Journal de bord du module
Probabilités et statistique pour la théorie de l'information
Les renvois de la table de matières sont cliquables.

Notes de cours

Version préliminaire; elle sera mise-à-jour au fur et à mesure de l'avancement du semestre.

Table des matières

1 Solutions du contrôle et de l'examen

— La solution du contrôle continu du 9 octobre 2018 (matin) est accessible ici.
— La solution de l'examen du 17 décembre 2018 est accessible ici.

La solution de l'examen du 20 décembre 2017 est accessible ici.

2 Recueil d'exercices

Les exercices pour le semestre sont consultables ici.

3 Introduction et motivation

— Quelques exemples d'application des probabilités et des statistiqueq en théorie de l'information.
  — Quantification de l'information.
  — Capacité du canal.
  — Chiffrement aléatoire.
  — Codes correcteurs d'erreur ; taux de compression ; borne de Shannon.
  — Estimation d'un paramètre.
— Définition intuitive (fréquentielle) de la probabilité.

4 Théorie élémentaire (i.e. sans théorie de la mesure) des probabilités

4.1 Espace probabilisé

— Espace des épreuves.
— Tribu des événements.
— Mesure de probabilité ; vecteur de probabilité dans le cas discret. ← Fin du cours du 7 septembre 2018.
— Mesure de probabilité dans le cas continu ; mesure à densité.
— Propriétés d'une probabilité.
— La probabilité
  — de l'ensemble vide est 0,
  — est croissante pour l'inclusion,
  — est σ-sous-additive,
  — est continue pour des suites monotones d'événements.

— Deux probabilités qui coïncident sur une famille génératrice stable par intersection finie sont identiques.

4.2 Variables aléatoires

— Motivation.
— Définition d'une variable aléatoire sur $(\Omega, \mathscr{F}, \mathbb{P})$ à valeurs dans $(\mathbb{X}, \mathscr{X})$.
— Loi $\mathbb{P}_X$ d'une variable aléatoire.
— Fonction de répartition F_X d'une variable aléatoire réelle X.
— Propriétés la de fonction de répartition. ← Fin du cours du 11 septembre 2018 (matin).
— TD du 11 septembre 2018 (après-midi): exercices 1 – 12.

4.3 Conditionnement, indépendance

— Propriétés plausibles de la probabilité conditionnelle.
— Définition de la probabilité conditionnelle.
— Formule de la probabilité totale et formule de Bayes. Exemples.
— Construction de modèles réalistes pour des probabilités conjointes. Exemples.
— Indépendance. ← Fin du cours du 13 septembre 2018.
— Indépendance ; indépendance 2-à-2, indépendance mutuelle d'une famille de variables aléatoires.
— TD du 14 septembre 2018: exercices 13 –21.

4.4 Espérance, variance, covariance

— Intégrabilité ; définition de l'espérance.
— Énoncé des propriétés de l'espérance : monotonie, linéarité, additivité dénombrable ; l'espérance du produit de deux variables aléatoires **indépendantes** est égale au produit des espérances.
— Variables aléatoires de carré intégrables, variance, covariance, coefficient de corrélation.
— Fonction génératrice et son utilisation. ← Fin du cours du 18 septembre 2018 (matin).
— TD du 18 septembre 2018 (après-midi) : exercices 22–30a.

4.5 Théorèmes des grands nombres

— Inégalités de Markov, de Bienaymé-Tchebychev.
— Convergence en probabilité.
— Énoncé et démonstration du théorème faible des grands nombres.
— Convergence presque sûre.
— Énoncé (sans démonstration) du théorème fort des grands nombres.
— Applications.

4.6 Théorème central limite

— Fonction caractéristique.
— Propriétés de la fonction caractéristique.
— La fonction caractéristique caractérise la loi. ← Fin du cours du 20 septembre 2018.
— Théorème central limite. ← Fin du cours du 20 septembre 2018 ayant légèrement débordé au 21 septembre.
— TD du 21 septembre 2018: exercices 30b–36.

4.7 Chaînes de Markov

— Définition d'une chaîne de Markov. Matrice stochastique. Probabilité initiale.
— Mesure conjointe sur l'ensemble des trajectoires infinies. Marginales fini-dimensionnelles.
— Expression algébrique de la probabilité au temps n.
— Accessibilité, communication.
— Classification des états, irréductibilité.
— Récurrence, transience. ← Fin du cours du 24 septembre 2018.
— Théorème de convergence (dans un cas d'espace fini avec une condition d'irréductibilité renforcée).
— Application algorithmique : PageRank (algorithme de classement des réponses à une requête sur internet). ← Fin du cours du 25 septembre (matin).
— TD du 25 octobre 2018 (matin et après-midi): exercices 37–42a.

