

# Practica 1

Juan Cantero Jimenez

2/27/2022

## Introducción

Los sistemas de funciones iteradas, SFI, son construcciones matemáticas que permiten representar conjuntos de puntos que presentan autosimilaridad, autoafinidad o autocorrelación. En este caso estamos interesados en SFI que apliquen transformaciones afijas, así estas SFI tomaran la forma:

$$f_1(x, y) = \begin{bmatrix} c_1 & c_2 \\ c_3 & c_4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} c_5 \\ c_6 \end{bmatrix} \quad (1)$$

$$f_2(x, y) = \begin{bmatrix} c_7 & c_8 \\ c_9 & c_{10} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} c_{11} \\ c_{12} \end{bmatrix} \quad (2)$$

de forma más específica se ha representado un SFI con dos atractores. Para la representación de el atractor codificado en este SFI se hace uso del algoritmo de iteración aleatoria. Primero se escoge un punto aleatorio del espacio métrico, en este caso de dos dimensiones, seleccionado el punto se escoge una de las funciones que pertenecen al SFI de forma aleatoria, se aplica esta función al punto inicial. Este mismo procedimiento se repetira sobre el resultado de la primera iteración. Las coordenadas de los puntos en las distintas iteraciones se representan en el espacio métrico obteniendose la representación del atractor codificado en el SFI.

En este trabajo se hace uso del algoritmo propuesto en (1) para la obtención de motivos visualmente relevantes. Así se generan los coeficientes del SFI de forma aleatoria, se obtiene la representación por el metodo de iteración aleatoria mientras se va evaluando la calidad visual de estos mediante dos coeficientes, la dimension fractal y el exponente máximo de Lyapunov. Siguiendo con las conclusiones obtenidas en (1) aquellos atractores visualmente relevantes serán los que posean una dimensión fractal mayor que 1 y un exponente máximo de Lyapunov menor que -0.2. Si cumple estos requisitos la representación se cosiderará visualmente relevante, exito, en caso contrario no se considera relevante, fracaso. Las Figuras 1, 2 y 3 muestran tres SFI visualmente relevantes.

En este trabajo se quiere hacer inferencia sobre la proporción de representaciones visualmente relevantes sobre el total de las generadas. Esta inferencias se realizará para SFI con 2, 3 y 4 atractores, con el objetivo de comprar las proporciones en función del número de atractores de sus SFI. Los datos han sido generados en Python 3.8.

## Inferencia Bayesiana sobre la proporción de representaciones visualmente relevantes

### Lectura de los datos

Primero se procede a cargar los datos en R:

```
data <- read.csv("datos_investigacion.csv")
```

y se separan los datos para 2, 3 y 4 atractores en distintas variables:



Figure 1: Ejemplo de representación gráfica de un IFS con dos atractores

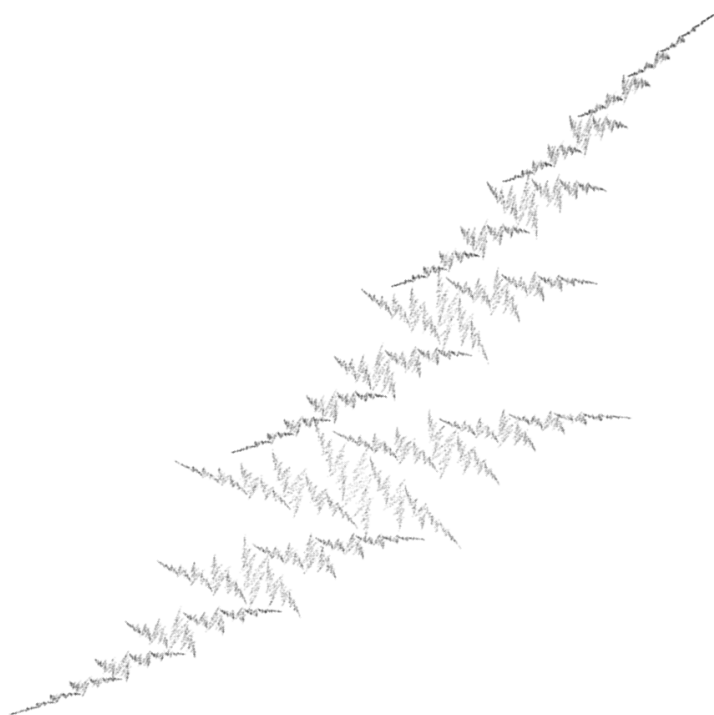


Figure 2: Ejemplo de representación gráfica de un IFS con dos atractores

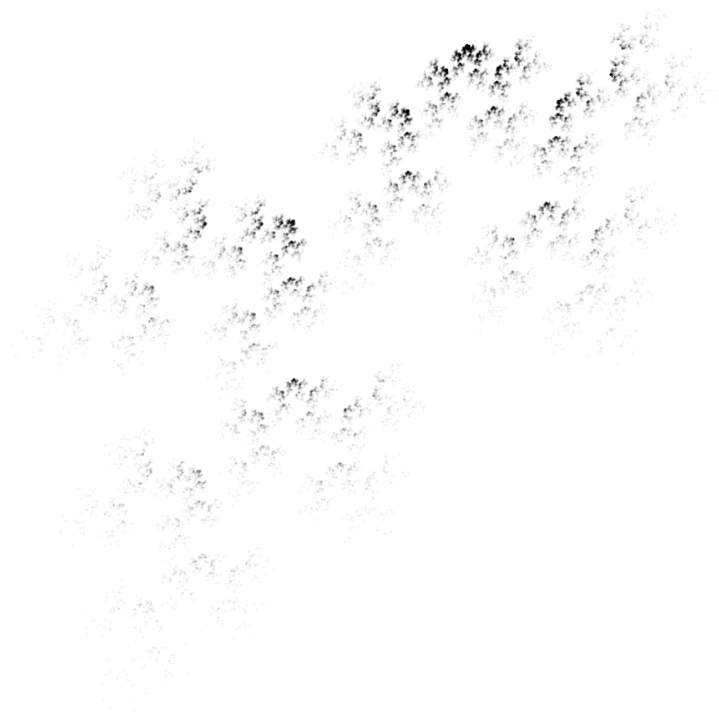


Figure 3: Ejemplo de representación gráfica de un IFS con dos atractores

```
head(data)
```

```
##   X          L          FD Atractores          Id
## 1 0  0.16250242 0.6518040          2          YVMPMRMJXTLY
## 2 1 -0.02183638 1.5843295          3          FEMSSNFWNXVPUNYNPM
## 3 2  0.20041153 1.3935737          4 JHEQSAKIYOBMKVXJRJBCLKKL
## 4 3          NA 0.0000000          2          UCYIJKTYGEYX
## 5 4 -0.67532628 0.7113785          3          PHTMMOPVKKL VUHTGKH
## 6 5  0.30137549 2.2329936          4 AFKJPWBQDTQKJHSYHCYNIKIG
```

```
str(data)
```

