

Les Probabilités Sans Peine?

Olivier RIOUL

Télécom ParisTech
Institut Telecom
CNRS LTCI

`olivier.rioul@telecom-paristech.fr`

Journées Télécom-UPS 2012
Paris, France
10 mai 2012

Parlons de ce qui fâche :

- 1 Les Probabilités : un domaine hasardeux ?
- 2 Les Probabilités : un lieu de paradoxes ?
- 3 Les Probabilités : une théorie trop difficile ?
- 4 Les Probabilités : un monde à part ?
- 5 Les Probabilités : ennuyeuses et trop abstraites ?
- 6 Les Probabilités : probable inflation aux concours ?
- 7 En guise de conclusion

Pas de culture mathématique du hasard

- à l'encontre du déterminisme des sciences (Laplace)
- incompetence à décrire complètement notre environnement
- lois des grands nombres, configurations moyennes
- mathématiques appliquées (c-à-d impures)

Distinguer deux problèmes

- logique : axiomes, théorèmes (basé sur l'analyse)
- physique : pertinence, sensations

Domaine
hasardeux ?

Lieu de
paradoxes ?

Trop difficile ?

Monde à part ?

Ennuyeuses / trop
abstraites ?

Inflation aux
concours ?

Conclusion

Coïncidence d'anniversaires

Je parie qu'au moins deux d'entre vous sont nés le même jour (il n'y a pas de jumeaux). Ai-je raison ?

Oui :

- Effectif N
- Interprétation linéaire (fausse) : $N > 365/2$
- $365 \cdot 364 \cdot (365 - N + 1)$ arrangements de N dates distinctes :

$$\mathbb{P}(\text{tort}) = \prod_{k=1}^N (1 - k/365) \approx \exp\left(-\frac{N^2}{2 \times 365}\right)$$

qui est < 0.0001 pour $N = 84$.

Vous êtes candidat au jeu du “Big Deal”

La voiture est derrière l'un des 3 rideaux. Vous choisissez au hasard un rideau ; le présentateur (qui sait où se trouve la voiture) fait ouvrir un autre rideau derrière lequel se trouve une chèvre, et vous offre la possibilité de changer d'avis.

Je change d'avis :

- l'argument d'uniformité ne s'applique pas :
- Problème d'inférence bayésienne : hypothèse H « la voiture est derrière le premier rideau » + événement E « la voiture n'est pas derrière au autre rideau »
- $\mathbb{P}(H) = \frac{1}{3}$ $\mathbb{P}(E|H) = \mathbb{P}(E|H^c) = \frac{1}{2}$ donc $\mathbb{P}(H|E) = \mathbb{P}(H)$ inchangé = $\frac{1}{3}$.

Les Probabilités
Sans Peine?

Olivier RIOUL

Domaine
hasardeux ?

Lieu de
paradoxes ?

Contre-Intuitif
Langagier

Trop difficile ?

Monde à part ?

Ennuyeuses / trop
abstraites ?

Inflation aux
concours ?

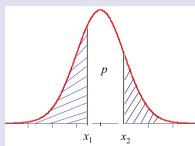
Conclusion

Vous êtes (encore) candidat au jeu TV

On présente deux enveloppes contenant des sommes d'argent, vous en choisissez une et prenez connaissance de son contenu. Vous avez la possibilité de garder celui-ci ou de choisir l'autre enveloppe.

Je prends ma décision sur un tirage :

- Soit x la somme connue, y une réalisation normale (par ex).
- Si $x \leq y$, on choisit l'autre enveloppe.
- $\mathbb{P}(\text{raison}) = \frac{1-p}{2} + p > 0.5$.



Domaine
hasardeux ?

Lieu de
paradoxes ?

Contre-Intuitif
Langagier

Trop difficile ?

Monde à part ?

Ennuyeuses / trop
abstraites ?

Inflation aux
concours ?

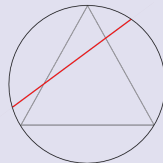
Conclusion

On choisit au hasard une corde d'un cercle

Quelle est la probabilité pour qu'elle soit de longueur supérieure au côté du triangle équilatéral inscrit ?

Tout dépend du « choisi au hasard »

- extrémités au hasard : $\mathbb{P} = 1/3$
- direction + hauteur au hasard : $\mathbb{P} = 1/2$.
- milieu au hasard : $\mathbb{P} = 1/4$.
- longueur au hasard : $\mathbb{P} = 1 - \sqrt{3}/2$.



