Available online at www.sciencedirect.com
ScienceDirect
Journal of Statistical Planning and Inference 138 (2008) 1167 – 1178
journal of statistical planning and inference
www.elsevier.com/locate/jspi

Shortening the distance between Edgeworth and Berry–Esseen in the classical case

José A. Adell*, Alberto Lekuona

Departamento de Métodos Estadísticos, Facultad de Ciencias, Universidad de Zaragoza, 50009 Zaragoza, Spain

Received 14 March 2006; received in revised form 2 February 2007; accepted 18 April 2007
Available online 25 May 2007

Abstract

We obtain near optimal Berry–Esseen bounds for standardized sums of independent identically distributed random variables. This is achieved by distinguishing the lattice and the non-lattice cases, as one-term Edgeworth expansions do. The main tool is an easy inequality involving the usual second modulus of continuity, in substitution of Esseen's smoothing inequality. An illustrative example concerning the exponential distribution is also considered.

MSC: 60F05

Keywords: Berry–Esseen bound; Edgeworth expansion; Moduli of smoothness

1. Introduction

Let us start by introducing some notation. Let X be a random variable with distribution function F, characteristic function φ, and such that $EX=0$, $EX^2=\sigma^2>0$, $EX^3=v$ and $E|X|^3=\beta<\infty$. If F is lattice, we denote by $s>0$ the maximal span of F and assume that the discontinuities of F are located at x_0+ks, $k=0,\pm 1,\pm 2,\ldots$, for some $x_0\in\mathbb{R}$. In such a case, we set $D(F)=\sup\{P(X=x):x\in\mathbb{R}\}$ for the maximum jump of F. Let $(X_k)_{k\geqslant 1}$ be a sequence of independent identically distributed random variables with common distribution function F. For any $n=1,2,\ldots$, denote by S_n the standardized sum $S_n=(X_1+\cdots+X_n)/(\sigma\sqrt{n})$ and by F_n the distribution function of S_n. Finally, let Z be a random variable having the standard normal distribution function Φ.

Usually, rates of convergence in the central limit theorem are described by means of Edgeworth expansions, Berry–Essen bounds or "leading terms" (cf. Hall, 1982; Hall and Wang, 2004). In the case at hand, and depending on whether or not F is lattice, we have the following one-term Edgeworth expansions for $F_n(x)$ as n tends to infinity (see, for instance, Esseen, 1945, pp. 49–56; Petrov, 1975, pp. 159–171)

$$F_n(x)-\Phi(x)=\frac{s}{\sigma\sqrt{2\pi}}R\left(\frac{(x-\zeta_n)\sigma\sqrt{n}}{s}\right)\mathrm{e}^{-x^2/2}n^{-1/2}+\frac{v}{6\sigma^3\sqrt{2\pi}}(1-x^2)\mathrm{e}^{-x^2/2}n^{-1/2}+\mathrm{o}(n^{-1/2}), \tag{1}$$

* Corresponding author.
E-mail addresses: adell@unizar.es (J.A. Adell), lekuona@unizar.es (A. Lekuona).

0378-3758/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.jspi.2007.05.032

or

$$F_n(x)-\Phi(x)=\frac{v}{6\sigma^3\sqrt{2\pi}}(1-x^2)\mathrm{e}^{-x^2/2}n^{-1/2}+\mathrm{o}(n^{-1/2}), \tag{2}$$

respectively, where $R(x)=\lfloor x\rfloor - x+1/2$ is the "rounding error" function, $\lfloor x\rfloor$ stands for the integer part of x, and

$$\zeta_n=\frac{s}{\sigma\sqrt{n}}\left(\frac{nx_0}{s}-\left\lfloor\frac{nx_0}{s}\right\rfloor\right)$$

is the least non-negative discontinuity point of $F_n(x)$. On the other hand, Berry–Esseen bounds take the usual form

$$\|F_n-\Phi\|\leqslant C\frac{\beta}{\sigma^3}n^{-1/2},\quad n=1,2,\ldots, \tag{3}$$

