Sequential Sampling Examples

Using the provided functions for sequential sampling

B. K. M. Case

April 7, 2022

Example of using some of the functions from this study. Please note, however, that this code was not intended to work on other datasets and would require minor changes to do so.

Load required packages and the data

Here we will use the data from the village El Paternito:

The `sampling_design` function

This function will take care of most of the required setup for us, so we can skip the usual steps of using SPDE and get results right away.

```
design <- sampling_design(dat_sub, init=30, pred="known", silent=FALSE, strat_arg=list(alpha=0.5))
```

```
## [1] "Begin iteration 1"
## [1] "Begin iteration 2"
## [1] "Begin iteration 3"
## [1] "Begin iteration 4"
## [1] "Begin iteration 5"
## [1] "Begin iteration 6"
## [1] "Begin iteration 7"
## [1] "Begin iteration 8"
## [1] "Begin iteration 9"
## [1] "Begin iteration 10"
## [1] "Begin iteration 11"
## [1] "Begin iteration 12"
## [1] "Begin iteration 13"
## [1] "Begin iteration 14"
## [1] "Begin iteration 15"
## [1] "Begin iteration 16"
## [1] "Begin iteration 17"
## [1] "Begin iteration 18"
## [1] "Begin iteration 19"
## [1] "Begin iteration 20"
## [1] "Begin iteration 21"
```

Here, we have used adaptive sampling with an initial sample of 30 houses, the globally available covariates, and exploration parameter $\alpha = 1$. We have also set `silent=FALSE` to print our progress each iteration of the algorithm.

We can leave `strat_arg` empty to use random selection:

```
design_random <- sampling_design(dat_sub, init=30, pred="global")
```

Scoring the designs

To evaluate the number of houses visited and the true infection rate remaining in the village, we run

```
design_score(design, dat_sub$truth)
```

```
## # A tibble: 1 x 4
##       m     n act_pct sel
##   <int> <int>   <dbl> <list>
## 1    90   107  0.0280 <list [21]>
```

```
design_score(design_random, dat_sub$truth)
```

```
## # A tibble: 1 x 4
##       m     n act_pct sel
##   <int> <int>   <dbl> <list>
## 1    99   107 0.00935 <list [24]>
```

The second argument of `design_score` is a vector of the true infection labels in the village.