

README

A design and analytic strategy for monitoring disease positivity and case characteristics in accessible closed populations

by Robert H. Lyles, Yuzi Zhang, Lin Ge, and Lance A. Waller

Illustrate calculations for case count and continuous mean estimations with a simulated dataset

The data was simulated under the anchor stream design with the setting in which ($N_{\text{tot}} = 500$, selection probability into the random sample $\psi = 0.1$, prevalence $p = 0.2$).

The simulated dataset has 196 observations, each observation includes a continuous random variable (\mathbf{x}) representing a hypothetical biomarker level, indicators of capture status (y_1 for Stream 1, and y_2 for Stream 2), symptom (`symp`), and disease (`case`). For the continuous random variable, it was drawn from a mixture of normal distributions with the mean and the variance differing according to symptom status and disease status. Specifically, normal distributions used for generating data are given by

	symptom = 1	symptom = 0
case = 1	$N(10, 0.75^2)$	$N(5, 0.5^2)$
case = 0	$N(2.5, 1.2^2)$	$N(1, 1.5^2)$

read in self-defined functions and the simulated dataset

```
source("FUN_AnchorStream.R")
load("toydata.rda")
```

take a look of the simulated individual-level data

```
##   y1 y2 case symp      x
## 1  1  1   0    0 0.5010140
## 2  1  1   0    0 -1.7333531
## 3  1  1   0    0  3.1168936
## 4  1  1   0    0 -0.2563737
## 5  1  1   0    0 -0.6856442
## 6  1  1   0    0  5.5656488
```

observed cell counts

n1	n2	n3	n4	n5	n6	n7
6	5	100	46	33	6	304

case count estimates based on the simulated data

```
Ntot = 500
p2 = 0.1
re.counts <- AnchorStream_CaseCount(dat.obs = dat.obs,
                                   Ntot = Ntot, p2 = p2,
                                   num.post = 10000,
                                   seed = 1234,
                                   data.type = "individual",
                                   cellcounts.vec = NULL)
```

results from using random sample alone

```
re.counts$pointest$Nhat.RS
```

```
## [1] 110
```

```
re.counts$SE$SE.RS.FPC
```

```
## [1] 28.07061
```

```
re.counts$CI$CI.RS.Jeffreys
```

```
## [1] 63.47027 171.54453
```

results from using the estimator \hat{N}_ψ

```
re.counts$pointest$Nhat.psi
```

```
## [1] 111
```

```
re.counts$SE$SE.Nhat.psi
```

```
## [1] 23.2379
```

```
re.counts$CI$CI.Nhat.psi.Diri
```

```
## [1] 76.01301 166.25019
```

results from using the estimator \hat{N}_{ψ^*}

```
re.counts$pointest$Nhat.psihatstar
```

```
## [1] 103.7692
```

```
re.counts$SE$SE.Nhat.psihatstar
```

```
## [1] 21.94541
```

```
re.counts$CI$CI.Nhat.psihatstar.Diri
```

```
## [1] 71.86727 162.77501
```

continuous mean estimates based on the simulated data

```
re.continuous <- AnchorStream_Continuous(dat.obs = dat.obs,
                                          Ntot = Ntot, seed = 12345,
                                          nboot = 1000)
```

results of estimating overall mean

```
re.continuous$overall
```

```
##           est           se           lci           uci
## xbar1dot 3.378704           NA           NA           NA
## xbar2dot 2.765692 0.4898039 1.830166 3.722037
## muhatx   2.552818 0.3188371 1.965768 3.194178
```

results of estimating continuous mean among cases

```
re.continuous$cases
```

```
##           est           se           lci           uci
## xbar1dot.cases 7.471739 0.3651921 6.782233 8.195105
## xbar2dot.cases 8.581201 0.8525923 6.771870 10.165044
## muhatx.cases   7.692362 0.5977918 6.529768 8.786034
```

results of estimating continuous mean among non-cases

```
re.continuous$noncases
```

```
##           est           se           lci           uci
## xbar1dot.noncases 1.409414 0.1610869 1.0859128 1.707119
## xbar2dot.noncases 1.125420 0.2890897 0.5713258 1.699620
## muhatx.noncases   1.206818 0.2090333 0.7799455 1.631326
```

results of estimating continuous mean difference for cases relative to non-cases

```
re.continuous$difference
```

```
##           est           se           lci           uci
## xbar1dot.diff 6.062325 0.3996517 5.279764 6.815613
## xbar2dot.diff 7.455780 0.9088060 5.518571 9.238345
## muhatx.diff   6.485544 0.6343667 5.230512 7.656846
```