Trabajo Practico 2 Parte 2

Ignacio Pardo & Luca Mazzarello

2022-08-31

Al igual que en la Actividad 1, en esta actividad generaremos muestras de una dis- tribución uniforme en el intervalo [0,3). Es decir, trabajaremos con v.a. $(X_i)_{i \geq 1}$ i.i.d., $X_i \sim U(0, \theta)$ con $\theta = 3$ y calcularemos el ECME de los estimadores.

$\hat{\theta}_n = \hat{\theta}_n(X_1, ..., X_n) = 2\overline{X_n}$, y $\tilde{\theta}_n = \tilde{\theta}_n(X_1, ..., X_n) = max\{X_1, ..., X_n\}$.

a) Calculen, en base a simulaciones con `Nrep = 10000`, el error cuadrático medio empírico (ECME) de cada estimador $\hat{\theta}_n$ y $\tilde{\theta}_n$, para un tamaño de muestra de $n = 30$. ¿Cuál tiene menor ECME? ¿Guarda relación con los visto en la teórica?

```
f_est1 <- function (q) {
  2 * mean(q)
}


f_est2 <- function (q) {
  max(q)
}
```

```
ECME <- function(Nrep, unif_v, f){
  mean((replicate(Nrep, f(unif_v))-3)^2)
}
```

```
unif_vals <- function(cant){
  runif(cant, 0, 3)
}
```

```
Nrep = 10000
```

```
vals_30 = unif_vals(30)

ECME_est1_30 <- ECME(Nrep, vals_30, f_est1)
ECME_est2_30 <- ECME(Nrep, vals_30, f_est2)

ECME_est1_30
```

```
## [1] 0.01920569
```

```
ECME_est2_30
```

```
## [1] 0.001305959
```

El estimador $est2$ tiene menor ECME y, como vimos en la teorica, se cumple que $ECM(\hat{\sigma}_n^2) < ECM(s_n^2) \forall n$

b) Repita el ítem anterior considerando ahora muestras de tamaño $n = 50$ y $n = 100$ y represente en una tabla, los ECME de cada estimador para cada tamaño de muestra $n = 30$, $n = 50$ y $n = 100$

```r
vals_50 = unif_vals(50)

ECME_est1_50 <- ECME(Nrep, vals_50, f_est1)
ECME_est2_50 <- ECME(Nrep, vals_50, f_est2)

ECME_est1_50
```

```
## [1] 0.03067346
```

```r
ECME_est2_50
```

```
## [1] 0.00339961
```

```r
vals_100 = unif_vals(100)

ECME_est1_100 <- ECME(Nrep, vals_100, f_est1)
ECME_est2_100 <- ECME(Nrep, vals_100, f_est2)

ECME_est1_100
```

```
## [1] 0.02828842
```

```r
ECME_est2_100
```

```
## [1] 0.01005392
```

```r
ECMEs_est_1 = c(ECME_est1_30, ECME_est1_50, ECME_est1_100)
ECMEs_est_2 = c(ECME_est2_30, ECME_est2_50, ECME_est2_100)

ECMEs = data.frame(c(30, 50, 100), ECMEs_est_1, ECMEs_est_2)

names(ECMEs)[names(ECMEs) == "ECMEs_est_1"] <- "2*promedio"
names(ECMEs)[names(ECMEs) == "ECMEs_est_2"] <- "max"
names(ECMEs)[names(ECMEs) == "c.30..50..100."] <- "cant"

ECMEs
```

```
##   cant 2*promedio         max
## 1   30 0.01920569 0.001305959
## 2   50 0.03067346 0.003399610
## 3  100 0.02828842 0.010053921
```

i. Para cada uno de los n's considerados, ¿qué estimador tiene menor EMCE? ¿Se mantiene la misma relación observada en el ítem anterior?

El estimador `f_est2` sigue siendo el que menor ECME tiene y se mantiene la relacion que en el item anterior

ii. Para cada estimador, ¿qué ocurre con el ECME cuando n aumenta? ¿Cómo se relaciona lo que observa con lo visto en la teórica?

