CHEAT SHEET
Notions de Probabilités Théoriques utiles pour le cours de Probabilités Appliquées et Créativité

La théorie des probabilités offre un socle d'outils mathématiques fondamental pour appréhender la modélisation des phénomènes aléatoires.

1 Concepts de Base

L'expérience aléatoire : elle est à la base de la question que l'on cherche à résoudre.
Univers : il s'agit de l'ensemble exhaustif des issues de l'expérience aléatoire. On le note traditionnellement Ω. Un élément $\omega \in \Omega$ est appelée issue ou réalisation de l'expérience aléatoire.
Evénement : il s'agit d'un sous-ensemble d'issues de l'expérience aléatoire. Si A est un événement alors $A \subseteq \Omega$.
Espace événementiel : il s'agit d'un ensemble d'événements, appelé tribu, et noté $\mathcal{F}$. Il est supposé non vide (on a toujours au moins $\Omega \in \mathcal{F}$), stable par passage au complémentaire (si $A \in \mathcal{F}$ alors $\overline{A} = \Omega \setminus A\mathcal{F}$) et par réunions au plus dénombrables (si $A_i \in \mathcal{F}$ pour tout $i \in \mathbb{N}$ alors $\bigcup_{i=0}^{+\infty} A_i \in \mathcal{F}$).
Probabilité : il s'agit d'une mesure, notée $\mathbb{P}$, permettant de comparer les événements. Mathématiquement il s'agit d'une fonction de $\mathcal{F}$ dans $[0, 1]$. Par définition tout événement $A \in \mathcal{F}$ a une probabilité $\mathbb{P}(A)$. Un événement de probabilité 1 est qualifié de presque sûrement certain (en particulier $\mathbb{P}(\Omega) = 1$). Un événement de probabilité 0 est qualifié de presque sûrement impossible (en particulier $\mathbb{P}(\emptyset) = 0$).
Espace de probabilité : Il s'agit du triplet $(\Omega, \mathcal{F}, \mathbb{P})$.

2 Règles de Base

Probabilité du complémentaire : $\mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A)$.
Probabilité d'une réunion disjointe : si $\forall i \neq j$ on a $A_i \cap A_j = \emptyset$ alors $\mathbb{P}(\cup_i A_i) = \sum_i \mathbb{P}(A_i)$.
Probabilité conditionnelle : si B est un événement de probabilité non nulle alors $\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$.
Indépendance : on dira que A et B sont deux événements indépendants si $\mathbb{P}(A|B) = \mathbb{P}(A)$. Cette relation est symétrique. En particulier on aura que $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.
Probabilités totales : pour A et B deux événements on a

$$\mathbb{P}(A) = \mathbb{P}(A|B)\mathbb{P}(B) + \mathbb{P}(A|\overline{B})\mathbb{P}(\overline{B})$$

et plus généralement si B_i est une partition Ω (i.e. $B_i \cap B_j = \emptyset$ si $i \neq j$ et $\cup_i B_i = \Omega$) alors

$$\mathbb{P}(A) = \sum_i \mathbb{P}(A|B_i)\mathbb{P}(B_i)$$

3 Variables aléatoires réelles : v.a.

Définition : Il s'agit de fonctions, notées par une majuscule, par exemple X, de Ω dans $\mathbb{R}$ vérifiant que $\{X \leq x\} = \{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{F}$ pour tout $x \in \mathbb{R}$.
Fonction de répartition (FDR) : si X est une v.a. alors la fonction F_X de $\mathbb{R}$ dans $[0, 1]$ définie pour tout $x \in \mathbb{R}$ par

$$F_X(x) = \mathbb{P}(X \leq x)$$

est appelée fonction de répartition de X. Elle caractérise la loi de la variable aléatoire et possède trois propriétés :

1. Elle est croissante au sens large : si $x < y$ alors $F_X(x) \leq F_X(y)$;
2. On a $\lim_{x\to-\infty} F_X(x) = 0$ et $\lim_{x\to+\infty} F_X(x) = 1$;
3. Elle est continue à droite : $\lim_{h\to 0^+} F_X(x+h) = F_X(x)$, $\forall x \in \mathbb{R}$.

