Alumno: Luis Donaldo Romero Tapia

***Tarea\_02\_seminario\_estadistica\_1***

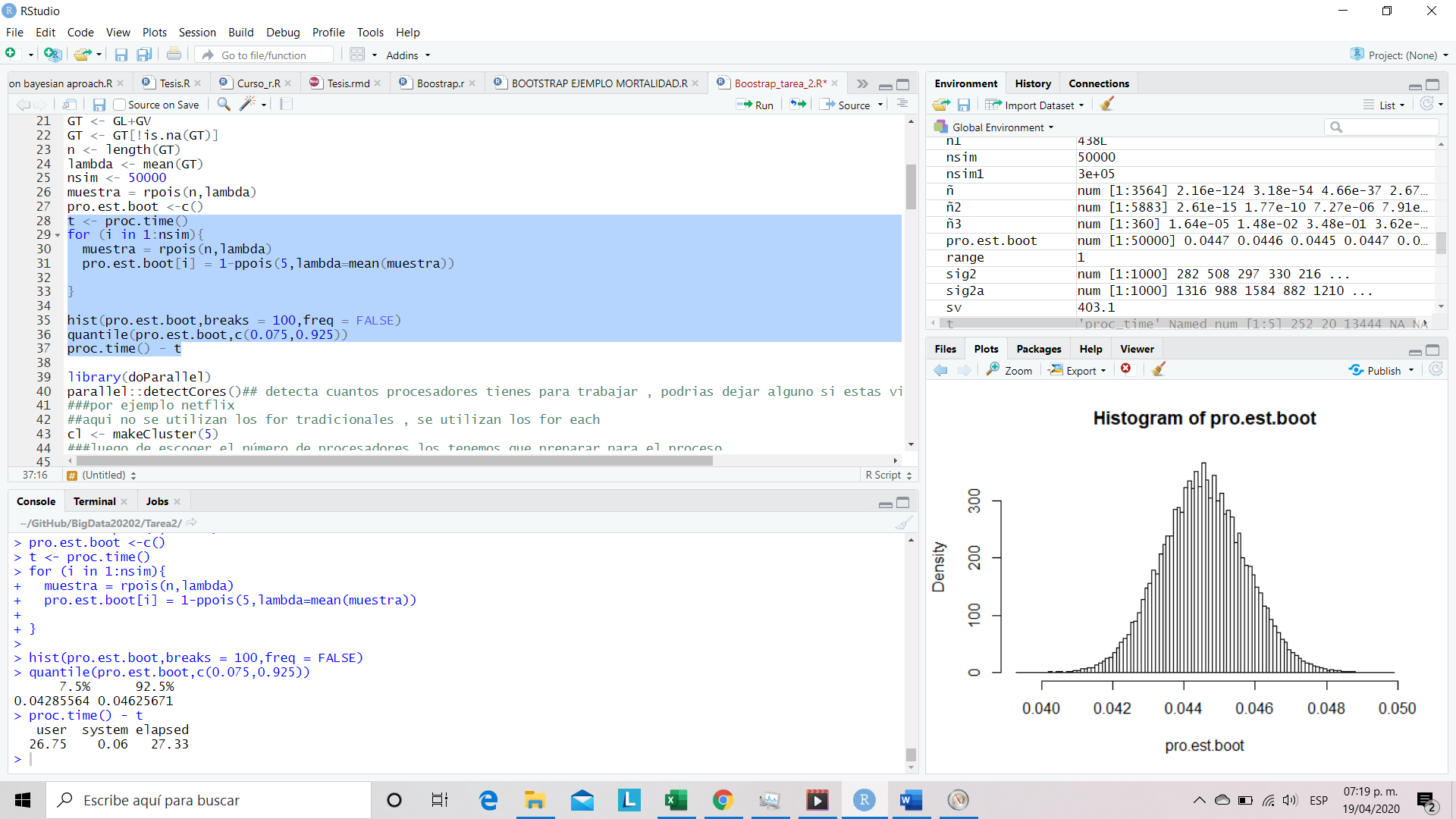
**Nota: El código que se corre esta subrayado en azul y la respuesta se muestra en la consola**

1. La base Fubol.csv contiene la información de n = 8698 partidos del futbol español donde la columna GL se refiere a lo número de goles que anotó el equipo local mientras que la columna GV corresponde a los goles anotados por el equipo visitante.

