricht door het team Zorgfonds binnen de afdeling, die op haar beurt deel is van het Agentschap Zorg en Gezondheid van de Vlaamse overheid.

Figuur 2 Kans op tenlasteneming (tln) naar leeftijd en geslacht, Vlaams Gewest, stand op 31/12/2007, in %



Bron: Afdeling Vlaamse Zorgverzekering. Bewerking: SVR

Per bestuurlijk arrondissement in het Vlaamse Gewest varieert de kans op een 'tenlasteneming' bij vrouwen van 0,026 (arrondissement Halle-Vilvoorde) naar 0,054 (arrondissement Aalst). Voor de arrondissementen Turnhout en Veurne noteren we een gelijke kans = 0,038. Vraag is of die gelijkheid standhoudt na standaardisatie voor leeftijd? Die vraag stelt zich in het bijzonder omdat de vrouwelijke bevolking van het arrondissement Veurne flink ouder is. Het aandeel 65-plussers is 27% voor het arrondissement Veurne tegenover 18% voor het arrondissement Turnhout; het aandeel 80-plussers is respectievelijk 8% versus 5%.

Er wordt gekozen voor indirecte standaardisatie omdat er in ieder arrondissement niet altijd een voldoende aantal waarnemingen (tln) zijn per leeftijdsgroep, waardoor de bepaling van de risicofactoren (in casu de kansen op een tenlasteneming per leeftijdsgroep en per arrondissement) onbetrouwbaar kan worden.

Bekijken we dit stapsgewijs.

1° Kies een relevante *standaard* waarbij voor elke gegeven leeftijdsgroep de kans op de gebeurtenis (in casu tln) *wel gekend* is.

Voor ons voorbeeld is dat het Vlaamse Gewest:

Tabel 4a Kans op tenlasteneming (tln) per 10-jaarsleeftijdsgroep bij vrouwen, Vlaams Gewest, stand op 31/12/2007 (= standaard)

<i>Leeftijdsgroep</i>	Bevolking	Aantal tln	Kans
0-9 jaar	316.995	484	0,0015
10-19 jaar	348.214	900	0,0026
20-29 jaar	368.355	1.313	0,0036
30-39 jaar	401.988	1.650	0,0041
40-49 jaar	471.678	3.149	0,0067
50-59 jaar	410.686	4.517	0,0110
60-69 jaar	326.178	8.032	0,0246
70-79 jaar	290.106	25.746	0,0887
80-89 jaar	161.726	52.077	0,3220
90+	25.718	19.086	0,7421
Totaal	3.121.644	116.954	0,0375

Bron: ADSEI & Afdeling Vlaamse Zorgverzekering. Bewerking: SVR

2° Pas de leeftijdsspecifieke kansen van de standaard toe op de studiepopulaties.

Concreet vermenigvuldigen we de waargenomen kansen met de overeenkomstige leeftijdsgroep van elke studiepopulatie (of indexpopulatie).

Voor de leeftijdsgroep 0-9 jaar van het arrondissement Turnhout bijvoorbeeld is dat: $0,0015 \times 21.205 = 32$ (afgerond). Dit geeft dus het verwachte aantal tenlastenemingen in die groep (naar arrondissement en leeftijdsgroep) gegeven de betreffende kansen/risico's van de standaard.

Tabel 4b Verwacht aantal tenlastenemingen (tln) in de arrondissementen Turnhout en Veurne volgens het Vlaamse leeftijdsgroepspecifieke regime op 31/12/2007

Leeftijdsgroep	Arr. Turnhout		Arr. Veurne	
	Bevolking	Verwacht # tln	Bevolking	Verwacht # tln
0-9 jaar	21.205	32	2.292	3
10-19 jaar	24.085	62	3.043	8
20-29 jaar	25.680	92	2.820	10
30-39 jaar	28.321	116	3.056	13
40-49 jaar	34.871	233	4.070	27
50-59 jaar	28.915	318	4.470	49
60-69 jaar	21.837	538	4.518	111
70-79 jaar	17.925	1.591	3.687	327
80-89 jaar	8.695	2.800	2.108	679
90+	1.352	1.003	358	266
Totaal		6.785		1.493

Bron: ADSEI & Afdeling Vlaamse Zorgverzekering. Bewerking: SVR

3° Sommeer over die verwachte aantallen (VERW). Dit levert 6.785 verwachte tenlastenemingen voor het arrondissement Turnhout, en 1.493 voor het arrondissement Veurne.

