MCSYLLABUS 1.8

J. KRIENS
F. GöBEL
W. MOLENAAR

LEERGANG BESLISKUNDE

DEEL 8

MINIMAXMETHODE

NETWERKPLANNING

SIMULATIE

MATHEMATISCH CENTRUM AMSTERDAM 1975

AMS(MOS) subject classification scheme (1970): 90A05, 90D05 / 90C35 / 62E25,
65C10, 90B99

ISBN 90 6196 034 7

1e druk: 1968
2e druk: 1972
3e druk: 1975

Inhoud

Voorwoord

Nadat in de deeltjes 1.6 en 1.7 van deze Leergang Besliskunde een aantal algemene technieken van de besliskunde behandeld zijn, volgen in deze aflevering drie los van elkaar staande, meer gespecialiseerde technieken. De afwezigheid van enige samenhang tussen de technieken komt ook tot uiting in de samenstelling van de tekst.

Het eerste hoofdstuk, dat handelt over de minimaxmethode, werd geschreven door Prof.J. KRIENS, die Drs.J. VAN LIESHOUT, medewerker aan de Katholieke Hogeschool te Tilburg, dank verschuldigd is voor talrijke suggesties ter verduidelijking van de inhoud. Het tweede hoofdstuk, getiteld Netwerkplanning, is van de hand van Drs.F. GÖBEL en het derde hoofdstuk van dit deeltje, dat gewijd is aan Simulatie, werd geschreven door Drs.W. MOLENAAR. De eindredactie van alle drie in dit deeltje behandelde onderwerpen werd verzorgd door Drs.B. DORHOUT.

Nu bij het verschijnen van dit deeltje de syllabusreeks Leergang Besliskunde een voorlopige afronding krijgt, willen de samenstellers eenieder die aan de totstandkoming heeft meegewerkt van harte bedanken. In het bijzonder zijn zij erkentelijk voor de redactionele werkzaamheden van de heer J. HILLEBRAND, het typewerk van mevr.R.M. BRUINS-WITKAMP, mevr.S.J. KUIPERS-HOEKSTRA en mevr.H. ROQUÉ-DE HOYER en de technische verzorging door de heren D. ZWARST en J. SUIKER.

DEEL 8

HOOFDSTUK 1

MINIMAXMETHODE DOOR J. KRIENS

1. Optimaliseren van verwachtingen.

Bij zeer veel besliskundige problemen is het gebruikelijk om dat alternatief te kiezen, waarvoor de verwachting van de opbrengsten maximaal is, resp. de verwachting van de verliezen minimaal.

In werkelijkheid wil men meestal de opbrengst zelf maximaliseren als funktie van één of meer grootheden, die men in de hand heeft. De relatie tussen de opbrengst en deze grootheden, de beslissingsvariabelen, is echter in veel gevallen niet deterministisch vastgelegd; wel is dan vaak bekend dat de opbrengst een kansverdeling volgt, waarin de beslissingsvariabelen als parameters voorkomen. Een dergelijke relatie is echter op zichzelf genomen niet geschikt om de keuze te doen en daarom leidt men uit de kansverdeling een andere grootheid af, in veel gevallen de verwachting, waarmee de keuze wel gedaan kan worden (vgl. o.a. deel 5, §6). In enkele gevallen neemt men ook wel de variantie of een andere grootheid als kriteriumfunktie (vgl. deel 6, voorbeeld 1.6).

Het kiezen van de verwachting als kriterium kan in veel gevallen gerechtvaardigd worden met behulp van limietstellingen, waarvan stelling 12.2 uit deel 2 een voorbeeld is, dat wij hier herhalen.

Stelling 1.1

Als $\underline{y}_1$, $\underline{y}_2$, ... onderling onafhankelijk verdeelde stochastische grootheden zijn, die alle dezelfde verwachting μ en variantie σ^2 bezitten, dan convergeert voor $n \to \infty$ het gemiddelde

$$\overline{\underline{y}}_n \overset{\text{def}}{=} \frac{1}{n} \sum_{i=1}^{n} \underline{y}_i \tag{1.1}$$

stochastisch naar μ; d.w.z. dat voor iedere $\varepsilon > 0$ geldt

$$\lim_{n \to \infty} P\left[\left| \overline{\underline{y}}_n - \mu \right| > \varepsilon \right] = 0. \tag{1.2}$$

Er bestaan nog tal van andere dergelijke stellingen, waarvan de inhoud ruw geformuleerd hierop neerkomt, dat onder bepaalde voorwaarden het gemiddelde van een aantal stochastische grootheden een grote kans heeft dicht te liggen bij het gemiddelde van de verwachtingen van deze variabelen. Wanneer men nu een gemiddelde $\frac{1}{n} \sum_{i=1}^{n} \underline{y}_i$ voor grote n wil maximaliseren, dan komt dit in deze gevallen vrijwel steeds op hetzelfde neer als het maximaliseren van het gemiddelde van de verwachtingen, dus van $\frac{1}{n} \sum_{i=1}^{n} \mathcal{E} \underline{y}_i$ (vgl. ook deel 4, §7 en §8).

Hangen vervolgens de verwachtingen $\mathcal{E} \underline{y}_i$ alle op dezelfde wijze af van de beslissingsvariabelen, dan wordt het maximum van $\frac{1}{n} \sum_{i=1}^{n} \mathcal{E} \underline{y}_i$ gevonden door de verwachting van $\underline{y}_i$ zelf, dus $\mathcal{E} \underline{y}_i$, te maximaliseren. In het geval, waarin de $\underline{y}_i$ alle van verschillende beslissingsvariabelen afhangen, is het maximaliseren van $\frac{1}{n} \sum_{i=1}^{n} \mathcal{E} \underline{y}_i$ eveneens gelijkwaardig aan het maximaliseren van alle verwachtingen afzonderlijk.

Bovenstaande resultaten kunnen ook als volgt geformuleerd worden: als men een groot aantal beslissingen moet nemen en men zich hierbij steeds baseert op de verwachting, dan kan bij iedere beslissing afzonderlijk de verwachting aanzienlijk afwijken van het uiteindelijke resultaat, doch voor alle beslissingen tezamen zal de einduitkomst weinig van de verwachting verschillen. De voorwaarden, waaraan bij de limietstellingen moet zijn voldaan, impliceren, dat er een voldoende aantal verwachtingen van vergelijkbare grootte dient te zijn.

De hier gegeven fundering om verwachtingen te optimaliseren is dus een fundering o.a. voor het geval, waarin we vele malen achtereen een beslissing willen nemen en voor het geval, waarin we een groot aantal beslissingen tegelijkertijd moeten nemen. Hoewel men, zoals gezegd, veelal op deze wijze te werk gaat, dient men toch voorzichtig te zijn, aangezien aan deze methode wel bezwaren verbonden kunnen zijn.

Zo kan men kritiek uitoefenen op het optimaliseren van winstverwachtingen in situaties, waarin limietstellingen geen betekenis hebben. Hoofdzakelijk is dit het geval, wanneer slechts één of een be-

perkt aantal beslissing(en) genomen moet(en) worden, omda dan in het geheel niet behoeft te gelden, dat bijvoorbeeld $\frac{1}{3}(\underline{y}_1 + \underline{y}_2 + \underline{y}_3)$ met grote kans dicht bij de verwachtingswaarde $\frac{1}{3}(\mathcal{E}\,\underline{y}_1 + \mathcal{E}\,\underline{y}_2 + \mathcal{E}\,\underline{y}_3)$ zal liggen. Men kan dan trachten een andere fundering te geven, hetgeen onder andere gedaan is door VON NEUMANN en MORGENSTERN *).

Bij deze funderingen wordt uitgegaan van een bepaalde consistentie in de voorkeuren van degene die moet kiezen. De formulering van deze consistentie geschiedt door middel van axiomastelsels, die bij verschillende auteurs in details uiteenlopen, doch waarvan de belangrijkste axioma's vrijwel overeenkomen. De axioma's beschrijven het gedrag van de mensen bij de keuze tussen alternatieven. Uitgaande van deze axioma's kan men een nutsfunktie opstellen met de eigenschap, dat het maximaliseren van de verwachting van het volgens deze funktie te verwerven nut precies leidt tot het aanwijzen van dat alternatief, waaraan men de voorkeur zou geven indien men de axioma's volgt. Een uitstekend overzicht van deze theorieën is gegeven door LUCE en RAIFFA **). Het belangrijkste bezwaar, dat men tegen deze argumentatie om verwachtingen te optimaliseren kan maken, is, dat in de praktijk het menselijk handelen niet steeds aan de gestelde axioma's voldoet. Wij zullen hier niet nader op ingaan.

Ter afsluiting van deze paragraaf merken wij op, dat het optimaliseren van verwachtingen goed te verdedigen valt wanneer

1) het gaat om een groot aantal beslissingen waarvan de uitkomsten min of meer vergelijkbaar zijn;
2) de kansen door middel van ervaring of experimenten goed geschat kunnen worden;
3) de voor- en nadelen van de verschillende alternatieven in dezelfde eenheid (bijv. de gulden) uitgedrukt kunnen worden.

---

*) J. VON NEUMANN en O. MORGENSTERN, Theory of Games and Economic Behaviour, Princeton University Press, Princeton, appendix, (1944 , 1ste druk) , (1947 ; 2de druk).

**) R.D. LUCE en H. RAIFFA, Games and Decisions, John Wiley and Sons, New York,(1957).

In sommige gevallen kan het maximaliseren van verwachtingen dwingen tot het lopen van grote risico's. Hiertegen kan men zich veilig stellen door het invoeren van één of meer bijvoorwaarden, zoals bijvoorbeeld gedaan is in voorbeeld 6.4 van deel 5.

Wanneer niet voldaan is aan voorwaarde 2) is het maximaliseren van een verwachte opbrengst niet mogelijk en is men genoodzaakt een ander kriterium te gebruiken. Hiervoor kan het zogenaamde minimax-kriterium in aanmerking komen, welk kriterium wij in de volgende paragraaf zullen illustreren aan de hand van de strategische spel-theorie.

2. De strategische speltheorie.

Het is niet mogelijk alle aspecten van de theorie der strategische spelen in kort bestek te belichten. Wij zullen ons daarom beperken tot het invoeren van enkele begrippen, het noemen van een paar belangrijke stellingen en het geven van een aantal voorbeelden.

Naast de zuivere kansspelen bestaan er spelen, waarbij de spelers zelf invloed kunnen uitoefenen op het resultaat, zoals bridge en het schaakspel. Dit type spelen noemt men strategische spelen. Zij bezitten een grote overeenkomst met allerlei situaties uit het dagelijks leven, waarin twee of meer partijen met elkaar in conflict zijn.

Voorbeeld 2.1

Een zeer eenvoudig strategisch spel is het spel, waarbij twee personen onafhankelijk van elkaar een getal aangeven en waarbij het van de combinatie van deze getallen afhangt hoeveel de één aan de ander moet betalen. Bij iedere combinatie van getallen is de som van de bedragen, die beide spelers ontvangen, dus nul; een spel waarvoor deze eigenschap geldt en waaraan, zoals in dit geval, twee personen deelnemen, is een twee personen nulspel. Laten wij aannemen, dat de ene speler (I) kan kiezen uit de getallen 1, 2 en 3 en de andere (II)

uit de getallen 1, 2, 3 en 4. Tabel 2.1 geeft aan hoeveel guldens II na afloop aan I moet betalen. Wijst I dus het getal 2 aan en II het getal 3, dan ontvangt I na afloop *f* 1,-; wijst I echter eveneens het getal 3 aan, dan ontvangt hij —*f* 1,-, m.a.w. hij moet II *f* 1,- betalen; enz.

Tabel 2.1

De winstfunktie

I \ II	1	2	3	4
1	3	5	0	-5
2	4	2	1	2
3	-2	0	-1	3

Iedere speelwijze die I of II kan kiezen, noemen wij een strategie *). Wij geven de strategieën aan met x voor I en met z voor II; in het voorbeeld kan x dus de waarden 1, 2 en 3 aannemen en z de waarden 1, 2, 3 en 4. De funktie, die aangeeft hoeveel II na afloop aan I moet betalen, noemen wij de winstfunktie; de notatie is y(x,z); zo is y(1,4) = -5. Een spel, waarin beide spelers slechts uit eindig veel strategieën kunnen kiezen, is een eindig of rechthoekig spel.

De vraag rijst nu, of het mogelijk is redelijke strategieën te vinden voor de deelnemers.

Wanneer I strategie 1 speelt, zal hij, afhankelijk van de keuze welke II doet, 3, 5, 0 of -5 ontvangen. De slechtste uitkomst voor I is dan -5. Als hij het getal 2 kiest, is de slechtste uitkomst +1 en, als hij het getal 3 kiest, -2. Nu kan I besluiten dié strategie uit te voeren, waarbij het bedrag dat hij minimaal ontvangt zo groot mogelijk is; hij moet dan het getal 2 aanwijzen. In het algemeen kan men deze strategie vinden door eerst voor iedere waarde van x het

*) In verband met het later in deze paragraaf in te voeren begrip gemengde strategie worden de hier bedoelde strategieën ook wel zuivere strategieën genoemd.

minimum van y(x,z) te berekenen als funktie van z en vervolgens te bepalen voor welke waarde van x de funktie $\min_z y(x,z)$ maximaal is.

Voor II geeft tabel 2.1 niet de opbrengst van het spel aan, maar het bedrag, dat hij moet betalen. Uitgaande van deze tabel kan hij trachten dit bedrag zo klein mogelijk te maken. Hij berekent dan voor ieder van de getallen 1, 2, 3 en 4 de maxima van de eventueel te betalen bedragen en kiest dan dié strategie, waarvoor dit maximum minimaal is; met andere woorden hij kiest dié strategie, waarvoor $\max_x y(x,z)$ minimaal is. In ons geval is $\max_x y(x,1) = 4$, $\max_x y(x,2) = 5$, $\max_x y(x,3) = 1$ en $\max_x y(x,4) = 3$; het minimum van deze maxima is dus 1 en het wordt bereikt voor z = 3.

Men noemt deze methode, waarbij men het maximaal mogelijke verlies minimaliseert de minimaxmethode. Zou ook voor II de winstfunktie gegeven zijn, dan zouden alle bedragen in de bijbehorende tabel de tegengestelden zijn van de bedragen in tabel 2.1. Gaat II uit van deze nieuwe tabel en zoekt hij, evenals I in het bovenstaande, de strategie, die zijn minimaal mogelijke winst maximaliseert, dan leidt dit tot hetzelfde resultaat als toepassing van de minimaxmethode op tabel 2.1. Hoewel de voor de spelers aangegeven methoden om geschikte strategieën te vinden dus verschillend lijken, zijn ze in feite identiek.

In het voorbeeld is $\max_x \min_z y(x,z) = 1$ en $\min_z \max_x y(x,z) = 1$. Speler I kan er dus voor zorgen, dat hij minstens *f* 1,- krijgt, onafhankelijk van hetgeen speler II doet. Evenzo kan speler II voorkomen, dat hij meer dan *f* 1,- moet betalen, onafhankelijk van wat I doet. Ook dit laatste is speler I bekend en hij weet dus dat er geen strategie kan bestaan, die hem een bedrag groter dan *f* 1,- garandeert. Anderzijds is er wél een strategie, waarbij hij het bedrag van *f* 1,- in ieder geval ontvangt en daarom kan hij besluiten het zekere voor het onzekere te nemen en strategie 2 spelen. Voor II gelden soortgelijke overwegingen, zodat men ook voor hem het spelen van de minimaxstrategie kan verdedigen.

Voor spelen, waarvan de winstfunktie $y(x,z)$ de eigenschap bezit, dat

$$\max_x \min_z y(x,z) = \min_z \max_x y(x,z), \tag{2.1}$$

kan men op de hierboven aangegeven wijze dus redelijke strategieën voor de spelers construeren. Wij noemen de corresponderende strategieën optimale strategieën *) van I en II en het door (2.1) bepaalde bedrag de waarde van het spel.

Het volgende voorbeeld illustreert dat niet alle winstfunkties aan de relatie (2.1) voldoen.

Voorbeeld 2.2. Het kruis- of muntspel.

In het kruis- of muntspel leggen de spelers I en II onafhankelijk van elkaar een munt op tafel. Blijken beide munten met dezelfde kant boven te liggen, dan ontvangt I *f* 1,- van II, terwijl in het andere geval II *f* 1,- van I ontvangt. Beide spelers kunnen in dit spel uit twee strategieën kiezen n.l. kruis of munt bovenleggen. De winstfunktie is vermeld in tabel 2.2.

Tabel 2.2

Winstfunktie in het kruis- of muntspel

I \ II	K	M
K	1	-1
M	-1	1

In dit spel is $\max_x \min_z y(x,z) = -1$ en $\min_z \max_x y(x,z) = +1$ en er geldt dus

$$\max_x \min_z y(x,z) < \min_z \max_x y(x,z). \tag{2.2}$$

---

*) De betekenis van het woord "optimaal" wijkt hier dus af van die, waarin het elders in deze leergang wordt gebruikt.

De vorm van de ongelijkheid (2.2) is niet toevallig, maar het gevolg van een stelling uit de wiskunde, die luidt:

Stelling 2.1

Zijn $\mathcal{X}$ en $\mathcal{Z}$ twee gegeven verzamelingen, is $f(x,z)$ gedefinieerd voor $x \in \mathcal{X}$ en $z \in \mathcal{Z}$ en bestaan $\max_x \min_z f(x,z)$ en $\min_z \max_x f(x,z)$, dan is

$$\max_x \min_z f(x,z) \leq \min_z \max_x f(x,z). \tag{2.3}$$

Bewijs:

Voor iedere x uit $\mathcal{X}$ en iedere z uit $\mathcal{Z}$ geldt

$$\min_z f(x,z) \leq f(x,z) \leq \max_x f(x,z). \tag{2.4}$$

Vergelijken wij nu $\min_z f(x,z)$ en $\max_x f(x,z)$, dan zien wij, dat in

$$\min_z f(x,z) \leq \max_x f(x,z) \tag{2.5}$$

het linkerlid onafhankelijk is van z en het rechterlid onafhankelijk van x. Bovendien geldt (2.5) voor iedere waarde van x en iedere waarde van z. Deze ongelijkheid blijft daarom juist als wij in het linkerlid het maximum over x en in het rechterlid het minimum over z nemen en dus geldt

$$\max_x \min_z f(x,z) \leq \min_z \max_x f(x,z).$$

Beperken wij ons tot eindige spelen, dan bevatten de verzamelingen $\mathcal{X}$ en $\mathcal{Z}$ van de strategieën van I en II beide eindig veel elementen en kan men laten zien dat de max min en de min max van de winstfunktie altijd bestaan. Toepassing van stelling 2.1 leidt dus voor ieder eindig spel tot de eigenschap

$$\max_x \min_z y(x,z) \leq \min_z \max_x y(x,z), \tag{2.3'}$$

of, als

$$\lambda_G \overset{\text{def}}{=} \max_x \min_z y(x,z) \tag{2.6}$$

en

(2.7) $$\nu_G \overset{\text{def}}{=} \min_z \max_x y(x,z),$$

tot

(2.8) $$\lambda_G \leq \nu_G.$$

Dat aan deze ongelijkheid moet zijn voldaan, volgt ook uit de interpretatie die wij in voorbeeld 2.1 aan λ_G en ν_G hebben gegeven en die wordt geïllustreerd in figuur 2.1: I kan er voor zorgen, dat hij minstens λ_G ontvangt; II kan zijn strategie zó kiezen, dat de wins van I in ieder geval kleiner dan of gelijk aan ν_G is.

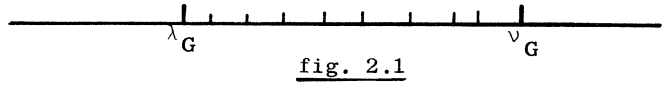


fig. 2.1

Bedrag, dat I ontvangt

Wanneer nu $\lambda_G < \nu_G$ is en het voor I dus niet vaststaat, dat II hem kan beletten meer dan λ_G te winnen, is nog niet in te zien, dat de max min strategie voor I een aanvaardbare strategie is. Hetzelfde geldt voor de min max strategie voor II. Voor dit type spelen hebben wij dus nog geen oplossing gevonden.

Om de later te geven oplossing voor te bereiden bekijken wij eerst een variatie op het kruis- of muntspel.

Voorbeeld 2.3. Het kruis- of muntspel met spioneren.

In dit spel moet eerst I zijn munt neerleggen en II pas, nadat hij gezien heeft, welke strategie I heeft gekozen. I kan weer kiezen uit twee strategieën, doch II heeft meer mogelijkheden gekregen, omdat hij zijn strategie kan laten afhangen van hetgeen I heeft gedaan. Geven wij kruis aan met K, munt met M en stelt (i,j) de strategie voor van II, waarbij hij i kiest als I kruis boven legt en j als I munt boven legt, dan kan II kiezen uit de strategieën:

(K,K); (K,M); (M,K); (M,M).

Komen er gelijke zijden boven te liggen, dan ontvangt I f 1,- van II en anders II f 1,- van I. Tabel 2.3 bevat de winstfunktie van dit spel. Hier geldt wel $\max_x \min_z y(x,z) = \min_z \max_x y(x,z)$; de waarde van het spel is -1.

Tabel 2.3

Winstfunktie in het kruis- of muntspel met spioneren

I \ II	(K,K)	(K,M)	(M,K)	(M,M)
K	1	1	-1	-1
M	-1	1	-1	1

In vergelijking met het gewone kruis- of muntspel is de situatie voor I slechter geworden, want II kan zó spelen, dat I steeds f 1,- aan II moet betalen. Dit komt uiteraard doordat II volledig is ingelicht over hetgeen I heeft gedaan.

Ook in het kruis- of muntspel zonder spioneren loopt I echter het risico, dat II in een reeks van spelen de speelwijze, die I toepast, doorziet en dan kan II steeds zijn winst vergroten door het tegenovergestelde te doen van hetgeen I speelt. Hetzelfde risico loopt ook de tweede speler. Het is daarom voor beide spelers van belang hun speelwijze voor de ander te verbergen. Zij kunnen dit doen door in achtereenvolgende spelen op niet-systematische wijze K of M boven te leggen. Een hulpmiddel hiertoe is een kansmechanisme, dat voor ieder afzonderlijk spel met een bepaalde kans aangeeft of K, dan wel M gespeeld moet worden. De spelers kiezen dan niet meer een strategie voor ieder spel afzonderlijk , doch zij kiezen de kans, waarmee de in een bepaald spel te spelen strategie wordt aangewezen.

In het volgende voorbeeld wordt deze gedachtengang uitgewerkt voor het in voorbeeld 2.2 genoemde kruis- of muntspel.

Voorbeeld 2.4. Het kruis- of muntspel met gemengde strategieën.

Stel dat I in het kruis- of muntspel een zodanig kansmechanisme kiest, dat de strategie K met kans a en de strategie M dus met kans 1-a optreedt en dat II een kansmechanisme kiest, zodanig dat de strategie K met kans b en de strategie M dus met kans 1-b optreedt. Aangezien I en II de strategieën onafhankelijk van elkaar kiezen, bedraagt de kans dat beide K kiezen ab, de kans dat I K en II M kiest a(1-b), de kans dat I M en II K kiest (1-a)b en de kans dat beide M kiezen (1-a)(1-b). Als de winstfunktie per spel de in tabel 2.2 vermelde is, dan is de winstverwachting voor I dus

$$y^*(a,b) = ab + (1-a)(1-b) - a(1-b) - b(1-a). \tag{2.9}$$

Ook in dit spel kan I nagaan hoe groot de winstverwachting is, die hij minimaal heeft voor een bepaalde a en dan a zodanig kiezen dat deze minimale winstverwachting maximaal is. Hij moet dan berekenen voor welke waarde van a de funktie $\min_b y^*(a,b)$ het maximum bereikt. Nu is

$$\min_b y^*(a,b) = \min_b \left[1-2a+2b(2a-1)\right] = \begin{cases} 2a-1 \text{ voor } 2a-1 \leq 0 \\ 1-2a \text{ voor } 2a-1 \geq 0 \end{cases} \tag{2.10}$$

en dus is

$$\max_a \min_b y^*(a,b) = 0; \tag{2.11}$$

een maximum dat bereikt wordt voor $a = \frac{1}{2}$ (vgl. fig. 2.2a).

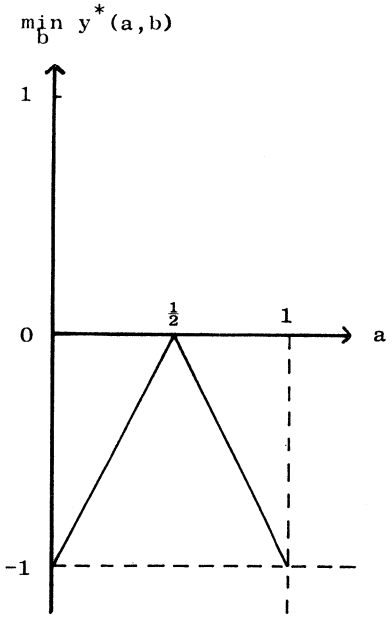


fig. 2.2a

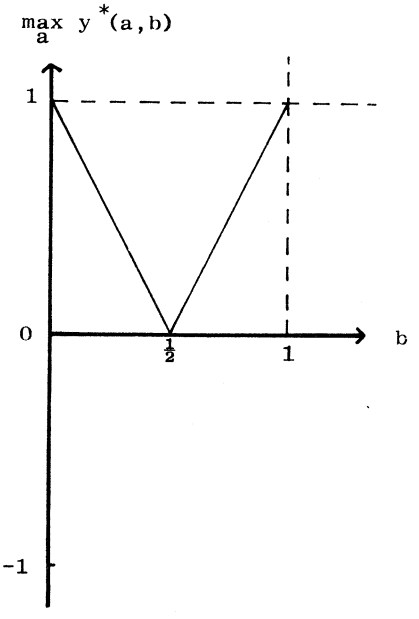


fig. 2.2b

De funkties $\min_b y^*(a,b)$ en $\max_a y^*(a,b)$

Evenzo zal II eerst $\max_a y^*(a,b)$ bepalen en vervolgens b zodanig kiezen, dat deze funktie minimaal is. Uit

$$(2.12)\quad \max_a y^*(a,b) = \max_a \left[1-2b+2a(2b-1)\right] = \begin{cases} 1-2b \text{ voor } 2b-1 \leq 0 \\ 2b-1 \text{ voor } 2b-1 \geq 0 \end{cases}$$

volgt

$$(2.13)\qquad \min_b \max_a y^*(a,b) = 0 \qquad \text{(vgl. fig. 2.2b).}$$

Voor dit nieuwe spel, waarin de spelers via een kansmechanisme bepalen, welke van de strategieën K en M zij in een bepaald spel spelen, geldt dus

(2.14) $$\max_a \min_b y^*(a,b) = \min_b \max_a y^*(a,b).$$

De strategieën a en b van dit nieuwe spel worden gemengde strategieën genoemd van het oude kruis- of muntspel, omdat zij, tenzij a en b nul of één zijn, ertoe leiden dat de spelers soms K,soms M spelen. Door zo'n gemengde strategie toe te passen kan I de winstverwachting, die hij in ieder geval kan bereiken, verhogen van -1 tot 0; II kan door het toepassen van een gemengde strategie beletten, dat de winstverwachting van I groter is dan nul.

Speelt I de strategie $a = \frac{1}{2}$, dan is zijn winstverwachting altijd gelijk aan 0, onafhankelijk van hetgeen II doet. Men kan dit eenvoudig nagaan door in (2.9) $a = \frac{1}{2}$ te substitueren. Gebruikt I een andere strategie, dan hangt het van de strategie van II af, of zijn winstverwachting positief is of negatief. Om dezelfde redenen als in voorbeeld 2.1 noemen wij daarom de strategie $a = \frac{1}{2}$ optimaal voor I en de strategie $b = \frac{1}{2}$ optimaal voor II.

Wij zullen nu nagaan wat de konsekwenties van gemengde strategieën zijn voor willekeurige twee personen nulspelen. Daartoe geven wij een strategie van I weer aan met x, de verzameling van al zijn mogelijke strategieën met $\mathfrak{X}$, een strategie van II met z en de verzameling van alle strategieën, waaruit II kan kiezen, met $\mathfrak{Z}$. Voorlopig nemen wij aan dat $\mathfrak{X}$ en $\mathfrak{Z}$ slechts eindig veel elementen bevatten. Een gemengde strategie van I is dan een kansverdeling f op $\mathfrak{X}$, een gemengde strategie van II een kansverdeling g op $\mathfrak{Z}$.

De kans, waarmee I het element x kiest, geven wij aan met f(x); de kans, waarmee II het element z kiest, met g(z). Moet II, wanneer I x en II z kiest, het bedrag y(x,z) aan I betalen, dan is de winstverwachting voor I, wanneer I de gemengde strategie f en II de gemengde strategie g speelt, gelijk aan

(2.15) $$y^*(f,g) = \sum_x \sum_z y(x,z)\, f(x)\, g(z).$$

Voeren wij in voorbeeld 2.4 de volgende notatie in:

$x = 1$ betekent speler I legt K boven,

$x = 2$ betekent speler I legt M boven,

$z = 1$ betekent speler II legt K boven,

$z = 2$ betekent speler II legt M boven,

dan krijgen de kansverdelingen f en g de volgende vorm

$$f(x) = \begin{cases} a & \text{voor } x = 1 \\ 1-a & \text{voor } x = 2 \\ 0 & \text{voor alle andere waarden van } x \end{cases} \tag{2.16}$$

en

$$g(z) = \begin{cases} b & \text{voor } z = 1 \\ 1-b & \text{voor } z = 2 \\ 0 & \text{voor alle andere waarden van } z. \end{cases} \tag{2.17}$$

Formule (2.15) wordt dan

$$y^*(f,g) = ab + (1-a)(1-b) - a(1-b) - b(1-a); \text{ (vgl. 2.9).}$$

De spelers kunnen ten opzichte van de winstfunktie (2.15) de te kiezen gemengde strategie op dezelfde wijze bepalen als in de voorgaande voorbeelden. De eerste speler zal dan voor iedere gemengde strategie f nagaan hoe groot het minimum van de winstverwachting is en dan die f kiezen, waarvoor $\min_g y^*(f,g)$ het maximum bereikt. Hij berekent dus de grootheid

$$\lambda_\Gamma \overset{\text{def}}{=} \max_f \min_g y^*(f,g) \tag{2.18}$$

en kiest die f, waarvoor λ_Γ wordt bereikt *).

Evenzo kan II die g kiezen, waarvoor het maximum over f minimaal is; de bijbehorende winstverwachting van I is dan

$$\nu_\Gamma \overset{\text{def}}{=} \min_g \max_f y^*(f,g). \tag{2.19}$$

---

*) Omdat f en g kansverdelingen zijn op eindige verzamelingen worden alle extremen ook werkelijk bereikt.

Wij gaan nu na wat het verband is tussen λ_G, ν_G, λ_Γ en ν_Γ. Daartoe voeren wij nog een nieuwe notatie in. Speler I kan voor f de kansverdeling kiezen, die met kans één de strategie x_0 aanwijst en de andere strategieën dus met kans nul. De winstverwachting bedraagt dan

$$\sum_x \sum_z y(x,z)\, f(x)g(z) = \sum_z y(x_0,z)g(z), \tag{2.20}$$

waarvoor wij schrijven $y^*(x_0,g)$. Voor de winstverwachting, die correspondeert met een willekeurige kansverdeling f voor I en de kansverdeling, die met kans één strategie z_0 aanwijst voor II, schrijven wij

$$y^*(f,z_0) = \sum_x y(x,z_0)f(x). \tag{2.21}$$

Stelling 2.2

Bevatten de verzamelingen $\mathcal{X}$ en $\mathcal{Z}$ van mogelijke strategieën voor I resp. II eindig veel elementen, dan is

$$\min_g y^*(f,g) = \min_z y^*(f,z) \quad \text{en} \quad \lambda_G \leq \lambda_\Gamma \tag{2.22}$$

en

$$\max_f y^*(f,g) = \max_x y^*(x,g) \quad \text{en} \quad \nu_G \geq \nu_\Gamma. \tag{2.23}$$

Bewijs:

Voor iedere f en g is

$$y^*(f,g) = \sum_z y^*(f,z)g(z) \geq \min_z y^*(f,z). \tag{2.24}$$

Immers, als

$$\min_z y^*(f,z) = y^*(f,z_0), \tag{2.25}$$

dan is

$$\sum_z y^*(f,z)g(z) = \sum_z \{y^*(f,z) - y^*(f,z_0) + y^*(f,z_0)\}g(z) =$$

$$= \sum_z \{y^*(f,z) - y^*(f,z_0)\}\, g(z) + y^*(f,z_0) \geq \tag{2.26}$$

$$\geq y^*(f,z_0) = \min_z y^*(f,z) .$$

Omdat betrekking (2.24) voor iedere g geldt, is zij ook juist voor het minimum van $y^*(f,g)$ over g en dus is

$$\min_g y^*(f,g) \geq \min_z y^*(f,z). \tag{2.27}$$

Onder de verzameling van alle mogelijke g's vallen ook die g's, die met kans één een bepaalde zuivere strategie aanwijzen, waaruit voor iedere z volgt

$$\min_g y^*(f,g) \leq y^*(f,z) \tag{2.28}$$

en dus

$$\min_g y^*(f,g) \leq \min_z y^*(f,z). \tag{2.29}$$

Uit de combinatie van (2.27) en (2.29) volgt

$$\min_g y^*(f,g) = \min_z y^*(f,z), \tag{2.30}$$

waarmee het eerste gedeelte van (2.22) is bewezen. Wanneer wij in (2.22) voor f de kansverdeling substitueren, die met kans één de zuivere strategie x aanwijst, dan vinden wij

$$\min_z y(x,z) = \min_g y^*(x,g). \tag{2.31}$$

Deze betrekking geldt voor iedere x en dus ook voor het maximum over x, zodat

$$\max_x \min_z y(x,z) = \max_x \min_g y^*(x,g). \tag{2.32}$$

Het linkerlid hierin is gelijk aan λ_G. Het rechterlid vergelijken we met $\lambda_\Gamma = \max_f \min_g y^*(f,g)$; λ_Γ is het maximum van $\min_g y^*(f,g)$ over alle

mogelijke verdelingsfunkties f op de verzameling $\mathcal{X}$. Hieronder komen de verdelingen voor die met kans één een bepaalde zuivere strategie x aanwijzen. Dit betekent, dat voor λ_Γ de funktie $\min_g y^*(f,g)$ gemaximaliseerd wordt over een grotere klasse van verdelingsfunkties dan in $\max_x \min_g y^*(x,g)$ en dus is

$$\max_x \min_g y^*(x,g) \leq \max_f \min_g y^*(f,g), \tag{2.33}$$

waaruit volgt $\lambda_G \leq \lambda_\Gamma$.

Het tweede deel van de stelling wordt op analoge wijze bewezen.

Deze stelling kan men niet alleen gebruiken voor het berekenen van optimale oplossingen, maar wij kunnen er ook het gezochte verband tussen λ_G, ν_G, λ_Γ en ν_Γ uit afleiden. Hierbij moet tevens stelling 2.1 toegepast worden, die men ook kan bewijzen voor funkties van het type $y^*(f,g)$. Uit deze generalisatie van stelling 2.1 volgt

$$\max_f \min_g y^*(f,g) \leq \min_g \max_f y^*(f,g), \tag{2.34}$$

of

$$\lambda_\Gamma \leq \nu_\Gamma .$$

Samen met stelling 2.2 volgt dan de betrekking

$$\lambda_G \leq \lambda_\Gamma \leq \nu_\Gamma \leq \nu_G. \tag{2.35}$$

Hieruit blijkt, dat de verwachting van de winst die I in ieder geval kan bereiken door het toepassen van gemengde strategieën wel hoger, maar nooit lager kan worden. Evenzo kan II met behulp van gemengde strategieën beletten, dat de verwachting van de winst die I kan bereiken, groter is dan ν_Γ, een bedrag dat kleiner is dan of gelijk aan ν_G. Dat er voorbeelden zijn, waarin de spelers hun situatie inderdaad kunnen verbeteren door gemengde strategieën toe te passen, is reeds gebleken in het eerder genoemde kruis- of muntspel. Ongelijkheid (2.35) leert ons bovendien, dat voor gevallen, waarin $\lambda_G = \nu_G$ is, ook $\lambda_\Gamma = \nu_\Gamma$ is.

Het introduceren van gemengde strategieën is nu daarom van veel belang omdat men de volgende stelling kan bewijzen:

Stelling 2.3

In ieder eindig twee personen nulspel is

(2.36) $$\lambda_\Gamma = \nu_\Gamma$$

en bezitten de spelers optimale strategieën.

Deze stelling is één van de versies van de hoofdstelling van de speltheorie en houdt in dat bij het toepassen van gemengde strategieën de winstverwachting, waarvan I zich in ieder geval kan verzekeren gelijk is aan de verliesverwachting, die II hoogstens behoeft te accepteren. Op grond van deze stelling kunnen wij dus ook optimale oplossingen vinden voor eindige spelen, waarin $\lambda_G < \nu_G$ is. Het besproken kruis- of muntspel is er een voorbeeld van.

In de vorm van stelling 2.3 treedt de hoofdstelling van de speltheorie op in het voor deze theorie baanbrekende boek van J. VON NEUMANN en O. MORGENSTERN, dat reeds op blz. 7 werd aangehaald. Van de vele bewijzen, die van deze stelling gegeven kunnen worden, is dat van G.B.DANTZIG een van de eenvoudigste (zie voetnoot op blz.23). Hij formuleert het probleem van I (het maximaliseren van λ_Γ) als een lineair programmeringsprobleem, waarbij de beslissingsvariabelen de kansen zijn, waarmee I de verschillende x-strategieën moet aanwijzen. Op dezelfde wijze geeft hij het probleem van II weer als een lineair programmeringsprobleem, waarbij ν_Γ geminimaliseerd moet worden. Het blijkt dan, dat het lineaire programmeringsprobleem van II het duale probleem is van dat van I, zodat $\lambda_\Gamma = \nu_\Gamma$ (vgl. deel 6, § 7).

Ter illustratie van de behandelde theorie geven wij het volgende voorbeeld.

Voorbeeld 2.5

Gevraagd wordt de optimale strategieën te berekenen van het spel, waarvan de winstfunktie gegeven is in tabel 2.4. De keuze van een gemengde strategie f is voor de eerste speler de keuze van de kans a,

waarmee in ieder spel strategie 1 wordt gekozen. Voor de tweede speler is een gemengde strategie een kansverdeling g voor de strategieën 1, 2, 3 en 4.

Tabel 2.4

De winstfunktie

I \ II	1	2	3	4
1	2	-2	-1	3
2	-2	4	0	-2

Volgens stelling 2.3 bezit het spel een waarde v, waarvoor geldt

$$v = \max_f \min_g y^*(f,g) = \min_g \max_f y^*(f,g). \tag{2.37}$$

Hierin is $y^*(f,g)$ de winstverwachting van de eerste speler. Doordat I uit slechts twee zuivere strategieën kan kiezen, kunnen wij $\max_f \min_g y^*(f,g)$ grafisch bepalen. Volgens stelling 2.2 is

$$\begin{aligned} \min_g y^*(f,g) &= \min_z y^*(f,z) = \min_z \left[y(1,z)f(1) + y(2,z)f(2)\right] = \\ &= \min_z \left[y(1,z)a + y(2,z)(1-a)\right]. \end{aligned} \tag{2.38}$$

De funktie $y(1,z)a + y(2,z)(1-a)$ is voor

$$\begin{aligned} z = 1 &: \quad 2a - 2(1-a) = 4a - 2, \\ z = 2 &: \quad -2a + 4(1-a) = -6a + 4, \\ z = 3 &: \quad -1a + 0(1-a) = -a, \\ z = 4 &: \quad 3a - 2(1-a) = 5a - 2. \end{aligned} \tag{2.39}$$

In fig. 2.3 zijn deze rechten geschetst voor het relevante gebied $0 \leq a \leq 1$. De dik getrokken, gebroken lijn geeft voor iedere waarde van a het minimum over z aan. Het maximum over a hiervan wordt bereikt in het snijpunt van de rechten $y^*(a,1) = 4a - 2$ en $y^*(a,3) = -a$.

Bij blz.22 : Zie bijv. G.B.DANTZIG, Linear Programming and Extensions, Princeton University Press, Princeton, (1963).

De bijbehorende waarde van a volgt dus uit

$$4a - 2 = -a$$

en is dus

$$a = \frac{2}{5} .$$

Voor $v = \max_a \min_g y^*(a,g)$ vinden wij $v = -\frac{2}{5}$.

Ook de optimale strategie voor II is zonder moeite op te sporen. Immers voor iedere waarde van z is $y^*(\frac{2}{5}, z) \geq -\frac{2}{5}$ (= v) en dus geldt voor de optimale strategie $\bar{g}$ van II

$$y^*(\tfrac{2}{5},\bar{g}) = \sum_z y^*(\tfrac{2}{5},z)\bar{g}(z) \geq -\frac{2}{5} . \qquad (2.40)$$

Anderzijds is de waarde van het spel $-\frac{2}{5}$, d.w.z. $y^*(\frac{2}{5},\bar{g}) = -\frac{2}{5}$; een bedrag dat slechts bereikt kan worden wanneer $\bar{g}(z) = 0$ voor iedere waarde van z, waarvoor $y^*(\frac{2}{5},z) > -\frac{2}{5}$ is. De optimale gemengde strategie voor II zal dus alleen positieve waarschijnlijkheden toekennen aan de strategieën z = 1 en z = 3 en het oorspronkelijke spel kan daarom voor de berekening van $\bar{g}$ vervangen worden door het in tabel

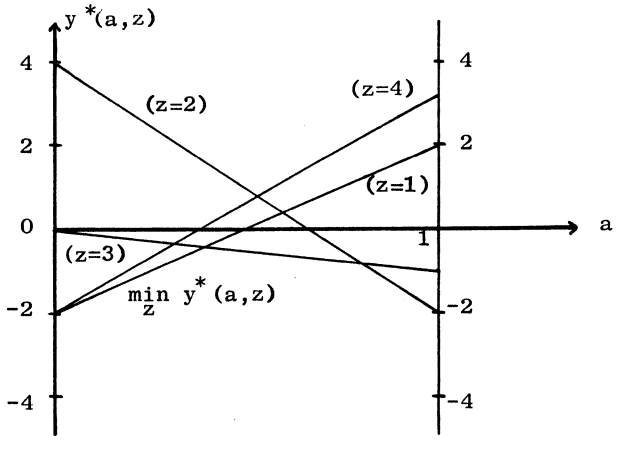


fig. 2.3 $y^*(a,z)$ voor z = 1, 2, 3, 4.
(op de assen zijn verschillende schalen gebruikt)

2.5 aangegeven spel. Dit is een spel waarvoor de optimale strategie voor II op dezelfde wijze gevonden kan worden als in het kruis- of

Tabel 2.5

Vereenvoudigde winstfunktie

I \ II	1	3
1	2	-1
2	-2	0

muntspel. Het resultaat is dan $\bar{g}(1) = \frac{1}{5}$ en $\bar{g}(3) = \frac{4}{5}$, zodat de optimale strategie van II in het in tabel 2.4 vermelde spel $\bar{g}(1) = \frac{1}{5}$, $\bar{g}(2) = 0$, $\bar{g}(3) = \frac{4}{5}$ en $\bar{g}(4) = 0$ is.

Tot nu toe hebben wij ons beperkt tot spelen, waarin $\mathcal{X}$ en $\mathcal{Z}$ slechts eindig veel elementen bevatten. Men kan echter ook spelen bedenken, waarin $\mathcal{X}$ en/of $\mathcal{Z}$ oneindig veel elementen bevatten. Zo zullen wij in § 3 en § 4 voorbeelden geven, waarin de strategieën van I of II de punten uit het interval $[0,1]$ zijn.

Onder bepaalde voorwaarden, die in praktijkproblemen vrijwel steeds vervuld zijn, gelden de stellingen 2.2 en 2.3 ook voor strategische spelen, waarin $\mathcal{X}$ en $\mathcal{Z}$ niet beide eindig veel elementen bevatten. Wij zullen hierop niet nader ingaan.

Naast de generalisatie tot spelen met oneindig veel strategieën voor de spelers zijn er ook theorieën ontwikkeld voor niet-nulspelen en spelen met meer dan twee personen. Men leze hiervoor bijvoorbeeld de reeds genoemde boeken van VON NEUMANN en MORGENSTERN en van LUCE en RAIFFA (vgl. de voetnoot op blz.7).

Hoewel de speltheorie vooral ontwikkeld is met de bedoeling economische conflictsituaties te onderzoeken, is het aantal praktische toepassingen op economisch terrein gering. Wel kan men bijvoorbeeld

eenvoudige situaties bij reclamecampagnes analyseren, evenals andere concurrentieproblemen en oligopolistische markten *). Bovendien verdiept de theorie het inzicht in bepaalde economische problemen. Hetzelfde geldt voor de in § 3 te behandelen spelen tegen de natuur en voor de speltheoretische analyses in de conflictologie en de polemologie. **)

Werkelijke toepassingen van de speltheorie zijn schaars en worden vooral gevonden in de krijgswetenschappen, waarvan voorbeeld 2.6 een eenvoudig geval illustreert. De minimaxmethode vindt uitgebreide toepassing, zoals twee voorbeelden in § 4 laten zien.

Voorbeeld 2.6

Een reeks artikelen over krijgskundige toepassingen van de strategische speltheorie is verschenen in Operations Research, het tijdschrift van de Operations Research Society of America (afgekort J.O.R.S.A. of ook wel O.R.). Zeer instructief is een artikel van O.G.HAYWOOD Jr. ***)

De auteur vergelijkt hierin de wijze van beslissingen nemen bij militaire operaties, zoals deze is goedgekeurd door de (Amerikaanse) Gezamenlijke Chefs van Staven met de oplossingsmethoden van de strategische speltheorie.

In principe kan een bevelhebber de situatie analyseren en zijn beslissingen nemen op grond van hetgeen de tegenstander in staat is te doen, of op grond van hetgeen hij vermoedt dat de tegenstander zal gaan doen. De eerste is de officieel goedgekeurde methode, aangezien de bevelhebber volgens de instructies dié acties moet uitvoeren, welke, gezien de mogelijkheden waarover de vijand beschikt, de grootste beloften tot succes geven. De geïnstrueerde analyse-methode bestaat

---

*) Men vergelijke bijv. : M.SHUBIK, Strategy and Market Structure, Wiley and Sons, New York, (1959).

**) Zie bijv. de artikelen van B.V.A.ROLING en Ph.P.EVERTS in Internationale Spectator, 20, (1966), p. 148-174.

***) O.G.HAYWOOD Jr., Military Decision and Game Theory, J.O.R.S.A., 2, (1954), p. 365-385.

uit vijf stappen en staat bekend onder de naam "estimate of the situation". Aan de hand van twee voorbeelden, ontleend aan de Tweede Wereldoorlog, laat de schrijver zien, dat bij de "estimate of the situation" en bij de speltheorie dezelfde overwegingen gebruikt worden. De speltheorie leidt in deze gevallen dan ook tot dezelfde beslissingen als de door de bevelhebbers genomen besluiten. In situaties waarin beide methoden dezelfde overwegingen gebruiken en de konsekwenties van de verschillende strategieën gekwantificeerd kunnen worden, is de speltheorie superieur, aangezien deze in tegenstelling tot de andere methode een procédé aangeeft om de optimale oplossing te berekenen.

Ter illustratie zullen wij het eenvoudigste voorbeeld hier in het kort behandelen. Het betreft de slag in de Bismarckzee. Een Japanse transportvloot moest in februari 1943 varen van Rabaul op New Britain naar Lae op Nieuw Guinea en kon hiervoor gebruik maken van de route ten noorden van New Britain of de route ten zuiden van het eiland. Beide zeewegen namen drie dagen varen in beslag. De bevelhebber van de geallieerde luchtmacht in het zuid-westen van de Grote Oceaan had de opdracht zoveel mogelijk schepen te vernietigen. Hiertoe moest uiteraard de transportvloot eerst worden opgespoord. Verder was bekend dat het zicht op de noordelijke route slecht was wegens regen.

De geallieerde bevelhebber beschouwde nu twee mogelijkheden: concentratie van het grootste gedeelte van zijn verkenningsvliegtuigen op de noordelijke route en slechts weinig op de zuidelijke, of juist andersom. Afhankelijk van de tijd, nodig om de vijandelijke vloot te ontdekken, kon men bepalen hoeveel dagen er beschikbaar waren om de transportvloot te bombarderen. Dit aantal is voor de verschillende gevallen opgegeven in tabel 2.6. Wij beschouwen deze tabel als de winstfunktie van een twee personen nulspel, waarin beide spelers twee strategieën ter beschikking hebben.

Tabel 2.6

Winstfunktie bij de slag op de Bismarckzee

Amerikanen \ Japanners	noordelijke route	zuidelijke route
noordelijke route	2	2
zuidelijke route	1	3

Men ziet direct, dat de waarde van het spel twee is en dat beide partijen een zuivere optimale strategie bezitten, namelijk de noordelijke route, die ook in werkelijkheid door beide is gekozen.

Naar aanleiding van het tweede voorbeeld, de slag bij Avranches in Frankrijk, behandelt HAYWOOD de betekenis van gemengde strategieen. Ook de verhouding van het nemen van beslissingen op grond van de "estimate of the situation" tot het handelen op grond van de vermoede plannen van de tegenstander komt ter sprake.

3. Statistische spelen.

A.WALD heeft in zijn boek Statistical Decision Functions *) aangetoond dat alle bekende statistische problemen, zoals het schatten en toetsen van parameters en het opstellen van betrouwbaarheidsintervallen, opgevat kunnen worden als strategische spelen, waarin de statisticus speelt tegen de natuur. In deze zgn. statistische spelen ("statistical games") krijgt de natuur de rol van de eerste speler en de statisticus

*) A.WALD, Statistical Decision Functions, Wiley and Sons, New York, (1950).

die van de tweede. De verzameling van alle mogelijke strategieën van de natuur wordt weer aangegeven met $\mathfrak{X}$; een bepaalde strategie met x. Weet men bijv. dat een bepaalde grootheid normaal verdeeld is met bekende standaardafwijking σ, maar onbekende verwachting μ, dan vormen alle mogelijke waarden van μ de verzameling $\mathfrak{X}$. Zijn zowel de verwachting μ als de standaardafwijking σ onbekend, dan bestaat $\mathfrak{X}$ uit de 2-dimensionale vectoren (μ,σ), $-\infty < \mu < +\infty$, $\sigma > 0$.

De statisticus kan een strategie a kiezen uit een gegeven verzameling van mogelijke strategieën $\mathfrak{A}$; in plaats van het woord strategie gebruikt men in dit geval ook wel het woord beslissing. Het verlies dat de statisticus lijdt, wanneer hij beslissing a neemt en de natuur strategie x speelt, is de <u>verliesfunktie</u> van het spel, welke wordt aangegeven met y(x,a).

<u>Voorbeeld 3.1</u>

In een keuringsprobleem moet de statisticus beslissen of hij een partij goederen goedkeurt (strategie a_1) of afkeurt (strategie a_2). De strategie van de natuur is de fractie defecte exemplaren die in de partij aanwezig is. De verzameling $\mathfrak{X}$ van alle mogelijke strategieën van de eerste speler is dus het interval $[0,1]$, zodat hij oneindig veel mogelijke strategieën heeft. Is de winst op ieder verkocht exemplaar c_1, maar wordt bij een verkocht defect exemplaar bovendien een verlies c_2 geleden, dan is de verliesfunktie voor een partij van N exemplaren

$$y(x,a) = \begin{cases} c_2 \times N - c_1 N & \text{voor } a = a_1, \\ 0 & \text{voor } a = a_2. \end{cases} \tag{3.1}$$

Geschreven in een tabel van dezelfde vorm als de winstfunkties in §2, ziet deze funktie er als volgt uit:

Tabel 3.1

Verliesfunktie van voorbeeld 3.1

I \ II	a_1	a_2
0	$-c_1N$	0
.	.	.
.	.	.
.	.	.
x	c_2xN-c_1N	0
.	.	.
.	.	.
.	.	.
1	c_2N-c_1N	0

Zouden wij in analogie met §2 op deze tabel voor de statisticus de minimaxmethode toepassen, dan zou hij altijd moeten goedkeuren, wanneer $c_2 < c_1$ is en altijd moeten afkeuren wanneer $c_2 > c_1$ is. De minimaxoplossing zou derhalve niet afhangen van de kwaliteit van de partij.

Men kan tegenwerpen, dat het onbevredigende resultaat in dit voorbeeld ontstaat doordat geen rekening is gehouden met de mogelijkheid eerst met behulp van een steekproef de kwaliteit te bepalen; i.e. de strategie van de natuur te onderzoeken. Inderdaad behoeft de statisticus niet zonder enige informatie omtrent de strategie van zijn tegenspeler een beslissing te nemen.

Onderstel daarom dat een steekproef van de omvang n genomen wordt. Het aantal goede exemplaren in de steekproef geven wij aan met k; de verzameling van alle denkbare steekproefresultaten met $\mathcal{K}$. Ieder element

x uit $\mathcal{X}$ legt een kansverdeling p(k|x) op $\mathcal{K}$ vast, die aangeeft met welke kans de waarde k optreedt, wanneer de natuur strategie x speelt. Op grond van het steekproefresultaat kiest de statisticus een element uit $\mathcal{A}$; m.a.w. aan ieder steekproefresultaat wordt een a_i toegevoegd en de statisticus kan bepalen op welke wijze deze toevoeging geschiedt. Hij kiest dus een funktie d, welke bij ieder element k uit $\mathcal{K}$ een bepaald element d(k) uit $\mathcal{A}$ aanwijst. De funktie d(k) wordt een beslissingsfunktie of besluitregel genoemd. De verzameling van alle funkties d waaruit de statisticus kan kiezen wordt aangegeven met $\mathfrak{D}$.

De betrekkingen die tussen de ingevoerde grootheden bestaan, zijn als volgt schematisch aan te geven

$$x \in \mathcal{X} \longrightarrow \text{kansverdeling } p(k|x) \text{ op } \mathcal{K};$$

$$k \in \mathcal{K} \xrightarrow{d \in \mathfrak{D}} a \in \mathcal{A}.$$

Voor de verliesfunktie y(x,a) kunnen wij nu ook schrijven y(x,d(k)) om het verband tussen het verlies en de gekozen beslissingsfunktie d tot uitdrukking te brengen. De verliesverwachting voor de statisticus bij de strategie x van de natuur en de beslissingsfunktie d is de risicofunktie r(x,d), waarvoor geldt

$$r(x,d) = \sum_{k \in \mathcal{K}} y(x,d(k))\ p(k|x). \tag{3.2}$$

Vergelijken wij het keuringsprobleem in de nu verkregen vorm met het gewone strategische spel, dan zien wij dat de ruimte $\mathcal{X}$ bestaat uit de onbekende waarden van een parameter en dat de ruimte $\mathcal{Z}$ vervangen is door de ruimte $\mathfrak{D}$ van beslissingsfunkties, waaruit de statisticus kan kiezen. De funktie y(x,z) die de winst voor de eerste en dus het verlies voor de tweede speler aangeeft, is overgegaan in de funktie r(x,d). Zowel $\mathcal{X}$ als $\mathfrak{D}$ zijn ruimten van zuivere strategieën. Dat de funktie r(x,d) hier in tegenstelling tot y(x,z) toch de vorm van een verwachting aanneemt, komt doordat een zuivere strategie van de natuur een kansverdeling voor het steekproefresultaat k vastlegt. Later in deze paragraaf en in § 4 zal op het vaststellen van een beslissingskriterium voor dit keuringsprobleem nader worden ingegaan.

De zojuist gegeven uitwerking van het keuringsprobleem kan ook bij andere problemen gebruikt worden. Wij illustreren dit eerst voor een eenvoudig schattingsprobleem.

Voorbeeld 3.2

Door middel van een steekproef met teruglegging van de omvang n wil men de fractie elementen met een bepaalde eigenschap in een populatie schatten. De verzameling $\mathcal{X}$ bestaat weer uit alle getallen in het interval $[0,1]$; de verzameling $\mathcal{K}$ uit de getallen 0, 1, ..., n. De kansverdeling $p(k|x)$ luidt

$$p(k|x) = \binom{n}{k} x^k (1-x)^{n-k} \quad (0 \leq x \leq 1;\ k = 0,1,\ldots,n). \tag{3.3}$$

De ruimte $\mathcal{A}$ bestaat uit het interval $0 \leq a \leq 1$, waarin a een schatting is voor de onbekende fractie x. Een beslissingsfunktie die men kan kiezen is bijv.

$$d(k) = \frac{k}{n}\ ; \tag{3.4}$$

men neemt dan als schatting van de fractie in de populatie de fractie, die men in de steekproef heeft gevonden. Een andere beslissingsfunktie, die men zou kunnen kiezen is

$$\begin{cases} d_1(k) = \dfrac{k-1}{n} & \text{voor } 1 \leq k \leq n, \\ d_1(k) = 0 & \text{voor } k = 0. \end{cases} \tag{3.5}$$

Is bovendien gegeven dat het verlies, dat men lijdt, evenredig is met het kwadraat van het verschil tussen de geschatte en de werkelijke fractie, dan is

$$y(x,a) = \lambda(x-a)^2, \tag{3.6}$$

waarin λ een constante is.

De risicofunktie wordt

$$r(x,d) = \sum_{k=0}^{n} y(x,d(k))\ p(k|x) = \sum_{k=0}^{n} \lambda(x - \frac{k}{n})^2 \binom{n}{k} x^k (1-x)^{n-k} =$$

(3.7)

$$= \frac{\lambda}{n^2} \sum_{k=0}^{n} (k-xn)^2 \binom{n}{k} x^k (1-x)^{n-k} = \frac{\lambda}{n^2}\, nx(1-x) = \frac{\lambda x}{n}\,(1-x)$$

(vgl. formule (10.29), deel 2).

Opmerking 3.1

Wij hebben hier bij het opstellen van de risicofunktie geen rekening gehouden met de kosten, die verbonden zijn aan het verrichten van waarnemingen. Dit is juist, wanneer de omvang van de steekproef van te voren vaststaat en de kosten onafhankelijk zijn van de uitkomsten der experimenten. In de meeste gevallen is aan deze voorwaarden echter niet voldaan en dan moeten ook deze kosten in de risicofunktie worden opgenomen.

Voor meer algemene statistische problemen kan de beschreven gedachtengang eveneens worden uitgewerkt. Dit is o.a. gedaan door E. L. LEHMANN*). Wel zijn de ruimten $\mathcal{X}, \mathcal{K}$ en $\mathcal{A}$ meestal ingewikkelder dan in de hier gegeven voorbeelden. Zo zal de beslissingsfunktie d meestal niet in de vorm d(k) geschreven kunnen worden, doch een meer algemene funktie zijn van de waarnemingen $w_1, \ldots, w_n$. Hij blijft echter een keuze a uit $\mathcal{A}$ toevoegen aan de waarnemingsresultaten.

In deze redenering kan men een tweezijdige toets van H_0: $\theta = \theta_0$ beschouwen als een besluitregel, die de beslissing "H_0 verwerpen ten gunste van $\theta < \theta_0$" toevoegt aan alle $(w_1, \ldots, w_n)$ waarvoor de toetsingsgrootheid $t(w_1, \ldots, w_n)$ ligt in de linker kritieke zone, de beslissing "H_0 verwerpen ten gunste van $\theta > \theta_0$" aan waarnemingen waarvoor $t(w_1, \ldots, w_n)$ ligt in de rechter kritieke zone en "H_0 niet verwerpen" aan de overige waarnemingen. De toetsingsgrootheid t neemt hier de

---

*) E.L. LEHMANN, Testing Statistical Hypotheses, John Wiley and Sons, New York, (1959).

plaats in van de beslissingsfunktie d; de onbekende parameter θ, die van de strategie x van de natuur.

