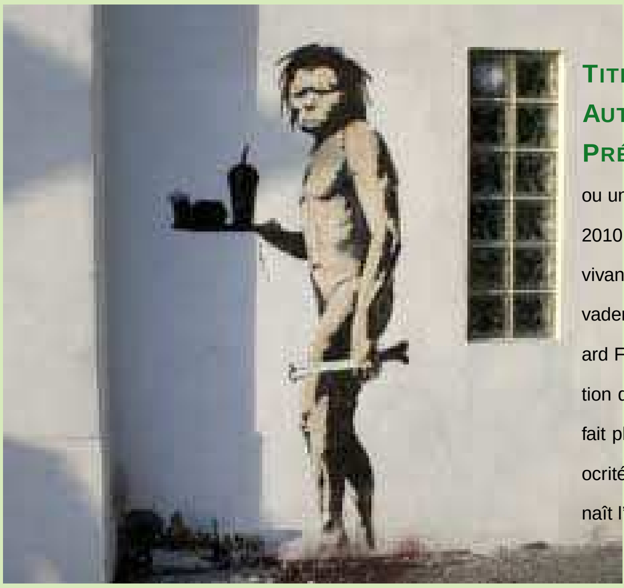
Chapitre 7

Dénombrement et loi de Probabilités

Hors Sujet



Titre : « Faites le mur »

Auteur : Banksy

Présentation succinte de l'auteur : Faites le mur ! est un film (?) ou un documentaire (documenteur !) réalisé par Banksy, sorti en salle le 15 décembre 2010. Thierry Guetta, un commerçant français excentrique, documentariste amateur vivant à Los Angeles, présenté dans le film comme le cousin de l'artiste Space Invader, aurait amassé une considérable archive d'interviews et d'action de Zevs, Shepard Fairey, André etc. A mesure qu'il filme de manière compulsive la nouvelle génération de l'art urbain, son obsession pour Banksy, le célèbre pochoiriste britannique se fait plus dévorante. Ils se rencontrent enfin. Banksy incite Guetta - au vu de la médiocrité de ses productions audiovisuelles - à se tourner vers l'art urbain. C'est alors que naît l'artiste Mr Brainwash.

Document réalisé à l'aide de LaTeX
Auteur : C. Aupérin
Site : wicky-math.fr.nf
Lycée : Jules Fil (Carcassonne)

Table des matières

« *Télécharger c'est tuer l'industrie, tuons les tous* »

THURSTON MOORE (SONIC YOUTH)

LEÇON 7

Dénombrement et loi de Probabilités

Résumé

Le but de ce chapitre est d'étudier certaines lois de probabilités particulières.
Tout d'abord, nous commencerons par une loi discrète. Mais pour l'appréhender, il nous faudra apprendre à compter ! Si, si ! Nous allons apprendre à déterminer le nombre d'issues possibles d'une expérience (ie à dénombrer) dans une première partie. Cela mènera à la découverte de nouvelles notations, méthodes, propriétés, formules, etc. Que de joie !
Puis la seconde partie sera consacrée à la loi qui nous intéresse.
Dans un troisième temps, nous étudierons des exemples de lois de probabilité dites « continues » (par opposition aux lois discrètes) portant sur des univers infinis non dénombrables (comme des intervalles de temps). Pour cela, tout ce que nous aurons fait sur la fonction exponentielle et les intégrales s'avèrera très utile (qui a dit que les mathématiques ne servaient pas ? ?).

I) Dénombrement

Travail de l'élève **1.**

1. Un sondage auprès de 150 personnes a donné les résultats suivants :
A la question « Consommez vous régulièrement de l'alcool ? », 50 personnes répondent oui.
A la question « Etes-vous fumeur ? », 80 personnes répondent oui.
A la question « Etes-vous un fumeur consommant régulièrement de l'alcool ? », 35 personnes répondent oui.

a. Représenter ces données par un diagramme, puis par un tableau à double entrée.

b. Combien de personnes sont des fumeurs ne consommant pas régulièrement de l'alcool ?

c. Combien de personnes consomment régulièrement de l'alcool et ne sont pas fumeurs ?

d. Combien de personnes ne sont pas fumeurs et ne consomment pas régulièrement de l'alcool ?

e. Combien de personnes sont fumeurs ou consomment régulièrement de l'alcool ?
Dans ce cas là, on travaille sur la réunion et l'intersection d'ensemble.

2. J'ai offert à mon neveu Ethan pour son anniversaire un jeu de cubes où sont inscrits les lettres de l'alphabet.
Très pédagogue, je lui donne d'abord les trois cubes A, B et C.

a. Combien de « mots » de trois lettres peut-il alors former ?

b. Et si je lui en donne 4 ? 26 ? 32 ?
Dans ce cas là, on cherche le nombre de possibilités d'ordonner, ou encore de ***permuter*** *les 3, 4, 26 ou 32 cubes.*

3. Au poker fermé, le jeu contient 32 cartes et une main est constituée de 5 cartes. Combien y a-t-il de mains différentes ?
Dans ce cas là, on cherche le nombre de possibilités de ***choisir 5 cartes parmi 32****, sans tenir compte de l'ordre.*

On peut dénombrer des cas possibles grâce à un diagramme ou un arbre (voire un tableau à double entrée).

I-1 Permutations

Définition 1 :
On appelle **permutation** sur un ensemble E toute liste *ordonnée* de tous les éléments de E. (*sans répétition)*

Exemple :
Si $E = \{1;2;3;4\}$, alors $(1,2,3,4)$, $(2,1,4,3)$ et $(1,2,4,3)$ sont des permutations de E, mais pas $(1,2,3,3)$

Proposition 1 :
Le nombre de permutations d'un ensemble à n éléments est le nombre noté $n!$ (se lit **factoriel** n) égal à

$$n! = n \times (n-1) \times (n-2) \times \cdots \times 1$$

Remarque : Par convention, $0! = 1$

Preuve **Idée**
On construit un arbre. On a n façons de choisir le premier élément donc la première étape comporte n branches. Puis il reste $(n-1)$ choix pour le deuxième élément, soit $n-1$ branches, etc jusqu'au dernier élément de la liste, où l'on n'a plus qu'un choix possible.

Exemple :
Il y a 26 élèves dans cette classe, et donc 26! positions différentes pour la photo de classe.

Exercice 1 **:**
Combien y a-t-il d'anagrammes du mot ZOÉ ? Du mot ANA ?