4° Vergelijk met de geobserveerde tenlastenemingen (OBS), respectievelijk 8.017 en 1.145, waaruit volgt:

- $SMR_{(arr. Turnhout)} = OBS/VERW = 8.017/6.785 = 1,18$
- $SMR_{(arr. Veurne)} = OBS/VERW = 1.145/1.493 = 0,77$

In vergelijking met de standaard (vrouwen Vlaams Gewest, eind 2007) ligt het aantal tenlastenemingen 18% hoger in het arrondissement Turnhout en 23% lager in het arrondissement Veurne.

Ergo: het arrondissement Veurne scoort flink lager dan het arrondissement Turnhout qua proportie tenlastenemingen als we rekening houden met het verschil in leeftijdsverdeling.

3.2.2. Voorbeeld 5: oversterfte in Vlaamse gemeenten

Het Agentschap Zorg en Gezondheid publiceert op haar website een statistisch overzicht van de sterfgevallen in het Vlaamse Gewest op basis van haar registratie van de ingezonden sterftecertificaten.¹³

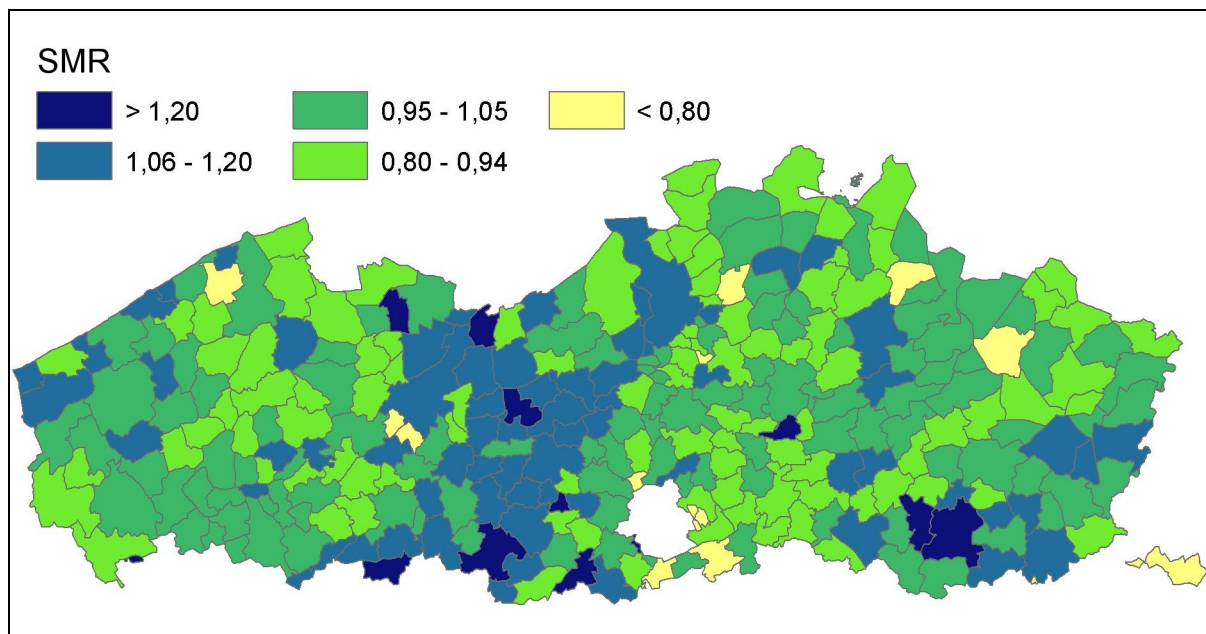
¹³ Sinds 1994 verwerkt het team Gegevensverwerking en Resultaatopvolging van het (Vlaams) Agentschap Zorg en Gezondheid de sterfgevallen die plaatsvinden in alle gemeenten van het Vlaamse Gewest, op basis van het sterftecertificaat ingevuld door een arts (de C-strook met medische gegevens). De cijfers op de website bevatten vanaf 1998 ook alle sterfgevallen van residenten van het Vlaamse Gewest, gestorven in het Brussels Hoofdstedelijk Gewest. Daarmee is er natuurlijk nog steeds een onderschatting van het aantal sterfgevallen omwille van sterfgevallen in het Waalse Gewest of in het buitenland. Jaarlijks worden zo rond de 58.500 sterfgevallen van inwoners van het Vlaams Gewest geregistreerd. Zie: http://www.zorg-en-gezondheid.be/data_sterfte.aspx

Wegens het kleine aantal sterfgevallen per gemeente, werden de data van vijf jaar samengeteld.¹⁴

De kaart met indirect gestandaardiseerde sterfteratio's (SMR's) per gemeente voor mannen toont dat er in een groot deel van Oost-Vlaanderen (zowel in het zuiden en Gent als in de Denderstreek en het Waasland) meer sterfgevallen genoteerd worden dan verwacht. Ook de streek rond Sint-Truiden en Tongeren is donkerder ingekleurd. West-Vlaanderen (vooral tussen Brugge en Roeselare) en Vlaams-Brabant kennen daartegenover veel gemeenten waar minder mannen sterven dan verwacht. Ook de randgemeenten rond de 3 grote steden hebben relatief lage sterftcijfers voor mannen.¹⁵

Op de kaart met SMR's per gemeente voor vrouwen is veel minder dan voor mannen een tendens vast te stellen.