5 Introduction à la théorie de l'information

5.1 Quantification de l'information d'une variable aléatoire (ou de sa loi)

— Postulats de la quantité d'information : monotonie, extensivité, regroupement, continuité.
— La seule fonction (à une constante près) satisfaisant les 4 postulats est le fonction entropie.
— Preuve du théorème précédent.
— Trois interprétations de l'entropie.
  — L'entropie est une espérance.
  — L'entropie est un minorant de l'espérance du nombre de questions binaires qu'il faut poser pour connaître le valeur d'une variable aléatoire.
  — Configurations typiques. L'entropie détermine le volume des configurations typiques. Interprétation de ce théorème. Sa démonstration sera faite lors de la prochaine séance. ← Fin du cours du 2 octobre 2018 (matin et début de l'après-midi).
— TD du 2 octobre 2018 (fin d'après-midi): exercices 42b–45 et 47 ; quelques indications pour 46.
— TD du 4 octobre 2018: fin de l'exercice 46 et 48–54a.
— TD du 5 octobre 2018: 54b–57e.
— Démonstration du théorème montrant que l'entropie détermine le volume des configurations typiques. ← Fin du cours du 9 octobre 2018 (matin).

5.2 Propriétés de la fonction entropie et diverses autres quantités provenant de l'entropie

— Si $\mathbf{p}$ et $\mathbf{q}$ sont deux vecteurs de probabilité et X une variable aléatoire dont la loi est décrite par $\mathbf{p}$, alors $H(\mathbf{p}) := -\mathbb{E}(\log p(X)) \leq -\mathbb{E}(\log q(X))$.
— Entropie relative ou contraste de Kullback-Leibler $D(\mathbf{p}\|\mathbf{q})$.
— Entropie conjointe $H(X,Y)$.
— Sous-additivité : $H(X,Y) \leq H(X)+H(Y)$.
— Entropie conditionnelle par rapport à un événement, par rapport à une variable aléatoire $H(X|Y)$. Le conditionnement réduit l'entropie : $H(X|Y) \leq H(X)$.
— Information mutuelle $I(X:Y)$. On peut écrire $I(X:Y) = D(\mathbb{P}_{X,Y}\|\mathbb{P}_X \otimes \mathbb{P}_Y)$. ← Fin du cours du 9 octobre 2018 (après-midi).

5.3 Codage de la source

— Description de la source à espace de symboles $\mathbb{X}$. Codage : $C : \mathbb{X} \to \mathbb{A}^+ = \cup_{n\geq 1}\mathbb{A}^n$, extension du code de $\mathbb{X}$ à $\mathbb{X}^+$ par concaténation.
— Codes non-singuliers, uniquement décodables, instantanés.
— Théorème de Kraft.
  — Si C est un code instantané de $\mathbb{X}$ dans un alphabet $\mathbb{A}$ de cardinal A, alors $\sum_{x\in\mathbb{X}} A^{-|C(x)|} \leq 1$.
  — Si $(l_x)_{x\in\mathbb{X}}$ est une famille d'entiers supérieurs à 1 qui vérifie $\sum_{x\in\mathbb{X}} A^{-l_x} \leq 1$, alors il existe un code instantané C dont les longueurs des mots de code vérifient $\forall x \in \mathbb{X}, |C(x)| = l_x$.
— Théorème de Shannon : Pour tout code instantané C, on a $\mathbb{E}|C(X)| \geq H(X)$.
— Plus précisément : $H(X) \leq \mathbb{E}|C(X)| \leq H(X)+1$ (avec logarithme en base A).
— Théorème de McMillan.
  — Tout code C uniquement décodable vérifie $\sum_x A^{-|C(x)|} \leq 1$.
  — Réciproquement, si (l_x) est une famille d'entiers vérifiant $\sum_x A^{-l_x} \leq 1$, alors il existe un code C uniquement décodable admettant (l_x) comme longueurs des mots de code. (La fin de la démonstration de ce théorème sera faite la prochaine fois). ← Fin du cours du 11 octobre 2018.
  — Fin de la démonstration du théorème de McMillan. (En début de séance du 12 octobre 2018).
  — TD du 12 octobre 2018: exercices 59–62.
  — Codes $\mathscr{K}$-optimaux pour $\mathscr{K}$ une classe de codes.
  — Si $C \in \mathscr{C}_{\text{inst}}$ est optimal, alors il est optimal dans la classe $\mathscr{C}_{\text{ud}}$.
  — Définition de la notion d'arbre binaire de type 0-2.
  — Description algorithmique du code de Huffman sous forme de pseudocode.
  — Critique du code de Huffman. ← Fin du cours du 16 octobre 2018 (matin).
  — Codage de Shannon-Fano-Elias (SFE).
  — Algorithmes de codage et de décodage SFE.
  — Critique du codage SFE. ← Fin du cours du 16 octobre 2018 (après-midi).
  — TD du 16 octobre 2018 (après-midi): Exercices : 63–65.

