비모수 검정

언제? N<30 , 정규성을 가정할 수 없는 경우

-관측된 자료가 특정분포를 따른다고 가정할 수 없는 경우.

-관측된 자료의 수가 30 개 미만으로 적은 경우.

-자료가 개체간의 서열관계를 나타내는 경우.

어떻게? 관측값의 절대적인 크기에 의존하지 않는 관측값들의 순위나 두 관측값 차이의 부호 등을 이용해 검정한다.

'분포의 형태가 동일한지' 또는 '분포의 형태가 동일하지 않은지'와 같이 분포의 형태에 대해 설정한다.

종류?

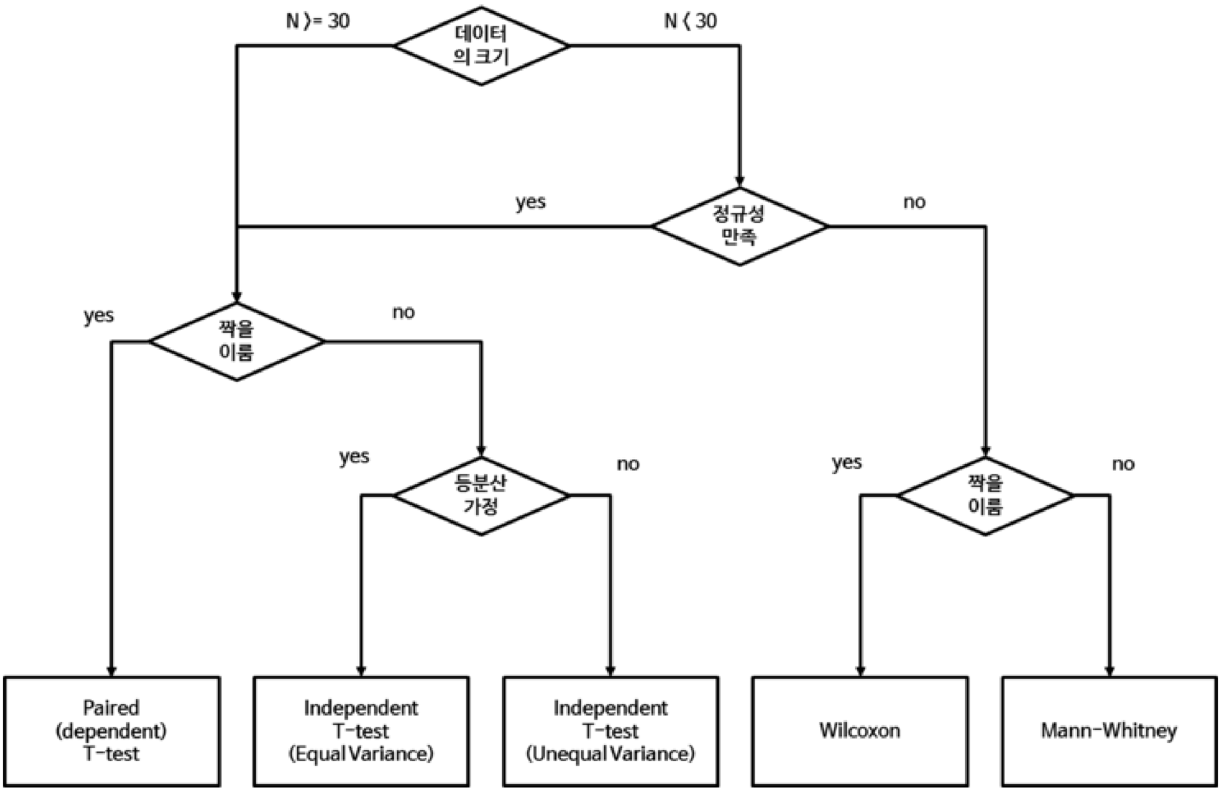
부호 검정(sign test)

월콕슨의 순위합검정(rank sum test) = 만-위트니의 U 검정

윌콕슨의 부호순위합검정(Wilcoxon signed rank test)

런 검정(run test)

스피어만의 순위상관계수



==

부호검정(sign test)

두 표본들의 분포가 동일한지 여부를 검정

부호검정은 +혹은 - 부호가 나올 확률 p 를 0.5 로 갖는 이항부호검정(Binominal sign test)라고도 불린다. 이 검정 방법은 관측치의 플러스와 마이너스 부호의 방향에 기반하며, 관측치의 수치에 기반하지 않는다.

