

Deconstruir el modelado no eStan complicado

Una introducción a Stan

Iván Barberá



Preguntas,
hipótesis

Preguntas,
hipótesis

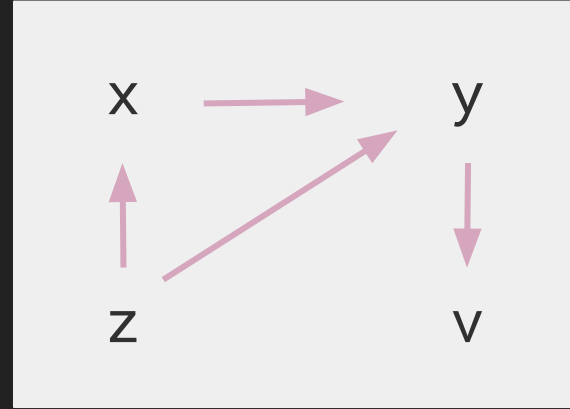


Modelos
cualitativos

Preguntas,
hipótesis



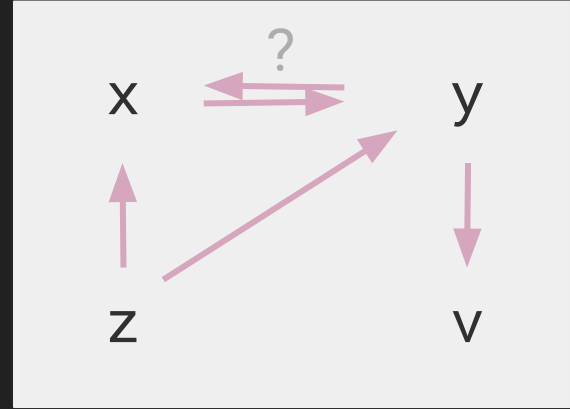
Modelos
cualitativos



Preguntas,
hipótesis



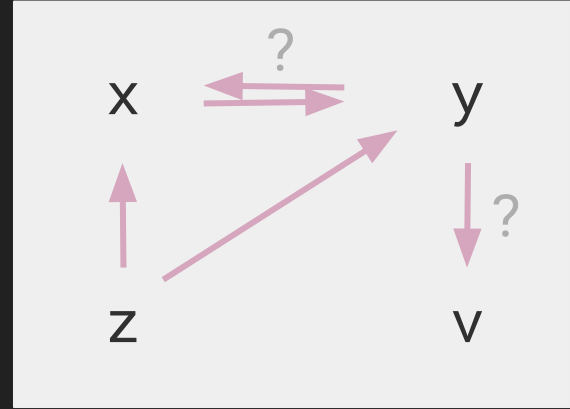
Modelos
cualitativos



Preguntas,
hipótesis



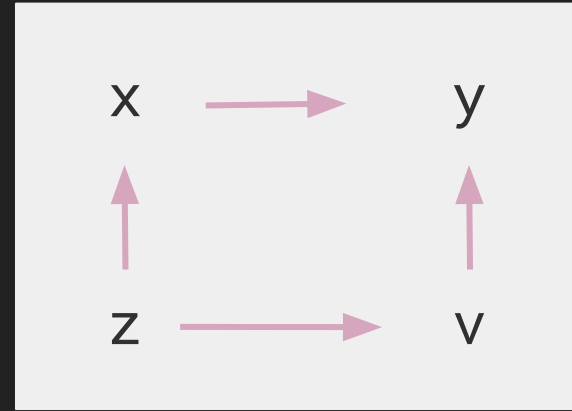
Modelos
cualitativos



Preguntas,
hipótesis



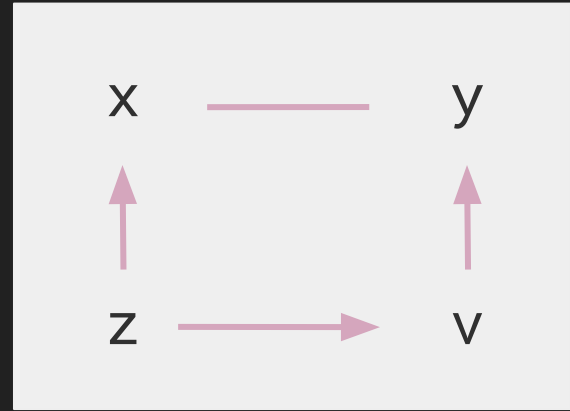
Modelos
cualitativos



Preguntas,
hipótesis



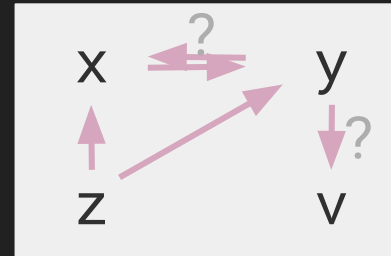
Modelos
cualitativos



Preguntas,
hipótesis



Modelos
cualitativos



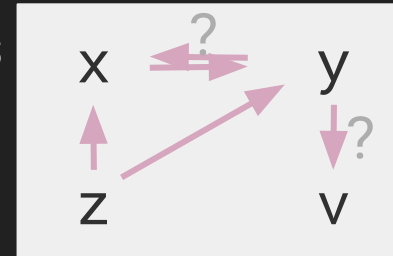
Preguntas,
hipótesis



Modelos
cualitativos



Datos

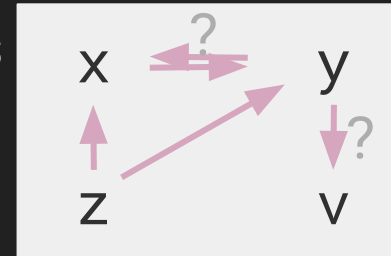


Modelos
cuantitativos

Preguntas,
hipótesis

Datos

Modelos
cualitativos



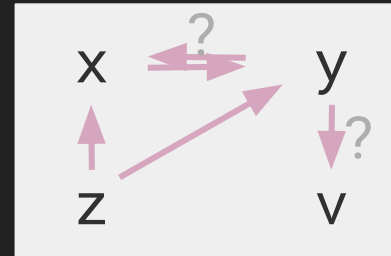
$$y = \beta_0 + \beta_x x + \beta_z z + \beta_v v$$

