8 Hypothesis testing and Confidence Intervals

8.1 Hypothesis testing: General Hypothesis

$\beta = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} \begin{matrix} p+1 \\ k+1 \end{matrix}$

$y = X\beta + \epsilon$, $\epsilon \sim N(0, \sigma^2 I)$

$\dim(\beta_2) = k-p$

special case

Let $\mathbf{y} \sim N(\mathbf{X}\boldsymbol{\beta}, \boldsymbol{\sigma}^2\mathbf{I})$, Then

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}, \qquad \hat{\boldsymbol{\beta}} \sim N(\boldsymbol{\beta}, (\mathbf{X}'\mathbf{X})^{-1}\boldsymbol{\sigma}^2)$$

.

$q * (k+1)$

We are now interested in testing the following hypothesis:

$H_0: \beta_2 = 0$

$$H_0 : \mathbf{K}'\boldsymbol{\beta} = \boldsymbol{\mu} \tag{8.1}$$

$\Longleftrightarrow K'\beta = 0$

where $\mathbf{K}'$ is $q \times (k+1)$, and $\mathbf{K}'$ is assumed to be full row rank.

$k+1$

$\operatorname{rank}(K) = q$

$\mu = 0$

$K' = (0 \;\; I_{k-p})$

$(k-p) * (k+1)$

Note that

1. $$\mathbf{K}'\hat{\boldsymbol{\beta}} - \boldsymbol{\mu} \sim N[\mathbf{K}'\boldsymbol{\beta} - \boldsymbol{\mu}, \mathbf{K}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{K}\boldsymbol{\sigma}^2]$$

$\hat\beta = (X'X)^{-1}X'y$

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

2. $(\mathbf{K}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{K})^{-1}$ is symmetric.

3. Let
$$Q = (\mathbf{K}'\hat{\boldsymbol{\beta}} - \boldsymbol{\mu})'[\mathbf{K}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{K}]^{-1}(\mathbf{K}'\hat{\boldsymbol{\beta}} - \boldsymbol{\mu})$$
(Q is sometimes denoted by SSH) then
$$\frac{Q}{\sigma^2} \sim \chi^2_{\{q, \frac{1}{2\sigma^2}(\mathbf{K}'\boldsymbol{\beta}-\boldsymbol{\mu})'[\mathbf{K}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{K}]^{-1}(\mathbf{K}'\boldsymbol{\beta}-\boldsymbol{\mu})\}}$$

$\hat\sigma^2 = \frac{SSE}{n-k-1}$

4. Q and SSE are independent

SSE $= (y-\hat y)'(y-\hat y) = y'(I-H)y$

$H = X(X'X)^{-1}X'$

SSE and $\hat\beta$ are indepent!

$\Rightarrow$ SSE and Q are indepent!

The test statistics

$$\begin{aligned} F(H) &= \frac{Q/q}{SSE/[N-r(\mathbf{X})]} \\ &= \frac{Q}{q\hat{\sigma}^2} \\ &\sim F_{[q, N-r(\mathbf{X}), \frac{1}{2\sigma^2}(\mathbf{K}'\boldsymbol{\beta}-\boldsymbol{\mu})'[\mathbf{K}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{K}]^{-1}(\mathbf{K}'\boldsymbol{\beta}-\boldsymbol{\mu})]} \end{aligned}$$

$\hat{\beta}$

λ

Here $\hat{\sigma}^2 = SSE/(N - r(\mathbf{X}))$ which is the unbiased estimator of σ^2.

