Abstract

We study in this thesis three subjects which are:

Cumulative Renewal Processes, Stochastic Control and Gradient Estimation.

The first subject is inspired from Smith's work on cumulative processes and the early part of Glynn and Heidelberger's paper on bias properties of budget constrained simulations.

We develop for this class of cumulative processes, new methods for evaluating explicitly and asymptotically arbitrary moments of a product of n distinct cumulative vector renewal reward processes termed $N(t)$ and $N(t)+1$ cases. A combinatorial approach is used to derive such moments. The analysis developed hinges on an expression of the moments in terms of the cumulants of the underlying time renewal sequence, and is founded on the recognition that certain sets of random variables are conditionally exchangeable. This gives rise to equivalence classes for the case $N(t)$, and to expectation summable classes for the case $N(t)+1$. Besides that, the theory of martingales is used for the case $N(t)+1$, where we generalize asymptotically Wald's fundamental equation in the discrete time.

The second subject is devoted to the study of optimal control problems of linear stochastic continuous-time systems, when the continuous time domain is decomposed into a finite set of N disjoint random interval of the form $[t_i, t_{i+1})$, where a complete state observation is taken at each instant $t_i, 0 \leq i \leq N-1$. We consider two optimal control problems termed (piecewise) time invariant control and time variant control. Concerning the observation point process, we consider first, the general situation where the increment intervals are *i.i.d.r.v.s* with unspecified probabilistic distribution. Second, the exponential distribution is considered. In this context, optimal control law are obtained for both control problems.

The third subject is concerned with the interplay between gradient estimation and ratio estimation. Given unbiased estimators for the numerator and the denominator of a ratio, as well as their gradients, joint central limit theorems for the ratio and its gradient are derived. The resulting confidence regions are of potential interest when optimizing such ratios numerically, as for sensitivity analysis, with respect to parameters whose exact value is unknown. We discuss briefly low-bias estimation for the gradient of a ratio.

Résumé

Nous étudions dans cette thèse trois sujets qui sont:

Processus Cumulatifs de Renouvellement, Contrôle Stochastique et Estimation du Gradient.

Le premier sujet est inspiré du travail de Smith sur les processus cumulatifs et la partie du début de l'article de Glynn et Heidelberger sur les propriétés du biais des simulations avec contraintes de budget. Nous développons pour cette classe de processus cumulatifs, des nouvelles méthodes pour évaluer explicitement et asymptotiquement des moments arbitraires d'un produit de n processus cumulatifs de renouvellement avec gain à valeur vectorielle appelés les cas $N(t)$ et $N(t)+1$. Une approche combinatoire est mise à contribution pour obtenir de tels moments. L'analyse développée dépend des expressions des moments en termes des cumulants de la suite des temps de renouvellement sous-jacents, et est basée sur la reconnaissance que certains ensembles de variables aléatoires sont conditionnellement interchangeables. Ceci donne lieu à des classes d'équivalences pour le cas $N(t)$, et à des classes sommables d'espérances pour le cas $N(t)+1$. En plus, la théorie des martingales est utilisée pour le cas $N(t)+1$, où nous généralisons asymptotiquement l'équation fondamentale de Wald à temps discret.

Le deuxième sujet est voué à l'étude des problèmes de contrôle optimal des systèmes stochastiques à temps continu, lorsque celui-ci est décomposé en un ensemble fini de N intervalles aléatoires disjoints de la forme $[t_i, t_{i+1})$, où une observation complète de l'état est prise à chaque instant $t_i, 0 \leq i \leq N-1$. Nous considérons dans ce cadre, deux problèmes de contrôle appelés contrôle invariant par rapport au temps (par morceaux) et contrôle variant par rapport au temps. Au sujet du processus ponctuel d'observations, nous considérons en premier lieu, la situation générale lorsque les accroissements des intervalles sont des *v.a.i.i.d.* sans spécifier la loi probabiliste. Ensuite, la distribution exponentielle est considérée. Dans ce cadre, les lois de contrôle optimal sont obtenues pour les deux problèmes.

Le troisième sujet traite des effets combinés entre l'estimation du gradient et l'estimation du quotient. Ayant des estimateurs non-biaisés pour le numérateur et le dénominateur d'un quotient, de même que pour leurs gradients, des théorèmes limites centraux conjoints pour le quotient et son gradient sont obtenus. Les régions de confiance qui en résultent sont d'un intérêt potentiel lorsqu'on optimise de tels quotients numériquement, ou pour analyse de sensitivité, par rapport aux paramètres dont la valeur exacte est inconnue. Nous discutons brièvement de l'estimation du faible-biais pour le gradient du quotient.

Dedications

To

My mother

The memory of my father

My brothers and sisters

and

Thérèse

who never lost faith

Acknowledgements

First, I would like to sincerely thank Professor Peter E. Caines my thesis supervisor for his excellent guidance, availability for lively and lovely discussions, humanity and kindness during these years. This thesis would have never been possible without your outstanding human dignity and exceptional comprehension for all. I am sincerely indebted to you.

Second, I would like to warmly thank the co-supervisor of this thesis, Professor Roland P. Malhamé for his confidence, animated and good debates, advice and hospitality at the École Polytechnique de Montréal. Thanks for your kindness and availability for passionated and enlightning discussions. This thesis is a result of your cumulative and renewal help during these years, and your marvellous understanding for all.

"Dear Peter and Roland, my deep and sincere thanks to you, Merci de tout coeur."

We, Professors Caines, Malhamé and myself share most of the fundamental ideas presented in this thesis.

I would like to thank very deeply my mother for her sincere and constant encouragements, particularly in the difficult time I lived through. You are one of the few people who have shared my pains and joys during these years. Thanks mother for your prayers and patience. This thesis would have never seen the day without your moral support. I am deeply indebted to you forever.

I wish also to deeply thank and from the bottom of my heart, my brothers and sisters for their confidence, moral support and fraternal love.

I sincerely thank Atef Harb for his fraternal friendship, good discussions, advice and confidence during these years.

I would like to warmly thank Professor Alain Haurie for providing me a financial support during the first years of this thesis. I am also grateful to the University of Québec at Montréal for having received, earlier during this thesis, the grant of excellence awarded to the teachers of this institution.

I would also like to sincerely thank Professor Pierre L'Ecuyer for having suggested the paper of Glynn and Heidelberger: "On Bias Propertier of Budget Constrained Simulations". Chapter 2 in this thesis is inspired from the early part of this article.

Over these years, I benefited from fruitful discussions with Professors Serge Alalouf, Jean-Pierre Dion and Gilbert Labelle from the University of Québec at Montréal, to whom I would like to address my deep thanks. In particular, my sincere and deep thanks to Professor Gilbert Labelle "Cher Gilbert, ton génie mathématique et tes qualités humaines exceptionnelles m'émerveillent depuis et pour toujours, merci du fond du coeur pour tout".

I warmly thank my colleague, Carlos Martínez-Mascarúa for his kindness and help to finalize the edition of the thesis.

Finally my sincere thanks to Ms Nicole Paradis and Francine Benoît from GERAD, and to Ms Mindle Levitt from McGill University for the excellent Tex work in producing the manuscripts of these chapters included in this thesis.

Claims of Originality

The following contributions are made in this thesis:

- Expectation of a product of summations of n distinct cumulative vector renewal reward processes, termed $N(t)$ and $N(t)+1$, are considered and analyzed in full detail.
- The product of random variables occurring in $P_n(t) = \prod_{\ell=1}^{n} \sum_{i=1}^{N(t)} Y_i^{(\ell)}$ is partitioned into equivalence classes for the case $N(t)$.
- An asymptotic expression for $\mathrm{E}[P_n(t)]$ is obtained, and a recursive scheme is given for generating monomials occurring in $\mathrm{E}[P_n(t)]$.
- The product of random variables occurring in $P_n^*(t) = \prod_{\ell=1}^{n} \sum_{i=1}^{N(t)+1} Y_i^{(\ell)}$ is partitioned into expectation summable classes for the case $N(t)+1$.
- An asymptotic expression for $\mathrm{E}[P_n^*(t)]$ is obtained, thus generalizing asymptotically Wald's fundamental equation in discrete time, and a recursive scheme is given for generating monomials in $\mathrm{E}[P_n^*(t)]$.
- Linear stochastic continuous systems are considered where the time domain is decomposed into a finite set of disjoint half-open random intervals, where observations are taken at the initial instant of each interval. For such systems, two control structures are considered; piecewise-time invariant and time variant controls. Optimal control laws are obtained for both cases.
- A confidence interval methodology for estimating the partial derivative of a ratio is developed and a joint central limit theorem for the simultaneous estimation of the entire gradient is obtained. A low bias estimation for the gradient of a ratio is given.

Contributions

My contribution to Chapters 2 through 4, co-authored with Professor Roland P. Malhamé (see the corresponding referenced papers), is of the order of 65%; in particular, I am entirely responsible for the combinatorics work in those chapters. Chapter 4 is mostly my own work; a few discussions were carried out with Professor R.P. Malhamé (see again the corresponding referenced paper).

My contribution to the co-authored work, with Professors Peter E. Caines and Roland P. Malhamé (see again the corresponding referenced paper), is of the order of 50%.

Finally, my contribution to the work of Chapter 6 is of the order of 35%.

TABLE OF CONTENTS

CHAPTER 1

Introduction

This thesis is the result of several papers in the area of applied probability and related topics. There are three subjects addressed in this thesis:

First: Cumulative Renewal Processes

Second: Stochastic Control

Third: Gradient Estimation

Chapters 2, 3 and 4 are motivated by applications of renewal reward processes in the areas of electrical engineering as well as in management science.

Indeed, the random variables introduced and analyzed in these chapters, symbolically written as $Y_i^{(\ell)}$, have a concrete interpretation and diverse practical applications. For example, consider the context where we have a dam, to which we associate random variables $Y_i^{(1)}, Y_i^{(2)}$ and $Y_i^{(3)}$ respectively representing the amount of rainfall, water extracted and electricity sale during the i-th subinterval of time.

The length of each subinterval is defined by the time between the onset of two successive rainfalls. We observe these variables in the interval $[0, t)$. Let $N(t)$ be the number of renewal times (number of distinct storms) in that interval; clearly $N(t)$ is a random integer variable which is known as a counting process.

It is of particular interest to evaluate the joint statistics of total (cumulative) rainfall, water extracted and electricity sale in the interval $[0, t)$; symbolically expressed as: $E\left[\sum_{i=1}^{N(t)} Y_i^{(1)}\right]$, $E\left[\sum_{i=1}^{N(t)} Y_i^{(2)}\right]$ and $E\left[\sum_{i=1}^{N(t)} Y_i^{(3)}\right]$; for estimation, forecasting and hydro-planing purposes.

While, ideally probability distribution functions would be highly desirable, they are difficult to obtain. However, much information is contained in different moments of these variables: averages, variances, correlation coefficients. Furthermore, despite the fact that it is the total amounts on interval $[0, t)$ which are of immediate interest, summations on $N(t)$ renewal cycles and $(N(t) + 1)$

renewal cycles respectively constitute, lower and upper bounds on these moments, which are easier to characterize. Thus Chapter 3 is concerned with the $\sum_{i=1}^{N(t)+1} Y_i^{(1)}$, $\sum_{i=1}^{N(t)+1} Y_i^{(2)}$ and $\sum_{i=1}^{N(t)+1} Y_i^{(3)}$ type variables.

In Chapters 2 and 3, we develop a rigorous theory and efficient method to asymptotically evaluate the above moments and more complicated forms.

We elaborate in Chapter 4 on a linearization technique to compute $E\left[\sum_{i=1}^{N(t)+1} Y_i^{(1)} \sum_{i=1}^{N(t)+1} Y_i^{(2)} \sum_{i=1}^{N(t)+1} Y_i^{(3)}\right]$, which seems to be very tedious. Consequently, one can appreciate the efficiency of the alternative methods developed in Chapters 2 and 3.

Specifically, Chapters 2, 3 and 4 are inspired from the work of Smith ([10, [11]) and Murthy [9] on cumulative processes, and the early part of Glynn and Heidelberger's paper on bias properties of budget constrained simulations [8]. Briefly, their article is concerned with the problems of analyzing and producing low bias estimates from Monte Carlo simulations, within a budget constraint t that represents the maximum amount of computer time allocated to the simulation.

We study in these chapters cumulative vector renewal reward processes. We develop there new methods for evaluating explicitly and asymptotically the expectation of a product of n distinct cumulative vector renewal reward processes. We consider two major classes of such processes termed $N(t)$ and $N(t)+1$.

We study respectively the cases $N(t)$ and $N(t)+1$ in Chapters 2 and 3, where we develop a combinatorial approach to derive explicit and asymptotic expressions for arbitrary moments of cumulative vector renewal reward processes. The analysis developed hinges on an expression of the moments in terms of the cumulants of the underlying time renewal sequence, and is founded on the recognition that certain sets of random variables are conditionally exchangeable. This gives rise to equivalence classes for the case $N(t)$, and to expectation summable classes for the case $N(t)+1$. Consequently, Smith's [11] asymptotic theory of cumulants is applied. Besides that, we need the theory of martingales for the case $N(t)+1$, to overcome the difficulty inherent to the analysis of the last renewal cycles involved in the summation part, which are probabilistically quite different from the other variables included in that summation part. This is the so-called renewal paradox. Chapters 2 and 3, are published as "Cahiers du GERAD" G-94-32 [2] and G-97-34 [3], and will be submitted respectively to Mathematics of Operations Research and to Advances in Applied Probability.

As we said, in Chapter 4 we elaborate on a standard linearization technique using the characteristic function, to evaluate explicitly the expectation of a product of triple summations of cumulative processes, as the time horizon goes to infinity. This chapter provides an independent check on the validity of the methods developed in Chapters 2 and 3. This chapter is published as "Cahier du GERAD" G-94-13 [1].

The motivation behind Chapter 5 is the inspection paradigm, where measurements of stochastic systems for example involve certain costs. In this chapter, we have the context where, one has to implement a closed-loop control of a stochastic system and random mode of sampling is chosen as a means of observing it. The total number of observations is fixed, but the control horizon is random. The solution of the linear quadratic Gaussian regulator is generalized to this situation.

Specifically, in Chapter 5 we study optimal control problems of linear stochastic continuous-time systems, when the continuous time domain is decomposed into a finite set of N disjoint random intervals of the form $[t_i, t_{i+1}), 0 \leq i \leq N-1$, at the start of each a state observation is taken. We consider two optimal control problems termed (piecewise) time invariant control and time variant control. Due to the information (i.e. observation structure), on each interval a state space system with complete initial state observation is defined. Concerning the observation point process, we, first, consider the general situation where the increment intervals are *i.i.d.r.v.s* with unspecified probabilistic distributions. Second, the exponential distribution is considered. In this context, optimal control laws are obtained for both control problems. The class of problems studied in this chapter are open to generalization to problems which appear to be significantly more difficult in particular, there is the case where the total number of observations N is random. There is also a possible link between that case and Chapters 2 and 3. A scalar version for the time variant case, which is the scalar example 5.1 of Chapter 5 in this thesis, has been presented for the Cyprus conference [4]. Chapter 5 will be submitted to IEEE Transactions on Automatic Control.

Chapter 6 is concerned with the interplay between gradient estimation and ratio estimation. Given unbiased estimators for the numerator and the denominator of a ratio, as well as their gradients, joint central limit theorems for the ratio and its gradient are derived. The resulting confidence regions are of potential interest when optimizing such ratios numerically or for sensitivity analysis with respect to parameters whose exact value is unknown. This chapter also briefly discusses low-bias estimation for the gradient of a ratio. There is a potential link between Chapters 5 and 6. Indeed, on the one hand, we showed in Sections 4 and 5 of Chapter 5, that the optimal cost-to-go of an arbitrary stage is expressed as infinite horizon discounted cost. On the other hand, it was shown by Fox and Glynn [6], that the infinite horizon discounted cost of a regenerative process can also be expressed in terms of an appropriately chosen ratio estimation problem. An earlier version of this chapter was given at the Optimization Days [5]. Chapter 6 is the theoretical part of a proceedings paper which was presented at the Winter Simulation Conference [7].

References

[1] Adès, M. and Malhamé, R.P. (1994). On the Moments of Cumulative Processes: A Preliminary Study. Les Cahiers du GERAD G-94-13, Montréal.

[2] Adès, M. and Malhamé, R.P. (1994). Asymptotics of the Moments of Cumulative Vector Renewal Reward Processes: The Case $N(t)$. Les Cahiers du GERAD G-94-32, Revised July 1997, Montréal.

[3] Adès, M. and Malhamé, R.P. (1997). Asymptotics of the Moments of Cumulative Vector Renewal Reward Processes: The Case $N(t)+1$. "Generalization of Wald's Fundamental Equation in the Discrete Time: An Asymptotic Study". Les Cahiers du GERAD G-97-34, Revised July 1997, Montréal.

[4] Adès, M., Caines, P.E. and Malhamé, R.P. (1997). Linear Gaussian Quadratic Regulation under Poisson Distributed Intermittent State Observations. Fifth IEEE Mediterranean Conference on Control and Systems, July 1997, Cyprus.

[5] Adès, M., Glynn, P.W. and L'Ecuyer, P. (1991). Confidence Intervals for Likelihood Ratio Derivative Estimators Over Infinite-Horizon: Discounted and Undiscounted Cases. Optimization Days, May 1991, Montréal.

[6] Fox, B.L. and Glynn, P.W. (1989). Simulating Discounted Costs. Management Science, 35, 1297-1315.

[7] Glynn, P.W., L'Ecuyer, P. and Adès, M. (1991). Gradient Estimation For Ratios. Proceedings of the Winter simulation Conference, 986-993, December 1991, Phoenix, Arizona.

[8] Glynn, P.W. and Heidelberger, P. (1990). Bias Properties of Budget Constrained simulations. Operations Research, 39, 801-814.

[9] Murthy, V.K. (1974). The General Point Processes. Addison-Wesley, Massachusetts.

[10] Smith, W.L. (1955). Regenerative Stochastic Process. Proceedings of the Royal Society A 232, 6-31.

[11] Smith, W.L. (1959). On the cumulants of Renewal Process. Biometrika, 46, 1-29.

CHAPTER 2

Asymptotics of the Moments of Cumulative Vector Renewal Reward Processses: The Case $N(t)$

1. Introduction: Classical Definitions and Notations

Consider a renewal sequence $\{t_i\}$ $i = 1, 2, \ldots$ of i.i.d. non negative variables (time intervals). To each t_i, we associate a random vector function $Y_i \in \mathbb{R}^d$, $d \geq 1$, which in general depends on t_i. We assume that $[t_i, Y_i']$, $i \geq 1$, is a sequence of i.i.d. random vectors (the components of Y_i could be interdependent), where $'$ indicates vector transposition. The general problem of interest here is that of obtaining asymptotic expression of $E[P_n(t)]$ as time t increases, where the products $P_n(t)$, $n \geq 1$, are of the form: $P_n(t) = \prod_{\ell=1}^{n} \Big(\sum_{i=1}^{N(t)}\Big) y_i^{(\ell)}$, (ℓ) refers to the ℓ^{th} component of the Y_i vector, $N(t)$ is the random integer such that $\sum_{i=1}^{N(t)} t_i \leq t \leq \sum_{i=1}^{N(t)+1} t_i$. Such sums appear in the study of cumulative processes (Smith [**16**]).

Interest in this problem stems from the relation that exists between the moments of a random variable vector and its (possibly multivariate) characteristic function under the form of a Taylor's expansion around the origin. Thus, for example, if the sequence of Y_i's is a scalar sequence, knowledge of the asymptotic behavior of $E\Big[\sum_{i=1}^{N(t)} Y_i\Big]^p$, $p = 1, 2, 3, \ldots$ could permit sharper estimates of the asymptotic distribution of $\Big[\sum_{i=1}^{N(t)} Y_i\Big]$ via inversion of rational approximants of the partial Taylor series expansion of the associated characteristic function $\phi(\alpha)$ around $\alpha = 0$, than could allow a central limit theorem based analysis.

We write:
$K_r^{(\ell)} = E\Big[[Y_i^{(\ell)}]^r\Big]$, $\mu_r = E[t_i^r]$, $\beta_{jk}^{(\ell)(m)} = E\Big[(Y_i^{(\ell)})^j (Y_i^{(m)})^k\Big]$, $\mu_{jk}^{(\ell)} = E\Big[t_i^j [Y_i^{(\ell)}]^k\Big]$, $\xi_{jkp}^{(\ell)(m)} = E\Big[t_i^j [Y_i^{(\ell)}]^k [Y_i^{(m)}]^p\Big]$; $\ell, m = 1, 2, \ldots, d$; $i, j, k, r, p = 1, 2, \ldots$.

We assume that the above moments are finite, and that the distribution of the $\{t_i\text{'s}\}$, $F(\cdot)$, is class $\mathcal{J}$ (Smith [17]), i.e., the class of distribution functions for which the k-fold convolution of the function with itself has an absolutely continuous component.

Obviously $E[P_n(t)]$ is the expectation of a product of summations of n (in general) distinct cumulative processes. For evaluating such expectations, we will require the concept of exchangeability of random variables. Also, we rely on Smith's [17] asymptotic analysis of factorial moments of $N(t)$, as well as the theory of combinatorics (Comtet [6]).

Indeed, inspired from the early part of Glynn and Heidelberger [9], our analysis hinges on recognizing that a number of products of random variables occurring in $P_n(t)$, have the same expectation when conditioned on $N(t)$. Subsequently we classify these random variables into conditional expectation equivalent classes, and the cardinality of each class is expressed in terms of $N(t)$. Finally, the asymptotic behaviour of the expectation is expressed in terms of the asymptotics of $N(t)$ as t goes to infinity.

This chapter is organized as follows. A reminder of the concept of exchangeability of random variables is presented in Section 2. A study of the asymptotic behaviour of $E[P_1(t)]$ and $E[P_2(t)]$ is given in Sections 3 and 4 respectively. The results are well known but these two sections constitute an illustration of the main steps of our methodology on relatively simple cases. We present in Section 5 useful elements of combinatorics and exchangeability results essential in partitioning the product of random variables in $P_n(t)$ into equivalence classes. In Section 6, we give an asymptotic expression for $E[P_n(t)]$. Detailed calculations for evaluating explicitly the asymptotic expression of $E[P_3(t)]$ are carried out in Section 7. Finally, a recursive scheme is given in Section 8 for generating $P_n(t)$ occurring in $E[P_n(t)]$. Note that Jensen's paper [10] addresses the same issues, but for $E\Big[\Big(\sum_{i=1}^{N(t)} Y_i\Big)^k\Big]$, and the analytical methods are quite different. Thus, this work is more general.

2. Preliminary Notions on Exchangeability

The random variables $\left(Y_1, Y_2, \ldots, Y_n\right)$ are said to be exchangeable if $\left(Y_{i_1}, Y_{i_2}, \ldots, Y_{i_n}\right)$ has the same joint distribution as $\left(Y_1, Y_2, \ldots, Y_n\right)$, whenever $(i_1, i_2, \ldots, i_n)$ is a permutation of $1, 2, \ldots, n$. That is, they are exchangeable if the joint distribution function $P(Y_1 \le y_1, Y_2 \le y_2, \ldots, Y_n \le y_n)$ is a symmetric function of $(y_1, y_2, \ldots, y_n)$; (Ross [14]). The concept of exchangeability was introduced by DeFinetti [8] in his classical paper in 1931. For more details and applications of this concept see e.g. Chow & Teicher [5], Kingman [11] and Koch & Spizzichino [12].

3. Asymptotic Form of $E\left[\sum_{i=1}^{N(t)} Y_i\right]$

We consider here the case where $N(t)$ and Y_i are dependent on each other, otherwise the problem of evaluating $E\left[\sum_{i=1}^{N(t)} Y_i\right]$ and $\text{Var}\left[\sum_{i=1}^{N(t)} Y_i\right]$ would be simple; see Cox [7].

Let $W_i' = (t_i, Y_i')$ and let $\left(\Pi(1), \ldots, \Pi(m)\right)$ be any permutation of the integers $1, 2, \ldots, m$. Then it can be shown that:

$$P\left[(W_1, \ldots, W_m) \in A | N(t) = m\right] = P\left[(W_{\Pi(1)}, \ldots, W_{\Pi(m)}) \in A | N(t) = m\right]. \tag{3.1}$$

See Glynn & Heidelberger [9] and Ross [14] for (3.1).

Therefore $\{(W_i), i = 1, \ldots, m\}$ are conditionally exchangeable random vectors; then following [9]:

$$E\left[Y_i | N(t) = m\right] = E\left[Y_1 | N(t) = m\right] \tag{3.2}$$

$$E\left[\sum_{i=1}^{N(t)} Y_i | N(t) = m\right] = \sum_{i=1}^{m} E\left[Y_i | N(t) = m\right] = mE\left[Y_1 | N(t) = m\right] \quad m \geq 1 \tag{3.3}$$

Using the properties of conditional expectations:

$$\begin{aligned} E\left[\sum_{i=1}^{N(t)} Y_i\right] &= E\left[E\left[\sum_{i=1}^{N(t)} Y_i | N(t)\right]\right] \\ &= E\left[N(t) E\left[Y_1 | N(t)\right] I\left(N(t) \geq 1\right)\right] \\ &= E\left[E\left[Y_1 N(t) | N(t)\right] I\left(N(t) \geq 1\right)\right] \\ &= E\left[Y_1 N(t)\right] \end{aligned} \tag{3.4}$$

where $I(\cdot)$ is the indicator function, and the last equality stems from the fact that $N(t) = 0$ if $I\left(N(t) \geq 1\right) = 0$. Under conditioning this time on t_1 we have using the conditional independence of Y_1, $N(t)$, given t_1:

$$E\left[Y_1 N(t)\right] = E\left[E\left[Y_1 | t_1\right] E\left[\left(1 + N(t - t_1)\right) I(t_1 < t) | t_1\right]\right] \tag{3.5}$$

In (3.5), one can utilize the asymptotic expression of $E\left[N(t) + 1\right]$ developed by Smith [17], as $\gamma_1(t - t_1) + \gamma_2 + \omega(t - t_1)$ where γ_1 and γ_2 are constants which can be calculated, and $\omega(t - t_1)$ is $o(1)$, for a given t_1.

Using Lemmas 5 and 6 in Section 6, we have respectively and asymptotically $E\Big[Y_1\Big(\gamma_1(t - t_1) + \gamma_2\Big)\, I(t_1 > t)\Big] = o(1)$ and $E[Y_1\omega(t - t_1)] = o(1)$. Thus, after substituting, Smith's asymptotic expressions, and getting rid of the $I(t_1 < t)$ factor, one obtains:

$$E\Big[\sum_{i=1}^{N(t)} Y_i\Big] = \frac{t\mathcal{K}_1}{\mu_1} - \frac{\mu_{11}}{\mu_1} + \frac{\mu_2\mathcal{K}_1}{2\mu_1^2} + o(1) \tag{3.6}$$

Note that Equation (3.6) coincides perfectly with Lemma 1 in Brown & Solomon [4] and Equation (3.3.6) in Murthy [13]. However, we have to underline that our approach, while not necessarily much easier than theirs, more easily generalizes to the bivariate $E[P_2(t)]$ and multivariate $E[P_n(t)]$ cases. Note also that a generalization of $E\Big[Y_1 E[1 + N(t)]\, I(t_1 > t)\Big] = o(1)$ will be required and proved in Lemma 5 of Section 6.

4. Asymptotic Form of $E\left[\sum_{i=1}^{N(t)} Y_i^{(1)} \sum_{i=1}^{N(t)} Y_i^{(2)}\right]$

In this section, we will deal with two types of scalar cumulative process.

Carrying out the ordinary multiplication of $\sum_{i=1}^{N(t)} Y_i^{(1)} \sum_{i=1}^{N(t)} Y_i^{(2)}$ results in the following two dimensional array valid for any $N(t) \geq 2$:

$$\begin{aligned}
& Y_1^{(1)}Y_1^{(2)} + Y_1^{(1)}Y_2^{(2)} + Y_1^{(1)}Y_3^{(2)} + \cdots + Y_1^{(1)}Y_{N(t)-1}^{(2)} + Y_1^{(1)}Y_{N(t)}^{(2)} \\
+ \; & Y_2^{(1)}Y_1^{(2)} + Y_2^{(1)}Y_2^{(2)} + Y_2^{(1)}Y_3^{(2)} + \cdots + Y_2^{(1)}Y_{N(t)-1}^{(2)} + Y_2^{(1)}Y_{N(t)}^{(2)} \\
+ \; & Y_3^{(1)}Y_1^{(2)} + Y_3^{(1)}Y_2^{(2)} + Y_3^{(1)}Y_3^{(2)} + \cdots + Y_3^{(1)}Y_{N(t)-1}^{(2)} + Y_3^{(1)}Y_{N(t)}^{(2)} \\
+ \; & \\
\vdots \; & \\
+ \; & Y_{N(t)-1}^{(1)}Y_1^{(2)} + Y_{N(t)-1}^{(1)}Y_2^{(2)} + Y_{N(t)-1}^{(1)}Y_3^{(2)} + \cdots + Y_{N(t)-1}^{(1)}Y_{N(t)-1}^{(2)} + Y_{N(t)-1}^{(1)}Y_{N(t)}^{(2)} \\
+ \; & Y_{N(t)}^{(1)}Y_1^{(2)} + Y_{N(t)}^{(1)}Y_2^{(2)} + Y_{N(t)}^{(1)}Y_3^{(2)} + \cdots + Y_{N(t)}^{(1)}Y_{N(t)-1}^{(2)} + Y_{N(t)}^{(1)}Y_{N(t)}^{(2)}
\end{aligned} \tag{4.1}$$

Note that from (4.1), we can observe two different cases; the first one is characterized by the random variables $Y_i^{(1)}Y_i^{(2)}$ for $i = 1, \ldots, N(t)$, which are i.i.d., and the other is derived from the random variables $Y_i^{(1)}Y_j^{(2)}, i \neq j$ where $i, j = 1, 2, \ldots, N(t)$, which may be dependent.

For this bivariate case, we wish to prove that:

$$\begin{aligned}
& E\Big[\sum_{i=1}^{N(t)} Y_i^{(1)} \sum_{i=1}^{N(t)} Y_i^{(2)}\Big] = E\Big[Y_1^{(1)}Y_1^{(2)} E\Big(N(t - t_1) + 1\Big) I(t_1 \leq t)\Big] \\
& + E\Big[Y_1^{(1)}Y_2^{(2)} E\Big(\{N(t - t_1 - t_2) + 2\}\{N(t - t_1 - t_2) + 1\}\Big) I(t_1 + t_2 \leq t)\Big]
\end{aligned} \tag{4.2}$$

To establish the above equation, we first show that

$$E\Big[\sum_{i=1}^{N(t)} Y_i^{(1)} \sum_{i=1}^{N(t)} Y_i^{(2)}\Big] = E\Big[N(t)Y_1^{(1)}Y_1^{(2)}I(t_1 \le t) + N(t)\big(N(t)-1\big)Y_1^{(1)}Y_2^{(2)}I(t_1+t_2 \le t)\Big] \tag{4.3}$$

Note that Equation (4.3) is based on a counting argument derived from the expansion in (4.1). Also we make use of the fact that under the prevailing assumptions we can show that:

$$E\Big[Y_i^{(1)}Y_i^{(2)}|N(t)=m\Big] = E\Big[Y_1^{(1)}Y_1^{(2)}|N(t)=m\Big] \quad \forall i = 1,2,\ldots,m \text{ for } m \ge 1 \tag{4.4}$$

and

$$E\Big[Y_i^{(1)}Y_j^{(2)}|N(t)=m\Big] = E\Big[Y_1^{(1)}Y_2^{(2)}|N(t)=m\Big] \quad \forall i \ne j;\ i,j = 1,2,\ldots,m \text{ for } m \ge 2 \tag{4.5}$$

Indeed, Equation (4.4) follows from (3.3) and assuming as we are about to show that (4.5) is true, then we would have for $m \ge 1$:

$$E\Big[\sum_{i=1}^{N(t)} Y_i^{(1)} \sum_{j=1}^{N(t)} Y_j^{(2)}|N(t)=m\Big] = mE\Big[Y_1^{(1)}Y_1^{(2)}|N(t)=m\Big] \tag{4.6}$$

and for $m \ge 2$:

$$\begin{aligned} E\Big[\sum_{i=1}^{N(t)} Y_i^{(1)} \sum_{j=1}^{N(t)} Y_j^{(2)}|N(t)=m\Big] &= \sum_{i=1}^{N(t)}\sum_{j=1}^{N(t)} E\Big[Y_i^{(1)}Y_j^{(2)}|N(t)=m\Big] \\ &= \sum_{i=1}^{N(t)} E\Big[Y_i^{(1)}Y_i^{(2)}|N(t)=m\Big] \\ &\quad + \sum_{i=1}^{N(t)}\sum_{j=1}^{N(t)} E\Big[Y_i^{(1)}Y_j^{(2)}|N(t)=m\Big] \text{ for } i \ne j \end{aligned} \tag{4.7}$$

Using Equations (4.4), (4.5) and (4.7) results for $m \ge 2$ in:

$$\begin{aligned} E\Big[\sum_{i=1}^{N(t)} Y_i^{(1)} \sum_{j=1}^{N(t)} Y_j^{(2)}|N(t)=m\Big] &= mE\Big[Y_1^{(1)}Y_1^{(2)}|N(t)=m\Big] \\ &+ m\big(m-1\big)E\Big[Y_1^{(1)}Y_2^{(2)}|N(t)=m\Big] \end{aligned} \tag{4.8}$$

Recognizing that:

$$E\Big[\sum_{i=1}^{N(t)} Y_i^{(1)} \sum_{j=1}^{N(t)} Y_j^{(2)}\Big] = E\Big[E\Big[\sum_{i=1}^{N(t)} Y_i^{(1)} \sum_{j=1}^{N(t)} Y_j^{(2)}|N(t)\Big]\Big] \tag{4.9}$$

and using (4.6) and (4.8) results in:

$$E\left[\sum_{i=1}^{N(t)} Y_i^{(1)} \sum_{j=1}^{N(t)} Y_j^{(2)}\right]$$
$$= E\Big[N(t)Y_1^{(1)}Y_1^{(2)}I\big(N(t) \geq 1\big) + N(t)\big(N(t)-1\big)Y_1^{(1)}Y_2^{(2)}I\big(N(t) \geq 2\big)\Big]$$

or equivalently:

$$E\left[\sum_{i=1}^{N(t)} Y_i^{(1)} \sum_{j=1}^{N(t)} Y_j^{(2)}\right] = E\Big[N(t)Y_1^{(1)}Y_1^{(2)}I(t_1 \leq t) + N(t)\big(N(t)-1\big)Y_1^{(1)}Y_2^{(2)}I(t_1+t_2 \leq t)\Big] \tag{4.10}$$

which is Equation (4.3). Note that (4.10) would still be valid without the indicator functions.

It remains to establish Equation (4.5), which essentially is a claim that, conditional on $N(t)$, the variables $Y_i^{(1)}Y_j^{(2)}, i = 1,\ldots,N(t), j = 1,\ldots,N(t)$, have the same expectation if $i \neq j$. Such variables will be termed *conditional expectation equivalent*. Note that conditional expectation equivalence is an equivalence relation. We establish (4.5), by first showing that under the prevailing assumptions, a given row or a given column of array (4.1) corresponds to conditionally exchangeable random variables, as long as $i \neq j$ in $Y_i^{(1)}Y_j^{(2)}$. Let us then state the following lemma:

LEMMA 1. *For $m \geq 2, E\Big[Y_i^{(1)}Y_j^{(2)}|N(t)=m\Big] = E\Big[Y_1^{(1)}Y_2^{(2)}|N(t)=m\Big]$ for $i \neq j$ and $i,j = 1,2,\ldots,m$.*

Proof. Define for $m \geq 2, Z_i = Y_1^{(1)}Y_i^{(2)}, i = 2,\ldots,m$, a sequence of random variables and let $\big(\Pi(2),\ldots,\Pi(m)\big)$ be any permutation of the integers $(2,\ldots,m)$. We have for $m \geq 2$, and for arbitrary scalars $Z_i, i = 1,\ldots,m$:

$$P\Big[Z_2 \leq z_2,\ldots,Z_m \leq z_m|N(t)=m\Big]$$
$$= \frac{P\Big[Y_1^{(1)}Y_2^{(2)} \leq z_2,\ldots,Y_1^{(1)}Y_m^{(2)} \leq z_m;\, t_1+t_2+\cdots+t_m \leq t < t_1+t_2+\cdots+t_m+t_{m+1}\Big]}{P\Big[t_1+t_2+\cdots+t_m \leq t < t_1+t_2+\cdots+t_m+t_{m+1}\Big]}$$
$$= \frac{E\Big[P[Y_2^{(2)} \leq z_2/y_1^{(1)},\ldots,Y_m^{(2)} \leq z_m/y_1^{(1)};\, t_2+\cdots+t_m \leq t-\tau_1 < t_2+\cdots+t_{m+1}|Y_1^{(1)}=y_1^{(1)},t_1=\tau_1]\Big]}{P\Big[t_1+t_2+\cdots+t_m \leq t < t_1+t_2+\cdots+t_m+t_{m+1}\Big]} \tag{4.11}$$
$$= \frac{E\Big[P\Big[Y_{\pi(2)}^{(2)} \leq z_2/y_1^{(1)},\ldots,Y_{\pi(m)}^{(2)} \leq z_m/y_1^{(1)};t_{\pi(2)}+\cdots+t_{\pi(m)} \leq t-\tau_1 < t_{\pi(2)}+\cdots+t_{\pi(m)}+t_{m+1}|Y_1^{(1)}=y_1^{(1)},t_1=\tau_1\Big]\Big]}{P\Big[t_1+t_{\Pi(2)}+\cdots+t_{\Pi(m)} \leq t < t_1+t_{\Pi(2)}+\cdots+t_{\Pi(m)}+t_{m+1}\Big]} \tag{4.12}$$
$$= \frac{E\Big[P\Big[Z_{\Pi(2)} \leq z_2,\ldots,Z_{\Pi(m)} \leq z_m;t_{\Pi(2)}+\cdots+t_{\Pi(m)} \leq t-\tau_1 < t_{\Pi(2)}+\cdots+t_{\Pi(m)}+t_{m+1}|Y_1^{(1)}=y_1^{(1)},t_1=\tau_1\Big]\Big]}{P\Big[t_1+t_{\Pi(2)}+\cdots+t_{\Pi(m)} \leq t < t_1+t_{\Pi(2)}+\cdots+t_{\Pi(m)}+t_{m+1}\Big]} \tag{4.13}$$

$$= \frac{P\left[Z_{\Pi(2)} \le z_2, \ldots, Z_{\Pi(m)} \le z_m;\, t_1 + t_{\Pi(2)} + \cdots + t_{\Pi(m)} \le t < t_1 + t_{\Pi(2)} + \cdots + t_{\Pi(m)} + t_{m+1}\right]}{P\left[t_1 + t_{\Pi(2)} + \cdots + t_{\Pi(m)} \le t < t_1 + t_{\Pi(2)} + \cdots + t_{\Pi(m)} + t_{m+1}\right]} \tag{4.14}$$

$$= P\left[Z_{\Pi(2)} \le z_2, \ldots, Z_{\Pi(m)} \le z_m | N(t) = m\right]$$

Note that the step from (4.11) to (4.12) is based on (3.1). Thus for $m \ge 2$, the conditional random variables $\left(Y_1^{(1)} Y_j^{(2)} | N(t) = m\right), j = 2, \ldots, m$ are exchangeable and in particular:

$$E[Y_1^{(1)} Y_j^{(2)} | N(t) = m] = E[Y_1^{(1)} Y_2^{(2)} | N(t) = m], j = 2, \ldots, m. \tag{4.15}$$

Using similar arguments for a column in the expansion of (4.1), i.e. if one considers the random variables $Y_i^{(1)} Y_2^{(2)}, i = 1, 3, \ldots, m$; it is possible to prove the interchangeability of the conditional random variables $\left(Y_i^{(1)} Y_2^{(2)} | N(t) = m\right)$ for $m \ge 2, i \ne 2$, and thus:

$$E[Y_i^{(1)} Y_2^{(2)} | N(t) = m] = E[Y_1^{(1)} Y_2^{(2)} | N(t) = m], i = 1, 3, \ldots, m,\; m \ge 2 \tag{4.16}$$

The lemma derives from (4.15) and (4.16).

Equation (4.10) now follows directly from (4.5).

From (4.10) we obtain:

$$\begin{aligned} E\left[\sum_{i=1}^{N(t)} Y_i^{(1)} \sum_{i=1}^{N(t)} Y_i^{(2)}\right] &= E\left[E\left[Y_1^{(1)} Y_1^{(2)} N(t) I(t_1 \le t) | t_1\right]\right] \\ &+ E\left[E\left[Y_1^{(1)} Y_2^{(2)} N(t)\left(N(t) - 1\right) I(t_1 + t_2 \le t) | t_1, t_2\right]\right] \end{aligned} \tag{4.17}$$

$$\begin{aligned} &= E\left[E\left[Y_1^{(1)} Y_1^{(2)} | t_1\right] E\left[N(t) I(t_1 \le t) | t_1\right]\right] \\ &+ E\left[E\left[Y_1^{(1)} Y_2^{(2)} | t_1, t_2\right] E\left[N(t)\left(N(t) - 1\right) I(t_1 + t_2 \le t) | t_1, t_2\right]\right] \end{aligned} \tag{4.18}$$

$$\begin{aligned} &= E\left[E\left[Y_1^{(1)} Y_1^{(2)} | t_1\right] E\left[\left(1 + N(t - t_1)\right) I(t_1 \le t) | t_1\right]\right] \\ &+ E\Big[E\left[Y_1^{(1)} Y_2^{(2)} | t_1, t_2\right] E\Big[\{N(t - t_1 - t_2) + 1\}\{N(t - t_1 - t_2) + 2\} \\ &\quad I(t_1 + t_2 \le t) | t_1, t_2\Big]\Big] \end{aligned} \tag{4.19}$$

4.1. Asymptotic Explicit Expression of $E\left[\sum_{i=1}^{N(t)} Y_i^{(1)} \sum_{i=1}^{N(t)} Y_i^{(2)}\right]$. Using Lemmas 5 and 6 in Section 6 we have asymptotically, $E\left[Y_1^{(1)} Y_1^{(2)} \phi_1(t - t_1) I(t_1 > t)\right] = o(1)$, $E\left[Y_1^{(1)} Y_2^{(2)} \phi_2(t - t_1 - t_2) I(t_1 + t_2 > t)\right] = o(1)$, $E\left[Y_1^{(1)} Y_1^{(2)} \omega_1(t_*^1)\right] = o(1)$ and $E\left[Y_1^{(1)} Y_2^{(2)} \omega_2(t_*^2)\right] = o(1)$, where $\phi_k(t) \triangleq E\left[\left(N(t)+1\right)\left(N(t)+2\right) \ldots \left(N(t)+k\right)\right]$, $k = 1, 2, \ldots$, are the cumulants of $N(t)$, $\omega_1(t_*^1) \triangleq \omega_1(t - t_1)$

and $\omega_2(t_*^2) \triangleq \omega_2(t - t_1 - t_2)$ are the remainders in the asymptotic expansions of $\phi_1(t_*^1)$ and $\phi_2(t_*^2)$ respectively, as established by Smith [17]. As in Section 3, we can replace the cumulants by their asymptotic polynomial expansion in terms of $(t - t_1)$ or $(t - t_1 - t_2)$, and ignore the action of the indicator functions.

Thus the asymptotic behaviour of $E\left[\sum_{i=1}^{N(t)} Y_i^{(1)} \sum_{i=1}^{N(t)} Y_i^{(2)}\right]$ derives directly from the asymptotics of $\phi(t_*^1)$ and $\phi(t_*^2)$ given in general form in Section 7. Using the fact that $Y_1^{(1)}$ and $Y_2^{(2)}$ are correlated respectively with t_1 and t_2, we find after some algebra:

$$\begin{aligned}
&E\left[\sum_{i=1}^{N(t)} Y_i^{(1)} \sum_{i=1}^{N(t)} Y_i^{(2)}\right] = \tfrac{t^2}{\mu_1^2} E[Y_1^{(1)}]E[Y_2^{(2)}] \\
&+t\left(\tfrac{E[Y_1^{(1)}Y_1^{(2)}]}{\mu_1} + \tfrac{2\mu_2^2}{\mu_1^3} E[Y_1^{(1)}]E[Y_2^{(2)}] - \tfrac{2}{\mu_1^2} E[Y_1^{(1)}t_1]E[Y_2^{(2)}] - \tfrac{2}{\mu_1^2} E[Y_2^{(2)}t_2]E[Y_1^{(1)}]\right) \\
&+\left(\tfrac{\mu_2}{2\mu_1^2} E[Y_1^{(1)}Y_1^{(2)}] - \tfrac{E[t_1 Y_1^{(1)}Y_1^{(2)}]}{\mu_1} - \tfrac{2\mu_2}{\mu_1^3} E[Y_1^{(1)}t_1]E[Y_2^{(2)}]\right. \qquad (4.20) \\
&-\tfrac{2\mu_2}{\mu_1^3} E[Y_1^{(1)}]E[Y_2^{(2)}t_2] + \tfrac{E[Y_1^{(1)}t_1^2]E[Y_2^{(2)}]}{\mu_1^2} + \tfrac{E[Y_2^{(2)}t_2^2]E[Y_1^{(1)}]}{\mu_1^2} \\
&\left.+2\tfrac{E[Y_1^{(1)}t_1]E[Y_2^{(2)}t_2]}{\mu_1^2} + \left(\tfrac{9\mu_2^2-4\mu_1\mu_3}{6\mu_1^4}\right)E[Y_1^{(1)}]E[Y_2^{(2)}]\right) + o(1)
\end{aligned}$$

If $Y_i^{(2)} \equiv Y_i^{(1)}$ then Equation (4.20) becomes

$$\begin{aligned}
&E\left[\sum_{i=1}^{N(t)} Y_i^{(1)}\right]^2 = \tfrac{t^2 K_1^2}{\mu_1^2} + t\left(\tfrac{K_2}{\mu_1} + \tfrac{2\mu_2^2 K_1^2}{\mu_1^3} - \tfrac{4\mu_{11}K_1}{\mu_1^2}\right) \\
&+\left(-\tfrac{4\mu_2\mu_{11}K_1}{\mu_1^3} + \tfrac{2\mu_{21}K_1}{\mu_1^2} + \tfrac{2\mu_{11}^2}{\mu_1^2} + \left(\tfrac{9\mu_2^2-4\mu_1\mu_3}{6\mu_1^4}\right)K_1^2\right. \qquad (4.21) \\
&\left.+\tfrac{3K_1^2\mu_2^2}{2\mu_1^4} - \tfrac{2K_1^2\mu_3}{3\mu_1^3} - \tfrac{\mu_{12}}{\mu_1} + \tfrac{\mu_2 K_2}{2\mu_1^2}\right) + o(1)
\end{aligned}$$

The result of $E\left[\sum_{i=1}^{N(t)} Y_i^{(1)}\right]^2$ is already known in the literature but has been obtained, however, using a different approaches amongst which a computation of the characteristic function of $E\left[\sum_{i=1}^{N(t)} Y_i^{(1)}\right]^2$; (Murthy [13]).

