

Uma aplicação de equações diferenciais à um problema elementar em probabilidade

Artur Lopes

Vamos apresentar a seguir uma aplicação interessante da Teoria dos Sistemas de Equações Lineares a um problema elementar em probabilidade.

Considere uma máquina que pode se apresentar em dois estados: funcionando ou estragada. Associamos o número 1 à máquina funcionar e o número 0 à máquina estar estragada. Vamos denotar por X_t este estado no tempo t e que será ou zero ou um.

A máquina em consideração pode estragar ou (sendo consertada) voltar a funcionar a qualquer momento. Sendo assim, analisando o comportamento da máquina a partir de um tempo $t = 0$ o estado da máquina seria descrito por um comportamento X_t do tipo que mostra a fig. 1

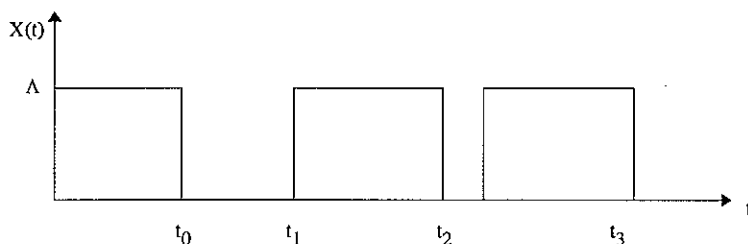


Fig. 1

Gostaríamos de saber a probabilidade da máquina estar ou não funcionando em um tempo t . Esta probabilidade será dada por $p_1(t)$ enquanto $p_0(t)$ vai denotar a probabilidade da máquina estar estragada.

Portanto, $p_1(t) + p_0(t) = 1, \forall t \geq 0$.

Esta probabilidade se torna importante quando estivermos considerando uma fábrica com várias máquinas de um certo tipo (digamos 10.000), e desejarmos saber a cada tempo t quantas delas (qual proporção ou equivalentemente qual probabilidade) estariam funcionando e quantas estariam estragadas.

Inicialmente, o único conceito que vamos necessitar de probabilidade é o conceito de probabilidade condicional.

Definição: A probabilidade condicional que ocorrendo B , vá acontecer A é dada por

$$P(A/B) = \frac{P(A \text{ e } B)}{P(B)}.$$

Por exemplo, podemos pensar que A significa ser aprovado no vestibular e que B denota ser aluno de um colégio X . Sendo assim, a probabilidade condicional de ser aprovado no vestibular, dado que estudou no colégio X é o quociente do número de alunos do colégio X aprovados no vestibular, dividido pelo número de alunos do colégio X que fizeram vestibular.

Vamos supor que o processo em consideração é Markoviano e homogêneo. Somente com estas hipóteses conseguiremos obter a expressão analítica das probabilidades da máquina estar ou não estar estragada em tempo t . Em breve, ao longo de nossas considerações vamos explicar o que exatamente significa isto.

Vamos denotar por

$$P_{ji}(t) = P[X_t = i \mid X_0 = j],$$

$i, j \in \{0, 1\}$ onde X_t denota o estado da máquina 1 (funcionando) ou 0 (estragada).

Note que

$$P_{j0}(t) + P_{j1}(t) = 1 \tag{1}$$

$$\forall j \in \{0, 1\}, \forall t \in \mathbf{R}.$$

Isto se deve ao fato que t fixo e j fixo

$$P_{j0}(t) + P_{j1}(t) = P(X_t = 0 \mid X_0 = j) + P(X_t = 1 \mid X_0 = j) =$$

$$\frac{P(X_t = 0 \text{ e } X_0 = j)}{P(X_0 = j)} + \frac{P(X_t = 1 \text{ e } X_0 = j)}{P(X_0 = j)} = \frac{P(X_0 = j)}{P(X_0 = j)} = 1$$

A penúltima igualdade se deve ao fato que os conjuntos $\{X_t = 0 \text{ e } X_0 = j\}$ e $\{X_t = 1 \text{ e } X_0 = j\}$ são disjuntos e sua união é o conjunto $\{X_0 = j\}$.

A evolução temporal da matriz

$$\begin{pmatrix} P_{00}(t) & P_{10}(t) \\ P_{01}(t) & P_{11}(t) \end{pmatrix}$$

é extremamente importante para identificar $p_0(t)$ e $p_1(t)$.

