

lib

Rianne Remmerde

Vakgroep Gezondheidsleer

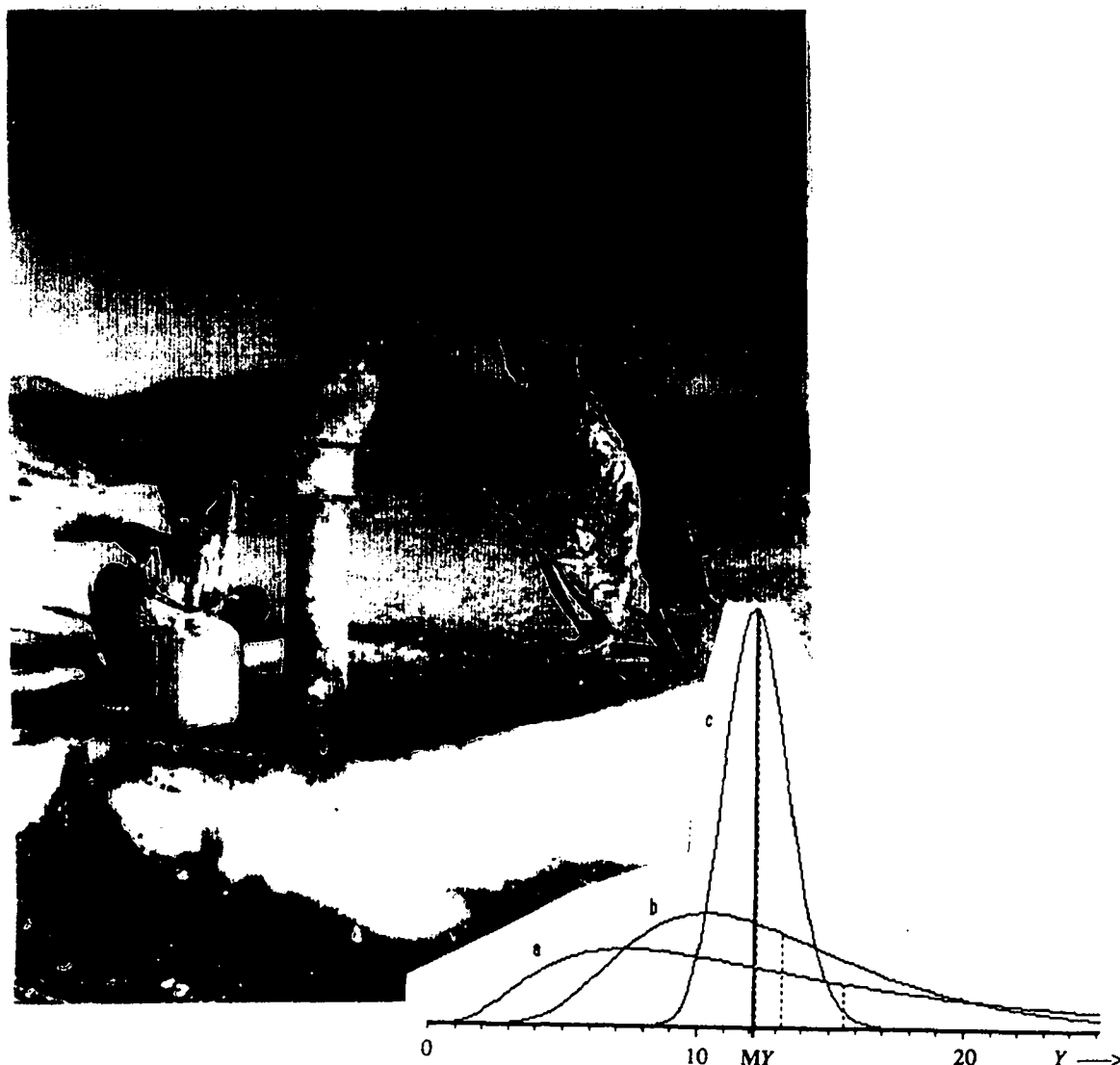
Methodologische problemen bij studies naar de effecten van drinkwatervoorzieningen en sanitaire faciliteiten op de gezondheid

203.1

90 M E

Stratificatie, als methode om versturende variabelen uit te schakelen

mei 1990



Landbouwniversiteit Wageningen

203.1-90ME-7291

**Methodologische problemen bij studies naar de effecten van
drinkwatervoorzieningen en sanitaire faciliteiten op de
gezondheid.**

**Stratificatie, als methode om versturende variabelen uit te
schakelen.**

Handwritten notes and stamps, including the number 7291 and the text 203.1.90ME.

maart 1990

Rianne Remmerde

Vakgroep Gezondheidsleer

Landbouwuniversiteit Wageningen

INHOUDSOPGAVE

voorwoord
samenvatting

1.	Algemene inleiding water/sanitatie	1
1.1	Gebrekkige voorzieningen en ziekten	1
1.2	Het water- en sanitatie-decennium	2
1.3	Gezondheids-effect-studies	2
2.	Methodologische problemen bij gedane epidemiologische onderzoeken	4
2.1	Volgens Blum en Feachem	4
2.2	Volgens Esrey en Habicht	8
2.3	Conclusies	12
2.3.1	Conclusie Blum + Feachem	12
2.3.2	Conclusie Esrey + Habicht	12
3.	Stratificatie, een methode om verstorende variabelen uit te schakelen	13
3.1	Een gestratificeerde steekproef	13
3.1.1	Algemene voordelen en aanbevelingen van de MEP	13
3.1.2	Naar welk kenmerk stratificeren	13
3.1.3	Het nemen van een steekproef	14
3.1.4	Notaties	16
3.2	Schatters voor populatiegrootheden bij een gestratificeerde aselechte steekproef	17
3.2.1	Eigenschappen van schatters	17
3.2.2	Additief versus multiplicatief model	17
3.2.3	Schatters voor het additieve model	20
3.2.4	Schatters voor het multiplicatieve model	24
3.2.5	Schatters als y_i kwalitatieve variabele is	28
3.3	Steekproefgrootte en verdeling van deze over de strata	30
3.3.1	De optimale toewijzing	30
3.3.2	De steekproefgrootte	32
4.	Conclusie	36
Annex 1	: enkele definities	37
Annex 2	: steekproef voor proporties en percentages	39
Annex 3	: enkelvoudige aselechte steekproef	41
Annex 4	: het verkrijgen van een g.a.s. uit een e.a.s.	42
	literatuur	43

Voorwoord

Dit is het verslag van een drie-maands-doctoraalvak Gezondheidsleer.

Het onderzoek had betrekking op de inventarisatie van methodologische problemen die optreden bij studies naar het gezondheidseffect van geïnstalleerde drinkwatervoorzieningen en sanitaire faciliteiten. Voorts is in het kader van deze studie aandacht geschonken aan stratificatie als methode om verstorende variabelen uit te schakelen.

De resultaten van dit onderzoek staan gepresenteerd in dit verslag.

Bij het uitvoeren van deze studie werd ik begeleid door drs S.Heisterkamp en ir B.Jansen die ik daarvoor bijzonder wil bedanken. Drs S.Heisterkamp (RIVM) voor ondersteuning van de statistische aspecten. Bij de besprekingen was de beschikbare tijd altijd de minst schaarse factor. Ir B.Jansen voor de coördinatie en algemene begeleiding. F Damen, JH Lobregt en ir J Pleyzier voor het bij nacht en ontij beschikbaar stellen van hun PC's.

Samenvatting

Het gebruik van schoon en voldoende water, goede sanitaire voorzieningen (pit-latrine en/of wasvoorziening), voorlichting over gebruik van de voorzieningen en over persoonlijke en huishoudelijke hygiëne kan in principe de opname en transmissie van ziekteverwekkers verkleinen en daardoor het optreden van diarreeziekten verminderen. Omdat, in de derde wereld, diarreeziekten tezamen met ondervoeding een grote bedreiging vormen voor het leven van kinderen onder de vijf jaar zou hierdoor de gezondheidstoestand van kleuters verbeteren. De meest gebruikte gezondheidsindicatoren bij water- en sanitatiestudies zijn diarree-morbiditeit, mortaliteit en de voedingsstatus.

De VN riep de jaren tachtig uit tot het decennium van de water- en sanitaire voorzieningen. Het doel was om in 1990 iedereen in de derde wereld voorzien te hebben van deugdelijk drinkwater en goede sanitaire faciliteiten. Het verband tussen geïnstalleerde drinkwatervoorzieningen en sanitaire faciliteiten en de gezondheid kan in het algemeen door onderzoek niet eenduidig worden aangetoond.

In 1983 publiceerden Blum en Feachem (15) een studie, waarin een groot aantal studies werden bekeken op de gebruikte onderzoeksmethoden. Zij vonden zeven veel voorkomende tekortkomingen. Water en sanitatie zijn dorp georiënteerde veranderingen en de onderzoekseenheden bestaan dan niet uit individuen of huishoudens maar uit dorpen. Verder concludeerden Blum en Feachem dat meer randomisatie noodzakelijk is om onbekende versturende variabelen uit te schakelen. Bekende versturende variabelen kan men uitschakelen door gebruik te maken van stratificatie en/of multivariantie analyse. Tevens worden 'opportunistic settings' aanbevolen.

Esrey en Habicht (14) publiceerden in 1985 een studie waarin zij een aantal tekortkomingen van gedane studies beschreven. Zij wezen er nog eens op dat het sterkste bewijs voor een hypothese wordt geleverd door experimenteel onderzoek.

Experimentele studies liefst dubbel-blind worden aanbevolen, gezien geen van de gedane studies als experimenteel kon worden aangemerkt. Tevens vonden zij dat het gezondheidseffect van sanitaire voorzieningen groter was dan het gezondheidseffect van watervoorzieningen.

Stratificatie kan verstorende variabelen uitschakelen en kan voor de hele populatie tot een grotere nauwkeurigheid leiden. Handboeken voor steekproeftheorie geven voor een gestratificeerde aselechte steekproef de schatters van de populatiekarakteristieken gebaseerd op een normale verdeling van de waarnemingen (21,22,23). Ecologische waarnemingen voldoen vaak niet aan de veronderstelling van normaliteit en homogeniteit van varianties: verdelingen zijn vaak scheef en varianties nemen vaak toe met het gemiddelde (20).

De aanname van normaliteit van de waarnemingen is gebaseerd op de meer fundamentele aanname dat onderzoeksfactoren additief werken. Als onderzoeksfactoren multiplicatief werken kan men gebruik makend van dezelfde argumenten zeggen dat de data lognormaal verdeeld zijn (20). Het additieve en het multiplicatieve model bezitten nog een kenmerk respectievelijk homogene varianties (standaarddeviaties) en homogene variatiecoëfficiënten. Dit betekent dat we uitgaande van het model kunnen voorspellen hoe de waarnemingen verdeeld zijn. Dit kan op verschillende manieren gecontroleerd worden. Een snelle en naar het lijkt effectieve methode is het uitzetten van (berekende) gemiddelden tegen de bijbehorende standaarddeviaties op dubbel-logaritmisch papier.

Gebruikelijke populatieparameters zijn centraliteit en dispersie. Om deze te schatten mbv een gestratificeerde aselechte steekproef (g.a.s.) wordt gebruik gemaakt van een enkelvoudige aselechte steekproef (e.a.s.) in ieder stratum. Bij een e.a.s. zijn de schatters voor het populatie-gemiddelde en de populatievariantie (dispersie) respectievelijk \bar{y} en s^2 . De standaardafwijking is per definitie de wortel uit de variantie. Het steekproefgemiddelde wijkt af van het

werkelijke gemiddelde; een maat voor de gemiddelde afwijking is de standaardfout: s/\sqrt{n} (of de variantie van de schatter s^2/n).

Naast puntschatters kan men ook betrouwbaarheidsintervallen schatten van de vorm $\bar{y} \pm t s/\sqrt{n}$, waarbij t de waarde is voor de Student of de normale verdeling welke correspondeert met de gekozen betrouwbaarheid.

Het stratengewicht $W_h = N_h/N$ ($h=1, \dots, L$) is in een g.a.s. een belangrijke factor. De populatiegrootte in ieder stratum moet dus bekend zijn.

In een g.a.s. zijn de puntschatters voor het populatiegemiddelde, de populatievariantie in een stratum en de standaardfout: $\Sigma_h W_h \bar{y}_h$, s_h^2 en $\sqrt{[\Sigma_h W_h^2 s_h^2/n_h]}$ (h duidt op het h -de-stratum).

Als de waarnemingen binomiaal verdeeld zijn dat wil zeggen dat de waarnemingen over twee klasse te verdelen zijn (bv ja of nee respondenten), dan is het aantal 'ja-respondenten' gelijk aan 'a' en de proportie gelijk aan a/n . De puntschatters voor de populatieproportie, voor de populatievariantie van 'a' en voor de steekproefvariantie van 'p' zijn gedefinieerd als respectievelijk $\Sigma_h W_h p_h$, s_h^2 en $\sqrt{[\Sigma_h W_h^2 s_h^2/n_h]}$.

Als de populatie niet erg groot is tov de steekproef is een correctiefactor voor de standaardfout noodzakelijk: $\sqrt{(N-n)/N}$ (voor de variantie van de schatter $(N-n)/N$). Zonder deze correctiefactor zou de standaardfout overschat worden.

Als blijkt dat de onderzoeksfactoren multiplicatief werken dus de verdeling scheef naar rechts is, wordt de mediaan gezien als de juiste centraliteitsparameter en de variatiecoëfficiënt als de juiste parameter voor dispersie (20). Een goede schatter voor de mediaan is: $\exp(\bar{z})$ (waarbij $z_i = \log y_i$ en \log : de natuurlijke logaritme). De variatiecoëfficiënt wordt het best geschat door $\sqrt{[\exp(s^2) - 1]}$ (waarbij s^2 is de variantie van $\log y_i$) (20).

De standaardfout geeft aan hoe betrouwbaar het geschatte gemiddelde is. In geval van een lognormale verdeling wordt geen gebruik gemaakt van het gemiddelde en voldoet de standaardfout dus niet. We kunnen nu een betrouwbaarheids-

interval aangeven aan weerszijde van de mediaan. Dit kan gebeuren door de betrouwbaarheidsgrenzen $g_1 = \bar{z} - t s/\sqrt{n}$ en $g_2 = \bar{z} + t s/\sqrt{n}$ uit te rekenen en terug te transformeren (s is de standaarddeviatie van $\log y_i$).

$\text{Exp}(g_1)$ en $\text{exp}(g_2)$ vormen dan de respectievelijke onder- en bovengrens voor de mediaan. Waarbij \bar{z} en s voor een g.a.s. op dezelfde wijze worden verkregen uit een e.a.s. als bij een normale verdeling.

Om zoveel mogelijk profijt te hebben van een g.a.s. is een passende steekproef in ieder stratum noodzakelijk. Echter een evenredige verdeling levert simpeler formule's en wijkt vaak niet veel af van de optimale toewijzing. Het vooraf bepalen van de steekproefgrootte is van belang om straks statistisch significante uitspraken te kunnen doen over de gevonden associaties.

