https://doi.org/10.26219/heal.aueb.6246

SCHOOL OF INFORMATION SCIENCES & TECHNOLOGY

DEPARTMENT OF STATISTICS

POSTGRADUATE PROGRAM

A Study of Poisson and Renewal Reward Processes with Applications.

By

Nektarios K. Makris

A THESIS
Submitted to the Department of Statistics
of the Athens University of Economics and Business
in partial fulfilment of the requirements for
the degree of Master of Science in Statistics

Athens, Greece
April 2021

https://doi.org/10.26219/heal.aueb.6246

https://doi.org/10.26219/heal.aueb.6246

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

ΣΧΟΛΗ ΕΠΙΣΤΗΜΩΝ & ΤΕΧΝΟΛΟΓΙΑΣ ΤΗΣ ΠΛΗΡΟΦΟΡΙΑΣ

ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ

ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ

Μελέτη της Διαδικασίας Poisson και των Ανανεωτικών Διαδικασιών με Αμοιβές.

Νεκτάριος Κ. Μακρής

ΔΙΑΤΡΙΒΗ

Που υποβλήθηκε στο Τμήμα Στατιστικής
του Οικονομικού Πανεπιστημίου Αθηνών
ως μέρος των απαιτήσεων για την απόκτηση
Διπλώματος Μεταπτυχιακών Σπουδών στη Στατιστική

Αθήνα
Απρίλιος 2021

https://doi.org/10.26219/heal.aueb.6246

DEDICATION

To my beloved parents…

https://doi.org/10.26219/heal.aueb.6246

https://doi.org/10.26219/heal.aueb.6246

ACKNOWLEDGEMENTS

I would like to give many thanks to my supervisor Epaminondas Kyriakidis for his valuable help and guidance towards the completion of this master thesis.

I would also like to thank my parents for their support.

https://doi.org/10.26219/heal.aueb.6246

https://doi.org/10.26219/heal.aueb.6246

Vita

I was born in Athens in 1995. After my graduation from High school in 2013, I enrolled the Mathematics Department of the National and Kapodistrian University of Athens. I concluded my studies at September 2018. I was admitted as a post-graduate student in the MSc. program of the Department of Statistics of the Athens University of Economics and Business in September 2019.

https://doi.org/10.26219/heal.aueb.6246

https://doi.org/10.26219/heal.aueb.6246

ABSTRACT

Nektarios Makris

A Study of Poisson and Renewal Reward Processes with Applications.

April 2021

In the present dissertation, we study extensively the Poisson Process and its various variants or generalizations such as the non-stationary Poisson and the Markov modulated batch Poisson process. We also study more generally various renewal processes where a reward structure lurks behind the stochastic process. Most theoretical results are also illustrated with applications with emphasis on inventory control problems and service control problems.

https://doi.org/10.26219/heal.aueb.6246

https://doi.org/10.26219/heal.aueb.6246

ΠΕΡΙΛΗΨΗ

Νεκτάριος Μακρής

Μελέτη της διαδικασίας Poisson και των ανανεωτικών διαδικασιών με αμοιβές.

Απρίλιος 2021

Στην παρούσα διπλωματική μελετάμε εκτεταμένα την στοχαστική διαδικασία Poisson και διάφορες παραλλαγές ή γενικεύσεις της όπως την non-stationary Poisson και την Markov modulated batch Poisson process. Επίσης μελετάμε πιο γενικά διάφορες ανανεωτικές διαδικασίες όπου μια δομή αμοιβών υποβόσκει πίσω από την στοχαστική διαδικασία. Στα περισσότερα θεωρητικά αποτελέσματα δίνονται και εφαρμογές με έμφαση σε προβλήματα ελέγχου αποθεμάτων και προβλήματα ελέγχου εξυπηρέτησης.

https://doi.org/10.26219/heal.aueb.6246

https://doi.org/10.26219/heal.aueb.6246

https://doi.org/10.26219/heal.aueb.6246

TABLE OF CONTENTS

https://doi.org/10.26219/heal.aueb.6246

https://doi.org/10.26219/heal.aueb.6246

LIST OF FIGURES

Figure

https://doi.org/10.26219/heal.aueb.6246

https://doi.org/10.26219/heal.aueb.6246

https://doi.org/10.26219/heal.aueb.6246

CHAPTER 1

INTRODUCTION

The Poisson process is a stochastic process that "counts" the number of occurrences of any particular event through time, therefore it is a counting (stochastic) process. Examples include the number of network failures per day, the number of file server virus infections at a data center during a 24-hour period, the arrivals of customers at a counter etc. This is the most common arrival process it is the most suitable model for systems with a great number of potential customers that use the service rarely and independently from each other. The models that will be examined in this chapter mostly are related to this arrival process. It is denoted by the capital letter (M) which is derived by the Memoryless or Markovian property of the Poisson. In various applied probability problems, the Poisson process is a natural modeling technique and it simulates several processes in the natural world. The process also facilitates tractable mathematical study because of its memoryless property. In the following chapter, the Poisson process and the memoryless property are discussed in detailed. To illustrate the utility of the model, illustrative examples are given. We will also deal with the compound Poisson process. Customers arrive individually in a Poisson arrival process, while customers arrive in batches in a compound Poisson arrival process. The non-stationary method of Poisson, which is discussed, is another generalization of the Poisson process. In the Poisson process the intensity at which events occur is time independent. In the non-stationary Poisson process, this assumption is removed.

https://doi.org/10.26219/heal.aueb.6246

https://doi.org/10.26219/heal.aueb.6246

CHAPTER 2

THE POISSON PROCESS AND RELATED PROCESSES

The content of this chapter is mainly based on chapter 1 of the book of Tijms (2003) and on chapter 2 of the book of Ross (1996).

2.1 THE POISSON PROCESS

First of all, by a stochastic process, we mean a family of random variables $\{X(t), t \in T\}$ where t is a point in a space T called the parameter space, and where for each $t \in T$, $X(t)$ is a point in a space S called the state space. T, S can be discrete or continuum.

Some basic types of stochastic processes include:

- Stationary processes
- Markov chains
- Renewal processes

Definition 2.1.1 A stochastic process $\{N(t), t \geq 0\}$ is said to be a counting process if $N(t)$ represents the total number of events that have occured up to time t. Hence, a counting process $N(t)$ must satisfy:

- $N(t) \geq 0, \forall t \geq 0$
- $N(t) \in \mathbb{N}$
- If $s < t$, then $N(s) \leq N(t)$
- For $s < t$, $N(t) - N(s)$ equals the number of events that have occurred in the interval $(s, t]$.

Note: A counting process $\{N(t), t \geq 0\}$ is said to possess **independent increments** if the number of events that occur in disjoint time intervals are independent.

A counting process is said to possess **stationary increments** iff:

The distribution of $N(t) - N(s), 0 < s < t$ depends only on the difference $t - s$ and it is independent of the specific values of s, t. The Poisson process is a special counting process and there are several equivalent ways to be defined. We start off with $X_1, X_2, \ldots$ independent, identically and positive random variables (we use positive random variables because we want to model time-

https://doi.org/10.26219/heal.aueb.6246

period until an event to be occurred). Let's consider X_n as the time before the occurrence of the nth event and after the occurrence of the $(n-1)$-th event. Let:

$$S_0 = 0 \quad and \quad S_n = \sum_{k=1}^{n} X_k \qquad n = 1,2,3, \dots$$

Then S_n is the specific moment that the nth event occurs. For every $t \geq 0$, we define the random variable $N(t)$ by

$$N(t) = \text{the largest integer } n \geq 0 \text{ for which } S_n \leq t$$

The random variable $N(t)$ is the number of occurrences up to the moment t.

There are many equivalent ways to define the Poisson process.

Definition 2.1.2 The counting process $\{N(t), t \geq 0\}$ is said to be a Poisson process having rate λ, $\lambda > 0$, if the interoccurrence times $X_1, X_2, \dots$ are independent and identically distributed with distribution function

$$F_{X_i}(x) = 1 - e^{-\lambda x} \text{ , } \forall i, x \geq 0$$

It seems to be conservative to assume exponentially distributed interoccurrence intervals, but it seems that for certain real world problems, the Poisson process is an excellent model. The reason lies in the following fundamental outcome. Suppose there are a very large number of independent stochastic processes at the micro-stage, where each separate microprocess seldom produces an event. The superposition of all these microprocesses then functions roughly as a Poisson process at the macro-stage. This insightful observation is similar to the well-known result that the number of successes in a very large number of independent Bernoulli trials with a very small probability of success approximately follows the Poisson distribution.

Definition 2.1.3 (equivalent definition of Poisson process)

The counting process $\{N(t), t \geq 0\}$ is said to be a Poisson process having rate λ, $\lambda > 0$, iff:

1) $N(0) = 0$
2) The process has independent increments
3) The number of events in any interval of length t is Poisson distributed with mean λt and for all $s, t \geq 0$

https://doi.org/10.26219/heal.aueb.6246

$$P(N(t+s)-N(s)=n)=e^{-\lambda t}\frac{(\lambda t)^n}{n!}, n=0,1,2\ldots$$

We can see that the above probability does not depend on the value of s.
So it is clearly that we have assumed stationary increments.

Definition 2.1.4 (another equivalent definition of Poisson process)
The counting process $\{N(t), t \geq 0\}$ is said to be a Poisson process having rate λ, $\lambda > 0$, iff:

(2.1.1)

1) $N(0)=0$
2) The process has stationary and independent increments.
3) $P(N(h)=0)=1-\lambda h+o(h)$, as $h \to 0$,
4) $P(N(h)=1)=\lambda h+o(h)$, as $h \to 0$,
5) $\mathrm{P}(N(h) \geq 2)=\mathrm{o}(h)$, as $h \to 0$,

where a function f is said to be $o(h)$ iff:

$$\lim_{h\to 0}\frac{f(h)}{h}=0$$

2.1.1 The memoryless Property

A Poisson process is a stochastic process without memory; that an occurrence has just happened or that an event has not occurred in a long time gives us no hint for the possibility that another event will occur shortly. The Poisson approach is used to model radioactive decay, web document queries, and order/call/show clients in queuing theory. Let T be the random variable representing the waiting time between the last event that has already occurred and the next event that is going to happen. The probabilities are related to this memoryless property of the Poisson process:

$$P(T>t+s|T>s)=P(T>t)$$

In other words, the likelihood that an occurrence will not occur in the next t minutes, $P(T>t+s|T>s)$, if we have already waited for a time s is the same if we had not already waited for time s. We will show that a bit later.
We first need to find out what is the distribution of the random variable $N(t)$.
We are going to use the fact that the sum of k independent and identically distributed random variables which follow the same exponential distribution has an Erlang distribution:

https://doi.org/10.26219/heal.aueb.6246

$$P(S_k \leq t) = 1 - \sum_{j=0}^{k-1} e^{-\lambda t} \frac{(\lambda t)^j}{j!}, \qquad t \geq 0$$

The Erlang (k, λ) distribution has the probability density function $\lambda^k t^{k-1} e^{-\lambda t} / (k-1)!$

Theorem 2.1.3 For every $t > 0$,

$$P(N(t) = k) = e^{-\lambda t} \frac{(\lambda t)^k}{k!}, \qquad k = 0,1,2, \dots \tag{2.1.2}$$

That is, $N(t)$ is Poisson distributed with mean = λt.

Proof: We use the fact that

$$\mathrm{P}(\mathrm{N(t)} \geq k) = (\text{At least } k \text{ events have occured up to time } t) =$$
$$= P(\text{The } k\text{th event happened before the moment } t)$$

Hence,

$$\mathrm{P}(N(t) \geq k) = \mathrm{P}(S_k \leq t) = 1 - \sum_{j=0}^{k-1} e^{-\lambda t} \frac{(\lambda t)^j}{j!},$$

$$\mathrm{P}(\mathrm{N}(t) \leq k-1) = \sum_{j=0}^{k-1} e^{-\lambda t} \frac{(\lambda t)^j}{j!}$$

which is the cumulative distribution function of Poisson with mean = λt.

Now we are going to show the memoryless property of the Poisson process. We will show that the time we have already waited for an event does not affect the distribution of the remaining time until the occurrence of the event.

Let γ_t denote the random variable which represents the remaining time.

$$\boldsymbol{\gamma_t = the\ waiting\ time\ from\ epoch\ t\ until\ the\ next\ event} \quad \forall \boldsymbol{t} \geq \mathbf{0}$$

One of the most important theorem for our analysis is the following:

Theorem 2.1.4 $\forall t \geq 0$, the random variable γ_t follows the same exponential distribution with parameter $1/\lambda$. That is

$$P(\gamma_t \leq x) = 1 - e^{-\lambda x}, \qquad x \geq 0 \tag{2.1.3}$$

independently of t.

Proof: Fit $t \geq 0$. The event $\gamma_t > x$ occurs only if one of the mutually exclusive events $X_1 > t + x, X_1 \leq t, X_1 + X_2 > t + x, X_1 + X_2 \leq t, X_1 + X_2 \leq t, X_1 + X_2 + X_3 > t + x, \dots$ occurs. This gives

$$P(\gamma_t > x) = P(X_1 > t + x) + \sum_{n=1}^{\infty} P(S_n \leq t, S_{n+1} > t + x)$$

By conditioning on S_n, we find out

$$P(S_n \leq t, S_{n+1} > t + x) = \int_0^t P(S_{n+1} > t + x | S_n = y)\} \lambda^n \frac{y^{n-1}}{(n-1)!} e^{-\lambda y} dy$$
$$= \int_0^t P(X_{n+1} > t + x - y) \lambda^n \frac{y^{n-1}}{(n-1)!} e^{-\lambda y} dy$$

This gives

$$\mathrm{P}(\gamma_t > \mathrm{x}) = e^{-\lambda(t+x)} + \sum_{n=1}^{\infty} \int_0^t e^{-\lambda(t+x-y)} \lambda^n \frac{y^{n-1}}{(n-1)!} e^{-\lambda y} \, dy$$
$$= e^{-\lambda(t+x)} + \int_0^t e^{-\lambda(t+x-y)} \lambda \, dy$$
$$= e^{-\lambda(t+x)} + e^{-\lambda(t+x)} \left(e^{\lambda t} - 1\right) = e^{-\lambda x}$$

It is correct to change the order of the integral because nothing is negative.

The theorem notes that the waiting time before the next arrival at each point in time has the same exponential distribution as the initial time of interarrival, regardless of how long ago the last event happened. The only renewal process having this memoryless-property is the Poisson and here **is why:** If we assume memoryless property, we will show that the waiting time (let's say T) is necessarily exponentially distributed, so $P(T > t) = e^{-\lambda t}$.

Using the definition of conditional probability we have:

$$P(T > t + s | T > s) = P(T > t + s \; and \; T > s) / P(T > s) \; \Rightarrow$$
$$P(T > t + s) = P(T > t + s | T > s) \cdot P(T > s)$$

And the memoryless property can be written as:

$$P(T > t + s) = P(T > t) \cdot P(T > s)$$

Now let's say $S(t) = P(T > t)$ the survival function. Since it is the probability that an event/arrival hasn't occurred until time t, then the memoryless-property in terms of the survival function can be written as:

$$S(t + s) = S(t) * S(s)$$

Let's say $s, t = 1/2$. Then,

$$S(1) = S\left(\frac{1}{2} + \frac{1}{2}\right) = S\left(\frac{1}{2}\right) \cdot S\left(\frac{1}{2}\right) = S\left(\frac{1}{2}\right)^2$$

https://doi.org/10.26219/heal.aueb.6246

If we divide 1 into another fraction, we see a pattern:

$$S(1) = S\left(3\frac{1}{3}\right) = S\left(\frac{1}{3}+\frac{1}{3}+\frac{1}{3}\right) = S\left(\frac{1}{3}\right)\cdot S\left(\frac{1}{3}+\frac{1}{3}\right) = S\left(\frac{1}{3}\right)^3$$

for integers $n > 0$:

$$S(1) = S\left(\frac{1}{n}\right)^n$$

And for any integer m:

$$S\left(m\frac{1}{n}\right) = S\left(\frac{1}{n}+\frac{1}{n}+\cdots+\frac{1}{n}\right) = S\left(\frac{1}{n}+\frac{m-1}{n}\right) = S\left(\frac{1}{n}\right)S\left(\frac{m-1}{n}\right) = \cdots = S\left(\frac{1}{n}\right)^m$$

This gives us:

$$\mathrm{S}\left(\frac{m}{n}\right) = \left[S(1)^{\frac{1}{n}}\right]^m = S(1)^{\frac{m}{n}}$$

We can approximate any actual number (by selecting m and n) $t > 0$ with $\frac{m}{n}$ so we can write $S(t) = S(1)^t$

If we define $\lambda = -\log S(1)$ we have $S(t) = P(T > t) = e^{-\lambda t}$

Note: $\lambda > 0$ because $S(1) \in (0,1)$

In conclusion we have: $P(T \leq t) = 1 - e^{-\lambda t}$

We can conveniently go back and show that the S = survival-function for the exponential distribution has the memoryless-property:

$$\mathrm{S}(t+s) = e^{-\lambda(t+s)} = e^{-\lambda t}e^{-\lambda s} = \mathrm{S}(t)\mathrm{S}(s)$$

For general arrival processes, this extraordinary property does not hold (e.g. like in the example with constant interoccurrence times). The absence of memory of the Poisson describes the process's computational tractability. The study will not require a state vector to maintain track of the time passed from the last arrival. Of note, the memoryless property of the Poisson process is directly linked to the exponential distribution's lack of memory.

