Pendant les vacances ? famille - repos
DM → 30 déc
révisions Sup.
planning

Fonctions génératrices

$X(\Omega) \subset \mathbb{N}$ pas ex : les lois usuelles.

Sauf mention contraire, $(\Omega, \mathscr{A}, P)$ est un espace probabilisé.
On s'intéresse dans ce chapitre uniquement aux variables aléatoires qui sont à valeurs dans $\mathbb{N}$. Typiquement, celles qui apparaissent dans des situations de comptage.

1 Définition

Lemme. Soit X une variable aléatoire à valeurs dans $\mathbb{N}$. La série entière :

$$\sum_{n \geqslant 0} P(X=n) t^n$$

converge normalement sur $[-1,1]$ (et même $DF(0,1)$ si on considère la variable complexe), et son rayon de convergence satisfait : $R_X \geqslant 1$.

Définition. Soit X une variable aléatoire à valeurs dans $\mathbb{N}$. On définit la **fonction génératrice** de X par :

$$G_X : t \mapsto \sum_{n \geqslant 0} P(X=n) t^n = \sum_{n \in X(\Omega) \subset \mathbb{N}} P(X=n) t^n$$

"série génératrice".

Remarque. *$G_X(1) = 1$ et $G_X(t) = \mathrm{E}(t^X)$ par la formule de transfert.*

Proposition. La loi d'une variable aléatoire à valeurs dans $\mathbb{N}$ est caractérisée par sa fonction génératrice.

Somme de SE

Si $X(\Omega) = \mathbb{N}$, la somme est complète

Si $X(\Omega) \subsetneq \mathbb{N}$ (par ex : $X(\Omega) = \{0, 1, \dots, n\}$)
la somme a des termes nuls

Preuve : $\forall t \in [-1,1]$, $|P(X=n)t^n| \leq P(X=n)$ indép de t

tg série cv (de somme 1)

$\sum f_n$ où $f_n(t) = P(X=n)t^n$ cv normalement sur $[-1,1]$

donc $R(\sum f_n) \geq 1$

G_X est continue sur $[-1, 1]$

Rmq : $G_X(1) = \sum_{n=0}^{+\infty} P(X=n)\, 1$
$= 1$

$$G_X(t) = \sum_{n \in X(\Omega)} t^n P(X=n)$$
$$= E(t^X) \quad \text{par transfert}$$

Fonctions génératrices des lois usuelles.

- 4 Si $X \sim \mathscr{U}(\llbracket 1, n \rrbracket)$, alors $G_X(t) = \dfrac{1}{n}(t + t^2 + \cdots + t^n)$.
- 3 Si $X \sim \mathscr{B}(p)$, alors $G_X(t) = pt + (1-p)$.
- Si $X \sim \mathscr{B}(n, p)$, alors $G_X(t) = (pt + (1-p))^n$.
- 2 Si $X \sim \mathscr{G}(p)$, alors $G_X(t) = \dfrac{pt}{1-(1-p)t}$.
- 1 Si $X \sim \mathscr{P}(\lambda)$, alors $G_X(t) = \mathrm{e}^{\lambda(t-1)}$.

Pour $X \sim \mathscr{P}(\lambda)$:

$$G_X(t) = \sum_{n=0}^{+\infty} P(X=n)\, t^n$$
$$= e^{-\lambda} \sum_{n=0}^{+\infty} \frac{\lambda^n}{n!} t^n$$
$$= e^{-\lambda} e^{\lambda t}$$
$$= e^{\lambda(t-1)}$$

Pour $X \sim \mathcal{G}(\lambda)$

$$G_X(t) = \sum_{n=1}^{+\infty} P(X=n)\, t^n$$

$$= \sum_{n=1}^{+\infty} (1-p)^{n-1} p\, t^n$$

$$= pt \sum_{n=1}^{+\infty} \left[(1-p)t\right]^{n-1}$$ Somme géom

$$= pt \cdot \frac{1}{1-(1-p)t}$$

Pour $X \sim \mathcal{B}(p)$

$$G_X(t) = \sum_{n=0}^{+\infty} P(X=n)\, t^n$$

$$P(X=n) = \begin{cases} 0 & \text{si } n \geq 2 \\ p & \text{si } n = 1 \\ (1-p) & \text{si } n = 0 \end{cases}$$

$$= \sum_{n \in X(\Omega)} P(X=n)\, t^n$$

$$= (1-p)\, t^0 + p\, t^1$$

Pour $X \sim \mathcal{B}(n,p)$

$$G_X(t) = \sum_{k=0}^{n} P(X=k)\, t^k$$

$$= \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k}\, t^k$$

$$= \left((1-p) + (pt)\right)^n$$

$X \sim \mathcal{U}([\![1, n]\!])$

$$G_X(t) = \sum_{k=0}^{+\infty} P(X=k)\, t^k$$
$$= \sum_{k=1}^{n} \frac{1}{n} t^k$$
$$= \frac{1}{n}\left(t + t^2 + \dots + t^n\right)$$

$$G_X(t) = 0.t^0 + \underbrace{\frac{1}{n}}_{P(X=1)} t + \underbrace{\frac{1}{n}}_{P(X=2)} t^2 + \dots + \underbrace{\frac{1}{n}}_{P(X=n)} t^n + 0\, t^{n+1}$$

Proposition. La loi d'une variable aléatoire à valeurs dans $\mathbb{N}$ est caractérisée par sa fonction génératrice.

