Initiation aux statistiques et aux probabilités

http://alexandre-mesle.com

26 mai 2009

Table des matières

Chapitre 1

Initiation aux probabilités

Étant donné un événement que vous ne pouvez pas prévoir de façon exacte (par exemple la face sur laquelle va tomber une pièce). Le calcul des probabilités est un ensemble de méthodes permettant d'évaluer les chances que cet événement se produise.

1.1 Définitions et terminologie

1.1.1 Expérience aléatoire, événement, univers

On évalue la **probabilité** que lors d'une **expérience aléatoire**, un **événement** se produise. On dit qu'une expérience est aléatoire si on ne connaît pas de façon certaine le résultat de cette expérience à l'avance. Par exemple, soit l'expérience "Lancer d'une pièce". Les deux événements que nous considérerons sont $P =$ "la pièce tombe sur pile" et $F =$ "la pièce tombe sur face". Si à la suite de l'expérience la pièce tombe sur face, on dira que l'événement F est **réalisé** et que l'événement P n'est pas réalisé. On appelle **univers** et on note Ω l'ensemble des événements. Dans l'exemple ci-avant, $\Omega = \{P, F\}$.

1.1.2 Probabilité

Définition 1.1.1 *Une **probabilité** est une application $p : \mathcal{P}(\Omega) \longrightarrow [0, 1]$ telle que :*
- $p(\Omega) = 1$
- $p(\emptyset) = 0$
- $\forall E, F \subset \Omega, p(E) + p(F) \geq p(E \cup F)$

Si $p(e) = 1$, alors e est un **événement certain**, ce qui signifie que l'on sait avant même le début de l'expérience qu'il sera réalisé. Si $p(e) = 0$, alors e est un **événement impossible**, ce qui signifie que l'on sait avant même le début de l'expérience qu'il ne sera pas réalisé. Plus $p(e)$ est élevé, plus il est probable que l'événement e soit réalisé lors de l'expérience aléatoire. Par exemple, lors du lancer d'une pièce, il a autant de "chances" (terminologie qui sera bientôt proscrite !) qu'elle tombe sur face que sur pile. Il y a donc une chance sur deux qu'une pièce tombe sur pile et une chance sur deux qu'elle tombe sur face, ce qui s'écrit $p(F) = p(P) = \frac{1}{2}$.

1.2 Calcul des probabilités

Les formules exposées dans les sections ci-dessous sont à connaître par coeur ou pouvoir être retrouvées en moins de 2 secondes !

1.2.1 Événement complémentaire

Définition 1.2.1 *Étant donné un événement A, on note $\bar{A}$ l'événement qui se réalise si et seulement si A ne se réalise pas. On dit que $\bar{A}$ est l'**événement complémentaire** de A.*

Deux événements sont complémentaires si à chaque expérience un et un seul des deux événements est toujours réalisé. Par exemple, considérons l'expérience "lancer d'un dé" et l'événement $A =$ "Le dé ne tombe pas sur un 6". Alors $\bar{A} =$ "Le dé tombe sur un 6". A et $\bar{A}$ sont liés par la propriété suivante.

Propriété 1.2.1 *Soit A un événement quelconque,*

$$p(A) + p(\bar{A}) = 1$$

Ainsi il est aisé de calculer la probabilité d'un événement quand on connaît la probabilité de son événement complémentaire. Reprenons l'expérience aléatoire "lancer d'un dé" : comme il y a une chance sur 6 que l'événement $\bar{A}$ se réalise, alors $p(\bar{A}) = \frac{1}{6}$. Donc $p(A) = 1 - p(\bar{A}) = 1 - \frac{1}{6} = \frac{5}{6}$.

1.2.2 Opérations ensemblistes sur les événements

Étant donnés deux événements A et B, l'**union** de A et B, notée $A \cup B$ est l'événement qui est réalisé si au moins un des deux événements A et B est réalisé. L'**intersection** de A et B, notée $A \cap B$ est l'événement qui est réalisé si les deux événements A et B sont réalisés. Prenons comme exemple l'expérience "tirage d'une carte dans un jeu de 52 cartes". Soit A l'événement "la carte choisie est de couleur pique" et B l'événement "la carte choisie est une reine". Alors $A \cap B$ est réalisé si la carte choisie est une dame de pique et $A \cup B$ est réalisé si la carte choisie est une dame ou est de couleur pique. On utilise dans les calculs la relation suivante.

Propriété 1.2.2 *Soit A et B deux événements, alors*

$$p(A) + p(B) = p(A \cup B) + p(A \cap B)$$

Reprenons l'expérience "tirage d'une carte dans un jeu de 52 cartes" : nous admettrons que $p(A) = \frac{1}{4}$, $p(B) = \frac{1}{13}$ et $p(A \cap B) = \frac{1}{52}$. On a donc $p(A \cup B) = p(A) + p(B) - p(A \cap B) = \frac{1}{4} + \frac{1}{13} - \frac{1}{52} = \frac{13+4-1}{52} = \frac{4}{13}$.

Exercice 1 - Bru-en-Creuse

Dans notre charmant village de Bru-en-Creuse, 60% des habitants jouent à la pétanque, 40% aux fléchettes et 20% jouent à la fois à la pétanque et aux fléchettes. On choisit un habitant au hasard. Soit P l'événement "l'habitant choisi joue à la pétanque", F l'événement "l'habitant choisi joue aux fléchettes".

1. Traduisez avec une phrase en français les événements $\bar{P}$, $\bar{F}$, $P \cup F$, $P \cap F$, $\bar{P} \cup F$, $\bar{P} \cap \bar{F}$.
2. Calculez les probabilités des événements $\bar{P}$, $\bar{F}$, $P \cup F$ et $P \cap F$.
3. Les événements P et F sont-ils complémentaires ?

Nous noterons bien que (A et B complémentaires) $\Longrightarrow$ $(p(A)+p(B) = 1)$, mais que la réciproque est fausse, comme nous l'avons constaté dans l'exemple ci-avant.

1.2.3 Événement incompatibles

Définition 1.2.2 *Soient A et B deux événements, A et B sont* ***incompatibles*** *si $p(A \cap B) = 0$.*

Deux événements sont incompatibles s'ils ne peuvent pas être réalisés simultanément. Par exemple, soit l'expérience "lancer d'un dé", considérons les événements $A =$ "le dé tombe sur un nombre impair" et $B =$ "le dé tombe sur 6". Comme A et B ne peuvent pas être réalisés en même temps, alors A et B sont incompatibles.

Propriété 1.2.3 *Soient A et B deux événements incompatibles, alors*

$$p(A) + p(B) = p(A \cup B)$$

Reprenons notre exemple : comme A et B sont incompatibles, alors $p(A \cup B) = p(A) + p(B) = \frac{1}{2} + \frac{1}{6} = \frac{2}{3}$.

On remarque que (A et B complémentaires) $\Longrightarrow$ (A et B incompatibles), mais que une fois de plus, et comme le corrobore l'exemple ci-avant, la réciproque est fausse.

Exercice 2 - Bru-en-Creuse (suite)

Les événements P et F de l'exercice précédent sont-ils incompatibles ?

1.2.4 Événements indépendants

Définition 1.2.3 *Soient A et B deux événements, A et B sont **indépendants** si*

$$p(A \cap B) = p(A)p(B)$$

Soit l'expérience "Évaluations simultanées de la météo papouasienne et du CAC 40 le premier Janvier 2009". Soit A l'événement "le CAC 40 est strictement croissant" et B l'événement "il pleut en Papouasie". Supposons que l'on ait $p(A) = \frac{3}{5}$, $p(B) = \frac{1}{11}$ et $p(A \cap B) = \frac{3}{55}$. Est-ce que A et B sont indépendants ? Vérifions : $p(A)p(B) = \frac{3}{5} \times \frac{1}{11} = \frac{3}{55} = p(A \cap B)$, A et B sont donc indépendants. Dans la plupart des cas, ce type de résultat est prévisible : la nature des expériences et des événements permet de savoir à l'avance si deux événements A et B sont indépendants.

Propriété 1.2.4 *Étant donnés deux événements indépendants A et B,*

$$p(A) + p(B) = p(A \cup B) + p(A)p(B)$$

On peut par exemple s'en servir pour calculer $p(A \cup B)$ et $p(A \cap B)$ à partir des formules $p(A \cap B) = p(A)p(B)$ et $p(A \cup B) = p(A) + p(B) - p(A)p(B)$.

Exercice 3 - Bru-en-Creuse (suite)

Les événements P et F de l'exercice précédent sont-ils indépendants ?

Exercice 4 - Gérard et Marcel

Gérard et Marcel vont tous les soirs au bar. 3 fois sur 4, Gérard prend du vin rouge, le reste du temps il prend du vin blanc. 4 fois sur 5, Marcel prend du pastis, le reste du temps il prend du whisky. Un soir pris au hasard, nous décidons de leur rendre visite. Soient R l'événement "Gérard a pris du vin rouge", P l'événement "Marcel a pris du Pastis". Nous admettons que les deux événements P et R sont indépendants. Calculer les probabilités des événements $\bar{P}$, $\bar{R}$, $P \cap R$, $P \cup R$. *L'abus d'alcool est dangereux pour la santé, à consommer avec modération.*

Propriété 1.2.5 *Étant donnés deux événements A et B, si A et B sont indépendants, alors*
- *$\bar{A}$ et B sont indépendants*
- *A et $\bar{B}$ sont indépendants*
- *$\bar{A}$ et $\bar{B}$ sont indépendants*

Exercice 5 - Gérard et Marcel (suite)

Calculer les probabilités des $P \cap \bar{R}$, $\bar{P} \cup R$.

1.2.5 Système complet d'événements

Définition 1.2.4 *Soit $E = \{E_1, \ldots, E_n\}$ un ensemble d'événements. E est un **système complet d'événements** si*
- *$\forall i, j \in \{1, \ldots, n\}$ tels que $i \neq j$, $p(E_i \cap E_j) = 0$. (Tous les événements de E sont incompatibles entre eux)*
- *$p(E_1 \cup \ldots \cup E_n) = 1$ (Il y a toujours un événement de E qui est réalisé)*

Autrement dit, E est un système complet d'événements s'il y a toujours un et un seul événement réalisé.

Par exemple, soit A un événement quelconque, alors $\{A, \bar{A}\}$ est un système complet d'événements, en effet A et $\bar{A}$ sont incompatibles car $p(A \cap \bar{A}) = 0$ et $p(A \cup \bar{A}) = 1$. Prenons comme autre exemple le lancer d'un dé, soit D_i l'événement "le dé tombe sur la face i", alors $\{D_1, \ldots, D_6\}$ est un système complet d'événements. En effet, considérons deux événements distincts D_i et D_j, comme il est impossible qu'un dé tombe sur deux faces différentes lors du même lancer, alors $p(D_i \cap D_j) = 0$. Comme un dé tombe nécessairement sur une des faces, alors $p(D_1 \cup \ldots \cup D_6) = 1$.

Propriété 1.2.6 *(probabilités totales) Soient A un événement et $E = \{E_1, \dots, E_n\}$ un système complet d'événements, alors on a*

$$p(A) = \sum_{i=1}^{n} p(A \cap E_i)$$

Considérons par exemple les événements A et B. Comme $E = \{B, \bar{B}\}$ forme un système complet d'événements avec $E_1 = B$ et $E_2 = \bar{B}$, alors $p(A) = \sum_{i=1}^{2} p(A \cap E_i) = p(A \cap E_1) + p(A \cap E_2) = p(A \cap B) + p(A \cap \bar{B})$.

Exercice 6 - Bru-en-Creuse (suite)

Calculez les probabilités des événements $\bar{P} \cap F$, $\bar{P} \cup F$, $P \cap \bar{F}$, $P \cup \bar{F}$, $\bar{P} \cap \bar{F}$ et $\bar{P} \cup \bar{F}$.

1.2.6 Tableaux des intersections

Pour s'épargner des calculs fastidieux et réduire les risques de faire des erreurs de calcul, on représente des probabilités dans des tableaux de la forme suivante :

	B	$\bar{B}$	
A	$p(A \cap B)$	$p(A \cap \bar{B})$	$p(A)$
$\bar{A}$	$p(\bar{A} \cap B)$	$p(\bar{A} \cap \bar{B})$	$p(\bar{A})$
	$p(B)$	$p(\bar{B})$	

On déduit de la formule des probabilités le fait que la dernière colonne contient les sommes de chaque ligne et que la dernière ligne contient les sommes de chaque colonne. Si par exemple, on a $p(A) = 0.3$, $p(B) = 0.5$ et $p(A \cap B) = 0.2$ alors on remplit le tableau de la sorte :

	B	$\bar{B}$	
A	0.2	$p(A \cap \bar{B})$	0.3
$\bar{A}$	$p(\bar{A} \cap B)$	$p(\bar{A} \cap \bar{B})$	$p(\bar{A})$
	0.5	$p(\bar{B})$	

On complète ensuite le tableau en utilisant les formules sur le complément et les probabilités totales :

	B	$\bar{B}$	
A	0.2	0.1	0.3
A	0.3	0.4	0.7
	0.5	0.5	

1.2.7 Équiprobabilité

Définition 1.2.5 *Soient A et B deux événements, A et B sont* ***équiprobables*** *si $p(A) = p(B)$.*

Propriété 1.2.7 *Si tous les événements (cas) d'un système complet d'événements E sont équiprobables, on calcule la probabilité qu'un événement $A \subset E$ se produise avec la formule*

$$p(A) = \frac{|A|}{|E|} = \frac{nombre\ de\ cas\ favorables}{nombre\ de\ cas\ possibles}$$

où les cas favorables sont les événements de E telles que A se produise et les cas possibles les événements de E.

Soient par exemple l'expérience "Sélection au hasard d'un élève de BTS IG en deuxième année" dans une classe comportant 45 élèves dont 15 développeurs et A l'événement "L'élève choisi suit l'option ARLE". Si le choix de l'élève est bien aléatoire, alors tous les élèves ont la même probabilité d'être choisis. $p(A)$ est donc le quotient entre le nombre de réseaux (cas favorables) et le nombre d'élèves (cas possibles), à savoir $p(A) = \frac{45 - 15}{45} = \frac{2}{3}$.

Reprenons l'exemple du tirage de cartes et calculons de façon propre les résultats admis précédemment. Si le tirage se fait effectivement au hasard, alors chaque carte a la même chance d'être tirée, on calcule dans ce cas les probabilités avec la relation $\frac{nombre\ de\ cas\ favorables}{nombre\ de\ cas\ possibles}$. Comme une carte sur 4 est de couleur "pique", alors $p(A) = \frac{1}{4}$. Comme une carte sur 13 est une reine, alors $p(B) = \frac{1}{13}$. Comme il y a une seule dame de pique, alors $p(A \cap B) = \frac{1}{52}$.

Exercice 7 - Dans les poches de Dédé

Dédé a 2 pièces de 1 euros et 2 pièces de 2 euros dans sa poche. Il choisit successivement et sans remise deux pièces au hasard et additionne leurs valeurs.

1. Soient U_i l'événement "Dédé a choisi la i-ème pièce de 1 euro", et D_j l'événement "Dédé a choisi la j-ème pièce de 2 euros". Décrire tous les couples pouvant être formés avec ces deux événements.
2. Soient A l'événement "Dédé a choisi deux pièces de 1 euro", B l'événement "Dédé a choisi deux pièces de 2 euros", C l'événement "Dédé a choisi une pièce de 1 euro puis une pièce de 2 euros" et D l'événement "Dédé a choisi une pièce de 2 euros puis une pièce de 1 euro". Calculez les probabilités de ces 4 événements.
3. Soit P_i l'événement "La somme des pièces choisies par Dédé est i". Calculez les probabilités des événements P_2, P_3 et P_4.

