Markovian modelling of gene products synthesis

Jean Peccoud
TIMC-IMAG
Faculté de médecine
Institut Albert Bonniot
38706 La Tronche Cedex
France

Bernard Ycart
LMC-IMAG
Université Joseph Fourier
BP53
38041 Grenoble Cedex 9
France

Abstract

We propose a Markovian model for the gene induction process. This model is a particular case of a birth and death process in Markovian environment. The environment is represented by the state of the gene (active or inactive) and the protein population size is ruled by the birth and death process. Proteins can be produced by the gene when it is activated (birth). They can disappear as well due to their denaturation (death). The distribution of the protein number is computed explicitly both for the transient and stationary case. It is shown that three out of the four parameters of the model can be estimated from an observed sample of protein numbers collected during the stationary regime.

1 Introduction

One of the major achievements of molecular biology during the last two decades has been the characterization of many phenomena ensuring the fine tuning of gene expression at the transcriptional level. One of the simplest kinds of gene regulations is induction. Basically, an inducible gene is a gene which is normally silent but can be expressed under the control of external signals such as hormone, sugar, temperature and so on. This reaction is usually mediated by a regulatory factor which can be activated by the external signal through ligand binding or phosphorylation for instance. The activation of this factor is reversible. The level of gene expression is proportional to the fraction of the regulatory factor in the activated state. Although the dissection of molecular mechanisms of gene regulation is presently a major concern for molecular biologists, quantitative observations of this phenomenon, in particular *in vivo*, are still strikingly scarce [11]. In 1990 Minoru Ko and coworkers conducted a pioneering experiment in which they were able to quantify on a cell-by-cell basis the induction of a reporter gene under the control of the glucocorticoid-inducible promoter/enhancer of the mouse mammary tumor virus [8]. They pointed out that the individual levels of reporter activity were very heterogeneous among a single cell population grown in a homogeneous concentration of inducer. When transcription studies are conducted *in vitro* the collected data are smooth since they result from the average behavior of large populations of transcription templates whereas in a single cell the action of the inducer results in a probability for the transcription template to be in the activated state, *i.e.* a probability for the gene to be in its "on" state [10]. In this case, one can imagine that large fluctuations of the gene activity can be observed if the half-life of the "on" state is large compared to the average time of a transcription cycle [15]. Ko initiated the mathematical modelling of gene induction mainly by doing Monte-Carlo simulations [9] of a discrete time two states Markov chain. The gene activity level was assumed to be proportional to the total time spent by the process in the "on" state.

In this paper a more detailed model of gene induction is presented. Both the random switch between the active and inactive states of the gene and the randomness of transcription initiation [15] are considered. The set of interactions considered in this model are

$$I \xrightarrow{\lambda} A \tag{1}$$

$$A \xrightarrow{\mu} I \tag{2}$$

$$A \xrightarrow{\nu} A + P \tag{3}$$

$$P \xrightarrow{\delta} \emptyset \tag{4}$$

where (1) and (2) express the change of the gene between inactivity and activity, (3) the reporter gene protein synthesis, and (4) the protein degradation. The symbol $\emptyset$ is used as a zero of chemical reaction meaning that P is transformed into something that is not considered in the model. The time evolution of the gene product synthesis is modelled by a Markov process in continuous time (see for instance [2] for a general reference). The state space of this process is

$$S = \{(i, n)\ ,\ i \in \{0, 1\}, n \in \mathbb{N}\}\,.$$

In any state (i, n) the first coordinate denotes the status of the gene.

$i = 0$ when the gene is inactive,

$i = 1$ when the gene is active.

The second coordinate n represents the number of molecules (copies) of the protein P in the cell. It is assumed that the time evolution of the state of the gene is Markovian. The gene remains active during a random time exponentially distributed with parameter μ ($1/\mu$ time units on average). Then, the gene jumps to its inactive state where it remains during a random time exponentially distributed with parameter λ. When the gene is inactive no protein is synthesized whereas during activity they are synthesized with a constant rate ν. Finally the life durations of all copies of P are assumed to be independent, identically distributed random variables, their common distribution being exponential with parameter δ. With these notations, it is possible to detail the rates of transition on the state space S.

