DEPARTAMENTO
DE COMPUTACION
Facultad de Ciencias Exactas y Naturales - UBA

Trabajo Práctico II

Probabilidad y Estadística
Primer Cuatrimestre de 2019

Integrante	LU	Correo electrónico
Facundo Linlaud	561/16	facundolinlaud@gmail.com
Marco Sotomayor	?	marco.soto1995@gmail.com

Facultad de Ciencias Exactas y Naturales
Universidad de Buenos Aires
Ciudad Universitaria - (Pabellón I/Planta Baja)
Intendente Güiraldes 2160 - C1428EGA
Ciudad Autónoma de Buenos Aires - Rep. Argentina
Tel/Fax: (54 11) 4576-3359
http://www.fcen.uba.ar

Índice

1. Ejercicio 1 **3**
- 1.1. Momentos de la muestra . . . 3
- 1.2. Estimador de Máxima Verosimilitud . . . 3

2. Ejercicio 2 **4**

3. Ejercicio 3 **4**

4. Ejercicio 4 **4**

5. Ejercicio 5 **4**

6. Ejercicio 6 **4**

7. Ejercicio 7 **5**

1. Ejercicio 1

1.1. Momentos de la muestra

Como sabemos que $X_1, \ldots, X_n$ es una muestra aleatoria con distribución $U[0, b]$, sabemos que el primer momento de una distribución uniforme es su esperanza, mientras que su segundo momento su varianza. De esta manera podemos estimar el primer y segundo momento de la muestra:

$$m_1 = E[X] = \frac{a+b}{2} \tag{1}$$

Sabiendo además que $a = 0$, podemos utilizar la equivalencia (1) para estimar el valor de b:

$$\begin{aligned}
\bar{X}_n &= \frac{1}{n} * \sum_{i=1}^{n} X_i \\
\iff \bar{X}_n &= \frac{1}{n} * n * \frac{a+b}{2} \\
\iff \bar{X}_n &= \frac{a+b}{2}
\end{aligned}$$

Luego, despejando b y reemplanzando a por 0, obtenemos su estimador:

$$\hat{b}_{mom} = 2 * \bar{X}_n$$

1.2. Estimador de Máxima Verosimilitud

Sabemos que la función de densidad de una distribución uniforme es:

$$f_X(x) = \frac{1}{b-a} I_{(a,b)}(x)$$

Sabiendo que $a = 0$ y que $a < b$ por propiedades de distribución uniforme, procederemos a maximizar $f_X(x)$ para estimar el valor de máxima verosimilitud de b:

$$\begin{aligned}
L(\theta) &= \prod_{i=1}^{n} \frac{1}{b-a} I_{(a,b)}(x) \\
\iff L(\theta) &= \prod_{i=1}^{n} \frac{1}{b-0} I_{(a,b)}(x) \\
\iff L(\theta) &= \prod_{i=1}^{n} \frac{1}{b} I_{(x_i,+\infty)}(\theta)
\end{aligned}$$

Por lo tanto:

$$L(\theta) = \begin{cases} \frac{1}{b^n} & \text{si } max\{x_1, \ldots, x_n\} < \theta \\ 0 & \text{sino} \end{cases}$$

Esto implica que $\frac{1}{b^n}$ es máximo cuando b^n es mínimo. Luego, el valor más chico posible de b es el valor más grande que haya tomado algún resultado en la muestra $X_1, \ldots, X_n$. Porque si b fuese menor a alguno de ellos, la probabilidad total sería nula (por contener un elemento fuera del rango de la distribución). Finalmente, tenemos un valor para nuestro estimador:

$$\hat{b}_{mv} = max\{x_1, \ldots, x_n\}$$

2. Ejercicio 2

Dada una muestra $X_1, ..., X_n$, estimar el doble de su mediana implica tomar los únicos datos que tenemos, es decir la muestra, calcular su mediana muestral y multiplicarla por dos. Su implementación puede ser observada en el archivo **tp2.py**.

3. Ejercicio 3

El procedimiento para resolver este ejercicio fue obtener el estimador $hat\theta_{mv}$ y restarlo por $b = 1$ para obtener el error de máxima verosimilitud. El mismo procedimiento fue realizado análogamente para el estimador de momento $\hat{\theta}_{mom}$. Al ejecutar **tp2.py**, se puede observar que módulo del error tiende a ser menor o igual a $0{,}1$.

4. Ejercicio 4

A continuación se exponen los datos requeridos por el ejercicio 4. Para más detalles, ver **tp2.py**.

Estimador	$\hat{\theta}_{mv}$	$\hat{\theta}_{mom}$	$\hat{\theta}_{med}$
Promedio	$0{,}935$	$1{,}003$	$1{,}005$
Sesgo	$0{,}064$	$-0{,}003$	$-0{,}005$
Varianza	$0{,}003$	$0{,}022$	$0{,}057$
ECM	$0{,}007$	$0{,}022$	$0{,}057$

5. Ejercicio 5

Ver **tp2.py**.