Exercice 8 - Lancers successifs d'un dé

On effectue deux lancers successifs (et indépendants) d'un même dé. On note A_i l'événement "au premier lancer, le dé tombe sur la face i" $(i = 1, \ldots, 6)$ et B_i l'événement "au deuxième lancer, le dé tombe sur la face i" $(i = 1, \ldots, 6)$. On note C_i l'événement "La somme des valeurs obtenue après les deux lancers est i" $(i = 2, \ldots, 12)$. Calculer

1. $A_i, i = 1, \ldots, 6$
2. $B_i, i = 1, \ldots, 6$
3. $A_i \cap B_i, i = 1, \ldots, 6$
4. $A_i \cup B_i, i = 1, \ldots, 6$
5. C_2
6. C_3
7. C_4

Exercice 9 - Questions de compréhension

1. Prouvez que $p(\bar{A} \cup \bar{B}) = 1 - p(A \cap B)$. Puis que $p(\bar{A} \cap \bar{B}) = 1 - p(A \cup B)$.
2. Prouvez que si A et B sont incompatibles, alors $p(A) + p(B) = p(A \cup B)$.
3. Prouvez que si A et B sont indépendants, alors $p(A \cup B) = 1 - (p(A) - 1)(p(B) - 1)$.
4. Prouvez que si $p(A \cup B) = p(A) + p(B) = 1$, alors $\{A, B\}$ est un système complet d'événements.
5. Prouvez que $p(A) = p(A \cap E) + p(A \cap \bar{E})$, vous utiliserez le fait que $A = A \cap (E \cup \bar{E})$.
6. Prouvez que si A et B sont indépendants, alors leurs complémentaires sont aussi indépendants.

1.3 Probabilités conditionnelles

1.3.1 Exemple

Étant donné une classe de 45 élèves dont 15 développeurs. 13 élèves dont 5 développeurs suivent maths option. Supposons que l'on choisisse un développeur au hasard, quelle est la probabilité qu'il suivent maths option ? Si l'on choisit au hasard un élève qui ne suit pas maths option, quelle est la probabilité qu'il soit en ARLE ?

On effectue ces calculs de façon informelle en utilisant le rapport entre le nombre de cas favorables en le nombre de cas possibles. Il y a 15 développeurs dont 5 qui suivent maths option, donc 1 développeur sur 3 suit maths option. Donc la probabilité qu'un développeur choisi au hasard suive maths option est $\frac{1}{3}$. Répondons maintenant à la deuxième question, Il y a $45 - 13 = 32$ élèves ne suivant pas maths option, $15 - 5 = 10$ d'entre eux sont développeurs et donc les 22 autres sont en réseau. Il y a donc 22 cas favorables sur 32 cas possibles. La probabilité qu'un élève ne suivant pas maths option soit en réseau est donc $\frac{11}{16}$.

1.3.2 Définition

Formalisons le raisonnement précédent :

Définition 1.3.1 *On note $p(A/B)$ la probabilité que l'événement A soit réalisé* ***sachant que*** *B est réalisé. La probabilité de l'événement A* ***sachant*** *B est définie par*

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

On dit que $p(A/B)$ est une ***probabilité conditionnelle.***

La première phase à accomplir dans un exercice de probabilités est l'expression de l'énoncé avec des notations mathématiques : soit D l'événement "l'élève choisi est développeur", M l'événement "l'élève choisi suit maths option". Comme il y a 15 développeurs sur 45 étudiants, alors $p(D) = \dfrac{1}{3} = \dfrac{15}{45}$. Comme parmi les 13 élèves suivant maths option, 5 sont développeurs, alors $p(D/M) = \dfrac{5}{13}$ et $p(\bar{D}/M) = \dfrac{8}{13}$. La probabilité qu'un développeur choisi au hasard suive maths option est $p(M/D)$ et la probabilité qu'un élève choisi parmi ceux qui ne suivent pas maths option soit en réseau est $p(\bar{D}/\bar{M})$. La deuxième phase est l'utilisation des formules pour effectuer les calculs : $p(M/D) = \dfrac{p(M \cap D)}{p(D)} = \dfrac{p(D/M)p(M)}{p(D)} = \dfrac{5}{13} \times \dfrac{13}{45} \times \dfrac{3}{1} = \dfrac{1}{3}$. Comme il y a $30 - 8$ réseaux qui ne suivent pas maths option sur les 45 élèves, alors $p(\bar{D} \cap \bar{M}) = \frac{22}{45}$. Comme il y a $45 - 13$ élèves ne suivant pas maths option, alors $p(\bar{M}) = \frac{32}{45}$, donc $p(\bar{D}/\bar{M}) = \dfrac{p(\bar{D} \cap \bar{M})}{p(\bar{M})} = \dfrac{22 \times 45}{32 \times 45} = \dfrac{11}{16}$.

Exercice 10 - De retour à Bru-en-Creuse (suite et fin)

Reprenez l'exercice 1 et répondez aux questions suivantes :

1. On choisit un habitant parmi ceux qui jouent à la pétanque, quelle est la probabilité qu'il joue aussi aux fléchettes ?
2. On choisit un habitant parmi ceux qui jouent aux fléchettes, quelle est la probabilité qu'il joue aussi à la pétanque ?
3. Quelle est la probabilité qu'un habitant qui ne joue pas à la pétanque ne joue pas non plus aux fléchettes ?
4. Quelle est la probabilité qu'un habitant joue aux fléchettes sachant qu'il ne joue pas à la pétanque ?

1.3.3 Représentation sous forme d'arbre

On évite bon nombre de confusions en représentant sous forme d'arbre les probabilités conditionnelles. Si par exemple $p(A) = 0.1$, $p(B/A) = 0.25$ et $p(B/\bar{A}) = 0.4$, on utilisera l'arbre de la figure 1.1.

Cet arbre est un arbre de choix dans le sens où l'emprunt de chaque branche correspond à un événement. Les deux branches partant de la racine correspondent respectivement aux événements A et $\bar{A}$, elles sont pondérées par $p(A)$ et $p(\bar{A})$. Les deux branches partant de A sont étiquetées par B et $\bar{B}$ et correspondent aux événements B/A et $\bar{B}/A$. Elles dont pondérées par $p(B/A)$ et $p(\bar{B}/A)$. On calcule $p(A \cap B)$ en multipliant les pondérations des arêtes A et B/A, donc $p(A \cap B) = p(A) \times p(B/A) = 0.1 \times 0.25 = 0.025$. Le calcul de $p(B)$ se fait en additionnant $p(B \cap A) = 0.025$ et $p(B \cap \bar{A}) = 0.36$, donc $p(B) = 0.385$. Il est aussi possible de représenter le même arbre plaçant A aux feuilles, pour cela il est nécessaire de connaître $p(A/B) = \dfrac{p(A \cap B)}{p(B)} = \dfrac{0.025}{0.385} = 0.065$ et $p(A/\bar{B}) = \dfrac{p(A \cap \bar{B})}{p(\bar{B})} = \dfrac{p(\bar{B}/A)p(A)}{p(\bar{B})} = \dfrac{0.75 \times 0.1}{1 - 0.385} = \dfrac{0.075}{0.615} = 0.122$. On a donc :

Exercice 11 - Exercice sans sujet

Étant donnés deux événements A et B, on a $P(B) = 0.4$, $P(A/B) = 0.2$ et $P(A/\bar{B}) = 0.3$.

1. Représentez cette situation avec un arbre dont les arêtes du premier niveau sont étiquetées avec B et $\bar{B}$, et les arêtes du deuxième niveau étiquetées avec A et $\bar{A}$.
2. Calculez les probabilités des événements $P(A \cap B)$, $P(A \cap \bar{B})$, $p(A)$, $P(\bar{A} \cap B)$, $P(\bar{A} \cap \bar{B})$, $p(\bar{A})$, $p(B/A)$, $p(B/\bar{A})$
3. Représentez cette situation avec un arbre dont les arêtes du premier niveau sont étiquetées avec A et $\bar{A}$, et les arêtes du deuxième niveau étiquetées avec B et $\bar{B}$.

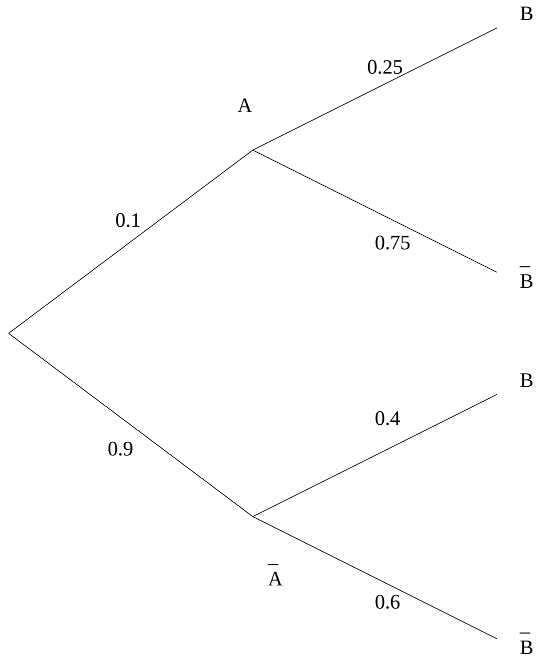
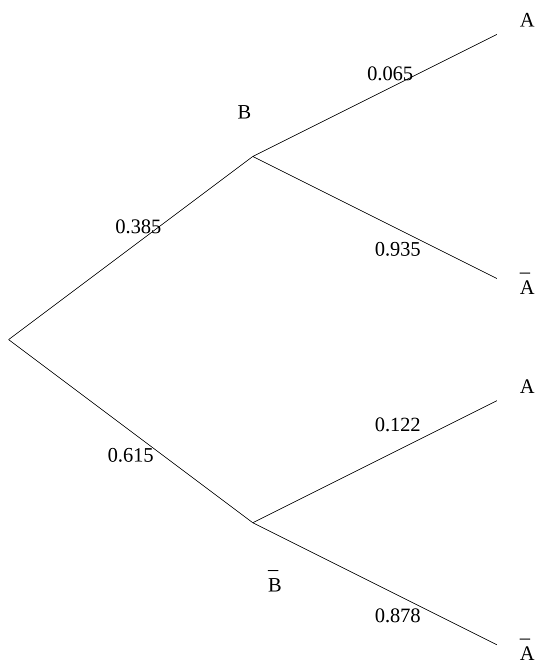


FIG. 1.1 – représentation sous forme d'arbre de probabilités conditionnelles



1.3.4 Théorème de Bayes

Dans la section précédente, nous avons exprimé $p(A/B)$ en fonction de $p(A)$, $p(B/A)$ et $p(B/\bar{A})$. La formule ci-dessous vous sera fort utile :

Théorème 1.3.1 *Étant donnés deux événements A et B,*

$$p(A/B) = \frac{p(B/A)p(A)}{p(B/A)p(A) + p(B/\bar{A})p(\bar{A})}$$

Prouvons-le,

$$\begin{aligned}
p(A/B) &= \frac{p(A \cap B)}{p(B)} \\
&= \frac{p(B/A)p(A)}{p(B \cap A) + p(B \cap \bar{A})} \\
&= \frac{p(B/A)p(A)}{p(B/A)p(A) + p(B/\bar{A})p(\bar{A})}
\end{aligned}$$

Vérifions avec l'exemple de la section précédente :

$$\begin{aligned}
p(A/B) &= \frac{p(B/A)p(A)}{p(B/A)p(A) + p(B/\bar{A})p(\bar{A})} \\
&= \frac{0.25 \times 0.1}{0.25 \times 0.1 + 0.4 \times 0.9} \\
&= \frac{0.025}{0.025 + 0.36} \\
&= \frac{0.025}{0.385} \\
&= 0.065
\end{aligned}$$

En fait, cette formule permet de calculer directement $p(A/B)$ sans passer par des étapes intermédiaires, qui sont à la fois une perte de temps et des sources d'erreurs.

Exercice 12 - Formulaire

1. Exprimer $p(A \cap B \cap C)$ en fonction de $p(A/B \cap C)$, $p(B/C)$ et $p(C)$.
2. Exprimer $p(A \cup B \cup C)$ en fonction de $p(A)$, $p(B)$, $p(C)$, $p(A \cap B)$, $p(B \cap C)$, $p(B \cap C)$ et $p(A \cap B \cap C)$.

1.3.5 Morceaux choisis

Exercice 13 - Les trois portes

A un jeu télévisé, un candidat doit choisir une porte parmi trois. Derrière une seule d'elles de trouve un lot. Une fois que le candidat a choisi une porte, le présentateur, qui lui sait derrière quelle porte se trouve le lot, ouvre une porte vide parmi les deux que le candidat n'a pas choisi et donne au candidat la possibilité de modifier son choix. Que feriez-vous à la place du candidat ?

Exercice 14 - La progéniture des Groseilles

Monsieur et Madame Groseille ont deux enfants.

1. Vous savez que l'aînée est une fille, quelle est la probabilité que le cadet soit une fille aussi ?
2. Vous savez que l'un des deux est une fille, quelle est la probabilité que l'autre soit aussi une fille ? (Ce n'est pas $\frac{1}{2}$!)

Chapitre 2

Sommations

2.1 Définitions

La notation en sigma est très utilisée pour noter les sommes, on utilise la lettre grecque

$$\sum$$

L'expression suivante

$$\sum_{i=1}^{n}$$

se lit : "somme de 1 à n de..." ou bien "somme pour i allant de 1 à n de...". A coté du sigma se trouve toujours une expression dépendant ou non de i, par exemple

$$\sum_{i=1}^{n} 2i+1$$

Cette expression a la même valeur que

$$(2(1)+1)+(2(2)+1)+\ldots+(2(i)+1)+\ldots+(2(n)+1)$$

On a additionné les différentes valeurs prises par l'expression $2i+1$, dans laquelle on a substitué à i toutes les valeurs entières comprises entre 1 et n. On dit que i est l'indice, que $2i+1$ est **sous la portée** de la somme. L'expression sous la portée du sigma peut être très variée, par exemple

- $\sum_{i=1}^{n} i = 1+2+\ldots+n$
- $\sum_{i=1}^{n+1} i = 1+2+\ldots+n+(n+1)$
- $\sum_{i=1}^{n+1} u_i = u_1+u_2+\ldots+u_n+u_{n+1}$
- $\sum_{i=0}^{n} u_{i+1} = u_{0+1}+u_{1+1}+\ldots+u_{n+1}$
- $\sum_{i=2}^{2n} iu_i = 2u_2+3u_3+\ldots+nu_n+\ldots+2nu_{2n}$

Exercice 1 - Notation en Σ

Ecrire les sommes suivantes sans Σ (vous utiliserez des ... à la place) :

1. $\displaystyle\sum_{i=1}^{n-1} i+2$
2. $\displaystyle\sum_{i=p}^{q} 2i$
3. $\displaystyle\sum_{i=1}^{n}(i-1)^2$
4. $\displaystyle\sum_{i=1}^{n}(u_n-u_{n-1})$
5. $\displaystyle\sum_{i=0}^{n-4} u_{n+4}$
6. $\displaystyle\sum_{i=n}^{2n}(2n-i)u_{2n+1}$
7. $\displaystyle\sum_{i=1}^{n}\sum_{j=i}^{n+1} j$
8. $\displaystyle\sum_{i=1}^{n} i \sum_{j=i-1}^{n-1} j^i$

Exercice 2 - Notation en Σ (suite)

Ecrire les sommes suivantes avec Σ.