Kaart 1 Indirect gestandaardiseerde sterfteratio (SMR) per gemeente - mannen, alle oorzaken, periode 2002-2006, vaste indeling

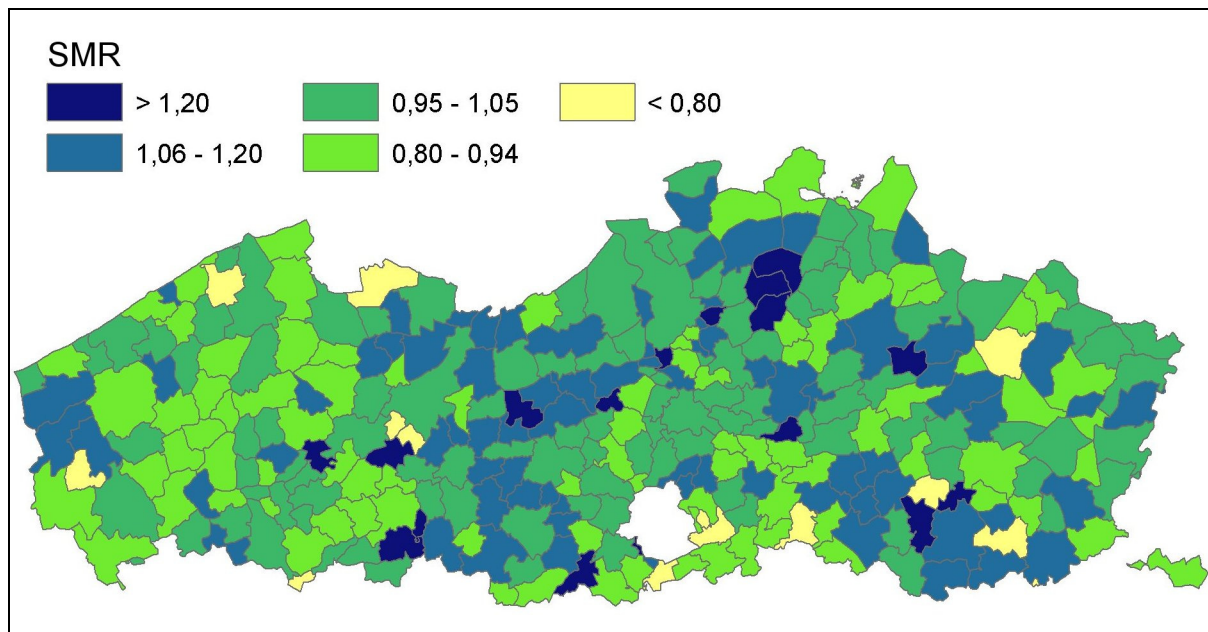


Bron: Vlaams Agentschap Zorg en Gezondheid

¹⁴ Op de website van het Agentschap Zorg en Gezondheid is toegevoegd dat ook dat meestal niet voldoende is om tot statistisch significante verschillen te komen.

¹⁵ Zie: <http://www.zorg-en-gezondheid.be/regiospreiding-sterfte.aspx#>

Kaart 2 Indirect gestandaardiseerde sterfteratio (SMR) per gemeente – vrouwen, alle oorzaken, periode 2002-2006, vaste indeling



