

Het Bayesiaanse Cox-Snell-model bij accountantscontroles

J.J.A. Moors

Samenvatting Accountants beschikken meestal over uitgebreide informatie omtrent de bedrijven waarvan zij de boeken moeten controleren. Het ligt voor de hand deze informatie te gebruiken bij het formuleren van statistische conclusies; de zgn. Bayesiaanse aanpak is hiervoor bij uitstek geschikt.

Van de Bayesiaanse modellen die zijn voorgesteld voor toepassing bij accountantscontroles lijkt dat van COX & SNELL (1979) het meest te beloven. Dit model wordt hier uitvoerig beschreven; een overzicht wordt gegeven van het verdere onderzoek dat aan dit model is gewijd.

Inhoud

1. Inleiding
 2. Populatieveronderstellingen
 3. Bayesiaanse aanpak
 4. Vergelijking met andere bovengrenzen
 5. Gemodificeerde Cox-Snell-modellen
 6. Robuuste modellen
 7. Analytische studie van het model
 8. Discussie
- Literatuur

1. Inleiding

Een van de centrale problemen bij accountantscontroles op steekproefbasis is het bepalen van een 95%- betrouwbaarheidsbovengrens voor de fout die het totaal van een populatie van posten vertoont. Met name voor populaties met slechts weinig foutieve posten leveren de klassieke statistische methoden geen bevredigende oplossing.

Van belang is verder dat de accountant die de populatie moet onderzoeken vaak op voorhand een globale indruk zal hebben van de grootte van de totale fout. Deze indruk zal gebaseerd zijn op de kwaliteit van de interne procedures (en de naleving daarvan) bij het bedrijf in kwestie of bij soortgelijke bedrijven. Het zou mooi zijn als deze voorkennis expliciet gebruikt zou kunnen worden.

Door een Bayesiaanse benadering van dit probleem wordt aan deze beide aspecten recht gedaan. Vanwege de grote flexibiliteit wordt hier in het bijzonder het model besproken dat is voorgesteld door COX & SNELL (1979).

Dit model wordt geïntroduceerd in de volgende twee paragrafen. In Par. 2 worden de veronderstellingen besproken die in het Cox-Snell-model worden gemaakt t.a.v. de populatie. De belangrijkste vooronderstelling is dat uitsluitend positieve fouten kunnen voorkomen: het opgegeven bedrag van een post is nooit lager dan het werkelijke bedrag. Verder wordt aangenomen dat de grootte van de fouten een exponentiële verdeling vertoont. Steeds wordt ervan uitgegaan dat de steekproeftrekking van de posten plaats vindt met kansen evenredig met de grootte. Zie voor een uitvoerige discussie van de statistische aspecten in de accountancy bijv. LESLIE e.a. (1979).

De Bayesiaanse aanpak komt in Par. 3 aan de orde. De priorverdeling wordt besproken; kort wordt weergegeven hoe hieruit, in combinatie met de waarnemingen, de posteriorverdeling volgt. Deze posteriorverdeling levert dan de gewenste Bayesiaanse 95%- betrouwbaarheidsbovengrens.

Aan het Cox-Snell-model is reeds vrij veel nader onderzoek verricht; een overzicht hiervan wordt gegeven in Par. 4-7. De (schaarse) vergelijkingen met meer klassieke bovengrenzen worden besproken in Par. 4.

Omdat de modeleisen van Cox en Snell nogal streng zijn en op sommige punten enigszins onrealistisch, bekeken GODFREY & NETER (1984) enkele varianten. Voor uiteenlopende parameterwaarden vergeleken zij de hieruit af te leiden bovengrenzen. Hun voornaamste conclusie is dat de invloed van de priorwaarden van de parameters veel groter is dan van de precieze modelformulering. Wat meer details zijn te vinden in Par. 5.

In Par. 6 worden de resultaten besproken van de simulatiestudies van GODFREY & NETER (1984) en NETER & GODFREY (1985). Daarbij waren negen verschil-

lende sets parameterwaarden betrokken voor het oorspronkelijke Cox-Snell-model, alsmede 21 populaties. Het voornaamste doel daarbij was een set priorwaarden voor de parameters op te sporen die voor een groot aantal uiteenlopende populaties bruikbaar zijn.

Het feit dat de 95%- betrouwbaarheidsbovengrens verkregen m.b.v. het Cox-Snell-model zo'n simpele uitdrukking is, wordt optimaal benut in Par. 7. De kansverdeling van deze bovengrens wordt voor enkele populaties bepaald volgens de methode die is aangegeven in MOORS (1983) en nader uitgewerkt in MOORS & JANSSENS (1987). Deze aanpak kan heel wat simulatiestudies overbodig maken.

Tenslotte worden in Par. 8 de bevindingen bediscussieerd en voorstellen voor verder onderzoek gedaan.

2. Populatieveronderstellingen

Een eindige populatie dient te worden gecontroleerd welke bestaat uit N posten, genummerd van 1 tot N . Laat de variabele X het aangegeven (nominale) bedrag voorstellen van een bepaalde post en de variabele Y de (niet-negatief veronderstelde) fout in dit bedrag, zodat $X-Y$ het correcte bedrag is van deze post. Dan stellen

$$\sum_{i=1}^N X_i, \quad \sum_{i=1}^N Y_i$$

het totale nominale bedrag voor van de N posten uit de populatie, resp. het totale ten onrechte opgevoerde bedrag; het correcte populatietotaal bedraagt dus $\sum_{i=1}^N X_i - \sum_{i=1}^N Y_i$. De populatieparameter ψ die hier centraal staat is de totale relatieve fout, de fractie van het totale nominale bedrag dat ten onrechte is opgevoerd:

$$(2.1) \quad \psi = \frac{\sum_{i=1}^N Y_i}{\sum_{i=1}^N X_i}$$

Steeds wordt uitgegaan van 'monetary unit sampling', hetgeen inhoudt dat steekprofelementen getrokken worden met kansen evenredig met het nominale bedrag X_i . Bij trekking van één steekprofelement geldt dus voor de trekkingskans p_i van populatiepost i

$$p_i = X_i / \sum_{i=1}^N X_i, \quad i=1,2,\dots,N$$

Laat de kansvariabele $\underline{S}$ (kansvariabelen worden onderstreept) de fractie aangeven van het bedrag van deze éne steekproefpost dat incorrect is; $\underline{S}$ neemt dan de waarden

$$Y_i/X_i, i=1,2,\dots,N$$

aan met kansen p_i . Voor de verwachting $E(\underline{S})$ van $\underline{S}$ geldt dan per definitie

$$(2.2) \quad E(\underline{S}) = \sum_{i=1}^N p_i Y_i/X_i = \sum_{i=1}^N Y_i / \sum_{i=1}^N X_i = \psi$$

ofwel $\underline{S}$ is een zuivere schatter voor ψ . Deze eigenschap blijft gelden, als een steekproef van n elementen getrokken wordt met kansen evenredig met X en met teruglegging. Voor de waarnemingen $\underline{S}_1, \underline{S}_2, \dots, \underline{S}_n$ met de definities

$$(2.3) \quad \underline{Z} := \sum_{i=1}^n \underline{S}_i, \quad \bar{\underline{Z}} := \underline{Z}/n$$

geldt nl. dat ook $\bar{\underline{Z}}$ zuiver is voor ψ :

$$E(\bar{\underline{Z}}) = E\left(\frac{1}{n} \sum_{i=1}^n \underline{S}_i\right) = \frac{1}{n} \sum_{i=1}^n E(\underline{S}_i) = \psi$$

Door gebruik te maken van de steekproefvariantie S_z^2 van $\bar{\underline{Z}}$ kan dan bijv. het klassieke rechtseenzijdige betrouwbaarheidsinterval

$$(2.4) \quad (0, \bar{\underline{Z}} + 1.645 S_z)$$

worden bepaald, dat ψ bevat met een betrouwbaarheid van ongeveer 95% voor grote n . Hierin is de factor 1.645 gebaseerd op de normale benadering van de kansverdeling van $\bar{\underline{Z}}$.

Aan deze klassieke aanpak kleven twee bezwaren.

(i) Veel populaties bevatten slechts weinig incorrecte posten, zodat de kans niet denkbeeldig is dat alle waarnemingen $\underline{S}_i$ de waarde nul aannemen en (2.4) onbruikbaar wordt. (Bij zeer weinig positieve $\underline{S}_i$ wordt voorts het gebruik van de normale verdeling twijfelachtig.)

(ii) De voorkennis waarover de accountant veelal beschikt, zowel m.b.t. het specifieke onderzochte bedrijf als m.b.t. het type bedrijf (branche, grootte) in zijn algemeenheid, wordt niet expliciet gebruikt.

