

Applied Statistical Analysis I

Multiple linear regression

Elena Karagianni, PhD Candidate

karagiae@tcd.ie

 November 19, 2025

Department of Political Science, Trinity College Dublin

t-test for individual coefficients

What are we testing?

t-test for individual coefficients

What are we testing?

- $H_0 : \beta = 0$ (no association between predictor and outcome)
- $H_1 : \beta \neq 0$ (association exists)

t-test for individual coefficients

What are we testing?

- $H_0 : \beta = 0$ (no association between predictor and outcome)
- $H_1 : \beta \neq 0$ (association exists)

Test statistic:

$$t = \frac{\hat{\beta} - \beta_{H_0}}{\hat{\sigma}_{\hat{\beta}}} = \frac{\hat{\beta}}{\hat{\sigma}_{\hat{\beta}}}$$

t-test for individual coefficients

What are we testing?

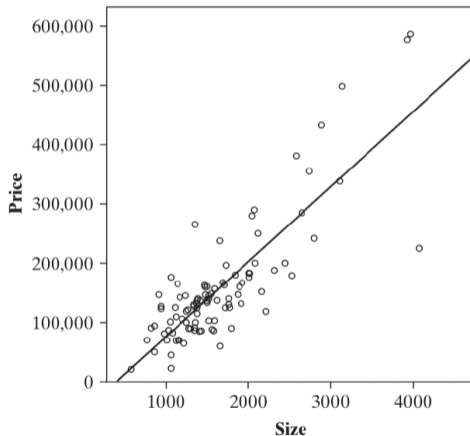
- $H_0 : \beta = 0$ (no association between predictor and outcome)
- $H_1 : \beta \neq 0$ (association exists)

Test statistic:

$$t = \frac{\hat{\beta} - \beta_{H_0}}{\hat{\sigma}_{\hat{\beta}}} = \frac{\hat{\beta}}{\hat{\sigma}_{\hat{\beta}}}$$

Interpretation: The t -value tells us how many *standard errors* the estimate is away from the null hypothesis value.

t-test example: house price and size



- Is there an association between house selling price and size?
- Estimated regression:

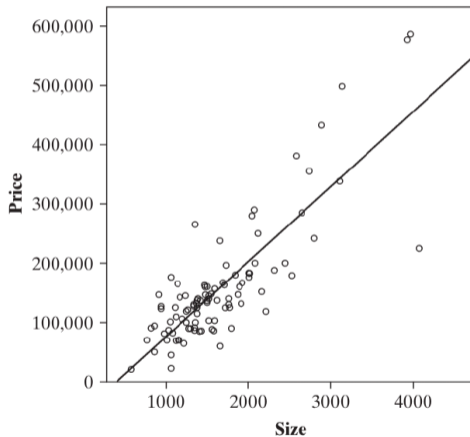
$$\widehat{\text{Price}} = 50,926.2 + 126.6 \times \text{Size}$$

- Test statistic:

$$t = \frac{126.6}{8.47} = 14.95$$

- How likely is this if $H_0: \beta = 0$ is true?

t-test example: house price and size



- Is there an association between house selling price and size?
- Estimated regression:

$$\widehat{\text{Price}} = 50,926.2 + 126.6 \times \text{Size}$$

- Test statistic:

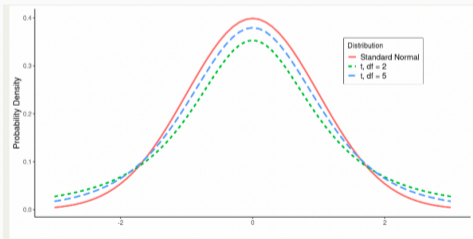
$$t = \frac{126.6}{8.47} = 14.95$$

- How likely is this if $H_0: \beta = 0$ is true?
- A t -value of 14.95 is extremely unlikely under $H_0 \Rightarrow \text{p-value} \approx 0$.

t-test: Intuition

Why does a large t lead us to reject H_0 ?

