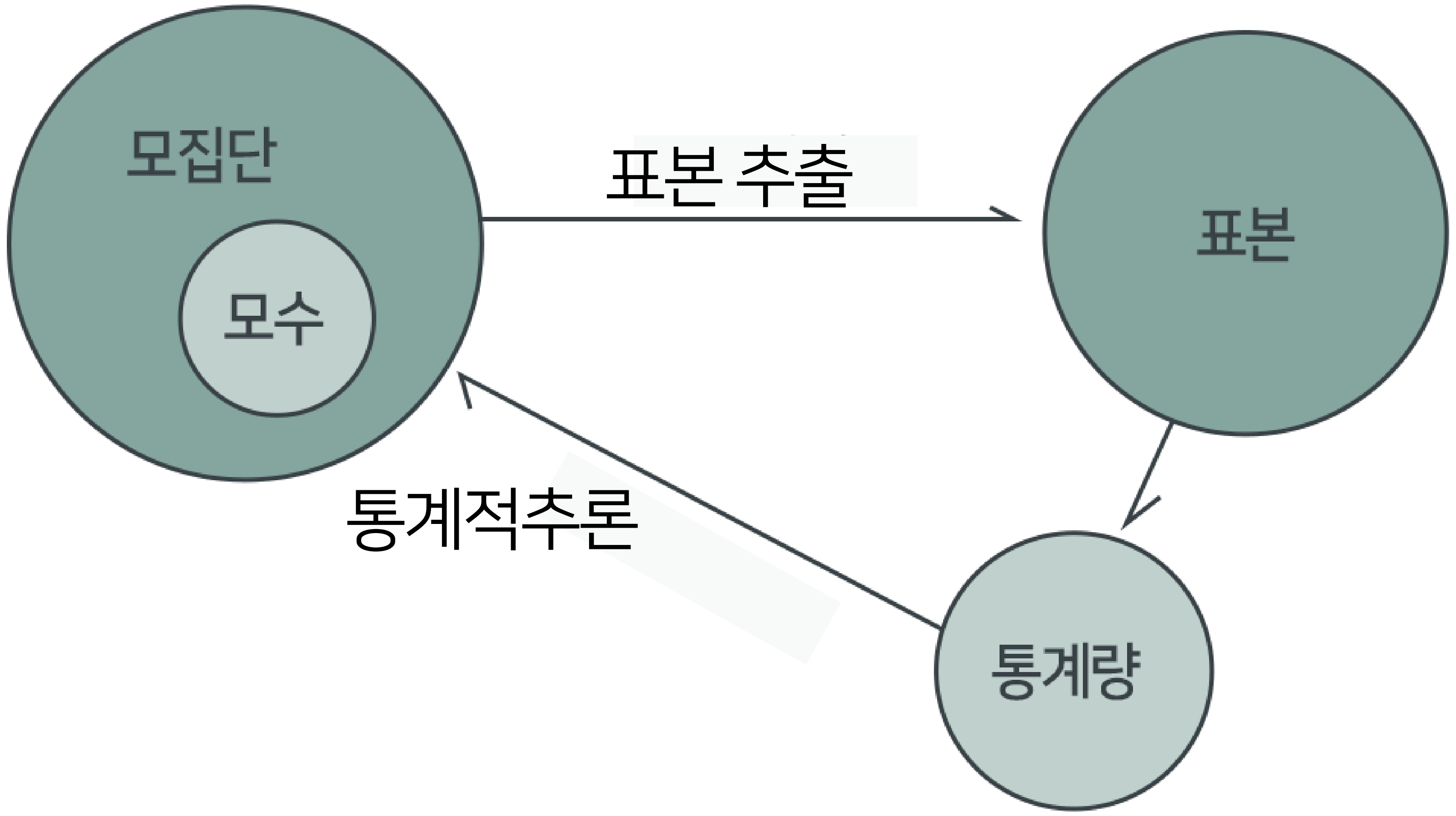
통계적 추론 개요,
표본추출법

Key words

#모집단 #모수 #확률표본
#통계량 #표본분포 #통계적 추론
#단순임의추출 #층화추출
#계통추출 #집락추출

통계적 추론 개요

통계적 추론 관련 기본 개념



확률표본, 통계량, 표본분포

모집단의 분포와 확률표본

- 모집단의 변수 X 의 확률분포함수를 $f(x)$ 로 가정.

- 모집단 $f(x)$로부터의 **확률표본** (iid 표본) $X_1, X_2, \ldots, X_n$은 다음의 두가지 성질을 만족하는 모집단 $f(x)$로부터의 표본을 뜻함.

  -$X_1, X_2, \ldots, X_n$ 은 서로 독립임.
  -$X_1, X_2, \ldots, X_n$ 는 모두 동일하게 $f(x)$의 분포를 따름.

확률표본, 통계량, 표본분포

통계량과 표본분포

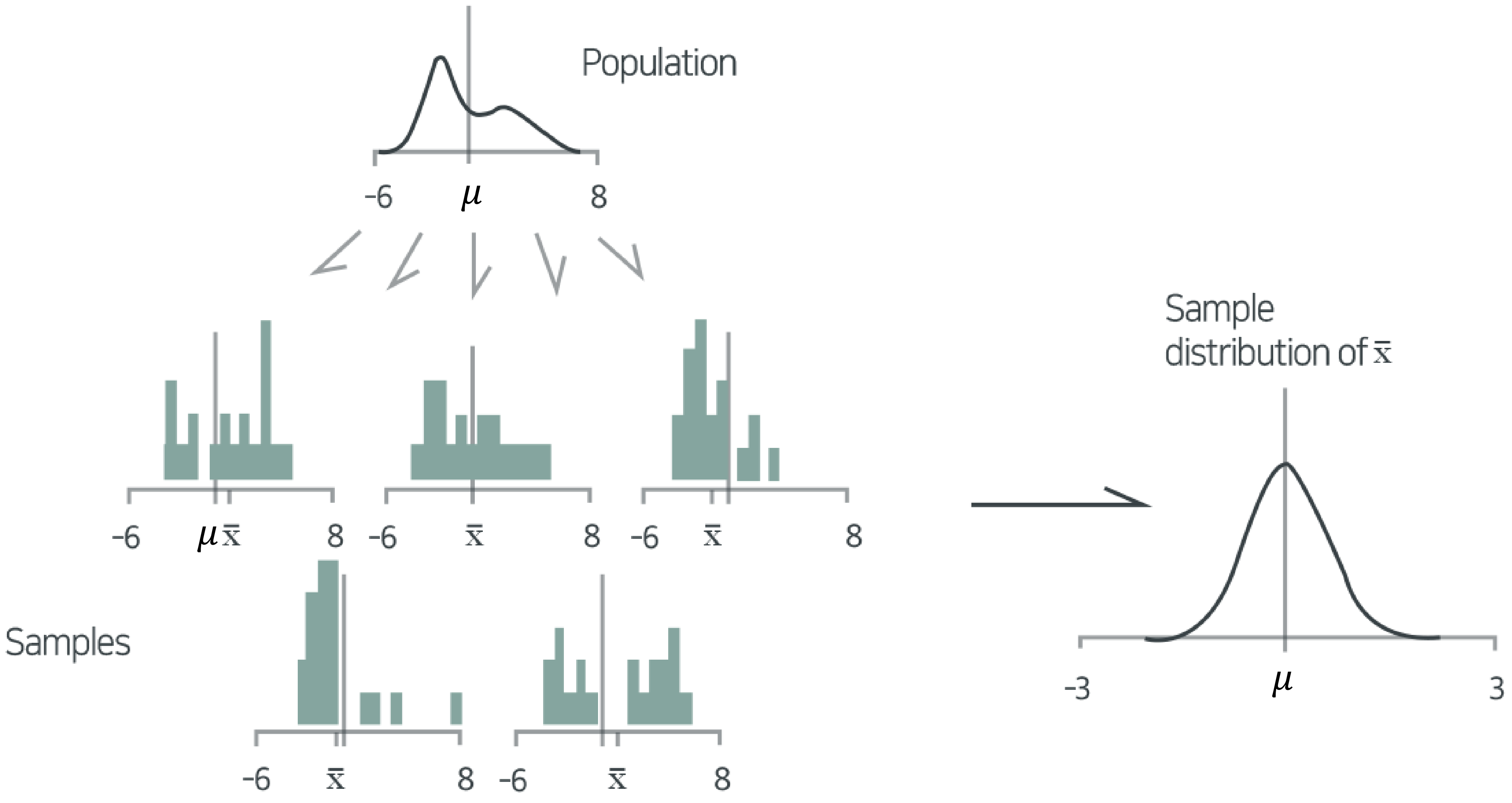
- 통계량 : 확률표본 $X_1, X_2, \ldots, X_n$ 의 함수.

