제 9장

가설검정

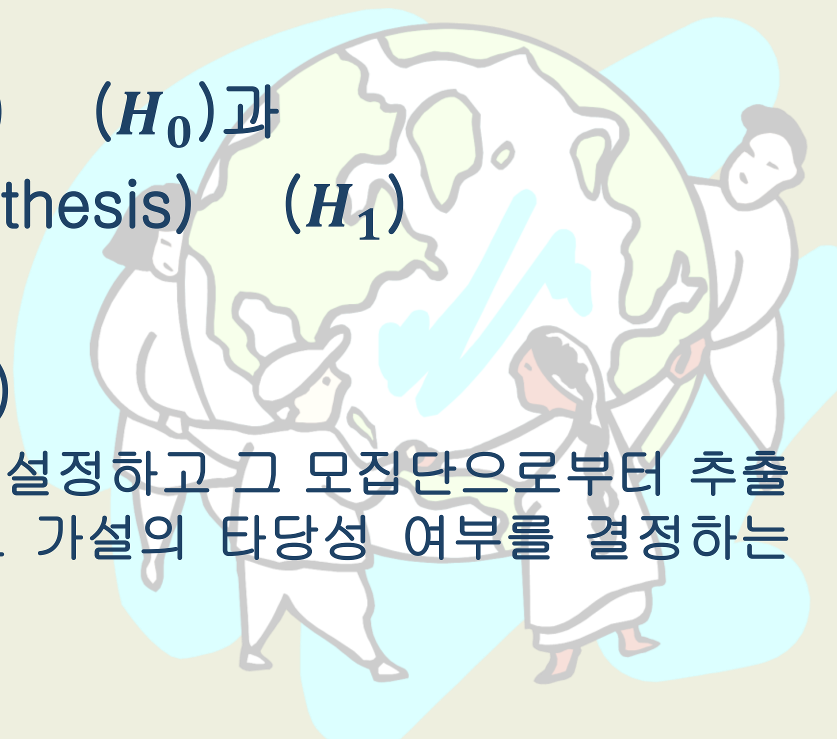
고려대학교 경영대학 박 광태

가설 검정의 개념

- 가설(hypothesis)
  - 과거의 경험, 지식, 연구의 결과 등으로 모수가 취할 것으로 알려진 값을 서술한 것

- 귀무가설(null hypothesis) (H_0)과 대립가설(alternative hypothesis) (H_1)

- 가설검정(hypothesis test)
  - 모집단에 대한 어떤 가설을 설정하고 그 모집단으로부터 추출된 표본을 분석함으로써 그 가설의 타당성 여부를 결정하는 것.



가설 검정의 개념

- 단순가설과 복합가설
  - 단순가설 : 단일치에 의하여 모수의 값을 서술
  - 복합가설 : 여러 개의 값 또는 실수구간에 의하여 모수의 값을 서술
- 검정통계량
  - 귀무가설의 기각여부를 결정하는데 사용되는 통계량
  - 모평균의 경우 표본평균, 모비율의 경우 표본비율, 모분산의 경우 표본분산이 검정통계량이 됨
- 기각치(critical value) : 기각영역과 채택영역을 분리시켜주는 값.
- 기각역: H_0 이 기각되는 검정통계량의 모든 값.
  - [그림 9-1] 참조

