A Multivariate Extension of Sign Test

Sahil Saini Shamriddha De Sonia Mehra

Department of Mathematics & Statistics
Indian Institute of Technology Kanpur

April 17, 2021

Outline

1. Univariate Sign Test
2. Multivariate Extension
   - Test Statistic
   - Asymptotic Null Distribution
3. Conclusions

Univariate Sign Test

Let us consider a population with distribution function F which is continuous and symmetric about its location parameter θ.

The testing problem can be stated as

$$H_0 : \theta = \theta_0 \quad \text{against} \quad H_1 : \theta \neq \theta_0,$$

which can be equivalently framed as

$$H_0 : \theta = 0 \quad \text{against} \quad H_1 : \theta \neq 0,$$

without any loss of generality.

Univariate Sign Test

Let $X_1, X_2, ..., X_n$ be a random sample from F. Define the indicator functions

$$I_i = \begin{cases} 0, & \text{if } X_i < 0 \\ 1, & \text{if } X_i > 0 \end{cases}, \quad i = 1, 2, ..., n.$$

The test statistic is formulated as

$$S_n = \sum_{i=1}^{n} I_i.$$

Under H_0, S_n should be close to $n/2$. H_0 is rejected if S_n is too larger or too smaller that $n/2$.
Moreover, S_n follows a Binomial distribution.

Multivariate Extension

Let us consider a multivariate population (of dimension p) with F_p as the distribution function.
Assumption: F_p is continuous and symmetric about its location parameter vector $\boldsymbol{\theta}$.

Testing problem:

$$H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0 \quad \text{against} \quad H_1 : \boldsymbol{\theta} \neq \boldsymbol{\theta}_0,$$

or equivalently,

$$H_0 : \boldsymbol{\theta} = 0 \quad \text{against} \quad H_1 : \boldsymbol{\theta} \neq 0,$$

without loss of generality.

Test Statistic

Let $\boldsymbol{X}_1, \boldsymbol{X}_2, ..., \boldsymbol{X}_n$ be a random sample from F_p.

Let us consider $\boldsymbol{I}_i = (I_{i1}, ..., I_{ip})^T$ $(i = 1, 2, ..., n)$, where

$$I_{ij} = \begin{cases} -1, & \text{if } X_{ij} < 0 \\ 1, & \text{if } X_{ij} > 0 \end{cases}, \quad i = 1, 2, ..., n, \quad j = 1, 2, ..., p.$$

We define

$$\boldsymbol{S}_n = \sum_{i=1}^{n} \boldsymbol{I}_i.$$

Test Statistic

Plausible test statistic:

$$T_n = \|\boldsymbol{S}_n\|^2 = \sum_{j=1}^{p} S_{nj}^2.$$

To reject H_0 at level of significance α if $t_n > (T_n)_\alpha$, where t_n is the observed value of T_n and $(T_n)_\alpha$ is the upper $100\alpha\%$ point of the distribution of T_n.

Asymptotic Distribution (under H_0)

Notes:

- $\boldsymbol{I}_i$'s are i.i.d random vectors
- The components of $\boldsymbol{I}_i$'s are bounded random variables
- Under H_0, $E(\boldsymbol{I}_1) = 0$

Using Multivariate Central Limit Theorem,

$$\frac{1}{\sqrt{n}} \boldsymbol{S}_n \xrightarrow{d} \boldsymbol{Z} \sim \mathcal{N}_p(0, \Sigma),$$

where Σ is the covariance matrix of $\boldsymbol{I}_1$.

(Σ need not be an identity matrix!)

Asymptotic Distribution (under H_0)

Continuous Mapping Theorem yields

$$\left\|\frac{1}{\sqrt{n}}\boldsymbol{S}_n\right\|^2 \xrightarrow{d} \|\boldsymbol{Z}\|^2,$$

or equivalently,

$$\frac{1}{n}T_n \xrightarrow{d} \|\boldsymbol{Z}\|^2,$$

where $\|\boldsymbol{Z}\|^2$ follows a Generalized Chi-Square distribution.

What's Next

- The exact distribution of the test statistic
- Finding explicit expressions for the parameters of the distributions while deducing asymptotic distribution
- Comparative analysis of the proposed test with some other prevailing tests.

Thank You!