Ejercicios del Sarndal

Pedro Leal

Problema 2.1

En la planificación de un estudio de red de oficinas, se propuso el siguiente esquema de muestreo secuencial para seleccionar una muestra aleatoria de dos intervalos horarios no adyacentes de los ocho intervalos 9–10, 10–11, ..., 16–17 (etiquetados 1, 2, ..., 8):

Seleccionar el primer intervalo con probabilidad uniforme de los ocho intervalos.

Seleccionar, sin reemplazo, el segundo intervalo con probabilidad uniforme de los intervalos no adyacentes al seleccionado en el primer paso.

a) Determine las probabilidades de inclusión de primer orden.

b) Determine las probabilidades de inclusión de segundo orden. ¿Es el diseño inducido por el esquema de muestreo medible?

c) Determine las covarianzas de los indicadores de pertenencia a la muestra.

d) Verifique que el Resultado 2.6.2 se cumple en esta aplicación.

Solución 2.1

(a) Probabilidades de inclusión de primer orden

La probabilidad π_i de que el intervalo i esté en la muestra es:

$$\pi_i = P(\text{seleccionar } i \text{ primero}) + P(\text{seleccionar } i \text{ segundo})$$

La probabilidad de seleccionar el intervalo i-esimo para todo i es $\frac{1}{8}$ dado que distribuye como uniforme, pero para ver las probabilidades de que el i-esimo sea escogido de segundas cambia y para ello es mas claro por medio de la siguiente tabla

Cuadro 1: Probabilidad de selección en segunda etapa para cada i

i	Conjunto de posibles j dado i	$P(\text{de cualquier } j \in A_j)$
1	$A_j = \{3,4,5,6,7,8\}$	$\frac{1}{6}$
2	$A_j = \{4,5,6,7,8\}$	$\frac{1}{5}$
3	$A_j = \{1,6,7,8\}$	$\frac{1}{5}$
4	$A_j = \{1,2,6,7,8\}$	$\frac{1}{5}$
5	$A_j = \{1,2,3,7,8\}$	$\frac{1}{5}$
6	$A_j = \{1,2,3,4,8\}$	$\frac{1}{5}$
7	$A_j = \{1,2,3,4,5\}$	$\frac{1}{5}$
8	$A_j = \{1,2,3,4,5,6\}$	$\frac{1}{6}$

Para el caso del primer intervalo se tiene

$$\begin{aligned}\pi_1 &= \underbrace{\frac{1}{8}}_{P(1)} + \underbrace{0 \cdot \frac{1}{5}}_{P(1|2)} + \underbrace{\frac{1}{8} \cdot \frac{1}{5}}_{P(1|3)} + \underbrace{\frac{1}{8} \cdot \frac{1}{5}}_{P(1|4)} + \underbrace{\frac{1}{8} \cdot \frac{1}{5}}_{P(1|5)} + + \underbrace{\frac{1}{8} \cdot \frac{1}{5}}_{P(1|6)} + + \underbrace{\frac{1}{8} \cdot \frac{1}{5}}_{P(1|7)} + \underbrace{\frac{1}{8} \cdot \frac{1}{6}}_{P(1|8)} \\ &= \frac{1}{8}\left(1 + \frac{5}{5} + \frac{1}{6}\right) = \frac{1}{8} \cdot \frac{13}{6} = \frac{13}{48}\end{aligned}$$

En el caso del segundo,

$$
\begin{aligned}
\pi_2 &= \underbrace{\frac{1}{8}}_{P(2)} + \underbrace{0 \cdot \frac{1}{6}}_{P(2|1)} + \underbrace{0 \cdot \frac{1}{5}}_{P(2|3)} + \underbrace{\frac{1}{8} \cdot \frac{1}{5}}_{P(2|4)} + \underbrace{\frac{1}{8} \cdot \frac{1}{5}}_{P(2|5)} + + \underbrace{\frac{1}{8} \cdot \frac{1}{5}}_{P(2|6)} + + \underbrace{\frac{1}{8} \cdot \frac{1}{5}}_{P(2|7)} + \underbrace{\frac{1}{8} \cdot \frac{1}{6}}_{P(2|8)} \\
&= \frac{1}{8}\left(1 + \frac{4}{5} + \frac{1}{6}\right) = \frac{1}{8} \cdot \frac{59}{30} = \frac{59}{240}
\end{aligned}
$$

