Índice

Contents

1. Introducción

El objetivo de este trabajo es crear un Estadístico para estimar la media poblacional de una distribución continua determinada (que no sea la Distribución Normal ni la Exponencial) y comparar su rendimiento con la media muestral (que es el estadístico más usado para aproximar la media poblacional). En nuestro caso hemos optado por la distribución Beta (con alpha y beta = 2) debido a que su función de densidad es simétrica y centrada, y nos parece que es sencillo encontrar un Estadístico fiable para aproximar su media poblacional.

2. Distribución Beta

La Distribución Beta es una distribución continua que depende de dos parámetros (alpha y beta) y que toma valores en el intervalo [0,1]. Debido a que solo está definida en [0,1] es una distribución muy usada para modelizar la probabilidad de que ocurra un evento, aunque también es usada para describir datos empíricos (debido a la variedad de formas que puede adoptar en función de los valores que tomen sus parámetros) y para modelar la fiabilidad de un sistema. Su función de densidad es distinta de cero solo cuando $0 < x < 1$ y es la siguiente:

$$f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\,\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1}$$

También se puede escribir como:

$$f(x) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1}$$

Dónde la función Beta es la siguiente:

$$B(\alpha, \beta) = \frac{\Gamma(\alpha)\,\Gamma(\beta)}{\Gamma(\alpha+\beta)} = \int_0^1 x^{\alpha-1} (1-x)^{\beta-1} dx$$

Propiedades de la función Beta

La función beta cumple las siguientes propiedades (las usaremos para encontrar su función generadora de momentos, etc):

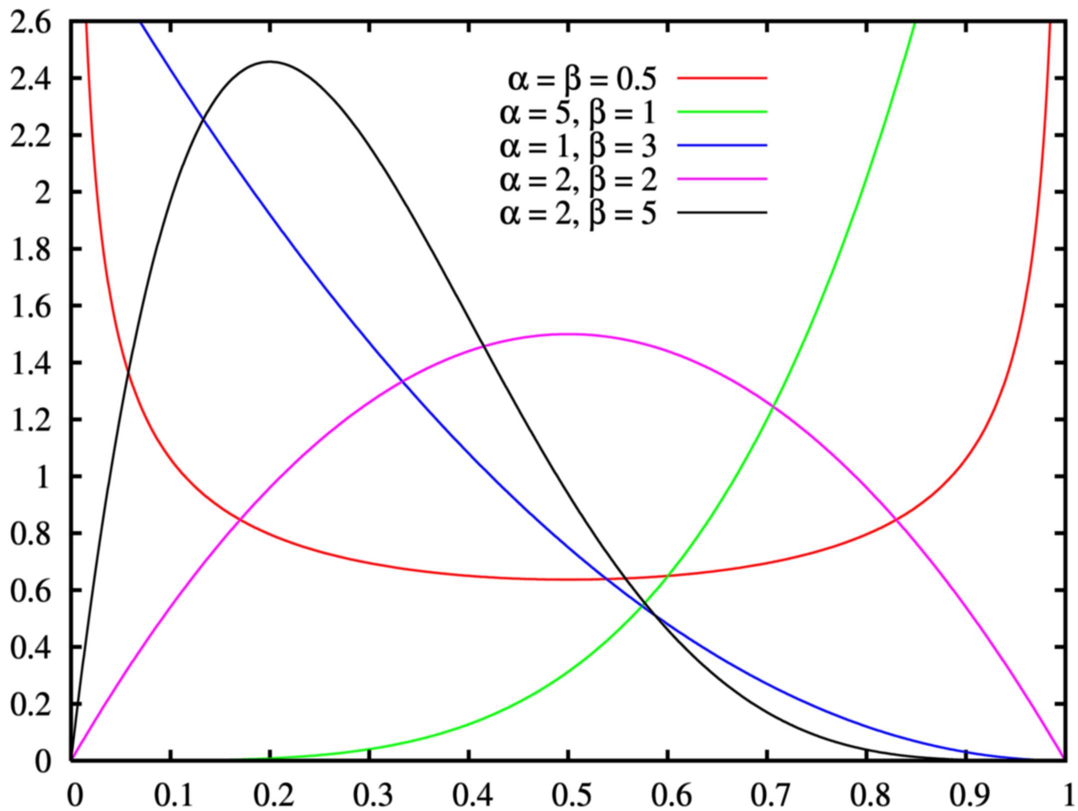
$$B(\alpha, \beta) = B(\beta, \alpha)$$

$$B(\alpha, 1) = \frac{1}{\alpha}$$

$$B(\alpha + 1, \beta) = \frac{\alpha}{(\alpha + \beta)} B(\alpha, \beta)$$

2.0 Forma de la distribución Beta

Esta distribución adopta formas muy diversas, en función de los valores de sus parámetros, por ello se utiliza mucho para modelar datos de manera empírica (es decir, observando la forma de la distribución de los datos, se intenta aproximar por la distribución beta con la forma más parecida)



2.1 Momentos y Función Generadora de Momentos

$$E[X] = \frac{\alpha}{(\alpha+\beta)}$$

$$Var[X] = \frac{\alpha\,\beta}{(\alpha+\beta)^2\,(\alpha+\beta+1)}$$

$$M_x(t) = \sum_{k=0}^{\infty} \frac{B(\alpha+k,\,\beta)}{B(\alpha,\,\beta)}\,\frac{t^k}{k!}$$

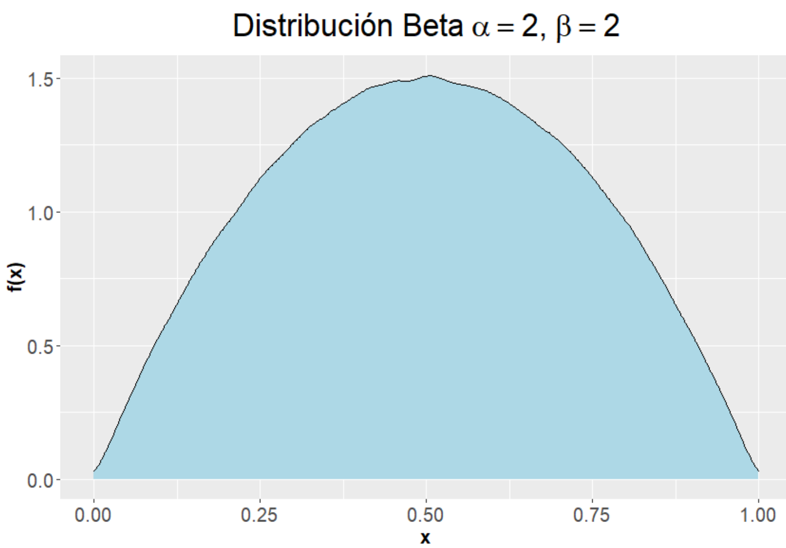
[1]