for some absolute positive constant C, where $\|\cdot\|$ means the supremum norm in $\mathbb{R}$. As far as we know, the smallest known constant fulfilling (3) is $C=0.7655$. This was shown by Shiganov (1986), improving an earlier constant found by van Beck (1972) (see also Zahl (1966) for a slightly different form of (3)). The proofs concerning such constants are based on Esseen's smoothing inequality and certain numerical computations. Most of these results actually hold with appropriate modifications for standardized sums of independent, not necessarily identically distributed, random variables. In the later context, Bentkus (1994) has considered the asymptotical behavior of the constant in the Berry–Esseen inequality, and Chen and Shao (2001) have obtained Berry–Esseen bounds with explicit upper constants under finite variance, using Stein's method in combination with concentration inequalities.

A comparison between one-term Edgeworth expansions and Berry–Esseen bounds reveals the following. Certainly, formulae (1) and (2) provide exact expressions for the main terms of the approximation for any $x\in\mathbb{R}$. It is also true that one-term Edgeworth expansions provide better accuracy than Berry–Esseen bounds even for moderate sample sizes, and this is confirmed by numerical computations (see, for instance, Brown et al., 2002; Zhou et al., 2004). However, with a very few exceptions (cf. Seoh and Hallin, 1997), Edgeworth expansions contain "big or little o" terms with unspecified upper bounds. This means that the question about how large must the sample size n be in order to guarantee some prescribed degree of accuracy in the approximation cannot be theoretically answered (in this respect, see the interesting discussion in Seoh and Hallin, 1997; Pfanzagl, 2000). Berry–Esseen bounds do answer this question, but the sharpness of C in (3) is limited by the intrinsic form of the Berry–Esseen inequality itself. In fact, according to Esseen (1956), such a constant C cannot be less than $(3+\sqrt{10})/(6\sqrt{2\pi})=0.4097\ldots$, because one can find a lattice distribution F for which the preceding constant is asymptotically attained. Nevertheless, if we look at (2), we see that in the non-lattice case the best asymptotic constant should be $1/(6\sqrt{2\pi})=0.06649\ldots$. Such problems could be remedied by reformulating the Berry–Esseen inequality in the sense of distinguishing the lattice and the non-lattice cases, as in (1) and (2).

The main purpose of this work is to approach the points of view coming from one-term Edgeworth expansions and Berry–Esseen bounds by showing the following. Depending on whether or not F is lattice, we prove for any $n=1,2,\ldots,$ that

$$\frac{1}{2}D(F_n)\leqslant\|F_n-\Phi\|\leqslant\frac{1}{2}D(F_n)+\frac{\beta}{6\sigma^3\sqrt{2\pi}}(1+\delta)n^{-1/2}+R(n), \tag{4}$$

or

$$\|F_n-\Phi\|\leqslant\frac{\beta}{6\sigma^3\sqrt{2\pi}}(1+\delta)n^{-1/2}+R(n), \tag{5}$$

respectively. In (4) and (5), δ is a positive parameter as close as we wish to zero and the remainders $R(n)$ are explicitly bounded, the upper bounds depending on some parameters arbitrarily chosen in certain ranges. No numerical computations are needed. It turns out that $R(n)=\mathrm{O}(n^{-1})$ as $n\to\infty$. Observe that, by the local limit theorem (cf. Petrov, 1975, p. 187), the term $D(F_n)/2$ behaves asymptotically as $sn^{-1/2}/(2\sigma\sqrt{2\pi})$. Therefore, in view of (1) and (2), the estimates in (4) and (5) are near optimal. There is a price to be paid, however. Indeed, if for a fixed n, we let $\delta\to 0$ in (4) and (5), the remainder terms could become just bounded or even tend to infinity (see Section 3 for more precise statements). In practical applications, the right thing to do is to choose the parameters depending on the sample size n. This is illustrated in the final section by means of a simple example involving the exponential distribution.