Agenda:

1. Testing linear restrictions on regression coefficients
2. Testing non-linear restrictions
3. Optimal linear prediction

Testing: $H_0: C\beta_* = b$ for some $C \in \mathbb{R}^{r\times p}$ & $b \in \mathbb{R}^r$

Ex: $(C = e_j^T = [0 \dots 0, 1, 0 \dots 0]$, $b = 0)$, then

$$\begin{array}{c} \uparrow \\ j\text{-th position} \end{array}$$

$$H_0: e_j^T \beta_* = 0 \iff [\beta_*]_j = 0$$

Ex: $(C = [1, -1, 0 \dots 0], \; b = 0)$

$$H_0: [1, -1, \dots 0 \dots]\, \beta_* = 0 \iff [\beta_*]_1 - [\beta_*]_2 = 0$$

Ex: $H_0: \beta_1^* = \beta_2^* = \beta_3^*$

$$C = \begin{bmatrix} 1 & -1 & \dots & 0 & \dots \\ 1 & 0 & -1 & \dots & 0 \dots \end{bmatrix} \quad b = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$C\beta_* = b \iff \begin{array}{l} \beta_1^* - \beta_2^* = 0 \\ \beta_1^* - \beta_3^* = 0 \end{array}$$

$$C = \begin{bmatrix} 1 & -1 & \dots & 0 & \dots \\ 0 & 1 & -1 & \dots 0 & \dots \end{bmatrix} \quad b = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\beta_1^* - \beta_2^* = 0$$
$$\beta_2^* - \beta_3^* = 0$$

Test $H_0: \underset{r\times p}{C}\,\underset{p\times 1}{\beta_*} - \underset{r\times 1}{b} = 0$

Recall: $\sqrt{n}(\hat\beta_n - \beta_*) \xrightarrow{d} Z$, $Z \sim N(0, \mathrm{Avar}[\hat\beta_n])$

Claim: $\underbrace{\sqrt{n}(C\hat\beta_n - b)} \xrightarrow{d} CZ \sim N(0, C\,\mathrm{Avar}[\hat\beta_n]\,C^T)$

$$\sqrt{n}(C\hat\beta_n - C\beta_*)$$
$$C \cdot \sqrt{n}(\hat\beta_n - \beta_*)$$

Recall delta method: $\sqrt{n}(Z_n - c) \xrightarrow{d} Z$

$$\sqrt{n}(f(Z_n) - f(c)) \xrightarrow{d} DF(c)\, Z$$

Intuition: pretend z_n, c, z are real numbers & f is a real-valued function

$$\sqrt{n}(f(z_n) - f(c)) \approx \sqrt{n}(f'(c)(z_n - c))$$

$$= f'(c) \cdot \underbrace{\sqrt{n}(z_n - c)}_{\Downarrow d \; Z}$$

$$\xrightarrow{d} f'(c)\, Z$$

Apply delta method w/ the function $f(\beta) = C\beta$

$\sqrt{n}(C\hat{\beta}_n - b) \xrightarrow{d} CZ$ $\qquad Df(\beta) = C$

Summary: $\sqrt{n}(C\hat{\beta}_n - b) \xrightarrow{d} N(0, C\,\mathrm{Avar}[\hat{\beta}_n]C^T)$

$$\sqrt{n}\,(C\,\mathrm{Avar}[\hat{\beta}_n]C^T)^{-\frac{1}{2}}(C\hat{\beta}_n - b) \xrightarrow{d} Z, \text{ where } Z \sim N(0, I_r)$$

$$n\,\|(C\,\mathrm{Avar}[\hat{\beta}_n]C^T)^{-\frac{1}{2}}(C\hat{\beta}_n - b)\|_2^2 \xrightarrow{d} \|Z\|_2^2 = \sum_{k=1}^r Z_k^2$$

Recall if $Z \sim N(0, I_r)$, $\|Z\|_2^2 = \sum_{k=1}^r Z_k^2 \sim \chi_r^2$

Wald-test statistic: $\underbrace{n\,\|(C\,\widehat{\mathrm{Avar}}[\hat{\beta}_n]C^T)^{-\frac{1}{2}}(C\hat{\beta}_n - b)\|_2^2}_{W_n}$

Claim: $W_n \xrightarrow{d} \chi_r^2$

Step 1: $\sqrt{n}\,(C\,\widehat{\mathrm{Avar}}\,C^T)^{-\frac{1}{2}}(C\hat{\beta}_n - b) \xrightarrow{d} N(0, I_r)$

$$\underbrace{(C\,\widehat{\mathrm{Avar}}\,C^T)^{-\frac{1}{2}}}_{\downarrow p \atop (C\,\mathrm{Avar}\,C^T)^{-\frac{1}{2}}}(C\,\mathrm{Avar}\,C^T)^{\frac{1}{2}}\underbrace{\sqrt{n}(C\,\mathrm{Avar}\,C^T)^{-\frac{1}{2}}(C\hat{\beta}_n - b)}_{\downarrow d \atop N(0, I_r)}$$

