1. Questão 1

   (a) Do formulário, temos que:

$$
\begin{aligned}
\widetilde{\widetilde{\tau}} &= \frac{1}{a}\sum_{\alpha=1}^{a}\widetilde{\tau}_\alpha = \frac{15288.82}{10} = 1528,88,\ \widetilde{\mu}_{C_1} = \frac{\widetilde{\widetilde{\tau}}}{\overline{\overline{B}}} = \frac{1528,882}{45,32} = 33,74,\\
\widetilde{\mu}_{C_2} &= \frac{\widetilde{\widetilde{\tau}}}{\widetilde{\widetilde{B}}} = \frac{1528,882}{33,70} = 45,37,\ \widetilde{\mu}_{C_3} = \frac{1}{a}\sum_{\alpha=1}^{a}\widetilde{\mu}_\alpha = \frac{385,14}{10} = 38,51,
\end{aligned}
$$

   As variâncias estimadas dos estimadores são dadas por:

$$
\begin{aligned}
\widetilde{\mathcal{V}}(\widetilde{\mu}_{C_1}) &= \frac{1}{a(a-1)}\sum_{\alpha=1}^{a}\left(\frac{B_\alpha}{\overline{\overline{B}}}\widetilde{\mu}_\alpha - \widetilde{\mu}_{C_1}\right)^2 = \frac{8541,626}{90} = 94,91,\\
\widetilde{\mathcal{V}}(\widetilde{\mu}_{C_2}) &= \frac{1}{a(a-1)}\sum_{\alpha=1}^{a}\left(\frac{b_\alpha}{\widetilde{\widetilde{B}}}\right)^2(\widetilde{\mu}_\alpha - \widetilde{\mu}_{C_2})^2 = \frac{1887,387}{90} = 20,97\\
\widetilde{\mathcal{V}}(\widetilde{\mu}_{C_3}) &= \frac{1}{a(a-1)}\sum_{\alpha=1}^{a}(\widetilde{\mu}_\alpha - \widetilde{\mu}_{C_3})^2 = \frac{1797,891}{90} = 19,98.
\end{aligned}
$$

   Já para os vícios temos que $\widetilde{B}(\widetilde{\mu}_{C_1}) = \widetilde{B}(\widetilde{\mu}_{C_2}) = 0$ e

$$
\begin{aligned}
\widetilde{\mu} &= \frac{1}{a}\sum_{\alpha=1}^{a}\left(\frac{b_\alpha}{\widetilde{\widetilde{B}}}\right)\widetilde{\mu}_\alpha = \frac{453,6742}{10} = 45,3674,\ \widetilde{B}(\widetilde{\mu}_{C_3}) = (\widetilde{\mu}_{C_3} - \widetilde{\mu}) = 38,514 - 45,3674\\
&= -6,85.
\end{aligned}
$$

   Por fim, temos que:

$$
\begin{aligned}
EQM(\widetilde{\mu}_{C_1}) &= \widetilde{\mathcal{V}}(\widetilde{\mu}_{C_1}) = 94,91\\
EQM(\widetilde{\mu}_{C_2}) &= \widetilde{\mathcal{V}}(\widetilde{\mu}_{C_2}) = 20,97\\
EQM(\widetilde{\mu}_{C_3}) &= \widetilde{\mathcal{V}}(\widetilde{\mu}_{C_3}) + \widetilde{B}(\widetilde{\mu}_{C_3})^2 = 66,95.
\end{aligned}
$$

   (b) Temos dois estimadores viciados, embora um deles tenda a ter viés próximo de zero. Assim, devemos comparar os EQM's. Do item a), temos que o estimador $\widehat{\mu}_{C_2}$ é o melhor.

   (c) Conforme os resultados do item c), escolheríamos o estimador $\widehat{\mu}_{C_2}$. Era esperado pois $\widetilde{Corre}(\widetilde{\mu}_\alpha, b_\alpha) = 0,83$.

