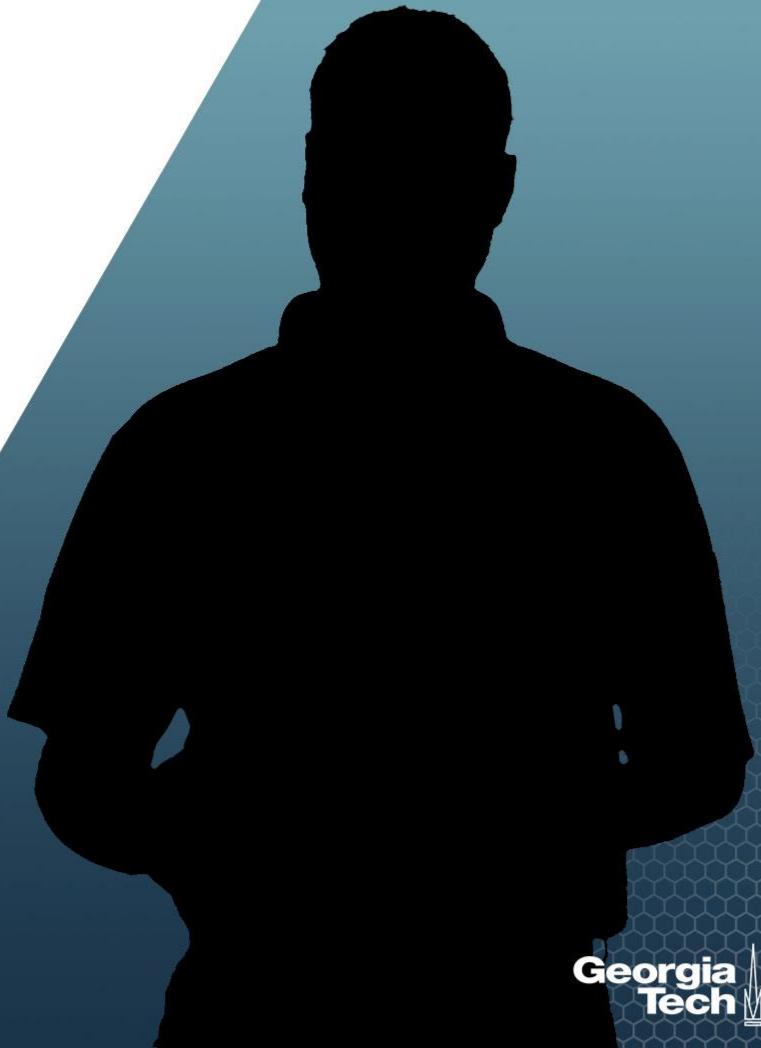
Regression Analysis

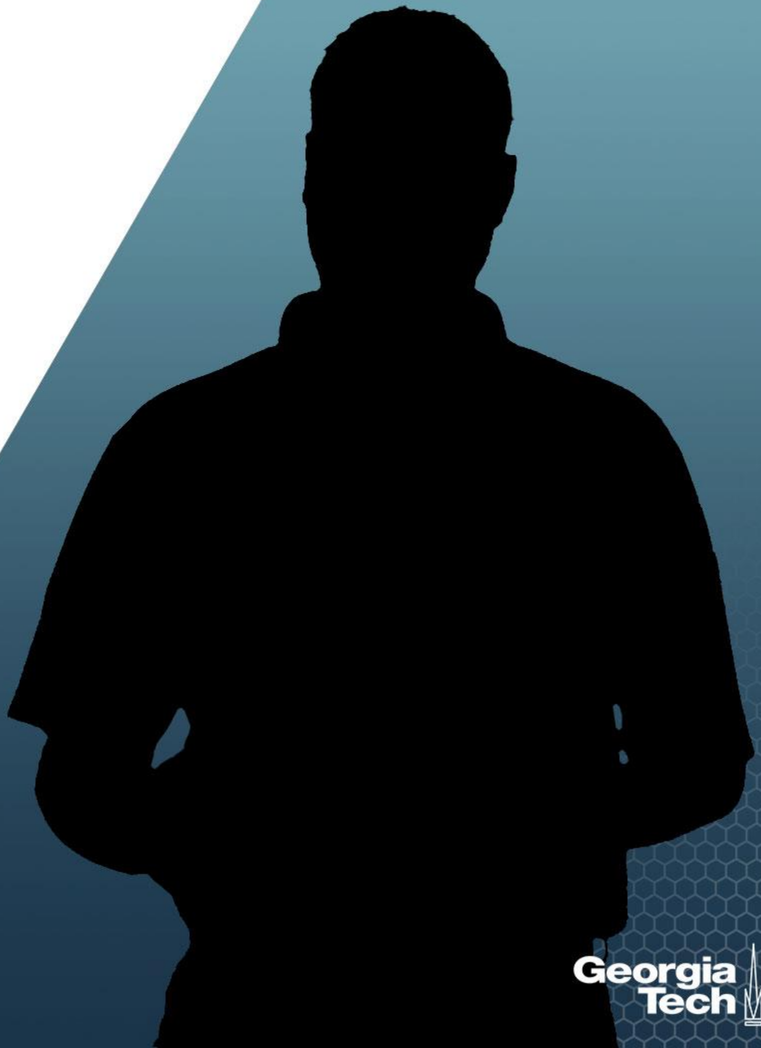
Multiple Linear Regression

Nicoleta Serban, Ph.D.
Professor
School of Industrial and Systems Engineering

Inference for Regression Parameters



About This Lesson



Sampling Distribution

$$\mathrm{E}\left(\hat{\beta}\right) = \beta$$

$$\mathrm{V}(\hat{\beta}) = \sigma^2\left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\right)^{-1} = \Sigma$$

