

UITBREIDEN VAN DE STEEKPROEF

door M. Vermaas

Inleiding

In mijn artikel in het maart-nummer van dit maandblad beriep ik mij op de mogelijkheid een steekproef bij het vinden van fouten uit te breiden, terwijl tevoren het maximaal toelaatbare foutenpercentage en de vereiste betrouwbaarheid waren aangegeven. Daarbij maakte ik gebruik van een tabel welke voorkomt in een publikatie in „De Accountant” nr. 8 van april 1975 (blz. 507). Die tabel luidt - verkort - als volgt:

aantal fouten in de steekproef	n.p. = produkt van steekproefomvang (n) en foutenfractie (p) in de populatie	
	minimaal	maximaal
0	0.00	3.00
1	0.05	4.74
2	0.36	6.30
3	0.82	7.75
4	1.37	9.15
5	1.97	10.51
6	2.61	11.84
	Betrouwbaarheid 95%	

Uit de tabel kan met een betrouwbaarheid van 95% worden afgeleid dat bij een steekproefomvang van b.v. 150 stuks waarbij 2 fouten worden gevonden, de populatie minimaal $\frac{0.36}{150} = 0.24\%$ en maximaal $\frac{6.30}{150} = 4.2\%$ fouten bevat.

Mij werd voorgehouden dat het uitbreiden van een steekproef bij het vinden van fouten, met gebruikmaking van deze tabel, bij handhaving van de eis van 95% betrouwbaarheid, statistisch niet toelaatbaar zou zijn. In een naschrift zei ik er geen kwaad in te zien; dat zal ik nu trachten te verantwoorden.

Het blijkt te gaan om een - door mij nog aangehangen - „oude opvatting” van de subcommissie steekproeven (CCS), ingesteld door de Commissie van Advies inzake Controlevraagstukken van het NIVRA; die opvatting komt voor in het interimrapport van die subcommissie CCS, dat werd gepubliceerd in „De Accountant” van juli/augustus 1974. Toen was de CCS nog van mening dat bedoelde handwijze toelaatbaar was.

Hoewel de CCS nog geen eindrapport heeft uitgebracht, is gebleken dat men binnen de CCS de „oude opvatting” heeft prijsgegeven voor een „nieuwe opvatting”. Bovendien stelt men dat de „oude opvatting” onjuist is. Dat laatste wil ik

weerleggen. Volgens mij hebben beide opvattingen recht van bestaan. Ik zal laten zien dat het verschil tussen beide opvattingen schuilt in een verschillende definitie van het begrip betrouwbaarheid. Ook hoop ik duidelijk te maken dat de „oude opvatting” de voorkeur verdient boven de „nieuwe opvatting”.

„Oude opvatting”

De „oude opvatting” van de CCS zal ik met een voorbeeld weergeven.

Het maximaal toegestane aantal fouten in een populatie is gesteld op 4%. Op grond van een steekproef moet worden gezien of met 95% betrouwbaarheid de uitspraak kan worden gedaan dat het werkelijke aantal fouten in de populatie de 4% niet overschrijdt.

Er zal worden begonnen met een steekproef van 75 stuks; worden daarbij geen fouten aangetroffen dan volgt de uitspraak: „met een betrouwbaarheid van 95% wordt verklaard dat de populatie niet meer dan 4% fouten bevat”.

Wordt in de steekproef van 75 stuks echter één fout gevonden, dan wordt een aanvullende steekproef van 44 stuks genomen; worden daarbij geen fouten aangetroffen, dan volgt alsnog *dezelfde* uitspraak als bovengenoemd, te weten: „met een betrouwbaarheid van 95% wordt verklaard dat de populatie niet meer dan 4% fouten bevat”.

De CCS stond met die opvatting niet alleen.

„Nieuwe opvatting”

Volgens een verslag van de „NIVRA-najaarsconferentie 1974”, gepubliceerd in „De Accountant” van april 1975 door de collega’s E. B. Tuitjer en J. M. Zuydervliet, beiden lid van de CCS, was achteraf gebleken dat de in het interimrapport van de CCS gehuldigde opvatting toch niet juist was.

De „nieuwe opvatting” zal ik beschrijven, uitgaande van het reeds gegeven voorbeeld.

