Un biberon comme outil de simulation au lycée

Nelly LASSALLE et Annette CORPART (IREM de Clermont-Ferrand)

II – Estimer une proportion ***(programme de terminale)***

Une entreprise d'agronomie produit en grande quantité de la semence pour gazon, qui est un mélange de deux types de graines :

- de l'Agropyrum Repens (chiendent, noté AR), dans une proportion *p*. Cette graine est adaptée aux conditions sèches mais n'offre pas une couverture très esthétique.
- du Lolium Perenne, dans une proportion 1 – *p*.

Le nombre de graines d'AR obtenu dans un échantillon aléatoire de taille *n* est variable. C'est une variable aléatoire notée X_n.
L'objectif est d'estimer *p* à partir de la fréquence $F_n = \frac{X_n}{n}$ d'apparition des graines d'AR dans **un** échantillon.

1° Expérience et recueil des données.

On va prélever un échantillon de *n* graines, successivement et avec remise, dans la production de l'entreprise contenant **une proportion *p* fixée mais inconnue de graines d'AR**. Cette production sera représentée par une bouteille opaque munie d'une tétine et contenant des boules rouges et des boules d'une autre couleur. A chaque tirage, on notera 1 à chaque fois que la couleur de la « graine » obtenue est rouge (c'est une graine d'AR) et 0 sinon.
Remarquons que dans la réalité, on prélèverait *n* graines au hasard mais sans remise ; l'échantillon serait pourtant considéré comme aléatoire car obtenu parmi un nombre considérable de graines.

➢ Réaliser alors 50 tirages de graines pour générer un échantillon de taille 50 et noter les résultats (0 ou 1) dans la fiche individuelle page 10.
Calculer le cumul des 1 puis la fréquence *f* des graines d'AR pour votre échantillon.

En comparant avec les résultats obtenus par les autres élèves, est-il raisonnable de prendre la valeur f de votre observation comme estimation de la proportion p ? Pourquoi ?

2° Estimer *p* par une fourchette.

On sait depuis la classe de seconde que :
pour une proportion *p* du caractère comprise entre 0,2 et 0,8 et pour des échantillons de taille *n* ≥ 25,
***f* appartient à l'intervalle de fluctuation** $\left[p-\frac{1}{\sqrt{n}}\,;p+\frac{1}{\sqrt{n}}\right]$ **avec une probabilité d'au moins 0,95.**

➢ *Construire un algorithme permettant d'afficher les bornes des intervalles de fluctuation pour des valeurs de p variant de 0,2 à 0,8 avec un pas de 0,05 puis représenter ces intervalles sur papier millimétré (page 11) : à chaque valeur de p en abscisse, on dessinera verticalement l'intervalle associé.*

➢ On peut remarquer que les bornes de tous les intervalles de fluctuation que l'on pourrait fabriquer se situent sur les droites d'équations $y = x - \frac{1}{\sqrt{50}}$ et $y = x + \frac{1}{\sqrt{50}}$.

Tracer ces droites.

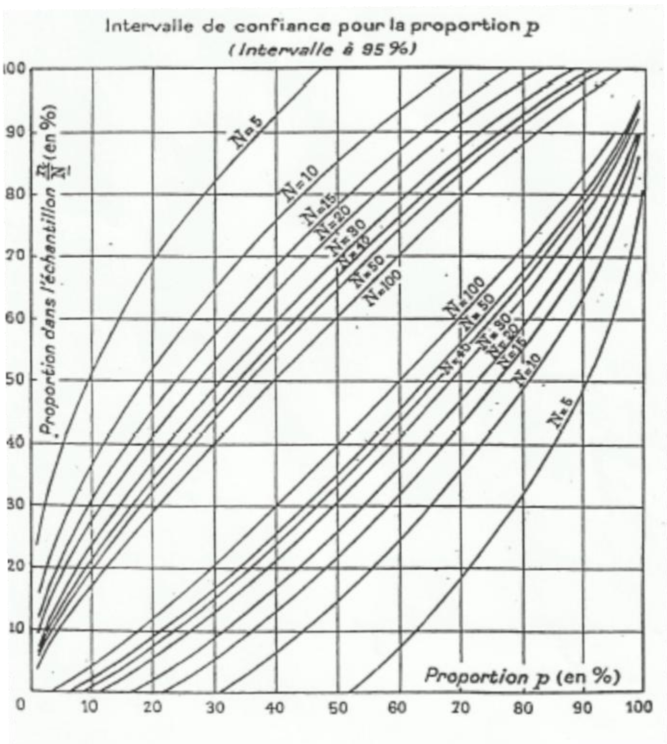
➢ Réciproquement, en plaçant en ordonnée la valeur de *f* obtenue au 1° avec un échantillon de 50 graines, on peut déterminer graphiquement les intervalles de fluctuation contenant *f* avec une probabilité d'au moins 0,95. Pour cela, chercher les abscisses *a* et *b* des points d'intersection des deux droites précédentes avec la droite d'équation *y* = *f*.

