A. Dénombrement

A-I. Qu'est ce que c'est ?

De très nombreuses situations peuvent être décrites à l'aide de paramètres prenant un nombre fini de valeurs. Il est souvent intéressant de compter combien de cas différents sont possibles suivant les valeurs des paramètres : on dit qu'on fait du dénombrement.

Par exemple, si au restaurant on a le choix entre trois entrées, deux plats principaux et deux desserts, combien de compositions différentes peut avoir un « repas » ?

Le but des techniques de dénombrement est de mettre en évidence les cas les plus typiques, ceux auxquels on pourra se référer pour traiter les cas plus réels et plus complexes qu'on rencontrera ensuite.

A-II. Factorielles

a. Présentation et notation

On dispose d'un ensemble à n éléments, et on souhaite savoir combien il y a de façons de former une liste ordonnée avec ces éléments. Ce nombre de façons de « ranger » un ensemble à n éléments s'appelle le nombre de permutations à n éléments. On le note $n!$

Par exemple, si un ensemble contient a et b, on peut le ranger de deux façons : soit (a ; b) soit (b ; a) et on a donc $2!=2$.

Autre exemple, si un ensemble est vide... il n'y a qu'une façon de faire une liste avec ces éléments, c'est une liste toute blanche (ou « vide ») et on a donc $0!=1$.

b. Formules de calcul

Pour créer une liste à partir de n éléments, on choisit le premier parmi les n possibles... il y a n possibilités, puis on choisit le deuxième parmi les $n-1$ qui restent... il y a $n-1$ possibilités etc. jusqu'à ce qu'on choisisse le dernier « parmi » le seul qui reste... il n'y a plus qu'une possibilité.

Comme à chaque choix du premier peuvent correspondre tous les choix du deuxième qui eux-mêmes peuvent être associés à tous les choix du troisième... on obtient le nombre total de listes possibles en multipliant n par $n-1$ puis par $n-2$ etc. jusqu'à 1 :

$$n!=n.(n-1).(n-2)...(2).(1)$$

On peut aussi remarquer que n! s'exprime très facilement par récurrence :

$$\begin{cases} 0! = 1 \\ n! = n.(n-1)! \quad pour \;\; n \geq 1 \end{cases}$$

c. Résumé fondamental

$n!$ est le nombre de façons de ranger un ensemble à n éléments :

0! = 1

pour n>0, n! = n.(n-1).(n-2).(n-3)...(2).(1)

ou n! = n.(n-1)!

A-III. Arrangements

a. Présentation et notation

Le problème est presque le même : on dispose encore d'un ensemble à n éléments mais on ne souhaite en ranger que p (avec $p \leq n$ bien sûr). Le nombre de façons de ranger p éléments parmi n est appelé nombre d'arrangements à p éléments parmi n et se note A_n^p.

On rappelle qu'une liste est ordonnée (ce n'est pas un ensemble) et se note entre parenthèses.

Par exemple, si on dispose de 3 éléments a, b c et si on souhaite n'en ranger que 2, on a les listes possibles (a ; b), (a ; c), (b ;a), (b ; c), (c ; a), (c ; b). Comme ces 6 listes sont les seules on a : $A_3^2 = 6$.

b. Formules de calcul

Le raisonnement fait pour les permutations reste valable, mais comme on ne range que p éléments on s'arrêtera lorsque le produit écrit comportera p termes... c'est à dire au facteur (n-p+1).

$$A_n^p = n.(n-1)\ldots(n-p+1)$$

Ce résultat peut aussi s'écrire, et plus simplement, avec la notation factorielle :

$$A_n^p = \frac{n!}{(n-p)!}$$

A-IV. Combinaisons

a. Présentation et notation

Il ne s'agit plus d'une question de rangement mais on part toujours d'un ensemble à n éléments. Le problème est de savoir combien de parties à p éléments on peut trouver dans cet ensemble, étant bien entendu que les parties ne sont pas ordonnées, ce sont des sous ensembles.

On rappelle qu'un ensemble (forcément non-ordonné) se note entre accolades.

Par exemple, si E est l'ensemble {a, b, c} on peut chercher toutes les parties de E en les triant d'après leurs tailles :

Parties à 0 élément	{ }		
Parties à 1 élément	{a}	{b}	{c}
Parties à 2 éléments	{a,b}	{a,c}	{b,c}
Parties à 3 éléments	{a,b,c}		

On trouve en tout 8 parties ayant de 0 à 3 éléments.

Le nombre de parties à p éléments dans un ensemble à n éléments se note C_n^p ou de façon plus « moderne » $\begin{pmatrix} n \\ p \end{pmatrix}$… attention, l'ordre de n et p n'est pas le même :

$C_3^0 = 1$	$C_3^1 = 3$	$C_3^2 = 3$	$C_3^3 = 1$

b. Formules de calcul et propriétés élémentaires

Une relation simple existe entre A_n^p et C_n^p : pour obtenir le nombre de listes (ordonnées) à p éléments parmi n, il suffit de multiplier le nombre de parties à p éléments par le nombre de façons de ranger ces p éléments. On obtient la formule de départ dont toutes les autres proviennent :

$$\boxed{A_n^p = C_n^p \, . \, p!}$$

On en déduit :

$$\boxed{C_n^p = \frac{n!}{p!.(n-p)!}}$$

c. Propriétés classiques

On remarque que, quel que soit l'entier n :

$$A_n^n = A_n^{n-1} = n! \text{ (pour n>0)}$$

$$A_n^0 = 0! = 1 \text{ (pour n≥0)}$$

et que :

Pour n positif ou nul :	$C_n^0 = 1 \; et \; C_n^n = 1$
Pour n positif ou nul et p inférieur à n :	$C_n^p + C_n^{p+1} = C_{n+1}^{p+1}$

Cette dernière propriété a une démonstration intéressante qui n'utilise pas de formule.

Soit E un ensemble contenant $n+1$ éléments dont un que l'on appelle X. Les parties à $p+1$ éléments de E sont de deux types : celles qui contiennent X et celles qui ne le contiennent pas.

- Les parties qui contiennent X s'obtiennent en prenant X et p autres éléments parmi les n qui restent : il y en a C_n^p .
- Les parties qui ne contiennent pas X s'obtiennent en prenant $p+1$ éléments parmi les n qui ne sont pas X : il y en a C_n^{p+1} .
- En tout le nombre de parties à $p+1$ éléments dans un ensemble à $n+1$ éléments s'obtient en additionnant les deux résultats précédents... d'où la formule.

Cette démonstration fait que la formule $C_n^p + C_n^{p+1} = C_{n+1}^{p+1}$ est appelée formule de partition.

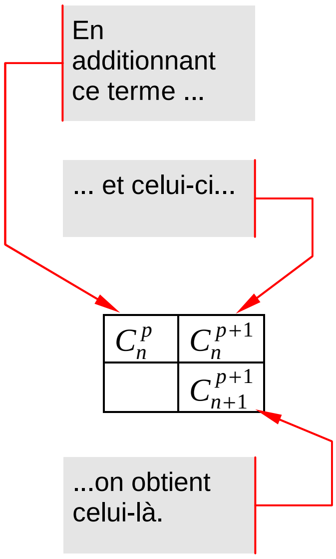
A-V. Applications classiques en mathématiques

a. Triangle dit « de Pascal »

Il s'agit d'une présentation commode de la propriété $C_{n+1}^{p+1} = C_n^p + C_n^{p+1}$ sous la forme d'un tableau triangulaire. Cette présentation connue plus de trois siècles avant Pascal mais popularisée sous son nom est très utilisée pour calculer rapidement, à la main, les nombres de combinaisons dans un ensemble ayant jusqu'à 10 éléments.

On présente en mettant n en tête de ligne et p en tête de colonne.

p =	0	1	2	3	4	5	6	7	8	9
n=0	1									
n=1	1	1								
n=2	1	2	1							
n=3	1	3	3	1						
n=4	1	4	6	4	1					
n=5										
n=6										
n=7										
n=8										
n=9										



La première colonne ne contient que des 1 car pour tout n on a $C_n^0 = 1$, de même la diagonale ne contient que des 1 car pour tout n on a $C_n^n = 1$. Pour les autre cases, on obtient son contenu en additionnant le terme situé juste au dessus et celui situé à la gauche de ce dernier.... autrement dit $C_{n+1}^{p+1} = C_n^p + C_n^{p+1}$.

