MODULE B4-MATHEMATIQUES
PROBABILITES et STATISTIQUES

Résumés détaillés des cours

Année 2013-2014

Cours201 : Théorie des ensembles - Probabilités

Le premier chapitre présente les bases de la théorie des ensembles. On verra au deuxième chapitre que les opérations logiques sur les évènements (en probabilités) sont comparables aux opérations habituelles de la théorie des ensembles.

On décrit ce qu'est une expérience aléatoire, et on s'intéresse aux résultats de cette expérience, les évènements élémentaires ω. L'ensemble de tous les résultats possibles de l'expérience est l'ensemble fondamental Ω.

Le résultat d'une expérience est un évènement, considéré comme une partie de l'ensemble fondamental Ω. On peut effectuer des opérations sur les évènements, de manière comparable aux opérations effectuées sur les parties d'ensemble en théorie des ensembles.

La loi empirique des grands nombres nous permet de donner une définition intuitive de la notion de probabilité. C'est le mathématicien Kolmogoroff qui en donna la définition axiomatique que l'on utilise aujourd'hui.

On donnera les règles de calcul des probabilités, notamment des probabilités conditionnelles, et des probabilités d'évènements dits indépendants.

Cours202 : Variables aléatoires discrètes

Une expérience aléatoire étant donnée, on peut définir une ou plusieurs fonctions du résultat de l'expérience appelées variables aléatoires. Les variables aléatoires qui ont leurs valeurs dans un ensemble discret (ensemble d'entiers...) sont dites variables aléatoires discrètes.

Une variable aléatoire prend ses valeurs avec une loi de probabilités. On cite en exemple plusieurs lois de probabilité, dont la loi uniforme discrète et la loi de Bernouilli.

Il arrive que la loi de probabilités soit déterminée pour plusieurs variables aléatoires simultanément. On cite le cas du couple de variables aléatoires dont on connait la loi conjointe. On peut alors en déduire la loi de probabilités dite marginale de chacune des variables aléatoires isolément. Enfin, on donne la définition de l'indépendance de deux variables aléatoires.

La loi de probabilités d'une variable aléatoire peut être caractérisée par la fonction de répartition associée. Chaque variable aléatoire est caractérisée également par deux grandeurs qui sont son espérance mathématique et sa variance.

Cours203 : Lois discrètes

Les lois discrètes les plus utilisées que nous présentons ici sont : la loi de Bernouilli ou loi du succès/échec, la loi binomiale ou somme d'une série d'épreuves de Bernouilli, la loi de Poisson qui convient particulièrement au phénomène de comptage d'évènements rares, la loi hypergéométrique ou loi du tirage sans remise.

Cours204 : Variables aléatoires continues

On décrit ce qu'est une variable aléatoire continue définie sur un espace de probabilité (Ω,P) pour laquelle on considère les évènements :

$$\{a \leq X \leq b\}$$

On donne alors les fonctions qui caractérisent au sens des probabilités une variable aléatoire continue : la fonction de répartition et la densité de probabilité.

Une variable aléatoire est totalement déterminée soit par sa fonction de répartition, soit par sa densité de probabilité. En statistiques, on se contente généralement de certaines valeurs caractéristiques qui ne décrivent pas entièrement la variable aléatoire mais qui sont en général suffisantes. Les valeurs caractéristiques les plus importantes sont l'espérance mathématique et la variance.

Cours205 : Lois continues

Dans ce chapitre, on décrit les lois des variables aléatoires continues les plus usuelles : la loi uniforme continue et la loi de Gauss. On en donne toutes les caractéristiques : densité de probabilité et fonction de répartition, espérance mathématique et variance.

La loi de Gauss est aussi connue sous le nom de loi normale. La loi normale est couramment utilisée pour des évaluations statistiques. Au début du XIX^{eme} siècle, les travaux de Gauss établirent l'aspect fondamental de la distribution normale, comme la forme de distribution résultante des erreurs de mesures.

Cours206 : Statistique descriptive

La partie des statistiques qui permet la description et l'analyse d'un groupe donné, sans vouloir en tirer de conclusions à propos d'un groupe plus important s'appelle la statistique descriptive.

On définit d'abord la population, composée d'individus, qui forme le champ d'analyse d'une étude particulière. Les individus sont étudiés suivant un ou plusieurs caractères que l'on observe. On dispose alors de données. L'objet de ce chapitre est de montrer comment étudier ces données, pour obtenir des caractéristiques de leur répartition.

Pour caractériser une répartition de données, on utilise une mesure de leur tendance centrale. Plusieurs mesures de ce type sont couramment utilisées : différentes formules de moyennes, la médiane et le mode.

Enfin, en complément de la mesure de tendance centrale, l'allure de la répartition des données est caractérisée par les mesures de dispersion.

Cours207 : Corrélation

Très souvent, il existe en pratique une relation entre deux variables ou plus. Il est souvent utile de pouvoir exprimer cette relation sous la forme d'une formule mathématique. Dans ce chapitre, on passe en revue les différentes méthodes qui permettent d'établir cette relation mathématique. La méthode la plus utilisée est la méthode des moindres carrés. Nous la présentons en détail.

Naturellement, il est nécessaire de prouver que la relation établie entre variables est juste. Pour cela, on dispose du coefficient de corrélation qui mesure la validité de l'ajustement réalisé. On donnera son expression dans le cas général et dans le cas particulier de la corrélation linéaire.

Cours208 : Ajustements statistiques et test du χ^2

Une série observée est le résultat d'observations où on dispose d'une liste de valeurs avec les effectifs observés correspondants.

On présente d'abord des ajustements de séries observées avec des lois de probabilités théoriques. Le principe en est le suivant. Pour une série observée donnée, on choisit une loi de probabilités adaptée. On calcule alors, pour chaque valeur de la série, l'effectif théorique donné par la loi de probabilités. Finalement, on compare l'ensemble des effectifs observés à celui des effectifs théoriques calculés et on juge de la validité de l'ajustement choisi.

Différents ajustements sont possibles, suivant les situations rencontrées. On retient, à titre d'exemple les ajustements suivant : ajustement d'une loi uniforme, ajustement d'une loi binômiale et ajustement d'une loi de Poisson.

Il y a de nombreux avantages à substituer une loi de probabilité connue à une série observée. En effet, cette loi a des propriétés connues, elle est en général tabulée. On présente alors, à la place de la série observée, quelques caractéristiques simples de la loi. Ceci facilite, entre autres, les comparaisons entre séries.

On peut observer directement si l'ajustement effectué semble donner de bons résultats. Mais on dispose également d'un outil très puissant pour affirmer si l'ajustement est valide ou non : le test du χ^2. Le test du χ^2 est précis et s'applique à toute loi théorique. Nous donnons de nombreux exemples où le test du χ^2 permet d'obtenir des conclusions pertinentes quand à la validité des ajustements effectués.

Cours209 : Echantillonnage

La théorie de l'échantillonnage vise à étudier les relations qui existent entre la population et les échantillons tirés de cette population. L'étude des caractéristiques d'une population à partir d'échantillons tirés de celle-ci, ainsi que l'estimation de la précision de ces estimations grâce à la théorie des probabilités s'appelle l'**inférence statistique**.

On passera en revue les différentes méthodes de sélection d'un échantillon au sein d'une population. Puis, on verra comment estimer un paramètre par une statistique. La valeur inconnue d'une population, à estimer à partir d'un échantillon, est appelée un **paramètre**. La paramètre de la population est estimé à partir d'une **statistique** calculée sur la base d'un échantillon.

A titre d'exemple, on considérera des distributions d'échantillonnage de moyenne ou de proportion.