STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

STATISTISCHE AFDELING

Rapport S 272

De statisticus achter het stuur

door

A.H. Haitsma

en

J.Th. Runnenburg

April 1960

Probleem:

Aan het begin van een weg vertrekken auto's. De tijdsintervallen tussen opeenvolgende vertrektijdstippen zijn onafhankelijke stochastische variabelen $\ldots, \underline{y}_{-n}, \ldots, \underline{y}_0, \ldots, \underline{y}_n, \ldots$ met

(1) $$P\{\underline{y}_i \leq y\} = 1 - e^{-\lambda y} \qquad \text{voor } y \geq 0.$$

Hierbij is λ een positieve constante en het vertrek-proces stationnair.

De rijtijd van een auto is de tijd, welke de auto nodig heeft om de hele weg af te leggen, rijdend met constante snelheid. De rijtijden der auto's zijn onderling onafhankelijke stochastische variabelen met dezelfde verdeling en verdelingsfunctie $F(t)$ (met $F(t)=0$ voor $t \leq 0$, de rijtijden zijn niet-negatief). De rijtijden en de vertrek-intervallen zijn eveneens onderling onafhankelijk. Een auto heeft verwaarloosbare lengte en houdt zich aan de uit zijn rijtijd voortvloeiende snelheid, totdat hij een eerder vertrokken wagen bereikt heeft. Deze blijft hij met de snelheid van zijn voorganger volgen tot het einde van de weg.

Gevraagd:

Wat is de verdeling van het aantal wagens in een stoet voor een willekeurige stoet aan het einde van de weg bij het stationnaire proces?

Oplossing:

Aan het einde van de weg komt telkens een stoet aan, die altijd bestaat uit een voorste wagen (een auto, die niet achter een langzamere "gevangen" is) en een stochastisch aantal volgwagens n, dat 0, 1, 2, ... kan zijn. Beschouw eerst

(2) $$Q_t(z) = \sum_{n=0}^{\infty} q_n(t) z^n,$$

waarin $|z| \leq 1$ is en $q_n(t)$ de kans voorstelt op precies n volgwagens, achter de voorste wagen aan het einde van de weg, onder de voorwaarde, dat de rijtijd van de voorste wagen t is.

Laat in het interval $(y, y+dy)$ na het vertrek op tijdstip 0 van een voorste wagen A met rijtijd t van het begin van de weg de volgende voorste wagen B vertrekken en laten er in het interval $(0,y)$ precies n wagens op de weg komen, die ten tijde t volgwagens bij A zijn. Er komen precies n auto's in $(0,y)$ op de weg met kans $e^{-\lambda y}\frac{(\lambda y)^n}{n!}$. Onder de voorwaarde "er komen in $(0,y)$ precies n auto's op de weg", kunnen de n vertrektijdstippen van die auto's beschouwd worden als evenzoveel onderling onafhankelijke trekkingen uit een homogene verdeling over het interval $(0,y)$. Laat $\underline{u}$ één zo'n trekking voorstellen. De op $\underline{u}$ vertrekkende auto (die rijtijd $\underline{t}$ heeft) is dan en slechts dan aan het einde van de weg volgwagen van A, als $\underline{u}+\underline{t}\le t$ is. De kans hierop is

$$(3) \qquad P\{\underline{u}+\underline{t}\le t\} = \int_0^y \frac{F(t-u)}{y}\,du = \frac{H(y)}{y},$$

als

$$(4) \qquad H(y) = \int_0^y F(t-u)\,du = \int_{t-y}^{t} F(x)\,dx$$

is.

De kans op "alle n wagens in $(0,y)$ vertrekkend zijn volgwagens ten tijde t" is dan $\left(\frac{H(y)}{y}\right)^n$.

Conclusie: met kans λdy vertrekt in $(y, y+dy)$ een auto B, die met kans $1-F(t-y)$ géén volgwagen wordt bij auto A. Met kans

$$(5) \qquad e^{-\lambda y}\frac{(\lambda y)^n}{n!}\left(\frac{H(y)}{y}\right)^n = e^{-\lambda y}\frac{\{\lambda H(y)\}^n}{n!}$$

vertrekken in $(0,y)$ precies n auto's, die volgwagens bij A zijn aan het einde van de weg. Dan is B dus inderdaad de eerstvolgende voorste wagen na A en we vinden, sommerend over de mogelijke waarden van y

$$(6) \qquad q_n(t) = \int_0^\infty e^{-\lambda y}\frac{\{\lambda H(y)\}^n}{n!}\{1-F(t-y)\}\lambda dy,$$

zodat .