Remarque : Les quotients de factoriels se simplifient toujours. Par exemple :

$$\frac{9!}{12!} = \frac{9 \times 8 \times \cdots \times 1}{12 \times 11 \times 10 \times 9 \times \cdots \times 1} = \frac{1}{12 \times 11 \times 10} = \frac{1}{1320}$$

Exercice 2 :

1. Démontrer que $6! \times 7! = 10!$ (sans calculer $10!$)
2. Simplifier
$$\frac{(n+1)!}{n!}$$
3. Démontrer que tout entier k : $(k+1)! - k! = k \times k!$

I-2 Combinaisons

Dans cette partie, n et p désignent deux nombres entiers naturels tels que $p \leq n$.

I-2.1 Dénombrement et coefficients binômiaux

Définition 2 :
Une partie formée de p éléments d'un ensemble E contenant n éléments s'appelle une **combinaison** de p éléments de E ou encore une $p-$combinaison (*sans ordre et sans remise*).

Exemple :
Donner toutes les combinaisons à deux éléments de l'ensemble $A = \{a; b; c; d\}$

Exemple :
Un jeu consiste à piocher trois boules parmi 5 numérotées de 1 à 5. Combien y a-t-il de tirages différents ?
Si $E = \{1; 2; 3; 4; 5\}$, alors $\{1, 2, 3\}$, est une $3-$combinaison de E. C'est la même que $\{2, 1, 3\}$.
Il y a $5 \times 4 \times 3$ combinaisons à 3 éléments possibles.

Théorème 1 : *Admis*

Le nombre de combinaisons à p éléments d'un ensemble E à n éléments est le nombre noté $\begin{pmatrix} n \\ p \end{pmatrix}$ (se lit « p parmi n » et s'appelle coefficient binômial) égal à
$$\begin{pmatrix} n \\ p \end{pmatrix} = \frac{n!}{p!(n-p)!}$$

A retenir
$\begin{pmatrix} n \\ p \end{pmatrix}$ représente en fait le nombre de façons de choisir p objets parmi n, sans ordre et sans remise.

Exemple :
Vous êtes 26 élèves et il y a 35 places dans la salle de classe. Combien y-a-t-il de façons de disposer les chaises vides ?

Exemple :
Combien y a-t-il de « mots » formé avec mes deux initiales et contenant exactement 4 fois celle de mon nom ? 2 fois celle de mon prénom ?

Remarques :

- $\binom{n}{0} = \binom{n}{n} = 1$
- $\binom{n}{1} = \binom{n}{n-1} = n$
- Calculatrice

Exemples :

- $\binom{4}{0} = \ldots$
- $\binom{4}{1} = \ldots$
- $\binom{4}{2} = \ldots$
- $\binom{4}{3} = \ldots$
- $\binom{4}{4} = \ldots$

Exemple :

Vérifier l'exemple du poker proposé en activité.

Combien y a-t-il de mains de 5 cartes avec l'as de coeur ?

Contenant exactement un as ?

Contenant exactement 3 as ?

Contenant au moins 3 as ?

Méthode pour l'outil à utiliser

Il faudra toujours se poser les deux questions suivantes :

- Les éléments peuvent-ils être répétés ?
- Les éléments sont-ils ordonnés ?

Critères	Avec répétition	Sans répétition
Avec ordre	Arbres	Arbres
Sans ordre	*Hors Programme*	Combinaisons

I-2.2 Propriétés des coefficients binômiaux

Propriété 1 :

Pour tout couple d'entiers naturels n et p tels que $0 \leq p \leq n$, on a

$$\binom{n}{p} = \binom{n}{n-p}$$

Et pour tout couple d'entiers naturels n et p tels que $0 \leq p \leq n-1$, on a

$$\binom{n}{p} + \binom{n}{p+1} = \binom{n+1}{p+1} \qquad \text{(Relation de Pascal)}$$

Preuve

Par le calcul :

1. $\binom{n}{p} = \dfrac{n!}{p!(n-p)!} = \dfrac{n!}{(n-p)!p!} = \dfrac{n!}{(n-p)!(n-(n-p))!} = \binom{n}{n-p}$
2. $\binom{n-1}{p-1} + \binom{n-1}{p} = \dfrac{(n-1)!}{(p-1)!(n-p)!} + \dfrac{(n-1)!}{p!(n-1-p)!} = \dfrac{(n-1)!p+(n-1)!(n-p)}{p!(n-p)!} = \dfrac{(n-1)![p+n-p]}{p!(n-p)!} = \dfrac{n!}{p!(n-p)!} = \binom{n}{p}$

On peut cependant retrouver ces résultats par des raisonnements ensemblistes :

Preuve (Suite) **à l'oral**

1. Si je forme un groupe de 3 élèves de la classe, j'obtiens du même coup un groupe de 23, puisque vous êtes 26.
   Si l'on compte le nombre de parties de A ayant p éléments dans un ensemble E à n éléments, il revient au même de compter le nombre de parties complémentaires de A.
2. (idée) Si je veux dénombrer tous les groupes de 3 parmi vous, soit 26 élèves. On peut distinguer deux types de groupes :
   - Ceux ne contenant pas Lise : il faut former des groupes de 3 parmi 25 élèves
   - Ceux contenant Lise : il faut compter les groupes de 2 parmi 25 (auquel on joint Lise pour obtenir 3 élèves)

   Un raisonnement similaire établit la relation de Pascal.

Exemple :

Le nombre de façons de choisir 2 délégués parmi 30 élèves est égal au nombre de façons de choisir 28 élèves non délégués parmi 30 :

$$\binom{30}{2} = \binom{30}{28}$$

Remarque : Ceci permet d'obtenir tous les $\binom{n}{p}$ de proche en proche à l'aide d'un tableau en forme de triangle, appelé triangle de Pascal (en hommage à ce dernier). On calcule de la façon suivante : pour trouver un certain coefficient, on additionne dans le tableau suivant les coefficients situés "juste au dessus" et "juste au dessus à gauche" entre eux.

n \ p	0	1	2	3	4	…	p-1	p	…	n-1	n
0	1										
1	1	1									
2	1	2	1								
3	1	3	3	1							
4	1	4	6	4	1						
5	1	5	10	10	5	1					
…											
$n-1$	1	$n-1$					$\binom{n-1}{p-1}$	$\binom{n-1}{p}$			1
n	1	n						$\binom{n}{p}$		n	1

Proposition 2 : *Formule du binôme*
Pour tous nombres réels a et b et pour tout entier natuel n non nul, on a

$$(a+b)^n = \sum_{k=0}^{n} \begin{pmatrix} n \\ p \end{pmatrix} a^{n-p} b^p$$

Exemples :

$$(a+b)^2 = \begin{pmatrix} 2 \\ 0 \end{pmatrix} a^{2-0}b^0 + \begin{pmatrix} 2 \\ 1 \end{pmatrix} a^{2-1}b^1 + \begin{pmatrix} 2 \\ 2 \end{pmatrix} a^{2-2}b^2 = 1\times a^2 + 2\times a\times b + 1\times b^2 = a^2+2ab+b^2$$

$$(a+b)^3 = a^3+3a^2b+3ab^2+b^3 \qquad \text{et} \qquad (a+b)^4 = a^4+4a^3b+6a^2b^2+4ab^3+b^4$$

Preuve **Admise**
Par récurrence, en remarquant que $(a+b)^{k+1} = a(a+b)^k + b(a+b)^k$ et en utilisant la relation de pascal au bon endroit.

