Go over the estimation of σ

The trick to finding the expectation of e_i^2 is writing them as a linear combination of uncorrelated variables, ϵ_i.

An unbiased estimate of $\sigma^2 = \frac{1}{n-p}\sum_{i=1}^{n} e_i^2$

Show $E\left(\frac{1}{n-p}\sum_{i=1}^{n} e_i^2\right) = \sigma^2$

1. Claim $\underset{\text{residuals}}{\underset{\uparrow}{\|e\|^2}} = \underset{\text{random errors}}{\underset{\uparrow}{\varepsilon^T}}(I-H)\varepsilon$

i) Show $e = (I-H)\varepsilon$

$$\begin{aligned}(I-H)\varepsilon &= (I-H)(Y-X\beta)\\ &= Y - HY + HX\beta - X\beta\\ &= Y - HY \qquad \text{since } HX = X\\ &= Y - \hat{Y} = e\end{aligned}$$

ii) Show $e^Te = \varepsilon^T(I-H)\varepsilon$

$$\begin{aligned}e^Te &= \varepsilon^T(I-H)^T(I-H)\varepsilon\\ &= \varepsilon^T(I-H)\varepsilon\end{aligned}$$

since $(I-H)^T = (I-H)$ and $(I-H)^2 = I-H$

2. Show $E(\varepsilon^T(I-H)\varepsilon) = \sigma^2 \text{trace}(I-H)$

Hint : $\underbrace{x^TAx}_{\text{quadratic form}} = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i x_j A_{ij}$ where $x = (x_1, x_2, \ldots, x_n)^T$

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots \\ A_{21} & & \\ \vdots & & \\ \vdots & & \end{pmatrix}_{n\times n}$$

$$E(\varepsilon^T(I-H)\varepsilon) = E\left(\sum_{i=1}^{n}\sum_{j=1}^{n}\varepsilon_i\varepsilon_j(I-H)_{ij}\right) \quad \text{by hint}$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} E(\varepsilon_i\varepsilon_j)(I-H)_{ij} \quad \text{by linearity of expectation}$$

turnover...

when $i=j$ $E(\varepsilon_i^2) = \sigma^2$ $E\varepsilon_i = 0$ $Var\,\varepsilon_i = \sigma^2 I$

$i \neq j$ $E(\varepsilon_i \varepsilon_j) = 0$ uncorrelated

$$E\left(\varepsilon^T (I-H)\varepsilon\right) = \sum_{i=1}^{n} \sigma^2 (I-H)_{ii}$$

$$= \sigma^2 \, \text{trace}(I-H)$$

③ Show $\text{trace}(I-H) = n-p$

Hint:	$\text{trace}(A+B) = \text{trace}(A) + \text{trace}(B)$
	$\text{trace}(AB) = \text{trace}(BA)$ $B_{m\times n}, A_{n\times m}$

$$\text{trace}(I-H) = \text{trace}\, I_{n\times n} - \text{trace}\, H$$

$$= n - \text{trace}\left(X(X^TX)^{-1}X^T\right)$$

$$= n - \text{trace}\left(\underset{p\times n\ n\times p}{(X^TX)^{-1}} \underset{p\times n\ n\times p}{X^T X}\right) \quad \text{by hint}$$

$$= n - \text{trace}(I_{p\times p})$$

$$= n-p$$

④ $E(\hat{\sigma}^2) = \sigma^2$ Put it all together

$$E(\hat{\sigma}^2) = E\left(\frac{1}{n-p} e^T e\right) = \frac{1}{n-p} \sigma^2 (n-p)$$

$$= \sigma^2$$

Summary

For the linear regression model

$$Y = X\beta + \epsilon$$

where $\mathsf{E}(\epsilon) = 0_{n\times 1}$, $\mathsf{Var}(\epsilon) = \sigma^2 I_n$, and the matrix $X_{n\times p}$ is fixed with rank p.

The least squares estimates are

$$\hat{\beta} = (X^T X)^{-1} X^T Y$$

Furthermore, the least squares estimates are BLUE, and

$$\mathsf{E}\left(\hat{\beta}\right) = \beta, \qquad \mathsf{Var}\left(\hat{\beta}\right) = \sigma^2 (X^T X)^{-1}$$

$$\mathsf{E}\left(\hat{\sigma}^2\right) = \mathsf{E}\left(\frac{1}{n-p}\sum_{i=1}^{n} e_i^2\right) = \sigma^2$$

We have not used any Normality assumptions to show these properties.

Normality assumption

multivariate normal, dimension n

Assume $\underset{\sim}{\epsilon} \sim N(0, \sigma^2 I)$. $\qquad \varepsilon_i \overset{iid}{\sim} N(0, \sigma^2)$

univariate Normal

Important reminders:

- Linear combinations of Normal r.v's are also Normal

  $x_{n\times 1} \sim N(\mu, \Sigma) \qquad A_{p\times n}$ fixed

  $Ax \sim N(A\mu, A\Sigma A^T)$

- Uncorrelated elements are independent

Leads to:

$$\underset{n\times 1}{Y} \sim N(X\beta, \sigma^2 I) \qquad Y = X\beta + \varepsilon$$

$$\hat{\beta} \sim N(\beta, \sigma^2 (X^T X)^{-1})$$

$$\hat{\beta} = \underbrace{(X^T X)^{-1} X^T}\, Y$$

Normality: $\hat{\beta}$ are also MLE of β $\qquad \left(\begin{array}{l} \text{MLE of } \tilde{\sigma}^2 \\ \text{is not } \hat{\sigma}^2 \end{array} \right)$

Inference on individual parameters

With the addition of the Normal assumption, it can be shown that

$$\frac{\hat{\beta}_j - \beta_j}{SE(\hat{\beta}_j)} \sim t_{n-p}$$

t-distributed n-p d.f.

leads to the usual construction of tests and confidence intervals for single parameters.