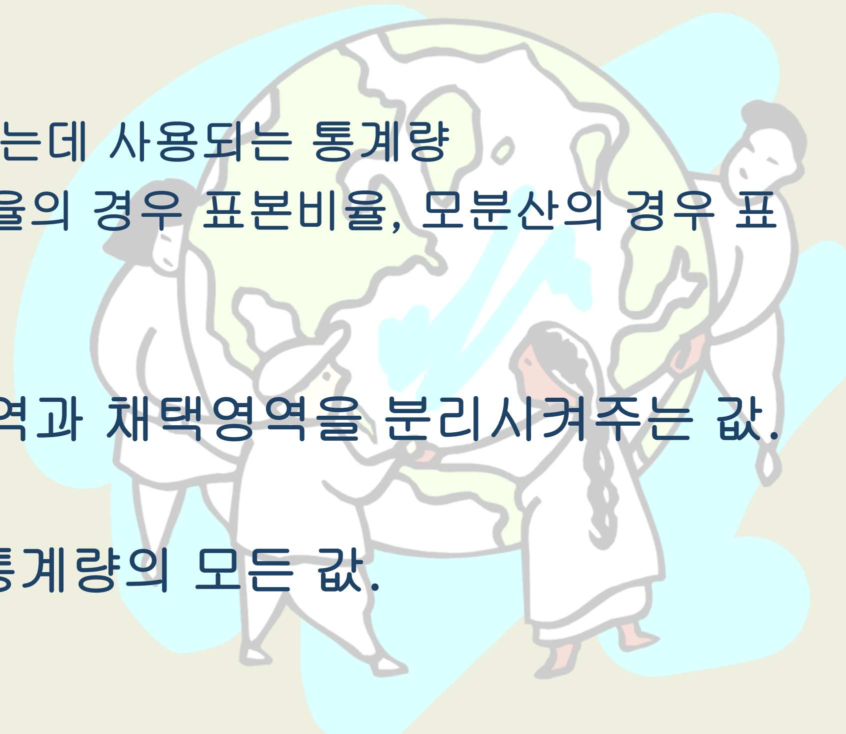
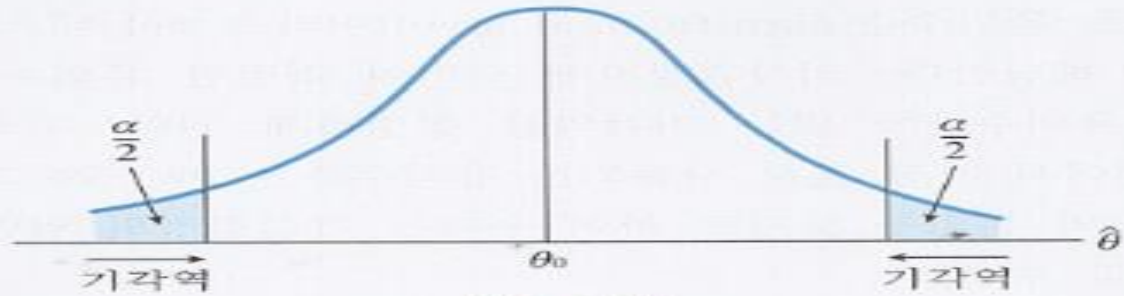


