# ECON 4003 Econometrics I

## Empirical Exercise 4

*By Duong Trinh*

University of Glasgow

# Picture the Scenario

- **Objective:** Predicting the US presidential elections using economic and political indicators.
- **Dataset:** `election.dta`
  - which contains 33 observations of elections from 1880 to 2008.
  - we will use data from 1916 to 2008.
- **Key variables:**
  - `vote`: Percentage share of the popular vote won by the incumbent party
  - `growth`: Growth rate of real GDP per capita in the first 3 quarters of the election year
  - `inflation`: Inflation rate in the first 15 quarters of the administration period

# Question 1

- Create a binary variable called majority such that:

$$\text{majority } = \begin{cases} 1 \text{ if votes } & > 50\% \\ 0 \text{ if votes } & \leq 50\% \end{cases}$$

- Compare the GDP growth rate per capita for those incumbents who won with majority and for those who didn't:

# Question 1

- Create a binary variable called majority such that:

$$\text{majority} = \begin{cases} 1 \text{ if votes} & > 50\% \\ 0 \text{ if votes} & \leq 50\% \end{cases}$$

- Compare the GDP growth rate per capita for those incumbents who won with majority and for those who didn't:

| | All Incumbents | With Majority Vote | Without Majority Vote |
|---|---|---|---|
| Mean | 1.4129 | 3.3817 | -1.8683 |
| *n* | 24 | 15 | 9 |

The GDP growth rate per capita in the previous 3 quarters of the elections' year was in average positive with 3.3817% for those who won with majority and negative with 1.8683% for those who didn't win with majority.

# Question 2

Consider the regression model:

$$vote_i = \beta_0 + \beta_1 growth_i + u_i$$

- (a) Point Estimate
- (b) Hypothesis Testing
- (c) Confidence Interval

# Question 2 (a) Point Estimate

- Estimate the regression model:

$$vote_i = \beta_0 + \beta_1 growth_i + u_i$$

# Question 2 (a) Point Estimate

- Estimate the regression model:

$$vote_i = \beta_0 + \beta_1 growth_i + u_i$$

Estimation results with heteroskedastic-robust standard errors:

$$\underset{(se)}{\widehat{vote}} = \underset{(1.0658)}{50.8484} + \underset{(0.1640)}{0.8859} \cdot growth \qquad R^2 = 0.5189 \quad SER = 4.798$$

- ☐ A 1 percentage point increase in the growth rate of real GDP per capita in the 3 quarters before the election *is associated with* a 0.8859 percentage points increase in the share of votes of the incumbent party *on average*.
- ☐ When real GDP growth is 0, the expected vote of the incumbent party is 50.85% (meaning that it will still maintain the majority vote).

# Question 2 (b) Hypothesis Testing

- Test the null hypothesis that economic growth has no effect on the percentage vote earned by the incumbent party at a 5% significance level.

Hypothesis

# Question 2 (b) Hypothesis Testing

Procedure includes 5 steps:

- Null hypothesis $H_0$
- Alternative hypothesis $H_1$
- Test statistic
- Decision rule
- Conclusion

# Question 2 (b) Hypothesis Testing

- Null hypothesis $H_0$:
  - $H_0$: GDP growth has no effect on the percentage vote
    $H_0 : \beta_1 = 0$
- Alternative hypothesis $H_1$
- Test statistic
- Decision rule
- Conclusion

# Question 2 (b) Hypothesis Testing

- Null hypothesis $H_0$:
  - $H_0$: GDP growth has no effect on the percentage vote
    $H_0 : \beta_1 = 0$
- Alternative hypothesis $H_1$:
  - $H_1$: GDP growth has a positive effect on the percentage of vote
    $H_1 : \beta_1 > 0$
  - We take the assumption that when the economy is doing well, the incumbent party has a better chance to win the elections.
- Test statistic
- Decision rule
- Conclusion

# Question 2 (b) Hypothesis Testing

- Null hypothesis $H_0$
- Alternative hypothesis $H_1$
- Test statistic:

$$\text{t-statistic} = \frac{\hat{\beta}_1 - 0}{SE\left(\hat{\beta}_1\right)} = \frac{0.8859 - 0}{0.1640} = 5.40$$

- Decision rule
- Conclusion

# Question 2 (b) Hypothesis Testing

- Test statistic:

$$\text{t-statistic} = \frac{\hat{\beta}_1 - 0}{SE\left(\hat{\beta}_1\right)} = \frac{0.8859 - 0}{0.1640} = 5.40$$

```
. reg vote growth, robust

Linear regression                               Number of obs     =         24
                                                F(1, 22)          =      29.19
                                                Prob > F          =     0.0000
                                                R-squared         =     0.5189
                                                Root MSE          =      4.798
```

| vote | Coef. | Robust Std. Err. | $\boxed{\text{t}}$ | P>\|t\| | [95% Conf. | Interval] |
|---|---|---|---|---|---|---|
| growth | .8859464 | .1639885 | $\boxed{5.40}$ | 0.000 | .545855 | 1.226038 |
| _cons | 50.8484 | 1.065803 | 47.71 | 0.000 | 48.63806 | 53.05874 |

