

MC SYLLABUS 1.4

G. DE LEVE

W. MOLENAAR

LEERGANG BESLISKUNDE

DEEL 4 MARKOVKETENS EN WACHTTIJDEN

MATHEMATISCH CENTRUM AMSTERDAM 1966

AMS (MOS) classificatie (1970): 60-01, 60J10, 60K25

Tweede (gecorrigeerde) bijdruk: 1969
Derde bijdruk: 1973

Inhoud

	blz.
Voorwoord	1
1. Systeem, toestand, proces, keten	3
2. Ketens	4
3. Enkelvoudige Markovketens	9
4. Fuiken	12
5. De invariante verdeling	17
6. Voorbeelden	19
7. De wet van de grote aantallen	27
8. Verdere behandeling van de voorbeelden	29
9. Een verzekeringsprobleem	33
10. Andere Markovprocessen	41
11. Wachttijdsituaties	43
12. Het wachttijdproces	44
13. Het Poissonproces	46
14. Laplace-Stieltjes-transformaties	55
15. Een loketprobleem	56
16. Een machineprobleem	66
Literatuur	78

Voorwoord

Bij een groot aantal beslissingsproblemen kan het te bestuderen object (voorraad, machinepark, rij wachtende klanten) beschreven worden als een systeem dat in verschillende toestanden kan verkeren. Wanneer veranderingen in de toestand van het systeem niet met zekerheid voorspelbaar zijn maar wel volgens een vast patroon van kansen verlopen, kan men voor het beschrijven van het gedrag van het systeem dikwijls met succes gebruik maken van de mathematische theorie van de stochastische processen.

In dit vierde deel van de Leergang Besliskunde bespreken wij in hoofdzaak enkelvoudige Markovketens, omdat dit soort stochastische processen voor het genoemde doel van het meeste belang is. Met de behandelde theorie worden enige optimaliseringsproblemen opgelost; als voornaamste toepassing echter behandelen wij een aantal wachttijdproblemen.

Dit deel is een bewerking door Drs. W. MOLENAAR van de hoofdstukken IX en XV van Rapport S 265 (C 13): Leergang Besliskunde, door Prof.Dr. G. DE LEVE en Drs. F. GÖBEL. Bij het samenstellen van het gedeelte over wachttijden is gebruik gemaakt van Rapport S 246 (C 11) deel I, door Ir. A.R. BLOEMENA.

Amsterdam, 1966.

1. Systeem, toestand, proces, keten.

Bij onze bespreking van Markovketens zullen wij uitgaan van het abstracte begrip systeem. Een systeem kan een dobbelsteen zijn, maar ook een voorraad bij een groot winkelbedrijf, of een rij wachtenden voor een loket. Een systeem kan in verschillende toestanden verkeren. Als het systeem een dobbelsteen is, ligt het voor de hand om zes toestanden te onderscheiden. Deze worden gekarakteriseerd door het aantal ogen op het bovenliggende zijvlak van de dobbelsteen. Men merke op dat de ruimtelijke positie van de dobbelsteen hierdoor geenszins is bepaald. Als het systeem een rij wachtenden voor een loket is, correspondeert met ieder aantal wachtenden een bepaalde toestand van het systeem.

Wij beschouwen alleen systemen waarvan de toestanden beschreven kunnen worden door één of meer kwantitatieve grootheden. En wel nemen wij aan dat men iedere mogelijke toestand van het systeem kan vastleggen door h getallen, bijv. $(a_1, a_2, \dots, a_h)$. Zo'n serie getallen noemt men een toestandsvector.

Een proces vindt plaats wanneer een systeem achtereenvolgens verschillende toestanden aanneemt. Indien de waarnemer zelf de oorzaak is van de toestandswijzigingen, zoals bij het werpen met een dobbelsteen, dan spreekt men soms van een experiment ^{*)}. Verder onderscheidt men deterministische processen, waarbij de waarnemer tevoren weet wat het resultaat van de toestandsverandering zal zijn, en stochastische processen, waarbij dit niet het geval is.

Wanneer het systeem slechts op discrete tijdstippen in een andere toestand overgaat, kunnen we het proces beschrijven met de reeks van de bijbehorende toestandsvectoren. Een dergelijke reeks noemen we een

^{*)} Wij gebruiken de term "experiment" ook wel in de betekenis van "teweegbrenging van een toestandsverandering".

keten *). Zo is de reeks van het aantal ogen op de achtereenvolgens bovenliggende zijden bij het experiment "werpen met een dobbelsteen" een keten. Men merke op dat in dit voorbeeld h gelijk is aan één; de toestandsvector heeft slechts één component. Het is misschien nuttig er op te wijzen dat niet ieder fysisch proces (of experiment) door een keten beschreven kan worden: wanneer het systeem een stuwmeer is en de toestand wordt aangegeven door het peil, dan is het niet mogelijk een keten aan te wijzen die de "levensloop" van het systeem volledig beschrijft. De wijzigingen vinden immers niet op discrete tijdstippen plaats, maar min of meer continu.

2. Ketens.

In deze paragraaf en de volgende beschouwen we uitsluitend processen die door een keten kunnen worden beschreven en waarbij het aantal toestanden aftelbaar **) is. Laten we aannemen dat deze toestanden de nummers 1, 2, ... dragen.

Stel dat het systeem in de toestand i is, dan is het mogelijk dat het systeem na het eerstvolgende experiment weer in de toestand i is. Het is echter ook mogelijk dat het systeem overgaat in een andere toestand j . We zullen de kans op dergelijke overgangen aangeven door de overgangswaarschijnlijkheid p_{ii} resp. p_{ij} . Er geldt dus voor alle i

*) Er zijn ook auteurs die van een keten spreken als het aantal toestanden aftelbaar **) is, terwijl de toestandswijzigingen op een continuum van tijdstippen mogen optreden (zie § 10).

**) Men spreekt van aftelbaar veel toestanden als er een voorschrift ontworpen kan worden, zodanig dat aan iedere toestand op een ondubbelzinnige wijze een nummer kan worden toegevoegd. Is het systeem bijv. de rij wachtenden voor een loket, dan kan men de toestanden nummeren in overeenstemming met de lengte van de rij. Indien echter de toestand van het systeem gekenmerkt wordt door twee grootheden, bijv. het aantal mannen en het aantal vrouwen in de rij, dan is het nummeren een minder eenvoudige opgave. Ook nu is het aantal toestanden aftelbaar (ga zelf na!).

$$(2.1) \quad \sum_j p_{ij} = 1 ,$$

waarbij j alle toestanden doorloopt (dus ook de toestand i). Vrijwel steeds zullen wij stilzwijgend aannemen dat de overgangswaarschijnlijkheden stationair zijn, d.w.z. niet beïnvloed worden door het tijdstip van de toestandsverandering.

Ter verduidelijking beschouwen we het volgende voorbeeld. Uit een doos die 7 knikkers bevat (4 rode, 2 zwarte en 1 witte) worden op aselechte wijze knikkers getrokken. De doos is in dit voorbeeld het systeem. We zeggen dat het systeem in de toestand 1 (2, 3) verkeert als de kleur van de laatstgetrokken knikker rood (zwart, wit) is. De kleur van de getrokken knikker geeft dus per definitie de toestand van het systeem aan.

Experiment A

Beschouw nu het volgende proces: we trekken een knikker, leggen deze terug, trekken opnieuw een knikker, leggen deze terug, enz. De kleur van de laatstgetrokken knikker geeft steeds de toestand van het systeem aan.

De bijbehorende toestanden vormen een keten en we willen nu de overgangswaarschijnlijkheid p_{31} berekenen. Dat wil dus zeggen: als de laatstgetrokken knikker wit was, hoe groot is dan de kans dat de volgende rood is? Deze kans is, omdat vier van de zeven knikkers rood zijn en ieder gelijke kans heeft om getrokken te worden, natuurlijk gelijk aan $\frac{4}{7}$. Doordat de getrokken knikker steeds wordt teruggelegd verandert de samenstelling van de doos niet. Bijgevolg geldt ook, dat p_{11} en p_{21} gelijk zijn aan $\frac{4}{7}$. Dus $p_{i1} = \frac{4}{7}$ en evenzo $p_{i2} = \frac{2}{7}$, $p_{i3} = \frac{1}{7}$ voor $i=1,2,3$.

Als we de trekkingen op deze wijze uitvoeren hangen de overgangswaarschijnheden niet af van de uitgangstoestand. We spreken dan van een keten van onafhankelijke experimenten.

*) wh = waarschijnlijkheid,
whn = waarschijnlijkheden.

Experiment B

Nu trekken wij een knikker en vervolgens nog één, zonder dat de eerste is teruggelegd. In dit geval hangen de overgangswah wel af van het resultaat van de eerste trekking, zoals uit de volgende tabel blijkt:

$$(2.2) \quad \begin{cases} p_{11} = \frac{3}{6} & p_{12} = \frac{2}{6} & p_{13} = \frac{1}{6} \\ p_{21} = \frac{4}{6} & p_{22} = \frac{1}{6} & p_{23} = \frac{1}{6} \\ p_{31} = \frac{4}{6} & p_{32} = \frac{2}{6} & p_{33} = 0. \end{cases}$$

Voor het berekenen van de kansen p_{1i} ($i=1,2,3$) bedenke men dat de doos in dit geval na de eerste trekking nog 3 rode, 2 witte en 1 zwarte knikker bevat. De overige kansen vindt men op analoge wijze.

Experiment C

Trekken we uit de doos met zeven knikkers achtereenvolgens drie knikkers zonder teruglegging, dan wordt de situatie nog iets ingewikkelder. Laat nl. $p_{(k)ij}$ de kans zijn op de overgang van toestand i na de tweede trekking naar toestand j na de derde trekking, als het systeem na de eerste trekking in toestand k was. Dan geldt

$$(2.3) \quad \begin{cases} p_{(1)32} = \frac{2}{5} \\ p_{(2)32} = \frac{1}{5} \end{cases},$$

terwijl we aan $p_{(3)32}$ geen waarde kunnen toekennen, omdat $p_{33} = 0$.

Bij het berekenen van de kansen $p_{(k)ij}$ gaat men uit van het feit dat er na de tweede trekking nog vijf knikkers in de doos zijn.

We zien dus dat $p_{(k)ij}$ niet altijd onafhankelijk is van k . Wel volgt uit de wijze van berekenen onmiddellijk: $p_{(k)ij} = p_{(i)kj}$.

Experiment D

Weer wijzigen we ons experiment; we trekken een knikker en daarna een tweede. Alvorens nu de derde knikker te trekken leggen we de eerste

weer in de doos. Evenzo leggen we de tweede knikker terug voor de trekking van de vierde, enz. Zoals men gemakkelijk kan nagaan geldt nu $P_{(k)ij} = p_{ij}$; d.w.z. de overgangswahngen hangen wel af van de uitgangstoestanden, maar niet van de daaraan voorafgaande toestand(en) van het systeem. Aangezien er bij iedere trekking steeds zes knikkers in de doos zijn, kunnen we hiermee een willekeurig lange keten construeren, wat bij de experimenten B en C niet mogelijk was.

Het experiment D kunnen we nog als volgt generaliseren:

Experiment E

We beginnen met het trekken van $m+1$ knikkers zonder teruglegging. Nu leggen we de eerst getrokken knikker terug, trekken de $(m+2)^{de}$, leggen de tweede terug, trekken de $(m+3)^{de}$, enz. Hierbij is m een willekeurig niet-negatief geheel getal^{*)}. Als $m = 1$ is, voeren we experiment D uit, terwijl de keuze $m = 0$ equivalent is aan experiment A.

Experiment F

Tenslotte beschouwen we het volgende voorbeeld: de eerste twee trekkingen zijn zonder teruglegging; na de derde trekking leggen wij drie knikkers in de doos van dezelfde kleur als de knikker van de derde trekking. De vierde en vijfde knikker worden weer zonder teruglegging getrokken, terwijl na de zesde trekking weer drie knikkers van dezelfde kleur als de zesde knikker in de doos worden gelegd, enz.

Aldus ontstaat een oneindig lange keten, waarbij de overgangswahngen afhangen van alle voorgaande toestanden.

Wij onderscheiden nu vier soorten ketens:

- a) De overgangswahngen hangen niet af van de uitgangstoestanden. Deze ketens noemen we ketens van onafhankelijke experimenten (vgl. deel 2, §4) [Experiment A].

*) m moet vanzelfsprekend kleiner zijn dan het aantal knikkers dat bij het begin van het experiment in de doos ligt. Ga zelf eens na wat er gebeurt als m gelijk is aan dit aantal minus 1. Dit is een bijzonder geval van het in §4 te behandelen begrip kringfuiik.

- b) De overgangswahngen hangen alleen af van de uitgangstoestand.
Deze ketens noemen we enkelvoudige Markovketens. [Experiment D].
- c) De overgangswahngen hangen af van de uitgangstoestand en een begrensd aantal voorafgaande toestanden.
Deze ketens worden meervoudige Markovketens genoemd. [Experiment E].
- d) De overgangswahngen hangen af van de uitgangstoestand en van een onbegrensd aantal voorafgaande toestanden.
Deze ketens zullen wij niet bespreken.

In geval a) heeft het bereiken van toestand j een kans p_j die niet van de uitgangstoestand i afhangt. Neemt men nu bij m experimenten de toestand na het experiment waar, dan heeft men een aselechte steekproef van m waarnemingen verkregen van een stochastische grootte die met kans p_j de waarde j aanneemt. Behalve als een speciaal geval van een enkelvoudige Markovketen kan men a) dus ook beschouwen als de produktie van een steekproef van een discreet verdeelde stochastische grootte en dit legt een verband met de in deel 3 beschreven toetsings- en schattingstheorie.

Van de overige soorten van ketens zijn de enkelvoudige Markovketens de belangrijkste. Dit hangt samen met de omstandigheid dat de onder c) genoemde ketens door het invoeren van andere toestanden herleid kunnen worden tot enkelvoudige Markovketens. Dit kan men als volgt bewijzen.

Stel dat de overgangswahngen afhangen van de uitgangstoestand, en van hoogstens $m-1$ voorafgaande toestanden. Deze wahn kunnen we dus aanduiden door $p_{(i_1 \dots i_{m-1}) i_m j_m}$.

We zeggen nu dat het systeem zich in de toestand (nieuwe stijl) $(i_1, \dots, i_m)$ bevindt als het systeem in de toestand (oude stijl) i_m is, terwijl $i_1, \dots, i_{m-1}$ de voorafgaande toestanden zijn.

Als $T = (i_1, \dots, i_m)$ en $U = (j_1, \dots, j_m)$ twee toestanden (nieuwe stijl) zijn en p_{TU} is de overgangswahngen van T naar U , dan geldt

$$(2.4) \quad p'_{TU} = \begin{cases} p_{(i_1 \dots i_{m-1}) i_m j_m} & \text{als } i_2=j_1, i_3=j_2, \dots, i_m=j_{m-1}, \\ 0 & \text{in alle overige gevallen.} \end{cases}$$

Immers, uit de definitie van de toestand (nieuwe stijl) volgt dat twee opeenvolgende toestanden (nieuwe stijl) van het systeem $m-1$ toestanden (oude stijl) gemeenschappelijk hebben en dus geldt $p'_{TU} = 0$ als T en U geen $m-1$ toestanden (oude stijl) gemeenschappelijk hebben.

Verder geldt

$$\begin{aligned}
 \sum_U p'_{TU} &= \sum_{j_1, \dots, j_m} p'_{(i_1 \dots i_m)(j_1 \dots j_m)} = \\
 (2.5) \quad &= \sum_{j_m} p'_{(i_1 \dots i_m)(i_2 \dots i_m j_m)} = \sum_{j_m} p_{(i_1 \dots i_{m-1}) i_m j_m} = 1,
 \end{aligned}$$

zodat de p'_{TU} inderdaad op te vatten zijn als overgangswah. We kunnen ons dus beperken tot de enkelvoudige Markovketens. Als oefening kan de lezer zelf de 19 toestanden (nieuwe stijl) opschrijven bij experiment E met $m = 3$.

3. Enkelvoudige Markovketens.

Voorlopig beschouwen wij alleen systemen die in slechts eindig veel toestanden kunnen verkeren. De overgangswah noemen we weer p_{ij} ($i=1, \dots, n$; $j=1, \dots, n$).

Stelling 3.1

Laat $p_{ij}^{(m)}$ de kans zijn dat het systeem vanuit toestand i na precies m experimenten in de toestand j komt ^{*)}. Dan geldt voor gehele $m \geq 2$

$$(3.1) \quad p_{ij}^{(m)} = \sum_{k=1}^n p_{ik}^{(m-1)} p_{kj}.$$

Bewijs:

De kans dat het systeem na precies $m-1$ stappen van de toestand i in een toestand k is gekomen bedraagt, volgens definitie, $p_{ik}^{(m-1)}$, terwijl de kans om van k in j over te gaan (in één stap) gelijk is aan p_{kj} .

^{*)} $p_{ij}^{(1)}$ is dus hetzelfde als p_{ij} .

Het produkt $p_{ik}^{(m-1)} p_{kj}$ is dus (onafhankelijkheid!) de kans dat het systeem in $m-1$ stappen van i in k komt en vandaar naar j gaat. Dus $\sum_{k=1}^n p_{ik}^{(m-1)} p_{kj}$ (exclusieve gebeurtenissen!) is de kans dat het systeem na $m-1$ stappen in één of andere toestand is en vandaar naar j gaat, hetgeen niets anders betekent dan dat het systeem vanuit i na m stappen in j is gekomen, wat juist bewezen moest worden.

De vorm van de vergelijking (3.1) vertoont grote overeenkomst met die van een matrixprodukt (vgl. (5.3) uit deel 1). Het ligt daarom voor de hand om de volgende matrix in te voeren

$$(3.2) \quad \mathcal{P}^{(m)} = \begin{pmatrix} p_{11}^{(m)} & p_{12}^{(m)} & \dots & p_{1n}^{(m)} \\ p_{21}^{(m)} & p_{22}^{(m)} & \dots & p_{2n}^{(m)} \\ \vdots & \vdots & & \vdots \\ p_{n1}^{(m)} & p_{n2}^{(m)} & \dots & p_{nn}^{(m)} \end{pmatrix} \quad \text{voor } m=1,2,\dots$$

Nemen we $m = 1$, dan zijn de matrixelementen $p_{ij}^{(m)}$ weer de overgangswahns p_{ij} . De matrix $\mathcal{P}^{(1)}$, meestal afgekort tot $\mathcal{P}$, noemen we dan ook de matrix van de overgangswahns. Na bovenstaande definitie van $\mathcal{P}^{(m)}$ kunnen we nu de vergelijking (3.1) ook als volgt schrijven:

$$(3.3) \quad \mathcal{P}^{(m)} = \mathcal{P}^{(m-1)} \mathcal{P}.$$

Substitueren we hierin achtereenvolgens $m = 2, 3, \dots$ dan vinden we dat de matrix $\mathcal{P}^{(m)}$ juist het matrixprodukt van m factoren $\mathcal{P}$ is, dus

$$(3.4) \quad \mathcal{P}^{(m)} = \mathcal{P}^m.$$

Voorbeeld 3.1

Gegeven een systeem dat in drie mogelijke toestanden (1,2,3) kan verkeren; de toestanden van het systeem vormen een Markovketen, waarvan de matrix der overgangswahns $\mathcal{P}$ gegeven wordt door:

$$(3.5) \quad \mathcal{P} = \begin{pmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{1}{4} & \frac{3}{4} \\ \frac{3}{4} & \frac{1}{4} & 0 \end{pmatrix}.$$

Gevraagd de kans te berekenen dat het systeem zich na 3 stappen in de toestand 2 bevindt, als de begintoestand resp. 1, 2 en 3 is.

Oplossing:

Uit (3.5) volgt

$$(3.6) \quad \left\{ \begin{aligned} \mathcal{P}^2 &= \begin{pmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{1}{4} & \frac{3}{4} \\ \frac{3}{4} & \frac{1}{4} & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{1}{4} & \frac{3}{4} \\ \frac{3}{4} & \frac{1}{4} & 0 \end{pmatrix} = \\ &= \begin{pmatrix} \frac{1}{16} + \frac{3}{16} & \frac{1}{8} + \frac{1}{8} + \frac{1}{16} & \frac{1}{16} + \frac{3}{8} \\ \frac{9}{16} & \frac{1}{16} + \frac{3}{16} & \frac{3}{16} \\ \frac{3}{16} & \frac{3}{8} + \frac{1}{16} & \frac{3}{16} + \frac{3}{16} \end{pmatrix} = \begin{pmatrix} \frac{4}{16} & \frac{5}{16} & \frac{7}{16} \\ \frac{9}{16} & \frac{4}{16} & \frac{3}{16} \\ \frac{3}{16} & \frac{7}{16} & \frac{6}{16} \end{pmatrix}. \end{aligned} \right.$$

$$(3.7) \quad \left\{ \begin{aligned} \mathcal{P}^3 &= \begin{pmatrix} \frac{4}{16} & \frac{5}{16} & \frac{7}{16} \\ \frac{9}{16} & \frac{4}{16} & \frac{3}{16} \\ \frac{3}{16} & \frac{7}{16} & \frac{6}{16} \end{pmatrix} \begin{pmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{1}{4} & \frac{3}{4} \\ \frac{3}{4} & \frac{1}{4} & 0 \end{pmatrix} = \\ &= \begin{pmatrix} \frac{4}{64} + \frac{21}{64} & \frac{4}{32} + \frac{5}{64} + \frac{7}{64} & \frac{4}{64} + \frac{15}{64} \\ \frac{9}{64} + \frac{9}{64} & \frac{9}{32} + \frac{4}{64} + \frac{3}{64} & \frac{9}{64} + \frac{12}{64} \\ \frac{3}{64} + \frac{18}{64} & \frac{3}{32} + \frac{7}{64} + \frac{6}{64} & \frac{3}{64} + \frac{21}{64} \end{pmatrix} = \begin{pmatrix} \frac{25}{64} & \frac{20}{64} & \frac{19}{64} \\ \frac{18}{64} & \frac{25}{64} & \frac{21}{64} \\ \frac{21}{64} & \frac{19}{64} & \frac{24}{64} \end{pmatrix}. \end{aligned} \right.$$

Uit (3.4) en (3.7) volgt:

$$(3.8) \quad \begin{pmatrix} p_{11}^{(3)} & p_{12}^{(3)} & p_{13}^{(3)} \\ p_{21}^{(3)} & p_{22}^{(3)} & p_{23}^{(3)} \\ p_{31}^{(3)} & p_{32}^{(3)} & p_{33}^{(3)} \end{pmatrix} = \begin{pmatrix} \frac{25}{64} & \frac{20}{64} & \frac{19}{64} \\ \frac{18}{64} & \frac{25}{64} & \frac{21}{64} \\ \frac{21}{64} & \frac{19}{64} & \frac{24}{64} \end{pmatrix}.$$

De gevraagde kansen kan men dus aflezen uit de tweede kolom van (3.8).

4. Fuiken.

Bij Markovketens komt het wel eens voor dat het systeem na een aantal experimenten in een zodanige toestand geraakt dat het sommige andere toestanden niet meer kan doorlopen. Men zegt dan dat het systeem zich in een fuik bevindt. Het bestaan van fuiken, en de algemenere kwestie van de onderlinge bereikbaarheid van de toestanden, zijn bepalend voor het gedrag van het systeem na vele experimenten. Het is bij de bestudering van de keten van belang, welke toestanden in een eindig aantal stappen vanuit een gegeven toestand bereikt kunnen worden.

Voorbeeld 4.1

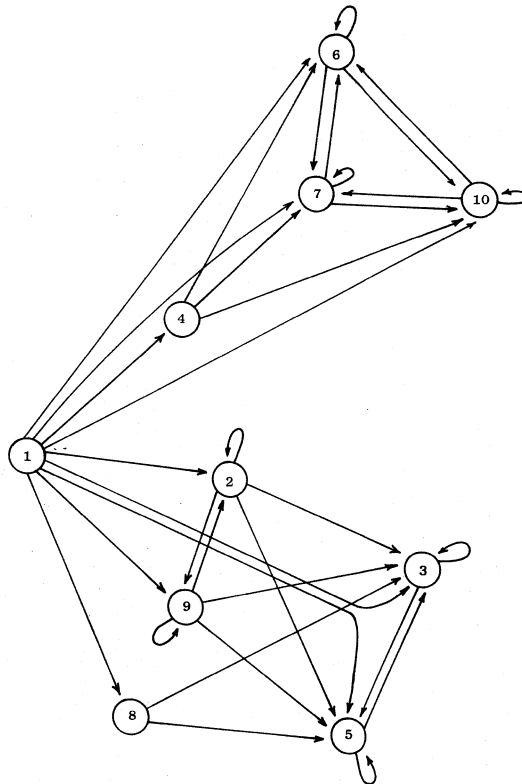
Laat een systeem dat slechts 10 toestanden kan aannemen, de in tabel 4.1 genoemde, bereikbare toestanden hebben.

Indien wij deze tabel en figuur 4.1 bestuderen, dan kunnen wij o.a. de volgende conclusies trekken:

- 1) Het systeem kan alleen aan het begin van de keten in de toestand 1 zijn. Daarna zal het nooit meer in deze toestand terugkeren.
- 2) Het systeem kan vanuit de toestand 2 wel in de toestanden 3 en 5 komen, maar niet omgekeerd. Vergelijk ook de toestanden 4 met 6, 7, 10; 8 met 3, 5; enz.
- 3) Indien het systeem eenmaal een toestand aangenomen heeft uit het toestandenpaar {3,5} of uit het drietal {6,7,10}, dan zal het systeem geen andere toestanden aannemen dan die uit {3,5} resp. {6,7,10}.

Tabel 4.1Overzicht van de in eindig veel stappen te bereiken toestanden

vanuit	zijn te bereiken in eindig veel stappen
1	alle overige
2	2,3,5,9
3	3,5
4	6,7,10
5	3,5
6	6,7,10
7	6,7,10
8	3,5
9	2,3,5,9
10	6,7,10

fig. 4.1Illustratie van tabel 4.1

Definitie 4.1

Toestand j heet bereikbaar vanuit toestand i , als voor minstens één eindige waarde van m geldt $p_{ij}^{(m)} > 0$. Toestand i heet een doorgangstoestand als er minstens één vanuit i bereikbare toestand j is, zodanig dat i niet bereikbaar is vanuit j . Toestanden die geen doorgangstoestanden zijn, worden terugkeertoestanden genoemd.

Een fuik (of stochastisch afgesloten verzameling) is een zodanige niet-lege verzameling van toestanden, dat vanuit deze toestanden geen toestanden buiten die verzameling bereikbaar zijn. Een priemfuik (of irreducibele fuik) is een fuik die geen twee of meer disjuncte (deel)-fuiken bevat. Een kernfuik (of ergodische klasse) is een fuik die geen kleinere fuik bevat.

In voorbeeld 4.1 is bijv. toestand 5 bereikbaar vanuit 8, maar niet omgekeerd; toestand 8 is dus een doorgangstoestand. Evenzo kan de lezer nagaan dat 2 en 4 doorgangstoestanden zijn en 3 en 5 terugkeertoestanden. Zijn er nog meer doorgangstoestanden?

Behalve de verzameling van alle toestanden (die bij elke Markovketen een fuik is; ga dit na) zijn er nog meer fuiken in dit voorbeeld, bijv. $\{4,6,7,10\}$, $\{2,3,5,9\}$, $\{3,5\}$, $\{3,5,6,7,8,10\}$. Hiervan zijn de eerste drie priemfuiken. De laatste bevat de disjuncte deelfuiken $\{3,5,8\}$ en $\{6,7,10\}$, zodat we hier geen priemfuik hebben. Kernfuiken zijn $\{3,5\}$ en $\{6,7,10\}$. Wij raden de lezer aan een lijst te maken van alle fuiken in dit voorbeeld en daarin de priemfuiken en kernfuiken aan te geven.

Om de namen doorgangs- en terugkeertoestand toe te lichten, vermelden wij een hier niet bewezen eigenschap van Markovketens met eindig veel toestanden : vanuit een terugkeertoestand keert het systeem met kans 1 na eindig veel stappen weer in dezelfde toestand terug; vanuit een doorgangstoestand is de kans op terugkeer binnen een eindig aantal stappen hetzij nul, hetzij een getal p met $0 < p < 1$. Men vergelijk de toestanden 3, 1 en 2 in fig. 4.1.

De volgende drie eigenschappen van Markovketens met eindig veel toestanden kan de lezer met enig nadenken zelf motiveren:

- a) elke kernfuik is een priemfuik, maar niet omgekeerd;
- b) een fuik kan niet uitsluitend uit doorgangstoestanden bestaan;
- c) een kernfuik bevat uitsluitend terugkeertoestanden, die allen onderling bereikbaar zijn.

Voorbeeld 4.2

Als de matrix der overgangswahns de volgende vorm zou hebben

$$(4.1) \quad \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \end{matrix} & \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix}$$

dan zijn de fuiken de volgende deelverzamelingen: $\{1,2,3,4\}$, $\{2,3,4\}$, $\{3,4\}$, $\{3\}$, $\{4\}$. Hiervan zijn $\{3\}$ en $\{4\}$ de enige priemfuiken en tevens de enige kernfuiken. Als een fuik uit slechts één toestand bestaat, noemt men deze toestand absorberend.

Voorbeeld 4.3

Beschouw de matrix

$$(4.2) \quad \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

(Het deterministische karakter van het bijbehorende proces doet niets af aan het feit dat we hier toch met een Markovproces te maken hebben).

We vinden de volgende fuiken: $\{1,2,3,4\}$, $\{2,3,4\}$, $\{3,4\}$ en $\{4\}$. Ieder van deze deelverzamelingen is zelfs een priemfuik; weliswaar bevat bijv. $\{2,3,4\}$ de fuiken $\{3,4\}$ en $\{4\}$, maar deze zijn niet disjunct. De enige kernfuik is $\{4\}$; 4 is de enige terugkeertoestand.

Voorbeeld 4.4

De matrix der overgangswahns heeft de volgende vorm, waarin alle getallen $p_{ij} \neq 0$ zijn.

$$(4.3) \quad \begin{pmatrix} p_{11} & p_{12} & p_{13} & p_{14} & p_{15} & p_{16} \\ 0 & p_{22} & p_{23} & p_{24} & 0 & 0 \\ 0 & p_{32} & p_{33} & p_{34} & 0 & 0 \\ 0 & p_{42} & p_{43} & p_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & p_{55} & p_{56} \\ 0 & 0 & 0 & 0 & p_{65} & p_{66} \end{pmatrix}$$

Bij deze keten zijn er twee kernfuiken: de eerste bestaat uit de toestanden 2, 3 en 4, de tweede uit de toestanden 5 en 6. Toestand 1 is een doorgangstoestand.

Definitie 4.2

Een kringfuik met periode M (waar M een geheel getal groter dan 1 is) is een fuik die men kan splitsen in deelverzamelingen $A_1, A_2, \dots, A_M$, zodanig dat voor $i=1, \dots, M$ vanuit een toestand behorend tot A_i alleen een direkte overgang mogelijk is naar een toestand die tot A_{i+1} behoort, waarbij we A_{M+1} met A_1 identificeren.

Voorbeeld 4.5

Beschouw de matrix

$$(4.4) \quad \begin{pmatrix} 1 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 4 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 5 & \frac{1}{4} & \frac{3}{4} & 0 & 0 & 0 & 0 & 0 \\ 6 & \frac{3}{4} & \frac{1}{4} & 0 & 0 & 0 & 0 & 0 \\ 7 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

De verzameling van alle toestanden is hier een kringfuik met periode 3, met als verzamelingen A_i achtereenvolgens $\{1,2\}$, $\{3,4\}$ en $\{5,6,7\}$. De verzameling van alle terugkeertoestanden $\{2,4,7\}$ is een kernfuik en tevens een kringfuik met periode 3.

5. De invariante verdeling.

In verband met de bestudering van de overige eigenschappen van Markovketens is het van belang te weten of de rij van getallen $p_{ij}^{(m)}$ een limiet *) heeft voor iedere i en j als m naar ∞ gaat.

Stelling 5.1

Indien de Markovketen geen kringfuiken bevat, dan bestaat voor alle i en j

$$(5.1) \quad \lim_{m \rightarrow \infty} p_{ij}^{(m)} = q_{ij}$$

en voor alle i geldt

$$(5.2) \quad \sum_{j=1}^n q_{ij} = 1 .$$

Zijn verder i en h toestanden die tot dezelfde priemfuike behoren, dan geldt voor alle j

$$(5.3) \quad q_{ij} = q_{hj} .$$

Als er geen disjuncte fuien bestaan, m.a.w. als er precies één kernfuike is, behoren alle toestanden tot dezelfde priemfuike en is q_{ij} dus onafhankelijk van i .

Stelling 5.2

Indien de Markovketen kringfuiken bevat, dan bestaat niet voor alle i en j $\lim_{m \rightarrow \infty} p_{ij}^{(m)}$, maar wel de zg. Cesarolimiet **)

$$(5.4) \quad \lim_{m \rightarrow \infty} \frac{1}{m} \sum_{k=1}^m p_{ij}^{(k)} = q_{ij}^c$$

en voor alle i geldt

$$(5.5) \quad \sum_{j=1}^n q_{ij}^c = 1 .$$

*) Zie deel 1, §8.

**) De Cesarolimiet van een rij kan blijkbaar bestaan zonder dat de gewone limiet bestaat. Maar als de gewone limiet bestaat, dan bestaat de Cesarolimiet eveneens en zijn de twee limieten aan elkaar gelijk.

Voor toestanden i en h uit dezelfde priemfuik geldt voor alle j

$$(5.6) \quad q_{ij}^c = q_{hj}^c .$$

In het bijzonder is q_{ij}^c onafhankelijk van i als er geen disjuncte fuiken zijn.

Wij zullen deze twee stellingen hier niet bewijzen *).

De kansen q_{ij} en q_{ij}^c kan men als volgt interpreteren. Stel dat men het systeem gedurende een zeer lange tijd niet heeft gadegeslagen en dat men alleen weet dat in deze tijd veel toestandswijzigingen hebben plaatsgevonden. Als het systeem het laatst in toestand i werd aangetroffen, dan zal de kans dat het systeem zich nu in toestand j bevindt bij benadering gelijk zijn aan q_{ij} of q_{ij}^c .

Beschouwen wij nog eens de vergelijkingen (3.1)

$$p_{ij}^{(m)} = \sum_{h=1}^n p_{ih}^{(m-1)} p_{hj} .$$

Hieruit volgt, indien er geen kringfuiken zijn,

$$(5.7) \quad \lim_{m \rightarrow \infty} p_{ij}^{(m)} = \sum_{h=1}^n \lim_{m \rightarrow \infty} p_{ih}^{(m-1)} p_{hj}$$

ofwel

$$(5.8) \quad q_{ij} = \sum_{h=1}^n q_{ih} p_{hj} ; \text{ d.w.z. } Q = QP .$$

Zijn er bovendien geen disjuncte fuiken, dan kunnen we deze vergelijking ook schrijven als

$$(5.9) \quad q_j = \sum_{h=1}^n q_h p_{hj} ; \text{ d.w.z. } Q = QP .$$

Als er wel kringfuiken zijn vinden wij uit (3.1)

$$\lim_{m \rightarrow \infty} \frac{1}{m} \sum_{k=1}^m p_{ij}^{(k)} = \sum_{h=1}^n \left\{ \lim_{m \rightarrow \infty} \frac{1}{m} \sum_{k=1}^m p_{ih}^{(k-1)} \right\} p_{hj} ,$$

*) Zie het in de literatuurlijst genoemde boek van FELLER, hoofdstuk XV, § 8.

ofwel

$$(5.10) \quad q_{ij}^c = \sum_{h=1}^n q_{ih}^c p_{hj} \quad ; \text{ d.w.z. } Q^c = Q^c \mathcal{P}.$$

Zijn er wel kringfuiken, maar geen disjuncte fuiken, dan geldt dus

$$(5.11) \quad q_j^c = \sum_{h=1}^n q_h^c p_{hj} \quad ; \text{ d.w.z. } Q^c = Q^c \mathcal{P}.$$

De kansverdeling met kansen q_{ij} , q_j , q_{ij}^c of q_j^c noemt men de invariante verdeling.

In de volgende paragraaf zullen wij in de gelegenheid zijn van enkele Markovketens de invariante kansen te berekenen naar aanleiding van enige praktijkvoorbeelden.

6. Voorbeelden.

Voorbeeld 6.1

Een handelaar in machines heeft van één type machine drie exemplaren verkocht. Het onderhoud van deze machines geschiedt onder toezicht van de handelaar. Hij heeft derhalve een voorraadjke reserve-onderdelen aangelegd, dat op maandagochtend kan worden aangevuld tot drie exemplaren ^{*)}. De kansverdeling van de vraag $\underline{d}$ per week naar een bepaald reserve-onderdeel wordt gegeven door

$$(6.1) \quad P[\underline{d}=0] = \frac{1}{6}, \quad P[\underline{d}=1] = \frac{1}{4}, \quad P[\underline{d}=2] = \frac{1}{3}, \quad P[\underline{d}=3] = \frac{1}{4}.$$

Verder wordt verondersteld dat de voorraadkosten per week evenredig zijn met de omvang van de voorraad aan het eind van die week. Zodra de voorraad niet toereikend is, worden noodinkopen verricht, waarvan de extra kosten per eenheid $10 \times$ de voorraadkosten per eenheid bedragen. Tenslotte wordt voor iedere aanvulling van de voorraad een vast bedrag

*) Wij nemen aan dat de opslagruimte een voorraad van meer dan 3 exemplaren niet toelaat. Wel zouden we kunnen toestaan dat de handelaar zijn voorraad steeds aanvult tot bijv. 2 exemplaren. Wij komen hierop nog terug.

gelijk aan $5 \times$ de voorraadkosten van één eenheid in rekening gebracht. Het probleem dat de handelaar nu stelt is het volgende: "Moet ik de voorraad aanvullen als ik nog één of misschien zelfs twee onderdelen in voorraad heb? Of moet ik pas bijbestellen als alle onderdelen gebruikt zijn?" *)

Oplossing:

De mogelijke toestanden aan het eind van iedere week zullen we aanduiden door symbolen (i) met de volgende betekenis:

- (3) $\equiv$ de voorraad bedraagt drie exemplaren,
- (2) $\equiv$ " " " twee exemplaren,
- (1) $\equiv$ " " " één exemplaar,
- (0) $\equiv$ er is geen voorraad, maar het was niet nodig om noodinkopen te doen,
- (-1) $\equiv$ er is éénmaal een noodinkoop verricht,
- (-2) $\equiv$ er is tweemaal een noodinkoop verricht.

De handelaar heeft nu drie strategieën tot zijn beschikking: bijbestellen (tot 3 stuks) wanneer de voorraad aan het eind van de week ten hoogste gelijk is aan 0, 1 of 2. Wij zullen deze drie strategieën aanduiden door resp. α , β en γ .

Wij vinden nu voor de overgangswah als strategie α wordt gevolgd:

$$(6.2) \quad \left\{ \begin{array}{l} p_{3,3} = p_{2,2} = p_{1,1} = P[\underline{d}=0] = \frac{1}{6} , \\ p_{3,2} = p_{2,1} = p_{1,0} = P[\underline{d}=1] = \frac{1}{4} , \\ p_{3,1} = p_{2,0} = p_{1,-1} = P[\underline{d}=2] = \frac{1}{3} , \\ p_{3,0} = p_{2,-1} = p_{1,-2} = P[\underline{d}=3] = \frac{1}{4} . \end{array} \right.$$

Aangezien de voorraad tot 3 wordt aangevuld als de voorraad aan het eind van de week gelijk is aan nul of minder, geldt

*) De handelaar kan ook besluiten nooit tevoren te bestellen en aan de gehele vraag te voldoen door noodinkopen. Deze mogelijkheid is hier niet beschouwd. Men kan zelf volgens de methoden van de paragrafen 6 en 8 nagaan dat het bij dit probleem zeer onvoordelig is uitsluitend met noodinkopen te werken. Bij welke gegevens omtrent vraag en kosten wordt dit wel voordelig?

$$(6.3) \quad \begin{cases} p_{0,-1} = p_{0,-2} = p_{-1,-1} = p_{-1,-2} = p_{-2,-2} = 0, \\ p_{0,3} = p_{-1,3} = p_{-2,3} = p_{3,3} = P[\underline{d}=0] = \frac{1}{6}, \\ p_{0,2} = p_{-1,2} = p_{-2,2} = p_{3,2} = P[\underline{d}=1] = \frac{1}{4}, \\ p_{0,1} = p_{-1,1} = p_{-2,1} = p_{3,1} = P[\underline{d}=2] = \frac{1}{3}, \\ p_{0,0} = p_{-1,0} = p_{-2,0} = p_{3,0} = P[\underline{d}=3] = \frac{1}{4}. \end{cases}$$

Verder zijn $p_{-2,-1} = p_{2,-2} = p_{3,-2} = p_{3,-1} = p_{1,2} = p_{1,3} = p_{2,3} = 0$.