From the state $(0, n)$, inactive gene and n copies of P, the process may jump to

$(1, n)$ with rate λ (the gene switches to the active state),

$(0, n-1)$ with rate $n\delta$ (degradation of one P).

From the state $(1, n)$, active gene and n molecules of P, the process may jump to

$(0, n)$ with rate μ (the gene switches to inactivity),

$(1, n-1)$ with rate $n\delta$ (degradation of one P),

$(1, n+1)$ with rate ν (synthesis of one P).

This set of transitions defines a Markov process on S that will be referred to as the IAP process.

Let $p_{0,n}(t)$ (respectively $p_{1,n}(t)$) be the probability that at time t the gene is inactive (resp. active) and that n molecules of P are present. It is easy to derive the Kolmogorov system of equations for the IAP process from the rates of transition [2].

$$\left\{ \begin{array}{rcl} \forall n \geq 0,\ p'_{0,n}(t) & = & -(\lambda + n\delta)p_{0,n}(t) + (n+1)\delta p_{0,n+1}(t) + \mu p_{1,n}(t) \\ p'_{1,0}(t) & = & -(\mu + \nu)p_{1,0}(t) + \delta p_{1,1}(t) + \lambda p_{0,0}(t) \\ \forall n \geq 1,\ p'_{1,n}(t) & = & -(\lambda + \nu + n\delta)p_{1,n}(t) + (n+1)\delta p_{1,n+1}(t) + \nu p_{1,n-1}(t) + \lambda p_{0,n}(t) \end{array} \right. \tag{5}$$

The model that has just been built is a particular case of a Markov process in random environment. These processes have first been introduced in different contexts by Eisen and Tainiter [4] for queuing processes and by Smith and Wilkinson [12] for branching processes. A review of processes in random environments and their applications to population studies and epidemiology can be found in [1].

More precisely, the IAP process can be seen as a birth and death process with a two states Markovian environment in continuous time [3]. In this perspective, the evolution in time of the gene between activity and inactivity is the environment process. The state of that process determines the instantaneous rates of the population of protein molecules birth and death process. When the gene is active and the decay rate δ is positive, the copy number of P evolves according to a linear growth birth and death process known in queuing theory as the $M/M/\infty$ queue [2]. When $\delta = 0$ the protein process is a Poisson process since no protein denaturation occurs. The asymptotic properties of birth and death processes in random environments have been extensively studied: stationarity conditions, central limit theorems, law of large numbers, diffusion approximations, and many other important results have been proved [3, 7, 14]. However, explicit solutions to the Kolmogorov system of equations, even in the stationary case are out of reach in general [3]. The main purpose of this paper is to show that the distribution of the IAP process can be explicitly computed in the transient case as well as in the stationary one. As a by-product of these computations, it will be shown that three out of the four parameters of the model can be estimated from a sample of the number of copies of P. This should allow an experimental validation of the IAP model of gene induction.

Section 2 deals with the transient case. The general case ($\delta > 0$) is set apart from the Poissonian case ($\delta = 0$) where no stationary regime exists. The asymptotics of the first two moments are derived and interpreted. Section 3 is devoted to the stationary regime. The generating function of the asymptotic distribution of the number of copies of P is shown to be a degenerate hypergeometric function. Explicit convergent estimators of the parameters in terms of the first three moments of that distribution are obtained.

2 Transient behavior

Finding the transient behavior of the IAP process amounts to solving the Kolmogorov system of differential equations (5). Two generating functions are introduced for this purpose.

$$G_0(z,t) \;=\; \sum_{n=0}^{\infty} z^n p_{0,n}(t) \quad ; \quad G_1(z,t) \;=\; \sum_{n=0}^{\infty} z^n p_{1,n}(t) \tag{6}$$

Multiplying by z^n the equations in $p'_{0,n}(t)$ and $p'_{1,n}(t)$ of the Kolmogorov system and summing them up results in the following system of partial differential equations in $G_0(z,t)$ and $G_1(z,t)$.

