Act-3000 Théorie du risque – A2018
Rappel sur les FGP

Professeur : Etienne Marceau

28 septembre 2018

1 Contexte

Ce document fournit un bref rappel sur les notions des fonctions génératrices de probabilité pour les v.a. discrètes.

2 Définition

Soit une v.a. discrète X définie sur l'ensemble $\mathbb{N}$. La fonction de masse de probabilité (f.m.p.) est notée par

$$f_X(k) = \Pr(X = k), \; k \in \mathbb{N}.$$

La fonction de génératice de probabilité (f.g.p.) de la v.a. X permet de représenter la f.m.p. de la v.a. X sous la forme d'une série de puissances.

Les coéfficients de cette série de puissances correspondent aux valeurs de la fonction de masse de probabilité.

Définition 1 *La fonction génératrice de probabilités de la v.a. X est définie par*

$$\mathcal{P}_X(t) = E\left[t^X\right] = \sum_{k=0}^{\infty} f_X(k)\, t^k,$$

pour tout nombre complexe t tel que $|t| \leq 1$ (en particulier pour des nombres réels $t \in [0,1]$).

3 Propriétés

Propriété 2 $\mathcal{P}_X(0) = f_X(0)$ *et* $\mathcal{P}_X(1) = 1$.

Propriété 3 Fonction de masse de probabilité. *La valeur de $f_X(k)$ est calculée à partir de $\mathcal{P}_X(t)$ avec*

$$f_X(k) = \frac{1}{k!} \frac{\mathrm{d}^k}{\mathrm{d}t^k} \mathcal{P}_X(t) \bigg|_{t=0}. \tag{1}$$

Propriété 4 *Soit deux v.a.* X *et* Y*. Si* $\mathcal{P}_X$ *et* $\mathcal{P}_Y$ *sont identiques. Alors, selon la Propriété 3, les v.a.* X *et* Y *ont la même distribution.*

Propriété 5 Espérance. *Supposons que $E[X] < \infty$. Alors, on a*

$$E[X] = \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{P}_X(t)\Big|_{t=1}$$

4 Exemples

Exemple 6 $\mathcal{P}_X(t) = t^c$ *(*$c \in \mathbb{N}$*) correspond à la f.g.p. d'une v.a.* X *où* $\Pr(X = c) = 1$*.*

Exemple 7 *Soit une v.a. discrète X avec*

$$\mathcal{P}_X(t) = 0.2 + 0.4t + 0.3t^2 + 0.1t^3, \text{ pour } t \in [0,1].$$

Sans recourir à (1), on déduit aisément les valeurs suivantes de la fonction de masse de probabilité de X :

k	0	1	2	3
$f_X(k)$	0.2	0.4	0.3	0.1

Exemple 8 *Soit une v.a. discrète X avec*

$$\mathcal{P}_X(t) = 0.2 + 0.4t^2 + 0.3t^7 + 0.1t^{15}, \text{ pour } t \in [0,1].$$

Il n'est pas nécessaire d'utiliser (1) pour obtenir les valeurs suivantes de la fonction de masse de probabilité de X :

k	0	2	7	15
$f_X(k)$	0.2	0.4	0.3	0.1

5 Exemples – Lois paramétriques

Pour certaines lois discrètes connues, la fonction de "t", $\mathcal{P}(t) = \sum_{k=0}^{\infty} f_X(k)\, t^k$, s'exprime sous une forme fermée. C'est notamment le cas des lois de Poisson, binomiale et binomiale négative.

En appliquant (1), on verra dans le cours que la fonction de masse de probabilité de ces trois lois paramétriques satisfait une relation récursive.

5.1 Loi de Poisson

Poisson ($\lambda \in \mathbb{R}^+$) :

$$\mathcal{P}(t) = \sum_{k=0}^{\infty} f_X(k)\, t^k = \mathrm{e}^{\lambda(t-1)},$$

pour tout nombre complexe t tel que $|t| \leq 1$ (en particulier pour des nombres réels $t \in [0, 1]$).

5.2 Loi de binomiale

Binomiale ($n \in \mathbb{N}^+$; $q \in (0,1)$) :

$$\mathcal{P}(t) = \sum_{k=0}^{\infty} f_X(k)\, t^k = (qt + 1 - q)^n ,$$

pour tout nombre complexe t tel que $|t| \leq 1$ (en particulier pour des nombres réels $t \in [0,1]$).

5.3 Loi de binomiale négative

Binomiale négative ($r \in \mathbb{R}^+$; $q \in (0,1)$) :

$$\mathcal{P}(t) = \sum_{k=0}^{\infty} f_X(k)\, t^k = \left(\frac{q}{1-(1-q)t}\right)^r,$$

pour tout nombre complexe t tel que $|t| \leq 1$ (en particulier pour des nombres réels $t \in [0,1]$).