$\Longleftrightarrow$

Si la loi est donnée, alors $G_X(t) = \sum_{n \in \mathbb{N}} P(X=n)\, t^n$

Récipr. Donner $G_X(t)$, c'est donner son D.S.E.

donc, par unicité des coeff, les $P(X=n)$

$\boxed{85.6}$ (a)(b)

On note $Y_1, Y_2, \dots, Y_n, \dots$ les va de Bernoulli

associées à chaque saut.

les Y_n sont indépendantes, et $P(Y_n=1) = \frac{1}{n}$

- $X(\Omega) \subset \mathbb{N}$ $\qquad$ $X+1$ est le rang du 1er échec

- Pour $n \in \mathbb{N}$

$$(X=n) = \bigcap_{k=1}^{n} (Y_k = 1) \cap (Y_{n+1} = 0)$$

Par indép $\quad P(X=n) = \prod_{k=1}^{n} \frac{1}{k} \times \left(1 - \frac{1}{n+1}\right)$

$$= \frac{n}{(n+1)!}$$

$$G_X(t) = \sum_{n=\cancel{0}\,1}^{+\infty} \frac{n+1-1}{(n+1)!} t^n \qquad \text{pour } t \in [-1,1] \setminus \{0\}$$

$$= \sum_{n=1}^{+\infty} \frac{1}{n!} t^n - \frac{1}{(n+1)!} t^n$$

$$= (e^t - 1) - \frac{1}{t} \sum_{n=1}^{+\infty} \frac{t^{n+1}}{(n+1)!}$$

$$= (e^t - 1) - \frac{1}{t}(e^t - 1 - t)$$

2 Propriétés, régularité

Proposition. On conserve les notations précédentes. G_X est continue sur $[-1,1]$ (et même sur $DF(O,1)$ si on considère la variable complexe).

Proposition. On conserve les notations précédentes.

X admet une espérance finie si et seulement si G_X est dérivable en 1 (à gauche).

Dans ce cas :

$$\mathrm{E}(X) = G'_X(1)$$

Exemple:

$X \sim \mathcal{B}(n, p)$ $\quad G_X(t) = (1-p+pt)^n$

polynomiale donc dérivable en 1

$G'_X(t) = np(1-p+pt)^{n-1}$

$G'_X(1) = np$

donc X est d'esp finie et $E(X) = np$.

Preuve: $\boxed{\Rightarrow}$ On suppose $\sum_{n=0}^{+\infty} n P(X=n) < +\infty$

$G_X(t) = \sum_{n=0}^{+\infty} \underbrace{P(X=n) t^n}_{f_n(t)}$ $\quad R \geq 1$

1 est peut être au bord de $]-R, R[$

G_X est continue sur $[-1, 1]$

Dérivons $\sum f_n$:

* $\sum f_n$ cv simplement sur $]-1, 1]$
* les f_n sont $\mathcal{C}^1$ $\quad f_n'(t) = n P(X=n) t^{n-1}$
* cv unif de $\sum f_n'$?

$\forall t \in]-1, 1]$ $\quad |f_n'(t)| \leq n P(X=n)$

indép de t

tg série cv car $E(X) < +\infty$

donc $\sum f'_n$ cv normalement sur $]-1, 1]$

donc unif

donc G_X est dérivable (et $\mathcal{C}^1$) sur $]-1, 1]$

et $G'_X(t) = \sum_{n=1}^{+\infty} n P(X=n) t^{n-1}$

en part $G'_X(1) = \sum_{n=1}^{+\infty} n P(X=n)$

$= E(X)$

$\boxed{\Leftarrow}$ [...]

$\boxed{85.6}$ $G_X(t) = (e^t - 1) - \frac{1}{t}(e^t - 1 - t)$

G_X est dérivable en 1 donc $X \in L^1$

et $G'_X(t) = e^t + \frac{1}{t^2}(e^t - 1 - t) - \frac{1}{t}(e^t - 1)$

donc $E(X) = G'_X(1)$

$= e + (e - 1 - 1) - 1(e - 1)$

$= e - 1$

$X \in L^2$

Proposition. On conserve les notations précédentes.
X admet une variance si et seulement si G_X est deux fois dérivable en 1 (à gauche).
Dans ce cas :

$$G''_X(1) = \mathrm{E}(X(X-1))$$

Remarque. *De cette égalité, il faut savoir retrouver rapidement l'expression de la variance à l'aide de G :*

