MS777 - Projeto Supervisionado I

Lei dos grandes números em sistemas aparentemente aleatórios

Instituto de Matemática, Estatística e Computação Científica

Universidade Estadual de Campinas

Aluno: Eric Lopes, R.A.: 076624

Orientadora: Prof$^{\underline{a}}$. Dra. Marina Vachkovskaia

Sumário

Resumo

O objetivo deste projeto é avaliar a Lei dos Grandes Números (LGN) em todas as suas formas tanto teórica como experimentalmente. Para isso foi dividido em duas partes:

Na primeira parte estudamos teoricamente a LGN analisando todas as suas formulações e hipóteses, Assim como todos os teoremas e resultados importantes para as verificações experimentais.

Na segunda parte do projeto avaliamos experimentalmente a validade da LGN para alguns casos em que nem todas as hipóteses são satisfeitas e avaliamos as peculiaridades de cada caso interessante.

1 Introdução e Motivações

Os teoremas limites são os resultados mais importantes da teoria da probabilidade, dos quais a Lei dos Grandes Números faz parte [13] , por causa de suas aplicações em todas as áreas da ciência é largamente usada por pesquisadores como ferramenta para obter conclusões, estimativas e previsões estatísticas [2, 14, 15, 9, 16].

Porém, para se enquadrar nas hipóteses da LGN muitas vezes são feitas suposições e simplificações em torno dos dados originais (como por exemplo, que todas as variáveis são independentes e identicamente distribuidas (i.i.d.) [9]).

A pergunta natural a partir das evidências é: mesmo após várias suposições em torno de um conjunto de dados, podemos esperar que os valores obtidos reflitam a verdade ?

Com essa motivação, iremos avaliar os possíveis relaxamentos das hipóteses das LGN e utilização em sistemas que não são totalmente aleatórios ou onde algumas hipóteses para os principais teoremas que garantem que sequências seguem a lei dos grandes números não são verificadas.

Estas abordagens são relevantes pois vários estudos em economia, física, medicina e previsões climáticas [19, 3, 6, 11, 5] tem comportamento modelado por sistemas dinâmicos (determinísticos, com variações em torno de um ponto), porém são simplificados e modelados como variáveis aleatórias. A segunda abordagem tem relevância quando não há certeza sobre o comportamento dos dados, podendo ser por exemplo, dependentes e usados como independentes, ou mesmo ilimitados.

2 Estudo Teórico

Para o estudo teórico omiti as provas mais longas dos teoremas por causa da extensa bibliografia que já aborda essa questão. Caso o leitor esteja interessado, [18] apresenta uma completa lista de referências disponíveis.

2.1 Lei dos Grandes Números (LGN)

Seja uma sequência $\{\chi_n, n \geq 1\}$ de variáveis aleatórias definidas no espaço de probabilidade $(\Omega, \mathcal{F}, \mathbf{P})$ vamos definir $S_n = \sum_{i=1}^{n} \chi_i$, $a_k = \mathbf{E}(\chi_k)$, $A_n = \mathbf{E}(S_n) = \sum_{i=1}^{n} a_i$

2.1.1 Lei Fraca dos Grandes Números

Dizemos que $\{\chi_n\}$ satisfaz a lei fraca dos grandes números se $\frac{1}{n} \cdot S_n - \frac{1}{n} \cdot A_n \xrightarrow{p} 0$ quando $n \to \infty$, ou seja, para todo $\epsilon > 0$ temos

$$\lim_{n\to\infty} \mathbf{P}\left(\left|\frac{1}{n} \cdot S_n - \frac{1}{n} \cdot A_n\right| \geq \epsilon\right) = 0 \tag{2.1}$$

2.1.2 Lei Forte dos Grandes Números

Dizemos que $\{\chi_n\}$ satisfaz a lei forte dos grandes números se $\frac{1}{n} \cdot S_n - \frac{1}{n} \cdot A_n \xrightarrow{q.c.} 0$ quando $n \to \infty$.

$$\mathbf{P}\left(\omega : \lim_{n\to\infty} \left[\frac{1}{n} \cdot S_n - \frac{1}{n} \cdot A_n\right] = 0\right) = 1 \tag{2.2}$$

2.2 Teorema de Khintchine

Seja $\{\chi_n, n \geq 1\}$ uma sequência de variáveis aleatórias i.i.d. com $\mathbf{E}(|\chi_1|) < \infty$. então a sequência satisfaz a lei fraca dos grandes números e $\frac{1}{n} \cdot S_n \xrightarrow{p} a$ quando $n \to \infty$ onde $a = \mathbf{E}(\chi_1)$.

2.3 Teorema de Markov

Suponha que $\{\chi_n, n \geq 1\}$ é uma sequência arbitrária de variáveis aleatórias em que se verifica a seguinte condição (Condição de Markov).

$$\left(\frac{1}{n^2}\right) \cdot \mathbf{Var}\left[\sum_{\mathbf{i=1}}^{\mathbf{n}} \chi_{\mathbf{i}}\right] \underset{\mathbf{n}\to\infty}{\to} \mathbf{0} \tag{2.3}$$

Então $\{\chi_n\}$ satisfaz a Lei fraca dos grandes números.

