Exercise 1.1 :

We consider a random variable x on $]-\infty, +\infty[$ following a Gauss's law with μ and σ parameters.

- Give the expression of Gauss's law.
- Give the conditions on parameters to make this law be a probability distribution function (pdf).
- What is the Characteristic function (definition and expression) ?
- Do the same for cumulative function.

Reminder :

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

Correction :

Let x be a random variable defined on $]-\infty, +\infty[$ and following a Normal distribution (Gaussian).
Distribution function is :

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

Conditions and parameters to make this formula be a pdf :

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

Characteristic function :

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

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

We do the following substitution variable :

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

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

We use this equality (demonstration can be found on Web) :

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

Taking $z = it\sqrt{2}\sigma$, one gets :

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

Cumulative function :

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

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

Another substitution variable $x" = x' - \mu$ gives :

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

For the second integral, doing the substitution $v = \dfrac{x"}{\sqrt{2}\sigma}$, we have :

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

So the final result :

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

Exercise 1.2 :

We do another substitution variable $y = exp(x)$, where x is defined above and follows Gauss's law.

- How will you calculate the pdf of y ?
- Give the expression of pdf(y).

- Calculate the expectation $E(y)$; first write its definition.
- Calculate the expression of $V(y)$ (variance of y).

Correction :

We have to find the distribution function of y recorded $g(y)$. Then,

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

Transfert theorem : taking into account that $x = ln\, y$ and there is conservation of probabilities with variable substitution, we have :

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

that implies :

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

So we deduce for $g(y)$:

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

Expectation $E(y)$:

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

with $g(y) = 0$ if $y \leq 0$.

Given $x = ln\, y$ and $dy = e^x\, dx$, we get :

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

One recognizes (6) equation for characteristic function :

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

Variance $V(y)$:

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

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

Following the same previous reasoning :

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

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

Finally :

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

Exercise 1.3 :

We have a set of random variables $x_i, i = 1, ..., n$ and we consider the random variable $z = \sum_{i=1}^{n} a_i x_i$ where a_i are constants.

- Write the expectation $E(z)$.
- Write the expression of $V(z)$ (in general case).

Correction :

Expectation $E(z)$:

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

Variance $V(z)$:

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

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

Exercise 2.1 :

One starts from 2 random variables x and y with respectively σ_x^2 and σ_y^2 standard deviations and ρ_{xy} correlation factor.

We do the following substitution of variable :

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

Calculate $V(u)$ and $V(v)$ variances and covariance $cov(u, v)$. What can you conclude ? What is the correlation factor ρ_{uv} ?

Correction :

We use the previous relation (27) so that :

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

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

In the same way for $V(v)$:

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

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

We can get for variances calculation of u and v :

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

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

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

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

We see that u and v are not correlated : conclusion is that we can made 2 no-correlated variables from 2 correlated ones.

Exercise 2.2 :

Both x and y random variables follow a Gauss's law (pdf= $f(x,y)$) with $\mu_x = 0$ and $\mu_y = 0$.

- Write its expression.
- Calculate Expectations $E(x)$ and $E(y)$.
- Calculate the distribution function for u and v (pdf= $h(u,v)$).
- What do you conclude about these two random variables u and v ?

Correction :

General formula of 2D Gaussian distribution function :

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

In our case, function is written with $\mu_x = \mu_y = 0$:

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

Expectations :

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

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

Calculate now the $h(u,v)$ pdf : from the transfert theorem, there is probabilities conservation for (x,y) and (u,v) couples :

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

then :

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

where $\Phi : (u,v) \longrightarrow (x,y)$ et $J_\Phi(u,v)$ is the Jacobian of Φ transformation.

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

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

Finally, one deduces $h(u,v)$ pdf :

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

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

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

Since we get from Exercise 2.1 :

$$\sigma_x^2(1+\rho_{xy}) = \frac{\sigma_u^2}{2}\sigma_x^2(1-\rho_{xy}) = \frac{\sigma_v^2}{2} \tag{46}$$

then,

$$h(u,v) = \frac{1}{\sqrt{2\pi}\sigma_u}\,exp\left(-\frac{u^2}{2\sigma_u^2}\right)\frac{1}{\sqrt{2\pi}\sigma_v}\,exp\left(-\frac{v^2}{2\sigma_v^2}\right) \tag{47}$$

$$h(u,v) = Gauss(u,0,\sigma_u)\quad\cdot\quad Gauss(v,0,\sigma_v) \tag{48}$$

We can see that u and v are independant because of pdf factorization $h(u,v) = \text{pdf}(u)\cdot\text{pdf}(v)$. If two variables X and Y are statiscally independent, then they are not correlated. The reverse is not necessarily true. Indeed, no-correlation implies independence only in particular cases and Gaussian random variables is one of this special cases.

