

Rapport 320103003/2004

**Probabilistische berekening van inname van stoffen
via incidenteel geconsumeerde voedingsproducten**

W. Slob en M.I. Bakker

corresponderend auteur:

W.Slob

Centrum voor Stoffen en Integrale Risicoschatting

w.slob@rivm.nl

Dit onderzoek werd verricht in opdracht en ten laste van de Keuringsdienst van Waren, Voedsel en Waren Autoriteit, in het kader van project 320103, 'Modellering humane blootstelling xenobiotica in voeding'

Abstract

Probabilistic calculation of intake of substances via incidentally consumed food products Supplement to the manual for modelling of intake of substances via food

In this report a model is developed to estimate the intake of substances occurring in foods that are consumed incidentally (e.g. pesticides on certain types of fruit). To date, only the intake of the substances that are part of our basal diet, such as dioxins in animal fat, could be estimated. This is carried out by combining food consumption survey data with concentrations of the substances in different food products. The resulting data may be analysed by a statistical model (STEM) to achieve a distribution of long-term intakes in the population. However, this model does not apply when the food products containing the substance are not consumed every day (resulting in 'zero' intakes). The new model takes into account that the consumption frequency of individual food products may vary between individuals. The amounts consumed on a single (consumption) day may vary between individuals as well. Combining the estimated distribution of consumption frequency with that of consumption amount results in a statistical model that may be used for comparing the intake in the population with certain (chronic or acute) limit values.

key words: dietary intake; food consumption; contaminants; statistical model; consumption frequency; interindividual variation

Rapport in het kort

Probabilistische berekening van inname via incidenteel geconsumeerde voedingsproducten

Aanvulling op het draaiboek voor modellering van inname van stoffen via voeding

In dit rapport wordt een model beschreven om de inname van stoffen te schatten die voorkomen in voedselproducten die incidenteel worden gegeten (bijvoorbeeld pesticiden op bepaalde soorten fruit). Tot op heden kon alleen de inname worden geschat van stoffen die in ons basale voedingspakket aanwezig zijn (bijvoorbeeld dioxines in dierlijk vet). Dit gebeurt door gegevens van een voedselconsumptiepeiling te combineren met concentraties van de stoffen in de voedselproducten. De uitkomsten van deze berekening worden geanalyseerd met een statistisch model waarmee een verdeling van de lange-termijn innames in de bevolking wordt verkregen. Dit model kan echter niet worden gebruikt als veel individuen een inname van 'nul' hebben, omdat deze mensen de relevante voedselproducten niet hebben gegeten tijdens de peiling. Het nieuwe model houdt er rekening mee dat de consumptiefrequentie van individuele voedselproducten alsmede de geconsumeerde hoeveelheid per dag kan variëren tussen individuen. Door de geschatte verdeling van de consumptiefrequentie te combineren met de verdeling van de consumptiehoeveelheid in een statistisch model kan de inname in de bevolking vergeleken worden met bepaalde (chronische of acute) limietwaarden.

trefwoorden: inname; voedselconsumptie; contaminanten; statistisch model; interindividuele variatie; consumptiefrequentie

Inhoud

Samenvatting	4
1. Inleiding	5
2. Schatten van consumptiefrequentie	7
2.1 Inleiding	7
2.2 Consumptiefrequentie als statistische Beta-verdeling	8
2.3 Afhankelijkheid met de leeftijd	9
2.4 Illustratie: rode wijn	10
3. Schatten van consumptiehoeveelheid	12
3.1 Verdeling consumptiehoeveelheid per dag	12
3.2 Relatie tussen consumptiehoeveelheid en –frequentie	12
3.3 Relatie tussen consumptiehoeveelheid en leeftijd	13
3.4 Lange-termijn inname	13
4 Toepassen van het model	15
4.1. Inleiding	15
4.2 Limietoverschrijding lange-termijn inname	15
4.3 Limietoverschrijding korte-termijn inname	16
4.4 Overzicht modelberekening	18
5 Discussie en conclusies	19
Lijst van afkortingen	21
Referenties	22
Appendix 1 Simulaties Beta-verdeling	23
Appendix 2 STEM.II, de software	25
Appendix 3 Berekenen limietoverschrijdende fractie van populatie	27
Appendix 4 Voorbeeldberekeningen	28

Samenvatting

Ons voedsel bevat stoffen zoals nutriënten, additieven en contaminanten. Sommige van deze stoffen zijn zo algemeen in de voeding aanwezig dat bijna iedereen elke dag de stof inneemt (bijvoorbeeld vitamine C via groente en fruit, of dioxines via de consumptie van dierlijk vet). De lange-termijn inname van deze stoffen kan goed worden geschat met een model dat voedselconsumptiepeilinggegevens combineert met de concentraties van de stof in de voedingsmiddelen. Andere stoffen bevinden zich in één of meer specifieke voedingsmiddelen, met als gevolg dat niet iedereen er dagelijks aan blootgesteld wordt (bijvoorbeeld pesticiden in bepaalde fruitsoorten). In het huidige model zijn dergelijke ‘nul-innames’ voor de inname niet of slechts beperkt toegestaan. Dit rapport beschrijft een nieuw model waarmee de inname van stoffen via individuele voedingsmiddelen, en de variatie daarin tussen individuen, wél kan worden geschat. Het model houdt rekening met het feit dat consumptiefrequenties tussen individuen kunnen verschillen en schat een verdeling van consumptiefrequenties. Op deze wijze beschrijft het model zowel situaties waarin een voedingsproduct door een groot deel van de populatie gegeten wordt maar met een lage frequentie, als situaties waarin een product door een kleine groep met grote frequentie wordt gegeten. In het model wordt de gemiddelde consumptiefrequentie beschreven als functie van de leeftijd, onder de aanname dat de (Beta-)verdeling rondom het gemiddelde voor alle leeftijden vergelijkbaar is. Onder deze aanname blijkt het redelijk goed mogelijk om op basis van informatie van slechts twee dagen de verdeling van consumptiefrequenties in de populatie te schatten. Vervolgens wordt de relatie van de consumptiehoeveelheid met de leeftijd (inclusief de variatie tussen individuen) geschat. Door de verdelingen van consumptiefrequentie en van consumptiehoeveelheden te combineren, kan de fractie van de populatie wier lange-termijn inname een bepaalde chronische limietwaarde (bijvoorbeeld acceptabele dagelijkse inname, ADI) overschrijdt, berekend worden. Een tweede toepassing van het model is het berekenen van de kans dat een dagelijkse inname een acute limietwaarde (bijvoorbeeld acute reference dose, ARfD) overschrijdt, als functie van de fractie in de populatie waarbij dat gebeurt. Beide resultaten worden gegeven als functie van de leeftijd.

Inleiding

In 2002 verscheen het RIVM-rapport 'Draaiboek voor modellering van inname van stoffen via voeding' (Bakker, 2002). Hierin is een protocol beschreven voor medewerkers van het Centrum voor Stoffen en Integrale Risicoschatting van het RIVM om de inname te berekenen van stoffen (xenobiotica, natuurlijke toxinen, nutriënten of voedseladditieven) in de mens via de voeding. Het hierbij gebruikte computermodel is geschikt voor het berekenen van de inname van stoffen die zo algemeen voorkomen in voedingsmiddelen dat iedereen vrijwel elke dag de stof inneemt.

De blootstelling aan die algemeen voorkomende stoffen via de voeding wordt volgens het genoemde draaiboek berekend door gegevens over concentraties in voedingsproducten te combineren met gegevens over voedselconsumptie. In een eerste ruwe schatting wordt in het draaiboek uit 2002 deze berekening gebaseerd op gemiddelde concentraties en gemiddelde voedselconsumptie-gegevens. Zolang deze ruwe schatting van de blootstelling ordes van grootte lager is dan het mogelijke effect-niveau is deze methode acceptabel. Wanneer dat niet zo is, dient de blootstelling nauwkeuriger vastgesteld te worden, waarbij onder andere rekening gehouden wordt met de variabiliteit in consumptiegedrag *tussen* mensen (maar niet met de variaties in de concentraties, dit is voor lange termijn inname niet relevant). Het draaiboek beschrijft het gebruik van het programma FRIDGE waarmee de individuele consumptiegegevens van de tweedaagse Voedsel-Consumptie-Peilingen (VCP) gecombineerd kunnen worden met de concentraties van een bepaalde stof in de genuttigde producten. Op deze manier wordt van elk individu de inname van de stof op die twee dagen geschat, resulterend in een verdeling van de waargenomen dagelijkse innames. Wanneer men de blootstelling bij de mens wil vergelijken met een *chronische* gezondheidslimietwaarde (ADI, TDI) dient deze met de lange termijn inname van die stof vergeleken te worden. Het draaiboek beschrijft hoe met behulp van statistische modellering - namelijk met het programma STEM (Slob, 1993) - het mogelijk is de variatie in lange termijn innames te schatten, op basis van VCP-consumptiegegevens van slechts twee dagen per individu. De output van dit model beschrijft de lange-termijn inname en de variatie tussen individuen daarin als functie van de leeftijd. Deze methode is voor diverse stoffen toegepast, bijvoorbeeld voor cadmium (Slob en Kranjc, 1994), dioxines (Freijer et al., 2001), ochratoxine A (Bakker en Pieters, 2002), DON (Pieters et al. 2002; Pieters et al., 2004) en indicator PCBs (Bakker et al., 2003).

Wanneer de stof voorkomt in diverse voedingsproducten die tot het basale voedingspakket horen, zodat de stof bij elk individu vrijwel elke dag tot een zekere inname leidt, is het draaiboek van 2002 goed toepasbaar. Een probleem ontstaat wanneer de stof slechts in voedingsproducten voorkomt die met een beperkte frequentie worden gegeten. Dit leidt tot 'nullen' in de berekende dagelijkse innames per individu (geen van de op die dag geconsumeerde producten bevat, voor zover bekend, de stof). In de huidige blootstelling-modellen leiden nullen tot een schending van de statistische aannames en tot minder adequate resultaten. Wanneer het aantal nullen beperkt is zal het effect verwaarloosbaar zijn, en kunnen zij eenvoudigweg genegeerd worden. In situaties waarbij het aantal nullen groter is, is het negeren van nullen niet meer geoorloofd en kan het draaiboek van 2002 niet gebruikt worden. Het doel van het huidige onderzoek is daarom een model te ontwikkelen voor de berekening van de inname van stoffen die slechts in voedingsmiddelen voorkomen die met een beperkte frequentie worden geconsumeerd.

Wanneer in de berekende dagelijkse innames nullen voorkomen, kunnen deze op twee wijzen geïnterpreteerd worden.