Indien in de steekproef van 75 stuks geen fouten worden aangetroffen, dan mag inderdaad de uitspraak luiden: „met een betrouwbaarheid van 95% wordt verklaard dat de populatie niet meer dan 4% fouten bevat”, mits men niet van plan was de steekproef voort te zetten.

Dezelfde uitspraak mag worden gedaan indien men van plan is een steekproef van 194 stuks te nemen, waarbij een tussentijdse beslissing of een voortzetting na de 194e trekking wordt uitgesloten en indien men daarbij niet meer dan 3 fouten vindt.

Indien echter in de steekproef van 75 stuks één fout wordt gevonden, in een tweede (aanvullende) steekproef van 44 stuks weer één fout, in een derde (aanvullende) steekproef van 39 stuks nogmaals één fout, maar in een vierde (aanvullende) steekproef van 36 stuks géén fout, dan mag niet dezelfde uitspraak worden gedaan. Toch omvat ook hier de totale steekproef 194 posten en werden niet meer dan 3 fouten gevonden. Was men in dit geval van plan de populatie te accepteren bij 0 fout in de eerste steekproef, 1 fout in de eerste en tweede steekproef tezamen, 2 fouten in de eerste tot en met de derde steekproef of 3 fouten in de eerste tot en met de vierde steekproef en verder van uitbreiding af te zien, dan

kan volgens de „nieuwe opvatting” de uitspraak slechts worden gedaan met een betrouwbaarheid van 89.3% in plaats van 95%.

Ook bij de „nieuwe opvatting” kan men een steekproefplan opstellen waarbij men bij het vinden van fouten tot uitbreiding van de steekproef overgaat en toch tot een uitspraak met een betrouwbaarheid van - in dit geval - 95% kan komen. Men moet dan echter beginnen met een grotere betrouwbaarheid, bijv. van 98% en dan een bepaald aantal mogelijke aanvullende steekproeven, bijv. 3, berekenen. Na die 3 aanvullende steekproeven is dan geen uitbreiding meer mogelijk; de 95% betrouwbaarheidsgrens is bereikt.

De „nieuwe opvatting” wordt blijkens een publikatie in het „Handboek Accountancy” (april 1976) van de hand van collega J. H. Blokdijk, voorzitter van de CCS, ook door hem aangehangen. Hij stelt eveneens dat de „oude opvatting” statistisch niet verantwoord is.

Een verhaal

De „nieuwe opvatting” houdt in dat het volgende verhaal op waarheid zou kunnen berusten.

In een land zijn over het belastingjaar 1977 5.000.000 aangiftebiljetten voor de inkomstenbelasting afgehandeld. Ze liggen keurig op volgorde van de bij binnenkomst aangebrachte, doorlopende nummering, in een centraal archief opgeslagen.

Het gerucht komt in omloop dat circa 1/3 van de aangiftebiljetten (opzettelijke) telfouten bevat, zonder dat zulks ontdekt zou zijn.

De verantwoordelijke bewindsman ontkent dit, beweert dat hoogstens 4% van de aanslagbiljetten onontdekte telfouten bevatten en zal op verzoek van het parlement een openbaar onderzoek doen instellen.

De bewindsman komt met vier accountants overeen dat zij een steekproefsgewijs onderzoek zullen verrichten. De accountants zullen hun uitspraken doen met 95% betrouwbaarheid.

Het onderzoek wordt per T.V. uitgezonden. Beelden van het archief, waar de 5.000.000 biljetten liggen opgeslagen, alsmede van de apparatuur waarmee de nummers van de voor de steekproef aan te wijzen biljetten zullen worden bepaald (geleend van de „Toto”) verschijnen op het scherm.

De regisseur introduceert de vier accountants - allen aanhangers van de „nieuwe opvatting” - bij de kijkers; elke accountant zet zijn plan uiteen.

Accountant A: zal zijn beslissing nemen nadat 75 geselecteerde biljetten zijn nagezien. Zijn er geen biljetten met telfouten geconstateerd, dan zal hij met een betrouwbaarheid van 95% verklaren dat niet meer dan 4% van de biljetten telfouten bevatten.

Accountant B: zal zijn beslissing nemen nadat 119 geselecteerde biljetten zijn nagezien. Is er slechts één biljet met telfouten gevonden, dan zal hij dezelfde verklaring afgeven als accountant A.