Rien de tel pour comprendre comment ça marche que de compléter le tableau ci dessus...

b. Formule du binôme de Newton

Newton s'est demandé comment calculer vite et bien les produits remarquables du genre $(a+b)^n$ et, en remarquant que ce n'est que le produit de n facteurs tous égaux à $(a+b)$, il a conclut que ce produit était $\sum_{p=0}^{p=n} C_n^p \, a^{n-p}.b^p$.

En effet, pour calculer ce produit il faut, en envisageant tous les choix possibles d'un terme dans chaque facteur effectuer les différents produits et les additionner :

- Si on choisit a dans chaque facteur on obtient a^n et cela une seule fois.
- Si on choisit a dans tous les facteurs sauf un, on obtient $a^{n-1}.b$ et cela autant de fois qu'il y a de façon de choisir le facteur où on ne prend pas a.

$\Rightarrow$ Il y a C_n^1 telles façons.

- Si on choisit a dans tous les facteurs sauf deux, on obtient $a^{n-2}.b^2$ et cela autant de fois qu'il y a de façons de choisir les deux facteurs où on ne prend pas a.

$\Rightarrow$ Il y a C_n^2 telles façons.

- Enfin, plus généralement, si on choisit a dans tous les facteurs sauf dans p d'entre eux, on obtient $a^{n-p}.b^p$ et cela autant de fois qu'il y a de façons de choisir les p facteurs où on ne prend pas a.

$\Rightarrow$ Il y a C_n^p telles façons.

$$\boxed{(a+b)^n = \sum_{p=0}^{p=n} C_n^p \, a^{n-p}.b^p = \sum_{p=0}^{p=n} C_n^p \, a^p.b^{n-p}}$$

Cette formule a des conséquences nombreuses et commodes, dont par exemple :

$$C_n^0 + C_n^1 + \ldots + C_n^n = \sum_{p=0}^{p=n} C_n^p = \sum_{p=0}^{p=n} C_n^p.1^{n-p}.1^p = (1+1)^n = 2^n$$

$$C_n^0 - C_n^1 + \ldots + (-1)^p C_n^p + \ldots + (-1)^n C_n^n = \sum_{p=0}^{p=n} (-1)^p C_n^p = \sum_{p=0}^{p=n} C_n^p.1^{n-p}.(-1)^p = (1-1)^n = 0$$

A-VI. Exercices résolus et à résoudre

a. Dénombrement usuel

1) Une course de chevaux comporte 15 chevaux au départ. Combien y a-t-il de paris différents possibles pour le tiercé (pari portant sur les trois chevaux arrivant en tête) d'une part dans l'ordre et d'autre part dans le désordre.

2) Au bridge, une donne est un partage de 52 cartes en quatre « mains » de 13 cartes chacune. Combien y a-t-il de donnes différentes au bridge ? Peut-on donner une évaluation de ce nombre avec une calculette ? En admettant qu'un joueur (intoxiqué) fasse dix parties par jour, chaque jour de sa vie, combien de temps lui faudrait-il pour faire toutes les parties possibles ?

3) Encore au bridge, combien y a-t-il de « mains » différentes qui contiennent :

- les quatre As ?
- l'As de cœur et le 4 de pique ?
- l'As de cœur et quatre piques ?

4) Une assemblée comportant huit hommes (dont Roméo et Rodrigue) et six femmes (dont Chimène et Juliette) choisit ses représentants. Les représentants doivent être quatre.

Combien de choix différents sont possibles si...

- il doit y avoir au moins une femme parmi les représentants ?
- Juliette n'accepte d'être représentante que si Roméo l'est aussi ?
- Chimène refuse d'être représentante si Rodrigue l'est aussi.

b. Formules...

1) Démontrer que, pour n supérieur à 2, $C_n^p = C_{n-1}^{p-2} + 2.C_{n-1}^{p-1} + C_{n-1}^p$.

2) Combien valent

a) $\sum_{p=0}^{p=n} C_n^p$ b) $\sum_{p=0}^{p=n} 2^p.C_n^p$ c) $\sum_{p=0}^{p=n} x^n.C_n^p$

3) Ecrire les formes développées de $(1+x)^n$ pour n dans {2, 3, 4, 5, 6}

4) Montrer que l'équation $C_n^5 = C_{n-1}^6$ n'a pas de solution.

c. Et quelques classiques

1) On appelle anagramme d'un mot tout arrangement des lettres de ce mot en supposant que des exemplaires différents d'une même lettre ne peuvent pas être distingués. On ne se préoccupe ni du sens ni de la possibilité de prononcer l'arrangement obtenu. Par exemple, un anagramme de DISQUE est QDSUIE.

- Combien d'anagrammes a le mot BIEN ?
- Combien d'anagrammes a le mot MALADE ? (Attention aux deux A qu'on ne peut pas distinguer l'un de l'autre)
- Combien d'anagrammes a le mot INFORMATIQUE ? Et le mot CREER ?

2) Combien y a-t-il de grilles de mots croisés de taille 8x8 comportant 10 cases noires ? Plus généralement, combien y a-t-il de grilles de mots croisés de taille $n \times p$ comportant k cases noires ?

3) Pour aller de chez moi à mon lieu de travail il y a sept feux tricolores (Vert, Orange, Rouge). Combien de situations différentes ces sept feux peuvent-ils me faire rencontrer ?

4) En C, en Pascal ou dans d'autres langages, un entier « classique » est codé sur 16 bits (« bit » = Binary Digit, à ne pas confondre avec un « bit » d'information…), combien d'entiers ont-ils au moins trois bits consécutifs à 1 ? Combien n'ont jamais trois bits consécutifs à 1 ?

B. Statistiques descriptives à une variable

B-I. Vocabulaire de base : population, individu, caractère

L'idée de base des statistiques est celle d'une enquête : on peut enquêter sur des gens, des animaux, des objets, des idées... Dans tous les cas, les « objets » soumis à enquête sont appelés des **individus** et l'ensemble des individus susceptibles d'être soumis à enquête forme la **population**.

En général, lors d'une enquête, on ne s'intéresse pas à toutes les caractéristiques des individus mais seulement à quelques-unes... et même souvent à une seule. La caractéristique étudiée est appelée **caractère** ou **variable statistique**.

Exemples d'enquêtes possibles:

- Demander aux étudiants de Mesures Physiques à Orsay combien de films ils ont vu en 96. Les individus sont les étudiants de Mesures Physiques à Orsay, la population est l'ensemble des étudiants de Mesures Physiques à Orsay, le caractère est le nombre de films vus en 96.
- Demander aux voitures stationnées sur les parkings de l'IUT combien elles ont de portes. Les individus sont les voitures...
- Demander à la secrétaire de 1ère année combien elle a comptabilisé d'absences excusées par semaine en 2007/2008. Les individus sont les semaines de l'année 2007/2008... La secrétaire de 1ère année n'est ici qu'un intermédiaire.

B-II. Caractères qualitatifs, caractères quantitatifs

a. Nature d'un caractère

Certains caractères correspondent à une quantité, ils s'expriment « naturellement » par des nombres et ces nombres sont « naturellement » ordonnés. Lorsqu'on obtient deux réponses à une enquête portant sur un tel caractère, il n'est pas stupide de faire des opérations entre les réponses, par exemple de calculer la moyenne.

D'autres caractères correspondent à une qualité, ils peuvent s'exprimer par des nombres mais alors ceux-ci ne sont que des codes, ils ne sont pas « naturellement ordonnés ». Lorsqu'on obtient deux réponses à une enquête portant sur un tel caractère, il est stupide de faire des opérations entre les réponses, par exemple de calculer la moyenne.

Prenons deux exemples utilisant la population des étudiants de MPhy :

Si on imagine l’enquête « Combien de livres avez-vous chez vous ? » le caractère étudié est quantitatif : calculer le nombre moyen de livres dont disposent les individus soumis à enquête a un sens.

