

Censo de derecho. Comentario al conteo de población

Iván Williams

Enero 2019

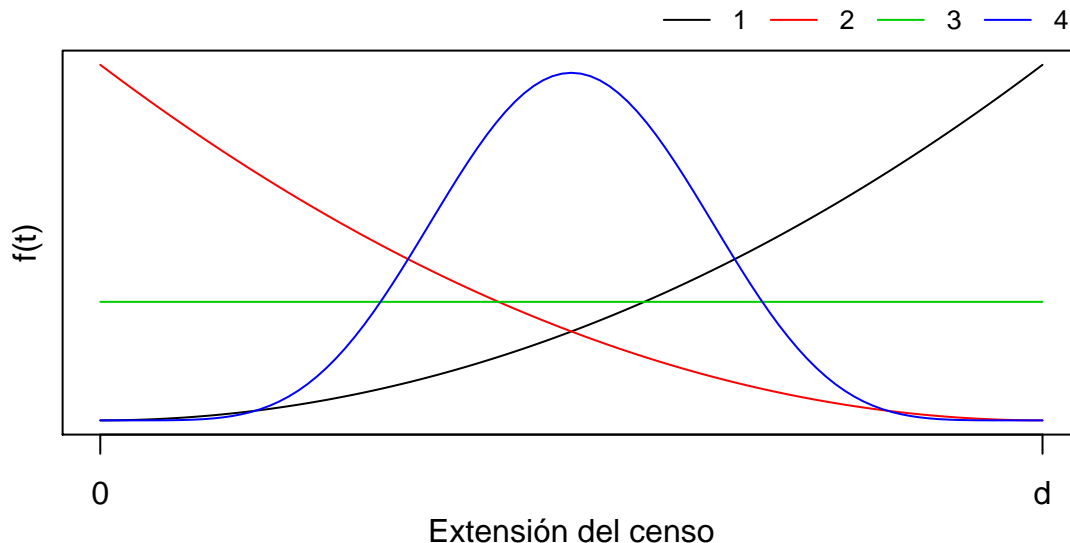
Si nos solicitan medir la superficie de un globo mientras otra persona lo esta inflando, ¿qué es lo que medimos? Si compro el mismo helado en Finlandia y en Río de Janeiro, ¿tomo la misma cantidad de helado? ¿Cuántas personas residen en una ciudad cualquiera, si en medio de un operativo censal cambia la intención de continuar residiendo de algunos de sus habitantes (y quizás se haga efectiva antes de que algun censista visite el hogar)? En todos estos casos hay un fenómeno dinámico (un globo que crece, un helado que se derrite y una población que cambia) y una necesidad de medirlo y resumir su dimensión en un punto temporal. Un censo de derecho implica un operativo que se extiende en una ventana de tiempo y territorio predefinidos. Se caracteriza principalmente por su definición de residencia (con su regla y múltiples excepciones) y por ser un conteo de “película” (más que una “foto”).

En estos párrafos daremos una primer aproximación a la implicancia de este carácter dinámico en el resultado más importante: ¿Cuántos somos?. Lo referido a la distribución de atributos (sexo, edad, residencia anterior, etc.) es más complejo, y requiere más profundidad de la que aquí se trata. Preguntas adicionales interesantes son: ¿cuán sensible es el conteo a las formas posibles de dinámica y relevamiento?, ¿dónde es conveniente colocar en el tiempo la estimación final? ¿Qué fuentes de variabilidad tiene? Aunque desde un punto de vista distinto, hace más de 50 años Keyfitz (1966) quiso resaltar el carácter estocástico de un censo cuando sugirió que podría considerarse como una realización de todas las poblaciones en el tiempo en donde las mismas condiciones demográficas prevalecieron (y prevalecerán, si se nos permite la adición).

Sin pretender abarcar la complejidad empírica del asunto, empecemos por suponer un censo que se prolonga $d \in (0, 1)$ años (menos de 12 meses), en una población distribuida homogéneamente a lo largo de un territorio accesible en todos sus puntos. Definimos una función de relevamiento $R_d \sim f(t)$ en $d \geq t \geq 0$, la cual indica de qué forma se cubre el territorio durante el censo, cumpliendo que $\int_0^d f(t)dt = 1$ (se recorre completamente).