Theorem 2.1.1 states that Poisson-distributed with mean λs is the number of arrivals during the time interval $(0, s)$. More commonly, the number of arrivals has a Poisson distribution of mean λs at any time interval of length s.

$$P(N(u+s) - N(u)) = e^{-\lambda s}\frac{(\lambda s)^k}{k!}, k = 0,1,2, \ldots$$

So we say that Poisson process possesses stationary increments. The proof of this comes from Theorem 2.1.4 the time spent between a given epoch u and the

 https://doi.org/10.26219/heal.aueb.6246

epoch of the first arrival after u has the same exponential distribution as the time elapsed between epoch 0 and the epoch of the first arrival after epoch 0. Next we use the proof of Theorem 2.1.1.

Poisson also possesses independent increments. This property is also an outcome of the memoryless-property.

Example 2.1.1 In a hockey game the score is modeled as a Poisson-process. We assume that the mean time between goals is 15 mins.

i. Find the probability that a 4th goal happens in the last 5 minutes in a 60 minute game.
ii. Assume 3 goals are scored (both teams) in a game. What is the average time of the 3rd goal?

Solution: The parameter of the hockey Poisson process is $\lambda = 1/15$.

i)

$$P(55 < S_4 \leq 60) = \frac{1}{6}\int_{55}^{60} (1/15)^4\, t^3 e^{-t/15} dt = 0.068$$

ii)

$$\begin{aligned} E(S_3|S_3 < 60) &= \frac{1}{P(S_3 < 60)}\int_0^{60} t\, f_{S_3}(t) dt = \\ &= \frac{1}{P(S_3 < 60)}\int_0^{60} t\,\frac{(1/15)^3 t^2 e^{-t/15}}{2}\, dt = \\ &= \frac{25.4938}{0.7619} = 33.461 \text{ minutes.} \end{aligned}$$

2.1.2 Merging and splitting of Poisson process

Many implementations include combining individual Poisson processes or separating events of the Poisson process into separate groups. Thinning or splitting a Poisson process refers to classifying each event/arrival, independently, into one of a finite number of different types. The events of a given type also form Poisson processes, and these processes are independent.

For example, suppose that the number of clients visiting a bank at a given time interval I is $N \sim Poisson(\lambda)$. Assume that each consumer takes a loan with a probability p, independently of other clients, and independently of the

https://doi.org/10.26219/heal.aueb.6246

valuation of N. Let N_1 be the number of clients that obtain the loan at that time period. Also let N_2 be the number of customers who deposit money.

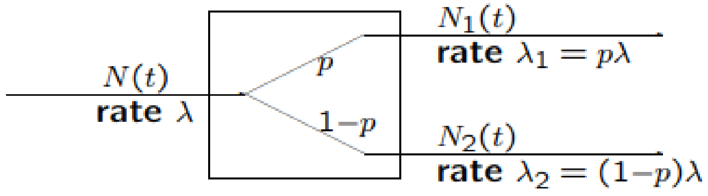

figure 2.1 splitting a Poisson process

Now we will show that: assume that $\{ N_1(t), t \geq 0 \}$ and $\{ N_2,\ t \geq 0 \}$ are independent Poisson processes with rates $\lambda_{1,}\, \lambda_2$ respectively, where the process $\{N_i(t), t \geq 0\}$ (i = 1, 2) corresponds to type i arrivals. Let $N(t)$ = $N_1(t) + N_2(t)$.

Then the merged process $\{N(t),\ t \geq 0\}$ is a Poisson process with rate λ = λ_1 + λ_2. Denoting by Z_k the interarrival time between the (k - 1)th and kth arrival in the merged process and lettin $I_k = i$ if the kth arrival in the merged process is a type i arrival, then for any $k = 1,2, \ldots$

$$P(I_k = i | Z_k = t) = \frac{\lambda_i}{\lambda_1 + \lambda_2}, \qquad i = 1,2 \tag{2.1.4}$$

independently of t.

The result (2.1.4) states that with probability $\lambda_i/(\lambda_1 + \lambda_2)$ the next arrival is of type i regardless of how long it takes until the occurrence of the next arrival.

It is obvious that the process $\{N(t), t \geq 0\}$ satisfies rules 1), 2) of the Poisson definition 2.1.4 To verify 4) we have:

$$P(\text{one arrival in } (t, t + \Delta t])$$

$$= \sum_{i=1}^{2} P(\text{one arrival of type } i \text{ and no arrival of the other type in } (t, t + \Delta t])$$

$$= [\lambda_1 \Delta t + o(\Delta t)][1 - \lambda_2 \Delta t + o(\Delta t)] + [\lambda_2 \Delta t + o(\Delta t)][1 - \lambda_1 \Delta t + o(\Delta t)]$$

$$= (\lambda_1 + \lambda_2)\Delta t + o(\Delta t) \text{ as } \Delta t \to 0$$

And to verify 3) we have:

$$P(\text{no arrival in } (t, t + \Delta t]) = [1 - \lambda_1 \Delta t + o(\Delta t)][1 - \lambda_2 \Delta t + o(\Delta t)]$$

$$= 1 - (\lambda_1 + \lambda_2)\Delta t + o(\Delta t) + o(\Delta t) \ as\ \Delta t \to 0$$

To verify 5): $\mathrm{P}(\text{at least two arrivals in } (\mathrm{t}, \mathrm{t} + \Delta t]) = 1 - \mathrm{P}(\mathrm{N}(\mathrm{h}) = 0) - \mathrm{P}(\mathrm{N}(\mathrm{h}) = 1) = 1 - \big(1 - \lambda\mathrm{h} + \mathrm{o}(\mathrm{h})\big) - \big(\lambda\mathrm{h} + \mathrm{o}(\mathrm{h})\big) = \mathrm{o}(\mathrm{h})$

https://doi.org/10.26219/heal.aueb.6246

To prove the other statement, denote by the random variable Y_i the interarrival time in the process $\{N_i(t), t \geq 0\}$. Then

$$P(Z_k > t, I_k = 1) = P(Y_2 > Y_1 > t)$$
$$= \int_t^\infty P(Y_2 > Y_1 > t | Y_1 = x)\lambda_1 e^{-\lambda_1 x}\, dx$$
$$= \int_t^\infty e^{-\lambda_2 x}\lambda_1 e^{-\lambda_1 x}\, dx = \frac{\lambda_1}{\lambda_1 + \lambda_2} e^{-(\lambda_1+\lambda_2)t}.$$

By taking $t = 0$, we find $\mathrm{P}(I_k = 1) = \lambda_1/(\lambda_1 + \lambda_2)$. Since $\{\mathrm{N(t)},\ t \geq 0\}$ is a Poisson process with rate $\lambda_1 + \lambda_2$, we have $\mathrm{P}(Z_k > t) = exp[-(\lambda_1 + \lambda_2)t]$.Hence

$$P(I_k = 1, Z_k > t) = P(I_k = 1)P(Z_k > t)$$

showing that

Theorem 2.1.5 Let $\{N(t), t \geq 0\}$ be a Poisson process with rate = λ. Suppose that each event-arrival of the process is graded as a type 1 or type 2 with the corresponding probabilities p_1 and p_2 regardless of any other arrivals. Let $N_i(t)$ be the type i number of arrivals up to time t. Then $\{N_1(t), t \geq 0\}$ and $\{N_2(t), t \geq 0\}$ are two separate independent Poisson processes with the corresponding λp_1 and λp_2 rates.

Proof: Obviously, the process $\{N_i(t), t \geq 0\}$ satisfies 1), 2), 5) of the Poisson Definition 1.1.5, to verify property 4), note that

$$P(\text{one arrival of type } i \text{ in } (t, t + \Delta t]) = (\lambda\Delta t)p_i + o(\Delta t) = (\lambda p_i)\Delta t + o(\Delta t)$$

So $P(\text{no arrivals of type } i \text{ in } (t, t + \Delta t])$= 1-$(\lambda p_i)\Delta t$ + o(Δt)

It remains to prove that the process $\{N_1(t), t \geq 0\}$ and $\{N_2(t), t \geq 0\}$ are independent. Fix $t > 0$. Then by conditioning $P(N_1(t) = k, N_2(t) = m)$

$$= \sum_{\mathrm{n}=0}^{\infty} \mathrm{P}(\, \mathrm{N}_1(\mathrm{t}) = \mathrm{k}, \mathrm{N}_2(\mathrm{t}) = \mathrm{m} | \mathrm{N}(\mathrm{t}) = \mathrm{n})\}\mathrm{P}(\mathrm{N}(\mathrm{t}) = \mathrm{n})$$
$$= P(N_1(t) = k, N_2(t) = m|\, N(t) = k + m)P(N(t) = k + m)$$
$$= \binom{k+m}{k} {p_1}^k {p_2}^m e^{-\lambda t} \frac{(\lambda t)^{k+m}}{(k+m)!}$$
$$= e^{-\lambda p_1 t}\frac{(\lambda p_1 t)^k}{k!} e^{-\lambda p_2 t}\frac{(\lambda p_2 t)^m}{m!},$$

showing that $P(N_1(t) = k, N_2(t) = m) = P(N_1(t) = k)P(N_2(t) = m)$

 https://doi.org/10.26219/heal.aueb.6246

Example 2.1.2 Births occur on a maternity ward according to a Poisson process with an average rate of λ = 24 births/day. Today is the 8th of March. Calculate the following probabilities:

i. P(10 births until today)
ii. P(10 births until tommorow|30 births until today)
iii. P(100 births from March 8 to 11|30 births until tomorrow)
iv. P(15 births today: 9 boys and 6 girls)
v. P(10 births today: boy-boy-boy-boy-boy-boy-girl-boy-girl-girl)
vi. P(5 boys today|30 births from today to tomorrow)

assume that the probability of having a boy is the same as of having a girl.

Solution: Let $\{N(t), t \geq 0\}$ be the Poisson process of births (we measure t in days). Based on theorem 1.1.1 the stochastic process that is counting the newborn boys is Poisson $\{N_1(t), t \geq 0\}$ and same way $\{N_2(t), t \geq 0\}$ is Poisson process counting the newborn girls. $S_1, S_2, \ldots$ the birth times starting from March 8$^{\text{th}}$.

i) $P(N(1) = 10) = e^{-24\cdot 1}\frac{(24\cdot 1)^{10}}{10!}$

ii) $P(N(2) - N(1) = 10|N(1) = 30) = independent\ increments =$
$P(N(2) - N(1)) = 10 = stationary\ increments = P(N(2 - 1) = 10) =$
$P(N(1) = 10) = same\ as\ i)$

iii) $P(N(4) = 100|N(2) - N(1) = 30) =$
$P(N(4) - N(2) + N(1) = 70|N(2) - N(1) = 30) = independent\ increments =$
$P(N(4) - N(2) + N(1) = 70) = stationary\ increments\ = P(N(3) = 70) =$
$e^{-72}\frac{72^{70}}{70!}$

iv) $P(N(1) = 15, N_1(1) = 9, N_2(1) = 6) = P(N_1(1) = 9, N_2(1) = 6) =$
$theorem\ 2.1.5 = P(N_1(1) = 9) \cdot P(N_2(1) = 6) = e^{-12\cdot 1}\frac{(12\cdot 1)^9}{9!} \cdot e^{-12\cdot 1}\frac{(12\cdot 1)^6}{6!}$

v) $P(N(1) = 10, Z_1 = boy, Z_2 = boy, \ldots, Z_8 = boy, Z_9 = girl, Z_{10} = girl =$
$\frac{e^{-24}\cdot 24^{10}}{10!}\left(\frac{1}{2}\right)^{10}$

vi) $P(N_1(1) = 5|N(2) = 30) = \frac{P(N_1(1)=5, N(2)=30)}{P(N(2)=30)} =$

$$= \sum_{k=5}^{30} P(N_1(1) = 5, N(1) = k, N(2) = 30)\ /P(N(2) = 30)$$

$$= \sum_{k=5}^{30} P(N_1(1) = 5, N_2(1) = k-5, N(2) - N(1) = 30-k)/P(N(2) = 30)$$

$$= \sum_{k=5}^{30} P(N_1(1) = 5)P(N_2(1) = k-5)P(N(1) = 30-k)/P(N(2) = 30)$$

$$= \sum_{k=5}^{30} e^{-12}\frac{12^5}{5!}e^{-12}\frac{12^{(k-5)}}{(k-5)!}e^{-24}\frac{24^{(30-k)}}{(30-k)!}/e^{-48}\frac{48^{30}}{30!}$$

$$= \frac{30!}{25!\,5!}\frac{12^5}{48^{30}}\sum_{k=5}^{30}\frac{25!\,12^{k-5}24^{30-k}}{(k-5)!\,(30-k)!}$$

$$= \binom{30}{5}\frac{12^5}{48^{30}}\sum_{k=5}^{30}\binom{25}{k-5}12^{k-5}24^{30-k}$$

$$= \binom{30}{5}\frac{12^5}{48^{30}}\sum_{j=0}^{25}\binom{25}{j}12^{j}24^{25-j} = \binom{30}{5}\frac{12^5 36^{25}}{48^{30}}$$

2.1.4 The Poisson Process and the uniform distribution

The possibility of a Poisson arrival is equally probable in any short interval of the same length. Poisson arrivals occur entirely spontaneously in time. We connect the Poisson process to the uniform distribution to make this point more clear.

Lemma 2.1.6 For any t > 0 and $n = 1,2,\ldots$

$$\mathrm{P}(S_k \le x|N(t) = n) = \sum_{j=k}^{n}\binom{n}{j}\left(\frac{x}{t}\right)^j\left(1-\frac{x}{t}\right)^{n-j} \qquad (2.1.5)$$

for $0 \le x \le t$ and $1 \le k \le n$.

Proof: $P(S_k = x|N(t) = n) = P(N(x) = k|N(t) = n)$

$$= \frac{P(N(x) = k, N(t) = n)}{P(N(t) = n)} = \frac{P(N(x) = k, N(t) - N(x) = n-k)}{P(N(t) = n)}$$

$$= independent\ increments = \frac{P(N(x) = k)P(N(t) - N(s) = n-k)}{P(N(t) = n)}$$

$$= stationary\ increments = \frac{P(N(x) = k)P(N(t-s) = n-k)}{P(N(t) = n)}$$

 https://doi.org/10.26219/heal.aueb.6246

$$= \frac{e^{-\lambda x}\frac{(\lambda x)^k}{k!}e^{-\lambda(t-x)}\frac{\left(\lambda(t-x)\right)^{n-k}}{(n-k)!}}{e^{-\lambda t}\frac{(\lambda t)^n}{n!}} = \binom{n}{k}\left(\frac{x}{t}\right)^k\left(1-\frac{x}{t}\right)^{n-k}$$

$$\text{So}, P(S_k \le x | N(t) = n) = P(N(x) \ge k | N(t) = n) = \sum_{j=k}^{n}\binom{n}{j}\left(\frac{x}{t}\right)^j\left(1-\frac{x}{t}\right)^{n-j}$$

Theorem 2.1.7 $\forall t > 0$ *and* $n = 1,2 \ldots$

$$f_{(S_1,S_2,\ldots S_n|N(t)=n)}(S_1,S_2,\ldots S_n) = f_{U_{1:n},U_{2:n},\ldots U_{n:n}}(S_1,S_2,\ldots S_n)$$

Proof:

$$f_{S_1,S_2,\ldots S_n|N(t)=n)}(S_1,S_2,\ldots S_n) =$$

$$= \lim_{h_1,\ldots h_n \to 0^+}\frac{P(s_1 \le S_1 \le s_1 + h_1, s_2 \le S_2 \le s_2 + h_2, \ldots s_n \le S_n \le s_n + h_n \mid N(t) = n)}{h_1 h_2 \cdots h_n}$$

$$= \lim_{h_1,\ldots h_n \to 0^+}\frac{P(s_1 \le S_1 \le s_1 + h_1, s_2 \le S_2 \le s_2 + h_2, \ldots s_n \le S_n \le s_n + h_n,\ N(t) = n)}{h_1 h_2 \cdots h_n \cdot P(N(t) = n)}$$

$$\lim_{h_1,\ldots h_n \to 0^+}\frac{e^{-\lambda s_1}e^{-\lambda h_1}\lambda h_1 e^{-\lambda(s_2-s_1-h_1)}e^{-\lambda h_2}\lambda h_2 \cdots e^{-\lambda h_n}\lambda h_n e^{-\lambda(t-s_n-h_n)}}{h_1 h_2 \cdots h_n e^{-\lambda t}\frac{(\lambda t)^n}{n!}}$$

$$= \frac{n!}{t^n} = f_{U_{1:n},U_{2:n},\ldots U_{n:n}}(S_1,S_2,\ldots S_n)$$

We illustrate the above by looking at arrivals in very small increments. Given t and $0 < s_1 < \cdot\cdot\cdot, < s_n < t$ we calculated the probability that there is a single arrival in each of the intervals $(s_i, s_i + h_i]$, $1 \le i \le n$ and no arrivals otherwise. $(s_0 = 0)$