Under $H_0 : \mathbf{K}'\mathbf{a} = \boldsymbol{\mu}$, we have

$$F(\mathbf{H}) \sim F_{\{q, N-r(\mathbf{X})\}}.$$

$n-k-1$

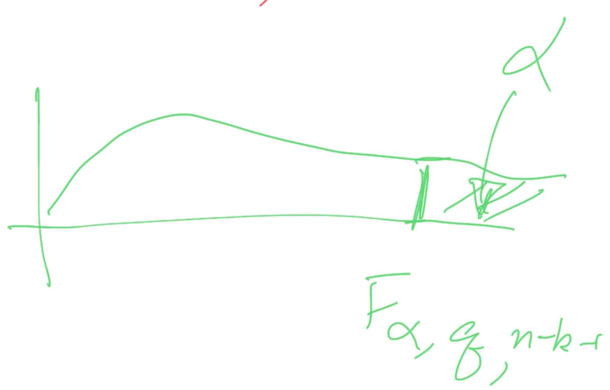
Under H_0: $\lambda = 0$

$$F(H) = \frac{Q/q}{SSE/(n-k-1)} \overset{H_0}{\sim} F_{q,n-k-1}$$

Reject H_0 if $F(H) \geq F_{\alpha,q,n-k-1}$ (one-sided)

$$\text{P-value} = P(F_{q,n-k-1} \geq F(H))$$

$q=1$



Estimation under constraints!

$\underset{\sim}{Y} = \underset{\sim}{X}\underset{\sim}{a} + \underset{\sim}{\varepsilon}, \quad \varepsilon \sim N(0, \sigma^2 I)$

$\underset{\sim}{K}'\underset{\sim}{a} = \mu$

Note Estimation of $\mathbf{a}$ under the null hypothesis H_0 : $\mathbf{K}'\mathbf{a} = \boldsymbol{\mu}$ or under the constraint. Denote the LS estimator of $\mathbf{a}$ by $\tilde{\mathbf{a}}$. To obtain the least squares estimator of $\mathbf{a}$, need to minimize

$$\min_{\underset{\sim}{a}} \quad (\mathbf{y} - \mathbf{Xa})'(\mathbf{y} - \mathbf{Xa}) + 2\theta'(\mathbf{K}'\mathbf{a} - \boldsymbol{\mu})$$

with respect to $\mathbf{a}$ and θ. Note that 2θ is a vector of Lagrange multipliers. After the minimization,

$$\begin{aligned} \tilde{\mathbf{a}} &= (\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{y} - \mathbf{K}(\mathbf{K}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{K})^{-1}(\mathbf{K}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} - \boldsymbol{\mu})) \\ &= \hat{\boldsymbol{\beta}} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{K}(\mathbf{K}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{K})^{-1}(\mathbf{K}'\hat{\boldsymbol{\beta}} - \boldsymbol{\mu}) \end{aligned} \tag{8.2}$$

- $\hat{\boldsymbol{\beta}} - \tilde{\mathbf{a}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{K}(\mathbf{K}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{K})^{-1}(\mathbf{K}'\hat{\boldsymbol{\beta}} - \boldsymbol{\mu})$;
- $\tilde{\mathbf{a}}$ is the BLUE.

$\underset{\sim}{\hat{\beta}} = (\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}'\underset{\sim}{Y}$

Let

$$\mathbf{L} = (\mathbf{y} - \mathbf{Xa})'(\mathbf{y} - \mathbf{Xa}).$$

$\forall \mathbf{a}_0$ with constraint $\mathbf{K}'\mathbf{a}_0 = \boldsymbol{\mu}$

$$\begin{aligned} &\mathbf{L}_0 \\ &= (\mathbf{y} - \mathbf{Xa}_0)'(\mathbf{y} - \mathbf{Xa}_0) \\ &= (\mathbf{y} - \mathbf{X}\tilde{\mathbf{a}} + \mathbf{X}\tilde{\mathbf{a}} - \mathbf{Xa}_0)'(\mathbf{y} - \mathbf{X}\tilde{\mathbf{a}} + \mathbf{X}\tilde{\mathbf{a}} - \mathbf{Xa}_0) \\ &= (\mathbf{y} - \mathbf{X}\tilde{\mathbf{a}})'(\mathbf{y} - \mathbf{X}\tilde{\mathbf{a}}) + (\mathbf{X}\tilde{\mathbf{a}} - \mathbf{Xa}_0)'(\mathbf{X}\tilde{\mathbf{a}} - \mathbf{Xa}_0) + 2(\mathbf{y} - \mathbf{X}\tilde{\mathbf{a}})'(\mathbf{X}\tilde{\mathbf{a}} - \mathbf{Xa}_0) \end{aligned}$$