1. $1+2+3+\ldots+2n$
2. $1+3+5+\ldots+101$
3. $3+7+11+\ldots+99$
4. $\dfrac{1}{2}+\dfrac{2}{3}+\dfrac{3}{4}+\ldots+\dfrac{900}{901}$
5. $1+2p+3p^2+4p^3+\ldots+100p^{99}$
6. $3p^7+5p^{10}+7p^{13}+9p^{16}+\ldots+103p^{160}$
7. $[1+2+\ldots+n]+[1+2+\ldots+(n-1)]+[1+2+\ldots+(n-2)]+\ldots+[1+2]+[1]$.
8. $[1+2+\ldots+2n]+[2+3+\ldots+(2n-1)]+\ldots+[(n-2)+(n-1)+n+(n+1)]+[(n-1)+n]$.

2.2 Propriétés

1. $\displaystyle\sum_{i=p}^{n} 1 = n-p+1$
2. $\displaystyle\sum_{i=p}^{n} u_i = \left(\sum_{i=p}^{n-1} u_i\right)+u_n = u_p+\sum_{i=p+1}^{n} u_i$
3. $\displaystyle\sum_{i=p}^{n}(a.u_i) = a\sum_{i=p}^{n} u_i$

4. $\displaystyle\sum_{i=p}^{n}(a_i+b)=\sum_{i=p}^{n}a_i+\sum_{i=p}^{n}b=\left(\sum_{i=p}^{n}a_i\right)+(n-p+1)b$
5. $\displaystyle\sum_{i=p}^{n}(a_i+b_i)=\sum_{i=p}^{n}a_i+\sum_{i=p}^{n}b_i$
6. $\displaystyle\sum_{i=p}^{n}u_i=\sum_{i=p+1}^{n+1}u_{i-1}=\sum_{i=p-1}^{n-1}u_{i+1}$
7. $\displaystyle\sum_{i=1}^{n}\sum_{j=1}^{m}u_{ij}=\sum_{j=1}^{m}\sum_{i=1}^{n}u_{ij}$

Bien que le $\sum$ soit prioritaire sur le + mais pas sur le $\times$, il usuel pour éviter toute confusion de rajouter des parenthèses, même inutiles. Parmi les sommes que l'on utilise fréquement figurent les deux suivantes :

1. $\displaystyle\sum_{i=1}^{n}i=\frac{n(n+1)}{2}$ (série arithmétique)
2. $\displaystyle\sum_{i=0}^{n}r^i=\frac{r^{n+1}-1}{r-1}$ (série géométrique)

Exercice 3 - Calculs de sommes

Calculer les sommes suivantes :

1. $\displaystyle\sum_{i=1}^{n}2i+1$
2. $\displaystyle\sum_{i=1}^{n}(i+1)^2$, on utilisera le fait que $\displaystyle\sum_{i=1}^{n}i^2=\frac{n(n+1)(2n-1)}{6}$
3. $\displaystyle\sum_{i=1}^{n}(\sum_{j=1}^{i}j)$
4. $\displaystyle\sum_{i=1}^{n}(\sum_{j=i}^{n}j)$
5. $\displaystyle\sum_{i=p}^{n}i$
6. $\displaystyle\sum_{i=p}^{n}r^i$
7. $\displaystyle\sum_{i=1}^{n}\frac{1}{i(i+1)}$, on montrera d'abord que $\displaystyle\frac{1}{i(i+1)}=\frac{1}{i}-\frac{1}{i+1}$

Exercice 4 - Sommes et puissances

1. Développer $(k+1)^3$
2. En déduire de $\displaystyle\sum_{i=1}^{n}(k+1)^3-\sum_{i=1}^{n}k^3$ une expression de $\displaystyle\sum_{i=1}^{n}k^2$ en fonction de $\displaystyle\sum_{i=1}^{n}k$ et de $\displaystyle\sum_{i=1}^{n}1$
3. Donner une expression sans Σ de $\displaystyle\sum_{i=1}^{n}(k+1)^3-\sum_{i=1}^{n}k^3$
4. Déduire des deux questions précédentes une expression sans Σ de $\displaystyle\sum_{i=1}^{n}k^2$

5. En adoptant une démarche similaire, donner une expression sans Σ de $\displaystyle\sum_{i=1}^{n} k^3$

Exercice 5 - Autres sommes

Donner des expressions les plus simples possibles des sommes suivantes :

1. $\displaystyle\sum_{i=1}^{n} ai + b$
2. $\displaystyle\sum_{i=1}^{n} (aq^{bi+c} + d)$

Exercice 6 - Série géométrique

1. Calculer la limite de la somme infinie $1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \dots$
2. Déterminer un critère permettant de déterminer en fonction de q si $\displaystyle\lim_{n \longrightarrow +\infty} \sum_{i=0}^{n} q^i$ est fini ou infini.

Chapitre 3

Raisonnements par récurrence

3.1 Principe

Ce type de raisonnement ne s'applique que dans des propriétés faisant intervenir des nombres entiers. Par exemple, “Quel que soit $n \geq 0$, $n^2 - n$ est pair”. Pour prouver qu'une propriété est vérifiée quelle que soit la valeur de n, on effectue une preuve par récurrence en procédant en deux temps :

- **l'initialisation**, On prouve que pour $n = 0$ (ou 1, ça dépend des cas), la propriété est vérifiée.
- **l'hérédité**, on prouve que si la propriété est vérifiée au rang n, alors elle l'est au rang $n + 1$.

3.2 Exemple

Montrons par récurrence sur n que “Quel que soit $n \geq 0$, $n^2 - n$ est pair” :

- **initialisation :** , $0^2 - 0$ est pair, c'est évident.
- **hérédité :** , supposons que $n^2 - n$ est pair, vérifions si $(n + 1)^2 - (n + 1)$ est pair lui aussi. Calculons

$$(n + 1)^2 - (n + 1) = n^2 + 2n + 1 - n - 1 = (n^2 - n) + 2n$$

  Comme (n^2-n) et $2n$ sont tous deux pairs, et que la somme de deux nombres pairs est paire, alors $(n+1)^2-(n+1)$ est pair.

Demandons-nous si $2^2 - 2$ est pair, on sait d'après l'initialisation que $0^2 - 0$ est pair. Posons $n = 0$, comme la propriété est vérifiée au rang n, elle est, d'après l'hérédité, nécessairement vérifiée au rang $n + 1$, donc $1^2 - 1$ est pair. Posons $n = 1$, comme $1^2 - 1$ est pair, la propriété est vérifiée au rang n, elle est, d'après l'hérédité, nécessairement vérifiée au rang $n + 1$, donc $2^2 - 2$ est pair. On peut généraliser ce raisonnement à n'importe quelle valeur de n.

Exercice 1

Prouver par récurrence les propriétés suivantes :

1. Quel que soit $n \geq 0$, $n^2 + n + 1$ est impaire.
2. Soit (u) une suite définie par $u_0 = -2$, $u_{n+1} = (1/2)u_n + 3$, alors $u_n < 6$
3. Soit (u) une suite définie par $u_0 = -2$, $u_{n+1} = (1/2)u_n + 3$, alors $u_n = 6 - \frac{8}{(2^n)}$
4. Soit (u) une suite définie par $u_0 = 2$, $u_{n+1} = 2u_n - n$, alors $u_n = 2^n + n + 1$
5. Pour tout $n > 0$, $\displaystyle\sum_{i=1}^{n} i = \frac{n(n + 1)}{2}$
6. Pour tout $n \geq 0$, $q \neq 1$, $\displaystyle\sum_{i=0}^{n} q^i = \frac{1 - q^{n+1}}{1 - q}$
7. Pour tout $n > 0$, $\displaystyle\sum_{i=1}^{n} i^2 = \frac{n(n + 1)(2n + 1)}{6}$

Exercice 2 - Dérivée de fonctions puissance n

On rappelle que $(fg)' = f'g + fg'$ et que f^n est définie par $(f^n)(x) = [f(x)]^n$. Prouvez par récurrence que $\forall n \in \mathbb{N}^*$

$$(f^n)' = nf'(f^{n-1})$$

Exercice 3 - Formule de Leibniz

On note $f^{(n)}$ la dérivée n-ième de f. Prouvez par récurrence la formule de Leibniz : $\forall n \in \mathbb{N}^*$

$$(fg)^n = \sum_{i=0}^{n} \mathcal{C}_n^i f^{(i)} g^{(n-i)}$$

Exercice 4 - Dérivée de fonctions composées

On rappelle que $f \circ g$ est défini par $(f \circ g)(x) = f(g(x))$ et que $(f \circ g)' = g'(f' \circ g)$. Etant donné un ensemble $\{f_1, \ldots, f_n\}$ de n fonctions, on note

$$\bigcirc_{i=1}^{n}[f_i] = f_1 \circ \ldots \circ f_n$$

la composition de ces fonctions. Démontrez par récurrence que

$$(\bigcirc_{i=1}^{n}[f_i])' = \prod_{i=1}^{n}[f_i' \circ (\bigcirc_{j=i+1}^{n}[f_j])]$$

Chapitre 4

Analyse Combinatoire

4.1 Rappels de théorie des ensembles

4.1.1 Définition

Définition 4.1.1 *Un **ensemble** est un regroupement en un tout d'objets. Etant donné un ensemble E, les objets regroupés dans E sont appelés **éléments** de E.*

On définit un ensemble notament en énumérant ses éléments et en les séparant par des virgules. Par exemple $E = \{1, 2, 3, 4\}$ définit un ensemble E contenant les **éléments** 1, 2, 3 et 4. Si e est un élément de E, on dit alors qu'il **appartient** à E, noté $e \in E$.

Définition à l'aide d'un prédicat

On définit aussi un ensemble par une **condition d'appartenance**. Par exemple, $E = \{e|(e \in N) \wedge (e \leq 5)\}$ définit l'ensemble des éléments e vérifiant $(e \in N) \wedge (e \leq 5)$, à savoir $\{0, 1, 2, 3, 4, 5\}$. Plus formellement,

Définition 4.1.2 *Soient $P(x)$ un prédicat et $E = \{x|P(x)\}$, alors*

$$x \in E \iff P(x)$$

$E = \{x|P(x)\}$ est alors l'ensemble E des éléments x vérifiant $P(x)$.

Ensemble vide

F est le prédicat constant prenant toujours la valeur "faux".

Définition 4.1.3 *On note $\emptyset$ l'**ensemble vide** défini comme suit*

$$\emptyset = \{i|F\}$$

$\emptyset$ est l'ensemble auquel n'appartient aucun élément. L'ensemble vide peut être défini par tout prédicat constant étant toujours faux. Par exemple,

$$\emptyset = \{i|(i \in N) \wedge (i = i + 1)\}$$

4.1.2 Opérations ensemblistes

On définit sur les ensembles les opérations suivantes :

Définition 4.1.4 ***L'union** de deux ensembles A et B, notée $\cup$, est définie par $A \cup B = \{x|(x \in A) \vee (x \in B)\}$.*

Un élément appartient à $A \cup B$ s'il appartient à A, **ou** s'il appartient à B. Par exemple,

$$\{1,3,7,8\} \cup \{2,4,6,8,10\} = \{1,2,3,4,6,7,8,10\}$$

Définition 4.1.5 ***L'intersection*** *de deux ensembles A et B, notée $\cap$, est définie par $A \cap B = \{x | (x \in A) \wedge (x \in B)\}$.*

Un élément appartient à $A \cap B$ s'il appartient à A, **et** s'il appartient à B. Par exemple,

$$\{1,3,4,6,7,9,10\} \cap \{2,4,6,8,10,12\} = \{4,6,10\}$$

Définition 4.1.6 ***La différence*** *de deux ensembles A et B, notée $\setminus$ ou $-$, est définie par $A \setminus B = \{x | (x \in A) \wedge (x \notin B)\}$.*

Un élément appartient à $A \setminus B$ (ou $A - B$) s'il appartient à A, **mais** pas à B. Par exemple,

$$\{1,2,3,4,6,7,8\} \setminus \{1,2,3\} = \{4,6,7,8\}$$

On note l'union de n ensembles $E_1 \cup E_2 \cup \ldots \cup E_n$ de la façon suivante :

$$\bigcup_{i=1}^{n} E_i$$

De même

$$E_1 \cap E_2 \cap \ldots \cap E_n = \bigcap_{i=1}^{n} E_i$$

Exercice 5

Etant donnés les ensembles $A = \{1,2,4,6,8,10\}$ et $B = \{1,2,3,4,5\}$. Calculer $A \cup B$, $A \cap B$, $A \setminus B$, $B \setminus A$.

4.1.3 Cardinal

Un ensemble est infini s'il contient un nombre infini d'éléments, sinon, on dit qu'il est fini.

Définition 4.1.7 *Le **cardinal** d'un ensemble fini E, noté $|E|$ est le nombre d'éléments qu'il contient.*

Par exemple, $|\emptyset| = 0$, $|\{1,2,4\}| = 3$, $|\{\{1,2\},\{4,6\}\}| = 2$, $|\{\emptyset\}| = 1$. Seul l'ensemble vide est de cardinal 0, un ensemble de cardinal 1 est appelé un **singleton**, un ensemble de cardinal 2 est appelé une **paire**.

Propriété 4.1.1 *Le cardinal d'un ensemble a les propriétés suivantes :*

- $|A \cap B| \leq |A|$ *et* $|A \cap B| \leq |B|$
- $|A \cup B| \geq |A|$ *et* $|A \cup B| \geq |B|$
- $A \setminus B \leq |A|$
- $|A \cup B| + |A \cap B| = |A| + |B|$, *d'où* $|A \cup B| \leq |A| + |B|$
- *Si* $A \cap B = \emptyset$, $|A \cup B| = |A| + |B|$

4.1.4 n-uplets

Un **n-uplet** est un regroupement ordonné de n éléments non nécessairement distincts. On note $X = (x_1, \ldots, x_n)$ le n-uplet composé des éléments $x_1, \ldots, x_n$, on les appelle **composantes** des E. Par exemple, $(4,8)$ est un 2-uplet, dit aussi **couple**, et $((3,2),(6,12),(0,3))$ est un **triplet** composé des trois couples $(3,2)$, $(6,12)$ et $(0,3)$.

Produit cartésien

Définition 4.1.8 *Le* ***produit cartésien de*** *n ensembles $E_1, \dots, E_n$, noté $E_1 \times \dots \times E_n$ est l'ensemble de tous les n-uplets $(x_1, \dots, x_n)$ tels que $\forall i \in \{1, \dots, n\}, x_i \in E_i$.*

Par exemple,

$$\{1, 2\} \times \{a, b, c\} = \{(1, a), (2, a), (1, b), (2, b), (1, c), (2, c)\}$$

Formellement,

$$E_1 \times \dots \times E_n = \{(x_1, \dots, x_n) | \forall i \in \{1, \dots, n\}, x_i \in E_i\}$$

Exercice 6

Enumérez les éléments de $\{1, 2, 3\} \times \{4, 5, 6\}$. Combien y en a-t-il ?

Exercice 7

Même question avec $\{1, 2\} \times \{1, 2\} \times \{1, 2\}$.