Bij een schatting van θ, zoals in voorbeeld 3.2, is de besluitregel de funktie t die aan elke rij waarnemingen de schatting $t(w_1, \ldots, w_n)$ toevoegt; elke beslissing is hier een getal. Bij het opstellen van een betrouwbaarheidsinterval vormen de funkties g_l en g_r, die de grenzen van het interval geven als funkties van de waarnemingen, de besluitregel; het interval $\left[g_l(w_1, \ldots, w_n), g_r(w_1, \ldots, w_n)\right]$ is de beslissing.

Voor alle drie genoemde situaties - toetsen, schatten d.m.v. een getal en schatten d.m.v. een interval - geldt dat de (mate van) juistheid van de beslissing afhangt van de werkelijke waarde van de onbekende parameter θ. LEHMANN neemt nu aan dat men bij de beslissing a, als θ de werkelijke waarde van de parameter is, een verlies $y(\theta,a)$ lijdt. Doel van de theorie is ook hier een besluitregel d te vinden, die het verwachte verlies minimaliseert. Is de simultane verdelingsfunktie van $\underline{w}_1, \ldots, \underline{w}_n$ bekend, dan is dit verwachte verlies, de risicofunktie, gelijk aan

$$r(\theta,d) = \mathcal{E}\left[y(\theta, d(\underline{w}_1, \ldots, \underline{w}_n) \mid \theta\right]. \tag{3.8}$$

Evenals in voorbeeld 3.2 is gedaan, kan men de verliesfunktie evenredig nemen aan het kwadraat van het verschil tussen de schatting a en de werkelijke waarde θ: $y(\theta,a) = \lambda(a-\theta)^2$. Wanneer wij ons tot zuivere schatters beperken, komt het minimaliseren van

$$r(\theta,d) = \mathcal{E}\left[(t(\underline{w}_1, \ldots, \underline{w}_n) - \theta)^2 \mid \theta\right] \tag{3.9}$$

neer op het kiezen van dié zuivere schatter $\underline{t}$ die de kleinste variantie heeft, de meest doeltreffende schatter dus (vgl. deel 3, § 3).

Bij het toetsen van $\theta = \theta_0$ zou men het verlies bij een fout van de eerste soort y_1 kunnen noemen en het verlies bij een fout van de tweede soort $y_2|\theta - \theta_0|$, als θ de ware waarde van de parameter is.

Dan is

$$r(\theta,d) = \begin{cases} y_1 \; P\left[\underline{t} \in Z \mid \theta_0\right] & \text{als } \theta = \theta_0, \\ y_2 \; |\theta - \theta_0| \; P\left[\underline{t} \notin Z \mid \theta\right] & \text{als } \theta \neq \theta_0, \end{cases} \tag{3.10}$$

waarin Z de kritieke zone is van de bij de besluitregel d behorende toets.

Tot nu toe hebben wij ons beperkt tot zuivere strategieën voor de natuur en voor de statisticus. Ook bij statistische spelen kunnen wij echter gemengde strategieën invoeren. Een gemengde strategie f van de natuur is dan een kansverdeling op $\mathfrak{X}$, welke verdeling men soms de a priori verdeling op de ruimte van mogelijke parameterwaarden noemt. Een gemengde strategie g voor de statisticus is een kansverdeling op de ruimte $\mathfrak{D}$ van beslissingsfunkties, waaruit hij kan kiezen. Zijn er eindig veel of aftelbaar oneindig veel zuivere strategieën, dan zijn f en g discrete verdelingsfunkties; in het andere geval kunnen ze discreet of continu zijn. Voor discrete verdelingsfunkties gaat de verliesverwachting van de statisticus over in

$$r(f,g) = \sum_x \sum_d r(x,d) f(x) g(d), \tag{3.11}$$

waarin f(x) de kans is, waarmee de natuur de strategie x aanwijst en g(d) de kans, waarmee de statisticus de beslissingsfunktie d kiest. Bij gemengde continue verdelingsfunkties moeten de sommen vervangen worden door integralen.

Zodra de statisticus de funktie r(f,g) kent, zal hij zich afvragen, welke strategie g voor hem de beste is. Om dit te onderzoeken heeft hij een maatstaf nodig; d.w.z. een kriterium ten opzichte waarvan hij een optimale strategie kan kiezen.

Aangezien het gelukt is het probleem te gieten in de vorm van een strategisch spel, ligt het voor de hand eerst na te gaan of het minimaxkriterium een bevredigende oplossing geeft. In tegenstelling tot spelen waarin de belangen van de spelers diametraal tegenover elkaar staan, is dit kriterium hier niet zo goed en kan het tot zeer merkwaardige konsekwenties leiden. Dit komt doordat men niet kan aannemen dat de

natuur tracht de statisticus zoveel mogelijk afbreuk te doen, terwijl bij de minimaxstrategie de beslissing gebaseerd wordt op de meest ongunstige strategie van de natuur. Een voorbeeld kan dit verduidelijken.

Stel dat een partij goederen alleen aan de kwaliteitseisen voldoet, wanneer de fractie defecten kleiner is dan of gelijk aan x_0 en anders niet. De statisticus keurt de partijen m.b.v. steekproeven voordat ze worden afgeleverd. Hierbij zal hij weinig onjuiste beslissingen nemen bij zeer goede en zeer slechte partijen. De natuur kan het de statisticus nu "moeilijk maken" door veel partijen te fabriceren, die een fractie defecten bezitten dicht bij x_0. Bij de minimaxstrategie gaat de statisticus uit van een a priori verdeling f van de natuur, waarbij vrijwel uitsluitend partijen met een fractie x_0 aan defecten worden gemaakt. In werkelijkheid beschikt men door ervaringen uit het verleden wel over enige voorkennis omtrent de a priori verdeling op $\mathcal{X}$ en doordat hiervan bij het minimaxkriterium geen gebruik wordt gemaakt, faalt dit kriterium.

Een beslissingskriterium, waarbij dit wel gebeurt is het <u>kriterium van Bayes</u>. Hierbij gaat men ervan uit dat de statisticus de kansverdeling f, die de natuur gebruikt, volledig kent. Het risico van de statisticus, wanneer hij als strategie de kansverdeling g op $\mathfrak{D}$ kiest, is dan de verwachting van $r(\underline{x},g)$ ten opzichte van de kansverdeling f; dus

$$r(f,g) = \sum_{x} r(x,g)f(x). \tag{3.12}$$

Hij zal dan dié zuivere of gemengde strategie g_0 kiezen, waarvoor de verwachting van r(f,g) minimaal is. Men noemt g_0 de <u>Bayes-oplossing</u> met betrekking tot de a priori verdeling f.

Deze methode is identiek aan de in §1 behandelde methode en zal dus tot goede resultaten leiden onder de aan het slot van die paragraaf genoemde voorwaarden. De a priori verdeling moet inderdaad bekend zijn. Bij veel problemen op het gebied van de kwaliteitsbeheersing is dit zo en kan men bijv. de momenten van de a priori verdeling schatten uit aanwezig waarnemingsmateriaal.

Sommige auteurs *) verdedigen ook het gebruik van Bayes-oplossingen, wanneer de a priori verdeling alleen het subjectieve oordeel van de onderzoeker beschrijft. Bayes-oplossingen hebben dan het bezwaar dat zij in dezelfde situatie tot geheel verschillende resultaten kunnen leiden, wanneer verschillende mensen deze situatie verschillend beoordelen. Men kan dan als statisticus alleen een uitspraak doen van de vorm: indien de beoordeling omtrent de a priori verdeling juist is, dan is deze g_0 de optimale beslissingsprocedure.

De conclusie is derhalve dat het opvatten van statistische problemen als spelen tegen de natuur weliswaar het inzicht in die problemen verdiept, maar in het algemeen toch weinig nieuwe gezichtspunten biedt om te komen tot praktisch bruikbare oplossingen van problemen, die ook niet met de klassieke methoden opgelost kunnen worden. De minimaxmethode is te pessimistisch omdat deze zich richt op de meest ongunstige omstandigheden, terwijl de Bayes-methode neerkomt op het optimaliseren van verwachtingen. Een groot aantal theoretische beschouwingen over het bestaan en de eigenschappen van optimale beslissingsfunkties kan men vinden in het reeds genoemde boek van LEHMANN (vgl. voetnoot op blz.33).

4. Twee toepassingen van de minimaxmethode.

Bij accountantscontroles moeten in veel gevallen lijsten met zeer grote aantallen kleine posten worden nagezien. Wil men vaststellen, of in een dergelijke lijst ernstige fouten voorkomen, dan kan men in principe natuurlijk alle posten afzonderlijk controleren. Praktisch gezien is dit echter niet altijd uitvoerbaar en economisch gezien niet altijd verantwoord. In sommige situaties kan dan gebruik worden gemaakt van aselecte steekproeven.

Stel bijv. dat men wil nagaan of een lijst posten geen ten onrechte opgevoerde (gefraudeerde) posten bevat, of posten, die voor een te hoog bedrag zijn genoteerd. Alle guldens die zich ten onrechte op de lijst bevinden, worden als foutieve of gefraudeerde guldens opgevat en fouten genoemd. Bij deze zogenaamde positieve controles eist men

---

*) Zie bijv.: L.J. SAVAGE, The Foundations of Statistics, John Wiley and Sons, New York, (1954).

veelal, dat in de steekproef één of meer fouten gevonden worden, wanneer het totale, te hoog opgevoerde, bedrag meer is dan een bepaald percentage van het opgegeven totale bedrag. Voor de steekproefgewijze controle zijn er nu twee uitgangspunten mogelijk: bevat de lijst N posten en is het opgegeven totaalbedrag T, dan kan men de lijst beschouwen als een populatie van N posten, òf als een populatie van T guldens. In het eerste geval neemt men een steekproef van posten, in het tweede een steekproef van guldens. Controle van een in de steekproef aangewezen gulden betekent dan dat de post, waartoe de desbetreffende gulden behoort, wordt gecontroleerd.

Statistisch gezien kunnen met beide methoden zinvolle antwoorden verkregen worden. Aangezien bij de guldensmethode in het algemeen met kleinere steekproeven kan worden volstaan om een bepaalde zekerheid te krijgen, zullen wij ons hiertoe beperken.

Voordat de steekproef genomen wordt nummert men alle guldens in de populatie met de getallen 1, ..., T. Een post van a guldens krijgt dus a nummers. Is de fractie ten onrechte opgevoerde guldens gelijk aan p en neemt men een aselecte steekproef zonder teruglegging van n guldens, dan is de kans P_0 dat geen van de aangewezen guldens fout is

$$P_0 = \frac{\binom{(1-p)T}{n}}{\binom{T}{n}} \tag{4.1}$$

(vgl. deel 2, voorbeelden 5.7 en 5.8). Meestal geldt $n << T$ en $n << pT$, zodat P_0 in goede benadering gelijk is aan $(1-p)^n$. Eist men nu dat P_0 hoogstens β_0 bedraagt, wanneer de fractie ten onrechte opgevoerde guldens p_0 is of meer, dan dient een steekproef genomen te worden van een omvang n, zodanig dat geldt

$$(1 - p_0)^n \leq \beta_0. \tag{4.2}$$

Tabel 4.1 geeft voor enige waarden van p_0 en β_0 de kleinste gehele waarden van n waarvoor aan (4.2) is voldaan. Heeft men n vastgesteld, dan maakt men een lijst van n aselect gekozen getallen tussen 1 en T;

Tabel 4.1

Minimaal vereiste steekproefomvang voor gegeven waarden van p_0 en β_0

p_0 \ β_0	0,05	0,02	0,01	0,005	0,001
0,05	59	77	90	104	135
0,02	149	194	228	263	342
0,01	299	390	459	528	688
0,005	598	781	919	1058	1379
0,001	2995	3911	4603	5296	6905

vervolgens worden de guldens met de op deze lijst voorkomende getallen in de populatie opgezocht en de posten waartoe deze guldens behoren, gecontroleerd.

De wijze waarop in de beschreven methode de waarden van p_0 en β_0 worden gekozen is betrekkelijk arbitrair. Een meer objectieve methode om n te bepalen is voorgesteld door VAN HEERDEN *). Het is een toepassing van het minimaxkriterium, die in enigszins van de oorspronkelijke formulering afwijkende bewoordingen op het volgende neerkomt.

Laten de controlekosten c per gecontroleerde eenheid bedragen; dan zijn de totale controlekosten bij een steekproef van de omvang n gelijk aan cn. Er ontstaat schade wanneer ten onrechte besloten wordt de lijst posten te aanvaarden. Keuren wij de lijst alleen dan goed wanneer in de steekproef geen enkele fout wordt aangetroffen, dan is de kans hierop bij een guldenssteekproef van n stuks uit een populatie waarin een fractie p ten onrechte is opgevoerd, gelijk aan $(1-p)^n$. Stel dat de schade die ontstaat wanneer deze lijst wordt aanvaard, gelijk is aan een bekende funktie S van p: S(p). De kans hierop is $(1-p)^n$, zodat de verwachting van de totale kosten, controlekosten plus schadekosten, gelijk is aan

*) A. VAN HEERDEN, Steekproeven als middel van accountantscontrole, Maandblad voor Accountancy en Bedrijfshuishoudkunde, 35, (1961), p. 453-475.

(4.3) $$cn + (1-p)^n S(p).$$

De waarde van p is echter niet bekend. Zou de fraudeur een gemengde strategie toepassen, die vastgelegd wordt door een onbekende verdelingsfunktie G(p) op het interval $[0,1]$, dan is de risicofunktie van de controleur

(4.4) $$r(G,n) = \int_0^1 (1-p)^n S(p)\,dG(p) + cn \quad *).$$

Men kan nu, evenals in § 2 en § 3 is gedaan, de vrijheid in de keuze van n gebruiken om het risico te minimaliseren voor het ongunstigste geval.

Daartoe wordt eerst

(4.5) $$r_m(n) = \max_{G(p)} r(G,n)$$

berekend. Het maximum $r_m(n)$ wordt bereikt voor dié G(p) die een kans 1 geeft aan de waarde p_m van p, waarvoor $(1-p)^n S(p)$ maximaal is, en een kans 0 aan alle andere waarden van p, zodat geldt

(4.6) $$r_m(n) = (1-p_m)^n S(p_m) + cn.$$

VAN HEERDEN stelt vervolgens de schade gelijk aan het te hoog opgevoerde bedrag en kiest derhalve

(4.7) $$S(p) = pT.$$

---

*) $\int_0^1 (1-p)^n S(p)\,dG(p)$ is een verkorte schrijfwijze, waarmee bedoeld wordt $\int_0^1 (1-p)^n S(p) g(p)\,dp$ als G(p) een continue verdelingsfunktie is en $\sum_i (1-p_i)^n S(p_i)\, P[\underline{p} = p_i]$ als G(p) een discrete verdelingskunktie is; g(p) is de verdelingsdichtheid van $\underline{p}$; $P[\underline{p} = p_i]$ is de kans dat $\underline{p}$ de waarde p_i aanneemt.

De waarde van p_m wordt dan gevonden door (4.7) in (4.3) te substitueren, naar p te differentiëren en de afgeleide gelijk aan nul te stellen. Het resultaat is

$$T(1 - p_m)^{n-1} (-np_m + 1 - p_m) = 0, \tag{4.8}$$

of

$$p_m = \frac{1}{n + 1} ; \tag{4.9}$$

$r_m(n)$ wordt

$$(1 - \frac{1}{n+1})^n \frac{T}{n+1} + cn. \tag{4.10}$$

Wij minimaliseren tenslotte dit maximale risico als funktie van n. De optimale waarde n_0 van n blijkt gelijk te zijn, òf aan het kleinste gehele getal $\geq V$, òf aan het grootste gehele getal $\leq V$, waarin

$$V = \sqrt{\frac{T}{ce} + \frac{1}{4}} - \frac{1}{2}. \quad *)$$

Voor grote waarden van T geldt dus

$$n_0 \approx \sqrt{\frac{T}{ce}} . \tag{4.11}$$

Opmerking 4.1

De risicofunktie (4.3) is alleen dan geheel correct, wanneer

a) de steekproef van n guldens ook een steekproef van n posten is (wegens de term cn);
b) het aantal elementen in de steekproef veel kleiner is dan het aantal elementen in de populatie en er slechts n elementen gezien worden (wegens $(1-p)^n$);
c) aan een gulden, besteed aan controle, dezelfde waarde wordt gehecht als aan een schadeverwachting van één gulden.

---

*) Voor de afleiding zij verwezen naar: J. KRIENS, De methoden van DE WOLFF en VAN HEERDEN voor het nemen van aselecte steekproeven bij accountantscontroles, Statistica Neerlandica, 17 (1963), p. 215-231 ; rapport S 300 van de afd.Math.Stat. van het Mathematisch Centrum.

Wanneer niet aan deze drie voorwaarden is voldaan moet nagegaan worden of (4.3) een voldoend nauwkeurige benadering van het risico vormt.

Opmerking 4.2

Het probleem vormt geen strategisch spel, zoals in §2 beschreven is, aangezien de verliesfunktie van de controleur niet gelijk is aan de winstfunktie van de fraudeur (o.a. heeft de fraudeur de kosten en niet).

De tweede toepassing van de minimaxmethode, die wij zullen behandelen, is eveneens ontleend aan het in het begin van deze paragraaf beschreven controleprobleem.

In veel gevallen is het onmogelijk om één steekproef uit de gehele te controleren populatie te nemen. Beschouw bijv. een firma, waar jaarlijks grote aantallen inkoopfakturen binnenkomen, die door middel van een steekproef gecontroleerd moeten worden op ernstige fouten. Om een steekproef uit de gehele populatie van alle fakturen te kunnen nemen moet men wachten tot het einde van het boekjaar. De fakturen zijn dan grotendeels opgeborgen, waardoor het opzoeken van de in de steekproef te controleren guldens (posten) zeer bemoeilijkt wordt. Dit bezwaar kan men ondervangen door het jaar te verdelen in r perioden en aan het einde van elke periode een steekproef te nemen uit de dan beschikbare deelpopulatie. De steekproef van n guldens, die men over het gehele jaar wil nemen, moet dan verdeeld worden over de verschillende perioden. Op welke wijze dient deze verdeling te geschieden?

Stel dat wij ook hier eisen, dat bij een te hoog bedrag, gelijk aan een fractie p van het totaal opgegeven bedrag T, behoudens een kans β_0, één of meer fouten gedurende het jaar worden gevonden. De kans dat in geen van de steekproeven iets gevonden wordt, is nu niet noodzakelijkerwijs gelijk aan $(1-p)^n$, omdat de fractie fouten niet in iedere periode gelijk aan p behoeft te zijn. Men kan de gezochte kans als volgt afleiden. Zij

a) T_i, de omvang van de populatie in periode i;

b) p_i, de fractie fouten in periode i;

c) n_i, de omvang van de steekproef in periode i.

Aangezien het gaat om een totale fout in het jaar, gelijk aan pT, en er gedurende het gehele jaar n guldens gecontroleerd worden, moet voldaan zijn aan

$$\sum_{i=1}^{r} p_i T_i = pT \tag{4.12}$$

en

$$\sum_{i=1}^{r} n_i = n. \tag{4.13}$$

De kans in de r steekproeven geen enkele fout te vinden is

$$(1-p_1)^{n_1}(1-p_2)^{n_2} \ldots (1-p_r)^{n_r} = \prod_{i=1}^{r} (1-p_i)^{n_i}. \tag{4.14}$$

Vatten wij het probleem op als een strategisch spel tussen de controleur en de fraudeur, dan is de eerste vrij om onder voorwaarde (4.13) de n_i te kiezen en de tweede om onder voorwaarde (4.10) de p_i te bepalen. Een zuivere strategie van de controleur bestaat dus uit de getallen $n_1, \ldots, n_r$ en een zuivere strategie van de fraudeur uit de fracties $p_1, \ldots, p_r$. De verliesfunktie (4.14) van de controleur geeft de kans aan dat hij bij een bepaalde hoeveelheid controlewerk en bij een bepaald te hoog opgevoerd bedrag pT geen enkele fout waarneemt. Past de controleur op deze funktie de minimaxmethode toe, dan minimaliseert hij de maximale waarde van deze kans. Voor de fraudeur geldt het tegenovergestelde. Wij bepalen dus

$$\min_{n_i} \max_{p_i} \prod_{i=1}^{r} (1 - p_i)^{n_i} \tag{4.15}$$

en

$$\max_{p_i} \min_{n_i} \prod_{i=1}^{r} (1 - p_i)^{n_i} \tag{4.16}$$

onder de voorwaarden (4.12) en (4.13).

De bepaling van $\max_{p_i} \prod_{i=1}^{r} (1 - p_i)^{n_1}$ onder de voorwaarde (4.12) kan geschieden door de multiplicatorenmethode van LAGRANGE (vgl. stelling 12.8, deel 1) toe te passen op de logaritme van $\prod_{i=1}^{r} (1 - p_i)^{n_i}$. De Lagrangefunktie is dan

$$F(p_i,\lambda) = \log \prod_{i=1}^{r} (1 - p_i)^{n_i} + \lambda \left(\sum_{i=1}^{r} p_i T_i - pT \right). \tag{4.17}$$

Verder is

$$\frac{\partial F}{\partial p_i} = - \frac{n_i}{1 - p_i} + \lambda T_i \tag{4.18}$$

en

$$\frac{\partial F}{\partial \lambda} = \sum_{i=1}^{r} p_i T_i - pT. \tag{4.19}$$

Nulstellen van de partiële afgeleiden leidt tot

$$\lambda^* = \frac{n}{(1 - p)T} \tag{4.20}$$

en

$$1 - p_i^* = \frac{(1 - p)T}{n} \quad \frac{n_i}{T_i}, \tag{4.21}$$

als λ^* en p_i^* de waarden van λ en p_i zijn, waarvoor het maximum wordt bereikt. De waarde van het maximum is

$$\left\{ \frac{(1 - p)T}{n} \right\}^n \prod_{i=1}^{r} \left(\frac{n_i}{T_i} \right)^{n_i}. \tag{4.22}$$

Passen wij vervolgens de methode van LAGRANGE toe op de logaritme van dit maximum en bijvoorwaarde (4.13), dan blijken de optimale waarden

$n_{i,0}$ van de n_i gelijk te zijn aan *)

$$n_{i,0} = \frac{T_i}{T}\, n \tag{4.23}$$

en de waarde van (4.15)

$$\min_{n_i} \max_{p_i} \prod_{i=1}^{r} (1 - p_i)^{n_i} = (1 - p)^n. \tag{4.24}$$

Het minimum van $\prod_{i=1}^{r} (1 - p_i)^{n_i}$ onder de bijvoorwaarde (4.13) als funktie van de n_i kan niet op de boven beschreven wijze met behulp van de multiplicatorenmethode gevonden worden (waarom niet?). Men kan echter direct inzien dat $\prod_{i=1}^{r} (1 - p_i)^{n_i}$ minimaal is, wanneer men dié n_i gelijk kiest aan n, waarvoor de bijbehorende p_i de grootste is en de andere n_i gelijk aan 0 zijn. Dus

$$\min_{n_i} \prod_{i=1}^{r} (1 - p_i)^{n_i} = (1 - p_i^{**})^n, \tag{4.25}$$

als

$$p_i^{**} = \max_i p_i. \tag{4.26}$$

De funktie $(1 - p_i^{**})^n$ is maximaal als p_i^{**} minimaal is. Daar bovendien aan (4.12) en (4.26) moet zijn voldaan geldt, dat dit maximum wordt bereikt voor $p_i^{**} = p$, dus als alle p_i de waarde p bezitten. Verder is

---

*) Strikt genomen kan naar de n_i niet worden gedifferentieerd, omdat deze alleen discrete waarden kunnen aannemen. Men kan de variabelen n_i echter vervangen door continu variërende variabelen x_i en dan de optimale waarden $x_{i,0}$ van de x_i bepalen. De optimale waarden van de n_i zijn dan hetzij de grootste gehele waarden $\leq x_{i,0}$, hetzij de kleinste gehele waarden $\geq x_{i,0}$ (vgl. ook deel 1, § 10). Ook dient men steeds na te gaan of de gezochte minima resp. maxima werkelijk gevonden worden; dit brengt echter geen moeilijkheden met zich mee.

$$\max_{p_i} \min_{n_i} \prod_{i=1}^{r} (1 - p_i)^{n_i} = (1 - p)^{n} . \tag{4.27}$$

Aangezien de min max en de max min van de verliesfunktie (4.14) aan elkaar gelijk zijn, bezit dit spel zuivere optimale strategieën : de controleur moet de steekproef over de deelpopulaties verdelen naar evenredigheid van hun omvang; de fraudeur moet overal dezelfde fractie frauderen. De waarde van het spel is $(1 - p)^n$, m.a.w. onafhankelijk van de wijze waarop de fraudeur de fraude verdeelt, is de kans een fraude pT niet te ontdekken hoogstens $(1 - p)^n$ (de kans die men ook zou lopen wanneer er één aselecte steekproef uit de gehele populatie genomen zou worden). Speelt de fraudeur niet de max min strategie, dan is de kans niets te zien kleiner dan de oude kans $(1 - p)^n$. Voor de fraudeur gelden soortgelijke overwegingen.

Opmerking 4.3

Uit formule (4.23) blijkt dat men T moet kennen om $n_{i,0}$ te berekenen. Deze waarde staat echter aan het begin van het jaar nog niet vast en moet derhalve geschat worden. Is deze schatting T' en werkt men hiermee gedurende het gehele jaar, dan is het totale aantal trekkingen dat men doet gelijk aan

$$\sum_{i=1}^{r} \frac{T_i}{T'} n = \frac{T}{T'} n. \tag{4.28}$$

Wil men derhalve aan de veilige kant zitten, dan dient men T' niet te hoog te schatten.

Voorbeeld 4.1

Bij een handelsonderneming wordt het totale bedrag aan inkoopfakturen voor een bepaald jaar geschat op ƒ 1.000.000,-. De accountant stelt voor de controle als eis, dat bij een foutentotaal van 0,5% of meer van dit bedrag de kans hiervan niets in de steekproef te zien hoogstens 1% mag bedragen. De steekproef zal worden verdeeld over 4 perioden.

Uit tabel 4.1 volgt dat gedurende het jaar in totaal 919 guldens gecontroleerd moeten worden. In de loop van het jaar blijken de deelpopulaties gelijk te zijn aan *f* 250.000,-, *f* 300.000,-, *f* 200.000,- en *f* 350.000,-. Toepassing van formule (4.23) leidt tot

$$
\begin{aligned}
n_{1,0} &= \frac{250.000}{1.000.000}\ 919 = 229{,}75, \\
n_{2,0} &= \frac{300.000}{1.000.000}\ 919 = 275{,}70, \\
n_{3,0} &= \frac{200.000}{1.000.000}\ 919 = 183{,}90, \\
n_{4,0} &= \frac{350.000}{1.000.000}\ 919 = 321{,}65.
\end{aligned}
\tag{4.29}
$$

Om aan de veilige kant te blijven, worden deze waarden naar boven afgerond en neemt men respectievelijk steekproeven van de omvang 230, 276, 184 en 322; in totaal worden dus 1012 guldens aangewezen.

5. De methode van de kleinste spijt.

De strekking van deze paragraaf wijkt enigszins af van het overige gedeelte van de leergang. Worden daar steeds methoden behandeld teneinde ze tegelijkertijd met vrucht te kunnen gebruiken, het doel van deze paragraaf is slechts te waarschuwen tegen toepassing van de te bespreken methode. Zou dit nooit ten onrechte geschieden, dan zou deze paragraaf niet geschreven zijn.

Naast het in §2 genoemde minimaxkriterium is een hieraan nauw verwant kriterium in de literatuur voorgesteld, het zogenaamde "minimax-regret"-kriterium of wel het kriterium van de kleinste spijt. Stel dat de natuur als tegenspeler de beschikking heeft over drie mogelijkheden, n.l. de toestanden S_1, S_2 en S_3. De statisticus kan kiezen uit vier alternatieven (f_1, f_2, f_3 en f_4), terwijl de winstfunktie gegeven is in tabel 5.1.*)

---

*) Deze tabel is ontleend aan D. VAN DANTZIG, Statistical Priesthood (Savage on personal probabilities), Statistica Neerlandica, 11, (1957), p. 1-16; rapport SP 51 van de afd.Math.Stat.van het Mathematisch Centrum.

Hierin is a een positief getal, dat groot is ten opzichte van 1.

Bij de methode van de kleinste spijt wordt de minimaxgedachte niet

Tabel 5.1

De winstfunktie

I \ II	f_1	f_2	f_3	f_4
S_1	0	-a	-2a	-1
S_2	0	-a	a	a-1
S_3	0	2a	a	2a-1

toegepast op de winstfunktie zelf, doch op een andere funktie, die hieruit als volgt wordt afgeleid. Wanneer de natuur "strategie" S_1 toepast, dan krijgt de statisticus de grootst mogelijke winst als hij f_3 speelt. Kiest hij een andere strategie, dan ontvangt hij minder dan mogelijk geweest zou zijn en dit verschil kan men zien als de spijt, die de statisticus heeft door het niet kiezen van de optimale handeling. In dit geval bedraagt de spijt bij de verschillende strategieën resp.

2a , -a , 0 en 2a-1.

Ook voor de andere strategieën van de natuur kan deze spijt berekend worden; de resulterende spijtfunktie is opgenomen in tabel 5.2. Bij de methode van de kleinste spijt wordt de minimaxmethode toegepast op

Tabel 5.2

De spijtfunktie behorende bij tabel 5.1

I \ II	f_1	f_2	f_3	f_4
S_1	2a	a	0	2a-1
S_2	a	0	2a	2a-1
S_3	0	2a	a	2a-1

deze spijtfunktie: men minimaliseert dus de maximaal mogelijke spijt. In deze zin is de optimale strategie voor de statisticus dus f_4. Past men de minimaxmethode toe op de winstfunktie zelf, dan is de optimale strategie voor de statisticus f_1. Evenals in dit geval leidt de methode van de kleinste spijt in het algemeen tot andere optimale strategieën dan de minimaxmethode.