  - 통계량의 예

    표본평균 $\bar{X} = \sum_{i=1}^{n} X_i / n$

- 표본분포 : 통계량의 확률분포.

확률표본, 통계량, 표본분포



표본추출법

표본추출 방법

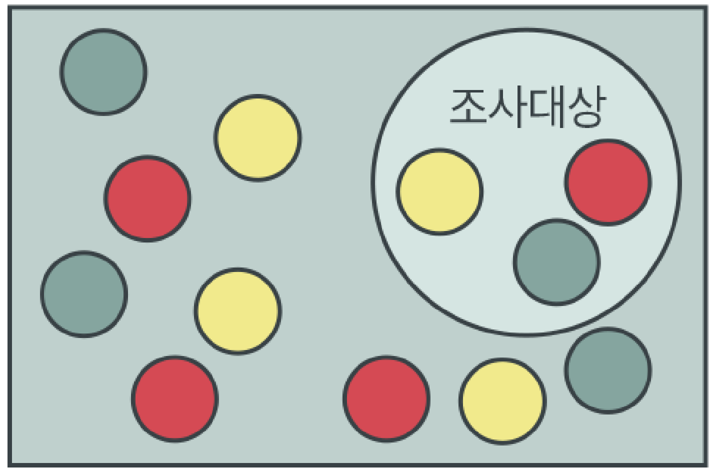
- 모집단을 대표할 수 있도록 표본을 추출하는 것이 가장 중요함.
- 표본추출방법에 따라 분석 결과 및 해석에
  큰 차이가 발생할 수 있음.
- 대표적인 4가지 표본 추출방법

  - 단순임의 추출법(simple random sampling)
  - 계통추출법(systematic sampling)
  - 집락추출법(cluster sampling)
  - 층화추출법(stratified sampling)

표본추출법

단순임의 추출법 (simple random sampling)

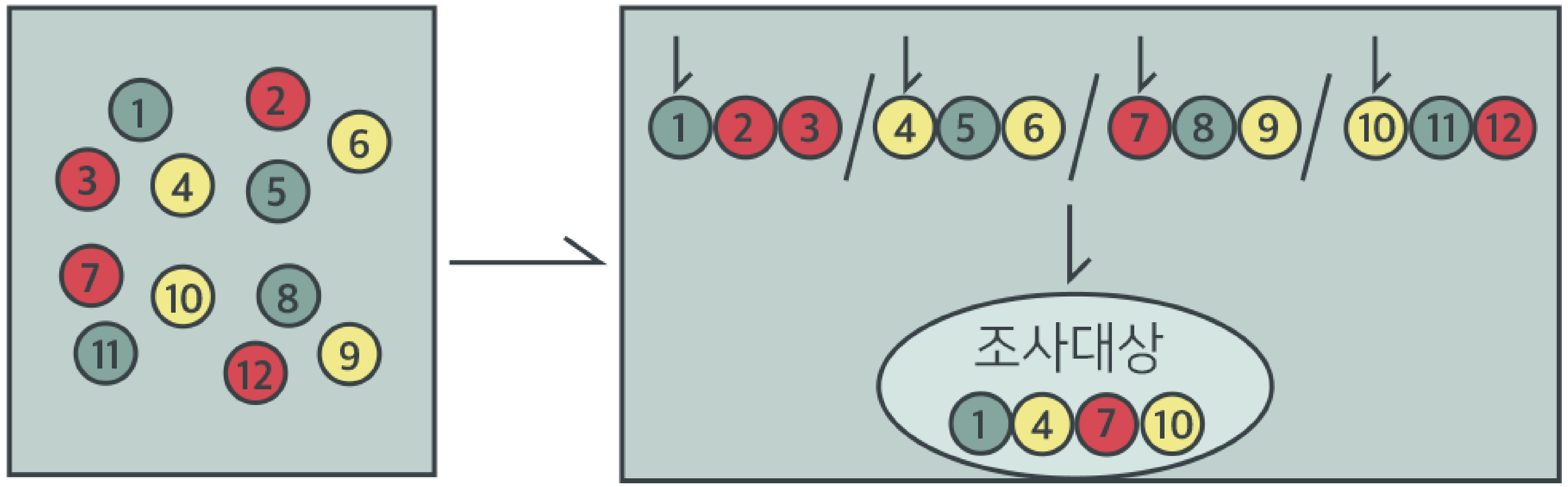
- 모집단을 구성하는 모든 원소에 대해 표본으로 추출될 확률을 동일하게 해주는 표본 추출법.
- 난수표에서 난수를 발생시켜 표본을 추출
  - N개의 원소로 구성된 모집단에서 $n(\leq N)$개의 표본을 추출할 때 각 원소에 1, 2, ..., N까지의 번호를 부여하고, 여기서 n개의 번호를 임의로 선택하여 그 번호에 해당하는 원소를 표본을 추출.



표본추출법

계통 추출법(systematic sampling)

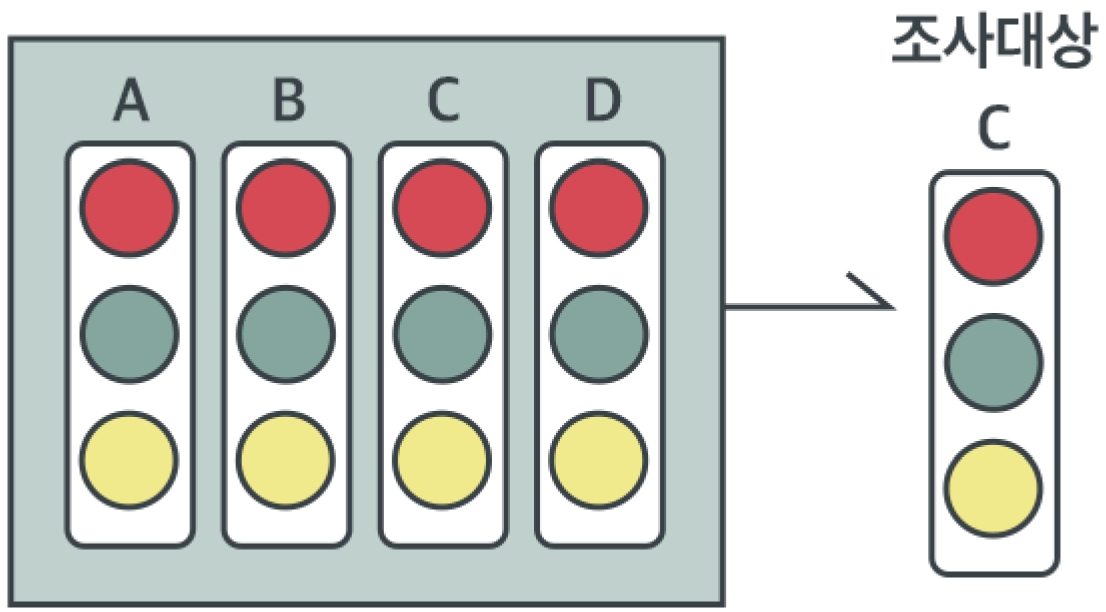
- N개의 원소로 구성된 모집단에서 $n(\leq N)$개의 표본을 추출할 때 각 원소에 1, 2, ⋯, N까지의 번호를 부여하고 이를 순서대로 나열함.
- $K(=\frac{N}{n})$개씩 n개의 구간으로 나눔.
- 첫번째 구간에서 K개의 원소 중 하나를 임의로 선택하고, 그 이후는 K개씩 띄어서 표본을 추출함.



