MOMENTOS E FUNÇÕES GERADORAS

HÉLIO BERNARDO LOPES[1]

Resumo. O presente texto visa mostrar, de um modo tão unificado quanto possível, a temática dos momentos e das funções geradoras, estas últimas muito ausentes, atualmente, das disciplinas de Teoria da Probabilidade, e, estranhamente, quase sempre mal recebidas quando apresentadas.
O texto pode também servir de ponto de partida para desenvolvimentos mais profundos do tema, tratado já então ao nível de pós-graduações diversas, até porque, manifestamente, ele vai mais longe do que o nível usualmente presente naquelas disciplinas.

No estudo da Teoria da Probabilidade está sempre presente o tratamento dos momentos das diversas distribuições estudadas, desde que os mesmos existam.

O seu significado pleno, contudo, fica longe de ser percebido e integrado na compreensão adequada das correspondentes distribuições.

Em contrapartida, de um modo muito geral, as funções geradoras de probabilidades e de momentos, a função caraterística e os cumulantes, só muito ligeiramente são abordados, passando completamente em claro o seu significado, as suas relações, as respetivas potencialidades, e, mais ainda que tudo isto, o seu potencial explicativo sobre o tipo de distribuição resultante de operações algébricas, ou mesmo transcendentes, entre variáveis aleatórias diversas.

Por tudo isto, esperando contribuir para a ultrapassagem desta situação, generalizadamente reconhecida, se decidiu escrever o presente texto.

Quando se trabalha com uma distribuição probabilística, seja discreta ou contínua, é essencial conhecer o domínio da correspondente variável aleatória, ou vetor aleatório, e a respetiva função massa de probabilidade ou função densidade de probabilidade, ou, em alternativa, a função distribuição.

O conhecimento do domínio da variável aleatória em estudo determina, de imediato, o âmbito de tratamento do problema, em cujo seio se consideram os acontecimentos que podem ter interesse para o problema.

Em contrapartida, este conhecimento não permite compreender completamente a estrutura da distribuição probabilística em causa. De um modo geral, essa estrutura está muito distante da uniformidade, sendo, pela natureza das coisas, expectável que a mesma se concentre em torno de um ponto distante dos extremos do domínio da variável aleatória em estudo, atenuando-se à medida que se caminha para esses extremos.

Significa tal que se impõe encontrar instrumentos que forneçam indicações, razoavelmente seguras e universalmente interpretáveis, sobre o modo como a unidade probabilística se distribui ao longo do domínio da variável aleatória em estudo.

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

Ora, os designados momentos são alguns desses instrumentos, que são, em essência, de dois tipos: momentos em relação a uma certa constante $c \in \mathbf{R}$, e momentos absolutos em relação a essa mesma constante. Note-se, contudo, que existem distribuições que não têm momentos.

Seja, então, X uma variável aleatória qualquer e $g(X)$ uma função mensurável da mesma, que seja ainda uma variável aleatória.

Admita-se, por igual, que $g(X)$ seja discreta ou contínua se X o for, respetivamente. Dá-se o nome de momento de ordem $n \in \mathbf{N}_1$ de $g(X)$ em relação à constante $c \in \mathbf{R}$ ao valor:

$$E\left[\left(g(X)-c\right)^n\right]=\sum_{x\in Z}\left[g(x)-c\right]^n p_X(x)$$

onde $p_X(x)$ é a função massa de probabilidade de X, no caso de X e de $g(X)$ serem variáveis aleatórias discretas, desde que a série anterior seja absolutamente convergente, ou seja, que exista em $\mathbf{R}$:

$$E\left[\left|g(X)-c\right|^n\right]=\sum_{x\in Z}\left|g(x)-c\right|^n p_X(x).$$

A esta última expressão dá-se também a designação de momento abasoluto de ordem $n \in \mathbf{N}_1$ de $g(X)$ em relação à constante $c \in \mathbf{R}$.