6 Somme de v.a. indépendantes

6.1 Relation de base

Soit les v.a. indépendantes X_1, ..., X_n définie sur $\mathbb{N}$ avec

$$f_{X_i}(k) = \Pr(X_i = k), \; k \in \mathbb{N}, \; .$$

et

$$\mathcal{P}_{X_i}(t) = \sum_{k=0}^{\infty} f_X(k) \, t^k = E\left[t^{X_i}\right], t \in [0,1], i = 1, 2, ..., n.$$

On définit $S = X_1 + ... = X_n$.

Alors, on a

$$\begin{aligned}\mathcal{P}_S(t) &= E\left[t^S\right] \\ &= E\left[t^{X_1+\ldots+X_n}\right] \\ &= E\left[t^{X_1}\right] \times \ldots \times E\left[t^{X_n}\right]\end{aligned}$$

qui devient

$$\mathcal{P}_S(t) = \mathcal{P}_{X_1}(t) \times \ldots \times \mathcal{P}_{X_n}(t), \ t \in [0,1]. \tag{2}$$

La relation en (2) est fort connue.

6.2 En coulisses

Habituellement, on se contente du formalisme et des expressions fermées pour les lois discrètes connues, tout en oubliant ce qui se passe en coulisses en multipliant les séries convergentes.

Dans le cas de v.a. définies sur des supports finis, $\mathcal{P}_{X_i}(t)$ correspondent à des polynômes de "t".

En appliquant la relation en (2), on multiplie des polynômes et le résultat de ce produit est, $\mathcal{P}_S(t)$, est lui-même un polynôme.

Les coefficients de ce polynôme correspondent aux valeurs de la fonction de masse de probabilité de la v.a. S.

Cette procédure est illustrée dans les prochains exemples.

Cette procédure est identique dans le cas de v.a. définies sur $\mathbb{N}$.

6.3 Exemple #1

Soit les v.a. indépendantes X_1 et X_2 avec

$$f_{X_i}(k) = \Pr(X_i = k) > 0,\ k \in \{0, 1, 2\},\ i = 1, 2.$$

On définit $S = X_1 + X_2$.

On vise à calculer les valeurs de

$$f_S(k) = \Pr(S = k) > 0,\ k \in \{0, 1, 2, 3, 4\}.$$

Pour y parvenir, on applique (2)

$$\begin{aligned}\mathcal{P}_S(t) &= \mathcal{P}_{X_1}(t) \times \mathcal{P}_{X_2}(t) \\ &= \left(f_{X_1}(0) + f_{X_1}(1)\, t + f_{X_1}(2)\, t^2\right) \times \left(f_{X_2}(0) + f_{X_2}(1)\, t + f_{X_2}(2)\, t^2\right)\end{aligned}$$

et, on obtient

$$\begin{aligned}\mathcal{P}_S(t) &= f_{X_1}(0) f_{X_2}(0) \\ &+ \left(f_{X_1}(0) f_{X_2}(1) + f_{X_1}(1) f_{X_2}(0)\right) \times t \\ &+ \left(f_{X_1}(0) f_{X_2}(2) + f_{X_1}(1) f_{X_2}(1) + f_{X_1}(2) f_{X_2}(0)\right) \times t^2 \\ &+ \left(f_{X_1}(1) f_{X_2}(2) + f_{X_1}(2) f_{X_2}(1)\right) \times t^3 \\ &+ \left(f_{X_1}(2) f_{X_2}(2)\right) \times t^4.\end{aligned} \tag{3}$$

Ensuite, puisque

$$\mathcal{P}_S(t) = f_S(0) + f_S(1)\, t + f_S(2)\, t^2 + f_S(3)\, t^3 + f_S(4)\, t^4. \tag{4}$$

Enfin, comme (3) et (4) sont identiques, on identifie les coefficients ($f_S(0)$, ..., $f_S(4)$) du polynôme en "t" de (4) à partir des coefficients du polynôme en "t" de (3).

6.4 Exemple #2

Cette procédure est identique dans le cas de v.a. définies sur $\mathbb{N}$.

Soit les v.a. indépendantes X_1 et X_2 avec

$$\mathcal{P}_{X_i}(t) = \sum_{k=0}^{\infty} f_{X_i}(k)\, t^k,\ i = 1, 2.$$

On définit $S = X_1 + X_2$.

Clairement, $\mathcal{P}_{X_1}(t)$ et $\mathcal{P}_{X_2}(t)$ sont des polynômes d'ordre m_1 et m_2 respectivement dont l'argument (la variable) est $t \in [0, 1]$.

