

STATISTIQUE DESCRIPTIVE – 1^{RE} ANNÉE¹

FEUILLE DE T.P. 3

Objectifs. Statistique univariée - synthèse. Premier T.P. noté.

Instructions. Vous travaillerez sur les données *lowbw.sta*. Copiez le fichier de données depuis *moodle* sur votre compte. Le compte rendu devra être déposé sur *moodle* pour le **mercredi 14 novembre**, avant minuit. Il devra être nommé :

TP3_Prénom_Nom.pdf

ATTENTION

- Aucun retard ne sera toléré.
- Pour des problèmes de compatibilité, les formats autres que “.pdf” ne seront pas acceptés.
- Vous devez **rédigier** votre compte-rendu **individuellement** pour ce T.P. noté, même si vous avez travaillé en binôme lors de la séance.

Rappels. Vous devez expliquer votre démarche et commenter les résultats obtenus. Il est fortement conseillé de relire le document *Conseils de rédaction pour les comptes rendus de T.P.* pour plus de détails.

Description du jeu de données

Le but de cette étude était d’identifier des facteurs de risque pour expliquer la naissance de bébés de faible poids (c’est à dire de poids inférieur ou égal à 2500 grammes). Les données ont été collectées au centre médical Baystate, à Springfield, dans le Massachusetts, en 1986, sur 189 femmes. On pensait que quatre variables pouvaient être importantes : l’âge, le poids de la mère à la dernière période menstruelle, la couleur de peau, et le nombre de visites chez un médecin pendant le premier trimestre de la grossesse. Pour chaque femme, les valeurs des 10 variables suivantes ont été collectées :

- id : identifiant
- age : âge de la mère en années
- lwt : poids de la mère à la dernière période menstruelle en kilogrammes
- race : couleur de peau de la mère
- smoke : tabagisme ou non pendant la grossesse
- ptl : antécédents d’accouchements prématurés
- ht : antécédent d’hypertension chez la mère
- ui : irritabilité utérine chez la mère
- ftv : nombre de visites chez le médecin pendant le premier trimestre de grossesse
- bwt : poids de naissance du bébé en grammes.

1. Enseignants responsables des TP : G. Chagny (gaelle.chagny@parisdescartes.fr) et C. Laclau (charlotte.laclau@parisdescartes.fr)

Objectifs

Votre travail consiste à explorer le jeu de donnée *lowbw.sta*. L'objectif est double :

Dans un premier temps, vous devrez décrire la population dans son ensemble.

Dans un second temps, vous chercherez à comprendre les différences entre les femmes ayant accouché de bébé de faible poids et celles ayant accouché de bébé de poids normal. A partir de votre étude, peut-on faire des hypothèses sur les variables qui peuvent avoir une influence sur la naissance de bébés de faible poids ? Peut-on proposer un "profil" de la mère ayant un bébé de faible poids et de la mère ayant un bébé de poids normal dans cette étude ? *Indication* : pour cette seconde étape, vous pourrez commencer par créer une nouvelle variable qualitative, prenant deux modalités, indiquant si le bébé est de faible poids (c'est-à-dire de poids $\leq 2500g$) ou de poids normal.

Consignes

Vous étudierez les variables qui vous semblent pertinentes, à l'aide des outils adaptés. Vos analyses et conclusions doivent s'appuyer sur des tableaux et des graphiques que vous présenterez et commenterez convenablement dans votre compte-rendu. Vous choisirez les éléments de vos analyses qui vous semblent les plus utiles pour répondre aux deux problèmes posés.

Les seules contraintes de ce T.P. sont les suivantes :

- vous devrez réaliser un histogramme de l'âge des mères, où l'âge est regroupé en classes d'amplitude 4 ans ;
- vous devrez, pour une variable que vous choisirez, tracer une boîte à moustaches avec *STATISTICA*, et la commenter précisément : notamment, vous rappellerez la signification de chaque élément de ce graphique. Pour la même variable, vous devrez tracer aussi sur un même graphique les boîtes à moustaches correspondant aux deux groupes étudiés dans la population (femmes ayant accouché de bébés de faible poids ou femmes ayant accouché de bébés de poids normal).
- vous devrez pour une variable quantitative discrète de votre choix, tracer la fonction de répartition empirique associée. Faites de même pour une variable quantitative continue.