Stein, Charles

Approximate computation of expectations. (English) Zbl 0721.60016

Institute of Mathematical Statistics Lecture Notes - Monograph Series 7. Hayward, CA: Institute of Mathematical Statistics (ISBN 0-940600-08-0). iv, 164 p., open access (1986).

The author presents a novel approach to the proof of the central limit theorem. This approach is described below in its simplest setting.

Let E denote a probability distribution on the Borel sets of $\mathbb{R}$; let $\mathcal{B}$ denote the space of all bounded measurable functions; and write $Eh = \int hdE$ for $h \in \mathcal{B}$. Then $\ker(E) = \{h : Eh = 0\}$. The approach requires finding a linear space $\mathcal{A}$ and a nontrivial linear transformation T: $\mathcal{A} \to \ker(E)$ for which $E \circ T = 0$, where $\circ$ denotes composition. For the special case in which E is the standard normal distribution, say $E = E_0$, convenient choices are to let $\mathcal{A}_0$ be the collection of all absolutely continuous f for which $f'(z)$ and zf(z) are bounded and to let $T_0 f(z) = f'(z) - zf(z)$ for $z \in \mathbb{R}$ and $f \in \mathcal{A}_0$. In fact, E_0 is the only probability distribution for which $E \circ T_0 = 0$. In this case, T_0 has an inverse transformation $U_0 : \ker(E_0) \to \mathcal{A}$ defined by $U_0 k(z) = [1/\phi(z)] \int_{-\infty}^{z} k(y)\phi(y)dy$ for $z \in \mathbb{R}$ and $k \in \ker(E_0)$, where ϕ denotes the standard normal density. Now let E_0 denote the standard normal distribution and let E be another distribution on $\mathbb{R}$ (e.g., the distribution of the standardized sum of independent random variables). If $h \in \mathcal{B}$, then $k = h - E_0 h \in \ker(E_0)$, so that $g = U_0 k \in \mathcal{A}_0$. Then $E \circ (T_0 - T)g = Ek - E \circ Tk = Ek = Eh - E_0 h$, since $E \circ T = 0$; and, therefore,

$$(*) \quad Eh - E_0 h = E \circ (T_0 - T) \circ U_0 k.$$

So, if T may be made to approximate T_0, then E_0 approximates E. This is the basic approach.

For the case in which E is the distribution of the sum S of independent random variables $X_1, ..., X_n$ with common mean 0 and variances with unit sum, a convenient choice of T may be constructed by a clever argument which involves replacing one of the X_i, selected at random; and T may be shown to approximate T_0 under the conditions for the central limit theorem. Moreover, the approach may be refined to handle Poisson limits, large deviations, local limit theorem, the Berry- Esseen theorem, and central limit theorems for sums of weakly dependent random variables. The basic approach is developed three times with increasing generality. The first two times are followed by applications to specific combinatorial problems which serve to illustrate the techniques. These include the number of ones in the binary expansion of a random integer, Latin rectangles, random allocations, and the number of isolated trees in a random graph. Most proofs are presented in sufficient detail that the book is readable by graduate students who have completed a course in (measure-theoretic) probability. In fact, when specialized to the discrete case (and perhaps stripped of some function-analytic language), the simplest form of the approach could be presented to advanced undergraduates.

{Reviewer's remarks: I find the author's approach interesting for its novelty and simplicity. It is unclear to me whether the approach will lead to important new discoveries about the distributions of sums of random variables. Early results on error bounds for the distributions of sums of dependent random variables (cited in the book) are encouraging, but more are required.}

MSC:

60F05 Central limit and other weak theorems
60-02 Research exposition (monographs, survey articles) pertaining to probability theory

Keywords:

Stein's method; central limit theorem; sums of weakly dependent random variables

Full Text: Euclid