Final Review

EC 320: Introduction to Econometrics

Kyle Raze
Fall 2019

Prologue

Housekeeping

Final Exam

Tuesday, December 10 at 10:15am in Chapman 220.

- Room open at 10:00am.

Problem Set 5

Due Saturday, December 7 by 11:59pm.

Housekeeping

Lab

Additional review session.

- Focus: coefficient interpretation.

Office hours

Regular hours this week.

Next week: extended office hours on Monday.

- Times and locations TBA.

Final Exam

Final Exam Topics

Anything from the lectures, labs, or problem sets **is fair game!**

1. Midterm Topics
2. Multiple Linear Regression: Inference
3. Categorical Variables
4. Interactive Relationships
5. Nonlinear Relationships

Final Exam Topics

1. Midterm Topics

Making predictions using fitted regression model

- *e.g.*, using a Hedonic model

Goodness of fit

Hypothesis testing

Omitted-variable bias

- Know when omitting a variable causes bias
- Sign the bias
- Back out correlations between explanatory variables
- Provide examples of problematic omitted variables

Final Exam Topics

2. Multiple Linear Regression: Inference

F tests (multiple parameters)

- State null hypothesis
- Identify restricted and unrestricted models
- Calculate *F* statistic
- Use table to find F_{crit}
- $F > F_{crit}$?
- State conclusion of the test

t tests (single parameter)

Q: Which test should you choose?

A: Depends on the null hypothesis!

Final Exam Topics

2. Multiple Linear Regression: Inference

Confidence intervals

- Formula, interpretation, and comparison of different intervals for the same coefficient

Proof: Show that the F statistic formula containing RSS implies the F statistic formula containing R^2

- For practice, you can also prove that the second formula implies the first

Final Exam Topics

3. Categorical Variables

How do you interpret coefficients on binary variables?

- **Note:** Depends on the presence of interaction terms and whether the outcome variable is transformed

Dummy variable trap

What is the reference category?

- How do you back out group-specific averages from a dummy variable regression?
- How do coefficient estimates change when you change the reference category?

Final Exam Topics

4. Interactive Relationships

How do you interpret interaction coefficients?

- Binary $\times$ binary
- Binary $\times$ continuous
- Continuous $\times$ continuous

How does an interaction term change how you interpret the effect of the variable of interest on the outcome variable?

- Marginal effects (partial derivative)

Final Exam Topics

5. Nonlinear Relationships

Identify nonlinear models

- OLS can handle nonlinear variables, but not nonlinear parameters

Transform nonlinear models

- Give OLS a chance

Final Exam Topics

5. Nonlinear Relationships

How do you interpret coefficients in the presence of logarithmic transformations?

- Level Y, level X
- Level Y, log X
- Log Y, level X
- Log Y, log X

Quadratic models: interacting X with itself

- Calculate marginal effects to understand how X affects Y

Final Exam Structure

No Multiple Choice!

No Fill-in the Blank!

No True-False!

Just Free Response

- 8 multi-part questions with varying numbers of points (200 points total)
- Explanations required for full credit

Final Exam Protocol

Materials

- Writing utensil
- 3-inch-by-5-inch note card
- Basic or scientific calculator (no graphing or programming capabilities)
- **Nothing else**

Procedure

- **Randomized** seating chart (penalty for non-compliance)
- 120 minutes from *"you may begin"* to *"pencils down"*
- First 30 minutes: **quiet period** (no questions, no getting up)
- Last 90 minutes: ask lots of questions

Practice

gapminder Package

Data on population, GDP per capita, and life expectancy

Unit of observation: country-year

- All countries, every 5th year between 1957 and 2007

```
p_load(gapminder)
data ← get('gapminder')
head(data)
```

```
#> # A tibble: 6 x 6
#>   country     continent  year lifeExp      pop gdpPercap
#>   <fct>       <fct>     <int>   <dbl>    <int>     <dbl>
#> 1 Afghanistan Asia       1952    28.8  8425333      779.
#> 2 Afghanistan Asia       1957    30.3  9240934      821.
#> 3 Afghanistan Asia       1962    32.0 10267083      853.
#> 4 Afghanistan Asia       1967    34.0 11537966      836.
#> 5 Afghanistan Asia       1972    36.1 13079460      740.
#> 6 Afghanistan Asia       1977    38.4 14880372      786.
```

GDP per Capita by Continent

```
reg ← lm(gdpPercap ~ continent, data = data)
tidy(reg)
```

```
#> # A tibble: 5 x 5
#>   term              estimate std.error statistic  p.value
#>   <chr>                <dbl>     <dbl>     <dbl>    <dbl>
#> 1 (Intercept)          2194.      347.      6.33 3.21e-10
#> 2 continentAmericas    4942.      609.      8.12 8.79e-16
#> 3 continentAsia        5708.      557.     10.3  5.43e-24
#> 4 continentEurope     12276.      573.     21.4  3.06e-90
#> 5 continentOceania    16428.     1802.      9.12 2.12e-19
```

```
nobs(reg)
```

```
#> [1] 1704
```

```
summary(reg)$r.squared
```

```
#> [1] 0.2295766
```

GDP per Capita by Continent

```
reg2 ← lm(log(gdpPercap) ~ continent, data = data)
tidy(reg2)
```

```
#> # A tibble: 5 x 5
#>   term              estimate std.error statistic   p.value
#>   <chr>                <dbl>     <dbl>     <dbl>     <dbl>
#> 1 (Intercept)          7.25     0.0373     195.  0.
#> 2 continentAmericas    1.37     0.0654      20.9 1.48e- 86
#> 3 continentAsia        0.823    0.0598      13.8 5.59e- 41
#> 4 continentEurope      2.10     0.0616      34.0 5.40e-194
#> 5 continentOceania     2.53     0.194       13.1 2.93e- 37
```

```
nobs(reg2)
```

```
#> [1] 1704
```

```
summary(reg2)$r.squared
```

```
#> [1] 0.4390461
```