Con la tabla anterior, construya las variables X := goles anotados en el partido y la variable Z := diferencia absoluta de goles entre los dos equipos.

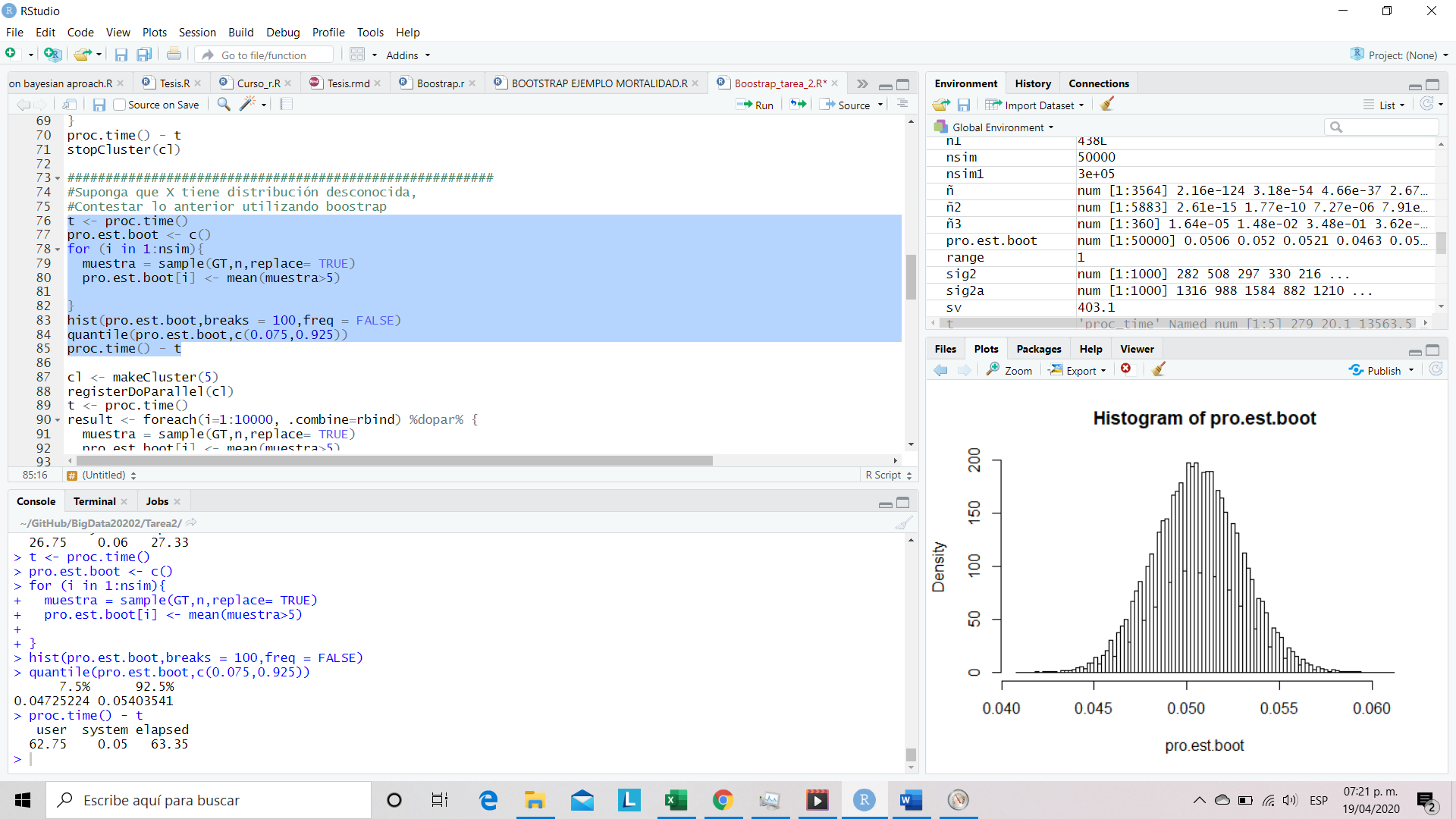
Con esta información realice lo siguiente:

* Asumiendo que X ∼ P ois(λ) y utilizando la técnica del Bootstrap paramétrico, encuentre un intervalo al 85 % de confianza para P (X > 5), es decir encuentre un intervalo de confianza al 85 % para la probabilidad de que un partido en el futbol español supere los 5 goles. (Realice al menos 50, 000 simulaciones bootstrap)



