

## Econometría Aplicada

División de Administración Pública  
Centro de Investigación y Docencia Económicas

### TEMARIO

**Horario:** Martes y jueves 8:00 – 9:30 vía Zoom.

**Laboratorio:** Miércoles 16:00 – 17:30 vía Microsoft Teams.

**Profesores:** César Rentería Marín (titular),  
Facundo Magaña Fajardo (adjunto).  
Vladimir Rodríguez Giménez (adjunto).

**Horario de atención a alumnos:** Lunes de 16:00 a 18:00 por Zoom.

**Presentación.** Los modelos estadísticos son una representación matemática de los procesos que generan o explican fenómenos sociales. Su utilidad, por lo tanto, radica en su capacidad de representar con fidelidad dichos procesos. El objetivo de este curso es que el estudiante aprenda técnicas y desarrolle habilidades para representar matemáticamente procesos correspondientes a fenómenos sociales.

El contenido del curso está organizado en cuatro partes. En la primera parte, *de la teoría al estudio empírico*, revisaremos técnicas para trasladar conceptos y teorías a mediciones y modelos matemáticos. En la segunda parte, estudiaremos con mucho detalle las características del estimador estrella en estadística, los *Mínimos Cuadrados Ordinarios*. En la tercera parte, *Modelación*, revisaremos herramientas para modelar con mayor precisión procesos sociales. Finalmente, revisaremos estimadores y técnicas estadísticas *más allá de Mínimos Cuadrados Ordinarios*.

Estoy convencido de que la estadística se aprende haciendo, por lo que este curso será primordialmente práctico. El curso requiere mucho trabajo práctico en Tareas y tiene un Taller para la aplicación de estadísticas en el software R. En clase, daré lecciones cortas y utilizaremos la mayor parte del tiempo para trabajar en preguntas o ejercicios. Evito a toda costa trabajo repetitivo y análisis mecánico de la estadística. Casi todos los ejercicios que haremos requieren un entendimiento de los principios estadísticos sobre los que se sustentan los estimadores y sus supuestos, los modelos estadísticos y la inferencia causal.

Este temario presenta más información de la que será necesaria para el curso, pues su objetivo principal es servir como recurso de consulta para la práctica estadística más allá de la clase. En cada sesión, por ejemplo, se listan *lecturas de apoyo*, cuyo objetivo es servir de referencia para los estudiantes que necesiten información adicional sobre estos temas en el futuro. En la sección *Recursos computacionales* se listan libros para el aprendizaje de funciones de R para la práctica estadística, que son más que suficientes para transitar a un nivel avanzado como usuarios de R. En la sección *Recursos complementarios a la clase*, se listan algunos recursos que varios estudiantes han encontrado útiles para complementar los temas impartidos en clase. La sección *Libros de estadística recomendados* presenta los libros de referencia más conocidos para diversos temas de estadística, que posiblemente deberán abordar durante sus carreras académicas. Finalmente, la sección *¿Cómo escribir un buen estudio empírico?* Ofrece artículos escritos por investigadores consolidados en los que ellos transmiten sus ideas de mejores prácticas en la elaboración de estudios empíricos.

**Formato del curso.** El curso está compuesto por cuatro horas de clase semanales y una hora y media taller de software estadístico semanal. La dinámica de las clases está diseñada bajo la técnica de [Aprendizaje Basado en Equipo](#) (ABE), en donde el trabajo fuera de clase de los estudiantes se concentra en la adquisición de conceptos, mismo que se pone a prueba en el salón de clases mediante dinámicas de equipo.

Una de las principales características del ABE es la Evaluación Rápida de Conocimiento (ERC), en las que los estudiantes deben aplicar los conocimientos adquiridos en su preparación para la sesión, primero individualmente, y posteriormente en equipos. Tendremos seis ERC durante el curso, que serán implementadas al inicio de la clase.

El curso será impartido vía remota. Con excepción de las sesiones de la clase y del taller, todo el contenido del curso, las interacciones y las entregas de trabajos se llevarán a cabo mediante Microsoft Teams. Las clases tomarán lugar en Zoom.

**Evaluación.** La evaluación del curso se basará en cinco componentes. El primer componente es la ERC, cuyas fechas de aplicación están en la sección de *Sesiones*. Durante el curso, se implementarán seis ERC, de las cuales se tomarán en cuenta cinco para la evaluación. Cada ERC se contesta dos veces. Primero se contesta de manera individual, cuya calificación se promediará con la calificación de la segunda etapa, en la que cada Equipo responde el ERC de manera colectiva. El segundo componente son las Tareas, cuyas fechas de entrega están en la sección de *Sesiones*. Se asignarán seis tareas, de las cuales se tomarán en cuenta cinco para la evaluación. Si el estudiante realizó las seis ERC o las seis tareas, se tomarán en cuenta las cinco calificaciones más altas. Existe además una Tarea 0, asignada antes de iniciar el curso. Esta tarea es opcional—sin valor en la calificación final—y su objetivo es nivelar las habilidades de programación de los estudiantes; también está diseñada como un preámbulo para la Replicación. Les recomiendo mucho que trabajen las tareas en grupos, aunque cada quién es responsable de entregar su propia tarea. El tercer componente es un examen parcial, cuya fecha de aplicación está en la sección de *Sesiones*. El cuarto componente es un arbitraje simulado de un artículo académico. Cada estudiante deberá evaluar críticamente los méritos teóricos y metodológicos de un artículo de investigación. En la sección *Arbitraje de artículos académicos empíricos* se provee información más detallada sobre esta actividad. Finalmente, los estudiantes realizarán una replicación de un artículo académico a lo largo de todo el semestre. Cada uno de los cinco componentes se evalúa con una escala del cero al diez.