Modelos
cuantitativos

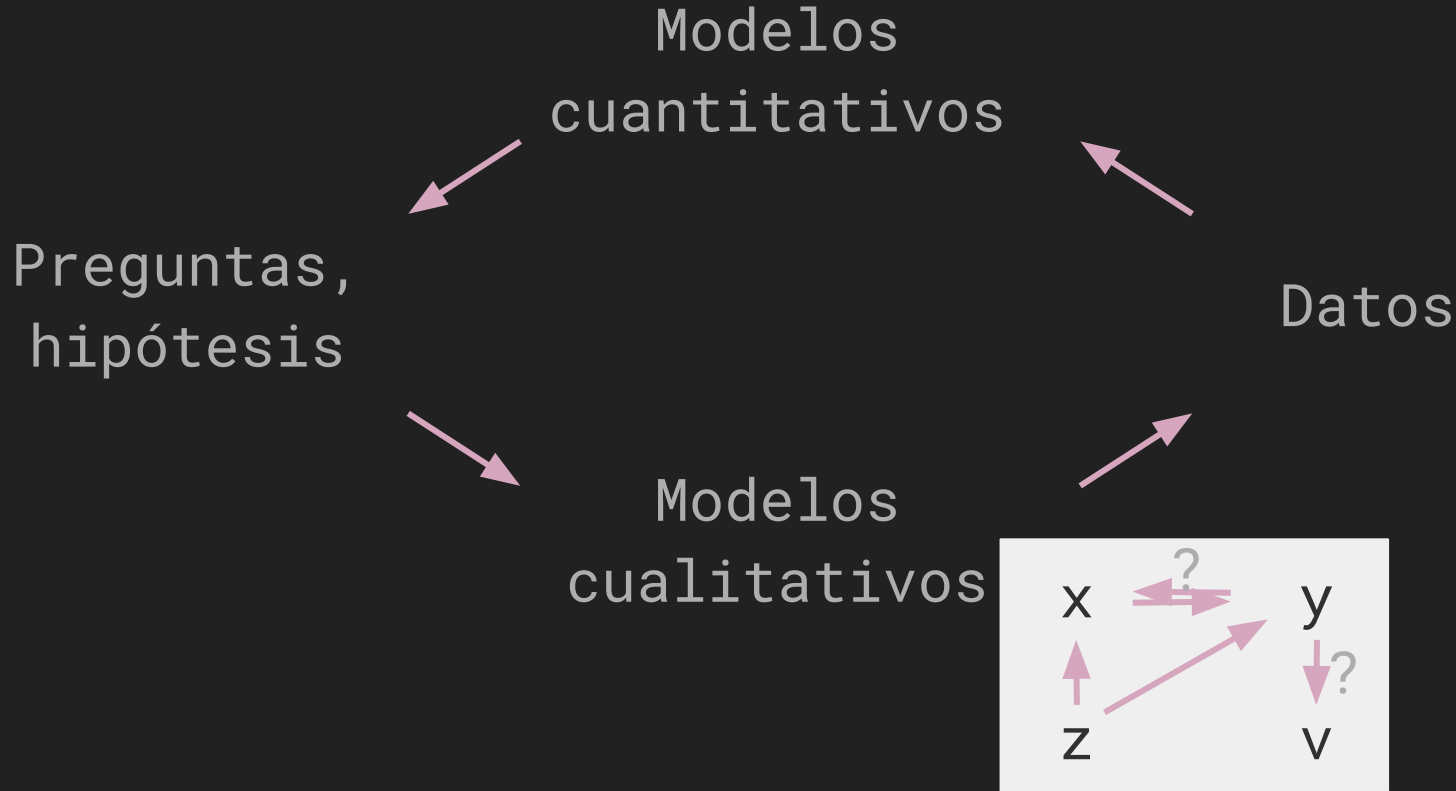
Preguntas,
hipótesis

Datos

Modelos
cualitativos

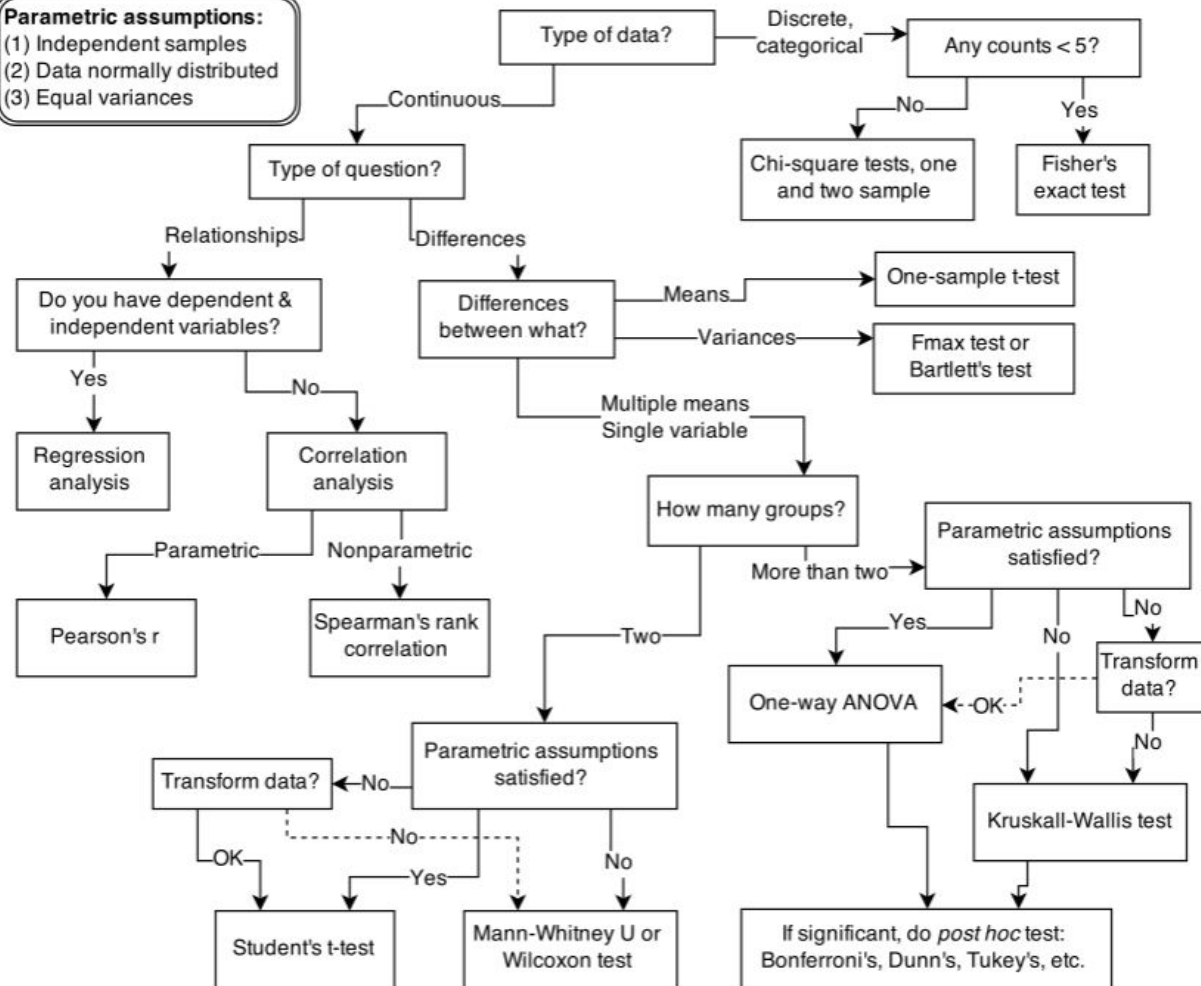


$$y = \beta_0 + \beta_x x + \beta_z z + \beta_v v$$



Parametric assumptions:

- (1) Independent samples
- (2) Data normally distributed
- (3) Equal variances



GLM

GAM

Support
Vector
Machines

LM

