STA 304H1F-1003H Fall 2019
Assignment 2-Question 4

Question 1. (5 marks)

Consider a stratified design composed of H strata of size $N_h,\ h = 1, \ldots, H$. We want to estimate the population mean μ_y of the characteristic y. Let $\mu_{x,h},\ h = 1, \ldots, H$ be the means in the strata (in the population) of an auxiliary characteristic x. The $\mu_{x,h}$ are supposedly known and we propose to estimate μ_y using the following estimator:

$$\widehat{\mu}_D = \overline{y}_{st} + \mu_x - \overline{x}_{st}$$

where $\overline{y}_{st}$ and $\overline{x}_{st}$ are the basic estimate of the population means μ_y and μ_x for y and x, respectively.

(a) (1 mark) Give an expression of μ_x in terms of $\mu_{x,h},\ h = 1, \ldots, H$.

(b) (1 mark) Show that $\widehat{\mu}_D$ is unbiased estimates for μ_y.

(c) (1 mark) Give the variance of $\widehat{\mu}_D$

(d) (1 mark) Let $n = \sum_{h=1}^{H} n_h$ be the sample size. What is the optimal allocation of the n_h in order to minimise the variance of $\widehat{\mu}_D$? We consider that the init cost of the survey does not depend on the stratum.

(e) (1 mark) In which favourable case is $\widehat{\mu}_D$ preferable to $\overline{y}_{st}$?