

8.4 입표본 비율

- 구간추정

모비율에 대한 $100 \times (1 - \alpha) \% 95\%$ 신뢰구간 $(\hat{p} - Z_{\alpha/2} \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \hat{p} + Z_{\alpha/2} \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}})$

- 가설검정 절차(Z-test)

(㉠) 가설 $H_0 : p = p_0$ v.s. $H_1 : p \neq p_0$

(㉡) 유의수준 $\alpha = 0.05$

(㉢) 검정통계량 $Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \simeq N(0,1)$

(㉣) H_0 를 기각 if $p\text{-value} = 2 \times P(Z \geq |z|) \leq \alpha(0.05)$

(1) 예제자료($n = 150, x = 110$)

`prop.test(x = 110, n=150, p=0.85)` # $H_1 : p \neq 0.85$

해석 :

1) 추정 : A후보의 지지율의 표본 지지율은 73.3%이며,
A후보의 지지율에 대한 95% 신뢰구간은 (65.3%, 80.1%)이다.

2) 가설검정

① 가설 H_0 : A후보의 지지율은 85%이다.

H_1 : A후보의 지지율은 85%가 다르다.

② 유의수준 $\alpha = 0.05$

③ 검정통계량 $\chi^2 = 15.1$

(R에서는 표준정규분포의 제공인 χ^2 -분포를 이용 => 같은 결과)

④ $P\text{값} = 0.0001 < \alpha \Rightarrow H_0$ 를 기각한다.

⑤ 결론 : 유의수준 5%에서 A후보의 지지율은 85%가 다르다고 할 수있다.

즉 A후보의 지지율은 85%보다 낮다고 할 수있다.

- 사용자 함수 정의(R 프로그래밍)

```
Z_prop_1 = function(x, n, p0) {
  phat = x / n ; se = sqrt ( phat*(1-phat) / n )
  lb = phat - qnorm(0.975) * se
  ub = phat + qnorm(0.975) * se
  se = sqrt ( p0*(1-p0) / n )
  Z = ( phat - p0) / se
  pvalue = 2 * ( 1 - pnorm( abs(Z) ))
  cat( " ===== 일표본 비율 검정 =====", "\n", "\n")
  cat( " 표본비율 = ", phat, " 95% 신뢰구간 ( " , lb, ub, " )" , "\n")
  cat( " Z = " , Z, " , P - value = " , pvalue , "\n")
}
n = 150 ; x = 110 ; p0 = 0.85
```

- R제공

```
prop.test(x = 110, n=150, p=0.85)          #  $H_1 : p \neq 0.85$ 
```

- R프로그래밍 사용

```
Z_prop_1(x = 110, n=150, p=0.85)
```

(2) 우리자료

- R제공

```
table (gender)          # 성별에 대한 도수분포표
prop.test(x = 33, length(gender), p=0.5)  #  $H_1 : p \neq 0.5$ 
```

- R프로그래밍 사용

```
Z_prop_1(x = 33, length(gender), p=0.5)
```

[과제22] (여러분은 따라서 해보시고 연습하시길)

동영상을 잘 보시고, 따라서 한번 해보시고,
 다양한 연습(phat을 변화시켜서)을 과제로 하시고, 해석을 해주시길..
 첨부파일 : 학번이름22.hwp (예 : 20192260홍길동22.hwp)