Chapitre 1

Intégrales doubles et probabilités

Dès lors que l'on traite d'un couple ou d'un n-uplet de variables aléatoires, l'intégration devient un outil incontournable. Nous allons traiter ici de quelques exemples de calcul d'intégrales multiples et nous introduirons l'utilité de ces calculs en théorie des probabilités.

1.1 Qu'est ce qu'une intégrale double ?

Soit une fonction réelle f à deux variables x et y. Le graphe de f est une surface qui représente les valeurs $f(x, y)$ pour tous les couples (x, y) sur le domaine de définition de la fonction. On va considérer que cette fontion est continue. A la rigueur, elle peut même être discontinue sur un nombre fini de points.

Soit maintenant une région quelconque $\mathcal{D}$ du plan et $\delta\mathcal{D}$ son bord. On souhaite calculer le volume du cylindre sous le graphe de f et dont les bords sont délimités par $\delta\mathcal{D}$. Ce volume noté $I_{\mathcal{D}}$ peut être négatif si le graphe de f est négatif et positif sinon. Pour calculer $I_{\mathcal{D}}$, on va partitionner l'espace horizontal en petits rectangles dont les côtés ont pour longueur Δx et Δy. On va également noter $M_i : (x_i, y_i)$ le centre (on peut en fait prendre n'importe quel point) des rectangles d'intersection non nulle avec le domaine $\mathcal{D}$. Si les rectangles sont suffisamment petits, le volume signé sous le graphe de f (l'intégrale $I_{\mathcal{D}}$) est approché comme une somme de parallélélipèdes rectangles de volume $\Delta x \times \Delta y \times f(x_i, y_i)$ et on obtient

$$\widehat{I_{\mathcal{D}}} = \sum_{M_i \in \mathcal{D}} f(x_i, y_i) \Delta x \Delta y,$$

comme valeur approchée de $I_{\mathcal{D}}$ (Fig. 1.1).

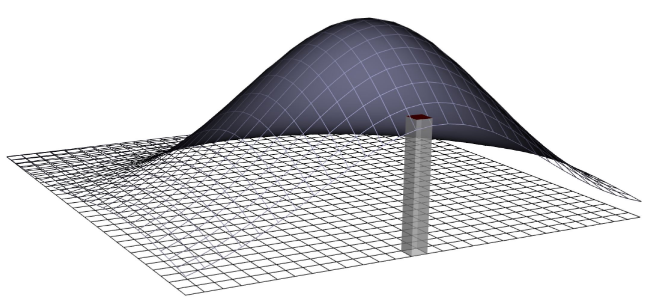

FIGURE 1.1 – Le volume sous la courbe en 3D est approximé en calculant les volumes de parallélépipèdes rectangles tels que celui représenté en gris.

En fait, l'intégrale que l'on cherche est la limite de la quantité ci-dessus quand les rectangles deviennent de plus en plus petits :

$$I_{\mathcal{D}} = \lim_{(\Delta x, \Delta y) \to 0} \sum_{M_i \in \mathcal{D}} f(x_i, y_i) \Delta x \Delta y,$$

et plutôt que de garder cette expression compliquée, on préfèrera l'écriture suivante :

$$I_{\mathcal{D}} = \iint_{\mathcal{D}} f(x, y) \, \mathrm{d}x \, \mathrm{d}y.$$

La somme est remplacée par les grands "s" de l'intégrale, et l'élément de surface $\Delta x \times \Delta y$ par l'élément différentiel $\mathrm{d}x \times \mathrm{d}y$ qui rappelle le fait que l'on travaille avec une limite.

1.2 Comment les calcule-t-on ?

La difficulté du calcul d'une intégrale double va essentiellement dépendre de la complexité du domaine $\mathcal{D}$ sur lequel on cherche à intégrer et de la difficulté à calculer la primitive de la fonction à intégrer.

1.2.1 Domaines rectangulaires

Ce sont les cas les plus simples, nous allons détailler la manière de procéder.

Exemple 1. Considérons la fonction $f(x, y) = x + x^2 y + 3$ définie sur $\mathbb{R}^2$ (Fig. 1.2). Soit le domaine $\mathcal{D} = \{(x, y) \in \mathbb{R}^2 \text{ t.q. } -1 \leq x \leq 1, 0 \leq y \leq 2\}$ qui correspond au rectangle $[-1, 1] \times [0, 2]$ sur lequel on souhaite intégrer. On veut calculer

$$I_{\mathcal{D}} = \iint_{\mathcal{D}} (x + x^2 y + 3) \, \mathrm{d}x \, \mathrm{d}y = \int_0^2 \left[\int_{-1}^1 (x + x^2 y + 3) \, \mathrm{d}x \right] \mathrm{d}y.$$

Le sens d'écriture indique ici que l'on va d'abord intégrer par rapport à x puis ensuite intégrer par rapport à y. On a :

$$\int_{-1}^1 (x + x^2 y + 3) \, \mathrm{d}x = \left[\frac{1}{2} x^2 + \frac{1}{3} x^3 y + 3x \right]_{-1}^1 = \frac{2}{3} y + 6.$$

Le résultat de cette intégrale dépend de y car cette variable était dans ce cas considérée comme une constante. L'intégrale est finalement calculée avec :

$$\int_0^2 \left[\int_1^1 (x - x^2 y + 3) \, \mathrm{d}x \right] \mathrm{d}y = \int_0^2 \left[\frac{2}{3} y + 6 \right] \mathrm{d}y = \left[\frac{1}{3} y^2 + 6y \right]_0^2 = \frac{40}{3}$$

■

On peut se poser la question de l'ordre d'intégration. On aurait en effet pu poser l'intégrale de la manière suivante :

$$I_{\mathcal{D}} = \iint_{\mathcal{D}} (x + x^2 y + 3) \, \mathrm{d}x \, \mathrm{d}y = \int_1^1 \left[\int_0^2 (x + x^2 y + 3) \, \mathrm{d}y \right] \mathrm{d}x.$$

Cela sous-entend que l'on aurait commencé à calculer l'intégrale en fonction de y. Le résultat aurait été le même (preuve laissée en exercice) grâce au théorème suivant dont nous admettrons l'existence :

Théorème 1:
Théorème de Fubini pour des domaines rectangulaires
Soit $\mathcal{D} = [a, b] \times [c, d]$ un domaine rectangulaire de $\mathbb{R}^2$ et f une fonction réelle, continue, de deux variables, alors f est intégrable sur $\mathcal{D}$ et l'on a :

$$I_{\mathcal{D}} = \iint_{\mathcal{D}} f(x, y)\,\mathrm{d}x\,\mathrm{d}y = \int_c^d \left[\int_{\boldsymbol{a}}^{\boldsymbol{b}} \boldsymbol{f(x, y)\,\mathrm{d}x}\right] \mathrm{d}y = \int_a^b \left[\int_{\boldsymbol{c}}^{\boldsymbol{d}} \boldsymbol{f(x, y)\,\mathrm{d}y}\right] \mathrm{d}x.$$

Le choix de l'ordre d'intégration est souvent conduit par la facilité avec laquelle les intégrales s'enchainent.

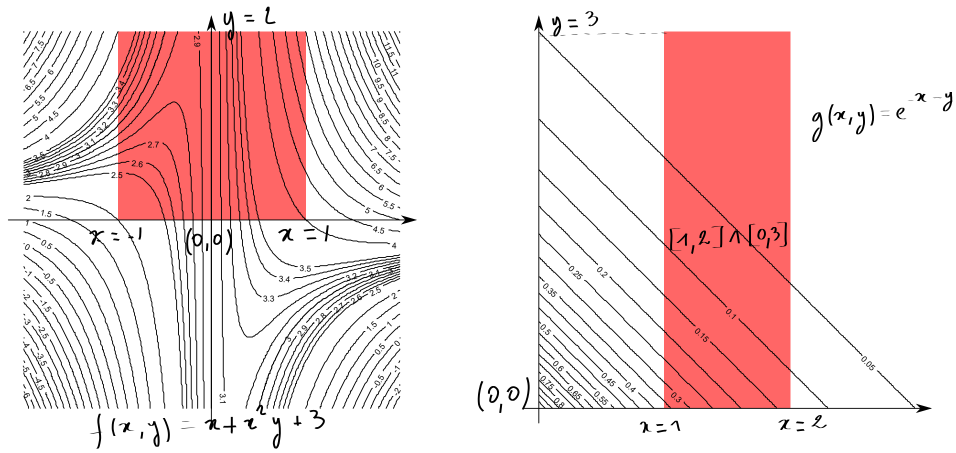

FIGURE 1.2 – Isolignes des fonctions des exemples 1. et 2. et domaines d'intégration.

