Introdução à Econofísica

Aula 6

O teorema limite central

Suponhamos N variáveis reais aleatórias

$$x_i, \text{ com } i = 1, 2, \ldots, N,$$

identicamente distribuídas na reta real, com densidade de probabilidade dada por $p(x_i)$, para cada $i = 1, 2, \ldots, N$. Também suponhamos que a densidade p seja arbitrária, mas tal que todos os valores médios das potências de x_i sejam finitos. Nesta aula calculamos a densidade de probabilidade para a variável

$$X = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{1}$$

no limite em que N é muito grande.

Como exemplo análogo do caso discreto, pensemos em uma série de N lançamentos de um dado. Cada lançamento pode resultar em uma de seis possíveis faces, com a probabilidade de ocorrência igual a um sexto. Assim, a ocorrência de cada um dos números de face de 1 a 6 é igualmente provável em cada lançamento. Depois de N lançamentos idênticos e independentes, podemos perguntar qual a média do número de face. Por exemplo, em três lançamentos, obtemos os números de face: $x_1 = 1$, $x_2 = 5$ e $x_3 = 4$ e a média é dada por

$$\begin{aligned} X &= \frac{1}{3}(1+5+4) \\ &= \frac{10}{3}. \end{aligned}$$

Podemos repetir esses três lançamentos muitas e muitas vezes e estimar a distribuição de probabilidades de que um valor X ocorra para a média dos valores de face. Essa distribuição, para três lançamentos, não é gaussiana. Mas se fizermos esse mesmo experimento com o número de lançamentos N cada vez maior, a distribuição de ocorrência de X será gaussiana, segundo o teorema do limite central. A seguir, segue uma prova do teorema.

Seja $P(X)$ a densidade de probabilidade de que a variável X, definida pela Eq. (1), resulte com valor entre X e $X + dX$. Seja $Q(K)$ a transformada de Fourier de $P(X)$:

$$Q(K) = \int_{-\infty}^{+\infty} dX\, P(X) \exp(iKX). \tag{2}$$

A exponencial pode ser escrita como uma série infinita:

$$\exp\left(iKX\right) = \sum_{n=0}^{\infty}\frac{i^{n}K^{n}}{n!}X^{n}.$$

Logo,

$$\begin{aligned} Q\left(K\right) &= \sum_{n=0}^{\infty}\frac{i^{n}K^{n}}{n!}\int_{-\infty}^{+\infty}dX\,P\left(X\right)X^{n} \\ &= \sum_{n=0}^{\infty}\frac{i^{n}K^{n}}{n!}\left\langle X^{n}\right\rangle, \qquad (3) \end{aligned}$$

onde denotamos o valor esperado de X^n por $\langle X^n\rangle$:

$$\left\langle X^{n}\right\rangle = \int_{-\infty}^{+\infty}dX\,P\left(X\right)X^{n}. \qquad (4)$$

Para $N=2$, por exemplo, temos a probabilidade $p\left(x_1\right)dx_1$ de que a variável estocástica x_1 assuma um valor entre x_1 e x_1+dx_1. Analogamente, temos a probabilidade $p\left(x_2\right)dx_2$ de que a variável estocástica x_2 assuma um valor entre x_2 e x_2+dx_2. De todos os valores independentes que x_1 e x_2 possam assumir, qual a probabilidade de que assumam valores tais que $\left(x_1+x_2\right)/2$ tenha um valor entre X e $X+dX$? A resposta para essa questão é simples: escrevemos

$$x_2 = 2X - x_1$$

e, para cada valor de X e x_1 a probabilidade é

$$p\left(x_1\right)p\left(x_2\right)dx_1dx_2.$$

Como agora queremos que as variáveis independentes sejam X e x_1, temos utilizamos a relação:

