Packet 5: Linear Regression

Chap 6.5 Linear Regression Model

From mathematics to statistics: Up to this point our construction of the least square estimators is completely out of mathematical intuition. Without a probabilistic model (i.e. a set of distributional assumptions), however, we cannot evaluate the statistical properties of these estimators. Under the sampling view, what happens if we repeat the experiment many many times? In particular, we cannot address pertinent questions such as the reliability of the slope and intercept estimates in the presence of "noise". Next, let us introduce some assumptions, and evaluate the properties of the estimators accordingly.

The simplest statistical model for the data pairs (x_i, y_i) is the linear regression model.

$$Y_i = \alpha + \beta(x_i - \bar{x}) + \epsilon_i, \quad i = 1, 2, \ldots, n.$$

where the ϵ_i are i.i.d. random errors with zero mean and common variance

$$\epsilon_i \sim N(0, \sigma^2), \quad \text{for } i = 1, 2, \ldots, n.$$

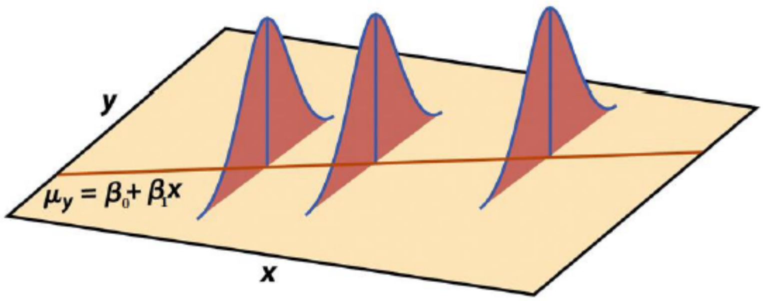
Let us assume that the x's are fixed (not random).

It is important to note the set of assumptions we are making here, i.e., LINE.

Under LINE Assumption

With the above assumptions, let us check properties of the least square estimators. The error terms ϵ_i's are i.i.d. Normal$(0, \sigma^2)$ r.v.s, i.e.,

$$Y_i = \alpha + \beta(x_i - \bar{x}) + \epsilon_i, \quad \epsilon_i \sim N(0, \sigma^2).$$



In this case, let us find the maximum likelihood estimators, $(\hat{\alpha}, \hat{\beta}, \hat{\sigma}^2)$.

The likelihood function is

$$L(\alpha, \beta, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-\frac{\sum_{i=1}^{n}(y_i - \alpha - \beta(x_i - \bar{x}))^2}{2\sigma^2}}.$$

Maximizing this log-likelihood with respect to α and β is the same as minimizing

$$\sum_{i=1}^{n}(y_i - \alpha - \beta(x_i - \bar{x}))^2.$$

More Examples: 6.5-1

So, the least square solutions must be the MLEs.

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

The MLE of σ^2 can be derived by taking a partial derivative on the log-likelihood with respect to σ^2.

We note that the distribution of the residual sum of squares divided by σ^2 is χ^2_{n-2}, i.e.,

Estimator properties: $\hat{\alpha}$ and $\hat{\beta}$ are both linear functions of random variables, Y_i's.

$$\hat{\alpha} = \bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$$

$$\hat{\beta} = \frac{\sum_{i=1}^{n}(Y_i - \bar{Y})(x_i - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$