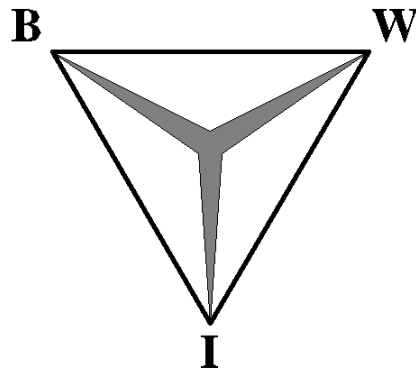


# Onderzoek Naar *Continuous Auditing Tools*

Jasper Roseboom  
Maart 2003



**Deloitte  
& Touche**



Stageverslag

Vrije Universiteit  
Faculteit der Exacte Wetenschappen  
Studierichting Bedrijfskunde en Informatica  
De Boelelaan 1081a  
1081 HV Amsterdam



<b>1</b>	<b>VOORWOORD</b> .....	<b>3</b>
<b>2</b>	<b>INLEIDING</b> .....	<b>5</b>
<b>3</b>	<b>DELOITTE &amp; TOUCHE</b> .....	<b>6</b>
<b>4</b>	<b>ALGEMEEN</b> .....	<b>9</b>
4.1	INLEIDING .....	9
4.2	CIJFERANALYSE .....	10
4.2.1	<i>Lineaire Regressie</i> .....	10
4.3	STEEKPROEVEN.....	10
4.3.1	<i>Transactieposten – Subposten</i> .....	10
4.3.2	<i>Overwaardering – Onderwaardering</i> .....	11
4.3.3	<i>Soorten steekproeven</i> .....	11
4.3.4	<i>Evaluatie steekproeven</i> .....	11
4.3.5	<i>Cellsampling</i> .....	12
<b>5</b>	<b>GELDSTEEKPROEVEN</b> .....	<b>13</b>
5.1	INLEIDING .....	13
5.2	STEEKPROEFGROOTTE.....	13
5.3	TOEVOEGINGEN .....	15
<b>6</b>	<b>POSTENSTEEKPROEVEN</b> .....	<b>17</b>
6.1	INLEIDING .....	17
6.2	KWANTITATIEVE POSTENSTEEKPROEVEN.....	17
6.2.1	<i>Steekproefgrootte</i> .....	17
6.3	EVALUATIE VAN KWANTITATIEVE STEEKPROEVEN .....	17
6.3.1	<i>Prestatiematen</i> .....	18
6.3.2	<i>Schatters</i> .....	19
6.3.3	<i>Betrouwbaarheidsinterval schatters</i> .....	20
6.4	EVALUATIE REGRESSIESCHATTER.....	21
6.4.1	<i>Excentriteit Interval</i> .....	21
6.5	STEEKPROEF VOORRADEN IN DE PRAKTIJK .....	22
<b>7</b>	<b>ONDERZOEKSVRAGEN</b> .....	<b>25</b>
<b>8</b>	<b>ONDERZOEKSVRAAG I</b> .....	<b>27</b>
8.1	INLEIDING .....	27
8.2	ALGEMEEN.....	27
8.3	‘OUDE’ TOEPASSING .....	28
8.4	CA-TOEPASSING .....	29
<b>9</b>	<b>ONDERZOEKSVRAAG II</b> .....	<b>31</b>
9.1	INLEIDING .....	31
9.2	ONDERZOEKSOPZET .....	31
9.2.1	<i>Regressieschatter-Plus</i> .....	32
9.2.2	<i>Vertalingsmechanisme</i> .....	32
9.2.3	<i>Simulatieprogramma</i> .....	34
9.2.4	<i>Foutverloop</i> .....	34
9.2.5	<i>Kenmerken van de voorraad</i> .....	35
9.3	UITVOERING EN RESULTATEN .....	35
9.3.1	<i>Basisinstellingen</i> .....	35
9.3.2	<i>Wijzigingen Instellingen</i> .....	37
9.3.3	<i>Samenvatting Resultaten</i> .....	45
9.3.4	<i>Kritische noot</i> .....	45
<b>10</b>	<b>CONCLUSIES</b> .....	<b>47</b>
<b>11</b>	<b>NOTATIELIJST</b> .....	<b>49</b>

<b>12</b>	<b>APPENDIX A</b> .....	<b>51</b>
12.1	ARM .....	51
12.2	BRAM.....	53
12.3	EVALUATIEMETHODEN GELDSTEEKPROEVEN .....	53
12.3.1	<i>Goed-of-Fout-methode</i> .....	54
12.3.2	<i>Stringerboud</i> .....	54
<b>13</b>	<b>APPENDIX B</b> .....	<b>57</b>
<b>14</b>	<b>APPENDIX C</b> .....	<b>59</b>
<b>15</b>	<b>APPENDIX C</b> .....	<b>67</b>
<b>16</b>	<b>LITERATUURLIJST</b> .....	<b>69</b>

# 1 Voorwoord

Voor u ligt een stageverslag waarin een onderzoek naar de bruikbaarheid van een tweetal technieken is behandeld. Deze onderzochte technieken waren gebaseerd op reeds bestaande (accountancy-) technieken. De nieuwe ontwikkelde technieken hadden als doel om een zogenaamde continue controle (continuous audit) mogelijk te maken over verschillende controleposten in een jaarrekening. Het door mij gedane onderzoek moest uitwijzen of deze nieuwe technieken inderdaad voor een continuous audit geschikt waren.

Mijn dank gaat naar Prof. Dr. Van der Vaart, drs. Paul van Batenburg (partner Deloitte & Touche) en de afdeling DQI van Deloitte & Touche voor het geven van de benodigde informatie. Verder wil ik mijn familie en vrienden bedanken voor de gegeven steun in de afgelopen periode.

Jasper Roseboom

Maart 2003



## 2 Inleiding

Het lopen van een stage bij een bedrijf is een onderdeel van de studie Bedrijfswiskunde & Informatica. De stage dient een aantal doeleinden. Naast het opdoen van ervaring op het gebied van het doen van een onderzoek, beoogt de stage ook om de student kennis te laten maken met het bedrijfsleven. Een student wordt als het ware een medewerker van het bedrijf.

De stage bij Deloitte & Touche duurde van oktober 2001 tot mei 2002. In deze periode heb ik onderzoek gedaan naar twee mogelijke toepassingen op het gebied van Continuous Auditing. De onderzoeksvragen betroffen de gegevensgerichte controles op de jaarrekening. Dit zijn controles die uitkomsten geven in kwantitatieve grootheden. Voor het berekenen van deze grootheden worden enkele 'tools' gebruikt. Deloitte wilde deze controle-tools, geschikt voor een eenmalige controle (jaarrekening), transformeren naar tools die geschikt zijn voor continu gebruik. Bedrijven hebben dan de mogelijkheid om continu inzicht te houden op de verschillende posten. Dit helpt onder andere de leiding bij het nemen van tussentijdse (betrouwbare) beslissingen.

In het verslag zal eerst een inleiding worden gegeven over de verschillende tools die Deloitte gebruikt om de jaarrekening te controleren. Op deze manier raakt men bekend met de termen en formules waarmee ik later in mijn onderzoeksvragen zal werken. Vervolgens zullen de onderzoeksvragen worden behandeld. Dit houdt onder ander de opzet, de resultaten en de gemaakte aannamen in.

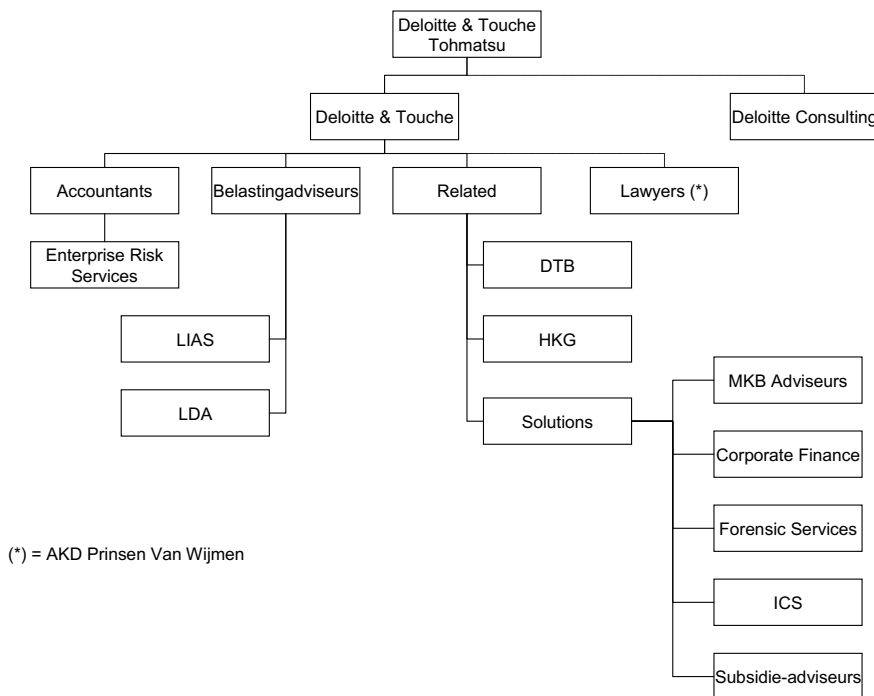
### 3 Deloitte & Touche

Deloitte Touche Tohmatsu (DTT) behoort tot de vier grootste accountantskantoren in de wereld. Het heeft ruim 95000 werknemers in dienst, verspreid over 140 landen. DTT biedt de volgende diensten aan:

- Accountancy
- Belastingadvies
- Consultancy
- Juridische dienstverlening.

De maatschap Deloitte & Touche in Nederland heeft 90 kantoren en zo'n 6000 medewerkers. De organisatie is marktleider op het gebied van overheid en non-profit-organisaties. Deloitte & Touche levert tevens diensten aan het Midden- en KleinBedrijf (MKB) via de regiokantoren.

De organisatiestructuur van DTT kan in het volgende organogram worden weergegeven:



In het organogram geven stippellijnen aan dat een bepaalde unit juridisch onafhankelijk is. Met Lawyers (AKD Prinsen Van Wijmen) is een strategische alliantie gesloten. Naar aanleiding van het Enron schandaal (Arthur Andersen) is Deloitte Consulting sinds kort een aparte unit geworden. Zo wil DTT de schijn van belangenverstrengeling voorkomen.

Mijn stage heeft op de afdeling Enterprise Risk Services (ERS) in Amsterdam (Sloterdijk) plaatsgevonden. De afdeling biedt naast ondersteuning aan het accountantsgedeelte van DT ook ondersteuning aan klanten bij het beheersen en controleren van allerlei soorten risico's. Deze risico's liggen onder andere op het gebied van strategie, organisatie en informatiemanagement. Het doel hiervan is om de belangrijkste bedrijfsprocessen te beschermen.

Het onderzoek werd begeleid door een specifieke groep binnen ERS, namelijk Data Quality en Integrity (DQI). DQI levert de volgende services:



- Audit services: het ondersteunen de accountants bij gegevensgerichte controles. Ook geeft DQI adviezen en cursussen met betrekking tot de gegevensgerichte controles en wordt er ook intern onderzoek en publicaties verricht.
- Client services: het ondersteunen van klanten bij het behouden en verbeteren van de kwaliteit van de data (Quality Improvement), het omzetten van gegevens in bruikbare informatie (File Interrogation) en datamining (het zoeken naar patronen in data).

Deze services hebben verscheidene toepassingsgebieden. Enkele hiervan zijn kredietonderzoek, voorspellingen, fraudedetectie en effectiviteitsonderzoek.



## 4 Algemeen

### 4.1 Inleiding

In de accountantscontrole worden controles gehouden die een uitspraak doen over de getrouwheid van de jaarrekening. Deze controles kunnen zowel systeem- als gegevensgericht zijn. Een systeemgerichte controle betreft een controle van het systeem waarin de gegevens verzameld en verwerkt worden. De gegevensgerichte controle doet een kwantitatieve uitspraak over deze gegevens. In dit verslag zullen de gegevensgerichte controles centraal staan.

Er zijn verschillende gegevensgerichte controletechnieken beschikbaar. Deze worden gebruikt om de verschillende onderdelen van de jaarrekening te controleren: de standgrootheden in de balansposten en de stroomgrootheden in de winst- en verliesrekening. Een post die pas gecontroleerd kan worden nadat alle onderdelen gecontroleerd zijn, is de post *Netto Winst*. Deze post komt in beide onderdelen van de jaarrekening voor:

#### *Winst- en verliesrekening*

<b>Post</b>	<b>Bedrag</b>
Opbrengsten	
Kosten	+ /-
Bruto Winst	=
Andere Kosten (o.a. Belasting)	+ /-
<i>Netto Winst</i>	=

#### *Balans*

<b>Activa</b>	<b>Passiva</b>
Kas	Lange Termijn
Voorraden	Korte Termijn
Debiteuren	Crediteuren
Gebouwen	Eigen Vermogen (o.a. <i>Netto Winst</i> )

De verschillende gegevensgerichte controletechnieken die gebruikt kunnen worden, zijn:

- Cijferanalyse: wordt bij voorkeur gebruikt om posten op onderwaarderingen te controleren. Een voorbeeld van zo een post is de winst- en verliesrekeningpost *Opbrengsten*.
- Steekproeven:
  - Geldsteekproeven: kan worden gebruikt om posten op overwaardering te controleren. De posten met betrekking tot kosten komen hiervoor in aanmerking.
  - Postensteekproeven: wordt gebruikt om posten te controleren die zowel een overwaardering als onderwaardering kunnen laten zien. Voorbeelden van zulke posten zijn *voorraden*, en in mindere mate *debiteuren* en *crediteuren*.

Deze controletechnieken zullen in de volgende paragrafen in het kort worden uitgelegd.

## 4.2 Cijferanalyse

Bij cijferanalyse maakt men gebruik van verbanden die de te controleren posten hebben met andere gegevens. Cijferanalyse wordt toegepast bij *Opbrengsten*. Aan de hand van de relevante kosten wordt er een voorspelling gemaakt van de betreffende opbrengsten.

Bij cijferanalyse kan het uitvoeren van *lineaire regressie* op de data een belangrijke hulpmiddel zijn om deze post te controleren. Het is een tool waarbij een model wordt opgezet waarin de afhankelijkheden tussen de data worden weergegeven. In zo een model vormt de opbrengst de afhankelijke variabele en vormen de kosten de onafhankelijke variabelen. De reden om op deze manier *Opbrengsten* te controleren heeft te maken met het feit dat de grootte van de *Opbrengsten* vaak moeilijk direct te achterhalen is, in tegenstelling tot de relevante kosten.

### 4.2.1 Lineaire Regressie

Met lineaire regressie wordt een zogenaamde kleinste kwadratenlijn gemaakt tussen de verschillende dimensies dataverzamelingen, de zogenaamde “line of best fit”. Deze lijn minimaliseert het totale kwadratische verschil tussen de datapunten en de lijn. Deze lijn is van de vorm:  $Y = aX + b$ .

De vergelijking probeert het werkelijke verband tussen de variabelen te benaderen. Aan de hand van verschillende statistische grootheden die men kan berekenen kan men uitspraken doen over het model en de afhankelijke variabele. De afdeling DQI maakt gebruik van het programma STAR (Statistical Techniques for Analytical Review) om deze grootheden te berekenen en zo een uitspraak te doen.

## 4.3 Steekproeven

Het doel van een steekproef is om een redelijke zekerheid te geven dat de jaarrekening geen materiële fouten bevat. Een 95% zekerheid wordt, indien deze te kwantificeren valt, vaak gezien als een redelijke zekerheid. Een 100% zekerheidsbeeld kan alleen verkregen worden als men alle data tot zijn beschikking heeft en als men bereid is om alles te controleren (integrale controle). Omdat deze punten een probleem vormen, wordt het toelaatbaar geacht om aan de hand van partiële waarnemingen een uitspraak te doen over de verschillende jaarrekeningposten. Door middel van een steekproef wordt dan een getrouwheidsbeeld verkregen. Met een bepaalde betrouwbaarheid wordt dan een uitspraak gedaan over de populatie.

### 4.3.1 Transactieposten – Subposten

In het verslag zullen in het geval van een stroomgrootte zoals de Winst-en Verliesrekeningpost *Kosten*, de *transactieposten* als steekproefelementen dienen. Een stroomgrootte heeft als kenmerk dat het wordt opgebouwd uit transacties door de tijd heen. De waarde van een post ‘stroomt’ als het ware. Hieronder een voorbeeld:

Tabel 1: voorbeeld van een stroomgrootheid, uitgedrukt in transacties.

Transactieno.	Datum	Waarde	Cumulatieve Waarde
1	01-04-'02	1000	1000
2	05-04-'02	100	1100
3	15-04-'02	500	1600

Voor een standgrootheid, ter sprake bij balansposten, zullen de elementen met *subposten* worden aangeduid. De subposten zullen vaak door de tijd heen in waarde veranderen maar hebben geen cumulatief (stromend) karakter zoals het geval was bij een stroomgrootheid. Bij de balanspost *Voorraden* zullen in feite de artikelnummers worden bedoeld wanneer de term 'subposten' wordt gebruikt.

### 4.3.2 Overwaardering – Onderwaardering

De waarde van een transactie- of een subpost zoals die in de boeken staat, de *boekwaarde*, hoeft niet overeen te komen met de werkelijke waarde, de *auditwaarde*. Wanneer de boekwaarde groter is dan de auditwaarde, is er sprake van *overwaardering*. Is het omgekeerde het geval dan spreekt men van *onderwaardering*. Het verschil tussen de boek- en de auditwaarde wordt het *residu* genoemd (residu = boekwaarde – auditwaarde).

### 4.3.3 Soorten steekproeven

In het algemeen zijn er twee soorten steekproeven te onderscheiden, namelijk:

- Postensteekproef. Elke subpost (of transactiepost) heeft een gelijke kans op selectie:  $\frac{1}{\# \text{subposten}}$
- Geldsteekproef. Elke geldeenheid heeft een gelijke kans op selectie:  $\frac{1}{\# \text{geldeenheden}}$

Beide technieken controleren uiteindelijk wel de sub- of transactieposten, maar bij de geldsteekproef bestaat de kans dat een sub- of transactiepost tweemaal gekozen kan worden. Een sub- of transactiepost bevat namelijk vaak meer dan één geldeenheid.

### 4.3.4 Evaluatie steekproeven

Bij de evaluatie van een steekproef, ongeacht de gebruikte variant, wordt de kwaliteit van een post in de volgende kwantitatieve termen uitgedrukt:

- Geschatte maximale fout (MF). De MF stelt een absoluut geldbedrag of een percentage fout voor die met een bepaalde betrouwbaarheid ontdekt zal worden. Deze MF wordt getoetst tegen de materialiteit. Onder materialiteit verstaan we de grootte van de fout in de accounting informatie die het waarschijnlijk maakt dat de beoordeling door de grootte van deze fout anders wordt. Is de MF kleiner dan de vooraf gespecificeerde materialiteit dan kan men concluderen dat de controlewerkzaamheden toereikend zijn geweest om de populatie goed te keuren. Is de MF echter groter dan deze vooraf gespecificeerde materialiteit dan is de steekproefcontrole nog niet toereikend en zullen aanvullende werkzaamheden moeten worden verricht.
- Bekende Fout (BF): dit is de totale fout die in de steekproef is gevonden. Deze fout wordt vaak gecorrigeerd in het teken van Quality Improvement.
- Waarschijnlijke Fout (WF): dit is de fout die men verwacht te vinden als men de rest van de populatie, populatie minus de steekproef, integraal zou controleren.
- Geprojecteerde fout (PF): deze fout is de som van de BF en WF. De PF is een goede indicator voor de kwaliteit van de populatie. Door de BF te extrapoleren naar de populatieomvang verkrijgt men een

puntschatting voor het meest waarschijnlijke foutbedrag.

De wijze waarop deze grootheden berekend worden hangt af van de steekproefmethode die men hanteert om de kwaliteit van de auditwaarden te bepalen.

#### 4.3.5 Cellsampling

Van de steekproefelementen wordt verondersteld dat ze een goede representatie zijn van de massa. Men wil namelijk een uitspraak doen over deze massa. Om er voor te zorgen dat de steekproefelementen met een zekere spreiding uit de massa worden getrokken, wordt in de accountancy *cellsampling* toegepast.

Cellsampling wordt vooral gebruikt bij stroomgrootheden. Men wil namelijk voorkomen dat transacties onevenredig vaak in de steekproef aanwezig zullen zijn. De situatie waarin bijvoorbeeld de helft van de steekproefelementen uit de maand januari komt terwijl het aantal transacties van januari niet meer dan 10% van het geheel uitmaakt, is ondenkbaar in de accountancy.

Volgens een aantal stappen wordt cellsampling uitgevoerd:

1. De massa wordt opgedeeld in gelijke delen. Deze delen worden 'cellen' genomen. De grootte van een cel wordt bepaald via de formule:  $cel\ grootte = \frac{massa}{steekproefgrootte}$ .
2. Uit iedere cel wordt een waarneming aselekt gekozen. Dit gebeurt onder voorbehoud dat de grootte van een cel een geheel getal is.
3. Sub- of transactieposten die groter zijn dan tweemaal de celgrootte worden automatisch meegenomen in de steekproef.

Cellsampling kan niet zomaar gebruikt worden. Het trekken van een steekproef element is onder cellsampling minder aselekt. Er wordt telkens maar één element aselekt gekozen uit een vaste cel. De rest van de steekproefelementen in de betreffende cel komt dan niet meer in aanmerking om tot de steekproef te behoren. Dit verhindert het gebruik van verdelingen die in de accountancy gebruikt worden. De verdelingen eisen namelijk aselekte trekking. Cellsampling zal dan alleen gebruikt worden wanneer er sprake is van een voldoende grote populatie. Het 'effect' van cellsampling wordt namelijk minder naarmate er meer elementen in de populatie zitten.

