Journal of Mathematical Chemistry, Vol. 41, No. 1, January 2007 (© 2006)
DOI: 10.1007/s10910-006-9069-x

Computation of one-sided probability density functions from their cumulants

Mário N. Berberan-Santos
Centro de Química-Física Molecular, Instituto Superior Técnico, 1049-001 Lisboa, Portugal
E-mail: berberan@ist.utl.pt

Received 10 November 2005; revised 10 December 2005

Explicit formulas for the calculation of a one-sided probability density function from its cumulants are obtained and discussed.

KEY WORDS: probability density function, cumulant, truncated Gaussian distribution, exponential distribution

AMS subject classification: 44A10 Laplace transform, 60E10 characteristic functions; other transforms, 62E17 approximations to distributions (nonasymptotic).

1. Introduction

The moment-generating function of a random variable X is by definition [1,2] the integral

$$M(t) = \int_{-\infty}^{\infty} f(x)\, \mathrm{e}^{tx} \mathrm{d}x, \tag{1}$$

where $f(x)$ is the probability density function (PDF) of X.

It is well known that if all moments are finite, the moment-generating function admits a Maclaurin series expansion [1–3],

$$M(t) = \sum_{n=0}^{\infty} m_n \frac{t^n}{n!}, \tag{2}$$

where the raw moments are

$$m_n = \int_{-\infty}^{\infty} x^n f(x)\, \mathrm{d}x \quad (n = 0,\ 1, \ldots). \tag{3}$$

The cumulant-generating function [2–4]

$$K(t) = \ln M(t), \tag{4}$$

admits a similar expansion

$$K(t) = \sum_{n=1}^{\infty} \kappa_n \frac{t^n}{n!}, \tag{5}$$

where the κ_n are the cumulants, defined by

$$\kappa_n = K^{(n)}(0) \qquad (n = 1, 2, \ldots). \tag{6}$$

The first four cumulants are

$$\begin{aligned}
\kappa_1 &= m_1 = \mu \\
\kappa_2 &= m_2 - m_1^2 = \sigma^2 \\
\kappa_3 &= 2m_1^3 - 3m_1m_2 + m_3 = \gamma_1 \sigma^3 \\
\kappa_4 &= -6m_1^4 + 12m_1^2m_2 - 3m_2^2 - 4m_1m_3 + m_4 = \gamma_2 \sigma^4,
\end{aligned} \tag{7}$$

where γ_1 is the skewness and γ_2 is the kurtosis. With the exception of the delta and Gaussian cases, all PDFs have an infinite number of non-zero cumulants.

The raw moments are explicitly related to the cumulants by

$$\begin{aligned}
m_1 &= \kappa_1 \\
m_2 &= \kappa_1^2 + \kappa_2 \\
m_3 &= \kappa_1^3 + 3\kappa_1\kappa_2 + \kappa_3 \\
m_4 &= \kappa_1^4 + 6\kappa_1^2\kappa_2 + 3\kappa_2^2 + 4\kappa_1\kappa_3 + \kappa_4. \\
&\cdots\cdots\cdots\cdots\cdots\cdots
\end{aligned} \tag{8}$$

Under relatively general conditions, the moments (or the cumulants) of a distribution define the respective PDF, as follows from the above equations. It is therefore of interest to know how to build the PDF from its moments (or cumulants). An obvious practical application is to obtain an approximate form of the PDF from a finite set of moments (or cumulants). Some aspects of the dependence of a PDF on its moments were reviewed by Gillespie [5].

In this work, explicit formulas for the calculation of a one-sided PDF from its cumulants are obtained, and their interest and limitations discussed.

2. Computation of a one-sided PDF from its cumulants

2.1. Laplace transform of the PDF

For a one-sided PDF (i.e., defined only for $x \geqslant 0$) it is convenient to consider not the moment-generating or characteristic functions, but instead the closely related Laplace transform,

$$G(t) = L[f(x)] = \int_0^\infty f(x)\, e^{-tx} dx. \tag{9}$$

The respective Maclaurin series is

$$G(t) = \sum_{n=0}^{\infty} \frac{g_n}{n!} t^n, \tag{10}$$

with

$$g_n = G^{(n)}(0) = (-1)^n m_n. \tag{11}$$

In this way, termwise Laplace transform inversion of equation (10) gives the PDF in terms of its moments,

$$f(x) = L^{-1}[G(t)] = \sum_{n=0}^{\infty} (-1)^n \frac{m_n}{n!} \delta^{(n)}(x), \tag{12}$$

where $\delta^{(n)}(x)$ is the nth order derivative of the delta function. Equation (12) again shows that a PDF is completely defined by its moments. This equation, previously obtained by Gillespie [5] in a different way, must nevertheless be understood as a generalized function representation of the PDF [5]. Indeed, equation (12) cannot be used to compute a PDF from its moments. To do so, one must resort to the cumulant expansion, as will be shown.