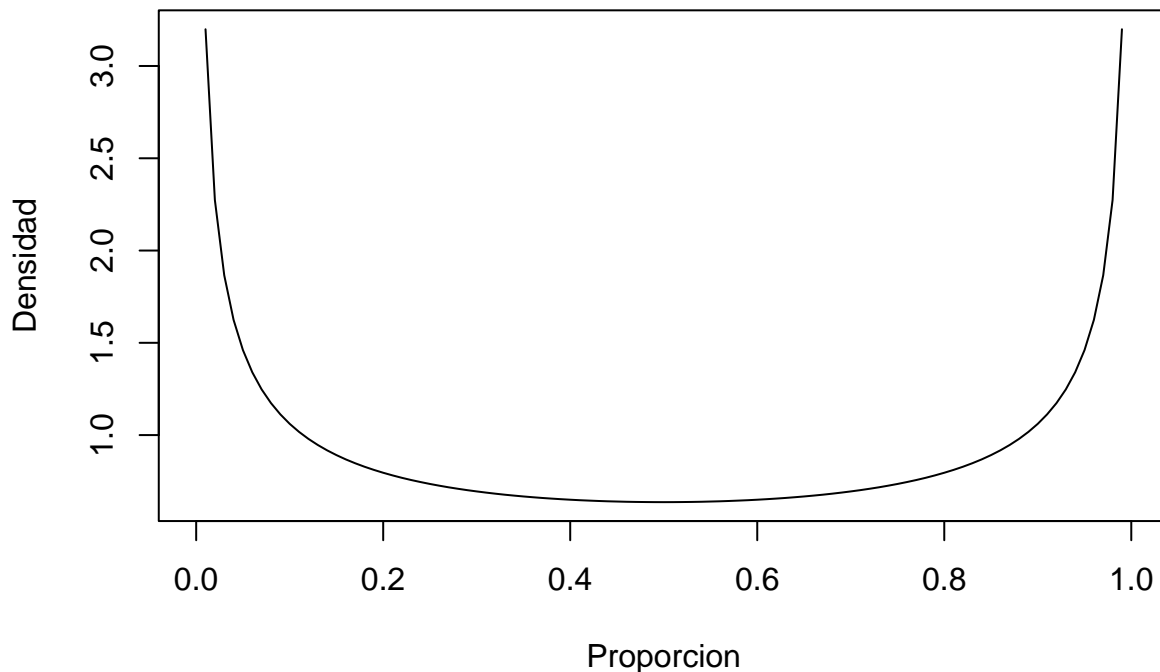
```
## 'data.frame':   300000 obs. of  5 variables:
## $ X           : int  0 1 2 3 4 5 6 7 8 9 ...
## $ L           : num  0.1625 -0.0218 0.2004 NA -0.6753 ...
## $ FD          : num  0.652 1.584 1.394 0 0.711 ...
## $ Atractores: int   2 3 4 2 3 4 2 3 4 2 ...
## $ Id          : chr   "YVMPMRMJXTLY" "FEMSSNFWNXVPUNYNPM" "JHEQSAKIYOBMKVXJRJBCLKKL" "UCYIJKTYGEYX" ..
atrac.2 <- data[which(data$Atractores==2),]
atrac.3 <- data[which(data$Atractores==3),]
atrac.4 <- data[which(data$Atractores==4),]
```

### Inferencia sobre la proporción de representaciones relevantes con 2 atractores.

La distribución previa sobre la proporción de representaciones visualmente relevantes será no informativa con la forma de  $\pi(\theta_{2A}) \sim Be(1/2, 1/2)$ . Esta toma la forma:

```
alpha0.2 <- 0.5
beta0.2 <- 0.5
curve(dbeta(x, alpha0.2, beta0.2), main = "Distribución previa sobre la proporción", xlab="Proporcion",
```

