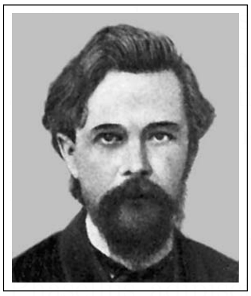
Variables aléatoires : compléments

Andreï Markov (1856–1922)



Mathématicien russe[1][2], est surtout connu pour ses travaux en probabilités.
Il est l'élève, entre autres, de Tchebychev et son étude de la distribution des lettres dans le roman *Eugène Onegin* d'Alexandre Pouchkine l'ont amené à développer l'outil qui sera appelé par la suite en son honneur « chaines de Markov ».
Ses travaux sont variés et touchent aussi bien l'algèbre (formes quadratiques binaires, c'est à dire les fonctions de la forme $q(x,y) = ax^2 + bxy + cy^2$), l'analyse (fractions continues et équations différentielles) que les probabilités.
On lui doit, entre autres :

- Les chaines de Markov qui permettent de modéliser des marches aléatoires :
  - à espace d'état discret : à chaque instant on est à une position choisie dans un ensemble ou on peut numéroter les éléments[3] ;
  - à temps discret : les instants sont eux aussi numérotables[4]
  - sans mémoire : toute l'information utile pour la prédiction du futur est contenue dans l'état présent.
- L'inégalité[5] de Markov (qu'il à co-découvert avec son frère Vladimir Andreïevich Markov)

Table des matières

[1] Merci qui ? Ben pas mamie Nova, mais MacTutor et Wikipedia...
[2] Le nom devrait être écrit en cyrillique, mais bon, euh, ben...
[3] Comme les postions sur un jeu de l'oie, un Monopoly ou (plus compliqué) les différentes configurations d'un échiquier.
[4] Comme les différents coups lors d'un jeu.
[5] Voir le IV de ce cours

Les questions

1 Quelles lois sont à connaître ?

Elles sont au nombre de 4 : uniforme, binomiale, géométrique et Poisson.

2 Que doit-on savoir sur elles ?

Tout ! Pour être précis : loi, moments, fonctions génératrices...

3 À quoi servent les fonctions génératrices ?

À calculer facilement des moments, à trouver des lois...

Les savoir faire

- Savoir calculer une fonction génératrice.
- Savoir utiliser une fonction génératrice pour trouver une loi, pour calculer une espérance ou une variance.
- Savoir calculer avec les lois usuelles.
- Savoir utiliser les inégalités de Markov ou de Bienaymé-Tchebychev pour majorer une probabilité.

I Vecteurs et indépendance

Dans tout ce paragraphe $(\Omega, \mathcal{A}, \mathbb{P})$ est un espace probabilisé.

1 Vecteurs

Propriété 1

Si $X_1, \ldots, X_p$ sont des variables aléatoires discrètes, alors U définie par :

$$\forall \omega \in \Omega,\ U(\omega) = (X_1(\omega), \ldots, X_p(\omega))$$

est une variable aléatoire discrète telle que $U(\Omega) \subset X_1(\Omega) \times \cdots \times X_p(\Omega)$.

Remarques

1. Par abus de notation, on écrira $U = (X_1, \ldots, X_p)$.
2. De même, si $U = (X_1, \ldots, X_p)$ et $V = (Y_1, \ldots Y_q)$ alors (U, V) sera noté sous la forme $(U, V) = (X_1, \ldots, X_p, Y_1, \ldots, Y_q)$.

2 Indépendance

Propriété 2

Si X_1, ..., X_p, Y_1,..., Y_q sont des variables aléatoires discrètes (mutuellement) indépendantes alors il en est de même pour U, Y_1, ..., Y_q où $U = (X_1, \ldots, X_p)$.

Propriété 3

Si U et V sont indépendantes et f et g deux fonctions définies respectivement sur $U(\Omega)$ et $V(\Omega)$ alors $f(U)$ et $g(V)$ sont indépendantes.

Exemple 1 Si $(X_k)_{n\geq 1}$ est une suite de variables aléatoires indépendantes alors pour tout n, $X_1 + \cdots + X_n$ et X_{n+1} sont indépendantes.