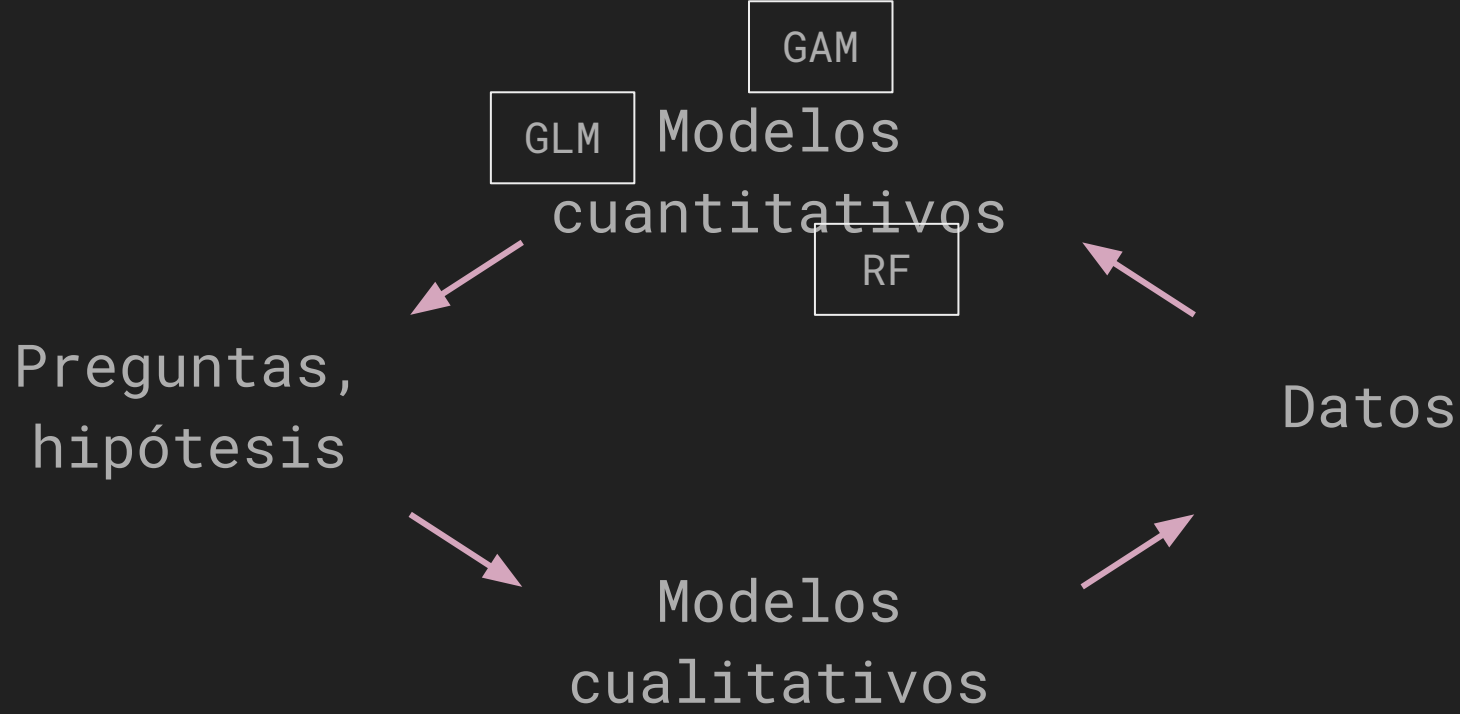
Gaussian
Process

Random
Forest

Deep
Neural
Network

Modelos
no lineales

K-Nearest
Neighbour



Preguntas,
hipótesis



Modelos
cuantitativos
(sin estimar)



Datos

(estimados)

Preguntas,
hipótesis



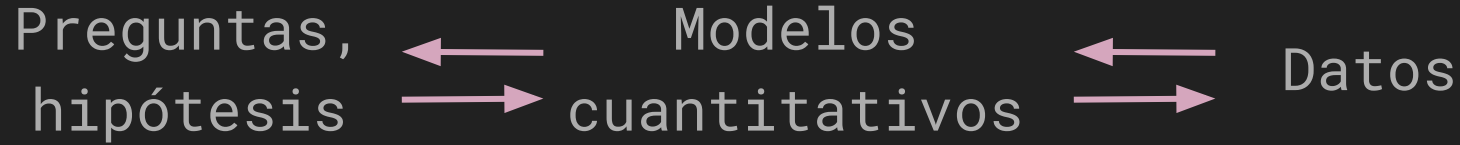
Modelos
cuantitativos



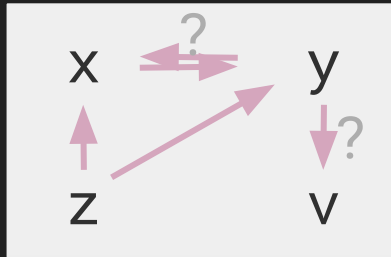
Datos

(sin estimar)

