



UNAM  
POSGRADO



Programa  
Universitario  
de Estudios  
del Desarrollo  
UNAM

Presentación del curso

# Introducción al análisis bayesiano de datos

Dr. Héctor Nájera  
Dr. Curtis Huffman



# ¿Qué entendemos por el estadística?

- No podemos hacer estadística sin datos
- Lo que hacemos es modelar los vínculos entre los datos y los constructos substantivos de interés
  - En estadística nos interesa elucidar la relación entre modelos científicos y ciertos datos para producir conocimiento
- En estadística aplicada continuamente tenemos que recordar el valor de integrar nuevos datos en cierto análisis teniendo en mente la “calidad” y “cantidad” de datos existentes
  - Calidad: Error de medición, sesgos en muestro, errores de recolección.
  - Cantidad: Poder estadístico para distinguir las relaciones de interés
- En algunos casos es posible responder las preguntas que nos interesa usando datos limpios pero regularmente nos damos cuenta que los ejemplos de nuestros libros de texto no nos sirve



# Estadística e información

- Recomendaciones de qué variables incluir en un modelo para ajustar por la representatividad de la encuesta
- Qué hacer con las observaciones aberrantes: regulares, errores o nuevos descubrimientos
- Problemas de medición (error de medición)
- Distribución poblacional de referencia y la de interés
- Regularización de las estimaciones cuando los datos son dispersos
- La forma funcional del modelo
- Información sobre los parámetros de interés

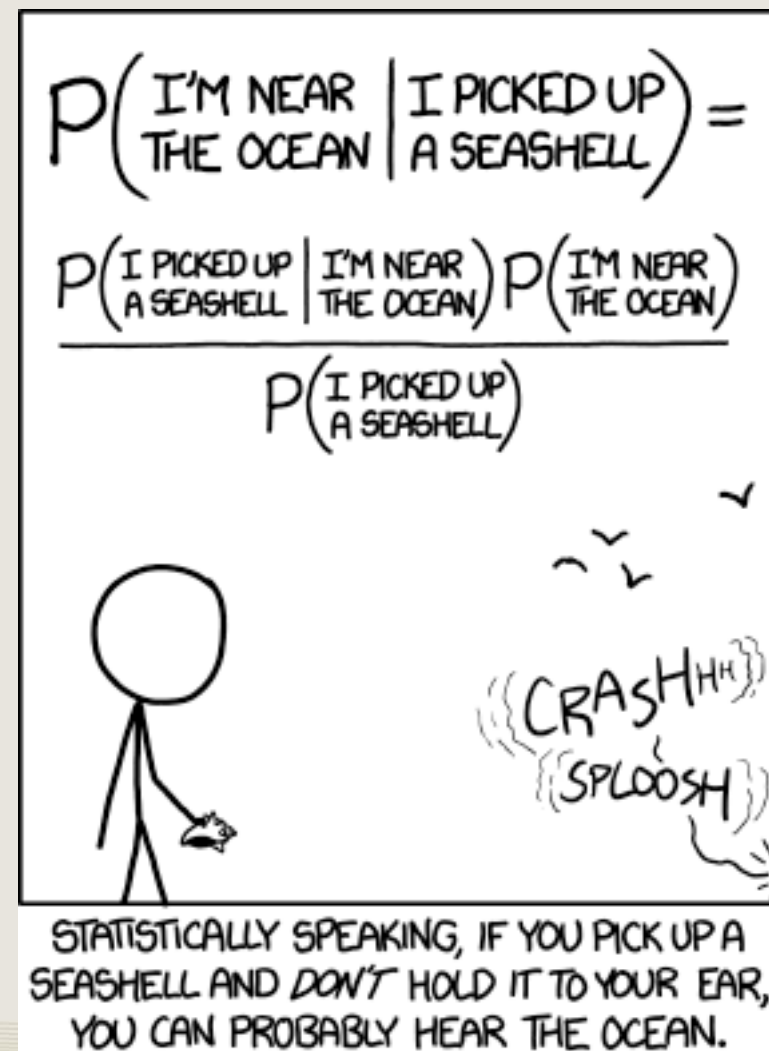
¿Hasta que punto nuestras inferencias son objetivas?

Si  $\beta = 3$  [1-5]  $p < .05$

¿Objetivamente descubrí que el efecto es 3?

# Análisis Bayesiano de datos

*O sobre cómo uso ciertos principios en estadística, bajo un marco computacional factible, para ajustar la incertidumbre asociada a mis conclusiones cuando combino nueva y vieja información*



# ¿Por qué este curso es importante?

- Vivimos en un mundo en el que la investigación empírica en la forma de inferencia estadística está jugando un papel de la mayor importancia, desde el pronóstico de elecciones, el descubrimiento de patrones en los hábitos de consumo (Amazon, Netflix, Facebook), el reconocimiento de voz para el control de cómputo, hasta encontrar las bases genéticas de enfermedades como el Covid-19.