## 5 Geldsteekproeven

### 5.1 Inleiding

Bij het nemen van een geldsteekproef heeft elke geldeenheid een even grote kans om gekozen te worden. Daarom is deze steekproefmethode uitermate geschikt om posten op overwaardering te controleren. Dit betreft vaak kostenposten.

In het onderstaande voorbeeld, *Tabel 2*, kan men zien dat geldsteekproeven niet geschikt zijn om ondergewaardeerde transactieposten, bijvoorbeeld transactieno. 3, te controleren. De ondergewaardeerde transactieposten hebben namelijk minder kans om gekozen te worden dan bij een postensteekproef waar elke transactiepost een even grote selectiekans krijgt toegewezen.

*Tabel 2: voorbeeld geldsteekproef*

Transactieno.	Boekwaarde	Auditwaarde	Selectiekans
1	1200	1000	1200/2200 = 55%
2	1000	1000	1000/2200 ≈ 45%
3	0	1000	0/2200 ≈ 0%

### 5.2 Steekproefgrootte

De accountant neemt een steekproef ter grootte  $n$  uit een populatie. Met deze steekproefgrootte wil de accountant met een bepaalde betrouwbaarheid een uitspraak kunnen doen over de aanwezige fout in de populatie: is de aanwezige fout groter, kleiner of gelijk aan de opgestelde foutnorm?

De steekproefgrootte zal men in principe via de hypergeometrische verdeling moeten bepalen. De interpretatie van de hypergeometrische verdeling: uit een populatie ter grootte  $r$  (de totale boekwaarde van een post) doet men een steekproef ter grootte  $n$ . In de populatie bevinden zich  $m$  fouten. De kans dat  $k$  fouten worden gevonden in de steekproef is dan gelijk aan:

$$(1) \quad P(X = k) = \frac{\binom{m}{k} \binom{r-m}{n-k}}{\binom{r}{n}}$$

Er zijn echter drie opmerkingen te maken met betrekking tot (1) in relatie met de steekproef die de accountant doet. Ten eerste is het werkelijk aantal fouten in een populatie onbekend. De interpretatie voor  $m$  zal dan ook moeten wezen: het maximaal aantal fouten dat gevonden mag worden bij een integrale controle wil de accountant de populatie goedkeuren.

Ten tweede is de cumulatieve kans voor de accountant interessant:

$$(2) \quad P(X \leq k) = \sum_{l=0}^k \frac{\binom{m}{l} \binom{r-m}{n-l}}{\binom{r}{n}}$$

De kans dat hoogstens  $k$  foute geldeenheden worden aangetroffen in de steekproef ter grootte  $n$  uit populatie ter grootte  $r$ . Minder dan  $k$  fouten vinden in de steekproef is niet zo erg want dat is het gevolg van de onnauwkeurigheid van de steekproef. Ten derde zal de kans bekend moeten zijn waaronder (2) geldt. Deze kans zal in het vervolg aangegeven worden met  $x$  en stelt de betrouwbaarheid voor waarmee de accountant een uitspraak doet omtrent de werkelijke fout in de populatie.

Bij gegeven  $k$ ,  $x$ ,  $m$  en  $r$  kan men de steekproefgrootte  $n$  met behulp van (2) uitrekenen. Deze afleiding is niet zo simpel. Men verkrijgt namelijk een open vergelijking waaruit de waarde voor  $n$  vrij moeilijk te bepalen is.

Bij een voldoende grote  $\frac{r}{n}$  is de binomiale verdeling uit de hypergeometrische verdeling af te leiden. De

fractie fout  $\frac{m}{r}$  kan dan namelijk als een constante beschouwd worden. De fractie zal nauwelijks veranderen

wanneer trekking zonder teruglegging plaatsvindt uit de populatie. De fractie wordt in de binomiale verdeling met de  $p$  aangegeven. Omdat de  $m$  het maximaal aantal fouten voorstelde dat de accountant in de populatie mocht vinden, wordt de  $p$  de kritieke foutfractie genoemd. De kritieke foutfractie zal worden aangeduid met de  $p_m$ . De kans dat men hoogstens  $k$  fouten vindt onder  $p_m$  in een steekproef ter grootte  $n$  wordt dan:

$$(3) \quad P(X \leq k) = x \Rightarrow \sum_{l=0}^k \binom{n}{l} \cdot p_m^l \cdot (1 - p_m)^{n-l} = x$$

Voor  $k = 0$  is de steekproefgrootte makkelijk te bepalen onder gegeven  $p_m$  en  $x$ :

$$(4) \quad P(X = 0) = x \Rightarrow (1 - p_m)^n = x \Rightarrow n = \frac{\ln(x)}{\ln(1 - p_m)}$$

Wanneer voor  $k$  een waarde groter dan nul wordt gekozen, zal net als in het hypergeometrische geval sprake zijn van een open vergelijking waaruit een waarde voor  $n$  lastig af te leiden is.

Bij een grote populatie en een kleine (kritieke) foutfractie kan de Poisson formule gebruikt worden. Deze Poisson benadering van de binomiale verdeling leidt tot een simpeler afleiding voor de waarde  $n$  voor  $k$  groter dan nul. De cumulatieve kans wordt in het geval van de Poisson formule:

$$(5) \quad P(X \leq k) = x \Rightarrow \sum_{i=0}^k \frac{e^{-np} \cdot (np)^i}{i!} = x$$

Hoewel net als in het geval van de binomiale verdeling voor  $k > 0$  geen gesloten uitdrukking voor  $n$  gevonden kan worden, kunnen de parameters  $n$  en  $p_m$  bij de Poisson-formule uitgedrukt worden in één waarde, namelijk de  $n$  maal  $p_m$  ( $np_m$ ). Dit houdt in dat men maar één waarde hoeft te extrapoleren om een oplossing voor de vergelijking te vinden. Omdat de waarde  $np_m$  afhangt van de waarde van  $k$  en  $x$ , kan (5) samengevat worden in:

$$(6) \quad np_m = R(k, x) \Rightarrow n = \frac{R(k, x)}{p_m}$$

De vraag die nu nog rest is welke waarde de accountant voor de parameters  $k$ ,  $x$  en  $p_m$  zal moeten kiezen?

De  $x$  stelt de betrouwbaarheid voor waarmee de accountant een uitspraak wilt doen over de fout in de populatie. Over het algemeen wordt de  $x = 0.95$  genomen.

De  $p_m$  wordt bepaald door het absolute foutbedrag, de *MP*. De *MP* staat voor Monetary Precision, een uit de materialiteit afgeleid foutbedrag. Het absolute foutbedrag wordt gedefinieerd als het bedrag dat niet overschreden mag worden in de massa  $M$ . De  $p_m$  is gelijk aan:  $MP/M$ .

De waarde van  $k$  hangt naast de kritieke foutfractie  $p_m$  ook af van de foutfractie die men vooraf verwacht. Deze foutfractie wordt aangegeven met  $p_a$ . De waarde voor  $k$  wordt bepaald door de breuk ( $p_a / p_m$ ):



$$(7) \quad \frac{p_a}{p_m} = \frac{k \cdot M/n}{R(k, x) \cdot M/n} = \frac{k}{R(k, x)}$$

Bij een gegeven waarde voor  $x$ , zal deze breuk bij vermenigvuldiging van 100% voor elke waarde van  $k$  (groter of gelijk aan 0) de verhouding tussen de  $p_a$  en  $p_m$  in een percentage uitdrukken. In *Tabel 3* (volgende bladzijde) worden voor verschillende percentages de bijbehorende (gehele)  $k$ -waarde gegeven.

*Tabel 3: keuze  $k$  wordt bepaald door  $p_a$  en  $p_m$ .*

$(p_a/p_m) \cdot 100\%$	$K$
(vrijwel) 0	0
Tot 21%	1
21 tot 32%	2
32 tot 39%	3
39 tot 44%	4
44 tot 48%	5

Neemt men voor de  $p_a$  een kleine waarde dan is het beste om voor  $k$  de waarde nul te kiezen. Men zal hiermee de steekproefgrootte minimaliseren. Om een te ongunstig beeld uit te sluiten bij een grotere foutfractie, is het handig om voor  $k$  een groter getal te kiezen. De steekproef leidt dan niet gelijk tot afkeuring van de massa bij het vinden van een fout. De foutfractie zal niet te hoog worden ingeschat.

In onderstaande tabel, *Tabel 4*, worden de  $R(k, x)$  waarden gegeven onder  $x = 0.95$  en voor  $k = 0, 1, \dots, 5$ .

*Tabel 4: waarde  $R(k, 0.95)$  voor  $k = 0, 1, \dots, 5$ . De waarden voor  $R(k, 0.95)$  zijn verkregen door gebruik te maken van extrapolatie.*

$K$	$R(k, 0.95)$
0	3.00
1	4.75
2	6.30
3	7.76
4	9.16
5	10.52

Wanneer de kritieke foutfractie gesteld wordt op 0.01, dus als niet meer dan 1% van de totale massa fout mag zijn, en er mogen geen fouten gevonden worden in de steekproef,  $k = 0$ , dan zal de  $n$  gelijk worden aan 300.

### 5.3 Toevoegingen

De waarde van  $R(k, x)$  kan mede bepaald worden door de mate van steun op het AO/IC. Er zijn twee modellen beschikbaar die met a priori gegevens werken met betrekking tot het AO/IC. Deze modellen zijn: het ARM (Audit Risk Model) en BRAM (Bayesiaanse Risico Analyse Model). Doel van beide methoden is een zo laag mogelijke waarde voor  $R(k, x)$  te onderbouwen. Bij het bepalen van die waarde wordt daarom uitgegaan van een  $k$  gelijk aan 0.

Een geldeenheid in de geldsteekproef kan deel uitmaken van een foute transactiepost waarin de auditwaarde niet overeenkomt met de boekwaarde. In de vorige evaluatiemethode werd de geldeenheid dan gelijk als één hele foute geldeenheid beschouwd en werd de waarde voor  $k$  met één opgehoogd. Een foute transactie is echter vaak niet in zijn geheel fout. De boekwaarde bevat vaak een deel dat terecht geboekt is. Om de consequenties van zo een gedeeltelijke foute transactie niet te overschatten, zijn er evaluatietechnieken bedacht die dit probleem enigszins ondervangen. Twee van zulke technieken zijn:

- De Goed-of-Fout methode
- De Stringerbound

Beide methoden gaan in op de manier van verdelen van de geldeenheden (naar foute en goede geldeenheden) in de foute transactie wanneer er via een trekking een foute transactie aan het licht is gekomen. Deze methoden zijn staan gedetailleerd beschreven in *Appendix A*.

## 6 Postensteekproeven

### 6.1 Inleiding

Omdat het hanteren van de postensteekproef vooral wordt toegepast bij enkele standgrootheden wordt in dit hoofdstuk de term *subposten* in plaats van *transactieposten* gebruikt. Bij de balanspost *Voorraden* die centraal staat in dit hoofdstuk, zal elk artikelnummer een afzonderlijke subpost vormen.

De postensteekproef is een steekproefvorm waarbij subposten in plaats van geldeenheden aselekt worden gekozen. Subposten die ondergewaardeerd zijn, krijgen zo meer kans om opgenomen te worden in de steekproef dan in het geval van een geldsteekproef, zie *Tabel 5*.

*Tabel 5: Selectiekans bij postensteekproef.*

Subpostno.	Boekwaarde	Auditwaarde	Selectiekans
1	1200	1000	≈ 33.33%
2	1000	1000	≈ 33.33%
3	0	1000	≈ 33.33%

### 6.2 Kwantitatieve postensteekproeven

Bij de kwantitatieve postensteekproeven worden kwantitatieve uitspraken gedaan over posten waarin relatief veel fouten kunnen voorkomen en deze fouten betrekking hebben op zowel over- als onderwaarderingen. De balanspost *Voorraden* is hiervan een goed voorbeeld.

De kwantitatieve uitspraken betreffen de grootte van de fout (geprojecteerde fout), de toereikendheid van de steekproef (maximale fout), de getrouwheid van de administratie als weergave van de populatie, en de rechtmatigheid van de transacties die zijn verantwoord in de populatie. Bij getrouwheid worden de positieve en negatieve fouten met elkaar gesaldeerd. Bij rechtmatigheid worden de fouten absoluut opgeteld zodat men inzicht in de totale absolute afwijking verkrijgt.

#### 6.2.1 Steekproefgrootte

De steekproefgrootte wordt bij een kwantitatieve postensteekproef niet bepaald door een  $p_m$  en een  $R(k,x)$ . In de regel wordt een steekproefgrootte van 200 gebruikt bij toepassing van verscheidene schatters. Deze omvang bleek uit een onderzoek<sup>1</sup> van voldoende niveau om een kwantitatief oordeel te kunnen geven over de te controleren post.

### 6.3 Evaluatie van kwantitatieve steekproeven

Kwantitatieve methoden geven uitkomsten op een verhoudingsschaal. Voorbeelden van kwantitatieve postensteekproef methoden zijn:

- Directe schatter
- Verschilschatter
- Regressieschatter

Naast een schatting wordt een tweezijdig betrouwbaarheidsinterval gegeven waarbinnen de opgegeven totale boekwaarde verwacht wordt te zitten.

---

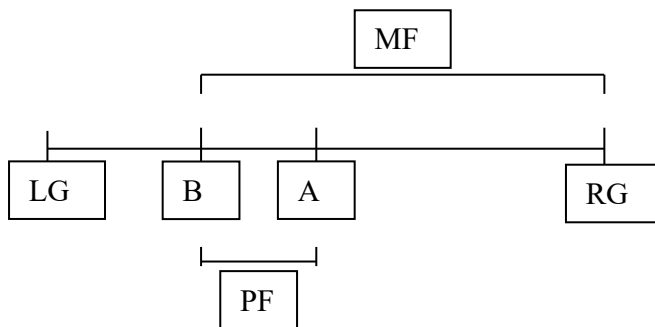
<sup>1</sup> Uit simulatiestudies van Kriens en Kleijnen kwam naar voren dat een aantal van 150-200 steekproefelementen voldoende is voor het schatten van het populatiegemiddelde.

### 6.3.1 Prestatiematen

De prestatie­maten waarnaar gekeken wordt bij de evaluatie zijn onder andere de geprojecteerde en maximale fout. De geprojecteerde fout kan men berekenen als de afstand tussen de puntschatting van de totale auditwaarde, en de totale boekwaarde. De maximale fout is de absolute afstand tussen de totale boekwaarde en de verst afgelegen intervalgrens in het geconstrueerde betrouwbaarheidsinterval.

Dit kan grafisch als volgt worden weergegeven:

Afkorting	Betekenis
LG	Linker Intervalgrens
RG	Rechter Intervalgrens
B	Totale Boekwaarde
A	Puntschatting Totale Auditwaarde
PF	Geprojecteerde Fout
MF	Maximale Fout



*Figuur 1: een schematische weergave voor het berekenen van de geprojecteerde en maximale fout. De B is willekeurig in het interval geplaatst.*

### 6.3.2 Schatters

*Directe schatters* zijn schatters waarbij alleen gebruik gemaakt wordt van de auditwaarden. Bij deze methode wordt eerst het gemiddelde en de variantie van de auditwaarden van de steekproef bepaald:

$$(13) \quad \bar{x}_{direct} = \frac{1}{n} \sum_{i=1}^n x_{i;audit}, \text{ en}$$

$$(14) \quad S_{direct}^2 = \frac{1}{n-1} \cdot \sum_{i=1}^n (x_{i;audit} - \bar{x}_{direct})^2$$

De variantie over het gemiddelde wordt dan:

$$(15) \quad S_{\bar{x}_{direct}}^2 = \left(1 - \frac{n}{N}\right) \cdot \frac{1}{n} \cdot S_{direct}^2$$

Aan de hand van deze statistische gegevens kan men het volgende betrouwbaarheidsinterval construeren voor het populatiegemiddelde:

$$(16) \quad \left[ \bar{x}_{direct} - t_{\alpha/2} \cdot S_{\bar{x}_{direct}} ; \bar{x}_{direct} + t_{\alpha/2} \cdot S_{\bar{x}_{direct}} \right]$$

Voor de schatting van de fout in de hele post wordt dit vermenigvuldigd met het totaal aantal subposten. De Student-verdeling<sup>2</sup> wordt gebruikt voor het maken van het betrouwbaarheidsinterval rondom het gemiddelde. Er worden  $n - 1$  vrijheidsgraden gebruikt. Het betrouwbaarheidsinterval bij de directe schatter is relatief groot. Dit geeft dus een relatief grote maximale fout.

Door gebruik te maken van een hulpvariabele kunnen de intervalgrenzen verkleind worden. De verschilschatter maakt gebruik van de *boekwaarde* als hulpvariabele. De grenzen zullen altijd kleiner worden wanneer er een correlatie  $> 0$  tussen de variabelen, de boek- en auditwaarde bestaat. Het populatiegemiddelde wordt dan als volgt geschat [Cochran 1966]:

$$(17) \quad \bar{x}_{verschil} = \bar{x}_{direct} + b(\bar{Y} - \bar{y}_{direct})$$

waarbij  $\bar{Y}$  de gemiddelde boekwaarde voorstelt van de hele populatie,  $\frac{M_{boek}}{N}$ , en  $\bar{y}_{direct}$  de gemiddelde boekwaarde van de steekproef. De  $b$  wordt bij de verschilschatter op 1 gesteld. Er wordt bij de verschilschatter namelijk van uitgegaan dat de verschillen tussen boek- en auditwaarden nul zullen zijn. Een correctie aan de gemiddelde boekwaarde,  $\bar{Y} - \bar{y}_{direct}$ , geeft dus een evenredige correctie aan de gemiddelde auditwaarde.

---

<sup>2</sup> De Student-verdeling wordt gebruikt bij stochasten die uit een zekere normale verdeling met onbekende standaarddeviatie komen. Bij een voldoende aantal stochasten is de eis van het normaal verdeeld zijn van de stochasten minder stringent. Omdat het aantal stochasten die voor het schatten van de totale auditwaarde gebruikt wordt groot genoeg is ( $n > 30$ ), kan hier de Student-verdeling worden gebruikt.

De variantie van het gemiddelde wordt dan, uitgaande van (17):

$$(18) \quad S_{\bar{x}_{verschil}}^2 = \frac{1 - \frac{n}{N}}{n(n-1)} \left( \sum_{i=1}^n x_i^2 - 2b \sum_{i=1}^n x_i y_i + b^2 \sum_{i=1}^n y_i^2 - n(\bar{x}_{direct} - b\bar{y}_{direct})^2 \right)$$

Het betrouwbaarheidsinterval zal een stuk korter zijn dan bij de directe schatter. De maximale fout zal daardoor vaak een stuk kleiner worden waardoor de steekproef eerder toereikend geacht zal worden. Bij de *regressieschatter* wordt er niet meer vanuit gegaan dat er geen fouten tussen boek- en auditwaarden aanwezig zullen zijn. Door de puntenwolk  $(x_i, y_i)$  kan er een kleinste kwadratenlijn getrokken die de totale afstand tussen de regressielijn en de afwijkingen ten opzichte van deze regressielijn minimaliseert.

De  $b$  wordt in dit geval:

$$(19) \quad b = \frac{\sum_{i=1}^n x_i y_i - n\bar{x}_{direct}\bar{y}_{direct}}{\sum_{i=1}^n y_i^2 - n\bar{y}_{direct}^2}$$

Deze  $b$  wordt verkregen door (18) te differentiëren naar  $b$  en deze afgeleide op nul te stellen. Formule (18) vormt een dalparabool. Er is dus sprake van een globaal *minimum*.

De variantie over het gemiddelde wordt:

$$(20) \quad S_{\bar{x}_{regressie}}^2 = S_{\bar{x}_{direct}}^2 - \frac{1 - \frac{n}{N}}{n(n-1)} \left( \frac{\left( \sum_{i=1}^n x_i y_i - n\bar{x}_{direct}\bar{y}_{direct} \right)^2}{\sum_{i=1}^n y_i^2 - n\bar{y}_{direct}^2} \right)$$

Deze variantie neemt de kleinste waarde aan vergeleken met de varianties van andere methoden die gebruik maken van een gecorreleerde hulpvariabele.

### 6.3.3 Betrouwbaarheidsinterval schatters

Met de Centrale Limiet Stelling kan voor de *directe schatter* een betrouwbaarheidsinterval rondom het populatiegemiddelde geconstrueerd worden. Omdat de sigma moet worden geschat uit de steekproef zal de Student-verdeling gebruikt worden. Bij voldoende grootte van de steekproef zou de tabel voor Normale verdeling gebruikt kunnen worden.

De *verschilschatter* maakt gebruik van een hulpvariabele, de boekwaarde. Om een betrouwbaarheidsinterval rondom het populatiegemiddelde te construeren met behulp van de normale verdeling (bij voldoende grote steekproef) is het van belang dat de parameters van de normale verdeling, de  $\mu$  en de  $\sigma^2$ , geschat kunnen worden door respectievelijk (17) en (18). Omdat zelfs in de meest ongunstige situatie (het gemiddelde niet normaal verdeeld en de sigma is onbekend) de normale verdeling gebruikt kan worden<sup>3</sup>, kan men ervan uitgaan dat bij voldoende grootte van de steekproef de verschilschatter in principe altijd kan worden toegepast.

---

<sup>3</sup> Uit een intern onderzoek van Deloitte zijn formele afleidingen gedaan waaruit dit bleek.

Bij het toepassen van regressie-analyse wordt meer geëist, namelijk dat de individuele afwijkingen rond de regressielijn normaal verdeeld moeten zijn. Deze eis komt voort uit het feit dat bij normaal verdeelde storingstermen de parameter  $b$  en de variantie van het regressiegemiddelde, de kleinste kwadraten lijn als beste schatter voor het regressiegemiddelde geldt<sup>4</sup>. De eis van normaliteit van de storingstermen is equivalent aan het stellen van de eis aan de normaliteit van de residuen. Dit is een grootheid die in tegenstelling tot de storingstermen makkelijk te toetsen valt. In praktijk blijkt echter dat het merendeel van de residuen gelijk aan nul is. De nulhypothese, de residuen zijn normaal verdeeld, zal dan ook door alle toetsen op normaliteit worden verworpen. Kaplan [1973] stelt voor om alleen de afwijkingen te testen. Dit brengt het aantal vrijheidsgraden bij gebruik van de Studentverdeling terug van  $n - 3$  naar: aantal afwijkingen  $- 3$ . Het zal dan over het algemeen grotere betrouwbaarheidsgrenzen geven dan de verschillenschatter.

