

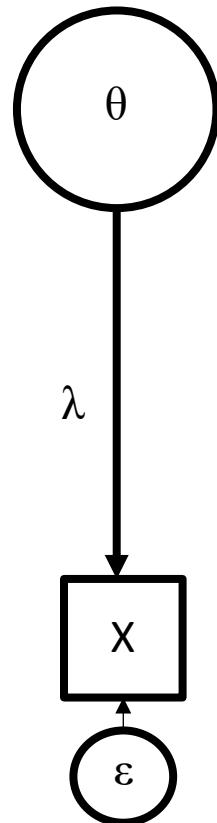
**Temas adicionales de SEM:**

# Modelación generativa

- SEM
- Inferencia bayesiana
- Algunas implementaciones de ML (Tienen algún nivel de asistencia)

# Modelación generativa

- Los modelos teóricos bien definidos son generativos:
  - Los parámetros (estimadores) son realizaciones de un conjunto de relaciones causales
  - Los parámetros (¿estimadores?) son variables aleatorias cuya realización es condicional a otros parámetros
  - En medición, los modelos de SEM/AFC suponen que los indicadores son producto de un modelo subyacente



# ¿Cuándo es importante migrar a SEM?

- ¿Tienes un modelo generativo?
  - ¿Cómo los parámetros son condicionales a supuestos y a los datos?
- Tienes por lo menos una variable latente (No se ubica obviamente en los observables)
- Modelo de medición (SEM 1): Errores de medición
- Modelo causal o múltiples relaciones dependientes (SEM 2)
- Modelo de clasificación/grupos (SEM 2)
- Problemas de multicolinealidad (Descomposición de varianza)

# Variables latentes



## Latent Variable Modelling: A Survey\*

ANDERS SKRONDAL, SOPHIA RABE-HESKETH

First published: 05 December 2007 | <https://doi.org/10.1111/j.1467-9469.2007.00573.x> | Citations: 89

✉ Anders Skrondal, Department of Statistics, London School of Economics, Houghton Street, London WC2A 2AE, UK.

E-mail: [a.skrondal@lse.ac.uk](mailto:a.skrondal@lse.ac.uk)

\* This paper was presented at the 21st Nordic Conference on Mathematical Statistics, Rebild, Denmark, June 2006 (NORDSTAT 2006).

[Read the full text >](#)

PDF TOOLS SHARE

### Abstract

**Abstract.** Latent variable modelling has gradually become an integral part of mainstream statistics and is currently used for a multitude of applications in different subject areas. Examples of ‘traditional’ latent variable models include latent class models, item-response models, common factor models, structural equation models, mixed or random effects models and covariate measurement error models. Although latent variables have widely different interpretations in different settings, the models have a very similar mathematical structure. This has been the impetus for the formulation of general modelling frameworks which accommodate a wide range of models. Recent developments include multilevel structural equation models with both continuous and discrete latent variables, multiprocess models and nonlinear latent variable models.

[Home](#) > [Behaviormetrika](#) > Article

## Beyond SEM: General Latent Variable Modeling

Open access | Published: 15 January 2002

Volume 29, pages 81–117, (2002) [Cite this article](#)

[Download PDF](#)

You have full access to this [open access](#) article



### Behaviormetrika

[Aims and scope](#) →

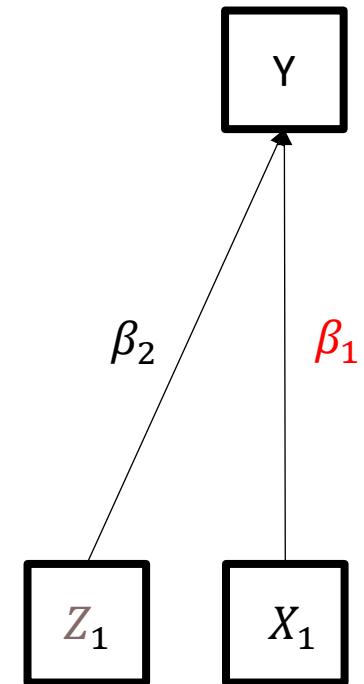
[Submit manuscript](#) →

[Use our pre-submission checklist](#) →

Avoid common mistakes on your manuscript.

# Modelación generativa

- La ecuación de Mincer:
  - $Y = \alpha + \beta_1 X_1 + \beta_2 Z_1 + e$
  - $Y$  Salario
  - $X_1$  Inversión en capital humano ~ Años de educación
  - $Z_1$  Confusores

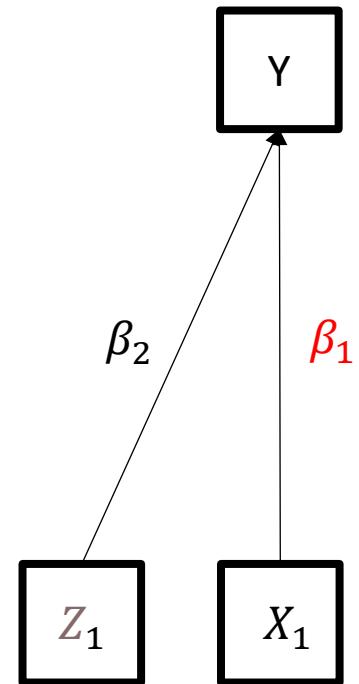


# Modelación generativa

- La ecuación de Mincer:
  - $Y = \alpha + \beta_1 X_1 + \beta_2 Z_1 + e$
  - $Y$  Salario
  - $X_1$  Inversión en capital humano ~ Años de educación
  - $Z_1$  Confusores

¿Capital humano?

