VARIABLES ALÉATOIRES DISCRÈTES

I Variables aléatoires discrètes

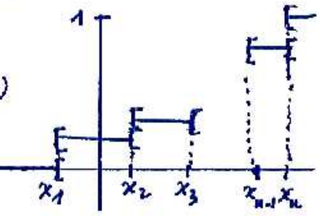
<u>ex</u>: On lance 3 pièces. $\Omega = \{P,F\}^3$, au résultat $\omega \in \Omega$, on associe $X(\omega)$ = nombre de faces dans ω

$X(PPP) = 0$, $X(PPF) = X(PFP) = X(FPP) = 3$, $X(PFF) = X(FPF) = X(FFP) = 2$, $X(FFF) = 3$.
$\mathbb{P}(X=2) = \mathbb{P}(\{PFF\} \cup \{FPF\} \cup \{FFP\}) = \mathbb{P}(X^{-1}(2)) = \frac{3}{8}$.

<u>Définition</u>: Soit $(\Omega, \mathcal{T}, \mathbb{P})$ un espace probabilisé et $X : (\Omega, \mathcal{T}) \to \mathbb{R}$.
X v.a. discrète $\iff$ $X(\Omega)$ fini ou dénombrable et $\forall x \in \mathbb{R}$, $X^{-1}(\{x\}) \in \mathcal{T}$.
On considère alors $\mathbb{P}(X^{-1}(x)) = \mathbb{P}(\{\omega \in \Omega : X(\omega) = x\})$ noté $\mathbb{P}(X=x)$ ou $\mathbb{P}_X(\{x\})$
et pour $A \subset \mathbb{R}$ $\mathbb{P}(X^{-1}(A)) = \mathbb{P}(\{\omega \in \Omega : X(\omega) \in A\})$ noté $\mathbb{P}(X \in A)$ ou $\mathbb{P}_X(A)$

$\mathbb{P}_X = \mathbb{P} \circ X^{-1}$ est une probabilité sur $\mathbb{R}$ appelée loi de X.
Pour une v.a. discrète X prenant les valeurs $x_1, x_2, \dots$, la loi de X est caractérisée par $\mathbb{P}_X(\{x_1\})$, $\mathbb{P}_X(\{x_2\})$, ...

<u>exemples fondamentaux</u>:
1) <u>loi de Bernoulli</u>; $X : \Omega \to \{0,1\}$, $\mathbb{P}(X=0) = p$, $\mathbb{P}(X=1) = q = 1-p$.
2) <u>loi binomiale</u> $\mathcal{B}(n,p)$: $X : \Omega \to \{0,1,\dots,n\}$, $\mathbb{P}(X=i) = C_n^i \, p^i q^{n-i}$, $q = 1-p$.
3) <u>loi de Poisson</u> $\mathcal{P}(\lambda)$: $X : \Omega \to \mathbb{N}$, $\mathbb{P}(X=i) = \frac{\lambda^i e^{-\lambda}}{i!}$, $\lambda > 0$.



<u>Fonction de répartition</u>: $F_X(x) = \mathbb{P}(X \leq x) = \mathbb{P}_X(]-\infty, x]) = \sum_{i : x_i \leq x} \mathbb{P}(X = x_i)$

<u>Proposition</u>: F_X est càdlàg, constante par morceaux, croissante $F(-\infty) = 0$, $F(+\infty) = 1$
$\mathbb{P}(X < x_i) = F_X(x_i^-)$, $\mathbb{P}(X = x_i) = F_X(x_i) - F_X(x_i^-)$

<u>Démonstration</u>: $\mathbb{P}(X < x_i) = \mathbb{P}\left[X^{-1}\left(\bigcup_{n=1}^{\infty}]-\infty, x_i - \tfrac{1}{n}]\right)\right] = \sup_{n \geq 1} \mathbb{P}(X \leq x_i - \tfrac{1}{n}) = F_X(x_i^-)$. $\square$

<u>Proposition</u>:
* Soit X, Y deux v.a. discrètes, $\alpha \in \mathbb{R}$. Alors $X+Y$, αX, XY, $\max(X,Y)$, $\min(X,Y)$ sont des v.a. discrètes.
* Soit $f : X(\Omega) \to \mathbb{R}$ une application. Alors $Y = f(X)$ est une v.a. discrète et $\mathbb{P}(Y = y) = \sum_{i : f(x_i) = y} \mathbb{P}(X = x_i)$.