Si on imagine l’enquête « Quel est le numéro de ligne du dernier bus que vous avez pris dans Paris ? » le caractère étudié est qualitatif : calculer la moyenne entre des numéros de ligne de bus est stupide.

b. Type d’un caractère quantitatif

Certains caractères quantitatifs sont tels que les réponses sont séparées les unes des autres : il arrive un moment où, entre deux réponses données il n’y a plus de possibilité intermédiaire. On dit que ces caractères sont discrets.

D’autres, au contraire, sont tels qu’entre deux réponses obtenues il y a toujours des possibilités intermédiaires. On dit que ces caractères sont continus.

Voici quelques exemples :

« Combien pesez-vous ? » [continu]

« Quelle est votre pointure de chaussures ? » [discret]

« Quel âge avez-vous ? » [continu]

« Combien d’anniversaires avez-vous passés ? » [discret]

B-III. Dépouillement, effectifs et fréquences

Lorsqu’une enquête a eu lieu, qu’on a collectionné les réponses des individus, il faut organiser, ranger, trier ces réponses pour pouvoir en extraire du sens. L’ensemble de ces actions forme ce qu’on appelle le DEPOUILLEMENT de l’enquête.

Le dépouillement peut prendre plusieurs formes, les unes purement numériques, n’utilisant que des tableaux de nombres, les autres utilisant différentes techniques de dessin. En général, on prépare le dépouillement avant de faire enquête... par exemple, les feuilles de dépouillement pour des élections législatives ou municipales sont prêtes avant que le dépouillement ait lieu, elles contiennent les noms des candidats, des cases pour noter le nombre de voix etc.. *Et si vous n’avez jamais participé à un dépouillement il serait grand temps de vous y mettre : les « vieux » commencent à fatiguer un peu !*

Cas d’un caractère discret

Le nombre d’individus qui ont donné une réponse est appelé EFFECTIF de cette réponse.

On compte pour chaque réponse obtenue l’effectif et on dresse un tableau dont la première ligne contient les valeurs du caractère et la deuxième les effectifs correspondants.

Par exemple, imaginons l'enquête « Combien d'ampoules électriques y a-t-il dans votre chambre ? ». Le caractère est quantitatif discret. En interrogeant dix personnes on peut (par exemple) obtenir les réponses successives 3, 5, 4, 4, 3, 6, 5, 5, 2, 6. Le dépouillement en tableau donne :

Caractère	2	3	4	5	6
Effectif	1	2	2	3	2

On peut aussi, pour chaque valeur du caractère, donner le pourcentage de ceux qui ont donné cette réponse, par rapport au nombre total d'individus. Ce pourcentage est appelé FREQUENCE de la réponse.

La même enquête que précédemment donne alors le dépouillement :

Caractère	2	3	4	5	6
Effectif	1	2	2	3	2
Fréquence	10%	20%	20%	30%	20%

Cas d'un caractère continu

Si un caractère est continu, il est impossible de prévoir toutes les réponses... car il y a une infinité de réponses possibles et on ne peut pas prévoir une infinité de cases, dont la plupart d'ailleurs resteraient inoccupées.

On regroupe les réponses en CLASSES en choisissant des frontières. Par exemple, si on fait l'enquête « Quel âge avez-vous ? » on obtiendra sûrement très vite des réponses du genre « 18 ans et 3 mois », « 18 ans 3 mois et 8 jours » etc. car il y a toujours des gens très fiers d'être un poil plus jeune (ou plus âgé) que le voisin. Pour préparer le dépouillement on peut imaginer de regrouper les réponses dans les classes de 0 à 16 ans, de 16 à 18 ans, de 18 à 25 ans, de 25 à 35 ans, plus de 35 ans. Ces classes sont celles correspondant aux diverses majorités suivant la loi Française.

Un problème se pose : que faire si une réponse se trouve juste sur une frontière ? On décide de comptabiliser les réponses d'un côté et pas de l'autre, c'est à dire qu'on utilise des intervalles semi-ouverts.

Supposons que l'enquête « Quel âge avez-vous ? » donne les réponses

19	19,5	32	24	21	28	20,5	35	24

On peut dépouiller de la façon suivante :

Age	0	16	18	25	35 (...et plus)
Effectif	0	0	6	3	1

Le fait que les valeurs du caractère soient placées aux bords des cases délimitant les classes indique la façon dont ces classes sont définies. Ici les classes sont [0;16[, [16;18[, [18;25[, [25;35[et [35;+∞ [

B-IV. Représentations graphiques

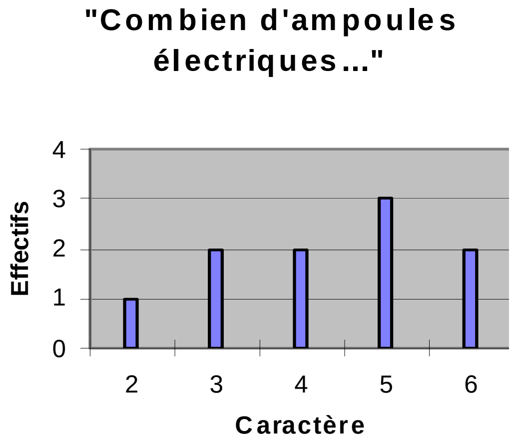
Le dépouillement étant fait on utilise divers procédés pour représenter le tableau d'effectifs ou de fréquences obtenus.

a. Diagrammes en bâtons

(caractère discret)

On représente les valeurs du caractère en abscisse et au dessus de chaque valeur atteinte par un individu on place un bâton dont la hauteur est proportionnelle à l'effectif.

L'enquête précédente sur les ampoules électriques donnerait alors le dessin ci-contre :



b. Histogrammes (caractère continu)

Si le caractère est continu on ne peut pas envisager de traiter les valeurs du caractère une par une (il risque d'y en avoir trop pour que ce soit compréhensible) et la plupart du temps l'effectif correspondant à une valeur du caractère serait 1 car il est peu probable que deux individus donnent exactement la même réponse si une infinité de réponses sont possibles.

On représente toujours les valeurs du caractère en abscisse mais au dessus de chaque classe on construit une colonne dont la SURFACE est proportionnelle à l'effectif de celle-ci.

Pour connaître la hauteur de la colonne on divise l'effectif de la classe par l'étendue de cette classe : le quotient obtenu est souvent appelé DENSITE de la classe.

Par exemple, si la classe [25 ; 30[a un effectif de 12, sa densité est 2,4.

$$\boxed{Densité = \frac{Effectif}{Etendue}}$$

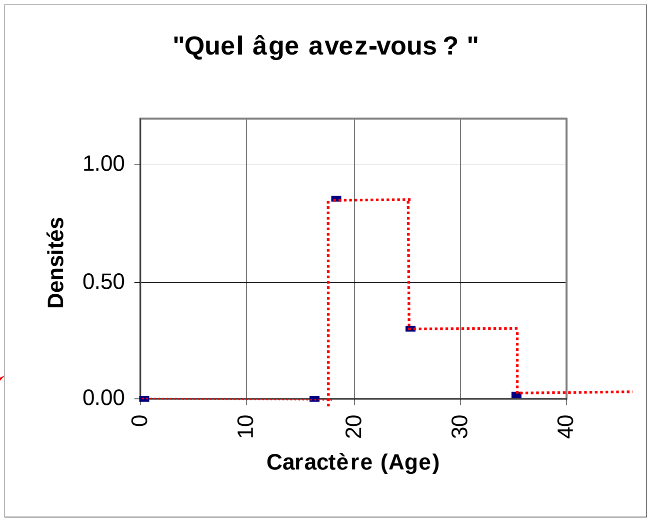
Il est à noter que la plupart des logiciels de Bureautique (Excel entre autres) appellent « Histogramme » des dessins qui ne sont en fait que des diagrammes à bâtons plus ou moins larges : ils ignorent en général la densité et représentent en ordonnées des grandeurs proportionnelles aux effectifs ce qui est trompeur dès que les classes n'ont pas toutes la même étendue. Pour cette raison, et pour d'autres du même genre, **on essaiera, dans la mesure du possible, de travailler avec 6 à 12 classes de même étendue (exceptionnellement 5 classes, jamais plus de 20).**

Voici ce qu'éventuellement on peut faire avec Excel pour obtenir, dans le cas des colonnes ayant des étendues différentes, un véritable histogramme... Les colonnes sont, ici, tracées à la main

Age	Classe	Effectif	Densité
0	[0 à 16[	0	0,00
16	[16 à 18[	0	0,00
18	[18 à 25[	6	0,86
25	[25 à 35[	3	0,30
35	[35 à 100[	1	0,02
100			

Tableau de calculs fait par Excel

Graphique « Excelo-Manuel »



c. Fonctions de cumul

Il est souvent intéressant de cumuler les effectifs (ou les fréquences) au fur et à mesure qu'on avance dans les valeurs du caractère.

