CHAPITRE 0^0
CONCEPTS PRÉALABLES SUR LES PROBABILITÉS

Le cours **STT289 — Probabilités** étant un préalable à notre cours, son contenu nous est considéré comme acquis. À titre de rappel, en voici ses objectifs et son contenu.

Objectifs

— Connaître les résultats fondamentaux et les méthodes de base du calcul des probabilités.
— Savoir quand et comment appliquer ces méthodes en situation de modélisation.

Contenu

— Espace de probabilité, probabilité conditionnelle, indépendance, formule de Bayes.
— Variables aléatoires discrètes et continues classiques : lois binomiale, de Poisson, binomiale négative, hypergéométrique, uniforme, normale, gamma, beta et autres.
— Vecteurs aléatoires et densités conjointes.
— Moments : espérance, variance, covariance, corrélation, fonction génératrice.
— Transformations de variables aléatoires.
— Distributions et espérances conditionnelles.
— Loi des grands nombres et théorème de la limite centrale.
— Génération de nombres pseudo-aléatoires.

Vous êtes encouragés à réviser vos notes de ce cours. Dans le présent chapitre, nous allons réviser certains concepts préalables sur les probabilités importants pour notre cours.

0.1 Probabilités, indépendance

Définition 0.1 Une **expérience aléatoire** est une expérience dont le **résultat** dépend du hasard, c'est-à-dire dont on peut déterminer parfaitement, par avance, tous les résultats possibles, mais dont on ne peut pas prévoir lequel de ces résultats sera réalisé. L'**espace échantillonnal** d'une expérience aléatoire, noté S (ou Ω), est l'ensemble de tous les résultats possibles de l'expérience. Un **évènement** est un sous-ensemble de l'espace échantillonnal. Un **évènement élémentaire** (ou **singleton**) est un évènement ayant un seul résultat.

Remarque Puisque les évènements sont des ensembles, il est possible d'effectuer les opérations usuelles de la théorie des ensembles sur les évènements, à savoir l'**union** ($\cup$), l'**intersection** ($\cap$) et le **complément** (^{c}) par rapport à l'espace échantillonnal S.

Définition 0.2 Soit A et B deux évènements d'une expérience aléatoire. Si $A \cap B = \emptyset$, alors A et B sont dits **disjoints** (ou **incompatibles**, **mutuellement exclusifs**). Si A et B sont disjoints et si $A \cup B = S$, alors A et B sont dits **complémentaires** (ou **contraires**). Par ailleurs, on appelle **système complet d'événements**

(ou **partition d'évènements**) tout ensemble $\{A_i\}_{i\in I}$ fini ou dénombrable d'événements incompatibles deux à deux dont $\bigcup_{i\in I} A_i = S$.

Exemple résolu 0.3 On lance une pièce de monnaie non biaisée à trois reprises et on note les résultats, à savoir la suite de piles (p) et de faces (f) obtenue. L'espace échantillonnal de cette expérience aléatoire est composé des $2^3 = 8$ résultats suivants :

$$S = \{\text{ppp, ppf, pfp, pff, fpp, fpf, ffp, fff}\}.$$

Sur cette expérience aléatoire on définit les évènements suivants :

- A : Obtenir exactement 2 piles.
- B : Obtenir pile au lancé 2 et face au lancé 3.
- C : Obtenir 0, 1 ou 3 piles.

Les résultats compris dans ces évènements sont

$$A = \{\text{ppf, pfp, fpp}\}, \quad B = \{\text{ppf, fpf}\} \quad \text{et} \quad C = \{\text{ppp, pff, fpf, ffp, fff}\}.$$

On remarque que

$$A \cup B = \{\text{ppf, pfp, fpp, fpf}\}, \quad A \cap B = \{\text{ppf}\} \quad \text{et} \quad A^c = \{\text{ppp, pff, fpf, ffp, fff}\}.$$

Par ailleurs, $A \cap C = \emptyset$, donc A et C sont disjoints, et $A \cup C = S$, donc ils sont également complémentaires. De fait, $\{A, C\}$ représente un système complet d'évènements.

