Journal of Mathematical Chemistry, Vol. 42, No. 3, October 2007 (© 2006)
DOI: 10.1007/s10910-006-9134-5

Expressing a probability density function in terms of another PDF: A generalized Gram-Charlier expansion

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

Received 21 March 2006; revised 11 April 2006

An explicit formula relating the probability density function with its cumulants is derived and discussed. A generalization of the Gram-Charlier expansion is presented, allowing to express one PDF in terms of another. The coefficients of this general expansion are explicitly obtained.

KEY WORDS: probability density function, cumulant, Gram-Charlier expansion, Hermite polynomials

AMS subject classification: 60E10 characteristic functions; other transforms, 62E17 approximations to distributions (non-asymptotic), 62E20 asymptotic distribution theory

1. Introduction

The moment-generating function of a random variable is by definition [1–3] 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 the random variable.

It is well known that if all moments are finite (this will be assumed throughout the work), 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

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

On the other hand, the cumulant-generating function [3,4]

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

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

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

The recurrence relation for the moments, obtained from

$$M(t) = \sum_{n=0}^{\infty} m_n \frac{t^n}{n!} = \exp\left(\sum_{n=1}^{\infty} \kappa_n \frac{t^n}{n!}\right) \tag{7}$$

by taking nth order derivatives at $t = 0$, is

$$m_{n+1} = \sum_{p=0}^{n} \binom{n}{p} m_{n-p}\, \kappa_{p+1} \ . \tag{8}$$

The raw moments are therefore explicitly related to the cumulants by

$$\begin{aligned}
m_1 &= \kappa_1,\\
m_2 &= \kappa_1^2 + \kappa_2,\\
m_3 &= \kappa_1^3 + 3\kappa_1\kappa_2 + \kappa_3,\\
m_4 &= \kappa_1^4 + 6\kappa_1^2\kappa_2 + 3\kappa_2^2 + 4\kappa_1\kappa_3 + \kappa_4,\\
m_5 &= \kappa_1^5 + 10\kappa_1^3\kappa_2 + 15\kappa_1\kappa_2^2 + 10\kappa_1^2\kappa_3 + 10\kappa_2\kappa_3 + 5\kappa_1\kappa_4 + \kappa_5,\\
m_6 &= \kappa_1^6 + 15\kappa_1^4\kappa_2 + 45\kappa_1^2\kappa_2^2 + 15\kappa_2^3 + 20\kappa_1^3\kappa_3 + 60\kappa_1\kappa_2\kappa_3 + 10\kappa_3^2\\
&\quad + 15\kappa_1^2\kappa_4 + 15\kappa_2\kappa_4 + 6\kappa_1\kappa_5 + \kappa_6,
\end{aligned} \tag{9}$$

Equation (8) can be solved for κ_{n+1},

$$\kappa_{n+1} = m_{n+1} - \sum_{p=0}^{n-1} \binom{n}{p} m_{n-p}\, \kappa_{p+1}. \tag{10}$$

An explicit formula giving the cumulants in terms of the moments is also known [3]. The first four cumulants are

$$\begin{aligned}
\kappa_1 &= m_1 = \mu,\\
\kappa_2 &= m_2 - m_1^2 = \sigma^2,\\
\kappa_3 &= 2m_1^3 - 3m_1m_2 + m_3 = \gamma_1\,\sigma^3,\\
\kappa_4 &= -6m_1^4 + 12m_1^2m_2 - 3m_2^2 - 4m_1m_3 + m_4 = \gamma_2\,\sigma^4,
\end{aligned} \tag{11}$$

where γ_1 is the skewness and γ_2 is the kurtosis. Of all cumulants, only the second is necessarily non-negative.

With the exception of the delta and normal cases, all PDFs have an infinite number of non-zero cumulants [3,5].

A third canonical function of the PDF is the characteristic function [1–3], defined by

$$\Phi(t) = \int_{-\infty}^{\infty} f(x)\, \mathrm{e}^{itx} \mathrm{d}x \tag{12}$$

that is simply the Fourier transform of the PDF. The PDF can therefore be written as the inverse Fourier transform of its characteristic function,

$$f(x) = F^{-1}[\Phi(t)] = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Phi(t)\, \mathrm{e}^{-itx} \mathrm{d}t. \tag{13}$$