Accountant C: zal dezelfde verklaring afgeven als A en B indien bij 88 geselecteerde biljetten geen biljet met telfouten is geconstateerd; zou er één biljet met telfouten gevonden zijn, dan zal hij zijn beslissing opschorten tot het 130e biljet is nagezien. Is het dan bij dat ene foute biljet gebleven, dan zal hij alsnog dezelfde verklaring afgeven.

Accountant D: zal zijn beslissing eerst nemen nadat 194 biljetten zijn nagezien. Hij zegt dezelfde verklaring toe als de andere accountants, indien niet meer dan 3 biljetten met telfouten worden ontdekt.

De „T.V.-show” is goed georganiseerd. Er zijn voldoende medewerkers ingeschakeld om de biljetten waarvan de nummers zijn aangewezen, uit de archiefrekken te halen en ter controle te overhandigen aan een groep assistenten die door de vier accountants zijn ingezet. Door vele „close-ups” zijn de kijkers vaak in staat zelf mee te tellen; een psychologische zet, want onder de bevolking zijn de gemoe-deren nogal verhit geraakt.

Om kort te gaan:

Het 60e geselecteerde biljet blijkt een grove telfout te bevatten en accountant A verlaat met een joviaal gebaar de archiefruimte (close-up!).

Nadat 119 geselecteerde biljetten zijn nagezien - en geen tweede biljet met telfouten is gesignaleerd - vertrekt ook accountant B.

In het 125e biljet zit een telfout en accountant C zwaait af.

Accountant D is overgebleven; hij maakt mee dat in het 144e biljet een telfout wordt gevonden. Dat is dus het derde foute biljet. Na controle van het 194e biljet - er zijn geen verdere fouten meer aan het licht gekomen - geeft accountant D een teken dat de „show” kan worden gestopt.

Na het middernachtelijk uur vindt de slotuitzending plaats vanuit de werk-kamer van de bewindsman (die door ambtsbezigheden verhinderd is geweest de uit-zending te volgen). De vier accountants zijn aldaar gearriveerd en leggen hun ver-klaringen af:

Accountant A: „aangezien ik niet met een betrouwbaarheid van 95% kan verkla-ren dat niet meer dan 4% van de biljetten telfouten bevatten, kan ik de door u gewenste verklaring niet afgeven”.

Accountant B: „met een betrouwbaarheid van 95% verklaar ik dat niet meer dan 4% van de biljetten telfouten bevatten”.

Accountant C: als accountant A.

Accountant D: als accountant B.

De bewindsman is nogal verbaasd en laat zich een en ander verklaren. Hij zegt dan (ook niet dom!): „maar heren A en C, kunt u de conclusie van uw collega D niet overnemen; die baseert zijn verklaring toch op het grootste aantal waarne-mingen”.

De accountants A en C zeggen dat niet te kunnen doen. A legt uit dat voor hem de mogelijkheid om met een betrouwbaarheid van 95% te kunnen verklaren dat niet meer dan 4% van de biljetten telfouten bevatten, bij het nazien van het 60e biljet al volkomen was opgesoupeerd en dat de verlangde conclusie door welke uitbreiding van de steekproef dan ook, nimmer meer bereikt kon worden (woor-den uit het verslag Tuitjer/Zuydervliet!).

De bewindsman - denkend aan de publieke opinie - vraagt nog eens: „heren, kunt u op grond van de thans bekende feiten werkelijk niet tot een eensluidend oordeel komen?” „Dat laten onze beroepsopvattingen niet toe”, antwoorden de accountants. „Vreemde beroepsopvattingen!”, laat de - duidelijk geïrriteerde - be-windsman zich nog ontvallen en de uitzending wordt beëindigd.

Nabeschuwing

Indien de „oude opvatting” van de CCS door de accountants A, B, C en D gehuldigd was, zou het verhaal anders geëindigd zijn. Alle vier de accountants zouden dan de „show” hebben gestopt op het moment dat 119 biljetten waren nagezien. Op dat moment was er nog maar 1 fout biljet gevonden en op grond daarvan zouden zij alle vier met 95% betrouwbaarheid hebben verklaard dat niet meer dan 4% van de biljetten telfouten bevatten.

Zou de „oude opvatting” dan toch . . .

We moeten nauwkeurig nagaan wat het „betrouwbaarheidspercentage” bij de „oude opvatting” zowel als dat bij de „nieuwe opvatting” betekent.

