# 2차 형태

# 정의

- 개요

1. n개 변수가 2차의 동차 다항식으로 이루어진 경우, 이를 2차 형태라고 한다.

1) 예를들어 분산 $\sum_{i=1}^{n}(x_i - \overline{x})^2$ = $\sum_{i=1}^{n}\left(x_i - \frac{\sum x_i}{n}\right)^2$ = $\sum_{i=1}^{n}({x_i}^2 + 2\overline{x}x_i + \overline{x}^2)$

$= \sum_{i=1}^{n} {x_i}^2 - 2\sum_{i=1}^{n}\frac{\sum x_i}{n}x_i + \mathrm{n}\sum_{i=1}^{n}\frac{\sum {x_i}^2}{n^2} = \frac{\mathrm{n}-1}{n}\sum_{i=1}^{n} {x_i}^2 + \frac{2}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} x_i\, x_j$

는 2차함수꼴로 나타낼 수 있다.

# 정의

- 2차형태 확률변수의 가법성

1. $Q = Q_1 + Q_2 + \cdots + Q_{k-1} + Q_k$가 각각 N($\mu_i$, $\sigma^2$)를 따르는 iid인 실2차형태라고 하자.
2. 이 때, $\frac{Q}{\sigma^2}, \frac{Q_1}{\sigma^2}, \cdots, \frac{Q_{k-1}}{\sigma^2}, \frac{Q_k}{\sigma^2}$는 각각 자유도가 $\mathrm{r}, r_1 \cdots r_k$ 를 따르는 $x^2$ 분포를 따른다.

1) 이 때, $\frac{Q}{\sigma^2} = \frac{Q_1}{\sigma^2} + \cdots + \frac{Q_{k-1}}{\sigma^2} + \frac{Q_k}{\sigma^2}$ 를 1.에서 다시 정의한다면

2) $\frac{Q_k}{\sigma^2}$는 $x^2(\mathrm{r} - r_1 - \cdots - r_{k-1})$인 카이제곱분포를 따른다.

# 예제

- F분포의 도출

1. (열차원의 도출) 전체 열의 통계량 $\bar{x}$를 각각 정의하면

1) 열평균 $\bar{x}_{cn} = \begin{bmatrix} \bar{x}_{.1} \\ \bar{x}_{.2} \\ \vdots \\ \bar{x}_{.b} \end{bmatrix} = \begin{bmatrix} \frac{X_{11}+X_{21}+\cdots+X_{a1}}{a} \\ \frac{X_{12}+X_{22}+\cdots+X_{a2}}{a} \\ \vdots \\ \frac{X_{1b}+X_{2b}+\cdots+X_{ab}}{a} \end{bmatrix} = \begin{bmatrix} \frac{\sum_{i=1}^{a} x_{i1}}{a} \\ \frac{\sum_{i=1}^{a} x_{i2}}{a} \\ \vdots \\ \frac{\sum_{i=1}^{a} x_{ib}}{a} \end{bmatrix}$

# 예제

- F분포의 도출

2) 이 때, 크기 n = ab인 확률표본의 분산 $S^2$은

(1) 분모인 (ab-1)를 $S^2$이 있는 항으로 이항하면

$(ab-1)S^2 = \sum_{i=1}^{a}\sum_{j=1}^{b}(x_{ij}-\bar{x})^2$

*(2)* $(x_{ij}-\bar{x})^2$ 를 $[(x_{ij}-\bar{x}_i)+(\bar{x}_i-\bar{x})]^2$ 로 분리하면

- $\sum_{i=1}^{a}\sum_{j=1}^{b}[(x_{ij}-\bar{x}_i)+(\bar{x}_i-\bar{x})]^2$

$= \sum_{i=1}^{a}\sum_{j=1}^{b}(x_{ij}-\bar{x}_i)^2 + \sum_{i=1}^{a}\sum_{j=1}^{b}(\bar{x}_i-\bar{x})^2 + 2\sum_{i=1}^{a}\sum_{j=1}^{b}(x_{ij}-\bar{x}_i)(\bar{x}_i-\bar{x})$