Y en el ultimo caso especifico,

$$
\begin{aligned}
\pi_3 &= \underbrace{\frac{1}{8}}_{P(3)} + \underbrace{\frac{1}{8} \cdot \frac{1}{6}}_{P(3|1)} + \underbrace{0 \cdot \frac{1}{5}}_{P(3|2)} + \underbrace{0 \cdot \frac{1}{5}}_{P(3|4)} + \underbrace{\frac{1}{8} \cdot \frac{1}{5}}_{P(3|5)} + + \underbrace{\frac{1}{8} \cdot \frac{1}{5}}_{P(3|6)} + + \underbrace{\frac{1}{8} \cdot \frac{1}{5}}_{P(3|7)} + \underbrace{\frac{1}{8} \cdot \frac{1}{6}}_{P(3|8)} \\
&= \frac{1}{8}\left(1 + \frac{3}{5} + \frac{2}{6}\right) = \frac{1}{8} \cdot \frac{58}{30} = \frac{58}{240}
\end{aligned}
$$

Observando la simetría del problema se puede afirmar:

$$
\begin{aligned}
\pi_1 = \pi_8 &= \frac{13}{48} \\
\pi_2 = \pi_7 &= \frac{59}{240} \\
\pi_3 = \pi_4 = \pi_5 = \pi_6 &= \frac{58}{240}
\end{aligned}
$$

Note que

$$
\sum_{i \in s} \pi_i = 2\left(\frac{13}{48}\right) + 2\left(\frac{59}{240}\right) + 4\left(\frac{58}{240}\right) = 2 = N
$$

(b) Probabilidades de inclusión de segundo orden

$$
\pi_{ij} = P(\text{seleccionar } i \text{ y } j) = \frac{1}{8}\left(\frac{1}{k_i} + \frac{1}{k_j}\right)
$$

donde k_i = número de intervalos no adyacentes a i.

Casos:

- **Pares adyacentes:** $\pi_{ij} = 0$ (ej. π_{12}).
- **Pares no adyacentes:**
  - **Bordes entre sí (1 y 8):**
  $$\pi_{18} = \frac{1}{8}\left(\frac{1}{6} + \frac{1}{6}\right) = \frac{1}{24} \approx 0{,}0417$$
  - **Borde y central (1 y 3):**
  $$\pi_{13} = \frac{1}{8}\left(\frac{1}{6} + \frac{1}{5}\right) = \frac{11}{240} \approx 0{,}0458$$
  - **Centrales entre sí (3 y 5):**
  $$\pi_{35} = \frac{1}{8}\left(\frac{1}{5} + \frac{1}{5}\right) = \frac{1}{20} = 0{,}05$$

Matriz completa π_{ij}:

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1	π_1	0	$\frac{11}{240}$	$\frac{11}{240}$	$\frac{11}{240}$	$\frac{11}{240}$	$\frac{11}{240}$	$\frac{1}{24}$
2	0	π_2	0	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{11}{240}$
3	$\frac{11}{240}$	0	π_3	0	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{11}{240}$
4	$\frac{11}{240}$	$\frac{1}{20}$	0	π_4	0	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{11}{240}$
5	$\frac{11}{240}$	$\frac{1}{20}$	$\frac{1}{20}$	0	π_5	0	$\frac{1}{20}$	$\frac{11}{240}$
6	$\frac{11}{240}$	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{1}{20}$	0	π_6	0	$\frac{11}{240}$
7	$\frac{11}{240}$	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{1}{20}$	0	π_7	$\frac{11}{240}$
8	$\frac{1}{24}$	$\frac{11}{240}$	$\frac{11}{240}$	$\frac{11}{240}$	$\frac{11}{240}$	$\frac{11}{240}$	$\frac{11}{240}$	π_8

Medibilidad:

El diseño es **no es medible** porque las probabilidades de inclusiones de intervalos adyacentes son nulos, es decir $\pi_{i,i\pm 1} = 0$

(c) Covarianzas de los indicadores de pertenencia a la muestra

La covarianza entre los indicadores I_i e I_j se calcula como:

$$\mathrm{Cov}(I_i, I_j) = \pi_{ij} - \pi_i \pi_j$$

Matriz de covarianzas completa:

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1	0,1975	−0,0666	−0,0654	−0,0654	−0,0654	−0,0654	−0,0654	−0,0113
2	−0,0666	0,0605	0,0000	−0,0123	−0,0123	−0,0123	−0,0123	−0,0666
3	−0,0654	0,0000	0,0584	0,0000	−0,0584	−0,0584	−0,0584	−0,0654
4	−0,0654	−0,0123	0,0000	0,0584	0,0000	−0,0584	−0,0584	−0,0654
5	−0,0654	−0,0123	−0,0584	0,0000	0,0584	0,0000	−0,0584	−0,0654
6	−0,0654	−0,0123	−0,0584	−0,0584	0,0000	0,0584	0,0000	−0,0654
7	−0,0654	−0,0123	−0,0584	−0,0584	−0,0584	0,0000	0,0605	−0,0666
8	−0,0113	−0,0666	−0,0654	−0,0654	−0,0654	−0,0654	−0,0666	0,1975

Explicación de los valores clave:

- **Diagonal principal (Varianzas):**

$$\mathrm{Var}(I_i) = \pi_i(1 - \pi_i)$$

Ejemplo para $i = 1$:

$$\text{Var}(I_1) = \frac{13}{48}\left(1 - \frac{13}{48}\right) = \frac{455}{2304} \approx 0{,}197$$

- **Pares adyacentes ($|i - j| = 1$):**

$$\text{Cov}(I_i, I_j) = -\pi_i \pi_j$$

Ejemplo para $(1, 2)$:

$$\text{Cov}(I_1, I_2) = -\frac{13}{48} \times \frac{59}{240} = -\frac{767}{11520} \approx -0{,}0666$$

- **Pares no adyacentes ($|i - j| > 1$):**

$$\text{Cov}(I_i, I_j) = \pi_{ij} - \pi_i \pi_j$$

Ejemplo para $(1, 3)$:

$$\text{Cov}(I_1, I_3) = \frac{11}{240} - \left(\frac{13}{48} \times \frac{29}{120}\right) = -\frac{377}{5760} \approx -0{,}0654$$

Note que la suma total de covarianzas satisface:

$$\sum_{i=1}^{8} \sum_{j=1}^{8} \text{Cov}(I_i, I_j) = 0$$

(d) Cálculo de Covarianza y Varianza por Separado

Calculamos $\text{Cov}(I_k, I_k)$ y $\text{Var}(I_k)$ por separado para verificar que sean iguales en la diagonal.

Fórmulas

Para una variable indicadora I_k:

$$\text{Cov}(I_k, I_k) = \pi_{k,k} - \pi_k \pi_k = \pi_k - \pi_k^2 = \pi_k(1 - \pi_k) = \text{Var}(I_k)$$

Cálculos específicos

- Para $k = 1$ y $k = 8$ ($\pi_1 = \pi_8 = \frac{13}{48}$):

$$\text{Cov}(I_1, I_1) = \frac{13}{48}\left(1 - \frac{13}{48}\right) = \frac{13}{48} \cdot \frac{35}{48} = \frac{455}{2304}$$

$$\text{Var}(I_1) = \frac{13}{48}\left(1 - \frac{13}{48}\right) = \frac{455}{2304}$$

$$\text{Cov}(I_1, I_1) = \text{Var}(I_1) = \frac{455}{2304}$$

- Para $k = 2$ y $k = 7$ ($\pi_2 = \pi_7 = \frac{59}{240}$):

$$\text{Cov}(I_2, I_2) = \frac{59}{240}\left(1 - \frac{59}{240}\right) = \frac{59}{240} \cdot \frac{181}{240} = \frac{10679}{57600}$$

$$\text{Var}(I_2) = \frac{59}{240}\left(1 - \frac{59}{240}\right) = \frac{10679}{57600}$$

$$\text{Cov}(I_2, I_2) = \text{Var}(I_2) = \frac{10679}{57600}$$

- Para $k = 3, 4, 5, 6$ ($\pi_3 = \pi_4 = \pi_5 = \pi_6 = \frac{29}{120}$):

$$\text{Cov}(I_3, I_3) = \frac{29}{120}\left(1 - \frac{29}{120}\right) = \frac{29}{120} \cdot \frac{91}{120} = \frac{2639}{14400}$$

$$\text{Var}(I_3) = \frac{29}{120}\left(1 - \frac{29}{120}\right) = \frac{2639}{14400}$$

$$\text{Cov}(I_3, I_3) = \text{Var}(I_3) = \frac{2639}{14400}$$

En todos los casos, $\text{Cov}(I_k, I_k) = \text{Var}(I_k) = \pi_k(1 - \pi_k)$, como se esperaba.