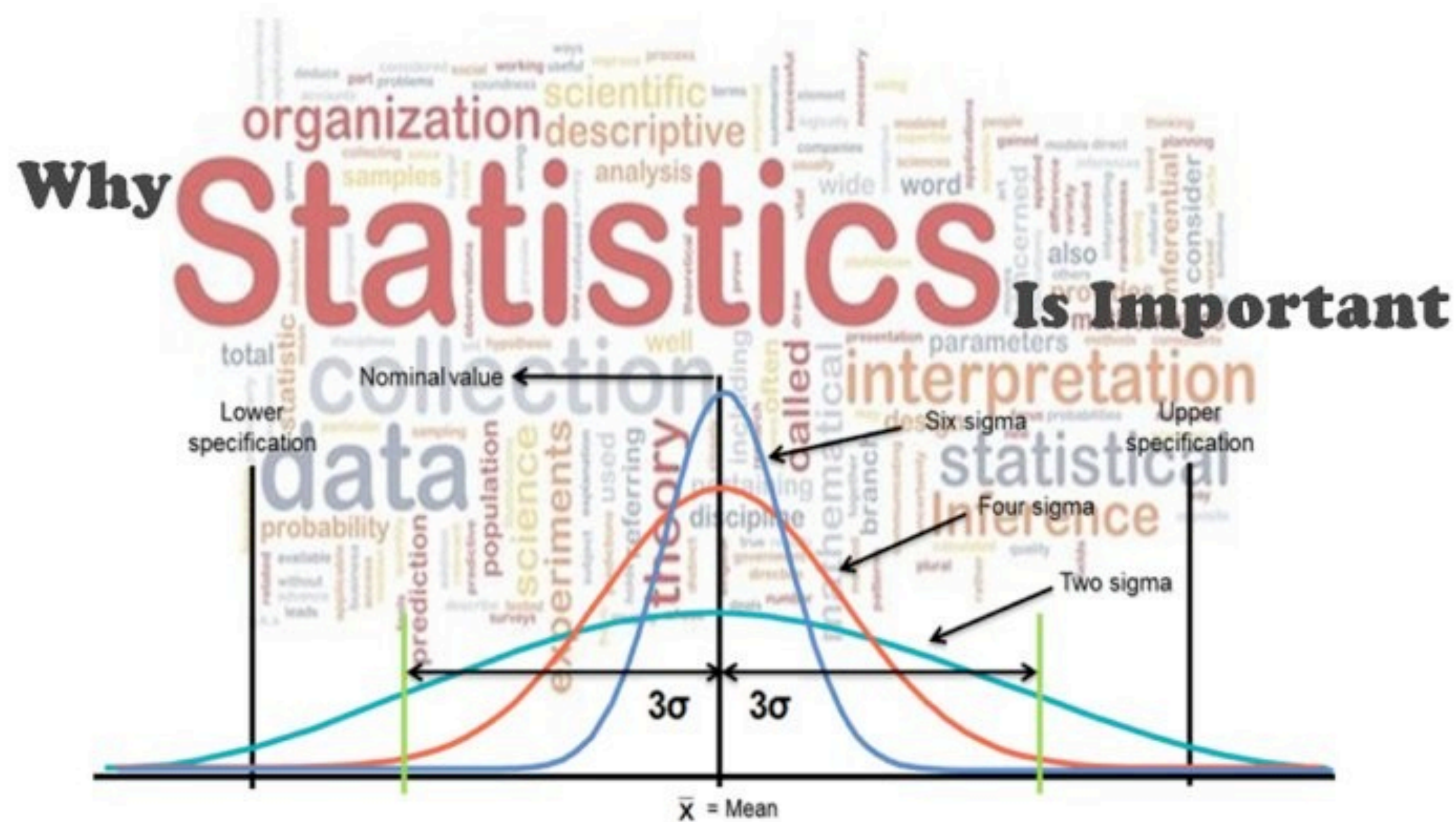


# BDA 7기 파문응 통계분석

20230916 7기 파문응

# 통계학



## 인구조사에서 생겨난 통계학