3.1 Variables aléatoires discrètes

Il s'agit de variables aléatoires X qui prennent leurs valeurs dans un ensemble discret E (fini ou au plus dénombrable).
Fonction de masse : leurs lois se caractérise également avec une fonction appelée fonction de masse que l'on note p_X définie par

$$p_X(k) = \mathbb{P}(X = k) \text{ pour tous } k \in E.$$

Il y a un double lien entre fonction de masse et FDR :

$$\begin{array}{rcl} p_X(k) & = & F_X(k) - \lim\limits_{h\to 0^-} F_X(k+h) \text{ i.e. la différence entre la FDR prise au point } k \text{ et sa limite à gauche} \\ F_X(x) & = & \sum\limits_{k\leq x,\ k\in E} p_X(k) \end{array}$$

En particulier on a $p_X(k) > 0$ pour tout $k \in E$ et $\sum_{k\in E} p_X(k) = 1$.

3.1.1 Quelques lois discrètes classiques

Loi de Bernoulli $B(p)$: elle a un paramètre $p \in [0, 1]$ et $E = \{0, 1\}$ on a

$$\begin{array}{rcl} p_X(0) & = & \mathbb{P}(X = 0) = 1 - p. \\ p_X(1) & = & \mathbb{P}(X = 1) = p \end{array}$$

On en déduit que

$$F_X(x) = (1 - p)\ \mathbf{1}_{[0,1[}(x) +\ \mathbf{1}_{[1,+\infty[}(x)$$

Loi Binomiale, $B(n, p)$: elle a deux paramètres $n \in \mathbb{N}^*$ et $p \in [0, 1]$ et $E = \{0, 1, ..., n\}$ on a pour $k \in E$

$$p_X(k) = \mathbb{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}.$$

On en déduit que $\forall x \in \mathbb{R}^+$

$$F_X(x) = \sum_{k=0}^{\min(\lfloor x\rfloor,n)} \binom{n}{k} p^k (1 - p)^{n-k}$$

où $\lfloor x\rfloor$ désigne la partie entière de x et $F_X(x) = 0$ si $x < 0$.
Loi Géométrique, $G(p)$: elle a un paramètre $p \in [0, 1]$ et $E = \mathbb{N}^*$ on a $\forall k \in E$

$$p_X(k) = \mathbb{P}(X = k) = (1 - p)^{k-1} p.$$

On en déduit que $\forall x \in [1, +\infty[$

$$F_X(x) = p \sum_{k=1}^{\lfloor x\rfloor} (1 - p)^{k-1} = 1 - (1 - p)^{\lfloor x\rfloor}$$

et $F_X(x) = 0$ si $x < 1$.
Loi de Poisson, $\mathcal{P}(\lambda)$: elle a un paramètre $\lambda > 0$ et $E = \mathbb{N}$, pour tout $k \in E$ on a

$$p_X(k) = \mathbb{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}.$$

On en déduit que $\forall x \mathbb{R}^+$

$$F_X(x) = e^{-\lambda} \sum_{k=0}^{\lfloor x \rfloor} \frac{\lambda^k}{k!}$$

et $F_X(x) = 0$ si $x < 0$.

3.1.2 Interprétations et Relations entre les lois classiques

- Une variable de Bernoulli représente le succès ($X = 1$) ou l'échec ($X = 0$) d'une expérience aléatoire.
- La loi Binomiale $B(n, p)$ représente le nombre de succès dans une répétition de n expériences de Bernoulli indépendantes de même paramètre p.
- La loi géométrique $G(p)$ représente le nombre d'expériences de Bernoulli indépendantes de même paramètre p qu'il faut répéter pour avoir un premier succès.
- La loi de Poisson $\mathcal{P}(\lambda)$ représente une approximation de loi Binomiale $B(n, p_n)$ lorsque $n \to +\infty$ et $np_n \to \lambda$ quand $n \to +\infty$. Elle peut donc être vue comme une loi de comptage d'événements rares. On pratique (statistique) on utilise l'approximation de Poisson lorsque $n \geq 25$ et $np < 5$ (ou $n(1-p) < 5$ dans ce cas c'est $n - X$ qui est de loi de Poisson).