4.1.5 Parties

Inclusion

Définition 4.1.9 *A est* ***contenu*** *dans B, noté $A \subset B$ si et seulement si $\forall e \in A, e \in B$,*

En d'autres termes si tout élément de A est aussi un élément de B. On dit aussi que A est un **sous-ensemble** de B, que A est **inclus** dans B, ou encore que A est une **partie** de B.

Parties de E

Définition 4.1.10 *L'ensemble $\mathcal{P}(E)$ des parties de E, est défini par*

$$\mathcal{P}(E) = \{e | e \subset E\}$$

$\mathcal{P}(E)$ est donc l'ensemble de tous les ensembles inclus dans E. Par exemple,

$$\mathcal{P}(\{1, 2, 3\}) = \{\{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \emptyset, \{1, 2, 3\}\}$$

On remarque que $\mathcal{P}(E)$ est un ensemble d'ensembles. Prennez note du fait que $\emptyset$ est un sous-ensemble de tous les ensembles, y compris de lui-même, et que le seul ensemble contenu dans $\emptyset$ est $\emptyset$. A votre avis, que vaut $\mathcal{P}(\emptyset)$? Cette question vous amènera à distinuer $\emptyset$ de $\{\emptyset\}$, qui est l'ensemble qui contient l'ensemble vide.

Exercice 8

Calculer $\mathcal{P}(\emptyset)$, $\mathcal{P}(\{1, 2\})$, $\mathcal{P}(\{\emptyset\})$, $\mathcal{P}(\mathcal{P}(\mathcal{P}(\emptyset)))$.

Propriété 4.1.2 *Soit E un ensemble à n éléments, on a $|\mathcal{P}(E)| = 2^n$*

4.1.6 Applications

Définition 4.1.11 *Une* ***relation*** *f entre deux ensembles A et B est un sous-ensemble de $A \times B$.*

Etant donné deux éléments $x \in A$ et $y \in B$, si $(x, y) \in f$, alors on dit que f associe y à x. L'ensemble des éléments qui sont associés à x est $\{y | (x, y) \in f\}$.

Définition 4.1.12 *Une relation f entre deux ensembles A et B est une* ***application*** *si $\forall a \in A, |\{b | (a, b) \in f\}| = 1$*

Plus explicitement, f est une application si à tout élément de A est associé exactement un élément de B. On dit alors que f est une application **de A dans B**, ce qui se note $f : A \longrightarrow B$. On dit par abus de langage que A est l'ensemble de départ et B l'ensemble d'arrivée.

Exercice 9

Soient $A = \{1, 2, 3, 4, 5\}$, $B = \{1, 2, 3\}$ deux ensembles et $f = \{(3, 2), (4, 1), (1, 1), (5, 2), (2, 3)\}$ une relation entre A et B. Représenter graphiquement f, est-ce une application ?

Pour tout $a \in A$, on note $f(a)$ l'élément de B qui lui est associé. On dit que $f(a)$ est **l'image** de a par f et a **un antécédent** de $f(a)$ par f.

Exercice 10

Soient $A = \{0, 1, 2, 3\}$ et $f : A \longrightarrow A$ telle que $f(x) = x \ mod\ 4$. Représenter f graphiquement et calculer $f(1)$, $f(f(1))$, $f(f(f(1)))$.

Propriété 4.1.3 *Soient A et B deux ensembles de cardinaux respectifs m et n,*
- *Il existe 2^{mn} relations entre A et B.*
- *Il existe n^m applications de A dans B.*

4.2 Arrangements

4.2.1 Factorielles

Pour tout $n \geq 0$, le nombre $n!$, appelé "factorielle n", est défini par
- $0! = 1$
- $n! = n \times (n-1)!$

4.2.2 Applications injectives

Soit $f : A \longrightarrow B$,

Définition 4.2.1 *f est injective si*

$$\forall a \in A, \forall a' \in A, f(a) = f(a') \Rightarrow a = a'$$

f est injective si deux éléments distincts ne peuvent pas avoir la même image.

4.2.3 Nombre d'applications injectives

Propriété 4.2.1 *Etant donnés deux ensembles A et B de cardinaux respectifs m et n, si $m \leq n$, il existe $n(n-1)\ldots(n-m+1)$ applications injectives de A dans B.*

4.2.4 Arrangements

Définition 4.2.2 *On appelle* ***arrangements*** *de m éléments dans n le nombre $\mathcal{A}_n^m = \dfrac{n!}{(n-m)!}$.*

Il existe $\mathcal{A}_n^m$ m-uplets d'éléments distincts choisis dans un ensemble de n éléments.

Exercice 11

Combien existe-t-il de façons de ranger 3 boules de billard numérotées dans 5 boîtes numérotées.

1. En s'autorisant à placer plusieurs boules dasn une même boîte.
2. En plaçant au maximum une boule par boîte.

Calculer dans les deux cas la probabilité que la première boîte ne soit pas vide.

Exercice 12

Quelle est la probabilité que dans un mot de 5 lettres, il y ait une lettre qui apparaisse plusieurs fois ?

Exercice 13

Calculer la probabilité que dans votre classe, toutes les dates d'anniversaire ne soient pas distinctes.

4.3 Permutations

Soit $f : A \longrightarrow B$,

Définition 4.3.1 *f est surjective si*

$$\forall b \in B, \exists a \in A, f(a) = b$$

f est surjective si tout élément de l'ensemble d'arrivée a un antécedent.

Définition 4.3.2 *f est bijective si elle est à la fois injective et surjective.*

f est bijective si tous les éléments de A et de B sont reliés deux à deux.

Propriété 4.3.1 *Soient A et B deux ensembles et $f : A \longrightarrow B$,*
- *Si f est injective, alors $|A| \leq |B|$*
- *Si f est surjective, alors $|A| \geq |B|$*
- *Si f est bijective, alors $|A| = |B|$*

Cette propriété est très importante ! En dénombrement, lorsque le cardinal d'un ensemble est difficile à déterminer, on passe par une bijection vers un autre ensemble dont il est davantage aisé de calculer le cardinal.

Définition 4.3.3 *Une **permutation** est une application bijective d'un ensemble dans lui-même.*

Propriété 4.3.2 *Le nombre de permutations d'un ensemble A à n éléments est $n!$.*

Exercice 14

Une boîte contient 7 jetons numérotés. On les sort de la boîte un par un en notant l'ordre de sortie.

1. De combien de façons différentes peut-on les sortir ?
2. Dans combien de cas le jeton numéro 1 sort-il en premier ?
3. Quelle est la probabilité que le jeton numéro 2 sorte en premier et le numéro 4 en dernier ?

Exercice 15

Le tiercé (resp. quarté, quinté) plus est un jeu dans lequel le parieur doit choisir dans l'ordre trois (resp. quatre, cinq) chevaux parmi 18. La parieur gagne si les chevaux sur lesquels porte son pari arrivent dans l'ordre qu'il a prévu. En supposant que tous les ordres d'arrivés soient équiprobables :

1. Quelle est la probabilité de gagner le tiercé plus ?
2. Quelle est la probabilité de gagner le quarté plus ?
3. Quelle est la probabilité de gagner le quinté plus ?

4.4 Combinaisons

Définition 4.4.1 *On appelle **combinaisons** de p éléments dans n et on note $\mathcal{C}_n^p$ le nombre de sous-ensembles à p éléments d'un ensemble à n éléments.*

Propriété 4.4.1 *Si $0 \leq p \leq n$, alors $\mathcal{C}_n^p = \dfrac{n!}{p!(n-p)!}$.*

On remarque que $\mathcal{C}_n^p = \dfrac{\mathcal{A}_n^p}{p!}$.

Exercice 16 - valeurs remarquables

Calculer les valeurs suivantes, vous vérifierez qu'elles sont correctes en utilisant les définitions ensemblistes :
- C_n^0
- C_n^1
- C_n^n
- C_n^2

4.5 Triangle de Pascal

Propriété 4.5.1 *Si* $1 \leq p < n$, *alors* $\mathcal{C}_n^p = \mathcal{C}_{n-1}^{p-1} + \mathcal{C}_{n-1}^p$

Cette propriété permet de calculer les valeurs des $\mathcal{C}_n^p$ en les disposant dans un tableau appelé **Triangle de Pascal**.
Le triangle de Pascal est un tableau triangulaire de taille infinie dont les lignes sont indicées par n et les colonnes par p, les indices commençant à 0. Les seules case du triangle renseignées sont celles dont les indices vérifient $0 \leq p \leq n$. Les valeurs se trouvant à la ligne d'indice n et la colonne d'indice p du triangle de Pascal est $\mathcal{C}_n^p$, on a ainsi

(n,p)		$p=$	0	1	2	3	4	5	6	...
$n=$	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		
	6		1	6	15	20	15	6	1	
	...		...	...	...	...	...	...	...	...

Vous remarquez que la propriété $\mathcal{C}_n^0 = 1$ traduit le fait que la première colonne ne comporte que des 1. Comme de même $\mathcal{C}_n^n = 1$, tous les éléments de la diagonale sont des 1. La propriété $\mathcal{C}_n^p = \mathcal{C}_{n-1}^{p-1} + \mathcal{C}_{n-1}^p$ traduit le fait que chaque élément $\mathcal{C}_n^p$ ne se trouvant pas dans la première colonne ou sur la diagonale est la somme de celui qui est au dessus ($\mathcal{C}_{n-1}^p$) et de celui qui est juste à gauche de ce dernier ($\mathcal{C}_{n-1}^{p-1}$).

4.5.1 Propriétés

Symétrie

La propriété suivante traduit le fait que chaque ligne est symétrique

Propriété 4.5.2 $\mathcal{C}_n^p = \mathcal{C}_n^{n-p}$

Somme des lignes

Si on somme les éléments sur chaque ligne, on obtient des puissances successives de 2 :

(n,p)		$p=$	0	1	2	3	4	5	6	...	$\sum_{p=0}^{n} \mathcal{C}_n^p$
$n=$	0		1								2^0
	1		1	1							2^1
	2		1	2	1						2^2
	3		1	3	3	1					2^3
	4		1	4	6	4	1				2^4
	5		1	5	10	10	5	1			2^5
	6		1	6	15	20	15	6	1		2^6
	...		...	...	...	...	...	...	...	...	...

On a donc

Propriété 4.5.3 $\sum_{i=0}^{n} \mathcal{C}_n^i = 2^n$.

Binôme de Newton

L'identité remarquable $(a+b)^n$ s'écrit à l'aide des coefficients su triangle de Pascal. Par exemple,

$$(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$$

Vous remarquez que le développement de $(a+b)^3$ s'ecrit avec un polynome à deux variables dont la somme des exposants de chaque terme est 3, chaque terme est donc de la forme a^ib^{3-i}. Les coefficients devant chaque terme sont issus de la ligne d'indice 3 du triangle de Pascal, à savoir $(1,3,3,1)$, autrement dit $(\mathcal{C}_3^0, \mathcal{C}_3^1, \mathcal{C}_3^2, \mathcal{C}_3^3)$. Donc

$$(a+b)^3 = \sum_{i=0}^{3} \mathcal{C}_3^i a^i b^{n-i}$$

On généralise cette formule avec la propriété suivante, appelée **Binôme de Newton**.

Propriété 4.5.4 $(a+b)^n = \displaystyle\sum_{p=0}^{n} \mathcal{C}_n^p a^p b^{n-p}$

Par exemple,

$$(a+b)^4 = a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4$$

On remarque par ailleurs que la formule $\displaystyle\sum_{i=0}^{n} \mathcal{C}_n^i = 2^n$ correspond au cas particulier du binôme de Newton où $a = b = 1$.

Exercice 17 - Damier

Etant donné un damier de quatre cases par quatre sur lequel sont disposés quatre jetons indiscernables.

1. Quelle est la probabilité que tous les jetons soient sur des cases blanches ?
2. Quelle est la probabilité que tous les jetons soient alignés (horizontalement, verticalement ou en diagonale) ?
3. Quelle est la probabilité que tous les jetons soient disposés sur les cases formant les deux diagonales ?
4. Quelle est la probabilité que deux jetons soient sur des cases blanches et deux sur des cases noires ?

Exercice 18 - Urne

Etant donné une urne contenant trois jetons blancs et deux jetons noirs. On extrait simultanément deux jetons.

1. Quelle est la probabilité d'extraire deux jetons blancs ?
2. Quelle est la probabilité d'extraire deux jetons noirs ?
3. Quelle est la probabilité d'extraire un jeton noir et un jeton blanc ?

Exercice 19 - Cartes

On sélectionne trois cartes simultanément dans un jeu de 32.

1. Quelle est la probablité que les trois cartes soient des as ?
2. Quelle est la probablité que les trois cartes soient de couleur rouge ?
3. Quelle est la probablité que deux cartes soient de couleur rouge et une de couleur noire ?
4. Quelle est la probablité qu'une des cartes soit un as et que les deux autres soient des têtes (valet, dame ou roi) ?

Chapitre 5

Variables aléatoires

5.1 Définitions

Définition 5.1.1 *Une* ***variable aléatoire*** *est une variable dont la valeur est inconnue avant une expérience aléatoire.*

Considérons par exemple l'expérience "Lancer d'un dé", soit X la variable aléatoire "valeur prise par le dé". La valeur de X est inconnue avant l'expérience et connue après.

Définition 5.1.2 *On note $X = k$ l'événement "X a pris la valeur k au cours de l'expérience".*

Par exemple, $X = 2$ est l'événement "Le dé est tombé sur un 2". On a donc $p(X = 2) = \frac{1}{6}$.

Définition 5.1.3 *Une* ***loi de probabilité*** *est une application qui à chaque valeur k pouvant être prise par la variable aléatoire X associe la probabilité $p(X = k)$*

Reprenons notre exemple, et faisons un tableau :

k	1	2	3	4	5	6
$p(X = k)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Ce tableau associe à chaque valeur pouvant être prise par X une probabilité. La quatrième colonne ($k = 3$, $p(X = 3) = \frac{1}{6}$) signifie que la probabilité que X prenne la valeur 3 est $\frac{1}{6}$. Il va de soi que la somme des valeurs de trouvant sur la deuxième ligne vaut 1.

Exercice 1 - Pile ou face

Vous jouez à pile ou face avec une pièce truquée qui tombe 2 fois sur 3 sur face. A chaque tour, lorsque cette pièce tombe sur face, vous gagnez 2 euros, sinon vous perdez 5 euros. Soit X la variable aléatoire "Somme gagnée lors d'un tour". Donnez la loi de probabilité de X.

Exercice 2 - Loterie

Vous participez à une loterie dans laquelle 100 billets sont vendus à 1 euro. Trois billets sont tirés au sort, les trois lots sont de valeur respective 40, 20 et 10 euros. Soit X vos bénéfices à la suite du tirage. Donnez la loi de probabilité de X.

Exercice 3 - Flechettes

Gégé manque la cible 1 fois sur 2 lorsqu'il joue aux flechettes. Il tente trois lancers. Soit X la variable aléatoire : "nombre de fléchettes ayant atteint la cible".

1. Donnez la loi de probabilité de X.
2. Marcel parie 20 euros avec Bernard qu'il n'atteindra pas la cible 3 fois et Bernard parie 15 euros avec Ginette qu'il manquera la cible les trois fois. Soit Y la variable aléatoire "Gains de Bernard", donnez sa loi de probabilité.

Exercice 4 - Dans les poches de Dédé

Dédé a 2 pièces de 1 euro et 3 pièces de 2 euros dans sa poche. Il choisit deux pièces au hasard. Soit X la variable aléatoire "Somme des valeurs des deux pièces". Donnez la loi de probabilité de X.

