ORIGINAL ARTICLE

Draft Version 0.1

A revisit of the Gram-Charlier and Edgeworth series expansions

Torgeir Brenn[1] | Stian Normann Anfinsen[1]

[1]University of Tromsø - The Arctic University of Norway, Department of Physics and Technology, Machine Learning Group, 9037 Tromsø, Norway

Correspondence
Stian Normann Anfinsen, UiT The Arctic University of Norway, Department of Physics and Technology, 9037 Tromsøø, Norway
Email: correspondingauthor@email.com

Funding information

In this paper we make several observations on the Gram-Charlier and Edgeworth series, which are methods for modeling and approximating probability density functions. We present a simplified derivation which highlights both the similarity and the differences of the series expansions, that are often obscured by alternative derivations. We also introduce a reformulation of the Edgeworth series in terms of the complete exponential Bell polynomials, which make both series easy to implement and evaluate. The result is a significantly more accessible methodology, in the sense that it is easier to understand and to implement. Finally, we also make a remark on the Gram-Charlier series with a gamma kernel, providing a novel and simple expression for its coefficients.

KEYWORDS
probability density functions, series expansions, Edgeworth, Gram-Charlier, Bell polynomials, Normal kernel, gamma kernel

1 | INTRODUCTION

The Gram-Charlier and Edgeworth series expansions provide attractive alternatives when it comes to probability density function (PDF) estimation. They combine the simplicity of fitting a two-parameter PDF with the flexibility of correcting for higher order moments, often resulting in fast and accurate approximations.

Abbreviations: CF, characteristic function; FT, Fourier transform; IID, independent and identically distributed; PDF, probability density function; RV, random variable.

The series expansion methods were introduced in the late 19th and early 20th century (Chebyshev, 1860; Gram, 1883; Thiele, 1889; Chebyshev, 1890; Thiele, 1903; Edgeworth, 1905; Charlier, 1905, 1906), and their history is summarized in (Wallace, 1958) and (Hald, 2000). The methods can be derived in terms of orthogonal polynomials (Kendall et al., 1994), but it is also possible to retrieve the PDF from the characteristic function (CF) (Lévy, 1925; Lukacs, 1970), as in (Wallace, 1958; Blinnikov and Moessner, 1998).

Traditionally, the series expansions used the normalian PDF as the kernel almost exclusively, but (Kendall et al., 1994) mentioned other possibilities and (Gaztanaga et al., 2000) presented an explicit expression for the Gram-Charlier series with a gamma PDF kernel. A tool which was not available to Gram, Charlier, Edgeworth etc. are the Bell polynomials, named after Eric Temple Bell, who introduced them under the name *partition polynomials* in (Bell, 1927). Among other things, they can be used to retrieve moments from cumulants (Pitman, 2002; Rota and Shen, 2000). Thiele, who introduced the cumulants in (Thiele, 1889), certainly mastered their relationship with the moments, but naturally did not have access to the Bell polynomials. Instead, he gave a recursion formula to compute cumulants from moments (Hald, 2000).

Today, polynomials like those named after Bell or Kummer (Kummer, 1837; Daalhuis, 2010) are readily available online and in mathematics software. The implication that using these polynomials to express the Gram-Charlier and Edgeworth series allows for easier and faster implementation, as demonstrated in (Withers and Nadarajah, 2009, 2015).

This paper is organized as follows. In Section 2 we briefly account for the necessary theoretical background, including moments, cumulants, CF, Bell polynomials and the traditional way of deriving the Gram-Charlier and Edgeworth series. In Section 3 we present our derivation of the same Gram-Charlier and Edgeworth series with the novel application of the Bell polynomials in this context. We make an observation on the Gram-Charlier series around the gamma kernel in Section 4 and present our conclusions in Section 5.

2 | THEORETICAL BACKGROUND

2.1 | The Hermite Polynomials and the Normal Distribution

The nth probabilists' Hermite polynomial $H_n(x)$ is defined in terms of the derivatives of the standardized (zero mean, unit variance) normal PDF $\phi(x) = (2\pi)^{-1/2}e^{-x^2/2}$, namely

$$(-\mathrm{D}_x)^n\phi(x) = H_n(x)\phi(x), \tag{1}$$

where $\mathrm{D}_x = \mathrm{d}/\mathrm{d}x$ is the differential operator and the factor $(-1)^n$ ensures that the leading coefficient of $H_n(x)$ is 1 (Kendall et al., 1994). For arbitrary mean μ and variance σ^2, we define

$$\phi(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} = \frac{1}{\sigma}\phi\left(\frac{x-\mu}{\sigma}\right), \tag{2}$$

and letting $y = (x-\mu)/\sigma$ we see that $\mathrm{D}_y = \sigma^{-1}\mathrm{D}_x$, i.e.

$$(-\mathrm{D}_x)^n\phi(x;\mu,\sigma) = \frac{1}{\sigma^n}H_n\left(\frac{x-\mu}{\sigma}\right)\phi(x;\mu,\sigma). \tag{3}$$