6. Ejercicio 6

A continuación, procederemos a graficar los sesgos, varianzas y errores cuadráticos medios de los estimadores en función de los distintos valores de **b** a estimar.

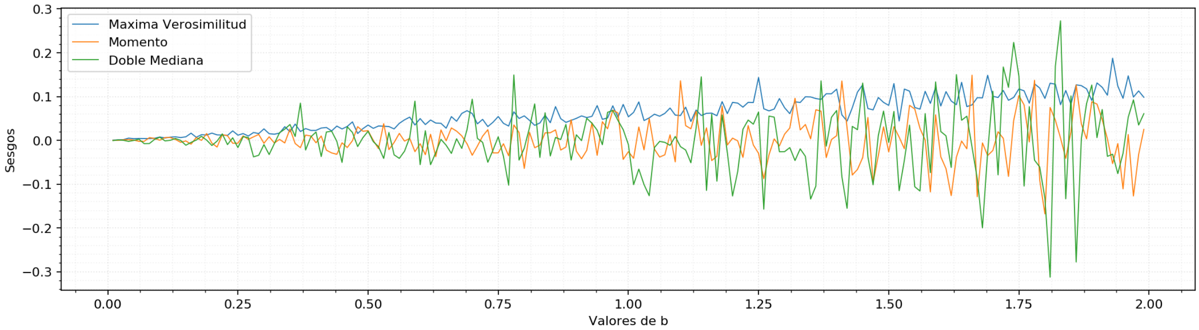

Figura 1: Sesgos de los estimadores en función de los valores de **b**. $a = 0, n = 15$

En la figura 1 se puede observar que, a medida que el valor de b aumenta, el valor absoluto del sesgo del estimador de b también incrementa. Esto tiene sentido porque, al observar la figura 2, a medida que b aumenta, la varianza de los estimadores también.

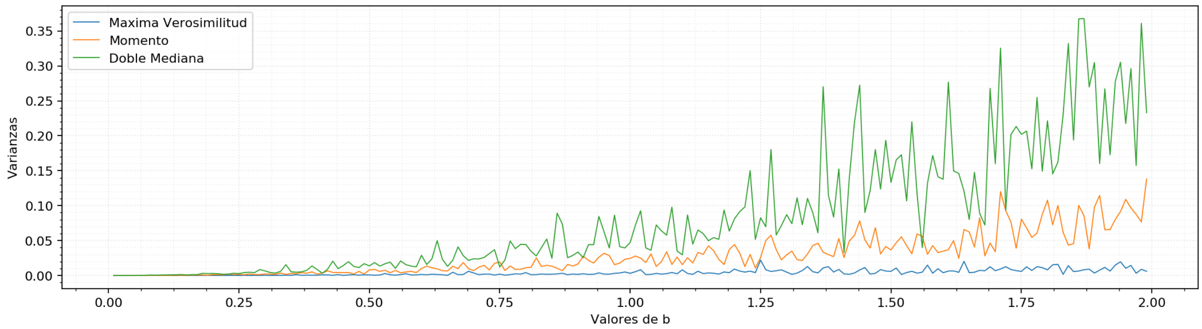

Figura 2: Varianzas de los estimadores en función de los valores de **b**. $a = 0, n = 15$

La figura 2 nos muestra el comportamiento de las varianzas. Podemos observar cómo la varianza del estimador de la doble mediana $\hat{\theta}_{med}$ tiene el crecimiento máximo de entre los tres, seguida por la varianza del estimador $\hat{\theta}_{mom}$ y en último lugar, la varianza del estimador $\hat{\theta}_{mv}$. Podemos inferir que el estimador de máxima verosimilitud es el estimador de mejor calidad de entre los tres, pues su varianza no crece tan agresivamente al aumentar el valor de b como en las otras alternativas.

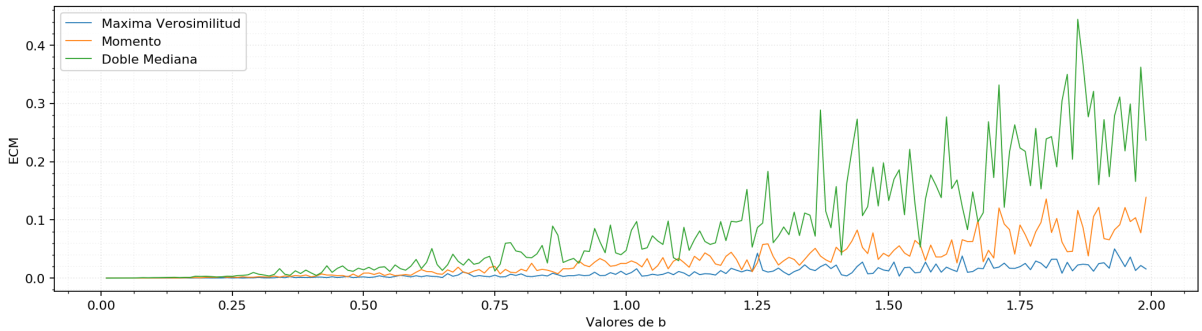

Figura 3: ECM de los estimadores en función de los valores de **b**. $a = 0, n = 15$

Finalmente, la figura 3, complementa la conclusión llegada en los análisis anteriores. A medida que el valor de b aumenta, el error cuadrático medio de cada estimador también. Podemos asegurar que el mejor estimador de los tres es el de **máxima verosimilitud**, que posee el menor crecimiento de los tres.

