

## Homework 3

Qianbo Wang  
uni: qw2180

### Problem 1

**What are the underlying assumptions for the capture-recapture estimation method mentioned in that paper?**

In this article, the author used an alternative two-sample capture-recapture method to estimate the population of the rats number in NYC, because the direct two-sample capture-recapture method was not allowed by NYC's Department of Health and Mental Hygiene. So in this alternative method, the underlying assumptions of the capture-recapture estimation were:

1. Rat-inhabited lots reported in the two sample periods are randomly and independently identified from the total population of rat-inhabited lots. This means that:
  - (a). In the first sampling period, rat-inhabited lots are equally likely to be reported.
  - (b). In the second sampling period, any lot identified as rat-inhabited during the first sample is as likely to be identified during the second sample period as any other inhabited lot.
2. The population of rat-inhabited lots being estimated is closed. That is, the total number of rat-inhabited lots does not change throughout the study period.
3. Every lot with at least one rat sighting as evidence that a full colony of rats inhabited the lot for the duration of the study period.
4. Every inhabited lot have approximately equal number of rats.

### Problem 2

**Give other potential application of the capture-recapture estimation method.**

The capture-recapture estimation method is often used in estimate the unknown size of wildlife or other animals in a particular region, such as, estimate the population of fish in a lake, estimate the population size of the wild ox on a steppe, and so on. And, this method can also be used to estimation some number which is hard to collect the population data, such as estimate the size of the undercount in censuses, estimate birth and death rates in an area, and estimate the number of duplicate records on a large list or a large database.

### Problem 3

**State some limitations of the design of this study.**

I think there are several limitations of the design of this study. And some of them the author has declared in the article, yet some not.

1. The validity of the first underlying assumptions. Actually, Rat-inhabited lots reported in the two sample periods are not randomly and independently identified from the total population of rat-inhabited lots, for example, the city may have instructed a property, or it may have hired someone to trap the rats on a specific region, which may causes the sighting not equally probable to be reported.
2. The validity of the second underlying assumptions, the population of rat-inhabited lots is not closed. Since there are always deaths, births and migrations within a population between samples. And also, there may occur some other incidents, such as clearance of rats in NYC. Though the author used a cool-down period to minimize this problem, but this is also a limitation of this study.
3. The author assumed every lot with at least one rat sighting as evidence that a full colony of rats inhabited the lot for the duration of the study period, but in fact there is much likely that territory of one rat colony encompasses several lots, which will cause the overestimation of the population of the rats in this experiment.
4. Because there is reality problem of the real experiment of this two-sample capture-recapture method, the author used an alternative method, which is using the sighting inhabited lots of rats in NYC to pretend the capture-recapture procedure. And use the average number times the total inhabited lots to estimate the population, but, in fact, this is not reality. First, Brooklyn sightings are more likely to happen in the same lots. Second, Brooklyn tends to have larger concentrations of smaller lots. In contrast, Manhattan sightings tend to come from different lots, and those lots make up a larger percentage of the total number of lots in each neighborhood.