일원배치 분산분석

정의

- 개요

1. 평균이 $\mu_1, \mu_2 \cdots \mu_b$로 다르고, 공통분산이 σ^2인 정규분포를 따르는 b개의 확률변수 $X_1, \cdots, X_b$ 를 정의하자.

1) 각각의 확률변수에서 추출한 확률표본 X_{ab} 을 $\mathrm{N}(\mu_i, \sigma^2)$ 을 따르는 iid라고 하자.

2) 이 때, 관측값에 대한 어떤 모형

(1) $x_{ij} = \mu_j + e_{ij}$를 고려하자.

(2) 단, 이 때 $e_{ij} \sim N(0, \sigma^2)$를 따른다.

3) 이제, 가설 <$H_0 : \mu_1 = \mu_2 = \cdots = \mu_b$ VS H_1 : 적어도 하나는 같지 않다> 를 검증한다.

(1) 총 모수공간 $\Omega = \{(\mu_1, \mu_2, \cdots, \mu_b, \sigma^2): -\infty < \mu_j < \infty,\ 0 < \sigma^2 < \infty\}$ 으로 놓고

(2) 가설공간 $\mathrm{w} = \{(\mu_1, \mu_2, \cdots, \mu_b, \sigma^2): -\infty < \mu_1 = \mu_2 = \cdots = \mu_b < \infty,\ 0 < \sigma^2 < \infty\}$

정의

- 개요

3) 우도비함수를 정의하면

(1) $L(\Omega) = \left[\frac{1}{\sqrt{2\pi}\sigma}\right]^{\frac{ab}{2}} \exp\left(-\frac{\sum\sum(x_{ij}-\mu)^2}{2\sigma^2}\right)$

(2) $L(w) = \left[\frac{1}{\sqrt{2\pi}\sigma}\right]^{\frac{ab}{2}} \exp\left(-\frac{\sum\sum(x_{ij}-\mu_j)^2}{2\sigma^2}\right)$

4) 우도비 검정을 정의하기 위해 MLE 추정량을 구하면

(1) $\log L(w)$

- $\frac{\partial \log L(w)}{\partial\mu} = \frac{\partial\left(\frac{ab}{2}log2\pi\sigma+\left[-\frac{1}{2\sigma^2}\sum\sum(x_{ij}-\mu)^2\right]\right)}{\partial\mu} = \frac{\sum\sum(x_{ij}-\mu)}{\sigma^2} = 0$

  따라서$\sum\sum x_{ij} - ab\mu = 0$에서 $\hat{\mu} = \frac{\sum\sum x_{ij}}{ab}$

정의

- 개요

4) 우도비 검정을 정의하기 위해 MLE 추정량을 구하면

(1) $\log \mathrm{L}(w)$

- $\frac{\partial \log \mathrm{L}(w)}{\partial \sigma^2} = \frac{\partial\left(\frac{ab}{2} log 2\pi\sigma + \left[-\frac{1}{2\sigma^2}\sum\sum(x_{ij}-\mu)^2\right]\right)}{\partial \sigma^2} = -\frac{ab}{2\sigma^2} + \frac{1}{2\sigma^4}\sum\sum(x_{ij}-\mu)^2 = 0$

따라서 $-\frac{ab}{2\sigma^2} + \frac{1}{2\sigma^4}\sum\sum(x_{ij}-\mu)^2 = 0$ 에서 $\widehat{\sigma^2} = \frac{\sum\sum(x_{ij}-\mu)^2}{ab}$

(2) 따라서 $\mathrm{L}(\widehat{w}) = \left[\frac{ab}{2\pi\sum\sum(x_{ij}-\mu)^2}\right]^{\frac{ab}{2}} \exp\left(-\frac{ab\sum\sum(x_{ij}-\mu)^2}{2\sum\sum(x_{ij}-\mu)^2}\right) = \left[\frac{ab}{2\pi\sum\sum(x_{ij}-\mu)^2}\right]^{\frac{ab}{2}} \exp\left(-\frac{ab}{2}\right)$

정의

- 개요

4) 우도비 검정을 정의하기 위해 MLE 추정량을 구하면

(3) $\log L(\Omega)$

- $\frac{\partial \log L(\Omega)}{\partial \mu_j} = \frac{\partial\left(\frac{ab}{2} log 2\pi\sigma + \left[-\frac{1}{2\sigma^2}\sum\sum(x_{ij}-\mu_j)^2\right]\right)}{\partial \mu_j} = \frac{\sum\sum(x_{ij}-\mu_j)}{\sigma^2} = 0$