   (d) Formulário está errado, pois o EPA é na verdade o ρ_{int}. Nesse caso tenho que considerar que AC teria vantagem e ser condescendente em relação às justificativas.

$$\widetilde{\sigma}^2_{dc} = \frac{1}{a}\sum_{\alpha=1}^{a}\left(\frac{b_\alpha}{\widetilde{\overline{B}}}\right)\widetilde{\sigma}^2_a = \frac{2616,475}{10} = 261,65$$

$$EPA_{C_2} = \frac{a\widetilde{\mathcal{V}}(\widetilde{\mu}_{C_2}) - \dfrac{\widetilde{\sigma}^2_{dc}}{\widetilde{\overline{B}}-1}}{a\widetilde{\mathcal{V}}(\widetilde{\mu}_{C_2}) + \widetilde{\sigma}^2_{dc}} = \frac{201,7075}{471,3565} = 0,43,$$

portanto como $EPA_{C_2} < 1$ então há um ganho de eficiência da amostragem por conglomerado com relação à AASs.

2. Temos que:

$$\mathcal{E}_{A_2}(\widehat{\mu}_D) = \mathcal{E}_{A_2}[\widehat{\mu}_Y + (\mu_x - \widehat{\mu}_X)] = \mathcal{E}_{A_2}(\hat{\mu}_Y) + \mu_X - \mathcal{E}_{A_2}(\hat{\mu}_X)_{A_2} = \mu_Y + \mu_X - \mu_X = \mu_Y, \tag{1}$$

$$\mathcal{V}_{A_2}(\widehat{\mu}_D) = \mathcal{V}_{A_2}\left(\frac{1}{n}\sum_{i=1}^{n} D_i\right) = (1-f)\frac{s_d^2}{n}, \mathcal{EP}_{A_2}(\widehat{\mu}_D) = s_d\sqrt{\frac{1-f}{n}}$$

em que $s_d^2 = \dfrac{1}{N-1}\displaystyle\sum_{i=1}^{N}\left(d_i - \overline{d}\right)^2$, $d_i = y_i - (x_i - \mu_x)$, $\overline{d} = \dfrac{1}{N}\displaystyle\sum_{i=1}^{N} d_i$. De (1), temos que $\mathcal{B}(\widehat{\mu}_D) = 0$ e, assim, $\mathcal{EQM}(\widehat{\mu}_D) = \mathcal{V}(\widehat{\mu}_D)$.

3. Temos que:

$$\mathcal{E}_{A_2}\left(\widehat{p}_A\right) = \mathcal{E}_{A_2}\left(\mathcal{E}_{A_2}\left(\widehat{p}_A|A\right)\right) = \mathcal{E}_{A_2}\left(p_A\right) = p$$

$$\begin{aligned}
\mathcal{V}_{A_2}\left(\widehat{p}\right) &= \mathcal{V}_{A_2}\left(\mathcal{E}_{A_2}\left(\widehat{p}_A|A\right)\right) + \mathcal{E}_{A_2}\left(\mathcal{V}_{A_2}\left(\widehat{p}_A|A\right)\right) = \mathcal{V}_{A_2}\left(p_A\right) + \mathcal{E}_{A_2}\left(\frac{N-A}{N-1}\frac{pq}{A}\right) \\
&= \mathcal{E}_{A_2}\left(\frac{N}{A} - 1\right)\left(\frac{pq}{N-1}\right) \approx \left(N\ln N + N - 1\right)\frac{pq}{N-1} \\
&= \left(\frac{N\ln N}{N-1} + 1\right)pq
\end{aligned}$$

Como ambos são não viciados, vamos comparar suas variâncias. Com efeito:

$$\frac{\mathcal{V}_{A_2}\left(\widehat{p}_A\right)}{\mathcal{V}_{A_2}\left(\widehat{p}\right)} \approx \frac{\left(N\ln N + N - 1\right)\frac{pq}{N-1}}{\frac{N-n}{N-1}\frac{pq}{n}} = \frac{n\left(N\ln N + N - 1\right)}{N-n} = g(n, N)$$

Note que:

$$\begin{aligned}
g(n, N) &= \frac{\left(N\ln N + N - 1\right)}{\frac{N}{n} - 1} \rightarrow \lim_{n\rightarrow\infty}\frac{\left(N\ln N + N - 1\right)}{\frac{N}{n} - 1} = \infty \\
g(n, N)(\text{L'Hopital}) &= n\left(\ln N + 1 + 1\right) \rightarrow n\left(\ln N + 2 > 1\right)
\end{aligned}$$

Assim, $\mathcal{V}_{A_2}(\widehat{p}_A) > \mathcal{V}_{A_2}(\widehat{p}_A)$ e, assim, como esperado o estimador $\widehat{p}$ é preferível ao estimador $\widehat{p}_A$