Grand classique
Fonction génératrice

On prendra comme convention : $0^0 = 1$.
$(\Omega, \mathcal{A}, \mathbb{P})$ désigne un espace probabilisé, X et Y sont deux variables aléatoires définies sur cet espace.

I Cas discret fini

Définition 1 — Fonction génératrice

Soit $n \in \mathbb{N}^*$. Supposons que $X(\Omega) \subset [\![0; n]\!]$.
La **fonction génératrice des probabilités de X**, notée G_X, est la fonction définie sur $\mathbb{R}$ par :

$$\forall t \in \mathbb{R},\ G_X(t) = \sum_{i=0}^{n} t^i \mathbb{P}\big([X = i]\big)$$

Petite remarque : Puisque la somme est finie, la fonction G_X est bien définie sur $\mathbb{R}$.

Propriétés 1

Avec les notations précédentes :

P1# La fonction G_X est $\mathscr{C}^\infty$ sur $\mathbb{R}$.

P2# $\forall t \in \mathbb{R}, G_X(t) = \mathbb{E}(t^X)$

P3#

$$G_X(1) = 1 \quad ; \quad G_X'(1) = \mathbb{E}(X) \quad ; \quad G_X''(1) = \mathbb{E}\big(X(X-1)\big)$$

P4# Si X et Y sont indépendantes, alors : $G_{X+Y} = G_X \times G_Y$.

P5# Pour tout $k \in \mathbb{N}$, $\mathbb{P}\big([X = k]\big) = \dfrac{G_X^{(k)}(0)}{k!}$

Petite remarque : Ainsi, d'après la formule de Koenig-Huygens :

$$\mathbb{V}(X) = G_X''(1) + G_X'(1) - \big(G_X'(1)\big)^2$$

Important ! La fonction génératrice caractérise donc la loi...

Démonstration :

P1# On remarque que, pour tout $t \in \mathbb{R}$, $G_X(t)$ est une expression polynomiale en t (de degré au plus n).
La fonction G_X est une fonction polynomiale, elle est donc $\mathscr{C}^\infty$ sur $\mathbb{R}$.

P2# Soit $t \in \mathbb{R}$. Puisque $X(\Omega)$ est fini, d'après le théorème de transfert, la variable aléatoire t^X admet une espérance et :

$$\begin{aligned}\mathbb{E}(t^X) &= \sum_{x \in X(\Omega)} t^x \mathbb{P}\big([X = x]\big) \\ &= \sum_{i=0}^{n} t^i \mathbb{P}\big([X = i]\big)\end{aligned}$$

$X(\Omega) \subset [\![0; n]\!]$ et $\mathbb{P}\big([X = i]\big) = 0$ si $i \notin X(\Omega)$

P3#
- Déjà :

$$\begin{aligned}G(1) &= \sum_{i=0}^{n} \mathbb{P}\big([X = i]\big) \\ &= 1\end{aligned}$$

$X(\Omega) \subset [\![0; n]\!]$

Petite remarque : On pouvait aussi procéder ainsi :

$$G_X(1) = \mathbb{E}(1^X) = \mathbb{E}(1) = 1$$

- On a, pour tout $t \in \mathbb{R}$:

$$\begin{aligned}G_X(t) &= \sum_{i=0}^{n} t^i \mathbb{P}\big([X = i]\big) \\ &= \mathbb{P}\big([X = 0]\big) + \sum_{i=1}^{n} t^i \mathbb{P}\big([X = i]\big)\end{aligned}$$

$0^0 = 1$

✘ Attention ! On pense à isoler le terme constant avant de dériver... Sinon, on arrive à écrire des choses confuses.

Or, G_X est $\mathscr{C}^\infty$ sur $\mathbb{R}$... Par linéarité de la dérivation, on obtient :

$$\forall t \in \mathbb{R},\ G_X'(t) = \sum_{i=1}^{n} i t^{i-1} \mathbb{P}\big([X = i]\big)$$

Par conséquent :

$$\begin{aligned}G_X'(1) &= \sum_{i=1}^{n} i \mathbb{P}\big([X = i]\big) \\ &= \sum_{i=0}^{n} i \mathbb{P}\big([X = i]\big) \\ &= \mathbb{E}(X)\end{aligned}$$

$X(\Omega) \subset [\![0; n]\!]$

- On sait déjà que G_X est $\mathscr{C}^\infty$ sur $\mathbb{R}$... Par linéarité de la dérivation, on obtient, en dérivant G'_X précédemment obtenue :

$$\forall t \in \mathbb{R},\ G''_X(t) = \sum_{i=2}^{n} i(i-1)t^{i-2}\mathbb{P}\big([X=i]\big)$$

Par conséquent :

$$\begin{aligned} G''_X(1) &= \sum_{i=2}^{n} i(i-1)\mathbb{P}\big([X=i]\big) \\ &= \sum_{i=0}^{n} i(i-1)\mathbb{P}\big([X=i]\big) && \text{théorème de transfert, car } X(\Omega)\subset [\![0;n]\!] \\ &= \mathbb{E}\big(X(X-1)\big) \end{aligned}$$

✘ Attention ! $G'_X(t) = \mathbb{P}\big([X=1]\big) + \sum_{i=2}^{n} t^i\mathbb{P}\big([X=i]\big)$, d'où l'expression de la dérivée seconde...

P4# Supposons que X et Y sont indépendantes. Montrons : $\forall t \in \mathbb{R},\ G_{X+Y}(t) = G_X(t) \times G_Y(t)$.
Soit $t \in \mathbb{R}$. Puisque X et Y sont indépendantes, d'après le lemme des coalitions, les variables aléatoires t^X et t^Y sont également indépendantes. Ainsi : $\mathbb{E}\big(t^X \times t^Y\big) = \mathbb{E}(t^X) \times \mathbb{E}(t^Y)$. Autrement dit :

$$G_{X+Y}(t) = G_X(t) \times G_Y(t)$$

Conclusion : X et Y sont indépendantes, alors $G_{X+Y} = G_X \times G_Y$.

P5#
- Remarquons déjà que :
  - ⋆ puisque $X(\Omega) \subset [\![0;n]\!]$, on a : pour tout $k \in [\![n+1;+\infty[\![$, $\mathbb{P}\big([X=k]\big) = 0$;
  - ⋆ puisque G_X est polynomiale de degré au plus n, on a : pour tout $k \in [\![n+1;+\infty[\![$, $\forall t \in \mathbb{R},\ G_X^{(k)}(t) = 0$.
    En particulier : $\forall k \in [\![n+1;+\infty[\![,\ G_X^{(k)}(0) = 0$

  Par conséquent : :

$$\forall k \in [\![n+1;+\infty[\![,\ \mathbb{P}\big([X=k]\big) = G_X^{(k)}(0)$$

- Montrons : $\forall k \in [\![0;n]\!],\ \forall t \in \mathbb{R},\ G_X^{(k)}(t) = \sum_{i=k}^{n} \frac{i!}{(i-k)!} t^{i-k}\mathbb{P}\big([X=i]\big)$. Procédons par récurrence finie...

Important ! $\frac{i!}{(i-k)!} = i(i-1)\dots(i-k+1)$

  - ⋆ **Initialisation.** Pour $k=0$:

$$\begin{aligned} G_X^{(0)}(0) &= G_X(0) \\ &= \sum_{i=0}^{n} 0^i\mathbb{P}\big([X=i]\big) \\ &= \mathbb{P}\big([X=0]\big) + \sum_{i=1}^{n} 0^i\mathbb{P}\big([X=i]\big) && \forall i \geqslant 1,\ 0^i = 0 \\ &= \mathbb{P}\big([X=0]\big) \end{aligned}$$

✘ Attention ! $0^0 = 1$

  L'initialisation est ainsi vérifiée.

  - ⋆ **Hérédité.** Soit $k \in [\![0;n-1]\!]$. Supposons : $\forall t \in \mathbb{R},\ G_X^{(k)}(t) = \sum_{i=k}^{n} \frac{i!}{(i-k)!} t^{i-k}\mathbb{P}\big([X=i]\big)$.

  Montrons : $\forall t \in \mathbb{R},\ G_X^{(k+1)}(t) = \sum_{i=k+1}^{n} \frac{i!}{(i-k-1)!} t^{i-k-1}\mathbb{P}\big([X=i]\big)$.

  Par hypothèse de récurrence, pour tout $t \in \mathbb{R}$:

$$\begin{aligned} G_X^{(k)}(t) &= \sum_{i=k}^{n} \frac{i!}{(i-k)!} t^{i-k}\mathbb{P}\big([X=i]\big) && 0^0 = 1 \\ &= k!\mathbb{P}\big([X=k]\big) + \sum_{i=k+1}^{n} \frac{i!}{(i-k)!} t^{i-k}\mathbb{P}\big([X=i]\big) \end{aligned}$$

  D'où, G_X étant $\mathscr{C}^\infty$ sur $\mathbb{R}$, en dérivant, on obtient par linéarité de la dérivation, pour tout $t \in \mathbb{R}$:

$$\begin{aligned} G_X^{(k+1)}(t) &= \big(G_X^{(k)}\big)'(t) \\ &= \sum_{i=k+1}^{n} \frac{i!}{(i-k)!}(i-k)t^{i-k-1}\mathbb{P}\big([X=i]\big) \\ &= \sum_{i=k+1}^{n} \frac{i!}{(i-k-1)!} t^{i-k-1}\mathbb{P}\big([X=i]\big) \end{aligned}$$

Pourquoi ? $(i-k)! = (i-k) \times (i-k-1)!$

  L'hérédité est ainsi établie.

Conclusion : $\forall k \in [\![0;n]\!],\ \forall t \in \mathbb{R},\ G_X^{(k)}(t) = \sum_{i=k}^{n} \frac{i!}{(i-k)!} t^{i-k}\mathbb{P}\big([X=i]\big)$.

En particulier, pour tout $k \in [\![0;n]\!]$:

$$\begin{aligned} G_X^{(k)}(0) &= \sum_{i=k}^{n} \frac{i!}{(i-k)!} 0^{i-k}\mathbb{P}\big([X=i]\big) && 0^0 = 1\text{, et la somme est nulle si } k = n \\ &= \frac{k!}{0!}\mathbb{P}\big([X=k]\big) + \sum_{i=k+1}^{n} \frac{i!}{(i-k)!} 0^{i-k}\mathbb{P}\big([X=i]\big) && \forall i \geqslant k+1,\ 0^{i-k} = 0 \\ &= k!\mathbb{P}\big([X=k]\big) \end{aligned}$$

Pourquoi ? Si $k = n$, la somme $\sum_{i=k+1}^{n} [\dots]$ est indexée sur un ensemble vide, elle est donc nulle par convention.

Conclusion : $\forall k \in \mathbb{N},\ \mathbb{P}\big([X=k]\big) = \frac{G_X^{(k)}(0)}{k!}$.

⋆

II Cas discret infini

Propriété 2

Supposons que $X(\Omega) \subset \mathbb{N}$.
Pour tout $t \in [-1;1]$, la série $\sum\limits_{i \geqslant 0} t^i \mathbb{P}\big([X=i]\big)$ est convergente.

♡ **Astuce du chef !** ♡
La série $\sum\limits_{i \geqslant 0} t^i \mathbb{P}\big([X=i]\big)$ n'est pas à terme général positif quand $t \leqslant 0$, on pense donc à étudier sa convergence absolue...

⋆ Démonstration : Soit $t \in [-1;1]$.
On a ainsi :

$$0 \leqslant |t| \leqslant 1$$

Puis, pour tout $i \in \mathbb{N}$, par croissance de la fonction $x \longmapsto x^i$ sur $\mathbb{R}^+$:

$$0 \leqslant |t|^i \leqslant 1$$

Et ainsi, une probabilité étant un réel positif :

$$\forall i \in \mathbb{N},\ 0 \leqslant |t|^i \mathbb{P}\big([X=i]\big) \leqslant \mathbb{P}\big([X=i]\big)$$

Autrement dit :

$$\forall i \in \mathbb{N},\ 0 \leqslant \big|t^i \mathbb{P}\big([X=i]\big)\big| \leqslant \mathbb{P}\big([X=i]\big)$$

Or $X(\Omega) \subset \mathbb{N}$, donc la série $\sum\limits_{i \geqslant 0} \mathbb{P}\big([X=i]\big)$ est convergente (de somme égale à 1).
Ainsi, d'après le critère de comparaison sur les séries à terme général positif, la série $\sum\limits_{i \geqslant 0} \big|t^i \mathbb{P}\big([X=i]\big)\big|$ est convergente.
Conclusion : pour tout $t \in [-1;1]$, la série $\sum\limits_{i \geqslant 0} t^i \mathbb{P}\big([X=i]\big)$ est absolument convergente, donc convergente.

⋆

Définition 2 — **Fonction génératrice**

Supposons que $X(\Omega) \subset \mathbb{N}$.
La **fonction génératrice des probabilités de X**, notée G_X, est la fonction définie sur $[-1;1]$ par :

$$\forall t \in [-1;1],\ G_X(t) = \sum_{i=0}^{+\infty} t^i \mathbb{P}\big([X=i]\big)$$

☞ **Pour info...**
Il se peut que G_X soit définie sur un intervalle plus grand que $[-1;1]$, même si $X(\Omega) = \mathbb{N}$.

On admet ensuite que les propriétés vues dans le cas discret se généralisent. Pour démontrer ces propriétés, il faudrait que l'on soit capables de dériver sous le signe $\sum\limits_{i=0}^{+\infty}$, ce qui n'est pas le cas en ECG...

Propriétés 3

Avec les notations précédentes :

P1# La fonction G_X est $\mathscr{C}^\infty$ sur $\mathbb{R}$

P2# $\forall t \in \mathbb{R}, G_X(t) = \mathbb{E}(t^X)$

P3#

$$G_X(1) = 1 \quad ; \quad G_X'(1) = \mathbb{E}(X) \quad ; \quad G_X''(1) = \mathbb{E}\big(X(X-1)\big)$$

P4# Si X et Y sont indépendantes, alors : $G_{X+Y} = G_X \times G_Y$.

P5# Pour tout $k \in \mathbb{N}$, $\mathbb{P}\big([X=k]\big) = \dfrac{G_X^{(k)}(0)}{k!}$

Important !
La fonction génératrice caractérise donc la loi...

Quelques cas particuliers...

- Soit $X \hookrightarrow \mathscr{B}(p)$. Dans ce cas : $X(\Omega) = \{0;1\}$ et, pour tout $t \in \mathbb{R}$:

$$\begin{aligned} G_X(t) &= t^0 \mathbb{P}\big([X=0]\big) + t^1 \mathbb{P}\big([X=1]\big) \\ &= 1 - p + tp \end{aligned}$$

- Soit $X \hookrightarrow \mathscr{B}(n;p)$. Dans ce cas : $X(\Omega) = [\![0;n]\!]$ et, pour tout $t \in \mathbb{R}$:

$$\begin{aligned} G_X(t) &= \sum_{i=0}^{n} t^i \mathbb{P}\big([X=i]\big) \\ &= \sum_{i=0}^{n} t^i \binom{n}{i} p^i (1-p)^{n-i} \\ &= \sum_{i=0}^{n} \binom{n}{i} (pt)^i (1-p)^{n-i} \\ &= \big(1-p+pt\big)^n \end{aligned}$$

formule du binôme de Newton

Petite remarque
On peut retrouver celle-ci en utilisant le fait qu'une VA suivant une $\mathscr{B}(n;p)$ est une somme de VA indépendantes suivant une $\mathscr{B}(p)$ et en utilisant P4.

- Soit $X \hookrightarrow \mathscr{G}(p)$. Dans ce cas : $X(\Omega) = \mathbb{N}^*$ et, pour tout $t \in [-1;1]$:

$$\begin{aligned} G_X(t) &= \sum_{i=1}^{+\infty} t^i (1-p)^{i-1} p \\ &= \sum_{k=0}^{+\infty} t^{k+1} (1-p)^k p \\ &= pt \sum_{k=0}^{+\infty} \big((1-p)t\big)^k \\ &= \frac{pt}{1-(1-p)t} \end{aligned}$$

changement d'indice $k i-1$

linéarité de la somme

Petite remarque

Valable sur $\left]\frac{-1}{1-p}; \frac{1}{1-p}\right[$, car il faut et il suffit que $(1-p)t \in]-1;1[$ pour assurer la CV de $\sum_{k\geqslant 0} \big((1-p)t\big)^k$ (série géométrique)...

- Soit $X \hookrightarrow \mathscr{P}(\lambda)$. Dans ce cas : $X(\Omega) = \mathbb{N}$ et, pour tout $t \in [-1;1]$:

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

linéarité de la somme

Petite remarque

Valable sur $\mathbb{R}$, car pour tout $t \in \mathbb{R}$, $\sum_{i\geqslant 0} \frac{(\lambda t)^i}{i!}$ est une série exponentielle convergente.

- Soient $X \hookrightarrow \mathscr{P}(\lambda)$ et $Y \hookrightarrow \mathscr{P}(\mu)$. Montrons que si X et Y sont indépendantes, alors $X+Y \hookrightarrow \mathscr{P}(\lambda+\mu)$.
  Supposons que X et Y sont indépendantes. D'après P4, on a alors :

$$G_{X+Y} = G_X G_Y$$

  Or $X \hookrightarrow \mathscr{P}(\lambda)$ et $Y \hookrightarrow \mathscr{P}(\mu)$, donc, pour tout $t \in \mathbb{R}$:

$$\begin{aligned} G_X(t) \times G_Y(t) &= \mathrm{e}^{\lambda(t-1)} \times \mathrm{e}^{\mu(t-1)} \\ &= \mathrm{e}^{(\lambda+\mu)(t-1)} \end{aligned}$$

  Ainsi :

$$\forall t \in \mathbb{R},\ G_{X+Y}(t) = \mathrm{e}^{(\lambda+\mu)(t-1)}$$

  On reconnaît la fonction génératrice d'une variable aléatoire suivant la loi $\mathscr{P}(\lambda+\mu)$.
  Puisque la fonction génératrice caractérise la loi, on en déduit que $X+Y \hookrightarrow \mathscr{P}(\lambda+\mu)$.