Compléments de Mathématiques Discrètes: Cours 6 .

Michaël Quisquater (Maître de Conférences)

UVSQ

Somme de variables aléatoires et indépendance stochastique de variables aléatoires.

Somme de variables aléatoires

Supposons que 10 étudiants passent deux examens et obtiennent donc chacun deux notes.

Ω est donc par exemple

$$\{(8,10),(9,9),(6,7),(5,8),(2,10),(8,4),(7,5),(8,5),(7,7),(8,6)\}.$$

On suppose que chaque étudiant (et donc son couple de notes) a la même probabilité d'être tiré au hasard.

On souhaite déterminer la distribution de probabilité de la somme des notes de chaque étudiant.

Par conséquent, pour tout $w \in \Omega$, $P(w) = 1/10$.

Somme de variables aléatoires
Variables aléatoires stochastiquement indépendantes
Loi (et distribution) de probabilité de la somme de v.a discrètes
Retour sur les propriétés de l'espérance et de la variance
Fonction génératrice des probabilités
Exemple introductif
Formalisation et probématique

Somme de variables aléatoires (suite)

Définissons la variable aléatoire $X : \Omega \to \mathbb{R} : w = (w_1, w_2) \mapsto X(w) = w_1$ qui représente la note du premier examen de chaque étudiant.

Définissons la variable aléatoire $Y : \Omega \to \mathbb{R} : w = (w_1, w_2) \mapsto Y(w) = w_2$ qui représente la note du premier examen de chaque étudiant.

La variable aléatoire $Z : \Omega \to \mathbb{R} : w \mapsto Z(w) = X(w) + Y(w)$ représente donc la somme des notes des deux examens.

On recherche la distribution de probabilité de Z (de laquelle on peut déduire la loi de probabilité), i.e. $P(Z = z)$ pour les différents z possibles.

Somme de variables aléatoires (suite)

On a :

$$\begin{aligned}
P(Z=12) &= P(\{(2,10),(8,4),(7,5)\}) \\
&= P(\{(2,10)\})+P(\{(8,4)\})+P(\{(7,5)\}) \\
&= 1/10+1/10+1/10 \\
&= 3/10
\end{aligned}$$

$$\begin{aligned}
P(Z=13) &= P(\{(6,7),(5,8),(8,5)\}) \\
&= P(\{(6,7)\})+P(\{(5,8)\})+P(\{(8,5)\}) \\
&= 1/10+1/10+1/10 \\
&= 3/10
\end{aligned}$$

Somme de variables aléatoires (suite)

$$\begin{aligned}
P(Z=14) &= P(\{(7,7),(8,6)\}) \\
&= P(\{(7,7)\})+P(\{(8,6)\}) \\
&= 1/10+1/10 \\
&= 2/10
\end{aligned}$$

$$\begin{aligned}
P(Z=18) &= P(\{(8,10),(9,9)\}) \\
&= P(\{(8,10)\})+P(\{(9,9)\}) \\
&= 1/10+1/10 \\
&= 2/10
\end{aligned}$$

P(Z=z)=0 pour les autres valeurs de z.

Somme de variables aléatoires (suite)

Plus généralement,

- Soit $(\Omega, \mathcal{P}(\Omega), P)$ un espace probabilisé.
- Soit $X : \Omega \to \mathbb{R}$ et $Y : \Omega \to \mathbb{R}$ deux variables aléatoires.
- Soit $Z : \Omega \to \mathbb{R} : w \mapsto Z(w) = X(w) + Y(w)$

Question : Est-il possible de déduire la distribution de probabilité de Z à partir de celles de X et Y. Y a-t-il des conditions pour cela ?

Variables aléatoires stochastiquement indépendantes

Supposons que l'on joue à pile ou face et que l'on réalise deux lancés.

$\Omega = \{(pile, pile), (pile, face), (face, pile), (face, face)\}$.

On suppose que tous ces résultats sont équiprobables, i.e. $P(w) = 1/4$ pour tout $w \in \Omega$.

Définissons la variable aléatoire $X : \Omega \to \mathbb{R} : w \mapsto X(w)$ qui vaut 1 si le premier lancé est pile et vaut 0 sinon.

