Post-estimation Simulation

POSC 3410 – Quantitative Methods in Political Science

Steven V. Miller

Department of Political Science

Goal for Today

Provide intuitive quantities of interest from your regression.

```r
library(tidyverse)    # for most things
library(stevemisc)    # for some various helper functions
library(modelr)       # for generating hypothetical data
library(modelsummary) # for regression tables
library(kableExtra)   # for some table formatting
```

Readable Regression Tables

Remember: your analysis should be as easily interpretable as possible.

- I should get a preliminary glimpse of effect size from a regression.
- Your *y*-intercept should be meaningful.

Standardizing variables helps.

- Creates meaningful zeroes (i.e. the mean).
- Coefficients communicate magnitude changes in *x*.
- Standardizing by two SDs allows for easy comparison with binary predictors.

Satisfy Your Audience

You need to relate your analysis to both me and your grandma.

- I will obviously know/care more about technical details.
- Grandma may not, but she may be a more important audience than me.

Her inquiries are likely be understandable. Examples:

- What's the expected tolerance of abortion for a religious Democrat?
- What's the increased probability of voting for a Republican for an increase of $20k in yearly income?

These are perfectly reasonable questions to ask of your analysis.

- If your presentation isn't prepared to answer her questions, you're not doing your job.

Statistical Presentations

Statistical presentations should:

1. Convey precise estimates of quantities of interest.
2. Include reasonable estimates of *uncertainty* around those estimates.
3. Require little specialized knowledge to understand Nos. 1 and 2.
4. Not bombard the audience with superfluous information.

We will do this with post-estimation simulation using draws from a multivariate normal distribution (King et al. 2000).

Estimating Uncertainty with Simulation

Any statistical model has a stochastic and systematic component.

- **Stochastic**: $Y_i \sim f(y_i \,|\, \theta_i, \alpha)$
- **Systematic**: $\theta_i = g(x_i, \beta)$

For a simple OLS model (i.e. a linear regression):

$$\begin{aligned} Y_i &= N(\mu_i, \, \sigma^2) \\ \mu_i &= X_i \beta \end{aligned}$$

Understanding our Uncertainty

We have two types of uncertainty.

1. **Estimation uncertainty**
    - Represents systematic components; can be reduced by increasing sample size.
2. **Fundamental uncertainty**
    - Represents stochastic component; exists no matter what (but can be modeled).

Getting our Parameter Vector

We want a **simulated parameter vector**, denoted as:

$$\hat{\gamma} \sim vec(\hat{\beta}, \hat{\alpha})$$

Central limit theorem says with a large enough sample and bounded variance:

$$\tilde{\gamma} \sim N(\hat{\gamma}, \hat{V}(\hat{\gamma}))$$

In other words: distribution of quantities of interest will follow a multivariate normal distribution with mean equal to $\hat{\gamma}$, the simulated parameter vector.

Another Way of Thinking About This

Subject to approximate regularity conditions and sample size:

- the conditional distribution of a quantity of interest, given the observed data,
- can be approximated with a multivariate normal distribution with
- parameters (coefficients, var-cov matrix) derived from the regression model.

Gelman and Hill (2007) call this a “pseudo-Bayesian” approach.

- i.e. there are no prior assumptions of model parameters
- this approach papers over it because dependence on priors fades through large samples.

Getting our Quantities of Interest

This is a mouthful! Let's break the process down step-by-step.

1. Run your regression. Look at your results.
2. Choose values of explanatory variable (as you see fit).
3. Grab the coefficient matrix and variance-covariance matrix.
4. Simulate outcomes by taking random draw from a multivariate normal distribution with those parameters.

Do this *m* times (typically *m* = 1000) to estimate full distribution of Y_c.

A Brief Introduction of Terms

Multivariate normal distribution: a generalization of the normal distribution to higher dimensions.

- every linear combination of its components has a normal distribution.
- used to describe correlated random variables each of which clusters around a mean value.

Variance-covariance matrix: a square matrix used for generating standard errors in regression

- I'll spare you the matrix algebra here.

