

Practica 3

Jose David Martinez Ruiz

August 28, 2017

1 Introducción

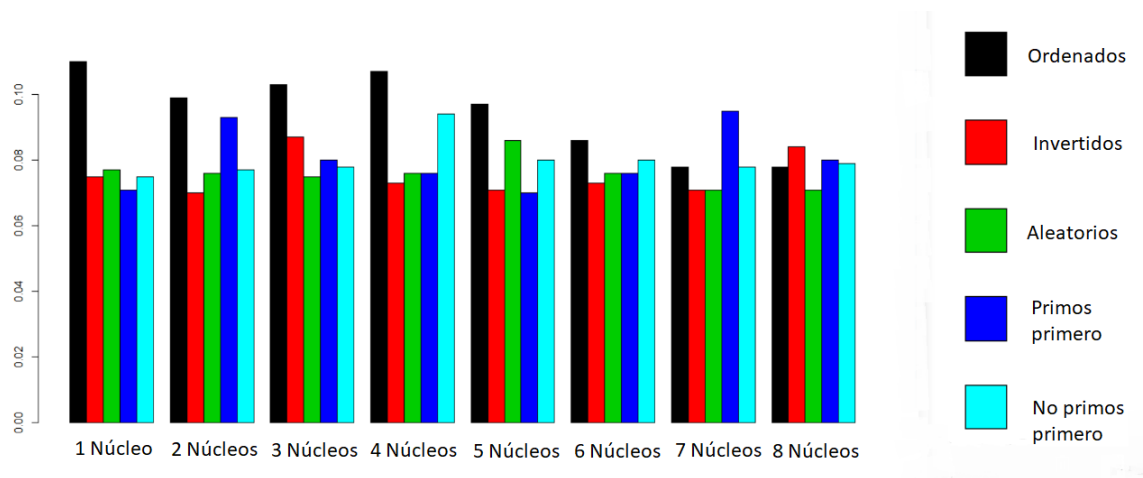
En esta práctica se calcularon los tiempos que le toma a una función determinar si un número es primo o no mediante algunas operaciones. Se realizó esta operación en paralelo para algunas secuencias de números, todas diferentes y en un mismo rango de números. Lo que se desea saber es como varían los tiempos de espera para cada ordenamiento cuando se variaban los núcleos que se le daba a la operación. Se especuló sobre las razones de éstas variaciones y se realizó una prueba estadística para corroborar las especulaciones.

2 Desarrollo

Se modificó el código proporcionado de modo que fuera variando la cantidad de núcleos dadas a las operaciones. Los núcleos fueron variando desde 1 núcleo hasta 8 núcleos. Se agregaron 2 ordenamientos más a los ya proporcionados, para poder tener más variedad a la hora de obtener los datos. Uno de estos ordenamientos era colocar todos primos que hay en el rango de números donde se trabajó al principio y todos los que no eran primos al final, el otro ordenamiento era algo similar al anterior, sólo que éste tenía todos los no primos del rango al principio y los que sí eran primos al final. Para hacer esto se utilizó la función "primo" que se nos fue proporcionada para saber cuáles números de mi rango eran primos y cuáles no.

Se hicieron 10 réplicas de los tiempos para cada ordenamiento. Los tiempos obtenidos para cada ordenamiento fueron guardados en una matriz. Al final esta matriz tenía 10 tiempos obtenidos para cada tipo de observación para cada uno de los núcleos disponibles.

