STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

STATISTISCHE AFDELING

Rapport S 274

Toepassingen van aselecte steekproeven bij accountantscontroles

door

J. Kriens

oktober 1960

1. Inleiding

In veel gevallen waarin men een groot aantal gelijksoortige dingen moet controleren, beperkt men zich tot het onderzoeken van een gedeelte daarvan. Soms is dit noodzakelijk omdat door het onderzoek het voorwerp vernietigd of onbruikbaar wordt, doch meestal geschiedt dit omdat men wil besparen op de kosten verbonden aan het keuren. Dat gedeelte, dat wel gecontroleerd wordt, noemt men de steekproef en het aantal exemplaren dat er in aanwezig is, de omvang van de steekproef.

Steekproeven, waarbij alle exemplaren uit de verzameling of populatie een gelijke kans hebben om getrokken te worden noemt men aselecte steekproeven en het zijn steekproeven van dit type, waarmee in de industrie (bijv. bij het keuren van gemaakte produkten), in het medisch onderzoek (bijv. bij de vergelijking van geneesmiddelen), bij de marktanalyse (bijv. enquetes), voor bevolkingsschattingen en op nog talloos veel andere terreinen gewerkt wordt. Men kan zich dan ook afvragen, of aselecte steekproeven, die elders zoveel besparingen opleveren, ook niet van nut kunnen zijn bij het werk van accountants.

Uiteraard is de situatie bij een accountantscontrole niet zonder meer vergelijkbaar met een keuring van een partij goederen. Een belangrijk punt van verschil ligt in de konsekwenties, die het maken van onjuiste conclusies op grond van de steekproef kan hebben. Zo kan een ten onrechte goedgekeurde partij goederen in veel gevallen zonder al te grote kosten vervangen worden door een andere, zodra de fouten ontdekt worden. Het niet ontdekken echter van bijvoorbeeld een fraude heeft daarentegen in het algemeen zeer aanzienlijke gevolgen. Een vergelijkbare situatie ontstaat soms bij het medisch onderzoek, waarin bij het vergelijken van geneesmiddelen door middel van een steekproef, het doen van een onjuiste keuze ernstige gevolgen zal hebben voor de volksgezondheid.

Het is duidelijk dat bij een steekproefonderzoek minder informatie verkregen wordt dan bij een volledige keuring.[1)]

----

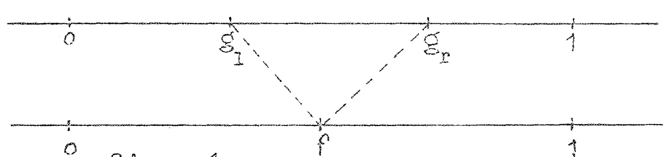
1) Strikt genomen geldt dit alleen, wanneer bij een onderzoek van alle eenheden, de keuring even zorgvuldig geschiedt als bij een zich tot een steekproef beperkend onderzoek. In gevallen waarin door monotonie van het werk of door vermoeidheid het onderzoeken van alle eenheden het risico met zich brengt van een minder gedegen keuring van iedere eenheid, behoeft het steekproefonderzoek niet noodzakelijk tot minder informatie te leiden dan volledig onderzoek.

De belangrijkste vraag betreft nu de aard van de conclusies, die wèl getrokken kunnen worden op grond van de informatie, verkregen door een steekproef. Deze vraag wordt in de mathematischestatistiek vanuit twee verschillende gezichtspunten onderzocht en beantwoord. De gedachtengang van beide redeneringen zullen wij hier in het kort weergeven.

Laten wij aannemen dat een lange rij posten gecontroleerd moet worden op onjuistheden. Een aantal posten zal in orde zijn en een aantal kan onjuist zijn. Er bestaat dus een bepaalde breuk p tussen nul en één, die aangeeft welk deel van de posten het kenmerk "onjuist" bezit. Men noemt deze breuk de fractie p (van de posten die het kenmerk "onjuist" bezitten) en het is deze fractie, waarover wij een uitspraak willen doen op grond van de resultaten van de steekproef. Stel eerst dat wij de grootte van p moeten schatten.

Indien wij een enkel getal op willen geven als schatting voor p, dan ligt het voor de hand hiervoor de fractie f van onjuiste getallen in de steekproef te nemen. Wanneer de steekproef een omvang van 100 had en er van de 100 posten 10 onjuist zouden zijn, dan zouden wij als schatting van p dus $f = \frac{10}{100} = 0,10$ opgeven. Zolang de steekproef niet alle posten omvat, behoeven f en p niet aan elkaar gelijk te zijn. Aan de andere kant zal p eerder een waarde in de omgeving van f bezitten, dan een waarde die sterk van f afwijkt. Men kan daarom een interval opgeven, waarbinnen p vermoedelijk ligt, een zogenaamd betrouwbaarheidsinterval.

Wij geven dus niet een bepaalde waarde op voor p, doch grenzen g_l en g_r waartussen de onbekende fractie p vermoedelijk ligt; vergelijk fig. 1. Zo zou men op grond van de gevonden f = 0,10 voor g_l kunnen opgeven 0,046 en voor g_r 0,186.



de fractie p in de populatie

de fractie f in de steekproef

fig. 1

Het verband tussen de fractie f in de steekproef en de fractie p in de populatie

Een betrouwbaarheidsinterval behoeft niet noodzakelijk een onder- en een bovengrens te bezitten. Men kan zich ook tot één van beide beperken, bijvoorbeeld door uit een steekproef de conclusie te trekken dat p niet groter is dan een bepaalde waarde. Zo zou de conclusie in het voorbeeld kunnen luiden: p is hoogstens gelijk aan 0,186.

Hoewel p vermoedelijk een waarde bezit, welke niet te veel verschilt van die van f kunnen wij toch andere waarden niet geheel uitsluiten. Zelfs is het in principe mogelijk dat onze steekproef alle onjuiste posten uit de lijst bevat; in dat geval kan p een waarde bezitten, die zeer veel kleiner is dan f. De conclusie die wij uit deze overwegingen trekken luidt, dat zelfs onze uitspraak in de vorm van een interval niet juist behoeft te zijn. Wel zullen wij in het algemeen minder onjuiste uitspraken doen, naarmate wij een groter interval opgeven, maar dit voordeel gaat dan samen met het nadeel van een minder precieze uitspraak over p. Wij zijn nu geïnteresseerd in het verband tussen de gevonden fractie f, de grenzen g_l en g_r van het betrouwbaarheidsinterval en de kans dat p werkelijk in dit interval ligt.

