

*Página www*

*Página de Abertura*

*Contenido*



*Página 1 de 56*

*Regresar*

*Full Screen*

*Cerrar*

*Abandonar*

# Estadística Bayesiana Clase 13

Juan Carlos Correa

15 de abril de 2021

## Familia Conjugada para la Distribución Multinomial

**Distribución Dirichlet** El vector aleatorio  $X = (X_1, \dots, X_k)'$  se distribuye como una Dirichlet con vector de parámetros  $\alpha = (\alpha_1, \dots, \alpha_k)'$  con  $\alpha_i > 0; i = 1, \dots, k$ , si la p.d.f.  $f(x|\alpha)$  para  $x = (x_1, \dots, x_k)$  y  $\sum_{i=1}^k x_i = 1$  está dada por:

$$f(x|\alpha) = \frac{\Gamma(\alpha_1 + \dots + \alpha_k)}{\Gamma(\alpha_1) \dots \Gamma(\alpha_k)} x_1^{\alpha_1-1} \dots x_k^{\alpha_k-1}$$

- La media de  $X_i$  es

$$E(X_i) = \frac{\alpha_i}{\alpha_0}$$

donde  $\alpha_0 = \sum_{i=1}^k \alpha_i$ .

- La varianza de  $X_i$  es

$$var(X_i) = \frac{\alpha_i(\alpha_0 - \alpha_i)}{\alpha_0^2(\alpha_0 + 1)}$$

- y la covarianza entre  $X_i$  y  $X_j$  es, ( $i \neq j$ ),

$$Cov(X_i, X_j) = -\frac{\alpha_i \alpha_j}{\alpha_0^2(\alpha_0 + 1)}$$

## Resultado:

- Suponga que  $Y = (Y_1, \dots, Y_k)'$  tiene una distribución multinomial con parámetros  $n$  (fijo) y  $\pi = (\pi_1, \dots, \pi_k)'$ , desconocidos con  $0 \leq Y_i \leq n$ , para  $i = 1, 2, \dots, k$ , y  $\sum_i Y_i = n$ .
- Suponga también que la distribución apriori de  $\pi$  es una Dirichlet con vector de parámetros  $\alpha = (\alpha_1, \dots, \alpha_k)'$  con  $\alpha_i > 0$ ;  $i = 1, \dots, k$ .
- Entonces la distribución posterior de  $\pi$  cuando  $Y_i = y_i$ ,  $i = 1, \dots, k$ , es una distribución Dirichlet con vector de parámetros  $\alpha^* = (\alpha_1 + y_1, \dots, \alpha_k + y_k)'$ .

## Ejemplo

La siguiente tabla presenta los datos sobre el tipo de sangre en una muestra de personas de la región central y oriental de Antioquia

	Tipo de Sangre			
	O	A	AB	B
Frecuencia	474	246	11	59

Si no tenemos un conocimiento apriori sobre las diversas proporciones, digamos  $\pi_O$ ,  $\pi_A$ ,  $\pi_{AB}$  y  $\pi_{AB}$ , entonces podemos escoger como apriori una  $Dirichlet(1, 1, 1, 1)$ . Entonces la aposteriori será  $Dirichlet(474 + 1, 246 + 1, 11 + 1, 59 + 1)$ .

```
library(bayesm)
res<-apply(matrix(c(475,247,12,60),nrow=1000,ncol=4,byrow=T),1,
rdirichlet)
```

```
par(mfrow=c(2,2))
plot(density(res[1,]),main='Tipo 0',xlim=c(0,1))
plot(density(res[2,]),main='Tipo A',xlim=c(0,1))
plot(density(res[3,]),main='Tipo AB',xlim=c(0,1))
plot(density(res[4,]),main='Tipo B',xlim=c(0,1))
```

```
plot(density(res[1,]),main='Tipo 0')
plot(density(res[2,]),main='Tipo A')
plot(density(res[3,]),main='Tipo AB')
plot(density(res[4,]),main='Tipo B')
```

[Página www](#)

[Página de Abertura](#)

[Contenido](#)



[Página 7 de 56](#)