$$\left\{ \begin{array}{l} \dfrac{\partial G_0}{\partial t}(z,t) = -\lambda G_0(z,t) + \mu G_1(z,t) + \delta(1-z)\dfrac{\partial G_0}{\partial z}(z,t) \\ \dfrac{\partial G_1}{\partial t}(z,t) = \lambda G_0(z,t) - \mu G_1(z,t) - \nu(1-z)G_1(z,t) + \delta(1-z)\dfrac{\partial G_1}{\partial z}(z,t) \end{array} \right. \tag{7}$$

From the expression of $G_0(z,t)$ and $G_1(z,t)$ the two components of the IAP process can be studied separately. The sum

$$G(z,t) = G_0(z,t) + G_1(z,t)$$

is the generating function of the protein number at time t. Of course,

$$\forall t \geq 0,\ G(1,t) = 1 \quad .$$

Moreover

$$\gamma_0(t) = G_0(1,t) \quad \text{and} \quad \gamma_1(t) = G_1(1,t)$$

are the probabilities for the gene to be inactive and active respectively at time t. Since the gene activity is a two states Markov process, its transient behavior is easily computed.

$$\begin{aligned} \forall t \geq 0,\ \gamma_0(t) = G_0(1,t) &= \frac{\mu}{\lambda+\mu} + \left(\gamma_0(0) - \frac{\mu}{\lambda+\mu}\right) e^{-(\lambda+\mu)t} \\ \gamma_1(t) = G_1(1,t) &= \frac{\lambda}{\lambda+\mu} + \left(\gamma_1(0) - \frac{\lambda}{\lambda+\mu}\right) e^{-(\lambda+\mu)t} \end{aligned} \tag{8}$$

Now for any given initial condition $(G_0(z,0), G_1(z,0))$, the system (7) admits a unique solution. Since the topic of this paper is gene induction, we assume that there is no copy of P in the cell at time $t = 0$.

$$\forall n \geq 1,\ p_{0,n}(0) = p_{1,n}(0) = 0 \tag{9}$$

It could seem mathematically reasonable to assume that the gene activity process is already at equilibrium at $t = 0$ and hence to start with

$$p_{0,0}(0) = \frac{\mu}{\lambda+\mu} \quad \text{and} \quad p_{1,0}(0) = \frac{\lambda}{\lambda+\mu} \ .$$

This assumption would lead to simpler expressions since the functions $\gamma_0(t)$ and $\gamma_1(t)$ would then be constant. However, this hypothesis does not seem to be biologically realistic. It is more consistent with the gene induction phenomenon to assume that the gene is inactive at time $t = 0$: $p_{0,0}(0) = 1$. Thus the initial values for the generating functions G_0 and G_1 are 1 and 0 respectively.

$$G_0(z,0) = 1 \text{ and } G_1(z,0) = 0 \tag{10}$$

At this stage the case where $\delta > 0$ has to be distinguished from $\delta = 0$ (no protein degradation considered). In the latter case, the system of partial differential equations

(7) simplifies into a linear system of differential equations with constant coefficients, z appearing only as a parameter.

$$\left\{ \begin{array}{l} \dfrac{\partial G_0}{\partial t}(z,t) = -\lambda G_0(z,t) + \mu G_1(z,t) \\ \dfrac{\partial G_1}{\partial t}(z,t) = \lambda G_0(z,t) - (\mu + \nu(1-z))\, G_1(z,t) \end{array} \right. \tag{11}$$

The solution of this system with initial condition (10) is easily computed by a symbolic calculation package such as *Mathematica* [6]. Since the number of copies of P is the only experimental data potentially available, only the expression of $G(z,t) = G_0(z,t) + G_1(z,t)$ is of interest.

$$G(z,t) = \frac{1}{2\beta(z)} \left((\alpha(z) + \beta(z))\, e^{(\alpha(z)-\beta(z))\,t} + (\beta(z) - \alpha(z))\, e^{(\alpha(z)+\beta(z))\,t} \right) \tag{12}$$

where

$$\begin{aligned} \alpha(z) &= \frac{1}{2}\left(\nu(z-1) - (\lambda+\mu)\right) \\ \beta(z) &= \frac{1}{2}\left(\nu^2 (z-1)^2 + 2\nu(\lambda-\mu)(z-1) + (\lambda+\mu)\right)^{1/2} . \end{aligned}$$