Tenzij wij ons beperken tot triviale mededelingen, zoals: p ligt tussen 0 en 1, zullen er in een lange reeks uitspraken over onbekende fracties (meestal) fouten gemaakt worden. Wij richten de vorm van onze uitspraken nu zo in, dat in deze reeks niet meer dan een bepaalde fractie uitspraken onjuist is. Hoe groot deze fractie gekozen wordt, hangt af van de aard van het onderzoek. Zo zal men deze fractie meestal lager kiezen bij proeven met een belangrijk geneesmiddel dan bij een vooronderzoek op het gebied van de marktanalyse. Veel gekozen waarden voor de toegelaten fractie onjuiste uitspraken zijn 0,01 respectievelijk 0,05. Deze fractie, die wij aangeven met α_0, noemt men de onbetrouwbaarheid. Staat deze eenmaal vast, dan stelt de mathematische statistiek ons weer in staat bij iedere steekproefomvang n en bij iedere gevonden fractie f in de steekproef, een betrouwbaarheidsinterval voor p te berekenen, zodanig dat de fractie onjuiste uitspraken gemiddeld gelijk is aan de gekozen onbetrouwbaarheid α_0. Op deze wijze is ook het bovenvermelde interval $0{,}046 \leq p \leq 0{,}186$ afgeleid bij een $\alpha_0 = 0{,}02$; dit interval kan opgegeven worden, wanneer de steekproefomvang n honderd bedraagt, de gekozen onbetrouwbaarheid 0,02 is en voor f gevonden is de waarde 0,10. Wanneer men alleen bovengrenzen g_r opgeeft, dan kan de uitspraak uiteraard slechts onjuist zijn doordat g_r te klein is. Trekt men derhalve uit de steekproef de conclusie dat p kleiner of hoogstens gelijk is aan 0,186, dan zal de kans op een onjuiste uitspraak minder worden dan 0,02 en wel kan men aantonen dat deze kans dan niet meer bedraagt dan de helft van 0,02, dus 0,01.

Bij de werkzaamheden van de accountant zijn in het algemeen slechts de bovengrenzen g_r voor de onbekende fracties p van belang. In de tabellen I en II zijn deze bovengrenzen opgegeven voor onbetrouwbaarheden

heden

van respectievelijk $\alpha_0 = 0{,}01$ en $\alpha_0 = 0{,}05$. Voor de steekproefomvang zijn gekozen de waarden n=50, n=100, n=200, n=300, n=400 en n=500. In de eerste kolom is het aantal onjuistheden k opgegeven, dat in de steekproef gevonden is; bij het gebruik van deze tabellen behoeft men dus niet eerst de steekproeffractie $f = \frac{k}{n}$ te berekenen. Er dient met nadruk op gewezen te worden, dat deze tabellen alleen gebruikt kunnen worden, wanneer de steekproef slechts een klein deel van de gehele populatie omvat.

Tabel I[1)]

Bovengrenzen voor de fractie onjuistheden in de populatie bij $\alpha_0 = 0{,}01$

aantal onjuistheden in de steekproef k	steekproefomvang n					
	50	100	200	300	400	500
0	0,053	0,027	0,014	0,009	0,007	0,006
1	0,106	0,054	0,028	0,018	0,014	0,011
2	0,141	0,073	0,037	0,025	0,019	0,015
3	0,172	0,089	0,045	0,031	0,023	0,018
4	0,201	0,104	0,053	0,036	0,027	0,021
5	0,228	0,119	0,061	0,041	0,031	0,024
6	0,254	0,133	0,068	0,046	0,034	0,027
7	0,279	0,147	0,075	0,050	0,038	0,030
8	0,304	0,160	0,082	0,055	0,042	0,033
9	0,328	0,173	0,089	0,060	0,045	0,036
10	0,351	0,186	0,095	0,064	0,048	0,039
11	0,374	0,198	0,101	0,068	0,051	0,042
12	0,396	0,210	0,108	0,073	0,054	0,044
13	0,418	0,222	0,114	0,078	0,058	0,046
14	0,440	0,234	0,121	0,082	0,061	0,049
15	0,462	0,246	0,127	0,086	0,064	0,052
16	0,483	0,258	0,133	0,090	0,067	0,054
17	0,504	0,270	0,140	0,094	0,070	0,056
18	0,524	0,281	0,146	0,098	0,073	0,059
19	0,544	0,292	0,151	0,102	0,076	0,062
20	0,564	0,303	0,157	0,106	0,079	0,064

----

1) De tabellen I en II zijn overgenomen uit: H.C. HAMAKER: "Average confidence" limits for binomial probabilities, Review of the International Statistical Institute 21 (1953), 17-27.

Tabel II

Bovengrenzen voor de fractie onjuistheden in de populatie bij $\alpha_0 = 0{,}05$

aantal onjuistheden in steekproef k	steekproefomvang n					
	50	100	200	300	400	500
0	0,037	0,019	0,010	0,006	0,005	0,004
1	0,074	0,038	0,020	0,013	0,010	0,008
2	0,106	0,054	0,028	0,018	0,014	0,010
3	0,134	0,068	0,034	0,024	0,018	0,014
4	0,161	0,082	0,042	0,028	0,021	0,016
5	0,186	0,096	0,048	0,032	0,024	0,020
6	0,211	0,108	0,055	0,037	0,028	0,022
7	0,235	0,122	0,062	0,041	0,030	0,025
8	0,258	0,134	0,068	0,046	0,034	0,028
9	0,282	0,146	0,074	0,050	0,038	0,030
10	0,304	0,158	0,080	0,054	0,040	0,032
11	0,326	0,170	0,086	0,058	0,044	0,035
12	0,348	0,182	0,093	0,062	0,046	0,038
13	0,370	0,192	0,099	0,066	0,050	0,040
14	0,392	0,204	0,104	0,070	0,052	0,042
15	0,414	0,216	0,110	0,074	0,056	0,044
16	0,434	0,227	0,116	0,078	0,058	0,047
17	0,455	0,238	0,122	0,082	0,062	0,049
18	0,476	0,250	0,128	0,086	0,064	0,052
19	0,496	0,262	0,134	0,090	0,068	0,054
20	0,516	0,273	0,139	0,094	0,070	0,056

Een vergelijking van tabel I met tabel II leert dat bij een gelijke steekproefomvang en een gelijk aantal onjuistheden, het interval voor een kleinere onbetrouwbaarheid langer is dan voor een grotere onbetrouwbaarheid. Ook kan men aan de hand van eenvoudige berekeningen laten zien dat bij gelijke onbetrouwbaarheid en een gelijke fractie onjuistheden in de steekproef een korter interval opgegeven wordt naarmate de steekproefomvang n groter is; zo lezen wij uit tabel I af dat vóor n=50 en k=2 de bovengrens voor p 0,141 bedraagt, voor n=100 en k=4 is deze bovengrens 0,104 en voor n=500 en k=20 vinden wij 0,064. Beide conclusies stemmen overeen met hetgeen men op intuïtieve gronden verwacht.

Resumerende kunnen wij vaststellen dat wij het volledige onderzoek teneinde een onbekende fractie p te bepalen, hebben vervangen door het nemen van een steekproef die ons in staat stelt een interval op te geven waarbinnen p vermoedelijk ligt, terwijl het bovendien mogelijk is gebleken deze intervallen zodanig te kiezen dat de fractie onjuiste uitspraken in een lange reeks gemiddeld gelijk is aan een van te voren vastgelegde onbetrouwbaarheid α_0. De tabellen I en II geven voor verschillende steekproefomvang bovengrenzen voor p bij onbetrouwbaarheden $\alpha_0=0{,}01$ en $\alpha_0=0{,}05$ en kunnen gebruikt worden zolang de steekproef klein is ten opzichte van de gehele populatie. Voor het probleem een schatting van p te geven op grond van een steekproef is dus een oplossing gevonden. In paragraaf 3 behandelen wij enige voorbeelden.