(estimados)



(sin estimar)

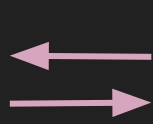


$$y = \beta_0 + \beta_x x + \beta_z z + \beta_v v$$

$$y = 0.3 + 1.1 x + 4 z - 2 v$$

(estimados)

Preguntas,
hipótesis

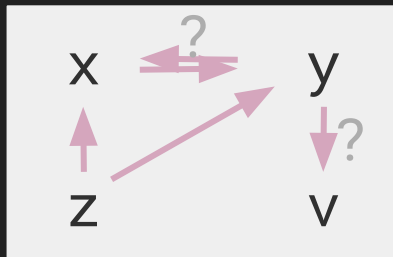


Modelos
cuantitativos

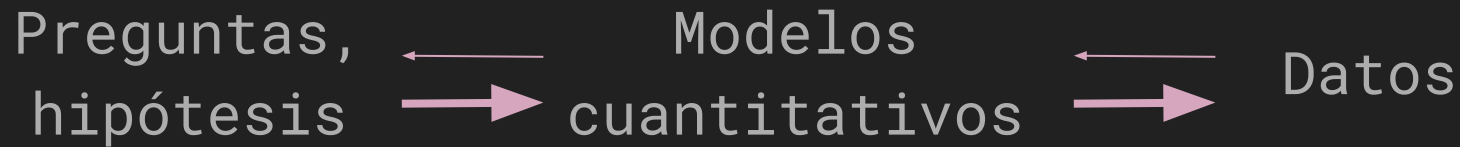


Datos

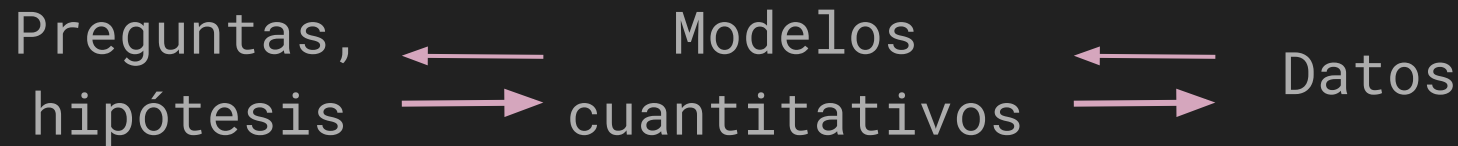
(sin estimar)



$$y = \beta_0 + \beta_x x + \beta_z z + \beta_v v$$

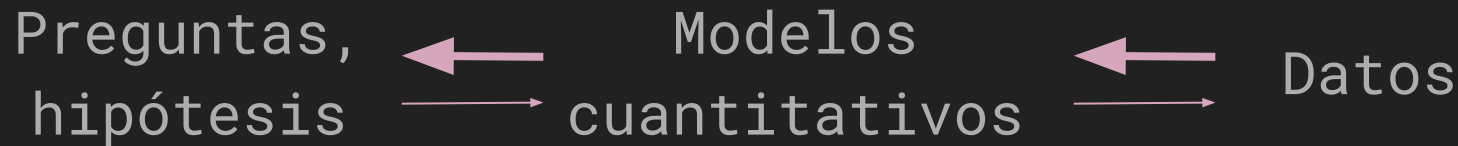


$$y = \beta_0 + \beta_x x + \beta_z z + \beta_v v$$



$$y = \beta_0 + \beta_x x + \beta_z z + \beta_v v$$

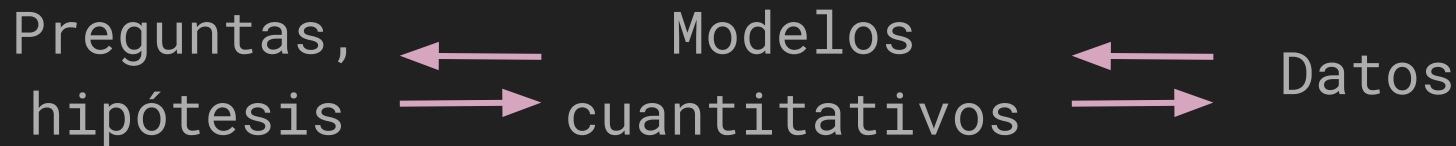
$$y = \beta_0 + \beta_x x + \beta_z z + f(v)$$



$$y = \beta_0 + \beta_x x + \beta_z z + \beta_v v$$

$$y = \beta_0 + \beta_x x + \beta_z z + f(v)$$

$$y = f(x, z, v)$$



$$y = \beta_0 + \beta_x x + \beta_z z + \beta_v v$$

LM

$$y = \beta_0 + \beta_x x + \beta_z z + f(v)$$

GAM

$$y = f(x, z, v)$$

DNN

Modelos estadísticos

$$y = \beta_0 + \beta_x x + \beta_z z + \beta_v v$$

Modelos estadísticos

$$y_i \sim \text{Normal}(\mu_i, \sigma)$$

$$\mu_i = \beta_0 + \beta_x x_i + \beta_z z_i + \beta_v v_i$$

Modelo genérico

$$y_i \sim \mathbf{D}_y(\mu_i, \phi_i)$$

Modelo genérico

$$y_i \sim \mathbf{D}_y(\mu_i, \phi_i)$$
$$\mu_i = f(\mathbf{x}_i, \boldsymbol{\theta}_\mu)$$