Le mécanisme est simple : on imagine que le dessin représentant le dépouillement (diagramme en bâtons ou histogramme) est entièrement recouvert par une feuille de papier et on fait petit à petit avancer cette feuille de papier vers la droite. A chaque instant on peut repérer la position de la feuille de papier par la valeur de l'abscisse, x, qui vient d'être découverte et on peut lui associer le nombre d'individus N(x) découverts. La fonction, qui à x associe N(x), est appelée FONCTION DE CUMUL. Il est en général commode d'avoir le dessin correspondant au dépouillement et le dessin correspondant au cumul sur la même page, l'un en dessous de l'autre, de sorte que les abscisses, c'est à dire les valeurs du caractère, se correspondent d'un dessin à l'autre.

Lorsque le caractère est discret, le dépouillement est représenté par un diagramme en bâtons et la fonction de cumul est une fonction en escalier : sa représentation graphique est discontinue, elle subit un saut au franchissement de chaque bâton.

Lorsque le caractère est continu, le dépouillement est représenté par un histogramme et la fonction de cumul est continue : sa représentation graphique est formée de segments de droites obtenus en joignant les points correspondants aux frontières des classes...

Dans tous les cas, la fonction de cumul est l'intégrale de la densité :

$$N(x) = \int_{-\infty}^{x} densité(x).dx$$

Voyons tout cela sur deux exemples.

Cas d'un caractère discret.

Une enquête a donné le dépouillement suivant :

Caractère	0	1	2
Effectifs	3	5	1

La représentation de ce dépouillement est faite par un diagramme en bâtons :

La représentation de la fonction de cumul est discontinue, elle est en escalier.

.

Cas d'un caractère continu

Une enquête donne le dépouillement suivant :

Caractère	0		1		2		3
Effectifs		3		5		1	

La représentation de ce dépouillement est faite par un histogramme :

La représentation de la fonction de cumul est continue, elle est formée de segments de droites reliés.

A vous de compléter les dessins… c'est le seul moyen de vous assurer que vous avez compris !

Si, dans le cas d'un caractère continu, on représente les effectifs cumulés par des segments de droite, c'est parce que, ne disposant d'aucun renseignement sur la répartition des individus dans les classes on suppose que cette répartition est **uniforme**. Dans les cas, rares, où on dispose d'informations sur cette répartition, on peut rencontrer d'autres façons de faire... ce ne sera pas notre cas ici mais attention, ce sera le cas dès le début de la métrologie.

Lorsqu'on représente les FREQUENCES CUMULEES, la valeur initiale, tant qu'on n'a pas découvert un seul individu, est évidemment 0 % et la valeur finale, quand on a découvert tous les individus, est 100 %.

B-V. Paramètres de position

a. Le mode (nom masculin)

C'est, dans le cas d'un caractère discret, la valeur du caractère la plus fréquente, celle qui correspond au plus grand effectif.

Si le caractère est continu, on ne peut plus savoir quel est le mode car l'effectif qui lui correspond aura, en général, été regroupé avec ceux des valeurs voisines. On définit alors la classe modale comme étant la classe la plus fréquente, celle qui a le plus grand effectif.

Bien entendu il arrive qu'une statistique discrète possède plusieurs modes et qu'une statistique continue possède plusieurs classes modales.

b. La médiane

C'est, si cela a un sens, la valeur « Me » du caractère telle que 50% de l'effectif donne des réponses inférieures à Me et 50% donne des réponses supérieures à Me.

Lorsque la statistique porte sur un petit nombre d'individus, cette notion globale est insuffisante et on définit alors la médiane comme la valeur de l'individu n° E (N/2)+1 si N est l'effectif total et les individus rangés en ordre croissant d'après leurs réponses.

Par exemple, pour une population de 10 individus, la médiane serait la valeur du 5ème, et pour une population de 11 individus ce serait la valeur du 6ème.

E(n) est la partie entière de n : c'est, parmi les entiers inférieurs ou égaux à n, celui qui est le plus grand.
E(3,14) = 3
E(-1.98) = -2

c. La moyenne

Si les valeurs du caractère sont notées x_i et les effectifs correspondants n_i la moyenne est le nombre noté $\overline{x}$ et tel que :

$$\boxed{\overline{x} = \frac{\sum_{i=1}^{k} n_i x_i}{\sum_{i=1}^{k} n_i}}$$

Cette notion s'applique évidemment mieux au cas des caractères discrets... mais rien n'interdit, s'il s'agit d'un caractère continu, de revenir à une représentation discrète en utilisant pour chaque classe la valeur centrale de celle-ci dans le calcul de moyenne. Ce sont ces petites modifications qui permettent aux statisticiens malhonnêtes de pousser les résultats dans un sens ou un autre et ce sont ces mêmes petites modifications qui font une mauvaise réputation aux statistiques... pourtant un calcul (juste !) en soi n'a rien de malhonnête !

B-VI. Paramètre de dispersion

a. La variance

Il s'agit maintenant de « mesurer » la façon dont les valeurs du caractère sont réparties et plus précisément la façon dont elles sont dispersées. En effet, si on sait par exemple que la moyenne des notes dans un groupe de TD est 11 cela peut recouvrir des réalités très différentes allant du « tout le monde a 11 » à « la moitié a 2 et l'autre moitié a 20 ». Dans le premier cas, l'enseignant peut trouver un rythme et un type d'explication qui convient à tous ou presque alors que dans le deuxième cas il est impossible de faire suivre les mêmes explications à l'ensemble du groupe.

Le principe général est le suivant : on utilise comme référence la moyenne et, pour chaque individu on calcule l'écart e_i entre sa réponse x_i et la moyenne. On fait enfin la moyenne des écarts et on regarde ce qu'on obtient.

Si l'écart est calculé par une simple différence (on appelle cet écart « écart algébrique ») on remarque (et on démontre, voir exercice ci-dessous) que la moyenne des écarts est toujours nulle... donc cette méthode ne présente aucun intérêt.

Si l'écart est la valeur absolue de la différence on remarque que le seul cas où la moyenne des écarts est nulle est celui où tous les individus ont donné la même réponse (concentration maximum) et plus les individus sont dispersés, plus la moyenne obtenue est grande... c'est bien mais les calculs avec les valeurs absolues sont trop lourds pour que cette méthode soit pratique : elle est très rarement employée.

Si l'écart est le carré de la différence on fait les mêmes remarques que pour les valeurs absolues et on évite les difficultés liées à l'emploi de ces valeurs absolues : c'est la méthode la plus fréquemment employée, le résultat obtenu est la **variance**.

Pour une série non classée (on dit aussi série énumérée) :

$$\boxed{Variance = \frac{1}{N}.\sum_{i=1}^{i=N}(x_i - \bar{x})^2}$$

Pour une série classée, comprenant k classes d'effectifs n_i avec $N = \sum_{i=1}^{i=k} n_i$:

$$\boxed{Variance = \frac{1}{N}.\sum_{i=1}^{i=k} n_i.(x_i - \bar{x})^2}$$

On peut calculer une variance « à la main » mais il faut savoir que toutes les calculettes (même la TI 30X) font ce calcul de façon automatisée.

En développant les carrés on peut aussi démontrer que la variance se calcule par une formule où la moyenne n'intervient qu'en dehors de la somme :

$$\boxed{Variance = \frac{1}{N}.\left(\sum_{i=1}^{i=k} n_i.x_i^2\right) - \bar{x}^2}$$

b. L'écart-type

La variance pose un problème d'origine physique : elle n'est pas homogène en **dimension** à la quantité mesurée. Si par exemple une enquête porte sur une longueur mesurée en mètres alors la variance de la distribution statistique obtenue est en mètres-carrés. On règle ce problème en prenant la racine carrée de la variance... cette racine étant appelée **ECART-TYPE**.