2.4 Teorema (Kolmogorov e Khintchine)

Dado $\mathbf{E}(\chi_n) = 0$. Se $\sum_{n=1}^{\infty} \mathbf{E}\left(\chi_n^2\right) < \infty$ então $\sum_{n=1}^{\infty} \chi_n < \infty$ com probabilidade 1.[17]

2.5 Teoremas de Kolmogorov

2.5.1 Teorema de Kolmogorov 1

Seja $\{\chi_n, n \geq 1\}$ uma sequência de variáveis aleatórias i.i.d. A existência de $\mathbf{E}(|\chi_1|)$ é condição necessária e suficiente para que a sequência $\{\chi_n\}$ satisfaça a lei forte dos grandes números e $\frac{1}{n} S_n \xrightarrow{q.c.} a$ quando $n \to \infty$ onde $a = \mathbf{E}(\chi_1)$

2.5.2 Teorema de Kolmogorov 2

Seja $\{\chi_n, n \geq 1\}$ uma sequência de variáveis aleatórias independentes com segundo momento finito. Com $\sigma_n^2 = \mathbf{Var}(\chi_\mathbf{n}) < \infty$, $n \geq 1$. Suponha que a seguinte condição (Condição de Kolmogorov) é satisfeita.

$$\sum_{n=1}^{\infty} \frac{\sigma_n^2}{n^2} < \infty \tag{2.4}$$

Então a sequência $\{\chi_n\}$ satisfaz a lei forte dos grandes números, ou seja $\frac{S_n - \mathbf{E}(S_n)}{n} \overset{q.c.}{\to} 0$.

Por ser um dos resultados mais utilizados neste trabalho provarei o Teorema de Kolmogorov 2. Para isso precisaremos de dois lemas:

Lema (Toeplitz): Seja $\{a_n\}$ uma sequência de números não negativos, $b_n = \sum_{i=1}^{n} a_i$, $b_n > 0$ e $b_n \underset{n\to\infty}{\to} \infty$. Seja $\{\chi_n\}$ uma sequência de números que convergem para χ então:

$$\frac{1}{b_n} \cdot \sum_{j=1}^{n} a_j \cdot \chi_j \to \chi \tag{2.5}$$

Prova: Seja $\epsilon > 0$ e $n_0 = n_0(\epsilon)$ de tal maneira que $|\chi_n - \chi| \leq \frac{\epsilon}{2}$ para todo $n \geq n_0$.
Escolha $n_1 > n_0$ daí

$$\frac{1}{b_{n_1}} \cdot \sum_{j=1}^{n_0} |\chi_n - \chi| < \frac{\epsilon}{2}$$

então, para $n > n_1$,
$\left|\frac{1}{b_n} \cdot \sum_{j=1}^{n} a_j \cdot \chi_j - \chi\right| \leq \frac{1}{b_n} \cdot \sum_{j=1}^{n} a_j \cdot |\chi_j - \chi| = \frac{1}{b_n} \cdot \sum_{j=1}^{n_0} a_j \cdot |\chi_j - \chi| + \frac{1}{b_n} \cdot \sum_{j=n_0+1}^{n} a_j \cdot |\chi_j - \chi|$
$\leq \frac{1}{b_{n_1}} \cdot \sum_{j=1}^{n_0} a_j \cdot |\chi_j - \chi| + \frac{1}{b_n} \cdot \sum_{j=n_0+1}^{n} a_j \cdot |\chi_j - \chi| \leq \frac{\epsilon}{2} + \frac{b_n - b_{n_0}}{b_n} \cdot \frac{\epsilon}{2} \leq \epsilon$
Completando a prova do lema.

Lema (Kronecker): Seja $\{b_n\}$ uma sequência de números crescentes, $b_n \underset{n\to\infty}{\to} \infty$, e seja $\{\chi_n\}$ uma sequência de números tal que $\sum \chi_n$ convirja. Então:

$$\frac{1}{b_n} \cdot \sum_{j=1}^{n} b_j \cdot \chi_j \underset{n\to\infty}{\to} 0 \tag{2.6}$$

Prova: Seja $b_0 = 0$, $S_0 = 0$, $S_n = \sum_{j=1}^{n} \chi_j$. Então:
$\sum_{j=1}^{n} b_j \cdot \chi_j = \sum_{j=1}^{n} b_j \cdot (S_j - S_{j-1}) = b_n \cdot S_n - b_0 \cdot S_0 - \sum_{j=1}^{n} S_{j-1} \cdot (b_j - b_{j-1})$
e então

$$\frac{1}{b_n} \cdot \sum_{j=1}^{n} b_j \cdot \chi_j = S_n - \frac{1}{b_n} \cdot \sum_{j=1}^{n} S_{j-1} \cdot a_j \to 0$$

Desde que $S_n \to \chi$, pelo lema de Toeplitz temos

$$\frac{1}{b_n} \cdot \sum_{j=1}^{n} S_{j-1} \cdot a_j \to \chi$$

Finalizando a prova.

Prova do Teorema de Kolmogorov 2: Dado que

$$\frac{S_n - \mathbf{E}(S_n)}{n} = \frac{1}{n} \cdot \sum_{k=1}^{n} k \left(\frac{\chi_k - \mathbf{E}(\chi_k)}{k} \right)$$

Uma condição suficiente para que $\frac{S_n - \mathbf{E}(S_n)}{n} \stackrel{q.c.}{\to} 0$, (pelo lema de Kronecker) é que a série $\sum \left[\frac{\chi_k - \mathbf{E}(\chi_\mathbf{k})}{k} \right] \stackrel{q.c.}{\to} C$ (C uma constante). Como, por hipótese $\sum \frac{\mathbf{Var}(\chi_n)}{n^2} < \infty$ e χ_n tem segundo momento finito, temos

$$\sum_{n=1}^{\infty} \frac{\mathbf{Var}(\chi_n)}{n^2} = \sum_{n=1}^{\infty} \frac{\mathbf{E}\left((\chi_k - \mathbf{E}(\chi_k))^2 \right)}{n^2} = \sum_{n=1}^{\infty} \mathbf{E}\left(\frac{(\chi_k - \mathbf{E}(\chi_k))^2}{n^2} \right) = \sum_{n=1}^{\infty} \mathbf{E}\left(\left(\frac{\chi_k - \mathbf{E}(\chi_k)}{n} \right)^2 \right) < \infty$$

$$\sum_{n=1}^{\infty} \mathbf{E}\left(\left(\frac{\chi_k - \mathbf{E}(\chi_k)}{n} \right)^2 \right) < \infty \stackrel{2.4}{\Longrightarrow} \sum_{n=1}^{\infty} \frac{\chi_k - \mathbf{E}(\chi_k)}{n} < \infty$$

Finalizando a demonstração.

