

A LEVEL Cambridge Topical Past Papers

FURTHER STATISTICS

2017— 2023

Chapter 18	Further Work On Distributions	Page 1
Chapter 19	Inference Using Normal And T-Distributions	Page 42
Chapter 20	X² Test	Page 97
Chapter 21	Bivariate Data Simple	Page 131
Chapter 24	Non Parametric Test	Page 151
Chapter 25	Continuous Random Variable	Page 171
Chapter 26	Probability Generating Function	Page 202
	ANSWERS	Page 231

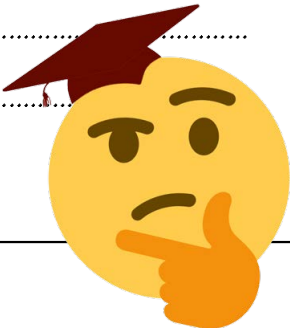


1 - (9231/21_Summer_2017_Q6) - Further Work On Distributions

A fair die is thrown repeatedly until a 6 is obtained.

(i) Find the probability that obtaining a 6 takes no more than four throws. [2]

(ii) Find the least integer N such that the probability of obtaining a 6 before the N th throw is more than 0.95. [3]



2 - (9231/21_Summer_2017_Q8) - Further Work On Distributions

The continuous random variable X has probability density function f given by

$$f(x) = \begin{cases} \frac{1}{4}(x-1) & 2 \leq x \leq 4, \\ 0 & \text{otherwise.} \end{cases}$$

- (i) Find the distribution function of X . [3]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

The random variable Y is defined by $Y = (X - 1)^3$.

- (ii) Find the probability density function of Y . [4]

.....

.....

.....

.....

.....

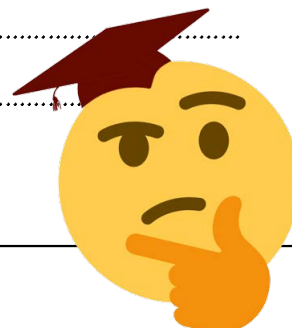
.....

.....

.....

.....

.....



(iii) Find the median value of Y .

[3]



A yellow thinking face emoji with a red graduation cap on its head, pointing to its chin.

The variable Y is related to X by $Y = 2^X$.

(iv) Find the probability density function of Y .

[5]



4 - (9231/21_Winter_2017_Q6) - Further Work On Distributions

A pair of fair dice is thrown repeatedly until a pair of sixes is obtained. The number of throws taken is denoted by the random variable X .

- (i) Find the mean value of X . [2]

.....

.....

.....

.....

.....

.....

.....

- (ii) Find the probability that exactly 12 throws are required to obtain a pair of sixes. [2]

.....

.....

.....

.....

.....

.....

.....

- (iii) Find the probability that more than 12 throws are required to obtain a pair of sixes. [2]

.....

.....

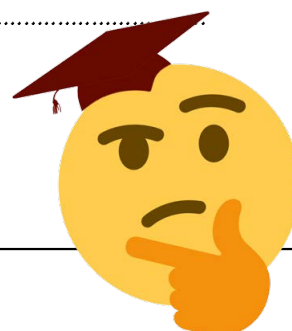
.....

.....

.....

.....

.....



5 - (9231/21_Winter_2017_Q7) - Further Work On Distributions

The random variable X has probability density function f given by

$$f(x) = \begin{cases} 0.2e^{-0.2x} & x \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

- (i) Find the distribution function of X . [2]

.....

.....

.....

.....

.....

.....

- (ii) Find $P(X > 2)$. [2]

.....

.....

.....

.....

.....

- (iii) Find the median of X . [3]

.....

.....

.....

.....

.....



6 - (9231/21_Summer_2018_Q6) - Further Work On Distributions

The continuous random variable X has distribution function given by

$$F(x) = \begin{cases} 1 - e^{-0.4x} & x \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

(i) Find $P(X > 2)$.

[2]

.....

.....

.....

.....

.....

.....

.....

.....

(ii) Find the interquartile range of X .

[4]



7 - (9231/23_Summer_2018_Q7) - Further Work On Distributions

The probability that a driver passes an advanced driving test has a fixed value p for each attempt. A driver keeps taking the test until he passes. The random variable X denotes the number of attempts required for the driver to pass. The variance of X is 3.75.

(i) Show that $15p^2 + 4p - 4 = 0$ and hence find the value of p . [4]

(ii) Find $P(X = 5)$. [1]

.....

.....

.....

.....

.....

(iii) Find $P(3 \leq X \leq 7)$. [2]

.....

.....

.....

.....

.....



8 - (9231/21_Summer_2018_Q9) - *Further Work On Distributions*

At a ski resort, the probability of snow on any particular day is constant and equal to p . The skiing season begins on 1 November. The random variable X denotes the day of the skiing season on which the first snowfall occurs. (For example, if the first snowfall is on 5 November, then $X = 5$.) The variance of X is $\frac{4}{9}$.

- (i) Show that $4p^2 + 9p - 9 = 0$ and hence find the value of p . [4]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

- (ii) Find the probability that the first snowfall will be on 3 November. [1]



- (iii) Find the probability that the first snowfall will not be before 4 November. [2]

.....

.....

.....

.....

.....

.....

.....

- (iv) Find the least integer N so that the probability of the first snowfall being on or before the N th day of November is more than 0.999. [4]



9 - (9231/23_Summer_2018_Q9) - *Further Work On Distributions*

The continuous random variable X has probability density function given by

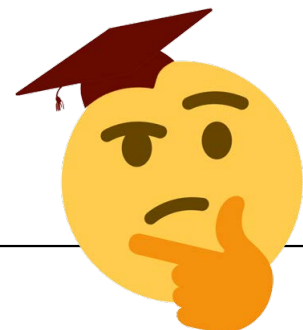
$$f(x) = \begin{cases} \frac{1}{20} \left(3 - \frac{1}{\sqrt{x}} \right) & 1 \leq x \leq 9, \\ 0 & \text{otherwise.} \end{cases}$$

The random variable Y is defined by $Y = \sqrt{X}$.

(i) Show that the probability density function of Y is given by

$$g(y) = \begin{cases} \frac{1}{10}(3y - 1) & 1 \leq y \leq 3, \\ 0 & \text{otherwise.} \end{cases} \quad [7]$$

(ii) Find the mean value of Y . [2]



10 - (9231/21_Winter_2018_Q6) - Further Work On Distributions

The continuous random variable X has probability density function f given by

$$f(x) = \begin{cases} \frac{1}{80} \left(3\sqrt{x} - \frac{8}{\sqrt{x}} \right) & 4 \leq x \leq 16, \\ 0 & \text{otherwise.} \end{cases}$$

(i) Find the distribution function of X .

[3]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

The random variable Y is defined by $Y = \sqrt{X}$.

(ii) Find the probability density function of Y .

[3]



11 - (9231/21_Winter_2018_Q7) - *Further Work On Distributions*

The random variable T is the lifetime, in hours, of a particular type of battery. It is given that T has a negative exponential distribution with mean 500 hours.

- (i) Write down the probability density function of T . [1]

.....

.....

.....

.....

- (ii) Find the probability that a randomly chosen battery of this type has a lifetime of more than 750 hours. [3]

.....

.....

.....

.....

.....

.....

.....

- (iii) Find the median value of T . [3]



12 - (9231/22_Winter_2018_Q7) - Further Work On Distributions

The continuous random variable X has distribution function given by

$$F(x) = \begin{cases} 0 & x < 0, \\ \frac{1}{90}(x^2 + x^4) & 0 \leq x \leq 3, \\ 1 & x > 3. \end{cases}$$

The random variable Y is defined by $Y = X^2$.

- (i) Find the probability density function of Y . [4]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

- (ii) Find the mean value of Y . [2]



13 - (9231/23_Winter_2018_Q7) - *Further Work On Distributions*

The random variable T is the lifetime, in hours, of a particular type of battery. It is given that T has a negative exponential distribution with mean 500 hours.

- (i) Write down the probability density function of T . [1]

.....

.....

.....

.....

- (ii) Find the probability that a randomly chosen battery of this type has a lifetime of more than 750 hours. [3]

.....

.....

.....

.....

.....

.....

.....

- (iii) Find the median value of T . [3]



14 - (9231/22_Winter_2018_Q8) - *Further Work On Distributions*

Lan starts a new job on Monday. He will catch the bus to work every day from Monday to Friday inclusive. The probability that he will get a seat on the bus has the constant value p . The random variable X denotes the number of days that Lan will catch the bus until he is able to get a seat. The probability that Lan will not get a seat on the Monday, Tuesday, Wednesday or Thursday of his first week is 0.4096.

- (i) Show that $p = 0.2$. [2]

.....

.....

.....

.....

.....

.....

.....

.....

.....

- (ii) Find the probability that Lan first gets a seat on Monday of the second week in his new job. [2]

- (iii) Find the least integer N such that $P(X \leq N) > 0.9$, and identify the day and the week that corresponds to this value of N . [4]



15 - (9231/21_Summer_2019_Q6) - Further Work On Distributions

A fair six-sided die is thrown until a 3 or a 4 is obtained. The number of throws taken is denoted by the random variable X .

(i) State the mean value of X . [1]

.....

.....

.....

.....

(ii) Find the probability that obtaining a 3 or a 4 takes exactly 6 throws. [1]

.....

.....

.....

.....

.....

(iii) Find the probability that obtaining a 3 or a 4 takes more than 4 throws. [2]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

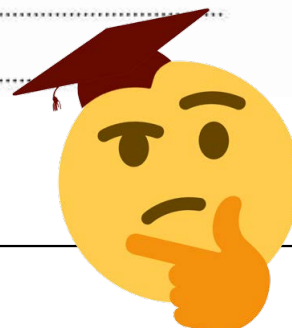
.....

.....

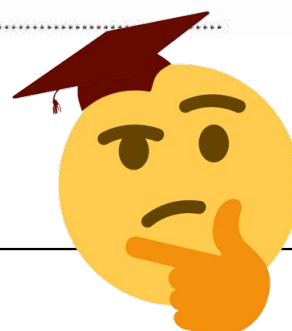
.....

.....

.....



- [illegible]



16 - (9231/23_Summer_2019_Q6) - *Further Work On Distributions*

The random variable T is the lifetime, in hours, of a randomly chosen battery of a particular type. It is given that T has a negative exponential distribution with mean 400 hours.

- (i) Write down the probability density function of T . [1]

.....

.....

.....

.....

- (ii) Find the probability that a battery of this type has a lifetime that is less than 500 hours. [2]

.....

.....

.....

.....

.....

.....

- (iii) Find the median of the distribution. [3]

.....

.....

.....

.....

.....

.....

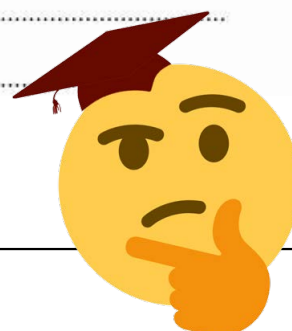
.....

.....

.....

.....

.....



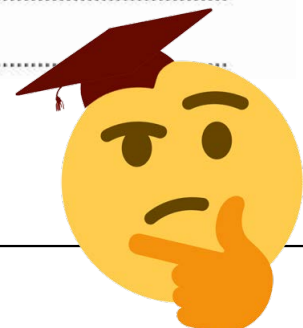
17 - (9231/21_Summer_2019_Q7) - *Further Work On Distributions*

The continuous random variable X has probability density function f given by

$$f(x) = \begin{cases} \frac{3}{4x^2} + \frac{1}{4} & 1 \leq x \leq 3, \\ 0 & \text{otherwise.} \end{cases}$$

(i) Find the distribution function of X .

[3]

[illegible]

[5]

[illegible]

18 - (9231/23_Summer_2019_Q7) - Further Work On Distributions

A pair of fair coins is thrown repeatedly until a pair of tails is obtained. The number of throws taken is denoted by the random variable X .

- (i) State the expected value of X . [1]

.....

.....

- (ii) Find the probability that exactly 3 throws are required to obtain a pair of tails. [2]

.....

.....

.....

.....

- (iii) Find the probability that fewer than 4 throws are required to obtain a pair of tails. [2]

.....

.....

.....

.....

- (iv) Find the least integer N such that the probability of obtaining a pair of tails in fewer than N throws is more than 0.95. [3]

.....

.....

.....

.....

.....

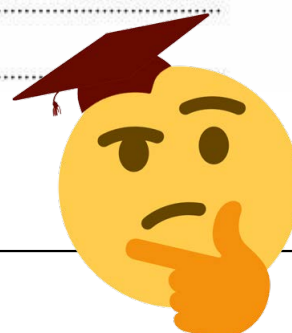
.....

.....

.....

.....

.....



19 - (9231/21_Winter_2019_Q7) - Further Work On Distributions

The time, T days, before an electrical component develops a fault has distribution function F given by

$$F(t) = \begin{cases} 1 - e^{-at} & t \geq 0, \\ 0 & \text{otherwise,} \end{cases}$$

where a is a positive constant. The mean value of T is 200.

- (i) Write down the value of a .

[1]

.....

.....

- (ii) Find the probability that an electrical component of this type develops a fault in less than 150 days.

[2]

.....

.....

.....

.....

.....

A piece of equipment contains n of these components, which develop faults independently of each other. The probability that, after 150 days, at least one of the n components has not developed a fault is greater than 0.99.

- (iii) Find the smallest possible value of n .

[4]

.....

.....

.....

.....

.....

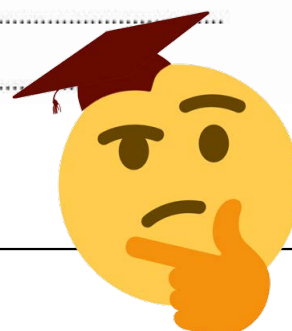
.....

.....

.....

.....

.....



20 - (9231/21_Winter_2019_Q10) - Further Work On Distributions

The random variable X has probability density function f given by

$$f(x) = \begin{cases} \frac{1}{30} \left(\frac{8}{x^2} + 3x^2 - 14 \right) & 2 \leq x \leq 4, \\ 0 & \text{otherwise.} \end{cases}$$

(i) Find the distribution function of X .

[3]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

The random variable Y is defined by $Y = X^2$.

(ii) Find the probability density function of Y .

[4]

.....

.....

.....

.....

.....

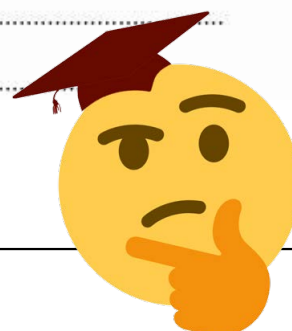
.....

.....

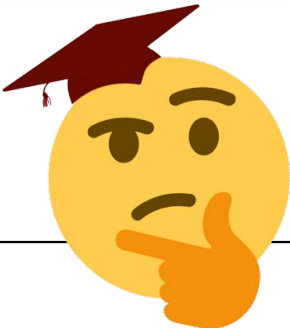
.....

.....

.....



(iii) Find the value of y such that $P(Y < y) = 0.8$. [3]



21 - (9231/31_Summer_2023_Q3) - *Circular Motion, Further Work On Distributions*

A particle P of mass m is attached to one end of a light inextensible string of length a . The other end of the string is attached to a fixed point O . The particle P is held at the point A , where OA makes an angle θ with the downward vertical through O , and with the string taut. The particle P is projected perpendicular to OA in an upwards direction with speed u . It then starts to move along a circular path in a vertical plane. The string goes slack when P is at B , where angle AOB is 90° and the speed of P is $\sqrt{\frac{4}{5}ag}$.

- (a) Find the value of $\sin \theta$. [2]

.....

.....

.....

.....

.....

.....

.....

.....

.....

- (b) Find, in terms of m and g , the tension in the string when P is at A . [5]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

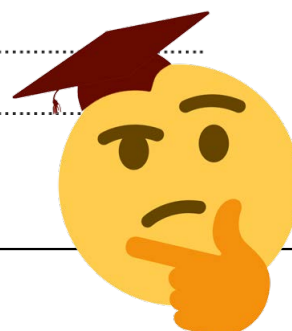
.....

.....

.....

.....

.....



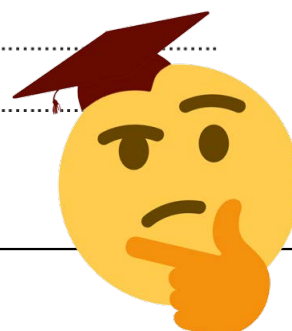
[illegible]

22 - (9231/31_Summer_2023_Q5) - *Circular Motion, Further Work On Distributions*

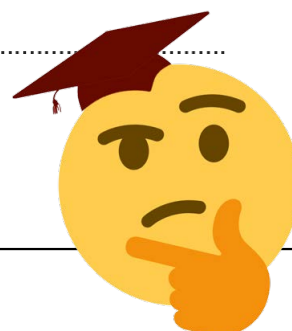
A light elastic string of natural length a and modulus of elasticity λmg has one end attached to a fixed point O on a smooth horizontal surface. When a particle of mass m is attached to the free end of the string, it moves with speed v in a horizontal circle with centre O and radius x . When, instead, a particle of mass $2m$ is attached to the free end of the string, this particle moves with speed $\frac{1}{2}v$ in a horizontal circle with centre O and radius $\frac{3}{4}x$.

(a) Find x in terms of a .

[5]

This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and extend across the width of the page. In the bottom right corner, there is a small, stylized illustration of a yellow notepad or sticky note with a red spiral binding. The overall appearance is that of a clean, unused piece of stationery.

[2]

[illegible]

23 - (9231/31_Summer_2023_Q7) - *Circular Motion, Further Work On Distributions*

At time t s, a particle P is projected with speed 40 m s^{-1} at an angle θ above the horizontal from a point O on a horizontal plane and moves freely under gravity. The greatest height achieved by P during its flight is H m and the corresponding time is T s.

- (a) Obtain expressions for H and T in terms of θ . [2]

.....

.....

.....

.....

.....

.....

During the time between $t = T$ and $t = 3$, P descends a distance $\frac{1}{4}H$.

- (b) Find the value of θ . [4]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

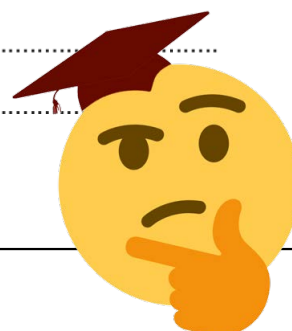
.....

.....

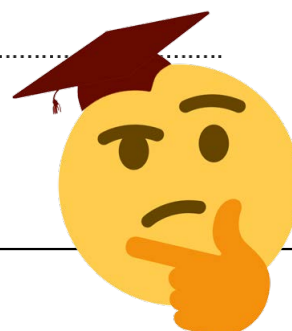
.....

.....

.....



[3]

[illegible]

24 - (9231/33_Summer_2023_Q2) - Circular Motion, Further Work On Distributions

One end of a light elastic string, of natural length a and modulus of elasticity λmg , is attached to a fixed point O . The string lies on a smooth horizontal surface. A particle P of mass m is attached to the other end of the string. The particle P is projected in the direction OP . When the length of the string is $\frac{4}{3}a$, the speed of P is $\sqrt{2ag}$. When the length of the string is $\frac{5}{3}a$, the speed of P is $\frac{1}{2}\sqrt{2ag}$.

Find the value of λ .

[4]



25 - (9231/41_Summer_2023_Q1) - *Further Work On Distributions*

The lengths of the leaves of a particular type of tree are normally distributed with mean μ cm. The lengths, x cm, of a random sample of 12 leaves of this type are recorded. The results are summarised as follows.

$$\sum x = 91.2 \qquad \sum x^2 = 695.8$$

Find a 95% confidence interval for μ .

[4]

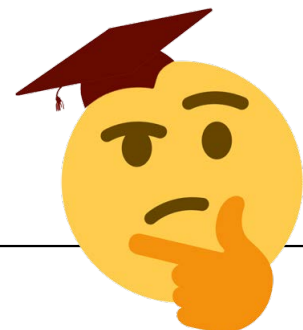


26 - (9231/43_Summer_2023_Q2) - Further Work On Distributions

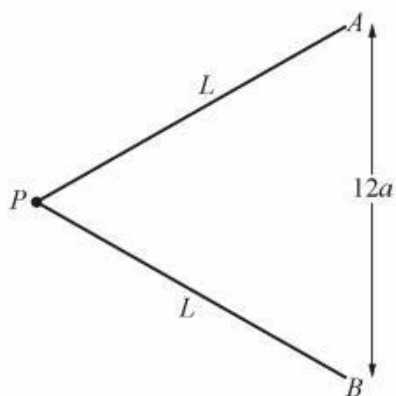
Shane is studying the lengths of the tails of male red kangaroos. He takes a random sample of 14 male red kangaroos and measures the length of the tail, x m, for each kangaroo. He then calculates a 90% confidence interval for the population mean tail length, μ m, of male red kangaroos. He assumes that the tail lengths are normally distributed and finds that $1.11 \leq \mu \leq 1.14$.

Find the values of $\sum x$ and $\sum x^2$ for this sample.

[6]



27 - (9231/31_Winter_2023_Q4) - Further Work On Distributions



A light elastic string has natural length $8a$ and modulus of elasticity $5mg$. A particle P of mass m is attached to the midpoint of the string. The ends of the string are attached to points A and B which are a distance $12a$ apart on a smooth horizontal table. The particle P is held on the table so that $AP = BP = L$ (see diagram). The particle P is released from rest. When P is at the midpoint of AB it has speed $\sqrt{80ag}$.

- (a) Find L in terms of a . [5]
- (b) Find the initial acceleration of P in terms of g . [3]



28 - (9231/33_Winter_2023_Q6) - Further Work On Distributions

A particle P of mass m is attached to one end of a light inextensible rod of length $3a$. An identical particle Q is attached to the other end of the rod. The rod is smoothly pivoted at a point O on the rod, where $OQ = x$. The system, of rod and particles, rotates about O in a vertical plane.

At an instant when the rod is vertical, with P above Q , the particle P is moving horizontally with speed u . When the rod has turned through an angle of 60° from the vertical, the speed of P is $2\sqrt{ag}$, and the tensions in the two parts of the rod, OP and OQ , have equal magnitudes.

- (a) Show that the speed of Q when the rod has turned through an angle of 60° from the vertical is $\frac{2x}{3a-x}\sqrt{ag}$. [2]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

- (b) Find x in terms of a . [5]

.....

.....

.....

.....

.....

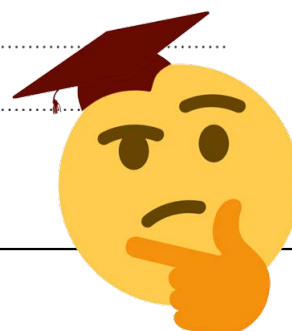
.....

.....

.....

.....

.....



.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

(c) Find u in terms of a and g . [4]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

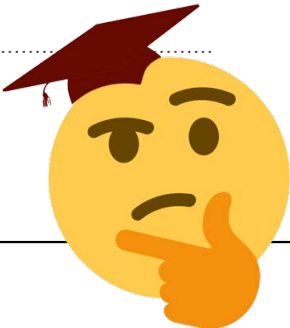
.....

.....

.....

.....

.....



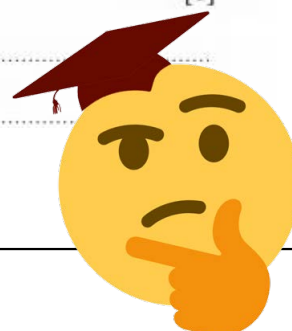
29 - (9231/41_Winter_2023_Q3) - *Further Work On Distributions*

Scientists are studying the effects of exercise on LDL blood cholesterol levels. Over a three-month period, a large group of people exercised for 20 minutes each day. For a randomly chosen sample of 10 of these people, the LDL blood cholesterol levels were measured at the beginning and the end of the three-month period. The results, measured in suitable units, are as follows.

	Person	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>	<i>G</i>	<i>H</i>	<i>I</i>	<i>J</i>
Cholesterol level	Beginning	72	84	120	90	102	135	64	75	80	88
	End	64	76	105	92	105	115	67	75	75	84

- (a) Test, at the 2.5% significance level, whether there is evidence that the population mean LDL blood cholesterol level has reduced by more than 2 units after the three-month period. [7]

- (b) State any assumption that you have made in part (a). [1]



30 - (9231/42_Winter_2023_Q1) - *Further Work On Distributions*

A factory produces small bottles of natural spring water. Two different machines, X and Y , are used to fill empty bottles with the water. A quality control engineer checks the volumes of water in the bottles filled by each of the machines. He chooses a random sample of 60 bottles filled by machine X and a random sample of 75 bottles filled by machine Y . The volumes of water, x and y respectively, in millilitres, are summarised as follows.

$$\sum x = 6345 \quad \sum (x - \bar{x})^2 = 243.8 \quad \sum y = 7614 \quad \sum (y - \bar{y})^2 = 384.9$$

$\bar{x}$ and $\bar{y}$ are the sample means of the volume of water in the bottles filled by machines X and Y respectively.

Find a 95% confidence interval for the difference between the mean volume of water in bottles filled by machine X and the mean volume of water in bottles filled by machine Y . [6]



31 - (9231/43_Winter_2023_Q1) - *Further Work On Distributions*

Maya is an athlete who competes in 1500-metre races. Last summer her practice run times had mean 4.22 minutes. Over the winter she has done some intense training to try to improve her times. A random sample of 10 of her practice run times, x minutes, this summer are summarised as follows.

$$\sum x = 42.05 \quad \sum x^2 = 176.83$$

Maya's new practice run times are normally distributed. She believes that on average her times have improved as a result of her training.

Test, at the 5% significance level, whether Maya's belief is supported by the data. [6]

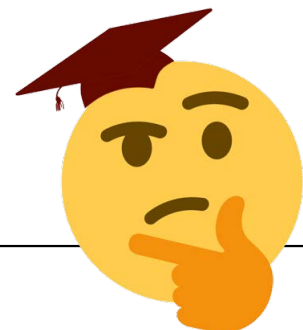


1 - (9231/23_Summer_2017_Q6) - Inference Using Normal And T-distributions

The independent variables X and Y have distributions with the same variance σ^2 . Random samples of N observations of X and $2N$ observations of Y are taken, and the results are summarised by

$$\Sigma x = 4, \quad \Sigma x^2 = 10, \quad \Sigma y = 8, \quad \Sigma y^2 = 102.$$

These data give a pooled estimate of 10 for σ^2 . Find N . [5]



2 - (9231/21_Summer_2017_Q7) - Inference Using Normal And T-distributions

A farmer grows a particular type of fruit tree. On average, the mass of fruit produced per tree has been 6.2 kg. He has developed a new kind of soil and claims that the mean mass of fruit produced per tree when growing in this new soil has increased. A random sample of 10 trees grown in the new soil is chosen. The masses, x kg, of fruit produced are summarised as follows.

$$\Sigma x = 72.0 \quad \Sigma x^2 = 542.0$$

Test at the 5% significance level whether the farmer's claim is justified, assuming a normal distribution. [7]

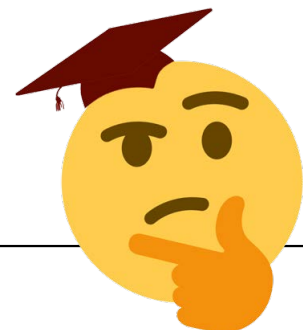


3 - (9231/23_Summer_2017_Q8) - *Inference Using Normal And T-distributions*

The number, x , of beech trees was counted in each of 50 randomly chosen regions of equal size in beech forests in country A . The number, y , of beech trees was counted in each of 40 randomly chosen regions of the same equal size in beech forests in country B . The results are summarised as follows.

$$\Sigma x = 1416 \quad \Sigma x^2 = 41\,100 \quad \Sigma y = 888 \quad \Sigma y^2 = 20\,140$$

Find a 95% confidence interval for the difference between the mean number of beech trees in regions of this size in country A and in country B . [9]



4 - (9231/21_Summer_2017_Q9) - *Inference Using Normal And T-distributions*

Two fish farmers X and Y produce a particular type of fish. Farmer X chooses a random sample of 8 of his fish and records the masses, x kg, as follows.

