가설검정
(Hypothesis Testing)

가설검정과 가설

- 가설(Hypothesis): 검정을 목적으로 제시된 모집단 모수의 수치에 대한 주장

– 두 가지 형태의 가설: 귀무가설(null hypothesis: H_0)과 대립가설(alternative hypothesis: H_1)

귀무가설

설정된 가설이 잘못되었다는 충분한 증거가 제시되기 전까지 참(true)으로 받아들여지는 가설

대립가설

귀무가설이 잘못되었다는 충분한 증거로 귀무가설을 기각할 때 받아들이는 가설

단측대립가설

귀무가설에서 설정된 값을 기준으로 어느 한쪽에 위치한 모든 값을 포함하는 대립가설

양측대립가설

귀무가설에서 설정된 값을 기준으로 양쪽에 위치한 모든 값을 포함하는 대립가설

가설검정과 가설

– 가설검정 유형: 단측검정과 양측검정

– 가설검정유형과 가설형태

1. 양측검정 $H_0 : \theta = \theta_0$

$H_1 : \theta \neq \theta_0$

2. 단측검정 $H_0 : \theta = \theta_0$ 또는 $\theta \leq \theta_0$

$H_1 : \theta > \theta_0$

3. 단측검정 $H_0 : \theta = \theta_0$ 또는 $\theta \geq \theta_0$

$H_1 : \theta < \theta_0$

가설검정과 가설

- 가설검정과정은 목적은 대립가설이 사실임을 밝힐 수 있는 통계적 증거가 충분히 있는지를 결정하는 것이다

•가설검정을 통해 두 가지 가능한 결론을 도출할 수 있다.

–결론1:귀무가설을 기각(reject)한다[대립가설을 채택(accept)한다]→ 귀무가설을 채택 할 수 없는 충분한 증거가 있다 (귀무가설을 기각할 충분한 증거가 있다)

–결론2: 귀무가설을 기각(reject)하지 못한다 [대립가설을 기각한다]→ 대립가설을 기각(reject) 할 수 있는 충분한 증거가 있다 (귀무가설을 거부 할 수 없는 충분한 증거가 있다)

가설검정과 검정통계치

- 검정통계량(test statistic)

–표본을 이용하여 검정통계량를 산출한다

→ 검정통계량의 표준화 변수= $\dfrac{\text{표본 통계량} - \text{귀무가설에서 설정된 모수값}}{\text{표본 통계량의 표준오차}}$

–귀무가설의 타당설 여부를 결정하는 기준이 된다

–만일 검정통계량의 값이 귀무가설과 일치하지 않으면, 귀무가설이 기각(reject)되어 대립가설이 사실이라고 추론한다

가설검정과 오류

- 가설검정에서 두 가지 가능한 오류가 발생한다:

→ 제1종오류(type I error)와 제2종오류(type II error)

검정결과	실제	
	H_0가 참	H_0가 거짓
채택	옳은 결정 확률 $= 1-\alpha$	제2종 오류 확률 $= \beta$(β 위험)
기각	제1종 오류 확률 $= \alpha$(유의수준)	옳은 결정 확률 $= 1-\beta$(검정력)

- 1종오류: 귀무가설이 사실임에도 불구하고 귀무가설을 기각하는 오류

- 2종오류: 귀무가설이 거짓임에도 불구하고 기각하지 못하는(채택)하는 오류

–오류의 확률: P(Type I error) = α / P(Type II error) = β

(α: 유의수준 *significance level*)

가설검정과 기각역(rejection region)

- 기각역: 가설검정시 귀무가설의 기각여부를 결정하는 기준을 설정하는 영역

–유의수준(α)에 의해 표준정규분포 또는 t-분포을 이용하여 산출되는 임계치(critical value)에 의해 결정된다

–만일 검정통계량이 임계치보다 클 경우(검정통계량이 기각역에 포함), 귀무가설을 기각한다(대립가설을 채택)

