Feuille de TD n° 5. Corrigé.

Propriétés générales

Exercice 1. On remarque d'abord que $\mathbb{E}[X|\mathcal{F}]$ satisfait l'égalité parce que $\mathbb{1}_C$ est $\mathcal{F}$-mesurable pour tout $C \in \mathcal{C}$. Supposons maintenant que Y est une v.a. intégrable et $\mathcal{F}$-mesurable telle que $\mathbb{E}[X\mathbb{1}_C] = \mathbb{E}[Y\mathbb{1}_C]$ pour tout $C \in \mathcal{C}$. Il suffit de montrer que cela implique $\mathbb{E}[X\mathbb{1}_A] = \mathbb{E}[Y\mathbb{1}_A]$ pour tout $A \in \mathcal{F}$, ce qui impliquerait que $Y = \mathbb{E}[X|\mathcal{F}]$. Pour cela, on pose

$$\mathcal{D} = \{A \in \mathcal{F} : \mathbb{E}[X\mathbb{1}_A] = \mathbb{E}[Y\mathbb{1}_A]\}.$$

On a alors $\mathcal{D} \supset \mathcal{C}$. On montre que $\mathcal{D}$ est une classe monotone :

1. Par hypothèse, $\Omega \in \mathcal{C} \subset \mathcal{D}$.
2. Soient $A, B \in \mathcal{D}$, $A \subset B$. On a alors

$$\mathbb{E}[X\mathbb{1}_{B\setminus A}] = \mathbb{E}[X\mathbb{1}_A] - \mathbb{E}[X\mathbb{1}_B] = \mathbb{E}[Y\mathbb{1}_A] - \mathbb{E}[Y\mathbb{1}_B] = \mathbb{E}[Y\mathbb{1}_{B\setminus A}],$$

   donc $B\setminus A \in \mathcal{D}$.
3. Soient $A_1, A_2, \ldots \in \mathcal{D}$, avec $A_n \subset A_{n+1}$ pour tout $n \in \mathbb{N}$. On note que $|X\mathbb{1}_{A_n}| \le |X| \in L^1$ et $|Y\mathbb{1}_{A_n}| \le |Y| \in L^1$ pour tout n. On a alors par convergence dominée,

$$\mathbb{E}[X\mathbb{1}_{\bigcup_n A_n}] = \mathbb{E}[\lim_{n\to\infty} X\mathbb{1}_{A_n}] = \lim_{n\to\infty} \mathbb{E}[X\mathbb{1}_{A_n}] = \lim_{n\to\infty} \mathbb{E}[Y\mathbb{1}_{A_n}] = \mathbb{E}[\lim_{n\to\infty} Y\mathbb{1}_{A_n}] = \mathbb{E}[Y\mathbb{1}_{\bigcup_n A_n}].$$

   Donc $\bigcup_n A_n \in \mathcal{D}$.

On conclut que $\mathcal{D}$ est une classe monotone. Puisque $\mathcal{C}$ est stable par intersection finie (par hypothèse), le lemme des classes monotones donne alors $\mathcal{D} \supset \sigma(\mathcal{C}) = \mathcal{F}$. Ceci montre que $\mathbb{E}[X\mathbb{1}_A] = \mathbb{E}[Y\mathbb{1}_A]$ pour tout $A \in \mathcal{F}$. CQFD.

Exercice 2. On sait que la fonction $g(x) = \mathbb{E}[f(x,Y)] = \int f(x,y)\,\mu_Y(dy)$ est mesurable, donc $g(X)$ est une v.a. $\sigma(X)$-mesurable. Il suffit alors de vérifier que $g(X)$ satisfait à la propriété caractéristique. Soit W une v.a. bornée $\sigma(X)$-mesurable. Par le lemme crucial il existe alors une fonction bornée mesurable h telle que $W = h(X)$. On a alors

$$\begin{aligned}
\mathbb{E}[h(X)g(X)] &= \int h(x)g(x)\,\mu_X(dx) \\
&= \int h(x)\left[\int f(x,y)\,\mu_Y(dy)\right]\mu_X(dx) \\
&= \int h(x)f(x,y)(\mu_X\otimes\mu_Y)(dx,dy) && \text{(par Fubini)} \\
&= \mathbb{E}[h(X)f(X,Y)].
\end{aligned}$$

Ceci montre que $g(X)$ satisfait à la propriété caractéristique et donc $g(X) = \mathbb{E}[f(X,Y)\,|\,X]$ p.s.

Exercice 3. On définit la classe

$$\mathcal{C} = \{B_1 \cap B_2 : B_i \in \mathcal{B}_i,\ i = 1,2\}.$$

La classe $\mathcal{C}$ est alors stable par intersection finie et engendre la tribu $\mathcal{B}_1 \vee \mathcal{B}_2$. De plus, pour tout $C = B_1 \cap B_2 \in \mathcal{C}$, $B_i \in \mathcal{B}_i,\ i = 1,2$, on a

$$\begin{aligned}
\mathbb{E}[Y\mathbb{1}_C] &= \mathbb{E}[Y\mathbb{1}_{B_1}\mathbb{1}_{B_2}] \\
&= \mathbb{E}[Y\mathbb{1}_{B_1}]\mathbb{E}[\mathbb{1}_{B_2}] && \text{par indépendance de } \sigma(Y)\vee\mathcal{B}_1 \text{ et } \mathcal{B}_2 \\
&= \mathbb{E}[\mathbb{E}[Y|\mathcal{B}_1]\mathbb{1}_{B_1}]\mathbb{E}[\mathbb{1}_{B_2}] \\
&= \mathbb{E}[\mathbb{E}[Y|\mathcal{B}_1]\mathbb{1}_{B_1}\mathbb{1}_{B_2}] && \text{par indépendance de } \sigma(Y)\vee\mathcal{B}_1 \text{ et } \mathcal{B}_2 \\
&= \mathbb{E}[\mathbb{E}[Y|\mathcal{B}_1]\mathbb{1}_C].
\end{aligned}$$

L'exercice 1 montre alors que $\mathbb{E}[Y|\mathcal{B}_1] = \mathbb{E}[Y|\mathcal{B}_1 \vee \mathcal{B}_2]$.