부호검정의 가정

- 데이터의 분포는 정규분포를 따를 필요가 없다.
- 서로 다른 모집단이나 동일한 모집단으로부터 나온 2 개의 표본이 필요하다.
- 2 개의 표본이 동일한 모집단에서 나왔다면, 'before-after' sample 과 같이 짝 지어진 샘플이어야 한다.

부호검정은 짝 지어진 관찰치들끼리 비교하여 위(+)인지 아래(-)인지만 검증하고 그 쌍의 차이를 측정하지는 않기 때문에 약한 검정 방법으로 본다. 반면, 이 방법은 최소한의 가정만을 사용하기 때문에 모수적 가정이 잘못되어 생기는 오류의 가능성이 적고, 부호에 기초한 방법이기 때문에 이상치에 민감하지 않다는 장점도 있다.

하나의 샘플로 부호검정을 하는 경우,

귀무가설 : "두 집단의 분포는 **차이가 없다 (동일하다)"**이다.

이 검정방법은 one sample t-test 의 대안법(alternative)이다.

모집단의 중위수(평균 또는 중앙값)와 표본의 관측치를 서로 빼서 나온 값이 양수이면 (+), 음수이면 (-)로 부호를 계산한다.

짝 지어진 샘플로 부호검정을 하는 경우,

귀무가설은 : "두 집단의 분포는 **차이가 없다 (동일하다)"** 이다.

이 검정방법은 independent two sample t-test 혹은 paired t-test 의 대안법이다.

짝 지어진 샘플들의 값을 빼서 나온 값이 양수이면 (+), 음수이면 (-)로 부호를 계산한다.

부호검정의 과정

1) 주어진 분포에 대해 **플러스(+)와 마이너스(-) 부호를 계산한다**. 계산한 값이 0 인 경우는 0 으로 입력한다.
2) 0 으로 입력한 경우는 제외하고, **전체 부호들의 개수를 n 으로 적는다**. 성공횟수를 x, 성공확률을 p 로 놓는다. 부호검정의 경우 p 는 0.5 이다. 즉, 성공확률과 실패확률이 동일하다.
   검정하고자 하는 대립가설을 설정한다.

   - 모집단과 표본집단의 분포는 차이가 있다 (H1 :M ≠ M0)
   - 표본 집단의 중위값이 더 작다 (H1 : M > M0)
   - 표본 집단의 중위값이 더 크다 (H1 : M < M0)

3) **p 값을 구하고, 유의수준과 비교해서** 귀무가설을 채택할 것인지 대립가설을 채택할 것인지를 판단한다. 일반적으로 유의수준은 5% 혹은 1%로 정한다.

부호검정 예제

set1 은 휴가 전의 생산성을 나열한 리스트이고, set2 는 휴가 후의 생산성을 나열한 리스트이다.
두 리스트는 전후 관계로 짝지어진 데이터이다.
numpy 를 사용하여 두 리스트를 array 로 만들고, 부호를 계산했다.

```python
#set1: 휴가 전 생산성, set2: 휴가 후 생산성
set1 =[83, 85, 75, 91, 80, 75, 90, 65, 78, 85, 83, 75, 78, 80, 82, 88, 85, 80, 78, 81, 70, 80]
set2 = [79, 87, 70, 93, 85, 75, 80, 71, 80, 88, 82, 71, 75, 85, 86, 85, 82, 87, 78, 84, 85, 81]
```

```python
#짝지어진 데이터 확인
print(len(set1), len(set2))
```

```
22 22
```

```python
import numpy as np
```

```python
set1=np.array(set1)
set2=np.array(set2)
```

```python
result=set2-set1
signs=[]
for value in result:
    if value <0:
        signs.append("-")
    elif value >0:
        signs.append("+")
    else:
        signs.append(0)
```