The characteristic function admits the Maclaurin expansion,

$$\Phi(t) = \sum_{n=0}^{\infty} m_n \frac{(it)^n}{n!}\,. \tag{14}$$

Taking into account that

$$\delta^{(n)}(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} (-it)^n \mathrm{e}^{-itx} \mathrm{d}x, \tag{15}$$

where $\delta^{(n)}(x)$ is the nth derivative of the delta function, termwise Fourier inversion of equation (14) yields the following formal PDF expansion, obtained in a different way by Gillespie [6]

$$f(x) = \sum_{n=0}^{\infty} (-1)^n \frac{m_n}{n!}\, \delta^{(n)}(x). \tag{16}$$

From the characteristic function a cumulant generating function can also be defined,

$$C(t) = \ln \Phi(t), \tag{17}$$

whose series expansion gives

$$C(t) = \sum_{n=1}^{\infty} \kappa_n \frac{(it)^n}{n!}\,. \tag{18}$$

Under relatively general conditions [3,7], the moments (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 (cumulants). An obvious practical application is to obtain an approximate form of the

PDF from a finite set of moments (cumulants). This is an important problem that has received much attention for a long time (see [3] and references therein).

In this work, an explicit formula relating a PDF with its cumulants is obtained, and its interest and limitations discussed. A generalization of the Gram-Charlier expansion that allows to express a PDF in terms of another PDF is obtained in two different ways. The procedure for the calculation of the coefficients of the general expansion is also given.

2. Calculation of a PDF from its cumulants

Using equations (13), (17), and (18), the PDF can be written as

$$f(x) = F^{-1}[\mathrm{e}^{C(t)}] = \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp\left(\sum_{n=1}^{\infty} \kappa_n \frac{(it)^n}{n!}\right) \exp(-ixt)\,\mathrm{d}t, \tag{19}$$

assuming that the series has a sufficiently large convergence radius. Taking into account that the PDF is a real function, equation (19) can be simplified to give

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

Finally, because the integrand is of even parity, equation (20) becomes

$$f(x) = \frac{1}{\pi}\int_{0}^{\infty} \exp\left(-\kappa_2\frac{t^2}{2!} + \kappa_4\frac{t^4}{4!} - \cdots\right) \cos\left(xt - \kappa_1 t + \kappa_3\frac{t^3}{3!} - \cdots\right) \mathrm{d}t. \tag{21}$$

Equations (20 and 21) explicitly relate a PDF with its cumulants and allow – at least formally – its calculation from the cumulants, provided the series involved are convergent in a sufficiently large integration range. Equation (20) was previously obtained for one-sided PDFs [8] by means of analytical Laplace transform real inversion [9].

3. Finite number of nonzero cumulants: dirac delta and normal PDFs

If all cumulants but the first are zero, equation (20) yields the delta function, as expected,

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \cos\left[(x-\kappa_1)t\right] \mathrm{d}t = \delta\left(x-\kappa_1\right). \tag{22}$$

If it is assumed that all cumulants but the first two are zero, the result is now the normal (or Gaussian) PDF, $\varphi(x)$,

$$\varphi(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\exp\left[-\frac{1}{2}(\sigma t)^2\right]\cos\left[(x-\mu)t\right]\,\mathrm{d}t=\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]. \tag{23}$$

No further PDFs correspond to a finite set of cumulants. With the exception of the delta and normal distributions, all PDFs have an infinite number of nonzero cumulants, as proved by Marcinkiewicz [3,5].

4. Derivatives of the normal function and an integral representation for Hermite polynomials

An interesting outcome of the integral representation of the normal function, equation (23), is a compact form for its derivatives:

$$\varphi^{(n)}(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}t^n\exp\left[-\frac{1}{2}(\sigma t)^2\right]\cos\left[(x-\mu)t+\frac{n\pi}{2}\right]\,\mathrm{d}t. \tag{24}$$

As a simple application of this result, take the Hermite polynomials, defined by the Rodrigues formula,

$$H_n(x)=(-1)^n\mathrm{e}^{x^2}\frac{\mathrm{d}^n}{\mathrm{d}x^n}\mathrm{e}^{-x^2}. \tag{25}$$

Using equation (24), a known integral representation for these polynomials follows immediately,

$$\begin{aligned}H_n(x)&=\frac{(-1)^n}{2\sqrt{\pi}}\mathrm{e}^{x^2}\int_{-\infty}^{\infty}t^n\exp\left(-\frac{t^2}{4}\right)\cos\left(xt+\frac{n\pi}{2}\right)\,\mathrm{d}t\\&=\frac{2^{n+1}}{\sqrt{\pi}}\mathrm{e}^{x^2}\int_{0}^{\infty}u^n\exp\left(-u^2\right)\cos\left(2xu-\frac{n\pi}{2}\right)\,\mathrm{d}u.\end{aligned} \tag{26}$$