1.2 1.4 0.8 2.1 1.8 2.6 1.5 2.0

Farmer Y chooses a random sample of 10 of his fish and summarises the masses, y kg, as follows.

$$\Sigma y = 20.2 \quad \Sigma y^2 = 44.6$$

You should assume that both distributions are normal with equal variances. Test at the 10% significance level whether the mean mass of fish produced by farmer X differs from the mean mass of fish produced by farmer Y . [10]



5 - (9231/21_Winter_2017_Q10) - Inference Using Normal And T-distributions

A factory produces bottles of an energy juice. Two different machines are used to fill empty bottles with the juice. The manager chooses a random sample of 50 bottles filled by machine X and a random sample of 60 bottles filled by machine Y . The volumes of juice, x and y respectively, measured in appropriate units, are summarised by

$$\Sigma x = 45.5, \quad \Sigma(x - \bar{x})^2 = 19.56, \quad \Sigma y = 72.3, \quad \Sigma(y - \bar{y})^2 = 30.25,$$

where $\bar{x}$ and $\bar{y}$ are the sample means of the volume of juice in the bottles filled by X and Y respectively.

- (i) Find a 90% confidence interval for the difference between the mean volume of juice in bottles filled by machine X and the mean volume of juice in bottles filled by machine Y . [7]

A test at the $\alpha\%$ significance level does not provide evidence that there is any difference in the means of the volume of juice in bottles filled by machine X and the volume of juice in bottles filled by machine Y .

- (ii) Find the set of possible values of α . [6]

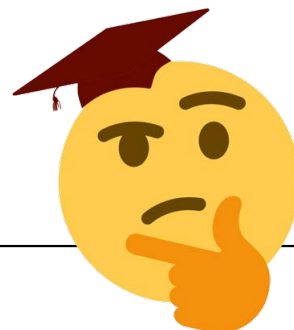


6 - (9231/21_Summer_2018_Q7) - *Inference Using Normal And T-distributions*

A large number of athletes are taking part in a competition. The masses, in kg, of a random sample of 7 athletes are as follows.

98.1 105.0 92.2 89.8 99.9 95.4 101.2

Assuming that masses are normally distributed, test, at the 10% significance level, whether the mean mass of athletes in this competition is equal to 94 kg. [7]



7 - (9231/21_Summer_2018_Q10) - Inference Using Normal And T-distributions

The times taken to run 400 metres by students at two large colleges P and Q are being compared. There is no evidence that the population variances are equal. The time taken by a student at college P and the time taken by a student at college Q are denoted by x seconds and y seconds respectively. A random sample of 50 students from college P and a random sample of 60 students from college Q give the following summarised data.

$$\Sigma x = 2620 \quad \Sigma x^2 = 138\,200 \quad \Sigma y = 3060 \quad \Sigma y^2 = 157\,000$$

- (i) Using a 10% significance level, test whether, on average, students from college P take longer to run 400 metres than students from college Q . [9]
- (ii) Find a 90% confidence interval for the difference in the mean times taken to run 400 metres by students from colleges P and Q . [3]



8 - (9231/23_Summer_2018_Q10) - *Inference Using Normal And T-distributions*

During the summer months, all members of a large swimming club take part in intensive training. The times taken to swim 50 metres at the beginning of the summer and at the end of the summer are recorded for each member of the club. The time taken, in seconds, at the beginning of the summer is denoted by x and the time taken at the end of the summer is denoted by y . For a random sample of 9 members the results are shown in the following table.

Member	A	B	C	D	E	F	G	H	I
x	38.5	40.2	32.3	35.1	36.2	41.4	32.0	38.2	38.2
y	37.4	38.1	31.6	34.7	34.2	38.6	31.8	36.3	36.8

The swimming coach believes that, on average, the time taken by a swimmer to swim 50 metres will decrease by more than one second as a result of the intensive training.

- (i) Stating suitable hypotheses and assuming a normal distribution, test the coach's belief at the 10% significance level. [8]
- (ii) Find a 95% confidence interval for the population mean time taken to swim 50 metres after the intensive training, assuming a normal distribution. [4]



9 - (9231/22_Winter_2018_Q6) - *Inference Using Normal And T-distributions*

The heights, in metres, of a random sample of 8 trees of a particular type are as follows.

14.2 11.3 10.8 8.4 12.8 11.5 12.1 9.2

Assuming that heights of trees of this type are normally distributed, calculate a 95% confidence interval for the mean height of trees of this type. [6]



10 - (9231/21_Winter_2018_Q8) - Inference Using Normal And T-distributions

The weekly salaries of employees at two large electronics companies, A and B , are being compared. The weekly salaries of an employee from company A and an employee from company B are denoted by $\$x$ and $\$y$ respectively. A random sample of 50 employees from company A and a random sample of 40 employees from company B give the following summarised data.

$$\Sigma x = 5120 \quad \Sigma x^2 = 531\,000 \quad \Sigma y = 3760 \quad \Sigma y^2 = 375\,135$$

- (i) The population mean salaries of employees from companies A and B are denoted by $\$ \mu_A$ and $\$ \mu_B$ respectively. Using a 5% significance level, test the null hypothesis $\mu_A = \mu_B$ against the alternative hypothesis $\mu_A \neq \mu_B$. [8]
- (ii) State, with a reason, whether any assumptions about the distributions of employees' salaries are needed for the test in part (i). [1]



11 - (9231/21_Winter_2018_Q9) - *Inference Using Normal And T-distributions*

There are a large number of students at a particular college. The heights, in metres, of a random sample of 8 students are as follows.

1.75 1.72 1.62 1.70 1.82 1.75 1.68 1.84

You may assume that heights of students are normally distributed.

- (i) Test, at the 5% significance level, whether the population mean height of students at this college is greater than 1.70 metres. [7]
- (ii) Find a 95% confidence interval for the population mean height of students at this college. [3]



12 - (9231/22_Winter_2018_Q11) - Simple Harmonic Motion, Inference Using Normal And T-distributions

Answer only **one** of the following two alternatives.

EITHER

One end of a light elastic spring, of natural length 0.8 m and modulus of elasticity 40 N, is attached to a fixed point O . The spring hangs vertically, at rest, with particles of masses 2 kg and M kg attached to its free end. The M kg particle becomes detached from the spring, and as a result the 2 kg particle begins to move upwards.

- (i) Show that the 2 kg particle performs simple harmonic motion about its equilibrium position with period $\frac{2}{5}\pi$ s. State the distance below O of the centre of the oscillations. [7]

The speed of the 2 kg particle is 0.4 m s^{-1} when its displacement from the centre of oscillation is 0.06 m.

- (ii) Find the amplitude of the motion. [3]

.....

.....

.....

.....

.....

.....

.....

.....

- (iii) Deduce the value of M . [4]

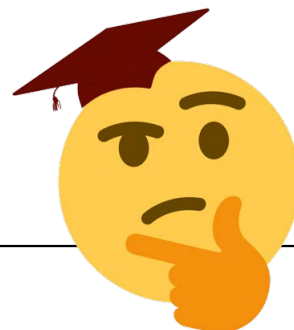
OR

In a particular country, large numbers of ducks live on lakes A and B . The mass, in kg, of a duck on lake A is denoted by x and the mass, in kg, of a duck on lake B is denoted by y . A random sample of 8 ducks is taken from lake A and a random sample of 10 ducks is taken from lake B . Their masses are summarised as follows.

$$\Sigma x = 10.56 \quad \Sigma x^2 = 14.1775 \quad \Sigma y = 12.39 \quad \Sigma y^2 = 15.894$$

A scientist claims that ducks on lake A are heavier on average than ducks on lake B .

- (i) Test, at the 10% significance level, whether the scientist's claim is justified. You should assume that both distributions are normal and that their variances are equal. [9]



A second random sample of 8 ducks is taken from lake A and their masses are summarised as

$$\Sigma x = 10.24 \quad \text{and} \quad \Sigma (x - \bar{x})^2 = 0.294,$$

where $\bar{x}$ is the sample mean. The scientist now claims that the population mean mass of ducks on lake A is greater than p kg. A test of this claim is carried out at the 10% significance level, using only this second sample from lake A . This test supports the scientist's claim.

(ii) Find the greatest possible value of p .

[5]

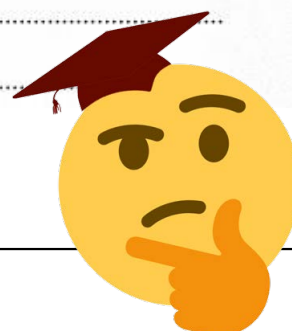


13 - (9231/21_Summer_2019_Q8) - *Inference Using Normal And T-distributions*

A large number of runners are attending a summer training camp. A random sample of 6 runners is chosen and their times to run 1500 m at the beginning of the camp and at the end of the camp are recorded. Their times, in minutes, are shown in the following table.

Runner	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>
Time at beginning of camp	3.82	3.62	3.55	3.71	3.75	3.92
Time at end of camp	3.72	3.55	3.52	3.68	3.54	3.73

The organiser of the training camp claims that a runner's time will improve by more than 0.05 minutes between the beginning and end of the camp. Assuming that differences in time over the two runs are normally distributed, test at the 10% significance level whether the organiser's claim is justified. [8]

[illegible]

This image shows a full page of a handwriting practice worksheet. It consists of multiple rows of horizontal dashed lines spaced evenly down the page, providing a guide for letter height and placement. The background is plain white, and there are no margins or additional markings.

14 - (9231/23_Summer_2019_Q9) - *Inference Using Normal And T-distributions*

A farmer grows large amounts of a certain crop. On average, the yield per plant has been 0.75 kg. The farmer has improved the soil in which the crop grows, and she claims that the yield per plant has increased. A random sample of 10 plants grown in the improved soil is chosen. The yields, x kg, are summarised as follows.

$$\Sigma x = 7.85 \quad \Sigma x^2 = 6.19$$

- (i) Test at the 5% significance level whether the farmer's claim is justified, assuming a normal distribution. [7]
- (ii) Find a 95% confidence interval for the population mean yield for plants grown in the new soil. [3]



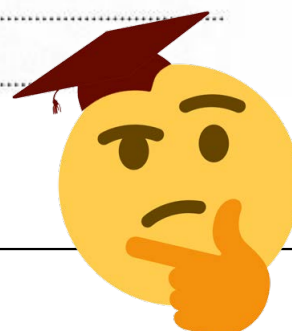
15 - (9231/21_Summer_2019_Q11) - *Momentum And Impulse, Circular Motion, Inference Using Normal And T-distributions*

Answer only **one** of the following two alternatives.

EITHER

A particle P , of mass m , is able to move in a vertical circle on the smooth inner surface of a sphere with centre O and radius a . Points A and B are on the inner surface of the sphere and AOB is a horizontal diameter. Initially, P is projected vertically downwards with speed $\sqrt{\frac{21}{2}ag}$ from A and begins to move in a vertical circle. At the lowest point of its path, vertically below O , the particle P collides with a stationary particle Q , of mass $4m$, and rebounds. The speed acquired by Q , as a result of the collision, is just sufficient for it to reach the point B .

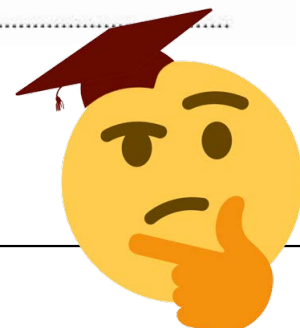
- (i) Find the speed of P and the speed of Q immediately after their collision. [7]

[illegible]

In its subsequent motion, P loses contact with the inner surface of the sphere at the point D , where the angle between OD and the upward vertical through O is θ .

(ii) Find $\cos \theta$.

[5]

This image shows a full page of a notebook or worksheet. It features approximately 20 horizontal dashed lines spaced evenly down the page, providing a guide for handwriting practice. The lines are light gray and extend across the entire width of the page. There is no text or other markings on the page.

OR

A farmer grows two different types of cherries, Type *A* and Type *B*. He assumes that the masses of each type are normally distributed. He chooses a random sample of 8 cherries of Type *A*. He finds that the sample mean mass is 15.1 g and that a 95% confidence interval for the population mean mass, μ g, is $13.5 \leq \mu \leq 16.7$.

- (i) Find an unbiased estimate for the population variance of the masses of cherries of Type *A*. [3]

.....

.....

.....

.....

.....

.....

.....

.....

.....

The farmer now chooses a random sample of 6 cherries of Type *B* and records their masses as follows.

12.2 13.3 16.4 14.0 13.9 15.4

- (ii) Test at the 5% significance level whether the mean mass of cherries of Type *B* is less than the mean mass of cherries of Type *A*. You should assume that the population variances for the two types of cherry are equal. [9]

.....

.....

.....

.....

.....

.....

.....

.....



[illegible]

16 - (9231/23_Summer_2019_Q11) - *Simple Harmonic Motion, Inference Using Normal And T-distributions*

Answer only **one** of the following two alternatives.

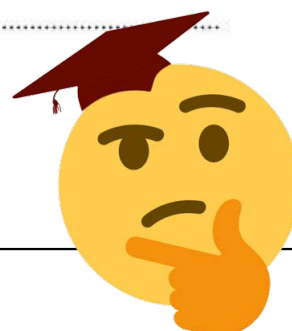
EITHER

A light spring has natural length a and modulus of elasticity kmg . The spring lies on a smooth horizontal surface with one end attached to a fixed point O . A particle P of mass m is attached to the other end of the spring. The system is in equilibrium with $OP = a$. The particle is projected towards O with speed u and comes to instantaneous rest when $OP = \frac{3}{4}a$.

- (i) Use an energy method to show that $k = \frac{16u^2}{ag}$. [2]

- (ii) Show that P performs simple harmonic motion and find the period of this motion, giving your answer in terms of u and a . [4]

[illegible]

[illegible]

[illegible]

17 - (9231/21_Winter_2019_Q6) - Inference Using Normal And T-distributions

A random sample of 9 members is taken from the large number of members of a sports club, and their heights are measured. The heights of all the members of the club are assumed to be normally distributed. A 95% confidence interval for the population mean height, μ metres, is calculated from the data as $1.65 \leq \mu \leq 1.85$.

- (i) Find an unbiased estimate for the population variance. [3]

.....

.....

.....

.....

.....

.....

.....

.....

- (ii) Denoting the height of a member of the club by x metres, find Σx^2 for this sample of 9 members. [4]

.....

.....

.....

.....

.....

.....

.....

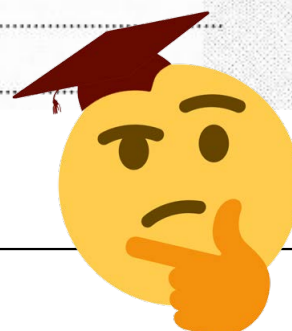
.....

.....

.....

.....

.....

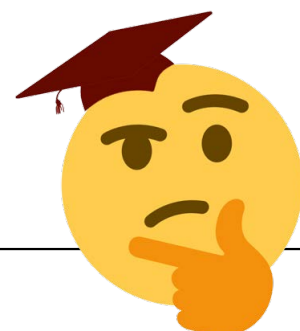


18 - (9231/21_Winter_2019_Q8) - Inference Using Normal And T-distributions

A random sample of 8 elephants from region A is taken and their weights, x tonnes, are recorded. (1 tonne = 1000 kg.) The results are summarised as follows.

$$\Sigma x = 32.4 \quad \Sigma x^2 = 131.82$$

A random sample of 10 elephants from region B is taken. Their weights give a sample mean of 3.78 tonnes and an unbiased variance estimate of 0.1555 tonnes². The distributions of the weights of elephants in regions A and B are both assumed to be normal with the same population variance. Test at the 10% significance level whether the mean weight of elephants in region A is the same as the mean weight of elephants in region B . [9]

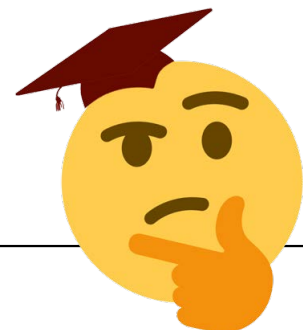


19 - (9231/43_Summer_2020_Q2) - *Inference Using Normal And T-distributions*

A random sample of 40 observations of a random variable X and a random sample of 50 observations of a random variable Y are taken. The resulting values for the sample means, $\bar{x}$ and $\bar{y}$, and the unbiased estimates, s_x^2 and s_y^2 , for the population variances are as follows.

$$\bar{x} = 24.4 \quad \bar{y} = 17.2 \quad s_x^2 = 10.2 \quad s_y^2 = 11.1$$

Find a 90% confidence interval for the difference between the population means of X and Y . [5]



20 - (9231/41_Summer_2020_Q4) - Inference Using Normal And T-distributions

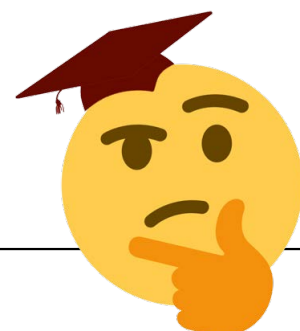
A company has two different machines, X and Y , each of which fills empty cups with coffee. The manager is investigating the volumes of coffee, x and y , measured in appropriate units, in the cups filled by machines X and Y respectively. She chooses a random sample of 50 cups filled by machine X and a random sample of 40 cups filled by machine Y . The volumes are summarised as follows.

$$\sum x = 15.2 \quad \sum x^2 = 5.1 \quad \sum y = 13.4 \quad \sum y^2 = 4.8$$

The manager claims that there is no difference between the mean volume of coffee in cups filled by machine X and the mean volume of coffee in cups filled by machine Y .

Test the manager's claim at the 10% significance level.

[9]

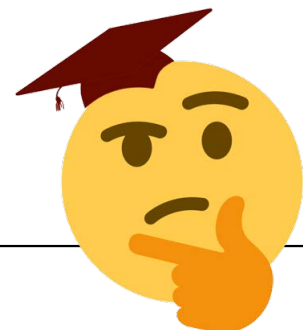


21 - (9231/41_Summer_2020_Q5) - *Inference Using Normal And T-distributions*

A large number of children are competing in a throwing competition. The distances, in metres, thrown by a random sample of 8 children are as follows.

19.8 22.1 24.4 21.5 20.8 26.3 23.7 25.0

- (a) Assuming that distances are normally distributed, test, at the 5% significance level, whether the population mean distance thrown is more than 22.0 metres. [7]
- (b) Find a 95% confidence interval for the population mean distance thrown. [3]



22 - (9231/43_Summer_2020_Q5) - Inference Using Normal And T-distributions

Students at two colleges, A and B , are competing in a computer games challenge.

- (a) The time taken for a randomly chosen student from college A to complete the challenge has a normal distribution with mean μ minutes. The times taken, x minutes, are recorded for a random sample of 10 students chosen from college A . The results are summarised as follows.

$$\sum x = 828 \quad \sum x^2 = 68622$$

A test is carried out on the data at the 5% significance level and the result supports the claim that $\mu > k$.

Find the greatest possible value of k . [4]

- (b) A random sample of 8 students is chosen from college B . Their times to complete the same challenge give a sample mean of 79.8 minutes and an unbiased variance estimate of 9.966 minutes².

Use a 2-sample test at the 5% significance level to test whether the mean time for students at college B to complete the challenge is the same as the mean time for students at college A to complete the challenge. You should assume that the two distributions are normal and have the same population variance. [7]



23 - (9231/41_Winter_2020_Q1) - Inference Using Normal And T-distributions

Kayla is investigating the lengths of the leaves of a certain type of tree found in two forests X and Y . She chooses a random sample of 40 leaves of this type from forest X and records their lengths, x cm. She also records the lengths, y cm, for a random sample of 60 leaves of this type from forest Y . Her results are summarised as follows.

$$\Sigma x = 242.0 \quad \Sigma x^2 = 1587.0 \quad \Sigma y = 373.2 \quad \Sigma y^2 = 2532.6$$

Find a 90% confidence interval for the difference between the population mean lengths of leaves in forests X and Y . [7]



24 - (9231/42_Winter_2020_Q1) - *Inference Using Normal And T-distributions*

The heights of the members of a large sports club are normally distributed. A random sample of 11 members of the club is chosen and their heights, x cm, are measured. The results are summarised as follows, where $\bar{x}$ denotes the sample mean of x .

$$\bar{x} = 176.2 \quad \Sigma(x - \bar{x})^2 = 313.1$$

Test, at the 5% significance level, the null hypothesis that the population mean height for members of this club is equal to 172.5 cm against the alternative hypothesis that the mean differs from 172.5 cm. [5]



25 - (9231/41_Winter_2020_Q4) - Inference Using Normal And T-distributions

Members of the Sprints athletics club have been taking part in an intense training scheme, aimed at reducing their times taken to run 400 m. For a random sample of 9 athletes from the club, the times taken, in seconds, before and after the training scheme are given in the following table.

Athlete	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>	<i>G</i>	<i>H</i>	<i>I</i>
Time before	48.8	48.2	50.3	49.6	49.4	48.9	47.6	50.3	48.4
Time after	47.9	47.8	49.6	49.1	49.6	48.9	47.7	49.1	48.1

The organiser of the training scheme claims that on average an athlete's time will be reduced by at least 0.3 seconds.

Test at the 10% significance level whether the organiser's claim is justified, stating any assumption that you make. [8]



26 - (9231/42_Winter_2020_Q6) - Inference Using Normal And T-distributions

Nassa is researching the lengths of a particular type of snake in two countries, A and B .

- (a) He takes a random sample of 10 snakes of this type from country A and measures the length, x m, of each snake. He then calculates a 90% confidence interval for the population mean length, μ m, for snakes of this type, assuming that snake lengths have a normal distribution. This confidence interval is $3.36 \leq \mu \leq 4.22$.

Find the sample mean and an unbiased estimate for the population variance. [4]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

- (b) Nassa also measures the lengths, y m, of a random sample of 8 snakes of this type taken from country B . His results are summarised as follows.

$$\Sigma y = 27.86 \quad \Sigma y^2 = 98.02$$

Nassa claims that the mean length of snakes of this type in country B is less than the mean length of snakes of this type in country A . Nassa assumes that his sample from country B also comes from a normal distribution, with the same variance as the distribution from country A .

Test at the 10% significance level whether there is evidence to support Nassa's claim. [8]

.....

.....

.....

.....

.....

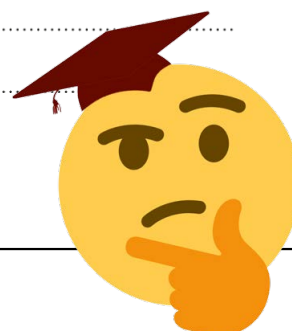
.....

.....

.....

.....

.....



27 - (9231/41_Summer_2021_Q1) - *Inference Using Normal And T-distributions*

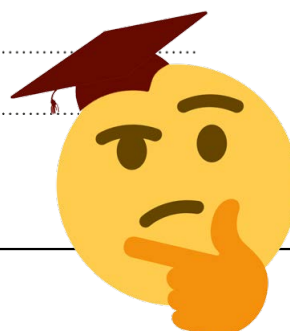
A random sample of 7 observations of a variable X are as follows.

8.26 7.78 7.92 8.04 8.27 7.95 8.34

The population mean of X is μ .

- (a) Test, at the 10% significance level, the null hypothesis $\mu = 8.22$ against the alternative hypothesis $\mu < 8.22$. [6]

- (b) State an assumption necessary for the test in part (a) to be valid. [1]

[illegible]

28 - (9231/43_Summer_2021_Q1) - Inference Using Normal And T-distributions

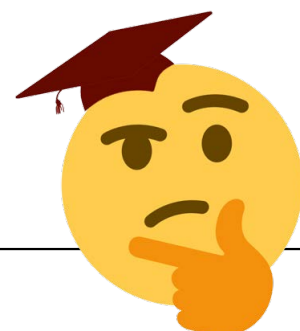
Farmer A grows apples of a certain variety. Each tree produces 14.8 kg of apples, on average, per year. Farmer B grows apples of the same variety and claims that his apple trees produce a higher mass of apples per year than Farmer A 's trees. The masses of apples from Farmer B 's trees may be assumed to be normally distributed.

A random sample of 10 trees from Farmer B is chosen. The masses, x kg, of apples produced in a year are summarised as follows.

$$\sum x = 152.0 \quad \sum x^2 = 2313.0$$

Test, at the 5% significance level, whether Farmer B 's claim is justified.

[6]



29 - (9231/43_Summer_2021_Q3) - Inference Using Normal And T-distributions

The heights, x m, of a random sample of 50 adult males from country A were recorded. The heights, y m, of a random sample of 40 adult males from country B were also recorded. The results are summarised as follows.

$$\sum x = 89.0 \quad \sum x^2 = 159.4 \quad \sum y = 67.2 \quad \sum y^2 = 113.1$$

Find a 95% confidence interval for the difference between the mean heights of adult males from country A and adult males from country B . [8]



30 - (9231/41_Summer_2021_Q4) - *Inference Using Normal And T-distributions*

A scientist is investigating the lengths of the leaves of birch trees in different regions. He takes a random sample of 50 leaves from birch trees in region A and a random sample of 60 leaves from birch trees in region B . He records their lengths in cm, x and y , respectively. His results are summarised as follows.

$$\Sigma x = 282 \quad \Sigma x^2 = 1596 \quad \Sigma y = 328 \quad \Sigma y^2 = 1808$$

The population mean lengths of leaves from birch trees in regions A and B are μ_A cm and μ_B cm respectively.

Carry out a test at the 5% significance level to test the null hypothesis $\mu_A = \mu_B$ against the alternative hypothesis $\mu_A \neq \mu_B$. [8]



31 - (9231/41_Winter_2021_Q1) - Inference Using Normal And T-distributions

The times taken for students at a college to run 200 m have a normal distribution with mean μ s. The times, x s, are recorded for a random sample of 10 students from the college. The results are summarised as follows, where $\bar{x}$ is the sample mean.

$\bar{x} = 25.6$

$\sum(x - \bar{x})^2 = 78.5$

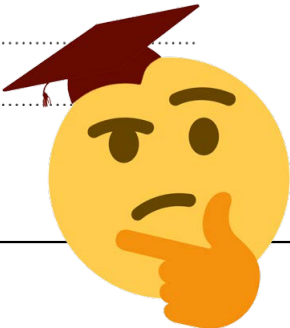
(a) Find a 90% confidence interval for μ .

[4]

A test of the null hypothesis $\mu = k$ is carried out on this sample, using a 10% significance level. The test does not support the alternative hypothesis $\mu < k$.

(b) Find the greatest possible value of k .

[3]



32 - (9231/42_Winter_2021_Q1) - *Inference Using Normal And T-distributions*

The number, x , of pine trees was counted in each of 40 randomly chosen regions of equal size in country A . The number, y , of pine trees was counted in each of 60 randomly chosen regions of the same equal size in country B . The results are summarised as follows.

$$\Sigma x = 752 \quad \Sigma x^2 = 14320 \quad \Sigma y = 1548 \quad \Sigma y^2 = 40200$$

Find a 95% confidence interval for the difference between the mean number of pine trees in regions of this size in countries A and B . [7]



33 - (9231/41_Winter_2021_Q4) - Inference Using Normal And T-distributions

Manet has developed a new training course to help athletes improve their time taken to run 800 m. Manet claims that his course will decrease an athlete's time by more than 2 s on average. For a random sample of 10 athletes the times taken, in seconds, before and after the course are given in the following table.

