Vecteurs aléatoires — partie 1

À l'intention des étudiants de STT-1500



Jérôme Soucy

Département de mathématiques et de statistique
Université Laval

Hiver 2025

Résumé

1. Distributions multivariées
2. Lois marginales
3. Lois conditionnelles
4. Variables aléatoires indépendantes
5. Distribution d'une fonction de plusieurs variables aléatoires
6. Somme de variables aléatoires et convolution

1
Distributions multivariées

Définition

Définition

Définition. Un **vecteur aléatoire** de dimension n est un vecteur de dimension n dont toutes les composantes sont des variables aléatoires.

Nous allons surtout considérer le cas $n = 2$, mais les autres cas se traitent de la même façon.

Cas discret

Cas discret

Définition. Lorsque l'ensemble des valeurs possibles pour le couple (X, Y) est fini ou dénombrable, la **distribution conjointe** des variables aléatoires X et Y est caractérisée par la **fonction de masse conjointe** $p(x, y)$ définie par

$$p(x, y) = \mathbb{P}((X, Y) = (x, y)) = \mathbb{P}(\{X = x\} \cap \{Y = y\}).$$

Remarque. La fonction p est toujours positive, et on a que

$$\sum_{(x,y)\in\mathbb{R}^2} p(x, y) = 1.$$

Cas absolument continu

Cas absolument continu

Définition. On dit que la distribution conjointe des variables aléatoires X et Y est **absolument continue** s'il existe une fonction f, appelée **fonction de densité**, vérifiant

$$\mathbb{P}(\{a < X \leq b\} \cap \{c < Y \leq d\}) = \int_c^d \int_a^b f(x, y) dx dy$$

pour tout choix de $a \leq b$ et $c \leq d$.

Cas absolument continu - suite

Cas absolument continu - suite

La fonction f satisfait ainsi

$$\mathbb{P}((X,Y) \in B) = \int \int_B f(x,y)dxdy$$

pour tout sous-ensemble B de $\mathbb{R}^2$ pour lequel l'intégrale est définie.

Cas absolument continu - suite

La fonction f satisfait ainsi

$$\mathbb{P}((X,Y) \in B) = \int \int_B f(x,y)dxdy$$

pour tout sous-ensemble B de $\mathbb{R}^2$ pour lequel l'intégrale est définie. Une fonction de densité conjointe de deux variables aléatoires absolument continues possède des propriétés analogues à la fonction de densité d'une variable aléatoire. Nous avons que

1. $$\int \int_{\mathbb{R}^2} f(x,y)dxdy = 1;$$
2. $$f(x,y) \geq 0 \ \forall (x,y) \in \mathbb{R}^2 .$$

Fonction de répartition

Fonction de répartition

Définition. La **fonction de répartition conjointe** de deux variables est définie par

$$F(x, y) = \mathbb{P}(\{X \leq x\} \cap \{Y \leq y\}).$$

Dans le cas absolument continu, nous avons que

$$F(x, y) = \int_{-\infty}^{y} \int_{-\infty}^{x} f(u, v) du dv = \int_{-\infty}^{x} \int_{-\infty}^{y} f(u, v) dv du;$$
$$f(x, y) = \frac{\partial^2 F}{\partial x \partial y} = \frac{\partial^2 F}{\partial y \partial x}.$$

Pour démontrer ces résultats, il faut utiliser le théorème fondamental du calcul différentiel et le théorème de Fubini.

Fonction de répartition - suite

Fonction de répartition - suite

Proposition. Soit X et Y des variables aléatoires dont la fonction de répartition conjointe est F. Alors

$$\mathbb{P}(\{a < X \leq b\} \cap \{c < Y \leq d\}) = F(b,d) - F(a,d) - F(b,c) + F(a,c)$$

pour tout choix de $a \leq b$ et $c \leq d$.

2

Lois marginales

Fonctions de masse marginales

Fonctions de masse marginales

Définition. Soit X et Y des variables aléatoires dont la fonction de masse conjointe est p. Les **fonctions de masse marginales** de X et Y sont données respectivement par p_X et p_Y, définies comme suit :

$$p_X(x) = \sum_{y \in \mathbb{R}} p(x, y) \qquad \text{et} \qquad p_Y(y) = \sum_{x \in \mathbb{R}} p(x, y).$$

Un exemple

On lance deux dés réguliers à 6 faces. On définit les v.a. suivantes :

$$X = \text{Le minimum des deux nombres obtenus ;}$$
$$X = \text{La somme des deux nombres obtenus.}$$

