Chi-square tests under models close to the normal distribution

A. García-Pérez

Abstract Many test statistics follow a $\chi^2$ distribution because a normal model is assumed as underlying distribution. In this paper we obtain good analytic approximations for the $p$-value and the critical value of $\chi^2$ tests when the underlying distribution is close but different from the normal model. With these approximations we study the robustness of validity of $\chi^2$ tests.

Keywords Robustness in hypotheses testing · von Mises expansion · tail area influence function · saddlepoint approximation · robustness of $\chi^2$ tests

1 Introduction

Many classical parametric tests were obtained assuming a normal distribution as underlying model. This is the reason why the $\chi^2$, Student’s $t$-, and $F$-distributions play a prominent role in Statistics as distributions for test statistics.

Nevertheless a complete and satisfactory understanding of the behaviour of these kind of tests when the underlying distribution is not normal but a slight deviation from it has not yet been reached (see for instance Benjamini 1983; Box 1953 and Rivest 1986).

In this paper we obtain good analytic approximations of the key elements of a $\chi^2$ test, such as the critical value and the $p$-value, when the underlying normality holds only approximately, developing a method proposed in García-Pérez (2003), which is based on considering all these elements as functionals of the model distribution, and that makes use of the von Mises (VOM) expansion of a functional plus, in some cases, saddlepoint approximations.
This method is specially useful in robustness studies where the model distribution is, frequently, a slight deviation from the normal distribution (for instance, a contaminated normal) but complicated enough to render impossible an exact calculation of these elements.

With the VOM approximations we propose in section 2, we obtain very easily that, for example, with \( n = 3 \), the actual level of the classical \( 0.05 \) one-sided variance test with known mean, is approximately \( 0.121 \) when the model distribution is a logistic distribution instead of the normal, or that the \( 0.01 \) critical value is \( 20.03 \) instead of \( 11.34 \) in this situation.

Also, with the VOM+SAD approximations of section 3, proved in the Appendix, and a real dataset we obtain, in a straightforward manner, the lack of robustness of validity of a decision based on the Bartlett’s test.

The robustness results we obtain are in line with those obtained by Box (1953) for large samples and by Rivest (1986) for small samples, in the sense that \( \chi^2 \) tests do not have robustness of validity even for small departures from a Gaussian population.

1.1 Preliminaries

Although the method that we are going to explain in this paper can be extended to a more general setting, we will consider in it a one-dimensional test based on a test statistic \( T_n = T_n(X_1, \ldots, X_n) \) that rejects the null hypothesis \( H_0 \) when \( T_n \) is larger than the critical value \( k_F^n \) and where \( F \) is the distribution that the \( X_i \)'s follow under \( H_0 \). If \( T_n = t \), the p-value will be then the tail probability \( p_F^n = P_F\{T_n > t\} \).

In particular, we will consider \( \chi^2 \) tests in the paper, i.e., tests where \( T_n \) follows a \( \chi^2 \) distribution under a normal model, studying here its behaviour under a model \( F \), close but different, from the normal.

We will suppose that \( T_n \) is real valued although the sample \( X_1, \ldots, X_n \) can be one- or multi-dimensional. The only restriction is that, under the null hypothesis, both the critical value \( k_F^n \) and the p-value \( p_F^n \) must be functionals of only one distribution function \( F \) that we will assume univariate. For instance, in a one-dimensional parametric test of the null hypothesis \( H_0 : \theta = \theta_0 \), if \( X_1, \ldots, X_n \) is a sample from a random variable \( X \) with distribution function \( F_{\theta} \) and \( F_{n;\theta} \) is the cumulative distribution function of the test statistic \( T_n \), the critical value of the level-\( \alpha \) test \( k_F^n = F_{n;\theta_0}^{-1}(1 - \alpha) \) and the p-value, \( p_F^n = P_{F_{\theta_0}}\{T_n > t\} \), will be considered functionals of \( F_{\theta_0} \) (throughout the paper, the inverse of any distribution function \( G \) is defined, as usual, by \( G^{-1}(s) = \inf\{y|G(y) \geq s\} \), \( 0 < s < 1 \)).

2 von Mises approximations for \( \chi^2 \) tests

Let \( T_n \) be a test statistic that follows a \( \chi^2_n \) distribution when the underlying model is a normal distribution, \( \Phi_{\mu, \sigma} \), and the null hypothesis holds. Under a model \( F \) close to \( \Phi_{\mu, \sigma} \) we have the next result.