Pour une série non-classée :

$$\boxed{EcartType = \sqrt{Variance} = \sqrt{\frac{1}{N}.\sum_{i=1}^{i=N}(x_i - \overline{x})^2}}$$

Pour une série classée :

$$\boxed{EcartType = \sqrt{Variance} = \sqrt{\frac{1}{N}.\sum_{i=1}^{i=k}n_i\,(x_i - \overline{x})^2}}$$

B-VII. Exercices à résoudre

1) Démontrer la formule $Variance = \frac{1}{N}.\left(\sum_{i=1}^{i=k} n_i.x_i^2\right) - \overline{x}^2$ en partant de la formule $Variance = \frac{1}{N}.\sum_{i=1}^{i=k} n_i.(x_i - \overline{x})^2$

2) Se plonger dans le mode d'emploi de sa calculette... puis, avec son aide...

Calculer la moyenne, la variance et l'écart-type des séries énumérées suivantes :

Série 1	3	4	5	6	7	8	9	10	11	12
Série 2	2	2	4	4	4	5	6	6	7	8

Calculer la moyenne, la variance et l'écart-type des séries classées suivantes :

Série 1	Valeurs	1	5	6	7	8	12	13	14	15
	Effectifs	2	3	4	4	8	5	4	1	0
Série 2	Valeurs	*3.4*	*3.8*	*3.9*	*4.0*	*4.1*	*4.2*	*4.5*	*4.6*	*5.0*
	Effectifs	5	14	22	84	78	61	21	8	1

3) Après avoir fait les calculs de l'exercice 2) représenter chacune des séries données par la courbe des fréquences cumulées, tracer les droites d'équations

$x=\bar{x}$, $x=\bar{x}+\sigma_x$ *et* $x=\bar{x}-\sigma_x$. En déduire le pourcentage de la population donnant des réponses entre $x=\bar{x}-\sigma_x$ *et* $x=\bar{x}+\sigma_x$

4) Démontrer que l'écart algébrique moyen de toute série statistique est nul.

C. Probabilités intuitives

C-I. Présentation

La notion de probabilité est issue de l'observation statistique des jeux de hasard et plus particulièrement de la recherche de stratégies pour gagner...

Exemples:

- Au poker, est-il plus fréquent d'avoir un brelan ou deux paires?
- Au bridge, si les adversaires ont quatre atouts, est-il plus probable qu'ils soient répartis en 1&3 ou en 2&2?
- A pile ou face, si face vient de sortir cinq fois de suite a-t-on plus de chances de gagner en pariant sur pile?

Dans toutes ces situations la notion de fréquence est utilisée de façon sous-jacente : on suppose (car l'expérience le montre) que si la même expérience est réalisée un grand nombre de fois, l'apparition d'un résultat particulier possède une fréquence qui lui est propre.

Par exemple, sur dix lancers d'une pièce de monnaie non-truquée on peut observer parfois neuf "face" pour un seul "pile", cependant il y aura aussi des cas où on rencontrera neuf "pile" pour un seul "face" et en moyenne, si le nombre de lancers devient très grand alors le quotient du nombre de "pile" par le nombre total de lancers est de plus en plus proche de 0,5. A la limite, si le nombre de lancers tend vers l'infini, la fréquence d'apparition de "pile" devient 0,5 et on dit que :

- "pile" arrive en moyenne une fois sur deux
- on a 50 chances sur cent d'avoir "pile"
- la probabilité de "pile" est 0,5

Pour que les choses soient claires, il faudra toujours essayer de bien distinguer l'expérience faite et la façon d'observer l'expérience... voyons quelques cas:

⇒ **Expérience 1** : On lance un dé « normal »...

- Observation : On note le nombre de points situés sur la face supérieure du dé.
- Analyse : Il n'y a que six possibilités pour l'observation et il n'y a pas de raisons pour que l'une de ces possibilités apparaisse plus fréquemment qu'une autre : on dira qu'il y a une chance sur six d'obtenir 1 (de même pour 2, 3,...) ou encore que la probabilité d'obtenir 1 est 1/6 (de même pour 2, 3,...)

⇒ **Expérience 2** : On tire un jeton d'un sac qui contient un jeton portant le n°1 et neuf jetons portant le n°2.

- Observation : On note le numéro porté par le jeton tiré.
- Analyse : Il y a dix tirages différents pour cette expérience, mais il n'y a que deux résultats possibles pour l'observation: soit on obtient 1 soit on obtient 2... Parmi les dix tirages différents qui peuvent avoir lieu un seul donne 1 alors que neuf donnent 2. Il est plus fréquent d'obtenir 2 que 1 et on dira qu'on a une chance sur dix d'obtenir 1 et neuf chances sur dix d'obtenir 2 c'est à dire:

  ∗ la probabilité de tirer 1 est 1/10

  ∗ la probabilité de tirer 2 est 9/10

  (…et, tant qu'on y est, la probabilité de tirer 3 ou 2,718 est 0)

⇒ **Expérience 3** : Un sac contient un jeton portant le n°1, deux jetons portant le n°2 et sept jetons portant le n°3. On tire simultanément de ce sac deux jetons.

- Observation : On note la somme des numéros obtenus, et on se demande si on a plus de chances d'obtenir 4 ou d'obtenir 5.
- Analyse : L'expérience fait apparaître en tout 45 tirages différents, et parmi ces tirages on dénombre ceux qui correspondent au résultat 4 puis ceux qui correspondent au résultat 5.

  On peut obtenir 4 en tirant 2&2 ou 1&3 mais on ne peut obtenir 5 qu'en tirant 2&3 ce qui pourrait laisser penser que 4 est plus probable que 5... en achevant les calculs de dénombrement on arrive aux conclusions suivantes :

  la probabilité de tirer 4 est 8/45 celle de tirer 5 est 14/45

  ... il y a donc plus de chances d'obtenir 5 que 4.

C-II. Vocabulaire tiré du concret

a. Univers, événement, éventualité

1) les résultats d'une expérience sont appelé **EVENTUALITES**
2) l'ensemble des éventualités est appelé **UNIVERS**
3) les résultats de l'observation sont appelés **EVENEMENTS**

b. Evénements particuliers

Certains événements se décomposent en plusieurs éventualités, d'autres en une seule éventualité et certains ne correspondent à aucune éventualité.

Les **EVENEMENTS ELEMENTAIRES** sont ceux qui se décomposent en une seule éventualité.

L' **EVENEMENT IMPOSSIBLE** est celui qui ne se décompose en aucune éventualité.

L' **EVENEMENT CERTAIN** est celui qui se décompose en toutes les éventualités de l'univers.

c. Probabilité intuitive des événements

L'événement impossible, n'arrive jamais, il a « 0 chance(s) » d'arriver : sa probabilité est 0.

L'événement certain arrive à tous les coups, il a « toutes les chances » d'arriver : sa probabilité est 1.

Tout événement a un certain nombre de chances d'arriver mais jamais plus que « toutes les chances » donc la probabilité d'un événement quelconque est toujours comprise entre 0 et 1.

C-III. Vocabulaire « technique » ou abstrait

Pour s'affranchir des descriptions de jeu de cartes ou de « pile ou face » on redéfinit le vocabulaire de façon plus abstraite... mais cela n'empêche pas de garder en mémoire l'origine de ce vocabulaire.

Soit E un ensemble quelconque, on appelle **éventualités** les éléments de E, **univers** l'ensemble E et **événement** toute partie de E.

(En toute rigueur, les événements sont les éléments d'une TRIBU… mais cette notion est trop matheuse pour l'IUT : ici, les événements sont les parties de l'univers)

La **partie vide** de E est l'événement impossible, la **partie pleine** de E est l'événement certain.

a. Evénements incompatibles (ou disjoints), événements complémentaires

Deux événements A et B tels que $A \cap B = \varnothing$ sont dits INCOMPATIBLES (ou DISJOINTS).

Deux événements A et B tels que $A \cap B = \varnothing \ et \ A \cup B = E$ sont dits COMPLEMENTAIRES.