부호를 계산한 결과와 데이터를 pandas 의 DataFrame 함수를 사용해 데이터프레임으로 정리하였다. 데이터프레임에 head 함수를 사용하면 지정한 개수만큼의 행들을 확인할 수 있다. 여기서 (+)의 개수를 x 로 놓고, 0 을 제외한 (+)와 (-)의 개수의 합을 n 으로 놓는다.

```python
from pandas import DataFrame
```

```python
# 부호 계산
df=DataFrame({"set1":set1, "set2":set2, "signs":signs})
df.head(6)
```

	set1	set2	signs
0	83	79	-
1	85	87	+
2	75	70	-
3	91	93	+
4	80	85	+
5	75	75	0

```python
# x = (+)개수, n = 0을 제외한 (+)와 (-)의 개수
x = df['signs'][df['signs']=="+"].count()
n = df['signs'][df['signs']=="-"].count() + x
print("x: %d, n: %d" %(x, n))
```

```
x: 12, n: 20
```

파이썬 math 라이브러리에는 팩토리얼(factorial)을 계산할 수 있는 함수가 있다. 팩토리얼 함수를 활용해서 해당 조건일 때의 p-value 를 계산할 수 있다. 일반적으로 n 이 30 이하이면 이항분포표를 기준으로 검정하고, 30 초과이면 근사 정규분포로 검정한다.

```python
#팩토리얼 함수
from math import factorial as fac
```

```python
#귀무가설: 휴가 후의 생산량이 더 크지 않다
#대립가설: 휴가 후 생산량이 더 크다 (오른쪽 검정)
```

```python
#방법1) p value를 계산해서 유의수준 0.05와 비교 (0.252는 0.05에 미치지 못하므로 대립가설 기각)
#결론: 휴가 후의 생산량이 더 크지 않다
goal1=0 #
for i in range(x, n+1):
    if i != n:
        print(n,"C", i, "x 0.5^%d"%i, "x 0.5^%d"%(n-i), '+')
        goal1 += fac(n)/fac(i)/fac(n-i)
    else:
        print(n,"C", i, "x 0.5^%d"%i, "x 0.5^%d"%(n-i))
        goal1 += fac(n)/fac(i)/fac(n-i)
print(goal1)
print("=> %d번 수행 시 +가 %d번 이상 나올 확률" %(n,x))
print("= 오른쪽 검정 시(right tail test) p-value: ", "%0.3f" %(goal1*(0.5)**n))
```

```
20 C 12 x 0.5^12 x 0.5^8 +
20 C 13 x 0.5^13 x 0.5^7 +
20 C 14 x 0.5^14 x 0.5^6 +
20 C 15 x 0.5^15 x 0.5^5 +
20 C 16 x 0.5^16 x 0.5^4 +
20 C 17 x 0.5^17 x 0.5^3 +
20 C 18 x 0.5^18 x 0.5^2 +
20 C 19 x 0.5^19 x 0.5^1 +
20 C 20 x 0.5^20 x 0.5^0
263950.0
=> 20번 수행 시 +가 12번 이상 나올 확률
= 오른쪽 검정 시(right tail test) p-value:  0.252
```

```python
#방법2) 이항분포표에서 20개 시행의 유의수준 0.05 우측검정의 임계치 확인: 14개 이상이 0.0577 (12개는 이에 미치지 못하므로 대립가설 기각)
```

부호검정: p-value 계산해서 유의수준과 비교하기 (math)

오른쪽 검정 기준으로 (+)가 12 이상일 때의 p-value 는 0.252 로 유의수준 0.05 의 기각역 안으로 들어오지 못했다. 이에 따라 대립가설을 기각하고 귀무가설을 채택한다. 결론은 "휴가 후의 생산량이 휴가 전보다 더 크지 않다"이다.

==

윌콕슨 순위합검정(Wilcoxon rank sum test) = Mann-Whitney U Test

윌콕슨 순위합검정 역시 모집단의 분포를 모르는 경우, 시행하는 비모수적 검정방법이다. 이 검정방법은 독립적인 두 표본으로 검정하며, 두 표본의 모집단의 중앙값이 동일한지를 검정한다. 이에 따라 귀무가설은 "두 표본의 중앙값은 차이가 없다"이고, 대립가설은 "두 표본의 중앙값은 차이가 있다"이다. 이 **검정은 만-위트니 U 검정 (Mann-Whitney U-test)와 일치하며**, independent two sample t-test 의 대안법이다.