[1] Nota: Los cálculos para encontrar los momentos y la función generadora están en el Anexo I y II (respectivamente)

3. Creación del Estadístico para estimar la Media Poblacional

3.1 Escogiendo los parámetros de la distribución

En nuestro caso hemos optado por seleccionar una distribución beta con alpha = 2, beta = 2, ya que pensamos que debido a su forma simétrica, seremos capaces de aproximar la media poblacional de manera más precisa que si fuera asimétrica. La forma de una distribución Beta(2,2) es la siguiente:



Los momentos de la distribución Beta(alpha = 2, beta = 2) son los siguientes:

$$E[X] = \mu = 0.5$$
$$Var[X] = \sigma^2 = 0.05$$

3.2 Fórmula del Estadístico

Nuestro estadístico es el promedio de los percentiles 40 y 60. Hemos escogido este estadístico debido a que la distribución por la que hemos optado (distribución beta(2,2)), es bastante simétrica, y nos parece que podemos llegar a estimar la media poblacional con bastante precisión. La fórmula de nuestro Estadístico es:

$$\text{Estadístico} = \frac{P_{40} + P_{60}}{2}$$

3.3 Código para obtener el estadístico

En primer lugar hemos creado la función *muestreo*. Esta función es la base de toda la práctica. Para crearla primeramente ejecutamos unas cuantas librerías de R.

```
library(tidyverse)
library(e1071)
library(moments)
library(gt)
library(ggpubr)
library(ggplotify)
library(grid)
```

A continuación creamos la función *muestreo*, que depende de 2 parámetros (alpha y beta, los mismos que determinan la distribución beta). Entonces creamos una lista vacía, una matriz de 40 columnas y 10 filas (donde guardaremos los datos de la muestra) y guardamos en la lista una matriz igual que la que hemos creado anteriormente. A continuación calculamos la media y varianza poblacionales, mediante la fórmula que hemos visto previamente.

Entonces mediante un bucle con j iterando por columnas, establecemos un set.seed j-ésimo, que nos crea j muestras aleatorias y reproducibles (para obtener siempre el mismo resultado), pero distintas entre sí, de una distribución beta, con parámetros alpha y beta introducidos por el usuario. A continuación cambiamos el nombre de las columnas, sustituyendo las V generadas por espacios vacíos. Utilizamos el sexto elemento de la lista para calcular nuestro estadístico de las 40 muestras, guardando los datos en las columnas de la matriz previamente creada.

Finalmente ordenamos la lista, creamos dataframes para resumir la información y hacemos que la función imprima dicha lista al ser llamada.

```
muestreo <- function(alpha, beta){
        lista <- list()
        tabla_muestras <- matrix(ncol = 40, nrow = 10) %>% as.data.frame()
        lista[[6]] <- matrix(ncol = 40, nrow = 1) %>% as.data.frame()
        media_poblacional <- alpha / (alpha + beta)
        varianza_poblacional <- alpha * beta / (((alpha + beta) ^ 2 ) * (alpha + beta + 1))
        for(j in 1:40){
                set.seed(j)
                tabla_muestras[, j] <- rbeta(10, shape1 = alpha, shape2 = beta)
                colnames(tabla_muestras)[j] <- as.numeric(gsub("V", "",
                                                           colnames(tabla_muestras)[j]))
                colnames(lista[[6]])[j] <-  as.numeric(gsub("V", "",
                                                        colnames(tabla_muestras)[j]))
                lista[[6]][, j] <- 0.5*(quantile(tabla_muestras[, j],
                                                 probs = c(0.6)) +
                                                quantile(tabla_muestras[, j],
                                                         probs = c(0.4)))
        }
        lista[[1]] <- tabla_muestras
        lista[[2]] <- media_poblacional
        lista[[3]] <- varianza_poblacional
        lista[[4]] <- data.frame(Muestra = names(colMeans(tabla_muestras)),
                                 mediamuestral = unname(colMeans(tabla_muestras)))
        lista[[5]] <- data.frame(Estimador = "Media Muestral",
                                 Media = mean(lista[[4]][, 2]),
```

```
                                  Mediana = median(lista[[4]][, 2]),
                                  SD = sd(lista[[4]][, 2]),
                                  IQR = IQR(lista[[4]][, 2]),
                                  MAD = mad(lista[[4]][, 2]),
                                  Curtosis = moments::kurtosis(lista[[4]][, 2]),
                                  Asimetría = moments::skewness(lista[[4]][, 2]))
        lista[[6]] <- gather(as.data.frame(lista[[6]]))
        colnames(lista[[6]]) <- c("Muestra", "Estadístico")
        lista[[7]] <- data.frame(Estimador = "Estadístico",
                                  Media = mean(lista[[6]][, 2]),
                                  Mediana = median(lista[[6]][, 2]),
                                  SD = sd(lista[[6]][, 2]),
                                  IQR = IQR(lista[[6]][, 2]),
                                  MAD = mad(lista[[6]][, 2]),
                                  Curtosis = moments::kurtosis(lista[[6]][, 2]),
                                  Asimetría = moments::skewness(lista[[6]][, 2]))
        lista[[8]] <- rbind(lista[[5]], lista[[7]])
        names(lista) <- c("tabla_muestras", "media_poblacional",
                           "varianza_poblacional", "media_muestral",
                           "medidas_media_muestral", "estadistico",
                           "medidas_estadistico", "tabla_comparacion")
        lista
}

resultados <- muestreo(alpha = 2, beta = 2)
datosgrafico <- inner_join(resultados$media_muestral , resultados$estadistico)
```

