

STAT3003 Problem Sheet 3

ZHENG Weijia (William, 1155124322)

April 11, 2021

1 Q1

1. (Adapted from Exercise 7.3 of Scheaffer et al. (2012)) A retail store with four departments has charge accounts arranged by department, with past-due accounts at the top of each departmental list. Suppose the departments average around ten accounts each, with approximately 40% past due. On a given day the accounts might appear as shown in the accompanying table (with account numbers 1 through 40). The store wishes to estimate the proportion of past-due accounts by systematic sampling.

Account numbers	1-11	12-20	21-28	29-40
Delinquent accounts	1,2,3,4	12, 13, 14	21, 22, 23, 24, 25	30, 31, 32, 33

Table 1: Retail Store Accounts Study Data

- (a) List all possible 1-in-10 systematic samples and compute the exact variance of the sample proportion.
- (b) List all possible 1-in-5 systematic samples and compute the exact variance of the sample proportion.
- (c) Compare the result in part (a) with an approximate variance obtained in a SRS sample of size $n = 4$ from this population. Similarly, compare the result in part (b) with that obtained from a SRS with $n = 8$. Can you explain what is going on?

1.1 (a)

We need to fix the starting point, when start at 1: the number of delinquent accounts is 3, the sample proportion is $\frac{3}{4}$.

When start at 2: the number of delinquent accounts is 4. The sample proportion is 1.

When start at 3: the number of delinquent accounts is 4. The sample proportion is 1.

When start at 4: the number of delinquent accounts is 3. The sample proportion is $\frac{3}{4}$.

When start at 5: the number of delinquent accounts is 1. The sample proportion is $\frac{1}{4}$.

When start at 6: the number of delinquent accounts is 0. The sample proportion is 0.

When start at 7: the number of delinquent accounts is 0. The sample proportion is 0.

When start at 8: the number of delinquent accounts is 0. The sample proportion is 0.
 When start at 9: the number of delinquent accounts is 0. The sample proportion is 0.
 When start at 10: the number of delinquent accounts is 1. The sample proportion is $\frac{1}{4}$.
 Therefore the exact variance of sample proportion is 0.165.

1.2 (b)

When start at 1: the number of delinquent accounts is 3. The sample proportion is 0.375.
 When start at 2: the number of delinquent accounts is 4. The sample proportion is 0.5.
 When start at 3: the number of delinquent accounts is 4. The sample proportion is 0.5.
 When start at 4: the number of delinquent accounts is 3. The sample proportion is 0.375.
 When start at 5: the number of delinquent accounts is 2. The sample proportion is 0.25.
 Hence the exact variance of sample proportion is 8.75×10^{-3} .

1.3 (c)

Using the formulae

$$Var(\hat{p}) = \frac{N - n}{N - 1} p(1 - p).$$

And we have $p = 0.4$, $N = 40$, therefore $Var(\hat{p}) = 0.0554$.

For $n = 8$, $Var(\hat{p}) = 0.0246$.

When $n=4$, SRS's variance is smaller and when $n=8$, SRS's is larger.