Exemple :
A l'aide de cette formule et du triangle pascal on retrouve des résultats bien utiles :

1. pour $n=2$ $(a+b)^2 = \binom{2}{0}a^2b^0 + \binom{2}{1}a^1b^1 + \binom{2}{2}a^0b^2 = a^2+2ab+b^2$
2. pour $n=3$ $(a+b)^3 = a^3+3a^2b+3ab^2+b^2$
3. pour $n=4$ $(a+b)^4 = a^4+4a^3b^+6a^2b^2+4ab^3+b^4$

Exemple :
Développer $(x+2)^5$ et $(1-\sqrt{2})^4$

Exercice 3 **:**
Une main au poker est constituée de 5 cartes tirées d'un jeu de 52 cartes. Combien y a-t- il de mains contenant des carrés (XXXXY) ? des fulls (XXXYY) ? des brelans (XXXYZ) ? des doubles paires (XXYYZ) ? des paires (XXYZA) ?
Deux lettres identiques (par exemple XX) correspondent à deux cartes de même hauteur (par exemple deux dames).

Exercice 4 **:**
Le capitaine des pompiers de New-York réside à l'angle de la 7ème rue et de la 33ème avenue. La caserne se trouve à l'angle de la 15ème rue et de la 40ème avenue. Il s'y rend tous les jours à pied et sans perdre de temps (i.e. dans le sens des numéros croissants aussi bien pour les rues que pour les avenues). Sachant qu'il a commencé à travailler le jour de ses 18 ans, et sachant qu'il n'est jamais passé deux fois par le même chemin, quel est l'âge maximum du capitaine ?

II) Loi de Bernoulli et loi Binomiale

Travail de l'élève **2.** La probabilité qu'un tireur atteigne sa cible est $p = \frac{3}{4}$. On appelle Succès, noté S cet événement.

1. Il procède à un unique tir. Modéliser cette expérience en fonction de S.
2. On suppose qu'il fait deux tirs et on note X la variable aléatoire associant à cette épreuve le nombre de succès obtenus. ($X = 0$, 1 ou 2)
   - **a.** Calculer la probabilité des événements $(X = 0)$, $(X = 1)$ et $(X = 2)$. (On pourra s'aider d'un arbre et on désignera par S les succès et E les échecs)
   - **b.** Calculer $\sum_{k=0}^{2} P(X = k)$.
3. On suppose maintenant qu'il fait six tirs et on note Y le nombre de succès obtenus. $Y \in \{0; 1; \ldots; 6\}$. On voudrait calculer la probabilité de l'événement $(Y = 4)$.
   - **a.** Peut-on encore raisonner à l'aide d'un arbre ?
   - **b.** Calculer la probabilité qu'il commence par quatre succès suivis de deux échecs.
   - **c.** Mais les succès et les échecs n'apparaissent pas nécessairement dans cet ordre. Parmi les "mots" de six lettres qui ne contiennent que des S et des E, combien contiennent exactement quatre fois la lettre S ?
   - **d.** En déduire la probabilité de l'événement $(Y = 4)$.

II-1 Loi de Bernoulli

Définition 3 : *Variable aléatoire suivant une loi de Bernoulli*

Une épreuve de Bernoulli est une expérience aléatoire ne comportant que 2 issues (Succès et Echec). On note p la probabilité de succès.

Soit X la variable aléatoire qui est égale à 1 en cas de succès et 0 sinon. Alors, on dit que X suit une loi de Bernoulli de paramètre p.

Exemple :

- Pile ou Face.
- Lancer un dé et regarder si l'on obtient un 6 ou non.
- Aimer ou non Mireille Matthieu
- Etre ou ne pas être

Propriété 2 : *Espérance et variance d'une variable aléatoire suivant une loi de Bernoulli*

Si $X \hookrightarrow B(1; p)$ alors :

$$E(X) = p \qquad \text{et} \qquad V(X) = p(1-p)$$

Preuve

$$E(X) = P(X = 0) \times 0 + P(X = 1) \times 1 = 0 + p = p$$

$$V(X) = E(X^2) - E(X)^2 = p - p^2 = p(1-p)$$

II-2 Loi Binomiale

Définition 4 : *Schéma de Bernoulli*

Soit $n \in \mathbb{N}^*$. Lorsqu'on répète, de manière indépendante, n fois une même épreuve de Bernoulli de paramètre p. on dit que l'on fait un schéma de Bernoulli de paramètres n et p.

Exercice 5 :

On lance n dés ($n \geq 1$). On note A l'événement « obtenir au moins un 4 (sur l'ensemble des n lancers) ».

1. Décrire l'événement $\overline{A}$ à l'aide d'une phrase.

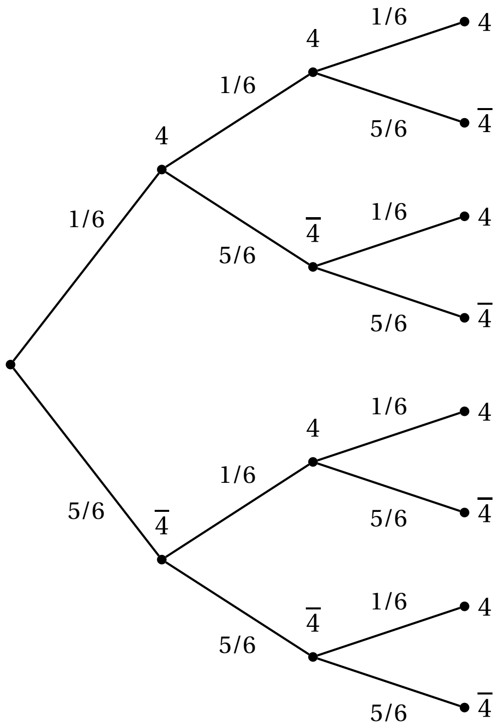
2. Faire un arbre et calculer $p(A)$ dans le cas où $n = 3$.