Note that equation (24) also results from the identity for Fourier transforms that enables to express the nth order derivative of a function in terms of its Fourier transform,

$$f^{(n)}(x)=\frac{(-1)^n}{2\pi}\int_{-\infty}^{\infty}(it)^n F\left[f(x)\right]\mathrm{e}^{-itx}\mathrm{d}t, \tag{27}$$

or, given that $f(x)$ is real,

$$f^{(n)}(x)=\frac{(-1)^n}{2\pi}\int_{-\infty}^{\infty}t^n Re\left[i^n F\left[f(x)\right]\mathrm{e}^{-itx}\right]\mathrm{d}t. \tag{28}$$

5. Gram-Charlier expansion

The usual form of the Gram-Charlier expansion (the so-called Type A series) is an expansion of a PDF about a normal distribution with common μ and σ (i.e., common cumulants κ_1 and κ_2) [3,10,11]. This expansion that finds applications in many areas, including finance [12], analytical chemistry [13,14], spectroscopy [15], fluid mechanics [16], and astrophysics and cosmology [17–19], can be straightforwardly obtained from equation (20). Indeed, taking into account equation (23), equation (20) can be rewritten as

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\exp\left[-\frac{1}{2}(\sigma t)^2\right]\cos(xt-\mu t)\cos\left(\kappa_3\frac{t^3}{3!}-\cdots\right)\exp\left(\kappa_4\frac{t^4}{4!}-\cdots\right)\mathrm{d}t + \\ -\frac{1}{2\pi}\int_{-\infty}^{\infty}\exp\left[-\frac{1}{2}(\sigma t)^2\right]\sin(xt-\mu t)\sin\left(\kappa_3\frac{t^3}{3!}-\cdots\right) \\ \exp\left(\kappa_4\frac{t^4}{4!}-\cdots\right)\mathrm{d}t. \tag{29}$$

Expansion of $\cos\left(\kappa_3\frac{t^3}{3!}-\cdots\right)\exp\left(\kappa_4\frac{t^4}{4!}-\cdots\right)$ and of $\sin\left(\kappa_3\frac{t^3}{3!}-\cdots\right)\exp\left(\kappa_4\frac{t^4}{4!}-\cdots\right)$ about $t=0$ gives

$$\cos\left(\kappa_3\frac{t^3}{3!}-\kappa_5\frac{t^5}{5!}+\cdots\right)\exp\left(\kappa_4\frac{t^4}{4!}-\kappa_6\frac{t^6}{6!}+\cdots\right) = 1+\frac{\kappa_4}{4!}t^4-\frac{1}{6!}\left(10\kappa_3^2+\kappa_6\right)t^6+\cdots \\ \sin\left(\kappa_3\frac{t^3}{3!}-\kappa_5\frac{t^5}{5!}+\cdots\right)\exp\left(\kappa_4\frac{t^4}{4!}-\kappa_6\frac{t^6}{6!}+\cdots\right) = \frac{\kappa_3}{3!}t^3-\frac{\kappa_5}{5!}t^5+\cdots \tag{30}$$

and therefore, using equation (24), equation (29) becomes the Gram-Charlier series,

$$f(x) = \varphi(x) - \frac{\kappa_3}{3!}\varphi^{(3)}(x) + \frac{\kappa_4}{4!}\varphi^{(4)}(x) - \frac{\kappa_5}{5!}\varphi^{(5)}(x) \\ +\frac{1}{6!}\left(10\kappa_3^2+\kappa_6\right)\varphi^{(6)}(x)+\cdots \tag{31}$$

The terms of this series, including the respective numerical coefficients, are frequently written as functions of the Hermite polynomials (cf. equations (24–26)), but the above form is much simpler for a general, unstandardized variable, compare, e.g. with the explicit form of the coefficients given in [16,18]. The general form of the coefficients in terms of cumulants, not found in [3,10,13–19], is given in the next section.