GDP per Capita by Continent

```
data ← data %>%
  mutate(continent = ifelse(as.character(continent) %in%
                              c("Europe", "Americas", "Oceania"),
                            "The West",
                            as.character(continent)))
reg3 ← lm(log(gdpPercap) ~ continent, data = data)
tidy(reg3)
```

```
#> # A tibble: 3 x 5
#>   term              estimate std.error statistic   p.value
#>   <chr>                <dbl>     <dbl>     <dbl>     <dbl>
#> 1 (Intercept)          7.25     0.0385     188.  0.
#> 2 continentAsia        0.823    0.0618      13.3 1.24e- 38
#> 3 continentThe West    1.79     0.0532      33.7 8.19e-191
```

```
summary(reg3)$r.squared
```

```
#> [1] 0.4007156
```

GDP per Capita by Continent

Null hypothesis: Europe, the Americas, and Oceania have the same average GDP per capita.

Setup *F* Test

```
# restricted model: reg3
# unrestricted model: reg2

# get r-squared from restricted and unrestricted models
r2_restrict ← summary(reg3)$r.squared
r2_unrestrict ← summary(reg2)$r.squared

# number of resrictions
n_restrict ← 2

# degrees of freedom in unrestricted model
dof ← summary(reg3)$df[2]

# significance level
alpha ← 0.01
```

GDP per Capita by Continent

Null hypothesis: Europe, the Americas, and Oceania have the same average GDP per capita.

Conduct *F* Test

```
# calculate F stat:
f_stat ← ((r2_unrestrict - r2_restrict)/n_restrict)/((1 - r2_unrestrict)/dof)

# find critical value of F:
f_crit ← qf(1-alpha, n_restrict, dof)

# reject null if:
f_stat > f_crit
```

```
#> [1] TRUE
```

Conclusion: At least one of the continents in "The West" has a different average GDP per capita than the others.

Association of GDP per Capita with Life Expectancy

	Life Expectancy	Life Expectancy	log(Life Expectancy)	log(Life Expectancy)
log(GDP/Capita)	**6.42**	5.69	0.112	0.111
	(0.183)	(0.325)	(0.004)	(0.006)
log(GDP/Capita) Americas		4.03		0.047
		(0.674)		(0.013)
log(GDP/Capita) Asia		0.561		-0.002
		(0.421)		(0.008)
log(GDP/Capita) Europe		0.614		-0.019
		(0.597)		(0.012)
log(GDP/Capita) Oceania		5.2		0.035
		(4.36)		(0.084)
Continent Dummies?	**Yes**	Yes	Yes	Yes
Observations	**1704**	1704	1704	1704
R-Squared	**0.704**	0.71	0.665	0.669

Association of GDP per Capita with Life Expectancy

	Life Expectancy	Life Expectancy	log(Life Expectancy)	log(Life Expectancy)
log(GDP/Capita)	6.42	**5.69**	0.112	0.111
	(0.183)	**(0.325)**	(0.004)	(0.006)
log(GDP/Capita) Americas		**4.03**		0.047
		(0.674)		(0.013)
log(GDP/Capita) Asia		**0.561**		-0.002
		(0.421)		(0.008)
log(GDP/Capita) Europe		**0.614**		-0.019
		(0.597)		(0.012)
log(GDP/Capita) Oceania		**5.2**		0.035
		(4.36)		(0.084)
Continent Dummies?	Yes	**Yes**	Yes	Yes
Observations	1704	**1704**	1704	1704
R-Squared	0.704	**0.71**	0.665	0.669

Association of GDP per Capita with Life Expectancy

	Life Expectancy	Life Expectancy	log(Life Expectancy)	log(Life Expectancy)
log(GDP/Capita)	6.42	5.69	**0.112**	0.111
	(0.183)	(0.325)	**(0.004)**	(0.006)
log(GDP/Capita) Americas		4.03		0.047
		(0.674)		(0.013)
log(GDP/Capita) Asia		0.561		-0.002
		(0.421)		(0.008)
log(GDP/Capita) Europe		0.614		-0.019
		(0.597)		(0.012)
log(GDP/Capita) Oceania		5.2		0.035
		(4.36)		(0.084)
Continent Dummies?	Yes	Yes	**Yes**	Yes
Observations	1704	1704	**1704**	1704
R-Squared	0.704	0.71	**0.665**	0.669

Association of GDP per Capita with Life Expectancy

	Life Expectancy	Life Expectancy	log(Life Expectancy)	log(Life Expectancy)
log(GDP/Capita)	6.42	5.69	0.112	**0.111**
	(0.183)	(0.325)	(0.004)	**(0.006)**
log(GDP/Capita) Americas		4.03		**0.047**
		(0.674)		**(0.013)**
log(GDP/Capita) Asia		0.561		**-0.002**
		(0.421)		**(0.008)**
log(GDP/Capita) Europe		0.614		**-0.019**
		(0.597)		**(0.012)**
log(GDP/Capita) Oceania		5.2		**0.035**
		(4.36)		**(0.084)**
Continent Dummies?	Yes	Yes	Yes	**Yes**
Observations	1704	1704	1704	**1704**
R-Squared	0.704	0.71	0.665	**0.669**