1 Algemene inleiding water/sanitatie

1.1 Gebrekkige voorzieningen en ziekten

Water is een basisbehoefte voor iedereen en heeft met veel dingen in het leven te maken. Maar het aller duidelijkste verband ligt er wel tussen water en gezondheid.

Er zijn ziekten welke veroorzaakt worden door pathogenen, die kunnen voorkomen in drinkwater (1). Vele van deze ziekten worden overgebracht door de faecaal-orale route. Besmette menselijke faeces komt in het water terecht direct of indirect via de grond en besmet zo de bronnen waaruit gedronken wordt. Behalve via het drinkwater kan besmetting ook plaatsvinden via handen en voedsel (2). De meeste van de faecaal-orale ziekten manifesteren zich in het maag-darmkanaal en het bekendste syndroom is diarree (3,4). Diarree-ziekten zijn een belangrijk probleem bij jonge kinderen. Er sterven in de derde wereld zo'n 10 à 20 miljoen kinderen aan deze ziekten (1,2).

Kinderen in de derde wereld zijn kwetsbaarder door een kleine voedselreserve en een zwak immuunsysteem (5,6). De sterfte zou nog hoger zijn als borstvoeding niet door een groot deel van de bevolking geaccepteerd was (7,8). Bijvoeding is echter na 6 maanden noodzakelijk en indien dit niet met schoon water bereid wordt, ontstaat er een belangrijke besmettingsbron (9,10,11). Diarree welke op deze leeftijd ontstaat wordt "weanling diarrhoea" genoemd. Ook worminfecties kunnen plaatsvinden doordat de eitjes in het drinkwater terecht komen (1).

De aanleg van sanitaire voorzieningen en een betere waterkwaliteit zou hierin verbetering kunnen brengen.

Gebrek aan voldoende water voor persoonlijke en huishoudelijke hygiëne kan ook tot ziekten leiden. Door een slechte lichaamshygiëne kunnen huidziekten zoals schurft ontstaan (1). Ook de faecaal-orale transmissie wordt verbroken door het wassen van de handen en huishoudelijke voorwerpen (12,13). Betere beschikbaarheid van water, een hoger gebruik per capita en voorlichting over persoonlijke en huishoudelijke hygiëne kunnen hier verbetering brengen.

De zogenaamde water-contact-ziekten worden overgebracht door huidcontact met de pathogenen of door contact met geïnfecteerd water. Mensen worden geïnfecteerd door door het water te lopen of door het te drinken. Schistosomiasis is zo'n ziekte. Een essentieel deel van de levenscyclus van de pathogeen vindt plaats in de waterslak (1).

Door het vermijden van huidcontact met onveilig water en het gebruik van drinkwater kan de kans op infectie worden verminderd.

Geïnfecteerde insecten broeden in vooral stilstaand water en steken personen die er in de buurt komen. Zo worden ziekten zoals malaria overgebracht (1).

De kans op het optreden van deze ziekten kan verminderd worden door de hoeveelheid open water te verminderen.

1.2 Het water- en sanitatiedecennium

De VN heeft de jaren tachtig uitgeroepen als het decennium van water en sanitatie. Het doel is (was) om in 1990 iedereen voorzien te hebben van schoon drinkwater en goede sanitaire voorzieningen. De helft van de wereldbevolking (2 miljard mensen (1)) heeft onvoldoende en vuil drinkwater en nog groter aantal bezit geen sanitaire voorzieningen. Op het platteland zijn de voorzieningen vaak nog slechter dan in de steden. In de steden zijn de mensen beter in staat politieke druk uit te oefenen om voorzieningen af te dwingen. Van alle mensen in de ontwikkelingslanden heeft 38% schoon drinkwater. Op het platteland is de situatie het slechts, daar heeft maar 22% schoon drinkwater (1). Om in 10 jaar iedereen voorzien te hebben zal een enorme financiële en arbeidsintensieve opgave zijn. Internationale en bilaterale organisaties werken daarom samen met de ontwikkelingslanden om deze doelstelling te verwezenlijken.

Vast is komen te staan dat het bij goed drinkwater ook om geld gaat. In landen met een inkomen per hoofd van de bevolking van minder dan 220 gulden per jaar ontbreekt het 80% van de mensen aan zuiver water en eigen riolering. Voor de meeste landen met een inkomen tussen 220-1000 gulden/houfd/jaar ligt dat tussen 20-70%. De waterproblematiek staat dus niet geïsoleerd: Zij is en moet beschouwd worden als een integraal onderdeel van de totale sociaal-ekonomische ontwikkeling van een land (1,2).

Dit jaar (1990) loopt het water- en sanitatie-decennium ten einde. De doelstelling om in 1990 iedereen voorzien te hebben van schoon drinkwater en goede sanitaire voorzieningen wordt zeker niet gehaald. Wel zou men kunnen zeggen dat door het decennium de waterproblematiek extra in de aandacht heeft gestaan.

1.3 Gezondheids-effect-studies

In de derde wereld zullen vrouwen en kinderen het meest profiteren van de verbeteringen in watervoorzieningen en sanitaire faciliteiten. Vrouwen zullen profiteren omdat ze veel energie spenderen aan het verzamelen en het transporteren van water. Kinderen profiteren omdat zij het zwaarst getroffen worden door diarree-ziekten (9,10,11,12,13). Om het gezondheids-effect van een geïnstalleerde voorziening te meten worden diarree-morbiditeit, mortaliteit en de voedingsstatus vaak gebruikt als indicator met name bij kinderen onder de 5 jaar. Er zijn al heel wat studies uitgevoerd om het verwachte effect van geïnstalleerde voorzieningen aan te tonen. Met name de donoren zijn gezien de grote investeringen welke gedaan worden erg geïnteresseerd. Gedane studies rapporteerden uitéénlopende resultaten. Er wordt soms een positief, soms een negatief en ook wel geen effect gevonden (14). Een gebrekkige studieopzet wordt gezien als één van de oorzaken. Echter veel studies blijken het gebruik en het functioneren van de voorzieningen niet vooraf gecontroleerd te hebben (15). Het

de hand dat als men geen gebruik maakt van de voorzieningen men ook geen effect kan verwachten. Ook moet vooraf worden onderzocht of de geïnstalleerde voorzieningen wel naar verwachting functioneren.

Schoon water kan opnieuw vervuild worden door transport of door de wijze waarop het bewaard wordt (12). Kindervoeding wordt meestal met water bereid. Als dit voedsel bij tropische temperaturen bewaard wordt of slechts even blijft staan, vormt het een ideale groeibodem voor allerlei micro-organismen. Voorlichting en onderwijs over voedselbereiding, persoonlijke en huishoudelijke hygiëne en over wateropslag zijn dus onlosmakelijk verbonden aan de introductie van nieuwe voorzieningen (12,13).

In deze scriptie zal gekeken worden naar de methodologische problemen van gezondheids-effect-studies.

2. Methodologische problemen bij gedane epidemiologische onderzoeken

2.1 Volgens Blum en Feachem

Er zijn al heel wat pogingen gedaan om de gezondheidseffecten van drinkwatervoorzieningen en sanitaire faciliteiten aan te tonen.

In 1983 hebben Blum en Feachem een studie gepubliceerd waarin zij 44 studies hebben bekeken op de gebruikte onderzoeksmethoden (15).

De meeste studies meldden namelijk wel een verbetering van de gezondheidsindicator. Maar een meer kritische beschouwing deed toch twifelen aan de betrouwbaarheid van deze conclusies.

Blum en Feachem bekeken de studies op een groot aantal criteria. De zeven meest voorkomende problemen werden er uitgelicht.

Deze waren:

1. gebrek aan contrôle
2. 1-1 vergelijking
3. verstorende variabelen
4. gezondheidsindicator
5. leeftijdsgroep
6. registratie gebruik faciliteiten
7. seizoensinvloed

Deze problemen zullen nu één voor één nader worden verklaard.

gebrek aan contrôle

In de onderzochte studies kwamen twee soorten problemen voor, welke betrekking hadden op de contrôle.

Ten eerste was er een totale afwezigheid van een adequate contrôlesteekproef. Deze is nodig om onderscheid te kunnen maken tussen het gezondheidseffect veroorzaakt door de aangebrachte voorzieningen en het gezondheidseffect dat plaatsvond door andere factoren zoals sociale en economische veranderingen. Zonder de voorzieningen zou de gezondheidssituatie verslechterd kunnen zijn. Met voorzieningen vindt men geen effect.

Ten tweede is er het probleem van een inadequate contrôlesteekproef. De interventiegroep en de contrôlegroep zijn niet goed vergelijkbaar omdat er geen base-line data verzameld zijn of als ze wel verzameld zijn, zijn deze niet gelijk aan het begin van de interventie. Een verandering in gezondheid kan dan niet toegeschreven worden aan de aangebrachte voorziening.

1-1 vergelijking

Als een contrôlegroep wordt geselecteerd is het algemeen gebruik deze groep met een interventiegroep te vergelijken. En meestal wordt voor iedere groep een dorp gekozen en hieruit wordt dan een steekproef genomen.

Watervoorzieningen en sanitaire faciliteiten zijn typisch

voorzieningen welke de hele gemeenschap aangaan. De hele gemeenschap is onderwerp van studie en niet het individu of het huishouden. Een vergelijking van twee dorpen (interventie + contrôle) is gelijk aan het vergelijken van 2 individuen. De steekproefgrootte is dan gelijk aan één en hieruit kan dus geen deugdelijke statistische conclusie worden getrokken. De cross-sectionele studies maken vaak wel gebruik van verschillende dorpen in beide groepen. Het individu of een huishouden wordt vaak genomen als onderzoekseenheid en er wordt aangenomen dat deze onafhankelijk van elkaar reageren op de voorzieningen. De statistische analyses gebruiken de inter-individuele (of intra-dorp) variatie om gemiddelden van dorpen te vergelijken. Er wordt geen rekening gehouden met de inter-dorp variatie. Deze benadering is alleen geldig als de inter-dorp variatie nul is. En dit gebeurt alleen als men verschillende dorpen aselekt selecteerd.

Een andere oplossing is om een dorp te gebruiken waaruit een interventiegroep en verscheidene contrôlegroepen geselecteerd worden. De contrôlegroepen worden geselecteerd op het geen gebruik maken van de voorzieningen. De inter-dorp variatie is dan niet relevant en de variatie tussen individuen of huishoudens in het dorp is dan de belangrijkste factor.

Met deze methode kunnen er andere problemen optreden:

1. Er kunnen verstorende verschillen zijn tussen de intra-dorp groepen in het vóórkomen van ziekten welke niet gerelateerd zijn aan de interventie. Base-line data voorafgaande aan de interventie zijn nodig om deze verschillen vast te leggen.
2. Een conclusie welke men getrokken heeft op grond van één dorp hoeft niet voor andere dorpen te gelden.

Sommige studies hebben één interventiedorp vergeleken met verschillende contrôledorpen. Deze vergelijking is toegestaan als men de variatie tussen de contrôledorpen berekend en deze gemiddelden van de contrôledorpen vergelijkt met het gemiddelde van het interventiedorp en kijkt of de variantie buiten de range van de verwachte variantie licht. Echter is het belangrijk de criteria te kennen waarbij het interventiedorp gekozen is.

Om effectief te kunnen omgaan met het 1-1 vergelijkingsprobleem moet de steekproefgrootte vergroot worden, als de steekproefeenheid een dorp is, omdat men een dorp gerichte interventie onderzoekt.

verstorende variabelen

Corrigeren voor de vele verstorende invloeden is een onmogelijke taak, behalve als men voldoende dorpen aselekt selecteerd. Onvermijdelijk is het echter nodig de variabelen welke de gezondheidsindicator het meest beïnvloeden te selecteren. Welke verstoorders men selecteert, hangt af van de gezondheidsindicator welke men gebruikt. Als men het effect van sanitaire faciliteiten op diarree bestudeert

zijn de watervoorziening en de sociaal-ekonomische status belangrijke verstoorders. Men kan verstoorders op verschillende wijzen uitschakelen. Men kan de interventiegroep en de contrôlegroep zodanig selecteren dat de versturende variabelen in beide groepen gelijk zijn. Interventiegroep en contrôlegroep kunnen worden geselecteerd op andere gronden en men kan corrigeren voor de verstoorders bij de data-analyse door gebruik te maken van "matched sub-samples" Deze twee methoden kunnen ook samen gebruikt worden. Het belangrijkste is de verstoorders te ontdekken en te meten bij aanvang van de studie en deze systematisch te meten tijdens het verloop van de studie. Het meten van verstoorders gedurende het onderzoek is belangrijk omdat niet te verwachten niet-omgevings-interventies kunnen plaatsvinden, welke de gekozen gezondheidsindicator kunnen beïnvloeden. Alle door Blum en Feachem bekeken studies faalde in het corrigeren voor de mogelijke versturende variabelen.

gezondheidsindicator

Een van de meest gebruikte indicatoren, vooral bij jonge kinderen, is diarree-morbiditeit. De informatie omtrent diarree-morbiditeit wordt meestal, verkregen door de onderzoekshuishoudens regelmatig te bezoeken. Men vraagt hoe vaak de leden diarree hebben gehad in de afgelopen periode. Men kan vragen naar het aan- of afwezig zijn van diarree of naar het aantal diarree-episodes. Er zijn veel redenen te noemen waarom deze informatie niet compleet is.

Ten eerste kan degene die ondervraagd wordt niet bekend zijn met de diarree-geschiedenis van de familieleden. Voor erg jonge kinderen zal dit niet gelden, gezien de moeder of verzorgster hier meestal wel mee bekend is.

Ten tweede kan er een schaamte zijn om dit bekend te maken ook al weet men dit wel. Door te beloven dat de diarree behandeld zal worden of/en door voorlichting omtrent diarree-preventie te beloven kan men de angst wegnemen. De motivatie waarom diarree ontkend kan worden, is dat deze in relatie gebracht wordt met een ziekte of angst dat een positief antwoord leidt tot het nemen van bloed- en/of faecesmonsters.

Gezondheidsindicatoren moeten exact gedefinieerd worden. Vele studies definiëren diarree onvoldoende. Diarree in het bijzonder leidt tot onjuiste interpretaties, wat varieert van respondent tot respondent als in een onderzoek niet een standaardisatie van de definitie voor dat speciale geval heeft plaatsgevonden. Bij infecties waarbij zowel blootstelling als immuniteit bijdragen tot de klinische verschijnselen, is de infectie een gevoeliger maat voor het resultaat van een interventie dan de klinische verschijnselen. Echter zelfs als indicatoren van infecties worden gebruikt, moeten zij eerst gedefinieerd worden.

Bv *Entamoeba histolytica* cysts in faeces wordt vaak niet gedefinieerd met respect voor het verschil van *Entamoeba hartmanni* cysts. Coliform tellingen worden gerapporteerd bij watervoorzieningen zonder dat gedefinieerd wordt wat wordt bedoeld met een coliform.

leeftijdsgroepen

De ziekten en infecties welke worden bekeken bij gezondheids-effect-studies zijn niet gelijk verdeeld over de leeftijden. De typen diarree welke worden bekeken, hebben de hoogste incidentie bij jonge kinderen. Het effect van een interventie hangt af van het gedrag en het gebruik van de nieuwe voorzieningen. Deze zijn beide afhankelijk van de leeftijd. Daarom is het noodzakelijk om een leeftijdsspecifieke benadering te hanteren. In sommige situaties kan men het hele onderzoek beter op jonge kinderen toepassen.

registratie van gebruikte voorzieningen

Watervoorzieningen of latrines hebben zelf geen effect op de gezondheid. Alle verbeteringen in de gezondheidstoestand hangen af van het gebruik van de voorzieningen. Terwijl deze bewering vanzelfsprekend lijkt, falen vele studies te vermelden of het gebruik gecontroleerd is. Men denkt dat aanwezigheid gelijk is aan gebruik.