# Question 2 (b) Hypothesis Testing

- Null hypothesis $H_0$
- Alternative hypothesis $H_1$
- Test statistic
- Decision rule:
  - Is this a two-sided test or an one-sided (left-tailed/right-tailed) test?
    $\Longrightarrow$ Right-tailed test (Look again $H_1 : \beta_1 > 0$).
  - What is the **significance level** $\alpha$?
    $\Longrightarrow$ 5% significance level.
  - Is the decision rule based on **critical values** or **p-value**?
    $\Longrightarrow$ Distinguish...
- Conclusion

# Question 2 (b) Hypothesis Testing

- Decision rule:
  - Right-tailed test
  - The significance level $\alpha = 0.05$
  - Is the decision rule based on **critical values** or **p-value**?

    *Approach 1: **Critical-value Test**: Reject $H_0$ if t-statistic $> t_{\alpha,n-2}$

    - The critical value is: $t_{0.05,22} = 1.7171$

```
. display invttail(22, 0.05)
1.7171444
```

    - Since $5.40 > 1.7171$, we *reject* $H_0$ at 5% level of significance.

# Question 2 (b) Hypothesis Testing

- Decision rule:
  - Right-tailed test
  - The significance level $\alpha = 0.05$
  - Is the decision rule based on **critical values** or **p-value**?

    *Approach 2: **P-value Test**: Reject $H_0$ if P-value $\leq \alpha$
    - Stata displays by default a two-sided p value:
      p-value one-sided = p-value two-sided/2 $\approx 0.000$
    - Since $0.000 < 0.05 = \alpha$, we reject $H_0$ at 5% level of significance.

```
. reg vote growth, robust

Linear regression                               Number of obs     =         24
                                                F(1, 22)          =      29.19
                                                Prob > F          =     0.0000
                                                R-squared         =     0.5189
                                                Root MSE          =      4.798
```

| vote | Coefficient | Robust std. err. | t | P>\|t\| | [95% conf. | interval] |
|---|---|---|---|---|---|---|
| growth | .8859464 | .1639885 | 5.40 | 0.000 | .545855 | 1.226038 |
| _cons | 50.8484 | 1.065803 | 47.71 | 0.000 | 48.63806 | 53.05874 |

# Question 2 (b) Hypothesis Testing

- Null hypothesis $H_0$
- Alternative hypothesis $H_1$
- Test statistic
- Decision rule
- Conclusion:
  - Real GDP growth rate per capita has a significant positive effect relation with the percentage votes of the incumbent party.

# Question 2 (c) Confidence Interval

- What is the 95% confidence interval for $\beta_1$? Interpret the result. What are the 90% and 99% confidence intervals for $\beta_1$?

CIs

# Question 2 (c) Confidence Intervals

- The 95% confidence interval for $\beta_1$ is

$$\left[\widehat{\beta}_1 - t_{\frac{0.05}{2},22} \cdot SE\left(\widehat{\beta}_1\right), \widehat{\beta}_1 + t_{\frac{0.05}{2},22} \cdot SE\left(\widehat{\beta}_1\right)\right]$$

```
. reg vote growth, robust

Linear regression                               Number of obs     =         24
                                                F(1, 22)          =      29.19
                                                Prob > F          =     0.0000
                                                R-squared         =     0.5189
                                                Root MSE          =      4.798
```

| vote | Coef. | Robust Std. Err. | t | P>\|t\| | [95% Conf. | Interval] |
|---|---|---|---|---|---|---|
| growth | .8859464 | .1639885 | 5.40 | 0.000 | .545855 | 1.226038 |
| _cons | 50.8484 | 1.065803 | 47.71 | 0.000 | 48.63806 | 53.05874 |

- By default Stata displays the 95% Confidence Interval which is $[0.5459; 1.2260]$

# Question 2 (c) Confidence Intervals

- The 95% confidence interval for $\beta_1$ is

$$[0.5459; 1.2260]$$

- In repeated samples, 95% of the intervals constructed in this way will contain the true value of parameter $\beta_1$.
- The lower bound of the 95% Confidence Interval is positive, referring that GDP growth rate and the share of votes for the incumbent party have a positive relationship.

