MAT 8480–MÉTHODES NON PARAMÉTRIQUES–HIV 2011

Horaire et local: mardi-jeudi, 10h30-12h; SH–3720.

Professeur: Sorana Froda; *e-mail*: froda.sorana@uqam.ca

Paramètres du prof: local PK-5520; tél: 987-3000, poste 3220.

SUJET

On attaque des questions statistiques par des méthodes dites non paramétriques quand les données suivent une loi continue, inconnue (par exemple, si l'hypothèse de normalité n'est pas satisfaite) ou quand les données sont de type qualitatif. (En particulier, les biologistes s'intéressent beaucoup à ces méthodes.) La plupart des tests rencontrés dans un cours de statistique élémentaire ou dans d'autres domaines (analyse de variance par exemple) ont un équivalent non paramétrique. Un deuxième volet du cours est le *lissage=smoothing* ou l'estimation fonctionnelle (estimation de densités, de fonctions de hasard, de fonctions de régression), qui est un domaine de recherche très actif présentement .

PRÉREQUIS: inférence statistique de base, estimation et surtout tests d'hypothèses, concepts de seuil (niveau), puissance, probabilité critique, tests de t pour en moyenne, comparaison de moyennes. Un livre de référence serait Rice, *Mathematical Statistics and Data Analysis*, ch. 9 et 11, ou Degroot and Schervish, *Probability and Statistics*, ch. 8.

STRUCTURE DU COURS

Partie 1: Statistique non paramétrique "classique".

Partie 2: Estimation de fonctions.

CONTENU DU COURS

Partie 1 (8 semaines environ). On développe en détail : statistiques d'ordre; test du signe, test de Wilcoxon-Mann-Whitney et test signé de Wilcoxon; efficacité asymptotique relative de deux tests: comparaison entre le test de Wilcoxon et le test t de Student; estimation robuste: moyennes tronquées.

Aperçu d'autres tests et procédures : analyse de variance à un facteur: test de Kruskal-Wallis; coefficient de corrélation de Spearman et coefficient τ de Kendall; test d' ajustement de Kolmogorov-Smirnov, tests de permutation.

Partie 2 (6 semaines environ). On introduit ces concepts : estimateur par le noyau de la densité et de la fonction de régression; estimation localement polynomiale; concept de smoothing (lissage) et choix du paramètre lissant; comparaison de plusieurs méthodes; quelques procédures simples du langage *R* (*lowess*, etc).

ÉVALUATION PROPOSÉE:

▷ séries d'exercices à 60% (environ 4 devoirs sur la première partie et un très grand devoir sur la deuxième);

▷ deux examens, dont un 'take home' (le final), total: 40%;

▷ dates examens : mardi 15 ou jeudi 17 mars (intra); mardi 26 avril–jeudi 28 avril (final).

Quelques LIVRES (à la réserve de la bibliothèque)

▷ **Livres utilisés dans la présentation**:

Partie 1.

R.H. Randles and D.A. Wolfe (1979). *Introduction to the theory of nonparametric statistics.* John Wiley. QA278.8R36 (ch. 1, 2, sect. 11.1, 12.3).

E.L. Lehmann (1975). *Nonparametrics: statistical methods based on ranks.* Holden-Day. QA278.8L43 (sect. 1.1-1.5; 2.1-2.5; 3.1, 3.2, 3.5; 4.1-4.3; 5.1-5.2; 6.1-6.2; 7.1-7.2).

P. Sprent (2001). *Applied nonparametric statistical methods*. Chapman and Hall. QA278.8.S74.2001 (contient les procédures de base rencontrées en pratique).

Partie 2.

J. R. Thompson and R.A. Tapia (1990). *Nonparametric function estimation, modelling and simulation*. SIAM. QA276.8T56 (jolie introduction à l'estimation de densités; sect. 2.5).

J. Fan and I. Gijbels (1996). *Local polynomial modelling and its applications*. Chapman and Hall. QA278.2F25 (général, bonnes références; ch. 2).

P.J. Green and B. W. Silverman (1994). *Nonparametric regression and generalized linear models*. Chapman and Hall. QA278.2G69 (contient une jolie introduction aux *splines*).

▷ **Autres livres**:

Partie 1.

J. Hájek and Z. Šidák (1967). *Theory of rank tests*. Academic Press. QA277.H24 (très avancé, un "classique").

J. Hájek (1968). *A course in nonparametric statistics*. Holden-Day. QA278.8H313 (complément appliqué du précédent).

M. Hollander and D.A. Wolfe (1973). *Nonparametric statistical methods*. John Wiley. QA278.8H65 (complément plus appliqué du premier livre de la liste).

Partie 2.

B. W. Silverman (1994) *Density estimation for statistics and data analysis*. Chapman and Hall. QA276.8S55

J. Hart (1997). *Nonparametric Smoothing and Lack-Of-Fit Tests*. Springer-Verlag. QA278.H357.1997 (avancé).

AUTRE DOCUMENTATION

Plusieurs résumés de théorie et sorties en "R" sont disponibles sur Moodle, cours MAT 8480.