REPASO

Bayes $$p(\theta|y) \propto p(y|\theta)p(\theta)$$

posterior

likelihood

prior

REPASO

Bayes

$$\underbrace{p(\theta|y)}_{\textit{posterior}} \propto \underbrace{p(y|\theta)}_{\textit{likelihood}}\underbrace{p(\theta)}_{\textit{prior}}$$

prior predictive

$$p(y) = \int p(y|\theta)p(\theta)d\theta$$

REPASO

Bayes

$$\underbrace{p(\theta|y)}_{\textit{posterior}} \propto \underbrace{p(y|\theta)}_{\textit{likelihood}}\underbrace{p(\theta)}_{\textit{prior}}$$

prior predictive $$p(y) = \int p(y|\theta)p(\theta)d\theta$$

posterior predictive $$p(\tilde{y}|y) = \int p(\tilde{y}|\theta)p(\theta|y)d\theta$$

REPASO

Bayes

$$\underbrace{p(\theta|y)}_{\textit{posterior}} \propto \underbrace{p(y|\theta)}_{\textit{likelihood}}\underbrace{p(\theta)}_{\textit{prior}}$$

prior predictive $$p(y) = \int p(y|\theta)p(\theta)d\theta$$

posterior predictive $$p(\tilde{y}|y) = \int p(\tilde{y}|\theta)p(\theta|y)d\theta$$

Modelo moneda

Modelo Gaussiana (media, varianza)

Otras distribuciones

Binomial $\longrightarrow$ Beta

Otras distribuciones

Binomial $\longrightarrow$ Beta

Gaussiana $\longrightarrow$ Gaussiana (media)

Gaussiana $\longrightarrow$ Gamma inversa (varianza)

Otras distribuciones

Binomial $\longrightarrow$ Beta

Gaussiana $\longrightarrow$ Gaussiana (media)

Gaussiana $\longrightarrow$ Gamma inversa (varianza)

Poisson (cuentas) $\longrightarrow$ Gamma

Otras distribuciones

Binomial $\longrightarrow$ Beta

Gaussiana $\nearrow$ Gaussiana (media)

Gaussiana $\searrow$ Gamma inversa (varianza)

Poisson (cuentas) $\longrightarrow$ Gamma

$$p(y|\theta) = \frac{\theta^y e^{-\theta}}{y!} \quad y = 0, 1, 2, \ldots$$

Otras distribuciones

Binomial $\longrightarrow$ Beta

Gaussiana $\longrightarrow$ Gaussiana (media)

Gaussiana $\longrightarrow$ Gamma inversa (varianza)

Poisson (cuentas) $\longrightarrow$ Gamma

$$p(y|\theta) = \frac{\theta^y e^{-\theta}}{y!} \quad y = 0, 1, 2, \ldots$$

$$p(y|\theta) = \prod_{i=1}^{n} \frac{\theta^{y_i} e^{-\theta}}{y_i!} \propto \theta^{n\bar{y}} e^{-n\theta}$$

Otras distribuciones

Binomial $\longrightarrow$ Beta

Gaussiana $\nearrow$ Gaussiana (media)

Gaussiana $\searrow$ Gamma inversa (varianza)

Poisson (cuentas) $\longrightarrow$ Gamma

$$p(y|\theta) = \frac{\theta^y e^{-\theta}}{y!} \quad y = 0, 1, 2, \ldots$$

$$p(y|\theta) = \prod_{i=1}^{n} \frac{\theta^{y_i} e^{-\theta}}{y_i!} \propto \theta^{n\bar{y}} e^{-n\theta}$$

$$p(\theta) = \Gamma(\alpha, \beta)$$

Otras distribuciones

Binomial $\longrightarrow$ Beta

Gaussiana → Gaussiana (media)

Gaussiana → Gamma inversa (varianza)

Poisson (cuentas) $\longrightarrow$ Gamma

$$p(y|\theta) = \frac{\theta^y e^{-\theta}}{y!} \quad y = 0, 1, 2, \ldots$$

$$p(y|\theta) = \prod_{i=1}^{n} \frac{\theta^{y_i} e^{-\theta}}{y_i!} \propto \theta^{n\bar{y}} e^{-n\theta}$$

$$p(\theta) = \Gamma(\alpha, \beta) \qquad\qquad p(\theta|y) = \Gamma(\alpha + n\bar{y}, \beta + n)$$

Otras distribuciones

Binomial $\longrightarrow$ Beta

Gaussiana $\longrightarrow$ Gaussiana (media)

Gaussiana $\longrightarrow$ Gamma inversa (varianza)

Poisson (cuentas) $\longrightarrow$ Gamma

$$p(y|\theta) = \frac{\theta^y e^{-\theta}}{y!} \quad y = 0, 1, 2, \ldots$$

$$p(y|\theta) = \prod_{i=1}^{n} \frac{\theta^{y_i} e^{-\theta}}{y_i!} \propto \theta^{n\bar{y}} e^{-n\theta}$$

$$p(\theta) = \Gamma(\alpha, \beta) \qquad p(\theta|y) = \Gamma(\underbrace{\alpha + n\bar{y}}_{\text{cuentas}}, \underbrace{\beta + n}_{\text{observaciones}})$$

Otras distribuciones

Binomial $\longrightarrow$ Beta

Gaussiana $\to$ Gaussiana (media)

Gaussiana $\to$ Gamma inversa (varianza)

Poisson (cuentas) $\longrightarrow$ Gamma

Otras distribuciones

Binomial → Beta

Gaussiana → Gaussiana (media)

Gaussiana → Gamma inversa (varianza)

Poisson (cuentas) → Gamma

Exponencial (tiempos) → Gamma

Otras distribuciones

Binomial $\longrightarrow$ Beta

Gaussiana $\nearrow$ Gaussiana (media)

Gaussiana $\searrow$ Gamma inversa (varianza)

Poisson (cuentas) $\longrightarrow$ Gamma

Exponencial (tiempos) $\longrightarrow$ Gamma

$$p(y|\theta) = \theta e^{-y\theta} \quad y > 0$$

Otras distribuciones

Binomial $\longrightarrow$ Beta

Gaussiana $\nearrow$ Gaussiana (media)

Gaussiana $\searrow$ Gamma inversa (varianza)

Poisson (cuentas) $\longrightarrow$ Gamma

Exponencial (tiempos) $\longrightarrow$ Gamma

$$p(y|\theta) = \theta e^{-y\theta} \quad y > 0$$

$$p(y|\theta) = \Gamma(1, \theta)$$

Otras distribuciones

Binomial $\longrightarrow$ Beta

Gaussiana $\nearrow$ Gaussiana (media)

Gaussiana $\searrow$ Gamma inversa (varianza)

Poisson (cuentas) $\longrightarrow$ Gamma

Exponencial (tiempos) $\longrightarrow$ Gamma

$$p(y|\theta) = \theta e^{-y\theta} \quad y > 0$$

$$p(y|\theta) = \Gamma(1, \theta)$$

$$p(y|\theta) = \theta^n e^{-n\bar{y}\theta}$$

Otras distribuciones

Binomial $\longrightarrow$ Beta

Gaussiana $\longrightarrow$ Gaussiana (media)

Gaussiana $\longrightarrow$ Gamma inversa (varianza)

Poisson (cuentas) $\longrightarrow$ Gamma

Exponencial (tiempos) $\longrightarrow$ Gamma

$$p(y|\theta) = \theta e^{-y\theta} \quad y > 0 \qquad\qquad p(\theta) = \Gamma(\alpha, \beta)$$

$$p(y|\theta) = \Gamma(1, \theta)$$

$$p(y|\theta) = \theta^n e^{-n\bar{y}\theta}$$

Otras distribuciones

Binomial $\longrightarrow$ Beta

Gaussiana → Gaussiana (media)

Gaussiana → Gamma inversa (varianza)

Poisson (cuentas) $\longrightarrow$ Gamma

Exponencial (tiempos) $\longrightarrow$ Gamma

$$p(y|\theta) = \theta e^{-y\theta} \quad y > 0$$

$$p(\theta) = \Gamma(\alpha, \beta)$$

$$p(y|\theta) = \Gamma(1, \theta)$$

$$p(\theta|y) = \Gamma(\alpha + n, \beta + n\bar{y})$$

$$p(y|\theta) = \theta^n e^{-n\bar{y}\theta}$$

Otras distribuciones

Binomial $\longrightarrow$ Beta

Gaussiana $\rightarrow$ Gaussiana (media)

Gaussiana $\rightarrow$ Gamma inversa (varianza)

Poisson (cuentas) $\longrightarrow$ Gamma

Exponencial (tiempos) $\longrightarrow$ Gamma

$$p(y|\theta) = \theta e^{-y\theta} \quad y > 0$$

$$p(\theta) = \Gamma(\alpha, \beta)$$

$$p(y|\theta) = \Gamma(1, \theta)$$

$$p(\theta|y) = \Gamma(\alpha + n, \beta + n\bar{y})$$

observaciones

tiempo de espera total

$$p(y|\theta) = \theta^n e^{-n\bar{y}\theta}$$

Otras distribuciones

Binomial → Beta

Gaussiana → Gaussiana (media)

Gaussiana → Gamma inversa (varianza)

Poisson (cuentas) → Gamma

Exponencial (tiempos) → Gamma

Otras distribuciones

Binomial $\longrightarrow$ Beta

Gaussiana $\rightarrow$ Gaussiana (media)

Gaussiana $\rightarrow$ Gamma inversa (varianza)

Poisson (cuentas) $\longrightarrow$ Gamma

Exponencial (tiempos) $\longrightarrow$ Gamma

Bloques para modelos más complejos

...¿más variables?