Aan deze beide bezwaren wordt tegemoet gekomen door het gebruik van Bayesiaanse methoden. Hierbij wordt immers de aanwezige voorkennis in het model geïncorporeerd, terwijl ook in het geval weinig of geen incorrecte posten worden waargenomen een betrouwbaarheidsinterval kan worden bepaald. In het volgende wordt uitsluitend aandacht besteed aan het model van COX & SNELL (1979). Weliswaar kwamen GARSTKA (1977) en VANECEK (1978) eerder met verwante modellen, maar

deze zijn minder flexibel dan het Cox-Snell-model, zoals aan het eind van de volgende paragraaf zal worden toegelicht.

Het Cox-Snell-model splitst allereerst de centrale parameter ψ uit (2.1) in twee afzonderlijke factoren. Laat de posten $1, 2, \dots, k$ precies de incorrecte posten uit de populatie zijn. Dan is (2.1) ook te schrijven als

$$(2.5) \quad \psi = \left(\sum_{i=1}^k X_i / \sum_{i=1}^N X_i \right) * \left(\sum_{i=1}^k Y_i / \sum_{i=1}^k X_i \right)$$

De eerste factor hierin is juist de kans op een incorrecte post (bij de gegeven trekkingsmethode); de tweede factor representeert de totale relatieve fout t.o.v. de incorrecte posten. De formule weerspiegelt het feit, dat hoge waarden van ψ kunnen ontstaan door veel kleine fouten, maar ook door weinige grote fouten.

Vervolgens wordt de populatie gestileerd door aan te nemen dat de kansverdeling van $\underline{S}$ beschreven kan worden door een eenvoudige analytische uitdrukking. Daarvoor is een gemengde kansverdeling genomen: bij de waarde $\underline{S} = 0$ behoort een puntmassa $1-\phi$, terwijl positieve waarden van S een exponentiële verdeling met parameter λ vertonen (notatie $\underline{S} | (\underline{S} > 0) \sim \text{Ne}(\lambda)$):

$$(2.6) \quad \begin{cases} P(\underline{S} = 0) = 1-\phi \\ P(s < \underline{S} < s + ds) = \phi \lambda \exp(-\lambda s) ds, s > 0 \end{cases}$$

Hierin is ds een infinitesimale afstand. Blijkbaar wordt de kansverdeling van $\underline{S}$ bepaald door de twee parameters ϕ en λ , die corresponderen met de twee factoren in (2.5). (Een hoge waarde van λ betekent daarbij dat de kans op grote fouten klein is.) Het verband met (2.5) kan ook worden getoond door de verwachting van (2.6) te berekenen:

$$E(\underline{S}) = \phi/\lambda$$

en combinatie met (2.2) geeft nu

$$(2.7) \quad \psi = \phi/\lambda$$

De twee factoren ϕ en $1/\lambda$ hierin hebben een soortgelijke interpretatie als in (2.5). Merk op dat de hier gepostuleerde kansdichtheid van $\underline{S}$ positief is voor alle $s > 0$, terwijl in werkelijkheid geldt $\underline{S} \leq 1$ zoals onmiddellijk uit de definitie volgt. Deze stilering is geen bezwaar zolang met name λ relatief grote waarden aanneemt. De kans

$$(2.8) \quad P(\underline{S} > 1) = \phi e^{-\lambda}$$

is dan immers te verwaarlozen; zo heeft (2.8) voor $\phi = \frac{1}{2}$ en $\lambda = 5$ de waarde 0.0034.

Dankzij (2.6) krijgt ook de kansverdeling van de centrale steekproefgrootheid $\underline{Z}$ uit (2.3) een eenvoudige vorm. Zij $\underline{M}$ het aantal incorrekte posten in de steekproef; dan heeft $\underline{M}$ de binomiale verdeling $B(n, \phi)$, die voor niet te grote waarden van ϕ goed te benaderen is door de Poissonverdeling $Po(n\phi)$. Voor $\underline{M} = 0$ geldt uiteraard $\underline{Z} = 0$, zodat $\underline{Z}$ hier een puntmassa heeft. Onder het gegeven $\underline{M} = m > 0$ is $\underline{Z}$ de som van m onafhankelijke $\underline{S}_i \sim Ne(\lambda)$; de voorwaardelijke verdeling van $\underline{Z}$ is dan de gamma-verdeling $\Gamma(\lambda, m)$:

$$(2.9) \quad \underline{Z} \mid (\underline{M}=m) \sim \Gamma(\lambda, m)$$

(zie bijv. van der GENUGTEN (1986), st. 4.4.10). Samenvattend wordt de gezamenlijke kansverdeling van $\underline{M}$ en $\underline{Z}$ gegeven door

$$(2.10) \quad \begin{cases} P(\underline{M}=0, \underline{Z}=0) = \exp(-n\phi) \\ P(\underline{M}=m, z < \underline{Z} < z+dz) = \frac{(n\phi\lambda z)^m \exp(-n\phi - \lambda z)}{zm! (m-1)!} dz \\ \text{voor } z > 0 \text{ en } m=1, 2, \dots \end{cases}$$

3. Bayesiaanse aanpak

Nu pas begint de Bayesiaanse aanpak. De twee parameters ϕ en λ worden niet meer beschouwd als constanten, maar als kansvariabelen; zij worden dan ook genoteerd als $\underline{\phi}$ en $\underline{\lambda}$. Hun kansverdeling, de zgn. priorverdeling wordt door de onderzoeker vastgelegd op basis van de kennis waarover hij op voorhand reeds beschikt. Er zijn twee gebruikelijke interpretaties van deze Bayesiaanse aanpak.

(i) De populatie die onderzocht moet worden, wordt beschouwd als een trekking uit een hele collectie van dergelijke populaties, elk met een specifiek paar parameters ϕ en λ . De relatieve frequentieverdeling van deze parameters over de hele collectie van populaties is dan juist de kansverdeling van $\underline{\phi}$ en $\underline{\lambda}$. De priorverdeling dient deze zo dicht mogelijk te benaderen. Voor de Bayesiaanse onderzoeker is het de kunst zijn kennis omtrent overeenkomstige populaties te gebruiken voor het kiezen van de priorverdeling.

(ii) De aanwezige voorkennis van de onderzoeker over de te onderzoeken populatie wordt gebruikt om een maat van vertrouwen aan te geven in de gebeurte-

nis dat $\underline{\phi}$ en $\underline{\lambda}$ bepaalde waarden aannemen. Door dit voor alle waarden te doen wordt de priorverdeling verkregen.

Merk op dat de tweede interpretatie puur subjectivistisch is, terwijl de eerste ook frequentistische aspecten heeft. De benodigde voorkennis betreft bij (ii) de specifieke onderzochte populatie en bij (i) ook overeenkomstige populaties. Het zal overigens duidelijk zijn dat beide interpretaties naast en door elkaar worden gehanteerd.

Vanwege de analytische en rekentechnische voordelen is het gebruikelijk slechts priorverdelingen te beschouwen van een vast type, waarbij slechts de parameterwaarden (eventueel) kunnen variëren. Mits het gekozen type voldoende flexibel is, is dit ook geen bezwaar, omdat vaak de precieze vorm van de priorverdeling van ondergeschikt belang is. In par. 4 wordt hierop teruggekomen.

COX & SNELL (1979) nemen nu als type priorverdeling voor zowel $\underline{\phi}$ als $\underline{\lambda}$ de gamma-verdeling en veronderstelden voorts onafhankelijkheid tussen $\underline{\phi}$ en $\underline{\lambda}$. In voor de hand liggende notatie is dit samen te vatten tot

$$(3.1) \quad (\underline{\phi}, \underline{\lambda}) \sim \Gamma(a, b) \# \Gamma(c, d)$$

