

ITAM
Departamento de Ciencia Política

Metodología Política Avanzada

Primavera 2020
Lunes 18:00-21:00
Río Hondo, Salón CC302

Adrián Lucardi
adrian.lucardi@itam.mx

Oficina: Martes 15:00-16:30 y Jueves 16:00-17:30 o previa cita

En este curso veremos los rudimentos del análisis cuantitativo en ciencias sociales, poniendo especial énfasis en (a) la *manipulación* y *visualización* de datos; (b) la *especificación* de modelos estadísticos; y (c) la *interpretación* de dichos modelos.

El curso parte de la idea de que la estadística no es una técnica ni una metodología sino un lenguaje, y la única forma de aprender un lenguaje es practicándolo. Por lo tanto, para aprobar el curso deberás completar numerosos ejercicios prácticos *fuera del horario de clases*. Además, tanto el material teórico como los ejercicios de manipulación y análisis de datos son acumulativos, por lo que es indispensable que te mantengas al día con las tareas.

Para tomar el curso no necesitas estar familiarizado con ningún lenguaje de programación. Sin embargo, en la presentación del material teórico daré por sentado que conoces los rudimentos de la teoría de probabilidades y el modelo de regresión por mínimos cuadrados ordinarios. A lo largo del curso discutiremos cuándo es posible interpretar los resultados de los modelos en términos causales, pero el foco estará puesto en la especificación e interpretación de modelos estadísticos, no en la inferencia causal como tal.

Calificación

- **Participación en clase (10%).** Debes leer la bibliografía obligatoria y participar activamente en las discusiones.
- **Tareas (33%).** A lo largo de la clase habrá un total de seis tareas. Puedes responderlas de forma individual o en conjunto con *un/a único/a compañero/a*, cuyo nombre deberá estar claramente especificado en la hoja de respuesta. La fecha límite para entregar las tareas es la indicada en este temario, *antes* del comienzo de la clase. Cada tarea será calificada como (i) entregada completa y aprobada (5.5% de la calificación final); (ii) entregada incompleta y/o deficiente (2%); o (iii) no entregada (0%).

- **Parciales (28%).** El 23 de marzo y el 11 de mayo habrá exámenes parciales. Ambos serán en clase, a libro abierto pero a computadora cerrada. El segundo parcial incluirá preguntas sobre lo visto en el primero. Cada parcial contará el 14% de la calificación final.
- **Trabajo final (29%).** Debes replicar los resultados –tablas y gráficos– de un artículo publicado en una revista académica de ciencias sociales. La calificación del trabajo final se distribuirá de la siguiente manera:

– *Selección.* A más tardar el 20 de abril deberás enviar por mail (a) el artículo que quieres replicar; (b) un texto de una página indicando qué tablas y figuras planeas replicar, y qué análisis adicional te gustaría llevar a cabo; y (c) la base de datos y/o el link donde descargar éstos. Puedes proponer más de una opción si lo deseas.

La replicación *puede* ser en equipo, sujeta a las siguientes restricciones:

- i. Ningún equipo puede tener más de dos integrantes;
- ii. El proceso de selección es individual: cada miembro debe proponer al menos un artículo; y
- iii. Las replications grupales deben ser más extensas –en el sentido de cubrir más figuras y tablas del artículo original– que las individuales.

Enviar tu selección a tiempo contará por el 3% de la calificación final.

– *Replicación.* En fecha a definir durante la semana de finales, cada estudiante o equipo debe enviar (a) un reporte detallando los resultados de la replicación; y (b) el código para reproducir los mismos. El reporte contará el 26% de la calificación final.