We can now work with the standardized $\phi(x)$ during the derivations for brevity and use (3) to generalize to arbitrary kernel mean and variance. The Hermite polynomials are orthogonal with respect the normal PDF in the sense that

(Kendall et al., 1994)

$$\int_{-\infty}^{\infty} H_k(x)H_n(x)\phi(x)\mathrm{d}x = \begin{cases} n! & ; \quad k = n\,, \\ 0 & ; \quad k \neq n\,. \end{cases} \tag{4}$$

2.2 | The Gram-Charlier Series with the Normal Kernel

Suppose now that the random variable (RV) X has unknown PDF $f_X(x)$, which can be written in terms of the derivatives of the normal PDF, i.e.

$$f_X(x) = \sum_{n=0}^{\infty} a_n H_n(x)\phi(x), \tag{5}$$

where we used (1). To find the coefficients a_n, we multiply both sides with $H_k(x)$, integrate from $-\infty$ to $+\infty$, swap the order of integration and summation, and use the orthogonal property from (4) to get

$$\int_{-\infty}^{\infty} f_X(x)H_k(x)\mathrm{d}x = \int_{-\infty}^{\infty} \sum_{n=0}^{\infty} a_n H_k(x)H_n(x)\phi(x) = a_n n!\,, \tag{6}$$

$$a_n = \frac{1}{n!}\int_{-\infty}^{\infty} f_X(x)H_n(x)\mathrm{d}x\,. \tag{7}$$

Since $H_n(x)$ is a polynomial in x, the coefficients a_n must be a linear combination of the moments $\mu_\nu = \mathrm{E}\{X^\nu\}$ of X. (Kendall et al., 1994) lists the first few coefficients both in terms of the moments and in terms of the cumulants, with the latter representation presented in (13).

2.3 | The Gram-Charlier Series with the Gamma Kernel

Let the gamma distribution PDF with shape ϕ and scale β be denoted

$$\gamma(x;\phi,\beta) = \beta^{\phi+1}x^{\phi}e^{-\beta x}/\Gamma(\phi+1) \tag{8}$$

for $x \geq 0$. The generalized Laguerre polynomial of degree n and order ϕ is implicitly defined as (Szeg, 1939)

$$L_n^{(\phi)}(x)\gamma(x;\phi) = \frac{1}{n!}\mathrm{D}^n[x^n\gamma(x;\phi)]\,, \tag{9}$$

where we take $\gamma(x;\phi)$ to mean that $\beta = 1$ for brevity, which is easily generalized by replacing x with βx as the argument of both $L_n^{(\phi)}(x)$ and $\gamma(x;\phi)$. The Laguerre polynomials have an orthogonality property[1] analogous to (4), namely

$$\int_{0}^{\infty} L_n^{(\phi)}(x)L_k^{(\phi)}(x)\gamma(x;\phi)\mathrm{d}x = \begin{cases} \dfrac{\Gamma(n+\phi+1)}{\Gamma(n+1)\Gamma(\phi+1)} & ; \quad k = n\,, \\ 0 & ; \quad k \neq n\,. \end{cases} \tag{10}$$

[1] This is found in terms of binomial coefficients in (Szeg, 1939), with (Fowler, 1996) providing a generalization to non-integer arguments. This can also be generalized by replacing x with βx on the left hand side in (10), leaving the right hand side unchanged.

From (Kendall et al., 1994) and (Gaztanaga et al., 2000) we know that a PDF $f_X(x)$ which is zero for negative x, can be written as

$$f_X(x) = \sum_{n=0}^{\infty} a_n L_n^{(\phi)}(\beta x)\gamma(x;\phi,\beta)\,, \tag{11}$$

where the coefficients a_n are found the same way as in Section 2.2, giving

$$a_n = \frac{n!}{\prod_{i=1}^{n}(\phi+i)} \int_0^{\infty} f_X(x) L_n^{(\phi)}(\beta x)\mathrm{d}x, \tag{12}$$

which clearly is a linear combination of the moments of X by the same reasoning as in Section 2.2. The first few terms of the series is given in (14).

2.4 | The Edgeworth Series

For the RV X with unknown PDF $f_X(x)$, its CF $\Psi_X(t)$ is the FT of $f_X(x)$ (Kendall et al., 1994), that is

$$\Psi_X(t) \equiv \mathcal{F}[f_X(x)](t) = \int_{-\infty}^{\infty} e^{itx} f_X(x)\,\mathrm{d}x = \mathrm{E}\{e^{itX}\}\,, \tag{16}$$

where j is the imaginary unit, the expectation $\mathrm{E}\{\cdot\}$ is with respect to x and t is a transform variable. The cumulant generating function is defined as

