

정규분포 가정

다변량 분석에서 왜도와 첨도에 의한 정규분포 기준

West et al(1995)의 정규분포 기준은 |왜도| < 3, |첨도| < 8

Hong et al(2003)의 정규분포 기준은 |왜도| < 2, |첨도| < 4

#01. 작업준비

패키지 참조

데이터 가져오기 + 회귀분석

왜도

첨도

Fisher = True

Fisher = False

정규분포 가정

최근 잔차분석을 통한 정규성 검정보다 더 많이 인용되고 있는 기준

탐색적 데이터 분석 과정에서 모든 독립변수에 대해 왜도와 첨도를 검사한다.

다변량 분석에서 왜도와 첨도에 의한 정규분포 기준

West et al(1995)의 정규분포 기준은 |왜도| < 3, |첨도| < 8

West, S. G., Finch, J. F., & Curran, P. J. (1995). Structural equation models with nonnormal variables: Problems and remedies. In R. H. Hoyle (Ed.), Structural equation modeling: Concepts, issues, and applications (p. 56–75). Sage Publications, Inc.

Hong et al(2003)의 정규분포 기준은 |왜도| < 2, |첨도| < 4

Hong, S., Malik, M. L., & Lee, M.-K. (2003). Testing configural, metric, scalar, and latent mean invariance across genders in sociotropy and autonomy using a non-Western sample. Educational and Psychological Measurement, 63(4), 636–654.

#01. 작업준비

패키지 참조

```
from pandas import read_excel
```

정규분포 가정

다변량 분석에서 왜도와 첨도에 의한
정규분포 기준

West et al(1995)의 정규분포 기준은 |왜도| < 3, |첨도| < 8

Hong et al(2003)의 정규분포 기준은 |왜도| < 2, |첨도| < 4

#01. 작업준비

패키지 참조

데이터 가져오기 + 회귀분석

왜도

첨도

Fisher = True

Fisher = False

```
from scipy.stats import skew, kurtosis
```

데이터 가져오기 + 회귀분석

```
df = read_excel("https://data.hossam.kr/E04/cars.xlsx")
df
```

	speed	dist
0	4	2
1	4	10
2	7	4
3	7	22
4	8	16
5	9	10
6	10	18
7	10	26
8	10	34
9	11	17
10	11	28
11	12	14

정규분포 가정

다변량 분석에서 왜도와 첨도에 의한
정규분포 기준

West et al(1995)의 정규분포 기준은 |왜도| < 3, |첨도| < 8

Hong et al(2003)의 정규분포 기준은 |왜도| < 2, |첨도| < 4

#01. 작업준비

패키지 참조

데이터 가져오기 + 회귀분석

왜도

첨도

Fisher = True

Fisher = False

	speed	dist
12	12	20
13	12	24
14	12	28
15	13	26
16	13	34
17	13	34
18	13	46
19	14	26
20	14	36
21	14	60
22	14	80
23	15	20
24	15	26
25	15	54
26	16	32
27	16	40
28	17	32
29	17	40

정규분포 가정

다변량 분석에서 왜도와 첨도에 의한
정규분포 기준

West et al(1995)의 정규분포 기준은 |왜도| < 3, |첨도| < 8

Hong et al(2003)의 정규분포 기준은 |왜도| < 2, |첨도| < 4

#01. 작업준비

패키지 참조

데이터 가져오기 + 회귀분석

왜도

첨도

Fisher = True

Fisher = False

	speed	dist
30	17	50
31	18	42
32	18	56
33	18	76
34	18	84
35	19	36
36	19	46
37	19	68
38	20	32
39	20	48
40	20	52
41	20	56
42	20	64
43	22	66
44	23	54
45	24	70
46	24	92
47	24	93

정규분포 가정

다변량 분석에서 왜도와 첨도에 의한 정규분포 기준

West et al(1995)의 정규분포 기준은 $|\text{왜도}| < 3$, $|\text{첨도}| < 8$

Hong et al(2003)의 정규분포 기준은 $|\text{왜도}| < 2$, $|\text{첨도}| < 4$

#01. 작업준비

패키지 참조

데이터 가져오기 + 회귀분석

왜도

첨도

Fisher = True

Fisher = False

	speed	dist
48	24	120
49	25	85

왜도

분포의 비대칭도.

구분	내용
정규분포	왜도 = 0
왼쪽으로 치우침	왜도 > 0
오른쪽으로 치우침	왜도 < 0

```
skew(df['speed'])
```

```
-0.11395477012828319
```

첨도

확률분포의 뾰족한 정도.

Fisher = True

- 첨도 기준이 Fisher (normal ==> 0.0) 이다.
- 정규분포의 첨도 = 0이다.

정규분포 가정

다변량 분석에서 왜도와 첨도에 의한 정규분포 기준

West et al(1995)의 정규분포 기준은 $|\text{왜도}| < 3$, $|\text{첨도}| < 8$

Hong et al(2003)의 정규분포 기준은 $|\text{왜도}| < 2$, $|\text{첨도}| < 4$

#01. 작업준비

패키지 참조

데이터 가져오기 + 회귀분석

왜도

첨도

Fisher = True

Fisher = False

구분	내용
정규분포	첨도 = 0
위로 뽀족함	첨도 > 0
아래로 뽀족함	첨도 < 0

Fisher = False

- 첨도 기준이 Pearson (normal ==> 3.0)
- 정규분포의 첨도 = 3이다.

구분	내용
정규분포	첨도 = 3
위로 뽀족함	첨도 > 3
아래로 뽀족함	첨도 < 3

```
kurtosis(df['speed'], fisher=True)
```

```
-0.5771474239437371
```