第10장 가설검정:한 모집단

- 가설검정의 기본개념
- 평균의 가설검정
- 모비율의 가설검정
- 모분산의 가설검정

Ⅰ. 가설검정의 기본개념

1. 가설검정이란?

   모집단에 대한 어떤 주장을 통계적 가설이라고 하며, 가설의 기각 여부에 대한 통계적 의사결정 과정

2. 가설의 형태

1) 귀무가설(Null hypothesis)(=영가설) (=H0)

   -모집단에 대한 지금까지의 주장 또는 믿음(기존가설)

   -검정하기 위하여 설정한 가설

   -모수값에 대하여 특정한 값을 나타내는 등호, 또는 등호를 포함하는 부등호의 형태로 표현됨(H_0: $\theta =, \geq, \leq \theta_0$)

2) 대립가설(Alternative hypothesis) (=연구가설)(=H1)

   - 기존상태로부터 새로운 변화 또는 효과가 존재하는 것

   - 귀무가설을 부정하고 연구자가 주장하는 새로운 가설

   - 귀무가설의 반대형태가 대립가설이 된다

   - 귀무가설의 부정형으로 표현된다. (H_0: $\theta \neq, <, > \theta_0$)

I. 가설검정의 기본개념

3. 가설형태에 따른 검정의 유형

① 양측검정

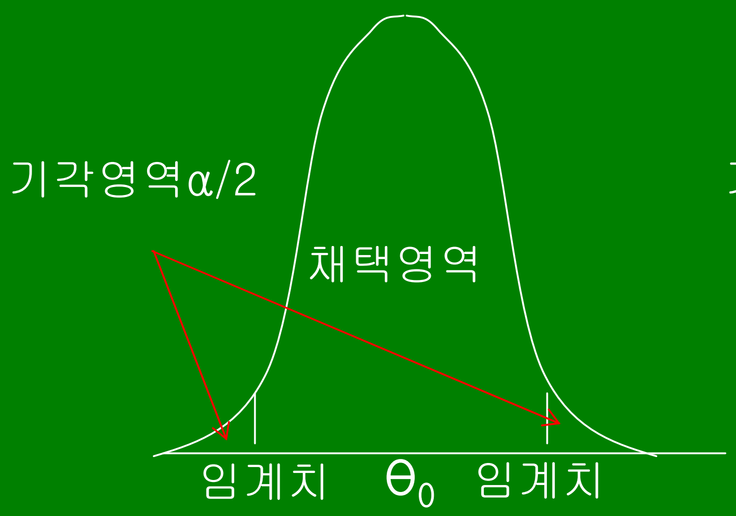
$$H_0 : \theta = \theta_0$$
$$H_1 : \theta \neq \theta_0$$

