20th March' 2024

Continuing with Sec 6.3 :

Two Sample Models:

Let's look at Figure 6.5 :→

There is clear overlap in medians, Boxes, whiskers and even outliers between the box plots. Thus, it seems like the work term salaries have similar distributions..

[★ Note : Take your time to practice these derivations of L.R.S ; All are considered ch 2, 3, 4 Materials. If they are not "obvious" to you. See me in office hours for help if needed ★]

Case 1 : Assume $\sigma_1^2 = \sigma_2^2 = \sigma^2$

Basic Model

$$L(\mu_1, \mu_2) = \prod_{i=1}^{n_1} \exp\left\{-\frac{1}{2\sigma^2}(y_{1i} - \mu_1)^2\right\} \cdot \prod_{j=1}^{n_2} \exp\left\{-\frac{1}{2\sigma^2}(y_{2j} - \mu_2)^2\right\}$$

$$\vdots$$

$$l(\mu_1, \mu_2) = -\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n_1}(y_{1i} - \mu_1)^2 + \sum_{j=1}^{n_2}(y_{2j} - \mu_2)^2\right]$$

If we take $\frac{\partial l}{\partial \mu_1}$ and $\frac{\partial l}{\partial \mu_2}$, and solve for $(\hat{\mu}_1, \hat{\mu}_2)$:

$$\hat{\mu}_1 = \sum_{i=1}^{n_1} y_{1i} / n_1 = \bar{y}_1$$

$$\hat{\mu}_2 = \sum_{j=1}^{n_2} y_{2j} / n_2 = \bar{y}_2$$

$$\boxed{K = 2}$$

Hypothesized Model : $\mu_1 = \mu_2 = \mu$ (unknown)

$$\boxed{q = 1}$$

$$l(\mu) = -\frac{1}{2s^2}\left[\sum_{i=1}^{n_1}(y_{1i}-\mu)^2 + \sum_{j=1}^{n_2}(y_{2j}-\mu)^2\right]$$

$$\hookrightarrow \left.\frac{\partial l}{\partial \mu}\right|_{\hat{\mu}} = 0 \Rightarrow \tilde{\mu} = \frac{\sum_{i=1}^{n_1} y_{1i} + \sum_{j=1}^{n_2} y_{2j}}{n_1 + n_2} \equiv \bar{y}$$

Average over both work terms.

Testing the Model

$$D = 2\left[l(\hat{\mu}_1, \hat{\mu}_2) - l(\tilde{\mu}, \tilde{\mu})\right]$$

$=$ ... in the lecture notes ...

$$= \frac{(\bar{y}_1 - \bar{y}_2)^2}{s^2(1/n_1 + 1/n_2)} = Z^2 \sim \chi^2_{(1)} \qquad \begin{matrix} k=2 \\ q=1 \end{matrix}$$

$n_1 = 352$, $\bar{y}_1 = \$3148.532$ $s^2 = 818^2$
$n_2 = 308$, $\bar{y}_2 = \$3354.483$
use these to get d_{obs}.

$$d_{obs} = 10.4129$$

$$p\text{-value} = P(D \geq d_{obs}) = P(\chi^2_{(1)} \geq 10.4129)$$

$$= P(|Z| \geq \sqrt{10.4129})$$

$$= 2\underbrace{P(Z \geq 3.23)}$$

Same. $\hookrightarrow P(Z \leq -3.23)$

$$= 2 * 0.0006$$

$$= 0.0012 \quad \text{via tables.}$$

95% C.I : $(\bar{y}_1 - \bar{y}_2) \pm Z_{0.975} * S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$

-205.95

$\Rightarrow (-331.04, -80.86)$

There is strong evidence against H_0, given p-value = 0.0012.
The estimated mean monthly increase in Work Term 2 over Work Term 1 is \$205.95 (95% C.I : \$80.86 – \$331.04)

The data suggest the increase in mean monthly salary is significantly different from zero, at 5% level of significance. (or 1% level)

[Here we looked at the other two examples in the lecture notes (unknown & unequal variances, unknown but equal variances)

↳ Corresponds to pages 20-22 of the ch-6 pdf under "Lecture Notes".