Murthy's approach becomes of exceeding complexity for the evaluation of $E\left[\sum_{i=1}^{N(t)} Y_i^{(1)}\right]^n$. We also mention Corollary 1 of Brown & Solomon [4] for $E\left[\sum_{i=1}^{N(t)} Y_i^{(1)}\right]^2$.

It seems nevertheless difficult to tackle the general case based on their paper. Before starting the next section and in order to facilitate understanding, let us give the analogue of Equation (4.10) for the trivariate case:

$$E\Big[\sum_{i=1}^{N(t)} Y_i^{(1)} \sum_{i=1}^{N(t)} Y_i^{(2)} \sum_{i=1}^{N(t)} Y_i^{(3)}\Big] = E\Big[N(t)Y_1^{(1)}Y_1^{(2)}Y_1^{(3)}\, I(t_1 \le t)$$

$$+N(t)\Big(N(t)-1\Big)Y_1^{(1)}Y_1^{(2)}Y_2^{(3)}\, I(t_1+t_2 \le t)$$

$$+N(t)\Big(N(t)-1\Big)Y_1^{(1)}Y_1^{(3)}Y_2^{(2)}\, I(t_1+t_2 \le t) \qquad (4.22)$$

$$+N(t)\Big(N(t)-1\Big)Y_1^{(1)}Y_2^{(2)}Y_2^{(3)}\, I(t_1+t_2 \le t)$$

$$+N(t)\Big(N(t)-1\Big)\Big(N(t)-2\Big)Y_1^{(1)}Y_2^{(2)}Y_3^{(3)}\, I(t_1+t_2+t_3 \le t)\Big]$$

On the R.H.S. of (4.22) we observe respectively from the first term to the last one the following:

a) $Y_1^{(1)}Y_1^{(2)}Y_1^{(3)}$ is one block of letters (or random variables) having the same lower index which is one.

b) $Y_1^{(1)}Y_1^{(2)}Y_2^{(3)}$ consists of two blocks which are $Y_1^{(1)}Y_1^{(2)}$ and $Y_2^{(3)}$. The cases of $Y_1^{(1)}Y_1^{(3)}Y_2^{(2)}$ and $Y_1^{(1)}Y_2^{(2)}Y_2^{(3)}$ are the same as that of $Y_1^{(1)}Y_1^{(2)}Y_2^{(3)}$.

c) $Y_1^{(1)}Y_2^{(2)}Y_3^{(3)}$ is considered as made-up of three blocks given by $Y_1^{(1)}, Y_2^{(2)}$ and $Y_3^{(3)}$. The lower indices are arranged in increasing order from one block to another.

Similar observations are applicable to Equations (4.10) and (3.4), the last one being the simplest case.

As we remark either in Equation (4.22) or (4.10) we have respectively a partition of 3 or 2 letters into distinct blocks. The generalization of this idea, i.e., a partition of n letters into m distinct blocks will be an essential notion in the forthcoming section.

5. Partitioning the Product of Random Variables in $P_n(t)$

Let us refer to n as the index in the product $\prod_{\ell=1}^{n}\Big(\sum_{i=1}^{N(t)} Y_i^{(\ell)}\Big)$. Thus in an index 1 case, we have $\sum_{i=1}^{N(t)} Y_i^{(1)}$; we are able here to distinguish one equivalence class (variables in a sum that have the same expectation conditional on $N(t)$). In the index 2 case, the number of distinct equivalence classes is inherited in part from the index 1 case, all $Y_i^{(1)}Y_i^{(2)}$ type variables, plus a new equivalence class specific to the index 2 when all indices are different i.e. $Y_i^{(1)}Y_j^{(2)}, i \ne j$ type variables. In the index 3 case, the equivalence classes are inherited in part from the index 2 case. Indeed we have all $Y_i^{(1)}Y_i^{(2)}Y_i^{(3)}$ type variables (one block, three letters, in this context, the number of blocks is given by the number of distinct lower (time) indices within a number of an equivalence class). Also $Y_i^{(1)}Y_i^{(2)}Y_j^{(3)}, Y_i^{(1)}Y_j^{(2)}Y_i^{(3)} \equiv Y_i^{(1)}Y_i^{(3)}Y_j^{(2)}$ and $Y_i^{(1)}Y_j^{(2)}Y_j^{(3)}, i \neq j$ type variables possess 2 blocks since $Y_i^{(1)}Y_i^{(2)}, Y_i^{(1)}Y_i^{(3)}$ and $Y_j^{(2)}Y_j^{(3)}$ behave as index 1 variables. Finally, we have a single new equivalence class $Y_i^{(1)}Y_j^{(2)}Y_k^{(3)}$ for i, j, k different, specific to the index 3 case, which obviously incorporates 3 blocks.

More generally, let $\prod^{(m,n)}$ be all partitions of n letters into m non-empty distinct blocks representative of all equivalence classes (conditional expectation equivalence) of random variables incorporating m blocks. Also let $\text{mon}\prod_\ell^{(m,n)}$ be the monomial associated with the ℓ^{th} distinct equivalence class with $\prod^{(m,n)}$. The lower indices from one block to another are arranged in increasing order, while the upper indices within each block are arranged in increasing order. The ordering of monomials within $\prod^{(m,n)}$ is lexicographic with the lower index order dominating the upper index order. Thus for the index 1 case, we have one monomial given by $\text{mon}\prod_1^{(1,1)} \triangleq Y_1^{(1)}$ representative of the equivalence class of $Y_i^{(1)}$ type variables. There are $N(t)$ type monomials represented by $\text{mon}\prod_1^{(1,1)}$. In the index 2 case, the number of equivalence classes is inherited from the index 1 case plus one equivalence class specific to the index 2. Indeed, we have one hereditary monomial $\text{mon}\prod_1^{(1,2)} \triangleq Y_1^{(1)}Y_1^{(2)}$ representative of $N(t), Y_i^{(1)}Y_i^{(2)}$ type random variables and one new index 2 specific monomial $\text{mon}\prod_1^{(2,2)} \triangleq Y_1^{(1)}Y_2^{(2)}$ representative of $N(t)\Big(N(t)-1\Big), Y_i^{(1)}Y_j^{(2)}$ type random variables for $i \neq j$. In the index 3 case, the number of equivalence classes is inherited from the index 2 case plus again one index 3 specific equivalence class. Indeed, we have one hereditary single block equivalence class denoted $\text{mon}\prod_1^{(1,3)} \triangleq Y_1^{(1)}Y_1^{(2)}Y_1^{(3)}$ representing $N(t)Y_i^{(1)}Y_i^{(2)}Y_i^{(3)}$ type variables and three hereditary two blocks equivalence classes denoted $\text{mon}\prod_1^{(2,3)} \triangleq Y_1^{(1)}Y_1^{(2)}Y_2^{(3)}$, $\text{mon}\prod_2^{(2,3)} \triangleq Y_1^{(1)}Y_1^{(3)}Y_2^{(2)}$ and $\text{mon}\prod_3^{(2,3)} \triangleq Y_1^{(1)}Y_2^{(2)}Y_2^{(3)}$ representing each $N(t)\Big(N(t)-1\Big)$ random variables. Finally, the equivalence class $\text{mon}\prod_1^{(3,3)} \triangleq Y_1^{(1)}Y_2^{(2)}Y_3^{(3)}$ is specific to the index 3 case. In order to show that, indeed the partition $Y_1^{(1)}Y_2^{(2)}Y_3^{(3)}$ is equivalent to $Y_i^{(1)}Y_j^{(2)}Y_k^{(3)}$, i, j, k pairwise different; we need here the following lemma which is a generalization of Lemma 1.

LEMMA 2. $E\Big[Y_{i_1}^{(1)}Y_{i_2}^{(2)}\ldots Y_{i_n}^{(n)}|N(t)\Big] = E\Big[Y_1^{(1)}Y_2^{(2)}\ldots Y_n^{(n)}|N(t)\Big]$ *for all pairwise distinct* $i_j, i_k; 1 \leq i_j, i_k \leq n$ *and* $N(t) \geq n$.

Proof. Using essentially the same approach as in Lemma 1, we can easily show that if one carries out index changes in one variable only at a time, the resulting random variables have the same expectations conditional on $N(t)$, as long as one verifies the pairwise distinctness of indices assumptions. Thus for example

$$E\Big[Y_{i_1}^{(1)}Y_2^{(2)}\ldots Y_n^{(n)}|N(t)\Big] = E\Big[Y_1^{(1)}Y_2^{(2)}\ldots Y_n^{(n)}|N(t)\Big]$$

By proceeding repeatedly for $Y_2^{(2)}, Y_3^{(3)}, \ldots, Y_n^{(n)}$ successively, we reach the equality in Lemma 2.

COROLLARY 1.

$$E\Big[\Big(\prod_{i\in I_1} Y_{j_1}^{(i)}\Big)\Big(\prod_{i\in I_2} Y_{j_2}^{(i)}\Big)\cdots\Big(\prod_{i\in I_m} Y_{j_m}^{(i)}\Big)|N(t)\Big]$$
$$= E\Big[\Big(\prod_{i\in I_1} Y_1^{(i)}\Big)\Big(\prod_{i\in I_2} Y_2^{(i)}\Big)\cdots\Big(\prod_{i\in I_m} Y_m^{(i)}\Big)|N(t)\Big]$$

for $j_1, j_2, \ldots, j_m$ pairwise distinct and indices from 1 to n partitioned into non intersecting blocks $I_1, I_2, \ldots, I_m$.

Proof. The result derives directly from Lemma 2. Indeed, the product $\prod_{i\in I_k} Y_{j_k}^{(i)}$ for a given k involves variables defined at the same time index j_k. Thus they could be considered as a single random variable denoted $\tilde{Y}_{j_k}^k$. Given that the indices $j_1, j_2, \ldots, j_k$ are pairwise distinct, the variables $\tilde{Y}_{j_k}^k$ for $k = 1, \ldots, m$ are mutually independant. Also, given that the I_k blocks, $k = 1, \ldots, m$ are non intersecting, the random variables $\tilde{Y}_{j_k}^k$ are all distinct. Thus they satisfy the conditions of Lemma 2, and Corollary 1 follows. ■

Note that Corollary 1 constitutes the basis for defining equivalence classes in an arbitrary index n case. Each time one is able to recognize a distinct partition of the indices of the n components of the $\underline{Y}_n' = (Y^{(1)}\ Y^{(2)}\ \ldots\ Y^{(n)})$ vector into m blocks, it will define a new equivalence class.

Let $S(n,m)$ be the number of ways one can partition a set of n distinct elements into m non-empty subsets (or blocks) [1].

$S(n,m)$ are Stirling's numbers of the second kind. Note that the total number of distinct equivalence classes for index n is given by:

$$B_n = \sum_{m=1}^{n} S(n,m)$$

The sequence $B_n, n = 1, 2, \ldots,$ is known as Bell's number (after Eric Temple Bell, see Andrews [3]). Finally, note that the cardinality of an equivalence class associated with m blocks is $A_m^{N(t)}$ if $N(t) \geq m$ which we also denote by $N(t)^{(m)} \triangleq N(t)\big(N(t)-1\big)\ldots\big(N(t)-m+1\big)$. This leads to the following lemma which is a generalization of Equations (4.6) and (4.8). Recall that mon $\Pi_\ell^{(m,n)}$ is the monomial representing the ℓ-th distinct equivalence class, within the equivalence classes associated with m blocks and index n, for $\ell = 1, 2, \ldots, S(n,m)$.

LEMMA 3. $$E\Big[P_n(t)|N(t)\Big] = \sum_{m=1}^{n}\Big[I\Big(N(t)\geq m\Big)\sum_{\ell=1}^{S(n,m)} N(t)^{(m)} E\Big[\mathrm{mon}\Pi_\ell^{(m,n)}|N(t)\Big]\Big]$$
$$\equiv \sum_{m=1}^{n}\sum_{\ell=1}^{S(n,m)} N(t)^{(m)} E\Big[\mathrm{mon}\Pi_\ell^{(m,n)}|N(t)\Big]$$

Proof. Lemma 3 follows from Corollary 1 and the remarks thereafter. Indeed in the expectation $E\Big[P_n(t)|N(t)\Big]$, one can subdivide monomials into equivalence classes corresponding to m blocks, $m = 1, \ldots, n$. Within the m blocks classes there will be $S(n,m)$ distinct ways of partitioning the

n components of vector $\underline{Y}'_m$, each represented by virtue of Corollary 1 by a single monomial mon $\Pi_\ell^{(m,n)}, \ell = 1, 2, \ldots, S(n,m)$. Within each such equivalence class, the total number of monomials is $N(t)^{(m)}$. Thus, the equality in Lemma 3 is a mathematical expression of these facts. ∎

Let us now state the following obvious consequence of Lemma 3.

COROLLARY 2. $E\Big[P_n(t)\Big] = E\Big[\sum_{m=1}^{n} \sum_{\ell=1}^{S(n,m)} N(t)^{(m)} \text{ mon } \Pi_\ell^{(m,n)} I\Big(\sum_{i=1}^{m} t_i \le t\Big)\Big]$

The next lemma is also easy to prove using nested conditioning and the independence of the first m reward vectors from $N(t)$, conditional on $t_1, t_2, \ldots, t_m$.

LEMMA 4.
$$E\Big[N(t)^{(m)} \text{mon} \Pi_\ell^{(m,n)} I\Big(\sum_{i=1}^{m} t_i \le t\Big)\Big]$$
$$= E\Big[E\Big[\text{ mon}\Pi_\ell^{(m,n)} | t_1, \ldots, t_m\Big] E\Big[\prod_{r=1}^{m}\Big(N(t_*^m) + r\Big) I\Big(\sum_{i=1}^{m} t_i \le t\Big) | t_1, \ldots, t_m\Big]\Big]$$

where $t_*^m \triangleq t - \sum_{i=1}^{m} t_i$.

The following result is a consequence of Corollary 4 and Lemma 4.

PROPOSITION 1.
$$E[P_n(t)] = \sum_{m=1}^{n} \sum_{\ell=1}^{S(n,m)} E\Big[E\Big[\text{mon } \Pi_\ell^{(m,n)} | t_1, \ldots, t_m\Big] E\Big[\prod_{r=1}^{m} \Big(N(t_*^m) + r\Big)$$
$$I\Big(\sum_{i=1}^{m} t_i \le t\Big) | t_1, \ldots, t_m\Big]\Big]$$

Note that Proposition 5 gives the exact expression for $E[P_n(t)]$. However, we are interested in an asymptotic expression for that expectation. Using Proposition 5 and Lemma 4, this question is tackled in the next section.

6. Asymptotic Behaviour of $E[P_n(t)]$

In order to characterize the asymptotic behaviour of $E[P_n(t)]$, we need to focus on the asymptotics of terms of the form:

$$E\Big[E\Big[\text{ mon}\Pi_\ell^{(m,n)} | t_1, \ldots, t_m\Big] E\Big[\prod_{r=1}^{m} \Big(N(t_*^m) + r\Big) I\Big(\sum_{i=1}^{m} t_i \le t\Big) | t_1, \ldots, t_m\Big]\Big] \qquad (6.1)$$

Now, recognizing that $E\Big[\prod_{r=1}^{m} \Big(N(t_*^m) + r\Big)\Big]$ is $\phi_m(t_*^m)$ with $t_1, t_2, \ldots, t_m$ treated as parameters, and where $\phi_m(t)$ is the factorial moment of $N(t)$ (as defined by Smith [17]), we can use the asymptotic theory of $\phi_m(t)$ (Smith [17]). We gather from [17] the following facts useful for our analysis.

"**Definition 1** *The function $\lambda(t)$ belongs to the class B if and only if it is bounded variation, tends to zero as t approaches $+\infty$ and satisfies the condition $\lambda(t)-\lambda(t-\alpha)=o(t^{-1})$ as $t\to+\infty$, for every $\alpha>0$.*

Theorem 1 If $\mu_{n+1}<\infty$ then $\phi_n(t)=\gamma_1 t^n+\gamma_2 t^{n-1}+\cdots+\gamma_n t+\gamma_{n+1}+\omega(t)$ where $\omega(t)\in B$".

Therefore it follows from Smith's [17] Theorem 1, that $\omega(t)=o(1)$ as $t\to\infty$. Note that the γ_i's $i=1,2,\dots,n$ represent finite rational functions of $\mu_1,\mu_2,\dots,\mu_i$.

Using the expansion of Theorem 1 in [17], and the properties of conditional expectation, we have:

$$E\Big[E\Big[\text{mon}\Pi_\ell^{(m,n)}\,|t_1,\dots,t_m\Big]E\Big[\Big(\gamma_1(t_*^m)^m+\gamma_2(t_*^m)^{m-1}+\cdots+\gamma_m(t_*^m)$$
$$+\gamma_{m+1}+\omega(t_*^m)\Big)I\Big(\sum_{i=1}^m t_i\le t\Big)|t_1,\dots,t_m\Big]\Big]$$

$$=E\Big[\text{mon}\Pi_\ell^{(m,n)}\Big[\gamma_1(t_*^m)^m+\gamma_2(t_*^m)^{m-1}+\cdots+\gamma_{m+1}+\omega(t_*^m)\Big]\Big(I\Big(\sum_{i=1}^m t_i\le t\Big)\Big)\Big] \qquad (6.2)$$

We proceed to show that under some moment finiteness assumptions to be specified:

$$E\Big[\text{mon}\Pi_\ell^{(m,n)}\Big[\gamma_1(t_*^m)^m\Big]I\Big(\sum_{i=1}^m t_i>t\Big)\Big]=o(1) \qquad (6.3)$$

and

$$E\Big[\text{mon}\Pi_\ell^{(m,n)}\,\omega(t_*^m)\Big]=o(1) \qquad (6.4)$$

6.1. Proof of Equation (6.3). We need the following lemma

LEMMA 5. *Under the hypothesis that*

$$E|\Big(\sum_{i=1}^m t_i\Big)^m\Big(\prod_{i\in I_1}Y_1^{(i)}\Big)\Big(\prod_{i\in I_2}Y_2^{(i)}\Big)\cdots\Big(\prod_{i\in I_m}Y_m^{(i)}\Big)|<\infty \qquad (6.5)$$

for any arbitrary partition of the n components of $\underline{Y}'_n$ into m blocks for $m=1,2,\dots,n$, then $E\Big[\text{mon}\Pi_\ell^{(m,n)}[\gamma_1(t_*^m)^m]I\Big(\sum_{i=1}^m t_i>t\Big)\Big]=o(1)$.

Proof. We have:

$$E\Big[|\,\text{mon}\Pi_\ell^{(m,n)}(\gamma_1)(t_*^m)^m I\Big(\sum_{i=1}^m t_i>t\Big)|\Big]\le|\gamma_1|E\Big[|\,\text{mon}\Pi_\ell^{(m,n)}\,t^m I\Big(\sum_{i=1}^m t_i>t\Big)|\Big].$$

But, $E\Big[|\,\text{mon}\Pi_\ell^{(m,n)}\,t^m I\Big(\sum_{i=1}^m t_i>t\Big)|\Big]\le E\Big[|\text{mon}\Pi_\ell^{(m,n)}\Big(\sum_{i=1}^m t_i\Big)I\Big(\sum_{i=1}^m t_i>t\Big)\Big]$; while the right-hand term is the tail of the expectation integral in (6.5), which as $t\to\infty$ must go to zero since the expectation in (6.5) is assumed to be finite. Thus, the lemma derives from the above two inequalities. ∎

6.2. Proof of Equation (6.4). Note that $\omega(t)$ is of bounded variation and thus $\omega^2(t)$ is of bounded variation. Also since $\omega(t)$ is $o(1)$, $\omega^2(t)$ is at least $o(1)$. We can assert that $E[\omega^2(t)]$ is

$o(1)$; indeed:

$$E[\omega^2(t_*^m)] = \int \omega^2(t-\beta)dF_m(\beta) \tag{6.6}$$

where $\beta \triangleq \sum_{i=1}^{m} t_i$ and $F_m(\beta)$ is the distribution of β.

Since $\omega^2(t)$ is of bounded variation, then it is bounded. Furthermore, $F_m(\beta)$ being a distribution function, is of bounded variation (or since $F_m(\beta)$ is class $\mathcal{J}$ (Smith [17]), then $F_m(\beta)$ is of bounded variation). Thus the conditions of Lemma 1 in Smith [15] are satisfied and $E[\omega^2(t_*^m)]$ is $o(1)$. We can now state the following lemma.

LEMMA 6. *Under the assumption that* $E[\text{ mon } \Pi_\ell^{(m,n)}]^2 < \infty$, *then* $E[\text{ mon } \Pi_\ell^{(m,n)}\ \omega(t_*^m)] = o(1)$.

Proof. By the Cauchy-Schwarz inequality, we have $|E[\text{ mon}\Pi_\ell^{(m,n)}\ \omega(t_*^m)]| \leq \sqrt{E[\text{ mon}\Pi_\ell^{(m,n)}]^2}\ \sqrt{E[\omega(t_*^m)]^2}$. As $E[\omega(t_*^m)]^2 = o(1)$, the result follows.

Finally we have the asymptotic behaviour of $E[P_n(t)]$ given by the following theorem.

THEOREM 1. *Under the finiteness of moments assumptions in Lemmas 5 and 6, and the assumption that distribution function* $F(\cdot)$ *is class* $\mathcal{J}$,

$$E[P_n(t)] = \sum_{m=1}^{n} \sum_{\ell=1}^{S(n,m)} E\Big[\text{mon } \Pi_\ell^{(m,n)} \Big(\sum_{i=1}^{m+1} \gamma_i (t_*^m)^{m+1-i}\Big)\Big] + o(1) \tag{6.7}$$

where the γ_i*'s above correspond to the asymptotic expression of* $\phi_m(t)$, *the* m^{th} *order cumulant of* $N(t)$ *as given in Theorem 1 in* [17], *and* $t_*^m = t - \sum_{i=1}^{m} t_i$.

Proof. This theorem follows from Proposition 5 and Equations (6.2) to (6.4).

7. Application

In this section, assuming the conditions of Theorem 6 are satisfied, we evaluate explicitly and asymptotically $E\Big[\prod_{\ell=1}^{3} \sum_{i=1}^{N(t)} Y_i^{(\ell)}\Big]$.

From Smith's [17] Lemma 6, we have

$$\phi_m^*(s) = m!\{1 - F^*(s)\}^{-m} \tag{7.1}$$

where $\phi_m^*(s)$ and $F^*(s)$ are Laplace-Stieltjes transform of $\phi_m(t)$ and $F(t)$ respectively, and $F^*(s)$ is given by Lemma 3 in Smith [17] as

$$F^*(s) = 1 - \mu_1 s + \frac{\mu_2 s^2}{2!} - \cdots + \frac{(-s)^n \mu_n}{n!} + o(s^n) \tag{7.2}$$

for real $s > 0$.

Equations (7.1) and (7.2) are the basis for computing $\phi_m(t_*^m)$ and the γ_i coefficients in Theorem 6. Applying Equation (6.7), we have

$$
\begin{aligned}
E\Big[\prod_{\ell=1}^{3}\sum_{i=1}^{N(t)} Y_i^{(\ell)}\Big] &= \sum_{m=1}^{3} E\Big[\sum_{\ell=1}^{S(3,m)} \mathrm{mon}\Pi_\ell^{(m,n)}\Big(\sum_{i=1}^{m+1}\gamma_i(t_*^m)^{m+1-i}\Big)\Big] + o(1) \\
&= \sum_{m=1}^{3} E\Big[\sum_{\ell=1}^{S(3,m)} \mathrm{mon}\Pi_\ell^{(m,n)}.\phi_m(t_*^m)\Big] + o(1)
\end{aligned}
\tag{7.3}
$$

$$
\begin{aligned}
E\Big[\prod_{\ell=1}^{3}\sum_{i=1}^{N(t)} Y_i^{(\ell)}\Big] &= E[Y_1^{(1)}Y_1^{(2)}Y_1^{(3)}\phi_1(t_*^1)] + E[Y_1^{(1)}Y_1^{(2)}Y_2^{(3)}\phi_2(t_*^2)] \\
&+ E[Y_1^{(1)}Y_1^{(3)}Y_2^{(2)}\phi_2(t_*^2)] + E[Y_1^{(1)}Y_2^{(2)}Y_2^{(3)}\phi_2(t_*^2)] \\
&+ E[Y_1^{(1)}Y_2^{(2)}Y_3^{(3)}\phi_3(t_*^3)] + o(1)
\end{aligned}
\tag{7.4}
$$

Recall that $t_*^m = t - \sum_{i=1}^{m} t_i$.

We assume that $Y_1^{(1)}$, $Y_2^{(2)}$ and $Y_3^{(3)}$ are correlated with t_1, t_2 and t_3 respectively. Using (7.1) and after some tedious algebra we find

$$
\begin{aligned}
&E\Big[\sum_{i=1}^{N(t)} Y_i^{(1)}\sum_{i=1}^{N(t)} Y_i^{(2)}\sum_{i=1}^{N(t)} Y_i^{(3)}\Big] = C_1^3E[Y_1^{(1)}]E[Y_1^{(2)}]E[Y_1^{(3)}]t^3 \\
&+\Big\{C_1^2E[Y_1^{(1)}Y_1^{(2)}]E[Y_1^{(3)}] + C_1^2E[Y_1^{(1)}Y_1^{(3)}]E[Y_1^{(2)}] + C_1^2E[Y_1^{(1)}]E[Y_1^{(2)}Y_1^{(3)}] \\
&-3C_1^3E[Y_1^{(1)}t_1]E[Y_1^{(2)}]E[Y_1^{(3)}] - 3C_1^3E[Y_1^{(1)}]E[Y_1^{(2)}t_1]E[Y_1^{(3)}] \\
&-3C_1^3E[Y_1^{(1)}]E[Y_1^{(2)}]E[Y_1^{(3)}t_1] + C_2E[Y_1^{(1)}]E[Y_1^{(2)}]E[Y_1^{(3)}]\Big\}t^2 \\
&+\Big\{C_1E[Y_1^{(1)}Y_1^{(2)}Y_1^{(3)}] - 2C_1^2E[Y_1^{(1)}Y_1^{(2)}t_1]E[Y_1^{(3)}] \\
&-2C_1^2E[Y_1^{(1)}Y_1^{(2)}]E[Y_1^{(3)}t_1] + C_3E[Y_1^{(1)}Y_1^{(2)}]E[Y_1^{(3)}] \\
&-2C_1^2E[Y_1^{(1)}Y_1^{(3)}t_1]E[Y_1^{(2)}] - 2C_1^2E[Y_1^{(1)}Y_1^{(3)}]E[Y_1^{(2)}t_1] \\
&+C_3E[Y_1^{(1)}Y_1^{(3)}]E[Y_1^{(2)}] - 2C_1^2E[Y_1^{(1)}t_1]E[Y_1^{(2)}Y_1^{(3)}] \\
&-2C_1^2E[Y_1^{(1)}]E[Y_1^{(2)}Y_1^{(3)}t_1] + C_3E[Y_1^{(1)}]E[Y_1^{(2)}Y_1^{(3)}] \\
&+3C_1^3E[Y_1^{(1)}t_1^2]E[Y_1^{(2)}]E[Y_1^{(3)}] + 3C_1^3E[Y_1^{(1)}]E[Y_1^{(2)}t_1^2]E[Y_1^{(3)}] \\
&+3C_1^3E[Y_1^{(1)}]E[Y_1^{(2)}]E[Y_1^{(3)}t_1^2] + 6C_1^3E[Y_1^{(1)}t_1]E[Y_1^{(2)}t_1]E[Y_1^{(2)}t_1]E[Y_1^{(3)}] \\
&+6C_1^3E[Y_1^{(1)}t_1]E[Y_1^{(2)}]E[Y_1^{(3)}t_1] + 6C_1^3E[Y_1^{(1)}]E[Y_1^{(2)}t_1]E[Y_1^{(3)}t_1]
\end{aligned}
$$

$$- C_4\left(E[Y_1^{(1)}t_1]E[Y_1^{(2)}]E[Y_1^{(3)}] + E[Y_1^{(1)}]E[Y_1^{(2)}t_1]E[Y_1^{(3)}] + E[Y_1^{(1)}]E[Y_1^{(2)}]E[Y_1^{(3)}t_1]\right)$$

$$+ C_5E[Y_1^{(1)}]E[Y_1^{(2)}]E[Y_1^{(3)}]\Big\}t$$

$$+ \Big\{-C_1E[Y_1^{(1)}Y_1^{(2)}Y_1^{(3)}t_1] + C_2E[Y_1^{(1)}Y_1^{(2)}Y_1^{(3)}] + C_1^2E[Y_1^{(1)}Y_1^{(2)}t_1^2]E[Y_1^{(3)}]$$

$$+ C_1^2E[Y_1^{(1)}Y_1^{(2)}]E[Y_1^{(3)}t_1^2] + 2C_1^2E[Y_1^{(1)}Y_1^{(2)}t_1]E[Y_1^{(3)}t_1]$$

$$- C_3E[Y_1^{(1)}Y_1^{(2)}t_1]E[Y_1^{(3)}] - C_3E[Y_1^{(1)}Y_1^{(2)}]E[Y_1^{(3)}t_1]$$

$$+ C_6E[Y_1^{(1)}Y_1^{(2)}]E[Y_1^{(3)}] + C_1^2E[Y_1^{(1)}Y_1^{(3)}]E[Y_1^{(2)}t_1^2]$$

$$+ 2C_1^2E[Y_1^{(1)}Y_1^{(3)}t_1]E[Y_1^{(2)}t_1] + C_1^2E[Y_1^{(1)}Y_1^{(3)}t_1^2]E[Y_1^{(2)}]$$

$$- C_3E[Y_1^{(1)}Y_1^{(3)}t_1]E[Y_1^{(2)}] - C_3E[Y_1^{(1)}Y_1^{(3)}]E[Y_1^{(2)}t_1]$$

$$+ C_6E[Y_1^{(1)}Y_1^{(3)}]E[Y_1^{(2)}] + C_1^2E[Y_1^{(1)}t_1^2]E[Y_1^{(2)}Y_1^{(3)}]$$

$$+ C_1^2E[Y_1^{(1)}]E[Y_1^{(2)}Y_1^{(3)}t_1^2] + 2C_1^2E[Y_1^{(1)}t_1]E[Y_1^{(2)}Y_1^{(3)}t_1]$$

$$- C_3E[Y_1^{(1)}t_1]E[Y_1^{(2)}Y_1^{(3)}] - C_3E[Y_1^{(1)}]E[Y_1^{(2)}Y_1^{(3)}t_1]$$

$$+ C_6E[Y_1^{(1)}]E[Y_1^{(2)}Y_1^{(3)}] - C_1^3E[Y_1^{(1)}t_1^3]E[Y_1^{(2)}]E[Y_1^{(3)}]$$

$$- C_1^3E[Y_1^{(1)}]E[Y_1^{(2)}t_1^3]E[Y_1^{(3)}] - C_1^3E[Y_1^{(1)}]E[Y_1^{(2)}]E[Y_1^{(3)}t_1^3]$$

$$- 3C_1^3E[Y_1^{(1)}t_1]E[Y_1^{(2)}t_1^2]E[Y_1^{(3)}] - 3C_1^3E[Y_1^{(1)}t_1]E[Y_1^{(2)}]E[Y_1^{(3)}t_1^2]$$

$$- 3C_1^3E[Y_1^{(1)}t_1^2]E[Y_1^{(2)}t_1]E[Y_1^{(3)}] - 3C_1^3E[Y_1^{(1)}]E[Y_1^{(2)}t_1]E[Y_1^{(3)}t_1^2]$$

$$- 3C_1^3E[Y_1^{(1)}t_1^2]E[Y_1^{(2)}]E[Y_1^{(3)}t_1] - 3C_1^3E[Y_1^{(1)}]E[Y_1^{(2)}t_1^2]E[Y_1^{(3)}t_1]$$

$$- 6C_1^3E[Y_1^{(1)}t_1]E[Y_1^{(2)}t_1]E[Y_1^{(3)}t_1]$$

$$C_7\left(E[Y_1^{(1)}t_1]E[Y_1^{(2)}t_1]E[Y_1^{(3)}] + E[Y_1^{(1)}t_1]E[Y_1^{(2)}]E[Y_1^{(3)}t_1] + E[Y_1^{(1)}]E[Y_1^{(2)}t_1]E[Y_1^{(3)}t_1]\right)$$

$$+ C_2\left(E[Y_1^{(1)}t_1^2]E[Y_1^{(2)}]E[Y_1^{(3)}] + E[Y_1^{(1)}]E[Y_1^{(2)}t_1^2]E[Y_1^{(3)}] + E[Y_1^{(1)}]E[Y_1^{(2)}]E[Y_1^{(3)}t_1^2]\right)$$

$$+ C_5\left(E[Y_1^{(1)}t_1]E[Y_1^{(2)}]E[Y_1^{(3)}] + E[Y_1^{(1)}]E[Y_1^{(2)}t_1]E[Y_1^{(3)}] + E[Y_1^{(1)}]E[Y_1^{(2)}]E[Y_1^{(3)}t_1]\right)$$

$$+ C_8E[Y_1^{(1)}]E[Y_1^{(2)}]E[Y_1^{(3)}]\Big\}$$

$$+ o(1) \tag{7.5}$$

where:

$$\begin{cases} C_1 = \frac{1}{\mu_1} \\ C_2 = \frac{9\mu_2}{2\mu_1^4} \\ C_3 = \frac{2\mu_2}{\mu_1^3} \\ C_4 = \frac{9\mu_2}{\mu_1^4} \\ C_5 = \frac{9\mu_2^2-3\mu_3\mu_1}{\mu_1^5} \\ C_6 = \frac{9\mu_2^2-4\mu_1\mu_3}{6\mu_1^4} \\ C_7 = \frac{9\mu_2}{\mu_1^4} \\ C_8 = \frac{-3\mu_1^2\mu_4+12\mu_2^3}{4\mu_1^6} \end{cases}$$

Note that in the development of (7.5) we used the following:

$$\begin{aligned} E[Y_m^{(\ell)}] &= E[Y_1^{(\ell)}] \qquad \ell = 1,2,3 \text{ and } m = 1,2,3 \\ E[Y_m^{(r)}Y_m^{(s)}] &= E[Y_1^{(r)}Y_1^{(s)}] \quad \forall\, r \neq s,\ r,s = 1,2,3 \\ E[Y_m^{(\ell)}t_m^k] &= E[Y_1^{(\ell)}t_1^k] \qquad \forall\, k < \infty \end{aligned}$$

Note also that Equation (7.5) cannot be further simplified. However, if $Y_i^{(1)} \equiv Y_i^{(2)} \equiv Y_i^{(3)} \equiv 1$, then (7.5) becomes exactly $E[N(t)]^3$ given by:

$$\begin{aligned} E[N(t)]^3 &= \frac{t^3}{\mu_1^3} + \left(\frac{9\mu_2-12\mu_1^2}{2\mu_1^4}\right)t^2 \\ &+ \left(\frac{9\mu_2^2-3\mu_3\mu_1+7\mu_1^4-12\mu_2\mu_1^2}{\mu_1^5}\right)t \\ &+ \left(\frac{-3\mu_1^2\mu_4-36\mu_2^2\mu_1^2+16\mu_3\mu_1^3+14\mu_2\mu_1^4+12\mu_2^3-4\mu_1^6}{4\mu_1^6}\right) \\ &+ o(1) \end{aligned} \tag{7.6}$$

Note also that Equation (7.6) coincides perfectly with $E[N(t)]^3$ in Adès, M. & Malhamé, R.P. [2], and Murthy (after correcting his algebra [13]). This provides an independent validation of the approach elaborated in the present chapter.

8. Recursive Generation of $P_n(t)$

The monomials occurring in $E[P_n(t)]$ are generated recursively and inherited from each others, following by that a specific pattern according to the analysis elaborated in this chapter. For example, $\mathrm{mon}\Pi_2^{(2,3)} = Y_1^{(1)}Y_1^{(3)}Y_2^{(2)}$ (monommial associated with a partition of 3 letters into 2 blocks), generates in the fourth generation the following three distinct monomials:
$Y_1^{(1)}Y_1^{(3)}Y_1^{(4)}Y_2^{(2)}, Y_1^{(1)}Y_1^{(3)}Y_2^{(2)}Y_2^{(4)}$ and $Y_1^{(1)}Y_1^{(3)}Y_2^{(2)}Y_3^{(4)}$

As we can remark, the first two monomials are associated with a partition of 4 letters into 2 blocks. In each of these two generated monomials, the "new" letter $Y^{(4)}$ is located at the end of each block, taking as lower index the position number of this block. The last monomial has, however, 3 blocks; this is an augmented (in the number of blocks) or innovated monomial, where $Y^{(4)}$ is located at the end of these two blocks and has 3 as lower index.

In general, $\mathrm{mon}\Pi_\ell^{(m,n-1)}$ generates in the n-th generation m monomials of n letters with m blocks, plus an augmented monomial of n letters with $m+1$ blocks, following essentially the same pattern described previously.

First Generation $\underline{Y}'_1 = (Y^{(1)})$

1.1.1 $\mathrm{mon}\Pi_1^{(1,1)} = Y_1^{(1)}$

Second Generation $\underline{Y}'_2 = (Y^{(1)}\ Y^{(2)})$

2.1.1 $Y_1^{(1)}$
- $\longrightarrow \mathrm{mon}\Pi_1^{(1,2)} = Y_1^{(1)}Y_1^{(2)}$
- $\longrightarrow \mathrm{mon}\Pi_1^{(2,2)} = Y_1^{(1)}Y_2^{(2)}$

Third Generation $\underline{Y}'_3 = (Y^{(1)}\ Y^{(2)}\ Y^{(3)})$

3.1.1 $Y_1^{(1)}Y_1^{(2)}$
- $\longrightarrow \mathrm{mon}\Pi_1^{(1,3)} = Y_1^{(1)}Y_1^{(2)}Y_1^{(3)}$
- $\longrightarrow \mathrm{mon}\Pi_1^{(2,3)} = Y_1^{(1)}Y_1^{(2)}Y_2^{(3)}$

3.2.1 $Y_1^{(1)}Y_2^{(2)}$
- $\longrightarrow \mathrm{mon}\Pi_2^{(2,3)} = Y_1^{(1)}Y_1^{(3)}Y_2^{(2)}$
- $\longrightarrow \mathrm{mon}\Pi_3^{(2,3)} = Y_1^{(1)}Y_2^{(2)}Y_2^{(3)}$
- $\longrightarrow \mathrm{mon}\Pi_1^{(3,3)} = Y_1^{(1)}Y_2^{(2)}Y_3^{(3)}$

Fourth Generation $\underline{Y}'_4 = (Y^{(1)}\ Y^{(2)}\ Y^{(3)}\ Y^{(4)})$

4.1.1 $Y_1^{(1)}Y_1^{(2)}Y_1^{(3)}$
- $\longrightarrow \mathrm{mon}\Pi_1^{(1,4)} = Y_1^{(1)}Y_1^{(2)}Y_1^{(3)}Y_1^{(4)}$
- $\longrightarrow \mathrm{mon}\Pi_1^{(2,4)} = Y_1^{(1)}Y_1^{(2)}Y_1^{(3)}Y_2^{(4)}$

4.2.1 $Y_1^{(1)}Y_1^{(2)}Y_2^{(3)}$
- $\longrightarrow \mathrm{mon}\Pi_2^{(2,4)} = Y_1^{(1)}Y_1^{(2)}Y_1^{(4)}Y_2^{(3)}$
- $\longrightarrow \mathrm{mon}\Pi_3^{(2,4)} = Y_1^{(1)}Y_1^{(2)}Y_2^{(3)}Y_2^{(4)}$
- $\longrightarrow \mathrm{mon}\Pi_1^{(3,4)} = Y_1^{(1)}Y_1^{(2)}Y_2^{(3)}Y_3^{(4)}$

4.2.2 $Y_1^{(1)}Y_1^{(3)}Y_2^{(2)} \longrightarrow$

$$\mathrm{mon}\Pi_4^{(2,4)} = Y_1^{(1)}Y_1^{(3)}Y_1^{(4)}Y_2^{(2)}$$
$$\mathrm{mon}\Pi_5^{(2,4)} = y_1 Y_1^{(3)}Y_2^{(2)}Y_2^{(4)}$$
$$\mathrm{mon}\Pi_2^{(3,4)} = Y_1^{(1)}Y_1^{(3)}Y_2^{(2)}Y_3^{(4)}$$

4.2.3 $Y_1^{(1)}Y_2^{(2)}Y_2^{(3)} \longrightarrow$

$$\mathrm{mon}\Pi_6^{(2,4)} = Y_1^{(1)}Y_1^{(4)}Y_2^{(2)}Y_2^{(3)}$$
$$\mathrm{mon}\Pi_7^{(2,4)} = Y_1^{(1)}Y_2^{(2)}Y_2^{(3)}Y_2^{(4)}$$
$$\mathrm{mon}\Pi_3^{(3,4)} = Y_1^{(1)}Y_2^{(2)}Y_2^{(3)}Y_3^{(4)}$$

4.3.1 $Y_1^{(1)}Y_2^{(2)}Y_3^{(3)} \longrightarrow$

$$\mathrm{mon}\Pi_4^{(3,4)} = Y_1^{(1)}Y_1^{(4)}Y_2^{(2)}Y_3^{(3)}$$
$$\mathrm{mon}\Pi_5^{(3,4)} = Y_1^{(1)}Y_2^{(2)}Y_2^{(4)}Y_3^{(3)}$$
$$\mathrm{mon}\Pi_6^{(3,4)} = Y_1^{(1)}Y_2^{(2)}Y_3^{(3)}Y_3^{(4)}$$
$$\mathrm{mon}\Pi_1^{(4,4)} = Y_1^{(1)}Y_2^{(2)}Y_3^{(3)}Y_4^{(4)}$$

There are B_4 monomials for this generation ($B_4 = 15$).

n-th Generation $\underline{Y}'_n = (Y^{(1)}\ Y^{(2)}\ \dots\ Y^{(n)})$

n.1.1 $Y_1^{(1)}Y_2^{(2)}\dots Y_1^{(n-1)} \longrightarrow$

$$\mathrm{mon}\Pi_1^{(1,n)} = Y_1^{(1)}Y_1^{(2)}\dots Y_1^{(n-1)}Y_1^{(n)}$$
$$\mathrm{mon}\Pi_1^{(2,n)} = Y_1^{(1)}Y_1^{(2)}\dots Y_1^{(n-1)}Y_2^{(n)}$$

n.2.1 $Y_1^{(1)}Y_1^{(2)}\dots Y_1^{(n-2)}Y_2^{(n-1)} \longrightarrow$

$$\mathrm{mon}\Pi_2^{(2,n)} = Y_1^{(1)}Y_1^{(2)}\dots Y_1^{(n-2)}Y_1^{(n)}Y_2^{(n-2)}$$
$$\mathrm{mon}\Pi_3^{(2,n)} = Y_1^{(1)}Y_2^{(2)}\dots Y_1^{(n-2)}Y_2^{(n-1)}Y_2^{(n)}$$
$$\mathrm{mon}\Pi_1^{(3,n)} = Y_1^{(1)}Y_1^{(2)}\dots Y_1^{(n-2)}Y_2^{(n-1)}Y_3^{(n)}$$

$\vdots \qquad\qquad \vdots \quad \vdots \quad \vdots$

$n.2.S(n-1,2)$

$Y_1^{(1)}\dots Y_1^{(n-3)}Y_2^{(n-2)}Y_2^{(n-1)} \longrightarrow$

$$\mathrm{mon}\Pi_{S(n,2)-1}^{(2,n)} = Y_1^{(1)}\dots Y_1^{(n-3)}Y_1^{(n)}Y_2^{(n-2)}Y_2^{(n-1)}$$
$$\mathrm{mon}\Pi_{S(n,2)}^{(2,n)} = Y_1^{(1)}\dots Y_1^{(n-3)}Y_2^{(n-2)}Y_2^{(n-1)}Y_2^{(n)}$$
$$\mathrm{mon}\Pi_{S(n-1,2)}^{(3,n)} = Y_1^{(1)}\dots Y_1^{(n-3)}Y_2^{(n-2)}Y_2^{(n-1)}Y_3^{(n)}$$

$\vdots \qquad\qquad \vdots \quad \vdots \quad \vdots$

$n.(n-1).1$

$$Y_1^{(1)}Y_2^{(2)}\dots Y_{n-1}^{(n-1)} \longrightarrow \begin{cases} \mathrm{mon}\Pi^{(n-1,n)}_{S(n,n-1)-(n-2)} = Y_1^{(1)}Y_1^{(n)}Y_2^{(2)}\dots Y_{n-1}^{(n-1)} \\ \mathrm{mon}\Pi^{(n-1,n)}_{S(n,n-1)-(n-1)} = Y_1^{(1)}Y_2^{(2)}Y_2^{(n)}\dots Y_{n-1}^{(n-1)} \\ \dots \quad . \quad . \\ \dots \quad . \quad . \\ \dots \quad . \quad . \\ \mathrm{mon}\Pi^{(n-1,n)}_{S(n,n-1)} = Y_1^{(1)}Y_2^{(2)}\dots Y_{n-1}^{(n-1)}Y_{n-1}^{(n)} \\ \mathrm{mon}\Pi^{(n,n)}_{S(n,n)} = Y_1^{(1)}Y_2^{(2)}\dots Y_{n-1}^{(n-1)}Y_n^{(n)} \end{cases}$$

Recall that for each generation, there are B_n monomials where B_n are Bell's numbers.

9. Conclusion

As we observe, the evaluation of $E[P_n(t)]$, using the present approach, is based only on $\phi_n(t_*^m)$, since the monomials of $P_n(t)$ are generated recursively. Note that the constants γ_1 to γ_{n+1} in the expansion of $\phi_n(t)$ can be computed using e.g. an algorithm of Teugels [19]. Thus, using a symbolic language of programming as Maple or Mathematica, the symbolic computation of $E[P_n(t)]$ can be performed efficiently.

References

[1] ABRAMOWITZ, M. AND STEGUN, I. (1972) Handbook of Mathematical Functions. *National Bureau of Standards, Applied Mathematics Series 55.* Tenth Printing, Washington.

[2] ADÈS, M. AND MALHAMÉ, R.P. (1994) On the Moments of Cumulative Processes: A Preliminary Study. *Les Cahiers du GERAD* **G-94-13**. École des Hautes Études Commerciales, Montréal.

[3] ANDREWS, G.E. (1976) *The Theory of Partitions.* Addison-Wesley, Massachusetts.

[4] BROWN, M. AND SOLOMON, H. (1975) A Second Approximation for the Variance of a Renewal Reward Process. *Stochastic Processes and their Applications* **3**, 303–314.

[5] CHOW, Y.S. AND TEICHER, H. (1988) *Probability Theory: Independance, Interchangeability, Martingales.* Springer-Verlag, New York.

[6] COMTET, L. (1974) *Advanced Combinatorics: The Art of Finite and Infinite Expansions.* D. Reidel Publishing Company, Boston, Massachusetts.