It is easy to derive the expectation and variance of the copy number at time t from its generating function. The expectation $\mathbf{E}(t)$ is the value of the derivative of $G(z,t)$ with respect to z at $z = 1$.

$$\mathbf{E}(t) = \frac{\lambda}{\lambda+\mu} \left(\nu t - \frac{\nu}{\lambda+\mu} \left(1 - e^{-(\lambda+\mu)t}\right) \right) \tag{13}$$

This expression can be interpreted as follows. If the gene remains permanently active and there is no protein denaturation the number of copies increases as a Poisson process with intensity ν. Therefore the average copy number of P at time t is νt. But if the gene alternates between activity and inactivity, the average proportion of time it spends being active in the long run is $\lambda/(\lambda+\mu)$. Therefore it could be anticipated that $\mathbf{E}(t)$ should be equivalent for large t's to

$$\frac{\lambda}{\lambda+\mu} \nu t .$$

The variance $\mathbf{V}(t)$ of the copy number at time t can be computed in a similar way from the second derivative of $G(z,t)$ with respect to z at $z = 1$.

$$\mathbf{V}(t) = t \left(\frac{\lambda\nu}{\lambda+\mu} + 2\frac{\lambda\mu\nu^2}{(\lambda+\mu)^3} \right)$$

$$+\frac{\lambda\nu}{(\lambda+\mu)^4}\left(\lambda\nu - 4\mu\nu - (\lambda+\mu)^2\right)$$
$$+\frac{\lambda\nu}{(\lambda+\mu)^4}e^{-(\lambda+\mu)t}\left(-\lambda\nu e^{-(\lambda+\mu)t} + 4\mu\nu + (\lambda+\mu)^2 - 2\nu t(\lambda^2-\mu^2)\right) \quad (14)$$

For large t's, the dominant term of the above expression is

$$t\left(\frac{\lambda\nu}{\lambda+\mu} + 2\frac{\lambda\mu\nu^2}{(\lambda+\mu)^3}\right).$$

In this equivalent, it is possible to recognize a term corresponding to the variance νt of a Poisson process with intensity ν multiplied by the asymptotic proportion of activity of the gene. This component of the variance is increased by another term in which one of the factors is the asymptotic variance of the gene activity process $\lambda\mu/(\lambda+\mu)^2$.

The second case where the denaturation rate δ of the P molecules is strictly positive has now to be considered. Now the system (7) is not as easy to solve as in the previous case since it is singular at $z = 1$. One possible method is to express the solution as a power series in $(z-1)$, the coefficients of which are the solution of a recurrence equation (see section 3). In our case, this leads rapidly to very complicated expressions. However, the first terms of the power series expansion of

$$G(z,t) = G_0(z,t) + G_1(z,t)$$

are worth computing since they are related to the moments of the distribution at time t of the copy number of P. At least the computations of the expectation $\mathbf{E}(t)$ and the variance $\mathbf{V}(t)$ are tractable.

$$\mathbf{E}(t) = \frac{\lambda}{\lambda+\mu}\frac{\nu}{\delta} + \frac{\lambda\nu}{(\lambda+\mu)(\lambda+\mu-\delta)}e^{-(\lambda+\mu)t} - \frac{\lambda\nu}{\delta(\lambda+\mu-\delta)}e^{-\delta t}$$

From this expression, it is clear that $\mathbf{E}(t)$ converges exponentially fast to

$$\mathbf{E} = \frac{\lambda\nu}{(\lambda+\mu)\delta} \quad .$$

This asymptotic mean number of copies can be interpreted as follows. With a gene permanently active, the copy number would evolve as an $M/M/\infty$ queue [2]. So its asymptotic distribution would be Poisson with mean ν/δ. But in the long run the gene is active only during a proportion $\lambda/(\lambda+\mu)$ of the time. Notice also that the exponential rates of convergence $-(\lambda+\mu)$ and $-\delta$ are those of the two states Markov environment (activity of the gene) and of the $M/M/\infty$ queue (number of copies) respectively.