Exercise 3.1 :

One considers a continous and positive random variable t, following an exponential law $\propto exp(-t/\tau)$. Give the expression of its pdf $f(t)$ and calculate its expectation $E(t)$ and variance $V(t)$. We have N measures t_i of this variable and choose the maximum likelihood method to determine τ parameter. Describe the principle of the maximum likelihood method and apply it to calculate this parameter and its variance. Now, we have 2 independants samples of N_1 and N_2 measures so that total sample is consisted of $N = N_1 + N_2$ measures. Calculate the estimation of τ from these N measures while making appear the contribution of τ_1 and τ_2.

Correction :

Let t be a random variable > 0 following an exponential law ; pdf is given by :

$$f(t) = \begin{cases} \frac{1}{\tau}\,exp\left(-\frac{t}{\tau}\right) & \text{if} \quad t \geqslant 0 \\ 0 & \text{if} \quad t < 0 \end{cases} \tag{49}$$

Moreover :

$$\int_0^{+\infty} f(t)dt = 1 \;\text{ and }\; f(t) \geqslant 0\;\forall t \tag{50}$$

To get expectation, we use integration rules :

$$E(t) = \int_0^{+\infty}\frac{t}{\tau}\,exp(\frac{-t}{\tau})\,dt = [\frac{t}{\tau}\,(-\tau)exp(-\frac{t}{\tau})]_0^{+\infty} - \int_0^{+\infty}\frac{1}{\tau}\,(-\tau)\,exp(-\frac{t}{\tau})\,dt \tag{51}$$

$$= 0 + \int_0^{+\infty} exp(-\frac{t}{\tau})\, dt = \tau \tag{52}$$

For variance :

$$E(t^2) = \int_0^{+\infty} \frac{t^2}{\tau}\, exp(\frac{-t}{\tau})\, dt = \big[-t^2\, exp(-\frac{t}{\tau}) \big]_0^{+\infty} - \int_0^{+\infty} \frac{1}{\tau}\, \frac{2t(-\tau)}{\tau}\, exp(-\frac{t}{\tau})\, dt \tag{53}$$

$$= 0 + 2 \int_0^{+\infty} t\, exp(-\frac{t}{\tau})\, dt = 2\tau^2 \tag{54}$$

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

We own N measures $t_i, i = 1, .., N$. Maximum likelihood allows to compute an estimation of τ parameter which is solution of the maximum or minimum for likelihood function L defined as :

$$L = \prod_{i=1}^{N} f(t_i) \tag{55}$$

where $f(t)$ is the pdf of t variable. τ parameter is found minimizing L :

$$\frac{\partial\, (-log(L))}{\partial \tau} = 0 \tag{56}$$

Calculate in our case the likelihood function L :

$$L = \prod_{i=1}^{N} (\frac{1}{\tau}\, e^{-\frac{t_i}{\tau}}) \tag{57}$$

We get :

$$-log\, L = -\sum_{i=1}^{N} log\big(\frac{1}{\tau}\, e^{-\frac{t_i}{\tau}}\big) = N\, log(\tau) + \sum_{i=1}^{N} \frac{t_i}{\tau} \tag{58}$$

So one has to solve :

$$\frac{\partial - log\, L}{\partial \tau} = 0 = \frac{N}{\tau} - \frac{1}{\tau^2} \sum_{i=1}^{N} t_i \tag{59}$$

$$\Rightarrow\ \tau = \frac{1}{N} \sum_{i=1}^{N} t_i \tag{60}$$

Variance of this estimation will be given by the second derivate :

$$V_\tau = -E\left(\frac{\partial^2(-log\,L)}{\partial\tau^2}\right)^{-1} \tag{61}$$

$$\begin{aligned}
\Leftrightarrow\ V_\tau &= -E\left(\frac{\partial^2}{\partial\tau^2}(\sum_{i=1}^{N} -log\,\tau - \sum_{i=1}^{N}\frac{t_i}{\tau})\right)^{-1} \\
\Leftrightarrow\ V_\tau &= -E\left(\frac{\partial}{\partial\tau}(-\frac{N}{\tau} + \frac{1}{\tau^2}\sum_{i=1}^{N} t_i)\right)^{-1} \\
&= -E(\frac{N}{\tau^2} - 2\frac{\sum_{i=1}^{N} t_i}{\tau^3})^{-1} \\
&= E(\frac{N}{\tau^2})^{-1} \\
\Rightarrow\ \sigma_\tau &= \frac{\tau}{\sqrt{N}}
\end{aligned} \tag{62}$$