(7) $$Q_t(z) = \sum_{n=0}^{\infty} \int_0^{\infty} \lambda e^{-\lambda y} \frac{\{\lambda z H(y)\}^n}{n!} \{1-F(t-y)\}\, dy =$$

$$= \int_0^{\infty} \lambda\{1-F(t-y)\} \exp\,[-\lambda\{y-zH(y)\}]\, dy =$$

$$= \int_0^{\infty} \lambda\{1-F(t-y)\} \exp\,[-\lambda \int_{t-y}^{t} \{1-zF(x)\}\, dx]\, dy.$$

Laat $dG(t)$ de kans zijn, dat de rijtijd van een voorste wagen tussen t en $t+dt$ ligt, d.w.z. laat $G(t)$ de verdelingsfunctie van die rijtijd zijn.

Ter toelichting beschouwen we eerst hetzelfde probleem bij een aantal discrete rijtijden $t_1 < \ldots < t_n$, optredend met kansen $p_1, \ldots, p_n$, waarbij $\sum_{i=1}^{n} p_i = 1$ is. Het aantal in een tijd x vertrekkende auto's met rijtijd t_i bezit een Poisson-verdeling met parameter $\lambda p_i x$. De kans, dat een op tijdstip 0 vertrekkende wagen met rijtijd t_k niet achter een andere wagen gevangen wordt, als die ander de rijtijd $t_i (> t_k)$ heeft, is dus

(8) $$e^{-\lambda p_i (t_i - t_k)} .$$

De kans, dat géén wagen de op tijdstip 0 vertrekkende auto "vangt", is dan

(9) $$\prod_{i=k+1}^{n} e^{-\lambda p_i (t_i - t_k)} = \exp\,[-\lambda \sum_{i=k+1}^{n} p_i (t_i - t_k)],$$

want een snellere wagen "vangt" nooit een tragere en de aantallen wagens met verschillende rijtijd zijn onderling onafhankelijk verdeeld.

De kans, dat als er op tijdstip 0 een auto vertrekt, deze een voorste wagen met rijtijd t_k zal zijn, is

(10) $$\exp\,[-\lambda \sum_{i=k+1}^{n} p_i (t_i - t_k)] \cdot p_k .$$

De kans, dat als er op tijdstip 0 een voorste wagen vertrekt, dit een voorste wagen met rijtijd t_k zal zijn, is

$$(11)\quad \Big\{\sum_{k=1}^{n} \exp\Big[-\lambda \sum_{i=k+1}^{n} p_i(t_i-t_k)\Big]\, p_k\Big\}^{-1} \exp\Big[-\lambda \sum_{i=k+1}^{n} p_i(t_i-t_k)\Big].p_k.$$

Algemener geldt: als op tijdstip 0 een voorste wagen met rijtijd t vertrekt, dan kan er in een interval met lengte x vóór 0 (d.w.z. het interval (-x,0) op de tijdas) géén auto met rijtijd in (x+t, x+t+dt) vertrokken zijn. De kans op "gedurende tijd x is géén auto met rijtijd in (x+t, x+t+dt) vertrokken" is $e^{-\lambda x dF(x+t)}$. De kans, dat als er bij 0 een voorste wagen vertrokken is, deze een rijtijd in het interval (t,t+dt) heeft, is dF(t). Maar dan is, sommerend over de mogelijke waarden van x (en met :: in de betekenis "evenredig")

$$(12)\qquad dG(t)\ ::\ \exp\Big[-\lambda\int_t^{\infty}\{1-F(x)\}\,dx\Big]\, dF(t),$$

zodat de verdelingsfunctie van de rijtijd van een voorste wagen gegeven wordt door

$$(13)\qquad G(t) = C\int_0^t \exp\Big[-\lambda\int_y^{\infty}\{1-F(x)\}\,dx\Big]\, dF(y),$$

waarin C volgt uit

$$(14)\qquad C^{-1} = \int_0^{\infty} \exp\Big[-\lambda\int_y^{\infty}\{1-F(x)\}\,dx\Big]\, dF(y)$$

en C^{-1} de kans voorstelt, dat een vertrekkende auto een voorste wagen is.

