Lecture 2

August 28, 2018

Simple Linear Regression

- Here there is only one explanatory variable x. We denote the data in simple linear regression as $(x_1, y_1), \ldots, (x_n, y_n)$.

Simple Linear Regression

- Here there is only one explanatory variable x. We denote the data in simple linear regression as $(x_1, y_1), \ldots, (x_n, y_n)$.
- We assume that these are independent with the relation of y_i to x_i expressed as

$$\mathbb{E}(y_i|x_i) = \beta_0 + \beta_1 x_i, \quad i = 1, \ldots, n.$$

Simple Linear Regression

- Here there is only one explanatory variable x. We denote the data in simple linear regression as $(x_1, y_1), \ldots, (x_n, y_n)$.
- We assume that these are independent with the relation of y_i to x_i expressed as

$$\mathbb{E}(y_i|x_i) = \beta_0 + \beta_1 x_i, \quad i = 1, \ldots, n.$$

Why can we also say E(Y)=Xb,ignoring conditional expectation?

- We can alternatively write the simple linear regression model as

$$y_i = \beta_0 + \beta_1 x_i + e_i, \quad \text{where } \mathbb{E}(e_i|x_i) = 0.$$

Parameter Estimation

- How does one estimate the parameters β_0 and β_1?

Parameter Estimation

- How does one estimate the parameters β_0 and β_1?
- The most popular way is to use the method of least squares.

Parameter Estimation

- How does one estimate the parameters β_0 and β_1?
- The most popular way is to use the method of least squares.
- The least squares estimates of β_0 and β_1 are those values b_0 and b_1 that minimize:

$$Q(b_0, b_1) = \sum_{i=1}^{n} (y_i - b_0 - b_1 x_i)^2$$

Parameter Estimation

- How does one estimate the parameters β_0 and β_1?
- The most popular way is to use the method of least squares.
- The least squares estimates of β_0 and β_1 are those values b_0 and b_1 that minimize:

$$Q(b_0, b_1) = \sum_{i=1}^{n} (y_i - b_0 - b_1 x_i)^2$$

- These estimates can be computed by simply differentiating Q with respect to b_0 and b_1 and setting the partial derivatives equal to zero. This gives

$$\frac{\partial Q}{\partial b_0} = -2 \sum_{i=1}^{n} (y_i - b_0 - b_1 x_i) = 0,$$

and

Parameter Estimation

- $$\frac{\partial Q}{\partial b_1} = -2\sum_{i=1}^{n} x_i(y_i - b_0 - b_1 x_i) = 0.$$

Parameter Estimation

- $$\frac{\partial Q}{\partial b_1} = -2\sum_{i=1}^{n} x_i(y_i - b_0 - b_1 x_i) = 0.$$

- These equations can be equivalently written as

$$nb_0 + b_1 \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i, \quad \text{and} \quad b_0 \sum_{i=1}^{n} x_i + b_1 \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i y_i.$$

Parameter Estimation

- $$\frac{\partial Q}{\partial b_1} = -2\sum_{i=1}^{n} x_i(y_i - b_0 - b_1 x_i) = 0.$$
- These equations can be equivalently written as
$$nb_0 + b_1\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i, \quad \text{and} \quad b_0\sum_{i=1}^{n} x_i + b_1\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i y_i.$$
- The above two equations are known as the normal equations. These can be easily solved to give
$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})y_i}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$
and
$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x},$$
where we used the notation $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$, and $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$.

- β_0 and β_1 are called the Least Squares estimates of β_0 and β_1 respectively. Sometimes they are also called OLS estimates (OLS stands for Ordinary Least Squares).

- β_0 and β_1 are called the Least Squares estimates of β_0 and β_1 respectively. Sometimes they are also called OLS estimates (OLS stands for Ordinary Least Squares).
- Note that for these estimates to make sense, we need $\sum_{i=1}^{n}(x_i - \bar{x})^2 > 0$ which means that the x-variables are not all the same.