Gaussiana “completa”

$$p(y|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2\sigma^2}(y-\mu)^2}$$

Gaussiana “completa”

$$p(y|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(y-\mu)^2}$$

prior conjugado

$$p(\mu,\sigma^2) = p(\mu|\sigma^2)p(\sigma^2)$$

Gaussiana “completa”

$$p(y|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(y-\mu)^2}$$

prior conjugado

$$p(\mu, \sigma^2) = p(\mu|\sigma^2)p(\sigma^2)$$

$$\mu|\sigma^2 \sim \mathrm{N}(\mu_0, \sigma^2/\kappa_0)$$

$$\sigma^2 \sim \text{Inv-}\Gamma(\alpha_0, \beta_0)$$

Gaussiana “completa”

$$p(y|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(y-\mu)^2}$$

prior conjugado

$$p(\mu,\sigma^2) = p(\mu|\sigma^2)p(\sigma^2)$$

$$\mu|\sigma^2 \sim \mathrm{N}(\mu_0, \sigma^2/\kappa_0)$$

$$\sigma^2 \sim \text{Inv-}\Gamma(\alpha_0, \beta_0)$$

prior semiconjugado

$$\mu|\sigma^2 \sim \mathrm{N}(\mu_0, \tau_0)$$

Gaussiana “completa”

$$p(y|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(y-\mu)^2}$$

prior conjugado

$$p(\mu,\sigma^2) = p(\mu|\sigma^2)p(\sigma^2)$$

$$\mu|\sigma^2 \sim \mathrm{N}(\mu_0, \sigma^2/\kappa_0)$$

$$\sigma^2 \sim \mathrm{Inv\text{-}\Gamma}(\alpha_0, \beta_0)$$

prior semiconjugado

$$\mu|\sigma^2 \sim \mathrm{N}(\mu_0, \tau_0)$$

$$p(\mu,\sigma^2) = p(\mu)p(\sigma^2)$$

Gaussiana “completa”

$$p(y|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(y-\mu)^2}$$

prior conjugado

$$p(\mu,\sigma^2) = p(\mu|\sigma^2)p(\sigma^2)$$

$$\mu|\sigma^2 \sim \mathrm{N}(\mu_0, \sigma^2/\kappa_0)$$

$$\sigma^2 \sim \text{Inv-}\Gamma(\alpha_0, \beta_0)$$

prior semiconjugado

$$\mu|\sigma^2 \sim \mathrm{N}(\mu_0, \tau_0)$$

$$p(\mu,\sigma^2) = p(\mu)p(\sigma^2)$$

No da un resultado estándar, pero puede ser útil, y más apropiado según el problema (si tenemos una estimación de μ independiente de la forma de producir muestras)

Gaussiana multivariada

$$p(y|\mu, \Sigma) = \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} e^{-\frac{1}{2}(y-\mu)^T \Sigma^{-1}(y-\mu)}$$

μ d-vector

Σ matriz $d \times d$ simétrica y definida positiva

Gaussiana multivariada

$$p(y|\mu, \Sigma) = \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} e^{-\frac{1}{2}(y-\mu)^T \Sigma^{-1} (y-\mu)}$$

μ $\quad d$-vector

Σ $\quad$ matriz $d \times d$ simétrica y definida positiva

Similar al caso $d = 1$

Covarianza conocida: *prior* conjugado Gaussiano

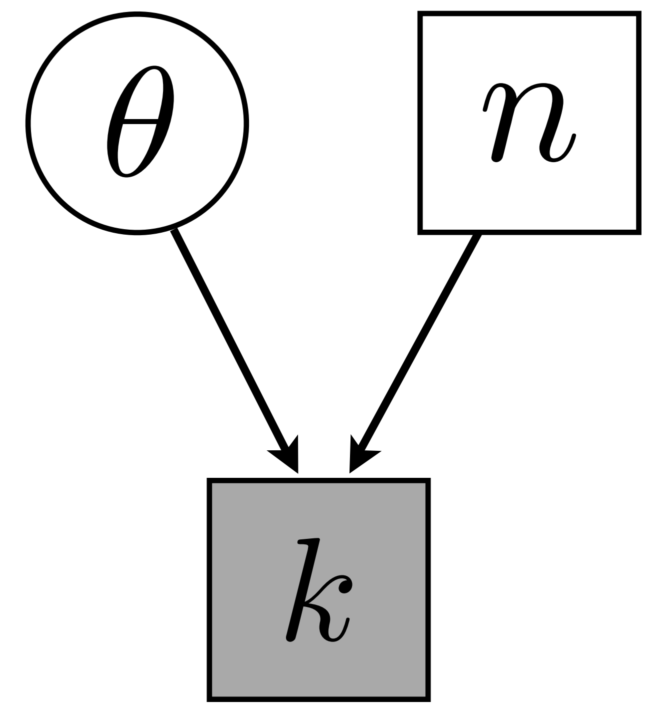
Posterior marginales: también Gaussianas

De vuelta a la binomial.. ahora con *n* desconocido

$$k \sim \text{Binomial}(\theta, n)$$

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

$$n \sim \text{Uniforme}(1, N_{max})$$



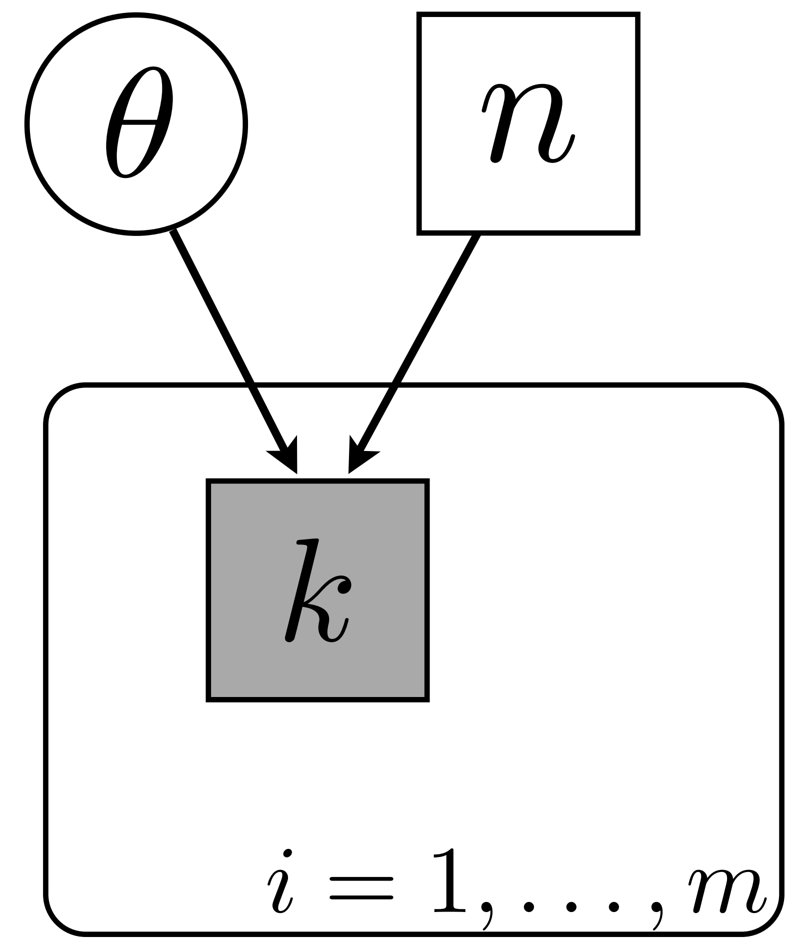
$$p(\theta, n|k) = \ldots$$

Ejemplo con varias observaciones (encuesta)