Al llamar la función, para alpha = 2 y beta = 2, este es el output generado por la función:

```
## $tabla_muestras
##            1         2         3         4         5         6         7
## 1  0.3275025 0.2594031 0.2439641 0.5609713 0.2730286 0.5757192 0.9598926
## 2  0.5516990 0.5520370 0.4179083 0.3497046 0.8452087 0.3265782 0.1918357
## 3  0.2743131 0.2059597 0.5727064 0.7392076 0.1797754 0.7337646 0.3098952
## 4  0.8814780 0.4773702 0.2012759 0.6644891 0.5197810 0.9011515 0.3848948
## 5  0.5923401 0.5373154 0.5551010 0.8877412 0.8989244 0.5068230 0.2419199
## 6  0.2780523 0.6936117 0.5084973 0.6888135 0.3336852 0.6029646 0.2474083
## 7  0.6355158 0.4326179 0.5240852 0.4400386 0.3686087 0.1691099 0.7038462
## 8  0.7013578 0.9329084 0.7910491 0.9230177 0.3251702 0.7014333 0.4670349
## 9  0.6591489 0.4608946 0.1872513 0.9079563 0.4197840 0.5126569 0.5430031
## 10 0.4143388 0.6165691 0.8226134 0.5044343 0.5389146 0.2240809 0.9523528
##            8         9        10        11         12        13        14
## 1  0.4761491 0.2914443 0.5052881 0.3368171 0.13752070 0.6534502 0.3182762
## 2  0.7268596 0.2790446 0.4481302 0.5075018 0.87859495 0.4213134 0.8999906
## 3  0.3709604 0.4601221 0.1572222 0.1588532 0.24518131 0.9077289 0.9464376
## 4  0.3474798 0.4220550 0.3346706 0.8043260 0.25387414 0.5527297 0.8653343
## 5  0.7007776 0.6216568 0.5826525 0.2505117 0.06569408 0.7966627 0.4898057
## 6  0.4695887 0.1939091 0.6089447 0.8336891 0.42353205 0.6159888 0.8154380
## 7  0.4520467 0.8071635 0.1894820 0.6720953 0.41157840 0.8149384 0.4817021
## 8  0.2151617 0.4948687 0.3982324 0.4871034 0.32703090 0.5665394 0.7805407
## 9  0.3362513 0.4302777 0.1136290 0.2341673 0.46998771 0.3977626 0.3945773
## 10 0.2932087 0.3984358 0.4279373 0.2760854 0.61939740 0.7885415 0.7747419
##           15        16        17        18        19        20        21
```

```
## 1  0.5727160 0.6325402 0.4775435 0.7476623 0.1933674 0.8009767 0.7151268
## 2  0.9150168 0.4646646 0.4344918 0.9139619 0.6086120 0.3381664 0.6448813
## 3  0.4048658 0.7865789 0.2788441 0.1161229 0.4035642 0.1645773 0.1765924
## 4  0.7406651 0.1439022 0.7098872 0.4199685 0.3482631 0.3755519 0.6207918
## 5  0.6356743 0.7978021 0.4533594 0.4041881 0.7604071 0.6575114 0.1225262
## 6  0.1797968 0.3696250 0.7585937 0.6024558 0.4334700 0.2661483 0.2503329
## 7  0.5064282 0.2337464 0.5717167 0.1656471 0.7192898 0.3714429 0.5179068
## 8  0.7853780 0.5179263 0.6025672 0.9681897 0.2970048 0.3462266 0.4993249
## 9  0.4627684 0.7706083 0.8048110 0.5179975 0.4270153 0.4943201 0.7061903
## 10 0.2181198 0.6584493 0.6768709 0.8734508 0.4484579 0.4576361 0.3481273
##           22        23        24         25         26        27         28
## 1  0.3578318 0.5543809 0.3488006 0.44040537 0.05282655 0.9242646 0.07632113
## 2  0.7666846 0.3787857 0.6487177 0.22561804 0.79781640 0.7971682 0.48186731
## 3  0.5821719 0.7445176 0.6171078 0.59013352 0.36392495 0.2920147 0.16485177
## 4  0.4412096 0.7890558 0.3387739 0.13415726 0.72339917 0.1415635 0.54581469
## 5  0.9183828 0.4219212 0.7285912 0.37583283 0.38553937 0.2140268 0.64733021
## 6  0.4813792 0.7692920 0.5747214 0.90210951 0.54191599 0.5828457 0.58292281
## 7  0.4541586 0.5128160 0.6239094 0.64194163 0.53609232 0.5019425 0.50581466
## 8  0.4437568 0.8783719 0.3701441 0.52809244 0.74504295 0.7998525 0.41270952
## 9  0.5843223 0.5614002 0.2711859 0.48367253 0.48873461 0.9476016 0.91632100
## 10 0.2921718 0.2245009 0.5006522 0.09667693 0.61724791 0.5668646 0.31975718
##           29        30        31        32         33        34        35
## 1  0.1741952 0.1731754 0.5156731 0.5041300 0.46169932 0.4608656 0.7796761
## 2  0.1781729 0.4026381 0.4481342 0.7348888 0.48849255 0.8087990 0.5374450
## 3  0.5603890 0.3552855 0.8815540 0.2287144 0.76730742 0.2962594 0.4903970
## 4  0.7525137 0.8238345 0.7567156 0.6878831 0.45542274 0.3409924 0.4873140
## 5  0.1964058 0.9141786 0.8668875 0.6252281 0.05046834 0.4259582 0.3901988
## 6  0.9454147 0.1322581 0.3760530 0.6136626 0.63853974 0.3731276 0.6146331
## 7  0.3553423 0.5311507 0.6091019 0.5799263 0.29435711 0.6839967 0.3444593
## 8  0.2514830 0.2929559 0.7459332 0.3280641 0.71151738 0.2710281 0.8070043
## 9  0.6170390 0.3161905 0.1759662 0.7266769 0.70547455 0.7800516 0.5977422
## 10 0.7898540 0.5770867 0.3005525 0.5872941 0.21268819 0.4978954 0.6420111
##           36        37        38        39        40
## 1  0.5874194 0.5351593 0.4287422 0.4477621 0.6328986
## 2  0.7291713 0.6068262 0.2224589 0.1851157 0.6378761
## 3  0.6929911 0.6600658 0.6880976 0.3808240 0.2684452
## 4  0.8972759 0.4175688 0.5070964 0.3355105 0.2759308
## 5  0.1620921 0.2761061 0.1370043 0.6300803 0.4098550
## 6  0.3424336 0.4067736 0.6200500 0.6176581 0.1701727
## 7  0.5167956 0.4459141 0.5022350 0.2301492 0.1479839
## 8  0.1455703 0.8412089 0.7379041 0.3285442 0.4190877
## 9  0.5552441 0.7306689 0.3856153 0.7260246 0.1691760
## 10 0.5075411 0.5607589 0.5171887 0.7559641 0.4804136
##
## $media_poblacional
## [1] 0.5
##
## $varianza_poblacional
## [1] 0.05
##
## $media_muestral
##    Muestra mediamuestral
## 1        1     0.5315746
## 2        2     0.5168687
```