Evaluación	N	Porcentaje
Evaluación Rápida de Conocimiento	5 de 7	20%
Tareas	4 de 5	30%
Examen parcial	1	20%
Arbitraje	1	10%
Replicación	1	20%
Total		100%

## SESIONES

Sesión y fecha	Descripción	<u>Evaluaciones</u> <u>Lecturas</u>
<b>Parte I. De la teoría al estudio empírico</b>		
08-18M	Conceptos	Goertz (2012) Caps. 1 y 2, Gerring (1999).  Lecturas de apoyo: Sartori (1970), Collier & Mahon (1993), Gerring & Barresi (2003), Collier & Levitsky (1996).
08-20J	Análisis exploratorio de datos	Brillinger (2011), Nuzzo (2016a), Sainani (2016).  Lecturas de apoyo: Gelman (2004), Chambers et al. (1983), Cleveland (1993), Hoaglin et al. (2003). Tukey (1977), Gerring (2012).
08-25M	Operacionalización	<u>Tarea 0. Juegos Olímpicos de Invierno</u>  Stevens (1946), Munck & Verkuilen (2002), Adcock & Collier (2001), O'Leary-Kelly & Vokurka (1998), Steckler & McLeroy (2008), Thomson, Perry, & Miller, (2009). *Vamos a utilizar Thompson, Perry & Miller (2009) para un ejercicio en clase.  Lecturas de apoyo: Goertz (2012) Cap. 4, Schram (2005), Acocck (2006) Cap. 12.
08-27J 09-01M 09-03J	- Funciones - Distribuciones - Regresión simple - Prueba de hipótesis	<u>ERC 1. Funciones y Distribuciones</u> <u>ERC 2. MCO</u>  Kennedy (2003) Cap. 5, Moore & Siegel (2013) Cap. 3. Pindyck & Rubinfeld (1998) Cap. 3, King (1998) Cap. 1. * De King, leer solo la sección 1.2 (pp. 6—13).  Regresión simple y prueba de hipótesis (leer sólo si necesitan repaso): Bailey (2016) Caps. 3 y 4.  Lecturas de apoyo: Gill (2006).
<b>Parte II. Mínimos Cuadrados Ordinarios</b>		
09-08M	Intervalos de confianza	Cohen (1992), Nuzzo (2014).  Lecturas de apoyo: Sainani (2009), Sainani (2011), Altman & Bland (2005), Cameron & Trivedi (2009) Cap 12, Kennedy (2003) Cap. 4, Cameron & Trivedi (2005) Cap. 7
09-10J	Poder estadístico	Krzywinski & Altman (2013), Nuzzo (2016b), Nakagawa & Cuthill (2007), Cohen (1994) .

		Lecturas de apoyo: Greenland et al. (2016).
09-15M	- Regresión Múltiple - Supuestos	<u>ERC 3. Prueba de Hipótesis y Poder Estadístico</u> <u>Tarea 1. MCO (Sept. 13)</u>  S: Lewis-Beck (1980) Cap. 2, Berry (1993) Cap 4. *De Lewis-Beck, leer sólo pp. 26—30.  Lecturas de apoyo: RM: Lewis-Beck (1980) Cap. 3, Pindyck & Rubinfeld (1998) Cap. 4. S: Fox (1991), Berry (1993).
09-17J	Sesgo en $\beta$	Bailey (2016) Cap. 5, Wooldridge (2012) Cap. 16 *De Wooldridge, leer sólo pp. 554—560.  Lecturas de apoyo: Cohen et al. (2002) Cap. 4.
09-22M	Datos atípicos	<u>ERC 4. Regresión Múltiple</u>  Fox (1991) Cap. 4, Aguinis, Gottfredson & Joo (2013).  Lecturas de apoyo: Cohen et al. (2002) Cap. 10.2-10.4 (pp. 391-419), Belsley, Kuh & Welsch (2005) Cap. 2.
09-24J 09-29M	- Sesgo en $var(\beta)$ - Multicolinealidad	<u>ERC 5. Supuestos</u> SEE: Pindyck & Rubinfeld (1998) Cap. 6. M: Fox (1991) Cap. 3.  Lecturas de apoyo: Belsley, Kuh & Welsch (2005), Cohen et al. (2002), Cap. 10.5-10.6 (pp. 419-30), Kennedy (2003) Cap. 12.
10-01J	<b>Examen parcial</b>	
<b>Parte III. Modelación</b>		
10-06M	Modelos estadísticos	<u>Tarea 2. Regresión Múltiple (Oct. 4)</u>  Lave & March (1993) Cap. 1, Kennedy (2002).  Lecturas de apoyo: Sainani (2014), Cameron & Trivedi (2005) Cap. 8, Varian (2016), Heckman (2015), Sparrowe & Mayer (2011), King (1991).
10-08J	Variables de control	McNamee (2003), Breaugh (2008). *De McNamee, leer sólo pp. 227—230.  Lecturas de apoyo: Bernerth & Aguinis (2016), Spector & Brannick (2011), Cinelli et al. (2020), Greenland (2008).
10-13M 10-15J	Variables independientes binarias	<u>ERC 6. Binarias e Interacciones</u>  Bailey (2016) Cap. 6.  Lecturas de apoyo: Kennedy (2003) Cap. 15.

