Introduction aux lois du zéro-un

Raphaël Rossignol – Bernard Ycart

MAP5

Intuition

Dans un gros espace

tout événement raisonnable

a une probabilité proche de zéro ou un.

Dans un gros espace . . .

$$\forall x \in \{0,1\}^n$$

$$\mu_{n,p}(x) = p^{\sum x(i)}(1-p)^{n-\sum x(i)}$$

- Graphes aléatoires
- Configurations cycliques aléatoires
- Images aléatoires
- k-satisfiabilité

Sortir de $\{0,1\}^n$?

- Equiprobabilité
- Structures combinatoires
- Espaces produits
- Faible dépendance

. . . tout événement raisonnable . . .

- Asymptotique : Kolmogorov, Hewitt-Savage
- Logique : Glebskii-Fagin
- Monotonie : Friedgut-Kalai

. . . a une probabilité proche de 0 ou 1

Inégalités de concentration

- Classiques : Chernov, Hoeffding, McDiarmid
- Elargissements : Talagrand
- Log-Sobolev : Ledoux, Boucheron et al.

Concentration de la mesure

Concentration de la mesure

Logique du premier ordre . . . Largeurs de seuil

Elargissements

- Elargissement

$$A \subset E \longrightarrow A_t = \{x \in E\,,\ \exists y \in A \quad d(x,y) \leq t\}$$

- Fonction de concentration

$$\alpha(t) = \inf \left\{ \alpha > 0 \text{ t.q. } P(A) \geq \frac{1}{2} \Longrightarrow 1 - P(A_t) \leq \alpha \right\}$$

Fonctions de concentration

- Sphère, loi uniforme

$$\alpha(t) \leq \left(\frac{\pi}{8}\right)^{\frac{1}{2}} e^{-\frac{n-1}{2}t^2}$$

- $\mathbb{R}^n$, gaussienne standard

$$\alpha(t) \leq \frac{1}{2}\, e^{-\frac{t^2}{2}}$$

- Espace produit, distance de Hamming

$$\alpha(t) \leq 2\, e^{-\frac{t^2}{n}}$$

Loi des grands nombres

$$\forall x \in \{0,1\}^n$$

$$\mu_{n,p}(x) = p^{\sum x(i)}(1-p)^{n-\sum x(i)}$$

- $X_1, \ldots, X_n$ v.a.i.i.d, Bernoulli $\mathcal{B}(1,p)$
- $S_n = X_1 + \cdots + X_n$ binomiale $\mathcal{B}(n,p)$
- Loi des grands nombres :

$$\lim_{n\to\infty} \frac{S_n}{n} = p$$

Combinaisons boolénnes

$$A_I = \left\{ x \in \{0,1\}\,,\ \frac{x(1) + \cdots + x(n)}{n} \in I \right\}$$

$$\lim_{n \to \infty} \mu_{n,p}(A_I) = \begin{cases} 0 & \text{si} \quad p \notin I \\ 1 & \text{si} \quad p \in I \end{cases}$$

Combinaisons booleénnes

$$A_I = \left\{ x \in \{0,1\}, \; \frac{x(1) + \cdots + x(n)}{n} \in I \right\}$$

$$\lim_{n \to \infty} \mu_{n,p}(A_I) = \begin{cases} 0 & \text{si} \quad p \notin I \\ 1 & \text{si} \quad p \in I \end{cases}$$

Soit $\mathcal{A}$ un ensemble de propriétés, $\widetilde{\mathcal{A}}$ sa fermeture booléenne.

- $\mathcal{A} \subset \widetilde{\mathcal{A}}$
- $A \in \widetilde{\mathcal{A}} \Longrightarrow \neg A \in \widetilde{\mathcal{A}}$
- $A, B \in \widetilde{\mathcal{A}} \Longrightarrow A \wedge B \in \widetilde{\mathcal{A}}$

Si $\mathcal{A}$ vérifie la loi du 0-1, alors $\widetilde{\mathcal{A}}$ aussi.