Exemple 2. Soit la fonction $g(x, y) = \exp(-x - y),\ x \geq 0, y \geq 0$ (Fig. 1.2). On souhaite l'intégrer sur le rectangle $\mathcal{D} = [1, 2] \times [0, 3]$. On a alors

$$I_{\mathcal{D}} = \iint_{\mathcal{D}} e^{-x-y}\,\mathrm{d}y\,\mathrm{d}x = \int_1^2 \int_0^3 e^{-x} e^{-y}\,\mathrm{d}y\,\mathrm{d}x = \int_1^2 e^{-x}\,\mathrm{d}x \times \int_0^3 e^{-y}\,\mathrm{d}y = \left(e^{-1} - e^{-2}\right)\left(1 - e^{-3}\right).$$

On voit ici que lorsque la fonction à intégrer peut s'écrire sous la fome d'un produit de fonctions à une variable le calcul de l'intégrale double se ramène au produit de deux intégrales simples.
■

1.2.2 Domaines non rectangulaires

Dans la plupart des cas, les domaines d'intégration sont des surfaces de forme variable. Nous allons montrer comment calculer des intégrales doubles dans ce cas.

Exemple 3. Domaine compris entre les graphes de deux fonctions et deux droites verticales. Soit le domaine

$$\mathcal{D} = \left\{(x, y) \in \mathbb{R}^2 \text{ t.q. } 1 \leq x \leq 4, (x - 2)^2 - 4 \leq y \leq -(x - 3)^2 + 4\right\}.$$

Il s'agit de calculer l'intégrale de la fonction $f(x, y) = y - x$ sur ce domaine. En posant $u(x) = (x-2)^2 - 4$ et $v(x) = -(x-3)^2 + 4$, on note immédiatement que $u(x) \leq v(x),\ 1 \leq x \leq 4$ (le graphe de u est toujours sous celui de v). On peut alors écrire :

$$\iint_{\mathcal{D}} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_1^4 \left[\int_{u(x)}^{v(x)} y - x\,\mathrm{d}y\right]\mathrm{d}x = \int_1^4 \left[\frac{1}{2}y^2 - xy\right]_{u(x)}^{v(x)}\mathrm{d}x = \int_1^4 \left(5x^2 - 25x + \frac{25}{2}\right)\mathrm{d}x = -45$$

■

Exemple 4. Nous allons calculer la surface d'une ellipse par une intégrale double de la fonction unité sur le domaine

$$\mathcal{D} = \left\{(x,y) \in \mathbb{R}^2 \text{ t.q. } \frac{x^2}{a^2} + \frac{y^2}{b^2} \leq 1\right\}.$$

Par symétrie de l'ellipse relativement à son axe horizontal, puis à son vertical, on en déduit que la surface totale de l'ellipse est égale à quatre fois la surface sous la courbe d'équation $y = b\sqrt{1 - x^2/a^2}$ (Fig. 1.3) pour x compris entre 0 et a, soit

$$I = 4 \times \int_0^a \left[\int_0^{b\sqrt{1-x^2/a^2}} \mathrm{d}y\right]\mathrm{d}x = 4\int_0^a b\sqrt{1 - \frac{x^2}{a^2}}\,\mathrm{d}x.$$

On voit ici que l'on a commencé à intégrer relativement à y car une des bornes de l'intégrale dépend de la variable x. On ne pouvait pas permuter l'ordre d'intégration dans ce cas.

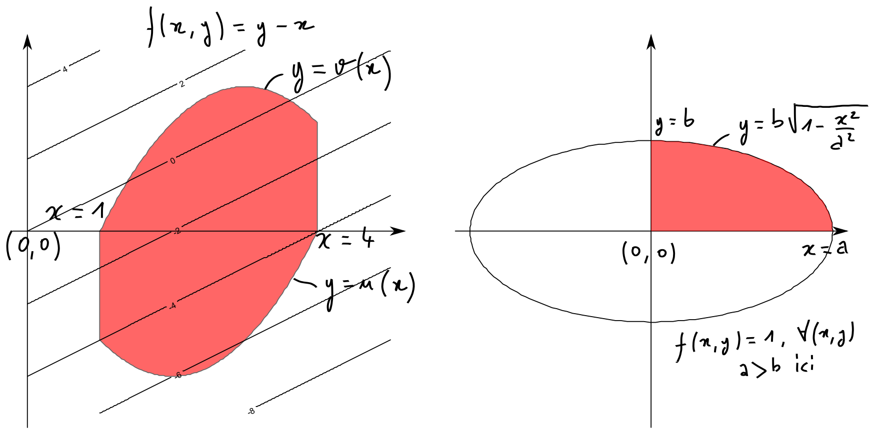

FIGURE 1.3 – Le volume délimité par l'ellipse et la fonction $f(x,y) = 1$ correspond à quatre fois la surface rouge.

La suite du calcul de cette intégrale nécessite un changement de variables. Posons $x = a\sin u$, on a $\mathrm{d}x = a\cos u\,\mathrm{d}u$ et pour les bornes : quand $x \to 0$ alors $u \to 0$ et quand $x \to a$ alors $u \to \pi/2$. La surface de l'ellipse se déduit du calcul suivant

$$\begin{aligned} I =& 4ab\int_0^{\frac{\pi}{2}} \sqrt{1 - \sin^2 u}\cos u\,\mathrm{d}u = 4ab\int_0^{\frac{\pi}{2}} \cos^2 u\,\mathrm{d}u \\ =& 4ab\int_0^{\frac{\pi}{2}} \frac{1 + \cos 2u}{2}\,\mathrm{d}u = 2ab\left[u + \frac{1}{2}\sin 2u\right]_0^{\frac{\pi}{2}} = ab\pi. \end{aligned}$$

Comme dans l'exemple précédent, on aurait pu calculer l'intégrale en notant que le domaine était délimité par les fonctions $y = b\sqrt{1 - x^2/a^2}$ d'une part et $y = -b\sqrt{1 - x^2/a^2}$, ce qui aurait permis d'écrire

$$I = \int_{-a}^a \left[\int_{-b\sqrt{1-x^2/a^2}}^{b\sqrt{1-x^2/a^2}} \mathrm{d}y\right]\mathrm{d}x = \int_{-a}^a \left(-2b\sqrt{1 - \frac{x^2}{a^2}}\right)\mathrm{d}x = 2ab\left[u + \frac{1}{2}\sin 2u\right]_{-\frac{\pi}{2}}^{\frac{\pi}{2}} = ab\pi,$$

ce qui conduit au même résultat.
■

Les résultats précédents ont été obtenu grâce au théorème suivant :

Théorème 2:
Soient $[a,b],\ (a<b)$ un intervalle fermé borné de $\mathbb{R}$, u et v deux fonctions continues sur $[a,b]$ telles que $\forall x\in[a,b],\ u(x)\leq v(x)$. Soit $\mathcal{D}$ le domaine de $\mathbb{R}^2$ défini par

$$\mathcal{D}=\left\{(x,y)\in\mathbb{R}^2 \text{ t.q } a\leq x\leq b, u(x)\leq y\leq v(x)\right\}$$

Si $f:\mathcal{D}\to\mathbb{R}$ est une fonction continue, alors f est intégrable sur $\mathcal{D}$ et on a

$$\int_{\mathcal{D}} f(x,y)\,\mathrm{d}x\,\mathrm{d}y=\int_a^b\left[\int_{u(x)}^{v(x)} f(x,y)\,\mathrm{d}y\right]\mathrm{d}x.$$

1.2.3 Quelques propriétés à retenir

1. Soit $h(x,y)=f(x)g(y)$ une fonction de deux variables qui s'exprime comme un produit de deux fonctions f et g à une variable. On souhaite intégrer h sur le pavé $[a,b]\times[c,d]$ de $\mathbb{R}^2$. Dans ce cas,
$$\int_c^d\int_a^b h(x,y)\,\mathrm{d}x\,\mathrm{d}y=\int_a^b f(x)\,\mathrm{d}x\times\int_c^d g(y)\,\mathrm{d}y.$$
2. Soit f et g deux fonctions continues à deux variables, alors
$$\iint_{\mathcal{D}}\lambda f(x,y)+\gamma g(x,y)\,\mathrm{d}x\,\mathrm{d}y=\lambda\iint_{\mathcal{D}} f(x,y)\,\mathrm{d}x\,\mathrm{d}y+\gamma\iint_{\mathcal{D}} g(x,y)\,\mathrm{d}x\,\mathrm{d}y.$$
3. Soit $\mathcal{D}$ un domaine fermé borné de $\mathbb{R}^2$. On suppose que $\mathcal{D}=\mathcal{D}_1\cup\mathcal{D}_2$ est la réunion de deux domaines fermés $\mathcal{D}_1$ et $\mathcal{D}_2$ et que $\mathcal{D}_1$ et $\mathcal{D}_2$ ont une intersection vide.
Dans ce cas, si $f:\mathcal{D}\to\mathbb{R}$ est une fonction continue sur $\mathcal{D}$, alors f est intégrable sur $\mathcal{D}$ et on a
$$\iint_{\mathcal{D}} f(x,y)\,\mathrm{d}x\,\mathrm{d}y=\iint_{\mathcal{D}_1} f(x,y)\,\mathrm{d}x\,\mathrm{d}y+\iint_{\mathcal{D}_2} f(x,y)\,\mathrm{d}x\,\mathrm{d}y$$

1.3 Changement de variables

Il arrive quelquefois qu'il faille changer le système de coordonnées initiales pour intégrer plus facilement. C'est le cas notamment avec des domaines qui sont circulaires. Nous allons considérer le changement de coordonnées cartésiennes en coordonnées polaires :

$$\left\{\begin{array}{lcl}\rho&=&\sqrt{x^2+y^2}\\ \theta&=&\arctan\left(\frac{y}{x}\right)\end{array}\right.\Rightarrow\left\{\begin{array}{lcl}x&=&\rho\cos\theta\\ y&=&\rho\sin\theta\end{array}\right. .$$

Un point M du plan n'est plus ici représenté par le couple de coordonnées (x,y) mais par le couple (ρ,θ) où ρ représente la distance du point considéré à l'origine O du repère et θ l'angle orienté entre l'axe des abcisses et le vecteur **OM** (Fig. 1.4).