10-20M	Efectos de interacción	<u>Tarea 3. Supuestos (Oct. 18)</u>  Brambor, Clark, & Golder (2006).  Lecturas de apoyo: Cohen et al. (2002) Caps. 7 y 9, Braumoeller (2004).
10-22J	Efectos fijos	Bailey (2016) Cap. 8.  Lecturas de apoyo: Schunck (2013), Cameron & Trivedi (2005) Cap. 21, Angrist & Pischke (2008) Cap. 5, Kennedy (2003) Cap. 18.
10-27M	Discusión de artículos académicos	Por definir.
10-29J	- Datos perdidos - Sesgo de selección de muestra	DP: Harel & Boyko (2013), Graham (2009). SSM: Winship & Mare (1992).  Lecturas de apoyo: DP: Sainani (2015), Cameron & Trivedi (2005) Caps. 16 y 27. SSM: Certo et al. (2016), King (1998) Cap. 8.
11-03M	Transformaciones y funciones polinomiales	Bailey (2016) Cap. 7  Lecturas de apoyo: Sainani (2012b), Bland & Altman (1996), Becker et al. (2019).
<b>Parte IV. Más allá de MCO</b>		
11-05J 11-10M 11-12J	Logit y Probit	<u>ERC 7. Variables Dependientes Cualitativas</u> <u>Tarea 4. Modelación (Nov. 1)</u>  Bailey (2016) Ch. 12.  Lecturas de apoyo: Long & Freese (2001) Cap. 4, Cameron & Trivedi (2009) Cap. 14, Aldrich and Nelson (1984) Cohen et al. (2002) Cap. 13.
11-17M	Poisson y Binomial Negativa	Coxe et al. (2009)  Lecturas de apoyo: Long & Freese (2001) Cap. 7, Cameron & Trivedi (2009) Cap. 17.
11-19J 11-24M	Inferencia Causal	<u>Tarea 5. Dependientes Cualitativas (Nov. 15)</u>  Angrist & Pischke (2008) Caps. 1 y 2, Ogden (2016) Cap. 3.  Lecturas de apoyo: Holland (1986), Bailey (2006) Cap. 1, Gerring (2005); Dague & Lahey (2019), Cohen et al. (2002) Cap. 12.
11-26J	- Selección de casos - Métodos mixtos	Mele & Belardinelli (2019).  Lecturas de apoyo: Soss (1999), Cresswell & Plano Clark (2011).

NA	Discusión de artículos académicos	Por definir.
12-1M	A elegir: <sup>1</sup> <ul style="list-style-type: none"> <li>• Análisis factorial. Acock (2006) Cap. 12.</li> <li>• Análisis de conglomerados. Kaufman &amp; Rousseeuw (1990) Cap. 1.</li> <li>• Series de tiempo. Jebb &amp; Tay (2017).</li> <li>• Modelos jerárquicos lineales. Woltman et al. (2012).</li> <li>• Modelos de ecuaciones estructurales. Williams, Vandenberg &amp; Edwards (2009)</li> <li>• Modelos de ecuaciones simultáneas. Pindyck &amp; Rubinfeld (1998) Cap. 12 o Woolridge (2012) Cap.16.</li> <li>• Estadística espacial. Darmofal (2015) Caps. 1 y 2.</li> <li>• Análisis de redes. Marin &amp; Wellman (2011).</li> <li>• Texto como datos. Grimmer &amp; Stewart (2013) o Hollibaugh (2019).</li> <li>• Experimentos con viñetas en Internet. Aguinis &amp; Bradley (2014) o Hansen &amp; Tummers (2020).</li> <li>• Aprendizaje de máquina. Anastasopoulos &amp; Whitford (2019).</li> <li>• Modelos no paramétricos. Jacoby (2000) o Andersen (2009).</li> </ul>	
12-16	<u>Entrega: Arbitraje de artículo académico.</u> <u>Entrega: Replicación Jennings &amp; Rubado (2015)</u>	

---

<sup>1</sup> En la última sesión, cada estudiante elegirá dos temas de y deberá hacer la lectura asignada a cada uno tema elegido. Durante la sesión, los estudiantes compartirán en grupos lo aprendido sobre las técnicas estadísticas elegidas.