3.2 Variables aléatoires continues

Il s'agit de variables aléatoires X qui prennent leurs valeurs dans des intervalles de (ou dans tout) $\mathbb{R}$.
Densité : leur loi est caractérisée par la fonction densité f_X de $\mathbb{R}$ dans $\mathbb{R}^+$ vérifiant

$$\int_{-\infty}^{+\infty} f_X(t)\, dt = 1.$$

Il y a un double lien entre densité et FDR :

$$\begin{array}{rcl} f_X(x) & = & F'_X(x) \\ F_X(x) & = & \int_{-\infty}^{x} f_X(t)\, dt \end{array}$$

On retrouve assez facilement que si $a < b$ alors

$$\mathbb{P}(X \in [a, b]) = \int_a^b f(t)\, dt = F_X(b) - F_X(a)$$

Particularité : contrairement au cas discret on a $\mathbb{P}(X = x) = 0$ quel que soit $x \in \mathbb{R}$.
Support : On appelle support d'une v.a. continue X l'ensemble noté $Supp(X)$ défini par

$$Supp(X) = \{x \in \mathbb{R} : f_X(x) > 0\}$$

3.2.1 Quelques lois continues classiques

Loi Uniforme $U(a,b)$: elle a deux paramètres $a < b$, son support est (a,b) on a

$$f_X(x) = \frac{1}{b-a}\ \mathbf{1}_{]a,b[}(x).$$

On en déduit que

$$F_X(x) = \frac{x-a}{b-a}\ \mathbf{1}_{[a,b[}(x) +\ \mathbf{1}_{[a,b[}(x)$$

Loi Exponentielle, $\mathcal{E}(\lambda)$: elle a un paramètre $\lambda > 0$ son support est $\mathbb{R}^{+,*}$ et on a

$$f_X(x) = \lambda e^{-\lambda x}\ \mathbf{1}_{]0,+\infty[}(x).$$

On en déduit que

$$F_X(x) = (1 - e^{-\lambda x})\ \mathbf{1}_{[0,+\infty[}(x)$$

Loi Gamma, $\Gamma(n,\lambda)$: elle a deux paramètres $n \in \mathbb{N}^*$ et $\lambda > 0$, son support est $\mathbb{R}^{+,*}$ et on a

$$f_X(x) = \frac{\lambda^n x^{n-1}}{(n-1)!} e^{-\lambda x}\ \mathbf{1}_{]0,+\infty[}(x).$$

On en déduit que

$$F_X(x) = \left(\int_0^x \frac{\lambda^n t^{n-1}}{(n-1)!} e^{-\lambda t}\ dt\right)\ \mathbf{1}_{[0,+\infty[}(x)$$

Loi Normale ou de Gauss, $\mathcal{N}(\mu,\sigma^2)$: elle a deux paramètres $\mu \in \mathbb{R}$ et $\sigma > 0$ son support est $\mathbb{R}$ on a

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

On en déduit que

$$F_X(x) = \int_{-\infty}^x \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right)\ dt$$

3.2.2 Interprétation et relations entre les lois classiques

- La loi uniforme ne privilégie aucun(e) point (zone) dans l'intervalle.
- La loi exponentielle modélise le temps d'occurence d'un événement rare. La somme de n variables aléatoires indépendantes de loi $\mathcal{E}(\lambda)$ est de loi $\Gamma(n,\lambda)$. Ce qui signifie que la loi $\Gamma(n,\lambda)$ modélise le temps de $n$ème occurence d'un événement rare.
- La loi de Gauss est utilisée pour modéliser une erreur dans la mesure d'une quantité.
- Le TCL (Théorème Central Limite) établit que si on considère des v.a. X_i indépendantes et de même loi (quelconque) de moyenne μ et de variance σ^2 alors la loi des v.a. définies par :

$$Z_n = \frac{\sum_{i=1}^n X_n - n\mu}{\sigma\sqrt{n}}$$

tend, lorsque $n \to +\infty$, vers une loi $\mathcal{N}(0,1)$.
- Théorème fondamental de la simulation : si X est une v.a. de FDR F_X et si U est une v.a. Uniforme $U(0,1)$ alors la v.a. définie par $Y = F_X^{-1}(U)$ est de même loi que X.

4 Espérance, Variance et Covariance

4.1 Espérance

Il s'agit d'un estimateur (de position) des valeurs que prend une variable aléatoire X. C'est une moyenne pondérée. On la note $\mathbb{E}(X)$.
Définition : Formellement on a

$$\mathbb{E}(X) = \int_\Omega X(\omega)\ d\mathbb{P}(\omega)$$

Linéarité : de la formule ci-dessus on tire aisément que si X et Y sont deux v.a. sur $(\Omega, \mathcal{F}, \mathbb{P})$ et si $a, b \in \mathbb{R}$ alors

$$\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y)$$

Le résultat qui suit permet de relier le calcul d'une espérance d'une v.a. (intégrale sur un espace abstrait Ω) à un calcul intégral (sur $\mathbb{R}$) avec la loi de cette v.a.
Théorème de transfert : Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité et soit X une variable aléatoire dont la FDR est notée F_X. Alors : pour toute fonction φ mesurable, on a

$$\int_\Omega \varphi(X(\omega))\ d\mathbb{P}(\omega) = \int_{-\infty}^{+\infty} \varphi(x)\ dF_X(x)$$

