Ecole Nationale Supérieure de Techniques Avancées ParisTech
PRB202 - Martingales et Algorithmes Stochastiques
Corrigé PC1 - 29 novembre 2017

Rappel : **Caractérisation** 1 **de l'espérance conditionnelle**.
Etant donné $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité; désignons par :

- $\mathbb{L}^+(\Omega, \mathcal{F}, \mathbb{P})$ l'espace des variables aléatoires X définies sur $(\Omega, \mathcal{F}, \mathbb{P})$ et à valeurs positives.
- $\mathbb{L}^1(\Omega, \mathcal{F}, \mathbb{P})$ l'espace des (classes de) variables aléatoires X définies sur $(\Omega, \mathcal{F}, \mathbb{P})$ et à valeurs réelles telles que $\mathbb{E}[|X|] < +\infty$.
- $\mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ l'espace des (classes de) variables aléatoires X définies sur $(\Omega, \mathcal{F}, \mathbb{P})$ et à valeurs réelles telles que $\mathbb{E}[X^2] < +\infty$.

Soit $\mathcal{G}$ une sous-tribu de $\mathcal{F}$.

- L'**espérance conditionnelle** d'une variable aléatoire $X \in \mathbb{L}^+(\Omega, \mathcal{F}, \mathbb{P})$ (respectivement d'une variable aléatoire $X \in \mathbb{L}^1(\Omega, \mathcal{F}, \mathbb{P})$) **relativement à** $\mathcal{G}$ est l'unique (à une égalité $\mathbb{P}$-presque-sûre près) variable aléatoire $\mathcal{G}$-mesurable, à valeurs positives (respectivement dans $\mathbb{L}^1(\Omega, \mathcal{F}, \mathbb{P})$), notée $\mathbb{E}[X|\mathcal{G}]$, telle que :
$$\forall A \in \mathcal{G}, \mathbb{E}[\mathbf{1}_A Y] = \mathbb{E}[\mathbf{1}_A X]\,. \tag{1}$$
- On notera que si $X \in \mathbb{L}^+(\Omega, \mathcal{F}, \mathbb{P})\,, \mathbb{E}[X|\mathcal{G}] \in \mathbb{L}^1(\Omega, \mathcal{F}, \mathbb{P})$ si et seulement si $X \in \mathbb{L}^1(\Omega, \mathcal{F}, \mathbb{P})$.
- $\mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ est un espace de Hilbert muni du produit scalaire :
$$< Z_1, Z_2 >= \mathbb{E}(Z_1\, Z_2)\,,\ \text{pour tout } (Z_1, Z_2) \in \mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{P}) \times \mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{P})\,.$$
La norme associée au produit scalaire $< .,. >$ vérifie donc : $\forall Z \in \mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{P})\,, ||Z||_2 = (\mathbb{E}(Z\,, Z))^{\frac{1}{2}}$.

  Notons $\mathbb{L}^2(\Omega, \mathcal{G}, \mathbb{P})$ l'ensemble des (classes de) variables aléatoires U définies sur $(\Omega, \mathcal{F}, \mathbb{P})$ et à valeurs dans $\mathbb{R}\,, \mathcal{G}$-mesurables telles que : $\mathbb{E}(U^2) < +\infty$.

  Si $X \in \mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{P})$, alors $\mathbb{E}[X|\mathcal{G}] \in \mathbb{L}^2(\Omega, \mathcal{G}, \mathbb{P})$ est la **projection orthogonale** de X sur le sous-espace fermé $\mathbb{L}^2(\Omega, \mathcal{G}, \mathbb{P})$ de l'espace de Hilbert $\mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{P})$.

  Le cadre géométrique de $\mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ permet de bien comprendre la nature de l'espérance conditionnelle : $\mathbb{E}[X|\mathcal{G}]$ est la meilleure approximation quadratique de X (au sens de la norme $||.||_2$) étant donnée l'information contenue dans $\mathcal{G}$.

Propriétés de l'espérance conditionnelle.

1. **Linéarité** :
$\forall (X_1, X_2) \in \mathbb{L}^+(\Omega, \mathcal{F}, \mathbb{P}) \times \mathbb{L}^+(\Omega, \mathcal{F}, \mathbb{P})\,, \forall (a_1, a_2) \in \mathbb{R}_+^2\,, \mathbb{E}[a_1 X_1 + a_2 X_2|\mathcal{G}] = a_1 \mathbb{E}[X_1|\mathcal{G}] + a_2 \mathbb{E}[X_2|\mathcal{G}]$, p.s.,
$\forall (X_1, X_2) \in \mathbb{L}^1(\Omega, \mathcal{F}, \mathbb{P}) \times \mathbb{L}^1(\Omega, \mathcal{F}, \mathbb{P})\,, \forall (a_1, a_2) \in \mathbb{R}^2\,, \mathbb{E}[a_1 X_1 + a_2 X_2|\mathcal{G}] = a_1 \mathbb{E}[X_1|\mathcal{G}] + a_2 \mathbb{E}[X_2|\mathcal{G}]$, p.s..
2. Si $X \geq 0$, p.s., alors $\mathbb{E}[X|\mathcal{G}] \geq 0$, p.s..
En conséquence, $\mathbb{E}[.|\mathcal{G}]$ est un opérateur **croissant**.
Soit X une variable aléatoire intégrable ou à valeurs positives; on a :
3. $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]] = \mathbb{E}[X]$.
4. Si X est $\mathcal{G}$-mesurable, alors $\mathbb{E}[X|\mathcal{G}] = X$, p.s..
5. Si X est indépendante de $\mathcal{G}$, alors $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X]$, p.s..
6. **Emboîtement** Si $\mathcal{H}$ est une sous-tribu de $\mathcal{G}$, alors :
$$\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}] = \mathbb{E}[X|\mathcal{H}],\ \mathbb{P}-\text{p.s.}\,. \tag{2}$$

7. **Sortir ce qui est connu** Soit X et Z deux v.a. réelles telles que Z est $\mathcal{G}$ - mesurable; on a alors :

$$\mathbb{E}[Z\,X|\mathcal{G}] = Z\,\mathbb{E}[X|\mathcal{G}],\ \mathbb{P}-\text{p.s.}, \tag{3}$$

dans chacun des deux cas suivants :

(a) les v.a. X, Z et $X\,Z$ sont intégrables,

(b) les v.a. X et Z sont positives.

Exercice 1 : Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité fixé.

1. Considérons deux évènements A et B attachés à $(\Omega, \mathcal{F}, \mathbb{P})$.

(a) Rappel :

- **Tribu engendrée par une famille de parties $\mathcal{U}$ d'un ensemble Ω.** Etant donnée $\mathcal{U}$, une famille de parties d'un ensemble Ω; il existe une plus petite (au sens de l'inclusion) tribu sur Ω, notée $\sigma(\mathcal{U})$, qui contient $\mathcal{U}$; $\sigma(\mathcal{U})$ est la tribu engendrée par $\mathcal{U}$.
- **Tribu engendrée par une variable aléatoire X.** Soit $(\Omega, \mathcal{F})$ et $(E, \mathcal{E})$ deux espaces mesurables et une variable aléatoire X définie sur $(\Omega, \mathcal{F})$ et à valeurs dans $(E, \mathcal{E})$.
La tribu engendrée par X , notée $\sigma(X)$, est la plus petite sous-tribu sur $(\Omega, \mathcal{F})$ qui rend X mesurable.
On a : $\sigma(X) = X^{-1}(\mathcal{E}) = (\{X \in C\}\,; C \in \mathcal{E})$.

□

Soit $C \in \mathcal{B}(\mathbb{R})$.
On obtient :

$$(\mathbf{1}_B)^{-1}(C) = \begin{cases} \emptyset, & \text{si } 0 \notin C\,, 1 \notin C \\ B\,, & \text{si } 0 \notin C\,, 1 \in C \\ B^c\,, & \text{si } 0 \in C\,, 1 \notin C \\ \Omega\,, & \text{si } 0 \in C\,, 1 \in C \end{cases}.$$

Ainsi $\sigma(\mathbf{1}_B) = \{\emptyset, B, B^c, \Omega\} = \sigma(\{B\}) = \sigma(\{B^c\})$.

