etsinf

Bachelor Degree in Computer Engineering

Statistics

SECOND PARTIAL EXAM

April 11th 2011

Surname, name	
Signature	

Instructions

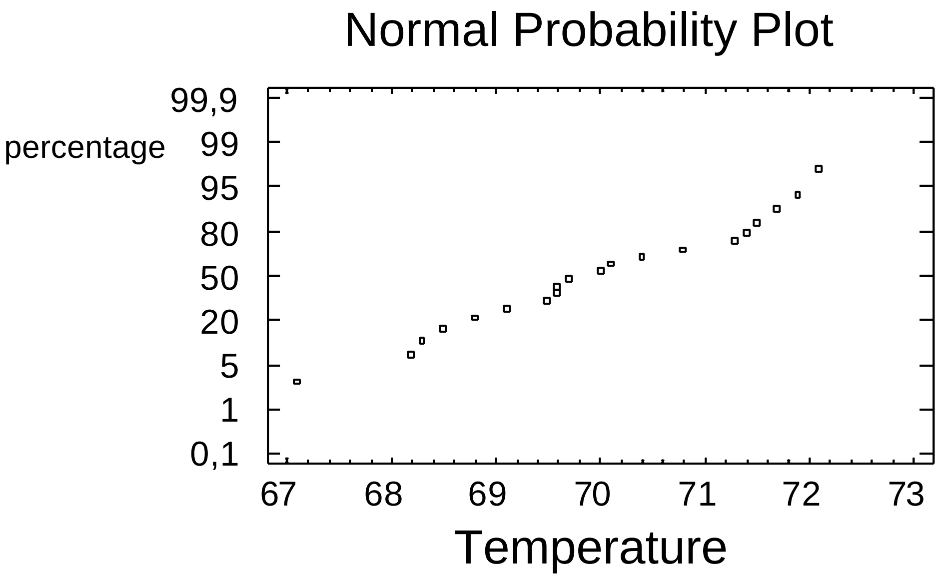
1. Write your name and sign in this page.
2. Answer each question in the corresponding page.
3. All answers must be justified.
4. Personal notes in the formula tables will note be allowed. Over the table it is only permitted to have the DNI (identification document), calculator, pen, and the formula tables.
5. Do not unstable any page of the exam (do not remove the staple).
6. All questions score the same (over 10).
7. At the end, it is compulsory to sign in the list on the professor’s table in order to justify that the exam has been handed in.
8. Time available: 2 hours

1. It has been checked that the life time of certain computer virus in a system follows an exponential distribution with median 80 hours.

a) What is the probability of that virus to remain in the system between 70 and 80 hours? (*5 points*)

b) If we know that the virus is remaining in the system 30 hours, what is the probability to exceed in total 90 hours? (*5 points*)

2. In one study about the maximum operative temperature that bears the CPU of certain type of computer equipment, a sample of 20 values of that temperature (in ºC) has been obtained. The sample values are represented below in a normal probability plot.



a) Does the maximum operative temperature follow approximately a normal distribution? (*2 points*)

b) Estimate approximately the mean and the standard deviation of the maximum operative temperature. (*4 points*)

c) What percentage of equipments will bear a maximum operative temperature higher than 68ºC? Carry out the calculation using the normal distribution table. (*4 points*)

3. One computing center offers computing services based on Cloud Computing techniques. For a certain process that inverts large data matrices in this computing center, the response time is distributed uniformly with time fluctuating between 15 and 25 seconds. One process of a weather forecasting company needs to access 30 times consecutively to this service of large matrix inversion. It is assumed that the data transfer time through the web can be regarded as negligible. What is the probability of the process to invert the 30 matrices before 9 minutes and 30 seconds? (*10 points*)

4. Samples with size 25 are extracted from a normal population with standard deviation 10. What percentage of these samples will have a variance lower than 117? (*10 points*)

5. One microchip factory assembles transistors over circular silica chips, with a diameter of 300 mm, that are purchased to third parties.

Due to variability causes in the chip manufacturing process, it is impossible to obtain all chips with a diameter of exactly 300 mm.