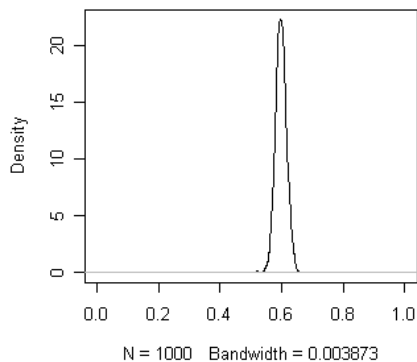
[Regresar](#)

[Full Screen](#)

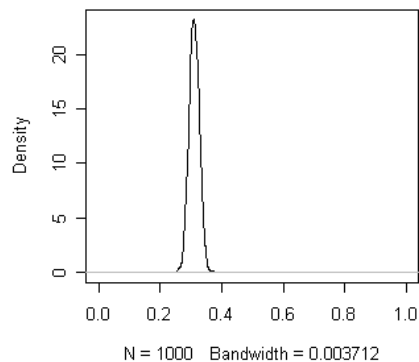
[Cerrar](#)

[Abandonar](#)

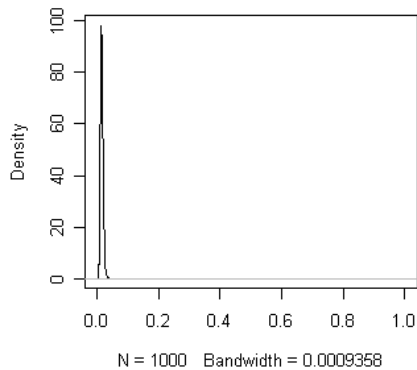
**Tipo O**



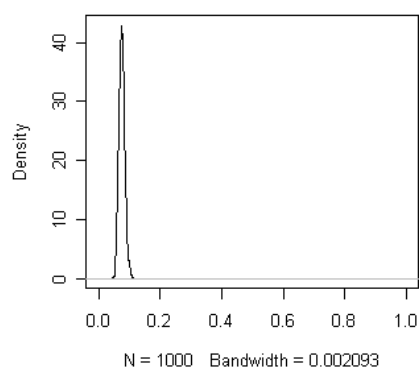
**Tipo A**



**Tipo AB**



**Tipo B**



[Página www](#)

[Página de Abertura](#)

[Contenido](#)



[Página 8 de 56](#)

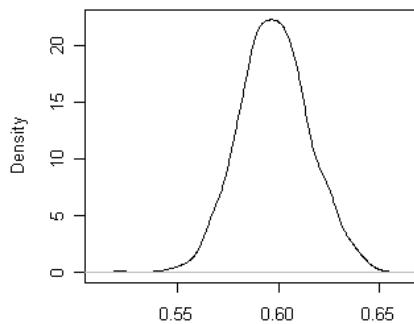
[Regresar](#)

[Full Screen](#)

[Cerrar](#)

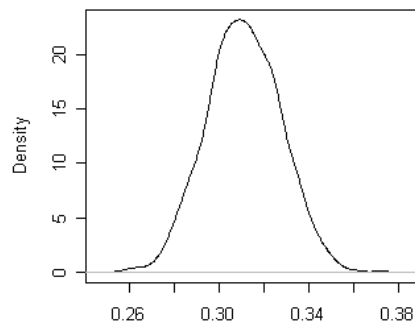
[Abandonar](#)