A. Inname is in werkelijkheid klein, maar niet nul

Hoewel de berekende inname op een dag bij een individu op nul uit kan komen, kan verondersteld worden dat deze in werkelijkheid, weliswaar laag, maar niet echt nul is. Wanneer het bijvoorbeeld gaat om contaminanten die in veel producten voorkomen, maar die niet in alle producten gemeten zijn, kan een inname van nul berekend worden, terwijl allerlei niet onderzochte producten toch minieme hoeveelheden van de stof bevatten. Met andere woorden, in werkelijkheid zal de inname klein, maar niet echt nul zijn. Het is dan beter om de analyse net zo uit te voeren als bij een set metingen waarvan een gedeelte onder de detectielimiet ligt. Op het RIVM is bij Slob een methode (plus software) beschikbaar om de verdeling van gegevens met waarnemingen beneden de detectielimiet te schatten. In deze methode wordt een berekende inname van nul geïnterpreteerd als: de inname ligt onder een bepaalde grens, maar het is onbekend waar precies. Deze methode lijkt direct toepasbaar op deze situatie.

B. Inname is in werkelijkheid nul

De inname kan ook in werkelijkheid nul zijn, namelijk wanneer een stof slechts in bepaalde producten voorkomt (bijvoorbeeld als additief), terwijl die producten door bepaalde personen (op bepaalde dagen) niet of nauwelijks genuttigd worden. Dit geldt bijvoorbeeld voor de zogenaamde liefhebbersproducten (paling, drop). Ten aanzien van de risicobeoordeling spelen twee vragen. Enerzijds is de vraag hoe groot de groep liefhebbers is, anderzijds hoe groot de inname is in die groep. Op het eerste gezicht lijken de VCP-gegevens niet geschikt om deze vragen te beantwoorden. Immers, op grond van informatie van twee dagen kan niet besloten worden of iemand een frequent gebruiker (liefhebber) van een product is of slechts een incidenteel gebruiker.

Situatie B is in dit rapport uitgewerkt. Dit rapport is een aanvulling op het draaiboek zoals het is verschenen in 2002. Het huidige document beschrijft een model voor het berekenen van de inname van stoffen die in minder frequent gegeten voedingsmiddelen voorkomen. De opzet van het model is zodanig dat de variatie in consumptiefrequenties tussen individuen geschat wordt. Het model is ook geschikt voor stoffen die voorkomen in een beperkt aantal verschillende voedingsmiddelen, zoals additieven. In dat geval worden de innames via de verschillende voedingsmiddelen per individu (en per dag) bij elkaar opgeteld.

Het model, aangegeven met de naam STEM.II¹, bestaat uit drie verschillende onderdelen. Het eerste onderdeel is het schatten van de verdeling van de consumptiefrequentie (Hoofdstuk 2). De consumptiefrequentie kan beschouwd worden als de 'kans op consumptie' op een bepaalde dag. Deze zal tussen individuen variëren, en dit wordt beschreven met een kansverdeling. Naast de consumptiefrequentie is het nodig de geconsumeerde hoeveelheid (op dagen dat er consumptie plaats vindt) te schatten (Hoofdstuk 3). Door de resultaten over de consumptiefrequentie en de consumptiehoeveelheid te combineren kunnen uitspraken gedaan worden over het overschrijden van limietwaarden. Met name kan de fractie van de populatie wier lange-termijn inname een bepaalde (chronische) limietwaarde overschrijdt worden geschat. Ook is het mogelijk om de kans dat een acute limietwaarde (op een dag) wordt overschreden, als functie van de fractie van de populatie waarbij dat gebeurt (Hoofdstuk 4).

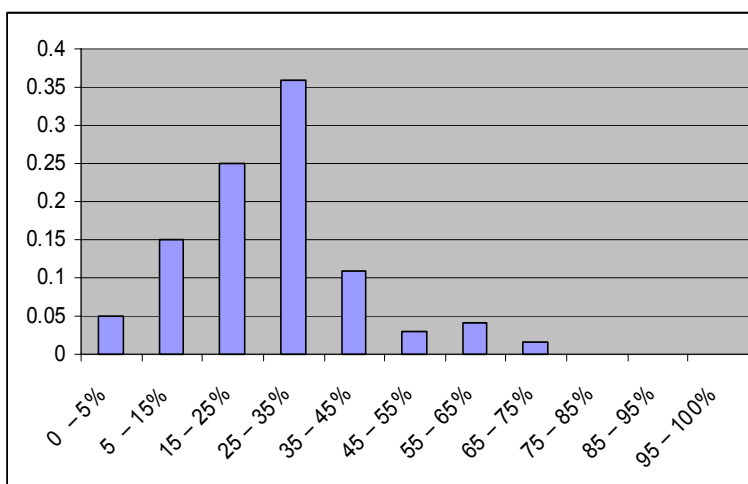
¹ STEM.II: Statistical Exposure Model for Incidental Intakes

2. Schatten van consumptiefrequentie

2.1 Inleiding

De meeste voedingsproducten worden niet elke dag geconsumeerd. De *consumptiefrequentie* zal hier gedefinieerd worden als de kans² dat een individu het product op een bepaalde dag nuttigt. Uiteraard is de consumptiefrequentie niet voor alle individuen gelijk. Sommige mensen zullen een bepaald product (bijvoorbeeld paling of wijn) zeer regelmatig gebruiken, anderen slechts incidenteel, en weer anderen nooit. Zoals al werd opgemerkt, lijkt het niet mogelijk om op basis van gegevens van slechts twee dagen een kwantitatieve schatting af te leiden van de grootte van de groep van regelmatige of incidentele gebruikers. Toch blijkt die voor de hand liggende conclusie onjuist. Van Erp-Baart et al. (2001) en Dodd (1996) beschrijven een methode om in te schatten welke fractie van de populatie binnen elk van een aantal categorieën van consumptiefrequentie valt voor een bepaald product. In tabel 1 en de bijbehorende figuur 1 wordt hiervan een voorbeeld gegeven. In dit voorbeeld heeft 5 % van de populatie een consumptiefrequentie van 0 - 5 % (dus 1 keer per ± 20 dagen), 15 % van de populatie heeft een frequentie van 5 - 15 %, etc. Bij deze methode werd gebruik gemaakt van de Engelse voedselconsumptiepeiling-gegevens, waarbij in vergelijking met de Nederlandse VCP weliswaar minder individuen geregistreerd zijn, maar wel op meer (namelijk zeven) achtereenvolgende dagen. Het nadeel van deze benadering is dat een vrij groot aantal onbekende parameters geschat moet worden, nl. één voor elke categorie (in het geval van tabel/figuur 1 in totaal 10). Deze aanpak is mogelijk voor de Engelse gegevens, omdat er per individu zeven achtereenvolgende dagen geregistreerd zijn, maar zou niet werken voor de Nederlandse VCP, met slechts twee dagen per individu.

Consumptiefrequentie	Fractie van de populatie
0 - 5%	0,05
5 - 15%	0,15
15 - 25%	0,25
25 - 35%	0,36
35 - 45%	0,11
45 - 55%	0,03
55 - 65%	0,04
65 - 75%	0,015
75 - 85%	0,00014
85 - 95%	0,00007
95 - 100%	0,00061



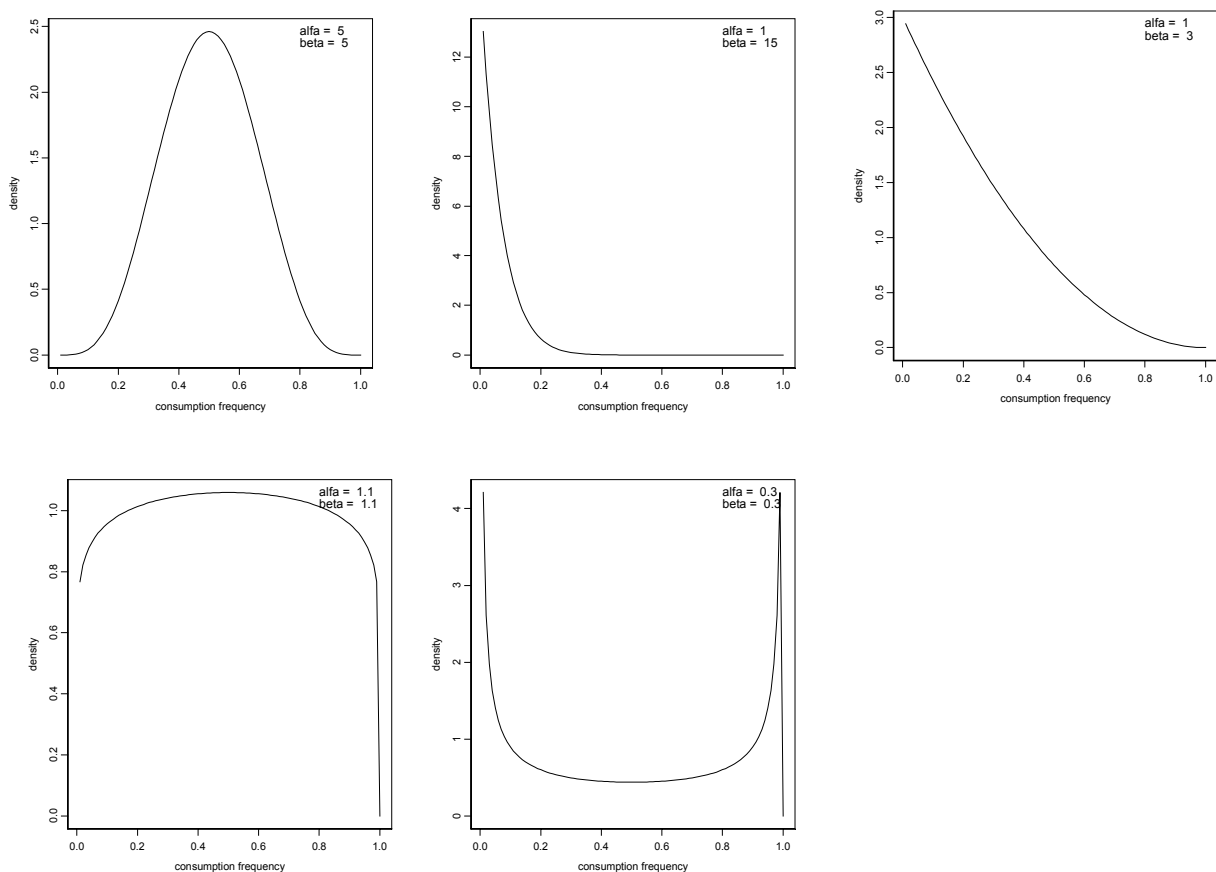
Tabel 1/ Figuur 1. Illustratie van geschatte fracties van de populatie per categorie consumptiefrequentie voor een bepaald voedingsproduct.