Bayesian



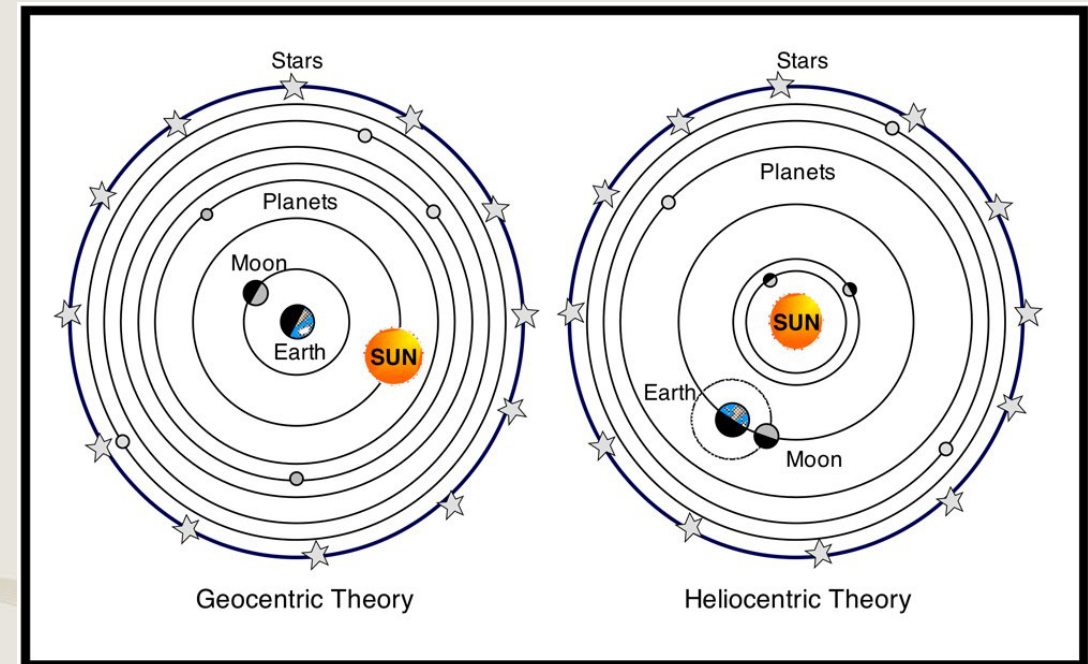


# ¿Usar bayes porque hay más datos?

## Estadística Bayesiana, IA y ML

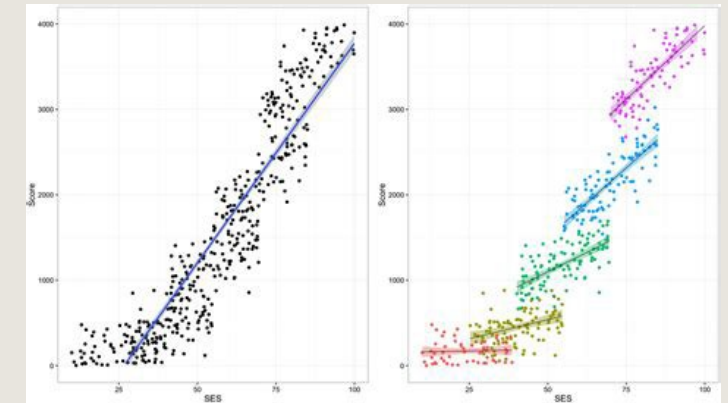
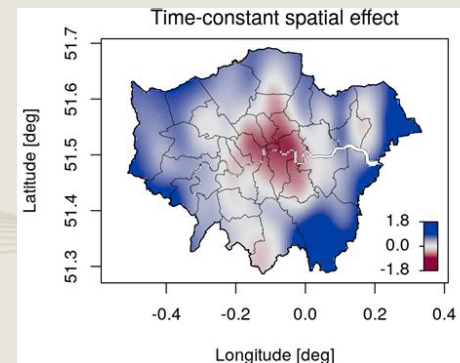
- ML no tiene una definición precisa, quizá la estadística bayesiana puede verse como un caso especial.
- ML usa modelos grandes para datos grandes haciendo el **mínimo posible de supuestos**: Muchos datos de **entrenamiento**
- La inferencia bayesiana, por el contrario, usa **estructuras fuertes** (teoría) y **supuestos fuertes** respecto a las hipótesis (priors).

“**Bayesian inference excels when your data are sparse and your model is dense.**”



# ¿Por qué este curso es importante?

- Sí, tenemos más datos y mayor capacidad de cómputo, pero algunos viejos problemas no han desaparecido.
- A veces es cuestión de que la estadística clásica entra en apuros.
  - Estudios pequeños con tamaños de efecto pequeños
  - Series cortas de tiempo
  - Pocas unidades en modelación multinivel
  - Queremos combinar datos de distintas fuentes
  - Nos gustaría utilizar la mayor cantidad de información possible
  - O viejas aproximaciones de modelación nos gustaría modernizarlas



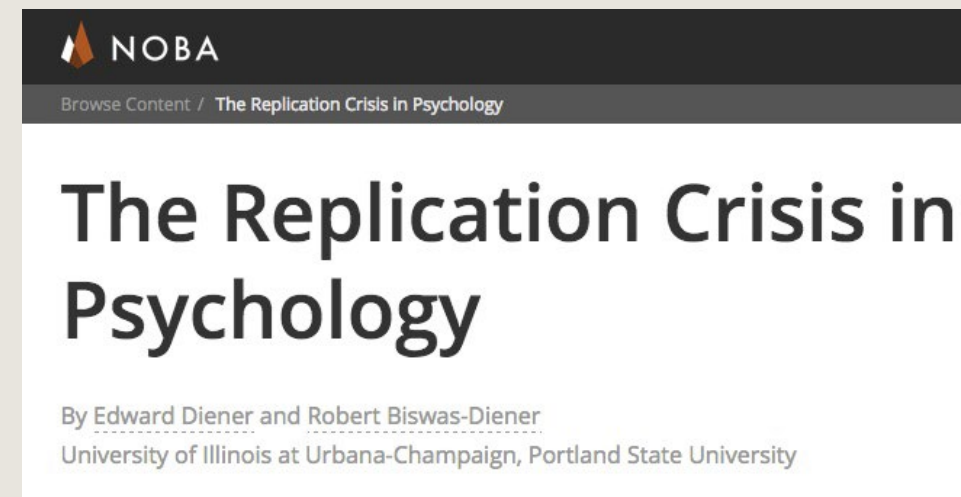


# ¿Por qué este curso es importante?

- Hoy en día no es posible “andar por la vida” sin “chocar” con la estadística bayesiana
  - se ha convertido en un tema de alfabetismo para la sociedad en su conjunto
  - de cómo interpretar información en la vida cotidiana
    - procesos electorales
    - salud pública
  - y en la vida profesional de la economista
    - cada vez más la profesión requiere un entendimiento cabal de las diferentes maneras en que es posible cuantificar **incertidumbre** usando medidas de probabilidad
      - las fortalezas y debilidades de lo que significa “cuantificar incertidumbre” para diferentes escuelas de estadística (estilos de inferencia)



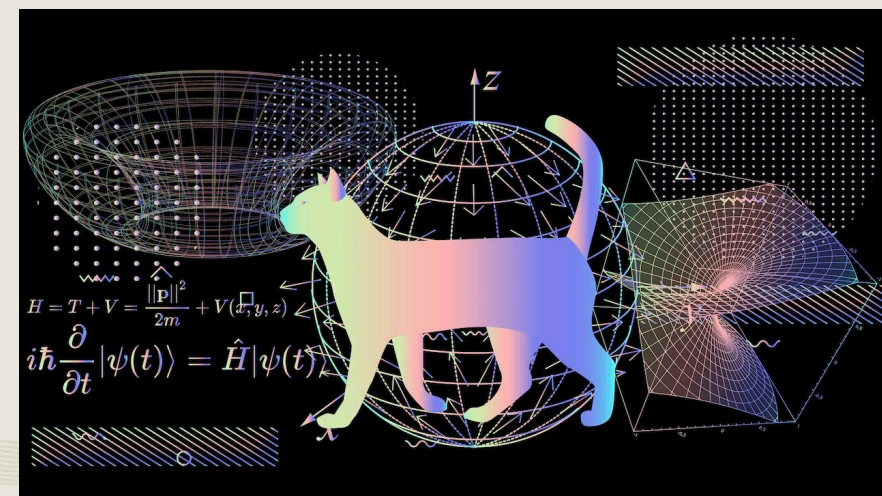
# ¿Por qué este curso es importante?



La estadística busca brindarnos herramientas para reducir la incertidumbre asociada a nuestras hipótesis de investigación. Descubrimos porque tenemos algún grado de certeza.

Necesitamos un aparato estadístico que nos permita incorporar la incertidumbre asociada a nuestros resultados.

