

## Modèles linéaires en actuariat – Exercices – Série 1

### Régression linéaire simple

1. Vous disposez des observations suivantes pour construire un modèle de prévision du salaire de vos employés en fonction de leur nombre d'années d'expérience de travail :

Années d'expérience	Salaire annuel
3	41 185
6	47 917
10	57 675
5	49 789
2	41 595
25	96 901
7	53 840
8	57 712
10	58 027
32	111 762

- Dans le contexte de la régression linéaire simple, définir la variable dépendante (ou réponse)  $Y$ , ainsi que la variable indépendante (ou prédictive/exogène)  $X$ .
- Utiliser Microsoft Excel ou R pour produire un graphique de  $Y$  en fonction de  $X$ .
- Écrire l'équation du modèle de régression linéaire simple et identifier les paramètres que l'on cherche à estimer.
- Démontrer de façon générale comment retrouver les estimateurs des paramètres du modèle de régression linéaire simple via la minimisation de la somme résiduelle des carrés.
- Trouver les estimateurs de l'ordonnée à l'origine et de la pente de la droite de régression dans le contexte salaire vs années d'expérience.
- Ajouter la droite de régression sur le graphique tracé précédemment.
- Identifier le point  $(\bar{X}, \bar{Y})$  sur le graphique tracé précédemment. Commenter sur l'emplacement de ce point par rapport à la droite de régression.
- Donner une interprétation aux coefficients de régression  $\beta_0$  et  $\beta_1$  du modèle de régression linéaire simple.
- Afin d'évaluer le « *parameter risk* » (incertitude liée à l'erreur d'estimation des paramètres), calculer  $Var(\hat{\beta}_0)$  et  $Var(\hat{\beta}_1)$ .

2. Utiliser les données et les résultats de l'exercice 1 de cette section pour répondre aux questions suivantes :
- Produire un tableau d'analyse de la variance (ANOVA).
  - Calculer le coefficient de détermination  $(R^2)$ , donner une explication de ce que cette statistique mesure, puis commenter la qualité du modèle de régression « Salaire » vs « Expérience ».
  - Obtenez un intervalle de confiance pour  $\beta_0$  et  $\beta_1$  au niveau de confiance  $(1 - \alpha) = 95\%$ .
  - Tester l'hypothèse  $H_0 : \beta_0 = 35000$  à l'aide de la statistique de Student avec un niveau de confiance de 95%.
  - Tester l'hypothèse  $H_0 : \beta_1 = 0$  à l'aide de la statistique de Student avec un niveau de confiance de 95%.
  - Tester la validité de la régression en utilisant la statistique F de Fisher avec un niveau de confiance de 95%. Comparer ce test avec le test effectué en (e).
  - Obtenez un intervalle de confiance pour la valeur moyenne, ainsi que pour la valeur prédite du salaire d'un employé ayant cumulé 15 ans d'expérience.