EXERCISES – UD 5: INFERENCE part 1

DISTRIBUTIONS IN SAMPLING

1) One sample of size N is randomly extracted from a Normal population with a standard deviation $\sigma=3$. Calculate the minimum value of N in order to have a probability lower than 1% to obtain a difference higher than 1 (in absolute value) between the sample mean and the population mean.

2) One quality parameters follows a distribution N(5; 2). If one sample of size 4 is taken, what is the probability to obtain a sample mean higher than 6? What is the probability if the sample size is 25?

3) Two random samples are extracted from one Normal population with variance 9, with sample sizes $n_1=22$ and $n_2=30$. If the difference of both sample means is computed, what is the probability to be greater than 1?

4) One quality parameter measured in a certain manufacturing process has a distribution N(m; 3). Determine the sample size to take for the estimation of the population mean of this parameter, so that the sample mean differs from the population mean less than 1 cm with a probability of 95%.

5) One company produces electronic circuits. The time of operation (in hours) follows a distribution N(2,000; 200). After certain changes in the manufacturing process, the time has a distribution N(2,200; 200). A sample of 10 circuits was taken before the change, and 30 circuits are taken after the change. If the difference of both sample means is computed, what is the probability to be between 195 and 205 ?

6) Calculate a and b so that $P(a < s^2 < b) = 0.8$, being s^2 the sample variance of a random sample with size 16 taken from a population with distribution N(8; 2).

7) The horizontal (X) and vertical deviations (Y) (both measured in cm) of certain shots with respect to the center of a target fluctuate independently according to a distribution N(0; 2). What percentage of the shots will be located on the target at a distance higher than 4 cm with respect to the center?

8) One hospital wants to estimate the average time of waiting at the urgency services by interviewing a random sample of patients attending these services. In a previous study, the standard deviation of this time was estimated as 12 minutes. Admitting a maximum error of ± 5 minutes and a confidence level of 95%, how many patients should be interviewed?

9) If 16 elements are randomly extracted from a population N(0; 5), what is the probability to obtain a sample mean greater than 3, in absolute value?

10) The transference time (milliseconds) of a certain type of data files through a network follows a distribution N(m, $\sigma=4$). If a simple random sample of size 25 is extracted, what is the probability to obtain a difference between the population mean and the sample mean higher than 2, in absolute value?

11) If the three components of a vector are independent values distributed normally with mean zero and variance 4, what is the mean value of the squared length of this vector?

12) In a chi-square distribution with 20 degrees of freedom, calculate $P(\chi^2 > 34.17)$, assuming that a χ^2_{20} can be approximated by means of a Normal distribution.

13) If a sample of size 10 is extracted from a normal population with variance 9, what is the probability to obtain a sample variance higher than 18?

14) One simple random sample of size 15 is extracted from a population N(m=8, σ=4). Obtain one interval with a probability of 99% to contain the sample variance.

15) In Student's t distribution with 12 degrees of freedom, what is the interval comprising 80% of the values?

16) The time of response in a computer system follows a normal distribution with mean m=5. Twenty trials were carried out to measure this time. If the sample variance is s^2=2.5, obtain an interval containing the value of the sample mean with a probability of 95%.

17) The voltage of certain electronic circuit has a normal distribution with average m=120. One sample of 10 circuits was extracted, resulting a sample standard deviation s=1.5. Calculate the probability to get a sample mean greater than 121.

18) In a group of 1500 people, how many subjects will have a score between 0.28 and 8.1 if the score is a random variable following an $F_{8,6}$ distribution?

19) The temperature of an electronic circuit after 1 hour of operation follows a normal distribution. If two samples are extracted from the population of circuits, with sizes N_1=10 and N_2=18, what is the probability to get a sample variance from the first sample more than 2.5 times the value of the second sample variance?

20) The time of data transfer (ms) through a network has a distribution N(m; 3). If one sample of size 16 is randomly taken, what is the probability to obtain a difference higher than 2, in absolute value, between the population mean and the sample mean?

21) What is the probability to obtain a sample variance greater than 10 when taking a sample of size 20 from a Normal population with a variance equal to 5 ?

22) One sample of size 9 is randomly extracted from a population with distribution N(30; 5). What is the probability to obtain a sample mean higher than 25? What is the probability to obtain a sample standard deviation below 3?