3. Dans cette question, on suppose n quelconque. Exprimer $p(A)$ en fonction de n.

4. Combien de dés faut-il lancer pour que la probabilité d'obtenir au moins un six soit supérieure à $\frac{3}{4}$?

Solution :

1. $\overline{A}$ est lévénement « ne pas obtenir de 4 ».
2.



Par conséquent

$$P(A) = 1 - P(\overrightarrow{A}) = 1 - \left(\frac{5}{6}\right)^3 = 1 - \frac{125}{216} = \frac{91}{216}$$

3. En raisonnement de la même manière on obtient :

$$P(A) = 1 - P(\overline{A}) = 1 - \left(\frac{5}{6}\right)^n$$

4. On cherche le plus petit entier n tel que :

$$\begin{aligned}
& & P(A) \geq \frac{3}{4} \\
&\iff & 1 - \left(\frac{5}{6}\right)^n \geq \frac{3}{4} \\
&\iff & -\left(\frac{5}{6}\right)^n \geq -\frac{1}{4} \\
&\iff & \left(\frac{5}{6}\right)^n \leq \frac{1}{4} \\
&\iff & \ln\left(\frac{5}{6}\right)^n \leq \ln\frac{1}{4} \\
&\iff & n\ln\left(\frac{5}{6}\right) \leq -\ln 4 \\
\iff \quad n \geq -\frac{\ln 4}{\ln\left(\frac{5}{6}\right)} = -\frac{\ln 4}{\ln 5 - \ln 6} \simeq 7,6 \iff & & n \geq 8
\end{aligned}$$

Nous devons donc lancer au moins 8 fois le dé pour être sûr à 75% d'obtenir au moins un 4.

Définition 5 : *Variable aléatoire suivant une loi binomiale*

Soit un schéma de Bernoulli de paramètres n et p. On note X la variable aléatoire égale au nombre de succès. (X est à valeurs dans $\{0;1;\dots;n\}$).

Dans ces conditions, on dit que la variable aléatoire X suit une loi binomiale de paramètres n et p.

Exemple :

Reprenons la situation précédente (lancer de 3 dés) et notons X le nombre de 4 obtenus.

X est à valeurs dans $\{0;1;2;3\}$.

Calculons la probabilité d'obtenir exactement deux 4. D'après les règles sur les arbres, on a :

$$P(X=2)=P(4-\overline{4}-4)+P(\overline{4}-4-4)+P(4-4-\overline{4})=\frac{1}{6}\times\frac{5}{6}\times\frac{1}{6}+\frac{5}{6}\times\frac{1}{6}\times\frac{1}{6}+\frac{1}{6}\times\frac{1}{6}\times\frac{5}{6}=3\times\left(\frac{1}{6}\right)^2\times\frac{5}{6}=\frac{5}{72}$$

Généralisons ce raisonnement :

Théorème 2 :

Soit X une variable aléatoire suivant une loi binomiale de paramètres n et p.

Pour tout $k\in\{0;1;\dots;n\}$:

$$P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}$$

Preuve

La probabilité d'avoir k succès suivis de $n-k$ échecs est :

$$p^k(1-p)^{n-k}$$

Mais les succès et les échecs n'apparaissent pas nécessairement dans cet ordre.... Voici un moyen de dénombrer toutes les possibilités d'apparition des succès et échecs : on considère l'ensemble des « mots »de n lettres qui ne contiennent que des S et des E. On sait qu'il y en a exactement $\binom{n}{k}$ qui contiennent k fois la lettre S (et donc $n-k$ fois la lettre E).

On en déduit :

$$P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}$$

Remarques :

- Si on note q la probabilité d'échec alors $P(X=k)=\binom{n}{k}p^kq^{n-k}$
- La probabilité d'avoir n succès est : $P(X=n)=p^n$
- La probabilité de n'avoir aucun succès est : $P(X=0)=q^n$

  Par conséquent, on retrouve des résultats déjà connus, comme la probabilité d'avoir au moins un succès est :

$$P(X\geq 1)=1-P(X=0)=1-q^n$$

Propriété 3 : *Espérance et variance d'une variable aléatoire suivant une loi de Bernoulli*

Si X suit une loi binomiale de paramètres $n\in\mathbb{N}^*$ et $p\in[0;1]$, alors :

$$E(X)=np \qquad \text{et} \qquad V(X)=np(1-p)$$

Preuve **Hors Programme**

$$E(X) = \sum_{k=0}^{n} P(X=k)k = \sum_{k=1}^{n} \binom{n}{k} p^k (1-p)^{n-k} k$$

Or

$$k\binom{n}{k} = k\frac{n!}{k!(n-k)!} = \frac{kn!}{k!(n-k)!} = \frac{n!}{(k-1)!(n-k)!}$$

et :

$$n\binom{n-1}{k-1} = \frac{n(n-1)!}{(k-1)!(n-1-k+1)!} = \frac{n!}{(k-1)!(n-k)!}$$

Par conséquent : $k\binom{n}{k} = n\binom{n-1}{k-1}$, ce qui donne :

$$E(X) = \sum_{k=1}^{n} n\binom{n-1}{k-1} p^k(1-p)^{n-k} = n\sum_{k=1}^{n}\binom{n-1}{k-1} p^k(1-p)^{n-k} = np\sum_{k=1}^{n}\binom{n-1}{k-1} p^{k-1}(1-p)^{n-1-(k-1)} = np\sum_{k=0}^{n-1}\binom{n-1}{k} p^k(1$$

Or d'après la formule du binôme de Newton on sait que :

$$(x+y)^{n-1} = \sum_{k=0}^{n-1}\binom{n-1}{k} x^k y^{n-1-k}$$

Ici donc :

$$E(X) = np(p+1-p)^{n-1} = np$$

Nous admettons la formule pour la variance.

Exemple :

Calculer l'espérance et la variance de Y dans l'activité sur le tireur.

Exercice 6 **:**

Dans lequel des cas suivants X suit-elle une loi binomiale ? Si oui, donner les paramètres de la loi et calculer l'espérance et la variance de X.