Hoewel de methode op het eerste gezicht niet zonder meer verwerpelijk lijkt, kan er toch ernstige kritiek op worden uitgeoefend (vgl.bijv. het op blz. 47 genoemde artikel van D. VAN DANTZIG). Vergelijkt men in tabel 5.1 strategie f_1 met strategie f_4, dan ziet men dat f_4 alleen iets beter is in situatie S_1, maar dat f_4 in de situaties S_2 en S_3 veel slechter is dan f_1. De methode van de kleinste spijt dwingt de statisticus in dit voorbeeld dus tot het lopen van grote risico's teneinde in een bepaald geval een kleine winst te krijgen. Bovendien is strategie f_4, als situatie S_1 zich voordoet, ongeveer het slechtste wat men doen kan, aangezien dan f_2 en f_3 veel betere strategieën zijn en f_1 slechts weinig minder goed dan f_4. Met andere woorden: men loopt enorme risico's om in een bepaald geval een kleine winst te maken, maar als dat geval zich voordoet, heeft men ongeveer het slechtste gedaan, wat men doen kan!

Een andere merkwaardige eigenschap komt aan het licht, wanneer wij de statisticus de beschikking geven over een vijfde alternatief f_5. Laat de winstfunktie overgaan in de in tabel 5.3 opgegeven funktie. De bijbehorende spijtfunktie is opgegeven in tabel 5.4.

Tabel 5.3

De uitgebreide winstfunktie

I \ II	f_1	f_2	f_3	f_4	f_5
S_1	0	-a	-2a	-1	a
S_2	0	-a	a	a-1	0
S_3	0	2a	a	2a-1	-a-1

Tabel 5.4

De spijtfunktie behorende bij tabel 5.3

I \ II	f_1	f_2	f_3	f_4	f_5
S_1	2a	a	0	2a-1	3a
S_2	a	0	2a	2a-1	a
S_3	a+1	3a+1	2a+1	3a	0

Minimaliseert men hier de maximale spijt, dan wordt strategie f_1 aangewezen. Dit betekent derhalve: heeft men de keuze tussen vier strategieën f_1 t/m f_4, dan kiest men f_4, maar wordt het arsenaal uitgebreid met een extra strategie f_5, dan kiest men niet de oude optimale strategie f_4 noch de nieuwe mogelijkheid f_5, doch strategie f_1; d.w.z. een strategie die men vroeger ook al ter beschikking had. Een niet gewenste strategie beïnvloedt dus de keuze tussen de andere strategieën.

Het laatste blijkt ook duidelijk, wanneer wij de winstfunktie voor strategie f_5 wijzigen. Kiest men respectievelijk de waarden a, -a en 0, dan vindt men voor de spijtfunktie de in tabel 5.5a opgegeven waarden. Brengt men een kleine verandering aan door de waarden a, -a, -2 in te vullen (a groot t.o.v. 1), dan vindt men tabel 5.5b. In het eerste geval blijft f_4 de optimale strategie, terwijl in het tweede geval f_1 en f_3 optimale strategieën zijn.

Tabel 5.5

Gewijzigde spijtfunkties

a

I \ II	f_1	f_2	f_3	f_4	f_5
S_1	2a	a	0	2a-1	3a
S_2	a	0	2a	2a-1	0
S_3	0	2a		2a-1	0

b

I \ II	f_1	f_2	f_3	f_4	f_5
S_1	2a	a	0	2a-1	3a
S_2	a	0	2a	2a-1	0
S_3	2	2a+2	a+2	2a+1	0

Een relatief kleine verandering in de niet ter zake dienende strategie f_5 beïnvloedt derhalve de bepaling van de optimale strategie, welke bepaling daardoor afhankelijk wordt van irrelevante alternatieven.

Uit het voorgaande is duidelijk geworden dat de methode van de kleinste spijt geen gelukkige oplossing biedt voor zogenaamde statistische spelen. Ook voor strategische spelen kan het als kriterium niet verdedigd worden.

Het kriterium is voorgesteld door L.J. SAVAGE *), die opmerkt dat het in sommige statistische spelen minder pessimistisch is dan het minimaxkriterium, maar die het vooral verdedigt voor de volgende van het voorafgaande afwijkende situaties. Een groep mensen (bijv. een jury of een directie) moet uit een aantal alternatieven kiezen. Het nut dat de i^{de} persoon toekent aan het j^{de} alternatief zij $U_i(j)$. Voor één of meer alternatieven zal dit maximaal zijn; wordt een ander alternatief uiteindelijk aangewezen, dan heeft het i^{de} lid een bepaalde spijt, daar het zijns inziens optimale niet bereikt is. Wanneer men uiteindelijk gezamenlijk één keuze moet doen, zal men, indien de meningen uiteenlopen, tot een compromis moeten komen. SAVAGE stelt nu voor dát alternatief als compromis te kiezen, waarvoor het maximum van de "spijten" die de leden van de groep hebben, minimaal is. De keuze wordt dus bepaald door de minimaxmethode toe te passen op de funktie

$$\max_k U_i(k) - U_i(j). \tag{5.1}$$

Ook hierop kan men kritiek uitoefenen.

De conclusie is derhalve, dat het niet raadzaam is de methode van de kleinste spijt bij statistische en strategische spelen toe te passen, terwijl men ook in andere situaties zeer voorzichtig dient te zijn.

Voor een verdere bespreking van het voor en tegen van verschillende beslissingskriteria verwijzen wij naar het op blz. 7 reeds genoemde boek van LUCE en RAIFFA (speciaal hoofdstuk 13).

---

*) L.J. SAVAGE, The Foundations of Statistics, John Wiley and Sons, New York, (1954).

Literatuur

J. VON NEUMANN and O. MORGENSTERN, Theory of Games and Economic Behaviour, Princeton University Press, Princeton; 1e druk (1944); 2e druk (1947).

R.D. LUCE and H. RAIFFA, Games and Decisions, John Wiley and Sons, New York, (1957).

J.D. WILLIAMS, Speltheorie, Het Spectrum, Utrecht, (1966).

H. CHERNOFF and L.E. MOSES, Elementary Decision Theory, John Wiley and Sons, New York, (1959).

DEEL 8

HOOFDSTUK 2

NETWERKPLANNING DOOR F. GÖBEL

1. Inleiding; het maken van een netwerk.

De term Netwerkplanning wordt hier gebruikt als samenvattende benaming voor technieken als C P M , PERT , LESS , enz. *) . Dit zijn, kort gezegd, technieken die de leiding van bijv. een bedrijf informatie verstrekken, waarmee ten aanzien van bepaalde projecten beslissingen kunnen worden genomen. Om deze vage omschrijving nader te preciseren zullen wij nagaan, waaraan een project moet voldoen, opdat een of andere vorm van netwerkplanning kan worden toegepast. Later zal de hierboven genoemde informatie ter sprake komen.

Het project moet kunnen worden verdeeld in een aantal onderdelen of deelbewerkingen **) , waartussen zekere ordeningsrelaties zijn gedefinieerd. Dat wil o.a. zeggen, dat voor iedere activiteit A is vastgesteld

1) welke activiteiten eraan vooraf dienen te gaan, en
2) aan welke activiteiten pas kan worden begonnen als A is voltooid.

Projecten, waarbij activiteiten voorkomen, die weliswaar niet tegelijkertijd kunnen plaatsvinden, maar overigens wel in een willekeurige volgorde kunnen worden uitgevoerd, voldoen niet aan deze eis, en komen niet in aanmerking voor een netwerkplanningmethode.

Wanneer aan bovengenoemde eis is voldaan, is het meestal mogelijk de ordeningsrelaties grafisch voor te stellen door middel van een netwerk. Een activiteit wordt hierbij voorgesteld door een pijl, zodanig dat begin- en eindpunt van de pijl corresponderen met begin resp. voltooiing van de activiteit. Hierbij kunnen zich bepaalde moeilijkheden voordoen, zoals uit het volgende voorbeeld blijkt.

Stel dat een project uit vier activiteiten A, B, C en D bestaat, waarbij A en B vooraf moeten gaan aan C, en bovendien B aan D. Wanneer men deze ordeningsrelaties grafisch zou voorstellen zoals in fig. 1.1 is gedaan, dan leest men hieruit af, dat A vooraf moet gaan aan D, wat niet de bedoeling is.

---

*) C P M = Critical Path Method
PERT = Program Evaluation Research Task; later is aan deze letters de betekenis "Program Evaluation and Review Technique" gegeven.
LESS = Least Cost Estimating and Scheduling

**) Deze onderdelen of deelbewerkingen worden gewoonlijk aangeduid met "karweien" of "activiteiten". Wij zullen de laatste term gebruiken.

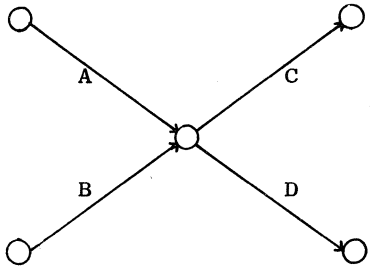


fig. 1.1

A en B gaan vooraf aan C en D

Deze moeilijkheid kan worden opge ost door het invoeren van een dummy-activiteit E, zoals in fig. 1.2 is gebeurd.

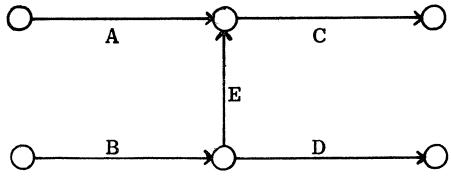


fig. 1.2

A gaat niet noodzakelijk vooraf aan D

Het is verder gebruikelijk om het netwerk zo te tekenen dat er slechts één beginknooppunt (oorsprong) en slechts één eindknooppunt (voltooiing is, en dat bovendien iedere activiteit eenduidig is bepaald door zijn begin en eind. Deze eisen vormen geen essentiële restricties voor het project; door eventueel weer dummy-activiteiten in te voeren kan er gemakkelijk aan worden voldaan Als bijv. ergens in een netwerk de in fig 1.3 getekende situatie optreedt, kan met behulp van de dummy H worden bereikt dat tussen ieder tweetal knooppunten niet meer dan één pijl loopt (zie fig. 1.4)

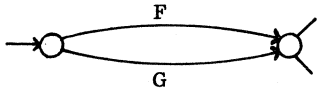


fig. 1.3

F en G hebben hetzelfde beginpunt en hetzelfde eindpunt

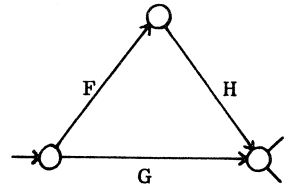


fig. 1.4

Iedere activiteit is door zijn begin- en eindpunt bepaald

In de praktijk zal het bij veel projecten voorkomen dat bewerkingen elkaar kunnen overlappen, in die zin dat aan B pas kan worden begonnen als aan A begonnen is, terwijl de voltooiing van A niet noodzakelijk vooraf behoeft te gaan aan het begin van B. Hoewel dergelijke projecten niet aan de gestelde eis voldoen, kan de volgende kunstgreep soms uitkomst brengen: A wordt verdeeld in deelactiviteiten A_1 en A_2, en het netwerk ziet er ter plaatse uit als in fig. 1.5.

Vaak gaat een dergelijke situatie gepaard met de eis dat A niet mag worden "ingehaald" door B. In het netwerk kan deze eis worden benaderd, zoals in fig. 1.6 is gedaan.

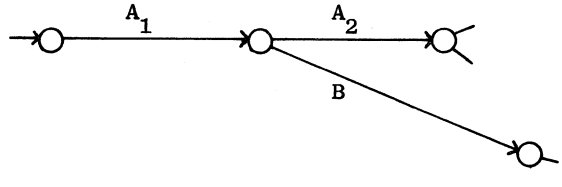


fig. 1.5

Grafische voorstelling van overlappende activiteiten

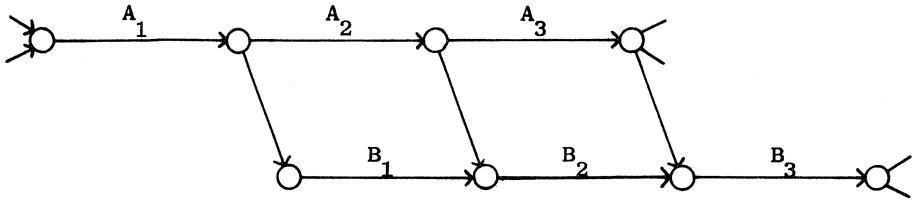


fig. 1.6

B mag A niet inhalen

Het is echter niet prettig wanneer een netwerk door herhaalde toepassing van deze kunstgreep moet worden gecompliceerd. Projecten waarbij deze situatie vaak optreedt, kunnen dan ook beter niet door middel van een netwerk worden gepland.

We merken tenslotte op dat in het netwerk nooit een cykel ("loop") kan optreden. Als bijv. voor twee activiteiten A en B geldt, dat A voorafmoet gaan aan B, en B aan A, zou men immers nooit aan dit gedeelte van het netwerk kunnen beginnen.

We hebben tot nu toe niets gezegd over de tijdsduren van de activiteiten, noch over de tijdstippen waarop de knooppunten plaatsvinden. Inderdaad houdt men bij C P M , PERT, e.d., de "planning", waarbij de ordeningsrelaties tussen de activiteiten worden vastgesteld, en de "scheduling", waarbij aan de knooppunten tijdstippen worden toegekend, scherp gescheiden. In de volgende paragraaf zal enige aandacht worden besteed aan de "scheduling".

2. Het kritieke pad.

Wanneer het netwerk is opgesteld, kan een begin worden gemaakt met de scheduling-fase van het project. Het is hierbij nuttig de knooppunten te nummeren van 1, ..., n, en wel zodanig dat $i < j$, indien i vooraf gaat aan j, d.w.z. indien er een geordende reeks van activiteiten bestaat van i naar j. Men ziet gemakkelijk in, dat deze

wijze van nummeren altijd mogelijk is *). Iedere activiteit is nu eenduidig bepaald door het paar (i,j). Omgekeerd behoeft niet bij ieder paar (i,j) met $i < j$ een activiteit te behoren, zodat de paren (i,j), die met activiteiten van het netwerk corresponderen, een deelverzameling vormen van de $\frac{1}{2}n(n-1)$ mogelijke paren (i,j) met $1 \leq i < j \leq n$. Een project kan dus worden opgevat als een deelverzameling $\mathcal{P}$ van de verzameling van alle paren (i,j).

We voegen nu aan de oorsprong het tijdstip 0 toe en nemen verder voorlopig aan dat de duur van iedere activiteit (i,j) een gegeven getal y_{ij} is.

Het is duidelijk dat aan de activiteiten die een gegeven knooppunt i als beginpunt hebben, pas kan worden begonnen als alle activiteiten die i als eindpunt hebben, zijn voltooid. Het vroegste moment e_i waarop dit kan gebeuren, wordt gegeven door

$$(2.1)\qquad \begin{cases} e_1 = 0, \\ e_i = \max \left\{ e_k + x_{ki} \;\middle|\; (k,i) \in \mathcal{P} \right\} \text{ voor } i = 2, \ldots, n. \end{cases}$$

Verder moet, indien een bepaald tijdstip T is gegeven waarop het project uiterlijk moet zijn voltooid, het knooppunt i uiterlijk zijn bereikt op het tijdstip f_i met

$$(2.2)\qquad \begin{cases} f_n = T, \\ f_i = \min \left\{ f_k - x_{ik} \;\middle|\; (i,k) \in \mathcal{P} \right\} \text{ voor } i = 1, \ldots, n-1. \end{cases}$$

Uit (2.1) volgt dat activiteit (i,j) niet eerder kan beginnen dan op tijdstip e_i, en niet voltooid kan zijn voor $e_i + x_{ij}$. Ook moet (i,j) uiterlijk zijn voltooid op f_j, en dus uiterlijk beginnen op $f_j - x_{ij}$. **)

---

*) Bijv. door met de volgende algorithme een oplossing te construeren:
a) geef de oorsprong nummer 1;
b) laat alle pijlen weg die in genummerde knooppunten ontspringen. Neem een willekeurig knooppunt waarop geen pijl uitkomt (een dergelijk knooppunt is altijd aanwezig, omdat er geen cykels zijn), en geef dit punt het laagste ongebruikte nummer;
c) pas stap b toe totdat de voltooiing genummerd is.

**) e_i en f_i worden in het Engels "earliest event time" resp. "latest event time" genoemd. Worden deze tijdstippen beschouwd met betrekking tot een activiteit, dan worden de namen "earliest start time" en "latest completion time" gebruikt. Verder heet $e_i + x_{ij}$ "earliest completion time" en $f_j - x_{ij}$ "latest start time".

Voor (i,j) is dus een tijdsinterval van de lengte $f_j - e_i$ beschikbaar en een speelruimte

(2.3) $$f_j - e_i - x_{ij} = \tau_{ij} .$$

Zo komen we tot de volgende definitie :

Definitie 2.1

Totale speelruimte ("total float"):

$$f_j - e_i - x_{ij} = \tau_{ij} .$$

Verder kent men de volgende soorten speelruimte:

Definitie 2.2

Vrije speelruimte ("free float"):

(2.4) $$e_j - e_i - x_{ij} = \tau^{v}_{ij} .$$

Onder de vrije speelruimte van een activiteit verstaan we de speelruimte die deze activiteit heeft, als alle andere activiteiten zo vroeg mogelijk beginnen.

Definitie 2.3

Onafhankelijke speelruimte ("independent float"):

(2.5) $$\max \{0, e_j - f_i - x_{ij} \} = \tau^{o}_{ij} .$$

De onafhankelijke speelruimte van een activiteit is de speelruimte die voor die activiteit overblijft ongeacht de begintijdstippen van de overige activiteiten.

Definitie 2.4

Afhankelijke speelruimte ("dependent float","interfering float"):

(2.6) $$f_j - e_j = \tau^{a}_{ij} .$$

De afhankelijke speelruimte van een activiteit is a.h.w. het complement van de vrije speelruimte en er geldt

(2.7) $$\tau^{a}_{ij} = \tau_{ij} - \tau^{v}_{ij}.$$

Van de drie laatstgenoemde soorten speelruimte is de vrije speelruimte de belangrijkste.

Een activiteit wordt kritiek genoemd als $\tau_{ij} = 0$. Het is eenvoudig in te zien, dat alleen dan kritieke activiteiten voorkomen als $e_n = T$. Is dit het geval, dan zijn er ook één of meer ketens van opeenvolgende activiteiten aan te wijzen, beginnend in knooppunt 1 en eindigend in n, waarvoor alle τ_{ij} nul zijn. Een dergelijke keten heet een kritiek pad van het netwerk.

Voorbeeld 2.1

Beschouw het in fig. 1.7 afgebeelde netwerk. Bij de pijlen zijn de activiteitsduren x_{ij} geplaatst; de knooppunten met hun nummers zijn door omcirkelde getallen aangegeven.

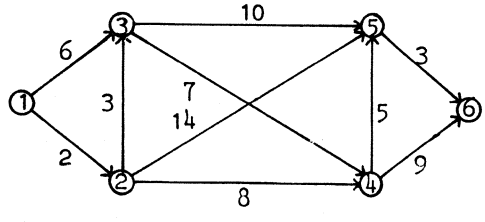


fig. 1.7

Gegevens bij voorbeeld 2.1

Met formule (2.1) vindt men gemakkelijk, dat

(2.8)
$$\begin{cases} e_1 = 0, \\ e_2 = 2, \\ e_3 = 6, \\ e_4 = 13, \\ e_5 = 18, \\ e_6 = 22, \end{cases}$$

zodat een kritiek pad alleen voorkomt als T = 22. Uitgaande van deze waarde voor T vinden we met formule (2.2)

$$(2.9)\qquad \begin{cases} f_1 = 0, \\ f_2 = 3, \\ f_3 = 6, \\ f_4 = 13, \\ f_5 = 19, \\ f_6 = 22. \end{cases}$$

Met formule (2.3) kan nu voor iedere activiteit de totale speelruimte worden berekend. Het resultaat hiervan is

$$(2.10)\qquad \begin{cases} \tau_{12} = 1, & \tau_{25} = 3, \\ \tau_{13} = 0, & \tau_{35} = 3, \\ \tau_{23} = 1, & \tau_{45} = 1, \\ \tau_{24} = 3, & \tau_{46} = 0, \\ \tau_{34} = 0, & \tau_{56} = 1. \end{cases}$$

Het kritieke pad bestaat dus uit de activiteiten (1,3), (3,4) en (4,6).

Blijkens de definitie van kritiek pad geldt voor de activiteiten die ertoe behoren, dat uitstel of vertraging van deze activiteiten een even grote verschuiving van het voltooiingstijdstip met zich meebrengt. Vanzelfsprekend zullen daarom de kritieke activiteiten door de leiding zorgvuldig in het oog moeten worden gehouden.

We zijn hiermee aangeland bij de eerste vorm van informatie die de netwerktechniek aan de leiding kan verschaffen: een netwerk, waarin is aangegeven, welke activiteiten kritiek, en bij voorkeur ook welke bijna-kritiek zijn (d.w.z. een betrekkelijk kleine τ_{ij} hebben). Zoals we later zullen zien, kunnen ook de activiteiten met een zeer grote speelruimte belangrijk zijn. Zelfs in deze zeer eenvoudige vorm kan netwerkplanning van nut zijn, afgezien nog van het inzicht in de samenhang van het project, dat reeds bij het maken van het netwerk wordt verkregen.

Uit het voorgaande blijkt dat het voortdurend bijwerken van de gegevens, naarmate het project vordert, van groot belang kan zijn. Immers, een activiteit die een grote τ_{ij} heeft bij de aanvang van het project, kàn kritiek worden als de eraan voorafgaande bewerkingen steeds maar worden uitgesteld. Wiskundig gezien levert het herplannen geen moeilijkheden op. Er ontstaat namelijk een nieuw netwerk, dat zich op dezelfde wijze laat behandelen als het oorspronkelijke. Wel kunnen zich natuurlijk allerlei organisatorische en administratieve problemen voordoen. Deze worden hier echter buiten beschouwing gelaten.

3. Uitbreidingen.

In de vorige paragraaf zijn (soms impliciet) bepaalde veronderstellingen gemaakt, waaraan echter in de praktijk niet altijd is voldaan. In dat geval is het nodig het model iets ingewikkelder te maken.

Als bijv. de onzekerheid in de duur der activiteiten groot is, laten de tijden zich beter beschrijven als stochastische variabelen. In het bijzonder is deze uitbreiding van het model vaak nodig bij een project dat voor het eerst wordt uitgevoerd. Dan is het dus noodzakelijk om de kansverdeling van de activiteitsduren te schatten. In §4 zal o.a. worden geschetst hoe dit in het zgn. PERT-systeem geschiedt.

Het kan ook gebeuren, dat de x_{ij} niet alleen zeer nauwkeurig bekend zijn, maar zelfs nog naar wens kunnen worden geregeld. Men kan bijv. de situatie hebben dat tegen hogere kosten (of ten koste van de kwaliteit van het eindprodukt) de activiteitsduur kan worden verminderd. We denken bijvoorbeeld aan het uitbesteden van kritieke activiteiten. In §5 en §6 zal een systeem worden besproken (de "Critical Path Method") dat sterk de nadruk legt op dit aspect.

Bij sommige research- en ontwikkelingsprojecten is het onmogelijk een netwerk te maken, omdat de gedaante ervan na een bepaalde

activiteit nog kan afhangen van het resultaat van die activiteit. Er is een poging gedaan om deze moeilijkheid op te lossen door het invoeren van een algemener soort netwerk, waarin knooppunten optreden met de eigenschap, dat niet alle eruit ontspringende activiteiten voltooid behoeven te worden.*) Gezien echter het speculatieve karakter van sommige daarbij gemaakte veronderstellingen zullen wij er niet verder op ingaan.

Tenslotte kan het voorkomen, dat men bij de planning rekening moet houden met de beperktheid van "hulpbronnen" (als men bijv. de beschikking heeft over 3 lassers, kunnen twee activiteiten die ieder 2 lassers vereisen, niet tegelijkertijd plaatsvinden). Aan dit punt zal in §7 enige aandacht worden besteed.

4. De "Program Evaluation and Review Technique".

In het PERT-systeem wordt dus verondersteld dat de activiteitsduren x_{ij} een bekende kansverdeling hebben. Men gaat bij het bepalen hiervan als volgt te werk. Aan de daarvoor in aanmerking komende personen wordt gevraagd een optimistische, een waarschijnlijkste en een pessimistische schatting van de activiteitsduur te geven. Stel dat deze schattingen a,m resp. b zijn. Men neemt nu aan dat de grootheid

$$\underline{x} = \frac{\underline{x}_{ij} - a}{b-a} \tag{4.1}$$

een bêta-verdeling heeft

$$P[\underline{x} \leq x] = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\,\Gamma(\beta)} \int_0^x v^{\alpha-1}(1-v)^{\beta-1}dv, \tag{4.2}$$

(vgl. deel 2, blz. 56, form. (7.28)).

Deze aanname berust niet op waarnemingen van activiteitsduren, maar heeft als enige redenen, dat de bêta-verdeling een eindige range heeft en dat de parameters α en β zo gekozen kunnen worden, dat de

---

*) H. EISNER, A Generalized Network Approach to the Planning and Scheduling of a Research Project, J.O.R.S.A., 10, (1962), p.115-125.

dichtheid eentoppig is *). Dit laatste wordt dan ook gedaan en wel zodanig dat de modus van x gelijk is aan $\frac{m-a}{b-a}$. Uit (4.2) leidt men gemakkelijk af, dat dan de volgende betrekking tussen α en β geldt

$$\alpha(b-m) - \beta(m-a) = a + b - 2m. \tag{4.3}$$

Aangezien de bêta-verdeling nu nog niet vastligt, maakt men nog een veronderstelling, welke a.v. luidt:

$$\sigma(\underline{x}_{ij}) = \frac{1}{6}(b-a). \tag{4.4}$$

Deze betrekking is equivalent met

$$\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} = \frac{1}{36}. \tag{4.5}$$

De verdeling is nu bepaald en de verwachting kan worden berekend. Hiervoor is o.a. het oplossen van de derdegraads-vergelijking (4.5) vereist. Gezien de benaderingen die al zijn gemaakt, verdient het de voorkeur dit oplossen evenmin exact te doen, maar gebruik te maken van de volgende formule

$$\mathcal{E}(\underline{x}_{ij}) \approx \frac{1}{6}(a+4m+b), \tag{4.6}$$

die, zoals uit numerieke berekeningen is gebleken, een redelijke benadering is.

Met behulp van deze gegevens kan men nu trachten om bijv. de kansverdeling van $\underline{e}_i$, het vroegste voltooiingstijdstip van knooppunt i, te berekenen. Het zal duidelijk zijn dat dit voor een netwerk van enige omvang een verre van eenvoudige taak is. Er zijn echter verschillende benaderingsmethoden bekend, waarvan in deze paragraaf de eenvoudigste zal worden behandeld.

Hierbij wordt allereerst de <u>verwachting</u> van het tijdstip waarop knooppunt i zal worden bereikt, benaderd door de grootheid $\hat{e}_i$, die als volgt gedefinieerd wordt **):

---

*) C.E.CLARK, The PERT-model for the Distribution of an Activity Time, J.O.R.S.A., <u>10</u>, (1962), p. 405-406.

**) De hier gebruikte notatie wijkt af van die in deel 3 (§ 4), waar een accent circumflex een meest aannemelijke schatter aanduidt.

$$(4.7)\qquad \begin{cases} \hat{e}_1 = 0, \\ \hat{e}_i = \max\{\hat{e}_k + \mathcal{E}(\underline{x}_{ki}) \mid (k,i) \in \mathcal{P}\} \quad \text{voor } i = 2, \ldots, n. \end{cases}$$

Iedere activiteitsduur wordt dus vervangen door zijn verwachting en op het resulterende netwerk wordt formule (2.1) toegepast. Dat (4.7) inderdaad een benaderingsformule is, blijkt uit het volgende eenvoudige voorbeeld.

Stel dat het netwerk de gedaante heeft die in fig. 4.1 is afgebeeld, en laat voldaan zijn aan

$$(4.8)\qquad \begin{cases} x_{12} = 0, \\ x_{13} = 2, \\ P\left[\underline{x}_{23} = 1\right] = P\left[\underline{x}_{23} = 4\right] = \frac{1}{2}. \end{cases}$$

Volgens (4.7) geldt nu $\hat{e}_3 = 2\frac{1}{2}$, terwijl $\mathcal{E}\underline{e}_3 = \frac{1}{2}\max(0+1,2) + \frac{1}{2}\max(0+4,2) = 3$.

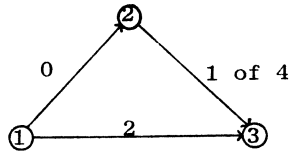


fig. 4.1

Een eenvoudig netwerk met $\hat{e}_3 \neq \mathcal{E}\underline{e}_3$

Voor de variantie van $\underline{e}_i$ gebruikt men de volgende benadering

$$(4.9)\qquad \hat{\sigma}^2(\underline{e}_i) = \sum \sigma^2(\underline{x}_{jk}),$$

waarbij gesommeerd wordt over alle activiteiten die liggen in een langste pad van 1 naar i in het corresponderende deterministische netwerk (waarin dus iedere activiteitsduur is vervangen door zijn verwachting).

In het hierboven genoemde voorbeeld is $\sigma^2(\underline{e}_3)$ gelijk aan 1, terwijl met (4.9) wordt gevonden $\hat{\sigma}^2(\underline{e}_3) = 2{,}25$. Blijkbaar is de be-

nadering tamelijk grof. Bovendien hangt de gevonden waarde nog af van de detaillering in het netwerk *).

Verder wordt in het PERT-systeem aangenomen, dat de grootheden $\underline{e}_i$ voor niet te kleine waarden van i bij benadering normaal verdeeld zijn. De verdeling van $\underline{e}_i$ ligt nu vast en men heeft de mogelijkheid om uitspraken van de vorm

$$P[\underline{e}_i \leq E_i] = p \tag{4.10}$$

te doen, waarin E_i ontleend is aan een gegeven "scheduling" van het project. Als we hun betrouwbaarheid even buiten beschouwing laten, zijn deze waarschijnlijkheidsuitspraken ("probability statements") vanzelfsprekend waardevolle inlichtingen voor de leiding.