Représentation d'une fonction de masse sous forme de tableau

$X\backslash Y$	2	3	4	5	6	7	8	9	10	11	12	$P_X(x)$
1	$1/36$	$2/36$	$2/36$	$2/36$	$2/36$	$2/36$						$11/36$
2			$1/36$	$2/36$	$2/36$	$2/36$	$2/36$					$9/36$
3					$1/36$	$2/36$	$2/36$	$2/36$				$7/36$
4							$1/36$	$2/36$	$2/36$			$5/36$
5									$1/36$	$2/36$		$3/36$
6											$1/36$	$1/36$
$P_Y(y)$	$1/36$	$2/36$	$3/36$	$4/36$	$5/36$	$6/36$	$5/36$	$4/36$	$3/36$	$2/36$	$1/36$	

Tab. : Fonction de masse conjointe des variables X et Y

Fonctions de densité marginales

Fonctions de densité marginales

Définition. Soit X et Y des variables aléatoires dont la fonction de densité conjointe est f. Alors les **fonctions de densité marginales** de X et Y sont données respectivement par les fonctions f_X et f_Y, définies ci-dessous :

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy \qquad \text{et} \qquad f_Y(y) = \int_{-\infty}^{\infty} f(x,y)dx.$$

3

Lois conditionnelles

Définitions

Il arrive parfois que nous connaissons la fonction de masse ou la fonction de densité conjointe de deux variables aléatoires, et qu'en plus de cela, nous connaissons aussi la valeur prise par l'une des deux variables aléatoires. Avec cette nouvelle information, on peut se demander quelle est la fonction de masse (ou la fonction de densité) conditionnelle de l'autre variable aléatoire.

Définition. Soit X et Y des variables aléatoires dont la fonction de masse conjointe est p, et soit $x \in \mathbb{R}$ tel que $\mathbb{P}(X = x) > 0$. La **fonction de masse conditionnelle** de Y sachant que $X = x$ est définie par l'équation

$$p_{Y|X=x}(y) = \frac{p(x, y)}{p_X(x)}.$$

Définition - suite

Définition. Soit X et Y des variables aléatoires dont la fonction de densité conjointe est f. Alors la **fonction de densité conditionnelle** de Y sachant que $X = x$ est donnée par l'équation

$$f_{Y|X=x}(y) = \frac{f(x,y)}{f_X(x)}.$$

La fonction de densité conditionnelle de X sachant que $Y = y$ est définie de façon analogue.

Fonction de répartition

Définition. Soit X et Y des variables aléatoires, et soit $x \in \mathbb{R}$. La **fonction de répartition conditionnelle** de Y sachant que $X = x$ est définie par l'équation

$$F_{Y|X=x}(y) = \mathbb{P}(Y \leq y | X = x).$$

La fonction de répartition conditionnelle de X sachant que $Y = y$ est définie de façon analogue.

Pour que cette définition soit valable dans le cas discret, il faut que $\mathbb{P}(X = x) > 0$. Dans le cas continu, il faut que $f_X(x) > 0$.

4

Variables aléatoires indépendantes

Définition

Définition. Les variables aléatoires X et Y sont **indépendantes** si pour tout choix de sous-ensembles de $\mathbb{R}$, disons C et D, les événements $\{X \in C\}$ et $\{Y \in D\}$ sont indépendants.

Au regard de la définition de l'indépendance de deux événements, deux variables aléatoires sont indépendantes si, pour tout choix de sous-ensembles de C et D, nous avons que

$$\mathbb{P}(\{X \in C\} \cap \{Y \in D\}) = \mathbb{P}(X \in C) \cdot \mathbb{P}(Y \in D).$$

Résultat

Proposition. Les variables aléatoires X et Y sont indépendantes si et seulement si pour tout choix de $a \leq b$ et $c \leq d$, on a que

$$\mathbb{P}(\{a < X \leq b\} \cap \{c < Y \leq d\}) = \mathbb{P}(a < X \leq b) \cdot \mathbb{P}(c < Y \leq d).$$

Puisque cette proposition donne une caractérisation des variables indépendantes, elle pourrait servir à *définir* l'indépendance de deux variables aléatoires.

Théorèmes

Théorème. Si X et Y sont deux variables aléatoires indépendantes, et si h et g sont des fonctions réelles d'une variable réelle, alors les variables aléatoires $h(X)$ et $g(Y)$ sont indépendantes.

Théorème. Deux variables aléatoires X et Y sont indépendantes si et seulement si la fonction de répartition conjointe F est égale au produit des fonctions de répartition marginales F_X et F_Y.

Théorèmes

Théorème. Deux variables aléatoires absolument continues sont indépendantes si et seulement si leur fonction de densité conjointe est égale au produit des fonctions de densité marginales de ces deux variables.

Théorème. Deux variables aléatoires discrètes sont indépendantes si et seulement si leur fonction de masse conjointe est égale au produit des fonctions de masse marginales de ces deux variables.