Uit een stageonderzoek (T. Lew, VU, 1999) blijkt het argument dat de residuen normaal verdeeld moeten zijn, kan worden afgezwakt. De OLS-regressie blijkt ook goede resultaten te geven bij non-normaliteit van de residuen. De afzwakking van de eis kan men doen wanneer de normale verdeling grotere overschrijdingskansen (dikkere staart) heeft dan de werkelijke verdeling. Het geeft dan een conservatievere, een breder, interval dan eigenlijk noodzakelijk is. Kortom, de storingen hoeven dus niet perse een normale verdeling te hebben.

## 6.4 Evaluatie Regressieschatter

Bij de analyse met behulp van de regressieschatter gelden net als bij de directe- en de verschillenschatter de volgende prestatie-maten:

- Het kwantificeren van de steekproefkwaliteit. Dit gebeurt met behulp van de geschatte maximale fout. De maximale fout zal vergeleken moeten worden met de opgestelde materialiteit.
- Het berekenen van de geprojecteerde fout

De waarden van de geprojecteerde en maximale fout bepalen welke acties ondernomen moeten worden wil men de post goedkeuren. Dit kan overzichtelijk in een tabel (Tabel 6) worden gezet.

Tabel 6: acties naar aanleiding van waarde maximale fout. In deze tabel is:  
*BF = Bekende Fout, de fout die in de steekproef gevonden is*  
*PF = Geprojecteerde Fout, de fout die waarschijnlijk nog in de populatie zit*

	Y in interval	Y buiten interval
MF < materialiteit	Accepteer Y; BF mag blijven	n.v.t.
MF > materialiteit	Acceptatie mogelijk na hertelling	Halve intervalbreedte < materialiteit => suggereer correctie van PF Halve intervalbreedte > materialiteit => toereikende controle niet bereikt

### 6.4.1 Excentriteit Interval

De intervallen rondom de schatting worden geconstrueerd door uit te gaan van een 90% betrouwbaarheid. Men wil namelijk een 95%  $((1 - \alpha / 2) * 100\%)$  betrouwbare ondergrens of een bovengrens met dezelfde betrouwbaarheid, voor het berekenen van de maximale fout aangezien men enkel geïnteresseerd is in het deel van het interval waar de boekwaarde van de massa zich bevindt. Wanneer men deze betrouwbaarheid aanneemt in plaats van 95% wordt het interval kleiner, de excentriteit van de normale verdeling verandert dan van 1.96 naar 1.65.

<sup>4</sup> De  $b$  is de beste schatter wanneer de storingstermen een verwachting van 0 hebben [Hoofdstuk 5 van scriptie T. Lew, 1999]. De lineaire kleinste kwadraten lijn geeft een zuivere en beste schatter voor de variantie als de kurtosiscoëfficiënt gelijk aan 3 is. Aangezien deze coëfficiënt bij de normale verdeling ook 3 is, kan hieruit worden afgeleid dat de normale verdeling voor de storingstermen moet gelden wil de kleinste kwadraten lijn als beste schatter gelden in de klasse van de lineaire schatters.

## 6.5 Steekproef Voorraden In De Praktijk

De post *Vorraden* op de balans kan op een willekeurig tijdstip met de tabel hieronder, *Tabel 7*, worden weergegeven:

*Tabel 7: een momentopname van een fictieve voorraad.*

Artikelno.	Aantal (q) (boek)	Prijs (p) p/s	Boekwaarde ( $p \cdot q$ )
10001	10	4.89	48.90
10002	132	1.30	171.60
10003	20	3.03	60.60
10004	66	2.99	197.34
10005	73	0.89	64.97

Een bedrijf heeft een aantal artikelnummers en van elk artikelnummer heeft het bedrijf het aantal in de boeken staan. De prijs van een artikel wordt bijvoorbeeld via de moving average methode bepaald. Dit is een methode waarbij de gemiddelde prijs van een bepaald artikel over een bepaalde periode wordt berekend. De boekwaarde van een artikelnummer is gelijk aan  $p \cdot q$ . Bij een postensteekproef worden de artikelnummers als steekproefelementen beschouwd. Per artikelnummer wordt het werkelijke aantal dat door telling door accountants wordt vastgelegd, met de moving average price vermenigvuldigd. Dit vormt de auditwaarde van het betreffende artikelnummer.

Fouten in voorraden worden voornamelijk veroorzaakt door:

- Verwisselingsfouten; een item wordt op een verkeerde plek in het magazijn gelegd. Dit maakt 90% van de foute subposten uit. Dit soort fouten kunnen verschillende oorzaken hebben zoals: producten die teveel op elkaar lijken qua vorm of qua barcode, oververmoeidheid bij personeel, etc.
- Administratieve fout: op systemen waar de voorraad handmatig bijgewerkt moet worden, kunnen fouten ontstaan door bijvoorbeeld typefouten.
- Telfout: de artikelen worden handmatig geteld, dit kan tot foute tellingen leiden

Bij het toepassen van de regressie-analyse bij *Vorraden* worden de volgende stappen ondernomen om de effectiviteit van de analyse te verbeteren:

1. Het bepalen van de *top-stratum grens*. Deze grens geeft aan vanaf welke waarde de artikelnummers integraal geteld moeten worden. De hoge voorraadwaarden kunnen wijzen op twee oorzaken: er is een hoge afwijking bij enkele artikelen naar boven aanwezig of het gaat om artikelen van hoge waarde. Wanneer er in de laatste soort artikelen een afwijking aanwezig zou zijn dan kan dit een relatief grote invloed uitoefenen op de steekproefwaarden. Het gemiddelde wordt dan aardig opgeschroefd. Daarom zullen dit soort posten integraal geteld worden. De top-stratumgrens wordt bepaald door de 35 - 50 hoogste waarden van de op boekwaarde gesorteerde voorraadlijst als artikelen te markeren die geteld moeten worden, en de eerstvolgende artikelwaarde als de grens te markeren. Eventueel kan dit bedrag naar boven afgerond worden naar een mooi geheel getal.
2. Het controleren van nulwaarden in de voorraadlijst. Deze nullen kunnen worden onderscheiden in toevallige en echte nullen. Toevallige nullen zijn nullen die bij het uitdraaien van de op waarde gesorteerde voorraadlijst net al dan niet terecht op de hoeveelheid nul stonden. De betreffende artikelen zijn wellicht uitverkocht of er is een foute positie in het bestand. Hierop dient een volledigheidscntrole te worden gedaan. De auditwaarden kunnen mogelijk hoger liggen. Echte nullen zijn nullen waarvan de bijbehorende artikelen niet meer geleverd worden door het bedrijf maar nog wel geregistreerd staan. Deze nullen zullen verwijderd moeten worden.



3. Het verwijderen van de negatieve getallen. Een negatief voorraadaantal kan theoretisch niet voorkomen. In de praktijk blijkt dit echter wel mogelijk in de voorraadadministratie. Het kan te maken hebben met een te vroeg afgeboekte verkoop of een te laat bijgeboekte inkoop. Als correctie worden de boekwaarden naar de boekhoudkundige waarden gecorrigeerd. Is dit echter niet mogelijk, dan worden deze waarden op nul gezet, de laagst mogelijke waarde die een voorraad van een artikel kan hebben. De boekwaarde van de totale voorraad dient wel aangepast te worden na deze correcties.
4. Uit de overige artikelnummers worden 225 steekproeven getrokken. Van dit aantal gelden 25 steekproeven als reserve, in geval dat sommige artikelnummers bijvoorbeeld out-of-date blijken te zijn, ze zijn fysiek niet meer aanwezig.

Zijn er tweehonderd artikelen gecontroleerd, dan kan men met behulp van regressie-analyse de kwantitatieve grootheden, de maximale en geprojecteerde fout, berekenen.



## 7 Onderzoeksvragen

Tijdens de stage zijn er twee onderzoeksvragen geformuleerd. De onderzoeksvragen liggen op het gebied van *Continuous Auditing (CA)*. Deze term wordt vaak in verband gebracht met dynamisering van de controle, dit ten onrechte. Bij dynamisering van de controle worden de accountantscontrole-inspanningen over een periode verspreid om aan het eind van deze periode een uitspraak te doen over het controle-object. Met CA is de doelstelling echter dat continu een uitspraak gedaan kan worden over een continu wijzigend controle-object.

In mijn scriptie zal ik twee huidige controle technieken transformeren naar een mogelijke CA-toepassing en onderzoeken naar het praktische nut van deze nieuwe toepassingen. De technieken die in de onderzoeksvragen centraal staan zijn:

1. Geldsteekproeven in zijn algemene vorm (*Onderzoeksvraag I*).
2. Kwantitatieve postensteekproeven. Specifieker: de tool waarmee met behulp van regressie-analyse de grootte van de fout in de balanspost *Voorraden* continu kan worden geschat (*Onderzoeksvraag II*).



# 8 Onderzoeksvraag I

## 8.1 Inleiding

In de eerste onderzoeksvraag staat de vraag centraal of het mogelijk is om een CA-toepassing te formuleren en te construeren op basis van een huidige geldsteekproefmethode. In een intern artikel van Deloitte & Touche staat geschreven dat een continu veranderende massa hierbij een obstakel zal vormen. Ik zal in dit hoofdstuk een CA-toepassing formuleren en construeren waarbij een continu veranderende massa geen probleem zal vormen. In 8.2 zullen twee geldsteekproefmethoden in het kort worden uitgelegd. Een van deze twee blijkt geschikt te zijn voor het maken van een CA-toepassing. Onder 8.3 zal ik de steekproefmethode in zijn huidige gebruik behandelen om vervolgens in 8.4 de methode te transformeren naar een CA-toepassing.

## 8.2 Algemeen

Er bestaan twee varianten om een geldsteekproef uit te voeren, namelijk:

1. Men bekijkt aan het eind van het jaar de totale massa en aan de hand van deze massa ( $M$ ), de gekozen strategie ( $R(k,x)$ ) en de kritieke foutfractie ( $p_m$ ), bepaalt men de monetary precision ( $MP$ ) en de grootte van de steekproef ( $n$ ).
2. Men raamt de massa voor het eind van het jaar, en gebruikt  $R(k,x)$  en  $p_m$  om de  $MP$  en de celgrootte ( $J$ ) te bepalen. De celgrootte wordt dan gebruikt om de steekproef over het hele jaar te verdelen.

Het voordeel van de variant 1 boven variant 2, is dat men bij variant 1 uitgaat van uiteindelijke massa en men dus ook makkelijk het exacte aantal steekproefelementen kan bepalen dat nodig is. Variant 2 heeft echter het voordeel dat de steekproefelementen verdeeld wordt over de tijd en men dus niet alles aan het eind hoeft te controleren. De schatting van de massa heeft consequenties voor het aantal steekproefelementen dat gebruikt zal worden. Dit kan dus ook in het nadeel uitvallen. Bij een te hoge inschatting van de uiteindelijke massa zal men uiteindelijk meer inspanningen verrichten dan in feite noodzakelijk was.

Voor een CA-toepassing bij de geldsteekproeven leent variant 2 zich uitstekend. Bij de CA-toepassing zal naast de eindmassa ook de tussentijdse massa's worden geschat. Hiermee verkrijgt men op verschillende momenten inzicht het effect van de schattingen op het aantal steekproefelementen: voldoet het geschatte aantal steekproefelementen op tijdstip  $t$  nog steeds? Men houdt dus een continue controle.

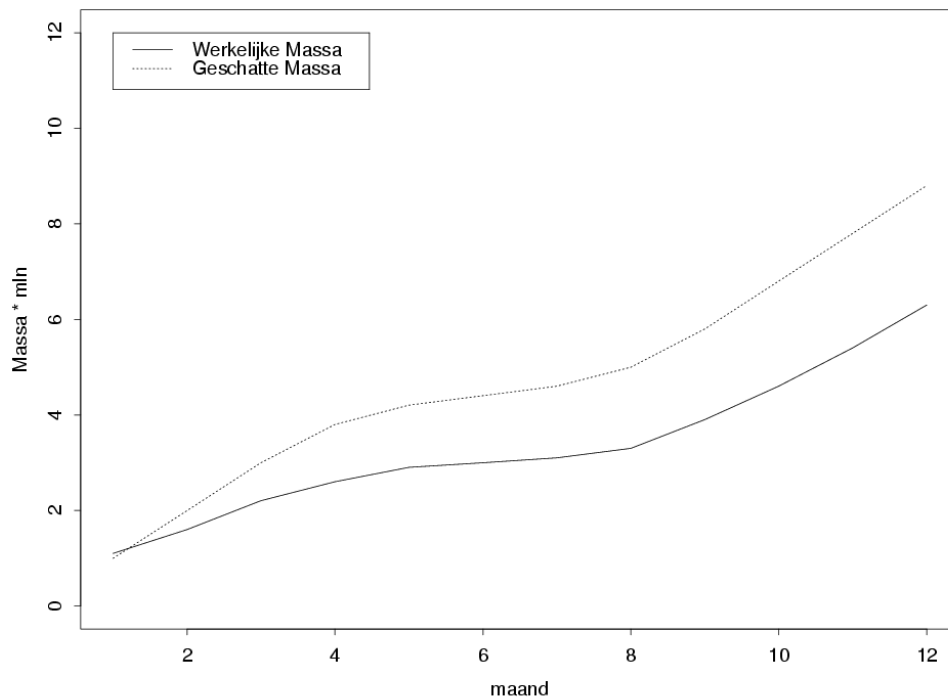
Ter illustratie van de werking van de 'oude' toepassing en de CA-toepassing zal ik gebruik maken van het volgende (fictieve) voorbeeld:

In *Tabel 8* is voor elke maand de massa tot en met die maand geschat. In *Figuur 2* zijn de waarden in *Tabel 8* grafisch weergegeven:

*Tabel 8: verloop geschatte en werkelijke massa.*

Massa(*mln)/Tijd(mnd)	1	2	3	4	5	6	7	8	9	10	11	12
Werkelijke massa	1.1	1.6	2.2	2.6	2.9	3.0	3.1	3.3	3.9	4.6	5.4	6.3
Geschatte massa	1.0	2.0	3.0	3.8	4.2	4.4	4.6	5.0	5.8	6.8	7.8	8.8

## Verloop Geschatte en Werkelijke Massa



Figuur 2: grafische weergave van Tabel 8

### 8.3 ‘Oude’ toepassing

In periode nul zal de massa op periode twaalf worden geschat. Als geschatte massa nemen we de massa in Tabel 12, deze is gelijk aan 8.8 miljoen. De kritieke foutfractie wordt op  $p_m = 0.050$  gesteld. De  $MP$  wordt dan:  $MP = p_m \cdot \hat{M}_{12} = 0.050 \cdot 8.8m \ln = 440000$  waarin  $\hat{M}_{12}$  de raming van de massa in maand 12 voorstelt, de eindmassa.

Om een steekproef over het hele jaar te doen en zo een gelijkmatige verspreiding over een hele massa te bewerkstelligen, maakt men gebruik van de celgrootte. De formule voor deze celgrootte kennen we als:

$J = \frac{\hat{M}_{12}}{n}$ , maar kunnen we ook schrijven als  $J = \frac{MP}{R}$  ( $n \cdot p_m = R$ ,  $p_m = \frac{MP}{M}$ ). In principe is het resultaat hetzelfde en kunnen ze van elkaar afgeleid worden. Maar in dit geval geven we de voorkeur aan de tweede formule omdat deze met de beschikbare gegevens snel afgeleid kan worden. De bedoeling van de celgrootte is, is dat men bij elke  $J^e$  gulden een aselechte trekking doet uit de laatste  $J$  opgeboekte geldeenheden en de bijbehorende transactiepost gaat controleren. Het steekproefaantal wordt dan bij elke controle opgehoogd met één. Dit geeft uiteindelijk een totaal aantal steekproeven aan het einde van het jaar die met vermenigvuldiging van de celgrootte:  $n_{\text{totaal}} * J$ , ongeveer overeenkomen zou moeten komen met de totale boekwaarde.

Wanneer men de raming met de werkelijke massa  $M$  vergelijkt, kunnen zich drie situaties voordoen:

1.  $M > \hat{M}_{12} \Rightarrow \frac{MP}{\hat{M}_{12}} \cdot 100\% < p_m$ , de werkelijke massa is groter gebleken dan van te voren geschat werd.

Dit houdt in dat de foutfractie kleiner wordt dan de veronderstelde kritieke foutfractie. Het aantal steekproefelementen die door het hele jaar gedaan zijn is dus van voldoende niveau. Een gedeelte was

zelfs overbodig. Is de werkelijke boekmassa meer dan tweemaal zo groot als de geraamde boekmassa dan wordt het alsnog niet goedgekeurd.

2.  $M = \hat{M}_{12} \Rightarrow \frac{MP}{\hat{M}_{12}} \cdot 100\% = p_m$ , de werkelijke massa is gelijk aan de geschatte massa. Het aantal steekproeven zal voldoende omdat dit aantal gelijk zal zijn aan het aantal geschatte steekproefelementen waar men vanuit ging.
3.  $M < \hat{M}_{12} \Rightarrow \frac{MP}{\hat{M}_{12}} \cdot 100\% > p_m$ , dit is een situatie die bij een te positieve raming voordoet. De  $MP$  blijft absoluut gezien gelijk maar wordt procentueel groter. Het aantal steekproefelementen hoeft dan niet per se gelijk als ontoereikend bestempeld te worden. Immers, de norm is niet een percentage ( $p_m$ ) maar een bedrag ( $MP$ ), en aan die norm blijft voldaan. Open vraag is natuurlijk of die norm bij een sterk gewijzigde waarde van  $M$  nog steeds een redelijke norm is.

In het voorbeeld is de werkelijke massa 6.3 mln. Wanneer men dit in verhouding ziet met de  $MP$ , wordt de kritieke foutfractie van 0.052 (330 000/ 6.3 mln) verkregen. Dit is dus te hoog wanneer men een foutfractie van 0.050 als maximaal toelaatbaar beschouwt. De gecontroleerde post wordt dus nog niet goedgekeurd.

## 8.4 CA-toepassing

De CA-toepassing is geënt op de ‘oude’ methode. Naast het schatten van de eindmassa zullen nu ook de massa’s op tussenliggende momenten worden geschat. Op elk moment kan dan worden gecontroleerd of het totaal aantal steekproefelementen waarop uiteindelijk uitgekomen zal worden, wel van voldoende niveau is.

Voor de CA-toepassing wordt de kritieke foutfractie constant geacht,  $p_{m;1} = p_{m;2} = \dots = p_{m;T}$ . Dit geeft voor elk moment  $t$  een kritiek foutbedrag  $MP_t$ :

$$(21) \quad MP_t = p_{m;t} \cdot \hat{M}_t$$

Wanneer de  $M_t$  bekend is, kan een ‘werkelijke’ foutfractie op moment  $t$  worden geformuleerd met behulp van  $n_t$  en  $J$ :

$$(22) \quad p_{m;t} = \frac{MP_t}{M_t} \approx \frac{MP_t}{n_t \cdot J}$$

Deze waarde kan dan getest worden tegen de toelaatbare kritieke foutfractie. In *Tabel 13* staan van elke maand de werkelijke foutfracties ( $p_{m;t}$ ):

*Tabel 9: de foutfracties op  $t = 1, 2, \dots, 12$ .*

	1	2	3	4	5	6	7	8	9	10	11	12
$p_{m;t}$	0.34	0.47	0.51	0.55	0.54	0.55	0.56	0.57	0.56	0.55	0.54	0.52

Zoals men kan zien wijkt de  $p_{m;t}$  teveel af vanaf maand drie, wanneer het wordt vergeleken met de toelaatbare kritiekfractie van 0.05.

Op deze manier kan continu gecontroleerd worden of de kritieke foutfractie van toelaatbaar niveau is. Het is dus een mogelijke CA-toepassing. De continu veranderende massa wordt door het maken van schattingen op verschillende momenten, als probleem ondervangen.

Wanneer de kritieke foutfractie in het jaar buiten de grens valt, heeft dit geen consequenties. Er worden geen maatregelen genomen om de kritieke foutfractie binnen de grenzen te laten vallen. Het doen van extra steekproeven bij een te hoge beginschatting is geen optie, de steekproeven worden dan onevenredig over de massa verdeeld. Het tweede deel van de massa, het deel dat komt nadat de kritieke foutfractie te hoog blijkt, zal dan relatief zwaarder vertegenwoordigt qua steekproeven. Bovendien kan net als in het voorbeeld, de kritieke foutfractie na verloop van tijd weer kleiner worden. Het is tenslotte op elk tijdstip een momentopname.



## 9 ONDERZOEKSVRAAG II

### 9.1 Inleiding

Is het mogelijk een tool te ontwikkelen waarmee men met minimale inspanningen, continu een audit op een voorraad kan uitvoeren, een zogenaamde Continuous Auditing tool (CA-tool) voor voorraden?