[*a* , *b*] est appelé un intervalle de confiance de *p* au niveau 95%. Ceci signifie que *p* est en dehors de cet intervalle pour seulement 5% des échantillons.

Donner votre estimation de la proportion de graines d'AR dans la production de semence par un intervalle de confiance au niveau 95% :

3° <u>Remarques.</u>

- Selon l'échantillon de 50 graines tiré par chaque élève, la valeur de *f* obtenue peut être différente. On ne peut donc pas parler de **<u>l</u>**'intervalle de confiance de *p* mais seulement d'**<u>un</u>** intervalle de confiance de *p*.
- En Terminale avec le travail sur la loi normale, on montre que l'intervalle de fluctuation de *p* au seuil 95% est de la forme $\left[p - 1{,}96\frac{\sqrt{p(1-p)}}{\sqrt{n}}, p + 1{,}96\frac{\sqrt{p(1-p)}}{\sqrt{n}}\right]$ dès que $n \geq 30, np \geq 5$ et $n(1-p) \geq 5$. On admettra que si l'on dessine les intervalles de fluctuation pour des valeurs de *p* variant de 0 à 1, leurs bornes se situent sur des arcs d'ellipses.



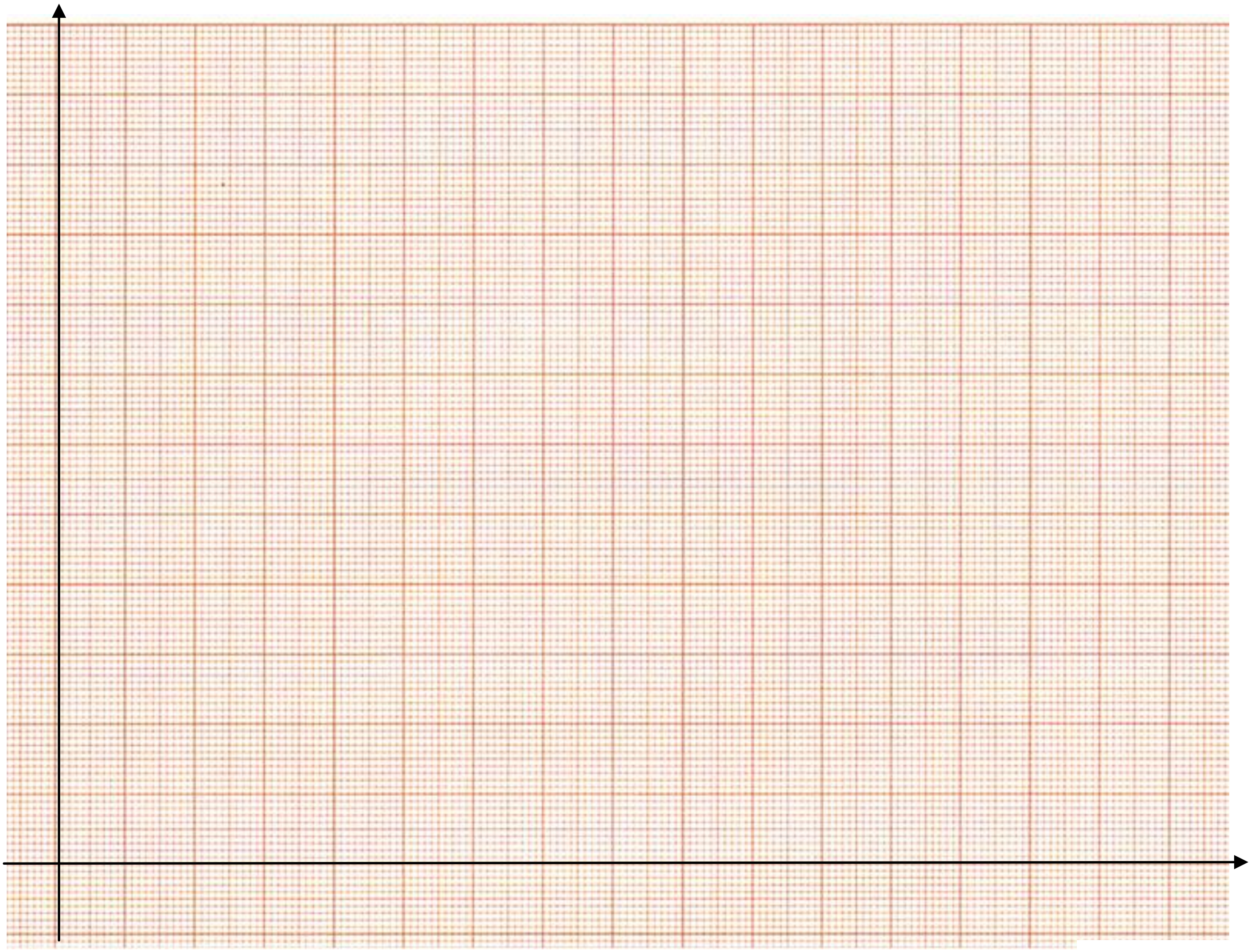
- Le professeur pourra dévoiler la composition des biberons et donner la valeur de la proportion de graines d'AR (par exemple *p* = 0,25) même si dans l'entreprise cette opération est impossible à réaliser.
  Les élèves peuvent ainsi s'apercevoir que leur propre intervalle de confiance ne contient pas forcément la valeur de *p*. Ils calculent alors le pourcentage de « bons » intervalles dans la classe et s'aperçoivent que ce résultat est proche de 95%.