$$\mathrm{V}(X) = G''_X(1) + G'_X(1) - (G'_X(1))^2$$

$$V(X) = E(X^2) - E(X)^2$$
$$= E(X(X-1)) + E(X) - E(X)^2$$

$\boxed{85.6}$

$$G'_X(t) = e^t + \frac{1}{t^2}(e^t - 1 - t) - \frac{1}{t}(e^t - 1)$$

dérivable en 1 donc X^2 est d'esp finie

$$G''_X(t) = e^t - \frac{2}{t^3}(e^t - 1 - t) + \frac{2}{t^2}(e^t - 1) - \frac{1}{t}(e^t)$$

$$G''_X(1) = e - 2(e-2) + 2(e-1) - e$$
$$= 2$$

Puis

$$V(X) = E(X(X-1)) + E(X) - E(X)^2$$
$$= G''_X(1) + G'_X(1) - G'_X(1)^2$$
$$= 2 + (e-1) - (e-1)^2$$

Exemple. Retrouver par les fonctions génératrices espérance et variance des lois usuelles.

3 Fonction génératrice et somme

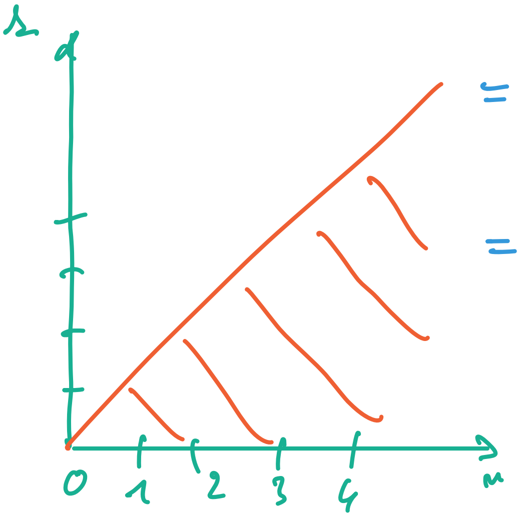
Proposition. Soit X et Y deux variables aléatoires à valeurs dans $\mathbb{N}$. Si X et Y sont indépendantes, alors pour tout $t \in]-1,1[$:

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

où $G_X(t)\, G_Y(t)$ est le produit de Cauchy des deux séries entières.

Preuve : pour $t>0$, dans $[0,+\infty[\cup \{+\infty\}$

$$G_{X+Y}(t) = \sum_{n\in\mathbb{N}} \underbrace{P(X+Y=n)}_{\sum_{k=0}^{n} P(X=k,\, Y=n-k)} t^n$$



$$= \sum_{n=0}^{+\infty} \sum_{k=0}^{n} P(X=k,\, Y=n-k)\, t^n$$

$$= \sum_{n=0}^{+\infty} \sum_{k=0}^{n} P(X=k)\, P(Y=n-k)\, t^n$$

par indép.

$$= \sum_{k=0}^{+\infty} \sum_{n=k}^{+\infty} P(X=k)\, P(Y=n-k)\, t^n$$

par Fubini positif.

$$= \sum_{k=0}^{+\infty} \left(P(X=k)\, t^k \underbrace{\sum_{n=k}^{+\infty} P(Y=n-k)\, t^{n-k}}_{\underbrace{\sum_{n=0}^{+\infty} P(Y=n)\, t^n}_{G_Y(t)}} \right)$$

$$= G_Y(t) \sum_{k=0}^{+\infty} P(X=k)\, t^k$$

$$= G_Y(t)\, G_X(t)$$

$$G_{X+Y}(t) = E(t^{X+Y}) = E(t^X t^Y) \overset{\text{indép.}}{=} E(t^X)\, E(t^Y) = G_X(t)\, G_Y(t)$$

Exemple. Soit X et Y deux variables aléatoires indépendantes suivant des lois de Poisson de paramètres respectifs λ et μ. Alors :

$$X + Y \sim \mathscr{P}(\lambda + \mu)$$

$$P(X+Y=n) = \sum_k P(X=k, Y=n-k)$$

[...]

X et Y indép donc

$$G_{X+Y}(t) = G_X(t)\, G_Y(t)$$
$$= e^{\lambda(t-1)}\, e^{\mu(t-1)}$$
$$= e^{(\lambda+\mu)(t-1)}$$

on reconnaît la fct génératrice de $\mathscr{P}(\lambda+\mu)$

donc $X+Y \sim \mathscr{P}(\lambda+\mu)$

Proposition. Soit $X_1, X_2, \ldots, X_n$ des variables aléatoires à valeurs dans $\mathbb{N}$. Si elles sont indépendantes, alors pour tout $t \in]-1, 1[$:

$$G_{X_1+X_2+\cdots+X_n}(t) = G_{X_1}(t)\, G_{X_2}(t) \ldots G_{X_n}(t)$$

Exemple:

$X_1 \ldots X_n$ des variables de Bernoulli i.i.d. $\mathcal{B}(p)$

$Y = X_1 + \cdots + X_n \sim \mathcal{B}(n, p)$

$$G_Y(t) = \prod_{k=1}^{n} G_{X_k}(t) \qquad \text{par indép}$$

$$= \prod_{k=1}^{n} (1-p+pt)$$

$$= (1-p+pt)^n$$