Définition 0.4 Une **probabilité** est une application $P : S \to \mathbb{R}$ satisfaisant aux trois axiomes suivants :

- Axiome 1 : Pour tout ensemble A, $0 \leq P(A) \leq 1$.
- Axiome 2 : $P(S) = 1$.
- Axiome 3 : Pour toute suite finie ou infinie d'évènements disjoints, $P(\bigcup_i A_i) = \sum_i P(A_i)$.

Proposition 0.5 Soit A, B et C des évènements d'une expérience aléatoire.

- Propriété 1 : $P(A) = 1 - P(A^c)$.
- Propriété 2 : $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
- Propriété 3 : Si $S = \{1, \ldots, n\}$ et $P(\{i\}) = p_i$, alors $P(A) = \sum_{i\in A} p_i$. Par ailleurs, si $p_i = 1/n$, alors les évènements élémentaires sont dits **équiprobables** et dans ce cas, $P(A) = |A|/|S|$.

Définition 0.6 Soit A et B deux évènements d'une expérience aléatoire. Si $P(B) > 0$, alors la **probabilité conditionnelle** de A étant donné la réalisation de B, notée $P(A \mid B)$, est

$$P(A \mid B) = P(A \cap B)/P(B).$$

Proposition 0.7 (règle de multiplication) On déduit de cette définition les équations suivantes :

$$\begin{aligned} P(A \cap B) &= P(B) \cdot P(A \mid B) \\ &= P(A) \cdot P(B \mid A). \end{aligned}$$

Définition 0.8 Soit A et B deux évènements d'une expérience aléatoire. Ces évènements sont dits **indépendants** si $P(A \mid B) = P(A)$. Dans le cas contraire, les évènements sont dits **dépendants**. Par ailleurs, les évènements $\{A_i\}_{i\in I}$ sont dits :

- **indépendants deux à deux** si, pour tous $i, j \in I$ avec $i \neq j$, on a $P(A_i \cap A_j) = P(A_i) \cdot P(A_j)$,
- **mutuellement indépendants** si, pour tout $J \subset I$, on a $P\left(\bigcap_{j\in J} A_j\right) = \prod_{j\in J} P(A_j)$.

Exemple résolu 0.9 On lance une pièce de monnaie non biaisée à n reprises et on note les résultats, à savoir la suite de piles (p) et de faces (f) obtenue. Sur cette expérience aléatoire on définit les évènements suivants :

— A_i : Obtenir pile au lancé i.

Les évènements A_i sont mutuellement indépendants car $P(A_i) = 1/2$ et, pour tout choix d'indices $1 \leq i_1 < \cdots < i_k \leq n$, on a

$$P\left(A_{i_1} \cap \cdots \cap A_{i_k}\right) = (1/2)^k = P\left(A_{i_1}\right) \cdot \cdots \cdot P\left(A_{i_k}\right).$$

Exemple résolu 0.10 On lance une pièce de monnaie non biaisée à trois reprises et on note les résultats, à savoir la suite de piles (p) et de faces (f) obtenue. Sur cette expérience aléatoire on définit les évènements suivants :

— A : Obtenir le même résultat aux lancés 1 et 2.
— B : Obtenir le même résultat aux lancés 1 et 3.
— C : Obtenir le même résultat aux lancés 2 et 3.

Les probabilités de ces évènements sont $P(A) = P(B) = P(C) = 1/2$. Les résultats compris dans les intersections deux à deux de ces évènements sont $A \cap B = A \cap C = B \cap C = \{\text{ppp}, \text{fff}\}$, ce qui représente un évènement de probabilité $1/4$.

Ainsi, les évènements A, B, C sont indépendants deux à deux car

$$P(A \cap B) = P(A) \cdot P(B), \quad P(A \cap C) = P(A) \cdot P(C) \quad \text{et} \quad P(B \cap C) = P(B) \cdot P(C).$$

Cependant, ils ne sont pas mutuellement indépendants car $A \cap B \cap C = \{\text{ppp}, \text{fff}\}$, d'où

$$P(A \cap B \cap C) = 1/4 \neq 1/8 = P(A) \cdot P(B) \cdot P(C).$$

Proposition 0.11 Soit A et B deux évènements d'une expérience aléatoire. A et B sont indépendants si et seulement si $P(A \cap B) = P(A) \cdot P(B)$.

