Statistical Inference and Multivariate Analysis (MA324)

LECTURE SLIDES
Lecture 30

Multiple Linear Regression

Indian Institute of Technology Guwahati

Jan-May 2023

Estimation of Error Variance (σ^2)

- It can be shown that

$$E(SS_{Res}) = (n - p - 1)\sigma^2$$

- Hence, $\hat{\sigma}^2 = \frac{SS_{Res}}{n-p-1} = MS_{Res}$ is an unbiased estimator of σ^2. Here MS_{Res} is residual mean square.
- Observed value of $\hat{\sigma}^2 = \frac{SS_{Res}}{n-p-1}$ is called ***Residual variance***. It's **square root is called *residual standard error*.**
- $\frac{(n-p-1)MS_{Res}}{\sigma^2} \sim \chi^2_{n-p-1}$.
- Here,

$$SS_{Res} = \sum_{i=1}^{n} e_i^2 = y^T y - \hat{\beta}_1^T X^T y,$$

Hypothesis Testing: Test for Significance of Regression ($\sim$ ANOVA)

- Assumptions: ϵ_i's are i.i.d $N(0, \sigma^2)$ Rvs.

$$\epsilon \sim N_n(\underset{\sim}{0}, \sigma^2 I_n)$$

- Want to test the hypothesis **if there is a linear relationship** between the response y and any of the regressor $x_1, \cdots, x_n$.

$H_0 : \beta_1 = \beta_2 = \cdots = \beta_p = 0$ ag. $H_1 : \beta_j \neq 0$ for atleast one $j = 1, \cdots, p$

- Therefore, the test statistic is

$$F_0 = \frac{SS_{Reg}/p}{\hat{\sigma}^2} = \frac{SS_{Reg}/p}{SS_{Res}/(n-p-1)} \sim F_{p,n-p-1}, \text{ under } H_0.$$

- Reject H_0 iff $F_0 > F_{p,n-p-1;\alpha}$ (at level α).

Hypothesis Testing for individual regression coefficients: β_j

- Want to test:

$$H_0 : \beta_j = 0 \text{ ag. } H_1 : \beta_j \neq 0$$

- Therefore, the test statistic is

$$t_0 = \frac{\hat{\beta}_j}{\sqrt{MS_{Res} C_{jj}}} \sim t_{n-p-1}, \text{ under } H_0,$$

where C_{jj} is the diagonal element of $(X^T X)^{-1}$

- Reject H_0 iff $|t_0| > t_{n-p-1,\alpha/2}$; (at level α).

Test of contribution of a subset of the regressors

- $$\begin{aligned} \underset{\sim}{y} &= \underset{\sim}{X}\underset{\sim}{\beta} + \underset{\sim}{\epsilon} \\ \implies \underset{\sim}{y} &= (\underset{\sim}{X}_1, \underset{\sim}{X}_2) \begin{pmatrix} \underset{\sim}{\beta}_1 \\ \underset{\sim}{\beta}_2 \end{pmatrix} + \underset{\sim}{\epsilon} \end{aligned}$$
- Want to test:
  $$H_0 : \underset{\sim}{\beta}_2 = 0 \text{ ag. } H_1 : \underset{\sim}{\beta}_2 \neq 0$$
- Based on the full model ($\underset{\sim}{y} = \underset{\sim}{X}\underset{\sim}{\beta} + \underset{\sim}{\epsilon}$), $\hat{\underset{\sim}{\beta}} = (X^T X)^{-1} X^T y$ and $SS_{Reg}(\underset{\sim}{\beta})$ has p degrees of freedom.
- To find the contribution of $\underset{\sim}{\beta}_2$, fit the model assuming $\underset{\sim}{\beta}_2 = 0$.
- The reduced model is $\underset{\sim}{y} = \underset{\sim}{X}_1\underset{\sim}{\beta}_1 + \underset{\sim}{\epsilon}$, $\hat{\underset{\sim}{\beta}}_1 = (X_1^T X_1)^{-1} X_1^T y$ and $SS_{Reg}(\underset{\sim}{\beta}_1)$ has $p - r$ degrees of freedom. Where r denotes the number of components in $\underset{\sim}{\beta}_2$

Test of contribution of a subset of the regressors

- $SS_{Reg}(\underset{\sim}{\beta}_2|\underset{\sim}{\beta}_1) = SS_{Reg}(\underset{\sim}{\beta}) - SS_{Reg}(\underset{\sim}{\beta}_1)$ can be used as a measure of contribution of $\underset{\sim}{\beta}_2$.

- Note that if $\underset{\sim}{\beta}_2$ has significant contribution then $SS_{Reg}(\underset{\sim}{\beta}_2|\underset{\sim}{\beta}_1)$ is large.