Exercice 4.

1. Si $\mathcal{G} = \{0, \Omega\}$, cela signifie que $\mathbb{E}[f(X)g(Y)] = \mathbb{E}[f(X)]\mathbb{E}[g(Y)]$ pour toutes fonctions f, g positives mesurables, donc que X et Y sont indépendantes. Si $\mathcal{G} = \mathcal{F}$, l'égalité est toujours vérifiée, car

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

puisque $f(X)$ et $g(Y)$ sont $\mathcal{F}$-mesurables.

2. Si $\mathbb{E}[f(X)g(Y)|\mathcal{G}] = \mathbb{E}[f(X)|\mathcal{G}]\mathbb{E}[g(Y)|\mathcal{G}]$, alors on a pour toute v.a. $Z \geq 0$, $\mathcal{G}$-mesurable,

$$\mathbb{E}[f(X)g(Y)Z] = \mathbb{E}[\mathbb{E}[f(X)g(Y)|\mathcal{G}]Z] = \mathbb{E}[\mathbb{E}[f(X)|\mathcal{G}]\mathbb{E}[g(Y)|\mathcal{G}]Z] = \mathbb{E}[f(X)\mathbb{E}[g(Y)|\mathcal{G}]Z],$$

où dans la première égalité on a appliqué la propriété caractéristique à $\mathbb{E}[f(X)g(Y)|\mathcal{G}]$, et dans la dernière égalité à $\mathbb{E}[f(X)|\mathcal{G}]$.
Inversement, pour toute v.a. $Z \geq 0$, $\mathcal{G}$-mesurable, si on a $\mathbb{E}[f(X)g(Y)Z] = \mathbb{E}[f(X)\mathbb{E}[g(Y)|\mathcal{G}]Z]$ alors on a

$$\mathbb{E}[f(X)g(Y)Z] = \mathbb{E}[f(X)\mathbb{E}[g(Y)|\mathcal{G}]Z] = \mathbb{E}[\mathbb{E}[f(X)|\mathcal{G}]\mathbb{E}[g(Y)|\mathcal{G}]Z].$$

Ceci montre que la v.a. $\mathcal{G}$-mesurable $\mathbb{E}[f(X)|\mathcal{G}]\mathbb{E}[g(Y)|\mathcal{G}]$ satisfait la propriété caractéristique pour $f(X)g(Y)$ et $\mathcal{G}$. Ceci implique que $\mathbb{E}[f(X)g(Y)|\mathcal{G}] = \mathbb{E}[f(X)|\mathcal{G}]\mathbb{E}[g(Y)|\mathcal{G}]$.
Supposons maintenant que $\mathbb{E}[g(Y)|\sigma(\mathcal{G}, \sigma(X))] = \mathbb{E}[g(Y)|\mathcal{G}]$. On a alors pour toute v.a. $Z \geq 0$, $\mathcal{G}$-mesurable, $\mathbb{E}[f(X)g(Y)Z] = \mathbb{E}[f(X)Z\mathbb{E}[g(Y)|\sigma(\mathcal{G}, \sigma(X))]] = \mathbb{E}[f(X)Z\mathbb{E}[g(Y)|\mathcal{G}]]$, ce qui donne la deuxième égalité. Inversement, supposons que pour toutes fonctions $f, g \geq 0$ mesurables et pour toute v.a. $Z \geq 0$, $\mathcal{G}$-mesurable, on ait $\mathbb{E}[f(X)g(Y)Z] = \mathbb{E}[f(X)Z\mathbb{E}[g(Y)|\mathcal{G}]]$. Par le "lemme crucial", chaque v.a. $W \geq 0$, $\sigma(X)$-mesurable, s'écrit que $f(X)$ pour une fonction $f \geq 0$ mesurable. En prenant $W = \mathbb{1}_B$ pour $B \in \sigma(X)$ et $Z = \mathbb{1}_C$ pour $C \in \mathcal{G}$, on a alors avec $A = B \cap C$, $\mathbb{E}[g(Y)\mathbb{1}_A] = \mathbb{E}[\mathbb{E}[g(Y)|\mathcal{G}]\mathbb{1}_A]$. La classe d'ensemble

$$\mathcal{C} = \{B \cap C : B \in \sigma(X),\ C \in \mathcal{G}\}$$

contient Ω et est stable par intersection finie. De plus, $\mathcal{C}$ contient $\sigma(X)$ et $\mathcal{G}$ et donc engendre $\mathcal{F}$. Par l'exercice 1, on a alors $\mathbb{E}[g(Y)\mathbb{1}_A] = \mathbb{E}[\mathbb{E}[g(Y)|\mathcal{G}]\mathbb{1}_A]$ pour tout $A \in \mathcal{F}$. De plus, $\mathbb{E}[g(Y)|\mathcal{G}]$ est $\sigma(\mathcal{G}, \sigma(X))$-mesurable. On a vu dans le cours que ceci implique que $\mathbb{E}[g(Y)|\sigma(\mathcal{G}, \sigma(X))] = \mathbb{E}[g(Y)|\mathcal{G}]$.

Exercice 5 (Théorème de la variance totale).