표본추출법

집락 추출법(cluster sampling)

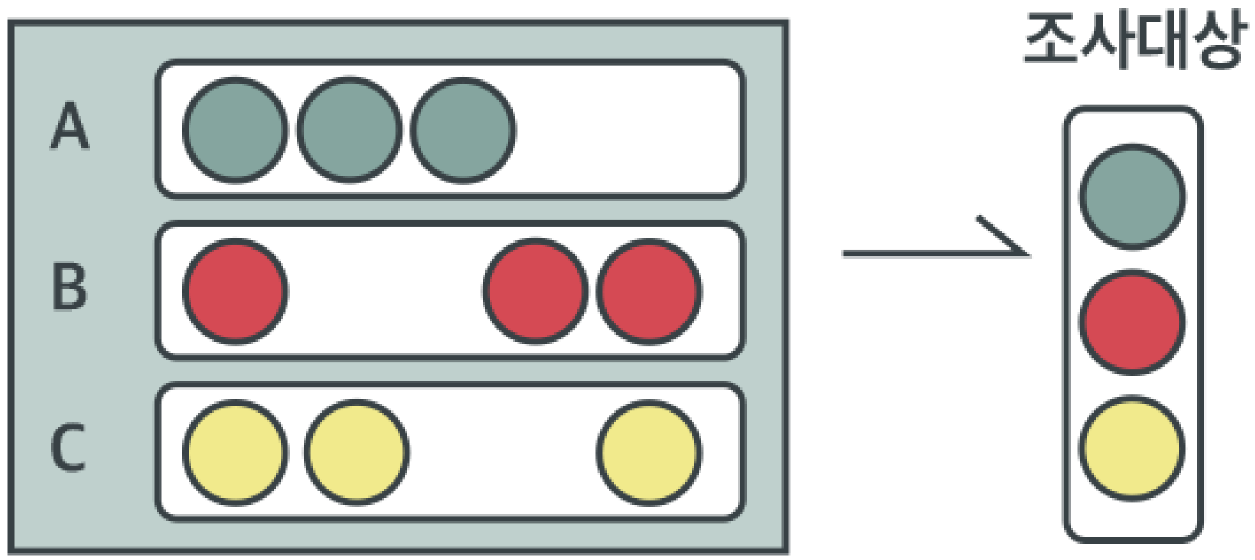
- 모집단이 몇 개의 집락(cluster)이 결합된 형태로 구성되어 있고, 각 집락에서 원소들에게 일련번호를 부여할 수 있는 경우에 이용됨.
- 일부 집락을 랜덤하게 선택한 뒤, 선택된 각 집락에서 표본을 임의로 선택함.



표본추출법

층화 추출법(stratified sampling)

- 상당히 이질적인 원소들로 구성된 모집단에서 각 계층을 고루 대표할 수 있도록 표본을 추출함.
- 이질적인 모집단의 원소들을 서로 유사한 것끼리 몇 개의 층(stratum)으로 나눈 후, 각 층에서 표본을 랜덤하게 추출함.



[추가자료] 기존 교안에 없지만, 강의 중 언급한 내용을 정리한 자료입니다.

통계적 추론 관련 주요 개념

모집단 관련

- 모집단 : 관심 대상 전체를 모아둔 집합.
- 모수 : 모집단 변수에 대한 요약.

표본 관련

- 표본 : 모집단에서 일부 대상을 추출한 것.
- 통계량 : 표본 자료에 대한 요약.
- 통계량의 표본분포 : 통계량은 확률변수이며, 통계량이 가지는 확률분포를 표본분포라고 함.

모집단의 모수	표본의 통계량
모평균 μ 또는 $E[X]$	표본평균 $\bar{X}$
모분산 σ^2 또는 $V[X]$	표본분산 S^2
모표준편차 σ 또는 $S[X]$	표본표준편차 S

<모집단 모수와 표본 통계량의 대응관계>

[추가자료] 영상에 없는 내용으로, 이해를 높이기 위한 보충학습자료입니다.

심화

표본평균 $\bar{X}$ 의 성질을 통해 통계량과 표본분포 개념 정리.

- 모집단 변수 X의 기대값이 μ 이고 분산이 σ^2이라고 할 때,
  이 모집단에서 추출된 n개의 확률표본 $X_1, X_2, \ldots, X_n$에 대한 표본평균 $\bar{X}$
- $\bar{X}$ 는 통계량.
- $\bar{X}$ 의 표본분포는 기대값이 μ 이고, 분산이 $\frac{\sigma^2}{n}$이 됨.
- $E[\bar{X}] = \mu$

  - 표본평균의 기대값은 모평균과 같음.

