

DE „OPTIMALE STEEKPROEFOMVANG”: EEN ILLUSIE?

door M. Vermaas

1. Inleiding

Bij een beschouwing over de optimale steekproefomvang ligt het voor de hand te beginnen bij de publikatie van Van Heerden in het MAB van december 1961. Collega J. H. Blokdijk noemde in het MAB van juni 1969 (blz. 298) de publikatie van Van Heerden nog een geslaagde poging de omvang van de steekproef op objectieve wijze vast te stellen, namelijk op de omvang waarbij de kosten verbonden aan het verder uitbreiden van de controle groter worden dan de daarmee corresponderende vermindering van de maximale „resterende onzekerheid”.

In 1974 verscheen het interim rapport van de subcommissie Controle en Steekproeven (CCS) van de Commissie Controlevraagstukken van het NIVRA; Blokdijk is voorzitter van de CCS. In dat interim-rapport valt te lezen: „Over de aanvaardbaarheid van de premissen van de formule van Van Heerden, en daarmee over de mogelijkheid tot objectieve vaststelling van de steekproefomvang wordt in de commissie . . . uiteenlopend gedacht”.

Nu zijn in 1979 publikaties verschenen waarbij leden van de CCS betrokken zijn en waarin de „methode Van Heerden” opnieuw ter sprake komt.

In de eerste plaats is er het boek „Steekproeven in de Accountantscontrole” van prof. drs. J. Kriens en drs. A. C. Dekkers · lid CCS. Op blz. 17 van dat boek wordt gesteld dat de betekenis die aan de formule van Van Heerden moet worden toegekend, voor de accountant die in zijn praktijk steekproeven wil toepassen, vrij beperkt is. In de tweede plaats is er het artikel van collega R. Metzger · lid CCS · in het april nummer 1979 van het MAB. In dat artikel doet Metzger een nieuwe poging door afweging van nut en kosten de optimale steekproefomvang te bepalen, voortbouwend op de „methode Van Heerden”.

2. Probleem niet opgelost

Kennelijk was met de „methode Van Heerden” het probleem van de optimale steekproefomvang niet voor iedereen opgelost. Ook voor mij niet; dat zal ik eerst demonstrenen. We kiezen daartoe het volgende voorbeeld:

- een te controleren populatie van 10 miljoen eenheden;
- een eenheid is één gulden;
- de kosten van de steekproef zijn per trekking f 6,03.

De omvang van de steekproef wordt volgens de formule van Van Heerden:

$$\sqrt{\frac{10 \text{ miljoen}}{2,71828 \times 6,03}} = 781 \text{ trekkingen}$$

Het steekproefschema luidt dan als volgt:

steekproefomvang	:	781 trekkingen
0 fouten gevonden	:	goedkeuren
1 of meer fouten gevonden	:	?

Uit het vraagteken in het steekproefschema blijkt wel dat er sprake is van een niet opgelost probleem. Van Heerden werpt (in paragraaf 12 van zijn publikatie) wel de vraag op wat gedaan moet worden indien één of meer fouten gevonden worden. Een werkelijk antwoord geeft hij echter niet. Hij wijst slechts op een *andere* methode, nl. die van „sequential sampling”. De eerste steekproefserie moet dan groter worden dan volgens zijn formule, zo zegt Van Heerden, en bij het vinden van fouten kunnen dan nog aanvullende steekproefseries worden genomen. Dan hebben we echter n.m.m. het terrein van de „optimale steekproefomvang” verlaten.

Is het niet mogelijk het vraagteken in het steekproefschema te vervangen door het woord „afkeuren”, zo zou men kunnen vragen. Neen, moet ik antwoorden, want dat zou een onlogisch steekproefschema zijn. Voor de motivering van deze uitspraak verwijs ik naar paragraaf 5 van mijn in het november nummer 1979 van het MAB verschenen artikel „Steekproeven bij controles op ernstige fouten”.

3. Op zoek naar een betere weg

De in paragraaf 1, de inleiding, genoemde poging van Metzger is naar mijn oordeel nuttig omdat de probleemstelling goed uit de doeken wordt gedaan. Bij de uitwerking komt Metzger echter volgens mij niet tot de juiste conclusie.