Modelo genérico

$$y_i \sim \mathcal{D}_y(\mu_i, \phi_i)$$

$$\mu_i = f(\mathbf{x}_i, \boldsymbol{\theta}_\mu)$$

$$\phi_i = f(\mathbf{x}_i, \boldsymbol{\theta}_\phi)$$

Modelo lineal

$$y_i \sim \mathbf{D}_y(\mu_i, \phi_i)$$

$$\mu_i = f(\mathbf{x}_i, \boldsymbol{\theta}_\mu)$$

$$\phi_i = f(\mathbf{x}_i, \boldsymbol{\theta}_\phi)$$

$$y_i \sim \text{Normal}(\mu_i, \sigma)$$

$$\mu_i = \alpha + \beta x_i$$

Modelo lineal generalizado

$$y_i \sim \text{D}_y(\mu_i, \phi_i)$$

$$\mu_i = f(\mathbf{x}_i, \boldsymbol{\theta}_\mu)$$

$$\phi_i = f(\mathbf{x}_i, \boldsymbol{\theta}_\phi)$$

$$y_i \sim \text{Gamma}(\mu_i, \phi)$$

$$\log(\mu_i) = \alpha + \beta x_i$$

Modelo lineal generalizado

$$y_i \sim \mathcal{D}_y(\mu_i, \phi_i)$$

$$\mu_i = f(\mathbf{x}_i, \boldsymbol{\theta}_\mu)$$

$$\phi_i = f(\mathbf{x}_i, \boldsymbol{\theta}_\phi)$$

$$y_i \sim \text{Bernoulli}(\theta_i)$$

$$\text{logit}(\theta_i) = \alpha + \beta x_i$$

Formulación del modelo

- Conocimiento del sistema
- Imaginación

Formulación del modelo

- Conocimiento del sistema
- Imaginación

Estimación del modelo

- Marco de inferencia
- Software flexible

Formulación del modelo

- Conocimiento del sistema
- Imaginación

Estimación del modelo

- Marco de inferencia
- Software flexible

Inferencia estadística

Frecuentista

probabilidad
como frecuencia

Bayesiana

probabilidad representa
incertidumbre

Inferencia estadística

Frecuentista

probabilidad
como frecuencia

- Estimadores puntuales
- Distribución bajo H_0
- Intervalos de confianza

Bayesiana

probabilidad representa
incertidumbre

Inferencia estadística

Frecuentista

probabilidad
como frecuencia

- Estimadores puntuales
- Distribución bajo H_0
- Intervalos de confianza

Bayesiana

probabilidad representa
incertidumbre

Distribución de
los parámetros
dados los datos

Inferencia estadística

Frecuentista

$$\hat{\theta} = \operatorname{argmax}_{\theta} [p(y|\theta)]$$

Inferencia estadística

Frecuentista

$$\hat{\theta} = \operatorname{argmax}_{\theta} \underbrace{[p(y|\theta)]}_{\text{Likelihood}}$$

MLE

Inferencia estadística

Frecuentista

Bayesiana

$$\hat{\theta} = \operatorname{argmax}_{\theta} \underbrace{[p(y|\theta)]}_{\text{Likelihood}}$$

MLE

$$p(\theta)$$

Inferencia estadística

Frecuentista

Bayesiana

$$\hat{\theta} = \operatorname{argmax}_{\theta} \underbrace{[p(y|\theta)]}_{\text{Likelihood}}$$

MLE

$$p(\theta|y) = \frac{p(y|\theta) p(\theta)}{\int p(y|\theta) p(\theta) d\theta}$$

Inferencia estadística

Frecuentista

Bayesiana

$$\hat{\theta} = \operatorname{argmax}_{\theta} [\underbrace{p(y|\theta)}_{\text{Likelihood}}]$$

MLE

Likelihood

$$p(\theta|y) = \frac{\overbrace{p(y|\theta)}^{\text{Likelihood}} p(\theta)}{\int p(y|\theta) p(\theta) d\theta}$$

Inferencia estadística

Frecuentista

Bayesiana

$$\hat{\theta} = \operatorname{argmax}_{\theta} [\underbrace{p(y|\theta)}_{\text{Likelihood}}]$$

MLE

Likelihood

Previa

$$\underbrace{p(\theta|y)}_{\text{Posterior}} = \frac{\underbrace{p(y|\theta)}_{\text{Likelihood}} \underbrace{p(\theta)}_{\text{Previa}}}{\underbrace{\int p(y|\theta) p(\theta) d\theta}_{\text{Prob. marginal de los datos}}}$$

Prob. marginal de los datos

Inferencia estadística

Frecuentista

Bayesiana

$$\hat{\theta} = \operatorname{argmax}_{\theta} [\underbrace{p(y|\theta)}_{\text{Likelihood}}]$$

MLE

Likelihood

Previa

$$p(\theta|y) \propto \underbrace{p(y|\theta)}_{\text{Likelihood}} \underbrace{p(\theta)}_{\text{Previa}}$$

Posterior

Likelihood prior posterior
en R

Inferencia estadística en la práctica

Frecuentista

Bayesiana

- MLE
- Perfiles de likelihood
- Aproximación Normal

Inferencia estadística en la práctica

Frecuentista

- MLE
- Perfiles de likelihood
- Aproximación Normal

Bayesiana

- Muestreo de la posterior (MCMC, SMC)
- Aproximaciones determinísticas (Laplace, EM, VI, INLA)

Formulación del modelo

- Conocimiento del sistema
- Imaginación

Estimación del modelo

- Marco de inferencia
- Software flexible

Formulación del modelo

- Conocimiento del sistema
- Imaginación

Estimación del modelo

- Marco de inferencia
- Software flexible

Software de estimación

Maximum likelihood

- `bbmle`
- `maxLik`
- `likelihoodAsy`
- `stats4 (mle)`
- `stats (optim)`
- `GA`

Distribución posterior

- `BayesianTools`
- `INLA`
- `Nimble`
- `JAGS`
- `Stan`
- `PyMC`
- `Turing`

Software de estimación

Maximum likelihood

- `bbmle`
- `maxLik`
- `likelihoodAsy`
- `stats4 (mle)`
- `stats (optim)`
- `GA`

Distribución posterior

- `BayesianTools`
- `INLA`
- `Nimble`
- `JAGS`
- `Stan`
- `PyMC`
- `Turing`

Software de estimación

Stan

- MCMC: Hamiltonian Monte Carlo (HMC) - No-U-Turn Sampler (NUTS)
- Optimización (MAP o MLE)
- Laplace Approximation
- ADVI
- Pathfinder