No caso de X e de $g(X)$ serem variáveis aleatórias contínuas, define-se momento de ordem $n \in \mathbf{N}_1$ de $g(X)$ em relação à constante $c \in \mathbf{R}$ como sendo:

$$E\left[\left(g(X)-c\right)^n\right]=\int_{-\infty}^{+\infty}\left(g(X)-c\right)^n f_X(x)dx$$

desde que o anterior integral seja absolutamente convergente, isto é, que exista em $\mathbf{R}$ o valor de:

$$E\left[\left|g(X)-c\right|^n\right]=\int_{-\infty}^{+\infty}\left|g(X)-c\right|^n f_X(x)dx.$$

E, à semelhança do caso discreto, esta última expressão toma o nome de momento absoluto de ordem $n \in \mathbf{N}_1$ de $g(X)$ em relação à constante $c \in \mathbf{R}$.

Se se considerar a função da variável aleatória X :

$$g(X)=X$$

também ela aleatória, as expressões anteriores passarão a ser:

$$E\left[(X-c)^n\right]=\sum_{x\in Z}(x-c)^n p_X(x) \quad \Leftarrow \quad E\left[|X-c|^n\right]=\sum_{x\in Z}|x-c|^n p_X(x)\in \mathbf{R}$$

e:

$$E\left[(X-c)^n\right]=\int_{-\infty}^{+\infty}(x-c)^n f_X(x)dx \quad \Leftarrow \quad E\left[|X-c|^n\right]=\int_{-\infty}^{+\infty}|x-c|^n f_X(x)dx\in \mathbf{R}$$

respetivamente, para os casos discreto e contínuo, e onde o valor de cada uma das primeiras expressões só existe se for absolutamente convergente, ou seja, se o valor da correspondente segunda expressão existir em **R**.

Como é evidente, quando, no caso discreto, o domínio da variável aleatória é finito, as séries consideradas degeneram em somas que, naturalmente, existem sempre em **R**.

Ora, quando $c = 0$ os momentos são designados por ordinários, ou seja, calculados em relação à origem do corpo real, vindo nesta situação:

$$\mu_n^{'} = E\left[X^n\right] = \sum_{x \in Z} x^n p_X(x)$$

desde que exista em **R** o valor:

$$E\left[|X|^n\right] = \sum_{x \in Z} |x|^n p_X(x)$$

onde a primeira expressão representa o momento ordinário de ondem $n \in \mathbf{N}_1$ e a segunda o momento absoluto ordinário de ordem $n \in \mathbf{N}_1$, no caso em que X é uma variável aleatória discreta, e tendo-se:

$$\mu_n^{'} = E\left[X^n\right] = \int_{-\infty}^{+\infty} x^n f_X(x)dx$$

se existir em **R** o valor:

$$E\left[|X|^n\right] = \int_{-\infty}^{+\infty} |x|^n f_X(x)dx$$

no caso contínuo.

Note-se, todavia, que se podem definir momentos de ordem $-n$, com $n \in \mathbf{N}_0$, ou seja, para valores de n inteiros e negativos, tendo-se, por exemplo, as expressões:

$$\mu_{-1}^{'} = E\left[\frac{1}{X}\right]$$

e:

$$\lim_{n \to 0} \frac{\mu_n^{'} - 1}{n}$$

onde a primeira representa a média harmónica e a segunda a média geométrica. E podem, por igual, definir-se momentos fracionários.

Se o domínio de X, designado intervalo de variação de X, tiver grande amplitude, os valores de x^n serão grandes, em módulo, para a generalidade dos valores da variável aleatória. Ao contrário, pois, do que se passa no caso em que o intervalo de variação seja bastante menor.

Porém, tal indicador pouca ajuda pode fornecer, porque o primeiro momento ordinário de X pode ser o mesmo, ou muito próximo. Mas o mesmo já não ocorre com o primeiro momento absoluto ordinário, que variará diretamente com a amplitude do intervalo de variação de X. Um exemplo esclarecerá o que acaba de dizer-se.

Seja, então, a variável aleatória contínua, definida por:

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

Tem-se, no presente caso:

$$E[X] = 0 \quad \wedge \quad E[|X|] = 1.$$

Em contrapartida, para a variável aleatória X , igualmente contínua, dada por:

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

vem:

$$E[X] = 0 \quad \wedge \quad E[|X|] = 2.$$

Ou seja, mau grado a unidade probabilística se encontrar distribuída em intervalos de variação distintos, o segundo com uma amplitude dupla da do primeiro, ambas as variáveis aleatórias apresentam o mesmo valor para o primeiro momento ordinário. Ao contrário, o primeiro momento absoluto ordinário cresceu com o crescimento do intervalo de variação.

