

Probabilités, Statistiques, Combinatoire

Philippe Duchon

Université de Bordeaux – Licence Informatique

<http://www.labri.fr/perso/duchon/Enseignements/Probas/>

2019-2020

Programme du jour

- ▶ On a une grande variété de lois de probabilités pour des variables aléatoires.
- ▶ Certaines lois sont bien connues, mais on ne “tombe” pas forcément dessus quand on modélise une situation.
- ▶ On utilise des “résumés” pour décrire de manière sommaire (et incomplète) les lois de variables aléatoires.
- ▶ **Espérance** : décrit l'ordre de grandeur des valeurs (c'est une moyenne)
- ▶ **Variance** : décrit la “dispersion” par rapport à la moyenne

Espérance

- ▶ **Définition** : X variable aléatoire à valeurs réelles, V son ensemble de valeurs :

$$\mathbf{E}(X) = \sum_{v \in V} v \mathbb{P}(X = v)$$

- ▶ **Définition alternative (théorème)** : si X est définie comme fonction sur l'univers Ω ,

$$\mathbf{E}(X) = \sum_{t \in \Omega} X(t) \mathbb{P}(\{t\})$$

- ▶ (Il se peut que l'espérance ne soit pas définie mathématiquement, lorsque Ω est infini ; dans ce cours, on ne s'intéresse pas à ces questions, on fait comme si c'était toujours défini)

Quelques propriétés de l'espérance

- ▶ Si X est une variable aléatoire *positive* (i.e. $\mathbb{P}(X \geq 0) = 1$), alors $\mathbf{E}(X) \geq 0$
- ▶ **Linéarité** : $\mathbf{E}(X + Y) = \mathbf{E}(X) + \mathbf{E}(Y)$ (**sans condition** sur X et Y , autre que l'existence des espérances) ;
- ▶ **Linéarité (bis)** : $\mathbf{E}(aX) = a\mathbf{E}(X)$ (pour a constante)
- ▶ **Attention** : en règle générale, il n'y a pas de raison pour que $\mathbf{E}(f(X))$ soit égale à $f(\mathbf{E}(X))$ si f n'est pas une fonction affine.
- ▶ En particulier, $\mathbf{E}(X^2) > (\mathbf{E}(X))^2$, sauf si X est une variable aléatoire constante.
- ▶ **Si** X et Y sont des variables aléatoires **indépendantes**, alors

$$\mathbf{E}(X.Y) = \mathbf{E}(X).\mathbf{E}(Y).$$

Espérances des lois usuelles

Loi	param.	conditions	espérance
Bernoulli	p	$0 < p < 1$	p
Binomiale	(n, p)	$n \in \mathbb{N}^*, 0 < p < 1$	$n.p$
Géométrique	p	$0 < p < 1$	$1/p$
Poisson	x	$x > 0$	x

- ▶ **Bernoulli** : calcul facile
- ▶ **Binomiale** : Bernoulli et linéarité (somme de Bernoulli)
- ▶ **Géométrique, Poisson** : cf. calculs en TD (manipulations de sommes)

Espérance de $f(X)$

- ▶ **Situation** : on a une variable aléatoire X , dont on connaît la loi ; et une fonction $f : \mathbb{R} \rightarrow \mathbb{R}$.
- ▶ On veut calculer $\mathbf{E}(f(X))$.
- ▶ Cas particulier : si f est de la forme $f(x) = ax + b$ (a et b constantes), $\mathbf{E}(aX + b) = a\mathbf{E}(X) + b$
- ▶ Sinon, il faut poser le calcul :

$$\mathbf{E}(f(X)) = \sum_{t \in \Omega} f(X(t))\mathbb{P}(\{t\})$$

- ▶ En regroupant les termes selon la valeur de $X(t)$, on obtient, en posant V pour l'ensemble des valeurs possibles de X :

$$\mathbf{E}(f(X)) = \sum_{v \in V} f(v)\mathbb{P}(X = v)$$

Un petit exemple de calcul

- ▶ X : tirage d'un dé, moins 3 ; valeurs $\{-2, -1, 0, 1, 2, 3\}$
- ▶ $Z = X^2$: valeurs $\{0, 1, 4, 9\}$
- ▶ On peut calculer la loi de Z :
 - ▶ $\mathbb{P}(Z = 0) = \mathbb{P}(Z = 9) = 1/6$
 - ▶ $\mathbb{P}(Z = 1) = \mathbb{P}(Z = 4) = 2/6$
- ▶ On en déduit le calcul de $\mathbf{E}(Z)$:

$$\mathbf{E}(Z) = \frac{0}{6} + \frac{9}{6} + \frac{1 \times 2}{6} + \frac{4 \times 2}{6} = \frac{19}{6}$$

- ▶ On a aussi l'autre formule :

$$\mathbf{E}(Z) = \frac{4 + 1 + 0 + 1 + 4 + 9}{6} = \frac{19}{6}$$

Un exemple de calcul

P une variable de Poisson de paramètre x , $Z = P^2$.

- ▶ On connaît la loi de P : $\mathbb{P}(P = k) = e^{-x} \frac{x^k}{k!}$ pour $k \in \mathbb{N}$.
- ▶ On a donc une formule pour $\mathbf{E}(Z)$:

$$\mathbf{E}(Z) = \sum_{k=1}^{+\infty} k^2 e^{-x} \frac{x^k}{k!}.$$

- ▶ Manipulation de la somme : en écrivant $k^2 = k(k-1) + k$, et en développant la somme, on obtient

$$\mathbf{E}(Z) = \sum_{k=0}^{\infty} e^{-x} \frac{k(k-1)x^k}{k!} + \sum_{k=0}^{\infty} e^{-x} \frac{kx^k}{k!},$$

- ▶ et donc (simplification) $\mathbf{E}(Z) = x^2 + x$.