$$\log \Psi_X(t)\,, \tag{17}$$

$$f_X(x) = \left[1 + \frac{\kappa_3}{6\sigma^3} H_3\left(\frac{x-\mu}{\sigma}\right) + \frac{\kappa_4}{24\sigma^4} H_4\left(\frac{x-\mu}{\sigma}\right) + \frac{\kappa_5}{120\sigma^5} H_5\left(\frac{x-\mu}{\sigma}\right) + \frac{\kappa_6 + 10\kappa_3^2}{720\sigma^6} H_6\left(\frac{x-\mu}{\sigma}\right) + \frac{\kappa_7 + 35\kappa_3\kappa_4}{5040\sigma^7} H_7\left(\frac{x-\mu}{\sigma}\right) + \cdots\right]\phi(x;\mu,\sigma) \tag{13}$$

$$f_X(x) = \left[1 + \left(-\frac{\beta^3\mu_3}{(\phi+1)(\phi+2)(\phi+3)} + 1\right) L_3^{(\phi)}(\beta x) + \left(\frac{\beta^4\mu_4}{(\phi+1)\cdots(\phi+4)} - \frac{4\beta^3\mu_3}{(\phi+1)(\phi+2)(\phi+3)} + 3\right) L_4^{(\phi)}(\beta x) + \cdots\right]\gamma(x;\phi,\beta) \tag{14}$$

$$\begin{aligned} f_X(x) = \phi(x;\mu,\sigma) &+ \frac{1}{r^{1/2}}\left[\frac{\lambda_3}{6\sigma^3} H_3\left(\frac{x-\mu}{\sigma}\right)\right]\phi(x;\mu,\sigma) + \frac{1}{r}\left[\frac{\lambda_4}{24\sigma^4} H_4\left(\frac{x-\mu}{\sigma}\right) + \frac{\lambda_3^2}{72\sigma^6} H_6\left(\frac{x-\mu}{\sigma}\right)\right]\phi(x;\mu,\sigma) \\ &+ \frac{1}{r^{3/2}}\left[\frac{\lambda_5}{120\sigma^5} H_5\left(\frac{x-\mu}{\sigma}\right) + \frac{\lambda_3\lambda_4}{144\sigma^7} H_7\left(\frac{x-\mu}{\sigma}\right) + \frac{\lambda_3^3}{1296\sigma^9} H_9\left(\frac{x-\mu}{\sigma}\right)\right]\phi(x;\mu,\sigma) + \mathrm{O}\left(\frac{1}{r^2}\right) \end{aligned} \tag{15}$$

i.e. the natural logarithm of the CF. The cumulants $\kappa_{X,\nu}, \nu \in \mathbb{Z}_{>0}$ can, if they all exist, be retrieved from

$$\log \Psi_X(t) = \sum_{\nu=1}^{\infty} \kappa_{X,\nu} \frac{(it)^\nu}{\nu!} . \tag{18}$$

We let $\Psi_\phi(t)$ denote the CF of $\phi(x)$ (Bryc, 2012). Using (17) and (18), we can write the CFs of X and the normal kernel as

$$\Psi_X(t) = \exp\left\{ \sum_{\nu=1}^{\infty} \kappa_{X,\nu} \frac{(it)^\nu}{\nu!} \right\} , \tag{19}$$

$$\Psi_\phi(t) = \exp\left\{ \sum_{\nu=1}^{\infty} \kappa_{\phi,\nu} \frac{(it)^\nu}{\nu!} \right\} . \tag{20}$$

These can be combined into

$$\Psi_X(t) = \exp\left\{ \sum_{\nu=1}^{\infty} [\kappa_{X,\nu} - \kappa_{\phi,\nu}] \frac{(it)^\nu}{\nu!} \right\} \Psi_\phi(t). \tag{21}$$

As (Wallace, 1958) notes, it is possible to retrieve the Gram-Charlier series at this point by using the power series expansion of $\exp\{\cdot\}$, sorting the terms by their power of $(-\mathrm{D}_x)$, and applying (1). Instead, Edgeworth assumed that the nearly-normal RV X was a standardized sum

$$X = \frac{1}{\sqrt{r}} \sum_{i=1}^{r} \frac{Z_i - \mu}{\sigma}, \tag{22}$$

where the RVs $Z_1, Z_2, \ldots, Z_r$ are independent and identically distributed (IID), each with mean μ, variance σ^2 and higher order cumulants $\kappa_{Z,\nu} = \sigma^\nu \lambda_\nu$. The dimensionless λ_ν will simplify the following derivation, and the properties of the cumulants gives (Hald, 2000)

$$\kappa_{X,\nu} = \frac{\lambda_\nu}{r^{\frac{\nu}{2}-1}} \ , \ \nu \geq 3 . \tag{23}$$

Since the mean and variance equal the first and second order cumulant, standardized X implies $\kappa_{X,\nu} - \kappa_{\phi,\nu} = 0$ for $\nu = 1, 2$. Unique to the normal distribution is the property that $\kappa_{\phi,\nu} = 0 \ \forall \ \nu \geq 3$ (Hald, 2000), giving

$$\kappa_{X,\nu} - \kappa_{\phi,\nu} = \begin{cases} 0 & ; \quad \nu = 1, 2 , \\ \frac{\lambda_\nu}{r^{\frac{\nu}{2}-1}} & ; \quad \nu \geq 3 . \end{cases} \tag{24}$$