5.2 Espérance mathématique

Nous considérerons dans cette partie une variable aléatoire arbitraire X pouvant prendre ses valeurs dans l'ensemble $\{x_1, \ldots, x_n\}$.

Définition 5.2.1 *Soit X une variable aléatoire, alors l'**expérance mathématique** de X, notée $E(X)$, est définie comme suit :*

$$E(X) = \sum_{i=1}^{n} p(X = x_i)x_i$$

Dans l'exemple, on a $x_i = i$ pour tout $i \in \{1, \ldots, 6\}$, donc $E(X) = \sum_{i=1}^{n} p(X = x_i)x_i = \sum_{i=1}^{6} \frac{1}{6}x_i = \frac{1}{6}\sum_{i=1}^{6} x_i = \frac{1}{6}(1+2+3+4+5+6) = \frac{21}{6} = \frac{7}{2}$.

Exercice 5 - Pile ou face

Calculez l'espérance mathématique de X.

Exercice 6 - Loterie

Donnez l'espérance mathématique de X. Est-il intéressant pour vous de participer à cette loterie ?

Exercice 7 - Flechettes

Donnez l'espérance mathématique de X et de Y.

5.3 Variance, écart-type

Définition 5.3.1 *Soit X une variable aléatoire, alors la **variance** de X, notée $V(X)$, est définie comme suit :*

$$V(X) = \sum_{i=1}^{n} p(X = x_i)(x_i - E(X))^2$$

Dans l'exemple, on a

$$\begin{aligned}
V(X) &= \sum_{i=1}^{n} p(X = x_i)(x_i - E(X))^2 \\
&= \sum_{i=1}^{6} \frac{1}{6}\left(i - \frac{7}{2}\right)^2 \\
&= \frac{1}{6}\sum_{i=1}^{6}\left(i^2 - 2i\frac{7}{2} + \frac{49}{4}\right) \\
&= \frac{1}{6}\left(\sum_{i=1}^{6} i^2 - 7\sum_{i=1}^{6} i\right) + \frac{49}{4} \\
&= \frac{1}{6}((1+4+9+16+25+36) - 7 \times 21) + \frac{49}{4}
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{6}(91-147)+\frac{49}{4} \\
&= \frac{-56}{6}+\frac{49}{4} \\
&= \frac{-56\times 22+49\times 23}{12} \\
&= \frac{35}{12}
\end{aligned}
$$

Définition 5.3.2 *Soit X une variable aléatoire, alors l'**écart-type** de X, noté $\sigma(X)$, est défini comme suit :*

$$\sigma(X) = \sqrt{V(X)}$$

Propriété 5.3.1 *Soit X une variable aléatoire alors*

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

Pour simplifier la démonstration, nous noterons p_i à la place de $p(X = x_i)$. Par définition, on a

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

D'une part on a pour tout i,

$$(x_i - E(X))^2 = x_i^2 - 2x_iE(X) + E(X)^2$$

Donc,

$$V(X) = \sum_{i=1}^{n} p_i(x_i^2 - 2x_iE(X) + E(X)^2)$$

De plus,

$$p_i(x_i^2 - 2x_iE(X) + E(X)^2) = p_ix_i^2 - p_i2E(X)x_i + p_iE(X)^2$$

Donc,

$$V(X) = \sum_{i=1}^{n}(p_ix_i^2 - p_i2E(X)x_i + p_iE(X)^2)$$

Décomposons la somme,

$$V(X) = \sum_{i=1}^{n} p_ix_i^2 - \sum_{i=1}^{n} p_i2E(X)x_i + \sum_{i=1}^{n} p_iE(X)^2$$

Mettons $2E(X)$ en facteur dans $\sum_{i=1}^{n} p_i2E(X)x_i$,

$$V(X) = \sum_{i=1}^{n} p_ix_i^2 - 2E(X)\sum_{i=1}^{n} p_ix_i + \sum_{i=1}^{n} p_iE(X)^2$$

Comme $\sum_{i=1}^{n} p_ix_i^2 = E(X^2)$ et $\sum_{i=1}^{n} p_ix_i = E(X)$, alors

$$V(X) = E(X^2) - 2E(X)^2 + E(X)^2\sum_{i=1}^{n} p_i$$

Comme $\sum_{i=1}^{n} p_i = 1$, alors

$$V(X) = E(X^2) - 2E(X)^2 + E(X)^2$$

Ce qui se simplifie

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

Vérifions ce résultat sur l'exemple, $V(X) = E(X^2) - E(X)^2$ avec $E(X)^2 = (\frac{7}{2})^2 = \frac{49}{4}$ et $E(X^2) = \dfrac{1}{6}\sum_{i=1}^{6} i^2 = \dfrac{1}{6}(1+4+9+16+25+36) = \dfrac{91}{6}$, donc $E(X^2) - E(X)^2 = \dfrac{91}{6} - \dfrac{49}{4} = \dfrac{35}{12}$.

Exercice 8 - Pile ou face

Calculez la variance de X.

Exercice 9 - Loterie

Donnez la variance de X.

Exercice 10 - Flechettes

Donnez les variances de X et de Y.

5.4 Opérations entre variables aléatoires

Soit X une variable aléatoire, posons $Y = aX + b$. Y est aussi une variable aléatoire, Y peut prendre les valeurs $\{ax_1 + b, \dots, ax_n + b\}$ et on a $p(Y = ax_i + b) = p(X = x_i)$. On a la propriété suivante :

Propriété 5.4.1 *Soit X et Y deux variables aléatoires telles que $Y = aX + b$, alors*

$$E(Y) = aE(X) + b$$

et

$$V(Y) = a^2 V(X)$$

En effet (et en notant $p_i = p(X = x_i)$),

$$\begin{aligned}
E(Y) &= \sum_{i=1}^{n} p_i(ax_i + b) \\
&= \sum_{i=1}^{n} (p_i a x_i + p_i b) \\
&= \sum_{i=1}^{n} p_i a x_i + b \sum_{i=1}^{n} p_i \\
&= a \sum_{i=1}^{n} p_i x_i + b \times 1 = aE(X) + b
\end{aligned}$$

Par ailleurs,

$$\begin{aligned}
V(Y) &= \sum_{i=1}^{n} p_i(ax_i + b - E(Y))^2 \\
&= \sum_{i=1}^{n} p_i(ax_i + b - aE(X) - b)^2
\end{aligned}$$

$$
\begin{aligned}
&= a^2 \sum_{i=1}^{n} p_i(x_i - E(X))^2 \\
&= a^2 V(X)
\end{aligned}
$$

Chapitre 6

Lois discrètes usuelles

6.1 Loi binomiale

6.1.1 Définition

Etant donné une suite de n expériences indépendantes 2 à 2 et dont chacune peut se solder par un succès avec une probabilité p ou un echec avec une probababilité $1-p$. Soit X la variable aléatoire "nombre d'expériences réussies", la variable X suit une loi binomiale de paramètres (n,p), ce qui se note X suit $\mathcal{B}(n,p)$. Alors, pour tout $k \in \{0,\ldots,n\}$, on a

$$p(X=k) = \mathcal{C}_n^k p^k (1-p)^{n-k}$$

Où $p(X=k)$ est la probabilité que k expériences soient réussies.

6.1.2 Exemple

Gégé lance 10 fléchettes sur une cible et atteint la cible 2 fois sur 3, soit $X =$ "nombre de lancers réussis".

- Pour chaque lancer il y a deux issues possibles : succès si la flèche atteint la cible et échec sinon.
- La probabilité qu'un lancer soit réussi ne dépend pas des lancers précédents, les lancers sont donc indépendants 2 à 2.

Les deux conditions sont vérifiées, donc X suit $\mathcal{B}(10, \frac{2}{3})$. La loi de probabilité de X est donc donnée par

$$p(X=k) = \mathcal{C}_{10}^k \left(\frac{2}{3}\right)^k \left(\frac{1}{3}\right)^{10-k}$$

La probabilité que Gégé réussisse exactement deux lancers est donc

$$p(X=2) = \mathcal{C}_{10}^2 \left(\frac{2}{3}\right)^2 \left(\frac{1}{3}\right)^{10-2}$$

On a $\mathcal{C}_{10}^2 \frac{10!}{2!8!} = \frac{10.9}{2} = 45$, par ailleurs $\frac{2}{3}^2 = \frac{4}{9}$ et $\frac{1}{3}^8 = \frac{1}{6561}$. D'où $p(X=2) = 45.\frac{4}{9}.\frac{1}{6561} = \frac{20}{6561}$.

6.1.3 Espérance mathématique et variance

Si X suit $\mathcal{B}(n,p)$ alors $E(X) = np$ et $V(X) = np(1-p)$. En effet ,

$$\begin{aligned} E(X) &= \sum_{i=0}^{n} i p(X=i) \\ &= \sum_{i=0}^{n} i \mathcal{C}_n^i p^i (1-p)^{n-i} \end{aligned}$$

$$\begin{aligned} &= n\sum_{i=0}^{n} \mathcal{C}_{n-1}^{i-1} p^i (1-p)^{n-i} \\ &= \dots \\ &= np \end{aligned}$$

Les étapes intermédiaires vous sont laissées en exercice, de même pour la relation

$$\begin{aligned} V(X) &= \sum_{i=0}^{n} (np-i)^2 p(X=i) \\ &= \dots \\ &= np(1-p) \end{aligned}$$

Exercice 1 - Calculs

Soit X une variable aléatoire suivant une loi binomiale de paramètres $n = 5$ et $p = 0.3$, calculer les probabilités suivantes :

1. $p(X = 2)$
2. $p(X = 1)$
3. $p(X \leq 1)$
4. $p(X \neq 4)$
5. $p(X > 1)$

Exercice 2 - Un jeu particulièrement intelligent

Gégé et Bernard joue fréquement à un jeu de cartes dans lequel le perdant est obligé de boire 20 centilitres de vokda à chaque tour. Ils jouent 5 tours et on a constaté par le passé que 2 fois sur 5, c'est Gégé qui gagnait.

1. Soit X la variable aléatoire "nombre de parties gagnées par Bernard". Quelle est la probabilité que Bernard gagne les cinq tours ?
2. Quelle est la probabilité que Bernard gagne exactement quatre tours ?
3. Quelle est la probabilité que Bernard gagne au plus trois tours ?
4. Donner l'espérance mathématique de X et sa variance.
5. Soit Y la variable aléatoire "Quantité d'alcool absorbée par Gégé". Donner la loi de probabilité de X, son espérance et sa variance.

Exercice 3 - Les départementales de fléchettes

Pour se qualifier aux épreuves départementales de fléchettes, Marcel doit effectuer trois séries de quatre lancers. Une série de lancers est réussie s'il ne manque pas la cible plus d'une fois. Marcel sera qualifié s'il réussit au moins deux séries sur les trois. Sachant que Marcel atteint la cible quatre fois sur cinq, quelle est la probabilité qu'il se qualifie ?

Exercice 4 - Un autre jeu très intelligent

Raymond, en raison de sa médiocre prestation au tournoi de pétanque de la veille, est soumis à un gage. Il doit tirer avec remise 3 boules dans une urne en contenant 10 dont 7 vertes, les autres sont bleues. S'il tire au moins deux boules vertes, il paie la tournée générale, qui s'élève 30 euros, sinon chacun paie son verre (y compris Raymond), chaque verre coûtant 6 euros. Raymond est soumis à ce rituel sadique 4 fois dans la soirée. Soit X la variable aléatoire "somme dépensée par Raymond" pendant la soirée. Donnez la loi de probabilité de X, son espérance mathématique et sa variance.

6.2 Loi de Poisson

6.2.1 Exemple

Etant donné un bar de province dans lequel viennent très peu de clients. Dans une heure, il y a en moyenne 4 arrivées de clients indépendantes deux à deux et il est extrêmement peu probable que deux clients arrivent exactement en même temps. Quelle est la probabilité qu'il y ait au moins 3 arrivées de clients entre 15 heures et 16 heures ?

6.2.2 Définitions

Soit X une varible aléatoire suivant une loi de Poisson de paramètre λ, ce que l'on note X suit $\mathcal{P}(\lambda)$, alors

$$p(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

A chaque fois que vous aurez à utiliser la loi de Poisson, cela sera précisé dans l'énoncé. Résolvons l'exemple : soit X la variable aléatoire "nombre de clients arrivés entre 15 heures et 16 heures". Nous admettrons que X suit $\mathcal{P}(4)$. Alors $p(X = k) = \frac{4^k e^{-4}}{k!}$. Calculons

$$\begin{aligned}
p(X \geq 3) &= 1 - p(X < 3) \\
&= 1 - [p(X = 0) + p(X = 1) + p(X = 2)] \\
&= 1 - \left[\frac{4^0 e^{-4}}{0!} + \frac{4^1 e^{-4}}{1!} + \frac{4^2 e^{-4}}{2!}\right] \\
&= 1 - e^{-4}\left[\frac{1}{1} + \frac{4}{1} + \frac{16}{2}\right] \\
&= 1 - e^{-4}[1 + 4 + 8] \\
&= 1 - 13e^{-4} \\
&= 0.76
\end{aligned}$$

à 10^{-2} près.

6.2.3 Espérance mathématique et variance

Si X suit $\mathcal{P}(\lambda)$ alors $E(X) = V(X) = \lambda$. On obtient ces résultats en calculant

$$\begin{aligned}
E(X) &= \lim_{n \longrightarrow +\infty} \sum_{i=0}^{n} i \frac{\lambda^i e^{-\lambda}}{i!} \\
&= \ldots \\
&= \lambda
\end{aligned}$$

Et

$$\begin{aligned}
V(X) &= \lim_{n \longrightarrow +\infty} \sum_{i=0}^{n} (i - \lambda)^2 \frac{\lambda^i e^{-\lambda}}{i!} \\
&= \ldots \\
&= \lambda
\end{aligned}$$

La compréhension des étapes intermédiaires nécessite quelque bagage mathématique (notamment sur les séries), libre à vous de vous y intéresser.

Exercice 5 - Calculs

Soit X une variable aléatoire suivant une loi de poisson de paramètre 0.2, calculer les probabilités suivantes :

1. $p(X = 0)$
2. $p(X = 1)$
3. $p(X \leq 2)$
4. $p(X > 1)$
5. $p(2 \geq X < 6)$

Exercice 6 - Passage de voitures

Raymond et Dédé regardent passer les voitures sur la place centrale du village. Ils constatent qu'il passe en moyenne 4 voitures par heure.

1. Soit X la variable aléatoire nombre de voitures passant aujourd'hui entre 16 heures et 17 heures. Donnez la loi de probabilité de X.
2. Raymond a parié avec Dédé qu'il passera strictement plus de 5 voitures, s'il gagne, Dédé lui paira un verre d'une valeur de 8 euros, s'il perd il paiera un verre d'un montant de 5 euros à Dédé. Soit Y la variable aléatoire "Gains (éventuellement de négatifs) de Raymond". Donnez la loi de probabilité de Y.
3. Raymond a-t-il été intelligent en faisant ce pari ?

