**Exercice3: Régression[4pts]**

1. La régression linéaire permet d’étudier les individus atypiques. Décrire les différents cas possibles d’anormalité. **[1pt]**

Atypique aberrant: valeur inhabituelle sur une variable ou combinaison de variables

Influent: pèse de manière exagérée dans la régression

Atypique régression: cible très mal expliquée, erreur très élevée

1. Quelles sont les méthodes pas à pas pour la construction de modèles réduits? Donner brièvement l’idée de chacune. **[1pt]**

méthode descendante ou élimination en arrière lorsqu’on élimine des variables

-méthode ascendante ou sélection en avant lorsque on ajoute des variables

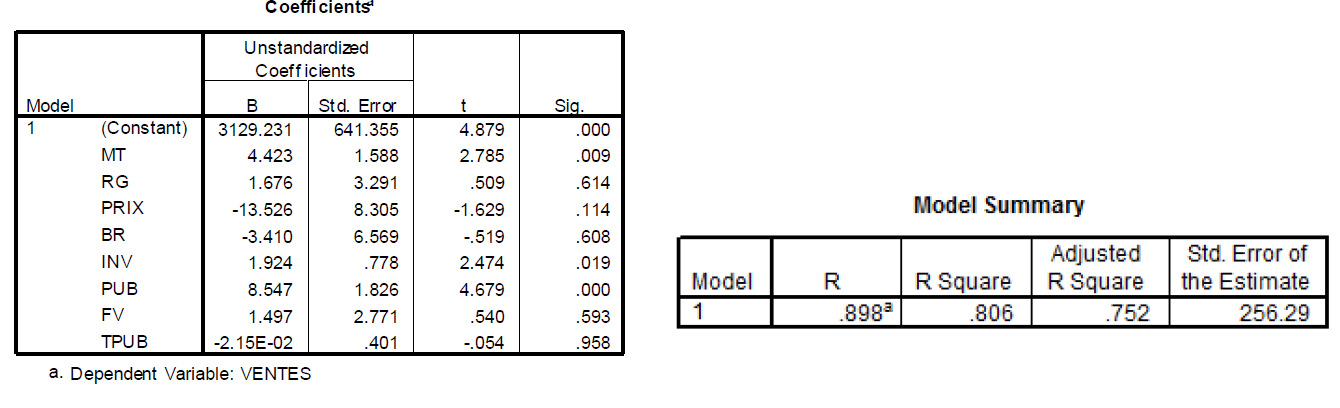
-méthode mixteest une combinaison de ces deux méthodes

**3.** Expliciter une méthode pour l’obtention des coefficients estimateurs dans la régression linéaire.**[1pt]**

Exercice2 : On propose de construire un modèle permettant de prédire les ventes semestrielles d’un produit agroalimentaire via la régression linéaire multiple.

Les variables explicatives sont: MT: Besoin du marché,RG: Remises aux grossistes,PRIX: Prix du produit, BR: Budget de Recherche, INV: Investissement, PUB: Publicité, FV = Frais de ventes, TPUB = Total budget publicité de la branche.

Les résultats du modèle sont affichés ci-dessus.



**1.** En étudiant individuellement les coefficients liés à chaque variable explicative, déterminer les variables significativement pertinentes dans l’augmentation des ventes semestrielles du produit. [1pt]

**PUB, MT ont les plus faibles valeurs significatives donc elles sont les plus pertinentes.**

**2.** En se basant sur les résultats des coefficients de détermination, conclure sur la qualité du modèle construit. [1pt]

**R² = 0.806 donc il s’agit d’un bon modèle.**

**3.** On propose de sélectionner les variables pertinentes du modèle en utilisant la méthode de sélection pas à pas descendante (Backward). Expliquer brièvement les étapes de cette méthode. [2pts]

**AIC, le critère de choix du modèle le plus performant : itération par itération on élimine la variable la moins pertinente.**

**4.** Donner l’équation du modèle permettant de prédire les ventes semestrielles en fonction de toutes les variables explicatives. [2pts]