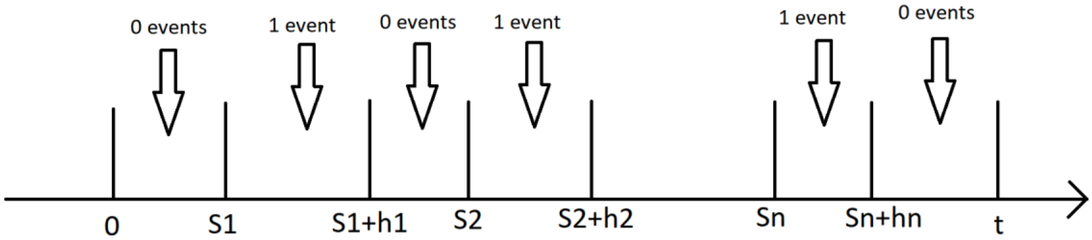

figure 2.2 illustration of the arrival process

Exercise 2.1.1 $\{N(t), t \ge 0\}$ Poisson processrate λ and $S_1, S_2, \ldots S_n$ the arrival times. Calculate $E(S_1|N(t) \ge 1)$.

https://doi.org/10.26219/heal.aueb.6246

Solution 1:

$$E(S_1|N(t) \geq 1) = \sum_{n=1}^{\infty} E(S_1|N(t) = n, N(t) \geq 1)P(N(t) = n|N(t) \geq 1)$$

$$= \sum_{n=1}^{\infty} \frac{E(S_1|N(t) = n)P(N(t) = n)}{P(N(t) \geq 1)} = theorem\ 2.1.7 = \sum_{n=1}^{\infty} \frac{\frac{1 \cdot t}{n+1} e^{-\lambda t} \frac{(\lambda t)^n}{n!}}{1 - e^{-\lambda t}}$$

$$= \frac{\frac{1}{\lambda} e^{-\lambda t} \sum_{n=1}^{\infty} \frac{(\lambda t)^{n+1}}{(n+1)!}}{1 - e^{-\lambda t}} = \frac{\frac{1}{\lambda} e^{-\lambda t}\left(e^{-\lambda t} - \lambda t - 1\right)}{1 - e^{-\lambda t}}$$

Solution 2: $E(S_1) = P(N(t) \geq 1)E(S_1|N(t) \geq 1) + P(N(t) = 0)E(S_1|N(t) = 0)$

$$\Rightarrow \frac{1}{\lambda} = 1 - e^{-\lambda t} E(S_1|N(t) \geq 1) + e^{-\lambda t}\left(t + \frac{1}{\lambda}\right)$$

$$\Rightarrow E(S_1|N(t) \geq 1) = \frac{\frac{1}{\lambda} e^{-\lambda t}\left(e^{-\lambda t} - \lambda t - 1\right)}{1 - e^{-\lambda t}}.$$

2.1.5 The M/G/∞/FCFS queue (Kendall's notation)

In this section we study the M/G/∞ queue. It assumes that at a service station a flow of customers arrives according to a Poisson process, let's say with rate λ. The service times of the customers are independent random variables having a common probability distribution with mean $\mu < \infty$. We denote the distribution function of the service time of a customer by B(t). When a customer enters the service station, he/she immediately starts to be served because the number of servers is assumed to be infinite, that is the third symbol in the name of the model. The assumption that the number of server is infinite makes this system easy to study in some of its characteristics as the processing of any customer in the server station does not affect the processing of other customers already present or arriving later to the station. In implementations, this flexible model is very helpful. A significant question is: what is the limiting-distribution of the number of customers in the system? The number of customers is always the same as the number of busy servers because the service facility has an ample number of servers, so every customer is always being served by exactly one (busy) server. The remarkably straightforward answer to this question is that the limiting distribution is a Poisson distribution with mean $\lambda\mu$.

 https://doi.org/10.26219/heal.aueb.6246

$$\lim_{t\to\infty} P(k \text{ servers are busy at time } t) = e^{-\lambda\mu}\frac{(\lambda\mu)^k}{k!} \quad \text{for } k = 0,1,2\ldots \tag{2.1.6}$$

The form of the above distribution does not depend on the "shape" of the distribution of the service times of the customers but it only depends on the mean μ. This result is extremely helpful for applications. The M/G/∞ has a variety of implementations. One of them is the (S - 1, S) inventory policy with back ordering. In this model, customers ask for a specific product and they arrive according to a Poisson process with rate λ. Each customer asks for one of the product units. Initially there is a number of s units in stock. Each time a customer demand occurs, a replenishment order is placed for exactly one unit of the product. A customer demand that occur when the on hand inventory is zero also triggers a replenishment order and the demand is back ordered until a unit becomes available to satisfy the demand. The lead times of the replenishment orders are independent random variables each having the same probability distribution with mean τ. Some reflections indicate that this inventory policy (S - 1, S) can be "converted" into the queueing model M/G/∞ like this: Identify the outstanding replenishment orders with customers in service and identify the lead times of the replenishment orders with the service times. So the limiting distribution of the number of outstanding replenishment orders is a Poisson distribution with mean $\lambda\tau$.

$$\text{the long run } (\mathrm{t} \to \infty) \text{ average on hand inventory } = \sum_{k=0}^{S}(S-k)e^{-\lambda\tau}\frac{(\lambda\tau)^k}{k!}.$$

Returning to the M/G/∞ queue, we give two proofs for the equation 2.1.6

Proof 1: The method of proof is actually based on partitioning the Poisson random variable N(t) into two types: those that are in service at time t, X(t) and those that have departed by time t (denoted by D(t)). Thus we need only figure out what is the probability pr(t) that a customer who arrived during (0, t] (that is, one of the N(t) arrivals) is still in service at time t. We first recall that conditional on N(t) $=$ n the n (unordered) arrival times are i.i.d. with a

Unif(0, t) distribution (theorem 2.1.7). Letting U denote a typical such arrival time, and C their service time, we conclude that this customer will still be in service at time t if and only if $U + C > t$ (arrival time + service time > t);$\Rightarrow$

$C > t - U$. Thus $pr(t) = P(C > t - U)$ where C and U are assumed independent. But (as is easily shown)) $t - U \sim \text{Unif}(0, t)$ if $U \sim \text{Unif}(0, t)$. Thus $pr = P(C > t - U) = P(C > U)$. Noting that $P(C > U|U = x) = P(C > x)$ we conclude that

$$pr(t) = P(C > U) = \frac{1}{t}\int_0^t P(C > x)dx$$

where we have conditioned on the value $U = x$ (which density 1/t) and integrated over all such values. This did not depend upon the value n and so we are done. Thus for fixed t, we can partition $N(t)$ into two independent Poisson processes, $X(t)$ and $D(t)$ (the number of departures by t), to conclude that $X(t) \sim Poisson(a(t))$ where

$$a(t) = \lambda t pr(t) = \lambda \int_0^t P(C > x)dx$$

Similarly $D(t) \sim \text{Poisson}(\beta(t))$ where $\beta(t) = \lambda t\big(1 - pr(t)\big)$.

Finally we have

$$P(X(t) = k) = e^{-a(t)} \frac{(a(t))^k}{k!}, k \geq 0 \qquad (2.1.7)$$

But $\lim_{t\to\infty} a(t) = \lambda\mu$ so $p_k \overset{\text{def}}{=} \lim_{t\to\infty} P(X(t) = k) = e^{-\lambda\mu} \frac{(\lambda\mu)^k}{k!}$ (1.1.6)■

The fact that $\lim_{t\to\infty} a(t) = \lambda\mu$ is obtained from the following theorem.

Theorem 2.1.8 Let T be a positive random variable and $S(t) = P(T \geq t)$ then $E(T) = \int_0^\infty S(t)dt$ (This theorem is very usefull in survival analysis)

Proof: Consider any $n > 0$,

$$\int_{t=0}^{n} t\, f_T(t)dt = \int_{t=0}^{n}\left(\int_{s=0}^{t} ds\right) f_T(t)dt = \int_{t=0}^{n}\int_{s=0}^{t} f_{T(t)}dsdt = \int_{s=0}^{n}\int_{t=s}^{n} f_T(t)dtds$$

$= \int_{s=0}^{n} F_T(n) - F_T(s)ds$ then as $n \to \infty, F_T(n) \to 1$ and we obtain

$E(T) = \int_{s=0}^{\infty} 1 - F_T(s)ds = \int_{s=0}^{\infty} S_T(s)ds$ ■

A comment about the proof 1 of (2.1.6): We have to keep in mind that (2.1.6) implies convergence in a time average sense also:

 https://doi.org/10.26219/heal.aueb.6246

$$p_k = \lim_{t\to\infty}\int_0^t P(X(s)=k)ds = e^{-\lambda\mu}\frac{(\lambda\mu)^k}{k!}$$

which is exactly the continuous time analog of the stationary distribution π for Markov chains:

$$\pi_k = \lim_{n\to\infty}\frac{1}{n}\sum_{j=1}^{n} P\left(X_j = k\right)$$

Thus we interpret p_j as the long run proportion of time that there are k busy servers. The average number of busy servers is given by the mean of the limiting distribution:

$$L = \sum_{k=0}^{\infty} k\, p_k = \lambda\mu$$

This result agrees with “Little’s Law”. We study “Little’s Law” in detail in the next chapter.

Proof 2 of (2.1.6): The differential equation approach can be used to give a rigorous proof of (2.1.6). Assuming that there are no customers present at epoch 0, define for any $t > 0$

$$p_j(t) = P(\text{there are } j \text{ busy servers at time } t),\ \ j = 0,1,2\ldots$$

Consider now $p_j(t+\Delta t)$ for Δt small. The event that there are j servers busy at time $t+\Delta t$ can occur in the following mutually exclusive ways:

a) no arrival occurs in (0, Δt) and there are j busy servers at time $t+\Delta t$ due to arrivals in (Δt, t + Δt)

b) one arrival occurs in (0, Δt), the service of the first arrival is completed before time $t+\Delta t$ and there are j busy servers at time $t+\Delta t$ due to arrivals in (Δt, t +Δt)

c) one arrival occurs in (0, Δt), the service of the first arrival is not completed before time $t+\Delta t$ and there are $j-1$ other busy servers at time $t+\Delta t$ due to arrivals in (Δt, t + Δt)

d) two or more arrivals occur in (0, Δt) and j servers are busy at time $t+\Delta t$.

Let $B(t)$ denote the probability distribution (as already mentioned) of the service time of a customer. Then, since a probability distribution function has

https://doi.org/10.26219/heal.aueb.6246

at most a countable number of discontinuity points, we find for “almost” all $t > 0$ that

$$p_j(t+\Delta t) = (1-\lambda\Delta t)p_j(t) + \lambda\Delta t B(t+\Delta t)p_j(t) + \lambda\Delta t\{1-B(t+\Delta t)\}p_{j-1}(t) + o(\Delta t)$$

Subtracting $p_j(t)$ $from\, p_j(t+\Delta t), dividing\ by\ \Delta t\ and\ letting\ \Delta t \to 0,$ we find that:

$$p'_0(t) = -\lambda\big(1-B(t)\big)p_0(t)$$

$$p'_j(t) = -\lambda\big(1-B(t)\big)p_j(t) + \lambda\big(1-B(t)\big)p_{j-1}(t), \quad j = 1,2,3,\dots$$

Next by induction on j, it is readily verified that

$$p_j(t) = e^{-\lambda\int_0^t(1-B(x))\,dx}\frac{\left[\lambda\int_0^t\big(1-B(x)\big)dx\right]^j}{j!}, \;\; j = 0,1,2\dots$$

By a continuity argument this relation holds for all $t \geq 0$.

Since $\int_0^\infty(1-B(x))\,dx = \mu$ (by theorem 2.1.7) the result (2.1.6) follows.

The $M/G/\infty$ queue has many different applications in different fields.

Example 1.1.2 Let’s say we have a computer-server and clients “arrive” according to a Poisson process with rate $\lambda = 500\text{ per hour}$. The clients stay connected to the server in average for $\mu = 180\text{ minutes} = 3\text{ hours}$. It is pretended that clients are infinite. Then the expected number of simultaneous clients L is

$$L = \sum_{k=0}^{\infty} k\,p_k = \lambda\mu = 500\cdot 3 = 1500$$

(which is based on the comment on the proof 1 of (2.1.6))

To ensure that the server will stay up for at least 95% of the time, the capacity c of the server must satisfy:

$$\sum_{n=0}^{c} P_n \geq 0.95 \overset{(2.1.6)}{=} \sum_{n=0}^{c}\frac{e^{-1500}(1500)^n}{n!} \geq 0.95$$

$c = 1564$ is the minimum number that satisfies the above relation.

https://doi.org/10.26219/heal.aueb.6246

2.2 COMPOUND POISSON PROCESSES

In a compound Poisson process, each arrival in an ordinary Poisson process comes with an associated real-valued random variable that represents the value of the arrival in a sense. These variables are independent and identically distributed, and are independent of the underlying Poisson process. Our interest centers on the sum of the random variables for all the arrivals up to a fixed time. This sum is a Poisson-distributed random sum of random variables. Distributions of this type are said to be compound Poisson distributions, and are important in their own right, particularly since some surprising parametric distributions turn out to be compound Poisson.

Definition 2.2.1 A stochastic process $\{X(t), t \geq 0\}$ is said to be a compound Poisson process if it can be written as:

$$X(t) = \sum_{i=1}^{N(t)} D_i, \;\; t \geq 0$$

where $\{N(t), t \geq 0\}$ is a Poisson process with rate λ, and $D_1, D_{2,\ldots}$ are independent and identically distributed non-negative random variables that are also independent of the process $\{N(t), t \geq 0\}$.

Suppose now that each arrival has an associated real-valued random variable that represents the *value* of the arrival in a certain sense. Here are some typical examples:

- The arrivals are customers at a store. Each customer spends a random amount of money.
- The arrivals are visits to a website. Each visitor spends a random amount of time at the site.
- The arrivals are failure times of a complex system. Each failure requires a random repair time.
- The arrivals are earthquakes at a particular location. Each earthquake has a random severity, a measure of the energy released.

In all cases we assumed that the sizes of the arrivals D_i are independent and identically distributed random variables, which are also independent of the arrival process $\{N(t), t \geq 0\}$.

https://doi.org/10.26219/heal.aueb.6246

Lemma 2.2.1 Let $D_1, D_2, \ldots$ be a sequence of independent and identically distributed random variables whose first two moments are finite. Also, let N be a non-negative and integer-valued random variable having finite first two moments. If the random variable N is independent of the random variables $D_1, D_{2,\ldots}$ then

$$E\left(\sum_{k=1}^{N} D_k\right) = E(N)E(D_1) \tag{2.2.1}$$

$$var\left(\sum_{k=1}^{N} D_k\right) = E(N)var(D_1) + var(N)E^2(D_1) \tag{2.2.2}$$

Proof: The proof uses the law of total expectation. By conditioning on N, we find

$$E\left(\sum_{k=1}^{N} D_k\right) = \sum_{n=0}^{\infty} E\left(\sum_{k=1}^{N} D_k \mid N = n\right) P(n = n)$$
$$= \sum_{n=0}^{\infty} E\left(\sum_{k=1}^{n} D_k\right) P(N = n) = \sum_{n=0}^{\infty} nE(D_1)P(N = n)$$

which verifies (2.2.1). Note that the second equality uses that the random variables $D_1, D_{2,\ldots}$ are independent of the event $\{N = n\}$. Similarly,

$$E\left[\left(\sum_{k=1}^{N} D_k\right)^2\right] = \sum_{n=0}^{\infty} E\left[\left(\sum_{k=1}^{N} D_k\right)^2 \mid N = n\right] P(N = n)$$
$$= \sum_{n=0}^{\infty} [nE(D_1^2) + n(n-1)E^2(D_1)]P(N = n)$$
$$= E(N)E(D_1^2) + E(N(N-1))E^2(D_1) \tag{2.2.3}$$

Using $\sigma^2(D_1) = E(D_1) - E^2(D_1)$, we obtain (2.2.2) from (2.2.1) and (2.2.3) ■

From the above lemma we also deduce that the mean and variance of the compound Poisson variable X(t) are given by

$$\mathrm{E}(\mathrm{X}(\mathrm{t})) = \lambda \mathrm{tE}(\mathrm{D}_1) \quad \text{and} \quad \sigma^2(\mathrm{X}(\mathrm{t})) = \lambda \mathrm{tE}(\mathrm{D}_1^2), \quad \mathrm{t} \geq 0 \tag{2.2.4}$$

Discrete compound Poisson distribution

We follow Tijms (2009, p.19)

 https://doi.org/10.26219/heal.aueb.6246

Consider first the case of discrete random variables $D_1, D_2, \ldots$:

$$\alpha_j = P(D_1 = j), \quad j = 0,1,2,\ldots$$

Then a simple algorithm can be given to compute the probability distribution of the compound Poisson variable $X(t)$. For $t \geq 0$, we define

$$r_j(t) = P(X(t) = j), j = 0,1,2 \ldots$$

Define the generating function $A(z)$ by

$$A(z) = \sum_{j=0}^{\infty} a_j z^j, \quad |z| \leq 1$$

Also, for any fixed $t > 0$, define the generating function $R(z,t)$ as

$$R(z,t) = \sum_{j=0}^{\infty} r_j(t) z^j, \quad |z| \leq 1.$$

Theorem 2.2.2 For any fixed t > 0 it holds that:

a) the generating function $R(z,t)$ is given by

$$R(z,t) = e^{-\lambda t(1-A(z))}, \quad |z| \leq 1 \qquad (2.2.5)$$

b) the probabilities $\{r_j(t), j = 0,1, \ldots\}$ satisfy the recursion

$$r_j(t) = \frac{\lambda t}{j} \sum_{k=0}^{j-1} (j-k)\, a_{j-k} r_k(t), \quad j = 1,2,\ldots \qquad (2.2.6)$$

starting with $r_0(t) = e^{-\lambda t(1-a_0)}$.

