MAT2377 – Exercises and Multiple Choice Questions II

(no answers for these ones)

1. Suppose that John and Tom are sitting in a classroom containing 9 students in total. A teacher randomly divides these 9 students into two groups: Group I with 4 students, Group II with 5 students.

(a) What is the probability that John is in Group I?

(b) If John is in Group I, what is the probability that Tom is also in Group I?

(c) What is the probability that John and Tom are in the same group?

2. In a certain group of people with two opposite-sex parents, it was found that 42% of them have alcoholic fathers, 8% of them have alcoholic mothers, and 48% of them have at least one alcoholic parent. If we randomly choose one individual from this group, what is the probability that:

(a) the selected individual has two alcoholic parents?

(b) the selected individual has an alcoholic mother but he/she does not have an alcoholic father?

(c) the selected individual has an alcoholic mother, if he/she has an alcoholic father?

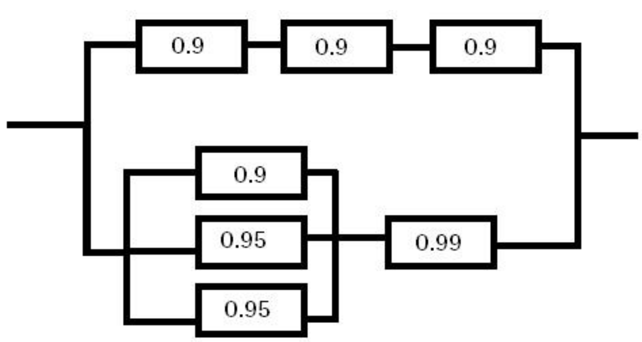
(d) the selected individual has an alcoholic mother, if he/she does not have an alcoholic father?

3. Assume that company A makes 75% of all electrocardiograph machines in the market, company B makes 20% of them, and company C makes the other 5%. The electrocardiographs machines made by company A have a 4% rate of defects, the company B machines have a 5% rate of defects, while the company C machines have a 8% rate of defects.

(a) If a randomly selected electrocardiograph machine is tested and is found to be defective. Find the probability that it was made by company A.

(b) Suppose we randomly select one electrocardiograph machine from the market. Find the probability that it was made by company A and it is not defective.

4. The following circuit operates only if there is a path of functional devices from left to right. The probability that each device functions is shown on the graph. Assume that devices fail independently. What is the probability that the circuit operates?



5. Let X be a discrete random variable. The following table shows its possible values x and the associated probabilities $P(X = x) = f(x)$.

x	-1	0	1	3
$f(x)$	2/8	3/8	2/8	1/8

(a) Verify that $f(x)$ is a probability mass function.

(b) Calculate $P(X < 1)$, $P(X \leq 1)$, and $P(X < 0.5 \text{ or } X > 2)$.

(c) Find the cumulative distribution function of X.

(d) Compute the mean and the variance of X.

6. About 1% of a certain type of light bulb fails during a 24-hour test. Failures are assumed to be independent. Consider a sign consisting of 10 such light bulbs. Let X be the number of light bulbs that fail during the 24-hour test.

(a) What distribution does X follow?

(b) What is the probability that the sign will burn for the full 24-hour test with no bulb failure?

(c) What is the probability that the sign will lose at least 3 bulbs during the 24-hour test?

(d) What is the probability that the sign will lose at least 2 and at most 4 bulbs during the test?

7. Water samples in European rivers are selected one at a time. Each sample has an independent probability of 3% of testing positive for algae blooms.

(a) What is the probability that the 7th sample is the first to test positive for algae blooms?

(b) Let T_2 be the number of samples required to obtain 2 positive algae blooms tests. What distribution does T_2 follow? What is the expected value and the standard deviation of T_2?

(c) What is the probability that the 5th sample is the second to test positive for algae blooms?

8. Suppose that a certain type of magnetic tape contains, on the average, 2 defects per 100 meters, according to a Poisson process.