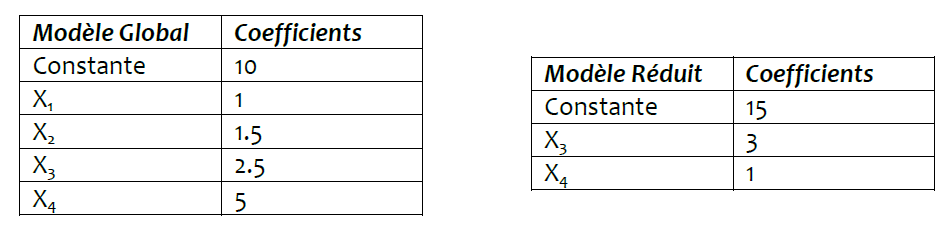
**VENTES = 3129,231 + 4,423 MT + 1,676 RG - 13,526 PRIX – 3,410 BR +1,924 INV + 8,328 PUB + 1,497 FV – 0,00215 TPUB**

**Exercice 3: Régression Linéaire Multiple[5pts]**

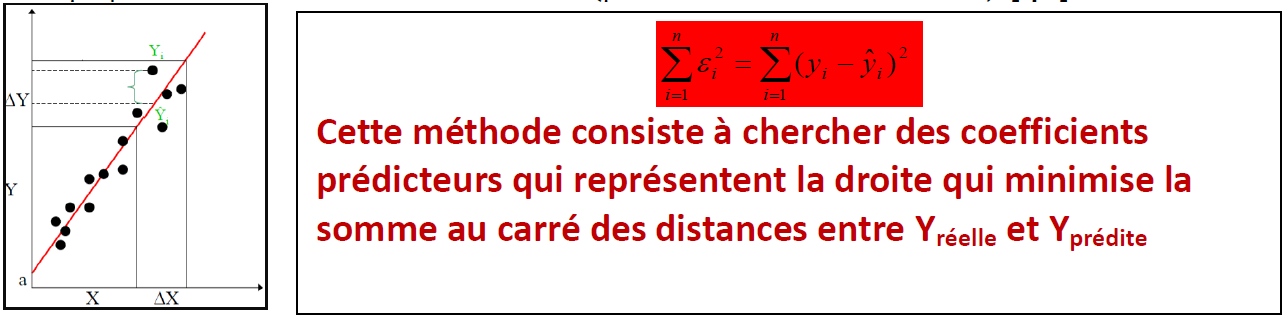
On propose de construire le meilleur modèle permettant de prédire une variable Y via la régression linéaire multiple.

Les variables explicatives sont:X1, X2, X3 et X4.

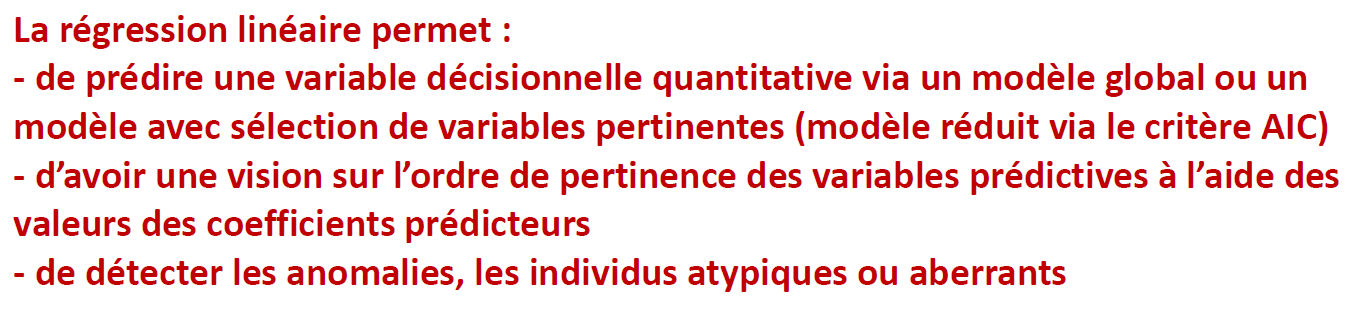
Les résultats des modèles global et réduit sont affichés ci-dessous.



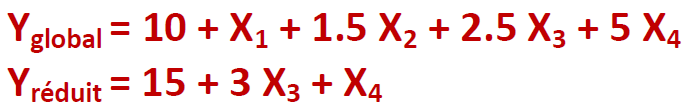
**1.** Expliquer la méthode des **moindres carrées** (possibilité d’utiliser un schéma)



**2.** Décrire l’**utilité** de la régression linéaire. **[2pts]**



1. Donner la **formule** du modèle global et celle du modèle réduit. **[1pt]**



**4.** Calculer les **prédictions** d’un individu I(X1=1, X2=1,X3=1,X4=2).

1. Sachant que la valeur réelle de la variable décisionnelle de l’individu I est **Yréelle = 20**, Comparer les deux prédictions obtenues et **Conclure**. **[0.5pt]**

