

Tarea 1

Ivan Santiago Rojas Martinez

Estudiante de Pregrado en Estadística

Docente

Victor Ignacio Lopez Rios

Asignatura

Muestreo Estadístico



Sede Medellín
Agosto 29 de 2023

Índice

1	Indique la forma como seleccionó la MAS y registre los respectivos valores.	2
2	Con la muestra seleccionada, estime la proporción de hogares que tienen adultos mayores y halle el respectivo intervalo de confianza 97% usando la aproximación Binomial y la Hipergeométrica.	3
2.1	Intervalo de confianza usando la aproximación Binomial	4
2.2	Intervalo de confianza usando la aproximación Hipergeometrica	5
3	Repita el punto anterior para el número de hogares en el barrio bajo estudio que tienen adultos mayores	6
4	Cuál de los intervalos es mejor? Argumente.	6
5	Anexo	6

Índice de figuras

1	Tabla de Datos	2
---	--------------------------	---

Índice de cuadros

1	Tabla de la Muestra Aleatoria	3
2	Tabla Adultos mayores	3
3	Limite Superior	4
4	Limite Inferior	4
5	Limite Superior	5
6	Limite Inferior	6

Tarea 1

Suponga que los datos mostrados en la siguiente tabla corresponde al registro de 120 hogares que conforman un Barrio a quienes se les preguntó si tenían adultos mayores (60 o más años).

1: Corresponde a un hogar que cumple con el atributo. Seleccione una **MAS** de **15** hogares.

CH	Y	CH	Y	CH	Y	CH	Y	CH	Y	CH	Y
1	0	21	1	41	0	61	0	81	0	101	1
2	1	22	1	42	0	62	0	82	1	102	0
3	0	23	0	43	1	63	1	83	0	103	0
4	1	24	0	44	1	64	0	84	1	104	1
5	0	25	1	45	1	65	0	85	0	105	0
6	0	26	0	46	1	66	1	86	0	106	0
7	0	27	1	47	0	67	0	87	1	107	1
8	1	28	0	48	1	68	1	88	0	108	0
9	0	29	0	49	0	69	0	89	0	109	0
10	0	30	1	50	0	70	0	90	0	110	1
11	1	31	0	51	0	71	1	91	0	111	0
12	0	32	0	52	1	72	0	92	0	112	1
13	0	33	0	53	0	73	0	93	1	113	0
14	0	34	0	54	0	74	0	94	0	114	1
15	0	35	0	55	0	75	1	95	1	115	0
16	0	36	1	56	0	76	0	96	0	116	0
17	1	37	0	57	0	77	1	97	0	117	0
18	1	38	0	58	1	78	0	98	1	118	1
19	0	39	1	59	0	79	1	99	0	119	0
20	1	40	0	60	0	80	0	100	0	120	0

Figura 1: Tabla de Datos

1 Indique la forma como seleccionó la MAS y registre los respectivos valores.

Usando el software estadístico R y la función `sample()`, se procede a sacar una muestra aleatoria sin remplazo, garantizando que cualquier hogar tendrá la misma probabilidad de ser seleccionado en el estudio. Además se fija una semilla = **1** con el fin de permitir reproducibilidad del estudio.

La muestra aleatoria es la siguiente:

Cuadro 1: Tabla de la Muestra Aleatoria

n	1	14	21	34	39	43	51	54	59	68	82	85	87	97	106
y	0	0	1	0	1	1	0	0	1	1	0	1	0	0	0

2 Con la muestra seleccionada, estime la proporción de hogares que tienen adultos mayores y halle el respectivo intervalo de confianza 97% usando la aproximación Binomial y la Hipergeométrica.

Se procede a filtrar los hogares que tienen adultos mayores los cuales son:

Cuadro 2: Tabla Adultos mayores

n	21	39	43	59	68	85
y	1	1	1	1	1	1

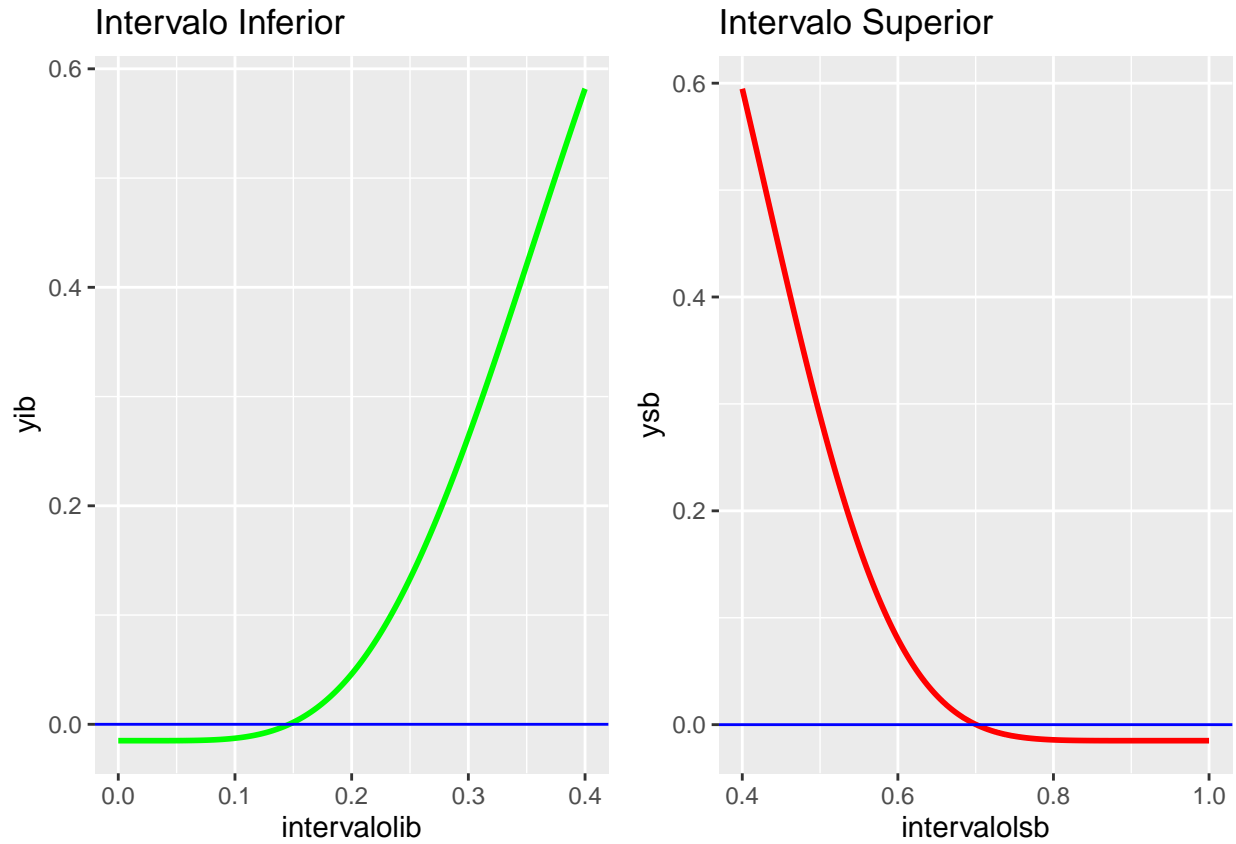
Como se puede observar en la muestra aleatorio cuentan con 6 hogares que tienen adultos mayores. Se procede a realizar los respectivos intervalos de confianza.

Tenemos:

- $N = 120$
- $n = 15$
- $\alpha = 0.03$

Basados en lo anterior, procedemos a calcular los intervalos de confianza aproximados.

2.1 Intervalo de confianza usando la aproximación Binomial



Cuadro 3: Limite Superior

Proporción	Error Absoluto
0.7005	6.95e-05
0.7006	3.50e-05
0.7007	7.00e-07
0.7008	-3.36e-05
0.7009	-6.79e-05

Se toma como limite superior = **0.7007**.