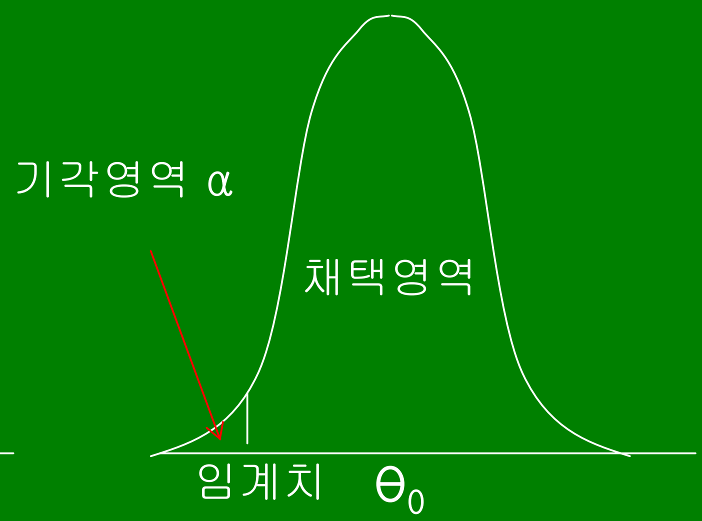
② 좌측검정

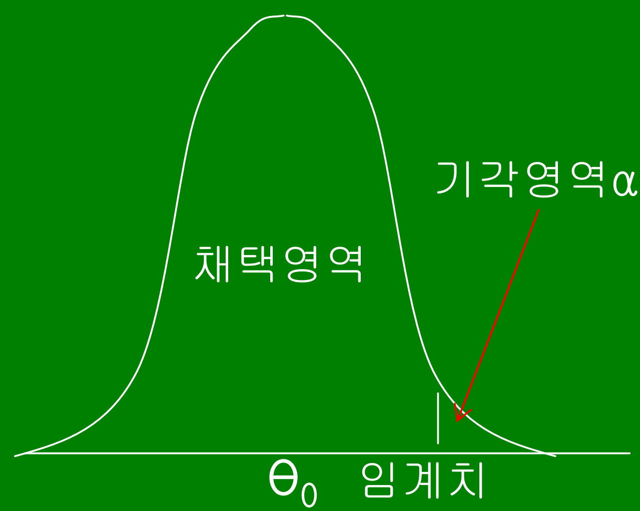
$$H_0 : \theta \geq \theta_0$$
$$H_1 : \theta < \theta_0$$

③ 우측검정

$$H_0 : \theta \leq \theta_0$$
$$H_1 : \theta > \theta_0$$







I. 가설검정의 기본개념

① 양측검정

$$H_0 : \theta = \theta_0$$
$$H_1 : \theta \neq \theta_0$$

② 좌측검정

$$H_0 : \theta \geq \theta_0$$
$$H_1 : \theta < \theta_0$$

③ 우측검정

$$H_0 : \theta \leq \theta_0$$
$$H_1 : \theta > \theta_0$$

1) <u>양측검정</u>: 귀무가설이 “모수의 값이 어떤 특정 값이다.”라는 형식의 경우 검정통계량의 값이 모수값과 같지 않은 경우는 크거나 작은 경우에 해당함으로 분포의 양극단의 2개의 임계치와 비교

2) <u>단측검정</u>

- 좌측검정:“모수의 값이 어떤 특정 값 이상이다.”라고 표현한 경우로서 검정통계량의 값이 모수의 값보다 작은 영역에 속하는지 여부를 좌측의 임계치와 비교
- 우측검정 :“모수의 값이 어떤 특정 값 이하이다.”라고 표현한 경우로서 검정통계량의 값이 모수의 값보다 큰 영역에 속하는지 여부를 우측의 임계치와 비교

예1)

가설1) “한국인 성년 남자의 평균신장이 170cm이다”

가설2) “한국인 성년 남자의 평균신장이 170cm보다 크거나 같을 것이다.”

가설3) “한국인 성년 남자의 평균신장이 170cm보다 작거나 같을 것이다.”

◆ 단순가설(양측검정)-가설1의 검정

- 표본평균의 값이 170보다 아주 크거나 아주 작은 경우 귀무가설을 기각하고 대립가설을 채택할 수 있다.
- 이의 판단기준은 표본평균의 분포를 이용하여 설정하는데 이는 특히 분포의 양쪽 꼬리부분에 사회적 관습이나 연구자의 주관에 의해 기각영역을 설정하여 검정통계량이 기각역에 해당하는지의 여부를 판단
- 기각역에 포함이 되면 귀무가설을 기각하고 대립가설을 채택하고 기각역에 포함되지 않으면 귀무가설을 기각할 만한 충분한 증거가 없음으로 판정한다.

◆ 복합가설(좌측검정)-가설2의 검정

- 모수값이 특정값보다 크다는 식으로 나타난 경우에는 좌측검정을 하며 임계치보다 표본통계량의 값이 작으면 귀무가설을 기각하며 즉, 표본평균의 값이 170보다 아주 작으면 귀무가설을 기각한다.
- 이는 표본평균의 분포로부터 유의수준에 따라 기준이 되는 임계치를 설정하며 이때의 임계치는 하나가 된다. 표본통계량의 값이 기각역 밖에 있으면 귀무가설을 기각할 만한 충분한 근거가 없다고 판단함.)

◆ 복합가설(우측검정)-가설3의 검정

- 모수값이 특정값보다 작다는 식으로 나타난 경우에는 우측검정을 하며 임계치보다 표본통계량의 값이 크면 귀무가설을 기각하며, 표본통계량의 값이 임계치보다 작으면 귀무가설을 기각할 만한 충분한 근거가 없다고 판단함.)
- 표본평균의 값이 170보다 아주 크면 귀무가설을 기각한다. 이는 표본평균의 분포로부터 유의수준에 따라 기준이 되는 임계치를 설정하며, 이때의 임계치는 하나가 된다.

