

# ORDRE DES INGÉNIEURS DU QUÉBEC

MAY 2012

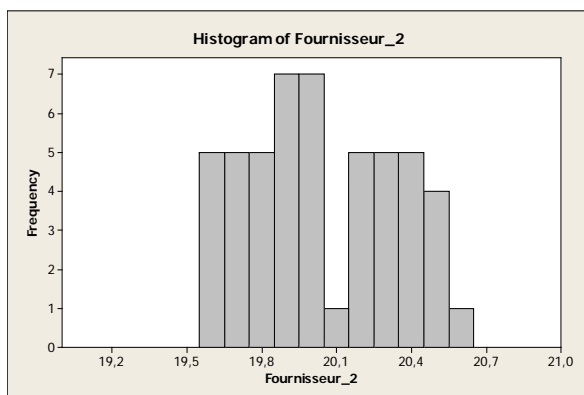
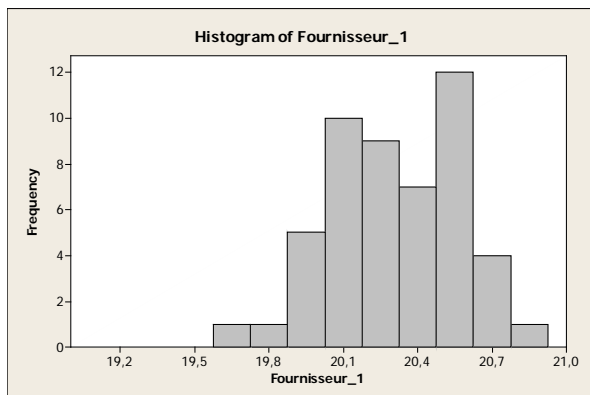
Toute documentation permise  
Calculatrices : modèles autorisés seulement  
Durée de l'examen : 3 heures

## 98-Ind-B1 Applied Probability and Statistics

### Question 1 (20 points)

You must estimate the quality of two different suppliers. You have considered a typical part and you asked each supplier to produce 50 samples. The requirements of the key characteristic are  $20 \pm 1$ . You calculated the following statistics for each supplier:

	Supplier 1	Supplier 2
Estimated Average (Expectancy)	20.307	20.049
Estimated standard deviation	0.274	0.285
Maximum Value	20.798	20.553
Minimum Value	19.623	19.603
Skewness	-0.26	0.16
Kurtosis	-0.60	-1.18



- (5 points) Are the **average** behaviors of the two suppliers identical? Justify your answer with a statistical test.
- (5 points) Can the **two variations** be considered equal? Justify your answer with a statistical test.
- (7 points) If the requirement is  $20 \pm 1$ , which supplier offers a better behavior? Justify your answer by estimating the level of defects to be produced by each machine.
- (3 points) Can we use kurtosis and skewness to estimate the quality of each machine? Why?

### Question 2 (20 points)

You are responsible for the estimation of hydraulic equipment **reliability**. The Marketing Department wants to offer a 3 year warranty (which corresponds to approximately **500 hours** of operation). We accept a return rate due to the guarantee of around **8%**. A test bench has done a cycling test of 12 equipments until they broke. The results of their lifetimes (in hours) are:

594	583	552	571	567	559
545	544	528	564	563	547

- (7 points) Estimate the average life (mean time) with 95% as confidence interval at.
- (7 points) Can we certify whether the equipment threshold of 500 hours at 92% (100%-8%)? Justifies your answer with a statistical test + a graph.
- (6 points) For accuracy in predicting the life of  $\pm 50$  hours, what is the minimum sample size you should consider?

### Question 3 (20 points)

Level of a basin  $Y$  (m) and the valve opening  $X$  (%) was measured. The results are shown in the table below.

$X$ (%)	10	20	30	40	50	60	70	80	90
$Y$ (m)	5,050	5,208	5,494	5,952	6,66	7,812	9,803	13,888	26,315

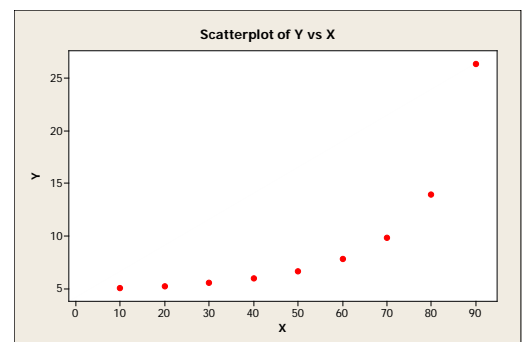
The chart is for reference only.

- (7 points) Considering the empirical model:

$$Y = a_0 + \frac{a_1}{1 - \left(\frac{X}{100}\right)^{1.5}} + \varepsilon$$

Where  $\varepsilon$  is a Gaussian noise with zero mean. You are asked to estimate the parameters and the of least squares method.

- (6 points) Does the value of  $X$  is significant or can it be considered "virtually" zero? Justify your answer.
- (3 points) Calculate the linear correlation coefficient (*Pearson coef.*) between  $X$  and  $Y$ . Can it be considered "significant"? Justify your answer by considering the sample size  $n = 9$  and an error  $\alpha = 0.05$ .



- d) (4 points) Repeat the previous question, but considering the two variables  $X$  and  $Z$ , with :  $Z = \frac{1}{1 - \left(\frac{X}{100}\right)^{1.5}}$ .

#### Question 4 (10 points)

Waiting time (in minutes) to get a response in an IT support company is distributed according to an exponential law (exponential distribution). An experimental study showed that the median is approximately 30 minutes:

- (4 points) Write the mathematical expression of the probability density governing the waiting time.
- (2 points) Determine the probability that the delay is less than 10 minutes.
- (2 points) Determine the probability that the delay is equal to 30 minutes.
- (2 points) Determine the probability that the delay is greater than 5 minutes.

#### Question 5 (10 points)

Scientific research is conducted to demonstrate a relation between **smoking** and the **lung cancer**. A statistical study was conducted on **13,420** people over five years. The results were classified into two categories: **Non-smoking** (2 months of cigarettes / week) and **Smoking**. The results are:

Profile	No Pathology	Pathology	Total
Non-smoking	10962	43	11005
Smoking	1874	541	2415
Total	12836	584	13420

- (5 points) With the assumption that the lung cancer is related to smoking. Formulate statistical hypotheses  $H_0$  and  $H_a$  for this contingency table.
- (5 points) Can we consider as probable the hypothesis that smoking is a risk for lung cancer? Justify your answer. Use error TYPE I = 5%.

**Question 6 (20 points)**

---

Consider  $X$  as a random variable which can be considered normal with average equal to 1 and a unit variance  $x \sim N(\mu_x = 1, \sigma_x^2 = 1)$ .  $Y$  is a new variable defined as:

$$Y = X^2 + 1$$

- a) (8 points) Calculate the mean and variance of the new variable  $Y$ .
- b) (12 points) A new transformation is define as  $z = a(x-1)$ , calculate the mean and variance of the new variable  $U$  :

$$U = Z^2 + 1$$

---

**End**