Echantillonnage de réseaux : une relecture de S.K.Thompson avec une nouvelle présentation et quelques nouveautés

Jean-Claude DEVILLE[1]

On s'intéresse à une population dont on sait qu'elle possède une structure de voisinages, c'est-à-dire qu'elle est munie d'un graphe non orienté (et réflexif). Dans cette population on cherche à mesurer une variable *y* dont on sait qu'elle est essentiellement intéressante dans un certain domaine, éventuellement défini par une condition sur *y* . Si un individu appartient au domaine, il est plausible qu'un ou plusieurs de ses voisins y soient aussi, d'où l'idée d'augmenter un échantillon initial par les voisins des individus qui sont dans le domaine, et ainsi de suite. Exemple : recherche de certaines roches en géologie, chasse aux rupins en statistique sociale, recherche dur le usage de drogues etc.

Cette procédure aboutit à des estimateurs sans biais, de variance calculable et estimable pour de nombreux plans de sondage comme le sondage aléatoire simple avec où sans remise, mais aussi pour les sondages à probabilités inégales avec remise où sans remise à entropie maximale (poissonniens ou poisonniens conditionnels).

Tous ces estimateurs sont améliorables avec la technique de Rao-Blackwell, et, once more, ont une variance estimable.

On examinera aussi les liens de cette technique avec l'échantillonnage indirect (ou méthode de partage des poids).

Bibliographie (entre autre) :

Steven K. Thompson: Adaptive Cluster Sampling, *Journal of the American Statistical Association, Vol. 85, No. 412 (Dec., 1990), pp. 1050-1059*

[1] Ecole Nationale de la Statistique et de l'Analyse de l'Information, Laboratoire de Statistique d'Enquête\crest, deville@ensai.fr