# *Un biberon comme outil de simulation au lycée*

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

## ***III – Prendre une décision dans un contrôle de qualité*** **(programme de première)**

Dans le service technique d'une commune, vous êtes chargé de réceptionner la commande d'un lot de semence (par exemple une tonne) pour engazonner les espaces verts. Le cahier des charges de la commande précise que **le mélange doit contenir 30% de graines d'Agropyrum Repens (chiendent, noté AR)**. Cette plante n'offre pas une couverture très esthétique mais elle est bien adaptée à la sécheresse.
La procédure de contrôle statistique, appelé plan de contrôle, fixe :

- ✓ la procédure d'échantillonnage en indiquant **la taille *n* de l'échantillon aléatoire** de graines à prélever dans le lot et les modalités de prélèvement. On compte le nombre de graines d'AR dans cet échantillon.
- ✓ **le critère de rejet du lot** : c'est le nombre de graines d'AR à partir ou en dessous duquel le lot doit être rejeté comme non conforme à la commande.

**Deux erreurs** peuvent survenir dans votre prise de décision : vous pouvez vous tromper en rejetant un lot conforme (risque fournisseur) ou en acceptant un lot non conforme (risque client). Ces risques proviennent de la variabilité des résultats issus d'une procédure d'échantillonnage : en échantillonnant deux fois de la même façon, on peut obtenir des résultats différents (voir programme de seconde).

Dans un premier temps, on réalise l'expérience pour recueillir les données. Dans un deuxième temps, on travaille sur un modèle probabiliste pour comprendre comment obtenir les critères de rejet pour un risque d'erreur fixé.

### 1° Protocole expérimental et recueil des données.

➢ Comme un contrôle réel est impossible en classe, un lot de semence sera représenté par une bouteille opaque munie d'une tétine et contenant des boules rouges et des boules d'une autre couleur, en proportion inconnue. Chaque élève expérimente avec une bouteille numérotée, assimilée au lot de semence.
A chaque tirage, on notera 1 si la couleur de la « graine » obtenue est rouge (c'est une graine d'AR) et 0 sinon.
Remarquons que pour un contrôle réel, 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 20 tirages de graines pour générer un échantillon A de taille 20 et noter les résultats (0 ou 1) dans votre fiche individuelle (page 16). De la même manière, noter les résultats d'un échantillon B obtenu avec 50 tirages. Calculer ensuite le nombre $X$ de graines d'AR obtenues dans chacun des deux échantillons.

➢ Le tableau récapitulatif de la classe (page 17) se remplit manuellement et est photocopié ensuite par l'enseignant.

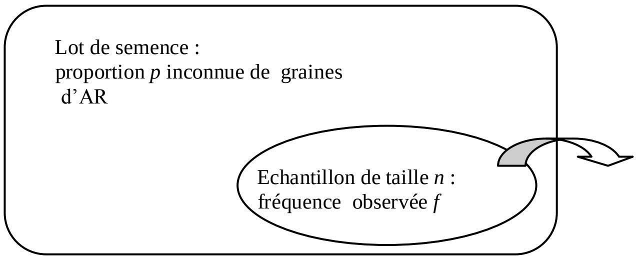
### 2° Retour sur le contrôle de qualité.

Vous devez prendre une décision concernant l'acceptation ou le refus du lot de semence livré.

➢ *Qu'est ce qu'un lot conforme au cahier des charges ?*

*Quels lots vous semblent non conformes à la commande au vu des échantillons A ? Des échantillons B ?*

3° **<u>Construction du critère de rejet à l'aide des probabilités.</u>**




Prenons comme **hypothèse** que **la proportion** ***p*** **est de 30%** car c'est celle qui est prévue par le cahier des charges. On rejettera cette hypothèse lorsqu'on observera une différence significative entre la valeur supposée de $p$ et la valeur observée $f$ (donc si $f$ est trop éloignée de $p$).
**On choisit un seuil de décision $\alpha$** : on souhaite que la probabilité de rejeter l'hypothèse alors qu'elle est vraie, soit au maximum de 5% si on choisit $\alpha$=0,05.

- On rappelle que $X$ est la variable aléatoire mesurant le nombre de graines d'AR observées dans un échantillon de taille $n$.