### Taller de estadística (impartido por Facundo y Vladimir)

Sesión	Tema
1. 08-19	Juegos Olímpicos de Invierno I
2. 08-26	Juegos Olímpicos de Invierno II
3. 09-2	Técnicas de <i>flujo de trabajo</i> I
4. 09-9	Sesgo en $\beta$ (Simulaciones)
5. 09-16	Poder estadístico
6. 09-23	Técnicas de <i>flujo de trabajo</i> II
7. 09-30	Datos atípicos
8. 10-7	Replicación I
9. 10-14	Efectos de interacción
10. 10-21	Efectos fijos, Transformaciones y funciones polinomiales
11. 10-28	Datos perdidos
12. 11-4	Replicación II
13. 11-11	Variables dependientes cualitativas
14. 11-18	Pendiente
15. 11-25	Propensity Score Matching
16. 12-2	Diferencias y Diferencias Regresión Discontinua

### Recursos computacionales

Vamos a privilegiar el aprendizaje del *software* de programación R para los laboratorios y las tareas. La siguiente tabla tiene vínculos a guías con funciones de R para cada tema a ser cubierto en el laboratorio.

Gestión de datos	<a href="#">Data import tutorial</a> , <a href="#">Introduction to data cleaning with r</a> , <a href="#">exporting data</a> .
Análisis exploratorio de datos	<a href="#">Exploring data with graphics</a> .
Modelos MCO	<a href="#">Simple and Multiple Regression</a> .
Diagnóstico de MCO	<a href="#">Regression Diagnostics</a> .
Efectos de interacción	<a href="#">Decomposing, Probing and Plotting interactions in R</a> , <a href="#">An introduction to margins</a>
Transformaciones	<a href="#">Transformations</a> .
Datos perdidos	<a href="#">Dealing with missing values</a> , <a href="#">Amelia II: A Program for Missing Data</a> .
Poder estadístico	<a href="#">Introduction to Power Analysis</a> , <a href="#">Intermediate Power Analysis</a> .
Modelos no paramétricos	<a href="#">Nonparametric regression</a> , <a href="#">The np package</a> .
Variables dependientes cualitativas	<a href="#">Logit regression</a> , <a href="#">Probit regression</a> , <a href="#">Poisson regression</a> , <a href="#">Negative Binomial regression</a> .

En la siguiente lista, comparto libros *abiertos* y cursos de programación de estadística en R para que continúen desarrollando sus habilidades de programación y análisis estadístico durante y después del curso. Estos materiales deberían ser suficientes para que alcances un nivel intermedio de programación.

#### Libros:

- Boehmke [Introduction to R](#)
- Diaz et al. [OpenIntro Statistics](#)
- Hanck et al. [Econometrics with R](#)
- Dalpiaz [Applied Statistics with R](#)
- Grolemond & Wickham [R for Data Science](#)
- Phillips [YaRrr! The Pirate's Guide to R](#)

#### Cursos en línea:

- [Introduction to R](#)
- [Exploratory Data Analysis in R](#)
- [Importing and cleaning data in R](#)
- Data Visualization, [Part I](#) y [Part II](#)

### Recursos complementarios a la clase

El curso requiere de bases sólidas de probabilidad (que asumo ya todos tienen). En caso de que alguien necesite fortalecer alguno de estos aspectos, recomiendo mucho el libro *Math Prefresher for Political Sciences*, en particular los siguientes capítulos:

- 1 Linear Algebra
- 2 Functions and Operations

Muchos estudiantes han encontrado los siguientes cursos en línea muy útiles para complementar los revisado en clase:

- Kahn Academy [Statistics and Probability](#)
- Ben Lambert: A Full Course in Econometrics (Licenciatura) [Parte I](#) y [Parte II](#)
- Ben Lambert: [A Graduate Course in Econometrics](#)
- David Siegel: [A Mathematics Course for Political and Social Research](#)

### Libros de estadística recomendados

La siguiente lista presenta los libros que considero vale la pena tener en el librero para los estudiantes que desean especializarse en métodos cuantitativos después del curso.

- Conocimiento básico de estadística aplicada
  - Bailey, M. A. (2016). Real Stats: using econometrics for political science and public policy. Oxford: Oxford University Press.
  - Kennedy, P. (2003). A guide to econometrics. MIT press.
  - Pindyck, R. S., & Rubinfeld, D. L. (1998). Econometric models and economic forecasts.
  - Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2002). Applied multiple regression/correlation analysis for the behavioral sciences. Routledge.
  - Cameron, A. C., & Trivedi, P. K. (2005). Microeconometrics: methods and applications. Cambridge University Press.