(b) Il sera démontré à la question **2.(b)** que si, $(A_n)_{n\in\mathbb{N}}$ est une partition de l'ensemble Ω , alors toute variable aléatoire réelle X , $\sigma((A_n)_{n\in\mathbb{N}})$ - mesurable, s'écrit : $X = \sum_{n\in\mathbb{N}} \lambda_n \mathbf{1}_{A_n}$, où $\lambda_n \in \mathbb{R}$, pour tout $n \in \mathbb{N}$.
Or, (B, B^c) , lorsque $B \neq \emptyset$, constitue une partition de l'ensemble Ω et $\sigma(\mathbf{1}_B) = \{\emptyset, B, B^c, \Omega\} = \sigma(\{B\}) = \sigma(\{B^c\}) = \sigma((B, B^c))$, d'après le résultat de la question **1.(a)**.
Ainsi les variables aléatoires $\sigma(\mathbf{1}_B)$ - mesurables sont exactement de la forme : $a\mathbf{1}_B + b\mathbf{1}_{B^c}$, avec $(a, b) \in \mathbb{R}^2$.

(c) On se rappellera que $\mathbb{E}[\mathbf{1}_A|\mathbf{1}_B]$ est une notation usuelle pour désigner $\mathbb{E}[\mathbf{1}_A|\sigma(\mathbf{1}_B)]$.
Comme $\mathbb{E}[\mathbf{1}_A|\mathbf{1}_B]$ est une variable aléatoire $\sigma(\mathbf{1}_B)$ - mesurable, on peut trouver, d'après la question précédente, $(a, b) \in \mathbb{R}^2$, tels que $\mathbb{E}[\mathbf{1}_A|\mathbf{1}_B] = a\mathbf{1}_B + b\mathbf{1}_{B^c}$.
Ainsi,

$$\begin{aligned} \mathbb{P}(A \cap B) = \mathbb{E}[\mathbf{1}_{A\cap B}] = \mathbb{E}[\mathbf{1}_A \mathbf{1}_B] &= \mathbb{E}[\mathbb{E}[\mathbf{1}_A \mathbf{1}_B|\mathbf{1}_B]]\,, \text{ en utilisant la propriété } \mathbf{3.} \text{ de l'espérance conditionnelle,} \\ &= \mathbb{E}[\mathbf{1}_B \mathbb{E}[\mathbf{1}_A|\mathbf{1}_B]]\,, \text{ car } \mathbf{1}_B \text{ est } \sigma(\mathbf{1}_B)\text{ - mesurable} \\ &\qquad \text{et d'après la règle “sortir ce qui est connu” rappelée en (3),} \\ &= \mathbb{E}[\mathbf{1}_B(a\mathbf{1}_B + b\mathbf{1}_{B^c})]\,, \\ &= a\mathbb{E}[\mathbf{1}_B]\,, \\ &= a\mathbb{P}(B)\,. \end{aligned}$$

Si $\mathbb{P}(B) \in]0, 1[$, on a alors : $a = \frac{\mathbb{P}(A\cap B)}{\mathbb{P}(B)} = \mathbb{P}(A|B)$.
On montrerait de la même façon que : $b = \mathbb{P}(A|B^c)$.
Il vient alors : $\mathbb{E}[\mathbf{1}_A|\mathbf{1}_B] = \mathbb{P}(A|B)\mathbf{1}_B + \mathbb{P}(A|B^c)\mathbf{1}_{B^c}$, lorsque $\mathbb{P}(B) \in]0, 1[$.

2. Soit $(A_n)_{n\in\mathbb{N}}$ une partition de l'ensemble Ω , c'est-à-dire que, pour tout $n \in \mathbb{N}, A_n \in \mathcal{F}$, $A_n \neq \emptyset$, $A_i \cap A_j = \emptyset$, lorsque $(i, j) \in \mathbb{N}^2$ et $i \neq j$, de plus, $\cup_{n\in\mathbb{N}} A_n = \Omega$.

(a) Posons $\mathcal{G} = \{\cup_{j\in J} A_j \, ; J \subset \mathbb{N}\}$, J décrivant toutes les parties possibles de $\mathbb{N}$, y compris l'ensemble vide, avec la convention que $\cup_{j\in\emptyset} A_j = \emptyset$.
Montrons que $\mathcal{G}$ est une sous-tribu de $\mathcal{F}$.
On a : $\Omega = \cup_{n\in\mathbb{N}} A_n$, prenant $J = \mathbb{N}$, ainsi $\Omega \in \mathcal{G}$.
Soit $B \in \mathcal{G}$. B s'écrit alors : $B = \cup_{j\in J} A_j$, où $J \subset \mathbb{N}$. Ainsi : $B^c = \cup_{k\in J^c} A_k$ et $B^c \in \mathcal{G}$.
Considérons $(B_n)_{n\in\mathbb{N}}$ une suite d'éléments de $\mathcal{G}$. On a : $B_n = \cup_{j\in J_n} A_j$, où $J_n \subset \mathbb{N}\,, n \in \mathbb{N}$.
Posons $J = \cup_{n\in\mathbb{N}} J_n$; il vient : $\cup_{n\in\mathbb{N}} B_n = \cup_{n\in\mathbb{N}} \cup_{j\in J_n} A_j = \cup_{j\in J} A_j$.
Ainsi, $\cup_{n\in\mathbb{N}} B_n \in \mathcal{G}$.
On conclut du développement précédent que $\mathcal{G}$ est une sous-tribu de $\mathcal{F}$.
En prenant $J = \{n\}\,, n \in \mathbb{N}$, il apparaît que, pour tout $n \in \mathbb{N}\,, A_n \in \mathcal{G}$, de sorte que : $(A_n)_{n\in\mathbb{N}} \subset \mathcal{G}$.
Mais, par définition, $\sigma((A_n)_{n\in\mathbb{N}})$ est la plus petite sous-tribu sur $(\Omega, \mathcal{F})$ contenant $(A_n)_{n\in\mathbb{N}}$; ainsi, $\sigma((A_n)_{n\in\mathbb{N}}) \subset \mathcal{G}$.
Par ailleurs, comme $A_m \in \sigma((A_n)_{n\in\mathbb{N}})$, pour tout $m \in \mathbb{N}$, la tribu $\sigma((A_n)_{n\in\mathbb{N}})$ étant stable par réunion dénombrable, $\cup_{j\in J} A_j \in \sigma((A_n)_{n\in\mathbb{N}})$, pour toute partie $J \subset \mathbb{N}$.
On a donc également : $\mathcal{G} \subset \sigma((A_n)_{n\in\mathbb{N}})$.
On conclut que : $\sigma((A_n)_{n\in\mathbb{N}}) = \{\cup_{j\in J} A_j \, ; J \subset \mathbb{N}\}$.

(b) Montrons tout d'abord que si Y est une variable aléatoire réelle $\sigma((A_n)_{n\in\mathbb{N}})$ - mesurable, alors Y est nécessairement constante sur chaque évènement $A_n\,, n \in \mathbb{N}$.
Soit $n \in \mathbb{N}$ et $\omega_0 \in A_n$ ($A_n \neq \emptyset$); posons : $\lambda_n = Y(\omega_0)$.
Comme $\{\lambda_n\} \in \mathcal{B}(\mathbb{R})$ et Y est une variable aléatoire réelle $\sigma((A_n)_{n\in\mathbb{N}})$ - mesurable, $Y^{-1}(\{\lambda_n\}) \in \sigma((A_n)_{n\in\mathbb{N}})$, soit il existe $J \subset \mathbb{N}$ tel que $Y^{-1}(\{\lambda_n\}) = \cup_{j\in J} A_j$, car, d'après la question précédente $\sigma((A_n)_{n\in\mathbb{N}}) = \{\cup_{j\in J} A_j \, ; J \subset \mathbb{N}\}$.
Puisque, $\omega_0 \in Y^{-1}(\{\lambda_n\}) \cap A_n$, on a $\cup_{j\in J} A_j \cap A_n \neq \emptyset$, de sorte que A_n est l'un des A_j , pour $j \in J$; en effet, on a affaire à une partition et les $A_m\,, m \in \mathbb{N}$ sont donc deux à deux distincts.
Ainsi, $Y^{-1}(\{\lambda_n\}) = A_n$ et pour tout $\omega \in A_n\,, Y(\omega) = \lambda_n$.
On en déduit la forme générale que doit avoir une variable aléatoire réelle $\sigma((A_n)_{n\in\mathbb{N}})$ - mesurable : il existe une famille de nombres réels $(\lambda_n)_{n\in\mathbb{N}}$ tels que : $Y = \sum_{n\in\mathbb{N}} \lambda_n \mathbf{1}_{A_n}$. Comme toute variable aléatoire qui a cette forme est $\sigma((A_n)_{n\in\mathbb{N}})$ - mesurable, il vient que l'ensemble des variables aléatoires réelles $\sigma((A_n)_{n\in\mathbb{N}})$ - mesurables est exactement l'ensemble des variables aléatoires s'écrivant :

$$Y = \sum_{n\in\mathbb{N}} \lambda_n \mathbf{1}_{A_n}\,, \text{ avec } \lambda_n \in \mathbb{R}\,.$$

