

Chapitre 9 : Variable aléatoire.

Eliot ROY

2022-2023

Points du programme.

Thème : Probabilité.

Contenus :

- Variable aléatoire discrète : loi de probabilité, espérance.
- Loi de Bernoulli $(0,1)$ de paramètre p , espérance.

Capacités attendues :

- Interpréter en situation les écritures $\{X = a\}$, $\{X \leq a\}$, où X désigne une variable aléatoire et calculer les probabilités correspondantes $P(X = a)$, $P(X \leq a)$.
- Calculer et interpréter en contexte l'espérance d'une variable aléatoire discrète.
- Reconnaître une situation aléatoire modélisée par une loi de Bernoulli.
- Simuler N échantillons de taille n d'une loi de Bernoulli et représenter les fréquences observées des 1 par un histogramme ou un nuage de points.
- Interpréter sur des exemples la distance à p de la fréquence observée des 1 dans un échantillon de taille n d'une loi de Bernoulli de paramètre p .

Commentaires :

- On s'abstient de tout formalisme sur les variables aléatoires. Elles sont essentiellement manipulées en contexte pour modéliser des situations dans lesquelles les issues sont des nombres aléatoires.
- La simulation d'échantillons de taille n d'une loi de Bernoulli de paramètre p permet d'observer la fluctuation d'échantillonnage.
- Sur des simulations de N échantillons (N de l'ordre de plusieurs centaines), on évalue le pourcentage d'échantillons dont la fréquence observée des 1 se situe à une distance de s , $2s$ ou $3s$ de p , où s désigne l'écart-type de la série des fréquences observées. Sans développer de théorie de décision ou de test, et en prenant appui sur des simulations et des représentations (histogramme, nuage de points), on fait percevoir, pour une observation donnée, la diversité des interprétations possibles de la distance à p (paramètre du modèle) de la fréquence des 1 : situation fréquente ou situation rare dans le cadre du modèle.
- Sur des simulations, on constate que la série des fréquences observées des 1 dans N échantillons de taille n d'une loi de Bernoulli a un écart-type de l'ordre de $\frac{1}{\sqrt{n}}$. Pour plusieurs valeurs de n ,

on représente $\frac{1}{\sqrt{n}}$ en abscisse et, en ordonnée, l'écart-type s des fréquences observées des 1 dans N échantillons (plusieurs centaines) de taille n . On peut commenter ce résultat en observant que pour diviser la dispersion par k , il faut multiplier la taille de l'échantillon par k^2 .

Commentaires :

- Simuler des échantillons de taille n d'une loi de Bernoulli à partir d'un générateur de nombres aléatoires entre 0 et 1.
- Représenter par un histogramme ou par un nuage de points les fréquences observées des 1 dans N échantillons de taille n d'une loi de Bernoulli.
- Compter le nombre de valeurs situées dans un intervalle de la forme $[p - ks; p + ks]$ pour $k \in \{1; 2; 3\}$.

I Variables aléatoires discrètes.

Définition : On appelle **variable aléatoire réelle** X une fonction qui, à chaque issue d'une expérience aléatoire, associe un nombre réel.

Exemple : On considère un sac contenant une boule noire, notée N , deux boules rouges, notées R_1 et R_2 , et deux boules blanches, notées B_1 et B_2 . L'univers de l'expérience aléatoire est l'ensemble $E = \{N; R_1; R_2; B_1; B_2\}$.

Si on tire :

- la boule noire, on perd 10 euros ;
- l'une des boules rouges, on gagne 1 euros ;
- l'une des boules blanches, on gagne 4 euros.

Notons X la variable aléatoire égale au gain du joueur. L'ensemble des valeurs prises par cette variable est $E' = \{-10; 1; 4\}$.

Définition : Soit X une variable aléatoire définie sur l'univers E et $E' = \{x_1; x_2; \dots x_n\}$ l'ensemble des valeurs prises par X . L'évènement $\{X = x_i\}$ est l'ensemble des issues de E auxquelles on associe le réel x_i . L'évènement $\{X < x_i\}$ est l'ensemble des issues de E auxquelles on associe un réel strictement inférieur à x_i .

Exemple : Reprenons l'exemple précédent. L'évènement $\{X = 4\}$ correspond à l'évènement : « Le joueur gagne 4 euros ». Il correspond donc à l'ensemble $\{B_1; B_2\}$.

