STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

AFDELING MATHEMATISCHE STATISTIEK

Rapport S 265 (C 13)

Leergang Besliskunde

Hoofdstuk XVI

Monte Carlo-methoden

door

A.H. Haitsma,

bewerkt door J. Oosterhoff

3e druk

juni 1964

Printed at the Mathematical Centre at Amsterdam,49,2nd Boerhaavestraat.
The Netherlands.

The Mathematical Centre, founded the 11th of February 1946, is a non - profit institution aiming at the promotion of pure mathematics and its applications, and is sponsored by the Netherlands Government through the Netherlands Organization for Pure Scientific Research (Z.W.O.) and the Central National Council for Applied Scientific Research in the Netherlands (T.N.O.), by the Municipality of Amsterdam and by several industries.

1. Inleiding

Wanneer men langs wiskundige weg vraagstukken uit de praktijk wenst op te lossen, is het noodzakelijk dat men eerst voor de situaties waarop deze vraagstukken betrekking hebben, een model formuleert dat toegankelijk is voor wiskundige analyse. Alle factoren die bij een vraagstuk een enigszins belangrijke rol spelen, moeten een plaats binnen het model vinden. Wanneer er echter veel factoren zijn waarmee men rekening moet houden, kan het model gemakkelijk te gecompliceerd worden voor een exacte mathematische behandeling. Vereenvoudiging van het model is dan dikwijls niet goed mogelijk, omdat daardoor de invloed van een of meer belangrijke factoren verwaarloosd moet worden.

Deze moeilijkheden doen zich o.a. voor bij allerlei technische problemen. Bij de vliegtuigbouw is het niet meer mogelijk langs wiskundige weg de grootheden te berekenen, die de verschillende eigenschappen van het vliegtuig bepalen. De wiskundige analyse wordt hier dan geheel of gedeeltelijk vervangen doordat men een fysisch model opstelt en met behulp van dit model de gewenste grootheden meet. Een model van het vliegtuig wordt hiervoor in een windtunnel geplaatst en de omstandigheden waarin het vliegtuig in werkelijkheid zal komen te verkeren, worden daarin zo goed mogelijk nagebootst.

Bovendien kan men, omdat men de invloed van verschillende factoren beheerst, processen bestuderen die in de praktijk pas na lange tijd zullen optreden. Men verkrijgt dus in het heden reeds informatie over in de toekomst plaats te vinden gebeurtenissen.

We zullen nu nagaan of een soortgelijke methode geschikt is voor het bestuderen van de processen die ons hier interesseren. Sommige van deze processen worden gekenmerkt door het optreden van gebeurtenissen, waarvan het tijdstip van optreden en het directe gevolg van het optreden van een gebeurtenis stochastisch zijn. In het mathematisch model corresponderen deze tijdstippen en gevolgen met waarden die stochastische variabelen aannemen. Bij een wachttijdprobleem kunnen dit de tijdstippen van binnenkomst van de klanten en de bedieningstijden van deze klanten zijn. Door bepaalde regels die het proces beheersen, zoals het aantal loketten, prioriteitsregels e.d., ontstaan nu nieuwe situaties zoals wachttijden var bepaalde lengte voor een loket, welke in het model corresponder

met functies van de stochastische variabelen en de parameters uit de regels die het proces bepalen. Gevraagd wordt nu meestal naar frequentieverdelingen van situaties van een bepaald type.

Is nu het model te ingewikkeld voor een wiskundige behandeling, dan kan men uitgaande van het model het gehele proces nabootsen. Wanneer men de beschikking heeft over processen waarvan de bijbehorende kansverdelingen der uitkomsten overeenstemmen met de verdelingen van de stochastische variabelen van het model, dan kan men deze uitkomsten beschouwen als mogelijke gebeurtenissen in het na te bootsen proces. In ons voorbeeld zou men zich kunnen voorstellen dat de tijdslengten tussen de aankomsttijdstippen en ook de bedieningstijden kunnen worden nagebootst door deze hulpprocessen. Met de uitkomsten van deze hulpprocessen, geïnterpreteerd dus als mogelijke gebeurtenissen van het proces, gaat men nu na met welke frequenties meer samengestelde gebeurtenissen plaatsvinden. Door dit procédé lang genoeg voort te zetten, verkrijgt men een beeld van de bijbehorende frequentieverdeling en op grond hiervan is het mogelijk schattingen te verkrijgen van grootheden die langs analytische weg niet te berekenen bleken. Wenst men bijv. het verschil in invloed tussen twee beleidsregels te kennen, dan kan men het proces nabootsen voor beide beleidsregels met dezelfde mogelijke gebeurtenissen, dus met dezelfde uitkomsten van de hulpprocessen. Het verschil in frequentie van de meer samengestelde gebeurtenissen behorende bij verschillende beleidsregels, vormt dan de grondslag waarop men de beleidsregels kan beoordelen.

Deze methode van nabootsing van het model heet Monte-Carlo-methode, welke benaming erop duidt, dat het eindresultaat evenals bij een gokspel stochastisch is, doordat het afhangt van een steekproef uit een of meer verdelingen.

Met deze methode kan men dus modellen onderzoeken die te ingewikkeld zijn voor wiskundige analyse. Een belangrijk nadeel van de Monte-Carlo-methode is echter de onmogelijkheid de invloed van de verschillende parameters die het model bepalen, te onderscheiden. Bij een wiskundige analyse verkrijgt men voor de grootheden die men wenst te berekenen, formules waarmee de invloed van de verschillende parameters op die grootheden is na te gaan. Bij een nabootsing zijn de uitkomsten echter numeriek, zodat men slechts een model kan onderzoeken waarin de parameters numeriek gespecificeerd zijn.

Een tweede nadeel van de Monte-Carlo methode is de onnauwkeurigheid van de berekeningsmethode. Bij een wiskundige analyse zijn de resultaten, afgezien van de onvolkomenheden van het model, exact. Bij een nabootsing van het model voeren we een steekproefexperiment op het model uit. De resultaten van dit steekproefexperiment zijn dus stochastisch en bezitten dus een variantie welke aanzienlijk kan zijn. Een van de moeilijkste problemen van de Monte-Carlo-methode is dan ook het nabootsingsproces zo uit te voeren, dat de variantie van de schattingen der verschillende grootheden binnen redelijke grenzen blijft.

Met een eenvoudig voorbeeld zullen we nu laten zien dat niet alle Monte-Carlo-schattingen van een bepaalde grootheid dezelfde variantie bezitten. Als $\underline{x}$ een homogene verdeling bezit op het interval (0,1), en men wenst de verwachting te berekenen van $g(\underline{x})$, waarbij we eenvoudigheidshalve $g(x)$ begrensd veronderstellen, dan geldt:

$$I = \mathcal{E}\, g(\underline{x}) = \int_0^1 g(x)dx \ . \qquad (1.1)$$

Immers de homogene verdeling op (0,1) wordt gedefinieerd als:

$$P(\underline{x} \leqq x) = \begin{cases} x & 0 \leqq x \leqq 1 \\ 0 & x < 0 \\ 1 & x \geqq 1 \ . \end{cases} \qquad (1.2)$$

Indien we niet in staat zijn om deze integraal rechtstreeks te bepalen, maar als we wel de beschikking hebben over een proces, waarvan de uitkomsten een homogene verdeling bezitten, dan kan de volgende Monte-Carlo-schatting van deze integraal worden gegeven. Stel $\{\underline{x}_i\}$ $(i=1,2,\ldots,N)$ zijn uitkomsten van dit hulpproces. Een schatting van (1.1) wordt dan gegeven door:

$$\underline{I}_1 = \frac{1}{N} \sum_{i=1}^{N} g(\underline{x}_i) \qquad (1.3)$$

en de variantie van $\underline{I}_1$ is dan:

$$\sigma^2_{\underline{I}_1} = \frac{1}{N} \left\{ \int_0^1 (g(x))^2 dx - (\mathcal{E}\, g(\underline{x}))^2 \right\}. \qquad (1.4)$$

Een tweede schatting van (1.1) verkrijgt men als verondersteld wordt, dat voor $0 \leqq x \leqq 1$ geldt $0 \leqq g(x) \leqq 1$. (Indien de functie $y=g(x)$

niet aan deze voorwaarde voldoet, kan men dit bereiken door een translatie zodanig dat $g(x) \geqq 0$ voor $0 \leqq x \leqq 1$, en een contractie zodanig dat $0 \leqq g(x) \leqq 1$ voor $0 \leqq x \leqq 1$.)

Zijn nu $\{x_i\}$ en $\{y_i\}$ $(i=1,\ldots,N)$ uitkomsten van het hulpproces, dan interpreteren we een paar (x_i,y_i) als een schot in het punt (x_i,y_i) van het vierkant met hoekpunten $(0,0)$, $(1,0)$, $(1,1)$ en $(0,1)$ in het cartesisch coördinatenstelsel. Een treffer wordt nu gedefinieerd als een schot onder de kromme $y=g(x)$. De kans op een treffer is dan (wegens de homogene verdeling van $\underline{x}$ en $\underline{y}$) gelijk aan de oppervlakte onder deze kromme binnen het eenheidsvierkant en deze kans wordt dus gegeven door (1.1). Het frequentiequotiënt van het aantal treffers onder deze N schoten is dan een schatting $\underline{I}_2$ van (1.1).

De variantie van $\underline{I}_2$ is dan de variantie van het frequentiequotiënt behorende bij een binomiale verdeling met $p=I$ en $n=N$, en dus gelijk aan

$$\sigma^2_{\underline{I}_2} = \frac{1}{N} I(1-I) = \frac{1}{N}\left[\int_0^1 g(x)dx - (\mathcal{E}\, g(\underline{x}))^2\right]. \qquad (1.5)$$

Vergelijken we nu (1.4) met (1.5), dan zien we

$$\sigma^2_{\underline{I}_2} \geqq \sigma^2_{\underline{I}_1} \qquad (1.6)$$

omdat

$$\int_0^1 g(x)dx \geqq \int_0^1 (g(x))^2 dx \qquad (1.7)$$

wegens

$$0 \leqq g(x) \leqq 1 \quad \text{voor} \quad 0 \leqq x \leqq 1.$$

Het gelijkteken in (1.7) en (1.6) geldt slechts dan als $g(x)$ alleen de waarde 0 of 1 aanneemt in het interval $0 \leqq x \leqq 1$.

De eerste methode leidt dus bij dezelfde steekproefomvang tot een kleinere variantie dan de tweede methode, of anders gezegd een bepaalde nauwkeurigheid kan met de eerste methode met een kleinere steekproefomvang bereikt worden.

Het verschil in variantie tussen (1.4) en (1.5) is gemakkelijk te verklaren. Bij de eerste methode wordt bij een x, $g(x)$ berekend, terwijl bij de tweede methode slechts gekeken wordt of $y < g(x)$, waarbij x en y uitkomsten zijn van het hulpproces. Doordat we dus bij de eerste schattingsmethode van meer informatie over de functie $g(x)$ gebruik maakten, verkregen we dus een kleinere variantie.

Vele problemen die met Monte-Carlo-methoden worden opgelost, zijn eigenlijk integraties, meestal meervoudige, waarbij men soms de integrand niet expliciet kent. In paragraaf 3 zullen enige algemede methoden besproken worden over variantiereductie bij schattingen van integralen.

In de eerste twee paragrafen zullen we methoden bespreken om reeksen getallen voort te brengen die beschouwd kunnen worden als een representatieve steekproef uit een gegeven verdeling. Dit probleem wordt in twee fasen opgelost.

1e. Men brengt een reeks getallen voort die als steekproef uit een homogene verdeling op het interval (0,1) beschouwd kan worden.

2e. De transformatie van een steekproef uit de homogene verdeling tot een steekproef uit een gegeven verdeling.

In de volgende paragraaf zullen we nu het voortbrengen van een steekproef uit de homogene verdeling bespreken.

2. De voortbrenging van aselecte getallen

Wanneer een proces uitkomsten oplevert, welke onderling onafhankelijk worden verkregen en die beschouwd kunnen worden als trekkingen uit een homogene verdeling op het interval (0,1), dan noemt men die uitkomsten aselecte getallen. De homogene verdeling is reeds gedefinieerd in (1.2). Uit (1.2) volgt nu voor een homogeen verdeelde stochastische variabele $\underline{x}$

$$P\left[x \leqq \underline{x} < x + \Delta x\right] = \Delta x \text{ voor } 0 \leqq x < x + \Delta x \leqq 1 \qquad (2.1)$$

waarin Δx de lengte van een deelinterval van (0,1) is. Schrijft men de aselecte getallen in een decimale vorm, dan is voor elke decimaal van een aselect getal de kans, dat deze een van de symbolen $0,1,2,\ldots,9$ zal zijn, gelijk $\frac{1}{10}$. Immers stel dat bijv. de eerste drie decimalen gegeven worden door 4 5 8, dan is de kans dat het vierde cijfer een 3 is wegens (2.1) gelijk aan:

$$P\left[0{,}4583 \leqq \underline{x} < 0{,}4584 \,\middle|\, 0{,}458 \leqq \underline{x} < 0{,}459\right] = \frac{10^{-4}}{10^{-3}} = 10^{-1} \; . \qquad (2.2)$$

Aangezien $\underline{x}$ een continu verdeelde variabele is, bestaat een aselect getal volgens deze definitie uit een oneindige rij cijfers, die alle een van de waarden $0,1,\ldots,9$ aannemen met een kans 10^{-1}. Men kan een aselect getal dus beschouwen als een rij aselecte

cijfers (random digits). Het voortbrengen van aselecte getallen is dus aequivalent met het voortbrengen van aselecte cijfers. In de praktijk is men uiteraard genoodzaakt een rij aselecte cijfers tot een eindig aantal n te beperken. De keuze van n wordt dan bepaald door de nauwkeurigheid, die voor het uit te voeren steekproefexperiment vereist is. De aldus verkregen aselecte getallen van n decimalen zijn dan dus trekkingen uit een discrete in plaats van uit een continue verdeling op het interval (0,1) (want er zijn slechts 10^n getallen tussen 0 en 1 met n decimalen). Voortaan zullen we een getal van n cijfers ook een aselect getal noemen, wanneer we bedoelen dat dat getal gedeeld door 10^n een aselect getal in het interval (0,1) is.

Wanneer men zou beschikken over een lotingsmechanisme, dat tien verschillende uitkomsten met gelijke kans kan opleveren, terwijl een bepaalde uitkomst onafhankelijk van de voordien waargenomen uitkomsten tot stand komt, dan zou men met dit lotingsmechanisme willekeurig veel aselecte cijfers kunnen voortbrengen. A priori kan men van geen enkel lotingsmechanisme echter onderstellen dat aan deze eis is voldaan. Alleen achteraf kan men door de uitkomsten statistisch te toetsen, nagaan of bepaalde vormen van afhankelijkheid niet voorkomen, of althans geen merkbare invloed hebben gehad. Tabellen met aselecte cijfers, verkregen met dergelijke lotingsmechanismen, zijn gepubliceerd door M.G. KENDALL en B. BABINGTON SMITH (1939) en door de RAND CORPORATION (1955). Deze aselecte cijfers zijn uitvoerig getoetst. Ook de tienzijdige dobbelsteen van H.C. Hamaker (1948) is een lotingsmechanisme, dat redelijke resultaten oplevert.

Deze tabellen zijn zeer goed bruikbaar voor Monte-Carlo-berekeningen met pen en papier, maar voor grote steekproefexperimenten, die met behulp van een automatische rekenmachine worden verricht, zijn ze minder geschikt, omdat het programmeren van een gehele tabel aselecte cijfers relatief veel tijd vergt. Bovendien wordt hierdoor beslag gelegd op een aanzienlijk deel van de geheugencapaciteit.