Acontece que as variáveis aleatórias acabadas de expor são simétricas em relação a $\mu_1^{'} = 0$, podendo mostrar-se que, numa tal circunstância, todos os seus momentos de ordem impar em relação a $\mu_1^{'}$ são nulos, embora se deva referir que a recíproca não é verdadeira.

Para as duas variáveis aleatórias consideradas, o seu primeiro momento ordinário é o mesmo, embora apresentem intervalos de variação muito distintos. O efeito desta última realidade, como se viu, fez-se sentir através do primeiro momento absoluto ordinário, onde o efeito da consideração do operador módulo foi o de tornar positivo o valor da variável que surge a multiplicar a respetiva função densidade de probabilidade, sempre não nula.

Estes factos permitem intuir que a simetria das anteriores variáveis aleatórias está ligada ao valor dos momentos de ordem impar em relação a $\mu_1^{'} = 0$, ao passo que a maior ou menos proximidade dos valores dessa variável se encontra ligada aos momentos de ordem par, que têm sobre os valores da variável aleatória, positivos ou negativos, o mesmo efeito que o criado pela aplicação do operador módulo.

Ao primeiro momento ordinário de uma variável aleatória, discreta ou contínua, dá-se o nome de valor médio dessa variável aleatória, sendo designado por $\mu_1^{'}$. Representa o centro da distribuição, em tono do qual a função massa de probabilidade, ou a função densidade de probabilidade, se distribui. E, como pôde já referir-se atrás, na grande generalidade das distribuições, qualquer daquelas funções apresenta o seu máximo nas proximidades do respetivo valor médio.

É essencial salientar que a existência de momentos de ordem elevada está ligada à baixa probabilidade de ocorrerem valores de X que, em módulo, sejam elevados, ou seja:

$$P(|X|>n)$$

é um infinitésimo de ordem superior a X^{-n}, com $n \in \mathbf{N}_1$.

Assim, a caraterização de uma distribuição probabilística começará pela consideração da respetiva natureza, ou seja, se se está perante uma variável aleatória discreta ou contínua.

A segunda caraterística dessa distribuição é o intervalo de variação da correspondente variável aleatória. Ele dará, ao menos, duas indicações: a região do eixo real onde a variável aleatória pode assumir valores, e a maior ou menor concentração da distribuição da unidade probabilística nesse intervalo.

A terceira caraterística de uma distribuição é, pois, o seu valor médio, que é, na enorme generalidade dos casos, e desde que exista, o principal indicador do centro da distribuição.

Nesta fase, torna-se essencial e já possível introduzir o novo conceito de momento central de ordem $n \in \mathbf{N}_1$ de uma variável aleatória, e que corresponde ao caso em que $c = \mu_1^{'}$:

$$\mu_n = E\left[\left(X-\mu_1^{'}\right)^n\right] = \sum_{x \in Z}\left(x-\mu_1^{'}\right)^n p_X(x) \quad \Leftarrow \quad E\left[\left|X-\mu_1^{'}\right|^n\right] = \sum_{x \in Z}\left|x-\mu_1^{'}\right|^n p_X(x) \in \mathbf{R}$$

e:

$$E\left[\left(X-\mu_1^{'}\right)^n\right] = \int_{-\infty}^{+\infty}\left(X-\mu_1^{'}\right)^n f_X(x)dx \quad \Leftarrow \quad E\left[\left|X-\mu_1^{'}\right|^n\right] = \int_{-\infty}^{+\infty}\left|X-\mu_1^{'}\right|^n f_X(x)dx \in \mathbf{R}$$

respetivamente, para os casos discreto e contínuo, e onde cada uma das primeiras expressões só existe se for absolutamente convergente, ou seja, se a correspondente segunda expressão existir em $\mathbf{R}$.

Ora, como pôde já referir-se, e como se terá intuído do exemplo anteriormente apresentado, os momentos de ordem par fornecem uma indicação da concentração da distribuição probabilística: se forem muito pequenos, essa concentração será grande, porque os valores da variável aleatória serão próximos; se forem grandes, será inversa a situação.

Isto mostra que deverá tomar-se como medida da concentração da distribuição em torno do seu valor médio um dos momentos centrais de ordem par. A questão a que tem de responder-se é esta: qual a ordem desse momento?