Domaine
hasardeux ?

Lieu de
paradoxes ?

Contre-Intuitif

Langagier

Trop difficile ?

Monde à part ?

Ennuyeuses / trop
abstraites ?

Inflation aux
concours ?

Conclusion

Votre voisin a deux enfants.

Vous sonnez à sa porte. Une fille ouvre (E). Quelle est la probabilité que l'autre enfant soit un garçon (H) ?

Sans autre information :

- $\mathbb{P}(H) = 2/3$ (sachant qu'il y a une fille).
- $\mathbb{P}(E|H) = 1/2$ (équiprobabilité), $\mathbb{P}(E|H^c) = 1$

$$\frac{\mathbb{P}(H|E)}{\mathbb{P}(H^c|E)} = \frac{\mathbb{P}(H)}{\mathbb{P}(H^c)} \cdot \frac{\mathbb{P}(E|H)}{\mathbb{P}(E|H^c)} = 2 \cdot \frac{1}{2} = 1$$

d'où $\mathbb{P}(H|E) = 0.5$.

Questions

- se farcir les *tribus* et autres σ -algèbres ?
- abandonner les lois continues ? (tout est discret)
- et l'hypothèse normale ? (TCL)

Quel est le but ?

- bénéficier (rapidement) des outils
- besoins pratiques de l'ingénieur
- sans technicité excessive
- sans abandonner la rigueur

Les Probabilités
Sans Peine?

Olivier RIOUL

Domaine
hasardeux ?

Lieu de
paradoxes ?

Trop difficile ?

Mystérieux Ω
Axer sur les « v.a. »

Discret & continu

Distributions

Tout le reste

Monde à part ?

Ennuyeuses / trop
abstraites ?

Inflation aux
concours ?

Conclusion

Quel est ce mystérieux Ω ?

L'univers de tous les possibles

- Théorie élémentaire des probabilités (ensembliste) :

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \quad (= \mathbb{P}(A) \quad \text{si } A \perp B)$$

\implies critiquable.

- Borel-Cantelli ($\limsup A_n$) \implies délicat.
- Choix judicieux de $(\Omega, \mathcal{A}, \mathbb{P}) \implies$ ad-hoc.
- Variables aléatoires $X_n(\omega), \omega \in \Omega \implies$ mystère !

\implies raisonner sur les quantités aléatoires plutôt que sur les événements.

Domaine
hasardeux ?

Lieu de
paradoxes ?

Trop difficile ?

Mystérieux Ω

Axer sur les « v.a. »

Discret & continu

Distributions

Tout le reste

Monde à part ?

Ennuyeuses / trop
abstraites ?

Inflation aux
concours ?

Conclusion

Qu'est qu'une variable aléatoire

- une quantité (X) et non une fonction ($\omega \mapsto X(\omega)$)
- caractérisée (définie !) par les probabilités $\mathbb{P}(X \in A)$
- “toujours” numérique : $X \in \mathbb{R}$ ou $X \in \mathbb{R}^n$ (quitte à transcoder)

Harald Cramér, Mathematical Methods of Statistics, Princeton Univ. Press, 1946 (réimpression 1999).

v.a. la plus générale (Lebesgue-Radon-Nikodym) :

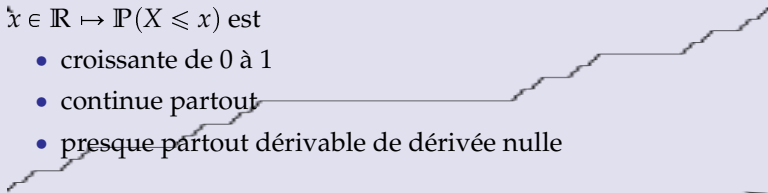
- Partie discrète (masses ponctuelles au plus dénombrable)
- Partie (absolument) continue (définie par une densité)
- Partie diffuse (dégénérée)

Escalier du Diable

$x \in \mathbb{R} \mapsto \mathbb{P}(X \leq x)$ est

- croissante de 0 à 1
- continue partout
- presque partout dérivable de dérivée nulle

Quelles applications ?



Formalisme unifié

- Tout est discret ou continu (ou un mélange)
- Notation sommatoire

$$\mathbb{P}(X \in A) = \sum_{x \in A} p(x)$$

où $p(x)$ est soit une probabilité, soit une densité.

- Comparer

$$\mathbb{E}(X) = \sum x p(x)$$

à

$$\mathbb{E}(X) = \int_{\Omega} X(\omega) d\mathbb{P}(\omega)$$

Domaine
hasardeux ?

Lieu de
paradoxes ?

Trop difficile ?

Mystérieux Ω

Axer sur les « v.a. »

Discret & continu

Distributions

Tout le reste

Monde à part ?

Ennuyeuses / trop
abstraites ?

Inflation aux
concours ?

Conclusion

On peut (on doit ?) faire rapidement

- loi (faible/forte) des grands nombres
- théorème central limite
- lois du tout ou rien (Borel-Cantelli, Kolmogorov)
- processus aléatoires : stationnarité, ergodicité

On peut ne pas (on ne pourra pas) faire rapidement

- l'espace de Hilbert L^2
- les mouvements browniens
- les martingales
- calcul différentiel stochastique
- intégrale de Itô, etc.

Les Probabilités
Sans Peine?

Olivier RIOUL

Domaine
hasardeux ?

Lieu de
paradoxes ?

Trop difficile ?

Mystérieux Ω

Axer sur les « v.a. »

Discret & continu

Distributions

Tout le reste

Monde à part ?

Ennuyeuses / trop
abstraites ?

Inflation aux
concours ?

Conclusion

Monde exotique ?

Langage exotique ?

- *expérience, événement, probabilité, espérance, moments, lois...*
- à comparer à : *corps, anneau, idéal, trace, base, noyau...*
- proximité avec le langage humain, les applications

Discipline exotique ?

- omniprésente pour la modélisation – applis innombrables
- plus utile pour les ingénieurs que pour les commerciaux (prépa HEC)

⇒ son absence en CPGE était incompréhensible pour tous.

Enseignement exotique ?

⇒ isolé par rapport au reste des mathématiques ?

Les Probabilités
Sans Peine?

Olivier RIOUL

Domaine
hasardeux ?

Lieu de
paradoxes ?

Trop difficile ?

Monde à part ?

En Analyse

En Algèbre

En Arithmétique

En Géométrie

Ennuyeuses / trop
abstraites ?

Inflation aux
concours ?

Conclusion

Théorème

Toute fonction continue sur un segment est limite uniforme d'une suite de polynômes.

Démonstration

Soit f continue sur $[0, 1]$, X une v.a. binomiale (p, n) et

$B_n(p) = \mathbb{E}\left(f\left(\frac{X}{n}\right)\right)$, polynôme en p . Si $\varepsilon > 0$:

$$\begin{aligned} |B_n(p) - f(p)| &\leq \mathbb{E}\left(\left|f\left(\frac{X}{n}\right) - f(p)\right| \mathbf{1}_{\left|\frac{X}{n} - p\right| < \varepsilon}\right) + \mathbb{E}\left(\left|f\left(\frac{X}{n}\right) - f(p)\right| \mathbf{1}_{\left|\frac{X}{n} - p\right| \geq \varepsilon}\right) \\ &\leq \max_{\left|\frac{k}{n} - p\right| < \varepsilon} \left|f\left(\frac{k}{n}\right) - f(p)\right| + 2\|f\|_\infty \mathbb{P}\left(\left|\frac{X}{n} - p\right| \geq \varepsilon\right) \end{aligned}$$

On conclut avec la continuité uniforme de f et l'inégalité de Bienaymé-Chebyshev □

Les Probabilités
Sans Peine?

Olivier RIOUX

Domaine
hasardeux ?

Lieu de
paradoxes ?

Trop difficile ?

Monde à part ?

En Analyse

En Algèbre

En Arithmétique

En Géométrie

Ennuyeuses / trop
abstraites ?

Inflation aux
concours ?

Conclusion

Théorème

$$\frac{n^{n+1/2}}{n!} e^{-n} \xrightarrow{n \rightarrow \infty} \frac{1}{\sqrt{2\pi}}.$$

Démonstration

Soit $(X_n)_n$ une suite i.i.d. de v.a. Poisson(1). La v.a. $S_n = X_1 + X_2 + \dots + X_n$ suit une loi de Poisson(n), et par le *théorème central limite*, $\frac{S_n - n}{\sqrt{n}} \xrightarrow{\text{loi}} Y \sim \mathcal{N}(0, 1)$. Comme $x \mapsto x^-$ est continue,

$$\mathbb{E} \left\{ \left(\frac{S_n - n}{\sqrt{n}} \right)^- \right\} \xrightarrow{n \rightarrow \infty} \mathbb{E}(Y^-) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^0 y e^{-y^2/2} dx = \frac{1}{\sqrt{2\pi}}$$

Mais dans $\mathbb{E} \left\{ \left(\frac{S_n - n}{\sqrt{n}} \right)^- \right\} = \sum_{k=0}^n \left(\frac{n-k}{\sqrt{n}} \right) \frac{n^k}{k!} e^{-n}$, tous les termes s'éliminent en cascade sauf $\frac{n^{n+1/2}}{n!} e^{-n}$. □

Théorème

On peut trouver au plus $\max_k \binom{n}{k} = \binom{n}{\lfloor n/2 \rfloor}$ parties de $\{1, 2, \dots, n\}$ dont aucune n'est contenue dans aucune autre.