2.6 Resultados importantes para as verificações experimentais

2.6.1 Condição de Kolmogorov, suficiente mas não necessária

A condição de Kolmogorov (2.5.2) é suficiente, mas não necessária para a lei forte dos grandes números.

Como por exemplo seja $\{\chi_n, n \geq 1\}$ uma sequência de variáveis aleatórias independentes em que seja verificado $\mathbf{P}[\chi_n = \pm 100] = \frac{1}{2} \cdot (1 - 2^{-n})$, $\mathbf{P}[\chi_n = 2^n] = \mathbf{P}[\chi_n = -2^n] = 2^{-(n+1)}$.

Neste caso $\mathbf{E}(\chi_n) = 0$, $\sigma_n^2 = \mathbf{Var}(\chi_n) = 100 \cdot (1 - 2^{-n}) + 2^n$, portanto a condição de Kolmogorov (2.5.2) não é satisfeita. Porém $\{\chi_n\}$ obedece a lei forte dos grandes números (a prova pode ser obtida em [18]) .

Como segundo exemplo temos $\{\chi_n, n \geq 1\}$ de forma que $\mathbf{P}\left[\chi_n = \pm\sqrt{n/\ln(n)}\right] = \frac{1}{2}$, dessa forma $\mathbf{Var}(\chi_\mathbf{n}) = \mathbf{n}/\ln(\mathbf{n})$ quebrando a condição de Kolmogorov (2.5.2)

2.6.2 Condição de Kolmogorov, a melhor condição possível para a lei dos grandes números

A condição de Kolmogorov (2.5.2) $\sum_{n=1}^{\infty} \sigma_n^2/n^2 < \infty$ é a melhor condição possível para a lei forte dos grandes números, pois podemos construir vários exemplos de sequências tais que $\sum_{n=1}^{\infty} \sigma_n^2/n^2 = \infty$ e que não satisfaçam a LGN, ou seja, se tentarmos uma situação mais geral que a condição de Kolmogorov, essa condição não será suficiente para garantir a LGN. [18]

2.6.3 A série $\sum_{n=1}^{\infty} 1/(\ln(n+1)\cdot(n+1))$ não é convergente

Prova: Sem perda de generalidade, podemos fazer $x = n + 1$ e verificar a não convergência de $\sum_{x=2}^{\infty} 1/(\ln(x)\cdot x)$.

$$\sum_{x=2}^{\infty} 1/(\ln(x)\cdot x) = \left(\frac{1}{2 \cdot \ln(2)} \right) + \left(\frac{1}{3 \cdot \ln(3)} + \frac{1}{4 \cdot \ln(4)} \right) + \left(\frac{1}{5 \cdot \ln(5)} + \cdots + \frac{1}{8 \cdot \ln(8)} \right) + \cdots$$

$$\sum_{x=2}^{\infty} 1/(\ln(x)\cdot x) \geq \left(\frac{1}{2 \cdot \ln(2)} \right) + \left(\frac{2}{4 \cdot \ln(4)} \right) + \left(\frac{4}{8 \cdot \ln(8)} \right) + \cdots$$

$$\sum_{x=2}^{\infty} {}^{1}\!/\!_{(\ln(x)\cdot x)} \geq \left(\frac{2^0}{2^1\cdot\ln(2^1)}\right)+\left(\frac{2^1}{2^2\cdot\ln(2^2)}\right)+\left(\frac{2^2}{2^3\cdot\ln(2^3)}\right)+\cdots=\sum_{n=1}^{\infty}\frac{2^{n-1}}{2^n\cdot n\cdot\ln(2)}$$

$$\sum_{x=2}^{\infty} {}^{1}\!/\!_{(\ln(x)\cdot x)} \geq \frac{1}{2\cdot\ln(2)}\cdot\sum_{n=1}^{\infty}\frac{2^n}{2^n\cdot n}=k\cdot\sum_{n=1}^{\infty}\frac{1}{n}\to\infty$$

Provei que a série em questão é maior que a série harmônica, portanto pelo teste da comparação [10] a série é divergente.

2.6.4 A série $\sum_{n=1}^{\infty} 1/\left(\ln(n+1)^2\cdot(n+1)\right)$ é absolutamente convergente

Prova: Sem perda de generalidade, podemos fazer $x = n+1$ e verificar a convergência de $\sum_{x=2}^{\infty} 1/\left(\ln(x)^2\cdot x\right)$.