Chapitre 7

Loi continues

7.1 Variables aléatoires continues

7.1.1 Fonction de répartition

Soit X une variable aléatoire quelconque, la **fonction de répartition** F associe à tout x la probabilité que X prenne une valeur inférieure à x. C'est à dire

$$F(x) = P(X \leq x)$$

Exercice 1 - Fonction de répartition

Etant donnée une variable aléatoire discrète X dont la loi de probabilité est le suivante : $p(X = 1) = 0.4$, $p(X = 2) = 0.3$ et $p(X = 3) = 0.3$. Donnez tous les valeurs pouvant être prises par la fonction de répartition de X. Représentez-là graphiquement.

7.1.2 Densité de probabilité

X est une **variable aléaoire continue** si elle prend ses valeurs dans un ensemble continu, dans ce cours, nous utiliserons $\mathbb{R}$. La dérivée de la fonction de répartition de X, notée f, s'appelle la **densité de probabilité**. On a donc

$$F(x) = \int_{-\infty}^{x} f(t)dt$$

f est une densité de probabilité si elle vérifie les conditions suivantes :
- f est positive ou nulle sur $\mathbb{R}$
- f est continue
- $\displaystyle\int_{-\infty}^{+\infty} f(t)dt = 1$

On se sert pour calculer $P(a < X < b)$, où a et b sont deux constantes, de l'intégrale

$$\begin{aligned} P(a < X < b) &= \int_{a}^{b} f(t)dt \\ &= F(b) - F(a) \end{aligned}$$

Exercice 2 - Lois continues

Considérons la fonction f définie par morceaux :
- si $x < 0$, $f(x) = 0$
- si $0 \leq x < \frac{1}{2}$, $f(x) = x$
- si $\frac{1}{2} \leq x < \frac{3}{2}$, $f(x) = \frac{1}{2}$

– si $\frac{3}{2} \leq x < 2$, $f(x) = 2 - x$
– si $x \geq 0$, $f(x) = 0$

Est-ce que f est une densité de probabilité ? Représenter graphiquement f. Soit X la variable aléatoire de densité de probabilité f. calculer $p(\frac{1}{2} \leq x < \frac{3}{2})$.

7.1.3 Espérance mathématique et variance

Il est, en règle général, plus commode d'utiliser la fonction de répartition pour effectuer des calculs de probabilités avec des variables continues. Comme toute variable aléatoire, X à une **espérance mathématique**

$$E(X) = \int_{-\infty}^{+\infty} tf(t)dt$$

et une **variance**

$$\begin{aligned} V(X) &= \int_{-\infty}^{+\infty} (t - E(X))^2 f(t)dt \\ &= E(X^2) - E(X)^2 \\ &= \int_{-\infty}^{+\infty} t^2 f(t)dt - (\int_{-\infty}^{+\infty} tf(t)dt)^2 \end{aligned}$$

Exercice 3 - Lois continues

Donnez l'espérance mathématique et la variance de la variable aléatoire X de la question précédente.

7.1.4 La loi normale

La loi continue la plus célèbre, et la plus courante pour des raisons énoncées dans le chapitre ci-après, est la **loi normale**, dite aussi loi de **Laplace-Gauss** (ces deux noms empêchent les apprentis mathématiciens de dormir la nuit depuis le 19-ème siècle). On dit que la variable aléatoire continue X suit une loi normale de paramètres m et σ, ce qui est noté X suit $\mathcal{N}(m, \sigma)$ si la densité de probabilité de X est

$$f(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(\frac{t-m}{\sigma})^2}$$

m est le mode, la moyenne et la médiane de X, σ est l'écart-type de X. Voici quelques exemples de courbes de Gauss :

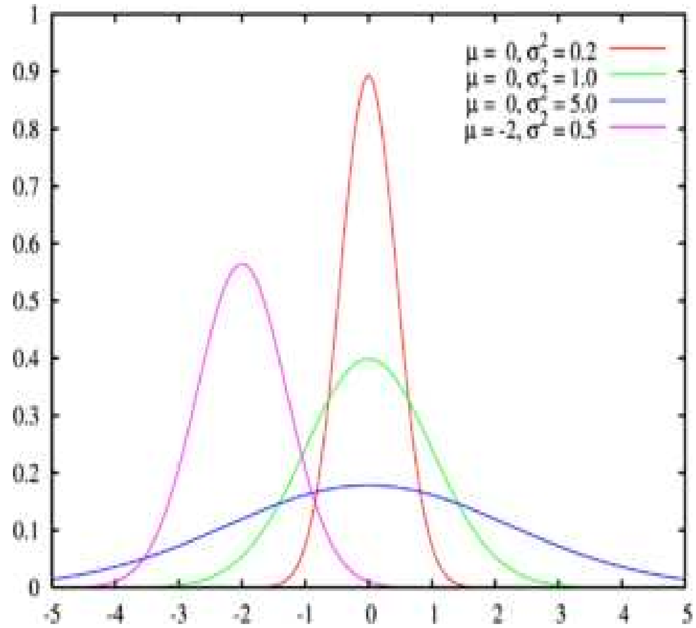
Vous observez que la courbe est symétrique par rapport à l'axe d'équation $x = m$, par ailleurs, si σ est élevé, la courbe décroit plus lentement quand on s'éloigne de m, si σ est bas, la courbe décroit plus vite quand on s'éloigne de m.

7.1.5 La loi normale centrée réduite

Observons que si X suit $\mathcal{N}(m, \sigma)$, alors

$$\frac{X - m}{\sigma} \ suit \ \mathcal{N}(0, 1)$$

dite aussi, loi normale centrée (de moyenne 0) réduite (d'écart-ype 1). La fonction de répartition de la loi normale centrée réduite est difficile à calculer manuellement. Il est d'usage d'utiliser soit des logiciels, soit des tableaux. Les tableaux donnent les valeurs de la fonction de répartition $F(x)$ de la loi normale centrée réduite pour x strictement positif.



Exercice 4 - Lecture dans la table

Etant donnée une variable aléatoire X suivant une loi normale centrée réduite. Lire les valeurs suivantes dans une table :

- $p(X \leq 1)$
- $p(X \leq 3)$
- $p(X \leq 0.5)$
- $p(X \leq 1.64)$
- $p(X \leq 3.39)$

Soit X une variable aléatoire suivant $\mathcal{N}(m, \sigma)$, alors la variable aléatoire $\dfrac{X-m}{\sigma}$ suit une loi normale centrée réduite. On ramène donc le calcul de $p(X < a)$ à celui de $p\left(\dfrac{X-m}{\sigma} < \dfrac{a-m}{\sigma}\right)$.

Exercice 5 - Changement de variable

Etant donnée une variable aléatoire X suivant une loi normale de moyenne 4 et d'écart-type 2. Calculer les probabilités suivantes en vous ramenant à une loi normale centrée réduite :

- $p(X \leq 5)$
- $p(X \leq 7)$
- $p(X \leq 4.5)$

On se ramène à une lecture de valeurs du tableau avec les propriétés suivantes :

1. $F(0) = 0.5$ (la courbe est symétrique)
2. $p(X < a) = F(a)$ (par définition)
3. $p(X > a) = 1 - p(X < a) = 1 - F(a)$ (complémentarité)
4. $p(X < -a) = p(X > a)$ (symétrie de la courbe)
5. Si $a < 0$, alors $p(-a < X < a) = 2F(a) - 1$.

Exercice 6 - Transformation de l'écriture

Etant donnée une variable aléatoire X suivant une loi normale centrée réduite. Calculer les probabilités suivantes en vous ramenant à des lectures dans une table :

- $p(X \geq 1)$

- $p(X \leq -1)$
- $p(1 \leq X \leq 2)$
- $p(-1 \leq X \leq 1)$

Exercice 7 - Débrouillez-vous

Etant donnée une variable aléatoire X suivant une loi normale de moyenne 8 et d'écart-type 4. Calculer les probabilités suivantes :
- $p(X \geq 9)$
- $p(X \leq 2)$
- $p(5 \leq X \leq 10)$
- $p(4 \leq X \leq 7)$

7.2 Echantillonnage

Cette section traite de la façon dont on peut se servir de la loi d'une population pour déterminer la loi de probabilité d'un échantillon.

7.2.1 La loi faible des grands nombres

Par exemple, étudions la variable qualitative P (résultat du lancer d'une pièce) prenant la valeur 0 si la pièce tombre sur face, 1 si la pièce tombe sur pile. Il va de soi que P suit $\mathcal{B}(\frac{1}{2}, 1)$. Nous décidons de lancer n fois la pièce, nous considérerons P_i le résultat du i-ème lancer (1 pour pile, 0 pour face). Posons $S_n = \sum_{i=1}^{n} P_i$, S_n suit $\mathcal{B}(\frac{1}{2}, n)$, d'où $E(S_n) = \frac{1}{2}n$, $V(S_n) = \frac{1}{4}n$. Posons Y_n la proportion de pile observé ans l'échantillon, à savoir $Y_n = \frac{1}{n}S_n$. On a

$$\begin{aligned}
E(Y_n) &= E(\frac{1}{n}S_n) \\
&= \frac{1}{n}E(S_n) \\
&= \frac{1}{n}\frac{1}{2}n \\
&= \frac{1}{2} \\
V(Y_n) &= V(\frac{1}{n}S_n) \\
&= \frac{1}{n^2}V(S_n) \\
&= \frac{1}{n^2}\frac{1}{4}n \\
&= \frac{1}{4n}
\end{aligned}$$

Vous remarquez la proportion théorique $E(Y_n) = 0.5$, et surtout le fait que plus n prendra des valeurs élevées, plus la variance diminuera. La loi faible des grands nombres peut s'énoncer de la sorte, Y_n **tend vers la proportion théorique quand n tend vers l'infini**. Cela signifie aussi qu'en prenant des valeurs de n suffisament grandes, Y_n a *presque* la valeur de la proportion théorique.

7.2.2 Le théorème de la limite centrée

Soit X une variable aléatoire quelconque, on effectue n expérience pour observer la valeur de X. Soit X_i la variable aléatoire : "valeur prise par X à la i-ème expérience". Nous nous intéressons à la loi de probabilité de

$$Y_i = \frac{1}{n} \sum_{i=1}^{n} X_i$$

Le **théorème de la limite centrée** s'énonce de la façon suivante : Y_n **converge vers une loi normale**. C'est à dire, si n est assez grand (en pratique : $n \geq 30$), Y_n suit *presque* une loi normale.

Chapitre 8

Statistiques

Statistique vient du latin *status*, qui signifie *État*. La statistique était, dans l'antiquité la tenue à jour et l'analyse de données, principalement économiques. Maintenant le sens de ce mot s'est élargi, on pourrait plus généralement le définir par *l'Art de faire parler les chiffres*. C'est-à-dire comment tirer des conclusions à partir de chiffres. Les statistiques trouvent des applications dans des domaines très variés : médecine, biologie, marketing, économie, productique, sondage, etc.

8.1 Définition et terminologie

8.1.1 Population, individu, échantillon

La **population** est l'ensemble que l'on étudie. Cette population est composée d'**individus**. Par exemple, si l'on décide d'étudier les revenus des étudiants, la population est l'ensemble des étudiants et chaque étudiant est un individu. Si l'on étudie le poids des plaques d'égout sortant d'une usine, la population es l'ensemble des plaques d'égout fabriquées dans cette usine et chaque plaque est un individu.

Lorsqu'il est impossible pour raisons pratiques d'observer la totalité des individus, on restreint l'étude à un sous-ensemble de la population. Un tel sous-ensemble s'appelle un **échantillon**. Par exemple, les sondages servant à estimer les votes des français sont effectués sur des échantillons. Si un échantillon est prélevé au hasard dans la population, on dit alors qu'il est **représentatif** de cette population. Par exemple, on n'aurait pas idée de baser des prévisions aux prochaines élections sur un sondage effectué en plein milieu d'une manifestation de la CGT ou d'un congrès du MEDEF...

8.1.2 Caractère quantitatif, qualitatif

Sur chacun des individus sondés, on observe un **caractère** (ou **variable**). Par exemple :
- âge
- revenus
- métier
- nombre d'enfants
- pression artérielle
- durée de bon fonctionnement,
- fumeur
- titulaire du permis B

Ce caractère est **quantitatif** s'il est possible de le mesurer, donc de le représenter avec un nombre :
- âge
- revenus
- nombre d'enfants
- pression artérielle
- durée de bon fonctionnement

Il est **qualitatif** dans le cas contraire :
- métier
- fumeur
- titulaire du permis B

Une valeur prise par une variable s'appelle une **modalité**.

Exercice 1 - Lancers de fléchettes

On a observé des lancers de fléchettes dans un bar et relevé pour chaque joueur le nombre de fléchettes ayant atteint la cible. Le caractère observé est-il quantitatif ou qualitatif ?

Exercice 2 - Capacités d'absorption

On a observé les clients dans un bar et relevé pour chacun d'entre eux le nombre le nombre de verres de whisky qu'il aura consommé. Le caractère observé est-il quantitatif ou qualitatif ?

8.1.3 Discret, continu

Dans le cas où les modalités d'un caractère quantitatif sont des valeurs réelles, on dit que ce caractère est **continu**. Par exemple,
- revenus
- âge
- pression artérielle
- durée de bon fonctionnement

Si au contraire les modalités sont des valeurs isolées, par exemple des valeurs entières, alors le caractère est **discret**. Par exemple,
- âge
- nombre d'enfants

La distinction est quelquefois floue et dépend à la fois du nombre d'individus observés et de la précision avec laquelle est mesurée le caractère. Par exemple, le nombre d'enfants est nécessairement entier, il est en effet impossible d'avoir $1,5$ enfants dans un foyer. L'âge d'un individu, si on le mesure en années, est un caractère discret car il est dans ce cas impossible qu'un individu ait un âge fractionnaire. Si par contre on mesure cet âge en minutes, alors on observe un tel nombre de valeurs qu'il est davantage approprié de considérer que l'on a une variable continue.

Exercice 3 - Lancers de fléchettes

Le caractère observé est-il discret ou continu ?

Exercice 4 - Capacités d'absorption

Le caractère observé est-il discret ou continu ?

8.2 Série statistiques à une variable

Une série statistique à une variable est un ensemble de couples $\{(x_1, n_1), \ldots, (x_p, n_p)\}$ modalités/effectifs. Les composantes $\{x_1, \ldots, x_p\}$ sont les modalités (pas nécessairement numériques) prises par le caractère observé. Les composantes $\{n_1, \ldots, n_p\}$ sont les effectifs. Pour chaque i, n_i est le nombre d'individus sur lesquels on a observé la modalité x_i. On pose $N = \displaystyle\sum_{i=1}^{p} n_i$ le nombre d'individus observés. Lorsque tous les n_i sont à 1, on les mentionne pas et on a $n = p$.

8.2.1 Exemples

- Observons le nombre d'enfants par foyer dans un immeuble arbitraire, le recueil des données est résumé par la série statistique suivante : $\{(0,4),(1,2),(2,2),(3,1)\}$. On a donc $p = 4$ modalités $\{0,1,2,3\}$ et p effectifs $\{4,2,2,1\}$. Cela signifie qu'on a observé 4 (n_1) foyers avec 0 (x_1) enfants, 2 (n_2) foyers avec 1 (x_2) enfants, 2 (n_3) foyers avec 2 (x_3) enfants et 1 (n_4) foyers avec 3 (x_4) enfants. Il y a en tout $n = n_1+n_2+n_3+n_4 = 4+2+2+1 = 9$ foyers (= individus).
- Observons la première boisson consommée dans la journée par les élèves d'une classe de BTS IG : on a comme modalités {eau, café, thé, chocolat, vodka, huile de vidange} et comme effectifs $\{5,14,3,2,11,0\}$. 5 élèves prennent de l'eau, 14 du café... Il y a $n = 5+14+3+2+11+0 = 35$ individus.