$$\begin{aligned} dx_1dx_2 &= \begin{vmatrix} \frac{\partial x_1}{\partial x_1} & \frac{\partial x_1}{\partial X} \\ \frac{\partial x_2}{\partial x_1} & \frac{\partial x_2}{\partial X} \end{vmatrix} dx_1dX \\ &= \begin{vmatrix} 1 & 0 \\ -1 & 2 \end{vmatrix} dx_1dX \\ &= 2dx_1dX \end{aligned}$$

e a probabilidade procurada pode ser escrita como:

$$P\left(X\right)dX = 2\int_{-\infty}^{+\infty}p\left(x_1\right)p\left(2X-x_1\right)dx_1dX,$$

já que qualquer valor de x_1 pode ocorrer e ainda assim termos o mesmo valor para X. Portanto, a densidade de probabilidade é dada por:

$$P\left(X\right) = 2\int_{-\infty}^{+\infty}p\left(x_1\right)p\left(2X-x_1\right)dx_1,$$

que também pode ser expressa como:

$$P(X) \;=\; 2\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} p(x_1)\,p(x_2)\,\delta\left[x_2-(2X-x_1)\right]dx_1dx_2,$$

onde δ é a chamada "função delta de Dirac". Se utilizamos as propriedades:

$$\delta(ay) \;=\; \frac{1}{|a|}\delta(y)$$

e

$$\delta(y) \;=\; \delta(-y),$$

obtemos:

$$P(X) \;=\; \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} p(x_1)\,p(x_2)\,\delta\left(X-\frac{x_1+x_2}{2}\right)dx_1dx_2.$$

Para $N>2$, podemos facilmente generalizar a fórmula acima:

$$P(X) \;=\; \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\ldots\int_{-\infty}^{+\infty} p(x_1)\,p(x_2)\ldots p(x_N)\,\delta\left(X-\frac{1}{N}\sum_{n=1}^{N}x_i\right)dx_1dx_2\ldots dx_N.$$

Usando esse resultado na Eq. (4), temos:

$$\begin{aligned}
\langle X^n\rangle &= \int_{-\infty}^{+\infty} dX\,P(X)\,X^n \\
&= \int_{-\infty}^{+\infty} dX\,X^n\int_{-\infty}^{+\infty}\ldots\int_{-\infty}^{+\infty} p(x_1)\ldots p(x_N)\,\delta\left(X-\frac{1}{N}\sum_{n=1}^{N}x_i\right)dx_1\ldots dx_N \\
&= \int_{-\infty}^{+\infty}\ldots\int_{-\infty}^{+\infty} dx_1\ldots dx_N p(x_1)\ldots p(x_N)\int_{-\infty}^{+\infty} dX\,X^n\delta\left(X-\frac{1}{N}\sum_{n=1}^{N}x_i\right) \\
&= \int_{-\infty}^{+\infty}\ldots\int_{-\infty}^{+\infty} dx_1\ldots dx_N p(x_1)\ldots p(x_N)\left(\frac{1}{N}\sum_{n=1}^{N}x_i\right)^n.
\end{aligned}$$

Utilizando essa igualdade na Eq. (3), vem:

$$\begin{aligned}
Q(K) &= \sum_{n=0}^{\infty}\frac{i^nK^n}{n!}\int_{-\infty}^{+\infty}\ldots\int_{-\infty}^{+\infty} dx_1\ldots dx_N p(x_1)\ldots p(x_N)\left(\frac{1}{N}\sum_{n=1}^{N}x_i\right)^n \\
&= \int_{-\infty}^{+\infty}\ldots\int_{-\infty}^{+\infty} dx_1\ldots dx_N p(x_1)\ldots p(x_N)\sum_{n=0}^{\infty}\frac{i^nK^n}{n!}\left(\frac{1}{N}\sum_{n=1}^{N}x_i\right)^n.
\end{aligned}$$

