

8 검정

Topics:

- 8.1 서론
- 8.2 검정의 원리
- 8.3 모평균의 검정
- 8.4 모비율의 검정
- 8.5 두 모평균의 차에 관한 검정
- 8.6 두 모비율의 차에 관한 검정
- 8.7 비모수적 검정

8.1 서론

Topics:

- 가설검정
- 가설검정의 단계

가설검정(hypothesis testing):

추정: 모집단의 어떤 미지의 값을 추론하는 과정

가설검정: 모집단에 대한 어떤 예상의 옳고 그름을 판정하거나, 주장의 채택/기각을 결정하는 과정

- 예: A 백열전구의 새 공법이 소개되고 평균수명이 1,200시간 이상이라고 광고

- 신뢰구간과 가설의 주장과의 관계:

- 가설검정의 단계:

8.2 검정의 원리

Topics:

- 가설검정의 요소들
- 가설검정의 절차

귀무가설(null hypothesis, H_0)과 대립가설(alternative hypothesis, H_1):

통계적 가설: 모집단에 대한 예상 또는 주장

귀무가설: 과거 이론이나 경험적으로 참이라고 믿어지지만 검증되어야 하는 가설

대립가설: 귀무가설에 대립하여 설정되는 또 다른 가설

- 예: 어느 유제품 제조회사에서 특정 아이스크림에 생우유가 50% 들어 있다고 주장:

검정통계량(test statistic), 기각역(rejection region), 검정오류(test error):

검정통계량: 귀무가설과 대립가설 중에서 하나를 선택하는 데 사용하는 표본자료의 통계량
기각역: 귀무가설을 기각시키는 검정통계량의 관측값의 영역
검정오류: 제1종 오류와 제2종 오류
제1종 오류(type 1 error): 귀무가설이 옳는데 대립가설을 채택하는 오류
제2종 오류(type 2 error): 대립가설이 옳는데 귀무가설을 채택하는 오류

- 검정통계량의 예:

- 검정의 **유의수준(significance level)**:

제1종 오류가 발생할 확률(의 허용한계)을 α 로 표기하고, 검정의 유의수준이라 한다.
또한 제2종 오류의 확률을 β 로 표기한다.

- 귀무가설의 진위와 의사결정:

- 유의수준 α 의 결정:

- 기각역(rejection region): H_0 가 사실이 아니라고 기각하게 되는 검정통계량의 영역

가설검정의 절차:

1. 귀무가설 H_0 와 대립가설 H_1 을 설정
2. 검정의 유의수준 α 를 설정
3. 표본에 근거한 검정통계량을 계산
4. 기각역을 설정
5. 유의성을 판정하고, H_0 의 기각 또는 채택을 결정

- 예: 백열전구 생산공정에서 평균수명이 1,200시간보다 긴 새로운 공법이 개발되었다고 주장.

8.3 모평균의 검정

Topics:

- 대표본에서 모평균의 검정
- 소표본에서 모평균의 검정

대표본에서 모평균의 검정:

$$\text{귀무가설 } H_0 : \mu = \mu_0$$
$$\text{검정통계량 } Z = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$

- 기각역
- (1) $H_1 : \mu > \mu_0$ 일 때, $Z \geq z_\alpha$
 - (2) $H_1 : \mu < \mu_0$ 일 때, $Z \leq -z_\alpha$
 - (3) $H_1 : \mu \neq \mu_0$ 일 때, $|Z| \geq z_{\alpha/2}$

- **Review:** 모평균의 관한 추론은 표본평균을 이용한다.

- 예: 내용물의 무게가 350g으로 표시된 어느 과일 통조림의 제조과정에서는 평균무게가 360g이 되도록 공정을 관리하고 있다. 이를 확인하기 위하여 랜덤하게 30개의 표본을 추하여 조사한 결과

$$\bar{x} = 356(\text{g}), s = 10(\text{g})$$

을 얻었다. 공정에 이상이 있다고 확신할 수 있는가? 유의수준 5%에서 검정하여라.

Ans.

1. 귀무가설:
대립가설
2. 유의수준:
3. 검정통계량:
4. 기각역:
5. 검정통계량의 관측값:

- 예: 어느 특수 컴퓨터칩 생산회사에서는 칩의 평균수명에 대하여 15,000시간 이상을 보증한다고 광고하고 있다. 이를 확인하기 위하여 100개의 칩을 단축실험으로 검사한 결과

$$\bar{x} = 15,200(\text{시간}), s = 1,220(\text{시간})$$

을 얻었다. 이 회사의 광고를 믿을 수 있는가? 유의수준 1%에서 검정하여라.

Ans.

1. 귀무가설:
대립가설
2. 유의수준:
3. 검정통계량:
4. 기각역:
5. 검정통계량의 관측값:

- 예: 어느 치즈 스낵 통조림에는 치즈함량이 12.5%라고 표시되어 있다. 이를 확인하기 위하여 64통을 조사한 결과 치즈함량이

$$\bar{x} = 11.5(\%), s = 2.0(\%)$$

을 얻었다. 이 통조림의 내용물 함량표시는 정당하다고 할 수 있는가? 유의수준 5%에서 검정하여라.