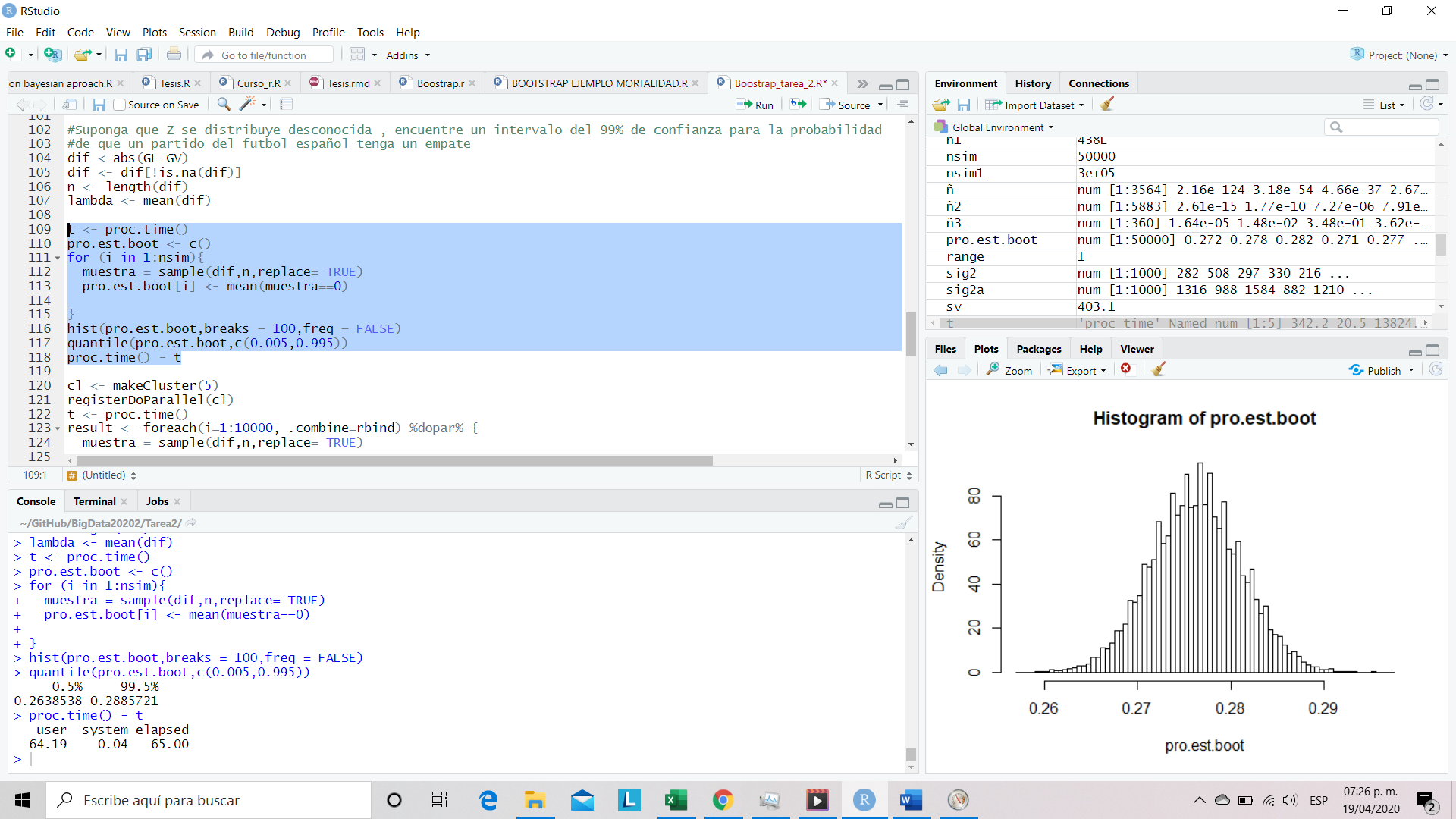
En este caso el intervalo del 85% de confianza es (0.04285564,0.4625671). Con un tiempo sin paralelizar de 27.33 s.

* Suponga ahora que X ∼ FX(x) con FX(x) una distribución desconocida, conteste la pregunta anterior utilizando la técnica de bootstrap no paramétrico. (Realice al menos 50, 000 simulaciones bootstrap).