In order to guarantee that the diameter of chips purchased by the factory to a new supplier is appropriate, one quality control is conducted by taking a sample of 100 chips and measuring their diameter. The average diameter of chips in the sample has resulted 299 mm and the standard deviation of the diameter 9 mm.

Answer the following questions with an appropriate justification:

a) Can we affirm, with a confidence level of 90%, that the average diameter of chips from that supplier is appropriate for the factory? (*5 points*)

b) The supplier affirms that the diameter of chips that he markets has a standard deviation of 8 mm. Does this statement seem reasonable from the sample analyzed (α=10%)? (*5 points*)

SOLUTION OF SECOND PARTIAL - STATISTICS

1) variable T: time of the virus in the system → $T \approx \exp(\alpha)$

$P(T>80)=0.5$; $e^{-\alpha \cdot 80}=0.5$; $-80\cdot\alpha = \ln 0.5$; $\alpha = -\ln 0.5 / 80 = 0.008664$

a) $P(70<T<80) = P(T<80)-P(T<70) = 0.5-(1-e^{-0.008664 \cdot 70}) = 0.5-0.454=$ **0.0452**

b) Taking into account the lack of memory property of the exponential:

$P(T>90/T>30) = P(T>60) = e^{-0.008664 \cdot 60} =$ **0.5946**

2) MOP: maximum operative temperature

a) Yes, because points tend to be aligned following approx. a straight line.

b) According to the plot, $P(MOP<70) \approx 0.5$ → median ≈ 70

As the distribution is approximately normal: mean = median ≈ **70**

In a normal distribution: $m \pm 2\sigma$ comprises 95% of the data. According to the plot:

$P(MOP<67) \approx 2.5\%$ → $m-2\sigma = 67$ → $\sigma = (70-67)/2 \approx$ **1.5**

c) $P(X>68)= P[N(70;\ 1.5) > 68] = P[N(0;1) > (68-70)/1.5]=P[N(0;1) > -1.33] =$ **0.908**

3) $X \approx U(15;\ 25)$ → $E(X)=20$; $\sigma^2(X)=(b-a)^2/12 = (25-15)^2/12 = 8.333$

T: time required to invert 30 matrices → $T = X_1+X_2+\ldots+X_{30}$

$E(T) = E(X_1)+\ldots+E(X_{30}) = 30\cdot E(X_i) = 30\cdot 20 = 600$

$\sigma^2(T) = \sigma^2(X_1)+\ldots+\sigma^2(X_{30}) = 30\cdot\sigma^2(X_i) = 30\cdot 8.333 = 250$

Central limit theorem: $T \approx N(600; \sqrt{250})$

$P(T<570)= P\left[N(600;\ \sqrt{250}) < 570\right] = P\left[N(0;\ 1) < \frac{570-600}{\sqrt{250}}\right]$ $=P[N(0;1)<-1.9]=$ **0.0287**

4) $P(s^2<117) = P\left(s^2\cdot\frac{n-1}{\sigma^2} < 117\cdot\frac{n-1}{\sigma^2}\right) = P\left(\chi^2_{n-1} < 117\cdot\frac{24}{10^2}\right) = P\left(\chi^2_{24} < 28.08\right) \approx$ **0.75**

5) H_0: $m=300$; H_1: $m \neq 300$; $\alpha=0.1$; $\bar{x}=299$; $s = 9$ We will accept H_0 if:

a) $\left|\frac{\bar{x}-m}{s/\sqrt{n}}\right| < t_{n-1}^{\alpha/2}$; $\left|\frac{299-300}{9/\sqrt{100}}\right| < t_{99}^{0.05}$; $1.111 < 1.66$ → Accept H_0: yes, we can affirm that.

b) H_0: $\sigma = 8$; H_1: $\sigma \neq 8$; $(n-1)\cdot s^2/\sigma^2 \approx \chi^2_{n-1}$; $99\cdot 9^2/8^2 = 125.3$

Critical values of the χ^2_{99} distribution (at $\alpha=0.1$), from the table: 77.93; 124.34

Given that 125.3 > 124.34 → reject H_0 → The statement does not seem reasonable.