II Variables aléatoires à valeurs dans $\mathbb{N}$

A Espérance

Propriété 4

Soit X une variable aléatoire à valeurs dans $\mathbb{N}$. X a une espérance si et seulement si $\sum\limits_{n\geq 1} \mathbb{P}(X \geq n)$ converge. Dans le cas où cette série converge,

$$\mathbb{E}(X) = \sum_{n=1}^{+\infty} \mathbb{P}(X \geq n)$$

Remarques

1. Une autre version de cette formule est $\mathbb{E}(X) = \sum\limits_{m=0}^{+\infty} \mathbb{P}(X > m)$.
2. Cette formule fournit un lien entre la fonction de répartition et l'espérance. On rappelle que $F_X(n) = \mathbb{P}(X \leq n)$ et donc la la formule donne

$$\mathbb{E}(X) = \sum_{n=1}^{+\infty} (1 - F_X(n-1)) = \sum_{m=0}^{+\infty} (1 - F_X(m)).$$

B Fonctions génératrices

Dans tout ce paragraphe, on s'intéresse principalement aux variables aléatoires prenant leurs valeurs dans $\mathbb{N}$. On peut cependant étendre les idées à d'autres ensembles image. Il y aura cependant des difficultés techniques supplémentaires.

1 Définition

Lemme 5

Soit X une variable aléatoire prenant ses valeurs dans $\mathbb{N}$. Le rayon de convergence R_X de la série entière $\sum_{n\geq 0} \mathbb{P}(X=n)t^n$ vérifie $R_X \geq 1$.

Remarques

1. Le théorème du transfert permet de dire que

$$\forall t \in]-R_X, R_X[, \quad \mathbb{E}(t^X) = \sum_{n=0}^{+\infty} \mathbb{P}(X=n)t^n$$

2. La somme de la série est $\mathcal{C}^\infty$ sur $]-R_X, R_X[$.

Définition 1

Soit X une variable aléatoire prenant ses valeurs dans $\mathbb{N}$. La *fonction génératrice* de X est la somme :

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

Remarques

1. La fonction génératrice est au moins définie sur $[-1,1]$ et est $\mathcal{C}^\infty$ sur $]-1,1[$.
2. Dans d'autres filières (PC par exemple) la terminologie est « série » génératrice.
3. Par unicité du développement en série entière, deux variables aléatoires on même fonction génératrice si et seulement si elles ont même loi.

Exemples 2

1. Fonction génératrice de la loi uniforme.
2. Fonction génératrice d'une loi de Bernoulli, d'une loi binomiale.

Propriété 6

La Fonction génératrice d'une variable aléatoire caractérise sa loi.

Remarque Le mot génératrice dit bien ce qui se passe : la fonction *engendre* la loi : on peut retrouver la loi à partir de la somme de la série entière.

2 Moments

Propriété 7

Soit X une variable aléatoire à valeurs dans N et G_X sa fonction génératrice.

(1) X a une espérance si et seulement si G_X est dérivable en 1.

(2) dans ce cas : $\mathbb{E}(X) = G'_X(1)$.

Remarques

1. L'hypothèse de dérivabilité est vérifiée dans le cas $R_X > 1$ (voir les remarques autour du lemme 5).
2. Si $R_X = 1$ alors il est préférable[6] de calculer la somme puis de voir si on peut dériver en 1.

Propriété 8

Soit X une variable aléatoire à valeurs dans N et G_X sa fonction génératrice.

(1) X a une variance si et seulement si G_X est deux fois dérivable en 1.

(2) dans ce cas : $\mathbb{V}(X) = G''_X(1) + G'_X(1) - \left(G'_X(1)\right)^2$.

Remarques

1. Les remarques faites au théorème précédent tiennent encore.
2. En fait si G_X est deux fois dérivable en 1, c'est $X(X-1)$ qui a une espérance et on a $G''_X(1) = \mathbb{E}(X(X-1))$.

Exemples 3

1. Calcul de l'espérance de la loi uniforme
2. Calcul de l'espérance et de la variance d'une loi de Bernoulli, d'une loi binomiale.