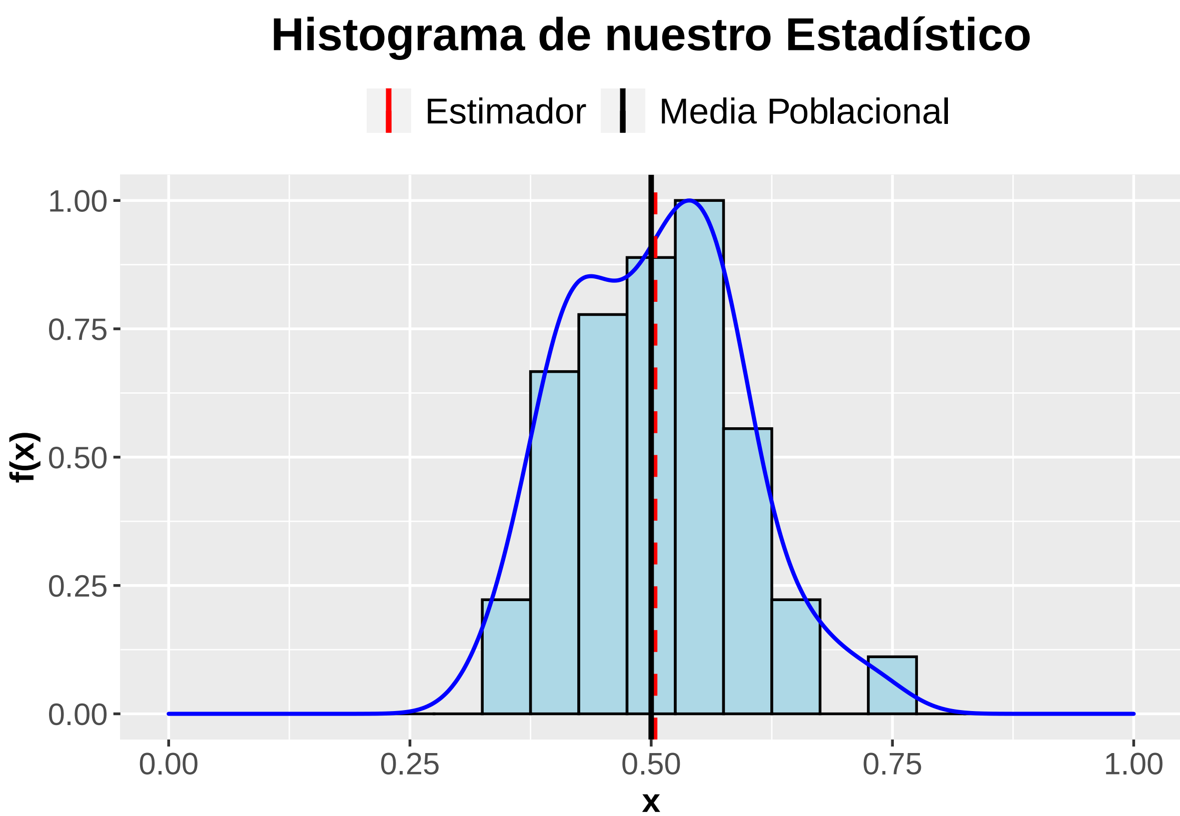
```
## 3        3     0.4824452
## 4        4     0.6666374
## 5        5     0.4702881
## 6        6     0.5254282
## 7        7     0.5002083
## 8        8     0.4388484
## 9        9     0.4398978
## 10      10     0.3766189
## 11      11     0.4561150
## 12      12     0.3832392
## 13      13     0.6515656
## 14      14     0.6766845
## 15      15     0.5421429
## 16      16     0.5375843
## 17      17     0.5768686
## 18      18     0.5729645
## 19      19     0.4639452
## 20      20     0.4272558
## 21      21     0.4601801
## 22      22     0.5322070
## 23      23     0.5835042
## 24      24     0.5022604
## 25      25     0.4418640
## 26      26     0.5252540
## 27      27     0.5768145
## 28      28     0.4653710
## 29      29     0.4820810
## 30      30     0.4518754
## 31      31     0.5676571
## 32      32     0.5616469
## 33      33     0.4785967
## 34      34     0.4938974
## 35      35     0.5690881
## 36      36     0.5136535
## 37      37     0.5481051
## 38      38     0.4746392
## 39      39     0.4637633
## 40      40     0.3611840
##
## $medidas_media_muestral
##        Estimador     Media   Mediana        SD        IQR        MAD Curtosis
## 1 Media Muestral 0.5072706 0.5012344 0.0721176 0.08862303 0.06388053 3.077518
##   Asimetría
## 1 0.3219243
##
## $estadistico
##    Muestra Estadístico
## 1        1   0.5531827
## 2        2   0.5069920
## 3        3   0.5043766
## 4        4   0.6660266
## 5        5   0.4072110
## 6        6   0.5484703
## 7        7   0.4261586
```

```
## 8        8   0.4103158
## 9        9   0.4274114
## 10      10   0.4044111
## 11      11   0.4038936
## 12      12   0.3570640
## 13      13   0.6518479
## 14      14   0.7276335
## 15      15   0.5434318
## 16      16   0.5697627
## 17      17   0.5831680
## 18      18   0.5696622
## 19      19   0.4285500
## 20      20   0.3848710
## 21      21   0.4989533
## 22      22   0.4858471
## 23      23   0.5862010
## 24      24   0.5200625
## 25      25   0.4580084
## 26      26   0.5445990
## 27      27   0.6047352
## 28      28   0.4880094
## 29      29   0.4484238
## 30      30   0.3968453
## 31      31   0.5762460
## 32      32   0.6013179
## 33      33   0.5038501
## 34      34   0.4402518
## 35      35   0.5615622
## 36      36   0.5406040
## 37      37   0.5393235
## 38      38   0.4919856
## 39      39   0.4392095
## 40      40   0.3432423
##
## $medidas_estadistico
##     Estimador    Media   Mediana         SD       IQR        MAD Curtosis
## 1 Estadístico 0.503593 0.5041133 0.08791211 0.1353218 0.09725723 2.656788
##   Asimetría
## 1 0.2855404
##
## $tabla_comparacion
##        Estimador     Media   Mediana         SD        IQR        MAD Curtosis
## 1 Media Muestral 0.5072706 0.5012344 0.07211760 0.08862303 0.06388053 3.077518
## 2    Estadístico 0.5035930 0.5041133 0.08791211 0.13532182 0.09725723 2.656788
##   Asimetría
## 1 0.3219243
## 2 0.2855404
```

