Lecture 30

30.1 Joint distribution of the estimates.

In our last lecture we found the maximum likelihood estimates of the unknown parameters in simple linear regression model and we found the joint distribution of $\hat{\beta}_0$ and $\hat{\beta}_1$. Our next goal is to describe the distribution of $\hat{\sigma}^2$. We will show the following:

1. $\hat{\sigma}^2$ is independent of $\hat{\beta}_0$ and $\hat{\beta}_1$.
2. $n\hat{\sigma}^2/\sigma^2$ has χ^2_{n-2} distribution with $n-2$ degrees of freedom.

Let us consider two vectors

$$a_1 = (a_{11}, \ldots, a_{1n}) = \Big(\frac{1}{\sqrt{n}}, \ldots, \frac{1}{\sqrt{n}}\Big)$$

and

$$a_2 = (a_{21}, \ldots, a_{2n}) \text{ where } a_{2i} = \frac{X_i - \bar{X}}{\sqrt{n(\overline{X^2} - \bar{X}^2)}}.$$

It is easy to check that both vectors have length 1 and they are orthogonal to each other since their scalar product is

$$a_1 \cdot a_2 = \sum_{i=1}^n a_{1i}a_{2i} = \frac{1}{\sqrt{n}}\sum_{i=1}^n \frac{X_i - \bar{X}}{\sqrt{n(\overline{X^2} - \bar{X}^2)}} = 0.$$

Let us choose vectors $a_3, \ldots, a_n$ so that $a_1, \ldots, a_n$ is orthonormal basis and, as a result, the matrix

$$A = \begin{pmatrix} a_{11} & \cdots & a_{n1} \\ a_{12} & \cdots & a_{n2} \\ \vdots & \vdots & \vdots \\ a_{1n} & \cdots & a_{nn} \end{pmatrix}$$

is orthogonal. Let us consider vectors

$$Y = (Y_1, \ldots, Y_n), \mu = \mathbb{E}Y = (\mathbb{E}Y_1, \ldots, \mathbb{E}Y_n)$$

and

$$Y' = (Y_1', \ldots, Y_n') = \frac{Y-\mu}{\sigma} = \Big(\frac{Y_1 - \ \mathbb{E}Y_1}{\sigma}, \ldots, \frac{Y_n - \mathbb{E}Y_n}{\sigma}\Big)$$

so that the random variables $Y_1', \ldots, Y_n'$ are i.i.d. standard normal. We proved before that if we consider an orthogonal transformation of i.i.d. standard normal sequence:

$$Z' = (Z_1', \ldots, Z_n') = Y'A$$

then $Z_1', \ldots, Z_n'$ will also be i.i.d. standard normal. Since

$$Z' = Y'A = \Big(\frac{Y-\mu}{\sigma}\Big)A = \frac{YA - \mu A}{\sigma}$$

this implies that

$$YA = \sigma Z' + \mu A.$$

Let us define a vector

$$Z = (Z_1, \ldots, Z_n) = YA = \sigma Z' + \mu A.$$

Each Z_i is a linear combination of Y_is and, therefore, it has a normal distribution. Since we made a specific choice of the first two columns of the matrix A we can write down explicitely the first two coordinates Z_1 and Z_2 of vector Z. We have,

$$Z_1 = \sum a_{i1} Y_i = \frac{1}{\sqrt{n}} \sum Y_i = \sqrt{n}\bar{Y} = \sqrt{n}(\hat{\beta}_0 + \hat{\beta}_1 \bar{X})$$

and the second coordinate

$$\begin{aligned} Z_2 &= \sum a_{i2} Y_i = \sum \frac{(X_i - \bar{X})Y_i}{\sqrt{n(\overline{X^2} - \bar{X}^2)}} \\ &= \sqrt{n(\overline{X^2} - \bar{X}^2)} \sum \frac{(X_i - \bar{X})Y_i}{n(\overline{X^2} - \bar{X}^2)} = \sqrt{n(\overline{X^2} - \bar{X}^2)}\hat{\beta}_1. \end{aligned}$$

Solving these two equations for $\hat{\beta}_0$ and $\hat{\beta}_1$ we can express them in terms of Z_1 and Z_2 as

$$\hat{\beta}_1 = \frac{1}{\sqrt{n(\overline{X^2} - \bar{X}^2)}} Z_2 \text{ and } \hat{\beta}_0 = \frac{1}{\sqrt{n}} Z_1 - \frac{\bar{X}}{\sqrt{n(\overline{X^2} - \bar{X}^2)}} Z_2.$$

Next we will show how $\hat{\sigma}^2$ can also be expressed in terms of Z_is.