- 이 때, $\sum_{i=1}^{a}\sum_{j=1}^{b}(x_{ij}-\bar{x}_i)(\bar{x}_i-\bar{x}) = \sum_{i=1}^{a}\left[(\bar{x}_i-\bar{x})\sum_{j=1}^{b}(x_{ij}-\bar{x}_i)\right]$에서

$\sum_{j=1}^{b}(x_{ij}-\bar{x}_i) = 0$ 이므로, <u>이 교차항은 소거된다</u>.

(3) 한편, $\sum_{i=1}^{a}\sum_{j=1}^{b}(\bar{x}_i-\bar{x})^2 = \mathrm{b}\sum_{i=1}^{a}(\bar{x}_i-\bar{x})^2$ 이므로, 정리하면

- $\sum_{i=1}^{a}\sum_{j=1}^{b}(x_{ij}-\bar{x}_i)^2 + \mathrm{b}\sum_{i=1}^{a}(\bar{x}_i-\bar{x})^2$

# 예제

- F분포의 도출

3) 위를 2차형식으로 표현하면

*(1)* $\sum_{i=1}^{a}\sum_{j=1}^{b}\left(x_{ij}-\overline{x}_i\right)^2 + \mathrm{b}\sum_{i=1}^{a}(\overline{x}_i-\overline{x})^2 = Q_1 + Q_2 = \mathrm{Q}$

- 이 때, 전체 $\frac{Q_1}{\sigma^2} = \sum_{i=1}^{a}[\sum_{j=1}^{b}\frac{\left(x_{ij}-\overline{x}_i\right)^2}{\sigma^2}]$ 는 $x^2(b-1)$의 a개의 선형결합과 같다.
- 따라서 $x^2$의 가법성에 따라 $\sum_{i=1}^{a}x^2(b-1) \sim x^2\big(a(b-1)\big)$

(2) 한편, $Q_1 = \mathrm{Q} - Q_2$ 이고, 가법성에 따라 $Q_2 \sim x^2(r-r_1)$ 이므로

- $(r-r_1)$ $=\mathrm{ab}-1-a(b-1) = \mathrm{a}-1$
- 따라서 $Q_2 \sim x^2(\mathrm{a}-1)$

# 예제

- F분포의 도출

2. (행차원의 증명) 한편, 마찬가지로 행 평균을 정의하면

1) 행평균 $\bar{x}_{nr} = \begin{bmatrix} \bar{x}_{1.} \\ \bar{x}_{2.} \\ \vdots \\ \bar{x}_{a.} \end{bmatrix} = \begin{bmatrix} \frac{X_{11}+X_{12}+\cdots+X_{1b}}{b} \\ \frac{X_{21}+X_{22}+\cdots+X_{2b}}{b} \\ \vdots \\ \frac{X_{a1}+X_{a2}+\cdots+X_{ab}}{b} \end{bmatrix} = \begin{bmatrix} \frac{\sum_{i=1}^{b} x_{1i}}{b} \\ \frac{\sum_{i=1}^{b} x_{2i}}{b} \\ \vdots \\ \frac{\sum_{i=1}^{b} x_{ai}}{b} \end{bmatrix}$

# 예제

- F분포의 도출

2) 열평균을 도출한 것과 동일한 수준에서 논의를 진행하면

(1) 중간 단계를 생략하고, 바로 $\sum_{i=1}^{a}\sum_{j=1}^{b}\left(x_{ij}-\overline{x}_j\right)^2 + \mathrm{a}\sum_{j=1}^{b}\left(\overline{x}_j-\overline{x}\right)^2$ 에서

- $\sum_{i=1}^{a}\sum_{j=1}^{b}\left(x_{ij}-\overline{x}_j\right)^2 + \mathrm{a}\sum_{i=1}^{b}\left(\overline{x}_j-\overline{x}\right)^2 = Q_3 + Q_4 = \mathrm{Q}$

(2) 이 때, $\frac{Q_3}{\sigma^2} = \sum_{j=1}^{b}\left[\sum_{i=1}^{a}[\frac{\left(x_{ij}-\overline{x}_j\right)^2}{\sigma^2}\right]$는 $x^2(a-1)$의 b개의 선형결합과 같다.