Deze resultaten zijn samengevat in matrix (6.4).

$$(6.4) \quad \begin{matrix} & (3) & (2) & (1) & (0) & (-1) & (-2) \\ \begin{matrix} (3) \\ (2) \\ (1) \\ (0) \\ (-1) \\ (-2) \end{matrix} & \begin{pmatrix} \frac{1}{6} & \frac{1}{4} & \frac{1}{3} & \frac{1}{4} & 0 & 0 \\ 0 & \frac{1}{6} & \frac{1}{4} & \frac{1}{3} & \frac{1}{4} & 0 \\ 0 & 0 & \frac{1}{6} & \frac{1}{4} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{6} & \frac{1}{4} & \frac{1}{3} & \frac{1}{4} & 0 & 0 \\ \frac{1}{6} & \frac{1}{4} & \frac{1}{3} & \frac{1}{4} & 0 & 0 \\ \frac{1}{6} & \frac{1}{4} & \frac{1}{3} & \frac{1}{4} & 0 & 0 \end{pmatrix} \end{matrix} \quad \begin{matrix} \text{Matrix van over-} \\ \text{gangswaarschijn-} \\ \text{lijkheden bij} \\ \text{strategie } \alpha. \end{matrix}$$

Indien strategie β wordt gevolgd zijn overgangen van (1) naar (-1) en (-2) uitgesloten. De toestand (-2) kan zich nu niet meer realiseren. Op analoge wijze te werk gaande als bij strategie α vindt men voor de matrix der overgangswah

$$(6.5) \quad \begin{matrix} & (3) & (2) & (1) & (0) & (-1) \\ \begin{matrix} (3) \\ (2) \\ (1) \\ (0) \\ (-1) \end{matrix} & \begin{pmatrix} \frac{1}{6} & \frac{1}{4} & \frac{1}{3} & \frac{1}{4} & 0 \\ 0 & \frac{1}{6} & \frac{1}{4} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{6} & \frac{1}{4} & \frac{1}{3} & \frac{1}{4} & 0 \\ \frac{1}{6} & \frac{1}{4} & \frac{1}{3} & \frac{1}{4} & 0 \\ \frac{1}{6} & \frac{1}{4} & \frac{1}{3} & \frac{1}{4} & 0 \end{pmatrix} \end{matrix} \quad \begin{matrix} \text{Matrix van overgangs-} \\ \text{waarschijnlijkheden} \\ \text{bij strategie } \beta. \end{matrix}$$

$$(6.8) \quad q_3 = \frac{200}{2148}, \quad q_2 = \frac{360}{2148}, \quad q_1 = \frac{588}{2148}, \quad q_0 = \frac{567}{2148}, \quad q_{-1} = \frac{286}{2148}, \quad q_{-2} = \frac{147}{2148}.$$

Strategie β

$$(6.9) \quad \left\{ \begin{array}{l} q_3 = \frac{1}{6} q_3 + \frac{1}{6} q_1 + \frac{1}{6} q_0 + \frac{1}{6} q_{-1} \\ q_2 = \frac{1}{4} q_3 + \frac{1}{6} q_2 + \frac{1}{4} q_1 + \frac{1}{4} q_0 + \frac{1}{4} q_{-1} \\ q_1 = \frac{1}{3} q_3 + \frac{1}{4} q_2 + \frac{1}{3} q_1 + \frac{1}{3} q_0 + \frac{1}{3} q_{-1} \\ q_0 = \frac{1}{4} q_3 + \frac{1}{3} q_2 + \frac{1}{4} q_1 + \frac{1}{4} q_0 + \frac{1}{4} q_{-1} \\ q_{-1} = \frac{1}{4} q_2 \end{array} \right.$$

Hieruit volgt

$$(6.10) \quad q_3 = \frac{20}{156}, \quad q_2 = \frac{36}{156}, \quad q_1 = \frac{49}{156}, \quad q_0 = \frac{42}{156}, \quad q_{-1} = \frac{9}{156}.$$

Strategie γ

$$(6.11) \quad \left\{ \begin{array}{l} q_3 = \frac{1}{6} (q_3 + q_2 + q_1 + q_0) = \frac{1}{6} \\ q_2 = \frac{1}{4} (q_3 + q_2 + q_1 + q_0) = \frac{1}{4} \\ q_1 = \frac{1}{3} (q_3 + q_2 + q_1 + q_0) = \frac{1}{3} \\ q_0 = \frac{1}{4} (q_3 + q_2 + q_1 + q_0) = \frac{1}{4} \end{array} \right.$$

Wij hebben nu alle invariante verdelingen berekend. Het gestelde probleem is hiermee overigens nog niet opgelost. Het is ons zelfs nog niet duidelijk waarom we de invariante verdelingen hebben berekend. In paragraaf 7 zal echter blijken dat het kennen van deze verdelingen van groot nut is bij de verdere oplossing, die in paragraaf 8 zal worden uitgevoerd.

Voorbeeld 6.2

Een vrachtauto rijdt leeg van A naar B, en beladen van B naar A. Elke rit duurt een uur. Met een gegeven kans p ($0 < p < 1$) treedt onderweg motorstoring op, wat geen tijdverlies maar wel extra kosten met zich meebrengt. Bij iedere aankomst in A of B mag de vrachtrijder

beslissen of hij in het volgende uur zal terugrijden, dan wel de auto in een garage brengen voor een één uur durende preventieve onderhoudsbeurt. Zowel na een rit met motorpech als na een onderhoudsbeurt geldt, dat de eerstvolgende rit zeker zonder motorpech zal verlopen; bij daarna volgende ritten is de kans op pech weer gelijk aan p .

Het systeem (de vrachtauto) kan gedurende elk uur in de volgende toestanden verkeren:

- toestand 1: rijdend van A naar B, met pech;
- toestand 2: rijdend van B naar A, met pech;
- toestand 3: onderhoudsbeurt in A;
- toestand 4: onderhoudsbeurt in B;
- toestand 5: rijdend van A naar B, zonder pech;
- toestand 6: rijdend van B naar A, zonder pech.

Het is duidelijk dat het geen zin heeft op een rit met pech een onderhoudsbeurt te laten volgen, omdat dan een storingsvrije volgende rit al gegarandeerd is. Omdat de lege rit van A naar B alleen kosten geeft en de beladen rit van B naar A bovendien (grotere) inkomsten met zich meebrengt, lijkt het misschien wel zinvol bij aankomst in A anders te besluiten dan bij aankomst in B. Wij vergelijken daarom de volgende vier strategieën:

- $\alpha \equiv$ nooit onderhoud;
- $\beta \equiv$ alleen onderhoud na elke storingsvrije rit van A naar B;
- $\gamma \equiv$ alleen onderhoud na elke storingsvrije rit van B naar A;
- $\delta \equiv$ onderhoud na elke storingsvrije rit.

Oplossing:

Ook hier berekenen wij voorlopig alleen de invariante verdelingen. Uit de gegevens volgt, dat ongeacht de strategie voor de overgangswahns geldt: $p_{16} = 1$, $p_{25} = 1$, $p_{35} = 1$, $p_{46} = 1$, zodat vanuit de toestanden 1, 2, 3 en 4 geen andere overgangen mogelijk zijn.

Wanneer strategie α wordt gevolgd, zullen de toestanden 3 en 4 niet optreden. Vanuit toestand 5 wordt nu in elk geval direkt de terugreis aanvaard; d.w.z. het systeem gaat over in toestand 2 met kans p (pech op

de terugreis) en in toestand 6 met kans $1-p$ (geen pech). Met een analoge redenering vindt men $p_{61} = p$ en $p_{63} = 1-p$, zodat de matrix der overgangswahns luidt

$$(6.12) \quad \begin{matrix} & (1) & (2) & (5) & (6) \\ \begin{matrix} (1) \\ (2) \\ (5) \\ (6) \end{matrix} & \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & p & 0 & 1-p \\ p & 0 & 1-p & 0 \end{pmatrix} \end{matrix} .$$

Deze matrix heeft geen disjuncte fuiken, wel een kringfuik met periode 2. De vergelijkingen (5.11) voor de invariante verdeling worden

$$(6.13) \quad \left\{ \begin{array}{l} q_1^c = pq_6^c \\ q_2^c = pq_5^c \\ q_5^c = q_2^c + (1-p)q_6^c \\ q_6^c = q_1^c + (1-p)q_5^c \end{array} \right. .$$

De oplossing onder de voorwaarde $\sum q_j^c = 1$ is

$$(6.14) \quad q_1^c = q_2^c = \frac{p}{2+2p}, \quad q_5^c = q_6^c = \frac{1}{2+2p} .$$

Bij strategie β komen de toestanden 2 en 3 niet voor (waarom?) en men vindt achtereenvolgens de matrix

$$(6.15) \quad \begin{matrix} & (1) & (4) & (5) & (6) \\ \begin{matrix} (1) \\ (4) \\ (5) \\ (6) \end{matrix} & \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ p & 0 & 1-p & 0 \end{pmatrix} \end{matrix}$$

zonder enige kring- of disjuncte fuiken, de vergelijkingen

$$(6.16) \quad \left\{ \begin{array}{l} q_1 = pq_6 \\ q_4 = q_5 \\ q_5 = (1-p)q_6 \\ q_6 = q_1 + q_4 \end{array} \right.$$

en hun oplossing

$$(6.17) \quad q_1 = \frac{p}{3-p}, \quad q_4 = q_5 = \frac{1-p}{3-p}, \quad q_6 = \frac{1}{3-p}.$$

Voor strategie γ is de matrix

$$(6.18) \quad \begin{matrix} & (2)(3)(5)(6) \\ \begin{matrix} (2) \\ (3) \\ (5) \\ (6) \end{matrix} & \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ p & 0 & 0 & 1-p \\ 0 & 1 & 0 & 0 \end{pmatrix} \end{matrix}$$

Deze is na verwisseling van de toestanden 5 en 6 identiek met (6.15), zodat de invariante verdeling wordt

$$(6.19) \quad q_2 = \frac{p}{3-p}, \quad q_3 = q_6 = \frac{1-p}{3-p}, \quad q_5 = \frac{1}{3-p}.$$

Tenslotte vindt men bij strategie δ

$$(6.20) \quad \begin{matrix} & (3)(4)(5)(6) \\ \begin{matrix} (3) \\ (4) \\ (5) \\ (6) \end{matrix} & \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \end{matrix}$$

De verzameling van alle toestanden is hier een kringfuiik met periode 4: het systeem doorloopt met kans 1 de cyclus $3 \rightarrow 5 \rightarrow 4 \rightarrow 6 \rightarrow 3 \rightarrow 5 \rightarrow 4 \rightarrow \dots$. Men moet in dit geval (5.11) gebruiken in plaats van (5.9) en dit geeft

$$(6.21) \quad q_3^c = q_6^c; \quad q_4^c = q_5^c; \quad q_5^c = q_3^c; \quad q_6^c = q_4^c;$$

dus

$$(6.22) \quad q_3^c = q_4^c = q_5^c = q_6^c = \frac{1}{4}.$$

7. De wet van de grote aantallen.

Beschouw een willekeurige Markovketen met rationale overgangswahrscheinlichkeiten p_{ij} . Wij tonen nu aan dat deze keten kan worden verkregen door knikkers uit bakjes te trekken. Stel nl. dat n bakjes met knikkers in n verschillende kleuren zijn gevuld; zowel de bakjes als de knikkers zijn genummerd van 1 tot n . De bakjes zijn op een zodanige wijze gevuld, dat de kans om uit het i^{de} bakje een knikker van de kleur j te trekken gelijk is aan de gegeven waarde p_{ij} . Wij voeren nu het volgende experiment uit.

Trek een knikker uit bakje i_1 ; de kleur hiervan zij i_2 ; leg deze knikker weer terug. Trek dan de tweede knikker uit bakje i_2 ; de kleur hiervan zij i_3 ; enz. De rij $i_1, i_2, i_3, \dots$ is nu een Markovketen waarvan de overgangswahrscheinlichkeiten juist gelijk zijn aan p_{ij} , zoals men gemakkelijk kan inzien. De beginstap i_1 kunnen we bijv. bepalen met behulp van een apart bakje, waarin van iedere kleur één knikker ligt.

Stel nu dat men bij iedere trekking van een knikker met kleur i een boete $x(i)$ moet betalen, voor $i=1, \dots, n$. Laat verder x_m de boete zijn die men bij de m^{de} trekking moet betalen, voor $m=1, 2, \dots$. Deze grootte is stochastisch. Immers men weet van tevoren niet welke kleur de m^{de} getrokken knikker heeft. Men realiseert zich dat de stochastische variabele x_m één van de waarden $x(i)$ ($i=1, \dots, n$) zal aannemen.

Als het systeem zich bij het begin van het experiment in de toestand i bevindt, dan geldt voor de boete die bij de m^{de} trekking moet worden betaald

$$(7.1) \quad P \left[x_m = x(j) \right] = p_{ij}^{(m)}.$$

Er geldt dus *)

$$(7.2) \quad \mathbb{E} x_m = \sum_{j=1}^n x(j) p_{ij}^{(m)}.$$

*) Wij gaan hierbij steeds uit van een vaste begintoestand i .

Hieruit volgt

$$(7.3) \quad \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{m=1}^k \mathcal{E} \underline{x}_m = \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{m=1}^k \sum_{j=1}^n x(j) p_{ij}^{(m)} = \\ = \sum_{j=1}^n x(j) \left\{ \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{m=1}^k p_{ij}^{(m)} \right\} = \sum_{j=1}^n x(j) q_{ij}^c.$$

Men kan bewijzen dat het eerste lid van deze betrekking gelijk is aan

$$(7.4) \quad \mathcal{E} \left\{ \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{m=1}^k \underline{x}_m \right\}.$$

Als toestand i tot een priemfuik behoort, kan men verder aantonen dat de stochastische grootheid tussen de accoladen in (7.4) variantie nul heeft, zodat dan uit (7.3) volgt

$$(7.5) \quad \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{m=1}^k \underline{x}_m = \sum_{j=1}^n q_{ij}^c x(j) \quad \text{met kans } 1.$$

Verder kan men dan bewijzen dat uit (7.5) voor iedere $\varepsilon > 0$ volgt

$$(7.6) \quad \lim_{k \rightarrow \infty} P \left[\left| \frac{1}{k} \sum_{m=1}^k \underline{x}_m - \sum_{j=1}^n q_{ij}^c x(j) \right| > \varepsilon \right] = 0.$$

Dit betekent dus: de kans dat $\left| \frac{1}{k} \sum_{m=1}^k \underline{x}_m - \sum_{j=1}^n q_{ij}^c x(j) \right|$ groter is dan ε , is een functie van k die naar nul nadert als k naar ∞ gaat.

De betrekking (7.6) is het analogon voor Markovprocessen van de wet van de grote aantallen (deel 2, stelling 12.3) voor onafhankelijke experimenten.

Stelling 13.2 van deel 2 leert ons dat voor onafhankelijke experimenten de grootheid $\frac{1}{m} \sum_{k=1}^m \underline{x}_k$ voor grote m bij benadering een $N \left(\underline{x}_1, \frac{\sigma(\underline{x}_1)}{\sqrt{m}} \right)$ -verdeling heeft, mits de grootheden $\underline{x}_k$ allen dezelfde verdeling hebben. Men kan bewijzen dat voor Markovketens een soortgelijk resultaat geldt.

8. Verdere behandeling van de voorbeelden.

Een handelaar of fabrikant die zijn kosten wil verminderen, komt al spoedig voor het probleem te staan over welke periode hij moet minimaliseren. Een periode van 1 jaar ligt nogal voor de hand, maar 2 jaar is waarschijnlijk beter en 3 jaar nòg beter. Wij zouden zo kunnen doorgaan, maar de kosten over een oneindig lange periode zijn altijd oneindig groot en hieraan valt niet veel te minimaliseren.

Beschouwen we echter achtereenvolgens:

de kosten gedurende één jaar,

de gemiddelde kosten per jaar over een periode van 2 jaar,

de gemiddelde kosten per jaar over een periode van 3 jaar,

enz.,

dan krijgen we een rij van grootheden waarvan het minimaliseren steeds zinvoller wordt, terwijl deze getallen bovendien de prettige eigenschap hebben dat ze niet oneindig groot worden. We zullen dus steeds trachten de volgende grootheid te minimaliseren:

$$(8.1) \quad \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{m=1}^k x_m,$$

waarin x_m de kosten zijn die in de m^{de} periode worden gemaakt. Nu is de uitdrukking (8.1) volgens formule (7.5) met kans 1 gelijk aan

$$(8.2) \quad \sum_{j=1}^n q_{ij}^c x(j).$$

Indien we dus niet (8.1) maar (8.2) minimaliseren, zullen we zeer vaak (nl. met kans 1) ons doel bereiken.

Dit passen we nu toe bij de voorbeelden van paragraaf 6, waar we de invariante verdelingen reeds hebben berekend. q_{ij}^c kunnen we in beide voorbeelden nog vervangen door q_j^c , omdat in de matrices van paragraaf 6 geen disjuncte fuiken voorkomen. We zullen dus steeds dié strategie gebruiken, waarvoor $\sum_{j=1}^n q_j^c x(j)$ zo klein mogelijk is.

*) De index c in q_j^c kan in dit voorbeeld, wegens het ontbreken van kring-fuiken, steeds worden weggelaten.

Voorbeeld 6.1 (vervolg)

Laten we eerst het geval beschouwen waarin de handelaar strategie α toepast. In paragraaf 6 hebben wij aangenomen dat de voorraadkosten per week evenredig zijn met de grootte van de voorraad aan het eind van die week. Kiezen wij de voorraadkosten per week per eenheid als geldeenheid, dan vinden wij voor strategie α

$$(8.3) \quad x(3) = 3, \quad x(2) = 2, \quad x(1) = 1.$$

Verder lezen wij in paragraaf 6, dat bij elke bestelling een bedrag van 5 geldeenheden aan onkosten wordt gemaakt, terwijl voor iedere noodinkoop 10 geldeenheden dienen te worden uitgegeven. Op grond van deze gegevens vinden wij

$$(8.4) \quad x(0) = 5, \quad x(-1) = 15, \quad x(-2) = 25.$$

Voor de uitdrukking ^{*)}

$$\sum_{j=-2}^3 q_j x(j)$$

vinden wij dus

$$(8.5) \quad \sum_{j=-2}^3 q_j x(j) = 3 \cdot \frac{200}{2148} + 2 \cdot \frac{360}{2148} + 1 \cdot \frac{588}{2148} + 5 \cdot \frac{567}{2148} + \\ + 15 \cdot \frac{286}{2148} + 25 \cdot \frac{47}{2148} = \frac{12708}{2148} \approx 5,9.$$

Past de handelaar strategie β toe, dan geldt

$$(8.6) \quad x(3) = 3, \quad x(2) = 2, \quad x(1) = 6, \quad x(0) = 5, \quad x(-1) = 15.$$

Hieruit vinden wij

$$(8.7) \quad \sum_{j=-1}^3 q_j x(j) = 3 \cdot \frac{20}{156} + 2 \cdot \frac{36}{156} + 6 \cdot \frac{49}{156} + 5 \cdot \frac{42}{156} + \\ + 15 \cdot \frac{9}{156} = \frac{771}{156} \approx 4,9.$$

Voor strategie γ geldt

$$(8.8) \quad x(3) = 3, \quad x(2) = 7, \quad x(1) = 6, \quad x(0) = 5,$$

zodat

$$(8.9) \quad \sum_{j=0}^3 q_j x(j) = 3 \cdot \frac{1}{6} + 7 \cdot \frac{1}{4} + 6 \cdot \frac{1}{3} + 5 \cdot \frac{1}{4} = 5,5.$$

De handelaar zal dus de voorkeur geven aan strategie β ; d.w.z. hij zal zijn voorraad aanvullen (tot drie stuks) zodra hij aan het einde van de week ten hoogste 1 reserve-onderdeel in voorraad heeft. Zouden we hebben toegelaten dat de handelaar ook tot 2 stuks mag bijbestellen, dan was het probleem hierdoor niet moeilijker geworden. De extra berekeningen zijn nl. van dezelfde aard als de hierboven uitgevoerde, en zelfs nog eenvoudiger. Uit het resultaat van deze extra berekeningen volgt, dat het voordeliger is als de handelaar toch steeds tot drie stuks bijbestelt. De besparing op de voorraadkosten weegt blijkbaar niet op tegen de nu veel hogere nood-inkoopkosten.

