A DESIGUALDADE DE CHEBYCHEV

HÉLIO BERNARDO LOPES[1]

Resumo. A desigualdade de Chebychev constitui um resultado de grande importância na estimação da probabilidade de acontecimentos oriundos de experiências aleatórias de que se desconhece a distribuição da correspondente população. Pode, em todo o caso, surgir de modos diversos, naturalmente equivalentes, como também em distintos domínios que têm no acaso a sua raiz essencial. O presente texto expõe, precisamente, a importância da Desigualdade de Chebychev, mas vista à luz da diversidade da sua aplicação.

Quando se pretende calcular a probabilidade de poder ocorrer determinado acontecimento e se conhece a distribuição probabilística que está em causa no problema, não se colocam dificuldades particulares. É o que sucede, por exemplo, com uma variável aleatória X, contínua, cuja função densidade de probabilidade seja:

$$f_X(x) = \begin{cases} \frac{1}{2}x & \Leftarrow \quad x \in [0,2] \\ \\ 0 & \Leftarrow \quad x \notin [0,2]. \end{cases}$$

O valor médio de X - o seu primeiro momento ordinário, portanto - e o seu segundo momento ordinário valem, respetivamente:

$$E[X] = \int_0^2 x \frac{1}{2} x dx = \frac{4}{3}$$

$$E\left[X^2\right] = \int_0^2 x^2 \frac{1}{2} x dx = 2$$

pelo que a variância de X toma o valor:

$$V[X] = E\left[X^2\right] - \left(E[X]\right)^2 = 2 - \left(\frac{4}{3}\right)^2 = \frac{2}{9}.$$

Admita-se, agora, que se pretende calcular a seguinte probabilidade:

$$P\left(\left|X - \frac{4}{3}\right| < \frac{\sqrt{2}}{3}\right).$$

Ora, tendo-se:

$$\left|X - \frac{4}{3}\right| < \frac{\sqrt{2}}{3} \quad \Leftrightarrow \quad \frac{4-\sqrt{2}}{3} < X < \frac{4+\sqrt{2}}{3}$$

o valor da probabilidade procurada vale:

$$\int_{\frac{4-\sqrt{2}}{3}}^{\frac{4+\sqrt{2}}{3}} \frac{1}{2} x dx \approx 0{,}629.$$

[1] Antigo Professor e Membro do Conselho Científico da Escola Superior de Polícia.

Esta é, pois, uma estimativa da probabilidade de que X assuma valores no intervalo:

$$\left]\frac{4-\sqrt{2}}{3}, \frac{4+\sqrt{2}}{3}\right[$$

centrado no valor médio de X :

$$E[X] = \mu_X' = \frac{4}{3}$$

e de semi-amplitude igual ao desvio-padrão de X :

$$\sigma_X = \frac{\sqrt{2}}{3}.$$

Neste caso foi possível obter o valor da probabilidade procurada, conseguido com a precisão que se entendeu, dado ser conhecida a distribuição da variável aleatória X em causa.

Pode, porém, acontecer que se conheçam o valor médio e o desvio-padrão da variável aleatória, mas se desconheça a correspondente distribuição, o que impossibilita o cálculo tal como anteriormente apresentado. É para uma situação deste tipo que a Desigualdade de Chebychev se mostra de enorme utilidade, já que, segundo Pestana (2004) a mesma envolve apenas o valor médio e a variância (de X), mostrando que o simples conhecimento de localização e escala permite fazer avaliações de probabilidades.

Este importante instrumento da Teoria da Probabilidade é válido para uma qualquer variável aleatória, com a única condição de ser finito o valor da respectiva variância, o que acarreta que os dois primeiros momentos ordinários o sejam também.

Este resultado é válido, por igual, para o caso de distribuições discretas, mas acarreta, em qualquer caso e como seria sempre de esperar, uma imprecisão na estimativa achada para a probabilidade do acontecimento em causa.

A Desigualdeade de Chebychev é um caso particular da Desigualdade de Markov, que se apresenta de seguida, sem demonstração, e que pode encontrar-se nos manuais dos autores portugueses mais consagrados.

