

IECD 2C2023: Parcial 1 - Ejercicio 3 Resuelto

Gonzalo Barrera

2023-09-26

3a - 2pts

Generar una muestra $n = 20$ de una dist exponencial con $\lambda = 1/2$ y evaluar los estimadores $T1$ y $T2$ en dicha muestra. Copiar los valores obtenidos:

```
n.x <- 20
lambda <- 1/2
muestra <- rexp(n.x, lambda)
t1 <- function(x) { mean(x)^2 }
t2 <- function(x) { mean(x^2) / 2 }

c(t1=t1(muestra), t2=t2(muestra))

##          t1          t2
## 3.438911 3.217109
```

3b - 3pts

Fijar la semilla en el valor 2023. Repetir el inciso anterior $Nrep = 1000$ veces, obteniendo para cada uno de los dos estimadores, 1000 replicaciones a guardar en sendos vectores.

```
set.seed(2023)
n.rep <- 1000
reps_t1 <- vector("numeric", n.rep)
reps_t2 <- vector("numeric", n.rep)
for (i in 1:n.rep) {
  muestra <- rexp(n.x, lambda)
  reps_t1[i] <- t1(muestra)
  reps_t2[i] <- t2(muestra)
}
```

3c - 10pts

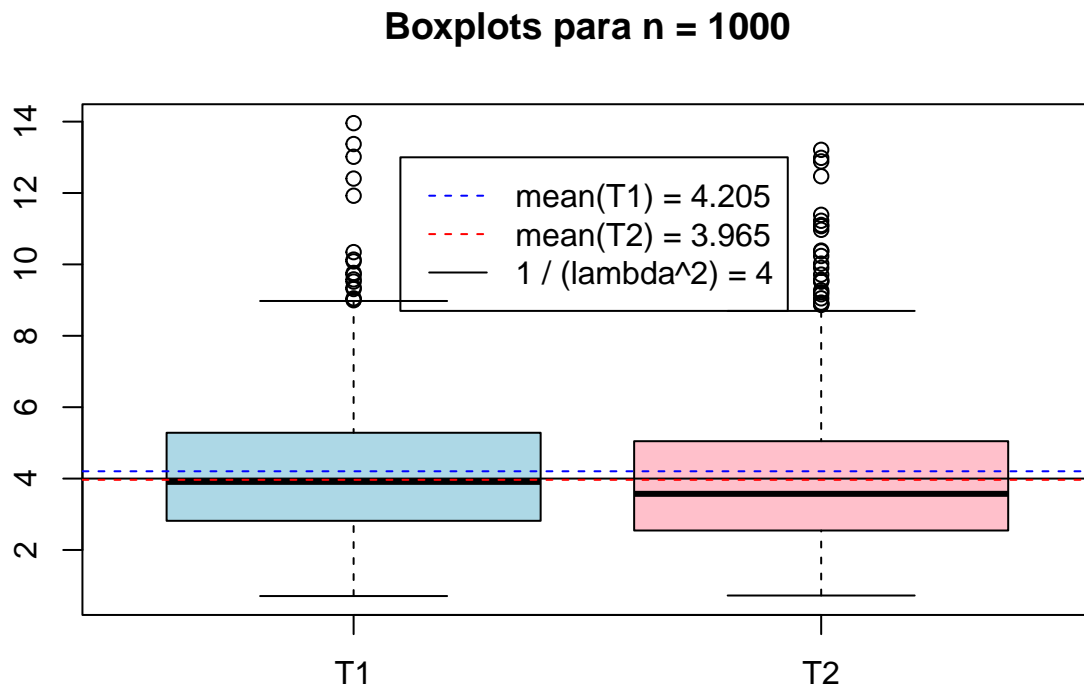
Graficar boxplots paralelos para los $Nrep$ valores obtenidos de $T1$ y $T2$ e interpretar.

```
rango <- range(c(reps_t1, reps_t2))
par(mfrow=c(1, 2))
boxplot(reps_t1, ylim=rango)
boxplot(reps_t2, ylim=rango)
```

También, aprovechando que cuando a `boxplot` se le pasa más de un vector grafica boxplots paralelos,

```
boxplot(reps_t1, reps_t2, names=c("T1", "T2"), col=c("lightblue", "pink"))
abline(h=lambda^-2)
abline(h=mean(reps_t1), col="blue", lty=2)
```

```
abline(h=mean(reps_t2), col="red", lty=2)
etiquetas <- c(
  paste0("mean(T", c(1, 2), ") = ", round(c(mean(reps_t1), mean(reps_t2)), 3)),
  "1 / (lambda^2) = 4"
)
legend(1.1, 13, legend=etiquetas, col=c("blue", "red", "black"), lty=c(2, 2, 1))
title(paste("Boxplots para n =", n.rep))
```



¿Sugieren simetría? ¿Hay valores atípicos?

Simétricos no son, en virtud de ser ambos estimadores funciones cuadráticas en las X_i , y se observan bastante valores atípicos a derecha, pasado el “bigote” de $q75 + 1.5 \text{ IQR}$.