II Loi de probabilité.

II.1 Définition.

Définition : Définir une loi de probabilité d'une variable aléatoire X , c'est associer une probabilité à chaque valeur que peut prendre X .

Traduction : ☺ La loi de probabilité d'une variable aléatoire est souvent présentée sous la forme d'un tableau comme ci-dessous :

x_i	x_1	x_2	$\dots$	x_n
$P(X = x_i)$	p_1	p_2	$\dots$	p_n

Pour la suite de la leçon, on conservera les notations ci-dessus.

Exemple : Dans l'exemple précédent, la loi de probabilité de X est :

x_i	-10	1	4
$P(X = x_i)$	0,2	0,4	0,4

II.2 Espérance d'une variable aléatoire.

Définition : On appelle espérance de la variable aléatoire X la quantité :

$$E(X) = x_1p_1 + x_2p_2 + \dots + x_np_n.$$

Traduction : ☺ L'espérance $E(X)$ correspond à la moyenne des valeurs prises par X quand l'expérience aléatoire est répétée un très grand nombre de fois. Elle a la même unité que les valeurs x_i .

Exemple : Dans l'exemple précédent, l'espérance de la variable aléatoire X vaut :

$$E(X) = 0,2 \times (-10) + 0,4 \times 1 + 0,4 \times 4 = 0$$

Cela signifie que si le joueur joue un très grand nombre de fois, il peut espérer gagner zéro euros. Dans ce cas, on dira que le jeu est équitable.

II.3 Loi de Bernoulli.

On rappelle qu'une épreuve de Bernoulli de paramètre p est une expérience présentant deux issues contraires :

- l'une est appelée « succès », notée S , de probabilité p ;
- l'autre est appelée « échec », noté $\bar{S}$, de probabilité $1 - p$.

Exemple : On lance une pièce truquée de sorte que la probabilité d'obtenir le côté « pile » soit égale à 0,6. On note P l'issue « obtenir pile », et $\bar{P}$ l'issue « obtenir face ». Cette expérience est une épreuve de Bernoulli de paramètre $p = 0,6$.

Définition : On appelle **loi de Bernoulli** la loi de probabilité qui prend :

- la valeur 1 en cas de succès, avec une probabilité p ;
- la valeur 0 en cas d'échec, avec une probabilité $1 - p$.

x_i	0	1
$P(X = x_i)$	$1 - p$	p

On dit que p est le paramètre de cette loi de Bernoulli.

Exemple : Considérons l'épreuve de Bernoulli qui consiste à lancer un dé à quatre faces, et donc le succès est S : « obtenir la face 3 ». Comme la probabilité d'obtenir la face 3 est égale à 0,25, la variable aléatoire X associée à cette épreuve de Bernoulli est telle que $P(X = 1) = 0,25$ et $P(X = 0) = 1 - 0,25 = 0,75$.

x_i	0	1
$P(X = x_i)$	0,75	0,25

Remarque : On dit que X suit une loi de Bernoulli lorsque sa loi de probabilité est une loi de Bernoulli.

Propriété : Soit X une variable aléatoire qui suit une loi de Bernoulli. Alors $E(X) = p$.

Traduction : ☀ L'espérance mathématique d'une variable aléatoire qui suit une loi de Bernoulli est égale à la probabilité de succès.

III Simulation.

III.1 Définition.

Définition : Lorsqu'on répète n fois une même épreuve de Bernoulli de façon identique et indépendante, on obtient une série de n résultats que l'on appelle **échantillon de taille n** .

Exemple : On lance une pièce 20 fois et on observe le nombre de faces Pile obtenues. La face Pile est considérée comme le succès, associé à la valeur 1. On peut simuler un tel échantillon :

Méthode Python :

La fonction `piece()` simule un lancer de pièce : elle retourne 1 pour l'obtention de Pile et 0 pour l'obtention de Face.

La fonction `echantillon(n)` simule n lancers de pièce et retourne la liste des résultats.

