Sampling Space and Statistics on a Sample as an Element of a Semigroup
Žaneta Popeska
Faculty of Natural Sciences and Mathematics, Skopje, Republic of Macedonia
E-mail: zaneta@pmf.ukim.edu.mk

1. Sampling design

Let $B=\{b_1, b_2, ..., b_n\}$ be a finite population, $S=S(B)$ or $S=\langle B \rangle$, a semigroup generated by $B$, and $U(B)$ a free semigroup generated by $B$. The elements of $U(B)$ will be denoted by $\sigma$, $\tau$, $\omega$, ..., and the elements of a semigroup $S(B)$ by $s, t, u, ...$ We define a contents of $\sigma \in U$, by $C(\sigma)=\{b \mid b \in \sigma\}$ and a length of $\sigma$ by $L(\sigma)=n$ if $\sigma=b_1b_2...b_n$, for $b \in B$.

If $S(B)$ is a semigroup generated by $B$, then there exists a unique homomorphism (which is epimorphism) $\psi: U(B) \rightarrow S(B)$ for which $\psi(b)=b$ for each $b \in B$. From now on we will use the symbol $\approx$ only for this epimorphism. We define a contents of $s \in S$, by $C(s)=\{C(\sigma) \mid \psi(\sigma)=s\}$ and a length of $s$ by $L(s)\{n \mid n=L(\sigma), \psi(\sigma)=s\}$.

Let $p:S(B)\rightarrow [0,1]$ be a real function.

Given $B, S(B)$, the triple $P=(B, S(B), p)$ is called a sampling design if

i) $\Sigma_1 p(s)=1$

ii) $\forall b \in B, \exists s \in S(B)$, such that $b \in s$ and $p(s)>0$.

The semigroup $S(B)$ is called a sampling set, and the real function $p$ - a design function. The elements of $S$ will be called $S$ - samples on $B$. We say that the unit $b \in B$ belongs to a sample $s \in S$, and write $b \in s$ if $s$ can be written as a product of elements of $B$ in which $b$ appears, i.e. $s=b_1b_2...b_n \in B$ and $\exists i, 1 \leq i \leq n, b=b_i$. In other words $b$ is a sample if and only if (iff) there is $\sigma \in U$ such that $b \in C(\sigma)$ and $\psi(\sigma)=s$.

2. Sampling space

Let $B$ be a finite population and $Y:B \rightarrow \mathbb{R}$ a mapping. The vector $Y=(Y_1, Y_2, ..., Y_N)$, where $Y_i=Y(b_i)$ is called a population parameter. Let $U$ be the free semigroup and $S$ a semigroup generated by $B$ and $\psi:U \rightarrow S$ the epimorphism defined earlier. If we consider $s \in U$ as a mapping $s: b_1 \rightarrow \sigma(b_1), \ldots, b_n \rightarrow \sigma(b_n))$. For $s \in S$ we define $Y_s=\{Y_\sigma \mid \sigma \in \psi^{-1}(s)\} \} and (s:Y)=\{(\sigma, Y_\sigma) \mid \sigma \in \psi^{-1}(s))\}$. Let $\Delta(S)=\{(s, y) \mid s \in S, y \in \mathbb{R}^n, \exists y \in \mathbb{R}^N, y \in Y_s\}$ and $\Delta^*(S)=\{(s, y) \mid s \in S, y \in \mathbb{R}^N\}$. We define a relation $\sim$ on $\Delta(U)$ by $(s, y) \sim (\tau, z)$ if $\psi(s)=\psi(\tau)$ and $s \tau \Rightarrow y=z)$. This relation satisfies the following properties:

1° $(s, y) \sim (\tau, z) \Rightarrow (\tau, z) \sim (s, y)$.

2° $\forall (s, y) \in \Delta(U), (s, y) \sim (s, y)$.

3° If $\psi(s)=\psi(\tau)$ and $y=Y_\sigma, z=Y_\tau$ for $y \in \mathbb{R}^N$, then $(s, y) \sim (\tau, z)$.

4° For any $\sigma$, $\tau$ for which $\psi(\sigma)=\psi(\tau)$, holds $(s, Y_\sigma) \sim (\tau, Y_\tau) \forall y \in \mathbb{R}^N$.

5° The relation $\sim$ does not being transitive.