(c) Soit X une variable aléatoire réelle à valeurs positives ou intégrable et cherchons $\mathbb{E}[X|\sigma((A_n)_{n\in\mathbb{N}})]$.
Comme $\mathbb{E}[X|\sigma((A_n)_{n\in\mathbb{N}})]$ est $\sigma((A_n)_{n\in\mathbb{N}})$ - mesurable, il existe une famille de nombres réels $(\lambda_n)_{n\in\mathbb{N}}$ tels que : $\mathbb{E}[X|\sigma((A_n)_{n\in\mathbb{N}})] = \sum_{n\in\mathbb{N}} \lambda_n \mathbf{1}_{A_n}$.
De plus,

$$\begin{aligned}
\mathbb{E}[X\mathbf{1}_{A_n}] &= \mathbb{E}[\mathbb{E}[X\mathbf{1}_{A_n}|\sigma((A_n)_{n\in\mathbb{N}})]]\,, \text{ en utilisant la propriété } \mathbf{3.} \text{ de l'espérance conditionnelle,}\\
&= \mathbb{E}[\mathbf{1}_{A_n}\mathbb{E}[X|\sigma((A_n)_{n\in\mathbb{N}})]]\,, \text{ car } \mathbf{1}_{A_n} \text{ est } \sigma((A_n)_{n\in\mathbb{N}})\text{ - mesurable}\\
&\qquad\qquad\text{et d'après la règle “sortir ce qui est connu” énoncée en (3),}\\
&= \mathbb{E}\left[\mathbf{1}_{A_n}\left(\sum_{m\in\mathbb{N}} \lambda_m \mathbf{1}_{A_m}\right)\right]\,,\\
&= \lambda_n \mathbb{E}[\mathbf{1}_{A_n}]\,,\\
&= \lambda_n \mathbb{P}(A_n)\,.
\end{aligned}$$

On a donc : $\lambda_n = \frac{\mathbb{E}[X\mathbf{1}_{A_n}]}{\mathbb{P}(A_n)}$, lorsque $\mathbb{P}(A_n) \in]0,1[$.
Ainsi,

$$\mathbb{E}[X|\sigma((A_n)_{n\in\mathbb{N}})] = \sum_{i\in I^*} \frac{\mathbb{E}[X\mathbf{1}_{A_i}]}{\mathbb{P}(A_i)} \mathbf{1}_{A_i}\,, \text{ où } I^* = \{i \in I; \mathbb{P}(A_i) > 0\}\,. \tag{4}$$

En particulier pour $X = \mathbf{1}_A$, il vient :

$$\mathbb{E}[\mathbf{1}_A|\sigma((A_n)_{n\in\mathbb{N}})] = \sum_{i\in I^*} \frac{\mathbb{P}(A \cap A_i)}{\mathbb{P}(A_i)} \mathbf{1}_{A_i} = \sum_{i\in I^*} \mathbb{P}(A|A_i)\mathbf{1}_{A_i}\,, \text{ avec } I^* = \{i \in I; \mathbb{P}(A_i) > 0\}\,.$$

3. (a) Rappelons qu'une variable aléatoire réelle Y définie sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ est dite **discrète** s'il existe un ensemble dénombrable $E \subset \mathbb{R}$ tel que $\mathbb{P}(Y \in E) = 1$.
Comme E est dénombrable, il est possible de trouver un ensemble d'indices $I\,, I \subset \mathbb{N}$ et $E = \{(b_i)_{i\in I}\}$.

La famille d'évènements $(A_i = Y^{-1}(b_i) = \{\omega \in \Omega; Y(\omega) = b_i\})_{i\in I}$ forme une partition de l'ensemble Ω et on a : $Y = \sum_{i\in I} b_i \, \mathbf{1}_{A_i}$.
La tribu engendrée par Y, $\sigma(Y)$, est la plus petite sous-tribu $\mathcal{G}$ sur $(\Omega, \mathcal{F})$ telle que Y soit $\mathcal{G}$-mesurable; ainsi, $\sigma(Y) = \sigma((\{Y = b_i\})_{i\in I})$.
Si X est une variable aléatoire à valeurs positives ou intégrable, on a alors, d'après (4) :

$$\mathbb{E}[X|Y] = \sum_{i\in I^*} \frac{\mathbb{E}[X\mathbf{1}_{\{Y=b_i\}}]}{\mathbb{P}(Y = b_i)} \mathbf{1}_{\{Y=b_i\}}, \ \mathbb{P} - \text{p.s.} , \tag{5}$$

où $I^* = \{i \in I; \mathbb{P}(Y = b_i) > 0\}$.

(b) Supposons que X soit une variable aléatoire suivant une loi de Poisson de paramètre $\lambda \in \mathbb{R}_+^*$.
Ainsi, X prend ses valeurs dans $\mathbb{N}$ et pour tout $k \in \mathbb{N}$:

$$\mathbb{P}(X = k) = e^{-\lambda} \frac{\lambda^k}{k!} \tag{6}$$

Posons $Y = 2\lfloor \frac{X}{2} \rfloor$, où $\lfloor x \rfloor$ désigne la partie entière de $x \in \mathbb{R}$.
Y s'écrit sous la forme : $Y = f(X)$ où $f \ : \mathbb{N} \to \mathbb{N}$ définie par : $f(x) = x \mapsto 2\lfloor \frac{x}{2} \rfloor$ est une fonction borélienne; Y est donc $\sigma(X)$-mesurable et $\mathbb{E}[Y|X] = Y$, en utilisant la propriété **4.** de l'espérance conditionnelle du rappel de cours liminaire.
Par ailleurs, remarquons que, pour tout $n \in \mathbb{N}, \{Y = 2n\} = \{X = 2n\} \cup \{X = 2n + 1\}$ et $\{X = 2n\}$ et $\{X = 2n + 1\}$ sont deux évènements incompatibles, de sorte que :

$$\begin{aligned}\mathbb{P}(Y = 2n) &= \mathbb{P}(X = 2n) + \mathbb{P}(X = 2n + 1) \\ &= e^{-\lambda} \frac{\lambda^{2n}}{(2n)!} + e^{-\lambda} \frac{\lambda^{2n+1}}{(2n + 1)!} \ , \text{ d'après (6)} \\ &> 0\end{aligned} \tag{7}$$

Appliquant la formule obtenue en (5), il vient alors :

$$\mathbb{E}[X|Y] = \sum_{n\in\mathbb{N}} \frac{\mathbb{E}[X\mathbf{1}_{\{Y=2n\}}]}{\mathbb{P}(Y = 2n)} \mathbf{1}_{\{Y=2n\}}, \ \mathbb{P} - \text{p.s.} , \tag{8}$$

Par ailleurs, quel que soit $n \in \mathbb{N}$, on a :

$$\begin{aligned}\mathbb{E}[X\mathbf{1}_{\{Y=2n\}}] &= \mathbb{E}[X\mathbf{1}_{\{X=2n\}}] + \mathbb{E}[X\mathbf{1}_{\{X=2n+1\}}] \\ &= 2n \, \mathbb{P}(X = 2n) + (2n + 1) \, \mathbb{P}(X = 2n + 1) \\ &= 2n \, e^{-\lambda} \frac{\lambda^{2n}}{(2n)!} + (2n + 1) \, e^{-\lambda} \frac{\lambda^{2n+1}}{(2n + 1)!}\end{aligned} \tag{9}$$

