

# Introducción a Estadística Bayesiana

Taller de filogenética bayesiana en RevBayes

Enero 2025

# ¿Qué es la estadística Bayesiana?

- Rama del estudio de la estadística que se enfoca en la inferencia y pronóstico de la probabilidades de eventos inciertos.
- Objetivo: Proporcionar una medida de la incertidumbre y no solamente un estimador puntual

Dos diferencias importantes de la estadísticas bayesiana con estadística tradicional

1.

2.

# ¿Cuándo hacemos estadística bayesiana y cuando no?

- En Biología, escogemos hacer inferencia bayesiana porque es una herramienta computacional poderosa
- Utilizar e implementar MCMCs es un arte, toma tiempo y esfuerzo, pero vale la pena por la calidad de la inferencia

# Conceptos bayesianos importantes

Dados dos eventos  $A$  y  $B$  la probabilidad condicional de  $A$  **dado**  $B$

se define como.  $P(A|B) = \frac{P(AB)}{P(B)}$

y aplicando dos veces el teorema de Bayes obtenemos que

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

## Ejemplo:

### El experimento de un reality show “El amor es ciego”

- Vamos a ver un reality show que se llama “El amor es ciego”, como estadísticos bayesianos vamos a estimar la probabilidad de que uno de los participantes se enamore y se case.
- El experimento: Silvia tiene una serie de citas a ciegas hasta que encuentre al amor de su vida y se case. Silvia debe tener un índice alto de carisma para casarse
- Como estadísticos, vamos a estimar la probabilidad de que Silvia tenga gran carisma.

# Definición de Eventos

# Definición de probabilidad *a priori*

- Como expectadores y estadísticos nosotros podemos definir de entrada algunas probabilidades. Por ejemplo, en el primer capítulo, se hace la introducción de Silvia y de entrada no sabemos mucho de Silvia. Dos decisiones que podemos tomar:

1.

2.

**La probabilidad a priori es un paso necesario de la estadística bayesiana que refleja nuestras creencias y es subjetiva, puede ser informativa o no informativa**

## El objetivo final: $P(A|B)$

- Silvia ¿acabará enamorándose?, en términos de probabilidad nos interesa lo que llamamos la probabilidad a posteriori, es decir

$P(A|B)$  = la probabilidad del carisma de Silvia dadas las citas exitosas durante el reality show.



## $P(A|B)$ es difícil de estimar

- Lo más sencillo es pensar en la probabilidad siguiente:

$$P(B|A)$$

Dado que conocemos el carisma de Silvia, la probabilidad de la cita es.

Esta segunda probabilidad condicional  $P(B|A)$  es muchísimo más sencilla de entender y de pensar. Afortunadamente el teorema de Bayes liga esta probabilidad más sencilla a la más difícil

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

# Variables aleatorias

Las variables aleatorias son una **función matemática** que nos lleva del espacio de eventos al espacio de los números.

$A$  : Silvia tiene carisma

$X$ :

$$X = \begin{cases} 0 & \text{cuando } A^C \text{ Silvia no tiene carisma} \\ 1 & \text{cuando } A \text{ Silvia tiene carisma} \end{cases}$$