Que se passe t-il lorqu'on doit intégrer une fonction dans ces conditions ?

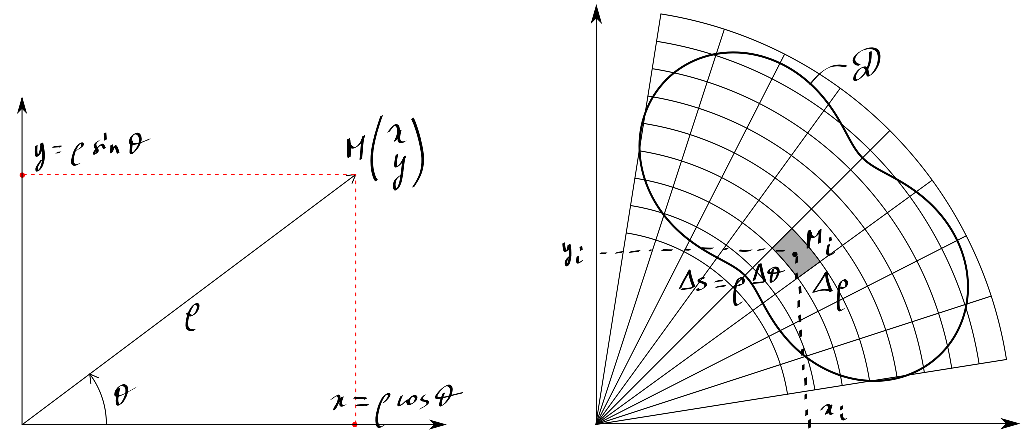

FIGURE 1.4 – Relation entre coordonnées cartésiennes et coordonnées polaires d'un point M du plan. Pavage du plan en coordonnées polaires. La surface du pavet gris est approximée par $\rho\Delta\theta\Delta\rho$.

1.3.1 Un point de vue empirique

Revenons sur la construction initiale de l'intégrale. Maintenant le domaine n'est plus partagé en petits rectangles mais en secteurs de côtés égaux à Δs et $\Delta\rho$ où Δs représente la longueur de l'arc qui délimite l'élément de surface et $\Delta\rho$ la longueur du segment entre deux arcs successifs (Fig. 1.4). Si pour un secteur donné $\Delta\rho$ reste invariant, la longueur Δs évolue en fonction de l'éloignement à l'origine. On va cependant considérer que pour de petites valeurs de Δs, la surface du secteur est presque celle d'un rectangle, égale à $\Delta s \times \Delta\rho$. Cette surface s'exprime également en fonction de θ et devient $\rho\Delta\theta \times \Delta\rho$. On utilise ici la relation $S = R\theta$: la longueur d'arc est égale au rayon R du cercle multipliée par l'angle θ parcouru (en radian). Si on parcourt tout le cercle, on obtient son périmètre $2\pi R$. En prenant un point quelconque M_i de coordonnées $(x_i = \rho_i \cos\theta_i, y_i = \rho_i \sin\theta_i)$ d'un secteur ayant une intersection non nulle avec le domaine, une valeur approchée de l'intégrale a maintenant pour expression :

$$\widehat{I_{\mathcal{D}}} = \sum_{M_i \in \mathcal{D}} f(\rho_i \cos\theta_i, \rho_i \sin\theta_i)\rho\Delta\rho\Delta\theta.$$

En prenant des secteurs de surface de plus en plus petite, on obtient :

$$I_{\mathcal{D}} = \iint_{\mathcal{D}} f(\rho\cos\theta, \rho\sin\theta)\rho \,\mathrm{d}\rho \,\mathrm{d}\theta.$$

On voit que l'élément différentiel $\mathrm{d}x\,\mathrm{d}y$ est devenu $\rho\,\mathrm{d}\rho\,\mathrm{d}\theta$ après changement de variable. On fera également attention au fait que le domaine $\mathcal{D}$ doit aussi s'exprimer en fonction de (ρ, θ).

Exemple 5. On veut calculer l'intégrale

$$I_{\mathcal{D}} = \iint_{\mathcal{D}} xy \,\mathrm{d}x \,\mathrm{d}y,$$

sur le domaine $\mathcal{D} = \{(x, y) \text{ t. q. } x \geq 0, y \geq 0, x^2 + y^2 \leq 1\}$. Ce domaine correspond au quart de disque sur le quadrant positif. En coordonnées polaires, il est équivalent à

$$\mathcal{D} = \left\{(\rho, \theta) \in [0, +\infty[\times[0, 2\pi[\text{ t. q. } \rho \leq 1, 0 \leq \theta \leq \frac{\pi}{2}\right\}$$

et l'intégrale devient

$$I_{\mathcal{D}} = \iint_{\mathcal{D}} \rho^2 \cos\theta \sin\theta \times \rho \,\mathrm{d}\rho \,\mathrm{d}\theta = \int_0^{\frac{\pi}{2}} \cos\theta \sin\theta \,\mathrm{d}\theta \int_0^1 \rho^3 \,\mathrm{d}\rho = \left[\frac{1}{2}\sin^2 x\right]_0^{\frac{\pi}{2}} \left[\frac{1}{4}\rho^4\right]_0^1 = \frac{1}{8}.$$

L'intégrale double peut se partager en produits d'intégrales simples car les variables ne sont pas liées entres elles, à la fois dans l'expression de la fonction mais également entre les bornes du domaine.
■

Exemple 6. Considérons la fonction $f(x, y) = \sqrt{1 - x^2 - y^2}$ que l'on veut intégrer sur son domaine de définition $\mathcal{D} = \{(x, y) \text{ t. q. } x^2 + y^2 \leq 1\}$. Il s'agit de l'équation de la demi-sphère sur son quadrant positif. L'intégrale double

$$I_{\mathcal{D}} = \iint_{\mathcal{D}} \sqrt{1 - x^2 - y^2} \,\mathrm{d}x \,\mathrm{d}y$$

est compliquée à calculer dans ce cas. Procédons au changement de variable précédent. Puisque $x^2 + y^2 = \rho^2$, le domaine de définition devient $\mathcal{D} = \{(\rho, \theta) \text{ t. q. } 0 \leq \rho \leq 1, 0 \leq \theta \leq 2\pi\}$. L'intégrale prend alors la forme :

$$I_{\mathcal{D}} = \iint_{\mathcal{D}} \sqrt{1 - \rho^2} \rho \,\mathrm{d}\rho \,\mathrm{d}\theta.$$

Comme ρ et θ ne sont pas liées dans l'équation, l'intégrale devient :

$$I_{\mathcal{D}} = \int_0^{2\pi} \mathrm{d}\theta \int_0^1 \sqrt{1 - \rho^2} \rho \,\mathrm{d}\rho = 2\pi \times \int_0^1 \sqrt{1 - \rho^2} \rho \,\mathrm{d}\rho.$$

En notant que $\left(\sqrt{1 - \rho^2}^3\right)' = -3\rho\sqrt{1 - \rho^2}$, on obtient :

$$I_{\mathcal{D}} = 2\pi \times \int_0^1 \sqrt{1 - \rho^2} \rho \,\mathrm{d}\rho = 2\pi \times \left[-\frac{1}{3}\sqrt{1 - \rho^2}^3\right]_0^1 = \frac{2}{3}\pi.$$

On retrouve bien la moitié du volume de la sphère $\left(\frac{4}{3}\pi R^3\right)$ de rayon 1. Il faut noter dans ce calcul que les bornes d'intégration ont également changé lors du changement de variables car quand $y \to 0$ alors $\rho \to 1$ et quand $y \to 1$ alors $\rho \to 1$.
■

En résumé, lorsqu'on effectue un changement de variable dans une intégrale, il faut veiller à trois choses :

1. Les bornes changent,
2. l'élément différentiel change,
3. les variables de la fonction changent également.

1.4 Quel rapport avec les statistiques ?

De la même manière que dans le cas univarié, toute fonction continue f de plusieurs variables, positive sur son domaine de définition et vérifiant :

$$\int \cdots \int_{\mathbb{R}^n} f(x_1, \cdots, x_n) \,\mathrm{d}x_1 \cdots \mathrm{d}x_n = 1,$$

est une densité de probabilité pour le n-uplet $\mathbf{V} = (X_1, \cdots, X_n)' \in \mathbb{R}^n$. Dans le cas bivarié, la probabilité qu'un couple (X, Y) de variables aléatoires prenne ses valeurs dans un domaine $\mathcal{D}$ sera donnée par :

$$P\left[(X,Y) \in \mathcal{D}\right] = \iint_{\mathcal{D}} f(x,y)\,\mathrm{d}x\,\mathrm{d}y.$$

Il faudra dans ce cas calculer une intégrale double.

