0173 SP80

Some numerical results on waiting-time distributions for dependent arrival-intervals *)

by J. Th. Runnenburg **)

U.D.C. 311:65.011

S a m e n v a t t i n g

Ter illustratie van een elders beschreven theorie worden hier voor een één-loket wachttijdprobleem enige numerieke resultaten gegeven en besproken. Hieruit blijkt, hoe de wachttijdverdeling (althans in het hier beschouwde geval) variëert met de afhankelijkheid *der aankomstintervallen.*

Usually it is assumed that arrivals (e.g. of customers at a counter, telephone calls in an exchange, ships in a harbour) can be described with sufficient accuracy in the following manner: the random variables $\underline{y}_1, \underline{y}_2, \ldots$ defined by

$$\underline{y}_n = \underline{t}_{n+1} - \underline{t}_n \qquad \text{for } n = 1,2,\ldots,$$

where $\underline{t}_n$ is the time at which the n^{th} arrival occurs ($\underline{t}_1=0$), have the properties

a) all $\underline{y}_n$ have the same distribution function $A(y)=P\{\underline{y}_1 \leqslant y\}$ with $A(0)=0$,
b) the $\underline{y}_n$ are independent random variables,
c) $A(y)=1-e^{-\lambda y}$ for all real $y \geqslant 0$ (λ is a positive constant).

If a), b) and c) hold, the arrival process is a stationary Poisson process. The latter is usually described in a different manner, but the present description is more convenient for our purposes. The Poisson process is often a very convenient model, but that does not concern us here. We wish to consider those situations in which there is a stationary arrival pattern, but where the Poisson process description is *not* satisfactory. We would like to have a theoretically and practically useful alternative to the Poisson model, which can be used for all kinds of problems.

This question has been considered before (in particular with respect to arrivals in waiting-time problems). The following alternative has been studied thoroughly: assumption c) was dropped and only assumptions a) and b) were retained, i.e. $A(y)$ now occurs in the model as an unknown but well-defined distribution function.

No doubt a very large number of stationary arrival patterns which are of practical importance are such that assumption a) applies. But if we drop condition c) it would seem that condition b) should be dropped as well! For the three

*) Report S 298 (SP80), Mathematical Centre, Amsterdam.
**) Professor in Theory of Probability and Analysis at the University of Amsterdam.

assumptions a), b) and c) together describe a physical situation with recognizable practical properties. But if we no longer assume c) to be true but do keep a) and b), then the physical situation described is of a far more complex nature and one may doubt whether it will ever be met with except under very special circumstances. Hence we may suspect that if a model is based on the combination of assumptions a) and b) (arrivals of *renewal* type) this is done for mathematical simplification and not because of the intuitive acceptability of these assumptions.

In my thesis and in a subsequent paper[1]) I have tried to weaken assumption b). To obtain a reasonably manageable mathematical model for arrivals one is bound to make *some* assumption resembling assumption b). Instead of independence I assumed Markov dependence, i.e. I replaced condition b) by b*), where

b*) $P\{\underline{y}_n \leqslant y_n \mid \underline{y}_1 = y_1, \ldots, \underline{y}_{n-1} = y_{n-1}\} = P\{\underline{y}_n \leqslant y_n \mid \underline{y}_{n-1} = y_{n-1}\}$

for $n = 2, 3, \ldots$ and all nonnegative real values of $y_1, y_2, \ldots, y_n$.
Here

$$A(z \mid y) = P\{\underline{y}_n \leqslant z \mid \underline{y}_{n-1} = y\}$$

is an unknown but well-defined conditional distribution function for each value of y, independent of n (we describe a *stationary* arrival pattern). Assumption b) here occurs as a special case: $A(z \mid y) = A(z)$ for all values of y.

