Inverse Regression

H.M. James Hung

The simple linear regression model commonly seen is

$$y = \alpha + \beta x + \varepsilon \tag{1}$$

where x is a fixed variable (e.g., can be controlled), $E(\varepsilon) = 0$ and $V(\varepsilon) = \sigma^2$.

The OLS estimators for (α, β) based on n independent data $(y_1, x_1), \ldots, (y_n, x_n)$ are

$$\hat{\beta} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})y_i}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \quad , \quad \hat{\alpha} = \bar{y} - \hat{\beta}\bar{x} \tag{2}$$

When x is a random variable, (1) can be considered as

$$E(\, y \mid x) = \alpha + \beta x$$

Also, (1) leads to

$$x = \frac{y-\alpha}{\beta} - \varepsilon = -\frac{\alpha}{\beta} + \frac{1}{\beta} y - \varepsilon$$

This is the inverse regression model from (1), where the new regression coefficients are $\frac{\alpha}{\beta}$ and $\frac{1}{\beta}$.

If we estimate $\frac{\alpha}{\beta}$ and $\frac{1}{\beta}$ by $\frac{\widehat{\alpha}}{\widehat{\beta}}$ and $\frac{1}{\widehat{\beta}}$, respectively, using (2), **is $\boldsymbol{\frac{1}{\widehat{\beta}}}$ unbiased for $\boldsymbol{\frac{1}{\beta}}$?** **No.** $\boldsymbol{E(\frac{1}{\widehat{\beta}}) \neq \frac{1}{\beta}}$.

Likewise, $\boldsymbol{E(\frac{\widehat{\alpha}}{\widehat{\beta}}) \neq \frac{\alpha}{\beta}}$.

Exercises

1. Create a hypothetical data set and then perform simple linear regression analysis and the corresponding inverse regression analysis.
2. Construct unbiased estimators for the new intercept and new slope in the regression of x on y.