Exemple 7. Intégrale de la loi normale. Nous l'avons déjà vu en cours, une variable aléatoire qui suit une loi normale centrée réduite a pour densité

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2},\ x \in \mathbb{R}.$$

On sait également que $P(a \leq X \leq b) = F_X(b) - F_X(a)$ ne peut pas être évaluée analytiquement car la fonction de répartition F_X n'a pas d'expression algébrique. C'est pour cette raison que l'on dispose de tables pour évaluer les probabilités qui font intervenir la loi normale.

Nous allons montrer dans cet exercice comment le calcul d'une l'intégrale double permet de connaître analytiquement la valeur de

$$I = \int_{-\infty}^{+\infty} e^{-u^2}\,\mathrm{d}u.$$

Nous allons considérer trois domaines différents : $\mathcal{D}_{\square}$, le carré de côté $2a$ centré sur 0, $\mathcal{D}_a$, le cercle centré de rayon a et $\mathcal{D}_{\sqrt{2}a}$, le cercle centré de rayon $\sqrt{2}a$. Les surfaces des trois domaines sont respectivement égales à $4a^2$, πa^2 et $2\pi a^2$. Calculons l'intégrale de la fonction

$$g(u,v) = e^{-u^2-v^2},\ (u,v) \in \mathbb{R}^2.$$

La fonction g étant positive sur $\mathbb{R}^2$, on peut d'ores et déjà en déduire que

$$\iint_{\mathcal{D}_a} g(u,v)\,\mathrm{d}u\,\mathrm{d}v < \iint_{\mathcal{D}_{\square}} g(u,v)\,\mathrm{d}u\,\mathrm{d}v < \iint_{\mathcal{D}_{\sqrt{2}a}} g(u,v)\,\mathrm{d}u\,\mathrm{d}v. \tag{1.1}$$

Le passage en coordonnées polaires donne

$$I_{\mathcal{D}_a} = \iint_{\mathcal{D}_a} e^{-u^2-v^2}\,\mathrm{d}u\,\mathrm{d}v = \int_0^{2\pi}\int_0^a e^{-\rho^2}\rho\,\mathrm{d}\rho\,\mathrm{d}\theta = 2\pi\left[-\frac{1}{2}e^{-\rho^2}\right]_0^a = \pi\left(1 - e^{-a^2}\right).$$

On en déduit aisément que

$$I_{\mathcal{D}_{\sqrt{2}a}} = \pi\left(1 - e^{-2a^2}\right).$$

On calcule également

$$I_{\mathcal{D}_{\square}} = \int_{-a}^{a}\int_{-a}^{a} e^{-u^2-v^2}\,\mathrm{d}u\,\mathrm{d}v = \int_{-a}^{a} e^{-u^2}\,\mathrm{d}u \times \int_{-a}^{a} e^{-v^2}\,\mathrm{d}v = \left(\int_{-a}^{a} e^{-u^2}\,\mathrm{d}u\right)^2.$$

En utilisant l'équation 1.1, les calculs précédents, on peut borner $I_{\mathcal{D}_{\square}}$

$$\sqrt{\pi\left(1 - e^{-a^2}\right)} < \int_{-a}^{a} e^{-u^2}\,\mathrm{d}u < \sqrt{\pi\left(1 - e^{-2a^2}\right)}.$$

Le passage à la limite et le théorème des Gendarmes

$$\lim_{a\to+\infty}\sqrt{\pi\left(1-e^{-a^2}\right)} < \lim_{a\to+\infty}\int_{-a}^{a} e^{-u^2}\,\mathrm{d}u < \lim_{a\to+\infty}\sqrt{\pi\left(1-e^{-2a^2}\right)},$$

nous permet d'en déduire que

$$I=\int_{-\infty}^{+\infty} e^{-u^2}\,\mathrm{d}u=\sqrt{\pi}.$$

Si l'on pose $u=\frac{1}{\sqrt{2}}x$ alors $\mathrm{d}u=\frac{1}{\sqrt{2}}\,\mathrm{d}x$ et l'on a

$$I=\int_{-\infty}^{+\infty} e^{-\frac{1}{2}x^2}\times\frac{1}{\sqrt{2}}\,\mathrm{d}x=\sqrt{\pi}.$$

En divisant par $\sqrt{\pi}$ les deux termes de l'équation, nous déduisons enfin que

$$\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}\,\mathrm{d}x=1.$$

On reconnait ici l'intégrale de la gaussienne égale à 1. Nous venons de montrer d'où vient le terme $\frac{1}{\sqrt{2\pi}}$ par le calcul d'une intégrale double.

■

Exemple 8. Considérons la fonction $f(x,y)=a(1-x^2-y^2)$ sur le domaine $\mathcal{D}=\{(x,y)\in\mathbb{R}^2 \text{ t.q. } x^2+y^2\leq 1\}$ (disque de rayon 1 centré sur O), 0 sinon. On veut calculer la valeur a pour que ce soit une densité. On sait déjà que $\forall\ (x,y),\ f(x,y)\geq 0$ si $a\geq 0$. On veut ensuite trouver a tel que

$$\iint_{\mathcal{D}} a(1-x^2-y^2)\,\mathrm{d}x\,\mathrm{d}y=1.$$

En passant en coordonnées polaires, le domaine d'intégration devient

$$\mathcal{D}=\{(\rho,\theta)\in\mathbb{R}\times[0;2\pi]\ \text{ t.q. } 0\leq\rho\leq 1, 0\leq\theta\leq 2\pi\} \tag{1.2}$$

et on obtient :

$$\int_0^{2\pi}\int_0^1 a(1-\rho^2)\rho\,\mathrm{d}\rho\,\mathrm{d}\theta=2\pi\times a\int_0^1\left(\rho-\rho^3\right)\mathrm{d}\rho=1.$$

Finalement :

$$2\pi\times a\times\left[\frac{1}{2}\rho^2-\frac{1}{4}\rho^4\right]_0^1=1 \iff \boxed{a=\frac{2}{\pi}}.$$

■

Exemple 9. Soit un couple de variables aléatoires (X,Y) de densité conjointe donnée par

$$g(x,y)=\begin{cases}\exp(-x) & \text{si} \quad (x,y)\in\mathcal{D}\\ 0 & \text{sinon}\end{cases},$$

où $\mathcal{D}=\{(x,y)\in\mathbb{R}^2 \text{ t.q. } 0\leq y\leq x\}$. On va vérifier que g est bien une densité. D'une part, $g(x,y)\geq 0,\ \forall(x,y)\in\mathbb{R}^2$. D'autre part,

$$\iint_{\mathbb{R}^2} g(x,y)\,\mathrm{d}x\,\mathrm{d}y=\int_0^{+\infty}\left[\int_0^x e^{-x}\,\mathrm{d}y\right]\mathrm{d}x=\int_0^{+\infty} xe^{-x}\,\mathrm{d}x=\left[-xe^{-x}-e^{-x}\right]_0^1=1.$$

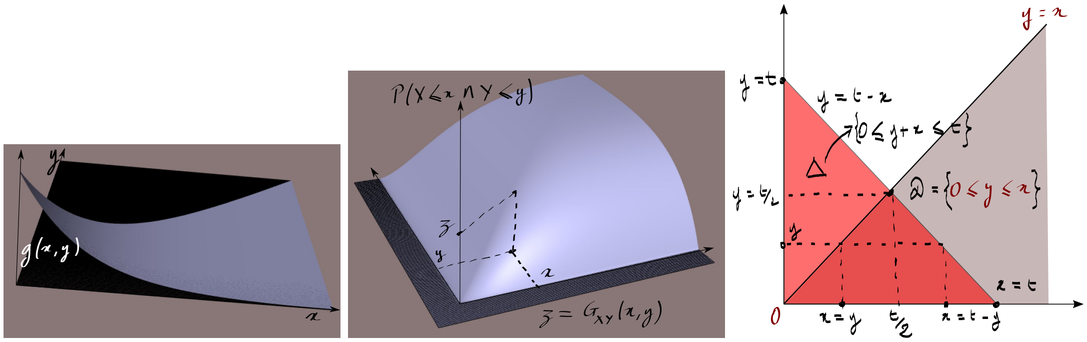

FIGURE 1.5 – Densité g sur $\mathcal{D}$, fonction de répartition G_{XY} et événement $\mathcal{D} \cap \Delta$