Se se tiver presente que o segundo momento, seja ordinário ou central, se exprime no quadrado das unidades usadas na medição dos valores de X, e que a sua raiz quadrada se exprime nessas mesmas unidades, de imediato se percebe que o momento central que deve ser usado para indicar o grau de concentração da distribuição de X ao redor do seu valor médio é o segundo, dado por:

$$\mu_2 = E\left[\left(X-\mu_1^{'}\right)^2\right] = \sum_{x \in Z}\left(x-\mu_1^{'}\right)^2 p_X(x)$$

$$\mu_2 = E\left[\left(X-\mu_1^{'}\right)^2\right] = \int_{-\infty}^{+\infty}\left(x-\mu_1^{'}\right)^2 f_X(x)dx$$

respetivamente, nos casos discreto e contínuo. Este segundo momento central de X toma o nome de variância de X. A sua raiz quadrada, tomada com sinal positivo, é designada por desvio-padrão de X, escrevendo-se σ_X, e exprime-se nas mesmas unidades de medida dos valores da variável aleatória.

Assim e em síntese, designar-se-ão os momentos ordinários de X por:

$$\mu_n^{'} = E\left[X^n\right]$$

e os momentos centrais de X por:

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

com $n \in \mathbf{N}_1$ e nas condições de convergência antes referidas.

Ora, como comparar as distribuições de variáveis aleatórias com intervalo de variação, valor médio e desvio-padrão distintos? O indicador preferível, quando é o desvio-padrão o parâmetro usado na medição da dispersão absoluta dos valores de X em torno do seu valor médio, é o designado coeficiente de variação:

$$CV = \frac{\sigma_X}{\mu_1^{'}}$$

que é um valor adimensional e que fornece uma indicação do "peso" do desvio-padrão face ao valor médio. Mede, pois, a dispersão relativa da distribuição de X, ao contrário de σ_X, que mede a dispersão absoluta.

Uma outra caraterística que importa quantificar, quando se estuda determinada distribuição probabilística, é o seu grau de assimetria face ao respetivo valor médio. Uma vez que no caso de ser a distribuição simétrica em relação a esse parâmetro todos os seus momentos centrais de ordem impar são nulos, é natural recorrer a estes com a finalidade de caraterizar o grau de assimetria da distribuição.

Neste sentido, é usual empregar como coeficiente de assimetria de uma variável aleatória X o parâmetro:

$$\gamma_1 = \frac{\mu_3}{\sigma_X^3}$$

que é, por igual, um parâmetro adimensional. Se $\gamma_1 = 0$, a distribuição é simétrica relativamente ao seu valor médio. Se o valor médio de X estiver mais próximo do extremo esquerdo do intervalo de variação, haverá uma predominância da probabilidade na região dos valores do domínio de X superiores a $\mu_1^{'}$, pelo que γ_1 será positivo, dizendo-se que a distribuição tem assimetria positiva. Se $\mu_1^{'}$ estiver mais próximo do extremo direito do intervalo de variação, haverá uma predominância da probabilidade na região dos valores do domínio de X menores que $\mu_1^{'}$, pelo que γ_1 será negativo, sendo a distribuição assimétrica negativa.

Finalmente, o indicador destinado a medir o grau de achatamento da distribuição de X, naturalmente ligado ao grau de concentração da mesma em torno do seu valor médio.

Viu-se já que essa concentração se encontra ligada aos momentos centrais de ordem par, sendo natural escolher os dois primeiros momentos dessa ordem, dado que o de ordem dois, ou variância, se encontra ligado à referida concentração, e que o de ordem quatro será maior que a variância quando a predominância da probabilidade surgir na região dos valores de X distantes de $\mu_1^{'}$ por um valor maior que a unidade, e será menor quando essa predominância surgir na região que satisfaça a condição:

$$\left|X - \mu_1^{'}\right| < 1.$$

Assim, o coeficiente destinado a medir o achatamento - ou kourtosis - da distribuição de uma variável aleatória X será o dado por:

$$\beta = \frac{\mu_4}{\mu_2^2} \quad \Leftrightarrow \quad \gamma_2 = \beta_2 - 3$$

dizendo-se mesocúrtica a distribuição se $\gamma_2 = 0$, leptocúrtica se $\gamma_2 > 0$ e platicúrtica se $\gamma_2 < 0$. E, como facilmente se percebe, também este parâmetro é adimensional.