Combinant (7) et (9), on trouve, pour tout $n \in \mathbb{N}$:

$$\begin{aligned}\frac{\mathbb{E}[X\mathbf{1}_{\{Y=2n\}}]}{\mathbb{P}(Y = 2n)} &= \frac{e^{-\lambda} \frac{\lambda^{2n}}{(2n-1)!} + e^{-\lambda} \frac{\lambda^{2n+1}}{(2n)!}}{e^{-\lambda} \frac{\lambda^{2n}}{(2n)!} + e^{-\lambda} \frac{\lambda^{2n}}{(2n+1)!}} \\ &= \frac{\frac{\lambda^{2n}(2n+\lambda)}{(2n)!}}{\frac{\lambda^{2n}((2n+1)+\lambda)}{(2n+1)!}} \\ &= \frac{(2n + \lambda)(2n + 1)}{2n + 1 + \lambda} \, .\end{aligned} \tag{10}$$

Utilisant (8) et (10), on obtient finalement :

$$\mathbb{E}[X|Y] = \sum_{n\in\mathbb{N}} \frac{(2n + \lambda)(2n + 1)}{2n + 1 + \lambda} \mathbf{1}_{\{Y=2n\}} = \frac{(Y + \lambda)(Y + 1)}{Y + 1 + \lambda}, \ \mathbb{P} - \text{p.s.} .$$

Exercice 2 : Rappel :

Théorème de Tonelli-Fubini Si f est une application $(\mathcal{B}(\mathbb{R}^2), \mathcal{B}(\bar{\mathbb{R}}_+))$-mesurable de $\mathbb{R}^2$ dans $\bar{\mathbb{R}}_+ = [0, +\infty]$, alors :

1. Pour tout $x \in \mathbb{R}$ (resp^t pour tout $y \in \mathbb{R}$), l'application $y \mapsto f(x, y)$ (resp^t l'application $x \mapsto f(x, y)$) est $(\mathcal{B}(\mathbb{R}), \mathcal{B}(\bar{\mathbb{R}}_+))$-mesurable (resp^t est $(\mathcal{B}(\mathbb{R}), \mathcal{B}(\bar{\mathbb{R}}_+))$-mesurable).
2. Les applications $y \in \mathbb{R} \mapsto \int_{\mathbb{R}} f(x, y)dx \in [0, +\infty]$ et $x \in \mathbb{R} \mapsto \int_{\mathbb{R}} f(x, y)dy \in [0, +\infty]$ sont $(\mathcal{B}(\mathbb{R}), \mathcal{B}(\bar{\mathbb{R}}_+))$-mesurables.
3. L'intégrale de f sur $\mathbb{R}^2$ est donnée par :

$$\begin{aligned}\int_{\mathbb{R}^2} f(x, y)\, dxdy &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x\,, y)dy \right) dx\,, \\ &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x\,, y)dx \right) dy\,.\end{aligned}$$

Dans le cas où la fonction numérique à intégrer est à valeurs dans $\mathbb{R}$, donc de signe quelconque, on a un théorème analogue au précédent à condition de se limiter aux fonctions intégrables.

Théorème de Fubini Si f est une application intégrable de $\mathbb{R}^2$ dans $\mathbb{R}$, alors :

1. Pour presque tout $x \in \mathbb{R}$ (resp^t pour presque tout $y \in \mathbb{R}$), l'application $y \mapsto f(x, y)$ (resp^t l'application $x \mapsto f(x, y)$) est $(\mathcal{B}(\mathbb{R}), \mathcal{B}(\mathbb{R}))$-mesurable et intégrable (resp^t est $(\mathcal{B}(\mathbb{R}), \mathcal{B}(\mathbb{R}))$-mesurable et intégrable).
2. Les applications $y \in \mathbb{R} \mapsto \int_{\mathbb{R}} f(x, y)dx \in \mathbb{R}$ et $x \in \mathbb{R} \mapsto \int_{\mathbb{R}} f(x, y)dy \in \mathbb{R}$ sont définies presque partout, sont $(\mathcal{B}(\mathbb{R}), \mathcal{B}(\mathbb{R}))$-mesurables et intégrables.
3. L'intégrale de f sur $\mathbb{R}^2$ est donnée par :

$$\begin{aligned}\int_{\mathbb{R}^2} f(x, y)\, dxdy &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x\,, y)dy \right) dx\,, \\ &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x\,, y)dx \right) dy\,.\end{aligned}$$

□

1. X admet pour densité la fonction $p\ :\ x \mapsto p(x) = \int_{\mathbb{R}} f_{X,Y}(x, y)dy \in [0, +\infty]$.
   En toute rigueur, il faut prendre $p(x) = 0$ pour les valeurs de $x \in \mathbb{R}$ telles que $\int_{\mathbb{R}} f_{X,Y}(x, y)dy = +\infty$ qui forment un ensemble de mesure nulle; nous négligerons cependant ce point de détail dans les calculs qui suivent.
   D'après le théorème de Tonelli-Fubini, la fonction $x \mapsto p(x) = \int_{\mathbb{R}} f_{X,Y}(x, y)dy$ est $(\mathcal{B}(\mathbb{R}), \mathcal{B}(\bar{\mathbb{R}}_+))$-mesurable.
   Par ailleurs, comme $h(X, Y)$ est intégrable, $q\ :\ x \mapsto \int_{\mathbb{R}} h(x, y)f_{X,Y}(x, y)dy$ est définie pour presque tout $x \in \mathbb{R}$ et est $(\mathcal{B}(\mathbb{R}), \mathcal{B}(\mathbb{R}))$-mesurable, d'après le théorème de Fubini.
   On en déduit que la fonction $g\ :\begin{cases} \frac{q(x)}{p(x)} & ,\text{ si } p(x) > 0 \\ 0 & ,\text{ si } p(x) = 0 \end{cases}$ est $(\mathcal{B}(\mathbb{R}), \mathcal{B}(\mathbb{R}))$-mesurable.
   Comme X est également $(\mathcal{B}(\mathbb{R}), \mathcal{B}(\mathbb{R}))$-mesurable, $g(X)$ est $(\mathcal{B}(\mathbb{R}), \mathcal{B}(\mathbb{R}))$-mesurable comme étant la composée de deux fonctions $(\mathcal{B}(\mathbb{R}), \mathcal{B}(\mathbb{R}))$-mesurables.
   De plus, on a :

$$\begin{aligned}\mathbb{E}[|g(X)|] &= \int_{\mathbb{R}} |g(x)| \left(\int_{\mathbb{R}} f_{X,Y}(x, z)dz \right) \mathbf{1}_{\{p(x)>0\}} dx\,,\ \text{car } X \text{ admet pour densité } p(x) = \int_{\mathbb{R}} f_{X,Y}(x, z)dz\,, \\ &\le \int_{\mathbb{R}} \left(\frac{\int_{\mathbb{R}} |h(x, y)| f_{X,Y}(x, y)dy}{\int_{\mathbb{R}} f_{X,Y}(x, y)dy} \right) \left(\int_{\mathbb{R}} f_{X,Y}(x, z)dz \right) \mathbf{1}_{\{p(x)>0\}} dx\,, \\ &= \int_{\mathbb{R}^2} |h(x, y)| f_{X,Y}(x, y)dxdy\,,\ \text{d'après le théorème de Tonelli-Fubini}, \\ &= \mathbb{E}[|h(X, Y)|] < +\infty\,.\end{aligned}$$

   On en déduit que $g(X)$ est alors une variable aléatoire intégrable.

