Statistiques et Probabilités

1 Statistiques

Il est souvent utile de résumer une série de données statistiques à l'aide de quelques quantités qui permettent de tester rapidement si une donnée extraite de la population étudiée est typique ou non (c'est à dire si elle ressemble aux autres ou si elle s'en écarte fortement).

1.1 Moyenne et Écart-type

Moyenne La moyenne représente une sorte de ≪ centre de gravité ≫ de la population. Si $\{x_i\}_{i=1\dots n}$ est une série statistique de n éléments, la moyenne notée $< x >$ ou $\bar{x}$ vaut :

$$< x >= \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

Variance La variance se calcule de deux manières : comme l'écart entre la moyenne des carrés et le carré de la moyenne, ou bien comme la moyenne des carrés des écarts à la moyenne.

$$\begin{aligned}\mathrm{Var}(x) &=< x^2 > - < x >^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)^2 \\ &=< (x- < x >)^2 >= \frac{1}{n}\sum_{i=1}^{n}(x_i- < x >)^2\end{aligned}$$

Démonstration On utilise l'identité remarquable $(a-b)^2$ et on réorganise les termes de la somme.

$$\begin{aligned}
< (x- < x >)^2 > &= \frac{1}{n}\sum_{i=1}^{n}(x_i- < x >)^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}(x_i^2+ < x >^2 -2x_i < x >) \\
&= \frac{1}{n}\sum_{i=1}^{n}x_i^2 + \frac{1}{n}\sum_{i=1}^{n} < x >^2 -2 < x > \frac{1}{n}\sum_{i=1}^{n}x_i \\
&=< x^2 > + \frac{1}{n} \times n < x >^2 -2 < x >< x > \\
&=< x^2 > + < x >^2 -2 < x >^2 \\
&=< x^2 > - < x >^2
\end{aligned}$$

Écart-type L'écart-type se calcule comme la racine carrée de la variance.

$$\sigma_x = \sqrt{\mathrm{Var}(x)}$$

C'est une mesure de la dispersion des données : plus σ est grand, plus les données sont largement répandues autour de la moyenne. Plus σ est petit au contraire, plus les données sont resserrées et proches de $< x >$.
Ainsi, une donnée qui se trouve à une distance de la moyenne inférieure ou du même ordre que σ ne sera pas surprenante. En revanche, une donnée qui s'écarte de plusieurs écarts-type de la moyenne sera plus inhabituelle et singulière.

1.2 Quantiles

Les quantiles divisent la population en sous-groupes de tailles égales, ordonnés les uns par rapport aux autres. Avant de pouvoir calculer des quantiles, il est donc toujours nécessaire de trier la liste de données par ordre croissant.

Médiane La médiane M divise la population en deux : la moitié de la population sera inférieure à la médiane et l'autre moitié sera supérieure. Son calcul dépend de la parité du nombre de données. Si $\{x_i\}_{i=1\ldots n}$ est une série statistique triée par ordre croissant, alors :

$$M = \begin{cases} \frac{1}{2}(x_{\frac{n}{2}} + x_{\frac{n}{2}+1}) & \text{si } n \text{ est pair} \\ x_{\frac{n+1}{2}} & \text{si } n \text{ est impair} \end{cases}$$

Assez souvent, la moyenne et la médiane seront proches. Lorsque ce n'est pas le cas, c'est généralement le signe que la distribution des données dans la population est relativement complexe et qu'il faut plus de détails pour identifier les valeurs typiques des valeurs inhabituelles.

Quartiles Les quartiles divisent la population en quatre : 25% de celle-ci sera inférieure au premier quartile Q_1, $Q_2 = M$ et 75% de la population sera inférieure au troisième quartile Q_3.
Exemple : considérons la série x={15 ;7 ;2 ;11 ;8 ;3 ;8 ;6 ;16 ;3 ;5 ;16 ;10} qui contient 13 valeurs. La liste triée s'écrit : x={2 ;3 ;3 ;5 ;6 ;7 ;8 ;8 ;10 ;11 ;15 ;16 ;16}. Les quartiles sont alors $Q_1 = x_4 = 5$, $Q_2 = M = x_7 = 8$ et $Q_3 = x_{10} = 11$.
La distance inter-quartile $Q_3 - Q_1$ est une autre mesure de la dispersion des données (comme l'écart-type). Par construction, 50% de la population est comprise entre Q_1 et Q_3. Plus les données sont étalées, plus la distance inter-quartile est grande.