23) In order to estimate the mean of a Normal population with variance 4, one sample of size N is extracted from the population and the sample mean is obtained. What is the minimum sample size required to have an error in the estimation above 0,5 units with a probability lower than 1%?

24) One random variable X is defined on a population Normally distributed with standard deviation 10. What is the minimum sample size required to guarantee with a probability of 95% that the sample mean will differ less than one unit from the population mean?

EXERCISES – UD 5: INFERENCE part 2:

INFERENCE ABOUT ONE NORMAL POPULATION

25) In the process of car painting, the thickness of the paint layer follows a normal distribution with average 105 μm and standard deviation 5 μm. The quality control of this process is conducted by obtaining the average of 4 measurements from 4 cars randomly selected. The process is considered as correct if the mean obtained is > 100 μm. What is the probability to reject the process?

26) The number of defects in one piece follows a Poisson distribution, Ps (λ). From a sample of size 10, we want to test the null hypothesis that λ of the Poisson distribution is 2 versus the alternative that is greater than 2. We will accept H_0 if the sample mean is < 2.5 and will reject H_0 otherwise.
a) What is the type I risk of this test?
b) What is the type II risk if λ is actually 3?
b) What is the type II risk if λ is actually 4?

27) From one random sample of size 21 from a normal population, the following hypothesis test is proposed: $H_0 : \sigma^2=4$ versus $H_1 : \sigma^2=9$. The null hypothesis is accepted if the sample variance s^2 is lower than 6, and it is rejected if s^2 results greater than 7. If s^2 is comprised between 6 and 7, the H_0 is accepted with a probability of 0.6 and rejected with a probability of 0.4. Obtain the probability of type I error (α) and type II error (β).

28) Indicate which of the following sentences is true.
a) One confidence interval of a certain parameter obtained at a confidence level of 90% contains with a probability of 90% the true unknown value of the population parameter to estimate.
b) One confidence interval of a certain parameter obtained at a confidence level of 90% is the acceptance region of a hypothesis test considering α=0.1.
c) None one of the above.

29) Certain characteristic X follows a Poisson distribution Ps(λ). In order to test the null hypothesis λ=1 versus the alternative λ=2, one observation is taken (i.e. sample size=1) and H_0 is rejected if x>3. Obtain the probability of type I error (α) and type II error (β).

30) One factory of computer circuits buys every month a large set of certain type of piece. In order to test the quality of the pieces prior to the purchase, a sample of 9 pieces is taken, and the whole set is accepted if all pieces in the sample are correct. Obtain the probability of type II error in the case of a set containing 10% of defective pieces.

31) One company manufactures 40 W light bulbs. The filament of these bulbs has an electric resistance following a distribution N(605, 1.18) ohms. The bulb is considered as correct when its resistance is comprised between 607.86 and 600.96 ohms. In order to test if

the population average is 605, which is the desired value, samples of 7 bulbs are periodically taken. The process is considered as out of control if the sample average is higher than 606.338 or lower than 603.662.

a) Obtain the probability of type I error.
b) Obtain the probability of type II error if the population average is 606.5.

32) The number of defects per square meter found in a certain type of marble follows a Poisson distribution with parameter λ = 0.3 defects/m^2. The quality of the marble is considered adequate (i.e. λ = 0.3 defects/m^2) if less than 3 defects are found in 4 m^2 randomly chosen. Otherwise, the number of defects is considered inappropriate.

a) Obtain the value α of this test.
b) If λ= 0.4 defects/m^2, what would be the value β?
c) How many square meters should be examined in order to obtain α=0.05?

33) The management of the subway service in Paris has determined that the average value of the delay with respect to the expected time of arrival is 15 seconds, with a standard deviation of 10 seconds. In summer, the management thinks that the time of delay might be different. In order to study this issue, the time of delay is measured in 20 cases, obtaining the following values (seconds): (10, 0, 3, -2, -4, 14, 20, 4, 30, 9, 3, 3, 6, 23, -10, 21, 2, 5, 23, -10). Assuming that data are normally distributed,

a) Obtain a confidence level at α=0.1 for the population mean time of delay, based only on the information from the sample.
b) Considering a significance level of 1%, is there enough evidence to affirm that the average time of delay in summer is not 15?
c) Calculate the probability of type II error of this test if the population mean delay time in summer would have been 10 seconds.