expandimos a série em:

$$\sum_{x=2}^{\infty} 1/\left(\ln(x)^2\cdot x\right)=\left(\frac{1}{2\cdot\ln(2)^2}+\frac{1}{3\cdot\ln(3)^2}\right)+\left(\frac{1}{4\cdot\ln(4)^2}+\cdots+\frac{1}{7\cdot\ln(7)^2}\right)+\left(\frac{1}{8\cdot\ln(8)^2}+\cdots+\frac{1}{15\cdot\ln(15)^2}\right)+\cdots$$

$$\sum_{x=2}^{\infty} 1/\left(\ln(x)^2\cdot x\right)\leq\left(\frac{1}{2\cdot\ln(2)^2}+\frac{1}{2\cdot\ln(2)^2}\right)+\left(\frac{1}{4\cdot\ln(4)^2}+\cdots+\frac{1}{4\cdot\ln(4)^2}\right)+\left(\frac{1}{8\cdot\ln(8)^2}+\cdots+\frac{1}{8\cdot\ln(8)^2}\right)+\cdots$$

$$\sum_{x=2}^{\infty} 1/\left(\ln(x)^2\cdot x\right)\leq\frac{2}{2\cdot\ln(2)^2}+\frac{4}{4.\cdot\ln(4)^2}+\frac{8}{8\cdot\ln(8)^2}+\frac{16}{16\cdot\ln(16)^2}\cdots=\sum_{n=1}^{\infty}\frac{2^n}{2^n\cdot\ln(2^n)^2}=\sum_{n=1}^{\infty}\frac{1}{\ln(2^n)^2}$$

$$\sum_{x=2}^{\infty} 1/\left(\ln(x)^2\cdot x\right)\leq\sum_{n=1}^{\infty}\frac{1}{\ln(2^n)^2}=\sum_{n=1}^{\infty}\frac{1}{(n\cdot\ln(2))^2}=\frac{1}{\ln(2)^2}\cdot\sum_{n=1}^{\infty}\frac{1}{n^2}=k\cdot\sum_{n=1}^{\infty}\frac{1}{n^2}\to c$$

Provei que a série em questão é menor que uma série convergente [10], portanto pelo teste da comparação a série é convergente. Como todos os valores dela são positivos, provei que é absolutamente convergente.

2.6.5 Esperança infinita implica em variância infinita

Seja χ uma variável aleatória tal que $\mathbf{E}(\chi)=\sum_{x\in\chi}x\cdot\mathbf{P}(x)=\infty$ por definição, temos que $\sigma^{\mathbf{2}}(\chi)=\mathbf{E}\left((\chi-E(\chi))^{\mathbf{2}}\right)=E\left(\chi^{\mathbf{2}}\right)-E(\chi)^{\mathbf{2}}=\sum_{x\in\chi}x^{\mathbf{2}}\cdot\mathbf{P}(x)-\left(\sum_{x\in\chi}x\cdot P(x)\right)^{\mathbf{2}}$, mas $\left(\sum_{x\in\chi}x\cdot\mathbf{P}(x)\right)^{\mathbf{2}}=\infty$ e $\sum_{x\in\chi}x^{\mathbf{2}}\cdot\mathbf{P}(x)\geq\left(\sum_{x\in\chi}x\cdot\mathbf{P}(x)\right)^{\mathbf{2}}$ portanto concluímos que $\sigma^{\mathbf{2}}(\chi)=\infty$. (Esse resultado é uma aplicação do Teorema 5.3 de [7])

Em particular, variáveis aleatórias com esperança infinita não obedecem a condição de Kolmogorov 2.5.2, assim, não há garantia que ela cumpra a LGN.

2.6.6 Uma variável com variância 0 é um valor determinístico

Seja χ uma variável aleatória com variância $\sigma^{\mathbf{2}}=0$, $\mathbf{k}>0$ e usando desigualdade de Chebyshev [13] temos

$\mathbf{P}(|\chi-\mu|\geq\mathbf{k})\leq\frac{\sigma^{\mathbf{2}}}{\mathbf{k}^{\mathbf{2}}}$ daí $\mathbf{P}(|\chi-\mu|>\mathbf{0})=\mathbf{0}$ portanto $\chi=\mu$ com probabilidade 1 e assim é determinístico.

3 Verificações experimentais

3.1 A LGN teórica funciona experimentalmente independente da distribuição

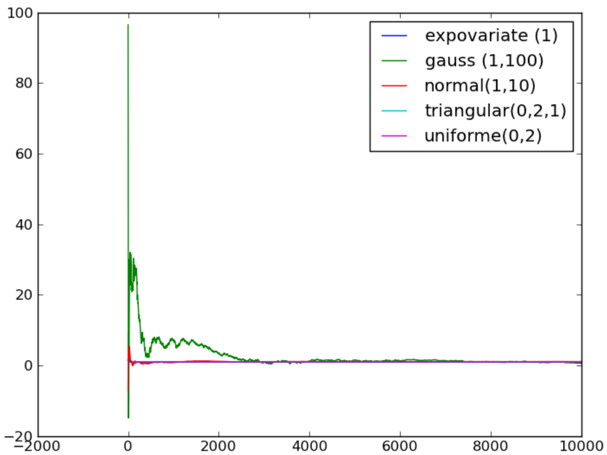

Figura 3.1: $n \text{x} \frac{S_n}{n}$ com sequências $\{\chi_n\}$ em várias distribuições i.i.d. e $n \in [1, 10000]$.

A figura 3.1 mostra o gráfico com aplicação experimental do Teorema de Kolmogorov 1 , do Teorema de Markov e do Teorema de Khintchine.

3.2 Verificando a condição de Kolmogorov

A condição de Kolmogorov (2.5.2) apresenta um ramo interessante para estudos experimentais, como veremos a seguir.