Proposition 0.12 (loi des probabilités totales) Soit A et B deux évènements d'une expérience aléatoire, où $P(B) > 0$. Alors

$$\begin{aligned} P(A) &= P(A \cap B) + P(A \cap B^c) \\ &= P(B) \cdot P(A \mid B) + P(B^c) \cdot P(A \mid B^c). \end{aligned}$$

De façon générale, si $\{B_i\}_{i \in I}$ est un système complet d'événements fini ou dénombrable où $P(B_i) > 0$ pour tout i, alors

$$\begin{aligned} P(A) &= \sum_{i \in I} P(A \cap B_i) \\ &= \sum_{i \in I} P(B_i) \cdot P(A \mid B_i). \end{aligned}$$

Par ailleurs, si C est un évènement, alors

$$\begin{aligned} P(A \mid C) &= \sum_{i \in I} P(A \cap B_i \mid C) \\ &= \sum_{i \in I} P(B_i \mid C) \cdot P(A \mid C \cap B_i). \end{aligned}$$

Théorème 0.13 (théorème de Bayes) Soit A et B deux évènements d'une expérience aléatoire, où $P(B) > 0$ et $P(A) > 0$. Alors

$$\begin{aligned} P(B \mid A) &= \frac{P(B) \cdot P(A \mid B)}{P(A)} \\ &= \frac{P(B) \cdot P(A \mid B)}{P(B) \cdot P(A \mid B) + P(B^c) \cdot P(A \mid B^c)}. \end{aligned}$$

De façon générale, si $\{B_i\}_{i \in I}$ est un système complet d'événements fini ou dénombrable où $P(B_i) > 0$ pour tout i, alors

$$\begin{aligned} P(B_j \mid A) &= \frac{P(B_j) \cdot P(A \mid B_j)}{P(A)} \\ &= \frac{P(B_j) \cdot P(A \mid B_j)}{\sum_{i \in I} P(B_i) \cdot P(A \mid B_i)}. \end{aligned}$$

0.2 Variables aléatoires et distributions

Définition 0.14 Soit une expérience aléatoire ayant S comme espace échantillonnal. Une **variable aléatoire** X associée à cette expérience aléatoire est une fonction $X : S \longrightarrow \mathbb{R}$ qui associe à chaque résultat de l'espace échantillonnal S une valeur réelle. L'ensemble des nombres réels que la variable aléatoire peut prendre s'appelle le **support** de X et on le note S_X. Une variable aléatoire est dite **discrète** si son support est un ensemble fini ou dénombrable de nombres réels. Une variable aléatoire est dite **continue** si son support est un intervalle de l'ensemble des nombres réels.

0.2.1 Variables aléatoires discrètes

Définition 0.15 Soit X une variable aléatoire discrète de support S_X. La **fonction de masse** (ou **loi de probabilité**, **fonction de masse marginale**) de X est la fonction $p_X : \mathbb{R} \longrightarrow [0,\ 1]$ telle que

$$p_X(x) = \mathrm{P}(X = x).$$

La **fonction de répartition** (ou **fonction de répartition marginale**) de X est la fonction $F_X : \mathbb{R} \longrightarrow [0,\ 1]$ telle que

$$\begin{aligned} F_X(x) &= P(X \leq x) \\ &= \sum_{a \leq x} p_X(a). \end{aligned}$$