Bij berekeningen met een automatische rekenmachine is het daarom gewenst, dat de machine zelf de voor het steekproefexperiment benodigde aselecte getallen berekent. Aangezien een rekenautomaat alleen vaste voorschriften kan opvolgen, zal het proces dat aselecte getallen voortbrengt deterministisch zijn. De voortgebrachte getallen zullen dus in feite niet aselect zijn, maar een zekere regelmaat vertonen, afhankelijk van het toegepaste proces. In plaats van aselecte getallen spreekt men daarom van pseudo-aselecte getallen (pseudo random numbers). Men heeft nu wel het voordeel, dat men ter controle van uitgevoerde berekeningen dezelfde pseudo-aselecte getallen opnieuw kan voortbrengen. Er blijken verschillende methoden te bestaan om rijen pseudo-aselecte getallen te berekenen, welke naar een aantal zinvolle criteria nauwelijks van representatieve steekproeven uit de homogene verdeling onderscheiden kunnen worden. Voor de meeste toepassingen zijn deze pseudo-aselecte getallen dan ook zeer goed bruikbaar.

Verschillende van deze methoden worden besproken door K.D. TOCHER (1954), O. TAUSSKY en J. TODD (1956).

Een van de gebruikelijkste methoden om pseudo-aselecte getallen te maken, is de multiplicatieve congruentiemethode, welke door D.H. LEHMER (1951) is geïntroduceerd. We laten hieraan de betekenis van enige termen voorafgaan. Twee positieve gehele getallen heten relatief priem (onderling ondeelbaar), als hun grootste gemene deler gelijk 1 is, bijv. 25 en 24. Als verder c, d en m > 0 gehele getallen zijn, dan betekent d=c (modulo m) of d=c (mod m): c-d is deelbaar door m; bijv. 23=38 (mod 5). Als behalve d=c (mod m) ook geldt $0 \leq d < m$, dan is d de rest bij deling van c door m.

De methode van LEHMER werkt nu als volgt. Wanneer a, b en m gehele positieve getallen zijn, met a < m en b < m en wel zo dat a.b en m relatief priem zijn, en als bovendien de rij $\{u_i\}$ (i=1,2,...) gedefinieerd wordt door:

$$u_0 = b$$
$$u_{n+1} = au_n \pmod m,\ 0 < u_{n+1} < m,\ (n=0,1,2,\ldots) \quad (2.3)$$

dan kan de rij $\{m^{-1} u_i\}$ als a en m niet te klein zijn in vele gevallen als een rij aselecte getallen beschouwd worden. De in (2.3) gegeven relatie tussen u_{n+1} en u_n stelt vast, dat u_{n+1} de rest is die ontstaat bij deling van au_n door m. Om deze deling snel te kunnen uitvoeren, wordt bij een machine, werkend in het tientallig stelsel, voor m meestal 10^s-1, 10^s of 10^s+1 gekozen, en voor een machine met tweetallig getallenstelsel 2^s-1, 2^s of 2^s+1.

Voorbeeld 1.

Nemen we b=5, a=27, m=32, dan verkrijgen we de volgende rij:

$u_0=5$ $u_1=7$ $u_2=29$ $u_3=15$ $u_4=21$ $u_5=23$
$u_6=13$ $u_7=31$ $u_8=u_0=5$.

We zien dus door rechtstreekse berekening dat $u_8=u_0$. Vanaf u_8 zal de rij zich dus herhalen en de gehele rij bevat dus slechts 8 verschillende elementen. Het kleinste getal $\delta > 0$, waarvoor bij een willekeurige $n \geqq 0$ geldt:

$$u_{n+\delta} = u_n \qquad (2.4)$$

noemen we de periode van de rij $\{u_i\}$. In ons voorbeeld is dus $\delta = 8$. Omdat bij deling van een geheel getal door m de rest slechts 0,1,... of (m-1) kan zijn, heeft de rij $\{u_i\}$ steeds hoogstens (m-1) verschillende elementen en dus altijd een periode $\delta \leqq (m-1)$. Het is duidelijk dat de periode noodzakelijk groot moet zijn, als men de rij $\{m^{-1} u_i\}$ wil gebruiken als aselecte getallen.

Over de periode van een rij pseudo-aselecte getallen volgens LEHMER zullen we in de Appendix enige stellingen geven. We vermelden hier slechts, dat bij de meest gebruikelijke keuze $m=2^s$ de periode van de rij $\{u_i\}$ nooit groter kan zijn dan 2^{s-2}.

Sedert 1960 wordt ook een iets gewijzigde methode om pseudo-aselecte getallen voort te brengen veel toegepast, namelijk de gemengde congruentie methode.

De voortgebrachte rij wordt hierbij gedefinieerd door

$$v_0 = \rho$$
$$v_{n+1} = \lambda v_n + \mu \pmod m,\ 0 \leq v_{n+1} < m\ (n=0,1,2\ldots), \quad (2.5)$$

waarin λ, μ en ρ gehele positieve getallen zijn, alle kleiner dan m, en $\lambda . \mu$ en m relatief priem zijn. Bij geschikte keuze van λ en μ kan de rij $\{ m^{-1} v_i \}$ wederom dienst doen als een rij aselecte getallen. De rij heeft de maximale periode m, indien $m=2^s$, μ oneven en $\lambda = 1 \pmod 4$, onafhankelijk van de startwaarde ρ. Voor een bewijs hiervan zij verwezen naar de Appendix. Voor $\mu=0$ gaat de methode over in de multiplicatieve congruentie methode van LEHMER.

De rij $\{ v_i \}$ heeft dus bij geschikte keuze van de parameters λ en μ een langere periode dan de rij $\{ u_i \}$, als men in beide gevallen dezelfde $m=2^s$ toepast. Bovendien leidt dan elke startwaarde ρ tot een rij $\{ v_i \}$ van maximale periode. Het voornaamste voordeel van de rij $\{ v_i \}$ is echter, dat men de seriële correlatie in de hand heeft. De seriële correlatie van een rij $\{ v_i \}$ met maximale periode m wordt gedefinieerd door

$$\rho(\underline{v}_n, \underline{v}_{n+1}) = \frac{\operatorname{cov}(\underline{v}_n, \underline{v}_{n+1})}{\operatorname{var} \underline{v}_n}, \quad (2.6)$$

waarin $\underline{v}_n$ een aselecte trekking uit de rij $\{ v_i \mid i=0,1,\ldots m \}$ voorstelt. De seriële correlatie is dus een maat voor de afhankelijkheid van opeenvolgend voortgebrachte getallen. Door R.R. COVEYOU (1960) en M. GREENBERGER (1961) is bewezen, dat bij verwaarlozing van termen van de orde m^{-1} of kleiner

$$\rho(\underline{v}_n, \underline{v}_{n+1}) \approx \frac{1}{\lambda} - \frac{6\mu}{\lambda m}\left(1 - \frac{\mu}{m}\right) + T_1, \text{ als } \mu \geq \lambda \quad (2.7)$$

en

$$\rho(\underline{v}_n, \underline{v}_{n+1}) \approx \frac{1}{\lambda} + T_2 \quad \text{als } \mu < \lambda, \quad (2.8)$$

waarin de correctietermen T_i voldoen aan

$$-\frac{\lambda}{m} \leq T_i \leq \frac{\lambda}{m} \qquad (i=1,2). \tag{2.9}$$

Een grote waarde van λ leidt dus tot een kleine seriële correlatie. Anderzijds dient men er echter zorg voor te dragen dat λ van kleiner grootte orde blijft dan m, zodat de termen T_i klein blijven (zie (2.9)). Het is ook mogelijk de seriële correlatie te beinvloeden door geschikte keuze van μ, maar dit is iets lastiger.

Herhaald toepassen van (2.5) geeft

$$v_{n+k} = \lambda^k v_n + \frac{\lambda^k - 1}{\lambda - 1}\mu \pmod m. \tag{2.10}$$

Deze formule is van dezelfde gedaante als (2.5), doch met iets andere parameters. Als de rij (2.5) maximale periode m heeft, dan heeft de rij $\{v_{n+ki} \mid i=0,1,2,\ldots\}$ eveneens periode m, zodat men de seriële correlatie met verschuiving k, $\rho(\underline{v}_n, \underline{v}_{n+k})$, eenvoudig kan berekenen door in (2.7) of (2.8) in plaats van λ en μ de parameters λ^k en $\frac{\lambda^k - 1}{\lambda - 1}\mu$ te substitueren.

De betekenis van een kleine seriële correlatie, hoe belangrijk ook, dient men niet te overschatten. De rijen $\{u_i\}$ en $\{v_i\}$ hebben ook vele andere eigenschappen die men bij practische problemen graag zou willen beinvloeden door de keuze van geschikte parameter waarden. In het algemeen is hier echter nog weinig over bekend.

Een eenvoudige toepassing van het gebruik van aselecte getallen geeft

Voorbeeld 2

Gevraagd wordt een schatting van

$$I = \int_0^1 12x^3(1-x)dx \ . \tag{2.11}$$

Aan de lezer wordt overgelaten, bijv. uit de "Tables of random sampling numbers" van M.G. KENDALL en B. BABINGTON SMITH 10 aselecte getallen $x_1,\ldots,x_{10}$ te kiezen (2 decimalen zijn al voldoende) en als schatting voor (2.5) te berekenen

$$\underline{I}_1 = \frac{1}{10} \sum_{i=1}^{10} 12x_i^3 (1-x_i). \qquad (2.12)$$

Men vergelijke deze uitkomst met de werkelijke waarde van I, die $\frac{3}{5}$ bedraagt, zoals analytisch door integratie in (2.5) gemakkelijk is te vinden.

De variantie van de schatting (2.6) bedraagt

$$\sigma^2(\underline{I}_1) = \frac{1}{10}\left[\int_0^1 \left\{12x^3(1-x)\right\}^2 dx - (\tfrac{3}{5})^2\right] \approx 0{,}02. \qquad (2.13)$$

Om een bepaalde keuze van a in (2.3) te rechtvaardigen, moet men nagaan of een rij pseudo aselecte getallen beschouwd kan worden als een representatieve steekproef uit de homogene verdeling. Om dit na te gaan brengt men een aantal rijen pseudo aselecte getallen voort, door voor elke rij een andere beginwaarde u_0=b te kiezen. Elke rij wordt dan in een aantal niet te grote groepen gesplitst (bijv. 1000 getallen per groep), en deze groepen worden aan verschillende statistische toetsen onderworpen. Men kan de pseudo aselecte getallen nu op twee verschillende wijzen toetsen:

1. Men kan de pseudo aselecte getallen beschouwen als een rij aselecte cijfers en deze aselecte cijfers aan de toetsen van M.G. KENDALL en B. BABINGTON SMITH (1938) onderwerpen. Dit zijn de volgende toetsen:

   a. De frequentietoets. Met deze toets worden de frequenties van alle cijfers 0 t/m 9 vergeleken met de verwachte waarden.

   b. De kettingtoets (serial test). Men vergelijkt met deze toets de frequenties van de paren 00, 01 t/m 99 met de verwachte waarden. Als de waarden die opeenvolgende cijfers aannemen afhankelijk zijn, kan men dat met de kettingtoets nagaan. Deze toets kan uitgebreid worden van tweetallen tot meertallen opeenvolgende cijfers. In het bijzonder is de frequentietoets een speciale kettingtoets. Een verbetering van de kettingtoets is door

I.J. GOOD (1953) gegeven.

c. De gatentoets (gap test). Hiermee wordt de verdeling van het aantal getallen tussen opvolgende gelijke cijfers, bijv. nullen, vergeleken met de theoretische verdeling.

2. Men kan de aselecte getallen in hun geheel beschouwen als trekkingen uit de homogene verdeling. Binnen elke groep kan men dan de frequentie van een bepaald patroon vergelijken met de bijbehorende verwachting of de theoretische verdeling Dergelijke patronen zijn bijv. het aantal runs (opvolgende getallen) onder en boven het gemiddelde, het aantal runs van oplopende getallen enz. Dit soort toetsen is in het bijzonder belangrijk voor Monte-Carlo-berekeningen omdat daar meestal verscheidene aselecte getallen worden gebruikt voor een steekproefpunt. Worden dan pseudo aselecte getallen gebruikt die regelmatigheden van dit type vertonen, dan kan dit tot onjuiste Monte-Carlo-schattingen leiden. Een overzicht van de hier aangestipte toetsen vindt men bij O. TAUSSKY en J. TODD (1956). Deze auteurs vermelden ook verschillende numerieke resultaten van de door hen uitgevoerde berekeningen.

3. Aselecte trekkingen uit een gegeven verdeling

Wanneer aan de verzameling van mogelijke uitkomsten van een experiment een verdelingsfunctie $F(x)$ kan worden toegevoegd, dan noemen we een waargenomen uitkomst x een aselecte trekking uit die verdeling. Voor de uitkomst x van het experiment geldt dus:

$$P\left[\underline{x} \leqq x\right] = F(x). \qquad (3.1)$$

Omdat $F(x)$ een monotoon niet dalende functie van x is, geldt:

$$P\left[F(\underline{x}) \leq F(x)\right] = F(x) \quad ^{1)} \tag{3.2}$$

Uit (3.2) volgt dat $\underline{y} = F(\underline{x})$ op het interval (0,1) homogeen verdeeld is. Bij een aselecte trekking y_1 uit de homogene verdeling (aselect getal) vindt men dus een aselecte trekking x uit de gegeven verdeling door de vergelijking

$$y_1 = F(x_1) \tag{3.3}$$

naar x_1 op te lossen. Deze oplossing kan volgens fig. 1 grafisch verkregen worden; x_1 blijkt slechts dan eenduidig bepaald te zijn, als F(x) monotoon stijgend is in x_1.

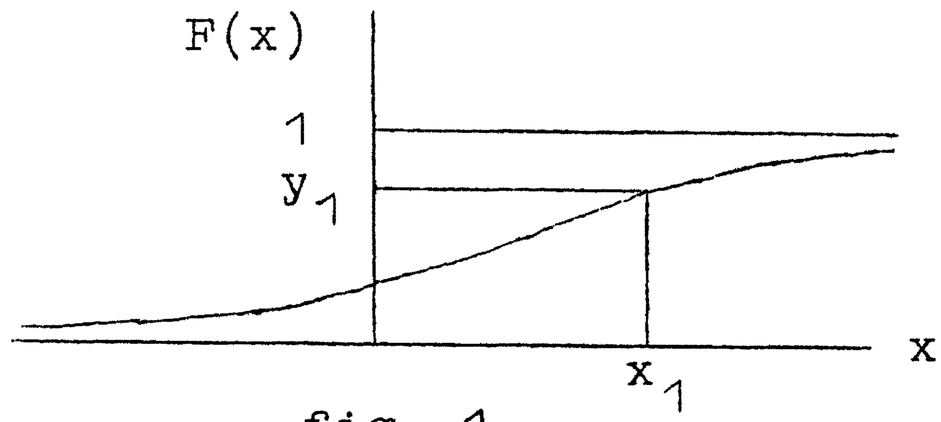

fig. 1

Bij een aselect getal y zoekt men in de grafiek de bijbehorende waarde x op. Wanneer F(x) een experimenteel bepaalde verdelingsfunctie is, komt de grafische oplossing van (3.3) in het bijzonder in aanmerking. F(x) is dan een trapfunctie en als aselecte trekkingen treden dan alleen de sprongpunten van deze

---

1) De afleiding van (3.2) is aldus: $P\left[F(\underline{x}) \leq F(x)\right] = P\left[\underline{x} \leq x\right] + P\left[\underline{x} > x \text{ en } F(\underline{x}) = F(x)\right]$. De eerste term in het rechterlid is F(x), de tweede term is 0, want als $g = \sup\left[\zeta \mid F(\zeta) = F(x)\right]$, dan is $P\left[\underline{x} > x \text{ en } F(x) = F(x)\right] = \lim_{\zeta \uparrow g} P\left[x < \underline{x} \leq \zeta\right] = \lim_{\zeta \uparrow g} \left[F(\zeta) - F(x)\right] = 0$.

trapfunctie op, dus alleen de waargenomen waarden.