3.1. De probleemstelling

In het artikel van Metzger komen de gebreken waaraan de „methode Van Heerden” mank gaat, duidelijk aan de orde:

- a. De populatie kan bij de „methode Van Heerden” slechts worden geaccepteerd indien geen enkele fout in de steekproef wordt ontdekt. Metzger streeft ernaar ook een steekproefplan mogelijk te maken waarbij wordt goedgekeurd bij het vinden van een beperkt aantal fouten.
- b. Het belang van het (niet) constateren dat een bepaald bedrag aan fouten in de populatie schuilt, wordt door Van Heerden gelijk gesteld aan dat bedrag. Metzger wijst er m.i. terecht op, dat het belang, in procenten uitgedrukt, toeneemt naarmate het bedrag aan fouten groter is.
- c. De kans dat een bepaald bedrag aan fouten optreedt, moet volgens Van Heerden ook in aanmerking worden genomen (daarmee is Blokdijk het niet eens; zie blz. 310 van zijn MAB-artikel, juni 1969), maar de praktische oplossing is door Van Heerden niet uitgewerkt. Volgens Metzger moet, om de optimale steekproefomvang te kunnen bepalen, de reeks van mogelijke foutenpercentages met hun respectieve kansen van optreden, worden geschat.
- d. Het risico dat een populatie op grond van de steekproefuitkomst ten onrechte wordt afgekeurd is niet in de formule van Van Heerden verwerkt. Metzger geeft aan hoe wel met dat risico rekening zou kunnen worden gehouden.

3.2. De voorgestelde methode

Aan de hand van een uitgewerkt voorbeeld introduceert Metzger een methode waarbij aan de sub 3.1. opgesomde bezwaren tegen de „methode Van Heerden” wordt tegemoetgekomen. Metzger pretendeert niet het laatste woord over het probleem van de optimale steekproefomvang te hebben gesproken; hij houdt rekening met discussie.

In zijn voorbeeld geeft Metzer, om de optimale steekproefomvang te kunnen becijferen, antwoord op de volgende vragen:

1. Wat is het meest waarschijnlijke foutenpercentage?
2. Wat is met 95% waarschijnlijkheid het maximale foutenpercentage?
3. Hoe groot is het geldelijk belang · in procenten aan te geven · bij het constateren dat een bepaald bedrag aan fouten in de populatie schuilt?
4. Wat zijn de controlekosten per te controleren eenheid?
5. Wat is het schadebedrag indien een als aanvaardbaar te beschouwen populatie toch wordt afgekeurd?

Verder wordt er door Metzer in zijn voorbeeld van uitgegaan dat de mogelijke foutenpercentages „normaal verdeeld” zijn.

3.3. Argwaan

In het door Metzer uitgewerkte voorbeeld heeft men · voordat met de steekproef wordt begonnen · al een duidelijke „verwachtingswaarde” toegekend aan de kwaliteit van de te onderzoeken populatie. De meest waarschijnlijke foutenfractie is 0,25% en de maximale foutenfractie is met 95% waarschijnlijkheid, m.a.w. „zo goed als zeker”, 1%. Daardoor sta ik argwanend tegenover de methode, want de vraag dringt zich aan mij op of een accountant die over middelen beschikt om een zo duidelijk beeld van de kwaliteit der populatie te verkrijgen, daarna nog behoefte heeft aan een statistische steekproef, vooral indien hij kosten en nut van de steekproef tegen elkaar wil afwegen.

3.4. De proef op de som

Metzer komt in zijn voorbeeld, via de door hem voorgestelde methode, tot een optimale steekproefomvang van 151 trekkingen, waarbij de populatie wordt geaccepteerd indien niet meer dan 2 fouten in de steekproef worden gevonden. Hij geeft bij die uitkomst een „proef op de som”, welke ons een goede gelegenheid biedt om de methode te beoordelen.

De „proef op de som” verloopt in principe als volgt:

Eventuele totale „fout” in de populatie van f 10 miljoen (dat is dus een foutenfractie van 1,457%!) Kans dat populatie toch wordt geaccepteerd	f 145.695 0,623 — ×
Resteert een risico van Kans dat de „fout” optreedt	f 90.768 0,415 — ×
Resteert een risico van Belang bij de „fout”	f 37.669 0,200 — ×
Resterend risico („restrisico”)	f 7.534 —

Het bedrag van f 7.534 is het „restrisico” dat bij de steekproef wordt gelopen. Dat „restrisico” is afhankelijk van de steekproefomvang. Bij de steekproefomvang van 151 trekkingen wordt een optimum bereikt zoals Metzer in de „proef op de som” laat zien:

<i>steekpr. omvang</i>	<i>„restrisico”</i>	<i>afname „restrisico”</i>	<i>Meerkosten steekproef (f 50 per trekking)</i>
140	f 8.134		
150	f 7.584	f 550	f 500
151	f 7.534	f 50	f 50
160	f 7.110	f 424	f 450

De „proef op de som” laat zien dat Metzger op de verkeerde weg is beland. Bij de weergegeven berekening van het „restrisico” ad f 7.534 wordt uitgegaan van een foutenfractie van 1,457%. Nu moet het wel verbazing wekken, dat de „kans dat de fout optreedt” in de berekening op 0,415 = 41,5% wordt gesteld. Immers, vooropgesteld is, dat de maximale foutenfractie met 95% waarschijnlijkheid 1% bedraagt, m.a.w. dat er slechts 5% kans is dat de foutenfractie meer dan 1% beloopt. Een foutenfractie van 1,457% heeft dan in werkelijkheid slechts 0,579% kans op te treden.

Ook de in de berekening opgenomen factor 0,200 = 20% voor het „belang van de fout” is niet juist. Bij een totale „fout” van f 145.695 dient het „belang van de fout” volgens de in het voorbeeld van Metzger gegeven cijfers op 96% te worden gesteld.

Het „restrisico” van f 7.534 in de „proef op de som” verandert zodoende in $f 90.768 \times 0,00579 \times 0,96 = f 505$.

We herhalen het in het voorgaande gegeven overzicht, maar nu met gecorrigeerde „restrisico’s”:

<i>steekpr. omvang</i>	<i>„restrisico”*)</i>	<i>toename „restrisico”</i>	<i>Meerkosten steekproef</i>
140	f 307		
150	f 488	f 181	f 500
151	f 505	f 17	f 50
160	f 752	f 247	f 450

*) Voor degene die het met het artikel van Metzger in de hand wil narekenen, geven we de opstelling in een bijlage.

De „proef op de som” blijkt nu helemaal niet meer te kloppen. Het „restrisico” zou stijgen (!) bij toename van de steekproefomvang. Dat zou betekenen dat de kosten van de steekproef toenemen, terwijl het nut ervan afneemt. Van een optimum bij een steekproefomvang van 151 trekkingen is geen sprake.

De fout in „de proef op de som” vinden we ook terug in de som van Metzger.

3.5. De foute som

Metzger gebruikt bij de aangeprezen methode ter bepaling van de optimale steekproefomvang de volgende formule:

$$n = \sqrt{(1-r)np \times (\text{kans} \times \text{belang}) \times \frac{T}{k}}$$

De factoren T (= omvang populatie) en k (= kosten per trekking) zijn geheel onafhankelijk.

De andere factoren zijn echter volgens Metzger aan voorwaarden gebonden; we zullen dat laten zien aan de hand van Metzger's voorbeeld, waarbij hij als eis voor acceptatie heeft gesteld, dat ten hoogste 2 fouten in de steekproef mogen worden gevonden:

1. (kans $\times$ belang) moet maximaal zijn; in het voorbeeld is dat het geval bij $41,32\% \times 20\% = 0,083$;
2. de sub 1 vermelde factoren behoren bij p (= de foutenfractie) = 0,5%; derhalve is $p = 0,5\%$;
3. ($1-r$) moet, bij de eis van ten hoogste 2 fouten in de steekproef, op 0,623 gesteld worden;
4. eveneens in verband met de eis van ten hoogste 2 fouten in de steekproef,

$$\text{moet } np \text{ 2,2 zijn; } p = 0,5\% \text{ (zie sub 2), dus } n = \frac{2,2}{0,005} = 440.$$

Bij invullen krijgen we dan de foute som:

$$440 = \sqrt{0,623 \times 440 \times 0,005 \times 0,4132 \times 0,2 \times \frac{10 \text{ mln}}{50}}$$

$$440 = 151$$

Metzger doet het als volgt:

$$n = \sqrt{0,623 \times 2,2 \times 0,4132 \times 0,2 \times \frac{10 \text{ mln}}{50}}$$

$$n = 151$$

De factor 2,2 onder het wortelteken, bestaat uit het produkt van n (= steekproef-omvang) en p (= foutenfractie); doordat Metzger het houdt op $n = 151$ in plaats van 440 verschijnt 1,457% als foutenfractie in plaats van 0,5%. Die foutenfractie van 1,457% hoort hier echter, zoals we in de „proef op de som” zagen, helemaal niet thuis:

- a. de foutenfractie van 1,457% behoort niet bij de ingezette factoren $0,4132 \times 0,2$ (zie sub 2 in het voorgaande);
- b. vooropgesteld was dat, met 95% waarschijnlijkheid, de maximale fout 1% bedroeg.