1. Dans une classe, on tire au sort sans remise 5 élèves, X est le nombre d'élèves abonné à Star'Ac mag dans le lot tiré au sort.
2. Dans un sac de 20 billes contenant 7 noires et 13 blanches, on tire avec remise 3 d'entre elles, X étant le nombre de billes noires obtenues.
3. On lance 10 dés, X est le nombre de 5 obtenus.
4. Un circuit comprend 2 lampes en série. Pour chacune d'elle, la probabilité qu'elle fonctionne est de 0.03. X est le nombres de lampes qui s'allument lorsqu'on appuie sur l'interrupteur. Même question avec cette fois des lampes en parallèles.

III) Lois continues

Il existe des variables aléatoires non discrètes, qui prennent toutes les valeurs d'un intervalle réel (borné ou non). On dit alors que la variable est continue. On s'intéresse alors à des événements du type X entre les réels a et b , $a \leq X$ ou encore $X \leq b$.

Exemples : le temps d'attente d'un bus, la durée de vie d'un appareil, la distance d'un point d'impact au centre d'une cible....

III-1 Du discret au continu

Travail de l'élève **3.** Activité p 366.

On choisit au hasard un nombre réel dans l'intervalle $[0;1]$. On cherche à modéliser cette expérience.

1. Préciser l'univers Ω de cette expérience.
   En quoi se distingue-t-il des univers déjà rencontrés ?
2. Lorsque Ω est un ensemble fini, définir une loi de probabilité, c'est donner la probabilité de chacun des événements élémentaires.
   Voyons si cette approche « point par point » est généralisable ici.
   a. Lequel des événements suivant paraît le plus probable :

   Obtenir 0 Obtenir $\dfrac{1}{2}$ Obtenir $\dfrac{\sqrt{3}}{2}$ Obtenir π ?

   b. Proposer une probabilité à leur associer.
   c. Supposons que la probabilité d'obtenir un nombre x de l'intervalle $[0;1]$ soit non nulle et notons $P(x) = k \neq 0$.
   On admet que les réels sont uniformément répartis. Les issues doivent évidemment être équiprobables.
   Considérons alors l'événement $E_n = \left\{\dfrac{1}{n};\dfrac{2}{n},\ldots,\dfrac{n}{n}\right\}$ un sous ensemble de Ω.

   i. Exprimer $P(E_n)$ en fonction de k et n.
   On suppose toujours valable le principe selon lequel la probabilité d'une réunion d'événements disjoints est la somme des probabilités de ces événements.
   ii. Montrer que pour n suffisamment grand, on a $P(E_n) > 1$
   iii. Que pouvez-vous en conclure sur k ?

   d. P étant une loi de probabilité modélisant le choix au hasard d'un réel dans $[0;1]$, donner $P(x)$ lorsque $x \notin [0;1]$ et $P(x)$ lorsque $x \in [0;1]$.
   e. Peut-on envisager de définir P en se donnant les $P(x)$ pour tout réel $x \in [0;1]$?
3. **a.** Quelle probabilité serait-il naturel d'associer à l'événement « Obtenir un réel inférieur à $\dfrac{1}{2}$?
   b. Que pourrait alors valoir de même : $P\left(\left[0;\dfrac{1}{4}\right]\right)$, $P\left(\left[\dfrac{1}{4};\dfrac{1}{2}\right]\right)$, $P([0.45;0.55])$ et $P([3;4])$?
   c. En généralisant ce raisonnement, proposer des probabilités pour $P([\alpha;\beta])$, $P(]\alpha;\beta])$, $P(]-\infty;\beta])$, $P([\alpha;+\infty[)$.
4. D'après ce que l'on a vu précédemment, la probabilité d'obtenir $\dfrac{1}{2}$ et celle d'obtenir π sont toutes les deux nulles. Pourtant, la probabilité d'obtenir un nombre voisin de $\dfrac{1}{2}$ ne l'est pas, et celle d'un nombre voisin de π si.
   On dira qu'il existe des réels affectés d'une densité de probabilité nulle (ceux en dehors de $[0;1]$) et d'autres d'une densité de probabilité non nulle (ceux dans $[0;1]$).
   Soit f la fonction constante égale à 1 sur $[0;1]$ et égale à 0 sinon.

Comparer la probabilité $P([\alpha;\beta])$ d'un intervalle $[\alpha;\beta]$ inclu dans $[0;1]$, et l'aire sous la courbe de la fonction f, limitée par les droites d'équation $x=\alpha$ et $x=\beta$.
Faire de même pour les intervalles $]-\infty;\beta]$ et $[\alpha;+\infty[$

Définition 6 :
Une densité de probabilité sur un intervalle I de $\mathbb{R}$ est une fonction f **continue**, **positive** sur I et qui vérifie

$$\int_I f(t)\,dt = 1$$

Une loi de probabilité p de densité f sur l'intervalle I est déterminée en associant à tout intervalle $[\alpha,\beta]$ contenu dans I, la probabilité de l'intervalle $[\alpha,\beta]$

$$P\big([\alpha,\beta]\big) = \int_\alpha^\beta f(t)\,dt$$

Remarques :

- Si I est un intervalle borné, on a simplement $\displaystyle\int_I f(t)\,dt = \int_a^b f(t)\,dt$
- Si I est du type $[a;+\infty[$ alors $\displaystyle\int_I f(t)\,dt = \lim_{x\to+\infty}\int_a^x f(t)\,dt$.
  La définition est analogue dans le cas $]-\infty;a]$
- Si $I=\mathbb{R}$ alors $\displaystyle\int_I f(t)\,dt = \lim_{x\to-\infty}\int_x^0 f(t)\,dt + \lim_{x\to+\infty}\int_0^x f(t)\,dt$ si elles existent.
- Grâce aux propriétés de l'intégrale, on retrouve les trois règles de définitions d'une probabilité :
  1. P à valeurs dans $[0;1]$
  2. $P(\Omega)=1$
  3. $P(A\cup B)=P(A)+P(B)$ pour tous événements A et B disjoints.
- La définition de la loi de probabilité P s'étend à des intervalles non bornés lorsque la limite existe.

Exemples :
Soit f une fonction constante sur un intervalle $[-1;3]$. Quelle doit être sa valeur pour que f soit une densité sur cet intervalle ? La loi associée à cette densité s'appelle la loi uniforme sur $[-1;3]$.

Démontrer que la fonction f définie sur sur $\mathbb{R}$ par $f(t)=3\exp^{-3t}$ est une densité de probabilité sur $\mathbb{R}^+$. La loi associée à cette densité s'appelle la loi exponentielle de paramètre 3.

