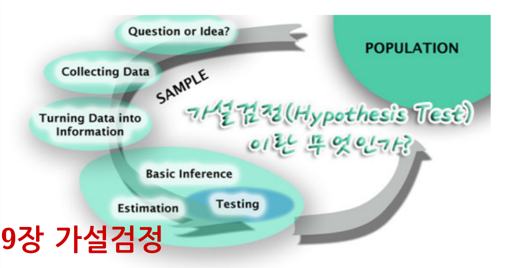


9장 가설검정

2021년 봄학기

전북대학교 통계학과

Introduction

- 멘델의 완두콩 실험
  - 가설: 우성:열성=3:1
  - 실험의 결과가 멘델의 법칙에 부합하는가?
  - 예측: 노랑콩 $1248 \times (3/4) = 936$개, 초록콩 $1248 \times (1/4) = 312$개
  - 검증: 최대예측오차$=\max\{|931-936|, |317-312|\} = 5$개
- 통계적 가설검정
  - 모집단을 관찰하여 그 분포에 관한 가설을 세운다.
  - 모집단의 일부를 표본으로 뽑아 관측값을 얻는다.
  - 가설이 참일 때 가설로부터 기대되는 예측값을 구한다.
  - 관측값과 예측값을 비교하여 가설의 진위를 가린다.

통계적 가설검정의 응용범위
자연·사회 현상에 대한 학문적 연구뿐만 아니라 실생활과 관련된 거의 모든 분야, 예를 들면

- 여론조사를 통한 정책 (후보자) 선호도 조사 ⇒ 정책결정
- 소비자 선호도 조사 ⇒ 제품설계·마케팅 전략 수립
- 공정 불량품 조사 ⇒ 생산성·품질 향상 도모
- 새로운 품종의 생산량·기후적응력 실험 ⇒ 신품종 보급 여부 결정
- 새로 개발된 약의 효능 실험 ⇒ 신약의 우수성 판단

Introduction

- 가설검정에 통계적 기법이 어떤 역할을 하는가?
    - 가설에 의한 예측결과와 실제 관측결과가 부합하는지를 판단하는 기준은?
    - 가설이 참일 때 거짓이라고 판단할 확률은?그 반대의 확률은?
    - 검정에 어떤 통계량을 쓸 것인가?
    - 기타

통계적 가설 검정 (statistical hypothesis testing)

입증되지 않은 주장에 대해 관측 결과가 일어날 확률을 기초로 주장의 타당성을 확인하는 일련의 과정

Object

We are interested in testing a statistical hypothesis

For this, suppose $X \sim f(x|\theta)$ and $\Theta_0 \subset \Theta$

Question Is it true $\theta \in \Theta_0$?

Answer We will give an answer for this question based on the given data

용어

가설 (hypothesis)
statement about the population parameter of the form $\theta \in \Theta_0$ where Θ_0 is a specified subset of Θ

대립가설 (alternative hypothesis)
조사된 표본이나 관측된 실험의 결과로부터 확실한 근거에 의하여 주장하고자 하는 가설 (H_1)

귀무가설 (null hypothesis)
대립가설과 상반되는 가설로 대립가설이 참이라는 근거가 없을 때 받아들일 수 있는 가설 (H_0)

용어

검정통계량 (test statistic)
귀무가설과 대립가설 중의 하나를 선택하는 기준으로 사용되는 통계량

기각역 (rejection region)
귀무가설 H_0를 기각시키는 검정통계량의 관측값의 영역

제 1종 오류 (type I error)
귀무가설이 참인데 귀무가설을 기각하는 오류

제 2종 오류 (type II error)
귀무가설이 거짓인데 귀무가설을 기각하지 않는 오류

용어

유의수준 (significance level)
제 1종 오류를 범할 확률의 최대 허용한계 (α)