3.6. Geen vergissing, maar wel fout

Nu moet niet worden gedacht dat Metzger per abuis een misrekening heeft gemaakt. Bovenaan blz. 163 van zijn artikel zegt hij, dat hij in de formule voor het produkt van kans $\times$ belang steeds dezelfde (maximale) waarde zal opnemen, ongeacht welke foutenfractie hij in zijn formule opneemt. Hij doet dat kennelijk omdat anders het invullen van de formule een „heksentoer” wordt. Metzger voegt er aan toe, dat zo „in ieder geval een conservatieve uitkomst wordt verkregen”.

Het is geen bezwaar in een berekening een vereenvoudiging aan te brengen, indien de uitkomst een acceptabele benadering van de formeel juiste uitkomst oplevert, zodat men toch tot de juiste conclusie komt. De vereenvoudiging van Met-

zer leidt echter tot een volkomen averechtse conclusie. We zullen straks zien dat een berekening waarbij aan de mogelijke foutenfracties de juiste daarbij behorende kansen van optreden alsmede de werkelijk daaraan toegekende „percentages belang” worden verbonden, aantoonst, dat bij de in het voorbeeld van Metzer gepresenteerde gegevens, geen enkele steekproef meer rendabel is. Metzer krijgt echter als uitkomst een steekproefomvang van 151 trekkingen à f 50 per trekking. Dat is geen conservatieve uitkomst, maar een foute uitkomst; een *optimale* steekproefomvang is het in ieder geval niet.

3.7. De argwaan bevestigd

We beginnen met een overzicht. De eerste drie kolommen zijn overgenomen uit het artikel van Metzer; de vierde kolom voegde ik er aan toe.

1 <i>Bedrag van de fout is minstens</i>	2 <i>kans op „fout”</i>	3 <i>belang bij „fout”</i>	4 <i>produkt van kolom 1, 2 en 3 = maximaal „restrisico”</i>
(f)	(%)	(%)	(f)
25.000	70,66	10	1.767
50.000	41,32	20	4.132
75.000	19,38	40	5.814
100.000	7,13	50	3.565
125.000	2,03	80	2.030
150.000	0,45	100	675
175.000	0,07	140	172

In het door Metzer uitgewerkte voorbeeld blijkt het nut van de steekproef niet boven f 5.814 te kunnen liggen (zie kolom 4). Aangezien een steekproef het „restrisico” nooit tot nul kan reduceren en bovendien nog rekening moet worden gehouden met het „afkeurrisico”, is het nut van de steekproef minder dan f 5.814. In het voorbeeld van Metzer zijn de kosten per trekking f 50, zodat de optimale

steekproefomvang kleiner moet zijn dan $\frac{5814}{50} = 116$ trekkingen.

We gaan na, met welke bedragen de „restrisico's” door een steekproef van 115 trekkingen worden gereduceerd en hoe groot daarbij het „afkeurrisico” is. Een groter aantal trekkingen zou immers - zoals we zagen - niet rendabel zijn. Als eis voor acceptatie houden we weer aan dat ten hoogste 2 fouten in de steekproef mogen worden gevonden.

De acceptatiekansen blijken uit het volgende overzicht:

<i>Bedrag van de „fout”</i>	<i>Foutenfractie</i>	<i>Acceptatiekansen</i>	
(f)	(%)		
25.000	0,25	$0,9975^{115} \times 0,0025^0 \times \binom{115}{0} = 0,7499$	
		$0,9975^{114} \times 0,0025^1 \times \binom{115}{1} = 0,2161$	
		$0,9975^{113} \times 0,0025^2 \times \binom{115}{2} = 0,0309$	
			0,9969
50.000	0,50	zelfde rekenwijze	0,9796
75.000	0,75	idem	0,9438
100.000	1,00	idem	0,8910
125.000	1,25	idem	0,8252
150.000	1,50	idem	0,7512
175.000	1,75	idem	0,6733