Définition 7 :
Soit P une loi de probabilité sur I de densité f . On dit qu'une variable aléatoire X , à valeurs dans I, suit une loi de probabilité P lorsque pour tout sous intervalle $[a;b]$ de I, on a $P\big([a\le X\le b]\big) = \int_a^b f(t)\,dt$

Exemples :

Si X est une variable aléatoire qui suit une loi uniforme sur $[-1;3]$, alors : $P(a\le X\le b) = \int_a^b \frac{1}{4}\,dt = \dfrac{d-c}{4}$.
Si X est une variable aléatoire qui suit une loi exponentielle de paramètre 3 sur $\mathbb{R}^+$, et x un réel positif, alors $P(0\le X\le x) = \int_0^x 3e^{-3t}\,dt = 1-e^{-3x}$
Et par complémentarité : $P(X\ge x) = 1-P(0\le X\le x) = e^{-3x}$.

Remarque : On définit la probabilité conditionnelle de la même manière, tout comme l'indépendance.

III-2 Exemples de lois continues

Définition 8 :
On appelle loi uniforme sur $[a, b]$ la loi de probabilité dont la densité est la fonction constante f définie sur $[a, b]$ par

$$f(x) = \frac{1}{b-a}$$

Exemple :
Dans la journée, un métro passe toutes les 6 minutes à la station n°14.
Soit X le temps d'attente d'une personne à cette station. On suppose que X suit la loi uniforme sur [0 ; 6].
Quelle est la probabilité que cette personne attende entre 3 et 5 minutes ?

Définition 9 :
Soit λ un réel strictement positif. On appelle loi exponentielle de paramètre λ la loi de probabilité dont la densité est la fonction f définie sur $[0, +\infty[$ par

$$f(t) = \lambda e^{-\lambda t}$$

Exemple :
On suppose que la durée de vie X d'une voiture suit une loi exponentielle de paramètre 0.1.

1. Calculer la probabilité qu'une voiture dépasse 10 ans de durée de vie.
2. On sait qu'une voiture a duré déjà 10 ans.
   Quelle est la probabilité qu'elle dépasse 12 ans de durée de vie ?
3. Comparer le résultat précédent avec la probabilité que la durée de vie de la voiture dépasse deux ans :
   On constate que la probabilité que la voiture dure deux ans de plus ne dépend pas de son âge. On dit que X est une loi de durée de vie sans vieillissement.

Exercices du livre :
TP p 375

IV) Adéquation à une loi équirépartie

IV-1 Principe du test sur un exemple détaillé

Cette partie vous introduit dans le monde des *Statistiques inférentielles* (**Inférer** : tirer une conséquence de quelque proposition, de quelque fait, etc.) auquel vous êtes en fait constamment confronté(e) en tant que citoyen(ne) d'une société hautement médiatisée. La statistique inférentielle est en effet la science qui permet de « *modéliser une partie observable du réel comme résultant d'un phénoméne aléatoire pour lequel on envisage non pas une mais toute une famille de lois de probabilités possibles* » (J.P. Raoult - dossier APMEP- Déc 2005)
C'est un sujet très intéressant, exigeant une réflexion approfondie sur les notions abordées, les conclusions à émettre, entraînant un débat intéressant, permettant d'avoir une démarche scientifique partant d'une expérimentation. Malheureusement, nous sommes loin de disposer du temps nécessaire. Nous nous contenterons donc d'une préparation pure et simple aux exercices d'examen qui sont tous identiques, mis à part le contexte expérimental.
Dans tous les cas, il s'agira d'étudier les résultats d'une série d'expériences ayant un nombre fini k d'issues possibles.

Regardons en détail un exemple.

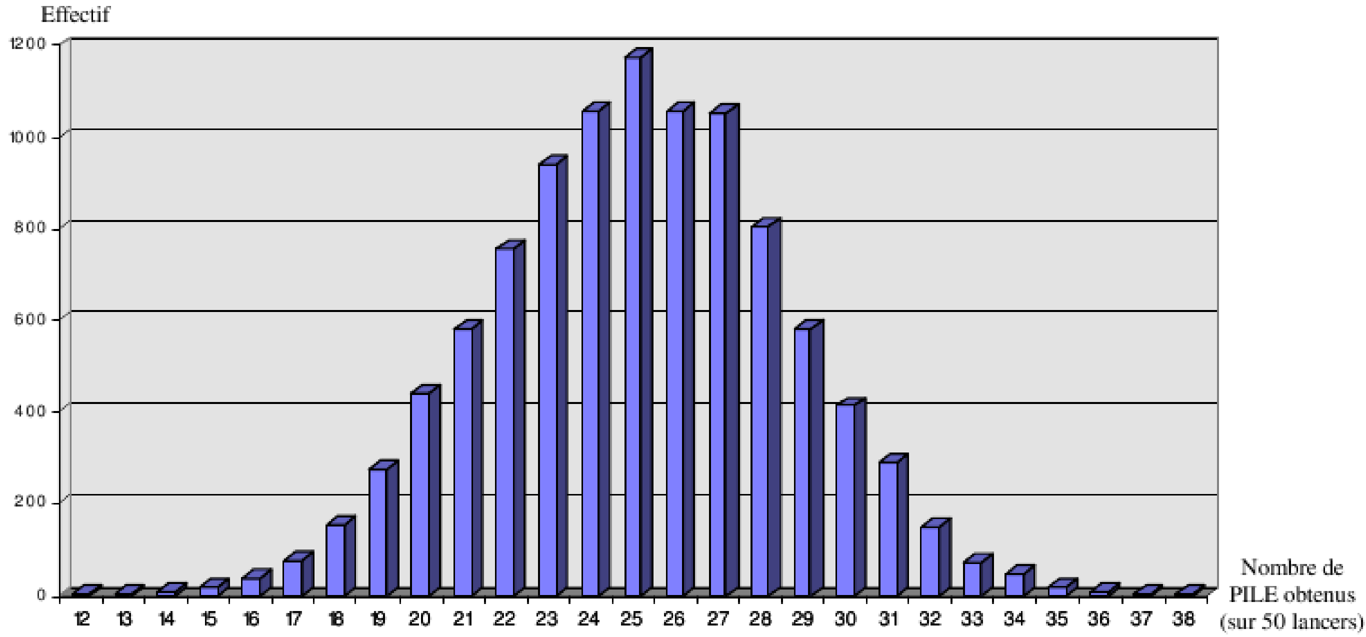
On dispose d'une pièce de monnaie et on souhaite savoir si elle est bien équilibrée. Mais comment le savoir ?

On la lance 50 fois (à la main, plus devient fastidieux). On a obtenu 30 fois Pile et 20 fois Face.

