ALEATOIRE - Les enjeux du cours de Probabilités en première année de l'Ecole Polytechnique

Télécom ParisTech, 09 mai 2012

http://www.mathematiquesappliquees.polytechnique.edu/accueil/programmes/cycle-polytechnicien/annee-1/

Organisation du cours pour 500 élèves.

Un cours en amphithéatre, 22 groupes de petites classes en parallèle, et un projet numérique.

Le projet numérique compte pour un tiers de l'évaluation.

Spécificité de l'audience: un niveau de mathématiques très hétérogène.

- 400 élèves issus des concours français, des différentes filières de classes préparatoires.
- 100 élèves étrangers (concours ou dossier)
- quelques étudiants de l'Université.

Profils mathématiques très variés.

Cours de Probabilités: **Le premier contact des élèves avec les Mathématiques Appliquées**
Fondalement tourné vers les applications et la simulation.

Les probabilités

Une théorie mathématique pour quantifier le HASARD.

Fondamentale dans de nombreux domaines d'applications.
Les applications développées à l'Ecole dans le cadre de la troisième année ou dans les masters cohabilités avec l'UPMC ou Orsay:

- Biologie, Ecologie
- Finance, Assurance.
- Informatique et Réseaux de télécommunications
- Médecine, Imagerie médicale
- Physique (physique quantique, physique des particules)
- Traitement du signal, de la parole

Modélisation, Abstraction, Simulation

Nouveauté pour les élèves à découvrir les Mathématiques Appliquées:

- Modélisation
- Analyse théorique
- Simulation numérique.

Le cours de probabilités est construit sur cette idée. Les probabilités sont en lien étroit avec la vie quotidienne.

- Traduction d'une situation concrète en un problème mathématique abstrait: **La modélisation.**

- Innombrables situations où le hasard intervient, de natures très différentes: nécessité d'une **abstraction mathématique** pour donner un cadre général d'étude.

Le modèle probabiliste

Méthodes numériques

- Les probabilités sont beaucoup utilisées à des fins numériques (gros codes de calcul, modélisation de la méconnaissance, turbulence):

  Méthodes de Monte-Carlo

  - Efficaces en grande dimension (mathématiques financières, météorologie, aérospatial)
  - résultats très rapides (en temps réel)
  - Simulation de phénomènes irréguliers.
- Méthodes probabilistes de simulation: mise en place d'expérimentations fictives sur machine appelées **simulations**.

Les élèves apprennent scilab et développent un projet de modélisation qui va jusqu'à la simulation.

Exemples de projets de simulation

Percolation et processus d'invasion

Loi de Mendel et évolution d'une population sous neutralité

Points fixes et cycles de permutations aléatoires

Restauration d'un signal structuré

Simulation d'une couverture de réseaux

Du comportement de certains gaz à Google

Débruitage et théorie de l'information

La value at Risk

L'offre de cours en Probabilités et Statistiques à l'Ecole Polytechnique

Le cours de tronc commun (Année 1):

Va de la construction du modèle probabiliste au théorème de la limite centrale.
Quelques ouvertures aux dynamiques stochastiques.

Année 2: trois cours proposés.

Chaînes de Markov et martingales
Introduction aux méthodes statistiques
Modal: Simulation numérique aléatoire

Année 3 - Master 1

Un programme de Mathématiques Appliquées avec plusieurs cours sur les modèles probabilistes, liés aux parcours que l'on propose. Ces parcours sont associés à un M2 cohabilité avec d'autres établissements:

Mathématiques et sciences du vivant

Mathématique, vision, apprentissage

Probabilités et finances

Probabilités et statistiques.

Les cours de 3ème année liés à l'aléatoire:

Modèles stochastiques en finance

Apprentissage statistique et estimation non paramétrique

Communication networks, algorithms and probability

Modèles aléatoires en écologie et évolution

Stochastic simulation and Monte-Carlo methods

Processes and estimation

Le cours de tronc commun

7 cours de 1h30 + un amphi introductif avec un peu d’histoire.

7 PC (TD) de 2h

L’importance de l’histoire des Probabilités

Un défi de l’homme face au divin: la théorie des probabilités a mis beaucoup de temps à émerger.