mc-stan.org

Software de estimación

Maximum likelihood

- `bbmle`
- `maxLik`
- `likelihoodAsy`
- `stats4 (mle)`
- `stats (optim)`
- `GA`

Distribución posterior

- `BayesianTools`
- `INLA`
- `Nimble`
- `JAGS`
- `Stan`
- `PyMC`
- `Turing`

EMOSIDO



ENGANADO

P

11



Formulación del modelo

- Conocimiento del sistema
- Imaginación

Estimación del modelo

- Marco de inferencia
- Software flexible

Formulación del modelo

- Conocimiento del sistema
- Imaginación

Deconstruir el
modelado

Estimación del modelo

- Freq vs. Bayes

Una introducción a Stan

$$y_i \sim \mathcal{D}_y(\mu_i, \phi_i)$$

$$\mu_i = f(\mathbf{x}_i, \boldsymbol{\theta}_\mu)$$

$$\phi_i = f(\mathbf{x}_i, \boldsymbol{\theta}_\phi)$$

$$y_i \sim \mathbf{D}_y(\mu_i, \phi_i)$$

$$\mu_i = f(\mathbf{x}_i, \boldsymbol{\theta}_\mu)$$

$$\phi_i = f(\mathbf{x}_i, \boldsymbol{\theta}_\phi)$$

$$\boldsymbol{\theta}_\mu \sim \mathbf{D}_\mu(a, b)$$

$$\boldsymbol{\theta}_\phi \sim \mathbf{D}_\phi(c, d)$$

$$y_i \sim D_y(\mu_i, \phi_i)$$

$$\mu_i = f(\mathbf{x}_i, \boldsymbol{\theta}_\mu)$$

$$\phi_i = f(\mathbf{x}_i, \boldsymbol{\theta}_\phi)$$

$$\boldsymbol{\theta}_\mu \sim D_\mu(a, b)$$

$$\boldsymbol{\theta}_\phi \sim D_\phi(c, d)$$

$$y_i \sim \text{Normal}(\mu_i, \sigma)$$

$$\mu_i = \alpha + \beta x_i$$

$$\alpha \sim \text{Normal}(0, 1)$$

$$\beta \sim \text{Normal}(0, 1)$$

$$\sigma \sim \text{Half-Normal}(2)$$

Caso de estudio:
comportamiento de ballenas en función del ataque
por gaviotas



Datos

	A	B	C	D
1	t	<u>grupo</u>	y	x
2	1	1	0	0
3	2	1	0	0
4	3	1	0	0
5	4	1	0	0
6	5	1	0	0
7	6	1	0	0
8	7	1	0	0
9	8	1	0	0
10	9	1	0	0
11	10	1	0	0
12	11	1	1	0
13	12	1	1	0
14	1	2	0	0
15	2	2	0	0
16	3	2	0	0
17	4	2	0	0
18	5	2	0	0
19	6	2	0	0
20	7	2	0	0
21	8	2	0	1
22	9	2	0	1
23	10	2	0	0



Modelo 1

$$y_i \sim \text{Bernoulli}(\theta)$$

$$\theta \sim \text{Unif}(0, 1)$$

$$y = 0 : \text{quieta}$$

$$y = 1 : \text{en movimiento}$$

Modelo 1

$$y_i \sim \text{Bernoulli}(\theta)$$

$$\theta \sim \text{Unif}(0, 1)$$

$$p(y|\theta)$$

$$p(\theta)$$

Modelo 1

Likelihood

$$y_i \sim \text{Bernoulli}(\theta)$$

$$p(y|\theta)$$

Previa

$$\theta \sim \text{Unif}(0, 1)$$

$$p(\theta)$$

Modelo 2

$$y_i \sim \text{Bernoulli}(\theta_i)$$

$$\theta_i = \text{logit}^{-1}(\alpha + \beta z_i)$$

$$\alpha \sim \text{Normal}(0, 5)$$

$$\beta \sim \text{Normal}(0, 2)$$

Modelo 3

$$y_t \sim \text{Bernoulli}(\theta_t)$$

$$\theta_t = \text{logit}^{-1}(\alpha + \beta z_t)$$

$$\begin{aligned} z_t = z_{t-1} \\ + \iota (1 - z_{t-1}) x_{t-1} \\ - \delta z_{t-1} (1 - x_{t-1}) \end{aligned}$$

