

# Hypothesis Tests and Confidence Intervals in Multiple Regression (SW Ch. 7) Part 1

Dragos Ailoae  
dailoae@gradcenter.cuny.edu

Advanced Economics and Business Statistics  
ECON-4400w - Spring 2022

Brooklyn College  
Apr 11, 2022

# Understanding regression output tables

## SUMMARY OUTPUT

<i>Regression Statistics</i>	
Multiple R	0.996
R Square	0.992
Adjusted R Square	0.992
Standard Error	0.990
Observations	100

## ANOVA

	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	2	11763.151	5881.576	5994.941	0.000
Residual	97	95.166	0.981		
Total	99	11858.317			

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>
Intercept	9.990	0.099	100.829	0.000	9.793	10.186
X	1.986	0.061	32.457	0.000	1.865	2.107
Z	3.019	0.047	64.493	0.000	2.926	3.112

# Outline

- 1. Hypothesis tests and confidence intervals for one coefficient**
- 2. Joint hypothesis tests on multiple coefficients**
- 3. Other types of hypotheses involving multiple coefficients**
4. Model specification: how to decide which variables to include in a regression model

# Hypothesis Tests and Confidence Intervals for a Single Coefficient (SW Section 7.1)

- Hypothesis tests and confidence intervals for a single coefficient in multiple regression follow the same logic and recipe as for the slope coefficient in a single-regressor model.
- $\frac{\hat{\beta}_1 - E(\hat{\beta}_1)}{\sqrt{\text{var}(\hat{\beta}_1)}}$  is approximately distributed  $N(0,1)$  (CLT).
- Thus hypotheses on  $\beta_1$  can be tested using the usual  $t$ -statistic, and confidence intervals are constructed as  $\{\hat{\beta}_1 \pm 1.96 \times \text{SE}(\hat{\beta}_1)\}$ .
- So too for  $\beta_2, \dots, \beta_k$ .

## *Example: The California class size data*

1.  $\widehat{TestScore} = 698.9 - 2.28 \times STR$   
(10.4) (0.52)

2.  $\widehat{TestScore} = 686.0 - 1.10 \times STR - 0.650PctEL$   
(8.7) (0.43) (0.031)

- The coefficient on  $STR$  in (2) is the effect on  $TestScores$  of a unit change in  $STR$ , holding constant the percentage of English Learners in the district
- The coefficient on  $STR$  falls by one-half
- The 95% confidence interval for coefficient on  $STR$  in (2) is  $\{-1.10 \pm 1.96 \times 0.43\} = (-1.95, -0.26)$
- The  $t$ -statistic testing  $\beta_{STR} = 0$  is  $t = -1.10/0.43 = -2.54$ , so we reject the hypothesis at the 5% significance level

# Standard errors in multiple regression in STATA

```
reg testscr str pctel, robust;
```

Regression with robust standard errors

```
Number of obs =      420
F(   2,   417) =  223.82
Prob > F       =   0.0000
R-squared      =   0.4264
Root MSE      =   14.464
```

-----						
		Robust				
testscr		Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
-----+-----						
str		-1.101296	.4328472	-2.54	0.011	-1.95213    -.2504616
pctel		-.6497768	.0310318	-20.94	0.000	-.710775    -.5887786
_cons		686.0322	8.728224	78.60	0.000	668.8754    703.189
-----						

$$\widehat{TestScore} = 686.0 - 1.10 \times STR - 0.650PctEL$$

(8.7)    (0.43)            (0.031)

We use [heteroskedasticity-robust standard errors](#) – for exactly the same reason as in the case of a single regressor.