Alors,

$$\begin{aligned}\mathcal{P}_S(t) &= \mathcal{P}_{X_1}(t) \times \mathcal{P}_{X_2}(t) \\ &= \left(\sum_{k_1=0}^{\infty} f_{X_1}(k_1)\, t^{k_1}\right) \times \left(\sum_{k_2=0}^{\infty} f_{X_2}(k_2)\, t^{k_2}\right)\end{aligned}$$

est un polynôme d'ordre $m_1 + m_2$ de "t", dont les coefficients sont les valeurs de f_S.

En effet, on a

$$\begin{aligned}\mathcal{P}_S(t) &= \left(\sum_{k_1=0}^{\infty} f_{X_1}(k_1)\, t^{k_1}\right) \times \left(\sum_{k_2=0}^{\infty} f_{X_2}(k_2)\, t^{k_2}\right) \\ &= \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} f_{X_1}(k_1)\, t^{k_1} \times f_{X_2}(k_2)\, t^{k_2} \qquad (5)\end{aligned}$$

La double somme en (5) peut se résumer en somme unique en rassemblant les coefficients de t^l pour chaque $l \in \mathbb{N}$, comme suit :

$$\begin{aligned}\mathcal{P}_S(t) &= \sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty} f_{X_1}(k_1)\, t^{k_1} \times f_{X_2}(k_2)\, t^{k_2} \\ &= \sum_{l=0}^{\infty} t^l \left(\sum_{j=0}^{l} f_{X_1}(j)\, f_{X_2}(l-j)\right), \qquad (6)\end{aligned}$$

où le coefficient de t^l, donné par

$$\sum_{j=0}^{l} f_{X_1}(j)\, f_{X_2}(l-j),$$

est la valeur de $f_S(l)$ par la définition de la f.g.p de S.

En effet, la définition de la f.g.p. de la v.a. S est

$$\mathcal{P}_S(t) = \sum_{l=0}^{\infty} f_S(l)\, t^l. \tag{7}$$

Puisque les deux représentations en (6) et (7) de la f.g.p. de la v.a. S sont identiques, il suffit d'identifier (pour chaque l) le coefficient de t^l dans (6) pour déterminer l'expression de $f_S(l)$ dans (7), permettant de conclure que

$$f_S(l) = \sum_{j=0}^{l} f_{X_1}(j)\, f_{X_2}(l-j),$$

ce qui correspond au produit de convolution de 2 fonctions de masse probabilité.

Cette "identification" des coefficients de t^l est essentielle pour comprendre les démonstrations des algorithmes récursifs.

7 Somme de v.a. i.i.d.

7.1 Contexte

Soit les v.a. $X_1, ..., X_n$ discrètes i.i.d. définies sur $\mathbb{N}$ avec

$$f_{X_i} = f_X \text{ et } \mathcal{P}_{X_i} = \mathcal{P}_X,\ i = 1, 2, ..., n.$$

On définit $S_n = X_1 + ... + X_n$ avec

$$f_{S_n}(k) = f_{X_1+...+X_n}(k) = f_X^{*n}(k)$$

où f_X^{*n} est le n-ième produit de convolution de f_X avec elle-même.

La f.g.p. de la v.a. S_n est

$$\begin{aligned} \mathcal{P}_{S_n}(t) &= E\left[t^{S_n}\right] \\ &= \mathcal{P}_X(t)^n = \sum_{k=0}^{\infty} f_{S_n}(k) \times t^k. \end{aligned}$$

Le calcul des valeurs de f_{S_n} via les f.g.p. est illustré par un exemple.

7.2 Exemple

Soit la v.a. $X \in \{0, 1, 2\}$ avec

$$\mathcal{P}_X(t) = \left(0.3 + 0.5t + 0.2t^2\right).$$

On définit $S_5 = X_1 + ... + X_5$ (une somme de v.a. i.i.d. avec $X_i \sim X$, $i = 1, 2, ..., 5$) avec

$$\mathcal{P}_{S_5}(t) = \mathcal{P}_X(t)^5 = \left(0.3 + 0.5t + 0.2t^2\right)^5.$$

Clairement,

$$f_{S_5}(0) = P_{S_5}(0) = 0.3^5.$$

De la Propriété 3, on déduit

$$\begin{aligned} f_{S_5}(1) &= \left.\frac{\mathsf{d}}{\mathsf{d}t}\mathcal{P}_{S_5}(t)\right|_{t=0} \\ &= \left. 5 \times \left(\mathcal{P}_X(t)^4\right) \times \mathcal{P}'_X(t)\right|_{t=0} \\ &= 5 \times f_{S_4}(0) \times f_X(1) \end{aligned}$$

avec

$$f_{S_4}(0) = f_X(0)^4 .$$

Toutefois, cette approche, quoique correcte, mais fastidieuse, ne conduit pas à un algorithme satisfaisant.