유의확률 (*p*-value)
검정통계량의 관측값에 기초하여 주어진 귀무가설 H_0를 기각할 수 있는 최소의 유의수준

단측검정 (한쪽검정, one-sided test)
대립가설의 내용이 한쪽 방향의 서술로 구성되는 경우

양측검정 (양쪽검정, two-sided test)
대립가설의 내용이 양쪽 방향의 서술로 구성되는 경우

용어

Summary

- Null hypothesis $H_0 : \theta \in \Theta_0$
- Alternative hypothesis $H_1 : \theta \in \Theta_0^c$
- Decide whether H_0 or H_1 is true based on data $\boldsymbol{X}$
- Testing Hypothesis: Specifies for which $\boldsymbol{X}$ makes the decision H_0 is true and for which $\boldsymbol{X}$ makes the decision H_1 is true
- Critical region (or rejection region)

$$R = \{\boldsymbol{X} : H_0 \text{ is rejected (i.e. decide } H_1 \text{ is true}\}$$

- Acceptance region: $\{\boldsymbol{X} : H_0 \text{ is accepted}\}$

Note: Any partition of sample space into two sets R (rejection region) and R^c (acceptance region) defines a hypothesis test.

용어

Results of statistical hypothesis testing

Test results\Real situation	H_0 true	H_0 false (H_1 true)
Can not reject H_0 (Accept H_0)	Correct decision	Type II error
Reject H_0 (Accept H_1)	Type I error	Correct decision

Probability of errors

$P(\boldsymbol{X} \in R|H_0)$: 제 1종 오류를 범할 확률 (보통 α로 나타냄)

$P(\boldsymbol{X} \in R^c|H_1)$: 제 2종 오류를 범할 확률 (보통 β로 나타냄)

용어

Test function (검정함수) $\phi(x)$를 이용한 표현

$\boldsymbol{X} = \boldsymbol{x}$일 때 검정함수 $\phi(\boldsymbol{x})$는 0 또는 1의 값을 가지는 함수

검정법의 기각역이 R일 때 $\phi(\boldsymbol{x})$는 보통 다음과 같이 정의함

$$\phi(\boldsymbol{x}) = \begin{cases} 1, & \boldsymbol{x} \in R \\ 0, & \boldsymbol{x} \notin R \end{cases}$$

$$\begin{aligned} E[\phi(\boldsymbol{X})|H_0] &= 1 \times P(\boldsymbol{X} \in R|H_0) + 0 \times P(\boldsymbol{X} \in R^c|H_0) \\ &= P(\boldsymbol{X} \in R|H_0) = \alpha \\ E[\phi(\boldsymbol{X})|H_1] &= 1 \times P(\boldsymbol{X} \in R|H_1) + 0 \times P(\boldsymbol{X} \in R^c|H_1) \\ &= P(\boldsymbol{X} \in R|H_1) = 1 - P(\boldsymbol{X} \in R^c|H_1) = 1 - \beta \end{aligned}$$

Level α test (수준 α 검정)

제 1종 오류를 범할 확률의 최대허용한계가 α인 검정

용어

- 검정력 (power) $1-\beta$: (단순)대립가설이 참일 때 귀무가설을 기각할 확률
  $P(\boldsymbol{X} \in R | H_1)$
- 검정력 함수 (power function) : 귀무가설을 기각할 확률, θ의 함수
  $K(\theta) = E[\phi((X)]$

용어

Example
$X_1, X_2, \ldots, X_{20}$은 Bernoulli (p)로부터의 랜덤샘플이다

$$H_0 : p \geq 1/2 \quad vs \quad H_1 : p < 1/2$$

- Test statistic (검정통계량): $Y = \sum_{i=1}^{20} X_i$
- Rejection region (기각역) :

$$R = \{y : y \leq 6\} = \{(x_1, x_2, \ldots, x_n) : \sum_{i=1}^{20} x_i \leq 6\}$$