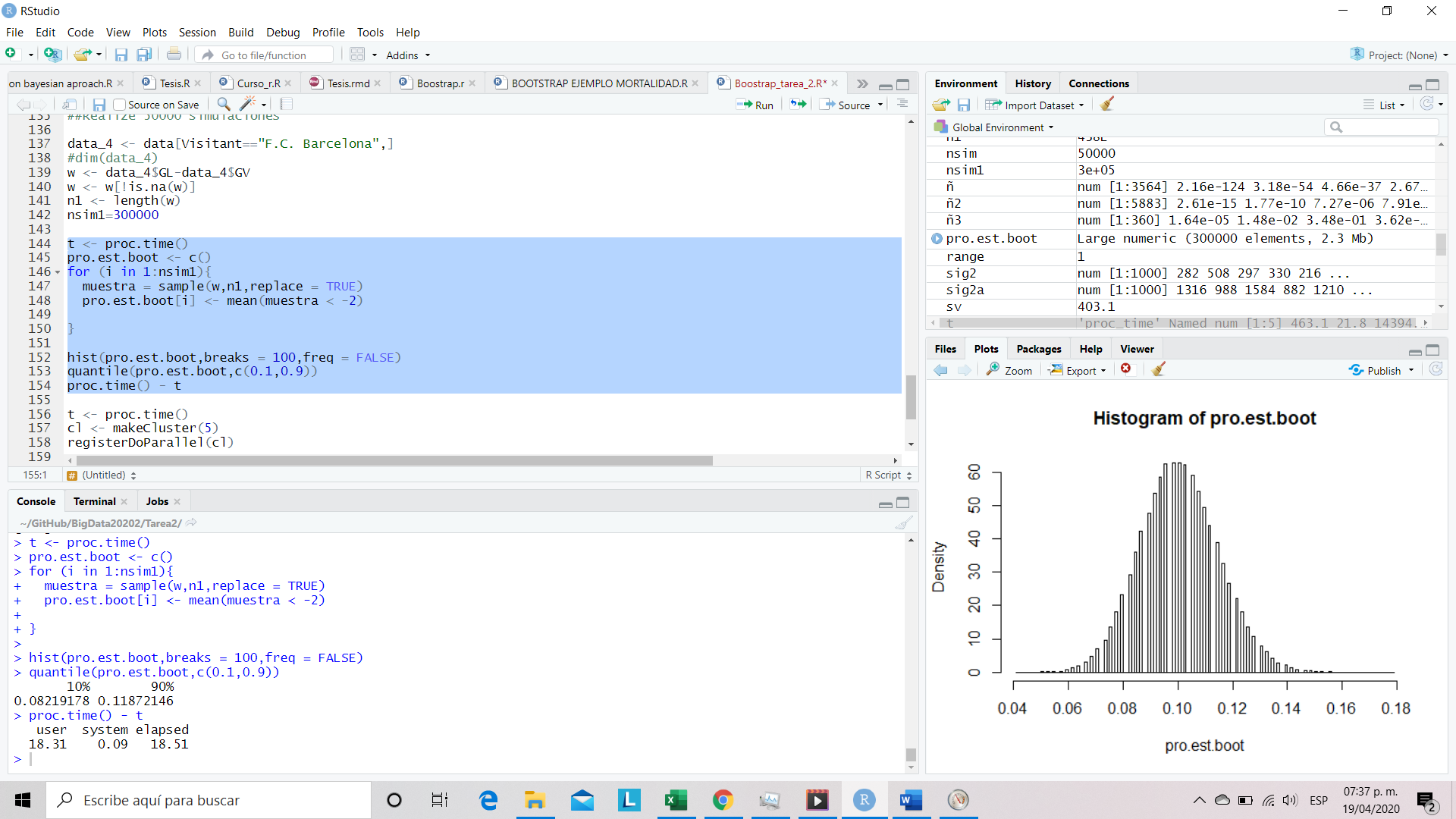
En este caso dio como resultado el intervalo del 85% de (0.04725224.0.05403541) y se tardo un tiempo de 63.55 s

* Suponga que Z ∼ FZ(z) con FZ(z) una distribución desconocida. Encuentre un intervalo al 99 % confianza para la probabilidad de que un partido del futbol español tenga un empate. (Realice al menos 50, 000 simulaciones bootstrap).