Naast het probleem een onbekende fractie p te schatten, doet zich soms een ander, hiermee verwant probleem voor. In veel gevallen moet van een partij goederen of een lijst met posten onderzocht worden of de kwaliteit zodanig is, dat ze aanvaard kan worden. Men is dan niet zozeer geïnteresseerd in een betrouwbaarheidsinterval voor de fractie p, doch meer voor het risico een partij te aanvaarden, wanneer deze eigenlijk afgekeurd moest worden. Zolang men niet alle eenheden keurt is de mogelijkheid een partij ten onrechte te aanvaarden steeds aanwezig, omdat de steekproef bijvoorbeeld geen enkel slecht of onjuist exemplaar kan bevatten, terwijl er in de hele partij of populatie toch een groot aantal aanwezig zijn.

Men kan nu alvorens de partij of lijst te accepteren eerst een aselecte steekproef nemen van de omvang n en nagaan hoeveel exemplaren daarin defect zijn. Dit aantal geven wij aan met k. De partij wordt afgekeurd wanneer dit aantal groter of gelijk is aan k_0 en goedgekeurd wanneer dit aantal kleiner is dan k_0. Wanneer een partij nu onaanvaardbaar is, als de fractie defecten p_0 of meer bedraagt, dan zal de kans β op het ten onrechte accepteren van de partij afhangen van de steekproefomvang n, de afkeurgrens k_0 en de waarde van p_0. Het verband tussen deze grootheden wordt weer berekend in de mathematische statistiek. Eén en ander illustreren wij met het volgende voorbeeld.

Stel dat men een lijst met posten onaanvaardbaar acht, wanneer er in een fractie groter of gelijk aan $p_0=0{,}10$ onjuistheden voorkomen. Stel verder dat de lijst wordt afgekeurd, wanneer men in een steekproef van $n=100$ posten $k_0=7$ of meer onjuiste posten aantreft; zijn er dus minder dan 7 onjuiste posten in de steekproef, dan aanvaardt men de lijst.

Het risico zit nu in de mogelijkheid dat de steekproef minder dan 7 onjuiste posten bevat, terwijl toch de fractie onjuiste posten in de gehele lijst meer dan 0,10 bedraagt. De lijst wordt dan ten onrechte geaccepteerd. De kans hierop kan berekend worden en bedraagt hoogstens $\beta = 0{,}12$.

Gewoonlijk gaat men andersom te werk en eist men dat de kans op het ten onrechte goedkeuren hoogstens $\beta = \beta_0$ (bijv. weer 0,01 of 0,05) mag bedragen en stelt dan vast wat de grenswaarde k_0 moet zijn, waarbij de lijst afgekeurd dient te worden. Wil men lijsten in hoogstens een fractie $\beta_0 = 0{,}01$ van de controles accepteren, wanneer de fractie onjuiste posten meer dan 0,10 bedraagt, dan zal men moeten afkeuren, zodra in steekproeven van de omvang 100 het aantal fouten 4 of meer bedraagt; laat men een fractie $\beta_0 = 0{,}05$ toe, dan zal afkeuren dienen te geschieden, wanneer $k \geq 5$ is. Er bestaan talrijke tabellen, waarin de waarde van k_0 wordt opgegeven voor verschillende waarden van β_0, n en p_0.

Resumerende kan vastgesteld worden dat het volledige onderzoek om na te gaan of een partij goederen of lijst met posten aanvaardbaar is, vervangen kan worden door een steekproefonderzoek, waarbij men van te voren kan vaststellen welk risico op het ten onrechte accepteren van slechte partijen men toelaatbaar acht. Bijvoorbeeld kan de keuring door de keuze van de steekproefomvang n en de afkeurgrens k_0 zo ingericht worden dat van de onaanvaardbare partijen of lijsten hoogstens een fractie 0,01 doorgelaten wordt.

In de appendix wordt nader ingegaan op de wiskundige formulering van dit probleem, terwijl in paragraaf 4 de controle op fraude behandeld wordt als toepassing van de bovenstaande gedachtengang.

Opmerkingen

1. Men dient goed onderscheid te maken tussen een uitspraak in de vorm van een betrouwbaarheidsinterval en het type uitspraken zoals in het tweede deel van deze paragraaf behandeld is. Geeft men een lange reeks betrouwbaarheidsintervallen op met een gemiddelde onbetrouwbaarheid van 0,01, dan betekent dit dat in 1% van alle uitspraken een interval opgegeven wordt, dat de ware p niet zal bevatten. Formuleert men de conclusie in de vorm van het goed- of afkeuren van partijen en eist men dat de kans een slechte partij goed te keuren hoogstens 0,01 is, dan betekent dit, dat van die gevallen, waarin de partij slecht is, hoogstens 1% niet ontdekt zal worden. In het eerste geval betreft het aantal onjuiste uitspraken dus 1% van alle uitspraken en in het tweede

geval 1% van alle uitspraken, waarbij een slechte partij betrokken is.

2. Soms kan men uitspraken van beide typen doen op grond van dezelfde steekproef. Men kan dan concluderen dat de partij wat het éne kenmerk betreft goed- of afgekeurd dient te worden, terwijl men tevens een betrouwbaarheidsinterval opgeeft voor de fractie elementen, die een ander kenmerk bezit (zie ook opmerking 6 in paragraaf 4).

2. Steekproeven bij accountantscontroles

In deze paragraaf willen wij in het kort nagaan of de in paragraaf 1 beschreven methoden ook bij de controlewerkzaamheden van de accountant van nut kunnen zijn.

De accountant stuit regelmatig op situaties, waarin een oordeel geveld moet worden over "partijen goederen". Een "partij goederen" kan dan bijvoorbeeld bestaan uit een saldolijst van debiteuren, de in een jaar door een schadeverzekeringsmaatschappij uitgekeerde schaden of de lijst van weeklonen die in een fabriek aan de arbeiders uitbetaald moeten worden.

Wanneer men nu moet controleren of een dergelijke partij goederen kan worden aanvaard, staat men voor de vraag of alle posten afzonderlijk gecontroleerd dienen te worden, of dat met een gedeelte daarvan kan worden volstaan. Uiteraard is het in theorie steeds mogelijk om alles te controleren, doch of dit practisch gezien uitvoerbaar is en economisch gezien verantwoord, is in veel gevallen de vraag. Er zijn tal van situaties, waarin grote aantallen kleine posten gecontroleerd moeten worden en waarin de eventuele risico's die men loopt niet opwegen tegen de hoge kosten van de controle. Terecht heeft men dan ook vaak betwijfeld of het nagaan van alle afzonderlijke eenheden van een partij wel onder alle omstandigheden vereist is. Vergelijk bijvoorbeeld verschillende artikelen in de laatste jaargangen van het Maandblad voor Accountancy en Bedrijfshuishoudkunde. Helaas heeft men zich in de practijk bij het nemen van steekproeven vaak laten verleiden tot het volgen van niet, of niet voldoende gefundeerde methoden, welke niet toelaten uit de resultaten van de steekproef ondubbelzinnig vastliggende conclusies te trekken.