# Variables aleatorias

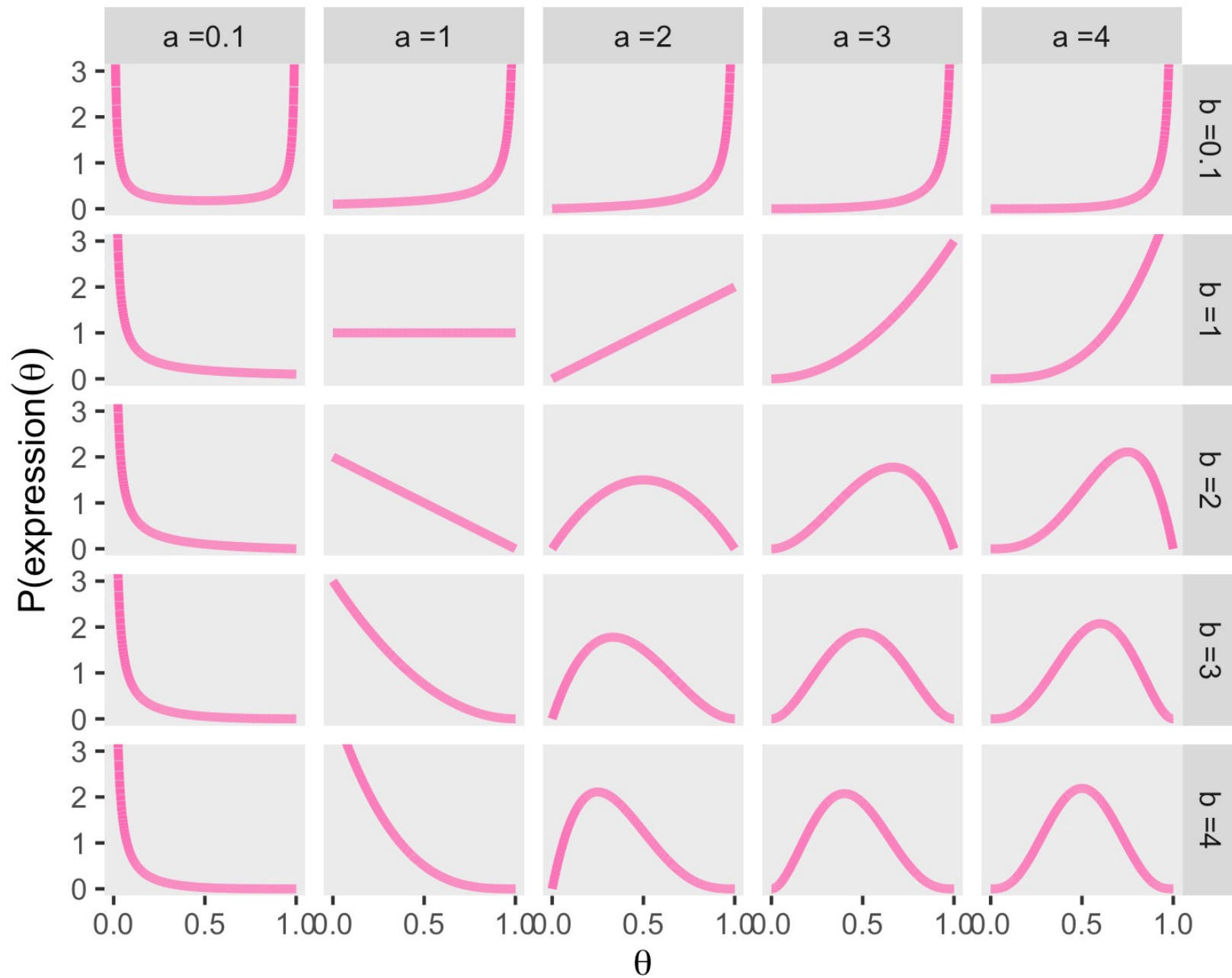
- $P(A) = P(X = 1) = \theta$
- $P(A^c) = P(X = 0) = 1 - \theta$

En el mundo real el carisma de una persona fluctua dependiendo de muchos factores.

Entonces una mejor aproximación para medir el carisma de Silvia es asumir que

# Elicitación de la distribución a priori del carisma de Silvia





**Los parámetros son desconocidos pero inciertos, necesitan una distribución de probabilidad**

# Los datos, las citas a ciegas

- Silvia tiene tres citas completamente a ciegas con la misma persona (ninguno se ve solo se escuchan el uno al otro). En nuestros términos estadísticos

$$N = 3$$

es el número total de citas con la misma persona

¿Cómo modelamos los resultados de estas tres citas, sabiendo el carisma de Silvia? Si recordamos el evento

