Le travail sur les séries statistiques et les probabilités mené en classe de seconde se poursuit avec la mise en place de nouveaux outils. Les sciences et techniques industrielles et du laboratoire fournissent un large éventail de sujets d'étude. Cette partie est organisée selon trois objectifs principaux :

(…)

- **Mettre en place la loi binomiale :** On s'appuie sur l'expérimentation et la simulation pour étudier le schéma de Bernoulli. On introduit la notion de variable aléatoire et on installe la loi binomiale dont les utilisations sont nombreuses dans les domaines technologiques.
- **Expérimenter la notion de différence significative par rapport à une proportion attendue :** L'acquisition de la loi binomiale permet de poursuivre la formation des élèves dans le domaine de l'échantillonnage et des procédures de prise de décision en contexte aléatoire. On fait remarquer que, pour une taille de l'échantillon importante, on conforte les résultats vus en classe de seconde.

CONTENUS	CAPACITES ATTENDUES	COMMENTAIRES
Echantillonnage Utilisation de la loi binomiale pour une prise de décision à partir d'une fréquence observée sur un échantillon.	- Déterminer à l'aide de la loi binomiale un intervalle de fluctuation, à environ 95 %, d'une fréquence. - Exploiter un tel intervalle pour rejeter ou non une hypothèse sur une proportion.	L'intervalle de fluctuation peut être déterminé à l'aide d'un algorithme ou d'un tableur. On peut traiter quelques situations liées au contrôle en cours de fabrication ou à la réception d'une production. Le vocabulaire des tests (test d'hypothèse, hypothèse nulle, risque de première espèce) est hors programme.

I. ECHANTILLON ET FREQUENCE

a. Définition

Par définition, un échantillon s'obtient par tirage avec remise. Un échantillon de taille n est donc la liste des n résultats obtenus par n répétitions indépendantes de la même expérience.

Exemples :

- 50 lancers d'une pièce de monnaie
- 200 lancers d'un même dé
- sondage auprès de 1000 personnes choisies au hasard…

b. Fluctuation des fréquences

Si on réalise plusieurs fois la même expérience sur des échantillons de même aille n, on constate que la fréquence varie (« fluctue ») en fonction de l'échantillon

On parle de « *fluctuation des fréquences* » ou «*fluctuation d 'échantillonnage* »

Exemple :

- Sur le premier échantillon de 50 premiers lancers de pièce on a 48% de PILE, sur le second échantillon on a 56%.

c. Loi des grands nombres

On a remarqué que quand n devient de plus en plus grand, la fréquence f de chaque résultat tend à se « stabiliser » vers une valeur théorique appelée **probabilité** de cette issue. C'est ce qu'on appelle la « *loi faible des grands nombres* ».

Exemple :

Si je lance « infiniment longtemps » ma pièce de monnaie, la fréquence de PILE devrait se stabiliser atours de 0,5 soit 50%.

II. ECHANTILLONNAGE

a. Fréquence suivant une loi binomiale

Dans le cadre d'une loi Binomiale, où on connaît la taille de l'échantillon n et la probabilité de « succès » n, on peut **sans réaliser l'expérience**, écrire la loi de distribution de X (le nombre de succès) ou F (la fréquence de succès).

x_i	0	1	2	3	4	5	…	95	96	97	98	99	100
f_i	0	0,01	0,02	0,03	0,04	0,05		0,95	0,96	0,97	0,98	0,99	1
P(X = x_i) ou P(F = f_i)	P_0	P_1	P_2	P_3	P_4	P_5		P_{95}	P_{96}	P_{97}	P_{98}	P_{99}	P_{100}

On peut facilement déterminer la probabilité que X (ou F) se trouve dans un certain intervalle.