In dit hoofdstuk staat een onderzoek beschreven waarin de mogelijkheden van een ontwikkelde CA-tool voor voorraden wordt getoetst. De betreffende CA-tool bestond uit twee onderdelen: een aangepaste regressieschatter (de regressieschatter-plus (*RP*)) en een zogenaamd ‘vertalingsmechanisme’ (*VM*). Met de *RP* werd met behulp van kleine steekproeftrekkingen uit massa’s afkomstig van verschillende controlemomenten, output geleverd die het foutverloop in de voorraad reflecteert. Het *VM* diende de output van de *RP* (continu) te vertalen naar één van de twee mogelijke uitspraken: “De fout in de voorraad zit onder de opgestelde norm” of “De fout in de voorraad heeft de opgestelde norm overschreden”. Het doel van het onderzoek was om die instellingen te vinden van de *RP* en het *VM* waarbij met een bepaalde zekerheid gesteld kan worden dat de uitkomst, één van de twee hierboven genoemde uitspraken, op een bepaald controlemoment<sup>5</sup> correct is. De waarden van deze instellingen vormen dan weer een indicatie voor de bruikbaarheid van de ontwikkelde CA-tool om de voorraad continu te kunnen auditen.

In paragraaf 9.2 komt de opzet van het onderzoek aan de orde. Daarin staat beschreven hoe de CA-tool werkt en hoe de CA-tool wordt getoetst. Paragraaf 9.3 beschrijft de daadwerkelijke uitvoering van het onderzoek en de resultaten ervan.

### 9.2 Onderzoeksopzet

De werking van de *RP* staat beschreven in 9.2.1. In 9.2.2 komt het *VM* aan bod. Om te meten hoe goed de CA-tool presteert onder verschillende instellingen van de *RP* en het *VM*, is een simulatieprogramma gebruikt. De plaats die het simulatieprogramma inneemt in het onderzoek staat beschreven in 9.2.3. In het simulatieprogramma wordt de CA-tool getoetst tegen een tweetal totaal verschillende fictieve foutverlopen. Deze foutverlopen staan beschreven in 9.2.4.

In het simulatieprogramma worden voorraden gesimuleerd. In het onderzoek zijn de voorraden naar een bestaand voorraadbestand gemodelleerd. Het modelleren komt in 9.2.5 aan bod.

---

<sup>5</sup> In het onderzoek beslaat (voor het overzicht) de periode tussen twee controlemomenten een week. In principe kan elk tijdsdeel (dag, maand, jaar) hiervoor ingevuld worden.

### 9.2.1 Regressieschatter-Plus

De *RP* (regressieschatter-plus) werkt als volgt:

1. Men neemt een steekproef ter grootte van  $n$  op controlemoment  $t = 0$ . Daarvoor verdeelt men de voorraad in een  $n$  aantal disjuncte cellen,  $c_1, c_2, \dots, c_{n-1}, c_n$ . De vereniging van deze cellen bevat dan alle artikelen die in de voorraad aanwezig zijn. Uit elke cel wordt vervolgens één steekproefelement (artikel) gekozen. Dit principe kennen we als cellsampling. Elke cel omvat voor alle controlemomenten ( $t = 0, 1, \dots, T$ ) dezelfde subposten. Gedurende de periode waarop wordt gecontroleerd verandert de voorraad niet wat betreft het aantal subposten. De steekproefelementen zijn afkomstig van de massa op  $t = 0$ , de  $M(0)$ . De steekproef op controlemoment  $t = 0$  wordt genoteerd met  $s(0)$ .
2. Op de daaropvolgende controlemomenten ( $t = 1, t = 2, \dots, t = T$ ) kiest men op elk tijdstip  $t$  at random een  $d$  ( $\ll n$ ) aantal elementen uit  $s(t-1)$ . Elk aangewezen element behoort tot een bepaalde cel,  $c_x$ . Uit elk van deze cellen, deel uitmakend van massa  $M(t)$ , wordt at random een element gekozen. Voordat dit element het element in  $s(t-1)$  afkomstig uit dezelfde cel vervangt, wordt het element in  $s(t-1)$  met de factor  $\frac{M(t)}{M(t-1)}$  vermenigvuldigd, waardoor de informatie van de steekproef zich aanpast aan de nieuwe massa. Zijn de aangewezen elementen in  $s(t-1)$  vervangen dan is er een nieuwe  $s(t)$ . De  $s(t)$  bestaat dus voor alle  $t$  uit  $n$  steekproefelementen. Aan de hand van deze  $s(t)$  worden de geprojecteerde en maximale fout op tijdstip  $t$  berekend. Dit wordt in het vervolg aangegeven met respectievelijk:  $PF'(t)$  en  $MF'(t)$ . Deze waarden zijn van wezenlijk belang voor het volgen van de fout.

Het basisidee achter deze toepassing is dat men op  $t > 0$  maar een kleine steekproef doet uit  $M(t)$ . De steekproefelementen uit deze steekproef worden vervolgens ‘vermengd’ met die van vorige massa’s. De  $PF'$  en  $MF'$  reflecteren dan het verloop van de *gesaldeerde* fout in de voorraad in de tijd.

De *RP* geeft dus als output:  $PF'$  en  $MF'$ . De  $PF'$  en  $MF'$  en de gesaldeerde fout worden in het vervolg in een percentage van de totale boekwaarde uitgedrukt. Het percentage gesaldeerde fout dat *werkelijk* in de voorraad aanwezig is, wordt aangeduid met ‘*PGF*’ (percentage gesaldeerde fout).

### 9.2.2 Vertalingsmechanisme

De  $PF'$ - en  $MF'$ -waarden zijn gebaseerd op een beperkt aantal steekproefelementen afkomstig van verschillende massa’s. De  $PF'$  en  $MF'$  zijn dus op zichzelf geen goede indicatie voor een fout op een controlemoment. Daarom is een vertaling van de output vertaling nodig om op grond van deze data een antwoord te kunnen geven op de vraag, of de fout de opgestelde norm op het betreffende controlemoment wel/niet heeft overschreden.

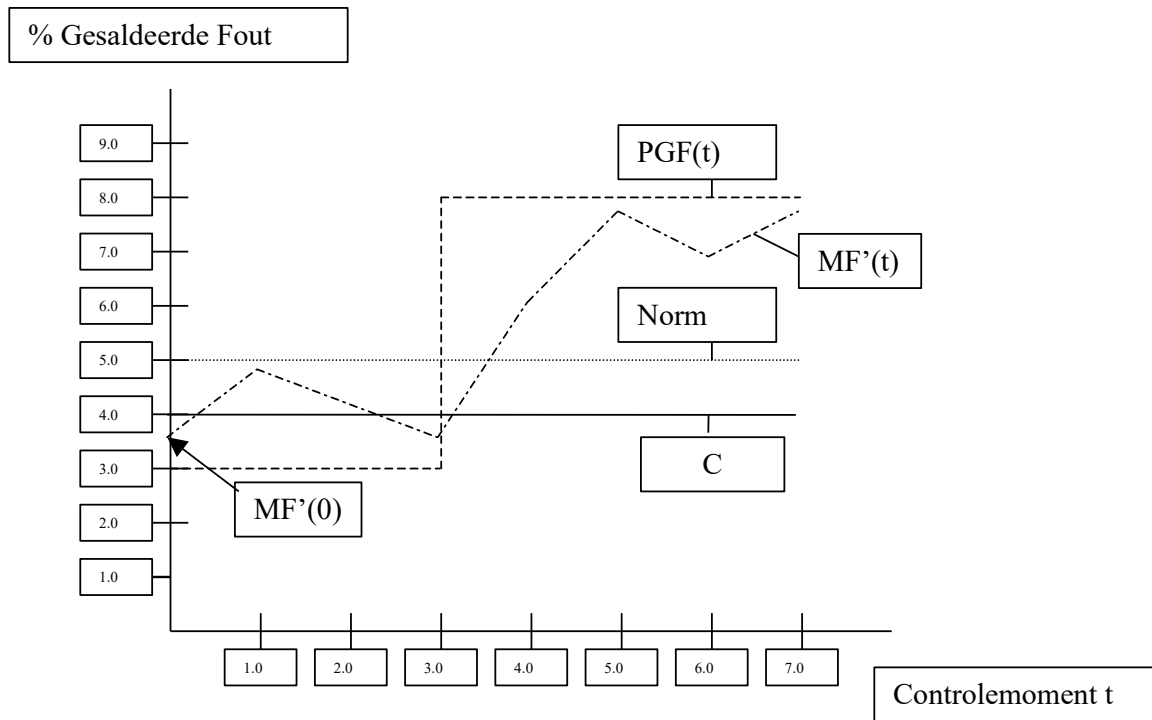
Het *VM* (vertalingsmechanisme) gaat uit van een simpel principe. Op elk controlemoment wordt de  $MF'$  berekend met behulp van de *RP*. Het berekenen van de  $MF'$  gebeurt onder een 95% betrouwbaarheidsinterval ( $\alpha = 0.05$ ). De  $MF'$  wordt in plaats van het  $PF'$  in het onderzoek gebruikt als indicator vanwege de volgende redenen:

1. De  $MF'$  vertoont een kleinere variantie dan de  $PF'$  en is daardoor meer geschikt als datapunt.
2. De  $MF'$  vormt op tijdstip 0 de bovengrens waar de absolute waarde van het *PGF* met 95% zekerheid onder zit. Dit is een bruikbaar gegeven in een mogelijk vervolgonderzoek.

De berekende  $MF'$  wordt op elk controlemoment ( $t = 0, 1, \dots, T$ ) getoetst tegen een vaste grens  $C$ . Als de  $MF'$ , een vastgesteld  $K$  aantal achtereenvolgende controlemomenten een hogere waarde aanneemt als de  $C$ , dan volgt de uitspraak dat het *PGF* (percentage gesaldeerde fout) in de voorraad een vastgestelde norm heeft overschreden. De waarden voor de  $C$ ,  $K$  en de norm dienen dus vooraf bepaald te worden. Bevindt de  $MF'$  over de gehele periode geen enkele keer een  $K$  achtereenvolgende controlemomenten boven de norm dan zal

de CA-tool aan het eind van de controleperiode de uitspraak doen dat het *PGF* op geen enkel controlemoment de opgestelde norm heeft overschreden. Waarom niet continu (op elk controlemoment) getoetst wordt of het *PGF* de norm niet heeft overschreden, komt in 9.2.3 aan de orde.

De werking van het *VM* in combinatie met de *RP*, samen dus de CA-tool vormend, kan geïllustreerd worden met het onderstaande figuur waarin met fictieve waarden is gewerkt:



Figuur 3: fictief verloop van het *PGF* en *MF'*

In *Figuur 3* zijn de instellingen van het *VM*:  $C = 4\%$ ,  $K = 3$  en  $\text{norm} = 5\%$ . Men kan zien dat de *MF'(t)* op controlemoment 6, al 3 achtereenvolgende controlemomenten een hogere waarde dan de 4% ( $C$ ) aanneemt. De CA-tool doet dus op controlemoment 6, onder deze instellingen, de uitspraak “het *PGF* heeft de norm overschreden”.

### 9.2.3 Simulatieprogramma

Om te meten hoe goed de CA-tool werkt onder een gegeven *PGF*-verloop, is een simulatieprogramma gebruikt. Per onderzochte instelling omtrent de parameters in de *RP* en het *VM*, zijn een 1000-tal simulatieruns uitgevoerd. In het simulatieprogramma wordt per run nagegaan of de CA-tool een juiste uitspraak levert. In het simulatieprogramma zijn vier uitkomsten mogelijk met betrekking tot de uitspraak van de CA-tool:

- CA-tool geeft correct de uitspraak dat de norm is overschreden door het *PGF*
- CA-tool geeft incorrect de uitspraak dat de norm is overschreden door het *PGF*
- CA-tool geeft correct de uitspraak dat de norm op geen enkel controlemoment is overschreden door het *PGF*
- CA-tool geeft incorrect de uitspraak dat de norm op geen enkel controlemoment is overschreden door het *PGF*

De laatste twee punten geven dus aan dat niet op elk controlemoment een uitspraak wordt gedaan. Dat voor deze formulering (uitspraak) is gekozen heeft te maken met het feit dat de CA-tool (in principe) een klein aantal steekproefelementen zal gebruiken om een uitspraak te doen. Een overschrijding van de norm door het *PGF* zal dus vaak niet direct terug te vinden zijn in vervangingselementen die gebruikt zijn. Met als gevolg dat het *MF'*-verloop in de meeste gevallen een (ver)late reactie is op het *PGF*-verloop. Het op elk controlemoment toetsen van de uitspraak “De norm is niet overschreden” is in het onderzoek dus niet van toepassing. Een *K*-waarde groter dan 1 maakt eveneens een continue toetsing op het niet overschrijden van de norm onmogelijk.

Het simulatieprogramma geeft uiteindelijk het percentage runs dat de CA-tool onder de instellingen een correcte uitspraak heeft gedaan. Dit percentage wordt in het vervolg afgekort met *CP* (correctheidspercentage). Dit is een belangrijke waarde in het onderzoek aangezien dit bij een oneindig aantal runs gelijk zal zijn aan het betrouwbaarheidspercentage. In het onderzoek worden die instellingen gezocht van de *RP* en het *VM* waarbij het *CP* minimaal 95% is.

De *CP*-waarden die in het onderzoek verkregen zijn, representeren de waarde die na 1000-tal runs zijn verkregen. Deze waarden zijn niet de werkelijke *CP*-waarden, ze zijn immers via een *beperkt* aantal runs verkregen. In het onderzoek worden de behaalde *CP*-waarden voor het gemak wel als werkelijke waarden beschouwd.

Verder bestaat een simulatierun uit 51 controlemomenten ( $t = 0, 1, \dots, 50$ ). In het onderzoek zal met een aantal *K*-waarden worden gewerkt. Een  $K = 1$  geeft aan dat er daadwerkelijk 51 controlemomenten worden gebruikt. Bij een hogere *K* zal dit aantal met  $K - 1$  verkleind worden. Dit heeft invloed op de *CP*-waarden. Deze invloed is echter voor het gemak in het onderzoek op nihil gesteld.

In het simulatieprogramma worden ook de controlemomenten bijgehouden waarop het *VM* aangeeft dat de norm is overschreden. Met deze gegevens kan men de reactiesnelheid testen van de *MF'*-verloop op een verandering van het *PGF*-verloop in het geval de CA-tool een juiste uitspraak levert..

### 9.2.4 Foutverloop

In het onderzoek zijn twee verschillende *PGF*-verlopen gebruikt. Een (specifiek) *PGF*-verloop wordt in het vervolg een scenario genoemd. De volgende twee scenarios zijn in het onderzoek gebruikt:

- Scenario 1: het *PGF* blijft constant op 1% over een periode van 51 controlemomenten.
- Scenario 2: het *PGF* is tot en met controlemoment  $t = 20$  gelijk aan 1%. Op controlemoment 21 vertoont het *PGF* een grote stijging,  $\Delta = 6\%$ . Deze stijging is eenmalig.

De gekozen scenarios zijn ‘extreem’ van aard. Bij scenario 1 zal de CA-tool zeker moeten aangeven dat het *PGF* de norm niet overschrijdt, dit in tegenstelling tot scenario 2. Als de CA-tool op deze scenarios slecht reageert, zal een CA-tool in zijn huidige vorm in praktijk zeker geen uitkomst bieden. In praktijk zijn veranderingen in het *PGF*-verloop namelijk grimmiger van aard.

In scenario 2 zal de stijging in het *PGF* veroorzaakt worden door een stijging in het aantal subposten met een residu ongelijk aan nul. De residuen zullen ten opzichte van de boekwaarden geen hogere waarden gaan aannemen. Het controlemoment waarop de stijging van het *PGF* plaatsvindt, is relatief vroeg in de controleperiode. Zodoende heeft de CA-tool voldoende controlemomenten tot zijn beschikking om te reageren. Bedenk dat het controlemoment waarop de verandering van het *PGF* plaatsvindt invloed zal hebben op de prestatie van de CA-tool. Omdat het met name om een indicatie gaat worden er geen andere scenario's gebruikt om deze invloed nader te bepalen.

Notabene, in het onderzoek is geen rekening gehouden met de verhouding tussen het aantal over- en onderwaarderingen. Volgens een intern artikel van Deloitte & Touche heeft de verhouding nauwelijks invloed op de prestaties van de regressieschatter

### 9.2.5 Kenmerken van de voorraad

De prestaties van de CA-tool worden sterk beïnvloed door de kenmerken van de voorraad. De kenmerken van de voorraad die wordt gebruikt in het simulatieprogramma zijn afgeleid van een bestaand voorraadbestand (*Appendix B*). Dit voorraadbestand is ontstaan van de artikelen die integraal geteld moesten worden (6.5), de hogere waarden in de voorraad. Voor het gemak zijn deze buiten beschouwing gelaten. In *Appendix C* staat de analyse van dit voorraadbestand beschreven. Deze analyse heeft modellen opgeleverd waarmee een theoretische voorraad is geconstrueerd. De modellen zijn van belang voor het simulatieprogramma. De modellen zelf geven inzicht in de boek- en residuwaarden (en daarmee ook de auditwaarden).

Hoe representatief het voorraadbestand is voor de voorraden is niet uitgebreid onderzocht. De prestaties van de CA-tool zijn dus gebaseerd op de voorraad gemodelleerd naar het voorraadbestand.

## 9.3 Uitvoering en resultaten

In het onderzoek naar de mogelijkheden van de ontwikkelde CA-tool, is de CA-tool onder verschillende *RP*- en *VM*-instellingen getoetst tegen een tweetal scenario's. Hierbij is gebruik gemaakt van het simulatieprogramma. Het doel van het onderzoek kon nader gespecificeerd worden in het vinden van die instellingen waaronder de uitspraak van de CA-tool met een 95% zekerheid juist is. Deze instellingen zijn in eerste instantie per scenario onderzocht. Vervolgens is nagegaan of er een (specifieke) *K*-waarde bestaat waarbij de CA-tool bij beide scenario's (9.2.4) een goed *CP*-waarde oplevert.

In het onderzoek dienen basisinstellingen van de *RP* en het *VM* als uitgangspunt. De basisinstellingen staan beschreven in 9.3.1. In deze paragraaf komen ook de resultaten onder de basisinstellingen aan bod. Vervolgens komen in 9.3.2 de resultaten aan de orde die onder veranderingen van enkele parameters in de *RP* en het *VM* zijn gehaald. Zo is de mogelijke invloed van een enkele parameter op het *CP* nagegaan. De resultaten die onder alle instellingen zijn gehaald, staan in *Appendix D*.

### 9.3.1 Basisinstellingen

#### *RP*

In de *RP* zijn de steekproefgrootte  $n$  en het aantal vervangingselementen  $d$  de belangrijkste variabelen. De  $n$  en de  $d$  lagen als basiswaarden op respectievelijk 200 en 40. Een  $n$  gelijk aan 200 is de grootte die men in de praktijk voor steekproeven ook gebruikt. In het onderzoek was niet de  $d$  maar de fractie  $d/n$  van belang. De  $d$  geeft namelijk geen inzicht in de verhouding nieuwe steekproefelementen die per controlemoment worden vervangen. De verhouding is bepalend voor de reactiesnelheid van de CA-tool.

Een andere variabele is de  $\alpha$ . Deze variabele heeft invloed op het betrouwbaarheidsinterval rondom de puntschatting  $PF'$  en dus ook op de prestaties van het CA-tool. De waarde voor deze parameter was in het gehele onderzoek gelijk aan 0.05 en verandert verder niet.

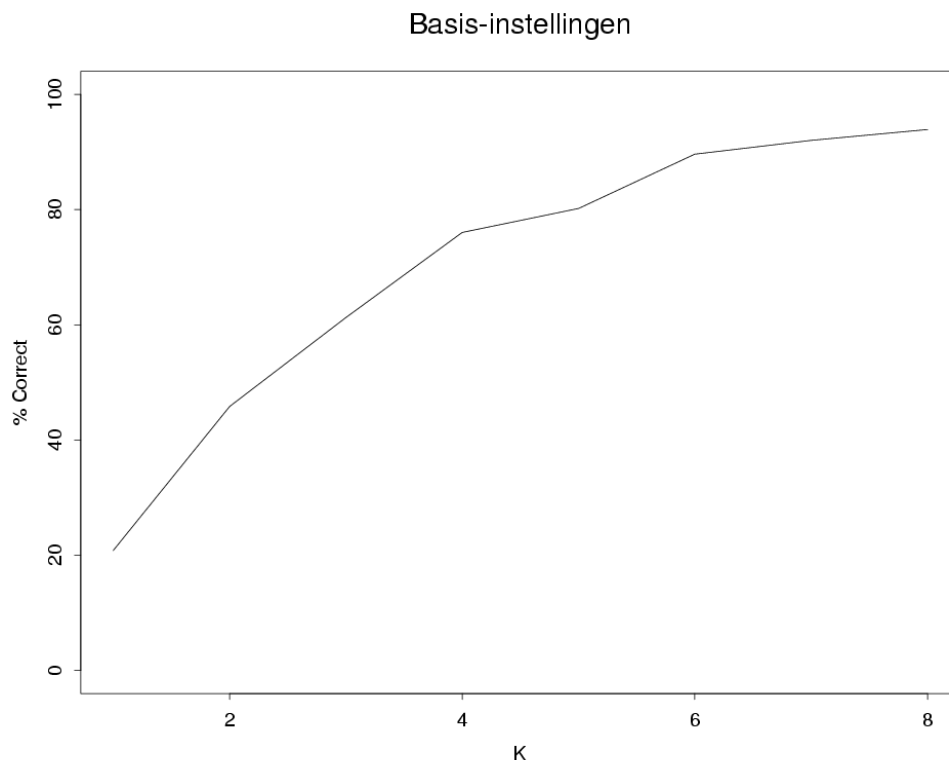
## VM

Het VM heeft als parameters de  $K$ ,  $C$  en de norm. De basiswaarden voor de norm en de  $C$  zijn voor beide parameters gelijkgesteld aan 5%. Per instelling van de parameterwaarden van de  $RP$  ( $n$  en de  $d$ ) en de  $VM$  ( $C$  en de norm) is de invloed van acht  $K$ -waarden (1,2,...,8) op het  $CP$  onderzocht.