[6] Les théorèmes généraux ne sont d'aucun intérêt dans ce cas : pour les utiliser il faut montrer la convergence de la série de somme $\mathbb{E}(X)$ pour obtenir la dérivabilité.

3 Somme

Propriété 9

Soient X et Y deux variables aléatoires **indépendantes** et à valeurs dans $\mathbb{N}$. Alors :

$$G_{X+Y} = G_X \times G_Y.$$

Remarques

1. Avec la propriété 6, on peut ainsi trouver la loi d'une somme.
2. En utilisant la remarque finale du I, on montre par récurrence sur $m \in \mathbb{N}^*$ que si X_1, …, X_m sont indépendantes et que $S = X_1 + \cdots + X_m$ alors $G_S = G_{X_1} \times \cdots \times G_{X_m}$.
3. Si en plus X_1,…, X_m sont indépendantes et de même loi[7], alors la fonction génératrice de $S = X_1 + \cdots + X_m$ vérifie :

$$G_S = \left(G_{X_1}\right)^m.$$

Exemples 4

1. Loi de la somme de deux variables suivant une loi uniforme.
2. Stabilité de la loi binomiale par la somme. Lien Bernoulli, binomiale.

III Lois discrètes usuelles

1 Loi uniforme

Cadre

C'est la loi du numéro tiré avec un dé non pipé ou une roulette ou tous les secteurs sont de même taille.

Loi

Définition 2

Soit $V \subset \mathbb{R}$ un ensemble **fini**. La variable aléatoire X suit la loi uniforme sur V, ce qu'on note $X \hookrightarrow \mathcal{U}_V$ si et seulement si :

$$\forall v \in V,\ \mathbb{P}(X = v) = \frac{1}{\text{card}(V)}.$$

Dans le cas où $V = [\![1, n]\!]$, on dira que $X \hookrightarrow \mathcal{U}_n$.

[7] On dit souvent « indépendantes et identiquement distribuées » abrégé en i.i.d.

Moments

Propriété 10

une variable aléatoire qui suit une loi uniforme a des moments de tous ordre. Notamment, si $X \hookrightarrow \mathcal{U}_n$ alors :

$$\mathbb{E}(X) = \frac{n+1}{2},$$
$$\mathbb{V}(X) = \frac{n^2-1}{12}.$$

Fonction génératrice

Propriété 11

Si $X \hookrightarrow \mathcal{U}_n$ alors :

$$R_X = +\infty, \quad \forall t \in \mathbb{R},\ G_X(t) = \frac{1}{n}(t + \cdots + t^n) = \begin{cases} \frac{1}{n}\frac{t - t^{n+1}}{1-t} & \text{si } t \neq 1, \\ 1 & \text{sinon.} \end{cases}$$

2 Loi binomiale

Cadre

Ici on compte le nombre de succès quand on répète la même épreuve indépendamment[8].

Loi

Définition 3

La variable aléatoire N suit la loi binomiale de paramètre $p \in]0,1[$, ce qu'on note $N \hookrightarrow \mathcal{B}(n,p)$ si et seulement si :

$$N(\Omega) = [\![0, n]\!], \quad \forall n \in [\![0, n]\!],\ \mathbb{P}(N = k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

Remarques

1. Le cas $n = 1$ est particulier : la loi $\mathcal{B}(1,p)$ est aussi appelée loi de Bernoulli et on l'abrège en $\mathcal{B}(p)$. Une variable aléatoire qui suit une loi de Bernoulli est en fait une indicatrice : celle du succès à l'unique épreuve du schéma de Bernoulli.

[8] On effectue un schéma de Bernoulli

2. Les cas $p = 0$ et $p = 1$ sont des cas extrêmes[9]. Ils correspondent aux cas N constante égale respectivement à 0 ou n.