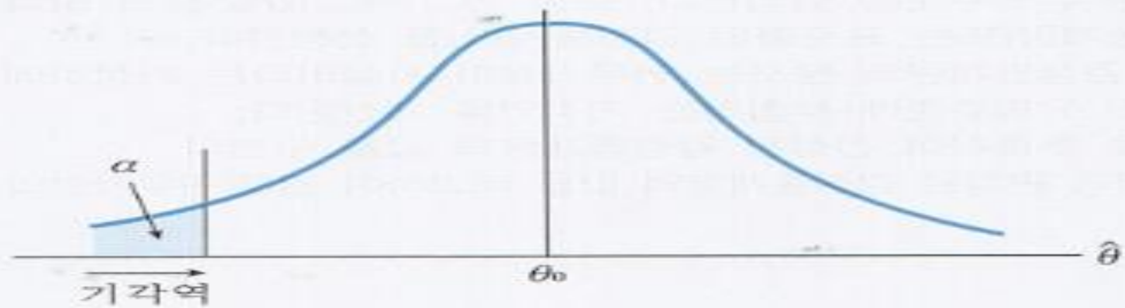
그림 9-1 표본통계량의 분포와 기각역의 설정

(a) $H_A : \theta \neq \theta_0$; $H_0 : \theta = \theta_0$



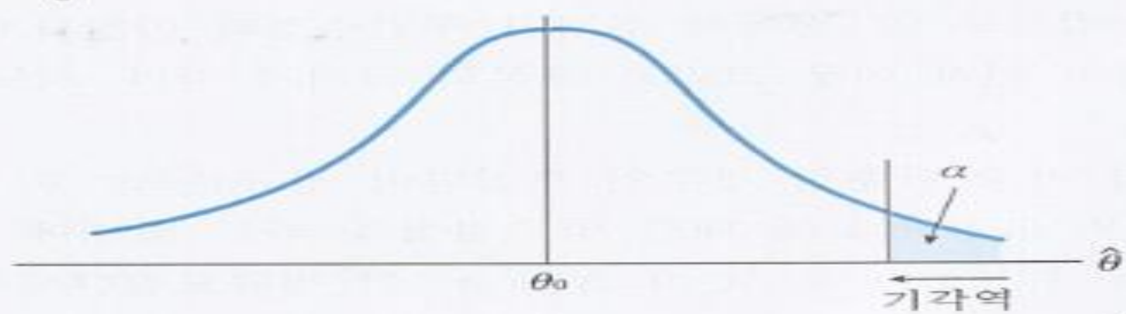
양측검정

(b) $H_A : \theta < \theta_0$; $H_0 : \theta \geq \theta_0$



좌측검정

(c) $H_A : \theta > \theta_0$; $H_0 : \theta \leq \theta_0$



우측검정

가설 검정 절차

- 가설검정 절차
  - 귀무가설($\boldsymbol{H_0}$)과 대립가설($\boldsymbol{H_1}$)의 설정
  - 유의수준 $\boldsymbol{\alpha}$의 선택
  - 검정통계량의 선정
  - 기각역의 계산
  - 검정통계량의 계산
  - 기각역과 검정통계량의 비교에 의한 결론

모평균에 대한 가설검증-1

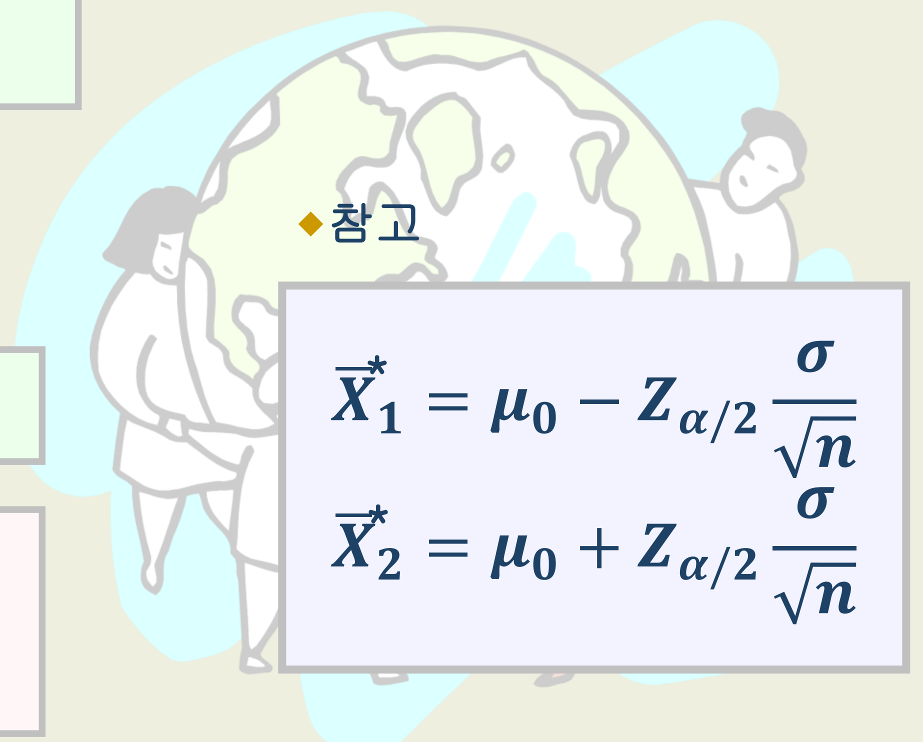
- 모평균의 가설검정(모분산이 기지이거나 표본이 큰 경우)

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

- 표본평균 $\bar{x}$ 이용

- 기각치 : $-z_{\alpha/2}, z_{\alpha/2}$

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



- 참고

$$\bar{X}_1^* = \mu_0 - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$
$$\bar{X}_2^* = \mu_0 + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

예제 9-3

천일기업에서 생산하는 백열전구의 수명은 평균(μ)이 1,000시간 그리고 분산(σ^2)이 100^2인 것으로 알려져 왔다. 천일의 경영진은 최근의 잦은 노사분규로 인한 종업원의 근로의욕 저하로 지금까지 알려져 온 백열전구의 평균수명이 여전히 유지되고 있는지 의문을 제기하였다. 이에 따라 64개의 전구를 무작위로 추출하여 수명검사를 한 결과 평균시간이 975시간으로 나타났다. 이러한 표본정보는 천일기업에서 생산되는 전구의 평균수명이 1,000시간과 다르다는 것을 의미하는가? 1%의 유의수준에서 가설검정을 하시오.

