Exercice 1.1 :

On considère une variable continue aléatoire x sur $]-\infty, +\infty[$ suivant une loi de Gauss de paramètre μ et σ.
-Donnez l'expression de la loi de Gauss.
-Donnez les conditions pour que cette loi soit une pdf ainsi que sur les paramètres.
-Quelle est la fonction caractéristique (définition générale et expression dans ce cas).
-Faites de même pour sa fonction cumulative.

Rappel :

$$\text{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z exp(-t^2)dt \tag{1}$$

Correction :

Soit x une variable aléatoire sur $]-\infty, +\infty[$ suivant une loi de Gauss. La fonction de densité s'écrit :

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

Conditions pour que cette loi soit une pdf ainsi que sur les paramètres :

- $f(x) \geq 0$
- $\forall x \int_{-\infty}^{+\infty} f(x)dx = 1$
- $\mu, \sigma \in \mathbb{R}\ et\ \sigma > 0$

Fonction caractéristique :

$$\Phi(t) = \int_{-\infty}^{+\infty} f(x)\, e^{itx} dx \tag{3}$$

$$\Longrightarrow \Phi(t) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} exp\left(-\frac{1}{2}(\frac{x-\mu}{\sigma})^2 + itx \right) dx \tag{4}$$

On fait le changement de variable suivant :

$$x = \sqrt{2}\sigma\, X + \mu \Rightarrow dx = \sqrt{2}\sigma\, dX \tag{5}$$

$$\Longrightarrow \Phi(t) = \frac{1}{\sqrt{\pi}} e^{it\mu} \int_{-\infty}^{+\infty} e^{-X^2}\, e^{it\sqrt{2}\sigma\, X} dX \tag{6}$$

On se sert de l'égalité suivante, elle peut être trouvée sur le Web :

$$\int_{-\infty}^{+\infty} e^{-t^2} e^{zt} dt = \sqrt{\pi} e^{z^2/4} \tag{7}$$

En prenant $z = it\sqrt{2}\sigma$, on obtient :

$$\Phi(t) = exp(it\mu)\, exp(-\frac{t^2\sigma^2}{2}) \tag{8}$$

Fonction cumulative :

$$F(x) = P(X \, \leq \, x) = \int_{-\infty}^{x} f(x')dx' = \int_{-\infty}^{\mu} f(x')dx' + \int_{\mu}^{x} f(x')dx' \tag{9}$$

$$F(x) = \int_{-\infty}^{\mu} \frac{1}{\sqrt{2\pi}\sigma} \, exp\Big(- \frac{1}{2} (\frac{x' - \mu}{\sigma})^2 \Big) dx' + \int_{\mu}^{x} \frac{1}{\sqrt{2\pi}\sigma} \, exp\Big(- \frac{1}{2} (\frac{x' - \mu}{\sigma})^2 \Big) dx' \tag{10}$$

On fait le changement de variable suivant : $x" = x' - \mu$ et l'on obtient :

$$\int_{-\infty}^{0} \frac{1}{\sqrt{2\pi}\sigma} \, exp\Big(- \frac{1}{2} (\frac{x"}{\sigma})^2 \Big) dx" = \frac{1}{2} \tag{11}$$

Pour la deuxième intégrale, on effetue le changement de variable suivant : $v = \dfrac{x"}{\sqrt{2}\sigma}$

ce qui donne :

$$\frac{1}{\sqrt{\pi}} \int_{0}^{\frac{x-\mu}{\sqrt{2}\sigma}} e^{-v^2} dv = \frac{1}{2} \, \mathrm{erf}(\frac{x - \mu}{\sqrt{2}\sigma}) \tag{12}$$

d'où le résultat final :

$$F(x) = \frac{1}{2} + \frac{1}{2} \, \mathrm{erf}(\frac{x - \mu}{\sqrt{2}\sigma}) \tag{13}$$

Exercice 1.2 :

On fait un changement de variable $y = exp(x)$, où x est défini plus haut et suit la loi de Gauss ci-dessus.
-Comment va-t-on calculer la loi de probabilité pour y ?
-Donnez l'expression de la pdf de y.
-Calculez l'espérance mathématique $E(y)$; donnez d'abord sa définition
-Donnez l'expression de la variance de y.