$$k_i \sim \text{Binomial}(\theta, n)$$

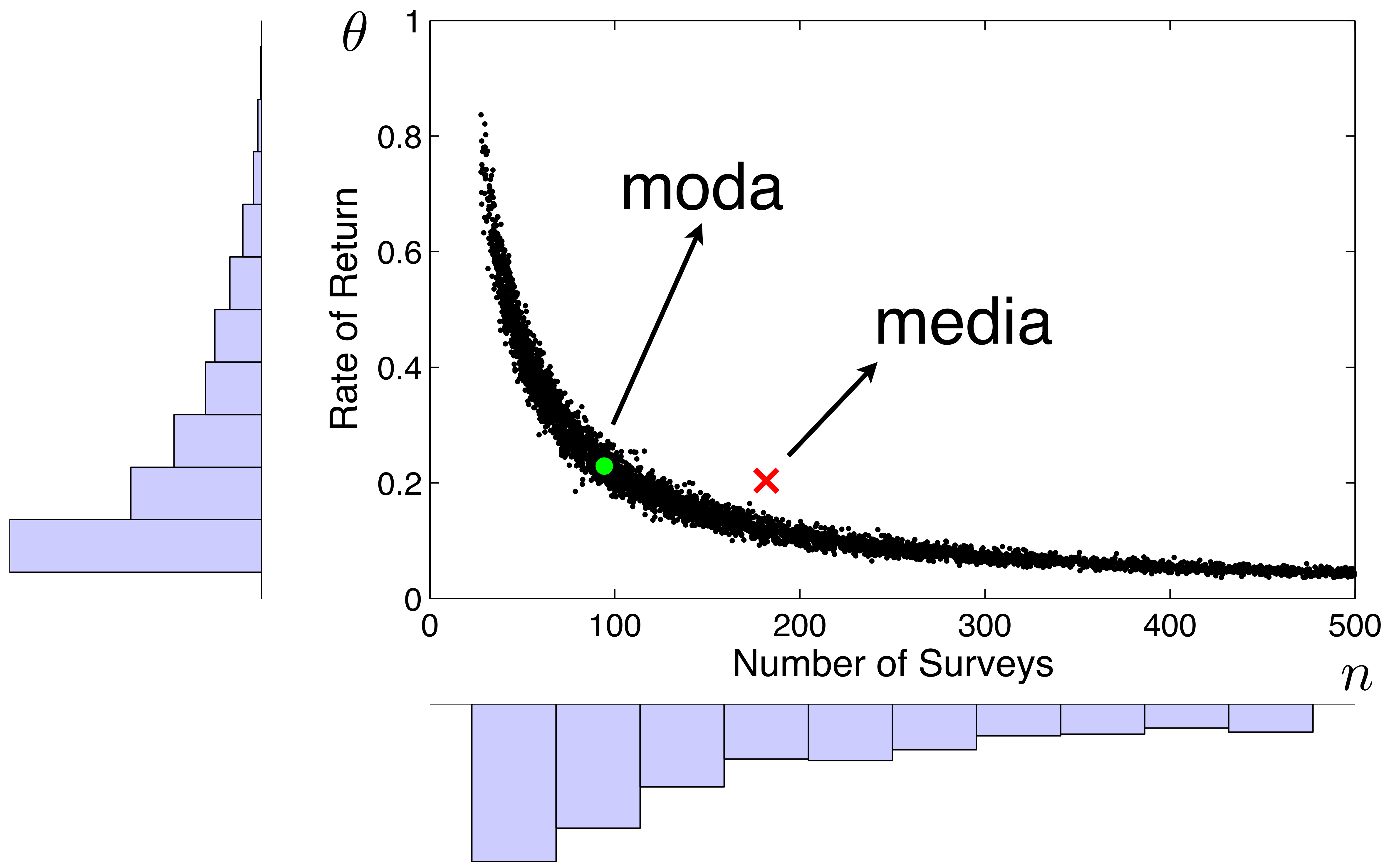
$$\theta \sim \text{Uniforme}(0, 1)$$

$$n \sim \text{Uniforme}(1, N_{max})$$

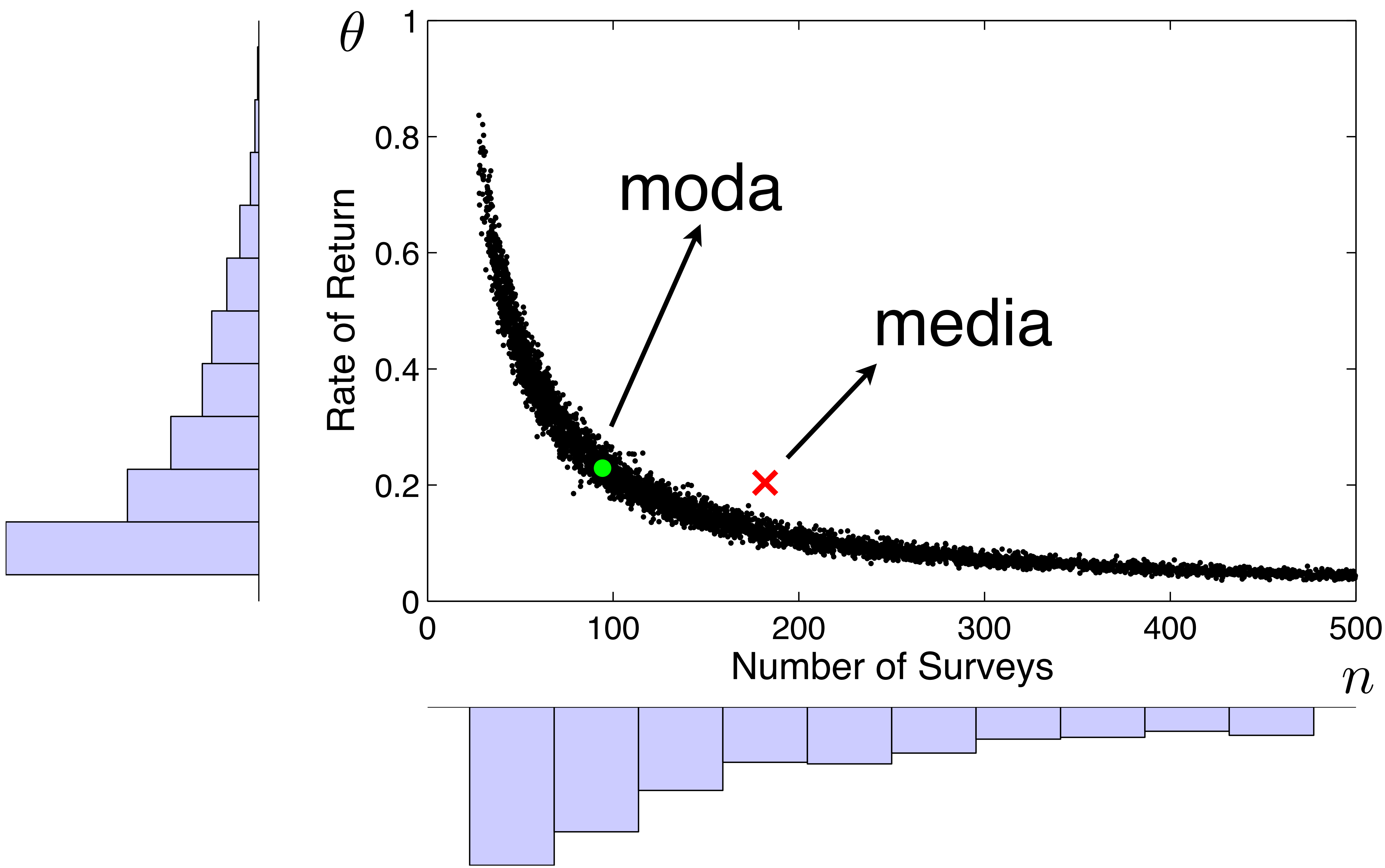


$$p(\theta, n | k) = \dots$$

$$k = \{16, 18, 22, 25, 27\}$$



$$k = \{16, 18, 22, 25, 27\}$$



¡Atención a la conjunta! (problemático con muchas variables..)

Distribución multinomial

$$
\begin{array}{ccccccc}
\theta_1, & \theta_2, & \ldots, & \theta_k \\
\downarrow & \downarrow & & \downarrow \\
y_1, & y_2, & \ldots, & y_k
\end{array}
$$

Distribución multinomial

probabilidades

$$\begin{array}{ccccccc} \theta_1, & \theta_2, & \ldots, & \theta_k \\ \downarrow & \downarrow & & \downarrow \\ y_1, & y_2, & \ldots, & y_k \end{array} \qquad \sum_{j=1}^{k} \theta_j = 1$$

Distribución multinomial

probabilidades

$$\theta_1, \ \theta_2, \ \ldots, \ \theta_k \qquad \sum_{j=1}^{k} \theta_j = 1$$

$$\downarrow \quad \downarrow \qquad\quad \downarrow$$

$$y_1, \ y_2, \ \ldots, \ y_k \qquad \sum_{j=1}^{k} y_j = n$$

cuentas

Distribución multinomial

probabilidades

$$\theta_1,\ \theta_2,\ \ldots,\ \theta_k$$

$$\downarrow \qquad \downarrow \qquad\qquad \downarrow$$

$$y_1,\ y_2,\ \ldots,\ y_k$$

cuentas

$$\sum_{j=1}^{k} \theta_j = 1$$

$$\sum_{j=1}^{k} y_j = n$$

$$p(y|\theta) \propto \prod_{j=1}^{k} \theta_j^{y_j}$$

Distribución multinomial

probabilidades

$$\theta_1,\ \theta_2,\ \ldots,\ \theta_k$$

$$\downarrow \qquad \downarrow \qquad\qquad \downarrow$$

$$y_1,\ y_2,\ \ldots,\ y_k$$

cuentas

$$\sum_{j=1}^{k} \theta_j = 1$$

$$\sum_{j=1}^{k} y_j = n$$

$$p(y|\theta) \propto \prod_{j=1}^{k} \theta_j^{y_j}$$

	k=2		
	Binomial	$\longleftrightarrow$	Multinomial
n=1	Bernoulli	$\longleftrightarrow$	Categórica

