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 1 (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 1

* "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:

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

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 in the different towns is constant.

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

$$P_{\overset{\circ}{j}\bigvee i} = P\big(\text{select circle } \overset{\circ}{j} \text{ within town } i\big) = \frac{A_c}{A_t} n_c$$

where A_c = area of the sampling circle (all circles will have the same radius defined by the "sampling radius" parameter set when running a sampler), A_t = area of a town (all towns have the same area) and n_c is the number of circles chosen within the town (including empty circles). The number of circles chosen will vary depending on the sampling radius and population density.

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

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

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

Therefore the probability of sampling household k is:

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

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

$$\begin{aligned} w_{\overset{\circ}{ijk}} &= \left[P_{\overset{\circ}{ijk}}\right]^{-1} \\ &= \left[P_i \times P_{\overset{\circ}{j}\bigvee i} \times P_{k\bigvee\overset{\circ}{j}}\right]^{-1} \\ &= \frac{n_{\overset{\circ}{ij}} N A_t}{n_i n_c A_c} \end{aligned}$$

where $n_{\overset{\circ}{ij}}$ = number of households in circle $\overset{\circ}{j}$, n_i = population of town i, n_c is the number of circles chosen within the town, N = total population, A_t = area of a town, A_c = area of the sampling circle

If all eligible people in the household are chosen, then this is the weight for all those individuals. When only one person per household is selected, let n_{kl} be the number of eligible people in the household (e.g.: the number of children or number of adults):

$$P_{l\bigvee k} = P\big(\text{select person}l\text{within household}k\big) = \frac{1}{n_{kl}}$$

Therefore the probability of sampling an individual l is:

$$P(\text{select individiual}) = P_{\overset{\circ}{ijkl}} = P_{l\bigvee k} \times P_{\overset{\circ}{ijk}}$$

Making the weight for individual $\overset{\circ}{ijkl}$ equal to:

$$w_{\overset{\circ}{ijkl}} = \frac{n_{\overset{\circ}{ij}} n_{kl} N A_t}{n_i n_c A_c}$$