Note que se no tempo $t = 0$ as probabilidades da máquina não estar ou estar funcionando são respectivamente $p_0(0) = p_0$ e $p_1(0) = p_1$, então

$$\begin{aligned} p_0(t) &= P(X_t = 0) = P(X_t = 0 \text{ e } X_0 = 0) + P(X_t = 0 \text{ e } X_0 = 1) = \\ &= \frac{P(X_t = 0 \text{ e } X_0 = 0)}{P(X_0 = 0)} P(X_0 = 0) + \frac{P(X_t = 0 \text{ e } X_0 = 1)}{P(X_0 = 1)} P(X_0 = 1) = \\ &= P_{00}(t)p_0 + P_{10}(t)p_1 \end{aligned}$$

Da mesma maneira

$$\begin{aligned} p_1(t) &= P(X_t = 1) = P(X_t = 1 \text{ e } X_0 = 0) + P(X_t = 1 \text{ e } X_0 = 1) = \\ &= \frac{P(X_t = 1 \text{ e } X_0 = 0)}{P(X_0 = 0)} P(X_0 = 0) + \frac{P(X_t = 1 \text{ e } X_0 = 1)}{P(X_0 = 1)} P(X_0 = 1) = \\ &= P_{01}(t)p_0 + P_{11}(t)p_1 \end{aligned}$$

De maneira sintética obtemos

$$\begin{pmatrix} p_0(t) \\ p_1(t) \end{pmatrix} = \begin{pmatrix} P_{00}(t) & P_{10}(t) \\ P_{01}(t) & P_{11}(t) \end{pmatrix} \begin{pmatrix} p_0 \\ p_1 \end{pmatrix} \quad (2)$$

Sendo assim, descobrir a evolução de

$$P(t) = \begin{pmatrix} P_{00}(t) & P_{10}(t) \\ P_{01}(t) & P_{11}(t) \end{pmatrix}$$

vai resolver o problema que estamos interessados que é afinal de contas encontrar $p_0(t)$ e $p_1(t)$.

A questão é, portanto, encontrar quem é $P(t)$.

Note que

$$P(0) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

pois $P(X_0 = j | X_0 = i) = \delta_{ij}$. Ainda, a soma de cada coluna de $P(t)$ é igual a 1 (ver (1)).

Vamos fazer algumas hipóteses sobre o modelo. A primeira é que o sistema é homogêneo, isto é,

$$P(X_{t_2} = i \mid X_{t_1} = j) = P(X_{t_2-t} = i \mid X_{t_1-t} = j)$$

$$\forall t, t_1, t_2 \in \mathbf{R}, t_2 > t_1 > 1.$$

O sistema é homogêneo, se a translação temporal por t não altera as propriedades de transição como descrito acima.

A segunda hipótese é que o sistema seja Markoviano, ou seja, que

$$P(X_t = i \mid X_{t_1} = j, X_{t_2} = k) = P(X_t = i \mid X_{t_2} = k)$$

onde $t > t_2 > t_1$. Em outras palavras, o sistema é Markoviano se não tem memória, e tudo o que acontece no futuro depende apenas do último instante (t_2) (do presente) e não do passado (t_1).

Com estas hipóteses podemos obter a equação de Chapman-Kolmogorov que é expressa por

$$P_{ij}(t + \tau) = P_{io}(t) \cdot P_{oj}(\tau) + P_{il}(t) P_{lj}(\tau) \quad \text{para todo } i, j \in [0, 1] \quad (3)$$

Uma forma matricial compacta, envolvendo o produto das matrizes $P(t)$ e $P(\tau)$, permite escrever a expressão acima como

$$P(t) P(\tau) = P(t + \tau).$$

Vamos agora demonstrar (3).

Ora

$$P_{ij}(t + \tau) = P(X_{t+\tau} = j \mid X_0 = i) =$$

$$P(X_{t+\tau} = j, X_t = 0 \mid X_0 = i) + P(X_{t+\tau} = j, X_t = 1 \mid X_0 = i) =$$

$$\frac{P(X_{t+\tau} = j, X_t = 0, X_0 = i)}{P(X_0 = i)} + \frac{P(X_{t+\tau} = j, X_t = 1, X_0 = i)}{P(X_0 = i)}$$

$$\frac{P(X_{t+\tau} = j, X_t = 0, X_0 = i)}{P(X_t = 0, X_0 = i)} \frac{P(X_t = 0, X_0 = i)}{P(X_0 = i)}$$

