Module 2

1. (a) $P(X=0)=\sum_{i=1}^{n} p_i^2$
   (b) $P(X=j)=2\sum_{i=1}^{n-j} p_i p_{i+j}$ avec $j \geq 1$.

   Pour (c) et (d), les formules suivantes sont utiles :

$$\begin{aligned}\sum_{i=1}^{n-1} i &= \frac{(n-1)n}{2}\\ \sum_{i=1}^{n-1} i^2 &= \frac{(n-1)n(2n-1)}{6}\\ \sum_{i=1}^{n-1} i^3 &= \frac{(n-1)^2n^2}{4}\end{aligned}$$

   (c) $\mu_X=(n^2-1)/3n$
   (d) $\sigma_X^2=(n^2-1)(n^2+2)/18n^2$.

2. (a) $k=2/\pi$
   (b) $f_X(x)=2\sqrt{1-x^2}/\pi$ si $|x|\leq 1$ et 0 sinon ; $f_Y(y)=4\sqrt{1-y^2}/\pi$ si $0<y\leq 1$ et 0 sinon.
   (c) $f_{Y|X}(y|x)=1/\sqrt{1-x^2}I_{0\leq y\leq\sqrt{1-x^2}}$ si $|x|\leq 1$ et 0 sinon.

3. (a) $k=1$
   (b) $f_X(x)=1-|x|$ si $|x|\leq 1$ et 0 sinon ; $f_Y(y)=2(1-y)$ si $0<y\leq 1$ et 0 sinon.
   (c) $f_{Y|X}(y|x)=1/(1-|x|)I_{0\leq y\leq 1-|x|}$ si $|x|\leq 1$ et 0 sinon.

4. (a) $f_{R\Theta}(r,\theta)=(r/2\pi)\,\exp(-r^2/2)$
   (b) Elles sont indépendentes.
   (c) $f_R(r)=r\exp(-r^2/2)$
   (d) $a=\sqrt{2\ln 2}$

5. $a=\frac{\rho_1(1-\rho_2)}{1-\rho_1^2}$ et $b=\frac{\rho_2-\rho_1^2}{1-\rho_1^2}$.

6. $\mu_{S_n}=n\mu$ et $\sigma_{S_n}^2=(n+2(n-1)\rho)\sigma^2$.

7.

8. Elles sont orthogonales, non corrélées et dépendantes.

9. On a

$$(X_n-a)^2=(X_n-a_n+a_n-a)^2\leq 2(X_n-a_n)^2+2(a_n-a)^2$$

   d'où en prenant les espérances,

$$E[(X_n-a)^2]\leq 2E[(X_n-a_n)^2]+2(a_n-a)^2$$

et ensuite les limites

$$0 \le \lim_{n\to\infty} E[(X_n - a)^2] \le 2\lim_{n\to\infty} E[(X_n - a_n)^2] + 2\lim_{n\to\infty}(a_n - a)^2 = 0.$$

10. $f_{Z|X}(z|x) = f_Y(z-x) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{(z-x)^2}{2}\right)$.

11. (a) On a

$$\begin{aligned} P(\Theta_n \ge k) &= P(\min\{X_1,\dots,X_n\} \ge k) = P(X_1 \ge k,\dots,X_n \ge k) \\ &= P(X_1 \ge k)P(X_2 \ge k)\dots P(X_n \ge k) \\ &= \left(1 - \frac{k-1}{N}\right)^n \end{aligned}$$

(b) $P(\Theta_n \ge k) = P(\Theta_n = k) + P(\Theta_n \ge k+1)$. Ce qui implique

$$\begin{aligned} P(\Theta_n = k) &= P(\Theta_n \ge k) - P(\Theta_n \ge k+1) \\ &= \left(1 - \frac{k-1}{N}\right)^n - \left(1 - \frac{k}{N}\right)^n. \end{aligned}$$

12. (a) On écrit, après quelques calculs

$$\begin{aligned} \frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu}) &= \frac{1}{2\sigma_X^2\sigma_Y^2(1-\rho^2)}\left[\sigma_Y^2(x-\mu_X)^2 + \sigma_X^2(y-\mu_Y)^2\right] \\ &\quad -\frac{1}{2\sigma_X^2\sigma_Y^2(1-\rho^2)}\left[2\rho\sigma_X\sigma_Y(x-\mu_X)(y-\mu_Y)\right] \\ &= \frac{1}{2(1-\rho^2)\sigma_X^2}\left[x - \left(\mu_X + \frac{\rho\sigma_X}{\sigma_Y}(y-\mu_Y)\right)\right]^2 + \frac{(y-\mu_Y)^2}{2\sigma_Y^2} \end{aligned}$$

d'où

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{\infty} f_{XY}(x,y)dx \\ &= \frac{1}{\sqrt{2\pi}\sigma_Y}e^{-\frac{(y-\mu_Y)^2}{2\sigma_Y^2}}\left[\frac{1}{\sqrt{2\pi}\sigma_X\sqrt{1-\rho^2}}\int_{-\infty}^{\infty} e^{-\frac{1}{2(1-\rho^2)\sigma_X^2}\left[x-\left(\mu_X+\frac{\rho\sigma_X}{\sigma_Y}(y-\mu_Y)\right)\right]^2}dx\right] \\ &= \frac{1}{\sqrt{2\pi}\sigma_Y}e^{-\frac{(y-\mu_Y)^2}{2\sigma_Y^2}} \end{aligned}$$

car la partie entre crochets correspond à une v.a. $N\left(\mu_X + \frac{\rho\sigma_X}{\sigma_Y}(y-\mu_Y), \sigma_X\sqrt{1-\rho^2}\right)$.

(b)

$$f_{X|Y}(x|y) = \frac{f_{XY}(x,y)}{f_Y(y)} = \frac{1}{\sqrt{2\pi}\sigma_X\sqrt{1-\rho^2}}e^{-\frac{1}{2(1-\rho^2)\sigma_X^2}\left[x-\left(\mu_X+\frac{\rho\sigma_X}{\sigma_Y}(y-\mu_Y)\right)\right]^2}$$

(c) A partir du résultat en b) on a directement

$$E[X|Y=y] = \mu_X + \frac{\rho\sigma_X}{\sigma_Y}(y-\mu_Y)$$

(d)

$$VAR[X|Y=y] = \sigma_X^2(1-\rho^2)$$

13. Pour $m = n$, on a évidemment $E\left[\frac{S_n}{S_n}\right] = E[1] = 1$. D'autre part, les v.a. X_i etant indépendantes est identiquement distribuées,

$$E\left[\frac{S_1}{S_n}\right] = E\left[\frac{X_1}{X_1+\ldots+X_n}\right] = E\left[\frac{X_2}{X_1+\ldots+X_n}\right] = \ldots = E\left[\frac{X_n}{X_1+\ldots+X_n}\right]$$

Comme

$$\frac{S_n}{S_n} = \frac{X_1+\ldots+X_n}{X_1+\ldots+X_n} = \frac{X_1}{X_1+\ldots+X_n} + \ldots + \frac{X_n}{X_1+\ldots+X_n}$$

en prenant les espérances on obtient

$$E\left[\frac{S_n}{S_n}\right] = E\left[\frac{X_1}{X_1+\ldots+X_n}\right] + \ldots + E\left[\frac{X_n}{X_1+\ldots+X_n}\right] = nE\left[\frac{S_1}{S_n}\right]$$

d'où $E\left[\frac{S_1}{S_n}\right] = \frac{1}{n}$.
Finalement,

$$E\left[\frac{S_m}{S_n}\right] = E\left[\frac{X_1}{X_1+\ldots+X_n}\right] + \ldots + E\left[\frac{X_m}{X_1+\ldots+X_n}\right] = mE\left[\frac{S_1}{S_n}\right] = \frac{m}{n}$$

14. (a) En appliquant le théorème des probabilités totales, on a

$$P(X=n) = \int_0^\infty P(X=n|\Lambda=\lambda)f_\Lambda(\lambda)d\lambda = \frac{1}{n!}\int_0^\infty \lambda^n e^{-2\lambda}d\lambda$$

Pour évaluer cette intégrale, soit on intègre n fois par parties, soit on consulte des tables, soit on pose $x = 2\lambda$ et on remarque que

$$\int_0^\infty \lambda^n e^{-2\lambda}d\lambda = \frac{1}{2^{n+1}}\int_0^\infty x^n e^{-x}dx = \frac{1}{2^{n+1}}E[X^n] = \frac{n!}{2^{n+1}}$$

où $X \sim expo(1)$. Dès lors

$$P(X=n) = \frac{1}{n!}\frac{n!}{2^{n+1}} = \frac{1}{2^{n+1}}$$

(b) *Méthode 1 :* Comme on connait la loi de probabilité $P(X = n)$, on peut calculer la fonction génératrice de X :

$$G_X(z) = \sum_{n=0}^{\infty} \frac{1}{2^{n+1}} z^n = \frac{1}{2} \sum_{n=0}^{\infty} (\frac{z}{2})^n = \frac{1}{2} \frac{1}{1 - \frac{z}{2}} = \frac{1}{2 - z}$$

D'où on tire que

$$E[X] = \left. \frac{\partial G_X(z)}{\partial z} \right|_{z=1} = 1$$

Méthode 2 : à partir du théorème des probabilités totales on a

$$E[X] = \int_0^{\infty} E(X|\Lambda = \lambda) f_\Lambda(\lambda) d\lambda = \int_0^{\infty} \lambda e^{-\lambda} d\lambda = 1$$

En effet $E(X|\Lambda = \lambda)$ est l'espérance d'une v.a. de Poisson de paramètre Λ fixé égal à λ. Remarquez que cette méthode ne nécessite pas la connaissance des $P(X = n)$.

(c) *Méthode 1 :*

$$E[X^2] = E[X(X-1)] + E[X] = \left. \frac{\partial^2 G_X(z)}{\partial z^2} \right|_{z=1} + 1 = \left. \frac{2}{(2-z)^3} \right|_{z=1} + 1 = 3$$

d'où

$$Var[X] = E[X^2] - E[X]^2 = 3 - 1 = 2$$

Méthode 2 : à partir du théorème des probabilités totales on a

$$E[X^2] = \int_0^{\infty} E(X^2|\Lambda = \lambda) f_\Lambda(\lambda) d\lambda = \int_0^{\infty} (\lambda + \lambda^2) e^{-\lambda} d\lambda = 3$$

d'où

$$Var[X] = E[X^2] - E[X]^2 = 3 - 1 = 2$$

15. (a) On a

$$\begin{aligned} P(\text{“Record au temps n”}) &= P(X_n > max\{X_1, \ldots, X_{n-1}\}) \\ &= P(X_n = max\{X_1, \ldots, X_{n-1}, X_n\}) \\ &= \frac{1}{n} \end{aligned}$$

car les $X_1, \ldots, X_n$ sont i.i.d (il y a exactement une chance sur n que n'importe lequel des X_i, en particulier X_n, soit le plus grand des n variables).

(b)

$$\begin{aligned}
&E[\text{“nombre de records au temps n”}]\\
&= E[\sum_{i=1}^{n} I(\text{“record au temps } i\text{”})]\\
&= \sum_{i=1}^{n} E[I(\text{“record au temps } i\text{”})] = \sum_{i=1}^{n} p_i = \sum_{i=1}^{n} \frac{1}{i}
\end{aligned}$$

16. On a

$$\begin{aligned}
\mu_Y &= E[Y] = aE[X] + b = a\mu_X + b\\
\sigma_Y^2 &= E[(Y-\mu_Y)^2] = E[(aX+b-(a\mu_X+b))^2] = a^2E[(X-\mu_X)^2] = a^2\sigma_X^2\\
COV[X,Y] &= E[(X-\mu_X)(Y-\mu_Y)] = E[(X-\mu_X)(aX+b-(a\mu_X+b))]\\
&= aE[(X-\mu_X)^2] = a\sigma_X^2
\end{aligned}$$

D'où

$$\rho(X,Y) = \frac{COV[X,Y]}{\sigma_X\sigma_Y} = \frac{a\sigma_X^2}{\sigma_X|a|\sigma_X} = \frac{a}{|a|} = \begin{cases} 1 & \text{si a>0} \\ -1 & \text{si a<0} \end{cases}$$

N.B. Si $a = 0$, ρ n'est pas défini.

17. *Méthode 1 :* on se souvient que toute transformation linéaire $\mathbf{Y} = A\mathbf{X}$ de v.a. gaussiennes donne encore des v.a. gaussiennes, telles que

$$\begin{aligned}
\mu_{\mathbf{Y}} &= A\mu_{\mathbf{X}} = A\begin{bmatrix}0\\0\\0\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix}\\
\Sigma_{\mathbf{Y}} &= A\Sigma_{\mathbf{X}}A^T = \begin{bmatrix}1&1&1\\1&-1&0\\1&0&-1\end{bmatrix}\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}\begin{bmatrix}1&1&1\\1&-1&0\\1&0&-1\end{bmatrix} = \begin{bmatrix}3&0&0\\0&2&1\\0&1&2\end{bmatrix}
\end{aligned}$$

D'où

$$\begin{aligned}
f_{\mathbf{Y}}(\mathbf{y}) &= \frac{1}{(2\pi)^{\frac{3}{2}}\sqrt{|det\Sigma_{\mathbf{Y}}|}} e^{-\frac{1}{2}(\mathbf{y}-\mu_{\mathbf{Y}})^T\Sigma_{\mathbf{Y}}^{-1}(\mathbf{y}-\mu_{\mathbf{Y}})}\\
f_{Y_1,Y_2,Y_3}(y_1,y_2,y_3) &= \frac{1}{(2\pi)^{\frac{3}{2}}3}\exp\left(-\frac{1}{2}[y_1\ y_2\ y_3]\frac{1}{9}\begin{bmatrix}3&0&0\\0&6&-3\\0&-3&6\end{bmatrix}\begin{bmatrix}y_1\\y_2\\y_3\end{bmatrix}\right)\\
&= \frac{1}{(2\pi)^{\frac{3}{2}}3} e^{-\frac{1}{6}(y_1^2+2y_2^2+2y_3^2-2y_2y_3)}
\end{aligned}$$

Méthode 2 : on ne se souvient pas de la propriété précédente des v.a. gaussiennes et on effectue un changement de variables classique

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \Leftrightarrow$$

$$\Leftrightarrow \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} y_1 + y_2 + y_3 \\ y_1 - 2y_2 + y_3 \\ y_1 + y_2 - 2y_3 \end{bmatrix}$$

D'où

$$\begin{aligned} f_{Y_1,Y_2,Y_3}(y_1, y_2, y_3) &= \frac{1}{3} f_{X_1,X_2,X_3}\left(\frac{1}{3}(y_1 + y_2 + y_3), \frac{1}{3}(y_1 - 2y_2 + y_3), \frac{1}{3}(y_1 + y_2 - 2y_3)\right) \\ &= \frac{1}{3(2\pi)^{\frac{3}{2}}} e^{-\frac{1}{18}(y_1+y_2+y_3)^2} e^{-\frac{1}{18}(y_1-2y_2+y_3)^2} e^{-\frac{1}{18}(y_1+y_2-2y_3)^2} \\ &= \frac{1}{3(2\pi)^{\frac{3}{2}}} e^{-\frac{1}{6}(y_1^2+2y_2^2+2y_3^2-2y_2y_3)} \end{aligned}$$

Remarquons qu'Y_1 est indépendante de Y_2 et Y_3, mais que Y_2 et Y_3 ne sont pas indépendantes car on peut écrire l'expression précédante comme

$$f_{Y_1,Y_2,Y_3}(y_1, y_2, y_3) = \frac{1}{\sqrt{3}\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{y_1}{\sqrt{3}})^2} \frac{1}{2\pi\sqrt{3}} e^{-\frac{1}{3}(y_2^2+y_3^2-y_2y_3)}$$

18. (a) On pose

$$\begin{array}{l} u = x + y \\ v = \frac{x}{x+y} \end{array} \Leftrightarrow \begin{array}{l} x = uv \\ y = u(1 - y) \end{array}$$

$$J = \begin{vmatrix} 1 & 1 \\ \frac{y}{(x+y)^2} & -\frac{x}{(x+y)^2} \end{vmatrix} = -\frac{x + y}{(x + y)^2} = -\frac{1}{x + y} = -\frac{1}{u}$$

$$\begin{aligned} f_{UV}(u, v) &= \frac{f_{XY}(uv, u(1 - v))}{\left|-\frac{1}{u}\right|} = u f_X(uv) f_Y(u(1 - v)) \\ &= u \frac{\lambda(\lambda uv)^{\alpha-1} e^{-\lambda uv}}{\Gamma(\alpha)} \frac{\lambda(\lambda u(1 - v))^{\beta-1} e^{-\lambda u(1-v)}}{\Gamma(\beta)} \\ &= \frac{\lambda(\lambda u)^{\alpha+\beta-1} e^{-\lambda u}}{\Gamma(\alpha)} \frac{v^{\alpha-1}(1 - v)^{\beta-1}}{\Gamma(\beta)} \\ &= \frac{\lambda(\lambda u)^{\alpha+\beta-1} e^{-\lambda u}}{\Gamma(\alpha + \beta)} \frac{v^{\alpha-1}(1 - v)^{\beta-1}\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \end{aligned}$$

(b) Comme le premier facteur ci-dessus est une densité de probabilité d'une v.a. $G(\lambda, \alpha+\beta)$, on a écrit $f_{UV}(u,v)$ comme le produit d'une fonction ne dépendant que de u par une autre ne dépendant que de v. De plus, chacune est normalisée pour être une densité de probabilité, et $f_{UV}(u,v) = f_U(u)f_V(v)$implique que U et V sont indépendantes.

(c) $U \sim G(\lambda, \alpha+\beta)$

19. Soit $\varepsilon > 0$. Comme $|X_n - X| \geq \varepsilon$ si et seulement si $(X_n - X)^2 \geq \varepsilon^2$,

$$P(|X_n - X| \geq \varepsilon) = P((X_n - X)^2 \geq \varepsilon^2)$$

et l'inégalité de Markov implique que

$$P(|X_n - X| \geq \varepsilon) = P((X_n - X)^2 \geq \varepsilon^2) \leq \frac{E[(X_n - X)^2]}{\varepsilon^2}.$$

En prenant la limite pour $n \to \infty$ dans l'équation précédente, la convergence en moyenne quadratique de $\{X_n\}$ implique que

$$\lim_{n\to\infty} P(|X_n - X| \geq \varepsilon) \leq \lim_{n\to\infty} \frac{E[(X_n - X)^2]}{\varepsilon^2} = 0,$$

et donc que pour tout $\varepsilon > 0$, $\lim_{n\to\infty} P(|X_n - X| \geq \varepsilon) = 0$, ce qui établit la convergence en probabilité de $\{X_n\}$.

Le contraire n'est pas vrai : la suite de v.a $\{X_n\}_{n\geq 1}$ avec

$$\begin{aligned} X_n &= \sqrt{n} \quad \text{avec probabilité } 1/n \\ &= 0 \quad \text{avec probabilité } 1 - 1/n. \end{aligned} \tag{49}$$

ne converge pas en moyenne quadratique vers 0 car $E[X_n^2] = 1$ pour tout $n \geq 1$ et donc

$$\lim_{n\to\infty} E[(X_n - 0)^2] = 1 \neq 0,$$

mais converge en probabilité vers 0 car pour tout $\varepsilon > 0$,

$$\lim_{n\to\infty} P(|X_n - 0| \geq \varepsilon) = \lim_{n\to\infty} P(X_n \geq \varepsilon) = \lim_{n\to\infty} \frac{1}{n} = 0.$$

20. $(\Leftarrow)$ On vérifie tout d'abord que

$$\begin{aligned} \frac{|X_n|}{1+|X_n|} \geq \frac{\varepsilon}{1+\varepsilon} \quad &\iff \quad |X_n|(1+\varepsilon) \geq \varepsilon(1+|X_n|) \\ &\iff \quad |X_n| \geq \varepsilon, \end{aligned}$$

et donc, en utilisant l'inégalité de Markov,

$$P\left(|X_n| \geq \varepsilon\right) = P\left(\frac{|X_n|}{1+|X_n|} \geq \frac{\varepsilon}{1+\varepsilon}\right) \leq E\left[\frac{|X_n|}{1+|X_n|}\right] \Big/ \left(\frac{\varepsilon}{1+\varepsilon}\right) \to 0$$

pour $n \to \infty$.

($\Rightarrow$) Soit $\varepsilon > 0$. Comme $|X_n|/(1+|X_n|) \leq 1$ et $|X_n|/(1+|X_n|) \leq \varepsilon/(1+\varepsilon)$ si et seulement si $|X_n| \leq \varepsilon$, le théorème des probabilités totales entraîne que

$$\begin{aligned}
E\left[\frac{|X_n|}{1+|X_n|}\right] &= E\left[\frac{|X_n|}{1+|X_n|} \mid |X_n| < \varepsilon\right] P(|X_n| < \varepsilon) + E\left[\frac{|X_n|}{1+|X_n|} \mid |X_n| \geq \varepsilon\right] P(|X_n| \geq \varepsilon) \\
&\leq \frac{\varepsilon}{1+\varepsilon} P(|X_n| < \varepsilon) + 1 \cdot P(|X_n| \geq \varepsilon) \\
&= \frac{\varepsilon}{1+\varepsilon} + \frac{1}{1+\varepsilon} P(|X_n| \geq \varepsilon).
\end{aligned}$$

Comme $\lim_{n\to\infty} P(|X_n| \geq \varepsilon) = 0$, et qu'on peut prendre $\varepsilon > 0$ aussi petit qu'on le souhaite, la dernière égalité ci-dessus implique que

$$\lim_{n\to\infty} E\left[\frac{|X_n|}{1+|X_n|}\right] = 0.$$