DÉPARTEMENT DE MATHÉMATIQUES ET DE GÉNIE INDUSTRIEL

MTH2302D - PROBABILITÉS ET STATISTIQUE

Devoir - Hiver 2025

Date de remise : 15 avril avant 23h59 (dans Moodle)

Veuillez remplir le tableau suivant et joindre cette page à votre rapport.

Identification de l’étudiant.e 1

|  |  |
| --- | --- |
| Nom : Saka | Prénom : Ndzana Missia Said |
| Groupe : 02 | Matricule : 2184805 |

Identification de l’étudiant.e 2

|  |  |
| --- | --- |
| Nom : Shair Zaie | Prénom : Safiullah |
| Groupe : 02 | Matricule : |

Placer les deux fichiers DevoirD\_H25.csv et charger.R dans le répertoire de travail de R. En utilisant votre matricule, exécuter ensuite (dans cet ordre) les deux commandes suivantes dans R pour générer votre ensemble de données personnalisées ’mondata’ :

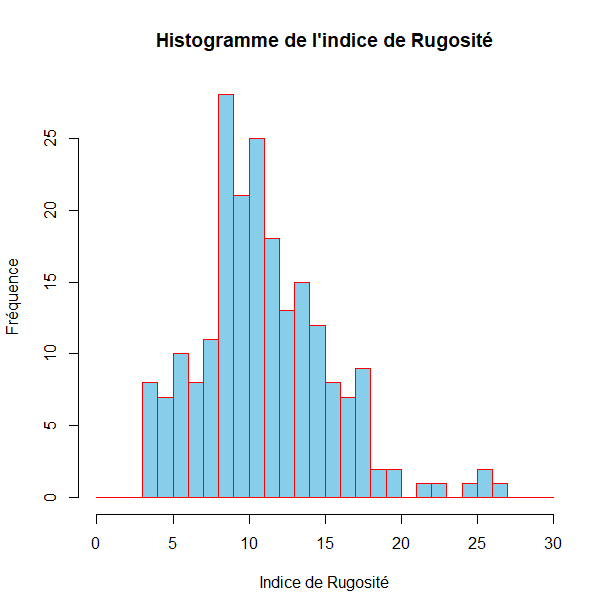
source(’**charger.R**’) mondata <- charger(**matricule**)

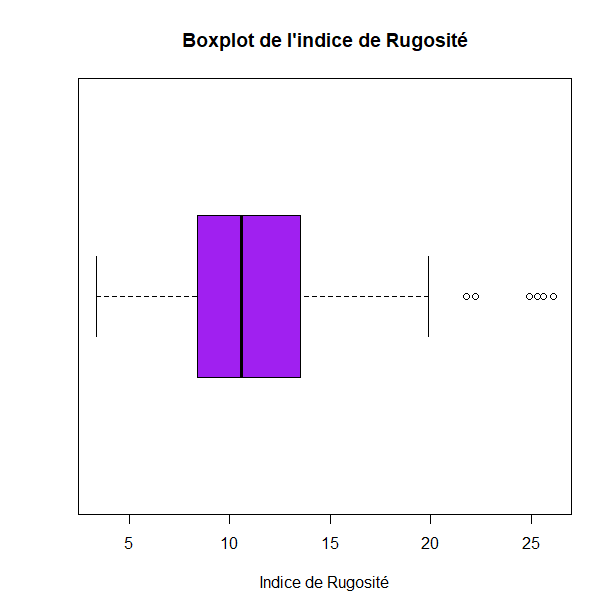
|  |  |
| --- | --- |
| Question | Note |
| a) | /7 |
| b) | /10 |
| c) | /15 |
| d) | /6 |
| Présentation | /2 |
| TOTAL | /40 |

Mardi le 15 avril 2025

**Phase 1 : Analyse statistique descriptive et inférence**

1. **Interprétation des résultats obtenus pour :**

* **Histogramme de Tukey :** Nous remarquons une grande quantité de matériaux dont l’IR est comprise 8 et 12. Par ailleurs, peu de matériaux on un IR supérieur à 25.
* **Diagramme de Tukey :** On conclut que l’IR est majoritairement concentré entre 8 et 12**.**

****

* **Droite de Henry :**

Les points s'écartent significativement de la droite de référence aux extrémités (pour les valeurs extrêmes de l'IR).

Cela confirme une distribution asymétrique, avec une queue de distribution plus épaisse que celle d'une loi normale (présence d'outliers ou de valeurs extrêmes).