- significance level of the test (검정의 유의수준)

$$\alpha = P\left(Y \leq 6 : p = \frac{1}{2}\right) = \sum_{y=0}^{6} \binom{20}{y} \left(\frac{1}{2}\right)^y \left(\frac{1}{2}\right)^{20-y}$$

Example (continued)

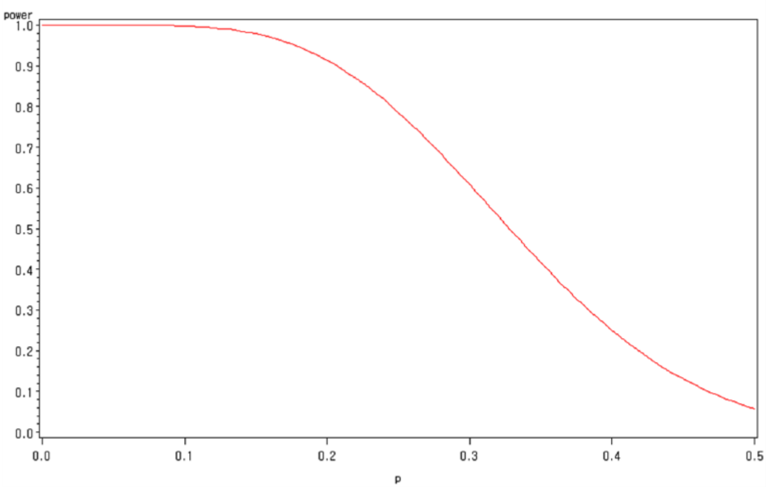
- probability of type II error (제 2종 오류를 범할 확률)
- power function of the test (검정력 함수) : $K(p)$

  Probability of rejecting H_0

$$K(p) = P(Y \leq 6 : p) = \sum_{y=0}^{6} \binom{20}{y} p^y (1-p)^{20-y}, \; 0 < p < 1$$

Note: 바람직한 검정력 함수는 H_0가 사실일 때는 작은 값을 가지고, H_0로부터 멀어질 때 큰 값을 가지는 함수

Example (continued)



최강력 검정 (Most powerful test)

- 귀무가설과 대립가설이 모두 단순가설인 경우 최적의 검정법을 구하는 방법

$$H_0 : \theta = \theta_0 \ \ vs \ \ H_1 : \theta = \theta_1$$

- 제1종 오류를 범할 확률이 일정한 값 (유의수준) 이하가 되는 검정법 중에서 검정력을 가장 크게 하는 검정
- 즉, 제1종 오류를 범할 확률을 기준으로 일정한 수준에 이른 검정법만으로 한정하고 이 중에서 제2종 오류를 범할 확률을 가장 작게 하는 검정법을 선택

정의 9.3: 최강력 검정
다음 조건을 만족하는 ϕ를 유의수준 α에서 최강력 검정이라고 함

(i) ϕ는 유의수준 α 검정, 즉 $E(\phi(\boldsymbol{X})|H_0) \leq \alpha$

(ii) 임의의 수준 α검정 ϕ'에 대하여 ϕ의 검정력이 ϕ'의 검정력보다 큼 즉, $E(\phi'(\boldsymbol{X})|H_0) \leq \alpha$인 검정 ϕ'에 대하여 $E(\phi(\boldsymbol{X})|H_1) \geq E(\phi'(\boldsymbol{X})|H_1)$

최강력 검정 (Most powerful test)

Theorem (Thm. 9.1 Neyman-Pearson lemma 네이만-피어슨 정리)

랜덤샘플 $\mathbf{X}$ 의 확률밀도함수가 $f(\mathbf{x}|\theta)$일 때, 다음의 가설검정에 대해,

$$H_0 : \theta = \theta_0 \quad vs \quad H_1 : \theta = \theta_1$$

다음과 같이 주어지는 검정법 ϕ는 유의수준 α에서 최강력검정법

$$\phi(\mathbf{x}) = \begin{cases} 1, & \dfrac{f(\mathbf{x}|\theta_1)}{f(\mathbf{x}|\theta_0)} \geq k \text{ 일 때} \\ 0, & \dfrac{f(\mathbf{x}|\theta_1)}{f(\mathbf{x}|\theta_0)} < k \text{ 일 때} \end{cases}$$