[7] COX, D.R. (1982) *Renewal Theory.* Methuen, New York.

[8] De Finetti, B. (1931) Funzione Caratteristica di un Fenomeno Aleatorio. *Mem. Acc. Lincei* **4**, 86–133.

[9] Glynn, P.W. and Heidelberger, P. (1990) Bias Properties of Budget Constrained Simulations. *Operations Research* **38**, 801–814.

[10] Jensen, U. (1984) On the Moments of Renewal Reward Process. In *Contributions to Operations Research and Mathematical Economics*. Eds.. G. Hamer and Pallaschke, 601–610.

[11] Kingman, J.F.C. (1978) Uses of Exchangeability. *The Annals of Probability* **6**, 183–197.

[12] Koch, G. and Spizzichino, F. (1982) *Exchangeability in Probability and Statistics*. North-Holland, Amsterdam.

[13] Murthy, V.K. (1974) *The General Point Process*. Addison-Wesley, Massachusetts.

[14] Ross, S.M. (1983) *Stochastic Processes*. John-Wiley, New York.

[15] Smith, W.L. (1953) Asymptotic Renewal Theorems. *Proceedings of the Royal Society of Edinburgh* A **64**, 9–48.

[16] Smith, W.L. (1955) Regenerative Stochastic Process. *Proceedings of the Royal Society* A **232**, 6–31.

[17] Smith, W.L. (1959) On the Cumulants of Renewal Process. *Biometrika* **46**, 1–29.

[18] Smith, W.L. (1988) Renewal Theory. In *Encyclopedia of Statistical Sciences* **8**, 30–36. John-Wiley, New York.

[19] Teugels, J.L. (1967) Exponential Decay in Renewal Theorems. *Bulletin de la Société Mathématique de Belgique* **19**, 259–276.

CHAPTER 3

Asymptotics of the Moments of Cumulative Vector Renewal Reward Processes: The Case $N(t)+1$ "Generalization of Wald's Fundamental Equation in the Discrete Time: An asymptotic Study"

1. Introduction: Classical Definitions and Notations

Consider a renewal sequence $\{t_i\}$ $i = 1, 2, \ldots$ of i.i.d. non negative variables (time intervals). To each t_i, we associate a random vector function $Y_i \in \mathbb{R}^d$, $d \geq 1$, which in general depends on t_i. We assume that $[t_i, Y_i']$, $i \geq 1$, is a sequence of i.i.d. random vectors (the components of Y_i could be interdependent), where ′ indicates vector transposition. The general problem of interest here is that of obtaining asymptotic expression of $E[P_n^*(t)]$ as time t increases, where the products $P_n^*(t)$, $n \geq 1$, are of the form: $P_n^*(t) = \prod_{\ell=1}^{n} \Big(\sum_{i=1}^{N(t)+1} Y_i^{(\ell)} \Big)$, (ℓ) refers to the ℓ^{th} component of the Y_i vector, $N(t)$ is the random integer such that $\sum_{i=1}^{N(t)} t_i \leq t \leq \sum_{i=1}^{N(t)+1} t_i$. More generally, we can evaluate the asymptotic expression of $E[\bar{P}_n(t)]$ where $\bar{P}_n(t) = \prod_{\ell=1}^{n} \Big(\sum_{i=1}^{N(t)+1} Y^{(\ell)} \Big)^{p_\ell}$ such that $\sum_{l=1}^{n} p_\ell = n$. Such sums appear in the study of cumulative processes (Smith [8]).

Clearly, the random integer variable $N(t)$ represents the number of events in the interval $(0, t]$ of the renewal process $\{t_i\}$ $i \geq 1$; note that Karlin & Taylor [4] and Ross [6] call $N(t)$ a renewal process, which is slightly different from Smith's [8] terminology. Let $\{N(t)+1\}$ be a stopping time (Ross [6], Shiryayev [7]) with respect to $\{t_i\}$ $i \geq 1$, i.e., the event $\{N(t)+1 = n\}$ depends only on $\{t_1, \ldots, t_n\}$ and therefore independent of $\{t_{n+1}, t_{n+2}, \ldots\}$, whereas $\{N(t) = n\}$ depends on $\{t_1, \ldots, t_n, t_{n+1}\}$. We write $\mathcal{K}_r^{(\ell)} = E\Big[[Y_i^{(\ell)}]^r\Big]$, $\mu_r = E[t_i^r]$, $\mu_{jk}^{(\ell)} = E\Big[t_i^j [Y_i^{(\ell)}]^k\Big]$, where $\ell = 1, \ldots, d$; $i, j, k, r = 1, 2 \ldots$; and $\mathcal{K}_r^{(\ell)} + \mu_r + \mu_{jk}^{(\ell)} < \infty$. We also assume that the distribution of the underlying renewal process is class $\mathcal{L}$ (Smith [9]) i.e., the class of distribution functions $F(\cdot)$ for which $F_n(\cdot)$, for some finite n, has an absolutely continuous component, where $F(\cdot)$ and $F_n(\cdot)$ are the distribution of $\{t_i\}\, i \geq 1$ and $T_n = \sum_{i=1}^{n} t_i$ respectively.

The fundamental difference between $N(t)$ and $N(t)+1$ is very crucial, indeed Corollary 2, Proposition 1 and Theorem 1 of Adès and Malhamé [1] are no longer valid if we replace $N(t)$ by $N(t)+1$; the present chapter is a consequence of that fact. Besides the interchangeability idea developed in [1], we shall need here the theory of martingales to overcome the difficulty inherent to the analysis of the last renewal cycles involved in the summation part of $P_n^*(t)$.

Our approach consists in classifying the product of random variables appearing in $P_n^*(t)$ into separate classes which we term *expectation summable class*. More precisely, such *expectation summable classes* of random variables, can each be associated with an auxiliary martingale sequence that can be shown via exchangeability type arguments to have the same expectation as that of the sum of random variables in this *expectation summable class*. Subsequently, using the martingale property, the asymptotic behaviour of the expectation of that sum is characterized.

The present chapter is organized as follows. In Section 2, we carry out a study of the asymptotic behaviour of $E[P_2^*(t)]$, while the results are well known, this section is, nevertheless, useful in illustrating our methodology on a relatively simple case. In Section 3, we present the analysis for $E[P_3^*(t)]$ wherein we introduce the notion of partition appropriate to our study. In Section 4, we present combinatorial elements essential in partitioning the product of random variables in $P_n^*(t)$ into *expectation summable classes*. In Section 5, we develop general asymptotic expressions for $E[P_n^*(t)]$. Detailed computations for evaluating explicitly the asymptotic expression in the case $n = 3$ are performed in Section 6. Finally, a recursive scheme is given in Section 7 for generating $P_n^*(t)$.

2. Asymptotics of $E\Big[\sum_{i=1}^{N(t)+1} Y_i^{(1)} \sum_{i=1}^{N(t)+1} Y_i^{(2)}\Big]$

This section deals with two types of scalar cumulative processes. Carrying out the ordinary multiplication of $\sum_{i=1}^{N(t)+1} Y_i^{(1)} \sum_{i=1}^{N(t)+1} Y_i^{(2)}$ results in the following two dimensional array valid for $N(t) \geq 2$:

$$\left.\begin{array}{l}
Y_1^{(1)}Y_1^{(2)} + Y_1^{(1)}Y_2^{(2)} + Y_1^{(1)}Y_3^{(2)} + \cdots + Y_1^{(1)}Y_{N(t)}^{(2)} + Y_1^{(1)}Y_{N(t)+1}^{(2)} \\
+Y_2^{(1)}Y_1^{(2)} + Y_2^{(1)}Y_2^{(2)} + Y_2^{(1)}Y_3^{(2)} + \cdots + Y_2^{(1)}Y_{N(t)}^{(2)} + Y_2^{(1)}Y_{N(t)+1}^{(2)} \\
+Y_3^{(1)}Y_1^{(2)} + Y_3^{(1)}Y_2^{(2)} + Y_3^{(1)}Y_3^{(2)} + \cdots + Y_3^{(1)}Y_{N(t)}^{(2)} + Y_3^{(1)}Y_{N(t)+1}^{(2)} \\
\vdots \\
+Y_{N(t)}^{(1)}Y_1^{(2)} + Y_{N(t)}^{(1)}Y_2^{(2)} + Y_{N(t)}^{(1)}Y_3^{(2)} + \cdots + Y_{N(t)}^{(1)}Y_{N(t)}^{(2)} + Y_{N(t)}^{(1)}Y_{N(t)+1}^{(2)} \\
+Y_{N(t)+1}^{(1)}Y_1^{(2)} + Y_{N(t)+1}^{(1)}Y_2^{(2)} + Y_{N(t)+1}^{(1)}Y_3^{(2)} + \cdots + Y_{N(t)+1}^{(1)}Y_{N(t)}^{(2)} + Y_{N(t)+1}^{(1)}Y_{N(t)+1}^{(2)}
\end{array}\right\} \quad (2.1)$$

Note that the distribution of $Y_{N(t)+1}^{(1)}$ is in general different from that of $Y_i^{(1)}$ $i = 1, 2, \ldots, N(t)$. This is the so-called renewal paradox; for more details and additional references see, e.g., Feller [3] and Ross [6]. Thus, conditional on $\{N(t) = n\}$ the random variables $(Y_1^{(1)}, \ldots, Y_n^{(1)})$ and $Y_{n+1}^{(1)}$ are not exchangeable. Consequently as a result of this fact, the approach developed for the case $N(t)$ in [1] cannot be directly transposed for the case $N(t) + 1$. From (2.1) we distinguish the following three cases:

(a) random variables on the diagonal $Y_i^{(1)}Y_i^{(2)}$ for $i = 1, 2, \ldots, N(t) + 1$.

(b) random variables on the right of this diagonal.

  i) $E\Big[Y_i^{(1)}Y_j^{(2)}|N(t)\Big] = E\Big[Y_1^{(1)}Y_2^{(2)}|N(t)\Big]$
  for $i \neq j$ and $i, j = 1, 2, \ldots, N(t)$; $N(t) \geq 2$.
  This is Lemma 1 in Adès and Malhamé [1].

  ii) $E\Big[Y_i^{(1)}Y_{N(t)+1}^{(2)}|N(t)\Big] = E\Big[Y_1^{(1)}Y_{N(t)+1}^{(2)}|N(t)\Big]$
  for $i = 1, 2, \ldots, N(t)$; $N(t) \geq 1$.
  This can be shown following an approach similar to Lemma 1 in Adès and Malhamé [1].

(c) random variables of the left of this diagonal.

For such variables we have:

  i) $E\Big[Y_i^{(2)}Y_j^{(1)}|N(t)\Big] = E\Big[Y_1^{(2)}Y_2^{(1)}|N(t)\Big]$
  for $i \neq j$ and $i, j = 1, 2, \ldots, N(t)$; $N(t) \geq 2$.

  ii) $E\Big[Y_i^{(2)}Y_{N(t)+1}^{(1)}|N(t)\Big] = E\Big[Y_1^{(2)}Y_{N(t)+1}^{(1)}|N(t)\Big]$
  for $i = 1, 2, \ldots, N(t)$; $N(t) \geq 1$.

The total number of products of random variables occurring in categories (b) i)–ii) is as follows: $\dfrac{(N(t)+1)^{(2)}}{2}$ for case (b) i) and $N(t)$ for case (b) ii), where

$$N(t)^{(k)} = k\sum_{i=1}^{N(t)} (i-1)(i-2)\cdots(i-k+1) \tag{2.2}$$

Thus, if we add expectations of all the variables in (b) i)–ii), we obtain the following expectation:

$$E\Big[\sum_{i=1}^{N(t)+1} (i-1)Y_1^{(1)}Y_i^{(2)}|N(t)\Big] \quad \text{for} \quad N(t)\geq 1.$$

As we show further, the above expectation can be evaluated using martingales techniques. We term the merge of product of random variables in (b) i)–ii) an *expectation summable class.*

Applying the same analysis for step (c) i)–ii) results in a total expectation of:

$$E\Big[\sum_{i=1}^{N(t)+1} (i-1)Y_1^{(2)}Y_i^{(1)}|N(t)\Big] \quad \text{for} \quad N(t)\geq 1.$$

Again the merge of product of random variables in (c) i)–ii) is also an *expectation summable class.* It follows that for $N(t)\geq 1$:

$$\begin{aligned} E\Big[\sum_{i=1}^{N(t)+1} Y_i^{(1)} \sum_{i=1}^{N(t)+1} Y_i^{(2)}|N(t)\Big] &= E\Big[\sum_{i=1}^{N(t)+1} Y_i^{(1)}Y_i^{(2)}|N(t)\Big] \\ &+E\Big[\sum_{i=1}^{N(t)+1}(i-1)Y_1^{(1)}Y_i^{(2)}|N(t)\Big] + E\Big[\sum_{i=1}^{N(t)+1}(i-1)Y_1^{(2)}Y_i^{(1)}|N(t)\Big]. \end{aligned} \tag{2.3}$$

Thus, the law of total probability entails:

$$\begin{aligned} E\Big[\sum_{i=1}^{N(t)+1} Y_i^{(1)} \sum_{i=1}^{N(t)+1} Y_i^{(2)}\Big] &= E\Big[\sum_{i=1}^{N(t)+1} Y_i^{(1)}Y_i^{(2)}\Big] \\ &+E\Big[\Big(\sum_{i=1}^{N(t)+1}(i-1)Y_1^{(1)}Y_i^{(2)}\Big)I(t_1\leq t)\Big] + E\Big[\Big(\sum_{i=1}^{N(t)+1}(i-1)Y_1^{(2)}Y_i^{(1)}\Big)I(t_1\leq t)\Big] \end{aligned} \tag{2.4}$$

where $I(\cdot)$ is the indicator function.

2.1. Evaluation of the Expectations in the RHS of (2.4). The first expectation can be evaluated via Wald's fundamental equation:

$$E\Big[\sum_{i=1}^{N(t)+1} Y_i^{(1)}Y_i^{(2)}\Big] = E\Big[Y_1^{(1)}Y_1^{(2)}\Big]E\Big[N(t)+1\Big]$$

In order to evaluate the next two expectations, we apply martingales techniques. We proceed as follows:

- Under conditioning on t_1, we have

$$\begin{aligned} E\Big[\Big(\sum_{i=1}^{N(t)+1}(i-1)Y_1^{(1)}Y_i^{(2)}\Big)I(t_1 \le t)\Big] \\ = E\Big[E\Big[\Big(\sum_{i=1}^{N(t)+1}(i-1)Y_1^{(1)}Y_i^{(2)}\Big)I(t_1 \le t)|t_1\Big]\Big] \\ = E\Big[E\Big[Y_1^{(1)}|t_1\Big]E\Big[\Big(\sum_{i=1}^{N(t-t_1)+2}(i-1)Y_i^{(2)}\Big)I(t_1 \le t)|t_1\Big]\Big] \end{aligned} \tag{2.5}$$

where $N(t-t_1)$ is defined only for $t \ge t_1$.

- In order to evaluate the second expectation in the RHS of (2.5), we need to consider the following sequence:

$$S_n = \sum_{i=1}^{n}(i-1)(Y_i^{(2)} - \mu) \tag{2.6}$$

for all finite $n \ge 2$ and $\mu = E[Y_2^{(2)}]$.
Note that S_n is a martingale since $E\Big[|S_n|\Big] < \infty$ for all finite $n \ge 2$ and $E\Big[S_{n+1}|S_2, S_3, \ldots, S_n\Big] = S_n$.

- Consider the random variable $N = N(t-t_1)+2$ for $t_1 \le t$, and assuming that $Y_i^{(2)}$ are measurable functions of t_i, N will be a stopping time with respect to the σ-field $\mathcal{F}_n^S = \sigma(\omega, S_2, S_3, \ldots, S_n)$. If we assume that the fundamental renewal cycles are such that $E\Big[N(t)\Big]^{\ell} < \infty$ for all finite $\ell \ge 0$ and for all finite t, then $E[N] < \infty$. Now define $\xi_i = (i-1)(y_i^{(2)} - \mu), i = 2, \ldots, n$. Note that $E[\sum_{i=2}^{N} E|\xi_i||t_1] \le \bar{\mu}E[\frac{N(N-1)}{2}|t_1] < \infty$, where $\bar{\mu}$ is the upper bound on the expectation of the absolute value of individual rewards.
- It follows from Shiryayev's [7] problem 6 on page 464, that $E\Big[S_N|t_1\Big] = E\Big[S_2|t_1\Big] = 0$, and after some algebra we have

$$\begin{aligned} E\Big[\Big(\sum_{i=1}^{N(t-t_1)+2}(i-1)Y_i^{(2)}\Big)I(t_1 \le t)|t_1\Big] \\ = E\Big[Y_2^{(2)}\Big]E\Big[\frac{\Big(N(t-t_1)+2\Big)\Big(N(t-t_1)+1\Big)}{2}I(t_1 \le t)|t_1\Big]. \end{aligned} \tag{2.7}$$

Using (2.7) in (2.5) yields:

$$\begin{aligned} E\Big[E\Big[\Big(\sum_{i=1}^{N(t)+1}(i-1)Y_1^{(1)}Y_i^{(2)}\Big)I(t_1 \le t)|t_1\Big]\Big] \\ = E\Big[E\Big[Y_1^{(1)}|t_1\Big]E\Big[Y_2^{(2)}\Big]E\Big[\frac{\Big(N(t-t_1)+2\Big)\Big(N(t-t_1)+1\Big)}{2}I(t_1 \le t)|t_1\Big]. \end{aligned} \tag{2.8}$$

Subsequently, adapting Lemmas 5 and 6 of Section 6 in Adès and Malhamé [1], we can write asymptotically:

$E\Big[E[Y_1^{(1)}|t_1]E[\phi_2(t-t_1)I(t_1 > t)|t_1]\Big] = o(1)$ and $E\Big[E[Y_1^{(1)}|t_1]E[\omega_2(t-t_1)|t_1]\Big] = o(1)$, where $\phi_m(t) = E\Big[\prod_{\ell=1}^{m}(N(t)+\ell)\Big]$ and $\omega(t-t_1)$ is the remainder in the asymptotic expansion of $\phi_2(t-t_1)$ given t_1. Thus, we can write:

$$E\left[\left(\sum_{i=1}^{N(t)+1}(i-1)Y_1^{(1)}Y_i^{(2)}\right)I(t_1 \le t)\right] = E[Y_2^{(2)}]E\left[\frac{Y_1^{(1)}\phi_2(t-t_1)}{2}I(t_1 \le t)\right] + o(1). \tag{2.9}$$

Following similar steps, the third expectation in the RHS of (2.4), can be written as:

$$E\left[\left(\sum_{i=1}^{N(t)+1}(i-1)Y_1^{(2)}Y_i^{(1)}\right)I(t_1 \le t)\right] = E[Y_2^{(1)}]E\left[\frac{Y_1^{(2)}\phi_2(t-t_1)}{2}I(t_1 \le t)\right] + o(1). \tag{2.10}$$

Thus:

$$\begin{aligned} E\left[\sum_{i=1}^{N(t)+1} Y_i^{(1)} \sum_{i=1}^{N(t)+1} Y_i^{(2)}\right] &= E[Y_1^{(1)}Y_1^{(2)}]E\Big[N(t)+1\Big] \\ &+E[Y_2^{(2)}]E\left[\frac{Y_1^{(1)}\phi_2(t-t_1)}{2}I(t_1 \le t)\right] \\ &+E[Y_2^{(1)}]E\left[\frac{Y_1^{(2)}\phi_2(t-t_1)}{2}I(t_1 \le t) + o(1).\right. \end{aligned} \tag{2.11}$$

2.2. Asymptotic Explicit Expressions for $E\left[\sum_{i=1}^{N(t)+1} Y_i^{(1)} \sum_{i=1}^{N(t)+1} Y_i^{(2)}\right]$ and $E\left[\sum_{i=1}^{N(t)+1} Y_i^{(1)}\right]^2$.
Using (2.11) and Smith's asymptotic analysis of factorial moments [9], we obtain:

$$\begin{aligned} E\left[\sum_{i=1}^{N(t)+1} Y_i^{(1)} \sum_{i=1}^{N(t)+1} Y_i^{(2)}\right] &= \frac{E[Y_1^{(1)}]E[Y_1^{(2)}]}{\mu_1^2}t^2 \\ &+\left(\frac{E[Y_1^{(1)}Y_1^{(2)}]}{\mu_1} + \frac{2\mu_2 E[Y_1^{(1)}]E[Y_1^{(2)}]}{\mu_1^3} - \frac{E[Y_1^{(1)}]E[t_1 Y_1^{(2)}]}{\mu_1^2}\right. \\ &\left.-\frac{E[Y_1^{(2)}]E[t_1 Y_1^{(1)}]}{\mu_1^2}\right)t + \frac{\mu_2 E[Y_1^{(1)}Y_1^{(2)}]}{2\mu_1^2} \\ &+E[Y_1^{(1)}]E[Y_1^{(2)}]\left(\frac{3\mu_2^2}{2\mu_1^4} - \frac{2\mu_3}{3\mu_1^3}\right) + \frac{E[Y_1^{(1)}]E[t_1^2 Y_1^{(2)}]}{2\mu_1^2} \\ &+\frac{E[Y_1^{(2)}]E[t_1^2 Y_1^{(1)}]}{2\mu_1^2} - \frac{\mu_2 E[Y_1^{(1)}]E[t_1 Y_1^{(2)}]}{\mu_1^3} \\ &-\frac{\mu_2 E[Y_1^{(2)}]E[t_1 Y_1^{(1)}]}{\mu_1^3} + o(1). \end{aligned} \tag{2.12}$$

When $Y_i^{(1)}$ and $Y_i^{(2)}$ are identical, (2.12) yields:

$$
\begin{aligned}
E\left[\sum_{i=1}^{N(t)+1} Y_i^{(1)}\right]^2 = \frac{E^2[Y_1^{(1)}]}{\mu_1^2}t^2 + \Bigg(& \frac{E[Y_1^{(1)}]^2}{\mu_1} \\
& + \frac{2\mu_2 E^2[Y_1^{(1)}] - 2\mu_1 E[Y_1^{(1)}]E[t_1 Y_1^{(1)}]}{\mu_1^3}\Bigg)t \\
& + \frac{\mu_2 E[Y_1^{(1)}]^2}{2\mu_1^2} + \frac{9\mu_2^2 E^2[Y_1^{(1)}]}{6\mu_1^4} - \frac{2\mu_3 E^2[Y_1^{(1)}]}{3\mu_1^3} \\
& + \frac{E[Y_1^{(1)}]E[t_1^2 Y_1^{(1)}]}{\mu_1^2} - \frac{2\mu_2 E[Y_1^{(1)}]E[t_1 Y_1^{(1)}]}{\mu_1^3} + o(1).
\end{aligned}
\tag{2.13}
$$

The result for $E\left[\sum_{i=1}^{N(t)+1} Y_i^{(1)}\right]^2$ is well known in the literature and coincides with (2.13); however, existing derivations are based on computing characteristic functions (Smith [8], Murthy [5]).

As shown in Adès and Malhamé [2], the evaluation of $E\left[\sum_{i=1}^{N(t)+1} Y_i^{(1)} \sum_{i=1}^{N(t)+1} Y_i^{(2)} \sum_{i=1}^{N(t)+1} Y_i^{(3)}\right]$ using characteristic functions and linearization technique is very long and tedious. By contrast, the approach herein presented for the bivariate case generalizes to the multivariate case $E[P_n^-(t)]$ much more easily.

Before tackling the general case, and for expository purposes, we yet present the combinatorial analysis for $n = 2, 3$.

3. Expression for $E\left[\prod_{\ell=1}^{3}\sum_{i=1}^{N(t)+1} Y_i^{(\ell)}\right]$

We give the analogue of equation (2.4) for the trivariate case.

$$
\begin{aligned}
E\left[\sum_{i=1}^{N(t)+1} Y_i^{(1)} \sum_{i=1}^{N(t)+1} Y_i^{(2)} \sum_{i=1}^{N(t)+1} Y_i^{(3)}\right] &= E\left[\sum_{i=1}^{N(t)+1} Y_i^{(1)} Y_i^{(2)} Y_i^{(3)}\right] \\
&+E\left[\left(\sum_{i=1}^{N(t)+1} (i-1)\, Y_1^{(1)} Y_1^{(2)} Y_i^{(3)}\right) I(t_1 \leq t)\right] \\
&+E\left[\left(\sum_{i=1}^{N(t)+1} (i-1)\, Y_1^{(3)} Y_i^{(1)} Y_i^{(2)}\right) I(t_1 \leq t)\right] \\
&+E\left[\left(\sum_{i=1}^{N(t)+1} (i-1)\, Y_1^{(1)} Y_1^{(3)} Y_i^{(2)}\right) I(t_1 \leq t)\right] \\
&+E\left[\left(\sum_{i=1}^{N(t)+1} (i-1)\, Y_1^{(2)} Y_i^{(1)} Y_i^{(3)}\right) I(t_1 \leq t)\right] \\
&+E\left[\left(\sum_{i=1}^{N(t)+1} (i-1)\, Y_1^{(1)} Y_i^{(2)} Y_i^{(3)}\right) I(t_1 \leq t)\right] \\
&+E\left[\left(\sum_{i=1}^{N(t)+1} (i-1)\, Y_1^{(2)} Y_1^{(3)} Y_i^{(1)}\right) I(t_1 \leq t)\right] \\
&+E\left[\left(\sum_{i=1}^{N(t)+1} (i-1)(i-2)\, Y_1^{(1)} Y_2^{(2)} Y_i^{(3)}\right) I(t_1 + t_2 \leq t)\right] \\
&+E\left[\left(\sum_{i=1}^{N(t)+1} (i-1)(i-2)\, Y_1^{(2)} Y_2^{(3)} Y_i^{(1)}\right) I(t_1 + t_2 \leq t)\right] \\
&+E\left[\left(\sum_{i=1}^{N(t)+1} (i-1)(i-2)\, Y_1^{(3)} Y_2^{(1)} Y_i^{(2)}\right) I(t_1 + t_2 \leq t)\right]
\end{aligned} \tag{3.1}
$$

To understand the structure of Equation (3.1), let us do the combinatorial and probabilistic analysis for the case $n = 2$. We consider the partition of the set $\{1,2\}$ into m distinct ordered blocks or classes ($m = 1,2$), that we term *expectation summable classes*. In Equation (2.4), we distinguish 3 distinct *expectation summable classes* of product of random variables which are $Y_i^{(1)}Y_i^{(2)}$, $Y_1^{(1)}Y_i^{(2)}$, and $Y_1^{(2)}Y_i^{(1)}$. The class $Y_i^{(1)}Y_i^{(2)}$ is inherited from the case $n = 1$ and corresponds to ordered partition of the set $\{1,2\}$ into one single block ($\{1,2\}$) where 1 and 2 represent respectively $Y^{(1)}$ and $Y^{(2)}$. The classes $Y_1^{(1)}Y_i^{(2)}$ and $Y_1^{(2)}Y_i^{(1)}$ are *expectation summable classes* of random variables specific to the case $n = 2$; they correspond to ordered partitions of the set $\{1,2\}$ into two blocks, thus yielding the partitions ($\{1\},\{2\}$) and ($\{2\},\{1\}$). As we have already seen in Section 2.(b)i)–ii) and 2.(c)i)–ii), these two expectations summable classes are made-up of a mix of random variables which have identical expectation, complemented with "edge" random variables associated with an index $N(t)+1$. Thus, the class ($\{1\},\{2\}$) is made-up of all variables of the type $Y_i^{(1)}Y_j^{(2)}$ for $i \neq j$ and $i,j = 1,\ldots,N(t)$; complemented with $Y_i^{(1)}Y_{N(t)+1}^{(2)}$ variables for $i = 1,\ldots,N(t)$; which are quite probabilistically distinct form the previous ones. Indeed, we have $E\left[Y_i^{(1)}Y_j^{(2)}|N(t)\right] = E\left[Y_1^{(1)}Y_2^{(2)}|N(t)\right]$ and $E\left[Y_i^{(1)}Y_{N(t)+1}^{(2)}|N(t)\right] = E\left[Y_1^{(1)}Y_{N(t)+1}^{(2)}|N(t)\right]$, for a total number of $\dfrac{N(t)(N(t)-1)}{2}$ and $N(t)$ random variables falling in these two categories, whose union results in one *expectation summable class* given by $E\left[\sum_{i=1}^{N(t)+1}(i-1)Y_1^{(1)}Y_i^{(2)}\right]$, where we used martingale techniques for computing this expectation.

We can now perform the combinatorial and probabilistic analysis for Equation (3.1) based on the case $n=2$.

We consider the partition of the set $\{1,2,3\}$ into m distinct blocks or classes ($m=1,2,3$), that we term *expectation summable classes*. Therefore, in (3.1) we can distinguish 10 distinct *expectation summable classes* of product of random variables. Note that a subset of these classes can be considered as inherited from the cases $n=1$ and $n=2$.

More specifically, the class $Y_i^{(1)}Y_i^{(2)}Y_i^{(3)}$ is inherited form the case $n=1$ and corresponds to ordered partition of the set $\{1,2,3\}$ into one single block $(\{1,2,3\})$, where 1,2 and 3 represent respectively $Y^{(1)}, Y^{(2)}$ and $Y^{(3)}$. The classes $Y_1^{(1)}Y_1^{(2)}Y_i^{(3)}$, $Y_1^{(3)}Y_i^{(1)}Y_i^{(2)}$, $Y_1^{(1)}Y_1^{(3)}Y_i^{(2)}$, $Y_1^{(2)}Y_i^{(1)}Y_i^{(3)}$, $Y_1^{(1)}Y_i^{(2)}Y_i^{(3)}$ and $Y_1^{(2)}Y_1^{(3)}Y_i^{(1)}$ are all inherited from the case $n=2$ and correspond to ordered partitions of the set $\{1,2,3\}$ into two blocks, thus yielding the following partitions:
$(\{1,2\},\{3\})$, $(\{3\},\{1,2\})$, $(\{1,3\},\{2\})$, $(\{2\},\{1,3\})$, $(\{1\},\{2,3\})$, and $(\{2,3\},\{1\})$.
The various above expectations summables classes in (3.1) are made-up of a mix of random variables which have identical expectation, complemented with "edge" random variables associated with an index $N(t)+1$. For example, to understand the *expectation summable class* $(\{1,2\},\{3\})$ we refer to (b)i)–ii) in the beginning of the previous section, thus this class is made-up of all variables of the type $Y_i^{(1)}Y_i^{(2)}Y_j^{(3)}$ for $i \neq j$ and $i,j=1,\ldots,N(t)$; complemented with $Y_i^{(1)}Y_i^{(2)}Y_{N(t)+1}^{(3)}$ variables for $i=1,\ldots,N(t)$; which are probabilistically distinct form the previous ones. Indeed, as in Section 2(b)i)–ii), we have $E\left[Y_i^{(1)}Y_i^{(2)}Y_j^{(3)}|N(t)\right] = E\left[Y_1^{(1)}Y_1^{(2)}Y_{N(t)+1}^{(3)}|N(t)\right]$ and $E\left[Y_i^{(1)}Y_i^{(2)}Y_{N(t)+1}^{(3)}|N(t)\right] = E\left[Y_1^{(1)}Y_1^{(2)}Y_{N(t)+1}^{(3)}|N(t)\right]$, for a total number of $\frac{N(t)(N(t)-1)}{2}$ and $N(t)$ random variables occurring in these two categories whose union results in one *expectation summable class* given by $E\left[\sum_{i=1}^{N(t)+1}(i-1)Y_1^{(1)}Y_1^{(2)}Y_i^{(3)}\right]$. Clearly, the advantage of grouping the variables in this manner is that martingales techniques can be used as in (2.5) to compute the expectation of this summable class.

Finally, there are *expectation summable classes* of random variables specific to the case $n=3$; they correspond to ordered partitions of the set $\{1,2,3\}$ into three blocks, thus yielding the partitions:
$(\{1\},\{2\},\{3\})$, $(\{2\},\{3\},\{1\})$, and $(\{3\},\{1\},\{2\})$.
Indeed, while all variables $Y_i^{(1)}Y_j^{(2)}Y_k^{(3)}$ for $i \neq j \neq k$ and i,j,k less than $N(t)+1$, have essentially the same expectation, it is not so for the class $Y_i^{(1)}Y_j^{(2)}Y_{N(t)+1}^{(3)}$ $i \neq j$ less than $N(t)+1$, and the class $Y_i^{(1)}Y_j^{(3)}Y_{N(t)+1}^{(2)}$ for $i \neq j$ less than $N(t)+1$. Note also that we complement the $\sum_{i=1}^{N(t)}(i-$

$1)(i-2)$ random variables of the type $Y_i^{(1)}Y_j^{(2)}Y_k^{(3)}$ for $i \neq j \neq k$ and i, j, k less than $N(t)+1$, with $N(t)\Big(N(t)-1\Big)$ random variables of the type $Y_i^{(1)}Y_j^{(2)}Y_{N(t)+1}^{(3)}$.

To conclude this subsection, note that the 10 distinct partitions or summable classes given in (3.1) are not the total ordered partitions of the set $\{1,2,3\}$; indeed there are 13 ordered partitions. To be complete, we write the partitions $(\{2\},\{1\},\{3\})$, $(\{1\},\{3\},\{2\})$ and $(\{3\},\{2\},\{1\})$ which correspond to $Y_1^{(2)}Y_2^{(1)}Y_i^{(3)}$, $Y_1^{(1)}Y_2^{(3)}Y_i^{(2)}$ and $Y_1^{(3)}Y_2^{(2)}Y_i^{(1)}$ respectively. However, in each of these three partitions, the last random variable has been already occurred in the previous partitions $Y_1^{(1)}Y_2^{(2)}Y_i^{(3)}$, $Y_1^{(3)}Y_2^{(1)}Y_i^{(2)}$ and $Y_1^{(2)}Y_2^{(3)}Y_i^{(1)}$. The events given in the partitions $Y_1^{(1)}Y_2^{(2)}Y_i^{(3)}$, $Y_1^{(2)}Y_2^{(3)}Y_i^{(1)}$ and $Y_1^{(3)}Y_2^{(1)}Y_i^{(2)}$ are probabilistically the same as the partitions $Y_1^{(2)}Y_2^{(1)}Y_i^{(3)}$, $Y_1^{(3)}Y_1^{(2)}Y_i^{(1)}$ and $Y_1^{(1)}Y_2^{(3)}Y_i^{(2)}$ respectively; for this reason we did not consider them in (3.1).

3.1. Generalization of Equation (2.9). As we will use very often the idea of partition of the set $\{1,2,\ldots,n\}$ into m distinct blocks, let us define $\mathrm{mon}\Pi_\ell^{(m,n)}$ as the ℓ-th monomial associated with a partition Π of n letters into m non-empty distinct blocks having the same lower index in each block.

We generalize now Equation (2.9) in two steps (a) and (b), which is the most general form of (2.9).

(a) Using martingales techniques, as in (2.7), we can show that:

$$\begin{aligned} &E\Big[\Big(\sum_{i=1}^{N(t)+1}(i-1)(i-2)\cdots(i-k+1)Y_1^{(j^1)}Y_2^{(j^2)}\ldots Y_{k-1}^{(j^{k-1})}Y_i^{(j^k)}\Big)I\Big(\sum_{i=1}^{k-1} t_i \le t\Big)\Big] \\ &= (1/k)E[Y_k^{(j^k)}]E\Big[Y_1^{(j^1)}Y_2^{(j^2)}\ldots Y_{k-1}^{(j^{k-1})}\,\phi_k(t_{k-1}^*)I\Big(\sum_{i=1}^{k-1} t_i \le t\Big)\Big] + o(1) \end{aligned} \tag{3.2}$$

where

- the index 2 in (j^2) represents the belonging of $Y_2^{(j^2)}$ to the second block.
- $E[Y_k^{(j^k)}] = E[Y_1^{(j^k)}]$
- $t_{k-1}^* = t - \sum_{i=1}^{k-1} t_i$
- $\phi_k(t_{k-1}^*) = E\Big[\prod_{\ell=1}^{k}\Big(N(t_{k-1}^*)+\ell\Big)\Big]$

Note that on the LHS of (3.2) there is only one monomial with k blocks given by $Y_1^{(j^1)}Y_2^{(j^2)}\ldots Y_{k-1}^{(j^{k-1})}Y_i^{(j^k)}$. In the general context we have different monomials with k blocks, thus for generating these distinct monomials, the indices $\{j^1,j^2,\ldots,j^{k-1},j^k\}$ could be any permutation of the set $\{1,2,\ldots,k\}$ resulting each time in different monomials.

However, on the RHS of (3.2) there are two distinct monomials given by $Y_k^{(j^k)}$ and $Y_1^{(j^1)}Y_2^{(j^2)}\dots Y_k^{(j^{k-1})}$.

We show now Equation (3.2) as follows:

(1) under conditioning on $t_1,\dots,t_{k-1}$ we have:

$$
\begin{aligned}
&E\Big[\Big(\sum_{i=1}^{N(t)+1}(i-1)(i-2)\cdots(i-k+1)Y_1^{(j^1)}Y_2^{(j^2)}\dots Y_{k-1}^{(j^{k-1})}Y_i^{(j^k)}\Big)\\
&I\Big(\sum_{i=1}^{k-1} t_i \le t\Big)\Big]\\
&= E\Big[E\Big[\Big(\sum_{i=1}^{N(t)+1}(i-1)(i-2)\cdots(i-k+1)Y_1^{(j^1)}Y_2^{(j^2)}\dots Y_{k-1}^{(j^{k-1})}Y_i^{(j^k)}\Big)\\
&\quad I\Big(\sum_{i=1}^{k-1} t_i \le t\Big)|t_1,t_2,\dots,t_{k-1}\Big]\Big]\\
&= E\Big[E\Big[Y_1^{(j^1)}Y_2^{(j^2)}\dots Y_{k-1}^{(j^{k-1})}|t_1,\dots,t_{k-1}\Big]\\
&\quad E\Big[\Big(\sum_{i=1}^{N}(i-1)(i-2)\cdots(i-k+1)Y_i^{(j^k)}|t_1,\dots,t_{k-1}\Big)I\Big(\sum_{i=1}^{k-1} t_i \le t\Big)\Big]\Big]
\end{aligned}
\tag{3.3}
$$

where $N = N(t^*_{k-1}) + k$ and $N(t^*_{k-1})$ is defined only for $t \ge \sum_{i=1}^{k-1} t_i$.

(2) Consider the following sequence

$$S_n = \sum_{i=1}^{n}(i-1)\cdots(i-k+1)(Y_i^{(j^k)} - \mu)$$

for all finite $n \ge k$ and $\mu = E[Y_k^{(j^k)}]$.
Note that S_n is a martingale since $E\Big[|S_n|\Big] < \infty$ for all finite $n \ge k$ and $E\Big[S_{n+1}|S_k,S_{k+1},\dots,S_n\Big] = S_n$

(3) Consider the random variable $N = N(t^*_{k-1}) + k$ for $\sum_{i=1}^{k-1} t_i \le t$, and assuming that $Y_i^{(j^k)}$ are measurable functions of t_i, N will be a stopping time with respect to the σ-field $\mathcal{F}_n^S = \sigma(\omega, S_k, S_{k+1},\dots,S_n)$. If we assume that the fundamental renewal cycles are such that $E\Big[N(t)\Big]^\ell < \infty$ for all finite $\ell \ge 0$ and for all finite t, then $E[N] < \infty$. Now define $\xi_i = (i-1)\dots(i-k+1)(y_i^{(j^k)} - \mu), i = k,\dots,n$. Note that $E[\sum_{i=k}^{N} E|\xi_i|$ $|t_1,\dots,t_{k-1}] \le \widetilde{\mu}E[\frac{N(N-1)\dots(N-k+1)}{k}|t_1,\dots,t_{k-1}] < \infty$, where $\widetilde{\mu}$ is the upper bound on the expectation of the absolute value of individual rewards.

(4) It follows from Shiryayev's [7] problem 6 on page 464 that $E\Big[S_N|t_1,\dots,t_{k-1}\Big] = E\Big[S_k|t_1,\dots,t_{k-1}\Big] = 0$,

and after some algebra, we find

$$E\Big[\Big(\sum_{i=1}^{N}(i-1)\cdots(i-k+1)Y_i^{(j^k)}\Big)I\Big(\sum_{i=1}^{k-1} t_i \le t\Big)|t_1,\ldots,t_{k-1}\Big]$$
$$= (1/k)E[Y_k^{(j^k)}]E\Big[\Big(N(t_{k-1}^*)+1\Big)\cdots\Big(N(t_{k-1}^*+k\Big) \qquad (3.4)$$
$$I\Big(\sum_{i=1}^{k-1} t_i \le t\Big)|t_1,\ldots,t_{k-1}\Big]$$

Using (3.4) in (3.3) yields:

$$E\Big[E\Big[\Big(\sum_{i=1}^{N(t)+1}(i-1)(i-2)\cdots(i-k+1)Y_1^{(j^1)}Y_2^{(j^2)}\ldots Y_{k-1}^{(j^{k-1})}Y_i^{(j^k)}\Big)$$
$$I\Big(\sum_{i=1}^{k-1} t_i \le t\Big)|t_1,t_2,\ldots,t_{k-1}\Big]\Big]$$
$$= (1/k)E\Big[E\Big[Y_1^{(j^1)}Y_2^{(j^2)}\ldots Y_{k-1}^{(j^{k-1})}|t_1,\ldots,t_{k-1}\Big]E[Y_k^{(j^k)}]E\Big[\Big(N(t_{k-1}^*)+1\Big)$$
$$\cdots\Big(N(t_{k-1}^*)+k\Big)I\Big(\sum_{i=1}^{k-1} t_i \le t\Big)|t_1,\ldots,t_{k-1}\Big]\Big]$$

Adapting again Lemmas 5 and 6 in Adès and Malhamé [1], we have asymptotically: $E\Big[E\Big[Y_1^{(j^1)}Y_2^{(j^2)}\ldots Y_{k-1}^{(j^{k-1})}|t_1,\ldots,t_{k-1}\Big]\Big[\phi_k(t_{k-1}^*)I\Big(\sum_{i=1}^{k-1} t_i > t\Big)|t_1,\ldots,t_{k-1}\Big]\Big] = o(1)$ and $E\Big[E\Big[Y_1^{(j^1)}Y_2^{(j^2)}\ldots Y_{k-1}^{(j^{k-1})}|t_1,\ldots,t_{k-1}\Big]E\Big[\omega(t_{k-1}^*)|t_1,\ldots,t_{k-1}\Big]\Big] = o(1)$, where $\omega(t_{k-1}^*)$ is the remainder in the asymptotic expansion of $\phi_k(t_{k-1}^*)$ given $\sum_{i=1}^{k-1} t_i$; thus, Equation (3.2) follows.

(b) Using similar steps as in (3.1) (a), we can easily generalize Equation (3.2):

$$E\Big[\Big(\sum_{i=1}^{N(t)+1}\Big(\prod_{s=1}^{k-1}(i-s)\Big)\Big(\prod_{r\in I_1} Y_1^{(r)}\Big)\Big(\prod_{r\in I_2} Y_2^{(r)}\Big)\cdots\Big(\prod_{r\in I_{k-1}} Y_{k-1}^{(r)}\Big)\Big(\prod_{r\in I_k} Y_i^{(r)}\Big)\Big)I\Big(\sum_{i=1}^{k-1} t_i \le t\Big)\Big]$$
$$= (1/k)E\Big[E\Big[\Big(\prod_{r\in I_1} Y_1^{(r)}\Big)\Big(\prod_{r\in I_2} Y_2^{(r)}\Big)\cdots\Big(\prod_{r\in I_{k-1}} Y_{k-1}^{(r)}\Big)|t_1,\ldots,t_{k-1}\Big]$$
$$E\Big[\prod_{r\in I_k} Y_k^{(r)}\Big]E\Big[\prod_{\ell=1}^{k}\Big(N(t_{k-1}^*)+\ell\Big)I\Big(\sum_{i=1}^{k-1} t_i \le t\Big)|t_1,\ldots,t_{k-1}\Big]\Big] \qquad (3.5)$$
$$= (1/k)E\Big[\prod_{r\in I_k} Y_k^{(r)}\Big]E\Big[\Big(\prod_{r\in I_1} Y_1^{(r)}\Big)\Big(\prod_{r\in I_2} Y_2^{(r)}\Big)\cdots\Big(\prod_{r\in I_{k-1}} Y_{k-1}^{(r)}\Big)$$
$$\phi_k(t_{k-1}^*)I\Big(\sum_{i=1}^{k-1} t_i \le t\Big)\Big] + o(1) \qquad (3.6)$$

where

- $I_1, \ldots, I_k$ are k distinct and non-overlap blocks, and the random variables in each block are quite different from one another.
- $E\Big[\prod_{r\in I_k} Y_k^{(r)}\Big] = E\Big[\prod_{r\in I_k} Y_1^{(r)}\Big]$

More specifically, Equation (3.6) should be written as:

$$
\begin{aligned}
&E\Big[\Big(\sum_{i=1}^{N(t)+1}\Big(\prod_{s=1}^{k-1}(i-s)\Big)\Big(\prod_{r_\ell\in I_1} Y_1^{(j^1_{r_\ell})}\Big)\Big(\prod_{r_\ell\in I_2} Y_2^{(j^2_{r_\ell})}\Big)\cdots \\
&\Big(\prod_{r_\ell\in I_{k-1}} Y_{k-1}^{(j^{k-1}_{r_\ell})}\Big)\Big(\prod_{r_\ell\in I_k} Y_i^{(j^k_{r_\ell})}\Big)\Big) I\Big(\sum_{i=1}^{k-1} t_i \le t\Big)\Big] \\
&\quad = (1/k)E\Big[\prod_{r_\ell\in I_k} Y_k^{(j^k_{r_\ell})}\Big]E\Big[\Big(\prod_{r_\ell\in I_1} Y_1^{(j^1_{r_\ell})}\Big)\Big(\prod_{r_\ell\in I_2} Y_2^{(j^2_{r_\ell})}\Big)\cdots \\
&\qquad \Big(\prod_{r_\ell\in I_{k-1}} Y_{k-1}^{(j^{k-1}_{r_\ell})}\Big)\phi_k(t^*_{k-1}) I\Big(\sum_{i=1}^{k-1} t_i \le t\Big)\Big] + o(1)
\end{aligned}
\tag{3.7}
$$

where

- $\prod_{r_\ell\in I_1} Y_1^{(j^1_{r_\ell})} = Y_1^{(j^1_1)} Y_1^{(j^1_2)} \ldots Y_1^{(j^1_{r_1})}$.

  The upper and lower indices in (j^1_2) represent respectively the belonging of $Y^{(j^1_2)}$ to the first block and its position inside this block which has r_1 different random variables.
- $E\Big[\prod_{r_\ell\in I_k} Y_k^{(j^k_{r_\ell})}\Big] = E\Big[\prod_{r_\ell\in I_k} Y_1^{(j^k_{r_\ell})}\Big]$

Obviously, on the LHS of (3.6) or (3.7) there is only one monomial which is made-up of k distinct blocks, within each there are r_ℓ random variables having the same lower index; for those k blocks there are n different random variables.