4. 가설검정 용어

- 검정통계량: 가설검정에서 이용하는 통계량으로 검정하려는 모수에 대응하는 표본통계량이 따르는 분포를 표준화한 통계량(예: t값, Z값, X^2값, F값 등)

- 유의수준(significant level): α
  귀무가설이 옳음에도 불구하고 귀무가설을 기각할 확률(1종 오류)로서 연구의 목적에 타당하도록 연구자의 자의에 의해 결정.(보통 사회과학에서는 1%, 5%, 10%수준 임) 유의수준에 따라 기각역과 임계치가 결정된다.

- <u>기각역(critical region)</u> : 통계적 가설검정에서 <u>검정통계량의 관측값이</u> 이 영역에 속하면 귀무가설를 기각하고 그렇지 않으면 기각할 수 없다는 결론을 내리게 되는 유의수준의 영역(=임계역)

- <u>임계치(critical value)</u> : 임계역(=기각역)의 하한과 상한값 양측검정의 경우는 음수구간의 한 값과 양수구간의 한 값으로 나타나며, 우측검정의 경우는 양수구간의 하한값 하나가 되며, 좌측검정의 경우 음수구간의 한 값이 된다. 이는 유의수준과 검정통계량에 의해 자동적으로 결정된다.

II. 가설검정의 오류

1. 제 I 종 오류(Type I error)란 귀무가설 $H0$가 실제로는 사실이어서 채택해야 함에도 불구하고 표본오차 때문에 이를 거부하는 오류를 말한다.(α 오류)

2. 제 II 종 오류(Type II error)란 귀무가설 $H0$가 허위라서 거부해야 됨에도 불구하고 표본오차 때문에 이를 채택하는 오류를 말한다.(β 오류)

실제상황 / 통계적결정	H_0가 사실	H_0가 허위
H_0 채택	옳은 결정 (1−α)	제II종 오류 (β)
H_0 기각	제 I 종 오류 (α)	옳은 결정 (1−β)

III. 가설검정의 순서

[가설검정의 순서]

가설의 설정

↓

유의수준 α 의 결정

↓

검정통계량의 선정

검정통계량을 사용할 때

- 임계치 및 기각영역의 결정
- ↓
- 검정통계량의 계산
- ↓
- 기각여부결정

P값을 사용할 때

- 검정통계량의 계산
- ↓
- p값 계산
- ↓
- 기각여부결정

IV. 모평균의 가설검정

◆ 모표준편차를 아는 경우

양측검정

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

검정통계량이 아래의 조건을 만족하면 H_0를 기각하고 H_1을 채택한다.

$$\frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} > Z_{\frac{\alpha}{2}}$$ 또는 $$\frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} < -Z_{\frac{\alpha}{2}}$$

[P값을 이용한 가설검정]

P값에 의한 결정규칙

만일 p 값 $< \alpha$ 이면 H_0를 기각

p 값 $\geq \alpha$ 이면 H_0를 채택

P값(p-value)이란 제 I 종 오류를 범할 실제로 정확한 확률. 즉 사실인 귀무가설을 거부하게 하는 유의수준 α 의 최소치를 말한다.

P값 계산방법

$H_1: \mu \neq \mu_0$ 이면(양측) p값 $= 2P(Z > |Z_S|)$

$H_1: \mu < \mu_0$ 이면(좌측) p값 $= P(Z < Z_S)$

$H_1: \mu > \mu_0$ 이면(우측) p값 $= P(Z > Z_S)$

단, $Z_S = \dfrac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}$

[좌측검정]

다음과 같이 가설이 설정된다고 하자.

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

검정통계량이 아래의 조건을 만족하면 H_0를 기각하고 H_1을 채택한다.

$$\frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} < -Z_\alpha$$

[우측검정]

다음과 같이 가설이 설정된다고 하자.

$$H_0 : \mu \le \mu_0$$
$$H_1 : \mu > \mu_0$$

검정통계량이 아래의 조건을 만족하면 H_0를 기각하고 H_1을 채택한다.

$$\frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} > Z_\alpha$$

[검정력을 구하는 절차]

1. 채택영역과 기각영역의 경계에 해당하는 표본평균 $\overline{X}$ 의 값을 계산한다.