Geheel analoog kan men $\mathcal{E}(\underline{f}_i)$, de verwachting van het tijdstip waarop knooppunt i uiterlijk moet zijn bereikt, bij gegeven T benaderen met

$$\begin{cases} \hat{f}_n = T, \\ \hat{f}_i = \min\{\hat{f}_k - \mathcal{E}(\underline{x}_{ik}) \mid (i,k) \in \mathcal{P}\} \quad \text{voor } i = 1,\ldots,n-1, \end{cases} \tag{4.11}$$

en de variantie van $\underline{f}_i$ met

$$\hat{\sigma}^2(\underline{f}_i) = \sum \sigma^2(\underline{x}_{jk}), \tag{4.12}$$

waarbij wordt gesommeerd over alle activiteiten die liggen in een langste pad van i naar n in het corresponderende deterministische netwerk. Ook van de grootheden $\underline{f}_i$ wordt aangenomen dat ze bij benadering normaal verdeeld zijn (voor niet te grote waarden van i) en ook hier kunnen dus waarschijnlijkheidsuitspraken worden gedaan.

Tot slot van deze paragraaf zullen we wat nader ingaan op een verbeterde schatting van $\mathcal{E}\underline{e}_i$, gevonden door FULKERSON **). De defi-

---

*) Zie T.L.HEALY, Activity Subdivision and PERT Probability Statements, J.O.R.S.A., 9, (1961), p.341-348.

**) D.R.FULKERSON, Expected Critical Path Lengths in PERT Networks, J.O.R.S.A., 10, (1962), p. 808-817.

nitie hiervan luidt

$$(4.13)\qquad \begin{cases} e_1^* = 0, \\ e_i^* = \mathcal{E}\ \max\{e_k^* + \underline{x}_{ki} \mid (k,i) \in \mathcal{P}\} \quad \text{voor } i = 2,\ldots,n. \end{cases}$$

We veronderstellen nu, dat de simultane verdeling van de duren van de activiteiten die in i ontspringen, onafhankelijk is van de verdelingen die met de overige knooppunten k ($i \neq k < n$) corresponderen. In het bijzonder is hieraan voldaan als de activiteitsduren $\underline{x}_{ij}$ onafhankelijk zijn. Onder de genoemde veronderstelling geldt

$$(4.14)\qquad \hat{e}_i \leq e_i^* \leq \mathcal{E}\underline{e}_i.$$

De linkerhelft van (4.14) kan men met volledige inductie als volgt bewijzen. Voor i = 1 is de bewering triviaal. Stel dat men al bewezen heeft, dat $\hat{e}_j \leq e_j^*$ voor $j = 1,\ldots,i-1$. Dan geldt

$$\begin{aligned} e_i^* &= \mathcal{E}\ \max\{\ e_k^* + \underline{x}_{ki} \mid (k,i) \in \mathcal{P}\} \geq \\ &\geq \max\ \mathcal{E}\{e_k^* + \underline{x}_{ki} \mid (k,i) \in \mathcal{P}\} = \\ &= \max\ \{e_k^* + \mathcal{E}\underline{x}_{ki} \mid (k,i) \in \mathcal{P}\} \geq \\ &\geq \max\ \{\hat{e}_k + \mathcal{E}\underline{x}_{ki} \mid (k,i) \in \mathcal{P}\} = \hat{e}_i. \end{aligned} \qquad (4.15)$$

De rechterhelft van (4.14) kan men op een dergelijke manier bewijzen

$$(4.16)\qquad \mathcal{E}\underline{e}_i = \mathcal{E}\mathcal{E}\ \max\{\underline{e}_k + \underline{x}_{ki} \mid (k,i) \in \mathcal{P}\}.$$

We hebben hier twee verwachting-tekens ($\mathcal{E}$'s) geschreven, omdat twee stochastische variabelen optreden die verschillend worden behandeld. Eén van deze $\mathcal{E}$'s wordt n.l. met de max-operator verwisseld

$$\begin{aligned} &\mathcal{E}\mathcal{E}\ \max\ \{\underline{e}_k + \underline{x}_{ki} \mid (k,i) \in \mathcal{P}\} \geq \\ &\geq \mathcal{E}\ \max\ \{\mathcal{E}\underline{e}_k + \underline{x}_{ki} \mid (k,i) \in \mathcal{P}\} \geq \\ &\geq \mathcal{E}\ \max\ \{\ e_k^* + \underline{x}_{ki} \mid (k,i) \in \mathcal{P}\} = e_i^*. \end{aligned} \qquad (4.17)$$

We zien dus dat e_i^*, evenals $\hat{e}_i$, altijd aan de optimistische kant is en dat e_i^* nooit slechter is dan $\hat{e}_i$. Uit de getallenvoorbeelden die in FULKERSON's rapport worden gegeven, krijgt men de indruk, dat e_i^* een reële verbetering betekent ten opzichte van $\hat{e}_i$. Hier staat tegenover dat de berekening van e_i^* ingewikkelder is. Het volgende aan FULKERSON ontleende voorbeeld geeft hiervan een indruk.

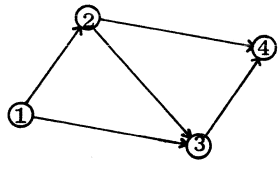


fig. 4.2

Voorbeeld van het berekenen van e_i^*

Stel dat ieder van de activiteitsduren in het in fig. 4.2 getekende netwerk met kans $\frac{1}{2}$ de waarden 0 en 1 aanneemt. Dan geldt volgens (4.13)

$$
(4.18)\qquad \left\{\begin{array}{l}
e_1^* = 0,\\
e_2^* = \frac{1}{2}(e_1^*+0) + \frac{1}{2}(e_1^*+1) = \frac{1}{2},\\
e_3^* = \frac{1}{4}\max(e_1^*+0, e_2^*+0) + \frac{1}{4}\max(e_1^*+0, e_2^*+1) +\\
\qquad + \frac{1}{4}\max(e_1^*+1, e_2^*+0) + \frac{1}{4}\max(e_1^*+1, e_2^*+1) = \frac{9}{8},\\
e_4^* = \frac{1}{4}\max(e_2^*+0, e_3^*+0) + \frac{1}{4}\max(e_2^*+0, e_3^*+1) +\\
\qquad + \frac{1}{4}\max(e_2^*+1, e_3^*+0) + \frac{1}{4}\max(e_2^*+1, e_3^*+1) = \frac{55}{32}.
\end{array}\right.
$$

In dit geval geldt $e_4^* = \mathcal{E}\,\underline{e}_4$, terwijl $\hat{e}_4 = \frac{3}{2}$. Indien de activiteitsduren met kans $\frac{1}{3}$ de waarden 0, 1 en 2 aannemen, geldt

$$
(4.19)\qquad \left\{\begin{array}{l}
\hat{e}_4 = 3,\\
e_4^* = 3{,}22,\\
\mathcal{E}\,\underline{e}_4 = 3{,}32.
\end{array}\right.
$$

Uit het voorgaande blijkt duidelijk dat we bij stochastische activiteitsduren eigenlijk niet meer van een kritiek pad kunnen spreken; ieder pad heeft een bepaalde kans om kritiek te worden.

5. De "Critical Path Method"

Zoals in §3 al is gezegd, wordt bij de Critical Path Method (C P M) in veel grotere mate dan bij PERT rekening gehouden met de kosten die aan het project zijn verbonden, in het bijzonder met de zgn. directe kosten. Deze zijn in het algemeen een dalende funktie van het voltooiingstijdstip s. Hoe kleiner s is, des te meer zal men moeten uitgeven aan overuren, uitbesteding, speciale voorzieningen, e.d. De indirecte kosten verminderen meestal als s kleiner wordt gemaakt. Men denke hier bijv. aan een project,dat tot doel heeft een nieuw produkt op de markt te brengen. Hoe eerder het produkt kan worden verkocht, hoe hoger de opbrengst zal zijn.

Als gewoonlijk tracht men de keuzevariabele (in dit geval s) zo te bepalen, dat de som van beide kostensoorten minimaal is. Bij C P M gaat het er nu om de directe kosten te minimaliseren bij iedere waarde van s, en wel door de activiteitsduren zo gunstig mogelijk te kiezen.

In het model nemen we voorlopig aan dat de directe kosten van iedere activiteit $(i,j)\in\mathcal{P}$ een lineaire niet-stijgende funktie zijn van de duur

(5.1) $$k_{ij}(x_{ij}) = k_{ij} - c_{ij}x_{ij}, \text{ waarin } c_{ij} \geq 0 \text{ is,}$$

en dat x_{ij} alle waarden tussen twee gegeven grenzen kan aannemen

(5.2) $$a_{ij} \leq x_{ij} \leq b_{ij}.$$

Voegen we aan ieder knooppunt i een tijdstip t_i toe, waarbij we zonder beperking $t_1 = 0$ kunnen nemen, dan moet gelden

(5 3) $$x_{ij} \leq t_j - t_i.$$

De totale kosten zijn

(5.4) $$\sum_{\mathcal{P}} k_{ij}(x_{ij})$$

en we moeten x_{ij} en t_i dus zodanig kiezen dat $\sum k_{ij}(x_{ij})$ minimaal is. Dit probleem is, omdat $\sum k_{ij} \equiv$ constante, equivalent met het volgende:

Probleem R:

Maximaliseer

$$\sum c_{ij} x_{ij} \tag{5.5}$$

onder de voorwaarden

$$\left.\begin{aligned} -x_{ij} &\leq -a_{ij} && (5.6)\\ x_{ij} &\leq b_{ij} && (5.7)\\ x_{ij} + t_i - t_j &\leq 0 && (5.8) \end{aligned}\right\} \quad (i,j) \in \mathcal{P}\,,$$

$$t_n \leq s. \tag{5.9}$$

Hierin is s de toegestane duur van het project. Het gevraagde maximum hangt in het algemeen van s af. We geven het daarom aan met c(s).

Als s voldoende groot is, kan $x_{ij} = b_{ij}$ worden gekozen. Voor zodanige s geldt dus

$$c(s) = \sum_{\mathcal{P}} c_{ij}\, b_{ij} \tag{5.10}$$

en dat is natuurlijk de grootste waarde die c(s) kan bereiken. De kleinste s waarvoor de keuze $x_{ij} = b_{ij}$ mogelijk is, duiden we aan met m_2. Men kan m_2 bepalen op dezelfde wijze als het tijdstip e_n in §2; dus

$$\begin{cases} t_1 = 0, \\ t_j = \max\{t_i + b_{ij} \mid (i,j) \in \mathcal{P}\} \quad \text{voor } j = 2,\ldots,n, \\ m_2 \overset{\text{def}}{=} t_n. \end{cases} \tag{5.11}$$

Als s echter kleiner is dan m_1, die als volgt gedefinieerd is

$$\begin{cases} t_1 = 0, \\ t_j = \max\{t_i + a_{ij} \mid (i,j) \in \mathcal{P}\} \quad \text{voor } j = 2,\ldots,n, \\ m_1 \overset{\text{def}}{=} t_n, \end{cases} \tag{5.12}$$

dan is het niet mogelijk om aan (5.2) en (5.3) te voldoen, zodat een schema met $s < m_1$ niet bestaat.

Er rest ons dus nog het probleem c(s) te bepalen voor $m_1 \leq s < m_2$. Dit gebeurt door uit te gaan van het schema met t_j, gedefinieerd door (5.11) en $x_{ij} = b_{ij}$, en door achtereenvolgens voor steeds kleinere waarden van s de optimale t_i en x_{ij} te bepalen. Hiervoor zijn ongeveer tegelijkertijd twee nauw verwante algorithmen gevonden, n.l. door KELLEY *) en door FULKERSON **). Beide algorithmen lossen het betreffende lineaire programmeringsprobleem op door naast het (oorspronkelijke) probleem R ook het duale probleem te beschouwen. In het Engels duidt men een dergelijke procedure aan met de term "primal-dual algorithm".

Alvorens de algorithme van FULKERSON in § 6 te bespreken, zullen we bij wijze van intuïtieve inleiding eerst nagaan wat er zoal kan gebeuren, als we de projectduur verkorten.

Allereerst merken we op, dat het hierboven genoemde beginschema (5.11) één of meer kritieke paden bevat. Dit zal ook blijken te gelden voor de later optredende schema's. Stel nu dat we de projectduur met een korte tijd θ willen bekorten. Hiertoe moeten dan bepaalde activiteiten worden ingekort. Het is duidelijk dat onze keus zal vallen op kritieke activiteiten, en wel dié, waarvoor de daling van c(s) zo klein mogelijk is. We beperken ons even tot het eenvoudige geval dat er slechts één kritiek pad is. De voordeligste wijze om de projectduur met θ te bekorten is dan de activiteit met de kleinste c_{ij} waarvoor een inkorting inderdaad mogelijk is, met θ in te korten (in het schema (5.11) worden door de genoemde restrictie alleen activiteiten met $a_{ij} = b_{ij}$ uitgesloten).

De vraag is nu: hoe groot kan θ zijn? Er zijn verschillende omstandigheden die maken dat θ niet willekeurig groot kan worden gekozen. In de eerste plaats mag de duur van een activiteit die wordt ingekort, niet kleiner worden dan a_{ij}. Ten tweede heeft het geen zin

---

*) J.E. KELLEY Jr., Critical Path Planning and Scheduling: Mathematical Basis, Journal of the Operations Research Society of America, 9, (1961), p. 296-320.

**) D.R. FULKERSON, A Network Flow Computation for Project Cost Curves, Management Science, 7,(1961), p. 167-178.

om de beschouwde activiteit zonder meer verder te verkorten, zodra door de initiële verkorting ook een ander pad kritiek is geworden. Ten derde mag θ niet zo groot zijn dat de resulterende verlenging van andere activiteiten een lengte groter dan b_{ij} zou opleveren. Het is verrassend dat een verkorting van de projectduur een verlenging van een activiteitsduur met zich kan meebrengen, maar men dient te bedenken dat aan c(s) steeds de maximale waarde moet worden gegeven. We zullen dit aan een eenvoudig voorbeeld toelichten.

Voorbeeld 5.1

Het netwerk heeft de gedaante die in figuur 5.1 is afgebeeld, terwijl bij de pijlen de grootheden a_{ij} en b_{ij} zijn vermeld. Verder geldt

(5.13) $$c_{12} = c_{34} = 2;\ c_{23} = 1;\ c_{13} = c_{24} = 0.$$

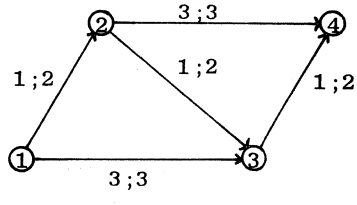


fig. 5.1

Netwerk waarin verlenging van een activiteit optreedt

In het beginschema ($x_{ij} = b_{ij}$) is het pad 1-2-3-4 kritiek. Het kan het voordeligst verkort worden door x_{23} te verkleinen, wat mogelijk is totdat $x_{23} = a_{23} = 1$. Nu zijn echter meteen alle paden kritiek geworden. Willen we de projectduur nog verder verkorten, dan moeten zowel x_{12} als x_{34} verkort worden (de enige activiteiten die nog verkort kunnen worden). Natuurlijk is het in principe mogelijk om x_{23} hierbij ongewijzigd te laten, maar dit zou geen optimale oplossing geven. Dus x_{23} moet verlengd worden en, als $x_{12} = x_{34} = 1$, is x_{23} weer gelijk aan 2.

De algorithme van FULKERSON wordt overigens niet rechtstreeks toegepast op het probleem R, maar op een probleem (S*) dat verwant is aan de duale R^* van R. Tot slot van deze paragraaf geven we nu de gedetailleerde herleiding van R tot S^*.

Dualisering van R geeft

Probleem R^*:

Minimaliseer

$$(5.14) \qquad -\sum_{\mathcal{P}} a_{ij}h_{ij} + \sum_{\mathcal{P}} b_{ij}g_{ij} + sv$$

onder de voorwaarden

$$(5.15) \qquad -h_{ij} + g_{ij} + f_{ij} = c_{ij}\ ,\ (i,j)\in\mathcal{P},$$

$$(5.16) \qquad -\sum_{i} f_{ij} + \sum_{k} f_{jk} = 0 \quad \text{voor } 1 < j < n,$$

$$(5.17) \qquad -\sum_{i} f_{in} + v = 0\ ,$$

$$(5.18) \qquad f_{ij},\ g_{ij},\ h_{ij} \text{ en } v \geq 0.$$

Als toelichting merken we in de eerste plaats op dat (5.15), (5.16) en (5.17) gelijkheden zijn, omdat van de variabelen in het oorspronkelijke probleem R niet expliciet wordt geëist dat ze niet-negatief zijn. Een tweede opmerking betreft de restricties (5.16). Deze zijn afkomstig van de variabelen t_j ($j < n$). Bekijken we de restricties (5.8) in probleem R voor een vaste j, dan zien we, dat t_j steeds optreedt als $(i,j)\in\mathcal{P}$. Aangezien echter voor $j < n$ deze restricties ook zijn te schrijven als $y_{jk} + t_j - t_k \leq 0$, komt t_j ook voor (nu met het plusteken) voor alle $(j,k)\in\mathcal{P}$. De restrictie (5.17) is op een dergelijke wijze verkregen door de variabele t_n te beschouwen.

Voor iedere keuze van de f_{ij} is wegens (5.15) het verschil $g_{ij} - h_{ij}$ bepaald. Daar verder in (5.14) de coëfficiënt van g_{ij} niet kleiner is dan de absolute waarde van de coëfficiënt van h_{ij}, doen we er goed aan g_{ij} en h_{ij} zo klein mogelijk te maken.

In verband met (5.18) vinden we dan

$$g_{ij} = (c_{ij} - f_{ij})^+, \tag{5.19}$$

$$h_{ij} = (f_{ij} - c_{ij})^+, \tag{5.20}$$

waarin x^+ een afkorting is van $\max(0,x)$. De kriteriumfunktie wordt nu

$$-\sum_{\mathcal{P}} a_{ij}(f_{ij} - c_{ij})^+ + \sum_{\mathcal{P}} b_{ij}(c_{ij} - f_{ij})^+ + sv. \tag{5.21}$$

Om deze niet-lineaire funktie weer lineair te maken, passen we de volgende kunstgreep toe. We definiëren variabelen f_{ij1} en f_{ij2} als volgt:

$$\left\{\begin{array}{ll} f_{ij1} = c_{ij} & \\ f_{ij2} = f_{ij} - c_{ij} & \text{als } f_{ij} > c_{ij}, \\ f_{ij1} = f_{ij} & \\ f_{ij2} = 0 & \text{als } f_{ij} \leq c_{ij}. \end{array}\right. \tag{5.22}$$

Blijkbaar geldt

$$\left\{\begin{array}{l} f_{ij1} + f_{ij2} = f_{ij}, \\ 0 \leq f_{ij1} \leq c_{ij}, \\ 0 \leq f_{ij2}. \end{array}\right. \tag{5.23}$$

In (5.21) zijn de termen in de tweede som gelijk aan 0 voor $f_{ij} > c_{ij}$, zodat we ons kunnen beperken tot $f_{ij} \leq c_{ij}$; maar dan is $f_{ij} = f_{ij1}$. Verder zijn de termen van de eerste som gelijk aan 0 voor $f_{ij} \leq c_{ij}$, terwijl voor $f_{ij} > c_{ij}$ geldt dat $(f_{ij} - c_{ij})^+ = f_{ij} - c_{ij} = f_{ij2}$. Dus (5.21) wordt

$$-\sum_{\mathcal{P}} a_{ij} f_{ij2} + \sum_{\mathcal{P}} b_{ij}(c_{ij} - f_{ij1}) + sv. \tag{5.24}$$

Afgezien van de constante $\sum_{\mathcal{P}} b_{ij}c_{ij}$ is dit te schrijven als

$$(5.25) \qquad -\sum_{r}\sum_{\mathcal{P}} a_{ijr}\, f_{ijr} + sv,$$

waarin $a_{ij1} = b_{ij}$, $a_{ij2} = a_{ij}$, $r = 1, 2$.

Ook de restricties laten zich op deze wijze herschrijven; aan (5.15) is <u>altijd</u> voldaan wegens de keuze (5.19) en (5.20), terwijl (5.16) en (5.17) overgaan in

$$(5.26) \qquad -\sum_{r}\sum_{i} f_{ijr} + \sum_{r}\sum_{k} f_{jkr} = 0,$$

$$(5.27) \qquad -\sum_{r}\sum_{i} f_{inr} + v = 0.$$

Tenslotte kunnen we de ongelijkheden van (5.23) schrijven als

$$(5.28) \qquad 0 \leq f_{ijr} \leq c_{ijr}$$

met $c_{ij1} = c_{ij}$, $c_{ij2} = \infty$.

Samenvattend kunnen we dus zeggen dat we het probleem R^* hebben herleid tot

<u>Probleem S^*</u>:

Minimaliseer

$$(5.25) \qquad -\sum_{r}\sum_{\mathcal{P}} a_{ijr}\, f_{ijr} + sv$$

onder de voorwaarden

$$(5.26) \qquad -\sum_{r}\sum_{i} f_{ijr} + \sum_{r}\sum_{k} f_{jkr} = 0 \quad \text{voor } 1 < j < n,$$

$$(5.27) \qquad -\sum_{r}\sum_{i} f_{inr} + v = 0,$$

$$(5.28) \qquad 0 \leq f_{ijr} \leq c_{ijr}.$$

Dit probleem heeft de volgende interpretatie. Als we in het netwerk iedere activiteit (i,j) vervangen door twee buizen (i,j,r), waarvan er één de capaciteit c_{ij1} en de ander capaciteit $c_{ij2} = \infty$ heeft, en f_{ijr} is een stroom materie door (i,j,r), dan zijn de restricties (5.26) juist behoudswetten voor de materie in ieder knooppunt

(behalve 1 en n), en (5.27) betekent dat de netto-stroom naar het eindknooppunt n gelijk is aan v. Het probleem is dan de stroom f_{ijr} zó te kiezen dat (5.25) minimaal is.

6. De algorithme van FULKERSON.

Aan het begin van de algorithme zijn v en de f_{ijr} alle 0, terwijl de t_j worden gegeven door (5.11). Op grond van de waarden van de verschillende constanten die in het probleem voorkomen, worden dan de knooppunten van merken ("labels") voorzien, waarna de grootheden f_{ijr} en t_j worden gewijzigd. Dan worden de knooppunten opnieuw gemerkt, enz., totdat de situatie $t_n = m_1$ is bereikt. Aan het begin van het merkproces zullen f_{ijr} en t_j steeds blijken te voldoen aan de volgende optimaliteitseisen:

$$(6.1)\qquad \begin{cases} \text{als } \bar{a}_{ijr} \overset{\text{def}}{=} a_{ijr} + t_i - t_j < 0, \text{ dan is } f_{ijr} = 0, \\ \text{als } \bar{a}_{ijr} > 0, \text{ dan is } f_{ijr} = c_{ijr}. \end{cases}$$

We laten nu zien dat deze eisen inderdaad voldoende zijn voor een optimale oplossing van probleem S^*; nauwkeuriger geformuleerd:

Hulpstelling 6.1

Als f_{ijr} en t_j voldoen aan (6.1) en aan de voorwaarden van probleem S^*, dan vormen de f_{ijr} een optimale oplossing van dit probleem voor $s = t_n$.

Bewijs:

Uit de definitie van $\bar{a}_{ijr}$ volgt

$$(6.2)\qquad \sum_r \sum_{\mathfrak{P}} \bar{a}_{ijr} f_{ijr} = \sum_r \sum_{\mathfrak{P}} a_{ijr} f_{ijr} + \sum_r \sum_{\mathfrak{P}} (t_i - t_j) f_{ijr}.$$

De tweede som van het rechterlid kan met behulp van (5.26) en (5.27) als volgt worden herleid:

$$\sum_r \sum_i \sum_j (t_i - t_j) f_{ijr} = \sum_r \sum_{i=2}^{n-1} t_i \left(\sum_j f_{ijr} \right) - \sum_r \sum_{j=2}^{n} t_j \left(\sum_i f_{ijr} \right) =$$

$$(6.3) \qquad = \sum_r \sum_{i=2}^{n-1} t_i \left(\sum_j f_{ijr} \right) - \sum_r \sum_{j=2}^{n-1} t_j \left(\sum_k f_{jkr} \right) - \sum_r t_n \left(\sum_i f_{inr} \right) =$$

$$= 0 - t_n v = 0 - sv = - sv \ .$$

Uit (6.2) en (6.3) zien we dat $\sum_r \sum_{\mathcal{P}} \bar{a}_{ijr} f_{ijr}$ gelijk is aan $\sum_r \sum_{\mathcal{P}} a_{ijr} f_{ijr} - sv$, zodat beide vormen tegelijk maximaal zijn. Kiezen we nu de f_{ijr} in overeenstemming met (6.1), dan is ten duidelijkste de eerste vorm maximaal en dus ook de tweede. Hiermee is hulpstelling 6.1 bewezen.

We geven nu eerst de algorithme en bewijzen daarna dat de uitkomsten de vereiste eigenschappen bezitten.

Begin :

Geef t_j de waarden gedefinieerd door (5.11) en neem alle f_{ijr} gelijk aan 0.

A1 (eerste fase van het merkproces):

Knooppunt 1 wordt gemerkt met *) (-, -, -, ∞). Verdere knooppunten j worden nu, zo mogelijk, als volgt gemerkt: als i gemerkt is, j ongemerkt, $(i,j) \in \mathcal{P}$ en $\bar{a}_{ij2} = 0$, geef dan alle knooppunten j met deze eigenschappen het merkje (i, 2, +, ∞); ga hiermee door totdat ðfwel n gemerkt is, ðfwel n niet gemerkt is terwijl geen verdere merkjes gegeven kunnen worden; in het eerste geval is de algorithme beëindigd ($s = m_1$); in het tweede geval gaan we verder bij A2.

A2 (tweede fase van het merkproces):

Als i gemerkt is en j niet, dan moet j worden gemerkt wanneer één van de volgende gevallen zich voordoet:

---

*) De streepjes geven aan dat de eerste drie "coördinaten" van het merk blanco gelaten worden.

a) $(i,j) \in \mathcal{P}$, $\bar{a}_{ijr} = 0$ en $f_{ijr} < c_{ijr}$ voor zekere r;
geef j het merk *) $(i, r, +, \min(\varepsilon_i, c_{ijr} - f_{ijr}))$.

b) $(j,i) \in \mathcal{P}$, $\bar{a}_{jir} = 0$ en $f_{jir} > 0$ voor zekere r;
geef j het merk $(i, r, -, \min(\varepsilon_i, f_{jir}))$.

Ga hiermee door totdat òfwel n gemerkt is, òfwel n niet gemerkt is terwijl geen verdere merkjes uitgedeeld kunnen worden. In het eerste geval gaan we verder bij B, in het tweede geval bij C.

B (wijziging van de f_{ijr}):

Bij het eindpunt n beginnend, bepalen we als volgt een route naar de oorsprong: als het beschouwde knooppunt p het merk $(i, r, +, \varepsilon_p)$ heeft, dan wordt f_{ipr} met ε_n vermeerderd en het volgende te beschouwen knooppunt is i; als het beschouwde knooppunt p het merk $(i, r, -, \varepsilon_p)$ heeft, dan wordt f_{pir} met ε_n verminderd; verwijder alle merkjes en ga terug naar A1.

C (wijziging van de t_j):

Definieer verzamelingen $\mathcal{D}_1$ en $\mathcal{D}_2$ en getallen δ_1, δ_2, δ als volgt:

$$
(6.4)\quad \begin{cases} \mathcal{D}_1 = \{(i,j,r) \mid i \text{ gemerkt, } j \text{ ongemerkt, } (i,j) \in \mathcal{P},\ \bar{a}_{ijr} < 0\}, \\ \mathcal{D}_2 = \{(i,j,r) \mid i \text{ ongemerkt, } j \text{ gemerkt, } (i,j) \in \mathcal{P},\ \bar{a}_{ijr} > 0\}, \\ \delta_1 = \min(-\bar{a}_{ijr} \mid \mathcal{D}_1), \\ \delta_2 = \begin{cases} \min(\bar{a}_{ijr} \mid \mathcal{D}_2) & \text{als } \mathcal{D}_2 \text{ niet-leeg is,} \\ \infty & \text{als } \mathcal{D}_2 \text{ leeg is,} \end{cases} \\ \delta = \min(\delta_1, \delta_2). \end{cases}
$$

Verminder nu alle t_i die bij ongemerkte knooppunten behoren met δ; verwijder alle merkjes, bepaal de nieuwe waarden van $\bar{a}_{ijr}$ en ga terug naar A1.

Tot zover de algorithme. We merken nu allereerst op dat de onder "Begin" bepaalde f_{ijr} en t_j aan (6.1) voldoen (want $\bar{a}_{ijr_*} \leq 0$ voor alle activiteiten) en natuurlijk ook aan de voorwaarden van S_*, en wel met $v = 0$.

---

*) De vierde coördinaat van het merk van knooppunt i wordt aangeduid met ε_i.

Als het merken eindigt via B, zodat de stroom f_{ijr} wordt gewijzigd, dan blijft (6.1) gelden, want alleen in activiteiten met $\bar{a}_{ijr} = 0$ is iets veranderd. Ook de voorwaarden van S^* blijven gelden met $v + \varepsilon_n$ in plaats van v in (5.22)

Als het merke eindigt via C, is ook in te zien dat zowel (6.1) als de voorwaarden van S^* blijven gelden. We bewijzen hiertoe de volgende hulpstelling:

Hulpstelling 6.2

Voor alle δ' met $0 \leq \delta' \leq \delta$ geldt, dat f_{ijr} benevens t'_j, gedefinieerd door

$$t'_j = \begin{cases} t_j & \text{als j gemerkt is,} \\ t_j - \delta' & \text{als j ongemerkt is,} \end{cases} \tag{6.5}$$

voldoen aan (6.1).

