Ex. Cap. 7.

1. Uma população com $N = 2000$ elementos foi dividida em $A = 200$ conglomerados de $N_0 = 10$ elementos. Desta população uma amostra de $a = 20$ conglomerados é selecionada de acordo com a AASc e todos os elementos nos conglomerados selecionados são observados com relação a determinada caracteristica populacional. O número de indivíduos que possuem a característica na amostra é apresentado na tabela abaixo.

C	I	C	I	C	I	C	I
1	5	6	1	11	2	16	7
2	3	7	6	12	3	17	0
3	2	8	10	13	6	18	7
4	9	9	4	14	1	19	2
5	3	10	4	15	1	20	1

C: No. do conglomerado - I: No. de indivíduos que possuem a característica

(i) Encontre uma estimativa para o número total de indivíduos na populacção que possuem a caracteristica de interesse e uma estimativa para a variância da estimativa do total.
(ii) Encontre uma estimativa para a proporção de indivíduos na população que possuem a caracteristica de interesse e uma estimativa para a variância da estimativa da proporção.
(iii) Encontre uma estimativa para o coeficiente de correlação intraclasse.

2. Conforme visto na seção 7.8, a amostragem sistemática usual não permite a obtenção de estimadores da variância da estimativa da média. Recordamos que na amostragem sistemática usual, a população é dividida em n grupos com k elementos cada, onde $N = kn$. Para poder contornar esta dificuldade, vamos considerar amostras sistemáticas replicadas. Nesta situação, a população com N elementos é dividida em n_s grupos com $k' = ks$ elementos cada grupo, de modo que

$$n = sn_s \quad \text{e} \quad N = nk = n_s sk = n_s k'.$$

No primeiro grupo selecionamos s elementos segundo a AASs (ou AASc) e, sistematicamente, selecionamos um elemento de cada grupo para cada amostra, sempre observando a ordem do elemento selecionado no primeiro grupo para cada amostra.

Exemplo: Para uma população com $N = 40$, $n = 8$ e $k = 5$, podemos sortear $s = 2$ amostras sistemáticas com $n_2 = 4$ elementos cada, de modo que $k' = ks = 2 \times 5 = 10$, ao invés de uma única amostra sistemática com $n = 8$ elementos.

Estimador de $\overline{Y}$: Dadas as médias $\overline{Y}_{si1}, \ldots, \overline{Y}_{sis}$ das s amostras sistemáticas, consideramos como estimador da média populacional a média das médias amostrais, isto é,

$$\overline{y}_{siR} = \frac{1}{s}\sum_{j=1}^{s}\overline{Y}_{sij}.$$

a) Usando resultados já vistos para amostragem por conglomerados de tamanhos iguais (identifique as amostras sistemáticas como conglomerados) com AASs, mostre que $\overline{y}_{siR}$ é não viciado para estimar $\overline{Y}$. Mostre que

$$Var_{AASs}[\overline{y}_{siR}] = (1-f_s)\frac{S^2_{siR}}{s},$$

com $f_s = s/k' = s/ks = 1/k$ e $S^2_{siR} = \sum_{j=1}^{k'}(\overline{Y}_{sij} - \overline{Y})^2/(k'-1)$. Mostre também que um estimador não viciado de $Var_{AASs}[\overline{y}_{siR}]$ é dado por

$$var_{AASs}[\overline{y}_{siR}] = (1-f_s)\frac{s^2_{siR}}{s},$$

onde $s^2_{sir} = \sum_{j=1}^{s}(\overline{Y}_{sij} - \overline{y}_{siR})^2/(s-1)$.

b) Considere a população dos $N = 180$ quarteirões do Exercício 2.2, onde a variável de interesse é Y_i : número de domicílios alugados no codomínio, $i = 1, \ldots, 180$, com esta ordenação. Estime $\overline{Y}$ usando amostragem sistemática com $n = 20$ e $s = 4$ réplicas. Estime a variância desta estimativa.