Calculons maintenant la probabilité de l'évènement $(X,Y) \in \Delta$ où $\Delta = \{0 \leq x + y \leq t\}\,,\ t \leq 0$. On a :

$$P[(X,Y) \in \Delta] = \iint_{\Delta} g(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_0^{t/2} \left[\int_y^{t-y} e^{-x}\,\mathrm{d}x \right] \mathrm{d}y.$$

La difficulté consiste ici à déterminer les bornes de l'intégrale pour faire le calcul. Le domaine d'intégration sur lequel $g(x,y) = \exp(-x)$ est donné par $\mathcal{D} \cap \Delta$ et correspond au triangle rouge foncé, intersection entre $\mathcal{D}$ (le triangle gris) et Δ (le triangle rouge clair)(Fig. 1.5). La figure montre que lorsque y varie de 0 à $t/2$ (la hauteur du triangle rouge), alors x varie de y à $t-y$. C'est ce qui permet d'établir l'expression de l'intégrale précédente. Le calcul donne :

$$P[(X,Y) \in \Delta] = \int_0^{t/2} \left[\int_y^{t-y} e^{-x}\,\mathrm{d}x \right] \mathrm{d}y = \int_0^{t/2} \left[-e^{-t+y} + e^{-y} \right] \mathrm{d}y = \left[-e^{-t+y} - e^{-y} \right]_0^{t/2} = 1 - 2e^{-t/2} + e^{-t}.$$

On peut également calculer la fonction de répartition du couple (X,Y) donnée par

$$G_{XY}(u,v) = P\left[X \leq u \cap Y \leq v \right] = \int_{-\infty}^{u} \int_{-\infty}^{v} g(x,y)\,\mathrm{d}y\,\mathrm{d}x,$$

soit

$$G_{XY}(u,v) = \begin{cases} 0 & \text{si } u \leq 0 \text{ ou } v \leq 0 \\ \iint_C e^{-x}\,\mathrm{d}y\,\mathrm{d}x & \text{si } 0 \leq u \leq v \\ \iint_A e^{-x}\,\mathrm{d}y\,\mathrm{d}x + \iint_B e^{-x}\,\mathrm{d}y\,\mathrm{d}x & \text{si } 0 \leq v \leq u \end{cases}.$$

L'expression de G dépend de la forme du domaine $\{(X \leq u \cap Y \leq v) \cap \mathcal{D}\}$. Ainsi, on obtient que

$$\iint_C e^{-x}\,\mathrm{d}y\,\mathrm{d}x = \int_0^u \int_0^x e^{-x}\,\mathrm{d}y\,\mathrm{d}x = \int_0^u \left[x e^{-x} \right] \mathrm{d}x = \left[-x e^{-x} - e^{-x} \right]_0^u = 1 - u e^{-u} - e^{-u},$$

et la seconde équation donne aussi que

$$\iint_A e^{-x}\,\mathrm{d}y\,\mathrm{d}x = \int_0^v \int_0^x e^{-x}\,\mathrm{d}y\,\mathrm{d}x = \left[-x e^{-x} - e^{-x} \right]_0^v = 1 - v e^{-v} - e^{-v},$$

$$\iint_B e^{-x}\,\mathrm{d}y\,\mathrm{d}x = \int_v^u \int_0^v e^{-x}\,\mathrm{d}y\,\mathrm{d}x = \left[-v e^{-x} \right]_v^u = -v e^{-u} + v e^{-v}$$

En additionnant ces deux dernières équations, on obtient finalement :

$$G_{XY}(u,v) = \begin{cases} 0 & \text{si } u \leq 0 \text{ ou } v \leq 0 \\ 1 - u e^{-u} - e^{-u} & \text{si } 0 \leq u \leq v \\ 1 - v e^{-u} - e^{-v} & \text{si } 0 \leq v \leq u \end{cases}.$$

■

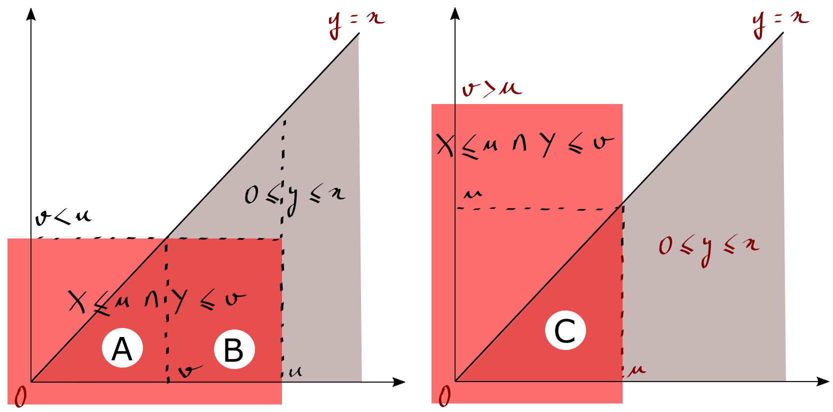

FIGURE 1.6 – La zone grise correspond à la surface sur laquelle la densité n'est pas nulle. La zone rouge est l'événement $\{X \leq u \,\cap\, Y \leq v\}$. Le calcul de la fonction de répartition G_{XY} dépend des valeurs de u et v. Lorsque $0 \leq u \leq v$, il faut calculer la valeur de la surface triangulaire notée C. Dans le cas $0 \leq v \leq u$, le calcul de l'intégrale est celui de la somme de la surface triangulaire A et du rectangle B.

1.4.1 Loi marginale

On peut toujours à partir de la loi conjointe de (X, Y) en déduire la loi de la variable X et celle de Y.

Variables continues

Soit g_{XY} la densité de probabilité d'un couple (X, Y) de variables aléatoires réelles. La fonction de répartition de la variable X est donnée par

$$F_X(t) = P(X \leq t) = P(X \leq t \cap Y \in \mathbb{R}) = \iint_{\mathcal{D}} g_{XY}(x, y)\,\mathrm{d}x\,\mathrm{d}y,$$

car les deux événements sont équivalents. Le domaine d'intégration est donc défini par l'ensemble

$$\mathcal{D} = \left\{(x, y) \in \mathbb{R}^2 \text{ t.q. } x \leq t, -\infty \leq y \leq +\infty\right\}.$$

En croisant la définition d'une fonction de répartition et l'expression ci-dessus, on peut écrire

$$F_X(t) = P(X \leq t) = \int_{-\infty}^{t} f_X(x)\,\mathrm{d}x = \int_{-\infty}^{t} \left[\int_{-\infty}^{\infty} g_{XY}(x, y)\,\mathrm{d}y\right] \mathrm{d}x,$$

et on en déduit le lien entre g_{XY} et f_X avec

$$f_X(x) = \int_{-\infty}^{\infty} g_{XY}(x, y)\,\mathrm{d}y.$$

De la même manière, la loi marginale de Y est donnée par

$$f_Y(y) = \int_{-\infty}^{\infty} g_{XY}(x, y)\,\mathrm{d}x.$$

Exemple 10. (Suite de l'exemple 8) On va calculer les lois marginales du couple (X, Y) de densité g. On a

$$f_X(x) = \int_{-\infty}^{+\infty} g(x, y)dy = \int_{0}^{x} e^{-x}dy = \left[ye^{-x}\right]_0^x = xe^{-x},\; x \geq 0,\; 0 \text{ sinon.}$$

Les bornes de l'intégrale ont ici changé en fonction de l'expression de g. En effet, por une valeur fixée de x, si $y \in]-\infty, 0]$ alors $g(x,y) = 0$, si $y \in [0, x]$ alors $g(x,y) = e^{-x}$ et si $y \in [x, +\infty[$ alors $g(x,y) = 0$.

De la même manière, la densité de Y est donnée par

$$f_Y(y) = \int_{-\infty}^{+\infty} g(x,y)dx = \int_{y}^{+\infty} e^{-x}dx = \left[-e^{-x}\right]_y^{+\infty} = e^{-y},\ y \geq 0,\ 0 \text{ sinon.}$$

On vérifiera ici que les deux densités marginales trouvées sont bien des densités (ce sont des fonctions positives et leur intégrale est égale à un sur $\mathbb{R}^+$).