Voorbeeld 6.2 (vervolg)

Om uit te maken welke van de vier strategieën α , β , γ en δ voor de vrachtrijder optimaal is, moeten we ons afvragen welke grootte we wensen te minimaliseren. Wij zullen hiervoor weer de kosten kiezen. Stel dat voor elke onderhoudsbeurt een bedrag G aan de garage wordt betaald en stel dat de kosten van elk pechgeval (spoedreparatie) gelijk zijn aan S . Laat verder elke rit (leeg of beladen) rijkosten R inhouden, terwijl voor elke beladen rit (d.w.z. elke rit van B naar A) een vrachtprijs V wordt ontvangen. Dan geldt, ongeacht de strategie, voor de kosten $x(i)$ in toestand i

$$(8.10) \quad \begin{cases} x(1) = R+S, & x(2) = R+S-V, \\ x(3) = x(4) = G, & x(5) = R, \quad x(6) = R-V. \end{cases}$$

Bij strategie α , dus invariante verdeling (6.14), zijn de verwachte kosten

$$(8.11) \quad K_{\alpha} = \sum_i q_i^c x(i) = \frac{p}{2+2p} (R+S) + \frac{p}{2+2p} (R+S-V) + \\ + \frac{1}{2+2p} R + \frac{1}{2+2p} (R-V) = R + \frac{p}{1+p} S - \frac{1}{2} V.$$

Voor strategie β , met invariante verdeling (6.17), geldt

$$(8.12) \quad K_{\beta} = \frac{p}{3-p} (R+S) + \frac{1-p}{3-p} G + \frac{1-p}{3-p} R + \frac{1}{3-p} (R-V) = \\ = \frac{1}{3-p} \{ (1-p)G + 2R + pS - V \}.$$

Voor strategie γ , met invariante verdeling (6.19), geldt

$$(8.13) \quad K_{\gamma} = \frac{p}{3-p} (R+S-V) + \frac{1-p}{3-p} G + \frac{1}{3-p} R + \frac{1-p}{3-p} (R-V) = \\ = \frac{1}{3-p} \{ (1-p)G + 2R + pS - V \}.$$

Tenslotte geldt voor strategie δ , met invariante verdeling (6.22),

$$(8.14) \quad K_{\delta} = \frac{1}{4} (G + G + R + R - V) = \frac{1}{2} G + \frac{1}{2} R - \frac{1}{4} V.$$

Blijkbaar zijn β en γ even winstgevend. Wij vergelijken eerst de twee extreme strategieën α en δ

$$(8.15) \quad K_{\alpha} - K_{\delta} = -\frac{1}{2} G + \frac{1}{2} R + \frac{p}{1+p} S - \frac{1}{4} V,$$

zodat α minder kostbaar is zodra

$$(8.16) \quad \frac{p}{1+p} < \frac{V + 2G - 2R}{4S};$$

d.w.z.

$$(8.17) \quad p < p_0 \stackrel{\text{def}}{=} \frac{V + 2G - 2R}{4S + 2R - V - 2G}.$$

Evenzo kunnen we K_{α} met K_{β} vergelijken en K_{β} met K_{δ} , maar dit eist wel enig rekenwerk. Nu zijn de strategieën β en γ een soort compromisoplossingen, waarbij we nu eens α volgen (geen onderhoud) en dan weer δ (wel onderhoud). We gaan daarom eerst na, of de kostenfunctie K_{β} wellicht een lineaire combinatie van K_{α} en K_{δ} is; d.w.z. we zoeken

getallen x en y zodat $K_\beta = xK_\alpha + yK_\delta$. Omdat in K_α geen term met G en in K_δ geen term met S voorkomt, zou dan moeten gelden $x = \frac{1+p}{3-p}$ en $y = \frac{2-2p}{3-p}$. Inderdaad levert deze keuze ook de juiste coëfficiënten voor R en V , zodat de strategieën β en γ dezelfde kosten veroorzaken als een strategie waarbij de vrachtrijder een fraktie $\frac{1+p}{3-p}$ van de tijd strategie α volgt en de rest van de tijd strategie δ .

Hieruit volgt direkt dat β en γ voor $p \neq p_0$ nooit optimaal kunnen zijn; voor $p < p_0$ geldt $K_\alpha < K_\beta = K_\gamma < K_\delta$ en voor $p > p_0$ geldt $K_\delta < K_\beta = K_\gamma < K_\alpha$. Alleen voor $p = p_0$ zijn alle strategieën even kostbaar.

De vrachtrijder moet dus strategie α kiezen als $p < p_0$ en strategie δ als $p > p_0$. Wij merken nog op dat de kosten G , R , S en V zich zo kunnen verhouden dat de waarde (8.17) van p_0 groter dan 1 uitvalt, en wel als $V + 2G > 2R + 2S$. In dat geval is α te verkiezen voor alle waarden van p . Is $V + 2G < 2R$, dan voldoet geen enkele niet-negatieve waarde van p aan (8.16) en (8.17) en verdient δ steeds de voorkeur.

9. Een verzekeringsprobleem.

Een transportonderneming heeft bij een assurantiemaatschappij een schadeverzekering afgesloten. De voorwaarden, welke in de polis zijn beschreven, kunnen als volgt worden samengevat:

- 1) De verzekeringsmaatschappij kent 4 klassen van verzekerden. De transportonderneming betaalt in de eerste klasse een premie van f 1.200,-. Voor de overige klassen bedraagt de premie resp. f 900,-, f 700,- en f 600,-.

Indien de transportonderneming gedurende een kalenderjaar geen schade heeft geclaimd, dan wordt zij het volgende jaar, tenzij zij reeds tot de hoogste klasse behoort, ingedeeld in de daarop volgende hogere klasse. Voor het geval zij reeds tot de hoogste klasse behoort, blijft zij tot die klasse behoren. Indien echter de transportonderneming een claim tot schadevergoeding heeft ingediend, dan wordt zij het

volgende jaar gerekend te behoren tot de eerste klasse.

2) De transportonderneming behoeft slechts aan het eind van ieder jaar de totale schade op te geven.

Er wordt dan en slechts dan een schadevergoeding uitgekeerd als het schadebedrag groter is dan

- a) f 750,- voor de eerste klasse,
- b) f 600,- voor de tweede klasse,
- c) f 500,- voor de derde klasse,
- d) f 450,- voor de vierde klasse.

De laatstgenoemde bedragen zijn de zg. eigen risico's.

3) De uit te keren schadevergoeding is gelijk aan het schadebedrag verminderd met het eigen risico.

Het is zonder meer duidelijk dat de transportonderneming in de eerste klasse geen schade zal claimen welke kleiner is dan f 1.050,-; immers zij derft een premieverlaging voor het komende jaar, welke groter is dan de schadeuitkering. Ook voor de overige klassen kan men tot soortgelijke uitspraken geraken. Wellicht is het voor de eerste klasse nog voordeliger om de "claimgrens" hoger te kiezen dan f 1.050,-.

In deze toepassing zullen wij trachten de optimale "claimgrenzen" α_i ($i=1, \dots, 4$) voor de vier klassen te berekenen. Een en ander hangt uiteraard ook af van de kansverdeling van de totale schade per jaar. Indien de omvang van de schade wordt aangeduid met t , dan zullen wij aannemen dat t een exponentiële verdeling heeft met parameter $\frac{1}{1000}$ (zie deel 2, formules (7.20) en (7.21)); d.w.z. voor de verdelingsdichtheid $f(t)$ en de verdelingsfunctie $F(t)$ geldt, voor $t \geq 0$

$$(9.1) \quad f(t) = \frac{1}{1000} e^{-\frac{t}{1000}}; \quad F(t) = 1 - e^{-\frac{t}{1000}}.$$

Oplossing:

Uit het gestelde volgt, dat de transportonderneming gezien als systeem, zich in 4 toestanden kan bevinden. Deze toestanden zullen wij overeenkomstig het nummer van de klasse aangeven met de getallen 1 t/m 4.

Aangezien men in de eerste klasse geen schade zal claimen, welke kleiner is dan α_1 , wordt de overgangswah p_{12} gegeven door $F(\alpha_1)$. Op dezelfde wijze vindt men de overgangswah p_{23} en p_{34} , welke resp. gegeven worden door $F(\alpha_2)$ en $F(\alpha_3)$. De overgangswah p_{44} is gelijk aan $F(\alpha_4)$. Indien de transportonderneming tot de eerste klasse van verzekerden behoort, zal zij het volgende jaar tot de eerste of tot de tweede klasse behoren. Dit betekent dat $p_{11} + p_{12} = 1$, zodat p_{11} gelijk is aan $1-F(\alpha_1)$. Op dezelfde wijze vinden we dat de overgangswah p_{21} , p_{31} en p_{41} resp. gelijk zijn aan $1-F(\alpha_2)$, $1-F(\alpha_3)$ en $1-F(\alpha_4)$. De matrix der overgangswah heeft dus de volgende gedaante:

$$(9.2) \quad \begin{matrix} & (1) & (2) & (3) & (4) \\ \begin{matrix} (1) \\ (2) \\ (3) \\ (4) \end{matrix} & \begin{pmatrix} 1-F(\alpha_1) & F(\alpha_1) & 0 & 0 \\ 1-F(\alpha_2) & 0 & F(\alpha_2) & 0 \\ 1-F(\alpha_3) & 0 & 0 & F(\alpha_3) \\ 1-F(\alpha_4) & 0 & 0 & F(\alpha_4) \end{pmatrix} & \end{matrix} .$$

De invariante kansen q_1 t/m q_4 moeten nu voldoen aan de volgende betrekkingen (vgl. (5.9)):

$$(9.3) \quad \left\{ \begin{array}{l} q_1 = \{1-F(\alpha_1)\}q_1 + \{1-F(\alpha_2)\}q_2 + \{1-F(\alpha_3)\}q_3 + \{1-F(\alpha_4)\}q_4 \\ q_2 = F(\alpha_1)q_1 \\ q_3 = F(\alpha_2)q_2 \\ q_4 = F(\alpha_3)q_3 + F(\alpha_4)q_4 \end{array} \right. .$$

Hieruit volgt

$$(9.4) \quad \left\{ \begin{array}{l} q_1 = \frac{1}{c} , \\ q_2 = \frac{F(\alpha_1)}{c} , \\ q_3 = \frac{F(\alpha_1) F(\alpha_2)}{c} , \end{array} \right.$$

$$(9.4) \quad q_4 = \frac{F(\alpha_1) F(\alpha_2) F(\alpha_3)}{c(1-F(\alpha_4))},$$

waar

$$c = 1 + F(\alpha_1) + F(\alpha_1) F(\alpha_2) + \frac{F(\alpha_1) F(\alpha_2) F(\alpha_3)}{1 - F(\alpha_4)}.$$

Als het systeem zich in de j^{de} toestand bevindt, dan worden de volgende kosten gemaakt:

- 1) de premie A_j ,
- 2) de schade $\underline{t}$ als $\underline{t} \leq \alpha_j$,
- 3) de schade β_j als $\underline{t} > \alpha_j$, waarin β_j het eigen risico voorstelt.

De verwachting van de kosten voor deze toestand bedraagt dus

$$(9.5) \quad A_j + \int_0^{\alpha_j} t f(t) dt + \beta_j \int_{\alpha_j}^{\infty} f(t) dt.$$

Voor de verwachting van de kosten voor een "heel ver in de toekomst liggend" jaar vinden wij dus

$$(9.6) \quad H(\alpha_1, \dots, \alpha_4) = \sum_{j=1}^4 q_j \left\{ A_j + \int_0^{\alpha_j} t f(t) dt + \beta_j \int_{\alpha_j}^{\infty} f(t) dt \right\}.$$

Een stel waarden α_i ($i=1, \dots, 4$) kan men een strategie noemen en zoals in paragraaf 8 is aangetoond, zijn met kans 1 die waarden van α_i optimaal, waarvoor de uitdrukking (9.6) minimaal wordt. De optimale waarden van α_i moeten voldoen aan de relaties

$$(9.7) \quad \frac{\partial H(\alpha_1, \dots, \alpha_4)}{\partial \alpha_i} = 0 \quad (i=1, \dots, 4).$$

Deze voorwaarden zijn wel noodzakelijk, echter niet voldoende.

Aangezien q_j een functie is van α_i , is (9.7) gelijkwaardig met

$$(9.8) \quad \sum_{j=1}^4 \frac{\partial q_j}{\partial \alpha_i} \{A_j + \int_0^{\alpha_j} t f(t) dt + \beta_j \int_{\alpha_j}^{\infty} f(t) dt\} + q_i (\alpha_i - \beta_i) f(\alpha_i) = 0.$$

Uit (9.4) volgt voor $\frac{\partial q_j}{\partial \alpha_i}$:

$$(9.9) \quad \left\{ \begin{array}{ll} \text{a)} & \frac{\partial q_j}{\partial \alpha_i} = q_j \sum_{h=1}^i q_h \frac{f(\alpha_i)}{F(\alpha_i)} \quad \text{als } j \geq i+1, i \leq 3, \\ \text{b)} & \frac{\partial q_j}{\partial \alpha_i} = q_j \sum_{h=1}^i q_h \frac{f(\alpha_i)}{F(\alpha_i)} - q_j \frac{f(\alpha_i)}{F(\alpha_i)} \quad \text{als } j \leq i, i \leq 3, \\ \text{c)} & \frac{\partial q_j}{\partial \alpha_4} = -q_j q_4 \frac{f(\alpha_4)}{1-F(\alpha_4)} \quad \text{als } j \leq 3, \\ \text{d)} & \frac{\partial q_4}{\partial \alpha_4} = q_4 (1-q_4) \frac{f(\alpha_4)}{1-F(\alpha_4)}. \end{array} \right.$$

Wij beperken ons nu eerst tot $i \leq 3$. Dan mogen wij voor (9.8) schrijven

$$(9.10) \quad \begin{aligned} & \frac{f(\alpha_i)}{F(\alpha_i)} \sum_{h=1}^i q_h \sum_{j=1}^4 q_j \{A_j + \int_0^{\alpha_j} t f(t) dt + \beta_j \int_{\alpha_j}^{\infty} f(t) dt\} + \\ & - \frac{f(\alpha_i)}{F(\alpha_i)} \sum_{j=1}^i q_j \{A_j + \int_0^{\alpha_j} t f(t) dt + \beta_j \int_{\alpha_j}^{\infty} f(t) dt\} + \\ & + q_i (\alpha_i - \beta_i) f(\alpha_i) = 0, \end{aligned}$$

ofwel

$$(9.11) \quad \frac{f(\alpha_i)}{F(\alpha_i)} \sum_{h=1}^i q_h \{H(\alpha_1, \dots, \alpha_4) - H_i(\alpha_1, \dots, \alpha_4) + \frac{q_i F(\alpha_i)}{\sum_{k=1}^i q_k} (\alpha_i - \beta_i)\} = 0$$

waarbij

$$(9.12) \quad H_i(\alpha_1, \dots, \alpha_4) \stackrel{\text{def}}{=} \sum_{j=1}^i \frac{q_j}{\sum_{h=1}^i q_h} \{A_j + \int_0^{\alpha_j} t f(t) dt + \beta_j \int_{\alpha_j}^{\infty} f(t) dt\}.$$

Uit (9.11) volgt:

$$(9.13) \left\{ \begin{array}{ll} \text{a)} & \frac{f(\alpha_i)}{F(\alpha_i)} = 0 \\ \text{of} & \\ \text{b)} & \sum_{h=1}^i q_h = 0 ; \quad \text{d.w.z.} \quad q_h = 0, \quad h \leq i \\ \text{of} & \\ \text{c)} & H(\alpha_1, \dots, \alpha_4) - H_i(\alpha_1, \dots, \alpha_4) + \frac{q_i F(\alpha_i)}{\sum_{h=1}^i q_h} (\alpha_i - \beta_i) = 0. \end{array} \right.$$

Aan (9.13a) en (9.1) voldoet geen enkele eindige waarde van α_i . Wel is $\lim_{x \rightarrow \infty} \frac{f(x)}{F(x)} = 0$ en de waarden $\alpha_i = \infty$ voldoen aan (9.7). Langs wiskundige weg kan men aantonen dat ze toch niet optimaal zijn, hetgeen ook op intuïtieve gronden voor de hand ligt; immers er wordt wel premie betaald, doch er worden nooit schaden geclaimd. Uit (9.13b) en (9.4) volgt $F(\alpha_4) = 1$, m.a.w. in de 4^{de} klasse wordt geen schade geclaimd. Ook deze beslissing is niet de beste. Na deze overwegingen blijft dus over de relatie (9.13c).

Voor $i=4$ gaat de relatie (9.8) over in

$$(9.14) \quad - q_4 \frac{f(\alpha_4)}{1-F(\alpha_4)} \sum_{j=1}^4 q_j \left\{ A_j + \int_0^{\alpha_j} t f(t) dt + \beta_j \int_{\alpha_j}^{\infty} f(t) dt \right\} +$$

$$+ q_4 f(\alpha_4) (\alpha_4 - \beta_4) + q_4 \frac{f(\alpha_4)}{1-F(\alpha_4)} \left\{ A_4 + \int_0^{\alpha_4} t f(t) dt + \beta_4 \int_{\alpha_4}^{\infty} f(t) dt \right\} = 0$$

ofwel

$$(9.15) \quad \frac{q_4 f(\alpha_4)}{1-F(\alpha_4)} \left\{ A_4 + \int_0^{\alpha_4} t f(t) dt + \beta_4 \int_{\alpha_4}^{\infty} f(t) dt - H(\alpha_1, \dots, \alpha_4) + \right.$$

$$\left. + (\alpha_4 - \beta_4) [1-F(\alpha_4)] \right\} = 0 .$$

Uit (9.15) volgt:

$$(9.16) \left\{ \begin{array}{ll} \text{a)} & \frac{f(\alpha_4)}{1-F(\alpha_4)} = 0 \\ \text{of} & \\ \text{b)} & q_4 = 0 \\ \text{of} & \\ \text{c)} & \{A_4 + \int_0^{\alpha_4} t f(t) dt + \beta_4 \int_{\alpha_4}^{\infty} f(t) dt - H(\alpha_1, \dots, \alpha_4) + \\ & + (\alpha_4 - \beta_4) [1-F(\alpha_4)]\} = 0 . \end{array} \right.$$

De relatie (9.16a) is in strijd met (9.1). Immers

$$1-F(\alpha_4) = \int_{\alpha_4}^{\infty} \frac{1}{1000} e^{-\frac{t}{1000}} dt = e^{-\frac{\alpha_4}{1000}} \quad \text{en dus}$$

$$\frac{f(\alpha_4)}{1-F(\alpha_4)} = \frac{1}{1000} \neq 0 \quad \text{voor iedere waarde van } \alpha_4.$$

De relatie (9.16b) is wegens (9.4) gelijkwaardig met de uitspraak:

"minstens één van de waarden α_i ($i=1,2,3$) is nul". Een dergelijke keuze kan niet optimaal zijn.

Na deze overwegingen blijft dus over de relatie (9.16c). Uit (9.13c) en (9.16c) volgt onmiddellijk:

$$(9.17) \left\{ \begin{array}{l} \alpha_i = \beta_i + \{H_i(\alpha_1, \dots, \alpha_4) - H(\alpha_1, \dots, \alpha_4)\} \frac{\sum_{h=1}^i q_h}{q_i F(\alpha_i)} \quad \text{voor } i \leq 3, \\ \alpha_4 = \beta_4 + \frac{H(\alpha_1, \dots, \alpha_4) - \{A_4 + \int_0^{\alpha_4} t f(t) dt + \beta_4 \int_{\alpha_4}^{\infty} f(t) dt\}}{1-F(\alpha_4)} \end{array} \right.$$

De relaties (9.17) bieden ons 4 vergelijkingen met 4 onbekenden, nl.

$\alpha_1, \dots, \alpha_4$. Deze vergelijkingen zijn helaas niet rechtstreeks oplosbaar.

Wij trachten daarom met behulp van een iteratie-procedure de optimale waarden van α_i te benaderen. Daartoe kiezen wij een willekeurig stel waarden $\alpha_i^{(0)}$ ($i=1, \dots, 4$) en substitueren deze waarden voor α_i in de rechterleden van (9.17). De op deze wijze verkregen numerieke waarden van de rechterleden van (9.17) noemen wij $\alpha_i^{(1)}$. Vervolgens substitueren wij deze waarden in de rechterleden van (9.17) en verkrijgen op analoge wijze de waarden $\alpha_i^{(2)}$. Zo voortgaande construeren wij vier getallenrijen $\{\alpha_i^{(n)}\}$ ($i=1, \dots, 4$). Indien ieder van deze getallenrijen convergeert tot een limiet α_i^* , dan voldoen de vier waarden α_i^* aan de relaties (9.17).

Het is niet gezegd dat deze iteratie-procedure voor iedere keuze van de beginschatting $\alpha_i^{(0)}$ tot een oplossing leidt. Wij hebben al opgemerkt dat α_1 minstens 1050 zal zijn, omdat de gederfde premieverlaging voor het volgende jaar anders de schadeuitkering zou overtreffen. Op dezelfde wijze beredeneert men dat

$$(9.18) \quad \alpha_i^* \geq A_1 - A_{i+1} + \beta_i \quad (\text{waarbij } A_5 \stackrel{\text{def}}{=} A_4)$$

voor $i=1, 2, 3, 4$ geldt; dus behalve $\alpha_1^* \geq 1050$ vindt men $\alpha_2^* \geq 1100$, $\alpha_3^* \geq 1100$, $\alpha_4^* \geq 1050$. Bij gebruik van deze ondergrenzen als beginschattingen vindt men de volgende tabel voor $\alpha_i^{(n)}$ en de bijbehorende verwachte kosten $H(\alpha_1^{(n)}, \alpha_2^{(n)}, \alpha_3^{(n)}, \alpha_4^{(n)})$.

Tabel 9.1

n	$\alpha_1^{(n)}$	$\alpha_2^{(n)}$	$\alpha_3^{(n)}$	$\alpha_4^{(n)}$	kosten
0	1050	1100	1100	1050	1381,8270
1	1308,799	1465,897	1476,538	1426,538	1363,1288
2	1319,661	1478,593	1490,145	1440,145	1363,1089
3	1319,667	1478,602	1490,155	1440,155	1363,1089
4	1319,667	1478,602	1490,155	1440,155	1363,1089

Strikt genomen moet men nu nog bewijzen dat voor de in de tabel bij $n = 4$ gevonden waarden de kostenfunctie inderdaad minimaal is. Wij zullen dit niet uitvoeren. Wel merken wij nog op dat uit (9.17) kan worden afgeleid $\alpha_4^* - \beta_4 = \alpha_3^* - \beta_3$, hetgeen in de tabel inderdaad vervuld blijkt te zijn ($\beta_3 = 500$, $\beta_4 = 450$).

