

★ 변수의 분류

1. 수학적 개념(척도)에 의한 분류

1.1 명칭, 명목 척도(nominal scale) : 범주로만 의미

ex) 성별, 혈액형

1.2 순위척도(ordinal scale) : 명목 + 대소관계, 수학적 조작은 불가능,

ex) 교육정도(국졸이하/중졸/고졸/대졸이상), 사회경제수준(상,중,하)

1.3 간격척도(interval scale) : 가감 가능, 승제 불가능

ex) 온도

1.4 비척도(ratio scale) : 가감승제 가능

ex) 연령

범주형 변수(categorical variable) = nominal scale + ordinal scale

연속형 변수 (continuous variable) = interval scale + ratio scale

2. 인과관계에 따른 변수의 분류

독립변수(설명변수) independent var., explanatory var., covariate (원인)

종속변수(반응변수) dependent var., response var., (결과)

3. 전산입력 형식에 따른 변수의 분류

3.1. 숫자형 변수

3.2. 문자형 변수

3.3. 날짜형 변수

★ 변수의 종류에 따른 통계분석법

종속변수	독립변수	통계분석법
연속변수 (혈압)	명칭척도(2개 종류)	T검정, paired T검정
연속변수 (혈압)	범주형(3개 종류 이상)	분산분석 (ANOVA)
범주형(병발생여부)	범주형(투약여부)	카이제곱, 로지스틱 회귀
연속형(태아의 체중)	연속형(재태 임신기간)	회귀분석
연속형(출생시 체중)	연속형+범주형	공분산분석
생존기간 (연속형, >0)	연속형+범주형	생존분석
종속변수와 독립변수의 구분이 없는 경우		주성분분석, 요인분석, 군집분석

▶ 종속관계가 명확하지 않을 때 -> 상관분석

- 1) 정규분포 가정시 : 피어슨 상관계수
- 2) 정규분포 비가정시 : 비모수 기법, 순위를 이용한 상관

▶ multiple -> ex) multiple regression

; 여러 개의 독립변수들을 한 모형에서 고려할 때

▶ multivariate -> ex) multivariate linear models

; 여러 개의 설명변수들을 모형화 할 때

▶ 모수적 검정법

▶ 비모수적 검정법, 분포무관 검정법

- 1) 표본의 수가 작을 때 (모집단의 분포가정 곤란)
- 2) 모집단의 분포가 특정분포(ex. 정규분포)가 아닌 것이 확실할 때
 - > A. 비모수적 방법
 - B. 변수 변환 후 모수적 방법

★ 자료의 성격에 따른 분석 방법

자료의 성격	모수적 방법	비모수적 방법
종속변수가 범주형	카이제곱	Fisher's exact test McNemar test Cochran's Q
< 종속변수가 연속형 >		
두 개의 독립된 집단	t-test	Wilcoxon rank sum test Man-Whitney median test
세 개 이상의 집단	ANOVA	Kruskal Wallis test
제3 변수 영향 고려	2-way ANOVA	Friedman's 2-way ANOVA
상관분석	Pearson Corr.	Spearman's corr Kendal's tau Stuart's tau