But for the general context, we have several distinct monomials, thus, for generating such monomials $\Big\{(j^1_1, \ldots, j^1_{r_1}), \ldots, (j^k_1, \ldots, j^k_{r_k})\Big\}$ could be any permutation of the set $\{1, 2, \ldots, n\}$ such that $\sum_{\ell=1}^{k} r_\ell = n$.

However, on the RHS of (3.6) or (3.7) there are two distinct monomials given by $\prod_{r_\ell\in I_k} Y_k^{(j^k_{r_\ell})}$ and $\prod_{t=1}^{k-1}\Big(\prod_{r_\ell\in I_t} Y_t^{(j^t_{r_\ell})}\Big)$.

4. Partitioning the Product of Random Variables in $P_n^*(t)$

As usual we refer to n in $P_n^*(t)$ as the index in the product $E\Big[\prod_{\ell=1}^{n} \sum_{i=1}^{N(t)+1} Y_i^{(\ell)}\Big]$ for $n \geq 2$, where for $n = 1$, we have $E\Big[\sum_{i=1}^{N(t)+1} Y_i^{(1)}\Big] = E\Big[Y_1^{(1)}\Big]E\Big[N(t)+1\Big]$ which is the well-known Wald fundamental equation in sequential analysis.

Using the definition of $\text{mon}\Pi_\ell^{(m,n)}$ introduced in the previous section, we can write Equations (2.4) and (3.1) respectively as follows:

(1) $$E\Big[\sum_{i=1}^{N(t)+1} Y_i^{(1)} \sum_{i=1}^{N(t)+1} Y_i^{(2)}\Big] = E\Big[\sum_{i=1}^{N(t)+1} \text{mon}_i\Pi_1^{(1,2)}\Big]$$

$$E\Big[\sum_{i=1}^{N(t)+1} (i-1)\Big(\sum_{\ell=1}^{2} \text{mon}_i\Pi_\ell^{(2,2)}\Big) I(t_1 \leq t)\Big]$$

where $\text{mon}_i\Pi_1^{(1,2)} = Y_i^{(1)}Y_i^{(2)}$, $\text{mon}_i\Pi_1^{(2,2)} = Y_1^{(1)}Y_i^{(2)}$ and $\text{mon}_i\Pi_2^{(2,2)} = Y_1^{(2)}Y_i^{(1)}$.

(2) $$E\Big[\sum_{i=1}^{N(t)+1} Y_i^{(1)} \sum_{i=1}^{N(t)+1} Y_i^{(2)} \sum_{i=1}^{N(t)+1} Y_i^{(3)}\Big] = E\Big[\sum_{i=1}^{N(t)+1} \text{mon}_i\Pi_1^{(1,3)}\Big]$$

$$+E\Big[\sum_{i=1}^{N(t)+1} (i-1)\Big(\sum_{\ell=1}^{6} \text{mon}_i\Pi_\ell^{(2,3)}\Big) I(t_1 \leq t)\Big] + E\Big[\sum_{i=1}^{N(t)+1} (i-1)(i-2)\Big(\sum_{\ell=1}^{3} \text{mon}_i\Pi_\ell^{(3,3)}\Big) I(t_1+t_2 \leq t)\Big]$$

where $\text{mon}_i\Pi_\ell^{(m,3)}$ are given in (3.1).

It is well known, see e.g., Adès and Malhamé [1] that the total number of distinct partitions of the set $\{1,2,\ldots,n\}$ is given by

$$B_n = \sum_{m=1}^{n} S(n,m)$$

where $S(n,m)$ are Stirling's numbers of the second kind and B_n are Bell's numbers. Following the combinatorial and probabilistic analysis for the cases $n = 2$ and 3, the total number of ordered partitions of the set $\{1,2,\ldots,n\}$ for the case $N(t)+1$ is given by:

$$B_{n_p} = \sum_{m=1}^{n} mS(n,m); \quad (p \text{ for probabilistic})$$

Note that in general, the total number of ordered partitions of the set $\{1,2,\ldots,n\}$ is given by:

$$B_{n_o} = \sum_{m=1}^{n} m!S(n,m) = \frac{1}{2}\sum_{k=0}^{\infty}\Big(\frac{1}{2}\Big)^k k^n$$

For the case $n = 4$, we have $B_{n_p} = 37$ and $B_{n_o} = 75$.

We interpret 37 as the number of *expectation summable classes* and 75 as the total number of ordered partitions of the set $\{1,2,3,4\}$ into $m = 1,2,3,4$ blocks; clearly $B_{n_p} \subset B_{n_o}$. However, for probabilistic arguments similar to the case $n = 3$, we considered only 37 from these 75 partitions.

To find a general expression for $E\Big[P_n^*(t)|N(t)\Big]$, we need the following two equations:

- $$E\Big[\Big(\prod_{i\in I_1} Y_{j_1}^{(i)}\Big)\Big(\prod_{i\in I_2} Y_{j_2}^{(i)}\Big)\cdots\Big(\prod_{i\in I_m} Y_{j_m}^{(i)}\Big)|N(t)\Big] = E\Big[\Big(\prod_{i\in I_1} Y_1^{(i)}\Big)\Big(\prod_{i\in I_2} Y_2^{(i)}\Big)\cdots\Big(\prod_{i\in I_m} Y_m^{(i)}\Big)|N(t)\Big] \quad (4.1)$$

for $j_1, j_2, \dots, j_m$ pairwise distinct and indices from 1 to n partitioned into non-intersecting blocks $I_1, I_2, \dots, I_m$. (This is Corollary 1 in [1].)

- $$E\Big[\Big(\prod_{i\in I_1} Y_{j_1}^{(i)}\Big)\Big(\prod_{i\in I_2} Y_{j_2}^{(i)}\Big)\cdots\Big(\prod_{i\in I_{m-1}} Y_{j_{m-1}}^{(i)}\Big)Y_{N(t)+1}|N(t)\Big] = E\Big[\Big(\prod_{i\in I_1} Y_1^{(i)}\Big)\Big(\prod_{i\in I_2} Y_2^{(i)}\Big)\cdots\Big(\prod_{i\in I_{m-1}} Y_{m-1}^{(i)}\Big)Y_{N(t)+1}|N(t)\Big] \quad (4.2)$$

which is a consequence of Corollary 1 in [1].

We note again that the merge of product of random variables in (4.1) and (4.2) is an *expectation summable class* given by:

$$E\Big[\sum_{i=1}^{N(t)+1}\Big(\prod_{s=1}^{m-1}(i-s)\Big)\Big(\prod_{r\in I_1} Y_1^{(r)}\Big)\Big(\prod_{r\in I_2} Y_2^{(r)}\Big)\cdots\Big(\prod_{r\in I_{m-1}} Y_{m-1}^{(r)}\Big)\Big(\prod_{r\in I_m} Y_i^{(r)}\Big)|N(t)\Big] \quad (4.3)$$

$$= E\Big[\sum_{i=1}^{N(t)+1}\Big(\prod_{s=1}^{m-1}(i-s)\Big)\mathrm{mon}_i\Pi_1^{(m,n)}|N(t)\Big] \quad (4.4)$$

Obviously, there is only one monomial in (4.3) or (4.4), which is representative of $\frac{(N(t)+1)^{(m)}}{m}$ random variables falling in this category. More specifically, the total number of product of random variables in (4.1) and (4.2) is respectively $\frac{N(t)^{(m)}}{m}$ and $\frac{N(t)^{(m)}}{(N(t)-m+1)}$.

If we add expectation of all variables in these two categories we obtain one *expectation summable class* given in (4.3) or (4.4). We state now the following lemma which is the generalization of Equation (2.3).

LEMMA 1.

$$E\Big[P_n^*(t)|N(t)\Big] = \sum_{m=1}^{n} E\Big[\sum_{i=1}^{N(t)+1}\Big(\prod_{s=1}^{m-1}(i-s)\Big)\sum_{\ell=1}^{mS(n,m)} mon_i\Pi_\ell^{(m,n)} I\Big(N(t)\geq m\Big)|N(t)\Big]$$

Proof This lemma follows from Equation (4.1) and (4.2) and the discussion in the beginning of Section 3. Indeed, in $E\Big[P_n^*(t)|N(t)\Big]$, one can subdivide monomials into *expectation summable classes* corresponding to m blocks, $m = 1, 2, \dots, n$. Within the m blocks, there are $mS(n,m)$ distinct ways of partitioning the n components of the vector $\underline{Y}_n' = (Y^{(1)}\ Y^{(2)}\ \dots\ Y^{(n)})$ into $mS(n,m)$ ordered partitions, each of them is represented by virtue of Equations (4.2) and (4.3) by a single monomial which is inside this *expectation summable class*. The monomial $\mathrm{mon}_i\Pi_\ell^{(m,n)}$ is, for fixed ℓ, representative of $\frac{\left(N(t)+1\right)^{(m)}}{m}$ of product of random variables occurring in this category (this is equivalent to state that there are, for fixed ℓ, $\frac{\left(N(t)+1\right)^{(m)}}{m}$ monomials having the structure $\mathrm{mon}_i\Pi_\ell^{(m,n)}$ which is made-up of product of random variables). Thus, Lemma 1 is a mathematical expression of these facts.

As a consequence of Lemma 1 and the law of total probability we state the following theorem which generalizes Equation (3.1).

THEOREM 1.

$$E\left[P_n^*(t)\right] = \sum_{m=1}^{n} E\Big[\sum_{i=1}^{N(t)+1}\Big(\prod_{s=1}^{m-1}(i-s)\Big)\sum_{\ell=1}^{mS(n,m)} \mathrm{mon}_i\Pi_\ell^{(m,n)} I\Big(\sum_{i=1}^{m-1} t_i \le t\Big)\Big]$$

For computational facilities only, we agree that for $m=1$, $\prod_{s=1}^{m-1}(i-s)=1$ *and* $I\Big(\sum_{i=1}^{m-1} t_i \le t\Big)=1$.

We can now state the following theorem.

THEOREM 2.

$$E\left[P_n^*(t)\right] = \sum_{m=1}^{n} E\Big[(1/m)\sum_{\ell=1}^{mS(n,m)}\Big\{E\Big[\Big(\prod_{r\in I_1} Y_1^{(r)}\Big)\cdots\Big(\prod_{r\in I_{m-1}} Y_{m-1}^{(r)}\Big)| \\ t_1,\ldots,t_{m-1}\Big]E\Big[\prod_{r\in I_m} Y_m^{(r)}\Big]\Big\}_\ell E\Big[\prod_{s=1}^{m}\Big(N(t_{m-1}^*)+s\Big)I\Big(\sum_{i=1}^{m-1} t_i\le t\Big)|t_1,\ldots,t_{m-1}\Big]\Big]$$

Proof This theorem follows from Theorem 1, Equation (3.5) and noting that

$$\mathrm{mon}_i\Pi_\ell^{(m,n)} = \Big\{\Big(\prod_{r\in I_1} Y_1^{(r)}\Big)\Big(\prod_{r\in I_2} Y_2^{(r)}\Big)\cdots\Big(\prod_{r\in I_{m-1}} Y_{m-1}^{(r)}\Big)\Big(\prod_{r\in I_m} Y_i^{(r)}\Big)\Big\}_\ell$$

where $\sum_{j=1}^{m}(r_j \in I_j) = n$ (see the beginning of this section for some examples).

Note that Theorem 2 gives the exact expression for $E[P_n^*(t)]$. However, we are interested in finding an asymptotic expression for that expectation; this will be performed in the next section.

5. Asymptotic Behaviour of $E\left[P_n^*(t)\right]$

To characterize the asymptotic behaviour of $E\left[P_n^*(t)\right]$, we have to study the asymptotics of terms of the form:

$$\begin{aligned} &E\Big[E\Big[\Big(\prod_{r\epsilon I_1} Y_1^{(r)}\Big)\cdots\Big(\prod_{r\epsilon I_{m-1}} Y_{m-1}^{(r)}\Big)|t_1,\ldots,t_{m-1}\Big]E\Big[\prod_{r\epsilon I_m} Y_m^{(r)}\Big] \\ &\quad E\Big[\prod_{s=1}^{m}\Big(N(t_{m-1}^*)+s\Big)I\Big(\sum_{i=1}^{m-1} t_i \le t\Big)|t_1,\ldots,t_{m-1}\Big]\Big] \end{aligned} \tag{5.1}$$

Now, as $E\Big[\prod_{r=1}^{m}\Big(N(t_m^*)+r\Big)\Big]$ is $\phi_m(t_m^*)$ with $t_1,t_2,\ldots,t_m$ treated as parameters, and where $\phi_m(t)$ is the factorial moment of $N(t)$ (as defined by Smith [9]). We can use the asymptotic theory of $\phi_m(t)$ (Smith [9]). We gather the following facts useful for our analysis.

"**Definition 1** *The function $\lambda(t)$ belongs to the class B if and only if it is bounded variation, tends to zero as t approaches $+\infty$ and satisfies the condition $\lambda(t)-\lambda(t-\alpha)=o(t^{-1})$ as $t\to+\infty$, for every $\alpha>0$.*

Theorem 1 *If $\mu_{n+1}<\infty$ then $\phi_n(t)=\gamma_1 t^n+\gamma_2 t^{n-1}+\cdots+\gamma_n t+\gamma_{n+1}+\omega(t)$ where $\omega(t)\,\epsilon\, B$".*

Therefore it follows from Smith's [9] Theorem 1, that $\omega(t)=o(1)$ as $t\to\infty$. Note that γ_i's $i=1,2,\ldots,n$ represent finite rational functions of $\mu_1,\mu_2,\ldots,\mu_i$.

Using essentially the same approach as in Section 6 of Adès and Malhamé [1], we write:

$$\begin{aligned} &E\Big[E\Big[\Big(\prod_{r\epsilon I_1} Y_1^{(r)}\Big)\cdots\Big(\prod_{r\epsilon I_{m-1}} Y_{m-1}^{(r)}\Big)|t_1,\ldots,t_{m-1}\Big]\Big[\Big(\gamma_1(t_{m-1}^*)^m \\ &\quad+\gamma_2(t_{m-1}^*)^{m-1}+\cdots+\gamma_{m+1}+\omega(t_{m-1}^*)\Big)I\Big(\sum_{i=1}^{m-1} t_i\le t\Big)|t_1,\ldots,t_{m-1}\Big]E\Big[\prod_{r\epsilon I_m} Y_m^{(r)}\Big]\Big] \\ &\quad= E\Big[\Big(\prod_{r\epsilon I_1} Y_1^{(r)}\Big)\cdots\Big(\prod_{r\epsilon I_{m-1}} Y_{m-1}^{(r)}\Big)\Big(\gamma_1(t_{m-1}^*)^m+\gamma_2(t_{m-1}^*)^{m-1}+\cdots \\ &\quad\quad+\gamma_{m+1}+\omega(t_{m-1}^*)\Big)I\Big(\sum_{i=1}^{m-1} t_i\le t\Big)E\Big[\prod_{r\epsilon I_m} Y_m^{(r)}\Big]\Big] \end{aligned} \tag{5.2}$$

Using Lemmas 5 and 6 in [1], we can show that:

- $E\Big[\Big(\prod_{r\epsilon I_1} Y_1^{(r)}\Big)\cdots\Big(\prod_{r\epsilon I_{m-1}} Y_{m-1}^{(r)}\Big)\gamma(t_{m-1}^*)^m I\Big(\sum_{i=1}^{m-1} t_i>t\Big)\Big]=o(1)$
- $E\Big[\Big(\prod_{r\epsilon I_1} Y_1^{(r)}\Big)\cdots\Big(\prod_{r\epsilon I_{m-1}} Y_{m-1}^{(r)}\Big)\omega(t_{m-1}^*)\Big]=o(1)$

Therefore we state without proof the following theorem for an asymptotic expression of $E\left[P_n^*(t)\right]$.

THEOREM 3.

$$E\Big[P_n^*(t)\Big] = \sum_{m=1}^{n}(1/m)\sum_{\ell=1}^{mS(n,m)}\Big\{E\Big[\Big(\prod_{r\in I_1} Y_1^{(r)}\Big)\cdots\Big(\prod_{r\in I_{m-1}} Y_{m-1}^{(r)}\Big) \phi_m(t_{m-1}^*) I\Big(\sum_{i=1}^{m-1} t_i \le t\Big)\Big] E\Big[\prod_{r\in I_m} Y_m^{(r)}\Big]\Big\}_\ell + o(1)$$

Note that the sum of random variables in $I_1,\ldots,I_{m-1}$ and I_m is n. We can also write Theorem 3 in a compact form as follows:

$$E\Big[P_n^*(t)\Big] = \sum_{m=1}^{n}(1/m)\sum_{\ell=1}^{mS(n,m)} E\Big[\text{mon}\Pi_\ell^{(m-1,n_1)}\phi_m(t_{m-1}^*) I\Big(\sum_{i=1}^{m-1} t_i \le t\Big)\Big] E\Big[\text{mon}\Pi_\ell^{(1,n_2)}\Big] + o(1) \tag{5.3}$$

such that $n_1 + n_2 = n$ and $n_1 = 1,\ldots,n$; finally note that for computation facilities, $t_0^* = t$ and $\Pi_\ell^{(0,0)} = 1$.

6. Applications

In this section, we evaluate explicitly and asymptotically $E\Big[\sum_{i=1}^{N(t)+1} Y_i^{(1)} \sum_{i=1}^{N(t)+1} Y_i^{(2)} \sum_{i=1}^{N(t)+1} Y_i^{(3)}\Big]$.

From Smith's [9] Lemma 6, we have

$$\phi_m^*(s) = m!\Big\{1 - F^*(s)\Big\}^{-m} \tag{6.1}$$

where $\phi_m^*(s)$ and $F^*(s)$ are Laplace-Stieltjes transform of $\phi_m(t)$ and $F(t)$ respectively, and $F^*(s)$ is given by Lemma 3 in Smith [9] as

$$F^*(s) = 1 - \mu_1 s + \frac{\mu_2 s^2}{2!} - \cdots + \frac{(-s)^n \mu_n}{n!} + o(s^n) \tag{6.2}$$

for real $s > 0$.

Equations (6.1) and (6.2) are the basis for computing $\phi_m(t_{m-1}^*)$.

It follows from our Theorem 3 that:

$$
\begin{aligned}
E\Big[\sum_{i=1}^{N(t)+1} Y_i^{(1)} \sum_{i=1}^{N(t)+1} Y_i^{(2)} \sum_{i=1}^{N(t)+1} Y_i^{(3)}\Big] &= E\Big[Y_1^{(1)}Y_1^{(2)}Y_1^{(3)}\Big]\phi_1(t_0^*) \\
&+(1/2)E\Big[Y_1^{(1)}Y_1^{(2)}\phi_2(t_1^*)I(t_1 \le t)\Big]E\Big[Y_2^{(3)}\Big] \\
&+(1/2)E\Big[Y_1^{(1)}Y_1^{(3)}\phi_2(t_1^*)I(t_1 \le t)\Big]E\Big[Y_2^{(2)}\Big] \\
&+(1/2)E\Big[Y_1^{(2)}Y_1^{(3)}\phi_2(t_1^*)I(t_1 \le t)\Big]E\Big[Y_2^{(1)}\Big] \\
&+(1/2)E\Big[Y_1^{(1)}\phi_2(t_1^*)I(t_1 \le t)\Big]E\Big[Y_2^{(2)}Y_2^{(3)}\Big] \\
&+(1/2)E\Big[Y_1^{(2)}\phi_2(t_1^*)I(t_1 \le t)\Big]E\Big[Y_2^{(1)}Y_2^{(3)}\Big] \\
&+(1/2)E\Big[Y_1^{(3)}\phi_2(t_1^*)I(t_1 \le t)\Big]E\Big[Y_2^{(1)}Y_2^{(2)}\Big] \\
&+(1/3)E\Big[Y_1^{(1)}Y_2^{(2)}\phi_3(t_2^*)I(t_1 + t_2) \le t\Big]E\Big[Y_3^{(3)}\Big] \\
&+(1/3)E\Big[Y_1^{(2)}Y_2^{(3)}\phi_3(t_2^*)I(t_1 + t_2) \le t\Big]E\Big[Y_3^{(1)}\Big] \\
&+(1/3)E\Big[Y_1^{(3)}Y_2^{(1)}\phi_3(t_2^*)I(t_1 + t_2) \le t\Big]E\Big[Y_3^{(2)}\Big] + o(1)
\end{aligned}
\tag{6.3}
$$

Recall that $t_{m-1}^* = t - \sum_{i=1}^{m-1} t_i$, $t_0^* = 1$ and $\phi_m(t) = E\Big[\Big(N(t)+1\Big)\cdots\Big(N(t)+m\Big)\Big]$.

On the RHS of (6.3) we have:

(1) One single partition of the vector $\underline{Y}_3' = \Big(Y^{(1)}\ Y^{(2)}\ Y^{(3)}\Big)$ in one block which is $\Big(\{Y_1^{(1)}Y_1^{(2)}Y_1^{(3)}\}\Big)$.

(2) Six ordered partitions of the vector $\underline{Y}_3' = \Big(Y^{(1)}\ Y^{(2)}\ Y^{(3)}\Big)$ in two blocks given by $\Big(\{Y_1^{(1)}Y_1^{(2)}\},\{Y_2^{(3)}\}\Big)$, $\Big(\{Y_1^{(1)}Y_1^{(3)}\},\{Y_2^{(2)}\}\Big)$, $\Big(\{Y_1^{(2)}Y_1^{(3)}\},\{Y_2^{(1)}\}\Big)$, $\Big(\{Y_1^{(1)}\},\{Y_2^{(2)}Y_2^{(3)}\}\Big)$, $\Big(\{Y_1^{(2)}\},\{Y_2^{(1)}Y_2^{(3)}\}\Big)$ and $\Big(\{Y_1^{(3)}\},\{Y_2^{(1)}Y_2^{(2)}\}\Big)$.

(3) Three ordered partitions of the vector $\underline{Y}_3' = \Big(Y^{(1)}\ Y^{(2)}\ Y^{(3)}\Big)$ in three blocks which are $\Big(\{Y_1^{(1)}\},\{Y_2^{(2)}\},\{Y_3^{(3)}\}\Big)$, $\Big(\{Y_1^{(2)}\},\{Y_2^{(3)}\},\{Y_3^{(1)}\}\Big)$ and $\Big(\{Y_1^{(3)}\},\{Y_2^{(1)}\},\{Y_3^{(2)}\}\Big)$.

We assume that $Y_1^{(1)}, Y_2^{(2)}$ and $Y_3^{(3)}$ are correlated with t_1, t_2 and t_3 respectively. Using (6.1) and (6.2) we obtain after a very tedious algebra the following:

$$E\Big[\sum_{i=1}^{N(t)+1} Y_i^{(1)} \sum_{i=1}^{N(t)+1} Y_i^{(2)} \sum_{i=1}^{N(t)+1} Y_i^{(3)}\Big] = \frac{E[Y_1^{(1)}]E[Y_1^{(2)}]E[Y_1^{(3)}]}{\mu_1^3}t^3$$

$$+\frac{1}{\mu_1^2}\Big\{E[Y_1^{(1)}]E[Y_1^{(2)}Y_1^{(3)}] + E[Y_1^{(2)}]E[Y_1^{(1)}Y_1^{(3)}] + E[Y_1^{(3)}]E[Y_1^{(1)}Y_1^{(2)}]$$

$$-\frac{2}{\mu_1^3}\Big(E[Y_1^{(1)}]E[Y_1^{(2)}]E[Y_1^{(3)}t_1] + E[Y_1^{(1)}]E[Y_1^{(3)}]E[Y_1^{(2)}t_1]$$

$$+E[Y_1^{(2)}]E[Y_1^{(3)}]E[Y_1^{(1)}t_1]\Big) + \frac{9}{2}\frac{\mu_2}{\mu_1^4}E[Y_1^{(1)}]EY_1^{(2)}]E[Y_1^{(3)}]\Big\}t^2$$

$$+\Big\{\frac{E[Y_1^{(1)}Y_1^{(2)}Y_1^{(3)}]}{\mu_1} + \frac{2E[Y_1^{(1)}]E[Y_1^{(2)}Y_1^{(3)}]\mu_2}{\mu_1^3}$$

$$+\frac{2E[Y_1^{(2)}]E[Y_1^{(1)}Y_1^{(3)}]\mu_2}{\mu_1^3} + \frac{2E[Y_1^{(3)}]E[Y_1^{(1)}Y_1^{(2)}]\mu_2}{\mu_1^3}$$

$$-\frac{E[Y_1^{(1)}]E[Y_1^{(2)}Y_1^{(3)}]t_1}{\mu_1^2} - \frac{E[Y_1^{(2)}]E[Y_1^{(1)}Y_1^{(3)}t_1]}{\mu_1^2}$$

$$-\frac{E[Y_1^{(3)}]E[Y_1^{(1)}Y_1^{(2)}t_1]}{\mu_1^2} - \frac{E[Y_1^{(1)}Y_1^{(2)}]E[Y_1^{(3)}t_1]}{\mu_1^2}$$

$$-\frac{E[Y_1^{(1)}Y_1^{(3)}]E[Y_1^{(2)}t_1]}{\mu_1^2} - \frac{E[Y_1^{(2)}Y_1^{(3)}]E[Y_1^{(1)}t_1]}{\mu_1^2}$$

$$+\frac{2E[Y_1^{(1)}]E[Y_1^{(3)}]E[Y_1^{(2)}t_1^2]}{\mu_1^3} + \frac{2E[Y_1^{(2)}]E[Y_1^{(3)}]E[Y_1^{(1)}t_1^2]}{\mu_1^3}$$

$$+\frac{2E[Y_1^{(1)}]E[Y_1^{(2)}]E[Y_1^{(3)}t_1^2]}{\mu_1^3} + \frac{2E[Y_1^{(1)}]E[Y_1^{(2)}t_1]E[Y_1^{(3)}t_1]}{\mu_1^3}$$

$$+\frac{2E[Y_1^{(2)}]E[Y_1^{(1)}t_1]E[Y_1^{(3)}t_1]}{\mu_1^3} + \frac{2E[Y_1^{(3)}]E[Y_1^{(1)}t_1]E[Y_1^{(2)}t_1]}{\mu_1^3}$$

$$-6E[Y_1^{(1)}]E[Y_1^{(3)}]E[Y_1^{(2)}t_1]\frac{\mu_2}{\mu_1^4} - 6E[Y_1^{(1)}]E[Y_1^{(2)}]E[Y_1^{(3)}t_1]\frac{\mu_2}{\mu_1^4}$$

$$-6E[Y_1^{(2)}]E[Y_1^{(3)}]E[Y_1^{(1)}t_1]\frac{\mu_2}{\mu_1^4}$$

$$+3\left(\frac{3\mu_2^2 - \mu_1\mu_3}{\mu_1^5}\right)E[Y_1^{(1)}]E[Y_1^{(2)}]E[Y_1^{(3)}]\Big\}t$$

$$+\Big\{\frac{\mu_2}{2\mu_1^2}E[Y_1^{(1)}Y_1^{(2)}Y_1^{(3)}] + \frac{E[Y_1^{(1)}]E[Y_1^{(2)}Y_1^{(3)}t_1^2]}{2\mu_1^2}$$

$$+\frac{E[Y_1^{(2)}]E[Y_1^{(1)}Y_1^{(3)}t_1^2]}{2\mu_1^2} + \frac{E[Y_1^{(3)}]E[Y_1^{(1)}Y_1^{(2)}t_1^2]}{2\mu_1^2}$$

$$
\begin{aligned}
&+\frac{E[Y_1^{(1)}Y_1^{(2)}]E[Y_1^{(3)}t_1^2]}{2\mu_1^2}+\frac{E[Y_1^{(1)}Y_1^{(3)}]E[Y_1^{(2)}t_1^2]}{2\mu_1^2}\\
&+\frac{E[Y_1^{(2)}Y_1^{(3)}]E[Y_1^{(1)}t_1^2]}{2\mu_1^2}-\frac{\mu_2E[Y_1^{(1)}]E[Y_1^{(2)}Y_1^{(3)}t_1]}{\mu_1^3}\\
&-\frac{\mu_2E[Y_1^{(2)}]E[Y_1^{(1)}Y_1^{(3)}t_1]}{\mu_1^3}-\frac{\mu_2E[Y_1^{(3)}]E[Y_1^{(1)}Y_1^{(2)}t_1]}{\mu_1^3}\\
&-\frac{\mu_2E[Y_1^{(1)}Y_1^{(2)}]E[Y_1^{(3)}t_1]}{\mu_1^3}-\frac{\mu_2E[Y_1^{(1)}Y_1^{(3)}]E[Y_1^{(2)}t_1]}{\mu_1^3}\\
&-\frac{\mu_2E[Y_1^{(2)}Y_1^{(3)}]E[Y_1^{(1)}t_1]}{\mu_1^3}-\frac{E[Y_1^{(1)}]E[Y_1^{(2)}t_1]E[Y_1^{(3)}t_1^2]}{\mu_1^3}\\
&-\frac{E[Y_1^{(1)}]E[Y_1^{(2)}t_1^2]E[Y_1^{(3)}t_1]}{\mu_1^3}-\frac{E[Y_1^{(2)}]E[Y_1^{(3)}t_1]E[Y_1^{(1)}t_1^2]}{\mu_1^3}\\
&-\frac{E[Y_1^{(2)}]E[Y_1^{(1)}t_1]E[Y_1^{(3)}t_1^2]}{\mu_1^3}-\frac{E[Y_1^{(3)}]E[Y_1^{(1)}t_1]E[Y_1^{(2)}t_1^2]}{\mu_1^3}\\
&-\frac{E[Y_1^{(3)}]E[Y_1^{(2)}t_1]E[Y_1^{(1)}t_1^2]}{\mu_1^3}-\frac{2E[Y_1^{(1)}]E[Y_1^{(2)}]E[Y_1^{(3)}t_1^3]}{3\mu_1^3}\\
&-\frac{2E[Y_1^{(2)}]E[Y_1^{(3)}]E[Y_1^{(1)}t_1^3]}{3\mu_1^3}-\frac{2E[Y_1^{(1)}]E[Y_1^{(3)}]E[Y_1^{(2)}t_1^3]}{3\mu_1^3}\\
&+\frac{3\mu_2E[Y_1^{(1)}]E[Y_1^{(2)}]E[Y_1^{(3)}t_1^2]}{\mu_1^4}+\frac{3\mu_2E[Y_1^{(2)}]E[Y_1^{(3)}]E[Y_1^{(1)}t_1^2]}{\mu_1^4}\\
&+\frac{3\mu_2E[Y_1^{(1)}]E[Y_1^{(3)}]E[Y_1^{(2)}t_1^2]}{\mu_1^4}+\frac{3\mu_2E[Y_1^{(1)}]E[Y_1^{(2)}t_1]E[Y_1^{(3)}t_1]}{\mu_1^4}\\
&+\frac{3\mu_2E[Y_1^{(2)}]E[Y_1^{(1)}t_1]E[Y_1^{(3)}t_1]}{\mu_1^4}+\frac{3\mu_2E[Y_1^{(3)}]E[Y_1^{(1)}t_1]E[Y_1^{(2)}t_1]}{\mu_1^4}\\
&+2\varphi_1E[Y_1^{(1)}]E[Y_1^{(2)}Y_1^{(3)}]+2\varphi_1E[Y_1^{(2)}]E[Y_1^{(1)}Y_1^{(3)}]\\
&+2\varphi_1E[Y_1^{(3)}]E[Y_1^{(1)}Y_1^{(2)}]-2\varphi_2E[Y_1^{(1)}]E[Y_1^{(2)}]E[Y_1^{(3)}t_1]\\
&-2\varphi_2E[Y_1^{(1)}]E[Y_1^{(3)}]E[Y_1^{(2)}t_1]-2\varphi_2E[Y_1^{(2)}]E[Y_1^{(3)}]E[Y_1^{(1)}t_1]\\
&+3\varphi_3E[Y_1^{(1)}]E[Y_1^{(2)}]E[Y_1^{(3)}]\Big\}+o(1) \qquad (6.4)
\end{aligned}
$$

where

$$
\begin{cases}
\varphi_1=\left(\dfrac{9\mu_2^2-4\mu_1\mu_3}{12\mu_1^4}\right)\\
\varphi_2=\left(\dfrac{3\mu_2^2-\mu_1\mu_3}{\mu_1^5}\right)\\
\varphi_3=\left(\dfrac{4\mu_2^3-\mu_4\mu_1^2}{4\mu_1^6}\right)
\end{cases}
$$

If $Y_i^{(1)}, Y_i^{(2)}$ and $Y_i^{(3)}$ are identical, (6.4) yields the special case:

$$
\begin{aligned}
E\Bigg[\sum_{i=1}^{N(t)+1} Y_i^{(1)}\Bigg]^3 = {} & \frac{E^3[Y_1^{(1)}]}{\mu_1^3}t^3 + \Bigg\{\frac{3E[Y_1^{(1)}]EY_1^{(1)}]^2}{\mu_1^2} \\
& -\frac{6E^2[Y_1^{(1)}]E[Y_1^{(1)}t_1]}{\mu_1^3} + \frac{9\mu_2 E^3[Y_1^{(1)}]}{2\mu_1^4}\Bigg\}t^2 \\
& +\Bigg\{\frac{E[Y_1^{(1)}]^3}{\mu_1} + \frac{6\mu_2 E[Y_1^{(1)}]E[Y_1^{(1)}]^2}{\mu_1^3} \\
& -\frac{3E[Y_1^{(1)}]E\Big[(Y_1^{(1)})^2 t_1\Big]}{\mu_1^2} - \frac{3E[Y_1^{(1)}]^2 E[Y_1^{(1)}t_1]}{\mu_1^2} \\
& +\frac{6E^2[Y_1^{(1)}]E[Y_1^{(1)}t_1^2]}{\mu_1^3} + \frac{6E[Y_1^{(1)}]E^2[Y_1^{(1)}t_1]}{\mu_1^3} \\
& -\frac{18E^2[Y_1^{(1)}]E[Y_1^{(1)}t_1]\mu_2}{\mu_1^4} + \frac{9\mu_2^2 E^3[Y_1^{(1)}]}{\mu_1^5} \\
& -\frac{3\mu_3 E^3[Y_1^{(1)}]}{\mu_1^4}\Bigg\}t + \Bigg\{\frac{\mu_2 E[Y_1^{(1)}]^3}{2\mu_1^2} \\
& +\frac{3E[Y_1^{(1)}]E\Big[(Y_1^{(1)})^2 t_1^2\Big]}{2\mu_1^2} + \frac{3E[Y_1^{(1)}]^2 E[Y_1^{(1)}t_1^2]}{2\mu_1^2} \\
& -\frac{3\mu_2 E[Y_1^{(1)}]E\Big[(Y_1^{(1)})^2 t_1\Big]}{\mu_1^3} - \frac{3\mu_2 E[Y_1^{(1)}]^2 E[Y_1^{(1)}t_1]}{\mu_1^3} \\
& +\frac{6E[Y_1^{(1)}]E[Y_1^{(1)}t_1]E[Y_1^{(1)}t_1^2]}{\mu_1^3} - \frac{2E^2[Y_1^{(1)}]E[Y_1^{(1)}t_1^3]}{\mu_1^3} \\
& +\frac{9\mu_2 E^2[Y_1^{(1)}]E[Y_1^{(1)}t_1^2]}{\mu_1^4} + \frac{9\mu_2 E[Y_1^{(1)}]E^2[Y_1^{(1)}t_1]}{\mu_1^4} \\
& +\frac{9\mu_2^2 E[Y_1^{(1)}]E[Y_1^{(1)}]^2}{2\mu_1^4} - \frac{2E[Y_1^{(1)}]E[Y_1^{(1)}]^2\mu_3}{\mu_1^3} \\
& +\frac{18E^2[Y_1^{(1)}]E[Y_1^{(1)}t_1]\mu_2^2}{\mu_1^5} + \frac{6E^2[Y_1^{(1)}]E[Y_1^{(1)}t_1]\mu_3}{\mu_1^4} \\
& +\frac{3E^3[Y_1^{(1)}]\mu_2^3}{\mu_1^6} - \frac{3E^3[Y_1^{(1)}]\mu_4}{4\mu_1^4}\Bigg\} + o(1)
\end{aligned}
\tag{6.5}
$$

Note that in the development of (6.5) we used the following

$E[Y_m^{(\ell)}] = E[Y_1^{(\ell)}]$ $\quad \ell = 1,2,3$ and $m = 1,2,3$

$E[Y_m^{(r)}Y_m^{(s)}] = E[Y_1^{(r)}Y_1^{(s)}]$ $\quad \forall r \neq s, r, s = 1,2,3$

$E[Y_m^{(\ell)}t_m^k] = E[Y_1^{(\ell)}t_1^k]$ $\quad \forall k < \infty$

Also note that Equations (6.4) and (6.5) are the same as those obtained in Adès and Malhamé [2]. This provides an independent validation of the approach elaborated in the present chapter.

7. Recursive Scheme of $P_n^*(t)$

The monomials occurring in $E\left[P_n^*(t)\right]$ are generated recursively and inherited from each other following by that a specific pattern according to the analysis elaborated in this paper.

Let us consider the following example:

$$\text{mon}\Pi_2^{(3,4)} = Y_1^{(4)}, Y_2^{(1)}Y_2^{(2)}, Y_3^{(3)}$$

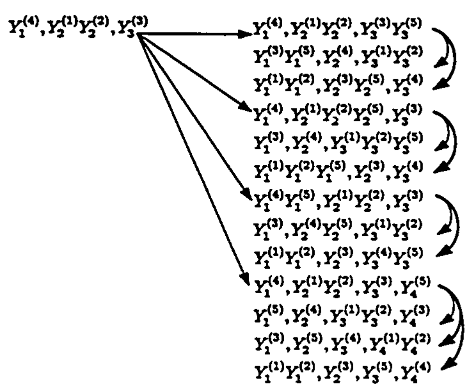
For this monomial, we have the following remarks:

- $\text{mon}\Pi_2^{(3,4)} = Y_1^{(4)}, Y_2^{(1)}Y_2^{(2)}, Y_3^{(3)}$ is inherited from $\text{mon}\Pi_1^{(2,3)} = Y_1^{(1)}Y_1^{(2)}, Y_2^{(3)}$
- $\text{mon}\Pi_2^{(3,4)}$ is one of 37 monomials occurring in the fourth generation according to $B_{n_p} = \sum_{m=1}^{4} mS(4,m) = 37$, as defined previously.
- $\text{mon}\Pi_2^{(3,4)}$ is associated with a partition of 4 letters into 3 blocks separated by 2 commas, where the last one has a special meaning in the context of Theorem 3.

  Indeed, we have by this theorem:

$$E\left[\text{mon}\Pi_2^{(3,4)}\right] = E\left[Y_1^{(4)}, Y_2^{(1)}Y_2^{(2)}, Y_3^{(3)}\right] \xrightarrow{th.3} E\left[Y_1^{(4)}Y_2^{(1)}Y_2^{(2)}\phi_3(t_2^*)I(t_1 + t_2 \le t)\right]E[Y_3^{(3)}].$$

- Thus, in the context of Theorem 3, we have 2 expectations associated with this monomial. The first expectation is applied to the first blocks located before the last comma, whereas the second expectation is specific to the block after this last comma. Therefore, it is clear by Theorem 3, that we have only 2 expectations applied to every monomial.
- $\text{mon}\Pi_2^{(3,4)} = Y_1^{(3)}, Y_2^{(1)}Y_2^{(2)}, Y_3^{(3)}$ generates in the fifth generation the following distinct monomials:



- As we can remark, $\text{mon}\Pi_2^{(3,4)}$ generates in the fifth generation:
  - (a) 9 distinct monomials $\triangleq \text{mon}\Pi_\ell^{(3,5)}$, which are associated with a partition of 5 letters into 3 blocks. For the fifth generation, we have $3S(5,3) = 75$ distinct monomials of 5 letters with 3 blocks.
  - (b) 4 distinct augmented monomials $\triangleq \text{mon}\Pi_\ell^{(4,5)}$. There are $4S(5,4) = 40$ distinct monomials of 5 letters with 4 blocks.
- The mechanism for generating these monomials is simple, which can be summarized by the next three steps:
  - (1) The new letter $Y^{(5)}$ has to be located at the end of each block, taking as lower index the position number of this block.
  - (2) When the new letter $Y^{(5)}$ is so fixed, we have a new monomial, from which we follow a cyclicity operation on the present blocks, generating thereafter the appropriate monomials. We illustrate this cyclicity by the rounded arrow (⤵).
  - (3) Once the new letter $Y^{(5)}$ has taken place at the end of each block, then $Y^{(5)}$ has its own or new block, where the position number of this new block is 4. We have then an augmented (in the number of blocks) monomial from which we follow step 2 as above.

Note that in practice, we do not generate monomials from $\text{mon}\Pi_2^{(3,4)}$⤵, since those monomials are already generated from $\text{mon}\Pi_1^{(3,4)}$.

