

Gaspard Palay / Regression Lineaire Simple ET Multiple

Gaspard Palay

29/01/2021

1. Critères d'évaluation

1. Comportement du Rmd lors de son exécution
2. Qualité de la rédaction du dossier
3. Didactisme et pertinence du dossier
4. Qualité et lisibilité du Rmarkdown
5. Qualité des explications du modèle mathématique

2. Lien vers le document commenté

En cliquant **ici**, vous trouverez le lien menant au GitHub de Nina ZOUMANIGUI hébergeant le fruit de sa réalisation.

3. Auteurs du document commenté

Le document évalué dans le cadre de ce rendu a été produit par Nina ZOUMANIGUI, étudiant en MSc Data Management à Paris School of Business.

4. Synthèse du document

Le document est une production RMD présentant les deux différents modèles de régression, à savoir : - La régression linéaire simple - La régression linéaire multiple

L'auteur commence par définir ce qu'est un modèle de régression linéaire et le comparer textuellement aux autres méthodes statistiques. Il explique au lecteur que ce modèle sert à établir un lien de corrélation entre une variable quantitative et une ou plusieurs variables quantitatives à l'aide d'un modèle statistique

5. Extrait commenté des parties de code

L'auteur écrit son document sur un RMD qu'il n'extrait pas en PDF. Il utilise le dataset Examen.csv afin d'étudier l'âge et le niveau d'éducation pour étudier l'attention des patients âgés. Il conduit une régression linéaire multiple sur ce jeu de données et le commente :

définir les variables

```
y=g$Age  
x=g$Elevel  
plot(y,x)
```

Autre application

```
data=read.csv("C:\\Users\\ninaz\\OneDrive\\Bureau\\R\\mtcars.csv",sep = ";", header= T)
attach(data)
head(data)
```

```
plot(mpg~wt,pch=20)
fit= lm(mpg~wt,data=data)
fit
abline(fit,col="red",lwd=2)
```

Il définit ensuite le modèle de régression multiple par : “est une généralisation du modèle de régression simple lorsque les variables explicatives sont en nombre fini”

Et suppose ensuite que les données collectuées suivent le modèle suivant : $y_i = b_0 + b_1x_{i1} + b_2x_{i2} + \dots + b_px_{ip} + \varepsilon_i, i = 1, \dots, n$.

6. Evaluation du travail suivant les 5 critères précités

1. Comportement du Rmd à l'exécution

L'auteur ne fournit pas le jeu de données sur lequel il travaille, impossible donc pour le lecteur de reproduire son RMD et de ploter ses résultats comme il le fait. Cependant, le RMD s'exécute bien et le fichier est clair et ordonné.

2. Qualité de la rédaction du dossier

La rédaction de ce document est de bonne qualité. Il est bien structuré, et les étapes de progression sont respectées. L'auteur ne réalise pas de régression linéaire simple sur son jeu de données. Il est cependant bien écrit, ne comporte pas de fautes d'orthographe et est titré par parties.

3. Didactisme et pertinence du dossier

La lecture de ce dossier est aisée et accessible. Il est didactique. Ce fichier est pertinent puisqu'il étudie les deux modèles de régression linéaires : simple et multiple, sur une même production. Il ne s'intéresse cependant pas à la régression logistique.

L'autrice arrive à faire adhérer le lecteur à sa production. Le document a vocation à transmettre une connaissance généraliste à un novice en statistique; Le but est atteint même si cela nécessite des recherches tierces si le lecteur souhaite approfondir sa connaissance.

4. Qualité et lisibilité du RMD

Le fichier est lisible et accessible sur son GitHub. Le fichier n'a pas été extrait en pdf ce qui empêche le lecteur de lire les plots de son modèle de régression multiple. Il est lisible et aéré.

5. Qualité des explications du modèle mathématique

Les applications sont simples. L'auteur n'a pas utilisé latex pour rédiger son expression mathématique. Il manque peut-être un peu d'explication algébriques du modèle de régression.

7. Conclusion

Selon moi, il s'agit globalement d'un bon travail.

L'auteur synthétise et définit le modèle de régression. Il s'appuie sur un dataset pour effectuer une régression logistique. La rédaction du dossier est claire et limpide. Sans artifice mais efficace.

Vous retrouvez ce document sur mon **GitHub**.