NonParaAnova: One R package Consists of Non-Parametric Statistics Analysis of Variance in Different Conditions

Haozhe Pang

April 23 2021

Abstract

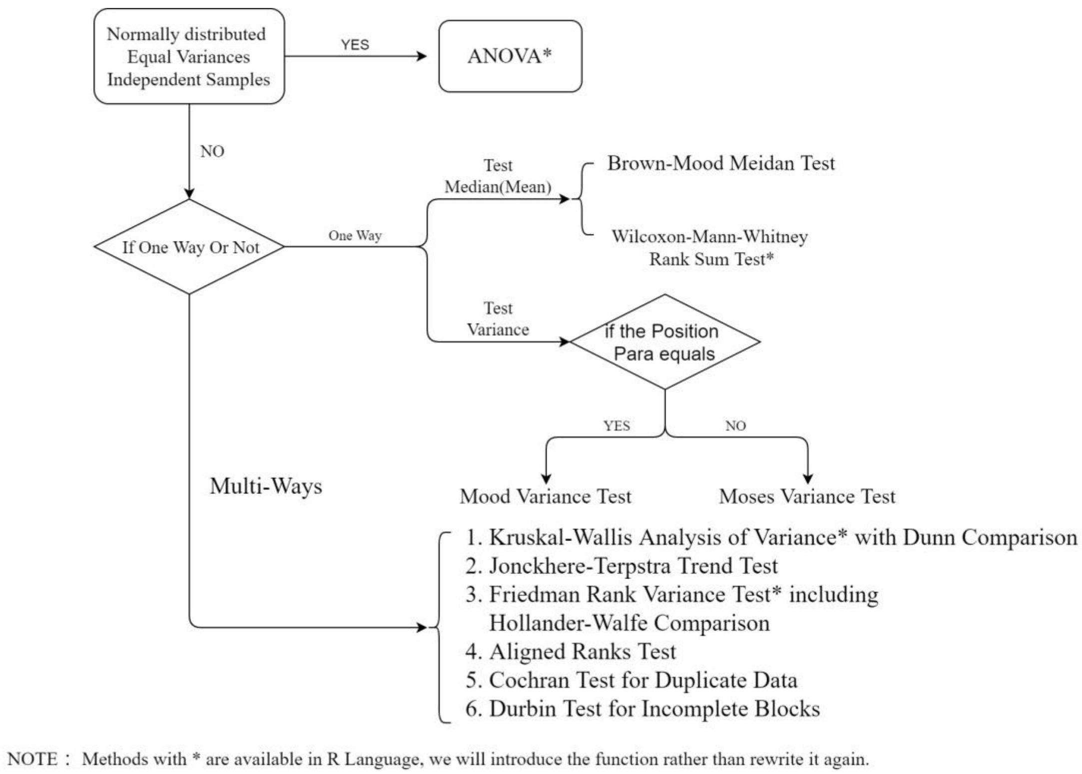
At the most time, we will can make the use of Fisher Test[1] to analysis variance of data from different treats. However, the data must satisfy some assumptions:

1. **Normally distributed experimental errors**
2. **Equal variances between treatments**
3. **Independence of samples**

Sometimes, after relevant tests, the data fail to satisfy these above. So, Non-parametric ANOVA will show up on the stage.

Right now, some methods are not available in R packages. What's more, different tests from different packages will have different Data Inputs, which gives rise to the programming difficulties. Therefore, we sincerely hope our R package can help more people from different backgrounds can carry out their works more efficiently. Here we give the overview of our package including our suggestions towards how to conduct Variance Analysis:

Also, we give an overview of ANOVA model to help users select the proper methods:



1. Analysis by Mean & Variance

1.1 Brown-Mood Median Test

Brown-Mood[2] test is a special type of χ^2 Test in the term of test median.
Compare two sets of data: X, Y, where $x_1, x_2, \dots, x_m \in X; y_1, y_2, \dots y_n \in Y$. Original Hypothesis is:

$$H_0: median_X = median_Y; H_1: median_x \neq median_Y$$

First, define the median of XY (which is mixed by X and Y) as M_{XY}

	X	Y	Total
$> M_{XY}$	A	B	t
$< M_{XY}$	C	D	$(m+n)-(A+B)$
Total	m	n	

In the condition of H_0, A satisfies the hypergeometric distribution:

$$P(A=k) = \frac{\binom{m}{k}\binom{n}{t-k}}{\binom{m+n}{t}}$$

When the data size is large, use the Normal Approximation of the Hypergeometric Distribution to estimate:

$$Z = \frac{A - mt/(m+n)}{\sqrt{(mnt(m+n-t)/(m+n)^3)}} \to N(0,1)$$

1.2 Mood ANOVA

When the data is scattered, it's hard to distinguish two groups in the angel of Rank. In 1954, Mood[2] used variance to differ two groups: For two groups X, Y, where $x_1, x_2, \dots, x_m \in X; y_1, y_2, \dots y_n \in Y$ which satisfy the distribution:

$$X_m \sim F\left(\frac{x}{\sigma_1}\right); Y_n \sim F\left(\frac{x}{\sigma_2}\right)$$