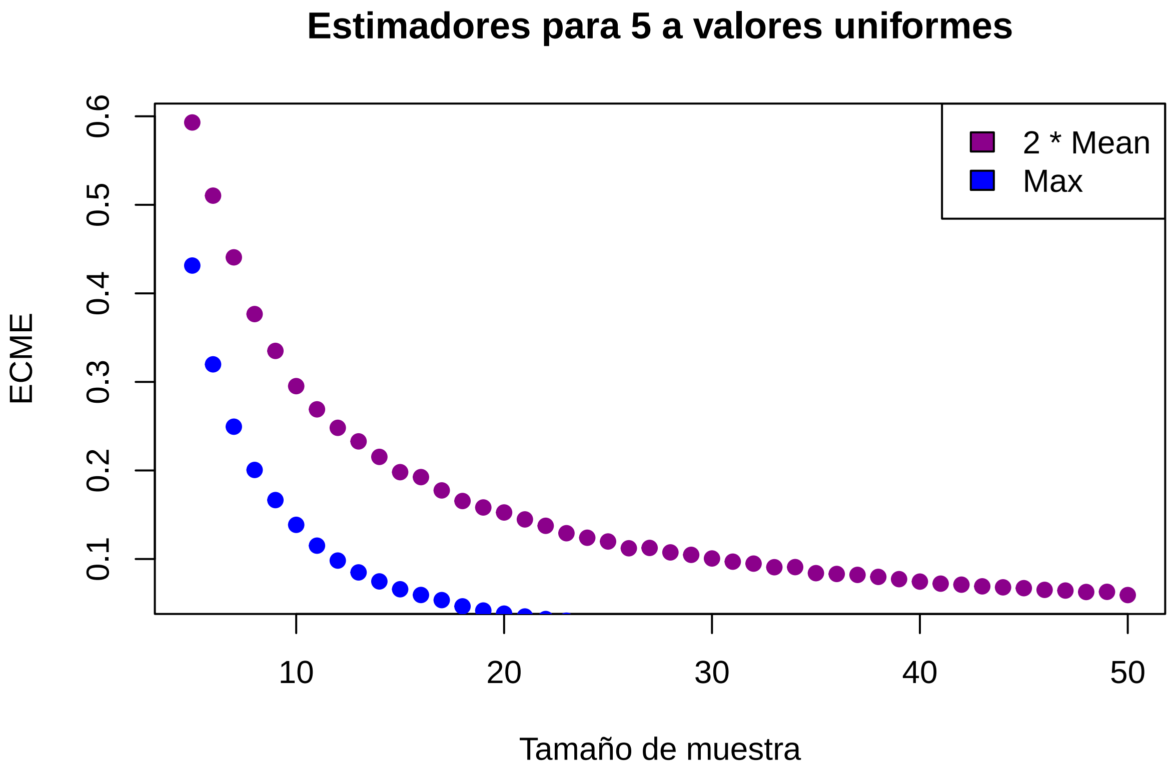
Cuanto mayor sea n, el ECME disminuye en ambos estimadores. Esto se explica, en el caso de `est1` con el teorema presentado en clase que dice que si $ECM(\hat{\theta}_n^2) \xrightarrow[n\to\infty]{} 0 \Rightarrow \hat{\theta}_n^2 \xrightarrow{p} 0$ y, con respecto al caso de `est2`, $ECM(\hat{\theta}_n^2) = \frac{2}{(n+2)(n+1)} \times \theta^2 \xrightarrow[n\to\infty]{} 0$

c. Ahora calcule el ECME de ambos estimadores para cada tamaño de muestra entre 5 y 50 (es decir, $n = 5, n = 6, n = 7, ..., n = 50$), siempre utilizando `Nrep = 10000`. Para cada estimador, guarde todos los ECMEs en un vector. Represente en un gráfico de dispersión el ECME de cada estimador versus el tamaño de muestra, utilizando colores distintos para cada estimador. ¿Qué observa? ¿Qué estimador elegiría?

```r
vals_est1 = rep(NA, 50)
vals_est2 = rep(NA, 50)

for (i in (5:50)){
  vals_est1[i] <- mean((replicate(Nrep, f_est1(runif(i, 0, 3)))-3)^2)
  vals_est2[i] <- mean((replicate(Nrep, f_est2(runif(i, 0, 3)))-3)^2)
}

plot(vals_est1, col="darkmagenta", pch=19, xlim=c(5, 50), ylab="ECME", xlab="Tamaño de muestra")
legend("topright", legend=c("2 * Mean", "Max"), fill=c("darkmagenta", "blue"))
points(vals_est2, col="blue", pch=19, xlim=c(5, 50))
title("Estimadores para 5 a valores uniformes")
```



Observamos que ambos estimadores tienden a cero a medida que n aumenta y elegiriamos el estimador `est2` (Max) dado a que tiende a cero mas rapido.