Approfondissement de cours sur les variables aléatoires : fonction génératrice et application

Valentin Melot — Lycée N.-D. de la Providence

15 janvier 2021

La *fonction génératrice* est un moyen de faire le lien entre étude d'une loi d'une variable aléatoire et analyse réelle.

La section 2 peut être traitée indépendamment des sections 3 et 4. La section 4 suppose d'avoir traité la section 3 au préalable.

Dans tout le problème, on se donne un espace probabilisé fini $(\Omega, \mathbf{P})$ quelconque.

1 Définition et premiers exemples

Soit X une variable aléatoire réelle discrète dont le support est inclus dans $\mathbf{N}$. On définit la *fonction génératrice* de X comme :

$$G_X : t \longmapsto \sum_{k=0}^{\infty} \mathbf{P}(X = k) t^k.$$

Deux cas de figure peuvent se présenter :

— Si le support de X est fini, c'est-à-dire si X ne peut prendre qu'un nombre distinct de valeurs, alors cette somme ne comporte qu'un nombre fini de termes non-nuls. Autrement dit, c'est un polynôme, dont le degré est la plus grande valeur que peut prendre X.

— S'il est infini, cette somme comporte un nombre infini de termes non-nuls. Elle est ce que l'on appelle une *série*, objet essentiel en mathématiques dont l'étude est au programme de L1.

Nous nous limitons pour ce problème au premier cas.

1. Soit $X \sim \mathscr{B}(p)$. Exprimer la fonction G_X sans utiliser le signe somme.

2. Soit $X \sim \mathscr{B}(n; p)$. Exprimer la fonction G_X sous la forme d'une somme finie. En utilisant la formule du binôme de Newton, exprimer G_X sous la forme d'une fonction simple.

Soit E un ensemble fini. On dit que X suit une loi uniforme sur un ensemble E si l'application $i \mapsto \mathbf{P}(X = i)$ est constante sur E, ou autrement dit si toutes les valeurs de E sont équiprobables pour $\mathbf{P}_X$. On note alors $X \sim \mathscr{U}(E)$.

3. Exprimer G_X lorsque $X \sim \mathscr{U}(\{1, 2, \ldots, 6\})$, puis lorsque $X \sim \mathscr{U}(\{2, \ldots, 12\})$.

4. Soit X une variable aléatoire entière à support fini, de loi inconnue, et $t \in \mathbf{R}$. Démontrer que $\mathbf{E}\left[t^X\right] = G_X(t)$. *On pourra notamment utiliser la formule de transfert.*

En pratique, c'est le plus souvent cette égalité qui est utilisée pour définir la fonction génératrice.

2 Génératrice et dérivation

On considère dans ce début de section une variable aléatoire entière à support fini X, de loi inconnue.

1. Calculer $G_X(0)$.

2. Démontrer que G_X est dérivable en 0, et que $G'_X(0) = \mathbf{P}(X = 0)$.

3. On appelle G''_X la dérivée de G'_X, et plus généralement $G_X^{(n+1)}$ la dérivé de $G_X^{(n)}$ (dérivée n-ième). Calculer $G''_X(0)$, $G_X^{(3)}(0)$, $G_X^{(4)}(0)$. Quelle conjecture peut-on faire ?

4. Soit n un entier naturel et $f : x \mapsto x^n$. Par récurrence, démontrer que

$$f^{(m)} : x \longmapsto \frac{n!}{(n-m)!} x^{n-m}$$

si $m \leqslant n$ et que $f^{(m)} = 0$ si $m > n$.

5. En déduire l'expression de $G_X^{(m)}$ quel que soit $m \in \mathbf{N}$.

6. Exprimer la loi de X en fonction de G_X et de ses dérivées.

On a démontré que la fonction génératrice comportait toute l'information sur une loi de probabilités, dans le cas où celle-ci est à support fini.

7. Calculer $G_X(1)$ et $G'_X(1)$.

8. Retrouver l'expression de l'espérance d'une variable aléatoire suivant une loi binomiale.

9. Montrer que $G''_X(1) = \mathbf{E}[X(X-1)]$.

10. En déduire une expression de $V(X)$ faisant intervenir $G'_X(1)$ et $G''_X(1)$.

11. Retrouver, l'expression de la variance d'une variable aléatoire suivant une loi binomiale.

12. Plus généralement, expliquer comment on peut calculer $\mathbf{E}[X^n]$ à partir des dérivées première, seconde, etc. et n-ième de G_X calculées en 1.