Bron: Vlaams Agentschap Zorg en Gezondheid

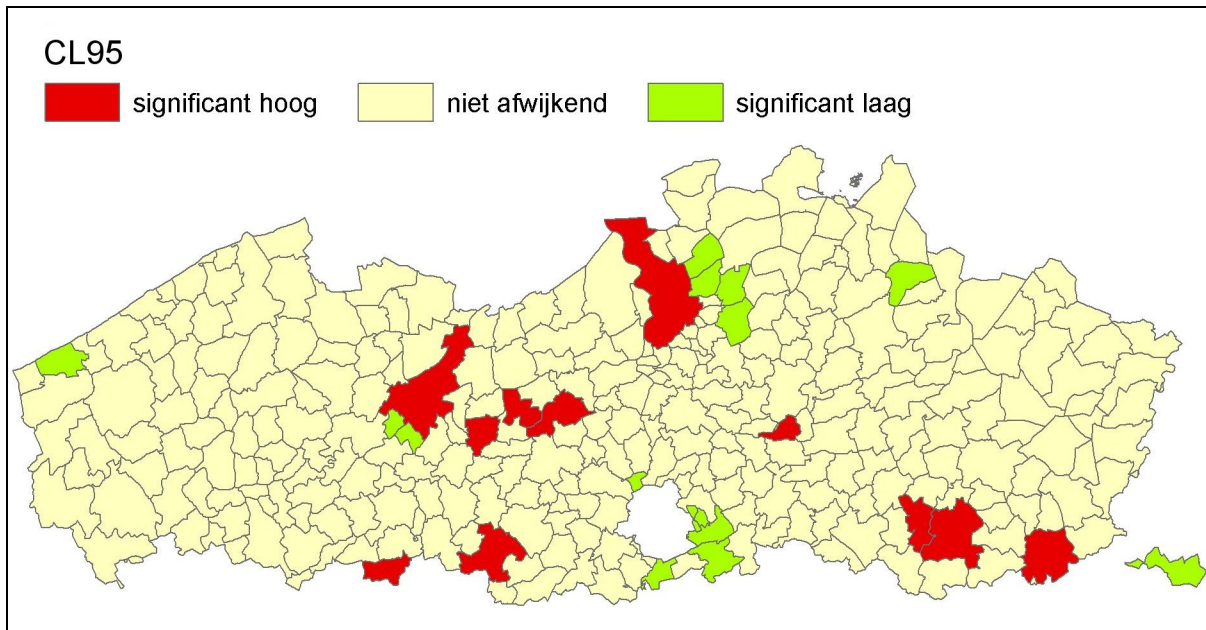
Op de website van Agentschap Zorg en Gezondheid wordt naast de display volgens vaste verschillen ook telkens een kaart met de indeling in kwintielen en een kaart met significante SMR meegegeven.

Hieronder nemen we de kaart over met de aanduiding van gemeenten met van de standaard significant afwijkende SMR (bij een gezamenlijk significantieniveau van $\alpha=0,050$, bepaald via de Bonferroni-correctie¹⁶) wat de sterfte voor alle oorzaken bij mannen betreft. Er is significante oversterfte in Gent, Wetteren, Berlare, Dendermonde, Ronse en Geraardsbergen, in Antwerpen, in Tremelo en in Zoutleeuw, Sint-Truiden en Tongeren. In een klein aantal gemeenten, voornamelijk randgemeenten van de grotere steden, is er significante ondersterfte.

Significante SMR worden – analoog als bij odds ratio's – bepaald door na te kijken of het daaromheen gecentreerde $(1-\alpha)\%$ -betrouwbaarheidsinterval al of niet de nulwaarde voor de ratio (= 1,0) insluit. Zo dit niet het geval is, dan is de ratio significant afwijkend van de nulwaarde op het aangegeven α -niveau.

¹⁶ Dit is een correctie voor 'veelvuldig gelijktijdig statistisch toetsen'. Ze gaat uit van de idee dat als je k afzonderlijke potentiële beslissingen wilt toelaten met een gezamenlijk significantieniveau van α , men deze gewoon elk afzonderlijk toetst, maar met een significantieniveau van α/k . Dit levert (veel) bredere intervallen, en dus minder kans op verwerping (Buijs, 2003, p.339).

Kaart 3 Indirect gestandaardiseerde sterfteratio (SMR) per gemeente – mannen, alle oorzaken, periode 2002-2006, significante SMR *



* Bepaald via inspectie van het 95%-betrouwbaarheidsinterval (CL95, van 95%-confidence limits) inclusief Bonferroni-correctie voor "multipel gelijktijdig statistisch toetsen", waarbij een significante afwijking aangeduid wordt bij het niet insluiten van de neutrale testwaarde (ratio = 1,0), met 'hoog' en 'laag' naargelang CL95 boven of onder die testwaarde ligt

Bron: Vlaams Agentschap Zorg en Gezondheid

Het team Gegevensverwerking en Resultaatopvolging van het agentschap vermeldt op haar website ook ondergrenzen (OG) en bovengrenzen (BG) van het 95%-betrouwbaarheidsinterval.¹⁷ Voor gemeenten met maximaal 100 getelde sterfgevallen zijn die bepaald aan de hand van de geëigende *Poisson-verdeling*; voor gemeenten met meer dan 100 getelde sterfgevallen zijn die asymptotisch bepaald, te weten via een benadering volgens een goedgekozen *normale verdeling*. Toegevoegd zijn tevens resultaten van de Bonferroni-correctie voor veelvuldig testen.

Twee methodes voor de bepaling van het betrouwbaarheidsinterval worden kort gepresenteerd in het addendum A3.