In de „nieuwe opvatting” betekent een betrouwbaarheid van 95% dat, *als* de onderzochte en geaccepteerde populatie een populatie is, waarin meer dan het maximaal toelaatbare foutenpercentage voorkomt, zij hoogstens 5% kans heeft gekregen te worden geaccepteerd.

In de „oude opvatting” betekent een betrouwbaarheid van 95%, dat er maar 5% kans is dat de onderzochte en geaccepteerde populatie een populatie is, waarin meer dan het maximaal toelaatbare foutenpercentage voorkomt.

Dat is niet hetzelfde; daarom leidt de „nieuwe opvatting” wèl tot zo’n vreemde situatie als geschetst in het voorgaande verhaal, en de „oude opvatting” niet.

Ik ben van mening dat het begrip betrouwbaarheid in de „oude opvatting” een verantwoord, op de kansrekening gebaseerd en voor de praktijk goed bruikbaar begrip is.

Belangrijk vind ik, dat men, uitgaande van de „oude opvatting” een *minimale* steekproefomvang kan vaststellen als het gewenste betrouwbaarheidspercentage is aangegeven.

Becijferingen

Ik wil nu met becijferingen laten zien hoe de berekening van de betrouwbaarheid bij de „oude opvatting” en bij de „nieuwe opvatting” van de subcommissie CCS verloopt.

Voorbeeld:

Maximaal toelaatbaar foutenpercentage 4.

Vereiste betrouwbaarheid 95%.

Steekproefplan bij de „oude opvatting”

1e steekproef 75 stuks:

0 fout, populatie accepteren

1 fout, steekproef uitbreiden

2 fout, idem

etc.

2e steekproef 44 stuks:

in totaal 1 fout, populatie accepteren

in totaal 2 fout, steekproef uitbreiden

in totaal 3 fout, idem

etc.

3e steekproef 39 stuks:

- in totaal 2 fout, populatie accepteren
- in totaal 3 fout, steekproef uitbreiden
- in totaal 4 fout, idem
- etc.

4e steekproef 36 stuks:

- in totaal 3 fout, populatie accepteren
- in totaal 4 fout, steekproef uitbreiden
- in totaal 5 fout, idem
- etc.

De steekproef kan worden gestaakt indien reeds zoveel fouten zijn gevonden dat de ondergrens van de foutenfractie met een betrouwbaarheid van 95% op 4% kan worden gesteld, of indien de onder- en bovengrens van de mogelijke foutenfractie in de populatie elkaar voldoende dicht zijn genaderd. Bij een steekproefomvang van b.v. 263 stuks ligt de foutenfractie in de populatie, indien 6 fouten werden gevonden, met 95% betrouwbaarheid tussen 1% en 4.5%. Dan verkeert men in de situatie dat men nog niet weet of de maximaal toelaatbare foutenfractie van 4% al dan niet is overschreden. Wil men de steekproef niet verder uitbreiden, dan zal men een keus moeten maken tussen goedkeuren of afkeuren van de populatie. Dat is echter een situatie welke niet kenmerkend is voor de „oude opvatting”; ook bij de „nieuwe opvatting” komt dat voor.

Er wordt dus begonnen met een *minimale* steekproef van 75 stuks; worden geen fouten gevonden, dan is men klaar.

Indien men door het vinden van fouten aan bijv. de 4e steekproef toekomt en bij het gereedkomen daarvan zijn 3 fouten in de totale steekproef van $75 + 44 + 39 + 36 = 194$ stuks gevonden, dan wordt de populatie geaccepteerd op grond van de volgende redenering:

Het feit dat van de 194 bij steekproef gecontroleerde posten er slechts 3 fout zijn betekent dat er slechts 5% kans is dat de populatie meer dan 4% fouten bevat; immers indien de populatie (iets meer dan) juist 4% fouten zou bevatten, dan zou er juist (iets meer dan) 95% kans zijn dat er (meer dan) 3 fouten waren gevonden. Indien er juist 4% fouten in de populatie voorkomen is n.l. de kans op het vinden van

0 fouten:

$$0,96^{194} = 0,000$$

1 fout:

$$0,96^{193} \times 0,04 \times 194 = 0,003$$

2 fouten:

$$0,96^{192} \times 0,04^2 \times \frac{194 \times 193}{2} = 0,016$$

3 fouten:

$$0,96^{191} \times 0,04^3 \times \frac{194 \times 193 \times 192}{3 \times 2} = \underline{\underline{0,031}}$$
$$\underline{\underline{0,05}}$$

De kans op het vinden van meer dan 3 fouten is dus $1 - 0,05 = 0,95 = 95\%$. Indien

er (iets) meer dan 4% fouten in de populatie voorkomen is de kans op het vinden van meer dan 3 fouten uiteraard (iets) groter dan 95%. Het feit dat er slechts 3 fouten werden gevonden zegt ons dan met een betrouwbaarheid van 95% dat er niet meer dan 4% fouten in de populatie voorkomen.