N.B. $\binom{115}{2}$ noemt men een binomiaalcoëfficiënt en het betekent

$$\frac{115 \times 114}{1 \times 2}$$

De „restrisico's” worden dan gereduceerd met:

$$(1-0,9969) \times f \ 1.767 = f \ 5$$

$$(1-0,9796) \times f \ 4.132 = f \ 84$$

$$(1-0,9438) \times f \ 5.814 = f \ 327$$

$$(1-0,8910) \times f \ 3.565 = f \ 389$$

$$(1-0,8252) \times f \ 2.030 = f \ 355$$

$$(1-0,7512) \times f \ 675 = f \ 168$$

$$(1-0,6733) \times f \ 172 = f \ 56$$

Het „afkeurrisico” is naar het voorbeeld van Metzger te waarderen op

$$(1-0,9969) \times f \ 20.000 = f \ 62.$$

Met een steekproef van 115 trekkingen wordt zodoende het volgende resultaat behaald:

$$\text{kosten } 115 \times f \ 50 \qquad f \ 5.750$$

$$\text{bij: „afkeurrisico”} \qquad f \ 62$$

$$\qquad \qquad \qquad f \ 5.812$$

$$\text{af: maximale reductie „restrisico”} \qquad f \ 389$$

$$\text{Nadelig saldo (!)} \qquad f \ 5.423$$

Nemen we een steekproef van minder dan 115 trekkingen, dan dalen zowel de kosten, als het „afkeurrisico” en de reductie van het „restrisico”. Het moet duidelijk zijn dat een steekproef groter dan $\frac{389}{50} = 7$ trekkingen, dan niet rendabel kan zijn.

Nu kunnen we weer - op analoge wijze als voor de 115 trekkingen - nagaan wat het resultaat wordt van een steekproef van 7 trekkingen. Dat resultaat luidt als volgt:

kosten $7 \times f\ 50$	f 350,00
bij: „afkeurrisico”	0,01
	f 350,01
af: maximale reductie „restrisico”	0,13
Nadelig saldo (!)	f 349,38

Er is geen steekproefomvang te vinden waarbij het „nut” de kosten overtreft. Dat is een bevestiging van onze intuïtieve argwaan: indien men reeds een duidelijk beeld heeft van de foutenfractie in de populatie, in de vorm van een kansverdeling der mogelijke foutenfracties, zoals Metzer die meent te kunnen geven, dan is een steekproef, welke dat beeld nog zou moeten verscherpen, niet meer rendabel.

4. Conclusie

Door de publikatie van Metzer wordt duidelijk over welke gegevens we zouden moeten beschikken om de optimale steekproefomvang te kunnen berekenen. Indien we echter over die gegevens beschikken, blijkt er geen optimale steekproefomvang meer te bestaan. Het lijkt op het zoeken naar de steen der wijzen.

De „optimale steekproefomvang” blijft dus (nog) een illusie! Daarom is het zo belangrijk een steekproef te beginnen met de *minimale* omvang en die steekproef, indien nodig, uit te breiden. Daarvoor moge ik verwijzen naar mijn publikaties in het mei- en november nummer 1979 van dit maandblad.

BIJLAGE

Becijfering „restrisico's”

a	<u>140</u>	<u>150</u>	<u>151</u>	<u>160</u>
	(f)	(f)	(f)	(f)
b	157.143	146.666	145.695	137.500
c	<u>0,623</u> ×	<u>0,623</u> ×	<u>0,623</u> ×	<u>0,623</u> ×
	97.900	91.373	90.768	85.662
d	<u>0,00283</u> ×	<u>0,00551</u> ×	<u>0,00579</u> ×	<u>0,00975</u> ×
	277	503	526	835
e	<u>1,11</u> ×	<u>0,97</u> ×	<u>0,96</u> ×	<u>0,90</u> ×
	307	488	505	752
	=====	=====	=====	=====

a = steekproefomvang

b = bedrag van de fouten in de populatie van 10 miljoen gulden waarbij de kans op acceptatie bij de steekproefomvang (a) gelijk is aan (c)

c = acceptatiekans bij acceptatie van hoogstens 2 fouten in de steekproef (a)

d = kans op een bedrag aan fouten van minstens (b)

e = belang van het bedrag aan fouten (b)