Variance

- ▶ La **variance** d'une variable aléatoire est une “mesure” de sa tendance à dévier de son espérance.
- ▶ **Définition** : si X est une variable aléatoire, et $m = \mathbf{E}(X)$ est son espérance, alors on définit la **variance** de X , notée $\mathbf{Var}(X)$, comme étant

$$\mathbf{Var}(X) = \mathbf{E}((X - m)^2)$$

- ▶ **Remarque** : d'après la définition, on a forcément $\mathbf{Var}(X) \geq 0$.
- ▶ **Proposition** : la formule suivante est également valable :

$$\mathbf{Var}(X) = \mathbf{E}(X^2) - (\mathbf{E}(X))^2.$$

- ▶ (C'est souvent cette formule qui est utilisée pour un calcul pratique ; mais par exemple, il n'est pas aisé sur cette formule de voir que la variance ne peut pas être négative)

Variances classiques

- ▶ Variable de Bernoulli : $B^2 = B$ donc $\mathbf{Var}(B) = p - p^2$.
- ▶ Variable de Poisson : on a vu que $\mathbf{E}(P^2) = x^2 + x$, or $\mathbf{E}(P) = x$, donc $\mathbf{Var}(P) = x$.
- ▶ Variable géométrique : le calcul de $\mathbf{E}(G^2)$ est du même acabit ; on va admettre que $\mathbf{E}(G^2) = \frac{2-p}{p^2}$, ce qui entraîne

$$\mathbf{Var}(G) = \frac{1-p}{p^2}$$

- ▶ Si a est une constante, $\mathbf{Var}(aX) = a^2\mathbf{Var}(X)$.

Variance d'une somme

- ▶ Soient X et Y deux variables aléatoires, on pose $m = \mathbf{E}(X)$ et $m' = \mathbf{E}(Y)$; essayons de calculer $\mathbf{Var}(X + Y)$.
- ▶ On va calculer $\mathbf{E}((X + Y)^2) - (\mathbf{E}(X + Y))^2$.
- ▶ Partie facile : $\mathbf{E}(X + Y) = m + m'$ donc $(\mathbf{E}(X + Y))^2 = m^2 + 2m.m' + m'^2$.
- ▶ Pour $\mathbf{E}((X + Y)^2)$, on développe : $(X + Y)^2 = X^2 + 2XY + Y^2$, donc par linéarité de l'espérance,

$$\mathbf{E}((X + Y)^2) = \mathbf{E}(X^2) + 2\mathbf{E}(XY) + \mathbf{E}(Y^2).$$

- ▶ Par soustraction :

$$\mathbf{Var}(X + Y) = \mathbf{Var}(X) + \mathbf{Var}(Y) + 2(\mathbf{E}(XY) - m.m')$$

- ▶ En règle générale, ça s'arrête là...
- ▶ **Si X et Y sont indépendantes**, on peut simplifier : $\mathbf{E}(XY) = m.m'$ et donc

$$\mathbf{Var}(X + Y) = \mathbf{Var}(X) + \mathbf{Var}(Y)$$

Variance d'une binomiale

- ▶ On généralise la formule de la variance à la somme de n variables indépendantes : **Si n variables $X_1, \dots, X_n$ sont indépendantes**, alors

$$\mathbf{Var}(X_1 + \dots + X_n) = \mathbf{Var}(X_1) + \dots + \mathbf{Var}(X_n)$$

- ▶ En particulier, en tenant compte de la description d'une binomiale comme somme de Bernoulli indépendantes, on obtient sa variance : $n.p.(1 - p)$.

Récapitulatif : espérances et variances

Loi	param.	espérance	variance
Bernoulli	p	p	$p(1 - p)$
Binomiale	(n, p)	$n.p$	$n.p.(1 - p)$
Géométrique	p	$1/p$	$(1 - p)/p^2$
Poisson	x	x	x

Inégalités basées sur espérance et variance

- ▶ **Idée générale** : trouver des propriétés de la forme, “la probabilité que telle variable aléatoire prenne une valeur [plus grande, plus petite] que [valeur] est petite”
- ▶ On va en voir deux, assez générales : l'**inégalité de Markov**, et l'**inégalité de Tchebycheff**.

Inégalité de Markov

- ▶ **Inégalité de Markov** : si X est une variable aléatoire à **valeurs positives**, d'espérance m , alors pour tout nombre a ,

$$\mathbb{P}(X \geq a.m) \leq \frac{1}{a}$$

- ▶ Autre formulation : $\mathbb{P}(X > b) \leq \frac{m}{b}$
- ▶ **Attention** : il faut que la variable soit positive ; ne marche pas si on n'a pas de telle garantie !
- ▶ (une v.a. positive a probabilité au plus 1/10 d'être 10 fois plus grande que son espérance)
- ▶ **Remarque** : l'inégalité ne dit rien d'intéressant pour $a \leq 1$.

Inégalité de Tchebycheff

- ▶ **Théorème** : si X est une variable aléatoire d'espérance m et de variance $v = \sigma^2$, alors on a, pour tout réel $x > 0$,

$$\mathbb{P}(|X - m| \geq x\sigma) \leq \frac{1}{x^2}.$$

- ▶ **Remarque** : la condition “ $|X - m| \geq x.\sigma$ ” correspond à “ $X \geq m + x.\sigma$ ou $X \leq m - x.\sigma$ ” : on regarde la probabilité d’être “loin de l’espérance”
- ▶ **Remarque** : pas de condition du type “ X positive”, mais en revance il faut connaître la variance.
- ▶ Cette inégalité va nous permettre de montrer la “loi des grands nombres”