Inserted into (21), this yields

$$\Psi_X(t) = \exp\left\{ \sum_{\nu=3}^{\infty} \frac{\lambda_\nu}{r^{\frac{\nu}{2}-1}} \frac{(it)^\nu}{\nu!} \right\} \Psi_\phi(t). \tag{25}$$

The PDF of X can be retrieved via the inverse FT (Wallace, 1958) as

$$f_X(x) = \exp\left\{ \sum_{\nu=3}^{\infty} \frac{\lambda_\nu}{r^{\frac{\nu}{2}-1}} \frac{(-\mathrm{D})^\nu}{\nu!} \right\} \phi(x). \tag{26}$$

Using the power series representation of the exponential function

$$\exp\{x\} = \sum_{m=0}^{\infty} \frac{x^m}{m!}, \tag{27}$$

we can collect the terms according to their power of $r^{-1/2}$ (instead of $(-\mathrm{D}_x)$) to get the Edgeworth series, presented in (15). It has been found to be superior to the Gram-Charlier series both with few terms and asymptotically (Blinnikov and Moessner, 1998).

2.5 | The Bell Polynomials

Named in honor of Eric T. Bell who introduced them as partition polynomials in (Bell, 1927), the partial exponential Bell polynomials are defined as (Mihoubi, 2008)

$$B_{n,r}(x_1, x_2, \dots, x_{n-r+1}) = \sum_{\Xi_r} n! \prod_{m=1}^{n-r+1} \frac{1}{j_m!} \left(\frac{x_m}{m!}\right)^{j_m}, \tag{28}$$

where the sum is over the set Ξ_r of all combinations of non-negative integers $j_1, \dots, j_n$ which satisfy $j_1 + 2j_2 + \cdots + (n - r + 1)j_{n-r+1} = n - r + 1$ and $r = j_1 + k_2 + \cdots + j_{n-r+1}$. The nth complete exponential Bell polynomial, which we will refer to from here simply as the Bell polynomial, is the sum

$$B_n(x_1, \dots, x_n) = \sum_{r=1}^{n} B_{n,r}(x_1, x_2, \dots, x_{n-r+1}). \tag{29}$$

The first Bell polynomials are

$$B_0 = 1, \tag{30}$$

$$B_1(x_1) = x_1, \tag{31}$$

$$B_2(x_1, x_2) = x_1^2 + x_2, \tag{32}$$

$$B_3(x_1, x_2, x_3) = x_1^3 + 3x_1x_2 + x_3, \tag{33}$$

$$B_4(x_1, \dots, x_4) = x_1^4 + 6x_1^2x_2 + 4x_1x_3 + 3x_2^2 + x_4, \tag{34}$$

$$B_5(x_1, \dots, x_5) = x_1^5 + 10x_1^3x_2 + 15x_1x_2^2 + 10x_1^2x_3 + 10x_3x_2 + 5x_4x_1 + x_5, \tag{35}$$

$$B_6(x_1, \dots, x_6) = x_1^6 + 15x_1^4x_2 + 20x_1^3x_3 + 45x_1^2x_2^2 + 15x_2^3 + 60x_1x_2x_3 + 15x_1^2x_4 + 10x_3^2 + 15x_2x_4 + 6x_1x_5 + x_6. \tag{36}$$

The Bell polynomials satisfy (Mihoubi, 2008)

$$\exp\left\{\sum_{\nu=1}^{\infty} x_\nu \frac{t^\nu}{\nu!}\right\} = \sum_{n=0}^{\infty} B_n(x_1, \dots, x_n)\frac{t^n}{n!}. \tag{37}$$

A well-known use of this result is to retrieve the nth order moment from all cumulants up to order n through (Rota and Shen, 2000)

$$\mu_n = B_n(\kappa_1, \dots, \kappa_n). \tag{38}$$

3 | REMARKS ON THE GRAM-CHARLIER AND EDGEWORTH SERIES EXPANSIONS AROUND THE NORMAL KERNEL

We now present our approach to the derivation of the Gram-Charlier and Edgeworth series. We focus on the similarity of the series by doing as much of the derivation along a single track before splitting in the last step to arrive at the two individual series, clearly emphasizing their differences. Our approach reveals a new and more compact representation of the Edgeworth series.

