

CONCEPTS DE BASE – RAPPELS DE PROBABILITÉ

Notations

Une donnée x est une valeur observée d'une variable aléatoire X . Ω représente l'univers. ω est un événement élémentaire et \Pr est une **probabilité** qui mesure les événements. En d'autres termes, les événements sont des sous-ensembles de Ω dont on sait "mesurer" la probabilité. Lorsque l'univers Ω est un ensemble fini, on peut considérer que n'importe quel sous-ensemble de Ω est dans \mathcal{F} .

Les **événements** sont les éléments de la **tribu** \mathcal{F} (sous-ensemble de $\mathcal{P}(\Omega)$, contenant Ω , stable par réunion dénombrable, par intersection finie et par passage au complémentaire).

Probabilité

\Pr est donc une fonction de \mathcal{F} dans $[0,1]$ dont les propriétés sont les suivantes :

$$\Pr \left\{ \begin{array}{l} \Pr \text{ est une mesure } > 0, \text{ avec } \Pr(\Omega) = 1 \\ \Pr(A \cup B) = \Pr(A) + \Pr(B) \text{ lorsque } A \cap B = \emptyset \\ \Pr(A^c) = 1 - \Pr(A) \text{ où } A^c \text{ est le complémentaire de } A \\ \Pr(\emptyset) = 0 \\ \Pr(A \setminus B) = \Pr(A) - \Pr(B) \text{ à condition que } B \subseteq A \\ \Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B) \text{ en général} \end{array} \right.$$

Variable aléatoire réelle

Une **variable aléatoire réelle** (v.a.r. X) se définit comme suit :

$$\text{v.a.r. } X \left\{ \begin{array}{l} X : \Omega \rightarrow \mathfrak{R} \\ \omega \rightarrow X(\omega) \end{array} \right. \left\{ \begin{array}{l} \text{tel que si } I \text{ est un intervalle,} \\ \{\omega / X(\omega) \in I\} \text{ est dans la tribu } \mathcal{F} \\ \text{On note } (X \in I) \text{ l'événement } \{\omega / X(\omega) \in I\} = X^{-1}(I) \end{array} \right.$$

X permet donc de quantifier l'observation (d'associer à ω) un nombre réel. X est une fonction. On ne peut pas "calculer" $X(\omega)$, mais on peut mesurer l'événement $(X \in I)$ pour I intervalle de \mathfrak{R} . On montre que pour "mesurer" ces événements, il suffit de pouvoir mesurer les événements du type $(X < x)$ pour tout réel x .

Fonction de répartition

La **loi** d'une variable aléatoire réelle est définie par sa **fonction de répartition** définie sur \mathfrak{R} par :

$$F_X(x) = \Pr(X < x) = \Pr(X \in]-\infty, x[), \text{ pour tout } x \in \mathfrak{R}$$

Les propriétés de F_X sont :

$$F_X \left\| \begin{array}{l} F \text{ est croissante} \\ F(-\infty) = 0, F(+\infty) = 1 \\ F \text{ est continue à gauche en tout point} \end{array} \right.$$

On étudiera ici deux cas :

- X est une variable aléatoire discrète

X prend des valeurs en nombre fini ou dénombrable $\{a_1, a_2, \dots, a_n\}$ ou N . La fonction de répartition est alors une fonction en escalier. Les événements sont du type $(X=a_k)$ ou $(X=k)$. Dans ce cas, on se donne la loi de X en se donnant les nombres :

$$p_k = \Pr(X=a_k) \text{ ou } p_k = \Pr(X=k), \text{ avec } \sum_k p_k = 1, 0 \leq p_k \leq 1$$

On obtient donc un tableau de valeurs type

x	a_1	a_2	a_3		a_n
Pr	p_1	p_2	p_3		p_n

On calcule alors $\Pr(A)$ comme $\sum_{a_k \in A} p_k$, ou $\sum_{k \in A} p_k$

Si $X(\Omega) \subseteq N$, on utilise la fonction génératrice (cfr définitions diverses).

- X est une variable aléatoire absolument continue

X prend ses valeurs dans un ou plusieurs intervalles, sa fonction de répartition est continue sur \mathfrak{R} . Elle est dérivable (sauf peut-être en un nombre fini ou dénombrable de points) et il existe une fonction positive f appelée **densité** telle que :

$$\Pr(X < x) = F_X(x) = \int_{-\infty}^x f(t) dt, \text{ pour tout } x \in \mathfrak{R}.$$

On a alors $\Pr(A) = \int_A f(t) dt$ et dans ce cas, il n'y a pas de probabilité positive (masse) en des points isolés : $\Pr(X = x) = 0$ car $\int_x^x (\dots) = 0$.

On définit donc la **densité** comme une fonction f :

$$f \text{ densité} \left\| \begin{array}{l} f \text{ est une fonction réelle positive ou nulle} \\ \int_{-\infty}^{+\infty} f(t) dt = 1 \end{array} \right.$$

Dans ce cas, on se donne la loi de X par sa densité ou par sa fonction de répartition. On calcule $\Pr(a \leq X \leq b) = \Pr(a \leq X < b) = \Pr(a < X \leq b) = \Pr(a < X < b)$ comme $F(b) - F(a)$.

Remarque : il existe des variables aléatoires qui ne sont ni discrètes ni absolument continues, mais nous ne considérerons que ces deux cas là.