

서강대학교 정보통신 대학원 통계기반 데이터 분석 10

2019년 5월 정화민 교수

데이터 신뢰성 검정

- 신뢰도(reliability):

- : 측정된 결과치의 일관성, 정확성, 의존가능성, 안정성, 예측가능성과 관련된 개념으로, 동일한 개념에 대해 측정을 되풀이했을 때 동일한 측정값을 얻을 가능성을 말함.
- : 하위요인들의 변수로 실시(즉 설문 측정값으로 실시. 만족에 대한 질문이 A,B,C 세 개라면 이 세가지 변수로 실시).
- : 신뢰도 분석의 결과는 Cronbach alpha와 같은 신뢰도척도를 계산한 값을 가지고 판단하고 신뢰도를 저해하는 항목이 있다면 제거해야 함.
- : 해석은 "데이터의 신뢰성이 있다"라고 해석

신뢰도 지수 Chronbach's Alpha:

- Chronbach's Alpha값은 내적일관성을 측정하기 위한 계수로 사용됨
- 내적일관성 : 동일한 개념을 측정하기 위해 여러 개의 항목을 이용하는 경우, 신뢰도를 저해하는 항목을 찾아 측정도구에서 제외시킴으로써 측정도구의 신뢰도를 높이기 위한 방법.

신뢰도 Chronbach's Alpha 값?

- 사회과학에서는 통상적으로 크론바하 알파 값이 0.6이상이면 데이터의 신뢰도가 있다고 판단한다.
- Fleiss(1986)는 >.75면 훌륭하다고 함
- Nunnally(1978)는 .80이상이면 기초연구에서는 충분

R 에서 Iris 데이터를 활용한 신뢰성 검정

```
1 install.packages("psych")
2 library(psych)
3 alpha(iris[,1:4])
4 |
```

```
Call: alpha(x = iris[, 1:4])
 raw alpha std.alpha G6(smc) average r S/N ase mean sd median r
     0.71
                     0.9
                             0.29 1.6 0.019 3.5 0.78
             0.62
                                                      0.35
                   95% confidence boundaries
lower alpha upper
0.67 0.71 0.75
Reliability if an item is dropped:
           raw alpha std.alpha G6(smc) average r S/N alpha se var.r med.r
Sepal.Length
                        0.15 0.72 0.056 0.18 0.0337 0.6176 -0.37
Sepal.Width
               0.88
                        0.96 0.96 0.884 22.90 0.0067 0.0054 0.87
Petal.Length
                       0.49
Petal.Width
               0.47
                        0.27
                              0.77
                                      0.109 0.37 0.0401 0.4610 -0.12
 Item statistics
            n raw.r std.r r.cor r.drop mean sd
Sepal.Length 150 0.94 0.940 0.93 0.89 5.8 0.83
Sepal.Width 150 -0.22 0.032 -0.15 -0.35 3.1 0.44
Petal.Length 150 0.97 0.880 0.91 0.86 3.8 1.77
Petal.Width 150 0.95 0.883 0.90 0.92 1.2 0.76
```

Raw alpha	0.71	크론바하알파값이 0.71
Std .alpha	0.62	표준화된 항목의 크론바하알파값

IRIS 데이터로 신뢰성분석 한 결과 크론바하 알 파값이 0.71로 나타나 그 값이 0.6 이상으로 데이터의 신뢰성이 있다고 할 수 있다.

- 항목이 drop(제거)되었을 때의 신뢰도 Sepal.Length를 제거하면 신뢰도가 0.45
- -만약 Sepal.Width를 제거하면 신뢰도는 0.88 이 된다는 얘기임
- * 문제: 신뢰도에서 가장 중요한 변수는?
- 전체 신뢰도를 저해하는 변수는?
- 그러면 여기서 신뢰도를 저해하는 변수를 제 거해야 하나?

- 각 조별 팀플 지도