Het gebruik van de voorzieningen kan worden bepaald door navraag en/of door observaties. Observationeel onderzoek naar gebruik is betrouwbaarder speciaal als het gaat om de gebruikte hoeveelheid water en het gebruik van sanitaire faciliteiten. Vragen omtrent welke bron wordt gebruikt, kunnen geverifieerd worden door observaties.

Weinig studies gebruiken observationele gegevens omtrent gebruik, vooral systematisch verkregen observationele gegevens worden weinig gebruikt. Als een gezondheidseffect wordt geconstateerd en men heeft geen juiste registratie van het gebruik van de vernieuwingen kan men niet zeggen waarom of waardoor het gezondheidseffect wordt veroorzaakt. Als geen gezondheidseffect werd waargenomen, kon dat zijn omdat de faciliteiten niet werden gebruikt. Observationele gegevens zijn vooral belangrijk bij gebruik van de latrines, daar zijn vragen beslist onbetrouwbaar. Het aanleggen van nieuwe latrines leidt niet altijd tot gebruik. Er zijn veel voorbeelden bekend, waar men totaal geen gebruik maakte van de nieuwe toiletten. Het is belangrijk dat defaecatie-gewoonten worden vastgelegd, vooral van die groep welke het meest lijdt aan de ziekte of infectie.

In geval van onderzoek naar diarree is het noodzakelijk het defaecatie-gedrag van jonge kinderen te bekijken, gezien deze het minst gebruik zullen maken van de faciliteiten.

seizoensinvloed

Diarree-ziekten en de infecties welke er aan gerelateerd zijn, zijn seizoengebonden in de meeste delen van de wereld. Dit geldt vaak ook voor infecties veroorzaakt door parasieten bv ascaris. De gezondheidsindicator wordt wel tegelijk gemeten in de interventie- en de contrôlegroep maar er wordt niet bij vermeld, waarom een bepaald seizoen is gekozen. Zelfs de interventie-variabelen zoals watergebruik, waterbron en

latrinegebruik zijn seizoengebonden. De meest betrouwbare informatie wordt verkregen door het hele jaar het gebruik te observeren. Dit is echter erg arbeidsintensief. Beter kan men de pieken in een jaar vaststellen en tijdens deze periode de observaties uitvoeren. De natte en droge periode(n) in een jaar zijn vaak zulke pieken. Hoe de gegevens ook verzameld zijn, ze moeten geanalyseerd worden per seizoen. Als slechts eenmaal per jaar observaties mogelijk zijn dan moeten deze tenminste ieder jaar in dezelfde periode worden verzameld. Bij voorkeur tijdens de periode waarin de gezondheidsindicator zijn piek bereikt. De gegevens moeten voor interventie- en contrôlegroep op hetzelfde tijdstip verzameld worden.

2.2 Volgens Esrey en Habicht

In 1985 publiceerden Esrey en Habicht een studie waarin zij alle na 1950 gepubliceerde studies over de relatie tussen watervoorzieningen/sanitaire faciliteiten en de gezondheid (mortaliteit, morbiditeit en voedingsstatus) bekeken (14).

Er werd onderscheid gemaakt tussen experimentele studies en observationele studies.

experimentele studies

Het belangrijkste criterium voor een experimentele studie is een aselechte toewijzing van de voorzieningen.

Dit vermindert de kans dat de gevonden associaties het gevolg zijn van versturende variabelen welke eigen zijn aan de interventie- of de contrôlegroep. Verder kan men zeggen dat de gevonden associatie niet het gevolg is van versturende variabelen, zelfs als men de verstoorders niet heeft kunnen vastleggen en meten.

Randomisatie alleen kan niet aangeven welke component verantwoordelijk is voor het gevonden effect.

Er is een studieontwerp nodig, dat voldoet aan de volgende punten:

1. Een steekproef moet groot genoeg zijn om statistische significantie aan te tonen.
2. Er moet een juiste onderzoekseenheid gekozen worden. Hier het dorp. De inter-kind variatie is kleiner dan de inter-dorp variatie. Als men dan de inter-kind variatie gebruikt in de analyse wordt het effect onderschat.
3. Als er ook voorlichting in het interventiedorp gegeven is, kan het effect niet alleen worden toegeschreven aan bv de watervoorziening.
4. Het analyseren van gegevens over diarree-morbiditeit moet zonder "bias" gebeuren. Deze gegevens moeten dus blind geanalyseerd worden.
5. Doordat men weet welk dorp de interventie heeft gekregen, kan er ook vertekening optreden. Men kan dit voorkomen door een groep een placebo-interventie toe te wijzen.

Studies waarin de interventie aselekt wordt toegewezen, de steekproef voldoende groot is, de schatting van de gezondheidsindicator blind is uitgevoerd en de toewijzing van de interventie blind is, worden dubbel-blind gerandomiseerde studies genoemd.

Experimentele studies moeten de bekende en de potentiële verstoorders meten, om te verifiëren of randomisatie succesvol was in het verdelen van de verstoorders. De waarschijnlijkheid dat er een causaal verband bestaat kan toenemen als verstoorders worden geïdentificeerd en in rekening gebracht in het studieontwerp of in de analyse. Zelfs als sommige verstoorders worden gemist is de kans groter dat de gevonden associatie niet uit verstoorders voortkomt.

Verstorende variabelen welke waarschijnlijk studies, naar het effect van water en sanitatie op de gezondheid, beïnvloeden, kunnen worden geïdentificeerd aan het begin van de biologische weg die leidt van de interventie naar de uitkomst. Als alle stappen veranderen als was voorspeld door de biologische theorie, dan zal de aannemelijkheid dat de veranderingen zijn toe te schrijven aan de interventie toenemen.

De bovenstaande criteria werden toegepast op water- en sanitaire interventies welke een statistisch significante associatie aangaven. Geen van de studies was gerandomiseerd. En omdat deze studies geen aselekte toewijzing aangaven, konden ze niet geëvalueerd worden als een experimentele studie.

Zij werden geëvalueerd volgens de criteria van een observationele studie.

observationele studies

Observationele studies kunnen alleen aangeven dat er een relatie is.

Cross-sectionele studies gebruiken de gegevens om de relatie tussen water en/of sanitatie en het effect te onderzoeken.

Voor en na surveys kijken naar de reductie in diarree welke door de interventie wordt veroorzaakt. Beide typen studies kunnen associaties meenemen welke niet direct veroorzaakt zijn door de veranderingen.

Bv De voor-na studie kan een vermindering van diarree constateren maar deze kan het gevolg zijn van de verbeterde levensstandaard. Deze kan een verbeterde gezondheid met zich meebrengen door een verbeterde hygiëne. Om een dergelijke verstoorder uit te sluiten, kan men er dorpen bijnemen welke geen verandering van levensstandaard hebben ondergaan.

Er kunnen twee technieken aangewend worden om voor verstorende variabelen te corrigeren. Vergelijkbare groepen kunnen "gematched" worden. Men kan ook gebruik maken van multivariantie analyse.

Evenals in experimentele studies moeten fouten van de metingen geminimaliseerd worden. Als blind toewijzen niet mogelijk is, blijft blind schatten over en is dit de enige manier om eventuele "bias" uit te sluiten. Metingen kunnen gestandaardiseerd worden door de verstoorders te stratificeren of te randomiseren over de vergelijkbare groepen.

Observationele studies moeten er van overtuigd zijn dat de vernieuwingen niet altijd gebruikt worden. Daarvoor zijn vele redenen te noemen: niet functioneren, niet weten hoe te gebruiken, andere bron is makkelijker bereikbaar en sanitatie problemen.

De aannemelijkheid van een gevonden associatie vermindert als er geen verslag wordt gedaan van het wel of niet gebruiken van de voorzieningen. Aan de andere kant neemt de geloofwaardigheid van een associatie toe als gezondheidseffecten worden gemeten en veranderen volgens de voorspelde theorie.

Een paar studies vertoonde een verandering in gezondheid maar berekende geen statistische associatie. Sommige onderzoekers rapporteerden zowel negatieve als wel positieve effecten voor verschillende gezondheidsindicatoren of voor verschillende leeftijdsgroepen.

Studies welke geen associatie aangaven, werden toch geëvalueerd in het licht van de volgende drie criteria.

Ten eerste was het te verwachten dat het resultaat werd veroorzaakt door de verbeteringen.

Bv In een gebied waar de meeste kinderen borstvoeding krijgen, zal het niet waarschijnlijk zijn dat de toevoeging van schoon water de kindersterfte doet afnemen. Gezien borstvoeding de blootstelling aan pathogenen vermindert en de immuniteit tegen binnen gekregen pathogenen verhoogt.

Ten tweede was de steekproefgrootte voldoende om statistisch significante uitspraken te kunnen doen.

Ten derde was er contrôle voor negatieve verstoorders. Een negatieve verstoorder werkt verstrend op beide uitkomsten en kan dus ieder effect van een interventie uitvegen.

Bv Een verbeterde watervoorziening brengt mensen samen in een bepaald gebied en veroorzaakt een overbevolkingsprobleem. Overbevolking vergroot de kans op verspreiding van pathogenen direct van persoon tot persoon en zal het effect van de verbeteringen maskeren.

Negatieven verstoorders kunnen resulteren in geen effect, terwijl positieve verstoorders een effect onafhankelijk van de interventie zal toestaan.

De meeste studies welke een negatieve associatie rapporteerden gaven niet de leeftijdsgroep aan. De overige studies corrigeerden niet voor verstorende variabelen of de interventies in de te vergelijken dorpen verschilden. Er waren geen studies welke een negatieve associatie rapporteerden en aannamen dat water- en/of sanitaire voorzieningen de gezondheid niet verbeterden.

2.3 Conclusies

2.3.1 Conclusie Blum + Feachem

Het is bij water- en sanitatieonderzoek van belang om te concentreren op "opportunistic settings". Dit zijn studies welke experimenteel van opzet zijn en aangrijpen bij bestaande of geplande watervoorzieningen of sanitatie programma's. Concentratie op "opportunistic settings" is niet inconsistent met het concept van random toewijzing, welke men altijd moet trachten toe te passen. Als effectevaluatie op het programma staat en aan het begin van het interventieprogramma in beschouwing wordt genomen dan is het mogelijk om gebieden te kiezen op random basis.

Tevens kan er bij "opportunistic settings" sprake zijn van een duidelijke oorzaak-effect relatie.

Verder concludeerden Blum en Feachem dat meer randomisatie noodzakelijk is om onbekende verstorende variabelen uit te schakelen. Bekende verstorende variabelen kan men uitschakelen door gebruik te maken van stratificatie en/of multivariantie analyse (16).

2.3.2 Conclusie Esrey + Habicht

Er wordt gesteld dat vele studies rapporteren dat er een experimenteel onderzoek heeft plaatsgevonden terwijl volgens de criteria geen van de onderzoeken als experimenteel kan worden aangemerkt.

Observationele onderzoeken kunnen wijzen op een verband tussen de aanwezigheid van een bepaalde factor en het vóórkomen van ziekten. Als het verband sterk en consistent is mag een bepaald verband vermoed worden. Dit kan echter door observationeel onderzoek niet bewezen worden, gezien de altijd aanwezige mogelijkheid van verstorende variabelen. Het sterkste bewijs voor de juistheid van een hypothese wordt geleverd door experimenteel onderzoek (16,17).

Bij experimenteel onderzoek heeft men volledige contrôle, men kan een interventie toewijzen en men kan deze weglaten. Bij een observationeel onderzoek bekijkt men iets zonder actief te sturen.

3. Stratificatie, een methode om verstorende variabelen uit te schakelen

3.1 Een gestratificeerde steekproef

3.1.1 Algemene voordelen en aanbevelingen van MEP

Het stratificeren van een onderzoekspopulatie is een bekende techniek. Er zijn een aantal redenen waarom men deze techniek kan toepassen.

Men wil bijvoorbeeld gegevens over een subpopulatie of als het onderzoeksgebied erg groot is kunnen subpopulaties organisatorische en/of administratieve voordelen opleveren.

In het algemeen dient het twee doelen:

1. Het bereiken van een grotere nauwkeurigheid met de dezelfde middelen.
2. Het uitschakelen van de invloed van verstorende variabelen.

De WHO heeft in 1982 een rapport uitgebracht (18) waarin een relatief goedkope en simpele methode wordt beschreven voor evaluatie van watervoorzieningen en/of sanitaire faciliteiten. Deze methode wordt de Minimum Evaluation Procedure genoemd, kort MEP. Een evaluatie naar het gebruik en het functioneren van de voorzieningen zou deel uit moeten maken van een gezondheids-effect-studie.

Bij zo'n evaluatie naar het gebruik en het functioneren van de voorzieningen wordt stratificatie vaak aanbevolen.

Ook zijn een groot aantal gezondheids-effect-studies geëvalueerd en één van de veel voorkomende tekortkomingen was dat men onvoldoende rekening hield met verstorende variabelen (zie 2.3.1).

Stratificatie zou hier van pas kunnen komen, gezien deze methode aangewend kan worden om de invloed van bepaalde populatie-karakteristieken uit te schakelen (19).

3.1.2 Naar welk kenmerk stratificeren

Het grootste voordeel qua nauwkeurigheid verkrijgt men, als de onderzoekspopulatie naar het onderzoekskenmerk wordt stratificeerd.

Men wil bijvoorbeeld weten hoeveel water de huishoudens in de onderzoekspopulatie per dag gebruiken. Uit een kleine steekproef vooraf blijkt, dat er grote verschillen bestaan in het waterverbruik per huishouden per dag.

Men zou dan een gestratificeerde steekproef kunnen doen, waarbij strata worden gevormd naar het waterverbruik.

Meestal blijkt zo'n stratificatie niet mogelijk en zal men

gebruik moeten maken van een toepassing welke deze situatie benadert. De populatie wordt dan gestratificeerd naar een kenmerk dat ongelijk is aan het onderzoekskenmerk maar er wel aan gerelateerd is.

Vooraf van belang is hierbij dat het kenmerk waarnaar men stratificeert, gerelateerd is aan het onderzoekskenmerk. Tevens moet iedere eenheid eenduidig in één van de strata zijn in te delen, bovendien moet dit eenvoudig zijn uit te voeren.

We gebruiken bijvoorbeeld de survey naar het watergebruik van het voorgaande voorbeeld. We hebben nu de beschikking over een eerder gedaan onderzoek in dit zelfde gebied. Hieruit blijkt dat er grote sociaal-ekonomische verschillen voorkomen.

Uit de literatuur weten we dat sociaal-ekonomische verschillen gerelateerd zijn aan het watergebruik van de huishoudens.