**Tipo O**



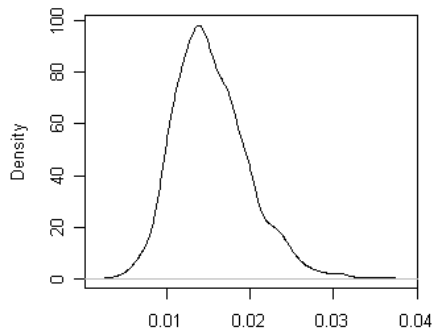
N = 1000 Bandwidth = 0.003873

**Tipo A**



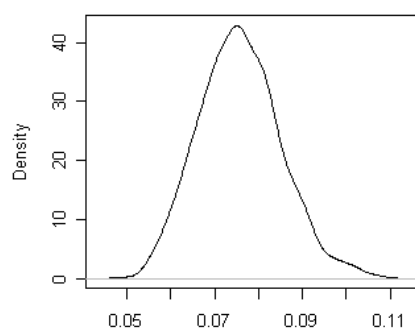
N = 1000 Bandwidth = 0.003712

**Tipo AB**



N = 1000 Bandwidth = 0.0009358

**Tipo B**



N = 1000 Bandwidth = 0.002093



Página www

Página de Abertura

Contenido



Página 9 de 56

Regresar

Full Screen

Cerrar

Abandonar

```
> library(hdrcde)
This is hdrcde 2.07
> hdr(res[1,])
$hdr
      [,1]      [,2]
99% 0.5556055 0.6435400
95% 0.5646887 0.6321565
50% 0.5856615 0.6083180
```

```
$mode
[1] 0.5965341
```

```
$falpha
      1%      5%      50%
1.104883 4.021291 18.497656
```

Página www

Página de Abertura

Contenido



Página 10 de 56

Regresar

Full Screen

Cerrar

Abandonar

```
> hdr(res[2,])
```

```
$hdr
```

```
[,1]
```

```
[,2]
```

```
99% 0.2711979 0.3517752
```

```
95% 0.2794612 0.3421650
```

```
50% 0.2991745 0.3219154
```

```
$mode
```

```
[1] 0.3091064
```

```
$falpha
```

```
1%
```

```
5%
```

```
50%
```

```
1.241335
```

```
4.484015
```

```
19.068098
```

Página www

Página de Abertura

Contenido



Página 11 de 56

Regresar

Full Screen

Cerrar

Abandonar

```
> hdr(res[3,])
```

```
$hdr
```

```
      [,1]
```

```
      [,2]
```

```
99% 0.005115702 0.02841837
```

```
95% 0.007569295 0.02454543
```

```
50% 0.011359559 0.01696535
```

```
$mode
```

```
[1] 0.01387235
```

```
$falpha
```

```
      1%
```

```
      5%
```

```
     50%
```

```
2.697576 13.571135 75.217231
```

```
> hdr(res[4,])
```

```
$hdr
```

```
          [,1]      [,2]  
99% 0.05381815 0.10095793  
95% 0.05772293 0.09302976  
50% 0.06889325 0.08154789
```

```
$mode
```

```
[1] 0.07507367
```

```
$falpha
```

```
          1%          5%          50%  
2.343003  7.706535 33.528008
```

```
> par(mfrow=c(1,1))
```

```
> hdr.boxplot.2d(res[1,],res[2,])
```

```
Loading required package: ash
```

```
> title(xlab='Tipo 0',ylab='Tipo A',  
main='Dist. Conjunta Bivariable')
```

[Página www](#)

[Página de Abertura](#)

[Contenido](#)



[Página 13 de 56](#)

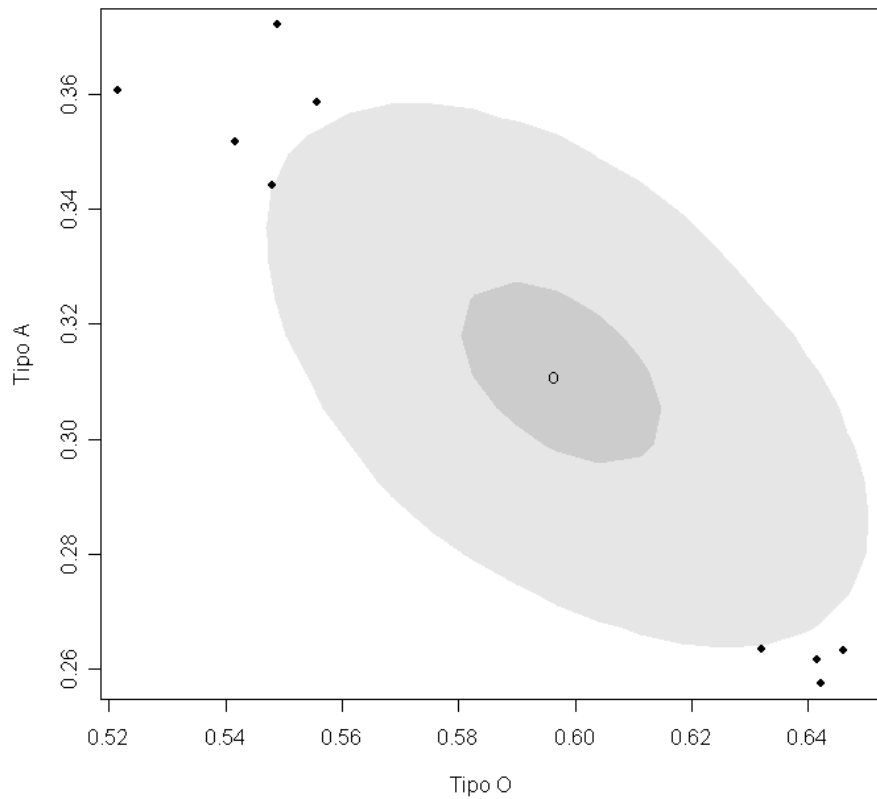
[Regresar](#)