1. Puisque $X \in L^2$, on a $\mathbb{E}[X|\mathcal{F}] = \Pi_{\mathcal{F}}(X) \in L^2$ également. Par l'inégalité de Cauchy–Schwarz, on a alors $\mathbb{E}[X|\mathcal{F}]Y \in L^1$. Puisque $\mathbb{E}[X|\mathcal{F}]$ est $\mathcal{F}$-mesurable, il vient

$$\mathbb{E}[\,\mathbb{E}[X|\mathcal{F}]Y \mid \mathcal{F}] = \mathbb{E}[X|\mathcal{F}]\,\mathbb{E}[Y|\mathcal{F}].$$

Par conséquent,

$$\mathbb{E}[\,\mathbb{E}[X|\mathcal{F}]Y \mid \mathcal{G}] = \mathbb{E}[\,\mathbb{E}[\,\mathbb{E}[X|\mathcal{F}]Y \mid \mathcal{F}] \mid \mathcal{G}] = \mathbb{E}[\,\mathbb{E}[X|\mathcal{F}]\,\mathbb{E}[Y|\mathcal{F}] \mid \mathcal{G}].$$

2. On a

$$\begin{aligned}
&\mathbb{E}\big[(X - \mathbb{E}[X|\mathcal{F}])(\mathbb{E}[X|\mathcal{F}] - \mathbb{E}[X|\mathcal{G}]) \,\big|\, \mathcal{G}\big] \\
&= \mathbb{E}[X\mathbb{E}[X|\mathcal{F}] - \mathbb{E}[X|\mathcal{F}]^2 - X\mathbb{E}[X|\mathcal{G}] + \mathbb{E}[X|\mathcal{F}]\mathbb{E}[X|\mathcal{G}] \mid \mathcal{G}] \\
&= \mathbb{E}[\mathbb{E}[X|\mathcal{F}]^2 - \mathbb{E}[X|\mathcal{F}]^2 - \mathbb{E}[X|\mathcal{G}]^2 + \mathbb{E}[X|\mathcal{G}]^2 \mid \mathcal{G}] && \text{(partie 1)} \\
&= 0 \quad \text{p.s.}
\end{aligned}$$

3. On a

$$\begin{aligned}
\mathrm{Var}(X|\mathcal{G}) &= \mathbb{E}[(X - \mathbb{E}[X|\mathcal{G}])^2 \mid \mathcal{G}] \\
&= \mathbb{E}[(X - \mathbb{E}[X|\mathcal{F}] + \mathbb{E}[X|\mathcal{F}] - \mathbb{E}[X|\mathcal{G}])^2 \mid \mathcal{G}] \\
&= \mathbb{E}[(X - \mathbb{E}[X|\mathcal{F}])^2 - 2(X - \mathbb{E}[X|\mathcal{F}])(\mathbb{E}[X|\mathcal{F}] - \mathbb{E}[X|\mathcal{G}]) + (\mathbb{E}[X|\mathcal{F}] - \mathbb{E}[X|\mathcal{G}])^2 \mid \mathcal{G}] \\
&= \mathbb{E}[\mathbb{E}[(X - \mathbb{E}[X|\mathcal{F}])^2 \mid \mathcal{F}] \mid \mathcal{G}] + \mathbb{E}[(\mathbb{E}[X|\mathcal{F}] - \mathbb{E}[\mathbb{E}[X|\mathcal{F}]|\mathcal{G}])^2 \mid \mathcal{G}] && \text{(partie 2)} \\
&= \mathbb{E}[\mathrm{Var}(X|\mathcal{F}) \mid \mathcal{G}] + \mathrm{Var}(\mathbb{E}[X|\mathcal{F}] \mid \mathcal{G}).
\end{aligned}$$

CQFD.

On remarque que $\mathbb{E}[X|\mathcal{F}]$ est la projection dans $L^2(\Omega, \mathbb{P})$ de X sur le sous-espace des v.a. $\mathcal{F}$-mesurables, si bien que $X - \mathbb{E}[X|\mathcal{F}]$ est orthogonale à ce sous-espace et en particulier, $X - \mathbb{E}[X|\mathcal{F}]$ et $\mathbb{E}[X|\mathcal{F}] - \mathbb{E}[X]$ sont orthogonales. Par conséquent,

$$\begin{aligned}\mathrm{Var}(X) &= \|X - \mathbb{E}[X]\|_2^2 \\ &= \|X - \mathbb{E}[X|\mathcal{F}] + \mathbb{E}[X|\mathcal{F}] - \mathbb{E}[X]\|_2^2 \\ &= \|X - \mathbb{E}[X|\mathcal{F}]\|_2^2 + \|\mathbb{E}[X|\mathcal{F}] - \mathbb{E}[X]\|_2^2 \\ &= \mathbb{E}[\mathrm{Var}(X|\mathcal{F})] + \mathrm{Var}(\mathbb{E}[X|\mathcal{F}])).\end{aligned}$$

Ceci est la loi de la variance totale pour $\mathcal{G} = \{\emptyset, \Omega\}$

Exercice 6. Soit $X \geq 0$ v.a. et $\mathcal{F}$ une tribu. On a pour tout $t > 0$,

$$\mathbb{P}(X \geq t \,|\, \mathcal{F}) = \mathbb{E}[\mathbb{1}_{X \geq t} \,|\, \mathcal{F}] \leq \mathbb{E}\Big[\frac{X}{t} \,\Big|\, \mathcal{F}\Big],$$

où la dernière inégalité provient de la positivité de l'intégrale conditionnelle. Par linéarité de l'espérance conditionnelle, ceci donne l'inégalité de Markov conditionnelle :

$$\mathbb{P}(X \geq t \,|\, \mathcal{F}) \leq \frac{\mathbb{E}[X|\mathcal{F}]}{t}.$$