Now we have 2 independant samples N_1 et N_2. Each produces its estimation :

$$\tau_1 = \frac{1}{N_1}\sum_{i=1}^{N_1} t_i \quad \sigma_{\tau_1} = \frac{\tau_1}{\sqrt{N_1}} \quad \tau_2 = \frac{1}{N_1}\sum_{j=1}^{N_2} t_j \quad \sigma_{\tau_2} = \frac{\tau_2}{\sqrt{N_2}} \tag{63}$$

Taking into account the total sample N_1+N_2 :

$$\tau = \frac{1}{N_1+N_2}\left(\sum_{i=1}^{N_1} t_i + \sum_{j=1}^{N_2} t_j\right) \tag{64}$$

$$\Rightarrow \tau = \frac{N_1\tau_1 + N_2\tau_2}{N_1+N_2} \tag{65}$$

This result is the weighted average from these 2 samples.

Exercise 3.2 :

Two independant experiments have measured (τ_1, σ_1) and (τ_2, σ_2) with σ_i representing errors on measures.

(1) From these two measures, assuming errors are gaussian, we want to get the extimation of τ and its error (i.e with a combination of two measures).
- Which method do you use ?
- Calculate the estimation of τ and its error.

(2) From these two measures (τ_1, σ_1) and (τ_2, σ_2) :
Define the equivalent number $\tilde{N}_1$ and $\tilde{N}_2$ of each measure ; give the relations defining them. We use the maximum likelihood method to find the estimation of τ from the definition of these 2 equivalent numbers. Calculate this estimation of τ in this case (Making appear τ_1, σ_1 et τ_2, σ_2 in the expression). Compare it to previous expression in (1).

Correction :

As the previous Exercise, we choose the maximum likelihood method with the pdf of 2 measures :

$$f(\tau, \sigma) = \frac{1}{\sqrt{2\pi}\,\sigma}\, exp\big(-\frac{1}{2}\frac{(\tau-\hat{\tau})^2}{\sigma^2}\big) \tag{66}$$

One has to maximize the likelihood function :

$$L = \prod_{i=1}^{2} \frac{1}{\sqrt{2\pi}\sigma_i}\, exp\big(-\frac{1}{2}\frac{(\tau_i-\hat{\tau})^2}{\sigma_i^2}\big) \tag{67}$$

taking the following condition :

$$\frac{\partial\,(-log(L))}{\partial\hat{\tau}} = 0 \tag{68}$$

We get :

$$\Rightarrow \hat{\tau} = \frac{\tau_1/\sigma_1^2 + \tau_2/\sigma_2^2}{1/\sigma_1^2 + 1/\sigma_2^2} \tag{69}$$

$\sigma_{\hat{\tau}}$ is deducted from second derivate :

$$\frac{1}{\sigma_{\hat{\tau}}} = \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} \tag{70}$$

For these both measures, equivalent number $\tilde{N}$ is defined by :

$$\frac{\sigma_1}{\tau_1} = \frac{1}{\sqrt{\tilde{N}_1}} \qquad \frac{\sigma_2}{\tau_2} = \frac{1}{\sqrt{\tilde{N}_2}} \tag{71}$$

This is the relative error of measure expressed by the statistical error due to the number of events.

If we apply the calculation of exercise 3.1 with the equivalent number $\tilde{N}$, we have :

$$\hat{\tau} = \frac{\tilde{N}_1 \tau_1 + \tilde{N}_2 \tau_2}{\tilde{N}_1 + \tilde{N}_2} \tag{72}$$

Finally :

$$\hat{\tau} = \frac{\tau_1/(\sigma_1/\tau_1)^2 + \tau_2/(\sigma_2/\tau_2)^2}{1/(\sigma_1/\tau_1)^2 + 1/(\sigma_2/\tau_2)^2} \tag{73}$$

In conclusion :

case (1) : weighted by the square of inverse error.
case (2) : weighted by the square of relative error.

Exercise 4.1 :

We are looking to compute the integral of a $f(x, y)$ function by Monte-Carlo method :

$$\int_{x=0}^{x=1} \int_{y=0}^{y=x} f(x, y)\, dx\, dy \tag{74}$$

For this, we own a random number and uniform generator between $[0, 1]$.
How will you proceed (Make a schema of the integration area on (x, y) plane) ?