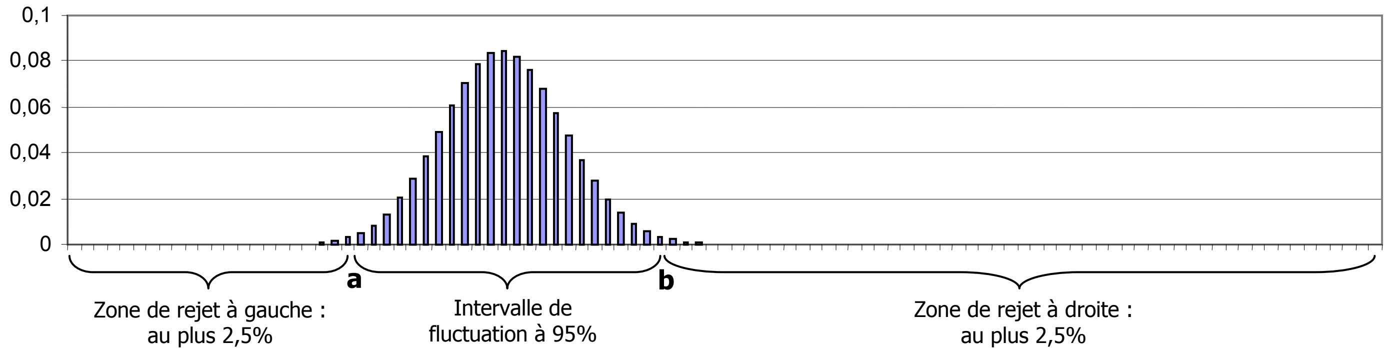
$$\mathbf{P}\,(\mathbf{X} \in [a\,;\,b]) = \mathbf{P}(a \leq \mathbf{X} \leq b) = \mathbf{P}_a + \mathbf{P}_{a+1} + \ldots + \mathbf{P}_{b-1} + \mathbf{P}_b$$

Un tel intervalle est appelé **intervalle de fluctuation** à un certain seuil.
Il représente l'intervalle dans lequel la fréquence doit se trouver, avec une certaine probabilité.

Exemple :
P(0,23 ≤ F ≤ 0,31) = 0,95 signifie que « F appartient à l'intervalle de fluctuation [0,23 ; 0,31] » avec un « **seuil de confiance** de 95% » (ou alors un « **seuil de risque** de 5% »).

b. Détermination d'un intervalle de confiance à 95%

Voici la représentation de la distribution binomiale d'une variable aléatoire X suivant une loi binomiale.



Le nombre **a** est le plus grand entier tel que P(X ≤ **a**) ≤ 0,025
Le nombre **b** est le plus petit entier tel que P(X ≤ **b**) ≥ 0,975

On peut alors écrire que : $\mathbf{P}(a \leq \mathbf{X} \leq b) \geq \mathbf{95\%}$ ou alors $\mathbf{P}\left(\frac{a}{n} \leq \mathbf{F} \leq \frac{b}{n}\right) \geq \mathbf{95\%}$

L'intervalle de fluctuation de F est $\left[\frac{a}{n}\,;\,\frac{b}{n}\right]$

c. Prise de décision

Cet intervalle de fluctuation sert à prendre des décisions avec un seuil de confiance de 95% (car on n'est jamais sûr de rien !), dans le cas d'une fréquence suivant une loi binomiale de paramètres n et p.
- on calcule la fréquence f dans un échantillon de taille n.
- on sait que cette fréquence appartient à l'intervalle **I** avec une probabilité de 95%.

On fait **l'hypothèse** suivante : *l'échantillon provient bien de la population* (« est représentatif »)
Si f appartient bien à I, alors on accepte l'hypothèse. Sinon on la rejette.

Exemple :
Dans la population, il parait qu'il y a 16% de gauchers. On va tester cette hypothèse sur un échantillon (qu'on suppose représentatif de la population) de 50 personnes.

1. X suit la loi binomiale de paramètres n = 50 et n = 0,16
2. On détermine l'intervalle de confiance :

x_i	0	1	2	3	4	…	10	11	12	13	14	…	49	50
f_i	0	0,02	0,04	0,06	0,08	…	0,20	0,22	0,24	0,26	1	…	1	1
P(F = f_i)	0	0,002	0,009	0,031	0,081	…	0,834	0,907	9,952	0,978	0,990	…	1	1
	Rejet à gauche			**Intervalle de fluctuation à 95%**							Rejet à droite			

L'intervalle de confiance est donc [0,06 ; 0,26] ce qui signifie que dans notre échantillon il y a entre 6% et 26% de gauchers, au seuil de confiance de 95%.
3. On calcule le pourcentage de gauchers dans notre échantillon. S'il appartient à l'intervalle on valide l'hypothèse de départ. Sinon, il est probable que celle-ci soit fausse, et qu'en réalité il n'y ait pas 16% de gauchers dans la popuation.