Correction :

On doit trouver la densité de probabilités de y notée g(y). On a :

$$(\ x \in]-\infty, +\infty[\ ) \Rightarrow (\ y = e^x \in]0, +\infty[\ ) \tag{14}$$

Théorème de transfert : en tenant compte que $x = ln\, y$ et qu'il y a conservation des probabilités dans le changement de variables, on a :

$$f(x)dx = g(y)dy \tag{15}$$

ce qui implique :

$$g(y) = f(x(y)) \mid \frac{dx}{dy} \mid \tag{16}$$

On obtient ansi pour $g(y)$:

$$g(y) = \begin{cases} \dfrac{1}{\sqrt{2\pi}\sigma} exp\left(-\dfrac{1}{2}\left(\dfrac{ln\, y - \mu}{\sigma}\right)^2\right)\dfrac{1}{y} & si \quad y > 0 \\ 0 & si \quad y \leq 0 \end{cases} \tag{17}$$

Espérance mathématique :

$$E(y) = \int_{-\infty}^{+\infty} y\, g(y)\, dy = \int_{0}^{+\infty} y\, g(y)\, dy \tag{18}$$

car $g(y) = 0$ pour $y \leq 0$.

En tenant compte du fait que $x = ln\, y$ et que $dy = e^x\, dx$, on a :

$$E(y) = \int_{-\infty}^{+\infty} e^x \frac{1}{\sqrt{2\pi}\sigma} exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right) dx \tag{19}$$

On reconnait l'équation (6) de la fonction caractéristique vue dans l'exercice 1.1 :

$$E(y) = \Phi(it = 1) = e^{\mu + \frac{\sigma^2}{2}} \tag{20}$$

Variance mathématique :

$$V(y) = E(y^2) - E(y)^2 \tag{21}$$

$$E(y^2) = \int_{-\infty}^{+\infty} y^2 g(y)\, dy \tag{22}$$

En suivant le même raisonnnement que précédemment :

$$E(y^2) = \int_{-\infty}^{+\infty} e^{2x} \frac{1}{\sqrt{2\pi}\sigma}\, exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right) dx \tag{23}$$

$$E(y^2) = \Phi(it = 2) \tag{24}$$

D'où le résultat final :

$$V(y) = exp(2\mu + 2\sigma^2) - exp(2\mu + \sigma^2) \tag{25}$$

Exercice 1.3 :

On dispose d'un ensemble de variables $x_i, i = 1, ..., n$, et on considère dans le cas le plus général, la variable aléatoire $z = \sum_{i=1}^{n} a_i x_i$ où les a_i sont des constantes.
-Donnez l'expression de l'espérance mathématique $E(z)$.
-Donnez l'expression de la variance $V(z)$ (cas le plus général).

Correction :

Espérance mathématique $E(z)$:

$$E(z) = E\big(\sum_{i=1}^{n} a_i x_i\big) = \sum_{i=1}^{n} a_i E(x_i) \tag{26}$$

Variance mathématique $V(z)$:

$$V(z) = \sum_{i=1}^{n} a_i^2 V(x_i) + \sum_{i \neq j} a_i a_j cov(x_i, x_j) \tag{27}$$

avec $cov(X, Y) = E[(X - E(X))(Y - E(Y))]$

Exercice 2.1 :

On se réduit à 2 variables aléatoires x et y de variance σ_x^2 et σ_y^2 et de corrélation ρ_{xy}.

On fait le changement de variables suivant :
- $u = x + y\sigma_x/\sigma_y$

- $v = x - y\sigma_x/\sigma_y$

Calculez les variances $V(u)$, $V(v)$ et la covariance $cov(u, v)$. Qu'en concluez-vous ? Quelle est la corrélation ρ_{uv} ?

Correction :

On utilise la relation de la covariance de l'exercice précédent (27) et on a :

$$V(u) = V(x) + (\frac{\sigma_x}{\sigma_y})^2 V(y) + 2\frac{\sigma_x}{\sigma_y} cov(x, y) \tag{28}$$

$$\Rightarrow V(u) = 2\sigma_x^2 + 2\sigma_x^2 \rho_{xy} = 2\sigma_x^2(1 + \rho_{xy}) \tag{29}$$

De la même manière pour $V(v)$:

$$\Rightarrow V(v) = 2\sigma_x^2 - 2\sigma_x^2 \rho_{xy} = 2\sigma_x^2(1 - \rho_{xy}) \tag{30}$$

Rappel : $\rho_{xy} = \dfrac{cov(x, y)}{\sigma_x \sigma_y}$

Faisons maintenant le calcul de la covariance de u et v :

$$cov(u, v) = E((u - \bar{u})(v - \bar{v})) = E((x + y\frac{\sigma_x}{\sigma_y} - \bar{x} - \bar{y}\frac{\sigma_x}{\sigma_y})(x - y\frac{\sigma_x}{\sigma_y} - \bar{x} + \bar{y}\frac{\sigma_x}{\sigma_y})) \tag{31}$$

$$= E((x - \bar{x})^2 - \frac{\sigma_x}{\sigma_y}(y - \bar{y})(x - \bar{x}) + (x - \bar{x})(y - \bar{y})\frac{\sigma_x}{\sigma_y} - \left(\frac{\sigma_x}{\sigma_y}\right)^2 (y - \bar{y})^2) \tag{32}$$

$$= V(x) - \left(\frac{\sigma_x}{\sigma_y}\right)^2 V(y) = 0 \tag{33}$$

$$\Rightarrow \rho_{uv} = 0 \tag{34}$$

u et v ne sont pas corrélées. La conclusion est que l'on peut fabriquer deux variables non corrélées à partir de deux variables corrélées