계산은 두 표본의 결과를 오름차순으로 정렬하고 순위를 부여 한다. 순위 부여 시 결과가 같으면 해당 순위의 평균값을 동일하게 적용한다. 그리고 표본의 순위들을 각각 합한다. 이 표본들의 순위합과 표본들의 개수를 사용하여 검정통계량을 계산하고 p-value 를 확인하여 어떤 가설을 채택할 것인지 판단한다.

파이썬의 scipy 모듈이 ranksums 와 mannwhitneyu 함수를 제공하고 있어 검정 통계량과 p-value 를 간단하게 얻을 수 있다. (from scipy.stats import ranksums, mannwhitneyu)

2) Mann-Whitney U 검정
두 개의 독립된 집단이 정규분포를 따르지 않을 때 사용한다.
연구 배경: V 반도체 회사는 공장 A 와 공장 B 2 개의 공장에서 반도체를 생산하고 있다.
이 때, 2 개 공장의 생산 효율성에 차이가 있는지를 Mann-Whitney U 검정을 통해 살펴본다.

가설검정 **귀무가설 = 공장 A 와 공장 B 의 생산효율성은 동일하다**
연구가설 = 공장 A 와 공장 B 의 생산효율성은 다르다.

```
import pandas as pd
from scipy.stats import mannwhitneyu

x = [12, 11, 13, 14, 15]
y = [16, 15, 17, 19, 20]

print(mannwhitneyu(x, y))

# 생산량 평균 순위 출력
xy = pd.DataFrame(x+y)
xy['생산량 순위'] = xy.rank(ascending=False)
xy['공장이름'] = ["A", "A", "A", "A", "A", "B", "B", "B", "B", "B"]
print(xy.groupby('공장이름').mean())
```

```
MannwhitneyuResult(statistic=0.5, pvalue=0.007985348176890053)
         0  생산량 순위
공장이름
A     13.0     7.9
B     17.4     3.1
```

- 검정통계량이 `0.5` 그리고 유의확률은 0.0079 이므로 귀무가설을 기각한다.
- 즉, A 와 B 의 공장 생산의 효율성은 다르며, A 의 효율성이 떨어지는 것을 볼 수 있다.

https://dschloe.github.io/python/python_edu/05_statistics/chapter_17_1_nonparametric_stat_analysis/

윌콕슨 부호순위검정(Wilcoxon signed-rank test)

이 검정방법은 짝지어진 두 표본으로 검정하며, **두 표본의 중앙값이 동일한지를 검정한다**. 이에 따라 **귀무가설은 "두 표본의 중앙값은 차이가 없다**"이고, 대립가설은 "두 표본의 중앙값은 차이가 있다"이다. 이 검정은 paired t-test 의 대안법이다.

계산은 두 표본들의 값의 차이의 부호와 절대값을 계산한다. 부호검정 때와 마찬가지로, 값이 0 인 경우는 제외하고 샘플 사이즈를 정의한다. 순위합검정 때와 마찬가지로, 두 표본들을 합치고 표본들의 값의 차이의 절대값을 오름차순으로 배열한 후 순위를 부여한다. (+) 부호인 것들의 순위합, (-) 부호인 것들의 순위합, 표본들의 개수를 사용하여 검정통계량을 계산하고 p-value 를 확인하여 어떤 가설을 채택할 것인지 판단한다.