Distribución multinomial

probabilidades

$$\theta_1,\ \theta_2,\ \ldots,\ \theta_k$$

$$\downarrow \quad \downarrow \qquad \downarrow$$

$$y_1,\ y_2,\ \ldots,\ y_k$$

cuentas

$$\sum_{j=1}^{k} \theta_j = 1$$

$$\sum_{j=1}^{k} y_j = n$$

$$p(y|\theta) \propto \prod_{j=1}^{k} \theta_j^{y_j}$$

	k=2		
	Binomial	$\longleftrightarrow$	Multinomial
n=1	Bernoulli	$\longleftrightarrow$	Categórica

$$(\theta_1, \theta_2, \ldots, \theta_k) = (1/6, 1/6, \ldots, 1/6)$$

prior conjuado: Dirichlet

$$p(\theta|\alpha) = \mathrm{Dirichlet}(\alpha_1, \ldots, \alpha_k) \propto \prod_{j=1}^{k} \theta_j^{\alpha_j - 1}$$

$$\theta_j \geq 0$$

$$\sum_{j=1}^{k} \theta_j = 1$$

prior conjuado: Dirichlet

$$p(\theta|\alpha) = \text{Dirichlet}(\alpha_1, \ldots, \alpha_k) \propto \prod_{j=1}^{k} \theta_j^{\alpha_j - 1}$$

$$\theta_j \geq 0$$

$$\sum_{j=1}^{k} \theta_j = 1$$

posterior

$$p(\theta|y) = \text{Dirichlet}(\alpha_1 + y_1, \ldots, \alpha_k + y_k)$$

prior conjuado: Dirichlet

$$p(\theta|\alpha) = \text{Dirichlet}(\alpha_1, \ldots, \alpha_k) \propto \prod_{j=1}^{k} \theta_j^{\alpha_j - 1}$$

$$\theta_j \geq 0$$

$$\sum_{j=1}^{k} \theta_j = 1$$

posterior

$$p(\theta|y) = \text{Dirichlet}(\alpha_1 + y_1, \ldots, \alpha_k + y_k)$$

Análogo a beta-binomial

prior conjuado: Dirichlet

$$p(\theta|\alpha) = \mathrm{Dirichlet}(\alpha_1, \ldots, \alpha_k) \propto \prod_{j=1}^{k} \theta_j^{\alpha_j - 1}$$

$$\theta_j \geq 0$$

$$\sum_{j=1}^{k} \theta_j = 1$$

posterior

$$p(\theta|y) = \mathrm{Dirichlet}(\alpha_1 + y_1, \ldots, \alpha_k + y_k)$$

Análogo a beta-binomial

parametrización alternativa:

$$(\alpha_1, \ldots, \alpha_k) \to \alpha(m_1, \ldots, m_k) \qquad \sum_j m_j = 1 \qquad \vec{m}: \ \mathrm{media}$$

prior conjuado: Dirichlet

$$p(\theta|\alpha) = \mathrm{Dirichlet}(\alpha_1, \ldots, \alpha_k) \propto \prod_{j=1}^{k} \theta_j^{\alpha_j - 1}$$

$$\theta_j \geq 0$$

$$\sum_{j=1}^{k} \theta_j = 1$$

posterior

$$p(\theta|y) = \mathrm{Dirichlet}(\alpha_1 + y_1, \ldots, \alpha_k + y_k)$$

Análogo a beta-binomial

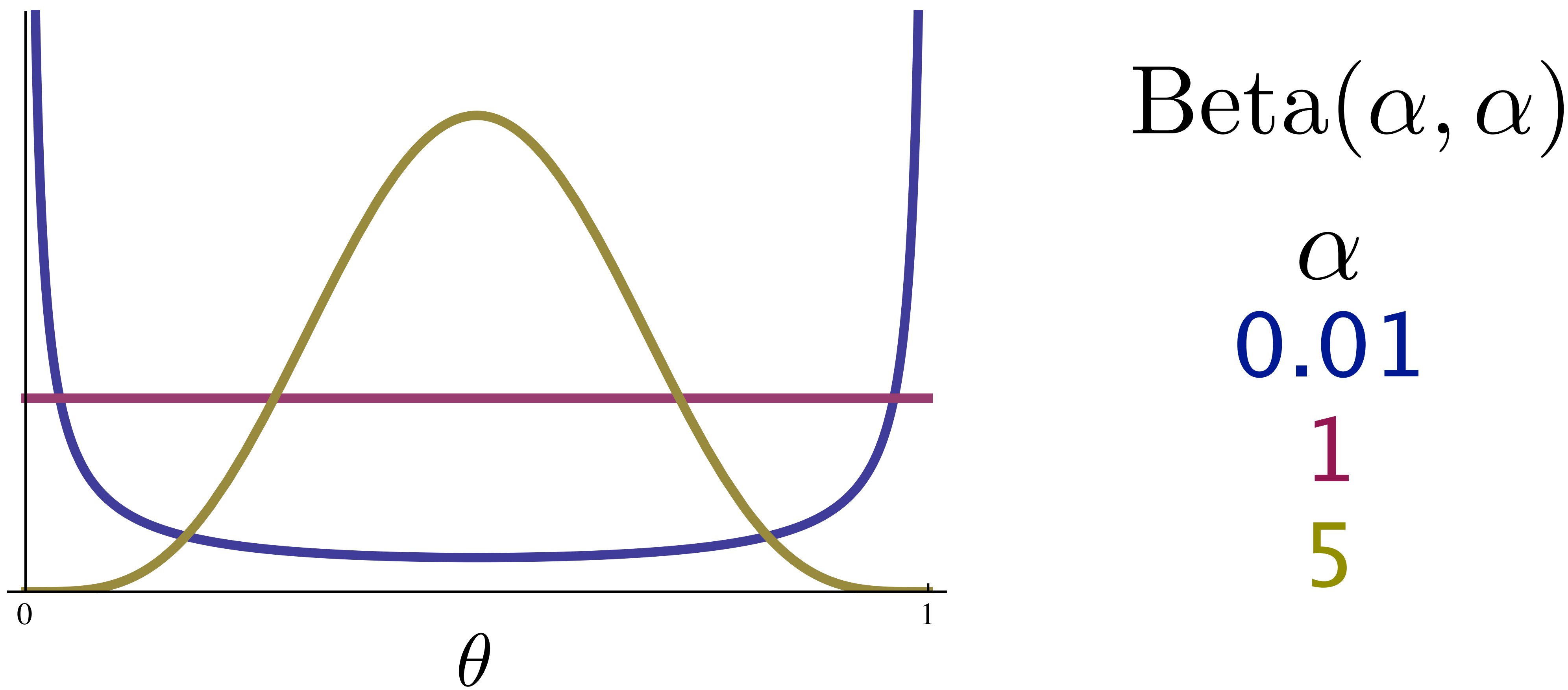
parametrización alternativa:

$$(\alpha_1, \ldots, \alpha_k) \to \alpha(m_1, \ldots, m_k)$$

$$\sum_j m_j = 1 \qquad \vec{m}: \ \text{media}$$

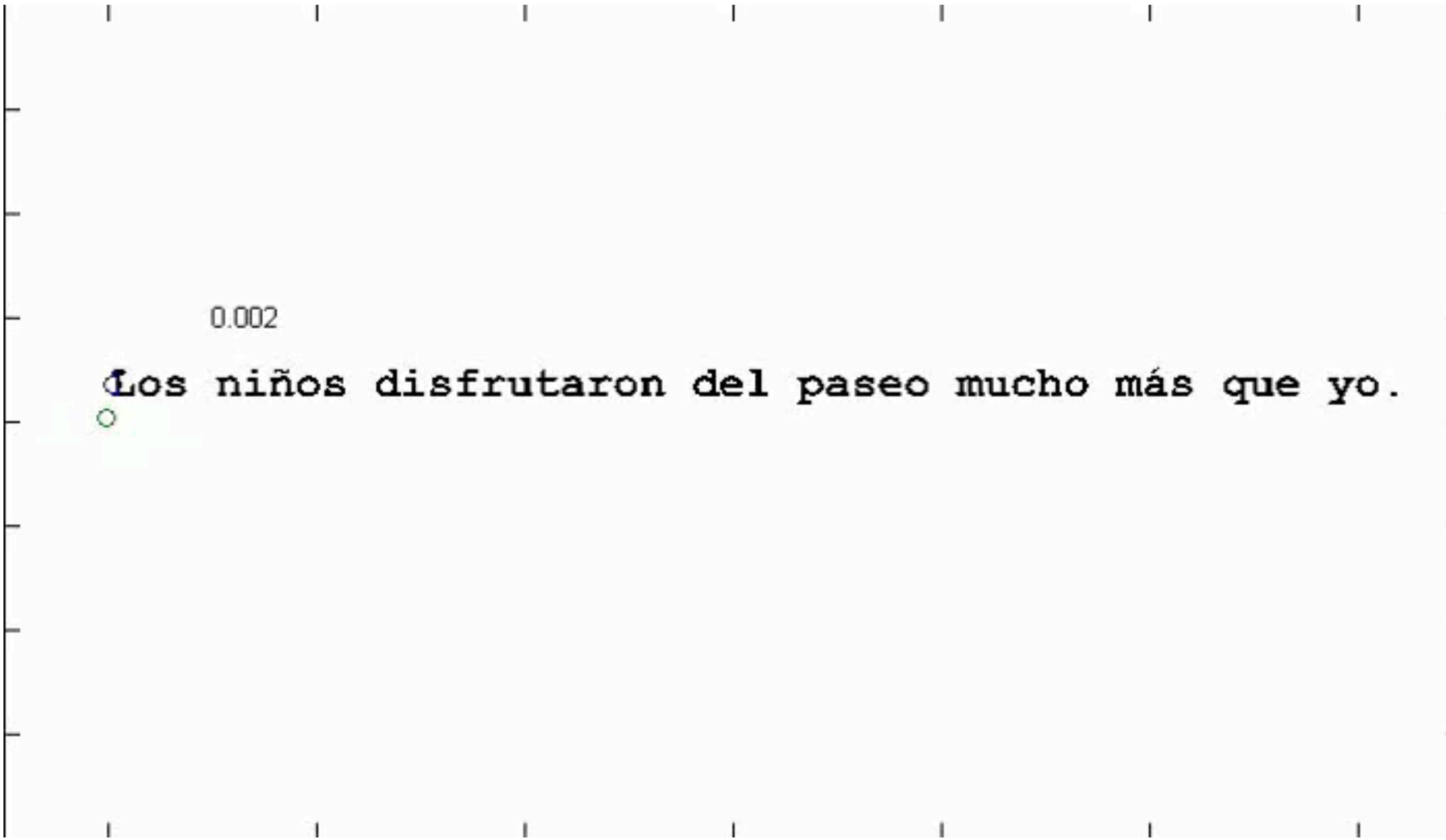
α *sharpness,* cuantifica similitud de las muestras con la media (¡análogo a la precisión!)

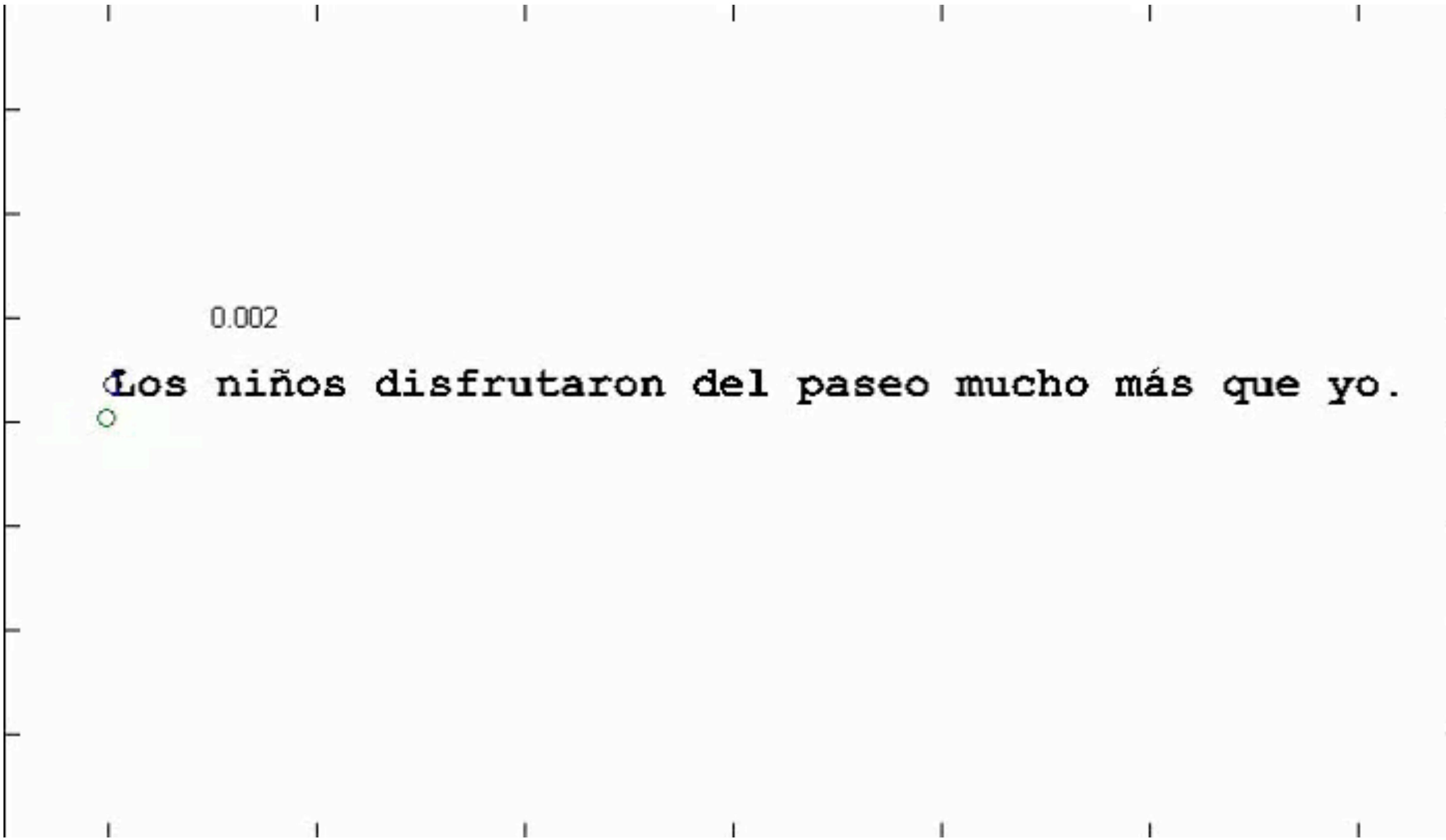
número de muestras para que los datos dominen sobre el *prior*



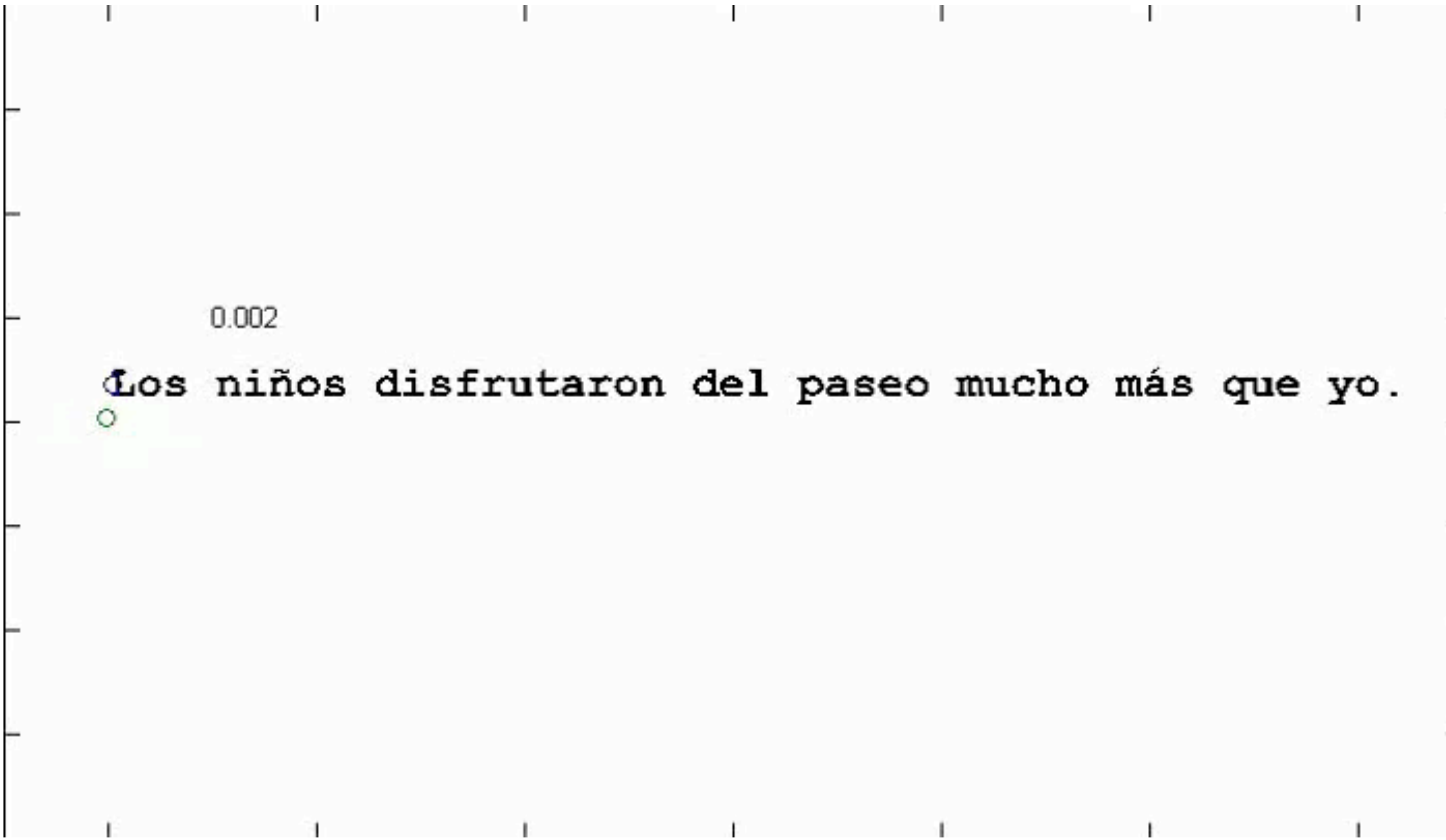
Beta, Dirichlet: distribuciones sobre *distribuciones*