모평균에 대한 가설검증-2

- 모평균의 가설검정(모분산이 미지이고 표본이 작은 경우)

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

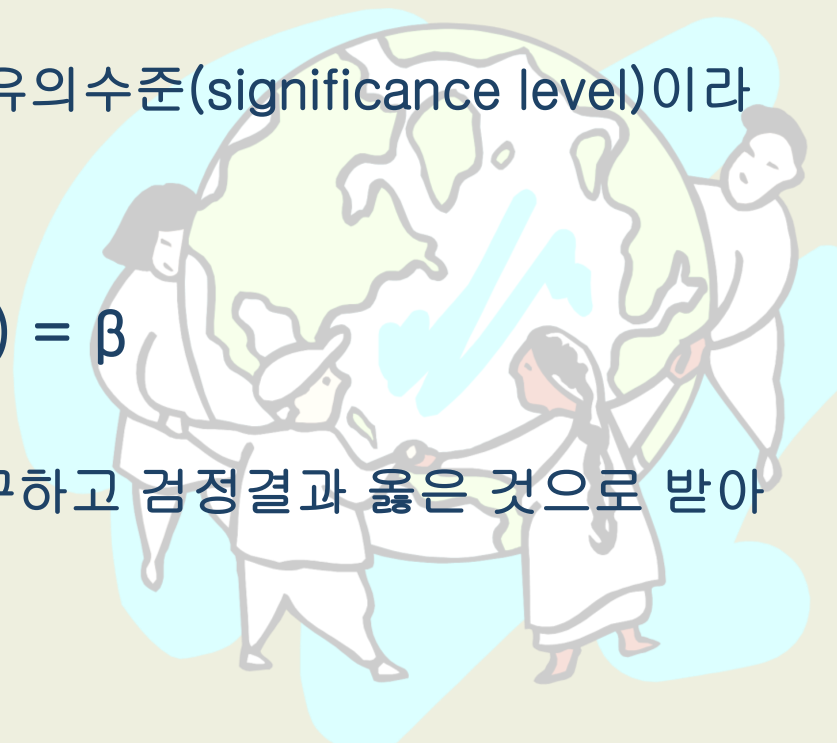
- 표본평균 $\bar{x}$ 이용
- 기각치 : $-t_{n-1,\alpha/2}, t_{n-1,\alpha/2}$
- 검정통계량 $t = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}}$

예제 9-4

안암산업의 부사장은 판매원의 성과에 불만을 표시하면서, 판매원의 계약체결실적이 평균적으로 많아야 주당 15건이라고 말하고 있다. 부사장의 주장을 검정하기 위하여 36명의 판매원을 임의로 추출하여 어느 한 주의 계약건수를 살펴보았더니, 평균이 16건이고 분산이 9로 나타났다. 이 표본자료는 부사장의 주장과 상반된다고 볼 수 있는가? 유의수준 0.05에서 검정하시오.

가설 검정의 오류

- 제1종 오류 (type Ⅰ error) = α
  - 귀무가설이 옳은데도 불구하고 검정결과 귀무가설을 기각하는 오류.
  - 제1종 오류를 범할 확률을 유의수준(significance level)이라 함
- 제2종 오류 (type Ⅱ error) = β
  - 귀무가설이 틀렸음에도 불구하고 검정결과 옳은 것으로 받아들이는 오류
- [표 9-1] 참조