양측검정 : $\overline{X}_{왼쪽} = \mu_0 - Z_{\frac{\alpha}{2}} \sigma_{\overline{X}} = \mu_0 - Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$

$\overline{X}_{오른쪽} = \mu_0 + Z_{\frac{\alpha}{2}} \sigma_{\overline{X}} = \mu_0 + Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$

좌측검정 : $\overline{X} = \mu_0 - Z_{\alpha} \sigma_{\overline{X}} = \mu_0 - Z_{\alpha} \frac{\sigma}{\sqrt{n}}$

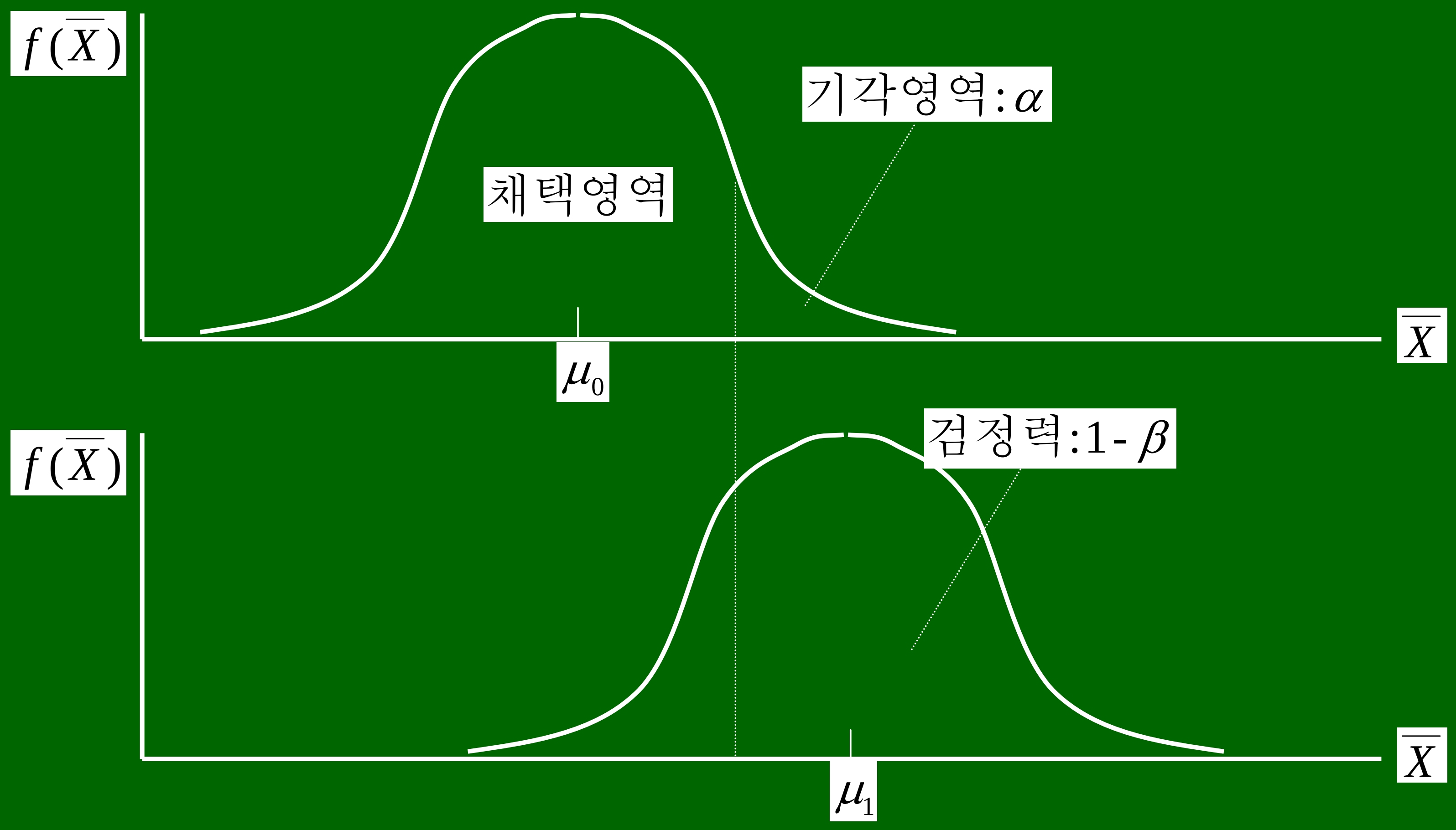
우측검정 : $\overline{X} = \mu_0 + Z_{\alpha} \sigma_{\overline{X}} = \mu_0 + Z_{\alpha} \frac{\sigma}{\sqrt{n}}$

2. 참 모평균이 μ_1 이라고 할 때 평균 μ_1 인 모집단으로부터 크기 n의 표본을 추출하여 구한 분포에서 1.에서 결정한 채택영역 속에 표본평균이 존재할 확률 β를 구한다.