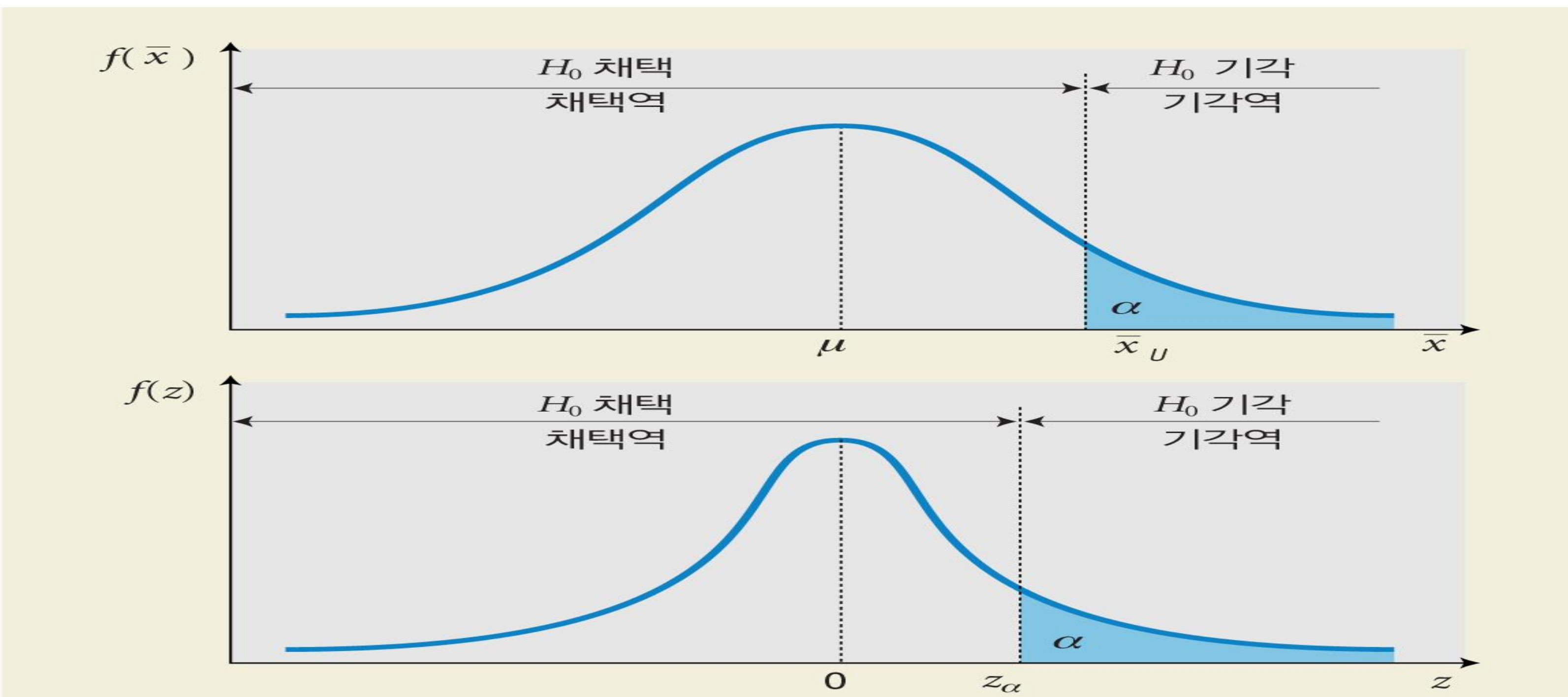
–만일 검정통계량이 임계치보다 작을 경우(검정통계량이 기각역에 불포함), 귀무가설을 기각하지 못한다(귀무가설 채택)

가설검정과 기각역(rejection region)

- 단측검정의 기각역

– H_0: $\mu = x_0$ vs. H_1: $\mu > x_0$
→ 만일 표본평균($\bar{x}$)이 주어진 유의수준(α)에 의해 설정된 특정값($\bar{x}_\alpha$)보다 크면 H_0를 기각

