Chap 9. 모평균과 모비율에 관한가설검정

Ho Sun Shon

목차

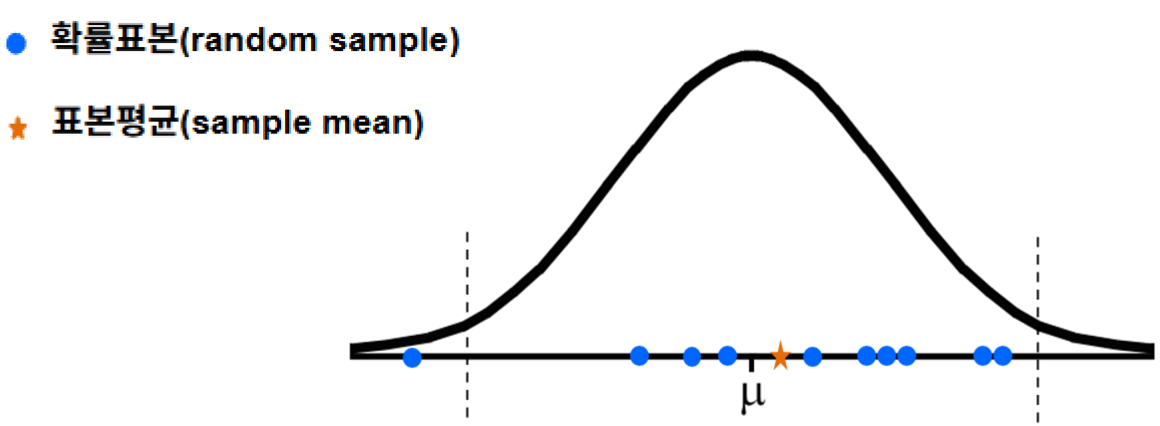
- 가설검정의 문제
- 귀무가설과 대립가설
- 제1종 오류와 제2종 오류
- 검정통계량과 기각력
- 모평균μ에 대한 가설검정: 표본의 크기가 큰 경우
- 유의확률$(p-value)$
- 모평균μ에 대한 가설검정: 표본의 크기가 작은 경우
- 모비율 p에 관한 가설검정

가설검정 문제

➢ 통계적 가설검정의 예

❖ 새 공정으로 생산된 전구의 평균수명을 μ라 한다면 μ가 75 시간보다 길다고 주장한다. 이 주장에 대해 표본을 통하여 어떻게 확인할 수 있는가?

- 평균에 관한 기존의 결과를 반영한 주장
- 상황의 변화에 따른 새로운 주장



모평균이 표본평균과 다르다고 할 수 있는가?

통계적 가설

- 통계적 가설
  - 통계적 가설은 모집단의 특성 값에 대한 주장
    - 기존 결과에 대한 주장과 상황의 변화에 따른 새로운 주장으로 구성
  - 표본분포의 성질을 이용하여 모집단 특성에 대한 가설의 진위를 가리는 것
  - 통계적 가설검정은 재판과정과 마찬가지로 간접적인 방법 즉, 주장(대립가설)과 반대되는 가설(귀무가설)을 사실로 가정한 상태에서 검정
  - 표본의 결과가 "귀무가설 H_0가 사실이 아니라는 증거"를 얼마나 충분히 제시하느냐에 따라 가설검정이 이루어짐

통계적 가설

- 통계적 가설검정의 예
  - 새 공정으로 생산된 전구의 평균수명을 μ라 한다면 μ가 75 시간보다 길다고 주장
  - 회사는 새 공정으로 생산된 전구의 평균수명이 75시간보다 길면 새 공정으로 전구를 생산하고, μ가 75시간과 같으면 기존 공정을 유지
    - 가설1: 새 공정으로 생산된 전구의 평균수명 μ가 75 시간과 같다($\mu = 75$)
    - 가설2: 새 공정으로 생산된 전구의 평균수명을 μ라 한다면 μ가 75 시간보다 길다($\mu > 75$)
  - 제약회사의 신약 개발
    - 가설1: 새로운 약품의 치료율 p가 기존 약품의 치료율과 같다($p = 0.3$)
    - 가설2: 새로운 약품의 치료율 p가 기존 약품의 치료율보다 크다($p > 0.3$)