파이썬의 scipy 모듈이 wilcoxon 함수를 제공하고 있어 검정 통계량과 p-value 를 간단하게 얻을 수 있다. (from scipy.stats import wilcoxon)

(1) Wilcoxon 부호-순위 검정

- 쌍체 표본 t-검정에서 표본의 수가 30 개 미만일 경우에 활용한다.
- 가설 검증
  - 배경: 기존의 물류 알고리즘보다 개선되었다고 알려진 새로운 물류 경로 최적화 알고리즘을 도입해 상품의 배송시간을 단축하고자 한다.
  - 귀무가설: 기존 물류 알고리즘과 신규 알고리즘간을 통한 평균 배송시간은 차이가 없다.
  - 대립가설: 기존 물류 알고리즘과 신규 알고리즘간을 통한 평균 배송시간은 차이가 있다.

```
from scipy.stats import wilcoxon

x = [10, 30, 9, 21, 35, 12, 17]
y = [8, 27, 16, 25, 30, 13, 11]

wilcoxon(x, y)
```

```
WilcoxonResult(statistic=12.0, pvalue=0.7353166906373405)
```

- 검정 통계량이 12.0, p-value 은 0.735 로 출력되었는데, 이는 귀무가설을 기각하지 못하므로 신규 알고리즘을 도입할 필요가 없다.

Spearman’s Rank Correlation

Tests whether two samples have a monotonic relationship.

Assumptions

- Observations in each sample are independent and identically distributed (iid).
- Observations in each sample can be ranked.

Interpretation

- H0: the two samples are independent.
- H1: there is a dependency between the samples.

Python Code

```
# Example of the Spearman's Rank Correlation Test
from scipy.stats import spearmanr
data1 = [0.873, 2.817, 0.121, -0.945, -0.055, -1.436, 0.360, -1.478, -1.637, -1.869]
data2 = [0.353, 3.517, 0.125, -7.545, -0.555, -1.536, 3.350, -1.578, -3.537, -1.579]
stat, p = spearmanr(data1, data2)
print('stat=%.3f, p=%.3f' % (stat, p))
if p > 0.05:
	print('Probably independent')
else:
	print('Probably dependent')
```

```
stat=0.855, p=0.002
Probably dependent
```

17 Statistical Hypothesis Tests in Python (Cheat Sheet)
https://machinelearningmastery.com/statistical-hypothesis-tests-in-python-cheat-sheet/

정규성 검정

정규분포 확인방법 : Shapiro 값의 유의확률이나 Kolmogorov-Smirnov 유의확률
각각의 변수가 **0.05 보다 낮은** 유의확률을 보일 때 귀무가설(정규분포이다)
기각하여 정규분포가 충족되지 않는다고 볼 수 있습니다

비모수 검정: Mann-Whitney U Test 외

Mindy 2021. 10. 1. 22:45

1. Nonparametric Statistics 비모수 통계학

통계학은 모수 통계학과 비모수 통계학으로 나눌 수 있습니다.

모수 통계학은 모집단이 정규분포라는 가정이 필요하지만,

비모수 통계학은 분포에 대한 가정 없이 가설검정을 수행할 수 있습니다.

- 모수 데이터: 알려진 확률분포를 지닌 모집단에서 나온 데이터
- 비모수 데이터: 확인되지 않은 확률분포를 지닌 모집단에서 나온 데이터

모수 통계학에서는 평균(Mean)이 중요한 반면,

비모수 통계학에서는 중위수(Median)이 중요합니다.

검정하는 내용도 둘 이상의 데이터셋이 같은 평균을 가지는지 확인하는 모수 통계학과는 달리,

비모수 통계학에서는 둘 이상의 데이터셋의 중심위치가 같은 지를 검정합니다.

데이터가 먼저 가우시안 분포 (정규분포)를 따르는지 확인하고

따르지 않으면 비모수 통계학의 방법을 사용하면 되는데,

정규성 확인에 사용할 수 있는 테스트의 종류들에는 대표적으로 다섯 개가 있습니다.

그래프를 이용한 방법으로는 히스토그램, Q-Q plot 이 있고

통계적 검정을 이용하는 방법으로는 Shapiro-Wilk test, D'Agostino's K^2 test, Anderson-Darling test 가 있습니다.

샤피로 윌크 검정을 귀무가설은 정규분포를 따른다는 것이고

유의확률이 0.05 보다 낮으면 귀무가설을 기각하면 됩니다.