Proof: Fix $t \geq 0$. By conditioning on the number of arrivals up to time t,

$$r_j(t) = \sum_{n=0}^{\infty} P(X(t) = j | N(t) = n) P(N(t) = n)$$

$$= \sum_{n=0}^{\infty} P(D_0 + \ldots + D_n = j) e^{-\lambda t} \frac{(\lambda t)^n}{n!}, \quad j = 0,1\ldots$$

with $D_0 = 0$. This gives, after an interchange of the order of summation,

$$\sum_{j=0}^{\infty} r_j(t)\, z^j = \sum_{n=0}^{\infty} e^{-\lambda t} \frac{(\lambda t)^n}{n!} \sum_{j=0}^{\infty} P(D_0 + \cdots + D_n = j) z^j.$$

https://doi.org/10.26219/heal.aueb.6246

Since the D_i are independent of each other, it follows that

$$\sum_{j=0}^{\infty} P\{D_0 + \ldots D_n = j\} z^j = E(z^{D_0+\ldots+D_n}) = E(z^{D_0}) \cdots E(z^{D_n}) = [A(z)]^n.$$

Therefore,

$$R(z,t) = \sum_{n=0}^{\infty} e^{-\lambda t} \frac{(\lambda t)^n}{n!} [A(z)]^n = e^{-\lambda t[1-A(z)]}$$

which proves (2.2.5). To prove part b) for fixed t, we write $R(z) = R(z,t)$ for ease of notation. It follows immediately from the definition of the generating function that the probability $r_j(t)$ is given by

$$r_j(t) = \frac{1}{j!} \frac{d^j R(z)}{dz^j} \bigg|_{z=0}$$

It is not possible to obtain (2.2.6) directly from this relation and (2.2.5). The following intermediate step is needed. By differentiation of (2.2.5) with respect to z, we find

$$R'(z) = \lambda t A'(z) R(z), |z| \leq 1.$$

This gives

$$\sum_{j=1}^{\infty} j r_j(t) z^{j-1} = \lambda t \left[\sum_{k=1}^{\infty} k a_k z^{k-1}\right] \left[\sum_{l=0}^{\infty} r_l(t) z^l\right]$$

$$= \sum_{k=1}^{\infty} \sum_{l=0}^{\infty} \lambda t k a_k r_l(t) z^{k+l-1}.$$

Substituting $k+1$ by j and interchanging the order of summation yields

$$\sum_{j=1}^{\infty} j\, r_j(t) z^{j-1} = \sum_{k=1}^{\infty} \sum_{j=k}^{\infty} \lambda t k a_k r_{j-k}(t) z^{j-1}$$

$$= \sum_{j=1}^{\infty} \left[\sum_{k=1}^{j} \lambda t k a_k r_{j-k}(t)\right] z^{j-1}.$$

Next equating coefficients gives the recurrence relation (2.2.6).
The recursion scheme for the $r_j(t)$ is easy to program and is numerically stable. It is often called Adelson's recursion scheme after Adelson (1966). In the

https://doi.org/10.26219/heal.aueb.6246

insurance literature the recursive scheme is known as Panjer's algorithm. Note that for the special case of $a_1 = 1$ the recursion (2.2.6) reduces to the known recursion scheme for computing Poisson probabilities. An alternative method to compute the compound Poisson probabilities $r_j(t), j = 0,1, \ldots$ is to apply the discrete fast Fourier transform method to the explicit expression (2.2.5) for the generating function of the $r_j(t)$.

2.3 NON-STATIONARY POISSON PROCESSES

A non-homogeneous Poisson process is similar to an ordinary Poisson process, except that the average rate of arrivals is allowed to vary with time. Many applications that generate random points in time are modeled more faithfully with such non-homogeneous processes. The mathematical cost of this generalization, however, is that we lose the property of stationary increments.

Definition 2.3.1 A counting process $\{N(t), t \geq 0\}$ is said to be a non-stationary Poisson process with intensity function $\lambda(t)$, $t \geq 0$, if it satisfies the following properties:

a) $N(0) = 0$

b) the process $\{N(t), t \geq 0\}$ has independent increments

c) $P(N(t + \Delta t) - N(t) = 1) = \lambda(t)\Delta t + o(\Delta t)$ as $\Delta t \to 0$

d) $P(N(t + \Delta t) - N(t) \geq 2) = o(\Delta t)$ as $\Delta t \to 0$.

The theorem below states that the total number of arrivals in a given time interval follows the Poisson distribution.

Theorem 2.3.1 For any $t, s \geq 0$,

$$P(N(t+s) - N(t) = k) = e^{-[M(t+s)-M(t)]} \frac{[M(t+s) - M(t)]^k}{k!}, \quad (2.3.1)$$

for $k = 0,1, \ldots where\ M(x) = \int_0^x \lambda(y) dy, x \geq 0$.

Proof: The proof is instructive. Fix $t \geq 0$. Put for abbreviation

$$p_k(s) = P(N(t+s) - N(t) = k), \qquad k = 0,1, \ldots$$

Consider now $p_k(s + \Delta s)$ for Δs small. Since the probability of two or more arrivals in a small time interval of length Δs is negligibly small compared with

Δs as $\Delta s \to 0$, it follows that the only possibility for the process to be in state k at time $t + s + \Delta s$ is that the process is either in state k – 1 or in state k at time $t + s$. Hence by conditioning on the state of the process at time $t + s$ and given that the process has independent increments,

$$p_k(s + \Delta s) = p_{k-1}(s)[\lambda(t + s)\Delta s + o(\Delta s)] + p_k(s)[1 - \lambda(t + s)\Delta s + o(\Delta s)]$$

as $\Delta s \to 0$. Substracting $p_k(s)$ from both sides of this equation and dividing by Δs, we have

$$p'_k(s) = -\lambda(t + s)[p_k(s) - p_{k-1}(s)], \;\; k = 1,2, \dots$$

For $k = 0$, we have $p'_0(s) = -(t + s)p_0(s)$. The initial conditions $p_0(0) = 1$ and $p_k(0) = 0$ for $k \geq 1$ hold. From the theory of differential equations we know that the solution of the first-order differential equation

$$y'(s) + a(s)y(s) = b(s), \;\; s \geq 0$$

is given by

$$y(s) = e^{-A(s)} \int_0^s b(x)\, e^{-A(x)} dx + ce^{-A(s)}$$

for some constant c, where $A(s) = \int_0^s a(x)dx$. The constant c is determined by a boundary condition on $y(0)$. This gives after some algebraic manipulations

$$p_0(s) = e^{-[M(s+t)-M(t)]}, \;\; s \geq 0.$$

By induction the expression for $p_k(s)$ next follows from

$p'_k(s) + \lambda(t + s)p_k(s) = \lambda(t + s)p_{k-1}(s)$. We do not give the details.

We point out that $M(t)$ represents the expected number of arrivals up to time t.

2.4 MARKOV MODULATED BATCH POISSON PROCESSES

We follow section 1.4 of the book of Tijms (2003)

The Markov modulated batch Poisson process is a compound Poisson process that has its parameter controlled by a Markov process (so-call phase process) independently of the arrivals. These arrival processes are typical in communications modeling where time-varying arrival rates capture some of the important correlations between inter-arrival times. The phase process can only assume a finite number of states $i = 1,2, \dots m$. The sojourn time of the phase

 https://doi.org/10.26219/heal.aueb.6246

process in state i is exponentially distributed with mean $1/\omega_i$. If the phase process leaves state i, it goes to state j with probability p_{ij}, independently of the duration of the stay in state i. It is assumed that $p_{ii} = 0$ for all i. The arrival process of customers is a compound Poisson process whose parameters depend on the state of the phase process. If the phase process is in state i, then batches of customers arrive according to a Poisson process with rate λ_i where the batch size has the discrete probability distribution $\left\{a_k^{(i)}, k = 1,2,\dots\right\}$. It is no restriction to assume that $a_0^{(i)=0}$; otherwise replace λ_i by $\lambda_i\left(1 - a_0^{(i)}\right)$ and $a_k^{(i)}$ by $a_k^{(i)}/\left(1 - a_0^{(i)}\right)$ for $k \geq 1$.

For any $t \geq 0$ and $i,\ j = 1,2,\dots,m$ define

$P_{ij}(k,t) = P(\textit{the total number of customers arriving in } (0,t) \textit{ equals } k \textit{ and the phase process is in state } j \textit{ at time } t \mid \textit{the phase process is in state } i \textit{ at the present time } 0), k = 0,1,\dots$

Also, for any $t > 0$ and $i, j = 1,2,\dots,m$ we define the generating function

$$P_{ij}^{*}(z,t) = \sum_{k=0}^{\infty} P_{ij}(k,t) z^k, \quad |z| \leq 1.$$

To derive an expression for $P_{ij}^{*(z,t)}$ it is helpful to use matrix notation. Let $Q = \left(q_{ij}\right)$ be the $m \times m$ matrix whose (i, j)th element is given by

$$q_{ii} = -\omega_i \quad \text{and} \quad q_{ij} = \omega_i p_{ij} \quad \text{for} \quad j \neq i.$$

Define the $m \times m$ diagonal matrices Λ and A_k by

$$\Lambda = diag(\lambda_1,\dots,\lambda_m) \quad \text{and} \quad A_k = diag(a_k^{(1)},\dots,a_k^{(m)}), \quad k = 1,2,\dots \qquad (2.4.1)$$

Let the $m \times m$ matrix D_k for $k = 0,1,\dots$ be defined by

$$\mathrm{D_0 = Q - \Lambda \quad and\ D_k = A_k\Lambda, \quad k = 1,2,\dots} \qquad (2.4.2)$$

Using $(D_k)_{ij}$ to denote the (i, j)th element of the matrix D_k, generating function of $(D_k)_{ij}$ is

$$D_{ij}(z) = \sum_{k=0}^{\infty} (D_k)_{ij} z^k, \qquad |z| \leq 1.$$

https://doi.org/10.26219/heal.aueb.6246

Theorem 2.4.1 Let $P^*(z,t)$ and $D(z)$ denote the $m \times m$ matrices whose (i, j)th elements are given by the generating functions $P_{ij}^*(z,t)$ and $D_{ij}(z)$. Then, for any $t > 0$,

$$P^*(z,t) = e^{D(z)t}, \qquad |z| \leq 1 \tag{2.4.3}$$

where e^{At} is defined by $e^{At} = \sum_{n=0}^{\infty} A^n t^n / n!$

Proof: The proof is based on deriving a system of differential equations for the $P_{ij}(k,t)$. Fix i, j, k and t. Consider $P_{ij}(k, t + \Delta t)$ for Δt small. By conditioning on what may happen in $(t,\ t + \Delta t)$, it follows that

$$P_{ij}(k,t+\Delta t) = P_{ij}(k,t)\big(1-\lambda_j \Delta t\big)\big(1-\omega_j \Delta t\big) + \sum_{s \neq j} P_{is}(k,t)\big[(\omega_s \Delta t) \times p_{sj}\big]$$

$$+ \sum_{l=0}^{k-1} P_{ij}(l,t)\left[\big(\lambda_j \Delta t\big) \times a_{k-l}^{(j)}\right] + o(\Delta t)$$

Using the definition of the q_{ij}, the above equation is rewritten as

$$P_{ij}(t,\Delta t) = P_{ij}(k,t)\big(1-\lambda_j \Delta t\big) + \sum_{s=1}^{m} P_{is}(k,t) q_{sj} \Delta t$$

$$+ \sum_{l=0}^{k-1} P_{ij}(l,t) \lambda_j a_{k-l}^{(j)} \Delta t + o(\Delta t)$$

From the above equation we deduce that

$$\frac{d}{dt} P_{ij}(k,t) = -\lambda_j P_{ij}(k,t) + \sum_{s=1}^{m} P_{is}(k,t) q_{sj} + \lambda_j \sum_{l=0}^{k-1} P_{ij}(l,t) a_{k-l}^{(j)}$$

Letting $P(k,t)$ be the $m \times m$ matrix whose (i,j)th element is $P_{ij}(k,t)$, we have in matrix notation that

$$\frac{d}{dt} P(k,t) = P(k,t)(Q - \Lambda) + \sum_{l=0}^{k-1} P(l,t) A_{k-l} \Lambda.$$

Using the definition of the matrices D_k, we deduce that

$$\frac{d}{dt} P(k,t) = P(k,t) D_0 + \sum_{l=0}^{k-1} P(l,t) D_{k-l}$$

 https://doi.org/10.26219/heal.aueb.6246

$$= \sum_{l=0}^{k} P(l,t) D_{k-l}$$

Multiply both sides of this matrix equation by z^k and sum over k. Since the generating function of the convolution of two sequences is the

product of the generating functions of the two sequences, it follows that

$$\frac{d}{dt} P^*(z,t) = P^*(z,t) D(z).$$

For each fixed i this equation gives a system of linear differential equations in $P_{ij}^*(z,t)$ for $j = 1, \dots, m$. Thus, using a known result from the theory of linear differential equations, we have that

$$P_i^*(z,t) = e^{D(z)t} P_i^*(z,0) \tag{2.4.4}$$

where $P_i^*(z,t)$ is the i-th row of the matrix $P^*(z,t)$. Since $P_i^*(z,0)$ is equal to the i-th unit vector $e_i = (0, \dots, 1, \dots, 0)$, it follows that $P^*(z,t) = e^{D(z)t}$. The proof is complete. In general it is an impressive task to obtain the numerical values of the probabilities $P_{ij}(k,t)$ from the expression (1.4.4), particularly when m is large. The numerical approach of the discrete fast Fourier transform method is only practically possible when the computation of the matrix $e^{D(z)t}$ is not too complicated. A great number of algorithms for the computation of the matrix exponential e^{-At} have been proposed, but they do not always provide high accuracy. The computational work becomes simple when the $m \times m$ matrix A has m different eigenvalues let's say $\mu_1, \dots, \mu_m$, as it usually happens in applications. It is well known from linear algebra that the matrix A can then be diagonalized as

$$A = S \chi S^{-1},$$

where the diagonal matrix χ is given by $\chi = \mathrm{diag}(\mu_1, \dots, \mu_m)$ and the column vectors of the matrix S are the linearly independent eigenvectors associated with the eigenvalues $\mu_1, \dots, \mu_m$. Moreover, by $A^n = S \chi^n S^{-1}$, it holds that

$$e^{At} = S \cdot \mathrm{diag}(e^{-\mu_1 t}, \dots, e^{-\mu_m t}) S^{-1}.$$

Fast codes for the computation of eigenvalues and eigenvectors of a matrix are widely available.

https://doi.org/10.26219/heal.aueb.6246

To conclude this section, it is remarked that the matrix $D(z)$ in the matrix exponential $e^{-D(z)t}$ has a very simple form for the important case of single arrivals. It then follows from (1.4.1) and (1.4.2) that

$$D(z) = Q - \Lambda + \Lambda z, \qquad |z| \leq 1.$$

The arrival process with single arrivals is called the Markov modulated Poisson process. A special case of this process is the switched Poisson process which has only two arrival rates (m=2). This model is frequently used in applications. In the special case of the switched Poisson process, the following explicit expressions can be given for the generating functions $P_{ij}^*(z,t)$:

$$P_{ii}^*(z,t) = \frac{1}{r_2(z) - r_1(z)}\Big[\{r_2(z) - (\lambda_i(1-z) + \omega_i)\}e^{-r_1(z)t}$$

$$- \{r_1(z) - (\lambda_i(1-z) + \omega_i)\}e^{-r_2(z)t}\Big], \quad i = 1,2$$

$$P_{12}^*(z,t) = \omega_1 \frac{e^{-r_1(z)t} - e^{-r_2(z)t}}{r_2(z) - r_1(z)} \quad and \quad P_{21}^*(z,t) = \omega_2 \frac{e^{-r_1(z)t} - e^{-r_2(z)t}}{r_2(z) - r_1(z)},$$

where

$$r_{1,2}(z) = \frac{1}{2}(\lambda_1(1-z) + \omega_1 + \lambda_2(1-z) + \omega_2)$$

$$\pm \frac{1}{2}[\{\lambda_1(1-z) + \omega_1 + \lambda_2(1-z) + \omega_2\}^2$$

$$-4\{(\lambda_1(1-z) + \omega_1)(\lambda_2(1-z) + \omega_2) - \omega_1\omega_2\}]^{1/2}.$$

It is a matter of straightforward but complicated algebra to derive these expressions. The probabilities $P_{ij}(k,t)$ can be easily computed from these expressions by applying the discrete fast Fourier transform method.

https://doi.org/10.26219/heal.aueb.6246

https://doi.org/10.26219/heal.aueb.6246

CHAPTER 3

Renewal-Reward Processes

The content of this chapter is mainly based on chapter 2 of the book of Tijms (2003).