  - 표본평균은 표본에 따라 매번 바뀌지만, 모평균 근처의 값들로 나옴.
- $V[\bar{X}] = \frac{\sigma^2}{n}$

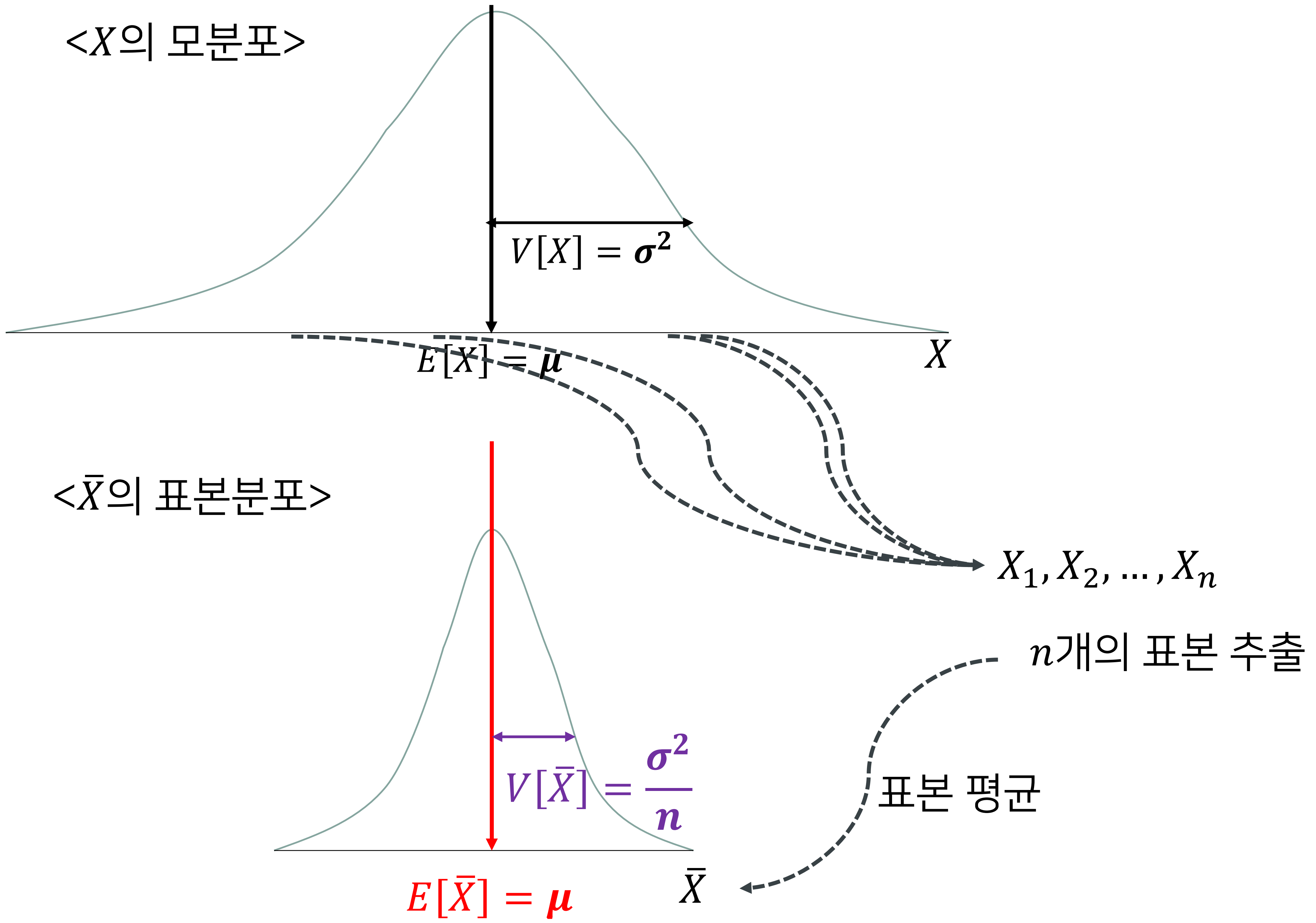
  - 표본평균의 분산은 '모분산/표본수'와 같음.

  - 표본수가 큰 경우 표본평균은 표본에 따라 변동하는 정도가 작아지므로,
  모평균에 더욱 가까운 값들이 나온다는 것.

[추가자료] 영상에 없는 내용으로, 이해를 높이기 위한 보충학습자료입니다.

심화

표본평균 $\bar{X}$ 의 성질을 통해 통계량과 표본분포 개념 정리.



[추가자료] 기존 교안에 없지만, 강의 중 언급한 내용을 정리한 자료입니다.

표본추출법 종류별 특징 요약

단순임의 추출법 (simple random sampling)

- 원하는 수의 표본을 무작위로 추출.

계통 추출법(systematic sampling)

- 동일 개수씩 구간을 나눈 뒤, 일정 간격으로 추출.

집락추출법(cluster sampling)

- 각 집락(cluster)이 모집단을 잘 대표하는 잘 섞여 있는 원소들로 구성됨.
  표본 추출이 상대적으로 쉽고 비용도 적게 듦.

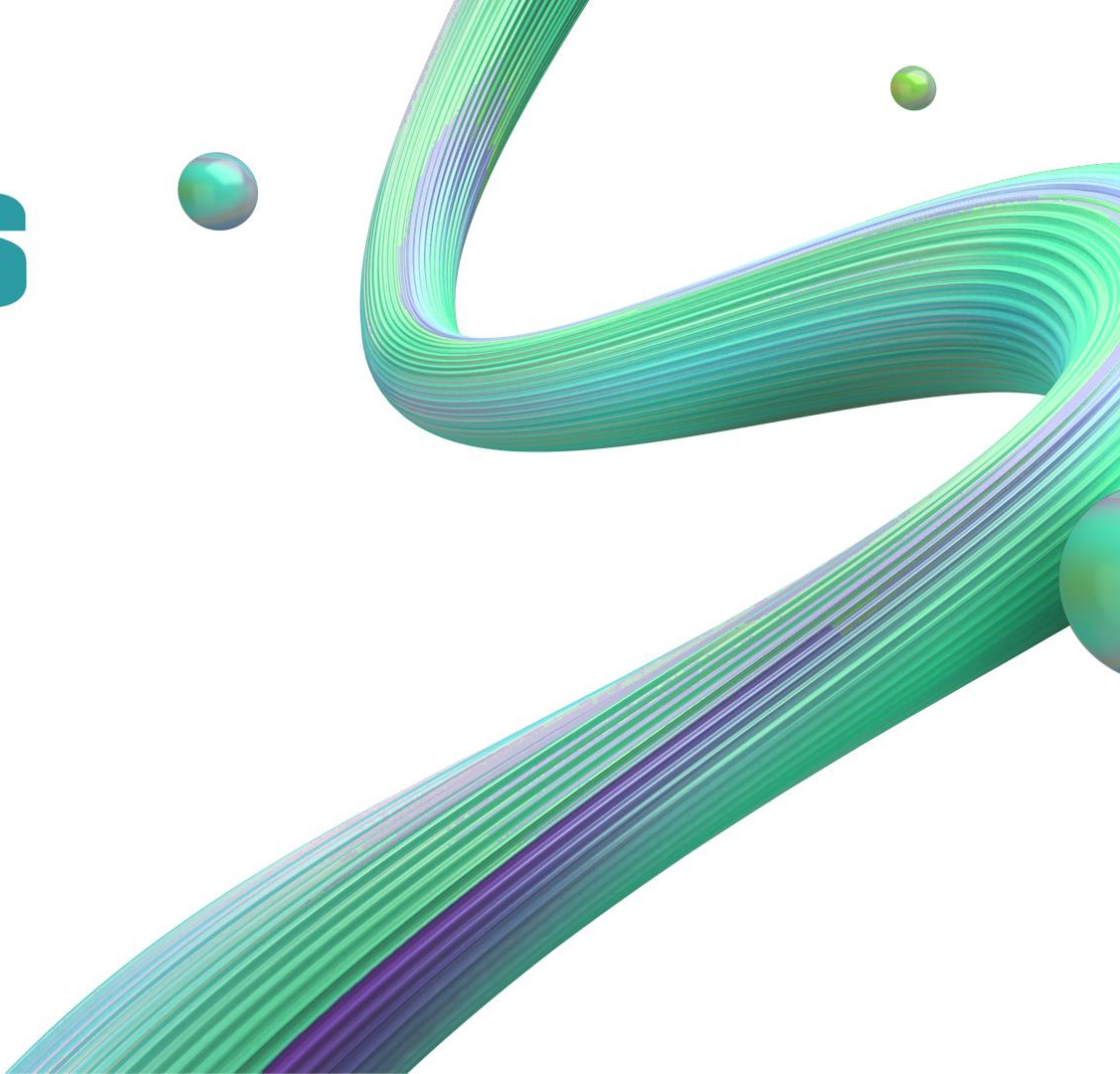
층화추출법(stratified sampling)

- 서로 이질적인 층(stratum)을 구성. 층 내 원소들은 서로 동질적임.
  각 층의 비중을 감안하여 층 별로 표본을 추출.

점추정과 구간추정

Key words

#점추정 #구간추정
#추정량 #추정치
#신뢰구간 #신뢰수준



점추정과 구간추정

통계적 추론 – 추정 개요

- 추정량

  - 모수 θ의 추정에 사용되는 통계량을 θ의 추정량($\hat{\theta}$ 으로 표기)이라고 함.
  - 관찰된 표본자료로 추정량의 값을 계산한 것을 추정치라고 함.