Lecturas

El foco de la clase estará puesto en cuestiones prácticas, no teóricas. Si deseas profundizar (algunas) cuestiones teóricas, puedes consultar los siguientes textos, ordenados de menor a mayor grado de dificultad:

- Joshua D. Angrist and Jörn-Steffen Pischke. *Mastering 'Metrics. The Path from Cause to Effect.* Princeton University Press, Princeton, 2015.
- Kosuke Imai. *Quantitative Social Science: An Introduction.* Princeton University Press, Princeton and Oxford, 2018.
- Peter Kennedy. *A Guide to Econometrics.* Wiley-Blackwell, Malden, MA, 6th edition, 2008.
- John Fox. *Applied Regression Analysis and Generalized Linear Models.* Sage, Los Angeles, 3rd edition, 2015.
- Joshua Angrist and Jörn-Steffen Pischke. *Mostly Harmless Econometrics. An Empiricist's Companion.* Princeton University Press, Princeton, 2009.

- Jeffrey M. Wooldridge. *Introductory Econometrics: A Modern Approach*. 5th edition, 2012.

Todas las lecturas se encuentran disponibles en [Google Drive](#).

Software

Junto con las tareas y el trabajo final, deberás adjuntar el correspondiente código de R ó Stata. No aceptaré código en otros lenguajes de programación. No obstante, ten en cuenta que todos los ejemplos de código que distribuiré a lo largo del curso serán en R; por lo que a menos que ya estés familiarizado con Stata, te recomiendo hacer tus tareas en R.

Puedes descargar R de manera gratuita del siguiente link: <https://cran.itam.mx/>. Te recomiendo usarlo en conjunto con RStudio (<https://www.rstudio.com/>). Si quieres usar Stata, tendrás que averiguar cómo instalarlo en tu computadora, o bien trabajar en las salas de cómputo del ITAM.

La manera más efectiva de aprender a programar es buscando código de otros autores en internet y adaptándolo a tus propias necesidades. Si eres nuevo en R, alguno de los siguientes recursos te pueden resultar útiles:

- Kelsey Moty, *workshop* de verano de NYU: <https://nyu-cdsc.github.io/learningr/>.
- Hadley Wickham and Garrett Grolemund. *R for Data Science. Visualize, Model, Transform, Tidy and Import Data*. O'Reilly, Sebastopol (CA), 2017: <https://r4ds.had.co.nz/>.

Integridad académica

Según el artículo 5 bis del Reglamento de Titulación del ITAM, “la reproducción parcial o total de textos o productos ajenos, de autor conocido o anónimo, publicados por cualquier medio o inéditos, sin entrecomillar los pasajes o elementos reproducidos o sin hacer la referencia bibliográfica pertinente,” constituye plagio o fraude académico. Las citas a otros autores son bienvenidas, pero deben ser reconocidas como tales. El uso de fragmentos y/o traducciones literales de textos ajenos sin la debida atribución será causal para reprobar la materia y merecerá la amonestación oficial al alumno por parte de la Dirección Escolar del Instituto.

Calendario

Sesión 1 (Enero/20). Introducción

- Presentación del curso.
- *Teoría*: Tipos de datos
 - *Lectura optativa*: Wooldridge, sec. 1.3.

- *Práctica*: introducción a R
 - Kieran Healy. *Data Visualization: A Practical Introduction*. Princeton University Press, Princeton and Oxford, 2018, secs. 2.2-2.7: <http://socviz.co/gettingstarted.html>
 - Kelsey Moty, *workshop NYU*: <https://nyu-cdsc.github.io/learningr/>. Lecturas y ejercicios de “Week 1: Intro to R” (“Part 1,” “Part 2” y “More about packages”).
- *Datos*:
 - Votación del matrimonio igualitario (2010) en Argentina.
 - Ideología en las votaciones de la ONU: Michael A. Bailey, Anton Strezhnev, and Erik Voeten. Estimating Dynamic State Preferences from United Nations Voting Data. *Journal of Conflict Resolution*, 61(2):430–456, 2017.

Sesión 2 (Enero/27). Manipulación de datos (I)

- *Práctica*: manipulación y creación de variables; selección de subconjuntos; combinación de bases de datos; datos agrupados
 - Kelsey Moty, *workshop NYU*: <https://nyu-cdsc.github.io/learningr/>. Lecturas y ejercicios de “Week 1” (“More about variables” y “More about vectors”); “Week 5” (“Relational data” y “Joining data using dplyr’s join verbs”); y “Week 4” (“Describing data” y “Data transformation,” excepto la parte correspondiente a `ggplot2`), en ese orden.
 - Consigna para la *Tarea #1*.
- *Datos*:
 - Ideología en las votaciones de la ONU: Michael A. Bailey, Anton Strezhnev, and Erik Voeten. Estimating Dynamic State Preferences from United Nations Voting Data. *Journal of Conflict Resolution*, 61(2):430–456, 2017.
 - Ideología de los presidentes latinoamericanos.