# Temas altamente controvertidos



- In 2020 most of the meaningful statistical research will be Bayesian (2000) (Donald Rubin)
- A Bayesian statistician is at least as good as a classical one (Gelfan, 2017)
- What's important about the statistical model is not what it does with the data, but what data it uses (Andrew Gelman)
- It is probably better to realize that the probability concept is in a sense *subjective*, that it is always based on uncertain knowledge, and that its quantitative evaluation is **subject to change as we obtain more information.** (Richard Feynman)



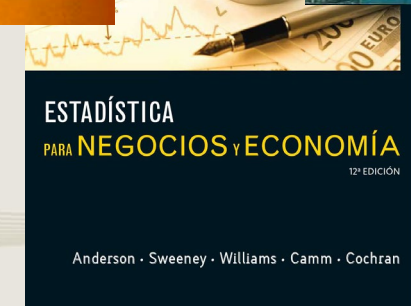
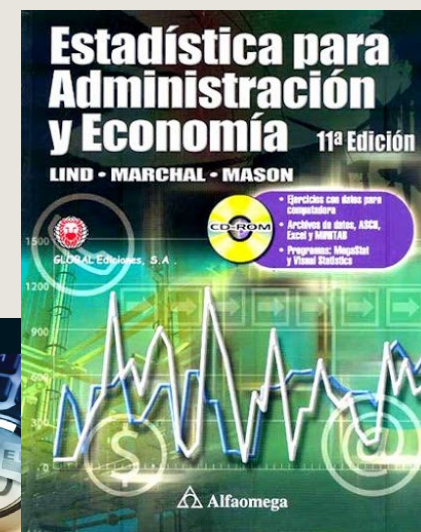
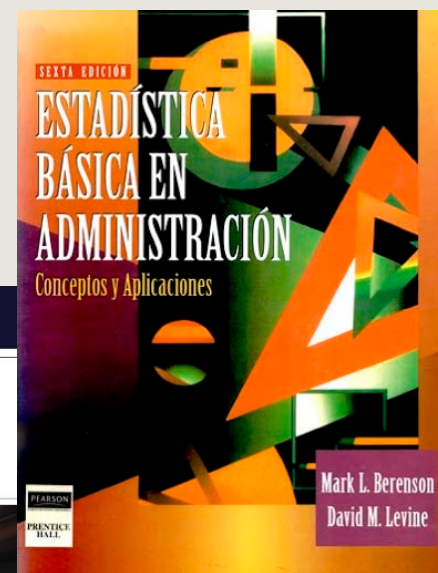
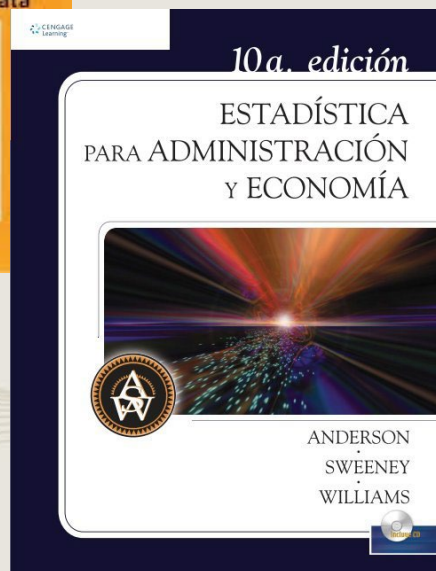
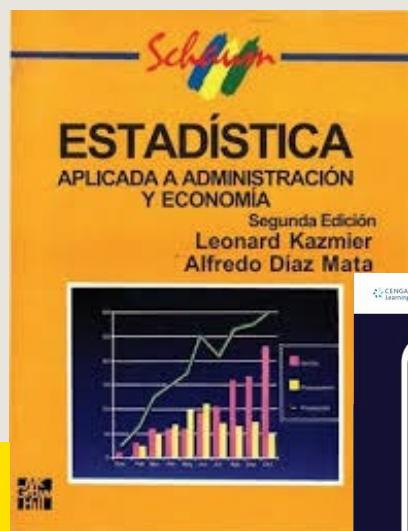
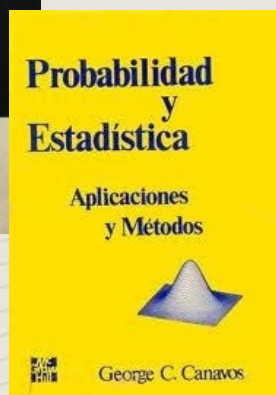
# Análisis bayesiano e investigación empírica

- La mayoría de los estadísticos contemporáneos más destacados son Bayesianos
- Las conferencias de punta en estadística son Bayesianas
- Los trabajos de punta (Tesla, Google, universidades top, bolsa de valores, cambio climático, NASA, MLB, NBA, Fórmula 1) se hacen con estadística Bayesiana
- Los programas de estadística en las universidades top en Estados Unidos en su mayoría Bayesianos
- Los modelos más complejos y exitosos que conocemos tiene generalmente un rasgo Bayesiano
- La escritura de nuevo software está pensada para hacer inferencia Bayesiana antes que clásica



# Introducción al análisis bayesiano

- A pesar de la aplicación exitosa a problemas actuales, existe una importante desconexión entre las mejores prácticas en la investigación empírica y la enseñanza típica a nivel pregrado en México (y algunos posgrados rezagados).





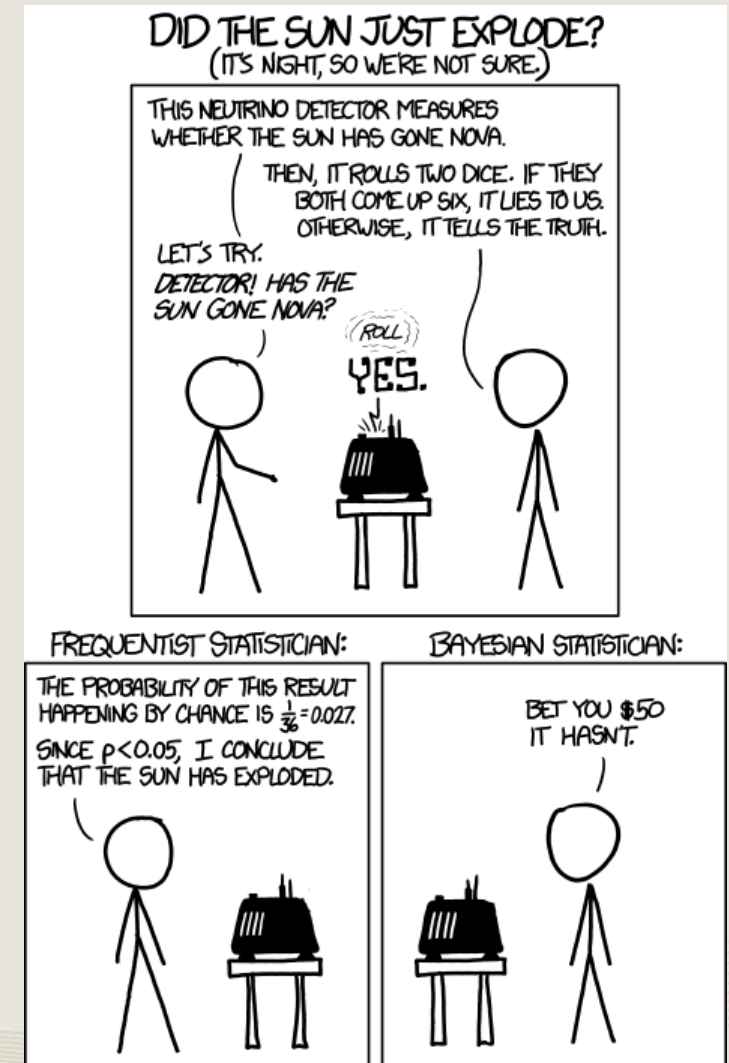


# Y es que... las cosas no siempre fueron así

- Aunque muchos de los conceptos clave que usamos hoy en día fueron desarrollados en los siglos XVII y XVIII
  - Esperanza matemática (Huygens 1657)
  - Pruebas de significancia (Arbuthnot 1711)
  - Aproximación a la binomial por la distribución normal (de Moivre 1718)
- Muchos de los primeros métodos estadísticos fueron desarrollados en la última parte del siglo XIX
  - Regresión lineal (Galton 1889)
  - Mínimos cuadrados (Legendre 1805, Gauss 1809)
  - Correlación estadística (Galton 1888, Edgeworth 1893, Yule 1897, Pearson 1896)
  - Medidas de bondad de ajuste (Pearson 1900)
- Pero el campo levantó en serio en la primera mitad del siglo XX
  - Estimación por verosimilitud (Fisher 1920, 1930)
  - Bases frecuentistas de la prueba de hipótesis (Neyman y Pearson 1933)
  - ANOVA (Yates y Cochran 1938)