Seja, então, $g(X)$ uma função mensurável da variável aleatória X , e que não assuma valores negativos, ou seja, $g(X) \geq 0$. Então, se existir o valor médio de $g(X)$, $E[g(X)]$, ter-se-á que:

$$\forall c \in \mathbf{R}^+ \quad , \quad P[g(X) \geq c] \leq \frac{E[g(X)]}{c}.$$

Como corolário desta propriedade, considere-se agora o caso em que a função considerada é:

$$g(X) = X.$$

Tem-se, neste caso, a desigualdade:

$$P[X \geq c] \leq \frac{E[X]}{c}.$$

Retomando o exemplo da distribuição inicial, facilmente se pode mostrar que:

$$P\left[X \geq \frac{5}{3}\right] = \int_{\frac{5}{3}}^{2} \frac{1}{2} x dx = \frac{1}{4}\left[x^2\right]_{\frac{5}{3}}^{2} = 0,30(5).$$

Em contrapartida, se se desconhecesse a distribuição da variável aleatória X , e se recorresse ao anterior resultado, corolário da Desgualdade de Markov, obter-se-ia:

$$P\left[X \geq \frac{5}{3}\right] \leq \frac{\frac{4}{3}}{\frac{5}{3}} = \frac{4}{5} = 0{,}8$$

o que mostra que o desconhecimento da distribuição de X determina a estimação de uma probabilidade do acontecimento em causa muito acima do seu valor real. A probabilidade estimada pelo recurso ao corolário da Desigualdade de Markov fornece um limite superior da probabilidade do acontecimento:

$$X \geq \frac{5}{3}$$

mas muito acima do valor real, calculável a partir do conhecimento da distribuição exacta.

Um segundo corolário da Desigualdade de Markov, mas que exige o conhecimento de mais informação, pode encontrar-se se se conhecer o momento absoluto ordinário de ondem $n \in \mathbf{N}_2$ de X, fornecendo o resultado:

$$P\left[|X| \geq c\right] \leq \frac{E\left[|X|^n\right]}{c^n}$$

onde $c \in \mathbf{R}^+$.

Tendo presente que para a variável aleatória X se tem:

$$E\left[|X|^5\right] \approx 9{,}14$$

este último corolário da Desigualdade de Markov permite a nova estimativa:

$$P\left[|X| \geq \frac{5}{3}\right] = P\left[X \geq \frac{5}{3}\right] \leq \frac{9{,}14}{\left(\frac{5}{3}\right)^5} \approx 0{,}711$$

que fornece um limite superior para a probabilidade do acontecimento em causa:

$$X \geq \frac{5}{3}$$

já mais próximo do verdadeiro valor da sua probabilidade, se fosse conhecida a distribuição exata de X.

Esta maior proximidade da probabilidade estimada através deste segundo corolário já requereu, contudo, o conhecimento do quinto momento absoluto ordinário de X, ou seja, uma informação maior que a requerida no caso do primeiro corolário.

A Desigualdade de Chebychev é um caso particular da Desigualdade de Markov, aplicada ao caso da função:

$$g(X) = \left(X - \mu_X'\right)^2$$

e tomando a constante c como sendo:

$$c = t^2 \sigma_X^2$$

onde μ_X' e σ_X^2 são, respetivamente, o valor médio e o desvio-padrão de X, e onde $t \in \mathbf{R}^+$. Virá, então, por substituição na Desigualdade de Markov:

$$P\left[\left(X - \mu_X'\right)^2 \geq t^2\sigma_X^2\right] \leq \frac{E\left[\left(X - \mu_X'\right)^2\right]}{t^2\sigma_X^2} = \frac{\sigma_X^2}{t^2\sigma_X^2} = \frac{1}{t^2}$$

ou seja:

$$P\left[\left|X-\mu_X^{'}\right|<t\sigma_X\right]\geq 1-\frac{1}{t^2}.$$

É esta expressão, ou a imediatamente anterior, que constitui a importante Desigualdade de Chebychev para o caso de uma única variável aleatória.

