

Técnicas de Análisis Cuantitativas y Cualitativas

Resolución del ejercicio de evaluación 1

Marcos Rial Docampo

22 de octubre de 2015

En el ejercicio se presentan dos tipos de resultados de forma gráfica aparecidos en prensa. Los primeros resultados son los referentes a la información extraída del resultado de las elecciones municipales de Madrid del 24 de mayo de 2015 en la que se muestra en forma de gráfico el porcentaje de votos recibidos por las diferentes formaciones políticas. Los segundos resultados son extraídos de una encuesta de estimación de voto en el municipio de Madrid hecha entre los días 4 y 10 de septiembre de 2015.

El primer error, en lo referente a la presentación de los datos, es la de no incluir en los datos de la encuesta los intervalos de confianza. En el primer caso no sería correcto incluir el intervalo de confianza puesto que se trata de datos obtenidos de unos resultados electorales donde se convocaba a participar al total de la población con derecho a voto (censo electoral) del municipio de Madrid. No se trata de una encuesta, como en el segundo caso, donde el tamaño muestral es de 502 personas consultadas.

Primeramente, podemos tomar el ejercicio como un conjunto de cuatro ejercicios individuales en los que cada uno trata un modelo discreto distinto bajo una distribución binomial. Por ejemplo, tomar para el caso del PP que un 33.7 % de la población lo votaría frente al 66.7 % que no lo haría. Tenemos así por tanto un éxito de 0.337.

Para completar los datos de límites superior e inferior del intervalo de confianza mostrados en el cuadro 1 utilizamos el comando *prop.test()* de R como se especifica en la figura 1. En él simplemente necesitamos especificar el número de elementos de éxito en la muestra frente al tamaño de la misma así como el nivel de confianza (en este caso del 95 %).

Se indica en la noticia que el error de muestreo es de $\pm 4,5\%$. Esto quiere decir que la diferencia entre los resultados de la encuesta y los reales preguntando a toda la población serían de un 4,5 %. Una forma de reducir este error podría ser la de ampliar el tamaño de la muestra. Pero es un valor que no debería generalizarse a toda la encuesta puesto que tenemos resultados porcentuales de lo que se espera obtener en la encuesta para cada uno de los partidos políticos (reflejados en los resultados electorales de mayo). Es decir que cada dato para cada partido político obtendría su propio dato de error de muestreo.

Un error muestral del $\pm 4,5\%$ se antoja bastante alto para afirmar con rotundidad, como dice el titular del artículo, que la candidata del Partido Popular (PP) recuperaría la alcaldía de haber un pacto con Ciudadanos (C's). Si a esto añadimos el hecho de

Partido	Elecciones (mayo)	Encuesta (sept.)		
	Valor	Valor	LI	LS
PP	34,6 %	33,7 %	29,6	38,0
Cs	11,4 %	13,8 %	10,9	17,1
AM	31,9 %	29,1 %	25,2	33,3
PSOE	15,3 %	17,4 %	14,2	21,0

LI y LS límites inferior y superior del intervalo de confianza al 95 %

Cuadro 1: Resumen de los resultados de las elecciones de mayo y la encuesta de septiembre.

```

1 #Aplicacion de prop.test para definir los intervalos de confianza con n
  = 502 y nivel de confianza del 95%
2 #PP con 33.7% de los votos
3 prop.test(round(0.337*502, digits = 0),502,
4           conf.level = 0.95)
5
6 #Cs con 13.8% de los votos
7 prop.test(round(0.138*502, digits = 0),502,
8           conf.level = 0.95)
9
10 #AM con 29.1% de los votos
11 prop.test(round(0.291*502, digits = 0),502,
12           conf.level = 0.95)
13
14 #PSOE con 17.4% de los votos
15 prop.test(round(0.174*502, digits = 0),502,
16           conf.level = 0.95)

```

Figura 1: Script en R con la aplicación de prop.test a los datos de la encuesta. Se aplica redondeo para ofrecer un número entero de votantes por cada partido.

que los intervalos de confianza son bastante amplios refuerza la incorrecta interpretación dada por el periodista.

Para asegurar lo dicho en el párrafo anterior podemos tomar los votos conjuntos a PP y C's para evaluarlos separadamente.