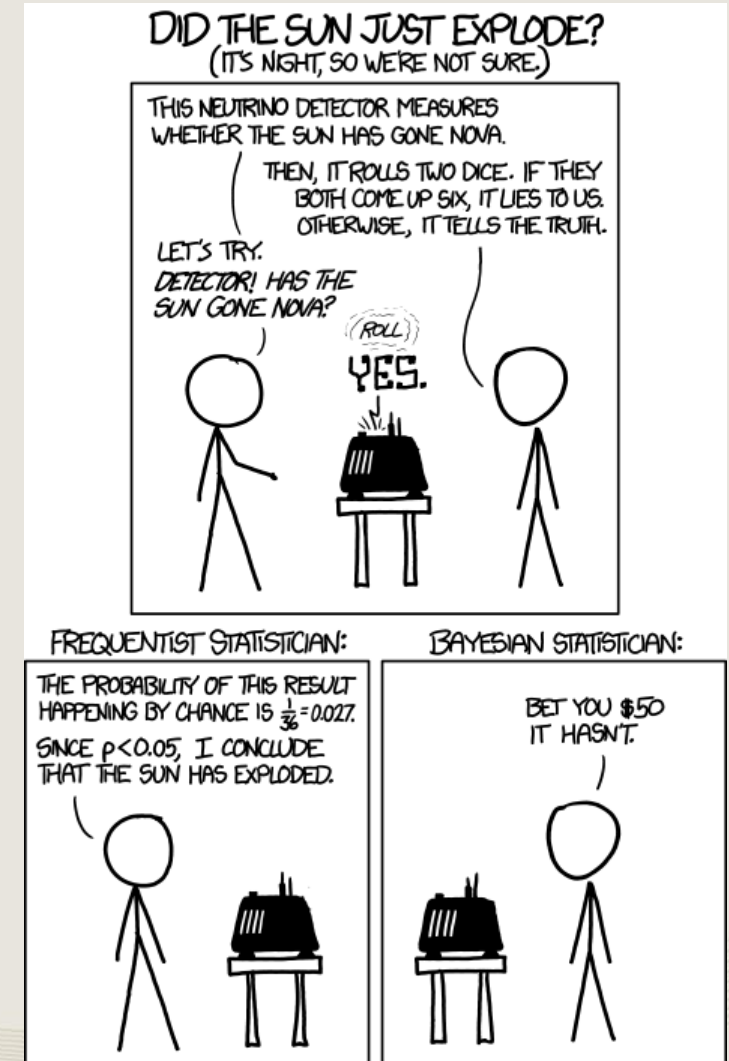
# Y es que... las cosas no siempre fueron así

- Aunque los métodos bayesianos se originan en el siglo XVII, en siglo XX los estadísticos bayesianos eran raros, excéntricos y subocupados
- Dominio de la estadística **clásica o frecuentista** en el siglo XX
  - Los datos dado  $\beta$  (un parámetro genera la realidad)
  - p-values,
  - errores estándar que no indican incertidumbre sino tamaño de muestra
  - MCO y Máxima verosimilitud



# Y es que... las cosas no siempre fueron así

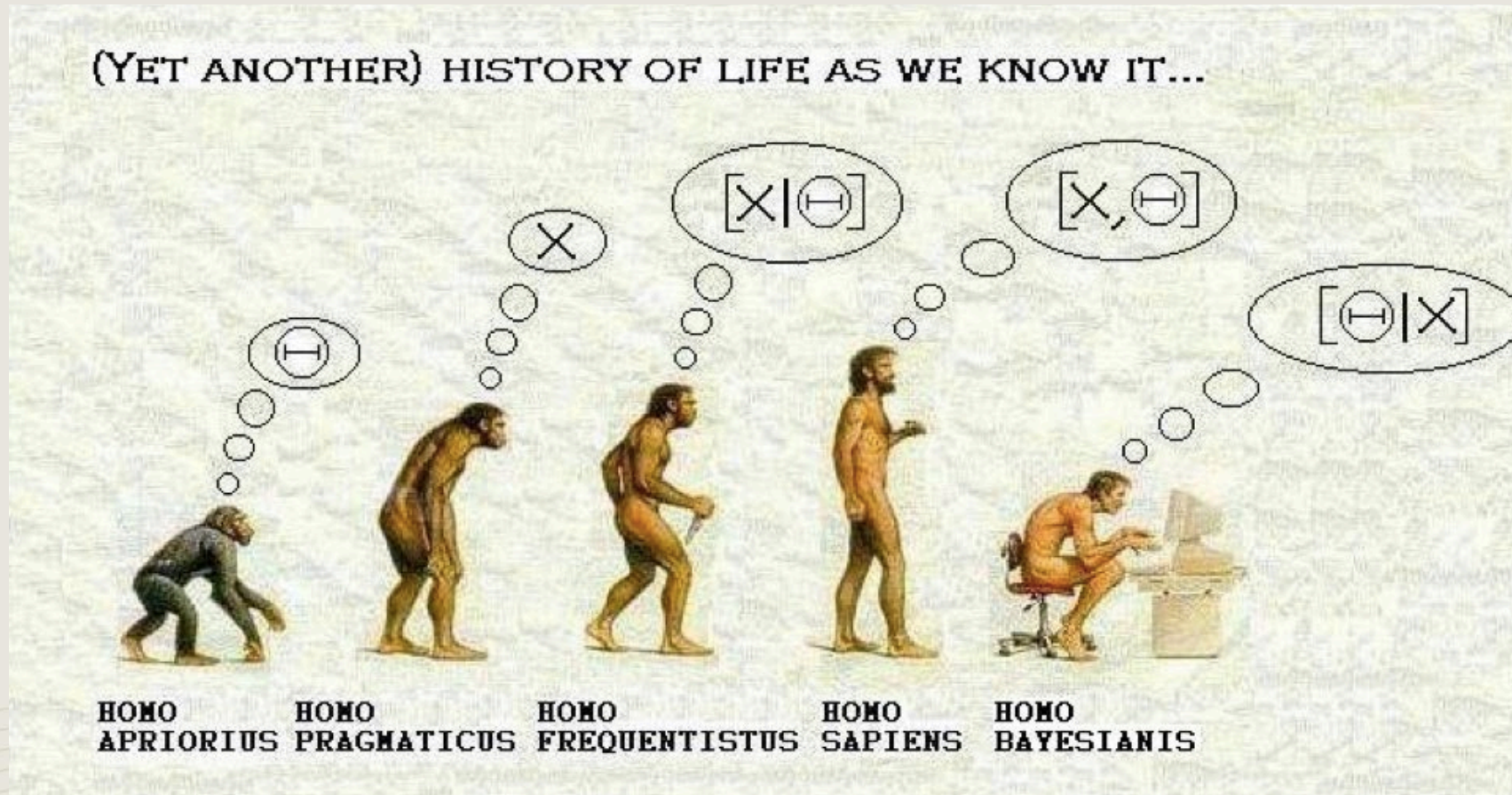
- *Fisherian statistics* (Por Fisher) es “objetiva” e inapelable: descubre la verdad.
- La inferencia basada en p-values e intervalos de confianza es válida y razonable
- No hay razones para desconfiar de la estadística clásica
- La estadística Bayesiana en tanto no es objetiva lleva a soluciones singulares
- El software estaba pensado a partir de la estadística clásica





# Análisis bayesiano

– ¿Qué originó esta revolución/evolución?