모수검정 설명	모수검정	각 모수검정에 대응하는 비모수검정
"두 집단의 관측치가 같은 분포? median 동일?" 두 개 그룹간의 평균에 차이가 나는지를 보는 검정. 샘플 각각의 평균값을 찾아서 그 두 개의 평균 값을 비교하는 것.	two sample t-test	만-위트니 U 검정 Mann-Whitney U test (= 만-위트니-윌콕슨 검정 Mann-Whitney-Wilcoxon test = 윌콕슨 순위 합 검정 Wilcoxon rank sum test)
"두 집단의 짝지어진 관측치가 같은 분포?" 데이터셋 2개의 차이 값들의 평균을 검정. 차이 (difference) 데이터셋을 one sample t-test 하는 것과 동일함.	paired sample t-test	윌콕슨 부호 순위 검정 Wilcoxon signed-rank test
"셋 이상 집단의 관측치가 같은 분포?" 분산을 이용해서 평균을 추론. 세 집단 이상 비교. 출력 변수에 대한 1개의 입력 변수의 3개 이상의 그룹의 평균들	one way ANOVA	크루스칼-왈리스 검정 Kruskal-Wallis H test (=순위 기반 일원분산분석 one-way ANOVA on ranks)
"셋 이상 집단의 관측치가 같은 분포?" 출력 변수에 대한 2 개의 입력 변수의 3개 이상의 그룹들의 상호간의 영향 검정	two way ANOVA	프리드먼 테스트 Friedman test

윌콕슨 부호 순위 검정 Wilcoxon signed-rank test

- 부호와 상대적 크기를 고려해 **중앙값을 검정함.**

- 분포의 연속성, 독립성, 대칭성을 가정함.

런 검정 Run Test

- **각 표본이 서로 독립적인지 검정.** 패턴/경향 없이 랜덤한지 검정.

* Run : 표본의 부호가 바뀔 때까지의 묶음. 런의 수가 상한치/하한치 범위를 벗어나면 H0 를 기각.

크루스칼-왈리스 H 검정

- k 개의 표본이 서로 다른 모집단에서 나왔는가를 검정

- 모집단에 대한 정규성, 등분산성의 가정 성립 여부에 대한 확신이 없음

- 측정치: 표본 내외 모두 독립, 연속적

- 귀무가설: k 개의 모집단의 중위수가 모두 같다.

프리드먼 테스트

- 측정값들이 동일한 모집단에서 나왔는가 검정

- 그룹 간 차이를 검정하는 것이 아니라, 레코드 간 차이가 없다는 귀무가설에 대해 검정

- 그룹과 레코드의 교호작용 (interaction effect)은 없다는 전제가 있음

- 자료를 순위로 표현한 후, 교차표를 작성하고 순위합을 구한다.

집단 1개	Run test, Kolmogorov-Smirnov 검정, 부호검정
집단 2개	Kolmogorov-Smirnov 검정, 윌콕슨 순위 합 검정(=만-위트니 U 검정), 윌콕슨 부호 순위 검정
집단 3개	크루스칼-왈리스 검정 (입력변수 1개), 프리드먼 테스트 (입력변수2개)

2. Normality Check 정규성 확인

먼저, 데이터가 정규분포를 따르는지 확인해야 합니다.
확인하는 방법에는 여러가지가 있는데,
선형적으로 나타내기 때문에 객관적으로 볼 수 있는 방법으로 Quantile-Quantile plot 이 있습니다.

확인해본 결과 정규 분포를 따르지 않는다고 나왔는데
모수 통계학을 사용하고 싶다면 정규분포를 따르도록 데이터를 바꿔줄 수도 있습니다.

정규 분포를 따르지 않는 데이터를 정규 분포를 따르도록 변환하는 방법에는
1) 샘플 크기를 증가 시키거나,
2) 데이터에 소수가 있으면 소수점 아래 자릿수를 늘리거나,
3) 아웃라이어를 제거하거나,
4) 롱테일을 제거하거나,
5) 멱변환(Power transform)
6) 로그 변환(log transformation)(로그 취하기)
7) 박스-칵스 Box-Cox 변환
등이 있습니다.

정규 분포를 따르지 않는 데이터를 비모수 통계학을 사용하여 가설검정을 수행할 때는
순위 (rank) 를 주로 활용할 수 있습니다.