Steekproefplan bij de „nieuwe opvatting”

1e mogelijkheid:

Men kiest één bepaalde steekproefomvang, waarbij dan geen verdere uitbreiding is toegestaan. Daarom zal men - als men van te voren verwacht dat in de steekproef wel fouten zullen voorkomen - niet de minimale steekproefomvang kiezen, maar een grotere steekproefomvang. Men kan bijv. de steekproef bepalen op 194 stuks; worden dan niet meer dan 3 fouten in de steekproef aangetroffen, dan kan men de populatie accepteren met 95% betrouwbaarheid.

Het is dus dezelfde steekproefomvang als bij de „oude opvatting” indien daarbij inderdaad 3 fouten in de steekproef worden aangetroffen. Maar . . . indien in de eerste 75 posten geen fouten worden gevonden controleert men bij de „nieuwe opvatting” onnodig 194 posten in plaats van 75.

2e mogelijkheid:

De tweede mogelijkheid bij de „nieuwe opvatting” is, dat men een minimale steekproefomvang bepaalt, waarbij men wel tot uitbreiding mag overgaan indien fouten worden aangetroffen. De betrouwbaarheid moet dan op meer dan de vereiste 95% worden gesteld. Ook de minimale steekproefomvang bij de „nieuwe opvatting” wordt derhalve groter dan de minimale steekproefomvang bij de „oude opvatting”.

Men kan beginnen met een betrouwbaarheid van bijv. 98% en tot bijv. het volgende steekproefplan komen:

1e steekproef 98 stuks:

- 0 fout, populatie accepteren
- 1 fout, steekproef uitbreiden
- 2 fout, idem
- 3 fout, idem

2e steekproef 51 stuks:

- in totaal 1 fout, populatie accepteren
- in totaal 2 fout, steekproef uitbreiden
- in totaal 3 fout, idem

3e steekproef 35 stuks:

- in totaal 2 fout, populatie accepteren
- in totaal 3 fout, steekproef uitbreiden

4e steekproef 32 stuks:

- in totaal 3 fout, populatie accepteren.

Verdere uitbreiding is dan - volgens de „nieuwe opvatting” - niet mogelijk, omdat de betrouwbaarheid dan onder de 95% zou dalen.

Om te laten zien hoe de acceptatiekans voor een populatie met juist 4% fouten op 5% wordt becijferd, duiden we de vier steekproeven aan met de letters a, b, c en d. Met de notatie a_2 b_0 c_0 geven we dan aan dat in de eerste steekproef 2 fouten en in de tweede en derde steekproef geen fouten werden gevonden.

Acceptatiemogelijkheden	Acceptatiekans (zie Poisson-verdeling)	
1e steekproef:		
a_0		= 0.02
2e steekproef:		
$a_1 b_0$	0.078×0.128	= 0.01
3e steekproef:		
$a_1 b_1 c_0$	$0.078 \times 0.263 \times 0.249$	= 0.0051
$a_2 b_0 c_0$	$0.153 \times 0.128 \times 0.249$	= 0.0049
4e steekproef:		
$a_1 b_1 c_1 d_0$	$0.078 \times 0.263 \times 0.346 \times 0.281$	= 0.0020
$a_1 b_2 c_0 d_0$	$0.078 \times 0.270 \times 0.249 \times 0.281$	= 0.0015
$a_2 b_0 c_1 d_0$	$0.153 \times 0.128 \times 0.346 \times 0.281$	= 0.0019
$a_2 b_1 c_0 d_0$	$0.153 \times 0.263 \times 0.249 \times 0.281$	= 0.0028
$a_3 b_0 c_0 d_0$	$0.200 \times 0.128 \times 0.249 \times 0.281$	= 0.0018
		<u>0.05</u>

Aangezien de acceptatie-kans van een populatie met juist 4% fouten inderdaad 5% is, is de betrouwbaarheid - in de betekenis van de „nieuwe opvatting” - 95%.

