

ÉCHANTILLON BIAISÉ (M1)

(18 / 03 / 2020, © Monfort, Dicostat2005, 2005-2020)

(i) On appelle **échantillon biaisé** un échantillon obtenu selon un **plan de sondage « biaisé »**, ie conduisant à des estimateurs biaisés des **paramètres d'intérêt** : **moyenne**, total, **proportion**, etc (cf **estimateur sans biais**).

De telles situations correspondent donc à des **biais de sélection**.

Ainsi, dans un cas extrême, si A est un **N-échantillon iid** extrait d'une **population finie** Ω selon un **plan de sondage** Π , et si Π est défini en sorte qu'il conduit toujours à tirer le même élément $\omega_0 \in \Omega$ (ie si $\Pi = \delta_{\omega_0}$, **mesure de DIRAC** placée en ω_0), tout N-échantillon A sera de la forme $A = (\omega_0, \dots, \omega_0) \in \Omega^N$ (appartenance à la diagonale de Ω^N) et conduit toujours à un vecteur (échantillon observé) de la forme (y_0, \dots, y_0) , avec $y_0 = \eta(\omega_0)$. Par suite, si Y n'est pas de la forme (Y_0, \dots, Y_0) , avec $Y_0 = y_0$ (ie si η n'est pas distribué avec la même valeur y_0 sur Ω), le plan sera biaisé.

(ii) Les sondages non aléatoires (eg **sondage par choix raisonné**, **sondage par quotas**) conduisent souvent à des biais, mais ceux-ci ne sont pas nécessairement importants.