단, α는 다음을 만족하는 값

$$E(\phi(\mathbf{X})|H_0) = P\left(\frac{f(\mathbf{x}|\theta_1)}{f(\mathbf{x}|\theta_0)} \geq k \middle| H_0\right) = \alpha$$

최강력 검정 (Most powerful test)

예 9.1
확률변수 X의 확률분포가 다음과 같을 때 X를 이용한 다음의 가설을 검정하고자 한다.

x	-3	-2	-1	0	1	2	3
$f(x\|\theta_0)$	0.4	0.2	0.1	0.1	0.1	0.05	0.05
$f(x\|\theta_1)$	0.05	0.05	0.1	0.1	0.2	0.2	0.3

- 유의수준 0.05에서 최강력 검정법
- 유의수준 0.1에서 최강력 검정법
- 유의수준 0.075에서 최강력 검정법

최강력 검정 (Most powerful test)

랜덤화검정(randomized test) : 이산형 확률분포에 대한 검정에서 특정한 관측값에서 귀무가설을 기각할 확률이 0이나 1이 아닌 값을 가지는 경우

예 9.4

$X_1, X_2, \ldots, X_n$이 N(θ,1)로부터의 랜덤샘플일 때 다음 가설에 대한 유의수준 α에서 최강력검정법의 기각역

$$H_0 : \theta = \theta_0 \ vs \ H_1 : \theta = \theta_1 \quad (\text{단, } \theta_1 > \theta_0)$$

예 9.5

$X_1, X_2, \ldots, X_n$이 Exp($1/\lambda$)로부터의 랜덤샘플일 때 다음 가설에 대한 유의수준 α에서 최강력검정법의 기각역

$$H_0 : \lambda = \lambda_0 \ vs \ H_1 : \lambda = \lambda_1 \quad (\text{단, } \lambda_1 > \lambda_0)$$

가능도비 검정 (Likelihood Ratio Test)

Recall 가능도 함수 (Likelihood function)

$$L(\theta|\boldsymbol{x}) = f(\boldsymbol{x}|\theta) = f(x_1|\theta)f(x_2|\theta)\cdots f(x_n|\theta)$$

가설: $H_0 : \theta \in \Theta_0 \;\; vs \;\; H_1 : \theta \in \Theta_1$

가능도비 검정 통계량 (LRT statistic)

$$\Lambda(\boldsymbol{x}) = \frac{\max_{\theta \in \Theta_0} L(\theta|\boldsymbol{x})}{\max_{\theta \in \Theta} L(\theta|\boldsymbol{x})}, \quad \Theta = \Theta_0 \cup \Theta_1$$

기각역 (Rejection region)

$$\{\boldsymbol{x} : \Lambda(\boldsymbol{x}) \le c\}, \;\; c\text{:상수}, 0 \le c \le 1, 0 \le \Lambda(\boldsymbol{x}) \le 1$$

가능도비 검정 (Likelihood Ratio Test)

LRT statistic -MLE와 밀접하게 관련됨

$\hat{\theta}$: 모수공간(parameter space) Θ에서의 MLE

$\hat{\theta}_0$: 모수공간(parameter space) Θ_0에서의 MLE

$$\Lambda(\mathbf{x}) = \frac{L(\hat{\theta}_0|\mathbf{x})}{L(\hat{\theta}|\mathbf{x})} \ or \ \lambda(\mathbf{x}) = l(\hat{\theta}_0|\mathbf{x}) - l(\hat{\theta}|\mathbf{x})$$