- Under H_0 , the sampling distribution of $\hat{\beta}$ is centered at 0.
- If the observed $\hat{\beta}$ is many standard errors away from 0, it is unlikely to occur by chance.
- The t -distribution lets us compute this probability (the p-value).



Large $|t| \rightarrow$ tiny p-value \rightarrow reject H_0 .

Interpreting the t-value

$$t = 14.95$$

What does this mean? (In other words:)

- The slope is **14.95 standard errors above 0**.
- Under H_0 , such an extreme value is virtually impossible.
- Therefore, we reject H_0 .

The F-test compares two models:

- Reduced Model (simpler)
- Full Model (includes all predictors)

F-test: Comparing model fit

General idea: Does adding predictors meaningfully improve the fit?

$$F = \frac{\text{SSE(RM)} - \text{SSE(FM)}}{p + 1 - k} \bigg/ \frac{\text{SSE(FM)}}{n - p - 1}$$

- SSE = sum of squared errors (lower = better fit)
- Numerator: improvement in fit
- Denominator: remaining unexplained variance
- $p \rightarrow$ number of IVs in full model
- $n \rightarrow$ number of observations
- $k \rightarrow$ number of parameters in reduced model

Understanding the components of the F-test

- $SSE(RM) - SSE(FM)$: improvement in residual error when adding predictors.
- Weighting by degrees of freedom ensures fair comparison.
- Large $F \rightarrow$ full model explains far more variance than reduced model.

Two versions of the F-test

1. **Global F-test:** All the regression coefficients are zero.
2. **Partial F-test:** Some of the regression coefficients are zero.

An example

Table 3.5 Regression Output for Supervisor Performance Data

Variable	Coefficient	s.e.	<i>t</i> -Test	<i>p</i> -value
Constant	10.787	11.5890	0.93	0.3616
X_1	0.613	0.1610	3.81	0.0009
X_2	-0.073	0.1357	-0.54	0.5956
X_3	0.320	0.1685	1.90	0.0699
X_4	0.081	0.2215	0.37	0.7155
X_5	0.038	0.1470	0.26	0.7963
X_6	-0.217	0.1782	-1.22	0.2356
$n = 30$	$R^2 = 0.73$	$R_u^2 = 0.66$	$\hat{\sigma} = 7.068$	$df = 23$

Table 3.2 Description of Variables in Supervisor Performance Data

Variable	Description
Y	Overall rating of job being done by supervisor
X_1	Handles employee complaints
X_2	Does not allow special privileges
X_3	Opportunity to learn new things
X_4	Raises based on performance
X_5	Too critical of poor performance
X_6	Rate of advancing to better jobs

F-Test: Are all coefficients zero?

“All regression slopes are zero.”

- **Reduced model (RM):** $Y = \beta_0 + \varepsilon$

Only the intercept; this is H_0 .

- **Full model (FM):** $Y = \beta_0 + \beta_1 X_1 + \cdots + \beta_p X_p + \varepsilon$

At least one slope $\neq 0$ under H_1 .

$$F = \frac{[\text{SSE(RM)} - \text{SSE (FM)}]/(p + 1 - k)}{\text{SSE(FM)} / (n - p - 1)} = \frac{[\text{SST} - \text{SSE}]/p}{\text{SSE} / (n - p - 1)} = \frac{\text{SSR}/p}{\text{SSE} / (n - p - 1)}$$

F-Test: Are all coefficients zero?

“All regression slopes are zero.”

- **Reduced model (RM):** $Y = \beta_0 + \varepsilon$

Only the intercept; this is H_0 .

- **Full model (FM):** $Y = \beta_0 + \beta_1 X_1 + \cdots + \beta_p X_p + \varepsilon$

At least one slope $\neq 0$ under H_1 .

$$F = \frac{[\text{SSE(RM)} - \text{SSE (FM)}]/(p + 1 - k)}{\text{SSE(FM)} / (n - p - 1)} = \frac{[\text{SST} - \text{SSE}]/p}{\text{SSE} / (n - p - 1)} = \frac{\text{SSR}/p}{\text{SSE} / (n - p - 1)}$$

Why? In RM, the best prediction is \bar{y} , so $\text{SSE(RM)} = \text{SST}$ and $\text{SST} = \text{SSR} + \text{SSE}$.