Wil men steekproeven toepassen, dan komen alleen die typen in aanmerking, welke gebaseerd zijn op moderne, wetenschappelijk gefundeerde statistische methoden. Hieronder vallen de in de eerste paragraaf genoemde aselecte steekproeven, die ons, zoals aldaar beschreven is, in staat stellen tot het trekken van exact geformuleerde conclusies. Wij leggen er de nadruk op dat deze conclusies mogelijk zijn in alle gevallen waarin men aselecte steekproeven neemt. De aard van het gecontroleerde materiaal heeft hierop geen enkele invloed, zodat het bijvoorbeeld onbelangrijk is of eventuele vergissingen alleen in een bepaald deel van een lijst voorkomen, of dat frauden op een bepaalde, systematische wijze verricht zijn.[1] De conclusie is derhalve dat vanuit statistisch standpunt gezien het nemen van aselecte steekproeven bij accountantscontroles volkomen gerechtvaardigd is. Of het in een concreet geval zinvol is dit te doen, hangt van de omstandigheden af. In de paragrafen 3 en 4 worden enige voorbeelden van toepassingen gegeven.

Opmerking

Reeds in paragraaf 1 werd opgemerkt dat bij een aselecte steekproef ieder element van de partij dezelfde kans moet hebben in de steekproef opgenomen te worden. Heeft men nu een lijst met posten, dan kan deze eis op twee verschillende manieren worden geïnterpreteerd.

Het ligt het meest voor de hand de posten van de lijst als elementen te beschouwen en er voor te zorgen dat iedere post een gelijke kans heeft in de steekproef gecontroleerd te worden. Alle posten worden dan als gelijkwaardig behandeld. Daarnaast kan men zich echter voorstellen dat men in sommige situaties de grote posten van meer belang acht dan de kleine en dat men die een grotere kans wil geven in de steekproef opgenomen te worden. Men kan dit doen en toch het karakter van een aselecte steekproef behouden door niet de afzonderlijke posten, maar alle in de lijst aanwezige guldens op te vatten als elementen van de lijst. Een aselecte steekproef is dan een steekproef, waarbij iedere gulden dezelfde kans heeft aangewezen te worden. Strikt genomen behoeft men bij het onderzoek dan alleen de aangewezen gulden te controleren; in de practijk zal men zich hier natuurlijk niet toe beperken, doch de gehele post onderzoeken, waartoe de betreffende gulden behoort. Dit betekent echter dat grote posten een grotere kans krijgen om gecontroleerd te worden dan kleinere en wel is het zo dat deze kans evenredig is met de grootte van de post. In de paragrafen 3 en 4 zullen wij zien, dat in het ene geval de opvatting van de lijst met posten meer aansluit bij de behoeften en in het andere geval de opvatting van de lijst met guldens meer geschikt is.

----

1) De door S. KLEEREKOPER in het artikel "De steekproeven als middel van accountantscontrole" in de literatuur, geformuleerde conclusies zijn dan ook onjuist; M.A.B. 10 (1933), blz. 54.

3. Accuratesse-controles

Men kan zich voorstellen, dat de eerste in paragraaf 1 beschreven gedachtengang, die van de betrouwbaarheidsintervallen, toegepast wordt bij een accuratesse-controle. Men neemt een aselecte steekproef van n posten uit het totale aantal, geeft op grond van de steekproefuitkomst een betrouwbaarheidsinterval op voor de fractie p van onjuistheden in alle posten en gaat vervolgens na of de gevonden grenzen aanleiding vormen tot een nader onderzoek of tot ingrijpen.

Wij geven nu enkele voorbeelden van de toepassing van betrouwbaarheidsintervallen bij accuratesse-controles.

1.Een fabriek heeft 2000 arbeiders in dienst, die wekelijks worden uitbetaald. Het loon moet iedere week opnieuw worden uitgerekend en om na te gaan of dit voldoende nauwkeurig geschiedt neemt men uit de 2000 berekeningen van een bepaalde week een aselecte steekproef van 50 stuks. Men treft twee fouten aan. In dit geval is dus n=50 en k=2. Wenst men te werken met een onbetrouwbaarheid van 0,01 (of 1%), dan kan uit tabel I worden afgelezen dat de fractie gemaakte fouten hoogstens 0,141 of ongeveer 14% bedraagt.

Wanneer men regelmatig steekproeven van dit type neemt bij hetzelfde bedrijf of bij verschillende bedrijven en steeds met behulp van tabel I een bovengrens opgeeft, dan betekent de onbetrouwbaarheid van 0,01 dus, dat gemiddeld in één op de honderd uitspraken een onjuiste grens wordt opgegeven. Stelt men lagere eisen en kiest men een onbetrouwbaarheid van 0,05, dan kan bij de gegeven steekproef als bovengrens voor p volgens tabel II opgegeven worden g_r=0,106 of 10,6%. Regelmatige toepassing van deze methoden leidt tot gemiddeld vijf onjuiste uitspraken op de honderd.

2. Wanneer men de kwaliteit van een loonadministratie iedere week controleert om na te gaan of het aantal gemaakte fouten toelaatbaar is, dan kan men zeer geschikt de tweede in paragraaf 1 geschetste gedachtengang toepassen. Laten wij aannemen dat iedere week een steekproef van 100 lonen gecontroleerd wordt en dat men een lijst met 5% foute berekeningen nog juist aanvaardbaar acht. Afhankelijk van het aantal fouten in de steekproef, zal de lijst met loonberekeningen geaccepteerd worden of niet, waarbij het laatste kan betekenen, dat men alle lonen laat narekenen. Keurt men nu een lijst af, wanneer er twee of meer fouten gevonden worden, dan zal men de lijst ten onrechte goedkeuren, wanneer de steekproef één of nul fouten bevat, terwijl toch de fractie fouten in de gehele lijst groter is dan 0,05. De kans hierop bedraagt 0,037 of 3,7%. Zou men lijsten met twee fouten ook nog accep-

teren, dan zou de kans op het ten onrechte goedkeuren 0,118 of ongeveer 12% bedragen. Dergelijke berekeningen kunnen uiteraard ook voor andere waarden van de steekproefomvang n en voor de toelaatbare grens p_0 gedaan worden.

Voor een uitvoeriger bespreking van deze problemen verwijzen wij naar een artikel van H.E.J. BOTJE. 1)

3. De penningmeester van een grote vereniging met meer dan 100.000 leden moet na 1 april aanmaningen sturen aan die leden, die hun contributie nog niet, of niet volledig betaald hebben. Om na te gaan of dit voldoende snel geschiedt en of de juiste bedragen op de aanmaningen vermeld staan, wordt omstreeks eind april een aselecte steekproef van 300 stuks genomen uit die leden die per 1 april nog niet betaald hadden. Men verricht de controles met behulp van de doorslagen van de aanmaningen en vindt hierbij vijf gevallen, waarin òfwel in het geheel niet gemaand is, òfwel een fout bedrag is opgegeven. Hieruit kan met een onbetrouwbaarheid van 0,05 de conclusie getrokken worden (vgl. tabel II), dat hoogstens een fractie $p \leqq 0{,}032$ van alle leden, die gemaand moesten worden geen of een onjuist bericht gekregen hebben.