Correction :

Reminder - Monte-Carlo method :

From the transfert theorem, we get the expression of expectation for a function g representing the random variable X as :

$$G = E(g(X)) = \int_a^b g(x) f_X(x)\, \mathrm{d}x \tag{75}$$

where f_X is a pdf on $[a, b]$ interval. One usually takes a uniform distribution on $[a, b]$:

$$f_X(x) = \frac{1}{b - a} \tag{76}$$

Principle is to generate a sample $(x_1, x_2, ..., x_N)$ with law of X (so we calculate an estimator called "the estimator of Monte-Carlo" from this sample).

Law of large numbers assumes to build this estimator from the empirical average :

$$\tilde{g_N} = \frac{1}{N} \sum_{i=1}^{N} g(x_i), \tag{77}$$

which is, by the way, an unbiased estimator of expectation. This is the estimator of Monte-Carlo. We can clearly see that, by replacing sample by a set of values located on an integral support, and the function to integrate, we can build an approximation of its value, statistically made.
Thanks to the "uniform generator", Monte-Carlo method gives a numerical value of the integral noticed I. Area of integration is represented on the figure 1 below :

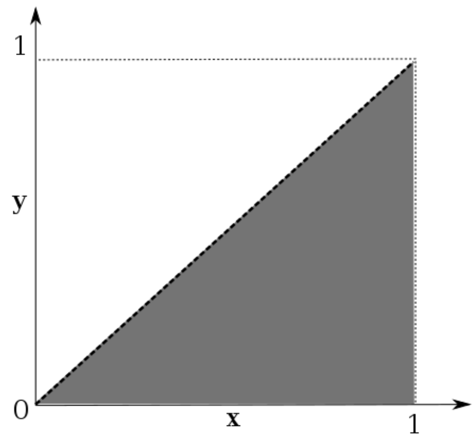

FIGURE 1 – Representation of integration area

So we can distinguish two cases :

case (1) - We do 2 random sampling, one for x and the other for y :

$$\text{Random sampling} = \left\{ \begin{array}{lcl} x_i & \in & [0,1] \\ y_i & \in & [0,1] \end{array} \right. \tag{78}$$

If $x_i > y_i$, then we increment I in the following way : $I = I + f(x_i, y_i)$, else we redo a random sampling. We have done 2 random sampling and the average to lose one is $1/2$.

case (2) - We do 2 random sampling like previous :

$$\text{Random sampling} = \left\{ \begin{array}{lcl} x_i & \in & [0,1] \\ y_i & \in & [0,1] \end{array} \right. \tag{79}$$

Then, as a function of results, we increment I in the following way :

$$\text{Incrementation} \left\{ \begin{array}{lll} if & x_i > y_i & I = I + f(x_i, y_i) \\ if & x_i < y_i & I = I + f(y_i, x_i) \end{array} \right. \tag{80}$$

The advantage here is that we use all the random sampling values unlike case (1). Finally, one has to multiply the I quantity by $(b - a)$ interval and divide by the size of random sampling N, so we get the numerical value of integral.

Exercise 4.2 :

A point source emits isotropically and covers an angle of θ_0. A disk detector is positioned perpendicular to this source. So we have a cylindrical symetry with 2 angles : φ between [0,2π] and θ as $\cos(\theta_0) < \cos(\theta) < 1$.

Calculate the pdf of $(\varphi, \cos\theta)$.

Having a generator of random values on [0,1], how will you random sample in acceptance disk a couple $(\varphi, \cos\theta)$?
With this detector, we want to do counting during equal interval time Δt. Express the distribution function for the number of recorded hits.

Correction :

Since point source emits isotropically, random variables φ and $\cos\theta$ follow a uniform law, respectively on [0,2π] et $[\cos\theta_0, 1]$.

We have for the φ pdf :

$$\mathrm{pdf}(\varphi) = \frac{1}{2\pi} \tag{81}$$

For the $\cos\theta$ random variable, we can write :

$$\mathrm{pdf}(\cos\theta)\, d\cos(\theta) = K\, d\cos(\theta) \quad \text{avec } K\text{=constante} \tag{82}$$

$$\Rightarrow \int_{\cos\theta_0}^{1} \mathrm{pdf}(\cos\theta)\, d\cos(\theta) = 1 = K \int_{\cos\theta_0}^{1} d\cos(\theta) = K\,(1 - \cos\theta_0) \tag{83}$$

$$\Rightarrow K = \frac{1}{(1 - \cos\theta_0)} \tag{84}$$

Given φ and $\cos\theta$ are independant, we conclude :

$$\mathrm{pdf}(\varphi, \cos\theta) = \mathrm{pdf}(\varphi)\,\mathrm{pdf}(\cos\theta) = \frac{1}{(1 - \cos\theta_0)}\,\frac{1}{2\pi} \tag{85}$$