(a) What is the probability that the next 100 meters of tape contain x defects, where $x = 0, 1, 2, \ldots$?

(b) What is the expected value and standard deviation of the number of defects per 100m?

(c) What is the expected number of defects in the next 300 meters of tape?

(d) What is the probability that there are more than 2 defects between meters 20 and 75?

9. The lifetime (in years) of a specific helicopter part follows an exponential distribution whose expectation is 3.2 years.

(a) What is the probability that one of these parts will stay operational for more than 4.4 years?

(b) What is the probability that one of these parts will stay operational between 3 and 9 months?

(c) If one of these parts is still operational after 3 years, what is the probability that it will remain operational for another 2 years?

10. Suppose that the number of cars passing a certain point of some road per minute between 8am and 10am on a Sunday morning follows a Poisson distribution with parameter $\lambda = 5$. A traffic inspector has arrived at the specified location during the aforementioned period. She notes the wait time W_1 for the next car to pass by the location.

(a) What distribution does W_1 follow? What are its expectation and variance?

(b) What is the probability she needs to wait more than 1 minute before the next car passes by?

(c) What is the probability that she needs to wait more than 1 minute for 2 cars to pass by?

11. Iron plates are required to have a certain thickness but each plate produced will differ slightly from each other due to properties of the material and uncertainties in the behaviour of the machines that make them. Let X be the plate thickness in mm of plates produced by a given machine. Using the machine's default setting, X follows a normal distribution with mean 10mm and standard deviation 0.02mm.

(a) What percentage of plates should be expected to be thinner than 9.97mm?

(b) What percentage of plates should be expected to be thicker than 10.05mm?

(c) What percentage of plates should be expected to deviate in thickness by more than 0.03mm from 10.00mm?

(d) Find $c > 0$ to ensure that 5% of plates are expected to deviate in thickness by more than c mm from 10.00mm?

(e) Given the c found in part (d), what percentage of plates are expected to deviate by more than c mm from 10.00mm if a slight adjustment in the machine shifts the expected value of X to 10.01mm.

12. A company manufactures resistors. The mean resistance for these resistors is 1000 ohms and the standard deviation is 200 ohms. We collect 50 resistors at random from the assembly line. Approximate the probability that the mean resistance of these 50 resistors will be 1005 ohms or larger.

13. The lifetime of a 75 watts lightbulb is normally distributed with mean μ hours and variance σ^2. From a random sample of 20 lightbulbs, the sample mean is 1014 hours and the sample standard deviation is 25 hours.

(a) Compute a 95% confidence interval for μ.

(b) Suppose that $\sigma = 25$ hours. We want to re-construct a 95% confidence interval for μ with the interval length no longer that 9 hours. Determine the required sample size.

14. The mean breaking strength of a certain type of fibre is required to be larger than 200 psi. Past experience has shown that we can assume that the breaking strength is normally distributed with standard deviation $\sigma = 4.5$ psi. A sample of 8 fibres yielded breakage at the following pressures (in psi):

$$210, 206, 198, 202, 201, 198, 199, 205.$$

(a) Formulate a null hypothesis versus an alternative hypothesis to verify that this type of fibre is acceptable.

(b) Compute a test statistic to test the hypotheses from part (a).

(c) Based on the value of the test statistic in part (a): (i) give the conclusion at $\alpha = 5\%$; (ii) give the conclusion at $\alpha = 10\%$.

(d) Suppose that a sample of 30 fibers yielded a mean breakage of $\bar{x} = 202.375$ psi. For this new study, compute a test statistic to test the hypotheses from part (a) and give the conclusion at $\alpha = 5\%$.

15. Suppose that the probability density function (p.d.f.) of the life (in weeks) of a certain part is

$$f(x) = \frac{3\,x^2}{(400)^3}, \quad 0 \leq x < 400.$$

(a) Compute the probability the a certain part will fail in less than 200 weeks.