II Couples aléatoires discrets

<u>Définition</u>: Soit X, Y deux v.a. discrètes. (X,Y) couple aléatoire. Vecteurs aléatoires : idem.
Fonction de répartition conjointe : $F_{(X,Y)}(x,y) = \mathbb{P}(X \leq x, Y \leq y) = \mathbb{P}_{(X,Y)}(]-\infty, x] \times]-\infty, y])$.
Distribution de proba conjointe : $\{\mathbb{P}(X = x, Y = y),\ x \in X(\Omega), y \in Y(\Omega)\}$.

<u>Indépendance</u>: X, Y indépendantes $\iff$ $\forall (x,y) \in \mathbb{R}^2$, $F_{(X,Y)}(x,y) = F_X(x) F_Y(y)$, i.e. $F_{(X,Y)} = F_X \otimes F_Y$
$\iff$ $\forall (x,y) \in \mathbb{R}^2$, $\{X \leq x\}$ et $\{Y \leq y\}$ indépendants.
$\iff$ $\forall (x,y) \in X(\Omega) \times Y(\Omega)$, $\{X = x\}$ et $\{Y = y\}$ indépendants.
$\iff$ $\forall (x,y) \in X(\Omega) \times Y(\Omega)$, $\mathbb{P}_{(X,Y)}(x,y) = \mathbb{P}_X(x) \mathbb{P}_Y(y)$, i.e. $\mathbb{P}_{(X,Y)} = \mathbb{P}_X \otimes \mathbb{P}_Y$

<u>Lois marginales</u>: $F_X(x) = F_{(X,Y)}(x, +\infty) = \sum_{y \in Y(\Omega)} \mathbb{P}(X \leq x, Y = y)$ (une fonction de répartition marginale)

$\mathbb{P}_X(x) = \sum_{y \in Y(\Omega)} \mathbb{P}_{(X,Y)}(\{(x,y)\})$ (une loi marginale)

exemple : loi géométrique $T : \Omega \to \overline{\mathbb{N}^*}$: $\mathbb{P}(T=n) = pq^{n-1}$, $q = 1-p$. $\mathbb{P}(T \geq n) = q^{n-1}$.
T est un temps d'attente. En effet, si $X_1, X_2, \dots$, sont des v.a. de Bernoulli de m̂ paramètre p indépendantes, $\mathbb{P}(T=n) = \mathbb{P}(X_1 = X_2 = \cdots = X_{n-1} = 0, X_n = 1)$. $\mathbb{P}(T=\infty) = 0$.

Somme de v.a. discrètes indépendantes : Soit X, Y deux v.a. discrètes indépendantes.

Alors $X+Y$ est une v.a. discrète dont la loi est donnée par

$$\mathbb{P}_{X+Y}(\{z\}) = \sum_{x+y=z} \mathbb{P}_X(\{x\})\mathbb{P}_Y(\{y\}) \quad \text{i.e.} \quad \mathbb{P}_{X+Y} = \mathbb{P}_X * \mathbb{P}_Y .$$

Cas $X(\Omega), Y(\Omega) \subset \mathbb{N}$: $\mathbb{P}_{X+Y}(\{n\}) = \sum_{k=0}^{n} \mathbb{P}_X(\{k\})\, \mathbb{P}_Y(\{n-k\})$

remarque : $\to$ série produit : $\mathbb{P}_{X+Y}(\mathbb{N}) = 1 = \mathbb{P}_X(\mathbb{N}) \times \mathbb{P}_Y(\mathbb{N})$.

exemple : loi binomiale = somme de n v.a. de Bernoulli de m̂ paramètre p indépendantes.
En effet : $X_1, \dots, X_n$ $\mathcal{B}(n,p)$ indépendantes.

$$\mathbb{P}(X_1 + \cdots + X_n = i) = \sum_{\substack{i_1+\cdots+i_n = i \\ i_1,\dots,i_n \in \{0,1\}}} \mathbb{P}(X_1 = i_1, X_2 = i_2, \dots, X_n = i_n) = C_n^i p^i q^{n-i}$$

($\to$ il y a i 1 et $(n-i)$ 0)

III Espérance mathématique, variance - stat. descriptive

Exemple d'introduction : - lancer de 2 pièces, S = nb de piles obtenus. $\{S=1\}$ a 2 fois plus de chance de se réaliser que $\{S=0\}$ et $\{S=2\}$ $\to$ moyenne pondérée.

Définition : Soit X une v.a. discrète ; $\mathbb{E}(X) = \sum_{x \in X(\Omega)} x\, \mathbb{P}(X=x)$ si la série est convergente.

