

Interrogación 3

EYP2417 - Muestreo

Camila Echeverría

Francisca Vilca

Análisis de la elección presidencial del año 2021

En el siguiente informe, se darán a conocer algunos resultados interesantes, con respecto a las elecciones presidenciales del año 2021, se analizará al votante de 1ra Vuelta con respecto a sus características como el Sexo, Edad y GSE, además se intentarán proyectar los resultados de 2da Vuelta, para ellos se usarán diferentes técnicas de muestreo y de balanceo para que los resultados sean lo más verosímiles a la realidad.

Diseño Muestral

En este caso se realizará un muestreo de conglomerados en dos etapas, estratificadas por zona (considera solo viviendas urbanas – censo 2017). Primeramente, se obtiene una muestra aleatoria simple de manzanas al interior de cada estrato y de cada manzana se obtiene una muestra aleatoria de viviendas. Al interior de cada vivienda se selecciona un elector según la técnica del “cumpleaños”. No se realiza actualización del número de viviendas de la manzana seleccionada, de aquí son seleccionadas 1007 personas de un total de 1.520084×10^7 .

Probabilidades de Selección y Factores de Expansión según Zona, Sexo y cat-Edad de acuerdo al diseño Muestral

Por un lado si se usa la información del censo realizado el 2017, es posible obtener la información de las manzanas y las viviendas totales para calcular las probabilidades de selección, por otro lado, si se usan las estimaciones y proyecciones de la población que se tienen para el 2021, es posible conocer la cantidad de personas por sexo y edad que habrá en Chile y usando la tabla 1 de las referencias, se puede conocer los GSE que existen al 2021.

Con toda esta información se tiene que las probabilidades de selección y factores de Expansión serán:

zona	sexo	cat.edad	Prob de Selección	Factores de Expansión
1	1	1	0.000133	7537.188
1	1	2	0.000086	11693.560
1	1	3	0.000076	13141.294
1	1	4	0.000045	22064.875

zona	sexo	cat.edad	Prob de Selección	Factores de Expansión
1	2	1	0.000091	11017.476
1	2	2	0.000094	10617.741
1	2	3	0.000080	12523.444
1	2	4	0.000043	23180.333
2	1	1	0.000086	11684.565
2	1	2	0.000037	27304.125
2	1	3	0.000051	19460.613
2	1	4	0.000058	17209.000
2	2	1	0.000095	10549.260
2	2	2	0.000084	11966.875
2	2	3	0.000055	18050.229
2	2	4	0.000062	16015.884
3	1	1	0.000197	5063.837
3	1	2	0.000065	15467.056
3	1	3	0.000062	16182.062
3	1	4	0.000058	17291.462
3	2	1	0.000069	14558.000
3	2	2	0.000074	13562.286
3	2	3	0.000078	12845.600
3	2	4	0.000053	18807.214
4	1	1	0.000070	14200.684
4	1	2	0.000037	27337.026
4	1	3	0.000044	22727.969
4	1	4	0.000045	22166.192
4	2	1	0.000068	14601.566
4	2	2	0.000076	13192.851
4	2	3	0.000072	13852.982
4	2	4	0.000048	20705.676

Cruce de votos 2da Vuelta vs 1ra Vuelta de la muestra sin ponderar

Para este caso, en general solo es de interés ver la redistribución de los votos tal y como es lo expresado por los ciudadanos de la muestra, los resultados pueden ser apreciados en la Tabla 2:

Boric Kast Provoste Sichel Ártes Enriquez Parisi No vote Blanco /Nulo No sabe
:—: :—: :—: :—: :—: :—: :—: :—: :—: :—: :—: :—: :—: :—: Boric 92.683
2.454 70.213 20.000 78.571 71.111 50.943 31.098 25 20.833 Kast 1.951 93.252 4.255
63.636 0.000 6.667 16.038 20.732 5 12.500 Nulo 0.976 0.613 8.511 1.818 7.143 2.222
12.264 8.232 25 8.333 No Votaré 0.000 0.613 2.128 3.636 7.143 4.444 4.717 4.573 0
0.000 No sabe 4.390 3.067 14.894 10.909 7.143 15.556 16.038 35.366 45 58.333 Total
205.000 163.000 47.000 55.000 14.000 45.000 106.000 328.000 20 24.000

Cruce de votos 2da Vuelta vs 1ra Vuelta de la muestra con ponderación

En la misma línea del ítem anterior, ahora se hará el cruce de votos pero haciendo una ponderación según la zona en la que habita la persona, y es hombre o mujer y en qué categoría se encuentra según su edad. Para poder lograr esto, debemos calcular las probabilidades de selección para cada uno de estos casos, y así podremos obtener los factores de expansión w_{ij} , cuyo procedimiento y detalle podemos encontrar en el script. Así, la distribución de los votos es la siguiente:

	Boric	Kast	Provoste	Sichel	Ártes	Enriquez	Parisi	No vote	Blanco /Nulo	No sabe
2da Vuelta	93.143	93.662	68.336	0.000	6.503	18.285	21.053	4.383	10.181	0.912
1ra Vuelta	93.143	93.662	68.336	0.000	6.503	18.285	21.053	4.383	10.181	0.912
Total	205.000	163.000	47.000	55.000	14.000	45.000	106.000	328.000	20.000	24.000

Cruce de votos con g-weights

- Obtenga las respectivas probabilidades de selección y los factores de expansión según Zona, sexo y cat.Edad de acuerdo con el diseño muestral. Contraste los resultados obtenidos con la información censal proyectada al año 2021. Para esta base, muestre el cruce voto 2da vuelta (P2v) vs 1era vuelta (P1v) ponderado y sin ponderar – en términos de porcentaje tal que permita leer la re-distribución de los votos. Comente los aspectos más llamativos, por ej. ¿Dónde van los votos de Parisi? (un párrafo). **(1.0 punto)**
- Para la base ponderada, obtenga los g-weights (%pob / %muestral) según zona y sexo-cat.edad. Para esta base, muestre el cruce voto 2da vuelta (P2v) vs 1era vuelta (P1v) – en términos de porcentaje tal que permita leer la re-distribución de los votos. Comente los aspectos más llamativos, por ej. ¿Dónde van los votos de Parisi? (un párrafo). **(1.0 punto)**
- Para la base sin ponderar, usando la distribución GSE según zona, más la proyección censal realice un rastrillado (raking) según **zona – sexo – cat.edad – GSE** (total) que le permita obtener una nueva ponderación. Reproduzca los cruces de b) y contraste con lo obtenido previamente y comente (un párrafo). **(1.0 punto)**
- Literalmente, hasta el momento se trabaja con el supuesto de cómo se comportarán los poco más de 15 millones de chilenos. Sin embargo, hubo una abstención importante y se pueden ver discrepancias entre los resultados de la elección en 1era vuelta y los valores obtenidos en la encuesta (P1v). Realice un rastrillado que considere estos resultados. Es decir, lleve a cabo un “raking” según **zona – sexo – cat.edad – Voto** (total ver tabla). Reproduzca los cruces de b) y contraste con lo obtenido previamente y comente (un párrafo). **(1.0 punto)**
- Considerando todos los resultados previos, realice predicciones para Total Poblacional (15.200.840), de tal manera que muestre el número de votos de cada candidato bajo cada ponderador (a, b, c y d). Comente (un párrafo). **(2.0 puntos)**