4. Evaluación de nuestro Estadístico

Para comparar nuestro estadístico con la media muestral usaremos las siguientes medidas de dispersión, centralización y forma:

Estadístico						
Medidas de Centralización		***Medidas de Dispersión***			***Medidas de Forma***	
Media	Mediana	SD	IQR	MAD	Curtosis	Asimetría
0.503593	0.5041133	0.08791211	0.1353218	0.09725723	2.656788	0.2748998

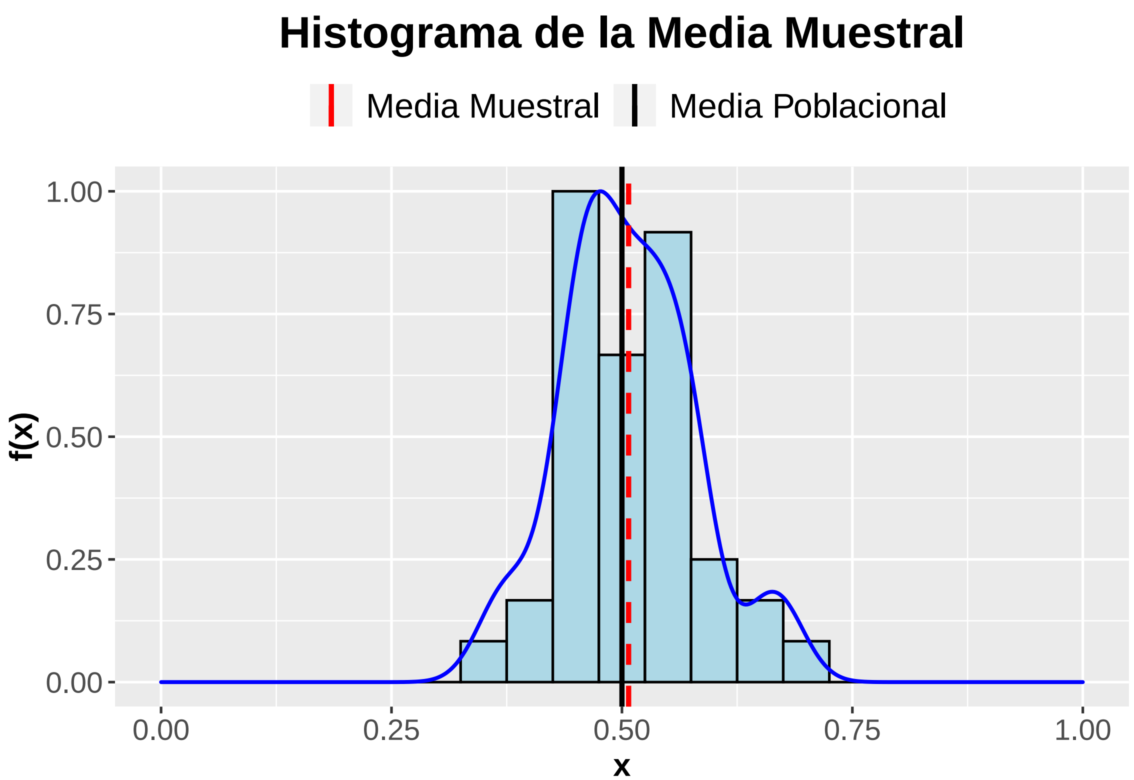


Podemos ver que en promedio nuestro estadístico se acerca bastante a la media poblacional (cuyo valor es 0.5) y por otro lado la mediana también está muy centrada, y también podría ser un buen estimador de la media poblacional. En cuanto a las medidas de forma, la curtosis nos indica que se trata de una distribución levemente platicúrtica (su valor es menor que 3), es decir menos apuntada y con colas menos gruesas que la normal y el coeficiente de asimetría nos indica una pequeña asimetría a la derecha, ya que su valor es mayor que 0. Finalmente en cuanto a la dispersión, la distancia entre el cuartil 3 y 1 (IQR) es muy pequeña, por lo que podemos concretar que la mayoría de los datos están concentrados en este rango. Además, la desviación absoluta mediana (MAD) y también la desviación típica (SD) son muy pequeñas, y eso nos indica que hay poca variabilidad en los datos, es decir, estan bastante concentrados alrededor de la media y la mediana. [2]

[2] Nota: El código para obtener el gráfico se encuentra en el Anexo III

5. Evaluación de la Media Muestral

Media Muestral						
Medidas de Centralización		Medidas de Dispersión			Medidas de Forma	
Media	Mediana	SD	IQR	MAD	Curtosis	Asimetría
0.5072706	0.5012344	0.0721176	0.08862303	0.06388053	3.077518	0.3099279



Con estos datos se confirmaría el que el valor de la esperanza de las medias muestrales es muy parecida a la media teórica. Como indica el coeficiente de curtosis, esta distribución es prácticamente mesocúrtica, es decir se parece mucho a una normal y es muy simétrica pero tiene un poco de asimetría a la derecha, por lo que también la mediana se parece mucho a la media al ser una distribución centrada. En cuanto a las medidas de dispersión, la desviación típica muestral (SD) es muy pequeña, así como el IQR y la MAD, indicando esto que los valores están muy concentrados en torno al centro, que en este caso está cerca de la media.
[3]