Athlete	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>	<i>G</i>	<i>H</i>	<i>I</i>	<i>J</i>
Before	150	146	131	135	126	142	130	129	137	134
After	145	138	129	135	122	135	132	128	127	137

Use a *t*-test, at the 5% significance level, to test whether Manet's claim is justified, stating any assumption that you make. [8]



34 - (9231/42_Winter_2021_Q6) - *Inference Using Normal And T-distributions*

A scientist is investigating the masses of a particular type of fish found in lakes A and B . He chooses a random sample of 10 fish of this type from lake A and records their masses, x kg, as follows.

2.1 1.8 0.9 3.0 2.4 2.6 1.8 2.2 1.9 2.5

The scientist also chooses a random sample of 12 fish of this type from lake B , but he only has a summary of their masses, y kg, as follows.

$$\Sigma y = 24.48 \quad \Sigma y^2 = 53.75$$

Test at the 10% significance level whether the mean mass of fish of this type in lake A is greater than the mean mass of fish of this type in lake B . You should state any assumptions that you need to make for the test to be valid. [10]



35 - (9231/41_Summer_2022_Q1) - Inference Using Normal And T-distributions

A manager is investigating the times taken by employees to complete a particular task as a result of the introduction of new technology. He claims that the mean time taken to complete the task is reduced by more than 0.4 minutes. He chooses a random sample of 10 employees. The times taken, in minutes, before and after the introduction of the new technology are recorded in the table.

Employee	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>	<i>G</i>	<i>H</i>	<i>I</i>	<i>J</i>
Time before new technology	10.2	9.8	12.4	11.6	10.8	11.2	14.6	10.6	12.3	11.0
Time after new technology	9.6	8.5	12.4	10.9	10.2	10.6	12.8	10.8	12.5	10.6

- (a) Test at the 10% significance level whether the manager's claim is justified. [7]
- (b) State an assumption that is necessary for this test to be valid. [1]



36 - (9231/43_Summer_2022_Q1) - *Inference Using Normal And T-distributions*

The times taken by members of a large quiz club to complete a challenge have a normal distribution with mean μ minutes. The times, x minutes, are recorded for a random sample of 8 members of the club. The results are summarised as follows, where $\bar{x}$ is the sample mean.

$$\bar{x} = 33.8 \quad \sum(x - \bar{x})^2 = 94.5$$

Find a 95% confidence interval for μ .

[4]



37 - (9231/41_Summer_2022_Q5) - *Inference Using Normal And T-distributions*

Raman is researching the heights of male giraffes in a particular region. Raman assumes that the heights of male giraffes in this region are normally distributed. He takes a random sample of 8 male giraffes from the region and measures the height, in metres, of each giraffe. These heights are as follows.

5.2 5.8 4.9 6.1 5.5 5.9 5.4 5.6

- (a) Find a 90% confidence interval for the population mean height of male giraffes in this region. [5]

Raman claims that the population mean height of male giraffes in the region is less than 5.9 metres.

- (b) Test at the 2.5% significance level whether this sample provides sufficient evidence to support Raman's claim. [4]



38 - (9231/43_Summer_2022_Q6) - *Inference Using Normal And T-distributions*

A company has two machines, A and B , which independently fill small bottles with a liquid. The volumes of liquid per bottle, in suitable units, filled by machines A and B are denoted by x and y respectively. A scientist at the company takes a random sample of 40 bottles filled by machine A and a random sample of 50 bottles filled by machine B . The results are summarised as follows.

$$\Sigma x = 1120 \quad \Sigma x^2 = 31400 \quad \Sigma y = 1370 \quad \Sigma y^2 = 37600$$

The population means of the volumes of liquid in the bottles filled by machines A and B are denoted by μ_A and μ_B .

- (a) Test at the 2% significance level whether there is any difference between μ_A and μ_B . [8]
- (b) Find the set of values of α for which there would be evidence at the $\alpha\%$ significance level that $\mu_A - \mu_B$ is greater than 0.25. [4]



39 - (9231/41_Winter_2022_Q1) - Inference Using Normal And T-distributions

Jasmine is researching the heights of pine trees in forests in two regions A and B . She chooses a random sample of 50 pine trees in region A and records their heights, x m. She also chooses a random sample of 60 pine trees in region B and records their heights, y m. Her results are summarised as follows.

$$\sum x = 1625 \quad \sum x^2 = 53200 \quad \sum y = 1854 \quad \sum y^2 = 57900$$

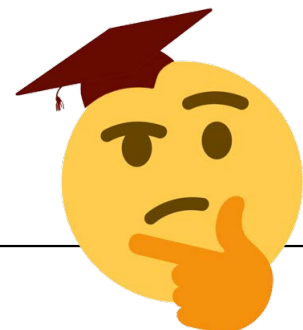
Find a 95% confidence interval for the difference between the population mean heights of pine trees in regions A and B . [7]



40 - (9231/42_Winter_2022_Q1) - Inference Using Normal And T-distributions

A basketball club has a large number of players. The heights, x m, of a random sample of 10 of these players are measured. A 90% confidence interval for the population mean height, μ m, of players in this club is calculated. It is assumed that heights are normally distributed. The confidence interval is $1.78 \leq \mu \leq 2.02$.

Find the values of $\sum x$ and $\sum x^2$ for this sample. [6]



41 - (9231/42_Winter_2022_Q3) - *Inference Using Normal And T-distributions*

A scientist is investigating the masses of birds of a certain species in country X and country Y . She takes a random sample of 50 birds of this species from country X and a random sample of 80 birds of this species from country Y . She records their masses in kg, x and y , respectively. Her results are summarised as follows.

$$\Sigma x = 75.5 \quad \Sigma x^2 = 115.2 \quad \Sigma y = 116.8 \quad \Sigma y^2 = 172.6$$

The population mean masses of these birds in countries X and Y are μ_x kg and μ_y kg respectively.

Test, at the 5% significance level, the null hypothesis $\mu_x = \mu_y$ against the alternative hypothesis $\mu_x > \mu_y$. State your conclusion in the context of the question. [8]



42 - (9231/41_Winter_2022_Q6) - Inference Using Normal And T-distributions

A company manufactures copper pipes. The pipes are produced by two different machines, A and B . An inspector claims that the mean diameter of the pipes produced by machine A is greater than the mean diameter of the pipes produced by machine B . He takes a random sample of 12 pipes produced by machine A and measures their diameters, x cm. His results are summarised as follows.

$$\sum x = 6.24 \quad \sum x^2 = 3.26$$

He also takes a random sample of 10 pipes produced by machine B and measures their diameters in cm. His results are as follows.

0.48 0.53 0.47 0.54 0.54 0.55 0.46 0.55 0.50 0.48

The diameters of the pipes produced by each machine are assumed to be normally distributed with equal population variances.

Test at the 2.5% significance level whether the data supports the inspector's claim. [9]



43 - (9231/41_Summer_2023_Q2) - *Inference Using Normal And T-distributions*

The children at two large schools, P and Q , are all given the same puzzle to solve. A random sample of size 10 is taken from the children at school P . Their individual times to complete the puzzle give a sample mean of 9.12 minutes and an unbiased variance estimate of 2.16 minutes². A random sample of size 12 is taken from the children at school Q . Their individual times, x minutes, to complete the puzzle are summarised by

$$\sum x = 99.6 \quad \sum (x - \bar{x})^2 = 21.5,$$

where $\bar{x}$ is the sample mean. Times to complete the puzzle are assumed to be normally distributed with the same population variance.

Test at the 5% significance level whether the population mean time taken to complete the puzzle by children at school P is greater than the population mean time taken to complete the puzzle by children at school Q . [8]

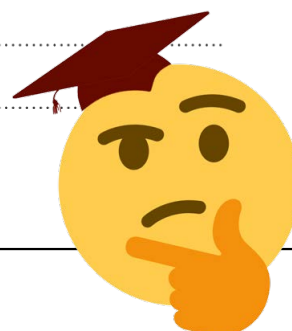


44 - (9231/43_Summer_2023_Q4) - Inference Using Normal And T-distributions

An inspector is checking the lengths of metal rods produced by two machines, X and Y . These rods should be of the same length, but the inspector suspects that those made by machine X are shorter, on average, than those made by machine Y . The inspector chooses a random sample of 80 rods made by machine X and a random sample of 60 rods made by machine Y . The lengths of these rods are x cm and y cm respectively. Her results are summarised as follows.

$$\Sigma x = 164.0 \quad \Sigma x^2 = 338.1 \quad \Sigma y = 124.8 \quad \Sigma y^2 = 261.1$$

- (a) Test at the 10% significance level whether the data supports the inspector's suspicion. [8]

[illegible]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

(b) Give a reason why it is not necessary to make any assumption about the distributions of the lengths of the rods. [1]

.....

.....

.....

.....

.....

.....

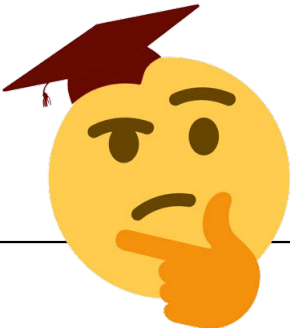
.....

.....

.....

.....

.....



45 - (9231/41_Winter_2023_Q1) - Inference Using Normal And T-distributions

Maya is an athlete who competes in 1500-metre races. Last summer her practice run times had mean 4.22 minutes. Over the winter she has done some intense training to try to improve her times. A random sample of 10 of her practice run times, x minutes, this summer are summarised as follows.

$$\sum x = 42.05 \quad \sum x^2 = 176.83$$

Maya's new practice run times are normally distributed. She believes that on average her times have improved as a result of her training.

Test, at the 5% significance level, whether Maya's belief is supported by the data. [6]



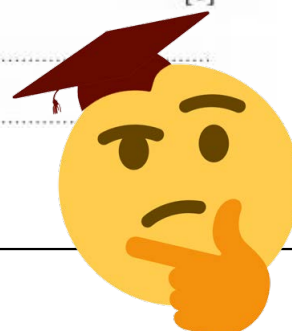
46 - (9231/43_Winter_2023_Q3) - Inference Using Normal And T-distributions

Scientists are studying the effects of exercise on LDL blood cholesterol levels. Over a three-month period, a large group of people exercised for 20 minutes each day. For a randomly chosen sample of 10 of these people, the LDL blood cholesterol levels were measured at the beginning and the end of the three-month period. The results, measured in suitable units, are as follows.

	Person	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>	<i>G</i>	<i>H</i>	<i>I</i>	<i>J</i>
Cholesterol level	Beginning	72	84	120	90	102	135	64	75	80	88
	End	64	76	105	92	105	115	67	75	75	84

- (a) Test, at the 2.5% significance level, whether there is evidence that the population mean LDL blood cholesterol level has reduced by more than 2 units after the three-month period. [7]

- (b) State any assumption that you have made in part (a). [1]



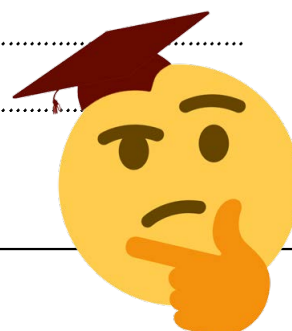
Roberto owns a small hotel and offers accommodation to guests. Over a period of 100 nights, the numbers of rooms, x , that are occupied each night at Roberto's hotel and the corresponding frequencies are shown in the following table.

Number of rooms occupied (x)	0	1	2	3	4	5	6	≥ 7
Number of nights	4	9	18	26	20	16	7	0

-
-
-

Number of rooms occupied (x)	0	1	2	3	4	5	6	≥ 7
Observed frequency	4	9	18	26	20	16	7	0
Expected frequency	3.88	12.60	20.48	22.18	18.02	11.72		

- 
- A decorative graphic in the bottom right corner featuring a red graduation cap (mortarboard) and a yellow diploma with a red ribbon, symbolizing the end of a course or achievement.



- (iii) Use a goodness-of-fit test at the 5% significance level to determine whether the Poisson distribution is a suitable model for the number of rooms occupied each night at Roberto's hotel. [7]



2 - (9231/21_Summer_2017_Q11) - Simple Harmonic Motion, X2 Test

Answer only **one** of the following two alternatives.

EITHER

A particle P of mass $3m$ is attached to one end of a light elastic spring of natural length a and modulus of elasticity kmg . The other end of the spring is attached to a fixed point O on a smooth plane that is inclined to the horizontal at an angle α , where $\sin \alpha = \frac{2}{3}$. The system rests in equilibrium with P on the plane at the point E . The length of the spring in this position is $\frac{5}{4}a$.

- (i) Find the value of k . [3]

.....

.....

.....

.....

.....

.....

.....

The particle P is now replaced by a particle Q of mass $2m$ and Q is released from rest at the point E .

- (ii) Show that, in the resulting motion, Q performs simple harmonic motion. State the centre and the period of the motion. [6]

.....

.....

.....

.....

.....

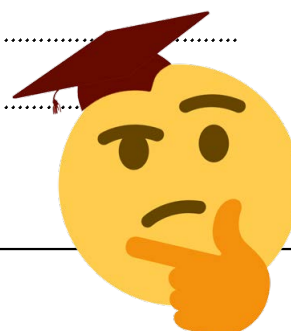
.....

.....

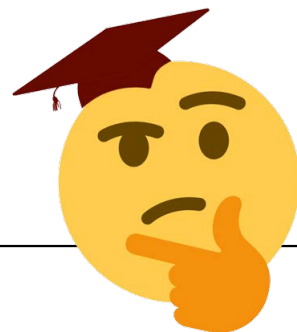
.....

.....

.....



(iii) Find the least tension in the spring and the maximum acceleration of Q during the motion. [5]



OR

A shop is supplied with large quantities of plant pots in packs of six. These pots can be damaged easily if they are not packed carefully. The manager of the shop is a statistician and he believes that the number of damaged pots in a pack of six has a binomial distribution. He chooses a random sample of 250 packs and records the numbers of damaged pots per pack. His results are shown in the following table.

Number of damaged pots per pack (x)	0	1	2	3	4	5	6
Frequency	48	69	78	32	22	1	0

- (i) Show that the mean number of damaged pots per pack in this sample is 1.656. [1]

.....

.....

.....

.....

.....

The following table shows some of the expected frequencies, correct to 2 decimal places, using an appropriate binomial distribution.

Number of damaged pots per pack (x)	0	1	2	3	4	5	6
Expected frequency	36.01	82.36	a	39.89	b	1.74	0.11

- (ii) Find the values of a and b , correct to 2 decimal places [5]

.....

.....

.....

.....

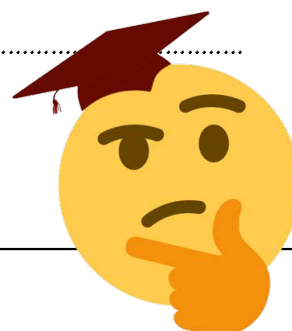
.....

.....

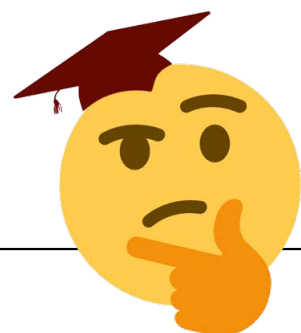
.....

.....

.....



- (iii) Use a goodness-of-fit test at the 1% significance level to determine whether the manager's belief is justified. [8]



3 - (9231/21_Winter_2017_Q8) - *X2 Test*

Members of a Statistics club are voting to elect a new president of the club. Members must choose to vote either by post or by text or by email. The method of voting chosen by a random sample of 60 male members and 40 female members is given in the following table.

		Method of voting		
		Post	Text	Email
Gender	Male	10	12	38
	Female	5	21	14

Test, at the 1% significance level, whether there is an association between method of voting and gender. [8]



4 - (9231/21_Summer_2018_Q8) - X2 Test

A manufacturer produces three types of car: hatchbacks, saloons and estates. Each type of car is available in one of three colours: silver, blue and red. The manufacturer wants to know whether the popularity of the colour of the car is related to the type of car. A random sample of 300 cars chosen by customers gives the information summarised in the following table.

		Colour of car		
		Silver	Blue	Red
Type of car	Hatchback	53	36	41
	Saloon	29	40	31
	Estate	28	24	18

Test at the 10% significance level whether the colour of car chosen by customers is independent of the type of car. [8]



5 - (9231/22_Winter_2018_Q10) - X2 Test

The number of accidents, x , that occur each day on a motorway are recorded over a period of 40 days. The results are shown in the following table.

Number of accidents	0	1	2	3	4	5	6	≥ 7
Observed frequency	3	5	8	10	5	7	2	0

- (i) Show that the mean number of accidents each day is 2.95 and calculate the variance for this sample. Explain why these values suggest that a Poisson distribution might fit the data. [3]

.....

.....

.....

.....

.....

.....

.....

A Poisson distribution with mean 2.95, as found from the data, is used to calculate the expected frequencies, correct to 2 decimal places. The results are shown in the following table.

Number of accidents	0	1	2	3	4	5	6	≥ 7
Observed frequency	3	5	8	10	5	7	2	0
Expected frequency	2.09	6.18	9.11	8.96	6.61	3.90	1.92	1.23

- (ii) Show how the expected frequency of 6.61 for $x = 4$ is obtained. [2]
- (iii) Test at the 5% significance level the goodness of fit of this Poisson distribution to the data. [7]



6 - (9231/21_Winter_2018_Q11) - *Circular Motion, X2 Test*

Answer only **one** of the following two alternatives.

EITHER

A particle P of mass m is free to move on the smooth inner surface of a fixed hollow sphere of radius a . The centre of the sphere is O and the point C is on the inner surface of the sphere, vertically below O . The points A and B on the inner surface of the sphere are the ends of a diameter of the sphere. The diameter AOB makes an acute angle α with the vertical, where $\cos \alpha = \frac{4}{5}$, with A below the horizontal level of B . The particle is projected from A with speed u , and moves along the inner surface of the sphere towards C . The normal reaction forces on the particle at A and C are in the ratio 8 : 9.

(i) Show that $u^2 = 4ag$. [6]

(ii) Determine whether P reaches B without losing contact with the inner surface of the sphere. [6]



OR

A machine is used to produce metal rods. When the machine is working efficiently, the lengths, x cm, of the rods have a normal distribution with mean 150 cm and standard deviation 1.2 cm. The machine is checked regularly by taking random samples of 200 rods. The latest results are shown in the following table.

Interval	$146 \leq x < 147$	$147 \leq x < 148$	$148 \leq x < 149$	$149 \leq x < 150$
Observed frequency	1	2	23	52

$150 \leq x < 151$	$151 \leq x < 152$	$152 \leq x < 153$	$153 \leq x < 154$
69	36	15	2

As a first check, the sample is used to calculate an estimate for the mean.

- (i) Show that an estimate for the mean from this sample is close to 150 cm. [2]

.....

.....

.....

.....

As a second check, the results are tested for goodness of fit of the normal distribution with mean 150 cm and standard deviation 1.2 cm. The relevant expected frequencies, found using the normal distribution function given in the List of Formulae (MF10), are shown in the following table.

Interval	$x < 147$	$147 \leq x < 148$	$148 \leq x < 149$	$149 \leq x < 150$
Observed frequency	1	2	23	52
Expected frequency	1.24	8.32	30.94	59.50

$150 \leq x < 151$	$151 \leq x < 152$	$152 \leq x < 153$	$153 \leq x$
69	36	15	2
59.50	30.94	8.32	1.24

- (ii) Show how the expected frequency for $151 \leq x < 152$ is obtained. [3]
- (iii) Test, at the 5% significance level, the goodness of fit of the normal distribution to the results. [7]



7 - (9231/23_Summer_2019_Q8) - X2 Test

Two salesmen, *A* and *B*, work at a company that arranges different types of holidays: self-catering, hotel and cruise. The table shows, for a random sample of 150 holidays, the number of each type arranged by each salesman.

		Type of holiday		
		Self-catering	Hotel	Cruise
Salesman	<i>A</i>	25	38	21
	<i>B</i>	28	21	17

Test at the 10% significance level whether the type of holiday arranged is independent of the salesman.
[8]



8 - (9231/21_Summer_2019_Q9) - X2 Test

A random sample of 50 observations of the continuous random variable X was taken and the values are summarised in the following table.

Interval	$0 \leq x < 0.8$	$0.8 \leq x < 1.6$	$1.6 \leq x < 2.4$	$2.4 \leq x < 3.2$	$3.2 \leq x < 4$
Observed frequency	18	16	8	6	2

It is required to test the goodness of fit of the distribution with probability density function f given by

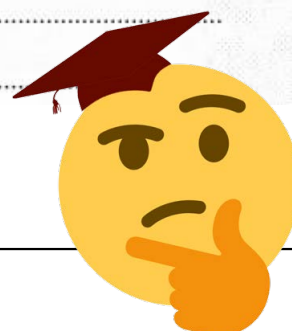
$$f(x) = \begin{cases} \frac{3}{16}(4-x)^{\frac{1}{2}} & 0 \leq x < 4, \\ 0 & \text{otherwise.} \end{cases}$$

The relevant expected frequencies, correct to 2 decimal places, are given in the following table.

Interval	$0 \leq x < 0.8$	$0.8 \leq x < 1.6$	$1.6 \leq x < 2.4$	$2.4 \leq x < 3.2$	$3.2 \leq x < 4$
Expected frequency	14.22	12.54	10.59	8.18	4.47

(i) Show how the expected frequency for $1.6 \leq x < 2.4$ is obtained.

[3]



[7]

[illegible]

9 - (9231/21_Winter_2019_Q11) - *Simple Harmonic Motion, X2 Test*

Answer only **one** of the following two alternatives.

EITHER

The points A and B are a distance 1.2 m apart on a smooth horizontal surface. A particle P of mass $\frac{2}{3}$ kg is attached to one end of a light spring of natural length 0.6 m and modulus of elasticity 10 N. The other end of the spring is attached to the point A . A second light spring, of natural length 0.4 m and modulus of elasticity 20 N, has one end attached to P and the other end attached to B .

- (i) Show that when P is in equilibrium $AP = 0.75$ m. [3]

.....

.....

.....

.....

.....

.....

.....

.....

The particle P is displaced by 0.05 m from the equilibrium position towards A and then released from rest.

- (ii) Show that P performs simple harmonic motion and state the period of the motion. [6]

.....

.....

.....

.....

.....

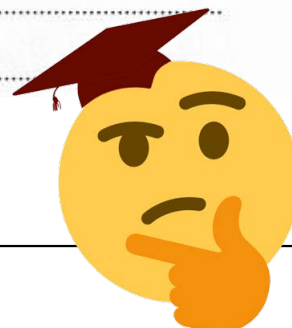
.....

.....

.....

.....

.....



.....

.....

.....

.....

.....

.....

.....

(iii) Find the speed of P when it passes through the equilibrium position. [2]

.....

.....

.....

.....

.....

(iv) Find the speed of P when its acceleration is equal to half of its maximum value. [3]

.....

.....

.....

.....

.....

.....

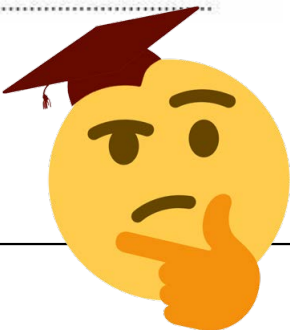
.....

.....

.....

.....

.....



OR

The number of puncture repairs carried out each week by a small repair shop is recorded over a period of 40 weeks. The results are shown in the following table.

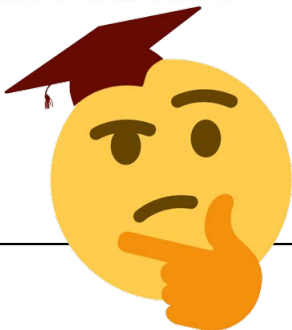
Number of repairs in a week	0	1	2	3	4	5	≥ 6
Number of weeks	6	15	9	6	3	1	0

- (i) Calculate the mean and variance for the number of repairs in a week and comment on the possible suitability of a Poisson distribution to model the data. [3]

Records over a longer period of time indicate that the mean number of repairs in a week is 1.6. The following table shows some of the expected frequencies, correct to 3 decimal places, for a period of 40 weeks using a Poisson distribution with mean 1.6.

Number of repairs in a week	0	1	2	3	4	5	≥ 6
Expected frequency	8.076	12.921	10.337	5.513	2.205	a	b

- (ii) Show that $a = 0.706$ and find the value of the constant b . [3]



[8]

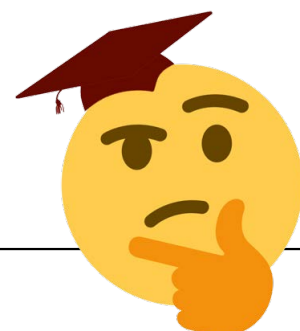


10 - (9231/41_Summer_2020_Q1) - X2 Test

Two randomly selected groups of students, with similar ranges of abilities, take the same examination in different rooms. One group of 140 students takes the examination with background music playing. The other group of 210 students takes the examination in silence. Each student is awarded a grade for their performance in the examination and the numbers from each group gaining each grade are shown in the following table.

	Grade awarded		
	A	B	C
Background music	49	51	40
Silence	93	68	49

Test at the 10% significance level whether grades awarded are independent of whether background music is playing during the examination. [6]



11 - (9231/43_Summer_2020_Q1) - X2 Test

Young children are learning to read using two different reading schemes, *A* and *B*. The standards achieved are measured against the national average standard achieved and classified as above average, average or below average. For two randomly chosen groups of young children, the numbers in each category are shown in the table.

	Standard achieved		
	Above average	Average	Below average
Scheme <i>A</i>	31	35	22
Scheme <i>B</i>	19	50	43

Test at the 5% significance level whether standard achieved is independent of the reading scheme used.
[6]



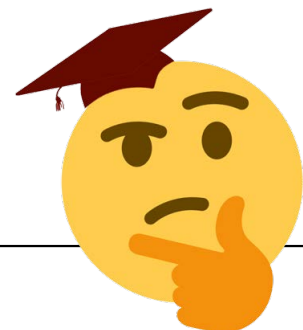
12 - (9231/41_Winter_2020_Q3) - X2 Test

Apples are sold in bags of 5. Based on her previous experience, Freya claims that the probability of any apple weighing more than 100 grams is 0.35, independently of other apples in the bag.

The apples in a random sample of 150 bags are checked and the number, x , in each bag weighing more than 100 grams is recorded. The results are shown in the following table.

x	0	1	2	3	4	5
Frequency	12	39	46	37	12	4

Carry out a goodness of fit test at the 5% significance level and hence comment on Freya's claim. [7]



13 - (9231/42_Winter_2020_Q3) - X2 Test

A random sample of 200 observations of the continuous random variable X was taken and the values are summarised in the following table.

Interval	$0 \leq x < 0.5$	$0.5 \leq x < 1$	$1 \leq x < 1.5$	$1.5 \leq x < 2$	$2 \leq x < 2.5$	$2.5 \leq x < 3$
Observed frequency	5	23	40	41	46	45

It is required to test the goodness of fit of the distribution with probability density function f given by

$$f(x) = \begin{cases} \frac{1}{9}x(4-x) & 0 \leq x \leq 3, \\ 0 & \text{otherwise.} \end{cases}$$

Most of the relevant expected frequencies, correct to 2 decimal places, are given in the following table.