¹⁷ <http://www.zorg-en-gezondheid.be/regiospreiding-sterfte.aspx>

4. Besluit

Standaardisatie voor één of meerdere kenmerken is een zinvolle statistische oefening om te corrigeren voor één of meer versturende variabelen voor de relatie tussen een studievariabele en een uitkomstvariabele. Ruwe waarnemingen worden daarmee gecontroleerde waarnemingen, te weten uitgezuiverd voor de effecten van 'potentiële verstoorders' (*potential confounders*).

De techniek lijkt ons vooral aangewezen voor het wegzuiveren van de vertekende effecten van "gekende" versturende variabelen. Dit heeft ermee te maken dat de aandacht meteen gericht wordt op het eindresultaat van de standaardisatie, waardoor mogelijke complicaties (een *specificatie* van het verband, *interactie-effecten*, enzovoort) niet meteen in beeld komen. In die zin is het minder een techniek van analyse.

Wat voor de techniek pleit, zeker als het om een *gekende* versturende variabele gaat (zoals leeftijd bij zaken als ziekte en dood), is de eenvoud van presentatie. De aan de standaard gewogen verhoudingsmaten zijn ook door de leek vrij goed te vatten en kunnen mooi in samenvattende overzichtskaarten gepresenteerd worden.

Keuze kan gemaakt worden tussen *directe* en *indirecte standaardisatie*. Het voordeel van directe standaardisatie is dat de techniek vertrekt van de waargenomen risico's op de uitkomst van belang in deelgroepen (vaak leeftijdsgroepen) van de studiepopulaties. In die zin leunt ze dichter aan bij de realiteit dan wat met de techniek van indirecte standaardisatie kan bereikt worden. Niettemin vormt die laatste techniek een zeer valabel alternatief, zeker als 1) de specifieke risico's van de deelgroepen niet gekend zijn, of 2) geheel onbetrouwbaar blijken tengevolge van het beperkte aantal waarnemingen (wat vaak het geval is bij zeldzame ziektes of doodsoorzaken).

Nu statistische pakketten wijd verspreid zijn, komt ook het *logistische model* als techniek van "meervoudig standaardiseren" in beeld. Strikt genomen is dat een gewone uitbreiding van de techniek van *directe standaardisatie* voor proporties naar meerdere versturende variabelen toe. Het is beslist een verleidelijke optie, maar meer nog dan bij de standaardisatie voor één (gekende) versturende variabele, moet in gedachte blijven dat ook hier vooral het globale eindresultaat voor elk der studiepopulaties naar voor komt, wat mogelijke fitnesses zoals een specificatie van het verband in (gecombineerde) deelgroepen (van de variabelen waarvoor gestandaardiseerd wordt), kan verhullen.

LITERATUUR

Agresti, A. (1996). *An introduction to categorical data-analysis*. New York: Wiley & Sons Inc.

Bayingana, K., Demarest, S., Gisle, L., Hesse, E., Miermans, P.-J., Tafforeau, J. & Van der Heyden, J. (2006). *Gezondheidsenquête België 2004, IPH/EPI Reports 2006/035*. Brussel: Wetenschappelijk Instituut voor Volksgezondheid.

Buchan I. (2004). *Calculating Poisson confidence Intervals in Excel*. Manchester: Public Health Informatics at the University of Manchester;
<http://www.nwpho.org.uk/sadb/Poisson%20CI%20in%20spreadsheets.pdf>

Buijs, A. (2003). *Statistiek om mee te werken*. Groningen/Houten: Stenfert Kroese (Zevende druk).

Dobson, A.J., Kuulasmaa, K., Eberle, E. & Scherer, J. (1991). Confidence intervals for weighted sums of Poisson parameters. In: *Statistics in Medicine*, 10, 457-462.

Hutcheson, G. & Sofroniou, N. (1999). *The multivariate social scientist: Introductory statistics using generalized linear models*. Thousand Oaks: Sage Publications.

Kint, J., Constales, D. & Vanderbauwhede A. (2006). Pierre-François Verhulst's Final Triumph. In: *The logistic map and the route to chaos. From the beginnings to modern applications (Book Series: Understanding complex systems)*. Berlin/Heidelberg: Springer Verlag, 13-28.

Pickery, J. (2006). *Een statistische analyse van een toenemende of dalende ongelijkheid in participatie. Van kruistabellen naar oddsratio's en van oddsratio's naar een logistische regressie (en terug), SVR-Technisch rapport 2006/3*. Brussel: Studiedienst van de Vlaamse Regering.

Pickery, J. (2008). *De interpretatie van interactie-effecten in regressiemodellen, SVR-Technisch rapport 2008/1*. Brussel: Studiedienst van de Vlaamse Regering.

Ulm, K. (1990). Simple method to calculate the confidence interval of a standardized mortality ratio (SMR). In: *Am. J. Epidemiol.*, vol. 131, no.2, 373-375.