But

$$\begin{aligned} &(\mathbf{y} - \mathbf{X}\tilde{\mathbf{a}})'(\mathbf{X}\tilde{\mathbf{a}} - \mathbf{Xa}_0) \\ &= (\mathbf{y}'\mathbf{X} - \tilde{\mathbf{a}}'\mathbf{X}'\mathbf{X})(\tilde{\mathbf{a}} - \mathbf{a}_0) \\ &= [\mathbf{y}'\mathbf{X} - \hat{\boldsymbol{\beta}}'(\mathbf{X}'\mathbf{X}) + (\mathbf{K}'\hat{\boldsymbol{\beta}} - \boldsymbol{\mu})'(\mathbf{K}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{K})^{-1}\mathbf{K}'(\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{X})](\tilde{\mathbf{a}} - \mathbf{a}_0) \\ &= (\mathbf{K}'\hat{\boldsymbol{\beta}} - \boldsymbol{\mu})'(\mathbf{K}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{K})^{-1}\mathbf{K}'(\tilde{\mathbf{a}} - \mathbf{a}_0) \\ &= 0 \end{aligned}$$

Because $(\mathbf{K}'\tilde{\mathbf{a}} = \mathbf{K}'\mathbf{a}_0 = \boldsymbol{\mu})$, therefore,

$$\mathbf{L}_0 = (\mathbf{y} - \mathbf{X}\tilde{\mathbf{a}})'(\mathbf{y} - \mathbf{X}\tilde{\mathbf{a}}) + (\tilde{\mathbf{a}} - \mathbf{a}_0)'\mathbf{X}'\mathbf{X}(\tilde{\mathbf{a}} - \mathbf{a}_0)$$

Hence,

$$\mathbf{a}_0 = \tilde{\mathbf{a}} \quad \text{minimize} \quad \mathbf{L}_0$$

Without the null hypothesis,

$$SSE = (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})$$

Under the null hypothesis, the sum of squares of residual is (Reduced Model)

$$\begin{aligned}
SSE_{H_0} &= (\mathbf{y} - \mathbf{X}\tilde{\mathbf{a}})'(\mathbf{y} - \mathbf{X}\tilde{\mathbf{a}}) \\
&= [\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{X}\hat{\boldsymbol{\beta}} - \mathbf{X}\tilde{\mathbf{a}}]'[\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{X}\hat{\boldsymbol{\beta}} - \mathbf{X}\tilde{\mathbf{a}}] \\
&= [\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{X}(\hat{\boldsymbol{\beta}} - \tilde{\mathbf{a}})]'[\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{X}(\hat{\boldsymbol{\beta}} - \tilde{\mathbf{a}})] \\
&= (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}) + (\hat{\boldsymbol{\beta}} - \tilde{\mathbf{a}})'\mathbf{X}'\mathbf{X}(\hat{\boldsymbol{\beta}} - \tilde{\mathbf{a}})
\end{aligned}$$

Since $((\hat{\boldsymbol{\beta}} - \tilde{\mathbf{a}})'\mathbf{X}'(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}) = 0)$

From (8.2),

$$\begin{aligned}
SSE_{H_0} &= SSE + (\mathbf{K}'\hat{\boldsymbol{\beta}} - \boldsymbol{\mu})'[\mathbf{K}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{K}]^{-1}\mathbf{K}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \\
&\quad \mathbf{K}[\mathbf{K}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{K}]^{-1}(\mathbf{K}\hat{\boldsymbol{\beta}} - \boldsymbol{\mu}) \\
&= SSE + (\mathbf{K}'\hat{\boldsymbol{\beta}} - \boldsymbol{\mu})'(\mathbf{K}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{K})^{-1}(\mathbf{K}'\hat{\boldsymbol{\beta}} - \boldsymbol{\mu}) \\
&= SSE + Q \\
&\geq SSE
\end{aligned}$$