34) In order to study the thickness of a painting layer for an automobile application, 10 samples from different vehicles were taken, resulting the following values measured in microns: { 45, 40, 44, 43, 45, 43, 46, 45, 44, 45 } Assuming that data are normally distributed,

a) Obtain a confidence interval for the average thickness considering α=0.05.
a) Obtain a confidence interval for the variance of thickness considering α=0.05.

35) For certain electronic components, the time of operation until failure is the most important quality parameter. One sample of 16 components of this type has presented an average time of 734 hours. It is assumed that data are normally distributed.

a) Can we accept that the population average is 740 hours with a significance level 0.05 considering σ=12 hours?
b) Can we accept that the population average is 740 hours with a significance level 0.05 considering s=12 hours?

36) One electric resistance has been measured 6 times, resulting the following values (in ohms): { 1.5, 1.6, 1.4, 1.5, 1.3, 1.1 } Obtain a confidence interval for the population mean considering α=0.05.

37) After many research studies, the number of heart beats per minute in the male population between 20 and 25 years follows a distribution N(72, 9). If a sample of 100 professional football players gives a sample average of 64 beats per minute, can we consider this difference as statistically significant at α=0.05?

38) One low-cost flight company offers the flight Madrid – Valencia. A random sample of 30 flights is selected, resulting an average price of 24 € and a variance of 25 $€^2$. This company is currently advertising that the average price of this flight is 20 €. Assuming that prices are normally distributed and based on the information from the sample, is this hypothesis acceptable at α=0.05?

39) One company has developed a new OCR (optic character recognition) program. In order to test the performance of this program, 10 sets of 1,000 characters are randomly selected and processed by the OCR program, resulting the following percentage of correct recognition: 99%, 97%, 98%, 98%, 96%, 99%, 97%, 95%, 94%, 92%.

a) Can the company guarantee that the average percentage of correct recognition is 95%, considering a significance level of 5%?
b) Obtain a confidence interval for the average percentage with a confidence level of 98%.

40) One manufacturing process produces USB cables with a length of 3 m. In order to study if the process works correctly, with a population average of 3 m, a sample of 20 cables is taken, resulting a sample average of 2.96 m and a sample standard deviation 0.1 m. Based on this information, should the process be readjusted considering α=0.05?

41) One computer program translates text from English into Spanish. In order to study the quality of this program, 14 paragraphs with 100 words in English are randomly selected and translated with the program. Afterwards, the number of errors is counted by an expert. The number of errors found in the sample of 14 paragraphs is the following: 10; 12; 7; 8; 6; 10; 9; 9; 10; 8; 9; 11; 10; 9.

a) Obtain a confidence interval for the mean number of errors per 100 words (α=0.05).
b) Can we affirm with a 95% of confidence that this mean is 8?

42) One computer company sells a certain laptop model stating that their average weight is 1000 gr. A sample of 5 laptops is examined, and the following values of weight are obtained: 995, 992, 1005, 998, 1000

a) Are the values of the sample consistent with the hypothesis that m=1000 with a significance level α= 10%?
b) Obtain a confidence interval at 95% for m.

SOLUTIONS

1) 49
2) 0.1587 ; 0.0062
3) 0.234
4) 36
5) 0.056
6) $a = 2.28$, $b = 5.95$
7) 13.53%
8) 23
9) 0.0164
10) 0.0124
11) 12
12) 0.0125
13) 0.035
14) [4.657, 35.793]
15) [-1.356, 1.356]
16) [4.26, 5.74]
17) 0.032
18) 1410
19) 0.05

25) 0.0228
26) a: 0.11 b: 0.21 c: 0.01
27) $\alpha = 0.042$ $\beta = 0.23$
28) a)
29) $\alpha = 0.019$ $\beta = 0.8571$
30) 0.3874
31) $\alpha=0.0027$, $\beta=0.3594$
32) a: 0.121 b: 0.783 c: 2.677
33) a: [3.17, 11.83] b: yes c: 0.633
34) a: [42.78, 45.22] b: [1.37, 9.63]
35) a: no b: yes
36) [1.21, 1.59]
37) yes
38) no
39) a: yes b: [94.874, 98,126]
40) no, $m \in$ [2.91, 3.01]
41) a: [8.24, 10.04] b: no