The expression of the variance of the copy number is somewhat more complicated and it would be pointless to reproduce it here in full extent. Its asymptotics is the following.

$$\lim_{t\to\infty} \mathbf{V}(t) = \mathbf{V} = \frac{\lambda}{\lambda+\mu}\frac{\nu}{\delta} + \frac{\lambda\mu}{(\lambda+\mu)^2}\frac{\nu^2}{\delta(\lambda+\mu+\delta)}$$

The interpretation of this expression is analogous to the one already given in the case where $\delta = 0$.

3 Asymptotic behavior

When δ is strictly positive, the expressions of the expectation and variance of the number of copies at time t suggest that the IAP process has a stationary distribution to which the distribution at time t converges exponentially fast.

This is indeed the case. That the IAP process is ergodic can be deduced from the general results of Cogburn and Torrez [3]. Only the existence of the stationary regime can be proved in this way. To go further, it is possible to use stochastic comparison (see Stoyan [13] for a general reference). The idea is to compare our model to the $M/M/\infty$ queue that would describe the copy number if the gene remained always active. In particular it is easy to see that the IAP process is stochastically dominated by a Poisson distribution with parameter ν/δ, and hence admits moments of any order. Now the environment process (status of the gene) as well as the underlying birth and death processes in any environment state have exponential rates of convergence to equilibrium. It can be proved that the IAP process inherits that property. Details of proofs are not given here for sake of brevity. This exponential rate of convergence to equilibrium is of high practical interest. It makes possible to collect experimental data the statistical distribution of which should be close to the asymptotic distribution.

In this section, an explicit expression for the generating function of the stationary distribution is first derived. Then the explicit expression of the parameters in terms of the first three moments of that distribution is given. Thus the parameters of the model admit convergent estimators in terms of a sample of copy numbers. Let

$$\{p_{i,n}\ ,\ i \in \{0,1\}, n \in \mathbb{N}\}$$

be the stationary distribution of the IAP process. The $p_{i,n}$'s are a constant solution to the Kolmogorov system of equations (5). The generating functions g_0 and g_1 are analogous to those of the previous section.

$$g_0(z) = \sum_{n=0}^{\infty} z^n p_{i,n} \quad \text{and} \quad g_1(z) = \sum_{n=0}^{\infty} z^n p_{0,n}$$

Naturally, the next relations hold.

$$g_0(z) = \lim_{t\to\infty} G_0(z,t) \text{ and } g_1(z) = \lim_{t\to\infty} G_1(z,t)$$

The functions g_0 and g_1 satisfy the following system of differential equations deduced from (6).

$$\begin{cases} \delta(z-1)g_0'(z) = -\lambda g_0(z) + \mu g_1(z) \\ \\ \delta(z-1)g_1'(z) = \lambda g_0(z) - \mu g_1(z) - \nu(1-z)g_1(z) \end{cases} \tag{15}$$

The values at $z=1$ of g_0 and g_1 are the asymptotic probabilities for the gene to be inactive and active respectively

$$g_0(1) = \frac{\mu}{\lambda+\mu} \quad \text{and} \quad g_1(1) = \frac{\lambda}{\lambda+\mu} \ .$$

The difficulty in solving this system comes from its singularity at $z=1$. From (15) the following second order differential equation in g_0 is derived.

$$-\delta^2(1-z)g_0''(z) + \delta\left(\lambda+\mu+\delta+\nu(1-z)\right)g_0'(z) - \lambda\nu g_0(z) = 0 \tag{16}$$

Through the change of variable $x = 1-z$ and setting $h(z) = g(1-x)$, (16) becomes

$$\delta^2 x h''(x) + \delta\left(\lambda+\mu+\delta+\nu x\right)h'(x) + \lambda\nu h(x) = 0. \tag{17}$$

The ratios a, b, and c are introduced to simplify the expressions.