And the hypothesis is:

$$H_0: \sigma_1 = \sigma_2; H_1: \sigma_1 \neq \sigma_2$$

Under the hypothesis, define the Statistics:

$$M = \sum_{i=1}^{m} \left(R_i - \frac{m+n+1}{2}\right)^2$$

Where $R_i, i \in [1, m]$ is the rank of data in X after mixing X and Y.
In the condition of **large sample**, M has the following statistical properties:

$$E(M) = \frac{m(m+n+1)(m+n-1)}{12}$$

$$Var(M) = \frac{mn(m+n+1)(m+n+2)}{180}$$

Obviously, $Z = \frac{M - E(M)}{\sqrt{Var(M)}} \to N(0,1)$

In the condition of **small sample**, Z can be corrected:

$$Z = \frac{M - E(M) + 0.5}{\sqrt{Var(M)}} \rightarrow N(0,1)$$

1.3 Moses Test

In 1963, Moses[3] proposed one method where assuming equal mean is unnecessary to test the variance.

Here is the algorithm:

Given $x_1, x_2, \dots x_m \in X;\ y_1, y_2, \dots, y_n \in Y$, the variances of two treats are δ_1, δ_2.

The hypothesis is:

$$H_0: \sigma_1 = \sigma_2; H_1: \sigma_1 \neq \sigma_2$$

The core of Moses Test is T Statistics, and it's defined as follows:

1. Divide data in Treat 1 into m_1 groups, while divide data in Treat 2 into m_2 groups. At the same time, they have k observations in each group.
2. Define the SST of each group:

$$SSA_r = \sum_{x_i \in A_r} (x_i - \bar{x})^2, r \in [1, m_1]$$

$$SSB_s = \sum_{y_j \in B_s} \left(y_j - \bar{y}\right)^2, s \in [1, m_2]$$

3. Mix all the SSA_r and SSB_s, and rank them.
4. Moses Statistics T_M is:

$$T_M = S - \frac{m_1(m_1 + 1)}{2}$$

Where $S = min\{\sum rank(SSA_r), \sum rank(SSB_s)\}$

5. When the sample is small (<20 in total), check Mann-Whitney W_α value. If $T_M < W_{\alpha, m_1, m_2}$, accept H_1. Else, if the sample is large, a p value can be calculated as:

$$Z_U = \frac{|U - n_1 n_2 / 2|}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}}$$

2. Data from Multi-Treats with Multi-Blocks

To test the effect of one work, sometimes the researchers would do experiments on different objects, which is defined as Blocks here. The most classical method is Kruskal-Wallis test. It has already installed inside R Language, and can be used in the form of *Kruskal.test()*. However, Kruskal's cannot suit all kinds of condition, hence, the Statistics Scientists did some further works.

2.1 Kruskal-Wallis Analysis of Variance with Dunn Comparison

Kruskal-Wallis Analysis[4] is the most famous non-parametric way, and it is widely used in different fields to test whether the difference exists. Function kruskal.test() is available in R.
Since there are more than one blocks, and if the result of Kruskal test shows that the relationship exists among blocks, a comparison is strongly suggested between blocks. Here, Dunn.test()[5] can be conducted.

2.2 Jonckheere–Terpstra Test

Sometimes, trend exists in the data like rising or descending, Jonckheere–Terpstra[6] Test proposed one way to help check this. Users can use function JT.test().
Given $X_{11}, \dots, X_{1n_1}, X_{21}, \dots, X_{2n_2}, \dots . X_{k1}, \dots, X_{kn_k}$ from k treats, and the Null Hypothesis is:

$$H_0: \theta_0 = \cdots . = \theta_k; H_1: \theta_0 < \cdots < \theta_k$$

1. Define $W_{ij} = Count(X_{iu} < X_{jv}), where: u \in [1, n_i]; v \in [1, n_j]$
2. Calculate Statistics $J = \sum_{i<j} W_{ij}$
3. When the sample is small (<30 in total for instance)

   When the sample is large:

$$Z = \frac{J - \left(N^2 - \sum_{i=1}^{k} n_i^2\right)/4}{\sqrt{\left[N^2(2N+3) - \sum_{i=1}^{k} n_i^2(2n_i+3)\right]/72}} \to N(0,1)$$

Especially, since when we calculate the Rank of sample, plots usually exist. Here is how it's collected.
If plots exist,

$$W_{ij}^* = count(X_{ik} < X_{jl}, where: k \in [1, n_i], l \in [1, n_j]) + \frac{1}{2} count(X_{ik} = X_{jl}, where: k \in [1, n_i], l \in [1, n_j])$$