(b) Compute the mean lifetime of a part and the standard deviation of the lifetime of a part.

(c) Suppose that we select $n = 50$ parts at random. Approximate the probability that the average lifetime for these 50 parts will be less than 275 weeks?

16. Cloud seeding has been studied for many decades as a weather modification procedure. The rainfall in acre-feet from 20 clouds that were selected at random and seeded with silver nitrate are recorded. We are displaying the data in an increasing order below :

18	18.8	19.8	21.2	21.8	22.3	23.4
24.7	25	26.7	26.9	27.1	27.1	27.9
29.2	30.7	31.6	31.8	31.9	34.8	

(a) Compute the sample mean and the sample standard deviation.

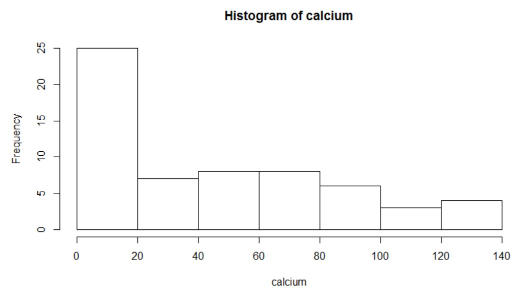
(b) Find the first, second, and third sample quartile.

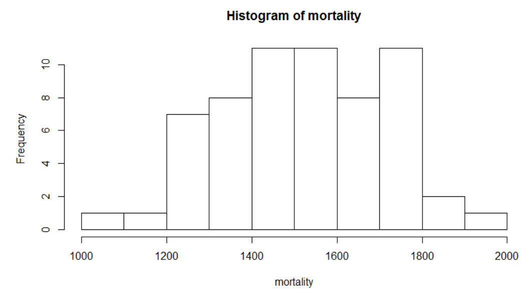
(c) Are there any outliers in this dataset? (Explain)

17. In an investigation of environmental causes of disease, data were collected on the annual mortality rate (deaths per 100 000) for males in 61 large towns in England and Wales. In addition, the water hardness was recorded as the calcium concentration (parts per million, ppm) in the drinking water. Below, we provide some descriptive statistics for both variables (mortality and calcium concentration).

```
> summary(calcium)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
   5.00   14.00   39.00   47.18   75.00  138.00
> summary(mortality)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
   1096    1379    1555    1524    1668    1987
```

Here are histograms for both variables.

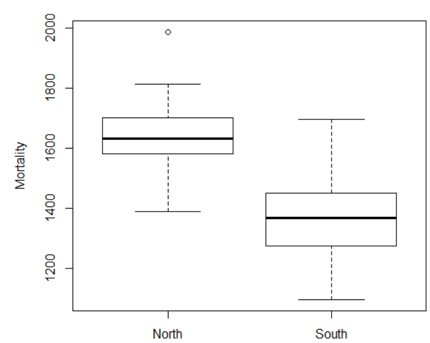




(a) For each variable (i.e. calcium concentration and mortality) describe its distribution in shape.

(b) For each variable, are there any outliers in the sample? (Explain.)

(c) Suppose that the towns in the sample are either South or North of Derby. To describe the mortality according to region, we produce the following side-by-side boxplots.



Answer the following questions based on the above boxplots.

(i) Which region (North or South) has the town with the largest mortality?

(ii) Which region (North or South) has the town with the smallest mortality?

(iii) In terms of the central tendency, which region has a higher mortality?

(iv) Is the mortality more dispersed in the North or the South?

18. Many students at uOttawa are fluent in both French and English. However, some students are only fluent in English and some are only fluent in French. Furthermore, all students at uOttawa are fluent in either French or English. Suppose that we randomly select a uOttawa student. Let A be the event that the student is **only** fluent in English and let B be the event that the student is **only** fluent in French. Which of the following statements is **incorrect**?

(a) $P(A|B) = P(B|A)$ **(b)** $P(A \cup B) = P(A) + P(B)$ **(c)** $P(A^c \cap B) = P(B)$
(d) A and B are mutually exclusive. **(e)** A and B are independent.