$$a = \frac{\lambda+\mu+\delta}{\delta} \ , \ b = \frac{\nu}{\delta} \ , \ c = \frac{\lambda\nu}{\delta^2}$$

The values of the derivatives of h at $x=0$ are obtained by induction from (17).

$$\forall n \geq 1, \ h^{(n)}(0) = (-1)^n \ \frac{c(c+b)\cdots(c+b(n-1))}{a(a+1)\cdots(a+n-1)} \cdot \frac{\mu}{\lambda+\mu} \tag{18}$$

Hence, the expansion of g_0 as a power series in $z-1$ is

$$g_0(z) = \frac{\mu}{\lambda+\mu}\left[1 + \sum_{n=1}^{\infty}\frac{1}{n!}\frac{c(c+b)\cdots(c+b(n-1))}{a(a+1)\cdots(a+n-1)}(z-1)^n\right] \ . \tag{19}$$

This power series is a particular case of a degenerate hypergeometric function [5]. The use of the classical notation ${}_1F_1$ for these functions leads to

$$g_0(z) = \frac{\mu}{\lambda+\mu} \ {}_1F_1\left(\frac{\lambda}{\delta}, \frac{\lambda+\mu+\delta}{\delta}; \frac{\nu}{\delta}(z-1)\right) \ .$$

The expression of $g_1(z)$ is deduced from the first equation of (15). Now the generating function of the asymptotic distribution of the copy number is

$$g(z) = g_0(z) + g_1(z) \quad .$$

We get

$$g(z) = 1 + \sum_{n=1}^{\infty} \frac{1}{n!} \frac{c(c+b)\cdots(c+b(n-1))}{(a-1)a\cdots(a+n-2)}(z-1)^n \quad . \tag{20}$$

Again, g is a degenerate hypergeometric function.

$$g(z) = {}_1F_1\left(\frac{\lambda}{\delta}, \frac{\lambda+\mu}{\delta}; \frac{\nu}{\delta}(z-1)\right) .$$

The expression of g as a power series in $(z-1)$ makes it particularly easy to compute the exponential moments, denoted by e_n, of the asymptotic distribution. Let X be the number of copies at equilibrium, and denote by $\mathbb{E}$ the mathematical expectation, then

$$e_n = \mathbb{E}\left[X(X-1)\cdots(X-n+1)\right] = g^{(n)}(1) = \frac{c(c+b)\cdots(c+b(n-1))}{(a-1)a\cdots(a+n-2)} .$$

Thus, the mean of the asymptotic distribution is

$$\mathbf{E} = e_1 = \frac{c}{a-1} = \frac{\lambda}{\lambda+\mu}\frac{\nu}{\delta} .$$

Its variance is

$$\mathbf{V} = e_2 + e_1 - e_1^2 = \frac{\lambda}{\lambda+\mu}\frac{\nu}{\delta} + \frac{\lambda\mu}{(\lambda+\mu)^2}\frac{\nu^2}{\delta(\lambda+\mu+\delta)} .$$

These two expressions have already been obtained in section 2 as the limits of $\mathbf{E}(t)$ and $\mathbf{V}(t)$.

Since the parameters of this model of gene expression cannot be directly measured, their estimation on the basis of a given sample of the asymptotic distribution (which is the only kind of experimental data potentially available) would be a significant step forward. *A priori* one should estimate four parameters λ, μ, ν, and δ. However, the model is homogeneous in time. Multiplying the scale of time by a constant is equivalent to dividing the four parameters by the same constant and this operation does not affect the asymptotic distribution. Hence, only three of the four parameters can be estimated from the asymptotic distribution. The scale of time is adjusted so as to have one of the parameters, say δ, equal to 1. This assumption is equivalent to choosing as time unit the average life time of the P molecule. Let r_1, r_2 and r_3 be the successive ratios of the first three exponential moments.

$$
\begin{aligned}
r_1 &= e_1 &= \frac{c}{a-1} &= \frac{\lambda}{\lambda+\mu}\cdot\nu \\
r_2 &= e_2/e_1 &= \frac{c+b}{a} &= \frac{\lambda+1}{\lambda+\mu+1}\cdot\nu \\
r_3 &= e_3/e_2 &= \frac{c+2b}{a+1} &= \frac{\lambda+2}{\lambda+\mu+2}\cdot\nu
\end{aligned}
$$