² Deze kans is dus gelijk aan de verwachte fractie dagen, d.w.z. ten opzichte van een oneindig aantal dagen. De beslissing van een individu op een bepaalde dag het product wel of niet te nuttigen kan beschreven worden als een random trekking met kans gelijk aan de verwachte fractie dagen.

De indeling in categorieën zoals in tabel 1 is arbitrair. In werkelijkheid bestaan er geen grenzen tussen consumptiefrequenties, en is er sprake van een ‘gladde’ verdeling in plaats van de discrete verdeling zoals in tabel 1. In dit rapport wordt een alternatieve benadering onderzocht, gebaseerd op de aanname dat de consumptiefrequentie een bepaalde statistische verdeling volgt.

2.2 Consumptiefrequentie als statistische Beta-verdeling

Een voor de hand liggende verdeling voor de consumptiefrequentie is de Beta-verdeling, omdat deze gedefinieerd is voor waarden tussen 0 en 1 (oftewel tussen 0 % en 100 %). Het voordeel is dat de Beta-verdeling bepaald wordt door twee parameters, zodat slechts twee onbekende parameters geschat behoeven te worden. Desondanks is de Beta-verdeling zeer flexibel, en kan allerlei vormen aannemen (zie figuur 2).



Figuur 2. Diverse Beta-verdelingen ter illustratie van haar flexibiliteit voor het beschrijven van allerlei soorten consumptiefrequenties en de variatie daarin tussen personen. Op y-as staat de density, i.e. de relatieve frequentie.

Links boven: consumptie treedt op in ca. de helft van de dagen, met matige spreiding tussen personen. Midden boven: incidentele consumptie met kleine spreiding. Rechts boven: incidentele consumptie maar met grote spreiding tussen personen. Links onder: alle consumptiefrequenties komen ongeveer even vaak voor. Rechts onder: populatie bestaat uit zowel relatief veel liefhebbers als relatief veel vermijders.

Wanneer de aanname juist is dat consumptiefrequenties een Beta-verdeling volgen, is het nog de vraag of deze verdeling accuraat kan worden geschat op basis van gegevens over slechts twee dagen per individu. Dit werd onderzocht met behulp van computersimulaties. De resultaten van de computersimulaties laten zien dat zelfs met twee dagen per individu de Beta-verdeling in het geval van een verdeling zoals links boven in figuur 2 goed wordt geschat (zie appendix 1, tabel A.1). Voor een verdeling zoals midden boven in figuur 2 (lage consumptiefrequentie) wordt de mediane consumptiefrequentie overschat (1,5 % in plaats van 0,65 %, zie appendix 1, tabel A.2). De fractie liefhebbers in de populatie wordt enigszins onderschat (0,8 % in plaats van 1,5 %, zie tabel A.2). Voor designs met meer dagen per individu (maar het totaal aantal dagen \times aantal personen blijft gelijk) wordt de schatting wat beter.

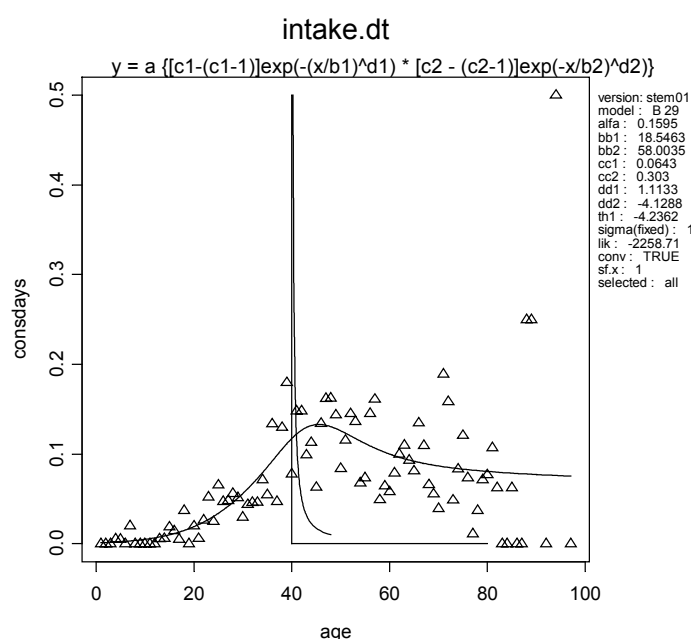
De conclusie is dat de verdeling van consumptiefrequenties redelijk tot goed ingeschat kan worden op basis van voedselconsumptiegegevens, zelfs wanneer slechts twee dagen per individu geregistreerd zijn, zoals in de Nederlandse VCP.

2.3 Afhankelijkheid van de leeftijd

Verwacht kan worden dat de consumptiefrequentie voor veel producten af zal hangen van de leeftijd. De VCP bevat hier informatie over, omdat alle leeftijdscategorieën vertegenwoordigd zijn. In principe zou dus voor elke leeftijdscategorie afzonderlijk de consumptiefrequentie, met de variatie tussen individuen (in de vorm van een Beta-verdeling), bepaald kunnen worden. Dit zou echter betekenen dat de Beta-verdeling geschat moeten worden op grond van veel minder personen dan 6000, waarvan uit werd gegaan in de simulaties. Met zo weinig individuen (circa 100) per leeftijdscategorie zullen de schattingen onbetrouwbaar worden. De oplossing hiervoor is de relatie tussen (gemiddelde) consumptiefrequentie en leeftijd als een continue functie op te vatten, en die functie te schatten. Tegelijkertijd kan aangenomen worden dat de variatie tussen individuen rondom deze functie tot uiting komt in de vorm van een Beta-verdeling. Op deze manier kan de gehele dataset van 6250 personen gebruikt worden om de Beta-verdeling te schatten. De Beta-verdeling heeft twee parameters: α en β . De aanname die hier gedaan wordt is dat één van de twee parameters van de Beta-verdeling, namelijk α , constant is, dus onafhankelijk van de leeftijd. De waarde van β volgt uit de waarde van de gemiddelde consumptiefrequentie en de waarde van α (en dus is β wel afhankelijk van de leeftijd). In STEM.II wordt een regressie-analyse uitgevoerd waarbij de waarde van α tegelijkertijd wordt geschat met de regressiefunctie die de gemiddelde consumptiefrequentie beschrijft als functie van de leeftijd. Hiermee wordt zowel de relatie tussen de (gemiddelde) consumptiefrequentie en de leeftijd als ook de variatie in de consumptiefrequentie daaromheen geschat.

2.4 Illustratie: rode wijn

Om het inschatten van de leeftijdsafhankelijkheid van de consumptiefrequentie inzichtelijker te maken, wordt in deze paragraaf een voorbeeldberekening gegeven van rode wijn.³ Figuur 3 geeft de waargenomen consumptiefrequentie van rode wijn per leeftijdscategorie (de driehoekjes), door alle individuen samen te nemen (dat wil zeggen: tel de consumptiedagen van alle individuen bij elkaar op, en deel door het totaal aantal registratiedagen). Bijvoorbeeld voor de leeftijd 40 jaar: 17 waargenomen consumptiedagen gedeeld door 218 (109 personen \times 2 registratiedagen per persoon) levert een gemiddelde consumptiefrequentie van 0,0078 op voor die leeftijdsklasse.



Figuur 3. Relatie tussen consumptiefrequentie en leeftijd beschreven met een regressie model (weergegeven door de curve).

De driehoekjes geven de waargenomen frequenties per leeftijdscategorie door de waarnemingen van alle individuen van die leeftijd samen te nemen. Voor leeftijd 40 jaar is de Beta-verdeling getekend, welke de variatie in consumptiefrequentie weergeeft voor de 40-jarigen. De geschatte parameters zijn b_1 , b_2 , c_1 , c_2 , d_1 , d_2 , θ_1 , en de parameter α van de Beta-verdeling. N.B. Boven een leeftijd van 80 jaar zijn er nog maar weinig VCP-participanten, en de waargenomen gemiddelde consumptiefrequentie kan daarom behoorlijk afwijken van de gefitte functie.

In figuur 3 is duidelijk een leeftijdsgerelateerde consumptiefrequentie zichtbaar. Met behulp van op het RIVM ontwikkelde software (PROAST, zie bijvoorbeeld Slob 2002) is het

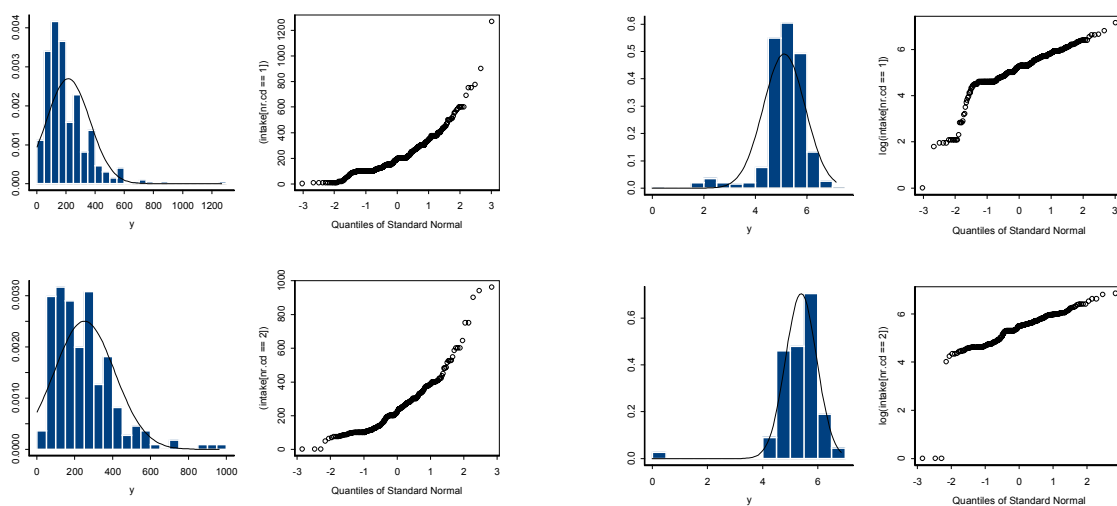
³ Er is gekozen voor de inname van rode wijn, dus het product zelf, en niet voor een bepaalde stof die aanwezig is in rode wijn. De berekening van de inname van een stof die in rode wijn aanwezig is, gaat op dezelfde wijze als in het gegeven voorbeeld, door de concentratie van de stof in rode wijn eerst te vermenigvuldigen met de geconsumeerde hoeveelheid (in feite zou je de inname van rode wijn kunnen beschouwen als een berekening van een stof in rode wijn met een concentratie van 1 per liter).

mogelijk om deze leeftijdsafhankelijkheid met een enkele functie de beschrijven, en tegelijkertijd de Beta-verdeling rondom deze functie te schatten. In dit geval bleek de (gemiddelde) consumptiefrequentie een niet-monotoon verband met de leeftijd te vertonen: na een te verwachten toename bij jong volwassenen blijkt de frequentie weer wat af te nemen bij oudere personen. De Beta-verdeling beweegt met de leeftijd mee. Met deze beschrijving ligt voor elke leeftijd de Beta-verdeling vast. Ter illustratie is er één getekend, namelijk die voor leeftijd 40 jaar. In appendix 2 wordt de uitvoering van de berekening met behulp van het computerprogramma STEM.II meer in detail uitgelegd.