Reconhecemos a série infinita acima como uma exponencial:

$$\sum_{n=0}^{\infty}\frac{i^nK^n}{n!}\left(\frac{1}{N}\sum_{n=1}^{N}x_i\right)^n \;=\; \exp\left(\frac{iK}{N}\sum_{n=1}^{N}x_i\right).$$

Logo,

$$\begin{aligned}
Q(K) &= \int_{-\infty}^{+\infty} \ldots \int_{-\infty}^{+\infty} dx_1 \ldots dx_N p(x_1) \ldots p(x_N) \exp\left(\frac{iK}{N} \sum_{n=1}^{N} x_i\right) \\
&= \int_{-\infty}^{+\infty} dx_1 p(x_1) \exp\left(\frac{iK}{N} x_1\right) \ldots \int_{-\infty}^{+\infty} dx_N p(x_N) \exp\left(\frac{iK}{N} x_N\right) \\
&= \left[\int_{-\infty}^{+\infty} dx p(x) \exp\left(\frac{iK}{N} x\right)\right]^N . \qquad (5)
\end{aligned}$$

Como estamos supondo que N é muito grande, podemos expandir:

$$\exp\left(\frac{iK}{N} x\right) = 1 + i\frac{K}{N} x - \frac{1}{2}\left(\frac{K}{N}\right)^2 x^2 + \cdots$$

e, portanto,

$$\begin{aligned}
\int_{-\infty}^{+\infty} dx p(x) \exp\left(\frac{iK}{N} x\right) &= \int_{-\infty}^{+\infty} dx p(x) \left[1 + i\frac{K}{N} x - \frac{1}{2}\left(\frac{K}{N}\right)^2 x^2 + \cdots\right] \\
&= \int_{-\infty}^{+\infty} dx p(x) + i\frac{K}{N} \int_{-\infty}^{+\infty} dx p(x) x \\
&- \frac{1}{2}\left(\frac{K}{N}\right)^2 \int_{-\infty}^{+\infty} dx p(x) x^2 + \cdots \\
&= 1 + i\frac{K}{N} \langle x \rangle - \frac{1}{2}\left(\frac{K}{N}\right)^2 \left\langle x^2 \right\rangle + \cdots,
\end{aligned}$$

onde

$$\int_{-\infty}^{+\infty} dx p(x) = 1,$$

$$\langle x \rangle = \int_{-\infty}^{+\infty} dx p(x) x$$

e

$$\left\langle x^2 \right\rangle = \int_{-\infty}^{+\infty} dx p(x) x^2.$$

Com isso, a Eq. (5) pode ser escrita como:

$$\begin{aligned}
Q(K) &= \left[1 + i\frac{K}{N} \langle x \rangle - \frac{1}{2}\left(\frac{K}{N}\right)^2 \left\langle x^2 \right\rangle + \cdots\right]^N \\
&= \exp\left\{N \ln\left[1 + i\frac{K}{N} \langle x \rangle - \frac{1}{2}\left(\frac{K}{N}\right)^2 \left\langle x^2 \right\rangle + \cdots\right]\right\}.
\end{aligned}$$

Para N muito grande, podemos expandir o logaritmo:

$$\begin{aligned}
\ln\left[1+i\frac{K}{N}\langle x\rangle-\left(\frac{K}{N}\right)^2\left\langle x^2\right\rangle+\cdots\right] &= \left[i\frac{K}{N}\langle x\rangle-\frac{1}{2}\left(\frac{K}{N}\right)^2\left\langle x^2\right\rangle+\cdots\right] \\
&- \frac{1}{2}\left[i\frac{K}{N}\langle x\rangle-\frac{1}{2}\left(\frac{K}{N}\right)^2\left\langle x^2\right\rangle+\cdots\right]^2+\cdots \\
&= i\frac{K}{N}\langle x\rangle-\frac{1}{2}\left(\frac{K}{N}\right)^2\left\langle x^2\right\rangle-\frac{1}{2}\left(i\frac{K}{N}\langle x\rangle\right)^2+\cdots \\
&= i\frac{K}{N}\langle x\rangle-\frac{1}{2}\left(\frac{K}{N}\right)^2\left(\left\langle x^2\right\rangle-\langle x\rangle^2\right)+\cdots \\
&= i\frac{K}{N}\langle x\rangle-\frac{1}{2}\left(\frac{K}{N}\right)^2\operatorname{var}(x)+\cdots.
\end{aligned}$$