For independent $\underline{y}_n$ (i.e. random variables $\underline{y}_1, \underline{y}_2, \ldots$ statisfying a) and b)) with finite variance σ^2, we have

$$\frac{1}{n} \operatorname{var}(\underline{y}_1 + \ldots + \underline{y}_n) = \operatorname{var} \underline{y}_1 = \sigma^2$$

and hence

$$\lim_{n \to \infty} \frac{1}{n} \operatorname{var}(\underline{y}_1 + \ldots + \underline{y}_n) = \sigma^2.$$

We may try to generalize the latter result to Markov-dependent $\underline{y}_n$ (i.e. random variables $\underline{y}_1, \underline{y}_2, \ldots$ satisfying a) and b*)). For it is well-known that under quite general conditions we have

$$\lim_{n \to \infty} \frac{1}{n} \mathscr{E}(\underline{y}_1 + \ldots + \underline{y}_n) = \mathscr{E}\underline{y}_1 = \int_0^\infty y \, \mathrm{d}A(y),$$

---

[1]) Thesis: J. Th. Runnenburg, Markov processes in waiting-time and renewal theory, Poortpers, Amsterdam 1960.
Paper: J. Th. Runnenburg, An example illustrating the possibilities of renewal theory and waiting-time theory for Markov-dependent arrival-intervals, Indag. Math. 23 (1961), 560-576.

both for independent $\underline{y}_n$ and for Markov-dependent $\underline{y}_n$ (in the latter case $A(y)$ is an invariant distribution function to the transition function $A(z|y)$, i.e. a distribution function for which

$$A(z) = \int_0^\infty A(z|y)\, \mathrm{d}A(y)$$

holds). Hence independent $\underline{y}_n$ and Markov-dependent $\underline{y}_n$ are in this respect indistinguishable.

Some technical difficulties necessitated a considerable restriction of the generality. The simplest results (some of which are quoted here) can be found in my paper. There I assumed as I now do here, that the $\underline{y}_n$ only take the values $1,2,\ldots,r$ (where r is a finite positive integer). Instead of the transition function $A(z|y)$ we can then consider the matrix P of transition probabilities p_{jk}, where

$$p_{jk} = P\{\underline{y}_n = k \,|\, \underline{y}_{n-1} = j\}.$$

It is of interest to compare the two models:

independent $\underline{y}_n$

p_{jk}	1	2	$\ldots$	r
1	a_1	a_2	$\ldots$	a_r
2	a_1	a_2	$\ldots$	a_r
$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$
r	a_1	a_2	$\ldots$	a_r

$a_j \geqslant 0, \quad \Sigma\, a_j = 1$

Markov-dependent $\underline{y}_n$

p_{jk}	1	2	$\ldots$	r
1	p_{11}	p_{12}	$\ldots$	p_{1r}
2	p_{21}	p_{22}	$\ldots$	p_{2r}
$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$
r	p_{r1}	p_{r2}	$\ldots$	p_{rr}

$p_{jk} \geqslant 0, \quad \sum_k p_{jk} = 1$ for all j

In our considerations we essentially need the powers of the matrix P with elements p_{jk} (or a_k). For ease of computation I further assume

$$p_{jk} = a_k + \frac{b_j b_k}{a_j} \quad \text{(where } a_j > 0 \quad \text{for each } j\text{).}$$

This means roughly that instead of assuming only one eigenvalue of the matrix P to be unequal zero (in the case of independent $\underline{y}_n$) or all r eigenvalues to be unequal zero (general case of Markov-dependent $\underline{y}_n$), we assume here that exactly *two* eigenvalues of the matrix P are unequal zero.

In order that the p_{jk} are transition probabilities, we must have (these are not further restrictions)

$$\Sigma\, a_j = 1, \quad \Sigma\, b_j = 0, \quad a_j > 0 \text{ for all } j, \quad 0 \leqslant a_k + \frac{b_j b_k}{a_j} \leqslant 1.$$

If we take $b_1 = b_2 = \ldots = b_r = 0$, we return to independent $\underline{y}_n$.

