EXERCICE 01 (07 pts):

I) 1)

$$\begin{aligned}
\hat{a} &= \bar{y} - \hat{b}\bar{x} \\
&= \bar{y} - \bar{x}\frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} \\
&= \bar{y} - \bar{x}\sum \frac{(x_i - \bar{x})}{\sum (x_i - \bar{x})^2} y_i, \text{car } \sum (x_i - \bar{x})\bar{y} = 0 \\
&= \frac{1}{n}\sum y_i - \sum \bar{x}\frac{(x_i - \bar{x})}{\sum (x_i - \bar{x})^2} y_i \\
&= \sum \left(\frac{1}{n} - \bar{x}\frac{(x_i - \bar{x})}{\sum (x_i - \bar{x})^2}\right) y_i \qquad (1 \text{ pt})
\end{aligned}$$

2)

$$\begin{aligned}
\sum a_i &= \sum \left(\frac{1}{n} - \bar{x}\frac{(x_i - \bar{x})}{\sum (x_i - \bar{x})^2}\right) = \sum \frac{1}{n} - \bar{x}\frac{\sum (x_i - \bar{x})}{\sum (x_i - \bar{x})^2}, \text{ de plus } \sum (x_i - \bar{x}) = 0 \\
&= 1 \qquad (1 \text{ pt})
\end{aligned}$$

$$\begin{aligned}
\sum a_i x_i &= \sum \left(\frac{1}{n} - \bar{x}\frac{(x_i - \bar{x})}{\sum (x_i - \bar{x})^2}\right) x_i \\
&= \frac{1}{n}\sum x_i - \bar{x}\frac{\sum (x_i - \bar{x}) x_i}{\sum (x_i - \bar{x})^2} \\
&= \bar{x} - \bar{x}\frac{\sum (x_i - \bar{x})^2}{\sum (x_i - \bar{x})^2}, \text{ car } \sum (x_i - \bar{x})\bar{x} = 0 \\
&= 0 \qquad (1\text{pt})
\end{aligned}$$

3) On va vérifier que $\hat{a}$ est un estimateur sans biais

$$\begin{aligned}
\hat{a} &= \sum a_i y_i = \sum a_i (a + bx_i + \varepsilon_i) \\
&= a\sum a_i + b\sum a_i x_i + \sum a_i \varepsilon_i \\
&= a + 0 + \sum a_i \varepsilon_i
\end{aligned}$$

Alors

$$E(\hat{a}) = a + \sum a_i E(\varepsilon_i) = a \qquad (1 \text{ pt})$$

II) On ajoute et on retranche $\hat{y}_i$

$$\begin{aligned}
\sum (y_i - \bar{y})^2 &= \sum (y_i - \hat{y}_i + \hat{y}_i - \bar{y})^2 \\
&= \sum (y_i - \hat{y}_i)^2 + \sum (\hat{y}_i - \bar{y})^2 + 2\sum (y_i - \hat{y}_i)(\hat{y}_i - \bar{y}) \qquad (0.5 \text{ pt})
\end{aligned}$$

Pour prouver que $SCT = SCE + SCR$, il suffit alors de montrer que $\sum (y_i - \hat{y}_i)(\hat{y}_i - \bar{y}) = 0$. Il est évident que

$$y_i - \hat{y}_i = y_i - \hat{a} - \hat{b}x_i = y_i - \left(\bar{y} - \hat{b}\bar{x}\right) - \hat{b}x_i = (y_i - \bar{y}) - \hat{b}(x_i - \bar{x}) \qquad (1 \text{ pt})$$

et

$$\hat{y}_i - \bar{y} = \left(\hat{a} + \hat{b}x_i\right) - \left(\hat{a} + \hat{b}\bar{x}\right) = \hat{b}(x_i - \bar{x}) \qquad (0.5 \text{ pt})$$

Donc

$$\begin{aligned} \sum (y_i - \hat{y}_i)(\hat{y}_i - \bar{y}) &= \hat{b}\sum (x_i - \bar{x})(y_i - \bar{y}) - \hat{b}^2 \sum (x_i - \bar{x})^2 \\ &= \frac{S_{xy}}{S_x} S_{xy} - \frac{S_{xy}^2}{S_x^2} S_x \\ &= 0 \qquad (1 \text{ pt}) \end{aligned}$$

On déduit alors que

$$SCT = SCE + SCR$$

EXERCICE 02 (10 pts):

I) On montre que la droite de régression passe par le point $(\bar{x}, \bar{y})$

$$\begin{aligned} \hat{y}_i &= \hat{\beta}_0 + \hat{\beta}_1\bar{x} = \bar{y} - \hat{\beta}_1\bar{x} + \hat{\beta}_1\bar{x} \\ &= \bar{y} \qquad (1 \text{ pt}) \end{aligned}$$

II) 1) Equation de la droite de régression : On sait que la droite de régression passe par $(\bar{x}, \bar{y}) = (50, 187.889)$. Connaissant le couple $(x_i, y_i) = (20, 347.296)$, on peut calculer la pente de la droite de régression

$$\begin{aligned} \hat{\beta}_1 &= \frac{347.296 - 187.889}{20 - 50} = -5.314 \\ &\Rightarrow \hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x} = 453.589 \qquad (2 \text{ pts}) \end{aligned}$$

Donc, la droite de régression est $\hat{y}_i = 453.589 - 5.314x_i$.

2) Estimation de la variance des paramètres :

$$\begin{aligned} \hat{\sigma}^2_{\hat{\beta}_1} &= \frac{\hat{\sigma}^2_\varepsilon}{S_x} = \frac{SCR/(n-2)}{S_x} = \frac{2747.825/7}{6000} = 0.065 \qquad (1\text{pt}) \\ \hat{\sigma}^2_{\hat{\beta}_0} &= \frac{\hat{\sigma}^2_\varepsilon \sum x_i^2}{nS_x} = \frac{\hat{\sigma}^2_\varepsilon\left(S_x + n\bar{x}^2\right)}{nS_x} = \frac{SCR}{n-2}\left(\frac{1}{n} + \frac{\bar{x}^2}{S_x}\right) = 207.18 \qquad (1 \text{ pt}) \end{aligned}$$

3) Test de la signification de $\hat{\beta}_0$

$$T = \frac{\left|\hat{\beta}_0\right|}{\hat{\sigma}_{\hat{\beta}_0}} = \frac{453.589}{14.393} = 31.515 \Rightarrow T > t_{1-\frac{\alpha}{2}}(n-2) = t_{0.975}(7) = 2.3646 \qquad (2 \text{ pt})$$

On rejete $(H_0 : \beta_0 = 0)$ alors le paramètre β_0 est significatif.

4) Coefficient de détermination D : On a

$$D = \frac{SCE}{SCT} = 1 - \frac{SCR}{SCT} = 1 - \frac{SCR}{S_y} = 1 - \frac{2747.825}{172136.9} = 0.984. \qquad (1 \text{ pt})$$

Ce qui implique une forte corrélation entre la température et la duré de vie des appareils et que le modèle est de très bonne qualité (0.5 pt)

5) Test de la signification globale du modèle à 95%.

$$F = T^2_{\hat{\beta}_1} = \frac{\hat{\beta}_1^2}{\hat{\sigma}^2_{\hat{\beta}_1}} = \frac{(-5.314)^2}{0.065} = 434.44 > f_{0.95}(1,7) = 5.59 \qquad (1.5)$$

On rejete $(H_0 : \beta_0 = 0$ et $\beta_1 = 0)$. Donc, le modèle est significatif globalement.