Ik wijs er nog op dat in dit voorbeeld de minimale steekproefomvang bij de „nieuwe opvatting” 98 stuks belooft, terwijl die bij de „oude opvatting” 75 stuks is.

Niet alleen het „grensgeval”

De acceptatie kans van het grensgeval d.w.z. een populatie welke - in ons voorbeeld - juist iets meer dan 4% fouten bevat, neemt toe van 5% (exacter 4.98%) tot 9.31%, als de steekproef van 75 stuks wordt uitgebreid met twee aanvullende steekproeven van resp. 44 en 39 stuks. De betrouwbaarheid zou dan dus van 95% tot 90.69% dalen.

Dat is de reden voor het loslaten van de „oude opvatting” door de CCS.

Die - gestegen - acceptatie kans van 9.31% zegt echter niet veel voor de accountant die één bepaalde populatie moet controleren. Die acceptatie kans slaat op het „grensgeval” en zij zou voor de accountant reële betekenis hebben indien hij *altijd* foute populaties ter controle kreeg voorgelegd, d.w.z. nooit populaties die wel goed zijn (in het voorbeeld: 4% of minder fouten). We kunnen wel stellen dat die situatie op dit ondermaanse niet voorkomt.

Er moet niet uitsluitend worden gelet op de - stijgende - acceptatie kans van het „grensgeval”, maar op de acceptatiekansen van alle mogelijke populaties. Ik zal een meer reële frequentieverdeling naar foutenpercentage geven van de populaties die aan een (grote groep van) accountant(s) in de loop der jaren ter controle kunnen worden voorgelegd, indien die populaties moeten voldoen aan de norm „niet meer dan 4% fouten”. Bedacht moet worden dat het stellen van die 4%-norm inhoudt, dat het normaal is dat de meeste populaties binnen de norm blijven.

De frequentie-verdeling van 1.000.000 ter controle voorgelegde populaties zou bijv. kunnen zijn:

108.100	populaties met 1	% fouten
216.200	populaties met 2	% fouten
351.400	populaties met 3	% fouten
87.850	populaties met 4	% fouten
87.850	populaties met 4.01	% fouten
94.600	populaties met 5	% fouten
40.500	populaties met 6	% fouten
13.500	populaties met 7	% fouten
<u>1.000.000</u>		

Nu geven we de acceptatie-kansen

Acceptatie-kansen in % bij max.toelaatb.foutenfr. van 4%

<i>Foutenfractie in de populatie</i>	<i>Bij 1 steekpr. van 75 st.</i>	<i>Bij 2 steekpr. van resp. 75 en 44 st.</i>	<i>Bij 3 steekpr. van resp. 75, 44 en 39 st.</i>
1 %	47.23	70.05	82.64
2 %	22.31	36.19	46.56
3 %	10.54	16.87	21.68
4 %	4.98	7.55	9.31
4.01%	4.94	7.49	9.23
5 %	2.35	3.33	3.90
6 %	1.11	1.47	1.64
7 %	0.52	0.65	0.70

Van de 1.000.000 populaties worden geaccepteerd

<i>Populaties met een foutenfractie van</i>	<i>Bij 1 steekpr. van 75 st.</i>	<i>Bij 2 steekpr. van resp. 75 en 44 st.</i>	<i>Bij 3 steekpr. van resp. 75, 44 en 39 st.</i>
1 %	51.056	75.724	89.334
2 %	48.234	78.243	100.663
3 %	37.038	59.281	76.184
4 %	4.375	6.633	8.179
4.01%	4.340	6.580	8.109
5 %	2.223	3.150	3.689
6 %	450	595	664
7 %	70	88	95
	147.786	230.294	286.917
=====			

Indien · zonodig · twee aanvullende steekproeven worden genomen, worden

dus in totaal 286.917 populaties geaccepteerd, in plaats van 147.786 zonder die aanvullende steekproeven.

Ten onrechte geaccepteerde populaties (met meer dan 4% fouten)

	<i>Bij 1 steekpr. van 75 stuks</i>	<i>Bij 2 steekpr. van resp. 75 en 44 stuks</i>	<i>Bij 3 steekpr. van resp. 75, 44 en 39 st.</i>
In % van het aantal onderzochte populaties	0.71%	1.04%	1.26%

Het is duidelijk dat het uitbreiden van de steekproeven de betrouwbaarheidsgrens van 95% in het geheel niet aantast.