To avoid further complications (Markov chains which are periodic or reducible) we assume that

$$c = \Sigma \frac{b_j^2}{a_j}$$

satisfies

$$|c| < 1.$$

The a_j now describe the invariant distribution of the Markov-dependent $\underline{y}_n$, i.e.

$$\lim_{n\to\infty} P\{\underline{y}_n = k \mid \underline{y}_1 = j\} = a_k \quad \text{for all } j \text{ and } k.$$

It turns out that

$$\lim_{n\to\infty} \frac{1}{n} \operatorname{var} (\underline{y}_1 + \ldots + \underline{y}_n) = \sigma^2 + 2 \frac{b^2}{1-c},$$

where

$$\sigma^2 = \Sigma j^2 a_j - (\Sigma j\, a_j)^2,\ b = \Sigma j\, b_j.$$

A further quantity to consider is $\underline{l}_n$, the number of arrivals in the interval $[0,n]$ on the time axis (by assumption at time 0 an arrival occurs). We find

$$\lim_{n\to\infty} \frac{1}{n} \mathscr{E} \underline{l}_n = \frac{1}{\Sigma j\, a_j}$$

and

$$\lim_{n\to\infty} \frac{1}{n} \operatorname{var} \underline{l}_n = \frac{1}{(\Sigma j\, a_j)^3} \left(\sigma^2 + 2 \frac{b^2}{1-c}\right).$$

One can use the arrival pattern described by Markov-dependent $\underline{y}_n$ as model for the arrivals in a waiting-time problem and so obtain some insight in the effect of making assumption b) rather than the more general b*).

We assume that customers arrive at a counter to be served in the order in which they arrive. The length $\underline{y}_n$ of the arrival-interval between the n^{th} and $(n+1)^{\text{st}}$ customer can only be 1 or 2 units of time (i.e. we take $r=2$). The random variables $\underline{y}_n$ are Markov-dependent with transition matrix

p_{jk}	1	2
1	ε	$1-\varepsilon$
2	$1-\varepsilon$	ε

where $0<\varepsilon<1$, $a_1=a_2=\frac{1}{2}$ and $b_1= -b_2=\frac{1}{2}\sqrt{2\varepsilon-1}$. This means that we consider exclusively arrival-intervals with invariant distribution described by $a_1=a_2=\frac{1}{2}$, i.e. a *fixed* invariant distribution, *independent* of ε (the *degree* of Markov-dependence). For $\varepsilon = \frac{1}{2}$ we have again independent $\underline{y}_n$, for ε close to 0 lengths

1 and 2 almost alternate and for ε close to 1 hardly any change-over from length 1 to length 2 (or from length 2 to length 1) occurs. In order to obtain a larger class of transition matrices we allow that b_1 and b_2 take purely imaginary values (for $0<\varepsilon<\frac{1}{2}$).

The service-times $\underline{s}_1, \underline{s}_2, \ldots$ of customers $1, 2, \ldots$ are assumed to be independent random variables, independent of the $\underline{y}_1, \underline{y}_2, \ldots$ and to have the same exponential distribution function

$$P\{\underline{s}_1 \leqslant s\} = 1-e^{-\mu s} \text{ for all real } s \geqslant 0 \quad (\mu \text{ is a positive constant}).$$

Let us now consider the stationary situation which will develop from the given conditions (provided μ is large enough), i.e. the distribution function

$$F_\varepsilon(w) = \lim_{n\to\infty} P\{\underline{w}_n \leqslant w\},$$

where $\underline{w}_n$ is the waiting-time of the n^{th} customer (with $\underline{w}_1=0$). The mean distance between two successive arrivals is in the stationary situation

$$\lim_{n\to\infty} \mathscr{E}\underline{y}_n = \tfrac{1}{2}\,.\,1 + \tfrac{1}{2}\,.\,2 = \tfrac{3}{2},$$

while

$$\sigma_\varepsilon^2 = \lim_{n\to\infty} \frac{1}{n} \operatorname{var}(\underline{y}_1+\ldots+\underline{y}_n) = \frac{\varepsilon}{4(1-\varepsilon)}.$$

Hence this quantity varies between 0 ($\varepsilon = 0$) and ∞ ($\varepsilon = 1$)!