Als afgeronde waarden voor de optimale claimgrenzen gebruiken wij

$$(9.19) \quad \alpha_1^* = 1320, \quad \alpha_2^* = 1479, \quad \alpha_3^* = 1490, \quad \alpha_4^* = 1440.$$

Wij zullen de transportonderneming dus adviseren om de totale schade van een jaar

in de eerste klasse te claimen als zij groter is dan f 1320,- ,
 in de tweede klasse te claimen als zij groter is dan f 1479,- ,
 in de derde klasse te claimen als zij groter is dan f 1490,- ,
 in de vierde klasse te claimen als zij groter is dan f 1440,- .

10. Andere Markovprocessen.

Vanaf paragraaf 3 hebben wij ons beperkt tot Markovketens met een eindig aantal toestanden. Wanneer we een aftelbaar oneindig aantal toestanden toelaten, krijgt de matrix van overgangswaarden een aftelbaar oneindig aantal rijen en kolommen. Een groot deel van de behandelde stellingen en begrippen blijft behouden. Maar een keten, en algemener een fuik, kan nu uit louter doorgangstoestanden bestaan; als een systeem met $p_{i,i+1} = 1$ voor elke i (dus $p_{ij} = 0$ voor $j \neq i+1$) begint in toestand 1, doorloopt het met kans 1 de toestandenreeks $1, 2, 3, \dots$. Van een invariante verdeling is in dit geval geen sprake. De eigenschap dat iedere Markovketen een invariante verdeling bezit geldt uitsluitend als het aantal toestanden eindig is. In paragraaf 15 zullen wij kennismaken met een Markovketen met aftelbaar oneindig veel toestanden, waarbij de invariante verdeling voor sommige waarden van de gegeven parameters bestaat en voor andere niet.

Het is ook mogelijk dat een systeem een continuüm van toestanden, bijv. alle reële getalwaarden, kan aannemen. In een dergelijk geval zal de kans op een overgang van toestand x naar toestand y in het algemeen nul zijn, net zoals $P[\underline{y}=y]$ nul is voor een stochastische grootheid met een continue verdeling (deel 2, paragraaf 7). Men definieert daarom liever voor elke toestand x en elke verzameling A van toestanden de functie $p(x,A)$ als de voorwaardelijke kans dat het systeem zich na de eerstkomende toestandswijziging in een tot A behorende toestand zal bevinden, gegeven dat het nu in toestand x verkeert. Voor Markovketens met een continuüm van toestanden kan een invariante verdeling bestaan, en men vindt deze door de redenering van paragraaf 5 tot dit geval uit te breiden.

Wanneer men bij een Markovketen de eis laat vallen dat de toestandsveranderingen slechts op discrete tijdstippen optreden, spreken wij van een Markovproces (zie echter voetnoot **) op blz. 4). Laat de toestandsverzameling aftelbaar zijn en laat $p_{ij}(t)$ de kans voorstellen dat een systeem dat begint in toestand i na een tijdsduur t in toestand j is geraakt. Dan geldt als analogon van (3.1), voor elke $s > 0$ en $t > 0$ en elk toestandenpaar i, j

$$(10.1) \quad p_{ij}(s+t) = \sum_h p_{ih}(s) p_{hj}(t) .$$

Een nauwkeurige formulering van de Markov-eigenschap (overgangswahrscheinlichkeit alleen afhankelijk van de uitgangstoestand) is hier wat lastiger. Ook voor Markovprocessen kan men het bestaan van doorgangstoestanden, fuiken, invariante verdelingen e.d., nagaan.

Voor de in deze paragraaf genoemde ketens en processen en ook voor stochastische processen die niet de Markov-eigenschap hebben, geldt dat de bestudering meer mathematische moeilijkheden oplevert dan bij de Markovketens met eindig veel toestanden het geval was. De meeste boeken (zie de literatuurlijst aan het eind van dit deel) eisen dan ook een grondige kennis van de maattheorie.

11. Wachttijdsituaties.

Iedereen kent de situatie waarin klanten aankomen bij een loket, waar zij door een bediende worden geholpen. Minder bekend zijn de situaties in fabrieken, waar een aantal gelijksoortige machines in bedrijf zijn die, indien zij door storing mochten uitvallen, worden hersteld door één van de aan het bedrijf verbonden reparateurs, zodra deze daarvoor beschikbaar komt.

De hierboven beschreven situaties zijn overeenkomstig in die zin, dat er in beide gewacht en bediend wordt. Het behoeft geen betoog dat wachten behalve vervelend ook kostbaar kan zijn. De wachttijd voor een loket of de tijdsduur welke voorafgaat aan een reparatie hangt uiteraard sterk af van het aantal ingeschakelde bedienden. Door het openstellen van een extra loket of door het toevoegen van een extra reparateur kan men de wachttijd vaak bekorten. Deze maatregelen brengen echter kosten met zich mee en de vraag is of deze offers opwegen tegen de vermindering van het aantal klanten dat ongeduldig wordt en de rij verlaat, resp. tegen de eventuele toename van de produktie. Ook kan men overwegen of het niet voordeliger is om kleine reparaties te laten voorgaan.

Dergelijke typen van keuzeproblemen kunnen met behulp van de theorie van de wachttijden wiskundig worden behandeld, hoewel de theorie zelf zich niet bezighoudt met optimaliteitsproblemen.

De te behandelen problemen geven aanleiding tot het opstellen van een wiskundig model, waarin o.a. de volgende aspecten een rol spelen:

- a) het aankomstenpatroon, bijv. de kansverdeling van de tijdsintervallen tussen de opvolgende aankomsten of van het aantal aankomsten in een tijdsinterval van gegeven lengte. Deze verdelingen kunnen nog van het tijdstip afhangen; men denke bijv. aan spitsuren.
- b) het bedieningsmechanisme (het aantal loketten, de kansverdeling van de bedieningstijd e.d.).
- c) de rijdiscipline. Tenzij anders vermeld wordt nemen wij aan dat de klanten in volgorde van aankomst worden geholpen. Maar men kan ook bepaalde klanten voorrang geven, bijv. omdat zij tevoren een afspraak hebben gemaakt of omdat zij een kortere bedieningstijd hebben.

Aan de hand van bepaalde onderstellingen over deze factoren tracht men informatie in te winnen over bijv. de verdelingsfunctie van de rijlengte, de verdelingsfunctie van de wachttijd, de vrije tijd van de bediende; kortom: grootheden die een maat zijn voor de efficiency van het systeem.

In het hierna volgende zullen wij spreken over loketten en klanten, over machines en reparateurs, maar men bedenke dat de resultaten ook geldig zijn in vele andere toepassingsgebieden. Zowel vliegvelden als telefoonlijnen kunnen dikwijls beschouwd worden als een loket waar de klant (vliegtuig of abonnee) op bediening wacht.

12. Het wachttijdproces.

Wanneer men als toeschouwer het komen en gaan van klanten bij een loket gadeslaat, dan ontdekt men dat

- a) niet van tevoren met zekerheid is vast te stellen op welke tijdstippen de rijlengte zich zal wijzigen;
- b) niet van tevoren met zekerheid is vast te stellen of op de in a) genoemde tijdstippen de rijlengte zal toenemen dan wel afnemen.

Wij beschouwen nu de rij van wachtenden voor het loket, inclusief degene die bediend wordt, als een systeem (zie paragraaf 1). De toestand van het systeem wordt bepaald door het aantal aanwezige klanten n : als $n = 0$ heeft de loketbediende niets te doen; als $n = 1$ is één klant aanwezig, die geholpen wordt; voor $n = k+1$ zijn er bovendien k klanten die wachten. De toestandsveranderingen van het systeem (aankomst en vertrek van klanten) voltrekken zich niet deterministisch; het systeem is onderworpen aan een stochastisch proces. Dit proces van wachtenden is alleen een keten, indien wij het proces op discrete tijdstippen aanschouwen.

Beperken wij onze aandacht tot discrete, op gelijke afstanden gelegen (equidistante) tijdstippen, dan behoeft het proces geen enkelvoudige Markovketen te zijn. Immers bij een enkelvoudige Markovketen

hangen de overgangswahrscheinlichkeiten slechts af van de laatst-aangenomen toestand en niet van de daaraan voorafgaande toestanden. Er is echter geen garantie dat het hierboven beschouwde proces zich op deze wijze laat beschrijven. Als er bijv. in de beschouwde wachtsituatie spitsuren zijn, is er bij waarnemen om de vijf minuten aan het begin van het spitsuur een grote kans dat de rijlengte toeneemt. Aan het eind van het spitsuur zal de rijlengte meestal afnemen. De overgangswahrscheinlichkeiten vanuit een vaste toestand zijn hier dus niet constant in de tijd.

Indien het beschouwde proces geen enkelvoudige Markovketen is, en er ook niet toe herleid kan worden door het invoeren van meer complexe toestanden, dan kan het proces nog wel eens beschreven worden door een Markovproces (vgl. paragraaf 10) met een invariante verdeling. In de wachttijdtheorie werkt men graag met de invariante verdeling, omdat deze niet afhangt van de begintoestand (er zijn nl. vrijwel nooit disjuncte fuiken). Men zegt dan dat het proces wordt beschouwd in de stationaire toestand.

Het is dikwijls ook zinvol om het proces niet te beschouwen op equidistante tijdstippen, maar juist op die tijdstippen waarop een klant binnenkomt of weggaat. Voor het verkrijgen van bepaalde informatie is het niet altijd nodig om het wachttijdproces op de voet te volgen, zoals in paragraaf 15 zal blijken.

Men onderscheidt in de wachttijdtheorie machineproblemen, waarbij het aantal potentiële klanten begrensd is, en loketproblemen, waarbij dit (in het model althans) niet het geval is (de toestandsvariabele n kan dan aftelbaar oneindig veel waarden aannemen). Wij zullen van beide groepen één probleem uitgebreid bespreken. In paragraaf 15 wordt een loketprobleem voor één loket behandeld, terwijl in paragraaf 16 een machineprobleem voor meer "loketten" aan de orde is.

13. Het Poissonproces.

In deze paragraaf zullen wij eerst aantonen dat onder betrekkelijk eenvoudige onderstellingen het aantal klanten dat in een tijdsinterval van een willekeurige maar vaste lengte aankomt, een Poissonverdeling heeft.

Laat $\underline{x}(t)$ het aantal klanten zijn dat in het interval $[0, t)$ aankomt, dan luiden de bovengenoemde onderstellingen

- a) in ieder eindig tijdsinterval komt slechts een eindig aantal klanten aan;
- b) op ieder tijdstip komt hoogstens één klant aan;
- c) als $0 \leq t_1 < \dots < t_n$ dan zijn de aantallen klanten die resp. in de perioden $t_1 \leq t < t_2, \dots, t_{n-1} \leq t < t_n$ aankomen en welke gegeven worden door $\underline{x}(t_2) - \underline{x}(t_1), \dots, \underline{x}(t_n) - \underline{x}(t_{n-1})$, onderling onafhankelijke stochastische grootheden (vgl. deel 2, def. 9.1);
- d) de verdelingsfunctie van het aantal klanten, dat in $[s, s+t)$ aankomt, $\underline{x}(s+t) - \underline{x}(s)$, hangt alleen af van de duur t en niet van het beginpunt s .

Indien aan deze vier onderstellingen is voldaan, noemt men de collectie stochastische grootheden $\underline{x}(t)$ ($0 < t < \infty$) een Poissonproces.

Hulpstelling 13.1

Als $Q(t)$ de kans is dat in het interval $[0, t)$ géén klanten aankomen, m.a.w.

$$(13.1) \quad Q(t) \stackrel{\text{def}}{=} P[\underline{x}(t) = 0],$$

dan geldt voor alle $t > 0$ en $s > 0$

$$(13.2) \quad Q(s+t) = Q(s) \cdot Q(t)$$

en

$$(13.3) \quad Q(s+t) \leq Q(t).$$

Voor elke $t > 0$ en elk natuurlijk getal n is bovendien

$$(13.4) \quad Q(t) = \left\{ Q\left(\frac{t}{n}\right) \right\}^n.$$

Bewijs:

$$Q(s+t) \stackrel{\text{def}}{=} P[\underline{x}(s+t)=0] = P[\underline{x}(s+t)-\underline{x}(s)=0 \wedge \underline{x}(s)=0].$$

Wegens onderstelling c) zijn de gebeurtenissen in het rechterlid stochastisch onafhankelijk, terwijl de kans op de eerste gebeurtenis $Q(t)$ is, volgens d). Dus is

$$Q(s+t) = P[\underline{x}(s+t)-\underline{x}(s)=0] \cdot P[\underline{x}(s)=0] = Q(t) \cdot Q(s).$$

Wegens $0 \leq Q(s) \leq 1$ volgt (13.3) uit (13.2). Ook volgt (13.3) direkt uit de definitie van $Q(t)$: als de periode langer wordt kan de kans op geen aankomst in die periode nooit toenemen.

Voor het bewijs van (13.4) kunnen we het interval $[0, t)$ als volgt in n deelintervallen verdelen:

$$(13.5) \quad \left[0, \frac{t}{n}\right); \left[\frac{t}{n}, \frac{2t}{n}\right); \dots; \left[\frac{(n-1)t}{n}, t\right).$$

De gebeurtenis {geen klant in $[0, t)$ } is nu de conjunctie van de onafhankelijke gebeurtenissen {geen klant in $\left[\frac{(j-1)t}{n}, \frac{jt}{n}\right)}$ voor $j=1, 2, \dots, n$. Elk van deze laatste gebeurtenissen heeft een kans $Q(\frac{t}{n})$; dus de kans $Q(t)$ op de eerstgenoemde gebeurtenis is het produkt van n factoren $Q(\frac{t}{n})$. Men kan (13.4) ook uit (13.2) bewijzen met volledige inductie.

Hulpstelling 13.2

De functie $Q(t) = P[\underline{x}(t)=0]$ is positief voor iedere positieve t .

Bewijs:

Stel dat voor zekere $t > 0$ geldt $Q(t) = 0$. Wegens (13.4) is dan ook $Q(\frac{t}{n}) = 0$. In elk der n deelintervallen (13.5) komt nu met kans nul géén klant, dus met kans één minstens één klant aan. Met kans één arriveren dus minstens n klanten in het eindige tijdsinterval $[0, t)$. Omdat dit geldt voor willekeurig grote n , zou het aantal aankomsten in $[0, t)$ oneindig groot zijn, zodat een tegenspraak ontstaat met onderstelling a). De onderstelling $Q(t) = 0$ kan dus voor geen enkele $t > 0$ juist zijn.

De eisen die aan $Q(t)$ blijken beide hulpstellingen moeten worden opgelegd, blijken deze functie volledig te bepalen

Stelling 13.1

Voor de kans $Q(t)$ op geen aankomst in $[0, t)$ geldt onder de onderstellingen a) $t/m d)$

$$Q(t) = e^{-\lambda t}.$$

Het is geen beperking te onderstellen dat de constante λ positief is.

Bewijs:

Formule (13.4) met $t=1$ geeft: $Q(1) = \{Q(\frac{1}{n})\}^n$. Stel $Q(1) = \alpha$, en $\log \alpha = -\lambda$, dan is

$$(13.6) \quad Q\left(\frac{1}{n}\right) = \alpha^{\frac{1}{n}} = e^{\frac{1}{n} \log \alpha} = e^{-\lambda \frac{1}{n}}.$$

Maar $t = \frac{n}{m}$ in (13.4) geeft

$$(13.7) \quad Q\left(\frac{n}{m}\right) = \{Q\left(\frac{1}{m}\right)\}^n = \left(e^{-\lambda \frac{1}{m}}\right)^n = e^{-\lambda \frac{n}{m}}.$$

De relatie $Q(r) = e^{-\lambda r}$ geldt dus voor elk positief rationaal getal r .

Om deze relatie voor ieder positief reëel getal ω te bewijzen gebruiken wij (13.3), d.i. de monotonie van $Q(t)$. Laten $r_1, r_2, \dots$ resp. $s_1, s_2, \dots$ rijen van rationale getallen zijn die monotoon stijgend resp. dalend naar ω convergeren. Dan geldt

$$(13.8) \quad Q(r_i) \geq Q(\omega) \geq Q(s_i)$$

of

$$(13.9) \quad \alpha^{s_i} \leq Q(\omega) \leq \alpha^{r_i}.$$

Omdat α^x een continue functie is van x , naderen zowel α^{s_i} als α^{r_i} tot α^ω als $i \rightarrow \infty$, zodat geldt

$$(13.10) \quad Q(\omega) = \alpha^\omega = e^{\omega \log \alpha} = e^{-\lambda \omega}.$$

Zou de constante λ negatief zijn dan was $Q(1) = e^{-\lambda} > 1$, maar de kans $Q(1)$ is ≤ 1 . Verder houdt $\lambda=0$ in dat $Q(t) = 1$ is voor elke eindige waarde van t . Met andere woorden: de kans op een aankomst van een klant is nul in ieder eindig tijdsinterval. Het is redelijk om aan te nemen, dat in dergelijke situaties geen wachttijdtheorie wordt toegepast.

Wij verdelen het interval $[0, t)$ weer in n gelijke deelintervallen volgens (13.5). Verder definiëren wij:

- 1) de gebeurtenis A vindt plaats als in het interval $[0, t)$ precies m klanten aankomen;
- 2) de gebeurtenis B_n vindt plaats als in m van de n deelintervallen minstens één klant aankomt en in de overige $n-m$ intervallen geen;
- 3) de gebeurtenis C_n vindt plaats als in minstens één van de n deelintervallen twee of meer klanten aankomen.

Wij kunnen nu het volgende vaststellen: als de gebeurtenis A heeft plaatsgevonden, dan is zeker ook minstens één van de gebeurtenissen B_n en C_n geschied; m.a.w. voor de kansen geldt (vgl. deel 2, (4.9) en (4.1)IV)

$$(13.11) \quad P[A] \leq P[B_n \vee C_n] \leq P[B_n] + P[C_n] .$$

Als de gebeurtenis B_n heeft plaatsgevonden, dan is zeker ook minstens één van de gebeurtenissen A en C_n geschied. Dus

$$(13.12) \quad P[B_n] \leq P[A \vee C_n] \leq P[A] + P[C_n] .$$

Uit (13.11) en (13.12) volgt

$$(13.13) \quad |P[A] - P[B_n]| \leq P[C_n] .$$

Volgens een stelling uit de maattheorie mogen P en $\lim$ verwisseld worden, mits $\lim C_n$ bestaat; dit levert

$$(13.14) \quad \begin{aligned} \lim_{n \rightarrow \infty} P[C_n] &= P[\lim_{n \rightarrow \infty} C_n] = \\ &= P[\text{minstens één tijdstip met 2 aankomsten}] = 0. \end{aligned}$$

Dus geldt

$$(13.15) \quad P[A] = \lim_{n \rightarrow \infty} P[B_n] .$$

Men kan nu gemakkelijk nagaan dat

$$(13.16) \quad \begin{aligned} P[B_n] &= \frac{n!}{m!(n-m)!} \left(e^{-\frac{\lambda t}{n}} \right)^{n-m} \left(1 - e^{-\frac{\lambda t}{n}} \right)^m = \\ &= \frac{1}{m!} e^{-\lambda t} \left(e^{\frac{\lambda t}{n}} - 1 \right)^m \frac{n!}{(n-m)!} = \\ &= \frac{1}{m!} e^{-\lambda t} \left[\frac{\lambda t}{n} + \frac{\lambda^2 t^2}{2! n^2} + \dots \right]^m \frac{n!}{(n-m)!} . \end{aligned}$$

Dus

$$(13.17) \quad \begin{aligned} P[A] &= \lim_{n \rightarrow \infty} P[B_n] = \\ &= \frac{1}{m!} e^{-\lambda t} (\lambda t)^m \lim_{n \rightarrow \infty} \left[1 + \frac{\lambda t}{2! n} + \dots \right]^m \frac{n!}{n^m (n-m)!} = \\ &= \frac{e^{-\lambda t} (\lambda t)^m}{m!} . \end{aligned}$$

Wij hebben A gedefinieerd als de kans op precies m aankomsten in het interval $[0, t)$. Wegens onderstelling d) verandert deze kans niet als we $[0, t)$ vervangen door $[s, s+t)$. Het maakt ook geen verschil of we het interval vervangen door $(0, t)$, $[0, t]$ of $(0, t]$, want voor elk tijdstip u geldt

$$(13.18) \quad \begin{aligned} &P[\text{geen aankomst precies op tijdstip } u] = \\ &= P \left[\lim_{\epsilon \downarrow 0} \{\text{geen aankomst in } [u, u+\epsilon)\} \right] = \\ &= \lim_{\epsilon \downarrow 0} P[\text{geen aankomst in } [0, \epsilon)] = \lim_{\epsilon \downarrow 0} e^{-\lambda \epsilon} = 1 , \end{aligned}$$

waarbij (13.17) gebruikt is met de gebeurtenis A voor $t = \epsilon$ en $m = 0$.