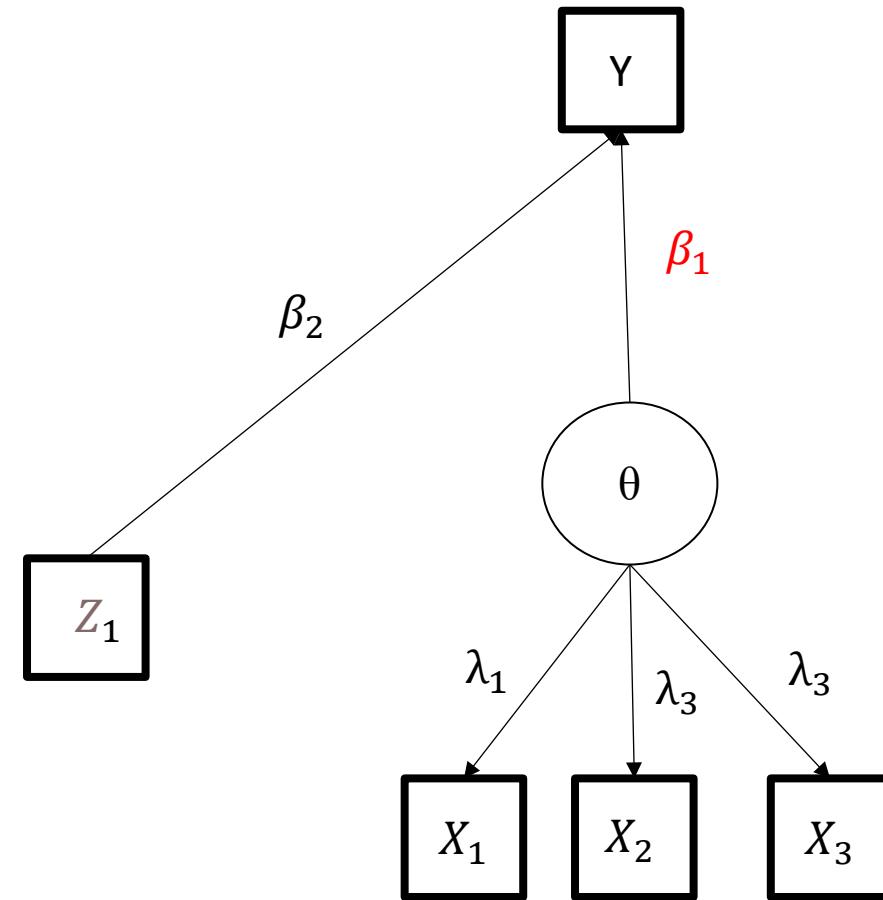
Parámetro/variable latente  $\theta$   $f=\{X\}$



# Modelación generativa

- La ecuación de Mincer:

- $Y = \alpha + \beta_1 \theta + \beta_2 Z_1 + e$
- $X_1 = \alpha_1 + \lambda_1 \theta + \varepsilon_1$
- $X_2 = \alpha_2 + \lambda_2 \theta + \varepsilon_2$
- $X_3 = \alpha_3 + \lambda_3 \theta + \varepsilon_3$
- $Y$  Salario
- $X_1$  Años de educación
- $X_2$  Nivel nutricional en la infancia
- $X_3$  Acceso a la salud
- $Z_1$  Confusores

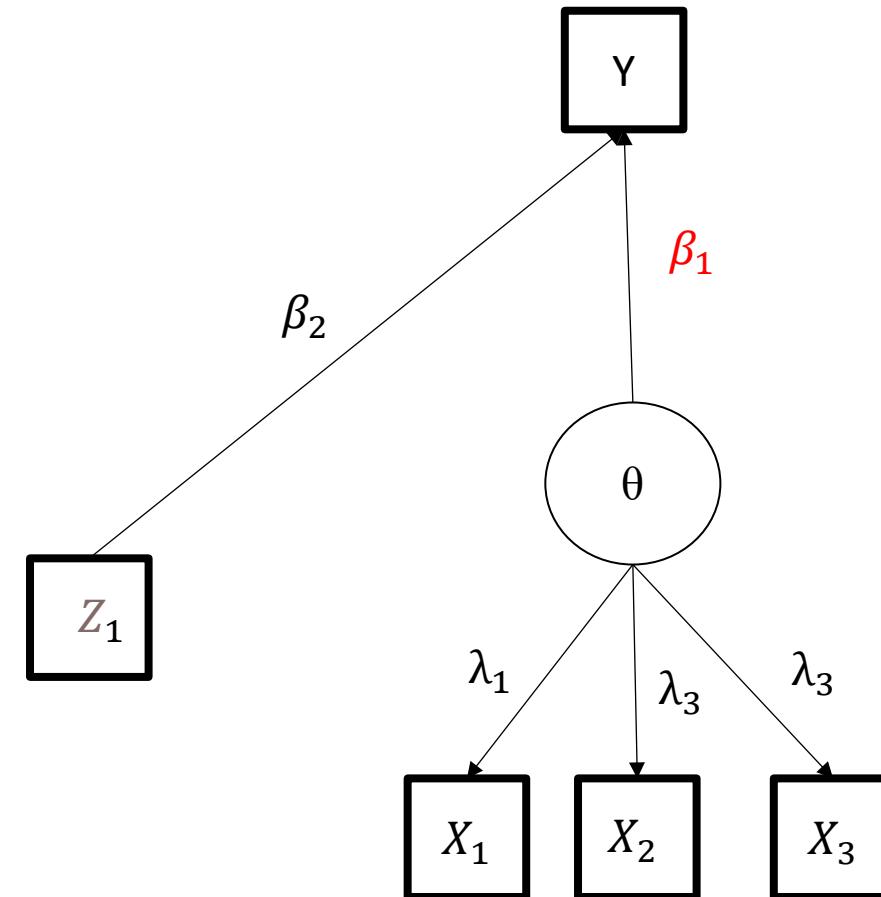


# Modelación generativa

- La ecuación de Mincer:

- $Y = \alpha + \beta_1\theta + \beta_2Z_1 + e$
- $X_1 = \alpha_1 + \lambda_1\theta + \varepsilon_1$
- $X_2 = \alpha_2 + \lambda_2\theta + \varepsilon_2$
- $X_3 = \alpha_3 + \lambda_3\theta + \varepsilon_3$

- $Y$  Salario
- $X_1$  ¿Años de educación? ¿Calidad?
- $X_2$  Nivel nutricional en la infancia
- $X_3$  Acceso a la salud
- $Z_1$  Confusores