Furthermore we find for the correlation coefficient

$$\lim_{n\to\infty} \varrho(\underline{y}_n, \underline{y}_{n+1}) = 2\,\varepsilon - 1.$$

In fig. 1 the waiting-time distribution $F_\varepsilon(w)$ has been drawn for different values of ε as a function of w for constant $\mu=1$. This means that the traffic intensity ϱ (mean service-time devided by mean arrival-interval) is $\frac{2}{3}$ for all curves. Only ε, the degree of Markov-dependence of the arrival-intervals, changes from curve to curve. For $\varepsilon = 0{,}5$ we find the traditional curve for independent arrival-intervals. For $\varepsilon = 0$ arrival-intervals with length 1 and length 2 alternate. If the first arrival-interval $\underline{y}_1$ has the invariant distribution, i.e. if $P\{\underline{y}_1=1\} = P\{\underline{y}_1=2\} = \frac{1}{2}$, then $\lim\limits_{n\to\infty} P\{\underline{w}_n \leqslant w\}$ exists and is given by $F_0(w)$. We have $F_0(w) = \lim\limits_{\varepsilon\downarrow 0} F_\varepsilon(w)$. For $\varepsilon = 1$ the curve $F_1(w) = \lim\limits_{\varepsilon\uparrow 1} F_\varepsilon(w)$ has been drawn. It is no longer a proper distribution function. Here $\lim\limits_{w\to\infty} F_1(w) = \frac{1}{2}$.

The curves for $\varepsilon=0{,}75$ and $\varepsilon=0{,}964$ have been drawn to cover the gap between $\varepsilon=0{,}5$ and $\varepsilon=1$. The curious 0,964 is due to an interpolation: one may start

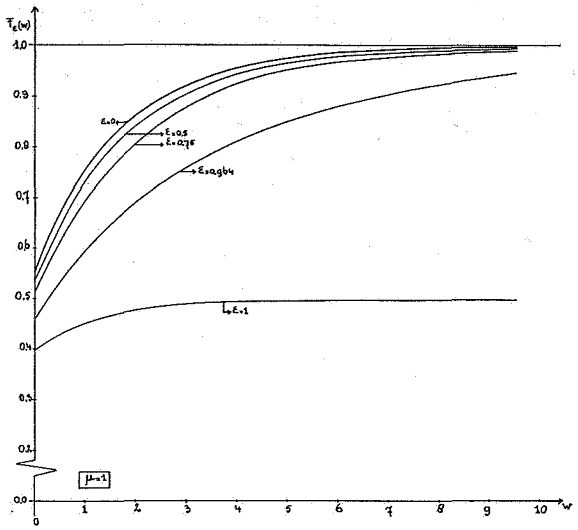

Fig. 1. Waiting-time distribution in the stationary situation.

from a positive number w_0 and a probability p_0 and compute ε from $F_\varepsilon\,(w_0) = p_0$ to obtain a curve passing through p_0 at w_0.

We note that the distribution functions for ε=0 and ε=0,5 differ but little, while a large gap occurs between ε=0,5 and ε=1. This can also be seen from the following table.

ε	$\varrho=2\,\varepsilon-1$	$\mathscr{E}\underline{w}$	$\sigma\,\underline{w}$
0	–1	0,75	1,4
0,25	–0,5	0,79	1,5
0,5	0	0,86	1,6
0,75	0,5	1,03	1,8
0,964	0,928	2,11	3,7
1	1	∞	∞

TABLE 1. *Some constants for the curves of fig. 1.*

Fig. 2 contains

$$F_\varepsilon^{(1)}\,(w) = \lim_{n\to\infty} \;\; P\,\{\underline{w}_{n+1} \leqslant w \mid \underline{y}_n = 1\}\,,$$

which may be described as the waiting-time distribution function in the stationary situation of a customer arriving one time-unit after the previous one.

In fig. 3 we find

$$F_\varepsilon^{(2)}(w) = \lim_{n\to\infty} \;\; P\,\{\underline{w}_{n+1} \leqslant w \mid \underline{y}_n = 2\},$$

which is the waiting-time distribution function in the stationary situation of a customer arriving two time-units after the previous one.