Naast de hier gegeven voorbeelden zijn er ook nog talrijke andere controles, waarbij door middel van steekproeven inzicht verkregen kan worden in de accuratesse waarmee gewerkt is; men denke bijvoorbeeld aan het controleren van facturen en van een voorraadadministratie.

4. Fraude controles

Bij een controle op fraude kan in principe evenals bij accuratessecontroles met betrouwbaarheidsintervallen gewerkt worden. Uiteraard zal men zich beperken tot het opgeven van een bovengrens voor de fractie p van gefraudeerde posten en bovendien zal men de onbetrouwbaarheid α_0 klein kiezen. Wanneer men systematisch op deze wijze te werk gaat, dan kan men dus concluderen, dat van een groot aantal controles slechts in hoogstens een fractie α_0 van de gevallen een interval opgegeven wordt, dat de ware p niet bevat.

Hoewel deze gedachtengang vanuit statistisch oogpunt gezien correct is, zijn er toch twee belangrijke bezwaren tegen aan te voeren. In de eerste plaats is men niet alleen geïnteresseerd in de fractie gefraudeerde posten, maar vooral in het totale gefraudeerde bedrag. Ten tweede zal men bij een controle, waarin een geval van

1) H.E.J. BOTJE, Steekproefcontrole in de loonadministratie, Sigma 4 Nr 5, 1958, 93-98.

fraude gevonden wordt zeker niet volstaan met het controleren van een steekproef uit de posten, doch overgaan tot het controleren van alle posten. Beide bezwaren kunnen worden ondervangen.

Laten wij aannemen dat alleen gefraudeerd kan worden door het opschrijven van te hoge bedragen, het toevoegen van niet bestaande posten, het opschrijven van hogere totaalsommen dan de werkelijke of het verkeerd transporteren van deeltotalen. Voor deze situatie zijn twee methoden ontwikkeld waarin aan de genoemde bezwaren tegemoet gekomen wordt.

P. DE WOLFF [1] heeft voorgesteld de grote posten allemaal te controleren en de kleinere slechts gedeeltelijk; zodra een geval van fraude gevonden wordt, gaat men tot het controleren van alle posten over en men doet dus slechts waarschijnlijkheidsuitspraken, wanneer in de steekproef geen enkel geval van fraude gevonden wordt.

Een andere methode om de bezwaren te ondervangen is afkomstig van A. VAN HEERDEN. Bepaalde moeilijkheden, die bij de methode van DE WOLFF optreden, zijn bij de methode van VAN HEERDEN niet aanwezig en daarom zullen wij ons hier beperken tot een uitvoeriger bespreking van de laatstgenoemde methode. De gedachtengang is als volgt.

Laat de op fraude te controleren lijst bestaan uit N posten en laat voor het totale bedrag van alle posten opgegeven zijn B gulden. Wij beschouwen nu niet de posten als eenheden, maar de afzonderlijke guldens en handelen dus alsof een lijst met B guldens gegeven is, waarover een uitspraak gedaan moet worden op grond van een steekproef. Uit de lijst worden n guldens aselect aangewezen en de steekproefcontrole kan zich derhalve beperken tot het nagaan of deze n guldens inderdaad aanwezig zijn. In de meeste gevallen zullen deze guldens echter behoren tot posten. Wij controleren nu niet alleen de aangewezen guldens, doch de gehele posten waarvan deze guldens deel uitmaken. Grote posten hebben op deze wijze een grotere kans om aangewezen te worden dan kleinere en kunnen zelfs verschillende keren aangewezen worden. Men controleert dus hoogstens n posten. Zodra een geval van fraude gevonden wordt, gaat men alle posten controleren. De vraag is nu hoe groot n gekozen moet worden om de risico's voldoende klein te houden. Het risico voor de accountant bestaat hier uit het <u>niet</u> ontdekken van fraude als er in werkelijkheid wel gefraudeerd is

----

1) P. DE WOLFF, Steekproeven bij administratieve controle, Statistica Neerlandica <u>10</u> (1956), 35-44.

id. Produktiviteitsverhoging bij accountantscontrole door toepassing van gelaagde steekproeven, Statistica Neerlandica <u>13</u> (1959), 215-232.

en men kan dus een eis stellen van de volgende vorm: als er meer dan een fractie p_0 van het totale aantal guldens gefraudeerd is, dient de kans dat dit niet tijdens de controle gevonden wordt hoogstens β_0 te bedragen.

Deze gedachtengang sluit direct aan bij de aan het eind van paragraaf 1 ontwikkelde redenering. Immers, ook hier interesseren wij ons voor de kans β, dat de lijst ten onrechte goedgekeurd wordt, dat wil zeggen, de kans dat fraude niet ontdekt wordt, terwijl er in werkelijkheid wel gefraudeerd is. Op intuïtieve gronden kan men inzien dat deze kans afneemt, naarmate er meer gefraudeerd is en naarmate de steekproefomvang n groter wordt. In figuur 2 is de kans β op het niet ontdekken van fraude voor verschillende waarden van de steekproefomvang n geschetst als functie van de fractie gefraudeerde guldens p. Daar deze kans van p afhangt, schrijven wij in plaats van β voor de duidelijkheid $\beta(p)$. Uit de figuur blijkt dat $\beta(p)$ inderdaad sterk afneemt voor toenemende p en toenemende n. De keuze van n