### Distribución previa sobre la proporción



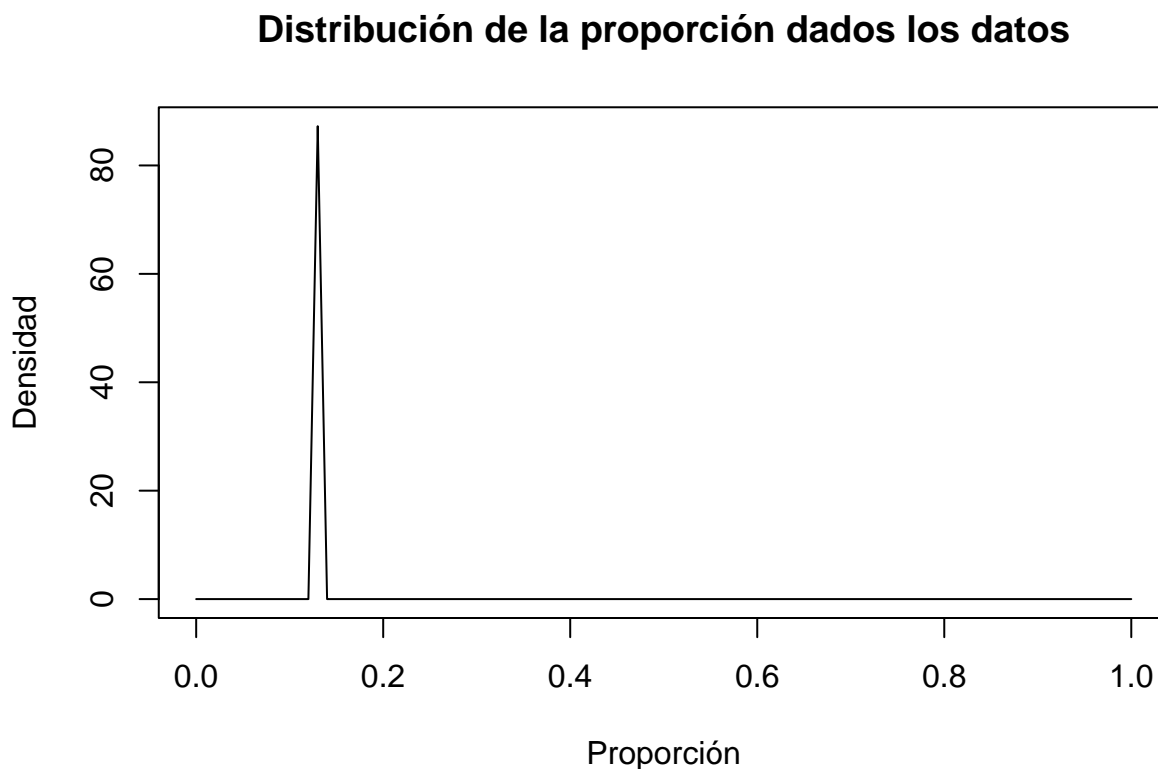
La distribución a posteriori de la proporción, es decir  $\pi(\theta_{2A} \mid \text{datos})$  se puede obtener como:

```
n.2 <- nrow(atrac.2)
#& atrac.2$FD != Inf & atrac.2$FD != -Inf & atrac.2$L != -Inf & atrac.2$L != Inf
 exitos.2 <- length(which(atrac.2$L <= -0.2 & atrac.2$FD >= 1 & atrac.2$FD != Inf & atrac.2$FD != -Inf &
 exitos.2

## [1] 13182

alpha.2 <- alpha0.2 + exitos.2
beta.2 <- beta0.2 + n.2 - exitos.2

### Representación gráfica
curve(dbeta(x, alpha.2, beta.2), main="Distribución de la proporción dados los datos", xlab="Proporción"
```



```
### Esperanza

media.2 <- alpha.2 / (alpha.2 + beta.2)
cat("Media", media.2)

## Media 0.1318237

dt.2 <- sqrt(media.2 * (1 - media.2) / (alpha.2 + beta.2 + 1))
cat("Desviación típica", dt.2)

## Desviación típica 0.001069785

### Intervalo de alta probabilidad 95 % sobre la media

result <- data.frame(low=qbeta(0.025, alpha.2, beta.2),
                     media=media.2,
                     upp=qbeta(0.975, alpha.2, beta.2))
result
```

```
##           low      media      upp
## 1 0.1297339 0.1318237 0.1339274
```

Así obtenemos que  $E(\theta_{2A} \mid \text{datos}) = 0.1318$ , y  $DT(\theta_{2A} \mid \text{datos}) = 0.001069785$  también sabemos que  $P(0.1297339 < E(\theta_{2A} \mid \text{datos}) = 0.1318 < 0.1339274) = 0.95$ .

## Inferencia sobre la proporción de representaciones relevantes con 3 atractores.

En este caso se utilizará la misma distribución a priori sobre la proporción que en el caso para 2 atractores. La distribución a posteriori dado los datos,  $\pi(\theta_{3A} \mid \text{datos})$ , puede ser obtenida como:

```
n.3 <- nrow(atrac.3)
alpha0.3 <- 0.5
beta0.3 <- 0.5
#& atrac.3$FD != Inf & atrac.3$FD != -Inf & atrac.3$L != -Inf & atrac.3$L != Inf
 exitos.3 <- length(which(atrac.3$L <= -0.2 & atrac.3$FD >= 1 & atrac.3$FD != Inf & atrac.3$FD != -Inf &
 exitos.3
```