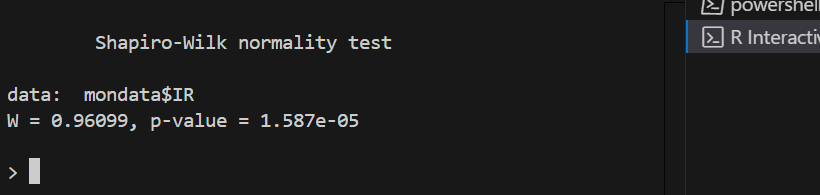
**Une image contenant texte, diagramme, ligne, capture d’écran

Le contenu généré par l’IA peut être incorrect.**

* **Test de normalité :**

D’après le test de normalité**: w = 0,96099 et p-value= 1.587e-5.**

**p-value < 0.05,** par conséquent, le test de Shapiro-Wilk confirme que l'IR n'est pas normalement distribué.

****

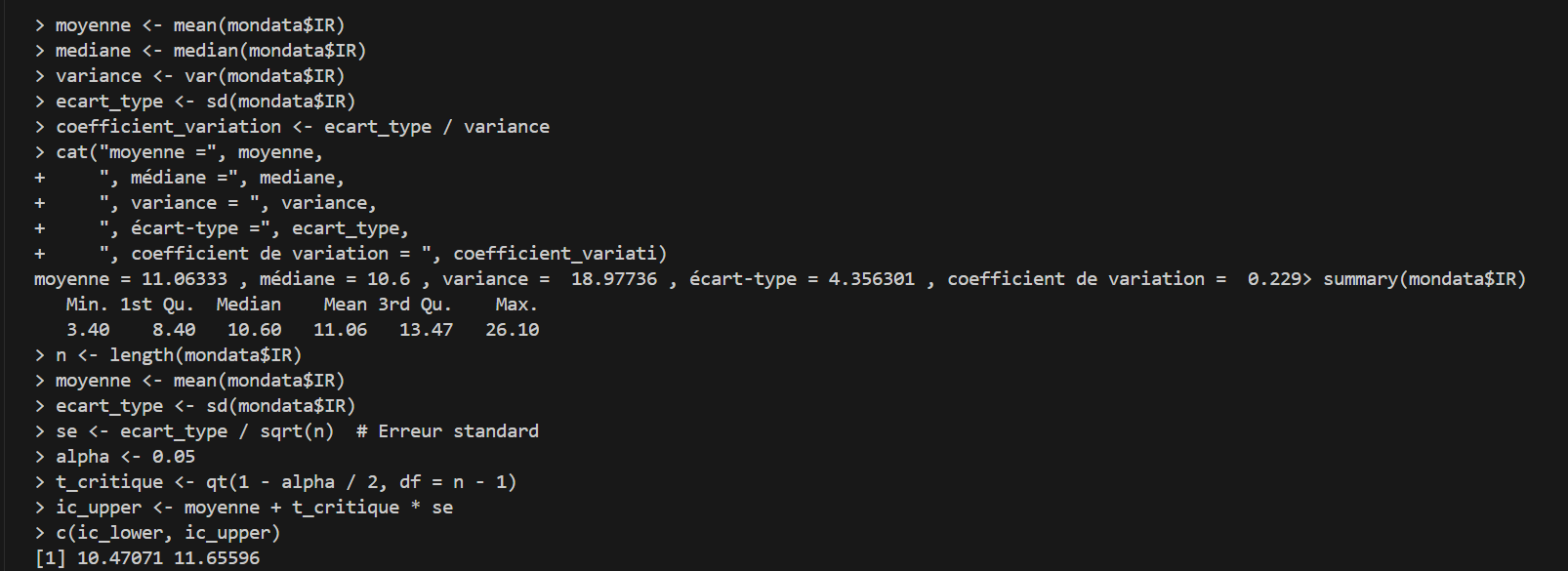
* **Tableau de statistiques descriptives :**

|  |  |
| --- | --- |
| **Moyenne** | 11.06333 |
| **1ER Quartile** | 8.40 |
| **Médiane** | 10.60 |
| **3ème Quartile** | 13.47 |
| **Écart-type** | 4.356301 |
| **IDC pour la moyenne** | [10.4707, 11.6559] |

On remarque que :

 La rugosité moyenne des trous percés est de 11.06;

50 % des trous ont un IR ≤ 10.60 ;

 L’écart-type étant de 4.36 alors il y a une dispersion modérée autour de la moyenne

1. **Conclusion à la suite de l’analyse des :**

Le tableau des statistiques montre que :

* Le matériau 0 a une rugosité légèrement plus élevée (11,39) que le matériau 1 (10,75), mais la différence est faible
* Les deux groupes ont des écarts-types similaires (~4.5), indiquant une variabilité comparable.
* Les intervalles de confiance (IC) se chevauchent largement :
  + Matériau 0 : [10.51, 12.27]
  + Matériau 1 : [9.94, 11.56]
* Les maximas sont proches (26.1 vs 25.6), suggérant des défauts de perçage dans les deux matériaux.