3.2.1 A condição de Kolmogorov não é necessária para que $\{\chi_n\}$ obedeça a LGN

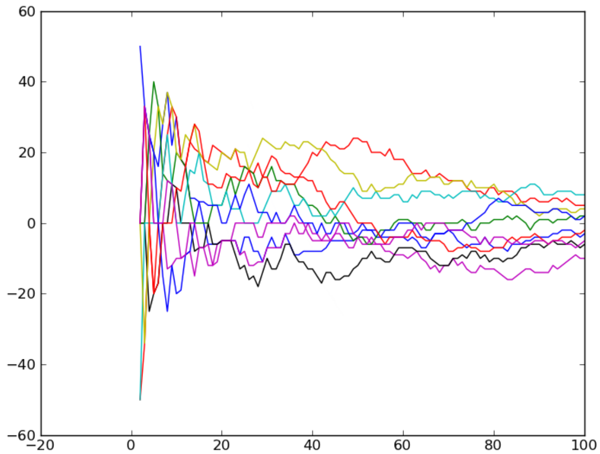

Figura 3.2: nx$\frac{S_n}{n}$ com sequência $\{\chi_n\}$ relançada diversas vezes, satisfazendo 2.6.1 .

A figura 3.2 mostra o gráfico de uma sequência que não satisfaz a condição de Kolmogorov (2.5.2), porém obedece a lei dos grandes números, dada por $\{\chi_n\}$ tal que $\mathbf{P}\left[\chi_n = \pm 100\right] = \frac{1}{2} \cdot (1 - 2^{-n})$, $\mathbf{P}\left[\chi_n = 2^n\right] = \mathbf{P}\left[\chi_n = -2^n\right] = 2^{-(n+1)}$ (2.6.1). Um ponto importante a ser ressaltado é que os valores retornados nessa distribuição são da ordem de 2^n o que justifica o número de lançamentos $n \propto 100$, valores maiores se tornam inviáveis, por isso relançamos a mesma distribuição diversas vezes, observa-se que a média dos valores realmente tende a média de $\{\chi_n\}$ apesar de não cumprir a condição de Kolmogorov (2.5.2).

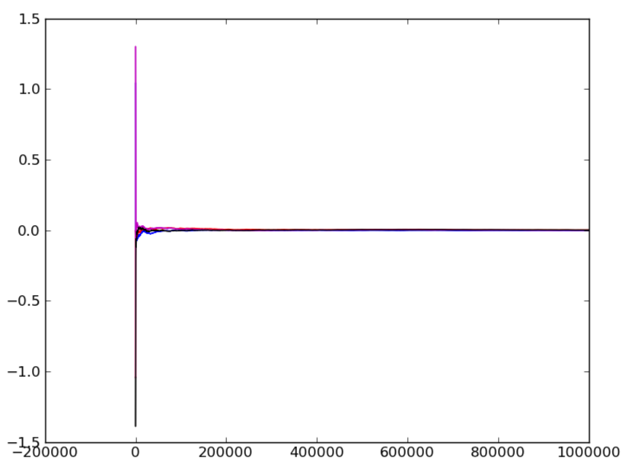

Figura 3.3: nx$\frac{S_n}{n}$ com sequência $\{\chi_n\}$ relançada diversas vezes, satisfazendo 2.6.1 .

A figura 3.3 tem distribuição de forma que $\mathbf{P}\left(\pm\sqrt{n/\ln(n)}\right) = 1/2$, não satisfazendo a condição de Kolmogorov 2.6.3 e 2.5.2, mas convergindo de forma adequada. (mais detalhes em [18]).

3.2.2 Na teoria $n \to \infty$, mas na prática $n \to K$, isso importa?

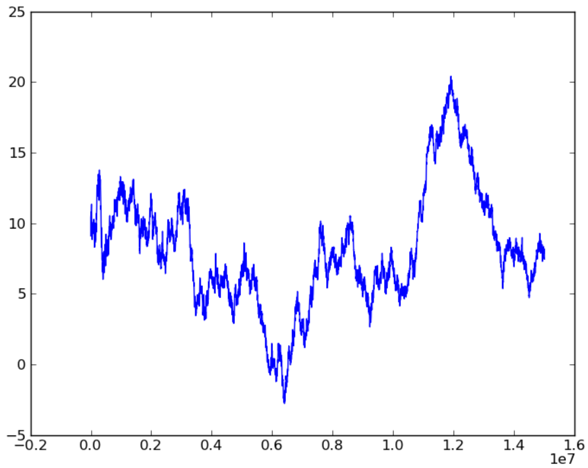

Figura 3.4: $n\text{x}\frac{S_n}{n}$ com sequência $\{\chi_n\}$ obedecendo 2.5.2 usando sequência da forma 2.6.4 e $n \propto 15000000$

Esse tópico é importante para mostrar que nem sempre é viável o tempo de convergência de uma sequência de variáveis aleatórias, nesse exemplo, usamos uma sequência de variáveis com distribuição normal com média 10 e variância $n/\ln(n)^2$ cumprindo a condição de Kolmogorov (2.5.2) $\sum_{n=2}^{\infty} \sigma_n^2/n^2 = \sum_{n=2}^{\infty} n/(n^2.\ln(n)^2) = \sum_{n=2}^{\infty} 1/(n.\ln(n)^2) < \infty$ (2.6.4).