- Wooldridge, J. M. (2012). Introductory econometrics: A modern approach. Nelson Education.
- Modelación estadística
  - Lave, C. A., & March, J. G. (1993). An introduction to models in the social sciences. University Press of America.
  - King, G. (1998). Unifying political methodology: The likelihood theory of statistical inference. University of Michigan Press.
- Conceptos, operacionalización y medición
  - Goertz, G. (2012). Social science concepts: A user's guide. Princeton University Press.
  - Blalock, H. M. (1982). Conceptualization and measurement in the social sciences (No. 04; H61, B5.).
- Variables cualitativas, variables de conteo, datos panel y modelos jerárquicos
  - Long, J. S. (1996). The analysis of categorical and limited dependent variables. Department of Sociology Indiana University.
  - Cameron, A. C., & Trivedi, P. K. (2013). Regression analysis of count data (Vol. 53). Cambridge university press.
  - Wooldridge, J. M. (2010). Econometric analysis of cross section and panel data. MIT Press.
  - Hsiao, C. (2014). Analysis of panel data. Cambridge University Press.
  - Gelman, A., & Hill, J. (2006). Data analysis using regression and multilevel/hierarchical models. Cambridge University Press.
- Datos perdidos
  - Little, R. J., & Rubin, D. B. (2019). Statistical analysis with missing data (Vol. 793). John Wiley & Sons.
- Análisis de poder estadístico
  - Cohen, J. (2013). Statistical power analysis for the behavioral sciences. Academic Press.
- Matemáticas para las ciencias sociales
  - Moore, W. H., & Siegel, D. A. (2013). A mathematics course for political and social research. Princeton University Press.
  - Gill, J. (2006). Essential mathematics for political and social research (pp. 271-277). Cambridge: Cambridge University Press.
- Inferencia causal
  - Angrist, J. D., & Pischke, J. S. (2008). Mostly harmless econometrics: An empiricist's companion. Princeton university press.
  - Ogden, T. N. (Ed.). (2016). Experimental conversations: Perspectives on randomized trials in development economics. MIT Press.
  - Morgan, S. L., & Winship, C. (2015). Counterfactuals and causal inference. Cambridge University Press.

Algunos de ustedes quizá deseen especializarse en métodos cuantitativos—a la par de la especialización en un tema sustantivo de administración pública. Los materiales que se presentan en este temario son una sólida base para comenzar, pero les recomiendo revisar con regularidad revistas académicas especializadas en métodos y los blogs de especialistas, como Andrew Gelman, Gary King, Sander Greenland o Kristin Sainani. En el siguiente [vínculo](#), comparto estos recursos y

comparto una lista de artículos metodológicos orientados a la comunidad de investigadores en administración pública.

### ¿Cómo escribir un *buen* estudio empírico?

Los mejores estudios empíricos están sustentados en un sólido razonamiento estadístico y en el conocimiento de las buenas y malas prácticas en el análisis estadístico o en la comunicación de resultados. A lo largo del curso, revisaremos buenas y malas prácticas, aunque quizá no podamos revisarlas con detalle. La siguiente es una bibliografía de discusiones sobre buenas y malas prácticas al elaborar un estudio empírico, que les podrá ser útil en el futuro.

- King, G. (1986). How not to lie with statistics: Avoiding common mistakes in quantitative political science. *American Journal of Political Science*, 666-687.
- Sainani, K. L. (2013). Avoiding careless errors: know your data. *PM&R*, 5(3), 228-229.
- Kennedy, P. E. (2002). Sinning in the basement: What are the rules? The ten commandments of applied econometrics. *Journal of Economic Surveys*, 16(4), 569-589.
- Green, J. P., Tonidandel, S., & Cortina, J. M. (2016). Getting through the gate: Statistical and methodological issues raised in the reviewing process. *Organizational Research Methods*, 19(3), 402-432.
- Sparrowe, R. T., & Mayer, K. J. (2011). Publishing in AMJ—part 4: grounding hypotheses.
- Zhang, Y., & Shaw, J. D. (2012). Publishing in AMJ—Part 5: Crafting the methods and results.
- Bailey (2016) Cap. 2
- Beck, N. (2010). Making regression and related output more helpful to users. *The Political Methodologist* 18, 4(9).
- Kastellec, J. P. and E. L. Leoni (2007). Using graphs instead of tables in political science. *Perspectives on Politics* 5, 755{771.
- Lange, D., & Pfarrer, M. D. (2017). Editors' comments: Sense and structure—The core building blocks of an AMR article.
- Woodside, A. G. (2016). The good practices manifesto: Overcoming bad practices pervasive in current research in business.