3.1 | The Double Infinite Sum

We start by pointing out that the moment representation of the CF,

$$\Psi_X(t) = \sum_{\nu=0}^{\infty} \mu_\nu \frac{(it)^\nu}{\nu!}, \tag{39}$$

is not fit to relate $\Psi_X(t)$ and $\Psi_\phi(t)$, since $\Psi_X(t)/\Psi_\phi(t)$ becomes a ratio of sums. By using the cumulant representation instead, we obtain (21), which by use of (27) can be expanded into

$$\Psi_X(t) = \sum_{m=0}^{\infty} \frac{\left[\sum_{\nu=1}^{\infty} [\kappa_{X,\nu} - \kappa_{\phi,\nu}] \frac{(it)^\nu}{\nu!}\right]^m}{m!} \Psi_\phi(t). \tag{40}$$

This reduces the exponential sum into a double sum of polynomials in (it), which will turn into a tractable expression to which we can apply (1). Like in the transition from (25) to (26), the inverse FT can be applied to produce

$$f_X(x) = \sum_{m=0}^{\infty} \frac{\left[\sum_{\nu=1}^{\infty} [\kappa_{X,\nu} - \kappa_{\phi,\nu}] \frac{(-\mathrm{D}_x)^\nu}{\nu!}\right]^m}{m!} \phi(x). \tag{41}$$

3.2 | The Gram-Charlier Series

It is perhaps a bit laborious, but not difficult to sort the terms by their power ν of $(-\mathrm{D}_x)$ and truncate the sum to get the Gram-Charlier series, or by their power of $r^{-1/2}$ to get the Edgeworth series, as pointed out in (Wallace, 1958).

We strongly support the proposal from (Withers and Nadarajah, 2009) to use the Bell polynomials as a simple and concise way of sorting the terms in the Gram-Charlier series. In practice, this amounts to applying (37) to (21), that is

$$\Psi_X(t) = \left[\sum_{n=0}^{\infty} B_n(\kappa_{X,1} - \kappa_{\phi,1}, \ldots, \kappa_{X,n} - \kappa_{\phi,n}) \frac{(it)^n}{n!}\right] \Psi_\phi(t), \tag{42}$$

where it must be stressed that (38) is not in general valid for the cumulant *differences*, i.e.

$$B_n(\kappa_{X,1} - \kappa_{\phi,1}, \ldots, \kappa_{X,n} - \kappa_{\phi,n}) \neq \mu_{X,n} - \mu_{\phi,n}. \tag{43}$$

For X standardized (zero mean, unit variance), $\kappa_{X,1} - \kappa_{\phi,1} = \kappa_{X,2} - \kappa_{\phi,2} = 0$ as before and since $\kappa_{\phi,n} = 0 \; \forall \; n \geq 3$ (Section

2.4), we have

$$\Psi_X(t) = \left[1 + \sum_{n=3}^{\infty} B_n(0, 0, \kappa_{X,3}, \ldots, \kappa_{X,n}) \frac{(it)^n}{n!}\right] \Psi_\phi(t). \tag{44}$$

Eqs. (30) through (36) demonstrate the extreme benefit of this simplification: The Bell polynomials of order zero through six have 30 terms in total, but $x_1 = x_2 = 0$ results in only six non-zero terms. The inverse FT and (3) now gives the Gram-Charlier series explicitly as

$$f_X(x) = \left[1 + \sum_{n=3}^{\infty} \frac{B_n(0, 0, \kappa_{X,3}, \ldots, \kappa_{X,n})}{n!\sigma^n} H_n\left(\frac{x-\mu}{\sigma}\right)\right] \phi(x; \mu, \sigma), \tag{45}$$

from which (13) is easily computed.

3.3 | The Edgeworth Series

In this section, we present a novel approach to deriving the Edgeworth series expansion approximation of the PDF.[2] It is more intuitive than previous derivations, but also reveals a new and simpler expression for the series, using the complete exponential Bell polynomials.

Starting from (25), we can view this as a power series in $r^{-1/2}$ instead of (it) by shifting the counting index $\nu \to \nu + 2$ to get

$$\Psi_X(t) = \exp\left\{\sum_{\nu=1}^{\infty} \frac{\lambda_{\nu+2}}{\left(r^{1/2}\right)^\nu} \frac{(it)^{\nu+2}}{(\nu+2)!}\right\} \Psi_\phi(t) = \exp\left\{\sum_{\nu=1}^{\infty} \tilde{a}_\nu \frac{\left(r^{-1/2}\right)^\nu}{\nu!}\right\} \Psi_\phi(t)\,, \tag{46}$$

where

$$\tilde{a}_\nu = \frac{\lambda_{\nu+2}(it)^{\nu+2}}{(\nu+1)(\nu+2)}. \tag{47}$$

We can again apply (37), since the coefficients $\tilde{a}_\nu$ do not depend on r, giving

$$\Psi_X(t) = \left[\sum_{n=0}^{\infty} B_n(\tilde{a}_1, \ldots, \tilde{a}_n) \frac{\left(r^{-1/2}\right)^n}{n!}\right] \Psi_\phi(t)\,. \tag{51}$$

Since $\tilde{a}_n$ is a power of (it), $B_n(\tilde{a}_1, \ldots, \tilde{a}_n)$ is a polynomial in (it) and we can apply the inverse FT to get the Edgeworth series

$$f_X(x) = \left[1 + \sum_{n=1}^{\infty} B_n(a_1, \ldots, a_n) \frac{\left(r^{-1/2}\right)^n}{n!}\right] \phi(x; \mu, \sigma)\,, \tag{52}$$

