Problem 1. Consider the simple linear regression $y_i = a + b\,x_i + \varepsilon_i$. Show that

A. The point $(\bar{x}, \bar{y})$ is on the OLS regression line.

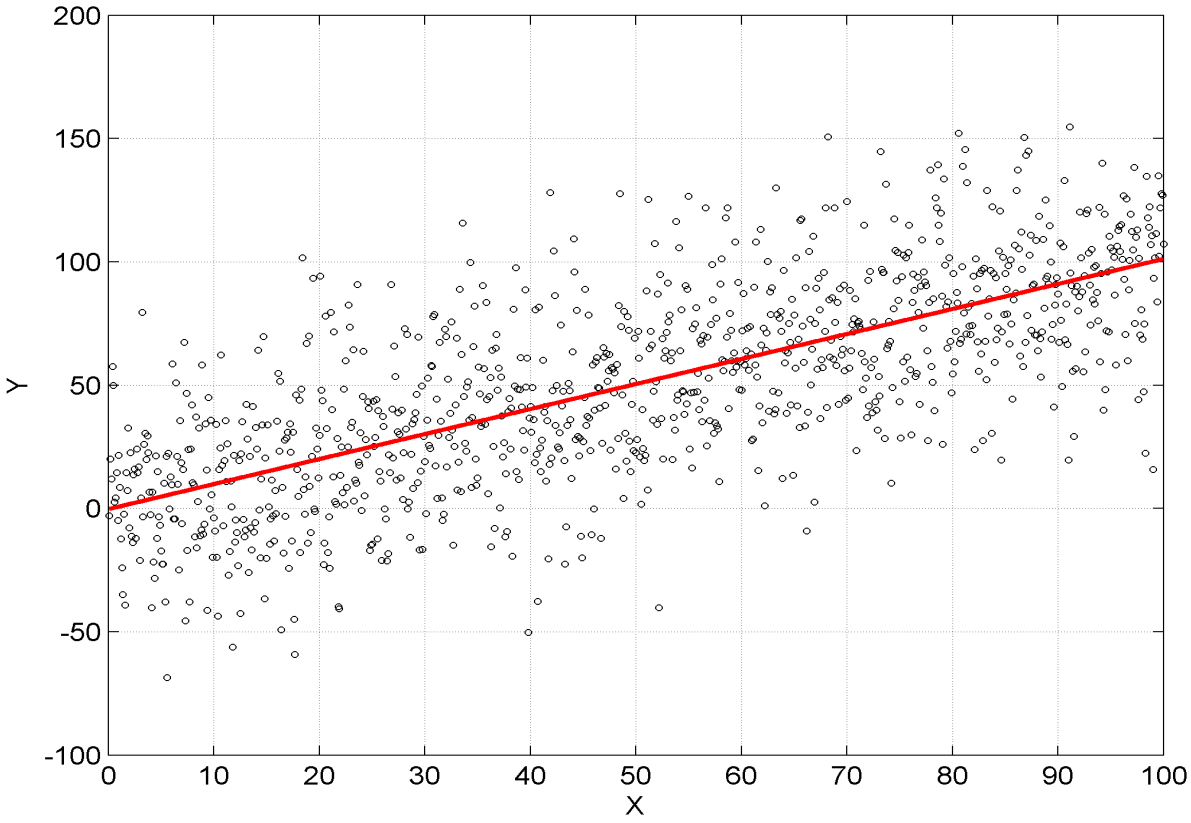
B. The OLS residuals add up to zero, i.e., $\sum_{i=1}^{n} \hat{\varepsilon}_i = 0$.

C. $\bar{y} = \bar{\hat{y}}$, i.e., the sample average of the actual y_i is the same as the sample average of the fitted values.

D. $\sum_{i=1}^{n} \hat{y}_i \hat{\varepsilon}_i = 0$.

Problem 2. Does the property B in Question 1 still hold if the linear specification is without the intercept, i.e., $y_i = b\,x_i + \varepsilon_i$? Explain your answer.

Problem 3. Consider an OLS regression of Y on X using 1,000 observations. The straight line through the plot below is $\widehat{Y} = \hat{a} + \hat{b}X$, and the standard error of the regression, typically denoted by $\hat{\sigma}_e$, is 29.



Now, another dot is going to be added to this chart, in line with the distribution of the plot. Choose the X value of the dot in such a way that a Y value of greater than zero is obtained. More precisely, at what value of X are you going to have a 95% chance of getting a dot such that it is in the positive territory of the Y axis? Note that all the information required to answer this question is already given in the chart (plus the fact that $\hat{\sigma}_e = 29$). Provide the arguments and workings by which you arrive at your answer.

Requirements:

(A) All answers are to be presented strictly in A4-size paper.

(B) You are free to either typeset in latex or write. If you write your answers, write nicely and legibly. The probability of inadvertently grading a sloppily written submission wrongly is quite high.

(C) Deadline: Upload your submission to elearn's dropbox by 2359 hours, October 31, 2018.