Mas esta desigualdade pode ainda assumir uma outra forma, se nela se fizer:

$$t\sigma_X=\alpha \quad\Leftrightarrow\quad t=\frac{\alpha}{\sigma_X} \quad\Rightarrow\quad t^2=\frac{\alpha^2}{\sigma_X^2} \quad\Leftrightarrow\quad \frac{1}{t^2}=\frac{V[X]}{\alpha^2}$$

forma essa que é:

$$P\left[\left|X-\mu_X^{'}\right|\geq\alpha\right]\leq\frac{V[X]}{\alpha^2}.$$

Tomando, mais uma vez, a variável aleatória inicialmente considerada, calcule-se a probabilidade:

$$P\left[\left|X-\frac{4}{3}\right|<1{,}1\cdot\frac{\sqrt{2}}{3}\right]=P\left[\frac{4-1{,}1\sqrt{2}}{3}<X<\frac{4+1{,}1\sqrt{2}}{3}\right].$$

Se se conhecer a distribuição de X, esta probabilidade vale:

$$\int_{\frac{4-1,1\sqrt{2}}{3}}^{\frac{4+1,1\sqrt{2}}{3}}\frac{1}{2}xdx\approx 0{,}691.$$

Contudo, se essa distribuição for desconhecida, e se recorrer à Desigualdade de Chebychev, virá, dado ser:

$$t=1{,}1$$

o valor da probabilidade em causa:

$$P\left[\left|X-\frac{4}{3}\right|<1{,}1\frac{\sqrt{2}}{3}\right]\geq 1-\frac{1}{1{,}1^2}\approx 0{,}174$$

que é um limite mínimo para a probabilidade procurada, embora muito distante do verdadeiro valor.

O que já pôde perceber-se é que a Desigualdade de Chebychev se mostra muito limitada ao nível das probabilidades estimadas. É o preço que a sua grande generalidade comporta.

O único caminho para melhorar o valor das suas contribuições é restringir o conjunto das distribuições a que se aplica, havendo necessidade de se conhecer, ao menos, que o seu comportamento tem maior proximidade com o de tipo gaussiano.

No caso da variável aleatória com valor médio nulo, $\mu_X^{'}=0$, e variância $\sigma_X^2=\sigma^2$, Murteira (1990) mostra que, se for conhecido o momento absoluto ordinário de quarta ordem:

$$\alpha_4^{'}=E\left[|X|^4\right]$$

se tem:

$$P\left[|X|\geq t\sigma\right]\leq\frac{\alpha_4^{'}-\sigma^4}{\alpha_4^{'}+t^4\sigma^4-2t^2\sigma^4}$$

com $t > 1$.

Admita-se agora que se possuem n variáveis aleatórias, semelhantes e independentes, cada uma com valor médio $\mu_1^{'}$ e variância σ^2, sendo $n \in \mathbf{N}_1$.

A média aritmética das n variáveis aleatórias é a variável aleatória:

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

cujo valor médio e variância são, respectivamente:

$$E\left[\overline{X}\right] = \mu_1^{'}$$

$$V\left[\overline{X}\right] = \frac{\sigma^2}{n}.$$

Recorrendo à Desigualdade de Markov, mas tomando agora a nova função g: **R**→**R**, definida por:

$$g\left(\overline{X}\right) = \left(\overline{X} - \mu_1^{'}\right)^2$$

para a qual:

$$E\left[\left(\overline{X} - \mu_1^{'}\right)^2\right] = \frac{\sigma^2}{n}$$

virá:

$$P\left[\left(\overline{X} - \mu_1^{'}\right)^2 \geq t^2\sigma^2\right] \leq \frac{\sigma^2}{nt^2\sigma^2} \quad \Leftrightarrow \quad P\left[\left|\overline{X} - \mu_1^{'}\right| > t\sigma\right] \leq \frac{1}{nt^2}.$$

Esta última expressão é, pois, a da Desigualdade de Chebychev, quando a variável aleatória é a média aritmética de n variáveis aleatórias, semelhantes e independentes, situação que se coloca frequentemente na prática.

Seja uma população normal, de valor médio, $\mu_1^{'} = 6$, e variância, $\sigma^2 = 0{,}36$, e suponha-se uma amostra de dimensão 100, oriunda dessa população. Ter-se-á, então:

$$E\left[\overline{X}\right] = 6$$

$$V\left[\overline{X}\right] = \frac{0{,}36}{100} = 0{,}0036$$

pelo que será:

$$\sigma_{\overline{X}} = 0{,}06.$$

Se neste caso se pretender estimar um valor para a probabilidade do acontecimento:

$$\left|\overline{X} - 6\right| < 1$$

virá:

$$P\left(\left|\overline{X}-6\right|<1\right)=P\left(\left|\overline{X}-6\right|<\frac{1}{0{,}06}\cdot 0{,}06\right)\geq 1-\frac{1}{100\left(\frac{1}{0{,}06}\right)^2}\approx 0{,}999964.$$

Esta é uma estimativa do mínimo da probabilidade procurada. De facto, se se soubesse que:

$$\overline{X}\sim N(6;0{,}0036)\Leftrightarrow Z=\frac{\overline{X}-6}{0{,}06}\sim N(0{,}1)$$

tirar-se-ia da tabela da distribuição Normal reduzida que:

$$P\left[\left|\overline{X}-6\right|<1\right]\approx 1.$$

A maior proximidade entre a anterior estimativa, 0,999964, e o valor real da probabilidade, quando se conhece a distribuição, deve-se ao facto de se ter usado uma amostra já grande, através da distribuição da sua média aritmética.