19. A study is conducted on male employees of age 50-64 working in a chemical plant. These workers are divided into three age groups as follows: 30% have age 50-54, 40% have age 55-59, and 30% have age 60-64. The annual national mortality rates for male employees working in similar conditions are: 5% in 50-54-year-old men, 7% in 55-59 year-old men, and 13% in 60-64-year-old men. If one of the workers in the chemical plant died during the past year, what is the probability that this man was in the age group 55-59?

20. Let X be a discrete random variable with the following probability mass function $f_X(x)$.

x	0	1	2	3	4
$f_X(x)$	0.05	0.25	0.33	0.09	0.28

Calculate $P(X - E[X] < 0.5)$.

21. The number of car accidents in an area follows a Poisson process with rate 2 accidents per week. Suppose that we observed one car accident in the first week. What is the probability that we will observe at most one car accident in the next 3 days?

22. Let X be a discrete random variable with mean $\mu_X = 12.5$ and variance $\sigma_X^2 = 0.36$. Let $Y = -2X + 30$. Find the mean and the standard deviation of Y.

23. Let X be a discrete random variable with support $\{0, 1, 2, 3, 4\}$. Its cumulative distribution function $F_X(x)$ is given as follows.

x	0	1	2	3	4
$F_X(x)$	0.2	0.3	0.4	0.8	1

Compute $E[X]$.

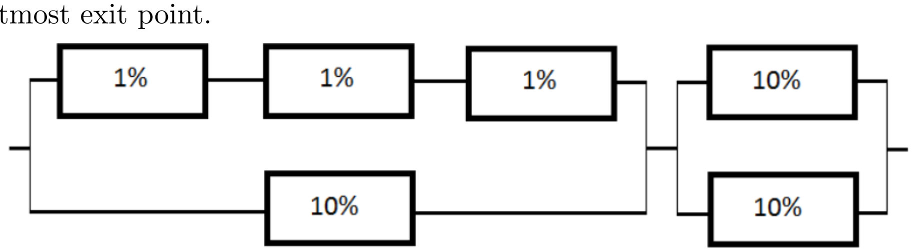
24. Suppose there are 14 boys and 6 girls in a classroom. A teacher randomly selects 5 students from the class without replacement. What is the probably that the selected students contain 2 girls?

25. A drone is tested prior to being marketed to the general public. The test results on 250 independent samples are shown below:

	Steering Defect	No Steering Defect
Propulsion Defect	15	30
No Propulsion Defect	35	170

Assume that these relative proportions are perfectly representative of the product (and of its flaws) outside the sample. Assuming that defects occur independently, what is the probability that each of 4 drones purchased by an individual would have at least one of the defects?

26. The following system operates only if there is a path of functional devices from the leftmost entry point to the rightmost exit point.



The probability that each device fails to function is as shown in the diagram above. Assuming independence of failure for the various devices, what is the probability that the system operates?

27. Ottawa Public Health collects beach water samples every day during the summer. A non-swimming advisory is issued for a beach if the bacteria level in water is too high. Assume that, on a randomly chosen day, Mooney‘s Bay beach is safe to swim with probability 0.9452. What is the probability that a non-swim advisory will be issued for Mooney‘s Bay beach at least twice in the next 5 years on July 15th?

28. An assembly line produces a specific auto part. The parts are independently sampled one at a time and tested for compliance. Historically, 10% of the parts have been found to be non-compliant. What is the probability that the 6th sampled part will be the 3rd part found to be non-compliant?

29. Let X and Y be a discrete random variable with joint probability mass function $f_{X,Y}(x,y)$ given by

x	y	$f_{XY}(x,y)$
0	0	1/12
0	2	3/12
1	1	1/12
1	2	4/12
2	0	2/12
2	2	1/12

Compute $P(X+Y \geq 2)$.

