SADDLEPOINT APPROXIMATIONS IN THE CASE OF INTRACTABLE CUMULANT GENERATING FUNCTIONS

John E. Kolassa
IBM T. J. Watson Research Center

Abstract

Saddlepoint Approximations have long been used to approximate densities and distribution functions of random variables with known cumulant generating function defined on an open interval about the origin. This approximation has very desirable asymptotic properties when approximating densities and tail probabilities for sums of random variables, and also often performs remarkably well for small sample sizes, including samples of one.

Calculating the saddlepoint approximation requires calculating the Legendre transform of the log of the cumulant generating function. In some cases this cumulant generating function may be unavailable; in other cases the Legendre transform is difficult to calculate analytically. This paper discusses modifications to the saddlepoint approximation necessary when the cumulant generating function is replaced by a similar but more tractable function whose Legendre transform can be given explicitly. Calculations for the logistic distribution are presented to illustrate the case of a known but intractable cumulant generating function, and an example involving an overdispersed binomial model is presented to illustrate the case of an unavailable cumulant generating function. An application to a random effects logistic linear model is discussed.

1. The Problem. Consider the following problem: $X_1, X_2, X_3, \ldots, X_n, \ldots$ are independent and identically distributed random variables. Assume that their common distribution is continuous, with probability density function $f_1(x)$ and cumulant generating function $K(t) = \log E[e^{tX}]$, defined on some convex set $I \subset \boldsymbol{R}$. Without loss of generality assume further that $E[X_i] = 0$. I wish to approximate the density function $f_n(s)$ and distribution function $F_n(s)$ of

$$S = \frac{X_1 + \ldots + X_n}{\sqrt{n}}$$

analytically, and am faced with three options:

a. Convolute the density n times.
b. Use the classical Edgeworth series approximation to $f_n(x)$.