Journal of Statistical Planning and Inference 167 (2015) 174–181

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

journal homepage: www.elsevier.com/locate/jspi

Expansions for multivariate densities

Ruby C. Weng

Department of Statistics, National Chengchi University, Taipei, Taiwan

ARTICLE INFO

Article history:
Received 23 April 2015
Received in revised form 4 May 2015
Accepted 5 May 2015
Available online 19 May 2015

Keywords:
Gram–Charlier series
Edgeworth series
Hermite polynomials
Woodroofe–Stein's identity

ABSTRACT

The Gram–Charlier and Edgeworth series are expansions of probability distribution in terms of its cumulants. The expansions for the multivariate case have not been fully explored. This paper aims to develop the multivariate Gram–Charlier series by Woodroofe–Stein's identity, and improve its approximation property by using the scaled normal density and Hermite polynomials. The series are useful to reconstruct the probability distribution from measurable higher moments.

1. Introduction

There exist several types of expansions of a univariate probability density function (pdf) in the orthogonal sets of Hermite polynomials. The expansions express the probability distribution in terms of its moments or cumulants. Three well-known expansions are the Gram–Charlier, Gauss–Hermite, and Edgeworth series. The first two series differ from each other in using different sets of Hermite polynomials, while the Edgeworth series differs from the first two series in that it collects terms of the same order. So, the Edgeworth series is an asymptotic expansion, whereas the other two are not.

It is known that the Gram–Charlier series may diverge in many cases of interest, and that even the asymptotic Edgeworth series may not converge. For the one-dimensional case, if the pdf $p(z)$ is of bounded variation in $(-\infty, \infty)$ and the integral $\int_{-\infty}^{\infty} p(z)\exp(z^2/4)dz$ exists, the Gram–Charlier series converges; on the other hand, if the conditions are not satisfied, the expansions may diverge (Cramér, 1946, Chapter 17). Blinnikov and Moessner (1998) gave a comparison of the Gram–Charlier, Gauss–Hermite, and Edgeworth series. They showed that the Edgeworth expansion is the best among them; however, for strongly non-Gaussian cases, like χ^2_ν with degrees of freedom $\nu = 2$, the Edgeworth series also diverges like the Gram–Charlier series.

Despite divergence in some cases, in many practical applications what we really concern is whether a small number of terms suffice to give a good approximation (Cramér, 1946). In fact, these expansions are useful to measure the deviations of a pdf from the normal pdf, to provide correction terms for density approximation, and to reconstruct a pdf by measurable higher order moments. They have been used in a variety of areas. For instance, Sargan (1975, 1976) introduced it into econometrics; Van Der Marel and Franx (1993), Scherrer and Bertschinger (1991), and Blinnikov and Moessner (1998) applied the Gram–Charlier, Gauss–Hermite, and Edgeworth expansions in astrophysics; Comon (1994) and Amari et al. (1996) used the Edgeworth series in Independent Component Analysis to approximate the one-dimensional differential entropy; Hall (1992) showed how Edgeworth expansion and bootstrap methods can help explain each other; Rubinstein (1998) employed the Edgeworth expansion to value derivatives, among others.

Though there are many applications based on the expansion for the one-dimensional probability density, relatively few studies are available for the multivariate densities. To the best of our knowledge, the most comprehensive account of the

E-mail address: chweng@nccu.edu.tw.

http://dx.doi.org/10.1016/j.jspi.2015.05.001
0378-3758/

expansions for the multivariate case is by Barndorff-Nielsen and Cox (1989). Basically, they take the traditional approach which inverts expansions of the characteristic functions by the inverse Fourier transform. Other studies for the multivariate cases include Csörgő and Hegyi (2000), Van Hulle (2005), among others.

Instead of inverting the characteristic function, Weng (2010) derived the Gram–Charlier type and Edgeworth type expansions of a univariate pdf based on a version of Stein's identity. The identity is essentially repeated integration by parts, which naturally leads to a series expansion. This version of Stein's identity was developed by Woodroofe (1989, 1992a) for obtaining integrable expansions for posterior distributions. It is closely related to the well-known Stein's lemma (Stein, 1981) — the latter considers the expectation with respect to a normal distribution, while the former the expectation with respect to a "nearly normal distribution" Γ in the sense of (7); and both are proved by an application of Fubini theorem. Stein's lemma (Stein, 1981) is famous for its applications to James–Stein estimator (James and Stein, 1961) and empirical Bayes methods. As for Stein's identity, it has been applied to set frequentist corrected confidence sets following sequentially designed experiments (e.g. Woodroofe, 1992b; Coad and Woodroofe, 1996; Weng and Woodroofe, 2006, among others), and to Bayesian inference (e.g. Weng, 2003; Baghishania and Mohammadzadeh, 2012, among others). Recently Weng and Lin (2011) applied this identity to derive Bayesian online algorithms for ranking players; and to distinguish it from the well-known Stein's lemma, they coined it as *Woodroofe–Stein*'s identity. The identity may be further explored.