En appliquant cette inégalité, on obtient,

$$\mathbb{P}(|X - \mathbb{E}[X|\mathcal{F}]| \geq t \,|\, \mathcal{F}) = \mathbb{P}((X - \mathbb{E}[X|\mathcal{F}])^2 \geq t^2) \leq \frac{\mathbb{E}[(X - \mathbb{E}[X|\mathcal{F}])^2 \,|\, \mathcal{F}]}{t^2},$$

ce qui donne lieu à l'inégalité de Chebychev conditionnelle :

$$\mathbb{P}(|X - \mathbb{E}[X|\mathcal{F}]| \geq t \,|\, \mathcal{F}) \leq \frac{\mathrm{Var}(X|\mathcal{F})}{t^2}.$$

Exemples

Exercice 7. X^2 est $\sigma(X)$-mesurable, donc $\mathbb{E}[X^2|X] = X^2$. Pour calculer $\mathbb{E}[X|X^2]$, soit W une v.a. bornée $\sigma(X^2)$-mesurable. Par le "lemme crucial", il existe alors une fonction f mesurable bornée telle que $W = f(X^2)$. On a alors par l'hypothèse de symétrie,

$$\mathbb{E}[Xf(X^2)] = \mathbb{E}[(-X)f((-X)^2)] = -\mathbb{E}[Xf(X^2)],$$

si bien que $\mathbb{E}[Xf(X^2)] = 0$. La variable aléatoire constante égale à 0 vérifie alors

- 0 est $\sigma(X^2)$-mesurable.
- $\mathbb{E}[0W] = 0 = \mathbb{E}[XW]$ pour toute v.a. bornée W, $\sigma(X^2)$-mesurable, i.e. 0 vérifie la propriété caractéristique.

Ceci montre que $\mathbb{E}[X|X^2] = 0$.

Exercice 8.

1. Par définition de l'espérance conditionnelle, on a pour tout i et toute fonction bornée mesurable f,
   - $\mathbb{E}[X_i|S_n]$ est $\sigma(S_n)$-mesurable,
   - $\mathbb{E}[\mathbb{E}[X_i|S_n]f(S_n)] = \mathbb{E}[X_i f(S_n)]$.

   Posons $F(x_1, \ldots, x_n) = x_1 f(x_1 + \cdots + x_n)$. On a alors pour tout i,

   $$\mathbb{E}[X_1 f(S_n)] = \mathbb{E}[F(X_1, \ldots, X_n)] = \mathbb{E}[F(X_i, X_2, \ldots, X_{i-1}, X_1, X_{i+1}, \ldots, X_n)] = \mathbb{E}[X_i f(S_n)].$$

   Ici, la deuxième égalité provient du fait que les vecteurs $(X_1, \ldots, X_n)$ et $(X_i, X_2, \ldots, X_{i-1}, X_1, X_{i+1}, \ldots, X_n)$ ont même loi (la loi $\mu_X^{\otimes n}$, qui est invariante par permutation des coordonnées). Ceci montre que $\mathbb{E}[X_i|S_n]$ vérifie également la propriété caractéristique pour X_1, donc $\mathbb{E}[X_i|S_n] = \mathbb{E}[X_1|S_n]$.

2. On a d'après la première partie,
$$n\mathbb{E}[X_1|S_n] = \sum_{i=1}^n \mathbb{E}[X_i|S_n] = \mathbb{E}[S_n|S_n] = S_n,$$
et donc
$$\mathbb{E}[X_1|S_n] = \frac{S_n}{n}.$$
3. Soit $\mathcal{G} = \sigma(X_{n+1}, X_{n+2}, \ldots)$ On a alors $\sigma(S_n, S_{n+1}, \ldots) = \sigma(S_n) \vee \mathcal{G}$. De plus, les tribus $\sigma(X_1, S_n)$ et $\mathcal{G}$ sont indépendantes. L'exercice 3 montre alors que $\mathbb{E}[X_1|S_n] = \mathbb{E}[X_1|\sigma(S_n) \vee \mathcal{G}] = \mathbb{E}[X_1|S_n, S_{n+1}, \ldots]$.

Exercice 9.

1. La fonction $x \mapsto |x|^p$, est convexe pour $p \geq 1$. Par l'inégalité de Jensen conditionnelle, on a alors
$$\mathbb{E}[|X+Y|^p] = \mathbb{E}[\mathbb{E}[|X+Y|^p \mid X]] \geq \mathbb{E}[|\mathbb{E}[X+Y \mid X]|^p] = \mathbb{E}[|X+\mathbb{E}[Y|X]|^p].$$
En prenant les racines p-ièmes, on obtient l'inégalité souhaitée.
2. Si X et Y sont indépendantes, on a $\mathbb{E}[Y|X] = \mathbb{E}[Y]$. L'inégalité découle alors de la précédente.

Exercice 10. Le fait que 4. implique 3. et que 3. implique 1. et 2. est évident. Montrons que 1. implique 3. Soient X, Y des v.a. intégrables telles que 1. est vérifiée. Alors on a
$$\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X|Y]] \leq \mathbb{E}[Y] = \mathbb{E}[\mathbb{E}[Y|X]] \leq \mathbb{E}[X],$$
donc $\mathbb{E}[X] = \mathbb{E}[Y]$. Par conséquent, on a d'après 1.,
$$\mathbb{E}[X] = \mathbb{E}[Y] = \mathbb{E}[Y - \mathbb{E}[X|Y] + \mathbb{E}[X|Y]] = \mathbb{E}[(Y - \mathbb{E}[X|Y])^+] + \mathbb{E}[X],$$
et donc $\mathbb{E}[(Y - \mathbb{E}[X|Y])^+] = 0$, ce qui donne $Y \leq \mathbb{E}[X|Y]$ p.s., et donc d'après 1., $Y = \mathbb{E}[X|Y]$ p.s. De la même façon, on montre que $X = \mathbb{E}[Y|X]$ p.s. Ceci montre que 1. implique bien 3.

