The Annals of Statistics
2003, Vol. 31, No. 4, 1376–1391

EDGEWORTH EXPANSION FOR U-STATISTICS UNDER MINIMAL CONDITIONS

BY BING-YI JING AND QIYING WANG

HKUST and Australian National University

Berry–Esseen bounds for U-statistics under the optimal moment conditions were derived by Koroljuk and Borovskich and Friedrich. Under the same optimal moment assumptions with an additional nonlattice condition, we establish a one-term Edgeworth expansion with remainder $o(n^{-1/2})$ for U-statistics.

1. Introduction and main results. Let $X_1, X_2, \dots, X_n, n \geq 2$, be independent and identically distributed (i.i.d.) random variables with common distribution function $F(x)$. Let $h(x, y)$ be a real-valued Borel measurable function, symmetric in its arguments with $Eh(X_1, X_2) = \theta$. Define a U-statistic by

$$U_n = \frac{2}{n(n-1)} \sum_{1 \leq i < j \leq n} h(X_i, X_j)$$

and

$$g(X_1) = E[h(X_1, X_2) \mid X_1] - \theta, \qquad \sigma_g^2 = \operatorname{Var}\bigl(g(X_1)\bigr).$$

Throughout this paper, we shall assume that $\sigma_g^2 > 0$. Our primary goal is to investigate the asymptotic distribution of the standardized U-statistic defined by

$$G_n(x) = P\biggl(\frac{\sqrt{n}(U_n - \theta)}{2\sigma_g} \leq x\biggr).$$

It is well known that $G_n(x)$ converges to the standard normal distribution function, $\Phi(x)$, provided $Eh^2(X_1, X_2) < \infty$ [see Hoeffding (1948)]. In fact, this moment condition can further be reduced to $Eg^2(X_1) < \infty$ and $E|h(X_1, X_2)|^{4/3} < \infty$; see Remark 4.2.4 of Koroljuk and Borovskich [(1994), page 131].

In recent years, there has been considerable interest in obtaining rates of convergence in the asymptotic normality for U-statistics, for instance, by Grams and Serfling (1973), Bickel (1974) and Chan and Wierman (1977). A sharper Berry–Esseen bound was given by Callaert and Janssen (1978), which states that

$$\sup_{t \in R} |G_n(x) - \Phi(x)| \leq A\sigma_g^{-3} E|h(X_1, X_2)|^3 n^{-1/2}$$

Received May 2001; revised July 2002.
AMS 2000 subject classifications. 62E20, 60F15.
Key words and phrases. U-statistics, Edgeworth expansion, optimal moments, nonlattice condition.