Bestaat de collectie elementen waaruit men trekkingen doet, uit getallen van n cijfers, dan kan een aselect getal y 10^n verschillende waarden aannemen. Ditzelfde geldt dan ook voor de door y = F(x) aan y toegevoegde variabele x. Dus ook voor een continue verdelingsfunctie F(x) is x dan een aselecte trekking uit een discrete verdeling met verdelingsfunctie G(x), welke de vorm heeft van een trapfunctie en in de punten

$$x = F^{-1}(k.10^{-n}) \quad (k = 0,1,2,\ldots,10^n) \tag{3.4}$$

de waarde $G(x) = F(x) = k.10^{-n}$ aanneemt en in deze punten een sprong maakt die gelijk is aan

$$\Delta G(x) = \Delta F(x) = 10^{-n} \tag{3.5}$$

(Met $x = F^{-1}(k.10^{-n})$ wordt bedoeld: x zodanig dat $F(x) = k.10^{-n}$).

Alleen die waarden $x=x_k$ waarbij G(x) een sprong maakt, kunnen als aselecte trekkingen voorkomen. Voor de intervallen (x_k,x_{k+1}) waarin geen aselecte trekkingen voorkomen, geldt: $10^{-n} = F(x_{k+1}) - F(x_k) = \int_{x_k}^{x_{k+1}} f(x)dx$; is de verdelingsdichtheid f(x) groot in (x_k,x_{k+1}), dan is het interval dus klein. In dat geval is het verschil tussen aselecte trekkingen uit F(x) en G(x) te verwaarlozen. Is daarentegen f(x) klein, dus heeft de verdelingsfunctie F(x) een vlak verloop, dan is de lengte van (x_k,x_{k+1}) groot. Wordt een beginpunt x_k van zo'n lang interval getrokken, dan verdient het aanbeveling niet dit beginpunt x_k, maar een punt binnen dat interval als aselecte trekking te aanvaarden. Men trekt in dat geval een nieuw aselect getal y en bepaalt hiermee als aselecte trekking $x_k+y\ \Delta x$, waarin Δx de lengte van het betreffende interval (x_k,x_{k+1}) is.

Rechtstreekse oplossing van (3.3) komt neer op het bepalen van de inverse van de verdelingsfunctie F(x).

Voor de meeste verdelingen leidt dit tot lastige berekeningen. Men maakt daarom vaak gebruik van speciale eigenschappen van een verdeling voor het verkrijgen van aselecte trekkingen uit die verdeling. Het voordeel van directe inversie van de verdelingsfunctie F(x) is echter, dat vele eigenschappen van de rij trekkingen uit de homogene verdeling $y_1, y_2, \ldots$, die men gebruikt om aselecte trekkingen uit de verdeling F(x) te verkrijgen, behouden blijven bij inversie, daar dit een monotone transformatie is. Beschikt men over een rij pseudo-aselecte getallen, die uitvoerig getoetst is, dan zullen de meeste toetsingsresultaten dus ook van toepassing zijn op de pseudo-aselecte rij trekkingen uit de verdeling F(x) die ontstaat door het proces van directe inversie. Het is dan gewoonlijk niet meer nodig de pseudo-aselecte trekkingen uit F(x) op "aselectheid" te toetsen. Past men echter andere methoden toe, dan is het geenszins zeker dat de ontstane trekkingen uit F(x) voldoen aan redelijke eisen van aselectheid, ook al voldoet de rij pseudo-aselecte getallen waar men vanuit gaat daar wel aan.

Voor de belangrijkste verdelingen volgt nu een overzicht van de methoden om aselecte trekkingen te verkrijgen.

1. <u>Binomiale verdeling met parameters p en N</u>

$$F(n) = \sum_{j=0}^{n} \binom{N}{j} p^j (1-p)^{N-j} \qquad (3.6)$$

De kans dat men uit de homogene verdeling op (0,1) een aselect getal $\leqq p$ trekt, is p. (3.6) geeft dus de kans dat bij trekken van N aselecte getallen uit (0,1), er n kleiner zijn dan p. Het aantal aselecte getallen $\leqq p$ onder N aselecte getallen is dus een aselecte trekking uit (3.6). Een tweede methode wordt beschreven in K.D. TOCHER (1954).

In dit geval verdient directe inversie van de verdelingsfunctie F(x) echter de voorkeur, daar dit zonder veel moeite geschieden kan.

2. <u>Exponentiële verdeling</u>

$$F(x) = 1-e^{-\lambda x} \qquad \lambda > 0. \tag{3.7}$$

a) Rechtstreekse methode. Aangezien met $\underline{y}$ ook $1-\underline{y}$ homogeen verdeeld is op (0,1), kan men in plaats van (3.3) ook

$$y = 1-F(x) \tag{3.8}$$

naar x oplossen. Substitutie van (3.7) in (3.8) geeft:

$$y = e^{-\lambda x}\ , \tag{3.9}$$

dus:

$$x = \frac{1}{\lambda} \mid \ln y \mid . \tag{3.10}$$

Aangezien het berekenen van een logarithme met een electronische rekenmachine relatief tijdrovend is, verdient deze methode niet altijd aanbeveling. We behandelen daarom onder b) een andere methode, die bij een eerste lezing zonder bezwaar kan worden overgeslagen. Het is echter geenszins zeker dat de onder b) uiteengezette methode sneller is dan de rechtstreekse methode, zoals in de literatuur soms wordt aangegeven. Bovendien is de methode b) geen monotone transformatie. Deze methode wordt tegenwoordig dan ook zelden meer toegepast, al is ze ongetwijfeld ingenieus.

b) Methode van J. VON NEUMANN (1951).
Achtereenvolgens trekken we uit (0,1) de aselecte getallen:

$$x_0, y_{01}, \ldots, y_{0i}, \ldots$$

en vormen de sommen:

$$1-x_0+y_{01},\ 1-x_0+y_{01}+y_{02}, \ldots, 1-x_0+y_{01}+\ldots+y_{0i}, \ldots \quad .$$

Zodra een van deze sommen groter wordt dan 1, bijv.:

$$1-x_0+y_{01}+\ldots+y_{0,i_0} \geqq 1 \tag{3.11a}$$

en

$$1-x_0+y_{01}+\ldots+y_{0,i_0-1} < 1 \qquad (3.11b)$$

beëindigen we de reeks sommaties. Wanneer i_0 oneven is, aanvaarden we x_0 als aselecte trekking uit de exponentiële verdeling met parameter $\lambda=1$ (de reden hiervoor zal straks duidelijk worden). Wanneer i_0 even is, herhalen we het procédé. We trekken dan uit (0,1) de aselecte getallen

$$x_1, y_{11}, \ldots, y_{1i}, \ldots$$

en vormen weer de sommen:

$$i-x_1+y_{11},\ 1-x_1+y_{12}, \ldots\ldots$$

Bij de eerste som die groter wordt dan 1 houden we weer op, bijv. als

$$1-x_1+y_{11}+\ldots+y_{1,i_1} \geqq 1 \qquad (3.12a)$$

en

$$1-x_1+y_{11}+\ldots+y_{1,i_1-1} < 1\ . \qquad (3.12b)$$

Wanneer i_1 oneven is, aanvaarden we nu $1+x_1$ als aselecte trekking uit de exponentiële verdeling met $\lambda=1$ (men zie het vervolg). Als i_1 even is, gaan we op dezelfde wijze voort totdat we, na t maal een even i_j ($j=0,1,\ldots,t-1$) gevonden te hebben, de $(t+1)^e$ maal een oneven i_t vinden. Dus:

$$1-x_t+y_{t1}+\ldots+y_{t,i_t} \geqq 1 \qquad (3.13a)$$

en

$$1-x_t+y_{t1}+\ldots+y_{t,i_t-1} < 1 \qquad (3.13b)$$

en

i_t oneven.

De bewering is dan dat $t+x_t$ een aselecte trekking is uit de exponentiële verdeling met parameter $\lambda=1$.

De trekkingsresultaten $x_0,\ldots,x_t,y_{01},\ldots,y_{t,i_t}$ in het bovenstaande schema kunnen we opvatten als waarden, aangenomen door de stochastische variabelen $\underline{x}_0,\ldots,\underline{x}_t,\underline{y}_{01},\ldots,\underline{y}_{t,i_t}$, die alle de homogene verdeling op het interval (0,1) be-

zitten. Door (3.11), (3.12) en (3.13) worden dan de stochastische variabelen $\underline{i}_0,\ldots,\underline{i}_t$ gedefiniëerd, waarbij t weer een bepaalde waarde is van de stochastische grootheid $\underline{t}$. Volgens het voorgaande wordt $\underline{t}$ gedefiniëerd door

$$\underline{t}=t \text{ als } \underline{i}_0,\underline{i}_1,\ldots,\underline{i}_{t-1} \text{ alle even, } \underline{i}_t \text{ oneven.} \qquad (3.14)$$

We definiëren voorts de stochastische grootheid $\underline{s}$:

$$\underline{s}=x_t \text{ als } \underline{x}_t=x_t, \text{ onder de voorwaarde dat } \underline{i}_t \text{ oneven is} \qquad (3.15)$$

Dus $P\left[\underline{s}=x_t\right] = P\left[\underline{x}_t=x_t \middle| \underline{i}_t \text{ oneven}\right]$. Het is duidelijk (zie (3.13)) dat de stochastische variabelen $\underline{x}_t$ en $\underline{i}_t$ onderling afhankelijk zijn, dus ook $\underline{s}$ en $\underline{i}_t$.

$\underline{s}$ is echter onafhankelijk van $\underline{t}$; weliswaar zetten we de reeks sommaties voort op grond van de resultaten van voorafgaande sommaties, maar de resultaten bij een nieuwe sommatie zijn onafhankelijk van de voorgaande sommaties. Om dezelfde reden zijn de stochastische variabelen $\underline{i}_0,\underline{i}_1,\underline{i}_2,\ldots,\underline{i}_t$ onderling onafhankelijk.

We zullen nu bewijzen dat $\underline{x}=\underline{t}+\underline{s}$ de verdelingsfunctie

$$F(x) = 1-e^{-x} \qquad (3.16)$$

bezit.

Bewijs:

$[x]$ is per definitie het grootste gehele getal $\leqq x$. Omdat $\underline{t}$ slechts gehele waarden $0,1,2,\ldots$ aanneemt, geldt:

$$P\left[\underline{x} \leqq x\right] = P\left[\underline{t}+\underline{s} \leqq x\right] = \sum_{k=0}^{[x]} P\left[\underline{t}=k, \underline{s} \leqq x-k\right] = \sum_{k=0}^{[x]} P\left[\underline{t}=k\right].P\left[\underline{s} \leqq x-k\right] =$$

$$= \sum_{k=0}^{[x]-1} P\left[\underline{t}=k\right] + P\left[\underline{t}=[x]\right] \cdot P\left[\underline{s} \leqq x-[x]\right] \text{ (want } P\left[\underline{s} \leqq 1\right]=1). \qquad (3.17)$$

Met behulp van (3.14) en (3.15) volgt hieruit:

$$P\left[\underline{x} \leqq x\right] = \sum_{k=0}^{[x]-1} P\left[\underline{i}_0,\underline{i}_1,\ldots,\underline{i}_{k-1} \text{ even, } \underline{i}_k \text{ oneven}\right] +$$

$$+ P\left[\underline{i}_0,\underline{i}_1,\ldots,\underline{i}_{[x]-1} \text{ even, } \underline{i}_{[x]} \text{ oneven}\right]. P\left[x_{[x]} \leqq x-[x] \middle| \underline{i}_{[x]} \text{ oneven}\right]. \qquad (3.18)$$

Aangezien $\underline{i}_0,\underline{i}_1,\ldots$ onderling onafhankelijk zijn en de-

zelfde verdelingsfunctie bezitten en verder de verdeling van $\underline{x}_{[x]}$ onafhankelijk is van de index $[x]$, kunnen we voor (3.18) schrijven:

$$P\left[\underline{x} \leqq x\right] = P\left[\underline{i} \text{ oneven}\right] . \sum_{k=0}^{[x]-1} \left(P\left[\underline{i} \text{ even}\right]\right)^k +$$

$$+ \left(P\left[\underline{i} \text{ even}\right]\right)^{[x]} . P\left[\underline{i} \text{ oneven}\right] . P\left[\underline{u} \leqq x-[x] \,\middle|\, \underline{i} \text{ oneven}\right] \qquad (3.19)$$

waarin $\underline{i}$ dezelfde verdeling heeft als $\underline{i}_0$ en $\underline{u}$ homogeen verdeeld is op $(0,1)$. Het verband tussen $\underline{u}$ en $\underline{i}$ wordt dan gegeven door:

$$1-\underline{u}+\underline{y}_1+\underline{y}_2+\ldots+\underline{y}_i \geqq 1 \qquad (3.20)$$

$$1-\underline{u}+\underline{y}_1+\underline{y}_2+\ldots+\underline{y}_{i-1} < 1. \qquad (3.21)$$

We definiëren nu p_i door:

$$p_i = P\left[1-\underline{u}+\underline{y}_1+\ldots+\underline{y}_i \geqq 1 \,\middle|\, \underline{u}=u\right] = P\left[\underline{y}_1+\ldots+\underline{y}_i \geqq u\right] =$$

$$= 1-P\left[\underline{y}_1+\ldots+\underline{y}_i \leqq u\right] . \qquad (3.22)$$

Door volledige inductie bewijzen we nu:

$$p_i = 1 - \frac{u^i}{i!} . \qquad (3.23)$$

Voor i=1 is (3.23) triviaal; stel nu dat (3.23) bewezen is voor i-1. Uit (3.22) volgt

$$p_i = 1 - \int_0^u P\left[\underline{y}_1+\ldots+\underline{y}_{i-1} \leqq u-y\right] dP\left[\underline{y}_i \leqq y\right] =$$

$$= 1 - \int_0^u \frac{(u-y)^{i-1}}{(i-1)!}\, dy = 1 - \frac{u^i}{i!} \qquad (3.24)$$

waarmee (3.23) bewezen is.

q_i wordt nu gedefinieerd door:

$$q_i = P\left[1-\underline{u}+\underline{y}_1+\ldots+\underline{y}_i \geqq 1,\ 1-\underline{u}+\underline{y}_1+\ldots+\underline{y}_{i-1} < 1 \,\middle|\, \underline{u}=u\right] \qquad (3.25)$$

$(\underline{y}_0=0)$

Hieruit volgt:

$$q_i = P\left[\underline{y}_1+\ldots+\underline{y}_i \geqq u,\ \underline{y}_i+\ldots+\underline{y}_{i-1} < u\right] = P\left[\underline{y}_1+\ldots+\underline{y}_i \geqq u\right] -$$

$$- P\left[\underline{y}_1+\ldots+\underline{y}_i \geqq u,\ \underline{y}_1+\ldots+\underline{y}_{i-1} \geqq u\right] = p_i - p_{i-1} \text{ wegens}$$

(3.22). Dus:

$$q_i = p_i - p_{i-1} \cdot \tag{3.26}$$

Volgens (3.24) en (3.26) is dus

$$q_i = \frac{u^{i-1}}{(i-1)!} - \frac{u^i}{i!} \tag{3.27}$$

$$\text{en } P\left[\underline{i} \text{ oneven} \middle| \underline{u}=u\right] = \sum_{j=0}^{\infty} P\left[\underline{i}=2j+1 \middle| \underline{u}=u\right] = \sum_{j=0}^{\infty} q_{2j+1} = e^{-u} \tag{3.28}$$

Omdat $\underline{u}$ homogeen verdeeld is op (0,1), volgt uit (3.28):

$$P\left[\underline{i} \text{ oneven}, \underline{u} \leqq u\right] = \int_0^u P\left[\underline{i} \text{ oneven} \middle| \underline{u}=v\right] dP\left[\underline{u} \leqq v\right] =$$

$$= \int_0^u e^{-v} dv = 1-e^{-u} \tag{3.29}$$

dus

$$P\left[\underline{i} \text{ oneven}\right] = \int_0^1 e^{-v} dv = 1-e^{-1} \tag{3.30}$$

$$\text{en} \quad P\left[\underline{i} \text{ even}\right] = e^{-1} \tag{3.31}$$

en uit (3.29) en (3.30):

$$P\left[u \leqq u \middle| \underline{i} \text{ oneven}\right] = \frac{P\left[\underline{u} \leqq u, \underline{i} \text{ oneven}\right]}{P\left[\underline{i} \text{ oneven}\right]} = \frac{1-e^{-u}}{1-e^{-1}} \cdot \tag{3.32}$$

Uit (3.19), (3.30), (3.31) en (3.32) volgt tenslotte:

$$P\left[\underline{x} \leqq x\right] = (1-e^{-1}) \sum_{k=0}^{[x]-1} e^{-k} + (1-e^{-1}) e^{-[x]} \frac{1-e^{-(x-[x])}}{1-e^{-1}} =$$

$$= 1-e^{-x},$$

q.e.d.