Does Prejudice Decrease Support for (American) Democracy?

Prejudice and Support for Democracy

Miller and Davis (2021) argue that “white social prejudice” decreases support for democracy in the U.S.

- Prejudice is a rejection of the equal status of outgroups.
- Democracy is a representation of multiculturalism, broadly understood (i.e. pluralism).
- Democracy enfranchises groups of people that prejudice don’t want empowered.

We expect white Americans who score high in prejudice are less receptive to democracy than those who score lower.

Data (Voter Study Group, Nov. 2019)

DVs: how good/bad is it for the U.S. to have...

- a strong leader who does not have to deal with Congress or elections
- the army rule the government
- a democratic political system

All three are on 1-4 scale, collapsed to binary 0/1.

- *Model*: logistic regression on people who self-identify as white.

Data (Voter Study Group, Nov. 2019)

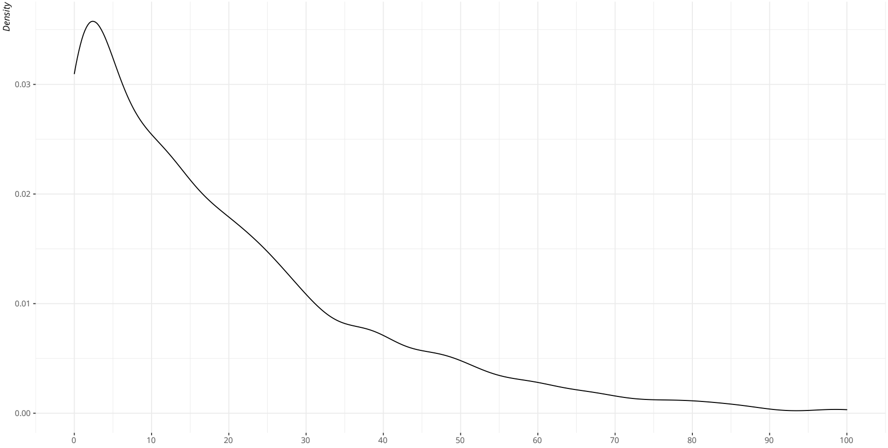
IVs: a thermometer based approach to prejudice

- i.e. therm. rating for white people - mean therm. rating for black people, Latinos, Asians, and Muslims
- Scale ranges from 0 (no social prejudice) to 100 (maximum social prejudice)

Controls: respondent's age, sex (female), college education, ideology, unemployment, partisanship

A Density Plot of White Social Prejudice

The data approximate an exponential distribution, with a noticeable right skew.



Data: Voter Study Group (Nov. 2019)

Table 1: The Effect of White Social Prejudice on Support for Democracy

	Strong Leader	**Army Rule**	**Oppose Democracy**
White Social Prejudice	1.298***	1.194***	0.903***
	(0.138)	(0.164)	(0.170)
Age	-0.860***	-1.446***	-0.936***
	(0.126)	(0.156)	(0.158)
Female	0.775***	1.013***	0.638***
	(0.121)	(0.158)	(0.154)
College Education	-0.804***	-0.747***	-0.741***
	(0.141)	(0.178)	(0.185)
Ideology (L to C)	0.077	0.327	0.353
	(0.178)	(0.218)	(0.229)
Unemployed	0.003	0.191	0.479
	(0.316)	(0.348)	(0.350)
Party ID (D to R)	0.431*	-0.080	1.111***
	(0.169)	(0.211)	(0.239)
Num.Obs.	2607	2608	2606

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Something Nana Might Ask

What is the probability of supporting a strong leader/army rule/opposing democracy by certain values of white social prejudice?

