Test U de Mann-Whitney (test de somme des rangs de Wilcoxon) en Python [pandas et SciPy]

Renesh Bedré   5 minutes de lecture

Ⓧ

Test Mann-Whitney U

- Le test U de Mann-Whitney est un test non paramétrique qui est une alternative au test *t* (https://www.reneshbedre.com/blog/ttest.html#two-sample-t-test-unpaired-or-independent-t-test) paramétrique à (https://www.reneshbedre.com/blog/ttest.html#two-sample-t-test-unpaired-or-independent-t-test)deux échantillons (https://www.reneshbedre.com/blog/ttest.html#two-sample-t-test-unpaired-or-independent-t-test) . Il est d'abord proposé par Frank Wilcoxon (1945) et plus tard travaillé par Henry Mann et Donald Whitney (1947). Par conséquent, le test U de Mann-Whitney est également connu sous le nom de test de somme des rangs de Wilcoxon ou test de Wilcoxon-Mann-Whitney (WMW) .

> *Le test de somme des rangs de Wilcoxon est différent du test de somme des rangs signés de Wilcoxon . Sur des données appariées, le test de somme des rangs signés de Wilcoxon est utilisé.*

- Test U de Mann-Whitney utilisé pour comparer les différences entre deux groupes indépendants. Il teste l'hypothèse que si les deux groupes proviennent de la même population ou ont les mêmes médianes. Il ne suppose aucune distribution spécifique (telle que la distribution normale des échantillons) pour le calcul des statistiques de test et des valeurs *p* . S'il y a plus de deux groupes à analyser, vous devriez envisager le test de Kruskal-Wallis (https://www.reneshbedre.com/blog/kruskal-wallis-test.html) .
- Les rangs moyens ou les médianes de l'échantillon (et non les moyennes) sont comparés dans le test U de Mann-Whitney en fonction de la forme de distribution de deux groupes indépendants, ce qui le distingue du test (https://www.reneshbedre.com/blog/ttest.html#two-sample-t-test-unpaired-or-independent-t-test)*t* (https://www.reneshbedre.com/blog/ttest.html#two-sample-t-test-unpaired-or-independent-t-test) , qui compare les moyennes de l'échantillon.
- Le test U de Mann-Whitney peut être appliqué sur de petits (5-20) et grands échantillons (n > 20). La puissance augmente avec la taille de l'échantillon.
- Bien que le test U de Mann-Whitney et le test *t aient une* puissance statistique (https://www.reneshbedre.com/blog/hypothesis-testing.html#type-i-%CE%B1-type-ii-errors-%CE%B2-and-power-1-%CE%B2) similaire , il est toujours judicieux d'utiliser le test *t* si ses hypothèses sont satisfaites.

Hypothèses du test Mann-Whitney U

- Les observations des deux groupes doivent être sélectionnées au hasard parmi les populations cibles
- Les observations sont indépendantes les unes des autres
- Les observations doivent être continues ou ordinales (par exemple, données d'éléments de Likert) (https://www.reneshbedre.com/blog/others.html#variable-types)

Test U de Mann-Whitney Hypothèses

Si nous avons deux groupes indépendants avec des observations $x_1, x_2, ..., x_m$ et $y_1, y_2, ..., y_n$ échantillonnées à partir des populations X et Y, alors le test U de Mann-Whitney compare chaque observation x_i à partir de l'échantillon x avec chaque observation (y_j) de l'échantillon y.

Ⓧ

Above two-sided *alternative hypothesis* tests that there is equal probability of x_i is greater or lesser than y_j (both groups came from same population),

One-sided *alternative hypothesis* tests probability of x_i is greater than y_j and vice versa.

We can also state the two-sided hypothesis in terms of *median* as (when two groups have same shape of distribution)

Null hypothesis: Two groups have equal median

Alternative hypothesis: Two groups does not have equal median

One-sided *alternative hypothesis* tests median from one group can be greater or lesser than other group.

Learn more about hypothesis testing and interpretation (https://www.reneshbedre.com/blog/hypothesis-testing.html)

Mann-Whitney U Test formula

The *p* value is calculated based on the comparison between the critical value (https://sphweb.bumc.bu.edu/otlt/mph-modules/bs/bs704_nonparametric/Mann-Whitney-Table-CriticalValues.pdf) and the *U* value. If *U* value <= critical value, we reject the *null hypothesis* and vice versa. If the sample is large (n>20), the *p* value is calculated based on the normal approximation using standardized test statistics.

How Mann-Whitney U Test works?

X

- Merge the data from two samples and rank them from smallest to largest

- Calculate the sum of rank for each sample (Rx and Ry)
- Calculate Mann-Whitney test statistic (U) using the formula (minimum of Ux and Uy)
- Calculate p value by comparing U with the critical value

Perform Mann-Whitney U test in Python

Mann-Whitney U test example

Suppose, there are two plant genotypes (A and B) differing in their yield phenotype. Mann-Whitney U test is appropriate to compare the yield of two genotypes under the assumption that yield output does not follow the normal distribution.