- 점추정과 구간추정

  - 점추정 (point estimation) : 하나의 모수를 한 개의 값으로 추정
  - 구간추정 (interval estimation) : 모수가 포함되리라 기대되는 구간으로 모수를 추정.

신뢰구간

- 모수 θ에 대한 신뢰구간 도출을 위해서는 다음을 알아야 함

  -추정량 $\hat{\theta}$

  -추정량 $\hat{\theta}$ 의 표본분포 : 일반적으로 모수 θ에 의존함.

신뢰구간의 정의

- θ의 추정량 $\hat{\theta}$을 적절하게 변형한 L과 U 이 있어
  다음을 만족하는 경우,

$$P[L \le \theta \le U] = 1 - \alpha, \qquad 0 \le \alpha \le 1,$$

- 구간 (L, U)을 θ의 $100(1-\alpha)\%$ 신뢰구간이라고 함.
- $100(1-\alpha)\%$: 신뢰수준.

모평균에 관한 신뢰구간

모평균 μ에 관한 신뢰구간(신뢰수준 : 1 - α)

- 모수
  - 모평균 μ
- 추정량
  - 표본평균 $\bar{X}$

모평균에 관한 신뢰구간

모평균 μ에 관한 신뢰구간(신뢰수준 : 1 - α)

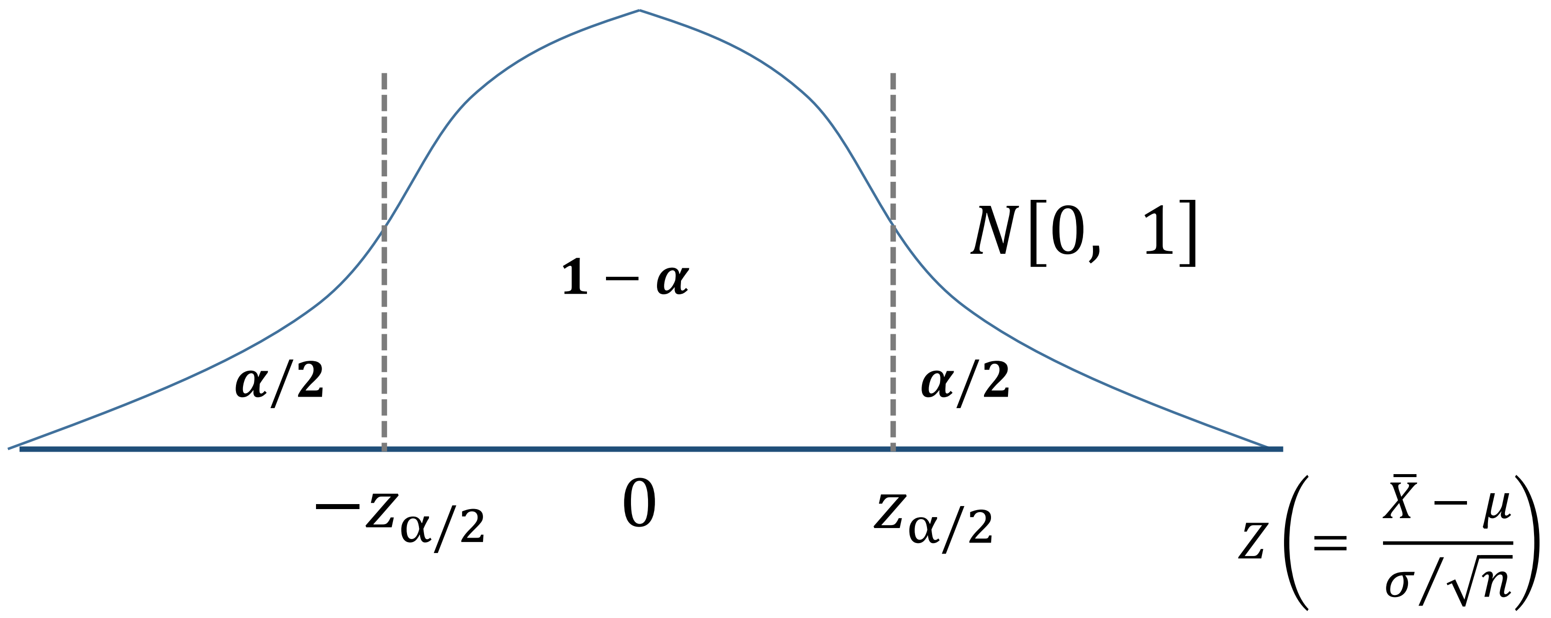
- 추정량의 표본분포

  -$X_1, \ldots, X_n$ 이 모분산 σ^2 이 알려진 정규 모집단 $N[\mu,\ \sigma^2]$ 으로부터의 확률표본인 경우.

$$\bar{X} \sim N\left[\mu, \frac{\sigma^2}{n}\right] \quad \leftrightarrow \quad Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N[0,\ 1]$$

모평균에 관한 신뢰구간

모평균 μ에 관한 신뢰구간(신뢰수준 : 1 - α)



모평균에 관한 신뢰구간

모평균 μ에 관한 신뢰구간(신뢰수준 : 1 - α)

$$P\left[-z_{\alpha/2} \le Z \le z_{\alpha/2}\right] = 1 - \alpha$$

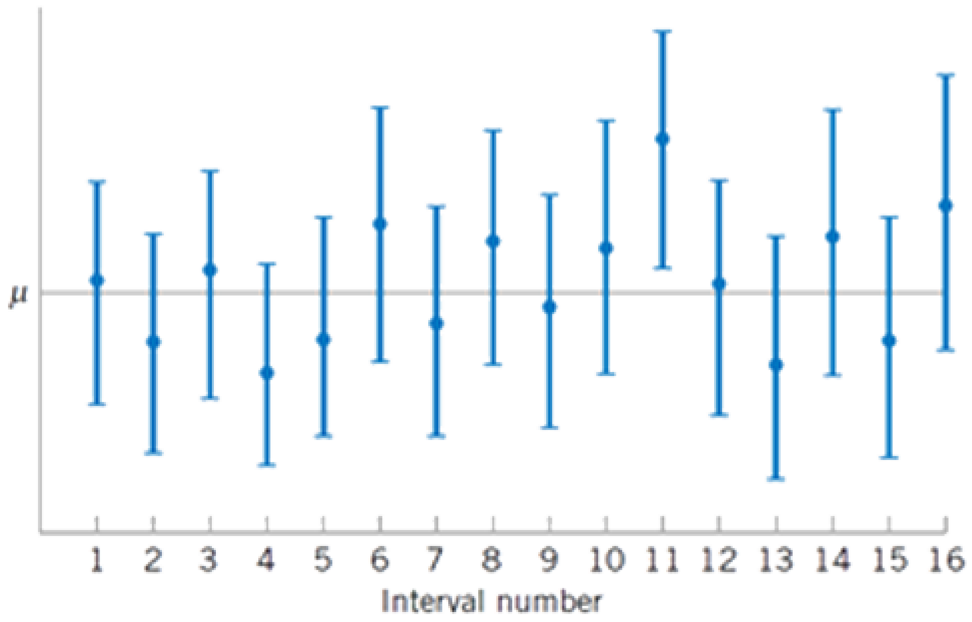
$$P\left[-z_{\alpha/2} \le \frac{\bar{X} - \boldsymbol{\mu}}{\sigma/\sqrt{n}} \le z_{\alpha/2}\right] = 1 - \alpha$$

$$\boxed{P\left[\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le \boldsymbol{\mu} \le \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right] = 1 - \alpha}$$

모평균에 관한 신뢰구간

신뢰구간의 해석

- 신뢰수준은 동일한 구간 추정법을 반복적으로 사용할 때 얻어지는 신뢰구간들이 참값 θ를 품을 확률을 의미함.