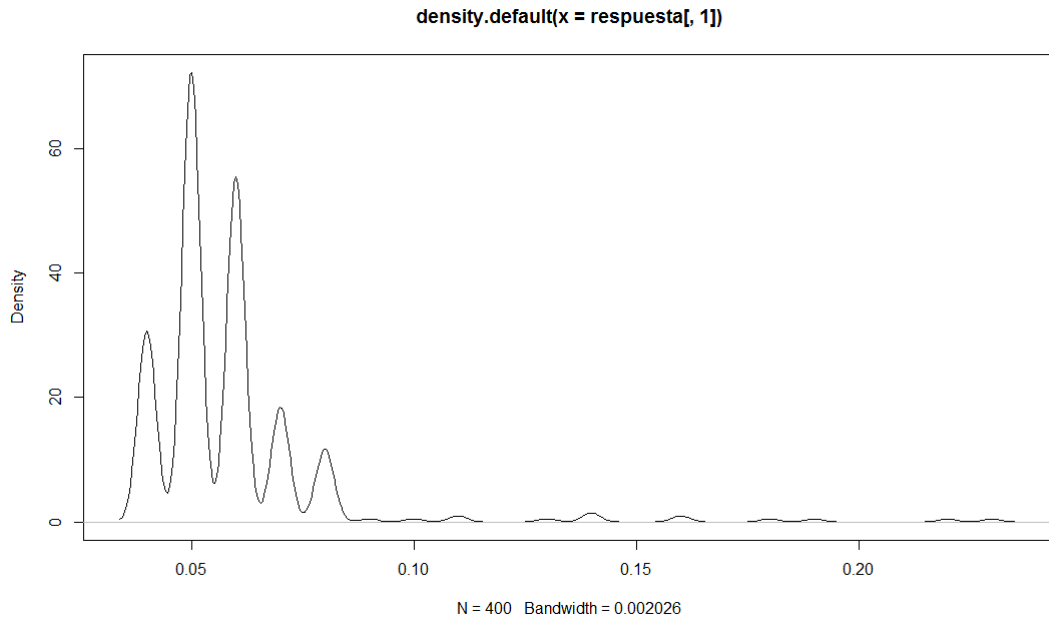
Una vez terminada la matriz, se procedió a sacar el tiempo promedio de las 10 réplicas de cada ordenamiento por cada núcleo. Estos resultados se muestran en la siguiente grafica de barras.



Como se puede observar en la gráfica hubo muchas variaciones de un ordenamiento a otro y de un núcleo a otro, en general parece que mientras más núcleos haya los tiempos promedio son menores, en otros casos parece lo contrario. Sin embargo, esto puede deberse a la misma computadora donde se ejecutó el programa. los ordenamientos reaccionaron de diferentes maneras

dependiendo de la cantidad de núcleos que tenían a su disposición, algunos se tardaban más con menos núcleos, otros con mas, esto puede deberse a que algunos ordenamientos requerían de un esfuerzo extra para realizarse, mientras que otros que no necesitaban tal cantidad de esfuerzo se hacían más lento por los núcleos que tenían asignado.

Nos interesa saber si hay una diferencia significativa entre los tipos de ordenamientos y la cantidad de núcleos utilizados, para hacer eso primero se verificó si la distribución de los datos era normal y así proceder a hacer el diseño de experimentos y saber si se usaran métodos paramétricos o no paramétricos. La gráfica de la densidad de los datos es la siguiente.



Como se puede ver, la gráfica de la densidad de los datos no se ajusta a la campana de la densidad normal, por lo que se utilizaran métodos no paramétricos para éste diseño de experimentos.

Para hacer el diseño de experimentos primero se colocaron todos los datos obtenidos de los tiempos de espera en una matriz con su correspondiente tipo de ordenamiento y el número de núcleos que fueron utilizados. una vez que la matriz fue terminada se procedió a hacer la prueba estadística de Kruskal-Wallis.

Primero se quiso verificar si el tipo de ordenamiento tenía un efecto significativo en los tiempos de espera. al realizar la prueba obtuvimos un p-valor mayor al 5 por ciento por lo que no hay una diferencia significativa entre los ordenamientos.

Después se verificó si hay una diferencia significativa entre la cantidad utilizada de núcleos a la hora de obtener los tiempos de espera. al realizar la prueba de Kruskal-Wallis se obtuvo un p-valor menos al 5 por ciento por lo que se concluye que si hay una diferencia significativa entre los núcleos.

3 Conclusiones

Hubo datos y resultados muy variados a la hora de analizar los tiempos de espera. Debido a las altas y bajas en el tiempo promedio de espera de los ordenamientos se concluye que se debió a la presencia de datos atípicos a la hora de sacar los tiempos. En cada ordenamiento el primer dato obtenido era particularmente más grande que los demás.

El hecho de que los ordenamientos no mostraran una diferencia significativa a la hora de hacer la prueba principalmente a la diferencia tan mínima que hay entre los tiempos promedios, los datos que se obtenían eran muy similares entre sí, esto último puede deberse al esfuerzo computacional

que se hizo para obtener los datos homogeneizó los tiempos de espera por núcleo. la diferencia obtenida entre los núcleos era una diferencia que se esperaba, debido a la carga que la computadora tiene a la hora de hacer el experimento, puede beneficiarse o perjudicarse de la cantidad de núcleos que tiene disponible.

En conclusión, se esperaba que hubiera diferencia tanto en los ordenamientos como en la cantidad de núcleos, pero debido a cómo la computadora funciona y procesa las cosas, las diferencia entre los ordenamientos no se hicieron notar.