Is $x = t+x_t$ een aselecte trekking uit de exponentiële verdeling met parameter $\lambda=1$, dan is $\frac{1}{\lambda} x$ een aselecte trekking uit de exponentiële verdeling met parameter λ. Immers:

$$P\left[\frac{1}{\lambda} \underline{x} \leqq x\right] = P\left[\underline{x} \leqq \lambda x\right] = 1-e^{-\lambda x}. \tag{3.33}$$

3. Poisson-verdeling.

$$F(n) = \sum_{k=0}^{n} \frac{(\lambda t)^k}{k!} e^{-\lambda t} \tag{3.34}$$

In hoofdstuk XV, pag. 9-10 (stelling nr 2) is bewezen, dat als het aantal klanten (bijv. voor een loket) dat in een tijdsinterval van lengte t aankomt, een Poisson-verdeling heeft met

parameter at, de lengten van de tijdsintervallen tussen opvolgende aankomsten onderling onafhankelijk en exponentiëel verdeeld zijn met dezelfde parameter a. Als a=1, dan is (3.34) de verdelingsfunctie van het aantal klanten, aankomende in een tijdsinterval van lengte λt. We vinden dus een aselecte trekking n uit (3.34) door n zo te bepalen dat

$$\sum_{k=1}^{n} x_k \leqq \lambda t < \sum_{k=1}^{n+1} x_k \qquad (3.35)$$

waarbij $x_1,\ldots,x_{n+1}$ aselecte trekkingen uit de exponentiële verdeling met parameter 1 zijn (de verdeling der tijden tussen opvolgende aankomsten).

In de meeste gevallen zal men ook hier liever tot rechtstreekse inversie van de verdelingsfunctie F(n) overgaan, te meer waar de boven beschreven methode betrekkelijk tijdrovend is, immers men dient voor elke trekking uit F(n) een aantal trekkingen uit de exponentiële verdeling te bepalen volgens 2.a) of b).

4. Γ_r <u>verdeling</u> met r geheel positief (Erlang-verdeling).

$$F(x) = \frac{\lambda^r}{\Gamma(r)} \int_0^x t^{r-1} e^{-\lambda t} dt \ . \qquad (3.36)$$

In hoofdstuk XV, pag. 10-11 (hulpstelling nr 5) is bewezen: wanneer de tijdsintervallen tussen opvolgende aankomsten aan een loket exponentieel verdeeld zijn met parameter λ, dan heeft het tijdstip van aankomst van de r^e klant de verdeling (3.36). De som van r aselecte trekkingen uit de exponentiële verdeling met parameter λ is dus een aselecte trekking uit (3.36).

5. <u>Normale verdeling</u>.

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}} \ . \qquad (3.37)$$

De verwachting en de variantie van een homogeen op het interval (0,1) verdeelde stochastische variabele <u>y</u> worden resp.

van

gegeven door:

$$\mathcal{E}\,\underline{y} = \int_0^1 y\,dy = \frac{1}{2} \qquad (3.38)$$

en

$$\sigma^2(\underline{y}) = \int_0^1 y^2\,dy - \frac{1}{4} = \frac{1}{12} \qquad (3.39)$$

De som van n homogeen op (0,1) verdeelde stochastische variabelen zal bij benadering normaal verdeeld zijn met gemiddelde $\frac{1}{2}$ n en spreiding $\sqrt{\frac{n}{12}}$ (Centrale limietstelling, zie hst. IV pag. 14; stelling 3;2). Is $\underline{t}$ de som van n aselecte getallen uit (0,1), dan zal dus

$$\underline{x} = \mu + \frac{\sigma(\underline{t} - \frac{n}{2})}{\sqrt{\frac{n}{12}}} \qquad (3.40)$$

bij benadering volgens (3.37) verdeeld zijn. De waarde van n die gekozen wordt, is afhankelijk van de vereiste nauwkeurigheid. In sommige gevallen is een waarde van n tussen 6 en 10 reeds voldoende nauwkeurig. In de beide staarten van de verdeling is de benadering echter slecht, en deze wordt met toenemende n slechts langzaam beter. Beschouw bv. het geval $\mu = 0$, $\sigma = 1$ en n=12, en zij $\underline{u}$ standaard-normaal verdeeld. Dan is $P\left[\underline{x} > 2\right] = P\left[\underline{u} > 2{,}0089\right]$ en $P\left[\underline{x} > 4\right] = P\left[\underline{u} > 4{,}3004\right]$. Teneinde de benadering in de staarten te verbeteren, is deze methode door TEICHROEW (1953) gewijzigd. Deze gewijzigde methode is, bij n=12, binnen 4σ - grenzen om μ zeer nauwkeurig, maar wordt toch weinig toegepast.

Door BOX en MULLER (1958) is een eenvoudige exacte methode aangegeven om normaal verdeelde grootheden voort te brengen. Als $\underline{y}_1$ en $\underline{y}_2$ twee aselecte getallen zijn, dan zijn

$$\underline{x}_1 = \sqrt{-2 \log \underline{y}_1}\,.\,\cos 2\pi\, \underline{y}_2 \qquad (3.41)$$

en

$$\underline{x}_2 = \sqrt{-2 \log \underline{y}_1}\,.\,\sin 2\pi\, \underline{y}_2$$

twee onderling onafhankelijke standaard-normaal verdeelde

grootheden. Deze methode is weliswaar ook in de staarten exact, doch is vrij tijdrovend en bovendien geen monotone transformatie.

Een snelle en nauwkeurige methode is ontwikkeld door MULLER (1958). Dit is een inversie-methode, waarbij de inverse normale verdelingsfunctie wordt benaderd door een groot aantal Chebyshev-polynomen van 1^e, 2^e of 4^e graad, terwijl in de staarten een afzonderlijke techniek wordt toegepast. Het bezwaar van deze methode is echter, dat een zeer groot aantal coefficienten in het geheugen van de computer moet worden opgeslagen.

Een eenvoudiger inversie-methode verkrijgt men door de inverse normale verdelingsfunctie te benaderen met een functie afkomstig van HASTINGS (1955). Als $\underline{y}$ een aselect getal is, dan is

$$\underline{u} = u(\underline{y}) = \begin{cases} \underline{\eta} - \dfrac{a_0+a_1\underline{\eta}+a_2\underline{\eta}^2}{1+b_1\underline{\eta}+b_2\underline{\eta}^2+b_3\underline{\eta}^3} & \text{met } \underline{\eta} = \sqrt{\log \dfrac{1}{\underline{y}^2}} \\ & \text{als } 0 < \underline{y} \leq \dfrac{1}{2} \\ - \underline{\eta} + \dfrac{a_0+a_1\underline{\eta}+a_2\underline{\eta}^2}{1+b_1\underline{\eta}+b_2\underline{\eta}^2+b_3\underline{\eta}^3} & \text{met } \underline{\eta} = \sqrt{\log \dfrac{1}{(1-\underline{y})^2}} \\ & \text{als } \dfrac{1}{2} \leq \underline{y} < 1 \end{cases} \qquad (3.42)$$

bij goede benadering standaard-normaal verdeeld. De coefficienten hebben de waarden:

$a_0 = 2,515517$ $\quad b_1 = 1,432788$
$a_1 = 0,802853$ $\quad b_2 = 0,189269$
$a_2 = 0,010328$ $\quad b_3 = 0,001308$

Bij dezelfde overschrijdingskans is het verschil tussen de waarde van $\underline{u}$ en de standaard-normaal verdeelde grootheid nooit groter dan 0,0006, behoudens in een gebied in de staarten met een kans kleiner dan 5.10^{-8}.

Aselecte trekkingen uit normale verdelingen zijn vereist voor vele soorten steekproefexperimenten. Tabellen van aselecte trekkingen uit de standaard-normale verdeling zijn dan ook beschikbaar, bijv. H. WOLD (1948).

6. χ^2_n-<u>verdeling</u>.

$$f(x) = \frac{1}{2^{n/2}\,\Gamma(\frac{n}{2})}\; x^{\frac{n}{2}-1}\; e^{-\frac{x}{2}} \quad (x > 0). \qquad (3.43)$$

De som van de kwadraten van n N(0,1) verdeelde variabelen heeft een χ^2_n-verdeling (3.43). Men verkrijgt dus een aselecte trekking uit (3.43) door de kwadraten van n aselecte trekkingen uit de normale verdeling met $\mu = 0$ en $\sigma = 1$ op te tellen.

Uit (3.43) is gemakkelijk af te leiden dat $\frac{1}{2}\chi^2_2$ een exponentiële verdeling met parameter $\lambda = 1$ bezit. Dit geeft dus een nieuwe methode om aselecte trekkingen uit de exponentiële verdeling te krijgen.

4. Variantiereductie van Monte Carlo-schattingen

Zoals reeds in paragraaf 1 werd opgemerkt, zijn de uitkomsten van de Monte Carlo-methode stochastische grootheden, omdat ze verkregen worden uit steekproeven uit de verdelingen die in het model voorkomen. Deze uitkomsten (schattingen voor bepaalde in het model optredende grootheden) hebben een variantie die een maat is voor hun onnauwkeurigheid, en die des te kleiner is naarmate de oorspronkelijke steekproefomvang groter is. Om de variantie beneden een vast gekozen grens te houden, moet men de variantie als functie van de steekproefgrootte kennen. Daar men echter juist de nabootsing van het model uitvoert als een analytische behandeling moeilijk is, zal men in het algemeen de gezochte variantie niet kunnen berekenen. Men kan dan de gevonden steekproefvariantie nemen als schatting van de werkelijke variantie en achteraf aangeven of de steekproefomvang groot genoeg is geweest.

De steekproefvariantie kan echter een zeer slechte schatting van de werkelijke variantie zijn en men moet dus voorzichtig zijn om uit de steekproefvariantie conclusies te trekken over de nauwkeurigheid van de gebruikte schattingsmethode.

Voorbeeld 1.

Als voorbeeld van een slechte schattingsmethode geven we een schatting van $\mu_{9,25}$ van een exponentiële verdeling met gemiddelde 1:

$$\mu_{9,25} = \mathcal{E}\underline{x}^{9,25} = \int_0^\infty x^{9,25} e^{-x} dx \qquad (4.1)$$

Doet men nu aselecte trekkingen $x_1, \ldots, x_n$ uit de exponentiële verdeling, dan is

$$m_{9,25} = \frac{1}{n} \sum_{i=1}^{n} x_i^{9,25} \qquad (4.2)$$

een schatting van (4.1). De steekproefvariantie is:

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i^{9,25} - m_{9,25})^2 \qquad (4.3)$$

Bij een berekening met n=100 werden op de volgende wijze aselecte getallen gebruikt uit M.G. KENDALL en B. BABINGTON's "table of random sampling numbers": van het 19^e duizendtal werden

van de voorste groep van 8 cijfers de eerste 5 cijfers als aselect getal gebruikt. Dit levert 25 aselecte getallen. Vervolgens werd hetzelfde procédé gevolgd voor de tweede, derde en vierde groep van 8 cijfers. De absolute waarden van de logarithmen van deze aselecte getallen zijn de aselecte trekkingen $\{x_i\}$, verwerkt in (4.2).

Deze schatting geeft dan:

$$m_{9,25} = 9364 \tag{4.4}$$

met steekproefvariantie:

$$s^2 = (63.456)^2 \tag{4.5}$$

De werkelijke waarden voor $\mu_{9,25}$ en $\sigma^2(\underline{m}_{9,25})$ zijn echter:

$$\mu_{9,25} = \Gamma(10,25) = 639.233 \tag{4.6}$$

en

$$\sigma^2(\underline{m}_{9,25}) = \frac{1}{100}\left\{\int_0^\infty x^{18,5}e^{-x}dx - \mu_{9,25}^2\right\} =$$

$$= \frac{1}{100}\left\{\Gamma(19,5) - (\Gamma(10,25))^2\right\} = (1,665.10^7)^2. \tag{4.7}$$

De verkregen schattingen zijn dus bijzonder slecht, zoals te begrijpen is, aangezien de grootheid $x^{9,25}$ waarvan hier het gemiddelde bepaald wordt, buitengewoon scheef verdeeld is. Immers:

$$P\left[\underline{x}^{9,25} > \mu_{9,25}\right] = 0,0145 \ . \tag{4.8}$$

We zullen straks een betere schattingsmethode voor $\mu_{9,25}$ aangeven.

Een middel om de variantie te verminderen, is vergroting van de steekproefomvang. Evenals in voorbeeld 1 verkrijgt men in vele gevallen echter een aanmerkelijke variantiereductie slechts bij een enorm veel grotere steekproefomvang. Om de variantie te verminderen heeft men daarom andere methoden ontwikkeld, die alle neerkomen op wijziging van het steekproefschema, d.w.z. geschikte wijziging van de verdeling waaruit een steekproef gedaan wordt.

Bij een praktisch probleem wordt dikwijls de grootste variantiereductie bereikt door een zo efficiënt mogelijk gebruik van de speciale gegevens van het probleem. We zullen echter slechts enige algemenere methoden bespreken. Deze zijn geen standaardtechnieken,

maar aanwijzingen om bij een gegeven probleem een weg te vinden om de variantie te reduceren. We zullen de betreffende ideeën dan ook slechts aan elementaire voorbeelden illustreren. Een overzicht van verscheidene van deze methoden vindt men bij H. KAHN (1956).

