



This assignment is due at my office, office 3.20 at the Department of Computing and Electrical Engineering, on Friday, March 7, before 4:30 pm. Complete the details below and staple this sheet to the front of your solutions. Please *underline* your family name.

NAME: \_\_\_\_\_ STUDENT NUMBER: \_\_\_\_\_

LECTURER: YVAN PETILLOT

1. CD's are examined for scratches. 150 CD's give rise to the following data:

Number of scratches per item	0	1	2	3	4	5	6
Frequency	11	45	38	35	14	5	2

Let  $x$  be the random variable 'number of scratches on the surface of a randomly chosen item', and suppose that  $x$  has a Poisson distribution with parameter  $\lambda$ .

- Find an estimator for  $\lambda$  and calculate the estimate for the above data.
  - Compute the estimated standard error and calculate the standard error of the data above.
2. Let  $x$  denote the proportion of allotted time that a randomly selected student works on a certain assignment. Suppose that the distribution of  $x$  has the following density function:

$$f(x) = \begin{cases} (\theta + 1)x^\theta & \text{if } 0 \leq x \leq 1 \\ 0 & \text{else} \end{cases}$$

with parameter  $\theta > -1$ . Find the maximum likelihood estimator of  $\theta$  and then compute an estimate for it with the following data:

0.91 0.89 0.67 0.45 0.98 0.81 0.70 0.78 0.95 0.86

3. The  $(1 - \alpha)100\%$  confidence interval for the mean of a normal population with known standard deviation  $\sigma$  is

$$\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}},$$

where  $\bar{x}$  is the sample mean.

- How much must the sample size  $n$  be increased if the length of the confidence interval is to be halved? Justify your answer.
  - If the sample size is increased by a factor of 25, what effect will this have on the length of the interval? Justify your answer.
4. A human resources officer in a large company wants to survey what proportion of the work force is aware of her position.
- What sample size is necessary if the 95% confidence interval for  $p$  is to have length at most 0.1, irrespective of  $p$ .
  - If the officer has strong reasons to believe that  $p = 2/3$ , what sample size should be recommended?

5. For the following sets of data calculate a 99% confidence interval for the difference between the true average.

sample ...	size	mean	standard dev.
set 1	68	26.99	4.89
set 2	74	35.76	6.43

6. Consider the data for EU Council of Ministers given on slides Unit 6-14 and Unit 6-15.
- (a) For the *combined list* of all 27 countries calculate the least squares regression line. For each country calculate the error between predicted number of seats and observed number of seats.  
(This question is not part of the assignment, but can you think of reasons why some data (like that of Germany) are so far off the least squares regression line?)
  - (b) Calculate the correlation factor for this set of data.
  - (c) Find a 95% confidence interval for the true slope of the least squares regression line. Give an interpretation of this confidence interval for this set of data.