통계적 가설의 종류

- 귀무가설(null hypothesis: H_0):
  - **기존의 결과 반영한 주장**
  - 영가설 이라고도 함
  - 대립가설이 기각되어 귀무가설이 채택되는 결과가 발생 하면 연구자의 입장에서는 얻는 것 없이 무위로 돌아가게 됨
- 대립가설(alternative hypothesis: H_1):
  - **상황의 변화에 따른 새로운 주장을 반영**
  - 기존의 결과에 대항하여 설립되는 가설, 연구가설 이라고도 함

가설의 설립

- 형식: 귀무가설(H_0) 대(against, vs.) 대립가설(H_1)
  - 예: H_0: $\mu = 3.2$(기존의 결과) $vs. H_1$: $\mu \neq 3.2$ (새로운 주장)
- 전통적 기본형식의 통계적 가설의 설립
  - 귀무가설과 대립가설이 중첩되지 않도록 설립
  - 귀무가설은 단순가설로 설립하고, 대립가설은 복합가설로 설립
  - 예: H_0: $\mu = 3.2$ $vs.$ H_1 : $\mu \neq 3.2$ (또는 $\mu < 3.2$ 또는 $\mu > 3.2$)
- 가설 별 주장의 반영 특성
  - 귀무가설: No Difference (or No Effect)
  - 대립가설: Difference (또는 Effect)

귀무가설과 대립가설

- 대립가설(alternative hypothesis: H_1)
  - 연구자가 표본을 통해 입증하려고 하는 모수에 대한 주장
- 귀무가설(null hypothesis: H_0)
  - 대립가설과 반대되는 모수에 대한 주장으로 기존부터 유지되어 왔던 보수적인 주장

귀무가설과 대립가설

➢ 가설 설정

❖ H_0: 새 공정으로 생산된 전구의 평균수명 μ가 75 시간과 같다**(μ=75)**

❖ H_1: 새 공정으로 생산된 전구의 평균수명을 μ라 한다면 μ가 75 시간보다 길다**(μ>75)**

❖ H_0: 새로운 약품의 치료율 p가 기존 약품의 치료율과 같다**(p=0.3)**

❖ H_1: 새로운 약품의 치료율 p가 기존 약품의 치료율보다 크다**(p>0.3)**

❖ 귀무가설(H_0): 피고인은 무죄다

❖ 대립가설(H_1): 피고인은 유죄다

가설의 검정

- 귀무가설과 대립가설 중 어느 것이 더 타당한지 표본에 의한 통계량의 분포에 의하여 결정하는 절차
- 타당하지 않은 가설은 기각(reject)하고, 다른 하나는 채택(accept)하게 됨
- 대립가설의 채택은 귀무가설의 기각을 의미
- 대립가설의 기각은 귀무가설의 채택을 의미

검정의 오류

- 제1종 오류(Type I error)
  - 귀무가설이 참인데 귀무가설을 기각하는 오류
- 제2종 오류(Type II error)
  - 대립가설이 참인데 대립가설을 기각하는 오류
- 가설 검정의 결과

결정 사실	H_0 채택(H_1 기각)	H_0 기각(H_1 채택)
H_0 참(H_1 거짓)	옳은 결정	잘못된 결정 (제 1종 오류)
H_0 거짓(H_1 참)	잘못된 결정 (제 2종 오류)	옳은 결정

검정의 오류

➢ 법정의 상황을 통계적 가설의 형태로 적용

❖ 귀무가설: 피고인은 무죄(자연인), 대립가설: 피고인은 유죄(범죄와 연루 등)

❖ 제1종 오류: 무죄인 피고인에게 유죄판결(처벌의 오류 발생)

❖ 제2종 오류: 유죄인 피고인에게 무죄 판결(무처벌의 오류 발생)

⇒무처벌의 오류보다 처벌의 오류를 더 심각하게 봄(무죄추정의 원칙)

casenote.kr
https://casenote.kr › 대법원 › 2005모472

대법원 2009. 7. 16.자 2005모472 전원합의체 결정 [재심기각 ...