$-\mu$ 에 관한 90% 신뢰구간 : $P\left[\bar{X} - 1.645\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{X} + 1.645\frac{\sigma}{\sqrt{n}}\right] = 0.90$

$\rightarrow \bar{X} \pm 1.645\frac{\sigma}{\sqrt{n}}$

모평균에 관한 신뢰구간

신뢰구간의 해석

- μ 에 관한 추정에서 신뢰수준 90%의 의미는 n 개의 표본을 추출하는 과정을 반복하여 매번 이 구간 추정식으로 신뢰구간을 계산한다면, 이들 중 90%에 해당하는 신뢰구간들이 참값 μ를 포함하는 것을 의미.

- n 개의 표본으로 구한 하나의 신뢰구간에 모수가 포함되었는지 여부는 알 수 없음.

모평균에 관한 신뢰구간

예제

- 어떤 수산물 상점에서 판매하는 고등어의 체장을 파악하고자 7마리를 추출하여 그 체장을 조사하였더니 다음과 같았다. 고등어의 체장은 정규분포를 따르며 고등어 체장의 표준편차가 1.5로 알려져 있다고 할 때, 90% 신뢰수준에서 고등어 체장의 신뢰구간을 구해보자.

$$28.9, 32.4, 29.8, 30.6, 27.8, 29.4, 31.3$$

모평균에 관한 신뢰구간

예제

- 표본평균 $\bar{x}$= 30.029
- 신뢰수준 90% = 100(1-0.10)% : $Z_{0.10/2} = 1.645$
- 모표준편차 σ=1.5
- 표본의 크기 n=7

모평균에 관한 신뢰구간

예제

- 모평균 μ에관한 90% 신뢰구간 :

$$\bar{X} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}} = 30.029 \pm 1.645\frac{1.5}{\sqrt{7}} = (29.676, 30.381)$$

▶ (29.676, 30.381)에 고등어체장의 모평균이 포함되며 이를 90% 신뢰할 수 있음.

[추가자료] 기존 교안에 없지만, 강의 중 언급한 내용을 정리한 자료입니다.

추정 기초 개념

점추정 vs 구간추정

- 점추정은 표본정보를 하나의 값으로 요약하여 모수를 추정하지만,
  구간추정은 모수가 포함될 것으로 예상되는 구간을 이용하여 모수를 추정함.

추정량

- 추정을 목적으로 하는 표본 통계량. (cf. 추정치는 추정량의 관찰값임)
- 관심모수를 θ, 그 모수를 추정하기 위한 추정량을 $\hat{\theta}$으로 표기.

[추가자료] 기존 교안에 없지만, 강의 중 언급한 내용을 정리한 자료입니다.

신뢰구간을 이용한 추정

신뢰구간 도출과정

1. 모수와 추정량을 정의
   - 모수는 μ, 추정량은 $\bar{X}$
2. 추정량의 표본분포
   - 정규모집단, 모분산 σ^2 이 알려진 경우, $\bar{X} \sim N\left[\mu, \frac{\sigma^2}{n}\right] \rightarrow Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N[0, \quad 1]$
3. 추정량의 표본분포를 이용하여 신뢰구간 1－α 인 구간을 정의.
   - $P\left[-z_{\alpha/2} \le Z \le z_{\alpha/2}\right] = 1 - \alpha$
4. 구간(추정량의 함수)이 모수를 포함하는 형태로 부등식 정리.
   - $P\left[\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le \boldsymbol{\mu} \le \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right] = 1 - \alpha$

[추가자료] 기존 교안에 없지만, 강의 중 언급한 내용을 정리한 자료입니다.

신뢰구간을 이용한 추정

신뢰수준 $1-\alpha$ 인 신뢰구간이란?

- 신뢰구간 추정량이 모수를 포함할 가능성이 $1-\alpha$가 되도록 만들어짐을 의미.
- 표본 자료로 구해진 신뢰구간 추정치에 모수가 실제로 포함되는지 여부는 알 수 없음.

(심화) 신뢰구간의 오차한계

- 오차한계 (margin of error) : 신뢰구간 길이의 절반을 의미.

- 신뢰수준이 높을수록, 오차한계가 작을수록 좋은 신뢰구간으로 볼 수 있음.
  - 신뢰수준은 추정이 얼마나 믿을만한지를 반영함.
  - 오차한계는 추정이 얼마나 정확한지를 반영함.

[추가자료] 영상에 없는 내용으로, 이해를 높이기 위한 보충학습자료입니다.

(심화) 신뢰구간의 오차한계

신뢰수준과 오차한계는 상충관계임.

- 신뢰수준을 높이고자 하면, 구간이 길어짐.
- 구간의 길이를 줄이고자 하면, 신뢰수준이 낮아짐.

신뢰구간 추정 시 오차한계를 줄이고자 한다면?

- 표본 수 n을 키우거나 (→ $\sqrt{n}$이 커짐.)
- 신뢰수준을 낮추면 됨 (→ $z_{\alpha/2}$ 작아짐.)

$$\bar{X} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

줄이려면?

가설검정의 원리

Key words

#귀무가설 #대립가설 #검정통계량
#기각역 #유의확률(p-value)
#유의수준 #제1종오류 #제2종오류

가설검정을 위한 기본개념

통계적 가설검정이란?

- 표본으로부터 주어지는 정보를 이용하여, 모수에 대한 예상, 주장 또는 추측 등의 옳고 그름을 판정하는 과정.

가설검정을 위한 기본개념

가설

- 통계적 가설은 모수에 대한 예상, 주장 또는 추측을 표현한 것.
  - 귀무가설(H_0) : 지금까지 사실로 알려져 있는 가설.
  - 대립가설(H_1) : 표본자료로부터의 강력한 증거에 의해 입증하고자 하는 가설.
  - 가설검정은 표본의 정보가 귀무가설 H_0에 대한 충분한 반증이 되는가를 보는 것임.
  - 결론도 "귀무가설 H_0 을 기각" 또는 "귀무가설 H_0 을 기각하지 못함" 중 하나로 표현.

가설검정을 위한 기본개념

가설 유형

- 관심 모수가 μ이고 검정하고자 하는 모수의 경계값이 μ_0라고 할 때,

단측(한쪽 꼬리)검정		양측(양쪽 꼬리) 검정
왼 꼬리 검정	오른 꼬리 검정	
$H_0: \mu = \mu_0$ $H_1: \mu < \mu_0$	$H_0: \mu = \mu_0$ $H_1: \mu > \mu_0$	$H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$