Van der Heyden, J., Demarest, S. & Van Oyen, H. (2003). De Belgische gezondheidsenquête. In: *Handboek Gezondheidseconomie*. Diegem: Kluwer, 225-256.

ADDENDUM

A1. Bepaling ASE van logit(p) via de variantie-covariantie matrix

Voor een model met één verklarende variabele is de bepaling:

$$ASE \text{ voor logit}(p) = \sqrt{Var(\hat{\alpha}) + x^2 Var(\hat{\beta}) + 2xCov(\hat{\alpha}, \hat{\beta})}$$

-> x = de verklarende variabele

$Var(\hat{\alpha})$ = variantie van de geschatte constante (intercept)

$Var(\hat{\beta})$ = variantie van de geschatte regressiecoëfficiënt van de verklarende variabele

$Cov(\hat{\alpha}, \hat{\beta})$ = covariantie tussen de geschatte constante en de geschatte regressiecoëfficiënt van de verklarende variabele.¹⁸

Voor een model met meerdere verklarende variabelen wordt dat:

$$ASE \text{ voor logit}(p) = \sqrt{\sum_{j=1}^k x_{j0}^2 Var(\hat{\beta}_j) + 2 \sum_{j=1}^k \sum_{h=1}^j x_{h0} x_{j0} Cov(\hat{\beta}_h, \hat{\beta}_j)}$$

-> $\hat{\beta}_j$ staat voor alle geschatte regressiecoëfficiënten inclusief de constante.¹⁹

Voor het voorbeeld van schatting van psychotrope geneesmiddelen in 2004 bij respondenten van 15 jaar of ouder en woonachtig in het Vlaamse Gewest (zie luik 2.2.2.) kan de variantie-covariantiematrix van de modelparameters opgevraagd worden:²⁰

Variantie-covariantiematrix van de modelparameters (N = 3.979)

	Cte	EDUC1	EDUC2	gender	age1	age2	age3	age4	alleen
		x1	x2	u1	u21	u22	u23	u24	u3
Waarde	1	0	0	0,525	0,291	0,280	0,146	0,161	0,227
Cte	0,071								
EDUC1	-0,008	0,017							
EDUC2	-0,010	0,011	0,019						
gender	-0,006	-0,0001	0,00002	0,010					
age1	-0,062	0,001	0,001	0,0004	0,077				
age2	-0,061	-0,001	-0,0005	0,001	0,061	0,069			
age3	-0,060	-0,004	-0,001	0,001	0,061	0,062	0,072		
age4	-0,059	-0,005	-0,001	0,001	0,061	0,062	0,063	0,072	
alleen	-0,0004	-0,002	-0,0004	-0,001	-0,0004	-0,001	-0,001	-0,004	0,012

¹⁸ Zie: Agresti, 1996, p.110 en Hutcheson & Sofroniou, 1999, p.126.

¹⁹ Zie: Hutcheson & Sofroniou, 1999, p.138.

²⁰ In SPSS voorzien de modules 'Binary Logistic Regression' en 'Multinomial Logistic Regression' de optie om de variantie-covariantie matrix op te slaan in een apart bestand, zie 'Export model information to XML file' onder 'Save', met daarbinnen de aan te vinken optie 'Include the covariance matrix'.

Voor opleidingsniveau = 'hoog' is EDUC1 = 0 en EDUC2 = 0, zodat:

$$\begin{aligned}
 ASE = & \sqrt{(1)^2(0,071) + (0)^2(0,017) + (0)^2(0,019) + (0,525)^2(0,010) \\
 & + (0,291)^2(0,077) + (0,280)^2(0,069) + (0,146)^2(0,072) + (0,161)^2(0,072) \\
 & + (0,227)^2(0,012) \\
 & + 2*((1)(0)(-0,008) + (1)(0)(-0,010) + (1)(0,525)(-0,006) + (1)(0,291)(-0,062) + (1)(0,280)(-0,061) + (1)(0,146)(-0,060) \\
 & + (1)(0,161)(-0,059) \\
 & + (1)(0,227)(-0,0004) \\
 & + (0)(0)(0,11) + (0)(0,525)(-0,0001) + (0)(0,291)(0,001) + (0)(0,280)(-0,001) + (0)(0,146)(-0,004) + (0)(0,161)(-0,005) \\
 & + (0)(0,227)(-0,002) \\
 & + (0)(0,525)(0,00002) + (0)(0,291)(0,001) + (0)(0,280)(-0,0005) \\
 & + (0)(0,146)(-0,001) + (0)(0,161)(-0,001) + (0)(0,227)(-0,0004) \\
 & + (0,525)(0,291)(0,0004) + (0,525)(0,280)(0,001) + (0,525)(0,146)(0,001) \\
 & + (0,525)(0,161)(0,001) + (0,525)(0,227)(-0,001) \\
 & + (0,291)(0,280)(0,061) + (0,291)(0,146)(0,061) + (0,291)(0,161)(0,061) \\
 & + (0,291)(0,227)(-0,0004) \\
 & + (0,280)(0,146)(0,062) + (0,280)(0,161)(0,062) + (0,280)(0,227)(-0,001) \\
 & + (0,146)(0,161)(0,063) + (0,146)(0,227)(-0,001) + (0,161)(0,227)(-0,004)] \\
 & = 0,1020
 \end{aligned}$$