- Très certainement d’origine arabe: az-zahr: le dé.
- En Inde, 4ème siècle, existence d’une science du dé et connaissance de ses rapports étroits avec une évaluation de type sondage. (Le Mahabharata)
- En Europe, aux 16-17èmes siècles, émergence d’une science du jeu de dé. Cardan, Kepler, Galilée.
  Théorie rigoureuse: Pascal.
- Résolution de controverses juridiques (Fermat, Leibniz).
- Autre impulsion motivée par des problèmes d’assurance (tables de mortalité et rentes viagères).
- Développement des Statistiques: outil puissant pour les organismes de décision...

Les développements mathématiques

Les probabilités sont un outil privilégié de modélisation des comportements humains, mais deviennent aussi un grand champ de développement des Mathématiques.

- 19ème siècle et début 20ème siècle: Essor des probabilités grâce aux méthodes d'analyse.
  - Calcul intégral et différentiel (Laplace, Gauss)
  - Théorie de la mesure (Borel, Lebesgue).

- A partir du 20ème siècle: Etude de phénomènes aléatoires qui évoluent au cours du temps. Processus de Markov

  - Problèmes de Physique statistique, Mouvement Brownien,

  - Problèmes de démographie. Processus de branchement, Processus de Poisson.

  - Théorie statistique, biométrie.

Aujourd'hui...

- Le modèle probabiliste actuel: Kolmogorov - 1933.
  Calcul stochastique: Itô, Doeblin, à partir de 1945.
- Les probabilités interviennent dans de nombreux développements mathématiques récents.
- L'Ecole Française de Probabilités est très active: Médaille Fields 2006, Wendelin Werner. (Probabilités et transformations conformes du plan).
- Les probabilités largement utilisées dans d'autres sciences (dures ou humaines), dans les entreprises et dans les banques.

Deux idées majeures incontournables

- Loi des grands nombres

  Hasard $=$ méconnaissance des paramètres intervenant dans une expérience.
  Les informations seront données par des répétitions de l'expérience.

- Conditionnement et indépendance
  La construction d'un modèle probabiliste repose sur l'information connue a priori sur l'expérience.
  Si l'information change, la probabilité de réalisation change.

  Si l'information donnée sur un phénomène n'a pas d'influence sur la probabilité de réalisation d'un autre phénomène, les phénomènes sont dits indépendants.
  L'hypothèse d'indépendance est fondamentale pour les calculs. **Notion liée à la probabilité.**

Une difficulté majeure: la notion de loi

Dans tout le cours, on adopte un point de vue fonctionnel: la notion de **variable aléatoire**.

Vocabulaire perturbant (et décalé) des probabilités.

Variable aléatoire = fonction de l'alea.

Intérêt d'une variable aléatoire: transporte l'espace de probabilité abstrait sur l'ensemble des valeurs de la variable aléatoire.

Loi de la variable aléatoire = mesure image de la probabilité abstraite par cette fonction.

Cela définit un nouvel espace de probabilité sur lequel on peut faire des calculs.

Simulations

Simulation: Expérimentation fictive sur machine d'un phénomène que l'on peut modélisé.

Elle permet de visualiser une expérience aléatoire, de calculer des quantités numériques et de vérifier certains résultats théoriques.

Chaque cours est émaillé de simulations et les élèves peuvent les retrouver sur le site internet et jouer avec.

Ils peuvent aussi en voir le code, en Scilab.

Marche aléatoire et mouvement brownien

Marche erratique sur la droite réelle: on avance ou l’on recule d’autant, avec même probabilité, à chaque pas de temps

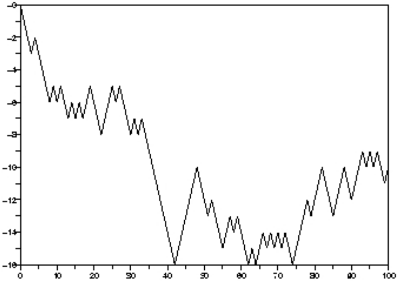


Figure: marche aléatoire en dimension 1

Vu d’une échelle plus macroscopique

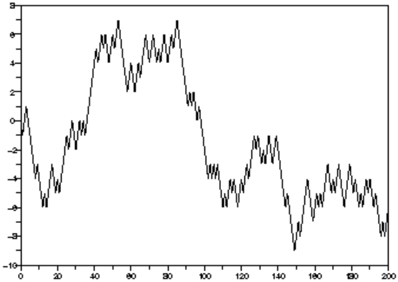


Figure: marche aléatoire en dimension 1

Encore plus...