3. Rank Correlation 순위 상관 계수

상관계수	데이터의 정규성			관계
Pearson 피어슨 상관계수	정규분포를 따르는 데이터	연속형의 두 변수		선형적
Spearman 스피어만 상관계수	정규분포를 벗어나는 데이터	연속형, 이산형	순위 rank 사용	비선형적. 단조적
Kendall's Tau 켄달의 타우 상관계수	정규분포를 벗어나는 데이터		순위 rank 사용	

피어슨 상관계수는 두 변수의 공분산을 표준편차의 곱으로 나눈 값으로 비선형 관계는 측정하지 못합니다.

스피어만 상관계수는 이산형 변수 간의 상관계수도 구할 수 있습니다.

4. 비모수 검정 (1) Mann-Whitney U Test

아래에 예시로 사용한 데이터셋 A 의 평균은 6 이고 데이터셋 B 의 평균은 6.25 입니다. 이 정도의 차이를 Mann-Whitney U Test 에서는 두 그룹의 평균에 차이가 없다고 할지 있다고 할지 검정해보겠습니다.

Mann-Whitney U Test 의 귀무가설 H0 는 두 그룹이 차이가 없다는 것이고
대립가설 H1 는 두 그룹에 차이가 있다는 것입니다.

먼저, Mann-Whitney 에서 사용하는 U 값을 구합니다.

U 값을 구하려면 각 데이터셋을 정렬하고 순위를 구한 뒤,
아래 테이블의 우측 2 개 열에 설명해둔 작업을 수행하여 각 셋 별 순위의 합을 구합니다.

data A 정렬	data B 정렬	data A 순위	data B 순위
3 (1위)		1위	
5 (2위)	5 (2위)	2위,3위,4위의 평균—> 3위	2위,3위,4위의 평균 —> 3위
	5 (2위)		2위,3위,4위의 평균 —> 3위
6 (5위)		5위	

7 (6위)	7 (6위)	6위,7위의 평균 —> 6.5위	6위,7위의 평균 —> 6.5위
	8 (8위)		8위
9 (9위)		9위	
데이터셋 A의 원소의 개수는 5개 (na 라고 하겠음)	데이터셋B의 원소의 개수는 4개 (nb 라고 하겠음)	순위의 합 = 1 + 3 + 5 + 6.5 + 9 = 24.5 (Ra라 하겠음)	순위의 합 = 3 + 3 + 6.5 + 8 = 20.5 (Rb라 하겠음)

데이터셋 A 의 U 값 = na*nb + (na*(na+1))/2 - Ra = 5*4 +5*6/2 - 24.5 = 10.5
데이터셋 B 의 U 값 = na*nb + (nb*(nb+1))/2 - Rb = 5*4 +4*5/2 - 20.5 = 9.5
두 U 값 중 더 작은 값이 최종 U 값이 됩니다. 즉, 여기서는 9.5 입니다.
그러면 이제 이 값을 Critical Values of the Mann-Whitney U 테이블에서 찾아봅니다.
n1 이 na 이고 n2 가 nb 로 보면 됩니다. 만나는 지점을 찾으면 됩니다.
아래 그림 테이블 왼편의 a 는 유의수준을 뜻합니다.

즉 위 예시에서는 5 개와 4 개였으므로,
n1= 5 이고 n2 = 4 인 지점을 찾으면
유의수준 0.05 일때는 1 이고 유의수준 0.01 일 때는 0 이란 값이
귀무가설을 기각할 수 있는 임계치라는 뜻입니다.

유의수준 0.05 일때 1 보다 작으면 귀무가설을 기각할 수 있는데 여기서는 9.5 였으므로
귀무 가설을 기각할 수 없습니다.
즉, 두 그룹은 평균에 차이가 없다고 보았습니다.

