Boîte à outils de probabilités discrètes

Bruno Grenet

1 Vocabulaire et définitions de base

Vocabulaire. Un *espace probabilisé* est une modélisation d'une expérience probabiliste : il est constitué d'un ensemble d'*évènements primitifs*, appelé l'*univers*, que l'on suppose fini ou dénombrable, et de leurs *probabilités* associées, telles que la somme des probabilités des évènements primitifs vaut 1. Un *évènement* est tout sous-ensemble de l'univers, c'est-à-dire un ensemble d'évènements primitifs.

Une *variable aléatoire (discrète)* est une fonction de l'ensemble des évènements primitifs dans un ensemble de valeurs. Si l'ensemble de valeurs est un sous-ensemble de $\mathbb{R}$, on parle de *variable aléatoire réelle*.

Définition 1.1. Si $X : \Omega \to V$ est une variable aléatoire à valeur dans V, on définit pour tout $v \in V$ l'évènement « $X = v$ » comme l'ensemble $\{\omega \in \Omega : X(\omega) = v\}$, c'est-à-dire qu'il s'agit de l'ensemble des évènements primitifs de valeur v. On définit de manière analogue les évènements « $X \neq v$ », « $X \geq v$ », « $X > v$ », etc.

Soit Ω un univers et $\Pr : \Omega \to \mathbb{R}_+$ la fonction qui associe à chaque évènement primitif sa probabilité. On étend Pr aux évènements E non primitifs en posant $\Pr[E] = \sum_{\omega \in E} \Pr[\omega]$. En particulier,

$$\Pr[X = v] = \sum_{\omega \in \Omega : X(\omega) = v} \Pr[\omega]$$

et de même pour $\Pr[X \geq v]$, etc.

Notations. Si E et F sont deux évènements, on note $E \wedge F$ leur intersection, c'est-à-dire l'ensemble $\{\omega \in \Omega : \omega \in E \wedge \omega \in F\}$ et $E \vee F$ leur union. On note de plus $\neg E$ le complémentaire de l'ensemble E, c'est-à-dire l'ensemble $\{\omega \in \Omega : \omega \notin E\}$.

Ces notations dans le langage de la logique correspondent aux notions intuitives : $\Pr[E \wedge F]$ est la probabilité que l'évènement E *et* l'évènement F se produisent ; $\Pr[E \vee F]$ est la probabilité que l'évènement E *ou* l'évènement F se produise ; $\Pr[\neg E]$ est la probabilité que l'évènement E *ne* se produise *pas*.

2 Espérance, variance, écart-type

Définition 2.1. L'*espérance* d'une variable aléatoire réelle discrète $X : \Omega \to V$ est définie par

$$\mathbb{E}[X] = \sum_{v \in V} v \times \Pr[X = v] = \sum_{\omega \in \Omega} X(\omega) \Pr[\omega].$$

Intuitivement, il s'agit de la moyenne pondérée des valeurs de tous les évènements possibles.

Définition 2.2. La *variance* d'une variable aléatoire réelle discrète X est

$$\mathrm{Var}(X) = \mathbb{E}\big[(X - \mathbb{E}[X])^2\big] = \mathbb{E}\big[X^2\big] - \mathbb{E}[X]^2.$$

Son *écart-type* $\sigma(X)$ est la racine carrée de la variance. Intuitivement, l'écart-type mesure la *moyenne (quadratique) des écarts à la moyenne*.

3 Probabilités conditionnelles et indépendance

Définition 3.1. La probabilité d'un évènement E *conditionnée à un évènement* F, ou *probabilité de* E *sachant* F, est

$$\Pr[E|F] = \frac{\Pr[E \wedge F]}{\Pr[F]}.$$

L'espérance d'une variable aléatoire $X : \Omega \to V$ *conditionnée à un évènement* F, ou *espérance de* X *sachant* F, est

$$\mathbb{E}[X|F] = \sum_{v \in V} v \times \Pr[X = v|F].$$

Théorème 3.2. *Si* E *et* F *sont deux évènements de probabilités non nulles,*

$$\Pr[E|F] = \frac{\Pr[F|E]\Pr[E]}{\Pr[F]}.$$

Définition 3.3. Deux évènements E et F sont dits indépendants si $\Pr[E \wedge F] = \Pr[E]\Pr[F]$.

Deux variables aléatoires $X : \Omega \to V$ et $Y : \Omega \to W$ sont dites indépendantes si pour tout $v \in V$ et $w \in W$, $\Pr[X = v \wedge Y = w] = \Pr[X = v]\Pr[Y = w]$.

Proposition 3.4. *Si* E *et* F *sont deux évènements de probabilités différentes de* 0 *et* 1*,* E *et* F *sont indépendants si et seulement si* $\Pr[E|F] = \Pr[E]$.

4 Quelques propriétés

Proposition 4.1. *Soit* $X, Y : \Omega \to V$ *des variables aléatoires réelles discrètes, et* E *et* F *deux évènements. Alors*

i. $\sum_{v \in V} \Pr[X = v] = 1$*;*
ii. $\Pr[\neg E] = 1 - \Pr[E]$*;*
iii. *Si* E *et* F *sont disjoints,* $\Pr[E \wedge F] = 0$*;*
iv. $\Pr[E \vee F] = \Pr[E] + \Pr[F] - \Pr[E \wedge F]$*;*
v. $\Pr[E \vee F] \leq \Pr[E] + \Pr[F]$ ***(inégalité de Boole)*****;**
vi. $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$ ***(linéarité de l'espérance)*****;**
vii. *Si* X *et* Y *sont indépendantes,* $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$ *et* $\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$*;*
viii. *Si* $V \subset \mathbb{N}$*,* $\mathbb{E}[X] = \sum_{v \geq 1} \Pr[X \geq v]$.