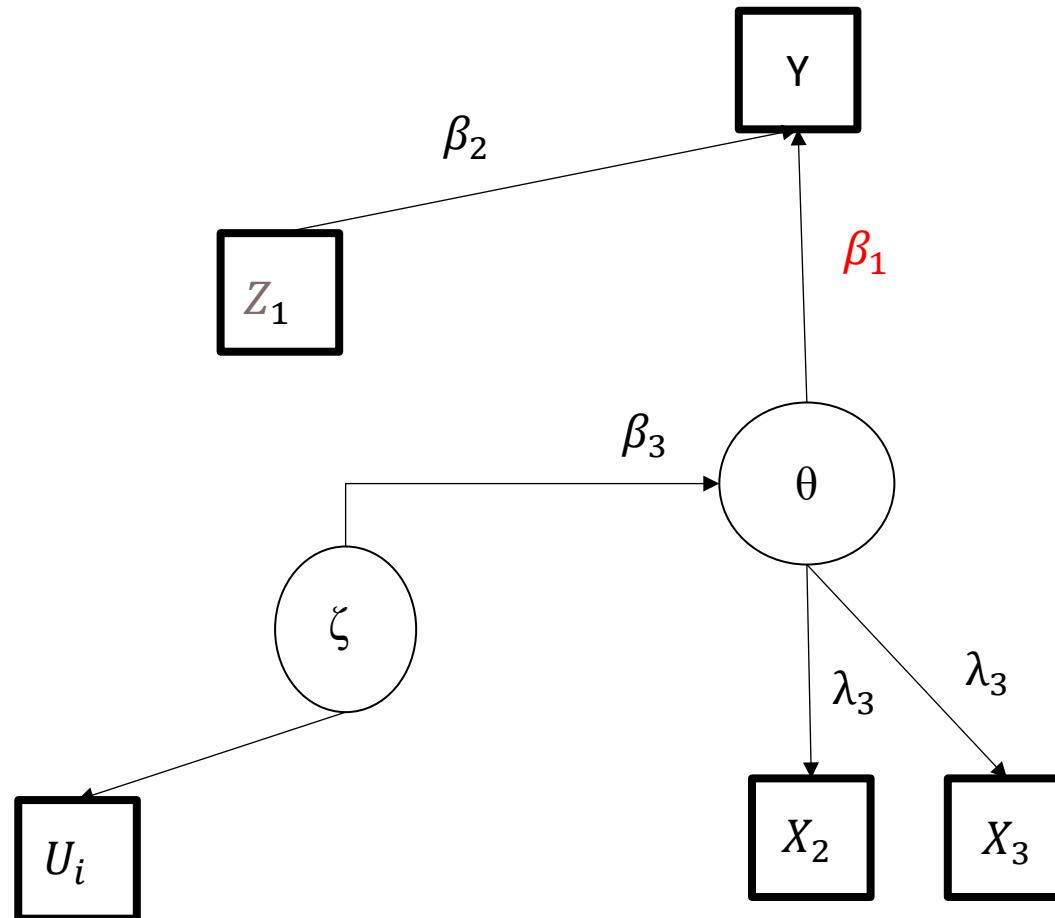


# Modelación generativa

- La ecuación de Mincer:

- $Y = \alpha + \beta_1 \theta + \beta_2 Z_1 + e$
- $\theta = \alpha + \beta_3 \zeta$  (Clases latentes de habilidad)
- $X_2 = \alpha_2 + \lambda_2 \theta + \varepsilon_2$
- $X_3 = \alpha_3 + \lambda_3 \theta + \varepsilon_3$

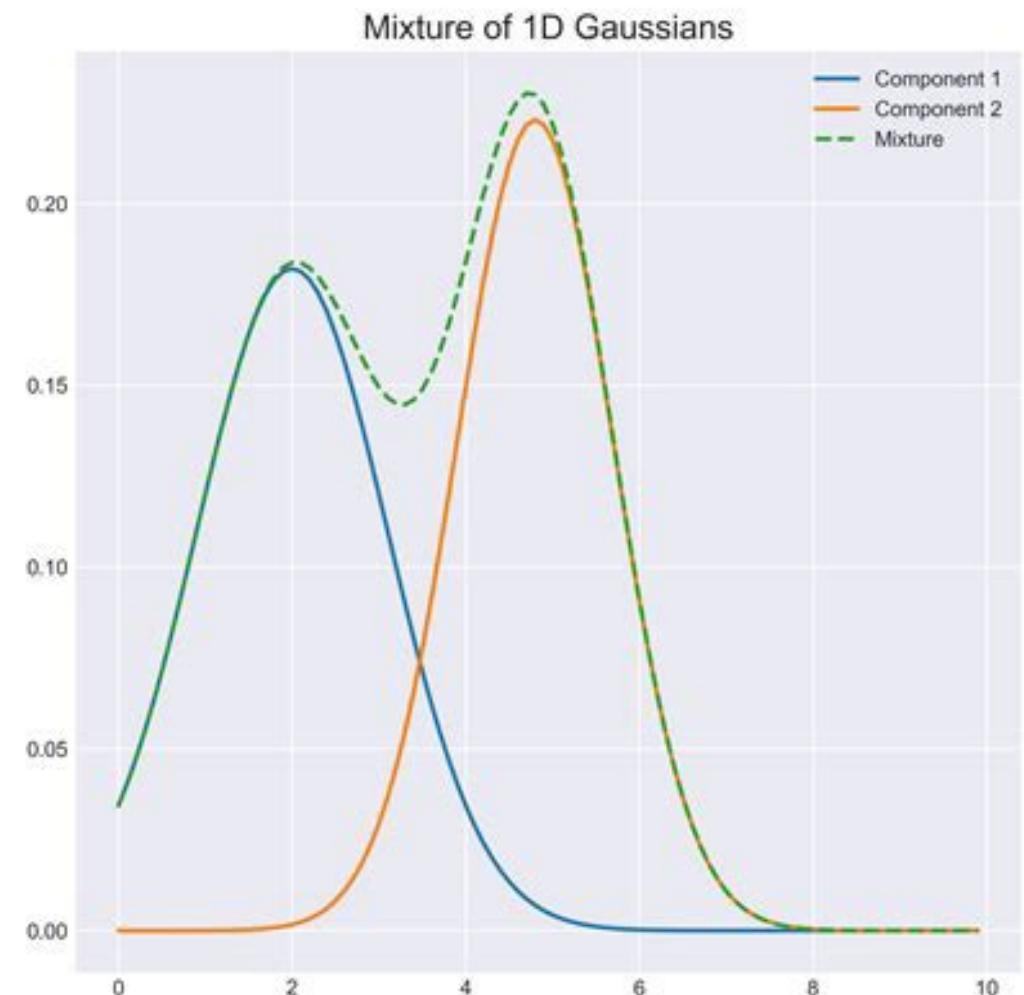
- $Y$  Salario
- $U_i$  Habilidad
- $X_2$  Nivel nutricional en la infancia
- $X_3$  Acceso a la salud
- $Z_1$  Confusores



# Clasificación

¿Cuántos grupos?

¿Cuál es la mejor separación posible?