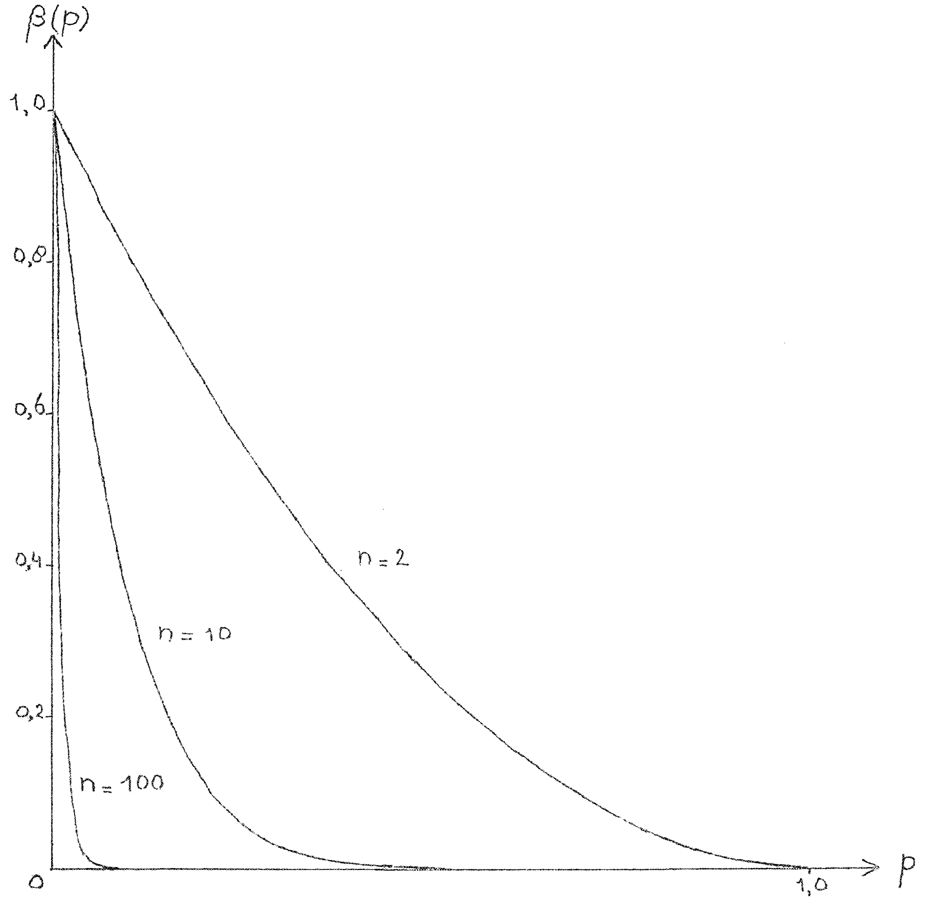


fig.2

De kans op het niet ontdekken van fraude bij een steekproefcontrole als functie van de fractie gefraudeerde guldens p en de steekproefomvang n.

is nog vrij en daarvan kunnen wij gebruik maken om aan $\beta(p)$ een bepaalde eis op te leggen. Het ligt voor de hand te eisen dat een fraude groter dan een fractie p_0 slechts een kleine kans heeft niet ontdekt te worden. Bijvoorbeeld kan men stellen dat deze kans hoogstens 0,01 mag bedragen voor $p_0 = 0{,}01$; voor deze keuze van β_0 en p_0 kan berekend worden dat n minstens 459 moet bedragen (zie de appendix). Aangezien ook andere waarden dan 0,01 voor p_0 en β_0 gekozen kunnen worden, zijn in tabel III waarden van n opgegeven voor $p_0=0{,}05$; 0,01 en 0,001 en $\beta_0=0{,}05$; 0,01 en 0,001.

Tabel III

Waarden van n voor verschillende waarden van p_0 en β_0

p_0 \ β_0	0,05	0,01	0,001
0,05	59	90	135
0,01	299	459	688
0,001	2995	4603	6905

Men leest uit tabel III af dat β_0 nog aanzienlijk verlaagd kan worden zonder een al te sterke toename van n, maar dat men bij hogere eisen ten opzichte van p_0 zeer omvangrijke steekproeven moet gaan nemen.

Conclusie

Wanneer wij aselect 459 guldens aanwijzen uit een totaal van B guldens, vervolgens de posten controleren, waartoe deze guldens behoren en tot het controleren van alle posten overgaan, zodra een geval van fraude gevonden wordt, dan is de kans dat een fraude groter dan een fractie 0,01 van het totaalbedrag niet ontdekt wordt hoogstens 0,01. Nog iets anders geformuleerd: passen wij de hier beschreven controlemethode met behulp van aselecte steekproeven toe, dan zal in een lange reeks controles hoogstens 1% van de gevallen waarin meer dan 1% gefraudeerd is, niet worden ontdekt. Hierbij kan nog opgemerkt worden dat het risico van het niet ontdekken van fraude in sommige gevallen nog aanzienlijk lager kan liggen (vergelijk opmerking 1).

Opmerkingen

1. Dáar wij niet alleen de aangewezen guldens controleren maar de gehele posten waartoe zij behoren, wordt er meer gecontroleerd dan alleen de n guldens, waarop de berekeningen zijn gebaseerd. De kans op het niet ontdekken van een fraude van meer dan 1% is dan ook

slechts in die gevallen precies 0,01, waarin voor alle posten geldt dat ze òf in het geheel niet, òf volledig gefraudeerd zijn. In die gevallen waarin posten gedeeltelijk gefraudeerd zijn (bijv. door het opgeven van een te hoog bedrag) kan de kans nog aanzienlijk lager zijn dan 0,01. Hoe klein deze kans precies is, valt zonder het maken van extra onderstellingen niet te berekenen.

2. De fraude is hier steeds uitgedrukt als fractie van het opgegeven totaalbedrag B. In gevallen waarin gefraudeerd is, bedraagt het werkelijke totaal B', een bedrag dat kleiner is dan B. De fraude moet dan uitgedrukt worden als een fractie van B! Dit kan zonder veel moeite, omdat een fraude-fractie p in B' gelijk is aan een fraude-fractie $\frac{p}{1+p}$ in B en wij zouden dus in het voorgaande p_0 kunnen vervangen door $\frac{p_0}{1+p_0}$. Voor kleine waarden van p_0 heeft deze substitutie echter slechts weinig invloed; zo zou voor $p_0=0,01$ en $\beta_0=0,01$ gevonden worden n=463 in plaats van n=459. Gezien het in de opmerkingen 1 en 5 gezegde is deze correctie weggelaten.

3. De steekproeven waarover voortdurend gesproken wordt, zijn aselecte steekproeven, dus steekproeven waarbij ieder van de B guldens dezelfde kans heeft om aangewezen te worden. Men kan hiervoor zorgen door te werken met lijsten met aselecte getallen; heeft men bijvoorbeeld een totaalbedrag B=30.000 gulden, dan gebruikt men een lijst met n aselecte getallen tussen nul en 30.000. Hoewel het in principe mogelijk is alle controles te verrichten met één lange lijst van aselecte getallen (die dan meer dan n aselecte getallen dient te bevatten) is het eenvoudiger aparte lijsten van n getallen op te stellen tussen bijvoorbeeld 0 en 10.000, tussen 0 en 15.000, enzovoorts.

In veel gevallen is van de lijst met posten niet alleen het totaalbedrag B bekend, maar zijn ook allerlei deeltotalen gemakkelijk te vinden. Het opzoeken van de aselect aangewezen guldens kan men dan vereenvoudigen door lijsten op te stellen met n aselecte getallen, die reeds naar grootte gerangschikt zijn. Men dient dus te beschikken over lijsten met n gerangschikte aselecte getallen tussen 0 en 10.000, en 15.000, enzovoorts tot en met 0 en 95.000. Deze lijsten kunnen ook gebruikt worden, wanneer het totaalbedrag meer dan 100.000 gulden is; bij bijv. B = 200.000 gebruikt men dan de lijst van 0 tot 20.000 en wijst men tientjes aan in plaats van guldens.

4. Wanneer allerlei deeltotalen uit de lijst bekend zijn, behoeft men bij het opzoeken van een aangewezen gulden niet vanaf het

begin te tellen. Laten wij onderstellen dat deze deeltotalen de van de ene naar de volgende bladzijde getransporteerde bedragen zijn. Moet men nu de 6.530^{ste} gulden controleren, dan zoekt men naar de bladzijde waarop bovenaan een getal onder de 6.530 staat en onderaan een getal dat hierboven ligt en men gaat op deze bladzijde van bovenaf tellen tot de gezochte gulden gevonden is.

Een fraude in de optelling van de pagina vindt men doordat de bladzijde, waarop bijv. bovenaan geschreven is 6.000 en onderaan 6.800 bij het tellen slechts 500 guldens blijkt te bevatten; de aangewezen gulden is dan een gulden uit een volledig gefraudeerde "post van 300 gulden". Een fraude, gepleegd door fout te transporteren wordt gevonden, wanneer bij het zoeken naar de 6.530^{ste} gulden blijkt dat de ene pagina met 6.500 eindigt, terwijl de volgende met bijv. 7.000 begint. Met een op de geschetste wijze verrichte controle worden dus ook fraudes in optellingen en bij het transporteren van bedragen gecontroleerd.