4.1. Importance sampling. We kiezen als voorbeeld het schatten van een bepaalde integraal

$$I = \int_{-\infty}^{\infty} g(x)f(x)dx \qquad (4.9)$$

waarin g(x) een bekende functie en f(x) een bekende verdelingsdichtheid is. De meest voor de hand liggende schatting van (4.9) is

$$\underline{I}_1 = \frac{1}{n} \sum_{i=1}^{n} g(\underline{x}_i) \qquad (4.10)$$

waarin $\{x_i\}$ aselecte trekkingen uit de verdeling met verdelingsdichtheid f(x) zijn.

We kunnen echter (4.9) ook anders schrijven:

$$I = \int_{-\infty}^{\infty} \frac{g(x)f(x)}{h(x)} h(x)dx \qquad (4.11)$$

waarin I de verwachting is van $\frac{g(x)f(x)}{h(x)}$ ten opzichte van een bekende verdelingsdichtheid h(x). Hieruit volgt dat we als schatting van I eveneens kunnen nemen:

$$\underline{I}_2 = \frac{1}{n} \sum_{i=1}^{n} \frac{g(\underline{y}_i)f(\underline{y}_i)}{h(\underline{y}_i)} \qquad (4.12)$$

waarbij $\{y_i\}$ aselecte trekkingen zijn uit de verdeling met verdelingsdichtheid h(x). We willen nu h(x) zodanig kiezen, dat de variantie van $\underline{I}_2$ minimaal is; dus moet

$$\sigma^2(\underline{I}_2) = \frac{1}{n} \left\{ \int_{-\infty}^{\infty} \frac{(g(x)f(x))^2}{h(x)} dx - I^2 \right\} \qquad (4.13)$$

minimaal zijn, met bijvoorwaarden

$$h(x) \geqq 0 \ (-\infty < x < \infty) \text{ en } \int_{-\infty}^{\infty} h(x)dx = 1 \ . \qquad (4.14)$$

Dit is een probleem uit de variantierekening dat kan worden opgelost met de multiplicatormethode van LAGRANGE (men zie bijv. R. COURANT, Diff. and integral calc. II, Chapter VII, 1948). De gezochte h(x) wordt verkregen door de integrand van

$$\int_{-\infty}^{\infty}\frac{(g(x)f(x))^2}{h(x)}\,dx + \lambda\int_{-\infty}^{\infty}h(x)dx = \int_{-\infty}^{\infty}\left\{\frac{(g(x)f(x))^2}{h(x)} + \lambda h(x)\right\}dx \qquad (4.15)$$

te differentiëren naar h(x) als onafhankelijke variabele, en de afgeleide gelijk nul te stellen. Dit geeft:

$$\left(\frac{g(x)f(x)}{h(x)}\right)^2 - \lambda = 0 \qquad (4.16)$$

of

$$h(x) = \frac{|g(x)|f(x)}{\sqrt{\lambda}} \; . \qquad (4.17)$$

Uit (4.14) volgt dan:

$$h(x) = \frac{|g(x)|f(x)}{\int_{-\infty}^{\infty}|g(x)|f(x)dx} \; . \qquad (4.18)$$

Door substitutie van (4.18) in (4.13) krijgen we de minimum-variantie:

$$\min \sigma^2(\underline{I}_2) = \frac{1}{n}\left\{\left(\int_{-\infty}^{\infty}|g(x)|f(x)dx\right)^2 - \left(\int_{-\infty}^{\infty}g(x)f(x)dx\right)^2\right\}. \qquad (4.19)$$

Als $g(x) \geqq 0$ voor $-\infty < x < \infty$, dan is (4.19) gelijk nul en $\underline{I}_2 = I$ wegens (4.12) en (4.18). Uit (4.18) blijkt dan echter ook, dat we langs deze weg de optimale h(x) niet kunnen bepalen, omdat we daarvoor juist de oorspronkelijke integraal (4.9) zouden moeten kennen. We kunnen echter opmerken, dat voor de optimale h(x) de functie

$$\frac{g(x)f(x)}{h(x)} \qquad (4.20)$$

waarvan het gemiddelde bepaald wordt t.o.v. de verdelingsdichtheid h(x), een constante (nl. I) is. In het algemeen wordt de variantie van (4.12) dus verminderd, als men h(x) zo kiest dat de functiewaarden van (4.20) een geringere variatie hebben dan de oorspronkelijke functie $g(x) = \frac{g(x)f(x)}{f(x)}$.

Het bovenstaande kan toegepast worden in voorbeeld 2 van paragraaf 2. We schrijven $I = \int_0^1 12x^3(1-x)dx = \int_0^1 x.12x^2(1-x)dx$ en geven hiervan de schatting $\underline{I}_2 = \frac{1}{10}\sum_1^{10}\underline{x}_i$, waarbij $\{\underline{x}_i\}$ aselecte

trekkingen zijn uit de B(3,2)-verdeling met verdelingsdichtheid $12x^2(1-x)$. Uit 10 aselecte getallen berekene de lezer $x_1,\ldots,x_{10}$ m.b.v. de vergelijking (3.4) uit paragraaf 3 en KARL PEARSON's "Tables of the incomplete beta-function". De variantie $\sigma^2(\underline{I}_2) = 0{,}004$ is kleiner dan 0,02, de variantie van de eerste schatting $\underline{I}_1$.

Nog veel duidelijker is het belang van bovenstaande methode echter in de volgende toepassing.

Voorbeeld 2.

We geven een tweede schatting van $\mu_{9,25}$ van de exponentiële verdeling. We schrijven (4.1) nu als

$$\mu_{9,25} = \int_0^\infty 9!\, x^{0,25}\, \frac{x^9 e^{-x}}{9!}\, dx \quad . \tag{4.21}$$

Hierbij wordt $\mu_{9,25}$ dus beschouwd als $9!\ \mathcal{E}\underline{x}^{0,25}$ behorend bij de Γ-verdeling met 10 vrijheidsgraden. Voor een Monte Carlo-schatting van (4.21) wordt gebruik gemaakt van dezelfde aselecte trekkingen uit de exponentiële verdeling als in voorbeeld 1. Volgens paragraaf 3 kan de som van 10 opvolgende van deze trekkingen gebruikt worden als aselecte trekking uit de Γ-verdeling. De 100 aselecte trekkingen $x_1,\ldots,x_{100}$ uit de exponentiële verdeling leveren dus 10 aselecte trekkingen $y_1,\ldots,{}_{10}$ uit de Γ-verdeling, en hierbij geldt

$$y_k = \sum_{i=1}^{10} x_{10(k-1)+i} \qquad k=1,\ldots,10 \quad . \tag{4.22}$$

Als schatting voor (4.21) wordt nu verkregen:

$$m'_{9,25} = \frac{9!}{10} \sum_{k=1}^{10} y_k^{0,25} = 631.331 \quad . \tag{4.23}$$

De steekproefvariantie bedraagt (n=10):

$$s^2 = \frac{1}{n-1} \sum_{k=1}^{n} (9!y_k^{0,25} - m'_{9,25})^2 = (58.510)^2 \quad . \tag{4.24}$$

De werkelijke variantie van $9!\underline{y}^{0,25}$ bedraagt voor een Γ-verdeling met 10 vrijheidsgraden:

$$\sigma^2(9!\underline{y}^{0,25}) = (9!)^2 \int_0^\infty y^{0,5}\, \frac{y^9 e^{-y}}{9!}\, dy - \mu_{9,25}^2 =$$

$$= 9!\,\Gamma(10,5) - (\Gamma(10,25))^2 = (51.242)^2 \tag{4.25}$$

dus de variantie van (4.23) is

$$\sigma^2(\underline{m}'_{9,25}) = \frac{1}{10}\sigma^2(9!\underline{y}^{0,25}) = \frac{(51.242)^2}{10} = (16.204)^2. \quad (4.26)$$

Vergelijkt men (4.26) met (4.7) dan blijkt door importance sampling de variantie te zijn verminderd met een factor

$$\left(\frac{1,665.107}{16.204}\right)^2 \sim 10^6 . \quad (4.27)$$

Bij het toepassen van importance sampling is de grootste moeilijkheid het vinden van een goede verdelingsdichtheid h(x). Voor het geval dat de keuze van h(x) beperkt wordt tot een bepaalde klasse verdelingsdichtheden h(x,Θ), die zich onderscheiden door verschillende waarden van de parameter Θ, is door A.W. MARSHALL (1956) een methode aangegeven om de optimale h(x.Θ) te benaderen.

Een andere methode om schattingen te verkrijgen van gemiddelden van grootheden die een bepaalde verdeling bezitten, door middel van steekproeven uit andere verdelingen, is de Conditional Monte Carlo. Deze methode is echter alleen geschikt voor steekproefexperimenten die met typisch statistische problemen samenhangen. Voor literatuur hierover zie men: H.F. TROTTER en J.W. TUKEY (1956) en J.M. HAMMERSLEY (1956).

4.2. Gecorreleerde steekproeven. We keren terug tot het schatten van de integraal (4.9). We veronderstellen dat we de integraal

$$I' = \int_{-\infty}^{\infty} w(y)v(y)dy \qquad (4.28)$$

kennen, waarin de functie w(y) en de verdelingsdichtheid v(y) op een bepaalde wijze samenhangen met de functie g(x) en de verdelingsdichtheid f(x). Is deze samenhang nauw, dan kunnen schattingen van I en I' sterk positief gecorreleerd zijn als zij worden uitgevoerd met dezelfde aselecte getallen. Deze eigenschap kunnen we gebruiken om de variantie van de schatting van I te verminderen.

Als nieuwe schatting voeren we in:

$$\underline{I}_3 = \frac{1}{n} \sum_{i=1}^{n} \{g(\underline{x}_i) - \alpha(w(\underline{y}_i) - I')\} \; . \qquad (4.29)$$

De $\{x_i\}$ en $\{y_i\}$ zijn aselecte trekkingen uit de verdelingen met dichtheid f(x) resp. v(y), gebaseerd op dezelfde aselecte getallen. De constante α in (4.29) moet nog nader bepaald worden.

Men ziet gemakkelijk in dat (4.29) een zuivere schatting is van I:

$$\mathcal{E}(\underline{I}_3) = I \qquad (4.30)$$

met variantie:

$$\sigma^2(\underline{I}_3) = \frac{1}{n} \iint \{(g(x)-I) - \alpha(w(y)-I')\}^2 \varphi(x,y)dxdy \qquad (4.31)$$

waarin $\varphi(x,y)$ de simultane verdelingsdichtheid van x en y voorstelt. Uitwerking van (4.31) geeft:

$$\sigma^2(\underline{I}_3) = \frac{1}{n} \{\sigma_1^2 - 2\alpha\rho\sigma_1\sigma_2 + \alpha^2\sigma_2^2\} \qquad (4.32)$$

met:

$$\sigma_1^2 = \int_{-\infty}^{\infty} (g(x))^2 f(x)dx - I^2 \qquad (4.33)$$

$$\sigma_2^2 = \int_{-\infty}^{\infty} (w(y))^2 v(y)dy - I'^2 \qquad (4.34)$$

$$\rho = \frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (g(x)-I)(w(y)-I')\,\varphi(x,y)dxdy}{\sigma_1\sigma_2} \qquad (4.35)$$

De beste waarde voor α in (4.29) is die waarvoor (4.32) minimaal wordt. Door de afgeleide naar α van (4.32) gelijk nul te stel-

len, vinden we voor de beste α:

$$\alpha = \rho \frac{\sigma_1}{\sigma_2} \quad . \tag{4.36}$$

Substitutie van (4.36) in (4.32) geeft nu de minimumvariantie:

$$\min \ \sigma^2(\underline{I}_3) = \frac{1}{n} \sigma_1^2(1-\rho^2) \quad . \tag{4.37}$$

De variantiereductiefactor t.o.v. de schatting $\underline{I}_1$ in (4.10) is $\dfrac{\sigma^2(\underline{I}_1)}{\sigma^2(\underline{I}_3)}$; deze bedraagt in dit geval $(1-\rho^2)^{-1}$ en is zeer groot als $\rho \sim 1$. Voor kleinere ρ wordt de betekenis van deze factor snel kleiner. Bij vele problemen komt de besproken methode echter in aanmerking om toegepast te worden.

In de praktijk kan men de optimale α niet bepalen, omdat men voor de berekening van ρ en σ_1 juist de waarde van I nodig heeft. Meestal wordt in dit geval $\alpha=1$ gekozen. Is dan $\sigma_1 \sim \sigma_2$ en $\rho \sim 1$, dan verschilt $\alpha=1$ niet veel van de beste α en wordt een aanzienlijke variantiereductie verkregen.

Het is echter ook mogelijk de beste α door een steekproefexperiment te schatten. Men kan daartoe de verkregen steekproef in twee gedeelten splitsen; met behulp van de trekkingen $\{x_i\}$ en $\{y_i\}$ in elk deel worden (op een of andere wijze) ρ, σ_1 en σ_2 geschat, wat een schatting geeft voor de beste α in (4.36). De α, verkregen uit het eerste deel, wordt nu in (4.29) gebruikt met de $\{x_i\}$ en $\{y_i\}$ uit het tweede deel, en omgekeerd. Aldus worden twee waarden $\underline{I}_3$ verkregen, waarvan het gemiddelde nog een zuivere schatting van I is.

Voorbeeld 3.

We geven een schatting van $\mu_{9,26}$ van de exponentiële verdeling in de veronderstelling dat we $\mu_{9,25} = 639.233$ kennen:

$$\mu_{9,26} = \int_0^\infty x^{9,26} \, e^{-x} dx \quad . \tag{4.38}$$

We gebruiken hierbij dezelfde aselecte trekkingen uit de Γ-verdeling als in voorbeeld 2 en maken gebruik van de daar gegeven schatting van $\mu_{9,25}$.

Analoog aan de schatting (4.23) van $\mu_{9,25}$ in voorbeeld 2

schatten we (4.38) door:

$$m'_{9,26} = \frac{9!}{10} \sum_{i=1}^{10} y_i^{0,26} = 645.600 \qquad (4.39)$$

De werkelijke waarde van $\mu_{9,26}$ bedraagt:

$$\mu_{9,26} = \Gamma(10,26) = 653.963 \; . \qquad (4.40)$$

De variantie van de schatting (4.39) bedraagt:

$$\sigma^2(\underline{m}'_{9,26}) = \frac{1}{10} \sigma^2(9!\underline{y}^{0,26}) = \frac{1}{10}\{9!\,\Gamma(10,52) - (\Gamma(10,26))^2\} =$$

$$= \frac{1}{10}(54.507)^2 = (17.236)^2. \qquad (4.41)$$

Een schatting overeenkomstig (4.29) met $\alpha = 1$, gebruikmakend van de schatting van $\mu_{9,25}$ in voorbeeld 2, levert:

$$\frac{9!}{10} \sum_{i=1}^{10} \{ y_i^{0,26} - (y_i^{0,25} - \frac{\mu_{9,25}}{9!}) \} = 653.492 \qquad (4.42)$$

wat goed overeenstemt met (4.40).