With a random generator between [0,1], one makes the following correspondences :

$$\varphi = 2\,\pi\,u \;\text{ with } u \text{ uniform on } [0, 1] \tag{86}$$

For $\cos\theta$, we use the relation, taking random variable v as uniform on [0,1] :

$$\int_{\cos\theta_0}^{\cos\theta} \frac{d\cos\theta}{1 - \cos\theta_0} = \int_0^v d\,v \tag{87}$$

$\Rightarrow$

$$\cos\theta - \cos\theta_0 = v(1 - \cos\theta_0) \Rightarrow \cos\theta = v + (1 - v)\cos\theta_0 \tag{88}$$

The number of hits recorded during interval Δt will follow a Poisson law.

Exercise 5.1 :

we have a set of measures y_i $i = 1, ..., n$ depending from coordinates x_i and whose theorical model is linear $y = ax + b$. Thanks to these data, we look for determining values of a and b.

Measures y_i have an error σ_i. Firstly, coordinates x_i are considered being without error.

- Express the χ^2 you have to use.
- Express the 2 equations from which you can deduce estimations for a and b.

Correction :

With n independant measures $y_i\, i = 1, ..., n$ and n coordinates x_i in a linear model $y = ax + b$, with σ_i errors on y_i and no errors on x_i, one can write the χ^2 as :

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

One has to minimize the χ^2 to compute a and b values : one gets 2 linear equations with 2 unknowns (a and b) :

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

and

$$\frac{\partial \chi^2}{\partial b} = \sum_{i=1}^{n} \frac{(y_i - a\, x_i - b)}{\sigma_i^2} = 0 \tag{91}$$

Exercise 5.2 :

Now, x_i coordinates have δ_i errors.

Express the χ^2 which has to be used in this case.
Write the 2 equations from which you calculate a and b estimation. What's the difference with previous case ?

Correction :

χ^2 formula must be modified because we take into account of δ_i errors on coordinates x_i. Indeed, variance of $(y_i - a\, x_i - b)$ is not only equal to $V(y_i) = \sigma_i^2$:

$$V(y_i - a\, x_i - b) = V(y_i^2) + a^2\, V(x_i) = \sigma_i^2 + a^2\, \delta_i^2 \tag{92}$$

The denominator of χ^2 depends on a parameter :

$$\chi^2 = \sum_{i=1}^{n} \frac{(y_i - (a\, x_i + b))^2}{\sigma_i^2 + a^2\, \delta_i^2} \tag{93}$$

Minimization of χ^2 is got by the 2 following equations :

$$\frac{\partial \chi^2}{\partial a} = 0 \;\text{ et }\; \frac{\partial \chi^2}{\partial b} = 0 \tag{94}$$

But we notice that all these equations are not linear since the second one which minimizes the χ^2 with b depends on powered a : we have not analytical solution in this case.

Exercise 6.1 :

χ^2 method can give estimations on parameters, $a \pm \sigma_a$ and $b \pm \sigma_b$, so a minimum value χ^2_{min}.

We want to draw in a, b plane the contour related to a given confidence level.

Express the distribution function that you use.

In this case, by fixing the confidence level, express the variation of χ^2 compared to χ^2_{min}, so we could write : $\chi^2(CL) = \chi^2_{min} + \Delta\chi^2$

What value do you get with $CL = 0.68$?

Correction :

$\chi^2(CL)$ term only contains second derivatives of χ^2 at lowest level :

$$\frac{\partial^2 \chi^2}{\partial a^2}\;,\; \frac{\partial^2 \chi^2}{\partial b^2} \;\text{ et }\; \frac{\partial^2 \chi^2}{\partial a \partial b} \tag{95}$$

Concerning $\Delta\chi^2$, distribution function is a χ^2 law with 2 freedom degrees; pdf is written as :

$$f(\Delta\chi^2) = \frac{1}{2}\, e^{-\frac{\Delta\chi^2}{2}} \tag{96}$$

So for a fixed confidence level CL, we have :

$$1 - CL = \int_{\Delta\chi^2_{CL}}^{+\infty} \frac{1}{2}\, e^{-\frac{\Delta\chi^2}{2}}\, d\,\chi^2 = e^{-\frac{\Delta\chi^2_{CL}}{2}} \tag{97}$$

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

For $CL = 0.68$, we have : $\Delta\chi^2_{CL} = 2.28$