### Resultaat Basisinstellingen

Onder de basisinstellingen is onder verschillende  $K$ -waarden (1-8) het  $CP$  uitgerekend. Wanneer de  $CP$  wordt uitgezet tegen de  $K$ -waarden onder scenario 1 ( $PGF$  blijft constant op 1%), geeft dit *Figuur 4*:

De grafiek moet men als volgt lezen: bij een  $K = 2$ , levert de CA-tool onder de huidige instellingen ( $n = 200$ ,  $d = 40$ ,  $C = 5\%$  en de norm = 5%) in ruim 45 % (45.8%) van de simulatieruns een correcte uitspraak op (“norm is op geen enkele controlemoment overschreden”). Het criterium van 95% wordt echter onder deze instellingen bij een  $K = 2$  niet gehaald.



*Figuur 4: het CP dat de CA-tool geeft onder scenario 1 en de basisinstellingen, uitgezet tegen de verschillende K-waarden*

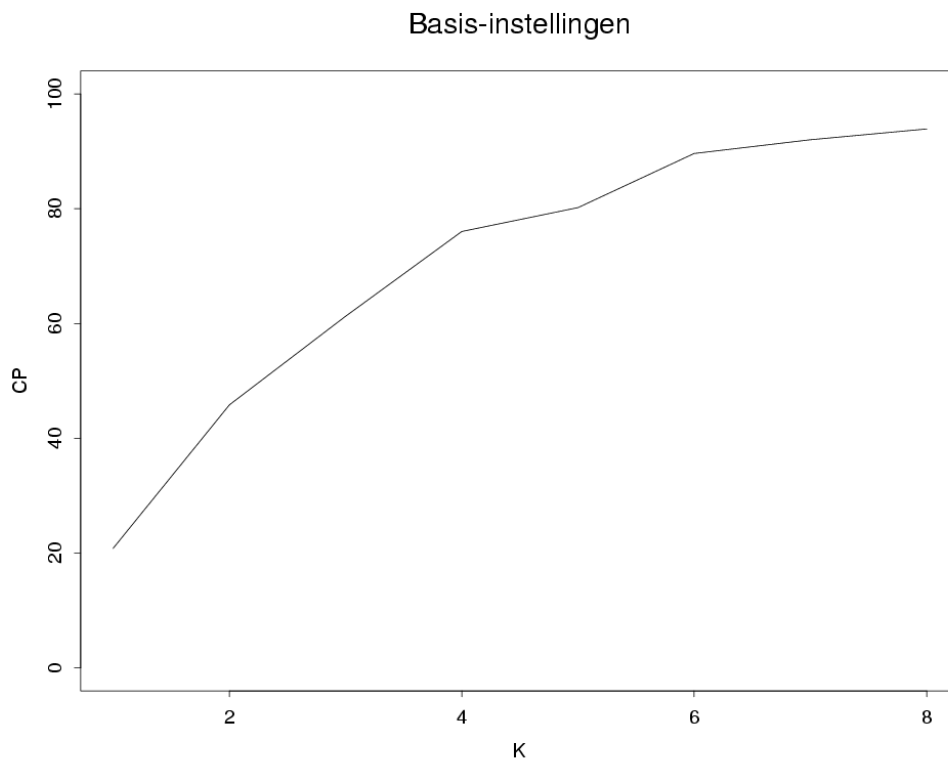
Onder de basisinstellingen wordt bij geen enkele  $K$ -waarde het criterium van 95% gehaald. Bij een maximum  $K$  die in het onderzoek geldt ( $K = 8$ ), geeft de CA-tool in 93.8% van de simulatieruns een correcte uitspraak. Het  $MF'$ -verloop blijft in de meeste runs langdurig boven de 5% hangen. Zelfs bij een  $K = 4$  heeft de CA-tool in ruim 20% van de runs de verkeerde uitspraak gedaan.

Het langdurig blijven hangen van de  $MF'$  boven de norm van 5%, terwijl het  $PGF$  toch gelijk aan 1% is, kan de volgende twee oorzaken hebben:

1. De  $MF'$  op controlemoment 0 is in feite de 95% bovengrens van het  $PGF$ . De  $MF'(0)$  zal dus met 95 % zekerheid het  $PGF$  op controlemoment 0 omvatten. Wanneer het betrouwbaarheidsinterval vrij ruim is, veroorzaakt door bijvoorbeeld ongelukkige steekproefelementen, wordt de gemiddelde afstand tussen de  $MF'$  en de  $PF'$  (en dus ook tussen de  $MF'$  en het  $PGF$ ) groter. Dit impliceert dat de  $MF'(0)$  vaker een veel hogere waarde aanneemt ten opzichte van het  $PGF(0)$ . In feite is er dus dan sprake van een slecht begin.

2. Het aantal steekproefelementen dat per controlemoment vervangen wordt ( $d$ ) heeft invloed op de flexibiliteit van de  $MF'$ . Bij een lage  $d$  kunnen 'ongelukkige' steekproefelementen langer in de steekproef blijven zitten. Dit kan als gevolg hebben dat gedurende een groot aantal controlemomenten de  $MF'$  een te hoge waarde aanneemt.

Onder scenario 2 ( $PGF$  stijgt op controlemoment 21 van 1% naar 7%) werd de volgende grafiek verkregen onder de basisinstellingen:



*Figuur 5: het CP dat de CA-tool geeft onder scenario 2 en de basisinstellingen, uitgezet tegen de verschillende K-waarden.*

Het resultaat is beduidend beter dan onder scenario 1. Bij een  $K = 6$  wordt een  $CP$  gehaald van 95.1%. De CA-tool reageert onder scenario 2 beter omdat het  $PGF$  extreem verhoogd wordt in dit scenario (van 1% naar 7%). Na het controlemoment waarop de stijging heeft plaatsgevonden, geeft de CA-tool in het resterende deel van de controleperiode ( $t = 21, 22, \dots, 50$ ), in alle runs aan dat de norm is overschreden. Dat de CA-tool niet een  $CP$  haalt van 100%, heeft te maken met de eerste 21 controlemomenten ( $t = 0, 1, \dots, 20$ ) waarin het  $PGF$  constant is (1%). De kans dat de  $MF'$  een aantal controlemomenten lang boven de  $C$  blijft hangen terwijl er geen overschrijding plaatsvindt van de norm, zal dan ook een stuk kleiner zijn dan in scenario 1. Toch doet de CA-tool bij een  $K = 6$  in 4.9% van de runs op één van de eerste 21 controlemomenten, incorrect de uitspraak dat de norm is overschreden.

De  $K$ -waarde waaronder beide scenarios voldoende reageert, kan in dit geval niet gegeven worden. Een  $K$ -waarde gelijk aan het maximum (8) gaf bij scenario 1 een  $CP$  van 93.9%, en dit is kleiner dan het criterium.

### 9.3.2 Wijzigingen Instellingen

In het onderzoek zijn vervolgens drie parameters in waarde gewijzigd. Twee parameters in de  $RP$  en één parameter in het  $VM$ . De betreffende parameters in de  $RP$  en het  $VM$  zijn de  $n$ ,  $d$  en de  $C$ . Welke wijzigingen precies zijn doorgevoerd ten aanzien van deze parameters, staan hieronder beschreven in *Wijziging N*,

*Wijziging D en Wijziging C.* In *Resultaten N*, *Resultaat D* en *Resultaat C* staan de resultaten beschreven van deze wijzigingen.

#### **Wijziging N**

De steekproefgrootte wordt verhoogd van 200 naar 300. Dit zal het betrouwbaarheidsinterval rondom de puntschatting kleiner maken. Dit houdt in dat a) de beginschatting  $MF'(0)$  gemiddeld genomen dichter bij het  $PGF(0)$  zit en b) de fluctuaties in het  $MF'$ -verloop kleiner worden. De CA-tool zal bij een verhoging van de  $n$  dus betere  $CP$ -waarden geeft.

#### **Wijziging D**

Het aantal vervangingselementen wordt in het onderzoek zowel verlaagd als verhoogd.

Onder een steekproefgrootte van 200 was het aantal vervangingselementen onder de basisinstelling gelijk aan 40. Dit komt neer op een fractie vervangingselementen,  $d/n$ , van 0.2 ( $d/n = 40/200$ ). In geval van verlaging van het aantal vervangingselementen wordt de fractie op 0.1 gesteld ( $d/n = 20/200$ ). In geval van verhoging wordt de fractie op 0.3 gesteld ( $d/n = 60/200$ ).

De fractie  $d/n$  is van belang voor de reactiesnelheid van de CA-tool op een verandering in het  $PGF$ -verloop. Hoe hoger deze fractie hoe 'actueler' de informatie in  $s(t)$  per controlemoment  $t$ . De fluctuatie in het  $MF'$ -verloop neemt hierdoor wel toe. De vraag is in welke mate het reactievermogen hoger wordt bij een hogere fractie  $d/n$  en welke invloed dit heeft op het  $CP$ .

#### **Wijziging C**

Het verlagen/verhogen van de grenswaarde zal de CA-tool sneller/langzamer doen reageren. De vraag is in welke mate dit invloed heeft op het reactievermogen en de reactiesnelheid van de CA-tool.

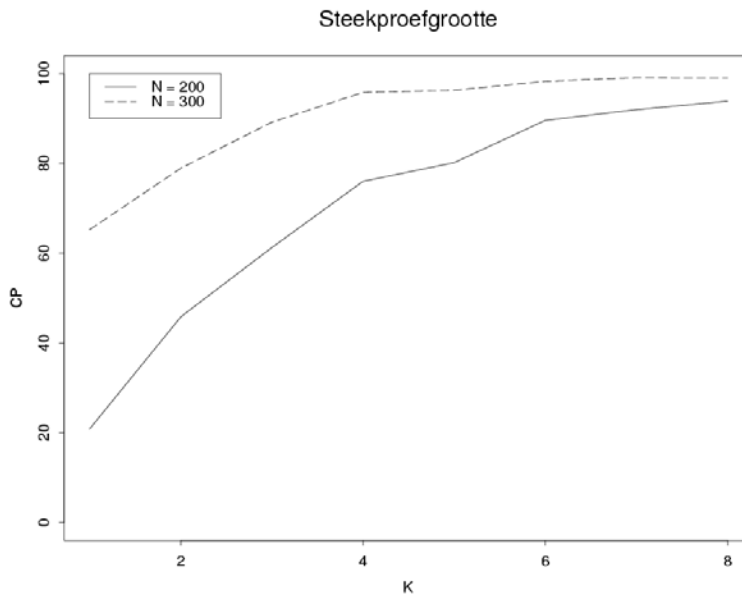
De waarden die gebruikt worden zijn 4%, 5% (basisinstelling) en de 6%.

Het verhogen van de  $C$  (de grenswaarde) laat de CA-tool langzamer maar ook beter presteren, tenminste in geval van scenario 1. Dit trucje zal niet voor alle scenarios gelden. En daarmee komen we gelijk bij een zwak gedeelte van het onderzoek. Onder de gevonden resultaten kunnen instellingen gevonden worden waarmee de CA-tool de twee scenarios (scenario 1 en scenario 2) onderling kan onderscheiden. In praktijk zijn de scenarios echter a) oneindig in aantal en b) grimmiger van aard. Dit onderzoek naar de bruikbaarheid van de CA-tool zal dus geen uitsluitsel kunnen geven over de bruikbaarheid van de CA-tool in praktijk.



### Resultaat N

Het vergroten van de steekproefgrootte van 200 naar 300 onder respectievelijk een  $d$  van 40 en 60 (de  $d/n$  fractie blijft dus hetzelfde) zorgt ervoor dat de  $MF'$  minder sterke fluctuaties kent. In *Figuur 6* ziet men het effect dat door middel van het verhogen van de steekproefgrootte onder scenario 1 en de verschillende  $K$ -waarden is verkregen.

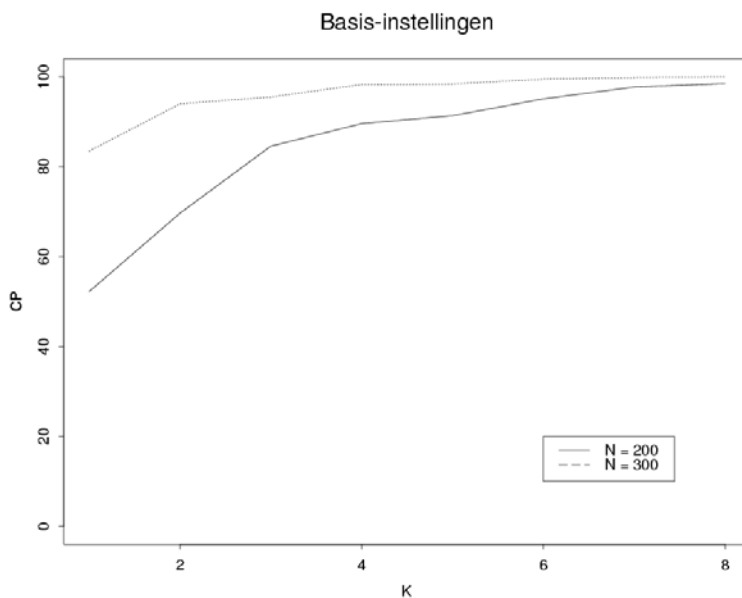


*Figuur 6: het CP dat de CA-tool geeft onder scenario 1 en steekproefgrootten 200 (basisinstelling) en 300, uitgezet tegen de verschillende  $K$ -waarden.*

Het verhogen van de steekproefgrootte resulteert inderdaad in hogere  $CP$ -waarden.

Het criterium wordt nu wel onder scenario 1 gehaald. Bij een  $K = 4$  is het  $CP$  gelijk aan 95.8 %.

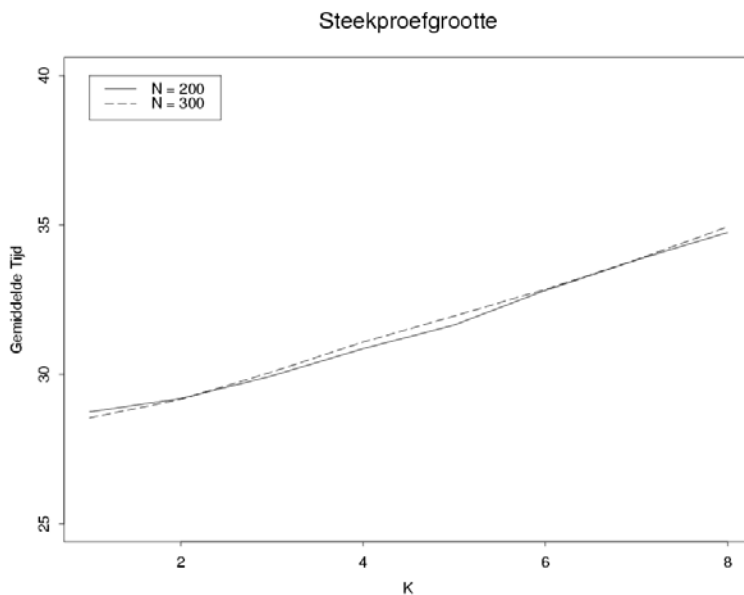
Onder scenario 2 wordt eveneens een sterke verbetering in de  $CP$ -waarden gerealiseerd.



*Figuur 7: het CP dat de CA-tool geeft onder scenario 2 en steekproefgrootten 200 (basisinstelling) en 300, uitgezet tegen de verschillende  $K$ -waarden.*

Een  $K = 3$  blijkt al voldoende te zijn om het criterium te halen. Het  $CP$  is dan gelijk aan 95.5%. In tegenstelling tot de basisinstellingen bestaat hier wel een  $K$ -waarde waaronder de CA-tool een voldoende  $CP$ -waarde haalt. Deze  $K$ -waarde moet dan gelijk zijn aan 4.

Het verhogen van de steekproefgrootten heeft overigens geen invloed op de gemiddelde reactiesnelheid van de CA-tool. Dit kan men in *Figuur 8* duidelijk zien. Hierin is het gemiddelde controlemoment dat de CA-tool onder scenario 2 een correcte uitspraak geeft, uitgezet tegen verschillende  $K$ -waarden:

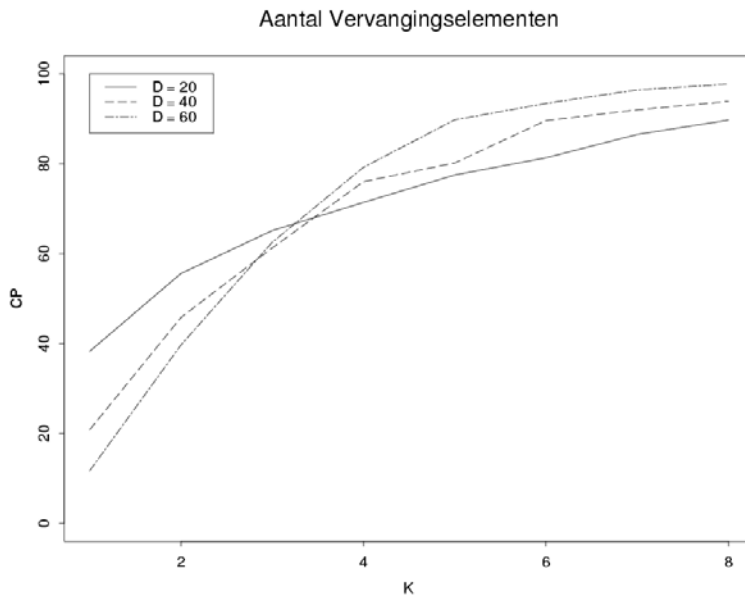


*Figuur 8: het gemiddelde controlemoment dat de CA-tool onder verschillende  $K$ -waarden een juiste uitspraak doet onder scenario 2*

Men kan zien dat het gemiddelde controlemoment waarop de CA-tool bij  $K = 3$  een correcte uitspraak geeft ongeveer 30 is. Ondanks een lage  $K$ -waarde is de reactietijd vrij lang te noemen. Het duurt gemiddeld zo'n 9 controlemomenten vanaf controlemoment 21, dat de CA-tool de uitspraak geeft dat de norm is overschreden. De reactietijd zou wellicht verlaagd kunnen worden door de  $C$  te verlagen. Dit zou echter weer een negatieve invloed kunnen hebben op de  $CP$ -waarden.

**Resultaat D**

De fractie  $d/n$  heeft invloed op de mate waarin de steekproef ‘ververst’ wordt met nieuwe (en meer actuele) steekproefelementen. *Figuur 9* (volgende bladzijde) laat zien welk effect de verschillende  $d/n$  waarden hebben op de  $CP$ -waarden onder scenario 1.



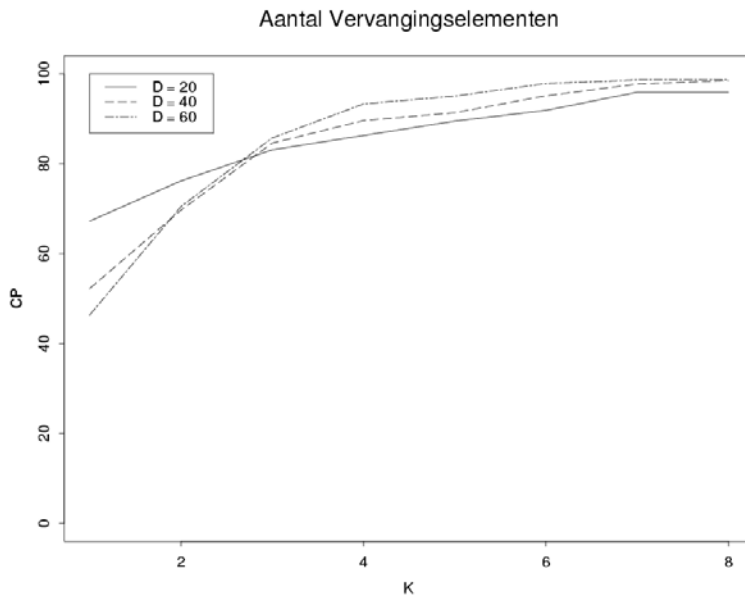
Figuur 9: CP-waarden dat de CA-tool geeft onder scenario 1 en verschillende fracties, uitgezet tegen verschillende K-waarden.

Uit Figuur 9 blijkt dat alleen de  $d/n$  fractie 0.3 het criterium haalt ( $K \geq 7$ ). Onder de  $d/n$  fracties 0.1 en 0.2 worden respectievelijk de waarden 89.7% en 93.9% gehaald bij een  $K = 8$ .

Een lagere  $d/n$  fractie leidt wel tot betere resultaten bij lagere  $K$ -waarden. Doordat de aanvoer van nieuwe steekproefelementen per controlemoment relatief klein is, zal de  $MF'(t)$  maar kleine fluctuaties kennen. De nieuwe elementen hebben dan een kleine invloed op de  $MF'(t)$ . Bij een goede beginschatting ( $MF'(0)$ ) zal dit betere resultaten geven onder scenario 1.

Echter een lage  $d/n$  fractie zorgt er ook voor dat ‘ongelukkige’ elementen (elementen die bijvoorbeeld een te hoge schatting veroorzaken) langer in de steekproef kunnen blijven zitten. Het  $MF'$ -verloop kan dan lang blijven ‘hangen’ in waarden die boven de  $C$  uitkomen. In geval van een  $d/n$  gelijk aan 0.1 bleek in 10.3% van de runs de  $MF'$  minimaal 8 achtereenvolgende controlemomenten boven de 5% te ‘hangen’ terwijl het  $PGF$  toch maar 1% is!