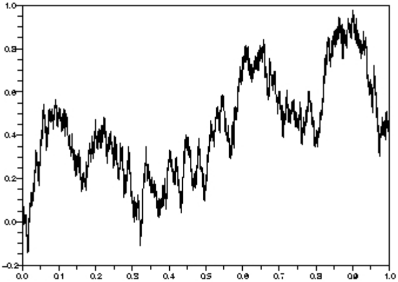
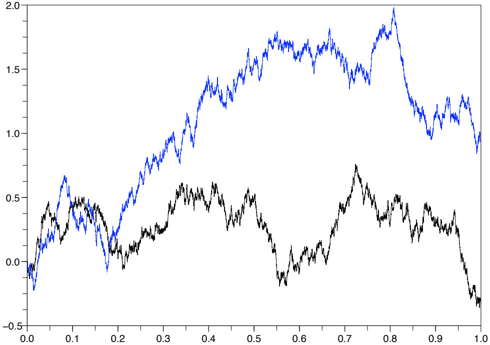


Figure: mouvement brownien en dimension 1

Trajectoires aléatoires

Réalisons plusieurs simulations:

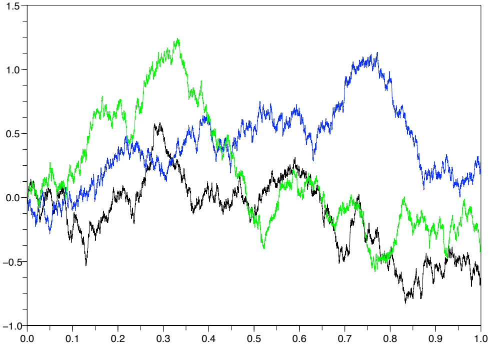
2



Trajectoires aléatoires

Réalisons plusieurs simulations:

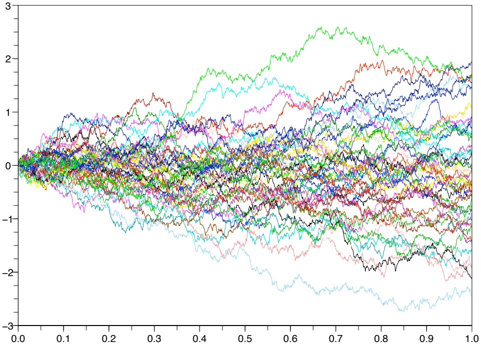
3



Trajectoires aléatoires

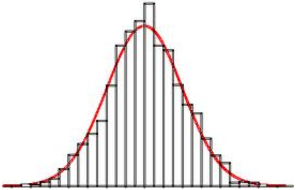
Réalisons plusieurs simulations:

50



La loi normale, loi de Gauss, courbe en cloche

Répartition des valeurs au temps $t = 1$ approchée par la courbe de Gauss: fonction de densité au temps 1.



Loi normale: une loi universelle

Les grands chapitres du cours

- Espace de Probabilité.
- Variables aléatoires sur un espace fini ou dénombrable.
- Variables aléatoires réelles.
- Vecteurs aléatoires, lois conditionnelles, calcul de loi.
- Convergences, Loi des grands nombres.
- Fonction caractéristique, vecteur gaussien, convergence en loi, théorème de la limite centrale.
- Introduction aux processus aléatoires

Espace de Probabilité

Expérience aléatoire et Espace d'états. Evénements aléatoires. **Notion d'information.**

Approche intuitive d'une probabilité: **limite de fréquences empiriques quand le nombre d'expériences tend vers l'infini.** Les premieres propriétés.

Exemples: les modèles d'urnes. Tirage avec ou sans remise, lois hypergéométrique et binomiale.