¿Qué es eso del análisis bayesiano y por qué me vengo enterando hasta ahora?





# ¿Qué es la estadística bayesiana?

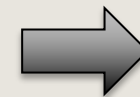
- La estadística bayesiana usa las reglas de la probabilidad y combina toda la información disponible (datos/muestra, conocimiento previo, incertidumbre) para producir inferencias más precisas respecto al caso del uso de información por separado
- En contraste, la estadística clásica evita usar información extra y se concentra en los datos. Puedes incluir en tu modelo un predictor, excluirlo o combinarlo con otros para crear una estimación más estable. Pero esto es lo único que puedes hacer.
- Puedes reproducir los métodos clásicos con inferencia bayesiana, pero con análisis bayesiano puedes hacer mucho más.

# ¿De dónde viene?

La teoría de la inferencia bayesiana se origina con el fraile Thomas Bayes.

En realidad el primer “verdadero” Bayesiano fue el matemático francés Pierre-Simon Laplace.

Laplace, P. S. (2012). *A philosophical essay on probabilities*. Courier Corporation.

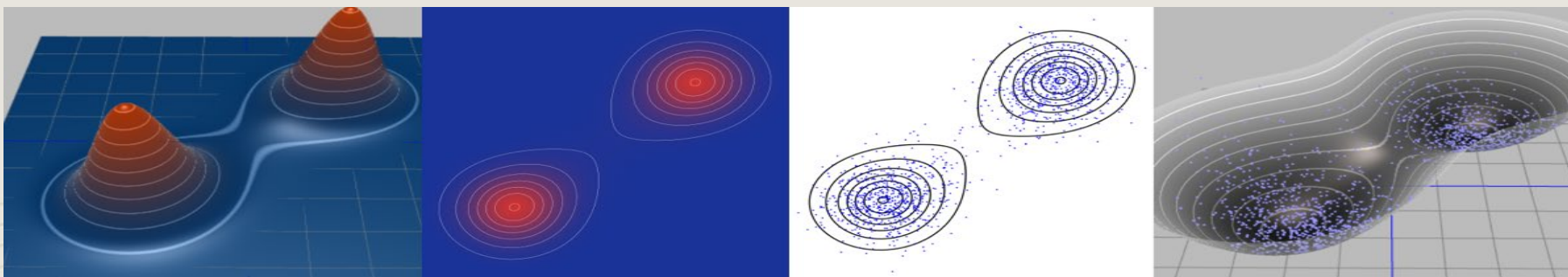


Bayesian in the flesh!

# ¿De dónde viene?

- Los métodos Bayesianos fueron usados tiempo después para resolver problemas específicos en ciencias, pero fue hasta la mitad del siglo XX que se propusieron como un marco general estadístico
- Algunas figuras clave fueron John Maynard Keynes y Frank Ramsey, quienes en los 1920s desarrollaron la teoría axiomática de la probabilidad
- Harold Jeffreys and Edwin Jaynes desarrollaron varios métodos para distintos problemas en la física
- Jimmie Savage and Dennis Lindley conectaron su investigación sobre inferencia Bayesiana en los 1950 y 1970s con la estadística clásica
- Alan Turing uso métodos probabilísticos Bayesianos para resolver el código enigma en la segunda guerra mundial

- Un factor clave que permitió la expansión de la inferencia bayesiana a finales del siglo XX fueron las mejoras en la infraestructura para el cómputo y el desarrollo de nuevos algoritmos
- Con excepción de problemas simples, la inferencia Bayesiana requiere cálculos complejos que implican problemas de altas dimensiones e integración numérica.
- La computación bayesiana usa métodos estocásticos de simulación. Esto se conoce como el método de Monte Carlo, que fue desarrollado por Stanislaw Ulam y colegas en los 1940s
- El potencial de estos métodos para resolver problemas irresolubles en estadística fue claro en los 80s. Desde entonces cada década hay saltos mayores en la sofisticación y eficiencia de los algoritmos existentes para resolver preguntas cada vez más complejas









# Somos más bayesianos que clásicos

- Considere  $\beta = 4$  IC 95 [2-6]. Entonces  $p\text{-value} < .05$
- La esencia de la estadística Bayesiana es la combinación de información de distintas fuentes. Esto se conoce como “prior information” o “información a priori”, o modelación jerárquica, actualización dinámica, o agrupación parcial.
- La estadística bayesiana junta todos los datos (información) para entender una estructura superior
- Los datos son muchos pero vienen de fuentes con error (muestreo y medición) y se necesita un marco que ponga a la inferencia en un mismo sitio.

# Algunos ejemplos

- Elecciones en USA (2020)
- Algoritmos de reconocimiento de fotografías
- Estimación de áreas pequeñas
- Las parejas *guapas* tienen más hijas...
  - ¡No! De hecho Bayes evita pifias de los métodos clásicos basádos en p-values.

