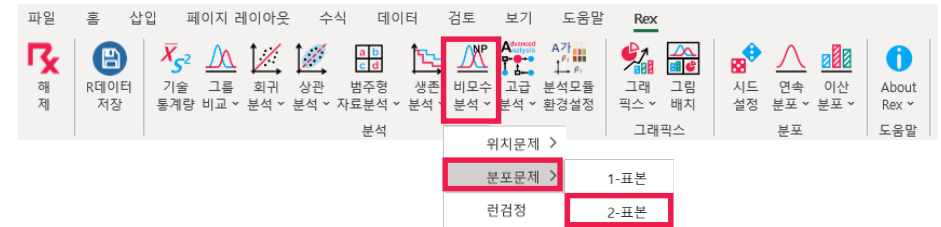


# 비모수 분포문제 2-표본

## 메뉴 호출하기

- Rex > 비모수분석 > 분포문제 > 2-표본



비모수 분석에서 2-표본 분포문제는 두 모집단이 서로 동일한 분포인지를 알아보기 위한 비모수 검정 방법입니다. 2-표본 콜모고로프-스미르노프 검정이라고도 합니다.

• 변수설정 탭

비모수 이표본분포문제

변수설정 분석옵션

데이터

전체변수

id  
bweight  
lowbw  
gestwks  
preterm  
matage  
hyp  
sex

1 종속변수(필수)

2 집단변수(필수)

도움말 재설정 확인 취소

메뉴 요소	설명
① 종속변수	분포를 비교하고자 하는 변수를 전체변수로부터 선택합니다. 반드시 1개 이상의 양적 변수가 선택되어야 합니다.
② 집단변수	집단변수에 해당하는 변수를 전체변수로부터 선택합니다. 반드시 1개의 변수가 선택되어야 하며, 2개의 수준으로 구성되어야 합니다.

• 분석옵션 탭

비모수 이표본분포문제

변수설정 분석옵션

① 대립가설

양측검정     좌측단측검정     우측단측검정

② 그래프

히스토그램     상자그래프

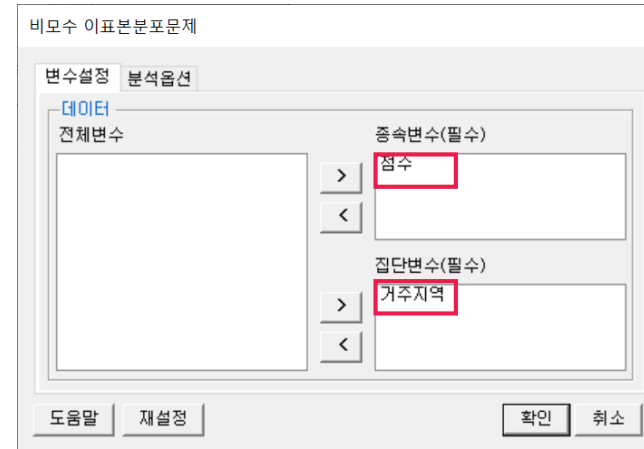
도움말    재설정    확인    취소

메뉴 요소	설명
① 대립가설	두 집단의 분포 $F_1(x)$ 와 $F_2(x)$ 에 대해 대립가설에 맞는 검정방법으로 다음 3가지 중 1개를 선택합니다. <ul style="list-style-type: none"> <li>양측검정 (Default) : <math>F_1(x) - F_2(x) \neq 0</math></li> <li>좌측단측검정 : <math>F_1(x) - F_2(x) &lt; 0</math></li> <li>우측단측검정 : <math>F_1(x) - F_2(x) &gt; 0</math></li> </ul>
② 그래프 > 히스토그램	집단별 종속변수의 히스토그램을 출력합니다.
③ 그래프 > 상자그래프	집단별 종속변수의 상자그래프를 출력합니다.