Febrero/3. Descanso obligatorio

- No hay clase.

Sesión 3 (Febrero/10). Manipulación de datos (II): Agrupación por categorías

- *Teoría*: variación “between” vs. “within”
 - *Lectura optativa*: Andrew Gelman. *Red State, Blue State, Rich State, Poor State: Why Americans Vote the Way They Do (Expanded Edition)*. Princeton University Press, 2009, ch. 2.
- *Práctica*: datos agrupados
 - Roger D. Peng. *R Programming for Data Science*, 2016, cap. 2: <https://bookdown.org/rdpeng/rprogdatascience/managing-data-frames-with-the-dplyr-package.html>.
 - Kelsey Moty, *workshop NYU*: <https://nyu-cdsc.github.io/learningr/>. Lecturas y ejercicios de “Week 3” (“Learning tidyr” y “Manipulating your data using tidyr”) y “Week 4” (“Learning the main 6 dplyr verbs,” esp. el código correspondiente a la base de datos `disgust`).

- Entrega y revisión de la *Tarea #1*.
- Consigna para la *Tarea #2*.
- *Datos*:
 - Votación del matrimonio igualitario (2010) en Argentina.
 - Número de hijas y liberalismo en el Congreso norteamericano: Ebonya L. Washington. *Female Socialization: How Daughters Affect Their Legislator Fathers' Voting on Women's Issues*. *American Economic Review*, 98(1):311–332, 2008.
 - Índice de Gini: SWIID (<https://fsolt.org/swiid/>).
 - Ideología en las votaciones de la ONU: Michael A. Bailey, Anton Strezhnev, and Erik Voeten. *Estimating Dynamic State Preferences from United Nations Voting Data*. *Journal of Conflict Resolution*, 61(2):430–456, 2017.
 - Ideología de los presidentes latinoamericanos.

Sesión 4 (Febrero/17). Manipulación de datos (III): Visualización

- *Teoría*: cómo funciona ggplot2.
 - *Lectura obligatoria*: Kieran Healy. *Data Visualization: A Practical Introduction*. Princeton University Press, Princeton and Oxford, 2018, caps. 3-4.
 - *Lectura optativa*: Kieran Healy. *Data Visualization: A Practical Introduction*. Princeton University Press, Princeton and Oxford, 2018, caps. 1 y 5.
 - *Lectura optativa*: Sarah Leo, “[Mistakes, We've Drawn a Few](#)”.
- *Práctica*: visualización de datos.
 - Kelsey Moty, *workshop NYU*: <https://nyu-cdsc.github.io/learningr/>. Lecturas y ejercicios de “Week 2” (“Making plots” y “Getting a better understanding...”).
 - Entrega y revisión de la *Tarea #2*.
 - Consigna para la *Tarea #3*.
- *Datos*:
 - Votación del matrimonio igualitario (2010) en Argentina.
 - Número de hijas y liberalismo en el Congreso norteamericano: Ebonya L. Washington. *Female Socialization: How Daughters Affect Their Legislator Fathers' Voting on Women's Issues*. *American Economic Review*, 98(1):311–332, 2008.
 - Ideología en las votaciones de la ONU: Michael A. Bailey, Anton Strezhnev, and Erik Voeten. *Estimating Dynamic State Preferences from United Nations Voting Data*. *Journal of Conflict Resolution*, 61(2):430–456, 2017.
 - Ideología de los presidentes latinoamericanos.