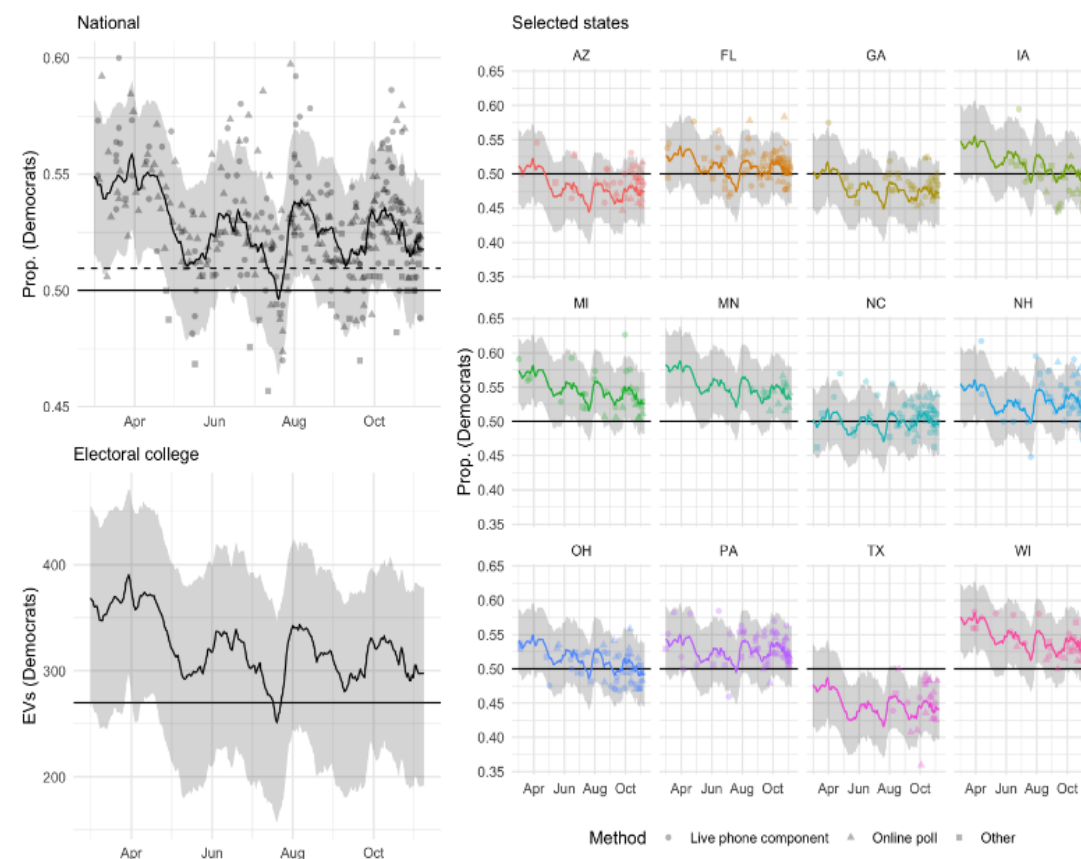
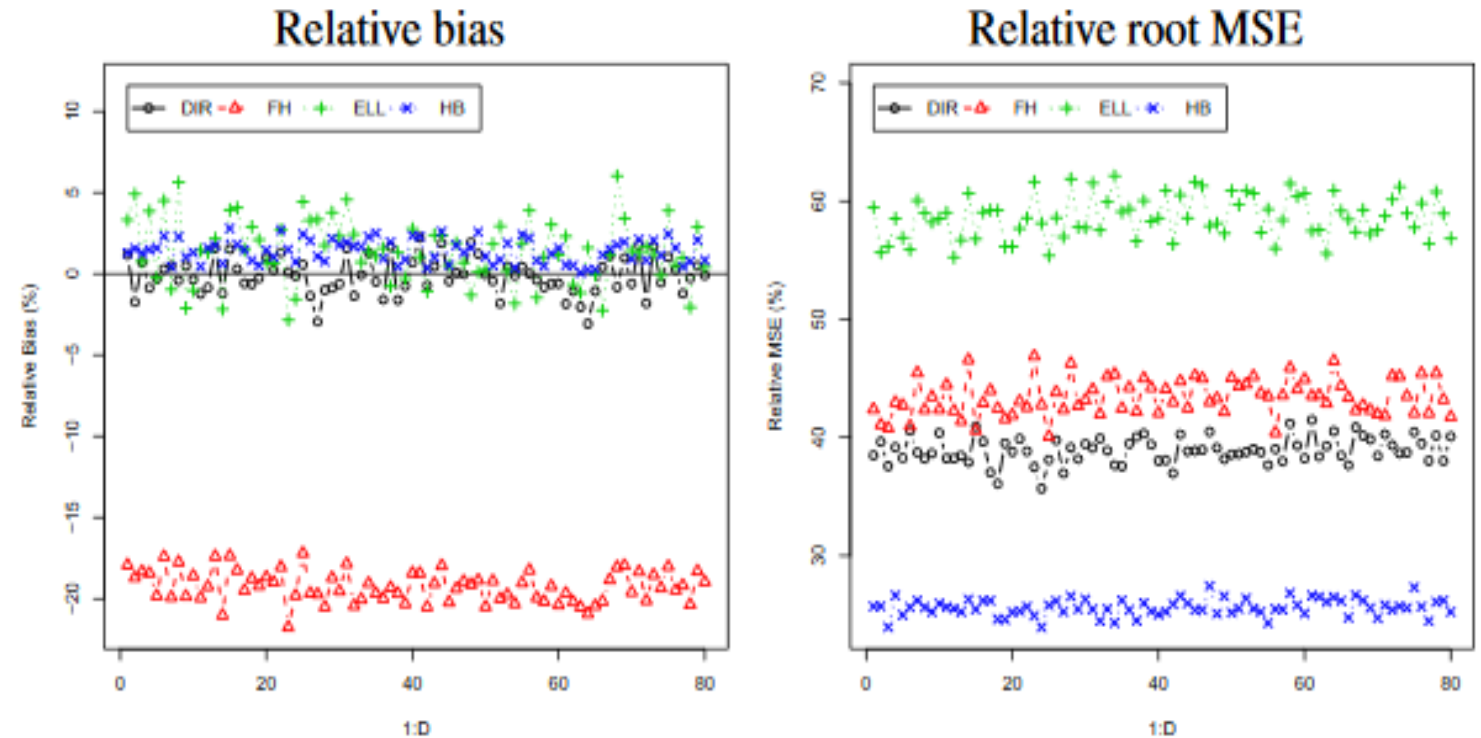


Figure 1: *Some summaries of the model, as fit retrospectively to using state and national polls from 2016. These graphs illustrate that our data and model are fitting national as well as separate state trends.*

# Estimación de áreas pequeñas

Error:

Bayes vs  
alternativas

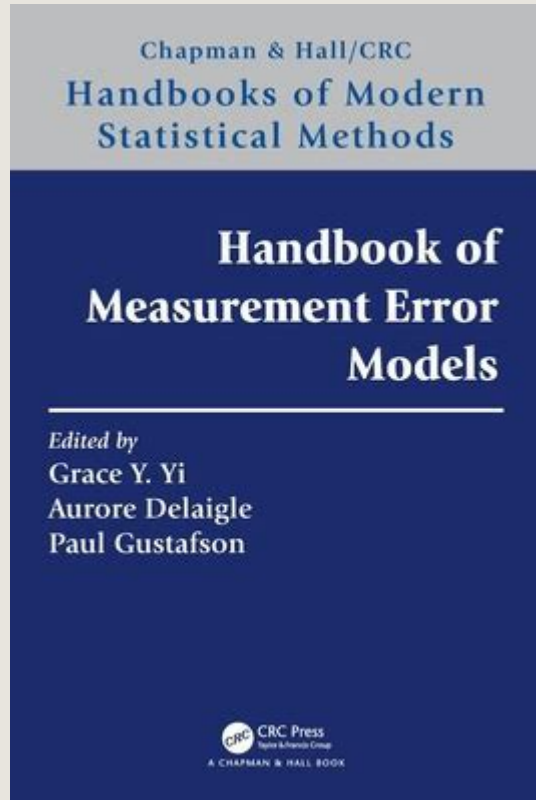


**Figure 4.** Percent relative bias (left) and relative root MSE (right) of direct, FH, HB and ELL estimators of poverty gap  $F_{1d}$  for each area  $d$  under low informativeness.



# Sus modelos categóricos o con error que no andan

number of quadrature points must be less than or equal to number of obs

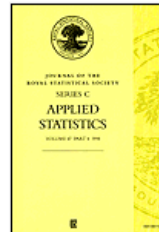


## JOURNAL ARTICLE

### Bayesian analysis of measurement error models using integrated nested Laplace approximations

Stefanie Muff, Andrea Riebler, Leonhard Held, Håvard Rue and Philippe Saner

Journal of the Royal Statistical Society.  
Series C (Applied Statistics)  
Vol. 64, No. 2  
(FEBRUARY 2015), pp.  
231-252 (22 pages)  
Published By: Wiley



<https://www.statalist.org/forums/forum/general-stata-discussion/general/1361744-t...>

#### The error of quadrature point when adding ... - Statalist

But when I run the command, Stata shows "number of quadrature points must be less than or equal to number of obs" (r198). There are totally 900 observations in my database. I really appreciate if anyone could tell me where the problem is, and if my way to add country dummy to the Random-effects logistic model is correct.

