Stat 565
1-9-18

Chapter 3 ANOVA

(1)

1-way analysis of variance

Model: $y_{ij} = \mu + \tau_i + \varepsilon_{ij}$ $\qquad i = 1, .., a$

$j = 1, .., n$

$N = na$

Assumptions: $\sum_{i=1}^{a} \tau_i = 0$

$E[\varepsilon_{ij}] \equiv 0$, $V[\varepsilon_{ij}] \equiv \sigma^2$, ε_{ij} are pairwise uncorrelated

Notation: $y_{i.} = \sum_{j=1}^{n} y_{ij}$ $\qquad y_{..} = \sum_{i=1}^{a}\sum_{j=1}^{n} y_{ij}$

$\bar{y}_{i.} = \frac{1}{n}\sum_{j=1}^{n} y_{ij}$ $\qquad \bar{y}_{..} = \frac{1}{N}\sum_{i=1}^{a}\sum_{j=1}^{n} y_{ij}$

(2)

Note: $\mu_i = \mu + \tau_i$

$H_0: \mu_1 = \mu_2 = \dots = \mu_a$

OR

$H_0: \tau_1 = \tau_2 = \dots = \tau_a = 0$

Total sum of squares

$$SS_T = \sum_{i=1}^{a}\sum_{j=1}^{n} (y_{ij} - \bar{y}_{..})^2$$

$$= \sum_i \sum_j \big(\underbrace{y_{ij} - \bar{y}_{i.}} + \underbrace{\bar{y}_{i.} - \bar{y}_{..}}\big)^2$$

3

$$SS_T = \sum_i \sum_j (y_{ij} - \bar{y}_{i\cdot})^2 + \sum_i \sum_j (\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})^2$$

$$+ 2 \sum_i \sum_j (y_{ij} - \bar{y}_{i\cdot})(\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})$$

(1) (2) (3)

(3): $$2 \sum_i \left[(\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot}) \underbrace{\sum_j (y_{ij} - \bar{y}_{i\cdot})}_{\cancel{y_{i\cdot} - n\bar{y}_{i\cdot}}\ 0} \right]$$

(2): $$n \sum_{i=1}^{a} (\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})^2 = SS_{TRT}$$

4

(1): $$\sum_{i=1}^{a} \sum_{j=1}^{n} (y_{ij} - \bar{y}_{i\cdot})^2 = SS_E$$

So $SS_T = SS_{TRT} + SS_E$

Find the expected value of each component

$$E[SS_E] = E\Big[\sum_{i=1}^{a} \underbrace{\sum_{j=1}^{n} (y_{ij} - \bar{y}_{i\cdot})^2}_{(n-1)S_i^2} \Big]$$

$$= \sum_{i=1}^{a} (n-1) \underbrace{E[S_i^2]}_{\sigma^2} = a(n-1)\sigma^2 = (N-a)\sigma^2$$

Define $MS_E = \frac{SS_E}{N-a}$ ⑤

Then $E[MS_E] = \sigma^2$

$$E[SS_{TRT}] = E\left[n\sum_{i=1}^{a}(\bar{y}_{i\cdot}-\bar{y}_{\cdot\cdot})^2\right]$$

$$= nE\left[\sum_{i=1}^{a}(\bar{y}_{i\cdot}^2+\bar{y}_{\cdot\cdot}^2-2\bar{y}_{i\cdot}\bar{y}_{\cdot\cdot})\right]$$

$$= nE\left(\sum_{i=1}^{a}\bar{y}_{i\cdot}^2 + a\bar{y}_{\cdot\cdot}^2 - 2\bar{y}_{\cdot\cdot}\underbrace{\sum_{i=1}^{a}\bar{y}_{i\cdot}}_{\frac{1}{n}\sum_{i=1}^{a}\sum_{j=1}^{n}y_{ij}}\right)$$

⑥

$$= nE\left[\sum_{i=1}^{a}\bar{y}_{i\cdot}^2 + a\bar{y}_{\cdot\cdot}^2 - 2\bar{y}_{\cdot\cdot}a\bar{y}_{\cdot\cdot}\right]$$

$$= nE\left[\sum_{i=1}^{a}\bar{y}_{i\cdot}^2 - a\bar{y}_{\cdot\cdot}^2\right]$$

$$\bar{y}_{i\cdot} = \frac{1}{n}\sum_{j=1}^{n}y_{ij} = \frac{1}{n}\sum_{j=1}^{n}(\mu+\tau_i+\varepsilon_{ij})$$

$$E[\bar{y}_{i\cdot}] = \frac{1}{n}\sum_{j=1}^{n}(\mu+\tau_i+0)$$

$$= \mu+\tau_i = \mu_i$$

⑦

$$V[\bar{y}_{i\cdot}] = V\left[\frac{1}{n}\sum_{j=1}^{n}(\mu + \tau_i + \varepsilon_{ij})\right]$$

$$= \frac{1}{n^2}\sum_{j=1}^{n}\underbrace{V(\varepsilon_{ij})}_{\sigma^2} = \frac{\sigma^2}{n}$$

Also, $\bar{y}_{\cdot\cdot} = \frac{1}{N}\sum_{i=1}^{a}\sum_{j=1}^{n} y_{ij} = \frac{1}{N}\sum_{i}\sum_{j=1}^{n}(\mu + \tau_i + \varepsilon_{ij})$

$$= \mu + 0 + \varepsilon_{\cdot\cdot}$$

$$E[\bar{y}_{\cdot\cdot}] = E[\mu + \varepsilon_{\cdot\cdot}] = \mu$$

$$V[\bar{y}_{\cdot\cdot}] = V[\mu + \varepsilon_{\cdot\cdot}] = V[\varepsilon_{\cdot\cdot}] = \frac{\sigma^2}{N}$$

⑧

Now $E[\bar{y}_{i\cdot}^2] = V[\bar{y}_{i\cdot}] + \left(E[\bar{y}_{i\cdot}]\right)^2$

$$= \frac{\sigma^2}{n} + \mu_i^2$$

and $E[\bar{y}_{\cdot\cdot}^2] = V[\bar{y}_{\cdot\cdot}] + \left(E[\bar{y}_{\cdot\cdot}]\right)^2$

$$= \frac{\sigma^2}{N} + \mu^2$$

We had $E[SS_{TRT}] = n\,E\left[\sum_{i=1}^{a}\bar{y}_{i\cdot}^2 - a\bar{y}_{\cdot\cdot}^2\right]$

$$= n\left[\sum_{i=1}^{a}\left(\frac{\sigma^2}{n} + \mu_i^2\right) - a\left(\frac{\sigma^2}{N} + \mu^2\right)\right]$$

9

$$= a\sigma^2 + n\sum_{i=1}^{a}(\mu+\tau_i)^2 - \sigma^2 - N\mu^2$$

$$= (a-1)\sigma^2 + N\cancel{\mu^2} + n\sum_{i=1}^{a}\tau_i^2 + n2\mu\overbrace{\sum_{i=1}^{a}\tau_i}^{0} - \cancel{N\mu^2}$$

So $E[SS_{TRT}] = (a-1)\sigma^2 + n\sum_{i=1}^{a}\tau_i^2$

Define $MS_{TRT} = \dfrac{SS_{TRT}}{a-1}$

$$E[MS_{TRT}] = \sigma^2 + \frac{n}{a-1}\sum_{i=1}^{a}\tau_i^2$$

Note: Under H_0, $E[MS_{TRT}] = \sigma^2$