가설검정을 위한 기본개념

검정통계량

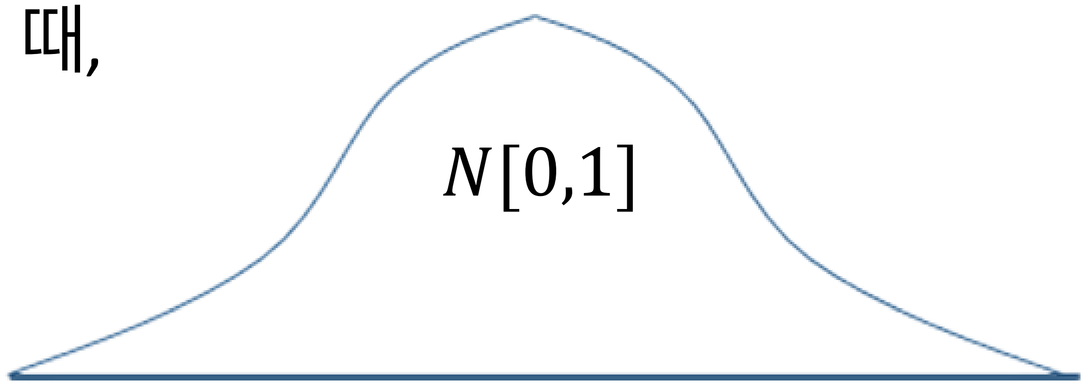
- 귀무가설 H_0 과 대립가설 H_1 중 어느 하나를 택하는 기준을 결정하는 통계량.
- 모집단이 모분산이 알려진 정규분포인 경우, 모평균 μ에 관한 검정의 경우,

- 검정통계량 : 표본평균 $\bar{X}$의 함수인 $Z = \dfrac{\bar{X} - \mu}{\sigma/\sqrt{n}}$

- 검정통계량 표본분포 : $Z = \dfrac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N[0,1]$

- 귀무가설 $H_0: \mu = \mu_0$이 사실일 때,

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim N[0,1]$$



$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$$

가설검정을 위한 기본개념

귀무가설 H_0의 기각여부를 판단하기 위한 2가지 접근 방식

- 기각역에 의한 검정.

  - 기각역 : 귀무가설 H_0 을 기각하게 하는 검정통계량 관측치의 영역.

- 유의확률에 의한 검정.

두 방식은 동일한 결론을 가짐.

가설검정을 위한 기본개념

유의확률에 의한 검정 방법

- 유의확률 (p-value)

-귀무가설 H_0가 사실인 경우의 검정통계량의 분포에서,
 대립가설 H_1의 방향으로, 검정통계량의 관찰값보다
 더 극단적인 값이 나올 확률로 계산됨.

가설검정을 위한 기본개념

유의확률에 의한 검정 방법

- 모분산 σ^2이 알려진 정규 모집단에서 모평균 μ에 관한 검정의 경우.

  귀무가설 H_0 가 사실일 때

  검정통계량 Z의 분포 $Z = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \sim N[0,1]$에서,

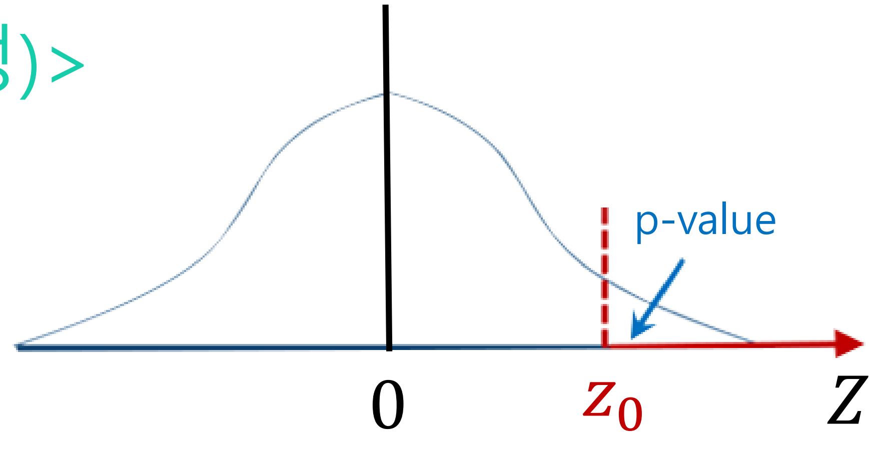
  z_0(= 표본 자료로부터 계산된 검정통계량의 값)을 이용하여 계산.