8 Somme aléatoire de v.a. i.i.d.

8.1 Contexte

Soit une v.a. X définie par

$$X = \begin{cases} \sum_{k=1}^{M} B_i & , \ M > 0 \\ 0, & , \ M = 0 \end{cases} , \tag{8}$$

où

- M est une v.a. discrète de fréquence ;
- $\underline{B} = \left\{ B_k, k \in \mathbb{N}^+ \right\}$ est une suite de v.a. positives i.i.d. définies sur $\mathbb{N}$, avec $B_k \sim B$; et

– $\underline{B}$ est indépendante de M.

Alors, la v.a. X prend des valeurs dans $\mathbb{N}$ avec

$$f_X(k) = \Pr(X = k), \; k \in \mathbb{N}.$$

La f.g.p. de la v.a. X est

$$\mathcal{P}_X(t) = \mathcal{P}_M(\mathcal{P}_B(t)).$$

L'objectif est de calculer les valeurs de f_X avec les f.g.p., comme on l'illustre par un exemple.

8.2 Exemple

Soit les v.a. $M \sim Binom\,(2, q)$ et $B \in \{0, 1, 2, 3\}$ avec

$$\Pr(B = k) = b_k,\ k = 0, 1, 2, 3,$$

et

$$\mathcal{P}_B(t) = \sum_{k=0}^{3} b_k t^k.$$

La f.g.p. de X est

$$\begin{aligned} \mathcal{P}_X(t) &= \mathcal{P}_M(\mathcal{P}_B(t)) \\ &= (1 - q + q\mathcal{P}_B(t))^2 \\ &= (1-q)^2 + 2(1-q)\,q\mathcal{P}_B(t) + q^2\mathcal{P}_B(t)^2 . \end{aligned}$$

Ensuite, la f.g.p. de X devient

$$\begin{aligned} \mathcal{P}_X(t) &= (1-q)^2 \\ &\quad +2(1-q)q \times \sum_{k=0}^{3} b_k t^k \\ &\quad +q^2 \times \left(\sum_{k=0}^{3} b_k t^k\right)^2 . \end{aligned}$$

Puis, on a

$$\begin{aligned} \mathcal{P}_X(t) &= (1-q)^2 \\ &\quad +2(1-q)q \times \sum_{k=0}^{3} b_k t^k \\ &\quad +q^2 \times \sum_{k=0}^{6} c_k t^k \end{aligned}$$

où

$$\sum_{k=0}^{6} c_k t^k = b_0^2 + 2b_1 b_0 t + \left(2b_2 b_0 + b_1^2\right) t^2 + (2b_3 b_0 + 2b_2 b_1)\, t^3$$
$$+ \left(2b_3 b_1 + b_2^2\right) t^4 + 2b_3 b_2 t^5 + b_3^2 t^6.$$

Finalement, en réarrangeant les termes, on obtient

$$\begin{aligned}
\mathcal{P}_X(t) &= \sum_{k=0}^{6} f_X(k)\, t^k \\
&= (1-q)^2 + 2(1-q)\, q \times b_0 + q^2 b_0^2 \\
&\quad + \Big(2(1-q)\, q \times b_1 + q^2 \times (2b_1 b_0)\Big) \times t \\
&\quad + \Big(2(1-q)\, q \times b_2 + q^2 \times \Big(2b_2 b_0 + b_1^2\Big)\Big) \times t^2 \\
&\quad + \Big(2(1-q)\, q \times b_3 + q^2 \times (2b_3 b_0 + 2b_2 b_1)\Big) \times t^3 \\
&\quad + q^2 \times \Big(2b_3 b_1 + b_2^2\Big) \times t^4 \\
&\quad + q^2 \times 2b_3 b_2 \times t^5 \\
&\quad + q^2 \times b_3^2 \times t^6.
\end{aligned}$$

Cette approche est applicable à la condition que $\mathcal{P}_M$ et $\mathcal{P}_B$ soient des polynômes de degrés qui ne sont pas trop elévés.

Il est possible d'établir une procédure plus efficace.

9 Conclusion

Les exemples ont permis d'illustrer le calcul des valeurs de la fonction de masse de probabilité de S, selon une approche dite de la "force brute" (*brute force*, en anglais).

L'objectif des algorithmes récursifs est d'identifiquer ces coefficients de façon efficace.