Bewijs:

We bewijzen dat iedere activiteit die voldoet aan

$$\bar{a}_{ijr} \overset{\text{def}}{=} a_{ijr} + t'_i - t'_j < 0, \tag{6.6}$$

ook voldoet aan $f_{ijr} = 0$; we onderscheiden drie gevallen, te weten

a) als $\bar{a}_{ijr} < 0$, dan was f_{ijr} al 0 en hij wordt niet gewijzigd; dus f_{ijr} blijft 0.

b) als $\bar{a}_{ijr} = 0$ (dus $a_{ijr} + t_i - t_j = 0$), dan is, in verband met (6.6), $t'_i < t_i$ en $t'_j = t_j$; dus i is ongemerkt en j is gemerkt. Dan moet echter gelden $f_{ijr} = 0$, anders zou i gemerkt zijn volgens A2, punt b.

c) als $\bar{a}_{ijr} > 0$, dan geldt weer dat i ongemerkt is en j gemerkt. Maar dan is (i,j,r) per definitie een element van de verzameling $\mathfrak{D}_2$. Hieruit zien we dat $\delta_2 \leq \bar{a}_{ijr}$ en dus, wegens $\delta' \leq \delta \leq \delta_2$, ook $\delta' \leq \bar{a}_{ijr}$. Dit geeft echter een tegenspraak; immers $\bar{a}'_{ijr} = a_{ijr} + t'_i - t'_j = a_{ijr} + t_i - \delta' - t_j = \bar{a}_{ijr} - \delta' \geq 0$ terwijl we juist hadden verondersteld (zie (6.6)) dat $\bar{a}'_{ijr} < 0$. Dit geval kan dus niet optreden.

Op precies dezelfde manier (ook met een splitsing in drie gevallen) bewijst men dat iedere activiteit (i,j,r) waarvoor $\bar{a}'_{ijr} > 0$ is, voldoet aan $f_{ijr} = c_{ijr}$.

We zien dus dat de grootheden f_{ijr} en t_j <u>steeds</u> aan de optimaliteitseisen voldoen.

<u>Hulpstelling 6.3</u>

De verzameling $\mathcal{D}_i$ is nooit leeg.

<u>Bewijs</u>:

Daar het merkproces via C is geëindigd, bestaat er een activiteit (i,j) waarvoor geldt: i is gemerkt, j is ongemerkt. Zou nu gelden $\bar{a}_{ij2} < 0$, dan is reeds $(i,j,2) \in \mathcal{D}_1$. In het geval $\bar{a}_{ij2} = 0$ zou j in de tweede fase gemerkt zijn, omdat altijd voldaan is aan $f_{ij2} < c_{ij2} = \infty$. In het resterende geval, $\bar{a}_{ij2} > 0$, geldt volgens (6.1) $f_{ij2} = c_{ij2} = \infty$. Dit zou alleen kunnen als n bij de vorige keer merken o.a. was voorzien van $\varepsilon_n = \infty$. Maar dan zou de algorithme volgens A1 beëindigd moeten zijn.

Uit deze hulpstelling volgt direct dat δ altijd eindig is.

Men kan aantonen dat de algorithme steeds na een eindig aantal stappen afloopt. FULKERSON geeft hiervan een eenvoudig bewijs, dat echter alleen opgaat in het door hem beschouwde geval, waarin alle a_{ijr} geheel zijn.

<u>Hulpstelling 6.4</u>

De optimale waarden van x_{ij} (dus de x_{ij} die probleem R oplossen) worden gegeven door

$$x_{ij} = \min(b_{ij}, t_j - t_i). \tag{6.7}$$

<u>Bewijs</u>:

Een vergelijking van de problemen R en R^* leert dat het voldoende is om de volgende drie beweringen te bewijzen:

a) als $x_{ij} + t_i - t_j < 0$, dan is $f_{ij} = 0$,

b) als $x_{ij} < b_{ij}$, dan is $g_{ij} = 0$,
c) als $x_{ij} > a_{ij}$, dan is $h_{ij} = 0$.

De bewijzen zijn eenvoudig. We geven ze hieronder achtereenvolgens.

a) Uit het gestelde en (6.7) volgt $b_{ij} + t_i - t_j < 0$, zodat ook $a_{ij} + t_i - t_j < 0$; volgens (6.1) geldt dan $f_{ijr} = 0$ (r = 1, 2), en dus $f_{ij} = 0$.

b) Uit het gestelde en (6.7) volgt $t_j - t_i < b_{ij} = a_{ij1}$, zodat $a_{ij1} + t_i - t_j > 0$; volgens (6.1) is dan $f_{ij1} = c_{ij1} = c_{ij}$, en hieruit volgt met (5.22) dat $f_{ij} \geq c_{ij}$. Uit (5.19) volgt nu $g_{ij} = 0$.

c) Uit het gestelde en (6.7) volgt $t_j - t_i > a_{ij}$, zodat $a_{ij2} + t_i - t_j < 0$; volgens (6.1) is dan $f_{ij2} = 0$, en hieruit volgt met (5.22) dat $f_{ij} \leq c_{ij}$. Uit (5.20) volgt nu $h_{ij} = 0$.

Uit deze hulpstelling zien we dat de getallen t_j, die door de algorithme steeds na het doorlopen van stap C worden geproduceerd, op eenvoudige wijze de grootheden x_{ij} bepalen en daarmee ook de waarde van de kriteriumfunktie $\sum_{\beta} c_{ij}x_{ij}$ in het punt t_n. Tussen de punten die corresponderen met de achtereenvolgende waarden van t_n, verloopt de kriteriumfunktie lineair. Met (6.5) en (6.7) laat dit zich gemakkelijk bewijzen.

Voorbeeld 6.1

We zullen de algorithme nu toelichten aan een voorbeeld en wel hetzelfde "project" als in voorbeeld 5.1. Het beginschema $x_{ij} = b_{ij}$ geeft aan t_j de volgende waarden:

(6.8) $$t_1 = 0,\ t_2 = 2,\ t_3 = 4,\ t_4 = 6.$$

Het verdient aanbeveling op dit punt de $\bar{a}_{ijr}$ te bepalen. Deze getallen zijn vermeld in tabel 6.1.

Tabel 6.1

i,j	1,2	1,3	2,3	2,4	3,4
$\bar{a}_{ij1}$	0	-1	0	-1	0
$\bar{a}_{ij2}$	-1	-1	-1	-1	-1

Alle f's zijn aan het begin gelijk aan 0.

Bij A1 krijgt alleen 1 een merk, n.l.(-, -, -, ∞). Onder A2 krijgen de knooppunten 2, 3, 4 respectievelijk de merken (1, 1, +, 2), (2, 1, +, 1) en (3, 1, +, 1), omdat langs het pad 1-2-3-4 overal geldt $\bar{a}_{ij1} = 0$ en $f_{ij1} < c_{ij1}$. Daar 4 een merk heeft, gaan we naar punt B, waar achtereenvolgens f_{341}, f_{231} en f_{121} gelijk aan 1 worden. Teruggaand naar A1 zien we dat 1 weer het merk (-, -, -, ∞) krijgt, en onder A2 kan alleen 2 worden gemerkt, n.l.met (1, 1, +, 1). We gaan nu naar punt C en vinden $\mathfrak{D}_1 = \{(1,3,1), (1,3,2), (2,3,2), (2,4,1), (2,4,2)\}$; $\mathfrak{D}_2$ is de lege verzameling en dus $\delta = \delta_1 = 1$. De nieuwe waarden van t_j zijn

(6.9) $$t_1 = 0,\ t_2 = 2,\ t_3 = 3,\ t_4 = 5,$$

terwijl de waarden van x_{ij} en $\bar{a}_{ijr}$ zijn vermeld in onderstaande tabel.

Tabel 6.2

i,j	1,2	1,3	2,3	2,4	3,4
y_{ij}	2	3	1	3	2
$\bar{a}_{ij1}$	0	0	1	0	0
$\bar{a}_{ij2}$	-1	0	0	0	-1

We gaan nu weer terug naar A1. Knooppunt 1 krijgt het merk (-, -, -, ∞) nu kan ook 3 gemerkt worden, n.l. met (1, 2, +, ∞). Verdere merkjes kunnen bij A1 niet gegeven worden. Bij A2 krijgen 2 en 4 respectievelijk de merken (1, 1, +, 1) en (2, 2, +, 1).

Volgens punt B wordt nu f_{242} verhoogd tot 1, en f_{121} tot 2. Weer keren we terug naar A1, en krijgen de knooppunten 1 en 3 de merken (-, -, -, ∞) en (1, 2, +, ∞). Bij A2 echter kan nu alleen 4 worden gemerkt en wel met (3, 1, +, 1). Als gevolg hiervan wordt dan onder punt B f_{341} verhoogd tot 2, en f_{132} tot 1. Ten derde male krijgen nu 1 en 3 resp. hun merken (-, -, -, ∞) en (1, 2, +,∞); nu kan bij A2 geen enkel merk worden uitgedeeld. Bij punt C vinden we $\delta = \delta_1 = \delta_2 = 1$ en

(6.10) $$t_1 = 0,\ t_2 = 1,\ t_3 = 3,\ t_4 = 4.$$

Tabel 6.3 geeft de nieuwe waarden van x_{ij} en $\bar{a}_{ijr}$.

Tabel 6.3

i,j	1,2	1,3	2,3	2,4	3,4
y_{ij}	1	3	2	3	1
$\bar{a}_{ij1}$	1	0	0	0	1
$\bar{a}_{ij2}$	0	0	-1	0	0

Bij A1 kunnen tenslotte 1, 2, 3 en 4 worden gemerkt met respectievelijk (-, -, -, ∞), (1, 2, +, ∞), (1, 2, +, ∞) en (2, 2, +, ∞), zodat 4 de kortst mogelijke projectduur is.

De funktie c(s) is nu geheel bekend: voor $s \geq 6$ geldt $c(s) = \sum c_{ij} b_{ij} = 10$; blijkens tabel 6.2 is c(5) = 9 en uit tabel 6.3 zien we dat c(4) = 6. De grafiek van c(s) is afgebeeld in figuur 6.1.

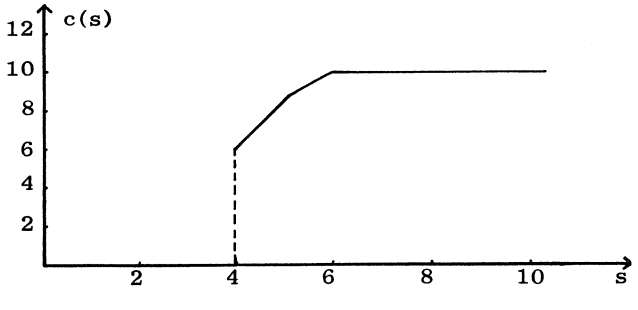


fig. 6.1

De funktie c(s) voor het project in voorbeeld 6.1

Eén van de bezwaren die men tegen CPM heeft ingebracht, is het weinig realistische karakter van de aanname (5.1). Aan dit bezwaar wordt tegemoet gekomen door de opmerking, dat de methode ook kan worden toegepast wanneer de $k_{ij}(x_{ij})$ convexe, stuksgewijs lineaire, niet-stijgende funkties zijn voor $a_{ij} \leq x_{ij} \leq b_{ij}$. We zullen hier niet verder op ingaan.

De "Critical Path Method" is voor het eerst toegepast bij de bouw van een chemische fabriek. Beknopte gegevens hierover en over niet-wiskundige aspecten van de methode kan men vinden in een brochure van J.E. KELLEY Jr. en M.R. WALKER *).

7. Beperkte hulpbronnen.

Veronderstel dat men bij het plannen volgens PERT een schema heeft gevonden dat een redelijke kans heeft om binnen een bepaalde tijd gereed te komen, of bij het plannen volgens CPM een schema dat de totale kosten minimaliseert. Het kan nu gebeuren dat het buiten beschouwing laten van de beperktheid der benodigde hulpbronnen een schema heeft opgeleverd dat in de praktijk onuitvoerbaar is. In §3 is hiervan al een eenvoudig voorbeeldje genoemd.

Het uitbreiden der hulpbronnen (meer arbeidskrachten aanstellen, meer machines aanschaffen, e.d.) om ervoor te zorgen dat zulke schema's toch uitvoerbaar zijn, is alleen dán efficiënt als deze hulpbronnen ook voor een belangrijk deel van de tijd gebruikt kunnen worden.

Men kan trachten de ideale situatie waarin van alle hulpbronnen steeds de maximale capaciteit wordt benut, te benaderen door bijv. activiteiten met een grote speelruimte enige tijd uit te stellen. Meestal zal het echter noodzakelijk zijn om de projectduur te verlengen. Het is natuurlijk lastig om een geschikt kriterium op te stellen dat het verlagen van de pieken afweegt tegen het uitstellen

---

*) J.E. KELLEY Jr. and M.R. WALKER, Critical Path Planning and Scheduling, (1959), Proceedings of the Eastern Joint Computer Conference.

van de voltooiing. Maar ook in het geval dat gevraagd wordt een schema op te stellen van minimale duur onder de bijvoorwaarde, dat het gebruik per tijdseenheid van iedere hulpbron een gegeven grens niet mag overschrijden, blijft een bijzonder moeilijk probleem bestaan.

Er is een algorithme *) bekend die volgens een probeermethode een benaderende oplossing geeft van dit probleem. Deze oplossing voldoet weliswaar aan de bijvoorwaarden, maar garandeert geen minimale projectduur.

Een verdere complicatie die kan optreden, is het feit dat hulpbronnen elkaar soms gedeeltelijk kunnen vervangen. Het probleem wordt nog complexer wanneer aan meer dan één project tegelijk wordt gewerkt. Ook hier doet zich de moeilijkheid voor dat eerst een geschikt kriterium moet worden gevonden, bijv. om de bekortingen van verschillende projecten tegen elkaar af te wegen.

---

*) W.A. GRAY en E.M. KIDD, Critical Path Scheduling with Resource Leveling on the I.B.M. 7090; Union Carbide Nuclear Company; A.E.C. Research and Development Report, K-1499, (1962).

Literatuur

J. BRANDENBERGER und R. KONRAD, Netzplantechnik, Industrielle Organisation, Zürich, (1967).

H.G. DERKSEN, Inleiding tot de planning met de netwerkmethode, Samsom, Alphen aan de Rijn, (1966).

A. KAUFMANN et G. DESBAZEILLE, La méthode du chemin critique, Dunod, Paris, (1966; 2e. edition).

J.J. MODER and C.R. PHILLIPS, Project Management with CPM and PERT, Reinhold Publishing Corp., New York, (1964).

Verder kan men de volgende bibliografie raadplegen:

S. LERDA-OLBERG, Bibliography on Network-based Project Planning and Control Techniques: 1962 - 1965, J.O.R.S.A., 14, (1966), p. 925 - 931.

DEEL 8

HOOFDSTUK 3

SIMULATIE DOOR W. MOLENAAR

1. Simulatie van deterministische processen.

Ter inleiding behandelen wij een tweetal gefingeerde voorbeelden.

Voorbeeld 1.1

Een fotograaf heeft aan vier opdrachtgevers beloofd dat hun foto's de komende woensdag klaar zouden zijn. Als hij die woensdagavond naar huis gaat is hij echter met geen van de vier opdrachten verder dan dat hij de negatieven op zijn werktafel heeft liggen. Alle cliënten hebben evenveel haast. Omdat ze alle vier nogal ver weg wonen besluit hij op donderdag eerst alle opdrachten te voltooien en ze dan samen weg te brengen.

De vraag is nu in welke volgorde onze fotograaf en zijn twee assistenten de vier opdrachten A, B, C en D moeten behandelen. Hierbij is gegeven, dat de eerste assistent alleen kan ontwikkelen en de tweede alleen afwerken. Afdrukken gebeurt alleen door de fotograaf zelf. Het afdrukken van een opdracht kan pas beginnen als het ontwikkelen voor die opdracht voltooid is en evenzo begint de afwerking van een opdracht pas na het afdrukken. Het materiaal moet in dezelfde volgorde elk der drie behandelingen ondergaan. Dus als B eerder ontwikkeld is dan A wordt B ook eerder afgedrukt enz. De benodigde tijden voor de diverse onderdelen zijn de fotograaf uit ervaring bekend en worden gegeven in tabel 1.1.

Tabel 1.1

Benodigde tijden in minuten per behandeling en per opdracht

opdracht	ontwikkelen	afdrukken	afwerken
A	120	100	40
B	50	90	60
C	20	60	120
D	90 *)	30	70

*) Wanneer opdracht D als eerste behandeld wordt, wordt dit echter 70; dit materiaal is n.l. al in de ontwikkeltrommel gespannen.

Bij elke volgorde van behandeling behoort nu een totale tijd vanaf het begin van het ontwikkelen van de eerste opdracht tot het voltooien van de afwerking van de laatste. Zo behoort bij de volgorde BACD een tijdsduur van 520 minuten, volgens het schema van tabel 1.2 *).

Tabel 1.2

Uitwerking voor de volgorde BACD

tijd in min.	bewerkingen
0 - 50	B ontw. , A wacht , C wacht , D wacht
50 - 140	B afdr. , A ontw , C wacht , D wacht
140 - 170	B afw. , A ontw. , C wacht , D wacht
170 - 190	B afw. , A afdr. , C ontw. , D wacht
190 - 200	B afw. , A afdr. , C wacht , D ontw.
200 - 270	B klaar , A afdr. , C wacht , D ontw.
270 - 280	B klaar , A afw. , C afdr. , D ontw.
280 - 310	B klaar , A afw. , C afdr. , D wacht
310 - 330	B klaar , A klaar , C afdr. , D wacht
330 - 360	B klaar , A klaar , C afw. , D afdr
360 - 450	B klaar , A klaar , C afw. , D wacht
450 - 520	B klaar , A klaar , C klaar , D afw.

Onze fotograaf vraagt zich nu op woensdagavond af, bij welke van de 4! rangschikkingen van de vier opdrachten de totale tijd minimaal is **) Het zou natuurlijk onverstandig zijn de volgorde BACD te kiezen, wanneer BCAD tot een kortere werktijd leidt. Hij gaat nu in gedachten het verloop bij elk van de 24 rangschikkingen na, d.w.z. hij maakt telkens een tabel zoals tabel 1.2 of schetst een grafische voorstelling van het verloop in de tijd. Het resultaat van zijn gedachtenexperiment vindt men in tabel 1.3.

---

*) Vraag: Is het probleem na de keuze van een vaste volgorde een netwerkprobleem geworden? Zo ja, teken dan het netwerk.

**) Dit is een "sequencing model for three stations and four jobs", zoals behandeld op blz. 456 van CHURCHMAN, ACKOFF and ARNOFF, Introduction to Operations Research, New York,(1957). Helaas is niet aan de aldaar genoemde condities voldaan, zodat de procedure om door eliminatie de snelste volgorde te vinden niet toepasbaar is.

Tabel 1.3

Benodigde totale tijd voor de volledige behandeling van de vier opdrachten

volgorde	tijd	volgorde	tijd	volgorde	tijd	volgorde	tijd
ABCD	560	BACD	520	CABD	460	DABC	560
ACBD	530	BCAD	430	CBAD	400	DBAC	520
ACDB	530	BCDA	430	CBDA	420	DBCA	430
ADCB	510	BDCA	430	CDBA	420	DCBA	400
ADBC	530	BDAC	540	CDAB	480	DCAB	460
ABDC	560	BADC	500	CADB	430	DACB	530

Wij zien, dat de fotograaf óf CBAD óf DCBA als volgorde moet kiezen, opdat de vier klanten zo snel mogelijk hun foto's ontvangen.

Voorbeeld 1.2

De firma Super en Co. gaat een bedrijf vestigen op het nieuwe industrieterrein van Dokkenburg. Men krijgt daar de beschikking over een eigen insteekhaven aan diep water. In eerste instantie is besloten tot de aanleg van 200 meter vrije kadelengte. De bouw van de fabriek en de installatie van voorzieningen op het terrein zijn nu zo ver voortgeschreden, dat binnenkort definitief beslist moet worden of de geprojecteerde lengte van 200 meter voldoende is. Nadien zal een uitbreiding van de kadelengte slechts met veel extra kosten mogelijk zijn.

De directie heeft nu aan de Afdeling Planning een schatting gevraagd van het te verwachten vervoer per schip. Deze prognose vindt men in tabel 1.4.

Tabel 1.4

Vervoersprognose

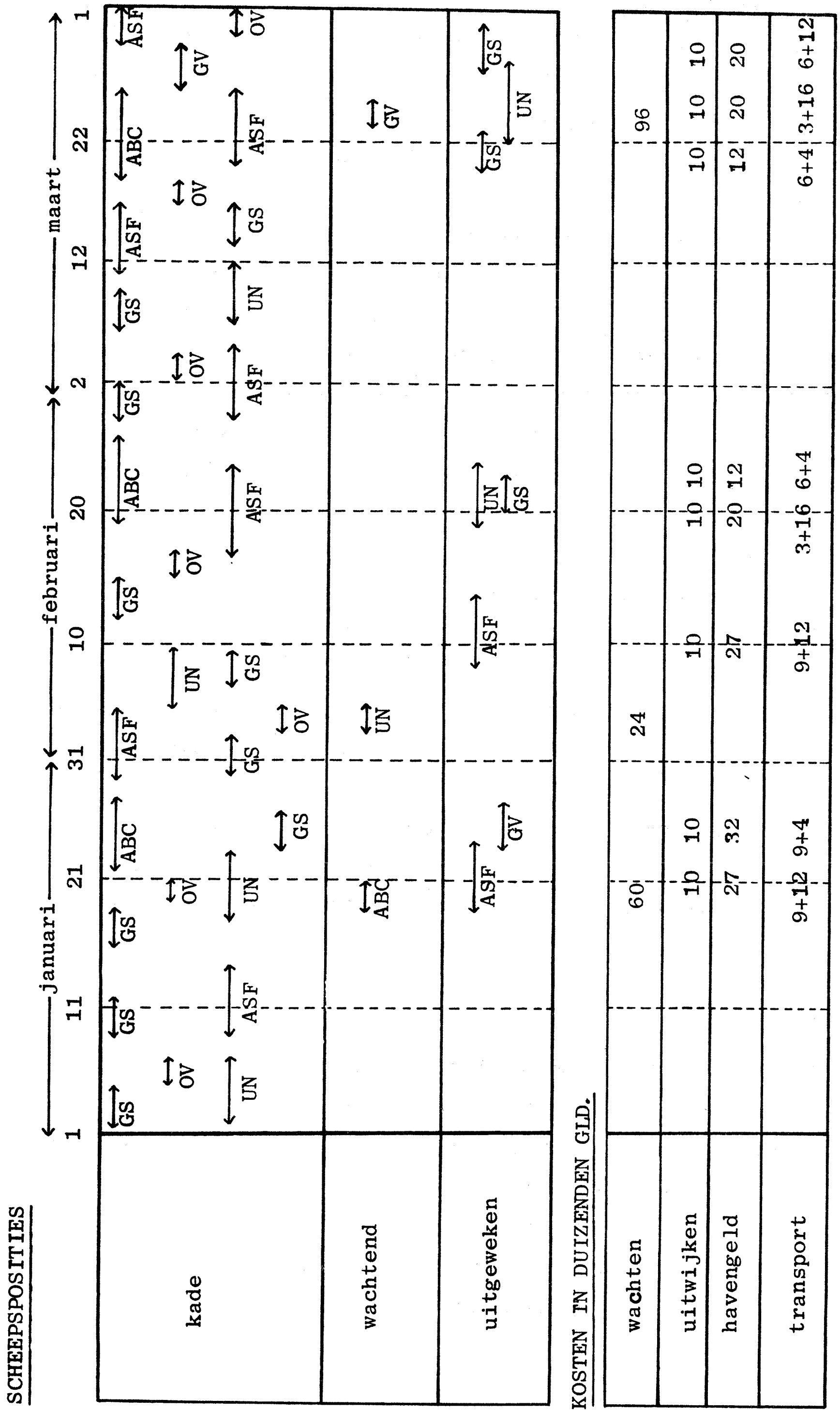
maatschappij	eerste aankomst in het jaar	frequentie	lostijd in dagen	laadtijd in dagen	vereiste lengte in meters
Gele Ster	2 jan.	1× per week	2	1	80
A.B.C.	19 jan.	1× per maand	4	3	100
Olievaart	5 jan.	1× per 14 dg.	2	0	100
Grootvervoer	24 jan.	1× per 2 mnd.	3	1	160
A.S.F.	9 jan.	1× per 10 dg.	3	3	90
United	2 jan.	1× per 16 dg.	1	4	80

Wanneer een schip bij aankomst niet voldoende vrije kadelengte vindt, moet het wachten of uitwijken naar een andere haven. Wanneer bekend is hoeveel extra kosten dit veroorzaakt, kan men berekenen welk deel van deze kosten vermeden had kunnen worden door de geprojecteerde kadelengte met L meter uit te breiden. Als die besparing voor een zekere waarde van L groot genoeg is in vergelijking tot de kosten van die uitbreiding, zal de directie van Super en Co. deze uitbreiding in nadere overweging nemen.

Bij de bepaling van de extra kosten nemen wij voorlopig aan, dat alle schepen precies volgens de dienstregeling varen, dat in de weekends wordt doorgewerkt en dat de laad- en lostijden niet aan toevalsfluctuaties onderhevig zijn. Elk schip dat een vrije plaats vindt, begint onmiddellijk met lossen en laden. De schepen van ASF kunnen niet wachten met lossen, omdat de lading dan zou bederven; de overige schepen kunnen maximaal 2 dagen wachten, behalve Gele Ster en Olievaart, waarvoor de maximale wachttijd 1 dag is.

Als bij aankomst van een schip niet binnen de maximale wachttijd een vrije plaats wordt verwacht, wijkt de kapitein uit naar een naburige gemeentelijke haven. Dit betekent vaste uitwijkkosten van *f* 10.000,- plus per ligdag een extra havengeld in guldens van 50 ×

Figuur 1.1 Scheepsposities en kosten



de vereiste lengte in meters, plus extra transportkosten van *f* 3.000,- per losdag en *f* 4.000,- per laaddag. Een wachtdag van een schip kost in guldens 300× de vereiste lengte in meters. Als het lossen van een Grootvervoer-schip x dagen vertraagd is, worden de eerstvolgende twee laadtijden van Gele Ster en de eerstvolgende laadtijd van United met evenveel dagen verlengd.

Men kiest nu een waarde voor de lengte L, waarmee de kadelengte van 200 meter zal worden uitgebreid. Vervolgens gaat men op papier na wat er onder deze omstandigheden zal gebeuren; d.w.z. men noteert voor elke dag welke schepen aan de kade liggen, welke er eventueel wachten of uitwijken en hoeveel extra kosten dit veroorzaakt.

Een grafische voorstelling van de ontwikkeling in januari tot en met maart bij een kadelengte van 200 meter (dus L = 0) vindt men in figuur 1.1.

Door deze grafiek voor bijvoorbeeld een heel jaar te maken en te vergelijken met een soortgelijke grafiek bij L = 50 kan men de besparingen schatten bij 50 meter extra kadelengte. Wij zullen het voorbeeld niet verder uitwerken.

Ook als de regelingen betreffende wachten en uitwijken ingewikkelder zijn, kan men de besparingen volgens de beschreven methode schatten. Met wat meer moeite kan men het hele proces behandelen voor het geval er in de weekends niet gewerkt wordt of tegen hogere beloning gewerkt wordt. Ook in de andere spelregels kan men zoveel verfijning aanbrengen, dat een redelijke beschrijving van de feitelijke toestand gewaarborgd is.

Wat is nu simulatie?

In de behandeling van beide voorbeelden zijn een aantal fasen te onderscheiden. Allereerst is het systeem afgebakend door een aantal veronderstellingen en eisen. Meestal blijft daarbij nog enige keuzevrijheid over: in voorbeeld 1.1 de volgorde van de foto-opdrachten en in voorbeeld 1.2 de lengte L waarmee de kade wordt uitgebreid. Daarna zijn spelregels opgesteld voor de onderlinge beïnvloeding van de onderdelen van het systeem. Deze regels komen misschien niet helemaal over-

een met de echte wisselwerkingen, maar ze beogen daarvan een betrouwbaar beeld te geven en toch zo eenvoudig mogelijk te zijn.

Vervolgens wordt in gedachten, of met potlood en papier, nagegaan hoe het systeem zich in de loop van de tijd zal ontwikkelen bij een vaste keuze van de beslissing (opdrachtenvolgorde, kadelengte). Men berekent daaruit de kosten, of de tijd, of welke andere doelfunktie dan ook. Vervolgens herhaalt men dit proces voor een andere gekozen beslissing en vergelijkt de resultaten.

Wij hebben dus in onze studeerkamer het complexe proces van gebeurtenissen nagespeeld, zoals een straaljagerpiloot in een linktrainer een vlucht nabootst of een waterloopkundige in een stroomgoot de invloed van wind en waterdiepte op de golfhoogte en de verplaatsing van zand. Dit naspelen zullen wij simuleren noemen. Deze techniek wordt in allerlei situaties gebruikt, waar het onmogelijk of te kostbaar is om met het "echte" systeem te experimenteren. Het behoeft geen betoog, dat de fotograaf er niet bij gebaat is diverse volgorden van de opdrachten uit te proberen en dat het onzinnig zou zijn eerst een jaar met de kadelengte van 200 meter aan te zien, hoeveel extra kosten deze toestand in feite zou veroorzaken. Geen enkele vliegtuigfabriek kan zich permitteren een prototype te bouwen zonder eerst een verkleind model in de windtunnel aan diverse windsnelheden, temperaturen enz. te onderwerpen. In zekere zin zou men dat ook een vorm van simulatie kunnen noemen.

Voor- en nadelen van simulatie.

Zoals wij gezien hebben moet het op papier naspelen herhaald worden voor iedere keuze van de beslissing (de volgorde resp. de kadelengte). Zo'n spel kost meestal niet veel tijd en geld, maar als het aantal beslissingen groot is wordt het toch nogal kostbaar het telkens te herhalen. In de meeste gevallen wordt de aard van het proces door de veranderde beslissing niet ingrijpend veranderd. Men kan dan het naspelen door een rekenautomaat laten doen, die zo geprogrammeerd is, dat met alle spelregels, afhankelijkheden en beperkingen rekening wordt gehouden.

Het gebruik van simulatie kan zonder meer worden afgeraden als er een wiskundig model bestaat, dat de situatie nauwkeurig beschrijft. Wie wil weten wat er gebeurt bij het samenpersen van een gas in een cylinder, is meer gebaat bij de formule PV = RT dan bij de resultaten van een aantal metingen. Evenzo zou het onzin zijn de aankomst van schepen bij diverse kadelengten te bestuderen, als men kan bewijzen, dat de besparing op de kosten in goede benadering gegeven wordt door $^{10}\log (L + 1)$. In het algemeen zal een mathematisch model er toe leiden dat de doelfunktie (tijd, kosten) expliciet op te schrijven is als funktie van de gegevens en van de gekozen beslissing. Naspelen onder steeds nieuwe omstandigheden is dan overbodig, omdat men de konsekwenties van die veranderde omstandigheden snel kan overzien.

