Chapter II: Bivariate Descriptive Statistics -Exercises

Bernardo D'Auria

Statistics Department

Universidad Carlos III de Madrid

GROUP 89 - COMPUTER ENGINEERING

September 14th, 2010



Exercise

Exam I.I. Sep'06 - C4

In a Statistics group of a first-year Computer Engineering, we know that 60% of the students passed the Continuous Evaluation (C.E.) that consists of practical tests made during all the semester. To get full approval of that course they require to be evaluated only on the theoretical part by passing a Final Exam (F.E.). 70% of the ones that did not pass the C.E. did not pass F.E. either that in this case consists of both a theoretical and practical part. Finally we know that the 18% of all students passed both the C.E. and the F.E.

- a) What is the percentage of students that **did not pass** both the C.E. and the F.E.?
- b) What percentage of students did not pass the F.E.?
- c) Among the students who **do not passed** the F.E. what percentage **did pass** the C.E.?

Chapter II: Bivariate Descriptive Statistics - Exercises



SOLUTION

Bivariate Relative joint-Frequency Table

F.E.	Passed	Not Passes	53 A B
Passed	42%	12%	54%
Not Passed	18%	28%	46%
	60%	40%	

a) 28%

b) 46%

c) 39.13%



Exercise

Т									
208	We have a co	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							
327									
352		of objects and we record its execution times T (seconds). As the CPU has to							
162									
296									
91		numbers of objects, n , in the list and the associated executions times T . The picture shows the scatterplot of these data.							
225	picture shows								
305									
85	n	67.2	t						
127	$\frac{n}{\overline{T}}$	203.67	ime						
138	corr(n, T)	0.972			•				
161	s_p^2	990.03	ecrit						
244	s_T^2	6770.76	ů.	• •					
178									
156				Lis	t Size				
	208 327 352 162 296 91 225 305 85 127 138 161 244 178	$\begin{array}{c c} \hline 208 \\ 327 \\ 352 \\ 552 \\ 352 \\ 352 \\ 352 \\ 352 \\ 352 \\ 352 \\ 352 \\ 352 \\ 352 \\ 363 \\ 352 \\ 305 \\ 3$	208 327We have a computer prog n objects. We execute the of objects and we record do other operations during the software with the same every time different values numbers of objects, n, in t picture shows the scatterp305 \bar{n} picture shows the scatterp 203.67305 \bar{n} 203.6785 \bar{n} 203.67138 161 244 s_T^2 s_T 990.03 s_T^2 6770.76	208 327We have a computer program that exect n objects. We execute the program 15 t of objects and we record its execution do other operations during the execution the software with the same values for every time different values of its execut numbers of objects, n, in the list and th picture shows the scatterplot of these d 30585 \bar{n} \bar{T} 127 \bar{T} \bar{T} 208 206 203.67 138 $corr(n, T)$ s_T^2 161 s_T^2 s_T^2 6770.76	208 327We have a computer program that executes a set of n objects. We execute the program 15 times using a of objects and we record its execution times T (see do other operations during the execution of the program the software with the same values for the number every time different values of its execution times. T numbers of objects, n, in the list and the associate picture shows the scatterplot of these data.305 \bar{n} \bar{T} 305 \bar{n} \bar{T} 305 \bar{n} \bar{T} 305 \bar{n} \bar{T} 305 \bar{n} \bar{T} 306 \bar{n} \bar{T} 307 \bar{n} \bar{T} 308 244 s_T^2 309 $\bar{0}$ $\bar{0}$ 301 \bar{n} $\bar{1}$ 302 \bar{n} \bar{n} 303 \bar{n} \bar{n} 304 \bar{n} \bar{n} 305 \bar{n} \bar{n} 305 \bar{n} \bar{n} 306 \bar{n} \bar{n} 307 \bar{n} \bar{n} 308 \bar{n} \bar{n} 309 \bar{n} \bar{n} 301 \bar{n} \bar{n} 302 \bar{n} \bar{n} 303 \bar{n} \bar{n} 304 \bar{n} \bar{n} 305 \bar{n} \bar{n} 306 \bar{n} \bar{n} 307 \bar{n} \bar{n} 308 \bar{n} \bar{n} 309 \bar{n} \bar{n} 309 \bar{n} \bar{n} 309 \bar{n} \bar{n} 309 \bar{n} \bar{n} 309 \bar{n} \bar{n} 300 \bar{n} \bar{n} 301	208 327We have a computer program that executes a set of tasks with n objects. We execute the program 15 times using each time a of objects and we record its execution times T (seconds). As do other operations during the execution the program, if we e the software with the same values for the number of objects every time different values of its execution times. The following numbers of objects, n, in the list and the associated execution picture shows the scatterplot of these data.305 \bar{n} T T 138 \bar{n} r_T 444 s_T^2 67.2 $203.67990.036770.76g_{\mu}g_{\mu}$			

- 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 be?



SOLUTION

a) How much time should we wait to execute the program with a list made of n = 90 elements?
We can compute the parameters of the regression line

 $\hat{T}(n) = b_T \cdot n + a_T$: $b_T = 2.54$ and $a_T = 32.87$

and find that

 $\hat{T}(90) = 261.62 \text{ sec};$

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 be? We can compute the parameters of the regression line

 $\hat{n}(T) = b_n \cdot T + a_n$: $b_n = 0.3716$ and $a_n = -8.488$ and find that

 $\lfloor \hat{n}(100) \rfloor = \lfloor 28.67 \rfloor = 28$ elements.

Chapter II: Bivariate Descriptive Statistics - Exercises



CONDITIONAL ABSOLUTE AND RELATIVE FREQUENCIES



 f_A stays for Absolute Frequency. $f_A[\cdot, \cdot]$ is the sample size. f_R stays for Relative Frequency, e.g.

 $f_R[2,1] = f_A[2,1]/f_A[\cdot,\cdot].$

Example of Conditional Frequency

 $f[2|_{11}] = f[2|_{21}]$

absolute

relative

$$f_{R}[2|y_{2}] = \frac{f_{A}[2, 2]}{f_{A}[\cdot|y_{2}]} = \frac{f_{A}[2, 2]}{f_{A}[\cdot, 2]}$$
$$= \frac{f_{A}[2, 2]}{f_{A}[\cdot, 2]} \frac{f_{A}[\cdot, 2]}{f_{A}[\cdot, 2]} = \frac{f_{R}[2, 2]}{f_{R}[\cdot, 2]}$$

 $\begin{aligned} f_A[2|x_1] &= f_A[1,2] \\ f_R[2|x_1] &= \frac{f_A[2|x_1]}{f_A[\cdot|x_1]} = \frac{f_A[1,2]}{f_A[1,\cdot]} \\ &= \frac{f_A[1,2]}{f_A[\cdot,\cdot]} \frac{f_A[\cdot,\cdot]}{f_A[1,\cdot]} = \frac{f_R[1,2]}{f_R[1,\cdot]} \end{aligned}$