Welk effect verschillende  $d/n$  waarden heeft op de CP-waarden onder scenario 2, kan men zien in Figuur 10 (volgende bladzijde)

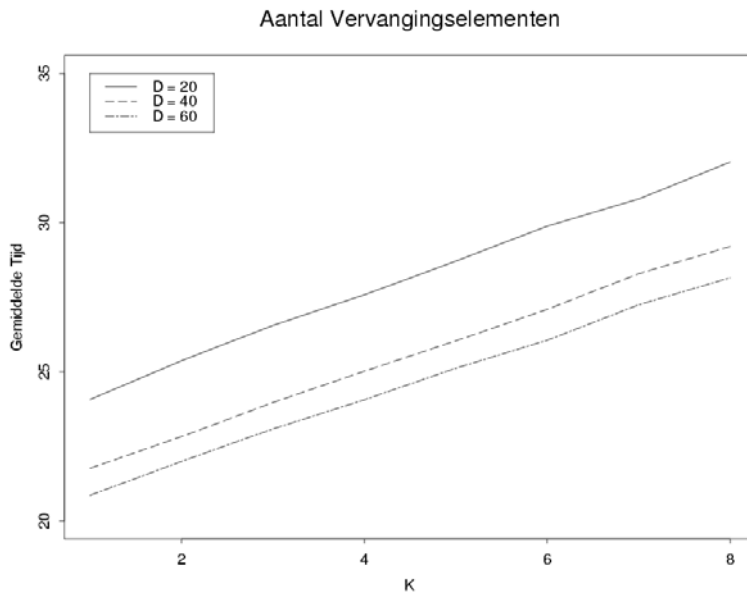


Figuur 10: het CP dat de CA-tool geeft onder scenario 2 en verschillende fracties, uitgezet tegen verschillende K-waarden.

Onder alledrie de  $d/n$  fracties werd het criterium gehaald. Onder de  $d/n$  fracties 0.1 0.2 en 0.3 worden respectievelijk de waarden 95.7% ( $K = 7$ ), 95.1% ( $K = 6$ ) (basisinstelling) en 95.0% ( $K = 5$ ) gehaald. Zoals eerder geschreven zijn de verbeterde resultaten ten opzichte van scenario 1 te danken aan de extreme ophoging in het  $PGF$ -verloop waardoor de CA-tool alleen een foute uitspraak kan geven op de eerste 21 controlemomenten ( $t = 0, 1, \dots, 20$ ). Voor het verloop van de CP-waarden geldt in wezen hetzelfde verhaal als bij scenario 1. Hoe lager de  $d/n$  fractie hoe hoger het CP bij een lage  $K$ . Bij hogere  $K$ -waarden presteert een hoge  $d/n$  fractie juist beter.

De  $K$ -waarde waaronder de CA-tool bij scenario 1 en scenario 2 goed reageert, is niet laag. Deze blijkt minimaal gelijk aan 7 te moeten zijn.

De invloed van de verschillende  $d/n$  fractie op de gemiddelde tijd dat de CA-tool reageert in scenario 2, staat in *Figuur 11* (zie volgende bladzijde).



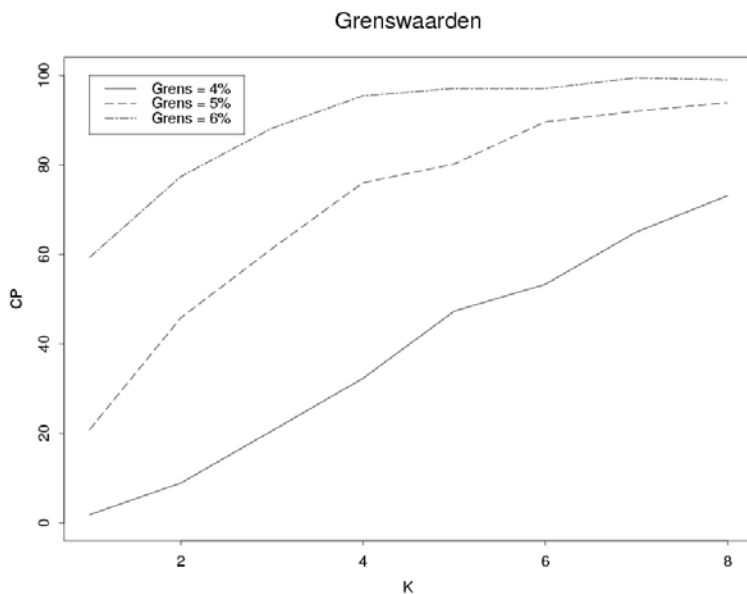
Figuur 11: het gemiddelde controlemoment dat de CA-tool onder verschillende K-waarden een juiste uitspraak doet onder scenario 2

Uit *Figuur 11* blijkt dat het heel weinig uitmaakt of men een  $d/n$  fractie van 0.2 of 0.3 gebruikt. Een  $d/n = 0.1$  geeft weer een grote ophoging op het gemiddelde controlemoment dat de CA-tool een correcte uitspraak geeft.

### Resultaat C

Het verhogen van de  $C$  doet de CA-tool minder snel reageren op veranderingen in het  $PGF$ -verloop. Dit komt zeker ten goede komen aan de  $CP$ -waarden in scenario 1. Een verlaging van de  $C$  geeft de CA-tool de gelegenheid om sneller op veranderingen te reageren, hetgeen van toepassing is bij scenario 2.

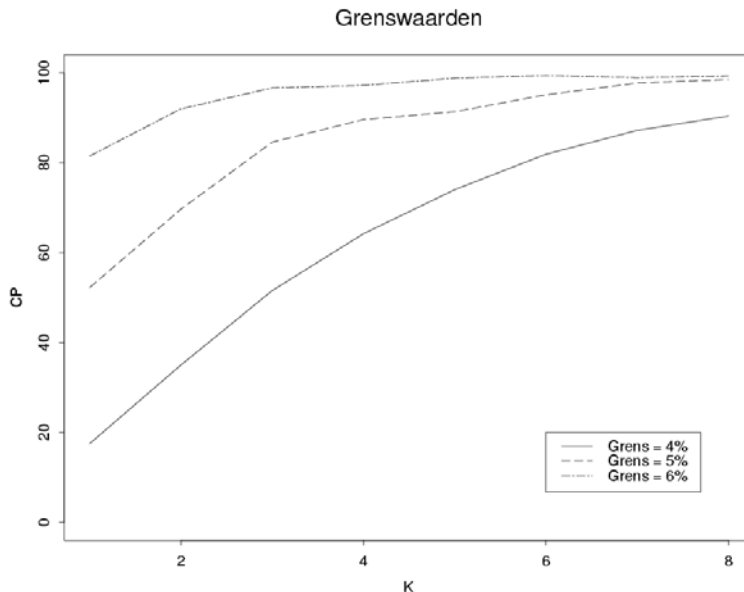
Het effect van drie verschillende grenswaarden (4, 5 en 6%) op het  $CP$  van de CA-tool staan in *Figuur 12* en *Figuur 13*.



Figuur 12: het CP dat de CA-tool geeft onder scenario 1 en verschillende fracties, uitgezet tegen verschillende K-waarden.

Het stellen van de  $C$  op 6% geeft onder scenario 1 een grote verbetering in de  $CP$ -waarden. Bij een  $K = 4$  wordt een  $CP$  gehaald van 95.5 %. Ter vergelijking, onder de basisinstellingen ( $C = 5\%$ ) werd slechts 76.0 % gehaald bij dezelfde  $K$ -waarde. Deze verbetering is natuurlijk niet zo vreemd. De  $CA$ -tool doet een verhoging van de  $C$  minder snel de uitspraak dat de norm is overschreden. Het verhogen van de  $C$  geeft automatisch hogere  $CP$ -waarden. Een  $C$  van 4% zal de  $CP$ -waarde flink doen dalen, de  $MF'$  lijn bevindt zich blijkbaar in een groot aantal runs langdurig boven de 4%.

Bij scenario 2 leidt het verhogen van de  $C$  niet automatisch tot hogere  $CP$ -waarden. Een hogere  $C$  zal de  $CA$ -tool minder snel laten reageren op een verandering in het  $PGF$ . Het gevaar bestaat dus dat de  $CA$ -tool een overschrijding van de norm door het  $PGF$  die in scenario 2 plaatsvindt, niet opmerkt.



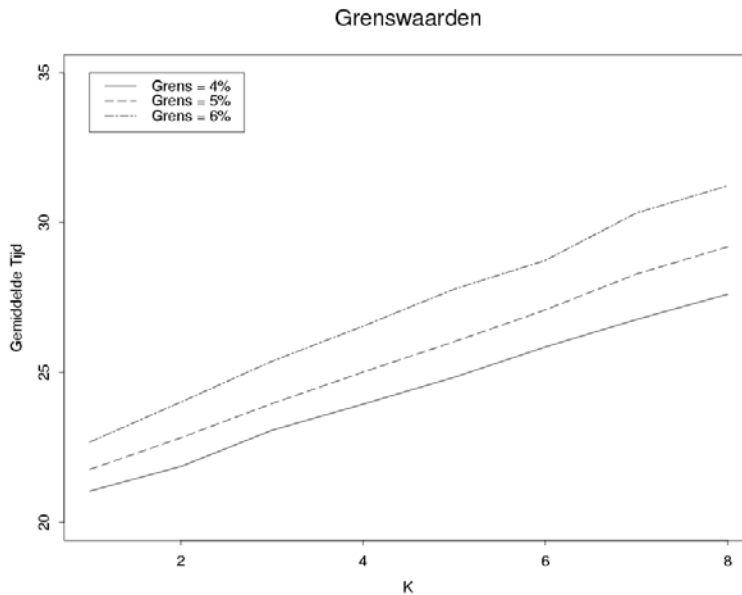
Figuur 13: het  $CP$  dat de  $CA$ -tool geeft onder scenario 2 en verschillende fracties, uitgezet tegen verschillende  $K$ -waarden.

Toch blijkt dat de  $C$  van 6% ook bij scenario 2 betere  $CP$ -waarden oplevert. Wel vertoont de  $CA$ -tool bij hogere  $K$ -waarden een stijgend aantal runs waarin de overschrijding niet wordt opgemerkt. Net als bij scenario 1 zal een verlaging van de  $C$  naar 4% tot slechtere  $CP$ -waarde leiden. Omdat een lagere  $C$  waarde bij scenario 2 alleen tot controlemoment 20 in het nadeel werkt,  $PGF$  zit met de 1% dan nog ruim onder de norm, in plaats van de 51 controlemomenten bij scenario 1, zijn de  $CP$ -waarden beter dan bij scenario 1. Een  $K$ -waarde van 3 blijkt bij scenario 2 al voldoende te zijn. Deze  $K$ -waarde levert een  $CP$ -waarde van 96.6%.

Het lijkt erop dat het verhogen van de  $C$  hetzelfde effect heeft als het verhogen van de steekproefgrootte. Bij beide wijzigingen vertonen de  $CP$ -waarden onder verschillende  $K$  veel hogere waarden. Dit komt omdat in beide gevallen het  $MF'$ -verloop minder ernstig wordt. Het verhogen van de steekproefgrootte zal de fluctuaties in het  $MF'$ -verloop absoluut verminderen. Het verhogen van de grenswaarde zal daarentegen de invloed van een fluctuatie verminderen. In dit geval levert dit ongeveer dezelfde grafieken en resultaten op.

Net als bij het verhogen van de steekproefgrootte is ook hier een  $K$  gelijk aan 4 voldoende om bij beide scenario's minimaal een  $CP$  van 95% te halen.

De invloed van de grenswaarden op de gemiddelde reactiesnelheid staat in *Figuur 14*.



*Figuur 14: het gemiddelde controlemoment dat de CA-tool onder verschillende K-waarden een juiste uitspraak doet onder scenario 2*

Het verlagen van de grenswaarde geeft een licht voordeel in de reactiesnelheid. Het verschil in tijd wordt groter naarmate de  $K$  toeneemt. Bij een  $K = 8$  zal bij een  $C = 4\%$  gemiddeld drie controlemomenten sneller worden gereageerd dan bij  $C = 6\%$ . Bij een  $K = 1$  is dit nog ongeveer gelijk aan twee controlemomenten.

### 9.3.3 Samenvatting Resultaten

Uit de resultaten bleek dat de  $MF'$  onder de basisinstellingen grote fluctuaties liet zien en bovendien langdurig kon blijven hangen in waarden die (relatief) ver afwaken van het werkelijke percentage fout in de voorraad. Om een  $CP$  van minimaal 95% te halen voor beide scenario's, zou de CA-tool een hoge  $K$ -waarde ( $>8$ ) moeten gebruiken. Een hoge  $K$ -waarde is echter niet gunstig omdat dit de CA-tool minder snel laat reageren op een verandering. En dat devalueert indirect weer het continuous-aspect in de CA-tool. Het aantal benodigde controlemomenten kon verlaagd worden door bijvoorbeeld het aantal steekproefelementen,  $n$ , te verhogen. De  $MF'$  op controlemoment 0 vormt dan een goede beginschatting van de fout. Bovendien is een grote steekproef stabielier aangezien ongelukkige steekproefelementen minder invloed hebben. De fractie  $d/n$  heeft eveneens invloed op de  $CP$ -waarden. Het verhogen van de  $d/n$  fractie zorgde ervoor dat elementen sneller 'ververst' werden. Dit kan zeker handig zijn wanneer relatief veel 'ongelukkige' artikelen in de voorraad zitten. Het verhogen van de  $C$  leidde onder beide scenario's eveneens tot betere  $CP$ -waarden.

### 9.3.4 Kritische noot

Ten aanzien van de CA-tool en het onderzoek zijn enkele kritische opmerkingen te maken. De CA-tool maakt gebruik van vaste grens waartegen de  $MF'(t)$  getoetst wordt. In de praktijk kan het stellen van een vaste grens alleen nuttig zijn als men van een beperkt aantal scenario's uitgaat waarvan men het verloop kent. Bovendien kunnen/zullen de instellingen van de CA-tool om verscheidene scenario's van elkaar te kunnen onderscheiden en herkennen alleen maar extremer worden ten opzichte van de gevonden instellingen onder de twee scenario's. En dit zal de CA-tool bedrijfseconomisch alleen maar minder interessant maken. Verder is er gebruik gemaakt van een voorraad die gemodelleerd was naar een voorbeeldbestand. Dit leverde een specifieke voorraad op. In praktijk kunnen voorraden sterk afwijken van deze voorraad. De prestaties van de CA-tool is mede afhankelijk van de voorraadinstellingen en de resultaten geven dus alleen een beeld onder de gebruikte voorraad.

Een ander punt betreft de verandering in het *PGF*-verloop. Een verandering werd bewerkstelligd door het aantal subposten te veranderen dat een fout bevatte. In werkelijkheid zou de fractie fout in een subpost ook kunnen veranderen. De invloed van een veranderende fractie op de prestatie van de *CA*-tool is verder niet onderzocht.



## 10 Conclusies

In *Onderzoeksvraag I* is onderzocht of een veranderende massa bij een geldsteekproef een probleem vormt voor het ontwikkelen van een CA-tool. Dit bleek niet het geval te zijn. Door de massa te schatten op verschillende controlemomenten, kon een bestaande geldsteekproefmethode makkelijk worden 'getransformeerd' naar een CA-toepassing. Het probleem van de veranderende massa kon hiermee worden omzeild. Maar door gebruik te maken van schattingen wordt ook de steekproefgrootte voor de hele controle periode vastgezet. Blijkt de werkelijke massa op een controlemoment kleiner te zijn dan de geschatte massa, dan wordt op hetzelfde moment de werkelijke foutfractie groter dan de kritieke foutfractie. Tot op dat moment zijn er dus in feite te weinig elementen getrokken. Het besluit om vanaf dit punt extra steekproefelementen te trekken heeft geen zin aangezien dit aan het eind een ongebalanceerde steekproef zou veroorzaken. Om dit probleem te voorkomen kan men aan het begin een lagere kritieke foutfractie vaststellen dan normaal wordt gehanteerd (5%). Het aantal steekproefelementen wordt hierdoor hoger. Blijkt dit aantal later te weinig blijken dan kan men het altijd nog toetsen tegen de normaal gehanteerde foutfractie. Men moet zich wel realiseren dat een tussentijds controlemoment, een momentopname is. De werkelijke massa kan altijd nog naar de geschatte massa toe groeien. En uiteindelijk gaat het om de eindmassa. Kortom de CA-tool is alleen geschikt voor het verkrijgen van een *indicatie* over een controlemoment.

In *Onderzoeksvraag II* werd getracht een antwoord te vinden op de vraag of de ontwikkelde CA-tool een *goede* CA-tool was. In het onderzoek dat hierop volgde werd de CA-tool onder een aantal instellingen getoetst tegen twee simpele scenario's. In eerste instantie was het onderzoek erop gericht om de instellingen te vinden waaronder de CA-tool met een grote zekerheid (95%) beide scenario's kon 'herkennen'. Uitgaande van deze instellingen kon worden ingeschat, of de CA-tool ook in praktijk te gebruiken is. Uit de resultaten bleek dat de CA-tool onder de basisinstelling zeker niet voldoet. Het 'continuous'-aspect in de CA-tool kwam onder grote druk te staan. Het aantal controlemomenten nodig om onderscheid te maken tussen de twee scenario's was onacceptabel groot. Weliswaar was het mogelijk middels aanpassingen in de instellingen van de CA-tool dit aantal benodigde controlemomenten te verlagen maar dergelijke aanpassingen zijn echter bedrijfseconomisch niet interessant aangezien ze een beduidend meer inspanning zullen kosten. Omdat het onderzoek net als de CA-tool zelf een aantal gebreken kende, was het onmogelijk om uitsluitel te geven op de vraag of de CA-tool een goede CA-tool was. Het onderzoek heeft hoogstens inzicht kunnen verschaffen in de CA-tool.

Ondanks genoemde beperkingen bevat de onderzochte CA-tool wel degelijk een indicatie van het foutverloop in de voorraad. Een grote verandering in de fout in de voorraad zal de CA-tool onder (bedrijfseconomisch) interessante instellingen gegarandeerd worden opgemerkt.



# 11 Notatielijst

$b$	= de regressiecoëfficiënt van de regressieschatter
$C$	= grenswaarde waartegen de $MF'$ -verloop getoetst wordt
$d$	= aantal steekproefelementen dat per controlemoment ( $> 0$ ) vervangen wordt door steekproefelementen van de nieuwe massa
$K$	= aantal achtereenvolgende controlemomenten dat het $MF'$ boven de $C$ moet verblijven wil de CA-tool de uitspraak doen dat de norm is overschreden door het $PGF$
<i>materialiteit</i>	= de gesaldeerde fout in de voorraad dat niet overschreden mag worden
$MF$	= maximale fout
$MF'(t)$	= de $MF$ op controlemoment $t$ , berekend m.b.v. de regressieschatter-plus
$M(t)$	= totale boekwaarde op controlemoment $t$
$n$	= aantal steekproefelementen
$N$	= totaal aantal subposten in een voorraad
$PF$	= geprojecteerde fout
$PF'(t)$	= de $PF$ op controlemoment $t$ , berekend m.b.v. de regressieschatter-plus
$PGF$	= percentage gesaldeerde fout
$r_{subpost:i}$	= residuwaarde van subpost $i$
$s(t)$	= aanduiding voor steekproef op controlemoment $t$
$T$	= totaal aantal controlemomenten in een controleperiode
$VM$	= vertalingsmechanisme
$WF$	= waarschijnlijke fout
$x_i$	= de auditwaarde van steekproefelement $i$
$x_{subpost:i}$	= de auditwaarde van subpost $i$
$y_i$	= boekwaarde van steekproefelement $i$
$y_{subpost:i}$	= boekwaarde van subpost $i$



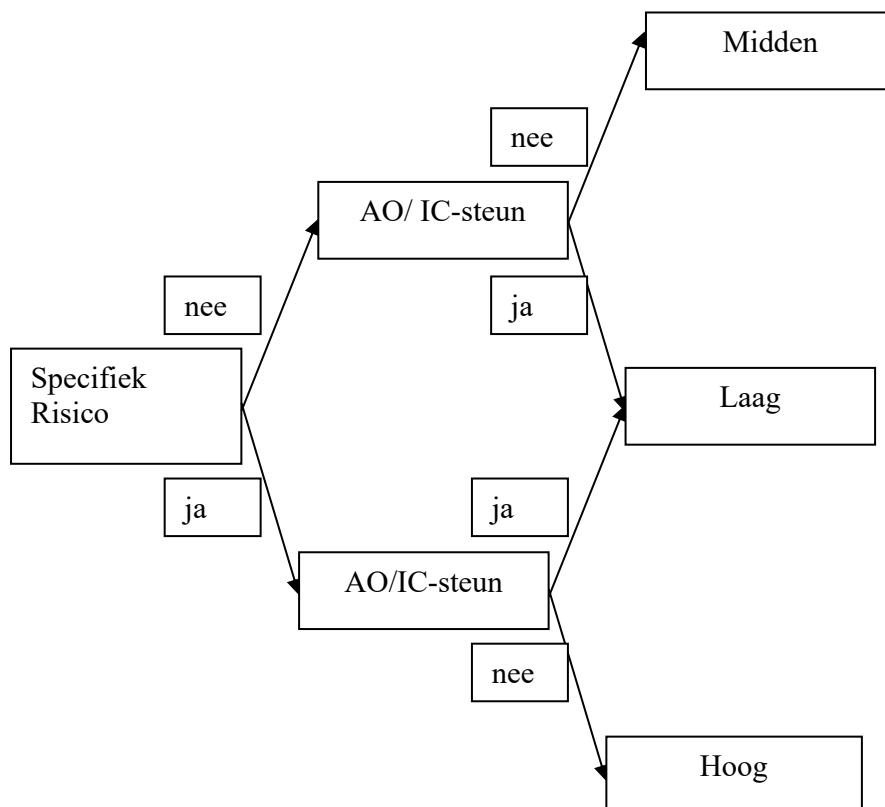
# 12 Appendix A

## 12.1 ARM

Het ARM maakt gebruik van de a priori gegevens omtrent het AO/IC. De mate waarop er gesteund wordt op het AO/IC beïnvloedt de  $x$  en daarmee ook direct de waarde voor  $R(\theta, x)$ . Dit heeft dus gevolgen voor de steekproefgrootte ( $n$ ). Door te steunen op het AO/IC wordt de  $R(\theta, x)$  waarde lager. Accountants hebben de keuze om op verschillende controle-niveaus te gaan werken. Hoe hoger het niveau hoe uitgebreider (hoe groter de steekproefgrootte) er gecontroleerd zal worden. De betreffende niveaus zijn:

1. Hoog niveau: er is geen sprake van een verlaagd risico, hoewel fouten niet aanwezig worden verondersteld, en men wil of kan niet steunen op de Interne Controle maatregelen die dat risico zouden kunnen indekken.
2. Midden niveau: er is een verlaagd risico onderkend maar er wordt niet gesteund op de Administratieve Organisatie.
3. Laag niveau: deze wordt gebruikt bij de volgende twee situaties:
  - Er is geen verlaagd risico onderkend en er wordt gesteund op de IC-maatregelen (goede werking is aangetoond door organisatie gerichte controles, de zogenaamde procescontroles) die het risico kunnen dekken
  - Er is een verlaagd risico onderkend en men steunt op de betrouwbaarheid van de AO (door organisatie besloten)

Dit kan keurig in een graaf worden weergegeven, zie volgende bladzijde.