Définissons la variable aléatoire $Y : \Omega \to \mathbb{R} : w \mapsto Y(w)$ qui vaut 1 si le second lancé est pile et vaut 0 sinon.

Variables aléatoires stochastiquement indépendantes

Calculons les distributions de probabilité de X et Y.

Nous avons :

- $P(X = 0) = P(\{(face, face), (face, pile)\}) = 1/4 + 1/4 = 1/2$ et $P(X = 1) = 1 - P(X = 0) = 1/2$
- $P(Y = 0) = P(\{(face, face), (pile, face)\}) = 1/4 + 1/4 = 1/2$ et $P(Y = 1) = 1 - P(Y = 0) = 1/2$

Aussi,

- $P(X = 0, Y = 0) = P(\{(face, face)\}) = 1/4$
- $P(X = 0, Y = 1) = P(\{(face, pile)\}) = 1/4$
- $P(X = 1, Y = 0) = P(\{(pile, face)\}) = 1/4$
- $P(X = 1, Y = 1) = P(\{(pile, pile)\}) = 1/4$

V.a. stochastiquement indépendantes (suite)

Par conséquent,

- $P(X = 0, Y = 0) = P(X = 0) \cdot P(Y = 0)$,
- $P(X = 0, Y = 1) = P(X = 0) \cdot P(Y = 1)$,
- $P(X = 1, Y = 0) = P(X = 1) \cdot P(Y = 0)$,
- $P(X = 1, Y = 1) = P(X = 1) \cdot P(Y = 1)$.

Par conséquent, X et Y sont stochastiquement indépendantes ; la connaissance d'une des variables ne modifie pas l'information probabiliste que l'on a de l'autre

$$P(X = x \mid Y = y) = P(X = x)\,.$$

Variables aléatoires stochastiquement indépendantes

Définition

Soit $(\Omega, \mathcal{P}(\Omega), P)$ un espace probabilisé. Soit $X : \Omega \to \mathbb{R}$ et $Y : \Omega \to \mathbb{R}$ deux variables aléatoires.
Alors, X et Y sont dites stochastiquement indépendantes si

$$P(X \in B_1, Y \in B_2) = P(X \in B_1) \cdot P(Y \in B_2) \qquad (1)$$

pour tout sous-ensemble B_1 et B_2 de $\mathbb{R}$ (à mieux préciser).

Intuitivement, cela signifie que n'importe quelle information sur une des variables ne donne aucune information probabiliste sur la seconde.

Variables aléatoires stochastiquement indépendantes (suite)

Remarque : Dénotons par S_1 et S_2 l'ensemble de valeurs que prennent les variables aléatoires discrètes X et Y respectivement, i.e.

$$S_1 = \{X(w) \mid w \in \Omega\} \text{ et } S_2 = \{Y(w) \mid w \in \Omega\}.$$

La définition précédente (1) est équivalente à la condition plus simple :

$$P(X = x, Y = y) = P(X = x) \cdot P(Y = y)$$

pour tout $x \in S_1$ et pour tout $y \in S_2$.

Somme de variables aléatoires indépendantes discrètes

Dans le cas de somme de variables aléatoires stochastiquement indépendantes, il est possible d'exprimer

la distribution de la somme de ces variables

en fonction de

la distribution de chacune de ces variables.

Somme de variables aléatoires indépendantes discrètes

Théorème

Soit $(\Omega, \mathcal{P}(\Omega), P)$ un espace probabilisé et soient les variables aléatoires discrètes $X, Y : \Omega \to \mathbb{R}$. Dénotons S_1 et S_2 l'ensemble des valeurs de X pour lesquels $P(X = x) > 0$ et $P(Y = y) > 0$ respectivement. De plus X et Y sont supposées stochastiquement indépendantes. Posons $Z : \Omega \to \mathbb{R} : w \mapsto X(w) + Y(w)$, alors pour tout $z \in \mathbb{R}$,

$$P(Z = z) = \sum_{x \in S_1} \left(P(Y = z - x) \cdot P(X = x) \right) .$$

Somme de variables aléatoires indépendantes discrètes (preuve)