Pour montrer que 2. implique 3., on peut faire une preuve analogue ou alors utiliser 1. pour les v.a. $\widetilde{X} = -X$ et $\widetilde{Y} = -Y$ (on note que $\sigma(\widetilde{X}) = \sigma(X)$ car $\widetilde{X} = -X$ et $X = -\widetilde{X}$ et de même $\sigma(\widetilde{Y}) = \sigma(Y)$).

Il suffit alors de montrer que 3. implique 4. On traite d'abord le cas où $X, Y \in L^2$. En supposant que 3. est vérifiée, on a alors
$$\begin{aligned}\mathbb{E}[(X-Y)^2] &= \mathbb{E}[X^2 + Y^2 - XY - XY] \\ &= \mathbb{E}[X^2 + Y^2 - X\mathbb{E}[Y|X] - Y\mathbb{E}[X|Y]] \\ &= \mathbb{E}[X^2 + Y^2 - X^2 - Y^2] \\ &= 0.\end{aligned}$$
Par conséquent, $\mathbb{E}[(X-Y)^2] = 0$ et donc $X = Y$ p.s.

Remarque : Une explication géométrique de cette preuve est la suivante : En supposant 3., on a $\mathbb{E}[X-Y|X] = \mathbb{E}[X-Y|Y] = 0$*, et donc* $X - Y \in X^\perp \cap Y^\perp$*. Par conséquent,* $X - Y$ *est orthogonal au sous-espace de* L^2 *engendré par* X *et* Y*, en particulier* $X - Y \in (X-Y)^\perp$*. Par conséquent,* $\mathbb{E}[(X-Y)^2] = \|X-Y\|_2^2 = 0$.

Supposons maintenant que $X, Y \in L^1$, $X \geq 0$, $Y \geq 0$, et que 3. est vérifiée. Alors $\sqrt{X}, \sqrt{Y} \in L^2$. De plus, $\sigma(\sqrt{X}) = \sigma(X)$ et $\sigma(\sqrt{Y}) = \sigma(Y)$, puisque $X = (\sqrt{X})^2$ et $Y = (\sqrt{Y})^2$. Puisque la fonction $x \mapsto \sqrt{x}$ est concave sur $\mathbb{R}_+$ (i.e. $x \mapsto -\sqrt{x}$ est convexe), l'inégalité de Jensen conditionnelle donne
$$\mathbb{E}[\sqrt{X}|\sqrt{Y}] = \mathbb{E}[\sqrt{X}|Y] \leq \sqrt{\mathbb{E}[X|Y]} \leq \sqrt{Y}.$$
Pareil, $\mathbb{E}[\sqrt{Y}|\sqrt{X}] \leq \sqrt{Y}$. Par l'équivalence entre 1. et 4. pour des v.a. L^2, cela donne $\sqrt{X} = \sqrt{Y}$ p.s., et donc $X = Y$ p.s.

Supposons maintenant que $X, Y \in L^1$, pas nécessairement positives, et telles que 3. est vérifiée. Puisque la fonction $x \mapsto x^+ = \max(x, 0)$ est convexe, l'inégalité de Jensen conditionnelle donne $\mathbb{E}[X^+|Y] \geq \mathbb{E}[X|Y]^+ = Y^+$, et pareil, $\mathbb{E}[Y^+|X] \geq X^+$. Puisque $\sigma(X^+) \subset \sigma(X)$ et $\sigma(Y^+) \subset \sigma(Y)$, cela donne $\mathbb{E}[X^+|Y^+] = \mathbb{E}[\mathbb{E}[X^+|Y]|Y^+] \geq \mathbb{E}[Y^+|Y^+] = Y^+$, et pareil, $\mathbb{E}[Y^+|X^+] \geq X^+$. En utilisant l'équivalence entre 2. et 4. pour des v.a. positives, on obtient alors $X^+ = Y^+$ p.s. En faisant le même raisonnement avec $\widetilde{X} = -X$ et $\widetilde{Y} = -Y$, obtient $X^- = Y^-$ p.s. Ceci donne finalement $X = Y$ p.s.

Exercice 11. Calcul direct : On sait que $X_1 + X_2$ suit la loi de Poisson de paramètre $\lambda_1 + \lambda_2$. Si $n \in \mathbb{N}$, $k \in \{0, \dots, n\}$, on a alors

$$\mathbb{P}(X_1 = k \,|\, X_1 + X_2 = n) = \frac{\mathbb{P}(X_1 = k,\, X_1 + X_2 = n)}{\mathbb{P}(X_1 + X_2 = n)} = \frac{\mathbb{P}(X_1 = k,\, X_2 = n-k)}{\mathbb{P}(X_1 + X_2 = n)} = \frac{\mathbb{P}(X_1 = k)\mathbb{P}(X_2 = n-k)}{\mathbb{P}(X_1 + X_2 = n)}$$
$$= \frac{e^{-\lambda_1}\frac{\lambda_1^k}{k!}e^{-\lambda_2}\frac{\lambda_2^{n-k}}{(n-k)!}}{e^{-(\lambda_1+\lambda_2)}\frac{(\lambda_1+\lambda_2)^n}{n!}} = \binom{n}{k}\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^k\left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{n-k}.$$

Il en suit que la loi de X_1 conditionnellement à $X_1 + X_2$ est la loi binomiale de paramètres $n = X_1 + X_2$ et $p = \lambda_1/(\lambda_1 + \lambda_2)$.