Het onderzoeksgebied kan nu gestratificeerd worden naar sociaal-ekonomische status. Maar voordat men dit uit gaat voeren, moeten de sociaal-ekonomische strata goed gedefinieerd worden. De eenheden kunnen dan eenduidig worden ingedeeld.

Nadat het stratificatie-kenmerk gekozen is, kan de populatie (N) worden ingedeeld in subpopulaties of strata zodat

$N = N_1 + N_2 + \dots + N_L$. De populatie-grootte van elk van de strata moet dus bekend zijn.

De populatie wordt altijd zo in strata verdeeld, dat binnen de strata het te onderzoeken kenmerk weinig varieert, terwijl de strata onderling veel variëren.

3.1.3 Het nemen van een steekproef

Nadat de populatie (N) is ingedeeld in subpopulaties of strata zodat $N = N_1 + N_2 + \dots + N_L$ kan in elk stratum een aselecte steekproef worden genomen.

De steekproefgrootte in een stratum wordt aangeduid met n_h en de steekproefgrootte (n) wordt gevormd door

$$n = n_1 + n_2 + \dots + n_L.$$

Voordat men in ieder stratum een steekproef gaat nemen, zal er eerst een beslissing genomen moeten over de grootte van de steekproef in ieder stratum (zie 3.3.1).

Echter een evenredige steekproef ligt het meest voor de hand.

Stel bijvoorbeeld een evenredige steekproef van 10%. Dan moet uit elk stratum een steekproef van 10% genomen worden. Grote strata zullen dan absoluut ook een groter aandeel leveren dan kleine strata; de verhoudingen tussen de steekproefaantallen zijn gelijk aan die tussen de populatieaantallen.

Om statistisch significante uitspraken te kunnen doen over de gevonden associaties is een steekproef van de juiste grootte noodzakelijk (zie 3.3.2).

3.1.4 Notaties

De letters symboliseren een bepaald begrip in de wiskundige statistiek en de hier gebruikte symbolen worden algemeen gebruikt.

Het achtervoegsel h duidt op een stratum en i op een eenheid in een stratum.

De hier onderstaande symbolen refereren allen aan het h-de stratum.

N_h		totaal aantal eenheden
n_h		aantal eenheden in de steekproef
Y_{hi}		waarde voor de i-de eenheid
$W_h = N_h/N$		stratumgewicht
$f_h = n_h/N_h$		steekproeffractie
$\bar{Y}_h = \Sigma_i Y_{hi}/N_h$		populatie-gemiddelde
$\bar{y}_h = \Sigma_i Y_{hi}/n_h$		steekproefgemiddelde
$S_h^2 = \Sigma_i (Y_{hi} - \bar{Y}_h)^2/(N_h - 1)$		werkelijke variantie
$s_h^2 = \Sigma_i (Y_{hi} - \bar{y}_h)^2/(n_h - 1)$		steekproefvariantie
$p_h = a_h/n_h$		proportie
$s_h^2 = p_h q_h * n_h/(n_h - 1)$		schatting voor de variantie als y_i kwalitatieve variabele is met slechts twee waarden (ja/nee etc)

3.2 Schatters voor populatiegrootheden bij een gestratificeerde aselechte steekproef

3.2.1 Eigenschappen van schatters

Bij het nemen van een steekproef is het niet alleen van belang dat men de juiste steekproefmethode gebruikt. Van zeker zo groot belang is dat de schatters van de populatieparameters correct worden berekend.

Twee belangrijke eigenschappen van schatters zijn:

a. Unbiasedness

Een schattingsmethode is zuiver (unbiased) als de gemiddelde waarde van de schatting genomen over alle mogelijke steekproeven bij een gegeven 'n' exact gelijk is aan de populatiewaarde.

b. Consistentie

De schatter is dan consistent als de kans, dat een fout groter is dan de toegestane fout, naar nul gaat als de steekproef groter wordt.

3.2.2 Additief versus multiplicatief model

Bij het schatten van de populatieparameters in een gestratificeerde aselechte steekproef gaat men uit van de aanname dat de waarnemingen normaal verdeeld zijn.

Ook bij analyses met verschillende variabelen zoals regressie analyse wordt uitgegaan van de aanname dat de waarnemingen normaal verdeeld zijn en de varianties homogeen zijn.

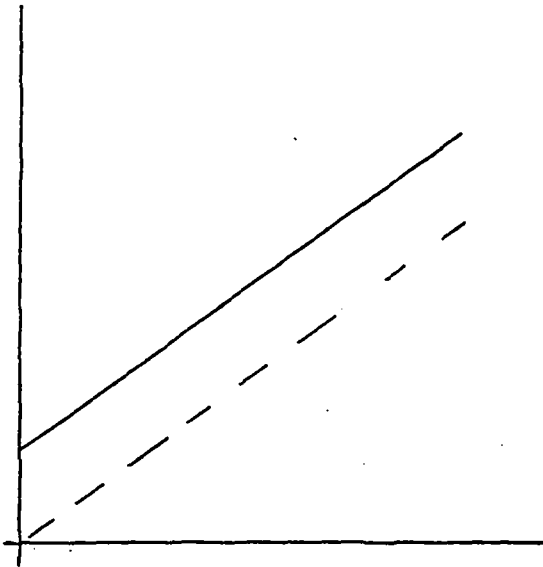
De aanname van normaliteit van de variabelen is gebaseerd op de meer fundamentele aanname dat variatiebronnen additief werken, op de gemeten variabelen.

Als variatiebronnen multiplicatief werken kan men gebruik makend van dezelfde argumenten zeggen dat de data lognormaal verdeeld zijn (20).

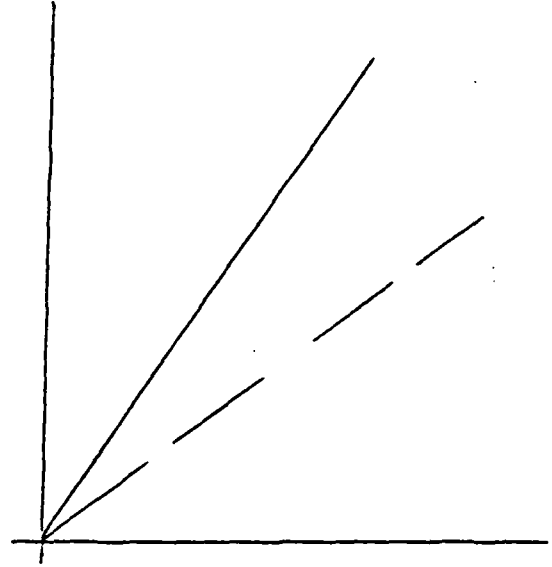
Dit betekent dat we uitgaande van het model kunnen voorspellen hoe de waarnemingen verdeeld zijn. Er kan nog een eigenschap van de waarnemingen voorspeld worden. In het additieve model zijn de varianties (of standaarddeviaties) homogeen. In het multiplicatieve model nemen de standaarddeviaties rechtevenredig toe met het gemiddelde. Omdat de variatiecoëfficiënt gedefinieerd is als de standaarddeviatie gedeeld door het gemiddelde, zijn in het laatste geval de variatiecoëfficiënten homogeen (20).

Of men met het additieve dan wel het multiplicatieve model te maken heeft, heeft gevolgen voor het analyseren van de waarnemingen.

Onder variatiebronnen verstaan we factoren welke van invloed zijn. Onbekende variatiebronnen zijn variatiebronnen welke niet zijn opgenomen in het experiment.



additief model



multiplicatief model

eigenschappen:

- normale verdeling
- variantie homogeen

- lognormale verdeling
- variatiecoëfficiënten homogeen

bv 1. additief model; $g = b + e$

2. multiplicatief model; $wg = b * a * p$.

1. De gezondheidstoestand van kinderen wordt bepaald mbv een gezondheidsindicator (g). Deze zou opgebouwd kunnen zijn uit de som van de base-line indicator (b) en een effectindicator (e).

2. Het watergebruik (wg) wordt bepaald door het gebruik van een gemiddeld huishouden: het basisgebruik (b), maal een factor welke de afstand tot de bron weergeeft (a) en een factor welke het aantal extra personen per huishouden weergeeft (p).

Hoe kunnen we nu bepalen met welk model we te maken hebben.

Ten eerste moeten waarnemingen welke voor het multiplicatieve model in aanmerking willen komen een natuurlijk nulpunt hebben. Immers als de variatiebronnen multiplicatief werken, kan men zeggen dat de data lognormaal verdeeld zijn. En de logaritme is alleen gedefinieerd voor positieve waarden.

Afmetingsmaten zoals lengten, gewichten en afstanden blijken vaak lognormaal verdeeld. Hoewel de lengte heel lang werd gezien als het schoolvoorbeeld voor een normale verdeling. Later vonden (Camp 1938 en Wachholder 1952, geciteerd (20)) dat de lengte van de mens bij grote aantallen data lognormaal verdeeld was.

Ten tweede zou men kunnen nagaan (mbv van literatuur of gedane onderzoeken) welke bijdragen te verwachten valt. Of de onderzoeksfactoren en andere variatiebronnen zijn opgebouwd uit een product ofwel uit een som.

Ten derde zou men de gevonden gemiddelden uit kunnen zetten tegen de varianties op dubbel-logaritmisch papier. Door de gevonden punten kan men dan een lijn trekken. Als de coëfficiënt van de lijn ongeveer gelijk is aan één is er sprake van homogene variatiecoëfficiënten en zijn de data lognormaal verdeeld. Als de coëfficiënt gelijk is aan nul zijn de data normaal verdeeld (20).

Als men gekozen heeft voor het additieve dan wel het multiplicatieve model kan men de populatieparameters schatten zoals in de volgende hoofdstukken wordt beschreven.

3.2.3 Schatters voor het additieve model

Een gestratificeerde random steekproef als de waarnemingen normaal zijn verdeeld.

Populatiekarakteristieken waarin men, bij het uitvoeren van een steekproef, meestal is geïnteresseerd.

PARAMETER	SCHATTER
centraliteit populatie-gemiddelde \bar{Y}	$\hat{Y} = \bar{Y}_{st}$
dispersie variantie van y_i	s_h^2
variantie van de schatter	$s(\bar{Y}_{st})^2$

Centraliteitsparameter

De schatter voor het populatie-gemiddelde (\hat{Y}) in een gestratificeerde aselechte steekproef (g.a.s.) is gelijk aan:

$$\bar{Y}_{st} = \sum_h N_h/N * \bar{Y}_h = \sum_h W_h \bar{Y}_h$$

y_{st} wordt gezien als een zuivere schatter voor het populatie-gemiddelde als het stratumgemiddelde y_h in ieder stratum zuiver is vastgesteld. Een stratumgemiddelde is per definitie zuiver bepaald als het verkregen is mbv een enkelvoudige aselechte steekproef (e.a.s.).

Bovenstaande schatter is ongelijk aan het steekproef-gemiddelde.

Voor het steekproefgemiddelde geldt:

$$\bar{y} = \sum_h n_h/n * \bar{y}_h = \sum_h w_h \bar{y}_h$$

Het verschil zit hem in het feit, dat bij de schatter voor het populatie-gemiddelde ieder stratum zijn correct gewicht krijgt. Het gevolg hiervan is dat ieder stratum een evenredige bijdrage levert aan het gemiddelde.

Bijvoorbeeld uit een populatie van 100 eenheden wordt een steekproef genomen van 30 eenheden.

$N = 100, n = 30.$

stratum	N_h	n_h	W_h	w_h	\bar{Y}_h
$h = 1$	50	10	0.5	0.33	8.3
2	30	10	0.3	0.33	7.9
3	20	10	0.2	0.33	15.8

$$\bar{Y}_{st} = 0.5 * 8.3 + 0.3 * 7.9 + 0.2 * 15.8 = 9,7$$

$$\bar{y} = 0.33 (8.3 + 7.9 + 15.8) = 0.33 * 32.0 = 10,6$$

Het derde stratum heeft een vrij hoog gemiddelde uit een relatief kleine steekproef. Als het steekproefgemiddelde de schatter zou zijn voor het populatie-gemiddelde zou aan het derde stratum teveel gewicht worden toegekend.

Een voor de hand liggende steekproef uit ieder stratum is dus een steekproef welke evenredig is aan de grootte van het stratum.

Men spreekt dan over een evenredige verdeling van de steekproef over de strata.

In formule: $n_h = N_h/N * n$

Uit $n_h = N_h/N * n$ volgt dat $n_h/n = N_h/N$.

Uit het bovenstaande blijkt ook dat bij een evenredige verdeling van de steekproef over de strata de steekproeffractie (f_h) in ieder stratum gelijk is aan de steekproeffractie (f).

Conclusie: In alle gevallen zal de schatter voor het populatie-gemiddelde welke bepaald is mbv het correcte stratumgewicht voldoen. Het steekproefgemiddelde voldoet alleen bij een evenredige toewijzing.

Dispersieparameter

De gebruikelijke parameter voor dispersie is de variantie. Als schatter voor de variantie van y_i in een stratum wordt een afleiding van de schatter voor de variantie in een e.a.s. gebruikt (Annex 3).

$$s_h^2 = \Sigma_i (Y_{hi} - \bar{Y}_h)^2 / (n_h - 1)$$

De variantie geeft de afwijking van het gemiddelde, hier dus

de afwijking van het stratumgemiddelde.

De standaardafwijking wordt in publicaties vaak gebruikt om aan te geven hoe betrouwbaar het gemiddelde is als centraliteitsparameter.

De standaardafwijking is per definitie de wortel uit de variantie van y_i .

Variantie van de schatter

We weten nu hoe we de schatter voor het populatie-gemiddelde kunnen bepalen. Maar we zijn ook geïnteresseerd in welke mate de schatter afwijkt van het werkelijke populatie-gemiddelde. Met andere woorden we willen weten hoe betrouwbaar de schatter wel is.

We kunnen deze vraag beantwoord krijgen door de variantie van de schatter te bepalen.

Als de steekproeven in de verschillende strata onafhankelijk zijn genomen dan geldt voor de variantie van de schatter:

$$V(\bar{y}_{st}) = \sum_h W_h^2 V(\bar{y}_h)$$

Dus als \bar{y}_h in ieder stratum een zuivere schatter is van \bar{Y}_h en de steekproeven in ieder stratum zijn onafhankelijk genomen, dan is \bar{y}_{st} een zuivere schatter van \bar{Y}_{st} met variantie $\sum_h W_h^2 V(\bar{y}_h)$.

De variantie van \bar{y}_{st} hangt slechts af van de variantie van de gemiddelden van de strata. Het stratumgewicht zorgt er ook hier voor dat ieder stratum een evenredige bijdrage levert.

De variantie van y_i in een e.a.s., toegepast op een stratum geeft:

$$V(\bar{y}_h) = S_h^2/n_h$$

Substitutie van $V(\bar{y}_h)$ in $V(\bar{y}_{st})$ geeft:

$$V(\bar{y}_{st}) = \sum_h W_h^2 S_h^2 / n_h$$

Een probleem bij het bepalen van de variantie van de schatter is echter dat men de variantie van y_i per stratum (S_h^2) niet weet. Hiervoor kan men echter de schatter (s_h^2) gebruiken.