# Tests of Joint Hypotheses (SW Section 7.2)

(1 of 2)

Let  $Expn$  = expenditures per pupil and consider the population regression model:

$$TestScore_i = \beta_0 + \beta_1 STR_i + \beta_2 Expn_i + \beta_3 PctEL_i + u_i$$

The null hypothesis that “school resources don’t matter,” and the alternative that they do, corresponds to:

$$H_0: \beta_1 = 0 \text{ *and* } \beta_2 = 0$$

$$\text{vs. } H_1: \text{*either* } \beta_1 \neq 0 \text{ *or* } \beta_2 \neq 0 \text{ *or both*}$$

$$TestScore_i = \beta_0 + \beta_1 STR_i + \beta_2 Expn_i + \beta_3 PctEL_i + u_i$$

# Tests of Joint Hypotheses (SW Section 7.2)

(2 of 2)

- $H_0: \beta_1 = 0$  *and*  $\beta_2 = 0$
- vs.  $H_1: \textit{either } \beta_1 \neq 0 \textit{ or } \beta_2 \neq 0 \textit{ or both}$
- A *joint hypothesis* specifies a value for two or more coefficients, that is, it imposes a restriction on two or more coefficients.
- In general, a joint hypothesis will involve  $q$  restrictions. In the example above,  $q = 2$ , and the two restrictions are  $\beta_1 = 0$  and  $\beta_2 = 0$ .
- A “common sense” idea is to reject if either of the individual  $t$ -statistics exceeds 1.96 in absolute value.
- But this “one at a time” test isn’t valid: the resulting test rejects too often under the null hypothesis (more than 5%)!



## *Why can't we just test the coefficients one at a time?*

Because the rejection rate under the null isn't 5%. We'll calculate the probability of incorrectly rejecting the null using the “common sense” test based on the two individual  $t$ -statistics. To simplify the calculation, suppose that  $\hat{\beta}_1$  and  $\hat{\beta}_2$  are independently distributed (this isn't true in general – just in this example). Let  $t_1$  and  $t_2$  be the  $t$ -statistics:

$$t_1 = \frac{\hat{\beta}_1 - 0}{SE(\hat{\beta}_1)} \quad \text{and} \quad t_2 = \frac{\hat{\beta}_2 - 0}{SE(\hat{\beta}_2)}$$

The “one at time” test is:

reject  $H_0: \beta_1 = \beta_2 = 0$  if  $|t_1| > 1.96$  and/or  $|t_2| > 1.96$

What is the probability that this “one at a time” test rejects  $H_0$ , when  $H_0$  is actually true? (It *should* be 5%.)

*Suppose  $t_1$  and  $t_2$  are independent*  
**(for this example).**

The probability of incorrectly rejecting the null hypothesis using the “one at a time” test

$$= \Pr_{H_0} [|t_1| > 1.96 \text{ and/or } |t_2| > 1.96]$$

$$= 1 - \Pr_{H_0} [|t_1| \leq 1.96 \text{ and } |t_2| \leq 1.96]$$

$$= 1 - \Pr_{H_0} [|t_1| \leq 1.96] \times \Pr_{H_0} [|t_2| \leq 1.96]$$

(because  $t_1$  and  $t_2$  are independent by assumption)

$$= 1 - (.95)^2$$

$$= .0975 = 9.75\% - \text{which is *not* the desired 5\%!!}$$

# The *size* of a test is the actual rejection rate under the null hypothesis.

- The size of the “common sense” test isn’t 5% !
- In fact, its size depends on the correlation between  $t_1$  and  $t_2$  (and thus on the correlation between  $\hat{\beta}_1$  and  $\hat{\beta}_2$ ).

## Two Solutions:

- Use a different critical value in this procedure – not 1.96 (this is the “Bonferroni method – see SW App. 7.1) (this method is rarely used in practice however)
- Use a different test statistic designed to test *both*  $\beta_1$  and  $\beta_2$  at once: the  $F$ -statistic (this is common practice)

## More on $F$ -statistics.

*There is a simple formula for the  $F$ -statistic that holds only under homoskedasticity (so it isn't very useful) but which nevertheless might help you understand what the  $F$ -statistic is doing.*

### **The homoskedasticity-only $F$ -statistic**

When the errors are homoskedastic, there is a simple formula for computing the “homoskedasticity-only”  $F$ -statistic:

- Run two regressions, one under the null hypothesis (the “restricted” regression) and one under the alternative hypothesis (the “unrestricted” regression).
- Compare the fits of the regressions – the  $R^2$ s – if the “unrestricted” model fits sufficiently better, reject the null

# *The “restricted” and “unrestricted” regressions*

*Example:* are the coefficients on  $STR$  and  $Expn$  zero?

