DSA 8020 Spring 2024

Statistical Methods II

Derivation of Least Squares Estimator

Instructor: Whitney Huang

We want to minimize the sum of squared errors:

$$\ell(\beta_0, \beta_1) = \sum_{i=1}^{n} \left[y_i - (\beta_0 + \beta_1 x_i)\right]^2,$$

in order to obtain the least squares estimators of β_0 and β_1, denoted by $\hat{\beta}_0$ and $\hat{\beta}_1$, respectively. Let's work with $\hat{\beta}_0$, the least squares estimator for β_0 by taking the partial derivative:

$$\begin{aligned}
\frac{\partial \ell}{\partial \beta_0} &= \frac{\partial \sum_{i=1}^{n} \left[y_i - (\beta_0 + \beta_1 x_i)\right]^2}{\partial \beta_0} \\
&= 2 \sum_{i=1}^{n} \left[y_i - (\beta_0 + \beta_1 x_i)\right](-1) \\
&= 2 \sum_{i=1}^{n} -y_i + \sum_{i=1}^{n} (\beta_0 + \beta_1 x_i) \\
&= 2 \left[\sum_{i=1}^{n} -y_i + n\beta_0 + \beta_1 \sum_{i=1}^{n} x_i\right]
\end{aligned}$$

To get $\hat{\beta}_0$, we need to set the above quantity to zero:

$$\begin{aligned}
&\cancel{2} \left[\sum_{i=1}^{n} -y_i + n\beta_0 + \beta_1 \sum_{i=1}^{n} x_i\right] \overset{set}{=} 0 \\
&\Rightarrow n\beta_0 = \sum_{i=1}^{n} (y_i - \beta_1 x_i) \\
&\Rightarrow \beta_0 = \frac{\sum_{i=1}^{n} (y_i - \beta_1 x_i)}{n} \\
&\Rightarrow \beta_0 = \frac{\sum_{i=1}^{n} y_i - \beta_1 \sum_{i=1}^{n} x_i}{n} \\
&\Rightarrow \hat{\beta}_0 = \bar{y} - \beta_1 \bar{x}
\end{aligned}$$

Next, let's work with $\hat{\beta}_1$:

$$\begin{aligned}
\frac{\partial L}{\partial \beta_1} &= \frac{\partial \sum_{i=1}^{n} \left[y_i - (\beta_0 + \beta_1 x_i)\right]^2}{\partial \beta_1} \\
&= \frac{\partial \sum_{i=1}^{n} \left[y_i - (\bar{y} - \beta_1 \bar{x} + \beta_1 x_i)\right]^2}{\partial \beta_1} \\
&= 2 \sum_{i=1}^{n} \left[y_i - \bar{y} - \beta_1 (x_i - \bar{x})\right] \left(-(x_i - \bar{x})\right)
\end{aligned}$$

Again, set this quantity to zero:

$$\begin{aligned}
&\cancel{2} \sum_{i=1}^{n} \left[y_i - \bar{y} - \beta_1 (x_i - \bar{x})\right] \left(-(x_i - \bar{x})\right) \overset{set}{=} 0 \\
&\beta_1 \sum_{i=1}^{n} \left(x_i - \bar{x}\right)^2 = \sum_{i=1}^{n} \left(y_i - \bar{y}\right)\left(x_i - \bar{x}\right) \\
&\Rightarrow \hat{\beta}_1 = \frac{\sum_{i=1}^{n} \left(y_i - \bar{y}\right)\left(x_i - \bar{x}\right)}{\sum_{i=1}^{n} \left(x_i - \bar{x}\right)^2}
\end{aligned}$$