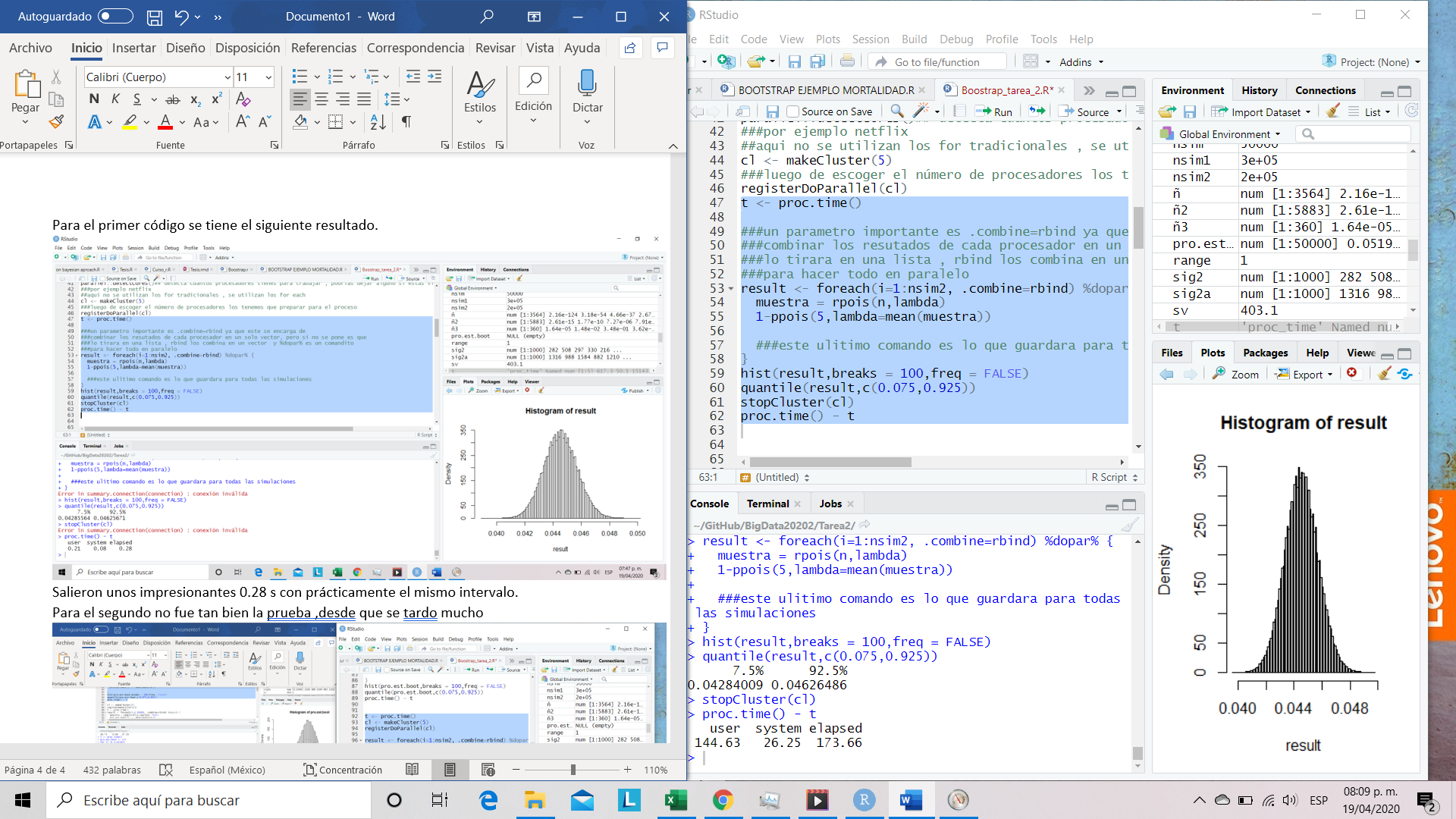
En este caso se obtuvo un intervalo de confianza (0.2638538,0.2885721) del 99% de confiabilidad y se tardo 65 s.