A renewal process is an idealized stochastic model for events that occur randomly in time (generically called renewals or arrivals). The basic mathematical assumption is that the times between the successive arrivals are independent and identically distributed. Renewal processes are regenerative; that is, from a probabilistic point of view, the process restarts itself whenever an arrival occurs. Renewal reward processes are renewal processes where each interarrival time is associated with a random variable that is generically thought of as the reward associated with that interarrival time. Our interest is in the process that gives the total reward up to time t. Renewal processes have a very rich and interesting mathematical structure and can be used as a foundation for building of more realistic models. Moreover, renewal processes are often found embedded in other stochastic processes, especially in notably Markov chains. Here are some typical interpretations and applications:

- The arrivals are customers arriving at a service station. A customer might be a person and the service station a store, but also a customer might be a file request and the service station a web server.
- A device is placed in service and eventually fails. It is replaced by a device of the same type and the process is repeated. We do not count the replacement time in our analysis; equivalently we can assume that the replacement is immediate. The times of the replacements are the renewals.
- The arrivals are times of some natural event, such as a lightening strike, a tornado or an earthquake, at a particular geographical point.
- The arrivals are emissions of elementary particles from a radioactive source.

 https://doi.org/10.26219/heal.aueb.6246

3.1 RENEWAL THEORY

Renewal processes are models of stochastic phenomena in which an event (or combination of events) occurs repeatedly over time. In chapter 2 a counting process was considered and the times between occurrences were exponential distributed. Renewal processes generalize this idea so that the times between occurrences are i.i.d. Thus the stringent distributional assumptions that were needed in chapter 2 do not hold and the memoryless property is also removed. The memoryless property was so essential in almost every theorem we used and it was also important even in the way we solved several problems. Since we removed that property, if we study the same things in the same way as in chapter 2 the mathematics will be harder to understand and to be proved. Renewal processes are very useful because they have a lot of applications in real world problems and lead to important generalizations such as renewal-reward processes. Renewal-reward processes are very useful in the computation of important performance measures such as long run cost and reward rates, and we discuss this topic in paragraph 3.2. In this section we begin the study or renewal processes.

Let $(X_n,\ n \geq 1)$ be a sequence of nonnegative, independent identically distributed (i.i.d.) random variables (r.v.'s) with distribution function (d.f.)

$$F(x) = P\{X_k \leq x\}, \qquad x \geq 0, k \in N$$

To avoid trivialities it is assumed that: $0 < \mu_1 < \infty$, where $\mu_1 = E(X_k)$.
We think of r.v. X_n as the time between the $(n-1)$th and nth event.
Since we do **not** allow $X_n = 0$, $\mathrm{n} \in N$, multiple events cannot occur simultaneously. Next define

$$S_0 = 0 \text{ and } S_n = \sum_{i=1}^{n} X_i, \quad n = 1,2,\ldots$$

The quantities $\mathrm{S}_0, \mathrm{S}_1, \mathrm{S}_2, \ldots$ are thought of as times of occurrence of some specific probability problem. For each $t \geq 0$, let

$$\mathrm{N}(t) = max\{n \geq 0 \colon S_n \leq t\}$$

Thus $N(t)$ is simply the number of events up to time t.

https://doi.org/10.26219/heal.aueb.6246

Definition 3.1.1 The counting process $\{N(t), t \geq 0\}$ is called the renewal process generated by the interoccurrence times $X_1, X_2, \ldots$

We say that the nth renewal/event occurs at time t and we denote $S_n = t$. Since the interoccurrence times are independent and identically distributed, it follows that the process probabilistically begins over at each renewal.

Theorem 3.1.1

$$P(N(t) < \infty) = 1, \quad t \geq 0$$

Proof: From the strong law of large numbers we have

$$\mathrm{P}\left(\frac{S_n}{n} \to \mu_1\right) = 1.$$

Since $\mu_1 > 0$, this implies that:

$$P(S_n \to \infty) = 1.$$

Thus, for $0 \leq t < \infty$

$$P(N(t) = \infty) = P(\lim_{n \to \infty} S_n \leq t) = 1 - P(S_n \to \infty) = 0.$$

which means that the number of events up to time t is finite with probability 1.

Remark 3.1.1 $\{N(t), t \geq 0\}$ and $\{S_n,\ n \geq 0\}$ mutually and univocally determine each other because for arbitrary $t \geq 0$ and $k \geq 0$ we have

$$N(t) \geq k \Leftrightarrow S_k \leq t.$$

From the definition of a Poisson process, we see that the renewal process generated by a sequence of i.i.d. $Exp(\lambda)$ r.v.'s is a Poisson process with parameter λ. Thus a renewal process is a direct generalization of a Poisson process when the inter-event times are i.i.d. nonnegative random variables that have a general distribution. We next list some phenomena, which it has been suggested can be modelled by renewal processes.

Example 3.1.1 Scheduled maintenance. An automobile is lubricated when its owner has driven it L miles or every M days, whichever comes first. Let $N(t)$ denote the number of lubrications up to time t. Suppose the numbers of miles driven in disjoint time periods are independent, and the number of miles in any time interval has the same distribution, regardless of where the

https://doi.org/10.26219/heal.aueb.6246

interval begins. Then it is reasonable to say that $N(t)$ is a renewal process. The inter-renewal distribution is $F(t) = P(M \le t, C \le t)$, where C denotes the time to accumulate L miles on the automobile.

3.1.1 The renewal function

An important role in renewal theory is played by the renewal function $M(t)$ which is defined by

$$M(t) = E[N(t)], \qquad t \ge 0 \tag{3.1.1}$$

For $n = 1,2, \dots$ define the probability distribution function

$$F_n(t) = P(S_n \le t), \qquad t \ge 0.$$

Note that $F_1(t) = F(t)$. From Remark 3.1.1 we have that

$$F_n(t) = P(N(t) \ge n) \tag{3.1.2}$$

Lemma 3.1.2 For any $t \ge 0$,

$$M(t) = \sum_{n=1}^{\infty} F_n(t) \tag{3.1.3}$$

Proof: Since for any non-negative integer-valued random variable N,

$$E(N) = \sum_{k=0}^{\infty} P(N > k) = \sum_{n=1}^{\infty} P(N \ge n)$$

the relation (3.1.3) is an immediate consequence of (3.1.2).

Remark 3.1.2

$$\boldsymbol{p_n(t) \coloneqq P(N(t) = n) = P(N(t) \ge n) - P(N(t) \ge n+1) = F_n(t) - F_{n+1}(t).}$$

Example 3.1.2 Poisson process. Let $F(t) = 1 - e^{-\lambda t}, t \ge 0$. Then $F_n(t)$ is the cdf of an $Erlang(n, \lambda)$ random variable. Hence

$$F_n(t) = 1 - \sum_{k=0}^{n-1} e^{-\lambda t} \frac{(\lambda t)^k}{k!}, t \ge 0$$

Hence from remark 3.1.2, we get

$$p_n(t) = F_n(t) - F_{n+1}(t) = e^{-\lambda t} \frac{(\lambda t)^k}{k!}$$

https://doi.org/10.26219/heal.aueb.6246

Thus $N(t) \sim P(\lambda t)$ and $M(t) = E\big(N(t)\big) = \lambda t, \quad t \geq 0$. This result is in accordance with previous Poisson process outcomes.

3.1.2 Recurrence times

For a renewal process $\{N(t), t \geq 0\}$, let

- $A(t) = t - S_{N(t)}, \quad t \geq 0$
- $\gamma_t = \mathrm{B}(t) = S_{N(t)+1} - \mathrm{t}, \quad t \geq 0$
- $C(t) = S_{N(t)+1} - S_{N(t)} = X_{N(t)+1} = A(t) + B(t), \quad t \geq 0$

Thus, after the most recent renewal at or before t; $A(t)$ is the cumulative time elapsed; it is called age or current life time or backward recurrence. The time from t to the first renewal after t is $B(t)$; it is referred to as the excess variable or excess life or residual life. The length of the inter-renewal cycle spanning time t is $C(t)$; it is named the total lifetime or just total life. The associations between $A(t), B(t)$ and $C(t)$ can be seen graphically in figure 2.1

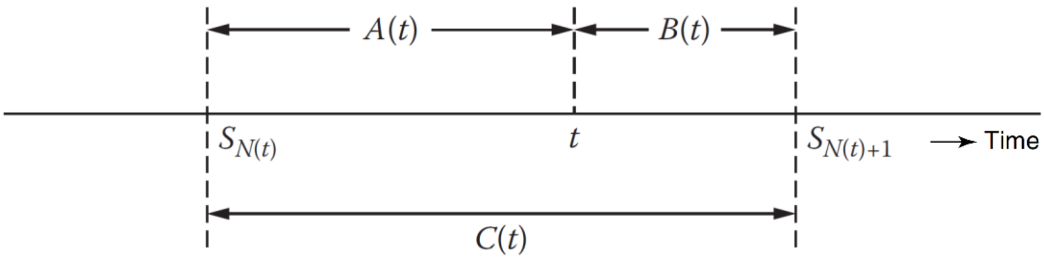

Figure 3.1 The random variables $A(t), B(t)$ and $C(t)$.

Figures 2.2, 2.3 and 2.4 show the sample paths of the three stochastic process.

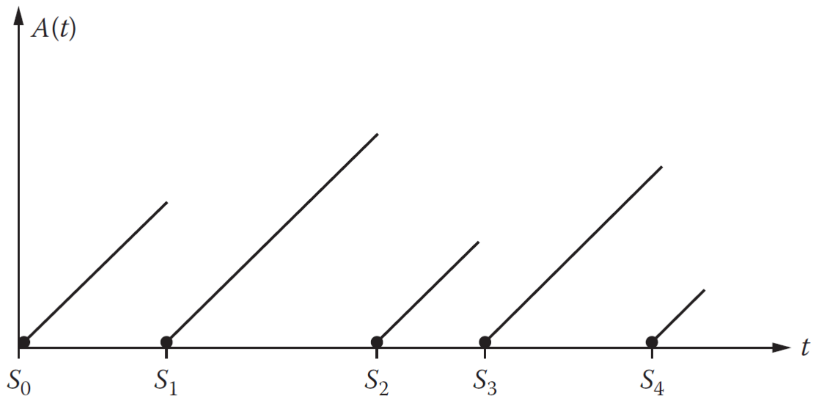

Figure 3.2 A typical sample path of the age process.

https://doi.org/10.26219/heal.aueb.6246

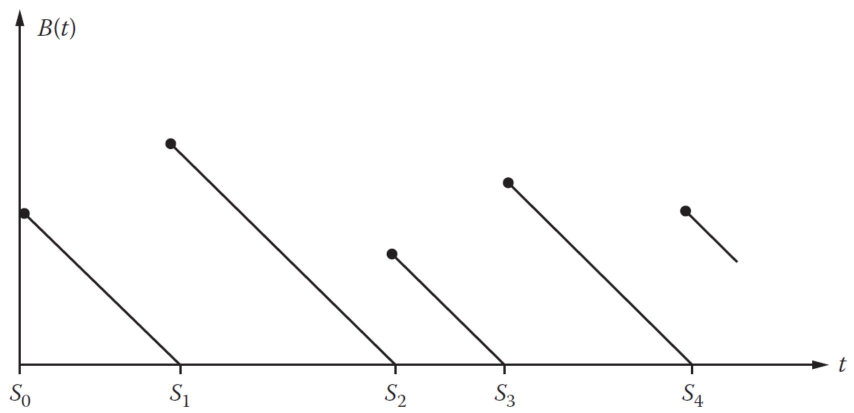

Figure 3.3 A typical sample path of the excess life process.

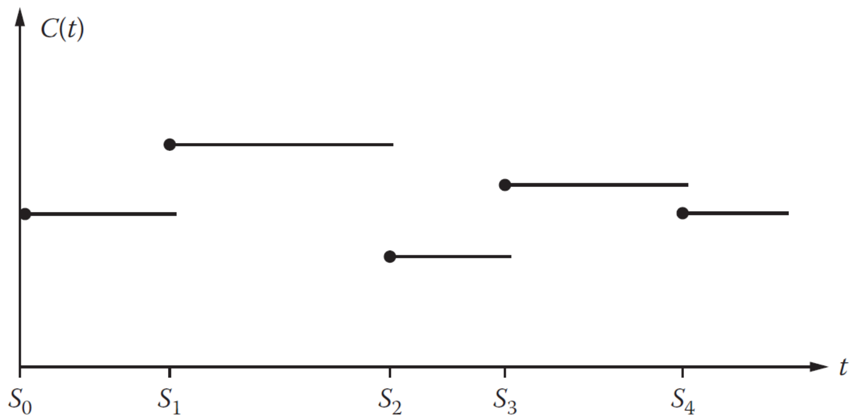

Figure 3.4 A typical sample path of the total life process.

The age process increases linearly with rate 1, jumps down to zero at every renewal epoch S_n. The remaining life process starts at X_1, and decreases linearly at rate 1. When it reaches zero at time S_n, it jumps up to X_{n+1}. The total life process has piecewise constant sample paths with upward or downward jumps at S_n.

Lemma 3.1.1 For $t \geq 0$,

$$E(\gamma_t) = \mu_1[1 + M(t)] - t. \tag{3.1.4}$$

Proof: Fix $t \geq 0$. To prove (3.1.4), we apply Wald's equation. To do so, note that $N(t) \leq n - 1$ if and only if $X_1 + \cdots + X_n > t$. Hence the event $\{N(t) + 1 = n\}$ depends only on $X_1, \ldots, X_n$ and is thus independent of $X_{n+1}, X_{n+2}, \ldots$. Hence:

$$E\left[\sum_{k=1}^{N(t)+1} X_k\right] = E(X_1)E[N(t) + 1],$$

https://doi.org/10.26219/heal.aueb.6246

which gives (3.1.4).

3.2 RENEWAL-REWARD PROCESSES

The renewal process is used to count the renewals that occur in an non-deterministic system. As the generalizations of renewal process, the renewal reward process models the cumulative rewards associated with the renewals up to some given time. A number of very interesting problems in queuing systems, reliability, stock control theory etc. are expressed by means of renewal-reward processes. Renewal-reward processes are also very important for ergodic theorems for Markov chains and other theoretical objectives. The model of renewal-reward is a simplistic and logically "attractive" model dealing with a regenerative process where a mechanism of cost or reward is placed. Certain stochastic processes do have the characteristic of regenerating themselves at many points in time $\{S_n, n \geq 1\}$ so that the probabilistic behavior of the stochastic process during the nth cycle, $\{X(t), t: S_{n-1} \leq t < S_n\}$, is the same for each $n \geq 1$, and independent from cycle to cycle. This idea is made precise in the following definition

Definition 3.2.1 A stochastic process $\{X(t), t \in T\}$ with time-index set T is called a regenerative process if there exists a (random) epoch S_1 such that:

1. $\{X(t + S_1), t \in T\}$ is independent of $\{X(t), 0 \leq t < S_1\}$
2. $\{X(t), t \in T\}$ and $\{X(t + S_1), t \in T\}$ are stochastically identical.

The index set T can be $T = \{0,1,2, \dots\}$ or $T = (0, \infty)$. We have a continuous time regenerative process in the second case, and a discrete time regenerative process in the first case. It is assumed that the state space of the process $\{X(t), t \in T\}$ is a subset of any Euclidean space. The existence of S_1 means the existence of an infinite sequence of increasing random variables $\{S_n, n \geq 1\}$ such that $\{X(t), \geq 0\}$ and $\{X(t + S_n), t \geq 0\}$ are stochastically identical, and $\{X(t + S_n), t \geq 0\}$ is independent of $\{X(t), 0 \leq t < S_n\}$. We say that S_n is the nth regeneration epoch, since the stochastic process loses all its memory and starts afresh at times S_n. The interval $[S_{n-1}, S_n)$ is called the nth regenerative cycle (we assume that $S_0 = 0$). Examples of regenerative processes are:

1. Consider an inventory model in which the daily demand is represented by random variables $Y_1, Y_2, \dots$ The Y_i are assumed to be i.i.d. integer

https://doi.org/10.26219/heal.aueb.6246

valued. The initial inventory is given by $X_1 = S$. The following (s,S) (with s < S) ordering policy is used. If the inventory X_n at the beggining of the day n is less than s, then we order enough to bring it back to the initial level S. It is easy to see that $\{X_n, n \geq 1\}$ is a discrete time regenerative process.