Proposition 4.2. *Soit* Ω *un espace probabilisé, et* $E_1 \sqcup E_2 \sqcup \cdots \sqcup E_n$ *une partition de* Ω*. Alors*

i. $\Pr[F] = \sum_{i=1}^{n} \Pr[F|E_i]\Pr[E_i]$ *pour tout évènement* F ***(formule des probabilités totales)*****;**
ii. $\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X|E_i]\Pr[E_i]$ *pour toute variable aléatoire* X ***(formule de l'espérance totale)*****.**

5 Inégalités

Inégalité de Markov. *Soit* X *une variable aléatoire discrète à valeur positives ou nulles. Alors*

$$\Pr[X \geq t] \leq \frac{\mathbb{E}[X]}{t} \textit{ et } \Pr[X \geq \lambda\mathbb{E}[X]] \leq \frac{1}{\lambda}.$$

Exercice. Pour doit-on supposer que X est à valeurs positives ou nulles ?

Inégalité de Tchebycheff. *Soit X une varaiable aléatoire et t un réel strictement positif. Alors*

$$\Pr[|X - \mathbb{E}[X]| \geq t] \leq \frac{\sigma(X)^2}{t^2}.$$

Inégalités de Chernoff. *Soit $X_1, \ldots, X_n$ des variables aléatoires indépendantes à valeur dans $\{0,1\}$, telles que $\Pr[X_i = 1] = p_i \in]0,1[$ pour tout i, et $X = X_1 + \cdots + X_n$. Alors*

i. $\Pr[X > (1+\delta)\mathbb{E}[X]] \leq e^{-\mathbb{E}[X]\delta^2/(2+\delta)}$ *pour tout $\delta > 0$, et*
ii. $\Pr[X < (1-\delta)\mathbb{E}[X]] \leq e^{-\mathbb{E}[X]\delta^2/2}$ *pour $0 < \delta \leq 1$.*

6 Distributions classiques

Définition 6.1. Une variable aléatoire X suit la
- **loi uniforme** si $X : \Omega \to V$ avec $\Pr[X = v] = 1/|V|$ pour tout $v \in V$;
- **loi de Bernoulli** (de paramètre p) si $X : \Omega \to \{0,1\}$ avec $\Pr[X = 1] = p$;
- **loi binomiale** (de paramètres p et n) si $X : \Omega \to \mathbb{N}$ avec $\Pr[X = k] = \binom{n}{k} p^k (1-p)^{n-k}$;
- **loi géométrique** (de paramètre p) si $X : \Omega \to \mathbb{N}$ avec $\Pr[X = n] = p(1-p)^{n-1}$;
- **loi de Poisson** (de paramètre $\lambda > 0$) si $X : \Omega \to \mathbb{N}$ avec $\Pr[X = n] = e^{-\lambda}\lambda^k/k!$.

Lemme 6.2. *Soit $X_1, X_2, \ldots$ des variables aléatoires* indépendantes *suivant la loi de Bernoulli de paramètre p. Alors $X = X_1 + \cdots + X_n$ suit la loi binomiale de paramètres p et n. Et $Y = \min\{i : X_i = 1\}$ suit la loi géométrique de paramètre p.*

Proposition 6.3. *Soit X une variable aléatoire. Si X suit la*
- ***loi uniforme** et si $V = \{1, \ldots, n\}$, $\mathbb{E}[X] = (n+1)/2$ et $\sigma^2(X) = (n^2-1)/12$;*
- ***loi de Bernoulli** (de paramètre p), $\mathbb{E}[X] = p$ et $\sigma^2(X) = p(1-p)$;*
- ***loi binomiale** (de paramètres p et n), $\mathbb{E}[X] = np$ et $\sigma^2(X) = np(1-p)$;*
- ***loi géométrique** (de paramètre p), $\mathbb{E}[X] = 1/p$ et $\sigma^2(X) = (1-p)/p^2$;*
- ***loi de Poisson** (de paramètre λ), $\mathbb{E}[X] = \sigma^2(X) = \lambda$.*

7 Autres résultats mathématiques utiles

Sommes diverses :

$$\sum_{i=1}^{n} i = \frac{n(n+1)}{2} \qquad \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} = (x+y)^n$$

$$\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6} \qquad \sum_{k=0}^{n} \binom{n}{k} k x^k y^{n-k} = nx(x+y)^{n-1}$$

$$\sum_{i=0}^{n} x^i = \frac{1-x^{n+1}}{1-x} \qquad \sum_{i \geq 0} x^i = \frac{1}{1-x}$$

$$\sum_{i \geq 0} i x^{i-1} = \frac{1}{(1-x)^2}$$

World's Most Useful Inequality[1] :

$$\forall x \in \mathbb{R}, 1 + x \leq e^x$$

Formule de Stirling :

$$\left(\frac{n}{e}\right)^n \sqrt{2\pi n} \leq n! \leq \left(\frac{n}{e}\right)^n \sqrt{e^2 n}$$

1. © Jeff Erickson : `http://jeffe.cs.illinois.edu/teaching/algorithms/`

Nombre harmonique H_n :

$$\gamma + \ln(n) < \sum_{i=1}^{n} \frac{1}{n} < \gamma + \ln(n+1) \text{ où } \gamma \simeq 0,577$$

Théorème des nombres premiers :

$$\pi(x) = \{p \leq x : p \text{ premier}\} \sim x / \ln(x)$$