Sesión 5 (Febrero/24). Manipulación de datos (IV): Sistemas de Información Geográfica

- *Teoría*: Estructura de datos con GIS
 - *Lectura optativa*: Kieran Healy. *Data Visualization: A Practical Introduction*. Princeton University Press, Princeton and Oxford, 2018, cap. 7.
 - *Lectura optativa*: Kosuke Imai. *Quantitative Social Science: An Introduction*. Princeton University Press, Princeton and Oxford, 2018, sec. 5.3.

- *Práctica*: coloreando y manipulando mapas
 - Mel Moreno y Mathieu Basille, “Drawing Beautiful Maps Programmatically with R, sf and ggplot2,” [Parte 1](#) y [Parte 2](#).
 - Continuación de la *Tarea #3*.
- *Datos*:
 - Resultados electorales en México y Bolivia.
 - Votación del matrimonio igualitario (2010) en Argentina.

Sesión 6 (Marzo/2). Regresión

- *Teoría*: regresión simple y múltiple
 - *Lectura optativa*: Wooldridge, secs. 2.1-2.4, 3.1-3.2 y 7.2-7.3.
- *Práctica*: interpretación y visualización de coeficientes
 - Entrega de la *Tarea #3*.
 - Consigna para la *Tarea #4*.
- *Datos*:
 - El Cuarteto de Anscombe y la Docena del Datasaurio.
 - Número de hijas y liberalismo en el Congreso norteamericano: Ebonya L. Washington. Female Socialization: How Daughters Affect Their Legislator Fathers’ Voting on Women’s Issues. *American Economic Review*, 98(1):311–332, 2008.

Marzo/9. Repaso

- Asistencia optativa.

Marzo/16. Descanso obligatorio

- Entrega de la *Tarea #4*.
- No hay clase.

Sesión 7 (Marzo/23). Primer parcial

- Examen parcial. En clase.

Sesión 8 (Marzo/30). Inferencia

- *Teoría*: significancia estadística + heteroscedasticidad + ataques de tiburones
 - *Lectura obligatoria*: Christopher H. Achen and Larry M. Bartels. *Democracy for Realists: Why Elections Do Not Produce Responsive Government*. Princeton University Press, Princeton, 2016, pp. 118-28.
 - *Lectura obligatoria*: Anthony Fowler and Andrew B. Hall. Do Shark Attacks Influence Presidential Elections? Reassessing a Prominent Finding on Voter Competence. *The Journal of Politics*, 80(4):1423–1437, 2018.

- *Lectura obligatoria*: Christopher H. Achen and Larry M. Bartels. Statistics as If Politics Mattered: A Reply to Fowler and Hall. *The Journal of Politics*, 80(4):1438–1453, 2018.
- *Lectura obligatoria*: Anthony Fowler and Andrew B. Hall. Politics as if Evidence Mattered: A Reply to Achen and Bartels. Unpublished manuscript, 2018.
- *Lectura optativa*: Wooldridge, secs. 4.1-4.3 y 8.1-8.2.
- *Lectura optativa*: Jeff Gill. The Insignificance of Null Hypothesis Significance Testing. *Political Research Quarterly*, 52(3):647–674, 1999.
- *Lectura optativa*: Carlisle Rainey. Arguing for a Negligible Effect. *American Journal of Political Science*, 58(4):1083–1091, 2014.
- *Práctica*: ajustando los errores estándar
- *Datos*:
 - Ideología en las votaciones de la ONU: Michael A. Bailey, Anton Strezhnev, and Erik Voeten. Estimating Dynamic State Preferences from United Nations Voting Data. *Journal of Conflict Resolution*, 61(2):430–456, 2017.
 - Ideología de los presidentes latinoamericanos.

Abril/6. Semana Santa

- No hay clase.

Sesión 9 (Abril/13). Interacciones

- *Teoría*: interpretación
 - *Lectura optativa*: Thomas Brambor, William R. Clark, and Matt Golder. Understanding Interaction Models: Improving Empirical Analyses. *Political Analysis*, 14(1):63–82, 2006.
 - *Lectura optativa*: Andrew Gelman, “[You need 16 times the sample size to estimate an interaction than to estimate a main effect](#)”.
- *Práctica*: incluyendo interacciones
 - Consigna para la *Tarea #5*.
- *Datos*:
 - Votación del matrimonio igualitario (2010) en Argentina.
 - Número de hijas y liberalismo en el Congreso norteamericano: Ebonya L. Washington. Female Socialization: How Daughters Affect Their Legislator Fathers’ Voting on Women’s Issues. *American Economic Review*, 98(1):311–332, 2008.