→ 표준화 변수로 전환: 검정통계량이 유의수준(α)에 의해 설정된 임계치(z_α)보다 크면 H_0를 기각



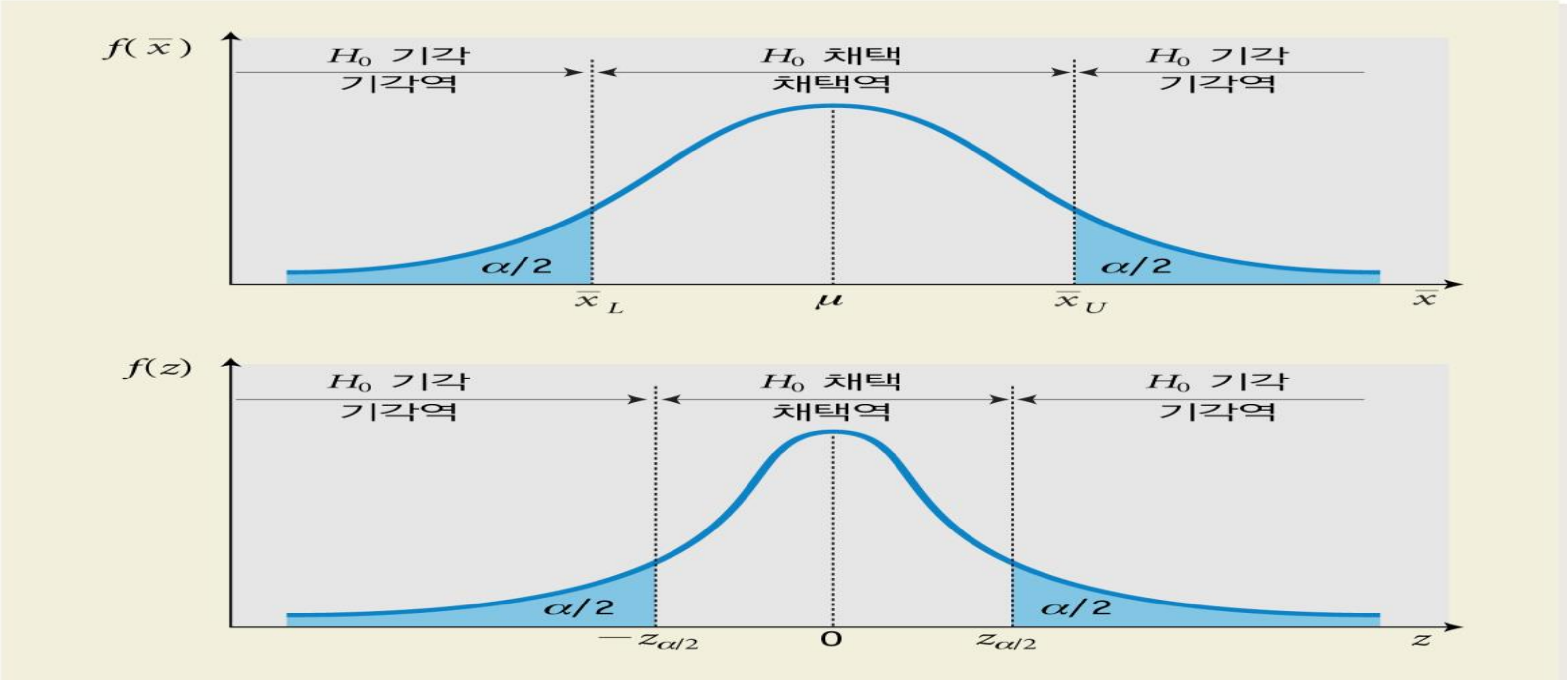
가설검정과 기각역(rejection region)

- 양측검정의 기각역

– H_0: $\mu = x_0$ vs. H_1: $\mu \neq x_0$

→ 만일 표본평균($\bar{x}$)이 주어진 유의수준(α)에 의해 설정된± $\bar{x}_\alpha$ 보다 크면 H_0를 기각

→ 표준화 변수로 전환: 검정통계량이 유의수준(α)에 의해 설정된 임계치(± $z_{\alpha/2}$)보다 크면 H_0를 기각



모집단 평균에 대한 가설검정

- 모집단이 정규분포하고 모분산을 아는 경우

표본평균의 값은 $\bar{x}$이며 가설검정의 유의수준은 α이다.