* Asumiendo que W es la variable que mide la deferencia de goles entre el equipo local y el visitante W := GL − GV y que W ∼ FW (w) con FW (w) una distribución desconocida; encuentre un intervalo al 80 % de confianza para la probabilidad de que el equipo F.C. Barcelona gane por mas de 2 goles como visitante.(Realice al menos 50, 000 simulaciones bootstrap).



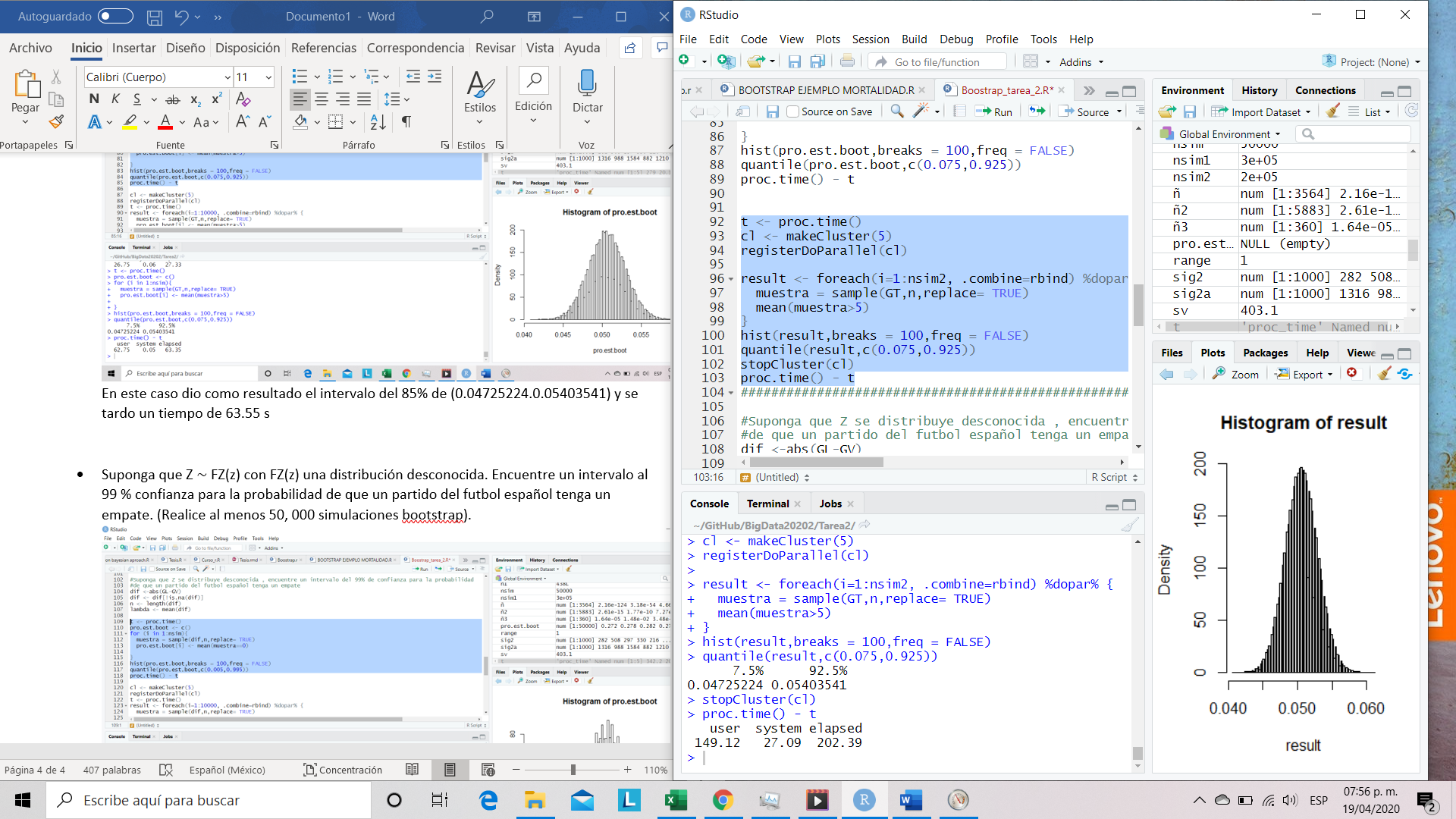
En este caso se obtuvo un intervalo de confianza (0.08219178,0.11872146) y se tardo 18.51 s.

