ORDRE DES INGÉNIEURS DU QUÉBEC

Examination Session of May 2020

All documentation allowed
Calculators : authorized models only
Duration of examination : 3 hours
The exam contains 5 questions, you must answer all of them.
Please write the algebraic formulae that apply before doing your calculations with numeric values. Show your whole work.

04-MB-2 - PROBABILITY AND STATISTICS

Question 1 (12 points)

According to the Ministry of Forests, Wildlife and Parks, 80% of the wood harvested in Quebec comes from public forests, the rest from private forests. In 2013, 17 millions m^3 of resinous wood (fir, pine, spruce, larch) were harvested in public forests, representing 82% of the wood volume coming from public land. From these forests, another 7% of poplars were harvested, and 11% of other species.

In private forests, 37% of resinous wood and 16% of poplar have been harvested.

(a) (6 points) What percentage of the total wood volume harvested in 2013 in Quebec is poplar?

(b) (6 points) A truckload of Quebec's resinous wood is intercepted at the border for random inspection before entering the United States. What is the probability that the wood comes from public forest?

Question 2 (25 points)

Minor paper wrinkling problems in a particular printing company occur randomly at a rate of 2 sheets per hour on each machine at full production. Consider that the Poisson distribution provides a good model for the frequency of wrinkling problems.

(a) (7 points) What is the probability that the next wrinkled sheet on machine 1 will occur in more than two hours?

(b) (8 points) What is the probability that at least two of the company's ten machines will undergo NO wrinkling problem in the next hour? Consider that the machines are all operating at full speed, and that they are independent of each other.

(c) (10 points) A typical promotionnal flyer order requires an average of 0.7 ml of ink per page, with a standard deviation of 0.3 ml. Under these conditions, what is the approximate probability that the last client's order (900 flyers of 4 pages each) will require more than 2.5 liters of ink? (Hint: Use the Central Limit Theorem.)

Question 3 (23 points)

An accelerated aging technique is used to test the shelf life of vacuum-packed food. The time before oxidation of 25 products chosen at random is measured, and an average duration of 88 days is obtained, with a standard deviation of 15 days. We assume that the normal distribution is a good model for this variable.

(a) (8 points) Calculate a 99% confidence interval for the average lifespan of vacuum-packed products subjected to accelerated aging.

(b) (8 points) We want to estimate the proportion of products that will end up with a damaged packaging. If we want a maximum margin of error of 0.05 on the estimation of the proportion, using a confidence level of 99%, how many products should be subjected to the accelerated aging process?

(c) (7 points) The real shelf life of refrigerated products (Y) is a function of the shelf life in accelerated aging (X) and of the storage temperature (T):

$$Y = 3\,X \;-\; 30\,T \;+\; 10$$

Considering that the mean temperature of refrigerators is $4\ ^{\circ}C$ with a standard deviation of $2\ ^{\circ}C$, and that the variables X and T are independent, give an estimate of the mean and of the standard deviation of the real shelf life of the sampled products.

Question 4 (20 points)

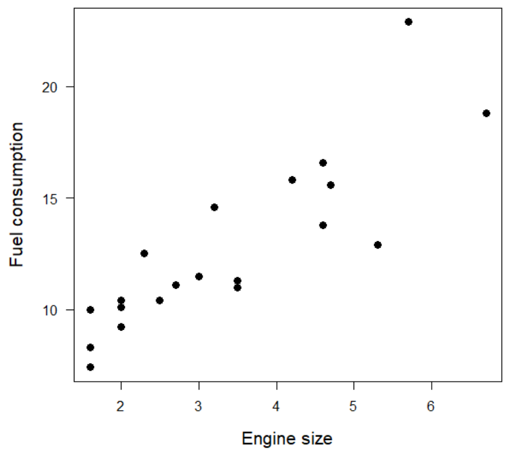
A company sells 20 kg bags of deicing salt. The quality control engineer suspects that the weight varies too much from one bag to another, and that the standard deviation would even exceed 0.9 kg, causing long-term losses for the company. He weighs 18 bags randomly selected from the warehouse, and obtains an average weight of 20.9 kg, with a standard deviation of 1.2 kg. (We assume that the normal distribution is a good model for the weight of salt bags.)

(a) (8 points) Perform a hypothesis test at a level of 5% to verify the engineer's suspicion.

(b) (12 points) In the summer, the same company sells bags of washed sand, and uses the same machine to fill them. The engineer wonders if the machine really puts the same weight of salt as sand in the bags. He weighs 10 sandbags and obtains an average of 20.2 kg, with a standard deviation of 1.1 kg. Is there a significant difference in means between the bags of salt and the bags of sand, with a 5% level? (The theoretical variability in the weight of the bags can be considered to be similar for sand and salt.)

Question 5 (20 points)

In the *Fuel Consumption Guide* published by the Canadian Department of Natural Resources, there are data on engine size (liters) and fuel consumption in the city ??(liters / 100 km) of several vehicle models. Here is the scatterplot of 22 pairs of measurements taken on cars chosen at random, as well as some statistics from this sample.



$$\overline{x} = 3.41$$

$$\overline{y} = 12.79$$

$$\sum_{i=1}^{22}(x_i - \overline{x})^2 = 45.18$$

$$\sum_{i=1}^{22}(y_i - \overline{y})^2 = 282.25$$

$$\sum_{i=1}^{22}(x_i - \overline{x})(y_i - \overline{y}) = 97.73$$

Corrélation(x, y)=0.865

(a) (4 points) What is the equation of the simple linear regression line fitted to these 22 observations?

(b) (7 points) Complete the analysis of variance table below and decide if the linear relationship is significant at a level of 1%.

Source of variation	Degrees of freedom	Sum of squares	Mean squares	F
Engine size	?	?	?	?
Error	?	?	3.541	
Total	?	?		

(c) (3 points) What is the proportion of the variability in fuel consumption explained by engine size?

(d) (6 points) Consider the Ford Focus, a car with an engine size that is 1 standard deviation below the sample mean $\overline{x}$.

- This cas has an engine size of ________________ liter(s).

- The linear model predicts that this type of car has a fuel consumption of ____________.

- The predicted fuel consumption is at

  (how many?) __________________ standard deviation(s)

  (above or below?) __________________ the average consumption.