[Full Screen](#)

[Cerrar](#)

[Abandonar](#)

**Dist. Conjunta Bivariable**



# Intervalos multinomiales clásicos

	Tipo de Sangre			
	O	A	AB	B
Frecuencia	474	246	11	59
$\hat{\pi}_i$	0.60000000	0.31139241	0.01392405	0.07468354
Intervalos simultáneos: TCL				
	0.556465534	0.270242626	0.003511297	0.051322879
	0.64353447	0.35254218	0.02433680	0.09804421
Intervalos simultáneos: Quesenberry y Hurst				
	0.53763946	0.25675391	0.00505438	0.04795771
	0.65927993	0.37184116	0.03776781	0.11451171
Intervalos simultáneos Bootstrap				
0.625 %	0.5509415	0.2746835	0.005063291	0.05094146
99.375 %	0.6414636	0.3518987	0.026582278	0.09842563
Intervalos simultáneos: Sison y Glaz				
	0.5658	0.2772	0.0000	0.0405
	0.6363	0.3477	0.0502	0.1109
Intervalos simultáneos: Sison y Glaz II				
	0.5646	0.2759	-0.0215	0.0392
	0.6354	0.3465	0.0494	0.1101

Página www

Página de Abertura

Contenido

◀

▶

◀

▶

Página 14 de 56

Regresar

Full Screen

Cerrar

Abandonar

*Página www*

*Página de Abertura*

*Contenido*



*Página 15 de 56*

*Regresar*

*Full Screen*

*Cerrar*

*Abandonar*

Cómo elicitamos los parámetros apriori de la distribución Dirichlet?

## Ejemplo usando muestras hipotéticas mentales

Recuerde la época en que no había pandemia...Cómo venían los estudiantes de pregrado a la universidad a una clase de 6 de la mañana en El Volador. Suponga que ud. se imagina una muestra mental de 1000 estudiantes:

Medio	A pie	en bus	en bicicleta	en taxi	en moto	en carro particular
Nro.	50	600	70	50	200	30

Nivel de seguridad: 30



```
Elicita.Apriori.Multinomial<-function(Valores.hipotéticos,
  Muestra.equivalente,Nsim=1000){
  Valores.hipotéticos<-ifelse(Valores.hipotéticos==0,0.5,
  Valores.hipotéticos)
  prob.estima<-Valores.hipotéticos/sum(Valores.hipotéticos)

  muestras<-rmultinom(Nsim,Muestra.equivalente,prob=prob.estima)

  aux1<-function(xx){
    xx<-ifelse(xx==0,0.5,xx)
    propor<-xx/sum(xx)
    return(propor)
  } # Fin aux1

  resu<-apply(muestras,2,aux1)

  library(sirt)
  alfas<-dirichlet.mle(t(resu))$alpha
  return(alfas)
}# Fin Elicita.Apriori.Multinomial
```

Página [www](#)

Página de Abertura

Contenido



Página **18** de **56**

Regresar

Full Screen

Cerrar

Abandonar

## Parámetros de la Dirichlet elicitados...

```
> Elicita.Apriori.Multinomial(c(50,600,70,50,200,30),30)
[1] 2.180651 22.957719 2.760708 2.153159 7.588523 1.633315
```

## Volviendo al tema de Simulación...

### Ejemplo para la proporción

Suponga que estamos interesados en determinar la proporción de estudiantes que sufren gastritis.