Interval	$0 \leq x < 0.5$	$0.5 \leq x < 1$	$1 \leq x < 1.5$	$1.5 \leq x < 2$	$2 \leq x < 2.5$	$2.5 \leq x < 3$
Expected frequency	p	q	37.96	43.52	43.52	37.96

(a) Show that $p = 10.19$ and find the value of q . [3]

(b) Carry out a goodness of fit test, at the 5% significance level, to test whether f is a satisfactory model for the data. [4]

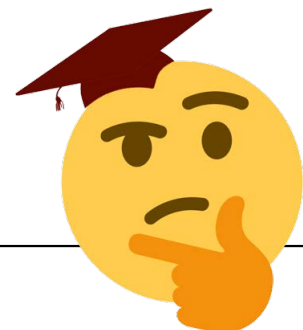


14 - (9231/41_Summer_2021_Q2) - X2 Test

A driving school employs four instructors to prepare people for their driving test. The allocation of people to instructors is random. For each of the instructors, the following table gives the number of people who passed and the number who failed their driving test last year.

	Instructor <i>A</i>	Instructor <i>B</i>	Instructor <i>C</i>	Instructor <i>D</i>	Total
Pass	72	42	52	68	234
Fail	33	34	41	58	166
Total	105	76	93	126	400

Test at the 10% significance level whether success in the driving test is independent of the instructor.
[7]



15 - (9231/43_Summer_2021_Q5) - X2 Test

Chai packs china mugs into cardboard boxes. Chai's manager suspects that breakages occur at random times and that the number of breakages may follow a Poisson distribution. He takes a small sample of observations and finds that the number of breakages in a one-hour period has a mean of 2.4 and a standard deviation of 1.5.

- (a) Explain how this information tends to support the manager's suspicion. [2]

.....

.....

.....

.....

.....

.....

.....

.....

.....

The manager now takes a larger sample and claims that the numbers of breakages in a one-hour period follow a Poisson distribution. The numbers of breakages in a random sample of 180 one-hour periods are summarised in the following table.

Number of breakages	0	1	2	3	4	5	6	7 or more
Frequency	21	33	46	31	23	16	10	0

The mean number of breakages calculated from this sample is 2.5.

- (b) Use the data from this larger sample to carry out a goodness of fit test, at the 10% significance level, to test the claim. [8]



16 - (9231/42_Winter_2021_Q2) - X2 Test

It is claimed that the heights of a particular age group of boys follow a normal distribution with mean 125 cm and standard deviation 12 cm. Observations for a randomly chosen group of 60 boys in this age group are summarised in the following table. The table also gives the expected frequencies, correct to 2 decimal places, based on the normal distribution with mean 125 cm and standard deviation 12 cm.

Height, x cm	$x < 100$	$100 \leq x < 110$	$110 \leq x < 120$	$120 \leq x < 130$	$130 \leq x < 140$	$x \geq 140$
Observed frequency	0	3	15	23	11	8
Expected frequency	1.12	5.22	13.97	19.38	13.97	6.34

- (a) Show how the expected frequency for $130 \leq x < 140$ is obtained. [2]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

- (b) Carry out a goodness of fit test, at the 5% significance level, to determine whether the claim is supported by the data. [6]

.....

.....

.....

.....

.....

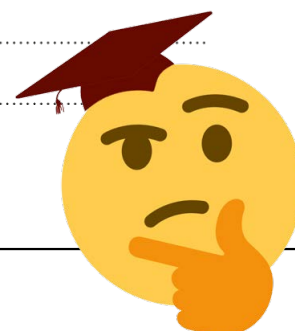
.....

.....

.....

.....

.....



17 - (9231/41_Winter_2021_Q3) - X2 Test

A supermarket sells pears in packs of 8. Some of the pears in a pack may not be ripe, and the supermarket manager claims that the number of unripe pears in a pack can be modelled by the distribution $B(8, 0.15)$.

A random sample of 150 packs was selected and the number of unripe pears in each pack was recorded. The following table shows the observed frequencies together with some of the expected frequencies using the manager's binomial distribution.

Number of unripe pears per pack	0	1	2	3	4	5	≥ 6
Observed frequency	35	48	43	15	6	3	0
Expected frequency	40.874	p	35.641	12.579	2.775	0.392	q

- (a) Find the values of p and q . [2]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

- (b) Carry out a goodness of fit test, at the 5% significance level, to test whether the manager's claim is justified. [6]

.....

.....

.....

.....

.....

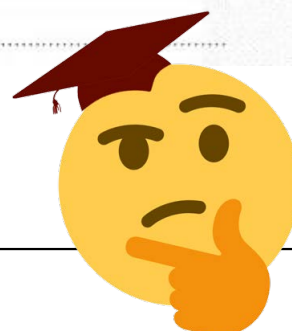
.....

.....

.....

.....

.....



18 - (9231/42_Summer_2022_Q4) - X2 Test

A scientist is investigating the numbers of a particular type of butterfly in a certain region. He claims that the numbers of these butterflies found per square metre can be modelled by a Poisson distribution with mean 2.5. He takes a random sample of 120 areas, each of one square metre, and counts the number of these butterflies in each of these areas. The following table shows the observed frequencies together with some of the expected frequencies using the scientist's Poisson distribution.

Number per square metre	0	1	2	3	4	5	6	≥ 7
Observed frequency	12	20	36	32	13	6	1	0
Expected frequency	9.85	24.63	30.78	25.65	p	8.02	3.34	q

- (a) Find the values of p and q , correct to 2 decimal places. [2]

.....

.....

.....

.....

.....

.....

.....

.....

.....

- (b) Carry out a goodness of fit test, at the 10% significance level, to test the scientist's claim. [6]

.....

.....

.....

.....

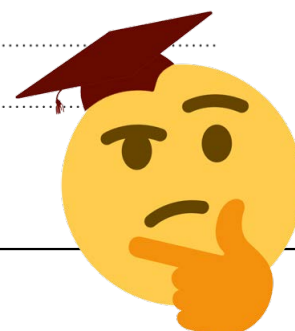
.....

.....

.....

.....

.....



19 - (9231/41_Winter_2022_Q2) - X2 Test

An organisation runs courses to train students to become engineers. These students are taught in groups of 8. The director of the organisation claims that on average 60% of the students in a group achieve a pass. A random sample of 150 groups of 8 students is chosen. The following table shows the observed frequencies together with some of the expected frequencies using the appropriate binomial distribution.

Number of passes per group	0	1	2	3	4	5	6	7	8
Observed frequency	0	0	8	24	45	36	26	10	1
Expected frequency	p	1.180	6.193	18.579	34.836	q	r	13.437	2.519

- (a) Find the values of p , q and r giving your answers correct to 3 decimal places. [2]

.....

.....

.....

.....

.....

.....

.....

.....

- (b) Carry out a goodness of fit test, at the 10% significance level, to test whether there is evidence to reject the director's claim. [6]

.....

.....

.....

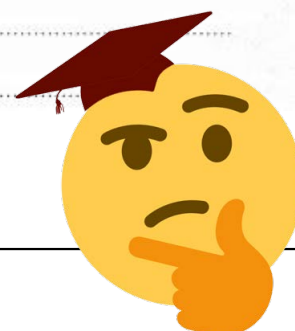
.....

.....

.....

.....

.....



20 - (9231/42_Winter_2022_Q2) - X2 Test

In the colleges in three regions of a particular country, students are given individual targets to achieve. Their performance is measured against their individual target and graded as 'above target', 'on target' or 'below target'. For a random sample of students from each of the three regions, the observed frequencies are summarised in the following table.

		Region			Total
		<i>A</i>	<i>B</i>	<i>C</i>	
Performance	Above target	62	41	44	147
	On target	102	94	95	291
	Below target	56	45	61	162
Total		220	180	200	600

Test, at the 10% significance level, whether performance is independent of region.

[7]

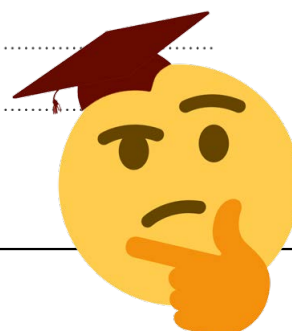


21 - (9231/43_Summer_2023_Q6) - X2 Test

A scientist is investigating whether the ability to remember depends on age. A random sample of 150 students in different age groups is chosen. Each student is shown a set of 20 objects for thirty seconds and then asked to list as many as they can remember. The students are graded *A* or *B* according to how many objects they remembered correctly: grade *A* for 16 or more correct and grade *B* for fewer than 16 correct. The results are shown in the table.

	Age of students		
	11–12 years	13–14 years	15–16 years
Grade A	25	16	19
Grade B	28	45	17

- (a) Carry out a χ^2 -test at the 2.5% significance level to test whether grade is independent of age of student. [7]



The scientist decides instead to use three grades: grade *A* for 16 or more correct, grade *B* for 10 to 15 correct and grade *C* for fewer than 10 correct. The results are shown in the following table.

	Age of students		
	11–12 years	13–14 years	15–16 years
Grade <i>A</i>	25	16	19
Grade <i>B</i>	12	27	11
Grade <i>C</i>	16	18	6

With this second set of data, the test statistic is calculated as 10.91.

- (b) Complete the χ^2 -test at the 2.5% significance level for this second set of data. [2]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

- (c) State, with a reason, whether you would prefer to use the result from part (a) or part (b) to investigate whether the ability to remember depends on age. [1]

.....

.....

.....

.....

.....

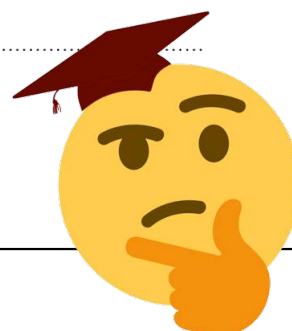
.....

.....

.....

.....

.....



22 - (9231/41_Winter_2023_Q2) - X2 Test

A town council has published its plans for redeveloping the town centre and residents are being asked whether they approve or disapprove. A random sample of 250 responses has been selected from residents in the four main streets in the town: North, East, South and West Streets. The results are shown in the table.

	North Street	East Street	South Street	West Street
Approve	33	54	42	26
Disapprove	19	39	28	9

Test, at the 5% significance level, whether the opinions of the residents are independent of the streets on which they live. [7]



23 - (9231/42_Winter_2023_Q2) - X2 Test

The number of breakdowns on a particular section of road is recorded each day over a period of 90 days. It is suggested that the number of breakdowns follows a Poisson distribution with mean 3.5. The data is summarised in the table, together with some of the expected frequencies resulting from the suggested Poisson distribution.

Number of breakdowns per day	0	1	2	3	4	5	6	7	8 or more
Observed frequency	0	5	13	17	21	16	9	5	4
Expected frequency	2.718	9.512	16.646		16.993	11.895		3.469	2.407

(a) Complete the table.

[2]

.....

.....

.....

.....

.....

.....

.....

(b) Carry out a goodness of fit test, at the 10% significance level, to determine whether or not $Po(3.5)$ is a good fit to the data.

[6]

.....

.....

.....

.....

.....

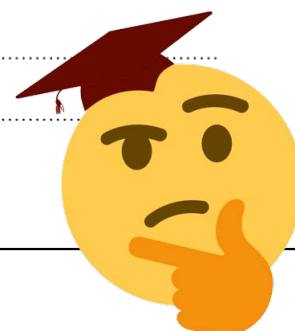
.....

.....

.....

.....

.....



This image shows a blank sheet of white paper with horizontal dashed lines, typical of primary school handwriting practice paper. The lines are evenly spaced and run across the width of the page. In the bottom right corner, there is a small, stylized illustration of a red pencil pointing towards the left. The rest of the page is empty, providing space for writing practice.

1 - (9231/23_Summer_2017_Q7) - Bivariate Data

A random sample of twelve pairs of values of x and y is taken from a bivariate distribution. The equations of the regression lines of y on x and of x on y are respectively

$y = 0.46x + 1.62$ and $x = 0.93y + 8.24$.

(i) Find the value of the product moment correlation coefficient for this sample. [2]

.....

.....

.....

.....

.....

.....

.....

(ii) Using a 5% significance level, test whether there is non-zero correlation between the variables. [4]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

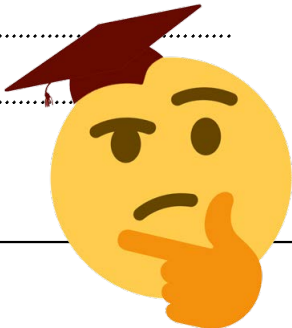
.....

.....

.....

.....

.....



2 - (9231/21_Summer_2017_Q10) - Bivariate Data

A random sample of 5 pairs of values (x, y) is given in the following table.

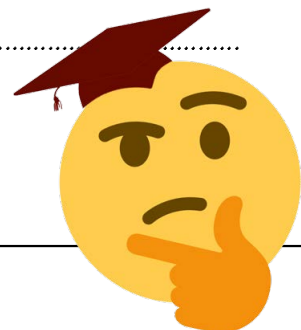
x	1	2	4	5	8
y	7	5	8	6	4

- (i) Find, showing all necessary working, the equation of the regression line of y on x . [4]



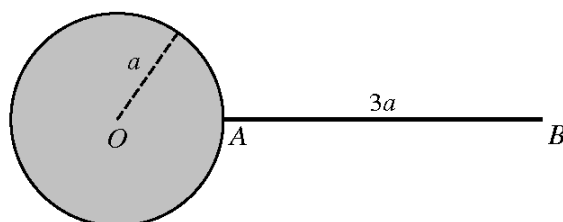
- [illegible]

-
- This image shows a blank sheet of white paper designed for handwriting practice. It features ten horizontal dashed lines spaced evenly across the page. In the bottom right corner, there is a small illustration of a red pencil pointing towards the center of the page. The entire sheet is otherwise empty, providing space for writing practice.



3 - (9231/23_Summer_2017_Q11) - Rotation Of A Rigid Body, Bivariate Data

Answer only **one** of the following two alternatives.

EITHER

The diagram shows a uniform thin rod AB of length $3a$ and mass $8m$. The end A is rigidly attached to the surface of a sphere with centre O and radius a . The rod is perpendicular to the surface of the sphere. The sphere consists of two parts: an inner uniform solid sphere of mass $\frac{3}{2}m$ and radius a surrounded by a thin uniform spherical shell of mass m and also of radius a . The horizontal axis l is perpendicular to the rod and passes through the point C on the rod where $AC = a$.

- (i) Show that the moment of inertia of the object, consisting of rod, shell and inner sphere, about the axis l is $\frac{289}{15}ma^2$. [6]

The object is free to rotate about the axis l . The object is held so that CA makes an angle α with the downward vertical and is released from rest.

- (ii) Given that $\cos \alpha = \frac{1}{6}$, find the greatest speed achieved by the centre of the sphere in the subsequent motion. [6]

OR

The times taken to run 200 metres at the beginning of the year and at the end of the year are recorded for each member of a large athletics club. The time taken, in seconds, at the beginning of the year is denoted by x and the time taken, in seconds, at the end of the year is denoted by y . For a random sample of 8 members, the results are shown in the following table.

Member	A	B	C	D	E	F	G	H
x	24.2	23.8	22.8	25.1	24.5	24.0	23.8	22.8
y	23.9	23.6	22.8	24.5	24.2	23.5	23.6	22.7

$$[\Sigma x = 191, \quad \Sigma x^2 = 4564.46, \quad \Sigma y = 188.8, \quad \Sigma y^2 = 4458.4, \quad \Sigma xy = 4510.99.]$$

- (i) Find, showing all necessary working, the equation of the regression line of y on x . [4]



The athletics coach believes that, on average, the time taken by an athlete to run 200 metres decreases between the beginning and the end of the year by more than 0.2 seconds.

- (ii) Stating suitable hypotheses and assuming a normal distribution, test the coach's belief at the 10% significance level. [8]



4 - (9231/21_Winter_2017_Q9) - Bivariate Data

The land areas x (in suitable units) and populations y (in millions) for a sample of 8 randomly chosen cities are given in the following table.

Land area (x)	1.0	4.5	2.4	1.6	3.8	8.6	7.5	6.5
Population (y)	0.8	8.4	4.2	1.6	2.2	10.2	4.2	5.2

[$\Sigma x = 35.9$, $\Sigma x^2 = 216.47$, $\Sigma y = 36.8$, $\Sigma y^2 = 244.96$, $\Sigma xy = 212.62$.]

- (i) Find, showing all necessary working, the value of the product moment correlation coefficient for this sample. [3]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

- (ii) Using a 1% significance level, test whether there is positive correlation between land area and population of cities. [4]

.....

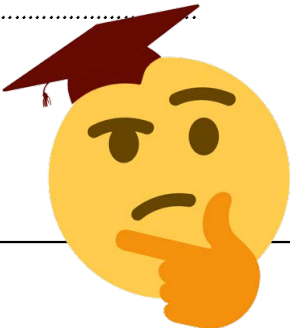
.....

.....

.....

.....

.....



The land areas and populations for another randomly chosen sample of cities, this time of size n , give a product moment correlation coefficient of 0.651. Using a test at the 1% significance level, there is evidence of non-zero correlation between the variables.

(iii) Find the least possible value of n , justifying your answer. [2]

.....

.....

.....

.....

.....

.....

.....

.....

.....

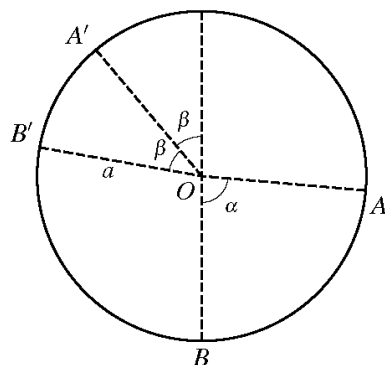
.....



5 - (9231/21_Winter_2017_Q11) - Circular Motion, Bivariate Data

Answer only **one** of the following two alternatives.

EITHER



A particle P of mass m is free to move on the smooth inner surface of a fixed hollow sphere of radius a . The centre of the sphere is O . The points A and A' are on the inner surface of the sphere, on opposite sides of the vertical through O ; the radius OA makes an angle α with the downward vertical and the radius OA' makes an angle β with the upward vertical. The point B is on the inner surface of the sphere, vertically below O . The point B' is on the inner surface of the sphere and such that OB' makes an angle 2β with the upward vertical through O (see diagram). It is given that $\cos \alpha = \frac{1}{16}$.

- (i) P is projected from A with speed u along the surface of the sphere downwards towards B . Subsequently it loses contact with the sphere at A' . Show that $u^2 = \frac{1}{8}ag(1 + 24 \cos \beta)$. [5]
- (ii) P is now projected from B with speed u along the surface of the sphere towards B' . Subsequently it loses contact with the sphere at B' . Find $\cos \beta$. [6]
- (iii) In part (i), the reaction of the sphere on P when it is initially projected at A is R . Find R in terms of m and g . [3]



OR

A large number of people attended a course to improve the speed of their logical thinking. The times taken to complete a particular type of logic puzzle at the beginning of the course and at the end of the course are recorded for each person. The time taken, in minutes, at the beginning of the course is denoted by x and the time taken, in minutes, at the end of the course is denoted by y . For a random sample of 9 people, the results are summarised as follows.

$$\Sigma x = 45.3 \quad \Sigma x^2 = 245.59 \quad \Sigma y = 40.5 \quad \Sigma y^2 = 195.11 \quad \Sigma xy = 218.72$$

Ken attended the course, but his time to complete the puzzle at the beginning of the course was not recorded. His time to complete the puzzle at the end of the course was 4.2 minutes.

- (i) By finding, showing all necessary working, the equation of a suitable regression line, find an estimate for the time that Ken would have taken to complete the puzzle at the beginning of the course. [5]

The values of $x - y$ for the sample of 9 people are as follows.

$$0.2 \quad 0.8 \quad 0.5 \quad 1.0 \quad 0.2 \quad 0.6 \quad 0.2 \quad 0.5 \quad 0.8$$

The organiser of the course believes that, on average, the time taken to complete the puzzle decreases between the beginning and the end of the course by more than 0.3 minutes.

- (ii) Stating suitable hypotheses and assuming a normal distribution, test the organiser's belief at the $2\frac{1}{2}\%$ significance level. [9]



6 - (9231/23_Summer_2018_Q6) - *Bivariate Data*

A random sample of 15 observations of pairs of values of two variables gives a product moment correlation coefficient of 0.430.

- (i) Test at the 10% significance level whether there is evidence of non-zero correlation between the variables. [4]

.....

.....

.....

.....

.....

.....

.....

.....

A second random sample of N observations gives a product moment correlation coefficient of 0.615. Using a 5% significance level, there is evidence of positive correlation between the variables.

- (ii) Find the least possible value of N , justifying your answer. [2]



7 - (9231/23_Summer_2018_Q8) - Bivariate Data

For a random sample of 6 observations of pairs of values (x, y) , the equation of the regression line of y on x is $y = bx + 1.306$, where b is a constant. The corresponding equation of the regression line of x on y is $x = 0.6331y + d$, where d is a constant. The values of x from the sample are

2.3	2.8	3.7	p	6.1	6.4
-----	-----	-----	-----	-----	-----

and the sum of the values of y is 46.5. The product moment correlation coefficient is 0.9797.

(i) Find the value of b correct to 3 decimal places.

[2]

.....

.....

.....

.....

.....

.....

.....

.....

(ii) Find the value of p .

[4]

(iii) Use the equation of the regression line of x on y to estimate the value of x when $y = 8.5$.

[3]



8 - (9231/21_Summer_2018_Q11) - Circular Motion, Bivariate Data

Answer only **one** of the following two alternatives.

EITHER

A particle P of mass m is attached to one end of a light inextensible string of length a . The other end of the string is attached to a fixed point O . The particle is held so that the string is taut, with OP horizontal. The particle is projected downwards with speed $\sqrt{\left(\frac{2}{5}ag\right)}$ and begins to move in a vertical circle. The string breaks when its tension is equal to $\frac{11}{5}mg$.

- (i) Show that the string breaks when OP makes an angle θ with the downward vertical through O , where $\cos \theta = \frac{3}{5}$. Find the speed of P at this instant. [6]
- (ii) For the subsequent motion after the string breaks, find the distance OP when the particle P is vertically below O . [6]

OR

The regression line of y on x , obtained from a random sample of 6 pairs of values of x and y , has equation

$$y = 0.25x + k,$$

where k is a constant. The values from the sample are shown in the following table.

x	4	5	7	8	10	14
y	5	8	p	7	p	9

- (i) Find the value of p and the value of k . [6]



(ii) Find the product moment correlation coefficient for the data.

[2]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

(iii) Test, at the 5% significance level, whether there is evidence of positive correlation between the variables.

[4]



9 - (9231/22_Winter_2018_Q9) - *Bivariate Data*

For a random sample of 5 observations of pairs of values (x, y) , the equation of the regression line of y on x is $y = 4.2 + cx$ and the equation of the regression line of x on y is $x = 10.8 + dy$, where c and d are constants. The product moment correlation coefficient is -0.7214 and the mean value of x is 7.018 .

- (i) Test at the 5% significance level whether there is evidence of non-zero correlation between the variables. [4]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

- (ii) Find the values of c and d . [5]

- (iii) Use an appropriate regression line to estimate the value of x when $y = 3.5$, and comment on the reliability of your estimate. [2]



10 - (9231/21_Winter_2018_Q10) - *Bivariate Data*

For a random sample of 10 observations of pairs of values (x, y) , the equation of the regression line of y on x is $y = 1.1664 + 0.4604x$. It is given that

$$\Sigma x^2 = 1419.98 \quad \text{and} \quad \Sigma y^2 = 439.68.$$

The mean value of y is 6.24.

(i) Find the equation of the regression line of x on y . [6]

(ii) Find the product moment correlation coefficient. [2]

.....

.....

.....

.....

.....

.....

.....

.....

.....

(iii) Test at the 5% significance level whether there is evidence of positive correlation between the two variables. [4]





(ii) Find the value of c . [3]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

(iii) Find the value of the product moment correlation coefficient. [3]

.....

.....

.....

.....

.....

.....

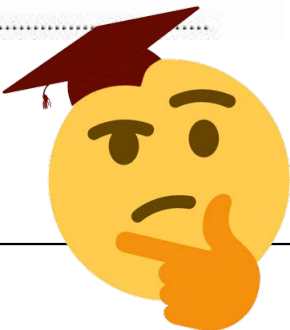
.....

.....

.....

.....

.....



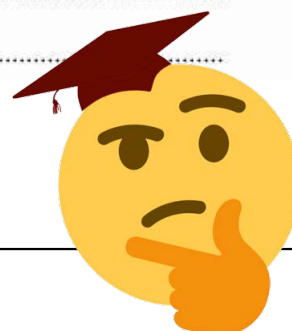
12 - (9231/23_Summer_2019_Q10) - *Bivariate Data*

The means and variances for a random sample of 8 pairs of values of x and y taken from a bivariate distribution are given in the following table.

	Mean	Variance
x	3.3125	3.3086
y	6.7375	7.9473

The product moment correlation coefficient for the sample is 0.5815, correct to 4 decimal places.

- (i) Find the equation of the regression line of y on x . [6]

This image shows a blank sheet of white paper with horizontal dashed lines, typical of primary school handwriting practice paper. The lines are evenly spaced and run across the width of the page. In the bottom right corner, there is a red arrow pointing towards the right edge of the page.

This image shows a full page of a handwriting practice worksheet. It consists of ten sets of horizontal dashed lines spaced evenly down the page, providing a guide for letter height and placement. The background is plain white, and there are no other markings or text present.

[illegible]

13 - (9231/21_Winter_2019_Q9) - *Bivariate Data*

A random sample of five pairs of values of x and y is taken from a bivariate distribution. The values are shown in the following table, where p and q are constants.

x	1	2	3	4	5
y	4	p	q	2	1

The equation of the regression line of y on x is $y = -0.5x + 3.5$.

(i) Find the values of p and q . [7]

(ii) Find the value of the product moment correlation coefficient. [3]



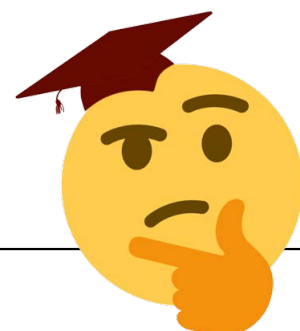
1 - (9231/41_Summer_2020_Q2) - *Non Parametric Test*

The times, in milliseconds, taken by a computer to perform a certain task were recorded on 10 randomly chosen occasions. The times were as follows.

6.44 6.16 5.62 5.82 6.51 6.62 6.19 6.42 6.34 6.28

It is claimed that the median time to complete the task is 6.4 milliseconds.

- (a) Carry out a Wilcoxon signed-rank test at the 5% significance level to test this claim. [6]
- (b) State an underlying assumption that is made when using a Wilcoxon signed-rank test. [1]



2 - (9231/43_Summer_2020_Q6) - Non Parametric Test

A biologist is studying the effect of nutrients on the heights to which plants grow. A random sample of 24 similar young plants is divided into two equal groups A and B . The plants in group A are fed with nutrients and water and the plants in group B are given only water. After four weeks, the height, in cm, of each plant is measured and the results are as follows.

Group A	12.3	11.8	12.1	13.2	11.1	10.6	13.8	12.0	12.2	12.4	13.5	13.9
Group B	11.7	10.8	10.9	11.3	11.2	12.6	11.0	10.5	11.9	12.5	10.7	11.6

The biologist decides to carry out a test at the 5% significance level to test whether the nutrients have resulted in an increase in growth.

- (a) She carries out a Wilcoxon rank-sum test. Give a reason why this is an appropriate choice of test. [1]
- (b) Carry out the Wilcoxon rank-sum test for these results. [10]



3 - (9231/41_Winter_2020_Q2) - Non Parametric Test

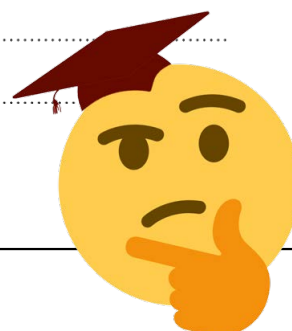
Metal rods produced by a certain factory are claimed to have a median breaking strength of 200 tonnes. For a random sample of 9 rods, the breaking strengths, measured in tonnes, were as follows.