Notation: Le complémentaire de A peut être noté $A^c \ ou \ \overline{A}$ ou encore d'autres façons...

Exemple : Pour le lancer d'un dé, l'univers serait {1,2,3,4,5,6} et cet univers, qui contient 6 éventualités, donnerait 64 événements parmi lesquels on trouve...

- {1,3,5} qu'on peut appeler « tirer un nombre impair »
- {1,2,3} qu'on peut appeler « tirer moins de 4 »
- {5,6} qu'on peut appeler « tirer plus de 4 » ou bien « tirer au moins 5 »
- $\varnothing$ qu'on peut appeler « tirer plus de 6 » ou bien « tirer un nombre non-entier » ou, en fait, n'importe quoi d'impossible...
- {1} et {2,3,5} seraient des événements incompatibles
- {1,3,4,5} et {2,6} seraient des événements complémentaires

Remarque : Avec cet exemple, si on pose A = {5,6}, il est assez tentant d'écrire la probabilité de {5,6} sous la forme p(A). Cette notation montre que la notion de fonction n'est pas loin. En clair, p doit être une fonction qui associe à chaque événement sa probabilité et les étapes intuitives ont montré que les résultats obtenus doivent tous être compris entre 0 et 1.

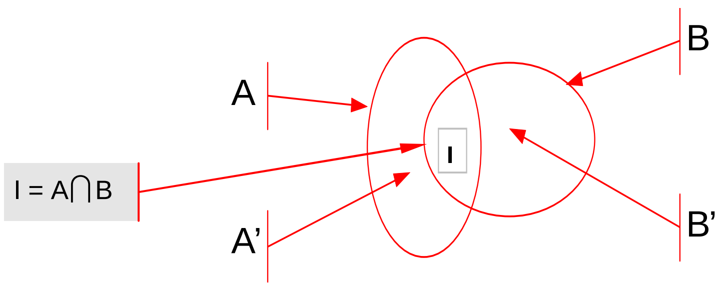
Définition : On appelle probabilité dans E (ensemble non-vide) toute application p de l'ensemble des parties de E vers [0,1] qui vérifie :

- ♦ Si A et B sont des parties de E alors
- ♦ $A \cap B = \varnothing \Rightarrow p(A \cup B) = p(A) + p(B)$
- ♦ $p(E) = 1$ ... et par conséquent $p(\varnothing) = 0$

Remarque : Le mot probabilité contient l'ambiguïté classique: il désigne aussi bien la fonction que le résultat obtenu en appliquant cette fonction : p est une fonction, p(A) est un nombre.

Propriété 1: Si A et B sont deux événements quelconques de l'univers E où la fonction-probabilité est notée p, on a :

$$\boxed{p(A \cup B) = p(A) + p(B) - p(A \cap B)}$$



En effet si on observe le découpage ci-dessus de $A \cup B$ on a :

$$A = A' \cup I \quad \text{et} \quad B = B' \cup I \quad \text{avec} \quad I = A \cap B$$

où les événements A' et I sont disjoints, de même que B' et I.

On en déduit : p(A) = p(A')+p(I) et p(B) = p(B')+p(I)

puis par addition : p(A) + p(B) = p(A') + p(B') + 2.p(I)

p(A) + p(B) - p(I) = p(A') + p(B') + p(I)

c'est à dire : p(A) + p(B) - p(A ⋂ B) = p(A ⋃ B).

Propriété 2 : A et B étant deux **événements complémentaires** dans E on a :

$$\boxed{p(A) = 1 - p(B) \quad \text{et} \quad p(B) = 1 - p(A)}$$

C'est presqu'une évidence... A et B sont tels que :

(1) $A \cap B = \varnothing$ donc p(A ⋃ B) = p(A) + p(B)

(2) $A \cup B = E$ donc p(A ⋃ B) = 1

on en déduit p(A) + p(B) = 1 ce qui donne les deux égalités espérées...

Remarques : Dans tout ce qui précède rien ne permet de calculer la probabilité d'un événement, et en particulier rien ne dit que les événements élémentaires ont la même probabilité... par exemple si on truque un dé en alourdissant une de ses faces on peut parvenir à une situation où le 1 serait beaucoup plus fréquent que le 6 c'est à dire que l'événement élémentaire {1} aurait une probabilité p({1}) plus grande que celle p({6}) de l'événement élémentaire {6}...

Le cas où les événements élémentaires ont tous la même probabilité est un cas particulier possédant de nombreuses applications, on l’appelle cas de la probabilité uniforme pour signaler que la probabilité est répartie uniformément sur les éventualités de l’univers.

b. Probabilité uniforme

Définition : On dit que la probabilité p dans l'univers E est UNIFORME lorsque les événements élémentaires de E ont tous la même probabilité. (On dit que les événements élémentaires sont équiprobables)

Remarque: Si E contient n éventualités, dire que la probabilité p dans E est uniforme c'est dire que la probabilité de chaque événement élémentaire est 1/n. Attention à bien comprendre que si l’univers contient un nombre infini d’éventualités on ne peut plus utiliser la même façon de faire : chaque événement élémentaire aurait une probabilité nulle.

Propriété: Si p est la probabilité uniforme dans un univers contenant n éventualités et si l'événement A se décompose en k éventualités alors la probabilité de A est k/n.

On retient souvent cette propriété sous la forme :

$$\boxed{p(A) = \frac{nombre\ de\ cas\ favorables\ à\ A}{nombre\ de\ cas\ au\ total}}$$

c. Evénements indépendants

Il arrive parfois que la réalisation d'un événement donne des indications sur la réalisation éventuelle d'un autre événement on dit alors que les événements sont dépendants... la dépendance peut être plus ou moins forte.

Il arrive aussi que la réalisation d'un événement ne donne aucun renseignement sur la réalisation éventuelle d'un autre événement et on dit alors que ces événements sont indépendants.

Exemple : On utilise le lancer d'un dé non-pipé (un dé normal...)

- On pose: A = "obtenir plus de 3" et B = "obtenir plus de 5"

  Il est clair que la réalisation de B entraîne celle de A : les événements A et B sont fortement liés.

- On pose: A = "obtenir 4 ou moins de 4" et B = "obtenir 5 ou plus de 5"

  Il est clair cette fois que la réalisation de l'un interdit celle de l'autre : les événements A et B, qui sont disjoints, sont encore fortement liés.

- On pose : A = "obtenir 5 ou plus" et B = "obtenir un nombre pair"

  Il faut être plus attentif...

⇒ si A est réalisé c'est qu'on a tiré 5 ou 6, et parmi ces deux cas l'un est pair, pas l'autre : la probabilité que B soit réalisé sachant que A l'est est donc 1/2

(…une chance sur deux, ou encore un cas favorable sur deux)

⇒ si on ignore l'événement A on constate que la probabilité de B est encore 1/2

(…trois chances sur six, trois résultats pairs sur six résultats en tout)

⇒ si on sait que A n'est pas réalisé c'est qu'on a tiré 1, 2, 3 ou 4 ... et la probabilité de B est encore 1/2...

Finalement, pour ce troisième cas, que A soit réalisé ou non ne donne aucun renseignement pour B : la probabilité de B reste la même. On dit que A et B sont indépendants.

d. Probabilité conditionnelle

Définition : On appelle probabilité de A conditionnée par B (ou probabilité conditionnelle de A si B est réalisé) le nombre noté p(A|B) tel que :

$$\boxed{p(A|B) = \frac{p(A \cap B)}{p(B)}}$$

Remarque : Dans le cas où la probabilité est uniforme, p(A | B) est le quotient du nombre de cas favorables à A ⋂ B par le nombre de cas favorables à B car le nombre de cas total « se simplifie ».

$$p(A|B)=\frac{Nb\ de\ cas\ \ favorables\ à\,(A\cap B)}{Nb\ de\ cas\ total}.\frac{Nb\ de\ cas\ total}{Nb\ de\ cas\ \ favorables\ à\ \ B}$$

On a dit que deux événements A et B étaient indépendants lorsque la réalisation de l'un ne donnait aucun renseignement sur l'autre c'est à dire, compte tenu de la notation de probabilité conditionnelle, lorsque p(A|B) = p(A) et cette égalité permet de comprendre la définition suivante choisie pour des événements indépendants...