Interpreting the global F-test

Table 3.5 Regression Output for Supervisor Performance Data

Variable	Coefficient	s.e.	t-Test	p-value
Constant	10.787	11.5890	0.93	0.3616
X_1	0.613	0.1610	3.81	0.0009
X_2	-0.073	0.1357	-0.54	0.5956
X_3	0.320	0.1685	1.90	0.0699
X_4	0.081	0.2215	0.37	0.7155
X_5	0.038	0.1470	0.26	0.7963
X_6	-0.217	0.1782	-1.22	0.2356
$n = 30$	$R^2 = 0.73$	$R_a^2 = 0.66$	$\hat{\sigma} = 7.068$	$df = 23$

Table 3.7 Supervisor Performance Data: Analysis of Variance (ANOVA) Table

Source	Sum of Squares	df	Mean Square	F-Test
Regression	3147.97	6	524.661	10.5
Residuals	1149.00	23	49.9565	

$$F = \frac{3147.97/6}{1149/23} = 10.50$$

Meaning: The full model fits the data significantly better than the null model. At least one predictor is associated with the outcome.

Partial F-Test: Testing a subset of coefficients

“Some regression slopes are zero.”

- Reduced model (RM):

$$Y = \beta_0 + \beta_1 X_1 + \beta_3 X_3 + \varepsilon$$

The slopes for the excluded variables are set to zero under H_0 .

- Full model (FM):

$$Y = \beta_0 + \beta_1 X_1 + \cdots + \beta_p X_p + \varepsilon$$

At least one of the excluded slopes differs from zero under H_1 .

$$F = \frac{[\text{SSE(RM)} - \text{SSE(FM)}]/q}{\text{SSE(FM)}/(n - p - 1)}$$

Partial F-Test: Testing a subset of coefficients

“Some regression slopes are zero.”

- **Reduced model (RM):**

$$Y = \beta_0 + \beta_1 X_1 + \beta_3 X_3 + \varepsilon$$

The slopes for the excluded variables are set to zero under H_0 .

- **Full model (FM):**

$$Y = \beta_0 + \beta_1 X_1 + \cdots + \beta_p X_p + \varepsilon$$

At least one of the excluded slopes differs from zero under H_1 .

$$F = \frac{[\text{SSE(RM)} - \text{SSE(FM)}]/q}{\text{SSE(FM)}/(n - p - 1)}$$

q = number of coefficients tested jointly (size of the subset).

Partial F-test: Testing a subset of predictors

Question: Test whether X_1 and X_3 contribute to the model.

Table 3.8 Regression Output from the Regression of Y on X_1 and X_3

ANOVA Table				
Source	Sum of Squares	df	Mean Square	F-Test
Regression	3042.32	2	1521.1600	32.7
Residuals	1254.65	27	46.4685	
Coefficients Table				
Variable	Coefficient	s.e.	t-Test	p-value
Constant	9.8709	7.0610	1.40	0.1735
X_1	0.6435	0.1185	5.43	< 0.0001
X_3	0.2112	0.1344	1.57	0.1278
$n = 30$	$R^2 = 0.708$	$R_a^2 = 0.686$	$\hat{\sigma} = 6.817$	df = 27

$$F = \frac{(1254.65 - 1149)/4}{1149/23} = 0.0528$$

Interpreting the partial F-test

$$F = 0.0528$$

This is very small. Why?

- Removing X_1 and X_3 barely increases SSE.

Interpreting the partial F-test

$$F = 0.0528$$

This is very small. Why?

- Removing X_1 and X_3 barely increases SSE.
- This means they do not explain meaningful variation beyond other predictors.
- p-value $> 0.05 \rightarrow$ retain H_0 .

Conclusion: The reduced model performs just as well.