$$+ \frac{P(X_{t+\tau} = j, X_t = 1, X_0 = i)}{P(X_t = 1, X_0 = i)} \frac{P(X_t = 1, X_0 = i)}{P(X_0 = i)}$$

$$= P(X_{t+\tau} = j \mid X_t = 0, X_0 = i) P(X_t = 0 \mid X_0 = i)$$

$$+ P(X_{t+\tau} = j \mid X_t = 1, X_0 = i) P(X_t = 1 \mid X_0 = i)$$

$$= P(X_{t+\tau} = j \mid X_t = 0) P(X_t = 0 \mid X_0 = i)$$

$$+ P(X_{t+\tau} = j \mid X_t = 1) P(X_t = 1 \mid X_0 = i).$$

A última igualdade segue do fato que o processo é Markoviano.

Sendo assim,

$$P_{ij}(t + \tau) = P(X_{t+\tau} = j \mid X_t = 0) P_{i0}(t) + P(X_{t+\tau} = j \mid X_t = 1) P_{i1}(t) =$$

$$P(X_\tau = j \mid X_0 = 0) P_{i0}(\tau) + P(X_\tau = j \mid X_0 = 1) P_{i1}(\tau).$$

A última igualdade segue do fato que o processo é homogêneo.

Concluimos, portanto, que vale (3), ou seja, para quaisquer $i, j \in \{0, 1\}$

$$P_{ij}(t + \tau) = P_{i0}(t) P_{0j}(\tau) + P_{i1}(t) P_{1j}(\tau).$$

A última hipótese que faremos é que existem as derivadas $P_{00}'(0)$, $P_{01}'(0)$, $P_{10}'(0)$ e $P_{11}'(0)$.

Para t fixo, derivando em relação a τ a equação de Chapman-Kolmogorov obteremos para $i, j \in \{0, 1\}$

$$\frac{d}{d\tau} P_{ij}(t + \tau) = P_{i0}(t) P_{0j}'(\tau) + P_{i1}(t) P_{1j}'(\tau).$$

Considerando $\tau = 0$, para qualquer $i, j \in \{0, 1\}$ temos que

$$P_{ij}'(t) = \left. \frac{dP_{ij}(t + \tau)}{d\tau} \right|_{\tau=0} = P_{i0}(t) P_{0j}'(0) + P_{i1}(t) P_{1j}'(0).$$

Estas equações podem ser escritas como

$$P_{00}'(t) = P_{00}(t) P_{00}'(0) + P_{01}(t) P_{10}'(0) \quad (4)$$

$$P_{01}'(t) = P_{00}(t) P_{01}'(0) + P_{01}(t) P_{11}'(0) \quad (5)$$

e

$$P_{10}'(t) = P_{10}(t) P_{00}'(0) + P_{11}(t) P_{10}'(0) \quad (6)$$

$$P_{11}'(t) = P_{10}(t) P_{01}'(0) + P_{11}(t) P_{11}'(0) \quad (7)$$

Note que vale $P_{00}'(0) \leq 0$ porque $P_{00}(0) = 1$ e por sua vez $P_{00}(t)$ tem que ser menor ou igual a 1.

Note também que $P_{11}'(0) \leq 0$ pela mesma razão.

Ora $\forall t \in \mathbf{R}$, $P_{00}(t) + P_{01}(t) = 1$ e $P_{10}(t) + P_{11}(t) = 1$, portanto derivando temos que se $P_{01}'(0) = \mu \geq 0$, concluiremos que $P_{00}'(0) = -\mu$. Da mesma maneira se $P_{10}'(0) = \lambda > 0$, teremos que $P_{11}'(0) = -\lambda$.

As equações (4) e (5) correspondem a

$$\begin{pmatrix} P_{00}'(t) \\ P_{01}'(t) \end{pmatrix} = \begin{pmatrix} P_{00}'(0) & P_{10}'(0) \\ P_{01}'(0) & P_{11}'(0) \end{pmatrix} \begin{pmatrix} P_{00}(t) \\ P_{01}(t) \end{pmatrix} = \begin{pmatrix} -\mu & \lambda \\ \mu & -\lambda \end{pmatrix} \begin{pmatrix} P_{00}(t) \\ P_{01}(t) \end{pmatrix},$$

e as equações (6) e (7) correspondem a

$$\begin{pmatrix} P_{10}'(t) \\ P_{11}'(t) \end{pmatrix} = \begin{pmatrix} P_{00}'(0) & P_{10}'(0) \\ P_{01}'(0) & P_{11}'(0) \end{pmatrix} \begin{pmatrix} P_{10}(t) \\ P_{11}(t) \end{pmatrix} = \begin{pmatrix} -\mu & \lambda \\ \mu & -\lambda \end{pmatrix} \begin{pmatrix} P_{10}(t) \\ P_{11}(t) \end{pmatrix}.$$

Sendo assim, $(P_{00}(t), P_{01}(t))$ e $(P_{10}(t), P_{11}(t))$ correspondem a soluções da mesma equação diferencial

$$x'(t) = \begin{pmatrix} -\mu & \lambda \\ \mu & -\lambda \end{pmatrix} x(t),$$

$x(t) \in \mathbf{R}^2$, com condições iniciais respectivamente $(1,0)$ e $(0,1)$ (isto porque $(P_{00}(0), P_{01}(0)) = (1,0)$ e $(P_{10}(0), P_{11}(0)) = (0,1)$).

Algumas palavras sobre sistemas de equações diferenciais lineares cabem aqui.

A solução $x(t)$ da equação diferencial linear (ver [S])

$$x'(t) = \begin{pmatrix} a & b \\ c & d \end{pmatrix} x(t)$$

com condição inicial $x(0) = x_0 \in \mathbf{R}^2$ é dada por

$$x(t) = e^{\begin{pmatrix} a & b \\ c & d \end{pmatrix} t} x(0).$$

Para saber o que é a exponencial e^A de uma matriz A (como mencionado acima) precisamos da próxima definição:

Definição: A definição de e^A , onde A é uma matriz n por n é

$$I + A + \frac{1}{2!} A^2 + \frac{1}{3!} A^3 + \dots + \frac{1}{n!} A^n + \dots$$

onde I é a matriz identidade.

Esta definição estende naturalmente a definição usual de exponencial e^x de um número x na reta real.

Exemplo: $e^{\begin{pmatrix} x_1 & 0 \\ 0 & x_2 \end{pmatrix}} = \begin{pmatrix} e^{x_1} & 0 \\ 0 & e^{x_2} \end{pmatrix}$

É de fácil verificação.

Para fins práticos de calcular exponencial de uma matriz, a seguinte proposição é muito útil.

Proposição: Se $A = BCB^{-1}$ então $e^A = Be^C B^{-1}$.

Esta proposição (de fácil demonstração) sugere que o problema de calcular a exponencial de uma matriz A diagonalizável, por exemplo suponha que

$$A = B \begin{pmatrix} x_1 & 0 \\ 0 & x_2 \end{pmatrix} B^{-1},$$

recai em saber encontrar a matriz B que diagonaliza a matriz A . Isto porque, pelo exemplo acima, já sabemos como calcular a exponencial de uma matriz diagonal. A matriz B pode ser obtida facilmente como a matriz cujas colunas são os autovetores de A .

No que vamos considerar a seguir a matriz A será a matriz

$$\begin{pmatrix} -\mu & \lambda \\ \mu & -\lambda \end{pmatrix} t.$$

O polinômio característico da matriz

$$\begin{pmatrix} -\mu & \lambda \\ \mu & -\lambda \end{pmatrix}$$

é $x^2 + (\lambda + \mu)x$. Logo os autovalores x da matriz acima são respectivamente $x_1 = 0$ e $x_2 = -(\lambda + \mu)$.

Um autovetor associado a 0 é

$$\left(\frac{\lambda}{\mu + \lambda}, \frac{\mu}{\mu + \lambda} \right)$$

e um autovetor associado a $-(\lambda + \mu)$ é $(1, -1)$.

A matriz das derivadas $(P'_{ij}(0))_{i,j \in \{0,1\}}$ é chamada de gerador infinitesimal do processo.

Logo

$$\begin{pmatrix} -\mu & \lambda \\ \mu & -\lambda \end{pmatrix} = \begin{pmatrix} \frac{\lambda}{\mu + \lambda} & 1 \\ \frac{\mu}{\mu + \lambda} & -1 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & -(\mu + \lambda) \end{pmatrix} \begin{pmatrix} \frac{\lambda}{\mu + \lambda} & 1 \\ \frac{\mu}{\mu + \lambda} & -1 \end{pmatrix}^{-1} =$$

$$\begin{pmatrix} \frac{\lambda}{\mu+\lambda} & 1 \\ \frac{\mu}{\mu+\lambda} & -1 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & -(\mu+\lambda) \end{pmatrix} \begin{pmatrix} 1 & 1 \\ \frac{\mu}{\mu+\lambda} & -\frac{\lambda}{\mu+\lambda} \end{pmatrix}.$$

Vamos agora usar alguns resultados mencionados acima sobre soluções de equações diferenciais lineares.