3. Schatten van geconsumeerde hoeveelheid

3.1 Verdeling consumptiehoeveelheid per dag

Naast de consumptiefrequentie is de hoeveelheid van het geconsumeerde product van belang. Uiteraard hangt de hoeveelheid ook af van de leeftijd. Om deze relatie met een regressie model te beschrijven, moet een keuze gemaakt worden over de statistische verdeling van de inname-gegevens. Figuur 4 laat de verdelingen zien van de waargenomen hoeveelheid wijn per dag voor de individuen die één, resp. twee dagen wijn geconsumeerd hadden. De normale verdeling voldoet duidelijk niet, terwijl de lognormale verdeling deze gegevens redelijk goed lijkt te beschrijven. Wel is er sprake van een aantal (lage) uitbijters, kleiner dan 50 ml, enkele bedragen zelfs 1 ml. Omdat het gaat om het bepalen van hoge percentielen van de verdeling, kunnen deze uitbijters beter worden weggelaten, of als het er veel zijn, als nullen beschouwd worden. In het laatste geval moet de analyse van consumptiefrequenties opnieuw gedaan worden omdat anders de fractie van de populatie die een limietwaarde overschrijdt niet correct ingeschat wordt.

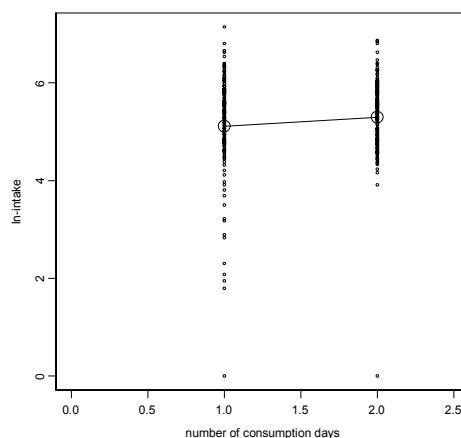


Figuur 4. Linker vier plots: verdeling van de geconsumeerde hoeveelheid wijn per dag, met de gefitte normale verdeling, voor individuen met één (boven) of twee (onder) consumptiedagen. De rechter plots geven de gegevens als QQ-plots voor de normale verdeling, welke een rechte lijn zouden moeten opleveren in geval van normaliteit. Rechter vier plots: idem, maar nu op de log-schaal. N.B. Voor de personen met twee consumptiedagen werd telkens één van de dagen random geselecteerd.

3.2 Relatie tussen consumptiehoeveelheid en –frequentie

In het model wordt aangenomen dat er geen relatie is tussen de consumptiefrequentie en consumptiehoeveelheid. Er wordt, met andere woorden, van uitgegaan dat liefhebbers die een product vaker gebruiken, niet méér per consumptiedag gebruiken dan niet-liefhebbers. Of deze aanname niet al te zeer ‘geschonden’ wordt, kan worden bekeken met de gegevens van

de VCP. De VCP bevat hieromtrent echter slechts weinig informatie, omdat met twee dagen alle individuen ofwel 50 % ofwel 100 % van de waargenomen dagen geconsumeerd hebben. De relatie tussen consumptiefrequentie en –hoeveelheid wordt dus slechts gedragen door twee punten (50 % en 100 %), en is moeilijk in te schatten. Bovendien is het niet goed mogelijk de variatie tussen individuen rondom deze relatie in te schatten. Daarom wordt in het model aangenomen dat consumptiehoeveelheid niet afhangt van consumptiefrequentie. De gegevens kunnen wel laten zien of deze aanname niet al te zeer geschonden wordt. Als illustratie laat figuur 5 zien dat er voor rode wijn inderdaad een licht positief verband tussen de consumptiehoeveelheid en de consumptiefrequentie bestaat.



Figuur 5. Hoeveelheden (op log-schaal) rode wijn geconsumeerd door personen met één, resp. twee consumptie dagen.

De grotere rondjes geven de gemiddelde waarden (op ln-schaal) weer. Op de oorspronkelijke schaal is het tweede (geometrisch) gemiddelde ruim 20 % hoger dan het eerste.

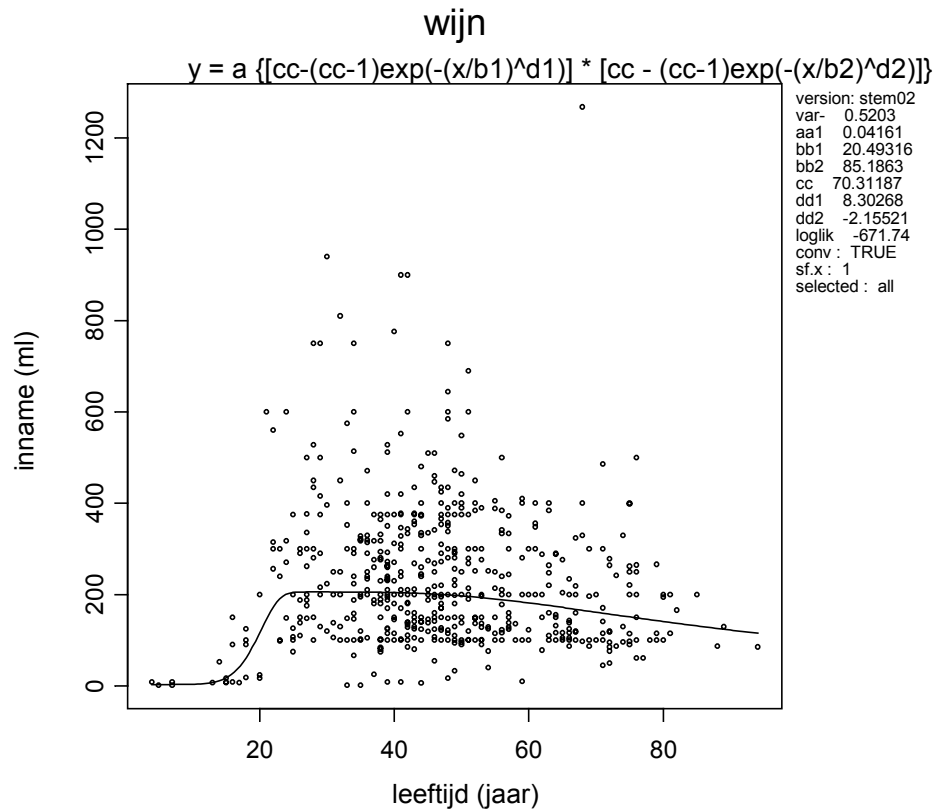
3.3 Relatie tussen consumptiehoeveelheid en leeftijd

De geconsumeerde hoeveelheid zal, net als de consumptiefrequentie, afhangen van de leeftijd. Die informatie is in principe in de VCP beschikbaar. Het verband tussen de consumptiehoeveelheid van rode wijn en leeftijd is in figuur 6 weergegeven. Het blijkt dat, na een stijging vanaf ± 16 jaar tot ± 25 jaar, er een lichte negatieve trend is met de leeftijd.

3.4 Variatie tussen personen en tussen dagen

De variatie in de consumptiehoeveelheid bestaat uit de variatie tussen de dagen (van één persoon) en de variatie tussen de personen. Als het gaat om lange-termijn inname willen we alleen de tussen-personen variatie en niet de tussen-dagen variatie meenemen in de berekening (overeenkomstig met het bestaande model STEM zie bijvoorbeeld Slob 1993, Bakker 2002). Hiertoe wordt de tussen-dagen variantie afgetrokken van de totale variantie (op de log-schaal). In STEM.II wordt de tussen-dagen variantie geschat op basis van de individuen met twee waargenomen consumptiedagen. Voor rode wijn resulteert dit in de waarde 0,18 (variantie op de ln-schaal). In de analyse van figuur 6 werd de totale variantie geschat op 0,36, zodat de variantie tussen individuen de waarden 0,18 heeft. Met deze

variantie en de gefitte regressiefunctie ligt de (lognormale) verdeling van de consumptiehoeveelheid als functie van de leeftijd vast.



Figuur 6. De waargenomen hoeveelheden geconsumeerde wijn per dag, uitgezet tegen de leeftijd. De curve geeft het gefitte model weer en weerspiegelt de mediane consumptie van wijn voor dagen dat er consumptie plaatsvindt.

4. Toepassen van het model

4.1 Inleiding

Er is nu een relatie van de consumptiefrequentie met de leeftijd geschat (Hoofdstuk 2, figuur 3) en één van de consumptiehoeveelheid met de leeftijd (Hoofdstuk 3, figuur 5). Beide relaties hebben ook ‘een derde dimensie’ in zich, in die zin dat er per leeftijdsklasse ook weer een distributie is geschat: bij de consumptiefrequentie is dat gedaan met een Beta-verdeling (zie figuur 3), bij de consumptiehoeveelheid met een lognormale verdeling (overeenkomstig met hoe dat in STEM gebeurt). De laatste verdeling weerspiegelt de variatie in consumptiehoeveelheid (per leeftijd) voor de dagen dat er consumptie is (zogenaamde conditionele verdeling).

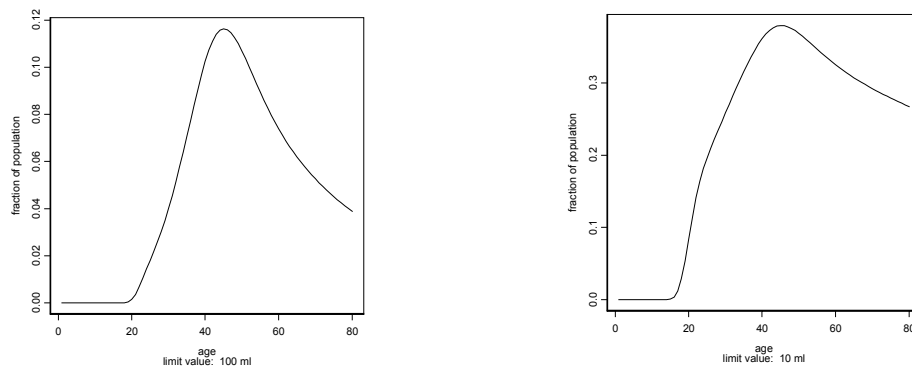
Nu worden de verdelingen van consumptiefrequentie en –hoeveelheid (als functie van de leeftijd) gecombineerd in een enkel model. Hiermee kunnen dan berekeningen gedaan worden ten aanzien van overschrijding van limietwaarden, hetzij een lange-termijn (bijvoorbeeld ADI) hetzij een korte-termijn (bijvoorbeeld Acute Reference Dose, ARfD) limietwaarde. Beide gevallen worden nu besproken.

