1. Questão 1

   a) Temos que $\mathcal{E}(\widehat{\mu}_{Reg}) = \mathcal{E}(\widehat{\mu}_y) + b_0 \left(\mu_x - \mathcal{E}(\overline{X})\right) = \mu_y$ e (ver formulário)

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

   em que

$$\begin{aligned}
s_d^2 &= \frac{1}{N-1}\sum_{i=1}^{N}\left(y_i - b_0\left(x_i-\mu_x\right)-\mu_d\right)^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left[\left(y_i-\mu_y\right)-b_0\left(x_i-\mu_x\right)\right]^2 \\
&= \frac{1}{N-1}\sum_{i=1}^{N}\left[\left(y_i-\mu_y\right)^2 - 2b_0\left(y_i-\mu_y\right)\left(x_i-\mu_x\right)+b_0^2\left(x_i-\mu_x\right)^2\right] \\
&= \frac{1}{N-1}\sum_{i=1}^{N}\left(y_i-\mu_y\right)^2 - 2b_0\frac{1}{N-1}\sum_{i=1}^{N}\left(y_i-\mu_y\right)\left(x_i-\mu_x\right) \\
&+ b_0^2\frac{1}{N-1}\sum_{i=1}^{N}\left(x_i-\mu_x\right)^2 = s_y^2 - 2b_0 s_{xy} + b_0^2 s_x^2 = s_y^2 - 2\frac{s_{xy}^2}{s_x^2}+\frac{s_{xy}^2}{s_x^2} \\
&= s_y^2\left(1-\rho_{xy}^2\right)
\end{aligned}$$

   Assim,

$$\mathcal{V}(\widehat{\mu}_{Reg}) = \frac{1}{n^2}\mathcal{V}\left(\sum_{i=1}^{n} D_i\right) = \frac{(1-f)\,s_y^2}{n}\left(1-\rho_{xy}^2\right)$$

   b) Temos que:

$$\begin{aligned}
EP(\widehat{\mu}_{Reg}) &= \sqrt{(1-f)\frac{s_d^2}{n}} = \delta \rightarrow \left(\frac{1}{n}-\frac{1}{N}\right) = \frac{\delta^2}{s_d^2} \rightarrow \frac{1}{n} = \frac{\delta^2}{s_d^2}+\frac{1}{N} \\
&\rightarrow\ n = \frac{s_d^2 N}{N\delta^2 + s_d^2} \qquad (1)
\end{aligned}$$

2. Questão 2

a) Temos que

Estimador	Probabilidade de seleção da amostra		
	12	13	23
$\widehat{p}_{C_1}$	0,75	0,50	0,25
$\widehat{p}_{C_2}$	0,60	0,67	0,25
$\widehat{p}_{C_3}$	0,67	0,50	0,17
$P(\widehat{p}_{C_i} = \widetilde{p}_{C_1})$	0,33	0,33	0,33

Estimador	Estatística			
	E(.)	V(.)	B(.)	EQM(.)
$\widehat{p}_{C_1}$	0,50	0,04	0,00	0,04
$\widehat{p}_{C_2}$	0,51	0,03	0,01	0,03
$\widehat{p}_{C_3}$	0,44	0,04	-0,06	0,05

b) Os estimadores $\widehat{p}_{C_2}$ e $\widehat{p}_{C_3}$ são viciados (pela teoria, o que também se reflete empiricamente). Contudo o EQM de $\widehat{p}_{C_2}$ é menor e, também pelo fato do seu vício ser pequeno $\approx 0,01$, escolhê-lo-ia. Tal resultado era esperado pois há uma correlação (tendência) positiva entre $\widetilde{p}_\alpha$ e B_α.

3. Questão 3

a) Temos que:

$$\mathcal{E}(\widehat{\mu}_A) = \mathcal{E}\left(\mathcal{E}\left(\widehat{\mu}_A | A = a\right)\right) = \mathcal{E}\left(\mathcal{E}\left(\frac{1}{a}\sum_{i=1}^{a} Y_i\right)\right) = \mathcal{E}(\mu) = \mu$$

$$\begin{aligned}
\mathcal{V}(\widehat{\mu}_A) &= \mathcal{E}(\mathcal{V}\left(\widehat{\mu}_A | A = a\right)) + \mathcal{V}(\mathcal{E}\left(\widehat{\mu}_A | A = a\right)) \\
&= \mathcal{E}\left(\mathcal{V}\left(\frac{1}{a}\sum_{i=1}^{a} Y_i\right)\right) + \mathcal{V}\left(\mu\right) = \mathcal{E}\left(\left(1 - f\right)\frac{s^2}{A}\right) + 0 \\
&= \mathcal{E}\left(\left(\frac{1}{A} - \frac{1}{N}\right)s^2\right) = \left(\mathcal{E}\left(\frac{1}{A}\right) - \frac{1}{N}\right)s^2 = \left(\frac{\ln(n_2) - \ln(n_1)}{n_2 - n_1} - \frac{1}{N}\right)s^2
\end{aligned}$$

Note que,

$$\frac{\mathcal{V}(\widehat{\mu}_A)}{\mathcal{V}(\widehat{\mu})} = \frac{\left(\frac{\ln(n_2) - \ln(n_1)}{n_2 - n_1} - \frac{1}{N}\right)}{\left(\frac{1}{n} - \frac{1}{N}\right)}$$

Por outro lado, temos que (dependendo dos valores de n_1, n_2 e n)

$$\frac{\frac{\ln(n_2) - \ln(n_1)}{n_2 - n_1}}{\frac{1}{n}} = \frac{n\left(\ln(n_2) - \ln(n_1)\right)}{n_2 - n_1} > 1$$

Portanto,

$$\begin{aligned}
\frac{1}{n} &< \frac{\ln(n_2) - \ln(n_1)}{n_2 - n_1} \to \left(\frac{1}{n} - \frac{1}{N}\right)s^2 < \left(\frac{\ln(n_2) - \ln(n_1)}{n_2 - n_1} - \frac{1}{N}\right)s^2 \\
&\to \mathcal{V}(\widehat{\mu}) < \mathcal{V}(\widehat{\mu}_A)
\end{aligned}$$

Eventualmente podemos ter que $\mathcal{V}(\widehat{\mu}) > \mathcal{V}(\widehat{\mu}_A)$. À rigor, para termos uma compração justa, os tamanhos de amostra n e a, tem de ser próximos.