Rappel : **Caractérisation 2 de l'espérance conditionnelle**.
Etant donné un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$, $\mathcal{G}$ une sous-tribu de $\mathcal{F}$ et X une variable aléatoire intégrable. L'espérance conditionnelle de X sachant $\mathcal{G}$, notée $\mathbb{E}[X|\mathcal{G}]$, est l'unique (au sens $\mathbb{P}$-presque sûr) variable aléatoire $\mathcal{G}$ −mesurable et intégrable telle que :

$$\text{Pour toute variable aléatoire } U\,, \mathcal{G} - \text{mesurable et bornée, } \mathbb{E}[U\mathbb{E}[X|\mathcal{G}]] = \mathbb{E}[UX]\,. \tag{11}$$

□

Soit U une variable aléatoire bornée $\sigma(X)$-mesurable, donc de la forme $\psi(X)$, où $\psi\ : \mathbb{R} \to \mathbb{R}$ est une fonction $(\mathcal{B}(\mathbb{R}), \mathcal{B}(\mathbb{R}))$-mesurable et bornée.
De plus, il vient :

$$\begin{aligned}
\mathbb{E}[\psi(X)g(X)] &= \int_{\mathbb{R}} \psi(x)g(x)\left(\int_{\mathbb{R}} f_{X,Y}(x,z)dz\right)\mathbf{1}_{\{p(x)>0\}}dx\,,\\
&= \int_{\mathbb{R}} \psi(x)\left(\frac{\int_{\mathbb{R}} h(x,y)f_{X,Y}(x,y)dy}{\int_{\mathbb{R}} f_{X,Y}(x,y)dy}\right)\left(\int_{\mathbb{R}} f_{X,Y}(x,z)dz\right)\mathbf{1}_{\{p(x)>0\}}dx\,,\\
&= \int_{\mathbb{R}^2} \psi(x)h(x,y)f_{X,Y}(x,y)dxdy\,, \text{ d'après le théorème de Fubini,}\\
&= \mathbb{E}[\psi(X)h(X,Y)]
\end{aligned} \tag{12}$$

Or, d'après la caractérisation 2 de l'espérance conditionnelle (cf (11)), $\mathbb{E}[h(X,Y)|X]$ est l'unique (au sens $\mathbb{P}$- presque sûr) variable aléatoire $\sigma(X)$-mesurable et intégrable vérifiant (12).
On en déduit que :

$$\mathbb{E}[h(X,Y)|X] = g(X), \quad \mathbb{P} - \text{p.s.}\,, \tag{13}$$

avec :

$$\forall x \in \mathbb{R}, \quad g(x) = \begin{cases} \frac{\int_{\mathbb{R}} h(x,y)f_{X,Y}(x,y)dy}{\int_{\mathbb{R}} f_{X,Y}(x,y)dy}\,, & \text{si } \int_{\mathbb{R}} f_{X,Y}(x,y)dy \neq 0,\\ 0\,, & \text{dans le cas contraire.}\end{cases} \tag{14}$$

2. Supposons que (X,Y) soit un couple de variables aléatoires réelles admettant la densité sur $\mathbb{R}^2$ définie, pour tout $(x,y) \in \mathbb{R}^2$, par :

$$f_{X,Y}(x,y) = n(n-1)(y-x)^{n-2}\,\mathbf{1}_{\Delta}(x,y)\,,$$

où $\Delta = \{(x,y) \in \mathbb{R}^2\,; 0 \leq x \leq y \leq 1\}$ et $n \in \mathbb{N}^*$.

(a) X admet une densité $f_X(x)$ donnée par :

$$\begin{aligned}
f_X(x) = \int_{\mathbb{R}} f_{X,Y}(x,y)\,dy &= \left(\int_x^1 n(n-1)(y-x)^{n-2}\,dy\right)\mathbf{1}_{0\leq x\leq 1}\\
&= n\Big[(y-x)^{n-1}\Big]_x^1 \mathbf{1}_{0\leq x\leq 1}\\
&= n(1-x)^{n-1}\mathbf{1}_{0\leq x\leq 1}
\end{aligned} \tag{15}$$

Il apparaît alors que X suit la loi bêta de paramètre $(n,1)$.
Par ailleurs,

$$\begin{aligned}
\int_{\mathbb{R}} y\,f_{X,Y}(x,y)dy &= \left(\int_x^1 n(n-1)y\,(y-x)^{n-2}\,dy\right)\mathbf{1}_{0\leq x\leq 1}\\
&= \left(\Big[n\,y(y-x)^{n-1}\Big]_x^1 - \int_x^1 n\,(y-x)^{n-1}\,dy\right)\mathbf{1}_{0\leq x\leq 1}, \text{ en intégrant par parties,}\\
&= \left(n\,(1-x)^{n-1} - \Big[(y-x)^n\Big]_x^1\right)\mathbf{1}_{0\leq x\leq 1}\\
&= \big(n\,(1-x)^{n-1} - (1-x)^n\big)\,\mathbf{1}_{0\leq x\leq 1}\\
&= (1-x)^{n-1}(n-1+x)\mathbf{1}_{0\leq x\leq 1}
\end{aligned} \tag{16}$$

D'après (13), (14), (15) et (17), on a :

$$\mathbb{E}[Y|X] = g(X), \quad \mathbb{P} - \text{p.s.}\,, \tag{17}$$

avec :

$$\forall x \in \mathbb{R}, \quad g(x) = \frac{(1-x)^{n-1}(n-1+x)}{n(1-x)^{n-1}} = \frac{n-1+x}{n}.$$

On en déduit que :

$$\mathbb{E}[Y|X] = \frac{n-1+X}{n}, \quad \mathbb{P}-\text{p.s.}. \tag{18}$$

(b) On a alors :

$$\begin{aligned}
\mathbb{E}[Y] &= \mathbb{E}[\mathbb{E}[Y|X]]\,, \text{ d'après la propriété } \mathbf{3.} \text{ de l'espérance conditionnelle,} && (19)\\
&= \frac{n-1+\mathbb{E}[X]}{n}, \text{ en utilisant (18),} && (20)
\end{aligned}$$

Or, d'après (15), il vient :

$$\begin{aligned}
\mathbb{E}[X] &= \int_{\mathbb{R}} x\, f_X(x)\, dx\\
&= \int_0^1 x\, n(1-x)^{n-1}\, dx, \text{ d'après (15),}\\
&= \Big[-x(1-x)^n\Big]_0^1 + \int_0^1 (1-x)^n\\
&= -\Big[\frac{(1-x)^n}{n+1}\Big]_0^1\\
&= \frac{1}{n+1} && (21)
\end{aligned}$$

Combinant (20) et (21), on obtient :

$$\mathbb{E}[Y] = \frac{n}{n+1}\,.$$

<u>Exercice 3</u> : 1. $\psi(Y)$ étant une v.a. $\sigma(Y)$-mesurable et intégrable (car ϕ est bornée), il suffit de vérifier (d'après la caractérisation 2 de l'espérance conditionnelle) que pour toute v.a. $\sigma(Y)$-mesurable et bornée, donc de la forme $h(Y)$, où h est une fonction borélienne bornée, on a :

$$\mathbb{E}[\psi(Y)h(Y)] = \mathbb{E}[\phi(X,Y)h(Y)]\,.$$

Or :

$$\mathbb{E}[\psi(Y)h(Y)] = \int_{\mathbb{R}} \psi(y)h(y)f_Y(y)dy\,,$$

et, puisque $\psi(y) = \mathbb{E}[\phi(X,y)] = \int_{\mathbb{R}} \phi(x,y)f_X(x)dx$, pour tout $y \in \mathbb{R}$, il vient :

$$\mathbb{E}[\psi(Y)h(Y)] = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \phi(x,y)f_X(x)dx\right) h(y)f_Y(y)dy\,.$$

Utilisant enfin le théorème de Fubini, on obtient :

$$\mathbb{E}[\psi(Y)h(Y)] = \int_{\mathbb{R}^2} \phi(x,y)h(y)f_X(x)f_Y(y)dx\, dy = \mathbb{E}[\phi(X,Y)h(Y)]\,,$$

puisque, comme X et Y sont indépendantes, $\forall (x,y) \in \mathbb{R}^2\,, f_{X,Y}(x,y) = f_X(x)f_Y(y)$, en dehors d'un ensemble négligeable.

On en déduit que :

$$\mathbb{E}[\phi(X,Y)|Y] = \psi(Y), \quad \mathbb{P}-\text{p.s.}\,, \tag{22}$$

où, pour tout $y \in \mathbb{R}$:

$$\psi(y) = \mathbb{E}[\phi(X,y)]\,. \tag{23}$$

2. Soit X et Y deux variables aléatoires réelles indépendantes; X est une variable aléatoire continue de densité notée f_X et Y suit une loi normale $\mathcal{N}(0,1)$.