[3]Nota: El código para obtener el gráfico se encuentra en el Anexo IV

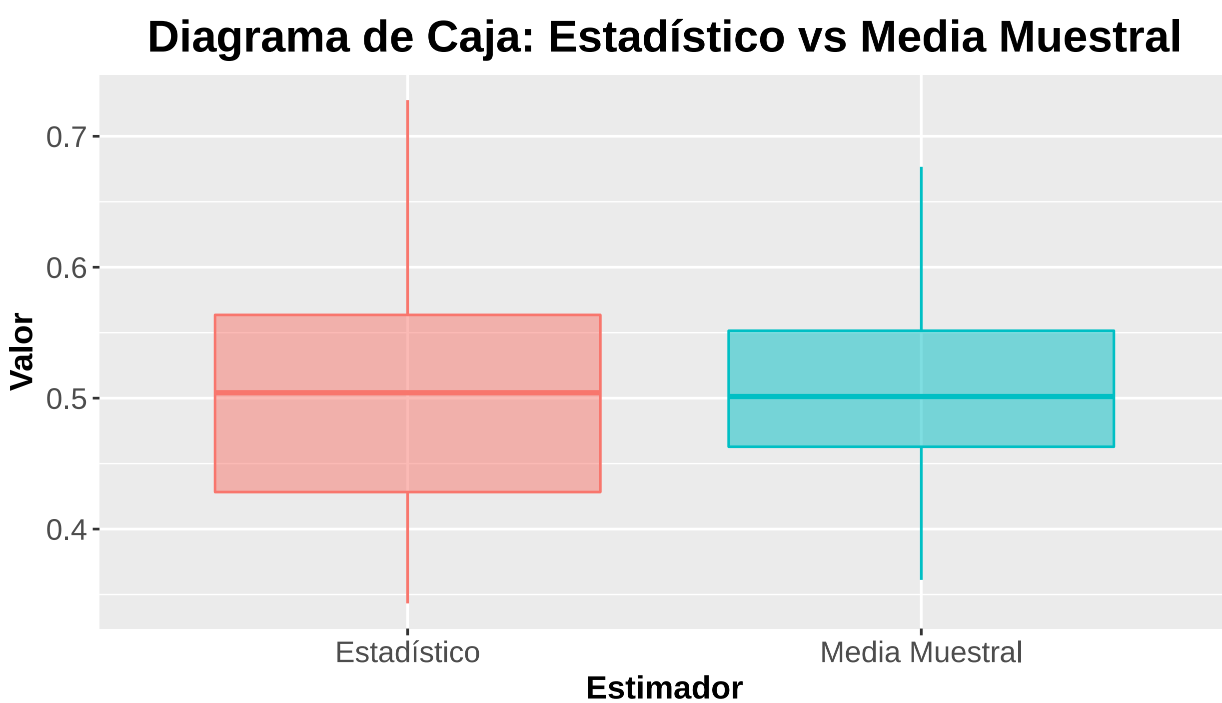
5.1 Varianza de la media muestral

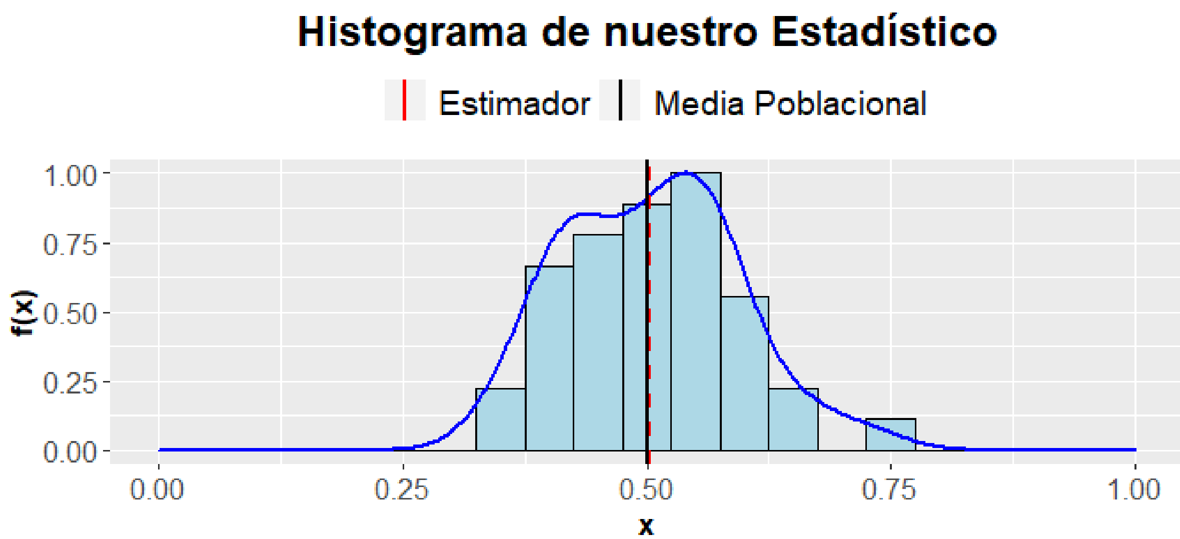
Varianza de la Media Muestral	
Valor Obtenido	Valor Teórico
0.005200948	0.005

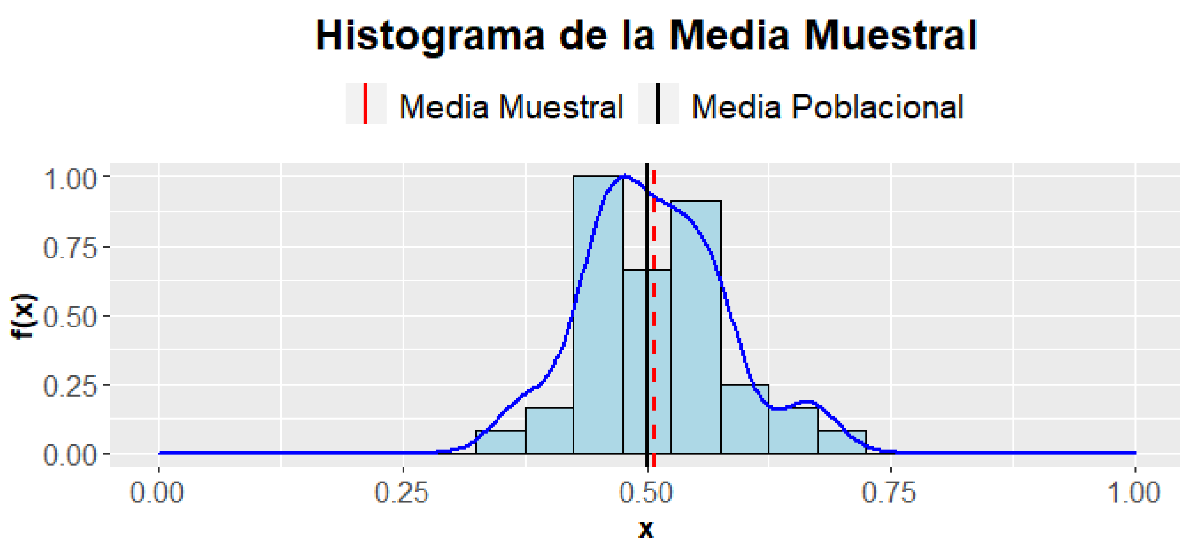
Podemos observar que la varianza de la media muestral que hemos obtenido, es muy parecida a la teórica, y si el tamaño muestral aumentará infinitamente, convergería a dicho valor teórico.