5. Aanknopende bij opmerking 3 kan men het risico een fraude niet te ontdekken nog verder verkleinen door in die gevallen, waarin de sprong tussen twee in grootte opeenvolgende aselecte getallen meer is dan 1% van het totaalbedrag B, aselect een getal bij te loten tussen de twee reeds aanwezige aselecte getallen in. In het algemeen zijn hiervoor slechts enkele nieuwe aselecte getallen nodig terwijl men de mogelijkheid uitsluit, dat een fraude van meer dan 1% van B in een enkele post, niet opgemerkt wordt.

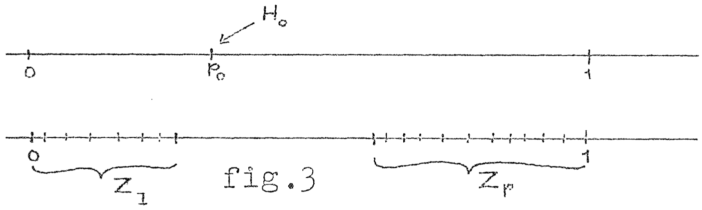
6. De in opmerking 2 van paragraaf 1 geuite suggestie eenzelfde steekproef te gebruiken voor het onderzoek naar twee verschillende kenmerken, kunnen wij toepassen door een controle tegelijkertijd te zien als fraude-controle en als accuratesse-controle. Werkt men met een steekproefomvang van n=459, dan kan men tegelijkertijd, zowel met de eerder vermelde risico's concluderen of er al of niet gefraudeerd is èn met behulp van tabel I of tabel II een bovengrens opgeven voor de fractie onjuistheden in de gehele lijst.

Appendix

Stel dat men een populatie (een partij goederen, een lijst posten, de bevolking van een stad) heeft, waarvan de elementen hetzij het kenmerk F, hetzij het kenmerk G bezitten. Er is dus een fractie p die het kenmerk F bezit en in veel situaties is het van belang na te gaan of deze fractie gelijk is aan een bepaalde waarde p_0. Wij beperken ons tot gevallen, waarin het onderzoek hiernaar met behulp van een steekproef geschiedt.

Alvorens de steekproef te nemen maakt men de onderstelling dat p de waarde p_0 bezit; men noemt dit de nulhypothese (H_0). De andere waarden die p aan kan nemen, dat zijn dus alle waarden tussen 0 en 1, ongelijk p_0, vormen samen de alternatieve hypothesen (H_1) en men zegt dat de nulhypothese H_0 getoetst wordt tegen de alternatieve hypothesen H_1.

Wanneer de fractie f van de elementen in de steekproef met het kenmerk A òf veel groter is dan p_0, òf veel kleiner dan p_0, dan zal men het steekproefresultaat niet in overeenstemming achten met de gemaakte onderstelling dat $p=p_0$ is en men zal dus de nulhypothese: $p=p_0$ verwerpen. De waarden van f die aanleiding tot verwerpen geven, vormen tezamen de kritieke zone Z, een zone die hier dus in twee gedeelten Z_l en Z_r uiteenvalt. Men spreekt van tweezijdige toetsing (vgl. fig.3).



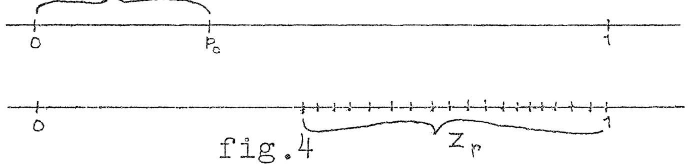
nulhypothese: $p=p_0$

steekproeffractie f

fig.3

Tweezijdige toetsing van de hypothese $p=p_0$

Behalve situaties, waarin men wil onderzoeken of p een bepaalde waarde p_0 bezit, komen er ook omstandigheden voor, waarin men na wil gaan of p een waarde $\leq p_0$ bezit. Men toetst dan de nulhypothese $p \leq p_0$ tegen de alternatieve hypothese $p > p_0$ en zal de nulhypothese verwerpen, wanneer de steekproeffractie een hoge waarde bezit. De kritieke zone bestaat dus nu uitsluitend uit hoge waarden in het interval 0-1, zodat er alleen een rechter kritiek gebied Z_r is en men spreekt dan ook van rechtseenzijdige toetsing (vgl. fig.4).



nulhypothese: $p \leq p_0$

steekproeffractie f

fig.4

Rechtseenzijdige toetsing van de hypothese $p \leq p_0$

Ook het omgekeerde geval kan zich voordoen, namelijk het toetsen van de nulhypothese $p \geqq p_0$ tegen de alternatieve hypothese $p < p_0$. Er is een <u>linker kritiek gebied</u>, bestaande uit lage waarden van f en men spreekt van <u>linkseenzijdige toetsing</u> (vgl. fig.5).

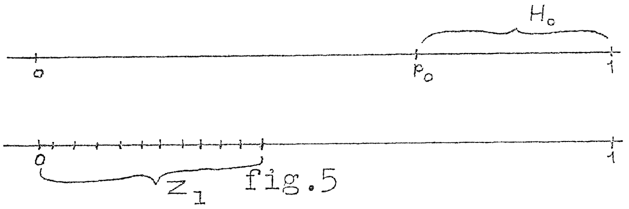


fig.5

<u>Linkseenzijdige toetsing van de hypothese $p \geqq p_0$</u>

Voor alle drie situaties geldt dat men ook als de nulhypothese juist is in het algemeen steekproefuitkomsten kan vinden in het kritieke gebied. Dit betekent dat de mogelijkheid de nulhypothese te verwerpen, wanneer ze juist is, dus ten onrechte te verwerpen, niet uitgesloten kan worden. Men noemt dit een <u>fout van de eerste soort</u>. De kans hierop wordt gewoonlijk aangegeven met α en <u>de onbetrouwbaarheid</u> van de toets genoemd, of de kans op een fout van de eerste soort.

Anderzijds is het ook niet uitgesloten dat in de steekproef een waarde buiten het kritieke gebied gevonden wordt, terwijl toch één van de alternatieve hypothesen de juiste waarde van p aangeeft en de nulhypothese dus onjuist is. De nulhypothese wordt dan ten onrechte niet verworpen, hetgeen bekend staat onder de naam <u>fout van de tweede soort</u>; de kans hierop wordt aangegeven met β . In tabel IV zijn de vier verschillende mogelijkheden die zich bij het toetsen kunnen voordoen, aangegeven.

<u>Tabel IV</u>

<u>Mogelijke combinaties bij het toetsen van een hypothese</u>

	H_0 juist	H_0 niet juist
H_0 niet verwerpen	terecht	fout 2de soort
H_0 verwerpen	fout 1ste soort	terecht

De kansen α en β hangen af van de keuze van de kritieke zone en van de omvang van de steekproef. Kiest men een kleine kritieke zone, dan is in het algemeen de kans op een fout van de eerste soort klein, doch de kans op een fout van de tweede soort groot. Bij een grote kritieke zone is in het algemeen het omgekeerde het geval.

