

Title: Normality: Who Needs It?
By Lynn R. LaMotte, Biostatistics Program

Abstract: The Ideal Sampling Model entails three conditions: independence, equal variances, and normality. Under these conditions, p-values for tests of hypotheses are exact. In every real-world application, though, the response does not follow a normal distribution, and so p-values are always approximations. P-values are probabilities from distributions of test statistics, not from distributions of the response itself. Thus examining the distribution of the response or residuals doesn't tell us anything directly about the accuracy of the approximation.

In a landmark paper, E. S. Pearson and N. W. Pleuse¹ examined the accuracy of normal-based p-value approximations in different settings. They concluded that “when the two sample sizes are equal, the two-sample t test seems in every case a safe one to use.” They briefly considered the efficacy of preliminary tests for skewness and normality, but found that “The gain was throughout minimal.”

I'll review these results and illustrate the robustness of normal-based p-values by simulation.

¹*Biometrika*(1972) 62(2): 223–241, readily available through JSTOR, <http://www.jstor.org/browse#Statistics>.