30. Consider an ordinary deck of 52 playing cards (13 cards – 2 to 10, jack, queen, king, ace in each suit; 2 suits in each colour – diamonds, hearts are red, clubs, spades are black). The deck is shuffled and a card is picked randomly. Consider the following events:

1. A: the card is red;
2. B: the card is a jack, queen, or king of diamonds;
3. C: the card is an ace.

What is $P((A \cap B^c) \cup C)$?

31. In a factory, machines 1, 2, and 3 produce screws of the same length, with 2%, 1%, and 3% defective screws, respectively. Of the total production of screws in the factory, the machines produces 35%, 25%, and 40%, respectively. If a screw is selected at random from the total screws produces in a day and is found to be defective, what is the probability that it was produced by machine 1?

32. Suppose that the probability of germination of a beet seed is 80%. If we plant 20 seeds and can assume that the germination of one seed is independent of another seed, what is the probability that 18 or fewer seeds germinate?

33. Assume that arrivals of small aircrafts at an airport can be modeled by a Poisson process with rate 2 aircrafts per hour. What is the probability that one has to wait at least 3 hours for the arrival of 3 aircrafts?

34. Let X be a random variable following a normal distribution with mean 14 and variance 4. Determine a value c such that $P(X - 2 < c) = 0.95$.

35. Let the joint probability mass function of two discrete random variables X, Y be

$$f(x, y) = \frac{xy^2}{30}, \quad x = 1, 2, 3 \text{ and } y = 1, 2.$$

Compute $\mathrm{E}[X + Y]$.

36. The concentration of nicotine was measured in a random sample of 40 cigars. The data are displayed below, from smallest to largest:

```
72, 85, 110,124,137,140,147,151,158,163,164,165,167,168,169,169,170,174,175,175,
179,179,182,185,186,188,190,192,193,197,203,208,209,211,217,228,231,237,246,256.
```

How many outliers do we have in this dataset?

37. A new type of electronic flash for cameras will last on average of $\mu = 5000$ hours with a standard deviation of $\sigma = 500$ hours. A quality control engineer selects a random sample of 100 flashes. What is the probability that the mean lifetime of these 100 flashes will be greater than 4928 hours?

38. A hockey puck manufacturer claims that its process produces pucks with a mean weight of 163 grams and a standard deviation of 5 grams. A random sample of n pucks is going to be collected. We plan to use the sample mean $\bar{X}$ to estimate the population mean. Determine the minimal sample size n so that $P(|\bar{X} - 163| < 2) = 0.95$. (Assume n is large.)

39. Two candidates (A and B) are running for an officer position. A poll is conducted: 120 voters are selected randomly and asked for their preference. Among the selected voters, 52% support A and 48% support B. Provide a 95% confidence interval for the true support rate of candidate A in the population.

40. The tensile strength of manila ropes follows a normal distribution. A random sample of 16 manila ropes has a sample mean strength 4450 kg and sample standard deviation 115 kg. Suppose that we want to test whether the mean strength of manila rope is less than 4500 kg. At a significance level $\alpha = 0.05$, the value of the test statistic and the conclusion for this test are:

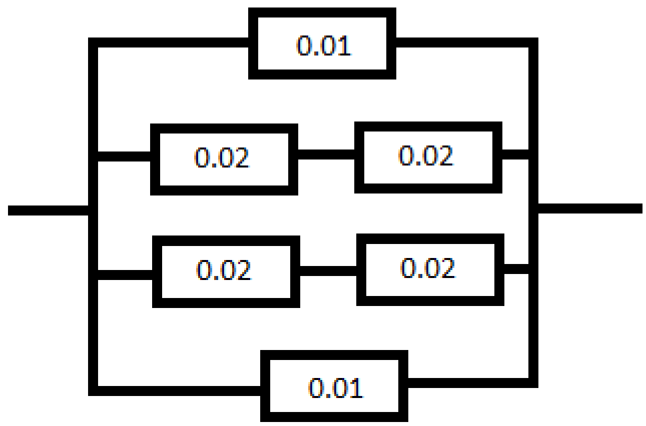
(a) -1.739; Do not reject H_0 **(b)** -1.739; Reject H_0 **(c)** -1.753; Do not reject H_0
(d) -1.753; Reject H_0 **(e)** 1.739; Reject H_0

