Section 4. Multiple Regression: Hypothesis Testing

1. **Interpreting the coefficient on X_2 (Frisch-Waugh Theorem)**
2. **Hypothesis Testing for a Single Coefficient**
3. **CI and Hypothesis Tests for Single Coefficient**
4. **Confidence Regions for Multiple Coefficients**
5. **Hypothesis Tests for Multiple Coefficients**

Review: Multiple Regression

The Multiple Regression Model

The multiple regression model is

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki} + u_i, \; i = 1, \ldots, n. \qquad (5.7)$$

where:

- Y_i is i^{th} observation on the dependent variable; $X_{1i}, X_{2i}, \ldots, X_{ki}$ are the i^{th} observations on each of the k regressors; and u_i is the error term.
- The population regression line is the relationship that holds between Y and the X's on average in the population:

$$E(Y|X_{1i} = x_1, X_{2i} = x_2, \ldots, X_{ki} = x_k) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k.$$

- β_1 is the slope coefficient on X_1, β_2 is the coefficient on X_2, etc. The coefficient β_1 is the expected change in Y_i resulting from changing X_{1i} by one unit, holding constant $X_{2i}, \ldots, X_{ki}$. The coefficients on the other X's are interpreted similarly.
- The intercept β_0 is the expected value of Y when all the X's equal zero. The intercept can be thought of as the coefficient on a regressor, X_{0i}, that equals one for all i.

1. Interpreting MVR coefficients: Frisch-Waugh

- In the equation $y = b_0 + b_1X_1 + b_2X_2 + e$, we interpret b_1 and b_2 as partial derivatives.
- Detrending question
- Frisch-Waugh theorem:
  A) Regress X_2 on X_1: $X_2 = c_1X_1 + e_2$
  B) Regress y on e_2 : $y = c_2e_2 + e_y$

  then.. $c_2 = b_2$.

  Proof

  Application to scatterplots

2. Hypothesis Testing for A Single Coefficient

Testing the Hypothesis $\beta_j = \beta_{j,0}$ Against the Alternative $\beta_j \neq \beta_{j,0}$

1. Compute the standard error of $\hat{\beta}_j$, $SE(\hat{\beta}_j)$.
2. Compute the t-statistic,

$$t = \frac{\hat{\beta}_j - \beta_{j,0}}{SE(\hat{\beta}_j)}. \tag{5.14}$$

3. Compute the p-value,

$$p\text{-value} = 2\Phi(-|t^{act}|), \tag{5.15}$$

where t^{act} is the value of the t-statistic actually computed. Reject the hypothesis at the 5% significance level if the p-value is less than 0.05 or, equivalently, if $|t^{act}| > 1.96$.

The standard error and (typically) the t-statistic and p-value testing $\beta_j = 0$ are computed automatically by regression software.