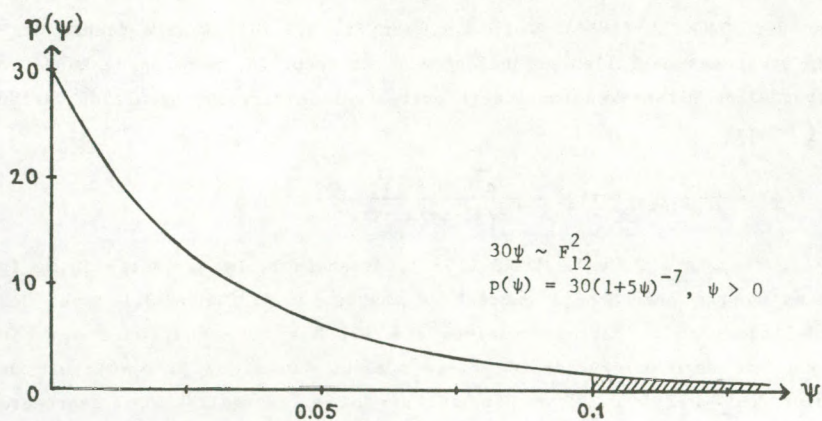
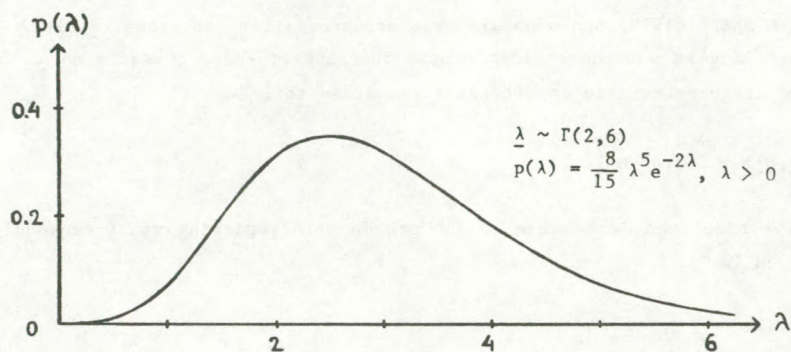
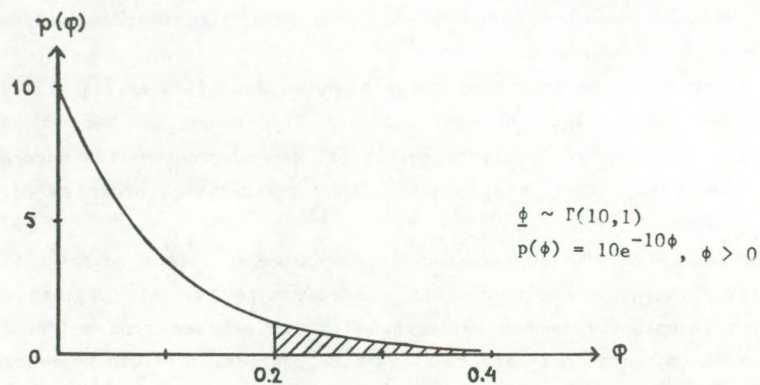
Het eerste voordeel van deze keuze is dat ook de priorverdeling van $\underline{\psi}$ eenvoudig te beschrijven is:

$$(3.2) \quad \frac{ad}{bc} \underline{\psi} \sim F_{2d}^{2b}$$

(vgl. van der GENUGTEN (1986), Def. 4.9.9 en St. 4.9.13), waarin rechts de F-verdeling staat met aantallen vrijheidsgraden $2b$ resp. $2d$. Merk op de keuze van de priorverdeling uitspraken impliceert over zowel de verwachting als de variantie van $\underline{\psi}$ wegens

$$E(\underline{\psi}) = \frac{bc}{a(d-1)}, \quad \text{Var}(\underline{\psi}) = \frac{c^2 b(b+d-1)}{a^2 (d-1)^2 (d-2)}$$

voor $d > 1$, resp. $d > 2$; vgl. MOORS (1983), Appendix B. De variantie is te interpreteren als de onzekerheid omtrent de waarde van $\underline{\psi}$. Figuur 3.1 toont deze priorverdelingen met de parameterwaarden $a = 10$, $b = 1$, $c = 2$, en $d = 6$. Een onderzoeker die deze priorverdeling kiest spreekt dus als zijn mening uit dat o.a. geldt $P(\underline{\phi} > 0.2) = 0.135$ en $P(\underline{\psi} > 0.1) = 0.088$. (Vergelijk de gearceerde oppervlakten in Figuur 3.1.) Precies zo geldt $P(\underline{\psi} < 0.130) = 0.95$; m.a.w. 0.13 is een a priori 95%- bovengrens is voor $\underline{\psi}$, die uitsluitend de voorkennis van de onderzoeker weergeeft. Merk op dat de kansverdeling van $\underline{\phi}$ weer het bezwaar heeft dat $\underline{\phi}$ niet beperkt is tot het eenheidsinterval; voor grote (kleine) waarden van a (b) is dit slechts van theoretisch belang.



Figuur 3.1. Priorverdelingen Cox-Snell-model.

De volgende stap is nu de priorverdeling te combineren met de informatie die de empirische waarnemingen leveren. In het licht van de waarnemingen zal de onderzoeker zijn mening kunnen wijzigen: de priorverdeling wordt vervangen door de zgn. posteriorverdeling. Deze weerspiegelt dus zowel de voorkennis van de onderzoeker als de nieuwe, empirische informatie. De technische redenering is als volgt.

In de Bayesiaanse optiek wordt de kansverdeling (2.9) van $\underline{M}$ en $\underline{Z}$ opgevat als een voorwaardelijke verdeling onder het gegeven dat $\underline{\phi}$ en $\underline{\lambda}$ de waarden ϕ en λ aannemen. De gezamenlijke kansverdeling van het viertal $(\underline{M}, \underline{Z}, \underline{\phi}, \underline{\lambda})$ wordt dan verkregen als het produkt van (2.9) met (3.1). Via standaardmethoden (zie bijv. van der GENUGTEN (1986), par. 4.3) wordt dan de voorwaardelijke verdeling van $(\underline{\phi}, \underline{\lambda})$ verkregen, gegeven de realisaties m en z van $\underline{M}$ en $\underline{Z}$. Deze voorwaardelijke verdeling is juist de bovenbedoelde posteriorverdeling. Concreet volgt:

$$(3.3) \quad (\underline{\phi}, \underline{\lambda}) \mid (\underline{M}=m, \underline{Z}=z) \sim \Gamma(a+n, b+m) * \Gamma(c+z, d+m)$$

Deze uitdrukking geldt ook nog in het speciale geval dat geen enkele incorrecte post wordt aangetroffen; wegens $z = m = 0$ geldt dan

$$(3.4) \quad (\underline{\phi}, \underline{\lambda}) \mid (\underline{M}=0, \underline{Z}=0) \sim \Gamma(a+n, b) * \Gamma(c, d)$$

Zelfs geldt dat (3.1) een speciaal geval is van (3.3)! Wanneer geen enkele waarneming gedaan wordt ($n=0$), vindt men vanzelfsprekend ook geen incorrecte posten: (3.4) gaat dan over in (3.1). Dat de vorm van de prior- en de posteriorverdeling dezelfde is, vereenvoudigt de analyse en is het tweede grote voordeel van het Cox-Snell-model. Uiteraard geldt deze eigenschap ook voor $\underline{\psi}$: uit (3.3) valt af te leiden

$$(3.5) \quad \frac{(a+n)}{(b+m)} \frac{(d+m)}{(c+z)} \underline{\psi} \sim F_{2(b+m)}^{2(a+n)} \frac{2(d+m)}{2(c+z)}$$

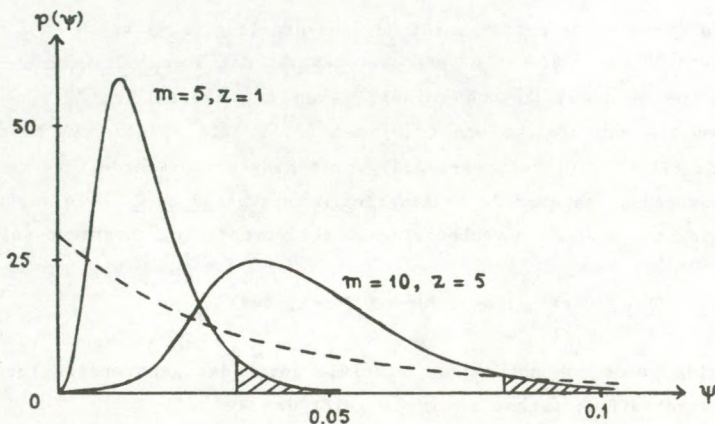
Deze uitdrukking geldt ook voor $m = 0$ en zelfs voor $n = 0$: vergelijk (3.2).

Bovenstaande afleiding berust in essentie op COX & SNELL (1979), dat echter nogal globaal is en bovendien een foutief resultaat geeft. MOORS (1983) geeft een complete afleiding van de correcte verdeling (3.5).

Beschouw als voorbeeld de priorverdeling uit Figuur 3.1, met de priorwaarden $a = 10$, $b = 1$, $c = 2$ en $d = 6$. Neem aan dat 100 waarnemingen zijn gedaan die vijf incorrecte posten opleverden; deze vijf posten waren gemiddeld 20% te hoog. In symbolen weergegeven geldt dus $n = 100$, $m = 5$ en $z = 1$, zodat (3.5) nu luidt

$$67.22 \psi \sim F_{22}^{12}$$

Figuur 3.2 toont deze posteriorverdeling, welke min of meer geconcentreerd is rond het steekproefgemiddelde $\bar{z} = 0.01$. De figuur toont nog een tweede numeriek voorbeeld, alsmede de priorverdeling (gestippeld).