Moments

Propriété 12

Si $N \hookrightarrow \mathcal{B}(n, p)$ alors N a des moments de tous ordres. Notamment :

$$\mathbb{E}(N) = np,$$
$$\mathbb{V}(N) = np(1-p).$$

Fonction génératrice

Propriété 13

Si $N \hookrightarrow \mathcal{B}(n, p)$ alors :

$$R_N = +\infty, \quad \forall t \in \mathbb{R},\ G_N(t) = \Big(1 - p + pt\Big)^n.$$

Et plus

Propriété 14

Si les variables B_1, ..., B_n sont indépendantes et suivent toutes la loi de Bernoulli de paramètre p alors :

$$N = B_1 + \cdots + B_n \hookrightarrow \mathcal{B}(n, p).$$

Remarque On a une forme de réciproque. Si $N \hookrightarrow \mathcal{B}(n, p)$ alors $X = B_1 + \cdots + B_n$, où B_k est la variable de Bernoulli indicatrice de l'événement « la k^{e} expérience est un succès ». Les variables (B_k) sont mutuellement indépendantes.

Propriété 15

Si X et Y sont deux variables aléatoires indépendantes suivant respectivement les lois $\mathcal{B}(n, p)$ et $\mathcal{B}(m, p)$ alors :

$$X + Y \hookrightarrow \mathcal{B}(n + m, p)$$

[9] Pour ne pas dire dégénérés...

3 Loi géométrique

Cadre

Cette loi est un *temps d'attente* : elle calcule le nombre d'essais indépendants jusqu'au premier succès. Il ne faut pas la confondre avec la loi binomiale. Ici le nombre d'expériences est aléatoire.

Loi

Définition 4

La variable aléatoire T suit la loi géométrique de paramètre $p \in]0,1[$, ce qu'on note $T \hookrightarrow \mathcal{G}(p)$ si et seulement si :

$$T(\Omega) = \mathbb{N}^*, \quad \forall n \in \mathbb{N}^*,\ \mathbb{P}(T = n) = (1-p)^{n-1}p.$$

Remarque Les cas $p = 0$ et $p = 1$ sont exclus car :

1. Si $p = 0$, alors il n'y a jamais de succès et donc « $T = +\infty$ » et on a donc pas de variable aléatoire.
2. Si $p = 1$ alors T est constante égale à 1.

Moments

Propriété 16

Si $T \hookrightarrow \mathcal{G}(p)$ alors T a des moments de tous ordres. Notamment :

$$\mathbb{E}(T) = \frac{1}{p},$$
$$\mathbb{V}(T) = \frac{1-p}{p^2}.$$

Fonction génératrice

Propriété 17

Si $T \hookrightarrow \mathcal{G}(p)$ alors :

$$R_T = \frac{1}{p}, \quad \forall t \in \left]-\frac{1}{p}, \frac{1}{p}\right[,\ G_T(t) = \frac{pt}{1-(1-p)t}.$$

Et plus

La loi géométrique est sans mémoire. Plus précisément :

Propriété 18

Si $T(\Omega) \subset \mathbb{N}$ alors $T \hookrightarrow \mathcal{G}(p)$ ssi :

$$\forall n, m \in \mathbb{N},\ \mathbb{P}(X > n + m \mid X > n) = \mathbb{P}(X > m)$$

4 Loi de Poisson

Cadre

Cette loi modélise le comptage d'événements rares. On sait qu'on est en présence d'une loi de Poisson quand on doit compter le nombre d'événements alors qu'on ne connaît que le nombre moyen de ceux-ci.
En fait, on est en présence[10] d'une loi de Poisson lorsque les événements qu'on compte sont rares, isolés et que leur nombre est peu ou prou proportionnel au temps d'étude.

Loi

Définition 5

La variable aléatoire U suit la loi de Poisson de paramètre $\lambda > 0$, ce qu'on note $U \hookrightarrow \mathcal{P}(\lambda)$ si et seulement si :

$$U(\Omega) = \mathbb{N}, \quad \forall n \in \mathbb{N},\ \mathbb{P}(U = n) = e^{-\lambda}\frac{\lambda^n}{n!}.$$

Remarque Le cas $\lambda = 0$ est exclu car il correspond à U constante égale à 0.

Moments

Propriété 19

Si $U \hookrightarrow \mathcal{P}(\lambda)$ alors U a des moments de tous ordres. Notamment :

$$\mathbb{E}(U) = \lambda,$$
$$\mathbb{V}(U) = \lambda.$$

[10] Voir le polycopié sur la loi de poisson.