The modified cumulant-generating function is

$$C_+(t) = \ln G(t), \tag{13}$$

the key point being that this modified cumulant-generating function admits the formal expansion

$$C_+(t) = \sum_{n=1}^{\infty} \frac{c_n}{n!} t^n, \tag{14}$$

where the coefficients are again directly related to the cumulants

$$c_n = C_+^{(n)}(0) = (-1)^n \kappa_n. \tag{15}$$

2.2. *Calculation of $f(x)$ from the cumulants*

The PDF can therefore be written as

$$f(x) = L^{-1}[\mathrm{e}^{C_+(t)}] = L^{-1}\left[\exp\left(\sum_{n=1}^{\infty}(-1)^n \frac{\kappa_n}{n!}\, t^n\right)\right], \tag{16}$$

assuming that the series has a sufficiently large convergence radius.

Application of the analytical inversion formula for Laplace transforms [6] (with $c = 0$) to equation (16) yields

$$f(x) = \frac{2}{\pi}\int_0^{\infty} \mathrm{Re}[\mathrm{e}^{C_+(it)}]\cos(xt)\,\mathrm{d}t, \tag{17}$$

and using equations (14–15), equation (17) becomes

$$f(x) = \frac{2}{\pi}\int_0^{\infty} \exp\left(-\kappa_2\frac{t^2}{2!} + \kappa_4\frac{t^4}{4!} - \cdots\right)\cos\left(\kappa_1 t - \kappa_3\frac{t^3}{3!} + \cdots\right)\cos(xt)\,\mathrm{d}t \quad (x > 0). \tag{18}$$

Two other equivalent inversion forms [6] give

$$f(x) = \frac{2}{\pi}\int_0^{\infty} \exp\left(-\kappa_2\frac{t^2}{2!} + \kappa_4\frac{t^4}{4!} - \cdots\right)\sin\left(\kappa_1 t - \kappa_3\frac{t^3}{3!} + \ldots\right)\sin(xt)\,\mathrm{d}t \quad (x > 0), \tag{19}$$

$$f(x) = \frac{1}{\pi}\int_0^{\infty} \exp\left(-\kappa_2\frac{t^2}{2!} + \kappa_4\frac{t^4}{4!} - \cdots\right)\cos\left(xt - \kappa_1 t + \kappa_3\frac{t^3}{3!} - \cdots\right)\mathrm{d}t \quad (x \geqslant 0). \tag{20}$$

Equations (18–20) allow – at least formally – the calculation of a one-sided PDF from its cumulants, provided the series are convergent in a sufficiently large integration range.

2.3. *Particular case*

If it is assumed that all cumulants but the first two are zero, equation (18) gives

$$\begin{aligned} f(x) &= \frac{2}{\pi}\int_0^{\infty} \exp\left[-\frac{1}{2}(\sigma t)^2\right]\cos(\mu t)\cos(xt)\,\mathrm{d}t \\ &= \frac{1}{\sqrt{2\pi\sigma^2}}\left[1 + \exp\left(\frac{2\mu x}{\sigma^2}\right)\right]\exp\left[-\frac{1}{2}\left(\frac{x+\mu}{\sigma}\right)^2\right], \end{aligned} \tag{21}$$

a PDF that for large μ/σ reduces to the Gaussian PDF. This result is however somewhat deceptive. Indeed, the Laplace transform of equation (21) is

$$G(t) = \frac{1}{2}\exp\left[-\mu t + \frac{1}{2}(\sigma t)^2\right]\left[1 + erf\left(\frac{\mu - \sigma^2 t}{\sqrt{2}\sigma}\right) + \exp(2\mu t)\, erfc\left(\frac{\mu + \sigma^2 t}{\sqrt{2}\sigma}\right)\right], \tag{22}$$

and $G(t)$ can be shown to possess an infinite number of nonzero cumulants. In this way, the parameters μ and σ used in equation (22) to generate the mentioned PDF are not its first two cumulants. With the exception of the delta and normal (Gaussian) distributions, all PDFs have an infinite number of nonzero cumulants, as proved by Marcinkiewicz [7].

3. Application to the truncated Gaussian and to the exponential probability density functions

3.1. Truncated Gaussian PDF

We now consider the truncated Gaussian (i.e., for $x \geqslant 0$ only) PDF. The truncated Gaussian PDF

$$f(x) = \sqrt{\frac{2}{\pi\sigma^2}}\frac{\exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]}{1 + erf\left(\frac{\mu}{\sqrt{2}\sigma}\right)}, \tag{23}$$

has the following Laplace transform,

$$G(t) = \frac{erfc\left(\frac{\sigma^2 t - \mu}{\sqrt{2}\sigma}\right)}{erfc\left(-\frac{\mu}{\sqrt{2}\sigma}\right)}\exp\left[-\mu t + \frac{1}{2}(\sigma t)^2\right], \tag{24}$$

and has an infinite number of non-zero cumulants. Its first cumulant (the mean) is

$$\kappa_1 = \mu + \sqrt{\frac{2}{\pi}}\sigma\frac{\exp\left[-\frac{1}{2}\left(\frac{\mu}{\sigma}\right)^2\right]}{1 + erf\left(\frac{\mu}{\sqrt{2}\sigma}\right)}. \tag{25}$$

For $t \ll \mu/\sigma^2$, $G(t)$ coincides with that of a normal distribution, and

$$G(t) \simeq \exp\left[-\mu t + \frac{1}{2}(\sigma t)^2\right]. \tag{26}$$

This equation has been used for the analysis of dynamic light-scattering data, in order to recover the distribution of particle sizes from the autocorrelation function [8,9], and applies to some luminescence decays for not too long times. The full Gaussian PDF (or a mixture of Gaussian PDFs) [10–12] and a Gaussian PDF truncated at $x_0 > 0$ [13] have indeed been used to describe fluorescence decays, although the mathematical reason invoked for truncation below a certain positive value x_0 [13] is not correct.