[2] In (Withers and Nadarajah, 2009, 2015), the authors applied the *partial* exponential Bell polynomials to the Edgeworth series expansion approximation of the cumulative distribution function and its derivatives of all orders, thus arriving at a very general result. We presently focus only on the PDF, arriving at a different expression.

where $(-\mathrm{D}_x)$ has replaced (it) in a_n compared to $\tilde{a}_n$, that is

$$a_n = \frac{\lambda_{n+2}(-\mathrm{D}_x)^{n+2}}{(n+1)(n+2)} \,. \tag{53}$$

We have included a definition which circumvents the use of coefficients in (48), where we also recovered the cumulants by multiplying all λ's with their corresponding powers of r. Simply using (3) recovers the workable expression (15). We have also reproduced the expressions for the Edgeworth series from (Blinnikov and Moessner, 1998) for comparison (alternative 1),[3] and from (Withers and Nadarajah, 2015) (alternative 2).[4]

4 | A REMARK ON THE GRAM-CHARLIER SERIES EXPANSION AROUND THE GAMMA KERNEL

4.1 | The Gram-Charlier Series With Arbitrary Kernel

Letting $\theta(x)$ be an arbitrary PDF with corresponding CF $\Psi_\theta(t)$. Now, if all its cumulants $\kappa_{\theta,n}$ exist,[5] (42) becomes

$$\Psi_X(t) = \left[\sum_{n=0}^{\infty} B_n(\kappa_{X,1} - \kappa_{\theta,1}, \ldots, \kappa_{X,n} - \kappa_{\theta,n}) \frac{(it)^n}{n!}\right] \Psi_\theta(t) \,. \tag{54}$$

Provided the inverse FT of $(it)^n \Psi_\theta(t)$ is permitted for all $n \in \mathbb{Z}_{\geq 0}$, the PDF $f_X(x)$ can be retrieved as

$$f_X(x) = \left[1 + \sum_{n=1}^{\infty} B_n(\kappa_{X,1} - \kappa_{\theta,1}, \ldots, \kappa_{X,n} - \kappa_{\theta,n}) \frac{(-\mathrm{D}_x)^n}{n!}\right] \theta(x) \,, \tag{55}$$

where $(-\mathrm{D}_x)^n \theta(x)$ must be evaluated in the same way as (1) for $\phi(x)$.

$$\text{Proposed: } f_X(x) = \left[1 + \sum_{n=1}^{\infty} \frac{1}{n!} B_n\left(\frac{\kappa_3(-\mathrm{D}_x)^3}{6}, \frac{\kappa_4(-\mathrm{D}_x)^4}{12}, \ldots, \frac{\kappa_{n+2}(-\mathrm{D}_x)^{n+2}}{(n+1)(n+2)}\right)\right] \phi(x;\mu,\sigma) \tag{48}$$

$$\text{Alternative 1: } \begin{cases} f_X(x) = \left[1 + \sum_{n=1}^{\infty} \sum_{\Xi_n} \frac{1}{\sigma^{n+2r}} H_{n+2r}\left(\frac{x-\mu}{\sigma}\right) \prod_{m=1}^{n} \frac{1}{j_m} \left(\frac{\kappa_{m+2}}{(m+2)!}\right)^{j_m}\right] \phi(x;\mu,\sigma) \\ \Xi_n = \left\{(j_1, \ldots, j_n; r) \,:\, j_m \in \mathbb{Z}_{\geq 0}, \sum_{m=1}^{n} m j_m = n, \; r = \sum_{m=1}^{n} j_m\right\} \end{cases} \tag{49}$$

$$\text{Alternative 2: } \begin{cases} f_X(x) = \left[1 + \sum_{n=1}^{\infty} \frac{r^{-n/2}}{n!} \sum_{k=1}^{n} H_{n+2k}(x) B_{n,k}\left(\eta_{n,k}\right)\right] \phi(x;\mu,\sigma) \\ \eta_{n,k} = \left\{\frac{\kappa_{j+2} r^{j/2}}{\sigma^{j+2}(j+1)(j+2)} \,:\, j = 1, 2, \ldots, n-k+1\right\} \end{cases} \tag{50}$$

---

[3] We have altered the representation slightly compared to (Blinnikov and Moessner, 1998) to account for our nomenclature, allow arbitrary mean and variance, and a slight simplification with respect to σ like the one in (Pastor et al., 2014). Note that the set Ξ_n is the union of the sets Ξ_r for $r \in \{1, 2, \ldots, n\}$ from Section 2.5.

[4] We have also altered this representation with respect to notation. Note that the jth element in the set $\eta_{n,k}$ is the same for any choice of n and k such that $n - k + 1 \geq j$.

[5] If the moments all exist, then so do the cumulants (Sundt et al., 1998).