# Question 2 (c) Confidence Intervals

- The 90% confidence interval for $\beta_1$ is

$$\left[\widehat{\beta}_1 - t_{\frac{0.1}{2},22} \cdot SE\left(\widehat{\beta}_1\right), \widehat{\beta}_1 + t_{\frac{0.1}{2},22} \cdot SE\left(\widehat{\beta}_1\right)\right]$$

```
. reg vote growth, robust level(90)

Linear regression                               Number of obs     =         24
                                                F(1, 22)          =      29.19
                                                Prob > F          =     0.0000
                                                R-squared         =     0.5189
                                                Root MSE          =      4.798
```

| vote | Coef. | Robust Std. Err. | t | P>\|t\| | [90% Conf. | Interval] |
|---|---|---|---|---|---|---|
| growth | .8859464 | .1639885 | 5.40 | 0.000 | .6043544 | 1.167538 |
| _cons | 50.8484 | 1.065803 | 47.71 | 0.000 | 49.01826 | 52.67853 |

# Question 2 (c) Confidence Intervals

- The 99% confidence interval for $\beta_1$ is

$$\left[\widehat{\beta}_1 - t_{\frac{0.01}{2},22} \cdot SE\left(\widehat{\beta}_1\right), \widehat{\beta}_1 + t_{\frac{0.01}{2},22} \cdot SE\left(\widehat{\beta}_1\right)\right]$$

```
. reg vote growth, robust level(99)

Linear regression                               Number of obs     =         24
                                                F(1, 22)          =      29.19
                                                Prob > F          =     0.0000
                                                R-squared         =     0.5189
                                                Root MSE          =      4.798

------------------------------------------------------------------------------
             |               Robust
        vote |      Coef.   Std. Err.      t    P>|t|     [99% Conf. Interval]
-------------+----------------------------------------------------------------
      growth |   .8859464   .1639885     5.40   0.000     .4237027     1.34819
       _cons |    50.8484   1.065803    47.71   0.000     47.84416    53.85264
------------------------------------------------------------------------------
```

# Question 3

Estimate the regression model:

$$vote_i = \gamma_0 + \gamma_1 inflation_i + u_i$$

- (a,b) Test the null hypothesis that inflation has no effect on the percentage vote earned by the incumbent party at a 1% significance level.
- (c) Test the null hypothesis that if inflation is zero the expected vote in favour of the incumbent party is 50% or more at a 1% significance level.

# Question 3 (c)

Hypothesis testing for the expected vote when inflation is 0

$$\mathbb{E}\left[vote_i \mid inflation_i = 0\right] = \gamma_0 + \gamma_1 \times 0 = \gamma_0$$

- Null hypothesis $H_0$:
  - $H_0$: Expected vote of the incumbent is 50% or more
    $H_0 : \gamma_0 \geq 50\%$
- Alternative hypothesis $H_1$:
  - $H_1$: Expected vote of the incumbent lower than 50%
    $H_1 : \gamma_0 < 50\%$
- Test statistic
- Decision rule
- Conclusion

# Testing Hypotheses About One of Regression Coefficients

- $\beta_1$ are **unknown** features of the population (population parameters), and we will never know them with certainty.
- Nevertheless, we can **hypothesize** about the value of $\beta_1$ and then use statistical inference to test our hypothesis.

# Testing Hypotheses About One of Regression Coefficients

Procedure includes 5 steps:

- Null hypothesis $H_0$
- Alternative hypothesis $H_1$
- Test statistic
- Decision rule
- Conclusion

# Testing Hypotheses About One of Regression Coefficients

Procedure includes 5 steps:

- Null hypothesis $H_0$:

$$H_0 : \beta_1 = \beta_{1,0}$$

  where $\beta_{1,0}$ is a hypothesized value.
- Alternative hypothesis $H_1$
- Test statistic
- Decision rule
- Conclusion

# Testing Hypotheses About One of Regression Coefficients

Procedure includes 5 steps:

- Null hypothesis $H_0$:

$$H_0 : \beta_1 = \beta_{1,0}$$

  where $\beta_{1,0}$ is a hypothesized value.

- Alternative hypothesis $H_1$:

| Test | $H_1$ |
|---|---|
| Two-sided | $\beta_1 \neq \beta_{1,0}$ |
| Left-tailed | $\beta_1 < \beta_{1,0}$ |
| Right-tailed | $\beta_1 > \beta_{1,0}$ |

- Test statistic
- Decision rule
- Conclusion

# Testing Hypotheses About One of Regression Coefficients

Procedure includes 5 steps:

- Null hypothesis $H_0$
- Alternative hypothesis $H_1$
- Test statistic:

$$\textbf{t-statistic} = \frac{\widehat{\beta}_1 - \beta_{1,0}}{SE\left(\widehat{\beta}_1\right)}$$

  follows a t-distribution with degrees of freedom $n-2$
- Decision rule
- Conclusion

# Testing Hypotheses About One of Regression Coefficients

Procedure includes 5 steps:

- Null hypothesis $H_0$
- Alternative hypothesis $H_1$
- Test statistic
- Decision rule:
  - Is this a two-sided test or an one-sided (left-tailed/right-tailed) test?
    $\Longrightarrow$ Look again $H_1$.
  - What is the **significance level** $\alpha$?
    $\Longrightarrow$ Usually chosen to be 0.01, 0.05 or 0.10.
  - Is the decision rule based on **critical values** or **p-value**?
    $\Longrightarrow$ Distinguish...
- Conclusion

# Decision Rule

Approach 1: **Critical-value Test**

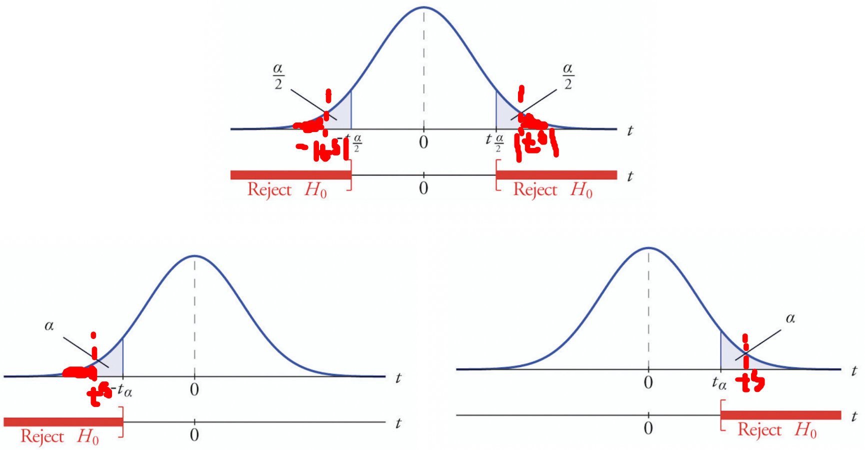
| Test | $H_1$ | Reject $H_0$ if |
|---|---|---|
| Two-sided | $\beta_1 \neq \beta_{1,0}$ | $t^s < -t_{\frac{\alpha}{2}}$ or $t^s > t_{\frac{\alpha}{2}}$ |
| Left-tailed | $\beta_1 < \beta_{1,0}$ | $t^s < -t_\alpha$ |
| Right-tailed | $\beta_1 > \beta_{1,0}$ | $t^s > t_\alpha$ |

Approach 2: **p-value Test**

| Test | $H_1$ | p-value | Reject $H_0$ if |
|---|---|---|---|
| Two-sided | $\beta_1 \neq \beta_{1,0}$ | sum probabilities to the right of $\lvert t^s \rvert$ and to the left of $-\lvert t^s \rvert$ | p-value $\leq \alpha$ |
| Left-tailed | $\beta_1 < \beta_{1,0}$ | probability to the left of $t^s$ | p-value $\leq \alpha$ |
| Right-tailed | $\beta_1 > \beta_{1,0}$ | probability to the right of $t^s$ | p-value $\leq \alpha$ |

*Note: p-value two-sided $=2\times$ p-value one-sided

# Decision Rule



# Testing Hypotheses About One of Regression Coefficients

Procedure includes 5 steps:

- Null hypothesis $H_0$
- Alternative hypothesis $H_1$
- Test statistic
- Decision rule
- Conclusion:
  - Do you *reject* or or *fail to reject* the null hypothesis at the significance level $\alpha$?
  - AVOID saying that you "*accept*" the null hypothesis, which can be very misleading

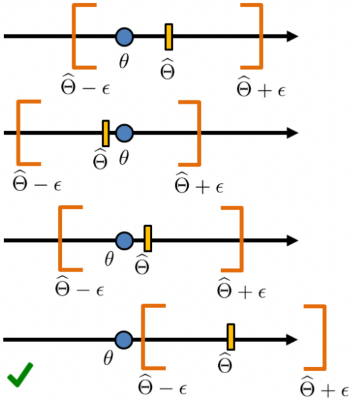
# Confidence Intervals

- The $100(1-\alpha)\%$ confidence interval for $\beta_1$ is given by

$$\left[\widehat{\beta}_1 - t_{\frac{\alpha}{2},n-2} \cdot SE\left(\widehat{\beta}_1\right), \widehat{\beta}_1 + t_{\frac{\alpha}{2},n-2} \cdot SE\left(\widehat{\beta}_1\right)\right]$$

  - Usually $\alpha = 0.01, 0.05$ or $0.10$, so that we obtain a 99%, 95% or 90% confidence interval, respectively.
  - $t_{\frac{\alpha}{2},n-2}$ : same critical value as two-sided hypothesis test.

# Confidence Intervals




If you have 100 random realizations of the confidence intervals, then 95 on average will include the true parameter.