Então,

$$\begin{aligned}
Q(K) &= \exp\left\{N\left[i\frac{K}{N}\langle x\rangle-\frac{1}{2}\left(\frac{K}{N}\right)^2\operatorname{var}(x)+\cdots\right]\right\} \\
&= \exp\left[iK\langle x\rangle-\frac{1}{2}\frac{K^2}{N}\operatorname{var}(x)+\cdots\right]. \qquad (6)
\end{aligned}$$

Essa é a trasformada de Fourier da densidade de probabilidade $P(X)$ que procuramos. Podemos inverter a Eq. (2):

$$\begin{aligned}
\int_{-\infty}^{+\infty} dK\, Q(K)\exp(-iKX) &= \int_{-\infty}^{+\infty} dK\left[\int_{-\infty}^{+\infty} dY\, P(Y)\exp(iKY)\right]\exp(-iKX) \\
&= \int_{-\infty}^{+\infty} dY\, P(Y)\int_{-\infty}^{+\infty} dK\exp[iK(Y-X)].
\end{aligned}$$

Como

$$\int_{-\infty}^{+\infty} dK\exp[iK(Y-X)] = 2\pi\delta(Y-X),$$

concluímos que podemos obter $P(X)$ a partir de $Q(K)$:

$$P(X) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} dK\, Q(K)\exp(-iKX).$$

Então, da Eq. (6), segue:

$$\begin{aligned}
P(X) &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} dK\ \exp\left[-iK(X-\langle x\rangle)-\frac{1}{2}\frac{K^2}{N}\operatorname{var}(x)+\cdots\right] \\
&\approx \frac{1}{2\pi}\int_{-\infty}^{+\infty} dK\ \exp\left[-iK(X-\langle x\rangle)-\frac{1}{2}\frac{K^2}{N}\operatorname{var}(x)\right]
\end{aligned}$$

$$= \frac{1}{2\pi}\int_{-\infty}^{+\infty} dK\ \exp\left\{-\frac{1}{2}\frac{\mathrm{var}\,(x)}{N}\left[K^2 + 2i\frac{NK}{\mathrm{var}\,(x)}\left(X - \langle x\rangle\right)\right]\right\}$$

e, completanto o quadrado no argumento da exponencial no integrando, vem:

$$\begin{aligned} P\,(X) &\approx \frac{1}{2\pi}\exp\left\{\frac{1}{2}\frac{\mathrm{var}\,(x)}{N}\left(\frac{iN\,(X-\langle x\rangle)}{\mathrm{var}\,(x)}\right)^2\right\} \\ &\times \int_{-\infty}^{+\infty} dK\ \exp\left\{-\frac{1}{2}\frac{\mathrm{var}\,(x)}{N}\left[K^2 + 2i\frac{NK}{\mathrm{var}\,(x)}\left(X-\langle x\rangle\right) + \left(\frac{iN\,(X-\langle x\rangle)}{\mathrm{var}\,(x)}\right)^2\right]\right\} \\ &= \frac{1}{2\pi}\exp\left\{-\frac{1}{2}\frac{N\,(X-\langle x\rangle)^2}{\mathrm{var}\,(x)}\right\} \\ &\times \int_{-\infty}^{+\infty} dK\ \exp\left\{-\frac{1}{2}\frac{\mathrm{var}\,(x)}{N}\left(K + \frac{iN\,(X-\langle x\rangle)}{\mathrm{var}\,(x)}\right)^2\right\}. \end{aligned}$$