3 Génératrice et indépendance

On admet le résultat suivant : si X et Y sont deux variables aléatoires indépendantes et f et g sont deux fonctions, alors $\mathbf{E}[f(X)g(Y)] = \mathbf{E}[f(X)]\mathbf{E}[g(Y)]$.

1. Soit X et Y deux variables aléatoires entières indépendantes exprimer G_{X+Y} en fonction de G_X et G_Y.

2. On suppose maintenant que X et Y sont indépendantes et de même loi. Exprimer G_{X+Y}.

3. Construire un contre-exemple au résultat admis dans le cas où X et Y ne sont pas indépendantes.

4 Application : dé pipé

On souhaite démontrer qu'il est impossible de truquer un dé de façon à ce que, si l'on lance ce dé deux fois indépendamment, le résultat du lancer suive une loi uniforme sur $\{2, \ldots, 12\}$.

Pour cela, on raisonne par l'absurde. Supposons qu'il existe une telle loi, c'est-à-dire des valeurs $p_1, p_2, \ldots, p_6$ telle que si X et Y sont deux variables aléatoires indépendantes vérifiant pour tout $k \in \{1, \ldots, 6\}$: $\mathbf{P}(X = k) = p_k$, alors $X + Y \sim \mathscr{U}(\{2, \ldots, 12\})$.

1. Exprimer G_X et G_Y.

2. Montrer qu'il existe un polynôme f de degré 10, relativement simple, tel que

$$G_{X+Y} : t \mapsto \frac{1}{11} t^2 f(t)$$

.

3. Justifier que G_X peut se factoriser par t.

4. En déduire que f est le carré d'un certain polynôme g. Quel est le degré de g ? Quel est son coefficient constant ? Quel est le coefficient de son monôme de plus grand degré ?

5. En déduire les limites de g.

6. On admet (et l'on démontrera dans le chapitre sur la continuité) que parce que g est un polynôme, le résultat précédent suffit à démontrer que g possède (au moins) une racine. Soit z une telle racine. Justifier que z est une racine de f.

7. En utilisant le résultat de première sur la somme des termes d'une suite géométrique, montrer que $f(x) = 0$ si et seulement si $t = 1$ ou $t^{2n+1} = 1$.

8. En déduire que f ne possède aucune racine.

9. Conclure.

10. Montrer qu'un tel résultat peut être généralisé à un dé à $2n$ faces, quel que soit $n \in \mathbf{N}^*$ (un dé à deux faces étant un pièce de monnaie).

5 Conclusion

Dans le cas fini, nous avons pu démontrer que la fonction génératrice d'une variable aléatoire portait en elle toute l'information disponible sur cette variable aléatoire. Des outils élémentaires d'analyse réelle (dérivation) permettent de déduire de cette fonction des propriétés simples de cette variable aléatoire. On peut aussi résoudre des problèmes plus complexes, faisant intervenir des sommes de variables aléatoires indépendantes.

La plupart des propriétés vues ici restent vraies dans le cas des variables aléatoires à support infini (c'est-à-dire pouvant prendre des valeurs arbitrairement grandes), comme la loi de Poisson. Néanmoins, des problèmes de convergence de la somme infinie utilisée dans la définition peuvent apparaître : il faut s'intéresser à la théorie des séries (niveau L1) pour les résoudre. Une fois ces problèmes identifiés, la série génératrice devient un outil puissant pour étudier des processus aléatoires complexes. Vous pouvez par exemple chercher, dans votre moteur de recherche préféré, ce qu'est le processus de Galton-Watson, qui modélise des dynamiques de populations en biologie. L'étude de ce processus repose essentiellement sur l'étude d'une série génératrice associée.

Plus généralement, l'idée consistant à introduire une somme de la forme de celle qui définit la série génératrice est fréquente en mathématiques. Elle permet de traduire de nombreux problèmes par une étude de certaines sommes infinies que l'on appelle les séries entières (dont l'étude générale est de niveau L2), dont le comportement est bien connu depuis le XIX^esiècle.