

**Ecole Polytechnique de Thiès**  
**Département Génie Informatique et Télécommunication**  
**TD1 : Analyse statistique**  
**Année 2018-2019**  
**N. F. NGOM**

**Exercice 1** *Tendances*

*On souhaite avoir une idée du choix des spécialités des étudiants de l'ept en fonction du genres.*

1. *Générer le tableau des spécialités en fonction des spécialités pour la TC2.*
2. *Calculer les tendances centrales et les mesures de variabilités.*
3. *Calculer les corrélations entre les différentes variables.*

**Exercice 2** *Estimation statistique*

1. *Loi binomiale*

(a) *L'objectif est de trouver une loi et ses paramètres pour un référendum. Puisque dans un référendum il y a uniquement deux réponses possibles, alors le modèle de Bernouilli peut être utilisé. Il nous reste maintenant à trouver les paramètres de la loi  $p$  à partir d'un échantillon de taille 100.*

- i. *Calculer la vraisemblance de l'échantillon.*
- ii. *En déduire l'estimateur au maximum de vraisemblance.*

(b) *Loi exponentielle est donnée par*

$$f(x, \theta) = \theta e^{-x}; E(X) = \frac{1}{\theta}; V(X) = \frac{1}{\theta^2}$$

- i. *Estimer les paramètres empiriques de cette loi à l'aide de la méthode des moments.*
- ii. *L'estimateur du paramètre  $\theta$  trouvé est-il biaisé ou non ?*

(c) *Lors d'une enquête policière, si un suspect de sexe inconnu mesure environ 1.5, on aura tendance à rechercher une femme tandis que s'il mesure environ 1.8, on recherchera plutôt un homme. La notion de **maximum de vraisemblance permet de formaliser cette intuition**. Supposons que l'on peut modéliser la distribution des tailles (en mètres) féminine par une loi gaussienne d'espérance  $\mu_1 = 1.62$  et d'écart-type  $\sigma_1 = 0.069$  et celle des tailles masculines par une loi gaussienne d'espérance  $\mu_2 = 1.76$  et d'écart-type  $\sigma_2 = 0.073$ .*

- i. *Quelle conclusion peut on tirer de ce modèle ?*
- ii. *A l'aide de la méthode des moments, estimer les paramètres du modèle.*

- iii. Calculer l'estimateur au maximum de vraisemblance des paramètres du modèle à partir d'un échantillon.

**Exercice 3** Estimation par intervalle de confiance

L'entreprise pour laquelle vous travaillez envisage de s'implanter sur le marché d'un pays voisin pour vendre un nouveau jus de fabrication. Ce marché ne l'intéresse que si la consommation moyenne par an et par habitant est de 0.20 litres. Si cette consommation est comprise entre 0.20 et 0.40, alors elle exportera à partir du Sénégal. Si la consommation moyenne dépasse 0.40 litres, elle envisage la création d'une unité de production dans ce pays. Une enquête auprès de 400 personnes montre, sur cet échantillon, la consommation moyenne est de 0.23 litre. Quelle décision faut-il prendre ?

**Exercice 4** Test d'indépendance du Chi Square

Un centre de formation délivre trois diplômes de fin d'études. Son directeur aimerait savoir si la durée de recherche d'un emploi après l'obtention du diplôme est liée au diplôme obtenu. Pour cela, il réalise un sondage auprès des diplômés d'une promotion de 100 étudiants et obtient les résultats donnés par la table

Durée de recherche	Diplôme I	Diplôme II	Diplôme III
< 1 mois	8	8	4
2-6 mois	9	11	20
> 6 mois	13	11	16

1. Calculer les tendances et les mesures de variabilités.
2. Quelle conclusion pouvez-vous faire sur l'hypothèse selon laquelle la durée de recherche d'un emploi après l'obtention du diplôme est liée au diplôme obtenu au **niveau de risque**  $\alpha = 5\%$ .