(a) Montrons que les propriétés suivantes sont équivalentes :

$$\text{(P1)} \quad e^{\frac{X^2}{2}} \text{ est intégrable} \quad \text{(P2)} \quad e^{XY} \text{ est intégrable} \quad \text{(P3)} \quad e^{|XY|} \text{ est intégrable}$$

On a : $|XY| \leq XY$, $\mathbb{P}-$p.s. et la fonction $x \mapsto e^x$ est croissante de sorte que : $e^{XY} \leq e^{|XY|}$, $\mathbb{P}-$p.s..
Si $e^{|XY|}$ est intégrable, alors e^{XY} l'est aussi.
Ainsi, (P3) implique (P2).

Supposons que e^{XY} soit intégrable. Alors, d'après le théorème de Fubini, il vient :

$$
\begin{aligned}
\mathbb{E}[e^{XY}] &= \int_{\mathbb{R}} \left(\int_{-\infty}^{+\infty} e^{xy} \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}}\, dy \right) f_X(x)\, dx, \\
&= \int_{\mathbb{R}} e^{\frac{x^2}{2}} \left(\int_{-\infty}^{+\infty} \frac{e^{-\frac{(y-x)^2}{2}}}{\sqrt{2\pi}}\, dy \right) f_X(x)\, dx, \\
&= \int_{\mathbb{R}} e^{\frac{x^2}{2}} \left(\int_{-\infty}^{+\infty} \frac{e^{-\frac{u^2}{2}}}{\sqrt{2\pi}}\, du \right) f_X(x)\, dx \\
&= \int_{\mathbb{R}} e^{\frac{x^2}{2}}\, f_X(x)\, dx \\
&= \mathbb{E}[e^{\frac{X^2}{2}}]
\end{aligned} \tag{24}
$$

On en déduit que (P1) et (P2) sont équivalentes.
Supposons maintenant que $e^{|XY|}$ soit intégrable. Utilisant à nouveau le théorème de Fubini, on a :

$$
\begin{aligned}
\mathbb{E}[e^{|XY|}] &= \int_{\mathbb{R}} \left(\int_{-\infty}^{+\infty} e^{|xy|} \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}}\, dy \right) f_X(x)\, dx \\
&= \int_{-\infty}^{0} \left(\int_{-\infty}^{+\infty} e^{-x|y|} \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}}\, dy \right) f_X(x)\, dx + \int_{0}^{+\infty} \left(\int_{-\infty}^{+\infty} e^{x|y|} \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}}\, dy \right) f_X(x)\, dx,
\end{aligned} \tag{25}
$$

Par ailleurs, pour tout $x \in \mathbb{R}^+$:

$$
\begin{aligned}
\int_{-\infty}^{+\infty} e^{x|y|} \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}}\, dy &= \int_{-\infty}^{0} e^{-xy} \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}}\, dy + \int_{0}^{+\infty} e^{xy} \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}}\, dy \\
&= e^{\frac{x^2}{2}} \int_{-\infty}^{0} \frac{e^{-\frac{(y+x)^2}{2}}}{\sqrt{2\pi}}\, dy + e^{\frac{x^2}{2}} \int_{0}^{+\infty} \frac{e^{-\frac{(y-x)^2}{2}}}{\sqrt{2\pi}}\, dy \\
&= e^{\frac{x^2}{2}} \int_{-\infty}^{0} \frac{e^{-\frac{u^2}{2}}}{\sqrt{2\pi}}\, du + e^{\frac{x^2}{2}} \int_{0}^{+\infty} \frac{e^{-\frac{u^2}{2}}}{\sqrt{2\pi}}\, du \\
&\leq e^{\frac{x^2}{2}} \int_{-\infty}^{+\infty} \frac{e^{-\frac{u^2}{2}}}{\sqrt{2\pi}}\, du + e^{\frac{x^2}{2}} \int_{-\infty}^{+\infty} \frac{e^{-\frac{u^2}{2}}}{\sqrt{2\pi}}\, du \\
&= 2\, e^{\frac{x^2}{2}}
\end{aligned} \tag{26}
$$

On obtient de la même façon que, pour tout $x \in \mathbb{R}^-$:

$$
\int_{-\infty}^{+\infty} e^{-x|y|} \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}}\, dy \leq 2\, e^{\frac{x^2}{2}} \tag{27}
$$

Combinant (25), (26) et (27), il vient :

$$
\mathbb{E}[e^{|XY|}] \leq 2\, \mathbb{E}[e^{\frac{X^2}{2}}],
$$

de sorte que : (P1) implique (P3).

(b) Supposons que $e^{\frac{X^2}{2}}$ soit intégrable; alors, d'après la question précédente, e^{XY} est intégrable.
Par ailleurs, la fonction $x \mapsto e^x$ étant convexe, il résulte de l'inégalité de Jensen conditionnelle que :

$$
\begin{aligned}
\mathbb{E}[e^{XY}|X] &\geq e^{\mathbb{E}[XY|X]}, \quad \mathbb{P}-\text{p.s.} \\
&= e^{X\mathbb{E}[Y|X]},\ \mathbb{P}-\text{p.s.}\,,\ \text{car } X \text{ est } \sigma(X)\text{-mesurable et en utilisant la règle “sortir ce qui est connu”}, \\
&= e^{X\mathbb{E}[Y]},\ \mathbb{P}-\text{p.s.}\,,\ \text{puisque } X \text{ et } Y \text{ sont indépendantes}, \\
&= 1, \quad \mathbb{P}-\text{p.s.}\,.
\end{aligned}
$$

(c) Si $e^{\frac{X^2}{2}}$ est intégrable, alors e^{XY} l'est aussi.
Comme X et Y sont indépendantes, on déduit de (22) et (23) que :

$$
\mathbb{E}[e^{XY}|X] = \psi(X), \quad \mathbb{P}-\text{p.s.}\,, \tag{28}
$$

où, pour tout $x \in \mathbb{R}$:

$$\psi(x) = \mathbb{E}[e^{xY}]\,. \tag{29}$$

Par ailleurs :

$$\begin{aligned}
\mathbb{E}[e^{xY}] &= \int_{-\infty}^{+\infty} e^{xy}\,\frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}}\,dy \\
&= e^{\frac{x^2}{2}}\left(\int_{-\infty}^{+\infty} \frac{e^{-\frac{(y-x)^2}{2}}}{\sqrt{2\pi}}\,dy\right), \\
&= e^{\frac{x^2}{2}}\left(\int_{-\infty}^{+\infty} \frac{e^{-\frac{u^2}{2}}}{\sqrt{2\pi}}\,du\right) \\
&= e^{\frac{x^2}{2}}
\end{aligned}$$

On conclut du calcul précédent que : $\mathbb{E}[e^{XY}|X] = e^{\frac{X^2}{2}}, \quad \mathbb{P}-\text{p.s.}\,.$

Exercice 4 : Soit (X,Y) un vecteur gaussien centré. On notera σ_1^2 la variance de X et σ_2^2 la variance de Y supposée strictement positive. ρ désignera le coefficient de corrélation de X et Y :

$$\rho = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}\,.$$

Rappel : **Vecteurs gaussiens**

- Un vecteur aléatoire $(X_0,\cdots,X_n)$ est dit **gaussien**, si pour tout $(u_0,\ldots,u_n) \in \mathbb{R}^{n+1}$, $\sum_{k=0}^{n} u_k X_k$ est une variable aléatoire gaussienne. Choisissant $u_k = 1$, quel que soit $k \in \{0,\ldots,n\}$ et $u_j = 0$, pour tout $j \in \{0,\ldots,n\}$ tel que $j \neq k$, X_k est alors une variable aléatoire gausienne.
- Si X est un vecteur aléatoire à valeurs dans $\mathbb{R}^d$, $d \geq 1$ et $Y = a + M\,X$, où $a \in \mathbb{R}^n$, $n \geq 1$ et M est une matrice à coefficients réels de taille $n \times d$, alors toute combinaison linéaire des coordonnées de X est une combinaison linéaire des coordonnées de Y à une constante près. Ainsi, si X est gaussien, Y l'est aussi et on obtient **la stabilité du caractère gaussien d'un vecteur aléatoire par transformation linéaire**.
- Si deux variables aléatoires X et Y à valeurs respectivement dans $\mathbb{R}^m$ et $\mathbb{R}^d$ forment un couple (X,Y) gaussien, elles sont indépendantes si et seulement si $\mathrm{Cov}(X_i,Y_j) = \mathbb{E}[X_iY_j] - \mathbb{E}[X_i]\mathbb{E}[Y_j] = 0$, pour tout $(i,j) \in \{1,\cdots,m\} \times \{1,\cdots,d\}$.