2. M/GI/k Queue. The M/G/k queue has a Poisson arrival process and k servers with general service time distribution. Whenever the queue is empty, all servers are idle and only the arrival process has an effect on the future. Thus the system process regenerates at the points S_n of the system becoming idle for the nth time. The durations $S_{n+1} - S_n$ between these points are i.i.d. Hence the M/G/k system process is a regenerative process.

Now let define the random variables $C_n = S_n - S_{n-1}$, $n = 1,2, \ldots$, where $S_0 = 0$ through convention. The random variables $C_1, C_2, \ldots$ are i.i.d. Actually, the sequence $\{C_1, C_2, \ldots\}$ underlies a renewal process in which the arrivals/events are the occurrences of the regeneration epochs. We can therefore interpret C_n as

$$C_n = \text{the length of the } n\text{th renewal cycle}, \quad n = 1,2, \ldots$$

Note also that the length of the C_n cycle assumes values from the index set T. In the following it is assumed that

$$0 < E(C_1) < \infty$$

A reward structure is given to the regenerative process $\{X(t), t \in T\}$ in several realistic applications. In general, it is assumed that the reward structure consists of reward amounts that are consistently accruing throughout time and batch rewards are earned at such state changes. Let us now introduce a reward structure on the regenerative process as follows:

$$R_n = \text{the total reward earned in the } n\text{th renewl cycle}, \quad n = 1,2, \ldots$$

It is assumed that $R_1, R_2, \ldots$ are i.i.d. In many applications R_n usually depends on C_n. If R_n is capable of taking both positive and negative values,
$\mathrm{E}(|R_1|) < \infty$ is assumed. Let

$$R(t) = \text{the cumulative reward earned up to time } t$$

https://doi.org/10.26219/heal.aueb.6246

The process $\{R(t), t \geq 0\}$ is called a renewal-reward process. We are now ready to prove the most important theorem in renewal-reward theory.

Theorem 3.2.1 Renewal-Reward theorem.

$$\lim_{t\to\infty} \frac{R(t)}{t} = \frac{E(R_1)}{E(C_1)} \quad \text{with probability 1.}$$

The above theorem is very intuitive and useful: it says that the long run expected rate of reward is simply the ratio of the expected reward in one cycle and the expected length of that cycle. What is surprising is that the reward does not have to be independent of the cycle length. This is what makes theorem so useful in applications. To prove this theorem we first mention the following lemma.

Lemma 3.2.2 (also known as the elementary renewal theorem) For any $t \geq 0$, let $N(t)$ be the number of cycles completed up to time t. Then

$$\lim_{t\to\infty} \frac{N(t)}{t} = \frac{1}{E(C_1)} \quad \text{with probability 1.}$$

Proof: By the definition of $N(t)$, we have

$$C_1 + \cdots + C_{N(t)} \leq t < C_1 + \cdots + C_{N(t)+1}.$$

Since $P(C_1 + \cdots + C_n < \infty) = 1$ for all $n \geq 1$, it is not difficult to verify that

$$P\left(\lim_{t\to\infty} N(t) = \infty\right) = 1$$

The above inequality gives

$$\frac{C_1 + \ldots + C_{N(t)}}{N(t)} \leq \frac{t}{N(t)} < \frac{C_1 + \ldots + C_{N(t)+1}}{N(t)+1} \cdot \frac{N(t)+1}{N(t)}.$$

By the strong law of numbers for a sequence of independent and identically distributed random variables, we have

$$P\left(\lim_{n\to\infty} \frac{C_1 + \ldots + C_n}{n} = E(C_1)\right) = 1$$

Hence, by letting $t \to \infty$ in the above inequality, the desired result follows.

Proof of theorem 3.2.1 For ease, let us first assume that the rewards are non-negative. Then for any $t > 0$,

https://doi.org/10.26219/heal.aueb.6246

$$\sum_{i=1}^{N(t)} R_i \leq R(t) \leq \sum_{i=1}^{N(t)+1} R_i.$$

This gives

$$\frac{\sum_{i=1}^{N(t)} R_i}{N(t)} \cdot \frac{N(t)}{t} \leq \frac{R(t)}{t} \leq \frac{\sum_{i=1}^{N(t)+1} R_i}{N(t)+1} \cdot \frac{N(t)+1}{t}.$$

By the strong law of large numbers for the sequence $\{R_n\}$, we have

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} R_i = E(R_1) \quad \text{with probability 1.}$$

As described in the Lemma 3.2.2, we have $P\left(\lim_{t\to\infty} N(t) = \infty\right) = 1$. Letting $t \to \infty$ in the above inequality and using Lemma 3.2.2, the desired result follows for the case that the rewards are non-negative. In the case where both positive and negative values are allowed, then the theorem can be proved by independently handling the positive and negative values of the rewards. We do not give the details.

If we think of a cycle being completed every time a renewal occurs, then the long run expected reward per time unit is given by the expected reward in a cycle divided by the expected length of one cycle.

Example 3.2.1 Alternating renewal process

An important special case of a regenerative process is the one for which the state space of the continuous-time process $\{X(t), t \geq 0\}$ contains only two elements, which will be denoted by 0 and 1. For example if $X(t) = 0$ then a (hypothetical) machine is down, and $X(t) = 1$ if it is operating at time t. Suppose that the machine is new at the initial time $S_0 = 0$, so that $X(0) = 1$ and that the time during which the machine operates, before breaking down, is a random variable U_1. Next, the repairman takes a random time D_1 to fix the machine, and then the machine starts again. Let

$$S_n = U_n + D_n \qquad \text{for } n = 1,2, \dots$$

where S_n is the operating time of the machine between the $(n-1)$th and the nth breakdown, and U_n is the repair time of the nth breakdown. We assume that $\{(U_n, D_n), n \geq 1\}$ is a sequence of i.i.d. bivariate random variables. Note that

 https://doi.org/10.26219/heal.aueb.6246

$\{U_n\}$ *and* $\{D_n\}$ are not required to be independent of each other. A sample path of such a process is shown in figure 3.5

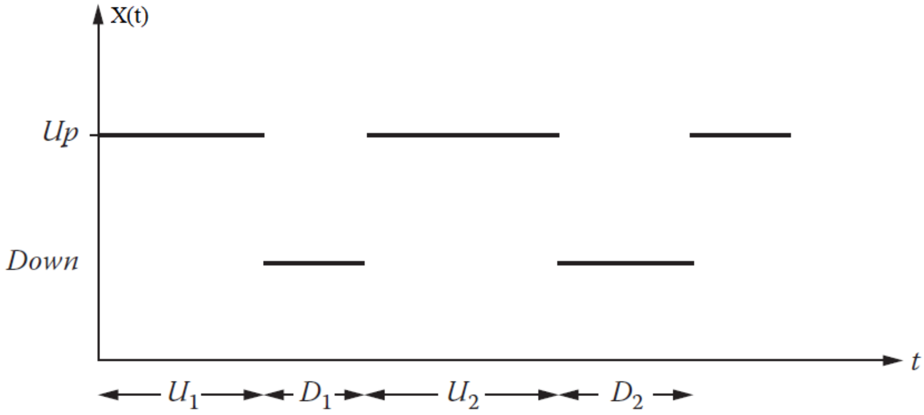

figure 3.5 A typical sample path of an alternating renewal process.

So, what is the long run fraction of time that the machine is down? The answer is

$$\text{the long run fraction of time the machine is down } = \frac{E(D_1)}{E(U_1) + E(D_1)} \qquad (3.2.1)$$

To confirm this, note that $\{X(t), t \geq 0\}$ is a regenerative process. The epochs where an operational period begins can be taken as regeneration epochs. The long run fraction of time that the machine is broken-down can be interpreted as a long run average cost per time unit by assuming that a cost of rate 1 is paid as long as the machine is broken and a cost of rate 0 while the machine is working properly. Therefore

$$E(\text{length of one cycle}) = E(U_1 + D_1)$$

and

$$E(\text{cost incurred during one cycle}) = E(D_1).$$

It follows that the long run average cost per time unit equals $E(D_1)/[E(D_1) + E(D_2)]$, by applying the renewal reward theorem, which proves the result (3.2.1)

Example 3.2.2 A bulb shines for a certain amount of time let's assume with a distribution F with mean μ and then burns out. At times, a janitor comes to inspect the bulb according to a Poisson process of rate λ, if the bulb is burnt out he will fix (or replace) it. A) At what rate are bulbs replaced? B) What is the limiting fraction of time that the bulb works? C) What is the limiting fraction of visits on which the bulb is working?

Solution: Assume that a brand-new bulb is put in at time $S_0 = 0$. It will last for an amount of time S_1. Using the lack of memory property of the exponential distribution, it follows that the amount of time until the next examination, u_1, will have an exponential distribution with rate λ. The bulb is then replaced and the cycle starts again, so we have an alternating renewal process.

A) We note that the expected length of a cycle $E(C_i) = \mu + 1/\lambda$, so if $N(t)$ is the number of bulbs replaced by time t, then is follows from Lemma 3.2.2 that

$$\frac{N(t)}{t} \to \frac{1}{\mu + 1/\lambda}$$

In words, bulbs are replaced on the average every $\mu + 1/\lambda$ units of time.

B) We let $R_i = C_i$, so (3.2.1) implies that in the long-run, the fraction of time the bulb has been working up to time t is

$$\frac{E(R_i)}{E(C_i)} = \frac{\mu}{\mu + 1/\lambda}$$

C) We note that if $V(t)$ is the number of visits the janitor has made by time t, then by the law of large numbers for the Poisson process we have

$$\frac{V(t)}{t} \to \lambda$$

Combining this with the result A) and we see that the fraction of visits on which bulbs are replaced is

$$\frac{N(t)}{V(t)} \to \frac{1/(\mu + 1/\lambda)}{\lambda} = \frac{1/\lambda}{\mu + 1/\lambda}$$

This answer is reasonable since it is also the limiting fraction of time the bulb is off.

Remark 3.2.1 An essential implementation of the renewal reward theorem is the characterization of the long run fraction of time a regenerative process $\{X(t), t \in T\}$ spends in some given set B of states. For the set B of states, define for any $t \in T$ the indicator variable

$$I_B(t) = \begin{cases} 1 & if X(t) \in B \\ 0 & if X(t) \notin \mathrm{B} \end{cases}$$

Also define the random variable

$T_B =$ the amount of time the process spends in the set B of states during one cycle.

Keep in mind that that $T_B = \int_0^{S_1} I_B(u)du$ for a continuous time process $\{X(t), t \geq 0\}$; otherwise, T_B equals the number of indices $0 \leq k \leq S_1$ with $X(k) \in B$. The following theorem is an immediate consequence of the renewal reward theorem.

Theorem 3.2.3 For the regenerative process $\{X(t), t \geq 0\}$ it holds that the long run fraction of time the process spends in the set B of states is $E(T_B)/E(C_1)$ with probability 1. That is,

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t I_B(u)du = \frac{E(T_B)}{E(C_1)} \text{ with probability 1}$$

for a continuous time process $\{X(t), t \geq 0\}$ and

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n} I_B(k) = \frac{E(T_B)}{E(C_1)} \text{ with probability 1}$$

for a discrete time process$\{X(n), n \in N\}$.

Proof: The long run fraction of time the process $\{X(t), t \geq 0\}$ spends in the set B of states can be interpreted as a long run average reward per time unit assuming that a reward at rate 1 is earned while the process is in the set B and a reward at rate 0 is earned otherwise. Then

$$E(\text{reward earned during one cycle}) = E(T_B)$$

Next, applying the renewal reward theorem, the desired outcome follows.

Since $E\big(I_B(t)\big) = P(X(t) \in B)$, we have as consequence of theorem 3.2.3 and the bounded convergence theorem that, for a continuous time process

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t P(X(u) \in B)du = \frac{E(T_B)}{E(C_1)}$$

Note that $(1/t)\int_0^t P(X(u) \in B)du$ can be interpreted as the probability than an outside observer arriving at a randomly chosen point in $(0, t)$ finds the process in the set B.
In certain cases, the proportion $E(T_B)/E(C_1)$ may be translated both as the long run fraction of time spent in the set B of states by the process $\{X(t), t \geq 0\}$ and as the probability of having the process in the set B when the process has achieved statistical balance (equilibrium). This raises the question whether $\lim_{t\to\infty} P(X(t) \in B)$ always

exists. This ordinary limit need not always exist. A counterexample is provided by periodic discrete time Markov chains.

Theorem 3.2.4 For the regenerative process $\{X(t), t \in T\}$

$$\lim_{t\to\infty} P(X(t) \in B) = \frac{E(T_B)}{E(C_1)}$$

provided that the probability distribution of the cycle length has a continuous part in the continuous time case and is aperiodic in the discrete time case.

A distribution function is said to have a continuous part if it has a positive density on some interval. A discrete distribution $\{\mathrm{a_j}, \mathrm{j} = 0,1, \dots\}$ is said to be aperiodic if the greatest common divisor of the indices $j \geq 1$ for which $a_j > 0$ is equal to 1. The proof of theorem 3.2.4 requires deep mathematics and we will not present it. Let's try to illustrate theorem 3.2.4. Consider again Example 3.2.1. In this example we analyzed the long run average behavior of the regenerative process $\{X(t), t \geq 0\}$, where $X(t) = 1$ if the machine is up at time t and $X(t) = 0$ otherwise. It was shown that the long run fraction of time the machine is down equals $E(D)/E(U) + E(D)$, where the random variables U and D denote the lengths of an up-period and a down-period. This result does not require any assumption about the shapes of the probability distributions of U and D. However, some assumption is needed in order to conclude that

$$\lim_{t\to\infty} P(\text{the system is down at time t}) = \frac{E(D)}{E(U) + E(D)}$$

It is sufficient to assume that the distribution function of the length of an up-period has a positive density on some interval.

Theorem 3.2.5 Let m and σ^2, assumed finite, represent the mean and variance of an interarrival time. Then

$$P\left(\frac{N(t) - t/m}{\sigma\sqrt{t/m^3}} < y\right) \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-x^2/2} dx \quad as \quad t \to \infty$$

Proof: Let $n_t = t/m + y\sigma\sqrt{t/m^3}$ Then

$$P\left(\frac{N(t) - t/m}{\sigma\sqrt{t/m^3}} < y\right) = P(N(t) < n_t) = P\left(S_{n_t} > t\right) = P\left(\frac{S_{n_t} - n_t m}{\sigma\sqrt{n_t}} > \frac{t - n_t m}{\sigma\sqrt{n_t}}\right)$$

https://doi.org/10.26219/heal.aueb.6246

$$= P\left(\frac{S_{n_t} - n_t m}{\sigma\sqrt{n_t}} > -y\left(1 + \frac{y\sigma}{\sqrt{tm}}\right)^{-t/2}\right)$$

Now, by the central limit theorem, $(S_{n_t} - n_t m)/\sigma\sqrt{n_t}$ converges to a normal random variable having mean 0 and variance 1 as t (and thus n_t) approaches $+\infty$. Also, since

$$-y\left(1 + \frac{y\sigma}{\sqrt{tm}}\right)^{-t/2} \to -y \quad \text{as} \quad t \to \infty$$

we see that

$$P\left(\frac{N(t) - t/m}{\sigma\sqrt{t/m^3}} < y\right) \to \frac{1}{\sqrt{2\pi}} \int_{-y}^{\infty} e^{-x^2/2} dx$$

and since

$$\int_{-y}^{\infty} e^{-x^2/2} dx = \int_{-\infty}^{y} e^{-x^2/2} dx \qquad \text{the result follows.}$$

Therefore $N(t)$ is asymptotically normal with mean t/m and variance $\sigma^2 t/m^3$.
For Poisson process $E\big(N(t)\big) = \lambda t = t/m$ and $V\big(N(t)\big) = \lambda t = t/m$ for exponential distribution, $m = 1/\lambda, \sigma^2 = 1/\lambda^2$

$$\frac{t}{m} = \frac{t}{1/\lambda} = \lambda t \quad \text{and} \quad \frac{\sigma^2 t}{m^3} = \frac{(1/\lambda)^2 t}{(1/\lambda)^3} = \lambda t$$

Thus the results hold exactly for a Poisson process.

Example 3.2.3 A stochastic clearing system

Suppose a manufacturing facility produces finished good one at a time according to a Poisson process with mean production time $\lambda < \infty$ per item and stores them in a warehouse. As soon as there are T items in the warehouse we clear it by shipping them to the retailers. The clearing is instantaneous. It costs h to hold an item in the warehouse per unit time, and it costs K to clear the warehouse. What is optimal value of T that minimizes the long run holding plus clearing cost per unit time?