■

Exemple 11. Soit un couple de variables aléatoires (X,Y) de densité conjointe

$$g_{XY}(x,y) = \begin{cases} kxy & \text{si} \quad (x,y) \in \mathcal{D} \\ 0 & \text{sinon} \end{cases},$$

where $\mathcal{D} = \{(x,y) \in \mathbb{R}^2 \text{ t. q. } x \geq 0; y \geq 0, x^2 + y^2 \leq a^2, a > 0\}$. Déterminons la valeur de k pour que g_{XY} soit bien une densité. D'une part, nous savons que pour que g_{XY} soit positive, k doit être positif puisque x et y le sont. D'autre part, on doit vérifier que

$$\iint_{\mathcal{D}} g(x,y)\,\mathrm{d}x\,\mathrm{d}y = 1.$$

Nous sommes sur un domaine circulaire, on va changer de variables. En coordonnées polaire, le domaine devient $\mathcal{D} = \left\{(\rho,\theta) \in \mathbb{R}^+ \times [0, \frac{\pi}{2}] \text{ t. q.} \rho \leq a\right\}$. L'intégrale précédente s'écrit alors de la façon suivante :

$$\iint_{\mathcal{D}} g(\rho\cos\theta, \rho\sin\theta)\rho\,\mathrm{d}\rho\,\mathrm{d}\theta = \int_0^{\frac{\pi}{2}} \left[\int_0^a k\rho\cos\theta \times \rho\sin\theta \times \rho\,\mathrm{d}\rho\right] \mathrm{d}\theta = k\int_0^{\frac{\pi}{2}} \cos\theta\sin\theta\,\mathrm{d}\theta \int_0^a \rho^3\,\mathrm{d}\rho = 1.$$

On en déduit que :

$$k\int_0^{\frac{\pi}{2}} \cos\theta\sin\theta\,\mathrm{d}\theta \int_0^a \rho^3\,\mathrm{d}\rho = k \times \left[\frac{1}{2}\sin^2 x\right]_0^{\frac{\pi}{2}} \times \left[\frac{1}{4}\rho^4\right]_0^a = k \times \frac{a^4}{8} = 1 \implies \boxed{k = \frac{8}{a^4}}.$$

■

La loi marginale de X peut ensuite être calculée avec

$$f_X(x) = \int_{-\infty}^{\infty} g(x,y)\,\mathrm{d}y = \frac{8}{a^4}\int_0^{\sqrt{a^2-x^2}} xy\,\mathrm{d}y = \frac{4}{a^4}x(a^2 - x^2) \text{ si } 0 \leq x \leq a,\ 0 \text{ sinon.}$$

La loi marginale de Y est de la même forme, du fait du rôle symétrique joué par x et y soit

$$f_Y(y) = \frac{4}{a^4}y(a^2 - y^2) \text{ si } 0 \leq y \leq a,\ 0 \text{ sinon.}$$

1.4.2 Indépendance de variables aléatoires

Définition 1:
Soit (X, Y) un couple de variables aléatoires réelles. On dit que les deux variables X et Y sont indépendantes si, pour tout couple $I \times J$ d'intervalles de $\mathbb{R}^2$, on a la propriété

$$P(X \in I \cap Y \in J) = P(X \in I)P(Y \in J). \tag{1.3}$$

Dans tous les autres cas, on dit que les variables X et Y sont dépendantes, ou liées, ou corrélées.

Variables réelles

Soit (X, Y) un couple de variables aléatoires réelles et de densité g_{XY}. On désigne par f_X la densité de X et par f_Y la densité de Y. Les variables aléatoires X et Y sont indépendantes si et seulement si la desnité conjonte est gale au produit des lois marginales soit :

$$g_{XY}(x, y) = f_X(x)f_Y(y). \tag{1.4}$$

En utilisant la définition, l'indépendance de X et de Y implique en particulier que

$$\forall (u, v) \in \mathbb{R}^2 : P(X \leq u \cap Y \leq v) = P(X \leq u)P(Y \leq v),$$

ce qui implique que la propriété précédente est également vérifiée entre les fonctions de répartition

$$\forall (u, v) \in \mathbb{R}^2 : F_{XY}(u, v) = F_X(u)F_Y(v).$$

■

Exemple 12. Soit $f_{XY}(x, y)$ la loi uniforme sur le domaine $\mathcal{D} = [-1, 1]^2$, carré centré de côté 2. Calculons les lois marginales. On a

$$f_X(x) = \int_{\mathbb{R}} f(x, y)\,\mathrm{d}y = \int_{-1}^{1} \frac{1}{4}\,\mathrm{d}y = \frac{1}{2},\ -1 \leq x \leq 1,\ 0 \text{ sinon.}$$

Par symétrie de la loi uniforme, on déduit la loi marginale de Y

$$f_Y(y) = \frac{1}{2}, -1 \leq y \leq 1,\ 0 \text{ sinon.}$$

On peut également en déduire les fonctions de répartition :

$$F_X(x) = \left\{ \begin{array}{cc} 0 & \text{si } x \leq -1 \\ \left(\frac{1}{2}x + \frac{1}{2}\right) & \text{si } -1 \leq x \leq 1 \\ 1 & \text{si } x \geq 1 \end{array} \right. ,$$

même expression pout F_Y. D'autre part, on vérifie aisément que

$$f_{XY}(x, y) = \frac{1}{4} = f_X(x) \times f_Y(y) = \frac{1}{2} \times \frac{1}{2},\ -1 \leq x \leq 1, -1 \leq y \leq 1,\ 0 \text{ sinon.}$$

Les deux variables aléatoires sont donc indépendantes. Calculons la fonction de répartition pour $-1 \leq u \leq v \leq 1$

$$F_{XY}(u,v) = \int_{-\infty}^{u}\int_{-\infty}^{v} f_{XY}(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_{-1}^{u}\frac{1}{2}\,\mathrm{d}x\int_{-1}^{v}\frac{1}{2}\,\mathrm{d}y = \left[\frac{1}{2}x\right]_{-1}^{u}\times\left[\frac{1}{2}y\right]_{-1}^{v} = \left(\frac{1}{2}u+\frac{1}{2}\right)\left(\frac{1}{2}v+\frac{1}{2}\right).$$

On obtient finalement que :

$$F_{XY}(x,y) = \begin{cases} 0 & \text{si } x \leq -1 \text{ ou } y \leq -1 \\ \left(\frac{1}{2}x+\frac{1}{2}\right)\left(\frac{1}{2}y+\frac{1}{2}\right) & \text{si } -1 \leq x \leq y \leq 1 \text{ ou } -1 \leq y \leq x \leq 1 \\ 1 & \text{si } x \geq 1 \text{ et } y \geq 1 \end{cases}.$$

On s'aperçoit également que $F_{XY}(x,y) = F_X(x) \times F_Y(y),\ \forall (x,y) \in \mathbb{R}^2$.
■

1.4.3 Fonctions de plusieurs variables aléatoires

Considérons une suite de N variables aléatoires continues $X_1, \ldots, X_N$ de distribution conjointe $f_{X_1,\ldots,X_N}$. On est souvent amené en pratique à traiter de fonction de ces variables de la forme $X_1 + X_2$, $X_1 - X_2$, X_1X_2 ou encore $\max X_1, \ldots, X_N$. La question qui se pose est ici de déterminer leur loi respective étant données la connaissance de la densité conjointe $f_{X_1,\ldots,X_N}$.

Variable somme

Soit la variable $Z = Y + X$. Déterminons la loi de cette variable en utilisant la fonction conjointe f_{XY} que l'on suppose connue.

Exemple 13. On considère un couple de variables discrètes dont la loi conjointe est donnée par le tableau de contingence :

$Y = y_j$ \ $X = x_i$	1	2	3
1	0	$\frac{1}{6}$	$\frac{1}{6}$
2	$\frac{1}{6}$	0	$\frac{1}{6}$
3	$\frac{1}{6}$	$\frac{1}{6}$	0

Dans ce cas la distribution de la variable $Z = X + Y$ est assez simple à calculer. En premier, on remarque que $Z \in \{2, \ldots, 6\}$: il s'agit des valeurs observables de la somme de X et de Y. Ensuite, nous allons déterminer $P\left(X = z\right)$ en utilisant le tableau de contingence :
$P(Z = 2) = P(X = 1 \cap Y = 1) = 0$
$P(Z = 3) = P([X = 1 \cap Y = 2] \cup [Y = 1 \cap X = 1]) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$
$P(Z = 4) = P([X = 1 \cap Y = 3] \cup [Y = 3 \cap X = 1] \cup [X = 2 \cap Y = 2]) = \frac{1}{6} + 0 = \frac{1}{3}$
$P(Z = 5) = P([X = 2 \cap Y = 3] \cup [Y = 3 \cap X = 2]) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$
$P(Z = 6) = P(X = 3 \cap Y = 3) = 0$
La distribution de Z est alors entièrement déterminée. On peut remarquer que les lois marginales en X et en Y sont identiques, données par $P(X = i) = P(Y = i) = \frac{1}{3},\ i = 1, \ldots, 3$ mais que les valeurs du tableau de contingence ne vérifient pas $P(X = i \cap Y = j) = P(X = i)P(Y = j),\ \forall (i,j) \in$