$\square$

1. Cherchons $\lambda \in \mathbb{R}$ tel que la variable aléatoire $Z = X - \lambda Y$ soit indépendante de Y.
   Comme $\begin{pmatrix} Z \\ Y \end{pmatrix} = \begin{pmatrix} 1 & -\lambda \\ 0 & 1 \end{pmatrix}\begin{pmatrix} X \\ Y \end{pmatrix}$, le couple $(Z,Y) = (X-\lambda Y, Y)$ est obtenu par transformation linéaire à partir du couple gaussien centré (X,Y), il est également gaussien centré.
   Ainsi Z est indépendant de Y si et seulement si $\mathrm{Cov}(Z,Y) = \mathbb{E}[ZY] = 0$ soit : $\mathbb{E}[XY] = \lambda\mathbb{E}[Y^2]$.
   Posons $\lambda_0 = \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(Y)} = \rho\frac{\sigma_1}{\sigma_2}$.
   $Z = X - \lambda_0 Y$ est alors indépendante de Y.
2. D'après la linéarité de l'espérance conditionnelle, $\mathbb{E}[X|Y] = \mathbb{E}[(Z+\lambda_0 Y)|Y] = \mathbb{E}[Z|Y] + \lambda_0\mathbb{E}[Y|Y]$.
   Or, $\mathbb{E}[Y|Y] = Y$, puisque Y est $\sigma(Y)-$ mesurable et $\mathbb{E}[Z|Y] = \mathbb{E}[Z]$, car Z est, par construction, indépendante de Y, c'est-à-dire, de $\sigma(Y)$.
   Rappel : **Indépendance de deux variables aléatoires**
   - Deux variables aléatoires réelles X et Y définies sur $(\Omega,\mathcal{F},\mathbb{P})$ sont indépendantes si et seulement si, pour toutes applications f et g boréliennes bornées ou à valeurs positives :

     $$\mathbb{E}[f(X)g(Y)] = \mathbb{E}[f(X)]\mathbb{E}[g(Y)]\,.$$

   - Soit $X\ : (\Omega,\mathcal{F}) \to (\mathbb{R},\mathcal{B}(\mathbb{R}))$ une variable aléatoire $\mathcal{F}$-mesurable et $\mathcal{G} \subset \mathcal{F}$ une sous-tribu de $\mathcal{F}$. On dit que X est indépendante de $\mathcal{G}$ si, pour toute variable aléatoire Y, $\mathcal{G}$-mesurable, X est indépendante de Y.

- Si X et Y sont deux variables aléatoires définies sur $(\Omega, \mathcal{F})$ et à valeurs dans $\mathbb{R}$, comme les variables aléatoires Z, $\sigma(Y)$-mesurables sont celles qui s'écrivent sous la forme $g(Y)$, avec $g : \mathbb{R} \to \mathbb{R}$, borélienne (soit $(\mathcal{B}(\mathbb{R}), \mathcal{B}(\mathbb{R}))$-mesurable), dire que X est indépendante de $\sigma(Y)$ équivaut à dire que X et Y sont indépendantes.

$\square$

De plus, X et Y étant centrées, $\mathbb{E}[Z] = \mathbb{E}[X] - \lambda_0 \mathbb{E}[Y] = 0$.
Ainsi :

$$\mathbb{E}[X|Y] = \lambda_0 Y = \rho \frac{\sigma_1}{\sigma_2} Y \text{ p.s.}. \tag{30}$$

3. Comme le couple (X, Y) est un vecteur gaussien, $Z = X - \lambda_0 Y$ est une variable aléatoire gaussienne. Z est centrée et :

$$\begin{aligned}
\operatorname{Var}(Z) = \mathbb{E}[(X - \lambda_0 Y)^2] &= \operatorname{Var}(X) + \lambda_0^2 \operatorname{Var}(Y) - 2\lambda_0 \operatorname{Cov}(X, Y), \\
&= \operatorname{Var}(X) + \rho^2 \frac{\sigma_1^2}{\sigma_2^2} \sigma_2^2 - 2\rho \frac{\sigma_1}{\sigma_2} \rho \sigma_1 \sigma_2, \\
&= \operatorname{Var}(X)(1 - \rho^2).
\end{aligned}$$

On en déduit, d'après (30) que :

$$X = \lambda_0 Y + Z = \mathbb{E}[X|Y] + Z \text{ p.s.},$$

où Z est une variable aléatoire gaussienne centrée indépendante de Y et de variance donnée par :

$$\operatorname{Var}(Z) = (1 - \rho^2)\sigma_1^2.$$

<u>Exercice 4</u> : Soit X une variable aléatoire de carré intégrable définie sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ et $\mathcal{G}$ une sous-tribu de $\mathcal{F}$. On pose :

$$\operatorname{Var}(X|\mathcal{G}) = \mathbb{E}\left[(X - \mathbb{E}[X|\mathcal{G}])^2 | \mathcal{G}\right].$$

D'après la linéarité de l'espérance conditionnelle, il vient :

$$\mathbb{E}\left[(X - \mathbb{E}[X|\mathcal{G}])^2 | \mathcal{G}\right] = \mathbb{E}[X^2|\mathcal{G}] - 2\mathbb{E}[X\, \mathbb{E}[X|\mathcal{G}]|\mathcal{G}] + \mathbb{E}[(\mathbb{E}[X|\mathcal{G}])^2|\mathcal{G}]. \tag{31}$$

Or, $\mathbb{E}[X|\mathcal{G}]$ étant $\mathcal{G}$-mesurable, on a :

$$\mathbb{E}[X\, \mathbb{E}[X|\mathcal{G}]|\mathcal{G}] = \mathbb{E}[X|\mathcal{G}]\, \mathbb{E}[X|\mathcal{G}] = (\mathbb{E}[X|\mathcal{G}])^2, \tag{32}$$

et, comme $(\mathbb{E}[X|\mathcal{G}])^2$ est encore $\mathcal{G}$-mesurable, on obtient :

$$\mathbb{E}[(\mathbb{E}[X|\mathcal{G}])^2|\mathcal{G}] = (\mathbb{E}[X|\mathcal{G}])^2. \tag{33}$$

Combinant (31), (32) et (33), on conclut que :

$$\operatorname{Var}(X|\mathcal{G}) = \mathbb{E}\left[(X - \mathbb{E}[X|\mathcal{G}])^2|\mathcal{G}\right] = \mathbb{E}[X^2|\mathcal{G}] - (\mathbb{E}[X|\mathcal{G}])^2. \tag{34}$$

En prenant l'espérance dans (34), on a :

$$\begin{aligned}
\mathbb{E}[\operatorname{Var}(X|\mathcal{G})] &= \mathbb{E}[\mathbb{E}[X^2|\mathcal{G}]] - \mathbb{E}[(\mathbb{E}[X|\mathcal{G}])^2], \\
&= \mathbb{E}[X^2] - \mathbb{E}[(\mathbb{E}[X|\mathcal{G}])^2], \\
&= \mathbb{E}[X^2] - (\mathbb{E}[X])^2 + (\mathbb{E}[X])^2 - \mathbb{E}[(\mathbb{E}[X|\mathcal{G}])^2], \\
&= \operatorname{Var}(X) + (\mathbb{E}[X])^2 - \mathbb{E}[(\mathbb{E}[X|\mathcal{G}])^2].
\end{aligned} \tag{35}$$

Or,

$$\begin{aligned}
\operatorname{Var}(\mathbb{E}[X|\mathcal{G}]) &= \mathbb{E}[(\mathbb{E}[X|\mathcal{G}])^2] - (\mathbb{E}[\mathbb{E}[X|\mathcal{G}]])^2, \\
&= \mathbb{E}[(\mathbb{E}[X|\mathcal{G}])^2] - (\mathbb{E}[X])^2.
\end{aligned} \tag{36}$$

Tenant compte de (35) et (36), on trouve la formule annoncée :

$$\operatorname{Var}(X) = \mathbb{E}[\operatorname{Var}(X|\mathcal{G})] + \operatorname{Var}(\mathbb{E}[X|\mathcal{G}]). \tag{37}$$

Exercice 5 : Etant donnée une suite $(X_n)_{n\geq 1}$ de variables aléatoires réelles i.i.d. et de carré intégrable. On considère une variable aléatoire τ à valeurs dans $\mathbb{N}^*$, et indépendante de $(X_n)_{n\geq 1}$.

