11 강

가설검정 II

통계·데이터과학과 이긍희 교수

한국방송통신대학교

목차

01

모비율의 가설검정

귀무가설과 대립가설

- 귀무가설 : $H_0 : p = p_0$
- 대립가설
  - $H_1 : p < p_0$
  - $H_1 : p > p_0$
  - $H_1 : p \neq p_0$

검정통계량

검정통계량

$$Z = \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}} \sim N(0,1)$$

검정방법

검정방법

가설	기각역을 이용한 검정	유의확률을 이용한 검정
$H_0: p = p_0$ $H_1: p > p_0$	$Z > z_\alpha$이면 H_0를 기각	p값 $= P(Z > z_{obs} \mid H_0)$가 α보다 작으면 H_0를 기각
$H_0: p = p_0$ $H_1: p < p_0$	$Z < -z_\alpha$이면 H_0를 기각	p값 $= P(Z < z_{obs} \mid H_0)$가 α보다 작으면 H_0를 기각
$H_0: p = p_0$ $H_1: p \neq p_0$	$\lvert Z \rvert > z_{\alpha/2}$이면 H_0를 기각	p값 $= P(\lvert Z \rvert > z_{obs} \mid H_0)$가 α보다 작으면 H_0를 기각

모비율 가설검정의 예

- 기존 치료제 치료율 60%, 새 치료제 50명 중 35명 치료. 새 치료제 치료율 향상되었는지 5% 유의수준에서 검정.

02

모분산의 가설검정

귀무가설과 대립가설

- 귀무가설 : $H_0: \sigma^2 = \sigma_0^2$
- 대립가설
  - $H_1: \sigma^2 < \sigma_0^2$
  - $H_1: \sigma^2 > \sigma_0^2$
  - $H_1: \sigma^2 \neq \sigma_0^2$

검정통계량

검정통계량

$$\frac{(n-1)S^2}{\sigma_0^2} \sim \chi^2_{n-1}$$

검정방법

검정방법

가설	기각역을 이용한 검정	유의확률을 이용한 검정
$H_0 : \sigma^2 = \sigma_0^2$ $H_1 : \sigma^2 > \sigma_0^2$	$(n-1)S^2/\sigma_0^2 > \chi^2_{n-1,\ \alpha}$이면 H_0를 기각	p값 $= P(\chi^2 > \chi^2_{obs} \mid H_0)$가 α보다 작으면 H_0를 기각
$H_0 : \sigma^2 = \sigma_0^2$ $H_1 : \sigma^2 < \sigma_0^2$	$(n-1)S^2/\sigma_0^2 < \chi^2_{n-1,\ 1-\alpha}$이면 H_0를 기각	p값 $= P(\chi^2 < \chi^2_{obs} \mid H_0)$가 α보다 작으면 H_0를 기각
$H_0 : \sigma^2 = \sigma_0^2$ $H_1 : \sigma^2 \neq \sigma_0^2$	$(n-1)S^2/\sigma_0^2 > \chi^2_{n-1,\ \alpha/2}$ 또는 $(n-1)S^2/\sigma_0^2 < \chi^2_{n-1,\ 1-\alpha/2}$이면 H_0를 기각	p값 $= P(\chi^2 < \chi^2_{obs}$ 또는 $\chi^2 > \chi^2_{obs} \mid H_0)$ 가 α보다 작으면 H_0를 기각

모분산의 가설검정 예

- 쌀포대 무게의 표준편차 10g 이하 안정적. 10개의 쌀포대 무게 표본 표준편차 5g. 쌀무게 정규분포. 쌀의 표준편차가 10g 이하인지 유의수준 5%에서 검정.

모분산의 가설검정 예

- 12명이 읽은 책 수 표준편차가 전년도 4권보다 큰지 유의수준 5%에서 검정. 모집단 정규분포.

03

가설검정과 구간추정

채택역과 신뢰구간

- $H_0: \theta = \theta_0$ 의 검정
- **채택역 : θ 에 대한 $\mathbf{100}(1-\alpha)\%$ 신뢰구간**

$$P(X \in A(\theta_0) | \theta = \theta_0) = 1 - \alpha$$

채택역과 신뢰구간

- $H_0: \mu = \mu_0,\ H_1: \mu \neq \mu_0$
- **채택역** : $|T| < t_{n-1,\alpha/2}$

  $\leftrightarrow \mu$ 의 $100(1-\alpha)\%$ 신뢰구간

$$\left[\bar{X} - t_{n-1,\alpha/2}\frac{S}{\sqrt{n}},\qquad \bar{X} + t_{n-1,\alpha/2}\frac{S}{\sqrt{n}}\right]$$

신뢰구간과 기각

- $H_0: \mu = \mu_0,\ H_1: \mu \neq \mu_0$
- μ_0가 $100(1-\alpha)\%$ 신뢰구간에 포함되지 않으면 H_0를 기각

04

R을 이용한 실습

모평균의 가설검정


```
> # 예제 6-6
> n = 10; s = 0.2; bar_x = 12.2
> alpha = 0.05
> ttest = (bar_x - 12)/(0.2/sqrt(10)) # 검정통계량
> ttest_cr = qt(1-alpha/2,n-1) # 기각역
> ttest_pv = (1-pt(ttest,n-1))*2 # 유의확률
> cat("검정통계량값: ", ttest, "기각역: ", ttest_cr, "유의확률: ", ttest_pv)
검정통계량값:  3.162278 기각역:  2.262157 유의확률:  0.01150799>
> # 예제 6-7
> book = c(5, 23, 20, 1, 10, 15, 15, 10, 9, 13, 18, 11, 18, 20, 19, 19)
> t.test(book,mu=11, alternative = "greater")

        One Sample t-test

data:  book
t = 2.0573, df = 15, p-value = 0.02874
alternative hypothesis: true mean is greater than 11
95 percent confidence interval:
 11.46216      Inf
sample estimates:
mean of x
   14.125
```

모비율의 가설검정


```
> # 예제 6-8
> p0 = 0.6; n = 50 ; hat_p = 0.7
> alpha = 0.05
> ptest =(hat_p - p0)/sqrt(p0*(1-p0)/n)
> ptest_cr = qnorm(1-alpha)
> ptest_pv = 1-pnorm(ptest)
> cat("검정통계량값: ", ptest, "기각역: ", ptest_cr, "유의확률: ", ptest_pv)
검정통계량값:  1.443376 기각역:  1.644854 유의확률:  0.07445734
```

모분산의 가설검정

```
> # 예제 6-10
> n=12
> alpha=0.05
> book = c(5, 23, 20, 1, 10, 15, 15, 10, 9, 13, 18, 11)
> vtest = var(book)*(12-1)/4^2
> vtest_cr = qchisq(1-alpha, n-1)
> vtest_pv = 1-pchisq(vtest, n-1)
> cat("검정통계량값: ", vtest, "기각역: ", vtest_cr, "유의확률: ", vtest_pv)
검정통계량값:  26.5625 기각역:  19.67514 유의확률:  0.005347859
```

정리하기

- 표본수가 클 때 1개 모집단의 모비율에 대한 검정은 z검정통계량을 이용한다.
- 1개 모집단의 모분산에 대한 검정은 χ^2 검정통계량을 이용한다.
- 가설검정의 채택역과 신뢰구간은 같다.

다음시간 안내

12강

통계적 비교 I

한국방송통신대학교