Proposition 0.16 Une fonction $p : \mathbb{R} \longrightarrow [0,\ 1]$ est possiblement une fonction de masse d'une variable aléatoire discrète si et seulement si

$$\sum_{x \in \mathbb{R}} p(x) = 1.$$

Définition 0.17 Soit X une variable aléatoire discrète de support S_X. L'**espérance mathématique** (ou **moyenne**) de X est la quantité

$$\begin{aligned} E(X) &= \sum_{x \in \mathbb{R}} x \cdot p_X(x) \\ &= \sum_{x \in \mathbb{R}} x \cdot P(X = x), \end{aligned}$$

sa **variance** est

$$\begin{aligned} Var(X) &= E\left(\left(X - E(X)\right)^2\right) \\ &= \sum_{x \in \mathbb{R}} \left(x - E(X)\right)^2 \cdot p_X(x) \\ &= \sum_{x \in \mathbb{R}} \left(x - E(X)\right)^2 \cdot P(X = x) \end{aligned}$$

et son **écart-type** est

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

Par ailleurs, son **moment d'ordre n** est

$$\begin{aligned} E(X^n) &= \sum_{x \in \mathbb{R}} x^n \cdot p_X(x) \\ &= \sum_{x \in \mathbb{R}} x^n \cdot P(X = x). \end{aligned}$$

Définition 0.18 Une variable aléatoire discrète X suit une **loi uniforme discrète** de paramètre $a, b \in \mathbb{Z}$ avec $a < b$, notée $X \sim \text{Unif}(a, b)$, si

$$p_X(x) = \begin{cases} 1/n & \text{si } x \in \{a, a+1, \dots, b-1, b\} \\ 0 & \text{sinon,} \end{cases}$$

où $n = b - a + 1$. Dans ce cas,

$$E(X) = (a + b)/2 \quad \text{et} \quad Var(X) = (n^2 - 1)/12.$$

La fonction de répartition de X est

$$F_X(x) = \begin{cases} 0 & \text{si } x < a, \\ (x - a + 1)/n & \text{si } x \in \{a, \dots, b\}, \\ 1 & \text{si } x > b. \end{cases}$$

Définition 0.19 Une variable aléatoire discrète X suit une **loi de Bernoulli** de paramètre $p \in [0,\ 1]$, notée $X \sim \text{Bern}(p)$, si

$$p_X(x) = \begin{cases} p & \text{si } x = 1, \\ q & \text{si } x = 0, \\ 0 & \text{sinon,} \end{cases}$$

où $q = 1 - p$. Dans ce cas,

$$E(X) = p \quad \text{et} \quad Var(X) = pq.$$

La fonction de répartition de X est

$$F_X(x) = \begin{cases} 0 & \text{si } x < 0, \\ q & \text{si } 0 \leq x < 1, \\ 1 & \text{si } x \geq 1. \end{cases}$$

Définition 0.20 Une variable aléatoire discrète X suit une **loi binomiale** de paramètres $n \in \mathbb{N}$ et $p \in [0,\ 1]$, notée $X \sim \text{Bin}(n, p)$, si

$$p_X(x) = \begin{cases} \binom{n}{x} p^x q^{n-x} & \text{si } x \in \{0, \dots, n\}, \\ 0 & \text{sinon,} \end{cases}$$

où $q = 1 - p$. Dans ce cas,

$$E(X) = np \quad \text{et} \quad Var(X) = npq.$$

Définition 0.21 Une variable aléatoire discrète X suit une **loi de Poisson** de paramètre $\lambda \in \mathbb{R}_+^*$, notée $X \sim \text{Poisson}(\lambda)$, si

$$p_X(x) = \begin{cases} \lambda^x e^{-\lambda}/x! & \text{si } x \in \mathbb{N}, \\ 0 & \text{sinon.} \end{cases}$$

Dans ce cas,

$$E(X) = \lambda \quad \text{et} \quad Var(X) = \lambda.$$

Définition 0.22 Une variable aléatoire discrète X suit une **loi géométrique** de paramètre $p \in [0,\ 1]$, notée $X \sim \text{Geom}(p)$, si

$$p_X(x) = \begin{cases} q^{x-1} p & \text{si } x \in \mathbb{N}^*, \\ 0 & \text{sinon,} \end{cases}$$

où $q = 1 - p$. Dans ce cas,

$$E(X) = 1/p \quad \text{et} \quad Var(X) = q/p^2.$$

La fonction de répartition de X est

$$F_X(x) = \begin{cases} 0 & \text{si } x \leq 0, \\ 1 - q^k & \text{si } x \in \mathbb{N}^*. \end{cases}$$

