

Bachelor in Computer Science and Engineering

Statistics Problems

VII Large Sample Inference

1. The life time in hours of a battery designed for electric vehicles is normally distributed. The engineers who designed the battery say the average life of the battery is 4 hours of driving time. However, the análisis of a random sample of 9 batteries has resulted in a mean life time of 3,5 hours and a quasi-standard deviation of 0,806 hours of driving.
 - a. Is there enough evidence to reject what the engineers say? (use $\alpha=0,05$)
 - b. Which is the minimum significance level for which we could reject the null hypothesis?
 - c. Build a 95% confidence Interval for the battery mean life time
 - d. The conclusión obtained in a. could it be derived analyzing the confidence Interval obtained in c? Justify your answer.
2. 270 managers are randomly selected from a group of public administration companies, 56 of whom were women. From another group of non-public companies, 203 managers were selected, also randomly, of which 151 were men. Answer the following questions justifying the answers and using $\alpha=0,05$:
 - a. Calculate a confidence interval for the proportion of female managers in the public administration companies.
 - b. Check by means of a hypothesis test if the proportion of male managers in the non-public companies is smaller than 78%
 - c. Check by means of a hypothesis test if there are significant differences between the proportion of female managers in public administration companies and no-public companies. Find the p-value for the hypothesis test.
3. To determine the proportion of computer attacks on a server that can be neutralized with a certain software simple random sampling is performed and it is found that from 50 attacks received 35 were neutralized.
 - a. Calculate a 90% confidence Interval for the proportion of attacks that can be neutralized.
 - b. If it is wanted to estimate that proportion with a precisión of 0,05 (Interval width 0,10) with the same confidence level, which should be the size of the sample?

4. A company is studying the percentage of defective items in its production line.
 - a. A simple random sample of 1000 items is taken, and 100 defective items are found. Calculate a 99% confidence Interval for the proportion of defective items.
 - b. In order to lower productions costs, a series of changes in the production process have been implemented. Subsequently, an external audit has carried out its own sampling of 1500 items, and has concluded that the proportion of defective items is 12%. With a significance level of 5%, can the external audit conclude that the changes implemented has increased the proportion of defective items?

5. The gross salary S of a population in a certain country follows a Pareto distribution with mean 30.000. A simple random sample of 100 workers from the country's capital yielded a sample mean salary of 32.000 with a sample standard deviation of 10.500. The interest is to study whether the sample mean salary in the capital is different from that of the national mean salary or not.
 - a. Which are the null and alternative hypothesis that define the hypothesis test of this problem?
 - b. Which is the value for the test statistic?
 - c. If we want a significance level of 5%, which is the rejection area?
 - d. Based on what was observed, is the null hypothesis rejected? Then what can be concluded for this problem?
 - e. What is the test p-value?

6. In the records of a computer maintenance company there are 68 failures of electronic type and 59 of mechanical type.
 - a. Could it be affirmed that the proportion of an electronic failure is greater than that of a mechanical one?
 - b. Calculate the p-value for the previous hypothesis test
 - c. 41 invoices have been issued, with an average amount of 94.38€ and a standard deviation of 13.93€. Does this data provide enough evidence to affirm that the average billing per invoice is greater than 90€? Establish conclusion to 5% significance level.

7. We want to analyze the time that Spanish children spend watching TV during the week. A simple random sample of 50 children has a weekly average time dedicated to watch TV of 18,36 hours, with a sample quasi-standard deviation of 3,92 hours. ($Z_{0,025} = 1,96$, and $Z_{0,05} = 1,64$).
 - a. Calculate a 95% confidence Interval for the weekly average time that Spanish children dedicate to watch TV.
 - b. Which should be the sample size to reduce the width of the previous Interval to a fourth of its width? (Assume that the rest of sample values do not change).
 - c. For a significance level of 5%, perform a hypothesis test considering the alternative that the weekly average time dedicated to watch TV is less than 19,3 hours.
 - d. Could we obtain the result of the previous hypothesis test analyzing the confidence Interval calculated in a). Justify your answer.