*Quelle loi suit X ? Quels sont ses paramètres ?*

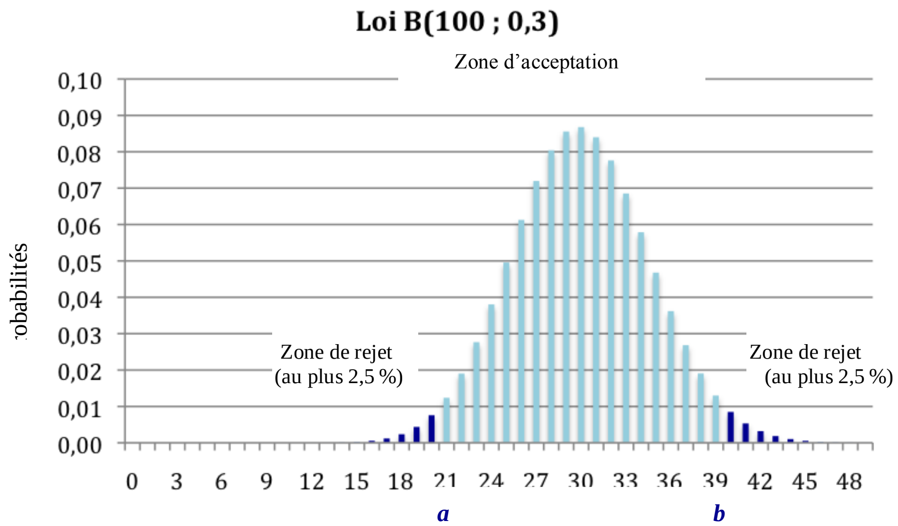
On cherche à partager l'intervalle [0, $n$], où $X$ prend ses valeurs, en trois intervalles [0 ; $a$-1], [$a$ ; $b$] et [$b$+1 ; $n$] de sorte que $X$ prenne ses valeurs dans chacun des intervalles extrêmes avec une probabilité cumulée d'au plus 0,025, ce qui est réalisé si $P(X \leq a-1) \leq 0{,}025$ et $P(X \geq b+1) \leq 0{,}025$.

Comme la loi de $X$ est discrète, on peut remarquer que
$P(X \leq a-1) \leq 0{,}025 \Leftrightarrow P(X < a) \leq 0{,}025$ et $P(X \geq b+1) \leq 0{,}025 \Leftrightarrow P(X > b) \leq 0{,}025$.

Il suffit donc de déterminer :

- le plus petit entier $a$ tel que $P(X \leq a) > 0{,}025 \Leftrightarrow \sum_{k=0}^{a} \binom{n}{k} \times 0{,}3^k \times 0{,}7^{n-k} > 0{,}025$
- le plus petit entier $b$ tel que $P(X \leq b) \geq 0{,}975 \Leftrightarrow \sum_{k=0}^{b} \binom{n}{k} \times 0{,}3^k \times 0{,}7^{n-k} \geq 0{,}975$

Exemple :




La règle de décision est la suivante : **si la fréquence observée *f* de graines d'AR appartient à l'intervalle** $\left[\frac{a}{n};\frac{b}{n}\right]$, **on considère que l'hypothèse selon laquelle le lot de semence comporte 30 % d'AR n'est pas remise en question et on l'accepte ; sinon, on rejette cette hypothèse en sachant qu'on a 5 % de risque de faire une erreur**.

Les entiers *a* et *b* se déterminent, soit à l'aide d'un algorithme, soit avec un tableur.

- Etudions l'algorithme suivant :

<u>Variables</u>
*n*, *p*, α, *k*, S
<u>Entrée</u>
Saisir α
Saisir *n*
Saisir *p*
*k* prend la valeur 0
S prend la valeur 0
<u>Traitement</u>
Tant que S $\leq \frac{\alpha}{2}$ faire
    S prend la valeur S + $\binom{n}{k} \times p^k \times (1-p)^{n-k}$
    *k* prend la valeur *k* + 1
Fin Tant que
<u>Sortie</u>
Afficher *k* – 1

*Identifier les différentes variables de cet algorithme.*

*Appliquer le pour α = 0,05, p = 0,30 et n = 20 (ou n = 50). Quelle valeur de a obtient-on ?*