Wij hebben nu de volgende stelling bewezen:

Stelling 13.2

Wanneer voldaan is aan de onderstellingen a) t/m d) heeft het aantal aankomsten in ieder tijdsinterval van lengte t een Poisson-verdeling met

$$(13.19) \quad P \left[\underline{x}(t) = m \right] = \frac{e^{-\lambda t} (\lambda t)^m}{m!}.$$

De positieve parameter λ stelt hier het verwachte aantal aankomsten per tijdseenheid voor. Immers de Poissonverdeling (13.19) heeft verwachting λt , zodat het verwachte aantal aankomsten in iedere tijdsduur t gelijk is aan λt .

Vervolgens zullen wij een stelling bewijzen die betrekking heeft op de kansverdeling van de lengte van het tijdsinterval tussen twee aankomsten.

Stelling 13.3

Als het aantal aankomsten in elk tijdsinterval van lengte t een Poissonverdeling heeft met verwachting λt , dan zijn de lengten van de tijdsintervallen tussen twee aankomsten stochastisch onafhankelijk en exponentieel verdeeld met verwachting $\frac{1}{\lambda}$.

Bewijs:

Laat $\underline{s}_1$ de lengte van het tijdsinterval voorstellen van tijdstip 0 tot de aankomst van de eerste klant, en $\underline{s}_i$ voor $i \geq 2$ de tijd welke verstrijkt tussen de aankomst van de $(i-1)^{\text{ste}}$ en de i^{de} klant. Nu geldt voor $i \geq 2$

$$(13.20) \quad P \left[\underline{s}_i > s \mid \sum_{j=1}^{i-1} \underline{s}_j = t \right] = P \left[\text{geen aankomst in } (t, s+t] \right] = e^{-\lambda s}.$$

Aangezien het rechterlid niet van t afhangt, is ook de onvoorwaardelijke kans $P \left[\underline{s}_i > s \right]$ gelijk aan $e^{-\lambda s}$. Dus $\underline{s}_i$ is exponentieel verdeeld (vgl. deel 2, (7.21)). Dit geldt ook voor $\underline{s}_1$: kies $t=0$ in (13.20) en laat de voorwaarde weg.

Voor willekeurige $u > 0$ en $v > 0$ kiest men nu δ met $0 < \delta < v$ en men definieert de gebeurtenis $A_u = \{u < \underline{s}_1 \leq u+\delta \wedge \underline{s}_2 > v\}$. Dan impliceert A_u de gebeurtenis $B_u = \{\text{geen aankomst in } [0, u] \text{ en één aankomst in } (u, u+\delta] \text{ en geen aankomst in } (u+\delta, u+v]\}$. Immers uit de ongelijkheid voor $\underline{s}_1$ volgt dat er geen aankomst in $[0, u]$ en minstens één in $(u, u+\delta]$ valt; wegens $\underline{s}_2 > v$ mag er geen tweede aankomst in dat interval, of zelfs vóór tijdstip $u+v$, plaats vinden. Men vergelijkte figuur 13.1, waar de gebieden zijn aangegeven waarin de tweede aankomst mag vallen.

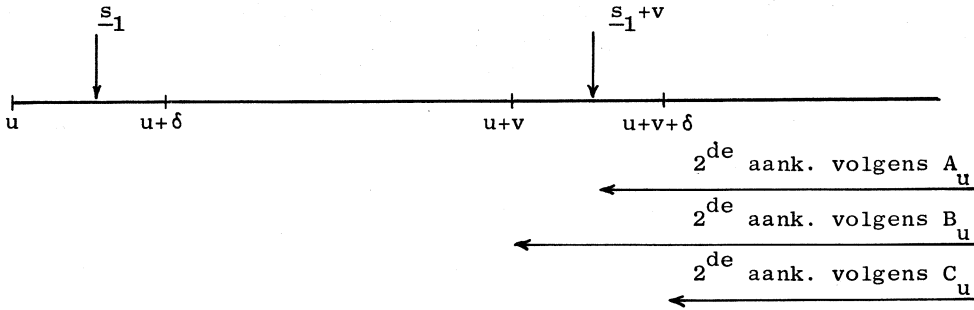


fig. 13.1

Uit $A_u \rightarrow B_u$ en de onafhankelijkheid van de drie beweringen in B_u volgt

$$(13.21) \quad P[A_u] \leq P[B_u] = e^{-\lambda u} \cdot \lambda \delta e^{-\lambda \delta} \cdot e^{-\lambda(v-\delta)}.$$

Evenzo gaat men na dat de gebeurtenis $C_u = \{\text{geen aankomst in } [0, u] \text{ en één aankomst in } (u, u+\delta] \text{ en geen aankomst in } (u+\delta, u+v+\delta]\}$ op zijn beurt A_u impliceert, zodat geldt

$$(13.22) \quad P[A_u] \geq P[C_u] = e^{-\lambda u} \cdot \lambda \delta e^{-\lambda \delta} \cdot e^{-\lambda v}.$$

Voor elke $s > 0$ kan nu de gebeurtenis $[\underline{s}_1 > s \wedge \underline{s}_2 > v]$ naar het tijdstip van de eerste aankomst gesplitst worden in de elkaar uitsluitende gebeurtenissen $A_{s+j\delta}$ ($j=0,1,\dots$), zodat men vindt

$$\begin{aligned}
 (13.23) \quad P \left[\underline{s}_1 > s \wedge \underline{s}_2 > v \right] &= \sum_{j=0}^{\infty} P[A_{s+j\delta}] \leq \\
 &\leq \sum_{j=0}^{\infty} e^{-\lambda(s+j\delta)} \cdot \lambda \delta e^{-\lambda v} = e^{-\lambda(s+v)} \frac{\lambda \delta}{1 - e^{-\lambda \delta}} = R_{\delta} ,
 \end{aligned}$$

waarbij (13.21) is gebruikt voor $u = s+j\delta$, en de somformule voor de meetkundige reeks. Evenzo vindt men met (13.22)

$$(13.24) \quad P \left[\underline{s}_1 > s \wedge \underline{s}_2 > v \right] \geq e^{-\lambda(s+v+\delta)} \frac{\lambda \delta}{1 - e^{-\lambda \delta}} = L_{\delta} .$$

Nu volgt uit

$$(13.25) \quad 1 - e^{-\lambda \delta} = 1 - 1 + \lambda \delta - \frac{(\lambda \delta)^2}{2!} + \dots$$

dat

$$(13.26) \quad \lim_{\delta \rightarrow 0} \frac{\lambda \delta}{1 - e^{-\lambda \delta}} = 1 .$$

Bovendien is $e^{-\lambda x}$ een continue functie van x , zodat L_{δ} en R_{δ} beide voor $\delta \rightarrow 0$ de limiet $e^{-\lambda(s+v)}$ hebben. Omdat $P \left[\underline{s}_1 > s \wedge \underline{s}_2 > v \right]$ niet van δ afhangt volgt dan uit (13.23) en (13.24)

$$(13.27) \quad P \left[\underline{s}_1 > s \wedge \underline{s}_2 > v \right] = e^{-\lambda(s+v)} .$$

Nu zijn $\underline{s}_1$ en $\underline{s}_2$ onafhankelijk, immers

$$\begin{aligned}
 (13.28) \quad P \left[\underline{s}_1 \leq s \wedge \underline{s}_2 \leq v \right] &= 1 - P \left[\underline{s}_1 > s \right] - P \left[\underline{s}_2 > v \right] + P \left[\underline{s}_1 > s \wedge \underline{s}_2 > v \right] = \\
 &= 1 - e^{-\lambda s} - e^{-\lambda v} + e^{-\lambda(s+v)} = (1 - e^{-\lambda s})(1 - e^{-\lambda v}) = P \left[\underline{s}_1 \leq s \right] \cdot P \left[\underline{s}_2 \leq v \right] .
 \end{aligned}$$

Met een eenvoudige generalisering van dit bewijs toont men aan dat elk k -tal $\underline{s}_{i_1}, \underline{s}_{i_2}, \dots, \underline{s}_{i_k}$ onafhankelijk is. Om de notatie eenvoudig te houden hebben wij ons hierboven tot $\underline{s}_1$ en $\underline{s}_2$ beperkt.

Uit stelling 13.3 volgt, dat de som van k verschillende aankomst-intervallen een gamma-verdeling met parameters k en λ heeft (deel 2, stelling 11.6).

Verder herinneren wij er aan dat voor elke exponentieel verdeelde grootheid $\underline{s}$ de voorwaardelijke verdeling van $\underline{s}$ -a onder de voorwaarde $\underline{s} > a$ wederom exponentieel is (deel 3, stelling 13.1). Deze merkwaardige eigenschap van de exponentiële verdeling hangt nauw samen met het model beschreven door de onderstellingen a) t/m d), waarbij de kans op het binnenkomen van een klant onafhankelijk van de voorgeschiedenis werd verondersteld.

Indien men bij een bushalte aankomt en men weet dat de vorige bus tien minuten geleden vertrokken is, dan zal men verwachten dat er nu wel eerder een bus zal komen dan indien men de vorige bus juist zou hebben gemist. Dit vermoeden kan blijkbaar alleen gegrond zijn, indien de tijdsintervallen tussen opeenvolgende aankomsten van bussen bij de halte niet exponentieel verdeeld zijn. Zijn deze tijdsintervallen wel exponentieel verdeeld, dan is de wachttijd tot de volgende bus arriveert niet afhankelijk van de tijd die verstreken is sinds het vertrek van de vorige bus.

In ieder geval zal men bij praktische toepassingen dienen na te gaan of de situatie zodanig is dat aan de genoemde eigenschap is voldaan. Dit is bijv. zeker niet het geval indien dienstregelingen bestaan.

Wij zullen deze eigenschap in de volgende paragrafen gebruiken, wanneer de bedieningstijd van een klant bij een loket exponentieel verdeeld is. Komt er dan tijdens deze bedieningstijd een andere klant binnen, dan is de tijd die deze klant moet wachten tot de klant die aan het loket geholpen wordt vertrekt, een kansgrootheid met dezelfde exponentiële verdeling.

Het omgekeerde van de stellingen 13.2 en 13.3 is ook juist; wij zullen dit niet bewijzen. De drie beweringen:

- 1) onderstellingen a) t/m d);
- 2) het aantal aankomsten in een tijdsduur t heeft een Poissonverdeling met verwachting λt ;
- 3) de aankomstintervallen zijn onderling onafhankelijk en exponentieel verdeeld met verwachting $\frac{1}{\lambda}$,

zijn onderling equivalent. Men vindt in de literatuur één van de drie, niet steeds dezelfde, als definitie en de andere twee als eigenschappen. De hier gegeven afleiding is van DOOB.

14. Laplace-Stieltjes-transformaties.

In de theorie der wachttijden wordt vrijwel steeds gewerkt met niet-negatieve stochastische variabelen (tijdsduren, aantallen klanten, enz.). Zo'n niet-negatieve stochastische variabele $\underline{x}$ kan behalve door zijn verdelingsfunctie $P[\underline{x} \leq x]$ ook worden vastgelegd door diens Laplace-Stieltjes-getransformeerde

$$(14.1) \quad Q(\tau) \stackrel{\text{def}}{=} \mathfrak{L} e^{-\tau \underline{x}} = \begin{cases} \int_0^{\infty} e^{-\tau x} f(x) dx \\ \sum_{k=0}^{\infty} p_k e^{-\tau x_k} \end{cases}$$

als $\underline{x}$ continu verdeeld is met dichtheid $f(x)$ resp. discreet verdeeld met $p_k = P[\underline{x} = x_k]$. Het is direkt duidelijk dat $Q(0) = 1$ is; men kan aantonen dat de integraal resp. de som voor alle positieve waarden van τ convergeert.

Men kan eenvoudig nagaan dat de n^{de} afgeleide van $Q(\tau)$, als deze bestaat, gegeven wordt door

$$(14.2) \quad Q^{(n)}(\tau) = (-1)^n \int_0^{\infty} x^n e^{-\tau x} f(x) dx ,$$

resp.

$$(14.3) \quad Q^{(n)}(\tau) = (-1)^n \sum_{k=0}^{\infty} x_k^n e^{-\tau x_k} p_k .$$

Bij substitutie $\tau = 0$ blijkt in beide gevallen

$$(14.4) \quad Q^{(n)}(0) = (-1)^n \mathfrak{L} \underline{x}^n .$$

Men kan bewijzen, dat een verdelingsfunctie door zijn Laplace-Stieltjes-getransformeerde éénéénduidig is bepaald. Het rekenen met Laplace-Stieltjes-getransformeerden is dikwijls eenvoudiger dan met verdelingsfuncties; soms is bijv. (14.4) een snelle methode om momenten van de verdeling te bepalen. In sommige wachttijdproblemen is een bij de stationaire toestand behorende kansverdeling lastig te bepalen, terwijl zijn getransformeerde eenvoudig te berekenen is. Het bepalen van de verdelingsfunctie uit de Laplace-Stieltjes-getransformeerde is theoretisch steeds mogelijk, maar niet altijd praktisch uitvoerbaar.

15. Een loketprobleem.

Laten wij aannemen dat klanten uit een onbegrensd grote populatie onafhankelijk van elkaar aankomen bij één loket, en dat daarbij aan de voorwaarden a) t/m d) van paragraaf 13 is voldaan: het aantal aankomsten in ieder tijdsinterval met lengte t heeft een Poissonverdeling met verwachting λt , terwijl de aankomstintervallen exponentieel verdeeld zijn met verwachting $\frac{1}{\lambda}$. Bediening geschiedt in volgorde van aankomst, terwijl de bedieningstijden onderling onafhankelijke stochastische variabelen zijn met dezelfde kansdichtheid $g(s)$ die niet nader wordt gespecificeerd.

Stel dat wij het systeem van wachtende klanten (inclusief degene die bediend wordt) beschouwen op de tijdstippen, waarop een klant na zijn bediening vertrekt. De toestand van het systeem wordt nu gekenmerkt door de lengte van de rij van wachtenden, die de vertrekkende klant achter zich laat.

Wij zullen nu aantonen dat de rij van toestanden welke het systeem doorloopt, een Markovketen vormt met aftelbaar veel toestanden. Het is duidelijk dat er sprake is van een keten, immers het proces wordt slechts op discrete tijdstippen beschouwd. Veronderstel tijdelijk dat de bedieningstijd voor iedere klant constant is en gelijk aan $s > 0$, dan zijn de volgende ontwikkelingen mogelijk:

a) Indien het aantal wachtenden op een beschouwd discreet tijdstip gelijk is aan $i \geq 1$, dan neemt de toestandsvariabele op het volgende tijdstip de waarde $i-1$ aan, als gedurende de bedieningstijd van de eerste der i wachtenden geen nieuwe klant arriveert. De kans op een overgang van i naar $i-1$ wordt dus gegeven door

$$(15.1) \quad p_{i,i-1} = P[\text{geen klanten in } s] = e^{-\lambda s} \quad (i \geq 1) .$$

b) Indien het aantal wachtenden op een beschouwd discreet tijdstip gelijk is aan $i \geq 1$, dan neemt de toestandsvariabele op het volgende tijdstip de waarde $i+k$ aan, als gedurende de bedieningstijd van de eerste der i wachtenden $(k+1)$ nieuwe klanten arriveren. De kans op een overgang van i naar $i+k$ wordt dus gegeven door

$$(15.2) \quad p_{i,i+k} = P[(k+1) \text{ klanten in } s] = \frac{e^{-\lambda s} (\lambda s)^{k+1}}{(k+1)!} \quad (i \geq 1, k \geq 0) .$$

c) Indien het aantal wachtenden op een beschouwd discreet tijdstip gelijk is aan 0, dan neemt de toestandsvariabele op het volgende tijdstip de waarde k aan, als gedurende de bedieningstijd van de eerstvolgende klant, die echter nog komen moet, k klanten arriveren. De kans op een overgang van 0 naar k wordt dus gegeven door

$$(15.3) \quad p_{0k} = P[k \text{ klanten in } s] = \frac{e^{-\lambda s} (\lambda s)^k}{k!} \quad (k \geq 0) .$$

d) Andere overgangen zijn niet mogelijk; dus voor de niet genoemde index-combinaties geldt $p_{jk} = 0$.

Men ziet gemakkelijk in dat informatie betreffende méér voorafgaande toestanden dan alleen de laatste geen invloed heeft op de kansverdeling van de eerstvolgende toestand. De beschouwde keten is dus een Markovketen met aftelbaar oneindig veel toestanden.

De oneindige matrix van overgangswhn ziet er schematisch als volgt uit:

$$(15.4) \quad \begin{pmatrix} e^{-\lambda s} & e^{-\lambda s} \lambda s & e^{-\lambda s} \frac{(\lambda s)^2}{2!} & e^{-\lambda s} \frac{(\lambda s)^3}{3!} & \dots \\ e^{-\lambda s} & e^{-\lambda s} \lambda s & e^{-\lambda s} \frac{(\lambda s)^2}{2!} & e^{-\lambda s} \frac{(\lambda s)^3}{3!} & \dots \\ 0 & e^{-\lambda s} & e^{-\lambda s} \lambda s & e^{-\lambda s} \frac{(\lambda s)^2}{2!} & \dots \\ 0 & 0 & e^{-\lambda s} & e^{-\lambda s} \lambda s & \dots \\ \vdots & & & & \ddots \end{pmatrix}$$

Ook bij Markovketens met aftelbaar oneindig veel toestanden kunnen fuiken optreden. Men gaat gemakkelijk na, dat in deze Markovketen noch disjuncte fuiken noch kringfuiken aanwezig zijn. Als de invariante kansen q_j bestaan, moeten zij voldoen aan (5.9), d.w.z.

$$(15.5) \quad q_0 = q_0 p_{00} + q_1 p_{10} = (q_0 + q_1) e^{-\lambda s}$$

en voor $k \geq 1$

$$(15.6) \quad \begin{aligned} q_k &= q_0 p_{0k} + \sum_{i=1}^{k+1} q_i p_{ik} = \\ &= q_0 e^{-\lambda s} \frac{(\lambda s)^k}{k!} + \sum_{i=1}^{k+1} q_i e^{-\lambda s} \frac{(\lambda s)^{k-i+1}}{(k-i+1)!} \end{aligned}$$

Tot dusver hebben wij voor de eenvoud aangenomen dat de bedieningstijd voor iedere klant constant is en wel gelijk aan s . Als de bedieningstijden $\underline{s}$ stochastisch zijn dan zijn de overgangswhn gegeven in (15.1) t/m (15.4) "voorwaardelijke" overgangswhn. Gaat men uit van de veronderstelling dat de bedieningstijd een stochastische grootheid is met kansdichtheid $g(s)$ dan vinden wij voor de overgangswhn

$$(15.7) \quad \left\{ \begin{array}{ll} p_{i,i-1} = \int_0^{\infty} e^{-\lambda s} g(s) ds & (i \geq 1), \\ p_{i,i+k} = \int_0^{\infty} \frac{e^{-\lambda s} (\lambda s)^{k+1}}{(k+1)!} g(s) ds & (i \geq 1, k \geq 0), \\ p_{0,k} = \int_0^{\infty} \frac{e^{-\lambda s} (\lambda s)^k}{k!} g(s) ds & (k \geq 0). \end{array} \right.$$

De invariante kansen q_j voldoen nu aan

$$(15.8) \quad q_0 = (q_0 + q_1) \int_0^{\infty} e^{-\lambda s} g(s) ds,$$

en

$$(15.9) \quad q_k = q_0 \int_0^{\infty} e^{-\lambda s} \frac{(\lambda s)^k}{k!} g(s) ds + \\ + \sum_{i=1}^{k+1} q_i \int_0^{\infty} e^{-\lambda s} \frac{(\lambda s)^{k-i+1}}{(k-i+1)!} g(s) ds \quad \text{voor } k \geq 1.$$

Bepaling van q_0

Wij zullen nu eerst de kans q_0 bepalen; daartoe berekenen wij $\sum_{k=1}^{\infty} k q_k$. Uit (15.9) volgt

$$\begin{aligned} \sum_{k=1}^{\infty} k q_k &= q_0 \sum_{k=1}^{\infty} k \int_0^{\infty} \frac{e^{-\lambda s} (\lambda s)^k}{k!} g(s) ds + \\ &+ \sum_{k=1}^{\infty} k \sum_{i=1}^{k+1} q_i \int_0^{\infty} \frac{e^{-\lambda s} (\lambda s)^{k-i+1}}{(k-i+1)!} g(s) ds = \\ (15.10) \quad &= q_0 \lambda \int_0^{\infty} s g(s) ds + \sum_{i=1}^{\infty} q_i \sum_{k=i-1}^{\infty} k \int_0^{\infty} \frac{e^{-\lambda s} (\lambda s)^{k-i+1}}{(k-i+1)!} g(s) ds = \\ &= q_0 \lambda \mathcal{E}_{\underline{s}} + \sum_{i=1}^{\infty} q_i \int_0^{\infty} e^{-\lambda s} g(s) \left[\sum_{h=0}^{\infty} (h+i-1) \frac{(\lambda s)^h}{h!} \right] ds = \\ &= q_0 \lambda \mathcal{E}_{\underline{s}} + \sum_{i=1}^{\infty} i q_i - (1-q_0) + (1-q_0) \lambda \mathcal{E}_{\underline{s}}. \end{aligned}$$

Uit (15.10) volgt nu

$$(15.11) \quad q_0 = 1 - \lambda \underline{\mathfrak{L}}_s .$$

Nu kunnen de kansen q_k achtereenvolgens worden berekend. Immers uit

(15.9) volgt

$$(15.12) \quad q_{k+1} = \frac{q_k - q_0 \int_0^\infty \frac{e^{-\lambda s} (\lambda s)^k}{k!} g(s) ds - \sum_{i=1}^k q_i \int_0^\infty \frac{e^{-\lambda s} (\lambda s)^{k-i+1}}{(k-i+1)!} g(s) ds}{\int_0^\infty e^{-\lambda s} g(s) ds} \quad \text{voor } k \geq 1$$

en uit (15.8)

$$(15.13) \quad q_1 = \frac{q_0 \left(1 - \int_0^\infty e^{-\lambda s} g(s) ds \right)}{\int_0^\infty e^{-\lambda s} g(s) ds}$$

Men definieert in de wachttijdtheorie de bezettingsgraad ^{*)} ρ als

$$(15.14) \quad \rho \stackrel{\text{def}}{=} \frac{\text{verwachte bedieningstijd}}{\text{verwachte aankomstinterval}} = \frac{\underline{\mathfrak{L}}_s}{\frac{1}{\lambda}} = \lambda \underline{\mathfrak{L}}_s .$$

Volgens (15.11) is q_0 alleen een kans als $0 \leq \rho \leq 1$. Het is duidelijk dat ρ zelfs positief is, want λ is een positieve parameter en bij $\underline{\mathfrak{L}}_s = 0$ zou de bedieningstijd nul zijn met kans 1. De eis $\rho \leq 1$ is intuïtief duidelijk: als de bediening gemiddeld langer duurt dan het aankomstinterval hopen zich steeds meer klanten in de wachtkamer op en ontstaat er geen stabiele stationaire verdeling. In het geval $\rho = 1$ volgt uit (15.11) $q_0 = 0$, uit (15.13) $q_1 = 0$, uit (15.12) voor $k = 1, 2, \dots$ achtereenvolgens $q_2 = 0$, $q_3 = 0$, etc. Ook bij de bezettingsgraad 1 is er dus geen invariante kansverdeling. Voor het geval $0 < \rho < 1$ kan men aantonen dat de invariante kansen q_j steeds bestaan en gegeven worden door de formules (15.11) t/m (15.13).

