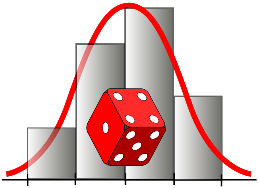
Estimation paramétrique ponctuelle & distribution d'échantillonnage



Module : Techniques d'estimation pour l'ingénieur

Plan

1. Introduction
   - Notion d'échantillonnage
   - Notion d'estimation paramétrique
2. Estimation paramétrique ponctuelle
   - Qualités des estimateurs
   - Méthodes d'estimation ponctuelle
     - Estimateur par méthode des moments (EMM)
     - Estimateur par la méthode du maximum de vraisemblance (EMV)
4. Distribution d'échantillonnage
   - Distribution échantillonnale de la moyenne $\overline{X_n}$
   - Distribution échantillonnale de la variance S_n^2
   - Distribution échantillonnale de la proportion F

Problématique ?

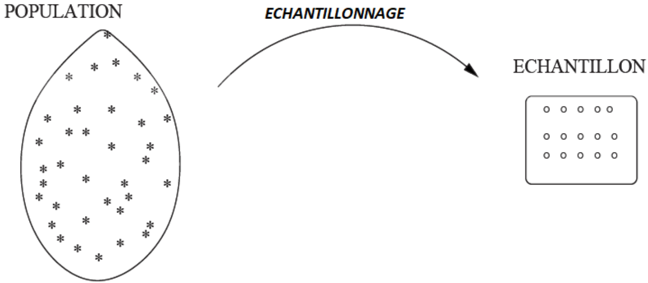
Question : combien y a-t-il de personnes atteintes de troubles de la vue parmi les conducteurs automobiles en Tunisie?

Réponse : 10%? 40%? 75%?

- Il est impossible de les compter toutes en examinant toute la population des conducteurs tunisiens.
- Il va être nécessaire d'utiliser une procédure particulière (l'échantillonnage) et des méthodes statistiques pour estimer la précision du résultat (incertitude) .

Terminologie

- **Population :** l'ensemble que l'on observe et qui sera soumis à une analyse statistique, chaque élément de cet ensemble est un individu ou unité statistique.
- **Echantillon :** Un sous-ensemble de la population étudiée.



Pourquoi prendre un échantillon ?

- **Le coût :** analyser toute la population coûte trop cher.
- **Le temps :** on souhaite obtenir l'information le plus rapidement possible.
- **L'impossibilité :** lorsque la population est infinie.

Types d'échantillonnage

Les techniques d'échantillonnage peuvent être regroupées en deux grandes familles :

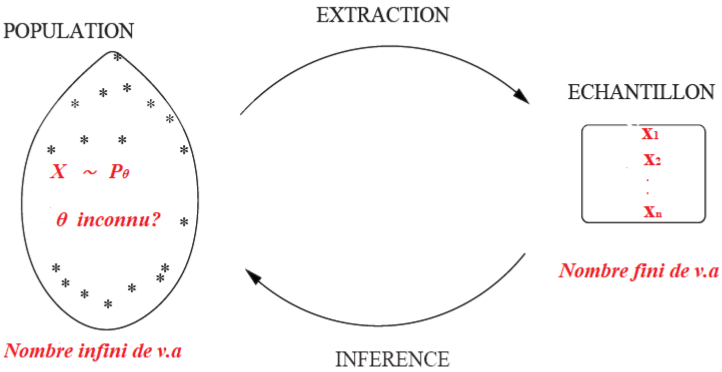
- Echantillonnage sur la base des méthodes empiriques :
  La méthode des quotas (se base sur la composition de la population pour certains critères) est la plus utilisée.
- Echantillonnage aléatoire simple :
  Tous les échantillons possibles de même taille ont la même probabilité d'être choisis et tous les éléments de la population ont une chance égale de faire partie de l'échantillon.

Contexte général

- Dans toute la suite du cours, on se place dans le cadre d'un échantillonnage aléatoire simple.
- On se propose d'étudier un caractère statistique X sur une population de taille N. On s'intéresse à la valeur inconnue d'un paramètre bien précis θ de la loi de X :
  - Si le caractère est quantitatif : Espérance, variance.
  - Si le caractère est qualitatif : Proportion.
- Si on prélève n individus dans cette population, on obtient n valeurs $x_1, x_2, ..., x_n$.
- L'observation x_i peut être considérée comme une observation d'une variable aléatoire X_i de même loi que X.

La situation peut être résumée par le schéma illustratif suivant :

Notion d'échantillonnage



La stratégie consiste à utiliser l'information obtenue sur un échantillon de taille n pour déduire de l'information sur la population : on extrait un échantillon de la population, on l'analyse et on infère sur la population.

Notion d'échantillonnage

On décris habituellement un échantillon ou une population à l'aide de mesures telles que le nombre d'unité (la taille), la moyenne, l'écart-type et le pourcentage (proportion).

Les mesures qu'on utilise pour décrire une population sont **des paramètres**.

Un paramètre est une caractéristique de la population.

Les mesures qu'on utilise pour dérire un échantillon sont appelées **des statistiques**.

Une statistique est une caractérisatique de l'échantillon.

esprit Se former autrement

Activité introductive :

"Comment on peut vérifier expérimentalement qu'une pièce de monnaie est équilibrée ?"

On jette cette pièce de monnaie $n-$fois et on associe pour chaque lancée une variable aléatoire $X \sim \mathcal{B}(p)$: Bernoulli de paramètre $p \in]0,1[$.
La $i^{\text{ème}}$ lancée correspond à la variable aléatoire X_i définie comme suit:

$$X_i \sim \mathcal{B}(p) \quad \Longrightarrow \quad \overset{\Omega}{\left\{ \begin{array}{c} \omega_1 = P \\ \omega_2 = F \end{array} \right\}} \quad \xrightarrow{X_i} \quad \overset{X_i(\omega)}{\left\{ \begin{array}{c} 1 \\ 0 \end{array} \right\}} \quad ; \quad \forall 1 \leq i \leq n$$

Notion d'estimation Paramétrique

Alors les X_i sont iid : indépendantes et de même loi que $X \sim \mathcal{B}(p)$. C'est ainsi, afin de vérifier si cette pièce de monnaie est équilibrée ou non, il suffit de vérifier si $p = \frac{1}{2}$ ou non.
Comme l'inconnu p est l'espérance de la loi de Bernoulli, $X \sim \mathcal{B}(p)$;

$$\mathbb{E}[X] = p$$

L'idée ici est de l'estimer ou de l'approcher par la variable aléatoire suivante:

$$\frac{S_n}{n} = \frac{X_1(\omega) + X_2(\omega) + \cdots + X_n(\omega)}{n}$$

Estimer un paramètre consiste à chercher une valeur approchée en se basant sur les résultats obtenus à partir d'un échantillon aléatoire.

Notion d'estimation paramétrique

Dans ce cours, on s'intéresse à estimer certaines caractéristiques statistiques (moyenne, variance, proportion) d'une certaine loi par différentes méthodes, où cette loi théorique on la connait mais on ignore son paramètre.

$\Longrightarrow$ C'est le cadre d'une **estimation paramétrique unidimensionnelle**.

On cite deux types d'estimations paramétriques:

- Estimation paramétrique ponctuelle : estimation est donnée par une seule valeur.
  - (E.M.M) La méthode des moments
  - (E.M.V) La méthode de maximum de vraisemblance
- Estimation paramétrique par intervalle de confiance.
  - Intervalle centré
  - Intervalle décentré

esprit Se former autrement