Near-exact distributions – what are they and why do we need them?

Carlos A. Coelho

Departamento de Matemática and Centro de Matemática e Aplicações,
Faculdade de Ciências e Tecnologia – Universidade Nova de Lisboa, Portugal
cmac@fct.unl.pt

Before answering the question set forth in the title, let us start by answering a related but simpler question which is: When may we need to use a near-exact distribution? The answer to this question might well be used as the subtitle for this talk, since it is: "when exact distributions are too complicated and common asymptotic distributions do not perform well"! Indeed, namely in multivariate analysis, most of the commonly used asymptotic distributions worsen their performance when the number of variables increase and even many of them are no longer proper distributions when the number of variables goes above a given threshold. These are facts that have been completely overlooked by most of the authors and these awkward behaviors of many of the common asymptotic distributions used in multivariate analysis are not easy to overcome, when we use the common asymptotic techniques. However, by using a different approach, which combines an adequate decomposition of the characteristic function of the statistic under study, most often a factorization, with the action of keeping then most of this characteristic function unchanged, and replacing the remaining smaller part by an adequate asymptotic approximation, it is possible to build manageable approximations, called 'near-exact' approximations, which yield distributions extremely close to the exact distribution, and which exhibit a very good performance for very small sample sizes and an asymptotic behavior not only for an increasing sample sizes but also for increasing number of variables involved. These near-exact distributions may then be applied to obtain very well-fitting near-exact quantiles and p-values and they have been, so far, successfully applied to a large number of statistics. Furthermore, near-exact distributions may be easily developed even for test statistics of highly complicated hypotheses, by considering the decomposition of the null hypothesis into a set of conditionally independent hypotheses. Examples will be given.

Key Words: characteristic function, factorization of characteristic functions, near-exact $p-$values, near-exact quantiles