Preuve : Par les probabilités totales,

$$\begin{aligned} P(Z=z) &= \textstyle\sum_{x\in S_1} \left(P(Z=z \mid X=x)\cdot P(X=x)\right) \\ &= \textstyle\sum_{x\in S_1} \left(P(X+Y=z \mid X=x)\cdot P(X=x)\right) \\ &= \textstyle\sum_{x\in S_1} \left(P(x+Y=z \mid X=x)\cdot P(X=x)\right) \\ &= \textstyle\sum_{x\in S_1} \left(P(Y=z-x \mid X=x)\cdot P(X=x)\right) \end{aligned}$$

Par l'indépendance de X et Y,

$$P(Y=z-x \mid X=x) = P(Y=z-x)\,.$$

Par conséquent,

$$P(Z=z) = \sum_{x\in S_1} \left(P(Y=z-x)\cdot P(X=x)\right).$$

Introduction

Si on souhaite calculer l'espérance ou la variance de la somme (resp. produit) de deux v.a. aléatoires,

- la méthode naturelle est de d'abord déterminer la distribution de probabilité de cette somme (resp. ce produit) et d'ensuite calculer l'espérance ou la variance au moyen de cette distribution.
- une autre méthode, plus simple, se base sur les théorèmes suivants.

Propriétés de l'espérance

Théorème

Soit X et Y des variables aléatoires discrètes.
Alors,

- $E(X + Y) = E(X) + E(Y)$
- *si X et Y sont stochastiquement indépendantes :* $E(X \cdot Y) = E(X) \cdot E(Y)$.

Somme de variables aléatoires
Variables aléatoires stochastiquement indépendantes
Loi (et distribution) de probabilité de la somme de v.a discrètes
Retour sur les propriétés de l'espérance et de la variance
Fonction génératrice des probabilités
Introduction
Propriétés de l'espérance
Propriétés de la variance

Propriétés de la variance

Théorème

Soit $X, Y : \Omega \to \mathbb{R}$ deux variables aléatoires discrètes.

- *si X et Y sont indépendantes,* $var(X + Y) = var(X) + var(Y)$

Pourquoi introduire l'outil des fonctions génératrices des probabilités ?

Soit X et Y deux variables aléatoires discrètes stochastiquement indépendantes de distributions de probabilité $P(X = x)$ et $P(Y = y)$ avec $x \in S_1$ et $y \in S_2$.

Question1 : Comment calculer la distribution de probabilité de la variable aléatoire $Z = X + Y$?

Réponse1 :

$$P(Z = z) = \sum_{x \in S_1} P(X = x) \cdot P(Y = z - x) \rightarrow \text{compliqué !}$$

Somme de variables aléatoires
Variables aléatoires stochastiquement indépendantes
Loi (et distribution) de probabilité de la somme de v.a discrètes
Retour sur les propriétés de l'espérance et de la variance
Fonction génératrice des probabilités
Motivation
Définition de fonction génératrices des probabilités
Exemple introductif : fgp et somme de v.a's discrètes finies

Idée ?

Observation :

Considérons les polynômes complexes

$$A(z) = a_0 + a_1 z + a_2 z^2 \cdots + a_n z^n$$

et

$$B(z) = b_0 + b_1 z + b_2 z^2 + \cdots + b_m z^m .$$

Considérons le produit de ces polynômes :

$$C(z) = A(z) \cdot B(z) = c_0 + c_1 z + \cdots c_{n+m} z^{n+m} ,$$

avec

$$c_j = \sum_{i=0}^{j} a_i b_{j-i}$$

Idée ? (suite)

Idée :

$\Rightarrow$ cette formule ressemble beaucoup à la formule permettant de calculer la distribution de somme de variables aléatoires indépendantes.

Question :

Calculer la distribution de la somme de v.a's indépendantes reviendrait-il à calculer le produit de polynômes particuliers ?