$B$  = es el número de citas que fueron bien

Proponemos la variable aleatoria

$Y$

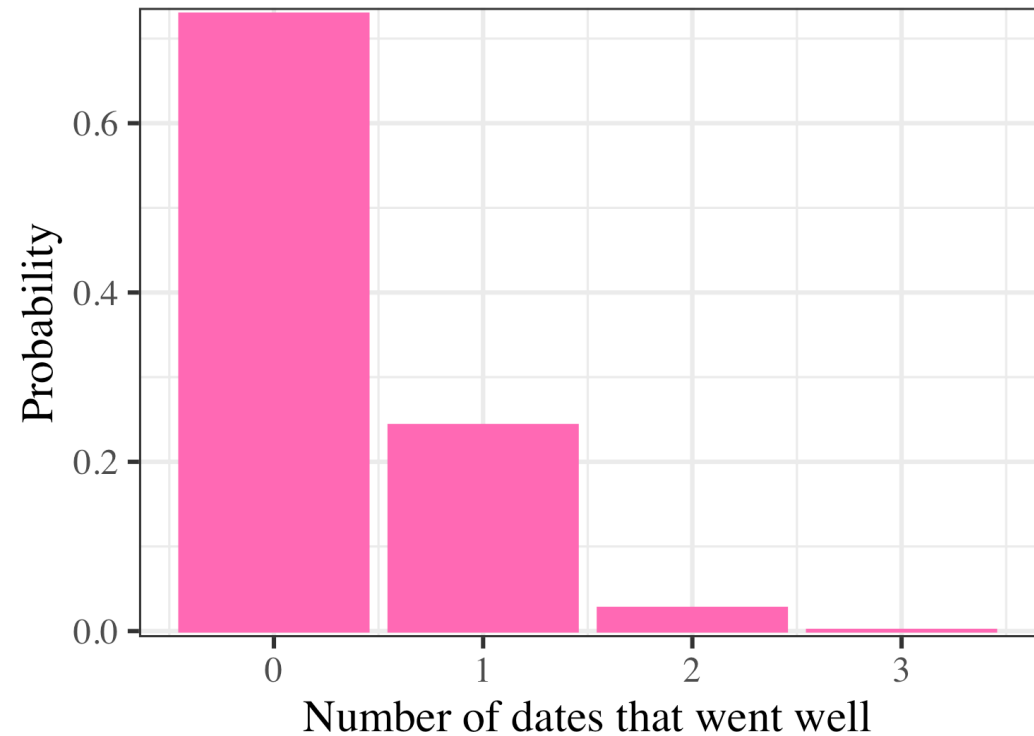
# Distribución Binomial

# Distribución Binomial

- [https://ixchelgzlzl.github.io/filo\\_bayes\\_UNAM/docs/intro\\_bayesian/Intro\\_bayesiana](https://ixchelgzlzl.github.io/filo_bayes_UNAM/docs/intro_bayesian/Intro_bayesiana)



Binomial distribution





# Verosimilitud de una cita

Primera persona dos de las tres citas salieron bien. La probabilidad de que dos citas salieron bien dado que sabemos el carisma de Silvia es

$$P(Y = 2|\theta) = \binom{3}{2} \theta^2 (1 - \theta)^1$$

Si

$$\theta = 0.1$$

entonces

$$P(Y = 2|\theta = 0.1) =$$

Este es un escenario **inverosímil**

# Verosimilitud con dos citas

$$y_1 = 2 \quad y_2 = 1$$

**La función de verosimilitud (likelihood function)**

# Verosimilitud con dos citas

**La función de verosimilitud** se define como la probabilidad de los datos dado el modelo.

$$P(\text{Datos}|\theta) = \prod_{y_i=1}^2 P(Y = y_i|\theta) \approx \theta^3(1 - \theta)^3$$

# Verosimilitud

- [https://ixchelgzlzl.github.io/filo\\_bayes\\_UNAM/docs/intro\\_bayesian/Intro\\_bayesiana](https://ixchelgzlzl.github.io/filo_bayes_UNAM/docs/intro_bayesian/Intro_bayesiana)



$$P(Datos|\theta) = \prod_{y_i=1}^2 P(Y = y_i|\theta) \approx \theta^3(1 - \theta)^3$$