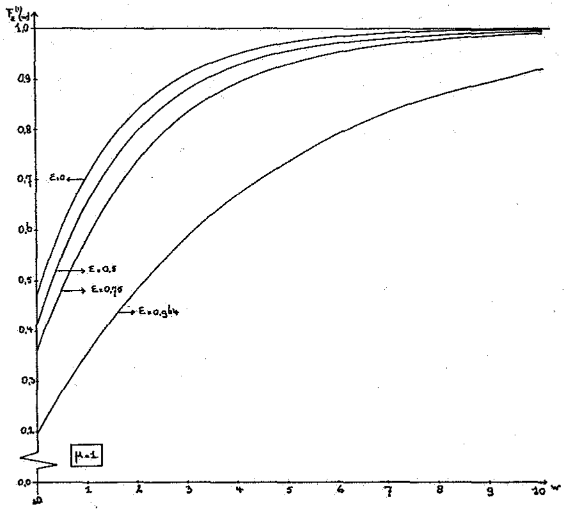

Fig. 2. Waiting-time distribution given the previous customer came one time-unit earlier.

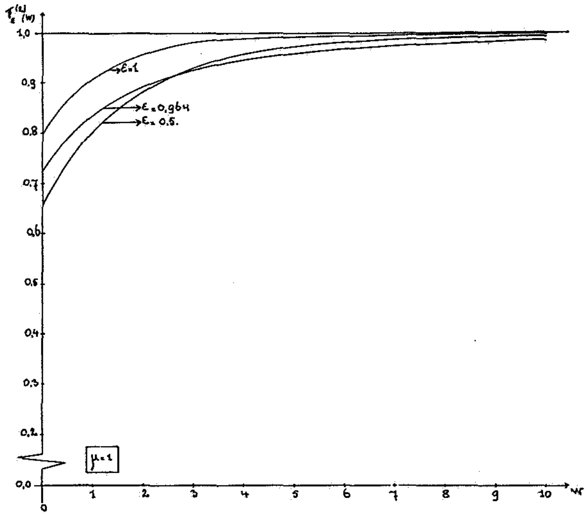

Fig. 3. Waiting-time distribution given the previous customer came two time-units earlier.

In fig. 2 and fig. 3 the curves for $\varepsilon=0$ and $\varepsilon=1$ are to be interpreted as limiting-curves, i.e.

$$F_0^{(\nu)}(w) = \lim_{\varepsilon \downarrow 0} F_\varepsilon^{(\nu)}(w) \text{ and } F_1^{(\nu)}(w) = \lim_{\varepsilon \uparrow 1} F_\varepsilon^{(\nu)}(w)$$

for $\nu=1$ and $\nu=2$. The curves for $\varepsilon=0$ correspond to waiting-time distributions for *periodic* arrivals. The theory contains periodic arrival patterns as a special case!

It is a rather curious fact that for $\varepsilon=0,5$ it is quite easy to obtain $F_{0,5}^{(1)}(w)$ and $F_{0,5}^{(2)}(w)$ from $F_{0,5}(w)$, while for $\varepsilon\neq 0,5$ this is no longer true. We have

$$F_\varepsilon(w) = \tfrac{1}{2}\, F_\varepsilon^{(1)}(w) + \tfrac{1}{2}\, F_\varepsilon^{(2)}(w).$$

Not all curves computed have been drawn. The ones for $\varepsilon=0,5$ (independent arrival-intervals) are always given.

ε	$\mathscr{E}\underline{w}^{(1)}$	$\sigma\underline{w}^{(1)}$
0	0,88	1,5
0,25	—	—
0,5	1,09	1,7
0,75	1,42	2,0
0,964	3,53	4,3
1	∞	∞

TABLE 2. *Some constants for the curves of fig. 2.*

ε	$\mathscr{E}\underline{w}^{(2)}$	$\sigma\underline{w}^{(2)}$
0	0,62	1,3
0,25	—	—
0,5	0,64	1,4
0,75	1,56	1,5
0,964	0,70	2,1
1	∞	∞