EJEMPLO


0.002
Los niños disfrutaron del paseo mucho más que yo.



Predictibilidad: buen predictor del tiempo pasado en cada palabra en lectura



Predictibilidad: buen predictor del tiempo pasado en cada palabra en lectura

Midiendo la predictibilidad de palabras

“Voy a sacudir el mantel porque está lleno de ______”

Análisis de *Maximum Likelihood*

(Predictabilidad = Fracción de sujetos)

Análisis de *Maximum Likelihood*

(Predictabilidad = Fracción de sujetos)

Limitaciones:

- Sin noción de error/incerteza (10/15 ~ 100/150)
- Predictibilidad del contexto ~ predictibilidad de palabra

Análisis de *Maximum Likelihood*

(Predictabilidad = Fracción de sujetos)

Limitaciones:

- Sin noción de error/incerteza (10/15 ~ 100/150)
- Predictibilidad del contexto ~ predictibilidad de palabra

migas migas migas migas hormigas hormigas hormigas
migas migas migas migas hormigas aceitunas árboles

Midiendo la predictibilidad de palabras 2

Indagando más allá:

“Voy a sacudir el mantel porque está lleno de ______”

Midiendo la predictibilidad de palabras 2

Indagando más allá:

“Voy a sacudir el mantel porque está lleno de ______”

______

______

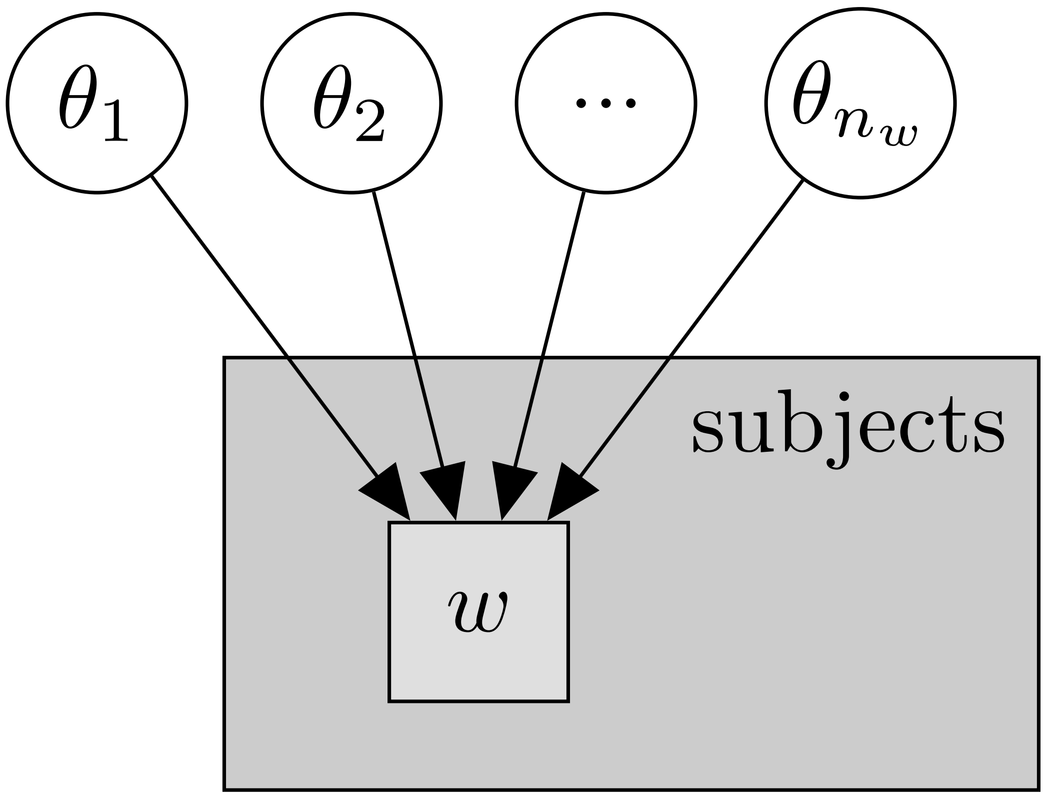
Análisis de *Maximum Likelihood*

(Predictabilidad = Fracción de sujetos)

Limitaciones:

- Sin noción de error/incerteza (10/15 ~ 100/150)
- Predictibilidad del contexto ~ predictibilidad de palabra
- No hace uso de palabras extra