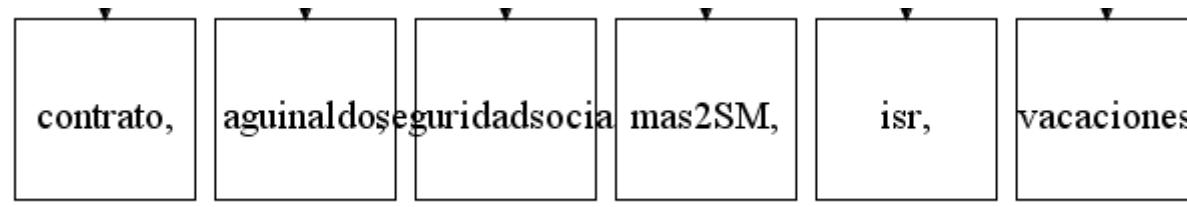
# Mixture modeling:

- Latent profile analysis: indicadores continuos
- Latent classs analysis: indicadores discretos

Supuestos:

- Hay un fenómeno subyacente
- Hay grupos distintivos del fenómeno subyacente
- Manifestaciones del fenómeno subyacente
- Es la pertenencia a cierto grupo (severidad del fenómeno) la que explica la probabilidad de respuesta a cada indicador

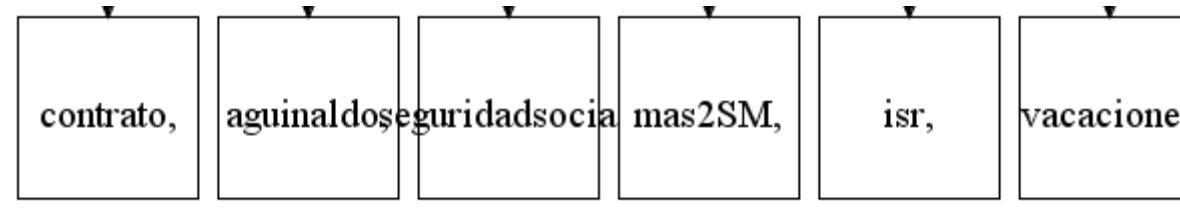
# Sin mixture modelling



Escala de 0-6

Formal >4  
Informal <=3

# Sin mixture modelling



Escala de 0-6

Formal >4  
Informal <=3

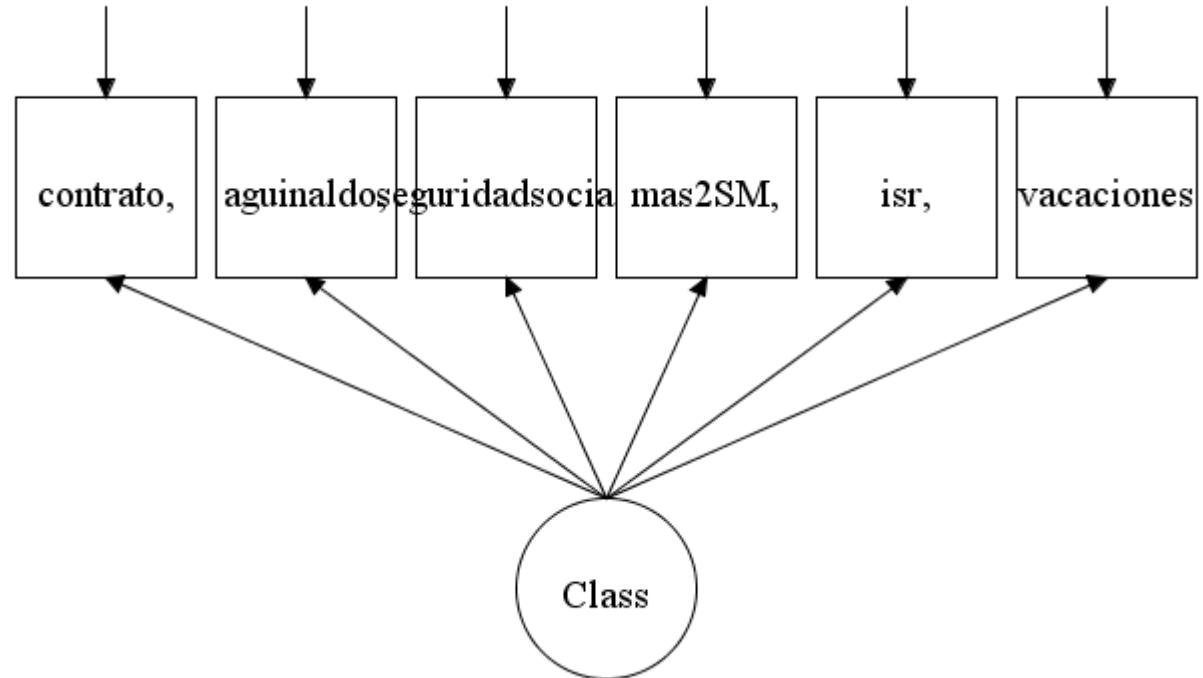
Logit con X's

# Con mixture modeling:

Hay un fenómeno  
subyacente: **formalidad**

Hay dos grupos  
**formales e informales**

Es la membresía al  
grupo (ser formal o  
informal) lo que  
determina tener o  
carece de las  
condiciones siguientes