Ainsi ce résultat se décline suivant

4.1.1 Cas discret

Si X est une v.a. discrète à valeurs dans E alors

$$\mathbb{E}(\varphi(X)) = \sum_{k \in E} \varphi(k)\mathbb{P}(X = k)$$

Exemples : pour les lois classiques discrète ci-dessus on peut calculer leurs espérances
- Bernoulli Si $X \sim b(p)$ alors $\mathbb{E}(X) = p$;
- Binomiale Si $X \sim B(n, p)$ alors $\mathbb{E}(X) = np$;
- Géométrique Si $X \sim G(p)$ alors $\mathbb{E}(X) = 1/p$;
- Poisson Si $X \sim \mathcal{P}(\lambda)$ alors $\mathbb{E}(X) = \lambda$.

4.1.2 Cas continu

Si X est une v.a. continue de densité f_X alors

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

Exemples : pour les lois classiques à densité ci-dessus on peut calculer leurs espérances
- Uniforme Si $X \sim U(a, b)$ alors $\mathbb{E}(X) = (a + b)/2$;
- Exponentielle Si $X \sim \mathcal{E}(\lambda)$ alors $\mathbb{E}(X) = 1/\lambda$;
- Gamma Si $X \sim \Gamma(n, \lambda)$ alors $\mathbb{E}(X) = n/\lambda$;
- Normale Si $X \sim \mathcal{N}(\mu, \sigma^2)$ alors $\mathbb{E}(X) = \mu$.

4.1.3 Moyenne empirique

Si X_i est une suite de v.a. indépendantes et identiquement distribuée (de moyenne $\mathbb{E}(X)$), on appelle moyenne empirique, et on la note $\overline{X}_n$, la variable aléatoire

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Par linéarité de l'espérance on observe que :

$$\mathbb{E}(\overline{X}_n) = \mathbb{E}(X)$$

Ceci se traduit par : l'espérance de la moyenne empirique de n v.a. de même moyenne μ est égale à cette moyenne.

4.2 Variance et Covariance

La variance d'une v.a., notée $\mathbf{Var}(X)$, mesure la dispersion des v.a. par rapport à leur moyenne. Il s'agit en fait de l'écart moyen quadratique entre la variable aléatoire et sa moyenne.
Définition : Etant donné deux v.a. X et Y on définit

1. La variance de X par
$$\mathbf{Var}(X) = \mathbb{E}\left([X - \mathbb{E}(X)]^2\right)$$
2. La covariance de X et Y par
$$\mathbf{Cov}(X,Y) = \mathbb{E}\left[(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))\right]$$

Propriétés de calcul : Pour effectuer les calculs on utilisera les formules équivalentes :

$$\mathbf{Var}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$$

on en déduit, au passage, qu'on a toujours $\mathbb{E}(X^2) \geq (\mathbb{E}(X))^2$ et on a

$$\mathbf{Cov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$$

On notera que la covariance quand à elle peut être de signe quelconque.
Autres propriétés : On observera que

1. $\mathbf{Cov}(X,X) = \mathbf{Var}(X)$;
2. $\mathbf{Cov}(X,Y) = \mathbf{Cov}(Y,X)$;
3. La covariance est bilinéaire : si X, Y, Z et W sont des v.a. et a, b, c, d des réels
$$\mathbf{Cov}(aX + bY, cZ + dW) = ac\mathbf{Cov}(X,Z) + ad\mathbf{Cov}(X,W) + bc\mathbf{Cov}(Y,Z) + bd\mathbf{Cov}(Y,W)$$
4. Attention pour la variance :
$$\mathbf{Var}(aX + bY) = a^2\mathbf{Var}(X) + 2ab\mathbf{Cov}(X,Y) + b^2\mathbf{Var}(Y)$$
5. Si X et Y sont indépendantes alors
$$\mathbf{Cov}(X,Y) = 0$$
on en déduit que $\mathbf{Cov}(X,1) = 0$. Attention : covariace nulle n'implique pas l'indépendance.
6. Si X et Y sont indépendantes alors $\mathbf{Var}(X+Y) = \mathbf{Var}(X) + \mathbf{Var}(Y)$, de même si $a \in \mathbb{R}$ alors $\mathbf{Var}(X+a) = \mathbf{Var}(X)$.

