Comparisons between the Binomial and CLT tests

For symmetric distributions the mean and median are equal, so we can compare tests for the mean and median to each other in this case. We will consider testing the null hypothesis $H_0 : \mu = \mu_0$ versus the alternative hypothesis $H_a : \mu > \mu_0$ for the binomial test (for the median) and a normal-theory-based parametric test (for the mean). Suppose that $X_1, X_2, ..., X_n$ are a random sample from a population with mean μ and standard deviation σ. When σ is known, the statistic

$$Z_\mu = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}}$$

has a standard normal distribution in large samples by the central limit theorem, and hence the test based on it is called the CLT test. If σ is unknown, then we can replace σ with the sample standard deviation S, and for samples from a normal distribution the CLT statistic then follows a t distribution with $n-1$ degrees of freedom. Recall that the binomial test for H_0 can be conducted via the test statistic:

$$Z_B = \frac{B - n/2}{\sqrt{n/4}}.$$

To compare the binomial and CLT tests for symmetric distributions, we need to consider two issues, i) the Type I error of the two tests, and ii) the power of the tests.

Type I error

For the CLT test, Z_μ is standard normal for large samples for any distribution with finite variance. Thus the Type I error of the CLT test is approximately correct, and turns out to be ok even for moderate sample sizes.

For the Binomial test, its distributional properties under H_0 only depend on the definition of the median, and do not depend on the true underlying continuous distribution. The use of the normal approximation to the binomial distribution with the statistic Z_B works well even for moderate sample sizes.

Power

The challenge in calculating power is that it depends on the distribution of the sample under H_a. Generally we decide upon a distribution (or set of distributions) of interest and use it (them) for power comparisons.

For the CLT, if the underlying distribution is normal then Z_μ is the uniformly most powerful (UMP) test - it is the best test. However, if the underlying distribution is not normal, then the UMP property does not hold, and there is no guarantee that the CLT test will perform well.

The text illustrates the power comparison of Z_μ to Z_B for a particular alternative hypothesis (with graphs on the website for more alternatives) for two situations: i) for a normally-distributed sample and ii) a sample from the longer-tailed Laplace distribution. These examples illustrate how the CLT test does better for a normal sample, but worse for a Laplace sample.