3. 따라서 검정력 $1-\beta$ 를 구할 수 있다.

[검정력]

$1-\beta$ 로 나타나는 검정력(power of the test)은 거짓인 귀무가설을 기각하는 확률을 말한다.



◆ 모표준편차를 모르는 경우

소표본의 경우

표본크기가 30 미만이면 t분포를 이용한다.

$$t = \frac{\bar{x} - \mu}{s_{\bar{x}}} = \frac{\bar{x} - \mu}{S/\sqrt{n}}$$

모평균의 가설검정 : 모표준편차를 모르는 경우(소표본)

양측검정 $H_0 : \mu = \mu_0$ $H_1 : \mu \neq \mu_0$	만일 $\dfrac{\bar{x}-\mu_0}{\dfrac{s}{\sqrt{n}}} > t_{n-1,\frac{\alpha}{2}}$ 또는 $\dfrac{\bar{x}-\mu_0}{\dfrac{s}{\sqrt{n}}} < -t_{n-1,\frac{\alpha}{2}}$ 이면 H_0를기각
좌측검정 $H_0 : \mu \geq \mu_0$ $H_1 : \mu < \mu_0$	만일 $\dfrac{\bar{x}-\mu_0}{\dfrac{s}{\sqrt{n}}} \leq -t_{n-1,\frac{\alpha}{2}}$ 이면 H_0를기각
우측검정 $H_0 : \mu \leq \mu_0$ $H_1 : \mu > \mu_0$	만일 $\dfrac{\bar{x}-\mu_0}{\dfrac{s}{\sqrt{n}}} \geq t_{n-1,\frac{\alpha}{2}}$ 이면 H_0를기각

대표본의 경우

표본크기가 30 이상이면 모집단의 분포가 어떤 분포이건 중심극한정리에 의하여 표본분포는 정규분포를 따르기 때문에 t값을 이용하든 Z값을 이용하든 별로 차이가 없게 된다.

간단하게 표준정규분포의 Z값을 이용할 수 있다.

$$Z = \frac{\bar{x} - \mu}{s_{\bar{x}}} = \frac{\bar{x} - \mu}{s / \sqrt{n}}$$

V. 모비율의 가설검정

모비율의 가설검정 : 대표본

양측검정
$H_0 : p = p_0$
$H_1 : p \neq p_0$

만일 $\dfrac{\hat{p} - p}{\sqrt{p(1-p)/n}} < -Z_{\frac{\alpha}{2}}$ 또는 $\dfrac{\hat{p} - p}{\sqrt{p(1-p)/n}} > Z_{\frac{\alpha}{2}}$ 이면 H_0를 기각

좌측검정
$H_0 : p \geq p_0$
$H_1 : p < p_0$

만일 $\dfrac{\hat{p} - p}{\sqrt{p(1-p)/n}} < -Z_{\alpha}$ 이면 H_0를 기각

우측검정
$H_0 : p \leq p_0$
$H_1 : p > p_0$

만일 $\dfrac{\hat{p} - p}{\sqrt{p(1-p)/n}} > Z_{\alpha}$ 이면 H_0를기각

VI. 모분산의 가설검정

모분산의 가설검정

양측검정

$$H_0 : \sigma^2 = \sigma_0^2$$
$$H_1 : \sigma^2 \neq \sigma_0^2$$

만일 $\frac{(n-1)S^2}{\sigma_0^2} > \chi^2_{n-1,\frac{\alpha}{2}}$ 또는 $\frac{(n-1)S^2}{\sigma_0^2} < \chi^2_{n-1,1-\frac{\alpha}{2}}$ 이면 H_0 를기각

좌측검정

$$H_0 : \sigma^2 \geq \sigma_0^2$$
$$H_1 : \sigma^2 < \sigma_0^2$$

만일 $\frac{(n-1)S^2}{\sigma_0^2} < \chi^2_{n-1,1-\alpha}$ 이면 H_0 를기각

우측검정

$$H_0 : \sigma^2 \leq \sigma_0^2$$
$$H_1 : \sigma^2 > \sigma_0^2$$

만일 $\frac{(n-1)S^2}{\sigma_0^2} > \chi^2_{n-1,\alpha}$ 이면 H_0 를기각