

INGENIERIA INFORMATICA
Escuela Politécnica Superior
Universidad Autónoma De Madrid

Statistics and Probability

Práctica 1

David Teófilo Garitagoitia Romero

2/21/2021

Tabla de contenido

Ejercicio 12

 A2

 B2

 C2

 D3

 E3

 F3

 G3

 H3

Ejercicio 24

 A4

 C5

 D5

Ejercicio 1

As part of their course assessment, a number of Computer Engineering students were asked to complete six simple programming tasks in two hours. The number of successfully completed programming tasks for these students were recorded and they are available in the file tasks_data.txt.

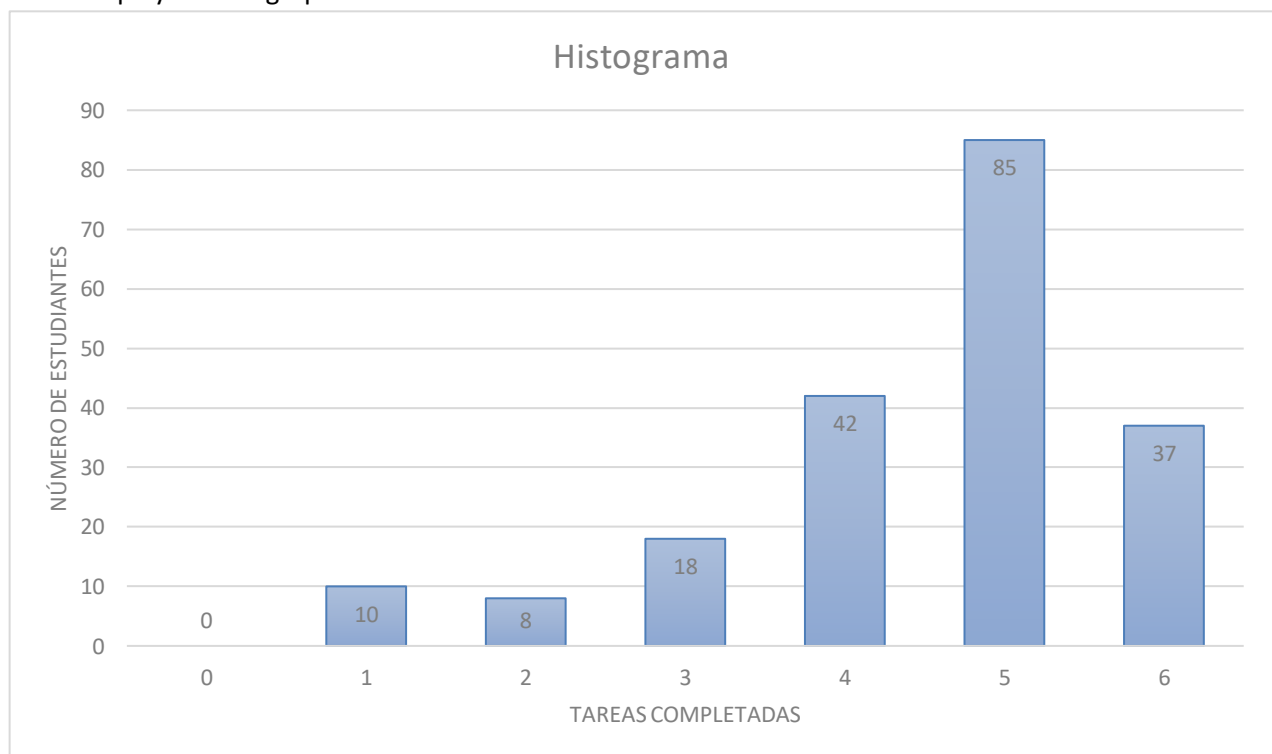
A How many students were asked to complete the tasks?

El tamaño de la muestra es de 200 estudiantes.

B Construct the frequency table of the dataset considered.

Valor	Frecuencia absoluta	Frecuencia Relativa
0	0	0
1	10	0,05
2	8	0,04
3	18	0,09
4	42	0,21
5	85	0,425
6	37	0,185

C Display the bar graph of this dataset.



D Describe the results that you obtained in (b) and (c).

En el b obtenemos la tabla de frecuencias absolutas (número de estudiantes en completar x número de tareas) y las frecuencias relativas (frecuencia absoluta/número total de estudiantes)

En el c, mostramos los datos obtenidos en el b mediante un gráfico.

E How many students were able to finish three tasks?

Con un vistazo rápido al gráfico podemos observar que el número de estudiantes que han completado exactamente 3 tareas son 18.

F How many students completed at least four tasks?

Para obtener el número de estudiantes que han completado al menos 4 tareas, basta con sumar las frecuencias absolutas desde los estudiantes que han superado 4 tareas, o lo que es lo mismo, restar al total aquellos que no han llegado a las 4 tareas completadas

$$N^{\circ} \text{ Est} = \text{Fabsoluta}(4) + \text{Fabsoluta}(5) + \text{Fabsoluta}(6) ;$$

$$N^{\circ} \text{ Est} = \text{Total} - (\text{Fabsoluta}(0) + \text{Fabsoluta}(1) + \text{Fabsoluta}(2) + \text{Fabsoluta}(3));$$

$$N^{\circ} \text{ Est} = 164.$$

G Find the proportion of students who only managed to finish one task.

Nuevamente basta con mirar el gráfico para observar que el número de estudiantes que únicamente lograron terminar una tarea es de 10 estudiantes.