* Repita los problemas anteriores pero paralelizando su codigo (Raelice al menos 200, 000 simulaciones bootstrap). Se hace con 5 cores y 200000

Para el primer código se tiene el siguiente resultado. 

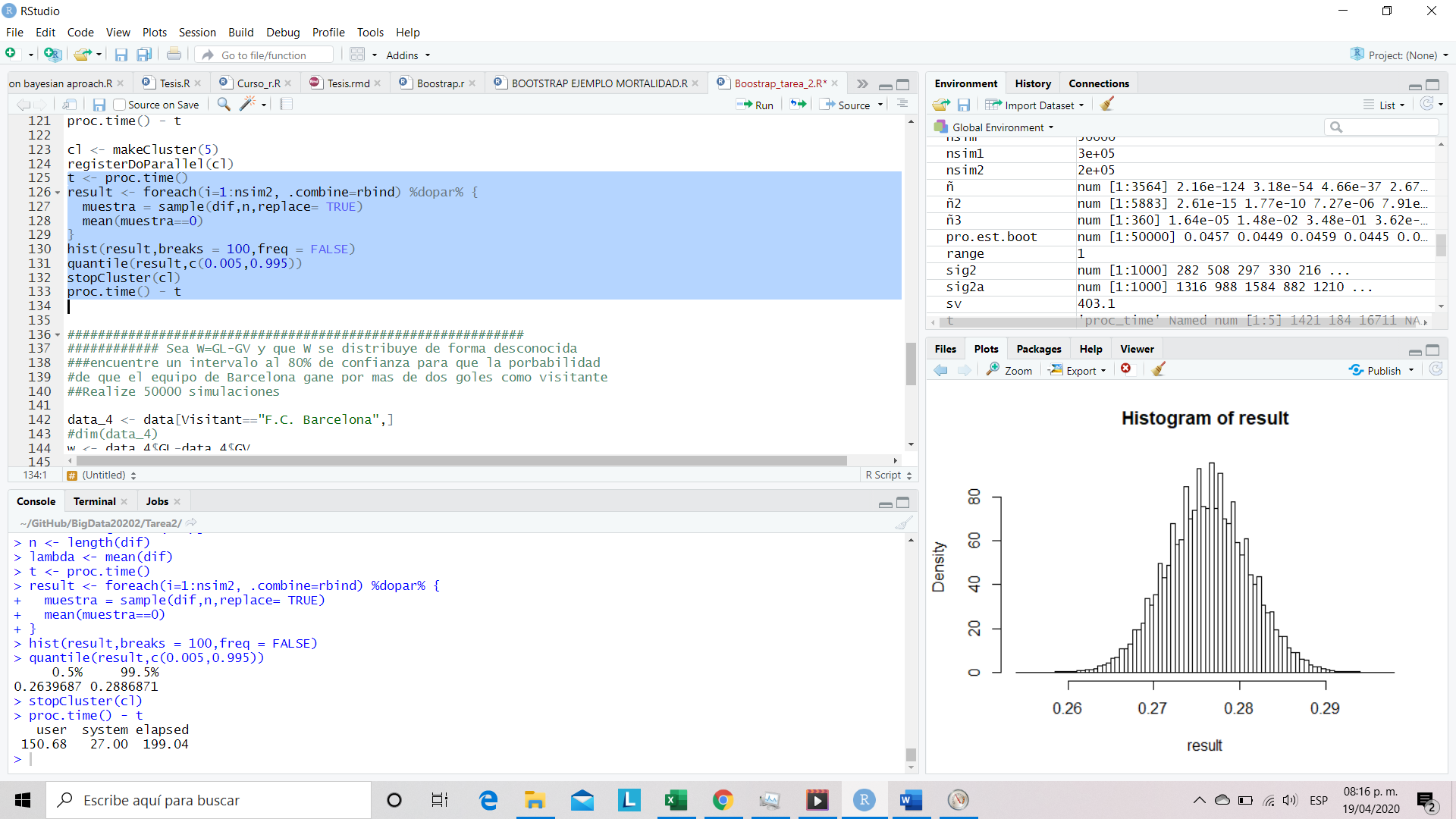
Salieron unos impresionantes 173.66s con prácticamente el mismo intervalo , contra los 27.33 s que se tardo con 50000 .