210 186 188 208 184 191 215 198 196

A scientist believes that the median breaking strength of metal rods produced by this factory is less than 200 tonnes.

- (a) Use a Wilcoxon signed-rank test, at the 5% significance level, to test whether there is evidence to support the scientist's belief. [6]

- (b) Give a reason why a Wilcoxon signed-rank test is preferable to a sign test, when both are valid. [1]

[illegible]

4 - (9231/42_Winter_2020_Q2) - Non Parametric Test

A large school is holding an essay competition and each student has submitted an essay. To ensure fairness, each essay is given a mark out of 100 by two different judges. The marks awarded to the essays submitted by a random sample of 12 students are shown in the following table.

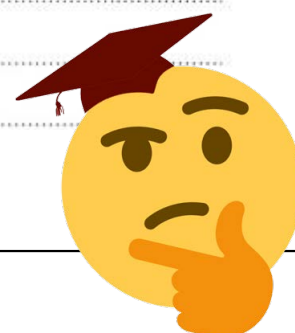
Student	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>	<i>G</i>	<i>H</i>	<i>I</i>	<i>J</i>	<i>K</i>	<i>L</i>
Judge 1	62	74	52	48	68	55	56	64	37	70	81	59
Judge 2	65	70	47	49	76	74	67	54	50	77	72	75

- (a) One of the students claims that Judge 2 is awarding higher marks than Judge 1.

Carry out a Wilcoxon matched-pairs signed-rank test at the 5% significance level to test whether the data supports the student's claim. [7]

It is discovered later that the marks awarded to student A have been entered incorrectly. In fact, Judge 1 awarded 65 marks and Judge 2 awarded 62 marks.

- (b) By considering how this change affects the test statistic, explain why the conclusion of the test carried out in part (a) remains the same. [2]



5 - (9231/43_Summer_2021_Q2) - Non Parametric Test

A company is developing a new flavour of chocolate by varying the quantities of the ingredients. A random selection of 9 flavours of chocolate are judged by two tasters who each give marks out of 100 to each flavour of chocolate.

Chocolate	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>	<i>G</i>	<i>H</i>	<i>I</i>
Taster 1	72	86	75	92	98	79	87	60	62
Taster 2	84	72	74	95	85	87	82	75	68

Carry out a Wilcoxon matched-pairs signed-rank test at the 10% significance level to investigate whether, on average, there is a difference between marks awarded by the two tasters. [7]



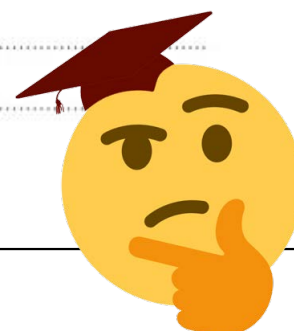
6 - (9231/41_Summer_2021_Q5) - Non Parametric Test

Georgio has designed two new uniforms X and Y for the employees of an airline company. A random sample of 11 employees are each asked to assess each of the two uniforms for practicality and appearance, and to give a total score out of 100. The scores are given in the table.

Employee	A	B	C	D	E	F	G	H	I	J	K
Uniform X	82	74	42	59	60	73	94	98	62	36	50
Uniform Y	78	75	63	56	67	82	99	90	72	48	61

- (a) Give a reason why a Wilcoxon signed-rank test may be more appropriate than a t -test for investigating whether there is any evidence of a preference for one of the uniforms. [1]

- (b) Carry out a Wilcoxon matched-pairs signed-rank test at the 10% significance level. [7]



7 - (9231/42_Winter_2021_Q4) - Non Parametric Test

Applicants for a particular college take a written test when they attend for interview. There are two different written tests, A and B , and each applicant takes one or the other. The interviewer wants to determine whether the medians of the distribution of marks obtained in the two tests are equal. The marks obtained by a random sample of 8 applicants who took test A and a random sample of 8 applicants who took test B are as follows.

Test A	46	32	29	12	33	18	25	40
Test B	36	28	49	37	48	35	41	31

- (a) Carry out a Wilcoxon rank-sum test at the 5% significance level to determine whether there is a difference in the population median marks obtained in the two tests. [6]

The interviewer considers using the given information to carry out a paired sample t -test to determine whether there is a difference in the population means for the two tests.

- (b) Give two reasons why it is not appropriate to use this test. [2]



8 - (9231/41_Winter_2021_Q6) - *Non Parametric Test*

The blood cholesterol levels, measured in suitable units, of a random sample of 11 women and a random sample of 12 men are shown below.

Women	51	55	242	167	152	256	75	137	98	238	235	
Men	311	262	170	302	175	320	220	260	72	351	86	333

Carry out a Wilcoxon rank-sum test, at the 5% significance level, to test whether, on average, there is a difference in cholesterol levels between women and men. [9]



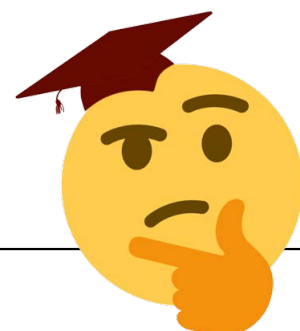
9 - (9231/43_Summer_2022_Q5) - *Non Parametric Test*

A manager claims that the lengths of the rubber tubes that his company produces have a median of 5.50 cm. The lengths, in cm, of a random sample of 11 tubes produced by this company are as follows.

5.56 5.45 5.47 5.58 5.54 5.52 5.60 5.35 5.59 5.51 5.62

It is required to test at the 10% significance level the null hypothesis that the population median length is 5.50 cm against the alternative hypothesis that the population median length is not equal to 5.50 cm.

Show that both a sign test and a Wilcoxon signed-rank test give the same conclusion and state this conclusion. [9]



10 - (9231/41_Summer_2022_Q6) - *Non Parametric Test*

A teacher at a large college gave a mathematical puzzle to all the students. The median time taken by a random sample of 24 students to complete the puzzle was 18.0 minutes. The students were then given practice in solving puzzles. Two weeks later, the students were given another mathematical puzzle of the same type as the first. The times, in minutes, taken by the random sample of 24 students to complete this puzzle are as follows.

18.2	17.5	16.4	15.1	20.5	26.5	19.2	23.2
17.9	18.8	25.8	19.9	17.7	16.2	17.3	16.6
17.1	20.1	20.3	12.6	16.0	21.4	22.7	18.4

The teacher claims that the practice has not made any difference to the average time taken to complete a puzzle of this type.

Carry out a Wilcoxon signed-rank test, at the 10% significance level, to test whether there is sufficient evidence to reject the teacher's claim. [10]

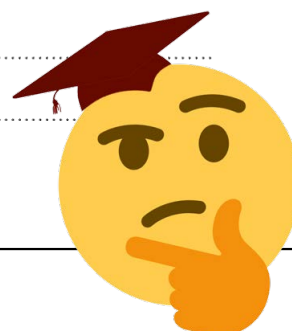


11 - (9231/41_Winter_2022_Q3) - Non Parametric Test

A large college is holding a piano competition. Each student has played a particular piece of music and two judges have each awarded a mark out of 80. The marks awarded to a random sample of 14 students are shown in the following table.

Student	A	B	C	D	E	F	G	H	I	J	K	L	M	N
Judge 1	79	54	63	74	69	52	50	57	55	42	63	55	56	48
Judge 2	75	62	60	73	76	41	31	51	45	55	49	50	65	36

- (a) One of the students claims that on average Judge 1 is awarding higher marks than Judge 2. Carry out a Wilcoxon matched-pairs signed-rank test at the 5% significance level to test whether the data supports the student's claim. [7]

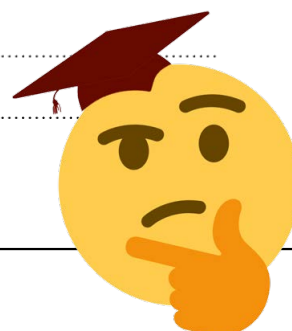


12 - (9231/42_Winter_2022_Q6) - Non Parametric Test

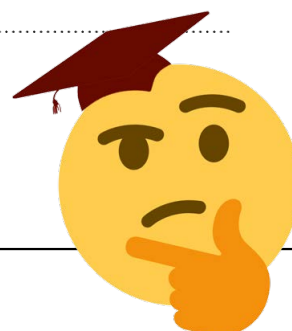
The manager of a technology company A claims that his employees earn more per year than the employees at technology company B . The amounts earned per year, in hundreds of dollars, by a random sample of 12 employees from company A and an independent random sample of 12 employees from company B are shown below.

Company <i>A</i>	461	482	374	512	415	452	502	427	398	545	612	359
Company <i>B</i>	454	506	491	384	361	443	401	472	414	342	355	437

- (a) Carry out a Wilcoxon rank-sum test at the 5% significance level to test whether the manager's claim is supported by the data. [9]

This image shows a blank sheet of white paper with horizontal blue ruling lines. The lines are evenly spaced and run across the width of the page. In the bottom right corner, there is a small, stylized illustration of a yellow graduation cap with a red tassel hanging from it. The cap is tilted slightly to the left.

- (b) Explain whether a paired sample t -test would be appropriate to test the manager's claim if earnings are normally distributed. [1]



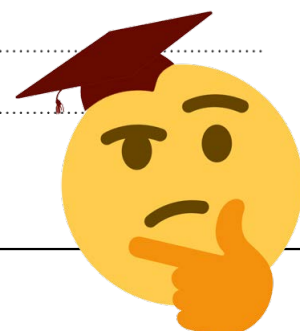
13 - (9231/41_Summer_2023_Q4) - *Non Parametric Test*

A random sample of 13 technology companies is chosen and the numbers of employees in 2018 and in 2022 are recorded.

Company	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>	<i>G</i>	<i>H</i>	<i>I</i>	<i>J</i>	<i>K</i>	<i>L</i>	<i>M</i>
Number in 2018	104	19	126	234	970	514	35	149	429	12	86	304	1104
Number in 2022	106	24	127	228	1012	525	32	156	449	24	78	294	1154

A researcher claims that there has been an increase in the median number of employees at technology companies between 2018 and 2022.

- (a) Carry out a Wilcoxon matched-pairs signed-rank test, at the 5% significance level, to test whether the data supports this claim. [7]

A yellow smiley face with a red graduation cap is located in the bottom right corner of the page. The smiley face is partially cut off by the edge of the page, and the graduation cap is a dark red color with a tassel.

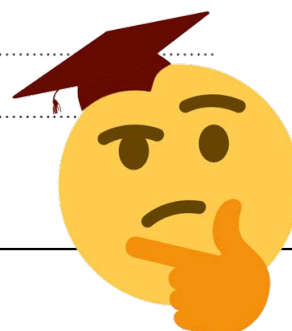
A yellow thinking face emoji with a red graduation cap on its head, pointing its right index finger upwards.

14 - (9231/43_Summer_2023_Q3) - *Non Parametric Test*

A large number of students took two test papers in mathematics. The teacher believes that the marks obtained in Paper 1 will be higher than the marks obtained in Paper 2. She chooses a random sample of 9 students and compares their marks. The marks are shown in the table.

Student	A	B	C	D	E	F	G	H	I
Paper 1	46	73	55	64	86	42	66	68	60
Paper 2	41	66	61	63	90	40	58	42	70

- (a) Carry out a Wilcoxon matched-pairs signed-rank test, at the 5% significance level, to test whether the data supports the teacher's belief. [7]



.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

(b) State an assumption that you have made in carrying out the test in part **(a)**. [1]

.....

.....

.....

.....

.....

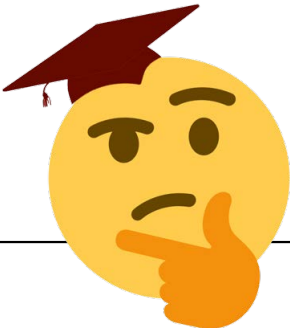
.....

.....

.....

.....

.....



15 - (9231/41_Winter_2023_Q6) - Non Parametric Test

A school is conducting an experiment to see whether the distance that children can throw a ball increases in hot weather. On a cold day, all the children at the school were asked to throw a ball as far as possible. The distances thrown were measured and recorded. The median distance thrown by a random sample of 25 of the children was 22.0m. The children were asked to throw the ball again on a hot day. The distances thrown by the same 25 children were measured and recorded and these distances, in m, are shown below.

21.2	23.5	22.9	18.6	19.4
22.1	26.5	20.2	25.7	20.6
22.3	17.4	22.2	27.0	23.9
28.2	22.6	27.2	23.0	23.7
19.8	22.7	23.3	21.5	24.3

The teacher claims that on average the distances thrown will be further when it is hot.

Carry out a Wilcoxon signed-rank test, at the 5% significance level, to test whether the data supports the teacher's claim. [10]



16 - (9231/42_Winter_2023_Q5) - Non Parametric Test

A company is deciding which of two machines, X and Y , can make a certain type of electrical component more quickly. The times taken, in minutes, to make one component of this type are recorded for a random sample of 8 components made by machine X and a random sample of 9 components made by machine Y . These times are as follows.

Machine X	4.0	4.6	4.7	4.8	5.0	5.2	5.6	5.8	
Machine Y	4.5	4.9	5.1	5.3	5.4	5.7	5.9	6.3	6.4

The manager claims that on average the time taken by machine X to make one component is less than that taken by machine Y .

- (a) Carry out a Wilcoxon rank-sum test at the 5% significance level to test whether the manager's claim is supported by the data. [6]
- (b) Assuming that the times taken to produce the components by the two machines are normally distributed with equal variances, carry out a t -test at the 5% significance level to test whether the manager's claim is supported by the data. [9]
- (c) In general, would you expect the conclusions from the tests in parts (a) and (b) to be the same? Give a reason for your answer. [1]



1 - (9231/41_Summer_2020_Q3) - Continuous Random Variable

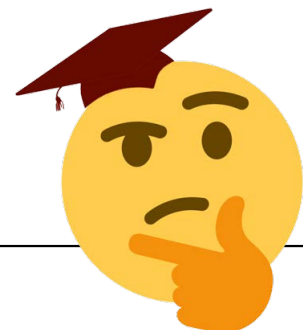
The continuous random variable X has probability density function f given by

$$f(x) = \begin{cases} \frac{3}{16}(2 - \sqrt{x}) & 0 \leq x < 1, \\ \frac{3}{16\sqrt{x}} & 1 \leq x \leq 9, \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find $E(X)$. [3]

The random variable Y is such that $Y = \sqrt{X}$.

(b) Find the probability density function of Y . [5]

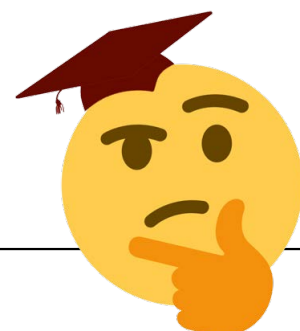


2 - (9231/43_Summer_2020_Q3) - Continuous Random Variable

The continuous random variable X has probability density function f given by

$$f(x) = \begin{cases} \frac{1}{5}x & 0 \leq x < 2, \\ \frac{2}{15}(5-x) & 2 \leq x \leq 5, \\ 0 & \text{otherwise.} \end{cases}$$

- (a) Find the cumulative distribution function of X . [3]
- (b) Find the median value of X . [2]
- (c) Find $E(X^2)$. [2]
- (d) Find $P(1 \leq X \leq 3)$. [2]



3 - (9231/42_Winter_2020_Q4) - Continuous Random Variable

The continuous random variable X has cumulative distribution function F given by

$$F(x) = \begin{cases} 0 & x < 2, \\ \frac{1}{60}x^2 - \frac{1}{15} & 2 \leq x \leq 8, \\ 1 & x > 8. \end{cases}$$

(a) Find $P(3 \leq X \leq 6)$.

[1]

.....

.....

.....

.....

.....

.....

.....

.....

(b) Find $E(\sqrt{X})$.

[3]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

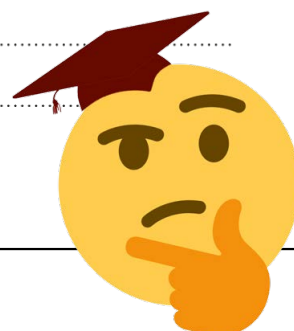
.....

.....

.....

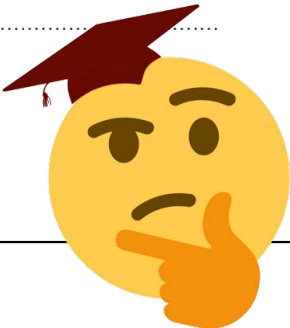
.....

.....



(c) Find $\text{Var}(\sqrt{X})$. [2]

(d) The random variable Y is defined by $Y = X^3$. Find the probability density function of Y . [3]



4 - (9231/41_Winter_2020_Q6) - *Continuous Random Variable*

The continuous random variable X has cumulative distribution function F given by

$$F(x) = \begin{cases} 0 & x < 0, \\ \frac{1}{60}(16x - x^2) & 0 \leq x \leq 6, \\ 1 & x > 6. \end{cases}$$

- (a) Find the interquartile range of X . [4]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

- (b) Find $E(X^3)$. [4]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

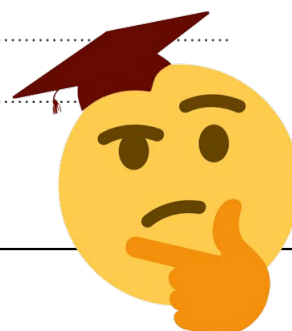
.....

.....

.....

.....

.....



The random variable Y is such that $Y = \sqrt{X}$.

(c) Find the probability density function of Y .

[3]



5 - (9231/41_Summer_2021_Q3) - Continuous Random Variable

The continuous random variable X has cumulative distribution function F given by

$$F(x) = \begin{cases} 0 & x < 0, \\ \frac{1}{81}x^2 & 0 \leq x \leq 9, \\ 1 & x > 9. \end{cases}$$

(a) Find $E(\sqrt{X})$. [3]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

(b) Find $\text{Var}(\sqrt{X})$. [2]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

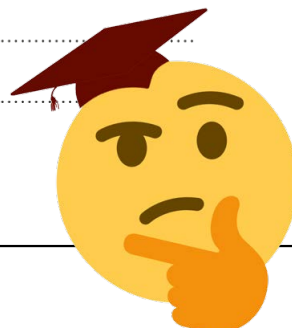
.....

.....

.....

.....

.....



- (c) The random variable Y is given by $Y^3 = X$. Find the probability density function of Y . [3]



A yellow thinking face emoji with a red graduation cap on its head, pointing its right index finger upwards.

The random variable Y is given by $Y = \sqrt[3]{X}$.

(c) Find the probability density function of Y .

[5]



7 - (9231/41_Winter_2021_Q2) - Continuous Random Variable

The continuous random variable X has cumulative distribution function F given by

$$F(x) = \begin{cases} 0 & x < -1, \\ \frac{1}{2}(1+x)^2 & -1 \leq x \leq 0, \\ 1 - \frac{1}{2}(1-x)^2 & 0 < x \leq 1, \\ 1 & x > 1. \end{cases}$$

(a) Find the probability density function of X .

[2]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

(b) Find $P\left(-\frac{1}{2} \leq X \leq \frac{1}{2}\right)$.

[2]

.....

.....

.....

.....

.....

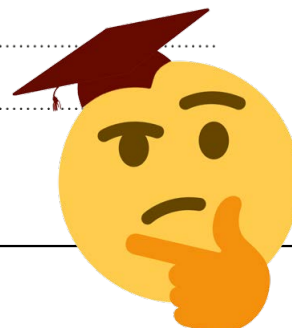
.....

.....

.....

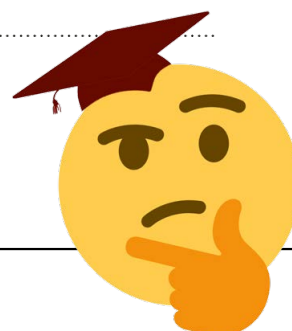
.....

.....



[2]

[2]



8 - (9231/42_Winter_2021_Q3) - Continuous Random Variable

The continuous random variable X has probability density function f given by

$$f(x) = \begin{cases} a + \frac{1}{5}x & 0 \leq x < 1, \\ 2a - \frac{1}{5}x & 1 \leq x \leq 2, \\ 0 & \text{otherwise,} \end{cases}$$

where a is a constant.

(a) Find the value of a .

[3]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

(b) Find $E(X^2)$.

[2]

.....

.....

.....

.....

.....

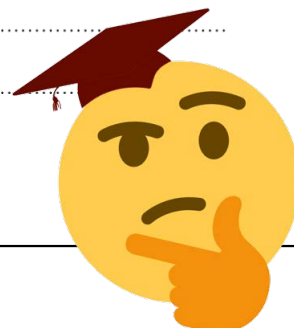
.....

.....

.....

.....

.....



(c) Find the cumulative distribution function of X .

[3]



9 - (9231/41_Summer_2022_Q3) - *Continuous Random Variable*

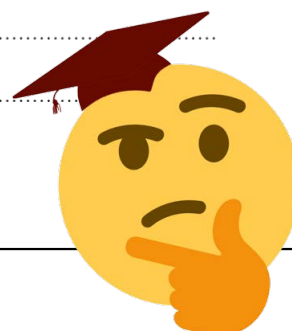
The continuous random variable X has probability density function f given by

$$f(x) = \begin{cases} kx(4-x) & 0 \leq x < 2, \\ k(6-x) & 2 \leq x \leq 6, \\ 0 & \text{otherwise,} \end{cases}$$

where k is a constant.

- (a) Show that $k = \frac{3}{40}$. [1]

- (b) Given that $E(X) = 2.5$, find $\text{Var}(X)$. [3]



(c) Find the median value of X .

[4]



10 - (9231/43_Summer_2022_Q4) - Continuous Random Variable

The continuous random variable X has probability density function f given by

$$f(x) = \begin{cases} \frac{3}{8} \left(1 + \frac{1}{x^2} \right) & 1 \leq x \leq 3, \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find $E(\sqrt{X})$. [3]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

The random variable Y is given by $Y = X^2$.

(b) Find the probability density function of Y . [4]

.....

.....

.....

.....

.....

.....

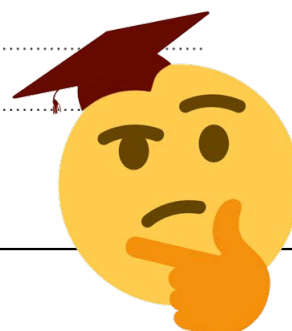
.....

.....

.....

.....

.....



(c) Find the 40th percentile of Y .

[3]



11 - (9231/42_Winter_2022_Q4) - Continuous Random Variable

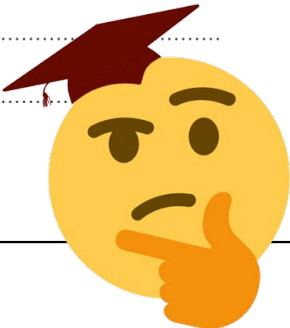
The continuous random variable X has probability density function f given by

$$f(x) = \begin{cases} k & 0 \leq x < 1, \\ kx & 1 \leq x \leq 2, \\ 0 & \text{otherwise,} \end{cases}$$

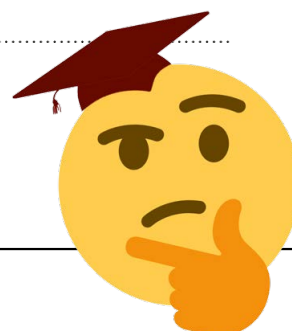
where k is a constant.

(a) Show that $k = \frac{2}{5}$. [1]

(b) Find the interquartile range of X . [5]



[4]



12 - (9231/41_Winter_2022_Q5) - Continuous Random Variable

The continuous random variable X has cumulative distribution function F given by

$$F(x) = \begin{cases} 0 & x < 0, \\ 1 - \frac{1}{144}(12-x)^2 & 0 \leq x \leq 12, \\ 1 & x > 12. \end{cases}$$

- (a) Find the upper quartile of X . [2]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

- (b) Find $\text{Var}(X^2)$. [5]

.....

.....

.....

.....

.....

.....

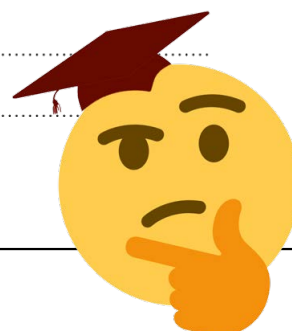
.....

.....

.....

.....

.....



.....

.....

.....

.....

.....

.....

.....

.....

The random variable Y is given by $Y = \sqrt{X}$.

(c) Find the probability density function of Y . [3]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

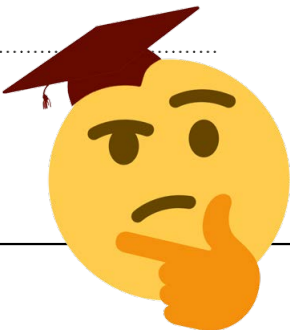
.....

.....

.....

.....

.....



13 - (9231/41_Summer_2023_Q3) - *Continuous Random Variable*

A random sample of 50 values of the continuous random variable X was taken. These values are summarised in the following table.

Interval	$1 \leq x < 1.5$	$1.5 \leq x < 2$	$2 \leq x < 2.5$	$2.5 \leq x < 3$	$3 \leq x < 3.5$	$3.5 \leq x \leq 4$
Observed frequency	3	3	8	11	13	12

It is required to test the goodness of fit of the distribution with probability density function f given by

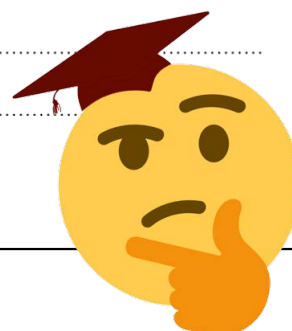
$$f(x) = \begin{cases} \frac{1}{24} \left(\frac{4}{x^2} + x^2 \right) & 1 \leq x \leq 4, \\ 0 & \text{otherwise.} \end{cases}$$

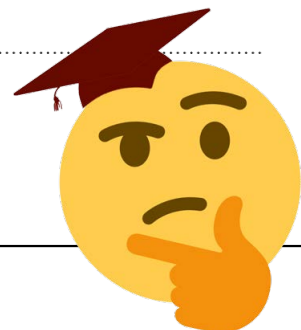
The expected frequencies, correct to 4 decimal places, are given in the following table.

Interval	$1 \leq x < 1.5$	$1.5 \leq x < 2$	$2 \leq x < 2.5$	$2.5 \leq x < 3$	$3 \leq x < 3.5$	$3.5 \leq x \leq 4$
Expected frequency	4.4271	a	6.1285	8.4549	b	14.9678

(a) Show that $a = 4.6007$ and find the value of b .

[3]





14 - (9231/41_Summer_2023_Q6) - Continuous Random Variable

The continuous random variable X has probability density function f given by

$$f(x) = \begin{cases} \frac{3}{28} \left(e^{\frac{1}{2}x} + 4e^{-\frac{1}{2}x} \right) & 0 \leq x \leq 2 \ln 3, \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find the cumulative distribution function of X .

[3]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

The random variable Y is defined by $Y = e^{\frac{1}{2}(X)}$.

(b) Find the probability density function of Y .

[3]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

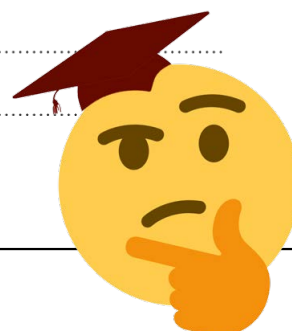
.....

.....

.....