A partir da última proposição, multiplicando por t e tomando exponencial de matriz de ambos os lados da última expressão obteremos

$$\begin{aligned} e^{\begin{pmatrix} -\mu & \lambda \\ \mu & -\lambda \end{pmatrix}t} &= \begin{pmatrix} \frac{\lambda}{\mu+\lambda} & 1 \\ \frac{\mu}{\mu+\lambda} & -1 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & e^{-(\mu+\lambda)t} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ \frac{\mu}{\mu+\lambda} & -\frac{\lambda}{\mu+\lambda} \end{pmatrix} = \\ &= \begin{pmatrix} \frac{\lambda}{\lambda+\mu} + \frac{\mu}{\lambda+\mu} e^{-(\mu+\lambda)t} & \frac{\lambda}{\lambda+\mu} - \frac{\lambda}{\lambda+\mu} e^{-(\mu+\lambda)t} \\ \frac{\mu}{\lambda+\mu} - \frac{\mu}{\lambda+\mu} e^{-(\mu+\lambda)t} & \frac{\mu}{\lambda+\mu} - \frac{\lambda}{\lambda+\mu} e^{-(\mu+\lambda)t} \end{pmatrix} \end{aligned}$$

Logo como $P_{00}(0) = 1$ e $P_{01} = 0$

$$\begin{pmatrix} P_{00}(t) \\ P_{01}(t) \end{pmatrix} = e^{\begin{pmatrix} -\mu & \lambda \\ \mu & -\lambda \end{pmatrix}t} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{\lambda}{\lambda+\mu} + \frac{\mu}{\lambda+\mu} e^{-(\mu+\lambda)t} \\ \frac{\mu}{\lambda+\mu} - \frac{\mu}{\lambda+\mu} e^{-(\mu+\lambda)t} \end{pmatrix}$$

e, portanto,

$$\lim_{t \rightarrow \infty} (P_{00}(t), P_{01}(t)) = \left(\frac{\lambda}{\mu + \lambda}, \frac{\mu}{\lambda + \mu} \right). \quad (8)$$

Este resultado pode ser interpretado da seguinte maneira: a máquina começando estragada, vai ter no limite quando t tende a infinito uma probabilidade

$$\frac{\lambda}{\mu + \lambda} \text{ de estar estragada e } \frac{\mu}{\mu + \lambda} \text{ de estar funcionando.}$$

Analogamente,

$$\begin{pmatrix} P_{10}(t) \\ P_{11}(t) \end{pmatrix} = e^{\begin{pmatrix} -\mu & \lambda \\ \mu & -\lambda \end{pmatrix} t} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{\lambda}{\lambda + \mu} - \frac{\lambda e^{-(\mu + \lambda)t}}{\lambda + \mu} \\ \frac{\mu}{\lambda + \mu} + \frac{\lambda e^{-(\mu + \lambda)t}}{\lambda + \mu} \end{pmatrix}$$

e da mesma maneira

$$\lim_{t \rightarrow \infty} (P_{10}(t), P_{11}(t)) = \left(\frac{\lambda}{\lambda + \mu}, \frac{\mu}{\lambda + \mu} \right). \quad (9)$$

A conclusão é que se a máquina estiver funcionando no tempo $t = 0$ ela vai ter para grandes valores de t uma probabilidade que será aproximadamente

$$\frac{\lambda}{\mu + \lambda} \text{ de estar estragada e } \frac{\mu}{\mu + \lambda} \text{ de estar funcionando.}$$

Dada as probabilidades iniciais da máquina não estar ou estar funcionando, ou seja, as condições iniciais p_0 e p_1 então as probabilidades $p_0(t)$ e $p_1(t)$ da máquina estar estragada ou funcionando em tempo t são dadas por

$$\begin{pmatrix} p_0(t) \\ p_1(t) \end{pmatrix} = e^{\begin{pmatrix} -\mu & \lambda \\ \mu & -\lambda \end{pmatrix} t} \begin{pmatrix} p_0 \\ p_1 \end{pmatrix}.$$