À primeira vista ela não obedece a LGN, mas repetindo a plotagem diversas vezes, observamos o seguinte:

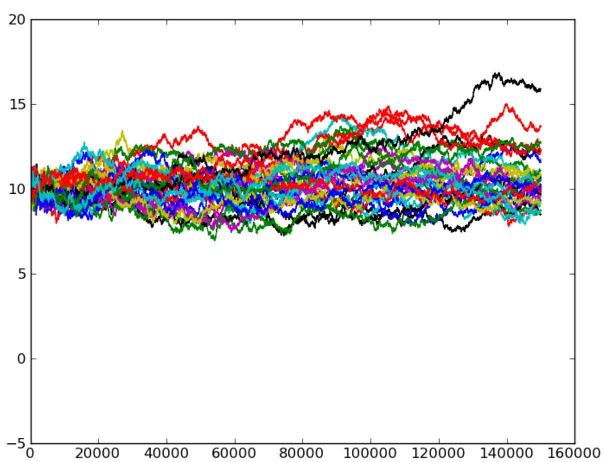

Figura 3.5: $n\text{x}\frac{S_n}{n}$ com sequências $\{\chi_n\}$ obedecendo 2.5.2 usando sequência da forma 2.6.4 e $n \propto 150000$

Agora observamos claramente na figura 3.5 que a média dos valores tende a média da distribuição (no caso $\mathbf{E}(\chi_i) = 10$), mas pelo gráfico, não podemos garantir que essa sequência segue a lei forte dos grandes números, apenas que aparentemente segue a lei fraca. Assim para a maioria das aplicações, como por exemplo encontrar a média do "grau de saúde" de um grupo de pacientes, não teríamos pacientes o suficiente para que n chegasse próximo o bastante de infinito e a média dos valores realmente refletisse alguma coisa, Mostrando que temos que tomar cuidado ao escolher a abordagem das médias para um experimento aleatório.

3.2.3 Quanto a variáveis aleatórias sem um limite definido.

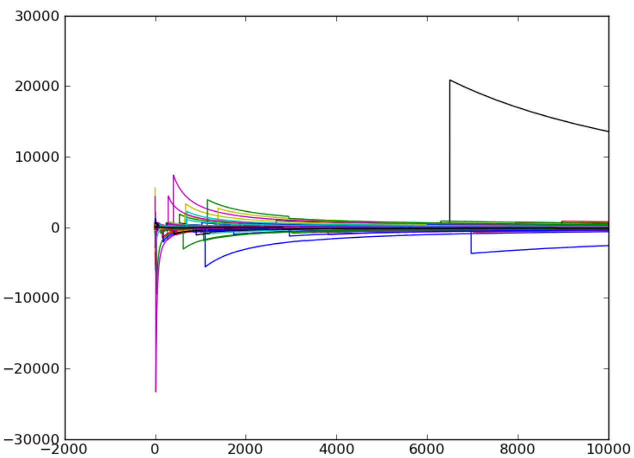

Figura 3.6: $nx\frac{S_n}{n}$ com sequências $\{\chi_n\}$ usando distribuição uniforme inversa no intervalo $[-0.1, 0.1]$ e aproximadamente quinhentos relançamentos com $n \propto 10000$

Neste exemplo utilizamos uma sequência $\{\chi_i\}$ definida como $\chi_i = \frac{1}{\gamma_i}$ e $\gamma_i \sim Uniforme(-0.1, 0.1)$, dessa forma temos uma variável que pode assumir ∞ um número limitado de vezes, com a propriedade de a cada novo lançamento se tornar mais difícil o novo valor aleatório alterar a distribuição de forma significativa (ϵ) ou seja, dado $max_k\{\chi_i\}$ sendo o maior valor obtido até o k-ésimo lançamento, então $\mathbf{P}(|\chi_{k+1}| > |\epsilon \cdot k \cdot max_k\{\chi_i\}|) = \mathbf{P}\left(|\gamma_{k+1}| < \left|\frac{1}{\epsilon \cdot k \cdot max_k\{\chi_i\}}\right|\right) \to 0$, o que pode gerar erros ocasionais para $n \ll \infty$. Como visto no gráfico 3.6 em pouco mais de quinhentos testes com a mesma distribuição, apenas duas vezes o resultado divergiu significativamente da média esperada e esse número tende a ser menor conforme n aumenta. Por fim, cabe observar que a sequência $\{\chi_i\}$ obedece a condição de Kolmogorov (2.5.2), por ser contínua e com um número finito de ∞ o que implica que satisfaz a lei forte dos grandes números, mesmo podendo assumir valores infinitos.

Dado $\{\lambda_i\}$ a mesma sequência $\{\chi_i\}$ discreta, a situação muda completamente, simplesmente não faz sentido um gráfico com o valor infinito, sua média será indefinida, terá variância indefinida (2.6.5) e não poderemos obter nenhuma informação útil.

O resultado prático deste tópido está na incapacidade de termos valores contínuos na maioria de nossos experimentos, deixando apenas duas alternativas para situações como essa.

A primeira alternativa é fazer uma linearização contínua dos valores do experimento, caso os trechos indefinidos sejam locais, eliminando assim pontos infinitos.

A segunda alternativa é eliminar os valores indefinidos do experimento, considerando-os "erro", caso faça sentido esse tipo de abordagem.

Caso eliminar os pontos indefinidos não fizer sentido no experimento em questão, então é necessário rever a abordagem e necessidade de algo como a lei dos grandes números para o estudo do experimento.

3.2.4 Quão dependente as variáveis do sistema podem ser e ainda satisfazer a lei dos grandes números ?

Em [1] o autor define sequências "$L^1 - Mix\acute{a}veis$" que sob certas circuntâncias (caso uniformemente integráveis) satisfazem a lei fraca dos grandes números, porém foge ao nosso objetivo entrar no ramo abstrato da matemática, com poucas relações experimentais, que é o foco deste trabalho.

Em [4] temos um resultado mais palpável, se a correlação entre as sequências de variáveis aleatórias for estritamente menor que 1 então o artigo apresenta condições suficientes de momento para que a sequência obedeça a lei dos grandes números.