The OLS Regression Line

- The OLS regression line is given by $y = \hat{\beta}_0 + \hat{\beta}_1 x$. This is the line with slope $\hat{\beta}_1$ and intercept $\hat{\beta}_0$. The regression line can be used to predict y given x.

The OLS Regression Line

- The OLS regression line is given by $y = \hat{\beta}_0 + \hat{\beta}_1 x$. This is the line with slope $\hat{\beta}_1$ and intercept $\hat{\beta}_0$. The regression line can be used to predict y given x.
- There is a very interesting alternative way of writing the regression line. This uses the following notation. Let

$$s_x := \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2} \quad \text{and} \quad s_y := \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2}.$$

The OLS Regression Line

- The OLS regression line is given by $y = \hat{\beta}_0 + \hat{\beta}_1 x$. This is the line with slope $\hat{\beta}_1$ and intercept $\hat{\beta}_0$. The regression line can be used to predict y given x.
- There is a very interesting alternative way of writing the regression line. This uses the following notation. Let

$$s_x := \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}, \quad \text{and} \quad s_y := \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2}.$$

- Also let

$$r := \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i - \bar{x}}{s_x}\right)\left(\frac{y_i - \bar{y}}{s_y}\right),$$

The OLS Regression Line

- s_x and s_y denote the sample standard deviations of x and y respectively and r denotes the sample correlation coefficient between x and y. It is always true that $0 \leq |r| \leq 1$ (why?).

The OLS Regression Line

- s_x and s_y denote the sample standard deviations of x and y respectively and r denotes the sample correlation coefficient between x and y. It is always true that $0 \leq |r| \leq 1$ (why?).
- With this notation, it is easy to see that $\hat{\beta}_1 = rs_y/s_x$.

The OLS Regression Line

- s_x and s_y denote the sample standard deviations of x and y respectively and r denotes the sample correlation coefficient between x and y. It is always true that $0 \leq |r| \leq 1$ (why?).
- With this notation, it is easy to see that $\hat{\beta}_1 = rs_y/s_x$. Hence, the regression line equals

$$\frac{y - \bar{y}}{s_y} = r\left(\frac{x - \bar{x}}{s_x}\right).$$

The OLS Regression Line

- s_x and s_y denote the sample standard deviations of x and y respectively and r denotes the sample correlation coefficient between x and y. It is always true that $0 \leq |r| \leq 1$ (why?).
- With this notation, it is easy to see that $\hat{\beta}_1 = rs_y/s_x$. Hence, the regression line equals

$$\frac{y - \bar{y}}{s_y} = r\left(\frac{x - \bar{x}}{s_x}\right).$$

- Note that the regression line passes through the point $(\bar{y}, \bar{x})$.

Expected values and variances of the OLS estimators

- Recall the linear regression model described before. The quantities β_0 and β_1 are parameters of the model. These are estimated by $\hat{\beta}_0$ and $\hat{\beta}_1$. How good are these estimators?

.

Expected values and variances of the OLS estimators

- Recall the linear regression model described before. The quantities β_0 and β_1 are parameters of the model. These are estimated by $\hat{\beta}_0$ and $\hat{\beta}_1$. How good are these estimators?
- We write X for the set of explanatory variable values $\{x_1, \ldots, x_n\}$. Under the assumptions of the linear model, it is easy to see that $\mathbb{E}(y_i|X) = \beta_0 + \beta_1 x_i$.

.

Expected values and variances of the OLS estimators

- Recall the linear regression model described before. The quantities β_0 and β_1 are parameters of the model. These are estimated by $\hat{\beta}_0$ and $\hat{\beta}_1$. How good are these estimators?
- We write X for the set of explanatory variable values $\{x_1, \ldots, x_n\}$. Under the assumptions of the linear model, it is easy to see that $\mathbb{E}(y_i|X) = \beta_0 + \beta_1 x_i$.
- The estimate $\hat{\beta}_1$ is unbiased for β_1.

.

