Statistics Revision

1. The weight of a fully harvest apple is normally distributed with mean 200 grams and standard deviation 12 grams. The weight of a windfall apple (those which fell from the trees before harvest) is independently normally distributed with mean 175 grams and standard deviation 9 grams. Determine, correct to 3 significant figures, the probability that

   (i) the weight of a randomly chosen windfall apple is between 172.5 grams and 174 grams.

   (ii) A sample of 2 randomly chosen fully harvest apples weighs more than two times the weight of a windfall apple.

   (iii) Find the maximum weight which will be exceeded by the total weight of a fully harvested apple and a windfall apple with a probability of at least 0.95.

2. An office block has a lift that can carry a maximum permitted load of 500kg. It is known that the weights of men and women using the lift can be modelled by normal distributions. The men's weights have mean 80kg and standard deviation 12kg. The women's weights have mean 56kg and standard deviation 6kg.

   On one occasion, 3 men and 4 women enter the lift. Calculate the probability that
   (i) their total weight exceeds the maximum permitted load,
   (ii) the total weight of the 3 men exceeds that of the 4 women.

3. In a country, 35% of the female population have blood pressure that is considered HIGH, and 20% of them with blood pressure considered LOW. Assuming the blood pressure of this female population is normally distributed with mean 130 and standard deviation 11, find the range (a,b) of the blood pressure for females that are considered neither HIGH or LOW. Give your answers correct to 3 significant figures.

   (i) Find the probability that the mean blood pressure of 2 females falls within the range (a,b).

   (ii) 10 females are chosen at random. What is the probability that there are not more than 3 females with blood pressure that is neither HIGH nor LOW?

4. Both Sue and Tom walk from home to school every morning. The time they take to travel to school, in minutes, may be assumed to follow independent Normal distributions with means and standard deviations as shown in the table below:

	Mean(Min)	Standard Deviation(Min)
Sue	10	1.5
Tom	18	4

   On a particular morning, Sue and Tom left home at 7.20 am and 7.15 am respectively.

(i) Find the probability that both Sue and Tom were in school by 7.30am.

(ii) Find the probability that Tom reached the school later than Sue.

(iii) Find the probability that Sue took more than half the time taken by Tom to travel to school.

5. There are two kinds of rare fishes found in the country of Aba: the Solomon fish and the Sheba fish. The masses of a Solomon fish and a Sheba fish are independent and normally distributed with means 2kg and 3 kg respectively, and standard deviation 1kg and 4 kg respectively.

Find the probability that the total mass of 2 Solomon fish differ from twice the mass of 1 Sheba fish by at most 2 kg.

A Solomon fish costs $5 per kg and a Sheba fish costs $8 per kg. A restaurant purchased 10 Solomon fish and 5 Sheba fish.

Find the probability that the restaurant has to pay more than $180 in total.

6. The mass, X grams, of a randomly chosen orange in a certain orchard is normally distributed with mean 320 and variance 50.

(i) Find the mass m such that $P(X<m)=15P(X>m)$.

(ii) 60 oranges are randomly selected from the orchard. Using a suitable approximation, find the probability that there are more than 56 oranges each having mass less than m grams.

7. At a games stall, an electronic device displays a three-digit number from 000 to 999. When a button is pressed, each number has an equal chance of being displayed. Find the probability that the three digits displayed

(i) have a sum of at least 24,

(ii) are identical, given that the sum of the three numbers is at least 24.

8. Six girls, including Janet, and five boys, including Carl and Dean, sat is a circle to play a game.

(i) Find the probability that all the boys are separated;

(ii) Find the probability that Janet sat between Carl and Dea, given that Janet, Carl and Dean sat together

*9. Aaron is an avid soccer fan who enjoys going to the stadium to watch the game.

There are soccer matches scheduled at the stadium next Friday, Saturday and Sunday. The probability that Aaron will go to the stadium on Friday is $\frac{1}{5}$. On each of the other days, the probability that he goes, given that he **has** gone the previous day, is $\frac{1}{6}$ and the probability that he goes, given that he **has not** gone the previous day, is $\frac{2}{3}$.

Find the probability that
(i) Aaron will go to the stadium on Sunday;

(ii) Aaron will not go to the stadium on Sunday given that he will go to the stadium on Friday;

(iii) Aaron will go to the stadium on Friday given that he will go to the stadium on Sunday.

10. The diameters, in cm, of 100 randomly chosen plastic doorknobs were measured. The results are summarised as $\sum x = 249$, $\sum x^2 = 779$.

(i) Find the unbiased estimates of the population mean and variance of the diameters of all knobs produced.

(ii) Find the probability that the mean diameter is less than 2.7cm.

(iii) 150 plastic doorknobs are chosen. Find the probability that the total diameter is less than 370cm.

