

DM M1 Maths-fi - à rendre pour le 6 décembre 2011

Les programmes sont à m'envoyer à mon adresse mail `chi.tran@univ-lille1.fr` (ce n'est pas la peine de les recopier dans votre copie).

Exercice 1 (Dettes publiques en 2010)

La table `dettes-pib2010.csv` est disponible sur :

<http://labomath.univ-lille1.fr/~tran/enseignements.html>

On s'intéresse aux dettes publiques en 2010 pour 127 pays. Dans les données, les individus sont les pays, les variables sont :

- `id` et `num` sont des identifiants dont on n'a pas besoin.
- `pib2010` : Le produit intérieur brut (PIB), en millions de dollars américains.
- `dette` : Les dettes publiques exprimées en pourcentage du PIB et recueillies sur :
http://fr.wikipedia.org/wiki/Liste_de_pays_par_dette_publicue
La définition retenue pour déterminer la dette publique est celle du *CIA World Factbook*.
- `UE` : qui indique si le pays est dans l'Union Européenne ou non. Pour les pays de l'UE, cette variable vaut 2 pour les pays qui bordent la Méditerranée, 1 sinon. On appellera les pays qui bordent la Méditerranée "pays du Sud de l'UE".
- `EUR` : indique pour les pays de l'UE s'ils sont passés à l'euro ou non.
- `entree` : indique pour les pays de l'UE l'année d'entrée dans l'UE.

1. Télécharger les données sur votre ordinateur, dans le répertoire `chemin`. Ouvrez R et lisez la base de données à l'aide de la commande :

```
donnees<-read.table(chemin,sep=";", header=TRUE)
```

Partie A : description de la dette publique au niveau mondial

2. Faire des statistiques descriptives et décrire la distribution des variables `dette` et `pib2010` (indicateur de tendance centrale et de dispersion, asymétrie et aplatissement de la distribution, minimum, maximum et commentaires).

3. Créer dans un vecteur `vectdettes` de taille 127 le montant des dettes en millions de dollars américains. On parlera dans la suite de la variable `dette` en valeur.

4. A quel quantile correspond la dette de la France en pourcentage de PIB ? en valeur ?

5. Tracer sur R la fonction de répartition empirique de la dette en valeur.

6. Placer une ligne verticale correspondant à la Grèce. Quel est le quantile correspondant ?

Partie B : Inégalité de l'endettement en Europe ?

On considère les 27 pays de l'Union Européenne (UE).

7. Calculer la dette moyenne, en valeur, des pays de l'UE. Comparer avec la question 1.

8. Quelle est la dette moyenne, en valeur, des pays du Sud de l'UE ? Quelle est la dette moyenne des autres pays ? Faire une analyse de la variance pour dire si l'écart de la dette moyenne peut s'expliquer par les caractères géographiques.

9. reprendre les questions 7 et 8 mais pour la dette en pourcentage.

Partie C : Liens entre le passage à l'euro et l'endettement

10. On veut savoir si le passage à l'euro est lié à l'endettement. Créer la table de contingence qui croise le fait d'être passé à l'euro ou non (EUR=1 ou EUR=2) et le fait d'avoir ou non une dette plus grande que 75% du PIB.

11. Calculer les fréquences marginales de la table de contingence précédente.

12. Suite aux questions 10 et 11, faire un test du χ^2 pour dire si le fait d'avoir un niveau d'endettement supérieur ou inférieur au PIB est ou non indépendant du fait d'être passé à l'euro. Quelle est la loi de ξ sous l'hypothèse H_0 d'indépendance ? Sous l'hypothèse alternative ? Conclusion ?

Exercice 2 (Prix du logement)

Le fichier `logements.txt` est téléchargeable sur

<http://math.univ-lille1.fr/~tran/enseignements.html>.

Nous disposons, pour un échantillon de $n = 30$ appartements de 3 pièces mis en location dans un même quartier de Paris :

- du loyer mensuel Y (en francs),
- de la surface X (en m^2).

L'échantillon est noté $(X_i, Y_i)_{i \in [1,30]}$.

1. Préciser la population, les variables (donner leur nature).
2. Décrire la distribution des deux variables X et Y (faire une étude avec les logiciels à votre disposition et résumer les résultats qui vous semblent intéressants).
3. Représenter graphiquement le loyer Y en fonction de la surface X . Premier commentaire ? Un modèle linéaire semble-t-il approprié ?

On s'intéresse au modèle $Y = aX + b + \varepsilon$, où a, b sont des coefficients à déterminer et où ε est un bruit gaussien centré et d'espérance σ^2 .

4. Quelle est la variable expliquée, quelle est la variable explicative ?
5. Estimer les coefficients a et b du modèle par la méthode des moindres carrés ordinaires :
 - en utilisant l'une des fonctions programmées dans **R**,
 - en calculant directement les coefficients à l'aide de leur expression théorique.Quelle est l'interprétation de a ? de b ? Commenter les valeurs numériques obtenues.
6. Superposer au nuage de points dessiné à la question 3 la droite de régression obtenue.
7. Calculer le coefficient de détermination. Commentaire ?
8. Etudier les hypothèses faites sur les résidus :
 - Calculer leur moyenne et leur variance empirique.
 - Donner l'estimateur MCO de la variance des résidus.
9. Tester avec le logiciel de votre choix la nullité du coefficient b .