Exercice 2.2 :

Ces deux variables aléatoires x et y suivent une loi de Gauss (pdf $f(x, y)$) avec $\mu_x = 0$ et $\mu_y = 0$.
* Donnez son expression.
* Donnez les espérances mathématiques $E(x)$ et $E(y)$.

* Calculez la loi de densité de probabilités pour les variables u et v (pdf $h(u,v)$).
* Que concluez-vous pour ces 2 variables aléatoires u et v ?

Correction :

Formule générale de la fonction de densité de la loi normale à deux dimensions :

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \\ exp\left(-\frac{1}{2}\frac{1}{(1-\rho^2)}\left(\left(\frac{x-\mu_x}{\sigma_x}\right)^2 + \left(\frac{y-\mu_y}{\sigma_y}\right)^2 - \frac{2\rho\,(x-\mu_x)(y-\mu_y)}{\sigma_x\,\sigma_y}\right)\right) \tag{35}$$

ici, la fonction de densité s'écrit avec $\mu_x = \mu_y = 0$:

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \\ exp\left(-\frac{1}{2}\frac{1}{(1-\rho^2)}\left(\left(\frac{x}{\sigma_x}\right)^2 + \left(\frac{y}{\sigma_y}\right)^2 - \frac{2\rho\,xy}{\sigma_x\,\sigma_y}\right)\right) \tag{36}$$

Espérances mathématiques :

$$E(x) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x\,f(x,y)\,dxdy = \mu_x = 0 \tag{37}$$

$$E(y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} y\,f(x,y)\,dxdy = \mu_y = 0 \tag{38}$$

Calculons maintenant la pdf h(u,v), nous avons la relation suivante : il y a conservation des probabilités pour le couple (x,y) et (u,v) :

$$h(u,v)\,dudv = f(x,y)\,dxdy \tag{39}$$

et donc :

$$h(u,v) = f(\Phi(u,v)) \mid det(J_\Phi(u,v)) \mid \tag{40}$$

où : $\Phi : (u,v) \longrightarrow (x,y)$ et $J_\Phi(u,v)$ est le Jacobien de la transformation Φ.

$$det(J_\Phi(u,v)) = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{vmatrix} \quad avec\ \ x = \frac{u+v}{2}\ \ et\ y = \frac{u-v}{2}\frac{\sigma_y}{\sigma_x} \tag{41}$$

$$det(J_\Phi(u,v)) = \begin{vmatrix} \dfrac{1}{2} & \dfrac{1}{2} \\ \\ \dfrac{1}{2}\dfrac{\sigma_y}{\sigma_x} & -\dfrac{1}{2}\dfrac{\sigma_y}{\sigma_x} \end{vmatrix} = -\frac{1}{2}\frac{\sigma_y}{\sigma_x} \tag{42}$$

On obtient donc la pdf $h(u,v)$:

$$\begin{gathered} h(u,v) = \frac{1}{2}\frac{\sigma_y}{\sigma_x}\frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \\ exp\left(-\frac{1}{2(1-\rho^2)}\left(\left(\frac{u+v}{2}\right)^2\frac{1}{\sigma_x^2} - \frac{1}{2}\left(\frac{(u-v)}{2}\frac{\sigma_y}{\sigma_x}\right)^2\frac{1}{\sigma_y^2} - 2\rho\frac{(u+v)(u-v)}{4\sigma_x\sigma_y}\frac{\sigma_y}{\sigma_x}\right)\right) \end{gathered} \tag{43}$$

$$\begin{gathered} h(u,v) = \frac{1}{4}\frac{1}{\pi\sigma_x^2\sqrt{1-\rho^2}} \\ exp\left(-\frac{1}{8(1-\rho^2)}\left(\frac{u^2+2uv+v^2}{\sigma_x^2}\right) - \frac{1}{8(1-\rho^2)}\left(\frac{u^2-2uv+v^2}{\sigma_x^2}\right) + \frac{\rho}{4(1-\rho^2)}\left(\frac{u^2-v^2}{\sigma_x^2}\right)\right) \end{gathered} \tag{44}$$

$$\begin{gathered} h(u,v) = \frac{1}{4}\frac{1}{\pi\sigma_x^2\sqrt{1-\rho^2}} \\ exp\left(-\frac{1}{4(1-\rho^2)}\frac{u^2(1-\rho)}{\sigma_x^2} - \frac{1}{4(1-\rho^2)}\frac{v^2(1+\rho)}{\sigma_x^2}\right) \end{gathered} \tag{45}$$