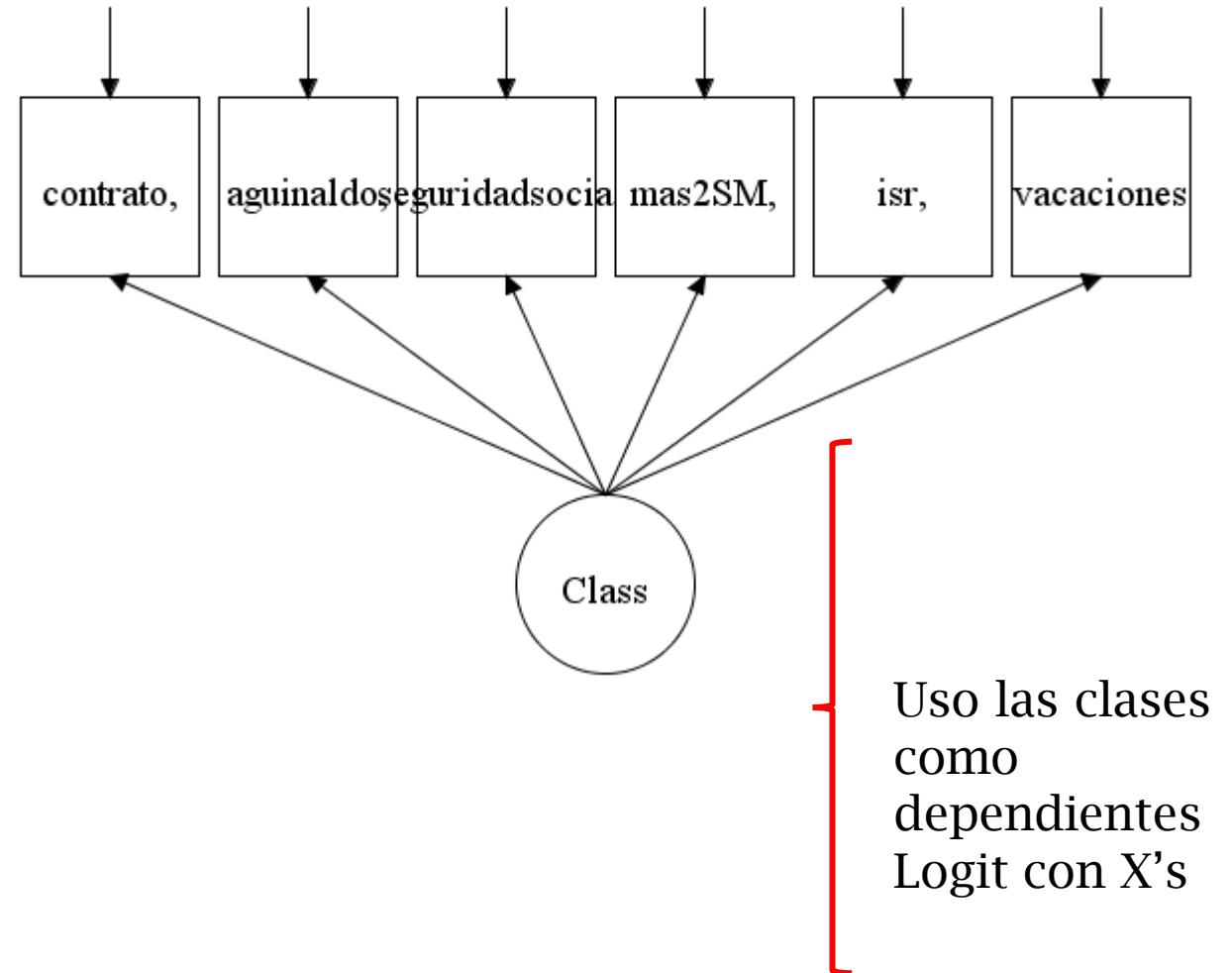


# Con mixture modeling:

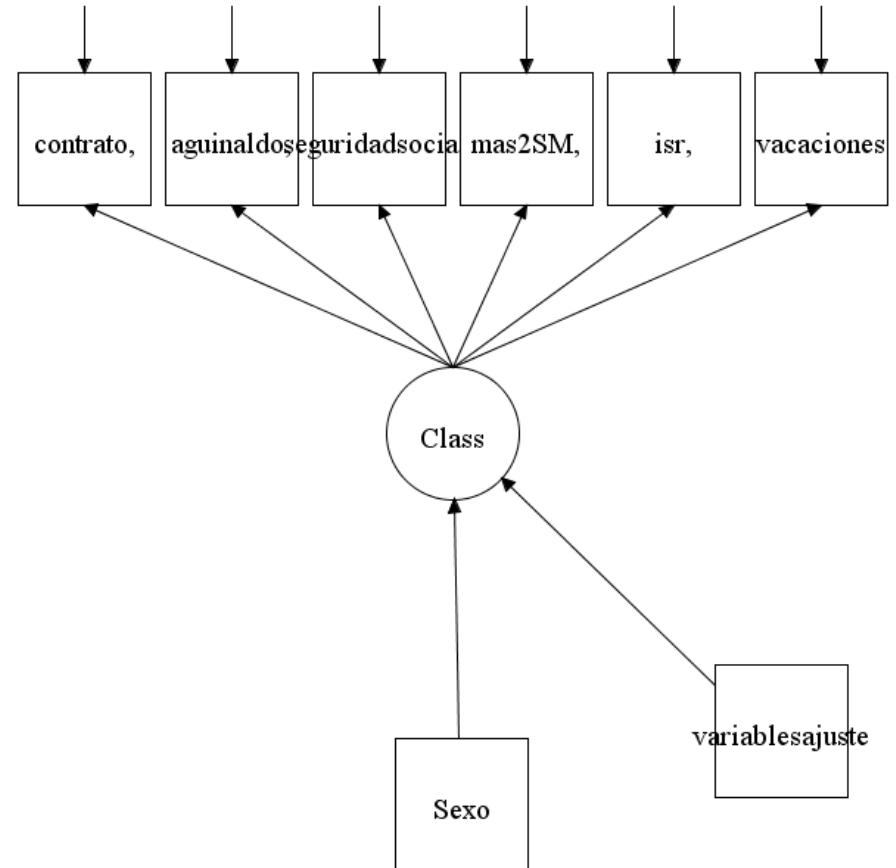
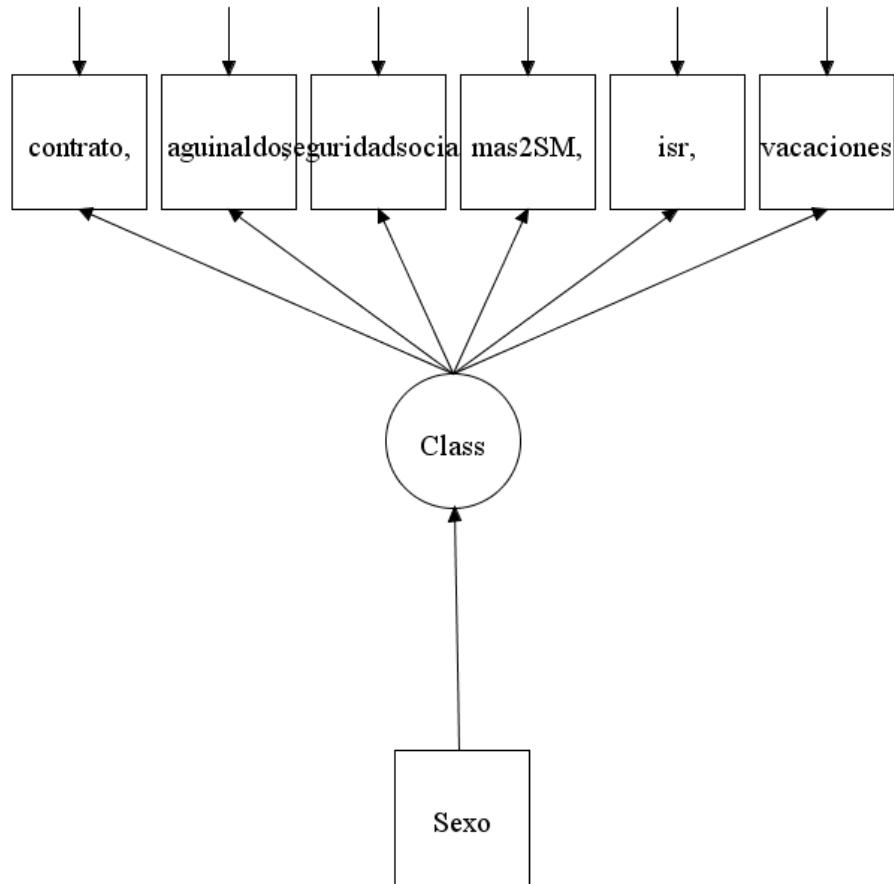
Hay un fenómeno subyacente: formalidad