It is now easy to solve the above equation in λ, μ, and ν and get

$$
\begin{aligned}
\lambda &= \frac{2r_1(r_3-r_2)}{r_1r_2-2r_1r_3+r_2r_3} , \\
\mu &= \frac{2(r_2-r_1)(r_1-r_3)(r_3-r_2)}{(r_1r_2-2r_1r_3+r_2r_3)(r_1-2r_2+r_3)} , \\
\nu &= \frac{-r_1r_2+2r_1r_3-r_2r_3}{r_1-2r_2+r_3} .
\end{aligned}
$$

Provided that some sample $(X_1,\ldots,X_N)$ of the asymptotic distribution is available the natural estimators of the first three moments are

$$
\begin{aligned}
\hat{e}_1 &= \frac{1}{N}\sum_{i=1}^{N} X_i , \\
\hat{e}_2 &= \frac{1}{N}\sum_{i=1}^{N} X_i(X_i-1) , \\
\hat{e}_2 &= \frac{1}{N}\sum_{i=1}^{N} X_i(X_i-1)(X_i-2) .
\end{aligned}
$$

By the law of large numbers, these are convergent estimators of e_1, e_2 and e_3 respectively. λ, μ and ν were shown to be a function, ϕ , of e_1, e_2 and e_3.

$$(\lambda,\mu,\nu) = \phi\,(e_1,e_2,e_3)$$

Since ϕ is a continuous function, the statistics

$$\left(\hat{\lambda},\hat{\mu},\hat{\nu}\right) = \phi\,(\hat{e}_1,\hat{e}_2,\hat{e}_3)$$

is a convergent estimator of the 3-tuple of the parameters (λ,μ,ν).

References

[1] M.S. Bartlett. Recent investigations involving stochastic population models. *Adv. in Appl. Probab.*, 16(3):449–470, 1984.

[2] A.T. Barusha-Reid. *Elements of the theory of Markov Processes and their Applications.* McGraw-Hill, New York, 1960.

[3] W.C. Cogburn, R.; Torrez. Birth and death processes with random environments in continuous time. *J. Appl. Probab.*, 18(1):19–30, 1981.

[4] M. Eisen, M.; Tainiter. Stochastic variations in queuing processes. *Operat. Res.*, 11:922–927, 1963.

[5] I.M. Gradsteyn, I.S.; Ryzhik. *Tables of Integrals, Series and Products, 4th ed.* Academic Press, New York, 1965.

[6] Wolfram Research Inc. *Mathematica.* Wolfram Research Inc, 1992.

[7] M.E. Karmeshu; Thompson. The one-server markov queue in a random environment. *Bul. Cal. Math. Soc.*, 85:203–208, 1993.

[8] M.S.H. Ko. The dose dependence of glucocorticoid-inducible gene expression results from changes in the number of transcriptionnaly active template. *EMBO J.*, 9:2835–2842, 1990.

[9] M.S.H. Ko. A stochastic model for gene induction. *J. Theor. Biol.*, 153:181–194, 1991.

[10] M.S.H. Ko. Induction mechanism of a single gene molecule : stochastic or deterministic ? *Bioessays*, 14:341–346, 1992.

[11] J. Maddox. Towards more measurement in biology. *Nature*, 368:9, 1994.

[12] W. Smith, W.L.; Wilkinson. On branching processes in random environments. *Ann. Math. Statist.*, 40:814–827, 1969.

[13] D. Stoyan. *Comparison Methods for Queues and other Stochastic Models.* Wiley, New York, 1983.

[14] W.C. Torrez. The birth and death chain in a random environment: instability and extinction theorems. *Ann. Probab.*, 6(6):1026–1043, 1978.

[15] T.D. von Hippel, P.H. ; Yager et al. *Quantitative aspects of the transcriptional cycle in E. Coli.*, chapter 7. Cold Spring Harbour Laboratory Press, 1992.