Circle GPS Sampling Algorithm

27 December 2016

Patrick D. Emond

Algorithm

Once a town has been selected for sampling the following algorithm is performed on that town:

1. select random point in town
2. make a set of all unselected buildings, B, which are within the "sampling radius"* of the selected point
3. if B is empty then go back to step 2 (if this happens 1000 times then quit with an error)
4. select a random building, b, in set B
5. add all individuals** in all households in building b to the sample
6. if number of individuals in the sample is less than the sample size return to step 2

* "sampling radius" is a parameter defined by the user when running the sampler and remains constant for all sampling circles across all towns

** selected individuals may be restricted to children only, or one individual per household based on input parameters defined by the user

Sampling Weights

Probability of selecting town i is:

$$P_i = P(\text{select town i}) = \frac{n_i}{N}$$

where n_i = population of town i, and N = total population. Strictly, it's number of buildings or households, but we ignore that, and assume the number of people per household is constant.

Probability of selecting circle $\overset{\circ}{j}$ is:

$$P_{\overset{\circ}{j}\bigvee i} = P(\text{select circle } \overset{\circ}{j} \text{ within town } i) = \frac{A_{\overset{\circ}{j}}}{A_i}$$

where $A_{\overset{\circ}{j}}$ = area of the circle $\overset{\circ}{j}$, and A_i = area of town i. Since all circles have the same area across all towns we do not have to include $P_{\overset{\circ}{j}\bigvee i}$ when determining sampling weights.

Probability of selecting household k in circle $\overset{\circ}{j}$ is:

$$P_{k\bigvee\overset{\circ}{j}} = P(\text{select household } k \text{ within circle } \overset{\circ}{j}) = \frac{1}{n_{\overset{\circ}{ij}}}$$

where $n_{\overset{\circ}{ij}}$ = number of households in circle $\overset{\circ}{j}$.

Therefore the probability of sampling household k is:

$$P_{\overset{\circ}{ijk}} = P(\text{select household } \overset{\circ}{ijk}) = P_i \times P_{\overset{\circ}{j}\bigvee i} \times P_{k\bigvee\overset{\circ}{j}}$$

Making the weight for household $\overset{\circ}{ijk}$ equal to:

$$\begin{aligned} w_{\overset{\circ}{ijk}} &= \frac{1}{P_{\overset{\circ}{ijk}}} \\ &= \frac{1}{P_i \times P_{k \bigvee \overset{\circ}{j}}} \quad \left(\text{since} P_{\overset{\circ}{j} \bigvee i} \text{ is identical in all towns}\right) \\ &= \frac{n_{\overset{\circ}{ij}}}{n_i} N \end{aligned}$$