6. Comparaciones y Conclusiones

Comparativa de Estimadores							
	Medidas de Centralización		***Medidas de Dispersión***			***Medidas de Forma***	
Estimador	Media	Mediana	SD	IQR	MAD	Curtosis	Asimetría
Media Muestral	0.5072706	0.5012344	0.07211760	0.08862303	0.06388053	3.077518	0.3099279
Estadístico	0.5035930	0.5041133	0.08791211	0.13532182	0.09725723	2.656788	0.2748998







Como se puede apreciar en los boxplots, histogramas de ambos y la tabla comparativa, ambos estadísticos aproximan bastante bien la media poblacional pero la distribución de las medias muestrales tiene bastante menos dispersión. Concretamente su desviación típica es menor que la de nuestro estadístico, por lo que podemos concluir que **la media muestral es un estimador más preciso de la media poblacional que nuestro estadístico** (la media entre el primer y el tercer cuartil de los datos). Aún así podemos concluir también que nuestro estadístico es un muy buen estimador de la media poblacional, en la medida en que su desviación típica no difiere demasiado de la de la media muestral. Es decir, **en caso de que no pudieramos usar la media muestral como estimador de la poblacional, nuestro estadístico sería una buena alternativa.**

Anexo I: Cálculo de los Momentos

Cálculo de la Esperanza

Sea $X \sim Beta(\alpha, \beta)$:

$$E[X] = \int_{-\infty}^{\infty} x \cdot f(x)\, dx = \int_0^1 x \cdot \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\cdot\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}\, dx =$$

$$= \int_0^1 x \cdot \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1}\, dx =$$

$$\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} = \frac{1}{B(\alpha,\beta)}$$ Propiedad

$$= \frac{1}{B(\alpha,\beta)} \int_0^1 x^{(\alpha+1)-1}(1-x)^{\beta-1}\, dx$$

Propiedad

$$B(\alpha+1,\beta) = \int_0^1 x^{(\alpha+1)-1}(1-x)^{\beta-1}\, dx$$

$$= \frac{B(\alpha+1,\beta)}{B(\alpha,\beta)} = \frac{\frac{\alpha}{\alpha+\beta}B(\alpha,\beta)}{B(\alpha,\beta)} = \frac{\alpha}{\alpha+\beta}$$

$$\boxed{E[X] = \frac{\alpha}{\alpha+\beta}}$$

$$B(\alpha+1,\beta) = \frac{\alpha}{\alpha+\beta} B(\alpha,\beta)$$

(propiedad)

Cálculo de la varianza (1)

$$E[x^2] = \int_{-\infty}^{\infty} x^2 \cdot p(x)\, dx = \int_0^1 x^2 \cdot \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\cdot\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1}\, dx =$$

$$= \int_0^1 x^2 \cdot \frac{1}{B(\alpha,\beta)} x^{\alpha-1} (1-x)^{\beta-1}\, dx =$$

$$= \frac{1}{B(\alpha,\beta)} \int_0^1 x^2 \cdot x^{\alpha-1} (1-x)^{\beta-1}\, dx =$$

$$= \frac{1}{B(\alpha,\beta)} \int_0^1 x^{(\alpha+2)-1} (1-x)^{\beta-1}\, dx$$

$$= \frac{B(\alpha+2,\beta)}{B(\alpha,\beta)} = \frac{B((\alpha+1)+1,\beta)}{B(\alpha,\beta)} = \frac{\alpha+1}{\alpha+1+\beta} \cdot \frac{B(\alpha+1,\beta)}{B(a,b)} =$$

$$\uparrow \quad B((\alpha+1)+1,\beta) = \frac{\alpha+1}{\alpha+1+\beta} \cdot B(\alpha+1,\beta)$$

$$= \left(\frac{\alpha+1}{\alpha+1+\beta}\right) \cdot \left(\frac{\alpha}{\alpha+\beta}\right) = \frac{\alpha+1}{\alpha+1+\beta} \cdot \frac{\alpha}{\alpha+\beta}$$

Es decir: $\boxed{E[x^2] = \frac{\alpha+1}{\alpha+1+\beta} \cdot \frac{\alpha}{\alpha+\beta}}$

Cálculo de la varianza (2)

Por lo tanto:

$$\mathrm{var}[X] = E[X^2] - E[X]^2 = \frac{(\alpha+1)\,\alpha}{(\alpha+1+\beta)(\alpha+\beta)} - \left(\frac{\alpha}{\alpha+\beta}\right)^2 =$$

$$= \frac{1}{(\alpha+\beta+1)(\alpha+\beta)^2}\Big((\alpha+1)\alpha(\alpha+\beta) - \alpha^2(\alpha+\beta+1)\Big) =$$

$$= \frac{1}{(\alpha+\beta+1)(\alpha+\beta)^2}\left(\cancel{\alpha^3} + \cancel{\alpha^2\beta} + \cancel{\alpha^2} + \alpha\beta - \cancel{\alpha^3} - \cancel{\alpha^2\beta}\right)$$

$$= \frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}$$

$$\boxed{\mathrm{var}[X] = \frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}}$$