*Construire à droite de l'algorithme précédent un autre algorithme permettant d'obtenir la valeur de b.*
*Programmer les deux algorithmes sur votre calculatrice.*

➢ Avec un tableur ou à la calculatrice, on peut également obtenir les valeurs de $a$ et $b$ en visualisant la distribution de la loi binomiale $\mathcal{B}(n\ ;\ 0{,}30)$ pour différentes valeurs de $n$.
On procède de la manière suivante, décrite pour le cas $n = 50$ :

- Dans la colonne A, on écrit les valeurs de k de 0 à 50. (par exemple 0 en A3, 1 en A4 …..)
- En entrant la formule = LOI.BINOMIALE (A3; 50; 0,30; FAUX) dans la cellule B3, on calcule la probabilité $P(X = \text{A3})$, c'est-à-dire $P(X = 0)$ dans notre exemple.
- En entrant la formule = LOI.BINOMIALE (A3; 50; 0,30; VRAI) dans la cellule C3, on calcule les probabilités cumulées $P(X \leq \text{A3})$, c'est-à-dire $P(X \leq 0)$ dans notre exemple.
- On recopie ces deux formules en colonne pour les valeurs de $k$ de 0 à 50.

Avec « Openofficecalc », il faut écrire 0 à la place de FAUX et 1 à la place de VRAI.

La valeur de $a$ est le plus entier $k$ tel que $P(X \leq k) > 0{,}025$ ; celle de $b$ le plus petit entier $k$ tel que $P(X \leq k) \geq 0{,}975$.

*En utilisant une des deux méthodes, compléter le tableau suivant :*

| Taille d'échantillon | Valeur de $a$ | Valeur de $b$ | Intervalle $\left[\frac{a}{n};\frac{b}{n}\right]$ |
|---|---|---|---|
| $n = 20$ | | | |
| $n = 50$ | | | |

*Quels sont les lots qui sont refusés ?*

➢ Remarque :

Pour $n$ assez grand et $p$ ni trop petit ni trop grand, on observe que l'intervalle $\left[\frac{a}{n};\frac{b}{n}\right]$ est sensiblement le même que l'intervalle de fluctuation $\left[p-\frac{1}{\sqrt{n}};p+\frac{1}{\sqrt{n}}\right]$ étudié en seconde.

*Vérifier le.*

L'intérêt de l'intervalle $\left[\frac{a}{n};\frac{b}{n}\right]$, calculé à partir de la loi binomiale, est de convenir pour **toutes les valeurs des paramètres *n* et *p*** (en particulier pour de petits échantillons) alors que l'intervalle $\left[p-\frac{1}{\sqrt{n}};p+\frac{1}{\sqrt{n}}\right]$ n'est adapté que sous certaines conditions de $n$ et de $p$.

**4° <u>Analyse des erreurs.</u>**

On dévoile la composition des bouteilles numérotées qui contiennent chacune vingt boules. Certaines ont une proportion de boules rouges égale à 30 %, d'autres une proportion de boules rouges inférieure ou supérieure à 30 %.

*Relever cette proportion pour chacune des bouteilles sur le tableau page 17.*

*Noter les cas où l'on s'est trompé. On fera bien la différence entre les deux types d'erreur :*

- *on a rejeté un lot conforme (surligner en vert) : c'est dommage pour le fournisseur.*
- *on a accepté un lot non conforme (surligner en rouge) : c'est dommage pour notre gazon !*

# Fiche individuelle ( III – 1°)

**NOM Prénom :**

**Biberon n°**

| | | Effectifs d'AR |
|---|---|---|
| Echantillon A | tirage 1 | |
| Echantillon A | tirage 2 | |
| Echantillon A | tirage 3 | |
| Echantillon A | tirage 4 | |
| Echantillon A | tirage 5 | |
| Echantillon A | tirage 6 | |
| Echantillon A | tirage 7 | |
| Echantillon A | tirage 8 | |
| Echantillon A | tirage 9 | |
| Echantillon A | tirage 10 | |
| Echantillon A | tirage 11 | |
| Echantillon A | tirage 12 | |
| Echantillon A | tirage 13 | |
| Echantillon A | tirage 14 | |
| Echantillon A | tirage 15 | |
| Echantillon A | tirage 16 | |
| Echantillon A | tirage 17 | |
| Echantillon A | tirage 18 | |
| Echantillon A | tirage 19 | |
| Echantillon A | tirage 20 | |
| | **Somme** | |