Peut-on raisonnablement estimer que la pièce est équilibrée ?

En fait, on ne peut jamais être sûr de rien, mais d'un point de vue statistique, on peut avoir une idée de la réponse, avec une certaine marge d'erreur, grâce aux simulations.

En effet, on peut simuler par un ordinateur 10000 fois l'expérience de 50 lancers, pour une pièce équilibrée (puisqu'on veut tester l'hypothèse d'équirépartition de notre propre pièce). On l'a fait sur tableur et on a obtenu :



On constate que sur cette simulation, à peine plus d'un dixième des expériences donnent exactement 25 Piles et 25 Faces. De plus, pour beaucoup d'entre elles, le nombre de Pile et de Face est assez éloigné de cette répartition, alors que la pièce est équilibrée. On parle de **fluctuation d'échantillonnage**.

On constate également qu'une certaine proportion (pas forcément négligeable) de ces 10000 expériences a obtenu 30 Pile et 20 Face. Cet événement obtenu lors de notre propre expérience n'est donc pas si improbable que cela pour une pièce équilibrée.

Mais on aimerait décider si notre pièce est, elle aussi, équilibrée, au risque de commettre une erreur.

Quels peuvent-être les critères de décisions ? Quelles sont alors les erreurs commises ?

Il est clair que plus les fréquences observées $\left(f_{Pile}, f_{Face}\right)$ sont éloignées des fréquences théoriques $\left(\frac{1}{2}, \frac{1}{2}\right)$, plus il est raisonnable de rejeter l'hypothèse d'équirépartition. Mais cela ne suffit pas pour décider.

On utilise comme indicateur d'erreur le carré de la distance euclidienne entre l'expérience et le modèle. En fait, on va comparer d^2 défini par

$$\mathrm{d}^2 = \sum_{i=1}^{k} \left(f_i - \frac{1}{2}\right)^2$$

à un paramètre de précision e arbitrairement choisi (et fourni par les énoncés de Bac), et dépendant de la simulation. Ici on obtient

$$d^2 = \left(f_{Pile} - p(Pile)\right)^2 + \left(f_{Face} - p(Face)\right)^2 = \left(\frac{30}{50} - \frac{1}{2}\right)^2 + \left(\frac{20}{50} - \frac{1}{2}\right)^2 = \frac{1}{50} = 0.02$$

Evidemment, plus une épreuve est proche des fréquences théoriques, plus son d^2 est proche de 0 et inversement.

Avec le tableur, on calcule d^2 pour chacune des simulations de 50 lancers, et on fait le tableau des effectifs cumulés croissants de ces d^2.

Répartition Pile-Face	25/25	24/26	23/27	22/28	21/29	20/30	19/31	18/32	17/33	16/34	15/35	14/36	13/37	12/38
Valeurs d^2	0	$\frac{1}{1250}$	$\frac{2}{625}$	$\frac{9}{1250}$	$\frac{8}{625}$	$\frac{1}{50}$	$\frac{18}{625}$	$\frac{49}{1250}$	$\frac{32}{625}$	$\frac{81}{1250}$	$\frac{2}{25}$	$\frac{121}{1250}$	$\frac{72}{625}$	$\frac{169}{1250}$
Effectif cumulés croissants	1171	3283	5277	6834	7996	8850	9412	9715	9862	9945	9980	9994	9998	10000

Maintenant, nous allons décider suivant une marge d'erreur fixée par l'énoncé, si nous considérons que notre pièce est équilibrée ou non. Pour cela, tout dépend en fait de la position de notre d^2 dans le tableau ci-dessus.
Par exemple, si on se donne une marge d'erreur de 10%, cela signifie que l'on considère que les 10% de d^2 les plus élevés obtenus par notre simulation sont marginaux.
Ainsi, si notre d^2 obtenu par notre expérience en fait partie, on rejettera notre hypothèse d'équirépartition, sinon, on ne pourra pas la rejeter (et par abus de langage, il est parfois dit qu'on la valide ou qu'on l'accepte).

Pour une marge d'erreur de 10%, on raisonne par rapport au 9^e décile. En effet, le 9^e décile D_9 est défini comme le réel tel qu'au moins 90% des d^2 soient inférieurs ou égaux à D_9. En général, l'énoncé fourni directement les quartiles, déciles, centiles ... utiles pour répondre. Mais on peut aussi nous donner le tableau ci-dessus, construire la courbe des effectifs cumulés croissants et le retrouver graphiquement (improbable), ou encore, l'énoncé peut fournir un diagramme à moustache.

Règle de décision

– Si $d^2 > D_9$, alors on rejette l'hypothèse d'équirépartition pour notre pièce, avec une marge d'erreur de 10%.
– Si $d^2 < D_9$, alors on ne peut pas rejeter l'hypothèse d'équirépartition pour notre pièce, avec une marge d'erreur de 10%.

Ici, $D_9 = 0.022$. Cela signifie donc qu'au moins $\frac{9}{10}$ des valeurs de d^2 sont inférieures ou égales à $D_9 = 0.022$.
Comme $d^2 < D_9$ donc on ne peut pas rejeter l'hypothèse d'équirépartition, avec une marge d'erreur de 10%.
Si on nous avait donné le 95^e centile, on aurait effectué un test avec une marge d'erreur de 5%, etc.

Et si nous avions obtenu 32 Pile et 18 Face avec une marge d'erreur de 10% ?

Même question avec une marge d'erreur de 1% (On nous donne le 99^e centile : $C_{99} = \frac{3}{5}$)

Remarque : Ceci paraît paradoxal, mais en fait, une marge d'erreur de 1% signifie que l'on considère que seulement 1% des résultats de la simulation sont marginaux. Il est donc moins probable que notre observation tombe dans ces 1% que dans 10%. La marge d'erreur à 1% a donc tendance à valider « trop » d'observations comme équirépartie. Et le cas marge d'erreur à 0% ?
A l'inverse, un modèle avec une marge d'erreur de 90%, on rejetterait « trop » d'observations.
Certaine statisticiens considèrent douteux et trop beaux pour être vrai de tels modèles.

IV-2 Exemples

On a lancé un dé 200 fois un dé cubique et on a obtenu les résultats suivants :

Face	1	2	3	4	5	6
Nombre de sorties de chaque face	28	33	34	29	26	50

$1\,000$ simulations de 200 lancers de dé équilibré ont été réalisées sur un tableur. Pour chaque expérience, on a calculé $d^2 = \sum_{k=1}^{6}\left(f_k - \frac{1}{6}\right)^2$ où f_k représente pour l'expérience, la fréquence observée du numéro obtenu k. On a alors obtenu une série statistique pour laquelle on a calculé le 99^e centile qui vaut 0,01258 et le 95^e centile qui vaut 0,00888.
Peut-on accepter l'hypothèse d'un dé équilibré au seuil de 1% ? 5% ?