1. 가설 설정 $H_0 : \mu = \mu_0,\ H_1 : \mu > \mu_0$

   검정통계량 $Z = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$

   결정규칙－검정통계량이 임계값 z_α보다 크면 H_0를 기각

2. 가설 설정 $H_0 : \mu = \mu_0,\ H_1 : \mu < \mu_0$

   검정통계량 $Z = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$

   결정규칙－검정통계량이 임계값 $-z_\alpha$보다 작으면 H_0를 기각

3. 가설 설정 $H_0 : \mu = \mu_0,\ H_1 : \mu \neq \mu_0$

   검정통계량 $Z = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$

   결정규칙－검정통계량이 임계값 $z_{\alpha/2}$보다 크거나 $-z_{\alpha/2}$보다 작으면 H_0를 기각

자주 사용되는 임계값 : $z_{0.10} = 1.283,\ z_{0.05} = 1.645,\ z_{0.025} = 1.960,\ z_{0.01} = 2.326$

예시1)
공장에서 기계가 멈추는 횟수가 정규분포를 따르고 표준편차가 1이었다. 25명의 기능공을 대상으로 조사한 평균회수가 4.5로 나타났다. 반면에 기계가 멈추는 전체 회수가 평균 5회라는 주장이 있다. 이에 대해 5회보다는 작다는 대립가설에 대한 검정을 유의수준 5%에서 실시하시오.

1. 가설설정: $H_0 : \mu = 5$ $H_1 : \mu < 5$

2. 유의수준: α=0.05

3. 임계치: $Z_a = Z_{0.05} = 1.645$

4.검정통계량: $$Z = \frac{\overline{X} - \mu}{\sigma / \sqrt{N}} = \frac{4.5 - 5}{1 / \sqrt{25}} = -2.5$$

$$\therefore Z = -2.5 < -Z_a = -Z_{0.05} = -1.645$$

따라서 귀무가설 기각(전체평균회수가 5회보다 작다)

예시2) 공장에서 기계가 멈추는 횟수가 정규분포를 따르고 표준편차가 1이었다. 25명의 기능공을 대상으로 조사한 평균회수가 5.5로 나타났다. 반면에 기계가 멈추는 전체 회수가 평균 5회라는 주장이 있다. 이에 대해 5회보다는 크다는 대립가설에 대한 검정을 유의수준 5%에서 실시하시오.

1. 가설설정: H_0 : $\mu = 5$ H_1 : $\mu > 5$

2. 유의수준: $\alpha = 0.05$

3. 임계치: $Z_a = Z_{0.05} = 1.645$

4.검정통계량:

$$Z = \frac{\overline{X} - \mu}{\sigma / \sqrt{N}} = \frac{5.5 - 5}{1 / \sqrt{25}} = 2.5$$

$$\therefore Z = 2.5 > Z_a = Z_{0.05} = 1.645$$

따라서 귀무가설 기각(전체평균회수 5회보다 크다)

예시3) 공장에서 기계가 멈추는 횟수가 정규분포를 따르고 표준편차가 1이었다. 25명의 기능공을 대상으로 조사한 평균회수가 6.0로 나타났다. 반면에 기계가 멈추는 전체 회수가 평균 5회라는 주장이 있다. 이에 대해 5회는 아니다는 대립가설에 대한 검정을 유의수준 5%에서 실시하시오.

1. 가설설정: $H_0 : \mu = 5$ $H_1 : \mu \neq 5$

2. 유의수준: α=0.05

3. 임계치: $Z_a = Z_{0.05/2} = Z_{0.025} = 1.96$

4.검정통계량: $$Z = \frac{\overline{X} - \mu}{\sigma / \sqrt{N}} = \frac{6-5}{1/\sqrt{25}} = 5$$

$$\therefore Z = |5| > \pm Z_{a/2} = \pm Z_{0.05/2} = \pm 1.960$$

따라서 귀무가설 기각(전체평균회수 5회가 아니다)

모집단 평균에 대한 가설검정

- 모집단의 분포는 모르나 모분산을 알고 표본크기가 큰 경우

→ 표본평균이 중심극한 이론에 의해 정규분포를 따른다

표본평균의 값은 $\bar{x}$ 이며 가설검정의 유의수준은 α이다.

