

# Statistique inférentielle

## TP bilan

Un compte rendu est à rendre à la fin du TP. Il s'agira d'un rapport au format PDF, incluant les programmes informatiques R. Le TP est à réaliser en binôme. La qualité de la rédaction (programmes et rapport) sera prise en compte. Le compte rendu sera envoyé à l'adresse **gwandji@polytech-lille.fr**.

- Nom du fichier : **GIS2A3\_TP\_SI\_NOM\_PRENOM.pdf**
- Intitulé du mail : **GIS2A3 CONTROLE TP SI**

### Exercice 1 : maximum de vraisemblance

9 pts

Les données suivantes représentent la durée de vie expérimentale d'un composant électronique observé sur un échantillon de mesures :

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
durée de vie	48.80	7.30	27.30	53.30	25.70	59.40	42.90	63.70	68.40	19.80

On suppose que cette durée de vie notée  $D \sim \mathcal{W}(\beta, k)$  de densité à support sur  $\mathbb{R}_+$ .

1. On suppose  $k$  connu, donner un estimateur de  $\beta$
2. Estimer  $\beta, k$  en utilisant la méthode du maximum de vraisemblance.

### Exercice 2 : Suivi de sinistres des clients d'une banque

Le fichier `incident.csv` contient des informations concernant des clients fragiles d'une institution financière. Un chargé d'études statistiques de l'organisme de crédit cherche à estimer la durée moyenne en mois entre le moment où un client a été jugé fragile et la survenance du premier incident de paiement. On suppose que ces durées sont distribuées normalement.

1. Importer le fichier de données et effectuer une analyse statistique (statistiques descriptives et représentations graphiques) des différentes variables.

3

2. Peut-on affirmer au risque  $\alpha = 5\%$  que la durée moyenne avant la survenance du premier incident de paiement est :

- supérieure à 5 mois ?
- inférieure à 7 mois ?

En déduire un encadrement de la durée moyenne avant le premier incident de paiement. Donner un intervalle de confiance de niveau 95% de la durée moyenne avant la survenance du premier incident de paiement. Discuter.

La population est constituée de deux catégories de clients, les clients dit “surendettés” et les clients dit “restructurés”.

3. Donner des intervalles de confiance de niveau 99% de la moyenne et de l'écart-type de la durée avant la survenance du premier incident de paiement pour chaque type de client.

4. Peut-on dire qu'il y a une différence entre la variance des durées avant la survenance du premier incident de paiement pour les clients surendettés et les clients restructurés ?

5. Peut-on affirmer qu'il y a une différence entre la durée moyenne avant la survenance du premier incident de paiement pour les clients surendettés et les clients restructurés ?

### Exercice 3 : Données accouchements prématurés

Le fichier `prematures.csv` contient des observations des facteurs prénataux (médicaux et de personne) liés à un accouchement prématuré chez des femmes déjà en travail prématuré.

Les variables du fichier sont les suivantes :

GEST : l'âge gestationnel en semaines à l'entrée dans l'étude

DILATE : la dilatation du col en cm

EFFACE : l'effacement de col (en pourcentages)

CONSIS : la consistance du col (1=mou, 2=moyen, 3=ferme)

CONTR : la présence (=1) ou non (=2) de contraction

MEMBRAN : les membranes ruptures (=1) ou non (=2) ou incertain (=3)

AGE : âge de la patiente

GRAVID : la gestite (nombre de grossesses antérieures y compris celle en cours).

DIAB : la présence (=1) ou non (=2) d'un problème de diabète, ou valeur manquante (=3).

TRANSF : transfert (1) ou non (2) dans un hôpital de soins spécialisés

GEMEL : grossesse simple (=1) ou multiple (2)

PREMATURE : accouchement prématuré (positif ou négatif)

1. Effectuer une analyse statistique (statistiques descriptives et représentations graphiques) des variables du jeu de données.

2. Estimer à partir des données brutes la proportion de femmes ayant une période de gestation de moins de 31 semaines.

3. Peut-on considérer que la moitié des femmes ayant déjà connues un accouchement prématuré ont une période de gestation inférieure à 31 semaines ? Donner la région critique du test, sa p-valeur.

- 1
4. La durée moyenne de gestation est-elle différente entre les femmes ayant déjà connues un accouchement prématuré et celle qui n'en ont jamais eu ?
- 2
5. Existe-il une relation entre :
- l'accouchement prématurée et la type de grossesses (gémellaires oui/non) ?
  - Les membranes rupturées et la consistance du col ?