En in een g.a.s. is $v(\bar{y}_{st})$ een zuivere schatting van $V(\bar{y}_{st})$.

Dus

$$v(\bar{y}_{st}) = s(\bar{y}_{st})^2 = \sum_h W_h^2 s_h^2 / n_h$$

Bij het bepalen van de variantie van de schatter is uitgegaan van een populatie welke erg groot is tov de steekproef. Als dit niet het geval is, zal de variantie gecorrigeerd moeten worden.

De factor $(N_h - n_h)/N_h$ is een correctie voor een eindige populatie. In de praktijk kan de fpc (finite population corrections) genegeerd worden als de steekproeffractie kleiner is dan 5% en vaak zelfs nog als de steekproeffractie 10% is (21). Het effect van het negeren van de fpc is het overschatten van de standaardfout.

In publicaties worden vaak de standaardfouten van het geschatte populatie-gemiddelde vermeld om aan te geven hoe betrouwbaar de gedane steekproef is.

Standaardfouten worden ook gebruikt om bepaalde steekproef methoden met elkaar te vergelijken. In geval van een gestratificeerde steekproef kan men de gevonden variantie van de schatter vergelijken met dezelfde parameter voor een enkelvoudige aselechte steekproef. Zo kan men vooraf bepalen welke methode het nauwkeurigst dus het meest geschikt is. Tevens wordt gebruik gemaakt van de standaardfout om de grootte van de te nemen steekproef te berekenen.

De standaardfout is per definitie de wortel uit de variantie van de schatter:

$$s(\bar{y}_{st}) = \sqrt{[\sum_h W_h^2 s_h^2 / n_h]}$$

Betrouwbaarheidsintervallen

De formule voor het betrouwbaarheidsinterval van het populatie-gemiddelde is als volgt:

$$\bar{y}_{st} \pm t s(\bar{y}_{st})$$

Waarbij t de waarde is voor de Student of de normale verdeling welke correspondeert met de gekozen betrouwbaarheid.

3.2.4 Schatters voor het multiplicatieve model

Een gestratificeerde random steekproef als de waarnemingen lognormaal verdeeld zijn.

Populatiekarakteristieken

PARAMETER	SCHATTER
centraliteit populatiemediaan	$MY = \exp(\bar{z})$
dispersie variatiecoëfficiënt	$vc = \sqrt{[(\exp(s^2) - 1)]}$
betrouwbaarheidsgrenzen	$z \pm t s/\sqrt{n}$

$$z_i = \log y_i$$
$$\bar{z} = 1/n \sum_i z_i$$

log: natuurlijke logaritme

Bovenstaande s^2 is de steekproefvariantie van $\log y_i$ in een stratum.

Als men tot de conclusie gekomen is dat de variatiebronnen multiplicatief werken op de gemeten variabelen kan men volgens eerder aangevoerde argumenten aannemen dat de data lognormaal verdeeld zijn.

Als de waarnemingen lognormaal verdeeld zijn dwz dat de verdeling scheef naar rechts is, kan men ter correctie, de logaritme van de waarnemingen nemen. De logaritme van een scheve verdeling geeft per definitie een normale verdeling. Na deze logtransformatie kunnen de gegevens geanalyseerd worden.

Centraliteitsparameter

Om de populatieparameters te kunnen schatten moeten de gegevens weer teruggetransformeerd worden. Onderzoekers laten transformatie vaak achterwege omdat teruggetransformeerde gegevens vaak moeilijk interpreteerbaar zijn.

Het teruggetransformeerde gemiddelde is in het algemeen niet gelijk aan het gewone (rekenkundige) gemiddelde. Het verschil tussen het teruggetransformeerde gemiddelde en het gewone

gemiddelde komt voort uit het feit dat ze niet dezelfde parameter schatten: Het teruggetransformeerde gemiddelde schat de mediaan, indien er sprake is van een lognormale verdeling en het gewone gemiddelde schat het populatie-gemiddelde (20).

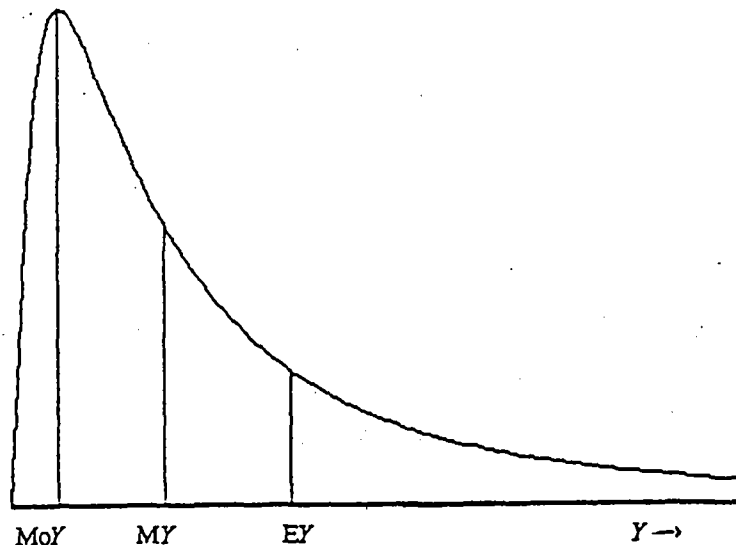


Fig.1. The three centrality parameters, MODE (MoY), MEDIAN (MY) and MEAN (EY) in a lognormal distribution with COEFFICIENT OF VARIATION of 150%.

bron: (20)

Hierboven is een scheve verdeling getekend met de drie centraliteitsmaten: de modus (de waarde waarvoor de verdeling zijn top heeft), de mediaan (de waarde die de verdeling in twee gelijke oppervlakken deelt) en het populatie-gemiddelde (de verwachtingswaarde) (20).

In zo'n scheve verdeling lijkt de mediaan de meest representatieve centraliteitsmaat (20).

De schatter voor de mediaan wordt verkregen door het op log-schaal verkregen gemiddelde: \bar{z} , terug te transformeren.

Dit geeft: $\exp(\bar{z})$.

De aldus verkregen schatter voor de mediaan wordt ook wel het geometrische gemiddelde genoemd ter onderscheiding van het rekenkundige gemiddelde dat een schatter is voor het populatie-gemiddelde.

Dispersieparameter

De variantie geeft de afwijking van het gemiddelde. En gezien bij een lognormale verdeling het teruggetransformeerde gemiddelde gelijk is aan de mediaan, blijkt de variantie hier geen goede maat.

Men zou dus liever over een maat beschikken welke de afwijking van de mediaan aangeeft.

Zo'n maat is de variatiecoëfficiënt.

Een schatter voor de variatiecoëfficiënt in een aselechte steekproef bij een normale verdeling van de waarnemingen is de steekproef-variantiecoëfficiënt.

Deze is gedefinieerd als de standaarddeviatie gedeeld door het berekende gemiddelde.

In formule: s/\bar{y}

Hoe schever de verdeling hoe groter de variatiecoëfficiënt wordt, terwijl de afstand tussen de mediaan en het gemiddelde toeneemt. Een grote variatiecoëfficiënt duidt op een groot verschil tussen het geometrische gemiddelde (de mediaan) en het berekende gemiddelde.

Een goede schatter van de variatiecoëfficiënt (vc) bij een lognormale verdeling wordt verkregen door

$$vc = \sqrt{[\exp(s^2) - 1]} \text{ te nemen.}$$

Waarbij s^2 is de steekproefvariantie van $\log y_i$.

Variantie van de mediaan

Standaardfouten worden door onderzoekers vaak gebruikt om aan te geven hoe betrouwbaar het gevonden gemiddelde wel is. Maar in geval de waarnemingen lognormaal verdeeld zijn, is ipv het gemiddelde de mediaan de juiste centraliteitsparameter. Standaardfouten voldoen hier dus niet.

Een manier om aan te geven hoe betrouwbaar de gevonden mediaan wel is, is het vermelden van de betrouwbaarheids-grenzen.

Men kan deze betrouwbaarheidsgrenzen vinden door ze op log-schaal te berekenen.

Voor de betrouwbaarheidsgrenzen op log-schaal geldt:

$$\begin{aligned} g_1 &= \bar{z} - t s/\sqrt{n} \\ g_2 &= \bar{z} + t s/\sqrt{n} \end{aligned}$$

Voor s wordt dan de standaarddeviatie van $\log y_i$ ingevuld.

De betrouwbaarheidsondergrens en de betrouwbaarheidsbovengrens worden dan gevonden door respectievelijk $\exp(g_1)$ en $\exp(g_2)$ te berekenen. Deze vertegenwoordigen dan de grenzen tov de mediaan op de originele schaal.

Merk op dat $F = \exp(t s/\sqrt{n})$ een vermenigvuldigingsfactor is die gebruikt kan worden om de betrouwbaarheidsgrenzen uit te rekenen. $\exp(\bar{z})/F$ respectievelijk $F * \exp(\bar{z})$ geven de betrouwbaarheidsgrenzen.

Bijvoorbeeld als $F=2$, dan is de ondergrens gelijk aan de mediaan/2 en de bovengrens aan $2 * \text{mediaan}$. De grenzen zijn dus scheef tov de mediaan

Dus door terugtransformatie geven de betrouwbaarheidsgrenzen aan hoe betrouwbaar de mediaan is, als schatter van het populatie-gemiddelde indien er sprake is van een lognormale verdeling.

Om de schatters voor de populatieparameters voor een g.a.s. te verkrijgen uit een e.a.s. als de waarnemingen lognormaal verdeeld zijn, kan men op dezelfde wijze te werk gaan als bij een normale verdeling (zie annex 4).

3.2.5 Schatters als y_i kwalitatieve variabele is

Een gestratificeerde steekproef als y_i kwalitatieve variabele is met slechts twee waarden (ja/nee etc).

Populatiekarakteristieken

PARAMETER	SCHATTER
proportie	$\hat{P} = p_{st}$
dispersie variantie van de aantallen	s_h^2
variantie van de schatter	s_p^2

We willen weten hoe groot het aantal eenheden, de proportie of het percentage is wat in een bepaalde klasse valt.

De meest ideale stratificatie vindt plaats als alle eenheden in C of C' vallen. Men spreekt dan meestal niet over een gestratificeerde steekproef maar over een steekproef voor proporties of percentages (zie annex 2).

Men wil bijvoorbeeld weten welk percentage van de bevolking in een bepaald gebied gebruik maakt van de getroffen en functionerende watervoorzieningen. Men heeft de ervaring uit eerder gedane evaluaties, dat naarmate men verder van de voorziening afwoont, men geneigd is deze minder te gebruiken gezien een andere bron dichterbij is. Het onderzoeksgebied kan dan verdeeld worden in strata naar afstand tot de voorziening.

De proportie

De schatter voor de proportie in het h-de stratum geeft

$$p_h = a_h/n_h$$

De schatter voor de proportie in de totale populatie geeft

$$p_{st} = \sum_h N_h/N * p_h = \sum_h W_h p_h$$

De som van de gewogen stratumproporties vormt hier de schatter voor de proportie in de totale populatie. Er wordt ook hier gebruik gemaakt van het correcte stratumgewicht.

Dispersie

De schatter voor de variantie van de aantallen geeft

$$s_h^2 = p_h q_h * n_h / (n_h - 1)$$

De standaarddeviatie is gedefinieerd als de wortel uit de variantie van de aantallen

Deze wordt weergegeven door:

$$s_h = \sqrt{s_h^2}$$

De variantie geeft de afwijking van het gemiddelde. Als y_i kwalitatieve variabele is, is er geen sprake van een gemiddelde maar van een proportie en een variantie van de aantallen.

De variantie van de geschatte proportie

Hoe betrouwbaar de proportie geschat wordt, anders gezegd hoeveel de geschatte proportie afwijkt van de werkelijke proportie in de totale populatie kan men berekenen door gebruik te maken van de volgende formule:

$$s_p^2 = \sum_h W_h p_h q_h / (n_h - 1)$$

De correctie voor een eindige populatie is $1 - f_h = (1 - n_h/N_h)$. Als in ieder stratum n/N erg klein is mag f_h verwaarloosd worden en ontstaat bovenstaande formule. Voor een eindige populatie moet dus nog een correctie toegepast worden ter grootte van $1 - f_h$.

3.3.1 De toewijzing van de steekproef over de strata

Wanneer er een populatie (N) gekozen is en de grootte van de strata (N_1, N_2, \dots, N_L) is vastgelegd, kan er uit ieder stratum "at random" een steekproef genomen worden. Maar, hoe groot moet die steekproef zijn ?

De stratumsteekproef moet dusdanig zijn dat de nauwkeurigheid van de totale steekproef zo groot mogelijk is en de kosten zo klein mogelijk zijn.

Aan deze eis wordt voldaan als de steekproef in ieder stratum

$$n_h = \frac{d/\sqrt{c_h}}{\sum_h d/\sqrt{c_h}} * n \text{ is.}$$

$$d = N_h S_h$$

c_h = kosten per eenheid in het h-de stratum

Zo'n verdeling van 'n' over de strata wordt een optimale toewijzing genoemd.

De steekproefgrootte (n_h) in ieder stratum is dus evenredig aan de totale steekproefgrootte (n) wat ook voor de hand ligt. De proportie wordt bepaald door de grootte van een stratum (N_h), door de standaarddeviatie in een stratum (S_h) en tevens door de kosten van een eenheid in het h-de stratum.

Dit betekent dat er in ieder stratum een grotere steekproef nodig is als; een stratum groter is
de standaarddeviatie groter is.

Als de kosten per eenheid in een stratum groter zijn, zal dit leiden tot een kleinere steekproef. De term c_h vervalt als de kosten per eenheid in ieder stratum gelijk zijn en dan geldt:

$$n_h = \frac{d}{\sum_h d} * n$$

$$d = N_h S_h$$

Deze verdeling wordt ook wel de Neyman toewijzing genoemd.

Een probleem is echter dat we de populatie-standaard-deviaties niet weten. Het vooraf bepalen van de populatie-standaard-deviaties mbv een kleine steekproef of door een "pilot study" uit te voeren is een duidelijk nadeel.

Ten eerste omdat de vooraf bepaalde standaarddeviaties een schatting zijn van de werkelijke standaarddeviaties en deze dus een zekere fout met zich meebrengen. Ten tweede kost het tijd en inspanning. En ten derde brengt het extra kosten met zich mee.

Een andere mogelijkheid om de grootte van een steekproef in ieder stratum te bepalen, is de evenredige toewijzing. De steekproefgrootte (n_h) in ieder stratum is dan ook evenredig aan n, maar de proportie is gelijk aan het stratumgewicht (W_h).

In formule : $n_h = N_h/N * n = W_h * n$.

Deze toewijzing vereist dus geen steekproef vooraf, maar is eenvoudig en direct toepasbaar.

Maar, geeft een eenvoudig toepasbare toewijzing ook de gewenste nauwkeurigheid.

Volgens Cochran (21) blijkt de optimale toewijzing de kleinste variantie van de schatter te geven: (V_{opt}). De evenredige toewijzing geeft een iets grotere variantie: (V_{prop}), maar deze is toch nog kleiner dan de variantie bij een enkelvoudige aselechte steekproef: (V_{ran}).

