

## Le problème des non-réponses dans les enquêtes

Le problème des non-réponses est une des grandes difficultés rencontrées dans le traitement statistique des enquêtes.

Prenons un exemple simple : parmi les 1000 étudiants d'un institut, 750 sont satisfaits de leur formation (soit 75%) et les autres sont non satisfaits. Une enquête de satisfaction auprès de l'ensemble des étudiants ne permet de recueillir que 300 réponses. Parmi ceux qui ont répondu, 100 se déclarent satisfaits (soit 33% seulement) et les autres se déclarent non satisfaits. On voit comment on peut observer un taux de satisfaction parmi ceux qui ont répondu à l'enquête (33%) qui ne reflète pas le taux réel dans la population cible de l'enquête (75%).

	réponses	non réponses	Ensemble
satisfaits	100	650	750
non satisfaits	200	50	250
Ensemble	300	700	1000

Dans la pratique, on ne connaît justement pas le taux réel et on souhaiterait le connaître grâce à une enquête. Si, pour les personnes enquêtées, le fait de répondre dépend de la question de l'enquête (ici la satisfaction), on peut avoir une distorsion (on appelle cela un *biais*) entre les résultats de l'enquête et les résultats que l'on aurait obtenus si tout le monde avait répondu.

Ce problème de non-réponses se pose aussi bien pour des enquêtes auprès de toute la population (enquêtes exhaustives) que pour des enquêtes auprès d'un échantillon de la population (enquêtes par sondage).

Une enquête par sondage aléatoire (procédure aléatoire de sélection de l'échantillon) avec un faible taux de non-réponses peut donner de bien meilleurs résultats qu'une enquête exhaustive avec un taux important de non-réponses. En effet, les statisticiens contrôlent beaucoup mieux les marges d'incertitude dues à l'échantillon aléatoire que les biais dus aux non-réponses.