4.2 Limietoverschrijding lange-termijn inname

Wanneer we de inname van personen willen vergelijken met een chronische blootstellingslimiet, zoals een ADI, TDI of RfD, dan is de lange-termijn inname (‘usual intake’) van een individu de relevante maat. Daarom worden voor stoffen die in het basale dieet voorkomen de korte-termijn fluctuaties binnen individuen weggemiddeld (via een indirecte statistische methode, met behulp van STEM). Voor stoffen die in incidenteel geconsumeerde producten voorkomen zal eveneens de (arithmetisch) gemiddelde lange-termijn inname als uitgangspunt gekozen worden voor vergelijking met een chronische limietwaarde. Men dient wel te bedenken dat hier een sterke aanname wordt gedaan, nl. dat de wet van Haber in extremo geldt. Dat wil zeggen dat de cumulatieve dosis (over een zekere periode) bepalend is voor het eventuele effect; als extreem geval geldt dat één keer een blootstelling aan een piekbelasting gelijk is aan constante (lage) blootstelling aan de ‘uitgesmeerde’ piek over een langere termijn. Met name voor gevallen waarbij de consumptiefrequentie erg laag is, is deze aanname twijfelachtig. Toch zal deze benadering toegepast worden, omdat er geen andere theorie voorhanden is die zegt wat de relevante lange-termijn inname is in geval van onregelmatige innames in de tijd (voor een bepaalde stof).

De fractie van de populatie die een lange-termijn inname vertoont die hoger is dan de limietwaarde kan in STEM.II berekend worden op basis van de geschatte verdelingen van frequentie en hoeveelheid. Dit kan voor elke leeftijd gedaan worden, zodat een figuur zoals figuur 8 verkregen wordt. De linker plot van figuur 8 geeft het resultaat voor de limietwaarde 100 ml, waaruit blijkt dat tot een leeftijd van 20 jaar vrijwel de gehele populatie eronder blijft, daarna neemt de fractie van de populatie die de limietwaarde overschrijdt toe, totdat bij leeftijd circa 40 jaar de maximale overschrijding plaatsvindt, namelijk een fractie van circa 11,5 % van de populatie. Met hogere leeftijd neemt de overschrijdingsfractie weer af.

Voor de limietwaarde 10 ml (figuur 8, rechter plot) wordt een vergelijkbaar patroon verkregen, maar uiteraard met hogere fracties overschrijding. De berekeningsmethode is nogal technisch en zal elders gerapporteerd worden. Voor een korte samenvatting zie appendix 3.



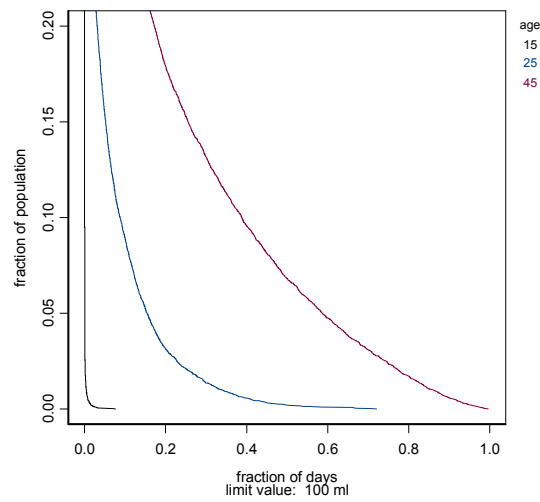
Figuur 8. Fractie van de populatie met een lange termijn inname van wijn boven de 100 ml (links) en boven de 10 ml (rechts).

4.3 Limietoverschrijding korte-termijn inname

Voor overschrijding van acute limietwaarden is de vraag wat de *kans* is dat een inname op een bepaalde dag hoger is dan die limietwaarde. Deze overschrijdingskans (of fractie dagen dat overschrijding plaats vindt) is niet voor alle individuen gelijk. Met STEM.II kan berekend worden wat de overschrijdingskans is als functie van de fractie in de populatie. Wanneer we denken aan een liefhebbersproduct dan zal voor een steeds kleiner wordende fractie van de populatie de overschrijdingskans nog kunnen toenemen (doordat de groep met een nog hogere consumptie steeds kleiner wordt).

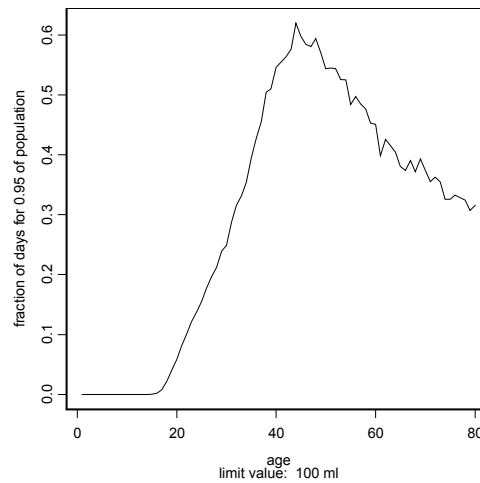
Figuur 9 illustreert dit voor de wijnconsumptie. Hier is voor drie leeftijden (15, 25 en 45 jaar) de fractie van de populatie uitgezet tegen de kans dat overschrijding plaats vindt (ofwel fractie dagen met overschrijding). Hieruit valt bijvoorbeeld af te lezen dat 10 % van de 45-jarigen een overschrijdingskans hebben van 40 % of hoger.

Figuur 10 toont de bovengrens van de overschrijdingskans (fractie dagen met overschrijding) die geldt voor 95 % van de bevolking. Met andere woorden, 95 % van de populatie zit onder de geplote overschrijdingskans. Deze is nu zichtbaar voor elke leeftijd.



Figuur 9. Fractie van de dagen dat de inname van wijn hoger is dan 100 ml (x-as) als functie van de fractie van de populatie (y-as).

Curven van links naar rechts: 15, 25 en 45 jaar. Voor de 45-jarigen kan afgelezen worden dat 10 % van de populatie de limietwaarde op circa 40 % (of meer!) van de dagen overschrijdt.



Figuur 10. Fractie dagen met overschrijding (van 100 ml) waar 95 % van de populatie onderblijft, als functie van de leeftijd.

4.4 Overzicht modelberekening

De modelberekening bestaat uit een aantal onderdelen (zie kader). De eerste twee verlopen volgens het draaiboek 2002, het laatste onderdeel bestaat uit het nieuwe model. In het nieuwe model (STEM.II) wordt de berekening stapsgewijs uitgevoerd, waarbij voor het fitten van functies het programma PROAST wordt aangeroepen (Slob, 2002). Zie appendix 2 voor een uitgebreide uitleg van het programma STEM.II.

Stappen in berekening van inname via individuele voedingsproducten

1. Concentraties van de stof aan de bijbehorende voedingsmiddelen (NEVOcodes) koppelen in database (zie draaiboek 2002)
2. Met FRIDGE de inputfiles voor STEM (en ook voor STEM.II) maken, zie draaiboek 2002
3. Met STEM.II:
 - relatie tussen consumptiefrequentie en leeftijd + Betaverdeling schatten
 - relatie tussen geconsumeerde hoeveelheid en consumptiefrequentie checken
 - relatie tussen geconsumeerde hoeveelheid en leeftijd
 - verdeling van consumptiehoeveelheid checken op basis van regressie-residuën (evt. uitbijters weglaten)
 - tussen-personen en binnen-personen variantie schatten
 - overschrijdingsfracties bepalen van chronische of acute limietwaarden

5. Discussie en conclusies

Dit rapport laat zien dat het, enigszins tegen de verwachting in, mogelijk is om op basis van de VCP-gegevens een redelijk beeld te krijgen van het consumentengedrag in de gehele populatie ten aanzien van individuele voedingsproducten, inclusief de zogenaamde liefhebbersproducten. Hierdoor is het mogelijk de inname van stoffen via individuele voedingsproducten te schatten. Dezelfde methode kan ook gebruikt worden wanneer stoffen in meerdere producten voorkomen welke incidenteel geconsumeerd worden, zodat de totale inname op veel dagen toch nul is.

Onzekerheden in het model

De computersimulaties lieten zien dat zelfs met twee registratiedagen per individu een redelijk goed beeld verkregen kan worden van de interindividuele variatie in consumptiefrequentie. Dit is mogelijk dankzij de statistische modelbenadering, waarbij een nogal sterke aanname gedaan wordt, nl. dat de consumptiefrequentie goed te beschrijven is met een Beta-verdeling, waarbij slechts één van de twee parameters leeftijdsafhankelijk is. Met een Monte Carlo benadering is het niet mogelijk om een indruk te krijgen van de variatie in consumptiefrequenties tussen individuen.

Het lijkt niet eenvoudig om te valideren of de genoemde aanname geldig is. Verder liet de simulatiestudie zien dat, als de consumptiefrequentie laag is, de mediane consumptiefrequentie enigszins over- en de fractie liefhebbers enigszins onderschat wordt (zie appendix 1). Als de consumptiefrequentie nog lager is, zal de schatting (nog) slechter worden. Men dient er daarom rekening mee te houden dat de resultaten van het model een fout kunnen bevatten, met name in gevallen met lage consumptiefrequenties.

Voor het bepalen van de relatie tussen consumptiefrequentie en consumptiehoeveelheid zijn de gegevens van de VCP met twee registratiedagen eigenlijk onvoldoende. Met meer registratiedagen zou de (gemiddelde) relatie beter geschat kunnen worden, maar zelfs dan blijft het een probleem om de variatie rond die relatie te schatten. Een voedselconsumptiepeiling met meer dagen per individu (zoals in Engeland en Frankrijk) zou ten aanzien van dit aspect daarom slechts een beperkte verbetering opleveren.

STEM.II beschrijft alleen de variabiliteit van het consumentengedrag en niet de variatie in concentraties. In het geval dat een berekening gedaan wordt van de inname van een stof via slechts één enkel voedingsproduct kan tevens rekening gehouden worden met de variatie in de concentraties in dat product. Die variatie wordt dan in termen van een variatiecoëfficiënt in STEM.II ingevoerd. Wanneer de stof in meerdere producten voorkomt, en men wenst de variatie in concentraties mee te nemen in de berekening, dan is dit met STEM.II niet mogelijk. In dat geval kan een combinatie van de Monte Carlo benadering met het statistische model van STEM.II een oplossing bieden. Daarvoor dient dan speciale software ontwikkeld te worden.