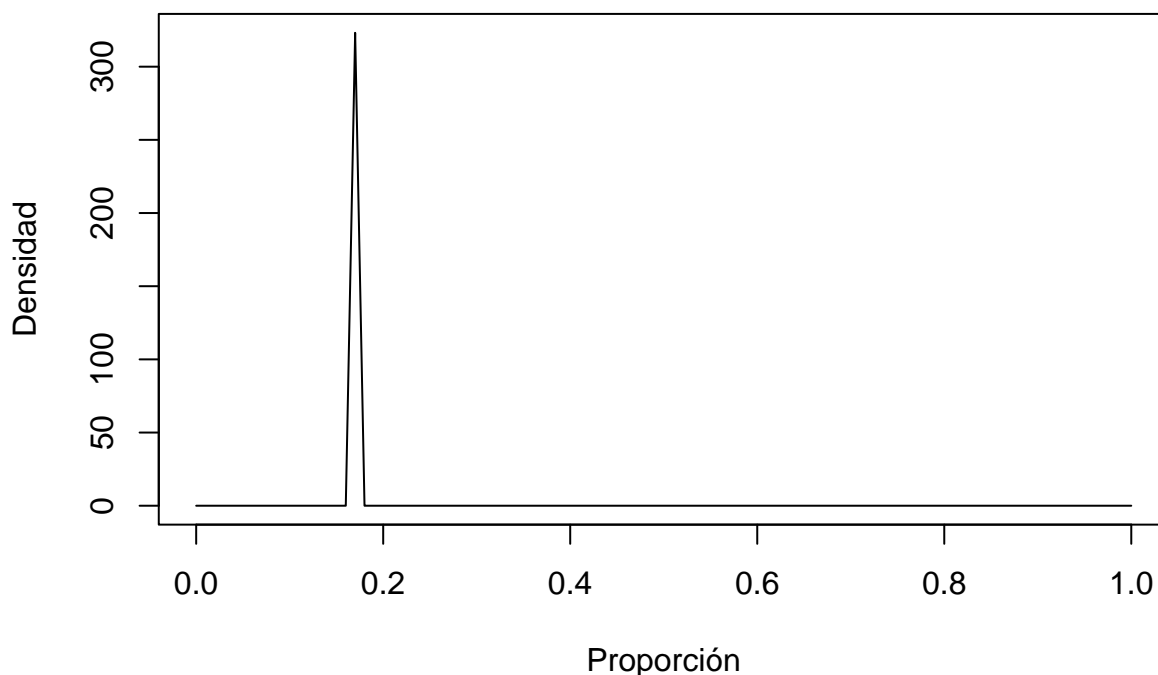
```
## [1] 16967
```

```
alpha.3 <- alpha0.3 + exitos.3
beta.3 <- beta0.3 + n.3 - exitos.3
```

```
### Representación gráfica
```

```
curve(dbeta(x, alpha.3, beta.3), main="Distribución de la proporción dados los datos", xlab="Proporción"
```

## Distribución de la proporción dados los datos



```
### Esperanza
```

```
media.3 <- alpha.3 / (alpha.3 + beta.3)
cat("Media", media.3)
```

```
## Media 0.1696733
```

```
dt.3 <- sqrt(media.3*(1-media.3)/(alpha.3 + beta.3 + 1))
cat("Deviación típica", dt.3)
```

```
## Deviación típica 0.001186935
```

```
### Intervalo de alta probabilidad 95 % sobre la media
```

```
result <- data.frame(low=qbeta(0.025, alpha.3, beta.3),
                     media=media.3,
                     upp=qbeta(0.975, alpha.3, beta.3))
result
```

```
##           low      media      upp
## 1 0.1673532 0.1696733 0.1720059
```

Así obtenemos que  $E(\theta_{3A} \mid \text{datos}) = 0.1696733$ , y  $DT(\theta_{3A} \mid \text{datos}) = 0.001186935$  también sabemos que  $P(0.1673532 < E(\theta_{3A} \mid \text{datos}) = 0.1696733 < 0.1720059) = 0.95$ .

