

# SÉLECTIONNER DEXTREMENT L'ÉCHANTILLON !

Yves Tillé<sup>1</sup>

<sup>1</sup> *Institut de statistique, Université de Neuchâtel, Avenue de Bellevaux 51, 2000 Neuchâtel, Suisse, yves.tille@unine.ch*

**Résumé.** Une stratégie d'estimation est composée d'un plan de sondage et d'un estimateur. Le choix de l'échantillon est ainsi la première étape d'une stratégie. Dans l'approche basée sur le plan de sondage, un échantillon est choisi aléatoirement parmi un ensemble d'échantillons. Les résultats obtenus à partir de l'échantillon sont ensuite extrapolés à la population en se basant sur la distribution des estimateurs induite par le plan de sondage. Dans l'approche basée sur un modèle, on se base sur un modèle qui régit la population sans tenir compte de la procédure de sélection de l'échantillon. Au moyen de ce modèle, on prédit alors les valeurs qui n'ont pas été échantillonnées.

Bien que ces deux approches soient complètement distinctes, elles aboutissent parfois aux mêmes estimateurs et aux mêmes intervalles de confiance. Dans l'approche basée sur le plan, le tirage aléatoire simple coïncide avec un modèle de régression ne contenant qu'une constante. Nous montrerons que l'étude des cas où les approches coïncident permet d'établir une méthodologie pour choisir dextrement l'échantillon. Le plan de sondage qui fait correspondre les deux approches est celui qui est le plus précis.

Cette méthodologie met en avant l'importance de modéliser l'ensemble des données disponibles pour pouvoir optimiser le choix de l'échantillon. Toute information peut ainsi être utile pour améliorer le choix de l'échantillon. Plusieurs exemples seront présentés. L'accent sera mis sur les modèles et les échantillons spatiaux et leurs nombreuses applications en statistique publique ou dans les sciences de l'environnement.