This result was presented in (Withers and Nadarajah, 2015), where the authors also defined the *generalized* Hermite polynomials as $[\theta(x)]^{-1}(-\mathrm{D}_x)^n\theta(x)$, with the case $\theta(x) = \phi(x)$ reducing to the usual Hermite polynomials, as evident in (1).

4.2 | The Gamma Kernel

A natural question is now whether the approach from Section 3.2 can be applied when the kernel is the gamma PDF. That is, is there an alternative approach which gives the same result as in Section 2.3? Simply put, the answer is no. To retrieve the PDF in (55) we required that the inverse FT can be applied to $(it)^n\Psi_\theta(t)$, but this is not the case with the gamma kernel, as $\mathrm{D}_x^n\gamma(x)$ is discontinuous at $x = 0$ for high enough n.[6]

Fundamentally, this is an example of the limitations of these classical series expansions when it comes to approximating non-negative RVs. The problem comes directly form attempting to apply the FT to a function with support on $(0, \infty)$. We show in (Brenn and Anfinsen, 2017) that better suited methods exist which are based on the Mellin transform.

Here instead, we will provide a simpler formula for the coefficients a_n in (12) by using the confluent hypergeometric function of the first kind $M(a; b; x)$, also known as the Kummer function, after the man who introduced it in (Kummer, 1837). It is defined in (Daalhuis, 2010) as

$$M(a;b;x) = \sum_{k=0}^{\infty} \frac{a(a+1)\cdots(a+k-1)x^k}{b(b+1)\cdots(b+k-1)k!}, \tag{56}$$

and the same reference provides the relationship with the Laguerre polynomials as

$$\begin{aligned} M(-n;\phi+1;x) &= \sum_{k=0}^{n} \frac{n(n-1)\cdots(n-k+1)(-1)^k x^k}{(\phi+1)(\phi+2)\cdots(\phi+k)k!} \\ &= \frac{(\phi+1)(\phi+2)\cdots(\phi+n)}{n!} L_n^{(\phi)}(x). \end{aligned} \tag{57}$$

To our knowledge, $M(\cdot)$ has not been applied to the Gram-Charlier series with gamma kernel, perhaps not surprising given the limited attention the series has received. We now propose a new formula for the coefficients, by using (57) in (12) resulting in

$$a_n = \int_0^\infty f_X(x) M(-n;\phi+1;\beta x)\mathrm{d}x = \mathrm{E}\{M(-n;\phi+1;\beta x)\} \tag{58}$$

$$a_n = \sum_{k=0}^{n} \mu_k \frac{n(n-1)\cdots(n-k+1)(-1)^k \beta^k}{(\phi+1)(\phi+2)\cdots(\phi+k)k!}. \tag{59}$$

In plain text the coefficients are just the expectation of $M(-n;\phi+1;\beta x)$, computed by replacing the powers of x with the (empirical) moments. Choosing ϕ, β such that $a_1 = a_2 = 0$,[7] we get

$$a_n = \frac{(n-1)(n-2)}{2} + \sum_{k=3}^{n} \mu_k \frac{n(n-1)\cdots(n-k+1)(-1)^k \beta^k}{(\phi+1)(\phi+2)\cdots(\phi+k)k!}, \tag{60}$$

[6] Formally, applying the inverse FT to the CF of the gamma distribution, $\Psi_\gamma(t)$, is not permitted because $(it)^n\Psi_\gamma(t)$ is not integrable for high enough n, as $\int_{\mathbb{R}} |(it)^n\Psi_\gamma(t)|\mathrm{d}t < \infty$ is not satisfied.

[7] In practice, this amounts to using the method of moments estimates resulting from solving $\mu_1 = (\phi+1)/\beta, \mu_2 = (\phi+1)(\phi+2)/\beta^2$ for ϕ, β with the first and second order empirical moments replacing μ_1, μ_2.

where the factor $(n-1)(n-2)/2$ is the result of the choices for ϕ, β that satisfy $a_1 = a_2 = 0$. The novel and compact representation of the Gram-Charlier series expansion around the gamma kernel is

$$f_X(x) = \left[1 + \sum_{n=3}^{\infty} \mathrm{E}\{M(-n;\phi+1;\beta x)\} L_n^{(\phi)}(\beta x)\right] \gamma(x;\phi,\beta). \tag{61}$$

5 | CONCLUSION

We have presented new and compact expressions for the Edgeworth series expansions of the normal kernel, and the Gram-Charlier series expansion of the gamma kernel. Compared to the previously available expressions, we have used the complete exponential Bell polynomials in the formulation of the Edgeworth series, with resulting simplifications. The availability of these polynomials in mathematical software means that practically anyone seeking to implement these series expansions will has immediate access to them. Using them also highlights the combinatorial relationships present in the series, which stems from the use of cumulants.