A partir de (2) obtemos

$$(p_0(t), p_1(t)) = p_0(P_{00}(t), P_{01}(t)) + p_1(P_{10}(t), P_{11}(t)).$$

Concluimos finalmente de (8) e (9) que independente da condição inicial (das probabilidades p_0 e p_1) o comportamento da máquina para valores grandes de t é tal que

$$\lim_{t \rightarrow \infty} (p_0(t), p_1(t)) = p_0 \left(\frac{\lambda}{\mu + \lambda}, \frac{\mu}{\mu + \lambda} \right) + p_1 \left(\frac{\lambda}{\mu + \lambda}, \frac{\mu}{\mu + \lambda} \right) = \left(\frac{\lambda}{\mu + \lambda}, \frac{\mu}{\mu + \lambda} \right),$$

porque $p_0 + p_1 = 1$.

Logo, a probabilidade da máquina estragar a longo prazo é

$$\frac{\lambda}{\mu + \lambda},$$

independente das probabilidades iniciais.

Note que o resultado acima é o análogo para tempo contínuo do que acontece com uma cadeia de Markov com tempo discreto (ver [L]).

Na figura 2 mostramos o espaço de fase (ou seja o desenho esquemático no plano das trajetórias $x(t) \in \mathbb{R}^2$) da equação diferencial

$$x'(t) = \begin{pmatrix} -\mu & \lambda \\ \mu & -\lambda \end{pmatrix} x(t).$$

O conjunto denotado por Δ representa o conjunto dos (p_1, p_2) tal que $p_1 + p_2 = 1$ e p_1, p_2 são positivos. O conjunto Δ como vimos antes é invariante pelo fluxo associado à equação diferencial.

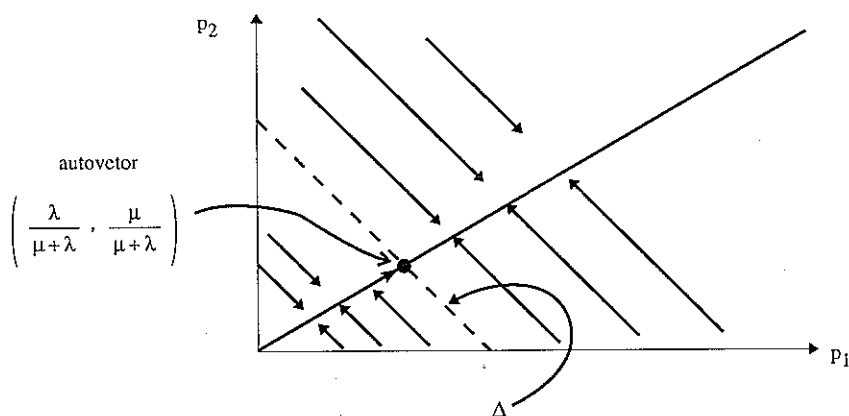


Fig. 2

Um resultado mais preciso sobre a velocidade de convergência de $(p_0(t), p_1(t))$ ao equilíbrio $\left(\frac{\lambda}{\lambda + \mu}, \frac{\mu}{\lambda + \mu}\right)$ é facilmente obtido como

$$\frac{1}{2} \left(\left| p_0(t) - \frac{\lambda}{\mu + \lambda} \right| + \left| p_1(t) - \frac{\mu}{\mu + \lambda} \right| \right) \leq C e^{-(\mu + \lambda)t},$$

onde C é uma constante.

Ou seja a convergência tem velocidade exponencial e o valor limite $\left(\frac{\lambda}{\lambda + \mu}, \frac{\mu}{\lambda + \mu}\right)$ é obtido (em valor aproximado) muito rapidamente.

Referimos ao leitor [C,D] para considerações mais gerais sobre o problema.

Referências

- [C,D] D. Clarke e A. Disney - Introdução aos Processos Estocásticos
- [L] A. Lopes, Matemática Universitária, N. 18, junho 1995, pp. 33-41
- [S] J. Sottomayor, Lições de Equações Diferenciais Ordinárias, Projeto Euclides, IMPA, RJ.

*Instituto de Matemática - UFRGS
Av. Bento Gonçalves, 9500
Prédio A1 - Campus do Vale
91.509-900 - Porto Alegre - RS*