## Inferencia sobre la proporción de representaciones relevantes con 4 atractores.

En este caso se utilizará la misma distribución a priori sobre la proporción que en el caso para 2 atractores. La distribución a posteriori dado los datos,  $\pi(\theta_{4A} \mid \text{datos})$ , puede ser obtenida como:

```
n.4 <- nrow(atrac.4)
```

```
alpha0.4 <- 0.5
```

```
beta0.4 <- 0.5
```

```
#& atrac.4$FD != Inf & atrac.4$FD != -Inf & atrac.4$L != -Inf & atrac.4$L != Inf
```

```
 exitos.4 <- length(which(atrac.4$L <= -0.2 & atrac.4$FD >= 1 & atrac.4$FD != Inf & atrac.4$FD != -Inf &
```

```
 exitos.4
```

```
## [1] 16127
```

```
alpha.4 <- alpha0.4 + exitos.4
```

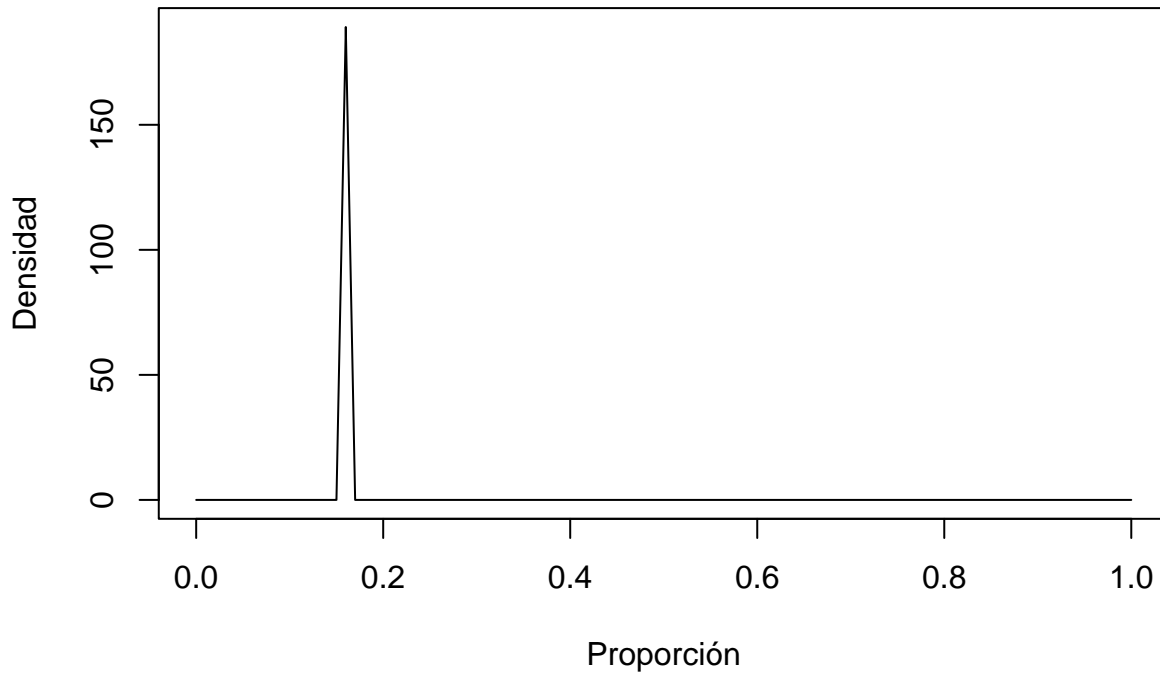
```
beta.4 <- beta0.4 + n.4 - exitos.4
```

```
### Representación gráfica
```

```
curve(dbeta(x, alpha.4, beta.4), main="Distribución de la proporción dados los datos", xlab="Proporción"
```



## Distribución de la proporción dados los datos



### ### Esperanza

```
media.4 <- alpha.4 / (alpha.4 + beta.4)
cat("Media", media.4)
```

```
## Media 0.1612734
```

```
dt.4 <- sqrt(media.4 * (1 - media.4) / (alpha.4 + beta.4 + 1))
cat("Desviación típica", dt.4)
```

```
## Desviación típica 0.00116302
```

### ### Intervalo de alta probabilidad 95 % sobre la media

```
result <- data.frame(low=qbeta(0.025, alpha.4, beta.4),
                     media=media.4,
                     upp=qbeta(0.975, alpha.4, beta.4))
result
```

```
##           low      media      upp
## 1 0.1590003 0.1612734 0.1635593
```

Así obtenemos que  $E(\theta_{4A} \mid \text{datos}) = 0.1612734$ , y  $DT(\theta_{4A} \mid \text{datos}) = 0.00116302$  también sabemos que  $P(0.1590003 < E(\theta_{4A} \mid \text{datos}) = 0.1612734 < 0.1635593) = 0.95$ .