Convém salientar agora que o operador E, designado esperança, é um operador linear, ou seja, satisfaz a condição:

$$E\left[\sum_{i=1}^{n} \alpha_i X_i\right] = \sum_{i=1}^{n} \alpha_i E\left[X_i\right]$$

com $n \in \mathbf{N}_1$, $\alpha_i \in \mathbf{R}$, e X_i variáveis aleatórias, $(i = 1, \ldots, n)$.

Neste ponto é já fácil estabelecer a relação existente entre os momentos ordinários e os momentos centrais de uma variável aleatória, desde que os mesmo existam.

Como se referiu atrás, o momento central de ordem $n \in \mathbf{N}_1$ de X, no caso de existir, é:

$$\mu_n = E\left[\left(X - \mu_n^{'}\right)^n\right].$$

Ora, desenvolvendo a expressão:

$$\left(X - \mu_1^{'}\right)^n$$

através da Fórmula do Binómio de Newton, virá:

$$\left(X - \mu_1^{'}\right)^n = {}^nC_0 X^n \left(\mu_1^{'}\right)^0 - {}^nC_1 X^{n-1}\left(\mu_1^{'}\right)^1 + {}^nC_2 X^{n-2}\left(\mu_1^{'}\right)^2 - \cdots + (-1)^n\, {}^nC_n X^0 \left(\mu_1^{'}\right)^n$$

pelo que se terá:

$$\mu_n = E\left[\left(X - \mu_1^{'}\right)^n\right] = E\left[\sum_{p=0}^{n} (-1)^p\, {}^nC_p X^{n-p} \left(\mu_1^{'}\right)^p\right] = \sum_{p=0}^{n} (-1)^p\, {}^nC_p \left(\mu_1^{'}\right)^p \mu_{n-p}^{'}.$$

Um caso de grande importância é aquele em que $n = 2$, ou seja:

$$\mu_2 = \mu_2^{'} - \left(\mu_1^{'}\right)^2$$

que surge usualmente com a simbologia:

$$V[X] = E\left[X^2\right] - \left(E[X]\right)^2$$

e que facilita bastante o cálculo do valor da variância de uma variável aleatória.

Mas é também possível exprimir os momentos ordinários em função dos absolutos. Tendo presente que se tem:

$$X^n = \left[\left(X-\mu_1^{'}\right)+\mu_1^{'}\right]^n = \sum_{p=0}^{n} {}^nC_p\left(X-\mu_1^{'}\right)^{n-p}\left(\mu_1^{'}\right)^p$$

virá:

$$\mu_n^{'} = E\left[X^n\right] = \sum_{p=0}^{n} {}^nC_p\,\mu_{n-p}\left(\mu_1^{'}\right)^p$$

que permite, pois, exprimir o momento ordinário de ordem $n \in \mathbf{N}_1$ de X em função dos momentos centrais, embora também de uma potência do valor médio de X .

Mau grado o tema ter sido já abordado atrás, convém salientar que a expressão:

$$\alpha_n^{'} = E\left[|X|^n\right] = \begin{cases} \sum_{x \in Z} |x|^n\, p_X(x) \\ \int_{-\infty}^{+\infty} |x|^n\, f_X(x)dx \end{cases}$$

representa o designado momento absoluto ordinário de ordem $n \in \mathbf{N}_1$ de X , sendo:

$$\alpha_n = E\left[\left|X-\mu_1^{'}\right|^n\right] = \begin{cases} \sum_{x \in Z} \left|x-\mu_1^{'}\right|^n p_X(x) \\ \int_{-\infty}^{+\infty} \left|x-\mu_1^{'}\right|^n f_X(x)dx \end{cases}$$

o momento absoluto central de ordem $n \in \mathbf{N}_1$ de X , respetivamente, nos casos discreto e contínuo.

Ora, o primeiro momento absoluto central de X , designado desvio médio:

$$\alpha_1 = E\left[\left|X-\mu_1^{'}\right|\right]$$

é usado, com grande frequência, como medida da dispersão da distribuição de uma variável aleatória.