De variantie van de schatting (4.42) vindt men volgens (4.32) met $\alpha=1$, $n=10$, $\sigma_1^2 = \sigma^2(9!\underline{y}^{0,26})$ uit (4.41) en $\sigma_2^2 = \sigma^2(9!\underline{y}^{0,25})$ uit (4.25) en:

$$\rho = \frac{9!\,\Gamma(10,51) - \Gamma(10,25).\Gamma(10,26)}{\sigma_1 \sigma_2} = 0,999860. \qquad (4.43)$$

Voor de variantie van (4.42) vindt men tenslotte:

$$\frac{1}{10}(\sigma_1^2 - 2\rho\sigma_1\sigma_2 + \sigma_2^2) = (1069)^2 \; . \qquad (4.44)$$

Vergelijking van (4.41) en (4.44) leert dat de variantie bij deze methode verminderd is met een factor:

$$\left(\frac{17.236}{1069}\right)^2 = 260 \; . \qquad (4.45)$$

Een schatting uitgevoerd met de optimale waarde van $\alpha = \rho\frac{\sigma_1}{\sigma_2} =$ $= 1,064$ geeft als schatting voor $\mu_{9,26}$: 654.019. De variantie van deze schatting bedraagt volgens (4.37): $(288)^2$. Hierdoor wordt de variantie dus nog verminderd met een factor

$$\left(\frac{1069}{288}\right)^2 = 14 \qquad (4.46)$$

want de totale vermindering is $\left(\frac{17.236}{288}\right)^2 =$

$$\left(\frac{17.236}{1069}\right)^2 \left(\frac{1069}{288}\right)^2 .$$

In het bovenstaande werd een integraal geschat waarbij men de waarde van een hiermee in verband staande integraal nauwkeurig kent. Als men echter beide integralen moet schatten en behalve in de waarde van elke integraal afzonderlijk, ook geïnteresseerd is in hun verschil, dan verdient het aanbeveling om beide schattingen te baseren op dezelfde aselecte getallen. Doet men dit niet, dan zal de variantie van het verschil de som zijn van de varianties der beide schattingen afzonderlijk, terwijl bij gebruikmaking van dezelfde aselecte getallen men de variantie volgens (4.32) vindt met $\alpha=1$. Deze laatste variantie bedraagt $\frac{1}{n}\{\sigma_1^2-2\rho\sigma_1\sigma_2+\sigma_2^2\}$ en is wegens $\rho > 0$ kleiner dan de eerste, die $\frac{1}{n}(\sigma_1^2+\sigma_2^2)$ bedraagt.

4.3. Antithetic variates. Deze methode is afkomstig van J.M. HAMMERSLEY en K.W. MORTON (1956). We geven een schets van een toepassing bij het schatten van de integraal

$$I = \int_0^1 g(x)dx \quad . \tag{4.47}$$

We kunnen hiervoor als schatting gebruiken:

$$\underline{I}_1 = \frac{1}{n}\sum_{i=1}^{n} g(\underline{x}_i) \tag{4.48}$$

waarbij $\{x_i\}$ aselecte getallen zijn. Een andere schatting is

$$\underline{t} = \sum_{i=1}^{n}(a_i-a_{i-1})g\{a_{i-1}+(a_i-a_{i-1})\underline{x}_i\} \tag{4.49}$$

waarin $0=a_0<a_1<\ldots<a_m=1$ en $\{x_i\}$ onafhankelijk verkregen aselecte getallen zijn. (4.49) is een zuivere schatting van I:

$$\mathcal{E}\,\underline{t} = I \quad . \tag{4.50}$$

Als de deelpunten $\{a_i\}$ geschikt gekozen worden, zal de variantie van (4.49) in het algemeen kleiner zijn dan die van (4.48).

We nemen nu de $\{x_i\}$ niet meer onafhankelijk, maar leggen er een zodanig verband tussen, dat als sommige termen in (4.49) relatief groot zijn, andere termen relatief klein zijn. Hierdoor kan de variantie verder gereduceerd worden, terwijl de schatting zuiver

blijft.

Voor n=2 gaat (4.49) over in:

$$\underline{t} = ag(a\underline{x}_1) + (1-a)g\{a+(1-a)\underline{x}_2\} \quad . \qquad (4.51)$$

Op grond van het verloop van g(x) in het interval (0,1), bijv. monotoon of met een extreem in (0,1), kunnen we nu het verband tussen $\underline{x}_1$ en $\underline{x}_2$ zo trachten te kiezen, dat $\underline{t}$ een geringe spreiding heeft. Voor n=4,8,... kan dit procédé telkens herhaald worden. Op de uitwerking hiervan gaan we niet in.

Het is mogelijk gebleken om met de methode van de antithetic variates op grond van een globale indruk van het verloop van functies aanzienlijke variantiereductie te bereiken voor zeer ingewikkelde schattingsproblemen.

5. Voorbeeld

We stellen ons voor dat we een voorraad massagoederen moeten beheren, waaruit door klanten op onregelmatige tijdstippen volgens hun behoeften zekere hoeveelheden worden betrokken. We veronderstellen dat het aantal klanten dat in een bepaald tijdsinterval een bestelling doet, een Poisson-verdeling bezit. Dit betekent dat er behoudens een kans nul slechts één klant tegelijk aankomt en dat de intervallen tussen twee opeenvolgende aankomsttijdstippen onderling onafhankelijk exponentiëel verdeeld zijn. Ook nemen we aan dat de grootten der bestellingen onderling onafhankelijk identiek en continu verdeeld zijn (dus niet noodzakelijk veelvouden van een of andere eenheid zijn).

Wanneer onze voorraad te zeer is verminderd of uitgeput, vullen we deze aan door een order te plaatsen bij een buitenlandse leverancier. Er is overeengekomen dat deze ons elke keer een constante hoeveelheid toezendt, bijv. met een schip. De schepen kunnen aan vertraging onderhevig zijn, maar halen elkaar niet in. Hun reisduur is dus stochastisch.

Als een klant meer bestelt dan onze voorraad bedraagt, verkopen we alvast wat we eventueel nog hebben en wachten dan aanvullingen uit het buitenland af, waaruit het ontbrekende geleverd wordt. We zijn contractueel verplicht om aan de klant een boete C_2 per ton te betalen voor elke dag dat we te laat afleveren. Verder zijn de opslagkosten C_1 per ton per dag.

Als we nu verder onder "voorraad" verstaan de hoeveelheid goederen in de opslagruimte, vermeerderd met de orders die eventueel al uit het buitenland naar ons onderweg zijn, dan stellen we nu de vraag tot hoever we deze voorraad steeds zullen laten dalen, alvorens een nieuwe order in het buitenland te plaatsen. Het criterium is dat bij deze grenswaarde van de voorraad de gemiddelde kosten, bestaande uit opslagkosten en boetes, op de lange duur minimaal zijn. Daar de in- en de verkoopsprijs van de goederen tevoren vaststaan, is hiermee bij een bepaalde verwachte vraag een constant bedrag gemoeid, dat we verder buiten beschouwing laten.

Uit statistische gegevens blijken er per maand gemiddeld 8 klanten te komen. De kans op n klanten in een periode van T maanden is dus:

$$p_{n,T} = \frac{(\lambda T)^n}{n!} e^{-\lambda T} \qquad (5.1)$$

met $\lambda = 8$.

Verder werd voor de levertijd uit het buitenland een Γ-verdeling met 4 vrijheidsgraden gevonden. De gemiddelde levertijd bedroeg 1 maand. De verdelingsfunctie van de levertijd T in maanden is dus:

$$L(T) = \frac{\alpha^\mu}{\Gamma(\mu)} \int_0^T t^{\mu-1} e^{-\alpha t}\, dt \qquad (5.2)$$

met $\alpha = \mu = 4$.

De verdeling van de vraag per klant sloot goed aan bij een Γ-verdeling met 2 vrijheidsgraden en gemiddelde 200 ton. Voor deze verdeling vinden we dus:

$$H(x) = \frac{\beta^\nu}{\Gamma(\nu)} \int_0^x t^{\nu-1} e^{-\beta t}\, dt \qquad (5.3)$$

met $\nu = 2$, $\beta = 1$. De grootte der bestellingen wordt dus uitgedrukt in eenheden van 100 ton.

De grootte van elke order bij de buitenlandse leverancier bedroeg q=500 ton.

We gaan nu de optimale grootte X van de voorraad bepalen. Zodra de voorraad beneden de grens X komt, wordt een nieuwe order ter grootte q geplaatst.

Om X te bepalen kan men het proces een aantal malen nabootsen met verschillende grenzen, tot waar men de voorraad laat dalen al-

vorens een nieuwe order te plaatsen. Op grond van de tijdens zo'n proces te betalen opslagkosten en boetes kan men door onderlinge vergelijking de optimale X benaderen. Deze methode is echter zeer moeizaam. Immers men moet bij een aantal grenzen $X_1, X_2, \ldots$ de gemaakte kosten bepalen en deze vergelijken. Bovendien vindt men bij elke $X_1, X_2, \ldots$ van de gemiddelde kosten slechts een schatting.

Daarom kan men het probleem beter eerst analytisch aanpakken en pas daar waar men met analytische methodes niet verder komt, een steekproefexperiment uitvoeren.

Uit formule (5.22) van hoofdstuk XI (Voorraadproblemen) van deze Leergang volgt dat de optimale X bepaald kan worden uit

$$M(X) = \frac{1}{q} \int_{X}^{X+q} F(s)ds = \frac{C_2}{C_1+C_2} \qquad (5.4)$$

waarin q, C_1, C_2 de bovengenoemde betekenis hebben en F(s) de verdelingsfunctie is van de vraag tijdens een levertijd (van een bestelling q in het buitenland).

Om X uit (5.4) op te lossen moeten we eerst F(s) berekenen. Het aantal klanten in een vaste levertijd T heeft volgens (5.1) de verdelingsfunctie

$$P(n|T) = \sum_{i=0}^{n} \frac{(\lambda T)^i}{i!} e^{-\lambda T} \quad . \qquad (5.5)$$

Als verdelingsfunctie van het aantal klanten tijdens een stochastische levertijd krijgen we wegens (5.2)

$$P(n) = \sum_{i=0}^{n} \int_{0}^{\infty} \frac{(\lambda T)^i}{i!} e^{-\lambda T} \, dL(T) =$$

$$= \sum_{i=0}^{n} \int_{0}^{\infty} \frac{\lambda^i \alpha^\mu}{i!\,\Gamma(\mu)} T^{i+\mu-1} e^{-(\lambda+\alpha)T} \, dT =$$

$$= \sum_{i=0}^{n} \binom{i+\mu-1}{\mu-1} \left(\frac{\lambda}{\lambda+\alpha}\right)^i \left(\frac{\alpha}{\lambda+\alpha}\right)^\mu . \qquad (5.6)$$

P(n) is dus een negatief binomiale verdeling.

De verdelingsfunctie van de vraag van één klant wordt gegeven door (5.3). De verdelingsfunctie van de vraag van n klanten, komende in een levertijd, is de verdelingsfunctie van $\underline{s} = \underline{s}_1 + \ldots + \underline{s}_n$, waarbij de $\underline{s}_i$ onderling onafhankelijk en alle volgens (5.3) verdeeld zijn. Wegens een bekende eigenschap van de Γ-verdeling heeft dan $\underline{s}$ een

Γ-verdeling met nv vrijheidsgraden, aan te geven door $H_n(s)$.

We vinden dus voor de verdelingsfunctie van de vraag in een stochastische levertijd:

$$F(s) = P\left[\underline{s} \leqq s\right] = P\left[\underline{s}=0\right] + P\left[0 < \underline{s} \leqq s\right] = P(0) +$$

$$+ \sum_{i=1}^{\infty} \binom{i+\mu-1}{\mu-1} \left(\frac{\lambda}{\lambda+\alpha}\right)^i \left(\frac{\alpha}{\lambda+\alpha}\right)^\mu H_i(s) = \left(\frac{\alpha}{\lambda+\alpha}\right)^\mu +$$

$$+ \sum_{i=1}^{\infty} \binom{i+\mu-1}{\mu-1} \left(\frac{\lambda}{\lambda+\alpha}\right)^i \left(\frac{\alpha}{\lambda+\alpha}\right)^\mu \frac{\beta^{iv}}{\Gamma(iv)} \int_0^s t^{iv-1} e^{-\beta t}\, dt. \qquad (5.7)$$

Met deze uitdrukking voor F(s) is in (5.4) niets te beginnen. Hoewel het mogelijk is een eenvoudiger vorm van F(s) af te leiden, zullen we nu eerst de Monte Carlo-methode toepassen. We doen een groot aantal trekkingen uit de verdeling F(s) (zie de volgende alinea) en berekenen hiervoor de cumulatieve frequentieverdeling G(s), d.i. de fractie der steekproefuitkomsten $\leqq s$. Substitueren we G(s) voor F(s) in (5.4), dan voldoet aan (5.4) een eenduidig bepaalde X, daar G(s) een monotoon niet-dalende trapfunctie is. De gevonden X is als resultaat van een steekproef stochastisch; notatie: $\underline{X}$. Om de nauwkeurigheid van de uitkomst te bepalen, moeten we dan nog de variantie van $\underline{X}$ kennen.

Een aselecte trekking uit de verdeling F(s) kan men opvatten als de totale vraag tijdens een (stochastische) levertijd. Door een trekking uit de negatief binomiale verdeling (5.6) wordt eerst het aantal klanten bepaald dat in een levertijd komt. Dit geschiedt doordat men de negatief binomiale verdeling als volgt kan voortbrengen: heeft men een alternatief met kans $\alpha/(\lambda+\alpha)$ op succes, dan is het aantal wansuccessen voorafgaande aan het μ^e succes, verdeeld volgens (5.6). Hier is $\mu=4$ en $\alpha/(\lambda+\alpha) = 4/12 = 1/3$.

Met de methode van LEHMER werden pseudo-aselecte getallen verkregen met

$$\left.\begin{aligned} u_0 &= 5^{11} \\ u_{n+1} &= 5^{11} u_n \pmod{2^{33}} \end{aligned}\right\} \qquad (5.8)$$

Het alternatief, te gebruiken in de verdeling (5.6), is het trekken van een aselect getal $< \frac{1}{3}$ of $> \frac{1}{3}$. Het eerste heeft een kans $\frac{1}{3}$ en wordt als een succes beschouwd. Is dus n het aantal aselecte getallen $> \frac{1}{3}$ voorafgaande aan het 4^e aselect getal $< \frac{1}{3}$, dan is n een

aselecte trekking uit (5.6) en geeft het aantal klanten dat in een levertijd een bestelling doet. Als n=0, is de vraag tijdens een levertijd , en dus de aselecte trekking uit F(s), gelijk nul. Als $n \geqq 1$, is voor een aselecte trekking uit F(s) nog een trekking uit de verdeling $H_n(s)$ nodig, dus uit een Γ-verdeling met nv=2n vrijheidsgraden en parameter $\beta=1$; d.i. volgens paragraaf 3 de som van 2n trekkingen uit een exponentiële verdeling met parameter =1, dus de absolute waarde van de logarithme van 2n aselecte getallen.

Op deze wijze werden 10.000 aselecte trekkingen uit F(s) berekend. Als speciale toets voor de aselecte getallen (5.8) werd nagegaan, hoeveel van de aselecte getallen die gebruikt werden voor het doen van aselecte trekkingen uit de negatief binomiale verdeling, groter dan $\frac{1}{3}$ en hoeveel kleiner dan $\frac{1}{3}$ waren. Er werd gevonden:

aselecte getallen $> \frac{1}{3}$: 80.472

" " $< \frac{1}{3}$: 40.000.

Dit resultaat komt zeer goed overeen met de verwachte aantallen bij de homogene verdeling op het interval (0,1)

De variantie van de stochastische variabele $\underline{X}$ kan men in principe als volgt bepalen. De functie M(X) in (5.4) is een verdelingsfunctie. Immers M(X) is een monotoon niet dalende functie van X, M(-q)=0, en

$$\lim_{X\to\infty} M(X) = \lim_{X\to\infty} \frac{1}{q} \int_{X}^{X+q} F(s)ds = 1 . \qquad (5.9)$$

Door middel van een steekproefexperiment wenst men nu het quantiel X te bepalen waarvoor $M(X) = \frac{C_2}{C_1+C_2}$. Men kan bewijzen dat de in het voorgaande uit de steekproef gevonden $\underline{X}$ asymptotisch normaal is met gemiddelde X en te berekenen spreiding.