Comme on a d'après l'exercice 2.1 : $\sigma_x^2(1+\rho_{xy}) = \dfrac{\sigma_u^2}{2}$ et $\sigma_x^2(1-\rho_{xy}) = \dfrac{\sigma_v^2}{2}$

$$h(u,v) = \frac{1}{\sqrt{2\pi}\sigma_u}\,exp\left(-\frac{u^2}{2\sigma_u^2}\right)\frac{1}{\sqrt{2\pi}\sigma_v}\,exp\left(-\frac{v^2}{2\sigma_v^2}\right) \tag{46}$$

$$h(u,v) = \text{Gauss}(u,0,\sigma_u)\ \ \cdot\ \ \text{Gauss}(v,0,\sigma_v) \tag{47}$$

On a donc u et v qui sont indépendantes car il y a factorisation de $\text{pdf}(u,v) = \text{pdf}(u)\cdot \text{pdf}(v)$. Si deux variables X et Y sont statistiquement indépendantes, alors elles sont non corrélées. L'inverse n'est pas nécessairement vrai. En effet, la non-corrélation implique l'indépendance que dans des cas particuliers et le cas des variables aléatoires gaussiennes est l'un de ces cas particuliers.

Exercice 3.1 :

On considère une variable continue t positive, distribuée suivant une loi exponentielle $\propto exp(-t/\tau)$: Donnez l'expression de la pdf $f(t)$ et calculez l'espérance mathématique $E(t)$ et la variance $V(t)$. On dispose de N mesures t_i de cette variable. On choisit la méthode du maximum de vraisemblance pour la détermination du paramètre τ. Rappelez le principe de la méthode et comment vous allez calculer ce paramètre et sa variance. Calculez par la méthode du maximum de vraisemblance l'estimation de τ et sa variance. On dispose de deux échantillons indépendants constitués de N_1 et N_2 mesures, de sorte que l'échantillon total contient $N = N_1 + N_2$ mesures. Calculez l'estimation de τ à partir de ces N mesures en faisant apparaître les deux contributions τ_1 et τ_2 des sous-échantillons N_1 et N_2.

Correction :

Soit t une variable aléatoire > 0 suivant une loi exponentielle, la pdf s'écrit :

$$f(t) = \begin{cases} \frac{1}{\tau} exp\left(-\frac{t}{\tau}\right) & si \quad t \geqslant 0 \\ 0 & si \quad t < 0 \end{cases} \tag{48}$$

On a de plus :

$$\int_0^{+\infty} f(t)dt = 1 \;\; et \;\; f(t) \geqslant 0 \;\forall t \tag{49}$$

Espérance mathématique, on intègre par partie :

$$E(t) = \int_0^{+\infty} \frac{t}{\tau}\, exp(\frac{-t}{\tau})\, dt = \big[\frac{t}{\tau}\,(-\tau) exp(-\frac{t}{\tau})\big]_0^{+\infty} - \int_0^{+\infty} \frac{1}{\tau}\,(-\tau)\, exp(-\frac{t}{\tau})\, dt \tag{50}$$

$$= \; 0 + \int_0^{+\infty} exp(-\frac{t}{\tau})\, dt = \tau \tag{51}$$

Variance mathématique :

$$E(t^2) = \int_0^{+\infty} \frac{t^2}{\tau}\, exp(\frac{-t}{\tau})\, dt = \big[-t^2\, exp(-\frac{t}{\tau})\big]_0^{+\infty} - \int_0^{+\infty} \frac{1}{\tau}\, \frac{2t(-\tau)}{\tau}\, exp(-\frac{t}{\tau})\, dt \tag{52}$$

$$= \; 0 + 2\int_0^{+\infty} t\, exp(-\frac{t}{\tau})\, dt = 2\tau^2 \tag{53}$$

$\Rightarrow \;\; V(t) = \tau^2 \;\; \Rightarrow \;\; \sigma = \tau$

Nous avons N mesures $t_i, i = 1, .., N$. La méthode du maximum de vraisemblance consiste à determiner l'estimateur τ qui est solution de la maximisation ou minimisation de la fonction de vraisemblance L définie ainsi :

$$L = \prod_{i=1}^{N} f(t_i) \tag{54}$$

où $f(t)$ est la pdf de la variable t. Le paramètre τ se trouve en minimisant ou maximisant L comme ceci :

$$\frac{\partial\,(-log(L))}{\partial\tau} = 0 \tag{55}$$

Calculons dans notre cas la fonction de vraisemblance L :

$$L = \prod_{i=1}^{N} \left(\frac{1}{\tau}\, e^{-\frac{t_i}{\tau}}\right) \tag{56}$$