¿Alrededor de que valor esperaríamos que estos boxplots estén centrados?

Todo boxplot está centrado en la *mediana* de la muestra, que puede o no coincidir con la *media* muestral. En nuestro caso, como la distribución es asimétrica con cola a derecha, esperaríamos que la mediana esté un poco por debajo de la media muestral.

```
quantile(reps_t1)
```

```
##          0%          25%          50%          75%         100%
## 0.7095291  2.8184083  3.9084380  5.2753948 13.9570433
```

```
quantile(reps_t2)
```

```
##          0%          25%          50%          75%         100%
## 0.7243921  2.5467953  3.5725046  5.0425645 13.2096007
```

¿Alguno de ellos parece insesgado de acuerdo con estos gráficos?

Cuando la distribución es asimétrica, que el boxplot (centrado en la mediana) no se corresponda con la media no es un problema, a priori. En este caso, la media de T1 es ~ 4.2 , mientras que la de T2 es ~ 4 . Pareciera ser que T2 es insesgado mientras que T1 tiene un ligero sesgo. Aún aumentando `n.reps` hasta $1e5$ el sesgo se mantiene, pero nada de esto se puede deducir directamente de los gráficos.

3d - 10pts

Obtener un estimador nopa de la densidad de T1 usando núcleo normal. ¿Cómo elige la ventana? Repetir para T2.

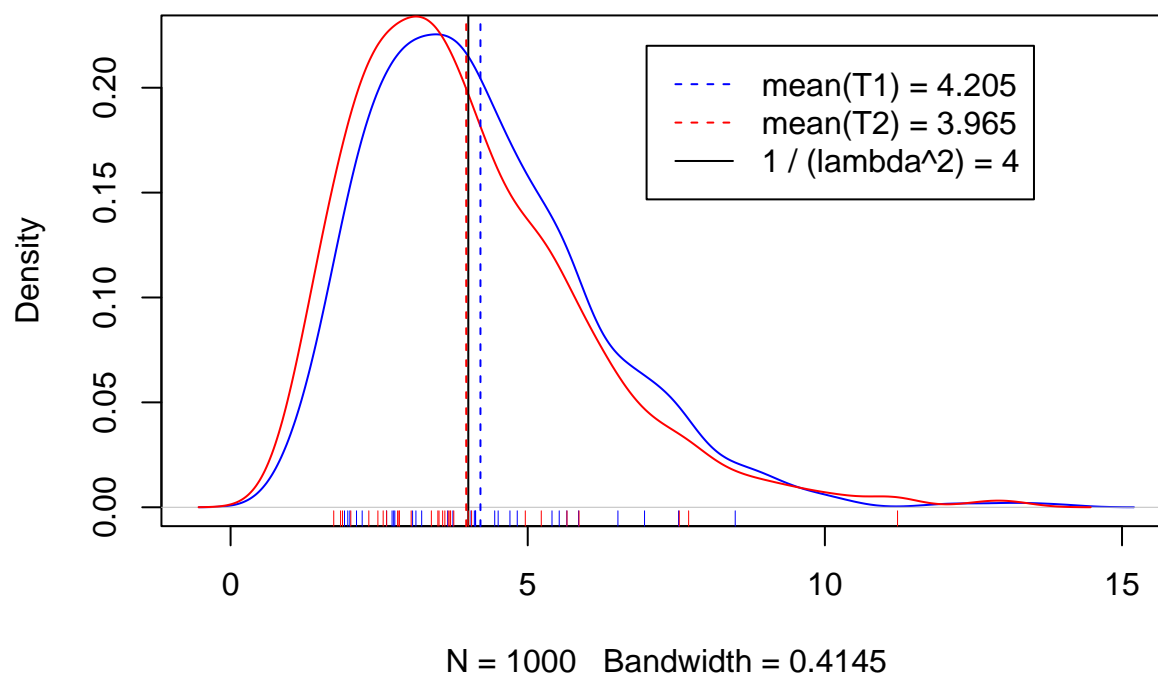
Elijo la ventana usando uno de las ventanas “óptimas” que ofrece R, la de Silverman (1986).

```
dens_1 <- density(reps_t1, "nrd0")
dens_2 <- density(reps_t2, "nrd0")
```

Graficar en un mismo plot los estimadores de la densidad de T1 y T2 obtenidos en distinto color.

```
plot(dens_1, col="blue", main="Densidades estimadas para T1 y T2")
lines(dens_2, col="red")
abline(v=lambda^-2)
abline(v=mean(reps_t1), col="blue", lty=2)
abline(v=mean(reps_t2), col="red", lty=2)
legend(7, 0.22, legend=etiquetas, col=c("blue", "red", "black"), lty=c(2, 2, 1))
rug(reps_t1[1:30], col="blue")
rug(reps_t2[1:30], col="red")
```

Densidades estimadas para T1 y T2



Comparar e interpretar el gráfico. Relacionar con 3c.

Con una elección adecuada de h , las dos densidades parecen tener la misma forma, pero la densidad de T1, que en 3c consideramos sesgada, está efectivamente desplazada hacia la derecha.