Fiche individuelle (II – 1°)

NOM Prénom :			Effectifs d'AR
	Echantillon E	tirage 1	
	Echantillon E	tirage 2	
	Echantillon E	tirage 3	
	Echantillon E	tirage 4	
	Echantillon E	tirage 5	
	Echantillon E	tirage 6	
	Echantillon E	tirage 7	
	Echantillon E	tirage 8	
	Echantillon E	tirage 9	
	Echantillon E	tirage 10	
	Echantillon E	tirage 11	
	Echantillon E	tirage 12	
	Echantillon E	tirage 13	
	Echantillon E	tirage 14	
	Echantillon E	tirage 15	
	Echantillon E	tirage 16	
	Echantillon E	tirage 17	
	Echantillon E	tirage 18	
	Echantillon E	tirage 19	
	Echantillon E	tirage 20	
	Echantillon E	tirage 21	
	Echantillon E	tirage 22	
	Echantillon E	tirage 23	
	Echantillon E	tirage 24	
	Echantillon E	tirage 25	
	Echantillon E	tirage 26	
	Echantillon E	tirage 27	
	Echantillon E	tirage 28	
	Echantillon E	tirage 29	
	Echantillon E	tirage 30	
	Echantillon E	tirage 31	
	Echantillon E	tirage 32	
	Echantillon E	tirage 33	
	Echantillon E	tirage 34	
	Echantillon E	tirage 35	
	Echantillon E	tirage 36	
	Echantillon E	tirage 37	
	Echantillon E	tirage 38	
	Echantillon E	tirage 39	
	Echantillon E	tirage 40	
	Echantillon E	tirage 41	
	Echantillon E	tirage 42	
	Echantillon E	tirage 43	
	Echantillon E	tirage 44	
	Echantillon E	tirage 45	
	Echantillon E	tirage 46	
	Echantillon E	tirage 47	
	Echantillon E	tirage 48	
	Echantillon E	tirage 49	
	Echantillon E	tirage 50	
		Somme	

Construction des intervalles :

Valeurs de f



Valeurs de p