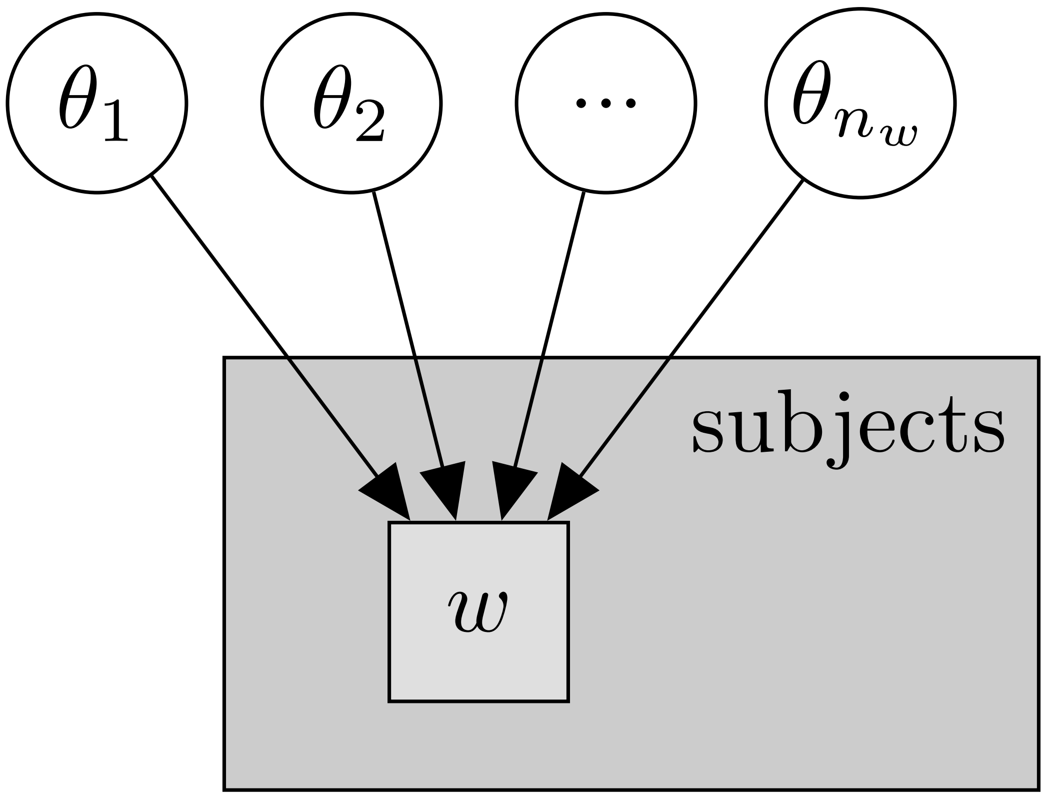
Modelo de una palabra



$$\theta \sim \mathrm{Dirichlet}(1, 1, ..., 1)$$

$$w \sim \mathrm{Categorical}(\theta_1, \theta_2, ..., \theta_{n_{words}})$$

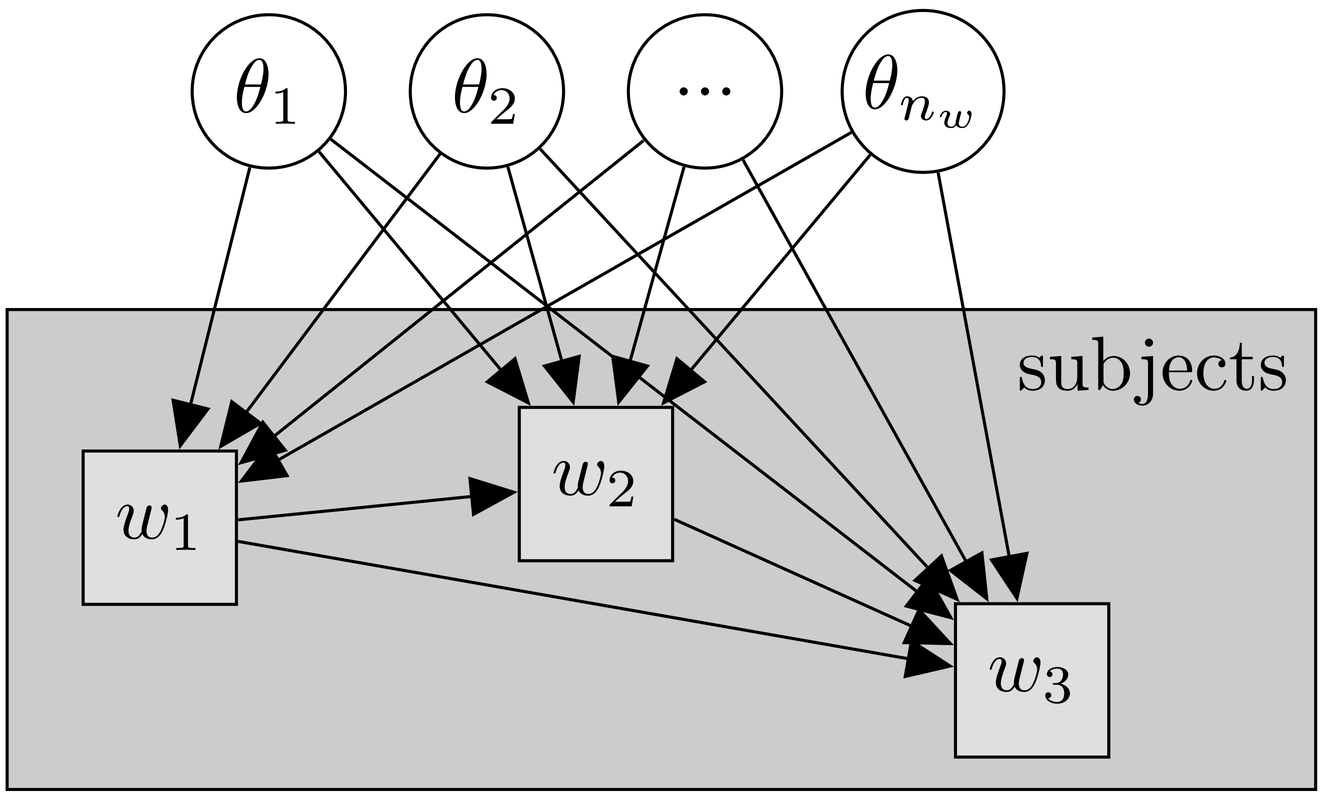
Modelo de una palabra



$$\theta \sim \text{Dirichlet}(1, 1, ..., 1)$$
$$w \sim \text{Categorical}(\theta_1, \theta_2, ..., \theta_{n_{words}})$$

- Noción de incerteza
- Predictibilidad del contexto ~ predictibilidad de palabra
- No hace uso de palabras extra

Modelo de muestreo secuencial



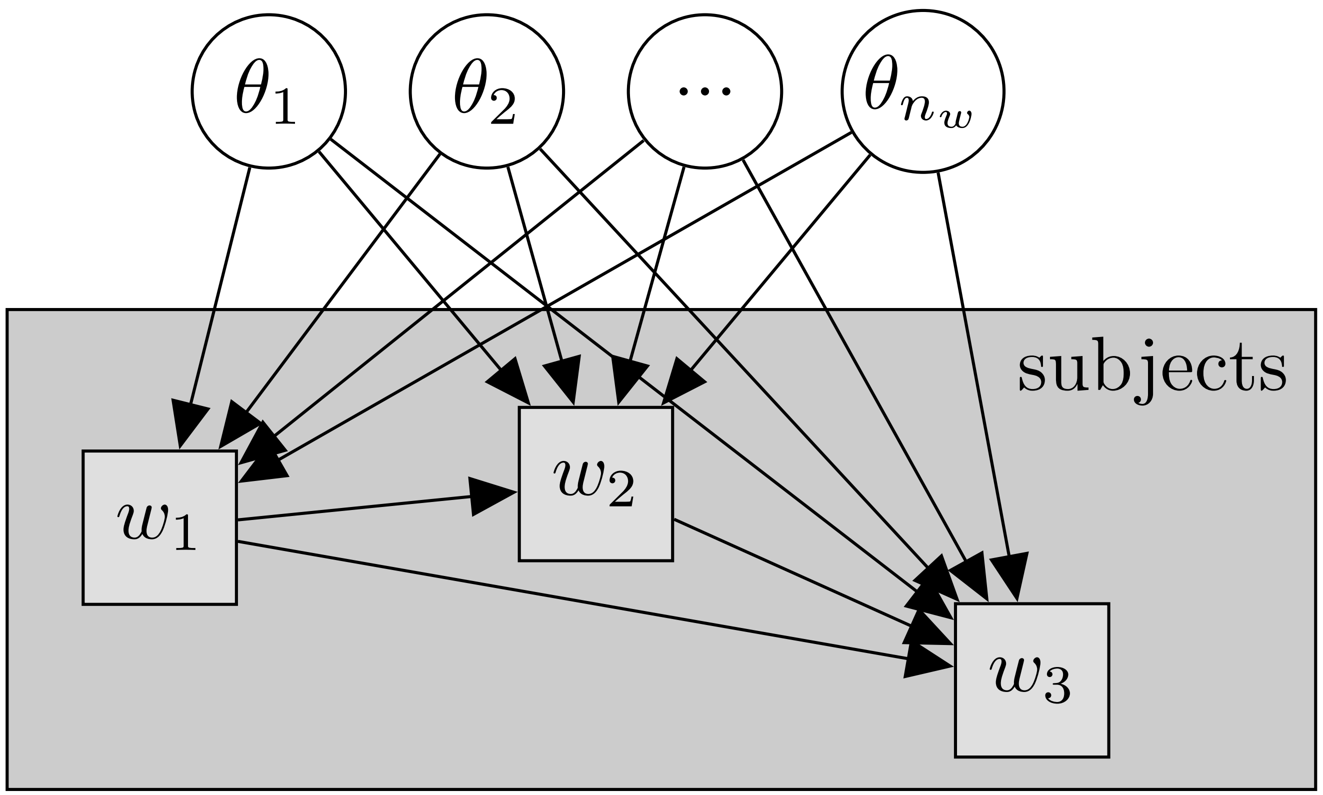
$$\theta \sim \mathrm{Dirichlet}(1, 1, ..., 1)$$

$$w_1 \sim \mathrm{Categorical}(\theta_1, \theta_2, ..., \theta_{n_{words}})$$

$$w_2 \sim \mathrm{Categorical}(\tilde{\theta}_1, \tilde{\theta}_2, ..., \tilde{\theta}_{n_{words}})$$

$$w_3 \sim \mathrm{Categorical}(\tilde{\tilde{\theta}}_1, \tilde{\tilde{\theta}}_2, ..., \tilde{\tilde{\theta}}_{n_{words}})$$

