Applied Regression (II)

Francisco Villamil

Applied Quantitative Methods II

IC3JM, Spring 2026

Today's goals

- Learn to model conditional relationships (interactions)
- Understand non-linear relationships
- Review hypothesis testing and confidence intervals
- Discuss how to present and interpret results

Roadmap

Interaction Effects

When effects depend on context

- Sometimes, the effect of X on Y depends on another variable Z
- Examples:
  - → Effect of education on income may differ by gender
  - → Effect of campaign spending may differ by incumbency status
  - → Effect of democracy on growth may depend on economic development
- We model this with **interaction terms**

The interaction model

$$Y = \beta_0 + \beta_1 X + \beta_2 Z + \beta_3 (X \times Z) + \varepsilon$$

- β_1: effect of X when $Z = 0$
- β_2: effect of Z when $X = 0$
- β_3: how the effect of X changes as Z increases

The marginal effect of X

$$\frac{\partial Y}{\partial X} = \beta_1 + \beta_3 Z$$

- The effect of X is no longer a single number
- It's a **function** of Z
- Need to report effects at meaningful values of Z

Common mistakes with interactions

- **Mistake 1**: Interpreting β_1 as "the effect of X"
  - $\rightarrow$ It's only the effect when $Z = 0$
  - $\rightarrow$ May not even be meaningful!
- **Mistake 2**: Omitting constitutive terms
  - $\rightarrow$ Always include X and Z separately, not just $X \times Z$
- **Mistake 3**: Not showing how the effect varies
  - $\rightarrow$ Plot the marginal effect across values of Z

Best practice: Visualize interactions

- Tables of coefficients are hard to interpret
- Better approach:
  - → Plot predicted values of Y for different combinations of X and Z
  - → Plot the marginal effect of X across values of Z
  - → Include confidence intervals
- Packages like `marginaleffects` make this easy in R

Roadmap

Non-linear Relationships

When linearity is not enough

- Linear regression assumes a constant slope
- But relationships are often curved
- Examples:
  - → Diminishing returns (logarithmic)
  - → U-shaped or inverted U-shaped relationships
  - → Threshold effects

Polynomial terms

$$Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \varepsilon$$

- Adding X^2 allows the slope to change with X
- Positive β_2: U-shaped relationship
- Negative β_2: inverted U-shape

- Marginal effect: $\frac{\partial Y}{\partial X} = \beta_1 + 2\beta_2 X$

Logarithmic transformations

- **Log-linear**: $Y = \beta_0 + \beta_1 \log(X)$
  → Effect of X decreases as X increases (diminishing returns)

- **Linear-log**: $\log(Y) = \beta_0 + \beta_1 X$
  → A one-unit change in X changes Y by $(\exp(\beta_1) - 1) \times 100\%$

- **Log-log**: $\log(Y) = \beta_0 + \beta_1 \log(X)$
  → β_1 is an elasticity: 1% change in $X \rightarrow \beta_1$% change in Y

When to use transformations

- Logarithms are common for:
  - → Variables with skewed distributions (income, GDP, population)
  - → Multiplicative relationships
- Polynomials are useful for:
  - → Relationships that curve in predictable ways
  - → Capturing turning points
- Always check if the transformation makes sense substantively

Roadmap

Inference: Standard Errors and Hypothesis Testing

What are standard errors?

- Our estimates ($\hat{\beta}$) are based on a sample
- A different sample would give different estimates

- The **standard error** measures this variability
- It tells us how precise our estimate is

- Smaller SE = more confidence in our estimate

What affects standard errors?

- **Sample size**: More data $\rightarrow$ smaller SE
- **Variation in** X: More spread in X $\rightarrow$ smaller SE
- **Error variance**: More noise in Y $\rightarrow$ larger SE

- Intuition: We learn more from larger, more varied samples with less noise

Confidence intervals

$$95\% \text{ CI} = \hat{\beta} \pm 1.96 \times SE(\hat{\beta})$$

- Interpretation: If we repeated the study many times...
- ...95% of the CIs would contain the true β

- **Not**: "There's a 95% probability β is in this interval"

Hypothesis testing

- Null hypothesis: $H_0 : \beta = 0$ (no effect)
- Alternative: $H_1 : \beta \neq 0$ (there is an effect)

- Test statistic: $t = \frac{\hat{\beta}}{SE(\hat{\beta})}$

- p-value: Probability of seeing this (or more extreme) result if H_0 is true

What p-values tell us (and don't)

- p-value **is**: probability of data given null hypothesis
- p-value **is not**: probability that the null is true

- Small p-value: data is surprising if H_0 is true
- Large p-value: data is compatible with H_0

- But H_0 compatible doesn't mean H_0 is true!

Statistical vs. practical significance

- Statistical significance: "Is the effect different from zero?"
- Practical significance: "Is the effect large enough to matter?"

- With enough data, tiny effects become "significant"
- A large sample can detect effects too small to care about

- Always consider the **size** of the effect, not just significance

Heteroskedasticity and robust standard errors

- OLS assumes constant error variance (homoskedasticity)
- Often violated in practice

- Consequence: Standard errors may be wrong
- Coefficients are still unbiased, but inference is off

- Solution: Use **robust standard errors** (HC1, HC2, etc.)
- Always report robust SEs unless you have good reason not to

Roadmap

Presenting Results

Principles for presenting regression results

- Focus on **substance**, not just coefficients
- Show the **uncertainty** in your estimates
- Use **visualizations** when possible
- Help readers understand the **size** of effects

Moving beyond coefficient tables

- Tables are useful but hard to interpret
- Better approaches:
  → Predicted values at meaningful scenarios
  → Marginal effects with confidence intervals
  → First differences (effect of moving from low to high)
- The `marginaleffects` package in R is excellent for this

Quantities of interest

- **Predicted values**: $\hat{Y}$ at specific X values
  - → "People with college education earn, on average, $X"
- **Marginal effects**: Change in $\hat{Y}$ for change in X
  - → "Each additional year of education increases income by $Y"
- **First differences**: $\hat{Y}|_{X_{high}} - \hat{Y}|_{X_{low}}$
  - → "Going from high school to college increases income by $Z"

Simulation for uncertainty

- Coefficients have a sampling distribution
- We can simulate from this distribution
- Steps:
  - → Draw from $\tilde{\beta} \sim N(\hat{\beta}, V(\hat{\beta}))$
  - → Calculate quantity of interest for each draw
  - → Report mean and percentiles (95% CI)
- Propagates uncertainty properly through calculations

Summary: Key takeaways

- Interactions model conditional relationships
- Non-linear terms allow curved relationships
- Standard errors measure estimation precision
- p-values are about data, not hypothesis probability
- Present results as interpretable quantities, not just coefficients

For next week

- Read King et al. (2000), "Making the most of statistical analyses"
- Read Brambor et al. (2006), "Understanding Interaction Models"
- Check out marginaleffects.com
- Complete Problem Set 2

- Next session: Model interpretation and diagnostics

Questions?