Portanto, podemos afirmar pouca coisa sobre variáveis dependentes, em geral há mais interesse e quase que a totalidade dos teoremas em variáveis independentes. Para variáveis dependentes, apenas temos o resultado experimental que se forem fracamente dependentes, ou seja, quando $n \to \infty$, a influência tende a k então ela geralmente obedece a lei dos grandes números, mas pequenas alterações de sua dependência podem influenciar de maneira significativa sua distribuição como observado abaixo:

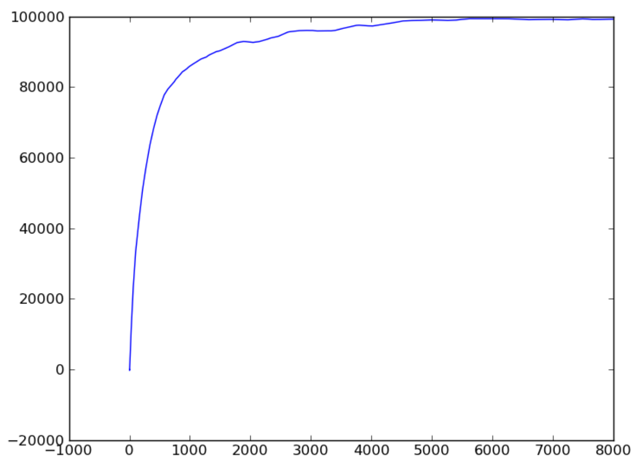

Figura 3.7: nx$\frac{S_n}{n}$ com sequência $\{\chi_n\}$ dependente de $\{\chi_{n-1}\}$ e $\{\chi_1\}$, de forma que $\mathbf{E}(\chi_i) = \mathbf{E}(\chi_1) + \mathbf{E}(\chi_{i-1})/1.01$ e $n \propto 8000$, Dado por $\chi_n = \varphi\left(\chi_1 + \frac{\chi_{n-1}}{1.01}, 1000\right)$ e $\varphi\left(\mu, \sigma^2\right) \sim Gauss\left(\mu, \sigma^2\right)$.

Algoritmo 1 Algoritmo de criação do gráfico 3.7

```
vSoma = 0
vListaY = vazia
vListaX = vazia
vXAnterior = 1
vAux = 0
vX1 = 1000
PARA n EM [1,8000] FAÇA:
    vAux = random.gauss(vX1+vXAnterior/1.01, 1000)
    vSoma += vAux
    vXAnterior = vAux
    ADICIONE vSoma/n EM vListaY
    ADICIONE n EM vListaX
MOSTRE O GRÁFICO (vListaX, vListaY)
```

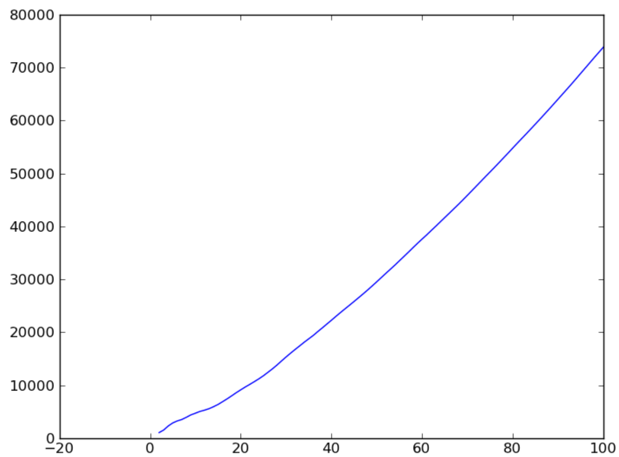

Figura 3.8: nx$\frac{S_n}{n}$ com sequência $\{\chi_n\}$ dependente de $\{\chi_{n-1}\}$ e $\{\chi_1\}$, de forma que $\mathbf{E}(\chi_i) = \mathbf{E}(\chi_1) + 1.01 \cdot \mathbf{E}(\chi_{i-1})$ e $n \propto 100$, Dado por $\chi_n = \varphi(\chi_1 + 1.01\chi_{n-1}, 1000)$ e $\varphi(\mu, \sigma^2) \sim Gauss(\mu, \sigma^2)$.

Algoritmo 2 Algoritmo de criação do gráfico 3.8

```
vSoma = 0
vListaY = vazia
vListaX = vazia
vXAnterior = 1
vAux = 0
vX1 = 1000
PARA n EM [1,100] FAÇA:
    vAux = random.gauss(vX1+vXAnterior*1.01, 1000)
    vSoma += vAux
    vXAnterior = vAux
    ADICIONE vSoma/n EM vListaY
    ADICIONE n EM vListaX
MOSTRE O GRÁFICO (vListaX, vListaY)
```

Estes resultados são interessantes, por isso vale a pena observar que para a primeira distribuição temos $\mathbf{E}\left(\chi_i\right) = \mathbf{E}\left(\chi_1\right) + \mathbf{E}\left(\chi_{i-1}\right)/1.01 = \mathbf{E}\left(\chi_1\right) + \frac{\mathbf{E}(\chi_1)+\frac{\mathbf{E}(\chi_1)+\cdots/1.01}{1.01}}{1.01}$ dessa forma, a influência de χ_{i-k} em χ_i equivale a $\frac{\mathbf{E}(\chi_1)+\mathbf{E}(\chi_{i-k-1})/1.01}{1.01^k}$, assim para valores de k grandes (variáveis distantes de χ_i), a influência tende a 0.

Analogamente, para a segunda distribuição temos que a influência tende ao infinito para valores de k grandes.

Como visto, devemos tomar cuidado especial quando considerar as variáveis do experimento como sendo independentes, pois pequenas alterações no modo em que age sua dependência podem fazer com que não obedeça nenhum teorema limite.