Calculs : Ainsi en utilisant à nouveau le théorème de transfert on a :

4.2.1 Cas discret

Si X est une v.a. discrète à valeurs dans E alors

$$\mathbb{E}(X^2) = \sum_{k \in E} k^2 \mathbb{P}(X = k)$$

Exemples : pour les lois classiques discrète ci-dessus on peut calculer leurs espérances
  Bernoulli Si $X \sim b(p)$ alors $\mathbb{E}(X^2) = p$ d'où $\mathbf{Var}(X) = p(1-p)$;
  Binomiale Si $X \sim B(n,p)$ alors $\mathbb{E}(X^2) = np(1-p) + n^2p^2$ d'où $\mathbf{Var}(X) = np(1-p)$;
  Géométrique Si $X \sim G(p)$ alors $\mathbb{E}(X^2) = (2-p)/p^2$ d'où $\mathbf{Var}(X) = (1-p)/p^2$;
  Poisson Si $X \sim \mathcal{P}(\lambda)$ alors $\mathbb{E}(X^2) = \lambda + \lambda^2$ d'où $\mathbf{Var}(X) = \lambda$.

4.2.2 Cas continu

Si X est une v.a. continue de densité f_X alors

$$\mathbb{E}(X^2) = \int_{-\infty}^{+\infty} x^2 f_X(x)\ dx$$

Exemples : pour les lois classiques à densité ci-dessus on peut calculer leurs espérances
  Uniforme Si $X \sim U(a,b)$ alors $\mathbb{E}(X^2) = (b^2 + ab + a^2)/3$ d'où $\mathbf{Var}(X) = (b-a)^2/12$;
  Exponentielle Si $X \sim \mathcal{E}(\lambda)$ alors $\mathbb{E}(X^2) = 2/\lambda^2$ d'où $\mathbf{Var}(X) = 1/\lambda^2$; ;
  Gamma Si $X \sim \Gamma(n,\lambda)$ alors $\mathbb{E}(X^2) = (n+n^2)/\lambda^2$ d'où $\mathbf{Var}(X) = n/\lambda^2$;
  Normale Si $X \sim \mathcal{N}(\mu, \sigma^2)$ alors $\mathbb{E}(X^2) = \mu^2 + \sigma^2$ d'où $\mathbf{Var}(X) = \sigma^2$.

4.2.3 Variance de la moyenne empirique

Si X_i est une suite de v.a. indépendantes et identiquement distribuée (de moyenne $\mu = \mathbb{E}(X)$ et de variance $\sigma^2 = \mathbf{Var}(X)$), alors

$$\mathbf{Var}(\overline{X}_n) = \mathbf{Var}\left(\frac{1}{n}\sum_{k=1}^n X_k\right) = \frac{1}{n^2}\sum_{k=1}^n \mathbf{Var}(X_k) = \frac{\sigma^2}{n}$$

i.e. la variance de la moyenne empirique diminue avec la taille de l'échantillon. Ce résultat est la clé de la loi des Grands Nombres : en effet lorsque $n \to +\infty$ on a $\mathbf{Var}(\overline{X}_n) \to 0$ ce qui traduit que $\overline{X}_n \to \mu$ lorsque $n \to \infty$ dans un certain sens.

5 Lois Jointes et lois Conditionnelles

Un vecteur aléatoire $(X_1, ..., X_n)$ sur $(\Omega, \mathcal{F}, \mathbb{P})$ est constitué de n composantes qui sont des variables aléatoires sur $(\Omega, \mathcal{F}, \mathbb{P})$.

5.0.1 Loi jointe

On appelle loi jointe la fonction décrivant les probabilités :

$$F_{(X_1,...,X_n)}(x_1, ..., x_n) = \mathbb{P}(X_1 \leq x_1, ..., X_n \leq x_n)$$

où $(x_1, ..., x_n)$ est un vecteur arbitraire de $\mathbb{R}^n$.
Lois marginales : étant donné un vecteur aléatoire $(X_1, ..., X_n)$ sur $(\Omega, \mathcal{F}, \mathbb{P})$ on appelle lois marginales les lois de ses composantes : c'est à dire les lois de X_1, ..., X_n.