Autres quantiles On utilise parfois les déciles, qui divisent la population en 10 groupes (10% de la population se trouve sous le premier décile D_1 et 90% sous le neuvième décile D_9), ou encore les centiles qui divisent la population en 100 groupes.

2 Probabilités

2.1 Événements

Un événement est une proposition qui est susceptible d'être réalisée ou non au cours d'une expérience aléatoire.
Par exemple, si on tire au hasard une carte dans un jeu de 52 cartes, on peut définir les événements suivants : T : ≪ La carte est un trèfle ≫, F : ≪ La carte est une figure ≫ ou encore R : ≪ La carte est rouge ≫. À un événement E, on peut associer une probabilité $P(E)$ entre 0 et 1, qui représente les chances que l'événement se produise quand on tente l'expérience.
Avec les exemples précédents, $P(T) = \frac{1}{4}$, $P(F) = \frac{3}{13}$ et $P(R) = \frac{1}{2}$. Un événement dont la probabilité est nulle est appelé impossible ; un événement dont la probabilité vaut 1 est appelé certain.

Contraire L'événement contraire de A est noté $\bar{A}$ et $P(\bar{A}) = 1 - P(A)$. Par exemple, $\bar{T}$ est l'événement ≪ La carte n'est pas un trèfle ≫.

Intersection L'intersection de deux événement A et B est notée $A \cap B$ et correspond à la réalisation conjointe de A et de B. Par exemple, $T \cap F$ est l'événement ≪ La carte est une figure de trèfle ≫.

Union L'union de deux événement A et B est notée $A \cup B$ et correspond à la réalisation de A ou de B (ou des deux). Par exemple, $T \cup R$ est l'événement ≪ La carte est un trèfle ou une carte rouge ≫.

propriété

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

Probabilités conditionnelles La probabilité de A sachant B, notée $P_B(A)$ ou $P(A|B)$, est la probabilité d'avoir l'événement A si on sait que l'événement B est réalisé.

$$P_B(A) = \frac{P(A \cap B)}{P(B)}$$

Indépendance Deux événements A et B sont dits indépendants si $P_B(A) = P(A)$, ce qui est équivalent à dire que $P_A(B) = P(B)$ ou encore $P(A \cap B) = P(A) \times P(B)$. En effet :

$$P_B(A) = P(A) \Leftrightarrow \frac{P(A \cap B)}{P(B)} = P(A) \Leftrightarrow P(A \cap B) = P(A) \times P(B)$$

Dire que deux événements sont indépendants revient à dire que le fait de savoir que B a eu lieu ne donne aucune information qui permette de modifier la probabilité associée à A.

Partition L'ensemble des issues possibles d'une expérience aléatoire est généralement noté Ω. Un ensemble d'événements $\{E_i\}_{i=1,...n}$ est appelé une partition de Ω si et seulement si :

- Leur union couvre Ω (dit autrement, la probabilité de leur union vaut 1) :
  $E_1 \cup E_2 \cup \cdots \cup E_n = \Omega$.
- toutes leurs intersections deux à deux sont vides (dit autrement, la probabilité de n'importe quelle intersection de deux événements distincts vaut 0) :
  Pour tout i, j tels que $i \neq j$, $E_i \cap E_j = \emptyset$

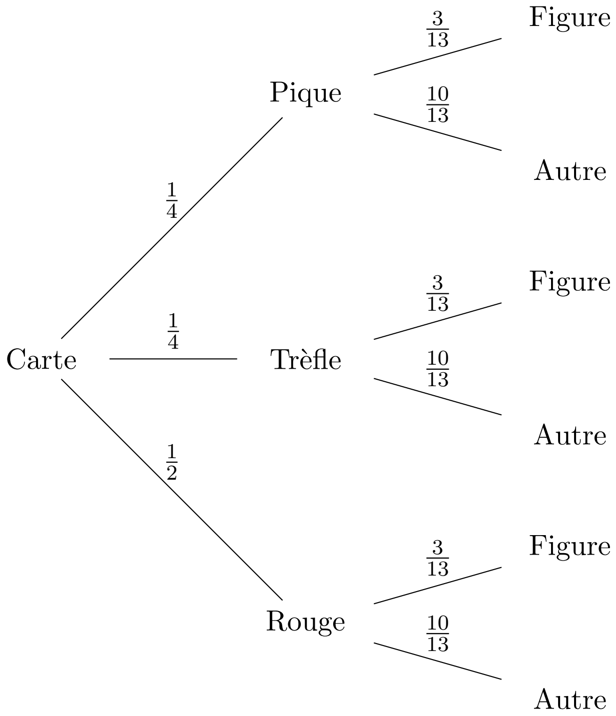
Par exemple, les événements T : « La carte est un trèfle », R : « La carte est rouge » et Q : « La carte est un pique » forment une partition.