1. Comme τ est une variable aléatoire à valeurs dans $\mathbb{N}^*, (\{\tau = k\}_{k\in\mathbb{N}^*})$ constitue une partition de Ω et $\sigma(\tau) = \sigma(\{\tau = k\}_{k\in\mathbb{N}^*})$.

   Appliquant la formule obtenue à la question **3.(a)** de l'**Exercice 1** de cette même PC, il vient :
$$\mathbb{E}[S_\tau|\tau] = \sum_{k\in\mathbb{N}^*} \frac{\mathbb{E}[S_\tau\,\mathbf{1}_{\{\tau=k\}}]}{\mathbb{P}(\tau = k)}\mathbf{1}_{\{\tau=k\}}\,. \tag{38}$$

   Rappel : **Lemme de regroupement** :
   Soit $Y_1, \cdots, Y_m$ des variables aléatoires indépendantes à valeurs réelles.
   Alors, si $0 < m_1 < m_2 < \cdots < m_p = m$, les variables aléatoires $f_1(Y_1, \cdots, Y_{m_1})\,, f_2(Y_{m_1+1}, \cdots, Y_{m_2})\,,$ $\cdots, f_p(Y_{m_{p-1}+1}, \cdots, Y_{m_p})$ sont indépendantes, pour toute application $f_1\ : \mathbb{R}^{m_1} \to \mathbb{R}\,, \cdots,$ $f_p\ : \mathbb{R}^{m_p-m_{p-1}} \to \mathbb{R}$, mesurables.
   Par exemple, si $Y_1, \cdots, Y_4$ sont des v.a. réelles indépendantes, les variables aléatoires $Y_1 + Y_2$ et Y_3Y_4 sont indépendantes.

   □

   D'après l'énoncé, τ est indépendante de la suite $(X_n)_{n\geq 1}$, soit les variables aléatoires $\tau, X_1, X_2, \cdots, X_n$ sont indépendantes, pour tout $n \geq 1$.
   Utilisant le lemme de regroupement, on en déduit que τ et $X_1 + X_2 + \cdots + X_n = S_n$ sont indépendantes, quel que soit $n \geq 1$.
   Rappel : **Indépendance de deux variables aléatoires réelles**
   X et Y sont deux variables aléatoires réelles indépendantes si et seulement si, pour toutes applications f et g boréliennes bornées ou à valeurs positives :
$$\mathbb{E}[f(X)g(Y)] = \mathbb{E}[f(X)]\mathbb{E}[g(Y)]\,. \tag{39}$$

   □

   Comme τ et $S_k\,, k \geq 1$ sont des variables aléatoires indépendantes, en prenant $\forall x \in \mathbb{R}$, $f(x) = x$ et $g(y) = \mathbf{1}_{\{y=k\}}, y \in \mathbb{R}$, on en déduit, d'après le rappel précédent que, pour tout $k \geq 1$,
$$\mathbb{E}[S_\tau\,\mathbf{1}_{\{\tau=k\}}] = \mathbb{E}[S_k\,\mathbf{1}_{\{\tau=k\}}] = \mathbb{E}[S_k]\mathbb{E}[\mathbf{1}_{\{\tau=k\}}]\,.$$
   Reprenant le calcul dans (38), on obtient :
$$\begin{aligned}
\mathbb{E}[S_\tau|\tau] &= \sum_{k\in\mathbb{N}^*} \mathbb{E}[S_k]\mathbf{1}_{\{\tau=k\}}\,,\\
&= \left(\sum_{k\in\mathbb{N}^*} k\,\mathbf{1}_{\{\tau=k\}}\right)\mathbb{E}[X_1]\,,\ \text{car les } X_k\,, k\geq 1\,, \text{ suivent la même loi,}\\
&= \tau\mathbb{E}[X_1]\,.
\end{aligned} \tag{40}$$
   Par ailleurs, en calquant la démarche précédente, on trouve :
$$\begin{aligned}
\mathrm{Var}(S_\tau|\tau) &= \mathbb{E}\left[(S_\tau - \mathbb{E}[S_\tau|\tau])^2|\tau\right]\,,\\
&= \mathbb{E}\left[(S_\tau - \tau\mathbb{E}[X_1])^2|\tau\right]\,,\ \text{d'après (40),}\\
&= \sum_{k\in\mathbb{N}^*} \frac{\mathbb{E}[(S_\tau - \tau\mathbb{E}[X_1])^2\,\mathbf{1}_{\{\tau=k\}}]}{\mathbb{P}(\tau = k)}\mathbf{1}_{\{\tau=k\}} \text{ en reprenant la formule } (\textbf{??}),\\
&= \sum_{k\in\mathbb{N}^*} \mathbb{E}\left[(S_k - k\mathbb{E}[X_1])^2\right]\,\mathbf{1}_{\{\tau=k\}}\,,\ \text{en utilisant l'indépendance des v.a. } \tau \text{ et } S_k\,, k\geq 1\\
&\qquad\qquad \text{et prenant } f(x) = (x - k\mathbb{E}[X_1])^2\,, x\in\mathbb{R} \text{ et } g(y) = \mathbf{1}_{\{y=k\}}, y\in\mathbb{R} \text{ dans (39),}\\
&= \sum_{k\in\mathbb{N}^*} \left(\mathrm{Var}\left(\sum_{k=1}^{n} X_k\right)\,\mathbf{1}_{\{\tau=k\}}\right)\,,\ \text{car les } X_k\,, k\geq 1\,, \text{ suivent la même loi,}\\
&= \sum_{k\in\mathbb{N}^*} \left(\left(\sum_{k=1}^{n} \mathrm{Var}(X_k)\right)\,\mathbf{1}_{\{\tau=k\}}\right)\,,\ \text{puisque les } X_k\,, k\geq 1\,, \text{ sont indépendantes,}\\
&= \mathrm{Var}(X_1)\left(\sum_{k\in\mathbb{N}^*} k\,\mathbf{1}_{\{\tau=k\}}\right)\,,\ \text{comme à nouveau les } X_k\,, k\geq 1\,, \text{ sont identiquement distribuées,}\\
&= \tau\mathrm{Var}(X_1)\,.
\end{aligned} \tag{41}$$

2. On a : $\mathbb{E}[S_\tau] = \mathbb{E}[\mathbb{E}[S_\tau]|\tau] = \mathbb{E}[\tau]\mathbb{E}[X_1]$, d'après (40).
De plus, $\mathbb{E}\left[(S_\tau - \mathbb{E}[S_\tau|\tau])^2\right] = \mathbb{E}[\mathbb{E}\left[(S_\tau - \mathbb{E}[S_\tau|\tau])^2\right]|\tau] = \mathbb{E}[\tau]\mathrm{Var}[X_1]$, utilisant le résultat montré en (41).
Enfin, en utilisant la formule (37) trouvée à l'Exercice 5 de cette même PC, il vient :
$\mathrm{Var}(S_\tau) = \mathbb{E}[\mathrm{Var}(S_\tau|\tau)] + \mathrm{Var}(\mathbb{E}[S_\tau|\tau]) = \mathbb{E}[\tau]\mathrm{Var}[X_1] + \mathrm{Var}(\tau)(\mathbb{E}[X_1])^2$, d'après les expressions (41) et (40).