Agenda:

1. Recap distributional results of OLS under linear model w/ Gaussian errors
2. distribution of t-statistic

Recall linear model w/ Gaussian errors:

(A1): (Linearity): $\underset{n\times 1}{\boxed{y}} = \underset{n\times p}{\boxed{X}}\underset{p\times 1}{\boxed{\beta^*}} + \underset{n\times 1}{\boxed{\epsilon}}$

(A2): (Gaussianity): $\underset{n\times 1}{\boxed{\epsilon}} \mid X \sim N(\underset{n\times 1}{\boxed{0_n}}, \underset{n\times n}{\boxed{\sigma^2 I_n}})$

Under (A1) + (A2), we showed

$\underset{p\times 1}{\boxed{\hat\beta}} - \underset{p\times 1}{\boxed{\beta^*}} \mid X \sim N(\underset{p\times 1}{\boxed{0_p}}, \underset{p\times p}{\boxed{\sigma^2 (X^T X)^{-1}}})$

Consider unbiased estimator of σ^2:

$s^2 := \frac{1}{n-p}\sum_{i=1}^n (y_i - x_i^T\hat\beta)^2$

Under the classical linear model, s^2 is unbiased: $E[s^2 \mid X] = \sigma^2$

Note this implies $E[s^2] = \sigma^2$

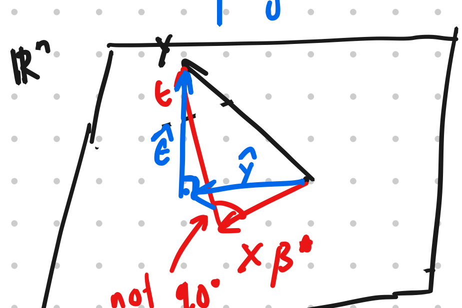
$s^2 = \frac{1}{n-p}\sum_{i=1}^n (y_i - \hat y_i)^2$ $\qquad \hat y_i \doteq x_i^T\hat\beta$

$= \frac{1}{n-p}\sum_{i=1}^n \hat\epsilon_i^2$ $\qquad \hat\epsilon_i \doteq y_i - \hat y_i$

$= \frac{1}{n-p}\|\hat\epsilon\|_2^2$ $\qquad \hat\epsilon = \begin{bmatrix}\hat\epsilon_1\\ \vdots \\ \hat\epsilon_n\end{bmatrix} \in \mathbb{R}^n$

Recall: $\hat\epsilon = (I_n - H)\epsilon$, where H is the projector onto $R(X)$.

$H \doteq X(X^TX)^{-1}X^T$

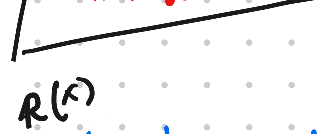


$y = X\beta^* + \epsilon$

$= \frac{1}{n-p}\|(I_n - H)\epsilon\|_2^2$

$= \frac{1}{n-p}\epsilon^T(I-H)^T(I-H)\epsilon$

$= \frac{1}{n-p} \epsilon \; (I_n ...)$



If H is a projector, then so is $I_n - H$. In particular $I_n - H$ is the projector onto the orthocomplement of $R(X)$.

$$= \frac{1}{n-p} \epsilon^T (I_n - H) \epsilon$$

If $I_n - H$ is a projector onto $\boxed{R(X)^{\perp}}$, then $I_n - H = \underset{n\times(n-p)}{\boxed{U}} U^T$, where U is an orthonormal basis of $R(X)^{\perp}$

$\hookrightarrow U^T U = I_{n-p}$

$$U = \begin{bmatrix} | & & | \\ u_1 & \dots & u_{n-p} \\ | & & | \end{bmatrix} \qquad [U^T U]_{k,\ell} = u_k^T u_\ell = \begin{cases} 0 & \text{if } k \neq \ell \\ 1 & \text{if } k = \ell \end{cases}$$

$$= \frac{1}{n-p} \epsilon^T U U^T \epsilon$$

$$= \frac{1}{n-p} (U^T \epsilon)^T (U^T \epsilon)$$

$$= \frac{1}{n-p} \| U^T \epsilon \|_2^2$$

Recall if $\underbrace{z \sim N(0, I_k)}_{\text{equiv}}$, then $\left\{ \begin{aligned} \|z\|_2^2 &= z^T z \\ &= \sum_{k=1}^{K} z_k^2 \end{aligned} \right\} \sim \chi^2_K$

$$z = \begin{bmatrix} z_1 \\ \vdots \\ z_K \end{bmatrix}, \quad z_k \overset{iid}{\sim} N(0,1)$$