And J can be collected as:

$$J^* = \sum_{i<j} W_{ij}^*$$

And it can be tested with Z test:

$$Z = \frac{J^* - E(J^*)}{\sqrt{(Var(J^*))}} \to N(0,1)$$

Where: $E(J^*) = \frac{N^2 - \sum_{i=1}^{k} n_i^2}{4}$

$$Var(J^*) = \frac{1}{72}[N(N-1)(2N+5) - \sum_{i=1}^{k}(n_i(n_i-1)(2n_i+5))) - \sum_{i=1}^{k}(\tau_i(\tau_i-1)(2\tau_i+5)]$$

$$+$$

$$\frac{1}{36N(N-1)(N-2)}\left[\sum_{i=1}^{k} n_i(n_i-1)(n_i-2)\right] * \left[\sum_{i=1}^{k} \tau_i(\tau_i-1)(\tau_i-2)\right]$$
$$+$$
$$\frac{1}{8N(N_1)}\left[\sum_{i=1}^{k} n_i(n_i-1)\right]\left[\sum_{i=1}^{k} \tau_i(\tau_i-1)\right]$$

2.3 Friedman Test with Hollander-Wolfe Comparison

Sometimes, it's necessary to consider the impact of different blocks, one of the classical methods is **Firedman test**[7], which is available in R Language as *Friedman.test()*
If the result of Friedman Test shows that the relationship exists between different treats, **Hollander-Wolfe Comparison**[8] can find the pairwise relationship.
First we define the Rank Sum from treat i as $R_{\cdot i}$: $R_{\cdot i} = \sum_{k=1}^{K} R_{ki}$ $(K\ is\ the\ length\ of\ Treat)$
And the Comparison between Treats can be described as:

$$D_{ij} = \frac{|R_{\cdot i} - R_{\cdot j}|}{SE}$$

Where: $SE = \sqrt{k(k+1)/6}$
If the tie exists, $SE = \sqrt{\frac{k(k+1)}{6} - \frac{b\sum(\tau_i^3-\tau_i)}{6(k-1)}}$,$\tau_i$ is the length of each tie.

3.3 Aligned Ranks Test

If between different blocks, a hug gap will affect the result of test towards treats. Hodges and Lehmmann[9] proposed the famous Hodges-Lehmmann Test where ranks would be aligned. The mean or median can be viewed as a kind of estimation of blocks' effect in some ways, so the effect of blocks can be easily erased by constructing its mean or median with the data.

2.4 Cochran Test

Sometimes, the data will show again and again, especially some indication data, which will make the test performs bad. So, the Corhan test[10] is suggested.

2.5 Durbin Test

Considering that if the blocks are incomplete, in 1951, Durbin[11] proposed one Rank method designed for this condition.

3 Examples

Please check the file: example.Rmd in the Github

4 References

[1] Anscombe, F. J. (1948). "The Validity of Comparative Experiments". Journal of the Royal Statistical Society. Series A (General). 111 (3): 181–211. doi:10.2307/2984159. JSTOR 2984159. MR 0030181

[2] Brown, George W., and Alexander M. Mood. "On median tests for linear hypotheses." *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*. The Regents of the University of California, 1951.

[3] Moses, Lincoln E. "Rank tests of dispersion." *The annals of mathematical statistics* (1963): 973-983.

[4] Kruskal; Wallis (1952). "Use of ranks in one-criterion variance analysis". Journal of the American Statistical Association. 47 (260): 583–621. doi:10.1080/01621459.1952.10483441

[5] Dunn, Olive Jean (1964). "Multiple comparisons using rank sums". Technometrics. 6 (3): 241–252. doi:10.2307/1266041

[6] Jonckheere, A. R. (1954). "A distribution-free k-sample test against ordered alternatives". Biometrika. 41: 133–145. doi:10.2307/2333011

[7] Friedman, Milton (December 1937). "The use of ranks to avoid the assumption of normality implicit in the analysis of variance". Journal of the American Statistical Association. 32 (200): 675–701

[8] Hollander, M.; Wolfe, D. A. (1973). Nonparametric Statistics. New York: J. Wiley. ISBN 978-0-471-40635-8

[9] Hodges Jr, J. L., and Erich L. Lehmann. "Comparison of the normal scores and Wilcoxon tests." Proc. Fourth Berkeley Symp. Math. Statist. Prob. Vol. 1. 1961.

[10] Tate, Merle W., and Sara M. Brown. "Note on the Cochran Q test." Journal of the American Statistical Association 65.329 (1970): 155-160.

[11] Durbin, James. "Incomplete blocks in ranking experiments." *British Journal of Statistical Psychology* 4.2 (1951): 85-90.