

## Modèles linéaires en actuariat – Exercice 1 – Série 4

Vous disposez de l'ensemble des données pour modéliser le nombre d'accidents de la route mortels par année au Québec ( $Y_t$ ) en fonction des investissements totaux du gouvernement en prévention (en million de dollars) par année ( $X_{t,1}$ ), du coût moyen des contraventions (en dollars) dans l'année ( $X_{t,2}$ ), et de la quantité totale de neige tombée (en cm) durant l'hiver ( $X_{t,3}$ ) :

| $Y_t$ | $X_{t,1}$ | $X_{t,2}$ | $X_{t,3}$ |
|-------|-----------|-----------|-----------|
| 685   | 1         | 150       | 325       |
| 573   | 1         | 350       | 285       |
| 629   | 2         | 225       | 300       |
| 567   | 2         | 525       | 410       |
| 552   | 2         | 100       | 222       |
| 500   | 3         | 155       | 271       |
| 580   | 3         | 235       | 316       |
| 260   | 4         | 455       | 251       |
| 437   | 4         | 382       | 321       |
| 327   | 5         | 125       | 210       |

De plus, on a que

$$(\mathbf{X}'\mathbf{X})^{-1} = \frac{1}{100\,000\,000} \begin{bmatrix} 637276565,70 & -52944107,80 & 335951,44 & -1975615,09 \\ -52944107,80 & 8657778,98 & -32733,63 & 131957,17 \\ 335951,44 & -32733,63 & 902,67 & -1688,33 \\ -1975615,09 & 131957,17 & -1688,33 & 7129,91 \end{bmatrix},$$

et que

$$\mathbf{X}'\mathbf{Y} = \begin{bmatrix} 5110 \\ 12417 \\ 1337609 \\ 1522631 \end{bmatrix},$$

où  $\mathbf{X}$  représente la matrice schéma du modèle de régression complet (avec ordonnée à l'origine  $\beta_0$ ) associée à ces données.

- (a) Valider les calculs de  $(\mathbf{X}'\mathbf{X})^{-1}$  et de  $(\mathbf{X}'\mathbf{Y})$  avec un logiciel comme Excel ou R.
- (b) Obtenir  $\hat{\beta}$  pour cette régression.
- (c) Effectuer un test de Student pour la significativité de chacune des variables dans le modèle, c'est-à-dire :

$$H_0 : \beta_i = 0$$

$$H_1 : \beta_i \neq 0, i = 1, 2, 3.$$

- (d) Obtenir le tableau ANOVA pour cette régression.
- (e) Tester la validité globale de cette régression à l'aide du test F de Fisher.
- (f) Calculer le coefficient de détermination du modèle ( $R^2$ ).
- (g) Calculer un intervalle de confiance au niveau  $(1-\alpha)=95\%$  pour la valeur prédite du nombre d'accidents de la route lorsque  $X_1^0 = 3$ ,  $X_2^0 = 250$  et  $X_3^0 = 300$ .