Departamento de Engenharia Informática
Faculdade de Ciências e Tecnologia da Universidade de Coimbra

Experimental Methods in Computer Science

Your company (a big multinational company) is installing a big datacenter to sell cloud services all over the world. The datacenter includes more than one hundred servers and has thousands of processors and tens of thousands of cores. The memory size needed is huge, including the main memory of the servers and memory to install state solid discs (SSD) in each node. In order to reduce the cost of memory, the company has decided to use the same memory board (a 4 GByte DRAM board) for both the main memory and for the SSDs, having ordered many thousands of memory modules.

Currently, the increased scale and the constant reduction of the size of the bit cells of dynamic memory (DRAM) chips are causing manufacturing failures that are very difficult to detect by the quality control. The most noticeable example is the data-dependent retention failures that consist of memory cells that may show errors due problems in the manufacturing process. The quality tests at the end of the manufacturing lines do not guarantee that the chips are 100% free of this problem, as the time needed to test each chip with all the data pattern combinations would be months (or more). The memory manufacturers estimate by sampling the maximum number of bit cells susceptible to data-dependent retention failures in each GBit. For the case of the chips your company has ordered, the manufacturer states an average of 23 bit cells susceptible to data-dependent retention failures in each Gbit, with standard deviation of 7.53.

Considering the huge amount of memory ordered, your company decided to test exhaustively a sample of 36 modules. After several months of test, the results were the following: An average of 26 bit cells susceptible to data-dependent retention failures in each Gbit.

Based on these results, do you think your company should accept the memory modules ordered or not, considering a confidence level of 95%? Your answer should indicate the hypothesis you are testing and should explain all steps to reach your conclusion.

2) Assume that you have obtained a new navigation system. The time to find a valid GPS signal follows a normal distribution with a mean of 10 seconds. In order to validate this information, you performed 9 time measurement in randomly chosen places. You have obtained a mean of 10.5 seconds and a standard deviation of 5 seconds. Would you trust on the results given by the GPS company with a significance level of 5%?