Etapa 2:

$$\begin{cases} \hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_i \\ y = \hat{\beta}_0 + \hat{\beta}_1 x_i + \varepsilon_i \end{cases}$$

a) $$SEQ = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$
$$= \sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2$$

$$\frac{dSEQ}{d\hat{\beta}_0} = 0$$

$$\frac{dSEQ}{d\hat{\beta}_0} = (-2) \sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \beta_1 x_i) = (-2) \left\{ \sum_{i=1}^{n} y_i - n\hat{\beta}_0 - \hat{\beta}_1 \sum_{i=1}^{n} x_i \right\} \begin{matrix} \div n \\ \div n \end{matrix}$$

$$= (-2)\, n \left\{ \frac{\sum_{i=1}^{n} y_i}{n} - \frac{\cancel{n}\hat{\beta}_0}{\cancel{n}} - \hat{\beta}_1 \frac{\sum_{i=1}^{n} x_i}{n} \right\} = (-2)\, n \left\{ \bar{y} - \hat{\beta}_0 - \hat{\beta}_1 \bar{x} \right\} \Rightarrow$$

$$\Rightarrow (-2)\, n \left\{ \bar{y} - \hat{\beta}_0 - \hat{\beta}_1 \bar{x} \right\} = 0 \Rightarrow \bar{y} - \hat{\beta}_0 - \beta_1 \bar{x} = 0$$

$$\boxed{\hat{\beta}_0 = -\hat{\beta}_1 \bar{x} + \bar{y}}$$

$$\frac{dSEQ}{d\hat{\beta}_1} = 0 = -2 \sum_{i=1}^{n} x_i (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = -2 \left\{ \sum_{i=1}^{n} x_i y_i - \hat{\beta}_0 \sum_{i=1}^{n} x_i - \beta_1 \sum_{i=1}^{n} x_i^2 \right\}$$

$$\sum x_i y_i - \hat{\beta}_0 \sum x_i - \hat{\beta}_1 \sum x_i^2 = 0$$

$$\sum x_i y_i + (\hat{\beta}_1 \bar{x} - \bar{y}) \sum x_i - \hat{\beta}_1 \sum x_i^2 = 0$$

$$\hat{\beta}_1 \left(\bar{x} \sum x_i - \sum x_i^2\right) = -\sum x_i y_i + \bar{y} \sum x_i$$

$$\hat{\beta}_1 = \frac{\bar{y} \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} x_i y_i}{\bar{x} \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} x_i^2} = \frac{-\sum x_i y_i + n \bar{y} \bar{x}}{-\sum x_i^2 + n \bar{x}^2}$$

$$\boxed{\hat{\beta}_1 = \frac{\sum x_i y_i - n \bar{y} \bar{x}}{\sum x_i^2 - n \bar{x}^2}} = \frac{Cov(X,Y)}{Var(X)}$$

b) A distribuição $\longrightarrow$ esperada para o erro é a normal pois seria o que melhor descreveria o comportamento dos erros.

O valor esperado é zero, já que em um estudo de dados não se espera ter erros.

A variância é igual e constante $(Var(\varepsilon_i) = \sigma^2)$

Ou seja: $\varepsilon_i \sim N(0, \sigma^2)$

Na prática, a suposição pode ser feita através da construção de gráficos.

c) O teste de hipóteses fica da seguinte forma:

$$\hat{\beta}_1 \sim N \begin{cases} H_0: \beta_1 = 0 \\ H_1: \beta_1 \neq 0 \end{cases}$$

A não rejeição de Ho faz com que $y = \beta_0 + \varepsilon_i$ e isso mostra que não existe relação entre as duas variáveis. Como $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$, caso Ho não seja refutado temos que $\beta_0 = \bar{y}$.

d) Sim. Na nova equação tem todas as variáveis inclusas e cada uma tem seu β. Para descobrir eles, deve-se derivar em função de cada β. Portanto a equação fica $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n + \varepsilon$, onde n é a quantidade de variáveis. As suposições do modelo continuam iguais, porém caso a correlação entre as variáveis sejam muito fortes, os parâmetros podem ser prejudicados. Por fim, o teste de hipóteses aumenta de acordo com a quantidade de variáveis analisadas. caso sejam duas variáveis, serão dois testes; três variáveis, três testes.