In het ARM wordt de keuze uit één van de drie niveaus doorgevoerd naar de keuze voor het getal  $R(0,x)$ . Deze keuze heeft invloed op de steekproefgrootte  $n$ . De volgende gelijkheid geldt immers:

$$(23) \quad R(k) = n \cdot p_m = n \cdot \frac{MP}{M}$$

De keuze voor de  $R(0,x)$  waarde wordt in grote mate subjectief bepaald. In eerste instantie werd voor laag, midden en hoog niveau respectievelijk de  $R(0)$  waarden 1, 2 en 3 gekozen. Omdat men de waarde 1 nog te hoog vond voor het lage niveau, is deze waarde verlaagd naar 0.7, de laagste waarde waarbij men nog iets zinnigs kan zeggen over de populatie. Een waarde kleiner dan 0.7 impliceert dat bij aanwezigheid van een materiele fout in de populatie, deze fout met minder dan 50% kans zou worden opgemerkt. De kans om de fout te ontdekken wordt dan kleiner dan dat de fout onopgemerkt blijft.

Hieronder een tabel waarin de  $R(0)$  waarden overzichtelijk gekoppeld zijn aan de graaf:

Tabel 10: waarden  $R(0)$  bij verschillende situaties

$\alpha = 0.05$	Steun AO/IC	Geen Steun AO/IC
Specifiek Risico	Laag: $R(0) = 0.70$	Hoog: $R(0) = 3.00$
Geen Specifiek Risico	Laag: $R(0) = 0.70$	Midden: $R(0) = 2.00$

De  $R(0)$  waarde kan ook bepaald worden via de volgende formule:

$$AR = IR \cdot ICR \cdot DR, \text{ waarbij}$$

- AR = accountantscontrole risico, de kans dat een materiële fout niet ontdekt wordt.
- IR = inherent risico, de kans dat er een materiële fout is opgetreden.
- ICR = intern controle risico, de kans dat de interne controle een materiele fout niet ontdekt

DR = detectierisico, de kans dat de accountant bij zijn gegevens gerichte werkzaamheden een materiële fout niet ontdekt.

De AR heeft in principe de vaste waarde 0.05, 5% kans dat een materiële fout bestaat, niet door de klant zelf wordt gevonden, en door controle niet ontdekt wordt.

De waarde voor ICR wordt in het geval van geen steun AO/IC op 1 worden gesteld, de IC zal zeker geen materiële fout kunnen ontdekken. De IR wordt op 1 gesteld wanneer er een specifiek risico is onderkend. De DR waarde kan bij een gegeven  $R(0)$  waarde als volgt worden berekend ( $DR = -\ln(1-x)$ ):

$$(24) \quad R(0) = -\ln(1-x) \stackrel{R(0)=w}{\Rightarrow} w = -\ln(1-x) \Rightarrow e^{-w} + 1 = x$$

De waarden voor de IC, ICR en DR kunnen in een tabel worden gezet:

Tabel 11

	Steun AO/IC $IR \cdot ICR \cdot DR$	Geen Steun AO/IC $IR \cdot ICR \cdot DR$
Specifiek Risico	100% · 10% · 50%	100% · 100% · 5%
Geen Specifiek Risico	37% · 28% · 50%	37% · 100% · 14%

Deze methode<sup>6</sup> voor het bepalen van de steekproefgrootte wordt niet door Deloitte in Nederland geadviseerd omdat de subjectief gekozen waarden niet gevalideerd kunnen worden.

## 12.2 BRAM

Het BRAM (Bayesiaanse Risico Analyse Model)<sup>7</sup> ondervangt het probleem van het maken van subjectieve waarden zonder deze te kunnen controleren. Met behulp van Bayesiaanse statistiek kan op toelaatbare wijze achterhaald worden of de gemaakte subjectieve keuzes inderdaad juist waren. Voor het bepalen van het de steekproefgrootte:

- Bepaal de steekproefomvang met de volgende formule:  $n = 3 \frac{M}{MP} - 3 \frac{M}{MP^*}$ . De  $MP^*$  is het bedrag waar de accountant zeker van is dat dit niet wordt overschreden. De formule herbergt een aantal karakteristieken, te weten:
  - Als  $MP^* = MP$  zal het aantal steekproeven gelijk aan nul zijn. Men is immers zeker dat de nauwkeurigheidseis is behaald.
  - De aanname  $MP^* = M$  leidt tot een afname van 3 in het steekproefaantal. Dit komt omdat de bovengenoemde formule een Poisson-benadering van de werkelijkheid is.
- Bepaal het controle-materialiteitsbedrag waar de accountant zeker van is dat deze niet wordt overschreden.

De waarde voor  $R(0)$  wordt gekoppeld aan een waarde voor  $MP^*$ . Dit  $MP^*$ -bedrag kan getoetst worden. Hiervoor is geen integraal onderzoek nodig, men volstaat met de steekproef op hoog niveau en het uitgaan van nul fouten. De  $MP^*$  wordt dan gelijk aan  $M$ ; er kan immers niet meer fout zijn dan de hele populatieomvang.

## 12.3 Evaluatiemethoden geldsteekproeven

Een geldeenheid in de geldsteekproef kan deel uitmaken van een foute transactiepost waarin de auditwaarde niet overeenkomt met de boekwaarde. In de vorige evaluatiemethoden werd de geldeenheid dan gelijk als één hele foute geldeenheid beschouwd en werd de waarde voor  $k$  met één opgehoogd. Een foute transactie is echter

<sup>6</sup> Steele

<sup>7</sup> Deze methode is ontwikkeld door P.C. van Batenburg.

vaak niet in zijn geheel fout. De boekwaarde bevat vaak een deel dat terecht geboekt is. Om de consequenties van zo'n gedeeltelijke foute transactie niet te overschatten, zijn er evaluatietechnieken bedacht die dit probleem enigszins ondervangen. Twee van zulke technieken zijn:

- De Goed-of-Fout methode
- De Stringerbound

Beide methoden gaan in op de manier van verdelen van de geldeenheden (naar foute en goede geldeenheden) in de foute transactie wanneer er via een trekking een foute transactie aan het licht is gekomen.

### 12.3.1 Goed-of-Fout-methode

De Goed-of-Fout-methode werkt als volgt:

1. Er wordt een geldeeenheid aselect getrokken
2. Als deze geldeeenheid bij een incorrecte post blijkt te horen, wordt deze post in twee delen opgesplitst, namelijk in een 'goed'-gelddeel en een 'fout'-gelddeel. Deze verdeling ligt al van te voren vast. Een verdelingsvorm kan bijvoorbeeld zijn dat het foute deel het eerste gedeelte van de transactie-geldeenheden vormt en het goede deel het tweede gedeelte.
3. Er wordt gekeken tot welk deel de getrokken geldeeenheid behoorde. Ligt deze in het goede deel dan wordt deze geldeeenheid als goed bestempeld, is echter het andere deel het geval dan wordt het als fout bestempeld. Wordt er een foute geldeeenheid gevonden dan zal de absolute maximale fout die in de populatie aanwezig zal zijn, gelijk worden aan  $4.75 \cdot \frac{M}{n}$ . Er geldt namelijk:  $MF = R(1) \cdot \frac{M}{n}$ . Wordt er geen foute geldeeenheid gevonden dan zal de  $MF$  gelijk worden aan  $3.00 \cdot \frac{M}{n}$ .

Het nadeel van deze methode is, is dat men bij een getrokken 'goede' geldeeenheid uit een foutieve transactiepost, de transactiepost alsnog goed gaat keuren. Dit zou betekenen dat het mogelijk is om een bekende fout ongelijk nul en een geprojecteerde fout van nul te hebben wanneer er wel foutieve transactieposten zijn aangetroffen.

### 12.3.2 Stringerbound

Door gebruik te maken van de zogenaamde Stringerbounds<sup>8</sup> wordt dit probleem ondervangen. Deze methode wordt ook wel de tainting-methode genoemd. In plaats van het opdelen van een post in goede en foute guldens wordt elke gulden 'besmet' met het foutpercentage in de betreffende transactiepost. Het foutpercentage ( $fp$ ) wordt als volgt gedefinieerd:

$$fp = \frac{\text{boekwaarde} - \text{auditwaarde}}{\text{boekwaarde}} \cdot 100\%$$

Het foutpercentage zal alleen een waarde tussen 0% en 100% aannemen. Er wordt namelijk alleen gekeken of alle geboekte geldeenheden juist zijn geboekt. Dit is altijd het geval bij overwaardering. Het feit dat er te weinig is geboekt is niet aan de orde. Het foutpercentage wordt dan ook op 0% gezet bij onderwaardering.

Het bepalen van de Stringerbounds gebeurt via de volgende stappen:

- Bepaal het aantal foutieve geldeenheden die in de steekproef gevonden zijn
- Bepaal per foute geldeeenheid de mate van aantasting.

---

<sup>8</sup> Verbaal geïntroduceerd door Stringer (1963), in formulevorm opgeschreven door Leslie, Teitlebaum en Andersen (1979).



- Bepaal per foute geldeenheid de bovengrens voor de foutfractie. Deze bovengrenzen en de ophoging per fout staan in *Tabel 12* (betrouwbaarheid = 95%):

*Tabel 12: voor verschillende waarde k, de  $R(k,0.95)$  en de  $\Delta(R(k,0.95) - R(k-1,0.95))$*

<b>K</b>	<b>R(k,0.95)</b>	<b><math>\Delta(R(k,0.95) - R(k-1,0.95))</math></b>
0	3.00	-
1	4.75	1.75
2	6.30	1.55
3	7.76	1.46
4	9.16	1.40

De Stringerbounds worden nu bepaald via de volgende formule:

$$(25) \quad SB = \{R(0) + t_1[R(1) - R(0)] + \dots + t_k[R(k) - R(k-1)]\} \cdot \frac{M}{n}$$

waarbij  $t_i$  de i-grootste tainting is. Uit deze formule blijkt dat de ophoging van de fout aan de hand van het taintingspercentage gebeurt. De R wordt dus in principe niet maximaal opgehoogd zoals dat het geval was bij de Goed-Fout-Methode bij een toegewezen foute geldeenheid. De volgorde van de tainting gebeurt in het geval van de Stringerbounds van hoog naar laag qua taintingpercentage. Zo verkrijgt men de meest conservatieve schatting, de hoogst mogelijke schatting van de maximale fout.

Ter illustratie, stel er zijn vier foute geldeenheden gevonden:

*Tabel 13*

<b>Foutno.</b>	<b>Boekwaarde</b>	<b>Auditwaarde</b>	<b><math>t_i</math></b>
1	1000	300	30%
2	100	25	25%
3	50	10	20%
4	10	5	50%

De Stringerboundwaarde wordt dan:

$$SB = \{3.00 + 50\% \cdot 1.75 + 30\% \cdot 1.55 + 25\% \cdot 1.46 + 20\% \cdot 1.40\} \cdot \frac{M}{n} = 4.985$$

Bij de EVA-methode<sup>9</sup> wordt deze grens aangescherpt door niet op de grootte van het taintingspercentage te letten maar op de absolute fout in een transactiepost. De formule blijft gehanteerd, alleen de volgorde van de tainting kan anders zijn. De  $t_i$  stelt niet meer de i-grootste tainting voor maar de i-grootste absolute fout. Bij gebruik van *Tabel 8* wordt de EVA- boundwaarde:

$$EB = \{3.00 + 30\% \cdot 1.75 + 25\% \cdot 1.55 + 20\% \cdot 1.46 + 50\% \cdot 1.40\} \cdot \frac{M}{n} = 4.905$$

Zoals men kan zien levert dit een scherpere grens.

---

<sup>9</sup> Intern ontwikkeld door Deloitte & Touche



## 13 Appendix B

Het voorraadbestand bestond uit de volgende boek- en auditwaarden:

Element no.	Boekwaarde	Auditwaarde	Element no.	Boekwaarde	Auditwaarde	Element no.	Boekwaarde	Auditwaarde
1	50,88	111,83	68	550,2	550,2	135	15,32	15,32
2	4,5	9,25	69	3,74	3,74	136	30,45	30,45
3	1937,48	968,74	70	61,39	61,39	137	58,87	58,87
4	54,39	31,08	71	11,1	11,1	138	26,28	26,28
5	7,06	4,06	72	145,92	145,92	139	1,72	1,72
6	8,74	11,4	73	29,97	29,97	140	91,2	91,2
7	244,08	284,76	74	31,61	31,61	141	8,14	8,14
8	1088,1	1263,6	75	1,62	1,62	142	62,9	62,9
9	14,04	11,88	76	21,93	21,93	143	4461,4	4461,4
10	86,71	98,67	77	46,98	46,98	144	2,2	2,2
11	117,52	102,83	78	368,93	368,93	145	51,31	51,31
12	71,19	63,28	79	46,32	46,32	146	8671,94	8671,94
13	25,47	22,64	80	2,48	2,48	147	72,6	72,6
14	87,22	78,32	81	83,45	83,45	148	1,95	1,95
15	25,99	23,46	82	4,71	4,71	149	0,98	0,98
16	104,83	114,36	83	195	195	150	13,91	13,91
17	16,96	15,52	84	1025,36	1025,36	151	0,25	0,25
18	1,2	1,1	85	249,48	249,48	152	266,98	266,98
19	6,76	7,28	86	37,7	37,7	153	127,34	127,34
20	12,9	12,04	87	57,96	57,96	154	45,54	45,54
21	78,72	73,8	88	1,13	1,13	155	19,5	19,5
22	87,78	92,4	89	8,6	8,6	156	0,5	0,5
23	578,6	549,67	90	3026,52	3026,52	157	876,06	876,06
24	625,8	594,51	91	135,96	135,96	158	53,46	53,46
25	25,22	24,25	92	21,01	21,01	159	5,14	5,14
26	585,75	569,25	93	10,4	10,4	160	13,87	13,87
27	502,55	514,05	94	140,72	140,72	161	12,06	12,06
28	286,88	280,36	95	36,9	36,9	162	92,89	92,89
29	213,3	218,04	96	419,84	419,84	163	68,31	68,31
30	113	110,74	97	23,22	23,22	164	10,08	10,08
31	270,54	275,55	98	1197,6	1197,6	165	3,08	3,08
32	51,23	52,17	99	107,16	107,16	166	8,48	8,48
33	135,52	137,94	100	29,92	29,92	167	0,34	0,34
34	15,73	15,47	101	1117,35	1117,35	168	7,4	7,4
35	473,2	479,96	102	347,69	347,69	169	15,77	15,77
36	160,5	162,64	103	49,7	49,7	170	0,5	0,5
37	664,4	671,95	104	2,07	2,07	171	122,96	122,96
38	265,68	262,98	105	254,28	254,28	172	1,54	1,54
39	1572	1556,28	106	524,79	524,79	173	1,38	1,38
40	222,74	224,96	107	10,5	10,5	174	10,8	10,8
41	237,6	239,4	108	18,15	18,15	175	273,6	273,6
42	83,4	82,8	109	80,43	80,43	176	0,65	0,65
43	42,84	43,12	110	47,4	47,4	177	41,48	41,48
44	378,42	380,8	111	0,56	0,56	178	7,14	7,14
45	112,5	113,1	112	88,66	88,66	179	6,48	6,48
46	51,12	51,3	113	15,39	15,39	180	7,29	7,29
47	1322,88	1326	114	58,56	58,56	181	11,7	11,7
48	11,47	11,47	115	26,2	26,2	182	1,26	1,26
49	14,95	14,95	116	144,32	144,32	183	134,06	134,06
50	8,41	8,41	117	7,74	7,74	184	4,56	4,56
51	218,46	218,46	118	885,92	885,92	185	18,62	18,62
52	2,84	2,84	119	260	260	186	48,72	48,72
53	32,55	32,55	120	48,36	48,36	187	22,76	22,76
54	49,77	49,77	121	277,2	277,2	188	1,59	1,59
55	5,48	5,48	122	43,89	43,89	189	12,24	12,24
56	3,38	3,38	123	967,64	967,64	190	1,32	1,32
57	1	1	124	4,98	4,98	191	2,06	2,06
58	4,67	4,67	125	7,93	7,93	192	21,69	21,69
59	8,82	8,82	126	6,6	6,6	193	214,32	214,32
60	2,16	2,16	127	1656	1656	194	137,64	137,64
61	2,08	2,08	128	47,08	47,08	195	259,82	259,82
62	3,91	3,91	129	0,48	0,48	196	1,06	1,06
63	1,28	1,28	130	66,82	66,82	197	189,24	189,24
64	2,24	2,24	131	31,78	31,78	198	1,92	1,92
65	23,6	23,6	132	1010,24	1010,24	199	298	298
66	104,25	104,25	133	4,5	4,5	200	0,13	0,13
67	32,21	32,21	134	32,22	32,22	201	85,92	85,92



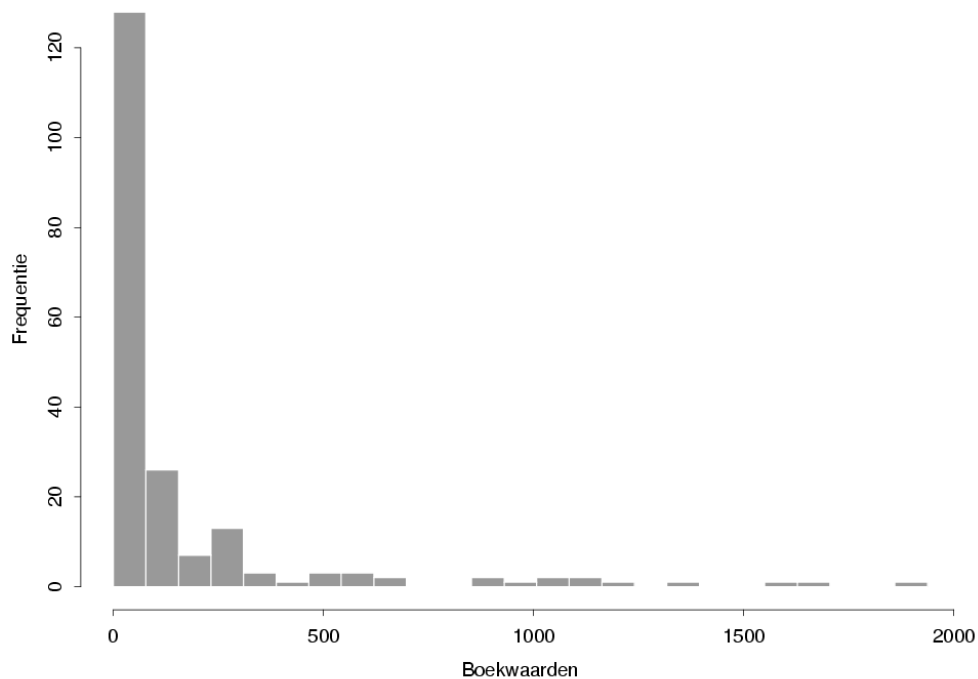
## 14 Appendix C

De analyse en de modellen die zijn afgeleid van deze analyse staan hieronder beschreven in *Model Boekwaarden* en *Model Residuwaarden*.

### **Model Boekwaarden**

Het voorraadbestand bestond uit 201 subposten. *Figuur 15* is een histogram van de 198 hoogste boekwaarden. De drie hoogste waarden zijn in het histogram buiten beschouwing gelaten omdat ze extreem hoog waren. Het histogram zou daardoor bij het meenemen van deze waarden onvoldoende beeld geven van de kleinere waarden die het overgrote deel uitmaken van het totaal aantal steekproefelementen.