Para el segundo no fue tan bien la prueba ,desde que se tardo mucho



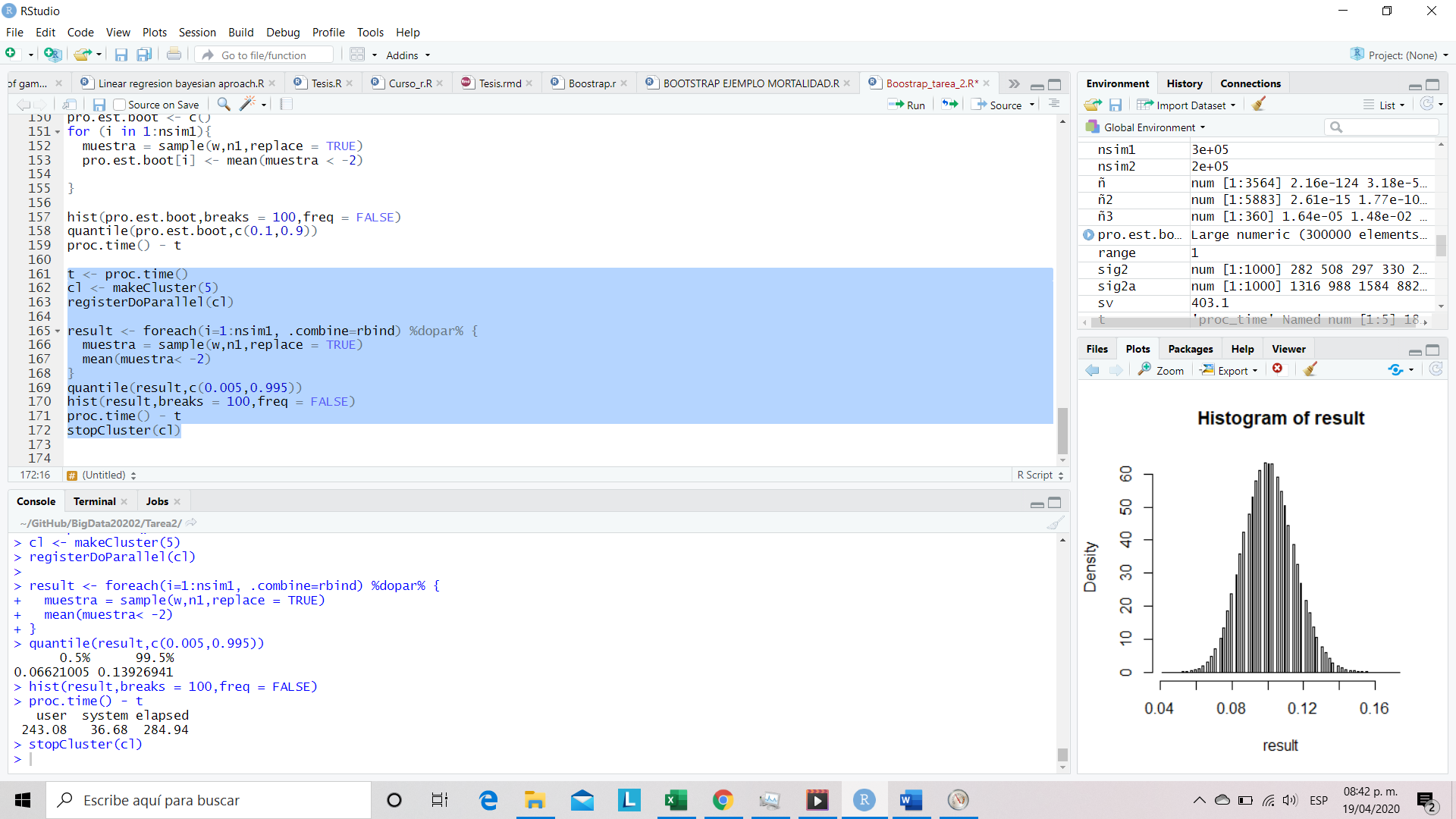
202.39 s contra los 63.55 s. que se había tardado anteriormente .

Para el tercer problema se tardo lo siguiente



199 s y con prácticamente el mismo intervalo , en comparación a los 65 s. que se tardo.

Por ultimo tenemos que



.Se tardo 284.94

Al final analice un poco el por que estos resultados y en algunos casos la computadora es mejor sin cluster cuando hay pocas iteraciones , cuando compare los anteriores ejercicios haciendo 200000 el tiempo mejoro en el cluster .