Expected values and variances of the OLS estimators

- Recall the linear regression model described before. The quantities β_0 and β_1 are parameters of the model. These are estimated by $\hat{\beta}_0$ and $\hat{\beta}_1$. How good are these estimators?
- We write X for the set of explanatory variable values $\{x_1, \ldots, x_n\}$. Under the assumptions of the linear model, it is easy to see that $\mathbb{E}(y_i|X) = \beta_0 + \beta_1 x_i$.
- The estimate $\hat{\beta}_1$ is unbiased for β_1. To see this, let us first compute

$$.$$

Expected values and variances of the OLS estimators

- Recall the linear regression model described before. The quantities β_0 and β_1 are parameters of the model. These are estimated by $\hat{\beta}_0$ and $\hat{\beta}_1$. How good are these estimators?
- We write X for the set of explanatory variable values $\{x_1, \ldots, x_n\}$. Under the assumptions of the linear model, it is easy to see that $\mathbb{E}(y_i|X) = \beta_0 + \beta_1 x_i$.
- The estimate $\hat{\beta}_1$ is unbiased for β_1. To see this, let us first compute

$$\mathbb{E}(\hat{\beta}_1|X) = \mathbb{E}\left(\frac{\sum_{i=1}^n (x_i - \bar{x}) y_i}{\sum_{i=1}^n (x_i - \bar{x})^2}|X\right) =$$

$$.$$

Expected values and variances of the OLS estimators

- Recall the linear regression model described before. The quantities β_0 and β_1 are parameters of the model. These are estimated by $\hat{\beta}_0$ and $\hat{\beta}_1$. How good are these estimators?
- We write X for the set of explanatory variable values $\{x_1, \ldots, x_n\}$. Under the assumptions of the linear model, it is easy to see that $\mathbb{E}(y_i|X) = \beta_0 + \beta_1 x_i$.
- The estimate $\hat{\beta}_1$ is unbiased for β_1. To see this, let us first compute

$$\begin{aligned}
\mathbb{E}(\hat{\beta}_1|X) &= \mathbb{E}\left(\frac{\sum_{i=1}^n (x_i - \bar{x}) y_i}{\sum_{i=1}^n (x_i - \bar{x})^2} \Big| X\right) = \frac{\sum_{i=1}^n (x_i - \bar{x}) \mathbb{E}(y_i|X)}{\sum_{i=1}^n (x_i - \bar{x})^2} \\
&= \qquad\qquad\qquad .
\end{aligned}$$

Expected values and variances of the OLS estimators

- Recall the linear regression model described before. The quantities β_0 and β_1 are parameters of the model. These are estimated by $\hat{\beta}_0$ and $\hat{\beta}_1$. How good are these estimators?
- We write X for the set of explanatory variable values $\{x_1, \ldots, x_n\}$. Under the assumptions of the linear model, it is easy to see that $\mathbb{E}(y_i|X) = \beta_0 + \beta_1 x_i$.
- The estimate $\hat{\beta}_1$ is unbiased for β_1. To see this, let us first compute

$$
\begin{aligned}
\mathbb{E}(\hat{\beta}_1|X) &= \mathbb{E}\left(\frac{\sum_{i=1}^n (x_i - \bar{x}) y_i}{\sum_{i=1}^n (x_i - \bar{x})^2} \middle| X\right) = \frac{\sum_{i=1}^n (x_i - \bar{x}) \mathbb{E}(y_i|X)}{\sum_{i=1}^n (x_i - \bar{x})^2} \\
&= \frac{\sum_{i=1}^n (x_i - \bar{x})(\beta_0 + \beta_1 x_i)}{\sum_{i=1}^n (x_i - \bar{x})^2}.
\end{aligned}
$$

- I will leave it as an exercise to verify that the last expression equals β_1.

- I will leave it as an exercise to verify that the last expression equals β_1.
- This gives $\mathbb{E}(\hat{\beta}_1|X) = \beta_1$.

- I will leave it as an exercise to verify that the last expression equals β_1.
- This gives $\mathbb{E}(\hat{\beta}_1|X) = \beta_1$. This means that $\hat{\beta}_1$ is conditionally (on X) unbiased for β_1.

