Chap 9 :

Estimations

I. Principe

1) Objectifs

Lorsqu'on cherche à déterminer le poids moyen des français il est bien sûr hors de question de peser ***tous*** les français. Par contre en choisissant ***judicieusement*** un petit nombre de personnes il est possible d'en obtenir une estimation.
On pratique beaucoup ces estimations dans les milieux industriels plutôt que d'étudier la population entière soit parce que cela prendrait trop de temps, soit parce que cela reviendrait trop cher, soit encore parce que cela serait illogique (contrôle qualité detruisant les pièces ...).

2) Loi des grands nombres

On considère une expérience aléatoire avec un événement A de probabilité p. Si on répète n fois et de manière aléatoire cette expérience on regarde la fréquence d'apparition f_n de l'événement A.

Théorème 1 : On obtient, avec une probabilité aussi grande que l'on veut, une fréquence f_n (pour n expériences indépendantes) aussi proche que l'on veut de p, lorsque n est suffisamment grand.

II. Estimation ponctuelle

Dans cette partie on fait les calculs sur un échantillon prélevé au hasard sur la population dont on cherche à faire l'étude.

1) Fréquence

Définition 1 : Pour estimer la fréquence p ***inconnue*** d'un caractère dans une population, on prélève un échantillon et on calcule la fréquence d'apparition de ce caractère dans l'échantillon.
Cette fréquence d'apparition est une estimation ponctuelle de la fréquence p.

2) moyenne

Définition 2 : Pour estimer la moyenne m ***inconnue*** d'une population, on prélève un échantillon et on calcule la moyenne de cet échantillon.
Cette moyenne d'échantillon est une estimation ponctuelle de la moyenne m.

3) Ecart-type

Pour l'écart-type, le calcul sur un échantillon donne une valeur potentiellement sous-évaluée, c'est pourquoi on corrige légèrement l'estimation :

Définition 3 : Pour estimer l'écart-type σ ***inconnu*** d'une population, on prélève un échantillon et on calcule l'écart-type σ' de cet échantillon. Le nombre $\sqrt{\dfrac{n}{n-1}}\,\sigma'$ est une estimation ponctuelle de l'écart-type σ.

III. Estimation par intervalle de confiance

1) Moyenne

On considère une population de moyenne m ***inconnue*** et d'écart-type σ qu'on suppose *connu*.
Si n est assez grand, la variable aléatoire X qui à chaque échantillon de n éléments associe sa moyenne suit approximativement la loi $\mathcal{N}\left(m; \dfrac{\sigma}{\sqrt{n}}\right)$.
On peut donc, à l'aide de la loi $\mathcal{N}(0;1)$, trouver un intervalle $[a;b]$ tel que $P(a \leqslant X \leqslant b) = 0,95$ par exemple.
Cet intervalle est appelé l'intervalle de confiance de la moyenne m avec le coefficient de confiance 0,95.
On a de manière plus générale :

Théorème 2 : L'intervalle $\left[\overline{x} - t\dfrac{\sigma}{\sqrt{n}}\,;\overline{x} + t\dfrac{\sigma}{\sqrt{n}}\right]$ est l'intervalle de confiance de la moyenne m de la population avec le coefficient de confiance $2\Pi(t) - 1$.
$\overline{x}$ est la moyenne de l'échantillon considéré.

Remarque : On ne peut déterminer un intervalle de confiance que si ***on connait déja*** l'écart-type σ.

Cas particuliers

- Pour le coefficient de confiance 0,95 il faut prendre t=1,96 (d'après le formulaire).
- Pour le coefficient de confiance 0,99 il faut prendre t=2,58 (d'après le formulaire).

2) Fréquence

On considère une population qui contient avec une fréquence p des individus ayant un certain caractère.
Si n est assez grand, la variable aléatoire F qui à chaque échantillon de n éléments associe la fréquence d'apparition des individus ayant ce caractère suit approximativement la loi $\mathcal{N}\left(p; \sqrt{\dfrac{p(1-p)}{n}}\right)$.
De manière analogue au cas pour une moyenne il est possible de déterminer un intervalle de confiance de la fréquence p avec un coefficient de confiance choisi.

Théorème 3 : L'intervalle $\left[f - t\sqrt{\dfrac{f(1-f)}{n-1}}; f + t\sqrt{\dfrac{f(1-f)}{n-1}}\right]$ est l'intervalle de confiance de la fréquence p avec le coefficient de confiance $2\Pi(t) - 1$.
f est la fréquence des individus ayant la caractère dans l'échantillon considéré.

Cas particuliers
Pour les coefficients de confiance 0,95 et 0,99 les valeurs de t sont encore 1,96 et 2,58.