$$\alpha \sim \text{Normal}(0, 5)$$

$$\beta \sim \text{Normal}(0, 2)$$

$$\iota \sim \text{Unif}(0, 1)$$

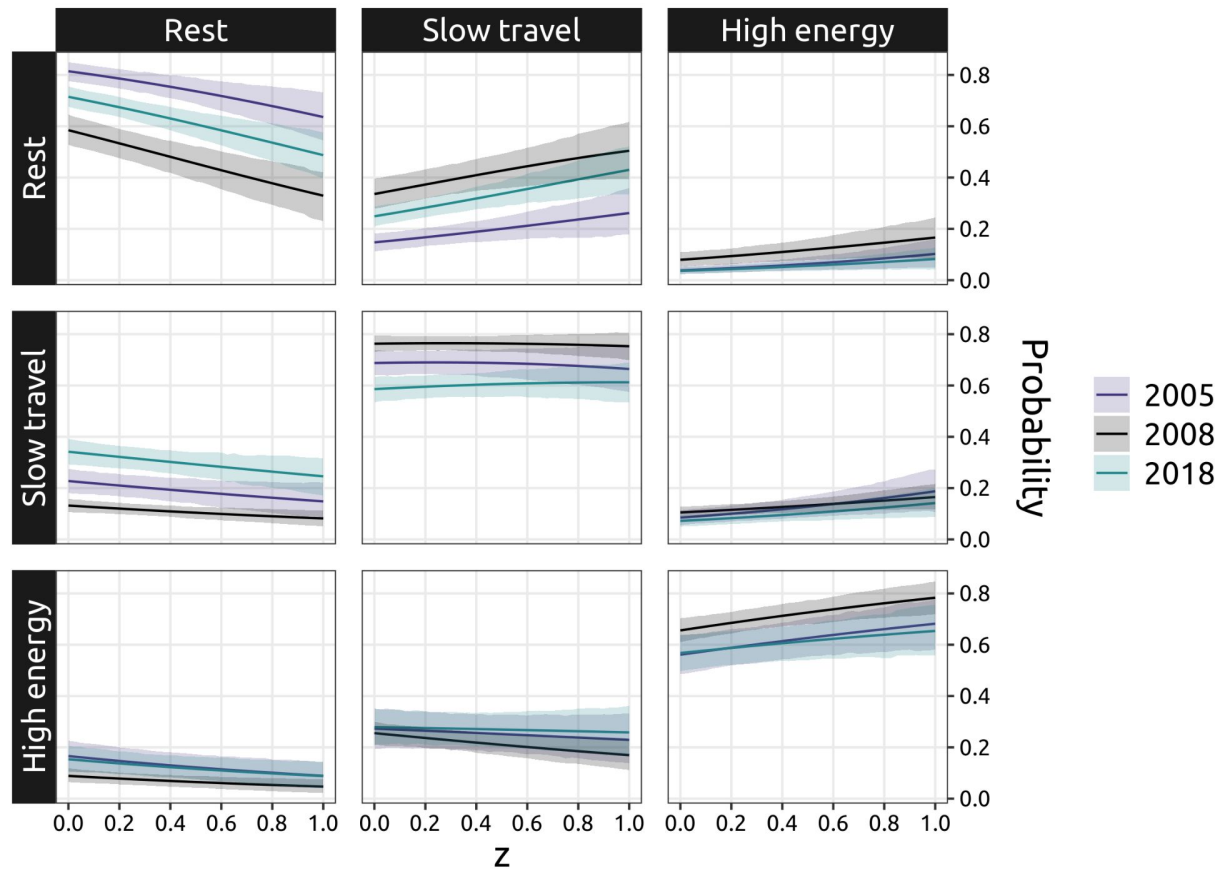
$$\delta \sim \text{Unif}(0, 1)$$

$$z_1 \sim \text{Unif}(0, 1)$$

R y Stan

Transition probabilities as a function of z

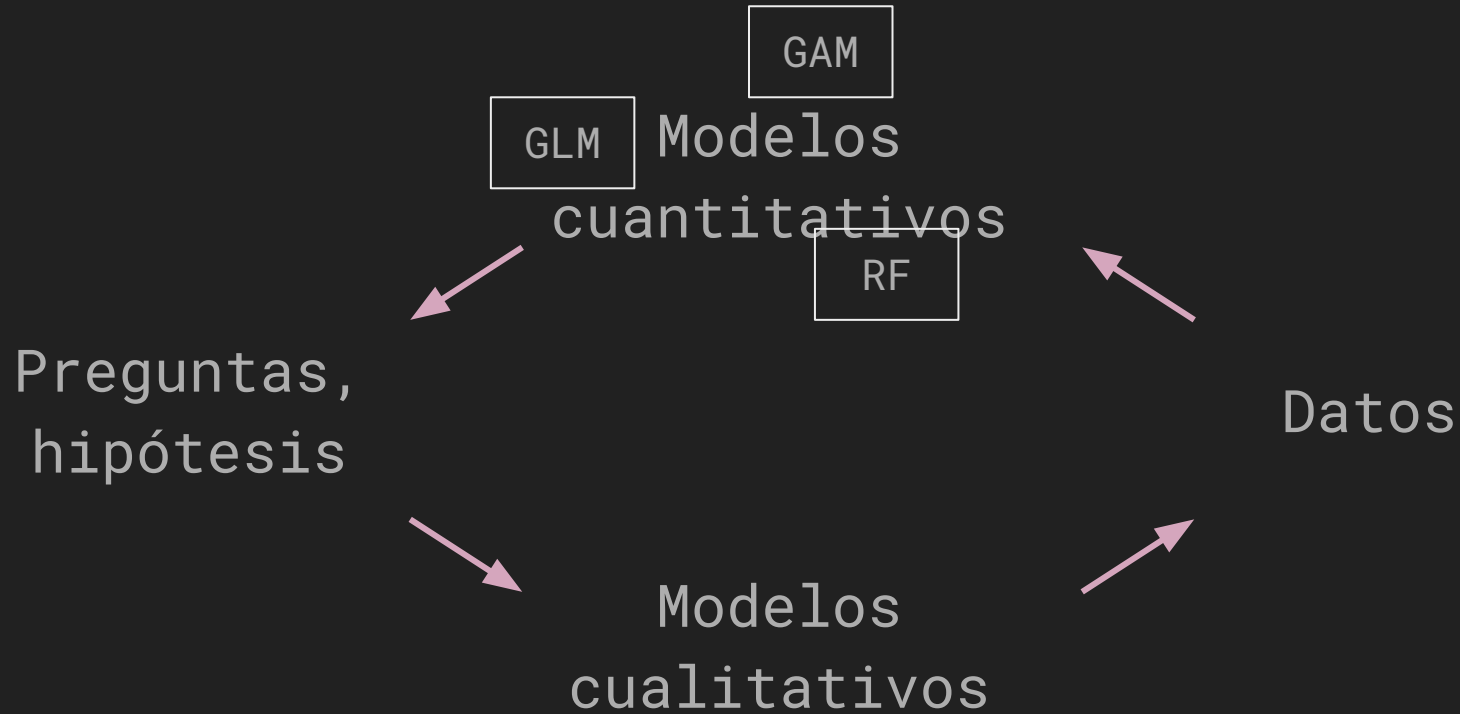
(row to column)