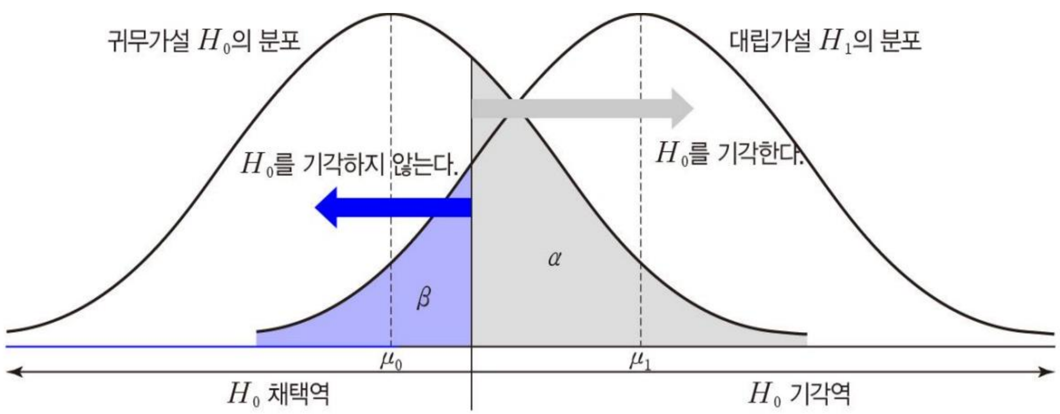
'10명의 범인을 놓치더라도 1명의 무고한 사람을 벌하여서는 아니 된다'는 법언(法諺)이 상징하는, 실체적 진실발견과 인권보호라는 명제는 민주적 형사절차의 ...

⇒검정 오류에서도 제1종 오류의 발생을 제2종 오류의 발생보다 더 심각하게 간주

오류의 크기

- $\alpha = P(type\ I\ error) = P(\boldsymbol{H_0}$가 참인데 $\boldsymbol{H_0}$를 기각$) = P(H_0$기각$|H_0$참$)$
- $\beta = P(type\ II\ error) = P(\boldsymbol{H_1}$가 참인데 $\boldsymbol{H_1}$를 기각$) = P(H_1$기각$|H_1$참$)$

- α와 β의 관계



❖ 둘 중 하나를 줄이면 다른 하나가 상승 => 상충효과 발생

❖ α와 β를 동시에 줄이기 위해서는 표본의 크기 n 을 증가시켜야 함

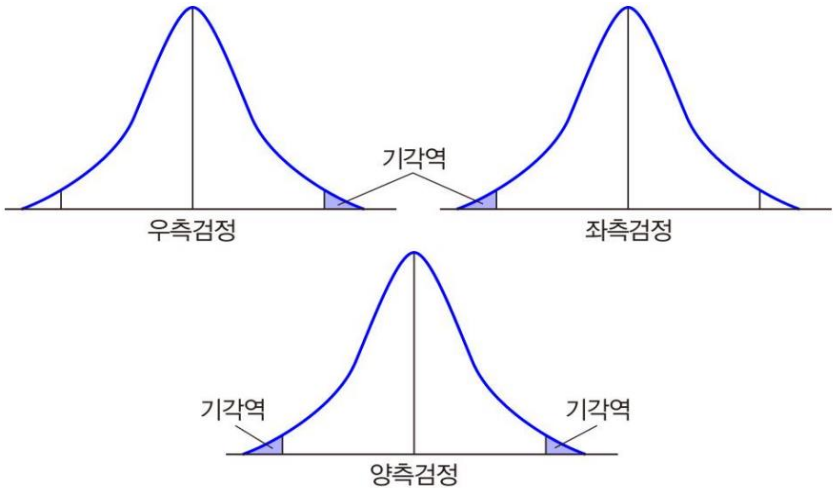
가설 검정의 원칙

- 적용 가설: 귀무가설($\boldsymbol{H_0}$): 단순 vs. 대립가설($\boldsymbol{H_1}$): 복합
- 조건: 표본 크기 n을 고정
- 기각역의 설정 원칙
  - 원칙1 : $\alpha + \beta$를 최소화
  - 원칙2 : 고정된 α에 대하여 β를 최소화(또는 $1 - \beta$를 최대화)
    - $1 - \beta$를 검정력(power of test) 라 함
    - 오류 중 더 심각한 것(제1종 오류)의 최대허용 한계를 정하여 통제함
  - 전통적인 방법은 원칙2를 따르며, 이때 **고정된 $\boldsymbol{\alpha}$를 유의수준 (significance level)** 이라 함.
    - 고정된 α의 값을 0.1, 0.05, 0.01 가 주로 사용됨(0.05는 Fisher에 의해 제안됨)

기각역 (Critical region)