Por fim, um outro conceito que importa introduzir, e que é o de momento fatorial de ordem $n \in \mathbf{N}_1$ de X , definido por:

$$\mu_{[X]} = E\left[X(X-1)\ldots(X-n+1)\right] = \begin{cases} \sum_{x \in Z} x(x-1)\ldots(x-n+1)\,p_X(x) \\ \int_{-\infty}^{+\infty} x(x-1)\ldots(x-n+1)\,f_x(x)dx \end{cases}$$

nos casos discreto e contínuo, e em que a série e o integral só existirão se forem absolutamente convergentes.

Ainda antes de abordar as diversas funções geradoras, é essencial fazer uma referência à designada Desigualdade de Jensen, segundo a qual:

$$E\left[g(X)\right] \geq g\left(E[X]\right)$$

onde X é uma variável aleatória e g uma função convexa e mensurável de X, que seja ainda uma variável aleatória. Se g for côncava, o sentido da desigualdade é o inverso.

Esta desigualdade permite, por exemplo, compreender que se tem:

$$E\left[X^2\right] \geq \left(E[X]\right)^2$$

correspondente ao caso em que $g(X) = X^2$, ou:

$$E\left[\frac{1}{X}\right] \geq \frac{1}{E[X]}$$

onde se tem $g(X) = X^{-1}$, ou também:

$$E[\ln X] \leq \ln\left(E[X]\right)$$

onde se tem $g(X) = \ln X$, que é uma função côncava.

Vejam-se, por fim, as diversas funções geradoras, de probabilidades e de momentos, começando pela própria noção de função geradora.

De um modo muito geral, se $(p_n), n \in \mathbf{N}_0$, for uma sucessão de termos em $\mathbf{R}$, à série de potências de t:

$$\Pi(t) = \sum_{n=0}^{+\infty} p_n t^n$$

dá-se o nome de função geradora dos termos da sucessão (p_n), e onde a série se supõe absolutamente convergente no conjunto definido pela condição $|t| < r$, com $r \in \mathbf{R}_0^+$.

Ora, se a anterior sucessão for limitada, é fácil concluir, por comparação com o que se passa com uma série geométrica, que a anterior função geradora é absolutamente convergente no conjunto definido pela condição $|t| < 1$.

Assim, quando se está perante uma variável aleatória discreta, definida em $\mathbf{N}_0$, para a qual se tem:

$$p_n \geq 0 \quad \wedge \quad \sum_{n=0}^{+\infty} p_n = 1 \quad \wedge \quad P\left(X = n\right) = p_n$$

a anterior função geradora toma a designação de função geradora de probabilidades, com a série absolutamente convergente se $|t| \leq 1$, designando-a por $\Pi_X(t)$. Ora, esta função pode definir-se por:

$$\Pi_X(t) = E\left[t^X\right] = p_0 t^0 + p_1 t^1 + p_2 t^2 + \cdots = \sum_{n=0}^{+\infty} p_n t^n .$$

Designa-se esta função por função geradora de probabilidades pelo facto dos coeficientes das potências t^X, serem as probabilidades, p_X, da variável aleatória assumir valores no domínio, $\mathbf{N}_0$.

Note-se, contudo, que a derivada de ordem $n \in \mathbf{N}_0$ de $\Pi_X(t)$, calculada em $t = 1$, representa, precisamente, o momento fatorial de ordem $n \in \mathbf{N}_0$ de X :

$$\Pi_X^{(n)}(1) = \mu_{[n]}.$$

A função geradora de probabilidades pode fornecer excelentes simplificações no cálculo dos valores de $p_n = P(X = n)$, $n \in \mathbf{N}_0$, em situações complicadas.

Conceito de maior interesse, pela muito maior latitude das suas aplicações, é o de função geradora de momentos, que se define por:

$$M_X(t) = E\left[e^{tX}\right]$$

sendo t um parâmetro real.

Uma vez que, como se conhece já:

$$e^{tX} = 1 + tX + \frac{(tX)^2}{2!} + \frac{(tX)^3}{3!} + \frac{(tX)^4}{4!} + \cdots$$

virá, por aplicação do operador esperança, E :

$$\begin{aligned} M_X(t) &= E\left[e^{tX}\right] = E\left[1 + tX + \frac{(tX)^2}{2!} + \frac{(tX)^3}{3!} + \cdots\right] \\ &= E[1] + tE[X] + \frac{t^2}{2!}E\left[X^2\right] + \frac{t^3}{3!}E\left[X^3\right] + \cdots \\ &= 1 + \mu_1' t + \mu_2' \frac{t^2}{2!} + \mu_3' \frac{t^3}{3!} + \cdots \end{aligned}$$

o que mostra que se tem:

$$M_X^{(n)}(0) = \mu_n'$$

com $n \in \mathbf{N}_1$ e $M_X(0) = 1$.

