Chapter II: Bivariate Descriptive Statistics

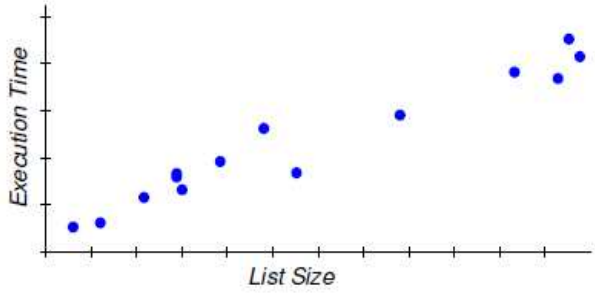
PROBLEMS

Proposed Problems

1. We have a computer program that executes a set of tasks with a list containing n objects. We execute the program 15 times using each time a different number of objects and we record its execution times T (seconds). As the CPU has to do other operations during the execution of the program, if we execute two times the software with the same values for the number of objects n we would get every time different values of its execution times. The following table shows the numbers of objects, n, in the list and the associated executions times T. The picture shows the scatter plot of these data.

n	65	118	116	66	114	30	60	106	25	38	45	44	85	52	44
T	208	327	352	162	296	91	225	305	85	127	138	161	244	178	156

$\bar{n}$	67.2
$\bar{T}$	203.67
corr(n, T)	0.972
s_n^2	990.03
s_T^2	6770.76



   a. How much time should we wait to execute the program with a list made of $n = 90$ elements?
   b. If we want that the execution time were less than $T = 100$ seconds, what should the maximum size of the list we could use?

   SOLUCTION:
   a. $\hat{T}(90) = 261.58$ seconds
   b. $\hat{n}(100) \approx 36$ objects

2. Show that the coefficient b of the slope of the simple regression line $\hat{y} = a + b\,x$ satisfies the following relation.

$$b = corr(x, y)\frac{s_y}{s_x}.$$

3. Let $(x_1, y_1), \cdots, (x_n, y_n)$ be a set of n pairs of points of positive variables (i.e., $x_i > 0, y_i > 0$), and which verifies that $corr(x, y) < 0$. Which of the following regression lines can match these data?
   a. $\hat{y} = 10 + 5\,x$
   b. $\hat{y} = -10 - 5\,x$
   c. $\hat{y} = 10 - 5\,x$
   d. $\hat{y} = -10 + 5\,x$

   SOLUCTION: Only the solution c. is possible.

4. Let $y = 10 + 2\,x$ be the least-squares simple regression model obtained from 20 pairs of data. Say whether the following statements are true or false.
    a. The correlation is 2.
    b. The expected value of y when x = 5 is 20.
    c. Both the covariance and the correlation are positive.
    d. The least-squares regression model that relates x with y is $x = -5 + 0.5\,y$.

    SOLUCTION:
    c. False
    d. True
    e. True
    f. False.