### Arbitraje de artículos académicos empíricos.

Para el arbitraje del artículo académico espero un documento de mínimo dos y máximo tres páginas completas, sin incluir referencias o apéndices (en caso de que decidan incluirlos). Para el arbitraje, espero que evalúen críticamente la teoría, la conexión entre la teoría y el estudio empírico y las decisiones metodológicas del artículo. El formato es libre. Les recomiendo que platiquen con sus asesores para que les compartan su experiencia en el arbitraje de artículos académicos. Para complementar eso, les comparto una lista de artículos que proveen consejos sobre cómo arbitrar un artículo académico:

- Spolarich, A. E., & Wilder, R. S. (2015). Becoming an effective journal reviewer. *American Dental Hygienists' Association*, 89(1), 51-53.
- Martin, B. (2008). Writing a helpful referee's report. *Journal of Scholarly Publishing*, 39(3), 301-306.

- LePine, J. A., & King, A. W. (Eds.). (2010). Editors' comments: Developing novel theoretical insight from reviews of existing theory and research.
- Rai, A. (2016). Editor's comments: writing a virtuous review. *MIS Quarterly*, 40(3), iii-x.
- Colquitt, J. A., & Ireland, R. D. (2009). Taking the mystery out of AMJ's reviewer evaluation form.

**Gratitudes.** Este temario es el resultado de la consulta de múltiples fuentes, por ejemplo, los materiales de los cursos que he tomado en estadística, los materiales disponibles en Internet de cursos en otras universidades, y mis anotaciones y aprendizajes a partir conversaciones con colegas y profesores, presentaciones en conferencias y *job talks*, de la elaboración o recepción de arbitrajes de pares y de la lectura crítica de artículos de investigación. Gran parte del contenido de este curso fue generado a partir de los cursos de estadística de Shawn Bushway y Matthew Ingram, con quienes tuve la fortuna de colaborar. También he adaptado mucho contenido—y estrategias de enseñanza—de Peter Kennedy, Michael Bailey, Andrew Gelman, Robert Pindyck y Daniel Rubinfeld.

## Bibliografia

- Acock, A. C. (2008). *A gentle introduction to Stata*. Stata press.
- Adcock, R., & Collier, D. (2001). Measurement validity: A shared standard for qualitative and quantitative research. *American political science review*, 95(3), 529-546.
- Aldrich, J. H., Nelson, F. D., & Adler, E. S. (1984). *Linear probability, logit, and probit models* (No. 45). Sage.
- Altman, D. G., & Bland, J. M. (2005). Standard deviations and standard errors. *Bmj*, 331(7521), 903.
- Anastasopoulos, L. J., & Whitford, A. B. (2019). Machine learning for public administration research, with application to organizational reputation. *Journal of Public Administration Research and Theory*, 29(3), 491-510.
- Andersen, R. (2009). Nonparametric methods for modeling nonlinearity in regression analysis. *Annual Review of Sociology*, 35, 67-85.
- Angrist, J. D., & Pischke, J. S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.
- Aguinis, H., & Bradley, K. J. (2014). Best practice recommendations for designing and implementing experimental vignette methodology studies. *Organizational research methods*, 17(4), 351-371.
- Aguinis, H., Gottfredson, R. K., & Joo, H. (2013). Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods*, 16(2), 270-301.
- Bailey, M. A. (2016). *Real stats: Using econometrics for political science and public policy*. Oxford: Oxford University Press.
- Becker, T. E., Robertson, M. M., & Vandenberg, R. J. (2019). Nonlinear transformations in organizational research: Possible problems and potential solutions. *Organizational Research Methods*, 22(4), 831-866.
- Belsley, D. A., Kuh, E., & Welsch, R. E. (2005). *Regression diagnostics: Identifying influential data and sources of collinearity* (Vol. 571). John Wiley & Sons.
- Bernerth, J. B., & Aguinis, H. (2016). A critical review and best-practice recommendations for control variable usage. *Personnel Psychology*, 69(1), 229-283.
- Berry, W. D. (1993). *Understanding regression assumptions*. Sage Publications.
- Bland, J. M., & Altman, D. G. (1996). Statistics notes: Transforming data. *Bmj*, 312(7033), 770.
- Braumoeller, B. F. (2004). Hypothesis testing and multiplicative interaction terms. *International organization*, 58(4), 807-820.
- Brambor, T., Clark, W. R., & Golder, M. (2006). Understanding interaction models: Improving empirical analyses. *Political analysis*, 14(1), 63-82.
- Breaugh, J. A. (2008). Important considerations in using statistical procedures to control for nuisance variables in non-experimental studies. *Human Resource Management Review*, 18(4), 282-293.

Brillinger, D. R. (2011). Data analysis, exploratory. International encyclopedia of political science. Thousand Oaks, CA, USA: SAGE Publications, Inc, 531-538.

Cameron, A. C., & Trivedi, P. K. (2005). Microeconometrics: methods and applications. Cambridge university press.

Cameron, A. C., & Trivedi, P. K. (2009). Microeconometrics using Stata (Vol. 5, p. 706). College Station, TX: Stata press.

Certo, S. T., Busenbark, J. R., Woo, H. S., & Semadeni, M. (2016). Sample selection bias and Heckman models in strategic management research. Strategic Management Journal, 37(13), 2639-2657.