Approche alternative : Soient $Y^1, Y^2, \dots$ des v.a. iid de loi $\mathbb{P}(Y^1 = i) = p_i = \lambda_i/(\lambda_1 + \lambda_2)$, $i = 1, 2$. Soit P une v.a. de loi $\mathrm{Po}(\lambda_1 + \lambda_2)$, indépendante de la suite $(Y^j)_{j=1,2,\dots}$. On définit le vecteur aléatoire

$$\vec{\xi} = (\xi_1, \xi_2) = \sum_{j=1}^{P} (\mathbf{1}_{(Y^j=1)}, \mathbf{1}_{(Y^j=2)}).$$

Par l'exercice 5 de la feuille TD n° 2, $(\xi_1, \xi_2) \overset{\text{loi}}{=} (X_1, X_2)$. Il suffit alors de montrer que conditionellement à $\xi_1 + \xi_2$, ξ_1 suit la loi binomiale de paramètres $n = \xi_1 + \xi_2$ et $p = \lambda_1/(\lambda_1 + \lambda_2)$. Or, $\xi_1 + \xi_2 = P$ et par définition, conditionnellement à P, $\vec{\xi}$ suit la loi multinomiale de paramètres P et p_1, p_2. En particulier, conditionnellement à $\xi_1 + \xi_2$, ξ_1 suit la loi binomiale de paramètres $n = X_1 + X_2$ et p_1.

Exercice 12. On rappelle que la loi de S est la loi de densité $p_1 * p_2$ par rapport à la mesure de Lebesgue sur $\mathbb{R}$. Soient $f, h : \mathbb{R} \to \mathbb{R}$ mesurables positives. On va montrer qu'on peut écrire

$$\mathbb{E}[f(X)h(S)] = \int \left(\int \varphi(s,x)f(x)\,dx\right)(p_1 * p_2)(s)h(s)\,ds$$
$$= \mathbb{E}\left[\left(\int \varphi(S,x)f(x)\,dx\right)h(S)\right],$$

pour une certaine fonction $\varphi(s,x)$. Il en découlera que la loi de X conditionnellement à S est la loi de densité $\varphi(S,x)$ par rapport à la mesure de Lebesgue sur $\mathbb{R}$.

On écrit

$$\mathbb{E}[f(X)h(S)] = \int \left(\int p_1(x)p_2(y)f(x)h(x+y)\,dy\right)dx.$$

En changeant de variables $y \mapsto s - x$ dans l'intégrale intérieure, il vient

$$\mathbb{E}[f(X)h(S)] = \int \left(\int p_1(x)p_2(s-x)f(x)h(s)\,ds\right)dx$$
$$= \int \left(\int p_1(x)p_2(s-x)f(x)\,dx\right)h(s)\,ds \quad \text{(par Fubini–Tonelli)}$$

Notons que si $0 = p_1 * p_2(s) = \int p_1(x)p_2(s-x)\,dx$, alors $p_1(x)p_2(s-x) = 0$ pour Lebesgue-presque tout x, par positivité de p, et donc $\int p_1(x)p_2(s-x)f(x)\,dx = 0$. Par conséquent, en posant

$$\varphi(s,x) = \begin{cases} \frac{p_1(x)p_2(s-x)}{p_1 * p_2(s)} & \text{si } p_1 * p_2(s) > 0, \\ 0 & \text{sinon,} \end{cases}$$

alors on a

$$\mathbb{E}[f(X)h(S)] = \int \left(\int \varphi(s,x)f(x)\,dx\right)(p_1 * p_2)(s)h(s)\,ds.$$

Donc la loi de X conditionnellement à S est la loi de densité $\varphi(S,x)$ par rapport à la mesure de Lebesgue sur $\mathbb{R}$.

Dans le cas où $X \sim \Gamma(\alpha_1, \beta)$, $Y \sim \Gamma(\alpha_2, \beta)$ avec $\alpha_1, \alpha_2, \beta > 0$, on a

$$\varphi(s,x) = c_s x^{\alpha_1 - 1}(s-x)^{\alpha_2 - 1}\mathbb{1}_{x \in [0,s]},$$

pour une constante c_s dépendant de s. Puisque $\int \varphi(s,x)\,dx = 1$, on a $c_s = (\int_0^s x^{\alpha-1}(s-x)^{\alpha-1})^{-1}$. Si Z_s est une v.a. de densité $\varphi(s,x)$ par rapport à Lebesgue sur $\mathbb{R}$, alors $s^{-1}Z_s$ est de densité

$$s\varphi(s,sx) = s^{\alpha_1+\alpha_2-1}c_s x^{\alpha_1-1}(1-x)^{\alpha_2-1}\mathbb{1}_{x\in[0,1]}.$$

On reconnaît cette loi comme la loi Bêta $B(\alpha_1, \alpha_2)$ (cf la feuille “Recueil d'exercices”).

Exercice 13.