Cuadro 4: Limite Inferior

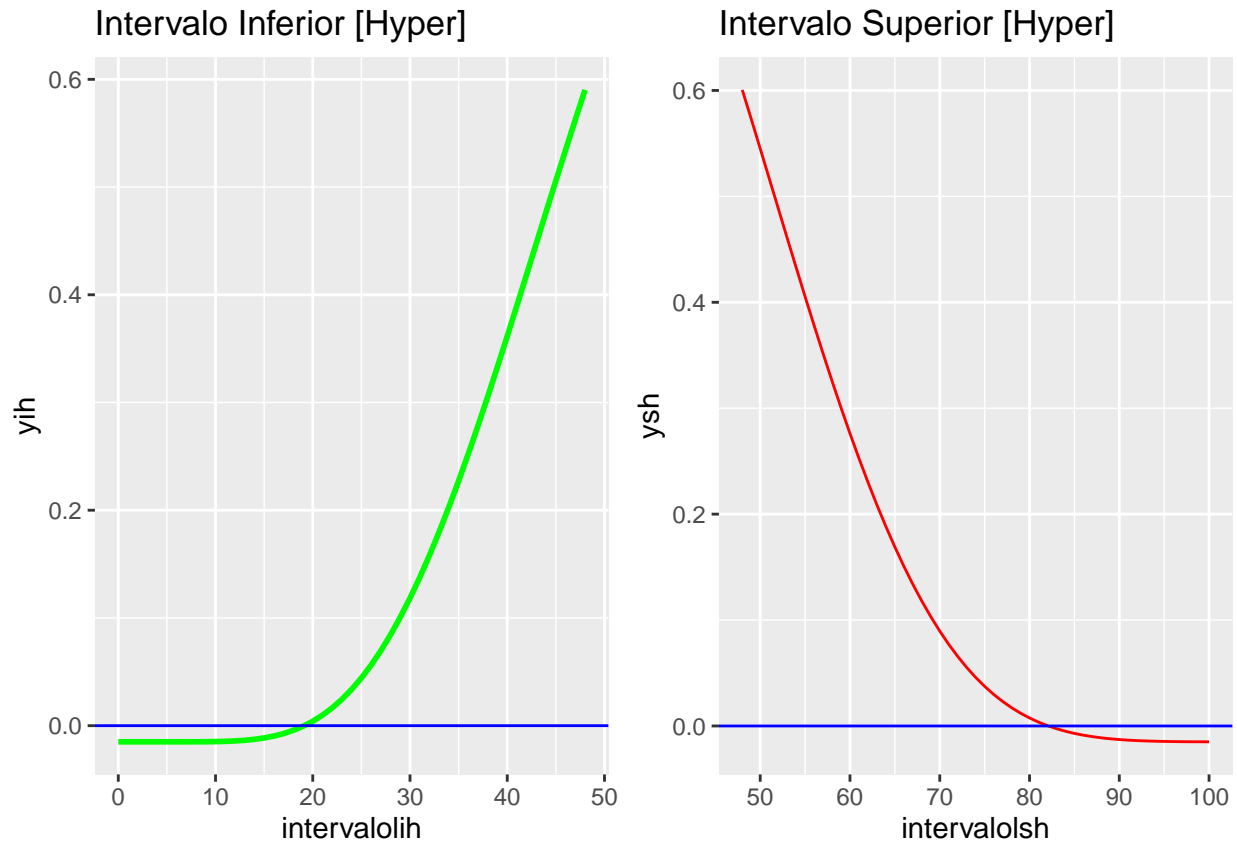
Proporción	Error Absoluto
0.5463	-6.32e-05
0.5464	-1.47e-05
0.5465	3.40e-05
0.5466	8.27e-05

Se toma como limite inferior = **0.5464**.

El intervalo de confianza del 97% usando la aproximación Binomial es:

(**0.5464** , **0.7007**)

2.2 Intervalo de confianza usando la aproximación Hipergeometrica



Cuadro 5: Limite Superior

A	Error Absoluto
80	0.0075560
81	0.0036477
82	0.0002927
83	-0.0025652
84	-0.0049796
85	-0.0070020
86	-0.0086806

Se toma como limite superior = **82**.

Cuadro 6: Limite Inferior

A	Error Absoluto
64	-0.0096645
65	-0.0073471
66	-0.0043510
67	-0.0005686
68	0.0041073
69	0.0097806

Se toma como limite inferior = **67**.

El intervalo de confianza del 97% usando la aproximación Binomial es:

(**0.5583** , **0.6833**)

3 Repita el punto anterior para el número de hogares en el barrio bajo estudio que tienen adultos mayores

Para estimar el numero total de hogares que cuentan con la característica de tener adultos mayores, bastaría con multiplicar **N** por cada limite del intervalo aproximado binomial seleccionado en el literal anterior. Dando como resultado:

- Utilizando la aproximación binomial, el intervalo de confianza está dado por:

(**66**, **85**)

- Utilizando la aproximación hypergeometrica, el intervalo de confianza está dado por:

(**67**, **82**)

4Cuál de los intervalos es mejor? Argumente.

Tomaremos como el mejor intervalo, el intervalo con mayor precisión, dicho de otra manera es el intervalo con menor longitud. Dicho intervalo es el intervalo aproximado con la hipergeometrica con una longitud de **15** comparado al de la Binomial con **19**.