Se na anterior expressão da Desigualdade de Chebychev para a média aritmética de n variáveis aleatórias se fizer:

$$t\sigma_{\overline{X}}=\alpha \quad \Rightarrow \quad t^2\sigma_{\overline{X}}^2=\alpha^2$$

a expressão da desigualdade assumirá a forma:

$$P\left[\left|\overline{X}-\mu_1^{'}\right|\geq\alpha\right]\leq\frac{\sigma_{\overline{X}}^2}{n\alpha^2}$$

onde $n\in \mathbf{N}_1$ é o número de variáveis aleatórias.

No caso de se estar perante uma sucessão de n provas de Bernoulli, sendo k o número de êxitos nessas n provas, a Desigualdade de Chebychev toma a forma, facilmente dedutível:

$$P\left[\left|\frac{k}{n}-p\right|\geq\alpha\right]\leq\frac{p(1-p)}{n\alpha^2}$$

onde p é a probabilidade de ocorrer um êxito num qualquer ensaio e $\alpha\in\mathbf{R}^+$.

Mas a Desigualdade de Chebychev pode ser ainda generalizada a situações mais amplas, como se mostra com as duas propriedades que se seguem.

Sejam X_i , $(i=1,\ldots,n)$, n variáveis aleatórias independentes, para as quais se tem:

$$E\left[X_i\right]=\mu_i^{'}$$

$$V\left[X_i\right]=\sigma_i^2$$

e seja:

$$L=\sup\left\{\sigma_1^2,\ldots,\sigma_n^2\right\}.$$

Então, sendo $\alpha\in\mathbf{R}^+$, tem-se que:

$$P\left[\left|\overline{X}-\frac{\sum_{i=1}^{n}\mu_i^{'}}{n}\right|\geq \alpha\right]\leq \frac{L}{n\alpha^2}\ .$$

Embora independentes, as n variáveis aleatórias não possuem necessariamente o mesmo valor médio e a mesma variância.

No caso da sucessão de n provas de Bernoulli, admita-se que a probabilidade de êxito na i-ésima prova é p_i . Então, sendo k o número de êxitos nas n provas, tem-se:

$$E\left[X_i\right]=p_i$$

$$V\left[X_i\right]=p_i\left(1-p_i\right)\leq \frac{1}{4}$$

pelo que virá a Desigualdade de Chebychev correspondente à presente situação:

$$P\left[\left|\frac{k}{n}-\frac{\sum_{i=1}^{n}p_i}{n}\right|\geq \alpha\right]\leq \frac{1}{4n\alpha^2}\ .$$

Ora, a Desigualdade de Chebychev a que se chegou inicialmente refere-se a um intervalo centrado no valor médio da variável aleatória em causa.

Podem considerar-se, contudo, intervalos centrados num valor real qualquer, τ , não necessariamente coincidente com o valor médio.

Retomando a Desigualdade de Markov e fazendo:

$$g(X)=(X-\tau)^2$$

ter-se-á:

$$P\left[(X-\tau)^2\geq t^2\sigma^2\right]\leq \frac{E\left[(X-\tau)^2\right]}{t^2\sigma^2}$$

ou seja:

$$P\left[|X-\tau|\geq t\sigma\right]\leq \frac{E\left[\left(\left(X-\mu_1^{'}\right)+\left(\mu_1^{'}-\tau\right)\right)^2\right]}{t^2\sigma^2}$$

ou ainda:

$$P\left[|X-\tau|\geq t\sigma\right]\leq \frac{E\left[\left(X-\mu_1^{'}\right)^2\right]+2\left(\mu_1^{'}-\tau\right)E\left[X-\mu_1^{'}\right]+E\left[\left(\mu_1^{'}-\tau\right)^2\right]}{t^2\sigma^2}$$