These expression were presented by using a new and more intuitive approach when deriving the Gram-Charlier and Edgeworth series. Our approach highlight the shared foundations of these two series expansions, and conveys in a clear and concise manner the assumptions and mathematical manipulations leading to each of them. This approach also revealed a highly significant simplification in the Edgeworth series, which is by far the most used of the methods discussed in this paper.

Regarding the Gram-Charlier series with gamma kernel, we used the Kummer function to reduce complexity in the computation of the coefficients, thus drastically reducing the effort required to implement the method.

REFERENCES

Bell, E. T. (1927) Partition polynomials. *Annals of Mathematics*, 38–46.

Blinnikov, S. and Moessner, R. (1998) Expansions for nearly Gaussian distributions. *Astronomy and Astrophysics Supplement Series*, **130**, 193–205.

Brenn, T. and Anfinsen, S. N. (2017) A framework for Mellin kind series expansion methods. *Tech. rep.*, UiT The Arctic University of Norway, Department of Physics and Technology. URL: `munin`.

Bryc, W. (2012) *The normal distribution: characterizations with applications*, vol. 100. Springer Science & Business Media.

Charlier, C. V. L. (1905) Über das fehlergesetz. *Ark. Mat. Astr. Och Fys.*, **Vol. 2 (1905-1906)**, 9.

— (1906) Über die darstellung willkürlicher funktionen. *Ark. Mat. Astr. Och Fys.*, **Vol. 2 (1905-1906)**, 35.

Chebyshev, P. L. (1860) Sur le developpement des fonctions à une seule variable. *Bull. Acad. Imp. Sci. St. Petersbourg*, **Series 3, Vol. 1**, 193–202.

— (1890) Sur deux théorèmes relatifs aux probabilités. *Acta Math.*, **14**, 305–315.

Daalhuis, A. O. (2010) *Confluent hypergeometric functions.* NIST Handbook of Mathematical Functions, p. 321–349. Cambridge University Press.

Edgeworth, F. Y. (1905) The law of error. *Cambridge Philos. Trans.*, **20**, 36–66, 113–141.

Fowler, D. (1996) The binomial coefficient function. *The American Mathematical Monthly*, **103**, 1–17.

Gaztanaga, E., Fosalba, P. and Elizalde, E. (2000) Gravitational evolution of the large-scale probability density distribution: The Edgeworth and gamma expansions. *The Astrophysical Journal*, **539**, 522.

Gram, J. P. (1883) Über die entwickelung reeler funktionen in reihen mittelst der methode der kleinsten quadrate. *J. Reine Angew. Math.*, **94**, 41–73.

Hald, A. (2000) The early history of the cumulants and the Gram-Charlier series. *International Statistical Review*, **68**, 137–153.

Kendall, M., Stuart, A. and Ord, J. (1994) *Kendall's Advanced Theory of Statistics*, vol. 1: Distribution theory. John Wiley & Sons, Ltd, 6th edn.

Kummer, E. E. (1837) De integralibus quibusdam definitis et seriebus infinitis. *Journal für die reine und angewandte Mathematik*, **17**, 228–242.

Lévy, P. (1925) *Calcul des probabilités*. Paris: Gauthier-Villars.

Lukacs, E. (1970) *Characteristic functions*. Griffin.

Mihoubi, M. (2008) Bell polynomials and binomial type sequences. *Discrete Mathematics*, **308**, 2450–2459.

Pastor, G., Mora-Jiménez, I., Caamano, A. J. and Jäntti, R. (2014) Log-cumulants-based Edgeworth expansion for skew-distributed aggregate interference. *2014 11th International Symposium on Wireless Communications Systems (ISWCS)*, *IEEE*, 390–394.

Pitman, J. (2002) Combinatorial stochastic processes. Technical Report 621, Dept. Statistics, UC Berkeley, 2002. Lecture notes for St. Flour course.

Rota, G.-C. and Shen, J. (2000) On the combinatorics of cumulants. *Journal of Combinatorial Theory, Series A*, **91**, 283–304.

Sundt, B., Dhaene, J. and De Pril, N. (1998) Some results on moments and cumulants. *Scandinavian Actuarial Journal*, **1998**, 24–40.

Szeg, G. (1939) *Orthogonal polynomials*, vol. 23. American Mathematical Soc.

Thiele, T. (1889) *Almindelig Iagttagelseslære: Sandsynlighedsregning og mindste Kvadraters Methode*. Kjøbenhavn: Reitzel.

— (1903) *Theory of Observations*. London: Layton.

Wallace, D. L. (1958) Asymptotic approximations to distributions. *The Annals of Mathematical Statistics*, **29**, 635–654.

Withers, C. S. and Nadarajah, S. (2009) Charlier and Edgeworth expansions for distributions and densities in terms of Bell polynomials. *Probability and Mathematical Statistics*, **29**, 271.

— (2015) Edgeworth–Cornish–Fisher–Hill–Davis expansions for normal and non-normal limits via Bell polynomials. *Stochastics: An International Journal of Probability and Stochastic Processes*, **87**, 794–805.