41. Twenty-four girls in Grades 9 and 10 are put on a training program. Their time for a 40-yard dash is recorded before and after participating in a training program. The differences between the before-training time and the after-training time for those 24 girls are measured, so that positive difference values represent improvement in the 40-yard dash time. Suppose that the values of those differences follow a normal distribution and they have a sample mean 0.079 min and a sample standard deviation 0.255 min. We conduct a statistical test to check whether this training program can reduce the mean finish time of 40-yard dash. What is the range of p-value for this test?

(a) $(0, 0.025)$ **(b)** $(0.025, 0.05)$ **(c)** $(0.05, 0.1)$ **(d)** $(0.1, 0.15)$ **(e)** $(0.15, 0.2)$

42. Three awards (research, teaching, and service) will be given to 18 graduate students in a math department. Suppose each student can receive at most one award. How many possible award outcomes are there?

43. The following circuit operates if and only if there is a path of functional devices from left to right. Assume that the devices fail independently and that the probability of ***failure*** of each device is as shown. What is the probability that the circuit does **not** operate?



44. There are two male students and two female students in a classroom. We randomly select two students from this classroom without replacement. Let X be the number of male students among the two selected students. Let $F(x)$ be the cumulative distribution function of X and $f(x)$ be the probability mass function of X. Which one of the following statements is **incorrect**?

(a) $F(x) > 0$ when $x = 0$
(b) $f(x) > 0$ when $x = 0$
(c) $F(x) = 0$ when $x = 3$
(d) $f(x) = 0$ when $x = 3$
(e) $f(x) = F(x)$ when $x < 0$

45. In a certain manufacture process, it is known that 1% of products are defective. Assume that products are manufactured one-by-one independently. What is the probability that the 3rd product will the first defective one?