Como

$$\int_{-\infty}^{+\infty} dK\ \exp\left\{-\frac{1}{2}\frac{\mathrm{var}\,(x)}{N}\left(K + \frac{iN\,(X-\langle x\rangle)}{\mathrm{var}\,(x)}\right)^2\right\} = \sqrt{\frac{2\pi N}{\mathrm{var}\,(x)}},$$

escrevemos:

$$P\,(X) \approx \sqrt{\frac{N}{2\pi\mathrm{var}\,(x)}}\exp\left\{-\frac{N}{2\mathrm{var}\,(x)}\left(X-\langle x\rangle\right)^2\right\}.$$

O valor esperado de X é dado por:

$$\begin{aligned} \langle X\rangle &= \left\langle \frac{1}{N}\sum_{n=1}^{N} x_n \right\rangle \\ &= \frac{1}{N}\sum_{n=1}^{N}\langle x_n\rangle \\ &= \langle x\rangle\,, \end{aligned}$$

já que as variáveis x_n são todas independentes e identicamente distribuídas. A variância de X é dada por:

$$\begin{aligned} \mathrm{var}\,(X) &= \left\langle X^2\right\rangle - \langle X\rangle^2 \\ &= \left\langle \frac{1}{N}\sum_{m=1}^{N} x_m \frac{1}{N}\sum_{n=1}^{N} x_n\right\rangle - \langle x\rangle^2 \\ &= \frac{1}{N^2}\sum_{m=1}^{N}\sum_{n=1}^{N}\langle x_m x_n\rangle - \langle x\rangle^2\,. \end{aligned}$$

Mas, como as variáveis x_m são todas independentes e identicamente distribuídas, temos:

$$\langle x_m x_n\rangle = \delta_{mn}\left\langle x^2\right\rangle + (1-\delta_{mn})\,\langle x\rangle^2\,,$$

onde δ_{mn} é a função de dois inteiros, m e n, chamada "delta de Kronecker". Assim,

$$\begin{aligned}
\mathrm{var}\left(X\right) &= \frac{1}{N^2}\sum_{m=1}^{N}\sum_{n=1}^{N}\left[\delta_{mn}\left\langle x^2\right\rangle + \left(1-\delta_{mn}\right)\left\langle x\right\rangle^2\right] - \left\langle x\right\rangle^2 \\
&= \frac{1}{N^2}\sum_{m=1}^{N}\left\langle x^2\right\rangle + \frac{1}{N^2}\sum_{m=1}^{N}\sum_{n=1}^{N}\left(1-\delta_{mn}\right)\left\langle x\right\rangle^2 - \left\langle x\right\rangle^2 \\
&= \frac{1}{N}\left\langle x^2\right\rangle + \frac{1}{N^2}\sum_{m=1}^{N}\sum_{n=1}^{N}\left\langle x\right\rangle^2 - \frac{1}{N^2}\sum_{m=1}^{N}\sum_{n=1}^{N}\delta_{mn}\left\langle x\right\rangle^2 - \left\langle x\right\rangle^2 \\
&= \frac{1}{N}\left\langle x^2\right\rangle + \left\langle x\right\rangle^2 - \frac{1}{N^2}\sum_{m=1}^{N}\left\langle x\right\rangle^2 - \left\langle x\right\rangle^2 \\
&= \frac{1}{N}\left\langle x^2\right\rangle - \frac{1}{N}\left\langle x\right\rangle^2 \\
&= \frac{\mathrm{var}\left(x\right)}{N}.
\end{aligned}$$

Podemos, então, escrever a densidade de probabilidade para a variável X, quando N é muito grande, como uma gaussiana:

$$P\left(X\right) \approx \frac{1}{\sqrt{2\pi \mathrm{var}\left(X\right)}} \exp\left\{-\frac{\left(X-\left\langle X\right\rangle\right)^2}{2\mathrm{var}\left(X\right)}\right\}.$$