Langs analytische weg kan uit (5.7) een eenvoudiger uitdrukking voor F(s) worden afgeleid. Substitueren we deze in (5.4), dan kunnen we hieruit X direct bepalen bij gegeven $\frac{C_2}{C_1+C_2}$. We zullen dit hier niet uitvoeren, maar vermelden slechts als resultaat bij enige waarden van $\frac{C_2}{C_1+C_2}$ de $\underline{X}$ uit de steekproef, de werkelijke X en de spreiding van $\underline{X}$.

c_2/c_1+c_2	$\underline{X}$ uit steekproef	X theoretisch	σ
0,75	19,5	19,4	0,2
0,80	21,7	21,55	0,2
0,85	24,5	24,2	0,2
0,90	28,3	27,9	0,3

Uit deze tabel volgt, dat onze Monte Carlo-schatting van X zeer nauwkeurig is.

Literatuur

J.L. ALLARD, A.R. DOBELL en T.E. HULL (1963): Mixed congruential random number generators for decimal machines. Journ.Ass.Comp. Mach.10,131.

V.D. BARNETT (1962):The behaviour of pseudo-random sequences generated on computers by the multiplicative congruential method. Math.Comp. 7,75.

E. en V.J. BOFINGER (1958): On a periodic property of pseudo random sequences. Journ.Ass.Comput.Mach. 5,261.

G.E.P. BOX en M.E. MULLER (1958): A note on the generation of random normal deviates. A.M.S. 29, 610.

J. CERTAINE (1958): On sequences of pseudo random numbers of maximal length. Journ.Ass.Comput.Mach. 5, 353.

R.R. COVEYOU (1960):Serial correlation in the generation of pseudo-random numbers. Journ.Ass.Comp. Mach. 7, 72.

H.J.A. DUPARC, C.G. LEKKERKERKER, A. NYENHUIS, W. PEREMANS (1952): Enige methoden om random numbers te maken. Rapport M.C. Z.W. 1952, 009.

H.J.A. DUPARC, C.G. LEKKERKERKER en W. PEREMANS (1953): Reduced sequences of integers and pseudo random numbers. Rapport M.C.Z.W. 1953, 002.

I.J. GOOD (1953): The serial test for sampling numbers and other tests for randomness. Proc.Cambridge Phil.Soc. 49, 276.

M. GREENBERGER (1961): Notes on a new pseudo-random number generator. Journ.Ass.Comp.Mach. 8, 163.

M. GREENBERGER (1961): An a priori determination of serial correlation in computer generated random numbers. Math.Comp. 15, 383.

H.C. HAMAKER (1948):Toevalscijfers, Statistica 2, 97.

J.M. HAMMERSLEY (1956): Conditional Monte Carlo. Journ.Ass. Comp.Mach. 3, 73.

J.M. HAMMERSLEY en J.G. MAULDON (1956): General principles of antithetic variates. Proc. Cambridge Phil. Soc. 52, 476.

J.M. HAMMERSLEY en K.W. MORTON (1954): Poor man's Monte Carlo. J.R.S.S. B 16, 23.

J.M. HAMMERSLEY en K.W. MORTON (1956): A new Monte Carlo technique: Antithetic variates. Proc. Cambridge Phil.Soc. 52, 449.

G.H. HARDY en E.M. WRIGHT (1945): Introduction to the theory of numbers. Oxford Univ.Press.

J. HARLING (1958): Simulation techniques in operations research - a review. J.O.R.S.A. 6, 307.

C. HASTINGS Jr. (1955): Approximations for digital computers, pag. 192. Princeton University Press.

D.L. JOHNSON (1956):Generating and testing pseudo-random numbers on the I.B.M.701. M.T.A.C. 10, 8.

H. KAHN (1956): Use of different Monte Carlo sampling techniques, in H.A. MEYER.

H. KAHN en A.W. MARSHALL (1953): Methods of reducing sample size in Monte Carlo computations. J.O.R.S.A. 1, 263.

M.G. KENDALL en B. BABINGTON SMITH (1938): Randomness and random sampling numbers. J.R.S.S. 101, 147.

M.G. KENDALL en B. BABINGTON SMITH (1939): Tables of random sampling numbers. Cambridge, Tracts for computers no. XXIV.

D.H. LEHMER (1951): Mathematical methods in large-scale computing units. 2d Symp. on large scale digital calcul. mach. Harvard Univ., 141.

A.W. MARSHALL (1956): Application of multiple stage sampling procedures to Monte Carlo problems, in H.A. MEYER.

G. MARSAGLIA (1961): Expressing a random variable in terms of uniform random variables. A.M.S. 32, 894.

G. MARSAGLIA (1961): Generating exponential random variables A.M.S. 32, 899.

H.A. MEYER (1956): Symposium on Monte Carlo methods. J. Wiley New York.

J. MOSHMAN (1954): The generation of pseudo random numbers on a decimal calculator. Journ.Ass.Comp. Mach. 1, 88.

J. MOSHMAN (1958): The application of sequential estimation to computer simulation and Monte Carlo procedures. Journ.Ass.Comput.Mach. 5, 343.

M.E. MULLER (1958): An inverse method for the generation of random normal deviates on large scale computers. M.T.A.C. 12, 167.

M.E. MULLER (1959): A comparison of methods for generating normal deviates on digital computers. Journ.Ass.Comp.Mach. 6, 376.

J. VON NEUMANN (1951): Various techniques used in connection with random digits, Monte Carlo methods. Nat.Bur. of Standards. Applied Math. Series no. 12.

P. PEACH (1961): Bias in pseudo-random numbers. J.A.S.A. 56, 610.

RANDCORPORATION: A million random digits. The Free Press, Glencoe, Illinois.

A. ROTENBERG (1960):A new pseudo-random number generator. Journ.Ass.Comp.Mach. 7, 75.

O. TAUSSKY en J. TODD (1956): Generation and testing of pseudo-random numbers, in H.A. MEYER.

D. TEICHROEW (1953):Distribution sampling with high speed computers. Ph.D. Thesis, Univ. of North Carolina.

H.F. TROTTER en J.W. TUKEY (1956): Conditional Monte Carlo for normal samples, in H.A. MEYER.

K.D. TOCHER (1954): The application of automatic computers to sampling experiments. J.R.S.S. B 16, 39.

J.G. WENDEL (1957): Groups and Conditional Monte Carlo, A.M.S. 28, 1048.

H. WOLD (1948): Random normal deviates. Cambridge, Tracts for computers no. XXV.

Appendix

Als a, b en m gehele positieve getallen zijn, $a < m$, $b < m$ en ab relatief priem met m, dan kan volgens paragraaf 2 de methode van LEHMER gebruikt worden om een rij pseudo-aselecte getallen $\{u_i\}$ te construeren:

$$u_0 = b$$
$$u_{n+1} = au_n \pmod m \qquad n=0,1,2,\ldots \tag{A.1}$$

Voorbeeld 1.

Met a=27, b=5, m=32 krijgen we de rij

$u_0=5$ $u_1=7$ $u_2=29$ $u_3=15$ $u_4=21$
$u_5=23$ $u_6=13$ $u_7=31$ $u_8=u_0=5$,

die al in paragraaf 2 besproken werd.

Over de periode van een aldus verkregen rij $\{u_i\}$ volgen hier enige stellingen. Daar deze behoren tot het terrein der getallentheorie, laten we van sommige het bewijs achterwege en verwijzen hiervoor naar: G.H. HARDY en E.M. WRIGHT (1945) en H.J.A. DUPARC, C.G. LEKKERKERKER en W. PEREMANS (1952) en (1953). Laatstgenoemden behandelen ook de periode van een ander type rijen. Verdere literatuur over de periodiciteitseigenschappen van de hier behandelde rijen geven E. en V.J. BOFINGER (1958), J. CERTAINE (1958) en J. MOSHMAN (1954).

Stelling 1.

Stel a, b en m gehele positieve getallen met $a < m$, $b < m$ en ab relatief priem met m. (Hieruit volgt dat a met m, en ook b met m relatief priem is.) Als de rij $\{u_i\}$ gedefinieerd wordt volgens (A.1) en als voor een bepaalde $n \geqq 0$ geldt:

$$u_{n+d} = u_n \tag{A.2}$$

dan is

$$a^d = 1 \pmod m \tag{A.3}$$

en aan relatie (A.2) is voldaan voor alle $n \geqq 0$.

Bewijs.

Uit (A.1) volgt: $u_1=ab \pmod m$. Stel nu $u_{n-1}=ba^{n-1} \pmod m$, dan is volgens (A.1):

$$u_n = au_{n-1} \pmod m = au_{n-1} - k_1 m = aba^{n-1} - ak_2 m - k_1 m = ba^n \pmod m \ . \qquad (A.4)$$

Hiermee is (A.4) bewezen voor alle gehele natuurlijke getallen n.

Uit (A.4) volgt verder:

$$u_{n+d} = ba^{n+d} \pmod m \qquad (A.5)$$

en uit (A.4) en (A.5):

$$\text{en} \quad \begin{matrix} m \mid ba^n - u_n \text{ (d.w.z. m deelbaar op } ba^n - u_n) \\ m \mid ba^{n+d} - u_{n+d} \ . \end{matrix} \qquad (A.6)$$

Uit (A.2) en (A.6):

$$m \mid ba^{n+d} - ba^n = ba^n(a^d - 1). \qquad (A.7)$$

Omdat m en ab, dus ook m en $a^n b$ relatief priem zijn, volgt uit (A.7):

$$m \mid a^d - 1$$

dus

$$a^d = 1 \pmod m. \qquad (A.8)$$

Omgekeerd kan uit (A.8) de relatie (A.2) afgeleid worden:

$m \mid (a^d - 1)ba^n = u_{n+d} - u_n \pmod m$ [wegens (A.4)] , dus (A.2) geldt voor alle $n \geqq 0$, q.e.d.

Uit de definitie van de periode δ volgt nu:

$$a^\delta = 1 \pmod m \qquad (A.9)$$

met als gevolg dat $u_{n+\delta} = u_n$ voor alle $n \geqq 0$ en alle keuzen van $u_0 = b$, m.a.w. de periode δ is onafhankelijk van b.

De voorwaarde dat ab en m relatief priem zijn, zullen we in het vervolg handhaven, maar niet meer expliciet vermelden.

Stelling 2.

$a^d = 1 \pmod m$ met $d > 0$ dan en slechts dan als d een veelvoud is van δ .

Bewijs.

Stel $d = g\delta + h$, met g en h geheel $\geqq 0$ en $0 \leqq h < \delta$.

Uit (A.9) volgt:

$$a^\delta = 1 + km$$

dus

$$a^{g\delta} = (1+km)^g = 1(\text{mod } m). \qquad (A.10)$$

Hieruit en uit het gegeven $a^d=1(\text{mod } m)$ volgt:

$$m \mid a^d - a^{g\delta} = a^{g\delta}(a^h-1).$$

Daar a en m relatief priem zijn, is dus

$$a^h=1(\text{mod } m) . \qquad (A.11)$$

Aangezien δ het kleinste gehele getal >0 is waarvoor (A.8) geldt, volgt uit (A.11) dat $h=0$ en $d=g\delta$, q.e.d.

Met het oog op enige volgende stellingen voeren we nu de functie van Euler $\varphi(m)$ in. $\varphi(m)$ wordt gedefinieerd als het aantal gehele positieve getallen $<m$ en relatief priem met m. In het bijzonder geldt als p een priemgetal is:

$$\varphi(p) = p-1 . \qquad (A.12)$$

Voor de functie $\varphi(m)$ gelden nu de volgende stellingen, waarvan we de eerste twee niet zullen bewijzen.

Stelling 3.

Als a en m relatief priem zijn, dan is

$$a^{\varphi(m)} = 1(\text{mod } m) \qquad (A.13)$$

Uit stelling 2 volgt dan dat δ een deler is van $\varphi(m)$.

Stelling 4.

Als m_1 en m_2 relatief priem zijn, dan geldt

$$\varphi(m_1m_2) = \varphi(m_1)\,\varphi(m_2) . \qquad (A.14)$$

Stelling 5.

Wanneer m de ontbinding in priemfactoren

$$m = p_1^{r_1} \ldots p_k^{r_k} \quad [r_i \text{ geheel } \geq 1,\ i=1,\ldots,k] \qquad (A.15)$$

bezit, dan geldt

$$\varphi(m) = m\left(1-\frac{1}{p_1}\right) \ldots \left(1-\frac{1}{p_k}\right). \qquad (A.16)$$

Bewijs:

Uit de definitie van de functie $\varphi(m)$ volgt dat

$$\varphi(p_1^{r_1}) = p_1^{r_1} - p_1^{r_1-1} = p_1^{r_1}(1-\frac{1}{p_1}) \qquad (A.17)$$

en uit (A.14) en (A.17) volgt dan (A.16), q.e d.

Steeds geldt, dat de periode δ deler is van $\varphi(m)$ (zie stelling 3). Nu volgt uit stelling 5 dat als m weinig, maar hoge priemfactoren bezit, $\varphi(m)$ groot zal zijn. Daar we volgens paragraaf 2 een rij $\{u_i\}$ met grote periode willen hebben, zal dan de gunstigste keuze van a, indien mogelijk, zodanig zijn dat $\delta = \varphi(m)$. In dat geval heet a een primitieve wortel mod m. Men kan bewijzen dat er alleen primitieve wortels mod m bestaan als m een van de vormen 2, 4, p^s, $2p^s$ (s geheel $\geqq 1$) heeft, waarin p een priemgetal >2 is. Bovendien is er in dat geval geen regel bekend die in staat stelt om een primitieve wortel te berekenen.

We duiden de bij a en m behorende periode δ verder aan met $\delta(a,m)$ en geven enige stellingen waarmee men voor $m=2^s$ en $m=10^s$ de maximale periode kan berekenen.

Stelling 6.

Als p een priemgetal >2 is en a relatief priem met p, dan is $p \mid a^{\delta(a,p)}-1$ volgens stelling 1. Is nu r de hoogste macht van p, deelbaar op $a^{\delta(a,p)}-1$ en s geheel $\geqq 1$, dan wordt $\delta(a,p^s)$ bepaald door

$$\delta(a,p^s) = \begin{cases} p^{s-r}\,\delta(a,p) & \text{als } s > r \\ \delta(a,p) & \text{als } s \leqq r. \end{cases} \qquad (A.18)$$

Voorbeeld 2.

Bij a=4 p=7 behoort $\delta(4,7)=3$. Nu is $\delta(4,49) = 7.3 = 21$, want 49 is niet deelbaar op 4^3-1.

Stelling 7.

Als a oneven en t de hoogste macht van 2 deelbaar op hetzij a+1, hetzij a-1, dan vinden we $\delta(a,2^s)$ uit:

$$\delta(a,2^s) = \begin{cases} 2^{s-t} & \text{voor } s > t \\ 2 & \text{voor } 1 < s \leqq t \text{ en } 2^t \mid a+1 \\ 1 & \text{voor } 1 < s \leqq t \text{ en } 2^t \mid a-1 \\ 1 & \text{voor } s=1\ . \end{cases} \qquad (A.19)$$

Omdat hetzij a-1, hetzij a+1 zeker deelbaar is door 4, is $t \geqq 2$, dus

$s-t \leqq s-2$, dus de maximale periode behorend bij $m=2^s$, is 2^{s-2}.

Voorbeeld 3.

