

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	Recueil d'exercices	1
2	Introduction et motivation	2
3	Théorie élémentaire (i.e. sans théorie de la mesure) des probabilités	2
3.1	Espace probabilisé	2
3.2	Variabes aléatoires	2
3.3	Conditionnement, indépendance	2
3.4	Espérance, variance, covariance	3
3.5	Théorèmes des grands nombres	3
3.6	Théorème central limite	3
3.7	Chaînes de Markov	3
4	Introduction à la théorie de l'information	4
4.1	Quantification de l'information d'une variable aléatoire (ou de sa loi)	4
4.2	Propriétés de la fonction entropie et diverses autres quantités provenant de l'entropie	4
4.3	Codage de la source	4

1 Recueil d'exercices

Les exercices pour le semestre sont [consultables ici](#).

2 Introduction et motivation

- Quelques exemples d'application des probabilités et des statistiques 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.

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

3.1 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.

3.2 Variables aléatoires

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

3.3 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.

3.4 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.

3.5 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.

3.6 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.

3.7 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.

4 Introduction à la théorie de l'information

4.1 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 la 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 la 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.
- La solution du contrôle continu du 9 octobre 2018 (matin) **est accessible ici**.
- Démonstration du théorème montrant que l'entropie détermine le volume des configurations typiques. ← **Fin du cours du 9 octobre 2018 (matin).**

4.2 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 $D(\mathbf{p} \parallel \mathbf{q})$.
- 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 $I(X : Y) = D(\mathbb{P}_{X,Y} \parallel \mathbb{P}_X \otimes \mathbb{P}_Y)$. ← **Fin du cours du 9 octobre 2018 (après-midi).**

4.3 Codage de la source

- Description de la source à espace de symboles \mathbb{X} . Codage : $C : \mathbb{X} \rightarrow \mathbb{A}^+ = \cup_{n \geq 1} \mathbb{A}^n$, 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 \mathcal{K} -optimaux pour \mathcal{K} une classe de codes.
 - Si $C \in \mathcal{C}_{\text{inst}}$ est optimal, alors il est optimal dans la classe \mathcal{C}_{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.