Bezien we de in het voorgaande gegeven tabel voor de acceptatie-kansen voor populaties met een foutenfractie van 1 t/m 7% nog eens, dan wordt het duidelijk dat bij een minimale steekproef van 75 stuks een echt goede populatie reeds een behoorlijke kans heeft als zodanig te worden herkend (indien de populatie slechts 1% fouten bevat, een kans van 47.23%). Daarom is het doelmatig met die minimale steekproef te beginnen.

Voorts blijkt uit de gemaakte opstellingen dat bij het uitbreiden van de steekproeven vooral de wel aan de norm voldoende populaties niet te verwaarlozen kansen krijgen alsnog terecht te worden geaccepteerd. Het meerdere malen uitbreiden van een steekproef kan dus wel degelijk nuttig zijn. Wat dat betreft ben ik het niet eens met een opmerking in het interimrapport van de CCS, dat verdere uitbreiding geen zin zou hebben indien in de eerste uitbreiding weer één of meer fouten worden aangetroffen. Ook daar heeft men m.i. te veel gelet op de kansen van het „grensgeval”, i.c. de populatie welke nog net goed is (4% fouten).

Ik merk nog op dat deze paragraaf zuiver is bedoeld als illustratie van de betrouwbaarheid van de „oude opvatting”. De ten tonele gevoerde frequentieverdeling maakt geen deel uit van de gehanteerde methode bij het volgen van de „oude opvatting”; die methode wordt geheel bepaald door de redenering weergegeven onder het hoofd „steekproefplan bij de „oude opvatting”” in de voorgaande paragraaf „Becijferingen”.

Sequential Sampling

De aanhangers van de „nieuwe opvatting” binnen de CCS kunnen stellen dat hun opvatting geheel spoort met de reeds lang in Amerika toegepaste methode van „sequential sampling”. Dat is juist.

Sinds de publikatie van P. C. Louwers in dit maandblad (oktober 1958) is die methode ook in ons land bekend onder de naam „likelihood-ratio-methode”. Mijn bedenkingen tegen de „nieuwe opvatting” binnen de CCS gelden echter ook tegen de „likelihood-ratio-methode”. Om dat te tonen volgt hier een voorbeeld volgens de „likelihood-ratio-methode”.

Omdat het gaat om een „ratio”, moeten 2 foutenfracties worden gesteld met een ruimte daartussen:

p_1 = een echt goede foutenfractie, bijv. 2%

p_2 = een net niet aanvaardbaar geachte foutenfractie, bijv. 4% (het „grensgeval” uit ons voorgaande voorbeeld).

Gaan we weer uit van een vereiste betrouwbaarheid van 95%, dan moet, om te kunnen goedkeuren, voldaan zijn aan de volgende „ratio”:

$$\frac{\text{acceptatiekans populatie met } p_2}{\text{acceptatiekans populatie met } p_1} \leq \frac{5\%}{95\%} = 0.053$$

De methode leidt tot de volgende tabel:

<i>Steekproefomvang</i>	<i>Aantal gevonden fouten</i>	
	<i>Populatie goedkeuren</i>	<i>Populatie afkeuren</i>
30		5
65		6
100		7
134		8
143	0	9
169	0	9
178	1	10
203	1	10
212	2	11

Voor goedkeuren is dus een minimale steekproefomvang van 143 stuks (met 0 fouten) nodig. Bij de methode volgens de „oude opvatting” van de CCS is de minimale steekproefomvang 75 stuks. Dus ook hier zien we dat de minimale steekproefomvang bij de „oude opvatting” belangrijk kleiner is.

Maar het volgende overzicht onthult nog meer:

<i>Steekproefomvang</i>	<i>Goedkeuren bij</i>	<i>Acceptatiekans</i>		<i>Waarde</i> <i>kolom 4</i> = <i>kolom 3</i>
		<i>Populatie met p_1</i>	<i>Populatie met p_2</i>	
143	0 fout	5.7%	0.3%	0.053
178	1 fout	13.8%	0.7%	0.051
212	2 fout	22.6%	1.2%	0.053

In de aanvang wordt gewerkt met een veel grotere betrouwbaarheid dan 95%. Men begint met een betrouwbaarheid van $100\% - 0.3\% = 99.7\%$, teneinde na herhaald uitbreiden binnen de 95% te kunnen blijven. Dat is in overeenstemming met hetgeen ik opmerkte bij de „nieuwe opvatting” binnen de CCS.