The main contribution of present paper is a theoretical advancement in generalizing Weng's (2010) approach to give a closed-form expression for the multivariate Gram–Charlier expansion. Furthermore, as the Gram–Charlier series suffers from poor convergence, a modified series is proposed for better convergence properties. The modification starts by suitably scaling the variable and applying the expansion to pdf of the scaled variable, and then converted back to the original variable.

The remaining of this paper is organized as the following. The next section provides some reviews. Section 3 presents the Gram–Charlier type expansion for multivariate densities and describes the proposed modified series. Section 4 concludes.

2. Reviews

2.1. Woodroofe–Stein's identity

We review the identity. Some results here will be generalized in Section 3.1. Some definitions and notations are needed. Let ϕ_p and Φ_p denote the density and distribution function of a standard p-variate normal distribution, and abbreviate Φ_1 and ϕ_1 as Φ and ϕ. For a function $h : R^p \to R$, we may write

$$\Phi_p h \equiv \int h d\Phi_p \tag{1}$$

for simplicity, provided the integral is finite. Let $h_0 = \Phi_p h$ be a constant, $h_p(\boldsymbol{z}) = h(\boldsymbol{z})$,

$$h_j(z_1, \ldots, z_j) = \int_{R^{p-j}} h(z_1, \ldots, z_j, \boldsymbol{w}) d\Phi_{p-j}(\boldsymbol{w}), \quad \text{and} \tag{2}$$

$$g_j(z_1, \ldots, z_p) = e^{z_j^2/2} \int_{z_j}^{\infty} [h_j(z_1, \ldots, z_{j-1}, w) - h_{j-1}(z_1, \ldots, z_{j-1})] e^{-w^2/2} dw, \tag{3}$$

for $-\infty < z_1, \ldots, z_p < \infty$ and $j = 1, \ldots, p$.

Lemma 2.1. *For a function $h : R^p \to R$, let h_j and g_j be as in (2) and (3). If $h(\boldsymbol{z})$ depends on $\boldsymbol{z}$ only through $z_1, \ldots, z_i$, then we have $h_j(z_1, \ldots, z_j) = h(\boldsymbol{z})$ for all $j \geq i$; and consequently, $g_j(\boldsymbol{z}) \equiv 0$ for all $j > i$.*

The proof of the lemma is straightforward from (2) and (3), so we omit it. Now we shall define operators $U(h)$ and $U^2(h)$ associated with h. Let $U(h)$ denote the vector of the functions g_j in (3),

$$U(h) = [g_1, \ldots, g_p]^T. \tag{4}$$

For example, for $\boldsymbol{z} \in \Re^p$, if $h(\boldsymbol{z}) = z_1$, then $Uh(\boldsymbol{z}) = (1, 0, \ldots, 0)^T$ and if $h(\boldsymbol{z}) = \|\boldsymbol{z}\|^2$, then $Uh(\boldsymbol{z}) = \boldsymbol{z}$. Further, let U^2 denote the composition of U with itself:

$$U^2(h) = U(U(h)) = [U(g_1), \ldots, U(g_p)]^T, \tag{5}$$

which is a $p \times p$ matrix whose jth row is $U(g_j)$ and g_j is as in (3). Since $g_j(\boldsymbol{z})$ defined in (3) depends only on $(z_1, \ldots, z_j)$, by Lemma 2.1 we have that $U^2(h)$ is a lower triangular matrix. Next, define

$$V(h) = \frac{U^2(h) + (U^2(h))^T}{2} = \frac{1}{2}\{[U(g_1), \ldots, U(g_p)]^T + [U(g_1), \ldots, U(g_p)]\}. \tag{6}$$

Then, $V(h)$ is a symmetric matrix.

A function $f : R^p \to R$ is said to be almost differentiable if there exists a function $\nabla f : R^p \to R^p$ such that

$$f(\boldsymbol{z} + \boldsymbol{x}) - f(\boldsymbol{z}) = \int_0^1 \boldsymbol{x}^T \nabla f(\boldsymbol{z} + t\boldsymbol{x}) dt$$