1. 가설 설정 $H_0 : \mu = \mu_0,\ H_1 : \mu > \mu_0$

   검정통계량 $Z = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$

   결정규칙 — 검정통계량이 임계값 z_α보다 크면 H_0를 기각

2. 가설 설정 $H_0 : \mu = \mu_0,\ H_1 : \mu < \mu_0$

   검정통계량 $Z = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$

   결정규칙 — 검정통계량이 임계값 $-z_\alpha$보다 작으면 H_0를 기각

3. 가설 설정 $H_0 : \mu = \mu_0,\ H_1 : \mu \neq \mu_0$

   검정통계량 $Z = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$

   결정규칙 — 검정통계량이 임계값 $z_{\alpha/2}$보다 크거나 $-z_{\alpha/2}$보다 작으면 H_0를 기각

자주 사용되는 임계값 : $z_{0.10} = 1.283,\ z_{0.05} = 1.645,\ z_{0.025} = 1.960,\ z_{0.01} = 2.326$

모집단 평균에 대한 가설검정

- 정규분포를 갖는 모집단의 모분산을 모를 경우

표본평균값은 $\bar{x}$이며 가설검정의 유의수준은 α이다.

1. 가설 설정 $H_0 : \mu = \mu_0$ 혹은 $\mu \leq \mu_0$

   $H_1 : \mu > \mu_0$

   검정통계량 $t_{n-1} = \dfrac{\bar{x} - \mu_0}{S/\sqrt{n}}$

   결정규칙−검정통계량이 $t_{(n-1,\ \alpha)}$보다 크면 H_0를 기각

2. 가설 설정 $H_0 : \mu = \mu_0$ 혹은 $\mu \geq \mu_0$

   $H_1 : \mu < \mu_0$

   검정통계량 $t_{n-1} = \dfrac{\bar{x} - \mu_0}{S/\sqrt{n}}$

   결정규칙−검정통계량이 임계값 $-t_{(n-1,\ \alpha)}$보다 작으면 H_0를 기각

3. 가설 설정 $H_0 : \mu = \mu_0$

   $H_1 : \mu \neq \mu_0$

   검정통계량 $t_{n-1} = \dfrac{\bar{x} - \mu_0}{S/\sqrt{n}}$

   결정규칙−검정통계량이 임계값 $t_{(n-1,\ \alpha/2)}$보다 크거나 $-t_{(n-1,\ \alpha/2)}$보다 작으면 H_0를 기각

t-분포를 이용한 가설검정: 예시

금융전문가들은 우리나라 금융기관들의 평균 배당률이 10%라고 예상한다. 이 예상이 타당한지 알아보기 위해 금융기관 40개를 표본추출하여 배당률을 계산해 보니 평균은 9.3%, 표준편차는 3%였다. 우리나라 금융기관들의 평균 배당률이 10%라는 전문가들의 주장이 타당한지 알아보기 위해 0.10의 유의수준에서 평균 배당률이 10% 미만이라는 대립가설을 가지는 단측검정과 평균 배당률이 10%가 아니라는 양측검정을 하라.

t-분포를 이용한 가설검정: 예시

- 단측검정: 금융기관들의 평균배당률이 정규분포를 따른다

1. 가설설정: $H_0 : \mu = 10$ $H_1 : \mu < 10$

2. 유의수준: $\alpha=0.10$

3. 임계치: $t_{(n-1,a)} = t_{(39,0.1)} = 1.303$

4. 검정통계량: $t = \dfrac{\overline{X} - \mu}{s/n} = \dfrac{9.3 - 10}{3/\sqrt{40}} = -1.4757$

$$\therefore t = -1.4757 < -t_{(n-1,a)} = -t_{(39,0.1)} = -1.303$$

따라서 귀무가설 기각(평균배당률이 10%미만이다)

t-분포를 이용한 가설검정: 예시

- 양측검정: 금융기관들의 평균배당률이 정규분포를 따른다

1. 가설설정: $H_0 : \mu = 10$ $H_1 : \mu \neq 10$

2. 유의수준: $\alpha=0.10$

3. 임계치: $t_{(n-1,a/2)} = t_{(39,0.1/2)} = 1.684$

4. 검정통계량: $t = \dfrac{\overline{X} - \mu}{s/n} = \dfrac{9.3 - 10}{3/\sqrt{40}} = -1.4757$

$$\therefore t = |1.4757| < \pm t_{(n-1,a/2)} = \pm t_{(39,0.1/2)} = \pm 1.684$$

따라서 귀무가설 기각하지 못함(평균배당률이 10%이다)