Définition 0.23 Une variable aléatoire discrète X suit une **loi binomiale négative** de paramètre $r \in \mathbb{N}^*$, notée $X \sim \text{Binneg}(r)$, si

$$p_X(x) = \begin{cases} \binom{x-1}{r-1} q^{x-r} p^r & \text{si } x \in \{r, r+1, \dots\}, \\ 0 & \text{sinon,} \end{cases}$$

où $q = 1 - p$. Dans ce cas,

$$E(X) = r/p \quad \text{et} \quad Var(X) = rq/p^2.$$

Définition 0.24 Une variable aléatoire discrète X suit une **loi hypergéométrique** de paramètre $n, N_1, N_2 \in \mathbb{N}^*$, notée $X \sim \text{Hypergeom}(n, N_1, N_2)$, si

$$p_X(x) = \begin{cases} \dfrac{\binom{N_1}{x}\binom{N_2}{n-x}}{\binom{N}{n}} & \text{si } x \in \{0,\ 1, \dots, \min\{n, N_1, N_2\}\}, \\ 0 & \text{sinon,} \end{cases}$$

où $N = N_1 + N_2$. Dans ce cas,

$$E(X) = np \quad \text{et} \quad Var(X) = npq\left(\frac{N-n}{N-1}\right),$$

où $p = N_1/N$ et $q = N_2/N$.

0.2.2 Variables aléatoires continues

Définition 0.25 Soit X une variable aléatoire continue de support S_X. La **fonction de densité** (ou **loi de probabilité**, **densité de probabilité**, **fonction de densité marginale**) de X est la fonction $f_X : \mathbb{R} \longrightarrow \mathbb{R}_+$ telle que

$$P(a \leq X \leq b) = \int_a^b f_X(x)\, dx.$$

La **fonction de répartition** de X est la fonction $F_X : \mathbb{R} \longrightarrow [0,\ 1]$ telle que

$$\begin{aligned} F_X(x) &= P(X \leq x) \\ &= \int_{-\infty}^{x} f(t)\, dt. \end{aligned}$$

Remarque La fonction de densité de X se retrouve à partir de sa fonction de répartition via la formule

$$f_X(x) = \frac{d}{dx}\big(F_X(x)\big).$$

Remarque À noter que

$$P(x \leq X \leq x + dx) \approx f_X(x) \cdot dx.$$

Proposition 0.26 (formule des probabilités totales) Soit A un évènement. Alors

$$P(A) = \int_{-\infty}^{\infty} f_X(x) \cdot P(A \mid X = x)\ dx.$$

Proposition 0.27 Une fonction $f : \mathbb{R} \longrightarrow \mathbb{R}_+$ est possiblement une fonction de masse d'une variable aléatoire discrète de support S_X si et seulement si

$$\int_{-\infty}^{\infty} f(x)\, dx = 1.$$

Définition 0.28 Soit X une variable aléatoire continue de support S_X. L'**espérance mathématique** (ou **moyenne**) de X est la quantité

$$E(X) = \int_{-\infty}^{\infty} x f_X(x)\, dx,$$

sa **variance** est

$$\begin{aligned} Var(X) &= E\left(\big(X - E(X)\big)^2\right) \\ &= \int_{-\infty}^{\infty} \big(x - E(X)\big)^2 f_X(x)\, dx, \end{aligned}$$

et son **écart-type** est

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

Par ailleurs, son **moment d'ordre $\boldsymbol{n}$** est

$$E(X^n) = \int_{-\infty}^{\infty} x^n f_X(x)\, dx.$$

Définition 0.29 Une variable aléatoire continue X suit une **loi uniforme continue** de paramètre $a, b \in \mathbb{R}$ avec $a < b$, notée $X \sim \mathrm{Unif}(a, b)$, si

$$f_X(x) = \begin{cases} 1/(b-a) & \text{si } x \in [a, b], \\ 0 & \text{sinon.} \end{cases}$$

