

Problema 1

A. Alonso Santaella

22 de julio de 2017

1. Idea 1:

El primer acercamiento conceptual a la solución del problema se basaba en la utilización del dato "*Hijos sobrevivientes, Fecha de nacimiento del último hijo, Sobrevivencia y Edad al morir.*" del cuestionario ampliado descrito en la documentación del C.P.V. 2010¹. Con el dato anterior se pretendía calcular la proporción de bebés de cero a seis meses con respecto a la población total de cada AGEB mediante la suma de los nacimientos en los últimos seis meses –y la resta de las muertes–. La proporción de bebés a la población total se utilizaría en conjunto con la tasa de crecimiento de la población (derivada de censos y conteos pasados mediante regresión) para calcular la población actual y determinar la proporción actual de niños de cero a seis meses. El acercamiento no pudo concretarse debido a que el dato requerido no apareció en los datos del .csv del portal de datos abiertos para el C.P.V. 2010.

2. Idea 2:

El segundo acercamiento consistía en utilizar los datos de tipo "*Población de 3 años y más*" para la construcción numérica de la función de distribución de la edad:

$$F(t) = P(\text{Edad} \leq t)$$

La construcción numérica de F permitiría la interpolación para la obtención de $F(t = 0.5)$, i.e., la probabilidad de que un miembro de la AGEB en 2010 sea menor a seis meses. La probabilidad podría luego ser aplicada para obtener el número de bebés esperados de acuerdo a la población actual,

¹<http://www.beta.inegi.org.mx/contenidos/proyectos/ccpv/2010/doc/presentacion.pdf>

obtenida mediante regresión. Problemas de la limpieza y consistencia de los datos, así como limitaciones de tiempo (esto no se me ocurrió hasta muy tarde), no permitieron la implementación de este método, por lo que en este momento no hay datos que presentar.

El Jupyter Notebook asociado a este problema se llama “Problema1.ipynb”, si bien no contiene nada demasiado útil; la última celda muestra que algunas AGEs de la Delegación Álvaro Obregón no tenían datos asociados, o había algún problema con la función de R split.