  따라서 $\sum\sum x_{ij} - a\mu_j = 0$에서 $\widehat{\mu_j} = \frac{\sum\sum x_{ij}}{a}$ (단, j = b인 경우)

- $\frac{\partial \log L(\Omega)}{\partial \sigma^2} = \frac{\partial\left(\frac{ab}{2} log 2\pi\sigma + \left[-\frac{1}{2\sigma^2}\sum\sum(x_{ij}-\mu_j))^2\right]\right)}{\partial \sigma^2} = -\frac{ab}{2\sigma^2} + \frac{1}{2\sigma^4}\sum\sum(x_{ij}-\mu_j)^2 = 0$

  따라서 $-\frac{ab}{2\sigma^2} + \frac{1}{2\sigma^4}\sum\sum(x_{ij}-\mu_j)^2 = 0$ 에서 $\widehat{\sigma^2} = \frac{\sum\sum(x_{ij}-\mu_j)^2}{ab}$

(4) $L(\hat{\Omega}) = \left[\frac{ab}{2\pi\sum\sum(x_{ij}-\mu_j)^2}\right]^{\frac{ab}{2}} \exp\left(-\frac{ab\sum\sum(x_{ij}-\mu_j)^2}{2\sum\sum(x_{ij}-\mu_j)^2}\right) = \left[\frac{ab}{2\pi\sum\sum(x_{ij}-\mu_j)^2}\right]^{\frac{ab}{2}} \exp\left(-\frac{ab}{2}\right)$

정의

- 개요

5) 정리해서 우도비를 정의하면

$$(1)\ \frac{L(\hat{w})}{L(\hat{\Omega})} = \frac{\left[\frac{ab}{2\pi\sum\sum(x_{ij}-\mu)^2}\right]^{\frac{ab}{2}} \exp\left(-\frac{ab}{2}\right)}{\left[\frac{ab}{2\pi\sum\sum(x_{ij}-\mu_j)^2}\right]^{\frac{ab}{2}} \exp\left(-\frac{ab}{2}\right)} = \left[\frac{\sum\sum(x_{ij}-\mu)^2}{\sum\sum(x_{ij}-\mu_j)^2}\right]^{\frac{ab}{2}}$$

(2) $\sum\sum(x_{ij}-\mu)^2$ 혹은 $\sum\sum(x_{ij}-\mu_j)^2$는 정규분포의 완비충분통계량 Y = $\sum\sum x_{ij}^2$의 함수이다.

정의

- 개요

6) 따라서 $g_1(Y) = \sum\sum(x_{ij} - \mu)^2$ 혹은 $g_2(Y) = \sum\sum(x_{ij} - \mu_j)^2$ 를 잘 알려진 어떤 분포로 변환하면

(1) $\frac{\sum\sum(x_{ij}-\mu)^2}{ab}$ 는 실2차형식 $\frac{Q}{ab}$ 와 같고

(2) $\frac{\sum\sum(x_{ij}-\mu_j)^2}{ab}$는 실2차형식 $\frac{Q_3}{ab}$와 같다.

7) $\left[\frac{\mathrm{L}(\widehat{w})}{\mathrm{L}(\widehat{\Omega})}\right]^{\frac{ab}{2}} = \left[\frac{\frac{\sum\sum(x_{ij}-\mu)^2}{ab}}{\frac{\sum\sum(x_{ij}-\mu_j)^2}{ab}}\right]^{\frac{ab}{2}} = \left[\frac{\sum\sum(x_{ij}-\mu)^2}{\sum\sum(x_{ij}-\mu_j)^2}\right]^{\frac{ab}{2}} = \left[\frac{Q_3}{Q}\right]^{\frac{ab}{2}} = \Lambda$

(1) 따라서 $[\Lambda]^{\frac{ab}{2}} = \frac{Q_3}{Q}$ = Z 이다.

(2) 이 때, $Q = Q_3 + Q_4$ 이므로, Z를 수정하면

- $\frac{Q_3}{Q_3 + Q_4} = \frac{1}{1+\frac{Q_4}{Q_3}}$ 이다.

8) 마지막으로, 최량기각역을 설정하면

(1) A = $P_{H_0}[\frac{1}{1+\frac{Q_4}{Q_3}} \leq Z] = P_{H_0}[\frac{Q_4/_{b-1}}{Q_3/_{[b(a-1)]}} \geq c(Z)]$

(2) 이 때, c(Z) = $\frac{[b(a-1)]}{b-1}[\Lambda]^{-\frac{2}{ab}-1}$ 이다.