Se em $M_X(t)$ se substituir X por $X - \mu_1'$, e tendo em conta que se tem:

$$e^{t\left(X - \mu_1'\right)} = 1 + t\left(X - \mu_1'\right) + \frac{\left[t\left(X - \mu_1'\right)\right]^2}{2!} + \frac{\left[t\left(X - \mu_1'\right)\right]^3}{3!} + \cdots$$

virá a nova função geradora de momentos centrais, após aplicação do operador esperança, E :

$$M_{X-\mu_1^{'}}(t) = 1 + \mu_1 t + \mu_2 \frac{t^2}{2!} + \mu_3 \frac{t^3}{3!} + \cdots$$

vindo, deste modo:

$$M_{X-\mu_1^{'}}^{(n)}(0) = \mu_n$$

com $n \in \mathbf{N}_1$ e $M_{X-\mu_1^{'}}(0) = 1.$

E torna-se fácil mostrar que se tem:

$$M_X(t) = e^{\mu_1^{'} t} M_{X-\mu_1^{'}}(t)$$

que permite deduzir os valores de $M_X(t)$ a partir dos de $M_{X-\mu_1^{'}}(t)$ e vice-versa.

Convém salientar, porém, que a existência de momentos para uma variável aleatória não garante a existência de função geradora de momentos, o que constitui, pois, uma limitação inerente a este conceito.

Mais importante, contudo, é a função caraterística de uma variável aleatória, uma vez que, contrariamente ao caso da função geradora de momentos, a função caraterística existe sempre, definindo-se como:

$$\psi_X(t) = E\left[e^{itX}\right]$$

com t parâmetro real e onde i é a unidade imaginária.

Dado que se tem:

$$e^{itX} = 1 + itX + \frac{(itX)^2}{2!} + \frac{(itX)^3}{3!} + \frac{(itX)^4}{4!} + \cdots$$

virá:

$$\psi_X(t) = E\left[e^{itX}\right] = 1 + it\mu_1^{'} - \frac{t^2}{2!}\mu_2^{'} - \frac{it^3}{3!}\mu_3^{'} + \frac{t^4}{4!}\mu_4^{'} + \cdots$$

pelo que se tem:

$$\psi_X^{(n)}(0) = i^n \mu_n^{'}$$

com $n \in \mathbf{N}_1$ e $\psi_X(0) = 1.$

Ora, a existência de função geradora de momentos determina a existência de momentos de todas as ordens, o que não se dá com a função caraterística, que pode existir sem que exista, ao menos, o primeiro momento ordinário. E, como se sabe, se existir o momento de ordem $n \in \mathbf{N}_1$ de X, existem os momentos de ordem inferior a n.

Esta função goza de propriedades diversas, a primeira das quais é:

$$\psi_X(0) = 1.$$

E, tendo presente que pela Fórmula de Euler se tem:

$$e^{itX} = \cos(tX) + isen(tX)$$

virá:

$$\psi_X(t) = E\left[e^{itX}\right] = E\left[\cos(tX) + isen(tX)\right]$$

pelo que:

$$\left|\psi_X(t)\right| = \left|E\left[\cos(tX) + isen(tX)\right]\right| \leq E\left[\left|\cos(tX) + isen(tX)\right|\right] = 1$$

como pode provar-se facilmente.

Além do mais, a função caraterística é hermítica, ou seja:

$$\psi_X(-t) = \overline{\psi_X(t)}$$

tendo-se, por igual:

$$\psi_{\alpha X+\beta}(t) = e^{it\beta}\psi(\alpha t)$$

como pode comprovar-se facilmente, e onde α e β são valores reais. Além do mais, a função caraterística é ainda uniformemente contínua em **R**.