## Comparación de los resultados y conclusiones

También se está interesado en inferir sobre la proporción de las proporciones para los distintos atractores, es decir estamos interesados en  $(\frac{\theta_{2A}}{\theta_{3A}} \mid \text{datos})$  para esto podemos simular una muestra de  $(\theta_{2A} \mid \text{datos})$  y de  $(\theta_{3A} \mid \text{datos})$  dividir las y así obtener los estimadores oportunos de dicha distribución simulada. Esto se realizará para las proporciones  $(\frac{\theta_{2A}}{\theta_{3A}} \mid \text{datos})$ ,  $(\frac{\theta_{2A}}{\theta_{4A}} \mid \text{datos})$  y  $(\frac{\theta_{3A}}{\theta_{4A}} \mid \text{datos})$ .

```
### inferencia sobre proporción 2 / 3 atractores
```

```
N <- 1000000  
prop.2 <- rbeta(N, alpha.2, beta.2)  
prop.3 <- rbeta(N, alpha.3, beta.3)  
result.2.3 <- prop.2/prop.3  
mean(result.2.3)
```

```
## [1] 0.7769616
```

```
sd(result.2.3)
```

```
## [1] 0.008320221
```

```
### intervalo de alta probabilidad 95 % sobre la media
```

```
confidence <- data.frame(low = quantile(result.2.3, c(0.025)),  
                          media = mean(result.2.3),  
                          upp = quantile(result.2.3, c(0.975)))  
confidence
```

```
##           low      media      upp  
## 2.5% 0.7607456 0.7769616 0.793392
```

```
### inferencia sobre proporción 2 / 4 atractores
```

```
N <- 1000000  
prop.2 <- rbeta(N, alpha.2, beta.2)  
prop.4 <- rbeta(N, alpha.4, beta.4)  
result.2.4 <- prop.2/prop.4  
mean(result.2.4)
```

```
## [1] 0.8174351
```

```
sd(result.2.4)
```

```
## [1] 0.008866146
```

```
### intervalo de alta probabilidad 95 % sobre la media
```

```
confidence <- data.frame(low = quantile(result.2.4, c(0.025)),  
                          media = mean(result.2.4),  
                          upp = quantile(result.2.4, c(0.975)))  
confidence
```

```
##           low      media      upp  
## 2.5% 0.8001889 0.8174351 0.8349512
```

```
### inferencia sobre proporción 3 / 4 atractores
```

```
N <- 1000000  
prop.3 <- rbeta(N, alpha.3, beta.3)  
prop.4 <- rbeta(N, alpha.4, beta.4)  
result.3.4 <- prop.3/prop.4  
mean(result.3.4)
```

```
## [1] 1.052146
```

```
sd(result.3.4)
```

```
## [1] 0.01058106
```

```
### intervalo de alta probabilidad 95 % sobre la media

confidence <- data.frame(low = quantile(result.3.4, c(0.025)),
                        media = mean(result.3.4),
                        upp = quantile(result.3.4, c(0.975)))

confidence

##           low      media      upp
## 2.5% 1.031542 1.052146 1.07309
```

A la luz de los resultados podemos concluir que los SFI con 3 atractores producen una mayor proporción de representaciones visualmente relevantes, seguidos de los SFI con 4 atractores y de los SFI con 2 atractores, con medias 0.1696733, 0.1612734 y 0.1318237 respectivamente. También se puede concluir que de media por cada SFI generado con 3 atractores visualmente relevantes, se generaran 0.77 SFI con 2 atractores visualmente relevantes; por cada SFI con 4 atractores visualmente relevantes, se generaran 0.82 SFI con 2 atractores visualmente relevantes y por cada SFI con 4 atractores visualmente relevantes, se generaran 1.05 SFI con 3 atractores visualmente relevantes.

## Bibliografía

1. Sprott, J. C. (1994). Automatic generation of iterated function systems. Computers & graphics, 18(3), 417-425.