H What is the percentage of students who completed the six tasks?

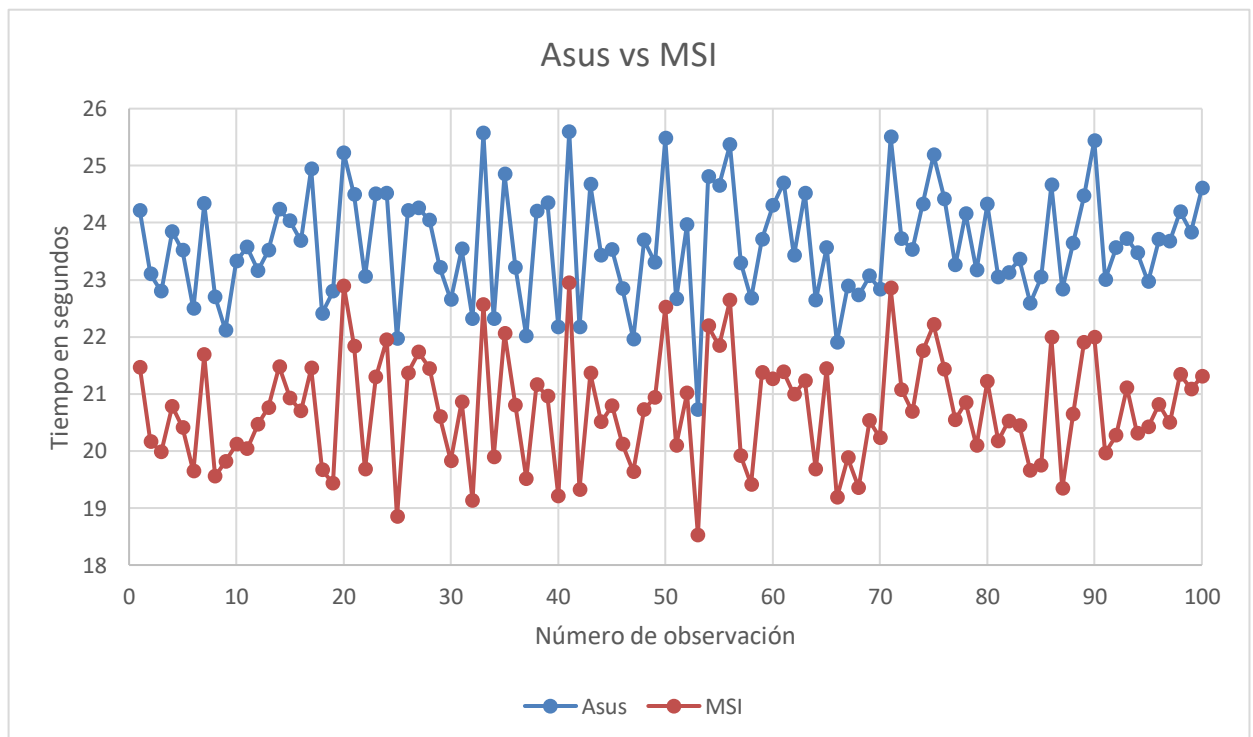
Volviendo a mirar el gráfico, podemos constatar que el número de estudiantes en completar las 6 tareas es de 37 estudiantes, lo que corresponde al 18.5%

Ejercicio 2

A group of computer scientists working in a certain research institute is interested in comparing the performance of two laptops with similar technical specifications, one manufactured by Asus and the other by Msi: both having Core i7 chips and 16Gb of RAM. To compare the performance of the two laptops, they install a program written to perform certain simulation-based calculations and record the times, in CPU seconds, needed to complete the calculations starting from 100 different initial values chosen at random. The times obtained for the 100 runs of the program are collected in the file `laptots_data.txt`.

A Draw a scatter diagram of these datasets and describe the relationship that you observe between them.

Observando el gráfico, se puede observar como la diferencia entre los valores que arroja el ordenador de Asus y el de MSI está siempre alrededor de los 2.87 segundos durante la totalidad de las pruebas, siendo el mínimo 2.188 segundos, y el máximo, 3.539 segundos. El ordenador de MSI es por ende aproximadamente un 12.2% más rápido que su rival



B Find the value of the sample (linear) correlation coefficient and interpret it.

Media X	Media Y	Covarianza x,y	Varianza x	Varianza y	Coefficiente de correlación
23,60959	20,73274	0,879178223	0,923953962	0,937722652	0,944527382

El coeficiente de correlación es muy próximo a 1, lo cual indica que existe una fuerte relación entre X e Y ya que el error cuadrático medio cometido es próximo a 0, además, indica que existe una asociación positiva entre las variables, es decir, al crecer X, crece a su vez Y

C Determine the regression line for predicting the time required with an Asus laptop from the time required with a Msi laptop.

$$A=4,171248; \quad B=0,937567437;$$

$$\text{Asus}=A+B\text{MSI} \quad \rightarrow \quad \text{ASUS}= 4,171248 + 0,937567437 \cdot X$$

D Find the value of the sample linear coefficient of determination

$$\text{sample linear coefficient of determination} = 0,944527382^2 = 8.92131975 \cdot 10^{-1}$$

E What can we say about the accuracy of the predictions made by using the regression line?

Como hemos comprobado por el ajuste de correlación lineal, el error cuadrático es próximo a 0 y por tanto la exactitud de la aproximación es elevada.

[FINAL DE DOCUMENTO]