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 Solution de l'examen du 20 décembre 2017

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

2 Sujets d'examens précédentes

Le sujet de l'examen du 15 décembre 2014 est accessible ici.

Le sujet de l'examen du 26 juin 2015 est accessible ici.

Le sujet d'examen du 18 décembre 2015 est accessible ici.

Le sujet du contrôle continu du 9 novembre 2016 est accessible ici.

Le sujet de l'examen du 14 décembre 2016 est accessible ici.

Le sujet du contrôle continu du 16 novembre 2017 est accessible ici.

3 Recueil d'exercices

Les exercices pour le semestre sont consultables ici.

4 Introduction et motivation

- Quelques exemples d'application des probabilités et de statistique 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.

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

5.1 Espace probabilisé

— Espace des épreuves.
— Tribu des événements.
— Mesure de probabilité ; vecteur de probabilité dans le cas discret. ← Fin du cours du 6 septembre 2017.
— 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. ← Fin du cours du 7 septembre 2017
— TD du 7 septembre 2017 (matin): exercice 1.
— TD du 7 septembre 2017 (après-midi): exercices 2 – (début de) 13.

5.2 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.

5.3 Conditionnement, indépendance

— Propriétés plausibles de la probabilité conditionnelle.
— Définition de la probabilité conditionnelle.
— Formule de la probabilité totale et formule de Bayes. Exemples. ← Fin du cours du 14 septembre 2017 (matin).
— TD du 14 septembre 2017 (après-midi): exercices 15–20. La solution de l'exercice 17 sera reprise la semaine prochaine.
— Construction de modèles réalistes pour des probabilités conjointes. Exemples.
— Indépendance ; indépendance 2-à-2, indépendance mutuelle d'une famille de variables aléatoires.

5.4 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.
— TD du 21 septembre 2017: exercices 21–27.

5.5 Théorèmes des grands nombres

— Inégalités de Markov, de Bienaymé-Tchebychev. ← Fin du cours du 21 septembre 2017. ← Fin du cours du 20 septembre 2017.
— 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.

5.6 Théorème central limite

— Fonction caractéristique.
— Propriétés de la fonction caractéristique.
— La fonction caractéristique caractérise la loi.
— Théorème central limite. ← Fin du cours du 28 septembre 2017.
— TD du 28 septembre 2017: exercice 17 re-visitée, exercices 28–33.

5.7 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 4 octobre 2017.
— 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 5 octobre 2017 (matin).
— TD du 5 octobre 2017 (matin et après-midi): exercices 34–36, début de 37.

6 Introduction à la théorie de l'information

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

— Postulats de la quantité d'information : monotonie, extensivité, regroupement, contiinuité.
— La seule fonction (à une constante près) satisfaisant les 4 postulats est le fonction entropie. La preuve de ce théorème sera donnée à la prochaine séance.

— 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 la semaine prochaine. ← Fin du cours du 12 octobre 2017.
— TD du 12 octobre 2017: exercices 37b–41, 44a.
— Démonstration du théorème montrant que l'entropie détermine le volume des configurations typiques.

6.2 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 18 octobre 2017.

6.3 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 $\forall x \in \mathbb{X}, |C(x)| = l_x$. ← Fin du cours du 19 octobre 2017.
— TD du 19 octobre 2017: exercices 44 (b), 43, 45, 47–51.
— 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.
— 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. ← Fin du cours du 20 octobre 2017.
— Description algorithmique du code de Huffman sous forme de pseudocode.
— Critique du code de Huffman. ← Fin du cours du 26 octobre 2017.

— TD du 26 octobre 2017: Exercices : 52–53, 58.
— TD du 8 novembre 2017: Exercices : 54–57d.
— Algorithme de codage de Lempel-Ziv (LZ78).
— Esquisse de la démonstration de l'optimalité de LZ78. ← Fin du cours du 9 novembre 2017.
— TD du 9 novembre 2017: Exercice : 57d–57f.
— TD du 10 novembre 2017: Exercices : 59–62a.
— Contrôle continu le 16 novembre 2017, 10:15–11:15.
— TD du 16 novembre 2017: Exercices : 62b–64.

6.4 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.← Fin du cours du 26 novembre 2017.
— Codage du canal.
— 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. ← Fin du cours du 22 novembre 2017.
— TD du 22 novembre 2017: Exercices : 65, 67, 68a.
— TD du 23 novembre 2017: Exercices : 68b–71.

7 Initiation aux statistiques

— Position du problème statistique.
— Modèle statistique.
— Estimation ponctuelle : estimateurs cohérents, asymptotiquement normaux.
— Estimateur du maximum de vraisemblance.
— Intervalles de confiance. ← Fin du cours du 24 novembre 2017.
— Principe général de test d'une hypothèse nulle contre une hypothèse alternative.
— Erreurs du 1er et du 2e type.
— Règle de décision ; puissance d'un test ; niveau de signification.
— Exemples.
— Test du χ^2. ← Fin du cours du 30 novembre 2017.
— TD du 24 novembre 2017: Exercices : 72, 78.