Définition : On dit que A et B sont indépendants lorsque
p(A ⋂ B) = p(A).p(B).

Attention, la notion d'événements disjoints et celle d'événements indépendants sont totalement différentes !!!!

e. Notion rudimentaire de variable aléatoire

Dans de nombreux cas on a été conduit à distinguer l'expérience réalisée et l'observation de cette expérience. La notion de variable aléatoire permet de théoriser cette distinction.

Définition : On appelle variable aléatoire dans l'univers E toute application de E vers $\mathbb{R}$.

Exemples :

- On lance deux dés et on note la somme des points obtenus : à chaque éventualité on associe un nombre... cette association est une variable aléatoire dont les valeurs sont tous les entiers compris entre 2 et 12.
- On lance deux dés et on note le produit des points obtenus : encore une variable aléatoire... avec des valeurs entières comprises entre 1 et 36, mais pas toutes ! (par exemple il est impossible d'obtenir 17)
- On « attrape » au hasard un élément chimique dans le tableau de Mendeleev et on compte ses électrons... On a une variable aléatoire prenant des valeurs depuis 1 (cas de l'hydrogène) à 103 (cas du Laurencium) en passant par toutes les valeurs intermédiaires.
- On « choisit » un nombre complexe au hasard et on lui associe son module... (attention, ici l'univers est infini)

Remarque : La notion de variable aléatoire permet de modéliser le passage des éventualités (éléments de l'univers) aux événements (parties de l'univers). En demandant « quelles sont les éventualités qui correspondent à un résultat x pour la variable aléatoire X », on obtient un ensemble d'éventualités qui forment l'événement noté $(X = x)$.

Notation : Si X est une variable aléatoire dans l'univers E et si x est un réel, on note $(X = x)$ l'événement formé de toutes les éventualités ε telles que $X(\varepsilon) = x$.

$$\boxed{(X = x) = \left\{\varepsilon \in E \mid X(\varepsilon) = x\right\}}$$

Définition : La fonction qui à tout réel x associe la probabilité de l'événement $(X = x)$ est appelée **loi de probabilité** de X.

Exemple : On lance deux pièces de monnaie « non-truquées » et on note à chaque tirage le nombre de "pile". La variable aléatoire utilisée, disons X, prend les valeurs 0, 1 ou 2. On obtient la loi de probabilité de cette variable X en cherchant les trois nombres p($X = 0$), p($X = 1$) et p($X = 2$).

En effectuant les calculs on obtient p($X = 0$)=1/4, p($X = 1$)=1/2 et p($X = 2$)=1/4.

Remarque : Dans le cas où un univers est muni de la probabilité uniforme, la façon d'observer cet univers peut faire que les observations ne soient pas équiprobables. La loi de probabilité d'une variable aléatoire rend compte de cet aspect, elle indique les probabilités « au travers de l'observation » et non pas directement au niveau de l'expérience.

f. Variable aléatoire et retour aux statistiques

Après avoir mis au point les définitions, le vocabulaire lié aux probabilités, on peut revenir aux statistiques en réutilisant les techniques de calculs mises au point depuis le collège en 2nde et en première.. et rappelées dans le chapitre précédent.

Définitions: Dans un univers E muni de la probabilité p, on utilise une variable aléatoire X qui prend les valeurs x_1 , x_2 , x_3 ..., x_n et on note p_i la probabilité $p(X=x_i)$:

On appelle FONCTION DE REPARTITION de X, la fonction F de R vers [0;1] telle que :

$$\boxed{F(x) = p(X \leq x)}.$$

(à relier à la notion de fréquences cumulées vue en statistiques)

On appelle ESPERANCE de X, le nombre E(X) tel que, en posant $p_i = P(X = x_i)$

$$\boxed{E(X) = p_1x_1 + p_2x_2 + ... + p_nx_n = \sum_{i=1}^{i=n} p_ix_i}$$

(à relier à la notion de moyenne en statistiques)

On appelle VARIANCE de X, le nombre noté Var(X) tel que :

$$\boxed{\begin{aligned} Var(X) &= p_1(x_1 - E(X))^2 + p_2(x_2 - E(X))^2 + \ldots + p_n(x_n - E(X))^2 \\ &= \sum_{i=1}^{i=n} p_i(x_i - E(X))^2 \end{aligned}}$$

(à relier à la notion de variance en statistiques)

En développant les carrés et en regroupant « astucieusement » les termes, on s'aperçoit, comme en statistiques, que la variance peut s'écrire :

$$\boxed{Var(X) = \sum_{i=1}^{i=n} p_i x_i^2 - E(X)^2}$$

On appelle ECART-TYPE de X, le nombre noté σ(X) tel que :

$$\boxed{\sigma\ (X) = \sqrt{Var(X)}}$$

(à relier à la notion d'écart-type en statistiques)

Exemple de calcul : On utilise deux jetons semblables dont les faces portent les nombres 1 et 2 et on attribue à chaque tirage le nombre de deux chiffres formé en écrivant pour les unités le plus petit et pour les dizaines le plus grand des deux nombres obtenus en lançant les deux jetons. La variable aléatoire X ainsi définie prend les valeurs 11, 21, 22. La probabilité de (X=11) est 0,25, celle de X=21 est 0,5 et celle de X=22 est 0,25.

- L'espérance de X est E(X) = 0,25*11 + 0,5*21 + 0,25*22 soit E(X) = 18,75
- La variance de X est Var(X) = $0{,}25*11^2 + 0{,}5*21^2 + 0{,}25*22^2 - 18{,}75^2 = 20{,}1875$
- L'écart-type de X est (X) = $(20{,}1875)^{1/2}$ = 4,49305....

g. Notion d'épreuve (invariance des conditions expérimentales)

On appelle EPREUVE toute expérience répétable de sorte que les conditions de réalisation restent identiques, la probabilité de n'importe quel événement concernant cette expérience n'étant jamais influencée par les réalisations précédentes.

En étant rigoureux, les situations concrètes sont rarement des épreuves, mais le modèle est tellement pratique, tellement commode pour les calculs qu'il est employé dès que l'indépendance entre les réalisations n'est pas complètement illusoire.

Par exemple, l'expérience consistant en la naissance d'un enfant n'est en aucun cas une épreuve (au sens statistique) ni pour la mère ni pour l'enfant : rien n'est pareil pour eux avant et après. Le lancer d'une pièce de monnaie est une épreuve : tout est pareil pour elle avant et après. L'explosion d'une bombe atomique dans l'atmosphère n'est pas une épreuve pour les êtres vivants proches... mais en est peut-être une pour les étoiles des autres galaxies !

h. Schéma de Bernouilli

Soit une épreuve ayant deux résultats possibles, le succès et l'échec.

On note p la probabilité du succès et q celle de l'échec donc, bien sûr, p + q = 1.

On cherche la probabilité d'obtenir k succès en n réalisations de l'épreuve (k ≤ n).

Il faut choisir, parmi n réalisations, les k réalisations qui vont donner un succès : il y a C_n^k possibilités. Pour chacune de ces C_n^k possibilités il y a k succès et n-k échec donc, grâce à l'indépendance, la probabilité est $p^k.q^{n-k}$. Les différentes possibilités correspondent à des événements disjoints donc on additionne leurs probabilités... et on obtient :

$$\boxed{p(k \text{ succès en } n \text{ réalisations}) = C_n^k \, p^k . q^{n-k}}$$

i. Système complet d'événements (ou partition de l'univers)

Si E est un ensemble, on a appris ce que sont deux parties complémentaires de E : deux parties sans éléments en commun et dont la réunion forme E en entier.

Cette notion de parties complémentaires peut être "généralisée" pour donner la notion de partition ou de "Système complet d'événements" comme on dit en probabilités.

Définition : Si E est un ensemble, on dit que F est une partition de E lorsque F est une famille de parties f_i de E vérifiant :

1. pour tout i, f_i est non-vide.
2. pour tout i différent de j, $f_i \cap f_j = \varnothing$.
3. la réunion de tous les f_i est l'ensemble E au complet.

Vocabulaire : Dans le domaine des probabilités, lorsque F est une partition de l'univers E on dit que F est un système complet d'événements.