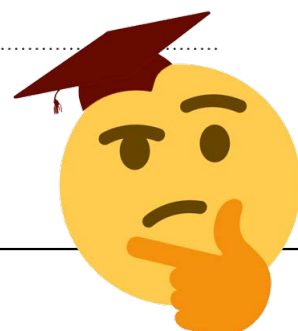
.....

.....



[3]

[2]



15 - (9231/43_Summer_2023_Q1) - Continuous Random Variable

The continuous random variable X has probability density function f given by

$$f(x) = \begin{cases} \frac{1}{6}(x^{-\frac{1}{3}} - x^{-\frac{2}{3}}) & 1 \leq x \leq 27, \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find the cumulative distribution function of X .

[3]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

The random variable Y is defined by $Y = X^{\frac{1}{3}}$.

(b) Find the probability density function of Y .

[3]

.....

.....

.....

.....

.....

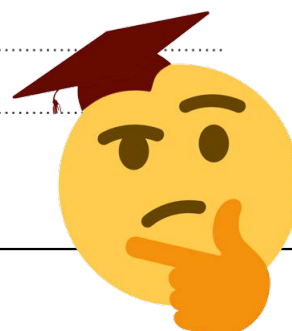
.....

.....

.....

.....

.....



.....

.....

.....

.....

.....

.....

.....

.....

.....

(c) Find the exact value of the median of Y . [2]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

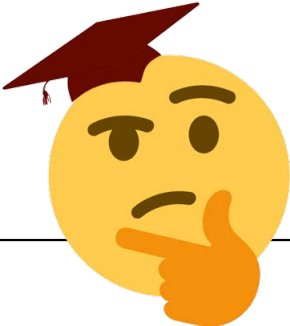
.....

.....

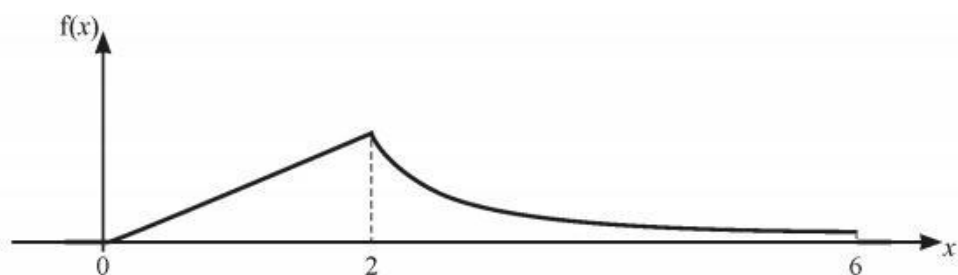
.....

.....

.....



16 - (9231/41_Winter_2023_Q4) - Continuous Random Variable



As shown in the diagram, the continuous random variable X has probability density function f given by

$$f(x) = \begin{cases} mx & 0 \leq x \leq 2, \\ \frac{k}{x^2} + c & 2 \leq x \leq 6, \\ 0 & \text{otherwise,} \end{cases}$$

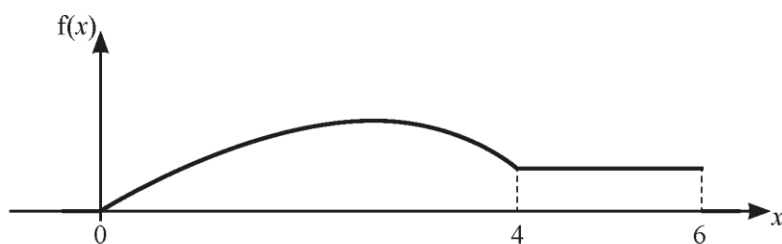
where m , k and c are constants.

(a) Given that $P(X \leq 2) = \frac{1}{3}$, show that $m = \frac{1}{6}$ and find the values of k and c . [4]

(b) Find the exact numerical value of the interquartile range of X . [5]



17 - (9231/42_Winter_2023_Q4) - *Continuous Random Variable*



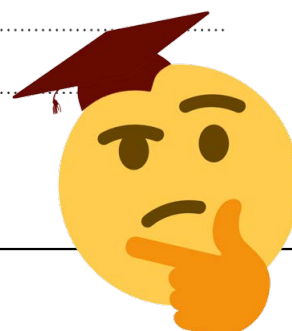
The diagram shows the continuous random variable X with probability density function f given by

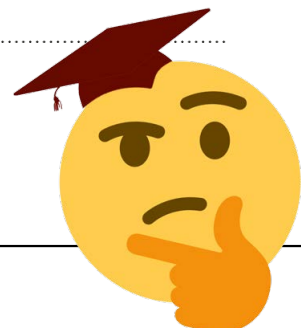
$$f(x) = \begin{cases} \frac{1}{128}(4ax - bx^3) & 0 \leq x \leq 4, \\ c & 4 \leq x \leq 6, \\ 0 & \text{otherwise,} \end{cases}$$

where a , b and c are constants.

The upper quartile of X is equal to 4.

(a) Show that $c = \frac{1}{8}$ and find the values of a and b . [4]





1 - (9231/43_Summer_2020_Q4) - *Probability Generating Function*

The discrete random variable X has probability generating function $G_X(t)$ given by

$$G_X(t) = 0.2t + 0.5t^2 + 0.3t^3.$$

The random variable Y is the sum of two independent observations of X .

(a) Find the probability generating function of Y , giving your answer as an expanded polynomial in t . [3]

(b) Use the probability generating function of Y to find $E(Y)$ and $\text{Var}(Y)$. [5]



2 - (9231/41_Summer_2020_Q6) - Probability Generating Function

A bag contains 4 red balls and 6 blue balls. Rassa selects two balls at random, without replacement, from the bag. The number of red balls selected by Rassa is denoted by X .

(a) Find the probability generating function, $G_X(t)$, of X . [2]

Rassa also tosses two coins. One coin is biased so that the probability of a head is $\frac{2}{3}$. The other coin is biased so that the probability of a head is p . The probability generating function of Y , the number of heads obtained by Rassa, is $G_Y(t)$. The coefficient of t in $G_Y(t)$ is $\frac{7}{12}$.

(b) Find $G_Y(t)$. [3]

The random variable Z is the sum of the number of red balls selected and the number of heads obtained by Rassa.

(c) Find the probability generating function of Z , expressing your answer as a polynomial. [3]

(d) Use the probability generating function of Z to find $E(Z)$. [2]



3 - (9231/41_Winter_2020_Q5) - Probability Generating Function

Keira has two unbiased coins. She tosses both coins. The number of heads obtained by Keira is denoted by X .

- (a) Find the probability generating function $G_X(t)$ of X . [1]

.....

.....

.....

.....

.....

.....

.....

Hassan has three coins, two of which are biased so that the probability of obtaining a head when the coin is tossed is $\frac{1}{3}$. The corresponding probability for the third coin is $\frac{1}{4}$. The number of heads obtained by Hassan when he tosses these three coins is denoted by Y .

- (b) Find the probability generating function $G_Y(t)$ of Y . [3]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

The random variable Z is the total number of heads obtained by Keira and Hassan.

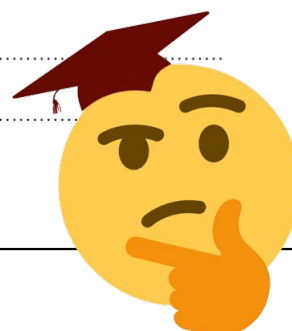
- (c) Find the probability generating function of Z , expressing your answer as a polynomial. [3]

.....

.....

.....

.....



.....

.....

.....

.....

.....

.....

(d) Use the probability generating function of Z to find $E(Z)$. [2]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

(e) Use the probability generating function of Z to find the most probable value of Z . [1]

.....

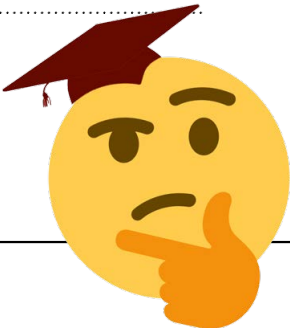
.....

.....

.....

.....

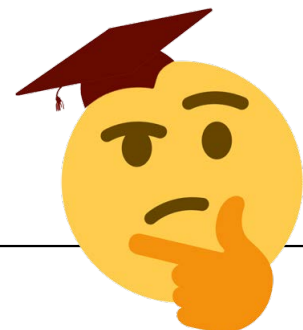
.....



4 - (9231/42_Winter_2020_Q5) - *Probability Generating Function*

The random variable X has the binomial distribution $B(n, p)$.

- (a) Write down an expression for $P(X = r)$ and hence show that the probability generating function of X is $(q + pt)^n$, where $q = 1 - p$. [3]
- (b) Use the probability generating function of X to prove that $E(X) = np$ and $\text{Var}(X) = np(1 - p)$. [5]



5 - (9231/43_Summer_2021_Q4) - Probability Generating Function

X is a discrete random variable which takes the values $0, 2, 4, \dots$. The probability generating function of X is given by

$$G_X(t) = \frac{1}{3 - 2t^2}.$$

(a) Find $E(X)$ and $\text{Var}(X)$. [5]

(b) Find $P(X = 4)$. [3]



6 - (9231/41_Summer_2021_Q6) - *Probability Generating Function*

Tanji has a bag containing 4 red balls and 2 blue balls. He selects 3 balls at random from the bag, without replacement. The number of red balls selected by Tanji is denoted by X .

- (a) Find the probability generating function $G_X(t)$ of X . [2]

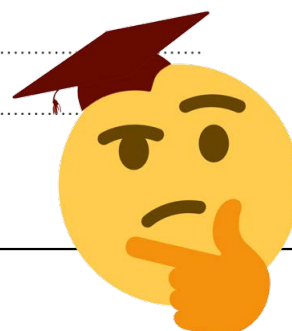
Tanji also has two coins, each biased so that the probability of obtaining a head when it is thrown is $\frac{1}{4}$. He throws the two coins at the same time. The number of heads obtained is denoted by Y .

- (b) Find the probability generating function $G_Y(t)$ of Y . [2]

The random variable Z is the sum of the number of red balls selected by Tanji and the number of heads obtained.

- (c) Find the probability generating function of Z , expressing your answer as a polynomial. [3]





(d) Use the probability generating function of Z to find $E(Z)$ and $\text{Var}(Z)$.

[5]



7 - (9231/41_Winter_2021_Q5) - Probability Generating Function

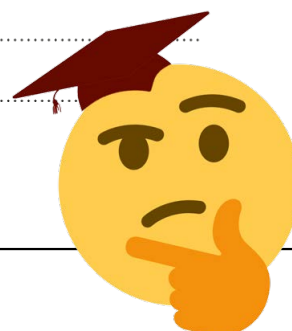
Nine balls labelled 1, 2, 3, 4, 5, 6, 7, 8, 9 are placed in a bag. Kai selects three balls at random from the bag, without replacement. The random variable X is the number of balls selected by Kai that are labelled with a multiple of 3.

- (a) Find the probability generating function $G_X(t)$ of X . [3]

The balls are replaced in the bag.

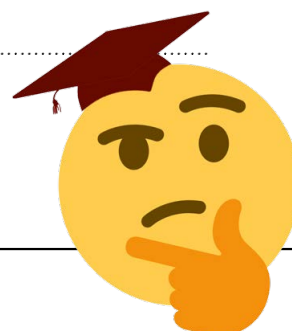
Jacob now selects two balls at random from the bag, without replacement. The random variable Y is the number of balls selected by Jacob that are labelled with an even number.

- (b) Find the probability generating function $G_Y(t)$ of Y . [2]



(c) Find the probability generating function of Z , expressing your answer as a polynomial. [3]

(d) Use the probability generating function of Z to find $E(Z)$. [2]



8 - (9231/42_Winter_2021_Q5) - Probability Generating Function

The random variable X is such that $P(X = r) = kr^2$ for $r = 1, 2, 3, 4$, where k is a constant.

- (a) Find the value of k . [1]

.....

.....

.....

.....

.....

.....

- (b) Find the probability generating function $G_X(t)$ of X . [2]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

The random variable Y has probability generating function $G_Y(t) = \frac{1}{4} + \frac{1}{2}t + \frac{1}{4}t^2$.

The random variable Z is the sum of X and Y .

- (c) Assuming that X and Y are independent, find the probability generating function $G_Z(t)$ of Z as a polynomial in t . [3]

.....

.....

.....

.....

.....

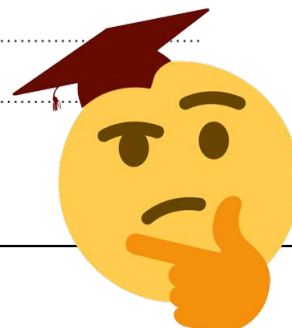
.....

.....

.....

.....

.....



(d) Given that $E(Z) = \frac{13}{3}$, use $G_Z(t)$ to find $\text{Var}(Z)$.

[3]



9 - (9231/41_Summer_2022_Q2) - *Probability Generating Function*

The probability generating function, $G_Y(t)$, of the random variable Y is given by

$$G_Y(t) = 0.04 + 0.2t + 0.37t^2 + 0.3t^3 + 0.09t^4.$$

- (a) Find $\text{Var}(Y)$. [4]

The random variable Y is the sum of two independent observations of the random variable X .

- (b) Find the probability generating function of X , giving your answer as a polynomial in t . [3]



10 - (9231/43_Summer_2022_Q2) - *Probability Generating Function*

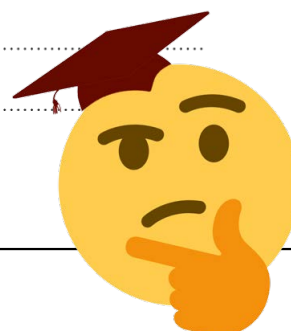
A scientist is investigating the size of shells at various beach locations. She selects four beach locations and takes a random sample of shells from each of these beaches. She classifies each shell as large or small. Her results are summarised in the following table.

		Beach location				Total
		<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	
Size of shell	Large	68	69	96	81	314
	Small	28	55	64	39	186
	Total	96	124	160	120	500

Test, at the 10% significance level, whether the size of shell is independent of the beach location. [7]

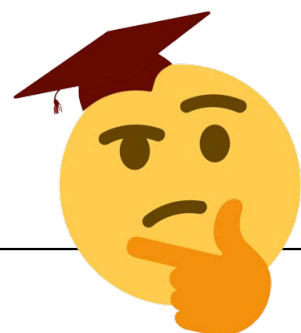


George throws two coins, A and B , at the same time. Coin A is biased so that the probability of obtaining a head is a . Coin B is biased so that the probability of obtaining a head is b , where $b < a$. The probability generating function of X , the number of heads obtained by George, is $G_X(t)$. The coefficients of t and t^2 in $G_X(t)$ are $\frac{5}{12}$ and $\frac{1}{12}$ respectively.



(c) Find $\text{Var}(Y)$.

[3]



12 - (9231/41_Summer_2022_Q4) - Probability Generating Function

A scientist is investigating the numbers of a particular type of butterfly in a certain region. He claims that the numbers of these butterflies found per square metre can be modelled by a Poisson distribution with mean 2.5. He takes a random sample of 120 areas, each of one square metre, and counts the number of these butterflies in each of these areas. The following table shows the observed frequencies together with some of the expected frequencies using the scientist's Poisson distribution.

Number per square metre	0	1	2	3	4	5	6	≥ 7
Observed frequency	12	20	36	32	13	6	1	0
Expected frequency	9.85	24.63	30.78	25.65	p	8.02	3.34	q

- (a) Find the values of p and q , correct to 2 decimal places. [2]

.....

.....

.....

.....

.....

.....

.....

.....

.....

- (b) Carry out a goodness of fit test, at the 10% significance level, to test the scientist's claim. [6]

.....

.....

.....

.....

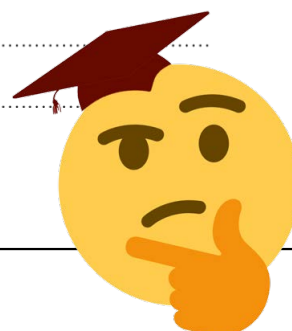
.....

.....

.....

.....

.....



13 - (9231/41_Winter_2022_Q4) - Probability Generating Function

Jason has three biased coins. For each coin the probability of obtaining a head when it is thrown is $\frac{2}{3}$. Jason throws all three coins. The number of heads obtained is denoted by X .

- (a) Find the probability generating function $G_X(t)$ of X . [3]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

Jason also has two unbiased coins. He throws all five coins. The number of heads obtained from the two unbiased coins is denoted by Y . It is given that $G_Y(t) = \frac{1}{4} + \frac{1}{2}t + \frac{1}{4}t^2$. The random variable Z is the total number of heads obtained when Jason throws all five coins.

- (b) Find the probability generating function of Z , expressing your answer as a polynomial. [3]

.....

.....

.....

.....

.....

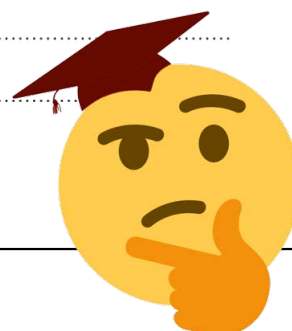
.....

.....

.....

.....

.....



.....

.....

.....

.....

.....

.....

.....

.....

.....

(c) Find $E(Z)$.

[2]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

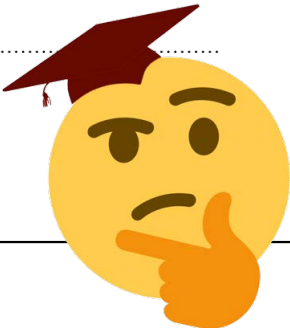
.....

.....

.....

.....

.....



14 - (9231/42_Winter_2022_Q5) - Probability Generating Function

A 6-sided dice, A , with faces numbered 1, 2, 3, 4, 5, 6 is biased so that the probability of throwing a 6 is $\frac{1}{4}$. The random variable X is the number of 6s obtained when dice A is thrown twice.

(a) Find the probability generating function of X . [2]

.....

.....

.....

.....

.....

.....

.....

.....

.....

A second dice, B , with faces numbered 1, 2, 3, 4, 5, 6 is unbiased. The random variable Y is the number of 6s obtained when dice B is thrown twice.

The random variable Z is the total number of 6s obtained when both dice are thrown twice.

(b) Find the probability generating function of Z , expressing your answer as a polynomial. [3]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

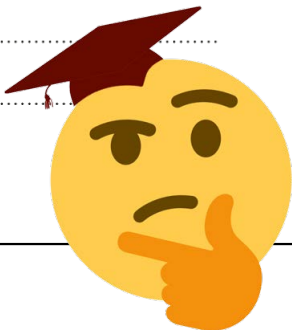
.....

.....

.....

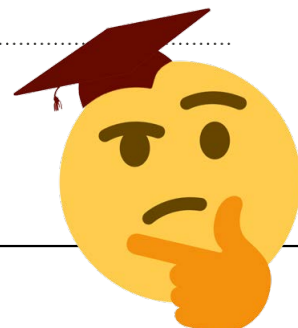
.....

.....



[3]

[1]



15 - (9231/41_Summer_2023_Q5) - *Probability Generating Function*

Harry has three coins.

- One coin is biased so that, when it is thrown, the probability of obtaining a head is $\frac{1}{3}$.
- The second coin is biased so that, when it is thrown, the probability of obtaining a head is $\frac{1}{4}$.
- The third coin is biased so that, when it is thrown, the probability of obtaining a head is $\frac{1}{5}$.

The random variable X is the number of heads that Harry obtains when he throws all three coins together.

- (a) Find the probability generating function of X . [3]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

Isaac has two fair coins. The random variable Y is the number of heads that Isaac obtains when he throws both of his coins together. The random variable Z is the total number of heads obtained when Harry throws his three coins and Isaac throws his two coins.

- (b) Find the probability generating function of Z , expressing your answer as a polynomial in t . [4]

.....

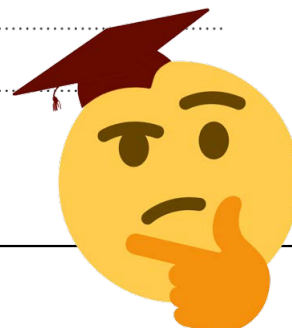
.....

.....

.....

.....

.....



.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

(c) Use the probability generating function of Z to find $E(Z)$. [2]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

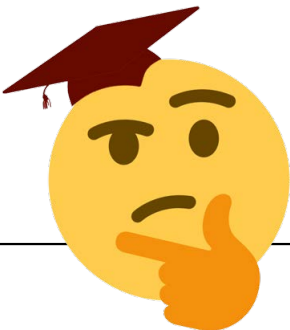
.....

.....

.....

.....

.....



16 - (9231/43_Summer_2023_Q5) - Probability Generating Function

The random variable X has probability generating function $G_X(t)$ given by

$$G_X(t) = k(1 + 3t + 4t^2),$$

where k is a constant.

- (a) Show that $E(X) = \frac{11}{8}$. [3]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

The random variable Y has probability generating function $G_Y(t)$ given by

$$G_Y(t) = \frac{1}{3}t^2(1 + 2t).$$

The random variables X and Y are independent and $Z = X + Y$.

- (b) Find the probability generating function of Z , expressing your answer as a polynomial in t . [2]

.....

.....

.....

.....

.....

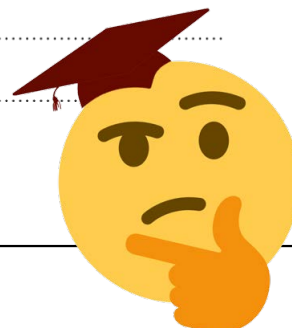
.....

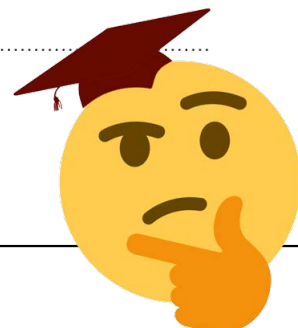
.....

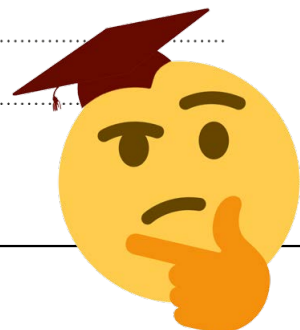
.....

.....

.....

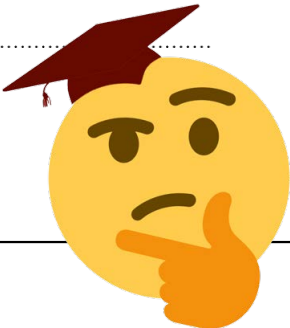






Kenny throws an ordinary fair 6-sided dice repeatedly. The random variable X is the number of throws that Kenny takes in order to obtain a 6. The random variable Z denotes the sum of two independent values of X .

(c) Find the probability generating function of Z . [2]



18 - (9231/42_Winter_2023_Q3) - Probability Generating Function

Toby has a bag which contains 6 red marbles and 3 green marbles. He randomly chooses 3 marbles from the bag, without replacement. The random variable X is the number of red marbles that Toby obtains.

(a) Find the probability generating function of X . [3]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

Ling also has a bag which contains 6 red marbles and 3 green marbles. He randomly chooses 2 marbles from his bag, without replacement. The random variable Y is the number of red marbles that Ling obtains. It is given that the probability generating function of Y is $\frac{1}{12}(1 + 6t + 5t^2)$.

The random variable Z is the total number of red marbles that Toby and Ling obtain.

(b) Find the probability generating function of Z , expressing your answer as a polynomial in t . [3]

.....

.....

.....

.....

.....

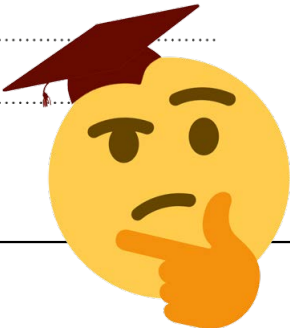
.....

.....

.....

.....

.....



.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

(c) Use the probability generating function of Z to find $\text{Var}(Z)$. [4]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

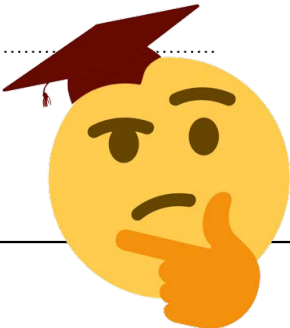
.....

.....

.....

.....

.....