Figuur 3.2 Posteriorverdelingen voor $n = 100$

Eenvoudig is nu een a posteriori 95%-bovengrens voor ψ af te leiden. Uit (3.5) volgt immers onmiddellijk dat met 95% kans geldt:

$$(3.6) \quad \psi < \frac{(b+m)}{(a+n)} \frac{(c+z)}{(d+m)} F_{2(d+m)}^{2(b+m)}; 0.05$$

waarin de laatste factor het 95%-kwantiel is, te vinden is in tabellen van de F-verdeling. Voor de twee verdelingen in Figuur 3.2 vinden we de bovengrenzen 0.0332 en 0.0823, zodat de gearceerde oppervlakten daar beide 5% bedragen.

Tabel 3.1 geeft een completer overzicht van deze bovengrenzen, steeds uitgaande van $n = 100$ en van dezelfde priorverdeling. Hierin geldt $\bar{z}^* = z/m$, het gemiddelde over de incorrecte posten.

Het Cox-Snell-model laat afzonderlijke priorverdelingen toe zowel voor de kans op een incorrecte post (ϕ) als voor de gemiddelde relatieve fout per incorrecte post ($1/\lambda$). Eerder gepubliceerde Bayesiaans modellen namen slechts één van deze beide parameters in ogenschouw. Zo behandelt GARSTKA (1977) een samengesteld Poissonmodel met een priorverdeling voor $1/\lambda$. Daarentegen poneert VANECEK (1978) juist een (β -)priorverdeling voor ϕ . Het hier besproken model is daarom te beschouwen als een algemenere en meer flexibele versie van deze twee voorgangers.

Tabel 3.1 A posteriori 95%- bovengrenzen voor ψ (in %) bij priorverdeling
 $(\phi, \lambda) \sim \Gamma(10,1) * \Gamma(2,6)$

m	$\frac{-*}{z}$						
	0	0.1	0.2	0.3	0.5	1	
0	1.18	---	---	---	---	---	
2	---	2.05	2.24	2.43	2.80	3.74	
5	---	2.76	3.32	3.87	4.98	7.74	
10	---	3.53	4.70	5.88	8.23	14.10	
20	---	4.76	7.14	9.52	14.27	26.17	

4. Vergelijking met andere bovengrenzen

Een betrouwbaarheidsgrens voor ϕ kan al verkregen worden door uitsluitend te letten op de aantallen aangetroffen fouten en niet op de grootte van deze fouten. Geef met ϕ_m de 95%-betrouwbaarheidsbovengrens aan als onder n waarnemingen m fouten worden aangetroffen; zie bijv. van der GENUGTEN (1986), par. 5.4.8 voor de berekening. Met de grootte van de waargenomen fouten kan als volgt rekening worden gehouden. Geef met T_i de naar grootte geordende relatieve fouten S_i aan, zodat $1 > T_1 > T_2 > \dots > T_m$. Dan is

$$(4.1) \quad \phi_0 + \sum_{i=1}^m T_i(\phi_i - \phi_{i-1})$$

de in de accountancy zeer bekende Stringer-bovengrens; zie voor een uitvoerige behandeling LESLIE e.a. (1979), Ch. 7.

Een vergelijking van de Cox-Snell bovengrens (3.6) met o.a. de Stringer-bovengrens werd gestart door SMITH (1979). Bij zijn onderzoek nam hij de volgende drie klassieke 95%- bovengrenzen mee (de laatste letters van de namen staan voor Upper Error Limit):

- ATUEL: de Stringer-bovengrens;
- MATUEL: de lichtelijk geamendeerde versie van ATUEL, die ontstaat door de aangetroffen foutieve posten te corrigeren;
- PBUEL: een pseudo-Bayes variant, waarbij de priorwaarde van 0.5 wordt gehanteerd voor de gemiddelde relatieve fout per incorrecte post.

Zie SMITH (1979) voor details.

Voor de bovengrens in (3.6) nam Smith de door COX & SNELL (1979) aanbevolen parameterwaarden. Helaas onderkende hij de eerder vermelde fout in de

afleiding niet, die tot gevolg heeft dat zijn parameterwaarde $a = \infty$ (in de notatie (3.1)) niet toegestaan is. MOORS (1983) herstelde deze fout door als parameterwaarden te nemen $a = 100$, $b = 1$, $c = 1$ en $d = 3$, in de lijn van de suggesties van Cox en Snell. Als vierde bovengrens werd zodoende

$$(4.2) \quad \ell_{cs}^* := \frac{(m+1)(z+1)}{(n+100)(m+3)} F_{2m+6; 0.05}^{2m+2}$$

in de vergelijking betrokken.

De vier bovengenoemde grenzen werden berekend voor een populatie, die door Smith speciaal voor dit doel was gegenereerd. Deze populatie bestaat uit 1000 posten met een totale boekwaarden van f 98165 ($=EX_1$); hiervan is f 972 ($=EY_1$) ten onrechte geboekt, veroorzaakt door 12 foutieve posten. Trekking van alle mogelijke systematische steekproeven leverde 981 waarden van deze bovengrenzen. Tabel 4.1 geeft de frequentieverdelingen bij een steekproef-omvang van 100 ($=n$). De laatste kolom is berekend uit Tabel 3 van SMITH (1979) en kan enigszins afwijken van de exacte waarden. Om dezelfde reden is de klassenindeling voor ℓ_{cs}^* nogal grof.

Tabel 4.1. Frequentieverdelingen van diverse 95%- bovengrenzen

%	ATUEL	MATUEL	PBUEL	ℓ_{cs}^*
< 2			171	553
2-2.5		61	315	338
2.5-3	61	56	160	90
3-3.5		54	96	
3.5-4	406	380	224	
4-4.5	84	2	15	
4.5-5		289		
5-5.5	280	49		
5.5-6	60	75		
6-7	90	15		
Totaal	981	981	981	981

Onmiddellijk springt in het oog dat met name de eerste twee grenzen erg conservatief zijn: hoewel de totale relatieve fout ψ slechts 1% bedraagt, zijn waarden boven 6% voor de bovengrenzen geen uitzondering. PBUEL doet het beter, maar de Cox-Snell bovengrens ℓ_{cs}^* neemt gemiddeld verreweg de laagste waarde aan. Verstreckende conclusies kunnen hieruit niet worden getrokken; het resultaat hangt af van de beschouwde populatie én van de gekozen a priori waarden van de

parameters. Wel illustreert dit dat het Cox-Snell model tot een lage 95%-boven-grens kan leiden, met name als de waarnemingen redelijk overeenstemmen met de priorverdeling.

Een meer gedetailleerde vergelijking van diverse 95%- bovengrenzen werd ondernomen door MENZEFRIKKE & SMIELIAUSKAS (1984). Zij vergeleken de prestaties van diverse Bayesiaanse grenzen - waaronder die van Cox en Snell - bij verschillende populaties. Bij hun simulaties gaan zij echter steeds uit van populaties waarin ook negatieve fouten kunnen voorkomen (met bijbehorende $Z_1 < 0$). Aangezien het Cox-Snell-model niet op deze situatie van toepassing is, blijven hun resultaten hier verder buiten beschouwing.

5. Gemodificeerde Cox-Snell-modellen

Zoals in Par. 3 bleek, is het grote voordeel van de modelveronderstellingen van Cox en Snell dat de verkregen posterior bovengrens een eenvoudige analytische vorm heeft. De studie van de eigenschappen van deze grens wordt daardoor aanzienlijk vergemakkelijkt. De keerzijde van de medaille is dat de gemaakte veronderstellingen tamelijk strikt zijn en op sommige punten zelfs onrealistisch. Hieronder volgt een opsomming van de zwakke en discutabele aspecten van het Cox-Snell-model; sommige daarvan zijn al eerder aangestipt.

(i) De ten onrechte opgevoerde fractie $\underline{S}$ van een incorrecte post heeft in dit model de exponentiële verdeling en dus een positieve dichtheid voor alle $S > 0$. In werkelijkheid geldt natuurlijk $\underline{S} < 1$.

(ii) Omdat $1/\lambda$ te interpreteren is als de gemiddelde relatieve fout in foutieve posten, moet gelden $\underline{\lambda} > 1$. De gekozen gammaprior geeft echter $0 < \underline{\lambda} < 1$ een positieve kans.

(iii) Iets dergelijks treedt ook op bij $\underline{\phi}$: een positieve dichtheid voor alle $\phi > 0$, terwijl in feite geldt $\underline{\phi} < 1$.

(iv) Zowel voor $\underline{\phi}$ als voor $\underline{\lambda}$ worden uitsluitend gammaverdelingen gebruikt; het is mogelijk dat hiermee de subjectieve voorkennis van de onderzoeker in onvoldoende mate beschreven kan worden.

(v) Met name lijkt het plausibel dat vaak posten in hun geheel niet in de populatie thuisshoren. Dit komt neer op optreden van een puntmassa bij $\underline{S} = 1$.