<https://www.stata.com/statalist/archive/2009-10/msg00053.html>

#### st: Time series quadrature error - Stata

number of quadrature points must be less than or equal to number of observations. I am not sure what's wrong ... I have more than 10,000 observations in the dataset.

[www.talkstats.com/threads/stata-error-r-198-quadrature-points.38698](http://www.talkstats.com/threads/stata-error-r-198-quadrature-points.38698)

#### Stata error r(198) quadrature points | Statistics Help ...

number of quadrature points must be less than or equal to number of obs. I understand that this can in part be caused as a result of an excess number of missing values. My questions are:

<https://www.stata.com/statalist/archive/2009-10/msg00056.html>

#### Re: st: Time series quadrature error - Stata

Mark Ludwig replied: > sort id date >> . tsset id date, monthly > panel variable: id (weakly balanced) > time variable: date, 2007m7 to 2008m11 > delta: 1 month >> . > . xtlogit war income male PID education L.war\_casualties L.dem\_iraq\_email > L.GOP\_iraq\_email L.dem\_iraq\_web L.GOP\_iraq\_web L.media\_iraq if tin(2007m9, 2008m3) > number of quadrature points must be less than or equal to ...

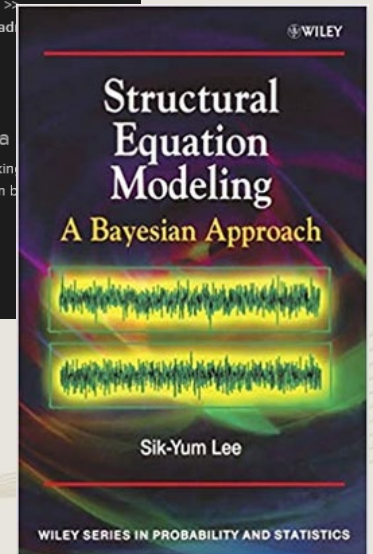
<https://www.stata.com/statalist/archive/2009-07/msg00535.html>

#### st: Re: xtlogit error "number of quadrature points ... - Stata

number of quadrature points must be less than or equal to number of obs I'm thinking something odd is going on because I have over 10 million observations as is shown by describe. . describe

<https://www.stata.com/statalist/archive/2003-10/msg00930.html>

#### st: Quadrature Points - Stata





# Hacia dónde va el análisis de datos bayesiano

- **Cómputo:** Variantes de Monte Carlo, procesamiento paralelo
- **Aplicación:** Bayesian workflow (replicabilidad, chequeos predictivos posteriores, evaluación, validación cruzada)
- **Filosofía:** Objetividad y subjetividad en el análisis bayesiano





# Objetivo

---

El objetivo del curso es introducir conceptos básicos, métodos y paquetería estadística de punta para llevar a cabo inferencia bayesiana, así como preparar al estudiante con experiencia práctica y nociones metodológicas para aplicar los métodos bayesianos a problemas reales.



# Objetivo

El objetivo del curso NO es convertirlos al bayesianismo, sino ayudarles a adquirir la claridad necesaria para distinguir las mejores herramientas para cuantificar incertidumbre cuando así lo requieran en su vida profesional.

- Dado un problema a resolver, si les parece que tiene sentido un abordaje particular, y expresan con claridad sus decisiones de modelaje teórico y estadístico, el curso habrá cumplido su objetivo.



# Expectativa

Al final del curso se espera que los alumnos sean capaces de:

- Mejorar sus habilidades de análisis de datos
- Identificar los usos apropiados e inapropiados de la estadística bayesiana
- Aquilatar las prácticas actuales de inferencia estadística y sus resultados
- Entender las ventajas de hacer investigación bajo el paradigma bayesiano
- Interpretar los resultados del análisis bayesiano de una forma crítica
- Reportar apropiadamente los resultados del análisis bayesiano
- Plantear problemas de toma de decisiones bajo incertidumbre en la tradición bayesiana
- Construir sus propios modelos para hacer inferencia probabilística en la tradición bayesiana

# Introducción al análisis bayesiano

- Es un curso introductorio sobre las formas y el análisis de datos bajo el esquema de la estadística bayesiana
- Es mucho más aplicado que teórico
- Nos basaremos en el libro de Kruschke “Doing Bayesian Data Analysis”. Second Edition
- Introduciremos la estimación de modelos bayesianos con **R-Software**

MODIFIED BAYES' THEOREM:

$$P(H|X) = P(H) \times \left( 1 + P(C) \times \left( \frac{P(X|H)}{P(X)} - 1 \right) \right)$$