Un résultat retourne par la fonction `echantillon(n)` est :

```
1 from random import *
2
3 def piece():
4     x=randint(0,1)
5     return x
6
7 def echantillon(n):
8     L=[]
9     for i in range(n):
10         L.append(piece())
11     return L
```

```
>>> echantillon(20)
[0, 0, 0, 0, 1, 1, 1, 1, 1, 0, 1, 0, 0, 0, 1, 0, 0, 1, 0, 0]
```

La fréquence de succès est égale à $\frac{8}{20}$ car le chiffre 1 apparaît 8 fois sur l'échantillon de taille 20.

Méthode tableur :

Avec le tableur, il est possible de simuler un échantillon de 0 et de 1, à l'aide de la fonction `ALEA.ENTRE.BORNES(0;1)`. Il suffit ensuite d'étendre la formule.

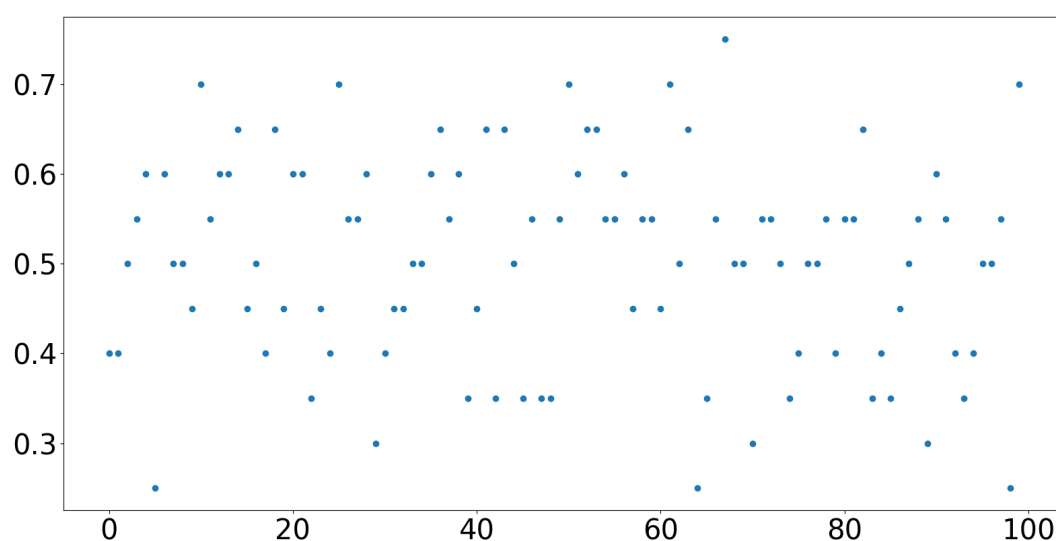
La fréquence de succès est égale à $\frac{10}{20}$ car le chiffre 1 apparaît 10 fois sur l'échantillon de taille 20.

A1	A	B	C
1	1		
2	1		
3	0		
4	0		
5	0		
6	0		
7	0		
8	0		
9	1		
10	0		
11	1		
12	1		
13	1		
14	0		
15	0		
16	1		
17	1		
18	1		
19	0		
20	1		

III.2 Fluctuation d'échantillonnage.

Définition : Si on réalise plusieurs échantillons de même taille, la fréquence du succès observée sur chaque expérience varie : c'est ce qu'on appelle la fluctuation d'échantillonnage.

Exemple : On réalise 100 échantillons de 20 lancers de pièce. La fréquence d'obtention de Pile dans chaque échantillon est représentée par le nuage de points ci-contre réalisé en langage Python. On voit que cette fréquence varie entre 0,1 et 0,9, alors que la proportion théorique de Pile est égale à 0,5.



Remarque : Plus la taille des échantillons est grande, plus le phénomène de fluctuation diminue : les fréquences se rapprochent de plus en plus de la proportion théorique.

III.3 Interprétation des résultats.

On considère une simulation de N échantillons de taille n . On note s l'écart-type de la série des fréquences obtenues et p la proportion théorique. En moyenne, on a :

- Environ 68% des fréquences dans l'intervalle $[p - s; p + s]$.
- Environ 95% des fréquences dans l'intervalle $[p - 2s; p + 2s]$.
- Plus de 99% des fréquences dans l'intervalle $[p - 3s; p + 3s]$.

Exemple :

Propriété : L'écart-type de la série des fréquences est de l'ordre de $\frac{1}{2\sqrt{n}}$. Autrement dit, la valeur de s est proche de $\frac{1}{2\sqrt{n}}$.