Por outro lado, esta função encontra-se ligada à função distribuição, $F_X(x)$, através da condição:

$$\psi_X(t) = E\left[e^{itX}\right] = \int_{-\infty}^{+\infty} e^{itx} dF_X(x)$$

e sendo:

$$\psi_X(t) = \begin{cases} \sum_{x \in Z} e^{itx} p_X(x) \\ \int_{-\infty}^{+\infty} e^{itx} f_X(x)dx \end{cases}$$

nos casos discreto e contínuo, respetivamente. Uma relação que se estende, ainda, à propriedade:

$$F_X(a+h) - F_X(a-h) = \lim_{T \to +\infty} \frac{1}{\pi} \int_{-T}^{T} \frac{sen(ht)}{t} e^{ita} \psi_X(t)dt$$

com $a+h$ e $a-h$ elementos do domínio de $F_X(x)$, aí contínua, e com $h \in \mathbf{R}^+$.

Assim, a função caraterística determina, de um modo único, a função distribuição que lhe corresponde. E vice-versa.

Por fim, e ainda em torno do conceito de função caraterística, note-se que, no caso de existirem, para certa variável aleatória com determinada distribuição, momentos de todas as ordens, e sendo:

$$\psi_X(t) = 1 + \sum_{n=1}^{+\infty} \mu_n^{'} \frac{(it)^n}{n!}$$

a função caraterística de X , pode definir-se a nova função:

$$K_X(t) = \log\left[\psi_X(t)\right]$$

cujo desenvolvimento em série de potências de it é:

$$K_X(t) = \sum_{n=1}^{+\infty} k_n \frac{(it)^n}{n!}$$

e onde aos coeficientes $k_n, n \in \mathbf{N}_1$, se dá a designação de cumulantes da função distribuição de X , $F_X(x)$. Ou seja, $K_X(t)$ é a função geradora dos cumulantes de $F_X(x)$.

Assim, nos termos do que acaba de referir-se, e de quanto se conhece já da Análise Matemática, virá:

$$\psi_X(t) = 1 + \sum_{n=1}^{+\infty} \mu_n^{'} \frac{(it)^n}{n!} \Leftrightarrow \sum_{n=1}^{+\infty} k_n \frac{(it)^n}{n!} = \log\left[1 + \sum_{n=1}^{+\infty} \mu_n^{'} \frac{(it)^n}{n!}\right] \Leftrightarrow e^{\sum_{n=1}^{+\infty} k_n \frac{(it)^n}{n!}} = 1 + \sum_{n=1}^{+\infty} \mu_n^{'} \frac{(it)^n}{n!}$$

por cujo desenvolvimento do segundo membro, e posterior identificação dos coeficientes dos termos:

$$\frac{(it)^n}{n!}$$

com $n \in \mathbf{N}_1$, se obtêm as expressões dos cumulantes, tendo-se, para os três primeiros:

$$k_1 = \mu_1^{'}$$

$$k_2 = \mu_2^{'} - \left(\mu_1^{'}\right)^2 = \mu_2$$

$$k_3 = \mu_3^{'} - 3\mu_2^{'}\mu_1^{'} + 2\left(\mu_1^{'}\right)^3 = \mu_3.$$

Relembrando a noção de transformação integral de uma função, tratada no domínio da Análise Matemática, e aplicando-a aqui à função distribuição, $F_X(x)$, a mesma assume a forma conhecida:

$$\int_{-\infty}^{+\infty} H(x,t)\,dF_X(x)$$

desde que exista em $\mathbf{R}$ este integral, que é o integral de Lebesgue-Stieltjes.

No caso em que $t \in \mathbf{N}_1$, têm interesse muito especial as situações:

$$H(x,t) = x^t$$

$$H(x,t) = |x|^t$$

$$H(x,t) = x_{[t]} = \begin{cases} 1 & \Leftarrow \quad t = 0 \\ x(x-1)\ldots(x-t+1) & \Leftarrow \quad t \in \{1,2,3,\ldots\} \end{cases}$$

ligadas, respetivamente, aos conceitos de momento ordinário, momento absoluto ordinário e momento fatorial.

Em contrapartida, se $t \in \mathbf{R}$, têm uma importância muito particular as três novas situações:

$$K(x,t) = t^x$$

$$K(x,t) = e^{tx}$$

$$K(x,t) = e^{itx}$$

que correspondem, respetivamente, às funções geradoras de probabilidades e de momentos, e à função caraterística.

Espera-se que o presente texto tenha conseguido tornar claros conceitos que são simples e, mais ainda que tudo, tenha contribuído para mostrar a unidade entre momentos e funções geradoras, bem como o respetivo papel na caraterização das distribuições probabilísticas.