• 예제

중학생의 정신건강이 거주지역(도시 - 1/시골 - 2)에 따라 차이가 있는지 조사한 자료이다. 두 분포가 동일한지 검정하라. (대립가설은 사전정보에 따라 검정력을 높이기 위해 우측단측검정으로 실시한다.) [데이터 : manual\_data.xlsx > 정신건강]

1. [변수설정] > [종속변수] > '점수' 지정
2. [변수설정] > [집단변수] > '거주지역' 지정
3. [분석옵션] > [대립가설] > [우측단측검정] 선택
4. [분석옵션] > [그래프] > [히스토그램] 체크
5. [분석옵션] > [그래프] > [상자그래프] 체크
6. [확인]



• 예제 - 결과창

## Two Sample Kolmogorov-Smirnov test

### Data Structure

No. of total observations	28
No. of used observations	28
No. of used variable(s)	2

### Variable List

Category	Variable	N	N.valid	(% .valid)	N.miss	(% .miss)
Dependent variable	점수	28	28	(100.00%)	0	(0.00%)
Group variable	거주지역	28	28	(100.00%)	0	(0.00%)

### Analysis Description

Test direction	Right-sided
H1	The distribution function for Group 2 is greater than that for Group 1 for at least one observation.

### Descriptive Statistics

	# of obs.	Mean	SD	Median	Q1	Q3
Group 1	16	31.9375	8.69842	27.0	25.75	38.75
Group 2	12	38.2500	8.86387	39.5	31.75	44.00

### Results of Two Sample Kolmogorov-Smirnov Test

	KS	P-value
점수	0.4792	0.0429

[Data Structure]

- No. of total observations : 전체 개체 수
- No. of used observations : 분석에 사용된 개체 수
- No. of used variable(s) : 분석에 사용된 변수의 수

[Variable List]

- 분석에 사용된 개별 변수들의 유효개수(N.valid), 유효퍼센트(% .valid), 결측수(N.miss), 결측퍼센트(% .miss)가 출력됩니다.

[Analysis Description]

- Test direction : [분석옵션] > [대립가설] 지정 항목 (우측단측검정)
- H1 : [분석옵션] > [대립가설] 지정 항목에 따른 대립가설 (적어도 한 관측치에 대해서는 그룹 1의 분포보다 그룹 2의 분포가 더 크다)

[Descriptive Statistics]

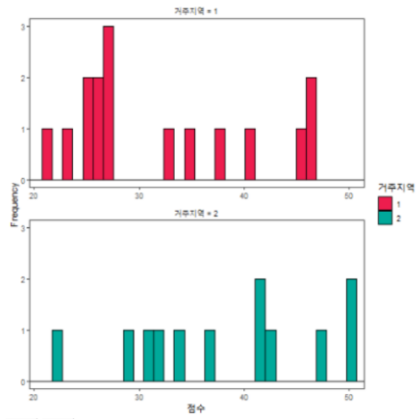
- 그룹 1(도시)의 중학생의 정신건강의 유효개체수는 16개이며, 평균은 31.9375, 표준편차는 8.69842, 중앙값은 27.0, 사분위수 범위는 (25.75, 38.75)입니다.
- 그룹 2(시골)의 중학생의 정신건강의 유효개체수는 12개이며, 평균은 38.2500, 표준편차는 8.86387, 중앙값은 39.5, 사분위수 범위는 (31.75, 44.00)입니다.

[Results of Two Sample Kolmogorov-Smirnov Test]

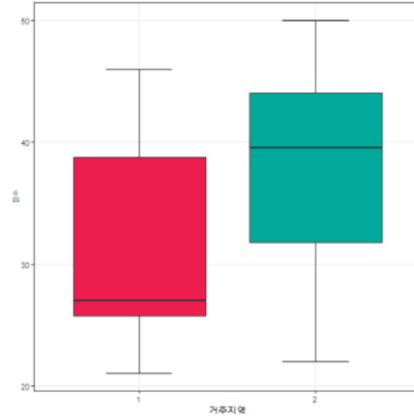
- $p=0.0429$ 로 유의수준 0.05하에서 영가설을 기각되어, 거주지역에 따라 남자 중학생의 정신건강 점수는 거주지역에 따라 다르고, 특히 도시지역 남자 중학생의 정신건강 점수의 분포가 시골지역 남자 중학생의 분포보다 우측에 위치한다고 볼 수 있습니다.

예제 - 결과창

Histogram



Box plot



[Histogram]

- 거주지역에 따른 중학생의 건강점수에 대한 히스토그램이 출력됩니다.

[Box Plot]

- 거주지역에 따른 중학생의 건강점수에 대한 박스그림이 출력됩니다.

[Used R packages]

- 본 분석에 사용된 R 패키지들이 나열되어 있습니다.

Used R Packages

- All results were written with basic functions of R.