(vi) Strikt genomen heeft het aantal incorrecte posten in de steekproef een binomiale verdeling, welke echter benaderd is m.b.v. een Poissonverdeling. Voor wat grotere waarden van ϕ is deze benadering onnauwkeurig.

(vii) Aangenomen wordt dat $\underline{\phi}$ en $\underline{\lambda}$ onderling onafhankelijk zijn, hetgeen betekent dat de relatieve fout in een post onafhankelijk is van het nominale bedrag. Dit behoeft in de praktijk niet zo te zijn.

Om aan deze bezwaren tegemoet te komen bedachten GODFREY & NETER (1983, 1984) een aantal modificaties van het oorspronkelijke Cox-Snell-model. In al hun alternatieve modellen is bezwaar (i) ondervangen door aan te nemen dat $\underline{S}$ een tot $[0,1]$ beperkte exponentiële verdeling heeft. (Bij een willekeurige kansdichtheid f is

$$(5.1) \quad g(x) := f(x) / \int_a^b f(x) dx, \quad a \leq x \leq b$$

de bijbehorende tot $[a,b]$ beperkte dichtheid.) Verder wordt bij alle alternatieven aangenomen dat - afgezien van eventuele puntmassa's - de priorverdeling van $1/\lambda$ de uniforme verdeling op $[0,1]$ is; notatie $U(0,1)$. Dit betekent dat λ zelf de Paretoverdeling $Pa(1,1)$ heeft met dichtheid

$$f(\lambda) = 1/\lambda^2, \quad \lambda > 1$$

Deze modificatie neemt bezwaar (ii) weg en komt tegemoet aan (iv).

De gemodificeerde modellen onderscheiden zich verder op de volgende punten. Voor de overzichtelijkheid wordt het originele Cox-Snell-model aangeduid met GG (naar de twee gamma-verdelingen die worden gehanteerd). Bij de hierna volgende vergelijking worden twee sets priorwaarden voor de parameters gehanteerd, corresponderend met een lage resp. hoge a priori verwachting van $\underline{\psi}$; de twee submodellen worden aangeduid met GGL resp. GGH.

De eerste variant van GG ontstaat door als priorverdeling van ϕ de tot $[0,1]$ beperkte gammaverdeling te nemen, waarmee bezwaar (iii) wordt weggenomen. Naar de twee gebruikte priorverdelingen wordt deze variant aangeduid met GU. Twee verschillende sets parameterwaarden leiden tot de subvarianten GUL en GUH.

De variant GU1 ontstaat uit GU door aan te nemen dat $\underline{S} = 1$ met kans 0.2ϕ optreedt; bezwaar (v) verdwijnt hierdoor.

In plaats van (2.6) wordt dus gebruikt

$$(5.2) \quad \begin{cases} P(\underline{S}=0) = 1-\phi \\ P(s < \underline{S} < s+ds) = 0.8 \phi \lambda \exp(-\lambda s) ds, \quad s > 0 \\ P(\underline{S}=1) = 0.2 \phi \end{cases}$$

Bij de variant BU wordt de exacte binomiale verdeling gebruikt voor het aantal incorrecte posten $\underline{M}$ - vergelijk bezwaar (vi) - en wordt een bètaverdeling aangenomen voor ϕ . Een lage resp. hoge a priori verwachting van $\underline{\psi}$ leidt tot de modellen BUL en BUH. Omdat de bètaverdeling de dichtheid 0 heeft buiten het een-

heidsinterval, wordt ook hier tegemoet gekomen aan bezwaar (iii). Door bovendien de verdeling (5.2) te gebruiken i.p.v. (2.6) ontstaat model BUL.

In totaal zijn nu acht verschillende modellen gedefinieerd, waarvan Tabel 5.1. een gedetailleerd overzicht geeft. De vier modellen met een lage priorverwachting voor ψ zijn eerst vermeld; merk op dat deze verwachting, net als het 95^e percentiel $k_{0.95}$, ruwweg gelijk is voor alle vier de modellen. Dit laatste geldt ook voor de vier modellen met hoge priorverwachting van ψ . Na te gaan valt dat voor alle modellen geldt $E(1/\lambda) = \frac{1}{2}$ en $V(1/\lambda) = 1/12$. Voor deze acht modellen werden bij een zevental uiteenlopende steekproefresultaten via numerieke integratie de posteriorverdelingen bepaald. Steeds werd uitgegaan van een overigens arbitraire steekproefomvang $n = 100$ uit de populatie beschreven in Par. 2. Tabel 5.2 geeft een overzicht van de hieruit berekende 95%-betrouwbaarheidsgrenzen van ψ ; m^* geeft het aantal aangetroffen posten aan met $S = 1$. Merk op dat de bovengrenzen in Tabel 5.2 niet vergeleken kunnen worden met de waarde van ψ in de populatie, omdat deze immers onbekend is. Uitsluitend onderlinge vergelijking van de bovengrenzen is mogelijk, waardoor de invloed van de diverse modificaties kan worden nagegaan

Tabel 5.1. Overzicht van de beschouwde modificaties van het Cox-Snell-model

Model	Verdeling $\underline{S}$		Verd. $\underline{M}$	Prior $\underline{\lambda}$	Priorverdeling $\underline{\phi}$			Prior $\underline{\psi}$	
	Dichth.	Puntm.			Type	$E(\underline{\phi})$	$\sigma_{\underline{\phi}}$	$E(\underline{\psi})$	$k_{0.95}$
GGL	exp.	0	Poisson	$\Gamma(2,5)$	$\Gamma(200,1)$	.005	.005	.0025	.0082
GUL	bep. exp.	0	Poisson	$Pa(1,1)$	bep. gamma	.005	.005	.0025	.0094
BUL	bep. exp.	0	Bin.	$Pa(1,1)$	$Be(4,791)$	.005	.0025	.0025	.0071
BUL	bep. exp.	0+1	Bin.	$Pa(1,1)$	$Be(4,791)$	.005	.0025	.0030	.0070
GGH	exp.	0	Poisson	$\Gamma(2,5)$	$\Gamma(100,25)$	.25	.05	.125	.2640
GUH	bep. exp.	0	Poisson	$Pa(1,1)$	bep. gamma	.25	.05	.125	.2967
GUL	bep. exp.	0+1	Poisson	$Pa(1,1)$	bep. gamma	.25	.05	.15	.2688
BUH	bep. exp.	0	Bin.	$Pa(1,1)$	$Be(5,15)$	.25	.0945	.125	.2975

Tabel 5.2 A posteriori 95%-betrouwbaarheidsbovengrenzen voor de verschillen de modellen

Steekproef			Model							
m	z	m*	GGL	GUL	BUL	BU1	GGH	GUH	GU1	BUH
0	0	0	0.55	0.64	0.56	0.64	13.2	13.0	13.4	5.2
1	.01	0	0.73	0.71	0.55	0.58	10.4	11.5	12.2	4.8
1	.5	0	0.91	1.13	0.77	0.78	13.0	14.1	14.7	6.5
1	1	1	1.09	1.21	0.82	0.73	15.6	14.9	14.0	6.9
10	3	1	2.33	4.00	1.67	1.46	10.1	17.3	15.5	12.6
10	5	0	3.26	4.77	1.96	1.99	14.1	19.6	20.0	15.2
10	7	3	4.20	5.03	2.06	1.99	18.1	20.6	21.4	16.0

Bron: GODFREY & NETER (1984).

Uit deze tabel kunnen een aantal voorlopige conclusies worden getrokken.

(i) Invloed priorwaarden parameters

Het meest opvallende aan Tabel 5.2 is dat de laatste vier modellen tot veel hogere bovengrenzen leiden dan de andere. Dit benadrukt hoe belangrijk het is of ψ tevoren hoog dan wel laag wordt geschat. Overigens is het verschil tussen de twee waarden van $E(\psi)$ in Tabel 5.1 ook wel erg groot.

(ii) Invloed steekproefuitkomsten

Als regel geldt dat meer (en grotere) fouten in de steekproef leiden tot een hogere bovengrens. Merkwaardig genoeg blijkt dit niet altijd te gelden: soms leidt het aantreffen van één foutje tot een lagere bovengrens dan een steekproef met uitsluitend correcte posten.

(iii) Invloed type priorverdeling

Voor lage priorverwachting van ψ is beperking van de priorverdeling niet van belang. Vergelijking van GGL en GUL toont dus het effect van de vervanging van de gammaverdeling door de Paretoverdeling. Alleen voor hogere aantallen fouten is het verschil enigszins substantieel. Bij de vergelijking van GGH en GUH kan ook de beperking van de kansdichtheden een rol spelen; het beeld is analoog. Bijna steeds leidt GU tot hogere bovengrenzen dan GG.