Par ailleurs, concernant les tests d’hypothèses nous avons :

* P-value = 0.2905 > 0.05 par conséquent, il n’y a pas de différence significative entre les moyennes de l’indice de rugosité entre les matériaux de type 0 et 1
* L’intervalle de confiance de la différence est : [-0.55, 1.82] inclut 0 donc il y a absence d’effet statistique

En outre, l’analyse des tests de normalité (Shapiro-Wilk) montre que pour:

* Le matériau 0 : p-value = 0.021 < 0.05 d’où on rejette la normalité
* Le matériau 1 : p-value = 0.0001 < 0.05 d’où on rejette aussi la normalité

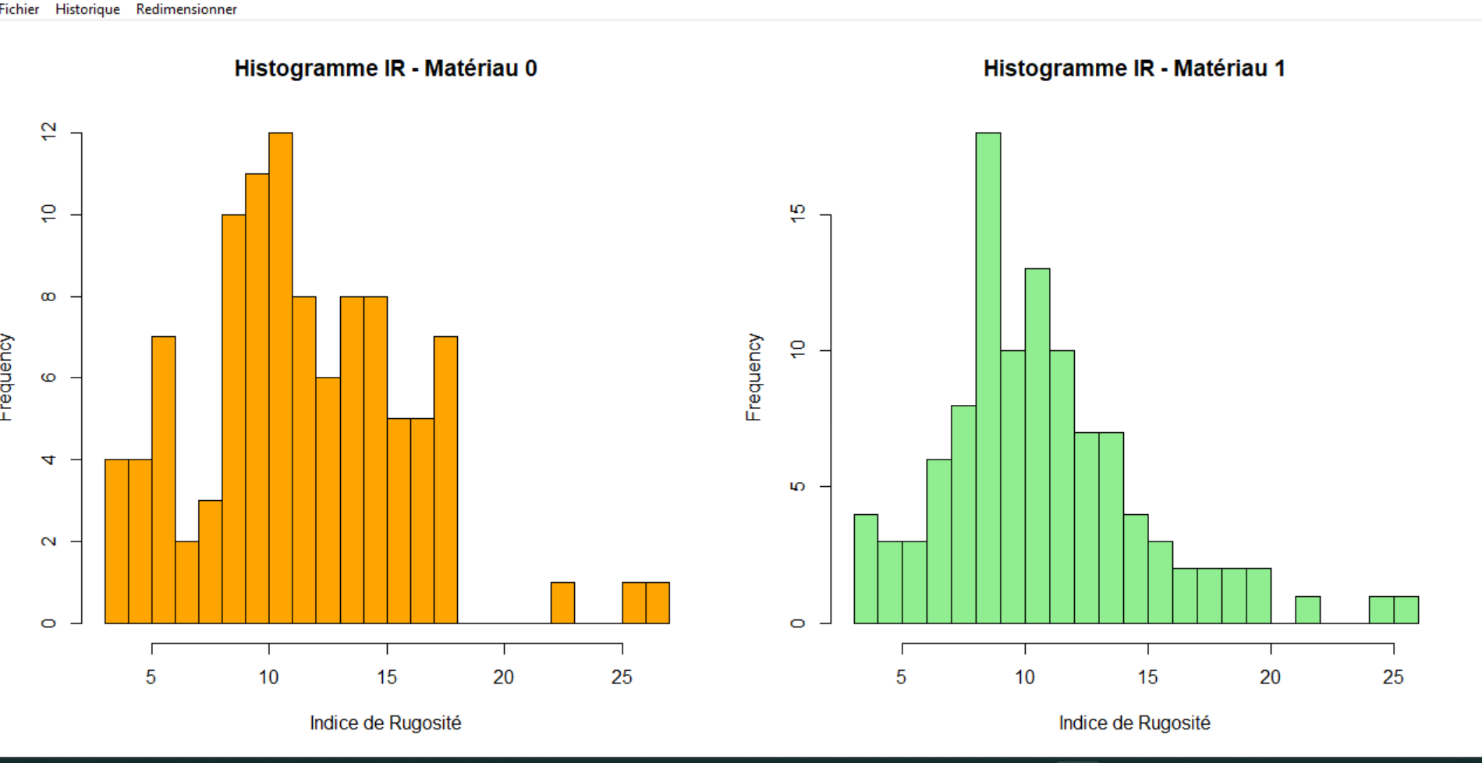
Conclusion :

Il faudra analyser les paramètres de perçage (vitesse, température) afin de déterminer le meilleur modèle.