- Como apriori supongamos que una normal truncada con parámetros  $\mu = 0,5$  y  $\sigma^2 = 0,2^2$ .
- Se saca una muestra al azar de 10 estudiantes y se les evalúa. De éstos solo dos tienen gastritis.

Página www

Página de Abertura

Contenido



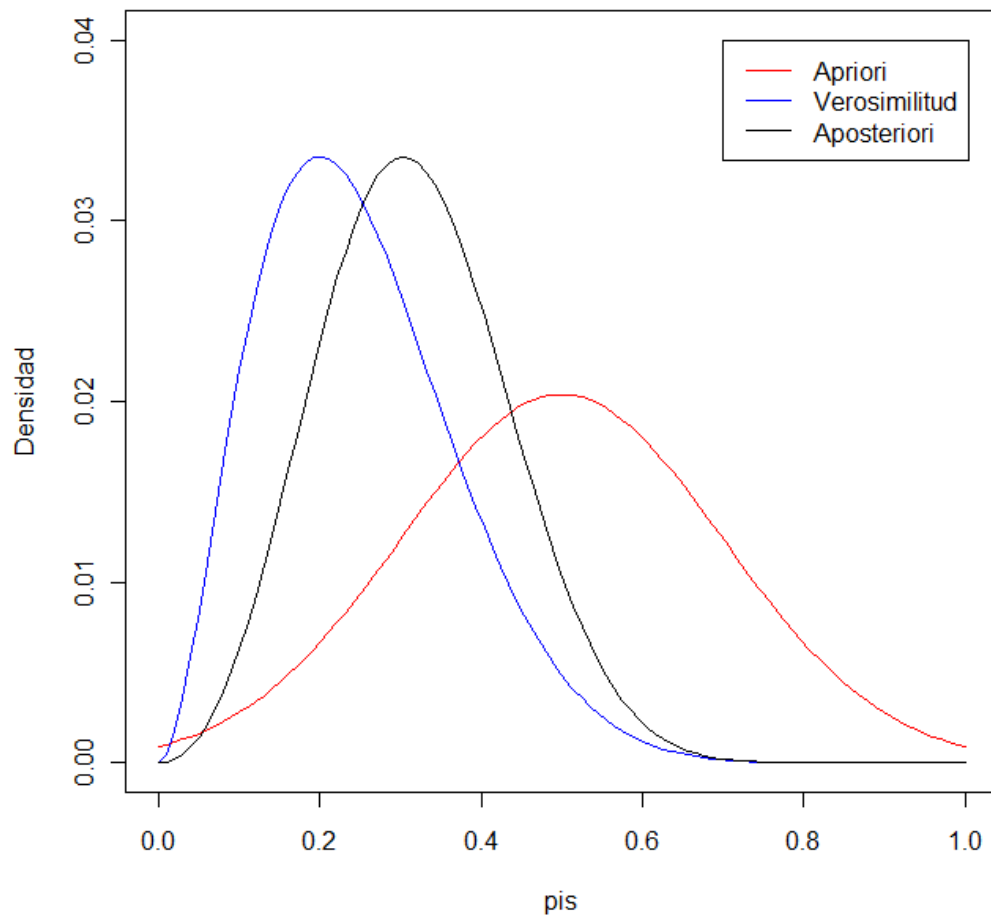
Página 20 de 56

Regresar

Full Screen

Cerrar

Abandonar



```
#Generación de muestra de una distribución aposteriori por medio  
#del método de discretización
```

```
#
```

```
densidad.posteriori <- function(x, media.apriori,  
    dt.apriori, n, nro.exitos){
```

```
#n = tamaño de muestra
```

```
#vero = verosimilitud
```

```
vero <- x^nro.exitos*(1-x)^(n - nro.exitos)
```

```
apriori <- exp(-(x - media.apriori)^2/(2*dt.apriori^2))
```

```
aposteriori <- vero*apriori
```

```
list(vero = vero, apriori = apriori, aposteriori = aposteriori)  
}
```

```
#Graficos de la verosimilitud, distribuciones apriori, aposteriori
```

```
pis <- seq(0.00001, 0.9999, length = 100)