De consequenties van vervanging van de gamma- door de bètaverdeling blijken door de modellen GU en BU te vergelijken. Steeds leidt gebruik van de gammaverdeling tot hogere grenzen; de verschillen zijn aanzienlijk als de steek-

proefresultaten slecht overeenstemmen met de priorverdeling (linksonder en rechtsboven in Tabel 5.2).

(iv) Invloed puntmassa in $S = 1$

Het als dan niet met positieve kans voorkomen van posten die in hun geheel niet thuishoren in de populatie blijkt steeds van zeer ondergeschikt belang te zijn.

Als voorlopige eindconclusies stelden GODFREY & NETER (1985) dat het Cox-Snell-model geen duidelijke tekortkomingen vertoont in vergelijking met de beschouwde modificaties. I.v.m. de analytische gemakken zal daarom de aandacht op dit model geconcentreerd blijven.

6. Robuuste modellen

In het Cox-Snell-model weerspiegelen de gekozen parameterwaarden de a priori mening van de onderzoekers over het voorkomen van fouten in de onderzochte populatie. Deze mening zal vrijwel nooit volledig correct zijn. Daarom zou het plezierig zijn als de Cox-Snell-bovengrens bruikbaar zou blijven, ook als er verschillen bestaan tussen de priorwaarden en de werkelijke parameterwaarden. De bovengrens wordt in dat geval robuust genoemd. In het algemeen noemt men een methode die ontworpen is voor toepassing onder bepaalde modelveronderstellingen robuust, als de methode ook bruikbaar is bij niet te grote afwijkingen van dit model. In zekere zin vormt zo'n breed toepasbare priorverdeling een compromis tussen de Bayesiaanse en de klassieke opvatting.

Teneinde een indruk te krijgen van de robuustheid van Cox-Snell-bovengrenzen, bekeken GODFREY & NETER (1984) en NETER & GODFREY (1985) het gedrag van een aantal van deze grenzen bij toepassing op uiteenlopende (kunstmatig gecreëerde) populaties. Zij deden dit door middel van simulatie. Daarnaast werden de simulatieresultaten gebruikt voor vergelijking met een tweetal klassieke bovengrenzen. Enkele typerende resultaten worden hieronder besproken.

Allereerst geeft Tabel 6.1 een overzicht van zes van de door Godfrey en Neter gebruikte modellen. De meest interessante zijn geselecteerd: de overige vertonen een lagere waarden van $E(\psi)$ en zijn daardoor minder algemeen toepasbaar. Let er op dat CS10 het model is dat in vorige paragrafen vaak als voorbeeld gehanteerd is.

Voorts bekeken Godfrey en Neter 21 fictieve populaties, waarbij in alle gevallen S in wezen een gammaverdeling vertoont beperkt tot $[0,1]$. Deze 21 populaties zijn echter niet alle verschillend: zes populaties blijken paarsgewijs identiek te zijn! In de populaties 4 t/m 6 heeft S voor incorrecte posten de exponentiële

verdeling met parameter 5, korter genoteerd: $S | (S > 0) \sim \text{Ne}(5)$, en in de populaties 10 τ/m 12 geldt $10 S | (S > 0) \sim \chi^2_2$. Wegens de algemene eigenschappen

Tabel 6.1 Diverse Cox-Snell-modellen

Model	Prior $\underline{\phi}$	$E(\underline{\phi})$	$\sigma_{\underline{\phi}}$	Prior $\underline{\lambda}$	$E(1/\underline{\lambda})$	$E(\underline{\psi})$	$\sigma_{\underline{\psi}}$	$k_{0.95}$
CS 5	$\Gamma(40,4)$	0.1	0.05	$\Gamma(3.4,18)$	0.2	0.02	0.0115	0.042
CS 9	$\Gamma(10,1)$	0.1	0.1	$\Gamma(0.4,3)$	0.2	0.02	0.0346	0.069
CS 10	$\Gamma(10,1)$	0.1	0.1	$\Gamma(2,6)$	0.4	0.04	0.0490	0.130
CS 11	$\Gamma(15/4,9/16)$	0.15	0.2	$\Gamma(15/32,41/16)$	0.3	0.045	0.1166	0.176
CS 20	$\Gamma(20,4)$	0.2	0.1	$\Gamma(17/320,33/16)$	0.05	0.01	0.0450	0.030
CS 23	$\Gamma(5,1)$	0.2	0.2	$\Gamma(5/9,34/9)$	0.2	0.04	0.0583	0.135

$$\chi^2_2 = \Gamma(\frac{1}{2}, 1) = \text{Ne}(\frac{1}{2})$$

$$\underline{x} \sim \text{Ne}(\lambda) \rightarrow \alpha \underline{x} \sim \text{Ne}(\lambda/\alpha), \alpha > 0$$

zijn de twee genoemde verdelingen identiek. Daardoor zijn de populaties 4 τ/m 6 dus identiek met resp. 10 τ/m 12. Inderdaad stemmen de simulatieresultaten voor deze paren populaties doorgaans zeer goed overeen; alleen tussen de populaties 4 en 10 in Tabel 6 van NETER & GODFREY (1985) treden forse verschillen op. Tabel 6.2 geeft een overzicht van de 21 beschouwde populaties exclusief deze dubblures.

Uit elk van deze populaties werden steekproeven van 100 posten getrokken; uit elke steekproef werden de Cox-Snell-bovengrenzen berekend voor de onderscheiden modellen, alsmede de Stringer-bovengrens (4.1). Dit alles werd 500 maal herhaald, anders gezegd het simulatie-experiment telde 500 replicaties. Als samenvattende maatstaven berekenden Godfrey en Neter voor alle combinaties van een model en een populatie de dekking en de zuinigheid.

De dekking (coverage) van een bepaald model bij een bepaalde populatie is gedefinieerd als het percentage van de replicaties waarbij de 95%-betrouwbaarheidsbovengrens voor ψ hoger is dan de populatiewaarde van ψ . Als de dekking van een model ten minste 95 is voor (bijna) alle populaties, is er sprake van een robuust model.

Hoewel een 95%-betrouwbaarheidsbovengrens voor ψ in 95% van de gevallen ψ dient te overschrijden, dient aan de andere kant deze overschrijding zo gering mogelijk te zijn. Is dat voor vele populaties het geval, dan noemt men de bovengrens krap (tight). Als maatstaf hiervoor maakten Godfrey en Neter de ver-

gelijking met de Stringer-bovengrens. De zuinigheid (rel-tightness) van een bepaald model bij een bepaalde populatie is gedefinieerd als de mediaan van de

Tabel 6.2 Diverse (fictieve) populaties met kansverdeling $\underline{S}$

No. pop.	Puntmassa ($1-\phi$) in 0	Puntmassa in 1	Beperkte dichtheid	ψ (in %)
1	0.9	-	Ne(20)	0.05
2	0.9	-	Ne(20)	0.5
3	0.8	-	Ne(20)	1
4	0.99	-	Ne(5)	0.19
5	0.9	-	Ne(5)	1.93
6	0.8	-	Ne(5)	3.86
7	0.99	-	Ne(1.23)	0.40
8	0.9	-	Ne(1.23)	4.00
9	0.8	-	Ne(1.23)	8.00
13	0.99	-	$\Gamma(\frac{1}{2}, 2)$	0.36
14	0.9	-	$\Gamma(\frac{1}{2}, 2)$	3.65
15	0.8	-	$\Gamma(\frac{1}{2}, 2)$	7.30
16	0.99	0.001	Ne(1.23)	0.27
17	0.9	0.01	Ne(1.23)	2.74
18	0.8	0.02	Ne(1.23)	5.48
19	0.99	0.001	$\Gamma(\frac{1}{2}, 2)$	0.43
20	0.9	0.01	$\Gamma(\frac{1}{2}, 2)$	4.28
21	0.8	0.02	$\Gamma(\frac{1}{2}, 2)$	8.57

Stringer-grens (over de replicaties), uitgedrukt als percentage van de mediaan van de Cox-Snell-bovengrens.

Tabel 6.3 geeft een overzicht van de resultaten; als Godfrey en Neter meer dan één waarde vermelden voor een combinatie model/populatie zijn deze waarden gemiddeld. Gemakshalve zijn de zes modellen gerangschikt naar oplopende $E(\underline{\psi})$ en de 18 populaties naar oplopende ψ .

Voor de eerste acht populaties (met $\psi < 1\%$) is de dekking 100 voor alle modellen, behalve CS 20 (met $E(\underline{\psi})=1\%$); alle modellen zijn veel tot zeer veel krapper dan de Stringer-bovengrens.