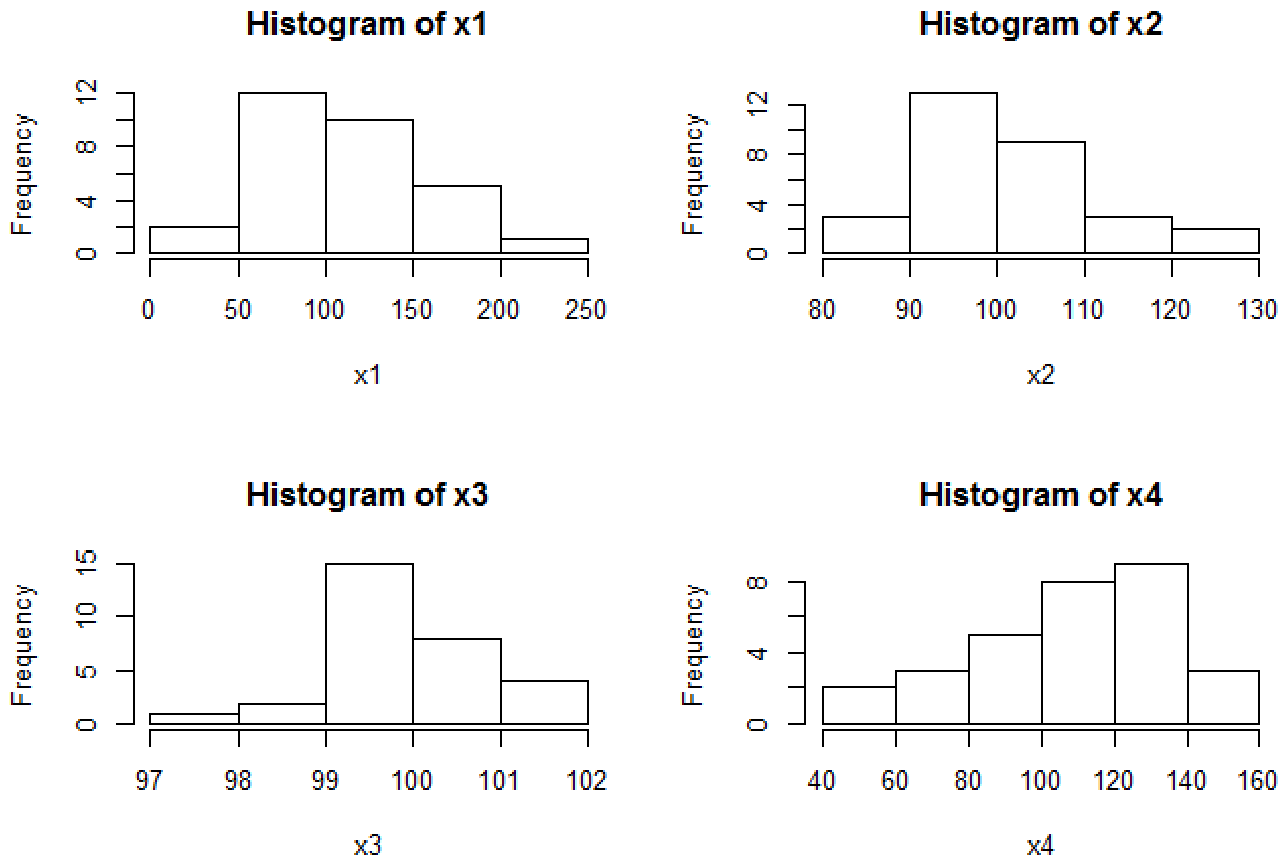
46. If the probability density function of a random variable X is given by

$$f(x) = k\,x^3, \quad 0 < x < 1.$$

Compute the probability that X will be between 1/4 and 3/4.

(a) 0.3125 **(b)** 0.0781 **(c)** 0.9267 **(d)** 0.4523 **(e)** 0.625

47. We have data which gives the speed of a vehicle at the time of accident for accidents that happened on 4 highways. The data are saved in `R` in variables x_1, x_2, x_3, x_4, respectively. Below are the histograms for these variables:



For these variables, the sample standard deviations are given below in an increasing order:

1.0 10.5 27.2 48.5.

Identify the sample standard deviation of each variable.

(a) $s_{x_1} = 1.0$; $s_{x_2} = 48.5$; $s_{x_3} = 27.2$; $s_{x_4} = 10.5$.
(b) $s_{x_1} = 48.5$; $s_{x_2} = 1.0$; $s_{x_3} = 27.2$; $s_{x_4} = 10.5$.
(c) $s_{x_1} = 27.2$; $s_{x_2} = 10.5$; $s_{x_3} = 1.0$; $s_{x_4} = 48.5$.
(d) $s_{x_1} = 48.5$; $s_{x_2} = 10.5$; $s_{x_3} = 1.0$; $s_{x_4} = 27.2$.
(e) $s_{x_1} = 10.5$; $s_{x_2} = 1.0$; $s_{x_3} = 48.5$; $s_{x_4} = 27.2$.

48. Among 2046 cars made by Company A in 1999, 56 had a problem in the brake system. Suppose that one wants to know whether the brake system defective rate for this type of car is less that 4%. Formulate the null and alternative hypotheses. Compute the p-value for this test.

(a) $H_0 : p = 0.04$ vs. $H_1 : p > 0.04$. p-value = 0.9982.
(b) $H_0 : p = 0.04$ vs. $H_1 : p < 0.04$. p-value = 0.0036.
(c) $H_0 : p = 0.04$ vs. $H_1 : p < 0.04$. p-value = 0.0018.
(d) $H_0 : p = 0.04$ vs. $H_1 : p \neq 0.04$. p-value = 0.0036.
(e) $H_0 : p = 0.04$ vs. $H_1 : p < 0.04$. p-value = 0.0155.

