제14장 분산에 대한 추정과 검정

분산에 대한 추정과 검정

- 모집단이 정규분포일 때
  - X^2 분포를 이용
- 모집단이 정규분포가 아닐 때
  - 표본자료의 변환(transform)을 통해 정규분포와 유사한 형태로 변환한다.
  - 비모수통계방법 사용: jackknife 방법

하나의 모집단의 분산에 대한 추정과 검정(1)

- s^2의 표본분포
  - 모집단이 정규분포이고 분산이 δ^2인 모집단에서 n개의 표본을 무작위로 추출하면

$$\frac{(n-1)s^2}{\delta^2} \sim x^2(n-1)$$

$$\text{단, } s^2 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1}$$

  - $x^2(n-1)$은 자유도가 n-1인 x^2분포
    - 연속확률분포로 모수는 자유도 v(누)
    - $x^2(v)$로 표기하고 x^2값은 항상 양수
    - 오른쪽 꼬리를 가진 비대칭분포
    - 자유도가 커지면 정규분포에 접근
  - P.315 그림 14.1 참고

하나의 모집단의 분산에 대한 추정과 검정(2)

- 모집단분산 δ^2에 대한 신뢰구간추정

$\frac{(n-1)s^2}{\delta^2} = x^2(n-1)$분포를 따르므로 x^2분포의 백분위수의 정의에 따라 다음과 같이 쓸 수 있다.

$$P\left[x^2(\frac{\alpha}{2};n-1) \le \frac{(n-1)s^2}{\delta^2} \le x^2(1-\frac{\alpha}{2};n-1)\right] = 1-\alpha$$

p.317 그림14.3 참고

$$P\left[\frac{(n-1)s^2}{x^2(1-\frac{\alpha}{2};n-1)} \le \delta^2 \le \frac{(n-1)s^2}{x^2(\frac{\alpha}{2};n-1)}\right] = 1-\alpha$$

  - p.317 예제 14.1
- 모집단분산 δ^2에 대한 가설검정
  - P. 319 예제 14.2

두 모집단의 분산의 비교(1)

- 두 모집단의 분산이 같은지의 여부
  - 두 모집단의 분산의 차이에 대해서가 아니라 비율에 의하여 추정과 검정(F분포)
  - 두 모집단이 정규분포라는 가정
  - 표본은 두 모집단에서 독립적을 추출되었다고 가정
- F분포
  - F분포는 연속확률변수로서 두 가지 양의 비율로 나타나며, 분자와 분모에 해당하는 두 개의 자유도(df)를 가진다.
  - F분포는 항상 양의 값을 가지며 오른쪽 꼬리를 가진 비대칭 분포
  - P.321 그림 14.5 참조

$$\text{두 모집단분산의비율}: \frac{\delta_2^2}{\delta_1^2}$$

$$F(\nu_1, \nu_2) \quad \nu_1: \text{분자의자유도}, \ \nu_2: \text{분모의자유도}$$

$$F(a;\, \nu_1, \nu_2) = \frac{1}{F(1-a;\, \nu_2, \nu_1)}$$

두 모집단의 분산의 비율에 대한 추정

분산이 δ_1^2과 δ_2^2인 정규모집단으로부터 n_1과 n_2개의 표본을 독립적으로 추출하면

$$\frac{s_1^2 / \delta_1^2}{s_2^2 / \delta_2^2} = F(n_1 - 1, n_2 - 1)$$

$$\text{단}, \ s^2 = \frac{\sum_{i=1}^{n} (X_i - \overline{X})^2}{n-1}$$

F분포의 백분위수의 정의에 따라 다음과 같이 쓸 수 있다.

$$\mathrm{P}\left[\mathrm{F}\left(\frac{\alpha}{2}; n_1 - 1, n_2 - 1\right) \le \frac{s_1^2 / \delta_1^2}{s_2^2 / \delta_2^2} \le \mathrm{F}\left(1 - \frac{\alpha}{2}; n_1 - 1, n_2 - 1\right)\right] = 1 - \alpha$$

p.323그림참조

$$\mathrm{P}\left[\mathrm{F}\left(\frac{\alpha}{2}; n_2 - 1, n_1 - 1\right) \cdot \frac{s_1^2}{s_2^2} \le \frac{\delta_1^2}{\delta_2^2} \le \mathrm{F}\left(1 - \frac{\alpha}{2}; n_2 - 1, n_1 - 1\right) \cdot \frac{s_1^2}{s_2^2}\right] = 1 - \alpha$$

$\frac{\delta_1^2}{\delta_2^2}$에 대한 신뢰구간 양쪽 값이 모두 1보다 크게나오면 δ_1^2이 δ_2^2보다 크다.

신뢰구간 양쪽 값이 모두 1보다 작게나오면 δ_1^2이 δ_2^2보다 작다.

신뢰구간에 1이 포함되어 있으면 차이가 있다고 할 수 없다.

p.324 예제14.3

두 모집단의 분산의 비율에 대한 가설검정

- 단측검정의 예

  $H_0 : \delta_1^2 \le \delta_2^2 \qquad H_1 : \delta_1^2 > \delta_2^2$

  위 가설은 다음과 같이 나타낼 수 있다.

  $H_0 : \dfrac{\delta_1^2}{\delta_2^2} \le 1 \qquad H_1 : \dfrac{\delta_1^2}{\delta_2^2} > 1$

  $\dfrac{\delta_1^2}{\delta_2^2}$ 을 검정하기 위해서는 먼저 표본분산 $\dfrac{s_1^2}{s_2^2}$ 을 계산

  $\dfrac{\delta_1^2}{\delta_2^2} = 1$ 일 때, 즉 $\delta_1^2 = \delta_2^2$ 는 $\dfrac{s_1^2}{s_2^2} = F(n_1 - 1, n_2 - 1)$

  그러므로 결정규칙은 다음과 같다.

  만약 $\dfrac{s_1^2}{s_2^2} \le F(1-\alpha; n_1 - 1, n_2 - 1)$, H_0를 선택

  만약 $\dfrac{s_1^2}{s_2^2} > F(1-\alpha; n_1 - 1, n_2 - 1)$, H_1을 선택

- P.326 예제 14.4