Special cases

1. $H_0 : \boldsymbol{\beta} = \boldsymbol{\beta_0} \quad \Rightarrow \quad \mathbf{K}' = \mathbf{I}, q = k+1, \boldsymbol{\mu} = \boldsymbol{\beta_0}$

   (a)

   $$Q \;=\; (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)'\boldsymbol{X}'\boldsymbol{X}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)$$

   (b) $F(H) = \frac{Q}{(k+1)\hat{\sigma}^2}$

   (c) Under the null hypothesis,

   $$F(H) \;\sim\; F_{\{k+1, n-(k+1)\}}$$

   (d) $\tilde{\mathbf{a}} = \hat{\boldsymbol{\beta}} - (\hat{\boldsymbol{\beta}} - \mathbf{a}_0) = \boldsymbol{\beta_0}$ (Under constraints, $\tilde{a} = \beta_0$)

2. $H_0 : \boldsymbol{\lambda}'\boldsymbol{\beta} = m \quad \Rightarrow \quad \mathbf{K}' = \boldsymbol{\lambda}', q = 1, \boldsymbol{\mu} = m$

(a)

$$\begin{aligned} Q &= (\boldsymbol{\lambda}'\hat{\boldsymbol{\beta}} - m)'[\boldsymbol{\lambda}'(\mathbf{X}'\mathbf{X})^{-1}\boldsymbol{\lambda}]^{-1}(\boldsymbol{\lambda}'\hat{\boldsymbol{\beta}} - m) \\ &= (\boldsymbol{\lambda}'\hat{\boldsymbol{\beta}} - m)^2/\boldsymbol{\lambda}'(\mathbf{X}'\mathbf{X})^{-1}\boldsymbol{\lambda} \end{aligned}$$

(b) $F(H) = \frac{Q}{\hat{\sigma}^2}$

(c) Under the null hypothesis,

$$F(H) \sim F_{(1,n-r(\mathbf{X}))}$$

Note: $\sqrt{F(H)} = \frac{\sqrt{Q}}{\hat{\sigma}} \sim t_{n-r(\mathbf{X})}$

(d)

$$\begin{aligned} \tilde{\mathbf{a}} &= \hat{\boldsymbol{\beta}} - (\mathbf{X}'\mathbf{X})^{-1}\boldsymbol{\lambda}[\boldsymbol{\lambda}'(\mathbf{X}'\mathbf{X})^{-1}\boldsymbol{\lambda}]^{-1}(\boldsymbol{\lambda}'\hat{\boldsymbol{\beta}} - \boldsymbol{\mu}) \\ &= \hat{\boldsymbol{\beta}} - \frac{(\boldsymbol{\lambda}'\hat{\boldsymbol{\beta}} - \boldsymbol{\mu})}{\boldsymbol{\lambda}'(\mathbf{X}'\mathbf{X})^{-1}\boldsymbol{\lambda}}(\mathbf{X}'\mathbf{X})^{-1}\boldsymbol{\lambda} \end{aligned}$$

Note: $\boldsymbol{\lambda}'\hat{\boldsymbol{\beta}} - \boldsymbol{\mu} \sim N(\boldsymbol{\lambda}'\boldsymbol{\beta} - \boldsymbol{\mu}, \boldsymbol{\lambda}'(\mathbf{X}'\mathbf{X})^{-1}\boldsymbol{\lambda}\sigma^2)$

special case $H_0 : \beta_j = 0$ $\qquad \underset{\sim}{\lambda} = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \rightarrow \beta_j$

$Q = \dfrac{\hat{\beta}_j^2}{g_{jj}}$

j-th diagonal element of $(X'X)^{-1}$ $\longrightarrow \underset{\sim}{\lambda}'(\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{\lambda} = g_{jj}$

$\Rightarrow \dfrac{\hat{\beta}_j}{g_{jj}^{1/2}\,\hat{\sigma}} \sim t_{n-k-1}$

$k+1-h$

Remark 8.1

3. $H_0 : \boldsymbol{\beta_2} = \mathbf{0}$ where $\boldsymbol{\beta} = \begin{bmatrix} \boldsymbol{\beta_1} \\ \boldsymbol{\beta_2} \end{bmatrix}$