Modelo de muestreo secuencial



$$\theta \sim \mathrm{Dirichlet}(1, 1, ..., 1)$$

$$w_1 \sim \mathrm{Categorical}(\theta_1, \theta_2, ..., \theta_{n_{words}})$$

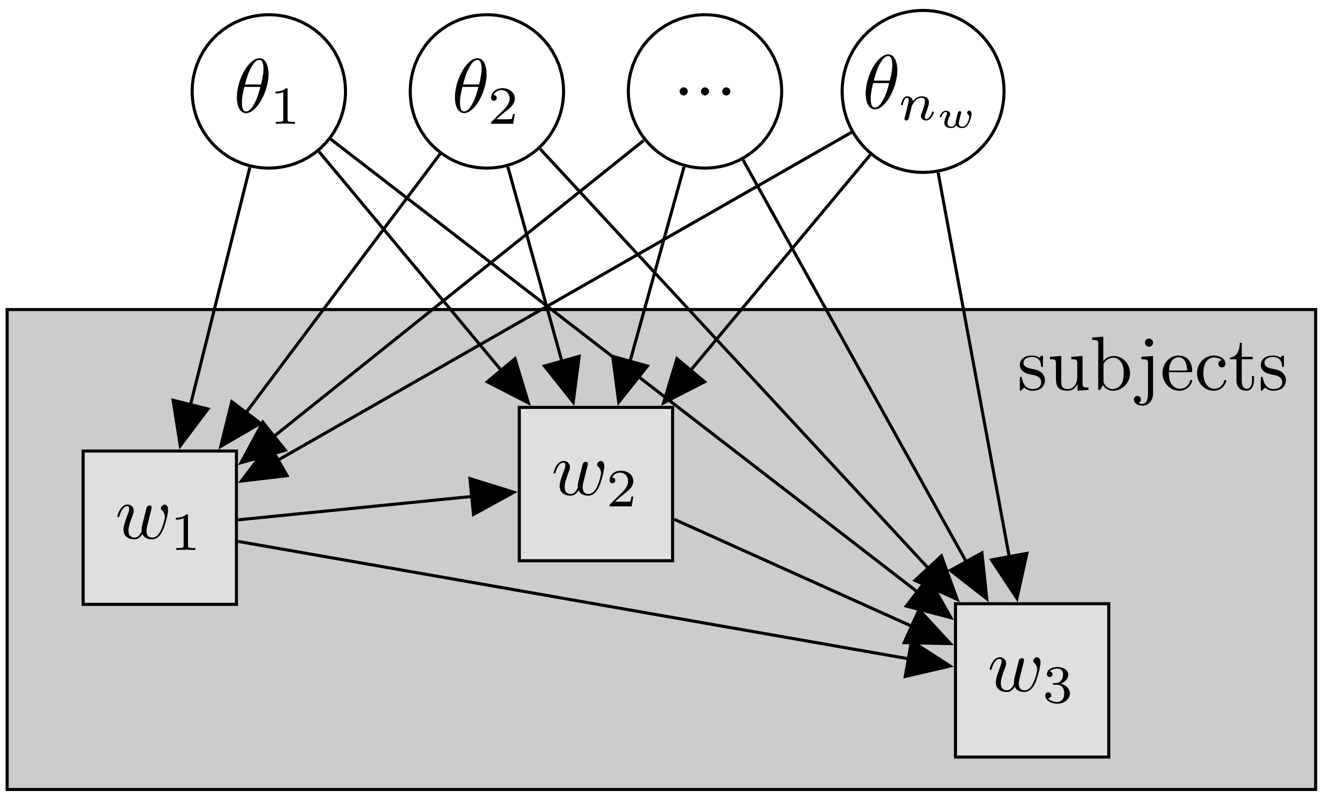
$$w_2 \sim \mathrm{Categorical}(\tilde{\theta}_1, \tilde{\theta}_2, ..., \tilde{\theta}_{n_{words}})$$

$$w_3 \sim \mathrm{Categorical}(\tilde{\tilde{\theta}}_1, \tilde{\tilde{\theta}}_2, ..., \tilde{\tilde{\theta}}_{n_{words}})$$

$$\tilde{\theta}_i = \begin{cases} \frac{\theta_i}{1-\theta_{w_1}} & \text{if } i \neq w_1 \\ 0 & \text{if } i = w_1 \end{cases}$$

$$\tilde{\tilde{\theta}}_i = \begin{cases} \frac{\tilde{\theta}_i}{1-\tilde{\theta}_{w_2}} & \text{if } i \neq w_2 \\ 0 & \text{if } i = w_2 \end{cases}$$

Modelo de muestreo secuencial



$$\theta \sim \text{Dirichlet}(1, 1, ..., 1)$$

$$w_1 \sim \text{Categorical}(\theta_1, \theta_2, ..., \theta_{n_{words}})$$

$$w_2 \sim \text{Categorical}(\tilde{\theta}_1, \tilde{\theta}_2, ..., \tilde{\theta}_{n_{words}})$$

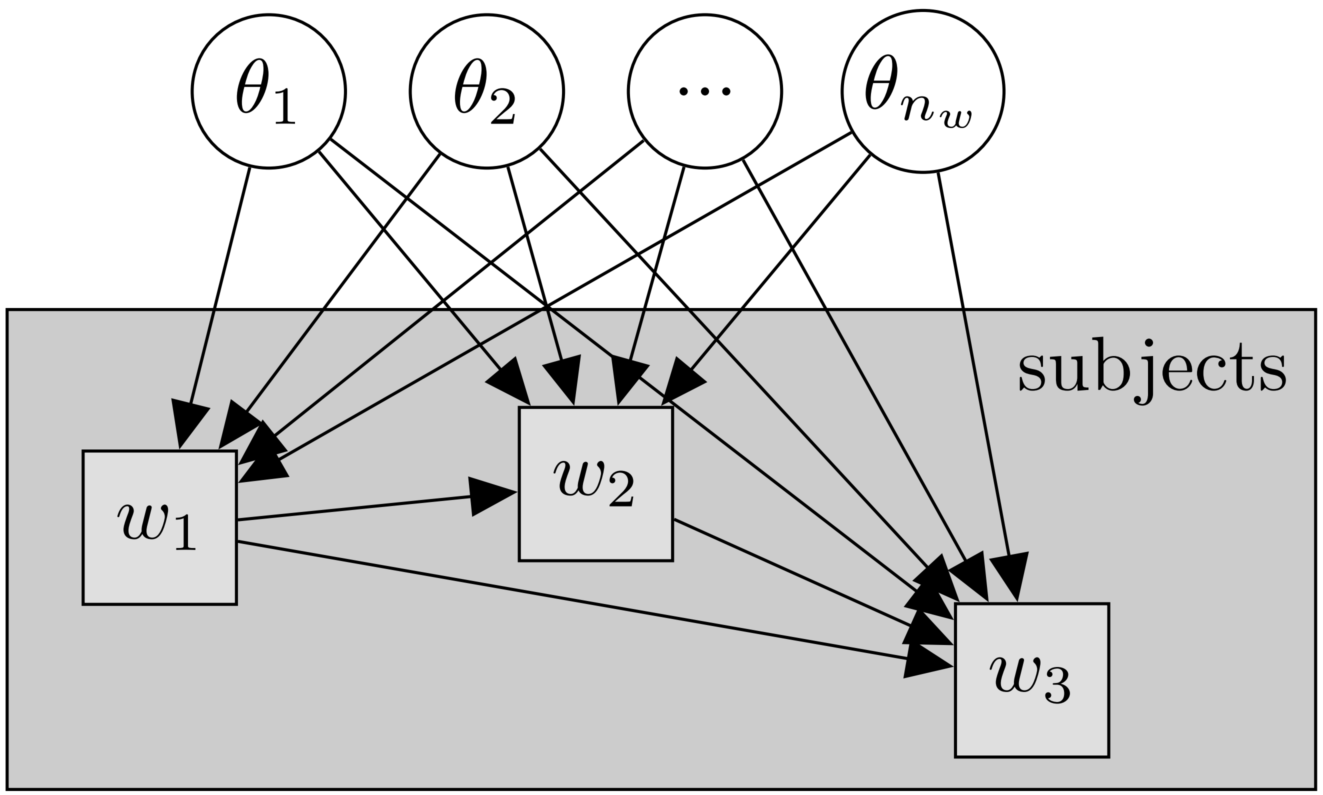
$$w_3 \sim \text{Categorical}(\tilde{\tilde{\theta}}_1, \tilde{\tilde{\theta}}_2, ..., \tilde{\tilde{\theta}}_{n_{words}})$$

$$\tilde{\theta}_i = \begin{cases} \frac{\theta_i}{1-\theta_{w_1}} & \text{if } i \neq w_1 \\ 0 & \text{if } i = w_1 \end{cases}$$

$$\tilde{\tilde{\theta}}_i = \begin{cases} \frac{\tilde{\theta}_i}{1-\tilde{\theta}_{w_2}} & \text{if } i \neq w_2 \\ 0 & \text{if } i = w_2 \end{cases}$$

- Noción de incerteza
- Predictibilidad del contexto ~ predictibilidad de palabra
- Usa palabras extra

Modelo de muestreo secuencial



$$\theta \sim \text{Dirichlet}(1, 1, ..., 1)$$
$$w_1 \sim \text{Categorical}(\theta_1, \theta_2, ..., \theta_{n_{words}})$$
$$w_2 \sim \text{Categorical}(\tilde{\theta}_1, \tilde{\theta}_2, ..., \tilde{\theta}_{n_{words}})$$
$$w_3 \sim \text{Categorical}(\tilde{\tilde{\theta}}_1, \tilde{\tilde{\theta}}_2, ..., \tilde{\tilde{\theta}}_{n_{words}})$$

$$\tilde{\theta}_i = \begin{cases} \frac{\theta_i}{1-\theta_{w_1}} & \text{if } i \neq w_1 \\ 0 & \text{if } i = w_1 \end{cases}$$

$$\tilde{\tilde{\theta}}_i = \begin{cases} \frac{\tilde{\theta}_i}{1-\tilde{\theta}_{w_2}} & \text{if } i \neq w_2 \\ 0 & \text{if } i = w_2 \end{cases}$$

- Noción de incerteza
- Predictibilidad del contexto ~ predictibilidad de palabra
- Usa palabras extra

Modelos jerárquicos