

ÉCHANTILLONNAGE

I - Échantillons

1°) - Échantillons de type "sondage"

Quand on doit décrire une population finie comportant un grand nombre d'individus, on ne peut pas ou on ne veut pas, en général pour des raisons économiques, en faire une étude exhaustive. Les observations ne portent alors que sur un nombre restreint d'individus à sélectionner selon un protocole expérimental. Les individus sélectionnés constituent un **échantillon**, leur nombre est la **taille** de l'échantillon.

Deux types de tirages d'échantillons aléatoires issus de populations finies peuvent être envisagés :

- Les tirages avec remise supposent que tout individu choisi est remplacé dans la population, après observation du caractère. Il peut donc être sélectionné plusieurs fois.
- Les tirages sans remise supposent que tout individu choisi ne peut jamais être repris une deuxième fois.

Remarque : Dans le cas d'échantillons constitués avec remise, les calculs sont plus simples. Les différences entre les résultats de calculs de probabilités obtenus pour les deux types d'échantillons sont négligeables lorsque l'effectif de la population est très grand et que le taux de sondage est faible (en général inférieur à 0,1). Dans ces conditions, pour les calculs de probabilités dans le cas d'échantillons constitués sans remise, on peut ainsi admettre que les échantillons sont constitués avec remise pour obtenir des valeurs approchées des résultats.

2°) - de type "expérimental"

Les observations d'un caractère, qualitatif ou quantitatif, sur les individus d'un échantillon de type "sondage" avec remise constituent un échantillon de type "expérimental".

Exemple : Couleurs de boules tirées d'une urne, intentions de vote d'un échantillon d'électeurs, défauts des pièces d'un échantillon issu d'une production lors d'un contrôle de qualité, mesures sur des pièces d'un échantillon issu d'une production...

On obtient aussi un échantillon de type "expérimental" lorsqu'on travaille sur des résultats d'expériences aléatoires répétées.

Exemple : Résultats de lancers successifs de pièces, de dés, échantillons de rendements de céréales sur des parcelles de terrain...

3°) - Tirages "au hasard"

L'expression *tirage au hasard* ne recouvre la notion d'équiprobabilité que pour des échantillons de type "sondage". Dans ce cas, la distribution de probabilité sur l'ensemble des échantillons est la loi équirépartie qui rend compte du fait que tous les échantillons d'individus issus de \mathcal{P} ont la même probabilité d'être choisis. Si ensuite on s'intéresse à une propriété C des individus, la distribution de probabilités de la fréquence de C dans les échantillons n'est pas équirépartie et elle dépend de la distribution de C dans \mathcal{P} .

Pour des échantillons de type "expérimental", la distribution de probabilités des résultats dépend de leurs probabilités d'observation lors de la réalisation de l'expérience aléatoire.

II - Échantillonnage

1°) - Échantillonnage

L'échantillonnage est l'étude des distributions de probabilités de variables aléatoires définies sur l'ensemble des échantillons (fréquence d'échantillonnage, moyenne d'échantillonnage, variance d'échantillonnage...).

La théorie de l'échantillonnage est utilisée en particulier dans deux situations :

- pour estimer un paramètre d'une population (pourcentage, moyenne, écart-type...) à partir de la valeur correspondante observée sur un échantillon (théorie de l'estimation) ;
- pour décider (avec un risque d'erreur fixé à l'avance) si, par exemple, la différence entre une valeur de référence et une moyenne observée sur un échantillon est due au hasard ou si elle est significative (théorie des tests).

2°) - Comment prélever un échantillon ?

Lors d'une prise de décision à partir d'un échantillon, pour que les résultats de la théorie des probabilités et de l'échantillonnage s'appliquent, il est important que l'échantillon soit aléatoire, c'est-à-dire prélevé selon une procédure aléatoire préétablie.

Pour cela, on peut entre autres moyens, utiliser une table de chiffres au hasard : on affecte un numéro d'identification à chacun des individus de la population, puis on sélectionne un échantillon de numéros en utilisant une telle table.