표본통계량에 대한 유의성 검증 (significance test)

- 유의성 검증:표본통계량의 통계적 유의성(statistical significance)을 파악하기 위한 검증

- 모평균의 추정량인 표본평균의 유의성 검증성
  - 가설: H_0: $\mu = 0$ vs. H_1: $\mu \neq 0$

→검정통계량: $z = \dfrac{\overline{X}}{\sigma/\sqrt{n}} \,/\, t = \dfrac{\overline{X}}{s/\sqrt{n}}$

→ 임계치: $z_{a/2} \,/\, t_{(n-1,\alpha/2)}$

→ 검정통계량와 임계치를 비교하여 귀무가설이 기각 여부 판정

표본통계량에 대한 유의성 검증 (significance test)

- 검정통계량와 임계치를 비교하여 귀무가설이 기각되는 경우, 모평균값이 0이 아닌다고 주장하는 대립가설이 사실이므로, 모평균의 추정량인 표본평균이 통계적으로 유의하다(statistically significant)는 의미를 갖고 있다

- 검정통계량와 임계치를 비교하여 귀무가설이 기각되지 못하는 경우, 모평균값이 0이 아닌다고 주장하는 대립가설이 사실이 아니므로, 모평균의 추정량인 표본평균이 통계적으로 유의하지 못하다(statistically insignificant)는 의미를 갖고 있다

유의성 검정: 예시

- Ex: 금융기관 평균배당률의 유의성 검증

1. 가설설정: $H_0 : \mu = 0 \quad H_1 : \mu \neq 0$

2. 유의수준: $\alpha=0.10$

3. 임계치: $t_{(n-1,a/2)} = t_{(39,0.1/2)} = 1.684$

4. 검정통계량: $t = \dfrac{\overline{X}}{s/\sqrt{n}} = \dfrac{9.3}{3/\sqrt{40}} = 19.62$

$$\therefore t = |19.60| > \pm t_{(n-1,a/2)} = \pm t_{(39,0.1/2)} = \pm 1.684$$

따라서 귀무가설 기각(표본평균 9.3%는 통계적으로 유의하다)

가설검정과 p-Value(유의확률)

- 가설검정을 위한 다른 방법은 *p-value* 를 이용하는 것이다

가설검정에서 귀무가설이 기각될 수 있는 최소 유의수준을 P-값이라고 한다. 즉 P-값은 귀무가설이 참일 때 검정통계량이 계산된 검정통계량의 값보다 더 극단에 위치할 확률이다.

-ex) 정규분포인 모집단(모표준편차:65)에서 표본(n=400)을 추출하여 표본평균이 178을 구했다. 이를 토대로 모평균에 대한 아래 가설을 검정하고자 한다