Formule des probabilités totales Si $\{E_i\}_{i=1,...n}$ est une partition de Ω, alors :

$$\sum_{i=1}^{n} P(E_i) = 1$$

2.2 Arbres

On peut représenter des probabilités conditionnelles entre événements indépendants à l'aide d'un arbre dont les nœuds représentent des événements formant une partition de l'univers. On note la probabilité de suivre une cranche au-dessus de celle-ci. Par exemple, si on tire au hasard une carte dans un jeu de 52 :



Pour obtenir la probabilité d'un événement, on repère d'abord tous les chemins qui réalisent cet événement, puis on applique les deux règles suivantes :

— La probabilité d'un chemin est le produit des probabilités des branches

— La probabilité de l'événement est la somme des probabilités des chemins.

Par exemple, l'événement F : « La carte est une figure noire » est réalisée par deux chemins : C_1 : Trèfle-Figure et C_2 : Pique-Figure. Ainsi :

$$\begin{aligned} P(F) &= P(C_1) + P(C_2) \\ &= \frac{1}{4} \times \frac{3}{13} + \frac{1}{4} \times \frac{3}{13} \\ &= \frac{3}{52} + \frac{3}{52} \\ &= \frac{3}{26} \end{aligned}$$

2.3 Lois de probabilité

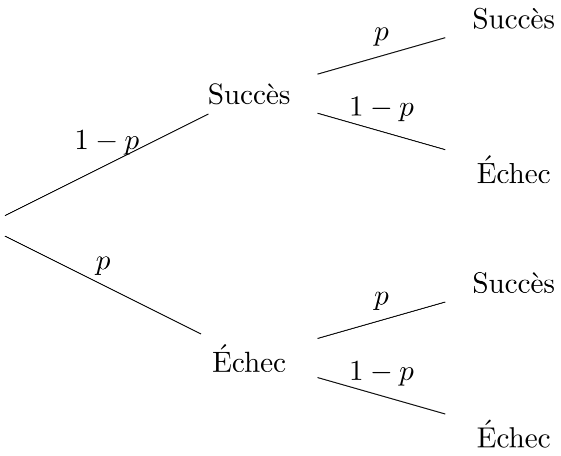
Une loi de probabilité associe une probabilité à à chaque événement d'une partition. On peut la résumer à l'aide d'un tableau, par exemple :

Evenement	T	R	C
Probabilité	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{2}$

Loi uniforme La loi uniforme correspond à des probabilités égales pour tous les événements. S'il y a n éléments, cette probabilité sera donc $\frac{1}{n}$ (du fait de la formule des probabilités totales). Par exemple, la loi de probabilité associée au lancement d'un dé à 6 faces est :

Face	1	2	3	4	5	6
Probabilité	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Loi binômiale On considère un jeu de hasard dans lequel la probabilité de gagner est p. Si on joue deux parties d'affilée (en supposant les parties indépendantes), on obtient l'arbre suivant :



La loi de probabilité associée au nombre total de parties gagnée est alors :

Succès	0	1	2
Probabilité	$(1-p)^2$	$2p(1-p)$	p^2

La loi binômiale de paramètres n et p est la loi de probabilité associée au nombre de succès de n parties indépendantes. Pour calculer la probabilité d'avoir k succès (avec $0 \leq k \leq n$), on s'intéresse aux chemins ayant k branches « succès » (qui vont contribuer d'un facteur p^k à la probabilité du chemin et $n-k$ branches « échec » (qui vont contribuer d'un facteur $(1-p)^{n-k}$). Le nombre de chemins est donné par le nombre de combinaisons de k éléments parmi n : $\binom{n}{k} = \frac{n!}{k!(n-k)!}$. La loi binômiale associe donc à un nombre k de succès la probabilité :

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