3.2. *Exponential PDF*

Consider next the function

$$G(t) = \frac{1}{1+t}. \tag{27}$$

Its inverse Laplace transform is immediate,

$$f(x) = \mathrm{e}^{-x}, \tag{28}$$

How is this exponential PDF recovered from its cumulants?

First, the Maclaurin expansion of $C(t)$ gives

$$C(t) = \ln G(t) = -t + \frac{t^2}{2} - \frac{t^3}{3} + \cdots, \tag{29}$$

hence the cumulants are

$$\kappa_n = (n-1)!. \tag{30}$$

In this way, equation (18) gives

$$f(x) = \frac{2}{\pi}\int_0^\infty \exp\left(-\frac{t^2}{2} + \frac{t^4}{4} - \cdots\right)\cos\left(t - \frac{t^3}{3} + \cdots\right)\cos(xt)\,\mathrm{d}t. \tag{31}$$

Using

$$-\frac{1}{2}\ln(1+t^2) = -\frac{t^2}{2} + \frac{t^4}{4} - \cdots, \tag{32}$$

and

$$\arctan t = t - \frac{t^3}{3} + \frac{t^5}{5}\cdots, \tag{33}$$

one indeed obtains

$$f(x) = \frac{2}{\pi}\int_0^\infty \frac{\cos xt}{1+t^2}\mathrm{d}t = \mathrm{e}^{-x}. \tag{34}$$

Note however that both series in equations (32) and (33) diverge for $t > 1$, and therefore the use of truncated series (i.e., with a finite number of cumulants, whatever their number) in equation (31) does not asymptotically yield the correct PDF. This difficulty may in principle be overcome by the use of Padé approximants [14,15].

4. Discussion and conclusions

In the above, it was implicitly assumed that all moments and cumulants were finite. For some PDFs, however, not all moments and cumulants are finite. For the Lévy PDFs, for instance, only m_1 can be (but is not always) finite. In these cases, the cumulant series expansion is not valid, and equations (18–20) do not apply. Even when all cumulants are finite, the Marcinkiewicz theorem and the eventual finite convergence radii of the Maclaurin series for the cumulants put some limitations on the practical use of equations (18–20). Nevertheless, equations (18–20) show the explicit connection between a one-sided PDF and its cumulants, and allow a formal calculation of the PDF. Direct approximate computation, on the other hand, may not be feasible without the use of further numerical techniques such as Padé approximants.

A question that immediately arises from the consideration of the present results is on their applicability or suitable extension to two-sided PDFs. This will be addressed in a forthcoming paper.

References

[1] H. Poincaré, *Calcul des Probabilités*, 2^{e} éd. (Gauthier-Villars, Paris, 1912).

[2] A. Papoulis, S.U. Pillai, *Probability, Random Variables, and Stochastic Processes*, 4th edn. (McGraw-Hill, 2001).

[3] R.A. Fisher, Proc. London Math. Soc. (Series 2) 30 (1929) 199.

[4] N.G. van Kampen, *Stochastic Processes in Physics and Astronomy*, 2nd edn. (North-Holland, Amsterdam, 1992).

[5] D.T. Gillespie, Am. J. Phys. 49 (1981) 552.

[6] M.N. Berberan-Santos, J. Math. Chem. 38 (2005) 165.

[7] C.W. Gardiner, *Handbook of Stochastic Methods*, 2nd edn. (Springer-Verlag, Berlin, 1985).

[8] D.E. Koppel, J. Chem. Phys. 57 (1972) 4814.

[9] B.J. Frisken, Appl. Opt. 40 (2001) 4087.

[10] D.R. James, Y.-S. Liu, P. de Mayo and W.R. Ware, Chem. Phys. Letters 120 (1985) 460.

[11] D.R. James and W.R. Ware, Chem. Phys. Letters 126 (1986) 7.

[12] J.R. Alcala, E. Gratton and F.G. Prendergast, Biophys. J. 51 (1987) 587.

[13] G. Verbeek, A. Vaes, M. Van der Auweraer, F.C. De Schryver and C. Geelen, Macromolecules 26 (1993) 472.

[14] C.M. Bender and S.A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers* (McGraw-Hill, Tokyo, 1978).

[15] G.A. Baker Jr. and P. Graves-Morris, *Padé Approximants*, 2nd edn. (Cambridge University Press, Cambridge, 1996).