Helaas is de gegeven situatie meestal veel gecompliceerder dan ieder eenvoudig wiskundig model. Wij moeten dan voor het toepassen van zo'n model de werkelijkheid geweld aandoen en het is niet bekend of onze conclusies volgens dat geforceerde model ook gelden voor de hiervan nogal afwijkende situatie, die in feite aanwezig is. Bij simulatie moeten ook wel zekere beperkende en idealiserende regels worden opgesteld, maar deze kunnen veel verfijnder en ingewikkelder zijn zonder dat de hanteerbaarheid in gevaar komt. Simulatie "kan situaties aan", waar een analytische oplossing niet door de wiskunde gepresteerd wordt.

Het is soms denkbaar dat de simulatie en de redenering vanuit een vereenvoudigd model elkaar aanvullen, net zoals een natuurkundige door een experiment een theoretische afleiding van een eigenschap zal controleren.

2. Monte Carlo methode : simulatie van stochastische processen.

Binnen het vereenvoudigde model van systeem en spelregels kenden wij bij de simulatie in de vorige paragraaf geen onzekerheid meer. Aankomstdata en ligtijden van de schepen waren exact bekend. Over de benodigde tijden voor ontwikkelen, afdrukken en afwerken bestond ook geen enkele twijfel. Het is duidelijk dat deze omstandigheden zich in de praktijk niet vaak zullen voordoen. Meestal zullen enige tijdsduren, hoeveelheden enz., die in een simulatiemodel optreden, aan toevalsfluctuaties onderhevig zijn. De ene keer vallen zij wat groter uit, de andere keer wat kleiner. Gelukkig is dikwijls de aard van deze schommelingen vrij goed bekend. Men kan (uit waarnemingen of uit theoretische overwegingen) de kansverdeling van deze grootheden bepalen of althans benaderen. Wij kwantificeren dan de onzekerheid. Ondanks de onzekerheid kunnen wij dan statistische uitspraken doen over de te verwachten ontwikkeling van het systeem.

Bij simulatie met stochastische gegevens zal de doelfunktie (tijd, kosten) eveneens een stochastische grootheid zijn. De kansverdeling van deze grootheid kan in theorie steeds worden afgeleid uit de kansverdelingen van de gegevens. Zo kunnen we bijvoorbeeld de verwachte kosten bepalen, of de kosten die behoudens een kans van 10% niet worden overschreden. Soms kan men de kansverdeling van de doelfunktie niet expliciet opschrijven, omdat de samenhang met de gegeven kansverdelingen te ingewikkeld is. Men gebruikt dan de zgn. Monte Carlo methode; d.w.z. men vervangt elk stochastisch gegeven door één van zijn mogelijke waarden en bepaalt de doelfunktie. Zou men dit één keer doen dan was het resultaat sterk afhankelijk van de gekozen waarde. Men herhaalt het proces daarom een groot aantal malen en kiest elke waarde van de gegevens met de juiste kans. Technieken daarvoor worden in de rest van dit hoofdstuk besproken. Men verkrijgt dan een serie waarnemingen van de doelfunktie. Volgens de wet van de grote aantallen geven deze waarnemingen, als ze maar talrijk genoeg zijn, een getrouw beeld van de kansverdeling.

In voorbeeld 2.1 wordt de kansverdeling van de doelfunktie expliciet bepaald. In voorbeeld 2.2 wordt de Monte Carlo methode toegelicht. Voorbeeld 2.3 is een bewerking van voorbeeld 1.1 met stochastische in plaats van constante bewerkingstijden. Wanneer men ook in voorbeeld 1.2 met stochastische tijdsduren zou werken, zou het systeem door de vele wisselwerkingen moeilijk hanteerbaar worden. In zulke gevallen adviseert men wel de tijdas in intervallen te verdelen (bijv. van een uur, of een halve dag) en na elk interval te kijken of de toestand veranderd is, en zo ja, hoe. Vooral als sommige tijdsduren continue verdelingen hebben, moet men de intervallen klein kiezen om een getrouw beeld van de feitelijke toestand te krijgen. Aan de andere kant leiden korte intervallen tot veel extra werk.

Voorbeeld 2.1

Een reiziger kan zowel per bus als per tram naar huis. Laat $\underline{x}$ resp. $\underline{y}$ de tijdsduur in uren zijn *) tot de aankomst van de eerstvolgende bus resp. tram. Er is gegeven dat $\underline{x}$ en $\underline{y}$ onafhankelijk en exponentieel verdeeld zijn met verwachting een kwartier (voor $\underline{x}$) en tien minuten (voor $\underline{y}$). Over hoeveel tijd verwacht de reiziger thuis te zijn, als zowel de tram als de bus een kwartier rijtijd hebben en hij daarna nog twee minuten moet lopen?

In uren uitgedrukt geldt $P[\underline{x} \leq x] = 1-e^{-4x}$ en $P[\underline{y} \leq y] = 1-e^{-6y}$. Als $\underline{w}$ de wachttijd bij de halte is tot de aankomst van het eerstkomend vervoermiddel (bus of tram), dan is $\underline{w} = \min(\underline{x},\underline{y})$. Dus geldt $P[\underline{w} > w] = P[\underline{x} > w \text{ en } \underline{y} > w] = P[\underline{x} > w] \cdot P[\underline{y} > w] = e^{-4w} \cdot e^{-6w} = e^{-10w}$. Kortom: $\underline{w}$ is exponentieel verdeeld met verwachting 0,1 uur = 6 minuten. De verwachte totale tijd is dus 6 + 15 + 2 = 23 minuten.

Voorbeeld 2.2

Een fabrikant weet uit ervaring dat het behandelen van x garantieclaims ongeveer $100x \cdot (1 + \frac{10}{x+1})$ gulden kost. Wat zijn de verwachte kosten voor een levering van 30 exemplaren, als elk exemplaar met kans 0,1 tot een garantieclaim leidt?

---

*) Stochastische grootheden worden hier genoteerd als onderstreepte letters.

Omdat het aantal claims binomiaal verdeeld is met n = 30 en p = 0,1, wordt het antwoord

$$(2.1) \qquad \sum_{j=0}^{30} \binom{30}{j} (0,1)^j \, (0,9)^{30-j} \cdot 100 \; j \; (1 + \frac{10}{j+1}).$$

Met een tabel van de binomiale verdeling kan men deze som vrij snel bepalen. Zonder tabel is het berekenen nogal bewerkelijk. Als men de beschikking heeft over aselecte getallen (zie §3) kan men ook als volgt te werk gaan.

Kies 30 aselecte cijfers. Laat hieronder j nullen voorkomen. Bepaal 100 j $(1 + \frac{10}{j+1})$. Herhaal dit N keer. Tel de resultaten op en deel door N. Elk aselect cijfer is met kans $\frac{1}{10}$ een nul, zodat elke j te beschouwen is als een aselecte trekking uit een binomiale verdeling met n = 30 en p = 0,1. Kortom: elke j is een realisering van het stochastische aantal garantieclaims. Omdat de gebeurtenis j = 0 een kans $(0,9)^{30}$ heeft, zal voor grote N ongeveer een fractie $(0,9)^{30}$ van de N realiseringen de waarde j = 0 opleveren (wet van de grote aantallen). Omdat voor j = 1 enz. soortgelijke overwegingen gelden, is

$$\sum_{j=0}^{30} \frac{n_j}{N} \, 100 \; j \; (1 + \frac{10}{j+1})$$

een goede benadering voor (2.1), als n_j het aantal van de N realiseringen voorstelt waarbij j nullen werden geteld.

Voorbeeld 2.3

Laten in voorbeeld 1.1 de tijdsduren van het ontwikkelen nog steeds gegeven zijn volgens tabel 1.1. De overige tijdsduren echter zijn onafhankelijk en normaal verdeeld. Zowel hun verwachting als hun variantie zijn gelijk aan de in tabel 1.1 gegeven waarden. De fotograaf weet dat bij constante duren de volgorde CBAD én de volgorde DCBA beide tot een minimale duur, n.l. 400 minuten (tabel 1.3), hebben geleid. Hij vraagt zich af welke van beide ordeningen onder de nieuwe gegevens de voorkeur verdient.

Men kan de tijdsduren aangeven door letters met indices. Laten a_1, a_2 en a_3 resp. de ontwikkeltijd, afdruktijd en afwerktijd van opdracht A voorstellen, en analoog voor B, C en D. Dan geldt voor de totale tijdsduur $\underline{t}$ bij volgorde CBAD

$$\underline{t} = c_1 + \underline{d}_3 + \underline{m},$$

waar $\underline{m}$ het maximum is van de tien grootheden

$a_1 + b_1 + d_1 + \underline{d}_2$; $a_1 + \underline{a}_2 + b_1 + \underline{d}_2$;

$\underline{a}_2 + \underline{b}_2 + \underline{c}_2 + \underline{d}_2$; $\underline{a}_2 + b_1 + \underline{b}_2 + \underline{d}_2$;

$a_1 + \underline{a}_2 + \underline{a}_3 + b_1$; $\underline{a}_2 + \underline{a}_3 + \underline{b}_2 + \underline{c}_2$;

$\underline{a}_2 + \underline{a}_3 + b_1 + \underline{b}_2$; $\underline{a}_3 + \underline{b}_2 + \underline{b}_3 + \underline{c}_2$;

$\underline{a}_3 + b_1 + \underline{b}_2 + \underline{b}_3$; $\underline{a}_3 + \underline{b}_3 + \underline{c}_2 + \underline{c}_3$.

Dit blijkt uit het netwerk voor de volgorde CBAD, dat in figuur 2.1 wordt gegeven.

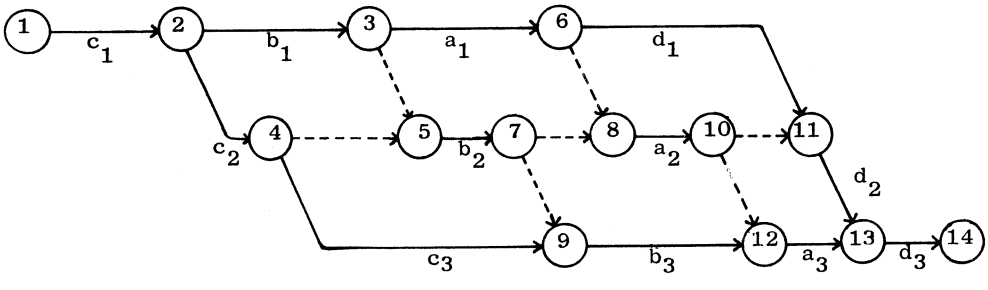


fig. 2.1

Netwerk voor de volgorde CBAD

Ieder van de tien grootheden is normaal verdeeld. Wegens de onafhankelijkheid van de termen binnen één grootheid kan men hun verwachtingen en varianties zonder moeite bepalen. Toch kan men de verdeling

van $\underline{m}$ niet vinden langs de weg van voorbeeld 2.1. Immers de tien grootheden zijn nu niet onafhankelijk verdeeld; ze bevatten paarsgewijs één of meer gelijke termen. Uit de simultane verdeling van de acht variabelen $\underline{a}_2$ t/m $\underline{d}_3$ mag dan in theorie die van $\underline{m}$ en die van $\underline{t}$ volgen, het expliciet opschrijven van deze verdeling is geen eenvoudige zaak. Wij volgen daarom de methode die bij voorbeeld 2.2 is geschetst.

Kies acht aselecte getallen. Transformeer deze tot acht aselecte trekkingen uit de standaardnormale verdeling (zie §4). Maak er realiseringen van $\underline{a}_2$ t/m $\underline{d}_3$ van door ze met de uit tabel 1.1 volgende standaardafwijking te vermenigvuldigen en er de verwachting uit tabel 1.1 bij op te tellen. Bepaal de waarde van m en vervolgens van t. Herhaal dit N keer.

Bij uitvoering van dit voorschrift ontstaan N waarnemingen van de stochastische tijdsduur bij de volgorde CBAD. Wij kunnen nu op dezelfde wijze N waarnemingen van de duur bij DCBA produceren. De resultaten worden daarna vergeleken, bijvoorbeeld naar verwachting, naar mediaan, of welk ander kriterium dan ook. Wij hebben het hier geschetste experiment laten uitvoeren door de Electrologica X8 rekenautomaat van het Mathematisch Centrum. Een samenvatting van de resultaten vindt men in tabel 2.1.

Tabel 2.1

Resultaat van 2 series van 500 bepalingen van de totale tijdsduur $\underline{t}$ voor de volgorde CBAD en de volgorde DCBA

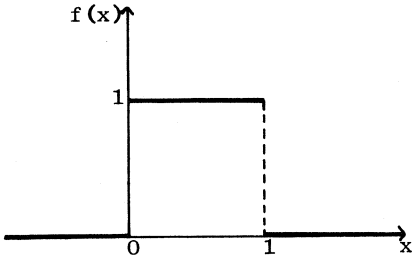
	volgorde CBAD		volgorde DCBA	
	1e serie	2e serie	1e serie	2e serie
$360 \leq t \leq 379$	22	42	4	8
$380 \leq t \leq 399$	205	211	161	181
$400 \leq t \leq 419$	235	207	287	255
$420 \leq t \leq 439$	38	40	45	51
$440 \leq t \leq 459$	0	0	3	5
Totaal	500	500	500	500
Gemiddelde	401,4	399,9	405,1	404,6
Mediaan	401	399	404	403
90-percentiel	417	418	419	420

Natuurlijk is de uitkomst van een Monte Carlo experiment altijd enigszins afhankelijk van de gebruikte aselecte getallen. Men ziet echter dat de twee series van 500 waarnemingen toch in sterke mate overeenstemmen. Bij constante tijdsduren (voorbeeld 1.1) was t = 400 zowel voor CBAD als voor DCBA. Men ziet dat de verwachte tijdsduur voor CBAD ongeveer 400 blijft, terwijl die voor DCBA tot ongeveer 405 oploopt. De fotograaf zal daarom aan CBAD de voorkeur geven, ook als hij let op de mediaan of de in 10% van de gevallen overschreden waarde van t. Per serie hebben wij dezelfde 8 × 500 aselecte getallen gebruikt voor beide volgorden. Het nut hiervan wordt in §5 toegelicht.

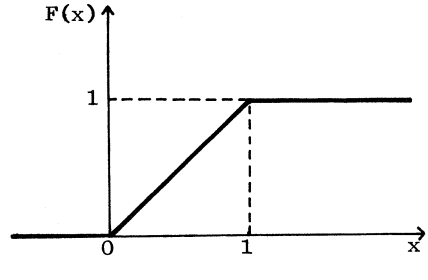
3. Aselecte getallen.

In de vorige paragraaf zijn we op het probleem gestuit van het nabootsen van een reeks onafhankelijke waarnemingen uit een gegeven kansverdeling. In §4 zal blijken dat het van speciaal belang is dit probleem op te lossen voor de homogene verdeling op het interval (0,1). Deze continue verdeling heeft een kansdichtheid f(x) en een verdelingsfunktie F(x), die er zeer eenvoudig uit zien

$$f(x) = \begin{cases} 0 & \text{als } x \leq 0, \\ 1 & \text{als } 0 < x < 1, \\ 0 & \text{als } x \geq 1, \end{cases} \qquad \text{en} \qquad F(x) = \begin{cases} 0 & \text{als } x \leq 0, \\ x & \text{als } 0 < x < 1, \\ 1 & \text{als } x \geq 1. \end{cases}$$



a. Verdelingsdichtheid



b. Verdelingsfunktie

fig. 3.1

De homogene verdeling op (0,1)

Wanneer een proces uitkomsten produceert die aselecte onafhankelijke trekkingen uit bovengenoemde verdeling voorstellen, dan noemt men die uitkomsten aselecte getallen. Schrijft men de aselecte getallen als decimale breuken, dan is voor elke decimaal elk van de cijfers 0, 1, ..., 9 even waarschijnlijk; zij hebben elk een kans $\frac{1}{10}$ om op te treden. Men noemt elke decimaal een aselect cijfer. Theoretisch zou een aselect getal nu een oneindig lange decimale breuk worden. Uiteraard is men in de praktijk tevreden met een breuk van bijvoorbeeld drie (of zes) decimalen, waarin elke decimaal één der getalwaarden 0 t/m 9 aanneemt met kans $\frac{1}{10}$. In feite vervangt men dan de continue homogene verdeling door een discrete, waarin elk der getallen 0,000; 0,001; ...; 0,999 een kans $\frac{1}{1000}$ heeft.

Bij het produceren van dergelijke aselecte getallen van drie cijfers heeft men vroeger gebruik gemaakt van hoge hoeden of vazen, waarin zich duizend papiertjes of schijfjes bevonden, waarop de getallen 0,000 t/m 0,999 waren geschreven. Na goed schudden werd één papiertje getrokken en het opgelezen getal werd genoteerd. Het papiertje werd teruggelegd, er werd opnieuw geschud en getrokken, enz. Uiteraard is dit een omslachtig en tijdrovend proces. Het kan enigszins vereenvoudigd worden door niet éénmaal uit een hoed met duizend briefjes te trekken, maar driemaal (met teruglegging) uit een hoed met tien. Elke trekking stelt dan één cijfer voor van het aselecte getal. Men zou verder aselecte trekkingen uit de getallen 0 t/m 36 kunnen produceren met een roulette en uit 1 t/m 6 met een zuivere dobbelsteen.

Omstreeks 1925 stelde TIPPETT voor tabellen van aselecte cijfers te maken. Hiertoe gebruikte men bijvoorbeeld de laatste vijf cijfers van elke in 20 decimalen getabelleerde logaritme. Er zijn ook tabellen geconstrueerd door een ronddraaiende schijf met tien gelijke sectoren op willekeurige tijdstippen stil te zetten en te noteren welke sector door een vaste wijzer werd aangewezen. Men kan zich in beide gevallen afvragen of de tabel inderdaad aselecte cijfers bevat. KENDALL en BABINGTON SMITH geven in hun bekende tabel*) een aantal statistische

---

*) M.G. KENDALL and B. BABINGTON SMITH, Tables of Random Sampling Numbers, Tracts for Computers no. 24, Cambridge University Press, Cambridge, (1939).

toetsen, bijv. op het even vaak voorkomen van elk cijfer of van elk paar opvolgende cijfers, op het aantal cijfers tussen twee nullen, enz. Bij de meeste tabellen van aselecte cijfers is aangegeven of aan de eisen van deze toetsen is voldaan.

Hierbij stuiten wij op een merkwaardige paradox. Wie vijf aselecte cijfers nodig heeft, kan de tabel op een willekeurige plaats openslaan en beginnen te lezen. Komt men dan bijvoorbeeld 0 2 3 8 5 tegen, dan zal men deze cijfers gebruiken. Zou men echter het vijftal 7 7 7 7 7 ontmoeten, dan zou men niet geneigd zijn dit als een groep aselecte cijfers te beschouwen; de kans op vijfmaal achter elkaar een 7 is $\frac{1}{10} \times \frac{1}{10} \times \frac{1}{10} \times \frac{1}{10} \times \frac{1}{10} = 10^{-5}$. Wanneer de tabel echter 100.000 aselecte cijfers bevat, dus 20.000 vijftallen, is de kans dat tenminste één van die vijftallen uitsluitend uit zevens bestaat gelijk aan $1 - (1 - 10^{-5})^{20000} \approx 0{,}18$. De kans dat minstens één van de 20.000 vijftallen uit vijfmaal hetzelfde cijfer bestaat is zelfs $1 - (1 - 10^{-4})^{20000} \approx 0{,}86$. Algemeen gezegd: wil de tabel als geheel een verzameling aselecte cijfers voorstellen, dan zullen ook zeer speciale combinaties zoals 1 2 3 4 5 of 7 7 7 7 7 hier en daar in de tabel moeten optreden. Is de tabel als geheel aselect, dan zullen bepaalde kleine onderdelen dat juist niet zijn. Het is dan ook gebruikelijk per groep van duizend cijfers te toetsen of aan de eisen van aselectheid is voldaan. Wanneer een enkel duizendtal niet voldoet, wordt het toch opgenomen in de tabel van aselecte cijfers, maar met een waarschuwende voetnoot. Wie slechts weinig duizendtallen aselecte cijfers wil gebruiken, moet deze gedeelten van de tabel overslaan. Het gebruik van minder dan duizend aselecte cijfers komt in de praktijk zelden voor. zodat toetsen voor kleinere gedeelten van de tabel niet gebruikelijk zijn.

Uit Monte Carlo methoden krijgt men natuurlijk steeds enigszins onzekere conclusies. Deze onzekerheid neemt ongeveer met een factor 10 af door 100 keer zoveel aselecte cijfers te gebruiken. Monte Carlo experimenten op grote schaal zijn dan ook pas in zwang gekomen sinds de ontwikkeling van elektronische rekenautomaten.

Stel dat wij zo'n rekenautomaat 100.000 aselecte cijfers willen laten gebruiken om bijv. een wachttijdprobleem te simuleren. Wij zouden dan een ponstypiste de 100.000 cijfers uit de tabel van KENDALL en BABINGTON SMITH op een ponsband kunnen laten zetten. Uiteraard is dit onzinnig tijdrovend. Zelfs als de computerbeheerder dit één keer had laten doen ten behoeve van alle toekomstige Monte Carlo gebruikers, resteert nog het probleem van het opbergen van deze 100.000 getallen in het geheugen van de machine. De werkruimte van zelfs een middelgrote computer laat dit niet toe en opbergen in een hulpgeheugen zou leiden tot groot tijdverlies bij het opzoeken.

Men heeft daarom al vroeg gezocht naar mogelijkheden om de machine zelf aselecte getallen te laten produceren. Omdat een rekenautomaat alleen op duidelijke en welgedefinieerde instructies reageert, is het principieel onmogelijk om hem strikt willekeurig uit de cijfers 0 t/m 9 te laten kiezen. Er zijn echter methoden om systematisch cijfers te produceren, waarbij het systeem zo gecompliceerd is dat het resultaat nauwelijks van een rij aselecte cijfers te onderscheiden is. Men noemt zo'n resultaat een verzameling pseudo-aselecte cijfers.

Een van de oudste voorbeelden van zo'n systeem om pseudo-aselecte cijfers te produceren is de "mid-square technique" van VON NEUMANN. Deze produceert eigenlijk geen aselecte cijfers maar aselecte getallen, en wel trekkingen uit de discrete homogene verdeling op 00, 01,..., 99. Laat y_i het i^{de} geproduceerde getal zijn; dan wordt y_{i+1} gevormd door de middelste twee cijfers van y_i^2. De keuze van y_i is hierbij willekeurig; d.w.z. moet door de programmeur gegeven worden. Twee voorbeelden vindt men in tabel 3.1.

Tabel 3.1

Mid-square techniek; twee voorbeelden

i	y_i	y_i^2	i	y_i	y_i^2
1	31	0961	1	79	6241
2	96	9216	2	24	0576
3	21	0441	3	57	3249
4	44	1936	4	24	0576
5	93	8649	5	57	3249
6	64	4096	6	24	0576
7	09	0081			
8	08	0064			
9	06	0036			
10	03	0009			
11	00	0000			

Twee bezwaren van de methode zijn hiermee gedemonstreerd. Zodra het tweede cijfer van y_i^2 een nul is convergeert de rij y_i in hoog tempo naar 0. Zodra een herhaling optreedt, is de rij verder periodiek (in de tweede kolom zelfs met periode 2). Hoewel allerlei verfijningen zijn aan te brengen om deze bezwaren weg te nemen, is de methode toch nergens meer in gebruik.

Een veel betere methode is de congruentiemethode van LEHMER. Men kiest hierbij een groot geheel getal m en vervolgens gehele getallen y_1, a en c tussen 0 en m-1 zodanig, dat c en m geen gemeenschappelijke factor hebben. Verder moet a-1 deelbaar zijn door elke priemfactor van m en door 4, als m een veelvoud van 4 is. Nu zijn y_i/m pseudo-aselecte trekkingen uit de homogene verdeling op (0,1), mits men voor y_{i+1} de rest kiest bij deling van $ay_i + c$ door m.

Ter illustratie kiezen wij m = 16, a = 3, c = 1 en $y_1 = 2$ (tabel 3.2), hoewel deze rij niet aan alle eisen voldoet.

Tabel 3.2

Congruentiemethode; voorbeeld

i	y_i	$ay_i + c$
1	2	7 = 0.16 + 7
2	7	22 = 1.16 + 6
3	6	19 = 1.16 + 3
4	3	10 = 0.16 + 10
5	10	31 = 1.16 + 15
6	15	46 = 2.16 + 14
7	14	43 = 2.16 + 11
8	11	34 = 2.16 + 2
9	2	7 = 0.16 + 7
10	7	...

Ook deze keuze leidt dus tot een periodieke rij (periode 8), in strijd met de eis van aselectheid. Iedere rij pseudo-aselecte getallen moet trouwens wel periodiek zijn, omdat na hoogstens m getallen herhaling gaat optreden. Men heeft echter vastgesteld, dat een geschikte keuze van m leidt tot een rij die wel periodiek is, maar een astronomisch lange periode heeft en ook overigens zeer bruikbaar is.

De keuzen $m = 2^{26}$, $a = 26.353.589$, $c = 1$ of $m = 2^{35}$, $a = 2^{18} + 1$, $c = 1$ zijn bijvoorbeeld bevredigend. Voor een snelle uitvoering van de delingen door m door de rekenmachine is het wenselijk, dat m van de vorm 2^p, 2^p-1 of 2^p+1 is (p willekeurig).

Er zijn allerlei elektrische en elektronische apparaten geconstrueerd, die op basis van ongeregelde verschijnselen, zoals ionen-emissie, aselecte getallen of cijfers produceren. Voor een beschrijving wordt verwezen naar K.D. TOCHER, The Art of Simulation, The English University Press, London, (1963).

4. Trekkingen uit verdelingen.

Bij het produceren van aselecte trekkingen uit een kansverdeling met gegeven verdelingsfunktie F(x) gaat men bijna steeds uit van (pseudo-) aselecte getallen. Wij zullen in deze paragraaf met de letter y steeds zo'n aselect getal aanduiden.

Omdat een verdelingsfunktie monotoon is, levert de instructie: "Trek een aselect getal y en zoek x, zodat F(x) = y" steeds een x met verdelingsfunktie F (zie fig. 4.1).

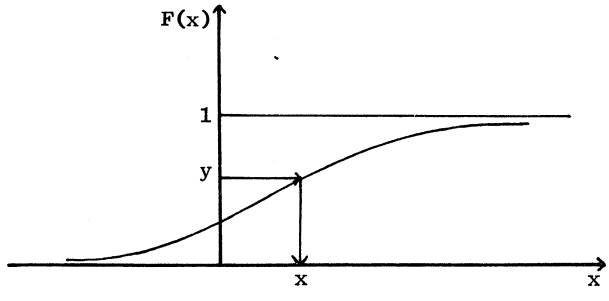


fig. 4.1

Directe inversie van een continue verdeling

Men noemt dit directe inversie en gebruikt het vooral in de volgende gevallen:

a) Het zoeken van x met F(x) = y is eenvoudig. Een voorbeeld is de exponentiële verdeling met $F(x) = 1-e^{-\lambda x}$, waarbij uit F(x) = y volgt $x = -\frac{1}{\lambda}\log(1-y)$. Omdat y en 1-y beide homogeen verdeeld zijn op (0,1) is het ook goed $x = -\frac{1}{\lambda}\log y$ te kiezen.

b) De verdelingsfunktie F is niet in formulevorm gegeven, maar wordt geschat uit tevoren verrichte waarnemingen $z_1, z_2, \ldots, z_n$. Als deze in opklimmende grootte zijn gerangschikt, zoekt men na het trekken van y het getal j, zodat $\frac{j-1}{n} < y \leq \frac{j}{n}$, en neemt $x = z_j$ (zie fig. 4.2).

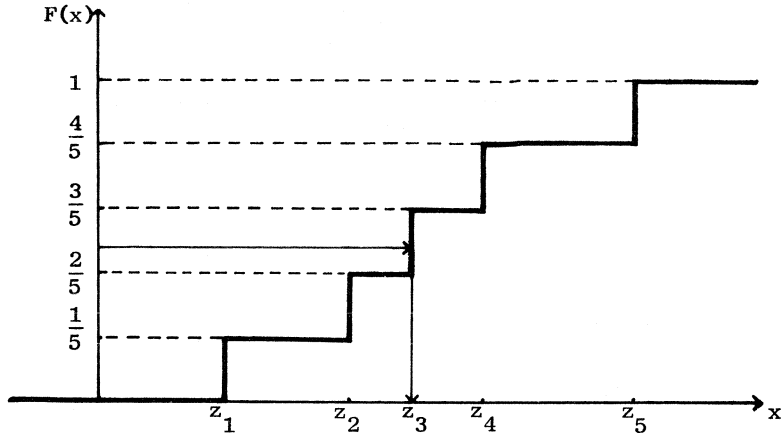


fig. 4.2

Directe inversie van een empirische verdeling

c) Discrete verdelingen waarbij een klein aantal waarden samen een zeer grote kans bezitten. Zo heeft de binomiale verdeling met n = 10 en p = 0,1 de kansen (zie fig. 4.3)

		individueel	cumulatief
$q(0) = \binom{10}{0} (0,9)^{10}$	=	0,349	0,349
$q(1) = \binom{10}{1} (0,1)\ (0,9)^{9}$	=	0,387	0,736
$q(2) = \binom{10}{2} (0,1)^{2}\ (0,9)^{8}$	=	0,194	0,930
$q(3) = \binom{10}{3} (0,1)^{3}\ (0,9)^{7}$	=	0,057	0,987
$q(4) = \binom{10}{4} (0,1)^{4}\ (0,9)^{6}$	=	0,011	0,998
$P(\underline{x} \geq 5) =$	=	0,002	

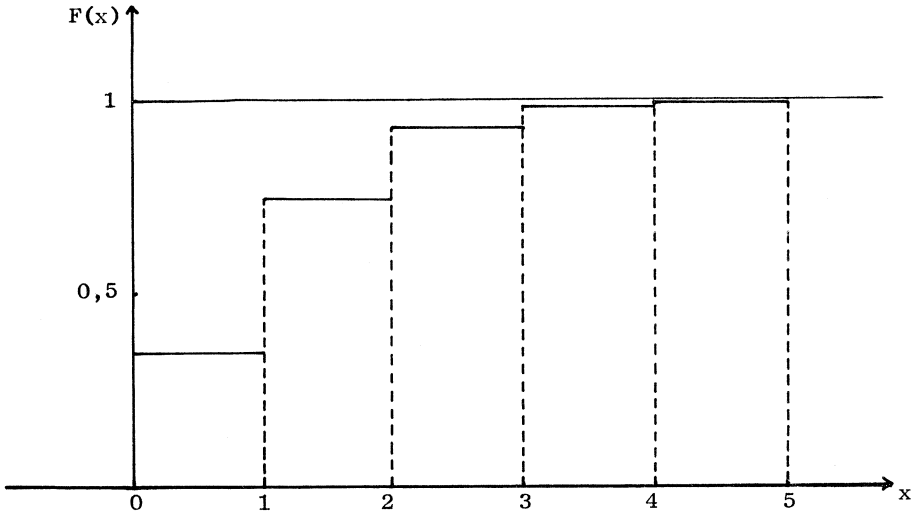


fig. 4.3

Binomiale verdeling met n = 10 en p = 0,1

Als het aselecte getal y kleiner is dan 0,349 noteert de computer dus 0 voor de "binomiale trekking" b. Voor $0{,}349 \leq y < 0{,}736$ noteert men b = 1, enz. Slechts in twee op de duizend gevallen behoeft de machine meer dan vijf getallen met y te vergelijken. De directe inversie zal hier snel en bevredigend verlopen.