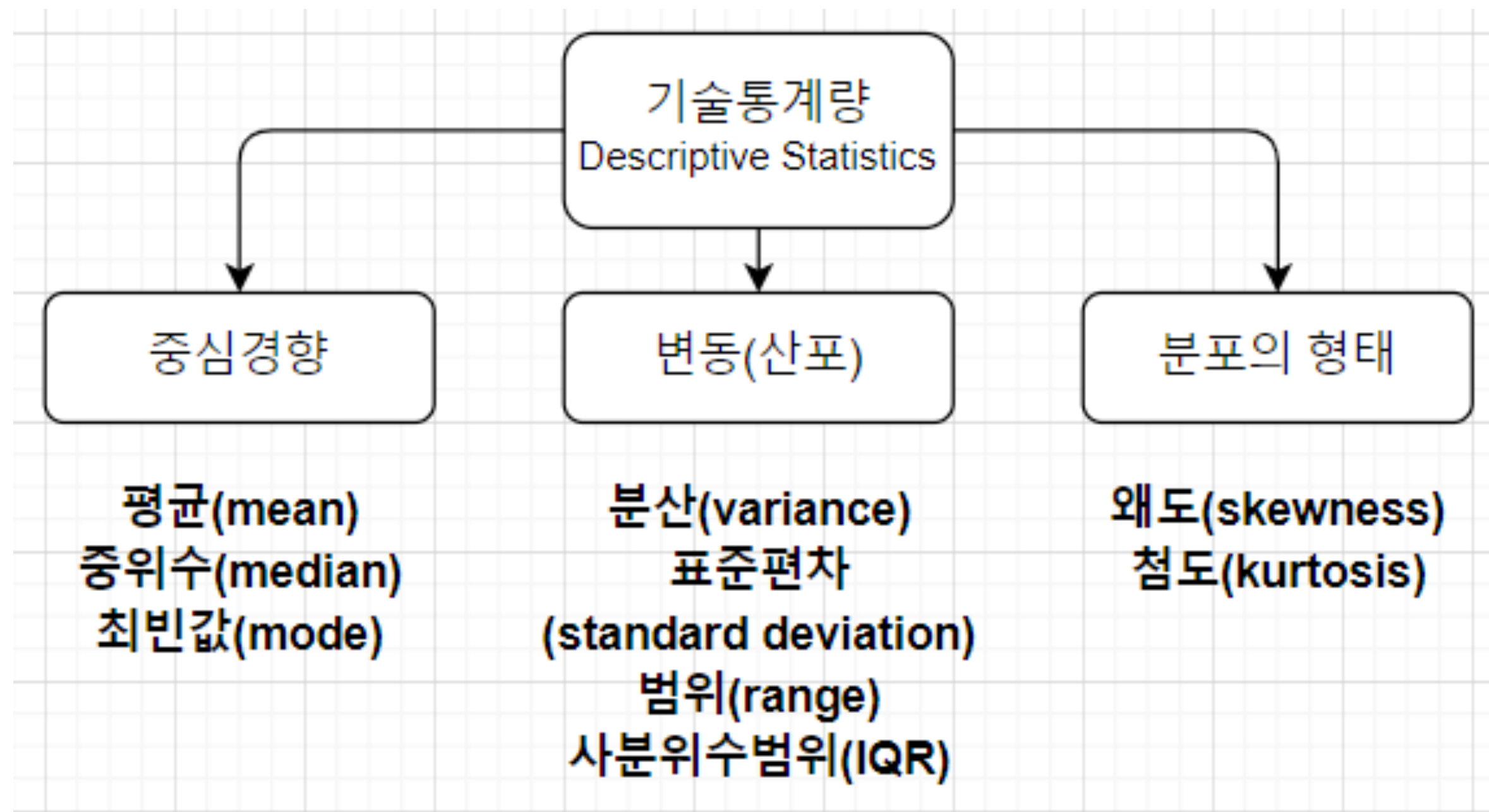
**독립변수 X**

**종속변수 Y**

## 귀무가설 H0

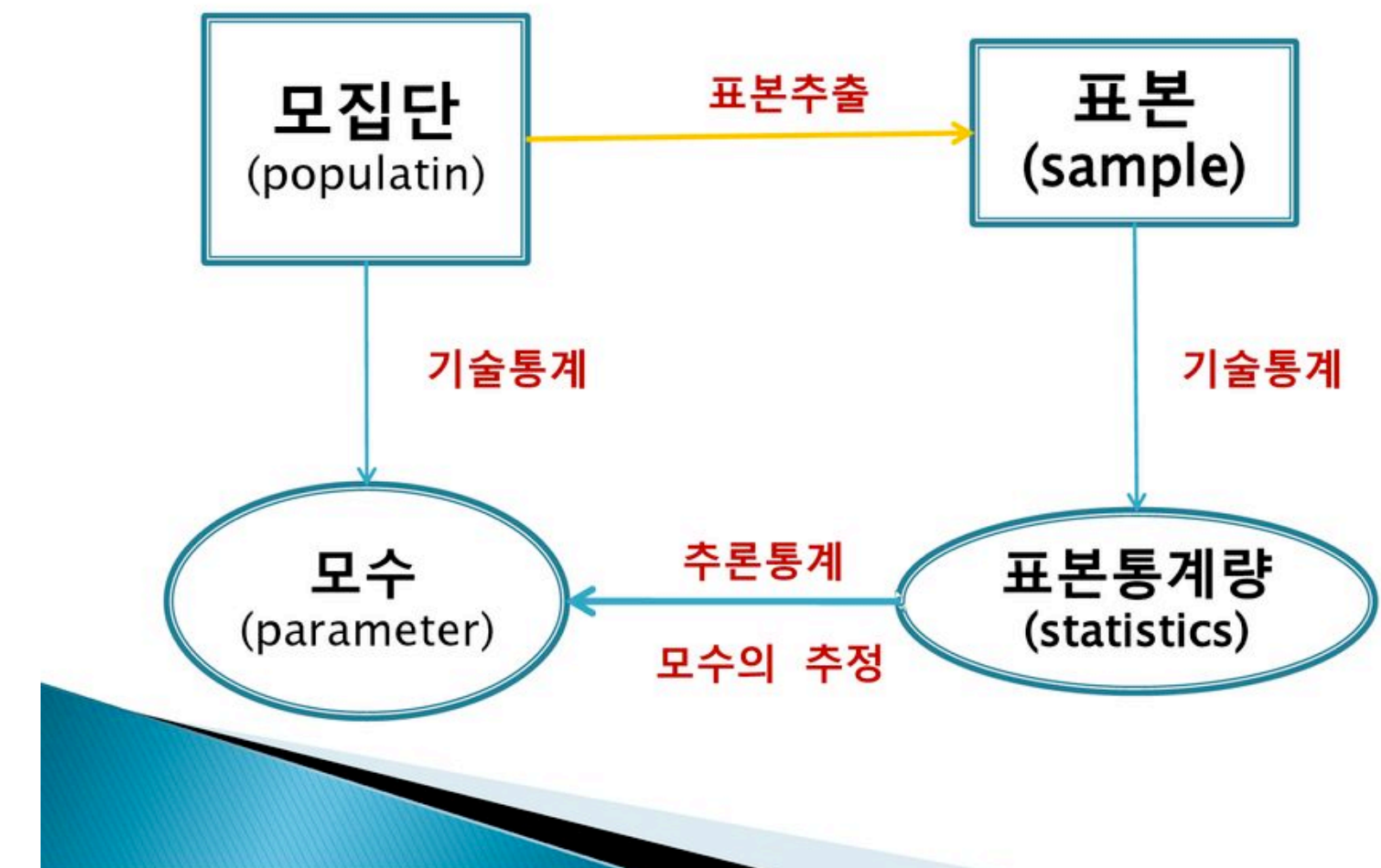
## 대립가설 H1

# 기술통계 추론통계



주어진 데이터의 특성을 사실에 근거하여 설명하고 묘사  
Descriptive = 서술하는, 사실에 근거한, 설명적인

## 추론통계의 기능



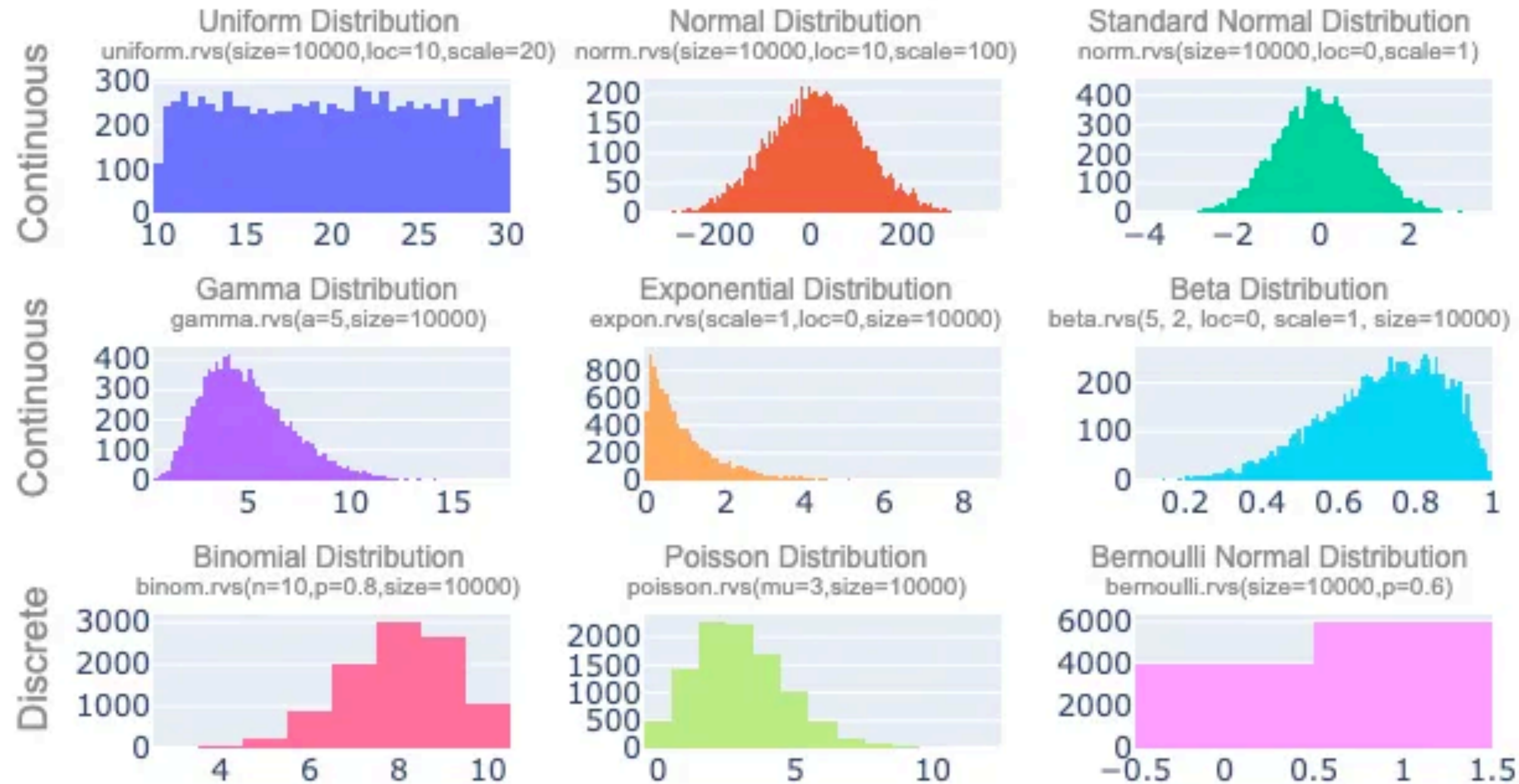
추론 통계는 표본 집단으로부터 모집단의 특성을 추론  
Inferential = 추리의, 추정에 의한