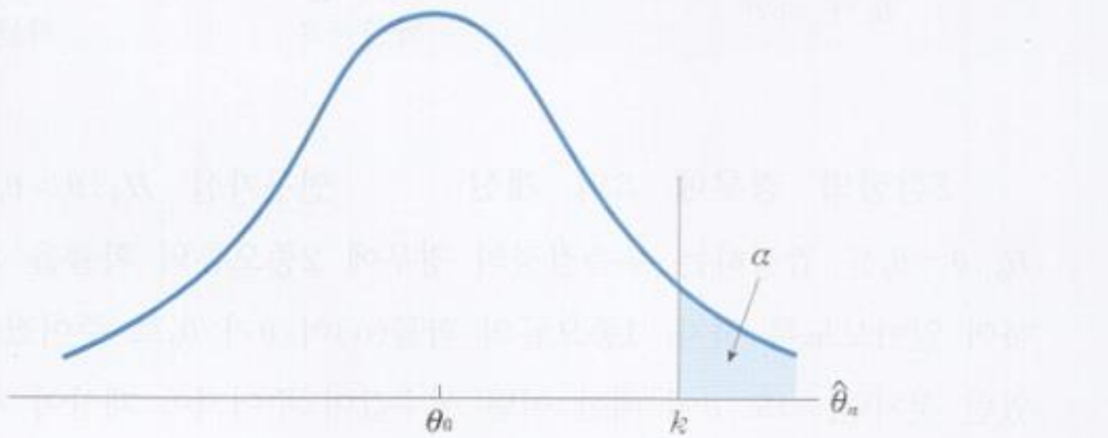
가설 검정의 오류

표 9-1 가설검정 네 가지 가능한 결과

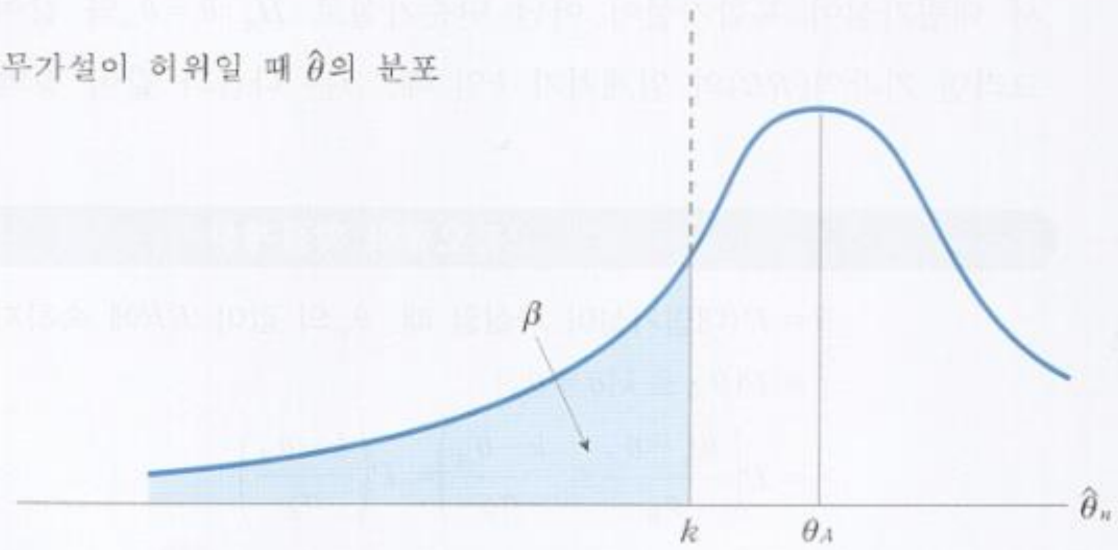
실제상황 \ 의사결정	H_0 채택	H_0 기각
H_0가 사실	올바른 결정 확률 $=1-\alpha$	1종오류 확률 $=\alpha$
H_0가 허위	2종오류 확률 $=\beta$	올바른 결정 확률 $=1-\beta=$ 검정력

그림 9-6 우측검정에서의 1종오류와 2종오류($H_A: \theta > \theta_0$; $H_0: \theta = \theta_0$)

(a) 귀무가설이 사실일 때 $\hat{\theta}$의 분포



(b) 귀무가설이 허위일 때 $\hat{\theta}$의 분포



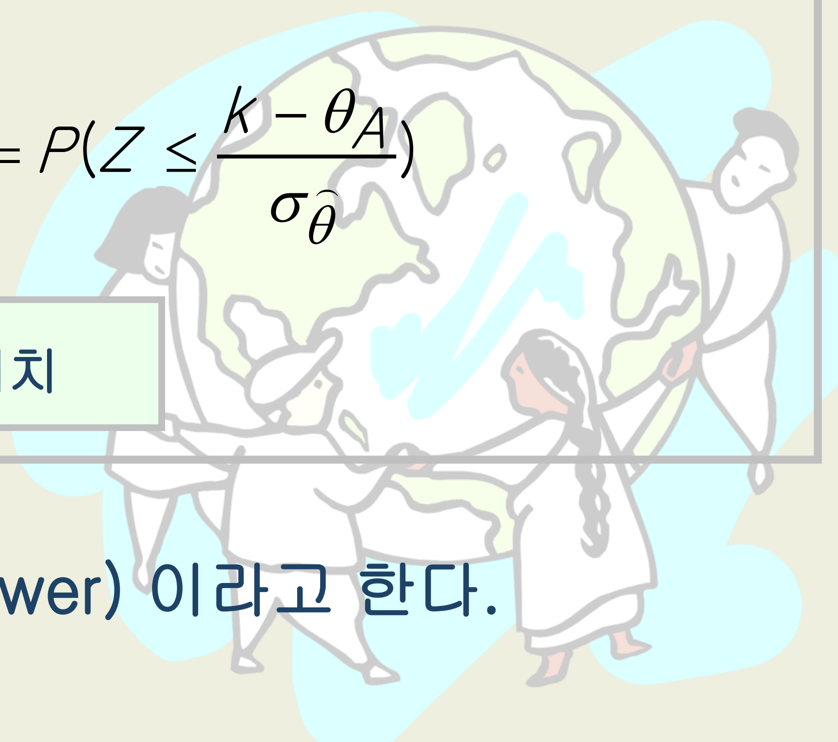
2종오류의 확률계산

β = P(대립가설이 사실일 때, $\hat{\boldsymbol{\theta}}$의 값이 RR에 속하지 않음)