Meestal richt men de toets nu zo in, dat de onbetrouwbaarheid α hoogstens gelijk is aan een van tevoren vastgelegde waarde α_0 (bijvoorbeeld 0,01 of 0,05). Eerst wordt dus α_0 gekozen en vervolgens bepaalt men de kritieke zone zodanig dat de kans op een fout van de eerste soort $\leqq \alpha_0$ is. Zijn er verschillende kritieke zones die aan deze eis voldoen, dan kan men van de overgebleven vrijheid nog gebruik maken door die zone te kiezen, waarvoor de kans β op een fout van de tweede soort zo klein mogelijk is. Ook kan men deze laatste kans meestal verkleinen door de steekproefomvang groter te maken.

Wij passen nu deze beschouwingen toe op het probleem van de controle op fraude door middel van steekproeven.

Gegeven is een populatie van B guldens, die twee kenmerken kunnen bezitten: gefraudeerd of niet gefraudeerd. De fractie gefraudeerde guldens geven wij aan met p. De controle of er gefraudeerd is, betekent dan statistisch gezien, dat wij willen toetsen of $p=0$ is. De nulhypothese luidt dus $p=0$ en deze wordt getoetst tegen de alternatieve hypothese $p>0$. De nulhypothese wordt verworpen als er één of meer gevallen van fraude gevonden zijn. Er is dus een rechter kritiek gebied en wij passen een rechtseenzijdige toets toe. Geven wij het aantal gefraudeerde guldens in de steekproef weer aan met k, dan bestaat het kritieke gebied uit alle waarden van k, groter of gelijk aan 1.

De kans op een fout van de eerste soort, dus de kans α op het ten onrechte verwerpen van de nulhypothese, is nul, daar de kans op één of meer fraudes in de steekproef nul bedraagt, wanneer er in het geheel niet gefraudeerd is.

De kans op het maken van een fout van de tweede soort, dus concluderen dat er niet gefraudeerd is, terwijl het in werkelijkheid wel het geval is, vinden wij door te bedenken, dat deze fout slechts gemaakt kan worden, wanneer er in de steekproef nul gevallen van fraude zijn, dus $k=0$ is. Verder betekent een fraude van een fractie p van het totaal bedrag B, dat er p B gefraudeerde guldens zijn en (1-p) B niet gefraudeerde. Wijst men nu aselect n guldens aan, dan kan men berekenen dat de kans dat geen ervan gefraudeerd is, gelijk is aan

$$\frac{\binom{(1-p)B}{n}}{\binom{B}{n}},$$

wat voor n veel kleiner dan B in voldoend goede benadering gelijk is aan $(1-p)^n$. De kans op een fout van de tweede soort

$$\beta(p) = (1-p)^n \qquad (1)$$

is dus nog een functie van p en van n. Het is deze kans die in figuur 2 als functie van p voor verschillende waarden van n werd getekend.

Wiskundig geformuleerd luidt de conclusie op blz.14 dus:

Wijst men aselect 459 guldens aan en gebruikt men als kritieke zone alle steekproefuitkomsten met één of meer fraudes, dan past men een rechtseenzijdige toets toe met onbetrouwbaarheid nul en een kans van hoogstens 0,01, dat een fraude groter dan 0,01 van het totaalbedrag niet ontdekt wordt.

De omvang van de steekproef wordt dus bepaald met behulp van formule (1). Hierbij wordt alleen gebruik gemaakt van het feit dat er n guldens aselect zijn aangewezen en dat bij controle gebleken is, dat deze geen van alle gefraudeerd waren. In werkelijkheid worden niet alleen deze n guldens gecontroleerd, doch de gehele posten, waartoe zij behoren. Wij kunnen ons nu afvragen of de kans op het niet ontdekken van gepleegde fraude niet kleiner wordt nu wij in het algemeen niet n doch veel meer dan n guldens controleren. Men kan aantonen dat dit inderdaad het geval is, wanneer er voor sommige posten te hoge bedragen werden geschreven, maar dat de kans niet kleiner wordt, wanneer alle posten òfwel correct, òfwel geheel gefraudeerd zijn. Beide beweringen illustreren wij aan de hand van het volgende voorbeeld.

Laat gegeven zijn een lijst met een totaal bedrag van B guldens, welke verdeeld zijn over $\frac{B}{10}$ posten van f 10,-. Wanneer er in totaal F guldens gefraudeerd zijn, dan bedraagt de fractie gefraudeerde guldens $p = \frac{F}{B}$. Wijzen wij nu n guldens aan en controleren wij alleen deze n guldens, dan is de kans op het niet ontdekken van fraude weer $(1-p)^n$.

Zouden wij posten aanwijzen en controleren, dan hangt de kans op het niet ontdekken van fraude af van de wijze waarop gefraudeerd is. Beschouw eerst het geval waarin een post òf niet, òf volledig is gefraudeerd. Dit betekent dat er van de $\frac{B}{10}$ posten $\frac{F}{10}$ gefraudeerd zijn, dus een fractie $\frac{F}{B} = p$. Is n ook klein t.o.v. het totale aantal posten, dan bedraagt de kans op het niet ontdekken van een fraude van een fractie p wederom $(1-p)^n$. In het geval waarin bij iedere gefraudeerde post de fraude slechts f 1,- bedraagt, zijn er van de $\frac{B}{10}$ posten in totaal F gefraudeerd. De fractie gefraudeerde posten bedraagt dan $\frac{F}{B/10} = \frac{10F}{B_n} = 10\ p$. De kans op het niet ontdekken van fraude is nu $(1-10p)^n$ en deze kans is aanzienlijk lager dan $(1-p)^n$. Een dergelijke redenering is van kracht, wanneer de fraude niet ƒ 1,- per gefraudeerde post is, doch een ander bedrag en verschillend van post tot post.

Bij de toegepaste methode wijst men guldens aan en controleert men posten. Zou een post nooit tweemaal aangewezen worden, dan zouden er n posten worden gecontroleerd en dus volgt uit bovenstaand voorbeeld, dat de kans op het niet ontdekken van fraude aanzienlijk kleiner dan $(1-p)^n$ kan zijn. Worden in ons voorbeeld niet n maar n_1 posten door de n guldens aangewezen (n_1 dus kleiner dan n), dan wordt bij een fraude van ƒ 1,- per gefraudeerde post de kans op het niet ontdekken van fraude $(1-10p)^{n_1}$ en deze kans is groter dan $(1-10p)^n$. Anderzijds blijft de oorspronkelijke redenering dat in ieder geval n losse guldens gecontroleerd zijn van kracht, zodat de kans op het niet ontdekken van fraude zeker kleiner of gelijk aan $(1-p)^n$ is.

De conclusie is dus, dat bij een controle, waarin men guldens aanwijst en posten controleert, de kans op het niet ontdekken van gepleegde fraude inderdaad kleiner kan zijn dan de gestelde grens; bij een steekproefomvang n=459 dus kleiner dan 0,01. In sommige gevallen kan deze kans zelfs aanzienlijk lager liggen.