3. Confidence Intervals with a Single Coefficient

Confidence Intervals for a Single Coefficient in Multiple Regression

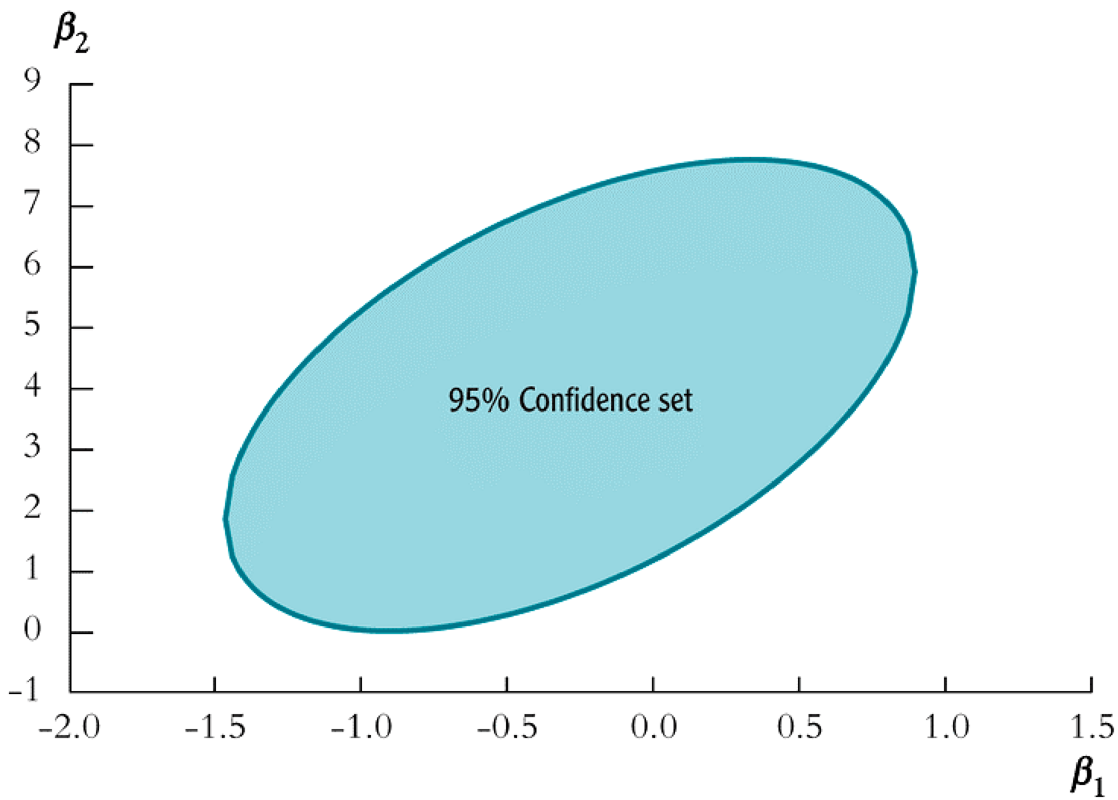
A 95% two-sided confidence interval for the coefficient β_j is an interval that contains the true value of β_j with a 95% probability; that is, it contains the true value of β_j in 95% of all possible randomly drawn samples. Equivalently, it is also the set of values of β_j that cannot be rejected by a 5% two-sided hypothesis test. When the sample size is large, the 95% confidence interval is:

$$95\% \text{ confidence interval for } \beta_j = (\hat{\beta}_j - 1.96SE(\hat{\beta}_j), \hat{\beta}_j + 1.96SE(\hat{\beta}_j)). \quad (5.17)$$

A 90% confidence interval is obtained by replacing 1.96 in Equation (5.17) with 1.645.

4. Confidence Regions for Multiple Coefficients

FIGURE 5.1 **95% Confidence Set for β_1 and β_2**



The 95% confidence set for β_1 and β_2 is an ellipse. The ellipse contains the pairs of values of β_1 and β_2 that cannot be rejected using the *F*-statistic at the 5% significance level.

Why F?

Because the product of random variables with standard N(.) distributions have F distributions.

Large-Sample Critical Values for the F-statistic from the $F_{m,\infty}$ Distribution

Reject if F > Critical Value

Degrees of Freedom (m)	Significance Level 10%	5%	1%
1	2.71	3.84	6.63
2	2.30	3.00	4.61
3	2.08	2.60	3.78
4	1.94	2.37	3.32
5	1.85	2.21	3.02
6	1.77	2.10	2.80
7	1.72	2.01	2.64
8	1.67	1.94	2.51
9	1.63	1.88	2.41
10	1.60	1.83	2.32
11	1.57	1.79	2.25
12	1.55	1.75	2.18
13	1.52	1.72	2.13
14	1.50	1.69	2.08
15	1.49	1.67	2.04
16	1.47	1.64	2.00
17	1.46	1.62	1.97
18	1.44	1.60	1.93
19	1.43	1.59	1.90
20	1.42	1.57	1.88
21	1.41	1.56	1.85
22	1.40	1.54	1.83
23	1.39	1.53	1.81
24	1.38	1.52	1.79
25	1.38	1.51	1.77
26	1.37	1.50	1.76
27	1.36	1.49	1.74
28	1.35	1.48	1.72
29	1.35	1.47	1.71
30	1.34	1.46	1.70

5. Hypothesis Testing for multiple coefficients

- H_o: $\$_j = \$_{j,0}, \ldots \$_m = \$_{m,0}$ for q restrictions
- H_A: At least one restriction doesn't hold.
- Test statistic will be $F_{q,\%}$