(* nultermen onderlijnd)

Voor opleidingsniveau = 'laag' is EDUC1=1 en EDUC2=0, zodat ASE = 0,0844

Voor opleidingsniveau = 'midden' is EDUC1=0 en EDUC2=1, zodat ASE = 0,0952

A2. Schatting van ASE van logit(p) via transformaties van verklarende variabelen en herberekening van het model

Deze methode geeft specifieke waarden voor ASE voor elke combinatie van gekozen waarden op de verklarende variabelen. In een eerste stap worden de verklarende variabelen gewijzigd door het *in mindering brengen* van de gekozen waarden; vervolgens wordt het model *herberekend*. De standaardfout van de *constante* in het nieuwe model is dan een goede benadering voor de ASE van logit(p).²¹

In SPSS werd de uitkomst PSYCHOTROP (geen/wel gebruik van psychofarmaca) geschat via logistische regressie uit de verklarende variabelen EDUC (laag, midden, hoog=*ref.cat.*), GESLACHT (vrouw/man = *ref.cat.*), LFT (leeftijd 15-24j = *ref.cat.* /25-44j /45-64j /65-74j /75+) en PH1 (wel/geen eenpersoonshuishouden = *ref.cat.*). Zie de tabel 3d. voor uitkomsten.

Stap 1: Transformatie van verklarende variabelen

Vooreerst werden nieuwe dummyvariabelen gecreëerd, naast GESLACHT en PH1.

Voor opleidingsniveau zijn dat EDUCL (*wel/geen 'laag' opleidingsniveau*) en EDUCM (*wel/geen 'midden' opleidingsniveau*), met als impliciete referentiecategorie het 'onderwijsniveau = hoger'.

Voor leeftijdsgroep zijn dat LFT25 (*wel/niet 25-44 jaar*), LFT45 (*wel/niet 45-64 jaar*), LFT65 (*wel/niet 65-74 jaar*), LFT75 (*wel/niet 75+*), met als impliciete referentiecategorie de 'leeftijdsgroep = 15-24 jaar'.

Voor elke set van waarden op de verklarende variabelen dient ASE opnieuw geschat te worden. Zoals aangegeven moeten die in een eerste stap *in mindering gebracht* worden:

```
COMPUTE r_GESLACHT = (geslacht - 0.525).
```

```
COMPUTE r_LFT25 = (LFT25 - 0.291).
```

```
COMPUTE r_LFT45 = (LFT45 - 0.280).
```

```
COMPUTE r_LFT65 = (LFT65 - 0.146).
```

```
COMPUTE r_LFT75 = (LFT75 - 0.161).
```

```
COMPUTE r_PH1 = (PH1 - 0.227).
```

Stap 2: herberekening van het model, apart per opleidingsgroep

a) Opleidingsgroep 'hoog':

```
TITLE 'voor EDUCH' .
```

```
LOGISTIC REGRESSION VARIABLES PSYCHOTROP
```

```
  /METHOD=ENTER EDUCL EDUCM r_GESLACHT r_LFT25 r_LFT45 r_LFT65 r_LFT75 r_PH1
```

```
  /PRINT=GOODFIT CI(95)
```

```
  /CRITERIA=PIN(0.05) POUT(0.10) ITERATE(20) CUT(0.5) .
```

=> S.E. (Constant) = 0,1020 -> ASE van logit(p)

²¹ Zie: Hutcheson & Sofroniou, 1999, p.127 en p.139.

b) Opleidingsgroep 'laag':

```
TITLE 'voor EDUCL'.  
COMPUTE r_EDUCL = (EDUCL - 1).  
LOGISTIC REGRESSION VARIABLES PSYCHOTROP  
  /METHOD=ENTER r_EDUCL EDUCL r_GESLACHT r_LFT25 r_LFT45 r_LFT65 r_LFT75 r_PH1  
  /PRINT=GOODFIT CI(95)  
  /CRITERIA=PIN(0.05) POUT(0.10) ITERATE(20) CUT(0.5) .
```