On a :

$$-log\, L = -\sum_{i=1}^{N} log\left(\frac{1}{\tau}\, e^{-\frac{t_i}{\tau}}\right) = N\, log(\tau) + \sum_{i=1}^{N} \frac{t_i}{\tau} \tag{57}$$

On a donc l'équation suivante à résoudre :

$$\frac{\partial - log\, L}{\partial\tau} = 0 = \frac{N}{\tau} - \frac{1}{\tau^2} \sum_{i=1}^{N} t_i \tag{58}$$

$$\Rightarrow \ \tau = \frac{1}{N} \sum_{i=1}^{N} t_i \tag{59}$$

La variance de cette estimation va être dans ce cas (1 seul paramètre) donnée par la dérivée seconde :

$$V_\tau = -E\left(\frac{\partial^2(-log\, L)}{\partial\tau^2}\right)^{-1} \tag{60}$$

$$\Leftrightarrow\ V_\tau = -E\Big(\frac{\partial^2}{\partial\tau^2}\big(\sum_{i=1}^{N} -log\,\tau - \sum_{i=1}^{N}\frac{t_i}{\tau}\big)\Big)^{-1}$$

$$\Leftrightarrow\ V_\tau = -E\Big(\frac{\partial}{\partial\tau}\big(-\frac{N}{\tau}+\frac{1}{\tau^2}\sum_{i=1}^{N} t_i\big)\Big)^{-1}$$

$$= -E\big(\frac{N}{\tau^2} - 2\frac{\sum_{i=1}^{N} t_i}{\tau^3}\big)^{-1}$$

$$= E\big(\frac{N}{\tau^2}\big)^{-1}$$

$$\Rightarrow\ \sigma_\tau = \frac{\tau}{\sqrt{N}} \tag{61}$$

On a 2 séries indépendantes N_1 et N_2. Chacune donne son estimateur :

$$\tau_1 = \frac{1}{N_1}\sum_{i=1}^{N_1} t_i \quad \sigma_{\tau_1} = \frac{\tau_1}{\sqrt{N_1}} \quad \tau_2 = \frac{1}{N_1}\sum_{j=1}^{N_2} t_j \quad \sigma_{\tau_2} = \frac{\tau_2}{\sqrt{N_2}} \tag{62}$$

En considérant l'échantillon total $N_1 + N_2$:

$$\tau = \frac{1}{N_1+N_2}\Big(\sum_{i=1}^{N_1} t_i + \sum_{j=1}^{N_2} t_j\Big) \tag{63}$$

$$\Rightarrow \tau = \frac{N_1\tau_1 + N_2\tau_2}{N_1+N_2} \tag{64}$$

Ce résultat est la moyenne pondérée des deux échantillons.

Exercice 3.2 :

Deux expériences indépendantes ont mesuré (τ_1, σ_1) et (τ_2, σ_2) les σ_i représentant l'erreur sur les deux mesures.

(1) A partir de ces deux mesures, en supposant les erreurs gaussiennes, on veut obtenir l'estimation de τ et son erreur. (Combinaison de deux mesures en tenant compte de leurs erreurs). -Quelle méthode utilisez-vous pour cela ? -Calculez l'estimation de τ et de son erreur.

(2) A partir des deux mesures indépendantes (τ_1, σ_1) et (τ_2, σ_2) : définissez le nombre équivalent $\tilde{N}_1$ et $\tilde{N}_2$ de chacune des deux mesures ; donnez les relations les définissant. On utilise la méthode du maximum de vraisemblance pour calculer l'estimation de τ à partir de la définition des nombres équivalents d'évènements des deux mesures. Qu'obtenez-vous pour l'estimation de τ dans ce cas. (Faites apparaître τ_1, σ_1 et τ_2, σ_2) dans l'expression. Comparez-la à l'estimation précédente calculée plus haut en (1).