Furthermore, $\widehat{\boldsymbol{\beta}}$ is a linear combination of $\{y_1,\ldots,y_n\}$. If we assume that $\varepsilon_i \sim$ Normal $(0, \sigma^2)$, then $\widehat{\boldsymbol{\beta}}$ is also distributed as

$$\widehat{\boldsymbol{\beta}} \sim \mathbf{N}(\boldsymbol{\beta}, \boldsymbol{\Sigma})$$

Georgia Tech

Properties of Regression Estimators

$$\widehat{\boldsymbol{\beta}} \sim \mathbf{N}(\boldsymbol{\beta}, \boldsymbol{\Sigma})$$

σ^2 is unknown!

Replace σ^2 with $\widehat{\sigma}^{\mathbf{2}} = \mathrm{MSE}$

$$\widehat{\sigma}^2 = \frac{\sum \widehat{\varepsilon}_i^2}{n-p-1} \sim \chi^2_{n-p-1} \quad \longrightarrow \quad \frac{\hat{\beta}_j - \beta_j}{\sqrt{\mathrm{V}\left(\hat{\beta}_j\right)}} \sim t_{n-p-1}$$

(chi-squared distribution with $n-p-1$ degrees of freedom)

(t-distribution with $n-p-1$ degrees of freedom)

Georgia Tech

Confidence Interval Estimation

We can derive confidence intervals for β_j using this *t* sampling distribution:

$$\boxed{\hat{\beta}_j \pm \left(t_{\alpha/2,\, n-p-1}\right)\left(\mathrm{SE}(\hat{\beta}_j)\right)}$$

Is $\boldsymbol{\beta_j}$ statistically significant?

- Check whether zero is in the confidence interval

Why is this a *t*-interval?

Confidence Interval Estimation

Why is this a *t*-interval?

$$\frac{\hat{\beta}_j - \beta_j}{\sqrt{\mathrm{V}\left(\hat{\beta}_j\right)}} = \frac{\hat{\beta}_j - \beta_j}{\mathrm{SE}\left(\hat{\beta}_j\right)} \sim \mathrm{T}_{n-p-1} \longrightarrow$$

t-interval for β_j

1-α Confidence Interval for β_j $\longrightarrow$

$$\underbrace{\hat{\beta}_j}_{\text{Estimate of } \beta_j} \pm \underbrace{\left(t_{\alpha/2,\, n-p-1}\right)}_{t\text{-critical point}} \underbrace{\left(\mathrm{SE}(\hat{\beta}_j)\right)}_{\text{Standard Deviation/Error of } \hat{\beta}_j}$$

Testing Statistical Significance

To test for statistical significance of β_j given all other predicting variables in the model, use a *t*-test for H_0 and H_a:

$$\boxed{\begin{array}{c}\mathrm{H}_0\colon \beta_j = 0 \\ \text{vs.} \\ \mathrm{H}_\mathrm{a}\colon \beta_j \neq 0\end{array}}$$

$$\boxed{t\text{−value} = \frac{\hat{\beta}_j - 0}{\mathrm{SE}(\hat{\beta}_j)}}$$

- Reject H_0 if |*t*-value| gets too large
- Interpret rejecting the null hypothesis as β_j being statistically significant

Testing Statistical Significance

How will the procedure change if

$$\begin{gathered} \mathrm{H_0}\colon \beta_j = b \\ \text{vs.} \\ \mathrm{H_a}\colon \beta_j \neq b \end{gathered}$$

for some known b?

we test

$$t\text{−value} = \frac{\hat{\beta}_j - b}{\mathrm{SE}(\hat{\beta}_j)}$$

- Reject $\mathrm{H_0}$ if |*t*-value| is large
  - For significance level α, if $|t\text{−value}| > t_{\alpha/2,\, n-p-1}$ $\longrightarrow$ reject $\mathrm{H_0}$
- Alternatively, compute a p-value based on the probability that the *t* distribution is greater than the *t*-value:

$$\text{p-value} = 2\mathrm{Prob}(\mathrm{T}_{n\text{-}p\text{-}1} > |t\text{-value}|)$$

- If p-value is small (e.g., < 0.01) $\longrightarrow$ reject $\mathrm{H_0}$

Testing Statistical Significance

How will the procedure change if we test whether a coefficient is statistically positive or negative?

Test for Statistically Positive

$$\boxed{\begin{array}{c} \mathrm{H_0}: \beta_j \leq 0 \\ \text{vs.} \\ \mathrm{H_a}: \beta_j > 0 \end{array}}$$

$$\text{p-value} = \mathrm{Prob}(\mathrm{T}_{n\text{-}p\text{-}1} > t\text{-value})$$

Test for Statistically Negative

$$\boxed{\begin{array}{c} \mathrm{H_0}: \beta_j \geq 0 \\ \text{vs.} \\ \mathrm{H_a}: \beta_j < 0 \end{array}}$$

$$\text{p-value} = \mathrm{Prob}(\mathrm{T}_{n\text{-}p\text{-}1} < t\text{-value})$$

Summary