STATISTIQUE INFÉRENTIELLE

Sous l'impulsion notamment du mouvement de la qualité, les méthodes statistiques sont aujourd'hui largement utilisées dans les milieux économique, social ou professionnel. Des procédures plus ou moins élaborées sont mises en œuvre, par exemple dans l'analyse des résultats d'expériences sur le vivant, les sondages, la maîtrise statistique des procédés, la fiabilité, les plans d'expériences. Des logiciels spécialisés exécutent automatiquement les calculs, suivant les normes AFNOR ou ISO.

Au-delà de l'exécution d'algorithmes ou de calculs dont le sens peut échapper, l'objectif essentiel de ce module est d'initier les étudiants, sur quelques cas simples, au raisonnement et méthodes statistiques et à l'interprétation des résultats obtenus.
Il s'agit de faire percevoir, à partir d'exemples figurant au programme, ce que sont les procédures de décision en univers aléatoire, ainsi que leur pertinence. Pour cela, la réalisation de simulations dans le cadre du modèle probabiliste de référence peut fournir un éclairage intéressant.
On soulignera que la validité d'une méthode statistique est liée à l'adéquation entre la réalité et le modèle la représentant.
On évitera les situations artificielles et on privilégiera les exemples issus de la vie économique et sociale ou du domaine professionnel, en liaison avec les enseignements d'autres disciplines ; dans le cadre de cette liaison, on pourra donner quelques exemples d'autres procédures que celles figurant au programme de mathématiques (par exemple utilisation de la droite de Henry, du test du χ^2, de la loi de Student), en privilégiant les aspects qualitatifs, mais aucune connaissance à leur sujet n'est exigible dans le cadre de ce programme.
On se placera dans le cadre d'échantillons considérés comme réalisations de variables aléatoires indépendantes.

a) Estimation ponctuelle d'un paramètre : - fréquence ; - moyenne et écart type.	Une illustration qualitative succincte des notions de biais et de convergence d'un estimateur peut être proposée, mais toute étude mathématique de ces qualités est hors programme.
b) Estimation par un intervalle de confiance d'un paramètre : - fréquence dans le cas d'une loi binomiale approximable par une loi normale ; - moyenne, dans le cas d'une loi normale quand son écart type est connu ou dans le cas de grands échantillons.	On distinguera confiance et probabilité : - avant le tirage d'un échantillon, la procédure d'obtention de l'intervalle de confiance a une <u>probabilité</u> $1-\alpha$ que cet intervalle contienne le paramètre inconnu, - après le tirage, le paramètre est dans l'intervalle calculé avec une <u>confiance</u> $1-\alpha$.
c) Tests d'hypothèse : - relatifs à une fréquence *p*, dans le cas d'une loi binomiale approximable par une loi normale, tester $p = p_0$ contre $p > p_0$, contre $p < p_0$; tester $p = p_0$ contre $p \neq p_0$; - relatifs à une moyenne *m*, dans le cas de la loi normale, tester $m = m_0$ contre $m > m_0$, contre $m < m_0$; tester $m = m_0$ contre $m \neq m_0$; - comparaison de deux proportions ou de deux moyennes.	La taille *n* de l'échantillon sera suffisamment grande ($n \geq 30$). On soulignera que la décision prise, rejet ou acceptation, dépend des choix faits *a priori* par l'utilisateur : choix de l'hypothèse nulle, choix du seuil de signification.

Travaux pratiques

1° Estimation ponctuelle et par intervalle de confiance de la fréquence, dans le cas d'une loi binomiale connue, à partir d'échantillons simulés.	La connaissance *a priori* de la loi sous-jacente permet de comparer le paramètre réel et les estimations obtenues à partir des échantillons. Aucune connaissance sur ce TP n'est exigible dans le cadre du programme de mathématiques.
2° Estimation ponctuelle et par intervalle de confiance de fréquences. Estimation ponctuelle de moyennes, d'écarts types et estimation par intervalle de confiance de moyennes, dans des situations relevant de la loi normale.	Quand *n* est grand, le théorème de la limite centrée étend la procédure mise au point pour les échantillons gaussiens à des cas plus généraux.
3° Construction et utilisation de tests : - unilatéraux et bilatéraux relatifs à une fréquence ; - unilatéraux et bilatéraux relatifs à une moyenne dans des situations relevant de la loi normale.	La construction d'un test comporte le choix des hypothèses nulle et alternative, la détermination de la région critique et l'énoncé de la règle de décision.
4° Construction et utilisation de tests de comparaison de deux proportions ou de deux moyennes.	Cette comparaison peut permettre, par exemple, d'apprécier une éventuelle amélioration dans un processus de fabrication.
5° Exemples d'utilisation de la droite de Henry, du test du χ^2, du test de Student (cas des petits échantillons).	Ce TP n'est à réaliser, en entier ou en partie, qu'en liaison avec les enseignants des disciplines professionnelles et seulement si, dans celles-ci, ces procédures sont utilisées. Aucune connaissance à son sujet n'est exigible dans le cadre du programme de mathématiques.