

Caso de Estudio R
95.07 Teoria de Lenguajes de Programación
Primer Cuatrimestre 2020

Federico Andres Ramos 101640

R para calculo de probabilidad de mayoria en las elecciones de senadores en USA

The New York Times creo un modelo estadístico llamado LEO para analizar la probabilidad de que el partido Demócrata o Republicano gane la mayoría en el senado en las elecciones de USA utilizando R.

Leo interpreta las encuestas entre dos candidatos creando distintas transformaciones para hacer dos encuestas distintas comparables entre si. Dependiendo de quien realice la encuesta se le hace ciertos ajustes para ser comparable, al igual que se hace distintas transformaciones a los resultados de la encuesta si el encuestado es un posible votante o un votante registrado. Por ejemplo, si Public Policy Polling, una firma que tiende a ser democrática, y Harper Polling, una firma que tiende a ser republicana, hacen una encuesta el mismo día a las mismas personas, es esperable que los resultados de encuestas de Public Policy Polling sean 1 o 2 puntos más favorables para los demócratas. A esta tendencia se le suele llamar "House Effects" y es la tendencia que tiene una firma a favorecer a un candidato democrático o republicano. Esta tendencias entre otras son necesarias a la hora de comparar los datos para que LEO tenga predicciones más precisas.

Cuando LEO tiene varias encuestas y debe combinarlas, le asigna un peso a los resultados de la encuesta dependiendo de la cantidad y variedad de encuestados además de la fecha en la que se realizo (una encuesta hecha un día antes de las elecciones tiene más importancia que una hecha un mes antes de las elecciones).

No solo importan las encuestas para el LEO, sino que tambien tiene peso variables como el porcentaje de aprobación de cada uno de los candidatos, además de su experiencia política, cantidad de plata recaudada, los resultados de las elecciones presidenciales en cada estado, etc. El peso que se le asigna a cada una de estas variables tambien depende de los datos obtenidos de las elecciones de los ultimos 20 años.

La relación entre el peso que tiene los resultados de encuestas y las variables mencionadas depende de la cantidad de encuestas hechas, además de la fecha. Si faltan 2 meses para las elecciones, el modelo le otorga más peso a variables como experiencia política o porcentaje de aprobación, mientras que en los días cercanos a las elecciones, se le da más peso a los resultados de las encuestas hechas recientemente.

Finalmente, para conseguir la predicción total de todos los asientos del senado, tambien se tiene en cuenta factores externos, como por ejemplo, un boom en la economia mientras el presidente electo es democrático puede beneficiar a los candidatos democráticos.