5 Anexo

Punto 1:

```

set.seed(1)
N <- 120 #Total de hogares
n <- sort(sample(1:N, 15)) #Muestra

y <- c(0, 0, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 0, 0)
data <- cbind(n, y) %>% as.data.frame()

```

Punto 2:

```

#Aprox Binomial
Ls.conf.binom.prop <- function(pp,a,alfa,n){
  pbinom(a,size = n, prob = pp)-alfa*0.5
}

Li.conf.binom.prop <- function(pp,a,alf,n){
  1-pbinom(a-1,size = n, prob = pp)-alf/2
}

N <- 120
n <- 15
alf <- 0.03

intervalolsb <- seq(a/n, 1, 0.0001)
intervalolib <- seq(0, a/n, 0.0001)

ysb <- apply(matrix(intervalolsb,ncol=1), 1,
  Ls.conf.binom.prop,a,alf,n)

yib <- apply(matrix(intervalolib,ncol=1), 1,
  Li.conf.binom.prop,a,alf,n)

library(ggplot2)
library(gridExtra)
# Hacemos los respectivos graficos de los intervalos de busqueda:
p1 <- ggplot(,aes(intervalolsb,ysb))+geom_line(size=1, colour="red")+
  geom_hline(yintercept = 0, colour="blue")+
  labs(title = "Intervalo Superior")

p2 <- ggplot(,aes(intervalolib,yib))+geom_line(size=1, colour="green")+
  geom_hline(yintercept = 0, colour="blue")+
  labs(title = "Intervalo Inferior")

```



```

grid.arrange(p2,p1,ncol=2)

ind1 <- which((abs(ysb)<0.0001))
mb1 <- matrix(c(intervalolsb[ind1],ysb[ind1]), ncol = 2)
kable(mb1, col.names = c("Proporción", "Error Absoluto"), caption = "Limite Superior") %>%
  kable_styling(latex_options = "HOLD_position")
maxb <- mb1[abs(mb1[,2])==min(abs(mb1[,2])),]
maxb <- maxb[1]

ind2 <- which((abs(yib)<0.0001))
mb2 <- matrix(c(intervalolsb[ind2],yib[ind2]), ncol = 2)
kable(mb2, col.names = c("Proporción", "Error Absoluto"), caption = "Limite Inferior") %>%
  kable_styling(latex_options = "HOLD_position")
minb <- mb2[abs(mb2[,2])==min(abs(mb2[,2])),]
minb <- minb[1]

#Aprox Hipergeometrica
Ls.conf.hiper.prop <- function(A,a,N,n,alfa){
  phyper(a,A,N-A,n)-alfa*0.5
}

Li.conf.hiper.prop <- function(A,a,N,n,alf){
  1-phyper(a-1,A,N-A,n)-alf/2
}

intervalolsh <- seq(a/n *N, 100,1)
intervalolih <- seq(0,a/n * N,1)

ysh <- apply(matrix(intervalolsh,ncol=1), 1,
  Ls.conf.hiper.prop,a,N,n,alf)

yih <- apply(matrix(intervalolih,ncol=1), 1,
  Li.conf.hiper.prop,a,N,n,alf)

p1 <- ggplot()+geom_line(aes(intervalolsh, ysh),colour="red")+
  geom_hline(yintercept = 0, colour="blue")+
  labs(title = "Intervalo Superior [Hyper]")

```

```

p2 <- ggplot(,aes(intervalolih,yih))+geom_line(size=1, colour="green")+
  geom_hline(yintercept = 0, colour="blue")+
  labs(title = "Intervalo Inferior [Hyper]")

grid.arrange(p2,p1,ncol=2)

ind1 <- which((abs(ysh)<0.01))

mh1 <- matrix(c(intervalolsh[ind1],ysh[ind1]), ncol = 2)
kable(mh1, col.names = c("A", "Error Absoluto"), caption = "Limite Superior") %>%
  kable_styling(latex_options = "HOLD_position")
maxh <- mh1[abs(mh1[,2])==min(abs(mh1[,2])),]
maxh <- maxh[1]

ind2 <- which((abs(yih)<0.01))

mh2 <- matrix(c(intervalolsh[ind2],yih[ind2]), ncol = 2)
kable(mh2, col.names = c("A", "Error Absoluto"), caption = "Limite Inferior") %>%
  kable_styling(latex_options = "HOLD_position")
minh <- mh2[abs(mh2[,2])==min(abs(mh2[,2])),]
minh <- minh[1]

```

Punto 3:

```

snt <- ceiling(maxb*N)
int <- ceiling(minb*N)

cat("(",int,",",snt,")")

snA <- maxh
inA <- minh

cat("(",inA,",",snA,")")

```

Punto 4:

```

abs(66-85)
abs(67-82)

```