- 앞면과 뒷면이 나올 확률이 같은 동전을 10번 던진다고 생각하여 보자. 이런 확률 실험에 대한 귀무가설은 '앞면과 뒷면이 나올 확률은 같다' 이다. 그런데 20번 확률 시행 결과 모두 앞면이 나왔다면 어떨까?
  - 귀무가설: 앞면과 뒷면이 나올 확률이 같다.
  - 20번 가까이 모두 앞면이 나올 확률: **극단적 상황의 확률은 극히 작다**
  - 어떤 가설에서 거의 우연히 일어날 수 없는 일이 일어나면 그 가설은 쓸모가 없고 문제는 있을 수 있는 것과 없는 것의 **경계 설정**
  - 거의 있을 수 없는 일을 간주하는 범위를 **기각역(critical region)**

기각역 (단측 검정, 양측 검정)



- 단측검정(One-sided test)
  - 우측검정: Ho: $\mu = \mu_0$, H_1: $\mu > \mu_0$
  - 좌측검정: Ho: $\mu = \mu_0$, H_1: $\mu < \mu_0$
- 양측검정(Two-sided test)
  - Ho: $\mu = \mu_0$, H_1: $\mu \neq \mu_0$

- 기각역으로 표현하는 일반적인 가설검정
  - 귀무가설 Ho: $\mu = \mu_0$의 검정통계량은 $Z = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$ 이고, z의 관측값 z라 할 때 유의수준 α 에서 기각역

검정통계량 (Test Statistic)

- 검정통계량(Test Statistic)
  - 귀무가설을 기각할 수 있는지 판단하는 기준 통계량, 즉 귀무가설과 대립가설 중의 하나를 선택하는 기준으로 사용되는 통계량
  - 모수의 점추정량에 근거하여 표본으로부터 계산된 통계량
    - 모평균 μ를 검정하기 위한 검정통계량: 표본의 크기가 클 때
    - 구무가설(H_0)의 기각여부를 검정하는데 사용하는 검정통계량으로 Z를 사용

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

- 기각역(Rejection region)
  - 귀무가설 H_0를 기각시키는 검정통계량의 관측값의 영역

기각역: 우측 검정(right-tailed test)

- 우측검정(right-tailed test)
  - 대립가설의 내용이 우측방향의 서술로 구성되는 경우
  - 예: 새 공정으로 생산된 전구의 평균수명을 μ라 한다면 μ가 75 시간보다 길다($\boldsymbol{Ho}$: $\mu = \mathbf{75}, \boldsymbol{H_1}$: $\mu > \mathbf{75}$)
- 검정 통계량과 유의수준 α 가 주어졌을 때 귀무가설의 **기각영역**
  - 전구의 예에서 Ho: $\mu = 75, H_1$: $\mu > 75$ 를 검정하기 위한 유의수준 $\alpha = 0.05$, 새로운 공정으로 생산된 전구의 전구수명의 모집단은 $\sigma^2 = 100$ 인 정규분포를 따른다고 가정하고, 임의로 추출한 25개의 전구들의 표본평균의 표본오차는 2가 될 때 기각역을 구해보자.

기각역: 우측검정(right-tailed test)

➤ 우측검정의 기각역 계산

❖ $0.05 = \alpha$

$= P(\text{옳은 } H_0\text{을 기각한다})$

$= P(Z > z^* | \mu_0 = 75)$

$= P\left(Z > \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \middle| \mu_0 = 75\right)$

$= P\left(Z > \frac{\bar{X}-75}{10/\sqrt{25}}\right)$

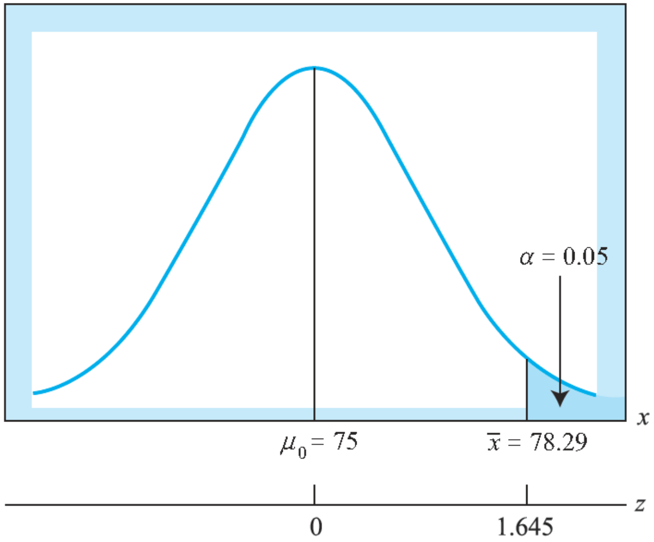


그림 9.1 H_0: $\mu = 75$와 H_1: $\mu > 75$에서 $\alpha = 0.05$일 때 기각역

(1) 검정통계량 Z의 값을 계산한 결과가 $z_{0.05} = 1.645$보다 크게 나타났을 때

(2) 임의로 선택된 25개 전구들의 평균수명 $\bar{X}$가 78.29시간보다 크게 나타났을 때

기각역: 좌측검정(left-tailed test)