1. Par l'Exercice 1, on a pour toute fonction bornée mesurable f,
$$\mathbb{E}[f(X+Y)\,|\,X] = \mathbb{E}[f(x+Y)]_{x=X},$$
et donc $X+Y$ suit la loi $\mathcal{N}(X,\sigma_Y^2)$ conditionnellement à X.
2. Première solution : On cherche $a,b\in\mathbb{R}$ tels que $aX+bY \perp X+Y$. Puisque $(X,Y)^T$ est un vecteur gaussien, $(X+Y,aX+bY)^T = \begin{pmatrix}1 & 1\\ a & b\end{pmatrix}(X,Y)^T$ en est un également. On a alors
$$X+Y\perp aX+bY \iff 0 = \mathrm{Cov}(X+Y,aX+bY) = a\sigma_X^2 + b\sigma_Y^2$$
Posons alors $a=\sigma_Y^2$, $b=-\sigma_X^2$. On a alors, $X = (b(X+Y)-(aX+bY))/(b-a) = U+V$, avec
$$U = \tfrac{b}{b-a}(X+Y),\quad V = -(b-a)^{-1}(aX+bY),\quad U\perp V$$
et donc
$$\sigma_U^2 = \left(\frac{\sigma_X^2}{\sigma_X^2+\sigma_Y^2}\right)^2(\sigma_X^2+\sigma_Y^2) = \frac{\sigma_X^4}{\sigma_X^2+\sigma_Y^2}$$
$$\sigma_V^2 = \frac{1}{(\sigma_X^2+\sigma_Y^2)^2}(\sigma_Y^4\sigma_X^2+\sigma_X^4\sigma_Y^2) = \frac{\sigma_X^2\sigma_Y^2}{\sigma_X^2+\sigma_Y^2}$$
Notons que U est $\sigma(X+Y)$-mesurable. Comme dans la première partie de l'exercice, la loi de X conditionnellement à $X+Y$ est alors la loi
$$\mathcal{N}(U,\sigma_V^2) = \mathcal{N}\left(\frac{\sigma_X^2}{\sigma_X^2+\sigma_Y^2}(X+Y), \frac{\sigma_X^2\sigma_Y^2}{\sigma_X^2+\sigma_Y^2}\right).$$
Deuxième solution :
On calcule d'abord $\mathbb{E}[X|X+Y]$. On sait d'après le cours qu'il existe $\alpha,\beta\in\mathbb{R}$, telles que $\mathbb{E}[X|X+Y] = \alpha+\beta(X+Y)$. En prenant des espérances, on obtient alors $0=\alpha$. De plus, on a
$$\mathbb{E}[\mathbb{E}[X|X+Y](X+Y)] = \mathbb{E}[X(X+Y)] = \mathbb{E}[X^2]+\mathbb{E}[XY] = \sigma_X^2,$$
et
$$\mathbb{E}[\mathbb{E}[X|X+Y](X+Y)] = \mathbb{E}[\beta(X+Y)(X+Y)] = \beta\mathbb{E}[(X+Y)^2] = \beta(\sigma_X^2+\sigma_Y^2),$$
si bien que $\beta = \sigma_X^2/(\sigma_X^2+\sigma_Y^2)$. Par conséquent,
$$\mathbb{E}[X|X+Y] = \frac{\sigma_X^2}{\sigma_X^2+\sigma_Y^2}(X+Y).$$
Maintenant on note que $X-\mathbb{E}[X|X+Y]$ est une variable gaussienne indépendante de $X+Y$, car $(X-\mathbb{E}[X|X+Y],X+Y)$ est un vecteur gaussien (car image d'un vecteur gaussien par une application linéaire) et $X-\mathbb{E}[X|X+Y]$ est orthogonale à $X+Y$ dans L^2, donc $\mathrm{Cov}(X-\mathbb{E}[X|X+Y],X+Y)=0$. Il suffit alors de déterminer la variance de $X-\mathbb{E}[X|X+Y]$. Or,
$$\begin{aligned}\mathrm{Var}\,(X-\mathbb{E}[X|X+Y]) &= \mathrm{Var}\left(\frac{(\sigma_X^2+\sigma_Y^2)X-\sigma_X^2(X+Y)}{\sigma_X^2+\sigma_Y^2}\right)\\ &= \frac{\mathrm{Var}(\sigma_Y^2X-\sigma_X^2Y)}{(\sigma_X^2+\sigma_Y^2)^2}\\ &= \frac{\sigma_Y^4\sigma_X^2+\sigma_X^4\sigma_Y^2}{(\sigma_X^2+\sigma_Y^2)^2}\\ &= \frac{\sigma_X^2\sigma_Y^2}{\sigma_X^2+\sigma_Y^2}.\end{aligned}$$
Donc, X est la somme de $\beta(X+Y)$ et d'une gaussienne indépendante, centrée, de variance $\frac{\sigma_X^2\sigma_Y^2}{\sigma_X^2+\sigma_Y^2}$. Comme dans la première partie de l'exercice, la loi de X conditionnellement à $X+Y$ est alors la loi gaussienne
$$\mathcal{N}\left(\frac{\sigma_X^2}{\sigma_X^2+\sigma_Y^2}(X+Y), \frac{\sigma_X^2\sigma_Y^2}{\sigma_X^2+\sigma_Y^2}\right).$$

Exercice 14 (Processus de Poisson homogène sur $\mathbb{R}_+$).

1. Pour tout $n \in \mathbb{N}$, $t \geq 0$,

$$\begin{aligned}
\frac{1}{n!}\int_t^\infty x^n e^{-x}\,dx &= e^{-t}\frac{1}{n!}\int_0^\infty (x+t)^n e^{-x}\,dx && x \to x+t \\
&= e^{-t}\frac{1}{n!}\int_0^\infty \sum_{k=0}^n \frac{n!}{k!(n-k)!} x^{n-k}t^k e^{-x}\,dx && \text{formule du binôme de Newton} \\
&= e^{-t}\sum_{k=0}^n \frac{t^k}{k!}\frac{1}{(n-k)!}\int_0^\infty x^{n-k}e^{-x}\,dx \\
&= e^{-t}\sum_{k=0}^n \frac{t^k}{k!} && \text{par la définition de la fonction } \Gamma.
\end{aligned}$$

2. On a pour tout $n \in \mathbb{N}$, $\Pi([0,t)) \leq n$ si et seulement si $X_{n+1} \geq t$. Or, $X_{n+1} \sim \Gamma(n+1,\lambda)$, donc $\lambda X_{n+1} \sim \Gamma(n+1,1)$. Cela donne,

$$\begin{aligned}
\mathbb{P}(\Pi([0,t)) \leq n) &= \mathbb{P}(X_{n+1} \geq t) = \mathbb{P}(\lambda X_{n+1} \geq \lambda t) \\
&= \frac{1}{n!}\int_{\lambda t}^\infty x^n e^{-x}\,dx && \text{car } \Gamma(n+1) = n! \\
&= e^{-\lambda t}\sum_{k=0}^n \frac{(\lambda t)^k}{k!} && \text{par la première partie de l'exercice.}
\end{aligned}$$

   Ceci montre que $\Pi([0,t))$ suit bien la loi de Poisson de paramètre λt.