Conclusies

- Het is mogelijk om op basis van de VCP-gegevens een redelijk beeld te krijgen van de interindividuele variatie in consumentenfrequenties van individuele voedingsproducten in de populatie.
- De blootstelling aan stoffen die specifiek voorkomen in een enkel voedingsproduct kan met STEM.II beschreven worden voor de gehele populatie, rekening houdend met de variatie tussen individuen in zowel frequentie als hoeveelheid van consumptie. Daarbij kan tevens rekening gehouden worden met de variatie in concentraties in de vorm van een variatiecoëfficiënt.
- Met dit model kan afgeleid worden welke fractie van de populatie een lange termijn inname heeft die een bepaalde limiet overschrijdt. Deze fractie wordt gegeven als functie van de leeftijd. Ook kan geschat worden hoeveel mensen hoe vaak een acute limietwaarde (per dag) overschrijden, eveneens in relatie tot de leeftijd.
- Dezelfde methodologie zou gebruikt kunnen worden voor het schatten van de overschrijdingskans van een (acute) gezondheidslimietwaarde, rekening houdend met de verdeling van de concentraties in de producten. Daartoe zou dan een koppeling van dit model met Monte Carlo methoden ontwikkeld moeten worden.

Lijst van afkortingen

ADI	-	Acceptable Daily Intake
ARfD	-	Acute Reference Dose
MCRA	-	Monte Carlo Risk Analysis
RfD	-	Reference Dose
PROAST	-	Computerprogramma voor dosis-respons modellering
SIR	-	Centrum voor Stoffen en Risicoanalyse
STEM	-	Statistical Exposure Model
STEM.II	-	Statistical Exposure Model for Incidental Intakes
TDI	-	Tolerable Daily Intake
VCP	-	Voedsel Consumptie Peiling
FRIDGE	-	Computerprogramma om de VCP-database te bevragen

Referenties

- Bakker, MI (2002) Draaiboek voor modellering van inname van stoffen via voeding (in Dutch). Bilthoven, RIVM-rapport 604502001.
- Bakker, M.I., Pieters, M.N. (2002) Risk assessment of ochratoxin A in the Netherlands. Bilthoven, RIVM-rapport 388802025.
- Bakker, M.I., Baars, A.J., Baumann, R.A., Boon, P.E., Hoogerbrugge, R. (2003). Indicator PCBs in foodstuffs: occurrence and dietary intake in The Netherlands at the end of the 20th century. Bilthoven, RIVM-rapport 639102025.
- Dodd, K.W. (1996) A Technical Guide to C-SIDE: Software for Intake Distribution Estimation. Technical Report 96-TR 32, Iowa State University.
- Freijer, J.I., Hoogerbrugge, R., Van Klaveren, J.D., Traag, W.A., Hoogenboom, L.A.P., and Liem, A.K.D. (2001) Dioxins and dioxin-like PCBs in foodstuffs: Occurrence and dietary intake in The Netherlands at the end of the 20th century. Bilthoven, RIVM-rapport 639102022.
- Pieters, M.N., Freijer, J.L., Baars, A.J., Fiolet, D.C.M., van Klaveren, J., Slob, W. Risk assessment of deoxynivalenol in food. Concentration limits, exposure and effects, In: Advances in Experimental Medicine and Biology, vol. 504: Mycotoxins and food safety (J.W. DeVries, M.W. Trucksess & L.S. Jackson, eds.), Kluwer Academic/Plenum Publishers, New York, USA, 2002.
- Pieters, M.N., Bakker, M.I., Slob, W. (2004) Reduced intake of deoxynivalenol in The Netherlands: a risk assessment update. Toxicology Letters 153: 145-153.
- Slob, W. (1993) Modeling long-term exposure of the whole population to chemicals in food. Risk Analysis 13:525-530.
- Slob, W., Krajnc, E.I. (1994) Interindividual variability in modeling exposure and toxicokinetics: a case study on cadmium. Environmental Health Perspectives 102: 78-81.
- Slob, W. (2002) Dose-response modeling of continuous endpoints. Toxicol. Sci., 66, 298-312
- Teunis P.F.M., Slob W. (1999) The Statistical Analysis of Fractions Resulting From Microbial Counts. Quantitative Microbiology 1: 63-88
- Van Erp-Baart, A.M.J., Kruizinga, A., Rubingh, C., Telman, J., Brantom, P. (2001) Development of probabilistic models for describing individual intakes of chemical residues in foods. Zeist, TNO-rapport V4289

Appendix 1 Simulaties Beta-verdeling

In een simulatie-studie werd bestudeerd of met een VCP van 6000 deelnemers en 2 dagen (of met 3000 deelnemers en 4 dagen, etc. aantal deelnemers \times aantal dagen blijft gelijk) een van te voren veronderstelde Beta-verdeling teruggeschat kan worden. Allereerst worden twee waarden voor de parameters van de Beta verdeling (α en β) gekozen, dat wil zeggen er wordt een bepaalde Beta-verdeling verondersteld. Dan worden er random gegevens getrokken uit die Beta-verdeling, en wel zodanig dat een voedselconsumptiepeiling gesimuleerd wordt. De simulatie gaat in twee stappen: eerst wordt voor een individu een random consumptiefrequentie getrokken uit de Beta-verdeling, en vervolgens wordt voor dat individu voor elke registratiedag een binomiale trekking gedaan gebaseerd op de voor dat individu getrokken consumptiefrequentie. Dit is te vergelijken met het gegeven dat elk individu zijn eigen consumptiefrequentie heeft, maar of dit individu het voedselproduct daadwerkelijk tijdens de consumptiepeiling eet, is een kansproces. Dit kansproces is te beschrijven als een binomiale verdeling met als parameters het aantal registratiedagen en de werkelijke consumptiefrequentie van dat individu.

Vervolgens worden de random getrokken gegevens geanalyseerd (voor de methode zie Teunis en Slob 1999), hetgeen een schatting oplevert voor de parameters α en β van de Beta-verdeling.

Deze simulatie is uitgevoerd voor twee verschillende Beta-verdelingen. Eén met waarden voor α en β van resp. 3,6 en 4,7, vergelijkbaar met de verdeling linksboven in figuur 1 (zie voor de resultaten Tabel A.1) en één met waarden van resp. 0,15 en 1,34, vergelijkbaar met de Beta-verdeling midden boven in figuur 1 (zie voor de resultaten tabel A.2). De geschatte waarden liggen in Tabel A.1. dicht bij de originele waarden, in tabel A.2 worden α , β en de fractie van de populatie met een consumptiefrequentie van $> 80\%$ overschat, terwijl de mediaan wordt onderschat. Bij meer waarnemingen per individu (maar minder individuen) wordt de schatting iets beter.

Tabel A.1. Simulatie resultaten (1000 runs per geval) voor het schatten van de Beta-verdeelde consumptiefrequentie. De hier veronderstelde Beta-verdeling is vergelijkbaar met de verdeling links boven in figuur 2. Designs zijn aangegeven als aantal personen : aantal dagen. Het totaal aantal waarnemingen (persoon-dagen) is constant.

TRUE	α		β		median		ffc	
	3.6		4.7		0.428		0.0112	
DESIGN	mean	var	Mean	var	mean	Var	mean	var
	estimate		estimate		estimate		estimate	
6000:02	3.56	0.27	4.65	0.46	0.428	2.8E-05	0.0117	2.6E-05
4000:03	3.58	0.14	4.67	0.25	0.428	3.1E-05	0.0115	1.4E-05
3000:04	3.57	0.11	4.66	0.19	0.428	3.2E-05	0.0115	1.1E-05
1200:05	3.60	0.22	4.68	0.36	0.428	7.1E-05	0.0113	2.0E-05
1000:06	3.59	0.18	4.71	0.32	0.428	7.5E-05	0.0112	1.9E-05
600:10	3.63	0.18	4.75	0.30	0.428	9.3E-05	0.0108	1.7E-05
100:60	3.68	0.43	4.79	0.74	0.429	3.6E-04	0.0105	4.7E-05

Median = consumptiefrequentie van de mediane persoon

Fc = fractie van de populatie met een consumptiefrequentie groter dan 80%.

Var = de variantie van de schattingen van de 1000 runs. Dit is een maat voor de precisie van de schatting.

Tabel A.2. Simulatie resultaten (1000 runs per geval) voor het schatten van de Beta-verdeelde consumptiefrequentie. De hier veronderstelde Beta-verdeling is vergelijkbaar met de verdeling midden boven in figuur 2. Designs zijn aangegeven als aantal personen : aantal dagen.

TRUE	α		beta		Median		ffc	
	0.15		1.34		0.0065		0.0153	
DESIGN	mean	var	mean	var	Mean	Var	mean	var
	est.		est.		est.		est.	
6000:02	0.21	1.55E-04	1.847	1.60E-02	0.017	5.29E-06	0.0077	2.29E-06
4000:03	0.19	1.14E-04	1.671	1.19E-02	0.013	3.97E-06	0.0099	3.01E-06
3000:04	0.18	1.02E-04	1.592	1.13E-02	0.012	3.72E-06	0.0111	3.92E-06
1200:05	0.18	2.07E-04	1.562	2.49E-02	0.012	7.53E-06	0.0116	9.84E-06
1000:06	0.18	2.08E-04	1.544	2.54E-02	0.011	7.80E-06	0.0119	1.04E-05
600:10	0.17	2.29E-04	1.507	3.25E-02	0.010	8.67E-06	0.0125	1.55E-05
100:60	0.17	7.17E-04	1.465	1.60E-01	0.009	3.07E-05	0.0133	7.54E-05

Median = consumptiefrequentie van de mediane persoon

Ffc = fractie van de populatie met een consumptiefrequentie groter dan 80%.

Var = de variantie van de schattingen van de 1000 runs. Dit is een maat voor de precisie van de schatting.