Propriétés symétriques

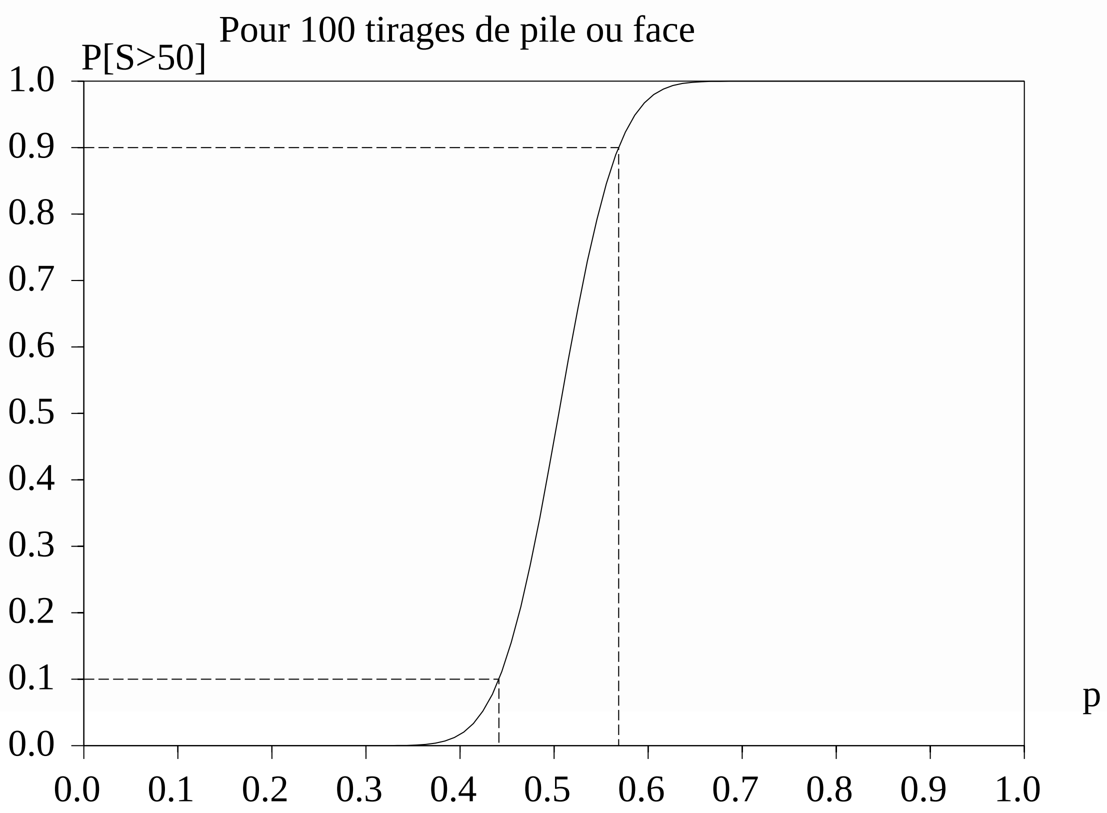
$A \subset \{0,1\}^n$ invariante par l'action transitive de $G \subset \mathcal{S}_n$

- $G = \mathcal{S}_n$: totalement symétrique
- Graphes aléatoires : permutation des sommets
- Configurations cycliques aléatoires : $G = \mathbb{Z}/n\mathbb{Z}$
- Images aléatoires : $G = \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$
- k-satisfiabilité : permutation des variables $\times$ négation des clauses

Propriétés croissantes

$$A \text{ croissante} \Longrightarrow \mu_{n,p}(A) \text{ fonction croissante de } p$$

$$p \mapsto \mu_{n,p}(A) \quad = \quad \alpha \mapsto p_\alpha$$



Inégalités exponentielles

$$\mu_{n,p}(S_n - np > c\sqrt{n}) \leq e^{-2c^2}$$

$$\mu_{n,p}(S_n - np < -c\sqrt{n}) \leq e^{-2c^2}$$

$$\mu_{n,p}(|S_n - np| > c\sqrt{n}) \leq 2e^{-2c^2}$$

Inégalités exponentielles

$$\mu_{n,p}(S_n - np > c\sqrt{n}) \leq e^{-2c^2}$$

$$\mu_{n,p}(S_n - np < -c\sqrt{n}) \leq e^{-2c^2}$$

$$\mu_{n,p}(|S_n - np| > c\sqrt{n}) \leq 2e^{-2c^2}$$

Si A est croissante et totalement symétrique alors

$$p_{1-\varepsilon} - p_\varepsilon \leq \frac{\sqrt{2\log(1/\varepsilon)}}{\sqrt{n}}$$

Théorème de Friedgut–Kalai

Si A est croissante et symétrique alors

$$p_{1-\varepsilon} - p_\varepsilon \leq 7.03\, \frac{\log \frac{1-\varepsilon}{\varepsilon}}{\log n}$$

Forte dérivée

Si pour tout $p \in [0,1]$:

$$\frac{d\mu_{n,p}(A)}{dp} \geq a\,\mu_{n,p}(A)(1-\mu_p(A))$$

Alors,

$$p_{1-\varepsilon} - p_\varepsilon \leq \frac{2}{a} \log \frac{1-\varepsilon}{\varepsilon}$$

Influence des coordonnées

$I_A(k)$: influence de la k-ième coordonnée sur A

Mesure en dimension $n-1$ des configurations pour lesquelles la coordonnée k est “sensible” pour A

$$I_A(k) = \mu_{n-1,p}(\{u \in \{0,1\}^{n-1},$$

$$\begin{array}{l}(u_1, \dots, u_{k-1}, 0, u_k, \dots, u_{n-1}) \notin A \\ (u_1, \dots, u_{k-1}, 1, u_k, \dots, u_{n-1}) \in A\end{array}$$

$$\frac{d\mu_{n,p}(A)}{dp} = \sum_{i=1}^{n} I_A(k)$$

Inégalité isopérimétrique

$$\sum_{i=1}^{n} I_A(k) \geq \frac{\log(n)}{\sqrt{2}\log(12)} \mu_p(A)(1-\mu_p(A))$$

$$a = \frac{\log(n)}{\sqrt{2}\log(12)} \Longrightarrow \frac{2}{a} \simeq \frac{7.028377}{\log(n)}$$

- $\sum I_A(k)$: mesure de surface
- $\mu_p(A)(1-\mu_p(A))$: mesure de volume