ou, finalmente:

$$\underbrace{P\left[|X-\tau|\geq t\sigma\right]\leq \frac{1}{t^2}+\frac{\left(\mu_1^{'}-\tau\right)^2}{t^2\sigma^2}}_{(2)}$$

dado que o primeiro momento central de X é nulo:

$$E\left[X-\mu\right]=0$$

e que:

$$E\left[\left(\mu_1^{'}-\tau\right)^2\right]=\left(\mu_1^{'}-\tau\right)^2$$

$$E\left[\left(X-\mu_1^{'}\right)^2\right]=\sigma^2$$

A expressão (2) pode tomar a forma:

$$P\left[\left|X-\tau\right|<t\sigma\right]\geq \underbrace{1-\frac{1}{t^2}-\frac{\left(\mu_1^{'}-\tau\right)^2}{t^2\sigma^2}}_{(3)}$$

onde (3) fornece uma estimativa do limite inferior da probabilidade de X assumir valores no intervalo:

$$\left]\tau-t\sigma,\tau+t\sigma\right[$$

centrado em τ e não em $\mu_1^{'}$.

De igual modo, se se tiver a função:

$$g\left(\overline{X}\right)=\left(\overline{X}-\tau\right)^2$$

a Desigualdade de Chebychev virá neste outro formato:

$$P\left[\left|\overline{X}-\tau\right|<t\sigma\right]\geq 1-\frac{1}{nt^2}-\frac{\left(\mu_1^{'}-\tau\right)^2}{t^2\sigma^2}$$

que é também de muito fácil obtenção.

A Desigualdade de Chebychev, que se tem vindo a tratar de um modo abrangente, pode apresentar-se de um outro modo mais geral. Considerem-se, de novo, $n \in \mathbf{N}_1$, variáveis aleatórias independentes, X_i, (i = 1,..., n), todas elas de média nula, $\mu_i^{'}=0$, e variância, σ_i^2, (i = 1,..., n).

Seja, agora, a variável aleatória:

$$X=X_1+\cdots+X_n=\sum_{i=1}^{n}X_i$$

para a qual se tem:

$$E\left[X^2\right]=E\left[\left(X_1+\cdots+X_n\right)^2\right]=\sum_{i=1}^{n}\sigma_i^2=\sum\nolimits_n^2 \quad .$$

Sejam, então, os acontecimentos:

$$\begin{array}{l}
D_1=\left|X_1\right|<t\sum\nolimits_n \\
D_2=\left|X_1+X_2\right|<t\sum\nolimits_n \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
D_n=\left|X_1+\cdots+X_n\right|<t\sum\nolimits_n
\end{array}$$

A Desigualdade de Chebychev garante, então, que:

$$P\left(D_1 \cap D_2 \cap \ldots \cap D_n\right) \geq 1 - \frac{1}{t^2} \Leftrightarrow P\left[\bigcap_{i=1}^{n} D_i\right] \geq 1 - \frac{1}{t^2}.$$

Trata-se de uma propriedade de essencial interesse para a obtenção de uma condição suficiente para a conhecida lei forte dos grandes números.

Finalmente, a Desigualdade de Chebychev está também presente no âmbito dos processos estocásticos, conceito este que constitui, pode dizer-se assim, uma generalização do de variável aleatória. De facto, o processo estocástico mais não é que um conjunto de variáveis aleatórias, todas igualmente distribuídas, mas cada uma delas dependente de um parâmetro definido em certo domínio.

Para certo valor desse parâmetro obtém-se uma variável aleatória, com a referida distribuição. Em contrapartida, para certo valor da variável aleatória, obtém-se uma função do parâmetro antes referido, definido no domínio considerado.

Ao domínio onde se encontra definido o parâmetro considerado dá-se o nome de conjunto-índice do processo estocástico correspondente.

De um modo geral, os casos mais importantes são aqueles em que o parâmetro do processo estocástico é a variável tempo. Se o conjunto-índice é o conjunto dos números naturais, **N**, ou o dos inteiros, **Z**, ou uma sua parte própria, o processo estocástico diz-se de parâmetro discreto. Se o conjunto-índice é o corpo real, ou uma sua parte própria, o processo estocástico designa-se de parâmetro contínuo.