Chambers, J. M., Cleveland, W. S., Kleiner, B., & Tukey, P. A. (1983). Graphical methods for data analysis. Wadsworth & Brooks. Cole Statistics/Probability Series.

Cinelli, C., Forney, A., & Pearl, J. (2020). A Crash Course in Good and Bad Controls.

Cleveland, W. S. (1993). Visualizing data. Hobart Press.

Cohen, J. (1992). Things I have learned (so far). In Annual Convention of the American Psychological Association, 98th, Aug, 1990, Boston, MA, US; Presented at the aforementioned conference.. American Psychological Association.

Cohen, J. (1994). The earth is round ( $p < .05$ ). American psychologist, 49(12), 997.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2002). Applied multiple regression/correlation analysis for the behavioral sciences. Routledge.

Collier, D., & Mahon, J. E. (1993). Conceptual “stretching” revisited: Adapting categories in comparative analysis. American Political Science Review, 87(4), 845-855.

Collier, D., & Levitsky, S. (1996). Democracy «With Adjectives»: Conceptual Innovation in Comparative Research.

Coxe, S., West, S. G., & Aiken, L. S. (2009). The analysis of count data: A gentle introduction to Poisson regression and its alternatives. Journal of personality assessment, 91(2), 121-136.

Cresswell, J. W., & Plano Clark, V. L. (2011). Designing and conducting mixed method research. 2nd Sage. Thousand Oaks, CA.

Dague, L., & Lahey, J. N. (2019). Causal inference methods: Lessons from applied microeconomics. Journal of Public Administration Research and Theory, 29(3), 511-529.

Darmofal, D. (2015). Spatial analysis for the social sciences. Cambridge University Press.

Fox, J. (1991). Regression diagnostics: An introduction (Vol. 79). SAGE Publications.

Graham, J. W. (2009). Missing data analysis: Making it work in the real world. Annual review of psychology, 60, 549-576.

Gelman, A. (2004). Exploratory data analysis for complex models. Journal of Computational and Graphical Statistics, 13(4), 755-779.

- Gerring, J. (1999). What makes a concept good? A criterial framework for understanding concept formation in the social sciences. *Polity*, 31(3), 357-393.
- Gerring, J. (2005). Causation: A unified framework for the social sciences. *Journal of theoretical politics*, 17(2), 163-198.
- Gerring, J. (2012). Mere description. *British Journal of Political Science*, 721-746.
- Gerring, J., & Barresi, P. A. (2003). Putting ordinary language to work: a min-max strategy of concept formation in the social sciences. *Journal of Theoretical Politics*, 15(2), 201-232.
- Gill, J. (2006). *Essential mathematics for political and social research* (pp. 271-277). Cambridge: Cambridge University Press.
- Goertz, G. (2012). *Social science concepts: A user's guide*. Princeton University Press.
- Greenland, S. (2008). Invited commentary: variable selection versus shrinkage in the control of multiple confounders. *American Journal of Epidemiology*, 167(5), 523-529.
- Greenland, S., Senn, S. J., Rothman, K. J., Carlin, J. B., Poole, C., Goodman, S. N., & Altman, D. G. (2016). Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. *European journal of epidemiology*, 31(4), 337-350.
- Grimmer, J., & Stewart, B. M. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political analysis*, 21(3), 267-297.
- Hansen, J. A., & Tummers, L. (2020). A Systematic Review of Field Experiments in Public Administration. *Public Administration Review*.
- Harel, O., & Boyko, J. (2013). Mining data: should we care? *American journal of public health*, 103(2), 200-201.
- Heckman, J. J. (2015). Gary Becker: model economic scientist. *American Economic Review*, 105(5), 74-79.
- Holland, P. W. (1986). Statistics and causal inference. *Journal of the American statistical Association*, 81(396), 945-960.
- Hollibaugh, G. E. (2019). The use of text as data methods in public administration: A review and an application to agency priorities. *Journal of Public Administration Research and Theory*, 29(3), 474-490.
- Hoaglin, D. C. (2003). John W. Tukey and data analysis. *Statistical Science*, 311-318.
- Jacoby, W. G. (2000). Loess: a nonparametric, graphical tool for depicting relationships between variables. *Electoral Studies*, 19(4), 577-613.
- Jebb, A. T., & Tay, L. (2017). Introduction to time series analysis for organizational research: Methods for longitudinal analyses. *Organizational Research Methods*, 20(1), 61-94.
- Kaufman, L., & Rousseeuw, P. J. (1990). *Finding groups in data: an introduction to cluster analysis*. John Wiley & Sons.