First Generation $\underline{Y}_1' = (Y^{(1)})$

1.1.1 $\text{mon}\Pi_1^{(1,1)} = Y_1^{(1)}$

Second Generation $\underline{Y}'_2 = (Y^{(1)}\ Y^{(2)})$

2.1.1 $Y_1^{(1)} \rightarrow$

$\mathrm{mon}\Pi_1^{(1,2)} = Y_1^{(1)} Y_1^{(2)}$

$\mathrm{mon}\Pi_1^{(2,2)} = Y_1^{(1)}, Y_2^{(2)}$

$\mathrm{mon}\Pi_2^{(2,2)} = Y_1^{(2)}, Y_2^{(1)}$

Third Generation $\underline{Y}'_3 = (Y^{(1)}\ Y^{(2)}\ Y^{(3)})$

3.1.1 $Y_1^{(1)} Y_1^{(2)} \rightarrow$

$\mathrm{mon}\Pi_1^{(1,3)} = Y_1^{(1)} Y_1^{(2)} Y_1^{(3)}$

$\mathrm{mon}\Pi_1^{(2,3)} = Y_1^{(1)} Y_1^{(2)}, Y_2^{(3)}$

$\mathrm{mon}\Pi_2^{(2,3)} = Y_1^{(3)}, Y_2^{(1)} Y_2^{(2)}$

3.2.1 $Y_1^{(1)}, Y_2^{(2)} \rightarrow$

$\mathrm{mon}\Pi_3^{(2,3)} = Y_1^{(1)}, Y_2^{(2)} Y_2^{(3)}$

$\mathrm{mon}\Pi_4^{(2,3)} = Y_1^{(2)} Y_1^{(3)}, Y_2^{(1)}$

$\mathrm{mon}\Pi_5^{(2,3)} = Y_1^{(1)} Y_1^{(3)}, Y_2^{(2)}$

$\mathrm{mon}\Pi_6^{(2,3)} = Y_1^{(2)}, Y_2^{(1)} Y_2^{(3)}$

$\mathrm{mon}\Pi_1^{(3,3)} = Y_1^{(1)}, Y_2^{(2)}, Y_3^{(3)}$

$\mathrm{mon}\Pi_2^{(3,3)} = Y_1^{(3)}, Y_2^{(1)}, Y_3^{(2)}$

$\mathrm{mon}\Pi_3^{(3,3)} = Y_1^{(2)}, Y_2^{(3)}, Y_3^{(1)}$

Fourth Generation $\underline{Y}'_4 = (Y^{(1)}\ Y^{(2)}\ Y^{(3)}\ Y^{(4)})$

4.1.1 $Y_1^{(1)} Y_1^{(2)} Y_1^{(3)} \rightarrow$

$\mathrm{mon}\Pi_1^{(1,4)} = Y_1^{(1)} Y_1^{(2)} Y_1^{(3)} Y_1^{(4)}$

$\mathrm{mon}\Pi_1^{(2,4)} = Y_1^{(1)} Y_1^{(2)} Y_1^{(3)}, Y_2^{(4)}$

$\mathrm{mon}\Pi_2^{(2,4)} = Y_1^{(4)}, Y_2^{(1)} Y_2^{(2)} Y_2^{(3)}$

4.2.1 $Y_1^{(1)} Y_1^{(2)}, Y_2^{(3)} \rightarrow$

$\mathrm{mon}\Pi_3^{(2,4)} = Y_1^{(1)} Y_1^{(2)}, Y_2^{(3)} Y_2^{(4)}$

$\mathrm{mon}\Pi_4^{(2,4)} = Y_1^{(3)} Y_1^{(4)}, Y_2^{(1)} Y_2^{(2)}$

$\mathrm{mon}\Pi_5^{(2,4)} = Y_1^{(1)} Y_1^{(2)} Y_1^{(4)}, Y_2^{(3)}$

$\mathrm{mon}\Pi_6^{(2,4)} = Y_1^{(3)}, Y_2^{(1)} Y_2^{(2)} Y_2^{(4)}$

$\mathrm{mon}\Pi_1^{(3,4)} = Y_1^{(1)} Y_1^{(2)}, Y_2^{(3)}, Y_3^{(4)}$

$\mathrm{mon}\Pi_2^{(3,4)} = Y_1^{(4)}, Y_2^{(1)} Y_2^{(2)}, Y_3^{(3)}$

$\mathrm{mon}\Pi_3^{(3,4)} = Y_1^{(3)}, Y_2^{(4)}, Y_3^{(1)} Y_3^{(2)}$

4.2.2 $Y_1^{(1)}, Y_2^{(2)} Y_2^{(3)} \rightarrow$

$\mathrm{mon}\Pi_7^{(2,4)} = Y_1^{(1)}, Y_2^{(2)} Y_2^{(3)} Y_2^{(4)}$
$\mathrm{mon}\Pi_8^{(2,4)} = Y_1^{(2)} Y_1^{(3)} Y_1^{(4)}, Y_2^{(1)}$
$\mathrm{mon}\Pi_9^{(2,4)} = Y_1^{(1)} Y_1^{(4)}, Y_2^{(2)} Y_2^{(3)}$
$\mathrm{mon}\Pi_{10}^{(2,4)} = Y_1^{(2)} Y_1^{(3)}, Y_2^{(1)} Y_2^{(4)}$
$\mathrm{mon}\Pi_4^{(3,4)} = Y_1^{(1)}, Y_2^{(2)} Y_2^{(3)}, Y_3^{(4)}$
$\mathrm{mon}\Pi_5^{(3,4)} = Y_1^{(4)}, Y_2^{(1)}, Y_3^{(2)} Y_3^{(3)}$
$\mathrm{mon}\Pi_6^{(3,4)} = Y_1^{(2)} Y_1^{(3)}, Y_2^{(4)}, Y_3^{(1)}$

4.2.3 $Y_1^{(1)} Y_1^{(3)}, Y_2^{(2)} \rightarrow$

$\mathrm{mon}\Pi_{11}^{(2,4)} = Y_1^{(1)} Y_1^{(3)}, Y_2^{(2)} Y_2^{(4)}$
$\mathrm{mon}\Pi_{12}^{(2,4)} = Y_1^{(2)} Y_1^{(4)}, Y_2^{(1)} Y_2^{(3)}$
$\mathrm{mon}\Pi_{13}^{(2,4)} = Y_1^{(1)} Y_1^{(3)} Y_1^{(4)}, Y_2^{(2)}$
$\mathrm{mon}\Pi_{14}^{(2,4)} = Y_1^{(2)}, Y_2^{(1)} Y_2^{(3)} Y_2^{(4)}$
$\mathrm{mon}\Pi_7^{(3,4)} = Y_1^{(1)} Y_1^{(3)}, Y_2^{(2)}, Y_3^{(4)}$
$\mathrm{mon}\Pi_8^{(3,4)} = Y_1^{(4)}, Y_2^{(1)} Y_2^{(3)}, Y_3^{(2)}$
$\mathrm{mon}\Pi_9^{(3,4)} = Y_1^{(2)}, Y_2^{(4)}, Y_3^{(1)} Y_3^{(3)}$

4.3.1 $Y_1^{(1)}, Y_2^{(2)}, Y_3^{(3)} \rightarrow$

$\mathrm{mon}\Pi_{10}^{(3,4)} = Y_1^{(1)}, Y_2^{(2)}, Y_3^{(3)} Y_3^{(4)}$
$\mathrm{mon}\Pi_{11}^{(3,4)} = Y_1^{(3)} Y_1^{(4)}, Y_2^{(1)}, Y_3^{(2)}$
$\mathrm{mon}\Pi_{12}^{(3,4)} = Y_1^{(2)}, Y_2^{(3)} Y_2^{(4)}, Y_3^{(1)}$
$\mathrm{mon}\Pi_{13}^{(3,4)} = Y_1^{(1)}, Y_2^{(2)} Y_2^{(4)}, Y_3^{(3)}$
$\mathrm{mon}\Pi_{14}^{(3,4)} = Y_1^{(3)}, Y_2^{(1)}, Y_3^{(2)} Y_3^{(4)}$
$\mathrm{mon}\Pi_{15}^{(3,4)} = Y_1^{(2)} Y_1^{(4)}, Y_2^{(3)}, Y_3^{(1)}$
$\mathrm{mon}\Pi_{16}^{(3,4)} = Y_1^{(1)} Y_1^{(4)}, Y_2^{(2)}, Y_3^{(3)}$
$\mathrm{mon}\Pi_{17}^{(3,4)} = Y_1^{(3)}, Y_2^{(1)} Y_2^{(4)}, Y_3^{(2)}$
$\mathrm{mon}\Pi_{18}^{(3,4)} = Y_1^{(2)}, Y_2^{(3)}, Y_3^{(1)} Y_3^{(4)}$
$\mathrm{mon}\Pi_1^{(4,4)} = Y_1^{(1)}, Y_2^{(2)}, Y_3^{(3)}, Y_4^{(4)}$
$\mathrm{mon}\Pi_2^{(4,4)} = Y_1^{(4)}, Y_2^{(1)}, Y_3^{(2)}, Y_4^{(3)}$
$\mathrm{mon}\Pi_3^{(4,4)} = Y_1^{(3)}, Y_2^{(4)}, Y_3^{(1)}, Y_4^{(2)}$
$\mathrm{mon}\Pi_4^{(4,4)} = Y_1^{(2)}, Y_2^{(3)}, Y_3^{(4)}, Y_4^{(1)}$

There are B_{4_p} monomials for this generation ($B_{4_p} = 37$).

n-th Generation $\underline{Y}'_n = (Y^{(1)}\ Y^{(2)}\ \ldots\ Y^{(n)})$

n.1.1 $\quad Y_1^{(1)}Y_1^{(2)}\dots Y_1^{(n-1)}$

$$\mathrm{mon}\Pi_1^{(1,n)} = Y_1^{(1)}Y_1^{(2)}\dots Y_1^{(n-1)}Y_1^{(n)}$$
$$\mathrm{mon}\Pi_1^{(2,n)} = Y_1^{(1)}Y_1^{(2)}\dots Y_1^{(n-1)}, Y_2^{(n)}$$
$$\mathrm{mon}\Pi_2^{(2,n)} = Y_1^{(n)}, Y_2^{(1)}Y_2^{(2)}\dots Y_2^{(n-1)}$$

n.2.1 $\quad Y_1^{(1)}Y_1^{(2)}\dots Y_1^{(n-2)}, Y_2^{(n-1)}$

$$\mathrm{mon}\Pi_3^{(2,n)} = Y_1^{(1)}Y_1^{(2)}\dots Y_1^{(n-2)}, Y_2^{(2)}Y_2^{(n)}$$
$$\mathrm{mon}\Pi_4^{(2,n)} = Y_1^{(n-1)}Y_1^{(n)}, Y_2^{(1)}Y_2^{(2)}\dots Y_2^{(n-2)}$$
$$\mathrm{mon}\Pi_5^{(2,n)} = Y_1^{(1)}Y_1^{(2)}\dots Y_1^{(n-2)}Y_1^{(n)}, Y_2^{(n-1)}$$
$$\mathrm{mon}\Pi_6^{(2,n)} = Y_1^{(n-1)}, Y_2^{(1)}Y_2^{(2)}\dots Y_2^{(n-2)}Y_2^{(n)}$$
$$\mathrm{mon}\Pi_1^{(3,n)} = Y_1^{(1)}Y_1^{(2)}\dots Y_1^{(n-2)}, Y_2^{(n-1)}, Y_3^{(n)}$$
$$\mathrm{mon}\Pi_2^{(3,n)} = Y_1^{(n)}, Y_2^{(1)}Y_2^{(2)}\dots Y_2^{(n-2)}, Y_3^{(n-1)}$$
$$\mathrm{mon}\Pi_3^{(3,n)} = Y_1^{(n-1)}, Y_2^{(n)}, Y_3^{(1)}Y_3^{(2)}\dots Y_3^{(n-2)}$$

n.2.2 $\quad Y_1^{(1)}, Y_2^{(2)}Y_2^{(3)}\dots Y_2^{(n-1)}$

$$\mathrm{mon}\Pi_7^{(2,n)} = Y_1^{(1)}, Y_2^{(2)}Y_2^{(3)}\dots Y_2^{(n-1)}Y_2^{(n)}$$
$$\mathrm{mon}\Pi_8^{(2,n)} = Y_1^{(2)}Y_1^{(3)}\dots Y_1^{(n-1)}Y_1^{(n)}, Y_2^{(1)}$$
$$\mathrm{mon}\Pi_9^{(2,n)} = Y_1^{(1)}Y_1^{(n)}, Y_2^{(2)}Y_2^{(3)}\dots Y_2^{(n-1)}$$
$$\mathrm{mon}\Pi_{10}^{(2,n)} = Y_1^{(2)}Y_1^{(3)}\dots Y_1^{(n-1)}, Y_2^{(1)}Y_2^{(n)}$$
$$\mathrm{mon}\Pi_4^{(3,n)} = Y_1^{(1)}, Y_2^{(2)}Y_2^{(3)}\dots Y_2^{(n-1)}, Y_3^{(n)}$$
$$\mathrm{mon}\Pi_5^{(3,n)} = Y_1^{(n)}, Y_2^{(1)}, Y_3^{(2)}Y_3^{(3)}\dots Y_3^{(n-1)}$$
$$\mathrm{mon}\Pi_6^{(3,n)} = Y_1^{(2)}Y_1^{(3)}\dots Y_1^{(n-1)}, Y_2^{(n)}, Y_3^{(1)}$$

$\vdots \qquad\qquad \vdots \quad \vdots \quad \vdots$

n.2.$S(n-1,2)$

$Y_1^{(1)}\dots Y_1^{(n-3)}, Y_2^{(n-2)}Y_2^{(n-1)}$

$$\mathrm{mon}\Pi_{2S(n,2)-3}^{(2,n)} = Y_1^{(1)}\dots Y_1^{(n-3)}, Y_2^{(n-2)}Y_2^{(n-1)}Y_2^{(n)}$$
$$\mathrm{mon}\Pi_{2S(n,2)-2}^{(2,n)} = Y_1^{(n-2)}Y_1^{(n-1)}Y_1^{(n)}, Y_2^{(1)}\dots Y_2^{(n-3)}$$
$$\mathrm{mon}\Pi_{2S(n,2)-1}^{(2,n)} = Y_1^{(1)}\dots Y_1^{(n-3)}Y_1^{(n)}, Y_2^{(n-2)}Y_2^{(n-1)}$$
$$\mathrm{mon}\Pi_{2S(n,2)}^{(2,n)} = Y_1^{(n-2)}Y_1^{(n-1)}, Y_2^{(1)}\dots Y_2^{(n-3)}Y_2^{(n)}$$
$$\mathrm{mon}\Pi_{3S(n-1,2)-2}^{(3,n)} = Y_1^{(1)}\dots Y_1^{(n-3)}, Y_2^{(n-2)}Y_2^{(n-1)}, Y_3^{(n)}$$
$$\mathrm{mon}\Pi_{3S(n-1,2)-1}^{(3,n)} = Y_1^{(n)}, Y_2^{(1)}\dots Y_2^{(n-3)}, Y_3^{(n-2)}Y_3^{(n-1)}$$
$$\mathrm{mon}\Pi_{3S(n-1,2)}^{(3,n)} = Y_1^{(n-2)}Y_1^{(n-1)}, Y_2^{(n)}, Y_3^{(1)}\dots Y_3^{(n-3)}$$

$\vdots \qquad\qquad \vdots \quad \vdots \quad \vdots$

$n.(n-1).1$

$$Y_1^{(1)}, Y_2^{(2)}, Y_3^{(3)}, \ldots, Y_{n-1}^{(n-1)} \longrightarrow$$

$$\mathrm{mon}\Pi_{c_{1,1}}^{(n-1,n)} = Y_1^{(1)}, Y_2^{(2)}, Y_3^{(3)}, \ldots, Y_{n-1}^{(n-1)} Y_{n-1}^{(n)}$$
$$\mathrm{mon}\Pi_{c_{1,2}}^{(n-1,n)} = Y_1^{(n-1)} Y_1^{(n)}, Y_2^{(1)}, Y_3^{(2)}, \ldots, Y_{n-1}^{(n-2)}$$
$$\vdots \qquad \vdots \qquad \vdots \qquad \vdots$$
$$\mathrm{mon}\Pi_{c_{1,n-1}}^{(n-1,n)} = Y_1^{(2)}, Y_2^{(3)}, \ldots, Y_{n-2}^{(n-1)} Y_{n-2}^{(n)}, Y_{n-1}^{(1)}$$
$$\mathrm{mon}\Pi_{c_{2,1}}^{(n-1,n)} = Y_1^{(1)}, Y_2^{(2)}, Y_3^{(3)}, \ldots, Y_{n-2}^{(n-2)} Y_{n-2}^{(n)}, Y_{n-1}^{(n-1)}$$
$$\mathrm{mon}\Pi_{c_{2,2}}^{(n-1,n)} = Y_1^{(n-1)}, Y_2^{(1)}, Y_3^{(2)}, \ldots, Y_{n-2}^{(n-3)}, Y_{n-1}^{(n-2)} Y_{n-1}^{(n)}$$
$$\vdots \qquad \vdots \qquad \vdots \qquad \vdots$$
$$\mathrm{mon}\Pi_{c_{2,n-1}}^{(n-1,n)} = Y_1^{(2)}, Y_2^{(3)}, \ldots, Y_{n-2}^{(n-2)} Y_{n-2}^{(n)}, Y_{n-1}^{(1)}$$
$$\cdots \quad . \quad .$$
$$\cdots \quad . \quad .$$
$$\cdots \quad . \quad .$$
$$\mathrm{mon}\Pi_{c_{n-1,1}}^{(n-1,n)} = Y_1^{(1)} Y_1^{(n)}, Y_2^{(2)}, Y_3^{(3)}, \ldots, Y_{n-2}^{(n-2)}, Y_{n-1}^{(n-1)}$$
$$\mathrm{mon}\Pi_{c_{n-1,2}}^{(n-1,n)} = Y_1^{(n-1)}, Y_2^{(1)} Y_2^{(n)}, \ldots, Y_{n-2}^{(n-3)}, Y_{n-1}^{(n-2)}$$
$$\vdots \qquad \vdots \qquad \vdots \qquad \vdots$$
$$\mathrm{mon}\Pi_{c_{n-1,n-1}}^{(n-1,n)} = Y_1^{(2)}, Y_2^{(3)}, \ldots, Y_{n-2}^{(n-1)}, Y_{n-1}^{(1)} Y_{n-1}^{(n)}$$
$$\mathrm{mon}\Pi_{c_{n,1}}^{(n,n)} = Y_1^{(1)}, Y_2^{(2)}, Y_3^{(3)}, \ldots, Y_{n-1}^{(n-1)}, Y_n^{(n)}$$
$$\mathrm{mon}\Pi_{c_{n,2}}^{(n,n)} = Y_1^{(n)}, Y_2^{(1)}, Y_3^{(2)}, \ldots, Y_n^{(n-1)}$$
$$\vdots \qquad \vdots \qquad \vdots \qquad \vdots$$
$$\mathrm{mon}\Pi_{c_{n,n}}^{(n,n)} = Y_1^{(n-1)}, Y_2^{(n)}, Y_3^{(1)}, Y_4^{(2)}, \ldots, Y_n^{(n-2)}$$

where the lower indices $c_{i,j}$ in $\mathrm{mon}\Pi_{c_{i,j}}^{(n-1,n)}$ are as follows:

$c_{1,1} = (n-1)S(n,n-1) - n(n-2)$

$c_{1,2} = (n-1)S(n,n-1) - n(n-2) + 1$

$\vdots$

$c_{1,n-1} = (n-1)S(n,n-1) - (n-1)(n-2)$

$c_{2,1} = (n-1)S(n,n-1) - (n-1)(n-2) + 1$

$c_{2,2} = (n-1)S(n,n-1) - (n-1)(n-2) + 2$

$\vdots$

$c_{2,n-1} = (n-1)S(n,n-1) - (n-1)(n-3)$

$\vdots$

$c_{n-1,1} = (n-1)S(n,n-1) - (n-2)$
$c_{n-1,2} = (n-1)S(n,n-1) - (n-2) + 1$
$\vdots$
$c_{n-1,n-1} = (n-1)S(n,n-1)$
$c_{n,1} = 1, c_{n,2} = 2, \ldots, c_{n,n} = n$

Concerning this recursive scheme, we note the following remarks:

- For the third generation, it is not necessary to generate monomials from $Y_1^{(2)}, Y_2^{(1)} \curvearrowright$, because these monomials are already generated in 3.2.1.
- For the fourth generation, we do not generate monomials from $Y_1^{(3)}, Y_2^{(1)}Y_2^{(2)} \curvearrowright$, $Y_1^{(2)}Y_1^{(3)}, Y_2^{(1)} \curvearrowright$, $Y_1^{(2)}, Y_2^{(1)}Y_2^{(3)} \curvearrowright$, $Y_1^{(3)}, Y_2^{(1)}, Y_3^{(2)} \curvearrowright$ and $Y_1^{(2)}, Y_2^{(3)}, Y_3^{(1)} \curvearrowright$, because these monomials are already generated in the previous steps.
- According to the analysis elaborated here, we do not generate monomials from those with the rounded arrow.
- The application of Theorem 3 to this recursive scheme is illustrated by the following examples:
  $E[Y_1^{(1)}Y_1^{(2)}Y_1^{(3)}, Y_2^{(4)}] \xrightarrow{th.3} E[Y_1^{(1)}Y_1^{(2)}Y_1^{(3)}\phi_2(t_1^*)I(t_1 \le t)]E[Y_2^{(4)}]$
  $E[Y_1^{(1)}, Y_2^{(2)}Y_2^{(3)}, Y_3^{(4)}] \xrightarrow{th.3} E[Y_1^{(1)}Y_2^{(2)}Y_2^{(3)}\phi_3(t_2^*)I(t_1 + t_2 \le t)]E[Y_3^{(4)}]$
  $E[Y_1^{(4)}, Y_2^{(1)}, Y_3^{(2)}, Y_4^{(3)}] \xrightarrow{th.3} E[Y_1^{(4)}Y_2^{(1)}Y_3^{(2)}\phi_4(t_3^*)I(t_1 + t_2 + t_3 \le t)]E[Y_4^{(3)}]$
- For each generation, there are B_{n_p} monomials, where $B_{n_p} = \sum_{m=1}^{n} mS(n,m)$.

8. Conclusion

As we observe, the computation of $E[P_n^*(t)]$ using the present approach, is based only on $\phi_m(t_m^*)$, since the monomials of $P_n^*(t)$ are generated recursively. Thus using a symbolic language of programming as Maple, the symbolic computation of $E[P_n^*(t)]$ can be performed efficiently.

References

[1] Adès, M. and Malhamé, R.P. (1994) Asymptotics of the Moments of Cumulative Vector Renewal Reward Processes: The Case $N(t)$. *Les Cahiers du GERAD* G–94–32. Revised April 1997.

[2] Adès, M. and Malhamé, R.P. (1994) On the Moments of Cumulative Processes: A Preliminary Study. *Les Cahiers du GERAD* G-94-13. École des Hautes Études Commerciales, Montréal.

[3] Feller, W. (1971) *An Introduction to Probability Theory and Its Applications*. Volume II, Second Edition, John Wiley & Sons, New York.

[4] KARLIN, S. AND TAYLOR, H.M. (1975) *A First Course in Stochastic Processes*. Second Edition, Academic Press, New York.

[5] MURTHY, V.K. (1974) *The General Point Processes*. Addison-Wesley, Massachusetts.

[6] ROSS, S.M. (1983) *Stochastic Processes*. John-Wiley & Sons, New York.

[7] SHIRYAYEV, A.N. (1984) *Probability*. Springer-Verlag, New York.

[8] SMITH, W.L. (1955) Regenerative Stochastic Process. *Proceedings of the Royal Society* A **232**, 6–31.

[9] SMITH, W.L. (1959) On the Cumulants of Renewal Process. *Biometrika* **46**, 1–29.

CHAPTER 4

On the Moments of Cumulative Processes: A Preliminary Study

1. Introduction: Classical Definitions and Notations

Let $\{t_i\}, i = 1, 2, \ldots$, be an infinite sequence of independent non-negative and identically distributed random variables, which are not zero with probability one. Such a sequence of random variables is called a renewal process.

The augmented sequence $\{t_i\}, i = 0, 1, 2, \ldots$, is called a general renewal process, where t_0 is a non-negative random variable, independent of the t_i's$, i = 1, 2, \ldots$, and not necessarily identically distributed like them.

Let $T_n = \sum_{i=1}^{n} t_i$ for the partial sums of t_i, where T_n is the time instant corresponding to the occurrence of the event E. Let $K(t)$ and $F(t)$ be the distribution functions of t_0 and $\{t_i\} i \geq 1$, respectively. For a given general renewal process $\{t_i\}$, we write $T_{-1} = 0$, $T_n = \sum_{i=0}^{n} t_i$ for $n = 0, 1, \ldots$, and we define for all $t \geq 0$, the random variable $N^*(t)$ as the greatest integer n such that $T_{n-1} \leq t$.

Let W_t be a real valued process which is defined to be a cumulative process if it satisfies the following two conditions:

(C1) $y_n = \Delta_n W_t = W_{T_n} - W_{T_{n-1}}$ is a sequence of independent and identically distributed random variables, where $n = 1, 2, \ldots$

(C2) W_t is, with probability one, of bounded variation in every finite t-interval (see Smith [15]).

Let us consider from here on the case where $t_0 = 0$, thus let $N(t)$ be a random variable which represents the number of events in the interval $(0, t]$ of the renewal process $\{t_i\}, i = 1, 2, \ldots$.

We assume that $\{t_n, y_n\}, n = 1, 2, \ldots$, is a sequence of independent and identically distributed random variables.

Let $G(t, y)$ denote the joint distribution function of $\{t_n, y_n\}$, and let $\kappa_r = E[y_n^r]$, $\mu_r = E[t_n^r]$ and $\mu_{ij} = E[t_n^i y_n^j]$, where $\kappa_r + \mu_r + \mu_{ij} < \infty$.

Let $G_s^*(\alpha) = \int\limits_{t=0}^{\infty} \int\limits_{y=-\infty}^{\infty} e^{-st+i\alpha y} d_{t,y} G(t,y)$,

then

$$G_0^*(\alpha) = \int\limits_0^{\infty} \int\limits_{-\infty}^{\infty} e^{i\alpha y} d_{t,y} G(t,y) = \int\limits_{-\infty}^{\infty} e^{i\alpha y} d_y G(y) \equiv \int\limits_{-\infty}^{\infty} e^{i\alpha y} dG(y)$$

and

$$G_s^*(0) = \int\limits_0^{\infty} \int\limits_{-\infty}^{\infty} e^{-st} d_{t,y} G(t,y) = \int\limits_0^{\infty} e^{-st} d_t G(t) \equiv \int\limits_0^{\infty} e^{-st} dG(t) = \int\limits_0^{\infty} e^{-st} dF(t)$$

where $G_0^*(\alpha)$ is the characteristic function of the random variable y_n and $G_s^*(0) = F^*(s) = \int\limits_0^{\infty} e^{-st} dF(t)$, is the Laplace-Stieltjes transform of $F(t)$. Recall that $F(t)$ is the distribution function of t_n.

This chapter is organized as follows. An explicit and asymptotic expression of the third moment of a cumulative processes is presented in section 2. A detailed procedure is given in section 3 to find the expectation of a product of triple summation. Finally, we study briefly $E[N(t)]^k$ in section 4.

2. Some Excerpts of Smith's Work on Renewal Theory and Further Development

One of the pioneers in the area of renewal theory is W.L. Smith.

In his paper of 1955, Smith [15] has considered the random variable Y_t with the related questions as follows:

a) $Y_t = \sum\limits_{i=1}^{N(t)+1} y_i = W_{t_1+t_2+\cdots+t_{N(t)+1}}$.

b) $\Psi_t(y) = P(Y_t \leq y)$ is the distribution function of Y_t.

c) For $Re(s) \geq 0$, let $\Psi_s^*(y) = s \int\limits_0^{\infty} e^{-st} \Psi_t(y) dt$, where for fixed s, $\Psi_s^*(y)$ is a distribution function in y.

d) When $Re(s) > 0$, $\psi_s^*(\alpha) = \int\limits_{-\infty}^{\infty} e^{i\alpha y} d_y \Psi_s^*(y) =$

$$\frac{G_0^*(\alpha) - G_s^*(\alpha)}{1 - G_s^*(\alpha)} = \frac{G_0^*(\alpha) - 1}{1 - G_s^*(\alpha)} + 1$$

where $\psi_s^*(\alpha)$ is the characteristic function of $\Psi_s^*(y)$ and $\alpha \in \mathbb{R}$. When $Re(s) > 0$, $\frac{G_0^*(\alpha)-1}{1-G_s^*(\alpha)}$ is $m-$times differentiable with respect to α, since $\kappa_m = E[y_n^m] < \infty$; recall for this purpose that $G_0^*(\alpha)$ is the characteristic function of the random variable y_n. Consequently, the first m moments of $\Psi_s^*(y)$ are finite (Theorem 3.2.1 page 142 of Laha & Rohatgi [12]).

2.1. Evaluation of the Third Derivative of $\psi_s^*(\alpha)$. We can show very easily that

$$\psi_s^{*\prime}(\alpha)=\frac{G_0^{*\prime}(\alpha)}{1-G_s^*(\alpha)}+\frac{G_s^{*\prime}(\alpha)-\left(G_0^*(\alpha)-1\right)}{\left(1-G_s^*(\alpha)\right)^2} \tag{2.1}$$

and

$$\begin{aligned}\psi_s^{*\prime\prime}(\alpha) &= \frac{G_0^{*\prime\prime}(\alpha)}{1-G_s^*(\alpha)}+\frac{2G_0^{*\prime}(\alpha)G_s^{*\prime}(\alpha)}{\left(1-G_s^*(\alpha)\right)^2}-\frac{G_s^{*\prime\prime}(\alpha)}{\left(1-G_s^*(\alpha)\right)^2}\\ &\quad-\frac{2\left(G_s^{*\prime}(\alpha)\right)^2}{\left(1-G_s^*(\alpha)\right)^3}+\frac{G_0^*(\alpha)G_s^{*\prime\prime}(\alpha)}{\left(1-G_s^*(\alpha)\right)^2}+\frac{2G_0^*(\alpha)\left(G_s^{*\prime}(\alpha)\right)^2}{\left(1-G_s^*(\alpha)\right)^3}\end{aligned} \tag{2.2}$$

Then after some algebra, one obtains:

$$\begin{aligned}\psi_s^{*\prime\prime\prime}(\alpha) &= \frac{G_0^{*\prime\prime\prime}(\alpha)}{\left(1-G_s^*(\alpha)\right)}+\frac{G_0^{*\prime\prime}(\alpha)G_s^{*\prime}(\alpha)}{\left(1-G_s^*(\alpha)\right)^2}\\ &\quad+\frac{2G_0^{*\prime\prime}(\alpha)G_s^{*\prime}(\alpha)+2G_0^{*\prime}(\alpha)G_s^{*\prime\prime}(\alpha)}{\left(1-G_s^*(\alpha)\right)^2}+\frac{4G_0^{*\prime}(\alpha)\left(G_s^{*\prime}(\alpha)\right)^2}{\left(1-G_s^*(\alpha)\right)^3}\\ &\quad-\frac{G_s^{*\prime\prime\prime}(\alpha)}{\left(1-G_s^*(\alpha)\right)^2}-\frac{2G_s^{*\prime}(\alpha)G_s^{*\prime\prime}(\alpha)}{\left(1-G_s^*(\alpha)\right)^3}\\ &\quad-\frac{4G_s^{*\prime}(\alpha)G_s^{*\prime\prime}(\alpha)}{\left(1-G_s^*(\alpha)\right)^3}-\frac{6\left(G_s^{*\prime}(\alpha)\right)^3}{\left(1-G_s^*(\alpha)\right)^4}\\ &\quad+\frac{G_0^{*\prime}(\alpha)G_s^{*\prime\prime}(\alpha)+G_0^*(\alpha)G_s^{*\prime\prime\prime}(\alpha)}{\left(1-G_s^*(\alpha)\right)^2}+\frac{2G_0^*(\alpha)G_s^{*\prime}(\alpha)G_s^{*\prime\prime}(\alpha)}{\left(1-G_s^*(\alpha)\right)^3}\\ &\quad+\frac{2G_0^{*\prime}(\alpha)\left(G_s^{*\prime}(\alpha)\right)^2+4G_0^*(\alpha)G_s^{*\prime}(\alpha)G_s^{*\prime\prime}(\alpha)}{\left(1-G_s^*(\alpha)\right)^3}\\ &\quad+\frac{6G_0^*(\alpha)\left(G_s^{*\prime}(\alpha)\right)^3}{\left(1-G_s^*(\alpha)\right)^4}\end{aligned} \tag{2.3}$$

Since $\psi_s^{*'''}(0) = \int_{-\infty}^{\infty} i^3 y^3 d_y \Psi_s^*(y)$, $G_0^*(0) = 1$, $G_0^{*'}(0) = i\kappa_1$, $G_0^{*''}(0) = i^2\kappa_2$, $G_0^{*'''}(0) = i^3\kappa_3$ and $G_s^*(0) = F^*(s)$, then we obtain:

$$\begin{aligned}\int_{-\infty}^{\infty} y^3 d_y \Psi_s^*(y) &= \frac{\kappa_3}{\left(1-F^*(s)\right)} + \frac{3\kappa_2 R_1^*(s)}{\left(1-F^*(s)\right)^2} \\ &\quad + \frac{3\kappa_1 R_2^*(s)}{\left(1-F^*(s)\right)^2} + \frac{6\kappa_1 R_1^{*2}(s)}{\left(1-F^*(s)\right)^3} \end{aligned} \tag{2.4}$$

where:

$$R_1^*(s) \triangleq \left[\frac{1}{i}\frac{\partial}{\partial\alpha} G_s^*(\alpha)\right]_{\alpha=0} \tag{2.5}$$

$$R_2^*(s) \triangleq \left[\frac{1}{i^2}\frac{\partial^2}{\partial\alpha^2} G_s^*(\alpha)\right]_{\alpha=0} \tag{2.6}$$

Obviously $\psi_s^{*'''}(0)$ gives the third moment of $\Psi_s^*(y)$. But as $\frac{\Psi_s^*(y)}{s}$ is the Laplace transform of $\Psi_t(y)$, then using the inverse of such transform one can obtain $\int_{-\infty}^{\infty} y^3 d_y \Psi_t(y) = E[Y_t]^3$.

2.2. Explicit Form for $R_1^*(s)$ and $R_2^*(s)$.

a) The function $R_1^*(s)$ is the Laplace Stieltjes transform of $R_1(t)$ which is of bounded variation in t (Smith [15]), where:

$$R_1(t) = \int_{-\infty}^{\infty} y d_y G(t, y) \tag{2.7}$$

$$R_1^*(s) = \int_0^{\infty} e^{-st} dR_1(t) \tag{2.8}$$

It can easily be shown that

$$\int_0^{\infty} dR_1(t) = E[y_n] = \kappa_1 \tag{2.9}$$

$$\int_0^{\infty} t dR_1(t) = E[t_n y_n] = \mu_{11} \tag{2.10}$$

where the integrals converge absolutely. We can expand e^{-st} by the Taylor-MacLaurin series which is given by:

$$e^{-st} = \sum_{i=0}^{\infty}(-1)^i\frac{(st)^i}{i!} \tag{2.11}$$

Then

$$R_1^*(s) = \int_0^{\infty}\Big(\sum_{i=0}^{\infty}(-1)^i\frac{(st)^i}{i!}\Big)dR_1(t) \tag{2.12}$$

$$= \sum_{i=0}^{\infty}\frac{(-s)^i}{i!}\Big(\int_0^{\infty}t^i dR_1(t)\Big) \tag{2.13}$$

$$= \sum_{i=0}^{\infty}\frac{(-s)^i}{i!}\mu_{i1} \tag{2.14}$$

The integral in (2.8) converges absolutely at a point $s = \sigma + i\tau$ if the integral

$$\int_0^{\infty} e^{-\sigma t}|dR_1(t)| = \int_0^{\infty} e^{-\sigma t}d\alpha(t)$$

converges, where $\alpha(x)$ is the total variation of the function $R_1(t)$ in the interval $0 \le t \le x$. Now as $\int_0^{\infty} e^{-st}dR_1(t)$ converges absolutely for $s = \sigma_0 + i\tau_0$, then it converges uniformly and absolutely in the half-plane $\sigma \ge \sigma_0$ (Theorem 3.1 page 46 of Widder [18]).

Consequently equation (2.13) relies on the fact that the integral in (2.8) converges uniformly, while equation (2.14) follows because $\mu_{i1} = E[t_n^i y_n] = \int_0^{\infty} t_n^i dR_1(t) = \int_0^{\infty} t^i dR_1(t)$.

Note also that $\mu_{01} = E[y_n] = \kappa_1$. To conclude this subsection, let us prove equation (2.5) as follows:

$$G_s^{*'}(\alpha) = \int_{t=0}^{\infty}\int_{y=-\infty}^{\infty} ie^{-st}e^{i\alpha y}y d_t d_y G(t,y)$$

$$G_s^{*'}(0) = \int_{t=0}^{\infty}\int_{-\infty}^{\infty} ie^{-st}y d_t d_y G(t,y)$$

$$G_s^{*'}(0) = i\int_{t=0}^{\infty} e^{-st}d_t\Big(\int_{-\infty}^{\infty} y d_y G(t,y)\Big)$$

$$G_s^{*'}(0) = i\int_{t=0}^{\infty} e^{-st}d_t R_1(t)$$

$$G_s^{*'}(0) = i\int_{t=0}^{\infty} e^{-st}dR_1(t)$$

b) The function $R_2^*(s)$ is the Laplace-Stieltjes transform of $R_2(t)$ which is of bounded variation in t, where:

$$R_2(t) = \int_{-\infty}^{\infty} y^2 d_y G(t,y) \quad (2.15)$$

$$R_2^*(s) = \int_0^{\infty} e^{-st}dR_2(t) \quad (2.16)$$

It can be shown that:

$$\int_0^{\infty} dR_2(t) = E[y_n^2] = \kappa_2 \quad (2.17)$$

$$\int_0^{\infty} t^2 dR_2(t) = E[t_n^2 y_n] = \mu_{22} \quad (2.18)$$

where the integrals converge absolutely.

Using the same development as previously (part 2.2.a), we can write

$$R_2^*(s) = \int_0^\infty \Big(\sum_{i=0}^\infty \frac{(-st)^i}{i!}\Big) dR_2(t) \tag{2.19}$$

$$= \sum_{i=0}^\infty \frac{(-s)^i}{i!} \Big(\int_0^\infty t^i dR_2(t)\Big)$$

$$= \sum_{i=0}^\infty \frac{(-s)^i}{i!} \mu_{i2} \tag{2.20}$$

where $\mu_{i2} = E[t_n^i y_n^2] = \int_0^\infty t_n^i dR_2(t) = \int_0^\infty t^i dR_2(t)$.

To end this subsection note that equation (2.6) can be proved very easily; indeed:

$$G_s^{*''}(\alpha) = \int_{t=0}^\infty \int_{y=-\infty}^\infty i^2 y^2 e^{-st} e^{i\alpha y} d_t d_y G(t,y)$$

$$G_s^{*''}(0) = \int_0^\infty \int_{-\infty}^\infty i^2 y^2 e^{-st} d_t d_y G(t,y)$$

$$= i^2 \int_0^\infty e^{-st} d_t \Big(\int_{-\infty}^\infty y^2 d_y G(t,y)\Big)$$

$$= i^2 \int_0^\infty e^{-st} dR_2(t)$$

2.3. Explicit Form for $\int_{-\infty}^\infty y^3 d_y \Psi_s^*(y)$. To obtain an explicit expression for $\int_{-\infty}^\infty y^3 d_y \Psi_s^*(y)$ given in equation (2.4), the knowledge of $F^*(s)$ is required. For this end we use Lemma 3 of Smith [17], where he gave an expansion for $F^*(s)$. This Lemma is stated as follows:

"A necessary and sufficient condition for $F(t)$ to have its first n moments finite is that, for real $s > 0$, $F^*(s)$ has an expansion
$F^*(s) = 1 - \mu_1 s + \frac{\mu_2 s^2}{2!} - \cdots + \frac{(-s)^n \mu_n}{n!} + o(s^n)$".

Using, on the one hand, the inversion techniques for a power series, see e.g. pages 506–515 of Knuth [10] Vol. 2, pages 436–444 of Markushevich [13] Vol. 1, Brent and Kung [5], Kung [11] and Beyer [4]. On the other hand, we consider $F^*(s)$ as stated above, $R_1^*(s)$, $R_2^*(s)$ as given by equations (2.14) and (2.20) respectively. We obtain, by applying a straightforward computation and after some steps of simplification, the following:

- $$\frac{1}{1-F^*(s)} = \frac{1}{s\mu_1} + \frac{\mu_2}{2\mu_1^2} + o(1) \quad \text{as } s \to 0 \tag{2.21}$$

- $$\begin{aligned} \frac{R_1^*(s)}{\left(1-F^*(s)\right)^2} &= \frac{\kappa_1}{s^2\mu_1^2} + \frac{1}{s}\left(\frac{\kappa_1\mu_2}{\mu_1^3} - \frac{\mu_{11}}{\mu_1^2}\right) \\ &+ \frac{\mu_{21}}{2\mu_1^2} - \frac{\mu_2\mu_{11}}{\mu_1^3} + \kappa_1\left(\frac{3\mu_2^2}{4\mu_1^4} - \frac{\mu_3}{3\mu_1^3}\right) \\ &+ o(1) \quad \text{as } s \to 0 \end{aligned} \tag{2.22}$$

- $$\begin{aligned} \frac{R_2^*(s)}{\left(1-F^*(s)\right)^2} &= \frac{\kappa_2}{s^2\mu_1^2} + \frac{1}{s}\left(\frac{\kappa_2\mu_2}{\mu_1^3} - \frac{\mu_{12}}{\mu_1^2}\right) \\ &+ \frac{\mu_{22}}{2\mu_1^2} - \frac{\mu_2\mu_{12}}{\mu_1^3} + \kappa_2\left(\frac{3\mu_2^2}{4\mu_1^4} - \frac{\mu_3}{3\mu_1^3}\right) \\ &+ o(1) \quad \text{as } s \to 0 \end{aligned} \tag{2.23}$$

- $$\begin{aligned} \frac{R_1^{*2}(s)}{\left(1-F*(s)\right)^3} &= \frac{1}{s^3}\frac{\kappa_1^2}{\mu_1^3} + \frac{1}{s^2}\left(\frac{3\kappa_1^2\mu_2}{2\mu_1^4} - \frac{2\kappa_1\mu_{11}}{\mu_1^3}\right) \\ &+ \frac{1}{s}\left(\frac{3\kappa_1^2\mu_2^2}{2\mu_1^5} - \frac{\kappa_1^2\mu_3}{2\mu_1^4} - \frac{3\kappa_1\mu_{11}\mu_2}{\mu_1^4} + \frac{\kappa_1\mu_{21}}{\mu_1^3} + \frac{\mu_{11}^2}{\mu_1^3}\right) \\ &+ \frac{1}{s^0}\left(\frac{\kappa_1^2\mu_2^3}{2\mu_1^6} - \frac{\kappa_1^2\mu_4}{8\mu_1^4} - \frac{3\kappa_1\mu_{11}\mu_2^2}{\mu_1^5} + \frac{\kappa_1\mu_{11}\mu_3}{\mu_1^4}\right. \\ &\quad \left. + \frac{3\kappa_1\mu_{21}\mu_2}{2\mu_1^4} + \frac{3\mu_{11}^2\mu_2}{2\mu_1^4} - \frac{\kappa_1\mu_{31}}{3\mu_1^3} - \frac{\mu_{11}\mu_{21}}{\mu_1^3}\right) \\ &+ o(1) \quad \text{as } s \to 0 \end{aligned} \tag{2.24}$$

Combining the last four equations we find:

$$\begin{aligned}
\int_{-\infty}^{\infty} y^3 d_y \Psi_s^*(y) &= \frac{3!}{s^3}\frac{\kappa_1^3}{\mu_1^3} + \frac{2}{s^2}\frac{3}{\mu_1^2}\left(\frac{3\kappa_1^3\mu_2}{2\mu_1^2} - \frac{2\kappa_1^2\mu_{11}}{\mu_1} + \kappa_1\kappa_2\right) \\
&+ \frac{1}{s}\frac{6}{\mu_1}\left(\frac{3\kappa_1^3\mu_2^2}{2\mu_1^4} - \frac{\kappa_1^3\mu_3}{2\mu_1^3} - \frac{3\kappa_1^2\mu_{11}\mu_2}{\mu_1^3} + \frac{\kappa_1^2\mu_{21}}{\mu_1^2}\right. \\
&\quad \left. + \frac{\kappa_1\mu_{11}^2}{\mu_1^2} + \frac{\kappa_3}{6} - \frac{\kappa_2\mu_{11}}{2\mu_1} + \frac{\kappa_1\kappa_2\mu_2}{\mu_1^2} - \frac{\kappa_1\mu_{12}}{2\mu_1}\right) \\
&+ \frac{1}{s^0}\left(\frac{\kappa_3\mu_2}{2\mu_1^2} + \frac{3\kappa_2\mu_{21}}{2\mu_1^2} - \frac{3\kappa_2\mu_2\mu_{11}}{\mu_1^3} + \frac{9\kappa_1\kappa_2\mu_2^2}{2\mu_1^4}\right. \\
&\quad - \frac{2\kappa_1\kappa_2\mu_3}{\mu_1^3} + \frac{3\kappa_1\mu_{22}}{2\mu_1^2} - \frac{3\kappa_1\mu_2\mu_{12}}{\mu_1^3} + \frac{3\kappa_1^3\mu_2^3}{\mu_1^6} \\
&\quad - \frac{3\kappa_1^3\mu_4}{4\mu_1^4} - \frac{18\kappa_1^2\mu_{11}\mu_2^2}{\mu_1^5} + \frac{6\kappa_1^2\mu_{11}\mu_3}{\mu_1^4} + \frac{9\kappa_1^2\mu_{21}\mu_2}{\mu_1^4} \\
&\quad \left. + \frac{9\kappa_1\mu_{11}^2\mu_2}{\mu_1^4} - \frac{2\kappa_1^2\mu_{31}}{\mu_1^3} - \frac{6\kappa_1\mu_{11}\mu_{21}}{\mu_1^3}\right) \\
&+ o(1) \quad \text{as } s \to 0
\end{aligned} \tag{2.25}$$

Evidently, equation (2.25) gives the third moment of $\Psi_s^*(y)$. But as mentioned previously $\frac{\Psi_s^*(y)}{s}$ is the Laplace transform of $\Psi_t(y)$.

Then using the inverse of such transform, we obtain:

$$\begin{aligned}
\int_{-\infty}^{\infty} y^3 d_y \Psi_t(y) &= E\left[\sum_{i=1}^{N(t)+1} y_i\right]^3 = \frac{\kappa_1^3}{\mu_1^3}t^3 + \frac{3}{\mu_1^2}\left(\frac{3\kappa_1^3\mu_2}{2\mu_1^2} - \frac{2\kappa_1^2\mu_{11}}{\mu_1} + \kappa_1\kappa_2\right)t^2 \\
&\quad + \frac{6}{\mu_1}\left(\frac{3\kappa_1^3\mu_2^2}{2\mu_1^4} - \frac{\kappa_1^3\mu_3}{2\mu_1^3} - \frac{3\kappa_1^2\mu_{11}\mu_2}{\mu_1^3} + \frac{\kappa_1^2\mu_{21}}{\mu_1^2}\right. \\
&\quad \left. + \frac{\kappa_1\mu_{11}^2}{\mu_1^2} + \frac{\kappa_3}{6} - \frac{\kappa_2\mu_{11}}{2\mu_1} + \frac{\kappa_1\kappa_2\mu_2}{\mu_1^2} - \frac{\kappa_1\mu_{12}}{2\mu_1}\right)t \\
&+ \left(\frac{\kappa_3\mu_2}{2\mu_1^2} + \frac{3\kappa_2\mu_{21}}{2\mu_1^2} - \frac{3\kappa_2\mu_2\mu_{11}}{\mu_1^3} + \frac{9\kappa_1\kappa_2\mu_2^2}{2\mu_1^4} - \frac{2\kappa_1\kappa_2\mu_3}{\mu_1^3}\right. \\
&\quad + \frac{\kappa_1\mu_{22}}{2\mu_1^2} - \frac{3\kappa_1\mu_2\mu_{12}}{\mu_1^3} + \frac{3\kappa_1^3\mu_2^3}{\mu_1^6} - \frac{3\kappa_1^3\mu_4}{4\mu_1^4} - \frac{18\kappa_1^2\mu_{11}\mu_2^2}{\mu_1^5} \\
&\quad \left. + \frac{6\kappa_1^2\mu_{11}\mu_3}{\mu_1^4} + \frac{9\kappa_1^2\mu_{21}\mu_2}{\mu_1^4} + \frac{9\kappa_1\mu_{11}^2\mu_2}{\mu_1^4} - \frac{2\kappa_1^2\mu_{31}}{\mu_1^3} - \frac{6\kappa_1\mu_{11}\mu_{21}}{\mu_1^3}\right) \\
&+ o(1) \quad \text{as } t \to \infty.
\end{aligned} \tag{2.26}$$

2.4. Evaluation of $E[N(t)]^3$. If we let $y_i \equiv 1$, then $\kappa_r = E[y_n^r] \equiv 1$. Consequently equation (2.26) becomes:

$$\begin{aligned} E[N(t)+1]^3 &= \frac{t^3}{\mu_1^3} + \frac{3t^2}{\mu_1^2}\left[\frac{3\mu_2}{2\mu_1^2} - 1\right] \\ &+ \frac{6t}{\mu_1}\left[\frac{3\mu_2^2}{2\mu_1^4} - \frac{\mu_3}{2\mu_1^3} - \frac{\mu_2}{\mu_1^2} + \frac{1}{6}\right] \\ &+ \left[\frac{\mu_2}{2\mu_1^2} + \frac{27\mu_2^2}{2\mu_1^4} + \frac{2\mu_3}{\mu_1^3} + \frac{3\mu_2^3}{\mu_1^6} - \frac{3\mu_4}{4\mu_1^4} - \frac{18\mu_2^2}{\mu_1^4}\right] \\ &+ o(1) \quad \text{as } t \to \infty. \end{aligned} \tag{2.27}$$

But as $E[N(t)]$, $E[N(t)]^2$ are already known (Smith [16]), then one can deduce $E[N(t)]^3$ from (2.27), since $E[N(t)+1]^3 = E[N(t)]^3 + 3E[N(t)]^2 + 3E[N(t)] + 1$. This small operation gives:

$$\begin{aligned} E[N(t)]^3 &= \frac{t^3}{\mu_1^3} + t^2\left[\frac{9}{2}\frac{\mu_2}{\mu_1^4} - \frac{6}{\mu_1^2}\right] + t\left[\frac{9\mu_2^2}{\mu_1^5} - \frac{3\mu_3}{\mu_1^4} - \frac{12\mu_2}{\mu_1^3} + \frac{7}{\mu_1}\right] \\ &+ \frac{14\mu_2\mu_1^4 - 36\mu_2^2\mu_1^2 + 16\mu_3\mu_1^3 + 12\mu_2^3 - 3\mu_4\mu_1^2 - 4\mu_1^6}{4\mu_1^6} \\ &+ o(1) \quad \text{as } t \to \infty. \end{aligned} \tag{2.28}$$

3. Expectation of a Product of Triple Summation

Murthy [14], gave an asymptotic expression for the covariance between $\sum_{i=1}^{N(t)+1} y_i^{(j)}$ and $\sum_{i=1}^{N(t)+1} y_i^{(k)}$, where $y_i^{(j)}$ and $y_i^{(k)}$ are the j-th and k-th components of the vector $\underline{Y}_i(n\times 1)$ and $i = 1, 2, \ldots, N(t)+1$. From this covariance one can deduce $E\left[\sum_{i=1}^{N(t)+1} y_i^{(j)} \sum_{i=1}^{N(t)+1} y_i^{(k)}\right]$.

In our present case, we want to evaluate

$$E\left[\prod_{r=\{j,k,\ell\}} \sum_{i=1}^{N(t)+1} y_i^{(r)}\right] = E\left[\sum_{i=1}^{N(t)+1} y_i^{(j)} \sum_{i=1}^{N(t)+1} y_i^{(k)} \sum_{i=1}^{N(t)+1} y_i^{(\ell)}\right] \tag{3.1}$$

where $y_i^{(j)}$, $y_i^{(k)}$ and $y_i^{(\ell)}$ are the j-th, k-th and ℓ-th components of the random vector $\underline{Y}_i(n \times 1)$.

We can interpret equation (3.1) as the three dimensional equation of Wald fundamental univariate equation which is given by:

$$E\left[\sum_{i=1}^{N(t)+1} y_i\right] = E[y_1]E[N(t)+1] \tag{3.2}$$

To find equation (3.1) let us define:

$$y_i = y_i^{(j)} + y_i^{(k)} + y_i^{(\ell)} \tag{3.3}$$

where $\{y_i\}$ are independently and identically distributed random variables, since $\{\underline{Y}_i(n \times 1\}$ are i.i.d random vectors. Obviously the variables $y_i^{(j)}$, $y_i^{(k)}$ and $y_i^{(\ell)}$ are dependant in general.