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

**Fiche récapitulative des effectifs et des fréquences d'AR ( III – 1°)**

| | | Echantillon A ($n$ = 20) | | Echantillon B ($n$ = 50) | |
|---|---|---|---|---|---|
| **Biberon n°** | **Nom Prénom** | effectif | fréquence | effectif | fréquence |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |

# Tables des lois binomiales

$\mathcal{B}(20\ ;\ 0{,}3)$

| k | P(X = k) | P(X ≤ k) |
|---|---|---|
| 0 | 0,0008 | 0,0008 |
| 1 | 0,0068 | 0,0076 |
| **2** | 0,0278 | 0,0355 |
| 3 | 0,0716 | 0,1071 |
| 4 | 0,1304 | 0,2375 |
| 5 | 0,1789 | 0,4164 |
| 6 | 0,1916 | 0,6080 |
| 7 | 0,1643 | 0,7723 |
| 8 | 0,1144 | 0,8867 |
| 9 | 0,0654 | 0,9520 |
| 10 | 0,0308 | 0,9829 |
| 11 | 0,0120 | 0,9949 |
| 12 | 0,0039 | 0,9987 |
| 13 | 0,0010 | 0,9997 |
| 14 | 0,0002 | 1,0000 |
| 15 | 0,0000 | 1,0000 |
| 16 | 0,0000 | 1,0000 |
| 17 | 0,0000 | 1,0000 |
| 18 | 0,0000 | 1,0000 |
| 19 | 0,0000 | 1,0000 |
| 20 | 0,0000 | 1,0000 |

$\mathcal{B}(50\ ;\ 0{,}3)$

| k | P(X = k) | P(X ≤ k) |
|---|---|---|
| 0 | 0,000 | 0,000 |
| 1 | 0,000 | 0,000 |
| 2 | 0,000 | 0,000 |
| 3 | 0,000 | 0,000 |
| 4 | 0,000 | 0,000 |
| 5 | 0,001 | 0,001 |
| 6 | 0,002 | 0,002 |
| 7 | 0,005 | 0,007 |
| 8 | 0,011 | 0,018 |
| 9 | 0,022 | 0,040 |
| 10 | 0,039 | 0,079 |
| 11 | 0,060 | 0,139 |
| 12 | 0,084 | 0,223 |
| 13 | 0,105 | 0,328 |
| 14 | 0,119 | 0,447 |
| 15 | 0,122 | 0,569 |
| 16 | 0,115 | 0,684 |
| 17 | 0,098 | 0,782 |
| 18 | 0,077 | 0,859 |
| 19 | 0,056 | 0,915 |
| 20 | 0,037 | 0,952 |
| 21 | 0,023 | 0,975 |
| 22 | 0,013 | 0,988 |
| 23 | 0,007 | 0,994 |
| 24 | 0,003 | 0,998 |
| 25 | 0,001 | 0,999 |
| 26 | 0,001 | 1,000 |
| 27 | 0,000 | 1,000 |
| 28 | 0,000 | 1,000 |
| 29 | 0,000 | 1,000 |
| 30 | 0,000 | 1,000 |
| 31 | 0,000 | 1,000 |
| 32 | 0,000 | 1,000 |
| 33 | 0,000 | 1,000 |
| 34 | 0,000 | 1,000 |
| 35 | 0,000 | 1,000 |
| 36 | 0,000 | 1,000 |
| 37 | 0,000 | 1,000 |
| 38 | 0,000 | 1,000 |
| 39 | 0,000 | 1,000 |
| 40 | 0,000 | 1,000 |
| 41 | 0,000 | 1,000 |
| 42 | 0,000 | 1,000 |
| 43 | 0,000 | 1,000 |
| 44 | 0,000 | 1,000 |
| 45 | 0,000 | 1,000 |
| 46 | 0,000 | 1,000 |
| 47 | 0,000 | 1,000 |
| 48 | 0,000 | 1,000 |
| 49 | 0,000 | 1,000 |
| 50 | 0,000 | 1,000 |

## Représentation des lois binomiales