=> S.E. (Constant) = 0,0844 -> ASE van $\text{logit}(p)$

c) Opleidingsgroep 'midden':

```
TITLE 'voor EDUCM'.  
COMPUTE r_EDUCM = (EDUCM - 1).  
LOGISTIC REGRESSION VARIABLES PSYCHOTROP  
  /METHOD=ENTER EDUCL r_EDUCM r_GESLACHT r_LFT25 r_LFT45 r_LFT65 r_LFT75 r_PH1  
  /PRINT=GOODFIT CI(95)  
  /CRITERIA=PIN(0.05) POUT(0.10) ITERATE(20) CUT(0.5) .
```

=> S.E. (Constant) = 0,0952 -> ASE van $\text{logit}(p)$

Bemerk dat de beide methodes, die van de variantie-covariantiematrix (A1) en die van de transformatie van verklarende variabelen annex herberekening van het model (A2), identieke waarden opleveren voor de ASE van $\text{logit}(p)$.

A3. Bepaling van het betrouwbaarheidsinterval voor SMR

Methode 1

Een praktische methode om de betrouwbaarheidsintervallen te berekenen wordt gesuggereerd in een webartikel van Buchan (2004). Daarbij wordt voor *grote steekproeven* verwezen naar een asymptotische benadering via de *normale verdeling*. De veronderstelling blijft wel een *Poisson-verdeling* voor tellingen, met 'het aantal waarnemingen' als verwachtingswaarde (\square) en als variantie.²²

Uit vergelijking (9): $SMR = \frac{d}{e}$

waarbij $LL = \frac{d - z_{1-\alpha/2} \sqrt{d}}{e}$; $UL = \frac{d + z_{1-\alpha/2} \sqrt{d}}{e}$

- ⇒ LL = lower limit (ondergrens) van SMR voor d waarnemingen
- UL = upper limit (bovengrens) van SMR voor d waarnemingen
- ⇒ $z_{1-\alpha/2}$ = waarde in termen standaardafwijking op de (standaard)normale verdeling die overeenkomt met het $(1-\alpha/2)^\circ$ centiel ($z_{1-0,025} = 1,96$);
- ⇒ d wordt hier behandeld als een *Poisson verdeelde kansveranderlijke*, e als een vaste waarde.

Methode 2

Een meer algemene én meer exacte methode steunt op het verband tussen de *Poisson-verdeling* en de *chi-kwadraat-verdeling* (Ulm, 1990; Dobson *e.a.*, 1991):

$$LL = \frac{\left(\chi^2_{\frac{\alpha}{2}, 2d}\right)}{2e} \quad ; \quad UL = \frac{\left(\chi^2_{1-\frac{\alpha}{2}, 2(d+1)}\right)}{2e}$$

-> $\chi^2(\alpha, v)$ = het $100 \cdot \alpha$ -de chi-kwadraatcentiel met v vrijheidsgraden
 $v=2d$ voor bepaling LL, $v=2(d+1)$ voor bepaling UL.

In MS Excel wordt dat, met in cel A1 het 'waargenomen aantal sterfgevallen' (*diagnosed*) en in cel B1 het 'verwachte aantal sterfgevallen' (*expected*), en bij $\alpha=0,050$:²³

[LL]=ALS(A1>0;(CHI.KWADRAAT.INV(0,975;2*A1)/2)/B1;0)

[UL]=ALS(A1>0;(CHI.KWADRAAT.INV(0,025;2*(A1+1))/2)/B1;(CHI.KWADRAAT.INV(0,025;2)/2))

Voor de gemeente Berlare bijvoorbeeld, met 416 waargenomen sterfgevallen bij de mannelijke bevolking versus 317 verwacht volgens de standaard, wordt dat:

$$SMR = 416/317 = 1,31, \text{ met } LL = 1,19 \text{ en } UL = 1,45$$

De benadering via de *normale verdeling* geeft: $LL=(416-1,96\sqrt{416})/317=1,19$ en $UL=(416+1,96\sqrt{416})/317=1,44$

Met Bonferroni-correctie wordt gerekend met $\alpha^*=\alpha/(k \text{ gemeenten}) = 0,050/308 = 0,00016$, waaruit $z^*=3,77$ bij tweezijdig toetsen.

²² Als X de kansveranderlijke is die het aantal voorvallen telt, dan is volgens de Poisson-verdeling de kans dat er precies k voorvallen zijn $= P(X=k) = e^{-\lambda} (\lambda^k/k!)$. De benadering via de normale verdeling gaat uit van verwachting en variantie $= \lambda$. Die kan al toegepast worden vanaf $\lambda > 10$ mits toepassing van de continuïteitscorrectie.

²³ Buchan (2004) geeft aan dat de functie in MS Excel in een aantal gevallen niet werkt, bv. voor $d = 466$.