```

```
res <- densidad.posteriori(pis, 0.5, 0.2, 10, 2)
```

```
res.apriori<-res$apriori/sum(res$apriori)
```

```
plot(pis, res.apriori, type = 'l', col='red', ylab='',ylim=c(0,0.04))
```

```
title(ylab='Densidad')
```

```
res.vero<-res$vero/sum(res$vero)
```

```
points(pis, res.vero, type = 'l', col='blue')
```

```
res.aposteriori<-res$aposteriori/sum(res$aposteriori)
```

```
points(pis, res.aposteriori, type = 'l', col='black')
```

```
legend(0.7,0.04,c('Apriori','Verosimilitud','Aposteriori'),  
lty=1,col=c('red','blue','black'))
```

Página www

Página de Abertura

Contenido

◀

▶

◀

▶

Página 21 de 56

Regresar

Full Screen

Cerrar

Abandonar

*Página www*

*Página de Abertura*

*Contenido*



*Página 22 de 56*

*Regresar*

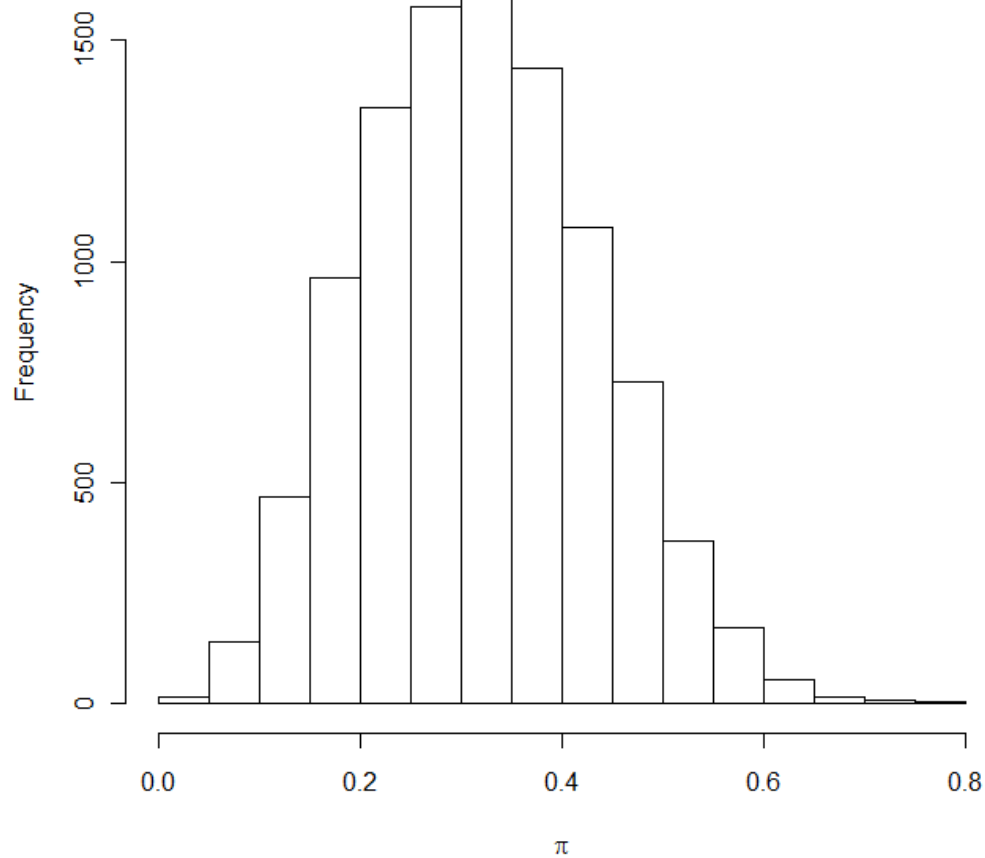
*Full Screen*

*Cerrar*

*Abandonar*

```
resu<-sample(pis,10000,prob=res.aposteriori,replace=T)
hist(resu,main='Distribución Simulada',xlab=expression(pi))
```