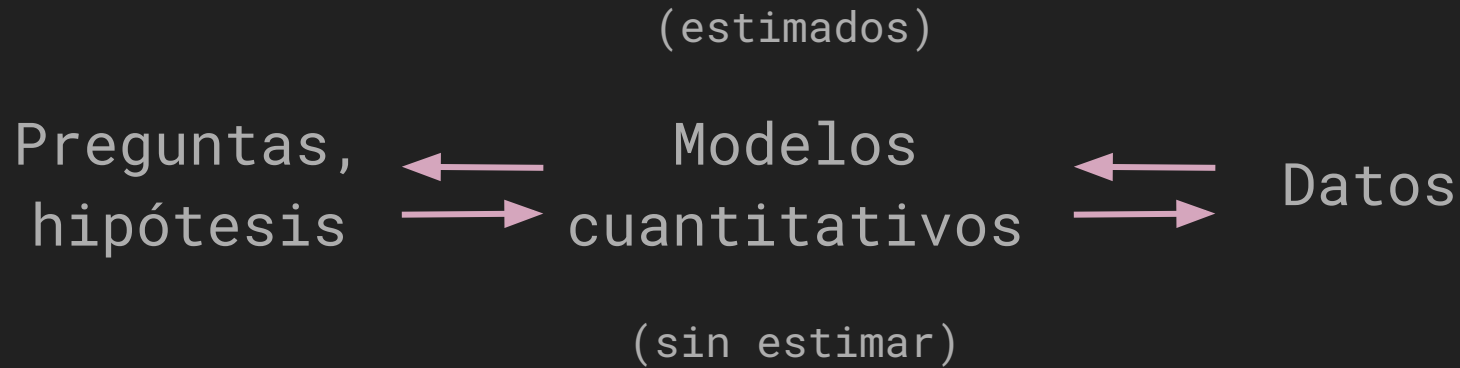


Modelado
prescriptivo



Modelado
emancipado





Modelado
prescriptivo



Modelado
prescriptivo



Formulación del modelo

- Conocimiento del sistema
- Imaginación

Deconstruir el
modelado

Estimación del modelo

- Freq vs. Bayes
- bbmle vs. Stan

Una introducción a Stan

Stan User's Guide

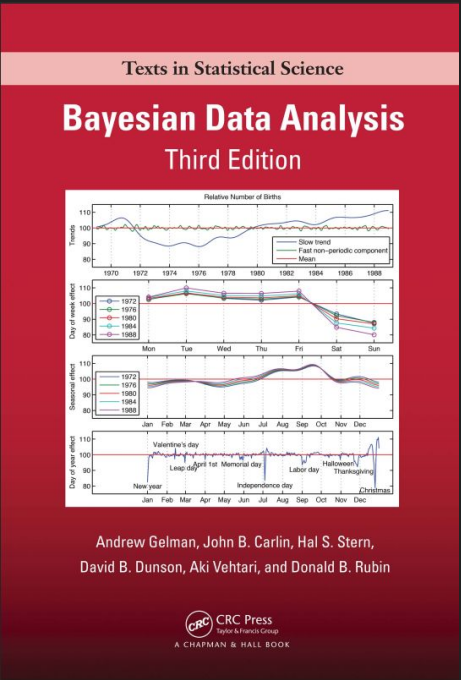
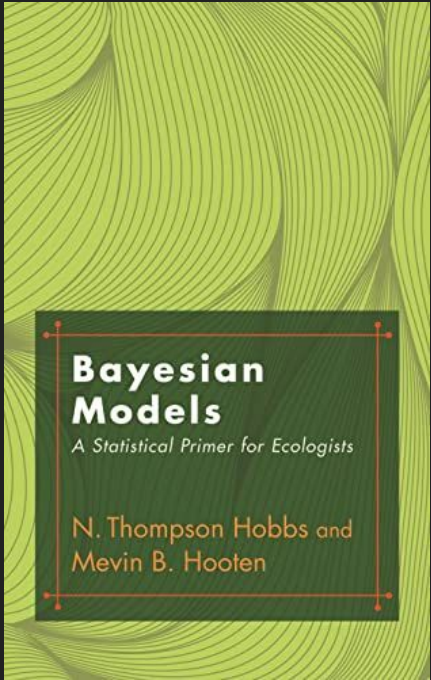
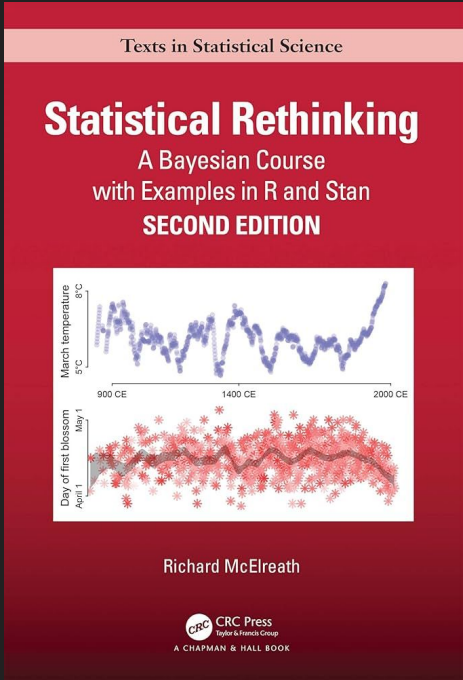
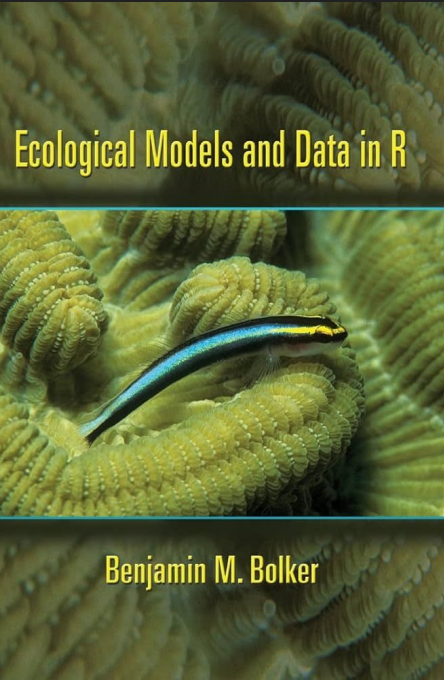
Version 2.35

This is the official user's guide for [Stan](#). It provides example models and programming techniques for coding statistical models in Stan.

- Part 1 gives Stan code and discussions for several important classes of models.
- Part 2 discusses various general Stan programming techniques that are not tied to any particular model.
- Part 3 introduces algorithms for calibration and model checking that require multiple runs of Stan.



mc-stan.org



Fin.