STA 104 Applied Nonparametric Statistics

Chapter 2: One-Sample Methods for Location Problem

Xiner Zhou

Department of Statistics, University of California, Davis

Table of contents

1. Wilcoxon Signed Rank Test
2. Signed Test
3. A Comparison of Statistical Tests
4. Paired Comparisons

A Comparison of Statistical Tests

Two statistical issues in choosing between tests

// P(reject H_0 | H_0 is true) = P(false positive)

- Type I error: The probability of a Type I error should be what we claim it to be.
  - t-test:

    from the central limit theorem, valid for large samples, so probabilities of Type I errors found by using the standard normal distribution will be approximately correct.
    Many studies have shown that the approximation is quite good even for moderate sample sizes and a range of population distributions.
  - Wilcox test or Fisher's signed test: stated probability of a Type I error will generally be essentially correct.

Two statistical issues in choosing between tests

1 - P(Type II error)

= P(reject H_0 | H_a is true) = P(true positive discovery)

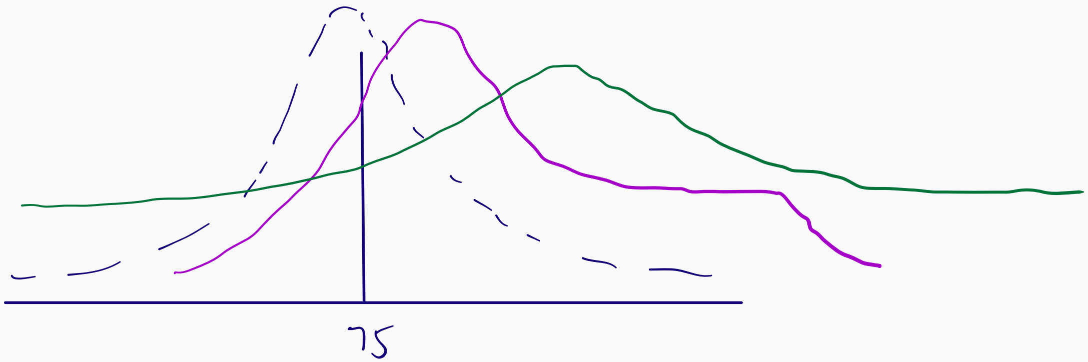
//

- Power: probability of rejecting H_0 if it is false.

  Power measures the ability of a test to detect a departure from the null hypothesis, i.e. true positive

  If two tests have the correct probability of a Type I error, then the one with the greater power is the preferred test.

  Power is the major concern since it's not controlled by design.



A researcher is testing the sodium contents in packages of food. The desired sodium content of $75\mathrm{mg}$ might not occur because the setting is too high on the machine that puts salt into the product.

For instance, if the amount of sodium per package has a normal distribution with mean $75\mathrm{mg}$ and standard deviation $2.5\mathrm{mg}$ under the correct setting of the salt machine, it may be reasonable to assume that the distribution is normal with mean $\mu > 75\mathrm{mg}$ and standard deviation $2.5\mathrm{mg}$ when the setting is too high.

However, the situation may be even more complicated. There may be a glitch in the machine that causes every tenth package, say, to have twice as much salt as it should.

Or an incorrect setting of the machine might cause both the mean and the standard deviation to change. One test may have greater power in one circumstance and the other test may be more powerful in another.