# Inferencia bayesiana: La distribución posterior

En el nuevo lenguaje de nuestras variables aleatorias

$$P(\theta|Datos)$$

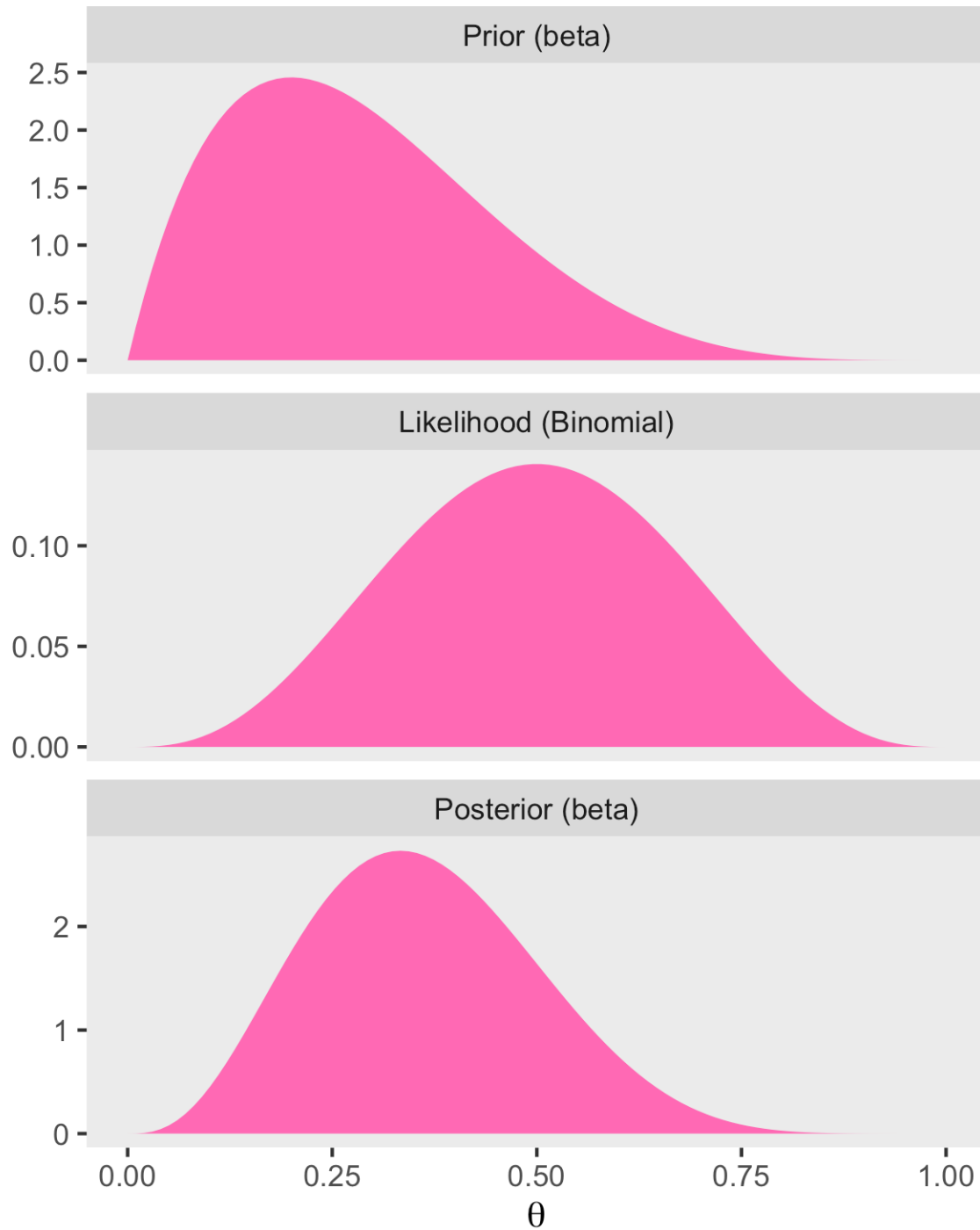
La distribución posterior es **proporcional** a la verosimilitud multiplicada por la distribución a priori

$$P(\theta|Data) \propto P(Data|\theta) \times P(\theta)$$

⏟

⏟

⏟



Comparemos las tres  
funciones que ya  
calculamos

# Inferencia Bayesiana utilizando el algoritmo MCMC

El algoritmo MCMC (Markov Chain Monte Carlo) nos permite optimizar la distribución posterior

Metropolis- Hastings- Un tipo de algoritmo del MCMC

$\theta$

- Propongo un valor de  $\theta$  : ¿incrementa el valor de la probabilidad posterior?-
- Si la respuesta es positiva, nos quedamos este valor, y si es negativa lo rechazamos.

# Propuesta- Un nuevo valor de $\theta$

```
proposalfunction <- function(nvals=1){  
  unif_val<-runif(nvals,min=0, max=1)  
  return(unif_val)}  
# Esta es una propuesta de una distribución uni  
# forme. Selecciona valores aleatorios entre 0 y  
# 1
```