Correction :

On choisit comme dans l'exercice précédent la méthode du maximum de vraisemblance avec la pdf des 2 mesures :

$$f(\tau,\sigma) = \frac{1}{\sqrt{2\pi}\,\sigma}\, exp\big(-\frac{1}{2}\frac{(\tau-\hat{\tau})^2}{\sigma^2}\big) \tag{65}$$

Nous devons maximiser la fonction de vraisemblance suivante :

$$L = \prod_{i=1}^{2} \frac{1}{\sqrt{2\pi}\sigma_i}\, exp\big(-\frac{1}{2}\frac{(\tau_i-\hat{\tau})^2}{\sigma_i^2}\big) \tag{66}$$

d'après la condition suivante :

$$\frac{\partial\,(-log(L))}{\partial\hat{\tau}} = 0 \tag{67}$$

On trouve :

$$\Rightarrow \hat{\tau} = \frac{\tau_1/\sigma_1^2 + \tau_2/\sigma_2^2}{1/\sigma_1^2 + 1/\sigma_2^2} \tag{68}$$

$\sigma_{\hat{\tau}}$ est donné par la dérivée seconde :

$$\frac{1}{\sigma_{\hat{\tau}}} = \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} \tag{69}$$

Pour ces deux mesures, le nombre équivalent d'évènements $\tilde{N}$ est défini par :

$$\frac{\sigma_1}{\tau_1} = \frac{1}{\sqrt{\tilde{N}_1}} \qquad \frac{\sigma_2}{\tau_2} = \frac{1}{\sqrt{\tilde{N}_2}} \tag{70}$$

C'est l'erreur relative de la mesure traduite par l'erreur statistique due au nombre d'évènements.

Si on applique le calcul de l'exercie 3.1 avec le nombre équivalent $\tilde{N}$, on obtient :

$$\hat{\tau} = \frac{\tilde{N}_1\tau_1 + \tilde{N}_2\tau_2}{\tilde{N}_1 + \tilde{N}_2} \tag{71}$$

On a finalement :

$$\hat{\tau} = \frac{\tau_1/(\sigma_1/\tau_1)^2 + \tau_2/(\sigma_2/\tau_2)^2}{1/(\sigma_1/\tau_1)^2 + 1/(\sigma_2/\tau_2)^2} \tag{72}$$

En conclusion :

cas (1) : pondération par le carré de l'inverse de l'erreur.
cas (2) : pondération par le carré de l'inverse de l'erreur relative.

Exercice 4.1 :

On veut calculer l'intégrale d'une fonction $f(x, y)$ par une méthode de Monte-Carlo :

$$\int_{x=0}^{x=1} \int_{y=0}^{y=x} f(x, y)\, dx\, dy \tag{73}$$

Pour cela on dispose d'un générateur de nombres aléatoires uniformes entre $[0, 1]$. Comment va-t-on procéder ? (faites le schéma de la région d'intégration dans le plan (x, y)).

Correction :

Rappel de la méthode Monte-Carlo :

Nous disposons de l'expression de l'espérance mathématique d'une fonction g de variable aléatoire X, résultant du théorème de transfert, selon lequel

$$G = E(g(X)) = \int_a^b g(x) f_X(x)\, \mathrm{d}x \tag{74}$$

où f_X est une fonction de densité sur le support $[a, b]$. Il est fréquent de prendre une distribution uniforme sur $[a, b]$:

$$f_X(x) = \frac{1}{b-a} \tag{75}$$

L'idée est de produire un échantillon $(x_1, x_2, ..., x_N)$ de la loi X (donc d'après la densité f_X) sur le support $[a, b]$, et de calculer un nouvel estimateur dit de Monte-Carlo, à partir de cet échantillon.

La loi des grands nombres suggère de construire cet estimateur à partir de la moyenne empirique :

$$\tilde{g_N} = \frac{1}{N} \sum_{i=1}^{N} g(x_i), \tag{76}$$

qui se trouve être, par ailleurs, un estimateur sans biais de l'espérance.
Ceci est l'estimateur de Monte-Carlo. Nous voyons bien qu'en remplaçant l'échantillon par un ensemble de valeurs prises dans le support d'une intégrale, et de la fonction à intégrer, nous pouvons donc construire une approximation de sa valeur,

construite statistiquement.

La méthode Monte-Carlo permet grâce au générateur "uniforme" d'obtenir la valeur numérique de l'intégrale ci-dessus notée I. La région d'intégration est représentée sur la figure 1 ci-dessous :

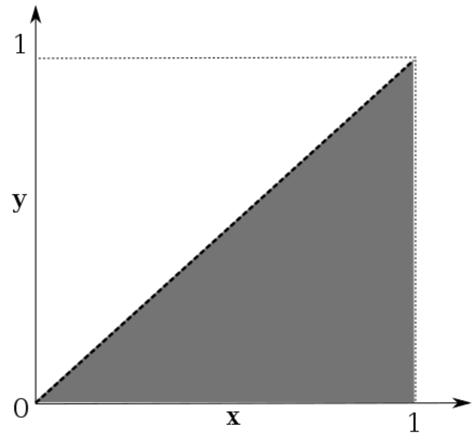

FIGURE 1 – Région d'intégration

Nous pouvons envisager 2 cas :

cas (1) - On fait 2 tirages, un pour x et un autre pour y :

$$\text{tirage} = \begin{cases} x_i & \in & [0,1] \\ y_i & \in & [0,1] \end{cases} \tag{77}$$