- 좌측검정(left-tailed test)
  - 기각역이 왼쪽에 나타나는 단측검정
  - 기존의 공정으로 생산되는 전구의 평균수명은 μ_0 로, 75시간이라고 가정하자(H_0: μ = 75). 그런데 생산 현장에서 작업자의 실수로 공정에 문제가 발생하여 이제부터 생산되는 전구들의 평균 수명 μ는 75시간보다 짧다고 함 (H_1: μ < 75)
  - 유의수준이 0.05로 주어졌고, 모집단은 분산이 σ^2인 정규분포를 따른다고 가정
  - 임의로 추출된 25개의 전구들의 평균수명도 현저히 짧게 나타나서 수명이 감소되었다는 증거가 되고 대립가설 H_1: μ < 75를 지지하게 되므로 Ho의 기각역은 μ_0 를 중심으로 왼쪽에 있음

기각역: 좌측검정(left-tailed test)

➢ 좌측검정 기각역 계산

$$\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} = -z_{0.05}$$

$$\bar{x} = \mu_0 - z_{0.05}\frac{\sigma}{\sqrt{n}}$$

$$= 75 - \left(1.645 \times \frac{10}{\sqrt{25}}\right) = 71.71$$

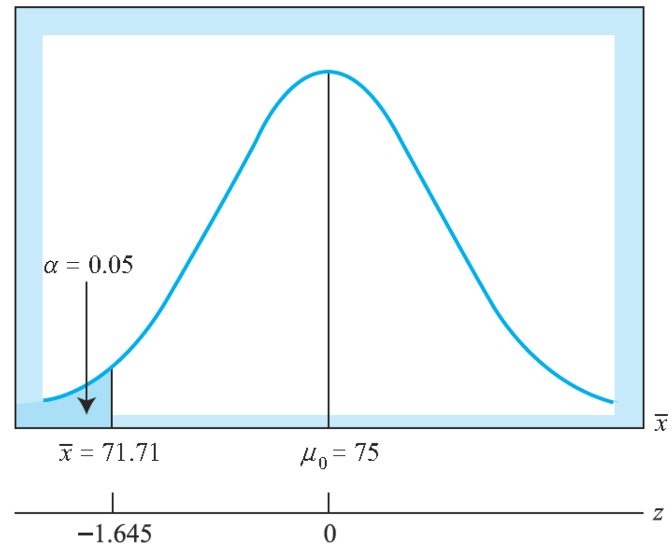


그림 9.2 H_0: $\mu = 75$와 H_1: $\mu < 75$에서 $\alpha = 0.05$일 때 기각역

(1) 검정통계량 Z의 값을 계산한 결과 $-z_{0.05} = -1.645$보다 작게 나타났을 때

(2) 임의로 선택된 25개 전구들의 평균수명 $\bar{X}$가 71.71시간보다 작게 나타났을 때

기각역: 양측검정(two-sided test)

- 양측검정(Two-sided test)
  - 기각역이 양쪽에 나타나는 검정
    - 기존의 공정으로 생산되는 전구의 평균수명 μ_0 로, 75시간이라고 가정 (H_0: $\mu = 75$),
    - 생산기계들의 부품 마모가 심각하여 부품들을 교환한 결과, 생산되는 전구수명의 모평균 μ 는 75시간과 차이가 있다고 함(H_1: $\mu \neq 75$).
  - $\alpha = 0.05$로 주어졌고, 모집단은 분산이 σ^2인 정규분포를 가정
  - 전구들의 수명이 기존의 전구들의 평균수명과 차이가 난다면,
  - 임의로 선택된 25개 전구들의 평균수명 $\bar{X}$도 기존의 전구 들의 평균수명 μ 와 달리 충분히 작거나 또는 크게 나타나서 기존의 전구수명과는 차이가 있다는 증거, 대립가설인 H_1 : $\mu \neq 75$ 를 지지하게 됨.
  - 따라서 H_0 의 기각역은 μ_0 를 중심으로 왼쪽이나 오른쪽에 있게 됨