Get example dataset and summary statistics

Load hypothetical plant genotypes (A and B) yield dataset,

Learn how to import data using pandas (https://www.reneshbedre.com/blog/import-data-pandas.html)

```
import pandas as pd
df = pd.read_csv("https://reneshbedre.github.io/assets/posts/mann_whitney/genotype.csv")
df.head(2)
     A   B
0   60  10
1   30  25


# get summary statistics
df.agg(["count", "min", "max", "median", "mean", "skew"])
                A          B
count   23.000000  23.000000
min     20.000000  10.000000
max     60.000000  32.000000
median  56.000000  28.000000
mean    47.695652  25.217391
skew    -0.710884  -1.270302


# generate boxplot to check data spread
import matplotlib.pyplot as plt
df.boxplot(column=['A', 'B'], grid=False)
plt.show()
```

X

Check data distribution

Check data distribution using Shapiro-Wilk test and histogram (https://www.reneshbedre.com/blog/anova.html#test-anova-assumptions) ,

```python
import scipy.stats as stats
w, pvalue = stats.shapiro(df['A'])
w, pvalue
(0.8239281177520752, 0.0009495539125055075)


w, pvalue = stats.shapiro(df['B'])
w, pvalue
(0.7946348190307617, 0.00031481595942750573)


# plot histogram
```

```python
ax2.hist(df['B'], bins=10, histtype='bar', ec='k')
ax1.set_xlabel("Yield")
ax2.set_xlabel("Yield")
plt.show()
```

As the p value obtained from the Shapiro-Wilk test is significant ($p < 0.05$), we conclude that the data is not normally distributed. Further, in histogram data distribution shape does not look normal. Therefore, Mann-Whitney U test is more appropriate for analyzing two samples.

Perform Mann-Whitney U test

Perform two-sided (https://www.reneshbedre.com/blog/hypothesis-testing.html#one--and-two-tailed-sided-alternate-hypothesis) (yield of two genotypes does not have equal medians) Mann-Whitney U test,

Note: We are comparing median as two genotypes have similar shape of distribution (see histogram and boxplot). If two groups do not have similar shape of distribution, you should compare mean ranks.

```
# SciPy v1.7.1
import scipy.stats as stats
# perform two-sided test. You can use 'greater' or 'less' for one-sided test
stats.mannwhitneyu(x=df['A'], y=df['B'], alternative = 'two-sided')
# output
MannwhitneyuResult(statistic=489.5, pvalue=7.004695394561267e-07)
```

Check online calculator (https://www.reneshbedre.com/blog/mann-whitney-u-test-calculator.html) for performing Mann-Whitney U test

Note: *In the above example, the p value obtained from mannwhitneyu is based on the normal approximation as the sample size is large (n > 20). If the sample size is small, a normal approximation is not appropriate. To get exact p value, set method="exact". The* `mannwhitneyu` *function automatically calculates the exact p value when one of the sample size is < 8. Both exact and normal approximation p values should be roughly similar.*

Mann-Whitney U test interpretation: As the *p* value obtained from the Mann-Whitney U test is significant (U = 489.5, p < 0.05), we conclude that the yield of the two genotypes significantly different from each other .

Perform one-sided (https://www.reneshbedre.com/blog/hypothesis-testing.html#one--and-two-tailed-sided-alternate-hypothesis) (median yield of A genotype is greater than median yield of genotype B) Mann-Whitney U test,

```python
import scipy.stats as stats
stats.mannwhitneyu(x=df['A'], y=df['B'], alternative = 'greater')
# output
MannwhitneyuResult(statistic=489.5, pvalue=3.5023476972806333e-07)
```

As the *p* value (https://www.reneshbedre.com/blog/how-to-calculate-p-value.html) obtained from the Mann-Whitney U test is significant (U = 489.5, p < 0.05), we conclude that the yield of the A genotype significantly greater than the genotype B.

Related reading

- Friedman test using R (with examples and code) (https://www.reneshbedre.com/blog/friedman-test.html)

- What is *p* value and how to calculate *p* value by hand (https://www.reneshbedre.com/blog/how-to-calculate-p-value.html)

References

2. Mann–Whitney Test (https://link.springer.com/referenceworkentry/10.1007%2F978-0-387-32833-1_243)

3. Mann Whitney U Test (Wilcoxon Rank Sum Test) (https://sphweb.bumc.bu.edu/otlt/mph-modules/bs/bs704_nonparametric/BS704_Nonparametric4.html)

If you enhanced your knowledge and practical skills from this article, consider supporting me on

(https://www.buymeacoffee.com/renesh)

X

Tags: | Statistics |

Updated: November 7, 2021