Fonctions génératrices des probabilités, fonctions génératrices des moments

I. Notion de fonction génératrice des probabilités

I.1. Cas des v.a.r. discrètes finies et entières

I.1.a) Définition

Définition

Soit X une v.a.r. discrète finie.

On suppose que :

× X est une v.a.r. à valeurs positives ($\forall x \in X(\omega),\ x \geqslant 0$).

× X est une v.a.r. à valeurs entières ($X(\Omega) \subset \mathbb{N}$).

Notons $X(\Omega) = \{x_1, \dots, x_r\}$ (avec $r \in \mathbb{N}^*$).

- On appelle **fonction génératrice des probabilités de** X et on note G_X la fonction $G_X : \mathbb{R} \to \mathbb{R}$ définie par :
$$\forall s \in \mathbb{R},\ G_X(s) = \mathbb{E}\big(s^X\big) = \sum_{k=1}^{r} s^{x_k}\ \mathbb{P}\big(\,[X = x_k]\,\big)$$
- La v.a.r. X étant à valeur entière, toutes les puissances apparaissant dans l'expression de la fonction G_X sont des puissances entières.
Ainsi, G_X est une fonction polynomiale de degré $m = \max\big(x_1, \dots, x_r\big)$.
En particulier, G_X est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$.
Pour faire mieux apparaître de résultat, on peut écrire G_X sous la forme :
$$\forall s \in \mathbb{R},\ G_X(s) = \sum_{j=0}^{m} s^j\ \mathbb{P}\big(\,[X = j]\,\big)$$
(on note que pour tout $j \notin X(\Omega)$, $\mathbb{P}\big(\,[X = j]\,\big) = 0$)

Démonstration.
Soit $s \in \mathbb{R}$.
Notons $g_s : x \mapsto s^x$, de sorte que : $s^X = g_s(X)$.
Par hypothèse, la v.a.r. X est finie. On en déduit, d'après le théorème de transfert, que la v.a.r. s^X admet une espérance donnée par :
$$\forall s \in \mathbb{R},\ G_X(s) = \mathbb{E}\big(s^X\big) = \sum_{k=1}^{n} s^{x_k}\ \mathbb{P}\big(\,[X = x_k]\,\big) \qquad \square$$

I.1.b) Illustration pour les lois usuelles

- Dans le cas où $X \hookrightarrow \mathcal{B}(p)$ (avec $p \in \,]0,1[$)

  On a : $X(\Omega) = \{0, 1\}$, ce qui correspond au cas $m = 2$, $x_1 = 0$ et $x_2 = 1$.
  Ainsi, l'expression de G_X est donnée par, pour tout $s \in \mathbb{R}$: :
$$\begin{array}{rcl} G_X(s) & = & s^0\ \mathbb{P}\big(\,[X = 0]\,\big) + s^1\ \mathbb{P}\big(\,[X = 1]\,\big) \\ & = & (1-p) + p\, s \end{array}$$

- Dans le cas où $X \hookrightarrow \mathcal{B}(n, p)$ (avec $n \in \mathbb{N}^*$ et $p \in \,]0,1[$)

  On a : $X(\Omega) = [\![0, n]\!]$, ce qui correspond au cas $m = n+1$, $x_1 = 0$, $\dots$, $x_{n+1} = n$. Ainsi, l'expression de G_X est donnée par, pour tout $s \in \mathbb{R}$:
$$\begin{array}{rcl} G_X(s) & = & \displaystyle\sum_{k=0}^{n} s^k\ \mathbb{P}\big(\,[X = k]\,\big) \\ & = & \displaystyle\sum_{k=0}^{n} \binom{n}{k} p^k\ (1-p)^{n-k}\ s^k \\ & = & \displaystyle\sum_{k=0}^{n} \binom{n}{k} \big(p\, s\big)^k\ (1-p)^{n-k} \\ & = & \Big((1-p) + p\, s\Big)^n \end{array}$$

  Évidemment, lorsque $n = 1$, on retrouve l'expression de la fonction génératrice des probabilités d'une v.a.r. X qui suit la loi $\mathcal{B}(1, p)$ $(= \mathcal{B}(p))$.

I.1.c) Résultat principal

Théorème 1.

Soit X une v.a.r. discrète finie.

On suppose que :

$\times$ *X est une v.a.r. à valeurs positives ($\forall x \in X(\omega),\ x \geqslant 0$).*

$\times$ *X est une v.a.r. à valeurs entières ($X(\Omega) \subset \mathbb{N}$).*

- *La fonction génératrice des probabilités de X caractérise la loi de X.*
- *Plus précisément :*

  1) $\boxed{\forall k \in \mathbb{N},\ \mathbb{P}\big([X = k]\big) \ = \ \dfrac{G_X^{(k)}(0)}{k!}}$

  où $G_X^{(k)}$ est la dérivée $k^{\text{ème}}$ de la fonction G_X.
  (la fonction G_X permet donc d'obtenir la loi de X)

  2) *si Y est une v.a.r. discrète finie, à valeurs positives et entières, on a :*

$$\boxed{\text{Les v.a.r. } X \text{ et } Y \text{ ont même loi} \ \Leftrightarrow \ G_X \ = \ G_Y}$$

Démonstration.

1) • Notons m la plus grande valeur entière prise par X. On a alors :

$$\forall s \in \mathbb{R},\ G_X(s) = \sum_{j=0}^{m} s^j\ \mathbb{P}\big([X = j]\big)$$

La fonction G_X est polynomiale de degré m.
En particulier, elle est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$.

• Démontrons par récurrence : $\forall k \in \mathbb{N},\ \mathcal{P}(k)$
où $\mathcal{P}(k) : \forall s \in \mathbb{R},\ G_X^{(k)}(s) = \sum\limits_{j=k}^{m} \big(j\,(j-1)\dots(j-k+1)\big)\ s^{j-k}\ \mathbb{P}\big([X = j]\big)$.

Remarque

- Remarquons tout d'abord que si $k > m$, la sommation s'effectue sur un ensemble d'indices vide. La somme considérée est donc nulle.
- On s'est permis d'utiliser le symbole « ... » pour décrire les coefficients de cette somme. Plus rigoureusement, on devrait écrire :

$$\Big(j\,(j-1)\dots(j-k+1)\Big) \ = \ \prod_{i=0}^{k-1} (j-i)$$

  En particulier, lorsque $k = 0$, le produit s'effectue sur un ensemble d'indices vide. Un tel produit vaut 1.

► **Initialisation** :
Soit $s \in \mathbb{R}$. Par définition : $G_X^{(0)}(s) = G_X(s) = \sum\limits_{j=0}^{m} 1 \times s^{j-0}\ \mathbb{P}\big([X = j]\big)$.
D'où $\mathcal{P}(0)$.

► **Hérédité** : soit $k \in \mathbb{N}$.
Supposons $\mathcal{P}(k)$ et montrons que $\mathcal{P}(k+1)$

$$\left(\textit{i.e. } \ \forall s \in \mathbb{R},\ G_X^{(k+1)}(s) = \sum_{j=k+1}^{m} \big(j\,(j-1)\dots(j-k)\big)\ s^{j-(k+1)}\ \mathbb{P}\big([X = j]\big) \right)$$

Soit $s \in \mathbb{R}$.
Rappelons tout d'abord que, par hypothèse de récurrence :

$$\begin{aligned} & G_X^{(k)}(s) \\ = \ & \sum_{j=k}^{m} \big(j\,(j-1)\dots(j-k+1)\big)\ s^{j-k}\ \mathbb{P}\big([X = j]\big) \\ = \ & k!\ \mathbb{P}\big([X = k]\big) + \sum_{j=k+1}^{m} \big(j\,(j-1)\dots(j-k+1)\big)\ s^{j-k}\ \mathbb{P}\big([X = j]\big) \end{aligned}$$

(le premier terme de la somme définissant $G_X^{(k)}$ est constant)

On a alors :

$$\begin{aligned} G_X^{(k+1)}(s) &= \left(G_X^{(k)}\right)'(s) \\ &= \sum_{j=k+1}^{m} \big(j\,(j-1)\dots(j-k+1)\,(j-k)\big)\ s^{j-k-1}\ \mathbb{P}\big([X=j]\big) \end{aligned}$$

D'où $\mathcal{P}(k+1)$.

Par principe de récurrence : $\forall k \in \mathbb{N}$, $\mathcal{P}(k)$.

- D'après ce qui précède, pour tout $s \in \mathbb{R}$:

$$\begin{aligned} & G_X^{(k)}(s) \\ = \ & k!\ \mathbb{P}\big([X=k]\big) + \sum_{j=k+1}^{m} \big(j\,(j-1)\dots(j-k+1)\big)\ s^{j-k}\ \mathbb{P}\big([X=j]\big) \\ = \ & k!\ \mathbb{P}\big([X=k]\big) + \Big(\sum_{j=k+1}^{m} \big(j\,(j-1)\dots(j-k+1)\big)\ s^{j-k-1}\ \mathbb{P}\big([X=j]\big)\Big)\ s \end{aligned}$$

et en particulier, en prenant $s=0$:

$$G_X^{(k)}(0) = k!\ \mathbb{P}\big([X=k]\big)$$

2) On procède par double équivalence.

($\Rightarrow$) Si X et Y ont même loi, alors, pour tout $j \in \mathbb{N}$:

$$\mathbb{P}\big([X=j]\big) = \mathbb{P}\big([Y=j]\big)$$

On en déduit, avec les notations précédentes :

$$G_X(s) = \sum_{j=0}^{m} s^j\ \mathbb{P}\big([X=j]\big) = \sum_{j=0}^{m} s^j\ \mathbb{P}\big([Y=j]\big) = G_Y(s)$$

($\Leftarrow$) Si $G_X = G_Y$ alors, d'après ***1)***, on a, pour tout $k \in \mathbb{N}$:

$$\mathbb{P}\big([X=k]\big) = \frac{G_X^{(k)}(0)}{k!} = \frac{G_Y^{(k)}(0)}{k!} = \mathbb{P}\big([Y=k]\big)$$

□

I.2. Cas des v.a.r. discrètes entières à ensemble image infini

I.2.a) Notion de série entière

Définition

Soit (a_n) une suite de réels.

Soit $x \in \mathbb{R}$.

- La série $\sum a_n\, x^n$ est appellée une série entière en la variable $x \in \mathbb{R}$.
- Si cette série est convergente, on note $\sum\limits_{j=0}^{+\infty} a_j\, x^j$ sa somme.
- Il est possible de démontrer que toute série entière est convergente sur un ouvert $]-R,R[$ où $R \in \mathbb{R} \cup \{+\infty\}$ est appelé **rayon de convergence** de cette série. On note par la suite f la fonction somme associée à la série $\sum a_n\, x^n$. On a alors :

$$\forall x \in]-R,R[,\ f(x) = \sum_{j=0}^{+\infty} a_j x^j$$

- On appelle série dérivée de la série $\sum a_n\, x^n$ la série entière $\sum n \geqslant 1 n\, a_n\, x^{n-1}$ (qu'on peut aussi noter $\sum\ (n+1)\, a_{n+1}\, x^n$).
La série dérivée admet le même rayon de convergence que la série initiale. De plus, sa somme n'est autre que la dérivée de la somme de la série initiale. Autrement dit :

$$\forall x \in\ ]-R,R[,\ f'(x) = \sum_{j=1}^{+\infty} j\ a_j\, x^{j-1} = \sum_{j=0}^{+\infty} (j+1)\, a_{j+1}\, x^j$$

Par une récurrence immédiate, on démontre que la fonction f est de classe $\mathcal{C}^\infty$ sur $]-R,R[$ et :

$$\forall j \in \mathbb{N},\ f^{(j)}(x) = \sum_{j=k}^{+\infty} \big(j\,(j-1)\dots(j-k+1)\big) s^{j-k}\ \mathbb{P}\big([X=j]\big)$$

Remarque

- L'écriture $\sum\limits_{k=0}^{+\infty} a_k x^k$ peut être vue comme une généralisation de la notion de polynômes : c'est un polynôme de degré ∞.
- Cette définition / théorème n'est pas à savoir démontrer. On la présente ici pour des aspects culturels. Les exercices ne demanderont pas de démontrer ces résultats dans le cas général. Il est par contre possible que l'on ait à démontrer des résultats similaires dans des cas particuliers (v.a.r. qui suivent des lois usuels notamment).
- La définition / théorème précédent démontre que l'on peut dériver sous le symbole somme infinie (dans l'intervalle $]-R, R[$ défini par le rayon de convergence R de la série) les sommes de séries entières. Encore une fois, ce n'est pas un résultat à connaître et il ne faudra jamais supposer que l'on peut effectuer ce genre de manipulations.
- Cette définition / théorème étant établi, les résultat principal sur les fonctions génératrices est aussi vérifié pour les v.a.r. discrètes à valeurs entières et à ensemble image infini.

I.2.b) Résultat principal

Théorème 2.

Soit X une v.a.r. discrète.

On suppose que :

× *X est une v.a.r. à valeurs positives ($\forall x \in X(\omega),\ x \geqslant 0$).*

× *X est une v.a.r. à valeurs entières ($X(\Omega) \subset \mathbb{N}$).*

- *La fonction génératrice des probabilités de X caractérise la loi de X.*

1) $$\boxed{\forall k \in \mathbb{N},\ \mathbb{P}\big([X = k]\big) \ = \ \dfrac{G_X^{(k)}(0)}{k!}}$$

2) *si Y est une v.a.r. discrète, à valeurs positives et entières, on a :*

$$\boxed{\text{Les v.a.r. } X \text{ et } Y \text{ ont même loi} \quad \Leftrightarrow \quad G_X \ = \ G_Y}$$

I.2.c) Illustration pour les lois usuelles

- Dans le cas où $X \hookrightarrow \mathcal{G}(p)$ (avec $p \in \,]0,1[$)

  L'expression de G_X est donnée par, pour tout $s \in \mathbb{R}$: :

$$\begin{aligned} G_X(s) &= \sum_{k=1}^{+\infty} s^k\,(1-p)\,p^{k-1} \\ &= s(1-p)\sum_{k=0}^{+\infty}(sp)^k \\ &= \frac{s(1-p)}{1-sp} \end{aligned}$$

- Dans le cas où $X \hookrightarrow \mathcal{P}(\lambda)$ (avec $\lambda > 0$)

  L'expression de G_X est donnée par, pour tout $s \in \mathbb{R}$: :

$$\begin{aligned} G_X(s) &= \sum_{k=0}^{+\infty} s^k\,\frac{\lambda^k}{k!}\,\mathrm{e}^{-\lambda} \\ &= \mathrm{e}^{-\lambda}\sum_{k=0}^{+\infty}\frac{(s\lambda)^k}{k!} \\ &= \mathrm{e}^{-\lambda}\,\mathrm{e}^{\lambda s} \ = \ \mathrm{e}^{\lambda(s-1)} \end{aligned}$$

I.2.d) Un résultat utile

Théorème 3.

Soit $n \in \mathbb{N}^*$.

Soient $X_1, \ldots, X_n$ des v.a.r. discrètes.

On suppose que les v.a.r. $X_1, \ldots, X_n$ sont indépendantes.

On suppose que pour tout $i \in [\![1, n]\!]$:

- *X_i est une v.a.r. à valeurs positives ($\forall x \in X_i(\omega),\ x \geqslant 0$).*
- *X_i est une v.a.r. à valeurs entières ($X_i(\Omega) \subset \mathbb{N}$).*

$$\boxed{G_{X_1+\ldots+X_n} \ = \ \prod_{i=1}^{n} G_{X_i}}$$

(la fonction génératrice d'une somme de v.a.r. c'est le produit des fonctions génératrices)

Corollaire 1.

Soit $n \in \mathbb{N}^*$ et soit $p \in \,]0, 1[$.

Soient $X_1, \ldots, X_n$ des v.a.r. discrètes.

On suppose que les v.a.r. $X_1, \ldots, X_n$ sont indépendantes.

On suppose enfin que pour tout $i \in [\![1, n]\!]$: $X_i \hookrightarrow \mathcal{B}(p)$.

Alors : $X_1 + \ldots + X_n \ \hookrightarrow \ \mathcal{B}(n, p)$.

Démonstration.
À démontrer ! □

Corollaire 2.

Soit $\lambda > 0$ et soit $\mu > 0$.

Soient X et Y deux v.a.r. indépendantes.

On suppose : $X \hookrightarrow \mathcal{P}(\lambda)$ et $Y \hookrightarrow \mathcal{P}(\lambda)$.

Alors : $X + Y \ \hookrightarrow \ \mathcal{P}(\lambda + \mu)$.

Démonstration.
À démontrer ! □