$k+1$ h-dimensional

$$\underset{h \times (k+1)}{\underset{\sim}{K}'} = (\underset{(k+1-h)}{\underset{\sim}{0}} \;\vdots\; \underset{\sim}{I}_h)$$

$$H_0 : \underset{\sim}{K}' \underset{\sim}{\beta} = 0$$

Likelihood Ratio Test

$\underset{\sim}{Y} = \underset{\sim}{X}\,\underset{\sim}{a} + \underset{\sim}{\varepsilon}\,,\ \underset{\sim}{\varepsilon} \sim N(0, \sigma^2 I)$

Theorem: If $\mathbf{y}$ is $N_n(\mathbf{xa}, \sigma^2\mathbf{I})$, where (rank of $\mathbf{x}$ is $k+1$), the likelihood ratio for H_0: $\mathbf{a} = \mathbf{0}$ can be based on

$$F = \frac{\hat{\mathbf{a}}'\mathbf{x}'\mathbf{y}/(k+1)}{(\mathbf{y}'\mathbf{y} - \hat{\mathbf{a}}'\mathbf{x}'\mathbf{y})/(n-k-1)}.$$

H_0 is rejected if $F > F_{\alpha,k+1,n-k-1}$.

$$LR = \frac{\max_{H_0} L(a, \sigma^2)}{\max L(a, \sigma^2)}$$

8.2 Confidence intervals and prediction intervals

OUTLINE

1. Confidence region for $\boldsymbol{\beta}$
2. Confidence interval for β_j
3. Confidence interval for $\boldsymbol{\lambda}'\boldsymbol{\beta}$
4. Confidence interval for $E(y^*)$ given $\boldsymbol{x} = \boldsymbol{x}^*$
5. Prediction interval for a future observation
6. Confidence interval for σ^2
7. Simultaneous intervals
   (a) Familywise confidence level
   (b) Bonferroni procedure
   (c) Scheffé procedure

$y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_k x_{ik} + \varepsilon_i, \quad \varepsilon_i \sim N(0, \sigma^2)$

$\text{rank}(X) = k+1$

Confidence region for $\boldsymbol{\beta}$

Since $P[\frac{(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})'\boldsymbol{X}'\boldsymbol{X}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})}{(k+1)\hat{\sigma}^2} \leq F_{\alpha,k+1,n-k-1}] = 1-\alpha$

a $100(1-\alpha)\%$ joint confidence region for $\beta_0, \beta_1, ..., \beta_k$ is defined to consist of all vectors $\boldsymbol{\beta}$ that satisfy

dim(β) = ?

$$(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})'\boldsymbol{X}'\boldsymbol{X}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}) \leq (k+1)\hat{\sigma}^2 F_{\alpha,k+1,n-k-1} \longrightarrow \text{ellipse}$$

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

$$(\hat{\beta}-\beta)' \frac{(X'X)}{\sigma^2} (\hat{\beta}-\beta) \sim \chi^2_{k+1}$$

$$\frac{(\hat{\beta}-\beta)'(X'X)(\hat{\beta}-\beta)}{(k+1)\hat{\sigma}^2} \sim F(k+1, n-k-1)$$

$\hat{\sigma}^2 = \frac{SSE}{n-k-1}$ SSE and $\hat{\beta}$ are indepedent

$$P\left(\frac{(\hat{\beta}-\beta)'(X'X)(\hat{\beta}-\beta)}{(k+1)\hat{\sigma}^2} \leq F_{\alpha,k+1,n-k-1}\right) \leq 1-\alpha$$

