Title: Introduction to the book Graph Sampling (CRC Press)

Finite population sampling has found numerous applications in the past century. The validity of sampling inference of real populations, hence its universal applicability, derives from the *known* sampling probabilities associated with the sample, “irrespectively of the unknown properties of the target population studied” (Neyman, 1934).

A *valued graph* is a more powerful representation, which allows one to incorporate the connections among the units in addition. The underlying structure is a graph given as a finite collection of nodes (for units/entities) *and* edges (for connections). Attaching measures to the nodes or edges or both yields a valued graph. Many technological, socio-economic and biological phenomena exhibit a graph structure that may be the central interest of study, or the edges may effectively provide access to those nodes that are the primary targets. Either way, graph sampling is a statistical approach to study real graphs.

Just like finite population sampling, *graph sampling* is universally applicable based on exploring the variation over all possible subgraphs (i.e. sample graphs), which can be taken from the given population graph, according to a specified method of probability sampling. On the one hand, graph sampling encompasses finite population sampling in the sense that, apart from element and cluster sampling, all the so-called “unconventional” techniques that make use of the connections between the relevant units can be more effectively studied as special cases of graph sampling, such as indirect, network, adaptive cluster, line-intercept or spatial sampling. On the other hand, graph sampling theory yields a rigorous approach to genuine graph problems, where the interest of estimation is given directly as graph parameters, allowing one to devise and make use of various *probabilistic* breadth- or depth-first non-exhaustive graph traversal algorithms.

The recently published book *Graph Sampling* establishes a theoretical framework that unifies the key elements in the pioneering works of Birnbaum & Sirken, O. Frank and S. Thompson, and develops new general sampling strategies including when one does not observe all the possible ways by which a given subgraph can be sampled. There is a wide range of application areas, such as social networks, internet, epidemiological and biomedical studies, graph-based machine learning, environmental and spatial statistics.