Janusz Wywiał
Katedra Statystyki
Akademia Ekonomiczna w Katowicach

Plany losowania prób

Definicja 1. Planem losowania próby nieuporządkowanej s nazywamy rozkład prawdopodobieństwa *P(s)* określony na przestrzeni , który dla każdej $s \in \Omega$ spełnia warunki:

$$P(s) \geq 0 \text{ i } \sum_{s \in \Omega} P(\underline{s}) = 1$$

Określamy zbiór:

$$A(k_1,...,k_r) = \{s: k_i \in s, \text{ dla } i=1,...,r\}$$

Prawdopodobieństwo doboru do próby o ustalonej liczebności k-tego elementu populacji *(k=1,...,N)* określa wyrażenie:

$$\pi_k = \sum_{s \in A(k)} P(s), \; A(k) = \{s: k \in s\}.$$

Prawdopodobieństwa inkluzji rzędu drugiego, czyli, że jednocześnie k-ty oraz t-ty elementy populacji należą do próby *(k=1,...,N, t=1,...,N* oraz *k≠t)*, są postaci:

$$\pi_{kt} = \sum_{s \in A(k,t)} P(s), \; A(k,t) = \{s: k \in s, t \in s, k \neq t\}.$$

$$\sum_{k=1}^{N} \pi_k = n \quad \text{i} \quad \sum_{t=1}^{N} \sum_{\substack{k=1 \\ k \neq t}}^{N} \pi_{k,t} = n(n-1).$$

Schemat losowania próby:

$$p(k_1) \prod_{i=2}^{n} p(k_i | k_{i-1},...,k_1) = P(\underline{s})$$

przy czym k_i=1,...,N oraz i=1,...,n.

Definicja 2. [Cassel i in. (1977)]: Mechanizm losowania realizujący plan wyboru elementów populacji do próby $\underline{s}$ według danych powyższym wzorem prawdopodobieństw warunkowych nazywamy *schematem losowania* próby.

Schemat losowania próby umożliwia więc techniczną realizację założonego planu losowania próby.

Twierdzenie 1. [T.V.H. Rao (1962)]: Dla każdego planu próbkowania *P(s)* istnieje przynajmniej jeden schemat losowania próby realizujący tenże plan.

Przykład 1. Plan zwrotnego losowania prostej próby uporządkowanej ma postać:

$$\bigwedge_{\underline{s}\in\underline{\Omega}} P_1(\underline{s}) = N^{-n} \qquad (1)$$

$$\pi_i^{(1)} = 1-(1-N^{-1})^n \qquad (2)$$

W szczególności plan losowania próby dwuelementowej z pięcioelementowej populacji ma postać: $\bigwedge_{\underline{s}\in\underline{\Omega}} P_1(\underline{s})=0.04$.

Przykład 2: Plan bezzwrotnego losowania (nie uporządkowanej) próby prostej o liczebności n elementów z populacji N-elementowej ma postać:

$$\bigwedge_{s\in\Omega} P_3(s) = \frac{1}{\binom{N}{n}} \qquad (3)$$

Prawdopodobieństwa wyboru każdego elementu populacji do próby są takie same i wynoszą:

$$\pi_i^{(3)} = \frac{n}{N}, \quad \textbf{\textit{dla i=1,...,N}} \qquad (4)$$

Przykład 3: Schemat losowania próby realizujący plany *$P_2(\underline{s})$* lub *$P_3(s)$* polega na wyborze bezzwrotnym kolejnych elementów do próby, tak aby prawdopodobieństwo wylosowania elementu o numerze k_i w i-tym losowaniu pod warunkiem, że wcześniej już wybrano elementy *$k_1,...,k_{i-1}$* wynosiło:

$$p(k) = \frac{1}{N}\text{, } k=1,...,N; \qquad p(k_i | k_{i-1},...,k_1) = \frac{1}{N-i+1}\text{, } i=2,...,n$$

Przykład 4. Plan Lahiriego bezzwrotnego losowania nieuporządkowanej próby proporcjonalny do sumy wartości cechy dodatkowej obserwowanej w tej próbie ma postać:

$$P(s) = \frac{(n-1)!(N-n)!}{(N-1)!}\frac{\sum_{i\in s} x_i}{\sum_{i\in U} x_i} = \binom{N}{n}^{-1}\frac{\overline{x}_s}{\overline{x}}.$$

$$\pi_k = \frac{N-n}{N-1}\frac{x_k}{N\overline{x}} + \frac{n-1}{N-1},$$

$$\pi_{kl} = \frac{n-2}{N-2}\left[\pi_k + \pi_l\right] - \frac{n(n-1)}{(N-1)(N-2)}.$$

W końcu prawdopodobieństwa warunkowe określające schemat losowania próby mają postać (Wywiał (1991)):

$$p(k_r \mid k_{r-1},...k_1) = \frac{(N-n)\sum_{i=1}^{r} x_{k_i} + N(n-r)\bar{x}}{(N-n)\sum_{i=1}^{r-1} x_{k_i} + N(n-r+1)\bar{x}} \frac{1}{N-r}$$

.

Lahiri (1951) zaproponował prostszy schemat losowania. Pierwszy element jest los. z prawd. prop. do:

$$p(k_1) = \frac{x_{k_1}}{N\bar{x}}.$$

k_1=1,2,...,N.

Przykład 5: Elementom populacji U={1,2,3,4,5} są przyporządkowane następujące wartości zmiennej y: y(1)=8, y(2)=1, y(3)=3, y(4)=1, y(5)=3. Średnia w populacji wynosi $\bar{y} = 3.2$, a wariancja $v_* = 8.2$. Niech plan bezzwrotnego losowania próby trój-elementowej określają wyrażenia:

$$P_1(1,2,3)=P_1(1,2,5)=P_1(1,3,4)=P_1(1,4,5)=\frac{1}{6}$$

$$P_1(1,2,4)=P_1(2,3,4)=P_1(2,3,5)=P_1(3,4,5)=\frac{1}{12}$$

$$P_1(1,3,5)=P_1(2,4,5)=0$$

Prawdopodobieństwa inkluzji:

$$\pi_1^{(1)} = \frac{3}{4}, \qquad \pi_2^{(1)} = \frac{7}{12}, \qquad \pi_3^{(1)} = \frac{7}{12}, \qquad \pi_4^{(1)} = \frac{7}{12}, \ \pi_5^{(1)} = \frac{1}{2},$$

$$\pi_{1,2}^{(1)} = \frac{5}{12}, \qquad \pi_{1,3}^{(1)} = \frac{4}{12}, \qquad \pi_{1,4}^{(1)} = \frac{5}{12}, \qquad \pi_{1,5}^{(1)} = \frac{4}{12}, \ \pi_{2,3}^{(1)} = \frac{4}{12},$$

$$\pi_{2,4}^{(1)} = \frac{2}{12}, \qquad \pi_{2,5}^{(1)} = \frac{3}{12}, \qquad \pi_{3,4}^{(1)} = \frac{4}{12}, \qquad \pi_{3,5}^{(1)} = \frac{2}{12}, \ \pi_{4,5}^{(1)} = \frac{3}{12},$$

Wyznaczyć wartość wariancji estymatora Horvitza-Thompsona.