- 따라서 $x^2$의 가법성에 따라 $\sum_{j=1}^{b} x^2(a-1) \sim x^2\left(b(a-1)\right)$

(3) $Q_4 = \mathrm{Q} - Q_3$ 에서 $Q_4 \sim x^2(r-r_3)$

- $(r-r_3)$ $=\mathrm{ab}-1-b(a-1) = \mathrm{b}-1$

- 따라서 $Q_4 \sim x^2(\mathrm{b}-1)$

# 예제

- F분포의 도출

3. 마지막으로 전체 평균을 정의하면

1) 전체 평균 $\overline{x}$ = $\frac{X_{11}+X_{12}+\cdots+X_{ab}}{a\cdot b}$ = $\frac{\sum_{i=1}^{n}\sum_{j=1}^{m}x_{ij}}{a\cdot b}$

2) 이 때 $x_{ij}-\overline{x}$ 를 $\overline{x}_{nr}$ 와 $\overline{x}_{cn}$ , $\overline{x}$ 로 다시 나타내면

(1) $(\overline{x}_{nr}-\overline{x})$ + $(\overline{x}_{cn}-x)$ + $(x_{ij}-\overline{x}_{nr}-\overline{x}_{cn}+\overline{x})$

(2) 따라서 (ab-1)$S^2$= $\sum_{i=1}^{a}\sum_{j=1}^{b}(x_{ij}-\overline{x})^2$ 는

- (ab-1)$S^2$ = $\mathrm{b}\sum_{i=1}^{a}(\overline{x}_i-\overline{x})^2+\mathrm{a}\sum_{i=1}^{b}\left(\overline{x}_j-\overline{x}\right)^2+\sum_{i=1}^{a}\sum_{j=1}^{b}\left(x_{ij}-\overline{x}_i-\overline{x}_j+\overline{x}\right)^2$

- 이를 2차형식 표현으로 표현하면 Q = $Q_2+Q_4+Q_5$

3) 이 때, $Q_2\sim\ x^2(\mathrm{a}-1)$ , $Q_4\sim\ x^2(\mathrm{b}-1)$ 이므로 $Q_5$는

(1) $(r-r_2-r_4)$ =$\mathrm{ab}-1-a+1-\mathrm{b}+1=(\mathrm{a}-1)(\mathrm{b}-1)$ 이다.

(2) 따라서 $Q_5\sim\ x^2((\mathrm{a}-1)(\mathrm{b}-1))$

# 예제

- F분포의 도출

4. 이제, 다음의 F통계량들을 정의할 수 있다.

1) 2차 형태의 통계량들이 서로 독립임을 가정했으므로

$$(1)\ \frac{Q_4}{Q_3} = \frac{Q_4 / \sigma^2(b-1)}{Q_3 / \sigma^2 b(a-1)} = \frac{Q_4 / (b-1)}{Q_3 / b(a-1)} \sim F(b-1, b(a-1))$$

$$(2)\ \frac{Q_4}{Q_5} = \frac{Q_4 / \sigma^2(b-1)}{Q_5 / \sigma^2((\mathrm{a}-1)(\mathrm{b}-1))} = \frac{Q_4 / (b-1)}{Q_5 / ((\mathrm{a}-1)(\mathrm{b}-1))} \sim F(b-1, (a-1)(b-1))$$

2) 이 통계량들은 충분통계량이 정규분포를 따르는 우도비 검정에서 활용할 수 있다.