In formule: $V_{opt} < V_{prop} < V_{ran}$.

Als de optimale toewijzing niet veel afwijkt van de evenredige toewijzing (dit is zo, als er geen grote verschillen zijn tussen de standaarddeviaties), kan men berekenen hoeveel de variantie van de schatter toeneemt als men de evenredige toewijzing gebruikt. Bij een optimale toewijzing heeft men de standaarddeviaties vooraf moeten bepalen en dat brengt fouten met zich mee. En zo zou het kunnen dat een evenredige toewijzing, ondanks dat deze een grotere variantie oplevert soms wel van 10 a 20% (21), toch nog te prefereren valt. Vooral ook omdat de formule's van de evenredige toewijzing veel simpeler zijn.

De optimale toewijzing voor n_h als y_i kwalitatieve variabele is:

$$n_h = \frac{d / \sqrt{c_h}}{\sum_h d / \sqrt{c_h}} * n$$

$$d = N_h P_h Q_h$$

De populatieproportie in een stratum (P_h) moet hier dus vooraf bepaald worden en Q_h kan men hiervan afleiden. Dit feit brengt de optimale toewijzing weer in discrediet tov de evenredige toewijzing.

Als de kosten per eenheid in ieder stratum gelijk zijn, vervalt de term c_h en kan men het volgende zeggen over de te gebruiken steekproef en toewijzing.

Als de populatieproporties in de strata geen grote verschillen vertonen, geeft een gestratificeerde aselechte steekproef weinig voordeel tov een enkelvoudige aselechte steekproef. Optimale toewijzing levert geen voordelen op boven een evenredige toewijzing als alle P_h 's tussen 0.1 en 0.9 liggen. De optimale toewijzing levert pas voordeel op als de P_h 's liggen tussen 0.001 en 0.5 (21).

3.3.2 De steekproefgrootte

In paragraaf 3.3.1 is gesproken over de toewijzing van de steekproef over de strata, terwijl de steekproefgrootte (n) nog niet bekend was.

De grootte van een te nemen steekproef wordt bepaald omdat men de fout welke gemaakt wordt door het schatten van de populatiekarakteristiek(en) wenst te begrenzen. Verder kan men zeggen dat een passende steekproefgrootte kostenbesparend is.

De steekproefgrootte wordt berekend door de term n uit de formule voor de variantie van de schatter te elimineren, nadat men een waarde gekozen heeft voor de toegestane variantie. In een g.a.s. komt echter ipv de term n de term n_h in de variantie voor. Men moet dus eerst n_h vervangen door een term waarin n wel voorkomt. Deze toewijzing van de steekproef over de strata wordt daarom in de voorgaande paragraaf besproken.

Om de steekproefgrootte (n) te kunnen bepalen moet er eerst een beslissing genomen worden over de fout welke men accepteert. Maw er zal een beslissing genomen moeten worden over de nauwkeurigheid en de betrouwbaarheid welke men verlangt.

Welke nauwkeurigheid verlangd wordt, wordt bepaald door de onderzoeker en deze zal zich vooral laten leiden door het doel waarvoor men de resultaten wil gaan gebruiken.

Onder de nauwkeurigheid (de toegestane fout) verstaan we de maximale afwijking tussen de schatter en de populatieparameter voor een gegeven steekproefgrootte en betrouwbaarheid, bij herhaalde steekproeven.

In formule: $|y - Y| \leq d$

d = de nauwkeurigheid van de schatter

Per definitie is de nauwkeurigheid gelijk aan het halve betrouwbaarheidsinterval.

In formule: $|y - Y| \leq d = t * \text{standaardfout}$

Waarbij t de waarde is van de Student of normale verdeling welke correspondeert met de gewenste betrouwbaarheid.

Bv gew betr b heid	%	80	90	95	99
	t	1,28	1,64	1,96	2,58

De betrouwbaarheid (α) is het risico dat de vereiste nauwkeurigheid (d) groter is.

In formule: $\Pr (|y - Y| \geq d) = \alpha$

Bv Als $\alpha = 0,05$ zal in 5% van de gevallen de populatieparameter niet binnen het interval liggen.
En in 95% van de gevallen ($1 - \alpha = 0,95$) wel. We spreken dan over een 95% betrouwbaarheidsinterval.

Als waarde voor α wordt vaak 0,05 of 0,01 gekozen, wat correspondeert met respectievelijk een 95% en 99% betrouwbaarheidsinterval.

Men wil nu de steekproefgrootte (n) bepalen voor een gestratificeerde random steekproef onafhankelijk van de toewijzing.

In paragraaf 3.2.3 vonden we voor de ongecorrigeerde variantie van de schatter

$$V(\bar{y}_{st}) = \sum_h W_h^2 S_h^2 / n_h$$

We staan een variantie toe voor welke geldt:

$$V = d^2 / t^2$$

Omdat we geen gebruik maken van een bepaalde toewijzing, kan n_h het best vervangen worden door $w_h * n$.

We moeten w_h dus vooraf bepalen evenals S_h .

We kunnen w_h en S_h verkrijgen door het uitvoeren van een pilot studie of door een kleine steekproef vooraf te doen. Eventueel zou men ook gebruik kunnen maken van reeds bestaande literatuur over het onderzoeksonderwerp.

Na substitutie van n_h door $w_h * n$, kan n geëlimineerd worden en dat geeft:

$$n_0 = 1/V * \sum_h W_h^2 S_h^2 / w_h$$

Als n_0/N niet te verwaarlozen is, dus er een correctie noodzakelijk is voor een eindige populatie dan geldt:

$$n = n_0 / (1 + 1/VN * \sum_h W_h S_h^2)$$

Voor een evenredige toewijzing geldt dat:

$$w_h = W_h = N_h/N$$

en hieruit volgt:

$$n_0 = \sum_h W_h S_h^2 / V \quad \text{en} \quad n = n_0 / (1 + n_0/N)$$

Als er gekozen is voor een optimale toewijzing van de steekproefgrootte (n) over de strata, kan men n verkrijgen door substitutie van de formule voor n_h bij optimale toewijzing (zie 3.3.1) in de formule voor de variantie (zie 3.2.3).

Hieruit volgt voor vastgestelde variantie

$$n = [(\sum_h W_h S_h c_h) \sum_h W_h S_h / c_h] / (V + 1/N * \sum_h W_h S_h^2)$$

Als de populatie groot is tov de steekproef vervalt de rechter term in de noemer.

Bij gelijke kosten voor iedere eenheid vervalt de term c_h .

Als y_i kwalitatieve variabele is, gaan we ook weer uit van een populatie die groot is tov de steekproef en berekenen de steekproefgrootte voor n_0 . Mocht dan blijken dat n_0/N niet te verwaarlozen is dan kan een correctie uitgevoerd worden voor een eindige populatie door mbv n_0 , n te berekenen.

Bij een evenredige toewijzing geldt:

$$n_0 = 1/V * \sum_h W_h P_h Q_h \quad \text{en} \quad n = n_0 / (1 + n_0/N)$$

Bij een optimale toewijzing geldt:

$$n_0 = 1/V * \sum_h W_h^2 P_h Q_h \quad \text{en} \quad n = n_0 / (1 + 1/NV * \sum_h W_h P_h Q_h)$$

Als p_h een benadering is voor P_h , volgt hieruit dat men Q_h kan vervangen door q_h .

Hierboven zijn we om n te kunnen vaststellen ervan uitgegaan dat we bij het schatten van het populatie-gemiddelde de absolute fout onder controle wilde houden.

Nu willen we de relatieve fout binnen gestelde grenzen houden.

$$\Pr [|(y - Y)/Y| \geq r] = \Pr [|y - Y| \geq rY] = \alpha$$

$$rY = t * \text{standaardfout}$$

Bij een evenredige toewijzing geldt:

$$n_0 = (t^2 * \sum_h W_h S_h^2) / r^2 Y_h^2 = 1/C * (\sum_h W_h * S_h^2 / Y^2)$$

$$n = n_0 / (1 + n_0/N)$$

De populatiekarakteristiek waarvan n afhangt is hier S_h/Y_h ipv S_h .

Voor een lognormale verdeling van de waarnemingen zijn de variatiecoëfficiënten homogeen.

Om de steekproefgrootte te berekenen als de waarnemingen scheef verdeeld zijn kan men beter gebruik maken van de relatieve fout.

De variatiecoëfficiënt voor een normale verdeling: S_h/Y_h , kan men vervangen voor de variatiecoëfficiënt van een lognormale verdeling: $[\exp(s^2) - 1]$.

Waarbij s^2 de steekproefvariantie is van $\log y_i$ en $C = r^2/t^2$.

4. Conclusie

Om het effect van watervoorzieningen en sanitaire faciliteiten op de gezondheid te onderzoeken, worden experimentele studies aanbevolen.

Bij een zuiver experiment dienen alle factoren, behalve de experimentele factor, zoveel mogelijk constant te blijven. Tevens dient de voorziening 'ad random' te worden toegewezen. Gezien bij gezondheids-effect-studies vaak niet aan deze eisen voldaan kan worden, zal men gebruik moeten maken van studies welke het experiment zoveel mogelijk benaderen.

Om onbekende versturende variabelen uit te schakelen zal men zoveel mogelijk gebruik moeten maken van randomisatie.

Bekende versturende variabelen kunnen worden uitgeschakeld met behulp van stratificatie.

Stratificatie is eenvoudig toepasbaar. Een gestratificeerde aselechte steekproef kan gezien worden als de gewogen som van een aantal enkelvoudige aselechte steekproeven. Echter bij een gestratificeerde steekproef dient men vooraf de grootte van de verschillende strata vast te stellen. Tevens dient men de steekproefgrootte en de verdeling van deze over de strata te bepalen. Echter een evenredige steekproef ligt het meest voor de hand: grote strata leveren dan absoluut ook een groter aandeel dan kleine strata.

Om de populatiekarakteristieken in een gestratificeerde aselechte steekproef te kunnen schatten is het van belang of men met het additieve dan wel het multiplicatieve model te maken heeft. De aanname van normaliteit van de waarnemingen is gebaseerd op de meer fundamentele aanname dat variatiebronnen additief werken. Als de variatiebronnen multiplicatief werken kan men zeggen dat de waarnemingen lognormaal verdeeld zijn.

Als men met het additieve model te maken heeft, wordt het rekenkundige gemiddelde als beste centraliteitsparameter gezien. Bij het multiplicatieve model wordt het geometrische gemiddelde gezien als de meest representatieve centraliteitsmaat.

Bij het additieve model geeft de standaardfout aan hoe betrouwbaar het geschatte gemiddelde is. Multiplicatief maakt men geen gebruik van het gemiddelde en voldoet de standaardfout dus niet. We kunnen nu een betrouwbaarheidsinterval aangeven aan weerszijde van de mediaan.

ENKELE DEFINITIES

Het nemen van een steekproef

Het selecteren van een deel van een populatie mbv een bepaalde steekproefmethode om hiermee bepaalde karakteristieken binnen een populatie te schatten.

Meestal toegepast als de financiële middelen of de beschikbare tijd onvoldoende is om de hele populatie te onderzoeken.

Dit schatten van populatiekarakteristieken gaat gepaard met het maken van fouten (variantie), welke niet gemaakt kunnen worden als men de hele populatie onderzoekt.

Verstorende variabelen

Variabelen anders dan de onderzoeksvariabelen, welke van invloed zijn op de indicator (onafhankelijke variabele).

Als men bijvoorbeeld de relatie tussen een watervoorziening en diarree-morbiditeit wil onderzoeken is de toiletvoorziening een verstorende variabele, tenzij de toiletvoorziening in de interventiegroep en de contrôlegroep gelijk is.

Variantie van y

Een maat voor de gespreidheid van de waarnemingen.

De standaarddeviatie is gedefinieerd als de wortel uit de variantie van y.

Variantie van de schatter

Is een maat voor de nauwkeurigheid van de steekproef.

De standaardfout is gedefinieerd als de wortel uit de variantie van de schatter.

s^2 versus σ^2

In de meeste literatuur wordt " s^2 " gebruikt als men de steekproefvariantie wil aanduiden. De term " σ^2 " wordt gebruikt om populatievariantie aan te geven.

De populatievariantie waarbij als noemer "n-1" wordt gebruikt, wordt met " S^2 " aangeduid.

Mediaan

De mediaan deelt het oppervlak onder een functie in twee gelijke delen.

De steekproefmediaan is de middelste waarneming in grootte als het aantal waarnemingen oneven is. Als het aantal waarnemingen even is, noemt men het midden van de twee middelste waarnemingen de mediaan.

STEEKPROEF VOOR PROPORTIES EN PERCENTAGES

Er wordt een steekproef genomen van een populatie, waarbinnen alle eenheden zijn op te delen in 2 klassen, C en C'.

Men wil bijvoorbeeld weten hoeveel huishoudens gebruik maken van een geïnstalleerde pomp.

Het antwoord op de vraag of men gebruik maakt van de pomp zal simpel "ja" of "nee" zijn.

"Ja" valt dan in C en "nee" in C'.

Men definieert;

$$\begin{aligned} y_i &= 1 && \text{als de eenheid in C valt} \\ y_i &= 0 && \text{als de eenheid in C' valt} \end{aligned}$$

Verder wordt er uitgegaan van de formule's, welke gebruikt worden bij een enkelvoudige aselechte steekproef.

Het aantal huishoudens in de steekproef, dat gebruik maakt van de pomp is dan:

$$y = \sum_i y_i = a$$

De proportie gebruikers in de steekproef vindt men door y te bepalen

$$\bar{y} = \sum_i y_i/n = a/n = p$$

Conclusie:

aantal eenheden in C	proportie in C
steekp populatie	steekp populatie
a A	p = a/n P = A/N

En ten overvloede N=populatiegrootte
n=steekproefgrootte

De andere parameter is de schatter voor populatievariantie van de aantallen

Deze ontstaat door substitutie van

$$\sum_i y_i^2 = a = np, \quad \bar{y} = p \quad \text{en} \quad 1 - p = q$$

in de formule voor de variantie van y_i bij een enkelvoudige aselechte steekproef.

Het resultaat geeft: $s^2 = pq/(n-1) * n$

De variantie van de geschatte proportie is dan:

$$s_p^2 = pq/(n-1)$$

Bovenstaande formule is de ongecorrigeerde variantie van de schatter. Deze kan gebruikt worden als men uit kan gaan van een oneidige populatie. Bij een eindige populatie moet bovenstaande formule gecorrigeerd worden met de factor $(N - n)/N$. Merk op dat als N erg groot is tov n , $1 - n/N$ tot één nadert.

EEN ENKELVOUDIGE ASELECTE STEEKPROEF

In een enkelvoudige aselecte steekproef is de schatter voor het populatie-gemiddelde gelijk aan het steekproefgemiddelde:

$$\bar{y} = 1/n \sum_i y_i$$

De dispersieparameter wordt weergegeven door de variantie van Y_i :

$$s^2 = \sum_i (y_i - \bar{y})^2 / (n - 1)$$

De variantie van de schatter wordt geschat door:

$$s^2/n$$

En de standaardfout is gedefinieerd als de wortel uit de variantie van de schatter:

$$s/\sqrt{n}$$

Bij een eindige populatie, dat wil zeggen als de steekproef fractie groter is dan 5%, dient de variantie van de schatter (ook de standaardfout) gecorrigeerd te worden door een factor $1 - f$ (of $\sqrt{1-f}$ voor de standaardfout), gezien anders de variantie dus ook de standaardfout overschat wordt.

Voor de betrouwbaarheidsgrenzen geldt:

$$\text{ondergrens } g_1 = \bar{y} - t s/\sqrt{n}$$

$$\text{bovengrens } g_2 = \bar{y} + t s/\sqrt{n}$$

Ook voor de betrouwbaarheidsgrenzen is er een correctie noodzakelijk voor een eindige populatie, omdat de standaardfout hier een rol speelt.

De waarde t staat voor de Student of de normale verdeling bij een gekozen betrouwbaarheid.

HOE VERKRIJGT MEN EEN G.A.S. UIT EEN E.A.S..

Een g.a.s. is eigenlijk de gewogen som van een aantal e.a.s. steekproeven. Het stratungewicht speelt hier een belangrijke rol.

Schatters voor populatiekarakteristieken van het additieve model:

gemiddelde	\bar{y}	$\sum_h W_h \bar{y}_h$
variantie per stratum	s^2	s_h^2
variantie van schatter	s^2/n	$\sum_h W_h^2 s_h^2 / n_h$
standaardfout (stf)	s/\sqrt{n}	$\sqrt{[\sum_h W_h^2 s_h^2 / n_h]}$
betrouwbaarheidsgrenzen	$\bar{y} \pm t \text{ stf}$	$\sum_h W_h \bar{y}_h \pm t \text{ stf}$

Schatters voor populatiekarakteristieken van het multiplicatieve model:

mediaan	$\exp(\bar{z})$	$\exp(\sum_h W_h \bar{z}_h)$
variatiecoëfficiënt per stratum	$\sqrt{[\exp(s^2)-1]}$	$\sqrt{[\exp(s_h^2)-1]}$
betrouwbaarheidsgr	$\exp(\bar{z}) / F$ en $\exp(\bar{z}) * F$	$\exp(\sum_h W_h \bar{z}_h) / F_{st}$ en $\exp(\sum_h W_h \bar{z}_h) * F_{st}$

Waarbij $z_i = \log y_i$, $\bar{z} = \sum_i z_i / n$ en s^2 is de steekproefvariantie van $\log y_i$.

$$F_{st} = \exp\{\sqrt{[\sum_h W_h^2 s_h^2 / n_h]} * t\}$$

$$F = \exp\{s/\sqrt{n} * t\}$$

log: natuurlijke logaritme

literatuur

- 1 McJunkin FE. Water and Human Health. USAID, 1983
- 2 Kusin JA, Sjoerdsma A, Roosmalen-Wiebenga M van. Ondervoeding in de tropen. Amsterdam:KIT, 1988:111 pag
- 3 Schneider RE, Schiffman M, Faigenblum J. The potential effect of water on gastrointestinal infections prevalent in developing countries. Am J Clin Nutr. 1978;31:2089-99
- 4 Cutting WAM, Hawkins P. The role of water in relation to diarrhoeal disease. J Trop Med Hyg. 1982;85:31-9
- 5 Benoterraf B, Unanue ER. Textbook of immunology. London:Williams&Wilkins, 1979
- 6 Black RE, Merson MH, Rowe B, et al. Enterotoxigenic E.coli:acquired immunity and transmission in an endemic area. Bull WHO. 1981;59:263-8
- 7 Cunningham AS. Morbidity in breastfed and artificially fed infants II. J Pediatr. 1979;95:685-9
- 8 WHO Contemporary patterns of breastfeeding. Report on the WHO collaborative study on breastfeeding. WHO Geneva:1981
- 9 Victoria CG, Smith PG, Vaughan JP, et al. Infant feeding and deaths due to diarrhea. A case-control study. Am J Epid. 1989;129(5):1032-41
- 10 Black RE, Brown KH, Becker S, et al. Contamination of weaning foods and transmission of enterotoxigenic E.coli diarrhoea in children in rural Bangladesh. Trans Roy Soc Trop Med Hyg. 1982;76:259-64
- 11 Barrel RAE, Rowland MGM. Infant food as a potential source of diarrhoeal disease in rural West Afrika. Trans Roy Soc Trop Med Hyg. 1979;73:85-90
- 12 Briscoe J. The role of water supply in improving health in poor countries (with special reference to Bangladesh) Am J Clin Nutr. 1978;31:2100-13
- 13 Feachem FG. Interventions for the control of diarrhoeal diseases among young children:promotion of personal and domestic hygiene. Bull WHO. 1984;62(3):467-76

- 14 Esrey SA, Habicht JP. The impact of improved Watersupplies and Excreta Disposal Facilities on Diarroal Morbidity, Growth and Mortality among children. USAID,1985
- 15 Blum D, Feachem RG. Measuring the impact of Watersupply and sanitation investments on Diarrhoeal Diseases:Problems of Methodology. Int J Epid.1983;12(3):357-65
- 16 Stuurmans F. Epidemiologie, theorie, methoden en toepassing. Nijmegen,1982:338 pag
- 17 Kleinbaum DG, Kupper LL, Morgenstern H. Epidemiologic Research. Principles and Quantitative methods.New York:Van Nostrand Reinhold Company,1982:529 pag
- 18 Minimum Evaluation Procedure (MEP) for water supply and sanitation projects. Unpublished document. WHO,1983:ETS/83.1
- 19 Miettinen OS. Stratification by a multivariate confounder score. Am J Epid.1976;104(6):609-20
- 20 Slob W. Strategies in Applying Statistics in Ecological Research. Amsterdam,1986:proefschrift
- 21 Cochran WG. Sampling Techniques. New York:Wiley,1977:428 pag
- 22 Konijn HS. Statistical theory of sample survey design and analysis. Amsterdam:NHC,1973:429 pag
- 23 Sudman S. Applied Sampling. Quantative studies in social relations. New York:Academic Press,1976:249 pag

De vakgroep Gezondheidsleer en Tropische Hygiëne geeft deze serie uit. Opgenomen worden scripties, nota's, verslagen en rapporten van studenten en medewerkers van de vakgroep Gezondheidsleer en Tropische Hygiëne. Een lijst van publicaties verschenen in de jaren 1974 t/m 1986 is op aanvraag verkrijgbaar bij het secretariaat van de vakgroep.

Publicaties verschenen sinds 1987:

- f 14,75 1988 - 311 Verslag
Binnenklimaat in woningen. Een benadering van gezondheidsklachten en woningkenmerken met behulp van vragenlijsten
Door: Monique Waegemaekers
- f 9,00 1988 - 312 Verslag
Inventarisatie van onderzoek in Nederland op het gebied van vitamines en mineralen met betrekking tot de gezondheid van de mens
Door: M. de Winkel, S. Pieterse, C. Hoitink en M. Wattel
- f 9,00 1988 - 313 Verslag
Longfunctieonderzoek in een meelfabriek
Door: Remko Houba en Tonnie Zweers
- f 6,00 1988 - 314 Verslag
Analyse van longfunctie en vóórkomen van luchtwegklachten bij schoolkinderen in Alkmaar en Egmond
Door: Dr. ir. B. Brunekreef
- f 13,50 1988 - 315 Verslag
Melk, goed voor elk? Een onderzoek naar overgevoeligheid voor melk en/of ei bij 839 kinderen
Door: Anneke Blokstra, Helmi Naber en Hanneke Oude Elberink
- f 12,75 1988 - 316 Verslag
De leefstijl en gezondheidssituatie in relatie tot de groei bij 5- en 6-jarige kinderen. Een onderzoek in de gemeentekring Almelo
Door: Marga Ocké en Annette Stolwijk
- f 21,00 1988 - 317 Verslag
Zware metalen in Utrecht
Door: Gon Ferdinandus en Carla van Egmond
- f 14,25 1988 - 318 Verslag
A survey on water use and hygiene in a rural area on mount Elgon, West-Kenya
Door: Dolores Brouwer

- f 8,00 1988 - 319 Verslag
Association between selenium status and cardiovascular risk factors in a healthy Dutch population
Door: Sandra Bukkens en Nelleke de Vos
- f 8,25 1988 - 320 Verslag
Magnesium in relatie tot bloeddruk en mortaliteit
Door: Margriet Penders
- f 8,25 1988 - 321 Verslag
IJzerstatus en het acute myocardinfarct
Door: Yvonne van der Schouw en Pauline van der Veeken
- ** 1988 - 322 Verslag
De ontwikkeling van een geïntegreerde zorgvoorziening voor bejaarden
Door: J.M.J. Ketelaar en J.A.J. Smets
- f 7,25 1988 - 323 Verslag
De relatie tussen polsfrequentie respectievelijk bloeddrukverandering en totale mortaliteit
Door: Manon J.E. Penris en Frederike C. Sieders
- ¹ 1988 - 324 Verslag
Validering van de schriftelijke vragenlijst van een cohortonderzoek naar voeding en kanker: de proefopzet
Door: Monique Al
- f 6,75 1988 - 325 Verslag
Inkomen en sterfte. Resultaten van een 25 jaar vervolgonderzoek bij mannelijke Amsterdamse ambtenaren
Door: Dianne Aben en Nicole Mares
- f 7,50 1988 - 326 Verslag
Onderzoek naar acute longfunctieveranderingen bij varkenshouders
Door: Roif van Zwieten
- f 8,50 1988 - 327 Verslag
Oriënterend onderzoek naar de blootstelling van spuiters in de chrysantenteelt onder glas aan het gewasbeschermingsmiddel Curamil
Door: Hillie Dillema en Angelo Rutten

¹ Dit verslag is verkrijgbaar bij:
CIVO-TNO
Postbus 360
3700 AJ ZEIST

- f 8,00 1988 - 328 Literatuuronderzoek
Huidtests en de RAST: een literatuuronderzoek naar factoren die de resultaten beïnvloeden
Door: Marti J. van Liere
- f 7,25 1988 - 329 Literatuuronderzoek
Het verkorten van protocollen voor de bepaling van a-specifieke bronchiale hyperreactiviteit
Door: Eric Zuidweg
- f 15,50 1988 - 330 Verslag
Verslag van een arbeidshygiënisch onderzoek in een loopvlakvernieuwingsbedrijf van vliegtuigbanden
Door: Wouter de Haan, Harm Ikink en Jan Koppejan
- f 9,50 1988 - 331 Verslag
Effektevaluatie van het onderwijsprogramma "Gezond Gedrag". Een onderzoek onder leerlingen van het lager agrarisch onderwijs
Door: Josien Breken en Koby van der Knaap
- f 19,75 1988 - 332 Verslag
Geboortegewicht gewogen. Een onderzoek naar determinanten van het geboortegewicht in Lelystad en Urk en geboortegewicht als gezondheids-indikator
Door: Marlene van der Star en Corrien Waardenburg
- f 7,50 1988 - 333 Verslag
Bloeddruk en mineralen bij Nederlandse ouderen
Door: Zandrie Hofman
- f 5,25 1988 - 334 Verslag
De relatie tussen het serum linolzuur, arachidonzuur, N-3 vetzuren en de m/v-ratio en acuut myocard infarct
Door: Peter Rokebrand
- ** 1988 - 335 Verslag
Intensieve veehouderij en gezondheid. Een inventariserend onderzoek naar de gevolgen van luchtverontreiniging door intensieve veehouderij op de (volks)gezondheid
Door: Catrien van Dam en Rob van Strien
- f 9,50 1988 - 336 Literatuuronderzoek
Literatuuronderzoek naar de registratie van medische gegevens, met name het pilgebruik, door huisartsen in Nederland.
Overeenkomst in gerapporteerd pilgebruik van voorschrijver en gebruikster; gerelateerd aan enkele kenmerken van de voorschrijver en van de vrouw
Door: Monique Mentjens
- f 5,00 1988 - 337 Verslag
Validiteit en betrouwbaarheid van het welzijn-formulier; een onderzoek naar de toepassing ervan bij 106 mensen uit zeven beroepen
Door: Lucas Wiessing
- f 8,00 1988 - 338 Verslag
Seroconversie voor Hepatitis A virus bij een groep reizigers in de periode 1979-1980
Door: Annegré de Roos
- f 11,25 1988 - 339 Verslag
Leerlingenbegeleiders en gezondheidsproblemen van leerlingen. Onderzoek onder schoolleiding en docenten van het voortgezet onderwijs in de regio Geldrop-Valkenswaard
Door: Alida Wapstra en Nynke van Zanden
- f 7,25 1988 - 340 Verslag
Laboratory studies on the behaviour of mosquitoes foraging on a bait protected by permethrin impregnated bednetting
Door: J. Voorham
- f 6,00 1988 - 341 Verslag
De Quetelet-index (gewicht/lengte²) als maat voor overgewicht bij Nederlandse ouderen
Door: Vera Minten
- f 17,75 1988 - 342 Literatuuronderzoek
Beroepsmatige blootstelling aan toxische stoffen in de graanoverslag en de daarmee samenhangende risico's voor de gezondheid
Door: Drs. A.M. Kremer
- f 9,75 1988 - 343 Verslag
De relatie tussen geïsoleerde systolische hypertensie en totale mortaliteit - geanalyseerd volgens twee analysemethoden
Door: Greetje van den Ban en Ellen Kampman
- f 18,50 1988 - 344 Verslag
Leefgewoonten en gezondheid van ouderen. Een vooronderzoek onder zelfstandig wonende ouderen in de gemeente Rhenen
Door: Paula Dijkema en Annette Stafleu
- f 11,25 1988 - 345 On the nature of time and aging
Door: Johannes J.F. Schroots
- f 7,50 1988 - 346 Verslag
Combinatie van risicofactoren voor hart- en vaatziekten bij ouderen in Nederland
Door: Sandra Borsboom en Clothilde Bun

f 6,75 1988 - 347 Literatuuroverzicht
Graanstof en gezondheid - een literatuuroverzicht
Door: Tjabe Smid

f 8,75 1988 - 348 Verslag
Vragen aan een vrouwengezondheidscentrum. 1986, 1987,en hoe verder?
Door: Meta Moerman

f 6,50 1988 - 350 Verslag
Domestic air pollution and respiratory infections in young Gambian children: analysis of data from a cross-sectional study
Door: Arnold Bergstra

f 9,75 1988 - 351 Verslag
Voeding, voedingstoestand en sociaal-economische variabelen bij 10-11 jarige jongens
Door: Ingrid Maas

** 1988 - 352 Scriptie
Ziektefrequentie- en associatiematen in de epidemiologie
Door: E.G. Schouten

f 13,50 1988 - 353 Rapport
Gezondheidsenquête in Noordelijk Noord-Holland: een vooronderzoek
Door: Wyanne Huitema en Erik Maarsingh

f 15,00 1988 - 354 Verslag
Lood om oud ijzer. Loodverontreiniging aan de Ruigenhil te Alblasterdam en het gezondheidsrisico voor kinderen
Door: Maarten Louer en Yvonne Tieleman

** 1988 - 355 Verslag
Ontwikkeling van een model ter beoordeling van het extra gezondheidsrisico bij bodemverontreiniging en uitwerking hiervan voor Rotterdamse volkstuinen
Door: Thomas Huy, Bert Scheeper, Kees de Schipper en Jan-Paul Zock

f 9,00 1988 - 356 Verslag
Aspecten van gezondheidszorg in Zuid-Soedan
Door: J.J.F. van Rooij

f 10,50 1988 - 357 Verslag
Vleesconsumptie in relatie tot de prevalentie van kanker
Door: Jeanne van Loon

f 10,25 1988 - 358 Verslag
De organisatie van en voorlichting over inenting, in het bijzonder bij negenjarigen
Door: Christine Bulsink en Irene Gosselink

f 11,00 1988 - 359 Report
Water, sanitation and health in Brazil. A health impact survey in Apiiai-Mirim and Ferreira dos Matos, Sao Paulo
By: Andrea Berghuizen and Inge Lardinois

f 6,75 1988 - 360 Verslag
De relatie tussen de consumptie van melk en melkproducten en borstkanker
Door: Jacqueline Dekker en Jos Lamers

f 8,00 1988 - 361 Verslag
Marginale voeding en borstkanker
Door: Ingrid Gelissen en Anja Huisman

f 8,00 1988 - 362 Verslag
Een longitudinaal onderzoek naar (effecten op) de longfunctiegroei bij basisschoolkinderen in de IJmondregio
Door: Marc Smeets

f 11,25 1988 - 363 Verslag
Determinanten van de voedingstoestand bij 10-11 jarige jongens
Door: Carola Schrijvers

f 8,00 1989 - 364 Verslag
Een schatting van de ingestie van bodem- en stofdeeltjes door jonge kinderen
Door: Pieter Clausing, Bert Brunekreef en Joop van Wijnen

f 12,00 1989 - 365 Verslag
Ziek zijn, dat is iets wat je overkomt. Onderzoek naar etiologie en geneesmogelijkheden in Noord-Benin
Door: Jacqueline Franssens

f 15,00 1989 - 366 Verslag
Werkloosheid en gezondheid: een aanzet tot onderzoek
Door: Liliame Begheijn

f 15,75 1989 - 367 Verslag
Onderzoek naar blootstelling en gezondheidseffecten van caseïnestof bij werknemers van "Het Kaasmerk" te Leiden
Door: Erik van Eek en Lonneke de Rijk

- f 19,75 1989 - 368 Report
Primary health care and hygiene education and extension in Indonesia. A case study in Indramayu
By: Inge Verdonk
- f 21,75 1989 - 369 Verslag
De relatie tussen selenium en borstkanker; een patiënt-controle onderzoek
Door: Reggy P.J. van der Wielen
- f 28,00 1989 - 370 Verslag
De leefsituatie van ouderen in de regio Geldrop-Valkenswaard. Een schriftelijke enquête onder zelfstandig wonende ouderen
Door: Cora Tabak en Marjon Theunissen
- f 19,25 1989 - 371 Scriptie
Gezondheidszorg in India. Welke? Onderzoek naar het functioneren van verschillende gezondheidszorgsystemen en het denken en handelen van de bevolking
Door: Annet Zwarteveen
- f 12,50 1989 - 372 Verslag
Verspreiding van HIV en gedragsbeïnvloeding
Door: Roland Prinsen
- ** 1989 - 373 Verslag
Arbeidshygiënisch onderzoek in de tabakverwerkende industrie naar de expositie van stof, nicotine en schimmelsporen en de effecten op de luchtwegen
Door: Evert Verberne, Helma Drost, Huub Agterberg, Mark Nieuwenhuijsen en Martie van Tongeren
- f 11,50 1989 - 374 Verslag
Selenium, α -tocoferol en meervoudig onverzadigde vetzuurstatus bij patiënten met coronair atherosclerose: de antioxidant hypothese
Door: Johan Melse en Ellen Verheul
- ** 1989 - 375 Rapport
Blootstelling aan stof in de mengvoederindustrie in Nederland
Door: Ir. G.B.M. Mensink, Ir. T. Smid en Ir. D.J.J. Heederik
- ** 1989 - 376 Rapport
Longfunctieveranderingen door blootstelling aan organisch stof en endotoxine in de Nederlandse mengvoederindustrie
Door: Ir. R. Houba, Ir. D. Heederik en Ir. T. Smid
- f 17,50 1989 - 377 Verslag
Arbeidshygiënische evaluatie van de toepassing van gewasbeschermingsmiddelen bij 10 loonwerkbodrijven
Door: Mark Nieuwenhuijsen en Martie van Tongeren
- f 22,00 1989 - 378 Verslag
Kom binnen, de koffie staat al klaar!! Een mondeling onderzoek naar de levensomstandigheden van zelfstandig wonende ouderen in de regio Geldrop-Valkenswaard
Door: Fieke van Casteren en Lisette Tiessens
- ** 1989 - 379 Verslag
De organisatie van dagopvangprojecten in de ouderenzorg. Een onderzoek naar het verband tussen de organisatiekenmerken van dagopvangprojecten en de deelnemerskenmerken
Door: T. Mandersloot
- f 8,00 1989 - 380 Verslag
Longfunctieonderzoek bij een aan organisch stof blootgestelde populatie bij een overslag bedrijf
Door: Dick Heederik, Remko Houba en Liesbeth Preller
- f 25,75 1989 - 381 Report
An inventory of watersupply, sanitation, health- and socio-economic status in four communities in the municipality of Capao Bonito, Sao Paulo Sate, Brazil
Door: Winfried Overbeek and Hennie van Vree
- ** 1989 - 382 Verslag
Schatten van blootstellingen: een onderzoek naar het nut van kwalitatieve blootstellingsschattingen en naar de (on)zin van indicatieve metingen
Door: Wendel K. Post
- f 15,00 1989 - 383 Verslag
Serumurinezuur spiegels bij ouderen; correlaten en gevolgen van verhoogd serumurinezuur
Door: Harriët Eshuis en Henriëtte Loenen
- f 10,75 1989 - 384 Verslag
Vergelijking van een buitenlandse job-exposure matrix met Nederlandse gegevens
Door: Wouter de Haan
- f 6,75 1989 - 385 Verslag
Acute longfunctieveranderingen door blootstelling aan organisch stof en endotoxine in de Nederlandse mengvoederindustrie
Door: Ir. R. Houba, Ir. D.J.J. Heederik & Ir. T. Smid

- | | | | | | |
|---------|------------|--|---------|------------|--|
| f 13,75 | 1989 - 386 | <u>Verslag</u>
Mantelzorg en gezondheidstoestand in de stad Utrecht. Een secundaire analyse van de Gezondheidspeiling Utrecht 1986-1987
Door: H.H. Marbus | f 14,25 | 1989 - 394 | <u>Verslag</u>
Serologisch onderzoek bij een kunststofgarenfabriek
Door: Leo van der Biessen |
| f 17,00 | 1989 - 387 | <u>Verslag</u>
Arbeidsomstandigheden in de fruitteelt. Een onderzoek naar de blootstelling van fruitteelers aan het gewasbeschermingsmiddel Captan
Door: L. van Amelsvoort, P. Duijzings en T. Huy | f 12,00 | 1989 - 395 | <u>Literatuuronderzoek</u>
Methoden van reproductie-epidemiologie. Een literatuuronderzoek met nadruk op de reproductieve effecten van pesticiden
Door: Karolien Westveer |
| | 1989 - 388 | <u>Verslag</u>
Alternatieven voor chemische gewasbescherming in de bollenteelt
Door: G.M. Bouman, H.A. Dillema, J.D. van der Laken, A.J.K. Sledsens en H.J. Weckselor | f 17,50 | 1989 - 396 | <u>Verslag</u>
Een ontwerp voor epidemiologisch onderzoek naar intolerantie voor voedseladditieven. De relatie tussen de consumptie van benzoaten (E210-E219) en het ontstaan van intolerantie voor deze stoffen
Door: Heleen Bot en Monique Sanders |
| f 14,50 | 1989 - 389 | <u>Verslag</u>
Metabole instelling van type 2 diabetes mellitus en cardiovasculaire complicaties. Een poging tot het ontwerpen van onderzoek
Door: Maribel van de Pas en Cecile Spiertz | f 18,50 | 1989 - 397 | <u>Verslag</u>
Wensen van ouders ten aanzien van de zuigenliniezorg van de kruisvereniging. Een aanzet tot het ontwikkelen van een instrument
Door: M.A.C. van Bree, J.P. Lania en M.C. de Winkel |
| f 16,00 | 1989 - 390 | <u>Verslag</u>
Vervolgonderzoek naar de longfunctie en luchtwegaandoeningen bij varkenshouders
Door: Leo van der Biessen en Nienke van Kerkhof | f 16,25 | 1989 - 398 | <u>Verslag</u>
Een schaal om riskantie in verband met HIV-besmetting en verandering van homoseksueel gedrag van mannen te bepalen
Door: Rosan Blijlevens en Ella Visser |
| f 15,75 | 1989 - 391 | <u>Verslag</u>
Het QT-interval in relatie tot de cardiovasculaire sterfte bij Amsterdamse ambtenaren. Een pilotstudie.
Door: Miriam Stoel | f 18,00 | 1989 - 399 | <u>Verslag</u>
Borstkanker en de inneming van voedingsvezel en B-caroteen
Corine Kolb en Petra Verhoef |
| gratis | 1989 - 392 | Organisatie en werkwijze van de vakgroep Gezondheidsleer. Studiejaar 1989/90 | f 16,00 | 1989 - 400 | <u>Verslag</u>
Ouderparticipatie in een instituut voor geestelijk gehandicapten. Een verkennend onderzoek naar inhoud en organisatorische vormgeving van ouderparticipatie
Door: Renske Miedema en Petra van Wezel |
| f 17,75 | 1989 - 393 | <u>Verslag</u>
Longfunctieonderzoek in een wollenstoffenfabriek
Door: André Holtrop, Fenneke Linker, Wendel Post en Jan-Paul Zock | f 20,50 | 1989 - 401 | <u>Rapport</u>
The management of confidentiality in general medical practice in the Netherlands and the USA.
Door: C.J. Lako |
| | | | f 15,50 | 1989 - 402 | <u>Verslag</u>
De combinatie van voedingsfactoren in relatie tot borstkanker
Door: Edith van Leer en Annelies Rietdijk |

¹ Dit verslag is verkrijgbaar via:
De Wetenschapswinkel van de
Landbouwuniversiteit
Postbus 101
6700 AC WAGENINGEN
08370 - 83908/84146/84661

onder vermelding van rapportnr. 35

f 15,50 1989 - 403 Verslag
De relatie tussen een verlengd QT-interval in het ECG en totale mortaliteit
Door: Mariette van Ommering

f 25,00 1990 - 404 Proefschrift
Epidemiological studies of the relationship between occupational exposures and chronic non-specific lung disease
Door: Dick Heederik

f 12,00 1990 - 405 Scriptie
Sociale netwerken, materiële hulp en welbevinden
Door: J.M. van Langen

f 17,00 1990 - 406 Verslag
Vrije en enzymatisch gecontroleerde radicaalreacties en kanker
Door: Annette Bergstra en Bernadette Nagelhout

f 16,25 1990 - 407 Verslag
Het gebruik van de zegels in het babyboekje voor migranten ter ondersteuning van voedingsadviezen
Door: Petra Dongelmans

f 9,25 1990 - 408 Verslag
De invloed van ozon op de peakflow van sportende kinderen
Door: René van den Berg en Paul Kosterink

f 12,00 1990 - 409 Report
Exposure to total dust in 6 terminal grain elevators
Door: Kees de Schipper

f 16,50 1990 - 410 Verslag
Aanzet tot epidemiologisch onderzoek naar de gezondheidseffecten van nitraat
Door: Ria de Peuter en Very Vlaar

f 17,75 1990 - 411 Verslag
Informatiebronnen welke een rol spelen bij de DKTP- en BMR-vaccinatie bij ouders met kinderen van 14 maanden
Door: Marjan de Kock en Margit Moorrees

f 14,75 1990 - 412 Verslag
Arbeidshygiënisch onderzoek aan een nieuw proces in een ijzergieterij
Door: Bert Scheeper en Rob van Strien

f 21,00 1990 - 413 Verslag
Het overblijven op basisscholen. Een exploratief onderzoek
Door: Jolande Leeuwenburg en Ingeborg Poorterman

f 13,25 1990 - 414 Verslag
Een onderzoek naar de blootstelling aan zwavelwaterstof bij het mixen van mengmest. "Een vlaag zwavelwaterstof"
Door: Peter Tamsma

f 10,75 1990 - 415 Verslag
Toetsing van het protocol voor gericht werkplekonderzoek in de rubberverwerkende industrie
Vergelijking van de resultaten van twee verschillende methoden van werkplekonderzoek
Door: M.J.A. van Tongeren en H. Kromhout

¹ 1990 - 416 Verslag
De beoordeling van ergoloc
Door: Dorien van der Drift

f 8,75 1990 - 417 Verslag
Reproduceerbaarheid van de beroepencodering voor een beroepen blootstellingen matrix
Door: Michiel Fraters

² 1990 - 418 Verslag
Validering van de proefversie van de schriftelijke vragenlijst van een cohortonderzoek naar voeding en kanker
Door: Hanneke den Breeijen

f 13,00 1990 - 419 Verslag
Een onderzoek naar huidandoeningen in de rubberindustrie
Door: Kees de Schipper

f 13,00 1990 - 420 Verslag
A health survey of the population of Ponan, Ivory Coast. Children under five and the population associated with them
Door: Vincent Koch en Marleen de Kok

¹ Dit verslag is verkrijgbaar bij:
IMAG
Postbus 43
6700 AA WAGENINGEN

² Dit verslag is verkrijgbaar bij:
CIVO-TNO
Postbus 360
3700 AJ ZEIST
03404 - 44144

- f 22,00 1990 - 421 Verslag
Luchtwegirriterende stoffen in een aramidegaren spinnerij. Karakterisering van de blootstelling aan zwaveldioxide, zoutzuur en zwavelzuur
Door: E.A.M. Hartmans, M.R. Roest en E.I.M. Tjoe Ny
- f 20,00 1990 - 422 Verslag
Het meten van de zorgbehoefte van verpleeghuisbewoners. Resultaten van een proefonderzoek naar de subjectieve zorgbehoefte van bejaarde verpleeghuisbewoners in de gemeente Utrecht
Door: Alie Fokkens en Anja Straathof
- 1990 - 423 Verslag
Ouder worden, gezondheid en variabiliteit. Een verkennend onderzoek
Door: Joke Hoogenboom
- 1990 - 424 Verslag
Methodologische problemen bij studies naar de effecten van drinkwatervoorzieningen en sanitaire faciliteiten op de gezondheid. Stratificatie, als methode om versturende variabelen uit te schakelen
Door: Rianne Remmerde

* = nog niet verschenen

** = uitverkocht

Voor de inhoud van de publikaties blijven de auteurs aansprakelijk. Inlichtingen over deze serie kan men verkrijgen bij de secretaresse van de vakgroep Gezondheidsleer, Dreijenlaan 1, Postbus 238, 6700 AE WAGENINGEN, tel. 08370-82080.

C.J. Lako, secretaris