Sesión 10 (Abril/20). Variables dependientes discretas

Último día para elegir el artículo a replicar.

- *Teoría*: *logit*, *probit* y el modelo de probabilidad lineal
 - *Lectura obligatoria*: Walter Sosa Escudero, “[Regresión logística ¿Para todos, algunos o nadie?](#)”

- *Lectura optativa*: Wooldridge, secs. 7.5 y 17.1.
- *Práctica*: convirtiendo coeficientes de modelos *logit/probit* en probabilidades
 - Continuación de la *Tarea #5*.
- *Datos*:
 - Votación del matrimonio igualitario (2010) en Argentina.
 - Ideología presidencial y asistencia a funerales.

Sesión 11 (Abril/27). Efectos fijos vs. aleatorios + datos de panel

- *Teoría*: variación “within” y “between”
 - *Lectura optativa*: Andrew Bell and Kelvyn Jones. Explaining Fixed Effects: Random Effects Modeling of Time-Series Cross-Sectional and Panel Data. *Political Science Research and Methods*, 3(1):133–153, 2015.
 - *Lectura optativa*: Wooldridge, cap. 14.
- *Práctica*: corriendo modelos con efectos fijos vs. aleatorios; visualización de datos de panel
 - Entrega y revisión de la *Tarea #5*.
 - Consigna para la *Tarea #6*.
- *Datos*:
 - Votación del matrimonio igualitario (2010) en Argentina.
 - Ideología en las votaciones de la ONU: Michael A. Bailey, Anton Strezhnev, and Erik Voeten. Estimating Dynamic State Preferences from United Nations Voting Data. *Journal of Conflict Resolution*, 61(2):430–456, 2017.
 - Ideología de los presidentes latinoamericanos.

Sesión 12 (Mayo/4). Diferencia-en-Diferencias

- *Teoría*: diferencia-en-diferencias
 - *Lectura obligatoria*: Joshua D. Angrist and Jörn-Steffen Pischke. *Mastering 'Metrics. The Path from Cause to Effect*. Princeton University Press, Princeton, 2015, ch. 5.
 - *Lectura optativa*: Andrew Goodman-Bacon. Difference-in-Differences with Variation in Treatment Timing. Working Paper 25018, National Bureau of Economic Research, 2018.
- *Práctica*: DiD con *pooled cross sections* y *paneles*
 - Continuación de la *Tarea #6*.
- *Datos*:
 - Ideología en las votaciones de la ONU: Michael A. Bailey, Anton Strezhnev, and Erik Voeten. Estimating Dynamic State Preferences from United Nations Voting Data. *Journal of Conflict Resolution*, 61(2):430–456, 2017.
 - Ideología de los presidentes latinoamericanos.
 - LAPOP: opiniones sobre liderazgo femenino en América Latina.

Sesión 13 (Mayo/11). Segundo parcial

- Entrega de la *Tarea #6*.
- Examen parcial. En clase.

Sesión 14 (Mayo/18). Regresión discontinua

- *Teoría*: El modelo de regresión discontinua
 - *Lectura optativa*: Matías D. Cattaneo, Rocío Titiunik, and Gonzalo Vázquez-Bare. The Regression Discontinuity Design. In Luigi Curini and Robert J. Franzese, editors, *Handbook of Research Methods in Political Science and International Relations*. Sage, forthcoming.
- *Práctica*: visualización y estimación de regresiones discontinuas
- *Datos*:
 - Resultados electorales en México.
 - Gobiernos islámicos y educación femenina en Turquía: Erik Meyersson. Islamic Rule and the Empowerment of the Poor and Pious. *Econometrica*, 82(1):229–269, 2014.