Exercice 5 - Lancers de fléchettes

Les données observées forment la série statistique suivante : $\{(0,1),(1,3),(2,5),(3,2),(4,1)\}$. Combien d'individus ont réussi exactement 3 lancers ?

Exercice 6 - Capacités d'absorption

Les données observées forment la série statistique suivante : $\{(0,2),(1,4),(2,10),(3,3),(4,4),(5,1)\}$. Combien d'individus ont bu 3 verres ou plus ?

8.2.2 Fréquence

Étant donnée une série statistique, la **fréquence** f_i de la modalité n_i est définie par $\frac{n_i}{n}$. f_i est la proportion d'individus sur lesquels on observe la modalité x_i. On remarque que la somme des fréquences doit être égale à 1. Dans l'exemple sur le nombre d'enfants, on a $f_1 = \frac{n_1}{n} = \frac{4}{9}$, donc dans les quatre neuvièmes des foyers qui ont été observés, il n'y avait pas d'enfants.

8.3 Tableaux

8.3.1 Caractère quantitatif discret

Lorsque l'on étudie un caractère quantitatif discret, il est usuel d'utiliser un tableau mettant en correspondance les modalités de la variable avec les effectifs (ou les fréquences). Représentons les nombres d'enfants par foyer de l'exemple ci-avant :

i	x_i	n_i	f_i
1	0	4	4/9
2	1	2	2/9
3	2	2	2/9
4	3	1	1/9

Exercice 7 - Lancers de fléchettes

Placez ces données dans un tableau dans lequel vous préciserez les fréquences.

Exercice 8 - Capacités d'absorption

Placez ces données dans un tableau dans lequel vous préciserez les fréquences.

8.4 Paramètres de position

Étant donné un caractère quantitatif, on détermine sa **position** en observant si les modalités prennent des valeurs élevées ou basses. Nous exposons ci-après trois critères : la moyenne, la médiane et le mode.

8.4.1 Moyenne

La moyenne $\bar{x}$ est donnée par la formule suivante :

$$\frac{1}{n}\sum_{i=1}^{p} n_i x_i$$

Ou bien

$$\sum_{i=1}^{p} f_i x_i$$

Exercice 9 - Lancers de fléchettes

Calculez la moyenne de cette série statistique.

Exercice 10 - Capacités d'absorption

Calculez la moyenne de cette série statistique.

8.4.2 Médiane

La médiane M_e est la valeur qui sépare les individus en deux groupes de même effectif (même s'il faut fractionner un individu), de sorte que la moitié des individus prennent des modalités inférieures à M_e et l'autre moitié des modalités supérieures. Selon le type de variable et l'effectif, il peut y avoir plusieurs façons de procéder :

1. Si le caractère est quantitatif discret et qu'il y a un nombre impair d'individus, alors on choisit la modalité de l'individu qui partage l'effectif en deux. Par exemple, la médiane de la série $\{1, 2, 7, 10, 14\}$ (où chaque modalité est d'effectif 1) est 7.
2. Si le caractère est quantitatif discret et qu'il y a un nombre pair d'individus, alors on utilise la moyenne des modalités des individus qui partagent l'effectif en deux. Dans l'exemple $\{1, 2, 7, 9, 10, 14\}$, on sélectionne les modalités des deux individus se trouvant au centre de la série, à savoir 7 et 9, et on utilise la moyenne de ces deux valeurs comme médiane : $M_e = (7 + 9)/2 = 8$.
3. Si le caractère est quantitatif continu, on détermine la classe dans laquelle se trouve la valeur séparant l'effectif en deux et
   - soit on prend le centre de cette classe,
   - soit on effectue une interpolation linéaire (documentez-vous pour les détails)

Exercice 11 - Lancers de fléchettes

Calculez la médiane de cette série statistique.

Exercice 12 - Capacités d'absorption

Calculez la médiane de cette série statistique.

8.4.3 Mode

Le **mode** est la classe (où la modalité) dont l'effectif est le plus élevé.

Exercice 13 - Lancers de fléchettes

Calculez le mode de cette série statistique.

Exercice 14 - Capacités d'absorption

Calculez le mode de cette série statistique.

8.5 Paramètres de dispersion

Étant donné un caractère quantitatif, on détermine sa **dispersion** en observant si les modalités prennent des valeurs regroupées ou éparpillées. Nous exposons ci-après deux critères : l'étendue et l'écart-type.

8.5.1 Étendue

L'étendue est un indicateur particulièrement médiocre du fait de sa sensibilité aux valeurs aberrantes. Elle est donnée par

$$\max\{x_i | i \in \{1, \ldots, n\} - \min\{x_j | j \in \{1, \ldots, n\}$$

Autrement dit, l'étendue est égale à la distance entre la plus petite valeur et la plus grande. Seules deux valeurs sont considérées dans le calcul, et ces valeurs sont souvent peu représentatives de population. Nous proscrirons donc cet indicateur.

Exercice 15 - Lancers de fléchettes

Calculez l'étendue de cette série statistique.

Exercice 16 - Capacités d'absorption

Calculez l'étendue de cette série statistique.

8.5.2 Variance

La variance $V(x)$ est un indicateur donné par la formule suivante :

$$V(X) = \frac{1}{n}\sum_{i=1}^{n} n_i(x_i - \bar{x})^2 = \sum_{i=1}^{n} f_i(x_i - \bar{x})^2$$

Une autre formule permet ce calcul :

$$V(X) = \frac{1}{n}\sum_{i=1}^{n} n_i x_i^2 - \bar{x}^2 = \sum_{i=1}^{n} f_i x_i^2 - \bar{x}^2$$

On encore

$$V(X) = \frac{\sum_{i=1}^{n} n_i x_i^2 - \left(\sum_{i=1}^{n} n_i x_i\right)^2}{n}$$

Exercice 17 - Lancers de fléchettes

Calculez la variance de cette série statistique.

Exercice 18 - Capacités d'absorption

Calculez la variance de cette série statistique.

8.5.3 Ecart-type

L'écart-type de la série statistique X, noté $\sigma(X)$, est la racine carrée de sa variance. On a donc $\sigma(X) = \sqrt{V(X)}$

Exercice 19 - Lancers de fléchettes

Calculez l'écart-type de cette série statistique.

Exercice 20 - Capacités d'absorption

Calculez l'écart-type de cette série statistique.

8.6 Séries statistiques à deux variables

Une série statistique à deux variables se forme en étudiant simultanément deux caractères sur la population. Par exemple, considérons l'expérience consistant à choisir 10 chats à mesurer leurs tailles et leurs poids, nous obtenons une série statistique à deux variables car les deux caractères observés sont la taille et le poids.

Nous noterons x_i la modalité du premier caractère prise par l'individu i et y_i la modalité du deuxième caractère prise par cet individu. Alors $\bar{x}$ et $\bar{y}$ sont les moyennes respectivement des x_i et des y_i.

Il est usuel de représenter une série à deux variable dans le plan, on obtient ainsi un nuage de points.

Exercice 21 - Relation poids/taille

Nous avons sur un échantillon de 10 individus de sexe masculin observé simultanément le poids et la taille. Les données suivantes ont été recueillies :

i	poids en kg (x_i)	taille en cm (y_i)
1	55	160
2	60	172
3	80	186
4	75	180
5	60	168
6	100	195
7	90	185
8	65	175
9	85	190
10	95	192

Représentez cette série statistique par un nuage de points.

Exercice 22 - Chiffre d'affaire

Nous observons depuis 2000 l'évolution du chiffre d'affaires d'une SSI. Voici les données recueillies :

i	année (x_i)	Ca en millio
1	2000	5
2	2001	6.5
3	2002	8.5
4	2003	10
5	2004	12
6	2005	13.5
7	2006	14.5
8	2007	15

Représentez cette série statistique par un nuage de points.

8.6.1 Coefficient de corrélation affine

La **covariance** de la série formée par l'observation simultanée des caractères X et Y est définie par

$$cov(X,Y) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})$$

On remarque que $cov(X,X) = V(X)$. Alors on définit le **coefficient de corrélation affine**

$$r = \frac{cov(X,Y)}{\sigma(X)\sigma(Y)}$$

Exercice 23 - Relation poids/taille

1. Calculez la covariance de X et Y.
2. Calculez le coefficient de corrélation affine

Exercice 24 - Chiffre d'affaire

1. Calculez la covariance de X et Y.
2. Calculez le coefficient de corrélation affine

8.6.2 Ajustement affine

Si $|r|$ est proche de 1, alors il est possible d'obtenir à partir de x_i une estimation convenable de y_i. A partir de x_i, on approche y_i avec une formule de la forme $ax_i + b$. Cela s'appelle une **régression linéaire**, ou encore un **ajustement affine**. Graphiquement, cela signifie qu'il existe une droite d'équation $y = ax + b$ passant près de chaque point du nuage de points. On calcule a et b de la façon suivante :

$$a = \frac{cov(X,Y)}{V(X)}$$

et

$$b = \bar{Y} - a\bar{X}$$

Les valeurs ci-dessus permettent d'effectuer la régression linéaire **de** Y **en** X.

Exercice 25 - Relation poids/taille

1. Un ajustement affine de Y en X est-il envisageable ?
2. Calculez les coefficients a et b de la droite d'ajustement.
3. Tracez la droite dans le repère.
4. A combien estimez-vous le taille d'un homme de 70 kilos ?

Exercice 26 - Chiffre d'affaire

1. Un ajustement affine de Y en X est-il envisageable ?
2. Calculez les coefficients a et b de la droite d'ajustement.
3. Tracez la droite dans le repère.
4. A combien estimez-vous le chiffre d'affaire en 2008 ?
5. Quand estimez-vous que le CA de cette entreprise dépassera 25 millions d'euros ?

8.6.3 Ajustement exponentiel

Quelquefois, lorsque $|r|$ est distant de 1, l'approximation est de mauvaise qualité, ce qui signifie que la droite ne passe pas au milieu du nuage de points. Une autre méthode est de poser $Z = \ln Y$, puis, si le coefficient de corrélation affine de X et Z a une valeur absolue suffisamment proche de 1, d'effectuer une régression de Z en X. On obtient donc une approximation de la forme

$$Z = aX + b$$

en composant les deux membres avec une fonction exponentielle, on a

$$e^Z = e^{aX+b}$$

comme $Z = \ln Y$, alors

$$e^{\ln Y} = e^{aX} e^b$$

ce qui donne
$Y = (e^a)^X e^b$
On appelle cela un **ajustement exponentiel**.

Exercice 27 - Ajustement exponentiel

Une base de données d'un site de discussion contient les dates d'inscriptions des utilisateurs. Nous avons noté depuis 2000 le nombre d'inscriptions par année.

i	année (x_i)	nombres d'inscriptions en milliers (y_i)
1	2000	5
2	2001	6.5
3	2002	8.5
4	2003	10
5	2004	12
6	2005	13.5
7	2006	14.5
8	2007	15

1. Représentez cette série statistique par un nuage de points.
2. Calculez la covariance de X et Y.
3. Calculez le coefficient de corrélation affine de X et Y.
4. Un ajustement affine de Y en X est-il envisageable ?
5. On pose $Z = \ln Y$, ajoutez les valeurs de Z dans une quatrième colonne.
6. Calculez le coefficient de corrélation affine X et Z
7. Un ajustement affine de Z en X est-il envisageable ?
8. Calculez les coefficients a et b de la droite d'ajustement de Z en X.
9. En déduire une approximation de Y de la forme ke^{pX}
10. Tracez la courbe dans le repère.
11. A combien estimez-vous le nombre d'inscription en 2008 ?
12. Quand estimez-vous que le nombre d'inscriptions annuel dépassera un million ?

Chapitre 9

Estimations

9.1 Estimation ponctuelle d'une moyenne

Cette section traite de la généralisation à une population d'observations effetuées sur un échantillon. Supposons que l'on souhaite déterminer la loi de la variable aléatoire X : "poids d'un français de sexe masculin ayant entre 20 et 30 ans". Il est impossible de tous les faire passer pour les peser, on va donc prendre les mesures sur un échantillon (de taille $n = 1000$ par exemple), puis les généraliser à une population, nous cherchons l'espérance mathématique m de X et son écart-type σ. Soit x_i le i-ème poids observé, alors on pose

$$m_e = \sum_{i=1}^{n} x_i$$

et

$$\sigma_e = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - m_e)^2}$$

Nous **admettrons** que m est estimé par m_e, et σ par $\sqrt{\frac{n}{n-1}}\sigma_e$ (si vous voulez savoir pourquoi, documentez-vous pour trouver ce qu'est le biais d'un estimateur).

9.2 Estimation ponctuelle d'une proportion

Supposons qu'une entreprise fabriquant des pétards veuille connaître la proportion de pétards défectueux fabriqués. Nous appellerons p cette proportion. Nous remarquons qu'il n'est possible de connaître cette proportion qu'en faisant sauter tous les pétards fabriqués. Nous devrons donc là-aussi nous restreindre à un échantillon. Soit p_e la proportion de pétards défecteux observé dans un échantillon de taille n, alors on estime p avec p_e.

9.3 Estimation par intervalle de confiance d'une moyenne

Il est intéressant, quand on effectue une estimation, de pouvoir mesurer la qualité de cette estimation. En donnant au passage la probabilité que l'on a de se tromper. Nous préfèrerons donc dorénavant des intervalles de confiance.

Reprenons notre problème de poids, nous souhaitons déterminer une intervalle $[a, b]$ tel que la probabilité que le poids moyen soit dans cet intervalle soit égale à $(1-p)$. p est le risque d'erreur, $1-p$ est le coefficient de confiance. Soit m_e la moyenne observée sur l'échantillon, σ $(= \sqrt{\frac{n}{n-1}}\sigma_e)$ l'estimation de l'écart-type. Alors **l'intervalle de confiance de la moyenne m avec le coefficient de confiance** $(1-p)$ est

$$[m_e - a\frac{\sigma}{\sqrt{n}}, m_e + a\frac{\sigma}{\sqrt{n}}]$$

avec a le réel tel que si X suit $\mathcal{N}(0,1)$, alors

$$P(-a < X < a) = (1-p)$$

Remarqons que, si F est la fonction de répartition de la loi normale centrée réduite, alors

$$\begin{aligned} P(-a < X < a) &= P(X < a) - P(X < -a) \\ &= F(a) - (1 - P(X > -a)) \\ &= F(a) - 1 + P(X < a) \\ &= 2F(a) - 1 \end{aligned}$$

Donc on cherche a tel que

$$F(a) = 1 - \frac{p}{2}$$

Par exemple, pour $p = 0.05$, on a $a = 1.96$.

9.4 Estimation par intervalle de confiance d'une proportion

Repenons notre problème de pétards, nous allons aussi déterminer un intervalle $[a,b]$ tel que la probabilité que le poids moyen soit dans cet intervalle soit égale à $(1-p)$. p est le risque d'erreur, $1-p$ est le coefficient de confiance. Soit f la proportion observée sur l'échantillon. Alors **l'intervalle de confiance de la fréquence f avec le coefficient de confiance** $(1-p)$ est

$$\left[f - a\sqrt{\frac{f(1-f)}{n-1}}, f + a\sqrt{\frac{f(1-f)}{n-1}}\right]$$

avec a le réel tel que

$$F(a) = 1 - \frac{p}{2}$$

Chapitre 10

Tests de validation d'hypothèses

10.1 Exemples introductifs

10.1.1 Comparaison d'une proportion observée et une proportion théorique

Soit un échantillon de 50 personnes, on observe un caractère c sur 32% d'entre elles. Nous souhaiterions comparer cette proportion à la proportion déjà connue d'un population de personnes présentant ce caractère, à savoir 30%. Une autre façon de formuler la question est : est-ce que ces personnes de l'échantillon proviennent cette population ? La différence, 2% ne semble pas très importante, mais il faudrait une méthode nous permettant de décider à partir de quand une différence est significative.