6. A general expansion

The derivation process of the preceding section can also be used to obtain the expansion about a given PDF $\psi(x)$ with cumulants κ_1^0, κ_2^0, $\kappa_3^0, \ldots$, by noting that the nth order derivative of a PDF $f(x)$ is (compare equation (24)),

$$f^{(n)}(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}t^n \exp\left(-\kappa_2\frac{t^2}{2!}+\kappa_4\frac{t^4}{4!}-\cdots\right)\cos\left(xt-\kappa_1 t+\kappa_3\frac{t^3}{3!}-\cdots+\frac{n\pi}{2}\right)\mathrm{d}t. \tag{32}$$

If the desired PDF $f(x)$ has cumulants given by

$$\kappa_n=\kappa_n^0+\delta_n, \quad (n=1,\ 2,\ldots,), \tag{33}$$

equation (20) becomes

$$f(x)=\sum_{n=0}^{\infty}(-1)^n\frac{\alpha_n}{n!}\,\psi^{(n)}(x), \tag{34}$$

where

$$\begin{aligned}
\alpha_0 &= 1,\\
\alpha_1 &= \delta_1,\\
\alpha_2 &= \delta_1^2+\delta_2,\\
\alpha_3 &= \delta_1^3+3\delta_1\delta_2+\delta_3,\\
\alpha_4 &= \delta_1^4+6\delta_1^2\delta_2+3\delta_2^2+4\delta_1\delta_3+\delta_4,\\
\alpha_5 &= \delta_1^5+10\delta_1^3\delta_2+15\delta_1\delta_2^2+10\delta_1^2\delta_3+10\delta_2\delta_3+5\delta_1\delta_4+\delta_5,\\
\alpha_6 &= \delta_1^6+15\delta_1^4\delta_2+45\delta_1^2\delta_2^2+15\delta_2^3+20\delta_1^3\delta_3+60\delta_1\delta_2\delta_3+10\delta_3^2\\
&\quad +15\delta_1^2\delta_4+15\delta_2\delta_4+6\delta_1\delta_5+\delta_6.
\end{aligned} \tag{35}$$

These relations have precisely the form of those that give the raw moments in terms of the cumulants [cf. equations (9)]. That this must be so, becomes clear if one chooses the delta function as the base PDF, in which case equation (34) reduces to the formal expansion equation (16).

The coefficients of an expansion where n cumulants are matched are obtained simply by dropping all terms containing δ_k, with $k \leqslant n$, in equations (35). This provides in particular a systematic way to obtain the coefficients of the Gram-Charlier expansion, equation (31), where $\psi(x)=\varphi(x)$ and $\delta_1=\delta_2=0$:

$$\begin{aligned}
\alpha_0 &= 1,\\
\alpha_1 &= 0,\\
\alpha_2 &= 0,\\
\alpha_3 &= \delta_3=\kappa_3,\\
\alpha_4 &= \delta_4=\kappa_4,\\
\alpha_5 &= \delta_5=\kappa_5,\\
\alpha_6 &= 10\delta_3^2+\delta_6=10\kappa_3^2+\kappa_6.
\end{aligned} \tag{36}$$

It has been argued [16] that in some cases it may not be appropriate to match the variances in the Gram-Charlier series. This should then correspond to the expansion with $n = 1$. In such a case, the form of the coefficients of equation (34) becomes identical to the relations that give the central moments in terms of the cumulants:

$$\begin{aligned}
\alpha_0 &= 1,\\
\alpha_1 &= 0,\\
\alpha_2 &= \delta_2,\\
\alpha_3 &= \kappa_3,\\
\alpha_4 &= 3\delta_2^2 + \kappa_4,\\
\alpha_5 &= 10\delta_2\kappa_3 + \kappa_5,\\
\alpha_6 &= 15\delta_2^3 + 10\kappa_3^2 + 15\delta_2\kappa_4 + \kappa_6
\end{aligned} \tag{37}$$

where $\delta_2 = \sigma^2 - \kappa_2$.

It may be noted that the PDF normalization condition applied to equation (34) gives

$$\sum_{n=1}^{\infty} (-1)^n \frac{\alpha_n}{n!} \int_{-\infty}^{\infty} \psi^{(n)}(x)\,\mathrm{d}x = 0. \tag{38}$$

This condition is automatically fulfilled as

$$\int_{-\infty}^{\infty} \psi^{(n)}(x)\,\mathrm{d}x = \psi^{(n-1)}(+\infty) - \psi^{(n-1)}(-\infty) = 0 \tag{39}$$

given that for a continuous $\psi(x)$ the function and all its derivatives $\psi^{(n)}(x)$vanish as x goes to infinity.

The series in equation (34) is not necessarily convergent, even when $\psi(x) = \varphi(x)$[10]. Convergence is assured for square integrable PDFs when equation (34) coincides with a generalized Fourier series whose basis functions are orthogonal eigenfunctions of a regular Sturm-Liouville problem [20].