Voor de zes daaropvolgende populaties (met $1.9 < \psi < 4.3$) schommelt de dekking praktisch tussen 90 en 100 voor alle modellen behalve CS 5. Het best voor de dag komt CS 10 (met $E(\underline{\psi})=4\%$), waarbij slechts voor populatie 20 de dekking beneden

Tabel 6.3 Dekking en zuinigheid van diverse Cox-Snell-modellen

No. pop.	Dekking						Zuinigheid					
	CS20	CS5	CS9	CS10	CS23	CS11	CS20	CS5	CS9	CS10	CS23	CS11
1	100	100	100	100	100	100	605	226	415	186	371	342
4	100	100	100	100	100	100	550	229	378	187	360	328
16	100	100	100	100	100	100	449	229	374	186	345	312
13	100	100	100	100	100	100	270	239	302	188	327	276
7	98	100	100	100	100	100	264	242	287	191	319	281
19	100	100	100	100	100	100	223	240	295	188	299	261
2	97	100	100	100	100	100	372	162	278	128	248	247
3	92	100	100	100	100	100	289	146	232	124	210	211
5	93	99	94	100	95	96	157	186	170	130	162	156
17	91	84	87	100	91	92	144	200	153	130	156	140
14	97	67	93	98	93	94	120	199	138	126	139	127
6	91	89	92	100	94	95	145	171	147	122	138	134
8	96	62	90	96	90	94	114	201	135	125	130	122
20	98	53	90	94	93	94	110	207	135	125	128	120
18	90	63	88	96	90	92	131	180	136	123	131	125
15	97	53	95	97	95	96	122	179	123	120	124	120
9	95	43	91	96	93	96	119	180	123	120	122	116
21	96	40	92	93	94	94	116	182	122	119	119	114

95 ligt. Opvallend is de relatief goede dekking van CS 20 (met $E(\psi)=1\%$), de zuinigheid ligt steeds boven 110 en meestal boven 130.

Dit beeld wordt versterkt bij de laatste vier populaties (met $5.5 < \psi < 8.6$); de beste blijft CS10 (met $E(\psi)=4\%$) met een vrijwel constante zuinigheid van 120. Voornaamste concurrenten zijn CS 11, CS 20 en CS 23. Echt verrassend is dat niet: het zijn immers de priorverdelingen uit Tabel 6.1 met de laagste σ_ψ , m.a.w. de minst specifieke.

De voorlopige conclusie is dat er blijkbaar Cox-Snell-modellen te construeren zijn die behoorlijk robuust zijn en krasser dan de Stringer-bovengrens. Hierdoor aangemoedigd betrokken NETER & GODFREY (1985) negen nieuwe populaties bij hun onderzoek, die nog (veel) meer fouten vertoonden. Het betrof opnieuw tot $[0,1]$ beperkte gammaverdelingen, al dan niet met een puntmassa 0.1ϕ in 1. Tabel 6.4 geeft een overzicht van de bevindingen voor de meest robuuste modellen.

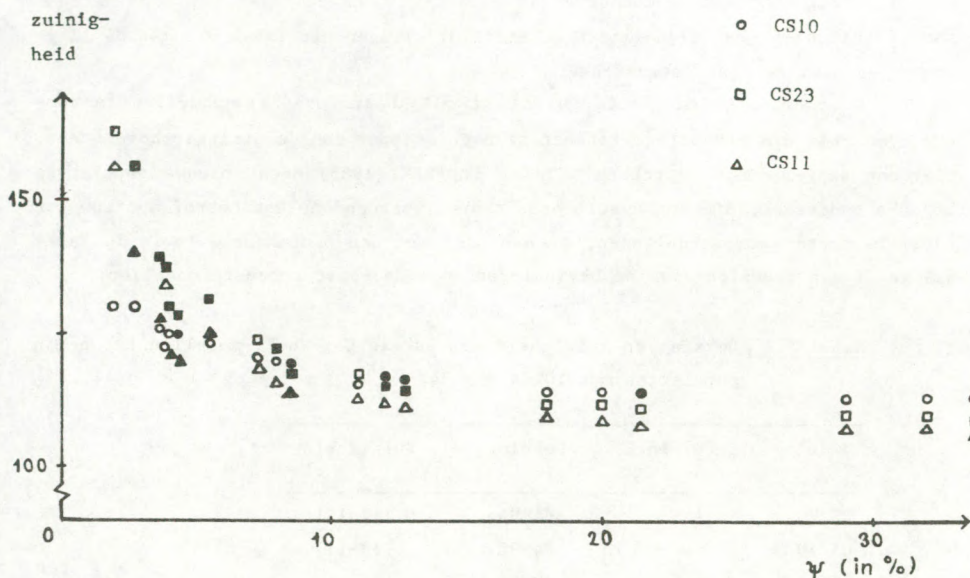
Tabel 6.4 Dekking en zuinigheid van enkele Cox-Snell-modellen bij negen populaties met $10\% < \psi < 34\%$

model	$E(\psi)$ in %	Dekking	Zuinigheid
CS 20	1	84-96	117-121
CS 10	4	94-98	113-117
CS 23	4	93-98	109-117
CS 11	4.5	96-99	107-113

Zelfs bij ψ -waarden boven 30% blijken de modellen CS 10, CS 23 en CS 11 nog zeer goed te voldoen. Figuur 6.1 laat dit nog eens grafisch zien. Hierin is voor deze drie modellen de zuinigheid afgezet tegen de ψ -waarden van de populatie. Een zwart gemaakt figuurtje geeft aan dat de dekking beneden de 95 ligt.

7. Analytische studie van het model

De simulatieresultaten die in de vorige paragraaf zijn beschreven berusten in totaal op het gigantische aantal van 1.5 mln. steekproeftrekkingen. De vraag klemmt daarom of simulatieonderzoek wel nodig is in de onderhavige situatie. In vele opzichten blijkt het antwoord ontkennend te zijn: dankzij hun analytische eigenschappen kunnen de Cox-Snell-bovengrenzen ook theoretisch worden bestudeerd. Uiteraard bespaart dit veel mens- en computerkracht en bovendien leidt deze aanpak tot exacte resultaten. De te volgen methode is al aangegeven in MOORS (1983) en wordt hier opnieuw uiteengezet aan de hand van een concreet voorbeeld.



Figuur 6.1. Zuinigheid (en dekking) van drie Cox-Snell-modellen voor 19 populaties.

De a posteriori bovengrens voor ψ wordt gegeven door het rechterlid van (3.6). Bij gegeven parameterwaarden van de priorverdeling hangt deze grens alleen af van de verkregen steekproefresultaten m en z . Voordat de waarnemingen verricht zijn is deze bovengrens dus op te vatten als een functie van de kansvariabelen $\underline{M}$ en $\underline{Z}$. Bij het eerder gebruikte model CS10 (zie ook Figuur 3.1 en Tabel 3.1) en steekproefomvang $n = 100$ ontstaat zo de stochastische Cox-Snell-bovengrens $\underline{L}_p$, gedefinieerd als

$$(7.1) \quad \underline{L}_p = \frac{(\underline{M}+1)(\underline{z}+2)}{110(\underline{M}+6)} F_{2\underline{M}+2; 0.05}$$

Dit is formule (3.2) uit MOORS (1983).

Onderzocht wordt nu het gedrag van $\underline{L}_p$ bij toepassing op populatie 4(=10) van Godfrey en Neter; hierbij heeft $\underline{S}$ voor incorrecte posten de tot $[0,1]$ beperkte verdeling $Ne(5)$, terwijl $\phi = 0.01$. De beperking tot $[0,1]$ maakt numeriek nauwelijks verschil en leidt theoretisch tot flinke (maar wel oplosbare) complicaties. Daarom wordt hier uitgegaan van de vereenvoudigende aanname $\underline{S} | \{\underline{S} > 0\} \sim Ne(5)$. Voor $m > 0$ volgt hieruit $\underline{Z} | (M=m) \sim \Gamma(5, m)$. Voor gegeven $M = m$ is $\underline{L}_p$ een lineaire functie van $\underline{Z}$, zodat de voorwaardelijke kansverdeling van $\underline{L}_p$ direkt volgt.

Heel eenvoudig kunnen de verwachting en de variantie van $\underline{L}_p$ worden berekend. Tabel 7.1 geeft enkele tussenresultaten die hiervoor nodig zijn. Er is gebruik gemaakt van het feit dat voor $\underline{Z} \sim \Gamma(5, m)$ geldt $E(\underline{Z}) = m/5$ en $V(\underline{Z}) = m/25$.