Solution: Let $X(t)$ be the number of items in the warehouse at time t. Suppose $X(0) = 0$ and S_n be the time of the nth clearing time of the warehouse. It is obvious that $\{X(t), t \geq 0\}$ is a regenerative process on state space $\{0,1, \dots, T-1\}$. A typical sample path of the $\{X(t), t \geq 0\}$ process is shown in figure 3.6. We see that

https://doi.org/10.26219/heal.aueb.6246

$$P(S_1 < \infty) = 1, \;\; E(S_1) = T\lambda$$

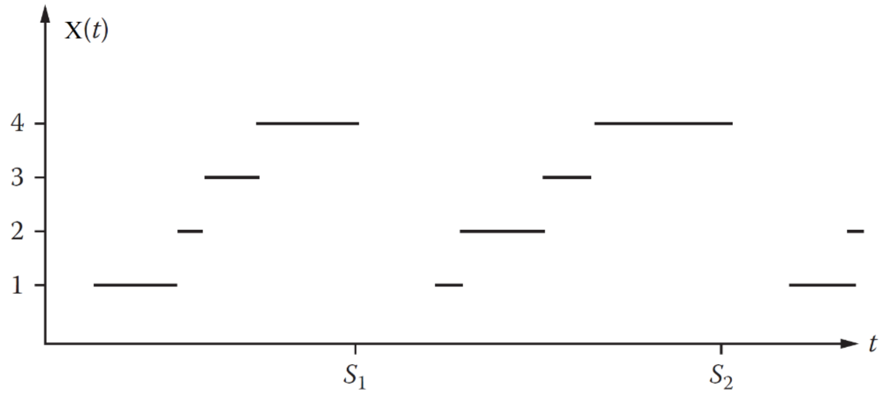

figure 3.6 A typical sample path of the $\{X(t), t \geq 0\}$ process

Hence, from theorem (3.2.4) we get

$$\lim_{t\to\infty} P(X(t) = j) = \frac{E(\text{the amount of time the process spends at state } j)}{E(\text{duration of } S_1)} = \frac{\lambda}{T\lambda} = \frac{1}{T}$$

$0 \leq j \leq T-1$. We compute C_h, the long run expected holding costs per unit time first. The system incurs waiting costs at a rate jh per unit time if there are j items in the warehouse. By the law of total expectation we get:

$$C_h = \sum_{j=0}^{T-1} jh \lim_{t\to\infty} P(X(t) = j) = \frac{h}{T}\sum_{j=0}^{T-1} j = \frac{1}{2}h(T-1).$$

Now let's compute C_K, the long run clearing cost per unit time, from renewal reward theorem we get:

$$C_K = \frac{E(\text{fixed cost for clearing the warehouse})}{E(\text{duration of } S_1)} = \frac{K}{T\lambda}.$$

Thus the $C(T)$, the long run total cost per unit time is given by

$$C(T) = \frac{1}{2}h(T-1) + \frac{K}{T\lambda}.$$

This is a convex function of T and is minimized at $T^* = \sqrt{\frac{2K}{h\lambda}}$. Note that T^* must be an integer, so we check the two integers near the above solution.

3.3 LITTLE'S LAW

One of the most important and useful relationships in queueing theory is what is commonly known as Little's law. To demonstrate this, we use two numerical examples.

 https://doi.org/10.26219/heal.aueb.6246

In the first instance consider a computer network node in which requests for data (users) arrive at an average rate of 3 per minute. Each user stays on the network for 20 mins on average. What is the average number of users on the network? Let $\lambda = 3$ denote the average number of new users on the network per minute and $W = 20$ the average number of minutes a user stays on the network. Then by Little's law or Little's formula we have $\mathrm{L} = \lambda \cdot W = 3 \times 20 = 60$ users on the network in average. In the second example consider a semiconductor factory. The manufacturing process entails starting with a silicon wafer and then building the electronic circuitry for multiple identical devices. Suppose that the semiconductor factory starts $\lambda = 1000$ wafers per day, on average; this is the input rate. Let's say the average amount of work in process inventory is $L = 45000$ wafers. Thus the wait time or expected time in the system for a random wafer is $W = L/\lambda = 45000/1000 = 45$ days. These instances demonstrate Little's law $L = \lambda W$. We all understand intuitively the Little's law so it can be said it is a "law of nature" and it has many applications in queueing theory. The Little's Law states that L the average number of customers (or users/wafers etc.), λ, the average arrival rate of customers to the system, and W, the average time a customer spends in the system are related to each other by the equation $L = \lambda W$. We will not give the proof of this equation. It is convenient to distinguish between queueing systems with infinite queue capacity and queueing systems with finite queue capacity.

Infinite capacity queues

Consider a queue with infinite queue capacity, thus every customer is allowed to wait in the queue until service can be provided. Define the following random variables:

$L_q(t) =$ the number of customers waiting in the queue at time t
(excluding those in service)

$L(t) =$ the number of customers in the system at time t
(including those in service)

$D_n =$ the amount of time spent by the nth customer in the queue
(excluding service time)

$U_n =$ the amount of time spent by the nth customer in the system
(including service time).

Assume that each of the stochastic processes $\{L_q(t)\}, \{L(t)\}, \{D_n\}$ and $\{U_n\}$ is regenerative and has a cycle length with a finite expectation. Then there are constants

https://doi.org/10.26219/heal.aueb.6246

L_q, L, W_q and W such that the following limits exist and are equal to the respective constants with probability 1:

$$\lim_{t\to\infty}\frac{1}{t}\int_0^t L_q(u)du = L_q \quad \text{(the long run average number in queue)}$$

$$\lim_{t\to\infty}\frac{1}{t}\int_0^t L(u)du = L \quad \text{(the long run average number in system)}$$

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n} D_k = W_q \quad \text{(the long run average delay in queue per customer)}$$

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n} U_k = W \quad \text{(the long run average sojourn time per customer)}$$

Now define the random variable

$A(t) =$ the number of customers arrived by time t

It is also assumed that, for some constant λ,

$$\lim_{t\to\infty}\frac{A(t)}{t} = \lambda \quad \text{with probability 1.}$$

The constant λ gives the long run average arrival rate of customers. The limit λ exists when customers arrive to a renewal process. The existence of the above limits is sufficient to prove the basic relations

$$L_q = \lambda W_q \tag{3.3.1}$$

and

$$L = \lambda W \tag{3.3.2}$$

These basic relations are the most familiar form of Little's formula. We will demonstrate the plausibility of this result. An intuitive explanation of Little's Law is as follows: Suppose each arriving customer pays the system 1€ per unit time the customer spends in the system. The long run rate at which the system earns revenue can be computed in two equivalent ways. First, the nth customer pays the system U_n€, the time spent by that customer in the system. Hence the average amount paid by a customer in steady state is W€. Since the average arrival rate is λ, the system earns λW€ per unit time in the long run. Second, since each customer in the system pays at rate 1€ per unit time, the system earns revenue at the instantaneous rate of $L(t)$€ per unit time at time

 https://doi.org/10.26219/heal.aueb.6246

t. Hence the long run rate of revenue is seen to be L, the steady state expected value of $L(t)$. Since these two calculations must provide the same answer, equation (3.3.2) follows. The formula (3.3.1) can be seen by a very similar reasoning: imagine that each customer pays 1€ per unit time while waiting in queue. Another interesting relation arises by imagining that each customer pays 1€ per unit time while in service. Denoting by $E(S)$ the long run average service time per customer, it follows that

the long run average number of customers in service = $\lambda E(S)$ (3.3.3)

If each customer requires only one server and each server can handle only one customer at a time, it follows that

the long run average number of busy servers $= \lambda E(S)$ (3.3.4)

Finite capacity queues

Now, assume that the number of customers permitted in the system is finite. That is to say, there are only a finite number of waiting spots and each arriving customer finding all waiting spots occupied is turned away. A rejected customer is assumed to have no further impact on the system (for example s/he could come back in the system after an hour). Let the rejection probability P_{rej} be defined by

$$P_{rej} = \text{the long run fraction of customers who turned away},$$

assuming that this long run fraction is well defined. The random variables $\mathrm{L}(t), L_q(t), D_n$ and U_n are defined as before, except that D_n and U_n now refer to the queueing time and sojourn time of the nth accepted customer. The constants W_q and W now represents the long run average queueing time per accepted customer and the long run average sojourn time per accepted customer. The formulas (3.3.1), (3.3.2) and (3.3.3) need only slight modification:

$L_q = \lambda(1 - P_{rej})W_q$ and $\mathrm{L} = \lambda(1 - P_{rej})\mathrm{W}$, (3.3.5)

the long run average number of customers in service $= \lambda(1 - P_{rej})E(S)$. (3.3.6)

Heuristically, these formulas follow by applying the reward principle we used for Little's Law and taking into account only the accepted customers as paying customers.

3.4 POISSON ARRIVALS SEE TIME AVERAGES

The Poisson arrivals see time averages property, abbreviated as PASTA, asserts that: if the arrival process in a queue is Poisson then the long run state distribution of the state of the queue seen by a customer just before his arrival is the same as the long run state distribution of the state of the queue at an arbitrary point of time. This is not true in general, as it can be seen by the following counterexample. Consider a single server queue into which customers arrive with exactly 2 time units difference deterministically, and suppose also that the arrival time of the customer is random, uniformly distributed on [0,2]. Suppose that customers need service for i.i.d. random times uniformly distributed between 0 and 2. If the queue is started empty, then the queue length $L(t)$ becomes 1 when the first customer arrives, stays at 1 for a random length of time of unit mean, and uniformly distributed between 0 and 2. Thus, the queue length is guaranteed to return to zero before the next arrival. This cycle repeats itself indefinitely. Now, if we consider any fixed time, then at this time the queue is equally likely to either be empty or to have exactly 1 customer in it. Thus, the steady state queue length distribution assigns probability 0.5 each to states 0 and 1, and probability 0 to all other states. On the other hand, new arrivals always see an empty queue. Thus, the queue length distribution seen by arrivals is not the same as the long run state queue length distribution.

Example 3.4.1 M/G/1/1 queue.

Customers arrive from outside to a single server according to a Poisson process with rate λ and require i.i.d service times with common distribution G and mean $\beta < \infty$. There is room only for the customer in service. Thus the capacity is 1. If an arriving customer finds the server idle, he immediately enters service otherwise he turns away. What is the long-run fraction of time the server is busy and what is the long-run fraction of customers that are lost? These two questions are easily answered by using the renewal reward theorem. Let us define the following random variables. For any $t \geq 0$, let $I(t) = \begin{cases} 1 & if\ the\ server\ is\ busy\ at\ time\ t \\ 0 & otherwise. \end{cases}$

Also, for any $n = 1,2, \ldots,$ let

$$I_n = \begin{cases} 1 & if\ the\ server\ is\ busy\ just\ prior\ to\ the\ nth\ arrival \\ 0 & otherwise. \end{cases}$$

Both process $\{I(t)\}$ and $\{I_n\}$ are regenerative. The arrival epochs occuring when the server is idle are regeneration epochs for the two processes. Let us say that a cycle starts each time an arriving customer finds the serve idle. The long-run fraction of time the server is busy is equal to the expected amount of time the server is busy during one

 https://doi.org/10.26219/heal.aueb.6246

cycle divided by the expected length of one cycle. The expected length of the busy period in one cycle equals β. Since the Poisson arrival process is memoryless, the expected length of the idle period during one cycle equals the mean interarrival time $1/\lambda$. Hence, with probability 1,

$$\text{the long}-\text{run fraction of time the server is busy} = \frac{\beta}{\beta + 1/\lambda}$$

The long-run fraction of customers that are lost equals the expected number of customers lost during one cycle divided by the expected number of customers arriving during one cycle. Since the arrival process is a Poisson process, the expected number of lost arrivals during the busy period in one cycle equal =

$= \lambda \times E(processing\ time\ of\ a\ service) = \lambda\beta$. Hence with probability 1,

$$\text{the long}-\text{run fraction of customers that are lost} = \frac{\lambda\beta}{1+\lambda\beta} \tag{3.4.2}$$

Thus, we obtain from (3.4.1) and (3.4.2) the remarkable result:

long-run fraction of arrivals finding the workstation busy =

= long-run fraction of time the workstation is busy (3.4.3)

Note that the form of the distribution G has no impact on the above results. We only used the mean β of this distribution.

Example 3.4.2 An inventory model

Consider a single-product inventory system in which customers asking for the products arrive according to a Poisson process with rate λ. Every client requests for one of the commodity units. Any demand that can not be fulfilled immediately from the inventory on hand is considered lost. Opportunities to replenish the inventory occur according to a Poisson process with rate μ. The two processes are assumed independent from each other. A replenishment can only be carried out for logistical purposes when the inventory is empty. Any time a replenishment is completed, the stock on hand is elevated to level Q. What is the long run fraction of time the system is out of stock? What is the long run fraction of demand that is lost?

In the same way as in Example 3.4.1, we define the random variables

$$I(t) = \begin{cases} 1 & if\ the\ system\ is\ out\ of\ stock\ at\ time\ t \\ 0 & otherwise \end{cases}$$

and

https://doi.org/10.26219/heal.aueb.6246

$$I_n = \begin{cases} 1 & \text{if the system is out of stock when the nth demand occurs} \\ 0 & \text{otherwise} \end{cases}$$

Both the continuous time process $\{I(t)\}$ and the discrete time process $\{I_n\}$ are regenerative. The regeneration cycles are the periods of demand under which the stock drops to nothing on hand. Why? Let us assume that every time we run out stock a regeneration epoch begins. The system is out of stock during the time elapsed from the beginning of a cycle until the next inventory replenishment. This length of time is distributed exponentially with a mean value of $1/\mu$. The length of time it required to go from stock level Q to 0 equals Q/λ on average. Therefore, with probability 1

$$\text{the long run fraction of time the system is out of stock } = \frac{1/\mu}{1/\mu + Q/\lambda} \qquad (3.4.4)$$

To find the fraction of demand that is lost, note that the expected amount of demand lost in one cycle equals $\lambda \times$ E(amount of time the system is out of stock during one cycle) $= \lambda/\mu$. Therefore, with probability 1,

$$\text{the long run fraction of demand that is lost } = \frac{\lambda/\mu}{\lambda/\mu + Q}. \qquad (3.4.5)$$

Together (3.4.4) and (3.4.5) lead to this remarkable result:

the long run fraction of customers finding the system out of stock

= the long run fraction of time the system is out of stock. (3.4.6)

The relations (3.4.3) and (3.4.6) are particular instances of the PASTA property. Roughly stated, this property expresses that in statistical equilibrium the distribution of the state of the system just prior to an arrival epoch is the same as the distribution of the state of the system at an arbitrary epoch when arrivals occur according to a Poisson process. For some specific problem let the continuous time stochastic process $\{X(t), t \geq 0\}$ describe the evolution of the state of a system and let $\{N(t), t \geq 0\}$ be a renewal process describing arrivals to that system. As examples:

a . $X(t)$ is the number of customers present at time t in a queueing system.

b . $X(t)$ describes jointly the inventory level and the prevailing production rate at time t in an inventory problem with a variable production rate. It is assumed tht the arrival process $\{N(t), t \geq 0\}$ can be seen as an exogenous factor to the system and is not affected by the system itself. More precisely, the following assumption is made.

https://doi.org/10.26219/heal.aueb.6246

Lack of anticipation assumption $\{\boldsymbol{N}(\boldsymbol{t}+\boldsymbol{u})-\boldsymbol{N}(\boldsymbol{u}), \boldsymbol{t} \geq \mathbf{0}\}$ and $\{X(t), 0 \leq t \leq u\}$ are independent. Note that $\{\boldsymbol{N}(\boldsymbol{t}+\boldsymbol{u})-\boldsymbol{N}(\boldsymbol{u}), \boldsymbol{t} \geq \mathbf{0}\}$ is independent of $\{N(t), 0 \leq t \leq u\}$ due to the independence of increments property of a Poisson process. Thus the lack of anticipation property says that the future arrivals after time u are independent of the system history up to time u. It is not necessary to specify how the arrival process $\{N(t)\}$ precisely interacts with the state process $\{X(t)\}$. Denoting by τ_n the nth arrival epoch, let the random variable X_n be defined by $X(\tau_n^-)$. In other words,

$$X_n = \text{the state of the system just prior to the } n\text{th arrival epoch.}$$

Let B be any set of states for the $\{X(t)\}$ process. For each $t \geq 0$, define the indicator variable

$$I_B(t) = \begin{cases} 1 & if\ X(t) \in B \\ 0 & otherwise \end{cases}$$

Also, for each $n = 1,2, \ldots$ define the indicator variable $I_n(B)$ by

$$I_n(B) = \begin{cases} 1 & if\ X_n \in B \\ 0 & otherwise \end{cases}$$

The technical assumption is made that the sample paths of the continuous time process $\{I_B(t), t \geq 0\}$ are right continuous and have left hand limits. In practical situations this assumption is always satisfied.

Theorem 3.4.1 Poisson arrivals see time averages.