- Kennedy, P. E. (2002). Sinning in the basement: What are the rules? The ten commandments of applied econometrics. *Journal of Economic Surveys*, 16(4), 569-589.
- Kennedy, P. (2003). *A guide to econometrics*. MIT press.
- King, G. (1991). " Truth" Is Stranger than Prediction, More Questionable than Causal Inference. *American Journal of Political Science*, 35(4), 1047-1053.
- King, G. (1998). *Unifying political methodology: The likelihood theory of statistical inference*. University of Michigan Press.
- Krzywinski, M., & Altman, N. (2013). *Points of significance: Power and sample size*.
- Lave, C. A., & March, J. G. (1993). *An introduction to models in the social sciences*. University Press of America.
- Lewis-Beck, C., & Lewis-Beck, M. (1980). *Applied regression: An introduction* (Vol. 22). Sage publications.
- Long, J. S., & Freese, J. (2001). *Regression models for categorical dependent variables using Stata*. Stata.
- Marin, A., & Wellman, B. (2011). *Social network analysis: An introduction*. The SAGE handbook of social network analysis, 11.
- McNamee, R. (2003). Confounding and confounders. *Occupational and environmental medicine*, 60(3), 227-234.
- Mele, V., & Belardinelli, P. (2019). Mixed methods in public administration research: Selecting, sequencing, and connecting. *Journal of Public Administration Research and Theory*, 29(2), 334-347.
- Moore, W. H., & Siegel, D. A. (2013). *A mathematics course for political and social research*. Princeton University Press.
- Munck, G. L., & Verkuilen, J. (2002). Conceptualizing and measuring democracy: Evaluating alternative indices. *Comparative political studies*, 35(1), 5-34.
- Nakagawa, S., & Cuthill, I. C. (2007). Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biological reviews*, 82(4), 591-605.
- Nuzzo, R. (2014). Scientific method: statistical errors. *Nature News*, 506(7487), 150.
- Nuzzo, R. L. (2016a). The box plots alternative for visualizing quantitative data. *PM&R*, 8(3), 268-272.
- Nuzzo, R. L. (2016b). Statistical power. *Pm&r*, 8(9), 907-912.
- Ogden, T. N. (Ed.). (2016). *Experimental conversations: Perspectives on randomized trials in development economics*. MIT Press.
- O'Leary-Kelly, S. W., & Vokurka, R. J. (1998). The empirical assessment of construct validity. *Journal of operations management*, 16(4), 387-405.
- Pindyck, R. S., & Rubinfeld, D. L. (1988). *Econometric models and economic forecasts*.
- Sainani, K. L. (2009). Putting P values in perspective. *PM&R*, 1(9), 873-877.



- Sainani, K. L. (2011). A closer look at confidence intervals. *PM&R*, 3(12), 1134-1141.
- Sainani, K. L. (2012). Dealing With Non-normal Data. *PM&R*, 4(12), 1001-1005.
- Sainani, K. L. (2014). Explanatory versus predictive modeling. *PM&R*, 6(9), 841-844.
- Sainani, K. L. (2015). Dealing with missing data. *PM&R*, 7(9), 990-994.
- Sainani, K. L. (2016). The Value of Scatter Plots. *PM&R*, 8(12), 1213-1217.
- Sartori, G. (1970). Concept misinformation in comparative politics. *American political science review*, 64(4), 1033-1053.
- Schram, A. (2005). Artificiality: The tension between internal and external validity in economic experiments. *Journal of Economic Methodology*, 12(2), 225-237.
- Schunck, R. (2013). Within and between estimates in random-effects models: Advantages and drawbacks of correlated random effects and hybrid models. *The Stata Journal*, 13(1), 65-76.
- Soss, J. (1999). Lessons of welfare: Policy design, political learning, and political action. *American Political Science Review*, 93(2), 363-380.
- Sparrowe, Raymond and Kyle Mayer. 2011. Publishing in AMJ Part 4: Grounding Hypotheses. *Academy of Management Journal* 54(6): 1098-1102.
- Spector, P. E., & Brannick, M. T. (2011). Methodological urban legends: The misuse of statistical control variables. *Organizational Research Methods*, 14(2), 287-305.
- Steckler, A., & McLeroy, K. R. (2008). The importance of external validity.
- Stevens, S. S. (1946). On the theory of scales of measurement.
- Thomson, A. M., Perry, J. L., & Miller, T. K. (2009). Conceptualizing and measuring collaboration. *Journal of Public Administration Research and Theory*, 19(1), 23-56.
- Tukey, J. W. (1977). *Exploratory data analysis* (Vol. 2, pp. 131-160).
- Williams, L. J., Vandenberg, R. J., & Edwards, J. R. (2009). Structural equation modeling in management research: A guide for improved analysis. *Academy of Management Annals*, 3(1), 543-604.
- Winship, C., & Mare, R. D. (1992). Models for sample selection bias. *Annual review of sociology*, 18(1), 327-350.
- Woltman, H., Feldstain, A., MacKay, J. C., & Rocchi, M. (2012). An introduction to hierarchical linear modeling. *Tutorials in quantitative methods for psychology*, 8(1), 52-69.
- Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT Press.
- Varian, H. R. (2016). How to build an economic model in your spare time. *The American Economist*, 61(1), 81-90.