Unrestricted population regression (under  $H_1$ ):

$$TestScore_i = \beta_0 + \beta_1 STR_i + \beta_2 Expn_i + \beta_3 PctEL_i + u_i$$

Restricted population regression (that is, under  $H_0$ ):

$$TestScore_i = \beta_0 + \beta_3 PctEL_i + u_i \quad (why?)$$

- The number of restrictions under  $H_0$  is  $q = 2$  (why?).
- The fit will be better ( $R^2$  will be higher) in the unrestricted regression (why?)

By how much must the  $R^2$  increase for the coefficients on  $Expn$  and  $PctEL$  to be judged statistically significant?

## *The homoskedasticity-only F-statistic:*

$$F = \frac{(SSR_{\text{restricted}} - SSR_{\text{unrestricted}})/q}{SSR_{\text{unrestricted}} / (n - k_{\text{unrestricted}} - 1)}$$

where:

$SSR_{\text{restricted}}$  = sum of squared residuals from the restricted regression

$SSR_{\text{unrestricted}}$  = sum of squared residuals from the unrestricted regression

$q$  = number of restrictions under the null hypothesis

$k_{\text{unrestricted}}$  = number of regressors in the unrestricted regression

## *Alternate formula for the homoskedasticity-only F-statistic:*

$$F = \frac{(R_{unrestricted}^2 - R_{restricted}^2)/q}{(1 - R_{unrestricted}^2)/(n - k_{unrestricted} - 1)}$$

where:

$R_{restricted}^2$  = the  $R^2$  for the restricted regression

$R_{unrestricted}^2$  = the  $R^2$  for the unrestricted regression

$q$  = the number of restrictions under the null

$k_{unrestricted}$  = the number of regressors in the unrestricted regression.

- The bigger the difference between the restricted and unrestricted  $R^2$ s – the greater the improvement in fit by adding the variables in question – the larger is the homoskedasticity-only  $F$ .

## Example:

Restricted regression:

$$\widehat{TestScore} = 644.7 - 0.671PctEL, \quad R^2_{restricted} = 0.4149$$

(1.0)    (0.032)

Unrestricted regression:

$$\widehat{TestScore} = 649.6 - 0.29STR + 3.87Expn - 0.656PctEL$$

(15.5)    (0.48)        (1.59)        (0.032)

$$R^2_{unrestricted} = 0.4366, \quad k_{unrestricted} = 3, \quad q = 2$$

So

$$F = \frac{(R^2_{unrestricted} - R^2_{restricted})/q}{(1 - R^2_{unrestricted})/(n - k_{unrestricted} - 1)}$$
$$= \frac{(.4366 - .4149)/2}{(1 - .4366)/(420 - 3 - 1)} = 8.01$$

**Note:** Heteroskedasticity-robust  $F = 5.43...$



## *The homoskedasticity-only F-statistic – summary*

$$F = \frac{(R_{unrestricted}^2 - R_{restricted}^2)/q}{(1 - R_{unrestricted}^2)/(n - k_{unrestricted} - 1)}$$

- The homoskedasticity-only  $F$ -statistic rejects when adding the two variables increased the  $R^2$  by “enough” – that is, when adding the two variables improves the fit of the regression by “enough”
- If the errors are homoskedastic, then the homoskedasticity-only  $F$ -statistic has a large-sample distribution that is  $\chi_q^2/q$ .
- But if the errors are heteroskedastic, the large-sample distribution of the homoskedasticity-only  $F$ -statistic is not  $\chi_q^2/q$

# The $F$ distribution

Your regression printouts might refer to the “ $F$ ” distribution.