Ook is het interessant te zien dat een echt goede populatie met een foutenfractie van 2% bij de eerste steekproef van 143 stuks nog maar 5.7% kans krijgt te worden geaccepteerd; bij een steekproefomvang van 212 stuks is die kans gestegen tot 22.6%. Maar in de voorgaande paragraaf kon worden afgelezen dat bij de methode volgens de „oude opvatting” bij de eerste steekproef van slechts 75 stuks die populatie met een foutenfractie van 2% reeds 22.31% kans kreeg om te worden geaccepteerd!

Ook bij toepassing van de „likelihood-ratio-methode” kan men een „vreemde situatie” schetsen.

Als een accountant zonder meer een steekproef neemt van 212 stuks en daarbij 3 fouten vindt, kan hij met 95% betrouwbaarheid verklaren dat de populatie niet meer dan $\frac{7.75}{212} = 3.65\%$ fouten bevat (zie tabel in de paragraaf „Inleiding”). In ons voorbeeld zal hij de populatie dus goedkeuren. Hiermee zijn ook de aanhangers van de „nieuwe opvatting” het eens. Maar had de accountant vorenstaande tabel volgens de „likelihood-ratio-methode” gebruikt, dan zou hij, met dezelfde controle-resultaten in handen, de populatie nog niet hebben mogen goedkeuren, dus niet dezelfde uitspraak hebben mogen doen! Dat komt niet doordat de „nieuwe opvatting” of de „likelihood-ratio-methode” fout zou zijn, maar het is de consequentie van het begrip betrouwbaarheid dat bij de „nieuwe opvatting” en bij de „likelihood-ratio-methode” hoort. Daarom o.a. geef ik de voorkeur aan het andere begrip betrouwbaarheid dat past bij de „oude opvatting” van de CCS.

Eindconclusie

Ik blijf van mening dat de „oude opvatting”, neergelegd in het interimrapport van de CCS, een juiste opvatting is.

Die opvatting houdt in dat een *minimale steekproefomvang* wordt gekozen waarbij men een populatie met een zekere tolerantie (foutenfractie) en een bepaalde betrouwbaarheid kan accepteren, indien in de steekproef geen fouten worden aangetroffen. Treft men wel fouten aan, dan wordt de mogelijkheid niet uitgesloten dat bij de uitbreiding van de steekproef het vereiste betrouwbaarheidsniveau toch nog wordt bereikt.

De bewering van Tuitjer en Zuydervliet en van Blokdijk in hun, in het voorgaande vermelde, publikaties, n.l. dat de „oude opvatting” van de CCS ten aanzien van het uitbreiden van de steekproef statistisch niet verantwoord zou zijn, is niet houdbaar; men zou hooguit kunnen stellen dat de „oude opvatting” zich niet verdraagt met een bepaalde definitie van betrouwbaarheid.

De „nieuwe opvatting” binnen de CCS (en ook de „likelihood-ratio-methode”) heeft belangrijke nadelen ten opzichte van de „oude opvatting” van de CCS:

- a. de minimale steekproefomvang bij de „nieuwe opvatting” is groter dan die bij de „oude opvatting”;
- b. het begrip betrouwbaarheid dat bij de „nieuwe opvatting” (en de „likelihood-ratio-methode”) behoort, is onbevredigend en verwarrend, omdat twee accountants die beiden over *dezelfde kennis* omtrent een onderzochte populatie beschikken, toch niet tot dezelfde uitspraak over die populatie mogen komen.

Zoals uiteengezet, kan het zijn dat beide accountants weten dat bij een steekproef van 194 stuks 3 fouten zijn gevonden; verder is aan niemand iets bekend omtrent die onderzochte populatie. Toch mogen volgens de „nieuwe opvatting” beide accountants niet dezelfde uitspraak over die populatie doen als één van hen tussentijds een beslissing heeft overwogen.

Het nut van het uitbreiden van de steekproef moet niet worden gezocht bij de kleine kans dat een „grensgeval” (nog net goede populatie) alsnog wordt geaccepteerd, maar bij de veel grotere kans dat een populatie welke ruim aan de gestelde norm voldoet, alsnog wordt geaccepteerd.