기각역: 양측검정(Two-sided test)

- 양측검정(Two-sided test) 기각역 계산
  - 기각역이 양쪽에 나타남
  - 생산된 전구의 모평균 μ는 75시간과 차이가 있다. (H_1: μ≠75)

$$\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} = \pm z_{0.05}$$

$$\bar{x} = \mu_0 \pm z_{0.025}\frac{\sigma}{\sqrt{n}}$$

$$= 75 \pm \left(1.96 \times \frac{10}{\sqrt{25}}\right) = (71.08, 78.92)$$

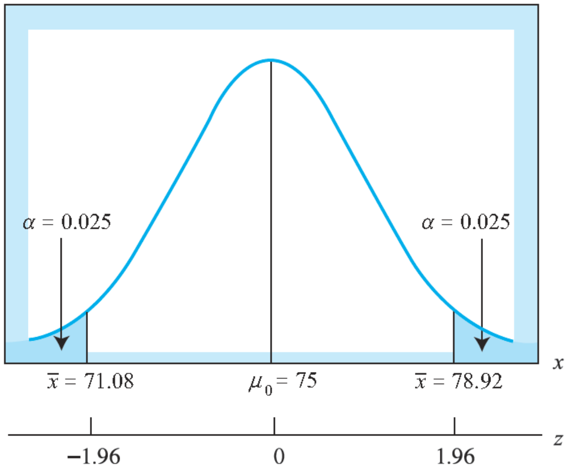


그림 9.3 H_0: $\mu = 75$와 H_1: $\mu \neq 75$에서 $\alpha = 0.05$일 때 기각역

(1) 검정통계량 Z의 값을 계산한 결과 $|z_{0.05}| = 1.96$보다 크게 나타났을 때
(2) 임의로 선택된 25개 전구들의 평균수명 $\bar{X}$가 71.08 보다 작게 나타나거나 78.92보다 크게 나타났을 때

가설검정절차

- 가설검정 절차

1. 가설 설정: $\boldsymbol{Ho}$: $\mu = \boldsymbol{\mu_0}$, $\boldsymbol{H_1}$: $\mu > \boldsymbol{\mu_0}$ ($\mu < \boldsymbol{\mu_0}$, $\mu \neq \boldsymbol{\mu_0}$)
2. 검정통계량 결정

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

3. 유의수준 α 결정 ($\alpha = 0.01$, $\alpha = 0.05$, $\alpha = 0.1$)
4. 판정 기준 결정: 기각역 결정
5. 검정통계량 값 계산 후 귀무가설($\boldsymbol{Ho}$)의 기각 여부 판정
6. 연구자의 주장에 대한 결론

모평균 μ에 대한 가설검정: 표본의 크기가 클 경우

- 모평균 μ를 검정하기 위한 검정통계량: 표본의 크기가 $n \geq 30$ 일 때

1) 검정통계량의 결정: 귀무가설의 기각 여부를 검정

$$\sigma \text{ 를 알 때}: Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}, \qquad \sigma\text{를 모를 때}: Z = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$$

2) 대립가설에 대하여 유의수준 α 일 때 기각역

- H_0: $\mu = \mu_0$ vs. H_1: $\mu > \mu_0$일 때, $R: Z > z_\alpha$ (우측검정)
- H_0: $\mu = \mu_0$ vs. H_1: $\mu < \mu_0$일 때, $R: Z < -z_\alpha$(좌측검정)
- H_0: $\mu = \mu_0$ vs. H_1: $\mu \neq \mu_0$일 때, $R: |Z| > z_{\alpha/2}$ (양측검정)

예제 9.1

- 어떤 타이어 제조회사에서 생산 중인 타이어의 평균수명이 37,000km로 알려져 있다. 이 회사의 연구개발팀에서는 타이어 수명을 연장시키기 위해서 새로운 공정을 개발하였다. 새공정에 의해 생산된 타이어 시제품 n=100 개의 수명을 측정한 결과, 평균수명이 $\bar{x}$ =37,500km이고 표준편차 s=5,000 였다. 이 표본 자료를 근거로 했을 때, 새로 개발된 공정이 성공적이라고 할 수 있는지를 유의수준 1%에서 검정하여라.
- 풀이:
  - 가설 설정: $H_0: \mu = 37{,}000 \quad vs. \quad H_1: \mu > 37{,}000$
  - 검정통계량 결정: 모표준편차를 모르기 때문에 S이용

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