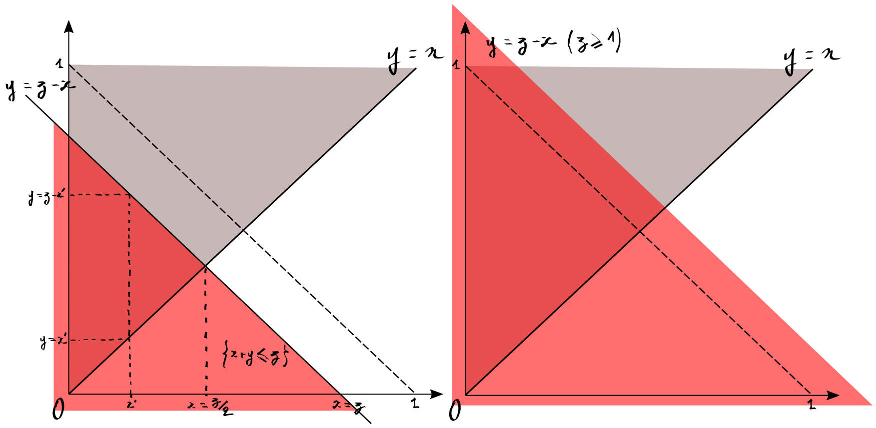

FIGURE 1.7 – Domaines triangulaires d'intégration de $Z = X + Y$. En fonction de la valeur de z, les domaines changent.

$\{1, \ldots, 3\} \times \{1, \ldots, 3\}$. Les variables X et Y ne sont pas indépendantes et la détermination de la loi de Z ne peut pas être simplement déterminée par la seule connaissance de leur loi marginale.
■

Exemple 14. On considère un couple de variables continues X et Y ayant une distribution conjointe uniforme sur le triangle $\{0 \leq x \leq y \leq 1\}$ telle que

$$f_{XY}(x,y) = \begin{cases} 2 & \text{si } 0 \leq x \leq y \leq 1 \\ 0 & \text{, sinon} \end{cases} .$$

Soit $Z = X + Y$, on remarque que $Z \in [0;2]$. Pour $z < 0$, $P(Z \leq z) = 0$ et pour $z > 2$, $P(Z \leq z) = 1$. Pour $0 \leq z \leq 2$, on a

$$F_Z(z) = P(Z \leq z) = P(X + Y \leq z) = \iint_{\substack{0 \leq x \leq y \leq 1 \\ x+y \leq z}} f_{XY}(x,y) \,\mathrm{d}x\,\mathrm{d}y.$$

Il y a deux cas à considérer (Fig. 1.7). Le premier cas où $0 \leq z \leq 1$ mène à l'intégrale

$$F_Z(z) = P(Z \leq z) = P(X + Y \leq z) = \int_{x=0}^{x=z/2} \left[\int_{y=x}^{y=z-x} 2\,\mathrm{d}y\right] \mathrm{d}x = \frac{z^2}{2}.$$

Le second cas où $1 \leq z \leq 2$ mène à la seconde intégrale

$$F_Z(z) = P(Z \leq z) = 1 - P(Z > z) = 1 - \int_{y=z/2}^{1} \left[\int_{x=z-y}^{y} 2\,\mathrm{d}x\right] \mathrm{d}y = 1 - \frac{(z-2)^2}{2}.$$

Finalement, en dérivant les différentes expression de F_Z, la densité est donnée par :

$$f_Z(z) = \begin{cases} z & \text{si } 0 \leq z \leq 1, \\ 2 - z & \text{si } 1 \leq z \leq 2, \\ 0 & \text{sinon.} \end{cases}$$

■

1.4.4 Méthodes des transformations

La détermination de la densité d'une somme, d'une différence peut être déterminée à l'aide d'une autre méthode. Considérons par exemple $Z = X + Y$ une variable aléatoire dont on souhaite déterminer la densité. Supposons f_{XY} la densité du couple (X, Y) définie sur $\mathbb{R}^2$. Soit la transformation $(U, V) = \mathbf{\Phi}(X, Y)$ et son inverse $(X, Y) = \mathbf{\Phi}^{-1}(U, V)$ telles que :

$$\left\{\begin{array}{rcl} U &=& X \\ V &=& X+Y \end{array}\right. \Rightarrow \left\{\begin{array}{rcl} X &=& U \\ Y &=& V-U \end{array}\right. .$$

La transformation $\mathbf{\Phi}$ est linéaire ici de sorte qu'elle peut s'écrire

$$\begin{pmatrix} U \\ V \end{pmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{pmatrix} X \\ Y \end{pmatrix}$$

où

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

est la forme matricielle de l'application $\mathbf{\Phi}$. L'application $\mathbf{\Phi}^{-1}$ est ici donnée par

$$\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} \begin{pmatrix} U \\ V \end{pmatrix},$$

où

$$\mathbf{A}^{-1} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$$

est la matrice jacobienne, inverse de $\mathbf{A}$. On veut déterminer $F_Z(z)$, soit

$$P(Z \leq z) = \iint_{x+y\leq z} f_{XY}(x, y)\, \mathrm{d}x\, \mathrm{d}y = \iint_{v\leq z} f_{XY}(u, v-u)|J|\, \mathrm{d}u\, \mathrm{d}v,$$

où $|J|$ est la valeur absolue du déterminant de A^{-1} soit $|J| = |1 \times 1 - (-1) \times 0| = 1$.

On a finalement

$$\begin{aligned} F_Z(z) &= \iint_{v\leq z} f_{XY}(u, v-u)\, \mathrm{d}u\, \mathrm{d}v \\ &= \int_{v\leq z} \left[\int_{-\infty}^{+\infty} f_{XY}(u, v-u)\, \mathrm{d}u\right] \mathrm{d}v \\ &= \int_{-\infty}^{z} f_Z(v)\, \mathrm{d}v. \end{aligned}$$

On en déduit que la densité de Z est donnée par

$$f_Z(z) = \int_{-\infty}^{+\infty} f_{XY}(u, v-u)\, \mathrm{d}u = \int_{-\infty}^{+\infty} f_{XY}(u-v, v)\, \mathrm{d}v,$$

la seconde expression ayant été obtenue par la même démarche si l'on avait considéré le changement de variables

$$\left\{\begin{array}{rcl} U &=& X+Y \\ V &=& Y \end{array}\right. .$$

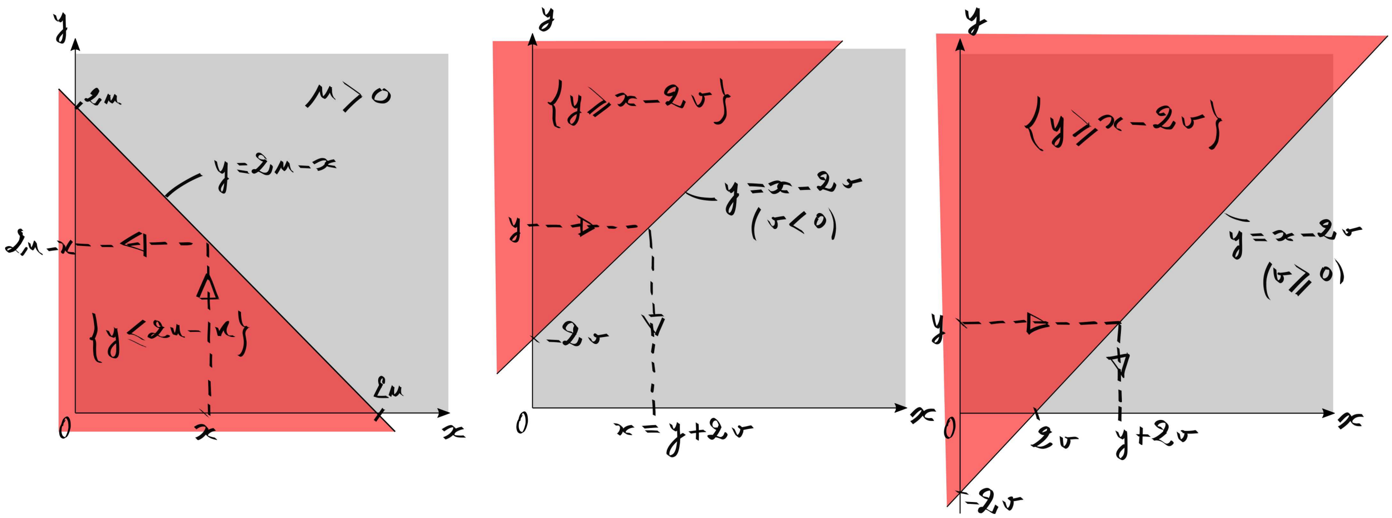

FIGURE 1.8 – Domaines triangulaires d'intégration de $X + Y \leq 2u$ et $X - Y \leq 2v$ (deux cas).

