**<기타 분석 용어 정리>**

* 다중 공선성의 정의

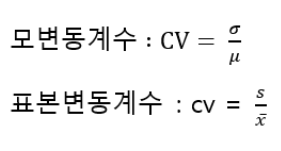
회귀 분석의 다중 공선성은 모형의 일부 예측 변수가 다른 예측 변수와 상관되어 있을 때 발생하는 조건이다. 중대한 다중 공선성은 회귀 계수의 분산을 증가시켜 불안정하고 해석하기 어렵게 만들기 때문에 문제가 된다.

* VIF(Variation Inflation Factor):분산 팽창 계수

다중공선성을 파악하기 위한 수치적 지표로 예측 변수들이 상관성이 있을 때 추정 회계 계수의 산포 크기를 측정하는 것이며, 산포가 커질 수 록 회귀 모형은 신뢰 할 수 없게 된다.

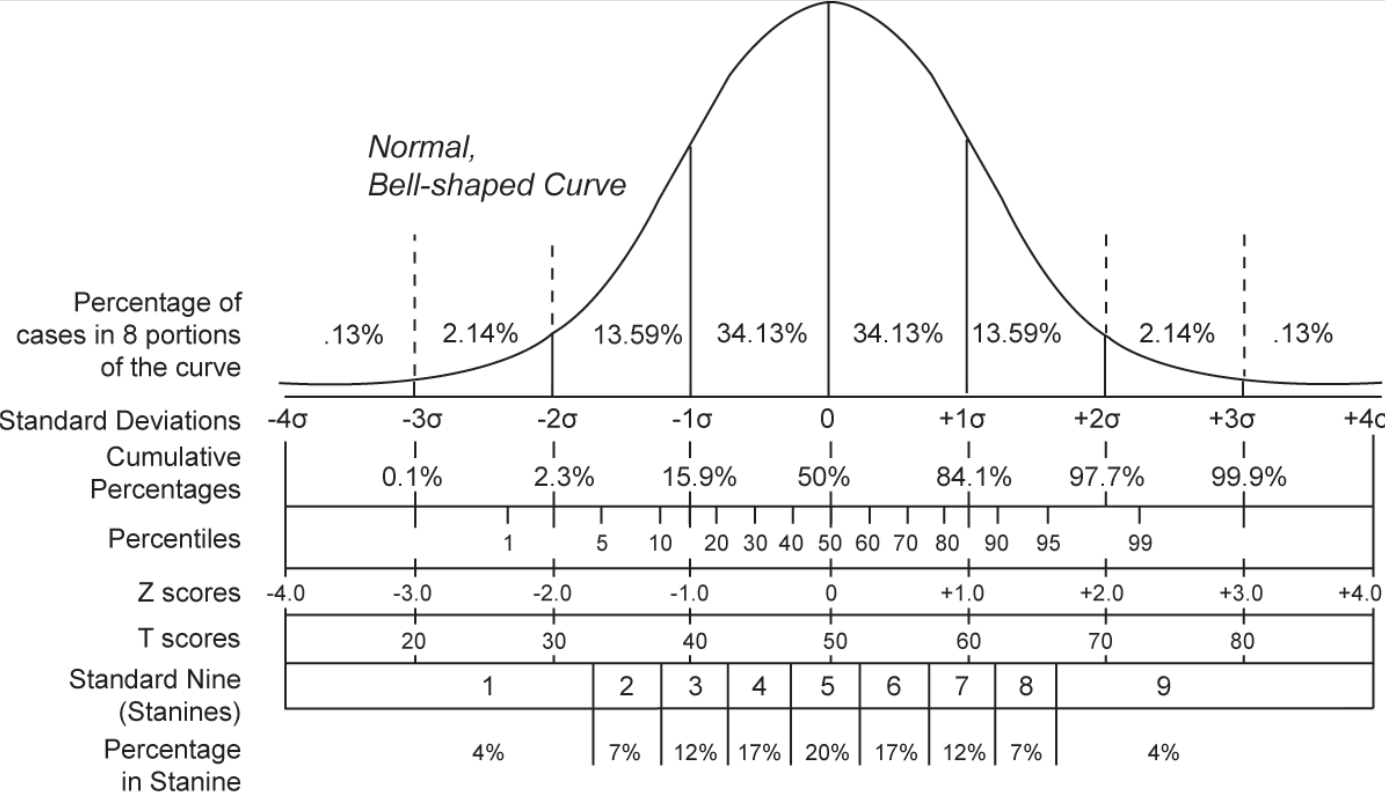
* 변동계수( coefficient of variation)

각 데이터 집단끼리 표준편차를 통해서 변동성을 나타내는 지표를 비교하려할 때 , 데이터 집단끼리 단위가 다르면 비교조차 할 수 없다. 이를 해결하기 위해서 변동계수를 사용한다. 변동계수는 표준편차를 평균으로 나누어 계산한다. 이를 통해 다른 데이터 관측치들로부터 서로 변동성에 대해 비교 할 수 있게된다.



* 표준 점수(Z score)

통계학적으로 정규분포를 만들고 개개의 경우가 표준편차상에 어떤 위치를 차지하는지를 보여주는 수치



* P-value