It follows from equation (3.3) that

$$\sum_{i=1}^{N(t)+1} y_i = \sum_{i=1}^{N(t)+1} y_i^{(j)} + \sum_{i=1}^{N(t)+1} y_i^{(k)} + \sum_{i=1}^{N(t)+1} y_i^{(\ell)}$$

$$\begin{aligned}
\text{Then} \quad E\left[\sum_{i=1}^{N(t)+1} y_i\right]^3 &= E\left[\sum_{i=1}^{N(t)+1} y_i^{(j)}\right]^3 + E\left[\sum_{i=1}^{N(t)+1} y_i^{(k)}\right]^3 \\
&+ E\left[\sum_{i=1}^{N(t)+1} y_i^{(\ell)}\right]^3 + 3E\left[\left(\sum_{i=1}^{N(t)+1} y_i^{(j)}\right)^2 \sum_{i=1}^{N(t)+1} y_i^{(k)}\right] \\
&+ 3E\left[\left(\sum_{i=1}^{N(t)+1} y_i^{(j)}\right)^2 \sum_{i=1}^{N(t)+1} y_i^{(\ell)}\right] \\
&+ 3E\left[\left(\sum_{i=1}^{N(t)+1} y_i^{(k)}\right)^2 \sum_{i=1}^{N(t)+1} y_i^{(j)}\right] \\
&+ 3E\left[\left(\sum_{i=1}^{N(t)+1} y_i^{(k)}\right)^2 \sum_{i=1}^{N(t)+1} y_i^{(\ell)}\right] \\
&+ 3E\left[\left(\sum_{i=1}^{N(t)+1} y_i^{(\ell)}\right)^2 \sum_{i=1}^{N(t)+1} y_i^{(k)}\right] \\
&+ 3E\left[\left(\sum_{i=1}^{N(t)+1} y_i^{(\ell)}\right)^2 \sum_{i=1}^{N(t)+1} y_i^{(j)}\right] \\
&+ 6E\left[\sum_{i=1}^{N(t)+1} y_i^{(j)} \sum_{i=1}^{N(t)+1} y_i^{(k)} \sum_{i=1}^{N(t)+1} y_i^{(\ell)}\right] \qquad (3.4)
\end{aligned}$$

3.1. Necessary Steps to Evaluate $E\left[\sum_{i=1}^{N(t)+1} y_i^{(j)} \sum_{i=1}^{N(t)+1} y_i^{(k)} \sum_{i=1}^{N(t)+1} y_i^{(\ell)}\right]$.

Since $E\left[\sum_{i=1}^{N(t)+1} y_i\right]^3$, $E\left[\sum_{i=1}^{N(t)+1} y_i^{(j)}\right]^3$, $E\left[\sum_{i=1}^{N(t)+1} y_i^{(k)}\right]^3$ and $E\left[\sum_{i=1}^{N(t)+1} y_i^{(\ell)}\right]^3$ can be evaluated by the result of equation (2.26) and computing the appropriate coefficients appearing in that formula, then to be able to evaluate $E\left[\sum_{i=1}^{N(t)+1} y_i^{(j)} \sum_{i=1}^{N(t)+1} y_i^{(k)} \sum_{i=1}^{N(t)+1} y_i^{(\ell)}\right]$, we have to know the following expressions:

- $$E\left[\left(\sum_{i=1}^{N(t)+1} y_i^{(j)}\right)^2 \sum_{i=1}^{N(t)+1} y_i^{(k)}\right] \quad \text{and} \quad E\left[\left(\sum_{i=1}^{N(t)+1} y_i^{(k)}\right)^2 \sum_{i=1}^{N(t)+1} y_i^{(j)}\right] \tag{3.5}$$
- $$E\left[\left(\sum_{i=1}^{N(t)+1} y_i^{(j)}\right)^2 \sum_{i=1}^{N(t)+1} y_i^{(\ell)}\right] \quad \text{and} \quad E\left[\left(\sum_{i=1}^{N(t)+1} y_i^{(\ell)}\right)^2 \sum_{i=1}^{N(t)+1} y_i^{(j)}\right] \tag{3.6}$$
- $$E\left[\left(\sum_{i=1}^{N(t)+1} y_i^{(k)}\right)^2 \sum_{i=1}^{N(t)+1} y_i^{(\ell)}\right] \quad \text{and} \quad E\left[\left(\sum_{i=1}^{N(t)+1} y_i^{(\ell)}\right)^2 \sum_{i=1}^{N(t)+1} y_i^{(k)}\right] \tag{3.7}$$

For this end, let us define the following auxiliary random variables:

$$y_i^{*} = y_i^{(j)} + y_i^{(k)} \tag{3.8}$$

$$y_i^{**} = y_i^{(j)} + y_i^{(\ell)} \tag{3.9}$$

$$y_i^{***} = y_i^{(k)} + y_i^{(\ell)} \tag{3.10}$$

With these definitions, we can find (3.5), (3.6) and (3.7). Indeed, we have:

- $$E\left[\sum_{i=1}^{N(t)+1} y_i^{*}\right]^3 - E\left[\sum_{i=1}^{N(t)+1} y_i^{(j)}\right]^3 - E\left[\sum_{i=1}^{N(t)+1} y_i^{(k)}\right]^3$$
$$= 3E\left[\left(\sum_{i=1}^{N(t)+1} y_i^{(j)}\right)^2 \sum_{i=1}^{N(t)+1} y_i^{(k)}\right] + 3E\left[\left(\sum_{i=1}^{N(t)+1} y_i^{(k)}\right)^2 \sum_{i=1}^{N(t)+1} y_i^{(j)}\right] \tag{3.11}$$
- $$E\left[\sum_{i=1}^{N(t)+1} y_i^{**}\right]^3 - E\left[\sum_{i=1}^{N(t)+1} y_i^{(j)}\right]^3 - E\left[\sum_{i=1}^{N(t)+1} y_i^{(\ell)}\right]^3$$
$$= 3E\left[\left(\sum_{i=1}^{N(t)+1} y_i^{(j)}\right)^2 \sum_{i=1}^{N(t)+1} y_i^{(\ell)}\right] + 3E\left[\left(\sum_{i=1}^{N(t)+1} y_i^{(\ell)}\right)^2 \sum_{i=1}^{N(t)+1} y_i^{(j)}\right] \tag{3.12}$$
- $$E\left[\sum_{i=1}^{N(t)+1} y_i^{***}\right]^3 - E\left[\sum_{i=1}^{N(t)+1} y_i^{(k)}\right]^3 - E\left[\sum_{i=1}^{N(t)+1} y_i^{(\ell)}\right]^3$$
$$= 3E\left[\left(\sum_{i=1}^{N(t)+1} y_i^{(k)}\right)^2 \sum_{i=1}^{N(t)+1} y_i^{(\ell)}\right] + 3E\left[\left(\sum_{i=1}^{N(t)+1} y_i^{(\ell)}\right)^2 \sum_{i=1}^{N(t)+1} y_i^{(k)}\right] \tag{3.13}$$

Therefore using equations (3.4), (3.11), (3.12) and (3.13), we obtain after simplification:

$$
\begin{aligned}
& E\left[\sum_{i=1}^{N(t)+1} y_i^{(j)} \sum_{i=1}^{N(t)+1} y_i^{(k)} \sum_{i=1}^{N(t)+1} y_i^{(\ell)}\right] \\
= \; & (1/6)\left\{E\left[\sum_{i=1}^{N(t)+1} y_i\right]^3 - E\left[\sum_{i=1}^{N(t)+1} y_i^{*}\right]^3 - E\left[\sum_{i=1}^{N(t)+1} y_i^{**}\right]^3\right. \\
- \; & E\left[\sum_{i=1}^{N(t)+1} y_i^{***}\right]^3 + E\left[\sum_{i=1}^{N(t)+1} y_i^{(j)}\right]^3 + E\left[\sum_{i=1}^{N(t)+1} y_i^{(k)}\right]^3 \\
+ \; & \left.E\left[\sum_{i=1}^{N(t)+1} y_i^{(\ell)}\right]^3\right\} \qquad (3.14)
\end{aligned}
$$

3.2. Evaluation of Equation (3.14).

In this section, we shall compute the appropriate coefficients of t^3, t^2, t^1 and t^0 appearing in formula (2.26) for each expectation involved in equation (3.14). To simplify the notation in the following cases which consist in the computation of those coefficients, let us introduce this notation:

$$\mu_{\alpha\beta\epsilon\gamma} = E\Big([y_i^{(j)}]^\alpha \, [y_i^{(k)}]^\beta \, [y_i^{(\ell)}]^\epsilon \, t_i^\gamma\Big) \qquad (3.15)$$

We are now ready to compute those coefficients for each of the following cases:

3.2.1. *The Case of* $E\left[\sum_{i=1}^{N(t)+1} y_i\right]^3$.

Since $y_i = y_i^{(j)} + y_i^{(k)} + y_i^{(\ell)}$, then

(1) $\kappa_1 \; = E[y_i] = \mu_{1000} + \mu_{0100} + \mu_{0010}$

(2)] $\kappa_2 \; = E[y_i^2] = \mu_{2000} + \mu_{0200} + \mu_{0020} + 2\mu_{1100} + 2\mu_{0110} + 2\mu_{1010}$

(3) $\mu_{21} = E[t_i^2 y_i] = \mu_{1002} + \mu_{0102} + \mu_{0012}$

(4) $\mu_{11} = E[t_i y_i] = \mu_{1001} + \mu_{0101} + \mu_{0011}$

(5) $\kappa_3 \; = E[y_i^3] = \mu_{3000} + \mu_{0300} + \mu_{0030} + 3\mu_{2100} + 3\mu_{2010} + 3\mu_{1200}$
$+ 3\mu_{0210} + 3\mu_{0120} + 3\mu_{1020} + 6\mu_{1110}$

(6) $\mu_{12} = E[t_i y_i^2] = \mu_{2001} + \mu_{0201} + \mu_{0021} + 2\mu_{1101} + 2\mu_{0111} + 2\mu_{1011}$

(7) $\mu_{22} = E[t_i^2 y_i^2] = \mu_{2002} + \mu_{0202} + \mu_{0022} + 2\mu_{1102} + 2\mu_{0112} + 2\mu_{1012}$

(8) $\mu_{31} = E[t_i^3 y_i] = \mu_{1003} + \mu_{0103} + \mu_{0013}$

3.2.2. *The Case of* $E\left[\sum_{i=1}^{N(t)+1} y_i^{(j)}\right]^3$.

(1) $\kappa_1 \; = E[y_i^{(j)}] = \mu_{1000}$

(2) $\kappa_2 \; = E[y_i^{(j)}]^2 = \mu_{2000}$

(3) $\mu_{21} = E[t_i^2 y_i^{(j)}] = \mu_{1002}$

(4) $\mu_{11} = E[t_i y_i^{(j)}] = \mu_{1001}$

(5) $\kappa_3 = E[y_i^{(j)}]^3 = \mu_{3000}$
(6) $\mu_{12} = E[t_i(y_i^{(j)})^2] = \mu_{2001}$
(7) $\mu_{22} = E[t_i^2(y_i^{(j)})^2] = \mu_{2002}$
(8) $\mu_{31} = E[t_i^3 y_i^{(j)}] = \mu_{1003}$

3.2.3. *The Case of* $E\left[\sum_{i=1}^{N(t)+1} y_i^{(k)}\right]^3$.

(1) $\kappa_1 = E[y_i^{(k)}] = \mu_{0100}$
(2) $\kappa_2 = E[y_i^{(k)}]^2 = \mu_{0200}$
(3) $\mu_{21} = E[t_i^2 y_i^{(k)}] = \mu_{0102}$
(4) $\mu_{11} = E[t_i y_i^{(k)}] = \mu_{0101}$
(5) $\kappa_3 = E[y_i^{(k)}]^3 = \mu_{0300}$
(6) $\mu_{12} = E[t_i(y_i^{(k)})^2] = \mu_{0201}$
(7) $\mu_{22} = E[t_i^2(y_i^{(k)})^2] = \mu_{0202}$
(8) $\mu_{31} = E[t_i^3 y_i^{(k)}] = \mu_{0103}$

3.2.4. *The Case of* $E\left[\sum_{i=1}^{N(t)+1} y_i^{(\ell)}\right]^3$.

(1) $\kappa_1 = E[y_i^{(\ell)}] = \mu_{0010}$
(2) $\kappa_2 = E[y_i^{(\ell)}]^2 = \mu_{0020}$
(3) $\mu_{21} = E[t_i^2 y_i^{(\ell)}] = \mu_{0012}$
(4) $\mu_{11} = E[t_i y_i^{(\ell)}] = \mu_{0011}$
(5) $\kappa_3 = E[y_i^{(\ell)}]^3 = \mu_{0030}$
(6) $\mu_{12} = E[t_i(y_i^{(\ell)})^2] = \mu_{0021}$
(7) $\mu_{22} = E[t_i^2(y_i^{(\ell)})^2] = \mu_{0022}$
(8) $\mu_{31} = E[t_i^3 y_i^{(\ell)}] = \mu_{0013}$

3.2.5. *The Case of* $E\left[\sum_{i=1}^{N(t)+1} y_i^*\right]^3$.

Since $y_i^* = y_i^{(j)} + y_i^{(k)}$, then

(1) $\kappa_1 = E[y_i^*] = \mu_{1000} + \mu_{0100}$
(2) $\kappa_2 = E[y_i^*]^2 = \mu_{2000} + \mu_{0200} + 2\mu_{1100}$
(3) $\mu_{21} = E[t_i^2 y_i^*] = \mu_{1002} + \mu_{0102}$
(4) $\mu_{11} = E[t_i y_i^*] = \mu_{1001} + \mu_{0101}$
(5) $\kappa_3 = E[y_i^*]^3 = \mu_{3000} + \mu_{0300} + 3\mu_{2100} + 3\mu_{1200}$
(6) $\mu_{12} = E[t_i(y_i^*)^2] = \mu_{2001} + \mu_{0201} + 2\mu_{1101}$
(7) $\mu_{22} = E[t_i^2(y_i^*)^2] = \mu_{2002} + \mu_{0202} + 2\mu_{1102}$
(8) $\mu_{31} = E[t_i^3 y_i^*] = \mu_{1003} + \mu_{0103}$

3.2.6. *The Case of* $E\left[\sum_{i=1}^{N(t)+1} y_i^{**}\right]^3$.

Recall that $y_i^{**} = y_i^{(j)} + y_i^{(\ell)}$, then

(1) $\kappa_1 = E[y_i^{**}] = \mu_{1000} + \mu_{0010}$

(2) $\kappa_2 = E[y_i^{**}]^2 = \mu_{2000} + \mu_{0020} + 2\mu_{1010}$

(3) $\mu_{21} = E[t_i^2 y_i^{**}] = \mu_{1002} + \mu_{0012}$

(4) $\mu_{11} = E[t_i y_i^{**}] = \mu_{1001} + \mu_{0011}$

(5) $\kappa_3 = E[y_i^{**}]^3 = \mu_{3000} + \mu_{0030} + 3\mu_{2010} + 3\mu_{1020}$

(6) $\mu_{12} = E[t_i (y_i^{**})^2] = \mu_{2001} + \mu_{0021} + 2\mu_{1011}$

(7) $\mu_{22} = E[t_i^2 (y_i^{**})^2] = \mu_{2002} + \mu_{0022} + 2\mu_{1012}$

(8) $\mu_{31} = E[t_i^3 y_i^{**}] = \mu_{1003} + \mu_{0013}$

3.2.7. *The Case of* $E\left[\sum_{i=1}^{N(t)+1} y_i^{***}\right]^3$.

Remember that $y_i^{***} = y_i^{(k)} + y_i^{(\ell)}$, then

(1) $\kappa_1 = E[y_i^{***}] = \mu_{0100} + \mu_{0010}$

(2) $\kappa_2 = E[y_i^{***}]^2 = \mu_{0200} + \mu_{0020} + 2\mu_{0110}$

(3) $\mu_{21} = E[t_i^2 y_i^{***}] = \mu_{0102} + \mu_{0012}$

(4) $\mu_{11} = E[t_i y_i^{***}] = \mu_{0101} + \mu_{0011}$

(5) $\kappa_3 = E[y_i^{***}]^3 = \mu_{0300} + \mu_{0030} + 3\mu_{0210} + 3\mu_{0120}$

(6) $\mu_{12} = E[t_i (y_i^{***})^2] = \mu_{0201} + \mu_{0021} + 2\mu_{0111}$

(7) $\mu_{22} = E[t_i^2 (y_i^{***})^2] = \mu_{0202} + \mu_{0022} + 2\mu_{0112}$

(8) $\mu_{31} = E[t_i^3 y_i^{***}] = \mu_{0103} + \mu_{0013}$

Note that $\mu_1 = E[t_i] = \mu_{0001}$, $\mu_2 = E[t_i^2] = \mu_{0002}$, $\mu_3 = E[t_i^3] = \mu_{0003}$ and $\mu_4 = E[t_i^4] = \mu_{0004}$ are the same for the cases 3.2.1 to 3.2.7.

3.3. Asymptotic Expression for $E\left[\sum_{i=1}^{N(t)+1} y_i^{(j)} \sum_{i=1}^{N(t)+1} y_i^{(k)} \sum_{i=1}^{N(t)+1} y_i^{(\ell)}\right]$.

By applying the result of equation (2.26) on each term appearing in equation (3.14), with the corresponding coefficients as computed in steps 3.2.1 to 3.2.7, we obtain after simplification the following asymptotic expression:

$$E\left[\sum_{i=1}^{N(t)+1} y_i^{(j)} \sum_{i=1}^{N(t)+1} y_i^{(k)} \sum_{i=1}^{N(t)+1} y_i^{(\ell)}\right] = \rho^3 \mu_{1000}\mu_{0100}\mu_{0010} t^3$$

$$+ \quad \Big\{ \rho^2 \mu_{1000}\mu_{0110} + \rho^2 \mu_{0100}\mu_{1010} + \rho^2 \mu_{0010}\mu_{1100}$$

$$-2\rho^3 \mu_{1000}\mu_{0100}\mu_{0011} - 2\rho^3 \mu_{1000}\mu_{0010}\mu_{0101} - 2\rho^3 \mu_{0100}\mu_{0010}\mu_{1001}$$

$$+\frac{9}{2}\rho^4 \mu_{1000}\mu_{0100}\mu_{0010}\mu_{0002} \Big\} t^2$$

$$+ \quad \Big\{ \rho\mu_{1110} + 2\rho^3 \mu_{1000}\mu_{0110}\mu_{0002} + 2\rho^3 \mu_{0100}\mu_{1010}\mu_{0002}$$

$$+2\rho^3 \mu_{0010}\mu_{1100}\mu_{0002} - \rho^2 \mu_{1000}\mu_{0111} - \rho^2 \mu_{0100}\mu_{1011}$$

$$-\rho^2 \mu_{0010}\mu_{1101} - \rho^2 \mu_{1100}\mu_{0011} - \rho^2 \mu_{1010}\mu_{0101} - \rho^2 \mu_{0110}\mu_{1001}$$

$$+2\rho^3 \mu_{1000}\mu_{0010}\mu_{0102} + 2\rho^3 \mu_{0100}\mu_{0010}\mu_{1002} + 2\rho^3 \mu_{1000}\mu_{0100}\mu_{0012}$$

$$+2\rho^3 \mu_{1000}\mu_{0101}\mu_{0011} + 2\rho^3 \mu_{0100}\mu_{1001}\mu_{0011} + 2\rho^3 \mu_{0010}\mu_{1001}\mu_{0101}$$

$$-6\rho^4 \mu_{1000}\mu_{0010}\mu_{0101}\mu_{0002} - 6\rho^4 \mu_{1000}\mu_{0100}\mu_{0011}\mu_{0002}$$

$$-6\rho^4 \mu_{0100}\mu_{0010}\mu_{1001}\mu_{0002}$$

$$+3\rho^5 \left(3\mu_{0002}^2 - \mu_{001}\mu_{0003}\right)\mu_{1000}\mu_{0100}\mu_{0010} \Big\} t$$

$$+ \quad \Big\{ \frac{\rho^2 \mu_{1110}\mu_{0002}}{2} + \frac{\rho^2 \mu_{1000}\mu_{0112}}{2} + \frac{\rho^2 \mu_{0100}\mu_{1012}}{2} + \frac{\rho^2 \mu_{0010}\mu_{1102}}{2}$$

$$+\frac{\rho^2 \mu_{1100}\mu_{0012}}{2} + \frac{\rho^2 \mu_{1010}\mu_{0102}}{2} + \frac{\rho^2 \mu_{0110}\mu_{1002}}{2}$$

$$-\rho^3 \mu_{1000}\mu_{0111}\mu_{0002} - \rho^3 \mu_{0100}\mu_{1011}\mu_{0002} - \rho^3 \mu_{0010}\mu_{1101}\mu_{0002}$$

$$-\rho^3 \mu_{1100}\mu_{0011}\mu_{0002} - \rho^3 \mu_{1010}\mu_{0101}\mu_{0002} - \rho^3 \mu_{0110}\mu_{1001}\mu_{0002}$$

$$-\rho^3 \mu_{1000}\mu_{0101}\mu_{0012} - \rho^3 \mu_{1000}\mu_{0102}\mu_{0011} - \rho^3 \mu_{0100}\mu_{0011}\mu_{1002}$$

$$-\rho^3 \mu_{0100}\mu_{1001}\mu_{0012} - \rho^3 \mu_{0010}\mu_{1001}\mu_{0102} - \rho^3 \mu_{0010}\mu_{0101}\mu_{1002}$$

$$-\frac{2\rho^3 \mu_{1000}\mu_{0100}\mu_{0013}}{3} - \frac{2\rho^3 \mu_{0100}\mu_{0010}\mu_{1003}}{3} - \frac{2\rho^3 \mu_{1000}\mu_{0010}\mu_{0103}}{3}$$

$$+3\rho^4 \mu_{1000}\mu_{0100}\mu_{0012}\mu_{0002} + 3\rho^4 \mu_{0100}\mu_{0010}\mu_{1002}\mu_{0002}$$

$$
\begin{aligned}
&+3\rho^4\mu_{1000}\mu_{0010}\mu_{0102}\mu_{0002}+3\rho^4\mu_{1000}\mu_{0101}\mu_{0011}\mu_{0002}\\
&+3\rho^4\mu_{0100}\mu_{1001}\mu_{0011}\mu_{0002}+3\rho^4\mu_{0001}\mu_{1001}\mu_{0101}\\
&+2\Phi_1\mu_{1000}\mu_{0110}+2\Phi_1\mu_{0100}\mu_{1010}+2\Phi_1\mu_{0010}\mu_{1100}\\
&-2\Phi_2\mu_{1000}\mu_{0100}\mu_{0011}-2\Phi_2\mu_{1000}\mu_{0010}\mu_{0101}\\
&-2\Phi_2\mu_{0100}\mu_{0010}\mu_{1001}+3\Phi_3\mu_{1000}\mu_{0100}\mu_{0010}\Big\}\\
&+\quad o(1) \quad \text{as } t\to\infty \qquad (3.16)
\end{aligned}
$$

where:

- $\rho = \dfrac{1}{\mu_{0001}}$
- $\Phi_1 = \dfrac{9\mu_{0002}^2-4\mu_{0001}\mu_{0003}}{12\mu_{0001}^4}$
- $\Phi_2 = \dfrac{3\mu_{0002}^2-\mu_{0001}\mu_{0003}}{\mu_{0001}^5}$
- $\Phi_3 = \dfrac{4\mu_{0002}^3-\mu_{0004}\mu_{0001}^2}{4\mu_{0001}^6}$
- $\mu_{\alpha\beta\epsilon\gamma} = E\left([y_i^{(j)}]^\alpha[y_i^{(k)}]^\beta[y_i^{(\ell)}]^\epsilon t_i^\gamma\right)$

It is appropriate to explain how we obtained the coefficient of t^3 in equation (3.16), since those of t^2, t^1 and t^0 are obtained using the same idea. The coefficient of t^3 is obtained as follows:

$$
\begin{aligned}
&\rho^3\Big[(\mu_{1000}+\mu_{0100}+\mu_{0010})^3-(\mu_{1000}+\mu_{0100})^3\\
&\quad -\ (\mu_{1000}+\mu_{0010})^3-(\mu_{0100}+\mu_{0010})^3\\
&\quad +\ \mu_{1000}^3+\mu_{0100}^3+\mu_{0010}^3\Big]/6\\
&\quad =\ \rho^3\mu_{1000}\mu_{0100}\mu_{0010} \qquad (3.17)
\end{aligned}
$$

4. Brief Study of $E\left[N(t)\right]^k$

Let $N(t)$ be a renewal process with $F(x)$ as its associated distribution function, then it follows from problems 16 and 17 page 233 in Karlin & Taylor [7], that $E[N(t)]^k = m_k(t)$ satisfies the following renewal equation

$$m_k(t) = Z_k(t) + \int_0^t m_k(t-\tau)dF(\tau) \tag{4.1}$$

$$\text{where} \quad Z_k(t) = \int_0^t \sum_{j=0}^{k-1} \binom{k}{j} m_j(t-\tau)dF(\tau) \tag{4.2}$$

$$\text{or} \quad Z_k(t) = (-1)^{k-1}F(t) + (-1)^k \sum_{j=1}^{k-1} (-1)^{j+1} \binom{k}{j} m_j(t) \tag{4.3}$$

4.1. Integral Equation for $E\left[N(t)\right]^k$. For this purpose, we state the following Lemma

LEMMA 4.1. *Let $N(t)$ be a renewal process then $E[N(t)]^k$ satisfies the following integral equation*

$$E[N(t)]^k = m_1(t) + \int_0^t \sum_{j=1}^{k-1} \binom{k}{j} m_j(t-\tau)dm_1(\tau) \tag{4.4}$$

Proof. We show the lemma by induction. It follows from theorem 5.1 in Karlin & Taylor [7] page 191, that

$$m_k(t) = Z_k(t) + \int_0^t Z_k(t-\tau)dm_1(\tau) \tag{4.5}$$

By substituting (4.3) in (4.5) results in

$$\begin{aligned} m_k(t) &= (-1)^{k-1}F(t) + (-1)^k \sum_{j=1}^{k-1} (-1)^{j+1} \binom{k}{j} m_j(t) \\ &\quad + \int_0^t (-1)^{k-1}F(t-\tau)dm_1(\tau) \\ &\quad + \int_0^t (-1)^k \sum_{j=1}^{k-1} (-1)^{j+1} \binom{k}{j} m_j(t-\tau)dm_1(\tau) \qquad (4.6) \\ &= (-1)^{k-1}m_1(t) + (-1)^k \left[(-1)^2 \binom{k}{1} \left\{m_1(t) + \int_0^t m_1(t-\tau)dm_1(\tau)\right\}\right. \\ &\quad \left. + \cdots + (-1)^k \binom{k}{k-1} \left\{m_{k-1}(t) + \int_0^t m_{k-1}(t-\tau)dm_1(\tau)\right\}\right] \qquad (4.7) \end{aligned}$$

Using the induction principle, equation (4.7) can be written as:

$$
\begin{aligned}
m_k(t) \;=\; & (-1)^{k-1}m_1(t) + m_1(t)\Big\{(-1)^k\Big((-1)^2\binom{k}{1} \\
& +\cdots+(-1)^k\binom{k}{k-1}\Big)\Big\} \\
& +(-1)^k\Big[(-1)^2\binom{k}{k-1}\int_0^t m_1(t-\tau)dm_1(\tau) \\
& +(-1)^3\binom{k}{k-2}\Big\{2\int_0^t m_1(t-\tau)dm_1(\tau)+\int_0^t m_2(t-\tau)dm_1(\tau)\Big\} \\
& +\cdots+(-1)^k\binom{k}{k-1}\Big\{\int_0^t\sum_{j=1}^{k-2}\binom{k-1}{j}m_j(t-\tau)dm_1(\tau) \\
& +\int_0^t m_{k-1}(t-\tau)dm_1(\tau)\Big\}\Big] \qquad (4.8)
\end{aligned}
$$

From (4.8) we have $\binom{k}{k-1}$ of $\int_0^t m_1(t-\tau)dm_1(\tau)$ and in general there is $(-1)^{k+1}\sum_{\ell=0}^{k-1}(-1)^{\ell}\binom{k}{j}\binom{\ell}{j}$ of $\int_0^t m_j(t-\tau)dm_1(\tau)$ $\forall j\geq 1$.

By using equation (34) page 63 of Knuth [9], we obtain

$$(-1)^{k+1}\sum_{\ell=0}^{k-1}(-1)^{\ell}\binom{k}{\ell}\binom{\ell}{j}=\binom{k}{j} \qquad (4.9)$$

Also from Abramowitz & Stegun [1] we have

$$1-\binom{k}{1}+\binom{k}{2}-\cdots+(-1)^k\binom{k}{k}=0 \qquad (4.10)$$

Applying (4.9) and (4.10) in (4.8), lemma 4.1 follows at once.

4.2. Laplace Transform of the Integral Equation (4.4). Taking the Laplace transform of (4.4) for $k=2$, Smith [16] found

$$m_2^*(s)=\int_0^\infty e^{-st}m_2(t)dt=\frac{F^*(s)}{s\left(1-F^*(s)\right)}+\frac{2\left(F^*(s)\right)^2}{s\left(1-F^*(s)\right)^2} \qquad (4.11)$$

where $F^*(s)$ is Laplace-Stieltjes transform of $F(t)$, given by Smith [17]:

$$F^*(s) = 1 - \mu_1 s + \frac{\mu_2 s^2}{2!} + \cdots + \frac{(-s)^n \mu_n}{n!} + o(s^n) \tag{4.12}$$

In fact

$$m_2^*(s) = \frac{1}{s}\sum_{r\geq 1}^{\infty} r^2\left[\left(F^*(s)\right)^r - \left(F^*(s)\right)^{r+1}\right] \tag{4.13}$$

In general we have

$$m_k^*(s) = \frac{1}{s}\sum_{r\geq 1}^{\infty} r^k\left[\left(F^*(s)\right)^r - \left(F^*(s)\right)^{r+1}\right] \tag{4.14}$$

$$m_k^*(s) = \frac{\left(1-F^*(s)\right)}{s}\sum_{r\geq 1}^{\infty} r^k\left(F^*(s)\right)^r \tag{4.15}$$

From theorem F page 245 of Comtet [6], which states that for each integer $k \geq 0$

$$\sum_{r\geq 0}^{\infty} r^k x^r = \frac{\sum_{j=1}^{k} A(k,j)x^j}{(1-x)^{k+1}} \tag{4.16}$$

Then

$$m_k^*(s) = \frac{\sum_{j=1}^{k} A(k,j)\left(F^*(s)\right)^j}{s\left(1-F^*(s)\right)^k} \tag{4.17}$$

where $A(k,j)$ are the Eulerian numbers, see e.g. Kimber [8].

Also note that from Abramowitz & Stegun [1] page 825, one has:

$$\sum_{k\geq 0}^{n} k^m x^k = \sum_{j=0}^{m} S(m,j)x^j \frac{\partial^j}{\partial x^j}\left\{\frac{1-x^{n+1}}{1-x}\right\} \tag{4.18}$$

where $S(n,m)$ are Stirling numbers of the second kind.

By applying equation (4.18), it can be easily shown that:

$$m_k^*(s) = \sum_{j=0}^{k} \frac{j!S(k,j)}{s}\left(\frac{F^*(s)}{1-F^*(s)}\right)^j \tag{4.19}$$

Consequently we derive from (4.19):

$$m_k^*(s) = \frac{\sum_{j=1}^{k}\left(F^*(s)\right)^j \sum_{\ell=0}^{j-1}(-1)^\ell j! S(k,j-\ell)\binom{k-j+\ell}{\ell}}{s\left(1-F^*(s)\right)^k} \tag{4.20}$$

Finally, it can be shown that:

$$m_k^*(s) = \frac{\left(1 - F^*(s)\right)}{s} m_k(s) \tag{4.21}$$

$$\text{where} \quad m_k(s) = \left(\frac{F^*(s)}{\frac{\partial F^*(s)}{\partial s}}\right)\left(\frac{\partial m_{k-1}(s)}{\partial s}\right) \tag{4.22}$$

$$\text{or} \quad m_k(s) = F^*(s) + 2^k\left(F^*(s)\right)^2 + 3^k\left(F^*(s)\right)^3 + \dots$$

5. Conclusion

As we observe, the method elaborated here, which is based on the characteristic function and the linearization technique for evaluating $E\left[\sum_{i=1}^{N(t)+1} X_i \sum_{i=1}^{N(t)+1} Y_i \sum_{i=1}^{N(t)+1} Z_i\right]$, is computationally very tedious. Such method is not efficient for higher moments.

Thus in our forthcoming papers [2] and [3], entitled "On the Moments of Randomly Stopped Cumulative Processes : The Cases $N(t)$ and $N(t) + 1$", which are respectively Chapter 2 and Chapter 3, we develop a new approach for evaluating the expectation of a product of n distinct cumulative processes. While computations remain complex, the new approach has the merit of making them much more systematic.

References

[1] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions*. National Bureau of Standards, Applied Mathematics Series 55. Tenth Printing (1972).

[2] M. Adès and R.P. Malhamé, *On the Moments of Randomly Stopped Cumulative Processes: The Case N(t)*, Les Cahiers du GERAD G-94-32. École des Hautes Études Commerciales, Montréal.

[3] M. Adès and R.P. Malhamé, *On the Moments of Randomly Stopped Cumulative Processes: The Case N(t) + 1*, in preparation.

[4] W.H. Beyer, *CRC Handbook of Mathematical Sciences*, 6*th* Edition, CRC Press (1987).

[5] R.P. Brent and H.T. Kung, $O\left((n \log n)^{3/2}\right)$ *Algorithms for Composition and Reversion of Power Series*, in Analytic Computational Complexity, edited by J.F. Traub, Academic Press (1976) 217-225.

[6] L. Comtet, *Advanced Combinatorics: The At of Finite and Infinite Expansions*, D. Reidel Publishing Company (1974).

[7] S. Karlin and H. Taylor, *A First course in Stochastic Processes*, Academic Press (1975).

[8] A.C. Kimber, *Eulerian Numbers*, in Encyclopedia of Statistical Sciences, Supplement Volume, John-Wiley & Sons (1989).

[9] D. Knuth, *the Art of Computer Programming. Volume 1. Fundamental Algorithms*, Second Edition, Addison-Wesley Publishing Company (1973).

[10] D. Knuth, *the Art of Computer Programming. Volume 2. Seminumerical Algorithms*, Second Edition, Addison-Wesley Publishing Company (1981).

[11] H.T. Kung, *On Computing Reciprocals of Power Series*, Numer. Math. **22** (1974), 341–348.

[12] R.G. Laha and V.K. Rohatgi, *Probability Theory*, John-Wiley & Sons (1979).

[13] A.I. Markushevich, *Theory of Functions of a Complex Variable*, Chelsea (1977).

[14] V.K. Murthy, *The General Point Process*, Addison-Wesley Publishing Company (1974).

[15] W.L. Smith, *Regenerative Stochastic Processes*, Proceedings of the Royal Society, Series A **232** (1955), 6–31.

[16] W.L. Smith, *On Renewal Theory, Counter Problems, and Quasi-Poisson Processes*, Proceedings of the Cambridge Philosophical Society **53** (1957), 175–193.

[17] W.L. Smith, *On the Cumulants of Renewal Processes*, Biometrika **46** (1959), 1–29.

[18] D.V. Widder, *The Laplace Transform*, Princeton University Press (1946).

CHAPTER 5

Stochastic Optimal Control Under Poisson Distributed Observations

1. Introduction

The optimal control of a partial observed stochastic system which evolves according to an Itô stochastic state system is well-known (Bagchi [1], Fleming and Rishel [5]). However, the optimal control of such continuous time systems, where the time domain is decomposed into a finite set of disjoint random intervals, where observations are taken at the initial instant of each interval, has not been carried out.

This optimal control problem is well motivated by potential applications to problems such as reservoir control. Consider a controller of a stochastic system, which may be taken to be observed at random times with a fixed total number of observations instants. This set may be regarded as a fund of observation actions which have been payed before the control exercise begins.

The inter-observation intervals will be taken to be generated by a sequence of i.i.d. $\mathbf{R}^+$ -valued random variables, i.e. the point process of observation times is a general independent increment process, and in particular, it will be given by the special case where the observations instants are Poisson distributed.

This chapter is organized as follows. In Section 2, we formulate two classes of optimal control problems, termed (piecewise) time invariant and time variant control, where the central issue is the control structure in each case. In Sections 3 and 4, we present in details the solution in the (piecewise) time invariant case, where the observation time instants are respectively a stochastic process with independent increments and Poisson distributed. In Section 5, we present the time variant case, where the mean time between observation instants is exponentially distributed in its parameter λ. A stochastic dynamic programming framework is used fr— the solution of these optimal control problems.

2. Problem Statement

We consider a stochastic system evolving according to the following time-invariant Itô stochastic state equation

$$dX(t) = AX(t)dt + BU(t)dt + GdW(t) \quad (2.1)$$

where X and U are respectively vector state and control processes with dimensions $d \times 1$ and $p \times 1$; while $A(d \times d), B(d \times p)$ and $G(d \times m)$ are respectively plant, control and disturbance distribution matrices. W is a normalized zero mean standard vector $(m \times 1)$ Brownian motion.

We decompose the time domain of the continuous equation in (2.1) by taking a finite number N of point observations at times t_i, where $t = \{t_i, i \geq 0\}$ is a stochastic process with independent increments, and we set $t_0 = 0$.

Following this process of decomposition of the time domain into a finite set of N disjoint random intervals $[t_i, t_{i+1})$ for $0 \leq i \leq N-1$; we associate on each interval a state space system equations with complete initial state observation, defined as follows:

$$dX(t) = AX(t)dt + BU(t)dt + GdW(t) \quad (2.2a)$$

$$X(t_i) = I.C. \text{ (initial conditions)}$$

$$Z(t_i) = X(t_i) \text{ for } t_i \leq t < t_{i+1} \text{ and } 0 \leq i < N. \quad (2.2b)$$

As $X(t)$ is a Markov process, it is known (Feller [4], Stam [10]), that the derived process $Z(t_i)$ is a Markov process. This fact will be used in the stochastic dynamic programming procedure (Caines [2], Davis [3], Fleming and Rishel [5]), and employed in the solution of the stochastic optimization problems in the following sections. We consider in sections (3, 4) and 5 two optimization control problems which we term time invariant control (piecewise-constant control), and time variant control problems respectively.

The objective in both cases is to construct a control law that satisfies the linear stochastic control system and minimizes the quadratic expected cost functional specified further down in (2.3) - (2.5) and (2.6) - (2.8).

The central issue of interest in the specification and solution of these problems is the particular information and control structure in each case; this will be clearly specified below for both control problems.

For notational convenience, we shall consider that conditioning on $X(t_i)$ corresponds in fact to the conditioning of both $X(t_i)$ and t_i. This telescoping technique of conditioning is used in the stochastic dynamic programming framework since $X(t_i)$ is of a Markovian nature.

Problem 1: (Piecewise-Constant) Time-Invariant Control

$$J(X(t_0)) = \min_{U\epsilon\mathcal{U}_{PI}} \mathrm{E}_{W,t}\Big[\int_{t_0}^{t_N} (X'(t)CX(t) + U'(\lfloor X\rfloor_t)DU(\lfloor X\rfloor_t))dt\Big] \tag{2.3}$$

$$= \min_{U\epsilon\mathcal{U}_{PI}} \mathrm{E}_W\Big[\sum_{i=0}^{N-1} \mathrm{E}_{t_{i+1}}\Big\{\int_{t_i}^{t_{i+1}} (X'(t)CX(t) + U'(\lfloor X\rfloor_t)DU(\lfloor X\rfloor_t))dt\big|X(t_i)\Big\}\Big] \tag{2.4}$$

which is associated with the following dynamics

$$dX(t) = AX(t)dt + BU(\lfloor X\rfloor_t)dt + GdW(t) \tag{2.5}$$

where

$U(\lfloor X\rfloor_t) \equiv U(X(t_i)) \,\epsilon\, \mathcal{U}_{PI}$ for $t_i \le t < t_{i+1}$.
$\mathcal{U}_{PI}$ is the class of admissible control laws (i.e. bounded or contained within a certain region), associated with this problem.
$U(\cdot) : \mathbf{R}^d \to \mathbf{R}^p$, where $U(\cdot)$ is Borel measurable, with respect to the σ-field $\mathcal{F}\{X(t_i)\}$, on each interval $[t_i, t_{i+1})$.
$C(d \times d), D(p \times p)$ are symmetric and respectively positive definite and positive semi-definite matrices.
$\{\cdot|X(t_i)\}$ corresponds to $\{\cdot|X(t_i), t_i\}$ for short.

See the Appendix for details about equations (2.3) to (2.4).

Problem 2: Time Variant Control

$$J(X(t_0)) = \min_{U\epsilon\mathcal{U}_{TV}} \mathrm{E}_{W,t}\Big[\int_{t_0}^{t_N} (X'(t)CX(t) + U'(t,\lfloor X\rfloor_t)DU(t,\lfloor X\rfloor_t))dt\Big] \tag{2.6}$$

$$= \min_{U\epsilon\mathcal{U}_{TV}} \mathrm{E}_W\Big[\sum_{i=0}^{N-1} \mathrm{E}_{t_{i+1}}\Big\{\int_{t_i}^{t_{i+1}} (X'(t)CX(t) + U'(t,\lfloor X\rfloor_t)DU(t,\lfloor X\rfloor_t))dt\big|X(t_i)\Big\}\Big] \tag{2.7}$$

which is associated with the following dynamics

$$dX(t) = AX(t)dt + BU(t,\lfloor X\rfloor_t)dt + GdW(t) \tag{2.8}$$

where

$U(t,\lfloor X\rfloor_t) \equiv U(t, X(t_i)) \,\epsilon\, \mathcal{U}_{TV}$ for $t_i \le t < t_{i+1}$.
$\mathcal{U}_{TV}$ is the class of admissible control laws associated with this problem.
$U(\cdot,\cdot) : \mathbf{R} \times \mathbf{R}^d \to \mathbf{R}^p$, where $U(\cdot,\cdot)$ is jointly Borel measurable, with respect to the σ-field $\mathcal{F}\{t, X(t_i)\}$, on each interval $[t_i, t_{i+1})$.

For the (piecewise-constant) time invariant control problem, we consider first the t_i's as a stochastic process with independent increments (termed an II process). For this problem, the class of admissible control laws is $\mathcal{U}_{PI_G}$, where "G" means a General II process.

Then, we take the sequence t_i's as Poisson distributed with mean inter-arrival time $\mu = 1/\lambda$ for both the first and second control problems. Thus, we associate $\mathcal{U}_{PI_P}$ and $\mathcal{U}_{TV_P}$ as the class of admissible control laws, respectively with the (piecewise-constant) time invariant and time variant control problems.

3. Optimal (Piecewise) Time Invariant Control: $\{t_i\}$ an II Process

The decomposition of the time domain of the continuous stochastic differential equation in (2.5), yields as solution the following:

$$X(t_{i+1}) = \alpha_{t_{i+1}} X(t_i) + \beta_{t_{i+1}} U(X(t_i)) + \gamma_{t_{i+1}} \tag{3.1}$$

where

$$\begin{aligned}
& i = 0, 1, \ldots, N-1 \ ; \ t_0 = 0 \\
& \alpha_{t_{i+1}} = e^{A(t_{i+1}-t_i)} \\
& \beta_{t_{i+1}} = \int_{t_i}^{t_{i+1}} e^{A(t_{i+1}-s)} B d_s = (e^{A(t_{i+1}-t_i)} - I) A^{-1} B \\
& \gamma_{t_{i+1}} = \int_{t_i}^{t_{i+1}} e^{A(t_{i+1}-s)} G dW_s
\end{aligned}$$

As we mentioned in the beginning, we use a dynamic programming formulation for the solution of this optimization problem, where the i-th stage starts with the occurrence of the i-th measurement, and the optimal cost-to-go for this i-th stage is given by

$$\begin{aligned}
V(X(t_i), i) &\triangleq V(X(t_i), t_i, i) \\
&= \min_{U \epsilon \mathcal{U}_{PI_G}} \mathrm{E}_W \Big[\sum_{j=i}^{N-1} \mathrm{E}_{t_{j+1}} \Big\{ \int_{t_j}^{t_{j+1}} (X'(t) C X(t) + U'(X(t_j)) D U(X(t_j))) dt \big| X(t_j) \Big\} \Big]
\end{aligned} \tag{3.2}$$

where $U(X(t_j)) \epsilon\, \mathcal{U}_{PI_G}$.

The terminal optimal cost-to-go associated with the Nth stage is given by:

$$V(X(t_N), N) = 0. \tag{3.3}$$

At the $(N-1)$th stage, the optimal cost-to-go is given by:

$$\begin{aligned}
V(X(t_{N-1}), N-1) = \min_{U_{N-1} \epsilon \mathcal{U}_{PI_{G,N-1}}} \mathrm{E}_W \Big[\mathrm{E}_{t_N} \Big\{ \int_{t_{N-1}}^{t_N} (X'(t) C X(t) \\
+ U'(X(t_{N-1})) D U(X(t_{N-1}))) dt \big| X(t_{N-1}) \Big\} \Big]
\end{aligned} \tag{3.4}$$

where $U_{N-1} \equiv U(X(t_{N-1})) \epsilon\, \mathcal{U}_{PI_{G,N-1}}$.

Using the independence between the Brownian process W and the point process $t = \{t_i, i \geq 0\}$, we obtain the following optimal control:

$$U^*(X(t_{N-1})) = -L_{N-1}X(t_{N-1}) \tag{3.5}$$

where

$$L_{N-1} = (q_{2,N-1} + D\mu_{N-1})^{-1} q_{1,N-1}$$
$$q_{1,N-1} = \mathrm{E}_{t_N}\Big[\int_{t_{N-1}}^{t_N} \beta_t' C\alpha_t dt \big| X(t_{N-1})\Big]$$
$$q_{2,N-1} = \mathrm{E}_{t_N}\Big[\int_{N-1}^{t_N} \beta_t' C\beta_t dt \big| X(t_{N-1})\Big].$$

After appropriate computations, the optimal cost-to-go value for the state $(N-1)$th stage is given by:

$$V(X(t_{N-1}), N-1) = X'(t_{N-1})\widetilde{L}_{N-1}X(t_{N-1}) + \widehat{\ell}_{N-1} \tag{3.6}$$

where

$$\widetilde{L}_{N-1} = q_{3,N-1} - q_{1,N-1}' L_{N-1} - L_{N-1}' q_{1,N-1}$$
$$+ L_{N-1}' q_{2,N-1} L_{N-1} + L_{N-1}' D L_{N-1} \mu_{N-1}$$
$$q_{3,N-1} = \mathrm{E}_{t_N}\Big[\int_{t_{N-1}}^{t_N} \alpha_t' C\alpha_t dt \big| X(t_{N-1})\Big]$$
$$q_{4,N-1} = \widehat{\ell}_{N-1} = \mathrm{E}_W\Big[\mathrm{E}_{t_N}\Big[\int_{t_{N-1}}^{t_N} \gamma_t' C\gamma_t dt \big| X(t_{N-1})\Big]\Big].$$

At this stage, it is clear that the optimal cost-to-go is quadratic in its initial state; we postulate then the hypotheses that the optimal cost structure will remain quadratic in its initial state.