C-IV. Formule des probabilités totales

Soit E un univers muni de la probabilité p et F un système complet d'événements de E formé des f_i où i varie de 1 à n.

Pour tout événement A de E on a :

$$p(A) = p(A|f_1).p(f_1) + p(A|f_2).p(f_2) + ... + p(A|f_n).p(f_n)$$

ou encore:

$$\boxed{p(A) = \sum_{i=1}^{i=n} p(A|f_i).p(f_i) = \sum_{i=1}^{i=n} p(A \cap f_i)}$$

Remarque: Le sens de cette propriété est simple, si on a trouvé un découpage d'une population en morceaux disjoints et sans oublier personne, la probabilité d'observer un caractère particulier est la somme des probabilités d'observer ce caractère dans chaque morceau de la population avec des poids correspondants à l'importance relative de chacun des morceaux...

Par exemple, on peut partager la population humaine suivant les continents, ce partage forme bien une partition. La probabilité qu'un humain ait le SIDA est alors la somme de :

- la probabilité qu'un Africain ait le SIDA multipliée par la probabilité qu'un humain soit Africain,
- la probabilité qu'un Européen ait le SIDA multipliée par la probabilité qu'un humain soit Européen,
- idem pour l'Asie
- idem pour l'Amérique
- idem pour l'Océanie...

Démonstration (pour les courageux) avec un peu de calcul ensembliste...

Préliminaire n°1 :

Si A, I, J, K,... sont des parties d'un ensemble E , et si on pose

$$X = A \bigcap (I \cup J \cup K \ldots) \quad et \quad Y = (A \cap I) \bigcup (A \cap J) \bigcup (A \cap K) \ldots$$

alors on a X = Y

On pourrait énoncer cette propriété en disant que l'opération « intersection » est distributive sur l'opération « union ».

Pour se persuader de cette égalité (c'est à dire pour la démontrer) il suffit de décrire en détail le contenu de chacun des deux membres de l'égalité pour être sûr que tout élément du membre de gauche est dans le membre de droite et réciproquement.

- ♦ **Etape (X ⊂ Y).**

Soit x un élément de $A \bigcap (I \cup J \cup K \ldots)$, cet élément est à la fois dans A et dans la réunion des I, J, K... Puisque x est dans la réunion des I, J, K... il doit être au moins dans l'un d'entre eux :

- s'il est dans I, il est alors dans A et I donc dans A⋂I et dans la réunion de A⋂I avec "n'importe quoi" c'est à dire dans $(A \cap I) \bigcup (A \cap J) \bigcup (A \cap K) \ldots$
- s'il est dans J, il est alors dans A et J donc dans A⋂J et dans la réunion de A⋂J avec "n'importe quoi" c'est à dire dans $(A \cap I) \bigcup (A \cap J) \bigcup (A \cap K) \ldots$
- etc...

- ♦ **Etape (Y ⊂ X).**

Soit x un élément de $(A\cap I)\bigcup(A\cap J)\bigcup(A\cap K)$... cet élément est au moins dans l'un des termes de la réunion. Tous ces termes sont formés d'éléments contenus dans A, il est donc certain que x est dans A. Chaque terme de la réunion est formé d'éléments de I ou de J ou de K... donc d'éléments qui sont dans la réunion de I, J, K... et il est alors certain que x est dans $(I\cup J\cup K\ldots)$

Comme x est à la fois dans A et dans $(I\cup J\cup K\ldots)$ il est dans l'intersection des deux c'est à dire dans $A\bigcap(I\cup J\cup K\ldots)$.

Conclusion: d'après les deux étapes précédentes, tous les éléments du membre de gauche sont dans le membre de droite, et tous les éléments du membre de droite sont dans le membre de gauche, on a bien :

$$A\bigcap(I\cup J\cup K\ldots) = (A\cap I)\bigcup(A\cap J)\bigcup(A\cap K)\ldots$$

Préliminaire n°2 :

> Si I, J, K sont des parties disjointes deux à deux de E alors :
> pour toute partie A de E, les parties $(A\cap I),\ (A\cap J),\ (A\cap K)$
> sont disjointes deux à deux.

Les éléments de $(A\cap I)$ sont tous dans I donc ils ne sont ni dans J, ni dans K... et ne peuvent pas être dans $(A\cap J)\ \ ou\ \ (A\cap K)\ldots$

De même en partant d'un élément de $(A\cap J)$, il ne peut être ni dans $(A\cap I)$ ni dans $(A\cap K)$.

Application aux probabilités totales

On a trivialement p(A) = p(A⋂E) puisque A⋂E=A et on en déduit:

$$\begin{aligned} p(A) &= p(A \bigcap (f_1\cup f_2\cup\ldots\cup f_n)) \\ &= p(\ (A\cap f_1)\bigcup(A\cap f_2)\bigcup\ldots\bigcup(A\cap f_n)\ ) \\ &= p(A\cap f_1) + p(A\cap f_2) + \ldots + p(A\cap f_n) \\ &= p(A|\,f_1).p(f_1) + p(A|\,f_2).p(f_2) + \ldots + p(A|\,f_n).p(f_n) \end{aligned}$$

Exemple de calcul avec les probabilités totales :

Dans un amphithéâtre il y a 35 filles et 122 garçons. Sachant que 80% des filles et 60% des garçons ont le polycopié du cours, quelle est la probabilité qu'un(e) étudiant(e) pris au hasard ait le polycopié ?

Donnons des noms brefs aux événements :

- G pour « c'est un garçon »,
- F pour « c'est une fille »,
- Poly pour « il ou elle a le polycopié »

p(Poly) = p(Poly et G) + p(Poly et F)

= p(Poly / G).p(G) + p(Poly / F).p(F)

= 0,6 * (122/157) + (0,8 * 35/157)

= 0,64458598...

C-V. Exercices à résoudre

a. Probabilités et dénombrement

1) A la belote on partage un jeu de 32 cartes en quatre « mains » de huit cartes. Sachant que le jeu comporte huit « atouts » quelle est la probabilité que j'en ai deux ou moins de deux ? Quelle est la probabilité que j'en ai plus de deux ?

2) On joue à pile ou face, quelle est la probabilité de tirer 3 fois pile en exactement 10 lancers ?

3) A chaque fois que j'arrive sur le quai du métro j'ai une chance sur cinq d'attendre moins d'une minute et quatre chances sur cinq d'attendre plus d'une minute. Sachant que je prend le métro dix fois dans la semaine quelle est la probabilité que trois fois exactement j'attende moins d'une minute ?

4) Un cadenas à chiffres comporte quatre rouleaux permettant chacun de choisir chacun un chiffre de 0 à 5. Quelle est la probabilité que le numéro 1234 soit le bon si...

- Je ne sais rien de plus.
- Je sais que le bon numéro est pair.
- Je sais que les chiffres du bon numéro sont rangés en ordre strictement croissant.
- Je sais que les chiffres du bon numéro sont tous distincts.

b. Probabilité non uniforme et variable aléatoire

1) Bernadette se dit : « En me levant le matin, il n'y a que deux possibilités, soit le prince charmant est au pied de mon lit, soit il n'y est pas. J'ai donc une chance sur deux de rencontrer le prince charmant demain matin ! ». Expliquer pourquoi Bernadette sera déçue.

2) Deux sacs A et B contiennent chacun 3 boules noires et 1 boule blanche. On tire une boule de chaque sac et on gagne 1 € par boule blanche obtenue. On note G la variable aléatoire correspondant au gain du joueur.

- Quelles sont les valeurs que peut prendre G ?
- Déterminer la loi de probabilité de G.
- Calculer l'espérance de G ?
- Si l'organisateur du jeu veut être sûr à long terme de ne pas perdre d'argent, quel tarif doit-il faire payer la partie ?

3) Une basse cour contient 40 poules, 25 canards, 6 oies et 50 lapins.

- On prend un animal au hasard et on compte ses pattes... On note P1 la variable aléatoire « nombre de pattes ». Calculer l'espérance et l'écart type de P1.
- On prend deux animaux au hasard (après avoir remis celui de la question précédente) et on compte les pattes... On note P2 la variable aléatoire « nombre de pattes ». Calculer l'espérance et l'écart-type de P2.