H: HYPOTHESIS

X: OBSERVATION

P(H): PRIOR PROBABILITY THAT H IS TRUE

P(X): PRIOR PROBABILITY OF OBSERVING X

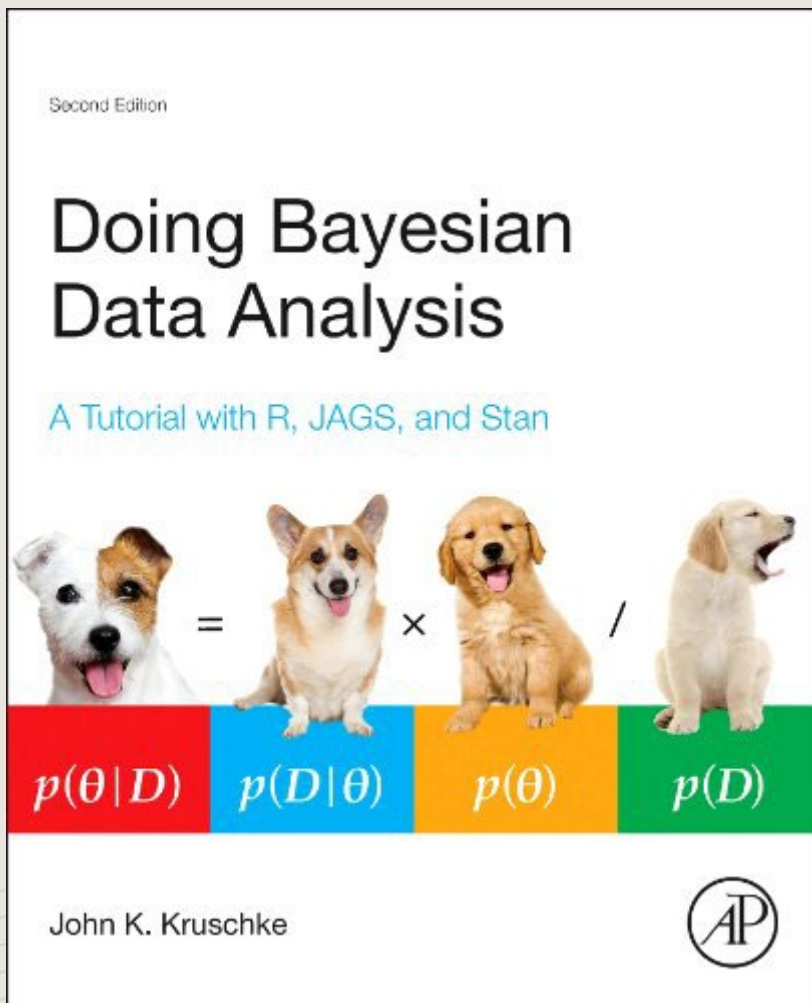
P(C): PROBABILITY THAT YOU'RE USING  
BAYESIAN STATISTICS CORRECTLY



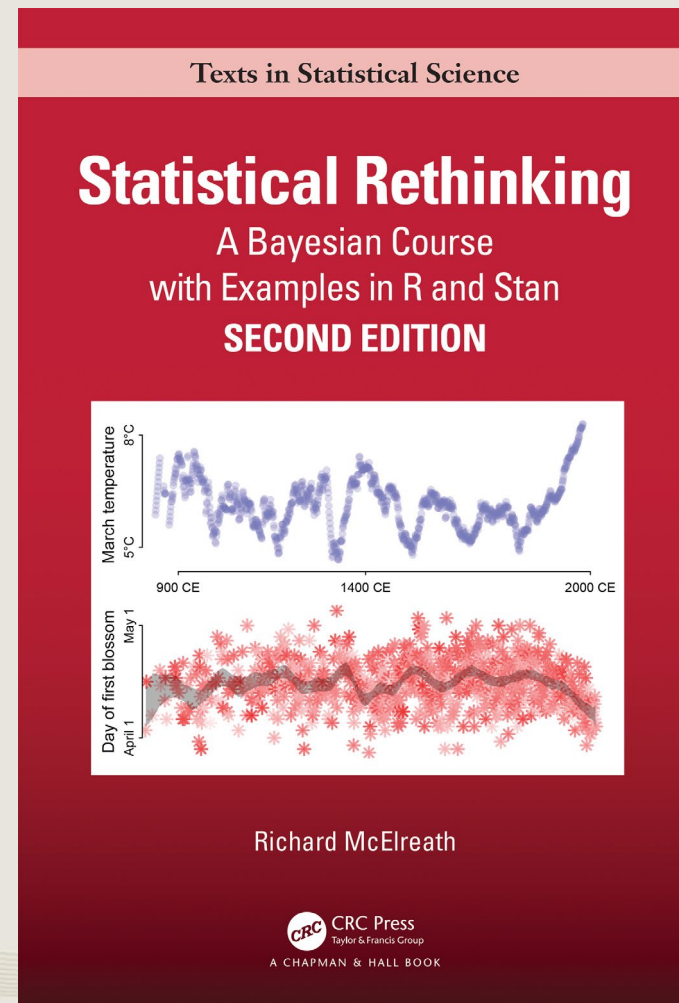
# Características de las sesiones

- Las sesiones combinan discusión, teoría y aplicación con el programa R.
- Antes de cada clase, los alumnos deberán leer una selección de capítulos/artículos para su discusión en grupo.
- Los docentes impartirán cada sesión (prepararán un archivo.ppt que subirán a Github después de cada clase) y se dedicará siempre un espacio para discusión, ejercicios en grupo y/o implementación de análisis usando el programa R.

## Básica



## Secundaria





# Evaluación

Se utilizarán dos ejercicios teórico-prácticos para valorar los contenidos que los alumnos manejan con confianza y aquellos que necesitan reforzarse.

- Nuestra prioridad es que aprendan y les sea útil el curso. Su aprendizaje tendrá una expresión numérica no tanto basada en resultados sino en el proceso.
- Los ejercicios nos indicarán la evolución del grupo e individual.



# Temario (General)

## 1. Conceptos básicos: modelos, probabilidad, regla de Bayes, y R

Credibilidad, modelos y parámetros; ¿qué es esa cosa llamada probabilidad?; regla de Bayes; el ambiente R (*r-project*)

## 2. Fundamentos aplicados a la inferencia de probabilidades binomiales

Inferir probabilidades binomiales; técnicas de Monte Carlo vía cadenas de Markov; modelos jerárquicos; abordaje bayesiano a las pruebas de significancia de hipótesis nulas; abordajes bayesianos a la prueba de hipótesis puntuales

## 3. Fundamentos de modelación bayesiana

Combinaciones lineales y predictores; regresión lineal simple, múltiple y logística; calibración bayesiana en modelos de regresión; evaluación bayesiana del modelo de regresión; expansión bayesiana del modelo de regresión; flujo de trabajo bayesiano





# Materiales: Github y drive

Los docentes utilizarán esta plataforma para compartir los materiales del curso (bibliografía, presentaciones, ejercicios). La dirección relevante es:

<https://github.com/hectornajera83/ClaseBayesUNAM>

## Lugar y hora

16 sesiones

Intentaremos no pasar de las 2:30 de clase

Intermedio de 10 minutos



## Próxima clase

---

Esencial:

–Kruschke Capítulo 2: Introduction, credibility models and parameters



# Próxima clase

## — Recomendada

10.1098/rsta.2003.1263



### Bayesian computation: a statistical revolution

BY STEPHEN P. BROOKS

*Statistical Laboratory, University of Cambridge, Centre for Mathematical Sciences,  
Wilberforce Road, Cambridge CB3 0WB, UK (steve@statslab.cam.ac.uk)*

*Published online 3 November 2003*

The 1990s saw a statistical revolution sparked predominantly by the phenomenal advances in computing technology from the early 1980s onwards. These advances enabled the development of powerful new computational tools which reunited



# Próxima clase

---

## — Recomendada

- Lennox, K. [Lawrence Livermore National Laboratory]. (2016, September 27). *All About that Bayes: Probability, Statistics, and the Quest to Quantify Uncertainty* [Video]. YouTube.  
<https://youtu.be/eDMGDhyDxuY>





# Referencias

- Andrew Gelman (2011), “Induction and Deduction in Bayesian Data Analysis“, Special Topic: Statistical Science and Philosophy of Science RMM Vol. 2, 2011, 67–78
- Brooks, S. P. (2003). Bayesian computation: a statistical revolution. Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences, 361(1813), 2681-2697.
- Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis*. CRC press.
- Hernández, D.R. 2007. *Introducción al análisis bayesiano*. Mar del Plata : Instituto Nacional de Investigación y Desarrollo Pesquero INIDEP. 45 p.
- Kruschke, J. (2014). *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan*. Academic Press.
- Laplace, P. S. (2012). *A philosophical essay on probabilities*. Courier Corporation.
- McElreath, R. (2020). *Statistical rethinking: A Bayesian course with examples in R and Stan*. CRC press.



# CONTACTO

Dres. Héctor Nájera y Curtis Huffman  
Investigadores (SNI-I)

Programa Universitario de Estudios del Desarrollo (PUED)  
Universidad Nacional Autónoma de México (UNAM)

Antigua Unidad de Posgrado (costado sur de la Torre II de Humanidades), planta baja.  
Campus Central, Ciudad Universitaria, Ciudad de México, México.

Tel. (+52) 55 5623 0222, Ext. 82613 y 82616

Tel. (+52) 55 5622 0889

Email: [hecatalan@hotmail.com](mailto:hecatalan@hotmail.com)  
[chuffman@unam.mx](mailto:chuffman@unam.mx)

*¡Bienvenidos  
estudiantes!*

