Université de Nice — L1MASS, année 2014-2015
Département de Mathématiques — Statistique
NOM : — Date : .
PRENOM : — Salle : .

Feuille-réponse 5
Régression et simulation linéaire

1 Reprise du cours 05

Former le vecteur `x` des tailles `Wr.hnd` du fichier `survey`, comme dans le cours :

```
library(MASS)
data(survey)
survey.cc<-survey[complete.cases(survey), ]
x<-survey.cc$Wr.Hnd
```

et former de même les vecteurs `y` et `z` des tailles `NW.Hnd` et `Height`.

1. Former un data.frame appelé `MesDonnées` regroupant `x`, `y`, et `z` ; quelles sont les trois premières lignes de `MesDonnées`

2. Former le "scatter-plot" des échantillons `x`, `y`, et `z`. Quels sont les nuages qui sont le mieux regroupés autour d'une droite ?

3. Faire dessiner (`plot`) le nuage des `x` et `y` ; que provoque le fait de choisir `pch=17` ? Recommencer avec `y` et `z`, dans cet ordre.

4. Former le modèle de régression linéaire `DteReg_z_sur_y` des `z` sur les `y`. Quelle est l'équation $z = ay + b$ que vous obtenez ? Quelle commande avez-vous utilisé.

5. Ajoutez la droite de régression à votre nuage. Quelle commande utilisez-vous ?

6. Ajoutez le "centre de gravité" de votre nuage. Appartient-il à la droite de régression ? Quelle commande utilisez-vous ?

7. Former les valeur `z2y` prédites par votre modèle linéaire. Quelle commande utilisez-vous ?

8. Ajoutez les (`y`,`z2y`) à votre dessin. Quelle commande utilisez-vous ?

9. Représentez **dans une nouvelle fenêtre** l'histogramme des résidus `z-z2y`. Esquissez ici le résultat ; commentez. Donnez ici la moyenne et l'écart-type de ces résidus.

2 Simulation

1. Formez un échantillon `Residu` de même taille que `z`, normal, centré, et d'écart-type 8. Commencez par la commande `set.seed(sssmm)` où `sssmm` désigne votre machine. Quelle commande utilisez-vous ensuite ?

2. Quelle est la moyenne et l'écart-type de l'échantillon `Residu` ; tracez son histogramme ; commentez.

3. Posez `zSim=3*y+118+Residu` ; expliquez le nom donné à ce nouvel échantillon.

4. Quelle est l'équation de la droite de régression des `zSim` sur les `y` ?

5. Quelle est la moyenne et l'écart-type des résidus de cette régression ?

6. Représentez le nuage des (`y`,`zSim`) et la droite de régression obtenue. Ajoutez la droite d'équation $z = 3y + 118$ au moyen de la commande `abline(c(118,3))`. Commentez le résultat obtenu.

.