Likelihood Ratio Tests 과정

1. MLE $\hat{\theta}_0, \hat{\theta}$ 구하기
2. $\hat{\theta}_0, \hat{\theta}$를 가능도 함수에 대입하여 $\Lambda(\mathbf{x})$를 구하기
3. $\{\Lambda(\mathbf{x}) \leq c\}$의 단순화
4. 유의수준 α를 만족하는 c 찾기

가능도비 검정 (Likelihood Ratio Test)

예 9.6 Normal LRT with known variance

$X_1, X_2, \ldots, X_n$이 N(θ,1)로부터의 랜덤샘플일 때 다음 가설에 대한 가능도비 검정의 유의수준 α 기각역

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

예 9.7 Normal LRT with unknown variance

$X_1, X_2, \ldots, X_n$이 $N(\mu, \sigma^2), \sigma^2$: unknown 로부터의 랜덤샘플일 때 다음 가설에 대한 가능도비 검정의 유의수준 α 기각역

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

예 Normal LRT for variance

$X_1, X_2, \ldots, X_n$이 $N(\mu, \sigma^2), \mu, \sigma^2$: unknown 로부터의 랜덤샘플일 때 다음 가설에 대한 가능도비 검정의 유의수준 α 기각역

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

가능도비 검정 (Likelihood Ratio Test)

정리 9.2
적절한 조건 하에서 $-2\log\Lambda(\boldsymbol{X})$는 귀무가설이 참일 때 점근적으로 카이제곱분포를 따름

$$-2\log\Lambda(\boldsymbol{X}) \sim \chi_d^2$$

단 d는 모수공간 $\Theta_0 \cup \Theta_1$의 차원과 모수공간 Θ_0의 차원의 차이

가능도비 검정의 유의수준 α 점근적인 기각역

$$\{\boldsymbol{x} : -2\log\Lambda(\boldsymbol{x}) \geq \chi_d^2(\alpha)\}$$

가능도비 검정 (Likelihood Ratio Test)

분할표자료: 적합도 검정(goodness-of-fit test)
$\boldsymbol{X} = (X_1, X_2, \ldots, X_m)$이 시행횟수 n이고 각 범주에 속할 확률이 $(p_1, p_2, \ldots, p_m)$ 인 다항분포를 따를 때 $\boldsymbol{X}$의 확률밀도 함수

$$f(x_1, x_2, \ldots, x_m | (p_1, p_2, \ldots, p_m) = \frac{n!}{x_1! x_2! \cdots x_m!} p_1^{x_1} p_2^{x_2} \cdots p_m^{x_m}$$

다음 가설에 대한 가능도비 검정

$$H_0 : (p_1, p_2, \ldots, p_m) \in \Theta_0 \;\; vs H_1 : (p_1, p_2, \ldots, p_m) \in \Theta_1$$

독립표본 t검정
서로 독립인 두 랜덤샘플

$$\boldsymbol{X} = (X_1, X_2, \ldots, X_m) \sim N(\mu_X, \sigma^2)$$

$$\boldsymbol{Y} = (Y_1, Y_2, \ldots, Y_n) \sim N(\mu_Y, \sigma^2)$$

을 이용하여 두 모집단의 평균에 대한 다음 가설의 가능도비 검정

$$H_0 : \mu_X = \mu_Y \;\; vs \;\; H_1 : \mu_X \neq \mu_Y$$

가능도비 검정 (Likelihood Ratio Test)

LRT for Exponential distribution mean
평균이 $\theta(0 < \theta < \infty)$인 지수분포로부터의 랜덤샘플 $X_1, X_2, \ldots, X_n$을 이용하여

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

를 검정할 때 유의수준 α인 가능도비 검정

LRT for uniform distribution
Uniform(0,θ)로부터의 랜덤샘플 $X_1, X_2, \ldots, X_n$을 이용하여

$$H_0 : \theta = 0.5 \;\; vs \;\; H_1 : \theta < 0.5$$

를 검정할 때 유의수준 α인 가능도비 검정