Inicialmente podemos considerar cuatro tipos de relevamiento según en qué momento se concentran los mayores esfuerzos:

Formas teóricas de relevamiento



Las funciones fueron pensadas con una distribución beta truncada en d , con valores 3, 1, 1 y 7 para α y 1, 3, 1 y 7 para β , en ese orden.

De esta manera, si la población a cada momento es N_t (función de t), el “conteo” censal puede ser considerado una esperanza matemática:

$$E(P) = \int_0^d N_t f(t) dt.$$

Donde se pondera la dinámica poblacional según la probabilidad (densidad) de que cada unidad (persona, hogar, vivienda, barrio, o ciudad, la escala es indiferente en esta instancia) sea seleccionada a cada momento t , y la cual es una porción de la exposición del período $P^* = \int_0^d N_t dt$. La varianza de P (variable aleatoria de “conteo”) se puede obtener por diferencia de momentos $E(P^2) - E(P)^2$. En este punto cabe aclarar que aquí consideramos a los problemas de declaración de edad u omisión como sesgos (y no varianza).

La segunda cuestión relevante es dónde colocar la estimación final. La respuesta adecuada pareciera ser en aquel valor t^* ($d \geq t^* \geq 0$) tal que $E(P) = N_{t^*}$, que podemos obtener teóricamente como $t^* = \ln(E(P)/N_0)/r$. En la práctica no contamos con N_0 (población al inicio del relevamiento), pero puede considerarse un conteo intercensal o el censo anterior con el fin de aproximarse.

Para mencionar un caso real, el censo de derecho de 2005 en Colombia se planificó con $d = 0.8$ (poco más de 9 meses), para el cual se agruparon áreas menores según similitud de dinámica demográfica, con relevamiento simultáneo al interior de cada una, concentrando en el mes central el 65% de la población total que se caracterizaba por una dinámica menos estable, y un 30% en los últimos 39 días (DANE 2006, pág. 40, y 2008, pág. 23). Por otro lado, Uruguay realizó un censo de derecho en 2 meses durante 2011, y colocó la estimación final en aquel momento donde el relevamiento acumuló el 50% del conteo final (INE de Uruguay, s/f). En nuestro ejercicio teórico este momento t^m es tal que $\frac{E(P)}{2} = \int_0^{t^m} N_t f(t) dt$.

Con motivo de conocer las características de P supongamos un crecimiento exponencial constante en una población de 100 personas en el momento 0 (el nivel no influye en las conclusiones), con escenarios de tasa anual $r = 0,013$ (promedio latinoamericano en 2010-2015, *UN-WPP17*) y $r = 0.04$ (Guinea Ecuatorial en mismo período, *UN-WPP17*), y un horizonte de relevamiento de $d = 0,2$ y $d = 0,75$ (2 y 9 meses, cercano a las experiencias de Uruguay en 2011 y Colombia en 2005). Para cada escenario nos interesa conocer cuánta población se “cuenta”, cuál es la varianza de esa estimación, y en qué proporción de la ventana censal se debería colocar lo primero.

Table 1: Escenarios teóricos de relevamiento

r	d	Rd	E(P)	CV	t^*/d	tm/d
0.01	0.20	1	1001501	0.04	0.75	0.79
0.01	0.20	2	1000927	0.06	0.46	0.45
0.01	0.20	3	1001001	0.06	0.50	0.50
0.01	0.20	4	1001713	0.02	0.86	0.89
0.01	0.75	1	1005642	0.15	0.75	0.79
0.01	0.75	2	1002415	0.18	0.32	0.27
0.01	0.75	3	1003759	0.22	0.50	0.50
0.01	0.75	4	1004941	0.12	0.66	0.66
0.04	0.20	1	1006019	0.15	0.75	0.79
0.04	0.20	2	1003714	0.23	0.46	0.45
0.04	0.20	3	1004011	0.23	0.50	0.50
0.04	0.20	4	1006870	0.10	0.86	0.89
0.04	0.75	1	1022772	0.58	0.75	0.79
0.04	0.75	2	1009716	0.73	0.32	0.27
0.04	0.75	3	1015151	0.87	0.50	0.50
0.04	0.75	4	1019921	0.49	0.66	0.66

^a CV: coeficiente de variación.