ANSWERS



1 - (9231/21_Summer_2017_Q6) - Further Work On Distributions

(i)	$P(X \leq 4) = 1 - q^4$	M1
	$= 671/1296 \text{ or } 0.518$	A1
	Total:	2
(ii)	$1 - q^{N-1} > 0.95$	M1
	$(5/6)^{N-1} < 0.05, N-1 > \log 0.05 / \log 5/6$	M1
	$N-1 > 16.4[3], N_{\min} = 18$	A1
	Total:	3

2 - (9231/21_Summer_2017_Q8) - Further Work On Distributions

(i)	$F(x) = \int f(x) dx = x^2/8 - x/4 [+ c]$	M1
	$= x^2/8 - x/4 \text{ or } \{(x-1)^2 - 1\}/8$ (AEF)	A1
	$F(x) = 0 (x < 2), F(x) = 1 (x > 4)$	A1
	Total:	3
(ii)	EITHER: $G(y) = P(Y < y) = P((X-1)^3 < y)$ $= P(X < 1 + y^{1/3}) = F(1 + y^{1/3})$ $= (1 + y^{1/3})^2/8 - (1 + y^{1/3})/4 \text{ or } (y^{2/3} - 1)/8$	(M1 A1)
	OR: Use $x = 1 + y^{1/3}$ to find $f(x) = 1/4 y^{1/3}$ and $dx/dy = 1/3 y^{-2/3}$	(M1 A1)
	$g(y) [= G'(y)] = (1/12) y^{-1/3} \text{ or } 1 / (12 y^{1/3})$	A1
	for $1 \leq y \leq 27$ [$g(y) = 0$ otherwise]	A1
	Total:	4
(iii)	$(m^{2/3} - 1)/8 = 1/2$	M1
	$m^{2/3} = 5, m = \sqrt{125} \text{ or } 5\sqrt{5} \text{ or } 11.2$	M1 A1
	Total:	3



3 - (9231/23_Summer_2017_Q9) - Further Work On Distributions

(i)	$(a / \ln 2) [-e^{-x \ln 2}]_0^\infty = a / \ln 2$ so $a = \ln 2$ or 0.693	M1 A1
	Total:	2
(ii)	$E(X) = 1 / \ln 2$ or 1.44	B1
	Total:	1
(iii)	$F(Q) = 1 - e^{-Q \ln 2} = 1/4$ or $3/4$	M1
	$Q_1 = (\ln 4/3) / (\ln 2) [= 0.415]$ (AEF)	A1
	$Q_3 = (\ln 4) / (\ln 2) [= 2]$ (AEF)	A1
	$Q_3 - Q_1 [= (\ln 3) / (\ln 2)] = 1.58$ [or 1.59]	A1
	Total:	4
(iv)	EITHER: $G(y) = P(Y < y) = P(2^X < y)$ $= P(X < (\ln y) / (\ln 2))$ $= F((\ln y) / (\ln 2))$ or $F(\log_2 y)$ (AEF)	(M1 A1)
	$= 1 - e^{-\ln y}$ or $1 - 1/y$	A1)
	OR: Use $x = (\ln y) / (\ln 2)$ to find both	(M1
	$f(x) = (\ln 2) e^{-x \ln 2} = (\ln 2) e^{-\ln y} = (1/y) \ln 2$	A1
	and $dx/dy = 1 / (y \ln 2)$	A1)
	$g(y) [= G'(y)] = 1/y^2$	A1
	for $y \geq 1$ [$g(y) = 0$ otherwise]	A1
	Total:	5

4 - (9231/21_Winter_2017_Q6) - Further Work On Distributions

(i)	$p = (1/6)^2$ or $1/36$	B1	Find (or imply) probability p of pair of sixes in one throw
	$1/p = 36$	B1	Find mean value of X
(ii)	$P(X = 12) = p(1-p)^{11} = 0.0204$	M1 A1	Find prob. of needing exactly 12 throws
(iii)	$P(X > 12) = (1-p)^{12} = 0.713$	M1 A1	Find prob. of needing more than 12 throws
		2	



5 - (9231/21_Winter_2017_Q7) - Further Work On Distributions

(i)	$F(x) = \int f(x) dx = -e^{-0.2x} + c = 1 - e^{-0.2x} \quad (x \geq 0)$	M1	State, or integrate and use $F(0) = 0$ or $F(x) \rightarrow 1$ as $x \rightarrow \infty$
	and $F(x) = 0 \quad (x < 0 \text{ or otherwise})$	A1	to find, $F(x)$ (A0 if case $x < 0$ omitted)
		2	
(ii)	$P(X > 2) = 1 - F(2) = e^{-0.4} = 0.670$	M1 A1	Find $P(X > 2)$: (M0 for $F(2)$)
		3	
(iii)	$1 - e^{-0.2m} = \frac{1}{2}, e^{-0.2m} = \frac{1}{2}$	M1	Find median value m from $F(m)$ or $1 - F(m) = \frac{1}{2}$
	$m = 5 \ln 2$ or 3.47	M1 A1	
		3	

6 - (9231/21_Summer_2018_Q6) - Further Work On Distributions

(i)	$P(X > 2) = 1 - F(2) = \exp(-0.8) = 0.449$	M1 A1	Find $P(X > 2)$. M0 for $F(2)$
		2	
(ii)	$F(Q) = 1 - \exp(-0.4Q) = \frac{1}{4}$ or $\frac{3}{4}$	M1	Formulate equation for either quartile value Q
	$Q_1 = (\ln 4/3) / (0.4) \quad [= 0.7192]$ (AEF)	A1	Find one [lower] quartile Q_1
	$Q_3 = (\ln 4) / (0.4) \quad [= 3.466]$ (AEF)	A1	Find other [upper] quartile Q_3
	$Q_3 - Q_1 = (\ln 3) / (0.4) = 2.75$	A1✓	Find interquartile range (FT on Q_1, Q_3 ; allow $Q_3 - Q_1$)
		4	

7 - (9231/23_Summer_2018_Q7) - Further Work On Distributions

(i)	$(1-p)/p^2 = 3.75, 15p^2 + 4p - 4 = 0$	AG	M1 A1	Find given eqn. for p using $\text{Var}(X) = (1-p)/p^2$
	$(5p-2)(3p+2) = 0, p = 2/5$ or 0.4		M1 A1	Solve quadratic for p (A0 if $p = -2/3$ not [implicitly] rejected)
			4	
(ii)	$P(X=5) = (1-p)^4 p = 0.6^4 \times 0.4 = 0.0518$ or 162/3125		B1	Find $P(X=5)$
			1	
(iii)	EITHER: $P(3 \leq X \leq 7) = (1-p)^2 - (1-p)^7$		M1	Find $P(3 \leq X \leq 7)$
	$= 0.6^2 - 0.6^7 = 0.36 - 0.028 = 0.332$		A1	M0 for $P(X \leq 7) - P(X \leq 3) [= 0.188]$ or similar error
	OR: $P(3 \leq X \leq 7) = \sum_{i=3}^7 (1-p)^{i-1} p$		(M1)	
	$= (0.6^2 + 0.6^3 + 0.6^4 + 0.6^5 + 0.6^6) \times 0.4$		(A1)	
	$= 0.830016 \times 0.4 = 0.332$			
			2	

8 - (9231/21_Summer_2018_Q9) - Further Work On Distributions

(i)	$(1-p)/p^2 = 4/9, 4p^2 + 9p - 9 = 0$	AG	M1 A1	Find given eqn. for p using $\text{Var}(X) = (1-p)/p^2$
	$(4p-3)(p+3) = 0, p = \frac{3}{4}$		M1 A1	Solve quadratic for p (A0 if $p = -3$ not [implicitly] rejected)
			4	
(ii)	$P(X=3) = (1-p)^2 p = (\frac{1}{4})^2 \frac{3}{4} = 3/64$ or 0.0469		B1	Find $P(X=3)$
			1	



(iii)	$P(X \geq 4) = (1-p)^3 = (\frac{1}{4})^3$ or $1 - (1 + \frac{1}{4} + (\frac{1}{4})^2) \frac{1}{4} = 1/64$ or 0.0156	M1 A1	Find $P(X \geq 4)$
		2	
(iv)	$1 - (1-p)^N > 0.999$	M1	Formulate condition for N ($1 - (1-p)^{N-1}$ is M0)
	$0.001 > (\frac{1}{4})^N$	A1	(< or = can earn M1 M1 only, max 2/4)
	$N > \log 0.001 / \log 0.25$	M1	Rearrange and take logs (any base) to give bound
	$N > 4.98, N_{\min} = 5$	A1	Find $N_{\min}$
		4	

9 - (9231/23_Summer_2018_Q9) - Further Work On Distributions

(i)	<i>EITHER:</i> $F(x) = \int f(x) dx = (1/20)(3x - 2\sqrt{x} + c)$ $c = -1$ so $F(x) = (1/20)(3x - 2\sqrt{x} - 1)$	M1	Find or state distribution function $F(x)$ for $1 \leq x \leq 9$ (may be implied by $G(y)$)
	or $(3/20)x - (1/10)\sqrt{x} - 1/20$ $G(y) [= P(Y < y) = P(\sqrt{X} < y) = P(X < y^2)]$	A1	Find or state $G(y)$ from $Y = \sqrt{X}$ for $1 \leq x \leq 9$ or $1 \leq y \leq 3$
	$= F(y^2)$ $= (1/20)(3y^2 - 2y - 1)$	M1	Allow A1 $\sqrt{\quad}$ as FT on expression found for $F(x)$
	or $(3/20)y^2 - (1/10)y - 1/20$	A2	Verify $g(y)$ (differentiation may be implied)
	$g(y) = G'(y) = (1/10)(3y - 1)$ [for $1 \leq y \leq 3, g(y) = 0$ otherwise]	AG M1 A1	SC Missing/incorrect c can earn M1 M1 A1 $\sqrt{\quad}$ M1 (max 4/7)
		7	
	OR: Use of $g(y) = f(x) \times dx/dy $	(*M1)	Reference to standard result required (not in syllabus)
	$f(x) = (1/20)(3 - 1/y)$ (dep *M1)	(M1 A1)	Find $f(x)$ using $x = y^2$
	$dx/dy = 2y$ (dep *M1)	(M1 A1)	Find dx/dy using $x = y^2$
	$g(y) = f(x) \times dx/dy = (1/10)(3y - 1)$ [for $1 \leq y \leq 3, g(y) = 0$ otherwise]	AG (M1 A1)	
		(7)	
(ii)	$E(Y) = (1/10) \int (3y^2 - y) dy$	M1	Find mean of Y from $\int y g(y) dy$
	$= (1/10) [y^3 - \frac{1}{2}y^2]_1^3 = 11/5$ or 2.2	A1	
		2	

10 - (9231/21_Winter_2018_Q6) - Further Work On Distributions

(i)	$F(x) = \int f(x) dx = (1/80)(2x^{3/2} - 16x^{1/2}) + c$	M1	Find or state distribution function $F(x)$ for $4 \leq x \leq 16$ using $F(4) = 0$ or $F(16) = 1$ to find c if necessary
	$= (1/80)(2x^{3/2} - 16x^{1/2} + 16)$ or $(1/40)(x^{3/2} - 8x^{1/2} + 8)$ or $x^{3/2}/40 - x^{1/2}/5 + 1/5$ (AEF)	A1	State $F(x)$ for other values of x
	$F(x) = 0$ ($x < 4$), $F(x) = 1$ ($x > 16$)	A1	
		3	
(ii)	<i>EITHER:</i> $G(y) [= P(Y < y) = P(\sqrt{X} < y) = P(X < y^2)]$ $= F(y^2) = (1/40)(y^3 - 8y + 8)$ (AEF)	M1A1	Find or state $G(y)$ for $2 \leq y \leq 4$ from $Y = \sqrt{X}$ (allow $<$ or $\leq$ throughout; FT on constant term)
	OR: Use $x = y^2$ to find $f(x) = (1/80)(3y - 8/y)$ and $dx/dy = 2y$	(M1A1)	Find $f(x)$ and dx/dy for use in $g(y) = f(x) \times dx/dy $
	$g(y) [= G'(y)] = (1/40)(3y^2 - 8)$ or $(3/40)y^2 - 1/5$ [for $2 \leq y \leq 4, g(y) = 0$ otherwise]	A1	Find $g(y)$ in simplified form for $2 \leq y \leq 4$
		3	



11 - (9231/21_Winter_2018_Q7) - Further Work On Distributions

(i)	$f(t) = (1/500) \exp(-t/500)$ [0 otherwise] (AEF)	B1	State pdf of T for $t \geq 0$
		1	
(ii)	$F(t) = 1 - \exp(-t/500)$	M1	Find or imply $F(t)$
	$P(T > 750) = 1 - F(750) = \exp(-750/500) = 0.223$	M1A1	Find $P(T > 750)$ (M0 for $F(750)$)
		3	
(iii)	$1 - \exp(-m/500) = \frac{1}{2}$ or $\exp(-m/500) = \frac{1}{2}$	M1	Find median value m of T from $F(m) = \frac{1}{2}$
	$m = 500 \ln 2 = 347$ [hours]	M1A1	
		3	

12 - (9231/22_Winter_2018_Q7) - Further Work On Distributions

(i)	<i>EITHER:</i> $G(y) [= P(Y < y) = P(X^2 < y)]$ $= P(X < y^{1/2}) = F(y^{1/2}) = (1/90)(y + y^2)$	M1A1	Find or state $G(y)$ for $0 \leq x \leq 3$ from $Y = X^2$ (allow $<$ or $\leq$ throughout)
	<i>OR:</i> Use $x = y^{1/2}$ to find $f(x) = (1/90)(2x + 4x^3) = (1/90)(2y^{1/2} + 4y^{3/2})$ and $dx/dy = 1 / 2y^{1/2}$	(M1A1)	Find $f(x)$ and dx/dy for use in $g(y) = f(x) \times dx/dy $
	$g(y) [= G'(y)] = (1/90)(1 + 2y)$	A1	Find $g(y)$ in simplified form
	for $0 \leq y \leq 9$ [$g(y) = 0$ otherwise]	A1	State corresponding range of y at any stage
		4	
(ii)	$E(Y) = (1/90) \int (y + 2y^2) dy$	M1	Find mean of Y from $\int y g(y) dy$
	$= (1/90) [\frac{1}{2}y^2 + \frac{2}{3}y^3]_0^9 = 117/20$ or 5.85	A1	
		2	

13 - (9231/23_Winter_2018_Q7) - Further Work On Distributions

(i)	$f(t) = (1/500) \exp(-t/500)$ [0 otherwise] (AEF)	B1	State pdf of T for $t \geq 0$
		1	
(ii)	$F(t) = 1 - \exp(-t/500)$	M1	Find or imply $F(t)$
	$P(T > 750) = 1 - F(750) = \exp(-750/500) = 0.223$	M1A1	Find $P(T > 750)$ (M0 for $F(750)$)
		3	
(iii)	$1 - \exp(-m/500) = \frac{1}{2}$ or $\exp(-m/500) = \frac{1}{2}$	M1	Find median value m of T from $F(m) = \frac{1}{2}$
	$m = 500 \ln 2 = 347$ [hours]	M1A1	
		3	

14 - (9231/22_Winter_2018_Q8) - Further Work On Distributions

(i)	$P(X \geq 5) = (1-p)^4 = 0.4096$ $p = 1 - 0.8 = 0.2$ AG	M1A1	Verify p using $P(X \geq 5) = 0.4096$
		2	
(ii)	$P(X = 6) = (1-p)^5 p = 0.8^5 \times 0.2 = 0.0655$	M1A1	Find $P(X = 6)$
		2	
(iii)	$1 - (1-p)^N > 0.9$	M1	Formulate condition for N ($1 - (1-p)^N$ is M0)
	$0.1 > 0.8^N$	A1	(< or = can earn M1 M1 only, max 2/4)
	$N > \log 0.1 / \log 0.8 = 10.3$	M1	Rearrange and take logs (any base)
	$N_{\min} = 11$ corresponding to Monday of 3rd week	A1	Find $N_{\min}$ and corresponding day and week
		4	



15 - (9231/21_Summer_2019_Q6) - Further Work On Distributions

(i)	Mean = 3	B1	State mean of X
			1
(ii)	$P(X=6) = q^5 p$ with $p = 1/3, q = 2/3$ $= 32/729$ or 0.0439 (AEF)	B1	Find probability of score of 3 or 4 on exactly 6 throws
			1
(iii)	$P(X > 4) = q^4 = 16/81$ or 0.1975 or 0.198	M1 A1	Find probability of score of 3 or 4 on more than 4 throws
			2
(iv)	$1 - q^{n-1} < 0.95$ (AEF)	M1	Formulate condition for n ($1 - q^n$ is M0)
			M1 Set $q = 2/3$, rearrange and take logs (any base) to give bound
	$0.05 < (2/3)^{n-1}, n-1 < \log 0.05 / \log 2/3$	M1	
	$n-1 < 7.39, n_{\max} = 8$	A1	Find $n_{\max}$ ($>$ or $=$ can earn M1 M1 A0, max 2/3)
		3	

16 - (9231/23_Summer_2019_Q6) - Further Work On Distributions

(i)	$f(t) = (1/400) \exp(-t/400)$ or $0.0025 \exp(-0.0025 t)$ [$= 0$ ($t < 0$)]	B1	State probability density function $f(t)$ for $t \geq 0$
			1
(ii)	$P(T < 500) = \int_0^{500} f(t) dt = [\exp(-t/400)]_0^{500}$ $= 1 - e^{-5/4} = 0.7135$ or 0.713 (allow 0.714)	M1 A1	Find $P(T < 500)$: (M0 for $1 - F(500) = 0.287$)
			2
(iii)	$1 - e^{-m/400}$ or $e^{-m/400} = 1/2, e^{m/400} = 2$	M1	Find median value m from $F(m) = 1/2$
			M1 A1
	$m = 400 \ln 2$ or 277	3	

17 - (9231/21_Summer_2019_Q7) - Further Work On Distributions

(i)	$F(x) = \int f(x) dx = -3/(4x) + x/4$ [$+ c$]	M1	Find or state distribution function $F(x)$ for $1 \leq x \leq 3$
			A1 using $F(1) = 0$ or $F(3) = 1$ to find c if necessary
			A1 State $F(x)$ for other values of x
			3
(ii)	$\int_1^Q f(x) dx = -3/4Q + Q/4 + 1/2 = 1/4$ [or $3/4$] (AEF)	M1	Formulate equation for either quartile value Q
			A1
	$Q^2 + [or -] Q - 3 = 0$	A1 A1	Find lower quartile Q_1 and upper quartile Q_3
	$Q_1 = 1/2(-1 + \sqrt{13}), Q_3 = 1/2(1 + \sqrt{13})$		Find interquartile range
	$Q_3 - Q_1 = 1$	5	



18 - (9231/23_Summer_2019_Q7) - Further Work On Distributions

(i)	$E(X) = 1 / (\frac{1}{2} \times \frac{1}{2}) = 4$	B1	Find or state $E(X)$
		1	
(ii)	$P(X=3) = q^2 p$ (with $p = \frac{1}{4}$, $q = \frac{3}{4}$) = $9/64$ or 0.141	M1 A1	Find prob. of exactly 3 throws needed
		2	
(iii)	$P(X < 4) = 1 - q^3 = 37/64$ or 0.578	M1 A1	Find prob. of fewer than 4 throws needed
		2	
(iv)	$1 - q^{N-1} > 0.95$ (AEF)	M1	Formulate condition for N ($1 - q^N$ is M0)
	$0.05 > (3/4)^{N-1}$, $N-1 > \log 0.05 / \log 0.75$	M1	Set $q = 3/4$, rearrange and take logs (any base) to give bound
	$N-1 > 10.4$, $N_{\min} = 12$	A1	Find $N_{\min}$ (< or = can earn M1 M1 A0, max 2/3)
		3	

19 - (9231/21_Winter_2019_Q7) - Further Work On Distributions

(i)	$a = 1/200$ or 0.005	B1	State a or find a by equating mean value to $1/a$
		1	
(ii)	$p = P(T < 150) = F(150) = 1 - e^{-150/a}$	M1	Find $P(T < 150)$
	$p = 1 - e^{-0.75} = 0.528$	A1	
		2	
(iii)	$1 - p^n > 0.99$	M1	Formulate condition for n
	$0.01 > (1 - e^{-0.75})^n$ or $0.01 > 0.528^n$	A1	
	$n > \log 0.01 / \log 0.528$	M1	Rearrange and take logs to give bound
	$n > 7.20$ [or 7.21] so $n_{\min} = 8$	A1	Find $n_{\min}$
		4	



20 - (9231/21_Winter_2019_Q10) - Further Work On Distributions

(i)	$F(x) = \int f(x) dx = (1/30) (-8/x + x^3 - 14x) [+c]$	M1	Find or state distribution function $F(x)$ for $2 \leq x \leq 4$
	$F(x) = (1/30) (-8/x + x^3 - 14x + 24)$	M1	Using $F(2) = 0$ or $F(4) = 1$ to find c if necessary. AEF
	$F(x) = 0$ ($x < \text{or} \leq 2$), $F(x) = 1$ ($x > \text{or} \geq 4$)	A1	State $F(x)$ for other values of x
		3	
(ii)	$G(y) = P(Y < y) = P(X^2 < y)$ $G(y) = P(X < \sqrt{y}) = F(\sqrt{y})$ $G(y) = (1/30) (-8/\sqrt{y} + y^{3/2} - 14\sqrt{y} + 24)$	M1 A1	Find or state $G(y)$ for $2 \leq x \leq 4$ from $Y = X^2$ (allow $<$ or $\leq$ throughout)
	Alternative method for question 10(ii)		
	Use $x = y^{1/2}$ to find $f(x) = (1/30)(8/y + 3y - 14), \frac{dx}{dy} = \frac{1}{2}y^{-1/2}$	(M1) A1)	Find $f(x)$ and $\frac{dx}{dy}$ for use in $g(y) = f(x) \times \left \frac{dy}{dx} \right $
	$g(y) = G'(y) = (1/30) (4/y^{3/2} + (3/2)y^{1/2} - 7/y^{1/2})$ for $4 \leq y \leq 16$ [$g(y) = 0$ otherwise]	A1 A1	Find $g(y)$. AEF State corresponding range of y for $G(y)$ or $g(y)$
(iii)	$(1/30) (-8/y^{1/2} + y^{3/2} - 14y^{1/2} + 24) = 0.8$	M1	Set $G(y) = 0.8$
	$-8 + y^2 - 14y = 0, y = 7 + \sqrt{57} \text{ or } 14.5[5]$ [rejecting $7 - \sqrt{57}$; allow 14.6]	M1 A1	Rearrange to give quadratic in y and solve to find value of y
		3	

21 - (9231/31_Summer_2023_Q3) - Circular Motion, Further Work On Distributions

(a)	At B, $mg \sin \theta = \frac{m4ag}{5a}$	M1	Allow cos instead of sin for M1. Do not award until tension = 0 used. Mass must be seen. No sign error.
	$\sin \theta = \frac{4}{5}$	A1	
		2	
(b)	At A, $T - mg \cos \theta = \frac{mv^2}{a}$	B1	
	Energy $\frac{1}{2}mv^2 - \frac{1}{2}m \times \frac{4ag}{5} = mga(\cos \theta + \sin \theta)$	M1 A1	Energy equation with 4 terms, dimensionally correct. Mass must be present, allow sign errors. Must see $\frac{1}{2}$ in the KE terms.
	Solve to find T	M1	Complete method leading to an expression in mg for T .
	$T = \frac{21}{5}mg$	A1	CWO
		5	

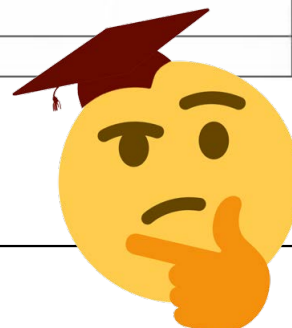


22 - (9231/31_Summer_2023_Q5) - Circular Motion, Further Work On Distributions

(a)	Hooke's law, $T_1 = \frac{\lambda mg}{a}(x-a)$ or $T_2 = \frac{\lambda mg}{a}\left(\frac{3}{4}x-a\right)$	B1	
	Also, $T_1 = \frac{mv^2}{x}$ and equate $\frac{mv^2}{x} = \frac{\lambda mg}{a}(x-a)$	M1	$v^2 = \frac{\lambda gx(x-a)}{a}$ Dimensionally correct terms.
	Similarly: $\frac{\lambda mg\left(\frac{3x}{4}-a\right)}{a} = \frac{2m\left(\frac{1}{2}v\right)^2}{\frac{3}{4}x}$	M1	$v^2 = \frac{3\lambda gx}{2a}\left(\frac{3}{4}x-a\right)$ Must have $\frac{1}{2}v$ and $\frac{3x}{4}$ on the RHS. Their dimensionally correct T_2 .
	Equate expressions for v^2 and solve for x in terms of a .	M1	
	$x = 4a$	A1	WWW
		5	SC B3 for answer of $4a$ using λ instead of λmg .
(b)	$\lambda = \frac{a}{xg(x-a)}v^2$ or $\lambda = \frac{2a}{3xg\left(\frac{3}{4}x-a\right)}v^2$ and substitute $x = 4a$, $v = \sqrt{12ag}$	M1	FT their expression for x .
	1	A1	CAO
		2	

23 - (9231/31_Summer_2023_Q7) - Circular Motion, Further Work On Distributions

(a)	$H = 80\sin^2\theta$ or $\frac{800\sin^2\theta}{g}$	B1	
	$T = 4\sin\theta$ or $\frac{40\sin\theta}{g}$	B1	
		2	
(b)	Between $t = T$ and $t = 3$ $\uparrow \frac{1}{4}H = \frac{1}{2} \times 10 \times (3-T)^2$	M1 A1	No extra terms.
	Use results from part (a) $\frac{1}{4}80\sin^2\theta = 5(3-4\sin\theta)^2$ $4\sin^2\theta - 8\sin\theta + 3 = 0$	M1	Substitute their expressions for H and T from part (a) and obtain a quadratic equation in $\sin\theta$ with no more than three terms.
	$\sin\theta = \frac{1}{2}$, $\theta = 30^\circ$	A1	Single answer. NFWW.
	Alternative method for question 7 part (b)		
	$\frac{3}{4}H = y(3) = 40 \times 3 \sin\theta - \frac{1}{2} \times 10 \times 3^2$	M1 A1	$120 \sin\theta - 45$
	Use results from (a): $\frac{3}{4}80\sin^2\theta = 120 \sin\theta - 45$ $4\sin^2\theta - 8\sin\theta + 3 = 0$	M1	Substitute their expressions for H and T from part (a) and obtain a quadratic equation in $\sin\theta$ with no more than three terms.
	$\sin\theta = \frac{1}{2}$, $\theta = 30^\circ$	A1	Single answer. NFWW.
		4	
(c)	When $t = 3$ speeds $\rightarrow 40\cos\theta$ and $\uparrow 40\sin\theta - 10 \times 3$	B1	
	Square and add to find square of speed: $v^2 = (20\sqrt{3})^2 + (-10)^2$	M1	Must be numerical.
	$v^2 = 1300$, $v = 10\sqrt{13}$ [= 36.1]	A1	
		3	



24 - (9231/33_Summer_2023_Q2) - Circular Motion, Further Work On Distributions

$\frac{1}{2}m\left(v^2 - \frac{v^2}{4}\right) = \frac{1}{2}\frac{\lambda mg}{a}\left[\left(\frac{2}{3}a\right)^2 - \left(\frac{1}{3}a\right)^2\right]$	M1	Kinetic energy = elastic potential energy, 4 terms, dimensionally correct, allow sign errors.
$\frac{1}{2}m\left(2ag - \frac{1}{2}ag\right) = \frac{1}{2}\frac{\lambda mg}{a}\left[\left(\frac{2}{3}a\right)^2 - \left(\frac{1}{3}a\right)^2\right]$	A1	With v substituted.
Solve $\left[\frac{3}{4}v^2 = \lambda g\left(\frac{3}{9}a\right)\right]$	M1	Solve to find value for λ dependent on energy equation with 3 or 4 terms
$\lambda = \frac{9}{2}$	A1	SCB2 for $\lambda = \frac{9}{2}mg$ if given λ not used
	4	

25 - (9231/41_Summer_2023_Q1) - Further Work On Distributions

$\bar{x} = \frac{91.2}{12} = 7.6$ $s_x^2 = \frac{1}{11}\left(695.8 - \frac{91.2^2}{12}\right) \left[= 0.2436 \left(\frac{67}{275}\right) \right]$	M1	Both.
CI: $7.6 \pm t_{11}(0.975) \times \sqrt{\frac{0.2436}{12}}$	M1	With a t -value.
2.201 seen	B1	
$7.6 \pm 0.3136 = [7.29, 7.91]$	A1	
	4	

26 - (9231/43_Summer_2023_Q2) - Further Work On Distributions

$\bar{x} + 1.771\sqrt{\frac{s^2}{14}} = 1.14$ or $\bar{x} - 1.771\sqrt{\frac{s^2}{14}} = 1.11$	M1	SO1 Allow incorrect t -value, not z -value.
[Add: $\bar{x} = \frac{1}{2}(1.14 + 1.11) = 1.125$] $\sum x = 15.75$	B1	Does not depend on use of a t -value.
Subtract or substitute: $\sqrt{\frac{s^2}{14}} = \frac{1}{2}\left(\frac{1.14 - 1.11}{1.771}\right)$	M1	Allow incorrect t -value, but not a z -value.
$s^2 = 0.00100[4]$ or $s = 0.0316[9]$	A1	$\frac{450}{448063}$, implied by correct final answer.
$\sum x^2 = 13s^2 + \frac{(\sum x)^2}{14}$	M1	OE
$\sum x^2 = 17.7(3)$	A1	CWO
	6	

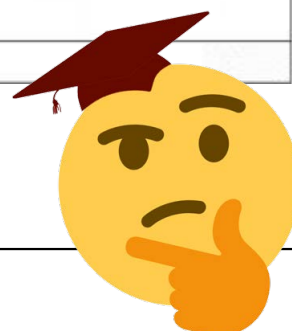


27 - (9231/31_Winter_2023_Q4) - Further Work On Distributions

(a)	Energy equation: gain in KE = loss in EPE	B1	At least one correct EPE.
	$\frac{1}{2}m \times 80ag = \frac{1}{2} \times \frac{5mg}{8a} \left((2L - 8a)^2 - (4a)^2 \right)$	M1	3-term energy equation.
	$4L^2 - 8aL - 20a^2 = 0$	A1	All correct.
	$L = 10a$	M1	Simplify to quadratic in L .
(b)	At P : $2T \cos \theta = m \times \text{acceleration}$	A1	Correct single answer.
		5	
	Hooke's law: $T = \frac{5mg}{8a} \times 12a = \frac{15}{2}mg$	M1	
	Solve, acceleration = 12g	A1	
		3	