Anexo II: Cálculo de la Función Generadora de Momentos

$$\boxed{M_X(t) = E[e^{tX}] = \int_{-\infty}^{\infty} e^{tx} \cdot f(x)\, dx =}$$

$$= \int_0^1 e^{tx} \cdot \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1}\, dx =$$

$$= \int_0^1 e^{tx} \cdot \frac{1}{B(\alpha,\beta)} x^{\alpha-1} (1-x)^{\beta-1}\, dx =$$

$$= \frac{1}{B(\alpha,\beta)} \int_0^1 e^{tx} \cdot x^{\alpha-1} (1-x)^{\beta-1}\, dx$$

$$\overset{\uparrow}{=} \frac{1}{B(\alpha,\beta)} \cdot \int_0^1 \sum_{k=0}^{\infty} \frac{(tx)^k}{k!} x^{\alpha-1} (1-x)^{\beta-1}\, dx$$

$$e^{tx} = \sum_{k=0}^{\infty} \frac{(tx)^k}{k!}$$

$$= \frac{1}{B(\alpha,\beta)} \sum_{k=0}^{\infty} \frac{t^k}{k!} \int_0^1 x^k \cdot x^{\alpha-1} (1-x)^{\beta-1}\, dx$$

$$= \frac{1}{B(\alpha,\beta)} \cdot \sum_{k=0}^{\infty} \frac{t^k}{k!} \underbrace{\int_0^1 x^{\alpha+k-1} (1-x)^{\beta-1}\, dx}_{B(\alpha+k,\,\beta)} =$$

$$\boxed{= \sum_{k=0}^{\infty} \frac{B(\alpha+k,\beta)}{B(\alpha,\beta)} \cdot \frac{t^k}{k!}}$$

Anexo III: Histograma de nuestro estadístico (código)

```
grafico_estadistico_histogram_density <- ggplot(data = datosgrafico, aes(x = Estadístico)) +
        geom_histogram(aes(y = ..ndensity..), fill = "lightblue",
                       binwidth = 0.05, color = "black",
                       position = "identity") +
        geom_density(aes(x = Estadístico, y = ..ndensity..), color = "blue", size = 0.75) +
        ggtitle("Histograma de nuestro Estadístico") +
        geom_vline(aes(xintercept = resultados[[7]]$Media, col = "Estimador"), linetype = "dashed", siz
        geom_vline(aes(xintercept = resultados$media_poblacional,  col = "Media Poblacional"), size = 1)
        scale_color_manual(name = "", values = c("Media Poblacional" = "black", Estimador = "red")) +
        xlab("x") + ylab("f(x)") +
        theme(legend.position = "top",
              legend.text = element_text(size = 14),
              axis.text.x = element_text(size = 12),
              axis.text.y = element_text(size = 12),
              axis.title = element_text(size = 13, face = "bold"),
              plot.title = element_text(size = 18, face = 'bold', hjust = 0.5)) + xlim(0, 1)
```

Anexo IV: Histograma de la media muestral (código)

```
grafico_mediamuestral_histogram_density <- ggplot(data = datosgrafico,
                                                   aes(x = mediamuestral)) +
        geom_histogram(aes(y = ..ndensity..), fill = "lightblue",
                       color = "black", binwidth = 0.05) +
        geom_density(aes(y = ..ndensity..),
                     size = 0.75, col = "blue") +
        ggtitle("Histograma de la Media Muestral") +
        geom_vline(aes(xintercept = resultados[[5]]$Media,
                       col = "Media Muestral"), linetype = "dashed", size = 1) +
        geom_vline(aes(xintercept = resultados$media_poblacional,
                       col = "Media Poblacional"), size = 1) +
        scale_color_manual(name = "",
                           values = c("Media Poblacional" = "black",
                                      "Media Muestral" = "red")) +
        xlab("x") + ylab("f(x)") +
        theme(legend.position = "top",
              legend.text = element_text(size = 14),
              axis.text.x = element_text(size = 12),
              axis.text.y = element_text(size = 12),
              axis.title = element_text(size = 13, face = "bold"),
              plot.title = element_text(size = 18, face = 'bold',
                                        hjust = 0.5)) +
        xlim(0, 1)
```

Bibliografía

- Asimetría estadística. (2020). Recuperado de: https://es.wikipedia.org/w/index.php?title=Asimetr%C3%ADa_estad%C3%ADstica&oldid=125822430

- Beta distribution. (2020). Recuperado de: https://en.wikipedia.org/w/index.php?title=Beta_distribution&oldid=981642297
- Brown, Kyle. (2020). Understanding R: Mathematical Expressions. Recuperado de: https://rpubs.com/kylewbrown/math-expressions
- Características de estimadores. (s. f.). Recuperado de: https://www.uv.es/webgid/Inferencial/42_caractersticas_estimadores.html#:~:text=Diremos%20que%20un%20estimador%20es,muestra%20aproxime%20al%20p
- Commonly used symbols in R Markdown. (s. f.). Recuperado de: https://people.ok.ubc.ca/jpither/modules/Symbols_markdown.html
- Curtosis. (2020). Recuperado de: https://es.wikipedia.org/w/index.php?title=Curtosis&oldid=128327400
- Curtosis - Definición, qué es y concepto. (2017). Recuperado de: https://economipedia.com/definiciones/curtosis.html
- Función gamma. (2020). Recuperado de: https://ast.wikipedia.org/w/index.php?title=Funci%C3%B3n_gamma&oldid=2985655
- Greek letters, symbols, and line breaks inside a ggplot legend label . (s. f.). Recuperado de: https://stackoverflow.com/questions/27690729/greek-letters-symbols-and-line-breaks-inside-a-ggplot-legend-label
- Mathematical Annotation in R. (s. f.). Recuperado de: https://stat.ethz.ch/R-manual/R-devel/library/grDevices/html/plotmath.html
- Medidas de dispersión. (2020). Recuperado de: https://es.wikipedia.org/w/index.php?title=Medidas_de_dispersi%C3%B3n&oldid=128745923
- Molina, Luis. (2009). Apuntes de Latex. Recuperado de: http://metodos.fam.cie.uva.es/~latex/apuntes/apuntes3.pdf
- Moment Generating Function of Beta Distribution. (2018). Recuperado de: https://proofwiki.org/wiki/Moment_Generating_Function_of_Beta_Distribution#:~:text=Let%20X%E2%88%BCBeta(%CE%B1,%2Br)tl
- Ramirez, Oscar. (2015). Escribiendo funciones en R. Recuperado de: https://rpubs.com/osoramirez/93049
- Reilly, Archer. (2020). Writing Mathematic Fomulars in Markdown. Recuperado de: http://csrgxtu.github.io/2015/03/20/Writing-Mathematic-Fomulars-in-Markdown/
- Remove space between two graphs in R. (s. f.). Recuperado de: https://stackoverflow.com/questions/22226358/remove-space-between-two-graphs
- Rincón, Luis. (2013). Distribución beta. Recuperado de: https://www.youtube.com/watch?v=cqvEph836gM&ab_channel=LuisRinc%C3%B3n