7. Compact derivation of the general expansion

Applying Fourier transforms to equation (34), one obtains

$$\Phi(t) = \sum_{n=0}^{\infty} (-1)^n \frac{\alpha_n}{n!} F\left[\psi^{(n)}(x)\right] \tag{40}$$

or

$$\Phi(t) = \left[\sum_{n=0}^{\infty} \alpha_n \frac{(it)^n}{n!}\right] F\left[\psi(x)\right] = \exp\left[\sum_{n=1}^{\infty} \delta_n \frac{(it)^n}{n!}\right] \Phi^0(t), \tag{41}$$

where $\Phi^0(t)$ is the characteristic function of $\psi(t)$, as could be expected. Indeed, a compact derivation of equation (34) is as follows:

$$\Phi(t) = \exp\left[\sum_{n=1}^{\infty} \kappa_n \frac{(it)^n}{n!}\right] = \exp\left[\sum_{n=1}^{\infty} \delta_n \frac{(it)^n}{n!}\right] \exp\left[\sum_{n=1}^{\infty} \kappa_n^0 \frac{(it)^n}{n!}\right]$$
$$= \left[\sum_{n=0}^{\infty} \alpha_n \frac{(it)^n}{n!}\right] F\left[\psi(x)\right]. \tag{42}$$

Inverse Fourier transform gives

$$f(x) = \sum_{n=0}^{\infty} \alpha_n \frac{(-1)^n}{n!} \delta^{(n)}(x) \otimes \psi(x), \tag{43}$$

where $\otimes$ stands for the convolution between two functions. Using

$$\delta^{(n)}(x) \otimes \psi(x) = \psi^{(n)}(x), \tag{44}$$

equation (34) is immediately obtained.

8. Conclusions

Equations (20) and (21) show the explicit connection between a PDF and its cumulants, and allow a formal calculation of the PDF. Using these equations, an integral representation of the Hermite polynomials was obtained, equation (26). A generalization of the Gram-Charlier expansion was derived, equation (34), and a general procedure for the determination of the respective coefficients given.

References

[1] H. Poincaré, *Calcul des Probabilités*, 2nd edn. (Gauthier-Villars, Paris, 1912).
[2] A. Papoulis and S.U. Pillai, *Probability, Random Variables, and Stochastic Processes*, 4th edn. (McGraw-Hill, New York 2001).
[3] A. Stuart and J.K. Ord, *Kendall's Advanced Theory of Statistics*, 6th edn., Vol. 1 (Hodder Arnold, London, 1994).
[4] R.A. Fisher, Proc. London Math. Soc. (Series 2) 30 (1929) 199.
[5] C.W. Gardiner, *Handbook of Stochastic Methods*, 2nd edn. (Springer-Verlag, Berlin, 1985).
[6] D.T. Gillespie, Am. J. Phys. 49 (1981) 552.
[7] P. Lévy, *Théorie de l'Addition des Variables Aléatoires* (Gauthier-Villars, Paris, 1954).
[8] M.N. Berberan-Santos, J. Math. Chem. in press.
[9] M.N. Berberan-Santos, J. Math. Chem. 38 (2005) 165.
[10] H. Cramér, *Mathematical Methods of Statistics* (Princeton University Press, Princeton, 1946, 19th printing, 1999).

[11] The expansion is named after the Danish mathematician Jorgen P. Gram (1850–1916) and the Swedish astronomer Carl V. L. Charlier (1862–1934). For historical accounts of the origin of the cumulants and of the Gram-Charlier expansion, see A. Hald, Int. Stat. Rev. 68 (2000) 137, and also A.W. Davies, Gram-Charlier Series, in: *Encyclopedia of Statistical Sciences*, 2nd edn., Vol. 3 (Wiley-Interscience, New York, 2005).
[12] E. Jondeau and M. Rockinger, J. Econ. Dyn. Control 25 (2001) 1457.
[13] J. Olivé and J.O. Grimalt, Anal. Chim. Acta 249 (1991) 337.
[14] V.B. Di Marco and G.G. Bombi, J. Chromatogr. A 931 (2001) 1.
[15] M.C.M. O'Brien, J. Phys.: Condens. Matter 4 (1992) 2347.
[16] A. Pelliccioni, Il Nuovo Cimento 20 (1997) 435.
[17] R. Juszkiewicz, D.H. Weinberg, P. Amsterdamski, M. Chodorowski and F. Bouchet, Astrophys. J. 442 (1995) 39.
[18] S. Blinnikov and R. Moessner, Astron. Astrophys. Suppl. Ser. 130 (1998) 193.
[19] C.R. Contaldi, R. Bean and J. Magueijo, Phys. Lett. B 468 (1999) 189.
[20] D.A. McQuarrie, *Mathematical Methods for Scientists and Engineers* (University Science Books, Sausalito, 2003).