Voor enige vaak voorkomende verdelingen geven wij hieronder een aantal trekkingsmethoden. Deze opsomming is verre van volledig. De keuze van de beste methode hangt af van de kosten, de vereiste nauwkeurigheid en het aantal gewenste trekkingen. Mede bepalend is de vereiste snelheid en, in verband met deze snelheid, de aard en uitrusting van de gebruikte rekenapparatuur, enz., enz.

a. Normale verdeling, $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$.

Als $\underline{u}$ standaardnormaal verdeeld is, heeft $\underline{u}\sigma+\mu$ de normale verdeling met verwachting μ en variantie σ^2. Het is dus voldoende trekkingen $\underline{u}$ uit N(0,1) te produceren.

a 1) BOX en MULLER produceren uit twee aselecte getallen y_1 en y_2 twee onafhankelijke "normal deviates"

$$\begin{cases} u_1 = \sqrt{-2 \log y_1}\ \cos 2\pi y_2, \\ u_2 = \sqrt{-2 \log y_1}\ \sin 2\pi y_2. \end{cases}$$

a 2) CENTRALE LIMIETSTELLING

$$u = (y_1 + y_2 + \ldots + y_n - \tfrac{1}{2}n)\sqrt{12/n}.$$

Bij de gangbare keuze n = 12 is dit in de staarten onnauwkeurig; echter wel zeer snel.

a 3) TEICHROEW *). Verbetering van a2) in de staarten; eist opslag van veel coëfficiënten in het computergeheugen; is wel nauwkeurig.

a 4) HASTINGS **). Numerieke benadering van de inverse verdelingsfunktie; werkt matig exact en matig snel.

a 5) WETHERILL ***). Polynomen met in totaal 10 vaste coëfficiënten, dus weinig belasting van het geheugen. De "normal deviate" is een eenvoudige funktie van een aselect getal; alleen in de staart wordt een logaritmische transformatie toegepast. De methode is minder exact, maar sneller dan a1) of a3).

a 6) Tabel van WOLD ****). Bevat 25.000 u-waarden. Zoals alle tabellen zeer geschikt voor verwerking met potlood en papier, maar minder handig voor een rekenmachine.

b. Exponentiële verdeling, $f(x) = \lambda e^{-\lambda x}$ voor $x \geq 0$. De directe inversie $x = -\frac{1}{\lambda} \log y$ is al genoemd.

---

*) D. TEICHROEW, Distribution Sampling with High Speed Computers, Ph.D.Thesis, University of North Carolina, (1953).

**) C. HASTINGS Jr., Approximations for Digital Computers (p. 192), Princeton University Press, Princeton, (1955).

***) G.B. WETHERILL, Generating Random Normal Deviates, Applied Statistics, 14, (1965), p. 201-205.

****) H.WOLD, Random Normal Deviates, Tracts for Computers no. 25, Cambridge University Press, Cambridge, (1948).

c. Chi-kwadraat-verdeling; Gammaverdeling; Erlangverdeling.
Som van onafhankelijke exponentiële trekkingen, of van kwadraten van onafhankelijke normale trekkingen.

d. Binomiale verdeling, $q(x) = \binom{n}{x} p^x (1-p)^{n-x}$ $(x=0, 1, \ldots, n)$.
Voor $p = \frac{1}{2}$ is x = aantal nullen onder n aselecte cijfers in het binaire stelsel, waarin bijna alle rekenmachines werken. Evenzo voor $p = \frac{1}{4}$: kies het aantal paren 00 onder n paren van twee aselecte binaire cijfers. Algemener voor succeskans p: kies x = het aantal van n aselecte getallen dat $\leq p$ is. Voor kleine n is directe inversie aan te bevelen.

e. Poisson-verdeling, $q(x) = e^{-\lambda} \frac{\lambda^x}{x!}$ $(x=0, 1, \ldots)$.

e 1) directe inversie

e 2) laten x_k $(k = 1, 2, \ldots)$ aselecte trekkingen uit de exponentiële verdeling met verwachting 1 zijn. Het getal m, zodanig dat

$$\sum_{k=1}^{m} x_k \leq \lambda < \sum_{k=1}^{m+1} x_k ,$$

is dan Poisson-verdeeld. Dit hangt samen met een eigenschap van Poisson-processen die bij de behandeling van wachttijden ter sprake is gekomen: de tussenaankomst-intervallen zijn dan en slechts dan exponentieel verdeeld als het aantal aankomsten in elk interval met vaste lengte Poisson-verdeeld is.

Voorbeeld 4.1

Ter illustratie van het aselect trekken uit kansverdelingen schetsen wij een probleem uit de wachttijdtheorie. Een reparateur heeft gedurende lange tijd genoteerd, op welke tijdstippen de opdrachten hem bereikten en hoeveel reparatietijd elke opdracht vergde. Uit statistische toetsen is gebleken, dat de intervallen tussen de opeenvolgende aankomsten van opdrachten kunnen doorgaan voor onafhankelijke trekkingen uit een gamma-verdeling met schaalparameter 1 en 2 vrij-

heidsgraden; dus met kansdichtheid $f(x) = x\,e^{-x}$ voor $x > 0$. De reparatietijden zijn voor opdrachten van dezelfde aard nagenoeg constant; en wel in 27% van de gevallen circa $\frac{1}{2}$, in 49% circa 1, in 14% circa 2 en in de resterende gevallen circa 3 tijdseenheden.

Men wenst door simulatie na te gaan met welke kans een binnenkomende opdracht moet wachten, omdat de reparateur al bezig is.

Van elk drietal aselecte getallen y_1, y_2, y_3 uit een tabel gebruiken we de eerste twee om met directe inversie twee exponentiële trekkingen te produceren: $x_i = -\lambda^{-1} \log y_i$ $(i = 1, 2)$, met in dat geval $\lambda = 1$. Hun som is dan een gamma-verdeeld aankomstinterval a. De derde trekking leidt, eveneens met directe inversie, tot een reparatietijd b; en wel volgens

$$b = \begin{cases} 0{,}5 & \text{als} \quad y_3 < 0{,}27, \\ 1 & \text{als} \quad 0{,}27 \leq y_3 < 0{,}76, \\ 2 & \text{als} \quad 0{,}76 \leq y_3 < 0{,}90, \\ 3 & \text{als} \quad 0{,}90 \leq y_3. \end{cases}$$

Voor een serie van elf drietallen y_1, y_2, y_3 vonden wij bijvoorbeeld het resultaat van tabel 4.1.

De eerste klant hoeft per definitie niet te wachten. Nu er van de resterende tien klanten drie moeten wachten kunnen we de onbekende kans op wachten schatten met 0,3. Hoe ongewis deze schatting is kan de lezer zelf nagaan door de berekening van hierboven met 33 andere aselecte getallen te herhalen! Het al dan niet moeten wachten van een klant is nu eenmaal in sterke mate afhankelijk van het al dan niet moeten wachten van zijn voorganger. Een paar opeenvolgende korte a-intervallen doen een lange rij ontstaan, die niet zo snel wordt weggewerkt. Pas als men een simulatie naar bovenstaand model over enige duizenden klanten uitstrekt, wordt het antwoord redelijk betrouwbaar.

Tabel 4.1

Simulatie van een wachtsituatie

Klant nr.	y_1	y_2	y_3	x_1	x_2	a	b	tijdstip van			wacht
								aankomst	begin rep.	eind rep.	
1	0,6558	0,0841	0,6288	0,422	2,476	2,898	1	2,898	2,898	3,898	--
2	0,6835	0,8157	0,1651	0,381	0,204	0,585	0,5	3,483	3,898	4,398	ja
3	0,8916	0,4899	0,8254	0,115	0,714	0,829	2	4,312	4,398	6,398	ja
4	0,4644	0,9027	0,8006	0,767	0,102	0,869	2	5,181	6,398	8,398	ja
5	0,0470	0,3070	0,7309	3,058	1,181	4,239	1	9,420	9,420	10,420	nee
6	0,1362	0,8427	0,3143	1,994	0,171	2,165	1	11,585	11,585	12,585	nee
7	0,3583	0,4472	0,2871	1,026	0,805	1,831	1	13,416	13,416	14,416	nee
8	0,5361	0,5937	0,4272	0,623	0,521	1,144	1	14,560	14,560	15,560	nee
9	0,5490	0,1532	0,5288	0,600	1,876	2,476	1	17,036	17,036	18,036	nee
10	0,7086	0,3686	0,1236	0,344	0,998	1,342	0,5	18,378	18,378	18,878	nee
11	0,3119	0,5347	0,2000	1,165	0,626	1,791	0,5	20,169	20,169	20,669	nee

5. Schatten van integralen; variantiereductie.

Een exact antwoord is bij het toepassen van Monte Carlo methoden nooit te bereiken. Onze berekeningen zijn immers gebaseerd op aselecte cijfers en het antwoord zal dus altijd enigszins afhankelijk zijn van de toevallige groep aselecte cijfers die gebruikt is. Wanneer tien proefpersonen elk honderd aselecte cijfers mogen produceren om daarvan het gemiddelde te bepalen, zullen er vermoedelijk tien verschillende antwoorden komen. Wel kunnen wij op grond van de statistische wetten van de grote aantallen verwachten, dat deze antwoorden geen van alle ver van 4,5 afliggen (het populatie-gemiddelde van de populatie bestaande uit 0, 1, 2, ..., 9) en dat zij om deze gemiddelde waarde verspreid liggen volgens een normale verdeling. Men kan direct narekenen, dat de populatievariantie $\sigma^2 = 8{,}25$ is; dus het gemiddelde van $n = 100$ trekkingen heeft een variantie $\sigma^2/n = 0{,}0825$ en een standaardafwijking $\sqrt{0{,}0825} = 0{,}29$. Ieder van de tien personen krijgt dus behoudens een kans van 5% een antwoord tussen $4{,}5 - 2 \times 0{,}29$ en $4{,}5 + 2 \times 0{,}29$. Als we voor elk antwoord een voorspellingsinterval met betrouwbaarheid $1-\alpha$ geven en α zó kiezen dat $(1-\alpha)^{10} = 0{,}95$, dan is $\alpha = 0{,}005$. Uit een tabel van de normale verdeling concluderen we, dat het interval $4{,}5 \pm 2{,}80 \times 0{,}29$ behoudens een kans van 5% alle tien antwoorden bevat. Zou hetzelfde Monte Carlo experiment dus tien keer zijn uitgevoerd, dan is het vrij zeker, dat de antwoorden telkens liggen tussen 3,69 en 5,31.

Het nadeel van de variabiliteit van de antwoorden wordt enigszins verminderd doordat wij over de te verwachten variabiliteit statistische uitspraken kunnen doen. Zou de opdrachtgever hierbij constateren, dat de variabiliteit (gemeten door de variantie) onaanvaardbaar groot is, dan kan hij besluiten het aantal gebruikte aselecte cijfers op te voeren. Omdat de variantie van het gemiddelde σ^2/n is, zouden tien gemiddelden van elk 10.000 trekkingen behoudens een kans van 5% alle liggen binnen $4{,}5 \pm 2{,}80 \times 0{,}029$; dus tussen 4,42 en 4,58. Wat in dit simpele geval exact juist is, geldt voor de meeste Monte Carlo experimenten als een vuistregel: door het aantal trekkingen k keer zo groot te maken wordt de standaardafwijking vermenigvuldigd met $1/\sqrt{k}$.

Tienmaal zo nauwkeurig betekent dus honderdmaal zoveel trekkingen. Deze methode tot opvoering van de nauwkeurigheid vindt dus al gauw zijn begrenzing in de kostbaarheid van het produceren van astronomische aantallen aselecte cijfers. Bij een ingewikkelde situatie is voor de bepaling van elk Monte Carlo antwoord een flink aantal aselecte trekkingen nodig. Het totale benodigde aantal voor een redelijke nauwkeurigheid is dan schrikbarend hoog. Gelukkig kan een middelgrote rekenautomaat in een redelijke tijd tienduizend of desnoods zelfs een miljoen aselecte getallen produceren.

Men heeft van het begin af gezocht naar middelen om zonder verhoging van het aantal trekkingen de variantie te verkleinen. Wij zullen straks een aantal van deze variantie reducerende methoden noemen. Bij het vergelijken van twee beslissingen (zoals twee volgorden van bewerking voor de fotograaf) is het aanbevelenswaardig dezelfde rij aselecte getallen voor beide Monte Carlo bepalingen te gebruiken. Een eventueel optredend verschil in het antwoord kan dan in geen geval ontstaan zijn doordat de ene rij toevallig anders uitviel dan de andere. Wij weten nu tenminste, dat het verschil bij deze rij trekkingen uitsluitend door het verschil in beslissing veroorzaakt is. Of het verschil bij tweemaal gebruiken van een andere rij niet juist andersom zou uitvallen, kan worden onderzocht door de variantie van de antwoorden te schatten.

Voor het toepassen van dit voorschrift is het prettig, dat er pseudo-aselecte getallen bestaan. Als wij bij de eerste beslissing "echte" aselecte getallen hadden gebruikt, zouden wij die allemaal moeten laten onthouden om dezelfde rij op de tweede situatie los te laten. Nu is het voldoende de computer twee keer dezelfde m, a, c en y van de congruentiemethode te geven. Wij weten dan zeker, dat de twee rijen exact overeenstemmen.

Allerlei technieken voor variantiereductie zullen wij nu illustreren aan de hand van het schatten van integralen. Voor de eenvoud kiezen we daarbij een enkelvoudige integraal, hoewel men in dat geval zelden Monte Carlo methoden gebruikt. Numerieke integratie levert doorgaans in minder tijd een beter resultaat. Bij meervoudige inte-

gralen worden Monte Carlo technieken aantrekkelijker. Zoals door HAMMERSLEY en HANDSCOMB *) wordt opgemerkt, is elk met aselecte trekkingen opgelost probleem te formuleren als het schattingsprobleem van een integraal. Het is dus nauwelijks een beperking als we de variantiereductie alleen bespreken voor het schatten van integralen.

Laat f(x) een gegeven funktie zijn, waarvoor

$$J = \int_0^1 f(x)\, dx$$

niet via een primitieve funktie te bepalen is. Wij zullen met diverse Monte Carlo methoden de waarde van J schatten. Iedere schatting is stochastisch, omdat het resultaat afhangt van de toevallig gebruikte aselecte cijfers. Iedere schatting heeft dus een variantie. Meestal hebben we de keus tussen snelle methoden met grote variantie en langzame met kleine variantie. Omdat de snelheid sterk afhankelijk is van de uitrusting van de gebruikte computer en de toelaatbare variantie van het doel van het onderzoek, is het niet mogelijk voor alle gebruikers tegelijk vast te stellen welke methode "de beste" is.

a. Ruw Monte Carlo

Zoals gewoonlijk stellen $y_1, y_2, \ldots, y_n$ aselecte getallen voor. Dan is

$$J_a = \frac{1}{n} \sum_{i=1}^{n} f(y_i)$$

een zuivere schatter voor J met variantie

$$\frac{1}{n} \int_0^1 (f(x) - J)^2\, dx.$$

Deze laatste integraal is uiteraard niet te bepalen omdat wij J niet kennen. Wij mogen de onbekende variantie van J_a echter schatten door

$$\frac{1}{n(n-1)} \sum_{i=1}^{n} (f(y_i) - J_a)^2.$$

---

*) Blz. 50 van J.M. HAMMERSLEY and D.C. HANDSCOMB, Monte Carlo Methods, Methuen, Londen, (1964). Dit belangrijke boekje wordt voortaan met H. en H. aangeduid.

b. Schietend Monte Carlo

Voor deze methode is het wenselijk dat f uitsluitend waarden tussen 0 en 1 aanneemt. Als dit niet het geval is, maar f is wel begrensd ($-M \leq f(x) \leq M$ voor alle x), dan gaat men over op $f^*(x) = \frac{f(x)+M}{2M}$. Deze funktie ligt wel tussen 0 en 1 en het is duidelijk, dat wegens $J^* = \int_0^1 f^*(x)\,dx = \frac{J}{2M} + \frac{1}{2}$ berekening van J^* voldoende is voor berekening van J.

Kies nu twee aselecte getallen y_{2i-1} en y_{2i} en scoor +1 (raak geschoten) als $y_{2i} \leq f^*(y_{2i-1})$. In het andere geval scoort men nul (misser). De fractie treffers onder n dergelijke schoten is een zuivere schatter van J^*. De hieruit bepaalde schatter voor J is eveneens zuiver en wordt met J_b aangeduid.

Ten opzichte van J_a heeft J_b twee nadelen. Ten eerste worden nu per bepaling twee aselecte getallen vereist in plaats van één, terwijl ook verder de bewerking meer tijd kost. Ten tweede kan men nagaan, dat de variantie van de schatter J_b aanzienlijk groter is dan die van J_a. Wij gaan om deze redenen niet verder op de methode in.

c. Stratificatie

Verdeel het integratie-interval in k stukken door de deelpunten $\alpha_0 = 0, \alpha_1, \alpha_2, \ldots, \alpha_{k-1}, \alpha_k = 1$. Kies voor elke $j = 1, 2, \ldots, k$ een groep van n_j aselecte getallen $y_{j,1}, y_{j,2}, \ldots, y_{j,n_j}$. Schat vervolgens J met

$$J_c = \sum_{j=1}^{k} \frac{\alpha_j - \alpha_{j-1}}{n_j} \sum_{i=1}^{n_j} f(\alpha_{j-1} + y_{j,i}(\alpha_j - \alpha_{j-1})).$$

De keuze van de deelpunten α_j en de aantallen n_j is hier vrij. Men moet deze vrijheid benutten om de variantie van de schatting J_c zo klein mogelijk te maken. Hoewel deelpunten op gelijke afstanden wel eens gebruikt worden, adviseert men meestal α_j zodanig te kiezen, dat de funktie f in elk stuk niet te veel varieert, en de keuze van n_j aan de lengte van de intervallen aan te passen. Men vindt verdere aanwijzingen in H. en H., blz. 55 *).

---

*) Zie vorige voetnoot (blz. 119).

d. Importantie

Voor elke funktie g geldt

$$J = \int_0^1 f(x)dx = \int_0^1 \frac{f(x)}{g(x)}\, dG(x), \text{ waar } G(x) = \int_0^x g(t)dt.$$

Wanneer wij zorgen, dat G een verdelingsfunktie is met G(1) = 1, zodanig dat men gemakkelijk aselecte trekkingen x_1, x_2, ... met deze verdelingsfunktie kan produceren, dan kan men methode a. toepassen op f/g en J schatten door

$$J_d = \frac{1}{n} \sum_{i=1}^{n} \frac{f(x_i)}{g(x_i)} .$$

Deze omweg is alleen voordelig als J_d een aanzienlijk kleinere variantie heeft dan J_a. Dit zal het geval zijn als de funktie f/g minder schommelingen vertoont dan f. Wanneer f positief is kunnen we proberen f/g ongeveer constant te houden, maar natuurlijk niet helemaal. Immers om de verdelingsfunktie G te kennen moeten we g integreren en dit gaat niet als g evenredig is met f, want f was een niet te integreren funktie. We zoeken dus een evenwicht tussen de tegenstrijdige eisen, dat g eenvoudig is en toch veel op f lijkt. In het Engels wordt deze techniek met "importance sampling" aangeduid.

e. Controlegrootheden

Voor elke funktie Q geldt

$$J = \int_0^1 f(x)dx = \int_0^1 Q(x)dx + \int_0^1 (f(x) - Q(x))dx.$$

Als de eerste term expliciet te integreren is en de tweede wordt geschat met methode a., dan ontstaat een schatting J_e voor J. Opdat J_e kleinere variantie heeft dan J_a, is nodig dat Q zo veel mogelijk op f lijkt en toch elementair integreerbaar is. Wij zien bij deze methode de algemene regel toegepast, dat men zo veel mogelijk wiskundig-theoretische methoden moet gebruiken en alleen waar

dat nodig is Monte Carlo technieken. Men noemt Q de controlegrootheid van f.

f. Diverse correlatiemethoden

In het reeds geciteerde boekje H. en H. worden nog diverse slimme methoden genoemd waarbij men paren aselecte trekkingen gebruikt die niet meer onafhankelijk zijn, maar sterk positief of negatief gecorreleerd. Zo schat men bij ruw Monte Carlo voor een monotone f wel met

$$\frac{1}{2n} \sum_{i=1}^{n} \{f(y_i) + f(1-y_i)\}.$$

Voor nadere details wordt naar de literatuur verwezen.

Slotwoord

De hier in het kort behandelde methoden voegen geheel nieuwe exemplaren toe aan de rijke gereedschapsvoorraad van de besliskundige. Vooral in de wachttijdtheorie zijn door Monte Carlo methoden problemen opgelost, die voordien de verdere voortgang van het onderzoek belemmerden. Men dient zich echter steeds bewust te zijn van de beperkte nauwkeurigheid en van de beperkte mogelijkheid tot generaliseren bij deze technieken. Een nooit slapend kostenbesef, een kritische geest en een inventief vermogen kunnen tezamen tot een nuttig gebruik van deze fascinerende hulpmiddelen leiden.

Literatuur

J.M. HAMMERSLEY and D.C. HANDSCOMB, Monte Carlo Methods, Methuen, London, (1964).

K.D. TOCHER, The Art of Simulation, The English University Press, London, (1963).

C.W. CHURCHMAN, R.L. ACKOFF and E.L. ARNOFF, Introduction to Operations Research, Wiley & Sons, New York, (1957).

M.G. KENDALL and B. BABINGTON SMITH, Tables of Random Sampling Numbers, Tracts for Computers no. 24, Cambridge University Press, Cambridge, (1939).

UITGAVEN IN DE SERIE MC SYLLABUS

Onderstaande uitgaven zijn verkrijgbaar bij het Mathematisch Centrum, 2e Boerhaavestraat 49 te Amsterdam-1005, tel. 020-947272.

MCS 1.1 F. GÖBEL & J. VAN DE LUNE, *Leergang Besliskunde, deel 1: Wiskundige basiskennis*, 1965. ISBN 90 6196 014 2.

MCS 1.2 J. HEMELRIJK & J. KRIENS, *Leergang Besliskunde, deel 2: Kansberekening*, 1965. ISBN 90 6196 015 0.

MCS 1.3 J. HEMELRIJK & J. KRIENS, *Leergang Besliskunde, deel 3: Statistiek*, 1966. ISBN 90 6196 016 9.

MCS 1.4 G. DE LEVE & W. MOLENAAR, *Leergang Besliskunde, deel 4: Markovketen, en wachttijden*, 1966. ISBN 90 6196 017 7.

MCS 1.5 G. DE LEVE & J. KRIENS, *Leergang Besliskunde, deel 5: Inleiding tot de mathematische besliskunde*, 1966. ISBN 90 6196 018 5.

MCS 1.6a B. DORHOUT & J. KRIENS, *Leergang Besliskunde, deel 6a: Wiskundige programmering 1*, 1968. ISBN 90 6196 032 0.

MCS 1.7a G. DE LEVE, *Leergang Besliskunde, deel 7a: Dynamische programmering 1*, 1968. ISBN 90 6196 033 9.

MCS 1.7b G. DE LEVE & H.C. TIJMS, *Leergang Besliskunde, deel 7b: Dynamische programmering 2*, 1970. ISBN 90 6196 055 X.

MCS 1.7c G. DE LEVE & H.C. TIJMS, *Leergang Besliskunde, deel 7c: Dynamische programmering 3*, 1971. ISBN 90 6196 066 5.

MCS 1.8 J. KRIENS et al., *Leergang Besliskunde, deel 8: Minimaxmethode, netwerkplanning, simulatie*, 1968. ISBN 90 6196 034 7.

MCS 2.1 G.J.R. FÖRCH et al., *Colloquium stabiliteit van differentieschema's, deel 1*, 1967. ISBN 90 6196 023 1.

MCS 2.2 L. DEKKER et al., *Colloquium stabiliteit van differentieschema's, deel 2*, 1968. ISBN 90 6196 035 5.

MCS 3.1 H.A. LAUWERIER, *Randwaardeproblemen, deel 1*, 1967. ISBN 90 6196 024 X.

MCS 3.2 H.A. LAUWERIER, *Randwaardeproblemen, deel 2*, 1968. ISBN 90 6196 036 3.

MCS 3.3 H.A. LAUWERIER, *Randwaardeproblemen, deel 3*, 1968. ISBN 90 6196 043 6.

MCS 4 H.A. LAUWERIER, *Representaties van groepen*, 1968. ISBN 90 6196 037 1.

MCS 5 J.H. VAN LINT et al., *Colloquium discrete wiskunde*, 1968. ISBN 90 6196 044 4.

MCS 6 K.K. KOKSMA, *Cursus ALGOL 60*, 1969. ISBN 90 6196 045 2.

MCS 7.1 *Colloquium moderne rekenmachines, deel 1*, 1969. ISBN 90 6196 046 0.

MCS 7.2 *Colloquium moderne rekenmachines, deel 2*, 1969. ISBN 90 6196 047 9.

MCS 8 H. BAVINCK & J. GRASMAN, *Relaxatietrillingen*, 1969. ISBN 90 6196 056 8.

MCS 9.1 T.M.T. COOLEN et al., *Colloquium elliptische differentiaalvergelijkingen, deel 1*, 1970. ISBN 90 6196 049 5.

MCS 10 J. FABIUS & W.R. VAN ZWET, *Grondbegrippen van de waarschijnlijkheidsrekening*, 1970. ISBN 90 6196 057 6.

MCS 11 H. BART et al., *Colloquium halfalgebra's en positieve operatoren*, 1971. ISBN 90 6196 067 3.

MCS 12 T.J. DEKKER, *Numerieke algebra*, 1971. ISBN 90 6196 068 1.

MCS 13 F.E.J. KRUSEMAN ARETZ, *Programmeren voor rekenautomaten; De MC ALGOL 60 vertaler voor de EL X8*, 1971. ISBN 90 6196 069 x.

MCS 14 H. BAVINCK et al., *Colloquium approximatietheorie*, 1971. ISBN 90 6196 070 3.

MCS 15.1 T.J. DEKKER et al., *Colloquium stijve differentiaalvergelijkingen, deel 1*, 1972. ISBN 90 6196 078 9.

MCS 15.2 P.A. BEENTJES et al., *Colloquium stijve differentiaalvergelijkingen, deel 2*, 1973. ISBN 90 6196 079 7.

*MCS 15.3 P.A. BEENTJES et al., *Colloquium stijve differentiaalvergelijkingen, deel 3*, 1972.

MCS 16.1 L. GEURTS, *Cursus programmeren, deel 1: De elementen van het programmeren*, 1973. ISBN 90 6196 080 0.

MCS 16.2 P.A. BEENTJES et al., *Cursus programmeren, deel 2: De programmeertaal ALGOL 60*, 1973. ISBN 90 6196 087 8.

MCS 17.1 P.S. STOBBE, *Lineaire algebra, deel 1*, 1974. ISBN 90 6196 090 8.

MCS 17.2 P.S. STOBBE, *Lineaire algebra, deel 2*, 1974. ISBN 90 6196 091 6.

MCS 18 F. VAN DER BLIJ et al., *Een kwart eeuw wiskunde, Syllabus van de Vakantiecursus 1971*, 1974. ISBN 90 6196 092 4.

MCS 19 A. HORDIJK et al., *Optimaal stoppen van Markovketens*, 1974. ISBN 90 6196 093 2.

MCS 20 T.M.T. COOLEN et al., *ALGOL 60 procedures voor begin- en randwaardeproblemen*, 1974. ISBN 90 6196 094 0.

MCS 21 J.W. DE BAKKER (Red.), *Colloquium Programmacorrectheid*, 1974. ISBN 90 6196 103 3.

*MCS 22 R. HELMERS et al., *Werkweek Statistiek*, 1973. ISBN 90 6196 104 1.

*MCS 23.1 J.W. DE ROEVER (Red.), *Colloquium Onderwerpen uit de Biomathematica, deel 1*, 1973. ISBN 90 6196 105 X.

*MCS 23.2 J.W. DE ROEVER (Red.), *Colloquium Onderwerpen uit de Biomathematica, deel 2*, 1973. ISBN 90 6196 115 7.

*MCS 24.1 P.J. VAN DER HOUWEN, *Numerieke integratie van differentiaalvergelijkingen, deel 1: Eenstapsmethoden*, 1974. ISBN 90 6196 106 8.

*MCS 25 L. GEURTS (Red.), *Colloquium Structuur van Programmeertalen*, 1974. ISBN 90 6196 116 5.

*MCS 26 N.M. TEMME (Red.), *Colloquium Niet-lineaire Analyse*, 1975. ISBN 90 6196 117 3.

De met een * gemerkte uitgaven moeten nog verschijnen.