Hay dos grupos formales e informales

Es la membresía al grupo (ser formal o informal) lo que determina tener o carece de las condiciones siguientes



# LCA condicional (confirmatorio):



Ho: La pertenencia a cierto grupo es condicional a cierto factor

# Lecturas de LCA en la práctica

## Latent Class Analysis: A Guide to Best Practice

Bridget E. Weller  , Natasha K. Bowen, and Sarah J. Faubert [View all authors and affiliations](#)

Volume 46, Issue 4 | <https://doi.org/10.1177/0095798420930932>

Contents

 PDF / ePub

 Cite article

 Share options

 Information, rights and permission:

## Abstract

Latent class analysis (LCA) is a statistical procedure used to identify qualitatively different subgroups within populations who often share certain outward characteristics. The assumption underlying LCA is that membership in unobserved groups (or classes) can be explained by patterns of scores across survey questions, assessment indicators, or scales. The application of LCA is an active area of research and continues to evolve. As more researchers begin to apply the approach, detailed information on key considerations in conducting LCA is needed. In the present article, we describe LCA, review key elements to consider when conducting LCA, and provide an example of its application.

## Teacher's Corner

## Recommended Practices in Latent Class Analysis Using the Open-Source R-Package tidySEM

C. J. Van Lissa  , M. Garnier-Villarreal  & D. Anadria 

Received 21 Apr 2023, Accepted 18 Aug 2023, Published online: 09 Oct 2023

 Cite this article  <https://doi.org/10.1080/10705511.2023.2250920> 

 Full Article

 Figures & data

 References

 Supplemental

 Citations

 Metrics

 Licensing

 Reprints & F

 View PDF

 View EPUB

# Paquetes en R

- OpenMX + tidySEM
- polCA

Para análisis más en forma y más complejos: Mplus <https://www.statmodel.com/>



# Conclusiones LCA

- Siempre que tenga en cuestión algún tipo de agrupación:
  - Grupos latentes de personas o de unidades territoriales
- Dirime disputas del siglo pasado sobre dónde poner la línea
- En un mismo marco atiende el problema de agrupación y explicación de los perfiles de los grupos

Próximo año

# Conceptual Grounding for Bayesian Inference for Latent Variables in Factor Analysis

Roy Levy  

Pages 195-214 | Published online: 02 Nov 2022

 Cite this article <https://doi.org/10.1080/15366367.2021.1996819> Full Article Figures & data References Citations Metrics Reprints & Permissions

Read this article

## ABSTRACT

Obtaining values for latent variables in factor analysis models, also referred to as factor scores, has long been of interest to researchers. However, many treatments of factor analysis do not focus on inference about the latent variables, and even fewer do so from a Bayesian perspective. Researchers may therefore be ill-acquainted with Bayesian thinking on this issue, despite the fact that certain existing procedures may be seen as Bayesian to some extent. The focus of this paper is to provide a conceptual grounding for Bayesian inference for latent variables, articulating not only what Bayesian inference has to say about values for latent variables, but why Bayesian inference is suited for this problem. As to why, it is argued that the notion of exchangeability motivates the form of factor analysis, as well as Bayesian inference for

Rela

Pe

Scale  
A LaTen  
Meas  
Persp  
Publi

# Más allá de la hipótesis nula

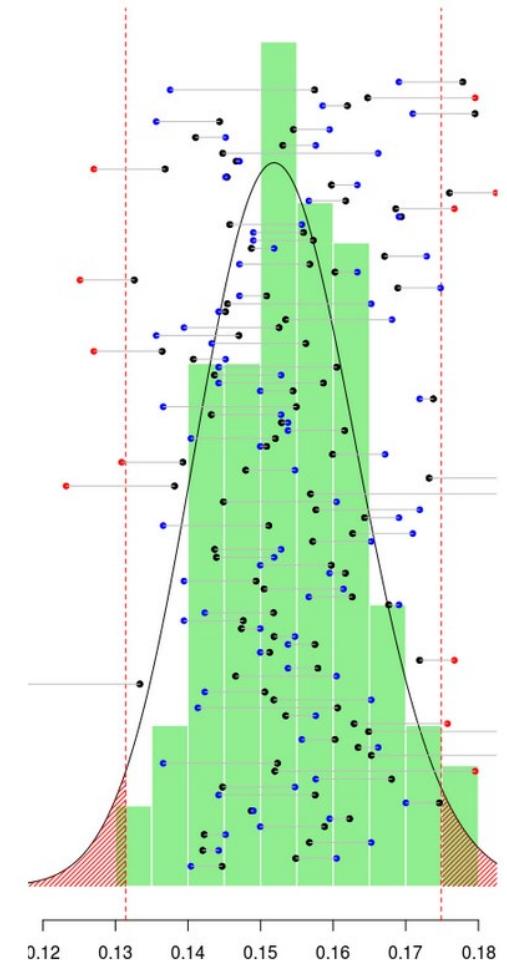
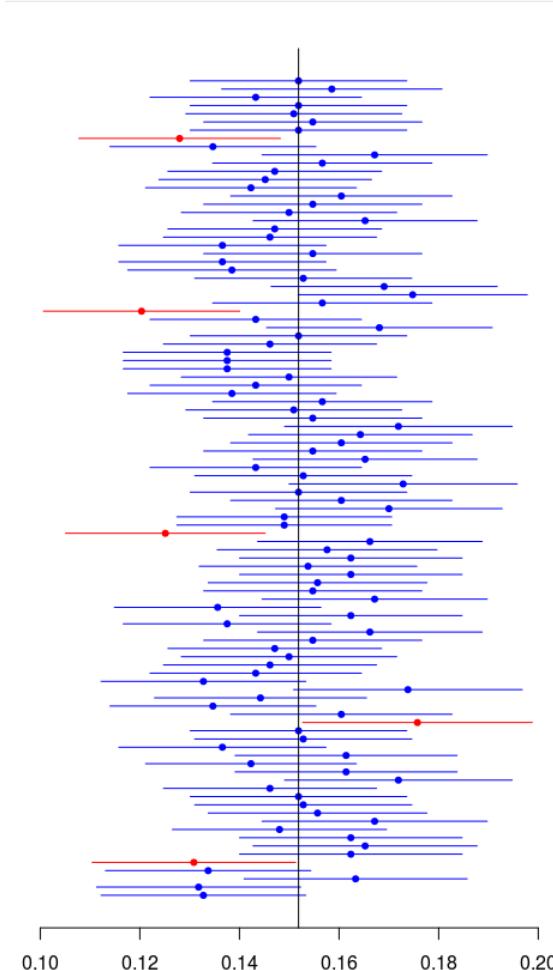
Beta = 1 [.5 – 1.2]

Es altamente probable (95%) que el efecto de interés se ubique entre .5 y 1.2.