If the four multiple regression LS assumptions hold *and if*:

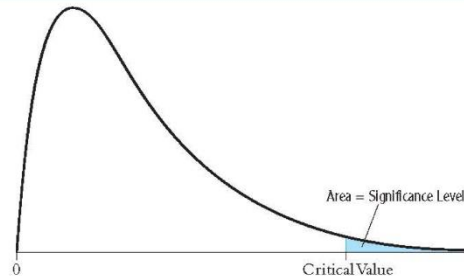
5.  $u_i$  is homoskedastic, that is,  $\text{var}(u|X_1, \dots, X_k)$  does not depend on  $X$ 's
  6.  $u_1, \dots, u_n$  are normally distributed then the homoskedasticity-only  $F$ -statistic has the “ $F_{q, n-k-1}$ ” distribution, where  $q$  = the number of restrictions and  $k$  = the number of regressors under the alternative (the unrestricted model).
- **The  $F$  distribution is to the  $\chi_q^2 / q$  distribution what the  $t_{n-1}$  distribution is to the  $N(0,1)$  distribution**

## *The $F_{q,n-k-1}$ distribution* (1 of 2)

- The  $F$  distribution is tabulated many places
- As  $n \rightarrow \infty$ , the  $F_{q,n-k-1}$  distribution asymptotes to the  $\chi_q^2/q$  distribution:
- **The  $F_{q,\infty}$  and  $\chi_q^2/q$  distributions are the same.**
- For  $q$  not too big and  $n \geq 100$ , the  $F_{q,n-k-1}$  distribution and the  $\chi_q^2/q$  distribution are essentially identical.
- Many regression packages (including STATA) compute  $p$ -values of  $F$ -statistics using the  $F$  distribution
- You will encounter the  $F$  distribution in published empirical work.

# The $F_{q,n-k-1}$ distribution (2 of 2)

TABLE 4 Critical Values for the  $F_{m,\infty}$  Distribution



Degrees of Freedom	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

This table contains the 90<sup>th</sup>, 95<sup>th</sup>, and 99<sup>th</sup> percentiles of the  $F_{m,\infty}$  distribution. These serve as critical values for tests with significance levels of 10%, 5%, and 1%.

# Testing Single Restrictions on Multiple Coefficients (SW Section 7.3) (1 of 2)

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + u_i, \quad i = 1, \dots, n$$

Consider the null and alternative hypothesis,

$$H_0: \beta_1 = \beta_2 \quad \text{vs.} \quad H_1: \beta_1 \neq \beta_2$$

This null imposes a *single* restriction ( $q = 1$ ) on *multiple* coefficients – it is not a joint hypothesis with multiple restrictions (compare with  $\beta_1 = 0$  and  $\beta_2 = 0$ ).

## *Rearrange (“transform”) the regression* (1 of 2)

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + u_i$$

$$H_0: \beta_1 = \beta_2 \quad \text{vs.} \quad H_1: \beta_1 \neq \beta_2$$

Add and subtract  $\beta_2 X_{1i}$ :

$$Y_i = \beta_0 + (\beta_1 - \beta_2) X_{1i} + \beta_2 (X_{1i} + X_{2i}) + u_i$$

or

$$Y_i = \beta_0 + \gamma_1 X_{1i} + \beta_2 W_i + u_i$$

Where

$$\gamma_1 = \beta_1 - \beta_2$$

$$W_i = X_{1i} + X_{2i}$$

## *Rearrange (“transform”) the regression* (2 of 2)

*(a) Original equation:*

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + u_i$$

$$H_0: \beta_1 = \beta_2 \quad \text{vs.} \quad H_1: \beta_1 \neq \beta_2$$

*(b) Rearranged (“transformed”) equation:*

$$Y_i = \beta_0 + \gamma_1 X_{1i} + \beta_2 W_i + u_i$$

where  $\gamma_1 = \beta_1 - \beta_2$  and  $W_i = X_{1i} + X_{2i}$

So  $H_0: \gamma_1 = 0 \quad \text{vs.} \quad H_1: \gamma_1 \neq 0$

- These two regressions ((a) and (b)) have the same  $R^2$ , the same predicted values, and the same residuals.
- The testing problem is now a simple one: test whether  $\gamma_1 = 0$  in regression (b).