Transformations of Y in a 1-way ANOVA

Stat 506
5-20-25

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

$$\varepsilon_{ij} \sim \text{iid } N(0, \sigma^2)$$

1

$H_0: \tau_i \equiv 0$ OR $H_0: \mu_1 = \mu_2 = \cdots \mu_a$

OR $H_0: E_1(y) = E_2(y) = \dots = E_a(y)$

Suppose we let $x_{ij} = f(y_{ij})$ and then

$$x_{ij} = \mu + \tau_i + \varepsilon_{ij}$$

all different now

2

$H_0: E_1(x) = \dots = E_a(x)$

OR $H_0: E_1(f(y)) = \dots = E_a(f(y))$

Note that $E(f(y)) \neq f(E(y))$

unless $f$ is linear

Box-Cox transformation

$$x_{ij} = \frac{y_{ij}^{\lambda} - 1}{\lambda}$$

(3)

Let $g(y_{ij})$ be the probability density of $y_{ij}$

Then $g(y_{ij}) = \underbrace{h(x_{ij})}_{\substack{\text{density} \\ \text{of } x_{ij}}} \left| \frac{\partial x_{ij}}{\partial y_{ij}} \right|$

(this works since the Box Cox transformation is monotonic)

We want $h(x_{ij})$ to be the normal density

(4)

$$g(y_{ij}) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x_{ij}-\mu-\tau_i}{\sigma}\right)^2} \left| \frac{\partial}{\partial y_{ij}} \frac{y_{ij}^{\lambda}-1}{\lambda} \right|$$

(5)

$$g(\vec{y}) = \prod_{i=1}^{a} \prod_{j=1}^{n} g(y_{ij})$$

$$= \sigma^{-an} (2\pi)^{-\frac{1}{2}an} e^{-\frac{1}{2\sigma^2} \sum_i \sum_j (x_{ij} - \mu - \tau_i)^2} \prod_{i=1}^{a} \prod_{j=1}^{n} |y_{ij}|^{\lambda-1}$$

$$= L(\mu, \sigma, \tau_1, \ldots, \tau_a, \lambda)$$

Goal: find the values of the parameters that will maximize $L$ (the likelihood function)

Maximize the log of $L$: (6)

$$l(\mu, \sigma, \tau_1, \ldots, \tau_a, \lambda) = \ln L(\cdots)$$

$$= -an \ln\sigma - \frac{1}{2} an \ln(2\pi) - \frac{1}{2\sigma^2} \sum_i \sum_j (x_{ij} - \mu - \tau_i)^2$$

$$+ (\lambda - 1) \sum_{i=1}^{a} \sum_{j=1}^{n} \ln|y_{ij}|$$

$$\frac{\partial l}{\partial \mu} = -\frac{\cancel{1}}{\cancel{2}\sigma^2} \sum_i \sum_j \cancel{2}(x_{ij} - \mu - \tau_i)(\cancel{-1}) \overset{set}{=} 0$$

$$x_{..} - an\mu - 0 = 0 \Rightarrow \hat{\mu} = \frac{x_{..}}{an} = \bar{x}_{..}$$

(7)

$$\frac{\partial l}{\partial \tau_i} = -\frac{\cancel{1}}{\cancel{2}\sigma^2} \sum_j \cancel{2}(x_{ij} - \mu - \tau_i)(\cancel{-1}) \overset{set}{=} 0$$

$$x_{i.} - n\mu - n\tau_i = 0$$

$$\hat{\tau}_i = \frac{x_{i.}}{n} - \hat{\mu}$$

$$\hat{\tau}_i = \bar{x}_{i.} - \bar{x}_{..}$$

$$\frac{\partial l}{\partial \sigma} = -\frac{an}{\sigma} - \frac{1}{2}\sum_i \sum_j (x_{ij} - \mu - \tau_i)^2 (-2\sigma^{-3}) \overset{set}{=} 0$$

$$-an\sigma^2 + \sum_i \sum_j (x_{ij} - \mu - \tau_i)^2 = 0$$

(8)

$$\hat{\sigma}^2 = \frac{\sum_i \sum_j (x_{ij} - \hat{\mu} - \hat{\tau}_i)^2}{an}$$

$$= \frac{\sum_i \sum_j (x_{ij} - \cancel{\bar{x}_{..}} - (\bar{x}_{i.} - \cancel{\bar{x}_{..}}))^2}{an}$$

$$\hat{\sigma}^2 = \frac{\sum_i \sum_j (x_{ij} - \bar{x}_{i.})^2}{an}$$

(9)

Now,

$$L(\hat{\mu}, \hat{\tau}_1, \ldots, \hat{\tau}_a, \hat{\sigma}, \lambda) =$$

$$\hat{\sigma}^{-an} (2\pi)^{-\frac{1}{2}an} e^{-\frac{1}{2\hat{\sigma}^2} \sum_i \sum_j (x_{ij} - \hat{\mu} - \hat{\tau}_i)^2} \left( \prod_i \prod_j |y_{ij}| \right)^{\lambda - 1}$$

$$= \hat{\sigma}^{-an} (2\pi)^{-\frac{1}{2}an} e^{-\frac{an}{2}} \left( \prod_i \prod_j |y_{ij}| \right)^{\lambda - 1}$$

$$= C \hat{\sigma}^{-an} \left( \prod_i \prod_j |y_{ij}| \right)^{\lambda - 1}$$

(10)

Maximizing this is equivalent to minimizing:

$$an \left[ \frac{\hat{\sigma}^{an}}{\left( \prod_i \prod_j |y_{ij}| \right)^{\lambda - 1}} \right]^{\frac{2}{an}}$$

$$= \frac{an \hat{\sigma}^2}{\left( \sqrt[an]{\prod_i \prod_j |y_{ij}|} \right)^{2(\lambda - 1)}} = \frac{an \hat{\sigma}^2}{\dot{y}^{2(\lambda - 1)}}$$

geometric mean

$$= \frac{\sum_i \sum_j (x_{ij} - \bar{x}_{i\cdot})^2}{\dot{y}^{2(\lambda-1)}}$$

$$= \sum_i \sum_j \left( \frac{x_{ij} - \bar{x}_{i\cdot}}{\dot{y}^{\lambda-1}} \right)^2 \qquad \text{Let } w_{ij} = \frac{x_{ij}}{\dot{y}^{\lambda-1}}$$

$$= \sum_i \sum_j (w_{ij} - \bar{w}_{i\cdot})^2 = \text{SSE for the } w_{ij}\text{'s}$$

$$\left( \text{where } w_{ij} = \frac{y_{ij}^{\lambda} - 1}{\lambda \dot{y}^{\lambda-1}} \right)$$

(12)

Choose $\lambda = 0, .1, .2, \ldots, .9$

Create the $w_{ij}$'s for each $\lambda$.

Run a 1-way ANOVA in each case
+ record the SSE.

Pick the value of $\lambda$ that gave the smallest SSE.

Finally, run your 1-way ANOVA on the

$$x_{ij} = \frac{y_{ij}^{\lambda} - 1}{\lambda}$$