Assignment 3. Q1.

1.(a) The total $N = 3374 + 1960 + 52 + 5 = 5391$

The stratified mean estimate $\hat{\mu}_s = \sum_{h=1}^{H} W_h \hat{\mu}_h$

$$= \frac{3374}{5391} \times 491 + \frac{1960}{5391} \times 1050 + \frac{52}{5391} \times 2568 + \frac{5}{5391} \times 18054$$

$$= 730.558$$

The the estimated variance $= \sum_{h=1}^{H} W_h^2 \left(\left(1 - \frac{m_h}{N_h}\right) \frac{\hat{\sigma}_h^2}{n_h} \right)$

$$= \left(\frac{3374}{5391}\right)^2 \left(1 - \frac{100}{3374}\right) \frac{921}{100} + \left(\frac{1960}{5391}\right)^2 \left(1 - \frac{30}{1960}\right) \frac{1278}{30}$$

$$+ \left(\frac{52}{5391}\right)^2 \left(1 - \frac{45}{52}\right) \frac{2343}{45} + \left(\frac{5}{5391}\right)^2 \left(1 - \frac{5}{5}\right) \frac{154054}{5}$$

$$= 6.8$$

(b) The 95% CI for population total τ

$$= N\hat{\mu}_s \pm C\sqrt{N^2 \times \hat{\mathrm{Var}}(\hat{\mu}_s)}$$

since for 95%, $C = 1.96$

$$= 5391 \times 730.558 \pm 1.96\sqrt{5391^2 \times 6.8}$$

The 95% CI is (3910884, 3965993)

Question 2:

(a) $\hat{n}_1 = 300 \times \frac{3374}{5391} = 187.75$

$\hat{n}_2 = 300 \times \frac{1960}{5391} = 109.07$

$\hat{n}_3 = 300 \times 52/5391 = 2.894$

$\hat{n}_4 = 300 \times 5/5391 = 0.278$

Sine the sum of n_h is 300 and we want integer number samples.

Thus $n_1 = 188$ $n_2 = 109$ $n_3 = 3$ $n_4 = 0$

$300-188-109-3=0$.

(b) For easy calculation

$\sum_{h=1}^{4} W_h \hat{\sigma}_h = \frac{3374}{5391} \times \sqrt{921} + \frac{1960}{5391}\sqrt{1278} + \frac{52}{5391}\sqrt{2343} + \frac{5}{5391}\sqrt{15400}$

$= 32.822$

$\hat{n}_1 = 300 \times \frac{\frac{3374}{5391} \times \sqrt{921}}{\sum_{h=1}^{4} W_h \hat{\sigma}_h} = 173.60$ $\hat{n}_2 = 300 \times \frac{\frac{1960}{5391}\sqrt{1278}}{\sum_{h=1}^{4} W_h \hat{\sigma}_h} = 118.80$

$\hat{n}_3 = 300 \times \frac{\frac{52}{5391} \times \sqrt{2343}}{\sum_{h=1}^{4} W_h \hat{\sigma}_h} = 4.27$ $\hat{n}_4 = 300 \times \frac{\frac{5}{5391} \times \sqrt{15400}}{\sum_{h=1}^{4} W_h \hat{\sigma}_h} = 3.23$

Round to integer: $n_1 = 174$ $n_2 = 119$ $n_3 = 4$ $n_4 = 3$

(c) $Var_{prop}(\hat{\mu}_{prop}) = \sum_{h=1}^{4} W_h (1 - \frac{n_h}{N_h}) \frac{\hat{S}^2_h}{n_h}$

$= (\frac{3374}{5391})^2 (1 - \frac{188}{3374}) \times \frac{921}{188} + (\frac{1960}{5391})^2 (1 - \frac{109}{1960}) \times \frac{1278}{109}$

$+ (\frac{52}{5391})^2 (1 - \frac{3}{52}) \times \frac{2343}{3} + (\frac{5}{5391})^2 (1 - \frac{0}{5}) \times \frac{15400}{0}$

Since no sample for n_4, we ignore it

$= 3.34$

$Var_{opt}(\hat{\mu}_{opt}) = (\frac{3374}{5391})^2 (1 - \frac{174}{3374}) \times \frac{921}{174} + (\frac{1960}{5391})^2 (1 - \frac{119}{1960}) \times \frac{1278}{119} + (\frac{52}{5391})^2 (1 - \frac{4}{52}) \frac{2343}{4} + (\frac{5}{5391})^2 (1 - \frac{3}{5}) \times \frac{15400}{3}$

$= 3.32$

The variance under optimal allocation is smaller that the proportional allocation, thus is experted since the optimal allocation is designed to minimize Variance