3. Soit $n \in \mathbb{N}$. On a égalité des événements

$$\{\Pi([0,t)) = n\} = \{X_n < t,\, X_{n+1} \geq t\} = \{X_n < t,\, Y_{n+1} \geq t - X_n\}.$$

   Il en suit pour tout $y \geq 0$,

$$\begin{aligned}
\mathbb{P}(X_1^{(t)} \geq y,\, \Pi([0,t)) = n) &= \mathbb{E}[\mathbf{1}_{X_1^{(t)} \geq y}\mathbf{1}_{\Pi([0,t))=n}] \\
&= \mathbb{E}[\mathbf{1}_{Y_{n+1} \geq t-X_n+y}\mathbf{1}_{X_n<t}\mathbf{1}_{Y_{n+1}\geq t-X_n}] \\
&= \mathbb{E}[\mathbb{E}[\mathbf{1}_{Y_{n+1}\geq t-X_n+y} \mid X_n]\mathbf{1}_{X_n<t}] \\
&= \mathbb{E}[e^{-\lambda y}\mathbb{E}[\mathbf{1}_{Y_{n+1}\geq t-X_n} \mid X_n]\mathbf{1}_{X_n<t}] && \text{car } Y_{n+1} \sim \mathrm{Exp}(\lambda) \text{ indép. de } X_n \\
&= e^{-\lambda y}\mathbb{E}[\mathbf{1}_{X_n<t}\mathbf{1}_{Y_{n+1}\geq t-X_n}] \\
&= e^{-\lambda y}\mathbb{P}(\Pi([0,t)) = n).
\end{aligned}$$

   Ceci montre que $X_1^{(t)}$ est indépendante de $\Pi([0,t))$ et suit la loi exponentielle de paramètre λ. Avec la dernière partie de l'exercice, ceci montre que $(X_1^{(t)}, \Pi([0,t))$ suit la loi $\mathrm{Exp}(\lambda) \otimes \mathrm{Po}(\lambda)$.

4. Soit $n \in \mathbb{N}$. Soit $f : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$ bornée mesurable. On note que sur l'événement $\Pi([0,t)) = n$, on a pour tout $k \geq 2$,

$$X_k^{(t)} - X_{k-1}^{(t)} = Y_{n+k}.$$

   Par conséquent,

$$\begin{aligned}
&\mathbb{E}[f((X_k^{(t)} - X_{k-1}^{(t)})_{k\geq 2}) \mid X_1^{(t)},\, \Pi([0,t))] \\
&= \sum_{n=0}^\infty \mathbb{E}[f((X_k^{(t)} - X_{k-1}^{(t)})_{k\geq 2}) \mid X_1^{(t)},\, \Pi([0,t))]\mathbf{1}_{\Pi([0,t))=n} \\
&= \sum_{n=0}^\infty \mathbb{E}[f((Y_{n+k})_{k\geq 2}) \mid X_{n+1},\, \Pi([0,t))]\mathbf{1}_{\Pi([0,t))=n} && X_1^{(t)} = X_{n+1} - t \text{ on } \{\Pi([0,t)) = n\} \\
&= \sum_{n=0}^\infty \mathbb{E}[f((Y_{n+k})_{k\geq 2})]\mathbf{1}_{\Pi([0,t))=n} && \sigma(Y_1,\ldots,Y_{n+1}) \text{ et } \sigma(Y_{n+2},\ldots) \text{ indépendantes.} \\
&= \mathbb{E}[f(Y_k)_{k\geq 2})]
\end{aligned}$$

Ceci montre l'indépendance entre $(X_k^{(t)} - X_{k-1}^{(t)})_{k\geq 2}$ et $X_1^{(t)}$, $\Pi([0,t))$ et de plus que $(X_k^{(t)} - X_{k-1}^{(t)})_{k\geq 2}$ est égale en loi à $(Y_k)_{k\geq 2}$.

5. Les deux dernières parties montrent que

$$(\Pi([0,t)), X_1^{(t)}, X_2^{(t)} - X_1^{(t)}, X_3^{(t)} - X_2^{(t)}, \dots) \sim \mathrm{Po}(\lambda) \otimes \mathrm{Exp}(\lambda)^{\otimes \mathbb{N}}.$$

Par conséquent, la suite $(X_i^{(t)})_{i\geq 1}$ est égale en loi à $(X_i)_{i\geq 1}$ et indépendante de $\Pi([0,t))$.

6. Preuve par récurrence sur k. $k=0$ est évident. Supposons que pour un certain $k \in \mathbb{N}$ l'énoncé soit vérifié pour tous les $0 = t_0 < t_1 < \ldots < t_k$. Soient $0 = t_0 < t_1 < \ldots < t_{k+1}$. Par la dernière partie de l'exercice, la suite $(X_i^{(t_1)})_{i\geq 1}$ est égale en loi à $(X_i)_{i\geq 1}$ et indépendante de $\Pi([0,t_1))$ qui suit la loi $\mathrm{Po}(\lambda t_1)$. En définissant $\Pi^{(t_1)}$ à partir de $(X_i^{(t_1)})_{i\geq 1}$ de manière analogue à Π et en utilisant l'hypothèse de récurrence pour $\Pi^{(t_1)}$, il en suit que le vecteur $(\Pi([t_1,t_2)), \ldots, \Pi([t_{k-1},t_k))) = (\Pi^{(t_1)}([0,t_2-t_1)), \ldots, \Pi^{(t_1)}([t_{k-1}-t_1, t_k-t_1)))$ est indépendant de $\Pi([0,t_1))$ et que ses composantes sont indépendantes et de lois respectives $\mathrm{Po}(\lambda(t_i - t_{i-1}))$, $i = 2, \ldots, k$. Ceci conclut la preuve.