10.1.2 Comparaison d'une moyenne observée et une moyenne théorique

Une entreprise fabrique des paquets de café de 250 grammes chacun. Lors d'un contrôle qualité, vous prélevez 100 paquets dans la chaîne de fabrication. Il vous est demandé de vérifier si les paquets fabriqués sont conformes à la spécification. La moyenne et l'écart-type du poids des paquets de l'échantillon sont respectivement 253 et 3. Peut-on supposer que les paquets fabriqués proviennent bien d'une usine dans laquelle les paquets fabriqués ont un poids moyen de 250 grammes ?

10.2 Généralités

10.2.1 Les hypothèses

Dans un test de validité d'hypothèse, on oppose deux hypothèses, l'**hypothèse nulle**, et l'**hypothèse alternative**.

- l'hypothèse nulle H_0 est vérifiée si la différence entre la moyenne théorique et la moyenne observée n'est pas **significative**
- l'hypothèse alternative H_1 est vérifiée si la différence entre la moyenne théorique et la moyenne observée est **significative**

Dans un test de comparaison entre une moyenne théorique m_t et une moyenne observée m_o, on pose

- $H_0 : m_t = m_o$
- $H_1 : m_t \neq m_o$

10.2.2 La règle de décision

Le but d'un test est de vérifier la validité d'une hypothèse, il nous faut donc une règle qui nous permettra de trancher. Par exemple (la règle suivante est complètement bidon, elle sert juste d'illustration),

- Si $m_o \in [39, 41]$, alors on accepte H_0
- sinon on accepte H_1

Nous verrons ultérieurement comment définir une règle de décision pertinente.

10.2.3 Les erreurs

Vous vous doutez certainement que l'on n'est jamais à l'abri d'une erreur, il y a deux type d'erreurs :

– rejeter à tort H_0, dite **erreur de première espèce**

– accepter à tort H_0, dite **erreur de deuxième espèce**

Le risque de première espèce est mesuré par la probabilité α. **Le risque α est la probabilité de rejeter à tort l'hypothèse nulle**, dans la plupart des tests, on détermine la règle de décsion en fonction du risque α. Le risque de deuxième espèce est de probabilité β, incalculable en pratique. **Le risque β est la probabilité d'accepter à tort l'hypothèse nulle**. La seule idée à avoir présente à l'esprit est que le risque α est inversement proportionnel au risque β.

L'élaboration d'une régle de décision est légèrement différente selon que l'on compare des moyennes ou des proportions.

10.3 Comparaison d'une moyenne observée et d'une moyenne théorique

Reprenons l'exemple introductif sur les paquets de cafés. Posons nos deux hypothèses : soit m le poids moyen en grammes d'un paquet de café pris dans la population (l'ensemble des paquets de café produits),

– $H_0 : m = 250$

– $H_1 : m \neq 250$

Soit X la variable aléatoire qui à tout échantillon de taille 100 pris dans la population, associe sa moyenne. On sait, d'après le cours sur l'échantillonnage, que

$$X \ suit \ \mathcal{N}\left(250, \frac{3}{\sqrt{100}}\right)$$

Notez que l'on a utilisé l'estimation de l'écart-type, cependant, quand n est grand, l'ajustement $\sqrt{\frac{n}{(n-1)}}$ devient négligeable, on se permet souvent de l'oublier... Donc,

$$\frac{10}{3}(X - 250) \ suit \ \mathcal{N}(0, 1)$$

Soit $\alpha = 0.05$, fixons t tel que

$$p(|X| > t) = \alpha$$

On se souvient que $t = 1.96$. Donc, si les poids des paquets fabriqués par cette usine vérifient bien l'hypothèse nulle, alors

$$p\left(250 - 1.96\frac{3}{10} < X < 250 + 1.96\frac{3}{10}\right) = 0.95$$

Autrement dit,

$$p(249.412 < X < 250.588) = 0.95$$

Cela signifie que si l'on renouvelle l'expérience un grand nombre de fois, les moyenne des échantillons tomberont 19 fois sur 20 dans l'intervalle ci-dessus. De ce fait, si une moyenne d'un échantillon sort de cette intervalle, on peut considérer

– soit que l'on est dans le cas sur 20 défavorable

– soit qu'il convient de rejeter l'hypothèse nulle

Comme la probabilité de sortir de l'intervalle est très petite, on applique la méthode du maximum de vraisemblance, et on rejette l'hypothèse nulle. On énonce donc la règle de décision au seuil de risque $\alpha = 0.05$ de la façon suivante :

– Si $X \in [249.412, 250.588]$, alors on accepte H_0

– sinon on accepte H_1

Dans l'exemple pris ci avant, la différence est significative et on rejette H_0.

10.3.1 En bref

Soit M la moyenne théorique à laquelle vous souhaitez comparer la moyenne observée sur un échantillon de taille $n > 30$, de moyenne m et d'écart-type σ. Soit α le seuil de risque préconisé, on applique alors la règle de décision suivante :

– Si
$$m \in \left[M - t\frac{\sigma}{\sqrt{n}}, M + t\frac{\sigma}{\sqrt{n}}\right]$$
où t est tel que si K *suit* $\mathcal{N}(0,1)$, alors $p(-t < K < t) = 1 - \alpha$ alors on accepte H_0

– sinon on accepte H_1

Exercice 1

Les évaluations des vitesses de 72 voitures par un radar au bord d'une autoroute a donné les résultats suivants :

Vitesse	Effectif
$[90, 100[$	1
$[100, 110[$	3
$[110, 120[$	6
$[120, 130[$	16
$[130, 140[$	24
$[140, 150[$	13
$[150, 160[$	9

On note μ et σ la moyenne et l'écart-type des vitesses de l'ensemble de la population considérée. On suppose que la variable aléatoire X qui à tout échantillon de taille 72 prélevé dans cette population associe sa vitesse moyenne suit $\mathcal{N}(\mu, \frac{\sigma}{\sqrt{72}})$, où σ est l'estimation ponctuelle calculée à partir de l'échantillon.

1. fournir une estimation ponctuelle de μ et de σ
2. construire une régle de décision permettant de décider, avec un risque α de 0.05, si la vitesse moyenne des véhicules circulant sur ce tronçon d'autoroute est égal à 130 km/h.
3. peut-on considérer que la vitesse moyenne des voitures circulant sur ce tronçon est égale à 130 km/h ?
4. renouveler le test avec un risque de première espèce égal à 0.01

Exercice 2

Un responsable d'agence bancaire souhaite estimer le nombre de comptes dont le solde est en permanence supérieur à 10000 euros, pour ce faire il tire 45 comptes au hasard. Il constate que pour 12 d'entre eux le solde est resté supérieur à 10000 euros pendant les trois derniers mois. Son bras droit estime que parmi les 432 comptes dont il s'occupe, 1 sur 2 à un solde supérieur à 10000 euros depuis plus de trois mois, son bras gauche estime que cette proportion n'est que de 5%.

1. Construire deux test permettant de vérifer chacune des hypohèses énoncées au seuil de risque de 5%. Appliquez-les.
2. Même question avec $\alpha = 0.01$

10.4 Comparaison d'une proportion observée et d'une proportion théorique

Soit F la fréquence théorique à laquelle vous souhaitez comparer la fréquence f observée sur un échantillon de taille $n > 30$. Soit α le seuil de risque préconisé, on applique alors la règle de décision suivante :

– Si
$$f \in \left[F - t\sqrt{\frac{F(F-1)}{n}}, F + t\sqrt{\frac{F(F-1)}{n}}\right]$$
où t est tel que si K *suit* $\mathcal{N}(0,1)$, alors $p(-t < K < t) = 1 - \alpha$ alors on accepte H_0

– sinon on accepte H_1

10.5 Test d'hypothèse unilatéral

10.5.1 Comparaison d'une moyenne observée et une moyenne théorique

Le responsable de production d'une usine de plaques d'égoût souhaite vérifier, pour des raisons techniques, que les diamètres des plaques fabriquées n'excèdent pas 80 centimètres. Nous prélevons pour ce faire un échantillon de $n = 50$ plaques dans la production, et nous calculons leur diamètre moyen μ. Nous adopterons comme hypothèse nulle $H_0 : \mu = 80$, et comme hypothèse alternative $H_1 : \mu > 80$. Le test est dit unilatéral (par opposition à bilatéral) car on n'accepte H_1 que si μ s'écarte de la moyenne théorique en prenant des valeurs supérieures. Notez bien que l'hypohèse nulle n'est pas $\mu \leq 80$! Fixons comme risque d'erreur $\alpha = 0.05$. Alors nous adoptons la règle de décision suivante :

– Si
$$\mu \leq 80 + a\frac{\sigma}{\sqrt{50}}$$
où a vérifie $P(T < a) = 1 - \alpha$ où T suit une loi normale centrée réduite, alors on accepte H_0
– sinon, on accepte H_1.

Récapitulons, soit M une moyenne théorique d'une variable X correspondant à un critère dans une population. Nous prélevons n éléments dans une population (on ne sait pas s'il s'agit de la même), et la question que l'on se pose est : la valeur moyenne de ce critère dans la population dans laquelle à été prélevé l'échantillon est-elle plus basse (ou plus haute, sans perte de généralité) que M ? Soit X la variable aléatoire : "moyenne des valeurs du critère sur un échantillon de taille $n > 30$ prélevé dans la population". Alors X suit une loi normale de paramètres $\left(M, \frac{\sigma}{\sqrt{n}}\right)$. Soit μ la moyenne observée, on traduit alors la question par l'hypothèse alternative $H_1 : \mu < M$. On prend comme hypotyhèse nulle $H_0 : \mu = M$. Alors la règle de décision du test de validité d'hypothèse au seuil de risque α que l'on adopte est :

– Si
$$\mu \geq M + a\frac{\sigma}{\sqrt{50}}$$
où a vérifie $P(a < T) = 1 - \alpha$ où T suit une loi normale centrée réduite, alors on accepte H_0
– sinon, on accepte H_1.

10.5.2 Comparaison d'une proportion observée et une proportion théorique

On procède de façon analogue avec une proportion. Soit F la proportion théorique des individus présentant un caractère c dans une population. Si on prélève dans cette population un échantillon de taille $n > 30$, alors la variable aléatoire X : "nombre d'individus de l'échantillon présentant le caractère c" suit une loi normale de paramètres $\left(F, \sqrt{\frac{F(1-F)}{n}}\right)$ Nous disposons d'un échantillon de $n > 30$, sur lequel on observe une proportion f d'individus présentant le caractère c. Nous voulons savoir si les individus de cet échantillon proviennent d'une population dans laquelle une proportion d'individus plus élevée que F présente le caractère c. Nous poserons comme hypothèse nulle $H_0 : f = F$ et comme hypothèse alternative $f > F$. Nous adoptons, au seuil de risque de première espèce α, la règle de décision suivante :

– Si
$$f \leq F + a\sqrt{\frac{F(1-F)}{n}}$$
où a vérifie $P(T < a) = 1 - \alpha$ où T suit une loi normale centrée réduite, alors on accepte H_0
– sinon, on accepte H_1.

10.6 Test de comparaison de deux valeurs observées

10.6.1 Comparaison de deux proportions observées

Soit deux groupes de patients A et B sur lesquels ont été appliqués deux traitements distincts. On note $n_A > 30$ l'effectif du groupe A et $n_B > 30$ l'effectif du groupe B. On note f_A la proportion d'individus du groupe A sur lequel le traitement est efficace, f_B la proportion d'individus du groupe B sur lequel le traitement est efficace. La question que nous nous posons est “la différence de fréquence entre les deux groupes est-elle significative ?”. La question peut se reformuler : ”Peut-on considérer que, après les traitements, les individus des groupes A et B proviennent de la même population”. On pose les hypothèses suivantes :

- $H_0 : f_A = f_B$
- $H_1 : f_A \neq f_B$

Soit C le caractère : “le traitement a été efficace”. Soit X_A (resp. X_B) : la variable aléatoire : “proportion des individus présentant le caractère c parmi un échantillon de n_A (resp. n_B)individus prelévés dans la population”, la fréquence estimée de personnes présentant ce caractère dans cette population est f_A (resp. f_B). Alors

$$X_A \text{ suit } \mathcal{N}\left(f_A, \sqrt{\frac{f_A(1-f_A)}{n_A}}\right)$$

et

$$X_B \text{ suit } \mathcal{N}\left(f_B, \sqrt{\frac{f_B(1-f_B)}{n_B}}\right)$$

De ce fait,

$$X_A - X_B \text{ suit } \mathcal{N}\left(f_A - f_B, \sqrt{\frac{f_A(1-f_A)}{n_A} + \frac{f_B(1-f_B)}{n_B}}\right)$$

Supposons, l’hypothèse nulle vérifiée, alors

$$X_A - X_B \text{ suit } \mathcal{N}\left(0, \sqrt{\frac{f_A(1-f_A)}{n_A} + \frac{f_B(1-f_B)}{n_B}}\right)$$

On applique, étant donné un risque d’erreur de première espèce α, la règle de décision suivante :

- Si
$$|f_A - f_B| < a\sqrt{\frac{f_A(1-f_A)}{n_A} + \frac{f_B(1-f_B)}{n_B}}$$
où a est telle que $P(-a < T < a) = 1 - \alpha$, avec T suivant une loi normale centrée réduite, alors on accepte H_0.
- sinon, on accepte H_1.

10.6.2 Comparaison de deux moyennes observées

Soient m_A et m_B deux moyennes observées sur deux échantillons de tailles respectives $n_A > 30$ et $n_B > 30$, et σ_A, σ_B les écart-types estimés (ou connus) des populations dans lesquelles ont été prélevés les échantillons. On se demande si la différence entre m_A et m_B est significative. On pose les hypothèses suivantes :

- $H_0 : m_A = m_B$
- $H_1 : m_A \neq m_B$

Soit X_A (resp. X_B) : la variable aléatoire : “Valeur moyenne du critère c sur les individus d’un échantillon d’effectif n_A (resp. n_B) prelévés dans une population”, la moyenne estimée dans cette population est m_A (resp. m_B). Alors

$$X_A \text{ suit } \mathcal{N}\left(m_A, \frac{\sigma_A}{\sqrt{n_A}}\right)$$

et

$$X_B \text{ suit } \mathcal{N}\left(m_B, \frac{\sigma_B}{\sqrt{n_B}}\right)$$

De ce fait,

$$X_A - X_B \text{ suit } \mathcal{N}\left(m_A - m_B, \sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}}\right)$$

Supposons, l'hypothèse nulle vérifiée, alors

$$X_A - X_B \ suit \ \mathcal{N}\left(0, \sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}}\right)$$

On applique, étant donné un risque d'erreur de première espèce α, la règle de décision suivante :

– Si

$$|m_A - m_B| < a\sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}}$$

où a est telle que $P(-a < T < a) = 1 - \alpha$, avec T suivant une loi normale centrée réduite, alors on accepte H_0.

– sinon, on accepte H_1.