Thus, let:

$$V(X(t_{k+1}), k+1) = X'(t_{k+1})\widetilde{L}_{k+1}X(t_{k+1}) + \widehat{\ell}_{k+1} \tag{3.7}$$

where $0 \leq k < N$, $\widetilde{L}_{k+1}$ and $\widehat{\ell}_{k+1}$ are some specific constants.

Using backwards induction, we show that:

$$V(X(t_k), k) = X'(t_k)\widetilde{L}_k X(t_k) + \widehat{\ell}_k. \tag{3.8}$$

By the Principle of Optimality in the dynamic programming framework, we have:

$$V(X(t_k),k) = \min_{U_k \epsilon \mathcal{U}_{PI_{G,k}}} \mathrm{E}_W \big[\mathrm{E}_{t_{k+1}} \{ \int_{t_k}^{t_{k+1}} (X'(t)CX(t) + U'(X(t_k))DU(X(t_k)))dt + V(X(t_{k+1}),k+1) | X(t_k) \} \big] \tag{3.9}$$

where $U_k \equiv U(X(t_k)) \epsilon \mathcal{U}_{PI_{G,k}}$.

Indeed,

$$V(X(t_k),k) = \min_{U_{t_k} \epsilon \mathcal{U}_{PI}} \mathrm{E}_W^{[t_k,t_{k+1})} \big[\mathrm{E}_{t_{k+1}} \{ \int_{t_k}^{t_{k+1}} \ell(X(t),U(X(t_k)))dt + \min_{\{U_{t_{k+1}}^{t_{N-1}} \epsilon \mathcal{U}_{PI}\}} \mathrm{E}_W^{[t_{k+1},t_N)} [\mathrm{E}_{t_{j+1}} \{ \sum_{j=k+1}^{N-1} \int_{t_j}^{t_{j+1}} \ell(X(t),U(X(t_j)))dt \} | X(t_j)] | X(t_k) \} \big] \tag{3.10a}$$

$$= \min_{U_{t_k} \epsilon \mathcal{U}_{PI}} \mathrm{E}_W^{[t_k,t_{k+1})} \big[\mathrm{E}_{t_{k+1}} \{ \int_{t_k}^{t_{k+1}} \ell(X(t),U(X(t_k))dt + V((X(t_{k+1}),U_{t_k}),k+1) | X(t_k) \} \big] \tag{3.10b}$$

where equation (3.10b) constitutes the Principle of Optimality of stochastic dynamic programming, see e.g. Chapter 11 in Caines [2].

After careful computations, we obtain the optimal control and cost-to-go respectively as:

$$U^*(X(t_k)) = -L_k X(t_k) \tag{3.11}$$

$$V(X(t_k),k) = X'(t_k)\overline{L}_k X(t_k) + \widehat{\ell}_k \tag{3.12}$$

where

$$
\begin{aligned}
L_k &= \big(q_{2,k} + D\mu_k + \frac{1}{2}\widetilde{q}_{2,k} + \frac{1}{2}\widehat{q}_{2,k}\big)^{-1}\big(q_{1,k} + \frac{1}{2}\widehat{q}_{1,k} + \frac{1}{2}\widetilde{q}_{1,k}\big) \\
\widetilde{L}_k &= q_{3,k} - q'_{1,k}L_k - L'_k q_{1,k} + L'_k q_{2,k} L_k + L'_k D L_k \mu_k + \widetilde{q}_{3,k} - \check{q}_{1,k} L_k \\
&\qquad - L'_k \widetilde{q}_{1,k} + L'_k \widetilde{q}_{2,k} L_k \\
\widehat{\ell}_k &= \sum_{j=N-1}^{k} q_{4,j} + \widetilde{q}_{4,k}; 0 \le k < N \\
q_{1,k} &= \mathrm{E}_{t_{k+1}}\big[\int_{t_k}^{t_{k+1}} \beta'_t C \alpha_t dt \big| X(t_k)\big] \\
q_{2,k} &= \mathrm{E}_{t_{k+1}}\big[\int_{t_k}^{t_{k+1}} \beta'_t C \beta_t dt \big| X(t_k)\big] \\
q_{3,k} &= \mathrm{E}_{t_{k+1}}\big[\int_{t_k}^{t_{k+1}} \alpha'_t C \alpha_t dt \big| X(t_k)\big] \\
q_{4,k} &= \mathrm{E}w\big[\mathrm{E}_{t_{k+1}}\big[\int_{t_k}^{t_{k+1}} \gamma'_t C \gamma_t dt \big| X(t_k)\big]\big] \\
\widetilde{q}_{1,k} &= \mathrm{E}_{t_{k+1}}\big[\beta'_t \widetilde{L}_{k+1} \alpha_t \big| X(t_k)\big] \\
\widetilde{q}_{2,k} &= \mathrm{E}_{t_{k+1}}\big[\beta'_t \widetilde{L}_{k+1} \beta_t \big| X(t_k)\big] \\
\widetilde{q}_{3,k} &= \mathrm{E}_{t_{k+1}}\big[\alpha'_t \widetilde{L}_{k+1} \alpha_t \big| X(t_k)\big] \\
\widetilde{q}_{4,k} &= \mathrm{E}w\big[\mathrm{E}_{t_{k+1}}\big[\gamma'_t \widetilde{L}_{k+1} \gamma_t \big| X(t_k)\big]\big] \\
\widehat{q}_{1,k} &= \mathrm{E}_{t_{k+1}}\big[\beta'_t \widetilde{L}'_{k+1} \alpha_t \big| X(t_k)\big] \\
\widehat{q}_{2,k} &= \mathrm{E}_{t_{k+1}}\big[\beta'_t \widetilde{L}'_{k+1} \beta_t \big| X(t_k)\big] \\
\check{q}_{1,k} &= \mathrm{E}_{t_{k+1}}\big[\alpha'_t \widetilde{L}_{k+1} \beta_t \big| X(t_k)\big].
\end{aligned}
$$

Note that for $k = N-1$ only, we have:

$$\widetilde{q}_{1,N-1} = \widetilde{q}_{2,N-1} = \widetilde{q}_{3,N-1} = \widetilde{q}_{4,N-1} = \widehat{q}_{1,N-1} = \widehat{q}_{2,N-1} = \check{q}_{1,N-1} = 0.$$

Thus we have proved equation (3.12), and this establishes our result for this section.

3.1. Numerical Example. We consider the following scalar stochastic differential equation:

$$dx(t) = -x(t)dt + u(\lfloor x \rfloor_t)dt + dw(t) \tag{3.13}$$

with the associated cost functional

$$J(x(t_0)) = \min_{u \in \mathcal{U}_{PI_G}} \mathrm{E}w\big[\sum_{i=0}^{N-1} \mathrm{E}_{t_{i+1}}\big\{\int_{t_i}^{t_{i+1}} (x^2(t) + u^2(x(t_i)))dt \big| x(t_i)\big\}\big] \tag{3.14}$$

where $u(x(t_i)) \in \mathcal{U}_{PI_G}$.

Equations (3.13) and (3.14) are a scalar version of equations (2.5) and (2.4).

Following the general framework solution elaborated through equations (3.3) to (3.12), we have:

$$V(x(t_N), N) = 0. \tag{3.15}$$

At the $(N-1)$th stage the optimal cost-to-go is

$$V(x(t_{N-1}), N-1) = \min_{u_{N-1} \in \mathcal{U}_{PI_G,N-1}} \mathrm{E}_W \big[\mathrm{E}_{t_N}\big\{\int_{t_{N-1}}^{t_N} (x^2(t) + u^2(x(t_{N-1})))dt | x(t_{N-1})\big\}\big] \tag{3.16}$$

$$\begin{aligned} &= \min_{u_{N-1} \in \mathcal{U}_{PI_G,N-1}} \mathrm{E}_W \big[\mathrm{E}_{t_N}\big\{\int_{t_{N-1}}^{t_N} (\alpha_t^2 x^2(t_{N-1}) + \beta_t^2 u^2(x(t_{N-1})) \\ &+ \gamma_t^2 + 2x(t_{N-1})u(x(t_{N-1}))\alpha_t\beta_t + 2x(t_{N-1})\alpha_t\gamma_t \\ &+ 2u(x(t_{N-1}))\beta_t\gamma_t + u^2(x(t_{N-1})))dt | x(t_{N-1})\big\}\big] \end{aligned} \tag{3.17}$$

where $u_{N-1} \equiv u(x(t_{N-1})) \in \mathcal{U}_{PI_G,N-1}$.

Using the independence between the Brownian motion ω and the point process $t = \{t_i, i \geq 0\}$, we obtain:

$$u^*(x(t_{N-1})) = -L_{N-1}x(t_{N-1}) \tag{3.18}$$

where

$$\begin{aligned} L_{N-1} &= \big(q_{2,N-1} + \mu_{N-1}\big)^{-1} q_{1,N-1} \\ q_{1,N-1} &= \mathrm{E}_{t_N}\big[\int_{t_{N-1}}^{t_N} \alpha_t\beta_t dt | x(t_{N-1})\big] \\ q_{2,N-1} &= \mathrm{E}_{t_N}\big[\int_{t_{N-1}}^{t_N} \beta_t^2 dt | x(t_{N-1})\big]. \end{aligned}$$

The resulting optimal cost-to-go is

$$V(x(t_{N-1}), N-1) = x^2(t_{N-1})\widehat{L}_{N-1} + \widehat{\ell}_{N-1} \tag{3.19}$$

where

$$\begin{aligned} \widehat{L}_{N-1} &= q_{3,N-1} - 2L_{N-1}q_{1,N-1} + L_{N-1}^2 q_{2,N-1} + L_{N-1}^2 \mu_{N-1} \\ q_{3,N-1} &= \mathrm{E}_{t_N}\big[\int_{t_{N-1}}^{t_N} \alpha_t^2 dt | x(t_{N-1})\big] \\ q_{4,N-1} &= \widehat{\ell}_{N-1} = \mathrm{E}_W\big[\mathrm{E}_{t_N}\big[\int_{t_{N-1}}^{t_N} \gamma_t^2 dt | x(t_{N-1})\big]\big]. \end{aligned}$$

Applying equations (3.7) to (3.12), we can compute for an arbitrary stage, $\widetilde{L}_k$ and $\widetilde{\ell}_k$ in

$$V(x(t_k),k) = x^2(t_k)\widetilde{L}_k + \widehat{\ell}_k \tag{3.20}$$

which is a scalar version of equation (3.8). This means that the structure of the optimal cost remains quadratic in its initial state from stage to another.

At the k-th stage, the optimal cost-to-go is:

$$\begin{aligned} V(x(t_k),k) = \min_{u_k \in \mathcal{U}_{PI_{G,k}}} \mathrm{E}_W\big[\mathrm{E}_{t_{k+1}}\big\{\int_{t_k}^{t_{k+1}} (x^2(t) + u^2(x(t_k)))dt \\ + V(x(t_{k+1}),k+1)|x(t_k)\big\}\big] \end{aligned} \tag{3.21}$$

$$\begin{aligned} = \min_{u_k \in \mathcal{U}_{PI_{G,k}}} \mathrm{E}_W\big[\mathrm{E}_{t_{k+1}}\big\{\int_{t_k}^{t_{k+1}} (\alpha_t^2 x^2(t_k) + \beta_t^2 u^2(x(t_k)) \\ + \gamma_t^2 + 2x(t_k)u(x(t_k))\alpha_t\beta_t + 2x(t_k)\alpha_t\gamma_t + 2u(x(t_k))\beta_t\gamma_t + u^2(x(t_k)))dt \\ + V(x(t_{k+1}),k+1)|x(t_k)\big\}\big] \end{aligned} \tag{3.22}$$

where $u_k \equiv u(x(t_k)) \in \mathcal{U}_{PI_{G,k}}$.

The optimal control law and the optimal cost-to-go are given respectively by:

$$u^*(x(t_k)) = -L_k x(t_k) \tag{3.23}$$

$$V(x(t_k),k) = x^2(t_k)\widetilde{L}_k + \widehat{\ell}_k \tag{3.24}$$

where

$$L_k = \left(q_{2,k} + \mu_k + \widetilde{q}_{2,k}\right)^{-1}\left(q_{1,k} + \widetilde{q}_{1,k}\right)$$

$$\widetilde{L}_k = q_{3,k} - 2q_{1,k}L_k + L_k^2 q_{2,k} + L_k^2\mu_k + \widetilde{q}_{3,k} + L_k^2\widetilde{q}_{2,k} - 2L_k\widetilde{q}_{1,k}$$

$$\widehat{\ell}_k = \sum_{j=N-1}^{k} q_{4,j} + \widetilde{q}_{4,k}.$$

For this scalar case, we note that:

$$\widetilde{q}_{1,k} = \widehat{q}_{1,k} = \check{q}_{1,k} = \widetilde{L}_{k+1}\mathrm{E}_{t_{k+1}}\left[\beta_t\alpha_t|x(t_k)\right]$$

and

$$\widetilde{q}_{2,k} = \widehat{q}_{2,k} = \widetilde{L}_{k+1}\mathrm{E}_{t_{k+1}}\left[\beta_t^2|x(t_k)\right].$$

If $(t_{k+1}-t_k)$ is exponentially distributed in its parameter λ, then we can evaluate explicitly $q_{1,k}$ to $\bar{q}_{4,k}$. Let us evaluate $q_{1,k}$.

$$
\begin{aligned}
q_{1,k} &= \mathrm{E}_{t_{k+1}}\Big[\int_{t_k}^{t_{k+1}} \alpha_t\beta_t dt | x(t_k)\Big] && (3.25)\\
&= \mathrm{E}_{t_{k+1}}\Big[\int_{t_k}^{t_{k+1}} e^{-(t-t_k)}\Big(\int_{t_k}^{t_{k+1}} e^{-(t-s)}ds\Big)dt | x(t_k)\Big]\\
&= \mathrm{E}_{t_{k+1}}\Big[\int_{t_k}^{t_{k+1}} e^{-(t-t_k)}\big(1-e^{-(t-t_k)}\big)dt | x(t_k)\Big]\\
&= \mathrm{E}_{t_{k+1}}\Big[\int_{t_k}^{t_{k+1}} \big(e^{-(t-t_k)}-e^{-2(t-t_k)}\big)dt | x(t_k)\Big] && (3.26)\\
&= \int_{t_k}^{\infty}\Big(\int_{t_k}^{t_{k+1}} \big(e^{-(t-t_k)}-e^{-2(t-t_k)}\big)dt\Big)\lambda\big(e^{-\lambda(t_{k+1}-t_k)}dt_{k+1} && (3.27)\\
&= \int_{t_k}^{\infty}\Big(\int_{t}^{\infty} \lambda e^{-\lambda(t_{k+1}-t_k)}dt_{k+1}\Big)\big(e^{-(t-t_k)}-e^{-2(t-t_k)}\big)dt && (3.28)\\
&= \int_{t_k}^{\infty} e^{-\lambda t}e^{\lambda t_k}\big(e^{-(t-t_k)}-e^{-2(t-t_k)}\big)dt\\
&= \big((\lambda+1)(\lambda+2)\big)^{-1}.
\end{aligned}
$$

We justify equations (3.26)-(3.28) by the distribution of the increment $(t_{k+1}-t_k)$ and Fubini theorem.

Using essentially the same method, we can evaluate the rest. Thus, we obtain:

$$
\begin{aligned}
q_{2,k} &= 2\big(\lambda(\lambda+1)(\lambda+2)\big)^{-1}\\
q_{3,k} &= (\lambda+2)^{-1}\\
q_{4,k} &= \big(\lambda(\lambda+2)\big)^{-1}\\
\bar{q}_{1,k} &= \bar{L}_{k+1}\lambda\big((\lambda+1)(\lambda+2)\big)^{-1}\\
\bar{q}_{2,k} &= 2\bar{L}_{k+1}\big((\lambda+1)(\lambda+2)\big)^{-1}\\
\bar{q}_{3,k} &= \bar{L}_{k+1}\lambda(\lambda+2)^{-1}\\
\bar{q}_{4,k} &= \bar{L}_{k+1}(\lambda+2)^{-1}.
\end{aligned}
$$

4. Optimal (Piecewise) Time Invariant Control: The Poisson Case

We consider here the quadratic expected cost functional and the dynamics of the stochastic system given in (2.3)-(2.5).

However, in this case we restrict the Π point process so that $t_0 = 0$ and the increments $t_{k+1} - t_k, k \geq 0$ are exponentially distributed with mean inter-arrival time $\mu_k = 1/\lambda$ for $0 \leq k < N$.

For this problem, recall that $\mathcal{U}_{PI_p}$ is the class of admissible control laws.

Using the same general framework as in the previous section, the terminal optimal cost-to-go associated with the N-th stage is given by:

$$V(X(t_N), N) = 0. \tag{4.1}$$

At the $(N-1)$th stage, the optimal cost-to-go is given by:

$$V(X(t_{N-1}), N-1) = \min_{U_{N-1}\in\mathcal{U}_{PI_{P,N-1}}} \mathrm{E}_W\Big[\mathrm{E}_{t_N}\Big\{\int_{t_{N-1}}^{t_N} (X'(t)CX(t) + U'(X(t_{N-1}))DU(X(t_{N-1})))dt \Big| X(t_{N-1})\Big\}\Big] \tag{4.2}$$

where $U_{N-1} \equiv U(X(t_{N-1})) \in \mathcal{U}_{PI_{P,N-1}}$.

Again using the independence between the Brownian motion W and the point process $t = \{t_i, i \geq 0\}$ in (4.2), and evaluating the expectation in its respect to t_N (given t_{N-1}; recall that $(\cdot|X(t_{N-1}))$ corresponds to $(\cdot|X(t_{N-1}), t_{N-1})$), we obtain:

$$V(X(t_{N-1}), N-1) = \min_{U_{N-1}\in\mathcal{U}_{PI_{P,N-1}}} \mathrm{E}_W\Big[\int_{t_{N-1}}^{\infty}\Big\{\int_{t_{N-1}}^{t_N} (X'(t)CX(t) + U'(X(t_{N-1}))DU(X(t_{N-1})))dt\Big\}\lambda e^{-\lambda(t_N - t_{N-1})}dt_N \Big| X(t_{N-1})\Big]. \tag{4.3}$$

Using Fubini's theorem in (4.3), one can interchange the order of integration to yield:

$$V(X(t_{N-1}), N-1) = \min_{U_{N-1}\in\mathcal{U}_{PI_{P,N-1}}} \mathrm{E}_W\Big[\int_{t_{N-1}}^{\infty}\Big\{\int_{t}^{\infty} \lambda e^{-\lambda(t_N - t_{N-1})}dt_N\Big\}(X'(t)CX(t) + U'(X(t_{N-1}))DU(X(t_{N-1})))dt \Big| X(t_{N-1})\Big] \tag{4.4}$$

$$V(X(t_{N-1}), N-1) = \min_{U_{N-1}\in\mathcal{U}_{PI_{P,N-1}}} \mathrm{E}_W\Big[\int_{t_{N-1}}^{\infty} e^{-\lambda(t - t_{N-1})}(X'(t)CX(t) + U'(X(t_{N-1}))DU(X(t_{N-1})))dt \Big| X(t_{N-1})\Big]. \tag{4.5}$$

We recognize (4.5) to be an infinite horizon discounted linear quadratic regulator problem with initial state $X(t_{N-1})$ known.

We can write (4.5) as:

$$V(X(t_{N-1}), N-1) = \min_{U_{N-1}\in\mathcal{U}_{PI_{P,N-1}}} \Big\{\mathrm{E}_W\Big[\int_{t_{N-1}}^{\infty} \mathrm{tr}[e^{-\lambda(t - t_{N-1})}CX(t)X'(t)]dt \Big| X(t_{N-1})\Big] + \lambda^{-1}U'(X(t_{N-1}))DU(X(t_{N-1}))\Big\}. \tag{4.6}$$

To evaluate the expectation in (4.6), we proceed as follows:

- Let $P(t)$ be the covariance matrix given by

$$P(t) = \mathrm{E}(X(t) - \mathrm{E}X(t))(X(t) - \mathrm{E}X(t))' \tag{4.7}$$

from which we have

$$P(t) + \mathrm{E}X(t)\mathrm{E}X'(t) = \mathrm{E}X(t)X'(t) \tag{4.8}$$

-

$$\begin{aligned} &\mathrm{E}_W\Big[\int_{t_{N-1}}^{\infty} \mathrm{tr}[e^{-\lambda(t-t_{N-1})}CX(t)X'(t)]dt \,|\, X(t_{N-1})\Big] \\ &= \int_{t_{N-1}}^{\infty} \mathrm{tr}[e^{-\lambda(t-t_{N-1})}CP(t)]dt + \int_{t_{N-1}}^{\infty} \mathrm{tr}[e^{-\lambda(t-t_{N-1})}C\mathrm{E}X(t)\mathrm{E}X'(t)]dt. \end{aligned} \tag{4.9}$$

- To compute the first integral in (4.9), we rely on the following differential equation (Gelb [7])

$$\dot{P}(\bar{t}) = AP(\bar{t}) + P(\bar{t})A' + GG', \qquad P(t_{N-1}) = 0. \tag{4.10}$$

where $\bar{t} = t - t_{N-1}$. Then:

$$\int_0^{\infty} e^{-\lambda\bar{t}}\dot{P}(\bar{t})d\bar{t} = \int_0^{\infty} e^{-\lambda\bar{t}}(AP(\bar{t}) + P(\bar{t})A' + GG')d\bar{t}. \tag{4.11}$$

Hence:

$$\begin{aligned} \int_0^{\infty} e^{-\lambda\bar{t}}\dot{P}(\bar{t})d\bar{t} &= sp(s)\ |_{s=\lambda} - P(t_{N-1}) \\ &= \int_0^{\infty} e^{-\lambda t}AP(t)dt + \int_0^{\infty} e^{-\lambda t}P(t)A'dt + \int_0^{\infty} e^{-\lambda t}GG'dt \qquad (4.12) \\ &= Ap(s)\ |_{s=\lambda} + p(s)A'\ |_{s=\lambda} + \frac{GG'}{s}\ |_{s=\lambda}. \qquad (4.13) \end{aligned}$$

By setting $s = \lambda$, we have:

$$(A - \frac{\lambda I}{2})p(\lambda) + p(\lambda)(A' - \frac{\lambda I}{2}) = -\lambda^{-1}GG' \tag{4.14}$$

where $p(\lambda)$ is given (Lancaster & Tismenetsky [8]) by:

$$p(\lambda) = \int_0^{\infty} e^{(A-\frac{\lambda I}{2})\tau}\lambda^{-1}GG'e^{(A'-\frac{\lambda I}{2})\tau}d\tau. \tag{4.15}$$

Therefore the first integral in (4.9) can now be expressed by:

$$\int_{t_{N-1}}^{\infty} \mathrm{tr}[e^{-\lambda(t-t_{N-1})}CP(t)]dt = \int_0^{\infty} \mathrm{tr}[e^{-\lambda\bar{t}}CP(\bar{t})]d\bar{t} = \mathrm{tr}[Cp(\lambda)]. \tag{4.16}$$

- To evaluate the second integral in (4.9), we rely on the following differential equation for $d\mathrm{E}X(t)$ and its solution:

$$d\mathrm{E}X(t) = (A\mathrm{E}X(t) + BU(X(t_{N-1})))dt \tag{4.17}$$

$$\mathrm{E}X(t) = \phi(t, t_{N-1})X(t_{N-1}) - Q(t, t_{N-1})U(X(t_{N-1})) \tag{4.18}$$

where

$$\begin{cases} \phi(t, t_{N-1}) = e^{A(t-t_{N-1})} \\ Q(t, t_{N-1}) = (I - \phi(t, t_{N-1}))A^{-1}B. \end{cases}$$

After appropriate computations, we obtain:

$$\begin{aligned} &\int_0^\infty \mathrm{tr}[e^{-\lambda(t-t_{N-1})}C\mathrm{E}X(t)\mathrm{E}X'(t)]dt \\ &= \int_0^\infty (X'(t_{N-1})e^{-\lambda t}\phi'(t)C\phi(t)X(t_{N-1}) - U'(X(t_{N-1}))e^{-\lambda t}Q'(t)C\phi(t)X(t_{N-1}) \\ &\quad - X'(t_{N-1})e^{-\lambda t}\phi'(t)CQ(t)U(X(t_{N-1})) \\ &\quad + U'(X(t_{N-1}))e^{-\lambda t}Q'(t)CQ(t)U(X(t_{N-1})))dt \end{aligned} \tag{4.19}$$

$$\begin{aligned} &= X'(t_{N-1})F_{\phi',C,\phi}X(t_{N-1}) - U'(X(t_{N-1}))F_{Q',C,\phi}X(t_{N-1}) \\ &\quad - X'(t_{N-1})F_{\phi',C,Q}U(X(t_{N-1})) + U'(X(t_{N-1}))F_{Q',C,Q}U(X(t_{N-1})) \end{aligned} \tag{4.20}$$

where

$$\begin{cases} F_{\phi',C,\phi} = \int_0^\infty e^{-\lambda t}\phi'(t)C\phi(t)dt \\ F_{Q',C\phi} = \int_0^\infty e^{-\lambda t}Q'(t)C\phi(t)dt \\ F_{\phi',C,Q} = \int_0^\infty e^{-\lambda t}\phi'(t)CQ(t)dt \\ F_{Q',C,Q} = \int_0^\infty e^{-\lambda t}Q'(t)CQ(t)dt. \end{cases}$$

We can now express $V(X(t_{N-1}), N-1)$ as:

$$\begin{aligned} V(X(t_{N-1}), N-1) = \min_{U_{N-1}\in\mathcal{U}_{PI_P,N-1}} &\{\mathrm{tr}[C_P(\lambda)] + X'(t_{N-1})F_{\phi',C,\phi}X(t_{N-1}) \\ &- U'(X(t_{N-1}))F_{Q',C,\phi}X(t_{N-1}) - X'(t_{N-1})F_{\phi',C,Q}U(X(t_{N-1})) \\ &+ U'(X(t_{N-1}))F_{Q',C,Q}U(X(t_{N-1})) + \lambda^{-1}U'(X(t_{N-1}))DU(X(t_{N-1}))\}. \end{aligned} \tag{4.21}$$

After some algebra, we obtain:

$$U^*(X(t_{N-1})) = (F_{Q',CQ} + \lambda^{-1}D)^{-1}F_{Q',C,\phi}X(t_{N-1}) \tag{4.22a}$$

$$= -L_{N-1}X(t_{N-1}) \tag{4.22b}$$

where L_{N-1} is explicitly given by:

$$L_{N-1} = \left(\int_0^\infty e^{-\lambda t} B'(A^{-1})'(I - e^{A't})C(I - e^{At})A^{-1}B dt + \lambda^{-1}D\right)^{-1}.$$
$$\left(\int_0^\infty e^{-\lambda t} B'(A^{-1})'(e^{A't} - I)Ce^{At} dt\right). \qquad (4.23)$$

At the $(N-1)$th stage, the optimal cost-to-go expression is given by:

$$V(X(t_{N-1}), N-1) = X'(t_{N-1})\widetilde{L}_{N-1}X(t_{N-1}) + \widehat{\ell}_{N-1} \qquad (4.24)$$

where

$$\begin{cases} \widetilde{L}_{N-1} = F_{\Phi',C,\Phi} + L'_{N-1}F_{Q',C,\Phi} + F_{\Phi',C,Q}L_{N-1} \\ \qquad + L'_{N-1}F_{Q',C,Q}L_{N-1} + \lambda^{-1}L'_{N-1}DL_{N-1} \\ \widehat{\ell}_{N-1} = \mathrm{tr}[Cp(\lambda)]. \end{cases}$$

Here the optimal cost-to-go expression is quadratic in its initial state $X(t_{N-1})$. We postulate then, that the structure of the optimal cost will remain quadratic in its initial state. We shall prove our hypotheses using backwards induction. Thus let:

$$V(X(t_{k+1}), k+1) = X'(t_{k+1})\widetilde{L}_{k+1}X(t_{k+1}) + \widehat{\ell}_{k+1} \qquad (4.25)$$

where $0 \leq k < N, \widetilde{L}_{k+1}$ and $\widehat{\ell}_{k+1}$ are some specific constants.

By the Principle of Optimality in the dynamic programming framework, we have:

$$V(X(t_k), k) = \min_{U_k \epsilon \mathcal{U}_{PI_{P,k}}} \mathrm{E}_W[\mathrm{E}_{t_{k+1}}\{\int_{t_k}^{t_{k+1}} (X'(t)CX(t)$$
$$+ U'(X(t_k))DU(X(t_k)))dt + V(X(t_{k+1}), k+1)\}|X(t_k)] \qquad (4.26)$$

where $U_k \equiv U(X(t_k)) \,\epsilon\, \mathcal{U}_{PI_{P,k}}$.

As for the $(N-1)$th stage, we can show for the present stage that:

$$\mathrm{E}_W[\mathrm{E}_{t_{k+1}}\{\int_{t_k}^{t_{k+1}} (X'(t)CX(t) + U'(X(t_k))DU(X(t_k)))dt|X(t_k)\}]$$
$$= \mathrm{E}_W[\int_{t_k}^\infty e^{-\lambda(t-t_k)}(X'(t)CX(t) + U'(X(t_k))DU(X(t_k)))dt|X(t_k)]. \qquad (4.27)$$

Now as $(t_{k+1} - t_k)$ is exponentially distributed in its parameter λ, and upon conditioning on t_k, we can write:

$$\begin{aligned}
&\mathrm{E}_W[\mathrm{E}_{t_{k+1}}\{X'(t_{k+1})\widetilde{L}_{k+1}X(t_{k+1})|X(t_k)\}] \\
&= \mathrm{E}_W\Big[\int_{t_k}^{\infty} X'(t_{k+1})\widetilde{L}_{k+1}X(t_{k+1})\lambda e^{-\lambda(t_{k+1}-t_k)}dt_{k+1}|X(t_k)\Big] \\
&= \mathrm{E}_W\Big[\int_{t_k}^{\infty} \lambda e^{-\lambda(t-t_k)}X'(t)\widetilde{L}_{k+1}X(t)dt|X(t_k)\Big]. \qquad (4.28)
\end{aligned}$$

Therefore (4.26) can be expressed as:

$$\begin{aligned}
V(X(t_k),k) = \min_{U_k \in \mathcal{U}_{PI_{P,k}}} &\mathrm{E}_W\Big[\int_{t_k}^{\infty} e^{-\lambda(t-t_k)}(X'(t)(C+\lambda\widetilde{L}_{k+1})X(t) \\
&+ U'(X(t_k))DU(X(t_k)))dt|X(t_k)\Big] + \widehat{\ell}_{k+1} \qquad (4.29)
\end{aligned}$$

$$\begin{aligned}
V(X(t_k),k) = \min_{U_k \in \mathcal{U}_{PI_{P,k}}} &\{X'(t_k)F_{\phi',\widetilde{C},\phi}X(t_k) - U'(X(t_k))F_{Q',\widetilde{C},\phi}X(t_k) \\
&- X'(t_k)F_{\phi',\widetilde{C},Q}U(X(t_k)) + U'(X(t_k))F_{Q',\widetilde{C},Q}U(X(t_k)) \\
&+ \lambda^{-1}U'(X(t_k))DU(X(t_k))\} + \mathrm{tr}[\widetilde{C}p(\lambda)] + \widehat{\ell}_{k+1} \qquad (4.30)
\end{aligned}$$

where

$$\begin{cases}
\widetilde{C} = C + \lambda\widetilde{L}_{k+1} \\
F_{\phi',\widetilde{C},\phi} = \int_0^\infty e^{-\lambda t}\phi'(t)\widetilde{C}\phi(t)dt \\
F_{Q',\widetilde{C}\phi} = \int_0^\infty e^{-\lambda t}Q'(t)\widetilde{C}\phi(t)dt \\
F_{\phi',\widetilde{C},Q} = \int_0^\infty e^{-\lambda t}\phi'(t)\widetilde{C}Q(t)dt \\
F_{Q',\widetilde{C},Q} = \int_0^\infty e^{-\lambda t}Q'(t)\widetilde{C}Q(t)dt.
\end{cases}$$

Note that $\widetilde{L}_{N-1}$ is symmetric, indeed:

$$\begin{aligned}
\widetilde{L}'_{N-1} = F_{\phi',C,\phi} + F_{\phi',C,Q}L_{N-1} + L'_{N-1}F_{Q',C,\phi} \\
+ L'_{N-1}F_{Q',C,Q}L_{N-1} + \lambda^{-1}L'_{N-1}DL_{N-1} = \widetilde{L}_{N-1}.
\end{aligned}$$

Using induction principle, we can show that $\widetilde{L}_{k+1}$ is symmetric.

Therefore, it follows from (4.30):

$$U^*(X(t_k)) = (F_{Q',\widetilde{C},Q} + \lambda^{-1}D)^{-1}F_{Q',\widetilde{C},\phi}X(t_k) \qquad (4.31a)$$

$$= -L_kX(t_k) \qquad (4.31b)$$

where L_k is given by:

$$L_k = \Big(\int_0^\infty e^{-\lambda t} B'(A^{-1})'(I - e^{A't})\bar{C}(I - e^{At})A^{-1}B dt + \lambda^{-1}D\Big)^{-1}.$$
$$\Big(\int_0^\infty e^{-\lambda t} B'(A^{-1})'(e^{A't} - I)\bar{C}e^{At} dt\Big). \tag{4.32}$$

At the k-th stage, the optimal cost-to-go expression is given by:

$$V(X(t_k), k) = X'(t_k)\widetilde{L}_k X(t_k) + \widehat{\ell}_k \tag{4.33}$$

where

$$\begin{aligned}
\widetilde{L}_k &= F_{\phi',\bar{C},\phi} + L_k' F_{Q',\bar{C},\phi} + F_{\phi',\bar{C},Q} L_k \\
&\quad + L_k' F_{Q',\bar{C},Q} L_k + \lambda^{-1} L_k' D L_k \\
\widehat{\ell}_k &= \mathrm{tr}[(C + \lambda \widetilde{L}_{k+1})p(\lambda)] + \widehat{\ell}_{k+1} \\
\widetilde{L}_N &= \widehat{\ell}_N = 0.
\end{aligned}$$

we establish with equation (4.33) our main result.

4.1. Scalar Example. We consider here the following scalar stochastic differential equation:

$$dx(t) = ax(t)dt + bu(\lfloor x \rfloor_t)dt + gdw(t) \tag{4.34}$$

with the associated cost functional

$$J(x(t_0)) = \min_{u \in \mathcal{U}_{PI_P}} \mathrm{E}_W\Big[\sum_{i=0}^{N-1} \mathrm{E}_{t_{i+1}}\Big\{\int_{t_i}^{t_{i+1}} (cx^2(t) + du^2(\lfloor x \rfloor_t))dt | x(t_i)\Big\}\Big] \tag{4.35}$$

where $u(\lfloor x \rfloor_t) \equiv u(x(t_i)) \in \mathcal{U}_{PI_P}$ for $t_i \leq t < t_{i+1}$.

Following the same general framework developed through equations (4.1) to (4.33), we have:

$$V(X(t_N), N) = 0. \tag{4.36}$$

At the $(N-1)$th stage, the optimal cost-to-go is

$$V(x(t_{N-1}), N-1) = \min_{u_{N-1} \in \mathcal{U}_{PI_{P,N-1}}} \Big\{\mathrm{E}_W\Big[\int_{t_{N-1}}^\infty e^{-\lambda(t-t_{N-1})} cx^2(t)dt | x(t_{N-1})\Big]$$
$$+ \lambda^{-1} du^2(x(t_{N-1}))\Big\}. \tag{4.37}$$

Using equations (4.7) to (4.20) and after some algebra, we obtain in our scalar context:

$$\begin{cases} \mathrm{tr}[Cp(\lambda)] = cg^2(\lambda(\lambda-2a))^{-1} \\ F_{\phi',C,\phi} = c(\lambda-2a)^{-1} \\ F_{Q',C,\phi} = F_{\phi',C,Q} = -bc((\lambda-a)(\lambda-2a))^{-1} \\ F_{Q',C,Q} = 2b^2c(\lambda(\lambda-a)(\lambda-2a))^{-1}. \end{cases}$$

Thus, we have:

$$\begin{aligned} V(x(t_{N-1}),N-1) = \min_{u_{N-1}\epsilon\mathcal{U}_{PI_{P,N-1}}} & \{cg^2(\lambda(\lambda-2a))^{-1} + c(\lambda-2a)^{-1}x^2(t_{N-1}) \\ & + 2bc((\lambda-a)(\lambda-2a))^{-1}u(x(t_{N-1}))x(t_{N-1}) \\ & + 2b^2c(\lambda(\lambda-a)(\lambda-2a))^{-1}u^2(x(t_{N-1})) + \lambda^{-1}du^2(x(t_{N-1}))\} \end{aligned} \quad (4.38)$$

where $u_{N-1} \equiv U(x(t_{N-1}))\,\epsilon\,\mathcal{U}_{PI_{P,N-1}}$.

$$u^*(x(t_{N-1})) = -L_{N-1}x(t_{N-1}) \quad (4.39)$$

where

$$L_{N-1} = bc\lambda(2b^2c + d(\lambda-a)(\lambda-2a))^{-1}. \quad (4.40)$$

The resulting optimal cost-to-go at the $(N-1)$th stage is expressed by:

$$V(x(t_{N-1}),N-1) = \widetilde{L}_{N-1}x^2(t_{N-1}) + \widehat{\ell}_{N-1} \quad (4.41)$$

where

$$\begin{aligned} \widetilde{L}_{N-1} &= c(\lambda-2a)^{-1} + 2bcL_{N-1}((a-\lambda)(\lambda-2a))^{-1} \\ &\quad + L_{N-1}^2(2b^2c(\lambda(\lambda-a)(\lambda-2a))^{-1} + \lambda^{-1}d) \\ \widehat{\ell}_{N-1} &= cg^2(\lambda(\lambda-2a))^{-1}. \end{aligned}$$

Applying equations (4.30) to (4.33) in our present scalar study, we can compute explicitly for an arbitrary stage, $\widetilde{L}_k$ and $\widehat{\ell}_k$ in

$$V(x(t_k),k) = \widetilde{L}_k x^2(t_k) + \widehat{\ell}_k. \quad (4.42)$$

The last equation means that the structure of the optimal cost-to-go remains quadratic from stage to another. We write now from equations (4.30) to (4.33) the main scalar steps:

- $$u^*(x(t_k)) = -L_k x(t_k) \tag{4.43}$$

- $$L_k = b\bar{c}\lambda(2b^2\bar{c} + d(\lambda - a)(\lambda - 2a))^{-1}. \tag{4.44}$$

- The resulting optimal cost-to-go is expressed by

$$V(x(t_k), k) = \widetilde{L}_k x^2(t_k) + \widehat{\ell}_k \tag{4.45}$$

where

$$\widetilde{L}_k = \bar{c}(\lambda - 2a)^{-1} + 2b\bar{c}L_k((a-\lambda)(\lambda - 2a))^{-1}$$
$$+ L_k^2(2b^2\bar{c}(\lambda(\lambda - a)(\lambda - 2a))^{-1} + \lambda^{-1}d)$$
$$\widehat{\ell}_k = \bar{c}g^2(\lambda(\lambda - 2a))^{-1} + \ell_{k+1}$$
$$\bar{c} = c + \lambda\widetilde{L}_{k+1}$$

$\widetilde{L}_{k+1}$ and $\widehat{\ell}_{k+1}$ are some constants specific to the stage $k+1$

$$\widetilde{L}_N = \widehat{\ell}_N = 0.$$

5. Optimal Time Variant Control

We consider in this section the expected quadratic cost functional and the stochastic differential equation given above in (2.6) - (2.8).

As we mentioned earlier in this chapter, for the time variant control, the t_i's are Poisson distributed with mean inter-arrival time $\mu_k = 1/\lambda$ for $0 \leq k < N$. As previously, we have:

$$V(X(t_N), N) = 0. \tag{5.1}$$

At the $(N-1)$th stage, the optimal cost-to-go is expressed by:

$$V(X(t_{N-1}), N-1) = \min_{U_{N-1} \epsilon \mathcal{U}_{TV_{P,N-1}}} \mathrm{E}_W \big[\mathrm{E}_{t_N} \{ \int_{N-1}^{t_N} (X'(t)CX(t) + U'(t, X(t_{N-1}))DU(t, X(t_{N-1})))dt | X(t_{N-1})\} \big] \tag{5.2}$$

$$= \min_{U_{N-1} \epsilon \mathcal{U}_{TV_{P,N-1}}} \mathrm{E}_W \big[\int_{t_{N-1}}^{\infty} e^{-\lambda(t - t_{N-1})} (X'(t)CX(t) + U'(t, X(t_{N-1}))DU(t, X(t_{N-1})))dt | X(t_{N-1}) \big] \tag{5.3}$$

where $U_{N-1} \equiv U(t, X(t_{N-1})) \, \epsilon \, \mathcal{U}_{TV_{P,N-1}}$.

We recognize (5.3) to be an infinite horizon discounted linear quadratic regulator problem with the initial state $X(t_{N-1})$ known and no further observations thereafter. This last fact is very crucial for our present time variant control case, contrary to the time invariant control case, where the control law is function only of the initial state on each interval.

Using Rishel's [9] result for (5.3), the control law is given for the moment by:

$$U^*(t, X(t)) = -D^{-1}B'K_{N-1}X(t) \tag{5.4}$$

where the matrix gain K_{N-1} is given below in equation (5.8).

However, as no further observations of the state $X(t)$ are available on the interval $[t_{N-1}, t_N)$, then by the separation principle (Bagchi [1], Fleming & Rishel [5]), we replace $X(t)$ by its optimal predictor $\widehat{X}(t|t_{N-1})$ with the initial state $X(t_{N-1})$ known.

Then the optimal control law is now expressed by:

$$U^*(t, \widehat{X}(t|t_{N-1})) = -D^{-1}B'K_{N-1}\widehat{X}(t|t_{N-1}). \tag{5.5}$$

This optimal filtering estimate of $X(t)$ is governed by the following dynamics:

$$d\widehat{X}(t|t_{N-1}) = (A - BD^{-1}B'K_{N-1})\widehat{X}(t|t_{N-1})dt \tag{5.6}$$

$$\widehat{X}(t|t_{N-1}) = e^{M_{N-1}(t-t_{N-1})}X(t_{N-1}) \tag{5.7}$$

where $M_{N-1} = A - BD^{-1}B'K_{N-1}$ and $\widehat{X}(t_{N-1}|t_{N-1}) = X(t_{N-1})$.

As we know, the matrix gain K_{N-1} is a solution of the Algebraic Riccati equation:

$$-K_{N-1}A - A'K_{N-1} - \lambda K_{N-1} + K_{N-1}BD^{-1}B'K_{N-1} - C = 0. \tag{5.8}$$

Thus, we have a closed-loop system with the following dynamics:

$$dX(t) = AX(t)dt - BD^{-1}B'K_{N-1}\widehat{X}(t|t_{N-1})dt + GdW(t) \tag{5.9}$$

$$d\widehat{X}(t|t_{N-1}) = (A - BD^{-1}B'K_{N-1})\widehat{X}(t|t_{N-1})dt \tag{5.10}$$

with the resulting cost functional expressed by:

$$\begin{aligned} V(X(t_{N-1}), N-1) = \mathrm{E}_W\Big[\int_{t_{N-1}}^{\infty} \mathrm{tr}\big[e^{-\lambda(t-t_{N-1})}(CX(t)X'(t) \\ + \widehat{M}_{N-1}\widehat{X}(t|t_{N-1})\widehat{X}'(t|t_{N-1}))\big]dt|X(t_{N-1})\Big] \end{aligned} \tag{5.11}$$

where $\widehat{M}_{N-1} = K'_{N-1}BD^{-1}B'K_{N-1}$.

We compute now the first integral in (5.11); for this end we rely on the same approach elaborated in the previous section, and we note that for this closed-loop system, we have obviously $\mathrm{E}X(t) = \widehat{X}(t|t_{N-1})$.

Therefore:

$$\int_{t_{N-1}}^{\infty} \mathrm{tr}[e^{-\lambda(t-t_{N-1})}C\mathrm{E}X(t)\mathrm{E}X'(t)]dt$$

$$= \int_{t_{N-1}}^{\infty} e^{-\lambda(t-t_{N-1})}X'(t_{N-1})e^{M'_{N-1}(t-t_{N-1})}Ce^{M_{N-1}(t-t_{N-1})}X(t_{N-1})dt$$

(where $X(t_{N-1})$ is already fixed)

$$= \int_{0}^{\infty} e^{-\lambda t}X'(t_{N-1})e^{M'_{N-1}t}Ce^{M_{N-1}t}X(t_{N-1})dt$$

$$= X'(t_{N-1})F_{M_{N-1}}X(t_{N-1}) \tag{5.12}$$

where $F_{M_{N-1}} = \int_0^\infty e^{-\lambda t}e^{M'_{N-1}t}Ce^{M_{N-1}t}dt$.

Finally, the last integral in (5.11) yields:

$$\int_0^\infty e^{-\lambda t}X'(t_{N-1})e^{M'_{N-1}t}\widehat{M}_{N-1}e^{M_{N-1}t}X(t_{N-1})dt = X'(t_{N-1})F_{\widehat{M}_{N-1}}X(t_{N-1}) \tag{5.13}$$

where $F_{\widehat{M}_{N-1}} = \int_0^\infty e^{-\lambda t}e^{M'_{N-1}t}\widehat{M}_{N-1}e^{M_{N-1}t}dt$. We know that $F = \int_0^\infty e^{A't}Ce^{At}dt$, is the solution of the algebraic Lyapunov matrix equation in the form $A'F + FA = -C$. For a detailed and systematic study of the well-known Lyapunov matrix equation, see Gajić and Qureshi [6].

At this stage, the optimal cost-to-go is expressed by

$$V(X(t_{N-1}, N-1) = X'(t_{N-1})(F_{M_{N-1}} + F_{\widehat{M}_{N-1}})X(t_{N-1})$$

$$+ \mathrm{tr}\Big[C\int_0^\infty e^{(A-\frac{\lambda I}{2})\tau}\lambda^{-1}GG'e^{(A'-\frac{\lambda I}{2})\tau}d\tau\Big] \tag{5.14}$$

$$V(X(t_{N-1}), N-1) = X'(t_{N-1})\widetilde{L}_{N-1}X(t_{N-1}) + \widehat{\ell}_{N-1} \tag{5.15}$$

where obviously $\widetilde{L}_{N-1} = F_{M_{N-1}} + F_{\widehat{M}_{N-1}}$ and $\widehat{\ell}_{N-1} = \mathrm{tr}[Cp(\lambda)]$.

Again the optimal cost is quadratic in its initial state $X(t_{N-1})$, we reformulate the hypotheses that the optimal cost value will remain quadratic in its initial state.