Une image contenant texte, capture d’écran, Police

Le contenu généré par l’IA peut être incorrect.Une image contenant texte, capture d’écran, Police, nombre

Le contenu généré par l’IA peut être incorrect.Une image contenant texte, capture d’écran, diagramme, ligne

Le contenu généré par l’IA peut être incorrect.

**Phase2 : Recherche du meilleur modèle**

1. **Conclusions Clés**

Variables significatives :

**La température (T)** montre un impact **hautement significatif** (p < 0.001) dans tous les modèles.

**La vitesse (V)** n'est **pas significative** (p > 0.4 dans tous les modèles).

Qualité d'ajustement :

Les modèles avec **T** expliquent **~63-65%** de la variance (R² ajusté).

Les modèles avec **V** sont inutiles (R² ≈ 0%).

Normalité des résidus :

Seuls les modèles basés sur **T** satisfont le test de Shapiro-Wilk (p > 0.05).

Comparaison par AIC :

**Le modèle 5 (puissance en T)** a l'AIC le plus bas (**21.34**), suivi du modèle 6 (**23.56**).

**Modèle à retenir** :

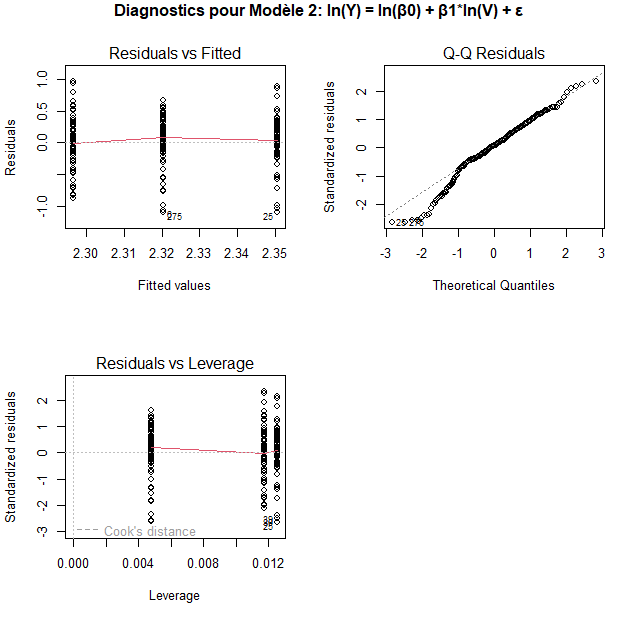
Par conséquent, le modèle à retenir est le **modèle 5 (puissance)** car il a les meilleurs AIC et résidus normaux. Son équation est :