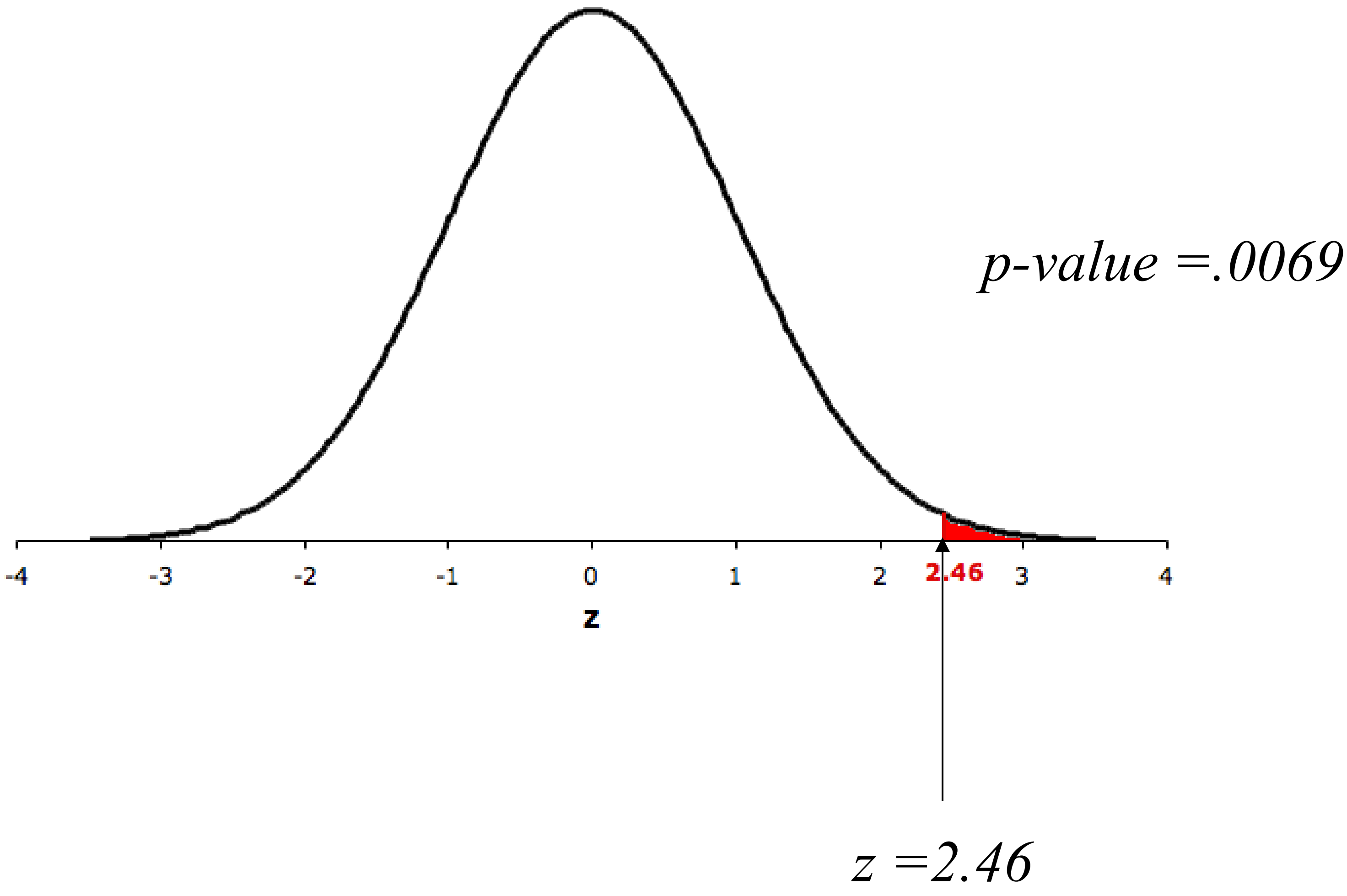
→ H_0: $\mu = 170$ vs. H_1: $\mu > 170$

→ p-value : $$P\left(\frac{\bar{x}-\mu}{\sigma/\sqrt{n}} > \frac{178-170}{65/\sqrt{400}}\right) = P(Z > 2.46) = .0069$$

→ 귀무가설(H_0)을 기각하기 위해서는 유의수준(α)이 최소한 0.0069 이상 이어야 한다 : p-value = 0.0069 이다

가설검정과 p-Value

p-value = $P(Z > 2.46)$



가설검정과 p-Value

- p-value를 이용한 가설검정은 유의수준(α)과p-value 를 비교함으로써 가능하다

– 만일 p-value가 α 보다 작을 경우, 귀무가설을 기각한다

– 만일 p-value가 α 보다 클 경우, 귀무가설을 기각하지 못한다

→ex) 위의 예시에서 p-value = .0069 이므로

1) 유의수준(α) = .05(5%)에서 귀무가설을 기각

2) 유의수준 (α) = 0.1(10%)에서 귀무가설 기각

가설검정의 결론

- 만일 가설검정 결과 귀무가설을 기각하면, 대립가설이 사실임을 보여주는 통계적 증거가 있음을 나타내는 것이다.

- 반면에, 귀무가설을 기각하지 못하면, 대립가설이 사실임을 보여주는 통계적 증거가 없음을 나타내는 것이다

- <u>주의:</u>귀무가설보다는 대립가설이 보다 중요하다. – 대립가설은 실제조사결과를 나타내고 있기때문..