Exercice 7 : Bac : Dé tétraédrique

Partie A

On dispose d'un dé en forme de tétraèdre régulier, possédant une face bleue, deux faces rouges et une face verte ; on suppose le dé parfaitement équilibré.
Une partie consiste à effectuer deux lancers sucessifs et indépendants de ce dé. A chaque lancer on note la couleur de la face cachée. On considère les évènements suivants :

E est l'évènement « à l'issue d'une partie, les deux faces notées sont vertes »,
F est l'évènement « à l'issue d une partie, les deux faces notées sont de la même couleur ».

1. Calculer les probabilités des évènements E et F ainsi que la probabilité de E sachant F.
2. On effectue dix parties identiques et indépendantes.
   Calculer la probabilité d'obtenir au moins deux fois l'évènement F au cours de ces dix parties (on en donnera une valeur approchée décimale à 10^{-3} près).

Partie B

On souhaite savoir si le dé utilisé peut être considéré comme parfaitement équilibré.
Pour cela on numérote de 1 à 4 les quatre faces de ce dé, puis on lance, ce dé 160 fois en notant le nombre n_i de fois où chaque face est cachée ; on obtient les résultats suivants :

face i	1	2	3	4
effectif n_i	34	48	46	32

On note f_i la fréquence relative à la face n_i et d^2_{obs} le réel $\sum_{i=1}^{4}\left(f_i - \frac{1}{4}\right)^2$.
On simule ensuite 1 000 fois l'expérience consistant à tirer un chiffre au hasard 160 fois parmi l'ensemble (1 ; 2 ; 3 ; 4) puis, pour chaque simulation, on calcule
$d^2 = \sum_{i=1}^{4}\left(F_i - \frac{1}{4}\right)^2$, où F_i est la fréquence d'apparition du nombre i. Le 9^{e} décile de la série statistique des 1 000 valeurs de d^2 est égal à 0,009 8.
Au vu de l'expérience réalisée et au risque de 10 %, peut-on considérer le dé comme parfaitement équilibré ?

Exercice 8 : Bac : boîte à moustache

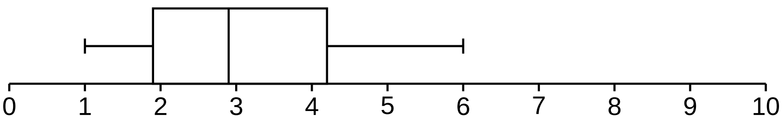
Les guichets d'une agence bancaire d'une petite ville sont ouverts au public cinq jours par semaine : les mardi, mercredi, jeudi, vendredi et samedi.
Le tableau ci-dessous donne la répartition journalière des 250 retraits d'argent liquide effectués aux guichets une certaine semaine.

Jour de la semaine	mardi	mercredi	jeudi	vendredi	samedi
Rang i du jour	1	2	3	4	5
Nombre de retraits	37	55	45	53	60

On veut tester l'hypothèse « le nombre de retraits est indépendant du jour de la semaine ». On suppose donc que le nombre des retraits journaliers est égal à $\frac{1}{5}$ du nombre des retraits de la semaine.

On pose $d^2_{\text{obs}} = \sum_{i=1}^{5}\left(f_i - \frac{1}{5}\right)^2$ où f_i est la fréquence des retraits du i-ème jour.

1. Calculer les fréquences des retraits pour chacun des cinq jours de la semaine.
2. Calculer alors la valeur de $1\,000 d^2_{\text{obs}}$ (la multiplication par $1\,000$ permet d'obtenir un résultat plus lisible).
3. En supposant qu'il y a équiprobabilité des retraits journaliers, on a simulé $2\,000$ séries de 250 retraits hebdomadaires. Pour chaque série, on a calculé la valeur du $1\,000 d^2_{\text{obs}}$ correspondant. On a obtenu ainsi $2\,000$ valeurs de $1\,000 d^2_{\text{obs}}$.
   Ces valeurs ont permis de construire le diagramme en boîte ci-dessous où les extrémités des « pattes » correspondent respectivement au premier décile et au neuvième décile.

   Lire sur le diagramme une valeur approchée du neuvième décile.
4. En argumentant soigneusement la réponse, dire si pour la série observée au début, on peut affirmer, avec un risque d'erreur inférieur à 10 %, que « le nombre de retraits est indépendant du jour de la semaine » ?

Exercice 9 **: Bac : π**

Les 1 000 premières décimales de π sont données ici par un ordinateur :

1415926535 8979323846 2643383279 5028841971 6939937510 5820974944 5923078164 0628620899 8628034825 3421170679 8214808651 3233066470 9384460959 0582235725 3594085234 8111745 ...

En groupant par valeurs entre 0 et 9 ces décimales, on obtient le tableau suivant :

Valeurs	0	1	2	3	4	5	6	7	8	9
Occurrences	93	116	102	102	94	97	94	95	101	106

Avec un tableur, on a simulé 1 000 expériences de 1 000 tirages aléatoires d'un chiffre compris entre 0 et 9.

Pour chaque expérience, on a calculé $d^2 = \sum_{k=0}^{k=9}\left(f_k - 0,1\right)^2$ où f_k représente, pour l'expérience, la fréquence observée du chiffre k.

On a alors obtenu une série statistique pour laquelle on a calculé le premier et neuvième décile (d_1 et d_9), le premier et troisième quartile (Q_1 et Q_3) et la médiane (Me) :

$d_1 = 0,000\,422$; $Q_1 = 0,000\,582$; Me $= 0,000\,822$; $Q_3 = 0,001\,136$; $d_9 = 0,001\,45$.

En effectuant le calcul de d_2 sur la série des 1 000 premières décimales de π, on obtient :

☐ $0,000\,456$ ☐ $0,004\,56$ ☐ $0,000\,314$

Un statisticien découvrant le tableau et ignorant qu'il s'agit des décimales de π, fait l'hypothèse que la série est issue de tirages aléatoires indépendants suivant une loi équirépartie. Il prend un risque de 10 % de rejeter cette hypothèse quand elle est vraie. Accepte-t-il cette hypothèse ?

☐ Oui ☐ Non ☐ Il ne peut pas conclure.