Si $x_i > y_i$, alors on incrémente I de la manière suivante : $I = I + f(x_i, y_i)$, sinon on réitère le tirage. On voit ici que l'on fait 2 tirages et qu'en moyenne, 1 sur 2 est perdu.

cas (2) - On fait 2 tirages comme précédemment :

$$\text{tirage} = \begin{cases} x_i & \in & [0,1] \\ y_i & \in & [0,1] \end{cases} \tag{78}$$

Ensuite, selon les résultats, on incrémente I de la manière suivante :

$$\text{Incrémentation} \begin{cases} si \quad x_i > y_i \quad I = I + f(x_i, y_i) \\ si \quad x_i < y_i \quad I = I + f(y_i, x_i) \end{cases} \tag{79}$$

L'avantage dans ce cas est que l'on utilise tous les tirages contrairement au cas (1). Il faut ensuite multiplier la quantité I obtenue par l'intervalle $(b - a)$ et la diviser par le nombre de tirages effectués N. Nous avons ainsi calculé la valeur demandée de l'intégrale.

Exercice 4.2 :

Une source ponctuelle émet de manière isotrope et un détecteur "disque" perpendiculaire à un axe passant par la source couvre un angle θ_0. On a donc une symétrie cylindrique avec 2 angles : φ entre $[0,2\pi]$ et θ tel que $\cos(\theta_0) < \cos(\theta) < 1$.

Exprimez la pdf de $(\varphi, \cos\theta)$.

Disposant d'un générateur de nombres aléatoires uniformes entre [0,1], comment allez-vous tirer de manière aléatoire dans l'acceptance du disque un couple $(\varphi, \cos\theta)$? A l'aide de ce détecteur on va faire des comptages pendant des intervalles de temps égaux Δt. Quelle va être la loi de distribution des nombres de coups enregistrés ?

Correction :

Puisque la source émet de façon isotrope, les variables aléatoires φ et $\cos\theta$ admettent une loi uniforme respectivement sur $[0,2\pi]$ et $[\cos\theta_0, 1]$.

On a donc pour la pdf de φ :

$$\text{pdf}(\varphi) = \frac{1}{2\pi} \tag{80}$$

Pour la v.a $\cos\theta$, on peut écrire :

$$\text{pdf}(\cos\theta)\, d\cos(\theta) = K\, d\cos(\theta) \quad \text{avec } K\text{=constante} \tag{81}$$

$$\Rightarrow \int_{\cos\theta_0}^{1} \text{pdf}(\cos\theta)\, d\cos(\theta) = 1 = K \int_{\cos\theta_0}^{1} d\cos(\theta) = K\,(1-\cos\theta_0) \tag{82}$$

$$\Rightarrow K = \frac{1}{(1-\cos\theta_0)} \tag{83}$$

Étant donné que φ et $\cos\theta$ sont indépendantes, on en conlut :

$$\text{pdf}(\varphi, \cos\theta) = \text{pdf}(\varphi)\, \text{pdf}(\cos\theta) = \frac{1}{(1-\cos\theta_0)}\, \frac{1}{2\pi} \tag{84}$$

Disposant d'un générateur de nombres aléatoires entre [0,1], nous faisons les correspondances suivantes :

$$\varphi = 2\,\pi\, u \;\text{ avec } u \text{ uniforme sur } [0,1] \tag{85}$$

Pour $\cos\theta$, on se sert de la relation suivante en prenant la variable aléatoire v uniforme sur [0,1] :

$$\int_{\cos\theta_0}^{\cos\theta} \frac{d\cos\theta}{1-\cos\theta_0} = \int_0^v d\,v \tag{86}$$

$\Rightarrow$

$$\cos\theta - \cos\theta_0 = v(1-\cos\theta_0) \Rightarrow \cos\theta = v + (1-v)\cos\theta_0 \tag{87}$$

Le nombre de coups enregistrés dans le compteur pendant Δt va suivre une loi de Poisson.

Exercice 5.1 :

On dispose de mesures y_i $i = 1, ..., n$ dépendant d'une coordonnée x_i dont le modèle théorique est linéaire $y = ax + b$. A l'aide de ces mesures, on cherche à déterminer les valeurs des paramètres a et b.