Tabel 7.1 Eigenschappen van $\underline{L}_p$ voor populatie 4

m	$\underline{L}_p$	$E(\underline{L}_p M=m)$	$V(\underline{L}_p M=m)$	$P(M=m)$
0	1.176	1.176	0	0.3679
1	0.808 ($\underline{Z}+2$)	1.778	0.026	0.3679
2	0.934 ($\underline{Z}+2$)	2.242	0.070	0.1839
3	1.014 ($\underline{Z}+2$)	2.634	0.123	0.0613
4	1.068 ($\underline{Z}+2$)	2.990	0.182	0.0153
5	1.106 ($\underline{Z}+2$)	3.318	0.245	0.0031
6	1.130 ($\underline{Z}+2$)	3.616	0.306	0.0006

Met de bekende eigenschap (van der GENUGTEN (1986), par. 4.3.5)

$$(7.2) \quad V(\underline{L}_p) = V[E(\underline{L}_p | M)] + E[V(\underline{L}_p | M)]$$

vindt men hieruit $E(\underline{L}_p) = 1.72$ en $V(\underline{L}_p) = 0.282$.

Tabel 7.2 toont enkele verdelingsfuncties behorend bij $\Gamma(5, m)$; het eerste deel van Tabel 7.3 geeft de hieruit berekende voorwaardelijke verdelingsfuncties van $\underline{L}_p$, d.w.z. de kansen $P(\underline{L}_p < \ell | M = m)$.

Tabel 7.2 Verdelingsfuncties van $\Gamma(5, m)$

m	Verdelingsfunctie
1	$1 - e^{-5z}$
2	$1 - (5z + 1) e^{-5z}$
3	$1 - (25z^2/2 + 5z + 1) e^{-5z}$
4	$1 - (125z^3/6 + 25z^2/2 + 5z + 1) e^{-5z}$

Tabel 7.3 Voorwaardelijke verdelingsfuncties van $\underline{L}_p$ voor populatie 4

ℓ (in %)	Voorwaardelijk				Onvoorwaardelijk
	$m = 1$	$m = 2$	$m = 3$	$m = 4$	
1.6	0				0.368*
1.8	0.680	0			0.618
2.0	0.907	0.158	0		0.765
2.2	1	0.530	0.055	0	0.837
2.4		0.777	0.279	0.037	0.896
2.8		0.959	0.732	0.377	0.963
3.2		0.993	0.927	0.732	0.987
3.6		1	0.983	0.910	0.996
4.0			1	0.974	0.999
4.4				1	1

* Puntmassa in 1.176.

De laatste kolom van Tabel 7.3 toont de onvoorwaardelijke verdelingsfunctie van $\underline{L}_p$. Deze is berekend met de formule

$$(7.4) \quad P(\underline{L}_p < \ell) = \sum_{m=0}^{\infty} P(\underline{L}_p < \ell \mid \underline{M}=m) P(\underline{M}=m)$$

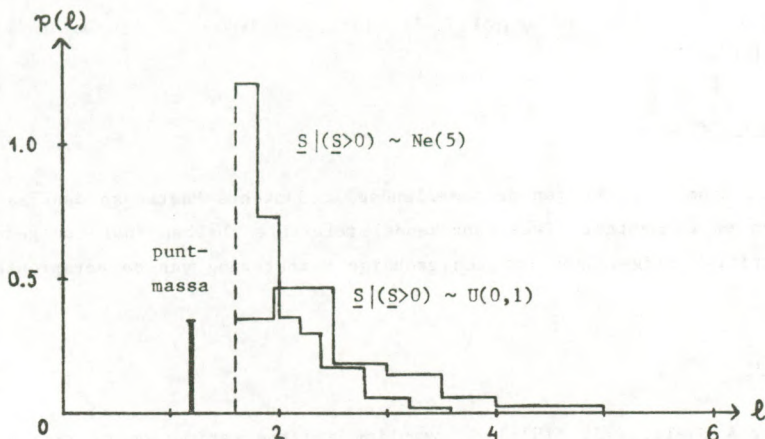
met $\underline{M} \sim \text{Po}(n\phi) = \text{Po}(1)$. De corresponderende dichtheid van $\underline{L}_p$ is weergegeven in Figuur 7.1. $\underline{L}_p > \psi = 0.2$ geldt met kans 1, terwijl de mediaan gelijk blijkt te zijn aan 1.69; beide resultaten stemmen goed overeen met de bevindingen in Tabel 9 van GODFREY & NETER (1984).

Figuur 7.1 toont eveneens de kansverdeling van $\underline{L}_p$ voor de populatie met $\phi = 0.01$ en $\underline{S} \mid (\underline{S} > 0) \sim U(0,1)$; de afleiding hiervan is te vinden in MOORS (1983).

De hier uiteengezette aanpak is nader uitgewerkt voor andere door Godfrey en Neter bestudeerde (gamma)priorverdelingen; de resultaten zijn te vinden in MOORS & JANSSENS (1987).

8. Discussie

Bij accountantscontroles bestaan tegen de klassieke betrouwbaarheidsbovengrenzen voor de totale relatieve fout ψ twee bezwaren:



Figuur 7.1 Kansverdeling L_p voor twee populaties

- de bovengrens kan niet bepaald worden als in de steekproef geen incorrecte posten worden aangetroffen.
- van de voorkennis van de accountant wordt geen expliciet gebruik gemaakt.

De Bayesiaanse aanpak komt aan deze beide bezwaren tegemoet. Van de in de literatuur voorgestelde modellen lijkt dat van Cox en Snell het geschiktst, omdat het zeer flexibel is en bovendien analytisch goed hanteerbaar. Deze analytische mogelijkheden zijn tot op heden overigens nog onvoldoende uitgebuit: in Par. 7 is uiteengezet hoe het onderzoek naar robuustheid ervan kan profiteren.

Het Cox-Snell-model gaat uit van gammaverdelingen als priors voor de grootheden ϕ en λ . Dit is echter geen wezenlijke inperking, zoals bleek uit de bestudering van alternatieve modellen in Par. 5 en uit het robuustheidsonderzoek in Par. 6. In deze laatste paragraaf kwam de bruikbaarheid naar voren van goed gekozen Cox-Snell-modellen voor populaties met sterk uiteenlopende waarden van ψ . Ook voor populaties met veel meer en veel grotere fouten dan vooraf werd verwacht bleek de 95%-bovengrens van het Cox-Snell-model ψ met ca. 95% kans te overschrijden, terwijl deze grens lager lag dan de klassieke Stringer-grens.

Verder onderzoek van het Cox-Snell-model zou de volgende richtingen op kunnen gaan.

(i) Uitbreiding van dit onderzoek m.b.t. anderssoortige (fictieve) populaties. In verband met limieteigenschappen komen met name normaal verdeelde populaties in aanmerking.

(ii) Toepassing van het model op echte praktijksituaties. Het is daarvoor nodig te beschikken over complete populaties, waarin ook alle fouten zijn

aangegeven. De analytische aanpak faalt hier: gebruik zal moeten worden gemaakt van simulatie.

Verantwoording

Ik ben de Commissie OKS van de Nederlandse Accountants Maatschap dankbaar voor hun steun en commentaar. Twee (onbekende) referenten hebben door hun gedetailleerde kritiek bijgedragen tot een grondige verbetering van de oorspronkelijke versie.

Literatuur

- COX, D.R. & SNELL, E.J. (1979) On sampling and the estimation of rare errors, *Biometrika* 66, 125-132.
- GARSTKA, S.J. (1977) Models for computing upper error limits in dollar-unit sampling, *Journal of Accounting Research* 15 (Autumn '77) 179-192.
- GENUGTEN, B.B. van der (1986), *Inleiding tot de waarschijnlijkheidsrekening en mathematische statistiek, deel I*, Stenfert Kroese, Leiden.
- GODFREY, J. & NETER, J. (1983) Bayesian bounds for monetary unit sampling in accounting and auditing. Report 83-001, Univ. of Georgia.
- (1984) idem, *Journal of Accounting Research* 22 (Autumn '84), 497-525.
- LESLIE, D.A., TEITLEBAUM, A.D. & ANDERSON, R.J. (1979), *Dollar unit sampling*, Copp, Clark & Pitman, Toronto.
- MENZEFRICKE, U. & SMIELIAUSKAS, W. (1984) A simulation study of the performance of parametric dollar unit sampling statistical procedures, *Journal of Accounting Research* 22 (Autumn '84), 588-604.
- MOORS, J.J.A. (1983) Bayes' estimation in sampling for auditing, *The Statistician* 32, 281-288.
- MOORS, J.J.A. & JANSSENS, M.J.B.T. (1987), Analytical properties of Bayesian Cox-Snell bounds in auditing, Reeks "Ter Discussie" 87.01, Katholieke Universiteit Brabant.

NETER, J. & GODFREY, J. (1984) Robust Bayesian bounds for monetary-unit sampling in auditing, Report 84-001, Univ. of Georgia.

- (1985) idem, Applied Statistics 34, 157-168.

SMITH, T.M.F. (1979) Statistical sampling in auditing: a statistician's viewpoint, The Statistician 28, 267-280.

VANECEK, M.T. (1978) Bayesian dollar unit sampling in auditing, Ph. D. Diss., Univ. of Texas.

Ontvangen: 11-03-1987
Geaccepteerd: 28-10-1987