Définition générale d'une probabilité. (Cas du jeu de Pile ou Face).
Notion de tribu.
L'espace de probabilité $(\Omega, \mathcal{A}, P)$.
Probabilité sur un espace dénombrable.
Probabilités conditionnelles. Formule de Bayes, Exemples.
Indépendance. Expériences aléatoires indépendantes.

Théorème de Borel Cantelli

Variables aléatoires sur un espace fini ou dénombrable

Variable aléatoire et sa loi.

Probabilité sur un espace dénombrable.

Variable aléatoire discrète usuelles. Tout est défini à partir d'un jeu de Pile ou Face: lois de Bernoulli, géométrique, binomiale, convergence vers la loi de Poisson.

Le processus de Poisson: leur donner l'idée d'une dynamique aléatoire qui évolue au cours du temps.

Espérance, variance, fonction génératrice, calcul de moments.

Couple de variables aléatoires, cas discret. Lois conditionnelles, variables aléatoires indépendantes. Somme de variables aléatoires indépendantes.

Information et Entropie.

Variables aléatoires réelles

Variable aléatoire réelle et sa loi

Fonction de répartition. Caractérisation d'une probabilité sur $\mathbb{R}$. (admis)

Variable aléatoire à densité. Variable aléatoire uniforme.

Simulation d'une variable aléatoire (inversion de la fonction de répartition).

Variable aléatoire exponentielle, normale.

Construction de l'espérance (va étagées, va positives comme limites de va étagées, va quelconques).

$$\mathbb{E}(g(X)) = \int_\Omega g(X(\omega))\mathbb{P}(d\omega) = \int_\mathbb{R} g(x)\mathbb{P}_\mathbf{X}(dx).$$

Calcul d'espérance, de variance pour les va à densité.

Calcul de loi: si X a la densité f_X, quelle est la loi de $h(X)$?

Inégalités

Vecteurs aléatoires, lois conditionnelles, calcul de loi

Vecteurs aléatoires. Loi d'un vecteur aléatoire à densité.

Couple de variables aléatoires, covariance. Moments d'un vecteur aléatoire.

Lois conditionnelles, cas à densité. Vecteurs aléatoires indépendants. Calculs de lois de vecteurs aléatoires. (Théorème du changement de variable).

Méthode de simulation dite du rejet.

Convergences, loi des grands nombres

Justifier "les lois empiriques du hasard".

Différents types de convergence Proximité des variables aléatoires: CV en moyenne, en probabilité, presque-sure. Les liens entre ces convergences, exemples et contre-exemples. Théorème de CV dominée (admis).

Loi des grands nombres: loi faible, loi forte (prouvée pour des variables aléatoires de carrée intégrable).

Applications et commentaires: convergence des histogrammes vers la densité, théorème de Weierstrass et polynômes de Bernstein.

Méthode de Monte-Carlo: simuler $\int_A f(z)dz$.

Fonction caractéristique, vecteur gaussien, convergence en loi, théorème de la limite centrale

Fonction caractéristique (transformée de Fourier).

Calcul pour les variables aléatoires usuelles. Caractérisation des variables aléatoires indépendantes. Fonction caractéristique d'une somme de variables aléatoires indépendantes.

Fonction caractéristique et moments.

Vecteurs gaussiens et fonction caractéristique. Matrice de covariance pour un vecteur gaussien.

Convergence en loi, lien avec les autres convergences. Théorème de Paul Lévy.

Théorème de la limite centrale, caractère universel de la loi normale.
Applications (contrôle d'erreur, vitesse de CV dans Monte-Carlo, Statistiques - intervalles de confiance).

Introduction aux processus aléatoires

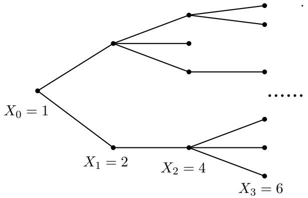
Suites récurrentes aléatoires.

Marche aléatoire.

Problème de la ruine d'un joueur.

Récurrence et transience de la marche aléatoire.

Processus de branchement.



processus de branchement

Files d'attente.