Thus, let:

$$V(X(t_{k+1}), k+1) = X'(t_{k+1})\widetilde{L}_{k+1}X(t_{k+1}) + \widehat{\ell}_{k+1} \tag{5.16}$$

where $0 \le k < N, \widetilde{L}_{k+1}$ and $\widehat{\ell}_{k+1}$ are some specific constants.

We have to prove the following equation using backwards induction

$$V(X(t_k), k) = X'(t_k)\widetilde{L}_kX(t_k) + \widehat{\ell}_k \tag{5.17}$$

for $0 \leq k < N$.

By the Principle of Optimality in the dynamic programming framework, we have:

$$V(X(t_k),k) = \min_{U_k \in \mathcal{U}_{TV_{P,k}}} \mathrm{E}_W\big[\mathrm{E}_{t_{k+1}}\{\int_{t_k}^{t_{k+1}} (X'(t)CX(t) + U'(t,X(t_k))DU(t,X(t_k)))dt + V(X(t_{k+1}),k+1)\}|X(t_k)\big] \quad (5.18)$$

where $U_k \equiv U(t,X(t_k)) \in \mathcal{U}_{TV_{P,k}}$.

On the one hand, using a similar approach than previously, we know that:

$$\mathrm{E}_W\big[\mathrm{E}_{t_{k+1}}\{\int_{t_k}^{t_{k+1}} (X'(t)CX(t) + U'(t,X(t_k))DU(t,X(t_k)))dt|X(t_k)\}\big] = \mathrm{E}_W\big[\int_{t_k}^{\infty} e^{-\lambda(t-t_k)}(X'(t)CX(t) + U'(t,X(t_k))DU(t,X(t_k)))dt|X(t_k)\big]. \quad (5.19)$$

On the other hand, as $(t_{k+1} - t_k)$ is exponentially distributed in its parameter λ, and upon conditioning on t_k, we can write:

$$\begin{aligned} &\mathrm{E}_W\big[\mathrm{E}_{t_{k+1}}\{X'(t_{k+1})\tilde{L}_{k+1}X(t_{k+1})|X(t_k)\}\big] \\ &= \mathrm{E}_W\big[\int_{t_k}^{\infty} X'(t_{k+1})\tilde{L}_{k+1}X(t_{k+1})\lambda e^{-\lambda(t_{k+1}-t_k)}dt_{k+1}|X(t_k)\big] \\ &= \mathrm{E}_W\big[\int_{t_k}^{\infty} \lambda e^{-\lambda(t-t_k)}X'(t)\tilde{L}_{k+1}X(t)dt|X(t_k)\big]. \end{aligned} \quad (5.20)$$

Therefore (5.18) can be written as:

$$V(X(t_k),k) = \min_{u_k \in \mathcal{U}_{TV_{P,k}}} \mathrm{E}_W\big[\int_{t_k}^{\infty} e^{-\lambda(t-t_k)}(X'(t)(C+\lambda\tilde{L}_{k+1})X(t) + U'(t,X(t_k))DU(t,X(t_k)))dt|X(t_k)\big] + \hat{\ell}_{k+1}. \quad (5.21)$$

Again, we recognize (5.21) to be an infinite horizon discounted linear quadratic regulator problem; then using Rishel's result and separation principle, we obtain:

$$U^*(t,\hat{X}(t|t_k)) = -D^{-1}B'K_k\hat{X}(t|t_k) \quad (5.22)$$

where K_k is the matrix solution of Riccati equation given by:

$$-K_kA - A'K_k - \lambda K_k + K_kBD^{-1}B'K_k - (C+\lambda\tilde{L}_{k+1}) = 0. \quad (5.23)$$

Then we have the following dynamics and cost functional:

$$dX(t) = (AX(t) - BD^{-1}B'K_k\widehat{X}(t|t_k))dt + GdW(t) \tag{5.24}$$

$$d\widehat{X}(t|t_k) = (A - BD^{-1}B'K_k)\widehat{X}(t|t_k)dt \tag{5.25}$$

$$\widehat{X}(t|t_k) = e^{M_k(t-t_k)}X(t_k) \tag{5.26}$$

where $M_k = A - BD^{-1}B'K_k$

$$\begin{aligned} V(X(t_k),k) = \mathrm{E}_W\Big[\int_{t_k}^{\infty} \mathrm{tr}\big[e^{-\lambda(t-t_k)}\big((C+\lambda\widetilde{L}_{k+1})X(t)X'(t) \\ + \widehat{M}_k\widehat{X}(t|t_k)\widehat{X}'(t|t_k)\big)\big]dt\,|X(t_k)\Big] + \widehat{\ell}_{k+1} \end{aligned} \tag{5.27}$$

where $\widehat{M}_k = K_k'BD^{-1}B'K_k$

$$\begin{aligned} V(X(t_k),k) &= X'(t_k)(F_{M_k} + F_{\widehat{M}_k})X(t_k) \\ &\quad + \mathrm{tr}\Big[(C+\lambda\widetilde{L}_{k+1})\int_0^{\infty} e^{(A-\frac{\lambda I}{2})\tau}\lambda^{-1}GG'e^{(A'-\frac{\lambda I}{2})\tau}d\tau\Big] + \widehat{\ell}_{k+1} \end{aligned} \tag{5.28}$$

$$V(X(t_k),k) = X'(t_k)\widetilde{L}_kX(t_k) + \widehat{\ell}_k \tag{5.29}$$

where

$$\begin{cases} \widetilde{L}_k = F_{M_k} + F_{\widehat{M}_k} \\ F_{M_k} = \int_0^{\infty} e^{-\lambda t}e^{M_k't}(C+\lambda\widetilde{L}_{k+1})e^{M_kt}dt \\ F_{\widehat{M}_k} = \int_0^{\infty} e^{-\lambda t}e^{M_k't}\widehat{M}_ke^{M_kt}dt \\ \widehat{\ell}_k = \mathrm{tr}[(C+\lambda\widetilde{L}_{k+1})p(\lambda)] + \widehat{\ell}_{k+1} \\ \widetilde{L}_N = \widehat{L}_N = 0. \end{cases}$$

We have then proved our main result.

5.1. Scalar Example. We consider the following scalar stochastic differential equation:

$$dx(t) = ax(t)dt + bu(t, \lfloor x \rfloor_t)dt + gdw(t) \tag{5.30}$$

with the associated cost functional

$$J(x(t_0)) = \min_{U \in \mathcal{U}_{TV_p}} \mathrm{E}_W\Big[\sum_{i=0}^{N-1} \mathrm{E}_{t_{i+1}}\Big\{\int_{t_i}^{t_{i+1}} (cx^2(t) + du^2(t, x(t_i)))dt\,|x(t_i)\Big\}\Big]. \tag{5.31}$$

Equations (5.30) and (5.31) are a scalar version of equations (2.8) and (2.7).

Following the general framework solution elaborated through equations (5.1) to (5.29), we have:

$$V(x(t_N), N) = 0. \tag{5.32}$$

At the $(N-1)$th stage, the optimal cost-to-go is

$$V(x(t_{N-1}), N-1) = \min_{u_{N-1} \epsilon \mathcal{U}_{TV_{P,N-1}}} \mathrm{E}_W \big[\mathrm{E}_{t_N}\big\{ \int_{t_{N-1}}^{t_N} (cx^2(t) + du^2(t, x(t_{N-1})))dt | x(t_{N-1})\big\}\big] \tag{5.33}$$

$$V(x(t_{N-1}), N-1) = \min_{u_{N-1} \epsilon \mathcal{U}_{TV_{P,N-1}}} \mathrm{E}_W \big[\int_{t_{N-1}}^{\infty} e^{-\lambda(t-t_{N-1})}(cx^2(t) + du^2(t, x(t_{N-1})))dt | x(t_{N-1})\big] \tag{5.34}$$

where $u_{N-1} \equiv u(t, x(t_{N-1})) \epsilon \mathcal{U}_{TV_{P,N-1}}$. Therefore

$$u^*(t, \widehat{x}(t|t_{N-1})) = -d^{-1}bk_{N-1}\widehat{x}(t|t_{N-1}) \tag{5.35}$$

where $\widehat{x}(t|t_{N-1})$ is the optimal predictor estimate of $x(t)$ according to:

$$d\widehat{x}(t|t_{N-1}) = M_{N-1}\widehat{x}(t|t_{N-1})dt \tag{5.36}$$

$$\widehat{x}(t|t_{N-1}) = e^{M_{N-1}(t-t_{N-1})}x(t_{N-1}) \tag{5.37}$$

where $M_{N-1} = a - b^2d^{-1}k_{N-1}$.

The gain k_{N-1} is a solution of the algebraic Riccati equation:

$$b^2d^{-1}k_{N-1}^2 - (2a+\lambda)k_{N-1} - c = 0. \tag{5.38}$$

Thus, we have a closed-loop system with the following dynamics:

$$dx(t) = (ax(t) - b^2d^{-1}k_{N-1}\widehat{x}(t|t_{N-1}))dt + gdw(t) \tag{5.39}$$

$$d\widehat{x}(t|t_{N-1}) = M_{N-1}\widehat{x}(t|t_{N-1})dt \tag{5.40}$$

with the resulting optimal cost-to-go is given by:

$$V(x(t_{N-1}), N-1) = \mathrm{E}_W\big[\int_{t_{N-1}}^{\infty} e^{-\lambda(t-t_{N-1})}(cx^2(t) + \widehat{M}_{N-1}\widehat{x}^2(t|t_{N-1}))dt | x(t_{N-1})\big] \tag{5.41}$$

where $\widehat{M}_{N-1} = k_{N-1}^2b^2d^{-1}$.

Using equations (5.12) - (5.14), we can compute easily the optimal cost expressed explicitly by:

$$V(x(t_{N-1}), N-1) = (\lambda - 2(a - b^2d^{-1}k_{N-1}))^{-1}(c + k_{N-1}^2b^2d^{-1})x^2(t_{N-1}) + cg^2(\lambda(\lambda-2a))^{-1} \tag{5.42}$$

which is of the form

$$V(x(t_{N-1}), N-1) = \widetilde{L}_{N-1}x^2(t_{N-1}) + \widehat{\ell}_{N-1}. \tag{5.43}$$

Applying equations (5.18) to (5.28), we can compute explicitly, for an arbitrary stage, $\widetilde{L}_k$ and $\widehat{\ell}_k$ in:

$$V(x(t_k), k) = \widetilde{L}_k x^2(t_k) + \widehat{\ell}_k \tag{5.44}$$

which is a scalar version of equation (5.17). This means that the structure of the optimal cost remains quadratic in its initial state from stage to another. Let us write from these equations ((5.18) - (5.28)), the major scalar steps:

- $$\begin{aligned} V(x(t_k), k) = \min_{u_k \epsilon \mathcal{U}_{TV_{P,k}}} \{ \mathrm{E}_W \Big[\int_{t_k}^{\infty} e^{-\lambda(t-t_k)} ((c+\lambda\widetilde{L}_{k+1})x^2(t) \\ + du^2(t, x(t_k))) dt | x(t_k) \Big] \} + \widehat{\ell}_{k+1} \end{aligned} \tag{5.45}$$

  where $u_k \equiv u(t, x(t_k)) \, \epsilon \, \mathcal{U}_{TV_{P,k}}, \widetilde{L}_{k+1}$ and $\widehat{\ell}_{k+1}$ are some given constants.

- $$u^*(t, \widehat{x}(t|t_k)) = -d^{-1}bk_k\widehat{x}(t|t_k) \tag{5.46}$$

  where k_k is the solution of Riccati equation

  $$b^2d^{-1}k_k^2 - (2a+\lambda)k_k - (c+\lambda\widetilde{L}_{k+1}) = 0. \tag{5.47}$$

- We have the following dynamics and cost functional:

  $$dx(t) = (ax(t) - b^2d^{-1}k_k\widehat{x}(t|t_k))dt + gdw(t) \tag{5.48}$$

  $$d\widehat{x}(t|t_{N-1}) = M_k\widehat{x}(t|t_k)dt \tag{5.49}$$

  $$M_k = a - b^2d^{-1}k_k$$

  $$\begin{aligned} V(x(t_k), k) = \mathrm{E}_W \Big[\int_{t_k}^{\infty} e^{-\lambda(t-t_k)} ((c+\lambda\widetilde{L}_{k+1})x^2(t) \\ + \widehat{M}_k\widehat{x}^2(t|t_k)) dt | x(t_k) \Big] + \widehat{\ell}_{k+1} \end{aligned} \tag{5.50}$$

  where $\widehat{M}_k = k_k^2b^2d^{-1}$.

- After appropriate computations, we obtain

$$V(x(t_k),k) = \left(\lambda - 2(a - b^2 d^{-1} k_k)\right)^{-1}(c + \lambda \widetilde{L}_{k+1} + k_k^2 b^2 d^{-1}) x^2(t_k) + (c + \lambda \widetilde{L}_{k+1}) g^2 (\lambda(\lambda - 2a))^{-1} + \widehat{\ell}_{k+1} \tag{5.51}$$

$$V(x(t_k),k) = \widetilde{L}_k x^2(t_k) + \widehat{\ell}_k \tag{5.52}$$

with $\widetilde{L}_N = \widetilde{\ell}_N = 0$.

6. Conclusion

The class of problems studied in this chapter are open to generalization to problems which appear to be significantly more difficult; for instance, we consider the case where:

(a) The total number of observations N is random and unknown to the controller.
(b) One has the option to pay for a fixed sum in order to obtain the value of an observation occurring at a random time, and where the total performance cost is the LQG cost given below, plus the total observation cost payed over the time horizon of the problem.
(c) The stochastic differential equation described in (2.1) has random and time variant coefficients.
(d) The observations instants are distributed according to another probabilistic distribution.

7. Appendix

$$J(X(t_0) = \min_{U \in \mathcal{U}_{PI}} \mathrm{E}_{W,t}\Big[\int_{t_0}^{t_N} \ell(X(t),U(t))dt\Big] \tag{7.1}$$

$$= \min_{U \in \mathcal{U}_{PI}} \mathrm{E}_{W,t}\Big[\sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \ell(X(t),U(t))dt\Big] \tag{7.2}$$

$$= \min_{U \in \mathcal{U}_{PI}} \mathrm{E}_{W}\Big[\sum_{i=0}^{N-1} \mathrm{E}_t\Big\{\int_{t_i}^{t_{i+1}} \ell(X(t),U(t))dt | X(t_0),\dots,X(t_i)\Big\}\Big] \tag{7.3}$$

$$= \min_{U \in \mathcal{U}_{PI}} \mathrm{E}_{W}\Big[\sum_{i=0}^{N-1} \mathrm{E}_t\Big\{\int_{t_i}^{t_{i+1}} \ell(X(t),U(t))dt | X(t_i)\Big\}\Big] \tag{7.4}$$

$$= \min_{U \in \mathcal{U}_{PI}} \mathrm{E}_{W}\Big[\sum_{i=0}^{N-1} \mathrm{E}_{t_{i+1}}\Big\{\int_{t_i}^{t_{i+1}} \ell(X(t),U(t))dt | X(t_i)\Big\}\Big]. \tag{7.5}$$

Concerning equations (7.1) to (7.5), note that:

(7.1) to (7.2) is justified by the decomposition of the continuous time domain into a finite set of N disjoint intervals $[t_i, t_{i+1})$.
(7.3) is a consequence of smoothing property of conditional expectation.
(7.4) relies on the fact that the stochastic process $\{X(t_i), i \geq 0\}$ is a Markov process.

References

[1] Bagchi, A., "Optimal Control of Stochastic Systems", Prentice-Hall, New York, 1993.

[2] Caines, P.E., "Linear Stochastic Systems", John-Wiley & Sons, New York, 1988.

[3] Davis, M.H.A. and Vinter, R.B., "Stochastic Modelling and Control", Chapman and Hall, London, 1985.

[4] Feller, W., "An Introduction to Probability Theory and its Applications", Vol.II. Second edition, John-Wiley & Sons, New York, 1971.

[5] Fleming, W.H. and Rishel, R.W., "Deterministic and Stochastic Optimal Control", Springer-Verlag, New York, Third Printing, 1986.

[6] Gajić, Z. and Qureshi, M.T.J., "Lyapunov Matrix Equation in System Stability and Control", Academic Press, New York, 1995.

[7] Gelb, A., "Applied Optimal Estimation", The M.I.T. Press, Massachusetts, 1994.

[8] Lancaster, P. and Tismenetsky, M., "The Theory of Matrices", Second Edition, Academic Press, New York, 1985.

[9] Rishel, R., "Controlled Continuous Time Markov Processes", in Stochastic Models, edited by Heyman, D.P. and Sobel, M.J., North-Holland, Amsterdam, 1990.

[10] Stam, A.J., "Derived Stochastic Processes", Compositio Mathematica, Vol.17, 102-140, 1965.

CHAPTER 6

Gradient Estimation for Ratios

1. Introduction

Let (A, B) be a pair of jointly distributed real-valued random variables. The estimation of the ratio $\alpha = \mathrm{E}[A]/\mathrm{E}[B]$ is known, in the simulation literature, as the *ratio estimation problem*. Such ratio estimation problems arise in many different applications settings. For example, it is well known that the steady-state mean of a positive recurrent regenerative stochastic process can be expressed as such a ratio of expectations; see, for example, Section 3.32 of [2], or Chapter 2 of [15]. In Section 2 of this chapter, we will discuss the ratio estimation problem in greater detail and offer additional examples. It will turn out that the infinite-horizon discounted cost of a non-delayed regenerative process can also be expressed in terms of an appropriately chosen ratio estimation problem. This fact was first pointed out by [3].

Recently, the simulation community has devoted a great deal of attention to the use of simulation as an optimization tool. An important component of this research effort has been the development of estimation methodology for computing the gradient of a real-valued performance measure with respect to a (finite-dimensional) decision parameter vector. Such gradients play an important role in many iterative algorithms for performing both constrained and unconstrained mathematical optimization. This chapter is intended as a study of the question of how to use this gradient estimation methodology in the setting of the ratio estimation problem.

The chapter is organized as follows. In Section 2, a number of different applications in which ratio estimation problems arise are discussed, and the mathematical framework for the remainder of the chapter is described. Section 3 is devoted to deriving a confidence interval methodology for estimating the partial derivative of a ratio. In addition, a joint central-limit theorem for the simultaneous estimation of the entire gradient is obtained. In Section 4, low-bias estimation issues are discussed. Section 5 concludes the chapter with a brief summary. The proof of our main theorem (Theorem 1) is given in the Appendix.

2. Examples of Ratio Estimation Problems

As discussed in the introduction, the ratio estimation problem is concerned with the estimation of the ratio

$$\alpha = \frac{\mathrm{E}[A]}{\mathrm{E}[B]},$$

where (A, B) is a pair of jointly distributed real-valued random variables. We now proceed to offer several examples of this estimation problem.

Example 1. Let $X = \{X(t), t \geq 0\}$ be a real-valued (possibly) delayed regenerative process with regenerative times $0 \leq T(0) < T(1) < \cdots$. For $i \geq 1$, let

$$\widetilde{A}_i = \int_{T(i-1)}^{T(i)} |X(s)|ds$$
$$A_i = \int_{T(i-1)}^{T(i)} X(s)ds$$
$$B_i = T(i) - T(i-1).$$

If $\mathrm{E}[\widetilde{A}_1 + B_1] < \infty$, then it can be shown (see, for example, [1], or [15]) that

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t X(s)ds \overset{a.s.}{=} \alpha = \mathrm{E}[A_1]/\mathrm{E}[B_1].$$

Hence, as discussed in the introduction, the steady-state mean of such a process can be expressed as the ratio of the two expectations $\mathrm{E}[A_1]$ and $\mathrm{E}[B_1]$.

Example 2. Let $X = \{X(t), t \geq 0\}$ be a non-delayed regenerative process, taking values in a state space S, with regenerative times $0 = T(0) < T(1) < \cdots$. Let f and g be two real-valued non-negative (measurable) functions defined on S, and set

$$V(t) = \int_0^t g(X(s))ds$$
$$\alpha = \mathrm{E}\Big[\int_0^\infty \exp[-V(t)]f(X(t))dt\Big].$$

Then, α is the infinite-horizon expected discounted cost, the process $g(X(t))$ corresponds to the (state-dependent) discount rate at time t, and $f(X(t))$ is the (undiscounted) rate at which cost is incurred at time t. A common choice for g is the one in which $g(\cdot)$ is constant and equal to $\rho > 0$, in which case

$$\alpha = \mathrm{E}\Big[\int_0^\infty \exp[-\rho t]f(X(t))dt\Big]$$

is the infinite-horizon ρ-discounted cost. Let

$$A_1 = \int_{T(0)}^{T(1)} \exp\Big[-\int_0^t g(X(s))ds\Big] f(X(t))dt$$
$$C_1 = \exp[-V(T(1))]$$
$$B_1 = 1 - C_1.$$

Because of the regenerative structure of X, it is evident that α satisfies the equation $\alpha = \mathrm{E}[A_1] + \mathrm{E}[C_1]\alpha$. Thus, if $\mathrm{E}[C_1] < 1$, it follows that α is finite and can be expressed as

$$\alpha = \frac{\mathrm{E}[A_1]}{\mathrm{E}[B_1]}.$$

Hence, the infinite-horizon discounted cost for a regenerative process can be expressed in terms of a ratio estimation problem; see [3] for further details.

Example 3. Let X be a regenerative process as in Example 2, and assume that X has right-continuous paths with left limits. Let F be a non-empty subset of the state space S, and let $\tau(F) = \inf\{t \geq 0 | X(t) \in F\}$ be the first hitting time of the subset F. Then,

$$\alpha = \mathrm{E}[\tau(F)]$$

is the mean hitting time of F. Such expectations are of interest, for example, in the reliability setting, in which case $\tau(F)$ would typically correspond to the system failure time, and $T(1)$ to a time at which the system is brought back to an "as good as new" state. Let

$$A_1 = \min[\tau(F), T(1)]$$
$$B_1 = I[\tau(F) < T(1)],$$

where I denotes the indicator function. If $P[\tau(F) < \infty] > 0$ (note that this is equivalent to requiring that $P[\tau(F) < T(1)] > 0$), it is easily shown that

$$\alpha = \frac{\mathrm{E}[A_1]}{\mathrm{E}[B_1]}.$$

See [7] for additional details. Thus, the mean hitting time of a regenerative process can be formulated in terms of the ratio estimation problem.

Example 4. Let X be a real-valued random variable and let C be an event with $P(C) > 0$. Suppose that we wish to estimate

$$\alpha = \mathrm{E}[X | C],$$

namely the conditional expectation of X, given that the event C has occurred. If $E[|X|] < \infty$, then we can express α in terms of the ratio $\alpha = E[A_1]/E[B_1]$, where

$$A_1 = XI(C)$$
$$B_1 = I(C).$$

Hence, conditional expectations are expressable in terms of the ratio estimation problem.

Thus, the ratio estimation problem arises in a variety of different applications contexts. We shall now introduce a decision parameter vector θ into the discussion. For each $\theta \in \mathbf{R}^d$, let P_θ be the probability measure associated with the parameter value θ, and let E_θ be its corresponding expectation operator. In addition, we shall permit the random variable $A(\theta)$ and $B(\theta)$ to depend explicitly on $\theta \in \mathbf{R}^d$. Then, for each $\theta \in \mathbf{R}^d$, the ratio of expectations can be expressed in the form

$$\alpha(\theta) = \frac{u(\theta)}{\ell(\theta)},$$

where $u(\theta) = E_\theta[A(\theta)]$ and $\ell(\theta) = E_\theta[B(\theta)]$. Given our above examples, computing the gradient of such a ratio $\alpha(\theta)$ is useful for sensitivity analysis or optimization of any of the following: steady-state costs or rewards in regenerative processes; infinite-horizon discounted costs; mean time to failure in reliability systems; conditional expectations and probabilities.

3. Confidence Intervals For Gradient Estimators of Ratios

Let $\theta_0 \in \mathbf{R}^d$ be fixed. In order for the gradient estimation problem to make sense, we shall require that both $u(\cdot)$ and $\ell(\cdot)$ have gradients at $\theta = \theta_0$. We shall further assume that there exists unbiased estimators for not only $u(\theta_0)$ and $\ell(\theta_0)$, but also their gradients $\nabla u(\theta_0)$ and $\nabla \ell(\theta_0)$. Focussing now on the i-th component of the gradient, we shall specifically assume that there exist jointly distributed random variables (A, B, C, D) such that

$$E[A] = u(\theta_0)$$
$$E[B] = \ell(\theta_0)$$
$$E[C] = \partial_i u(\theta_0) \stackrel{def}{=} \frac{\partial}{\partial \theta_i} u(\theta) \Big|_{\theta=\theta_0}$$
$$E[D] = \partial_i \ell(\theta_0) \stackrel{def}{=} \frac{\partial}{\partial \theta_i} \ell(\theta) \Big|_{\theta=\theta_0}$$

where ∂_i denotes the partial derivative with respect to θ_i, and θ_i is the i-th component of θ.

There is now a great deal of literature on various ways of constructing unbiased estimators for $\partial_i u(\theta_0)$ and $\partial_i \ell(\theta_0)$. The two principal approaches that have been explored are likelihood ratio gradient estimation (see [5] for a survey) and infinitesimal perturbation analysis (see [4]). For links between the two methods and for a general survey, see [10] and [11].

We shall now assume that it is possible for the simulator to generate a sequence $\{(A_j, B_j, C_j, D_j), j \geq 1\}$ of i.i.d. replicates of the random vector (A, B, C, D). In each of the problem settings described in Section 2, this is typically straightforward.

To estimate

$$\begin{aligned}\partial_i\alpha(\theta_0) &= \frac{\ell(\theta_0)\partial_i u(\theta_0) - u(\theta_0)\partial_i\ell(\theta_0)}{\ell^2(\theta_0)} \\ &= \frac{\partial_i u(\theta_0) - \alpha(\theta_0)\partial_i\ell(\theta_0)}{\ell(\theta_0)},\end{aligned}$$

the natural estimator to use is

$$\delta_i(n) = \frac{\overline{C}_n - \alpha_n\overline{D}_n}{\overline{B}_n},$$

where

$$\begin{aligned}\overline{A}_n &= \frac{1}{n}\sum_{j=1}^{n} A_j \\ \overline{B}_n &= \frac{1}{n}\sum_{j=1}^{n} B_j \\ \overline{C}_n &= \frac{1}{n}\sum_{j=1}^{n} C_j \\ \overline{D}_n &= \frac{1}{n}\sum_{j=1}^{n} D_j\end{aligned}$$

and

$$\alpha_n = \overline{A}_n/\overline{B}_n.$$

Our first proposition states that under reasonable conditions, $\delta_i(n)$ is a consistent estimator for $\partial_i\alpha(\theta_0)$. The proof is straightforward and therefore omitted.

PROPOSITION 1. *Suppose that* $\mathrm{E}[|A_1| + |B_1| + |C_1| + |D_1|] < \infty$ *and that* $\mathrm{E}[B_1] \neq 0$. *Then,*

$$\lim_{n\to\infty} \delta_i(n) \overset{a.s.}{=} \partial_i\alpha(\theta_0). \blacksquare$$

To develop a confidence interval methodology for $\partial_i(n)$, we need a central-limit theorem (CLT) for the estimator. Let

$$\begin{aligned}Z_j &= A_j - \alpha(\theta_0)B_j \\ W_j &= C_j - \alpha(\theta_0)D_j - \partial_i\alpha(\theta_0)B_j\end{aligned}$$

and note that under the assumptions of Proposition 1, $E[Z_j] = E[W_j] = 0$. This observation is an important element in the proof of the following theorem.

THEOREM 1. *Assume that* $E[Z_1^2 + W_1^2] < \infty$. *If, in addition, the conditions of Proposition 1 are in force, then*

$$\sqrt{n}[\delta_i(n) - \partial_i\alpha(\theta_0)] \Rightarrow \sigma N(0,1)$$

as $n \to \infty$, *where*

$$\sigma^2 = \frac{E[W_1 - (E[D_1]/E[B_1])Z_1]^2}{(E[B_1])^2}. \blacksquare$$

Theorem 1 has been previously established, using different methods, by [14] in the context of likelihood ratio gradient estimation for regenerative steady-state simulation. Their expression for the variance constant σ^2 is formally different, but algebraically identical.

The final step need to develop a confidence interval methodology for $\delta_i(n)$ is the construction of an appropriate estimator for σ^2. Let

$$v(n) = \frac{\frac{1}{n}\sum_{j=1}^{n}[\widehat{W}_j - (\overline{D}_n/\overline{B}_n)\widehat{Z}_j]^2}{(\overline{B}_n)^2}$$

where

$$\widehat{Z}_j = A_j - \alpha_n B_j$$
$$\widehat{W}_j = C_j - \alpha_n D_j - \delta_i(n)B_j.$$

The next proposition gives conditions under which $v(n)$ is strongly consistent for σ^2. The proof is straightforward and therefore omitted.

PROPOSITION 2. *Suppose that* $E[A_1^2 + B_1^2 + C_1^2 + D_1^2] < \infty$. *If* $E[B_1] \neq 0$, *then*

$$\lim_{n\to\infty} v(n) \overset{a.s.}{=} \sigma^2. \blacksquare$$

We note that if $v(n)$ is computed via a two-pass approach in which α_n and $\delta_i(n)$ are computed in the first pass through the data $\{(A_j, B_j, C_j, D_j), 1 \le j \le n\}$ and the sum of squares computed in the second pass, then it is essentially guaranteed that $v(n)$ will be computed as a non-negative quantity on any finite-precision computer. More importantly, this means of computing $v(n)$ is likely to be more stable numerically than that associated with the computation described in [14].

We are now ready to describe a general confidence interval methodology for estimating partial derivatives of ratios.. Suppose that we wish to compute a $100(1-\delta)\%$ confidence interval for $\partial_i\alpha(\theta_0)$. We use the following procedure:

Algorithm CI.

1 Generate $\{(A_j, B_j, C_j, D_j), j \geq 1\}$.

2 Compute α_n and $\delta_i(n)$.

3 Compute $v(n)$ (using the two-pass approach described above).

4 Find $z(\delta)$ such that $P[N(0,1) \leq z(\delta)] = 1 - \delta/2$.

5 Compute

$$L_n = \delta_i(n) - z(\delta)\sqrt{v(n)/n}$$
$$U_n = \delta_i(n) + z(\delta)\sqrt{v(n)/n}. \blacksquare$$

Then, $[L_n, U_n]$ is an (approximate) $100(1-\delta)\%$ confidence interval for $\partial_i\alpha(\theta_0)$. In particular, if the conditions of Proposition 2 are in force and $\sigma^2 > 0$, then

$$\lim_{n\to\infty} P[\partial_i\alpha(\theta_0) \in [L_n, U_n]] = 1 - \delta.$$

We conclude this section with a brief discussion of the problem of generating a confidence region for the vector $(\alpha(\theta_0), \partial_1\alpha(\theta_0), \ldots, \partial_d\alpha(\theta_0))$. A joint confidence region could be of potential interest in a number of optimization settings, since virtually all iterative (deterministic) optimization algorithms choose their search direction, at each iteration, by considering the full gradient.

Let $C(i)$ and $D(i)$ be unbiased estimators for $\partial_i u(\theta_0)$ and $\partial_i \ell(\theta_0)$, so that

$$\mathrm{E}[C(i)] = \partial_i u(\theta_0)$$
$$\mathrm{E}[D(i)] = \partial_i \ell(\theta_0).$$

If $\{(A_j, B_j, C_j(1), D_j(1), \ldots, C_j(d), D_j(d)), 1 \leq j \leq n\}$ is a set of n i.i.d. replicates of the random vector $(A, B, C(1), D(1), \ldots, C(d), D(d))$, then the estimators $\alpha_n, \delta_1(n), \ldots, \delta_d(n)$ can be constructed from the sample in the obvious way, namely

$$\alpha_n = \overline{A}_n/\overline{B}_n$$
$$\delta_i(n) = (\overline{C}_n(i) - \alpha_n\overline{D}_n(i))/\overline{B}_n.$$

Define

$$W_j(i) = C_j(i) - \alpha(\theta_0)D_j(i) - \partial_i\alpha(\theta_0)B_j.$$

We are now ready to state a joint CLT for $(\alpha_n, \delta_1(n), \ldots, \delta_d(n))$.

THEOREM 2. *Assume that* $\mathrm{E}[A_1^2 + B_1^2 + C_1^2(1) + D_1^2(1) + \cdots + C_1^2(d) + D_1^2(d)] < \infty$. *If* $\mathrm{E}[B_1] \neq 0$, *then*

$$\sqrt{n}[\alpha_n - \alpha(\theta_0), \delta_1(n) - \partial_1\alpha(\theta_0), \ldots, \delta_d(n) - \partial_d\alpha(\theta_0)]\mathrm{E}[B_1] \Rightarrow N(0, C)$$

as $n \to \infty$, *where* $C = (C_{ij}, 0 \le i, j \le d)$ *is a covariance matrix whose elements are given by*

$$
\begin{aligned}
C_{00} &= \mathrm{E}[Z_1^2] \\
C_{0i} &= C_{i0} \\
&= \mathrm{E}\big[(W_1(i) - \frac{\mathrm{E}[D_1(i)]}{\mathrm{E}[B_1]} Z_1) Z_1\big] \\
C_{ij} &= C_{ji} \\
&= \mathrm{E}\big[(W_1(i) - \frac{\mathrm{E}[D_1(i)]}{\mathrm{E}[B_1]} Z_1)(W_1(j) - \frac{\mathrm{E}[D_1(j)]}{\mathrm{E}[B_1]} Z_1)\big]
\end{aligned}
$$

for $1 \le i, j \le d$.■

The proof of this theorem mirrors that of Theorem 1 and is therefore omitted.

A procedure for producing asymptotically valid confidence regions for $(\alpha(\theta_0), \partial_1\alpha(\theta_0), \ldots \partial_d\alpha(\theta_0))$ can now easily be derived, using the same ideas as those described earlier in this section for $\partial_i\alpha(\theta_0)$.

4. Low Bias Estimation for the Gradient of a Ratio

Since the gradient of the ratio is a nonlinear function of the expectations $\mathrm{E}[A], \mathrm{E}[B], \mathrm{E}[C(1)], \mathrm{E}[D(1)], \ldots, \mathrm{E}[C(d)], \mathrm{E}[D(d)]$, it follows that the estimator $\delta_i(n)$ is, in general, biased for $\partial_i\alpha(\theta_0)$.

We will now proceed to (formally derive a bias expansion for $\delta_i(n)$. The proof of Theorem 1 shows that

$$\delta_i(n) - \partial_i\alpha(\theta_0) = \frac{\overline{W}_n - (\overline{D}_n/\overline{B}_n)\overline{Z}_n}{\overline{B}_n}. \tag{4.1}$$

We would like to approximate the expectation of that. We note that since $\overline{B}_n$ is close to $\mu \stackrel{def}{=} \mathrm{E}[B_1]$ for large n, we can use the power series expansion for $f(x) = (1-x)^{-1}$ to obtain

$$
\begin{aligned}
\frac{1}{\overline{B}_n} &= \frac{1}{\mu}[1 - (1 - \frac{\overline{B}_n}{\mu})]^{-1} \\
&= \mu^{-1}[1 + (1 - \frac{\overline{B}_n}{\mu}) + (1 - \frac{\overline{B}_n}{\mu})^2 + \cdots] \\
&\approx \mu^{-1}[1 + (1 - \frac{\overline{B}_n}{\mu})] \\
&= \frac{2\mu - \overline{B}_n}{\mu^2}.
\end{aligned}
$$

Using this approximation in (4.1), we find that

$$\begin{aligned}
\delta_i(n) - \partial_i\alpha(\theta_0) &\approx \frac{2\mu - \overline{B}_n}{\mu^2}\left(\overline{W}_n - \frac{\overline{Z}_n\overline{D}_n}{\overline{B}_n}\right) \\
&= \frac{2\overline{W}_n}{\mu} + \frac{\overline{Z}_n\overline{D}_n - \overline{B}_n\overline{W}_n}{\mu^2} - \frac{2\overline{Z}_n\overline{D}_n}{\mu\overline{B}_n} \\
&\approx \frac{2\overline{W}_n}{\mu} + \frac{\overline{Z}_n\overline{D}_n - \overline{B}_n\overline{W}_n}{\mu^2} - \frac{2\overline{Z}_n\overline{D}_n(2\mu - \overline{B}_n)}{\mu^3} \\
&= \frac{2\overline{W}_n}{\mu} - \frac{\overline{Z}_n\overline{D}_n + \overline{B}_n\overline{W}_n}{\mu^2} + \frac{2\overline{Z}_n\overline{D}_n\widetilde{B}_n}{\mu^3},
\end{aligned} \tag{4.2}$$

where $\widetilde{B}_n = \overline{B}_n - \mu$. Recall that $\mathrm{E}[W_j] = \mathrm{E}[Z_j] = 0$. Observe that for $i \neq j$, $\mathrm{E}[B_iW_j] = \mathrm{E}[B_i]\mathrm{E}[W_j] = 0$, since B_i and W_j are independent. Therefore

$$\mathrm{E}[\overline{B}_n\overline{W}_n] = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\mathrm{E}[B_iW_j] = \frac{\mathrm{E}[B_1W_1]}{n}.$$

Similarly, $\mathrm{E}[\overline{Z}_n\overline{D}_n] = \mathrm{E}[Z_1D_1]/n$. Also, $\mathrm{E}[Z_iD_j(B_k - \mu)] = 0$ whenever $i \neq k$. Therefore,

$$\begin{aligned}
\mathrm{E}[\overline{Z}_n\overline{D}_n\widetilde{B}_n] &= \frac{\mathrm{E}[Z_1D_1(B_1 - \mu)]}{n^2} + \frac{(n-1)\mathrm{E}[Z_1(B_1 - \mu)]\mathrm{E}[D_1]}{n^2} \\
&= \frac{\mathrm{E}[Z_1B_1]\mathrm{E}[D_1]}{n} + o(1/n).
\end{aligned}$$

Now taking the expectation in (4.2) yields

$$\mathrm{E}[\delta_i(n)] - \partial_i\alpha(\theta_0) \approx \frac{2\mathrm{E}[Z_1B_1]\mathrm{E}[D_1] - \mu\mathrm{E}[B_1W_1 + Z_1D_1]}{n\mu^3}.$$

This bias approximation suggests an obvious means of reducing the bias of gradient estimators for ratios. The idea is to estimate the bias term and correct for it by subtracting off the estimated bias. In this case, this approach leads to the estimator

$$\widetilde{\delta}_i(n) = \delta_i(n) - \frac{2\overline{D}_n\sum_{j=1}^{n}\widehat{Z}_jB_j}{n^2\overline{B}_n^3} + \frac{\sum_{j=1}^{n}(B_j\widehat{W}_j + \widehat{Z}_jD_j)}{n^2\overline{B}_n^2},$$

where $\widehat{Z}_j$ and $\widehat{W}_j$ are defined just before the statement of Proposition 2 in Section 3.

Under the appropriate regularity hypotheses, and by applying techniques similar to those used in [6], one can rigorously prove that $\widetilde{\delta}_i(n)$ reduces the asymptotic bias, in the sense that

$$\mathrm{E}[\widetilde{\delta}_i(n)] = \partial_i\alpha(\theta_0) + o(1/n).$$

A second approach that is frequently used to correct for "nonlinearity bias" of the above type is to "jackknife" the estimator. Specifically, for $1 \leq j \leq n$, let

$$\alpha_{n(j)} = \frac{\sum_{k=1,k\neq j}^{n} A_k}{\sum_{k=1,k\neq j} B_k}$$

$$\beta_{n(j)} = \frac{\left(\sum_{k=1,k\neq j}^{n} C_j\right) - \alpha_{n(j)} \sum_{k=1,k\neq j}^{n} D_j}{\sum_{k=1,k\neq j}^{n} B_j}$$

$$\delta_{n(j)} = n\alpha_n(j) - (n-1)\beta_{n(j)}.$$

Then,

$$\delta_i^J(n) = \frac{1}{n} \sum_{j=1}^{n} \delta_{n(j)}$$

is the jackknife estimator for $\partial_i \alpha(\theta_0)$. Also,

$$\sqrt{n} \frac{(\delta_i^J(n) - \partial_i \alpha(\theta_0))}{s^J(n)} \Rightarrow N(0,1),$$

where

$$(s^J(n))^2 = \frac{1}{n-1} \sum_{j=1}^{n} (\delta_{n(j)} - \delta_i^J(n))^2$$

is a consistent variance estimator. As in the case of the estimator $\bar{\delta}_i(n)$, one can prove rigorously (under suitable regularity hypotheses) that the estimator $\delta_i^J(n)$ reduces asymptotic bias, in the sense that

$$\mathrm{E}[\delta_i^J(n)] = \partial_i \alpha(\theta_0) + o(1/n).$$

It turns out that the improved bias characteristics of these estimators are cost-less relative to the variance, in the sense that the estimators $\bar{\delta}_i(n)$ and $\delta_i^J(n)$ obey precisely the same CLT as does $\delta_i(n)$. Hence, the estimators exhibit the same degree of asymptotic variability.

THEOREM 3. *Assume that* $\mathrm{E}[A_1^2 + B_1^2 + C_1^2 + D_1^2] < \infty$ *and that* $\mathrm{E}[B_1] \neq 0$. *Then,*

$$\sqrt{n}(\bar{\delta}_i(n) - \partial_i \alpha(\theta_0)) \Rightarrow \sigma N(0,1)$$

$$\sqrt{n}(\delta_i^J(n) - \partial_i \alpha(\theta_0)) \Rightarrow \sigma N(0,1)$$

where σ^2 *is the same constant as in Theorem 1.* ■

5. Conclusion

Ratio estimation problems arise in many different applications settings. When estimation is to be used to analyze the sensitivity of (or to optimize) a system in which the ratio estimation problem occurs, the results of this chapter become pertinent. We have derived a numerically stable confidence interval procedure for computing partial derivatives of such ratios, and have developed the appropriate joint CLT's necessary to extend this methodology to the computation of confidence regions for the full gradient of the ratio. In addition, we have discussed low-bias estimators for computing such partial derivatives.

6. Appendix

Proof of Theorem 1. We note that

$$\begin{aligned}\overline{B}_n[\delta_i(n) - \partial_i\alpha(\theta_0)] &= \overline{C}_n - \alpha_n\overline{D}_n - \partial_i\alpha(\theta_0)\overline{B}_n \\ &= \overline{W}_n - (\alpha_n - \alpha(\theta_0))\overline{D}_n \\ &= \overline{W}_n - (\overline{D}_n/\overline{B}_n)\overline{Z}_n \\ &= \overline{W}_n - (\mathrm{E}[D_1]/\mathrm{E}[B_1])\overline{Z}_n - (\overline{D}_n/\overline{B}_n - \mathrm{E}[D_1]/\mathrm{E}[B_1])\overline{Z}_n.\end{aligned}$$

Clearly, $\sqrt{n}\overline{Z}_n \Rightarrow (\mathrm{E}[Z_1^2])^{1/2}N(0,1)$ as $n \to \infty$ and $\overline{D}_n/\overline{B}_n \xrightarrow{a.s.} \mathrm{E}[D_1]/\mathrm{E}[B_1]$ as $n \to \infty$. It follows, by the converging-together principle, that

$$\sqrt{n}\,(\overline{D}_n/\overline{B}_n - \mathrm{E}[D_1]/\mathrm{E}[B_1])\overline{Z}_n \Rightarrow 0$$

as $n \to \infty$. The CLT for i.i.d. random variables also proves that

$$\sqrt{n}(\overline{W}_n - (\mathrm{E}[D_1]/\mathrm{E}[B_1])\overline{Z}_n) \Rightarrow \mathrm{E}[B_1]\sigma N(0,1)$$

as $n \to \infty$. A second application of the converging-together principle then yields

$$\sqrt{n}\,\overline{B}_n(\delta_i(n) - \partial_i\alpha(\theta_0)) \Rightarrow \mathrm{E}[B_1]\sigma N(0,1).$$

One final application of the converging-together principle (note that $\overline{B}_n \xrightarrow{a.s.} \mathrm{E}[B_1]$ as $n \to \infty$) proves the theorem.

References

[1] Asmussen, S. 1987. *Applied Probability and Queues*, Wiley.

[2] Bratley, P., B.L. Fox and L.E. Schrage. 1987. *A Guide to Simulation*, Springer-Verlag, New York, second edition.

[3] Fox, B.L. and P.W. Glynn, 1989. Simulating Discounted Costs. *Management Science*, 35, 1297-1315.

[4] Glasserman, P. 1991. *Gradient Estimation via Perturbation Analysis*, Kluwer Academic.

REFERENCES

[5] Glynn, P.W. 1990. Likelihood Ratio Gradient Estimation for Stochastic Systems. *Communications of the ACM*, 33, 10, 75-84.

[6] Glynn, P.W. and Heidelberger, P. 1991. Jackknifing Under a Budget Constraint. *ORSA Journal on Computing*, to appear.

[7] Goyal, A., P.Shahabuddin, Heidelberger, P., Nicola, V. F., and Glynn, P.W. 1991. A Unified Framework for Simulating Markovian Models of Highly Dependable Systems. *IEEE Transactions on Computers*. To appear.

[8] Heidelberger, P., X.-R. Cao, M.A. Zazanis, and R. Suri. 1988. Convergence Properties of Infinitesimal Perturbation Analysis Estimates. *Management Science*, 34, 11, 1281-1302.

[9] Iglehart, D.L. 1975. Simulating Stable Stochastic Systems, V: Comparison of Ratio Estimators. *Naval Research Logistics Quarterly*, 22, 553-565.

[10] L'Ecuyer, P. 1990. A unified Version of the IPA, SF, and LR Gradient Estimation Techniques. *Management Sciences*, 36, 11, pp.1364-1383.

[11] L'Ecuyer, P. 1991. An Overview of Derivative Estimation. In these proceedings.

[12] L'Ecuyer, P. , N. Giroux and P.W. Glynn. 1991. Stochastic Optimization by Simulation: Convergence Proofs and Experimental Results for the GI/G/1 Queue in Steady-State. In preparation.

[13] Miller, R.G. 1974. The Jackknife - A Review. *Biometrika*, 61, 1-15.

[14] Reiman, M.I and Weiss, A. 1989. Sensitivity Analysis for Simulations via Likelihood Ratios. *Op.Res.*, 37, 5, pp.830-844.

[15] Wolff, R. 1989. *Stochastic Modeling and the Theory of Queues*, Prentice-Hall.

CHAPTER 7

General Conclusion and Future Research

In this thesis, a class of problems in statistics and control, both with applications in electrical engineering has been studied, whereby an underlying time renewal sequence plays a key role in the evolution of the dynamic quantities of interest.

It will be of particular interest to study the moments of the same general forms; as detailed in Chapters 2, 3, and 4; but with dependent random variables. This type of moments involving dependent variables has many practical applications, mainly in random cumulative fatigue.

For the stochastic optimal control part, it will be very relevant, as subsequent research, to study the stochastic optimal control problem, where the framework is a fixed finite horizon with Poisson distributed observation instants, and the total number of observations is random.

We hope that the mathematical theory developed in this thesis will help the formulation and eventual resolution of new problems deriving from a cross between the problems of operations research and those of main stream in stochastic control theory.