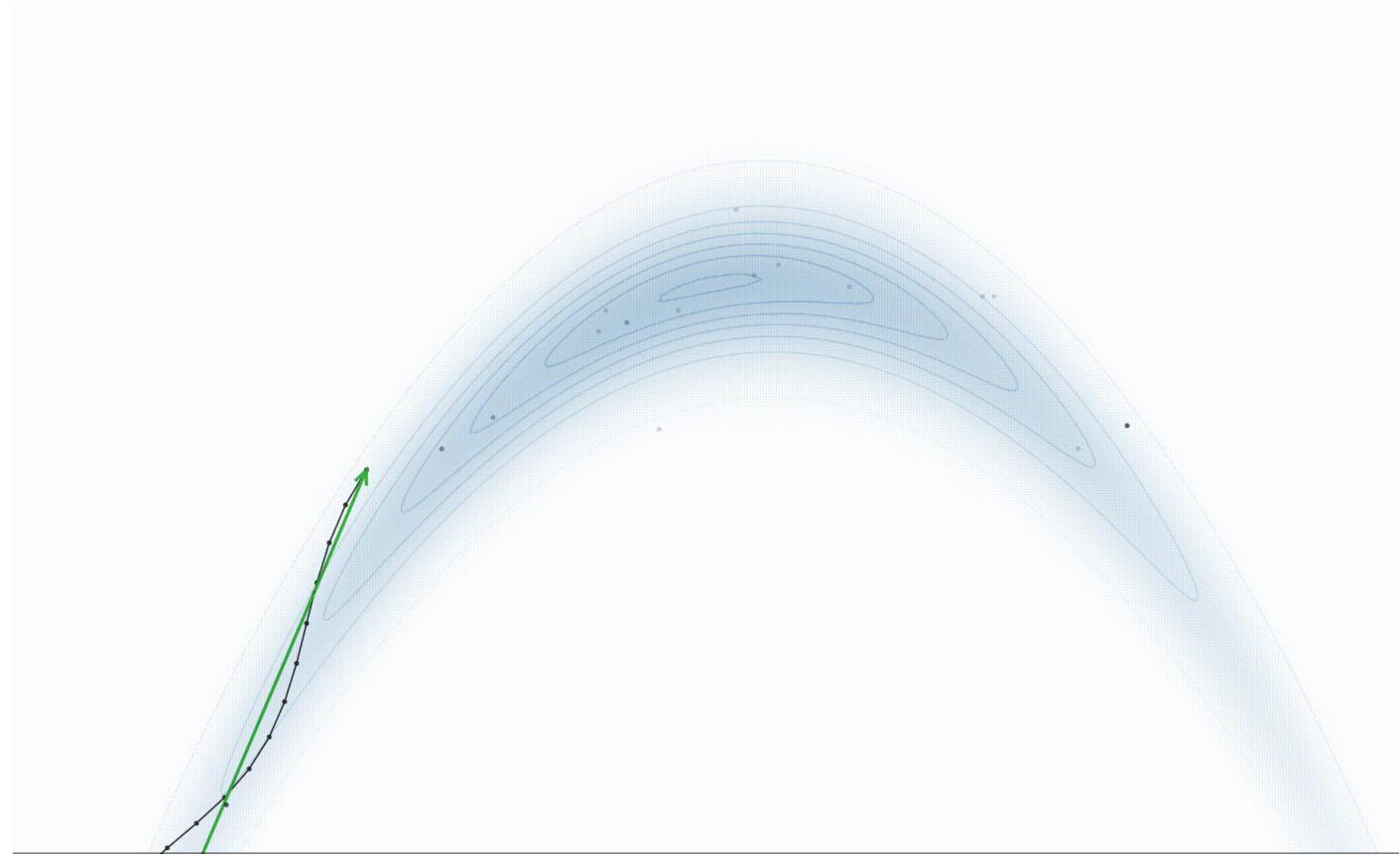
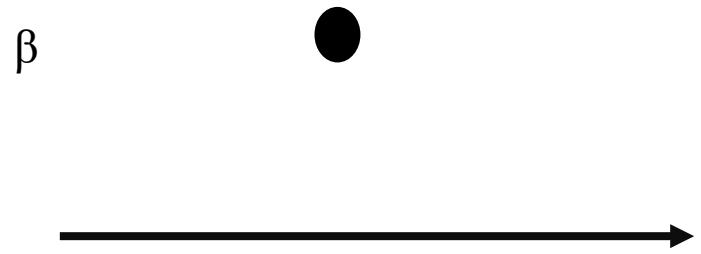
No puedo hacer este tipo de inferencia

95% de los intervalos de confianza contienen a 1!

SEM + inferencia bayesiana



# Punto vs distribuciones



# Problemas de cómputo e información previa

## Cómputo

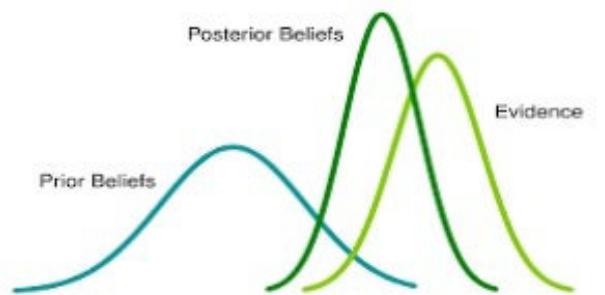
- La complejidad de algunos modelos reduce la factibilidad de su estimación
- Hay que moverse a cómputo bayesiano

## Inferencia e información

- La hipótesis nula y los p-values les hacen poco sentido (descubrimos rechazando la posibilidad de que beta sea un producto de un proceso aleatorio)
- Hay información que te gustaría incorporar en el modelo:
  - Información sobre el efecto de ciertas variables encontrado en otros estudios
  - Información sobre el proceso de pérdida de datos
- Bajo poder en el estudio