Ans.

1. 귀무가설:
대립가설
2. 유의수준:
3. 검정통계량:
4. 기각역:
5. 검정통계량의 관측값:

- 예: 어떤 베어링은 지름이 1.20cm로 표시되어 있다. 이를 확인하기 위하여 36개의 베어링을 조사한 결과

$$\bar{x} = 1.22(\text{cm}), s = 0.16(\text{cm})$$

을 얻었다. 이 베어링의 크기 표시는 정당하다고 할 수 있는가? 유의수준 5%에서 검정하여라.

Ans.

1. 귀무가설:
대립가설
2. 유의수준:
3. 검정통계량:
4. 기각역:
5. 검정통계량의 관측값:

소표본에서 모평균의 검정:

$$\begin{aligned} \text{귀무가설 } H_0 : \mu &= \mu_0 \\ \text{검정통계량 } t &= \frac{\bar{X} - \mu}{S/\sqrt{n}} \end{aligned}$$

- 기각역 (1) $H_1 : \mu > \mu_0$ 일 때, $t \geq t_\alpha(n-1)$
 (2) $H_1 : \mu < \mu_0$ 일 때, $t \leq -t_\alpha(n-1)$
 (3) $H_1 : \mu \neq \mu_0$ 일 때, $|t| \geq t_{\alpha/2}(n-1)$

- **Review:** 모집단의 분포가 정규분포일 때 t -통계량을 이용하여 모평균에 관한 검정을 한다.

- 예: 어느 지역 청소년의 평균 IQ가 110이라고 알려져 있다. 이 지역에서 약물 복용으로 체보된 청소년의 IQ가 평균 IQ보다 떨어지는지를 알아보기 위하여 15명의 약물복용 청소년에 대한 IQ 검사를 한 결과

$$\bar{x} = 108, s = 12.0$$

을 얻었다. 약물복용자의 평균 IQ는 이 지역의 IQ보다 떨어지는지를 유의수준 5%에서 검정하여라. (단, 약물복용자의 IQ는 정규분포를 따른다고 가정)

Ans.

1. 귀무가설:
대립가설
2. 유의수준:
3. 검정통계량:
4. 기각역:
5. 검정통계량의 관측값:

- 예: 우리 나라 농가에서 1987년 가구당 평균소득이 654(만원)이라는 발표가 있었다. 어느 마을에서 그 마을의 가구당 평균소득이 전국의 평균소득과 차이가 있는지를 알아보기 위하여 10가구를 랜덤하게 추출하여 조사한 결과

$$\bar{x} = 620(\text{만원}), s = 120(\text{만원})$$

을 얻었다. 이 마을의 평균소득이 전국의 평균소득과 다른지를 유의수준 10%에서 검정하여라. (단, 이 마을의 농가소득은 정규분포를 따른다고 가정)

Ans.

1. 귀무가설:
대립가설
2. 유의수준:
3. 검정통계량:
4. 기각역:
5. 검정통계량의 관측값:

8.4 모비율의 검정

대표본에서 모비율의 검정:

귀무가설 $H_0 : p = p_0$

$$\text{검정통계량 } Z = \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}}$$

기각역 (1) $H_1 : p > p_0$ 일 때, $Z \geq z_\alpha$

(2) $H_1 : p < p_0$ 일 때, $Z \leq -z_\alpha$

(3) $H_1 : p \neq p_0$ 일 때, $|Z| \geq z_{\alpha/2}$

- **Review:** 모비율의 관한 추론은 이항확률변수나 표본비율을 이용한다.

- 예: 어떤 제약회사에서는 무좀약을 새로 개발하여, 일주일간 치료하면 치유율이 70% 이상이라고 광고하였다. 이를 확인하기 위하여 30명의 무좀환자에게 이 약을 바르게 한 결과 이 중에서 27명이 완치되었다. 이 제약회사의 광고는 신빙성이 있다고 할 수 있는가? 유의수준 5%에서 검정하여라.

Ans.

1. 귀무가설:
대립가설
2. 유의수준:
3. 검정통계량:
4. 기각역:
5. 검정통계량의 관측값:

- 예: 지난 달에 어느 TV 방송국의 저녁 뉴스시간의 시청률이 38%로서 상당히 높은 것으로 나타났다. 이 시청률이 이 달에도 변함이 없는지를 알아보기 위해 200세대를 전화번호부에서 랜덤하게 뽑아 조사한 결과, 그 중 64세대가 이 방송국의 뉴스를 시청하고 있는 것으로 나타났다. 시청률이 변동되었다는 근거가 있는가. 유의수준 5%에서 검정하여라.

Ans.

1. 귀무가설:
대립가설
2. 유의수준:
3. 검정통계량:
4. 기각역:
5. 검정통계량의 관측값: