Chapitre 7

Fonctions génératrices

7.1 Fonction génératrice des probabilités d'une v.a. entière

Soit $X \rightsquigarrow \mathcal{B}(n,p)$ une v.a. binômiale ; on a $p_X(k) := \mathbb{P}\{X = k\} = \binom{n}{k} p^k(1-p)^{n-k}$. On observe que les nombres $p_X(k)$ sont précisément les coefficients du polynôme (de la variable s) suivant :

$$G_X(s) := ((1-p)+ps)^n = \sum_{k=0}^{n} \binom{n}{k} p^k(1-p)^{n-k}s^k = \sum_{k=0}^{n} p_X(k)s^k = \mathbb{E}(s^X).$$

Cette dernière expression, $\mathbb{E}(s^X)$, se prête à la généralisation :

Définition : Soit X une v.a. entière. On appelle *fonction génératrice des probabilités* (fgp) de X, et on note G_X, la fonction

$$G_X(s) := \mathbb{E}(s^X) = \sum_{k=0}^{\infty} p_X(k)s^k, \text{ toujours avec } p_X(k) := \mathbb{P}(\{X = k\}).$$

Remarques :
- Comme $\sum_{k=0}^{\infty} p_X(k) = 1$, le rayon de convergence R_G de la série entière $\sum_{k=0}^{\infty} p_X(k)s^k$ est au moins égal à 1.
- $\boxed{\text{Si } X \text{ et } Y \text{ ont même fgp, elles ont même loi.}}$ En effet

$$\mathbb{P}\{X = k\} = p_X(k) = \frac{G_X^{(k)}(0)}{k\ !} = \frac{G_Y^{(k)}(0)}{k\ !} = p_Y(k) = \mathbb{P}\{y = k\}.$$

Théorème 7.1 *Si $\mathbb{E}X^r$, le moment d'ordre r de la v.a. X, existe, la dérivé $G_X^{(r)}$ de la fgp de X vérifie*

$$G_X^{(r)}(1) = \mathbb{E}[X(X-1)\dots(X-r+1)], \tag{7.1}$$

où, lorsque $R_G = 1$, $G_X^{(r)}(1)$ désigne $\lim_{s\to1^-} G^{(r)}(s) = \lim_{s\to1^-} \sum_{k=0}^{\infty} \frac{d^r}{ds^r}(p(k)s^k)$.

Comme on sait seulement que $R_G \geq 1$, on est assuré de la dérivabilité de G_X que pour $|s| < 1$. Comme $G_X(1)$ est toutefois bien défini, le résultat suivant s'applique ; il servira dans la preuve du théorème 7.1.

Lemme 7.2 (Abel) *Si la série entière de $f(x) := \sum_{n=0}^{\infty} a_n x^n$ converge pour tout $|x| < 1$, et si $\sum_{n=0}^{\infty} a_n$ converge, alors $\lim_{x\to1^-} f(x) = \sum_{n=0}^{\infty} a_n = f(1)$.*

Preuve : Comme les dérivées terme à terme d'une série entière ont le même rayon de convergence que cette série entière, le rayon de convergence de la série de $G_X^{(r)}(s)$ pour $|s| < R_G$ est lui aussi, au moins égal à 1. Comme $\mathbb{E}X^r$ existe, il en est de même de $\mathbb{E}X^s$ pour $0 \leq s \leq r$, et donc $\mathbb{E}[X(X-1)\dots(X-r+1)]$ existe bien. Montrons le théorème pour $r = 1$; le cas général se montrerait de même par récurrence. Pour $|s| < 1$, on a, en dérivant terme à terme,

$$G'_X(s) = \sum_{k=0}^{\infty} \frac{d}{ds}(p(k)s^k) = \sum_{k=0}^{\infty} kp(k)s^{k-1},$$

d'où, par le lemme d'Abel 7.2 , $\lim_{s\to1^-} G'_X(s) = \sum_{k=0}^{\infty} kp(k) = \mathbb{E}X$. □

Exemples :

- *Fonction génératrice des probablités d'une v.a. de Poisson $X \rightsquigarrow \mathcal{P}(\lambda)$:*
  On a $\mathbb{P}\{X = k\} = e^{-\lambda}\frac{\lambda^k}{k\,!}$, donc

$$G_X(s) = \mathbb{E}s^X = \sum_{k=0}^{\infty} e^{-\lambda}\frac{\lambda^k}{k\,!}s^k = e^{-\lambda}\sum_{k=0}^{\infty}\frac{(\lambda s)^k}{k\,!} = e^{-\lambda}e^{\lambda s} = e^{\lambda(s-1)}.$$

  D'où $G_X'(s) = \lambda e^{-\lambda(s-1)}$, et donc $\mathbb{E}X = G_X'(1) = \lambda$. Plus généralement, $G_X^{(r)}(s) = \lambda^r e^{-\lambda(s-1)}$, et donc $\mathbb{E}[X(X-1)\dots(X-r+1)] = G_X^{(r)}(1) = \lambda^r$. En particulier $\lambda^2 = \mathbb{E}[X(X-1)] = \mathbb{E}X^2 - \mathbb{E}X$, donc $\mathbb{E}X^2 = \lambda^2 + \lambda$, d'où $\operatorname{Var} X = \mathbb{E}X^2 - (\mathbb{E}X)^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$.
- *Fonction génératrice des probablités d'une v.a. de loi géométrique $X \rightsquigarrow \mathcal{G}(\alpha)$:*
  On a $\mathbb{P}\{X = k\} = (1-\alpha)\alpha^{k-1}$, $k = 1, 2, \dots$, donc, comme $\sum_{k=1}^{\infty} q^k = \frac{q}{1-q}$,

$$G_X(s) = \mathbb{E}s^X = \sum_{k=1}^{\infty}(1-\alpha)\alpha^{k-1}s^k = \frac{1-\alpha}{\alpha}\sum_{k=1}^{\infty}(\alpha s)^k = (1-\alpha)\frac{s}{1-\alpha s},$$

  donc $G_X'(s) = \frac{1-\alpha}{(1-\alpha s)^2}$, et $\mathbb{E}X = G_X'(1) = \frac{1-\alpha}{(1-\alpha)^2} = \frac{1}{1-\alpha}$; par ailleurs

$$\mathbb{E}X^2 - \mathbb{E}X = G_X''(s) = (1-\alpha)\frac{+2\alpha}{(1-\alpha)^3} = \frac{2\alpha}{(1-\alpha)^2},$$

  d'où $\mathbb{E}X^2 = \frac{2\alpha}{(1-\alpha)^2} + \frac{1}{1-\alpha} = \frac{1+\alpha}{(1-\alpha)^2}$, et

$$\operatorname{Var} X = \mathbb{E}X^2 - (\mathbb{E}X)^2 = \frac{1+\alpha}{(1-\alpha)^2} - \frac{1}{(1-\alpha)^2} = \frac{\alpha}{(1-\alpha)^2}.$$

Proposition 7.3 *Si X et Y sont deux v.a. indépendantes, $G_{X+Y}(s) = G_X(s)G_Y(s)$.*

Preuve : Comme X et Y sont indépendantes, il en est de même pour les v.a. s^X et s^Y, d'où

$$G_{X+Y}(s) = \mathbb{E}[s^{X+Y}] = \mathbb{E}[s^X s^Y] = \mathbb{E}[s^X]\mathbb{E}[s^Y] = G_X(s)G_Y(s).$$

□

Exemples :

- *V.a. de Bernoulli :* on sait que la loi binômiale est la loi de la somme de n v.a de Bernoulli indépendantes : ceci explique que la fgp d'une v.a. binômiale, $G(s) = ((1-p)+ps)^n$, est la puissance n-ième de la fgp $(1-p) + ps$ d'une v.a. de Bernoulli.
- *Somme aléatoire de v.a. i.i.d. :* Soit N, X_1, …, X_k, …des v.a. entières indépendantes, les X_k étant identiquement distribuées. Notons respectivement G_N et G_X les fgp de N et des X_k. Soit $T := \sum_{k=1}^{N} X_k$ la somme aléatoire des X_k (un modèle possible pour le coût annuel des sinistres pour une compagnie d'assurance : N désigne le nombre de sinistres par an, et X_k représente le coût du k-ième). Alors $G_T = G_N \circ G_X$. En effet, $G_T(s) = \mathbb{E}(\mathbb{E}(s^T|N)) = \mathbb{E}g(N)$, avec

$$g(n) := \mathbb{E}(s^T|N=n) = \mathbb{E}\left(s^{\sum_{k=1}^n X_k}|N=n\right) \stackrel{\sharp}{=} \mathbb{E}\left(s^{\sum_{k=1}^n X_k}\right) \stackrel{\flat}{=} \prod_{k=1}^{n}\mathbb{E}(s^{X_k}) = (G_X(s))^n,$$

  où l'égalité $\sharp$ résulte de l'indépendance de $X_1, \dots, X_n$ de N, et l'égalité $\flat$ résulte de l'indépendance mutuelle des $X_1, \dots, X_n$. En posant $y := G_X(s)$, on a donc $G_T(s) = \mathbb{E}y^N = G_N(y) = G_N \circ G_X(s)$.

Théorème 7.4 (admis) *Soient $\overline{X}$, X_1, X_2, …une suite de v.a.. Si, pour tout $|s| \le 1$, $\lim_{n\to\infty} G_{X_n}(s) = G_{\overline{X}}(s)$, alors la suite des X_n* tend en loi *vers $\overline{X}$, c'est-à-dire que*

$$\lim_{n\to\infty}\mathbb{P}\{X_n = k\} = \mathbb{P}\{\overline{X} = k\} \text{ pour tout } k.$$

Dans l'exemple $Y_n \rightsquigarrow \mathcal{U}(1..n)$ (loi *discrète* uniforme sur $1..n$) on devine que si X est la limite des $X_n = \frac{1}{n}Y_n$, X suivra une loi *continue* uniforme $\mathcal{U}[0,1]$. Nous souhaitons donc une notion de fonction génératrice qui s'applique aussi au cas de v.a. absolument continues.

7.2 Fonction génératrice des moments (v.a. continue ou discrète)

Soit $X \rightsquigarrow \mathcal{E}(\lambda)$ une v.a. suivant une loi exponentielle. Son k-ième moment $\boxed{\mu_X(k) := \mathbb{E}X^k}$, se calcule facilement au moyen de k intégrations par parties et est donc égal à

$$\mu_X(k) = \mathbb{E}X^k = \int_0^\infty x^k \lambda e^{-\lambda x} dx = \frac{k\,!}{\lambda^k}.$$

On observe que les nombres $\frac{1}{k\,!}\mu_X(k)$ sont précisément les coefficients du développement en série entière de la fonction

$$\begin{aligned} M_X(t) \quad &:= \quad \frac{\lambda}{\lambda - t} = \frac{1}{1-(\frac{t}{\lambda})} = \sum_{k=0}^{\infty} \frac{t^k}{\lambda^k} = \sum_{k=0}^{\infty} \frac{\mu_X(k)}{k\,!} t^k \\ &= \quad \sum_{k=0}^{\infty} \frac{\mathbb{E}X^r}{k\,!} t^k = \sum_{k=0}^{\infty} \mathbb{E}\left[\frac{X^r}{k\,!} t^k\right] \overset{\natural}{=} \mathbb{E}\left[\sum_{k=0}^{\infty} \frac{X^r}{k\,!} t^k\right] = \mathbb{E}e^{tX}, \end{aligned}$$

où l'égalité surmontée d'un bécarre $\natural$ impose la circonspection usuelle vis-à-vis du caractère infini de la somme considérée. La dernière expression, $\mathbb{E}e^{tX}$, se prête à la généralisation :

Définition : On appelle *fonction génératrice des moments* de X la fonction M_X définie par

$$\boxed{M_X(t) := \mathbb{E}e^{tX}.}$$

Théorème 7.5 *Supposons que tous les moments $\mu_X(k) := \mathbb{E}X^k$ d'une v.a. X soient définis, et que la série $\sum_{k=0}^{\infty} \frac{\mu_X(k)}{k\,!} t^k$ converge, avec un rayon de convergence R non nul. Alors, pour $|t| < R$, $\mathbb{E}e^{tX}$ existe et est égal à $\sum_{k=0}^{\infty} \frac{\mathbb{E}X^r}{k\,!} t^k$, et*

$$M_X(t) = \mathbb{E}e^{tX} = \sum_{k=0}^{\infty} \frac{\mu_X(k)}{k\,!} t^k \;,\; (|t| < R).$$

Remarques :

- Nous admettons ce théorème et retenons sa conclusion sous la forme que $\mathbb{E}e^{tX} =: M_X(t) = \sum_{k=0}^{\infty} \frac{\mathbb{E}X^r}{k\,!} t^k$ pour $|t| < R$. Sa preuve est simple dès lors qu'on dispose de la théorie de Lebesgue de l'intégrale.
- Si X est à valeurs entières positives, en posant $s = e^t$, nous voyons que $\boxed{M_X(t) = G_X(e^t)}$, puisque

$$M_X(t) = \mathbb{E}e^{tX} = \mathbb{E}(e^t)^X = \mathbb{E}s^X = G_X(s) = G_X(e^t).$$

  Cette dernière expression est définie pour $|s| \leq 1$, donc $M_X(t)$ est définie pour $t \leq 0$ (au moins) ; les hypothèses sur la convergence de la série formée à partir des moments n'est utile que pour assurer une représentation de M_X par une série en puissances entières de t.
- Si X est absolument continue, de (fonction de) densité de probabilité (fdp) f_X, nous avons

$$M_X(t) = \mathbb{E}e^{tX} = \int_{-\infty}^{+\infty} e^{tx} f_X(x) dx.$$

  L'opération fonctionnelle associant à une fonction $(x \mapsto f(x))$ la fonction $(t \mapsto M(-t) = \int_{-\infty}^{+\infty} f(t)e^{-tx}dx)$ s'appelle la *transformation de Laplace (bilatère)* et est utilisée dans biens des branches des Mathématiques, allant de méthodes pour l'ingénieur aux théories les plus récentes de sommation des séries divergentes qui apparaissent notamment en physique théorique.
- Par la formule de Taylor nous savons que le coefficient de t^k dans la série entière de M_X n'est autre que la valeur de $\frac{M_X^{(k)}(0)}{k\,!}$, donc $\mu_X(k) = M^{(k)}(0)$, d'où

$$\boxed{\mathbb{E}X^k = M_X^{(k)}(0).} \tag{7.2}$$

Exemples :

- *Fonction génératrice des moments d'une v.a. uniforme* continue $X \rightsquigarrow \mathcal{U}[a,b]$:
  On a $f_X(x) = \frac{1}{b-a}\mathbb{I}_{[a,b]}$, d'où $\mu_X(k) = \int_a^b \frac{x^k}{b-a}dx = \frac{b^{k+1}-a^{k+1}}{b-a}$, une expression guère commode pour calculer $M_X(t)$ à partir de sa série entière. En fait
  $$M_X(t) = \mathbb{E}e^{tX} = \int e^{tx}\frac{1}{b-a}\mathbb{I}_{[a,b]}(x)dx = \frac{1}{b-a}\int_a^b e^{tx}dx = \frac{1}{b-a}\left[\frac{1}{t}e^{tx}\right]_{x=a}^{b} = \frac{e^{bt}-e^{at}}{(b-a)t}.$$
- *Fonction génératrice des moments d'une v.a. uniforme* discrete $X_n \rightsquigarrow \mathcal{U}]0..1]_{\frac{1}{n}} = \frac{1}{n}Y_n$ avec $Y_n \rightsquigarrow \mathcal{U}]1..n]$:
  On a $p_{X_n}(\frac{i}{n}) = \mathbb{P}\{X = \frac{i}{n}\} = \mathbb{P}\{Y = i\} = \frac{1}{n}$, pour $i = 1..n$, d'où
  $$M_{X_n}(t) = \mathbb{E}e^{tX} = \sum_{i=1}^{n}\frac{1}{n}e^{t\frac{i}{n}} = \frac{1}{n}\sum_{i=1}^{n}\left(e^{\frac{t}{n}}\right)^i = \frac{1}{n}e^{\frac{t}{n}}\frac{\left(e^{\frac{t}{n}}\right)^n - 1}{e^{\frac{t}{n}}-1} = \frac{e^t-1}{n(1-e^{-\frac{t}{n}})}.$$
- *Fonction génératrice des moments d'une v.a. exponentielle* $X \rightsquigarrow \mathcal{E}(\lambda)$:
  On a $f_X(x) = \lambda e^{-\lambda x}\mathbb{I}_{[0,+\infty[}(x)$. Là encore, inutile de calculer les moments. Pour $t < \lambda$ on a
  $$M_X(t) = \mathbb{E}e^{tX} = \int e^{tx}\lambda e^{-\lambda x}\mathbb{I}_{[0,+\infty[}(x) = \int_0^{+\infty}\lambda e^{(t-\lambda)x}dx = \frac{\lambda}{t-\lambda}\left[e^{(t-\lambda)x}\right]_{x=0}^{+\infty} = \frac{\lambda}{\lambda - t}.$$
- *Fonction génératrice des moments d'une v.a. normale* $X \rightsquigarrow \mathcal{N}(0,1)$:
  On a $f_X(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$, d'où
  $$M_X(t) = \mathbb{E}e^{tX} = \int_{-\infty}^{+\infty}e^{tx}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-\frac{x^2}{2}+tx}dx.$$
  C'est une intégrale gaussienne ; il convient de mettre l'exposant de l'intégrand sous la forme $-\frac{v^2}{2}+c$. On a
  $$-\frac{x^2}{2}+tx = -\frac{1}{2}(x^2-2tx) = -\frac{1}{2}\left[(x-t)^2-t^2\right] = -\frac{1}{2}(x-t)^2+\frac{t^2}{2} = -\frac{v^2}{2}+\frac{t^2}{2}$$
  pour $v = x-t$, et finalement
  $$M_X(t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-\frac{x^2}{2}+tx}dx = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-\frac{v^2}{2}+\frac{t^2}{2}}dv = e^{+\frac{t^2}{2}}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-\frac{v^2}{2}}dv = e^{+\frac{t^2}{2}}.$$
  On pourrait procéder de manière similaire pour $X \rightsquigarrow \mathcal{N}(\mu,\sigma)$ et obtenir
  $$M_X(t) = e^{\mu t}e^{\frac{1}{2}\sigma^2t^2}.$$
  Les calculs sont toutefois plus laborieux, et sont plus aisés avec le recul des propriétés de la transformation de Laplace. Notons qu'on trouve $M'_X(t) = (\mu+\sigma^2t)M_X(t)$ et $M''_X(t) = (\sigma^2+(\mu+\sigma^2t)^2)M_X(t)$, d'où $M'_X(0) = \mu$ et $M''_X(0) = \sigma^2+\mu^2$, d'où $\mathbb{E}X = M'_X(0) = \mu$, $\mathbb{E}X^2 = M''_X(0) = \sigma^2+\mu^2$, et $\operatorname{Var}X = \sigma^2+\mu^2-\mu^2 = \sigma^2$. On retrouve bien l'espérance et la variance d'une v.a. suivant une loi $\mathcal{N}(\mu,\sigma)$.

En raisonnant comme dans le cas de la fonction génératrice des probabilités on voit qu'on a

Proposition 7.6 *Si X et Y sont deux v.a. indépendantes, $M_{X+Y}(t) = M_X(t)M_Y(t)$.*

Théorème 7.7 (admis) *Soient $\overline{X}$, X_1, X_2, ... une suite de v.a.. Si, pour tout $|t| \leq R$, $\lim_{n\to\infty}M_{X_n}(t) = M_{\overline{X}}(t)$, alors la suite des X_n* tend en loi *vers $\overline{X}$, c'est-à-dire que*
$$\lim_{n\to\infty}M_{X_n}(t) = M_{\overline{X}}(t) \text{ pour tout } t \text{ implique que } \lim_{n\to\infty}F_{X_n}(x) = F_{\overline{X}}(x) \text{ pour tout } x,$$
où $F_{\overline{X}}(x) := \mathbb{P}\{\overline{X} \leq x\}$ et $F_{X_n}(x) := \mathbb{P}\{X_n \leq x\}$ désignent les fonctions caractérisques des v.a. considérées. En particulier pour deux v.a. X et Y, $\boxed{\textit{si on a } M_X = M_Y\textit{, alors } X \textit{ et } Y \textit{ ont même loi.}}$

Reprenons l'exemple $Y_n \rightsquigarrow \mathcal{U}[1..n]$ (loi *discrète* uniforme sur $1..n$). Nous avons vu que $M_{X_n}(t) = \frac{e^t-1}{n(1-e^{-\frac{t}{n}})}$. Un développement limité à l'ordre 1 du dénominateur $n(1-e^{-\frac{t}{n}}) = n(1-(1-\frac{t}{n}+\ldots))$ montre que $\lim_n M_{X_n}(t) = \frac{e^t-1}{t}$, la fonction génératrice des moments d'une v.a. continue uniforme sur $[0,1]$, ce qui montre que la suite $(Y_n)_{n\geq 1}$ converge en loi vers une loi $\mathcal{U}[0,1]$.