5.4 Codage du canal (discret)

— Définition d'un canal discret, d'un canal discret sans mémoire.
— Alphabets d'entrée et de sorte, matrice de transmission.

— Classes principales de canaux : sans perte, déterministes, sans bruit, inutiles, symétriques.
— Capacité du canal. Calcul explicite de la capacité pour un canal binaire symétrique.
— Probabilités d'erreur.
— Codage du canal. ← Fin du cours du 06 novembre 2018.
— Codes (K,n) du canal.
— Taux de transmission atteignable.
— Formulation du théorème fondamental de la transmission : Soit C la capacité d'un canal. Tout taux de transmission $R < C$ est atteignable. Toute transmission avec un taux $R > C$ est non fiable.
— Idéé de la démonstration de ce théorème ← Fin du cours du 08 novembre 2018.
— TD du 08 novembre 2018: Exercices : 66, 67, 68.

5.5 Codes correcteurs d'erreurs

— Identification des alphabets $\mathbb{X}$ et $\mathbb{Y}$ du canal avec des corps finis $\mathbb{F}_q$.
— Identification de l'ensemble des mots $\mathbb{X}^n$ avec l'espace vectoriel $\mathbb{F}_q^n$.
— Métrisation par le distance de Hamming.
— (N,L)-codes linéaires.
— Matrices génératrice G et de test de parité H d'un code linéaire. Forme standard de ces matrices.
— Exemples : code à répétition ; code de test de parité ; code de Hamming.
— Identification du décodage par maximum de vraisemblance avec minimum de distance de Hamming.
— Distance minimale entre les mots d'un code.
— Un (N,L,d)-code linéaire corrige jusqu'à $\lceil \frac{d-1}{2} \rceil$ erreurs.
— Si un (N,L)-code linéaire sur $\mathbb{F}_q$ permet de corriger t erreurs, alors $\sum_{i=0}^{t} C_N^i (q-1)^i \leq q^{N-L}$.
— Un code linéaire C avec $d_{\min}(C) = d$ peut corriger t erreurs si, et seulement si, $d \geq 2t+1$. ← Fin du cours du 13 novembre 2018.
— TD du 13 novembre 2018: Exercice 71 ; algorithme de Lempel-Ziv 78.
— Construction du tableau standard pour des (N,L)-codes linéaires ; partition en classes de congruence ; représentant dominant.
— Décodage par localisation dans le tableau standard.
— Syndrome. Décodage par syndrome. ← Fin du cours du 15 novembre 2018.
— TD du 15 novembre 2018: Exercice 72 ; début de 73.

5.6 Application des probabilités au chiffrement de messages

— Divers niveaux de sécurité : sécurité calculatoire, sécurité inconditionnelle.
— Code de Vernam.
— Théorème : Si la clé a la même longueur que le message, elle est utilisée une seule fois et est équidistribuée sur l'ensemble de mots, alors le chiffrement de Vernal offre une sécurité inconditionnelle.
— Démonstration du théorème.
— Quelques remarques sur la génération de nombres aléatoires. ← Fin du cours du 20 novembre 2018.
— TD du 20 novembre 2018: Fin de l'algorithme LZ78, Début de l'exercice 78.
— TD du 22 novembre 2018: Fin de l'exercice 78, exercices 75 et 76.

6 Introduction aux statistiques

- Problème de la statistique.
- Modèle statistique ; exemples.
- Statistiques paramétrique et non paramétrique.
- Les trois types d'inférence statistique (cas paramétrique) :
  - Estimation ponctuelle.
  - Intervalle de confiance.
  - Test d'hypothèses.
- Définition d'une statistique.
- Estimation ponctuelle : principe et exemples. Estimateurs cohérents. Biais, variance, erreur quadratique moyenne.
- Intervalles de confiance : principe et exemples. Estimateurs asymptotiquement normaux, quantile supérieur de la loi normale.
- Test d'hypothèses : principe et exemples.
- Estimation paramétrique multi-dimensionnelle. Estimateurs des moments.
- Estimateur du maximum de vraisemblance. Exemples. ← Fin du cours du 27 novembre 2018.
- Principe général de test d'une hypothèse nulle contre une hypothèse alternative.
- Erreurs du 1er et du 2e types.
- Règle de décision.
- Exemples.
- Test du χ^2. Statistique de Pearson.
- Application de la statistique de Pearson pour tester l'hypothèse d'uniformité d'un générateur de nombres aléatoires. ← Fin du cours du 29 novembre 2018 et fin du module.
- TD du 29 novembre 2018 et fin des TD: Exercices de révision sur les canaux en cascade.