*) Engels: traffic intensity.

Bepaling van de verwachting van de rijlengte in de stationaire toestand

Om de verwachting van de toestandsgrootte $\underline{n}$ te bepalen, berekent men de uitdrukking $\sum_{k=1}^{\infty} k(k-1)q_k$. Uit (15.9) volgt

$$\begin{aligned} \sum_{k=1}^{\infty} k(k-1)q_k &= q_0 \sum_{k=1}^{\infty} k(k-1) \int_0^{\infty} \frac{e^{-\lambda s} (\lambda s)^k}{k!} g(s) ds + \\ &\quad + \sum_{k=1}^{\infty} k(k-1) \sum_{i=1}^{k+1} q_i \int_0^{\infty} \frac{e^{-\lambda s} (\lambda s)^{k-i+1}}{(k-i+1)!} g(s) ds = \\ &= q_0 \lambda^2 \underline{\mathcal{E}}_{\underline{s}}^2 + \sum_{i=1}^{\infty} q_i \sum_{k=i-1}^{\infty} k(k-1) \int_0^{\infty} \frac{e^{-\lambda s} (\lambda s)^{k-i+1}}{(k-i+1)!} g(s) ds = \\ &= q_0 \lambda^2 \underline{\mathcal{E}}_{\underline{s}}^2 + \sum_{i=1}^{\infty} q_i \sum_{h=0}^{\infty} (h+i-1)(h+i-2) \int_0^{\infty} \frac{e^{-\lambda s} (\lambda s)^h}{h!} g(s) ds . \end{aligned}$$

Na de splitsing $(h+i-1)(h+i-2) = h(h-1) + 2hi - 2h + i(i-1) - 2i + 2$ behandelt men elke term zoals in (15.10), waardoor blijkt

$$\begin{aligned} \sum_{k=1}^{\infty} k(k-1)q_k &= \lambda^2 \underline{\mathcal{E}}_{\underline{s}}^2 + \sum_{i=1}^{\infty} i(i-1)q_i + 2(1-q_0)(1-\lambda \underline{\mathcal{E}}_{\underline{s}}) + \\ (15.15) \quad &\quad - 2 \sum_{i=1}^{\infty} i q_i (1-\lambda \underline{\mathcal{E}}_{\underline{s}}) . \end{aligned}$$

Uit (15.15) volgt

$$(15.16) \quad \underline{\mathcal{E}}_{\underline{n}} = \sum_{i=1}^{\infty} i q_i = 1 - q_0 + \frac{\lambda^2 \underline{\mathcal{E}}_{\underline{s}}^2}{2(1-\lambda \underline{\mathcal{E}}_{\underline{s}})} = \rho + \frac{\lambda^2 \underline{\mathcal{E}}_{\underline{s}}^2}{2(1-\rho)} .$$

De uitdrukking (15.16) geeft ons de verwachting van het aantal wachtenden in de stationaire toestand op het moment dat een klant zojuist bediend is. Het aantal wachtenden vermindert dus niet alleen als de gemiddelde bedieningstijd korter wordt, maar ook als de variatie in bedieningstijd vermindert bij gelijkblijvend gemiddelde. Voor een exponentieel verdeelde bedieningstijd met verwachting μ is $\underline{\mathcal{E}}_{\underline{s}}^2 = 2\mu^2$; dus blijkt in (15.16) met $\lambda\mu = \rho$

$$(15.17) \quad \underline{\mathcal{E}}_{\underline{n}} = \rho + \frac{2\rho^2}{2(1-\rho)} = \frac{\rho}{1-\rho} .$$

Is de bedieningstijd daarentegen constant gelijk aan μ dan is $\underline{\mathcal{E}}_s^2 = \mu^2$ en

$$(15.18) \quad \underline{\mathcal{E}}_n = \rho + \frac{\rho^2}{2(1-\rho)} = \frac{\rho}{1-\rho} \left(1 - \frac{1}{2}\rho\right).$$

Voor een geringe bezettingsgraad ($\rho \approx 0$) geldt $1 - \frac{1}{2}\rho \approx 1$ en de verwachte rijlengte blijft nagenoeg gelijk. Maar voor $\rho \approx 1$ (druk verkeer; Engels: heavy traffic) wordt de verwachte rijlengte bijna gehalveerd door de variatie in de bedieningstijd weg te nemen bij gelijkblijvende verwachting.

Bepaling van de verwachting van de wachttijd in de stationaire toestand

Wij definiëren dat de wachttijd van een klant begint op het moment van binnenkomst en eindigt op het moment waarop zijn bediening begint. COX en SMITH ^{*)} onderscheiden queueing-time (excl. bedieningstijd) en waiting-time (incl. bedieningstijd).

Omdat de bediening geschiedt in volgorde van aankomst, zijn de personen die de zojuist bediende klant achter zich laat, aangekomen tijdens de wacht- en bedieningstijd van deze klant. Bij een gegeven wacht- en bedieningstijd is het aantal aankomsten verdeeld volgens een Poissonverdeling met verwachting $\lambda(w+s)$ als de w de wachttijd voorstelt. Indien de beide tijden niet constant zijn maar kansverdelingen hebben, dan wordt de verwachting van het aantal aankomsten gegeven door $\underline{\mathcal{E}}_{\lambda(\underline{w}+\underline{s})}$. Uit (15.16) volgt

$$(15.19) \quad \rho + \frac{\lambda^2 \underline{\mathcal{E}}_s^2}{2(1-\lambda \underline{\mathcal{E}}_s)} = \lambda \underline{\mathcal{E}}_w + \lambda \underline{\mathcal{E}}_s$$

en dus geldt

$$(15.20) \quad \underline{\mathcal{E}}_w = \frac{\lambda \underline{\mathcal{E}}_s^2}{2(1-\rho)}.$$

Zoals boven ziet men in dat deze grootte gehalveerd wordt bij overgang van exponentiële bedieningstijd op constante bedieningstijd met dezelfde verwachting.

*) Zie de literatuurlijst aan het eind van dit deel.

Bepaling van de Laplace-Stieltjes-getransformeerde voor de rijlengte in de stationaire toestand

Volgens (14.1) is de Laplace-Stieltjes-getransformeerde van de stationaire verdeling van $\underline{n}$, die we $Q(\tau)$ zullen noemen, gelijk aan $\sum_{k=0}^{\infty} e^{-\tau k} q_k$. Na substitutie van (15.9) en enig rekenen vindt men

$$(15.21) \quad Q(\tau) = \frac{(1-\rho)(1-e^{-\tau})\psi(\lambda-\lambda e^{-\tau})}{\psi(\lambda-\lambda e^{-\tau}) - e^{-\tau}},$$

waarbij $\psi(\cdot)$ de Laplace-Stieltjes-getransformeerde is van de verdelingsfunctie van de bedieningstijden.

Bepaling van de Laplace-Stieltjes-getransformeerde voor de wachttijd in de stationaire toestand

De kans op r aankomsten in de bedienings- en wachttijd van een klant wordt gegeven door

$$(15.22) \quad \int_0^{\infty} \int_0^{\infty} \frac{e^{-\lambda(w+s)} \lambda^r (w+s)^r}{r!} g(s) ds h(w) dw \quad *).$$

De Laplace-Stieltjes-getransformeerde van deze kansverdeling wordt gegeven door

$$(15.23) \quad \sum_{r=0}^{\infty} \int_0^{\infty} \int_0^{\infty} \frac{e^{-\lambda(w+s)} \lambda^r (w+s)^r e^{-\tau r}}{r!} g(s) ds h(w) dw.$$

Indien wij (15.23) uitwerken, verkrijgen wij achtereenvolgens

$$\begin{aligned} & \int_0^{\infty} \int_0^{\infty} e^{-\lambda(w+s) + \lambda(w+s)e^{-\tau}} g(s) ds h(w) dw = \\ (15.24) \quad & = \int_0^{\infty} e^{-\lambda w(1-e^{-\tau})} h(w) dw \int_0^{\infty} e^{-\lambda s(1-e^{-\tau})} g(s) ds = \\ & = \sigma(\lambda-\lambda e^{-\tau}) \psi(\lambda-\lambda e^{-\tau}), \end{aligned}$$

*) $h(w)$ is de kansdichtheid van de wachttijd.

als $\sigma(\cdot)$ en $\psi(\cdot)$ de Laplace-Stieltjes-getransformeerden zijn van de kansverdelingen van de wachttijden resp. de bedieningstijden. Omdat het aantal aankomsten in de wacht- en bedieningstijd van een klant gelijk is aan het aantal wachtenden dat dezelfde klant na bediening achter zich laat, is het rechterlid van (15.21) gelijk aan dat van (15.24), zodat geldt

$$(15.25) \quad \sigma(\lambda - \lambda e^{-\tau}) = \frac{(1-\rho)(1-e^{-\tau})}{\psi(\lambda - \lambda e^{-\tau}) - e^{-\tau}}.$$

Stel nu $\lambda - \lambda e^{-\tau} = \tau'$, dan gaat (15.25) over in

$$(15.26) \quad \sigma(\tau') = \frac{(1-\rho)\tau'}{\lambda\psi(\tau') + \tau' - \lambda}.$$

Of men uit $\sigma(\tau')$ gemakkelijk de verdelingsfunctie van de wachttijd kan terugvinden hangt af van de vorm van $\psi(\tau')$, m.a.w. van de aard van de verdeling van de bedieningstijd. Als de bedieningstijd exponentieel verdeeld is met parameter μ , kan men bewijzen dat de wachttijd nul is met kans $1-\rho$ en met kans ρ exponentieel verdeeld is met parameter $\mu-\lambda$. Wie zich afvraagt waarom $\mu > \lambda$ is herleze de opmerkingen onder (15.14).

De verdeling van de werktijd en de vrije tijd van de bediende

Indien de bediende vrij is en er komt een klant binnen, dan begint een werktijd, die eindigt op het eerstvolgende tijdstip waarop een klant vertrekt en geen rij achterlaat. Het aantal klanten dat in een werktijd bediend wordt noemen wij de treinlengte. De eerste klant van een "trein" heeft dus een wachttijd nul. Nu kan geconstateerd worden, dat de lengte van de werktijd en de treinlengte niet veranderen indien de volgorde waarin de klanten worden bediend, gewijzigd wordt.

De werktijd begint met het binnenkomen van een klant op een tijdstip (zeg $t=0$) waarop de bediende vrij is. Komen er tijdens de bedieningstijd $\underline{s}$ van deze klant geen klanten aan, dan is de lengte van de werktijd $\underline{s}$ en de treinlengte 1. Komen er m klanten aan in de bedieningstijd $\underline{s}$ ($m \geq 1$), dan doen wij alsof deze klanten op de volgende wijze worden bediend. De m klanten worden allen afgezonderd. Aan het

eind van de tijd $\underline{s}$ is, afgezien van de m klanten, het loket leeg. Eén van de m klanten wordt dan toegelaten om bediend te worden. Als er in zijn bedieningstijd klanten binnenkomen, dan worden die meteen na hem bediend; komen in die bedieningstijd weer klanten binnen, dan worden ook die meteen daarna bediend, enz. Tot op een ogenblik het loket leeg komt (terwijl er dus nog steeds $m-1$ klanten in afzondering wachten). De tijd die verlopen is sinds het einde van de bedieningstijd $\underline{s}$ is een tijd $\underline{d}_1$. Deze tijd $\underline{d}_1$ verschilt alleen daarin van een werktijd, dat de bediende niet vrij was vóór het begin ervan. De verdelingsfunctie van de tijdsduur $\underline{d}_1$ is evenwel identiek aan die van een werktijd. Aan het einde van die tijd $\underline{d}_1$ begint een tijdsduur $\underline{d}_2$ door weer één van de $m-1$ afgezonderde klanten toe te laten voor bediening, terwijl alle klanten die in zijn bedieningstijd binnenkomen weer direkt na hem geholpen worden, enz. Weer heeft $\underline{d}_2$ de verdelingsfunctie van een werktijd. Dit procédé wordt voortgezet tot de laatste van de afgezonderde klanten bediend is en ook alle klanten die in zijn bedieningstijd binnenkwamen, enz. Zo komt tenslotte het loket vrij, en dit betekent het einde van de werktijd $\underline{d}$, die op tijdstip 0 begon. Hieruit volgt dus:

$$(15.27) \quad \begin{cases} \underline{d} = \underline{s} & \text{als } \underline{m} = 0, \\ \underline{d} = \underline{s} + \underline{d}_1 + \dots + \underline{d}_m & \text{als } \underline{m} = m > 0, \end{cases}$$

waarbij $\underline{m}$ het aantal klanten is dat in de bedieningstijd van de eerste klant aankomt. Bepaalt men de verwachting van $\underline{d}$ onder de voorwaarde $\underline{m} = m$, dan vindt men

$$(15.28) \quad \mathcal{E}(\underline{d} \mid \underline{m} = m) = \mathcal{E}_{\underline{s}} + m \mathcal{E}_{\underline{d}},$$

daar $\underline{d}$, $\underline{d}_1$, $\underline{d}_2$, ... alle dezelfde verdelingsfunctie bezitten. Wij vinden nu voor de onvoorwaardelijke verwachting

$$(15.29) \quad \mathcal{E}_{\underline{d}} = \mathcal{E}_{\underline{s}} + \mathcal{E}_{\underline{m}} \mathcal{E}_{\underline{d}}$$

of

$$(15.30) \quad \mathcal{E}_{\underline{d}} = \frac{\mathcal{E}_{\underline{s}}}{1 - \mathcal{E}_{\underline{m}}} = \frac{\mathcal{E}_{\underline{s}}}{1 - \rho}.$$

Voor de treinlengte $\underline{h}$ vindt men analoog aan (15.27)

$$(15.31) \quad \begin{cases} \underline{h} = 1 & \text{als } \underline{m} = 0, \\ \underline{h} = 1 + \underline{h}_1 + \dots + \underline{h}_m & \text{als } \underline{m} > 0, \end{cases}$$

waaruit men op dezelfde wijze als hiervoor voor $\mathcal{E}_{\underline{d}}$ is geschied, kan afleiden

$$(15.32) \quad \mathcal{E}_{\underline{h}} = \frac{1}{1-\rho}.$$

Na elke werktijd komt een tijdsinterval $\underline{v}$ waarin de bediende vrij is. De lengte van het tijdsinterval is de wachttijd tot de eerstvolgende klant binnenkomt. Dit is volgens de bekende eigenschap van de exponentiële verdeling weer een exponentieel verdeelde grootte, zodat

$$(15.33) \quad P[\underline{v} \leq v] = 1 - e^{-\lambda v} \quad (\text{als } v > 0)$$

en dus

$$(15.34) \quad \mathcal{E}_{\underline{v}} = \frac{1}{\lambda}.$$

De fractie van de tijd gedurende welke de bediende gemiddeld bezig is met het bedienen van klanten wordt nu gegeven door

$$(15.35) \quad \frac{\mathcal{E}_{\underline{d}}}{\mathcal{E}_{\underline{d}} + \mathcal{E}_{\underline{v}}} = \frac{\frac{\mathcal{E}_{\underline{s}}}{1-\rho}}{\frac{\mathcal{E}_{\underline{s}}}{1-\rho} + \frac{1}{\lambda}} = \lambda \mathcal{E}_{\underline{s}} = \rho.$$

16. Een machineprobleem.

In een fabriek staan n machines opgesteld (bijv. automatische draaibanken) die door m reparateurs worden bediend. Als we aannemen dat nooit twee of meer reparateurs tegelijk aan dezelfde machine werken, dan is het zinloos om meer dan n reparateurs in één werkploeg te hebben. Wel kan het voordeliger zijn om $m < n$ te kiezen. We zullen nu trachten de optimale grootte van m te bepalen.

Stel dat de kansverdeling van de tijd dat een machine die nu in bedrijf is, nog zonder storing in werking zal zijn, gegeven wordt door de exponentiële verdelingsfunctie

$$(16.1) \quad F(t) = 1 - e^{-\lambda t}.$$

Stel verder dat de kansverdeling van de nog resterende reparatieduur van een machine die nu in reparatie is, gegeven wordt door de eveneens exponentiële verdelingsfunctie

$$(16.2) \quad G(t) = 1 - e^{-\mu t}.$$

De kansverdelingen van de resterende loop- en reparatietijden hangen bij dit model dus niet af van de tijd dat de machine al loopt of al in reparatie is. Tenslotte nemen wij nog aan dat alle looptijden en reparatietijden onafhankelijke grootheden zijn.

Op ieder ogenblik verkeert ieder der machines in één van de volgende drie toestanden:

- a) de machine is in bedrijf,
- b) de machine wacht op reparatie,
- c) de machine wordt gerepareerd.

Als nu i het aantal lopende machines is, zal het aantal machines dat tot de categorieën b) of c) behoort, gelijk zijn aan $n-i$. We nemen uiteraard aan dat de machines die niet lopen onmiddellijk worden gerepareerd, voorzover er een reparateur beschikbaar is. Het aantal machines dat in reparatie is, is dus gelijk aan

$$(16.3) \quad j = j(i) = \min(m, n-i)$$

en het aantal machines dat op reparatie wacht, wordt gegeven door

$$(16.4) \quad n-i-j = n-i-\min(m, n-i) = \max(n-i-m, 0).$$

Wij mogen hieruit concluderen dat de toestand van het systeem "machinepark" gekarakteriseerd kan worden door een één-dimensionale toestandsvector, waarvan de component gelijk is aan i , het aantal lopende machines.

Het is duidelijk dat ook nu het systeem onderworpen is aan een

stochastisch proces. Indien wij dit proces slechts aanschouwen op de tijdstippen dat de toestand van het systeem verandert (d.w.z. òf een zojuist gerepareerde machine wordt in werking gesteld òf een machine raakt defect), dan zijn wij getuige van een Markovketen met eindig veel toestanden. We zullen van deze Markovketen de overgangswah p_{ij} van toestand i naar toestand j berekenen.

De kans dat een looptijd t groter is dan een waarde t wordt blijkens (16.1) gegeven door

$$(16.5) \quad P[\underline{t} > t] = e^{-\lambda t}.$$

Stel dat op een gegeven ogenblik i machines in werking zijn, dan is de kans dat alle i looptijden groter zijn dan t gelijk aan

$$(16.6) \quad (e^{-\lambda t})^i = e^{-i\lambda t}.$$

De kans op minstens één looptijd kleiner dan t wordt gegeven door $1 - e^{-i\lambda t}$. De verdelingsfunctie van de kleinste van i looptijden is dus

$$(16.7) \quad F_i(t) = 1 - e^{-i\lambda t}.$$

Stel dat op een gegeven ogenblik j machines in reparatie zijn, dan vindt men op analoge wijze dat de kleinste van deze j reparatietijden de volgende verdelingsfunctie heeft:

$$(16.8) \quad G_j(t) = 1 - e^{-j\mu t}.$$

De kansdichtheden behorende bij de verdelingsfuncties (16.7) en (16.8) worden resp. gegeven door

$$(16.9) \quad \begin{cases} f_i(t) = i\lambda e^{-i\lambda t}, \\ g_j(t) = j\mu e^{-j\mu t}. \end{cases}$$

Stel dat de kleinste van i looptijden gelijk is aan x , dan is de kans dat er eerder een machine gerepareerd wordt dan dat er één defect raakt gelijk aan

$$(16.10) \quad G_j(x) = 1 - e^{-j\mu x}.$$

De onvoorwaardelijke kans dat er eerder een machine wordt gerepareerd dan dat er één defect raakt, als er i machines in werking zijn, is dus

$$(16.11) \quad P[\text{kleinste van } j \text{ reparatietijden} < \text{kleinste van } i \text{ looptijden}] = \\ = \int_0^{\infty} (1 - e^{-j\mu x}) \lambda e^{-i\lambda x} dx = \frac{j\mu}{i\lambda + j\mu},$$

waarin j gegeven wordt door (16.3).