$$= P(\hat{\theta} \le k \;|\theta = \theta_A)$$

$$= P(\frac{\hat{\theta} - \theta_A}{\sigma_{\hat{\theta}}} \le \frac{k - \theta_A}{\sigma_{\hat{\theta}}}) = P(Z \le \frac{k - \theta_A}{\sigma_{\hat{\theta}}})$$

여기서 $\boldsymbol{k}$ 는 $\boldsymbol{RR}$ 의 임계치



- 참고 (1-β)는 검정력(Power) 이라고 한다.

2종오류의 확률계산

예제 9-5

모분산이 16으로 알려진 모집단에 대하여 다음의 가설을 검정하려고 한다.

$$H_0 : \mu = 18; \; H_A : \mu > 18$$

$\mu_A = 19.5$일 때, 2종오류의 확률은 다음 각각의 경우에 얼마인가?

(a) $\alpha = 0.10$, $n = 36$

(b) $\alpha = 0.05$, $n = 36$

(c) $\alpha = 0.05$, $n = 64$

p값에 의한 가설검정

- p값에 의한 가설검정
  - p값은 귀무가설이 기각되는 최소수준의 α값으로 해석할 수 있으며, 따라서 p값이 α 값보다 작으면 주어진 귀무가설을 기각하게 된다.
- 검정력 (1−β)

예제 9-6

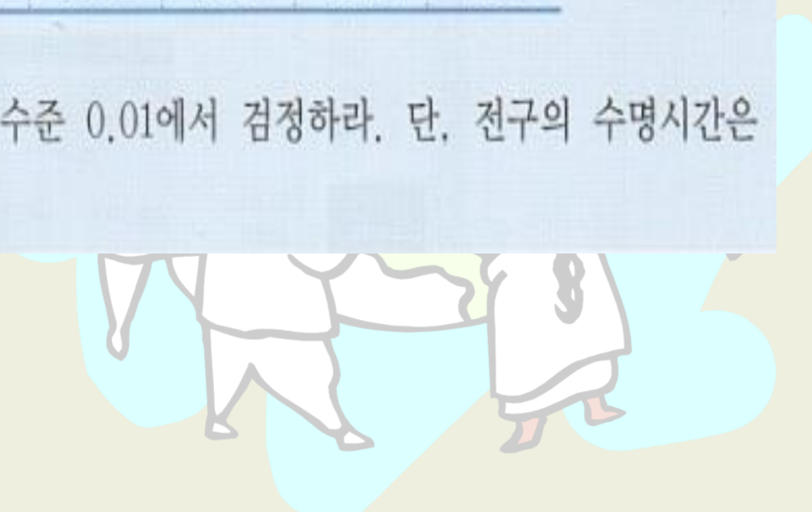
예제 4에 대해서 p값을 계산하여 유의수준 0.05에서 검정하시오.

예제 9-7

천일기업은 새로 개발한 백열전구의 평균수명이 1,000시간을 상회한다고 주장한다. 신제품의 실제 평균수명(시간)을 측정하기 위하여 생산된 전구 중 9개를 임의로 추출하여 검사하였더니 각 전구의 수명시간은 다음과 같았다.

1112	1053	990	984	1103	1005	1010	1152	1066

위의 표본자료를 이용하여 천일기업의 주장을 유의수준 0.01에서 검정하라. 단, 전구의 수명시간은 정규분포를 따른다고 가정한다.



모비율에 대한 가설검정

$$H_0 : p = p_0$$
$$H_A : p \neq p_0$$

- 표본평균 $\hat{P}$ 이용
- 기각치 : $-z_{\alpha/2}, z_{\alpha/2}$
- 검정통계량

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 q_0}{n}}}$$

모분산에 대한 가설검정

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

$$H_A : \sigma^2 \neq \sigma_0^2$$

- 표본평균 σ^2 이용

- 기각치 : $\chi^2_{n-1,1-\alpha/2}, \chi^2_{n-1,\alpha/2}$

- 검정통계량

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$$