Fonction génératrice

Propriété 20

Si $U \hookrightarrow \mathcal{P}(\lambda)$ alors :

$$R_U = +\infty, \quad \forall t \in \mathbb{R},\ G_U(t) = e^{\lambda(t-1)}.$$

Et plus

Propriété 21

Soient X et Y deux variables aléatoires indépendantes suivant chacune une loi de poisson de paramètres respectifs λ et μ. Alors :

$$X + Y \hookrightarrow \mathcal{P}(\lambda + \mu).$$

Remarque En utilisant la remarque suivant la propriété 9, on peut étendre la propriété précédente à une somme quelconque de variables indépendantes suivant une loi de Poisson.

IV Théorèmes d'approximation

1 Théorèmes généraux

Théorème 22 (Inégalité de Markov)

Soit X une variable aléatoire réelle discrète à valeurs positives. Alors, pour tout $\lambda > 0$,

$$\mathbb{P}(X \geq \lambda) \leq \frac{\mathbb{E}(X)}{\lambda}$$

Remarques

1. On voit que, puisque X est positive, on a pour tout $p \in \mathbb{N}^*$ l'égalité d'événements :

$$(X \geq \lambda) = (X^p \geq \lambda^p).$$

   De ce fait, si X^2 a une espérance, l'inégalité de Markov donne :

$$\mathbb{P}(X \geq \lambda) \leq \frac{\mathbb{E}(X^2)}{\lambda^2}$$

2. Si X n'est pas à valeurs positives, on remplace X par $|X|$ pour majorer $\mathbb{P}(|X| \geq \lambda)$.

Propriété 23

(Inégalité de Bienaymé-Tchebichev) Si X une variance alors pour tout $\lambda > 0$:

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq \lambda) \leq \frac{\mathbb{V}(X)}{\lambda^2}.$$

2 Résultats asymptotiques

Propriété 24

Si $X_n \hookrightarrow \mathcal{B}(n, p_n)$ et $np_n \xrightarrow[n \longrightarrow +\infty]{} \lambda$ alors :

$$\forall k \in \mathbb{N}, \quad \lim_{n \longrightarrow +\infty} \mathbb{P}(X_n = k) = e^{-\lambda} \frac{\lambda^k}{k!}.$$

Théorème 25 (Loi faible des grands nombres)

Soit $\left(X_n\right)_{n \geq 1}$ une suite de variables aléatoires indépendantes et ayant toutes la même espérance μ et la même variance σ^2. Alors :

$$\mathbb{P}\left(\left|\frac{X_1 + \cdots + X_n}{n} - \mu\right| \geq \varepsilon\right) \leq \frac{\sigma^2}{n\varepsilon^2}$$

Par suite

$$\mathbb{P}\left(\left|\frac{X_1 + \cdots + X_n}{n} - \mu\right| \geq \varepsilon\right) \xrightarrow[n \longrightarrow +\infty]{} 0.$$

Remarques

1. Attention ce théorème ne dit pas que la moyenne $\frac{X_1 + \cdots + X_n}{n}$ admet pour limite μ. Ce qu'il dit c'est qu'à un certain instant n, la probabilité d'être dans l'intervalle $[\mu - \varepsilon, \mu + \varepsilon]$ est un $O(1/n)$. Donc, si on suit l'évolution de la moyenne quand n varie, il est possible (et probable) que celle-ci sorte de l'intervalle.

2. On peut montrer, et c'est plus difficile[11], que sous certaines hypothèses simples[12], il est presque certain que la moyenne tende vers μ. C'est la *loi forte des grands nombres* qui dit que :

$$\mathbb{P}\left(\frac{X_1 + \cdots + X_n}{n} \xrightarrow[n \longrightarrow +\infty]{} \mu\right) = 1$$

[11] C'est un résultat dû à Kolmogorov. En prenant des hypothèses simples : même loi, moments centrés d'ordre 4, on arrive à un théorème pas trop douloureux (i.e. faisable en exercice) à démontrer...

[12] Indépendantes, de même loi et ayant une espérance. Donc moins que pour le théorème 25.