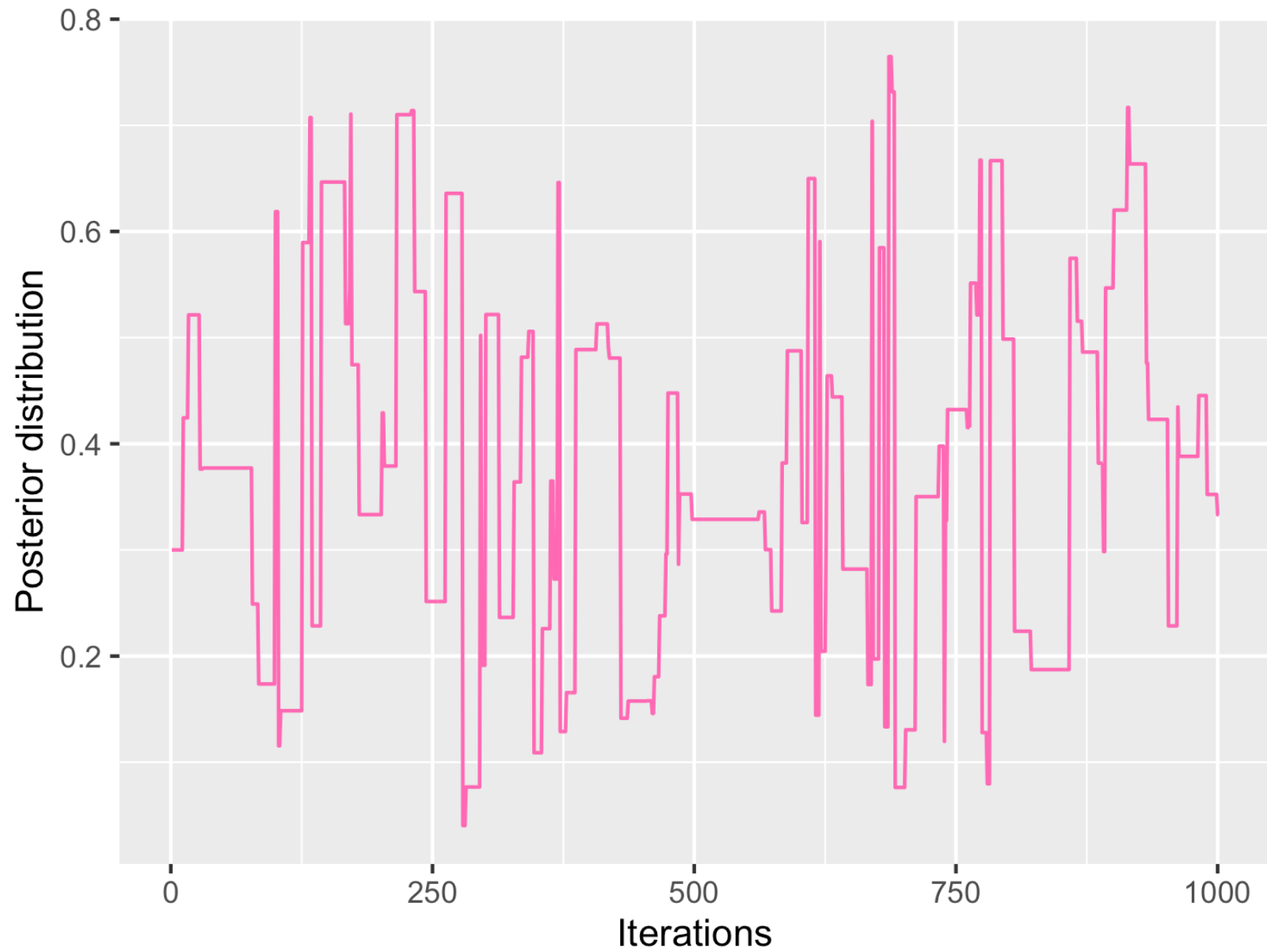
¿Qué está pasando en esta parte del código?



# Metropolis-Hastings

1. Empieza con un valor para  $\theta$  (startvalue) llamado  $\theta_0$
2. Haz  $\theta_{old} = \theta_0$
3. Calcula la distribución posterior  $P(\theta_{old}|Datos)$
4. Proponer una distribución aleatoria  $g(\theta)$  para obtener un nuevo valor  $\theta_{new}$
5. Calcula la distribución posterior  $P(\theta_{new}|Datos)$
6. Calcula los momios  $momios = \frac{P(\theta_{new}|Datos)}{P(\theta_{old}|Datos)}$
7. Calcula un valor aleatorio  $u$  entre 0 y 1
8. Si  $u < momios$  entonces acepta.  $\theta_{new}$ , guárdalo, sino rechaza y no lo guardes.
9. Si lo aceptas haz  $\theta_{old} = \theta_{new}$  y vuelve al paso dos hasta acabar las iteraciones. Sino continua al paso dos con el mismo  $\theta_{old}$  hasta acabar las iteraciones.

## MCMC run



¿Cómo saber si el MCMC  
encontró la posterior?

## Inferencia bayesiana: La distribución posterior

$$P(\theta|D_{\text{atos}})$$