49. The blood pressure X and the calcium level Y were measured on a random sample of 38 persons. Based on this data, the estimated regression line is given by:

$$\widehat{y} = -2.2 + 1.725x.$$

The sample standard deviations are $s_x = 0.35$ and $s_y = 1.667$. Find the sample correlation between the variables X and Y.

50. Let A and B be two events such that $P(A) > 0$ and $P(B) > 0$. Which one of the following statements is **false**?

(a) $(A \cup B)^c = A^c \cap B^c$
(b) $A^c \cap B$ and $A \cap B^c$ are mutually exclusive.
(c) $P(A|B) + P(A^c|B) = 1$
(d) If $P(A \cap B) = 0$, then A and B are independent.
(e) If A and B are independent, then $P(A|B) = P(A)$.

51. It takes a Christmas tree about 10 years to grow from seed to a size ready for cutting. We want to estimate the average height μ of a 4-year Christmas tree which has been grown from a seed. Assume that the height of a 4-year tree is normally distributed. A sample of 20 trees has a mean height 25.25 cm and a sample standard deviation 4.5 cm. This sample produces a confidence interval (C.I.) for μ of length 2.673. Determine the confidence level of this C.I.

52. According to a nationwide survey conducted by Statistics Canada, the mean birth weight in Canada is 3.4kg. A doctor would like to gain evidence for the hypothesis that urban mothers deliver babies whose birth weights are greater than 3.4kg. She conducted a statistical test based on 125 Canadian urban newborns with a sample standard deviation 0.78kg. Suppose that the p-value of this test is 0.0158. What is the mean weight (in kg) for those 125 Canadian urban newborns?

53. The mean weight of a newborn baby in North America is 120 ounces (oz). We want to test the hypothesis that mothers with low socioeconomic status have babies whose weight at birth is lower than 120 oz. Let μ be the mean weight of a newborn baby whose mother has a low socioeconomic status. Set-up a test of hypotheses and explain when type I error or type II error occur by choosing the correct statement from the list below.

(a) $H_0 : \mu = 120$ versus $H_1 : \mu < 120$. A Type II error occurs when we conclude that the mean weight of a newborn baby whose mother has a low socioeconomic status is lower than 120 oz, when in fact it is not true.
(b) $H_0 : \mu = 120$ versus $H_1 : \mu < 120$. A Type I error occurs when we conclude that the mean weight of a newborn baby whose mother has a low socioeconomic status is lower than 120 oz, when in fact it is not true.
(c) $H_0 : \mu \geq 120$ versus $H_1 : \mu < 120$. A Type I error occurs when we conclude that the mean weight of a newborn baby whose mother has a low socioeconomic status is lower than 120 oz, when in fact it is not true.
(d) $H_0 : \mu \geq 120$ versus $H_1 : \mu < 120$. A Type II error occurs when we conclude that the mean of a newborn baby whose mother has a low socioeconomic status is 120 oz, but in fact this weight is lower than 120 oz.
(e) $H_0 : \mu = 120$ versus $H_1 : \mu < 120$. A Type I error occurs when we conclude that the mean of a newborn baby whose mother has a low socioeconomic status is 120 oz, but in fact this weight is lower than 120 oz.

54. A company produces orange juice bottles with a volume of approximately 2 litres each. One machine fills half of each bottle with concentrate, and another machine fills the other half with water. Assume the two machines work independently. The volume (in litres) of concentrate poured by the first machine follows a normal distribution with mean 0.98 and variance 0.0009. The volume of water (in litres) poured by the second machine follows a normal distribution with mean 1.02 and variance 0.0016. A bottle of orange juice produced by this company is therefore a mixture of water and concentrate. What is the probability that a bottle contains more than 1.98 litres of juice?