Suppose that the arrival process $\{N(t), t \geq 0\}$ is a Poisson process with rate λ. Then:

1 . For any $t > 0$,

E[number of arrivals in (0,t) finding the system in the set B]=$\lambda E\left[\int_0^t I_B(u)\, du.\right]$

2 . With probability 1, the long run fraction of arrivals that find the system in the set B of states equals the long run fraction of time the system is in the set B of states. That is, with probability 1,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} I_K(B) = \lim_{t\to\infty} \int_0^t I_B(u) du.$$

The proof of the PASTA property will not be given and we only give an intuitive explanation. Suppose the process has been running since time -∞ and hence the system is in equilibrium. If we condition on there being arrival at time t, then the times of the

 https://doi.org/10.26219/heal.aueb.6246

previous arrivals are a Poisson process with rate λ. Thus knowing that there is an arrival at time t does not affect the distribution of what happened before time t. Theorem 3.4.1 has a useful corollary when it is assumed that the continuous time process $\{X(t)\}$ is a regenerative process whose cycles length has a finite positive mean. Define the random variables T_B and N_B by

$T_B =$ amount of time the process $\{X(t)\}$ is in the set B of states during one cycle

$N_B =$ number of arrivals during a cycle who find the $\{X(t)\}$ in the set of B states

Corollary 3.4.2 If the arrival process $\{N(t)\}$ is a Poisson process with rate λ, then

$$\mathrm{E}(N_B) = \lambda E(T_B).$$

Proof: Denote by the random variables T and N the length of one cycle and the number of arrivals during one cycle. Then, by theorem 3.2.3,

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t I_B(u)du = \frac{E(T_B)}{E(T)} \quad \text{with probability 1}$$

and

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} I_k(B) = \frac{E(N_B)}{E(N)} \quad \text{with probability 1}$$

It now follows from part 2 . of theorem 3.4.1 that $E(N_B)/E(N) = E(T_B)/E(T)$. Thus the corollary follows if we can verify that $E(N)/E(T) = \lambda$. To do so, note that the regeneration epochs for the process $\{X(t)\}$ are also regeneration epochs for the Poisson arrival process. Thus, by the renewal reward theorem, the long run average number of arrivals per time unit equals $E(N)/E(T)$, showing that $E(N)/E(T) = \lambda$. To conclude this section, we use the PASTA property to derive in a heuristic way one of the most famous formulas from queueing theory.

3.5 THE POLLACZEK-KHINCHIN MEAN VALUE FROMULA

Let customers arrive in a Poisson process with rate λ and get served by a single server. The service times of the customers are independent random variable having a common probability distribution S with finite mean and variance, so $E(S), E(S^2) < \infty$ and there is ample waiting room. This model is a queue model abbreviated as $M/G/1$ queue based on Kendall's notation. The offer load ρ is define by

$$\rho = \lambda E(S)$$

and it is assumed that $\rho < 1$. Noting that ρ is the expected number in service, which is the same as the probability the server is busy. Important performance measures are

$$L_q = the\ long\ run\ average\ number\ of\ customers\ waiting\ in\ queue$$

$$W_q = the\ long\ run\ average\ time\ spent\ per\ customer\ in\ queue$$

The Pollaczek-Khinhin mean value formula states that

$$W_q = \frac{\lambda E(S^2)}{2(1-\rho)} \tag{3.5.1}$$

This formula also implies an explicit expression for L_q by Little's Law

$$L_q = \lambda W_q \tag{3.5.2}$$

The Pollaczek-Khinhin formula gives not only an explicit expression for W_q, but more importantly it gives useful qualitative insights as well. It shows that the average delay per customer in the M/G/1 queue uses the service time distribution only through its first two moments.

A rigorous derivation of the Pollczek-Khinhin formula can be given by using the powerful generating-function approach. Define first the random variables

$L(t) =$ the number of customers at time t

$Q_n =$ the number of customers present just after the nth service

$L_n =$ the number of customers present just before the nth arrival epoch

The processes $\{L(t)\}, \{Q_n\}$ and $\{L_n\}$ are regenerative stochastic processes with finite expected cycle lengths. Denote the corresponding limiting distributions by

$\mathrm{p_j} = \lim_{t\to\infty} \mathrm{P(L(t) = j)}$, $q_j = \lim_{n\to\infty} P(Q_n = j)$ and $\pi_j = \lim_{n\to\infty} P(L_n = j)$ $j = 0,1,\ldots$

It is possible to deduce the existence of the limiting distributions from theorem 3.2.4. It is also possible to view the limiting probabilities as long term averages. For instance, $p_j * 100\%$ is the long run average percent of time there are j customers in the system. It holds the following significant identity:

$$\boldsymbol{\pi_j = p_j = q_j,\ \ j = 0, 1, \ldots} \tag{3.5.3}$$

Because the arrival process is a Poisson process, from theorem 3.2.4, the relation

$\pi_j = p_j$ is readily verified. To prove that the equality $\pi_j = q_j$, define the random variable $L_n^{(j)}$ as the number of customers over the first n arrivals who see j other customers present upon arrival and define the random variable $Q_n^{(j)}$ as the number of service completion epochs over the first n service completions at which j customers are left behind. Customers come separately and are served individually. Therefore, between (3.5.4) any two arrivals that find j other customers present there must be a service completion at which j customers are left behind and between any two service completions at which j customers are left behind there must be an arrival that sees j other customers present. By this up- and downcrossing argument, we have for each j that

$$\left|L_n^{(j)} - Q_n^{(j)}\right| \leq 1, \quad \mathrm{n} = 1,2,\ldots$$

Consequently, $\pi_j = \lim_{n\to\infty} L_n^{(j)}/n = \lim_{n\to\infty} Q_n^{(j)}/n = q_j$ for all j. We are now ready to prove that

$$\lim_{n\to\infty} E(z^{Q_n}) = \frac{(1-z)q_0 A(z)}{A(z)-z}$$

where

$$\mathrm{A}(z) = \int_0^\infty e^{-\lambda t(1-z)} b(t)dt$$

or in other words

$$\varphi(z) = \sum_{j=0}^{\infty} q_j z^j = (1-\rho)\frac{(1-z)\tilde{G}(\lambda-\lambda \mathrm{z})}{\tilde{G}(\lambda-\lambda \mathrm{z})-z} \, (3.5.4*) \text{ where } \tilde{G}(s) = \int_0^\infty e^{-st} dG(t)$$

with $\mathrm{b(t)}$ denoting the probability density of the service time of a customer. Before proving this result, we note that the unknown q_0 is determined by the fact that the left-hand side of (3.5.4) equals 1 for z = 1. By applying L'Hospital's rule, we find $\mathrm{q}_0 = 1-\rho$, in agreement with Little's Law $1 - p_0 = \rho$. By the bounded convergence theorem we have

$$\lim_{n\to\infty} E(z^{Q_n}) = \lim_{n\to\infty} \sum_{j=0}^{\infty} P(Q_n = j) z^j = \sum_{j=0}^{\infty} q_j z^j, \quad |z| \leq 1.$$

Hence by (3.5.3), (3.5.4),

(3.5.5)

$$\sum_{j=0}^{\infty} p_j z^j = \frac{(1-\rho)(1-z)A(z)}{A(z)-z}$$

Since the long run average queue size L_q is given by

$$L_q = \sum_{j=1}^{\infty}(j-1)p_j = \sum_{j=0}^{\infty} jp_j - (1-p_o) \tag{3.5.6}$$

the Pollaczek-Khinhin formula for L_q follows by differentiating the right-hand side of (3.5.5) and taking z = 1 in the derivative. It remains to prove (3.5.4). To do so, note that

$$Q_n = Q_{n-1} - \delta(Q_{n-1}) + A_n, \quad \mathrm{n} = 1,2,\ldots$$

where $\delta(x) = 1$ for $x > 0$, $\delta(x) = 0$ for $x = 0$ and A_n is the number of customers arriving during the nth service time. By the law of total probability,

$$\mathrm{P}(A_n = k) = \int_0^{\infty} e^{-\lambda t}\frac{(\lambda t)^k}{k!} b(t)dt, \quad k = 0,1,\ldots$$

and so

$$\sum_{k=0}^{\infty} P(A_n = k)z^k = \int_0^{\infty} e^{-\lambda t(1-z)} b(t)dt.$$

Since the random variables $Q_{n-1} - \delta(Q_{n-1})$ and A_n are independent of each other,

$$\mathrm{E}(z^{Q_n}) = \mathrm{E}\left(z^{Q_{n-1}-\delta(Q_{n-1})}\right)\mathrm{E}(z^{A_n}).$$

We have

$$\mathrm{E}\left(z^{Q_{n-1}-\delta(Q_{n-1})}\right) = \mathrm{P}(Q_{n-1} = 0) + \sum_{j=1}^{\infty} z^{j-1} P(Q_{n-1} = j)$$

$$= \mathrm{P}(Q_{n-1} = 0) + \frac{1}{z}[E(z^{Q_{n-1}}) - P(Q_{n-1} = 0)].$$

Substituting this in (3.5.6), we find

$$\mathrm{zE}(z^{Q_n}) = [E(z^{Q_{n-1}}) - (1-z)P(Q_{n-1} = 0)]\mathrm{A}(z).$$

Letting n → ∞, we next obtain the desired result (3.5.4). This completes the proof. Before concluding this section, we give an application of the Pollaczek-Khinhin formula

Example 3.5.1 The M/M/1 queue. If the service time distribution is

https://doi.org/10.26219/heal.aueb.6246

$$G(x) = 1 - e^{-\mu x}, \quad x \geq 0$$

the M/G/1 queue reduces to M/M/1 queue with $E(s) = 1/\mu$ and $\rho = \lambda/\mu$. $\tilde{G}(s)$ is the Laplace–Stieltjes transform at point s, thus

$$\tilde{G}(s) = \frac{\mu}{s + \mu}$$

Substituting in relation (3.5.4*) we get

$$\varphi(z) = (1 - \rho)\frac{(1 - z)\tilde{G}(\lambda - \lambda z)}{\tilde{G}(\lambda - \lambda z) - z}$$

$$= (1 - \rho)\frac{(1 - z)\mu/(\lambda + \mu - \lambda z)}{\mu/(\lambda + \mu - \lambda z) - z}$$

$$= (1 - \rho)\frac{(1 - z)\mu}{(1 - z)(\mu - \lambda z)}$$

$$= \frac{1 - \rho}{1 - \rho z}.$$

By expanding the last expression in a power series in z we get

$$\varphi(z) = \sum_{j=0}^{\infty} q_j z^j = \sum_{j=0}^{\infty} p_j z^j = (1 - \rho)\sum_{j=0}^{\infty} \rho^j z^j$$

Hence we get

$$p_j = (1 - \rho)\rho^j, \quad j \geq 0.$$

3.6 AN UP AND DOWNCROSSING TECHNIQUE

In this chapter, we analyze a commonly applicable up and downcrossing technique that can be used in partnership with the PASTA property to create connections, in queueing systems, between time average and customer average probabilities. We consider the G/M/1 queue to demonstrate that. In this case, customers arrive one at a time and the interarrival times are i.i.d. random variables with common cdf G (with $G(0) = 0$) and mean $1/\lambda$. The arriving customers form a single queue in an infinite waiting room in front of a single server and demand independent and identical exponentially distributed service times with mean $= 1/\beta$. Let $X(t)$ be the number of customers in the system at time t and X_n the number of customers present just prior to the nth arrival epoch. Both the stochastic processes $\{X(t)\}$ and $\{X_n\}$ are regenerative. The regeneration cycles are the epochs in which the queue is considered vacant by an incoming customer. Without confirmation, it is claimed that the $\lambda/\beta < 1$ assumption implies that the processes have

 https://doi.org/10.26219/heal.aueb.6246

a finite mean cycle length. The time-average and customer-average probabilities p_j and π_j can then be defined by

$$p_j = the\ long - run\ fraction\ of\ time\ that\ j\ customers\ are\ present$$

$$\pi_j = the\ long - run\ fraction\ of\ customers\ who\ find\ j\ customers\ upon\ arrival$$

for $j = 0,1, \dots$. Time averages are averages over time, and customer averages are averages over customers. To be accurate, $p_j = \lim_{t\to\infty}(1/t)\int_0^t I_j(u)du$ and $\pi_j = \lim_{n\to\infty}(1/n)\sum_{k=1}^n I_k(j)$, where $I_j(t) = 1$ if j customers are present at time t and $I_j(t) = 0$ otherwise, and $I_n(j) = 1$ if j other customers are present just before the nth arrival epoch and $I_n(j) = 0$ otherwise.

The probabilities p_j and π_j are related to each other by

$$\lambda\pi_{j-1} = \beta p_j, \quad j = 1,2, \dots \tag{3.6.1}$$

The justification for this outcome is instructive and it relies on three observations. Before giving the three steps, let us say that the continuous time process $\{X(t)\}$ makes an upcrossing from state $j-1$ to state j if a customer arrives and finds $j-1$ other customers present. The process $\{X(t)\}$ makes a downcrossing from state j to state $j-1$ if the service of a customer is finished and $j-1$ other customers are left behind.

Observation 1: Since customers arrive individually and are served individually, the long run average number of upcrossings from $j-1$ to j per time unit equals the long run average number of downcrossings from j to $j-1$ per time unit. This follows by noting that in any finite time interval the number of upcrossings from $j-1$ to j and the number of downcrossings from j to $j-1$ can differ at most by 1.

Observation 2: The long run fraction of customers seeing $j-1$ other customers upon arrival is equal to

$$\frac{the\ long\ run\ average\ number\ of\ upcrossings\ from\ j\ -\ 1\ to\ j\ per\ time\ unit}{the\ long\ run\ average\ number\ of\ arrivals\ per\ time\ unit}$$

for $j = 1,2, \dots$. In other words, the long run average number of upcrossings from $j-1$ to j per time unit equals $\lambda\pi_{j-1}$.

In fact, the above relation for fixed j is a special case of the Little's Law by assuming that each customer is charged 1€ by finding $j-1$ other customers present on arrival.

https://doi.org/10.26219/heal.aueb.6246

Observation 3: For exponential services, the long run average number of downcrossings from j to $j-1$ per time unit equals βp_j with probability 1 for each $j \geq 1$. The proof of this result relies on PASTA property. To make this simple, fix j and note that as long as the server is busy, service completions occur according to a Poisson process with rate β. Equivalently, we can assume that an exogenous Poisson process generates events at a rate of β, where a Poisson event results in a service completion only when there are j customers present. So, from PASTA propert,

$$\beta E\left[I_j(t)\right] = \mathrm{E}\left[D_j(t)\right] \text{ for t} > 0 \tag{3.6.2}$$

for any $j \geq 1$, where $I_j(t)$ is defined as the amount of time that j customers are present during $(0, t]$ and $D_j(t)$ is defined as the number of downcrossings from j to $j-1$ in (0, t]. Letting the constant d_j. denote the long run average number of downcrossings from j to $j-1$ per time unit, we have by the renewal reward theorem that $\lim_{t\to\infty} D_j(t)/t = d_j$ with probability 1. Similarly, $\lim_{t\to\infty} I_j(t)/t = p_j$ with probability 1. The renewal reward theorem also holds in the expected value version. Thus, for any $j \geq 1$,

$$\lim_{t\to\infty} \frac{E\left[D_j(t)\right]}{t} = d_j \quad \text{and} \quad \lim_{t\to\infty} \frac{E\left[I_j(t)\right]}{t} = p_j.$$

Hence relation (3.6.2) gives that $d_j = \beta p_j$ for all $j \geq 1$, as was to be proved.

https://doi.org/10.26219/heal.aueb.6246

CHAPTER 4

CONCLUSIONS

This dissertation is divided into two basic chapters. The first chapter deals with the Poisson process and its modifications. More specifically, we have sufficiently analyzed the memoryless property of the exponential distribution and its consequences in the Poisson process. We also studied the superposition and decomposition of this stochastic process, the compound Poisson process, the non-stationary case and the case where we have group of arrivals of events. The mathematical results of the above notions have many applications in many fields such as telecommunications, transport, finance, information technology, etc. We mainly dealt with examples of queue systems and inventory management. In the second chapter we dealt with the generalization of the Poisson process, the renewal processes, where events occur not after exponential time but after some time that follows a general distribution. This generalization allows us to analyze more problems but has a high cost in terms of computing and mathematics in the sense that now that mathematical formulas are more complicated, they are more difficult to understand and may not be able to be written in closed form. Then we incorporated a profit-loss structure, so our main concern was to calculate the expected long-term average cost in an evolving process. Fortunately, the elementary reward theorem simplifies such problems a lot. Furthermore, in this chapter, we worked on similar applications such as inventory management, control of queue systems and the calculation of the average expected cost when maintaining a machine.

https://doi.org/10.26219/heal.aueb.6246

https://doi.org/10.26219/heal.aueb.6246

References

1) H.C. Tijms (2003), *A First Course in Stochastic Models*, Wiley, Chichester
2) Sheldon M. Ross, *Stochastic Processes*, John Wiley & Sons, Inc., 1996
3) V. Kulkarni. Modeling and Analysis of Stochastic Systems. Chapman & Hall, 1995
4) J. D. C. Little and S. Graves, Little’s Law, 2008.
5) Δ. Φακίνος, *Στοχαστικά μοντέλα στην επιχειρησιακή έρευνα,* 2007.
6) Δ. Φακίνος, *Ουρές αναμονής*, 2008.
7) ADELSON, R. M., "Compound Poisson Distributions", Oper. Res. Quart., 17 (1966)

https://doi.org/10.26219/heal.aueb.6246