- I will leave it as an exercise to verify that the last expression equals β_1.
- This gives $\mathbb{E}(\hat{\beta}_1|X) = \beta_1$. This means that $\hat{\beta}_1$ is conditionally (on X) unbiased for β_1. This automatically implies unconditional unbiasedness of $\hat{\beta}_1$ as well (why?).

- I will leave it as an exercise to verify that the last expression equals β_1.
- This gives $\mathbb{E}(\hat{\beta}_1|X) = \beta_1$. This means that $\hat{\beta}_1$ is conditionally (on X) unbiased for β_1. This automatically implies unconditional unbiasedness of $\hat{\beta}_1$ as well (why?).
- Unbiasedness (conditionally on X and unconditionally) also holds for $\hat{\beta}_0$:

- I will leave it as an exercise to verify that the last expression equals β_1.
- This gives $\mathbb{E}(\hat{\beta}_1|X) = \beta_1$. This means that $\hat{\beta}_1$ is conditionally (on X) unbiased for β_1. This automatically implies unconditional unbiasedness of $\hat{\beta}_1$ as well (why?).
- Unbiasedness (conditionally on X and unconditionally) also holds for $\hat{\beta}_0$:

$$\mathbb{E}(\hat{\beta}_0|X) =$$

- I will leave it as an exercise to verify that the last expression equals β_1.
- This gives $\mathbb{E}(\hat{\beta}_1|X) = \beta_1$. This means that $\hat{\beta}_1$ is conditionally (on X) unbiased for β_1. This automatically implies unconditional unbiasedness of $\hat{\beta}_1$ as well (why?).
- Unbiasedness (conditionally on X and unconditionally) also holds for $\hat{\beta}_0$:

$$\mathbb{E}(\hat{\beta}_0|X) = \mathbb{E}(\bar{y} - \bar{x}\hat{\beta}_1|X) =$$

- I will leave it as an exercise to verify that the last expression equals β_1.
- This gives $\mathbb{E}(\hat{\beta}_1|X) = \beta_1$. This means that $\hat{\beta}_1$ is conditionally (on X) unbiased for β_1. This automatically implies unconditional unbiasedness of $\hat{\beta}_1$ as well (why?).
- Unbiasedness (conditionally on X and unconditionally) also holds for $\hat{\beta}_0$:

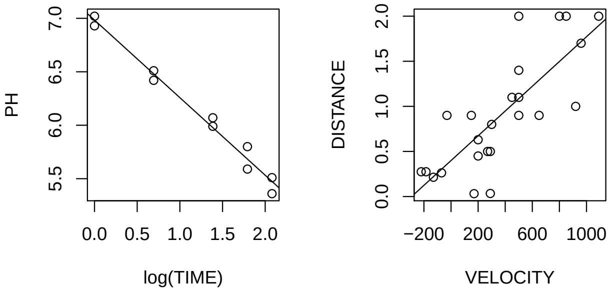
$$\mathbb{E}(\hat{\beta}_0|X) = \mathbb{E}(\bar{y} - \bar{x}\hat{\beta}_1|X) = \mathbb{E}(\bar{y}|X) - \bar{x}\beta_1.$$

- I will leave it as an exercise to verify that the last expression equals β_1.
- This gives $\mathbb{E}(\hat{\beta}_1|X) = \beta_1$. This means that $\hat{\beta}_1$ is conditionally (on X) unbiased for β_1. This automatically implies unconditional unbiasedness of $\hat{\beta}_1$ as well (why?).
- Unbiasedness (conditionally on X and unconditionally) also holds for $\hat{\beta}_0$:

$$\mathbb{E}(\hat{\beta}_0|X) = \mathbb{E}(\bar{y} - \bar{x}\hat{\beta}_1|X) = \mathbb{E}(\bar{y}|X) - \bar{x}\beta_1.$$

- Show that the last expression equals β_0.

Examples



PH
7.0
6.5
6.0
5.5
0.0
0.5
1.0
1.5
2.0
log(TIME)
DISTANCE
2.0
1.5
1.0
0.5
0.0
−200
200
600
1000
VELOCITY