

# Projet - MST

Nicolas Brunel, Anastase Charantonis, Julien Floquet

04/05/2018

*Ce projet est à déposer sur [exam.ensiie.fr](http://exam.ensiie.fr) **avant** le 3 juin.*

## CONSIGNES

1. Vous devez effectuer OBLIGATOIREMENT votre rapport en BINOME.
2. Nous vérifierons durant la séance du 14 Mai l'avancement de votre travail et si vous avez un binôme. Cette vérification entrera dans la notation du projet.
3. Le rendu doit être un document .pdf de 5 à 10 pages, limites strictes. Il ne peut pas contenir des annexes.
4. La notation tiendra également compte de la qualité de rédaction ainsi que de la qualité et de la pertinence des interprétations des graphiques.
5. Le rapport doit avoir comme titre : Binome \_Numero\_ Nom1 \_Nom2.pdf

---

Nous étudions ici la quantité de pêche (en kg) obtenue dans le port de Perdika, en fonction de l'intensité du vent (en noeuds), de la tranche d'âge (jeune : 0, adulte : 1, agé : 2) et du sexe (homme : 0 femme : 1) du pêcheur. Les quantités négatives de pêche correspondent à des préventes non honorées.

Le but de ce projet est de réaliser une petite étude en utilisant ce qui a été vu durant le module de modélisation statistique. Le projet comprend plusieurs parties : statistiques descriptives, inférence, intervalles de confiance et tests.

Quand vous avez votre binôme, allez sur <https://tinyurl.com/ycu8yenb>. Vous y trouverez un document de type spreadsheet à remplir avec les noms de vos binômes. Inscrivez votre nom dans les cases adaptées, puis téléchargez la base de données correspondante à votre binôme, présente dans le dossier BD ici : <https://tinyurl.com/ycbebc52>.

## 1. Statistiques Descriptives

- (a) Tracer les boîtes à moustaches de la quantité de pêche pour les hommes et pour les femmes sur un même graphique. Que remarquez-vous ?
- (b) Tracer la quantité de pêche en fonction de la tranche d'âge. Que remarquez-vous ?
- (c) Tracer l'histogramme l'intensité de vent. Quelles distributions l'intensité de vent pourrait-elle suivre (justifiez) ?
- (d) Tracer l'histogramme la quantité de pêche. Quelles distributions la quantité de pêche pourrait-elle suivre (justifiez) ?

## 2. Statistique Inferentielles

On considère que l'intensité du vent suit une loi de Poisson  $P(\lambda)$ . Dans cette partie nous cherchons à estimer  $\lambda$ .

- (a) Vérifier que les observations suivent une loi de poisson.
- (b) Ecrire la vraisemblance (ou la log vraisemblance) théorique du modèle.
- (c) Déterminer le maximum de vraisemblance du paramètre  $\lambda$ . Tracer la log-vraisemblance de l'échantillon en fonction de la valeur du paramètre  $\lambda$ .
- (d) Donner une estimation du paramètre  $\lambda$  et la faire apparaître sur le graphique.

On suppose que la quantité de pêche suit une loi Normale  $N(\frac{100}{1 + \lambda_i}, \sigma^2)$ , ou  $\lambda_i$  est l'intensité du vent en noeuds au moment de la pêche.

- (e) Ecrire la vraisemblance (ou la log vraisemblance) théorique du modèle de la quantité de pêche. Pour les calculs numériques, on peut considérer que les probabilités de réalisation de vagues  $< 0.000001$  peuvent être omises du calcul.
- (f) Utilisez le T.C.L. pour écrire la vraisemblance (ou la log vraisemblance) de la loi de repartition de la quantité de pêche.
- (g) Déterminer le maximum de vraisemblance du paramètre  $\sigma$ . Tracer la log-vraisemblance de l'échantillon en fonction de la valeur du paramètre  $\sigma$  avec chaque une des méthodes.
- (h) Donner une estimation du paramètre  $\sigma$  avec les deux approches et la faire apparaître sur les graphiques.

### 3. Intervalles de confiance

- (a) Ecrire mathématiquement l'intervalle de confiance asymptotique pour le paramètre  $\lambda$  en justifiant.
- (b) Calculer avec l'aide de R les bornes de l'intervalle de confiance asymptotique pour ce paramètre.

### 4. Tests

- (a) Nous souhaitons comparer la quantité de pêche obtenue par les femmes et les hommes. Proposez un test et l'implémenter sous R. Que concluez-vous ?
- (b) Nous souhaitons examiner si la quantité de pêche des jours d'intensité 3 Noeuds sont issus d'une distribution avec un écart-type  $\sigma = 20$  quelle est la valeur de alpha (à 0.01 près) à partir de laquelle nous ne pouvons pas rejeter cette hypothèse ?