De la même manière que pour les variables aléatoires on distingue deux cas
Cas discret
Dans ce cas la loi jointe est exprimée par une fonction de masse jointe :

$$p_{(X_1,...,X_n)}(k_1,...,k_n) = \mathbb{P}(X_1 = k_1,...,X_k = k_n)$$

et on a

$$F_{(X_1,...,X_n)}(x_1,...,x_n) = \sum_{k_1 \leq x_1} \cdots \sum_{k_n \leq x_n} p_{(X_1,...,X_n)}(k_1,...,k_n)$$

Cas continu
Dans ce cas la loi jointe est exprimée par une fonction de densité jointe :

$$f_{(X_1,...,X_n)}(x_1,...,x_n) \geq 0 \text{ vérifiant } \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f_{(X_1,...,X_n)}(x_1,...,x_n)\, dx_1 \cdots\, dx_n = 1$$

et on a

$$F_{(X_1,...,X_n)}(x_1,...,x_n) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f_{(X_1,...,X_n)}(t_1,...,t_n)\, dt_1 \cdots\, dt_n$$

5.1 Loi conditionnelle

Etant donné un vecteur aléatoire (X, Y) on définit la loi conditionnelle de X sachant $\{Y = y\}$ de la façon suivante :
Cas discret
On définit la fonction de masse condtionnelle de X sachant $\{Y = y\}$ par

$$p_{X|Y=y}(k) = \frac{\mathbb{P}(X = k, Y = y)}{\mathbb{P}(Y = y)}$$

Cas continu
On définit la densité conditionnelle de X sachant $\{Y = y\}$ par

$$f_{X|Y=y}(x) = \frac{f_{(X,Y)}(x, y)}{f_Y(y)}$$

On étend sans difficulté (si ce n'est d'écriture) ces définitions aux vecteurs aléatoires de dimension supérieure.

5.2 Indépendance

Si X et Y sont des v.a. indépendantes alors
Cas discret

$$p_{X|Y=y}(k) = p_X(k)$$

en conséquence $p_{(X,Y)}(k, y) = p_X(k)p_Y(y)$.
Cas continu

$$f_{X|Y=y}(x) = f_X(x)$$

en conséquence $f_{(X,Y)}(x, y) = f_X(x)f_Y(y)$.
Remarque : De fait si on ne connait que les lois marginales d'un vecteur, on n'est pas capable d'écrire les lois jointes. En revanche si on sait que les composantes sont indépendantes alors dans ce cas la loi jointe est le produit des lois marginales.

5.3 Espérance conditionnelle

Etant donné deux variables aléatoires X et Y on définit l'espérance conditionnelle de X sachant $\{Y = y\}$ comme la fonction

$$\varphi(y) = \mathbb{E}(X|Y = y)$$

Son calcul explicite se fait suivant :
Cas discret

$$\varphi(y) = \mathbb{E}(X|Y = y) = \sum_{k\in E} kp_{X|Y=y}(k)$$

Cas continu

$$\varphi(y) = \mathbb{E}(X|Y = y) = \int_{-\infty}^{+\infty} xf_{X|Y=y}(x)$$

Définition : Une fois ce calcul réalisé on peut alors définir l'espérance conditionnelle de X sachant Y notée $\mathbb{E}(X|Y)$ par

$$\mathbb{E}(X|Y) = \varphi(Y)$$

Remarque importante : Bien noter que l'espérance conditionnelle de X sachant Y est une variable aléatoire. En particulier on a les résultats suivants :

1. L'espérance conditionnel est linéaire : si X, Y, Z sont des v.a. et $a, b \in \mathbb{R}$ alors
$$\mathbb{E}(aX + bY|Z) = a\mathbb{E}(X|Z) + b\mathbb{E}(Y|Z)$$
2. Si X et Y sont des v.a. indépendantes alors
$$\mathbb{E}(X|Y) = \mathbb{E}(X)$$
3. L'espérance conditionnelle $\mathbb{E}(X|Y)$ est la meilleure prédiction (au sens des moindres carrés) de X connaissant la v.a. Y.

5.4 Formule des espérances totales

Ce résultat traduit que $\mathbb{E}(\mathbb{E}(X|Y)) = \mathbb{E}(X)$
Cas discret : si Y est à valeurs dans E

$$\mathbb{E}(X) = \sum_{y\in E} \mathbb{E}(X|Y = y)\mathbb{P}(Y = y)$$

Cas continu :

$$\mathbb{E}(X) = \int_{-\infty}^{+\infty} \mathbb{E}(X|Y = y)f_Y(y)\ dy$$

6 Théorèmes Importants

Formule des Probas totales
Formule de Bayes
Formule des Espérances totales
Loi des Grands Nombres
Théorème Central Limite