Também no caso de um processo estocástico:

$$\{X(t): t \in T\}$$

onde t é o parâmetro do processo, com valores no domínio T, se pode considerar uma função de valor médio do processo estocástico.

Em torno desta função de valor médio dispõem-se, para um e outro lado, as diversas realizações do processo estocástico, cada uma definida para um certo valor de $t \in T$.

É, então, possível mostrar que, se o processo estocástico:

$$\{X(t): t \in [a,b]\}$$

for diferenciável em média quadrática, e fazendo:

$$g_1(t) = \left\{E\left[\left|X(t)\right|^2\right]\right\}^{\frac{1}{2}}$$

$$g_2(t) = \left\{E\left[\left|X'(t)\right|^2\right]\right\}^{\frac{1}{2}}$$

se tem:

$$E\left[\sup_{t \in [a,b]} X^2(t)\right] \leq \frac{1}{2}\left[g_1^2(a) + g_1^2(b)\right] + \int_a^b g_1(t) \cdot g_2(t) dt.$$

E desta propriedade se pode obter, como corolário, a Desigualdade de Markov para o caso dum processo estocástico nas condições indicadas:

$$\forall c \in \mathbf{R}^+ \quad , \quad P\left[\sup_{t \in [a,b]} \left|X(t)\right| > c\right] \leq \frac{E\left[\sup_{t \in [a,b]} X^2(t)\right]}{c^2}$$

Se for $m(t)$ a função de valor médio do processo estocástico $X(t)$, pode obter-se a Desigualdade de Chebychev para o caso de um processo estocástico nas condições referidas, ou seja:

$$P\left[\left|X(t)-m(t)\right| \leq c\right] \geq 1-\left[\frac{\sigma_{X(a)}^{2}+\sigma_{X(b)}^{2}}{2c^{2}}+\frac{\int_{a}^{b}\sigma_{X(t)}\cdot\sigma_{X'(t)}dt}{c^{2}}\right]$$

onde $t \in [a,b]$ e $c \in \mathbf{R}^{+}$. Trata-se, pois, de um limite inferior para a probabilidade de o processo estocástico se situar no interior de certa região centrada na sua função de valor médio.

Se se considerarem duas realizações do processo estocástico em causa, sejam X e Y, ambas com valor médio nulo e variância unitária, e se for ρ o coeficiente de correlação entre as duas realizações - variáveis aleatórias, portanto -, pode mostrar-se que se tem:

$$E\left[max\{X,Y\}\right] \leq 1+\sqrt{1-\rho^{2}}$$

e também que:

$$P\left[\left|X-E[X]\right| \geq t\sigma_{X} \vee \left|Y-E[Y]\right| \geq t\sigma_{Y}\right] \leq \frac{1+\sqrt{1-\rho^{2}}}{t^{2}}.$$

E é claro que se for Y constante, será $\rho = 0$, obtendo-se, então, a expressão já antes achada para a Desigualdade de Chebychev no caso de uma só variável aleatória:

$$P\left[\left|X-E[X]\right| \geq t\sigma_{X}\right] \leq \frac{1}{t^{2}}.$$

Fica assim tratada a Desigualdade Chebychev, mas numa variedade muito mais vasta de situações que as normalmente contempladas nos textos de uso corrente ao nível dos cursos de licenciatura onde o tema está usualmente presente.

BIBLIOGRAFIA

GNEDENKO, B. V. (1976): *The Theory of Probability*, MIR, Moscovo.

MELLO, F. Galvão de (1993): *Probabilidades e Estatística, Conceitos e Métodos Fundamentais - Volume I*, Escolar Editora, Lisboa.

MURTEIRA, Bento José Ferreira (1990): *Probabilidades e Estatística - Volume I*, 2ª Edição Revista, Editora McGraw-Hill de Portugal, Lda..

OLIVEIRA, J. Tiago de (1990): *Probabilidades e Estatística: Conceitos, Métodos e Aplicações, Volume I*, Editora McGraw-Hill de Portugal, Lda..

PARZEN, Emanuel (1972): *Processos Estocasticos*, Paraninfo, Madrid.

PESTANA, Dinis Duarte e Sílvio Filipe Velosa, (2006): *Introdução à Probabilidade e à Estatística,* Volume I, 2ª Edição Revista e Actualizada, Fundação Calouste Gulbenkian, Serviço de Educação e Bolsas.

VENTZEL, H. (1973): *Théorie des Probabilités*, MIR, Moscovo