Critical Values of the Mann-Whitney U
(Two-Tailed Testing)

n_2	α	n_1: 3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
3	.05	--	0	0	1	1	2	2	3	3	4	4	5	5	6	6	7	7	8
	.01	--	0	0	0	0	0	0	0	0	1	1	1	2	2	2	2	3	3
4	.05	--	0	1	2	3	4	4	5	6	7	8	9	10	11	11	12	13	14
	.01	--	--	0	0	0	1	1	2	2	3	3	4	5	5	6	6	7	8
5	.05	0	1	2	3	5	6	7	8	9	11	12	13	14	15	17	18	19	20
	.01	--	--	0	1	1	2	3	4	5	6	7	7	8	9	10	11	12	13
6	.05	1	2	3	5	6	8	10	11	13	14	16	17	19	21	22	24	25	27
	.01	--	0	1	2	3	4	5	6	7	9	10	11	12	13	15	16	17	18
7	.05	1	3	5	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34
	.01	--	0	1	3	4	6	7	9	10	12	13	15	16	18	19	21	22	24
8	.05	2	4	6	8	10	13	15	17	19	22	24	26	29	31	34	36	38	41
	.01	--	1	2	4	6	7	9	11	13	15	17	18	20	22	24	26	28	30
9	.05	2	4	7	10	12	15	17	20	23	26	28	31	34	37	39	42	45	48
	.01	0	1	3	5	7	9	11	13	16	18	20	22	24	27	29	31	33	36
10	.05	3	5	8	11	14	17	20	23	26	29	33	36	39	42	45	48	52	55
	.01	0	2	4	6	9	11	13	16	18	21	24	26	29	31	34	37	39	42
11	.05	3	6	9	13	16	19	23	26	30	33	37	40	44	47	51	55	58	62
	.01	0	2	5	7	10	13	16	18	21	24	27	30	33	36	39	42	45	48

출처 : https://ocw.umb.edu/psychology/psych-270/other-materials/RelativeResourceManager.pdf

4. 비모수 검정 (2) Wilcoxon (Matched Pairs) Signed-Rank Test

Wilcoxon signed-rank test 는 Mann-Whitney U test 와 유사합니다.

다만, paired t-test 와 같이, 두 개 데이터셋의 차이를 구하는 것이고,
여기서는 차이의 크기를 부호가 있는 순위로 매겨서 사용합니다.

그 후 양의 부호인 순위들의 합을 구하고, 음의 부호인 순위들의 합을 구한 후,
둘 중에 작은 값을 최종 검정 통계량 값으로 취합니다.

검정 통계량 값을 구했으면 Mann-Whitney U test 에서와 같이 기각역 값들이 적힌 테이블을 확인합니다.
만약, 유의수준이 0.05 이면, a = 0.025 인 기각역과, a = 0.975 인 기각역을 찾아서,
위에서 구했던 최종 검정 통계량 값과 비교합니다.

기각역에 속한다면 귀무가설을 기각합니다.
Wilcoxon signed-rank test 에서도 H0 는 '차이가 없다'이고, 대립가설 H1 은 '차이가 있다'입니다.

즉, 귀무가설을 기각하면 두 개 데이터셋의 차이가 있다는 결론을 내리게 됩니다.

4. 비모수 검정 (3) Kruskal-Wallis H test

Kruskal-Wallis H test 는 2 개 표본에 사용하는 Mann-Whitney U Test 를 다변량 분석으로 확장시킨 것입니다.

여기서의 귀무가설은 'k 개의 모집단의 중위수가 모두 같다'는 것이고
대립가설은 'k 개의 모집단의 중위수가 모두 같지는 않다'입니다.

Kruskal-Wallis H test 는 각 표본을 정렬하여 순위를 매긴 후, 순위를 합하여
각각의 데이터셋들에 대한 순위합을 구합니다.

그리고 Kruskal-Wallis H test 의 검정 통계량인 H 값을 구하는 계산 공식을 통해 H 값을 구합니다.
(그런데 이 공식에서 분자에 12 라는 상수는 왜 12 를 쓰는 것일까요?)

그 후 그 H 값이 기각역에 포함된다면 귀무가설을 기각합니다.

4. 비모수 검정 (4) Friedman test

Friedman test 도 다변량 분석입니다.

이 검정은 교호작용이 없다는 가정을 전제로 하는데,

이 부분은 Friedman test 의 한계점이기도 합니다.

프리드먼 검정은 세 개 이상 그룹을 비교할 때 사용합니다.