LECTURE - 33

1. ANOTHER FORM OF $\hat{\sigma}^2$
2. UNBIASEDNESS OF $\hat{\sigma}^2$
3. DISTRIBUTION OF $\hat{\beta}_0, \hat{\beta}_1, \hat{\sigma}^2$ UNDER NORMALITY

Recall from the last lecture that our estimator for the error variance σ^2 is given by

$$\hat{\sigma}^2 = \frac{1}{n-2} \sum_{i=1}^{n} (Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i)^2$$

We will first derive an alternate expression for $\hat{\sigma}^2$ using the form of $\hat{\beta}_0$ and $\hat{\beta}_1$. Note that

$$\hat{\sigma}^2 = \frac{1}{n-2} \sum_{i=1}^{n} \left(Y_i - (\bar{Y} - \hat{\beta}_1 \bar{X}) - \hat{\beta}_1 X_i\right)^2$$

$$= \frac{1}{n-2} \sum_{i=1}^{n} \left((Y_i - \bar{Y}) - \hat{\beta}_1 (X_i - \bar{X})\right)^2$$

$$= \frac{1}{n-2} \left[\sum_{i=1}^{n} (Y_i - \bar{Y})^2 - 2\hat{\beta}_1 \sum_{i=1}^{n} (Y_i - \bar{Y})(X_i - \bar{X}) + \hat{\beta}_1^2 \sum_{i=1}^{n} (X_i - \bar{X})^2 \right]$$

$$= \frac{1}{n-2} \left[\sum_{i=1}^{n} (Y_i - \bar{Y})^2 - 2\hat{\beta}_1 \hat{\beta}_1 \sum_{i=1}^{n} (X_i - \bar{X})^2 + \hat{\beta}_1^2 \sum_{i=1}^{n} (X_i - \bar{X})^2 \right]$$

$$\left(\because \hat{\beta}_1 = \sum_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y}) \Big/ \sum_{i=1}^{n} (X_i - \bar{X})^2 \right)$$

$$= \frac{1}{n-2}\left[\sum_{i=1}^{n}(Y_i-\bar{Y})^2 - \hat{\beta}_1^2\sum_{i=1}^{n}(X_i-\bar{X})^2\right]$$

$$= \frac{1}{n-2}\left[\sum_{i=1}^{n}(Y_i-\bar{Y})^2 - \frac{\left(\sum_{i=1}^{n}(X_i-\bar{X})(Y_i-\bar{Y})\right)^2}{\sum_{i=1}^{n}(X_i-\bar{X})^2}\right]$$

We now show that $\hat{\sigma}^2$ is unbiased for σ^2. Note that

$$E[\hat{\sigma}^2] = \frac{1}{n-2}\quad E\left[\sum_{i=1}^{n}(Y_i-\bar{Y})^2\right] - \sum_{i=1}^{n}(X_i-\bar{X})^2 E[\hat{\beta}_1^2]$$

$$\text{---}\ (*)$$

Let us focus on $E[\sum_{i=1}^{n}(Y_i-\bar{Y})^2]$.

$$E\left[\sum_{i=1}^{n}(Y_i-\bar{Y})^2\right] = E\left[\sum_{i=1}^{n}Y_i^2 - n\bar{Y}^2\right]$$

$$= \sum_{i=1}^{n} E[Y_i^2] - nE[\bar{Y}^2]$$

$$= \sum_{i=1}^{n}\left(V(Y_i) + (E[Y_i])^2\right) - n\left(V(\bar{Y}) + (E[\bar{Y}])^2\right)$$

$$= \sum_{i=1}^{n}\left(\sigma^2 + (\beta_0+\beta_1 X_i)^2\right) - n\left(\frac{\sigma^2}{n} + (\beta_0+\beta_1\bar{X})^2\right)$$

$$= (n-1)\sigma^2 + \sum_{i=1}^{n}(\beta_0+\beta_1 X_i)^2 - n(\beta_0+\beta_1\bar{X})^2$$

$$\text{---}\ (**)$$

Similarly,

$$E[\hat{\beta}_1^2] = V(\hat{\beta}_1) + (E[\hat{\beta}_1])^2$$

$$= \frac{\sigma^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2} + \beta_1^2$$

Hence,

$$\sum_{i=1}^{n}(x_i - \bar{x})^2 \, E[\hat{\beta}_1^2] = \sigma^2 + \beta_1^2 \sum_{i=1}^{n}(x_i - \bar{x})^2 \qquad (***)$$

It follows by (*), (**), and (***) that

$$E[\hat{\sigma}^2] = \frac{[(n-1)\sigma^2 - \sigma^2]}{n-2} + \left\{ \sum_{i=1}^{n}(\beta_0 + \beta_1 x_i)^2 - n(\beta_0 + \beta_1 \bar{x})^2 - \beta_1^2 \sum_{i=1}^{n}(x_i - \bar{x})^2 \right\} / n-2$$

$$= \sigma^2 + \frac{1}{n-2}\left\{ \sum_{i=1}^{n}\{\beta_0^2 + 2\beta_0\beta_1 x_i + \beta_1^2 x_i^2\} - n\beta_0^2 - n\beta_1^2\bar{x}^2 - 2n\beta_0\beta_1\bar{x} - \beta_1^2\left(\sum_{i=1}^{n}x_i^2 - n\bar{x}^2\right) \right\}$$

$$= \sigma^2 + \frac{1}{n-2}\left\{ \cancel{n\beta_0^2} + \cancel{2n\beta_0\beta_1\bar{x}} + \cancel{\beta_1^2\sum_{i=1}^{n}x_i^2} - \cancel{n\beta_0^2} - \cancel{n\beta_1^2\bar{x}^2} - \cancel{2n\beta_0\beta_1\bar{x}} - \cancel{\beta_1^2\sum_{i=1}^{n}x_i^2} + \cancel{n\beta_1^2\bar{x}^2} \right\}$$

$$= \sigma^2.$$

Finally, if the errors $\varepsilon_1, \varepsilon_2, ---, \varepsilon_n$ are assumed to be independent with a common $N(0, \sigma^2)$ distribution, then the following results can be proved.

(1) $\hat{\beta}_1 \sim N\left(0, \dfrac{\sigma^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}\right)$

(2) $\hat{\beta}_0 \sim N\left(0, \dfrac{\sigma^2}{n} + \dfrac{\sigma^2 \bar{x}^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}\right)$

(3) $\hat{\sigma}^2 \sim \dfrac{1}{n-2} \chi^2_{n-2}$

CHI-SQUARE DISTRIBUTION WITH n-2 DEGREES OF FREEDOM