Université de Kairouan
Institut Supèrieur des Mathématiques Appliquées & Informatique

Session Janvier 2016

Module	**Statistiques paramétriques & non paramétriques**
Auditoire	**$1^{\text{iére}}$ Année Mastère Ingénierie financière**
Enseignant	**Mohamed Essaied Hamrita**
Durée	**Deux heures**

La présentation, la lisibilité, l'orthographe, la qualité de la rédaction et la clarté entreront pour une part importante dans l'appréciation des copies. Aucun document n'est autorisé.

Exercice 1 : Soit $(X_1, X_2, \ldots, X_n)$ un $n-$échantillon de variables aléatoires iid issu de la loi Binomial de paramètres (n, p) de densité de probabilité :

$$f(x,p) = p[X = x] = C_x^n p^x (1-p)^{n-x}, \qquad x = 0, 1, \ldots, n.$$

1) Construire un estimateur pour $\widehat{p}$ par la méthode du maximum de vraisemblance. Calculer cet estimateur pour l'échantillon suivant : $(3.6, 2, 0, 0, 3)$ et $n = 10$.

2) Étudier les propriétés statistiques de $\widehat{p}$.

Exercice 2 : Le coût d'un certain sinistre est une variable aléatoire suivant une loi $N(m, \sigma^2)$. On observe, dans une compagnie d'assurance, n dossiers de sinistres indépendants.

1) Proposer un estimateur pour le paramètre m. Donner sa loi.

2) On suppose que la variance est connue, est égale à 15^2 dinars.

i) Calculer l'intervalle de confiance 95% pour m.

ii) Combien de dossiers doit-on examiner pour que la longueur de l'intervalle de confiance soit au plus égale à 10 dinars.

3) Déterminer l'intervalle de confiance pour m lorsque la variance est inconnue et lorsqu'on a observé 20 dossiers, on a trouvé $\overline{X} = 120$ et $\sum_{i=1}^{20} X_i^2 = 290281$.

Exercice 3 : Les parties A et B sont indépendantes.

Le tableau suivant illustre les valeurs de 8 biens consommés par un individu en deux dates différentes (X : 31/12/2014 ; Y : 31/12/2015), choisis au hasard parmi l'ensemble des biens consommés par cet individu.

Biens	31/12/2014 (X)	31/12/2015 (Y)
Logements	180	183
Habillement	90	82
Frais de déplacement	110	118
Dépenses loisirs	80	85
Dépenses de soins	70	71
Dépenses fournitures scolaires	80	67
Dépenses de télécommunication	60	67
Dépenses en nourritures	200	198

Partie A : Dans cette partie, on suppose que $X \sim N(m_X, \sigma_X = 52.5)$ et $Y \sim N(m_Y, \sigma_Y = 52.5)$.

1) On veut tester $H_0 : m_X = 109$ vs $H_1 : m_X = 115$.

a) Déterminer la région critique au sens de Neyman-Pearson correspondant au seuil de signification $\alpha = 5\%$ (sans démonstration).

b) Quelle est la décision à prendre ?

c) Calculer β le risque de deuxiéme espéce. En déduire la puissance du test.

2) a) Déterminer un intervalle de confiance pour un niveau de signification 95% de $m_X - m_Y$.

b) Peut-on admettre, au seuil $\alpha = 5\%$, que la valeur moyenne des biens n'a pas changé ?

Partie B : Dans cette partie, on suppose que les lois des deux variables aléatoires, X et Y, sont inconnues. On veut tester $H_0 : m_X = m_Y$ vs $H_1 : m_X = m_Y$.

1) Quel test peut-on appliquer dans ce cas ? Justifier.

2) Appliquer le test proposé et donner la décision au seuil $\alpha = 5\%$.

Exercice 4 : 395 personnes sont sélectionnées au hasard, et sont regroupées selon deux caractères :

X : Soutiennent-ils ou non la légalisation de la consommation du canabis ("zatla").

Y : La classe d'âge dans laquelle ils tombent.

Les résultats sont donnés dans le tableau suivant :

	18-24	25-34	35-49	50-64	**Total**
Oui	60	54	46	41	201
Non	40	44	53	57	194
Total	100	98	99	98	395

Tester au seuil $\alpha = 5\%$, l'indépendance de deux caractères.

Bon Travail