<단측검정 (오른꼬리 검정)>

$H_0 : \mu = \mu_0$

$H_1 : \mu > \mu_0$

▸ p-value $= P[Z > z_0]$



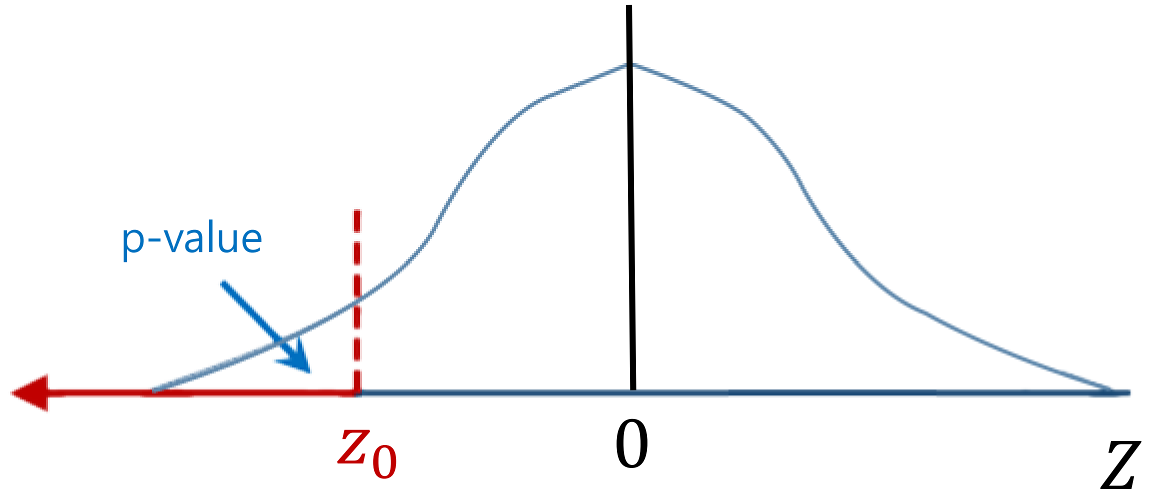
가설검정을 위한 기본개념

유의확률에 의한 검정 방법

<단측검정 (왼꼬리 검정)>

$H_0 : \mu = \mu_0$
$H_1 : \mu < \mu_0$

- p-value = $P[Z < z_0]$



가설검정을 위한 기본개념

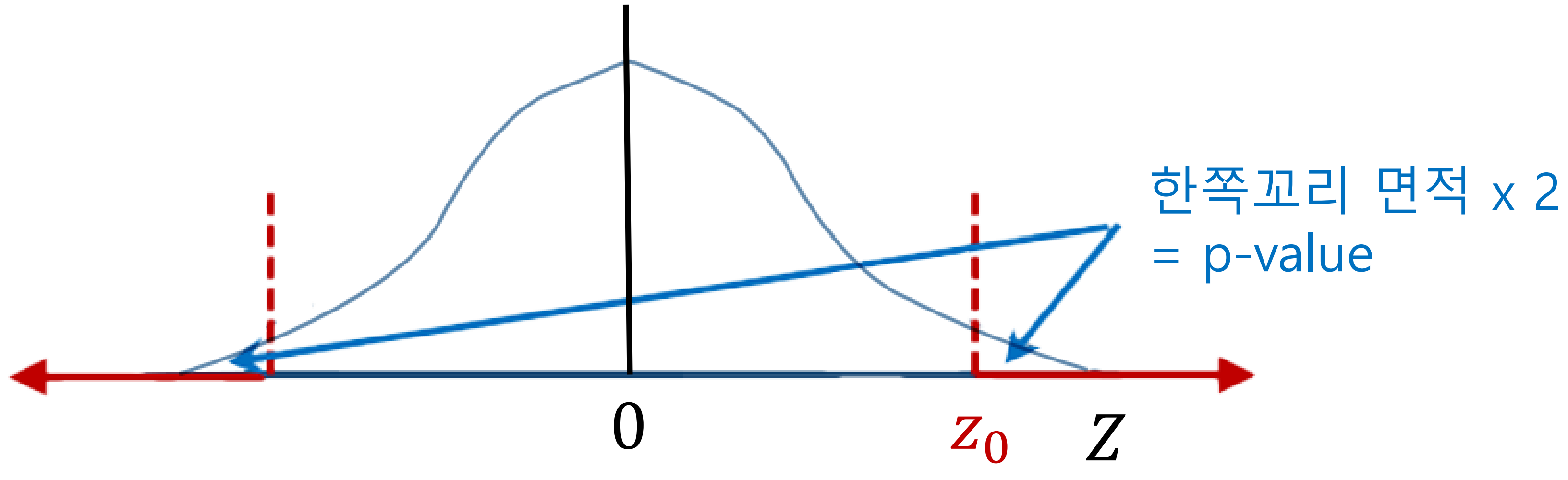
유의확률에 의한 검정 방법

<양측검정 (양쪽꼬리 검정)>

$H_0: \mu = \mu_0$

$H_1: \mu \neq \mu_0$

- p-value = $\min(P[Z > z_0], P[Z < z_0]) \times 2$



가설검정을 위한 기본개념

유의확률에 의한 검정 방법

- 유의확률 (p-value)의 해석

- 유의확률은 값이 작을수록 해당하는 표본은 귀무가설 H_0가 사실인 경우의 표본으로 보기 어려우며, 대립가설 H_1을 더욱 지지하는 것으로 해석할 수 있음.
- 유의확률이 일정수준보다 작으면, 귀무가설 H_0를 기각.

가설검정을 위한 기본개념

가설검정에서의 의사결정에서 발생할 수 있는 오류.

		실제 현상	
		귀무가설 H_0이 사실	대립가설 H_1이 사실
검정결과	귀무가설 H_0를 기각하지 못함	올바른 의사결정	**제2종 오류**
	귀무가설 H_0를 기각	**제1종 오류**	올바른 의사결정

가설검정을 위한 기본개념

가설검정에서의 의사결정에서 발생할 수 있는 오류.

- 제 1종 오류 :

  - 실제로는 귀무가설 H_0가 사실인데, 귀무가설 H_0을 기각하게 되는 오류.
  - 제 1종 오류의 확률을 α로 표기함.

- 제 2종 오류 :

  - 실제로는 대립가설 H_1이 사실인데, 귀무가설 H_0을 기각하지 못하게 되는 오류.
  - 제 2종 오류의 확률을 β로 표기함.

가설검정을 위한 기본개념

유의수준 (significance level) $100\,\alpha\,\%$의 검정법

- 검정의 오류
  - 두 종류의 오류는 서로 상충관계로 동시에 줄이기 어려움.
  - 더 위험한 제1종 오류를 일정수준을 넘지 못하도록 제한하면서, 제2종 오류를 최소로 하도록 하는 검정법을 찾고자 함.
- 유의수준
  - 제1종 오류를 범할 확률의 최대 허용 한계.

가설검정을 위한 기본개념

유의수준 (significance level) $100\,\alpha\,\%$의 검정법

- 자료로부터 계산된 유의확률 (p-value)이 주어진 유의수준 α 보다 작은 경우에 귀무가설 H_0를 기각함.
  - 유의확률(p-value) $\leq$ 유의수준(α) 면, 귀무가설 H_0를 기각.

가설검정을 위한 기본개념

예제

- 어느 도넛 회사에서 새로 출시된 도넛의 평균 칼로리가 100kcal보다 작은가에 대해 알아보고자 한다.
  해당 도넛의 칼로리가 정규분포를 따르며, 칼로리의 모표준편차는 10이라고 하자. 검정을 위해 64개의 표본을 추출하였더니 표본들의 칼로리의 평균이 95kcal었다고 하자.
  유의수준 5%에서 검정하여라.
- 가설

-귀무가설 $H_0: \mu = 100$
-대립가설 $H_1: \mu < 100$

가설검정을 위한 기본개념

예제

- 검정통계량 및 분포

- 정규분포로부터의 64개의 표본에 대한 표본평균을 $\bar{X}$ 라고 하면,

$Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} = \frac{\bar{X}-\mu}{10/\sqrt{64}} \sim N[0, 1]$ 이므로,

귀무가설 H_0가 사실인 경우, $Z = \frac{\bar{X}-100}{\sigma/\sqrt{n}} = \frac{\bar{X}-100}{10/\sqrt{64}} \sim N[0, 1]$를 따름.

가설검정을 위한 기본개념

예제

- 유의확률(p-value)에 의한 검정

- 유의확률은 귀무가설 H_0 가 사실일 때,

$$P\left[Z < \frac{95-100}{10/\sqrt{64}}\right] = P[Z < -4] = 0.00003$$

- 유의확률(p-value)=0.00003 가 유의수준 $\alpha = 0.05$보다 작으므로 귀무가설을 기각함.

[추가자료] 기존 교안에 없지만, 강의 중 언급한 내용을 정리한 자료입니다.

가설검정 개념정리

가설

- 모수에 관한 상반된 두 개의 주장
  - 귀무가설(H_0) : 지금까지 사실로 알려짐. 모수를 특정값(=)으로 표현
  - 대립가설(H_1) : 연구자의 새로운 주장. 모수를 특정값을 경계로 하는 범위(>, <, !=)로 표현.
- 표본 정보가 귀무가설 H_0에 대한 충분한 반증이 되는가를 확인하여,
  두 가설 중 하나를 선택하는 의사결정을 내림 → 'H_0기각' 또는 'H_0기각하지 못함'.

검정통계량과 표본분포

- 검정통계량 : 검정에 활용하기 위한 표본통계량.
- 표본 정보가 귀무가설 반증의 근거가 충분한가? → 귀무가설이 사실일 때
  표본의 검정통계량이 가질 수 있는 값들의 분포(검정통계량의 표본분포)의 정보가 필요.

가설검정 개념정리

2가지 접근법

- 기각역에 의한 접근

- 표본의 검정통계량 관찰값이, 귀무가설 하 분포에서 나오기 힘든 극단적인 "값"인가?

- 유의확률(p-value)에 의한 접근

- 표본의 검정통계량 관찰값이, 귀무가설 하 분포에서 나올 "확률"이 얼마나 희박한가?

가설검정 개념정리

유의확률(p-value)

- 산출방법

  - 귀무가설 하의 검정통계량의 분포에서, 검정통계량의 관찰값보다, 대립가설 방향으로 더 극단적인 값이 나올 확률.
  - 대립가설 방향에 따라 오른꼬리, 왼꼬리, 양쪽꼬리 확률에 해당함.

해석

- 유의확률은 작을수록 귀무가설에 대한 반증이자, 대립가설에 대한 지지를 나타냄.

[추가자료] 기존 교안에 없지만, 강의 중 언급한 내용을 정리한 자료입니다.

가설검정 개념정리

검정의 오류

- 제1종 오류 : 귀무가설이 사실인데 귀무가설을 기각하는 오류.
- 제2종 오류 : 대립가설이 사실인데 귀무가설을 기각하지 않는 오류.
- 두 종류의 오류는 서로 상충관계로 동시에 줄일 수 없음.
- 제1종 오류가 일정수준(유의수준)을 넘지 못하도록 통제하며
제2종 오류를 줄이고자 하는 것이 가설검정에서 기본적인 아이디어임.

유의수준 α 검정법

- 유의수준 : 제1종 오류의 최대허용한계
- 유의확률(p-value) < 유의수준(α)면 귀무가설을 기각함.