# 확률과 확률분포

## Statistical Distributions

Arrangement of values of a variable showing their frequency of occurrence



이산형 확률분포

연속형 확률분포

#30DayChartChallenge - statistics - 2021/04/09

Datasets created using scipy.stats

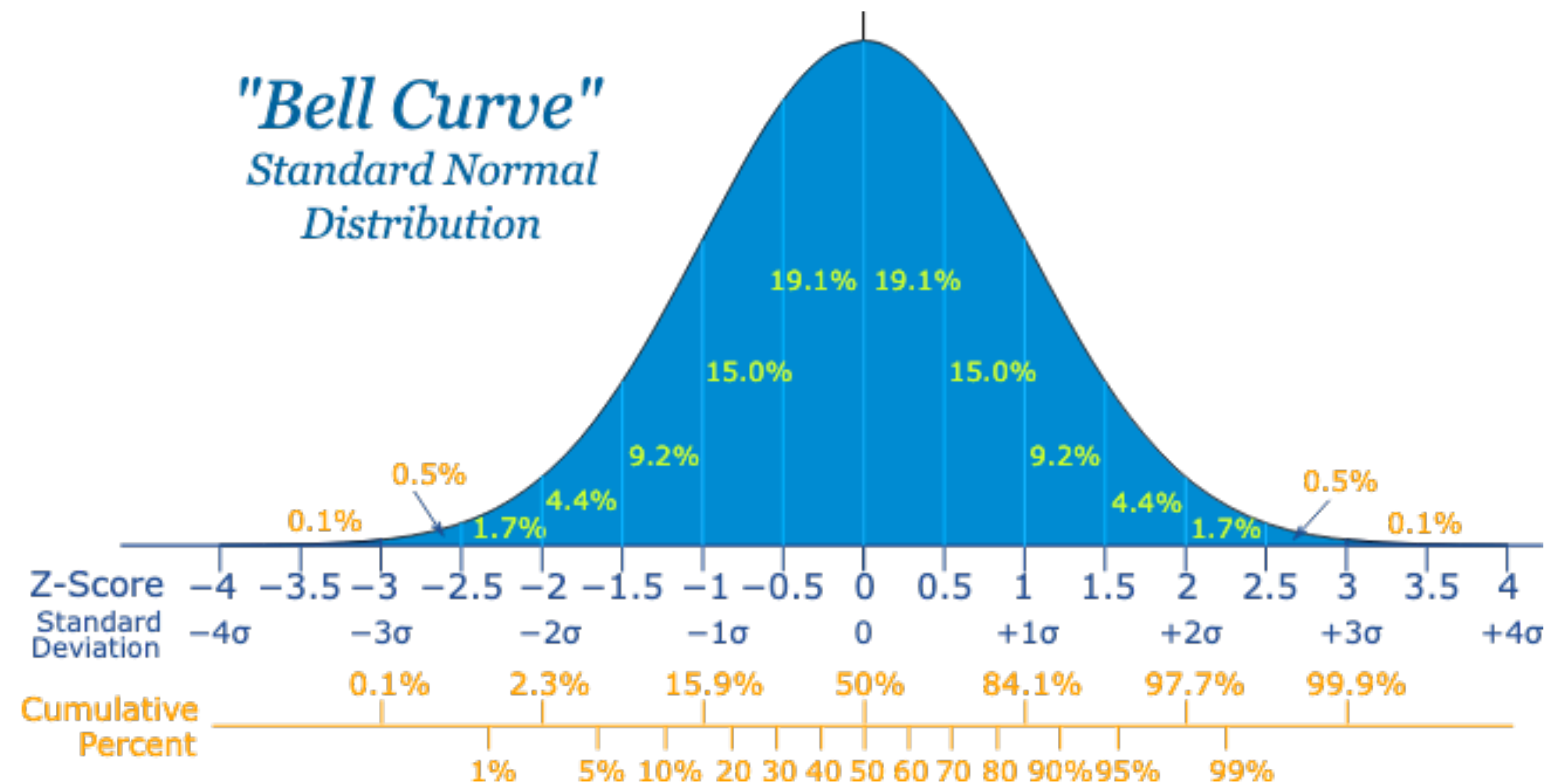
[twitter.com/vivekparasharr](https://twitter.com/vivekparasharr) | [github.com/vivekparasharr](https://github.com/vivekparasharr) | [vivekparasharr.medium.com](https://vivekparasharr.medium.com)

# 왜 정규분포인가?

자연현상의 대부분의 값은 거의 다 근사적으로 정규분포 및 19세기와 20세기 지나면서  
정규분포는 자연과학뿐만 아니라 다양한 사람들이 모인 사회, 정치, 경제 등 모든 분야 사용

모집단의 분포는 여러가지 모양을 가질 수 있다. 그런데 모집단의 분포가 정규분포를 가진다고 가정하면 여러가지 통계분석을 손쉽게 수행할 수 있다. 그리고 실제로 현실적인 모집단들은 정규분포와 유사한 모양을 가지는 경우들이 대부분이다. 따라서 모집단이 정규분포를 갖는다는 가정은 현실적인 많은 문제에서 타당성을 인정 받을 수 있다

정규분포는 표본분포에 매우 잘 적용되며, 우리가 사회적, 자연적 현상에서 보통 접하는 여러 통계치들의 분포가 정규분포와 비슷한 형태를 띠고 있다. 또한 이항분포, 포아송분포, 초기하분포와, 카이제곱분포, t분포 등은 표본의 크기를 무한히 증가시킴에 따라 정규분포에 접근하게 된다는 것이 알려져 있다.



# 정규분포의 중요성

정규분포는 평균, 표준편차만 주어진다면 정의할 수 있고 계산도 간단

대부분 자연현상 및 사회현상 등 정규분포를 따르고 있음

표본의 크기가 커질수록 중심극한정리에 따라 정규분포를 하지 않는 집단에서 추출한 표본의 경우도  
그 평균의 분포 ( 표본의 평균 분포)는 정규분포를 따르게 됨이 알려져 있으므로, 표본수 조건만 충족되면 거의 모든 것에 대한 분석시 정규분포 활용 가능

CLT : Central Limit Theorem