Confidence interval for β_j

Since $P[-t_{\alpha/2,n-k-1} \leq \frac{\hat{\beta}_j - \beta_j}{S\{\hat{\beta}_j\}} \leq t_{\alpha/2,n-k-1}] = 1 - \alpha$

hence, a $100(1-\alpha)\%$ confidence interval for β_j is

$s^2(\hat{\beta}_j) = g_{jj}\hat{\sigma}^2$

$$\hat{\beta}_j \pm t_{\alpha/2,n-k-1} S\{\hat{\beta}_j\}$$

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

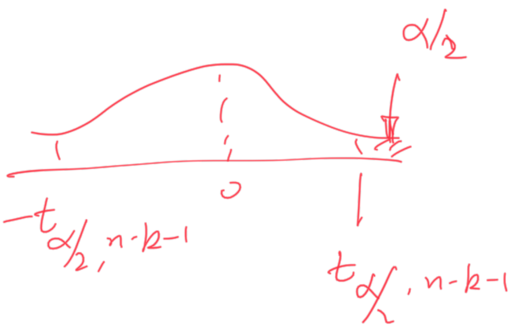
$\underbrace{\hat{\beta}_j \sim N(\beta_j, g_{jj}\sigma^2)}$ $*$

g_{jj} — j th diagonal element of $(\underset{\sim}{X}'\underset{\sim}{X})^{-1}$

$$\frac{\hat{\beta}_j - \beta_j}{(g_{jj}\sigma^2)^{1/2}} \sim N(0,1)$$

$$\hat{\sigma}^2 = \frac{SSE}{n-k-1}$$

$$\frac{\hat{\beta}_j - \beta_j}{g_{jj}^{1/2}\hat{\sigma}} \sim t_{n-k-1}$$



C.I.

$$P\left(\left| \frac{\hat{\beta}_j - \beta_j}{g_{jj}^{1/2}\hat{\sigma}} \right| \leq t_{\alpha/2,n-k-1} \right) = 1 - \alpha$$

Confidence interval for $\boldsymbol{\lambda}'\boldsymbol{\beta}$

Note that

$$t = \frac{\boldsymbol{\lambda}'\hat{\boldsymbol{\beta}} - \boldsymbol{\lambda}'\boldsymbol{\beta}}{\hat{\sigma}\sqrt{\boldsymbol{\lambda}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{\lambda}}} \sim t_{n-k-1},$$

hence, a $100(1-\alpha)\%$ confidence interval for $\boldsymbol{\lambda}'\boldsymbol{\beta}$ is

$$\boldsymbol{\lambda}'\hat{\boldsymbol{\beta}} \pm t_{\alpha/2,n-k-1}\hat{\sigma}\sqrt{\boldsymbol{\lambda}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{\lambda}}$$

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

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

$y^* = \underset{\sim}{x}^* \underset{\sim}{\beta} + \underset{\sim}{\varepsilon}^*$

$E(y^*) = \underset{\sim}{x}^* \underset{\sim}{\beta}$

$\underset{\sim}{\lambda} = \underset{\sim}{x}^*$

Confidence interval for $E(y^*)$ given $\boldsymbol{x} = \boldsymbol{x}^*$

Given that $\boldsymbol{x} = \boldsymbol{x}^*$

$$E(y^*) = \boldsymbol{x}^{*'}\boldsymbol{\beta}$$

$$\widehat{E(y^*)} = \boldsymbol{x}^{*'}\hat{\boldsymbol{\beta}}$$

$$Var(E(y^*) - \widehat{E(y^*)}) = [\boldsymbol{x}^{*'}(\boldsymbol{X'X})^{-1}\boldsymbol{x}^*]\sigma^2$$

hence, a $100(1-\alpha)\%$ confidence interval for $E(y^*)$ is

$$\boldsymbol{x}^{*'}\hat{\boldsymbol{\beta}} \pm t_{\alpha/2,n-k-1}\hat{\sigma}\sqrt{\boldsymbol{x}^{*'}(\boldsymbol{X'X})^{-1}\boldsymbol{x}^*}$$

Prediction interval for a future observation

PI of

Given that $\boldsymbol{x} = \boldsymbol{x}^*$. Let $y^* = \boldsymbol{x}^{'}\boldsymbol{\beta} + \epsilon_0$ be a future value of y when $\boldsymbol{x} = \boldsymbol{x}^*$ that needs to be predicted.