7. Ejercicio 7

En este experimento observaremos el error cuadrático medio para cada estimador en función de n, con n siendo la cantidad de muestras de tamaño 15. En el gráfico siguiente, podemos observar que errores cuadráticos medios de cada estimador convergen a valores no nulos. En particular, observando los datos adquiridos en el **ejercicio 4**, podemos observar que el error cuadrático medio de $\hat{\theta}_{mv}$, $\hat{\theta}_{mom}$ y $\hat{\theta}_{med}$ tienden a $0{,}003$, $0{,}022$ y $0{,}057$ respectivamente, que no coincidentalmente aproximan al valor de sus varianzas con $n \to \infty$. ¿Esto significa que los estimadores son insesgados? Probablemente. Lo analizaremos adelante.

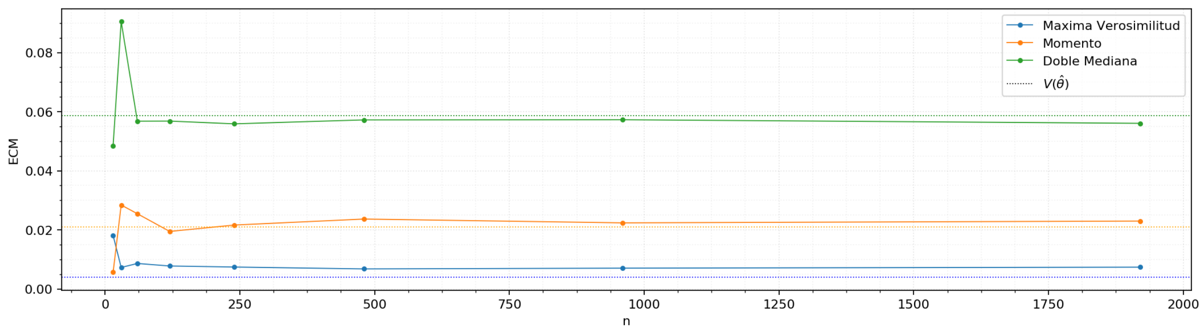

Figura 4: ECM de los estimadores en función de n. $a = 0, b = 1$

Observando el gráfico 4, podemos sospechar que los errores cuadráticos medios de los tres estimadores tienden a la varianza cuando n es muy grande. Podemos hipotetizar que los estimadores son consistentes, sobretodo porque en la figura 5 se aprecia que el sesgo de los estimadores tiende a cero a medida que n crece, es decir, los estimadores son – por lo menos – asintoticamente insesgados. Veamos los sesgos de los estimadores a medida que $n \to \infty$:

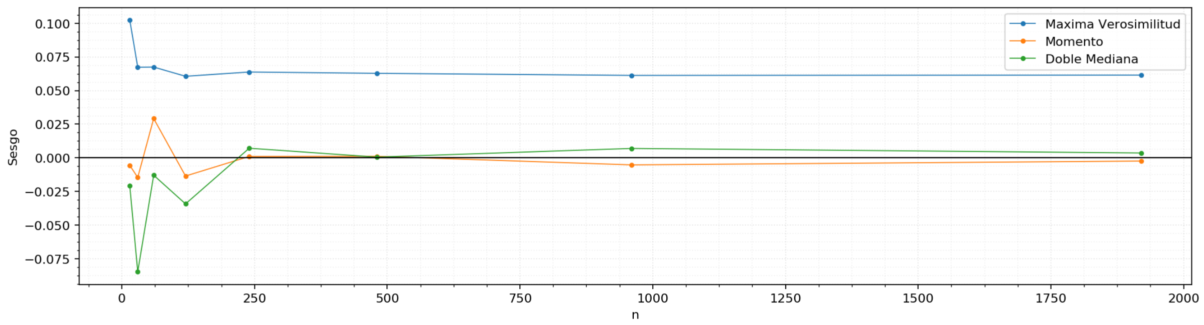

Figura 5: Sesgos de los estimadores en función de n. $a = 0, b = 1$

Podemos conjeturar, entonces, que los sesgos de los estimadores $\hat{\theta}_{mom}$ y $\hat{\theta}_{med}$ tienden a cero. ¿Pero qué hay del estimador de máxima verosimilitud? Si calculamos $E(\hat{\theta}_{mv})$, llegaremos a que equivale a $\frac{n}{n+1} * \theta$, que si bien no es igual a θ, sí tiende a este valor cuando $n \to \infty$. Luego el estimador de máxima verosimilitud es asintóticamente insesgado por lo tanto los errores cuadráticos medios de los estimadores efectivamente tienden a sus varianzas.