3.3 Em situações modeladas por sistemas dinâmicos, podemos usar o sistema como variáveis aleatórias?

A resposta é sim, devido a 2.6.6, podemos por exemplo transformar o sistema discreto:

$X_{n+1} = f(X_n, Y_n, \ldots)$, $Y_{n+1} = g(X_n, Y_n \ldots), \ldots$ em

$\chi_{n+1/\chi_n,\psi_n,\ldots} = N(f(\chi_n, \psi_n, \ldots), \sigma_\chi^2)$, $\psi_{n+1/\chi_n,\psi_n,\ldots} = N(g(\chi_n, \psi_n, \ldots), \sigma_\psi^2) \ldots$

Assim modelamos um sistema dinâmico (discreto) de forma probabilística em torno de valores próximos ao esperado pelo sistema, o que é útil (e faz mais sentido) para a maioria das situações reais, porém, cabe a observação que dificilmente seguirá a lei dos grandes números (3.2.4), por causa de sua dependência estrutural forte de todas as variáveis anteriores. Exceto quando porventura estamos interessados em médias localizadas, como por exemplo "média de temperatura às segundas feiras do mês de maio em dias nublados", obviamente uma informação localizada não necessita ser tão exata, basta que tenhamos variáveis iniciais o suficiente para que a influência das demais se torne irrelevante (3.2.4), em situações reais é suficiente por exemplo "temperatura média no mês de maio" .

3.4 E quando o sistema dinâmico chega ao limite do caos?

Como observado em 3.2.4 não podemos esperar que variáveis altamente dependentes sigam a lei dos grandes números, por convenção um sistema caótico é um sistema dinâmico em que pequenas variações nas condições iniciais alteram de forma praticamente imprevisível a solução do sistema, ou seja, é um sistema como em 3.3 que depende fortemente das variáveis de que é dependente.

Em geral, o resultado esperado coincide com a experimentação [12, 8, 19, 15, 20]. Mostrando que não podemos esperar que sistemas caóticos sigam a lei dos grandes números.

4 Conclusões

Com toda a base experimental e teórica, podemos concluir que os teoremas limite, em especial a LGN explorada neste projeto, é uma ferramenta sólida para auxiliar pesquisas em todas as áreas, desde que sejam cumpridas as hipóteses necessárias para sua utilização. E mesmo quando essas hipóteses são cumpridas é necessário verificar a viabilidade do experimento para se chegar a afirmações concisas a respeito do objeto de estudo.

Dentre os pontos principais cabe destacar:

- Deve-se evitar usar variáveis dependentes, e somente quando estritamente necessário, é essencial verificar se essa dependência é pequena o suficiente (para que não influencie nos resultados de forma inesperada). Neste contexto, suficiente depende do experimento em questão.
- Se for necessário usar variáveis dependentes, verificar se não faz mais sentido ao sistema como um todo, medir valores localizados.
- Ao iniciar um experimento deve-se tomar uma estimativa da convergência, utilizando métodos de inferência e outros não abordados neste texto, para verificar se é viável fazer o experimento. Caso não seja, deve-se utilizar outra abordagem.
- Se a informação da variância do sistema estiver acessível, sempre utilize a condição de Kolmogorov para garantir a convergência para a lei forte dos grandes números.

Referências

[1] Donald W. K. Andrews. Laws of large numbers for dependent non-identically distributed random variables. April 1987.

[2] Matteo Barigozzi. On the method of social sciences and statistics. April 2007.

[3] Richard T. Boylan. Laws of large numbers for dynamical systems with randomly matched individuals. November 1990.

[4] W. Bryc and W. Smolenski. Moment conditions form almost sure convergence of weakly correlated random variables. October 1993.

[5] Marco A. Palumbo Cabral and Felipe S. Figueiredo. A dynamical systems approach for malaria epidemiology in the amazon region. 2006.

[6] A. Hodgkin and A. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. 1952.

[7] Barry R. James. *Probabilidade: Um curso em nível intermediário*. LTC, 2 edition, 1996.

[8] Kunihiro Kaneko. Globally coupled chaos violates the law of large numbers but not the central-limit theorem. September 1990.

[9] Haya Kaspi and Kavita Ramanan. Law of large numbers limits for many-server queues.

[10] Elon Lages Lima. *Análise Real*, volume 1. IMPA, Rio de Janeiro, 10 edition, 2009.

[11] Raymond S. Bradley Michael E. Mann and Malcolm K. Hughes. Global-scale temperature patterns and climate forcing over the past six centuries. 1998.

[12] Arkady S. Pikovsky and Jürgen Kurths. Do globally coupled maps really violate the law of large numbers? March 1994.

[13] Sheldon M. Ross. *A First Course in Porbability*, volume 1. Prentice-Hall, United States of America, 5 edition, 1998.

[14] Manuel S. Santos and Adrian Peralta-Alva. Accuracy of simulations for stochastic dynamic models.

[15] Manuel S. Santos and Adrian Peralta-Alva. Generalized laws of large numbers for the simulation of dynamic economies. March 2010.

[16] Peter Sedlmeier and Gerd Gigerenzer. Was bernoulli wrong? on intuitions about sample size. December 2000.

[17] A. N. Shiryaev. *Probability*. Springer.

[18] Jordan M. Stoyanov. *Counter Examples in Probability*. John Wiley and Sons LTDA., United States of America, 2 edition, 1996.

[19] M. Azam Sudeshna Sinha, D. Biswas and S. V. Lawande. Nonstatistical behavior of higher-dimensional coupled systems. September 1992.

[20] Guo-Wei Wei Ying-Cheng Lai, Zonghua Liu and Choy-Heng Lai. Shadowability of statistical averages in chaotic systems. *Physical Review Letters*, 89(18), October 2002.