**Université d'Evry Val-d'Essonne Paris-Saclay** **Année universitaire 2019-2020**
**L2 SDV - MSV31**

# TD5 - Estimation par intervalle et théorème central limite

---

**Exercice 1** *Élection présidentielle*

Deux candidats $A$ et $B$ s'affrontent au second tour de l'élection présidentielle de 2022. Une estimation du résultat est toujours annoncée avant la fin du dépouillement. On note $p$ la vraie proportion de vote pour $A$ et donc $1-p$ pour $B$ à la fin du dépouillement. On peut introduire la variable aléatoire $X$ de Bernoulli telle que

| $x$ | $A$ | $B$ |
|---|---|---|
| $\mathbb{P}(X=x)$ | $p$ | $1-p$ |

On fait deux hypothèses : (1) le choix du candidat est une variable indépendante de l'heure de la journée (le ratio votes pour $A$ par votes pour $B$ est une constante au cours de la journée) et (2) le nombre total de sondés reste toujours très inférieur au nombre total de votants.
À 13h, $n = 1000$ personnes ont été interrogées dans un sondage de sortie des urnes : l'estimation de $p$ est égale à $p_0 = 51.3\%$.

1. Construire un intervalle de confiance de niveau 95% autour de la proportion $p$ en utilisant le théorème central limite et l'approximation $p \approx p_0$.
2. Construire un intervalle de confiance de niveau 95% autour de la proportion $p$ en utilisant le théorème central limite mais cette fois sans l'approximation $p \approx p_0$.

   L'équipe de sondeurs étant la même toute la journée, chaque heure, $n$ soudages sont effectués à partir de 13h (inclus). On suppose que la proportion mesurée à 13h $p_0$ reste toujours la même en cours de la journée.
3. À combien doit être égal $n$ pour qu'une estimation à 20h (après $8n$ sondages) donne le nom du vainqueur avec une certitude de 95%. Utiliser seulement l'intervalle de confiance de la question 1.
4. Même question avec un niveau de confiance relevé à 99%.

5. Votre budget alloué au sondage vous place en capacité de faire au plus $n = 500$ sondages par heure. Déterminer l'écart $\pi$ limite en deçà duquel vous ne pouvez plus donner le vainqueur avec une confiance de 95%. L'écart limite définit l'intervalle d'indécision $[0.5-\pi, 0.5+\pi]$.

**Exercice 2** *L'enquête anonyme*

On cherche à connaître la proportion $p$ de personnes qui ne trient pas leur déchet par une enquête téléphonique. Comme une réponse honnête du sondé n'est pas évidente, on utilise la procédure suivante :

Chaque personne qui répond au téléphone joue à pile ou face avec une pièce équilibrée.
- Si elle obtient pile : elle doit répondre honnêtement par oui on non à la question "triez-vous le verre et les cartons ?"
- Si elle obtient face : elle doit lancer à nouveau la pièce et répondre honnêtement à la question

"Avez-vous obtenu face au 2ème tirage ?"

Que la réponse soit oui ou non, il est impossible de savoir à quelle réponse le sondé a répondu. On note $p$ la vraie proportion de personnes triant leur déchet dans la population. On note $\pi$ la proportion de oui obtenue à la fin du sondage.

1. Montrer que
$$\pi = \frac{p}{2} + \frac{1}{4}$$
2. Soit $X_i$ la variable aléatoire égale au résultat du $i$ème sondage parmi $n$ sondages. Quelle est la loi des $X_i$ ? et la loi de $X = X_1 + \cdots + X_n$ ?
3. Donner un estimateur de $\pi$ noté $\Pi$ puis un estimateur de $p$ noté $P$.
4. Calculer $\mathbb{E}[P]$ et $V(P)$.
5. En déduire un intervalle de confiance de niveau $1 - \alpha$ pour $p$.
6. Avec $n = 1000$ et $x = 600$ donner l'intervalle de confiance de niveau 95%. Quel est la conséquence de la confidentialité sur la longueur de l'intervalle ?

**Exercice 3** *Incertitudes*

Calculer l'incertitude (la variance) sur le résultat $R$ pour les formules suivantes connaissant la variance sur les variables indépendantes $A$, $B$ et $C$ :

1. $R = A \times B^2$
2. $R = A \times B + C^3$