Appendix 2 STEM.II: de software

- Type na opstarten van Splus: f.use() in. Op de vraag 'Do you want to use PROAST?' antwoord je 'yes' (kies de nieuwste versie) en op de vraag 'Do you want to use STEM?', ook. Vervolgens kies je de nieuwste versie van stem.nul (= STEM.II). Statistical applications heb je niet nodig. Als je in de gewenste directory bent (met behulp van f.attach()), type je de functie: f.stem.nul(argument1, argument2). De argumenten van deze functie zijn 1) de contaminantfile en 2) een bijbehorende basisfile. Deze zijn aan te maken met FRIDGE mbv de opties 'STEM contaminant intake' en 'STEM food intake' (zie Bakker, 2002).
- STEM.nul bestaat uit vier onderdelen:
 - 1) analysis of consumption frequency
 - 2) analysis of intake amounts
 - 3) estimation of variance between days (within individuals)
 - 4) calculations with the models

1. Analysis of consumption frequency

- Na de opdracht f.stem.nul gaat het programma PROAST draaien om de relatie tussen consumptiefrequentie en leeftijd te schatten. (zie ook handleiding PROAST). Je moet eerst een model kiezen (optie 2). Voor een functie die alleen stijgt of daalt volstaan de 'classical models', voor de ingewikkeldere functies kies je 'latent variable models'. Na het kiezen van het model kies je startwaarden voor de parameters (optie 3). Daarna moet je de regressiefunctie fitten (optie 4).
- Probeer verschillende modellen, als de loglikelihood winst van extra parameters groter is dan de 'critical difference' (zie tabel A.3) dan is het model met de extra parameters beter.

increase number parameters	in of	critical difference in log-likelihoods
1		1.92
2		3.00
3		3.91
4		4.74
5		5.54
6		6.30

- Als je de grafiek wilt veranderen (bijv. andere schaal op de as), kun je dat met 'plot results' doen. Ook kun je interactief de grafiek verfraaien door in het menu Options – Graph Options in het rechterbovengedeelte van de window 'Create editable graphics' aan te klikken. Nu kun je door te dubbelklikken in de grafiek van alles veranderen. Kopiëren van de grafiek naar Word gaat het handigst met Copy-Paste Special.
- Vervolgens kun je de Beta-verdeling berekenen door op de vraag: 'Do you want betabinomial analysis?' 'yes' te antwoorden. Je moet een startwaarde voor alfa (alfa van Betaverdeling, zie figuur 1) opgeven, vervolgens kun je gaan optimaliseren.
- NB. Als je 'False convergence' krijgt, is dat niet zo erg, het gaat erom dat de functie de data goed beschrijft.

- Als je tevreden bent met de fit antwoord je 'yes' op de vraag 'Do you consider this your definite fit?' De file krijgt dan de naam temp.cf, dit is te gebruiken als argument 3 in de beginfunctie f.stem.nul. In dat geval begint het programma aan het einde van stap 1 (analyse van consumptiefrequentie).

2. Analysis of intake amounts

- Het programma gaat nu eerst de relatie tussen consumptiefrequentie en -hoeveelheid analyseren. Het geeft de (ratio van de) gemiddelde innames bij een frequentie van 0,5 en 1 en geeft het resultaat ook in grafiekvorm.
- Vervolgens wordt de verdeling van de consumptiehoeveelheid bekeken, op de gewone en op de log-schaal. Met behulp van de qq-plots kun je kijken welke verdeling het beste is. Tevens vraagt het programma of je outliers wilt verwijderen. Hierop kun je het beste Nee antwoorden, het verwijderen van uitbijters dient te gebeuren ná de regressie.
- Dan begint het programma PROAST weer, om de relatie tussen consumptiehoeveelheid en leeftijd te fitten. Hier moet je een vergelijkbare procedure afwerken als bij het fitten van de consumptiefrequentie met de leeftijd. Let op dat als je de innamehoeveelheid per eenheid lichaamsgewicht wilt uitdrukken, je begint met optie 1 (Start new or modified analysis).
- Na de regressie vraagt het programma: Do you want qqplot of residuals? Indien met ja beantwoord, komt het qqplot in beeld en worden vervolgens een formeel criterium uitbijters aangewezen. De volgende vraag is of je deze wilt verwijderen. Zo ja, dan komt het gewijzigde qqplot in beeld.
- Opnieuw vraagt het programma 'Do you consider this your definite fit?' en als je bevestigend antwoordt, wordt de fit opgeslagen onder de naam temp.ca. Dit kun je gebruiken als vierde argument in de functie f.stem.nul, dan begint het programma aan het eind van stap 2. Kies vervolgens optie 6 om door te gaan met onderdeel 3.

3. Estimation of variance between days (within individuals)

- Het programma berekent de variantie tussen de individuen (VAR_{between}) door de tussen-dagen/binnen individuen variantie (VAR_{within}) van de totale variantie (MS_{total}) af te trekken. De VAR_{between} wordt gebruikt in stap 4 van het programma.

4. Calculations with the model

- Het programma vraagt of je de fractie van de populatie die een chronische grenswaarde overschrijdt wilt berekenen, bijvoorbeeld een ADI/TDI. Deze moet je zonder eenheid opgeven, die wordt namelijk in de volgende vraag gevraagd.
- Vervolgens vraagt het programma of je de plots van de consumptiefrequentie (Beta-verdeling) voor elke leeftijd wil zien.
- Het programma berekent de fractie van de populatie die de chronische RfD overschrijdt, afhankelijk van de leeftijd.
- De volgende vraag is 'Do you want to calculate the fraction of the population exceeding an acute limit value?'. Zo ja, dan geef je weer de waarde en vervolgens de eenheid. Tevens moet je een coefficient of variation (relatieve standaarddeviatie) van de concentraties in de producten opgeven, een range van leeftijden waarin je geïnteresseerd bent, en een waarde voor de maximum fractie van de populatie in de plot. Dit is ten behoeve van de figuur als figuur 9. Tenslotte moet een waarde voor de specifieke fractie van de populatie worden opgegeven om te berekenen bij welke fractie dagen de limietwaarde wordt overschreden (zie figuur 10).

Appendix 3 Berekenen (lange-termijn) limietoverschrijdende fractie van de populatie

De limietoverschrijdende fractie van de populatie wordt berekend door (per leeftijdsklasse) de consumptiefrequentie en de bijbehorende kansdichtheid (= fractie van de populatie, gegeven door de Betaverdeling) te vermenigvuldigen met de fractie van de populatie die in de conditionele verdeling van de consumptiehoeveelheid de limietwaarde overschrijdt. Dit product wordt geïntegreerd over de hele range (van 0 tot 1) van mogelijke consumptiefrequenties :

$$\Pr(ca > \text{lim}) = \int_0^1 \Pr(ca_{\text{conditioneel}} > \text{lim}) \times cf \times \Pr(cf) \, dcf \quad (1)$$

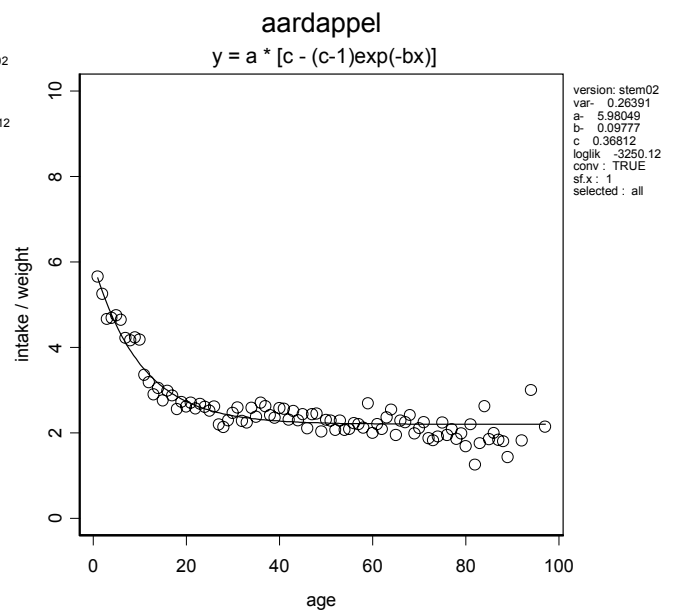
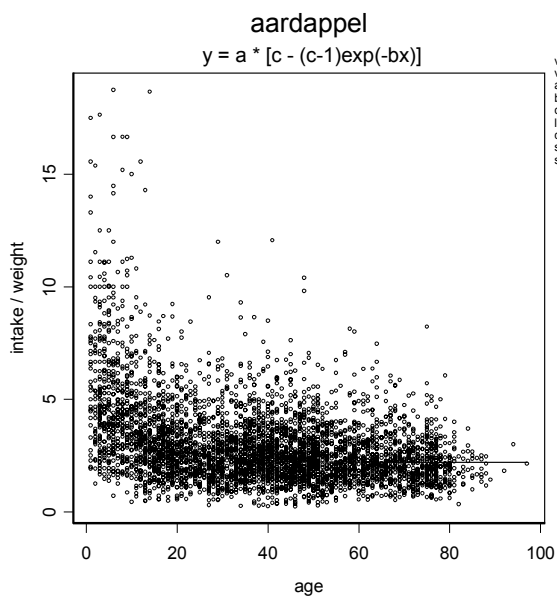
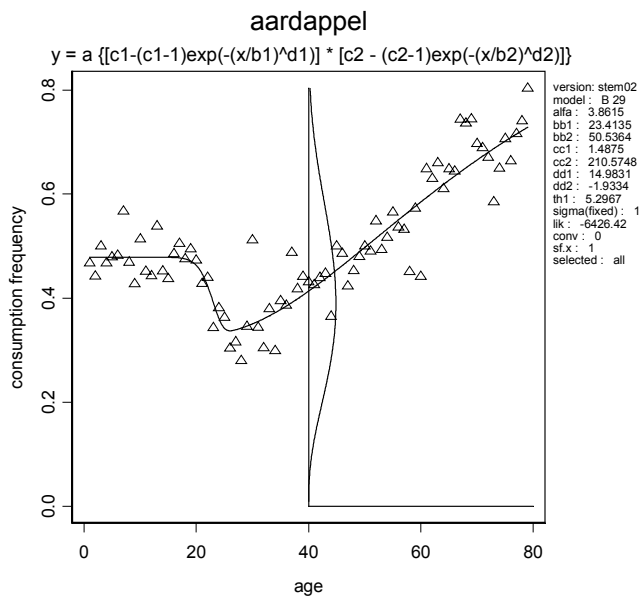
waarin Pr = probability (waarschijnlijkheid); ca = consumptiehoeveelheid; lim = limietwaarde; cf = consumptiefrequentie.

Door deze berekening voor elke leeftijdsklasse uit te voeren, kan de overschrijdingsplot getekend worden (figuur 8).