- Let's see!

```r
# find the z_prejudice values corresponding with raw values we want
Data %>% filter(prejudice %in% c(0, 25, 50, 100)) %>%
  distinct(z_prejudice) %>%
  arrange(z_prejudice) %>%
  pull(z_prejudice) -> z_prejudice_vals

# generating hypothetical data, everything else at typical value
# except for the prejudice values
Data %>%
  data_grid(.model = M1, z_prejudice = z_prejudice_vals,
            strongleaderd = 0, armyruled = 0, havedemd = 0) -> newdat_prejudice
```

```r
Data %>%
  filter(prejudice %in% c(0, 25, 50, 100)) %>%
  distinct(prejudice, z_prejudice) %>%
  arrange(prejudice)
```

```
## # A tibble: 4 x 2
##   prejudice z_prejudice
##       <dbl>       <dbl>
## 1         0      -0.181
## 2        25       0.323
## 3        50       0.827
## 4       100       1.84
```

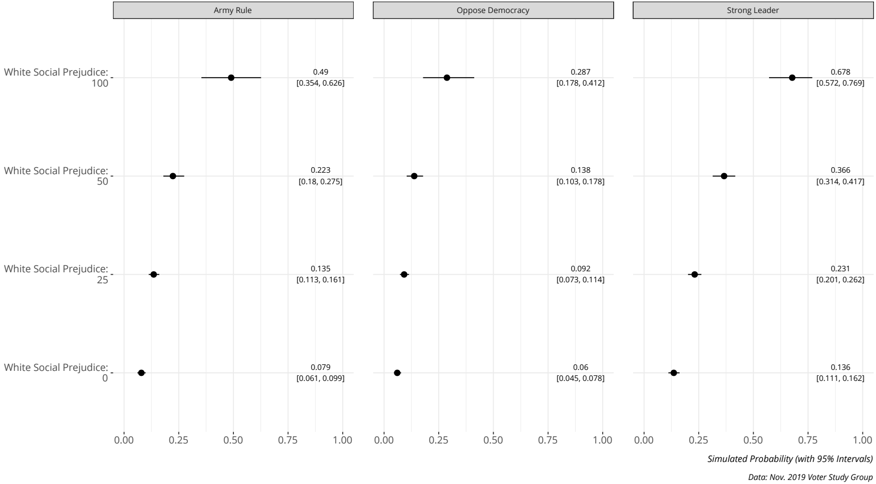
```
newdat_prejudice %>%
  select(-strongleaderd:-havedemd)
```

```
## # A tibble: 4 x 7
##   z_prejudice  z_age female collegeed  z_ideo unemployed z_pid7
##         <dbl>  <dbl>  <dbl>     <dbl>   <dbl>      <dbl>  <dbl>
## 1      -0.181 0.0727      1         0 -0.0376          0 0.0217
## 2       0.323 0.0727      1         0 -0.0376          0 0.0217
## 3       0.827 0.0727      1         0 -0.0376          0 0.0217
## 4       1.84  0.0727      1         0 -0.0376          0 0.0217
```

```r
Sims <- list() # Store in a list

# For each model, run 1,000 simulations for these hypothetical data
# with reproducible seed
Sims[[1]] <- get_sims(M1, newdat_prejudice, 1000, 8675309) %>%
  mutate(cat = "Strong Leader", prejudice = rep(c(0, 25, 50, 100), 1000))

Sims[[2]] <- get_sims(M2, newdat_prejudice, 1000, 8675309) %>%
  mutate(cat = "Army Rule", prejudice = rep(c(0, 25, 50, 100), 1000))

Sims[[3]] <- get_sims(M3, newdat_prejudice, 1000, 8675309) %>%
  mutate(cat = "Oppose Democracy", prejudice = rep(c(0, 25, 50, 100), 1000))
```

The Effect of White Social Prejudice on Attitudes about Democracy, Nov. 2019

In all cases, a min-to-max effect of white social prejudice results in a percentage change in the probability of an anti-democratic attitude by at least 200%.



Conclusion

Regression provides all-else-equal effect sizes across the range of the data.

- You can extract meaningful quantities of interest from regression output itself.
- Typically, you'll need more to answer substantive questions and provide meaningful quantities of interest.

Post-estimation simulation from a multivariate normal distribution does this.

- When you start doing this yourselves, be prepared to provide quantities of interest for your audience.
- Never forget: *you're trying to tell a story.* Tell it well.

Table of Contents