Les mesures y_i ont une erreur σ_i et dans un premier temps, les coordonnées x_i sont considérées sans erreur.

- Exprimer le χ^2 que vous allez utiliser.
- Donner les 2 équations vous permettant de calculer les estimations de a et b.

Correction :

Quand nous avons n mesures indépendantes y_i $i = 1, ..., n$ ainsi que n coordonnées x_i dans un modèle linéaire $y = ax + b$, avec une erreur σ_i sur y_i et sans erreur sur x_i, le χ^2 s'écrit :

$$\chi^2 = \sum_{i=1}^{n} \frac{(y_i - (a\,x_i + b))^2}{\sigma_i^2} \tag{88}$$

Il faut alors minimiser le χ^2 pour calculer les valeurs a et b, ce qui s'obtient avec les 2 équations linéaires suivantes à 2 inconnues a et b :

$$\frac{\partial\chi^2}{\partial a} = \sum_{i=1}^{n} \frac{(y_i - a\,x_i - b)\,x_i}{\sigma_i^2} = 0 \tag{89}$$

et

$$\frac{\partial\chi^2}{\partial b} = \sum_{i=1}^{n} \frac{(y_i - a\,x_i - b)}{\sigma_i^2} = 0 \tag{90}$$

Exercice 5.2 :

Les coordonnées x_i sont elles aussi connues avec une erreur δ_i.

Exprimez le χ^2 qui doit être utilisé dans ce cas.
Donnez les 2 équations vous permettant de calculer les estimations de a et b. Que constatez-vous par rapport à la situation précédente ?

Correction :

La formule du χ^2 va être modifiée car il faut tenir compte de l'erreur δ_i sur les coordonnées x_i. En effet, la variance de $(y_i - a\,x_i - b)$ n'est plus seulement égale à $V(y_i) = \sigma_i^2$:

$$V(y_i - a\,x_i - b) = V(y_i^2) + a^2\,V(x_i) = \sigma_i^2 + a^2\,\delta_i^2 \tag{91}$$

Le dénominateur du χ^2 dépend donc du paramètre a :

$$\chi^2 = \sum_{i=1}^{n} \frac{(y_i - (a\,x_i + b))^2}{\sigma_i^2 + a^2\,\delta_i^2} \tag{92}$$

La minimisation du χ^2 s'obtient toujours par les 2 équations suivantes :

$$\frac{\partial \chi^2}{\partial a} = 0 \text{ et } \frac{\partial \chi^2}{\partial b} = 0 \tag{93}$$

Mais nous voyons que ces équations ne sont pas linéaires car la deuxième qui minimise le χ^2 par b dépend d'une puissance de a. Nous n'avons pas de solution analytique dans ce cas là.

Exercice 6.1 :

La méthode du χ^2 vous a permis de déterminer les valeurs de 2 paramètres $a \pm \sigma_a$ et $b \pm \sigma_b$ et donc la valeur de χ^2_{min}.

On veut tracer dans le plan a, b le contour correspondant à un niveau de confiance donné.

Donnez la loi de probabilité que vous utilisez.
Dans ce cas, en fixant le niveau de confiance CL, donnez l'expression de la variation du χ^2 par rapport à χ^2_{min}, c'est-à-dire le contour correspondant à une valeur du χ^2 telle que : $\chi^2(CL) = \chi^2_{min} + \Delta\chi^2$

Qu'obtenez-vous pour $CL = 0.68$?

Correction :

Le terme $\chi^2(CL)$ ne fait intervenir que les dérivées secondes suivantes du χ^2 à l'ordre le plus bas :

$$\frac{\partial^2 \chi^2}{\partial a^2} \;,\; \frac{\partial^2 \chi^2}{\partial b^2} \text{ et } \frac{\partial^2 \chi^2}{\partial a \partial b} \tag{94}$$

Pour $\Delta\chi^2$, la loi de distribution est une loi du χ^2 à 2 degrés de liberté dont la pdf est :

$$f(\Delta\chi^2) = \frac{1}{2}\, e^{-\dfrac{\Delta\chi^2}{2}} \tag{95}$$

Donc pour un niveau de confiance donné CL, on a :

$$1 - CL = \int_{\Delta\chi^2_{CL}}^{+\infty} \frac{1}{2}\, e^{-\dfrac{\Delta\chi^2}{2}}\, d\,\chi^2 = e^{-\dfrac{\Delta\chi^2_{CL}}{2}} \tag{96}$$

D'où : $\Delta\chi^2_{CL} = -2\, ln(1 - CL)$.

Pour $CL = 0.68$, on obtient : $\Delta\chi^2_{CL} = 2.28$