## Distribución Simulada



*Página www*

*Página de Abertura*

*Contenido*

◀◀

▶▶

◀

▶

*Página 23 de 56*

*Regresar*

*Full Screen*

*Cerrar*

*Abandonar*

```
> mean(resu)
[1] 0.3171478
> median(resu)
[1] 0.3131069
> quantile(resu, probs=c(0.025, 0.975))
      2.5%      97.5%
0.1111089 0.5555044

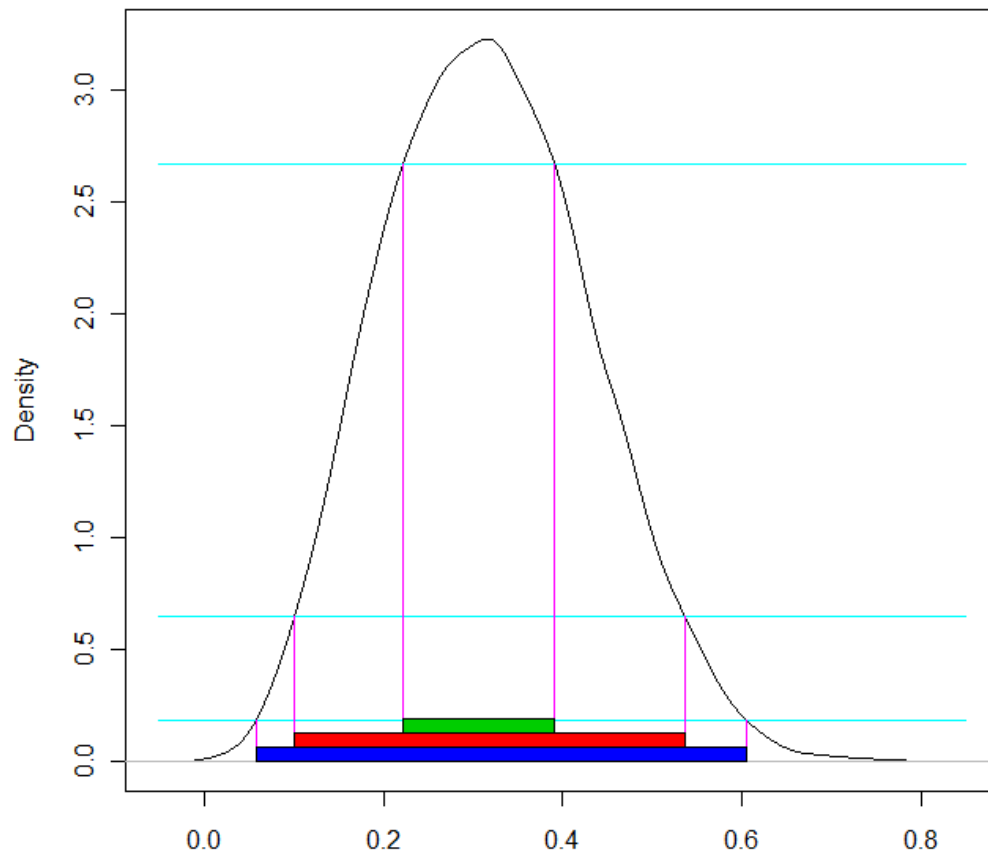
> require(hdrcde)
> hdr(resu)
$hdr
      [,1]      [,2]
99% 0.05800344 0.6060039
95% 0.10100899 0.5371268
50% 0.22220778 0.3912115

$mode
[1] 0.317054

$falpha
      1%      5%      50%
0.1858224 0.6524188 2.6730446
```



**density.default(x = x, bw = h)**



[Página www](#)

[Página de Abertura](#)

[Contenido](#)

◀◀

▶▶

◀

▶

[Página 25 de 56](#)

[Regresar](#)

[Full Screen](#)

[Cerrar](#)

[Abandonar](#)