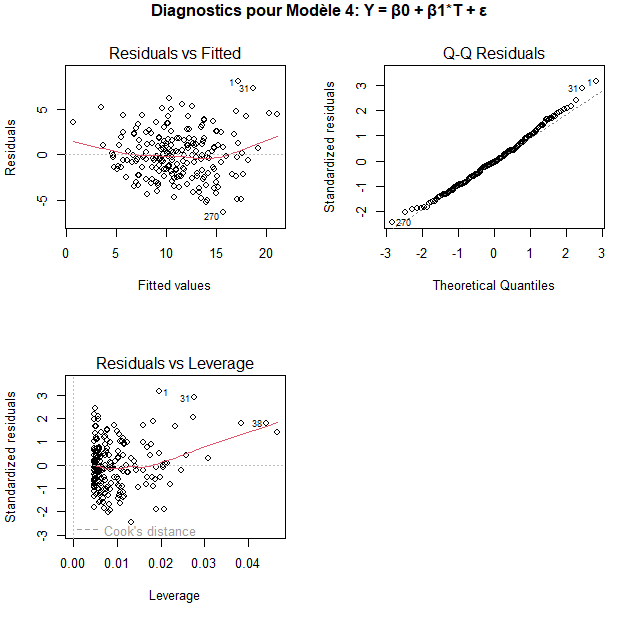
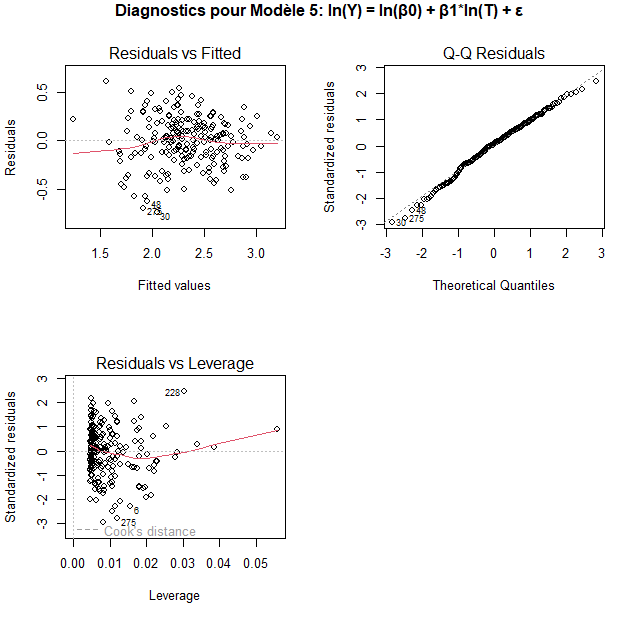
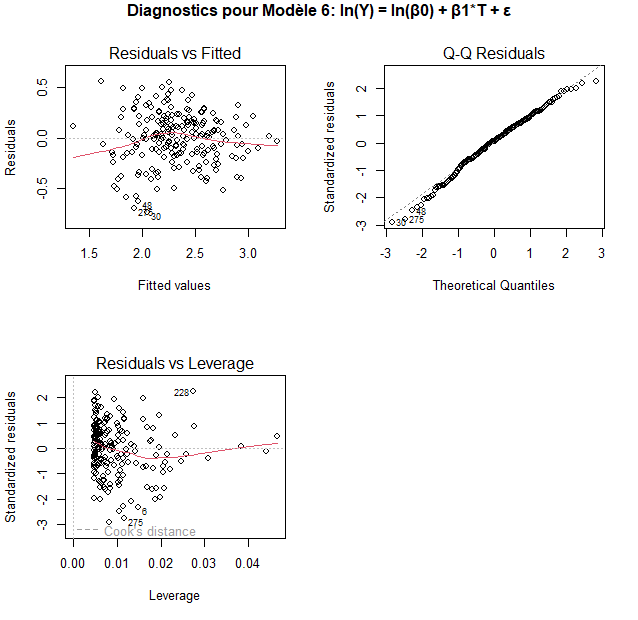
**ln (IR) = -15.92 + 4.90·ln(T) + ε**

**Preuve :**

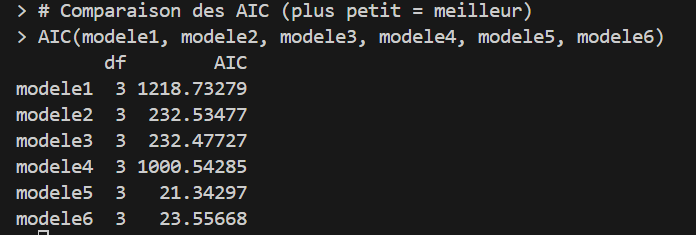
**Une image contenant texte, diagramme, capture d’écran, ligne

Le contenu généré par l’IA peut être incorrect.**



Une image contenant texte, diagramme, ligne, Parallèle

Le contenu généré par l’IA peut être incorrect.



1. Le modèle retenu étant le modèle 5 dont l’équation est ln (IR) = -15.92 + 4.90·ln(T) + ε nous avons :

Pour un perçage a une Température (T=40) :

Une valeur prédite de l’IR ≈ **10.2** ;

Un intervalle de confiance = [8.6, 12.1]

Commentaire :

La plage est relativement étroite, indiquant une bonne précision du modèle. L'IR prédit se situe dans la zone de qualité acceptable (car proche de la médiane globale de 10.6).

Le graphique de visualisation a été effectuer dans notre code charger.R afin de garantir une concordance avec les analyses que nous avions effectuées dans ce présent rapport.

Une image contenant texte, capture d’écran, ligne, diagramme

Le contenu généré par l’IA peut être incorrect.Une image contenant texte, capture d’écran, Police

Le contenu généré par l’IA peut être incorrect.