11. A wholesaler grades cabbages according to their weights. Cabbages with a weight exceeding 1.6kg are classified as large and cabbages with a weight less than 1kg are classified as small. A large batch of cabbages is graded and it is found that 5% are large and 7% are small. Assuming a normal distribution, find the mean weight of a randomly chosen cabbage.

Three cabbages are randomly selected. Find the probability that exactly one of them is large and exactly one of them is small.

12. The mass of coffee in a randomly chosen jar sold by a certain company may be taken to have a normal distribution with standard deviation 2.8g. It is claimed that the mean mass of coffee per jar is 57.5g. To investigate the company's claim, a random sample of 50 jars are examined and the sample mean is found to be 56.9g.
(i) Test at the 2% significance level whether the company is over-stating the mass of coffee per jar.

(ii) What sample mean value would we need in order to get a different conclusion to the test in (i)?

(iii) In order to obtain a more accurate estimate of the mean weight, it is proposed to take a larger random sample. Estimate the sample size that will be needed if it is intended that the probability of the sample mean being within 0.5g off the population mean is about 0.98.

(iv) If the standard deviation 2.8g was actually derived from the sample of 50, state the test statistic and calculate test statistic value for the test in (i).

13. The height, x mm, and the weight, y kg, of each boy in a school were recorded and a scatter diagram was then plotted.

The equation of the line of regression of Y on X is found to be Y=0.09X-90. The sample heights are found distributed about a mean of 1600mm with standard deviation 120 mm and the standard deviation of the weights is 12kg.

(i) Show that the sample mean of the weights is 54.

(ii) Find the equation of the line of regression of X on Y.

(iii) Find the linear (product moment) correlation coefficient between X and Y. Comment on what this value implies about the regression lines.

(iv) Find the expected weight of the boy whose height is 1.5m.

14. (a)The following data relate to the percentage scores on a physical fitness test and the weights (in kilograms) of ten primary school pupils.

Pupil	1	2	3	4	5	6	7	8	9	10
Weight (x)	41.8	38.6	54.1	38.6	44.9	52.4	30.2	41.4	38.4	32.1
Score(y)	58	60	59	72	62	54	81	62	86	94

Given that $\sum\frac{x-30}{0.1}=1125,\ \sum(\frac{x-30}{0.1})^2=179671,$

$\sum(y-50)=188,\ \sum(y-50)^2=5226,\ \sum(\frac{x-30}{0.1})(y-50)=13927,$

Calculate the product-moment correlation coefficient between the weights and the scores in the physical fitness test and comment on your result.

(b) The length (L mm) and width (W mm) of each 20 individuals of a single species of fossil are measured. A summary of the results is:

$\sum L=400.20,\ \sum W=176.00,\ \sum LW=3700.20,\ \sum L^2=8151.32,$
$\sum W^2=1780.52.$

Obtain an equation of the line of regression from which it is possible to estimate the length of a fossil of the same species whose width is known, giving the values of the coefficients to 2 decimal places.

15. The random variable X has a binomial distribution with mean 4 and variance 3.6. A large random sample of n observations of X is taken. Given that the probability

that the sample mean exceeds 4.1 is at most 0.01, find the set of possible values of n.

State clearly whether the central limit theorem has been used, and if so, where.

16. 65 boys and 60 girls are examined, of which 20 boys and 25 girls are found to be shortsighted.

(a) A child is selected at random.
Let A represent the event that the child selected is a girl and
B represent the even that the child selected is shortsighted.
Find (i) P(A)
(ii)P($A \cup B$)
(iii) P($A \cap B$)
(iv) P(A |B)

(b) If 3 children are selected at random, find the probability that
(i) all 3 children selected are shortsighted.
(ii) exactly one child is shortsighted.
(iii) there are 2 girls and 1 boy, where exactly 1 girl and 1 boy are both shortsighted.

17. In a large population of students, there is a 0.9% chance that a student will be late on any school day.

(i) In a week of 5 school days from Monday to Friday, find the probability that Thursday is the second day that a randomly chosen student is late.

(ii) Show that the probability in which any particular class of 25 students has no latecomers for 5 days is 0.323, correct to 3 significant figures.

(iii) 80 weeks of 5 school days are chosen at random. Using a suitable approximation, find the probability that, in a class of 25 students, there is at least 1 latecomer for more than 50 weeks.

18. The length, in cm, of a floor tile is a normal variable with mean 30 and standard deviation 0.2. The length, in cm, of a wall tile is a normal variable with mean 10 and standard deviation 0.15.

(a) Find the probability that the sum of the lengths of five randomly chosen floor tiles is between 150.5cm and 151.5cm.

(b) The random variable S is the sum of the lengths of 6 randomly chosen wall tiles and the random variable W is twice the length of a randomly chosen floor tile. Find $P(S<W+1)$.

(c) The average length of 25 randomly chosen floor tiles is denoted by $\overline{L}$.
Find the value of a such that $P(|\overline{L}-30| \geq a)=0.02$.