Tenslotte volgt nu voor de genererende functie Q(z) van q_n, de kans op een voorste wagen met n volgwagens (steeds is $|z| \leq 1$):

$$(15)\qquad Q(z) = \sum_{n=0}^{\infty} q_n z^n = \int_0^{\infty} Q_t(z)\, dG(t)\ .$$

Slechts in vrij triviale gevallen volgt voor Q(z) een <u>eenvoudige</u> formule, bijvoorbeeld voor <u>twee</u> mogelijke rijtijden a en b($>$ a), die met kans p respectievelijk q(p+q=1) optreden. De algemene oplossing kan voor numerieke benaderingen gebruikt worden. Men kan proberen, tenminste de <u>verwachting</u> van het aantal volgwagens $\underline{n}$ te bepalen. Uit (7) volgt:

$$1 - zQ_t(z) = (1-z)R_t(z), \tag{16}$$

waarbij

$$R_t(z) = \int_0^{\infty} \lambda \exp\left[-\lambda \int_{t-y}^{t} \{1-zF(x)\}dx\right] dy. \tag{17}$$

Differentiatie van (16) naar z en substitutie van z=1 levert

$$E\,\underline{n}_t = \sum_{n=1}^{\infty} nq_n(t) = \left(\frac{dQ_t(z)}{dz}\right)_{z=1} = R_t(1)-1. \tag{18}$$

Hieruit volgt, als $\tau = E\,\underline{t}$,

$$\begin{aligned} E\,\underline{n} &= \int_0^{\infty}\{R_t(1)-1\}\,dG(t) = \\ &= c\int_0^{\infty}\left(\int_0^{\infty}\lambda\exp\left[-\lambda\int_{t-y}^{t}\{1-F(x)\}dx\right]dy\right)\exp\left[-\lambda\int_t^{\infty}\{1-F(x)\}dx\right]dF(t)-1 = \\ &= c\int_0^{\infty}\int_0^{\infty}\lambda\exp\left[-\lambda\int_{t-y}^{\infty}\{1-F(x)\}dx\right]dy\,dF(t)-1 = c\int_0^{\infty}\left\{\int_0^{t}+\int_t^{\infty}\right\}-1 = \\ &= c\int_0^{\infty}\int_0^{t}\lambda\exp\left[-\lambda\int_y^{\infty}\{1-F(x)\}dx\right]dy\,dF(t)+c\int_0^{\infty}\int_t^{\infty}\lambda e^{-\lambda\tau}\exp[-\lambda(y-t)]\,dy\,dF(t)-1 = \\ &= c\int_0^{\infty}\lambda\{1-F(y)\}\exp\left[-\lambda\int_y^{\infty}\{1-F(x)\}dx\right]dy+ce^{-\lambda\tau}-1 = \\ &= c\left[\exp\left[-\lambda\int_y^{\infty}\{1-F(x)\}dx\right]\right]_{y=0}^{y=\infty}+ce^{-\lambda\tau}-1 = c-1. \end{aligned} \tag{19}$$

<u>De verwachting van de lengte van een stoet is dus juist C wagens!</u> Dit kan men ook rechtstreeks bewijzen.

Als

$$F(t) = 1 - e^{-\mu(t-a)} \quad \text{voor} \quad t \geq a, \tag{20}$$

dan kan C berekend worden (neem $z = e^{-\mu y}$):

$$C^{-1} = \int_a^\infty \exp\left[-\lambda \int_y^\infty e^{-\mu(x-a)}\,dx\right] d(1 - e^{-\mu(y-a)}) =$$
$$= \int_0^\infty \exp\left[-\lambda \int_y^\infty e^{-\mu x}\,dx\right] d(1 - e^{-\mu y}) =$$
$$= \int_0^\infty \exp\left[-\frac{\lambda}{\mu} e^{-\mu y}\right] d(1 - e^{-\mu y}) =$$
$$= \int_0^1 \exp\left[-\frac{\lambda}{\mu} z\right] dz = \left(\frac{\lambda}{\mu}\right)^{-1} (1 - e^{-\frac{\lambda}{\mu}}). \tag{21}$$