Statistics Tutorial – Questions

Exercises:

(i) A full set of test scores are (56, 78, 62, 38, 90, 77, 86, 51, 84, 92). Calculate the mean score and the standard deviation.

(ii) Calculate the correlation coefficient to 3 decimal places between the following datasets.

Event A	Event B
25	36
28	41
20	29
22	37
26	36
19	26
31	45

(iii) Models of cars are classified according to their engine capacity, measured in cubic centimetres (cc). Seven new cars are tested to find the distance they will travel, from rest, during the first 10 seconds. The results are given in the table below.

Engine Capacity (cc)	800	1000	1200	1500	2000	2400	3000
Distance (m)	85	120	140	175	190	235	280

(a) Draw a scatter plot for these results.
(b) Calculate the mean engine capacity.
(c) Calculate the mean distance.
(d) Plot and label the point M (mean engine capacity, mean distance) on the scatter diagram.
(e) Draw a line of best fit through the point M.
(f) By using the scatter diagram and showing your method clearly, estimate the distance that a car with an engine capacity of 1,800 cc will travel in the first 10 seconds.
(g) Calculate the correlation coefficient, r.

(iv) Find the following probabilities:
   a. $P(z < 1.63)$
   b. $P(z < -2.1)$
   c. $P(z > -0.96)$

(v) A random sample of 300 is taken from a population having mean 50 and standard deviation of 8. Find the probability that the sample mean is greater than 50.8.

(vi) The life of a certain make of light bulb is known to be normally distributed with a mean life of 2,000 hours and a standard deviation of 120 hours. The quality control department in the factory that manufacturers the light bulbs test a random selection of 600 bulbs to destruction. How many of those bulbs would you expect to last:
   a. Less than 1,800 hours
   b. More than 2,200 hours
   c. Between 1,800 hours and 2,200 hours

Statistics Tutorial – Questions

Past Exam Questions:

Question 1: **2023 Paper 2 Q5**

Rohan has a large number of small cubes. The cubes are identical in size. Some of the cubes are red, some are green, and the rest are blue. Rohan carries out an experiment in which he picks out 5 different cubes at random. He records the number of cubes of each colour and replaces the 5 cubes. He then repeats this process a number of times. The table below shows the number of cubes of each colour the first 7 times Rohan does this, labelled Trial A to Trial G

Trial	A	B	C	D	E	F	G
No. Red Cubes	0	3	2	2	4	5	1
No. Green Cubes	4	2	0	3	0	0	2
No. Blue Cubes	1	0	3	0	1	0	2

(i) Work out the mean and standard deviation of the number of red cubes per trial for these 7 trials. Give each answer correct to 1 decimal place.

(ii) Work out the correlation coefficient between the number of red cubes and the number of green cubes per trial, for these 7 trials. Give you answer correct to 3 decimal places.

(iii) Rohan repeats this experiment a large number of times. Explain why you would expect the correlation coefficient between the number of red cubes and number of green cubes per trial to be negative?

Question 2: **2023 Paper 2 Q8**

An online word game involves trying to guess a five-letter word in as few attempts as possible. For this game, each player is given a score s, where $s \in \mathbb{R}$, based on how many attempts it takes them to guess the word.

(a) In Ireland, players' scores are approximately normally distributed, with a mean of 3·87 and a standard deviation of 0·36. A player is selected at random from the players in Ireland. Find the probability that this player has a score of less than 3·5.

(b) A random sample of 64 Galway players has a mean score of 3·74. Based on this, a local newspaper claims that Galway players have a different mean score to players in Ireland.

i. Use the information about this sample to construct a 95% confidence interval for the mean score of all Galway players. Use the standard deviation of 0·36 in your calculations.

ii. Carry out a hypothesis test at the 5% level of significance to test the newspaper's claim that Galway players have a different mean score to players in Ireland. State your null hypothesis, state your alternative hypothesis, state your conclusion, and give a reason for your conclusion.

(c) A national newspaper conducts a survey on a random sample of n teenagers in Ireland. 35% of the sample said they play the online word game every day. Based on this, a 95% confidence interval for the percentage, p, of all teenagers in Ireland who play the game every day was calculated, as accurately as possible. Correct to one decimal place, this interval was: 26·5% $\leq p \leq$ 43·5%. Use this to work out the value of n, the number of teenagers surveyed.

Statistics Tutorial – Questions

Question 3: **2021 Paper 2 Q8**

(a) In a school all First Years sat a common maths exam. The results, in integer values, were normally distributed with a mean of 176 marks and a standard deviation of 36 marks. The top 10% of students will go forward to a county maths competition.

i. Find the minimum mark needed on the exam to progress to the county stage.

ii. The school awarded a Certificate of Merit to any student who achieved between 165 marks and 210 marks. Find the percentage of First Years who received the Certificate of Merit.

(b) A news report claimed that 6th year students in Ireland studied an average of 21 hours per week, outside of class time. A Leaving Cert class surveyed 60 students in 6th year, chosen at random, from different schools. It found that the average study time was 19·8 hours and the standard deviation was 5·2 hours.

i. Find the test statistic (the z-score) of this sample mean.

ii. Find the p-value of this test statistic. Comment on what can be concluded from its value, in a two-tailed hypothesis test at the 5% level of significance, in relation to the news report claim.

Question 4: **2019 Paper 2 Q8**

(a) A motoring magazine collected data on cars on a particular stretch of road. Certain details on 800 cars were recorded.

i. The ages of the 800 cars were recorded. 174 of them were new (less than 1 year old). Find the 95% confidence interval for the proportion of new cars on this road. Give your answer correct to 4 significant figures.

ii. The data on the speeds of these 800 vehicles is normally distributed with an average speed of 87·3 km per hour and a standard deviation of 12 km per hour. What proportion of cars on this stretch of road would you expect to find travelling at over 95 km per hour?

iii. The driver of a car was told that 70% of all the speeds recorded were higher than his speed. Find the speed at which this driver was recorded. Give your answer correct to the nearest whole number.

(b)

i. A road safety programme was carried out in the area using posters, signs and radio slots. After the programme the motoring magazine recorded the speeds of 100 passing cars. The magazine carried out a hypothesis test, at the 5% level of significance, to determine whether the average speed had changed. The p-value of the test was 0·024. What can the magazine conclude based on this p-value? Give a reason for your answer.

ii. The magazine found that the average speed of this sample was lower than the previously established average speed of 87·3 km per hour. Find the average speed of the cars in this sample, correct to 1 decimal place.