De kans die we hier hebben berekend is voor de Markovketen juist de kans op een overgang van toestand i naar toestand $i+1$, dus

$$(16.12) \quad p_{i,i+1} = \frac{j\mu}{i\lambda + j\mu} \quad (1 \leq i \leq n-1).$$

Omdat van i alleen overgang naar $i-1$ of $i+1$ mogelijk is, geldt

$$(16.13) \quad p_{i,i-1} = 1 - \frac{j\mu}{i\lambda + j\mu} = \frac{i\lambda}{i\lambda + j\mu} \quad (1 \leq i \leq n-1).$$

De matrix $\mathcal{M}$ der overgangswahns heeft dus de volgende gedaante:

$$(16.14) \quad \mathcal{M} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ p_{10} & 0 & p_{12} & \cdots & 0 \\ 0 & p_{21} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & p_{n-1,n-2} & 0 & p_{n-1,n} \\ 0 & \cdots & 0 & 0 & 1 & 0 \end{pmatrix}$$

waarbij de kansen p_{ij} gegeven worden door (16.12) en (16.13). Terwille van de overzichtelijkheid hebben we deze uitdrukkingen voor de p_{ij} niet ingevuld in $\mathcal{M}$.

De matrix $\mathcal{M}$ heeft geen disjuncte of kringfuiken en het aantal toestanden is eindig. De stationaire verdeling bestaat dus en wordt weer bepaald door de vergelijkingen (5.9), die hier de volgende eenvoudige vorm hebben:

$$(16.15) \quad \left\{ \begin{array}{lcl} q_0 & = & p_{10} \cdot q_1 \\ q_1 & = & 1 \cdot q_0 + p_{21} \cdot q_2 \\ q_2 & = & p_{12} \cdot q_1 + p_{32} \cdot q_3 \\ & \vdots & \\ & \vdots & \\ q_{n-2} & = & p_{n-3, n-2} \cdot q_{n-3} + p_{n-1, n-2} \cdot q_{n-1} \\ q_{n-1} & = & p_{n-2, n-1} \cdot q_{n-2} + 1 \cdot q_n \\ q_n & = & p_{n-1, n} \cdot q_{n-1} \end{array} \right.$$

We lossen uit dit stelsel $q_0, q_1, \dots, q_n$ op door eerst $q_1, \dots, q_n$ in q_0 uit te drukken. Voor q_1 is dit wel zeer eenvoudig; q_2 kan met behulp van de tweede vergelijking van het stelsel (16.15) geschreven worden als

$$(16.16) \quad q_2 = \frac{q_1 - q_0}{p_{21}},$$

en aangezien we reeds q_1 in q_0 hebben uitgedrukt, kan ook q_2 als "functie van q_0 " worden geschreven. Hiermee gaan we door totdat tenslotte q_n is uitgedrukt in q_0 . Vullen we dan voor de grootheden p_{ij} , die natuurlijk ook in deze uitdrukkingen voorkomen, de door (16.12) en (16.13) gegeven vormen in, dan blijkt het resultaat te zijn

$$(16.17) \quad q_i = \frac{(m\mu)^{i-1} (i\lambda + m\mu)}{i! \lambda^i} q_0 \quad \text{als } 0 \leq i \leq n-m,$$

$$(16.18) \quad q_i = \frac{m^{n-m} (m-1)! \mu^{i-1} \{i\lambda + (n-i)\mu\}}{(n-i)! i! \lambda^i} q_0 \quad \text{als } n-m+1 \leq i \leq n.$$

q_0 kunnen we nu berekenen door deze $n+1$ vergelijkingen op te tellen; links komt er dan $q_0 + q_1 + \dots + q_n = 1$ en rechts

$$(16.19) \quad q_0 \left\{ \sum_{i=0}^{n-m} \frac{(m\mu)^{i-1} (i\lambda + m\mu)}{i! \lambda^i} + \sum_{i=n-m+1}^n \frac{m^{n-m} (m-1)! \mu^{i-1} (i\lambda + n\mu - i\mu)}{(n-i)! i! \lambda^i} \right\}.$$

Blijkens het voorgaande kan men nu de door (16.17) en (16.18) gegeven stationaire kansen als volgt interpreteren:

q_i is de kans op het waarnemen (in de stationaire toestand) van een overgang van het systeem naar de toestand i , onder de voorwaarde dat een overgang wordt waargenomen.

Maar als we het systeem op willekeurige tijdstippen beschouwen zal de kansverdeling van de waargenomen toestanden een geheel andere zijn. Immers: een toestand waarin het systeem zelden overgaat heeft een lage waarde van q_i , maar indien zo'n "zeldzame" toestand gemiddeld bijv. 30x zo lang blijft bestaan als alle andere toestanden, zal deze "zeldzame" toestand juist vaak worden waargenomen wanneer we niet alleen naar tijdstippen kijken waarop een verandering plaatsvindt.

Om dit nu te corrigeren vermenigvuldigen we de kansen q_i ieder met de gemiddelde "levensduur" L_i van de bijbehorende toestand i . De produkten zijn in het algemeen geen kansen meer, want hun som is gelijk aan $\sum_{i=0}^n q_i L_i$ en dit is niet noodzakelijk gelijk aan 1. Delen we echter iedere $q_i L_i$ door deze som

$$(16.20) \quad p_i = \frac{q_i L_i}{\sum_{k=0}^n q_k L_k},$$

dan is p_i juist de kans dat het systeem op een willekeurig tijdstip in de toestand i zal worden aangetroffen *).

Om de kansen p_i te kennen, moeten we dus allereerst L_i berekenen. Hiertoe bepalen we de kansverdeling van de tijd t_i gedurende welke de toestand i blijft bestaan **).

$$(16.21) \quad \begin{aligned} P[t_i > t] &= P[i \text{ looptijden} > t; j \text{ reparatietijden} > t] = \\ &= e^{-i\lambda t} \cdot e^{-j\mu t}, \end{aligned}$$

dus

$$(16.22) \quad P[t_i \leq t] = 1 - e^{-(i\lambda + j\mu)t}.$$

*) Het exakte bewijs van deze bewering zullen we hier achterwege laten.

**) j is hier (nog steeds) gelijk aan $\min(m, n-i)$.

t_i is dus exponentieel verdeeld met verwachting

$$L_i = \mathcal{E}_{t_i} = \frac{1}{i\lambda + j\mu}.$$

Eliminatie van j levert ons de volgende formules voor L_i :

$$(16.23) \quad \begin{cases} L_i = \frac{1}{i\lambda + m\mu} & \text{voor } 0 \leq i \leq n-m, \\ L_i = \frac{1}{i\lambda + (n-i)\mu} & \text{voor } n-m+1 \leq i \leq n. \end{cases}$$

Met behulp van de formules (16.18) en (16.23) vinden we nu

$$(16.24) \quad \begin{cases} q_i L_i = \frac{(m\mu)^{i-1}}{i! \lambda^i} q_0 = \frac{m^i \mu^i}{i! \lambda^i} q_0 L_0 & \text{voor } 0 \leq i \leq n-m, \\ q_i L_i = \frac{m^{n-m} (m-1)! \mu^{i-1}}{(n-i)! i! \lambda^i} = \frac{m^{n-m} m! \mu^i}{(n-i)! i! \lambda^i} q_0 L_0 & \text{voor } n-m+1 \leq i \leq n, \end{cases}$$

zodat

$$(16.25) \quad \begin{cases} p_i = \frac{m^i \mu^i}{i! \lambda^i} p_0 & \text{voor } 0 \leq i \leq n-m, \\ p_i = \frac{m^{n-m} m! \mu^i}{(n-i)! i! \lambda^i} p_0 & \text{voor } n-m+1 \leq i \leq n, \end{cases}$$

waarin

$$(16.26) \quad p_0^{-1} = \sum_{i=0}^{n-m} \frac{m^i \mu^i}{i! \lambda^i} + \sum_{i=n-m+1}^n \frac{m^{n-m} m! \mu^i}{(n-i)! i! \lambda^i}.$$

De uitdrukking (16.25) geeft ons de kans dat op een willekeurig tijdstip i machines in werking zijn, mits er sinds het systeem in een vaste toestand begon veel toestandswijzigingen hebben plaats gehad.

De gemiddelde rijlengte R van de machines die niet in werking zijn, wordt gegeven door

$$(16.27) \quad R = \sum_{i=0}^n (n-i) p_i.$$

Toepassing

In een fabriek staan 20 machines die in 4 groepen van 5 verdeeld zijn. Voor iedere groep is er een reparateur, maar geen van de 4 reparateurs is in staat om machines te repareren die niet tot zijn groep behoren. Het gebeurt dus wel eens dat in de ene groep alle machines in werking zijn (zodat de bijbehorende reparateur niets te doen heeft), terwijl in een andere groep bijv. twee machines op reparatie wachten. De directeur van de fabriek vindt dit bijzonder onprettig en overweegt om de reparateurs te verzoeken de cursus A.R.R. *) te volgen. De reparateurs, die deze cursus zelf moeten bekostigen, zullen hiervoor natuurlijk een tegenprestatie verlangen en de directeur besluit dat deze zal bestaan uit een loonsverhoging. Zijn probleem is nu: hoever kan ik met deze verhoging gaan, zonder dat de totale kosten (lonen + verliezen doordat één of meer machines defect zijn) zullen stijgen?

Berekening

Wij nemen aan dat uit waarnemingen gebleken is dat het hier behandelde model de toestand redelijk goed beschrijft; d.w.z. de waargenomen tijdstippen van defect raken en weer in bedrijf stellen van de machines zijn verenigbaar met de hypothese van onafhankelijke, exponentieel verdeelde looptijden en reparatietijden met voor elke machine dezelfde verwachting $\frac{1}{\lambda}$ resp. $\frac{1}{\mu}$. Uit de waarnemingen volgen tevens schattingen voor λ en μ . Laat c_1 het huidige uurloon zijn van de reparateurs, c_2 het nieuwe uurloon en c_3 de verliezen per uur per machine die stilstaat. Laat verder $\frac{\lambda}{\mu}$ gelijk zijn aan 0,1.

In de huidige situatie is dan de verwachting van de kosten per uur per groep gelijk aan $c_1 + c_3 \cdot R(5;1)$, waarin $R(5;1)$ de gemiddelde rijlengte is bij 5 machines en 1 reparateur. De verwachte kosten per uur worden dus gegeven door

$$(16.28) \quad A = 4c_1 + 4c_3 \cdot R(5;1) .$$

*) All Round Reparateur.

Wanneer ieder van de reparateurs de machines van iedere groep kan repareren zijn de kosten per uur

$$(16.29) \quad B = 4c_2 + c_3 \cdot R(20;4) .$$

Het nieuwe uurloon c_2 moet dus voldoen aan

$$4c_2 + c_3 \cdot R(20;4) \leq 4c_1 + 4c_3 \cdot R(5;1)$$

ofwel

$$(16.30) \quad c_2 \leq c_1 + c_3 \left\{ R(5;1) - \frac{R(20;4)}{4} \right\} .$$

Voor de berekening van $R(5;1)$ en $R(20;4)$ kan men de definitie (16.27) toepassen, na de kansen p_i te hebben opgelost uit (16.25) en (16.26). Vooral voor $n=20$ vergt dit veel rekenwerk. Men kan zich deze moeite besparen door tabellen te raadplegen die bij gegeven n , m en $\frac{\lambda}{\mu}$ *) de relevante gegevens afdrukken. Zo vindt men bij PECK en HAZELWOOD een tabel van de kans D dat een machine die defect raakt, op reparatie moet wachten, met als ingangen: het aantal machines N (bij ons n), het aantal reparateurs M (bij ons m) en de "service faktor" $X = \frac{\lambda}{\lambda + \mu}$. De tabel geeft ook de "efficiency faktor" F die gedefinieerd is door

$$(16.31) \quad F = 1 - \frac{\text{gemiddeld aantal wachtende machines}}{\text{totaal aantal machines}} .$$

De tabel bevat de waarden $N = 4$ (1) 26 (2) 70 (5) 170 (10) 250 **) met bijpassende M en X -waarden. In de inleiding geven PECK en HAZELWOOD aan hoe uit F en de drie gegevens N , M en X de overige grootheden (bijv. gemiddeld aantal lopende machines) worden berekend. Zo is

$$(16.32) \quad N - R(N;M) = NF(1-X) .$$

In ons voorbeeld is $\frac{\lambda}{\mu} = 0,1$; dus $X = \frac{1}{11}$. Dit geeft in de tabel voor $N = 5$ en $M = 1$ de waarde $F = 0,960$ en voor $N = 20$ en $M = 4$ de waarde $F = 0,997$. Dus is

*) L.G. PECK & R.N. HAZELWOOD, Finite queueing tables, Wiley, New York (1958).

**) Met deze notatie bedoelt men: alle gehele waarden van 4 t/m 26, alle even waarden van 26 t/m 70, alle vijfvoudigen t/m 170, enz. De getallen tussen haakjes geven de stapgrootte aan, de getallen zonder haakjes de grenzen.

$$(16.33) \quad \begin{cases} R(5;1) = 5 - 5 \cdot 0,960 \cdot \frac{10}{11} = 5 - 4,36 = 0,64 , \\ R(20;4) = 20 - 20 \cdot 0,997 \cdot \frac{10}{11} = 20 - 18,13 = 1,87 . \end{cases}$$

Substitutie van deze waarden in (16.30) doet deze voorwaarde overgaan in

$$(16.34) \quad c_2 \leq c_1 + c_3 \cdot 0,17 .$$

De maximale loonsverhoging die de directeur onder de genoemde beperkingen kan geven is dus $c_3 \cdot 0,17$.

Berekening van de verwachte wachttijd w bij één reparator

Stel dat het systeem, vlak voordat een machine defect raakt, in de toestand i verkeert. Er zijn dus i machines in werking en van de overige $n-i$ is er één in reparatie. De verwachting van de tijd die zal verstrijken voordat deze reparatie klaar is, is op ieder tijdstip gelijk aan $\frac{1}{\mu}$. Hetzelfde geldt voor de verwachting van de reparatietijden van ieder der $n-i-1$ machines die op reparatie wachten.

De voorwaardelijke verwachting van de wachttijd van een machine die defect raakt op een tijdstip waarop het systeem van toestand i in toestand $i-1$ overgaat, bedraagt dus

$$(16.35) \quad \mathcal{E}(\underline{w} \mid i \rightarrow i-1) = \frac{n-i}{\mu} .$$

Om de onvoorwaardelijke verwachting van de wachttijd te berekenen, moeten we nu eerst de kans bepalen dat het systeem van i naar $i-1$ overgaat als we alleen die tijdstippen beschouwen waarop een machine defect raakt. Het gegeven dat een machine defect raakt (dus de toestands-grootte daalt) geven we met "dalend" aan. Dan is de gevraagde kans

$$(16.36) \quad P[i \rightarrow i-1 \mid \text{dalend}] = \frac{P[i \rightarrow i-1 \text{ en dalend}]}{P[\text{dalend}]} = 2q_i p_{i,i-1} .$$

Immers: het systeem kan alleen van i naar $i-1$ overgaan als het eerst in toestand i verkeert, d.w.z. als het bij de vorige toestandsverandering overging in i . De kans hierop is q_i , terwijl $p_{i,i-1}$ de kans is

dat het daarna in toestand $i-1$ overgaat. Dit alles onder de voorwaarde dat er een gebeurtenis plaatsvindt. Verder moet in de stationaire toestand even vaak een machine defect raken als weer in werking worden gesteld, omdat het systeem anders niet stabiel is. De kans $P[\text{dalend}]$ (waarbij niet bekend is vanuit welke $i!$) is dus $\frac{1}{2}$.

De onvoorwaardelijke verwachting van $\underline{w}$ is dus gelijk aan

$$(16.37) \quad \mathcal{E} \underline{w} = \sum_{i=1}^n 2q_i p_{i,i-1} \frac{n-i}{\mu}.$$

Om deze uitdrukking te herleiden, merken we allereerst op dat uit de formules (16.13), (16.17) en (16.18) volgt dat

$$(16.38) \quad q_i p_{i,i-1} = \left(\frac{\mu}{\lambda}\right)^{i-1} \frac{q_0}{(i-1)!} \quad \text{voor } i=1, \dots, n.$$

Verder geldt

$$(16.39) \quad 2 \sum_{i=1}^n q_i p_{i,i-1} = 1,$$

want op ieder tijdstip waarop een machine defect raakt, gaat het systeem van toestand i over in toestand $i-1$ voor één der waarden $i=1, \dots, i=n$.

Met behulp van de betrekkingen (16.38) en (16.39) kunnen we nu het rechterlid van (16.37) als volgt herleiden:

$$\begin{aligned} \sum_{i=1}^n 2q_i p_{i,i-1} \frac{n-i}{\mu} &= \frac{2}{\mu} \sum_{i=1}^n \{(n-1)-(i-1)\} q_i p_{i,i-1} = \\ &= \frac{2}{\mu} (n-1) \frac{1}{2} - \frac{2}{\mu} \sum_{i=2}^n (i-1) q_i p_{i,i-1} = \frac{n-1}{\mu} - \frac{2}{\mu} \sum_{i=2}^n (i-1) \left(\frac{\mu}{\lambda}\right)^{i-1} \frac{q_0}{(i-1)!} = \\ (16.40) \quad &= \frac{n-1}{\mu} - \frac{2}{\mu} \frac{\mu}{\lambda} \sum_{k=1}^{n-1} \left(\frac{\mu}{\lambda}\right)^{k-1} \frac{q_0}{(k-1)!} = \frac{n-1}{\mu} - \frac{2}{\lambda} \left\{ \frac{1}{2} - \left(\frac{\mu}{\lambda}\right)^{n-1} \frac{q_0}{(n-1)!} \right\} = \\ &= \frac{n-1}{\mu} - \frac{1}{\lambda} \{1 - 2q_n\}. \end{aligned}$$

Hier is $k = i-1$ als sommatie-index ingevoerd. De som is die van (16.39), maar de laatste term ontbreekt. Zo vindt men tenslotte

$$(16.41) \quad \xi_w = \frac{n-1}{\mu} - \frac{1}{\lambda} (1-2q_n) ,$$

waarbij q_n nog moet worden opgelost uit (16.17), (16.18) en $\sum_j q_j = 1$.

Slotwoord

Onder vrij stringente veronderstellingen omtrent aankomsten, bediening en rijdiscipline zijn in de voorafgaande paragrafen een aantal belangrijke begrippen en technieken van de wachttijdtheorie behandeld. Wij hopen hierdoor tevens diè literatuur ietwat toegankelijk gemaakt te hebben, waarin de wiskundige methoden en resultaten staan voor minder beperkende veronderstellingen. Naast de op de volgende pagina genoemde boeken van COX & SMITH, RIORDAN en SAATY zijn er nog vele andere. Men beschouwt o.a. meer loketten (in serie of parallel), andere aankomstpatronen, klanten die met een zekere kans uit ongeduld weglopen, voorrangsregelingen e.d.

In de praktijk van het wachten treden vaak situaties op waarin een analytisch-wiskundige aanpak (nog?) niet mogelijk is. Men bedient zich dan vaak van simulatie (vgl. deel 8) om het gedrag van het systeem te onderzoeken.

Literatuur

D.R. COX & W.L. SMITH, Queues, Methuen, London, 1961 (een zeer duidelijke inleiding, met uitvoerige literatuurlijst).

K.L. CHUNG, Markov chains with stationary transition probabilities, Springer, Berlin, 1960 (een boek met veel wiskundige theorie en weinig toepassingen).

J.L. DOOB, Stochastic processes, Wiley, New York, 1953 (een abstract algemeen leerboek voor gevorderden).

W. FELLER, An introduction to probability theory and its applications, deel 1, Wiley, New York, 3^{de} druk 1968 (een duidelijk en gedegen leerboek, waarin enige hoofdstukken aan Markovprocessen en Poissonprocessen zijn gewijd).

J.G. KEMENY & J.L. SNELL, Finite Markov chains, Van Nostrand, Princeton, 1960 (een uitvoerige inleiding).

J. RIORDAN, Stochastic service systems, Wiley, New York, 1962 (geeft in beknopte vorm veel resultaten over wachttijden).

T.L. SAATY, Elements of queueing theory, McGraw-Hill, New York, 1961 (overzicht met toepassingen en uitvoerige literatuurlijst).