TABLE 3. *Some constants for the curves of fig. 3.*

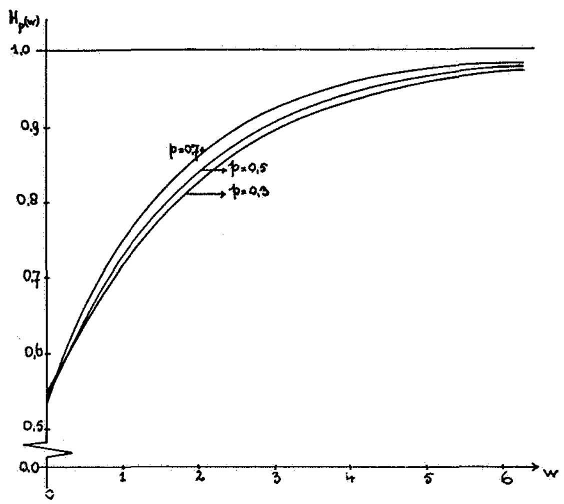

Fig. 4. Waiting-time distribution for different invariant arrival-interval distributions for constant ϱ.

In fig. 4 some curves for *independent* arrival-intervals have been drawn. If length 1 has probability p and length 2 probability $q=1-p$ to occur, then $\mathscr{E}\underline{y}_1=p+2\cdot(1-p)=2-p$. We have drawn $H_p(w)$, the waiting-time distribution function in the stationary situation, for different values of p, while the traffc density ϱ was kept constant. Hence $1/\mu = \mathscr{E}\underline{s}_1$ has *not* been kept constant. Therefore fig. 4 does not really belong in the series of figures we wish to consider, for we want to study the effect of varying the *dependence* of the arrival-intervals, without changing the invariant distribution of either the arrival-intervals or the service-times.

μ	ϱ	p	$\mathscr{E}\underline{w}$	$\sigma\underline{w}$
0,8824	$\frac{2}{3}$	0,3	0,93	1,7
1	$\frac{2}{3}$	0,5	0,86	1,6
1,1538	$\frac{2}{3}$	0,7	0,75	1,4

TABLE 4. *Some constants for the curves of fig. 4.*

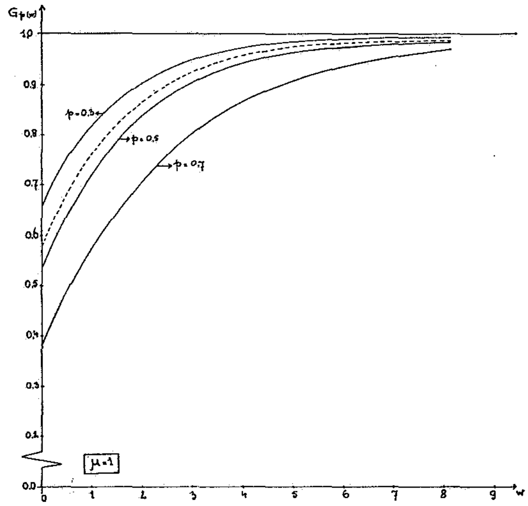

Fig. 5. Waiting-time distribution for different invariant arrival-interval distributions for constant μ.

In fig. 5 again some curves for *independent* arrival-intervals have been drawn. Again each arrival-interval has length 1 with probability p and length 2 with probability $q=1-p$. We have now drawn $G_p(w)$, the waiting-time distribution function in the stationary situation, for different values of p, while the service-time parameter μ was kept constant ($\mu=1$).

The dotted curve in fig. 5 corresponds to the waiting-time distribution function in the stationary situation, which is obtained if all arrival-intervals have *constant* length $\frac{3}{2}$ (again $\mu=1$).

μ	ϱ	p	$\mathscr{E}\underline{w}$	$\sigma\underline{w}$
1	0,5882	0,3	0,52	1,2
1	$\frac{2}{3}$	0,5	0,86	1,6
1	0,7693	0,7	1,60	2,4
1	dotted curve		0,72	1,4

TABLE 5. *Some constants for the curves of fig. 5.*

Finally one may obtain a rough picture of the importance of the different influences on the waiting-time distribution functions in the stationary situation by comparing fig. 1 and fig. 5. Roughly the effect of changing $\varepsilon=0{,}5$ to $\varepsilon=0{,}75$ is of the same order as changing $p=0{,}5$ to $p=0{,}55$.

Formulae to compute these and more complicated results may be found in the literature mentioned.