Démonstration

Soit \mathcal{A} une telle famille de parties, la chaîne aléatoire : $\mathcal{C}_\sigma = \{\emptyset, \{\sigma_1\}, \{\sigma_1, \sigma_2\}, \dots, \{1, 2, \dots, n\}\}$ et $X = |\mathcal{A} \cap \mathcal{C}_\sigma|$ ($= 0$ ou 1).

$$\begin{aligned} 1 &\geq \mathbb{E}(X) = \mathbb{E}\left(\sum_{A \in \mathcal{A}} 1_{A \in \mathcal{C}_\sigma}\right) = \sum_{A \in \mathcal{A}} \mathbb{E}(1_{A \in \mathcal{C}_\sigma}) = \sum_{A \in \mathcal{A}} \mathbb{P}(A \in \mathcal{C}_\sigma) \\ &= \sum_{A \in \mathcal{A}} \frac{1}{\binom{n}{|A|}} \geq \frac{|\mathcal{A}|}{\max_k \binom{n}{k}}. \end{aligned}$$



Théorème

$\forall A, B \in S_n^{++}(\mathbb{R})$ et $\forall \lambda > 0, \mu > 0$ tels que $\lambda + \mu = 1$,

$$\det(\lambda A + \mu B) \geq (\det A)^\lambda (\det B)^\mu.$$

Démonstration (1988)

Soit $X_0 \sim \mathcal{N}(0, A)$ et $X_1 \sim \mathcal{N}(0, B)$, Θ une v.a. binaire $\perp (X_0, X_1)$ de loi de Bernoulli (λ, μ) . La matrice de covariance de $Y = X_\Theta$ est $= \lambda A + \mu B$. Inégalités sur les entropies différentielles :

$$\lambda h(X_0) + \mu h(X_1) = h(Y|\Theta) \leq h(Y) \leq h(Z)$$

où $Z \sim \mathcal{N}(0, \lambda A + \mu B)$. On conclut en calculant l'entropie de la loi $\sim \mathcal{N}(0, C)$; c'est $= - \int p \ln p = \frac{n}{2} \ln(2\pi e) + \frac{1}{2} \ln \det C$. \square

Théorème

$\forall \varepsilon > 0,$

$$\frac{1}{n} \left| \left\{ N \leq n ; (1 - \varepsilon) \ln \ln N < \omega(N) < (1 + \varepsilon) \ln \ln N \right\} \right| \xrightarrow{n \rightarrow \infty} 1.$$

Démonstration

Soit $N = \prod_p p^{X_p}$ une v.a. entière uniforme sur $\{1, 2, \dots, n\}$. Alors $\omega(N) = \sum_{p \leq n} X_p^*$ où $X_p^* = \min(X_p, 1)$ suit une loi de Bernoulli de paramètre $\frac{1}{n} \lfloor \frac{n}{p} \rfloor$. Mais $E(\omega(N)) = \sum_p \frac{1}{n} \lfloor \frac{n}{p} \rfloor = \ln \ln n + O(1)$, et un calcul un peu long montre que $\text{Var}(\omega(N)) = O(\ln \ln n)$. On conclut avec l'inégalité de Bienaymé-Chebyshev :

$$\mathbb{P} \left(\left| \frac{\omega(N) - E(\omega(N))}{\ln \ln n} \right| \geq \varepsilon \right) \leq \frac{\text{Var}(\omega(N))}{(\ln \ln n)^2 \varepsilon^2} \xrightarrow{n \rightarrow \infty} 0. \quad \square$$

Cette preuve est reprise par Hardy... en effaçant les traces !

Théorème

Tout ensemble $A = \{a_1, a_2, \dots, a_n\}$ de n entiers non nuls contient $> n/3$ entiers $\{a_{i_k}\}_k$ tels que $a_{i_k} + a_{i_l} \neq a_{i_m}$ pour tous k, l, m .