ϵ^*

$$\hat{y}^* = \boldsymbol{x}^{*'}\hat{\boldsymbol{\beta}}$$

$$Var(y^* - \hat{y}^*) = [1 + \boldsymbol{x}^{*'}(\boldsymbol{X'X})^{-1}\boldsymbol{x}^*]\sigma^2$$

hence, a $100(1-\alpha)\%$ prediction interval for y^* is

$$\hat{y}^* \pm t_{\alpha/2,n-k-1}\hat{\sigma}\sqrt{1 + \boldsymbol{x}^{*'}(\boldsymbol{X'X})^{-1}\boldsymbol{x}^*}$$

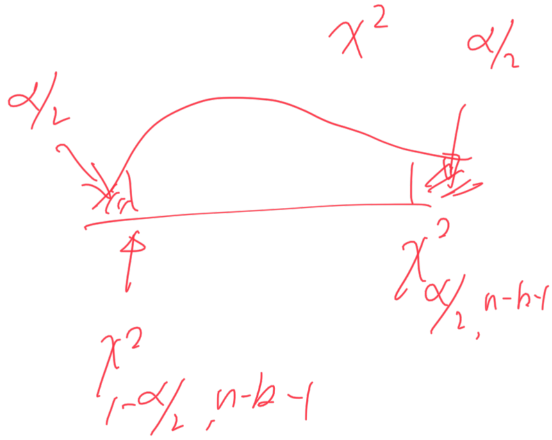
$$y^* - \hat{y}^* \sim N\left(0,\ \sigma^2 + x^{*'}(X'X)^{-1}x^*\sigma^2\right)$$

Confidence interval for σ^2

Note that $(n-k-1)\hat{\sigma}^2/\sigma^2 \sim \chi^2_{(n-k-1)}$. Therefore,

$$P[\chi^2_{1-\alpha/2,n-k-1} \leq \frac{(n-k-1)\hat{\sigma}^2}{\sigma^2} \leq \chi^2_{\alpha/2,n-k-1}] = 1-\alpha$$

hence, a $100(1-\alpha)\%$ confidence interval for σ^2 is



$$\left(\frac{(n-k-1)\hat{\sigma}^2}{\chi^2_{\alpha/2,n-k-1}}, \frac{(n-k-1)\hat{\sigma}^2}{\chi^2_{1-\alpha/2,n-k-1}}\right)$$

Simultaneous intervals

1. Familywise confidence level: $1 - \alpha_f$ implies that we are $100(1 - \alpha_f)\%$ confident that every interval contains its respective parameter.

2. Bonferroni confidence intervals

Remark 8.2

   - Individual confidence level $1 - \alpha_c$
   - m intervals
   - If we choose $\alpha_c = \alpha_f/m$, familywise confidence level $\geq 1 - \alpha_f$
   - For m linear functions $\boldsymbol{\lambda}_1'\boldsymbol{\beta}, \boldsymbol{\lambda}_2'\boldsymbol{\beta}, ..., \boldsymbol{\lambda}_m'\boldsymbol{\beta}$, the $100(1-\alpha)\%$ Bonferroni confidence intervals are

α/m

$$\boldsymbol{\lambda}_i'\hat{\boldsymbol{\beta}} \pm t_{\alpha/2m, n-k-1}\hat{\sigma}\sqrt{\boldsymbol{\lambda}_i'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{\lambda}_i}$$

   for $i = 1, ..., m$.

$\frac{\alpha}{\cancel{m}}/2$

3. Scheffé confidence intervals for all possible linear functions $\boldsymbol{\lambda}'\boldsymbol{\beta}$:

   The $100(1 - \alpha)\%$ conservative confidence interval for any and all $\boldsymbol{a}'\boldsymbol{\beta}$ is

λ

$$\boldsymbol{\lambda}'\hat{\boldsymbol{\beta}} \pm \hat{\sigma}\sqrt{(k+1)F_{\alpha,k+1,n-k-1}\boldsymbol{\lambda}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{\lambda}}$$