Under linear model w/ Gaussian errors: $\epsilon \mid X \sim N(0_n, \sigma^2 I_n)$

$$U^T \epsilon \mid X \sim N(\underbrace{U^T \cdot 0_n}_{0_{n-p}}, \underbrace{U^T (\sigma^2 I_n) U}_{\substack{= \sigma^2 U^T U \\ = \sigma^2 I_{n-p}}})$$

Define $z \triangleq \frac{1}{\sigma} U^T \epsilon$, then $z \sim N(0_{n-p}, I_{n-p})$

so $z^T z \sim \chi^2_{n-p}$

(Recall $\sum_{k=1}^{K} n_k^2 \sim \chi^2_K$, where

Recall from above $s^2 = \frac{1}{n-p} \| U^T \epsilon \|_2^2$

$$\frac{s^2}{\sigma^2} = \frac{1}{n-p} \left\| \frac{U^T \epsilon}{\sigma} \right\|_2^2 \sim \frac{1}{n-p} \chi^2_{n-p}$$

$$E\left[\frac{s^2}{\sigma^2}\right] = \frac{1}{n-p} \boxed{E[\chi^2_{n-p}]} = n-p$$

$$= 1$$

This implies $E[s^2] = \sigma^2$

$n_k \sim N(0,1)$

$$E\left[\sum_{k=1}^{K} n_k^2\right] = \sum_{k=1}^{K} E[n_k^2]$$

$$= \sum_{k=1}^{K} 1$$

$$= K$$

---

distribution of t-statistic:

Fact: If if $N \sim N(0,1)$ and $V \sim \chi^2_K$ (and N and V are ind),

the $\frac{N}{\sqrt{V/K}} \sim t_K$

Recall $\hat{\beta}_j - \beta_j^* \mid X \sim N(0, \overbrace{\sigma^2 [(X^T X)^{-1}]_{j,j}}^{SE[\hat{\beta}_j]^2})$ under linear model w/ Gaussian error terms

Under $H_0: \beta_j^* = 0$, we have $\hat{\beta}_j \mid X \sim N(0, \sigma^2 [(X^T X)^{-1}]_{j,j})$

t-statistic: $\frac{\hat{\beta}_j}{\widehat{SE}[\hat{\beta}_j]}$, when $\widehat{SE}[\hat{\beta}_j]^2 = s^2 [(X^T X)^{-1}]_{j,j}$

$$s^2 = \frac{1}{n-p} \sum_{i=1}^{n} (y_i - x_i^T \hat{\beta})^2$$

Recall:

z-statistic: $\frac{\hat{\beta}_j}{SE[\hat{\beta}_j]} \sim N(0,1)$

$$\underbrace{\boxed{\frac{\hat{\beta}_j}{\widehat{SE}[\hat{\beta}_j]}}}_{\text{t-statistic}} = \frac{\hat{\beta}_j}{SE[\hat{\beta}_j]} \cdot \frac{SE[\hat{\beta}_j]}{\widehat{SE}[\hat{\beta}_j]} = \frac{\left(\sigma^2 [(X^T X)^{-1}]_{j,j}\right)^{\frac{1}{2}}}{\left(s^2 [(X^T X)^{-1}]_{j,j}\right)^{\frac{1}{2}}} = \frac{\sigma}{s}$$

$$= \frac{\left[\frac{\hat{\beta}}{SE[\hat{\beta}_j]}\right]}{s/\sigma}$$ z-statistic

Recall from above: $\frac{s^2}{\sigma^2} \sim \frac{\chi^2_{n-p}}{n-p}$

Recall from ... $n-p$

This implies $\frac{\hat{\beta}}{\hat{SE}[\hat{\beta}_j]} \sim t_{n-p}$