Retour sur l'exemple du lancer de deux dés réguliers

Rappelons que X correspond au minimum des deux nombres obtenus alors que Y est la somme des deux nombres obtenus.

$X \backslash Y$	2	3	4	5	6	7	8	9	10	11	12	$P_X(x)$
1	$1/36$	$2/36$	$2/36$	$2/36$	$2/36$	$2/36$						$11/36$
2			$1/36$	$2/36$	$2/36$	$2/36$	$2/36$					$9/36$
3					$1/36$	$2/36$	$2/36$	$2/36$				$7/36$
4							$1/36$	$2/36$	$2/36$			$5/36$
5									$1/36$	$2/36$		$3/36$
6											$1/36$	$1/36$
$P_Y(y)$	$1/36$	$2/36$	$3/36$	$4/36$	$5/36$	$6/36$	$5/36$	$4/36$	$3/36$	$2/36$	$1/36$	

Tab. : Fonction de masse conjointe des variables X et Y

Est-ce que X et Y sont indépendantes ?

Pour que X et Y soient indépendantes, on doit avoir que

$$p(x, y) = p_X(x)p_Y(y) \qquad \forall (x, y) \in \mathbb{R}^2 .$$

Selon le tableau de la page précédente, nous avons que :

$$\begin{aligned} p(2, 6) &= {}^2\!/_{36}, \\ p_X(2) &= {}^9\!/_{36}, \\ p_Y(6) &= {}^5\!/_{36}. \end{aligned}$$

Comme $p(2, 6) \neq p_X(2)p_Y(6)$, X et Y ne sont pas indépendantes.

L'indépendance et les distributions conditionnelles

Remarque. On peut montrer que les variables aléatoires X et Y sont indépendantes si et seulement si les distributions conditionnelles de l'une sachant l'autre sont égales aux distributions marginales correspondantes. Ainsi, dans le cas absolument continu, nous avons que

$$f_{X,Y}(x,y) = f_X(x) \times f_Y(y)$$

si et seulement si

$$f_{Y|X=x}(y) = f_Y(y) \text{ et } f_{X|Y=y}(x) = f_X(x).$$

5

Distribution d’une fonction de plusieurs variables aléatoires

Contexte

Soit $g : \mathbb{R}^2 \to \mathbb{R}$ une fonction, et soit X et Y des variables aléatoires dont la distribution conjointe est connue. On veut déterminer la distribution de la variable aléatoire W, définie par $W = g(X, Y)$.

Cas discret

1. On détermine l'ensemble des valeurs possibles pour W ;
2. Pour chaque valeur de w on calcule la fonction de masse $p_W(w) = \mathbb{P}(W = w)$ en observant que

$$p_W(w) = \sum_{\{(x,y)\in\mathbb{R}^2 : g(x,y)=w\}} p(x,y),$$

ou p est la fonction de masse conjointe de X et Y.

Cas absolument continu

1. On détermine l'ensemble des valeurs possibles pour W ;
2. On calcule la fonction de répartition de W en observant que

$$F_W(w) = \int \int_{\{(x,y)\in\mathbb{R}^2 : g(x,y)\leq w\}} f(x,y)dxdy;$$

3. On dérive $F_W(w)$ par rapport à w pour obtenir la fonction de densité.

6

Somme de variables aléatoires et convolution

Plusieurs problèmes nécessitent de répondre à la question suivante : Soient X et Y des variables aléatoires indépendantes dont la distribution est connue. Quelle est la distribution de la variable aléatoire T définie par $T = X + Y$? Dans le cas où les variables X et Y sont à valeurs entières non négatives, on montre que la fonction de masse de T satisfait

$$p_T(k) = \begin{cases} \sum_{i=0}^{k} p_X(i) p_Y(k-i) & \text{si } k \in \mathbb{N}, \\ 0 & \text{sinon.} \end{cases}$$

Dans le cas où les variables X et Y sont absolument continues, on montre que la fonction de densité de T satisfait

$$f_T(t) = \int_{-\infty}^{\infty} f_X(x) f_Y(t-x) dx.$$

La dernière intégrale s'appelle la **convolution** des fonctions f_X et f_Y. De plus, si les variables X et Y sont à valeurs non négatives, nous avons que

$$f_T(t) = \begin{cases} \int_0^t f_X(x) f_Y(t-x) dx & \text{si } t \geq 0, \\ 0 & \text{sinon.} \end{cases}$$

Remarques

Remarques.

- Ce document est rédigé avec Beamer, une classe de LaTeX.

Pour signaler une erreur dans ce document, veuillez écrire à jerome.soucy@mat.ulaval.ca.