We keren terug naar voorbeeld 1 en schrijven de daar gevonden rij $\{u_i\}$ in binaire vorm:

$u_0 = 0\ 0\ 1\ 0\ 1$, $u_1 = 0\ 0\ 1\ 1\ 1$, $u_2 = 1\ 1\ 1\ 0\ 1$,
$u_3 = 0\ 1\ 1\ 1\ 1$, $u_4 = 1\ 0\ 1\ 0\ 1$, $u_5 = 1\ 0\ 1\ 1\ 1$,
$u_6 = 0\ 1\ 1\ 0\ 1$, $u_7 = 1\ 1\ 1\ 1\ 1$, $u_8=u_0=0\ 0\ 1\ 0\ 1$,
(dus bijv. $u_2 = 1.2^0 + 0.2^1 + 1.2^2 + 1.2^3 + 1.2^4 = 29$).

In dit geval is s=5 en t=2, dus volgens stelling 7 is in voorbeeld 1 de periode $8=2^{5-2}$ tevens de maximale periode, behorend bij $m=32=2^5$.

Het meest rechtse cijfer in de binaire schrijfwijze van $\{u_i\}$ is steeds 1 en heeft dus de periode 1; de periode van het groepje der laatste 2,3,4,5 cijfers (van rechts geteld) is respectievelijk 2,2,4,8. Dit volgt ook uit stelling 7, want het komt erop neer dat men bij a=27 respectievelijk neemt s=1,2,3,4,5.

Alleen de rij $\{u_i\}$ zelf heeft dus de periode 2^{s-t} als $s > t$; laat men van links naar rechts telkens een cijfer van de binaire vorm der u_i weg, dan ontstaan binair geschreven getallen waarvan de periode telkens met een factor 2 afneemt, totdat het aantal cijfers in een groep tot t gedaald is. Het is dus bij dit procédé aan te bevelen om zoal niet de $\{u_i\}$ zelf, dan toch een der cijfergroepen met grote periode te gebruiken, en de minder significante groepjes (met korte periode) buiten beschouwing te laten.

Stelling 8.

Als m_1 en m_2 relatief priem zijn en $m=m_1m_2$, dan is $\delta(a,m)$ het K.G.V. van $\delta(a,m_1)$ en $\delta(a,m_2)$.

Bewijs: uit

$$m \mid a^{\delta(a\ m)_{-1}} \tag{A.20}$$

volgt

$$m_1 \mid a^{\delta(a,m)_{-1}}, \text{ en } m_2 \mid a^{\delta(a,m)_{-1}}. \tag{A.21}$$

Volgens stelling 2 is $\delta(a,m)$ dus een gemeen veelvoud van $\delta(a,m_1)$ en $\delta(a,m_2)$. Zij nu v een willekeurig gemeen veelvoud van $\delta(a,m_1)$

en $\delta(a,m_2)$, dan is volgens stelling 2 $m_1 \mid a^v-1$ en $m_2 \mid a^v-1$, dus $m_1 m_2 \mid a^v-1$, want m_1 en m_2 zijn relatief priem. Dus $\delta(a,m)$ is een deler van v, q.e.d.

Voor $\delta(a,10^s)$ volgt hieruit:

$$\delta(a,10^s) = \text{K.G.V.}\ (\delta(a,2^s),\ \delta(a,5^s)) \ . \qquad (A.22)$$

Als r de hoogste macht van 5 deelbaar op $a^{\delta(a,5)}-1$ en t de hoogste macht van 2, deelbaar op $a\pm1$ is, terwijl bovendien $s > r,t$ dan volgt uit de stellingen 6 en 7 en (A.22):

$$\delta(a,10^s) = \text{K.G.V.}(2^{s-t},5^{s-r}\ \delta(a,5)) \ . \qquad (A.23)$$

Uit stelling 3 volgt dat $\delta(a,5)$ deler is van $4=\varphi(5)$. Stellen we dus:

$$\delta(a,5) = 2^i \qquad (i = 0,1,2) \qquad (A.24)$$

dan volgt uit (A.23) en (A.24) voor $s > r$:

$$\delta(a,10^s) = \begin{cases} 5^{s-r}\ 2^{s-t} & \text{als } s-t \geqq i \\ 5^{s-r}\ 2^{i} & \text{als } s-t < i \ . \end{cases} \qquad (A.25)$$

De periode behorend bij 10^s zal dus maximaal zijn als r=1, t=2 en $\delta(a,5)=4$, dus i=2.

Volgens (A.25) vinden we dan voor deze maximale periode:

$$\begin{aligned} 5^{s-1}.2^{s-2} &= 5.10^{s-2} && \text{als } s \geqq 4 \\ 5^2.2^2 &= 100 && \text{als } s = 3 \\ 5.2^2 &= 20 && \text{als } s = 2 \\ 2^2 &= 4 && \text{als } s = 1. \end{aligned} \qquad (A.26)$$

Men ziet gemakkelijk in, dat a=3 aan de bovengenoemde voorwaarden voldoet. In decimale vorm geschreven, bestaan nu de getallen $\{u_i\}$ uit s cijfers, als $m=10^s$. Volgens (A.26) zal hierbij het laatste cijfer de periode 4 hebben, de laatste twee cijfers tezamen de periode 20, enz. Immers de laatste k cijfers vormen de rest bij deling van de getallen $\{u_i\}$ door 10^k. Wil men van de minder significante cijfergroepjes met korte periode geen gebruik maken en is men alleen geïnteresseerd in waarden $s \geqq 4$, dan is het niet nodig a zo te kiezen dat $\delta(a,5)=4$. Bijv. a=11 voldoet eveneens aan de voorwaarden r=1 en t=2 en komt dus ook in aanmerking.

Wanneer behalve $a < m$, ook $a^2 < m$, $a^3 < m, \ldots, a^k < m$, dan zal als

$u_n < a$, $u_n < u_{n+1} < \ldots < u_{n+k-1} < m$. In de rij $\{u_i\}$ treedt dan een regelmatigheid op met als gevolg dat de rij $\{m^{-1}u_i\}$ dan niet meer als een rij aselecte getallen beschouwd kan worden. Opdat $\{u_i\}$ deze regelmatigheid niet zal vertonen, moet a dus groot gekozen worden. Bij het zoeken naar geschikte grote waarden van a kan de volgende stelling gebruikt worden.

Stelling 9.

Als de G.G.D. van $\delta = \delta(a,m)$ en n gelijk is aan g, dan is

$$\delta(a^n,m) = \frac{\delta(a,m)}{g} \quad . \tag{A.27}$$

Bewijs:

Uit (A.9) en stelling 2 volgt:

$$(a^n)^{\frac{\delta}{g}} = a^{\frac{n}{g}\delta} = a^{k\delta} = 1 \pmod m,$$

dus $\delta(a^n,m) \mid \frac{\delta(a,m)}{g}$. (A.28)

Verder is

$(a^n)^{\delta(a^n,m)} = a^{n.\,\delta(a^n,m)} = 1 \pmod m$, dus $\delta(a,m) \mid n.\,\delta(a^n,m)$.

Hieruit volgt:

$$\frac{\delta(a,m)}{g} \mid \frac{n}{g}\ \delta(a^n,m) \quad . \tag{A.29}$$

$\frac{\delta(a,m)}{g}$ en $\frac{n}{g}$ zijn geheel en bevatten geen gemeenschappelijke factoren, dus

$$\frac{\delta(a,m)}{g} \mid \delta(a^n,m) \quad . \tag{A.30}$$

Uit (A.28) en (A.30) volgt dan het gestelde in (A.27).

Het is gemakkelijk, althans bij een m van de vorm $m=10^s$, een kleine a te vinden waarbij de maximale periode behoort. Stelling 9 kan ons dan een grote a met deze maximale periode geven. Bij $m=10^{20}$ bijv. behoort volgens (A.26) de maximumperiode 5.10^{18}, die zoals we zagen wordt voortgebracht door a=3. Volgens stelling 9 brengt dan ook 3^{19} deze maximumperiode voort.

In één periode zal een element van de rij $\{u_i\}$ eenmaal voor-

komen. We gaan nu na wat er gebeurt als men in de decimale vorm van elke u_i de laatste cijfers weglaat. Als $m=10^{k+j}$, dan bestaan alle u_i uit k+j cijfers. Is $j \geq 4$, dan bedraagt volgens (A.26) de bij 10^j behorende periode $\lambda = 5.10^{j-2}$ en de bij 10^{k+j} behorende periode $5.10^{k+j-2} = 10^k . \lambda$.

Nu zijn er in het decimale stelsel 10^k verschillende getallen van k cijfers, nl. $0,1,2,\ldots,10^k-1$. Elk getal u_i bestaat uit één dezer getallen, gevolgd door een staartgroepje van j cijfers. Laten we omgekeerd elk getal uit $0,1,\ldots,10^{k-1}$ volgen door elk van de λ verschillende in $\{u_i\}$ bevatte staartgroepjes van j cijfers, dan ontstaan $10^k.\lambda$ verschillende getallen van k+j cijfers. Dit moeten juist alle u_i zijn, daar anders het aantal verschillende $u_i < 10^k.\lambda$ zou zijn. Dus elk getal $0,1,\ldots,10^{k-1}$ komt als beginstuk van k cijfers λ keer voor in één periode van de rij $\{u_i\}$.

In paragraaf 2 werd ook een andere methode besproken om pseudo-aselecte getallen voort te brengen, namelijk de gemengde congruentie methode. We zullen deze appendix besluiten met een stelling over de periode van de met behulp van deze methode voort gebrachte rij $\{v_i\}$.

Stelling 10.

Als $\lambda < 2^s$, $\mu < 2^s$ en $s > 1$ is, dan is de periode van de rij $\{v_i\}$ gedefinieerd door

$$\begin{aligned} v_0 &= \rho \\ v_{n+1} &= \lambda v_n + \mu \pmod{2^s}, \quad n = 0,1,2,\ldots \end{aligned} \tag{A.31}$$

dan en slechts dan gelijk aan 2^s, als μ oneven is en $\lambda = 1 \pmod 4$.

Bewijs.

Het is eenvoudig in te zien dat λ en μ beide oneven moeten zijn, daar anders alle $v_i (i=1,2,\ldots)$ òf oneven òf even zijn.
Herhaald toepassen van (A.31) geeft

$$v_{n+k} = \lambda^k v_n + \frac{(\lambda^k - 1)\mu}{\lambda - 1} \pmod{2^s} . \tag{A.32}$$

Zij δ de periode van de rij $\{v_i\}$, dit is het kleinste natuurlijk getal d dat voldoet aan $v_n = v_{n+d}$. Dan volgt uit (A.32)

$$0 = \frac{\lambda^d - 1}{\lambda - 1}\{(\lambda - 1)v_n + \mu\} \pmod{2^s} \qquad (A.33)$$

Daar λ en μ oneven zijn, is (A.33) equivalent met

$$0 = \frac{\lambda^d - 1}{\lambda - 1} \pmod{2^s}. \qquad (A.34)$$

Zij 2^b de grootste macht van 2, die als faktor van $\lambda - 1$ optreedt ($b \geq 1$), dan is (A.34) ook te schrijven als

$$1 = \lambda^d \pmod{2^{s+b}}. \qquad (A.35)$$

Daar λ en 2^{s+b} relatief priem zijn, volgt uit stelling 3 dat $d = \varphi(2^{s+b})$ een oplossing van (A.35) is, terwijl δ, dit is de kleinste d die aan (A.35) voldoet, een deler is van $\varphi(2^{s+b})$. Uit stelling 5 volgt dat

$$\varphi(2^{s+b}) = 2^{s+b-1}, \qquad (A.36)$$

zodat

$$\delta \mid 2^{s+b-1}. \qquad (A.37)$$

We kunnen δ dus schrijven als 2^ν met $\nu \leq s+b-1$.
Stel $\lambda = \kappa + 1$, dan is $\kappa = c2^b$ met c oneven. (A.35) is nu te schrijven als

$$1 = 1 + \sum_{j=1}^{d} \binom{d}{j} \kappa^j \pmod{2^{s+b}},$$

zodat δ moet voldoen aan

$$0 = \sum_{j=1}^{2^\nu} \binom{2^\nu}{j}(2^b c)^j = 2^{\nu+b} c \sum_{j=0}^{2^\nu - 1} \binom{2^\nu - 1}{j} \frac{2^j}{j+1} (2^{b-1}c)^j \pmod{2^{s+b}}. \qquad (A.38)$$

Beschouw de uitdrukking

$$f(j) = \binom{2^\nu - 1}{j} \frac{2^j}{j+1} = \frac{2^j(2^\nu - 1)(2^\nu - 2)\ldots(2^\nu - j)}{(j+1)!} \qquad (j=0,1,\ldots). \qquad (A.39)$$

We zullen bewijzen, dat $f(j)$ geheel is voor $j=0,1,\ldots,2^\nu-1$. Triviaal is $f(0)=1$ en $f(1)=2^\nu-1$. Bovendien is $\binom{2^\nu}{j+1}$ een geheel getal, zodat ook

$$(j+1)! \mid 2^\nu(2^\nu-1)\ldots(2^\nu-j). \qquad (A.40)$$

Schrijf $(j+1)! = 2^p.q$, waarin q oneven is (dit is een eenduidige schrijfwijze).
Dan volgt uit (A.40), dat

$$q \mid (2^\nu-1)(2^\nu-2)\ldots(2^\nu-j). \qquad (A.41)$$

Aangezien voor $j \geq 2$ het product $(2^\nu-1)(2^\nu-2)\ldots(2^\nu-j)$ minstens één faktor 2 bevat, is het voldoende aan te tonen, dat $2^p \leq 2^{j+1}$ is, of equivalent $p \leq j+1$.
Men verifieert eenvoudig, dat $p=\left[\frac{j+1}{2}\right] + \left[\frac{j+1}{4}\right]+\left[\frac{j+1}{8}\right]+\ldots$, waarin $[x]$ het grootste gehele getal $\leq x$ voorstelt. Hieruit volgt:

$$p \leq \frac{j+1}{2} + \frac{j+1}{4} + \frac{j+1}{8} +\ldots = j+1,$$

waarmee bewezen is dat $f(j)$ geheel is.
Zij nu $b \geq 2$. Dan is $(2^{b-1}c)^j$ even voor $j \geq 1$, zodat de termen van de in het rechterlid van (A.38) optredende som alle even zijn met uitzondering van de eerste term (met $j=0$), die 1 is. Dan is de som zelf dus oneven, en we kunnen (A.38) schrijven als

$$0 = 2^{\nu+b} \pmod{2^{s+b}} \quad \text{mits } b \geq 2. \qquad (A.42)$$

Hieruit volgt direct, dat $\nu = s$. De voorwaarde $b \geq 2$ is equivalent met $\lambda = 1 \pmod 4$, zodat we bewezen hebben dat de rij $\{v_i\}$ inderdaad de maximale periode $\delta = 2^s$ heeft als μ oneven en $\lambda = 1 \pmod 4$ is.

We zullen nog aantonen, dat als $b=1$, d.w.z. $\lambda = 3 \pmod 4$, de periode δ van de rij kleiner is dan 2^s. Stel $b=1$ en $\delta = 2^s$. Dan is dus $\delta = \varphi(2^{s+b})$ volgens (A.36) en (A.37). Uit (A.35) volgt dan, dat λ een primitieve wortel mod 2^{s+b} is, zie pag. A4. Dit is echter onmogelijk, daar 2^{s+b} geen primitieve wortels bezit, zoals we al eerder opmerkten (mits $s > 1$). Een tegenspraak is dus bereikt, zodat $\delta < 2^s$ moet zijn, q.e.d.