Si la respuesta es sí a alguna de estas condiciones:

## BAYESIAN INFERENCE



# Software

- SEM. Modelos convencionales: Pocos parámetros, estructuras razonablemente simples, corte transversal
  - Lavaan, MPLUS, stata, AMOS. ML
- SEM. Modelos convencionales pero con datos provenientes de muestras complejas
  - MPLUS. ML
- SEM Modelos complejos: Muchos parámetros, grandes datos, datos panel
  - MPLUS. BAYES
- SEM. Modelos complejos con distintas distribuciones
  - MPLUS y STAN. BAYES
- SEM. Modelos complejos con distintas distribuciones y efectos espaciales
  - STAN. BAYES



RStan

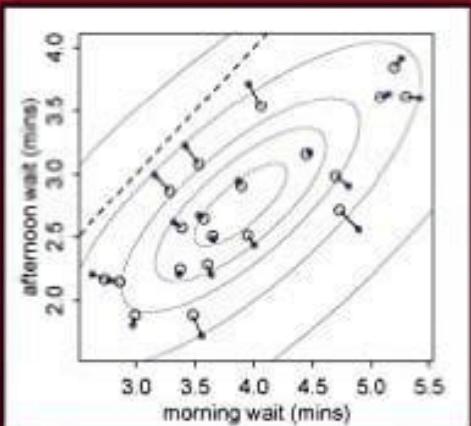


Copyrighted Material

Texts in Statistical Science

# Statistical Rethinking

A Bayesian Course with  
Examples in R and Stan



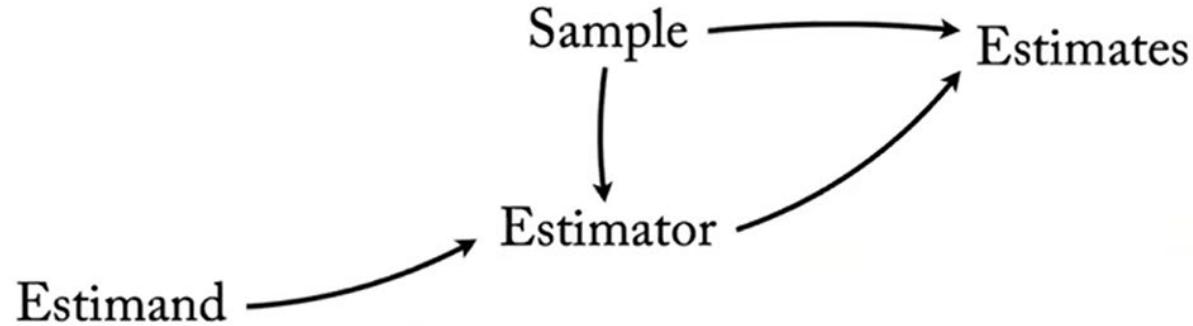
Richard McElreath



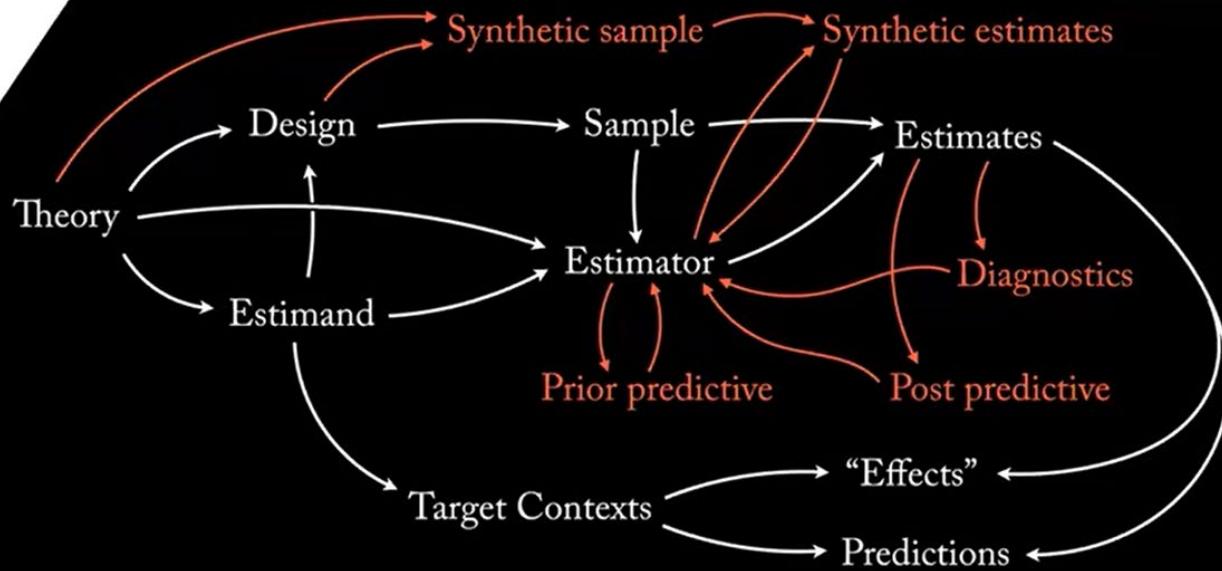
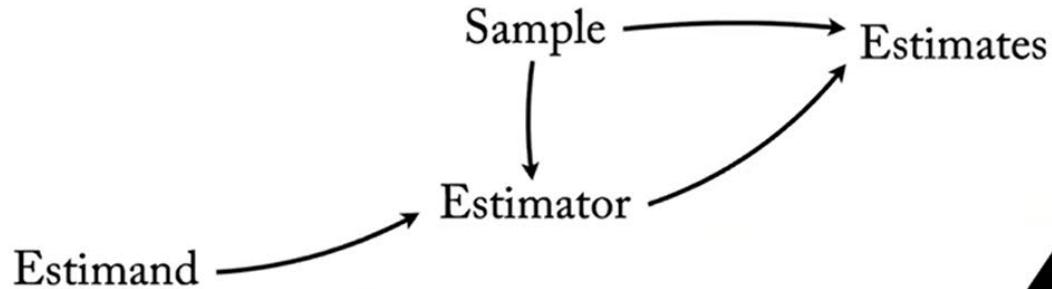
CRC Press  
Taylor & Francis Group

A CHAPMAN & HALL BOOK  
Copyrighted Material

# Bayesian workflow + SEM



# Bayesian workflow + SEM



# Statistical Rethinking

2022



Gracias por  
participar en este  
curso!