^b t^*/d y tm/d indican en qué porción de d se debería colocar la estimación final.

En base a los escenarios simulados, como principales conclusiones podemos inferir que:

- Extender el período aumenta tanto el “conteo” final como su variabilidad, y el impacto depende de la relación entre r y R_d . Si repetimos el anterior ejercicio para pequeños intervalos en d , con una tasa de crecimiento del 1%, cada mes adicional de extensión implica incrementar algo menos de 0,2% el coeficiente de variación.
- El momento donde colocar la estimación parece estar afectado solamente por la forma de relevamiento durante d . El único escenario que pareciera justificar colocar la estimación a mitad de período es cuando el relevamiento es uniforme (forma 3 de relevamiento) o si la población no crece en el intervalo. El criterio de t^m (INE Uruguay, 2011) es una muy buena aproximación a t^* .

Es posible que esta variabilidad tenga un impacto en las tasas de mortalidad específicas. Una de las metodologías de estimación de la varianza de m_x más utilizada por los demógrafos sólo considera la varianza de P (a mitad de período) como parte del estimador de la población a edad exacta, lo que permite un tratamiento de cohorte y hacer uso de una distribución binomial (Chiang, 1961). El hecho de que variabilidad de P sugerida aquí no dependa de la cantidad de población a “contar”, contrasta con la metodología mencionada, la que aumenta la precisión a medida que se acumula “exposición”. Una mayor profundización merece este posible impacto en las tasas.

Por último, un comentario merece la estrategia de Colombia, la cual tiene como fin principal la buena captación de una estructura de atributos a referenciar temporalmente a mitad de período. Como vimos en los escenarios simulados, subpoblaciones con fuerte cambio demográfico deben ser objeto de estrategias censales concentradas. Con operativos sectorizados de este tipo, cada estimación de estas subpoblaciones luego puede trasladarse en el tiempo hacia un punto en común, utilizando las respectivas tasas r (en función de una estimación censal anterior) y sus respectivos puntos de referencia t^* (o t^m). Adicionalmente, como elementos de análisis más innovadores, modelos de microsimulación (Zagheni, 2017) pueden ser una herramienta para cuantificar la variabilidad estadística del conteo en áreas de especial movilidad y crecimiento.

El motivo de estos párrafos fue exponer algunas ideas sobre la naturaleza de aquello que contamos en un censo de derecho, sistema complejo donde se interrelacionan la dinámica demográfica y la planificación censal

(con sus objetivos y restricciones), tratando de dimensionar un comportamiento dinámico, como aquel que mide un globo mientras otro lo esta inflando.

Bibliografía y fuente de datos:

- *Chiang et. al (1961). "Standard error of the age-adjusted death rate". U.S. Dept. of Health, Education, and Welfare, Public Health Service, National Vital Statistics Division.*
- *DANE (2008). "Síntesis del proceso de cierre del Censo General 2005". Colombia*
- *DANE (2009). "Metodología General del Censo 2005". Colombia*
- *INE de Uruguay (s/f). "Resultados del Censo de Población 2011". Consultado en <http://www.ine.gub.uy/documents/10181/35289/analisispais.pdf/cc0282ef-2011-4ed8-a3ff-32372d31e690>*
- *Keyfitz N. (1966). Sampling variance of standarized mortality rates. Human Biology, Vol. 38, No. 3*
- *UN-WPP17: United Nations. World Population Propects 2017. Consultado en <https://population.un.org/wpp/>*
- *Zagheni E. 2015. "Microsimulation in Demographic Research.". International Encyclopedia of the Social and Behavioral Sciences, 2nd edition, Vol 15.*

Interesados en replicar los resultados e intercambiar ideas pueden escribir a: ivanwilliams1985@gmail.com