Histogram Boekwaarden (1:198)



*Figuur 15: verdeling van de boekwaarden van het voorraadbestand.*

In het histogram is duidelijk te zien dat de kleinere waarden (waarde kleiner dan 75) qua aantal het grootste deel van de voorraad vormen. De kenmerken van de verdeling, hoge kansdichtheid van dlagere waarden en de aanwezigheid van een brede 'staart', doen vermoeden dat de verdeling van de boekwaarden met de Weibullverdeling kan worden benaderd. De verdelingsfunctie van de Weibull is:

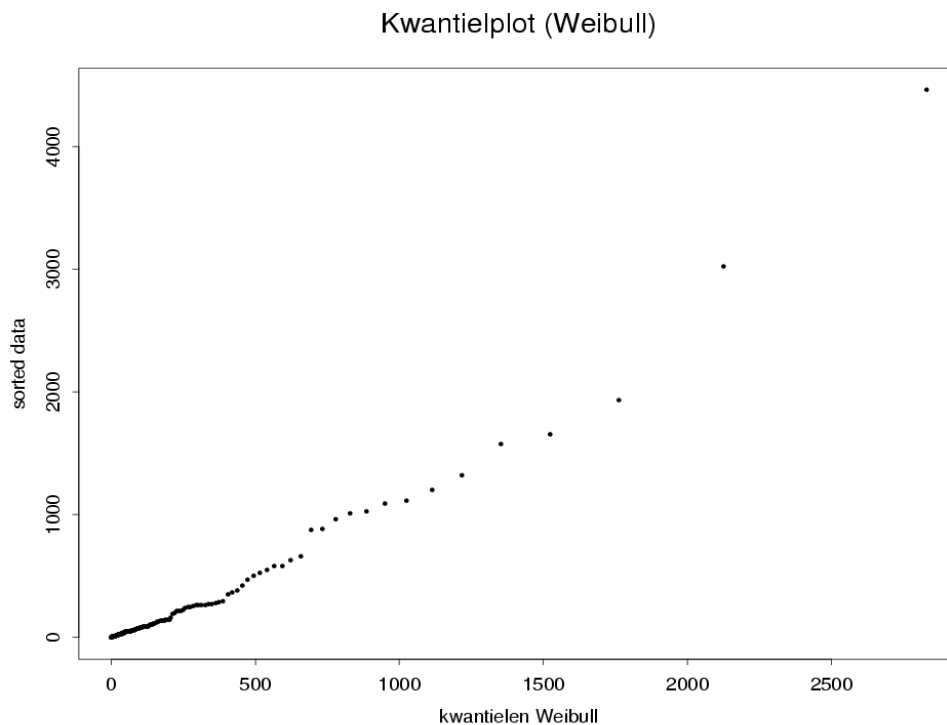
$$(26) \quad F(x) = 1 - e^{-\left(\frac{x}{\beta}\right)^\alpha}$$

Om te onderzoeken of de verdeling een goede benadering is van de onderliggende verdeling van de dataset maak ik gebruik van een kwantielplot. De parameterwaarde van de theoretische verdeling kunnen worden benaderd met de Meest Aannemelijke Schatter (MAS). Bij het berekenen van de MAS voor de vorm- en schaalparameter van de Weibullverdeling,  $\hat{a}$  en  $\hat{b}$ , komt men op respectievelijk de volgende vergelijkingen:

$$(27) \quad \frac{\partial X}{\partial \hat{b}} = -\frac{\hat{a}}{\hat{b}} \cdot n + \frac{\hat{a}}{\hat{b}} \sum_{i=1}^n \left( \frac{X_i}{\hat{b}} \right)^{\hat{a}} = 0$$

$$(28) \quad \frac{\partial X}{\partial \hat{\alpha}} = \frac{n}{\hat{\alpha}} + \sum_{i=1}^n \ln \left( \frac{X_i}{\hat{\beta}} \right) - \sum_{i=1}^n \left( \left( \frac{X_i}{\hat{\beta}} \right)^{\hat{\alpha}} \cdot \ln \left( \frac{X_i}{\hat{\beta}} \right) \right) = 0$$

Deze vergelijkingen is numeriek op te lossen. Op basis van de hele dataset, levert dat de waarden  $\hat{a} = 0.49$  en  $\hat{b} = 94.17$  op. De kwantielplot ziet er dan als volgt uit:

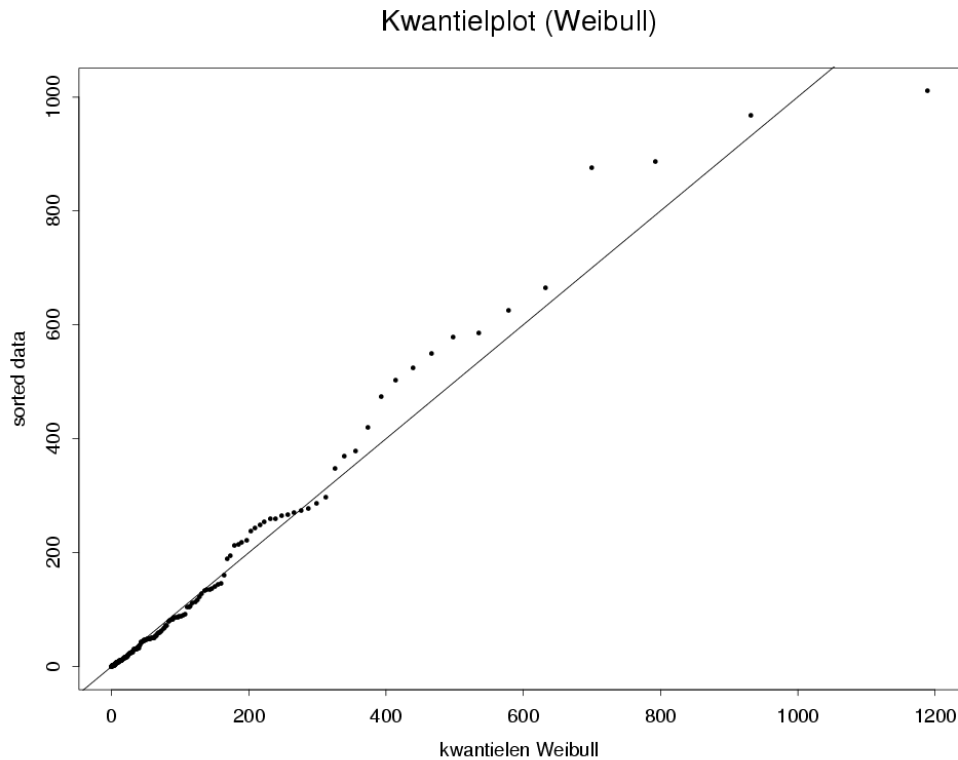


*Figuur 16: kwantielplot tegen de Weibullverdeling.*

In *Figuur 16* is te zien is dat de datapunten in de hogere kwantielen afwijken van een eventueel aanwezige  $y = x$  lijn. Wellicht dat het splitsen van de voorraad in een deel hogere waarden (uitschieters) en een deel met daarin de overige waarden uitkomst kan bieden.

De verdeling kan wellicht met een gemengde verdeling worden benaderd. Omdat de datapunten van het voorraadbestand moeilijk te scheiden zijn in twee afzonderlijke verdelingen, is het bepalen van de gemengde verdeling geen gemakkelijke opgave. Daarom is in het onderzoek het doel gesteld om een voorraad te construeren die lijkt op een Weibullverdeling met alleen een dikkere staart.

De tien hoogste waarden worden beschouwd als uitschieters. De MAS worden nu  $\hat{a} = 0.58$  en  $\hat{b} = 68.21$  voor de overige datapunten  $X_{(1)}, X_{(2)}, \dots, X_{(191)}$ . De kwantielplot van deze partiële dataset wordt:



Figuur 17: kwantielplot tegen de Weibullverdeling van de partiële dataset. De doorgetrokken lijn stelt de lijn  $y = x$  lijn voor.

In de kwantielplot ziet men dat de lijn  $y = x$  voornamelijk door de lagere waarden wordt gevolgd. De hogere waarden wijken af van deze lijn en geven dus aan dat de Weibull(0.58, 68.61) niet de onderliggende verdeling is.

Men kan tevens de nulhypothese: *dataset is Weibull(0.58,68.21) verdeeld*, toetsen met de Kolmogorov-Smirnov-toets. Dit levert een toetsingsgrootte van 5.00 op. De overschrijdingskans is 0.28 en daarmee groter dan de (standaard) significantiegrens van 0.05. Dit betekent dat de nulhypothese niet verworpen wordt en kan de Weibullverdeling niet worden uitgesloten. Daarom wordt uiteindelijk toch de Weibull(0.58,68.21) gekozen om de boekwaarden te construeren.

De overige waarden,  $X_{(192)}, X_{(193)}, \dots, X_{(201)}$ , bestaan uit waarden die niet allemaal extreem hoog zijn. De eerste zeven waarden verschillen niet eens zo veel van de  $X_{(191)}$ . Ze veroorzaken in wezen de bredere staart van de verdeling. Het modelleren van deze bredere staart gebeurt met de uniforme verdeling.

Het model waarmee de boekwaarden gegenereerd kunnen worden, maakt dus gebruik van de Weibullverdeling en de uniforme verdeling. Met een kans  $p_1$  zal de boekwaarde van een subpost via de Weibullverdeling bepaald worden. Dit betreft vooral de kleinere subpostwaarden. Met kansen  $p_2, \dots, p_n$  worden de overige (hogere) subpostwaarden bepaald. Deze worden getrokken uit een uniforme verdeling. Als de  $y_{subpost;i}$  de boekwaarde van subpost  $i$  voorstelt dan kan model (29) gebruikt worden:

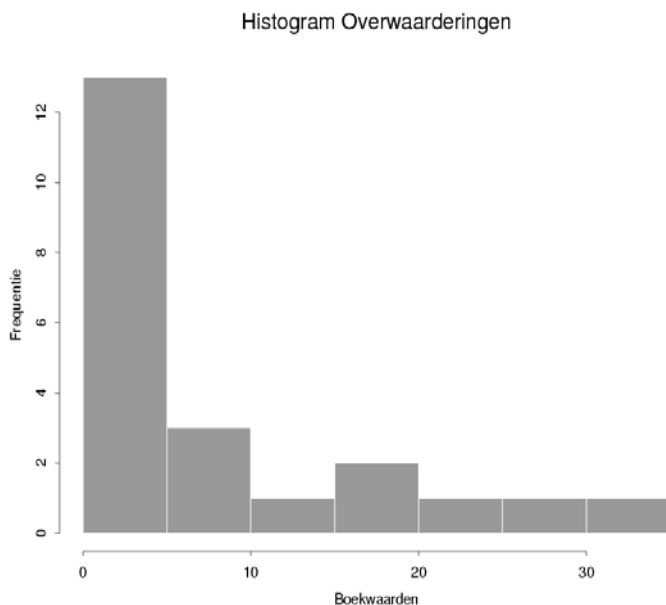
$$(29) \quad y_{subpost:i} \sim \begin{cases} p_1 = 0.965 : Weibull(0.58, 68.61) \\ p_2 = 0.025 : U(600, 1600) \\ p_3 = 0.010 : U(2000, 20000) \end{cases} \quad \text{met } \sum_{j=1}^3 p_j = 1$$

De  $p$ -waarden zijn net als de parameterwaarden van de gebruikte uniforme verdeling globaal ontleend van het voorraadbestand.

### Model Residuwaarden

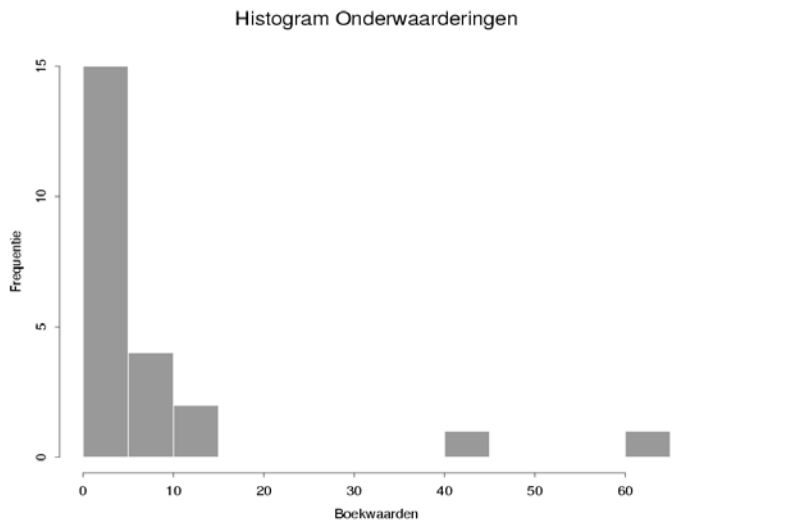
Het is niet interessant de verdeling van de auditwaarden te analyseren daar deze nagenoeg gelijk zal zijn aan de verdeling van de boekwaarden. Het is interessanter om te onderzoeken welke verdeling de residuen hebben en hoe deze residuen in waarde zich verhouden tot de (bijbehorende) boekwaarden. Voor het onderzoek zijn de residuen in twee delen gesplitst, namelijk residuen met een positieve waarde (overwaarderingen) en residuen met een negatieve waarde (onderwaarderingen).

In *Figuur 18* en *Figuur 19* staan respectievelijk de histogrammen van de over- en onderwaarderingen. In geval van onderwaardering is het absolute verschil *boek – audit* gebruikt.



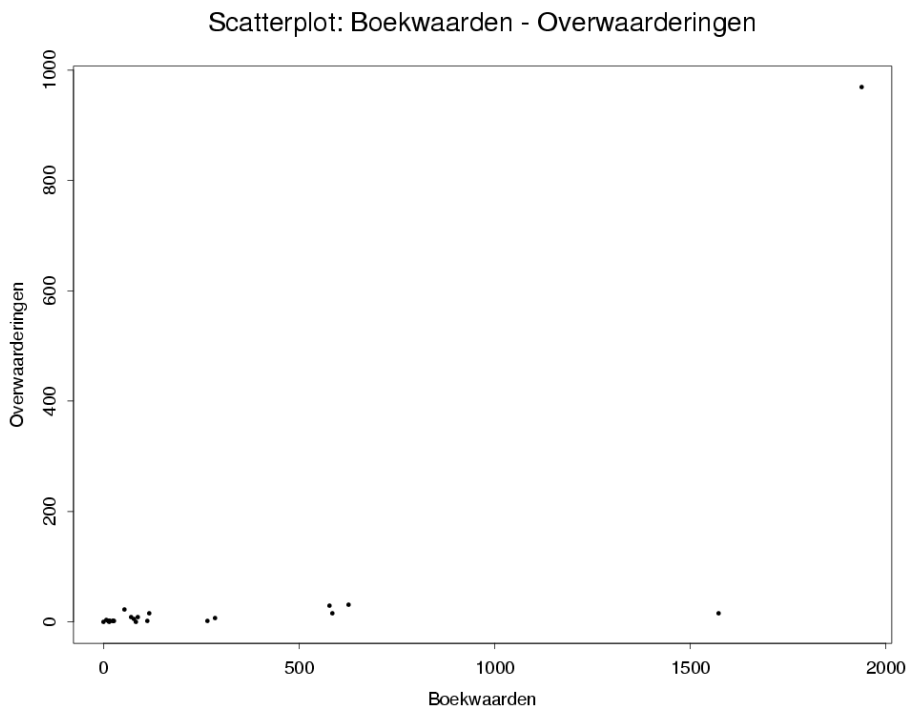
*Figuur 18: de overwaarderingen (22 datapunten) in een histogram.*





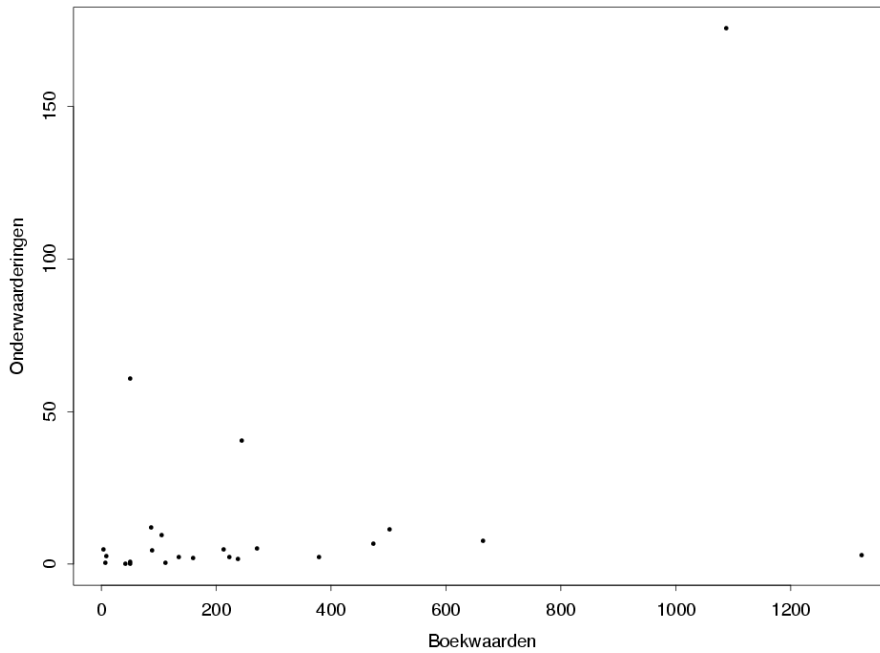
Figuur 19: de onderwaarderingen (absolute waarden)(23 datapunten) in een histogram

De verdeling van de residuen lijkt net als de boekwaardeverdeling sterk op een Weibullverdeling. Om te onderzoeken of er een sterk verband bestaat tussen de boek- en residuwaarden kan men gebruik maken van scatterplots. In *Figuur 20* en *Figuur 21* zijn respectievelijk de over- en de onderwaarderingen geplot tegen de boekwaarden.



Figuur 20: overwaarderingen uitgezet tegen de boekwaarden

Scatterplot: Boekwaarden - Onderwaarderingen



Figuur 21: onderwaarderingen (absolute waarde) uitgezet tegen de boekwaarden.

De scatterplots tonen aan dat de residuen voornamelijk kleine waarden aannemen. Er zijn echter ook uitzonderingen zoals de overwaardering van 1000 op een boekwaarde van 2000, of de onderwaardering van 175 op een boekwaarde van 1100.

In het model voor het construeren van de residuen, worden de boekwaarden verdeeld in een aantal onafhankelijk, aan elkaar grenzende intervallen. Per interval kan een residu behorend bij subpost  $i$  ( $r_{subpost;i}$ ) verschillende homogene fracties aannemen ten opzichte van de boekwaarde van subpost  $i$  ( $y_{subpost;i}$ ). In formulevorm:

$$(30) \quad r_{subpost;i} = y_{subpost;i} (Hom(0,1) \cdot a + b)$$

In (30) kunnen de waarden voor  $a$  en  $b$  variëren per interval.

Na analyse zijn met betrekking tot model (30) globale waarden gekozen voor de verhoudingen tussen de overwaarderingen - boekwaarden en de verhouding onderwaarderingen - boekwaarden. Deze waarden staan in Tabel 14 en Tabel 15.

Tabel 14: overwaarderingen – boekwaarden

$y_{subpost;i}$	Fractie	$a$	$b$
0 - 100	$p_1 = 1.00$	0.20	0.05
100 - $\infty$	$p_1 = 0.90$	0.05	0.00
	$p_2 = 0.10$	0.30	0.50

Tabel 15: onderwaarderingen – boekwaarden

$y_{subpost;i}$	Fractie	$a$	$b$
0 – 100	$p_1 = 0.50$	0.03	0.00
	$p_2 = 0.50$	0.10	0.10
100 - $\infty$	$p_1 = 0.85$	0.03	0.00
	$p_2 = 0.15$	0.03	0.15

Met de waarden in *Tabel 10* en *Tabel 11* wordt dus de hoogte van de residuen bepaald.



## 15 Appendix C

Het percentage runs waarin de CA-tool een correcte uitspraak heeft geleverd onder alle gebruikte instellingen staan in *Tabel 16* (scenario 1) en *Tabel 17* (scenario 2). De tabellen zijn als volgt te lezen:

- Rij 1, Kolom 1 - 3: betreffen de parameters in de regressieschatter-plus (het aantal steekproefelementen  $n$  en het aantal vervangingselementen  $d$ ) en het vertalingsmechanisme (de grenswaarde  $C$ ) die van waarde zijn gewijzigd in het onderzoek.
- Rij 2 – 6, Kolom 1 – 3: betreffen de waarden van de gewijzigde parameters ( $n$ ,  $d$  en  $C$ ) die in het onderzoek zijn gebruikt.
- Rij 1, Kolom 4 – 11: betreffen de  $K$ -waarden die onder specifieke instellingen van  $C$ ,  $n$  en de  $d$  zijn gebruikt bij het uitrekenen van de  $CP$ -waarde.
- Rij 2 – 6, Kolom 4 – 11: betreffen de  $CP$ -waarden die zijn gehaald onder de specifieke instellingen rondom de  $C$ ,  $n$ ,  $d$  en de  $K$ -waarde.

*Tabel 16: CP-waarden onder verschillende instellingen onder scenario 1*

$C$	$N$	$D$	$K = 1$	$K = 2$	$K = 3$	$K = 4$	$K = 5$	$K = 6$	$K = 7$	$K = 8$
5	200	40	20.8	45.8	61.2	76.0	80.2	89.6	92.0	93.9
5	300	60	65.2	78.9	89.2	95.8	96.3	98.3	99.1	99.0
4	200	40	1.8	8.9	20.5	32.3	47.3	53.3	65.0	73.1
6	200	40	59.3	77.4	88.2	95.4	97.1	97.1	99.5	99.0
5	200	20	38.3	55.6	65.1	71.4	77.5	81.3	86.5	89.7
5	200	60	11.7	39.7	62.5	79.2	89.8	93.4	96.4	97.7

*Tabel 17: CP-waarden onder verschillende instellingen onder scenario 2*

$C$	$N$	$D$	$K = 1$	$K = 2$	$K = 3$	$K = 4$	$K = 5$	$K = 6$	$K = 7$	$K = 8$
5	200	40	52.2	69.7	84.6	89.6	91.4	95.1	97.7	98.5
5	300	60	83.5	94.0	95.5	98.3	98.4	99.5	99.8	100.0
4	200	40	17.5	35.0	51.5	64.2	74.0	81.9	87.2	90.4
6	200	40	81.5	92.0	96.6	97.2	98.8	99.4	98.9	99.3
5	200	20	67.2	76.2	83.1	86.3	89.5	91.9	95.9	95.9
5	200	60	46.3	70.5	85.7	93.3	95.0	97.8	98.6	98.6



## 16 Literatuurlijst

- [1] P. van Batenburg Over de toepasbaarheid van regressieschatter en verschilschatter in de accountantscontrole. Onderzoeksverslag Deloitte & Touche, december 1994.
- [2] R.S. Kaplan Statistical Sampling in Auditing with Auxiliary Information Estimators. Journal of Accounting Research, najaar 1973.
- [3] M.C.M. de Gunst en A.W. van der Vaart Statistische Data Analyse, 2001.
- [4] T.R. Lew Regressieschatter: literatuurstudie naar de praktijkaannames voor toepassen van normaliteit
- [5] H. Timmermans en H. van den Wildenberg Simulatie van steekproeven tbv accountancy, september 1997.
- [6] I.Roest Stageverslag

