

Unidad 5: Inferencia Estadística: Estimación

Mg. J. Eduardo Gamboa U.

Table of contents

Descripción del caso	1
Carga de paquetes	2
Lectura de datos	2
Intervalo de confianza para la media	2
Intervalo de confianza para la varianza	3
Intervalo de confianza para la desviación estándar	4
Intervalo de confianza para la proporción	5
Intervalo de confianza para la mediana	6
Intervalo de confianza para la mediana	7

Descripción del caso

Una universidad desea evaluar el servicio de su biblioteca en términos de tiempos de espera, uso del servicio y nivel de satisfacción de los estudiantes. Para ello, se encuestó a 70 estudiantes al finalizar el semestre. Se recolectaron las siguientes variables

- **Tiempo_Espera** (en minutos): tiempo que un estudiante espero para ser atendido.
- **Libros_Prestados**: número de libros en préstamo por parte del estudiante en el ciclo.
- **Satisfecho**: respuesta a si están satisfechos con el servicio.

Carga de paquetes

```
library(pacman)
p_load(dplyr, magrittr, EnvStats, binom)
```

Lectura de datos

```
datos <- read.csv('datos_u5.csv')
datos |> head(10)
```

	ID	Tiempo_Espera	Libros_Prestados	Satisfecho
1	1	8.4	3	Sí
2	2	9.0	2	No
3	3	5.8	5	Sí
4	4	8.4	1	Sí
5	5	8.1	2	Sí
6	6	5.1	3	No
7	7	8.9	5	Sí
8	8	9.4	1	Sí
9	9	10.7	4	Sí
10	10	10.6	3	No

Intervalo de confianza para la media

Existe una confianza del 90% de que el verdadero tiempo medio de espera se encuentre entre los 7.61 y 8.70 minutos.

```
datos |>
  pull(Tiempo_Espera) |>
  t.test(conf.level = 0.90) |>
  use_series(conf.int)
```

```
[1] 7.613140 8.704003
attr(,"conf.level")
[1] 0.9
```

Con un 99% de confianza, se puede afirmar que la cantidad media poblacional de libros prestados en el ciclo se encuentra entre 2.84 y 3.54 libros.

```
datos |>
  pull(Libros_Prestados) |>
  t.test(conf.level = 0.99) |>
  use_series(conf.int)
```

```
[1] 2.835162 3.536267
attr(,"conf.level")
[1] 0.99
```

Intervalo de confianza para la varianza

Se tiene un 95% de confianza de que la varianza poblacional de los tiempos de espera se encuentre entre los 5.51 y 10.79 minutos².

```
datos |>
  pull(Tiempo_Espera) |>
  varTest(conf.level = 0.95) |>
  use_series(conf.int)
```

```
      LCL      UCL
5.507663 10.786414
attr(,"conf.level")
[1] 0.95
```

Existe una confianza del 90% de que la verdadera varianza de la cantidad de libros prestados esté contenida en el intervalo de 0.95 a 1.66 libros ²

```
datos |>
  pull(Libros_Prestados) |>
  varTest(conf.level = 0.90) |>
  use_series(conf.int)
```

```
      LCL      UCL
0.946242 1.662480
attr(,"conf.level")
[1] 0.9
```

Intervalo de confianza para la desviación estándar

Se tiene un 95% de confianza de que la desviación estándar poblacional de los tiempos de espera se encuentre entre los 2.35 y 3.28 minutos.

```
datos |>
  pull(Tiempo_Espera) |>
  varTest(conf.level = 0.95) |>
  use_series(conf.int) |>
  sqrt()
```

```
      LCL      UCL
2.346841 3.284268
attr(,"conf.level")
[1] 0.95
```

Existe una confianza del 90% de que la verdadera desviación estándar de la cantidad de libros prestados esté contenida en el intervalo de 0.97 a 1.29 libros.

```
datos |>
  pull(Libros_Prestados) |>
  varTest(conf.level = 0.90) |>
  use_series(conf.int) |>
  sqrt()
```

```
      LCL      UCL
0.9727497 1.2893719
attr(,"conf.level")
[1] 0.9
```

Intervalo de confianza para la proporción

```
(datos |> nrow() -> n)
```

```
[1] 70
```

```
(datos |> filter(Satisfecho=="Sí") |> nrow() -> x)
```

```
[1] 46
```

```
(x/n -> p)
```

```
[1] 0.6571429
```

```
n*p
```

```
[1] 46
```

```
n*(1-p)
```

```
[1] 24
```

```
binom.confint(x, n, methods = "asymptotic", conf.level = 0.96) # Aproximación Normal
```

	method	x	n	mean	lower	upper
1	asymptotic	46	70	0.6571429	0.5406271	0.7736587

```
binom.confint(x, n, methods = "wilson", conf.level = 0.96) # Aproximación de Wilson
```

	method	x	n	mean	lower	upper
1	wilson	46	70	0.6571429	0.5347039	0.7617206

```
binom.confint(x, n, methods = "exact", conf.level = 0.96) # Método exacto
```

	method	x	n	mean	lower	upper
1	exact	46	70	0.6571429	0.5283751	0.7710224

Con un nivel de confianza del 96%, se estima que la verdadera proporción de estudiantes usuarios de la biblioteca que están satisfechos se encuentra entre 0.5347 y 0.7617.

Intervalo de confianza para la mediana

```
# Estimación puntual
datos |>
  pull(Tiempo_Espera) |>
  median()
```

```
[1] 8.3
```

```
# Estimación intervalar
library(boot)
library(simpleboot)
datos |>
  pull(Tiempo_Espera) |>
  one.boot(FUN = median, R = 2000) |>
  boot.ci(type = "perc")
```

BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS

Based on 2000 bootstrap replicates

CALL :

```
boot.ci(boot.out = one.boot(pull(datos, Tiempo_Espera), FUN = median,
  R = 2000), type = "perc")
```

Intervals :

Level	Percentile
-------	------------

95%	(7.2, 9.2)
-----	--------------

Calculations and Intervals on Original Scale

Existe una confianza del 95% de que la verdadera mediana del tiempo de espera esté contenida en el intervalo de 7.2 a 9.2 minutos.

Intervalo de confianza para la mediana

```
# Estimación puntual
library(sjstats)
datos |>
  pull(Libros_Prestados) |>
  cv()
```

```
[1] 0.3475499
```

```
# Estimación intervalar
datos |>
  pull(Libros_Prestados) |>
  one.boot(FUN = cv, R = 2000) |>
  boot.ci(type = "perc", conf = 0.99)
```

BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS

Based on 2000 bootstrap replicates

CALL :

```
boot.ci(boot.out = one.boot(pull(datos, Libros_Prestados), FUN = cv,
  R = 2000), conf = 0.99, type = "perc")
```

Intervals :

Level	Percentile
-------	------------

99%	(0.2708, 0.4164)
-----	--------------------

Calculations and Intervals on Original Scale

Existe una confianza del 99% de que el verdadero coeficiente de variación del número de libros prestados esté contenido en el intervalo de 27.22% a 42.24%