Dans ce cas,

$$E(X) = (a+b)/2 \quad \text{et} \quad Var(X) = (b-a)^2/12.$$

La fonction de répartition de X est

$$F_X(x) = \begin{cases} 0 & \text{si } x < a, \\ (x-a)/(b-a) & \text{si } a \le x < b, \\ 1 & \text{si } x \ge b. \end{cases}$$

Définition 0.30 Une variable aléatoire continue X suit une **loi normale** de paramètre $\mu \in \mathbb{R}$ et $\sigma^2 \in \mathbb{R}_+^*$, notée $X \sim \mathrm{N}(\mu, \sigma^2)$, si

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/(2\sigma^2)}.$$

Dans ce cas,

$$E(X) = \mu \quad \text{et} \quad Var(X) = \sigma^2.$$

Définition 0.31 Une variable aléatoire continue X suit une **loi exponentielle** de paramètre $\lambda \in \mathbb{R}_+^*$, notée $X \sim \mathrm{Exp}(\lambda)$, si

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si } x \in \mathbb{R}_+^*. \\ 0 & \text{sinon.} \end{cases}$$

Dans ce cas,

$$E(X) = 1/\lambda \quad \text{et} \quad Var(X) = 1/\lambda^2.$$

La fonction de répartition de X est

$$F_X(x) = \begin{cases} 0 & \text{si } x < 0, \\ 1 - e^{-\lambda x} & \text{si } x \ge 0. \end{cases}$$

Définition 0.32 Une variable aléatoire continue X suit une **loi gamma** de paramètre $\lambda \in \mathbb{R}_+^*$, notée $X \sim \Gamma(k, \theta)$, si

$$f_X(x) = \begin{cases} \dfrac{x^{k-1} e^{-x/\theta}}{\Gamma(k)\theta^k} & \text{si } x \in \mathbb{R}_+^*. \\ 0 & \text{sinon,} \end{cases}$$

où $\Gamma(k)$ est la fonction gamma. Dans ce cas,

$$E(X) = k\theta \quad \text{et} \quad Var(X) = k\theta^2.$$

0.2.3 Généralités

Définition 0.33 La fonction **indicatrice** (ou **caractéristique**) d'un sous-ensemble F d'un ensemble E est la fonction $I_F : E \longrightarrow \{0,\ 1\}$ telle que

$$I_F(x) = \begin{cases} 1 & \text{si } x \in F \\ 0 & \text{si } x \notin F \end{cases}.$$

Proposition 0.34 Soit $a, b \in \mathbb{R}$ et X, Y deux variables aléatoires telles que $Y = aX + b$. Alors

$$E(Y) = a \cdot E(X) + b \quad \text{et} \quad Var(Y) = a^2 \cdot Var(X).$$

L'espérance est donc une fonction linéaire. Par ailleurs, soit n variables aléatoires discrètes indépendantes $X_1, \dots, X_n$ et leur somme $S = \sum_{i=1}^{n} X_i$. Alors

$$E(S) = \sum_{i=1}^{n} E(X_i) \quad \text{et} \quad Var(S) = \sum_{i=1}^{n} Var(X_i).$$

De façon générale, si l'on considère la fonction $g(X)$, alors

$$E\big(g(X)\big) = \sum_{x \in \mathbb{R}} g(x) \cdot p_X(x).$$

Proposition 0.35 La variance est la différence entre l'espérance des carrés et le carré de l'espérance :

$$Var(X) = E(X^2) - E(X)^2.$$

Remarque L'espérance mathématique d'une variable aléatoire est son moment d'ordre 1 tandis que sa variance est le moment d'ordre 2 de l'écart entre la variable et son espérance.

Proposition 0.36 Soit $X \sim \text{Bin}(n, p)$. Si $n \geq 30$, $np \geq 5$ et $nq \geq 5$, où $q = 1 - p$, alors X suit approximativement une loi normale d'espérance np et de variance npq.

Proposition 0.37 Soit $X \sim \text{Poisson}(\lambda)$. Si $\lambda \geq 5$, alors X suit approximativement une loi normale d'espérance λ et de variance λ.