Démonstration

Soit $p = 3k + 2$ premier, plus grand que tous les $2|a_i|$, X un v.a. uniforme à valeurs dans \mathbb{Z}_p^* . Alors $X_i = a_i \cdot X \bmod p$ suit aussi une loi uniforme. Soit $B \subset A$ constitué des entiers a_i tels que $X_i \in \{k + 1, k + 2, \dots, 2k + 1\}$. Sa taille moyenne est

$$\mathbb{E}(|B|) = \sum_{i=1}^n \mathbb{P}(k < X_i \leq 2k + 1) = \sum_{i=1}^n \frac{k+1}{p-1} = \frac{n(k+1)}{3k+1} > \frac{n}{3}$$

donc il existe une valeur $X = x$ donnant un $|B| > n/3$. Les éléments $a_i \in B$ sont t.q. $a_i x \bmod p \in \{k + 1, k + 2, \dots, 2k + 1\}$. Si $a_{i_k} + a_{i_l} = a_{i_m}$, en multipliant par x modulo p on trouverait une impossibilité. □

Théorème de Gram

Soit α_i ($0 \leq i \leq n$) la somme des angles solides intérieurs des i -faces d'un polyèdre convexe en n dimensions. Alors

$$\sum_{i=0}^n (-1)^i \alpha_i = 0.$$

Démonstration (1994)

Choisir la direction d'un hyperplan H au hasard et considérer le projeté orthogonal du polyèdre P sur H ...

Domaine
hasardeux ?

Lieu de
paradoxes ?

Trop difficile ?

Monde à part ?

En Analyse

En Algèbre

En Arithmétique

En Géométrie

Ennuyeuses / trop
abstraites ?

Inflation aux
concours ?

Conclusion

Théorème

Soit $C = [0, 1]^n$ l'hypercube unité de \mathbb{R}^n , H_t l'hyperplan d'équation $x_1 + x_2 + \dots + x_n = t$, et S le simplexe défini par l'enveloppe convexe des $n + 1$ points 0 et $e_1, 2e_2, \dots, ne_n$. Alors $\text{vol}(C \cap H_t) = \text{vol}(S \cap H_t)$.

Démonstration

Soit X_1, X_2, \dots, X_n des v.a. i.i.d. exponentielles de paramètre $\lambda > 0$, réordonnés par ordre croissant. Ecrire de deux manières $\mathbb{P}(\forall i, X_i \leq 1)$ et conclure par injectivité de la transformée de Laplace...

Domaine
hasardeux ?

Lieu de
paradoxes ?

Trop difficile ?

Monde à part ?

En Analyse

En Algèbre

En Arithmétique

En Géométrie

Ennuyeuses / trop
abstraites ?

Inflation aux
concours ?

Conclusion

Probabilités ennuyeuses ?



Premiers pas en probabilité...

Combinatoire

- les boules ... dans les urnes, les cartes à jouer, etc.
- raisonnements de pure dénombrement

Statistique

- les écarts-type et inter-quartiles sur tableur
- pure statistique descriptive

Théorie abstraite

- mesure
- probabilités pour elles-mêmes

Les Probabilités
Sans Peine?

Olivier RIOUL

Domaine
hasardeux ?

Lieu de
paradoxes ?

Trop difficile ?

Monde à part ?

Ennuyeuses / trop
abstraites ?

Inflation aux
concours ?

Conclusion

Modélisation

- simulations en scilab
- T.I.P.E.

Domaine
hasardeux ?

Lieu de
paradoxes ?

Trop difficile ?

Monde à part ?

Ennuyeuses / trop
abstraites ?

Inflation aux
concours ?

Conclusion

Théorie enseignée

- doit être utile aux sciences physiques
- formatrice (apprentissage de la logique, du raisonnement, de la rigueur)

Probabilités déjà présentes, mais cachées : (au CCMP)

- PC 2011 un théorème ergodique
- PC/PSI 2010 théorème central limite
- MP 2009 problème des moments et loi log-normale
- PSI 2009 matrices aléatoires
- PC/PSI 2007, MP 2006 matrices stochastiques
- PC/PSI 2004 séries génératrices de v.a. de Poisson

À l'avenir

- une introduction explicite (moins d'hypocrisie...)
- gagner du temps
- enrichir la palette du candidat

Il faudra bien s'adapter... (bon gré mal gré)

Formation des professeurs

- livres pédagogiques
(niveau intermédiaire entre Bac et Master)
- formations LIESSE
- ?