Use Slutsky's to combine limits:

Step 2: CMT to conclude

$$\|\sqrt{n}\,(C\,\widehat{\mathrm{Avar}}\,C^T)^{-\frac{1}{2}}(C\hat{\beta}_n - b)\|_2^2 \xrightarrow{d} \chi_r^2$$

Wald-test: reject if $W_n \geq \tau$, where $\tau = 1-\alpha$ quantile of χ_r^2 r.v (of level α)

Testing $H: c(\beta^*) = 0$ where $c(\beta)$ is a smooth function $c: \mathbb{R}^p \to \mathbb{R}^r$

Ex: $c(\beta) = \|\beta\|_2^2 = \sum_{j=1}^p |\beta_j^*|^2$

Wald-test: reject if $V_n = \|\sqrt{n}\,(J\,\widehat{\mathrm{Avar}}\,J^T)^{-\frac{1}{2}}\,c(\hat{\beta}_n)\|_2^2 \geq \Phi^{-1}(1-\alpha)$, where Φ is the CDF of a χ_r^2 r.v.

Here $J \triangleq Dc(\beta^*)$

Special case of linear $c(\beta) = C\beta - b$

$$J = Dc(\beta^*) = C$$

Claim: $\|\sqrt{n}\,(J\,\widehat{\mathrm{Avar}}\,J^T)^{-\frac{1}{2}}c(\hat{\beta}_n)\|_2^2 \xrightarrow{d} \chi_r^2$

$\underline{\text{step 1}}$: $\sqrt{n}(J\,Avar\,J^T)^{-\frac{1}{2}} c(\beta_n) \xrightarrow{d} N(0, I_r)$

Recall $\sqrt{n}(\hat{\beta}_n - \beta_*) \xrightarrow{d} N(0, Avar)$

delta method: $\sqrt{n}(c(\hat{\beta}_n) - \cancel{c(\beta_*)}^{\,0}) \xrightarrow{d} \underbrace{JZ}_{\text{has dist}}$, when $Z \sim N(0, Avar)$

$N(0, J\,Avar\,J^T)$

$$\sqrt{n}\, c(\hat{\beta}_n) \xrightarrow{d} N(0, J\,Avar\,J^T)$$

$$\sqrt{n}\,(J\,Avar\,J^T)^{-\frac{1}{2}} c(\hat{\beta}_n) \xrightarrow{d} N(0, I_r)$$

Step 2: $\sqrt{n}\,(J\,\widehat{Avar}\,J^T)^{-\frac{1}{2}} c(\hat{\beta}_n) \xrightarrow{d} N(0, I_r)$

Step 3: $\| \sqrt{n}\,(J\,\widehat{Avar}\,J^T)^{-\frac{1}{2}} c(\hat{\beta}_n) \|_2^2 \xrightarrow{d} \chi^2_r$

---

what happens if linearity is violated:

$$\hat{\beta}_n \triangleq \operatorname{argmin}_{\beta \in \mathbb{R}^p} \frac{1}{n}\sum_{i=1}^n \frac{1}{2}(y_i - x_i^T\beta)^2$$

$$= \operatorname{argmin}_{\beta \in \mathbb{R}^p} \frac{1}{n}\sum_{i=1}^n \frac{1}{2}y_i^2 - y_i x_i^T\beta + \frac{1}{2}(x_i^T\beta)^2$$

$$= \operatorname{argmin}_{\beta \in \mathbb{R}^p} \cancel{\frac{1}{n}\sum_{i=1}^n \frac{1}{2}y_i^2} - \frac{1}{n}\sum_{i=1}^n y_i x_i^T\beta + \frac{1}{n}\sum_{i=1}^n \frac{1}{2}(\beta^T x_i)(x_i^T\beta)$$

$$= \operatorname{argmin}_{\beta \in \mathbb{R}^p} \frac{1}{2}\beta^T\left(\frac{1}{n}\sum_{i=1}^n x_i x_i^T\right)\beta - \left(\frac{1}{n}\sum_{i=1}^n x_i y_i\right)^T\beta$$

suggest $\hat{\beta}_n$ is close to

$$\boxed{\beta_*} = \operatorname{argmin}_{\beta \in \mathbb{R}^p} \frac{1}{2}\beta^T E[x_1 x_1^T]\beta - E[x_1 y_1]^T\beta$$

$$= \operatorname{argmin}_{\beta \in \mathbb{R}^p} \underbrace{E\left[\frac{1}{2}(y_1 - x_1^T\beta)^2\right]}_{\shortparallel}$$

$$= E\left[\frac{1}{2}y_1^2 - y_1 x_1^T\beta + \frac{1}{2}(x_1^T\beta)^2\right]$$

$$= E\left[\frac{1}{2}y_1^2\right] - E[x_1 y_1]^T\beta + \frac{1}{2}\beta^T E[x_1 x_1^T]\beta$$