Moments d'ordre r : $\mathbb{E}(X^r) = \sum_{x \in X(\Omega)} x^r\, \mathbb{P}(X=x)$

Moments centré d'ordre r : $\mathbb{E}[(X-\mathbb{E}(X))^r] = \sum_{x\in X(\Omega)} (x - \mathbb{E}(X))^r\, \mathbb{P}(X=x)$

Variance : $\mathrm{var}(X) = \mathbb{E}[(X-\mathbb{E}(X))^2] = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 \geq 0$

Ecart-type : $\sigma(X) = \sqrt{\mathrm{var}(X)}$.

dém "$\mathbb{E}$ additive" : $\mathbb{E}(X) + \mathbb{E}(Y) = \sum x\, \mathbb{P}(X=x, Y=y) + \sum y\, \mathbb{P}(X=x, Y=y) = \sum_s s \sum_{x+y=s} \mathbb{P}(X=x, Y=y)$
$= \sum_s s\, \mathbb{P}(X+Y=s) = \mathbb{E}(X+Y)$.

Proposition : $\mathbb{E}$ est linéaire ; $\mathbb{E}(X^2) < +\infty \Rightarrow \mathbb{E}(|X|) < +\infty$ (résulte de $|X| \leq X^2 \vee 1$) ; $\mathbb{E}(\mathbb{1}_A) = \mathbb{P}(A)$

Exemples : 1) Loi de Bernoulli $\mathcal{B}(p)$: $\mathbb{E}(X) = p$, $\mathbb{E}(X^2) = p$, $\mathrm{var}(X) = pq$.

2) Loi binomiale $\mathcal{B}(n,p)$: $\mathbb{E}(X) = \sum_{i=0}^{n} i\, C_n^i\, p^i q^{n-i} = np$ (dérivée du binôme ou somme de Bernoulli)

$$\mathbb{E}(X^2) = \sum_{i=0}^{n} \underbrace{i^2}_{i(i-1)+i} C_n^i\, p^i q^{n-i} = n(n-1)p^2 + np \quad \text{(dérivée seconde)}$$

$$\Rightarrow \mathrm{var}(X) = npq .$$

3) Loi de Poisson $\mathcal{P}(\lambda)$: $\mathbb{E}(X) = \sum_{n=0}^{\infty} n e^{-\lambda} \frac{\lambda^n}{n!} = \lambda$, $\mathrm{var}(X) = \lambda$.

Siméon-Denis Poisson (1781-1840)

4) Loi géométrique : $\mathbb{E}(T) = \sum_{n=1}^{\infty} n p q^{n-1} = \frac{p}{(1-q)^2} = \frac{1}{p}$

$$\mathbb{E}(T^2) = \sum_{n=1}^{\infty} n^2 p q^{n-1} = \frac{2p}{(1-q)^3} + \frac{p}{(1-q)^2} = \frac{2}{p^2} + \frac{1}{p} \Rightarrow \mathrm{var}(T) = \frac{q}{p^2}.$$

Inégalités de Markov, Bienaymé-Tchebychev :

Andreï Markov 1856-1922

Jules Bienaymé 1796-1878

Tchebychev 1821-1894

$$\mathbb{P}(|X| \geq x) \leq \frac{\mathbb{E}(|X|)}{x} \quad \left(\text{ou } \frac{\mathbb{E}(|X|^r)}{x^r},\ \mathbb{E}(e^X)e^{-x}\right)$$

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq \alpha) \leq \frac{\mathrm{var}(X)}{\alpha^2} \quad \left(\text{ou } \frac{\mathbb{E}(X-\mathbb{E}X)^2}{\alpha^2}, \dots\right)$$

<u>Définition</u> : Soit X, Y deux v.a. discrètes admettant des moments d'ordre 2 finis.

$cov(X,Y) = \mathbb{E}[(X-\mathbb{E}(X))(Y-\mathbb{E}(Y))] = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$

$\rho(X,Y) = \dfrac{cov(X,Y)}{\sigma(X)\sigma(Y)}$ → coefficient de corrélation

Matrice de covariances $\begin{pmatrix} \sigma_X^2 & \rho(X,Y)\sigma_X\sigma_Y \\ \rho(X,Y)\sigma_X\sigma_Y & \sigma_Y^2 \end{pmatrix}$ $\qquad \sigma_X^2 = var(X) = cov(X,X)$

Soit $r_{ij} = \mathbb{P}(X=x_i, Y=y_j)$. $\qquad \mathbb{E}(XY) = \sum_{i,j} x_i y_j r_{ij}$, $\mathbb{E}(X) = \sum_i x_i p_i$, $\mathbb{E}(Y) = \sum_j y_j q_j$.

$p_i = \mathbb{P}(X=x_i) \quad q_j = \mathbb{P}(Y=y_j)$

$cov(X,Y) = \sum_{i,j} x_i y_j (r_{ij} - p_i q_j)$

<u>Proposition</u> :

* inégalité de Cauchy-Schwarz. $\mathbb{E}(|XY|) \leq \sqrt{\mathbb{E}(X^2)\mathbb{E}(Y^2)}$, et donc $|\rho(X,Y)| \leq 1$.

* X, Y indépendantes $\Rightarrow \mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y) \Leftrightarrow cov(X,Y) = 0$.

* $var(aX+b) = a^2 var(X)$, $cov(aX+b, cY+d) = ac\, cov(X,Y)$

  $var(X \pm Y) = var(X) + var(Y) \pm 2cov(X,Y)$ (cov : forme bilinéaire, var : f. quad.)

* $X_1, \dots, X_n$ indépendantes $\Rightarrow var(X_1 + \dots + X_n) = var(X_1) + \dots + var(X_n)$.

dém de $E(XY) = E(X)E(Y)$: $E(X)E(Y) = \sum x\, \mathbb{P}(X=x) \sum y\, \mathbb{P}(Y=y) = \sum xy\, \mathbb{P}(X=x, Y=y) = \sum_z z \sum_{xy=z} \mathbb{P}(X=x, Y=y) = \sum_z z\, \mathbb{P}(XY=z) = E(XY)$

Contre-exemple : X, X' : Bernoulli (p) indép., $Y = X+X'$: $B(2,p)$. Alors $E(XY) = E(X)E(Y)$ et X, Y non indép.

<u>Annexe</u> : <u>médiane</u>

<u>Définition</u> : une médiane de X est un nombre m tel que $\mathbb{P}(X \geq m) \geq \frac{1}{2}$ et $\mathbb{P}(X \leq m) \geq \frac{1}{2}$

ou encore $\mathbb{P}(X < m) \leq \frac{1}{2} \leq \mathbb{P}(X \leq m)$.

<u>Proposition</u> : une médiane minimise $m \mapsto E(|X-m|)$.

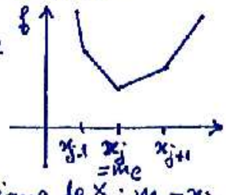
<u>Démonstration</u> : Posons $f(m) = \mathbb{E}(|X-m|) = \sum_{i=1}^{n} p_i |m - x_i|$ où $p_i = \mathbb{P}(X = x_i)$.

f est une fonction continue, affine par morceaux, convexe.

Chaque segment a pour pente $\sum_{i=1}^{k} p_i - \sum_{i=k+1}^{n} p_i = 2\sum_{i=1}^{k} p_i - 1$ sur $]x_k, x_{k+1}[$.

1^er^ cas : si on n'a jamais $\sum_{i=1}^{k} p_i = \frac{1}{2}$, il existe un unique j tel que

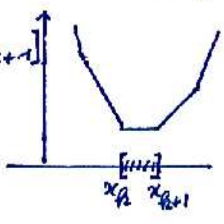
$\sum_{i=1}^{j-1} p_i < \frac{1}{2} < \sum_{i=1}^{j} p_i$, donc $2\sum_{i=1}^{j-1} p_i - 1 < 0 < 2\sum_{i=1}^{j} p_i - 1$



Donc x_j est un minimum strict pour f et c'est la médiane de X : $m_e = x_j$.

2^e^ cas : si $\exists k$ tel que $\sum_{i=1}^{k} p_i = \frac{1}{2}$, alors f est constante sur $[x_k, x_{k+1}]$

et c'est l'intervalle médian de X. □



<u>Remarque</u> : difficile à déterminer même dans les cas classiques → peu utilisé.

IV Fonction génératrice

<u>Définition</u> : soit X une v.a. discrète, $X : \Omega \to \mathbb{N}$.
Fonction génératrice de X : $\forall z \in \mathbb{C},\ G_X(z) = \sum_{n=0}^{+\infty} \mathbb{P}(X=n)\, z^n = \mathbb{E}(z^X)$

Le rayon de convergence est ≥ 1. $G_X(1) = \sum_{n=0}^{\infty} \mathbb{P}(X=n) = \mathbb{P}(X<+\infty) = 1$
(peut être ≤ 1 si $X : \Omega \to \overline{\mathbb{N}}$)

<u>Proposition</u> :
- Si $\mathbb{E}(X)$ existe, $\mathbb{E}(X) = \sum_{n=0}^{\infty} n\, \mathbb{P}(X=n) = G_X'(1)$
- Si $\mathbb{E}(X^2)$ existe ($\mathbb{E}(X)$ existe alors aussi), $\mathbb{E}(X(X-1)) = G_X''(1)$
$\Rightarrow \operatorname{var}(X) = G_X''(1) + G_X'(1) - G_X'(1)^2$

<u>Théorème</u> : G_X caractérise entièrement $X : \Omega \to \mathbb{N}$: la loi de X est donnée par les coefficients de G_X.

<u>Théorème</u> : Si $X_1, \dots, X_n$ sont indépendantes, $G_{X_1+\dots+X_n} = G_{X_1} \times \dots \times G_{X_n}$.

<u>exemples</u> :
1) <u>Loi de Bernoulli</u> : $G_X(z) = pz + q$
2) <u>Loi binomiale $B(n,p)$</u> : $G_X(z) = (pz+q)^n$ (somme de Bernoulli indép.)
somme de deux lois binomiales $B(n_1,p), B(n_2,p)$ indép. $G_X(z) = (pz+q)^{n_1+n_2}$ $\to B(n_1+n_2,p)$
3) <u>Loi de Poisson $\mathcal{P}(\lambda)$</u> : $G_X(z) = e^{\lambda(z-1)}$
somme de deux lois de Poissons $\mathcal{P}(\lambda_1), \mathcal{P}(\lambda_2)$ indép. $G_X(z) = e^{(\lambda_1+\lambda_2)(z-1)} \to \mathcal{P}(\lambda_1+\lambda_2)$
4) <u>Loi géométrique</u> : $G_X(z) = \dfrac{pz}{1-qz}$

<u>Application</u> : <u>formule de Poincaré généralisée</u> :

rappel : $\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{k=1}^{n} (-1)^k S_k$ où $S_k = \sum_{1 \leq i_1 < \dots < i_k \leq n} \mathbb{P}(A_{i_1} \cap \dots \cap A_{i_k})$

c'est la probabilité qu'au moins l'un des A_i soit réalisé.

On cherche maintenant la probabilité qu'exactement m A_i soient réalisés.
Soit X la v.a. représentant le nombre d'A_i réalisés.

On a $S_k = \mathbb{E}\left[\sum_{1 \leq i_1 < \dots < i_k \leq n} \mathbb{1}_{A_{i_1} \cap \dots \cap A_{i_k}}\right] = \mathbb{E}\left[\omega \mapsto \operatorname{card}\{(i_1, \dots, i_k) ; \omega \in A_{i_1} \cap \dots \cap A_{i_k}\}_{1 \leq i_1 < \dots < i_k \leq n}\right]$

$= \mathbb{E}(C_X^k) = \sum_{j=k}^{n} C_j^k\, \mathbb{P}(X=j)$

$S_k = \sum_{j=k}^{n} C_j^k\, \mathbb{P}(X=j)$ $\to$ système de $(n+1)$ équations aux inconnues $\mathbb{P}(X=j)$.

On va résoudre ce système à l'aide des fonctions génératrices :

On pose $G_S(z) = \sum_{k=0}^{n} S_k z^k$ et $G_X(z) = \sum_{j=k}^{n} \mathbb{P}(X=j) z^j$.

On a $G_S(z) = \sum_{0 \leq k \leq j \leq n} C_j^k z^k\, \mathbb{P}(X=j) = \sum_{j=0}^{n} (1+z)^j\, \mathbb{P}(X=j) = G_X(z+1)$

d'où $G_X(z) = G_S(z-1)$, et l'on en tire le coefficient de z^m :

$$\boxed{\mathbb{P}(X=m) = \sum_{k=m}^{n} (-1)^{k-m} C_k^m S_k}$$

ex : * $m = 0$: $\mathbb{P}(X \geq 1) = 1 - \mathbb{P}(X=0) = \sum_{k=1}^{n} (-1)^k S_k$ (Poincaré)

* Probabilité qu'au moins m A_i soient réalisés :
$\mathbb{P}(X \geq m) = \sum_{m \leq \ell \leq k \leq n} (-1)^{k-\ell} C_k^\ell S_k \Rightarrow \boxed{\mathbb{P}(X \geq m) = \sum_{k=m}^{n} (-1)^{k-m} C_{k-1}^{m-1} S_k}$.

V Loi conditionnelle

Définition : Soit X, Y deux v.a. discrètes sur $(\Omega, \mathcal{E}, \mathbb{P})$. On suppose $\mathbb{P}(Y=y_j)>0$ (j fixé).

1) Fonction de répartition conditionnelle de X sachant $Y=y_j$:
$$F_{X|Y}(x|y_j) = \mathbb{P}(X\leq x|Y=y_j)$$

2) Loi de probabilité conditionnelle de X sachant $Y=y_j$: $\{\mathbb{P}(X=x_i|Y=y_j),\ i\in\mathbb{N}\}$.

3) Espérance mathématique conditionnelle de X sachant $Y=y_j$:
$$\mathbb{E}(X|Y=y_j) = \sum_{i\in\mathbb{N}} x_i\, \mathbb{P}(X=x_i|Y=y_j).$$

On a $\mathbb{P}(X=x_i|Y=y_j) = \dfrac{r_{ij}}{q_j}$ avec $r_{ij} = \mathbb{P}(X=x_i, Y=y_j)$ et $q_j = \mathbb{P}(Y=y_j) = \sum_i r_{ij}$

Donc $\sum_i \mathbb{P}(X=x_i|Y=y_j) = 1$ et $\mathbb{P}(\cdot|Y=y_j)$ est bien une probabilité.

Théorème :
- $\mathbb{P}(X=x) = \sum_{y\in Y(\Omega)} \mathbb{P}(X=x|Y=y)\,\mathbb{P}(Y=y)$, (formule de désintégration) ou encore :
- $\mathbb{E}(X) = \sum_{y\in Y(\Omega)} \mathbb{E}(X|Y=y)\,\mathbb{P}(Y=y)$

$\mathbb{P}_X = \mathbb{P}_Y\,\mathbb{P}_{X|Y}$ avec $\mathbb{P}_X = (\mathbb{P}(X=x_1), \dots, \mathbb{P}(X=x_m))$
$\mathbb{P}_Y = (\mathbb{P}(Y=y_1), \dots, \mathbb{P}(Y=y_n))$
matrice stochastique : $\mathbb{P}_{X|Y} = (\mathbb{P}(X=x_j|Y=y_i))_{\substack{1\leq i\leq n\\ 1\leq j\leq m}}$

exemples :

1) Pendant un temps donné, le nombre N de véhicules passant le péage est une v.a. de Poisson $\mathcal{P}(\lambda)$. Si $N=n$, le nombre F de conducteurs féminins est une v.a. Binomiale $\mathcal{B}(n,p)$. Calculer $\mathbb{E}(F)$.
$$\mathbb{E}(F) = \sum_n \underbrace{\mathbb{E}(F|N=n)}_{np}\,\mathbb{P}(N=n) = p\,\mathbb{E}(N) = \lambda p.$$

2) Loi géométrique : $\mathbb{P}(T=n) = pq^{n-1}$. T est une v.a. sans mémoire car
$$\underbrace{\mathbb{P}(T\geq n+n_0 \mid T>n_0)}_{\sum_{i=n+n_0}^{\infty} pq^{k-1} / \sum_{i=n_0+1}^{\infty} pq^{k-1} = q^{n-1}} = \underbrace{\mathbb{P}(T\geq n)}_{=\sum_{k=n}^{\infty} pq^{k-1}} \text{ et } \underbrace{\mathbb{P}(T=n+n_0|T>n_0)}_{pq^{n+n_0-1}/q^{n_0}} = \mathbb{P}(T=n).$$

et c'est la seule : on doit avoir $\mathbb{P}(T>m+n) = \mathbb{P}(T>m)\,\mathbb{P}(T>n)$.
$$\Rightarrow \mathbb{P}(T>n) = \mathbb{P}(T>1+\dots+1) = \mathbb{P}(T>1)^n = q^n \text{ avec } q = \mathbb{P}(T>1)$$

Bibliographie : N. Boccara ; Probabilités, ch 2 p 31-48
G. Demengel et al. : Probabilités, statistique inférentielle, fiabilité, ch 4, p 108-126.
J.J. Droesbeke : éléments de statistique, ch 5 p 183-230
S.M. Ross : Initiation aux probabilités, ch 4 p 103-131