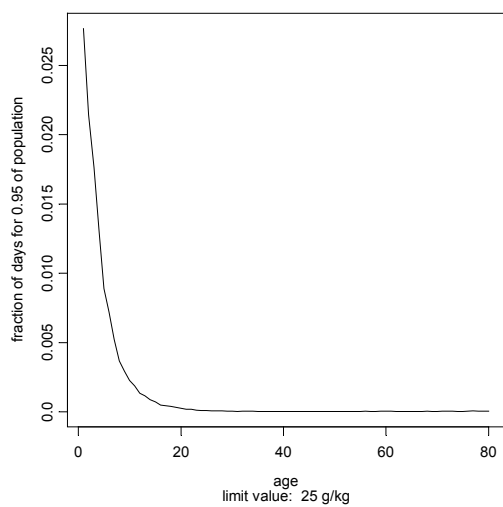
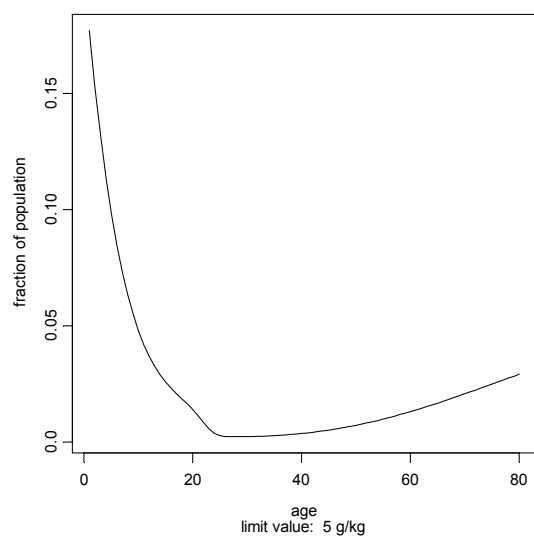
Appendix 4 Voorbeeldberekeningen

1. Aardappel

Boven links: consumptiefrequentie gefit met het latente variable model nr. 29 uit Proast.
 Onder, links: consumptiehoeveelheid per kg lichaamsgewicht gefit met Proast model 4.
 Onder, rechts: idem, alleen gemiddelden per leeftijdsklasse geplot.



Boven: limietoverschrijdende fractie van de populatie voor lange-termijn inname.
Onder: Maximale fractie van dagen met overschrijding van een acute limietwaarde welke geldt voor 95% van de populatie.

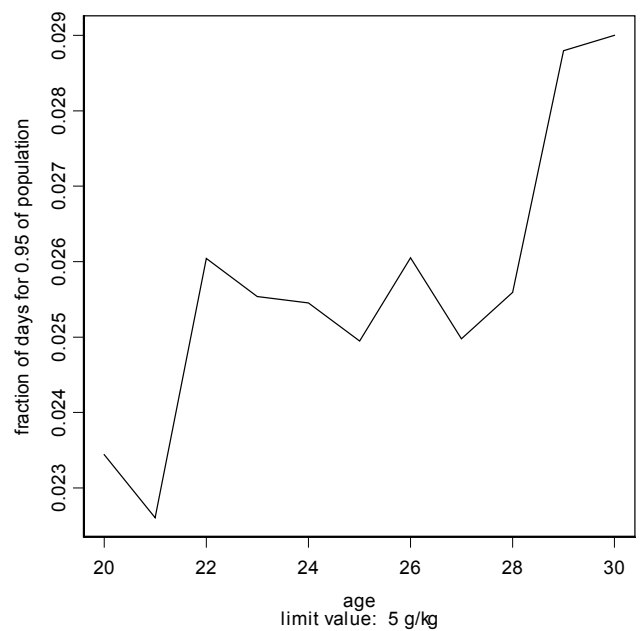
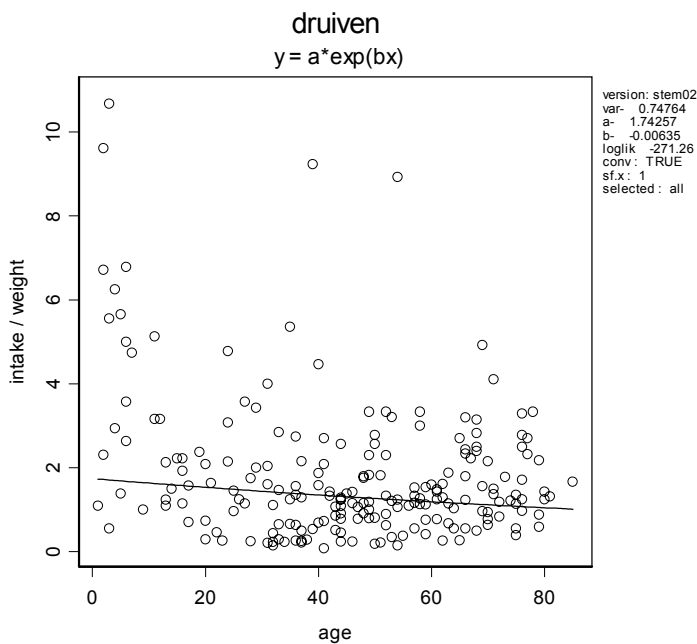
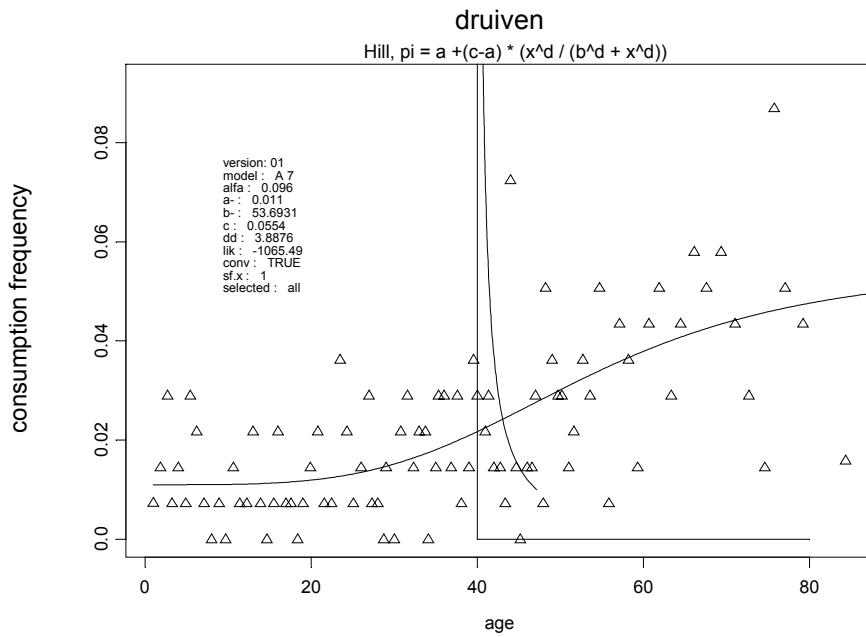


2. Druiven

Boven: consumptiefrequentie gefit met het Hill-model.

Onder, links: consumptiehoeveelheid gefit met Proast model 2.

Onder, rechts: maximale fractie van dagen met overschrijding van acute limietwaarde die geldt voor 95 % van de de populatie; limietwaarde 5 g/kg.

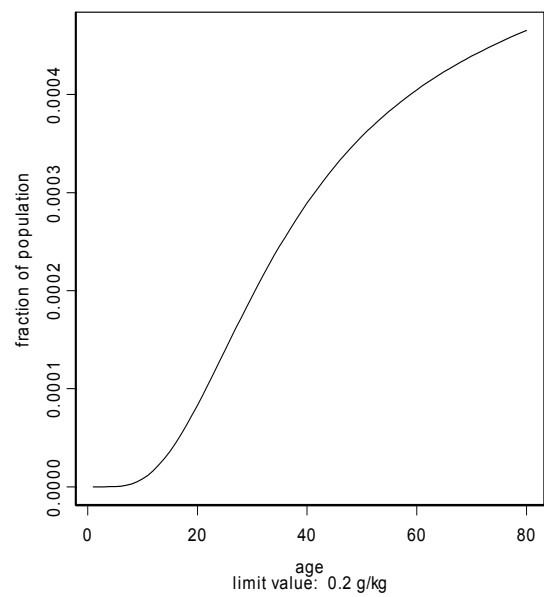
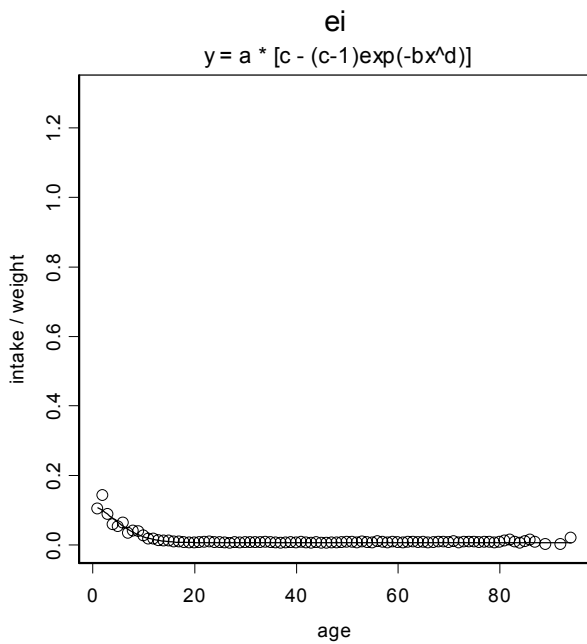
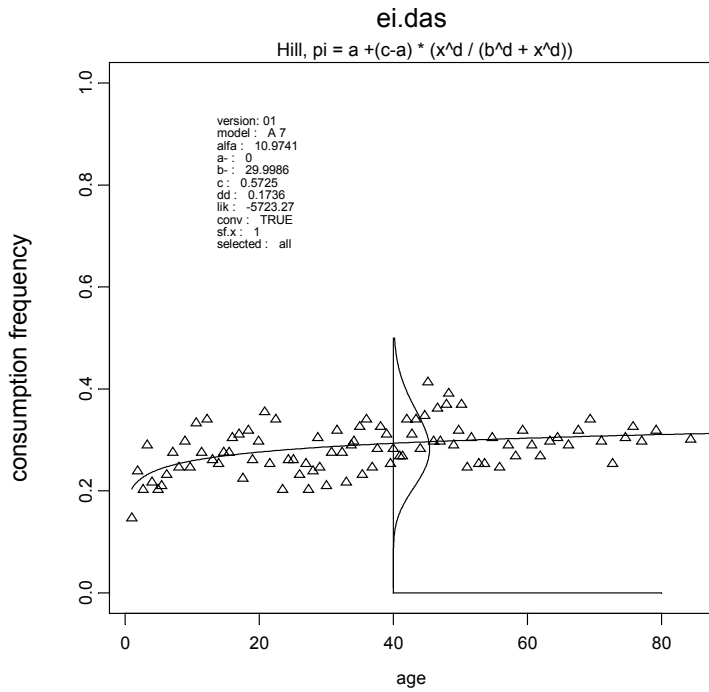


3. Ei

Boven: consumptiefrequentie gefit met het Hill-model.

Onder, links: consumptiehoeveelheid (gemiddelde waarden) gefit met Proast model 5.

Onder, rechts: fractie van de populatie die de chronische limietwaarde (gesteld op 2 g/kg) overschrijdt



Peer

Boven: consumptiefrequentie gefit met het Hill-model.

Onder, links: consumptiehoeveelheid (gemiddelde waarden) gefit met model 5.

Onder, rechts: limietoverschrijdende fractie van de populatie; chronische limietwaarde 2 g/kg.