Introduction

Nous allons considérer un polynôme particulier :

Définition (Fonction génératrice des probabilités d'une v.a. discrète finie)

Soit X une variable aléatoire à valeur dans $\{0, \cdots, n\}$ de loi de probabilité $P(X = x)$ pour $x \in \{0, \cdots, n\}$. On appelle fonction génératrice des probabilités (fgp) de X, notée $G_X(z)$, le polynôme complexe défini par

$$G_X(z) = E(z^X) = \sum_{x=0}^{n} P(X = x) z^x \,.$$

où $E(\cdot)$ est l'espérance.

Introduction

Remarques :

- A chaque distribution de probabilité correspond une et une seule fonction génératrice !
- La définition peut être étendue à des expressions qui ne sont plus des polynômes (série ou expression avec des exposants négatifs), i.e.

$$G_X(z) = E(z^X) = \sum_{x \in S} P(X = x) z^x$$

si S désigne l'ensemble des valeurs prises par la v.a. X.

Somme de variables aléatoires
Variables aléatoires stochastiquement indépendantes
Loi (et distribution) de probabilité de la somme de v.a discrètes
Retour sur les propriétés de l'espérance et de la variance
Fonction génératrice des probabilités
Motivation
Définition de fonction génératrices des probabilités
Exemple introductif : fgp et somme de v.a's discrètes finies

Exemple :

Soit X_i des v.a's indépendantes de loi Bernoulli de paramètre p, i.e. $P(X_i = 0) = 1 - p$ et $P(X_i = 1) = p$

$$X_i \sim Bernoulli(p) .$$

Considérons les fonctions génératrices des probabilités de ces variables aléatoires :

$$G_{X_i}(z) = E(z^{X_i}) = \sum_{x=0}^{1} P(X_i = x) z^x = (1 - p) + pz .$$

Considérons le produit de ces polynômes :

$$\prod_{i=1}^{n} G_{X_i}(z) = ((1 - p) + pz)^n .$$

Exemple (suite) :

Par le binôme de Newton, on a :

$$((1-p)+pz)^n = \sum_{x=0}^{n} C_n^x (pz)^x (1-p)^{n-x} = \sum_{x=0}^{n} (C_n^x p^x (1-p)^{n-x}) z^x \, .$$

Par conséquent,

$$\prod_{i=1}^{n} G_{X_i}(z) = G_X(z)$$

où X est une variable aléatoire binomiale de paramètre p et n.

Introduction

Question ?

Nous venons de voir que le produit de fonctions génératrices des probabilités de v.a's Bernoulli de paramètre p correspond à la fonction génératrice des probabilités d'une variable aléatoire binomiale.

Est-il vrai en général que le produit de fonctions génératrices des probabilités de variables aléatoires indépendantes est la fonction génératrice des probabilités de la somme de ces variables aléatoires ?

Fonctions génératrices des probabilités et somme de v.a's indépendantes discrètes finies

Théorème

Soient X_i des variables aléatoires indépendantes à valeurs dans des ensembles finis et $G_{X_i}(z)$ leurs fonctions génératrices des probabilités. Alors, si $X = \sum_{i=1}^{n} X_i$ et $G_X(z)$ est la fgp de X, on a

$$\prod_{i=1}^{n} G_{X_i}(z) = G_X(z) \text{ où } X = \sum_{i=1}^{n} X_i\,.$$

Fonctions génératrices des probabilités et somme de v.a's indépendantes discrètes finies

Preuve : Il suffit d'observer que si X_i sont des v.a's indépendantes, alors z^{X_i} le sont également (à montrer !). Par conséquent, comme $E(X \cdot Y) = E(X) \cdot E(Y)$ pour X et Y indépendantes, on a :

$$\prod_{i=1}^{n} G_{X_i}(z) = \prod_{i=1}^{n} E(z^{X_i}) = E(\prod_{i=1}^{n} z^{X_i}) = E(z^{\sum_{i=1}^{n} X_i}) = G_X(z)\,.$$

$\square$

Conclusion

- Somme de variables aléatoires,
- Variables aléatoires stochastiquement indépendantes,
- Expression distribution de la somme de variables aléatoires stochastiquement indépendantes,
- Retour sur les propriétés de l'espérance et de la variance,
- Fonction génératrice des probabilités et somme de variables aléatoires stochastiquement indépendantes.