- Therefore, the test statistic is

$$F_0 = \frac{SS_{Reg}(\underset{\sim}{\beta}_2|\underset{\sim}{\beta}_1)/r}{\hat{\sigma}^2} = \frac{SS_{Reg}(\underset{\sim}{\beta}_2|\underset{\sim}{\beta}_1)/r}{SS_{Res}/(n-p-1)} \sim F_{r,n-p-1}, \text{ under } H_0.$$

- Reject H_0 iff $F_0 > F_{r,n-p-1;\alpha}$ (at level α).

Testing of general linear hypothesis

- Want to test:
$$H_0 : T\underset{\sim}{\beta} = 0 \text{ ag. } H_1 : T\underset{\sim}{\beta} \neq 0,$$
where T is a $m \times (p+1)$ matrix of constants.
- Examples: $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \epsilon$
    - $H_0 : \beta_1 = \beta_3$ ag. $H_1 : \beta_1 \neq \beta_3$. Take, $T = \begin{bmatrix} 0 & 1 & 0 & -1 \end{bmatrix}$.
    - $H_0 : \beta_1 = \beta_3, \beta_2 = 0$ ag. $H_1 : \beta_1 \neq \beta_3$ or $\beta_2 \neq 0$. Take,
$$T = \begin{bmatrix} 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$
.
- The full model (FM) is $\underset{\sim}{y} = \underset{\sim}{X}\underset{\sim}{\beta} + \underset{\sim}{\epsilon}$, $\hat{\underset{\sim}{\beta}} = (X^T X)^{-1} X^T y$ and
$SS_{Res}(FM) = y^T y - \hat{\beta}^T X^T y$ has $n - p - 1$ degrees of freedom.
- Now assume that T has $r(\leq m)$ independent rows. Then $T\underset{\sim}{\beta}$ can be solved and r of the β_j's in FM can be written in terms of other $(p+1-r)$ β_j's.

Testing of general linear hypothesis

- This lead to the reduced model (RM)

$$\underset{\sim}{y} = \underset{\sim}{Z}\underset{\sim}{\gamma} + \underset{\sim}{\epsilon}^*,$$

where Z is $n \times \overline{p+1-r}$ matrix.

$$\hat{\underset{\sim}{\gamma}} = (Z^T Z)^{-1} Z^T y, \text{ and } SS_{Res}(RM) = y^T y - \hat{\gamma}^T Z^T y$$

has $n-p-1+r$ degrees of freedom.

- $SS_{Res}(FM) \leq SS_{Res}(RM)$
- Consider $SS_H = SS_{Res}(RM) - SS_{Res}(FM)$ with d.f $(n-p-1+r) - (n-p-1) = r$
- Therefore, the test statistic is

$$F_0 = \frac{SS_H/r}{SS_{Res}(FM)/(n-p-1)} \sim F_{r,n-p-1}, \text{ under } H_0.$$

- Reject H_0 iff $F_0 > F_{r,n-p-1;\alpha}$ (at level α).

Confidence Intervals (CIs):

- Confidence Interval of individual regression coefficient $\hat{\beta}_j$
- Pivot is

$$\frac{\hat{\beta}_j - \beta_j}{\sqrt{MS_{Res} C_{jj}}} \sim t_{n-p-1},$$

where C_{jj} is the diagonal element of $(X^T X)^{-1}$ matrix.

- A $100(1-\alpha)\%$ CI for β_j is

$$\left[\hat{\beta}_j \pm t_{n-p-1,\alpha/2} \sqrt{MS_{Res} C_{jj}} \right].$$

CI for mean response:

- Let $\underset{\sim}{x_0} = \begin{pmatrix} 1 \\ x_{01} \\ \vdots \\ x_{0p} \end{pmatrix}$ be a value of the regressor vector. The mean response at $\underset{\sim}{x_0}$ is $\underset{\sim}{x_0}^T \underset{\sim}{\beta}$.

- Confidence Interval of individual regression coefficient $\hat{\beta}_j$

- Pivot is $\frac{\hat{\underset{\sim}{y_0}} - \underset{\sim}{x_0}^T \underset{\sim}{\beta}}{\sqrt{MS_{Res} \underset{\sim}{x_0}^T (X^T X)^{-1} \underset{\sim}{x_0}}} \sim t_{n-p-1}$.

- A $100(1-\alpha)\%$ CI for β_j is

$$\left[\hat{\underset{\sim}{y_0}} \pm t_{n-p-1,\alpha/2} \sqrt{MS_{Res} \underset{\sim}{x_0}^T (X^T X)^{-1} \underset{\sim}{x_0}} \right].$$