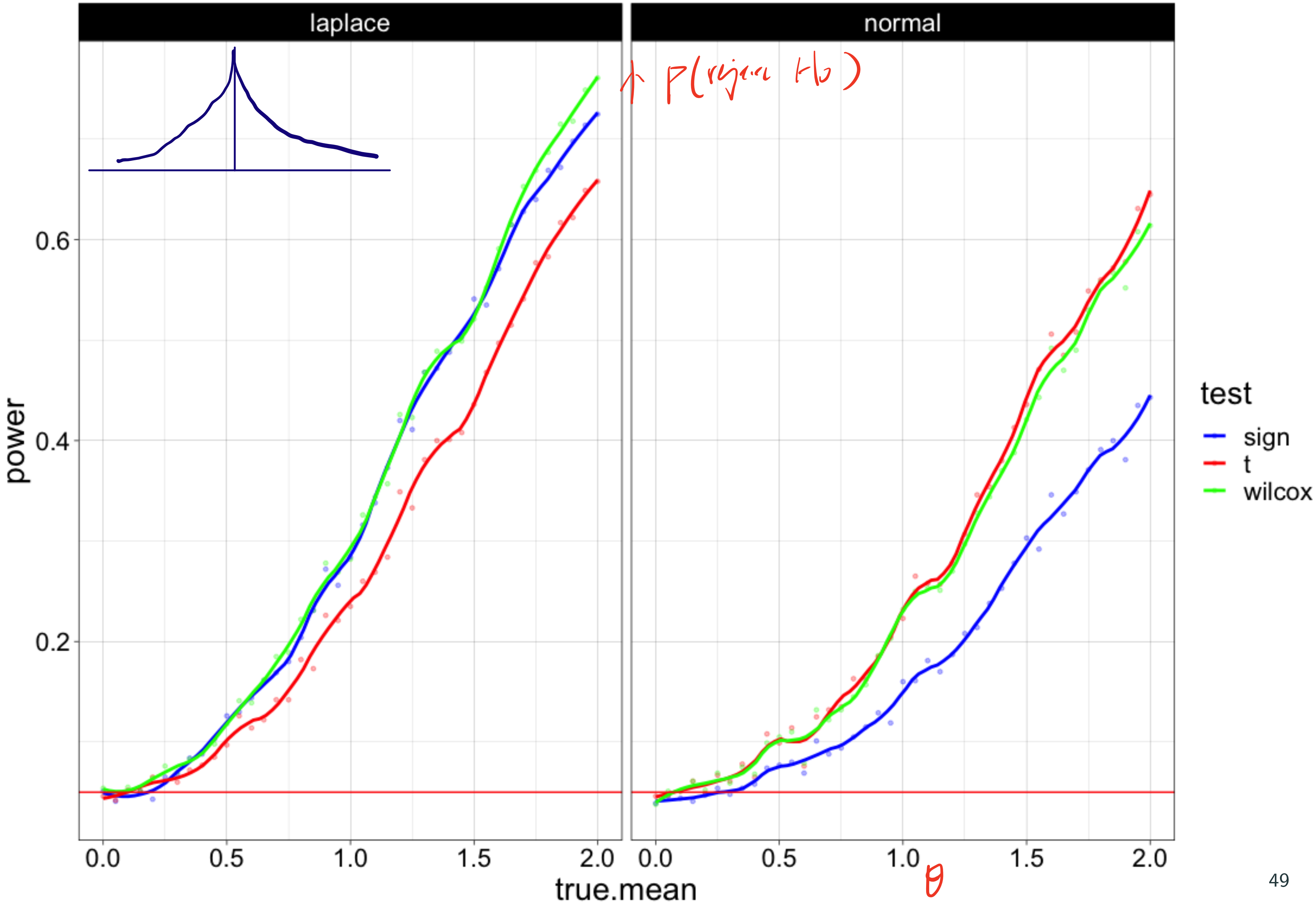
$\Rightarrow$ The strategy is to choose a test that has good power under the alternatives that seem plausible for the problem at hand and/or underlying distribution of the data.

A Simulation Study

$H_0: \theta = 0$ vs $H_1: \theta \neq 0$

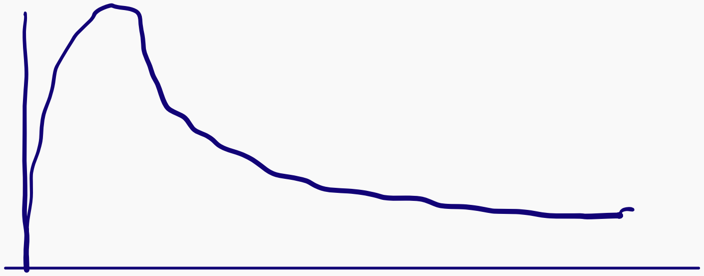
- Type I error ?
- Power

- t. dist $df = 1, 2, 3$



A Simulation Study

A heavy-tailed probability distribution is a type of probability distribution in which the tail of the distribution (i.e., the portion representing rare events) is "heavy" or thick, meaning that the probability of such events is much higher than in a distribution with a thinner tail, such as the normal distribution.



Examples of heavy-tailed distributions include

- the Pareto distribution
- the Lévy distribution

These types of distributions are often used to model phenomena such as

- income inequality
- popularity of items
- stock market fluctuations

where extreme events (such as large market crashes or a small number of individuals having a large proportion of wealth) occur more frequently than expected based on a normal distribution.

Takeaway

Generally, the nonparametric test will have higher power than the Z/t-test for heavier-tailed population distributions, but the opposite will be true for lighter-tailed distributions.

Statistical theory tells us that Z/t- test has the greatest power among all tests that have the same level of significance, and this advantage holds for all alternatives. The test is said to be uniformly most powerful when the underlying population is truly normal.

Some have mistakenly assumed that the optimal power properties of the (parametric) Z/t-test carry over to the case of sampling from nonnormal populations. This is not the case, not even for large samples.