Exemple 15. On considère deux variables aléatoires à valeurs réelles, X et Y , indépendantes et de même loi donnée par la densité de la loi exponentielle d'une variable T soit :

$$f_T(t) = \begin{cases} 0 & \text{si } t < 0 \\ \exp(-t) & \text{si } t \geq 0 \end{cases}.$$

Soit les variables aléatoires $U = \frac{1}{2}(X + Y)$ et $V = \frac{1}{2}(X - Y)$. Calculons espérance et variance de U et de V.
Nous avons

$$E(U) = E(\frac{1}{2}(X + Y)) = \frac{1}{2}E(X) + \frac{1}{2}E(Y) = 1$$

car les deux espérances sont égale à celle de la loi exponentielle de paramètre 1 soit $E(T) = 1$. De la même manière,

$$E(V) = E(\frac{1}{2}(X - Y) = \frac{1}{2}E(X) - \frac{1}{2}E(Y) = 0.$$

La variance de U est donnée par

$$\text{Var}(U) = \text{Var}(\frac{1}{2}(X + Y)) = \frac{1}{4}\text{Var}(X + Y) = \frac{1}{4}(\text{Var}(X) + \text{Var}(Y))$$

car les variables X et Y sont indépendantes. On sait de plus que $\text{Var}(T) = 1$ et on en déduit que $\text{Var}(U) = \frac{1}{2}$. De la même manière, $\text{Var}(V) = \frac{1}{4}\text{Var}(X - Y) = \frac{1}{4}(\text{Var}(X) + \text{Var}(Y))$ et $\text{Var}(V) = \text{Var}(U) = \frac{1}{2}$. Contrairement à l'espérance, les variances s'additionnent également dans le cas d'une différence entre variables aléatoires.
Nous allons maintenant déterminer les densités de probabilité de U et de V de deux manières différentes.
Par méthode directe, on cherche $F_U(u) = P(U \leq u)$ soit

$$P(U \leq u) = P(X + Y \leq 2u) = \iint\limits_{\substack{x \geq 0, y \geq 0 \\ y \leq 2u - x}} f_{XY} \, \mathrm{d}x \, \mathrm{d}y.$$

Le domaine est celui de la figure 1.8, intersection entre la surface grise et la rouge. On en déduit que $F_U(u) = 0$ si $u < 0$ et que, pour $u \geq 0$,

$$
\begin{aligned}
F_U(u) &= \int_0^{2u} \left[\int_0^{2u-x} \exp(-x) \exp(-y) \, \mathrm{d}y \right] \mathrm{d}x \\
&= \int_0^{2u} \exp(-x) \left[-\exp(-y) \right]_0^{2u-x} \mathrm{d}x \\
&= \int_0^{2u} \exp(-x) \left(1 - \exp(-2u + x) \right) \mathrm{d}x \\
&= \left[-\exp(-x) - x \exp(-2u) \right]_0^{2u} \\
&= 1 - \exp(-2u) - 2u \exp(-2u).
\end{aligned}
$$

La densité est obtenue en dérivant la fonction de répartition et on obtient :

$$
f_U(u) = \left\{ \begin{array}{cc} 0 & \text{si} \quad u < 0, \\ 4u \exp(-2u) & \text{si} \quad u \geq 0 \end{array} \right. .
$$

Pour la variable V, on cherche $F_V(v) = P(X - Y \leq 2v)$. On remarque ici que la variable $V \in \mathbb{R}$ alors que $U > 0$. On a :

$$
P(V \leq v) = P(X - Y \leq 2v) = \iint\limits_{\substack{x \geq 0, y \geq 0 \\ y \geq x - 2v}} f_{XY} \, \mathrm{d}x \, \mathrm{d}y.
$$

On va distinguer deux cas (Fig. 1.8). Le cas $v \leq 0$ mène à la fonction de répartition :

$$
\begin{aligned}
F_V(v) &= \int_{-2v}^{+\infty} \left[\int_0^{y+2v} \exp(-x) \exp(-y) \, \mathrm{d}x \right] \mathrm{d}y \\
&= \int_{-2v}^{+\infty} \exp(-y) \left[-\exp(-x) \right]_0^{y+2v} \mathrm{d}y \\
&= \int_{-2v}^{+\infty} \exp(-y) \left(1 - \exp(-y - 2v) \right) \mathrm{d}y \\
&= \left[-\exp(-y) + \frac{1}{2} \exp(-2y - 2v) \right]_{-2v}^{+\infty} \\
&= \exp(2v) - \frac{1}{2} \exp(2v) \\
&= \frac{1}{2} \exp(2v),
\end{aligned}
$$

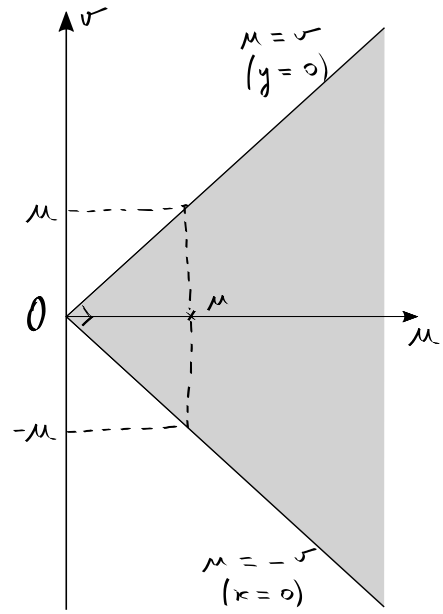

FIGURE 1.9 – Domaine d'intégration dans le repère (u, v).

et le cas $v \geq 0$:

$$\begin{aligned}
F_V(v) &= \int_0^{+\infty} \left[\int_0^{y+2v} \exp(-x)\exp(-y)\,\mathrm{d}x \right] \mathrm{d}y \\
&= \int_0^{+\infty} \exp(-y)\left[-\exp(-x)\right]_0^{y+2v} \mathrm{d}x \\
&= \int_0^{+\infty} \exp(-y)\left(1 - \exp(-y-2v)\right) \mathrm{d}x \\
&= \left[-\exp(-y) + \frac{1}{2}\exp(-2y-2v) \right]_0^{+\infty} \\
&= 1 - \exp(-2v).
\end{aligned}$$

On obtient finalement la densité suivante :

$$f_V(v) = \begin{cases} \exp(2v) & \text{si} \quad v \leq 0, \\ \exp(-2v) & \text{si} \quad v > 0 \end{cases} .$$

On peut également calculer ces densités en utilisant le changement de variables et en remarquant que :

$$\iint_{\mathbb{R}^+ \times \mathbb{R}^+} f_{XY}(x, y)\,\mathrm{d}x\,\mathrm{d}y = \iint_{\mathcal{D}_{UV}} f_{XY}(u+v, u-v) \times 2\,\mathrm{d}u\,\mathrm{d}v.$$

En effet, la transformation inverse nous donne $X = U + V$ et $Y = U - V$ et la valeur absolue de son jacobien est 2. La difficulté est ici de déterminer le domaine d'intégration $\mathcal{D}_{UV}$. Il correspond en fait à une rotation de $-\pi/4$ du quart de plan positif. En fait, V peut être négatif mais $U - V$ doit toujours être positif car $Y > 0$. Le domaine correspond au domaine hachuré de la figure 1.9.

On va d'abord calculer la loi conjointe du couple (U, V), soit

$$f_{UV}(u, v) = 2 \times f_{XY}(u+v, u-v) = 2\exp(-2u),\ \forall (u, v) \in \mathcal{D}_{UV}, 0 \text{ sinon.}$$

On en déduit la densité de U en remarquant que pour une valeur $u > 0$ fixée, alors v varie de $-u$ à u, ce qui donne

$$f_U(u) = \int_{-u}^{u} 2\exp(-2u)\,\mathrm{d}v = [2v\exp(-2u)]_{-u}^{u} = 4u\exp(-2u),\ u > 0, 0 \text{ sinon},$$

et l'on retrouve le résultat précédent. On fait la même chose pour la variable V en distinguant deux cas. Pour une valeur $v > 0$ fixé, u varie de v à $+\infty$, ce qui nous amène à la densité

$$f_V(v) = \int_{v}^{+\infty} 2\exp(-2u)\,\mathrm{d}u = [-\exp(-2u)]_{v}^{+\infty} = \exp(-2v),\ v > 0,$$

et dans le cas où $v \leq 0$, alors

$$f_V(v) = \int_{-v}^{+\infty} 2\exp(-2u)\,\mathrm{d}u = [-\exp(-2u)]_{-v}^{+\infty} = \exp(2v),\ v \leq 0.$$

On retrouve également les résultats obtenus par l'autre méthode.
■