$$\begin{aligned}
n\hat{\sigma}^2 &= \sum_{i=1}^n (Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i)^2 = \sum_{i=1}^n \Big((Y_i - \bar{Y}) - \hat{\beta}_1(X_i - \bar{X})\Big)^2 \ \{\text{since } \hat{\beta}_0 = \bar{Y} - \hat{\beta}_1\bar{X}\} \\
&= \sum (Y_i - \bar{Y})^2 - 2\hat{\beta}_1 n(\overline{X^2} - \bar{X}^2) \underbrace{\frac{\sum (Y_i - \bar{Y})(X_i - \bar{X})}{n(\overline{X^2} - \bar{X}^2)}}_{\hat{\beta}_1} + \hat{\beta}_1^2 \sum (X_i - \bar{X})^2 \\
&= \sum (Y_i - \bar{Y})^2 - \hat{\beta}_1^2 n(\overline{X^2} - \bar{X}^2) = \sum Y_i^2 - \underbrace{n(\bar{Y})^2}_{Z_1^2} - \underbrace{\hat{\beta}_1{}^2 n(\overline{X^2} - \bar{X}^2)}_{Z_2^2} \\
&= \sum_{i=1}^n Y_i^2 - Z_1^2 - Z_2^2 = \sum_{i=1}^n Z_i^2 - Z_1^2 - Z_2^2 = Z_3^2 + \cdots + Z_n^2.
\end{aligned}$$

In the last line we used the fact that $Z = YA$ is an orthogonal transformation of Y and since orthogonal transformation preserves the length of a vector we have,

$$\sum_{i=1}^n Z_i^2 = \sum_{i=1}^n Y_i^2.$$

If we can show that $Z_1, \ldots, Z_n$ are i.i.d. with distribution $N(0, \sigma^2)$ then we will have shown that

$$\frac{n\hat{\sigma}^2}{\sigma^2} = \Big(\frac{Z_3}{\sigma}\Big)^2 + \ldots + \Big(\frac{Z_n}{\sigma}\Big)^2 \sim \chi^2_{n-2}$$

has χ^2 distribution with $n-2$ degrees of freedom, because $Z_i/\sigma \sim N(0,1)$. Since we showed above that

$$Z = \mu A + \sigma Z' \Rightarrow Z_i = (\mu A)_i + \sigma Z_i',$$

the fact that $Z_1', \ldots, Z_n'$ are i.i.d. standard normal implies that Z_is are independent of each other and $Z_i \sim N((\mu A)_i, \sigma)$. Let us compute the mean $\mathbb{E}Z_i = (\mu A)_i$:

$$\begin{aligned}
(\mu A)_i &= \mathbb{E}Z_i = \mathbb{E}\sum_{j=1}^n a_{ji} Y_j = \sum_{j=1}^n a_{ji}\mathbb{E}Y_j = \sum_{j=1}^n a_{ji}(\beta_0 + \beta_1 X_j) \\
&= \sum_{j=1}^n a_{ji}(\beta_0 + \beta_1\bar{X} + \beta_1(X_j - \bar{X})) \\
&= (\beta_0 + \beta_1\bar{X})\sum_{j=1}^n a_{ji} + \beta_1 \sum_{j=1}^n a_{ji}(X_j - \bar{X}).
\end{aligned}$$

Since the matrix A is orthogonal its columns are orthogonal to each other. Let $a_i = (a_{1i}, \ldots, a_{ni})$ be the vector in the ith column and let us consider $i \geq 3$. Then the fact that a_i is orthogonal to the first column gives

$$a_i \cdot a_1 = \sum_{j=1}^{n} a_{j1} a_{ji} = \sum_{j=1}^{n} \frac{1}{\sqrt{n}} a_{ji} = 0$$

and the fact that a_i is orthogonal to the second column gives

$$a_i \cdot a_2 = \frac{1}{\sqrt{n(\overline{X^2} - \bar{X}^2)}} \sum_{j=1}^{n} (X_j - \bar{X}) a_{ji} = 0.$$

This show that for $i \geq 3$

$$\sum_{j=1}^{n} a_{ji} = 0 \text{ and } \sum_{j=1}^{n} a_{ji}(X_j - \bar{X}) = 0$$

and this proves that $\mathbb{E} Z_i = 0$ for $i \geq 3$ and $Z_i \sim N(0, \sigma^2)$ for $i \geq 3$. As we mentioned above this also proves that $n\hat{\sigma}^2/\sigma^2 \sim \chi^2_{n-2}$.

Finally, $\hat{\sigma}^2$ is independent of $\hat{\beta}_0$ and $\hat{\beta}_1$ because as we showed above $\hat{\sigma}^2$ can be written as a function of $Z_3, \ldots, Z_n$ and $\hat{\beta}_0$ and $\hat{\beta}_1$ can be written as functions of Z_1 and Z_2.