6° If $\psi(s)=\psi(\tau) \Rightarrow C(\sigma)=C(\tau)$ then the relation $\sim$ is an equivalence relation.

Let $\approx$ be the transitive closure of $\sim$. Then $\approx$ is an equivalence relation on $\Delta$. The set $\Delta^*(S)=\Delta(U)/\approx$ is called a data space over the semigroup $S$ and each equivalence class $(s, y)^\approx$ is a data and will be denoted by $[s, y]$.

If for each $s \in S$, $y \in \mathbb{R}^n$, with $\exists L(s)$, for which there is $s \in \psi^{-1}(s)$, such that ker $s \subset ker y$ we define a set $[s, y]=\{[\tau, y] \mid \tau \in \psi^{-1}(s)\}$ then $\Delta^*(S)=\bigcup_{m \in S} [s, y]$. For each $s \in \psi^{-1}(s)$, $(s, Y_s) \subset [s, Y_\sigma]$.
We say that a data \( d = [\sigma, y] \in \Delta^* (S) \) is *consistent* with the parameter \( Y \in \mathbb{R}^N \) if and only if \( (\sigma, Y \sigma) \in [\sigma, y] \). For each \( Y \in \mathbb{R}^N \) we define \( p_Y^*: \Delta^* (S) \to [0, 1] \) by

\[
p_Y^* (s : Zs) = \begin{cases} 
p(s) & \text{if } Zs = Ys \\
0 & \text{otherwise}
\end{cases}.
\]

8° The mapping \( p_Y^* \) is well defined and induces a probability measure on the algebra of subsets of \( \Delta^* (S) \) defined by

\[
P_Y^* (A) = \sum_{(s,Ys) \in A} p_Y^* (s : Zs) = \sum_{(s,Ys) \in A} p(s).
\]

Now we can define a mapping \( p_Y: \Delta^* (S) \to [0, 1] \) with

\[
p_Y ([\sigma, y]) = p_Y^* \{ (s : V) | (s : V) \subseteq [\sigma, y] \}.
\]

9° \( p_Y^* ([\sigma, y]) = \begin{cases} 
p(s) & (\sigma, y) = (\sigma, Y \sigma) \\
0 & \text{otherwise}
\end{cases} = \begin{cases} 
p((\sigma, y)) & [\sigma, y] = [\sigma, Y \sigma] \\
0 & \text{otherwise}
\end{cases}.
\]

10° The mapping \( p_Y \) induces a probability measure on the algebra of subsets of \( \Delta^* (S) \), \( P_Y \), defined for each \( A \subseteq \Delta^* (S) \) by:

\[
P_Y (A) = \sum_{[\sigma, y] \in A} p_Y ([\sigma, y]).
\]

For each \( Y \in \mathbb{R}^N \) we define subset \( \Delta^*_Y \subseteq \Delta^* (S) \) with

\[
\Delta^*_Y = \{ [\sigma, y] | [\sigma, y] \in \Delta^* (S), p_Y ([\sigma, y]) > 0 \}.
\]

11° The set \( \Delta^*_Y \) is at most countable set.

3. Statistics

Any mapping \( F: \Delta^* \to \mathbb{R}^m \) is a statistics over the design \( P = (B, S(B), p) \). Now, having in mind that the set \( \Delta^*_Y \) is at most countable and so is \( F(\Delta^*_Y) \) one can define mathematical expectation of \( F \), \( E_Y (F) \) by

\[
E_Y (F) = \sum_{w \in F(\Delta^*_Y)} w \cdot P_Y (F^{-1} (w))
\]

Then

\[
E_Y (F) = \sum_{s \in S \text{ for some } \tau \epsilon (\Delta^*_Y)} F([\tau, Y \tau]) \cdot p(s).
\]

13° If \( F \) and \( G \) are statistics on \( P = (B, S(B), p) \), for each \( Y \in \mathbb{R}^N \), and \( \alpha, \beta \in \mathbb{R} \) if the both sides exist,

\[
E_Y (\alpha F + \beta G) = \alpha E_Y (F) + \beta E_Y (G).
\]

REFERENCES


Zaneta Popeska, 1994 Quotient sampling design, Macedonian academy of sciences & arts, XV 2, 37-48


Jaroslav Hajek, 1981 Sampling from a finite population, Marcel Dekker