28 - (9231/33_Winter_2023_Q6) - Further Work On Distributions

(a)	Angular speeds of P and Q are equal, so $\frac{v_Q}{x} = \frac{v_P}{3a-x}$	M1	
	$v_Q = \frac{2x\sqrt{ag}}{3a-x}$	A1	Shown convincingly: angular speeds equal stated. AG
		2	
(b)	For P : $T + mg \cos 60^\circ = \frac{m \times 4ag}{3a-x}$	B1	
	For Q : $T - mg \cos 60^\circ = \frac{mv_Q^2}{x}$	B1	
	Eliminate T : $-mg \cos 60^\circ + \frac{m \cdot 4ag}{3a-x} = mg \cos 60^\circ + \frac{mv_Q^2}{x}$	M1	
	$\frac{m \times 4ag}{3a-x} = 1 + \frac{mv_Q^2}{(3a-x)^2}$	M1	Solve to find x . Obtain 3-term quadratic equation.
	$4a(3a-x) = (3a-x)^2 + 4ax, \quad x^2 + 2ax - 3a^2 = 0$		
	$(x-a)(x+3a) = 0, \quad x = a$	A1	
		5	
(c)	Energy changes from initial position: Gain in KE of P : $\frac{1}{2}m(4ag - u^2)$ Loss in KE of Q : $\frac{1}{2}m\left(\left(\frac{u}{2}\right)^2 - v_Q^2\right)$ Loss in GPE of $P = mg(3a-x)(1 - \cos 60^\circ) (= mga)$ Gain in GPE of $Q = mgx(1 - \cos 60^\circ) \left(= \frac{1}{2}mga\right)$	B1	KEs correct.
		B1FT	GPEs correct.
	$\frac{1}{2}m(4ag - u^2) - \frac{1}{2}m\left(\left(\frac{u}{2}\right)^2 - v_Q^2\right) = -mgx(1 - \cos 60^\circ) + mg(3a-x)(1 - \cos 60^\circ)$	M1	Energy equation.
	Simplify: $4ag - \frac{5}{4}u^2 + ag = ag$ $u^2 = \frac{16}{5}ag, \quad u = \frac{4}{5}\sqrt{5ag}$	A1	AEF
		4	



29 - (9231/41_Winter_2023_Q3) - Further Work On Distributions

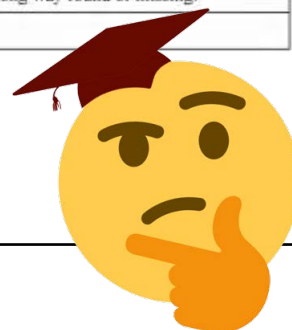
(a)	$H_0: \mu_B - \mu_G = 2 \quad H_1: \mu_B - \mu_G > 2$	B1	May use μ_d , but must be consistent with direction of differences found.										
	<table border="1"><tr><td>8</td><td>8</td><td>15</td><td>-2</td><td>-3</td><td>20</td><td>-3</td><td>0</td><td>5</td><td>4</td></tr></table>	8	8	15	-2	-3	20	-3	0	5	4	M1	Differences, at most 2 errors.
	8	8	15	-2	-3	20	-3	0	5	4			
	$\Sigma d = 52, \Sigma d^2 = 816, s^2 = \frac{1}{9} \left(816 - \frac{52^2}{10} \right) \left[= \frac{2728}{45} = 60.62 \right]$	M1	<i>Their</i> values but must see 9 and 10 used correctly.										
	$t = \frac{\frac{52}{10} - 2}{\sqrt{\frac{s^2}{10}}}$	M1											
	$t = 1.29967, 1.30$	A1											
	Tabular value = 2.262, '1.30' < 2.262, accept H_0	M1	Allow 'not significant'.										
	Insufficient evidence to suggest that cholesterol level has reduced by more than 2.	A1	CWO. Correct conclusion in context, following correct work, level of uncertainty in language. A0 if hypotheses wrong way round or missing.										
	7												
(b)	Population differences are normally distributed	B1											
		1											

30 - (9231/42_Winter_2023_Q1) - Further Work On Distributions

$s_x^2 = \frac{243.8}{59} \left[\frac{1219}{295} = 4.132 \right], s_y^2 = \frac{384.9}{74} \left[\frac{3849}{740} = 5.201 \right]$	B1	Implied by correct s or pooled estimate $\frac{243.8 + 384.9}{60 + 75 - 2} = 4.727$
$s^2 = \frac{4.132}{60} + \frac{5.201}{75} = 0.1382$ or $s = 0.3718$	M1	Using their sample variances. Pooled estimate M0.
	A1	
CI: $\frac{6345}{60} - \frac{7614}{75} \pm 1.96 \times 0.3718$	M1	With a z-value
	A1	With 1.96 (with their s) Pooled $\frac{6345}{60} - \frac{7614}{75} \pm 1.96 \times 2.174 \sqrt{\frac{1}{60} + \frac{1}{75}}$
$[3.5[0], 4.96]$ or $(3.5, 4.96)$	A1	4.23 ± 0.729 is A0, condone $[4.96, 3.5]$ etc.
	6	

31 - (9231/43_Winter_2023_Q1) - Further Work On Distributions

$H_0: \mu = 4.22 \quad H_1: \mu < 4.22$	B1	
$\bar{x} = \frac{42.05}{10} = 4.205, s^2 = \frac{1}{9} \left(176.83 - \frac{42.05^2}{10} \right) = \frac{13}{12000} = 0.0010833$	M1	
$t = \frac{4.205 - 4.22}{\sqrt{\frac{s^2}{10}}}$	M1	Their 4.205, their s^2 .
$t = -1.44$	A1	Condone sign.
Tabular value = 1.833: '1.44' < 1.833, accept H_0	M1	Compare their value with correct tabular value, signs consistent, allow 'not significant'.
There is insufficient evidence to support Maya's belief.	A1	Correct conclusion in context, following correct work, level of uncertainty in language. A0 if hypotheses wrong way round or missing.
	6	



1 - (9231/23_Summer_2017_Q6) - Inference Using Normal And T-distributions

$(10 - 4^2/N + 102 - 8^2/2N) / (N + 2N - 2)$	(AEF)	M1 A1
$112 - 48/N = 10(3N - 2), 15N^2 - 66N + 24 = 0$		M1 A1
$N = (66 \pm 54) / 30 = 4$		A1
Total:		5

2 - (9231/21_Summer_2017_Q7) - Inference Using Normal And T-distributions

$\bar{x} = 7.2$		B1
$s^2 = (542 - 72^2/10) / 9$ [= 118/45 or 2.622 or 1.619 ²]		M1
$H_0: \mu = 6.2, H_1: \mu > 6.2$	(AEF)	B1
$t_{9, 0.95} = 1.83[3]$		B1
$t = (\bar{x} - 6.2) / (s/\sqrt{10}) = 1.95$ [Accept H_1 :]		M1 A1
Claim (of mean mass increased) is justified	(AEF)	B1 FT
Total:		7



3 - (9231/23_Summer_2017_Q8) - Inference Using Normal And T-distributions

$\bar{x} = 28.32$ and $\bar{y} = 22.2$	B1
$s_A^2 = (41\ 100 - 1416^2/50) / 49$ and $s_B^2 = (20\ 140 - 888^2/40) / 39$	M1
$s_A^2 = 20.39$ and $s_B^2 = 164/15$ or 10.93 (to 3 s.f.)	A1
EITHER: $s^2 = s_A^2/50 + s_B^2/40 = 0.681$ or 0.825^2 ($19.98/50 + 10.66/40 = 0.666$ is M1 A0, max 8/9)	(M1 A1
$\bar{x} - \bar{y} \pm z s$	M1
$z_{0.975} = 1.96$	A1
$z s = 1.62$	A1
6.12 ± 1.62 or [4.50, 7.74]	A1)
OR: Assume equal [population] variances $s^2 = (49 s_A^2 + 39 s_B^2) / 88$ or $(41\ 100 - 1416^2/50 + 20\ 140 - 888^2/40) / 88$	(B1
$= 16.2$ or 4.02^2	B1
$\bar{x} - \bar{y} \pm z s \sqrt{(1/50 + 1/40)}$	M1
$z_{0.975} = 1.96$	A1
($t_{88, 0.975} = 1.9873$ or 1.99) 1.67 (1.70)	A1
6.12 ± 1.67 or [4.45, 7.79] (6.12 ± 1.70 or [4.42, 7.82])	A1)
Total:	9



4 - (9231/21_Summer_2017_Q9) - Inference Using Normal And T-distributions

$H_0: \mu_X = \mu_Y, H_1: \mu_X \neq \mu_Y$	(AEF)	B1
$\bar{x} = 13.4/8 \text{ or } 1.67[5], \bar{y} = 2.02$ (all to 3 s.f.)		B1
$s_X^2 = (24.7 - 13.4^2/8) / 7$ $= 451/1400 \text{ or } 0.3221 \text{ or } 0.5678^2 \text{ and}$ $s_Y^2 = (44.6 - 20.2^2/10) / 9$ $= 949/2250 \text{ or } 0.4218 \text{ or } 0.6494^2$		M1
$s^2 = (7 s_X^2 + 9 s_Y^2) / 16$ $\text{or } (24.7 - 13.4^2/8 + 44.6 - 20.2^2/10) / 16$	(AEF)	M1 A1
$= 6051/16\ 000 \text{ or } 0.3782 \text{ or } 0.6150^2$		A1
$t_{16, 0.95} = 1.746$		*B1
$[-] t = (\bar{y} - \bar{x}) / s \sqrt{(1/8 + 1/10)} = 1.18$		M1 A1
$t < 1.75$ so mean masses are the same	(AEF)	DB1 FT
SR: $Z = (\bar{y} - \bar{x}) / \sqrt{(s_X^2/8 + s_Y^2/10)}$ $= 0.345 / \sqrt{(0.078)} = 1.20$		(B1)
$Z < 1.645$ so mean masses are the same	(AEF)	(B1FT)
Total:		10



5 - (9231/21_Winter_2017_Q10) - Inference Using Normal And T-distributions

(i)	$\bar{x} = 0.91, \bar{y} = 1.205$ $s_x^2 = 19.56 / 49 [= 489/1225 \text{ or } 0.3992]$ and $s_y^2 = 30.25 / 59 [= 121/236 \text{ or } 0.5127]$	B1 M1	Find both sample means Estimate both population variances (allow biased here: 0.3912 and 0.5042)
	EITHER: $s^2 = s_x^2/50 + s_y^2/60$ $= 0.01653 \text{ or } 0.1286^2$ (to 3 s.f. throughout)	(M1) A1	Estimate or imply combined variance
	$[\pm] (\bar{y} - \bar{x}) \pm z s$	M1	Find confidence interval for difference $Y - X$ or $X - Y$
	$z_{0.95} = 1.645$ $[\pm] 0.295 \pm 0.211$ (allow 0.212)	A1	Use appropriate tabular value Evaluate confidence interval (either form)
	or $[\pm] [0.084, 0.506]$ (allow $[\pm] [0.083, 0.507]$)	A1)	
	OR: Assume equal [population] variances $s^2 = (49 s_x^2 + 59 s_y^2) / 108$ or $(19.56 + 30.25) / 108$	(B1)	State assumption Find or imply pooled estimate of common variance (note s_x^2 and s_y^2 not needed explicitly so first M1 may be implied by result)
	$= 4981/10800 \text{ or } 0.461 \text{ or } 0.679^2$	B1	
	$[\pm] (\bar{y} - \bar{x}) \pm z s \sqrt{1/50 + 1/60}$	M1	Find confidence interval for difference $Y - X$ or $X - Y$
	$z_{0.95} = 1.645$	A1	Use appropriate tabular z-value (or t-value from calculator)
	$[\pm] 0.295 \pm 0.214$ or $[\pm] [0.081, 0.509]$	A1)	Evaluate confidence interval (either form)
		7	

(ii)	$z = (1.205 - 0.91) / s = 0.295 / s$ $= 2.29[4] \quad [\text{or } 2.26[9]]$ (to 3 s.f.)	M1 A1	Find value of z (either sign)
	$(z) = 0.989[1] \quad [\text{or } 0.988[4]]$	A1	Find $\Phi(z)$
	$100 \times (1 - 0.989) \times 2 = 2.2 \quad [\text{or } 2.3]$ (to 1 d.p.)	M1 A1	Find limiting value for α , based on two-tail test (M0 for basing on one-tail test)
	$\alpha < (\text{or } \leq) 2.2 \quad [\text{or } 2.3]$	A1	Find set of possible values of α (Treat α instead of $\alpha\%$ as misread)
		6	

6 - (9231/21_Summer_2018_Q7) - Inference Using Normal And T-distributions

	$\bar{x} = 681.6/7 = 97.37$ (allow 97.4 for this B1)	B1	Find sample mean
	$s^2 = (66\,536.1 - 681.6^2/7) / 6$ $[= 27.96 \text{ or } 5.287^2]$	M1	Estimate population variance (allow biased here: 23.96 or 4.895 ²)
	$H_0: \mu = 94, H_1: \mu \neq 94$ (AEF)	B1	State hypotheses (B0 for $\bar{x} \dots$)
	$t = (\bar{x} - 94)/(s/\sqrt{7}) = 1.69$	M1 A1	Find value of t
	$t_{7,0.95} = 1.94[3]$	B1	State or use correct tabular t -value (or can compare $\bar{x}$ with $94 + 3.88 = 97.88$)
	[Accept H_0] Mean mass is equal to 94 kg (AEF)	B1✓	Consistent conclusion (FT on both t -values)
		7	



7 - (9231/21_Summer_2018_Q10) - Inference Using Normal And T-distributions

(i)	$H_0: \mu_P = \mu_C, H_1: \mu_P > \mu_C$	B1	State hypotheses (B0 for $\bar{x} \dots$)
	$s_P^2 = (138\,200 - 2620^2/50) / 49 = 18.61$ or $912/49$ and $s_C^2 = (157\,000 - 3060^2/60) / 59 = 15.93$ or $940/59$	M1 A1	Estimate both popn. variances (to 3 s.f.) (allow biased here: 18.24 and 15.67)
	$s^2 = s_P^2/50 + s_C^2/60 = 0.6378$ or 0.7986^2	M1 A1	Estimate combined variance (to 3 s.f.; may be implicit)
	$z_{0.9} = 1.28[2]$	*B1	State or use correct tabular z value
	$z = (\bar{y} - \bar{x}) / s = (52.4 - 51) / s = 1.75$ $z >$ tabular value so [accept H_1 and]	M1 A1	Calculate value of z (or $-z$) (or can compare $\bar{y} - \bar{x} = 1.02$ with 1.4)
	college P students take more time (AEF)	B1✓	Correctly stated conclusion (FT on z , dep *B1)
	[SC: $s^2 = (912 + 940)/108 = 17.15$ or 4.141^2 $z = 1.4 / s \sqrt{(1/50 + 1/60)} = 1.77$]		SC: Using pooled estimate of common variance can earn B1 M1 A1 (may be implied) M0 *B1 M0 B1✓ (max 5/9)
			9
(ii)	$\bar{x} - \bar{y} \pm z s$ (or $\bar{y} - \bar{x} \pm z s$)	M1	Find confidence interval for difference
	$z_{0.95} = 1.64[5]$	A1	Use appropriate tabular value
	1.4 ± 1.31 or $[0.09, 2.71]$ or -1.4 ± 1.31 or $[-2.71, -0.09]$ [SC: e.g. $\bar{x} - \bar{y} \pm z s \sqrt{(1/50 + 1/60)} = 1.4 \pm 1.30$]	A1	Evaluate confidence interval (either form) SC: Allow M1 A1 if common variance used (max 2/3)
			3

8 - (9231/23_Summer_2018_Q10) - Inference Using Normal And T-distributions

(i)	$H_0: \mu_c - \mu_h = 1, H_1: \mu_c - \mu_h > 1$ (AEF)	B1	State both hypotheses (B0 for $\bar{x} \dots$)
	$d_i: 1.1 \ 2.1 \ 0.7 \ 0.4 \ 2.0 \ 2.8 \ 0.2 \ 1.9 \ 1.4$	M1	Consider differences d_i , e.g. $x_i - y_i$
	$\bar{d} = 12.6 / 9 = 1.4$	B1	Find sample mean
	$s^2 = (23.72 - 12.6^2/9) / 8$ [= 0.76 or 0.8718 ²]	M1	Estimate population variance (allow biased here: [0.6756 or 0.8219 ²])
	$t_{8,0.9} = 1.397$ or 1.40	B1	State or use correct tabular t -value
	$t = (\bar{d} - 1) / (s/\sqrt{9}) = 1.38$ [Accept H_0] No evidence for coach's belief	M1 A1	Find t (or compare $\bar{d} - 1 = 0.4$ with $t_{8,0.9} s / \sqrt{9} = 0.406$) Consistent conclusion (FT on both t -values)
	or time will not decrease by more than 1 s (AEF)	B1✓	
			SC Wrong (hypothesis) test can earn only B1 for hypotheses B1✓ for conclusion (max 2/8)
			8
(ii)	$s_y^2 = (11\,395.79 - 319.5^2/9) / 8$ = 6.693 or 2.587 ² (to 3 sf)	B1	Estimate population variance using y 's sample (allow use of biased here: 5.950 or 2.439 ²)
	$319.5/9 \pm t \sqrt{(s_y^2/9)}$	M1	Find confidence interval (M0 if s^2 not s_y^2)
	$t_{8,0.975} = 2.306$ or 2.31	A1	State or use correct tabular value of t
	35.5 ± 1.99 or $[33.5, 37.5]$ (allow 35.5 ± 2.0)	A1	Evaluate C.I. (either form) SC B1 (max 1/4) for 1.4 ± 0.67 or $[0.73, 2.07]$
			4



9 - (9231/22_Winter_2018_Q6) - Inference Using Normal And T-distributions

$\bar{x} = 90.3 / 8 = 11.2875$ (to 4 s.f.)	B1	Find sample mean
$s^2 = (1043.67 - 90.3^2/8) / 7$ $= 19\,527 / 5600$ or 3.487 [or 1.867^2] (to 4 s.f.)	M1	Estimate population variance
$90.3 / 8 \pm t \sqrt{s^2/8}$	A1	(allow biased here: 3.051 or 1.747^2)
$t_{7,0.975} = 2.365$ (to 4 s.f.)	M1	Find confidence interval
11.3 ± 1.6 or $[9.7, 12.8[5]]$	A1	State or use correct tabular value of t
	A1	Evaluate C.I. (either form)
	6	

10 - (9231/21_Winter_2018_Q8) - Inference Using Normal And T-distributions

(i)	$s_A^2 = (531\,000 - 5120^2/50) / 49$ and $[= 6712/49]$ $s_B^2 = (375\,135 - 3760^2/40) / 39$ $[= 21695/39]$	M1	Estimate both population variances
	$s_A^2 = 136.98$ and $s_B^2 = 556.28$ (to 3 s.f. throughout)	A1	(allow biased here: 134.24 and 542.4)
	<i>EITHER:</i> $s^2 = s_A^2/50 + s_B^2/40 = 16.65$ or 4.08^2	M1A1	<i>EITHER:</i> Estimate combined variance
	$z_{0.975} = 1.96$	*B1	State or use correct tabular z (or t) value
	$z = (\bar{y} - \bar{x}) / s = (102.4 - 94) / s = 2.06$ $z >$ tabular value [accept H_1]	M1A1	Calculate value of z (or $-z$) (or can compare $\bar{y} - \bar{x} = 8.4$ with 8.0)
	so $\mu_A \neq \mu_B$ or salaries do differ (AEF)	B1	Correct conclusion (FT on z , dep *B1)
	<i>OR:</i> Assume equal [population] variances $s^2 = (49 s_A^2 + 39 s_B^2) / 88$ or $(531\,000 - 5120^2/50 + 375\,135 - 3760^2/40) / 88$	(B1)	<i>OR:</i> State assumption about variances (may be in part (ii)) Find pooled estimate of common variance (M1 A1 for s_A^2 and s_B^2 may be implied here)
	$= 28407/88$ or 322.8 or 17.97^2	B1	
	$z_{0.975} = 1.96$	*B1	State or use correct tabular z (or t) value
	$z = 8.4 / s \sqrt{1/50 + 1/40} = 2.20$	M1A1	Calculate value of z (or $-z$) (or can compare $\bar{y} - \bar{x} = 8.4$ with 7.47)
	$z >$ tabular value [accept H_1] so $\mu_A \neq \mu_B$ or salaries do differ (AEF)	B1)	Correct conclusion (FT on z , dep *B1)
		8	
(ii)	No assumption [of normality] needed since large samples or due to central limit theorem	B1	Valid comment on required assumptions (AEF) (assumption of equal variances may earn B1 above)
		1	



11 - (9231/21_Winter_2018_Q9) - Inference Using Normal And T-distributions

(i)	$\bar{x} = 13.88/8 = 1.735$ (allow 1.73 or 1.74 for this B1)	B1	Find sample mean
	$s^2 = (24.1182 - 13.88^2/8) / 7$ [= 13/2500 or 0.0052 or 0.07211 ²]	M1	Estimate population variance (allow biased here: 0.00455 or 0.06745 ²)
	$H_0: \mu = 1.7, H_1: \mu > 1.7$ (AEF)	B1	State hypotheses (B0 for $\bar{x} \dots$)
	$t = (\bar{x} - 1.7) / (s/\sqrt{8}) = 1.37$	M1A1	Find value of t
	$t_{7, 0.95} = 1.89[5]$	B1	State or use correct tabular t -value (or can compare $\bar{x}$ with $1.7 + 0.048 = 1.75$)
	[Accept H_0 :] Mean ht, not greater than 1.7 m (AEF)	B1	Consistent conclusion (FT on both t -values)
		7	
(ii)	$13.88/8 \pm t \sqrt{s^2/8}$	M1	Find confidence interval
	$t_{7, 0.975} = 2.36[5]$	A1	State or use correct tabular value of t
	$1.73[5] \pm 0.06$ or $[1.67, 1.80]$	A1	Evaluate confidence interval (either form)
		3	

12 - (9231/22_Winter_2018_Q11) - Simple Harmonic Motion, Inference Using Normal And T-distributions

A(i)	EITHER: $40 e_0 / 0.8 = 2g, e_0 = 0.4$ [or $OP_0 = 1.2$] [m]	M1A1	Find extension e_0 [or OP_0] at equilibrium position P_0
	$2 \frac{d^2x}{dt^2} = -40(e_0 + x) / 0.8 + 2g$ or $= +40(e_0 - x) / 0.8 - 2g$	M1A1	Use Newton's law at general point (e.g. x below or above P_0) (ignore LHS sign here only)
	OR: $2 \frac{d^2y}{dt^2} = -40(y - 0.8) / 0.8 + 2g$ $= -50y + 60$	(M1A1)	Use Newton's law at general point in terms of $y = OP$ (ignore LHS sign here only)
	Take $x = y - 1.2$ to give	M1A1	Change variable to give standard form of SHM eqn
	$\frac{d^2x}{dt^2} = -25x$	A1	Hence SHM (A0 if wrong sign or LHS unclear) (B1 only for stating SHM eqn, without proof)
	$T = 2\pi / \sqrt{25} = 2\pi/5$ AG	A1	Verify period T using $T = 2\pi/\omega$ with $\omega = \sqrt{25}$
	$OP_0 = 1.2$ [m]	B1	State OP_0 explicitly (may imply first M1 A1)
		7	
A(ii)	$0.4^2 = 25(a^2 - 0.06^2)$	M1A1	Find amplitude a from $v^2 = \omega^2(a^2 - x^2)$
	$a = \sqrt{(0.0064 + 0.0036)} = 0.1$ [m]	A1	
		3	
A(iii)	$40 e_1 / 0.8 = (2 + M)g, [e_1 = 0.4 + 0.2 M]$ or $40(e_1 - e_0) / 0.8 = Mg [e_1 - e_0 = 0.2 M]$	M1A1	Find extension e_1 [or OP_1] at first equilibrium position P_1 or equate additional extension to M by Hooke's Law
	$e_1 - e_0 = a, M = a/0.2 = 0.5$	M1A1	Find M by relating e_0, e_1 and a
		4	



B(i)	$H_0: \mu_1 = \mu_2, H_1: \mu_1 > \mu_2$ (or in terms of μ_A, μ_B)	B1	State hypotheses (B0 for $\bar{x} \dots$)
	$s_1^2 = (14.1775 - 10.56^2/8) / 7 = 0.03404$ and $s_2^2 = (15.894 - 12.39^2/10) / 9 = 0.06031$ (to 3 s.f.)	M1A1	Estimate both population variances (allow biased here: 0.02979 and 0.05428)
	$s^2 = (7 s_1^2 + 9 s_2^2) / 16$ or $(14.1775 - 10.56^2/8 + 15.894 - 12.39^2/10) / 16$ $= 0.78\ 109 / 16$ or 0.04882 or 0.22095 ² (to 3 s.f.)	M1A1	Find pooled estimate of common variance (M1 A1 for s_1^2 and s_2^2 may be implied here)
	$t_{15, 0.9} = 1.337$ (to 3 s.f.)	*B1	State or use correct tabular t value
	$t = (1.32 - 1.239) / \sqrt{(1/10 + 1/8)} = 0.773$	M1A1	Calculate value of t (or $-t$) (or can compare $\bar{y} - \bar{x} = 0.081$ with 0.140)
	$t < \text{tabular value}$ so claim not justified or A's not heavier than B's (AEF)	B1	Correct conclusion (FT on t , dep *B1)
	SC: $z = (1.32 - 1.239) / \sqrt{(s_1^2/8 + s_2^2/10)}$ $= 0.081 / \sqrt{(0.078)} = 0.799$	(B1	SC: Implicitly taking s_1^2, s_2^2 as unequal popln. variances (may also earn first B1 M1 A1)
	$z < 1.282$ so claim is not justified (AEF)	B1)	Comparison with $z_{0.9}$ and conclusion (FT on z ; max 5/9)
		9	
B(ii)	$\bar{x} = 1.28$ and $s^2 = 0.294 / 7 = 0.042$ (or 0.205 ²)	M1	Find sample mean & estimate popn. var (allow M1 if 0.294 / 8)
	$t = (1.28 - p) / \sqrt{(0.042/8)}$ (AEF)	M1A1	Find value of t
	$t > 1.415$ (< is A0), $p < 1.28 - 0.1025, p_{\text{max}} = 1.18$	M1A1	Find p_{max} by comparison with tabular value, here $t_{7, 0.9}$
		5	

13 - (9231/21_Summer_2019_Q8) - Inference Using Normal And T-distributions

$H_0: \mu_0 - \mu_e = 0.05, H_1: \mu_0 - \mu_e > 0.05$ (AEF)	B1	State both hypotheses (B0 for $\bar{x} \dots$)
$d_i: 0.10 \ 0.07 \ 0.03 \ 0.03 \ 0.21 \ 0.19$ (or in sec)	M1	Consider differences d_i from e.g. $x_0 - y_e$
$\bar{d} = 0.63 / 6 = 0.105$ (or 6.3 sec)	B1	Find sample mean
$s^2 = (0.0969 - 0.63^2/6) / 5$ [$= 123/20\ 000$ or 0.00615 or 0.0784 ²] (or 22.14)	M1	Estimate population variance (allow biased here: [41/8000 or 0.005125 or 0.0716 ²])
$t_{5, 0.9} = 1.476$ (to 3 sf)	B1	State or use correct tabular t -value
$t = (\bar{d} - 0.05) / (s/\sqrt{6}) = 1.72$	M1 A1	Find value of t (or compare $\bar{d} - 0.05 = 0.055$ with $(t_{5, 0.9})s/\sqrt{6} = 0.0473$)
[Reject H_0 :] Evidence for organiser's belief or times improve by more than 0.05 min (AEF)	B1	FT on both t -values Consistent conclusion
		SC Wrong type of hypothesis test can earn only B1 for hypotheses B1FT for conclusion (max 2/8)
	8	



14 - (9231/23_Summer_2019_Q9) - Inference Using Normal And T-distributions

(i)	$\bar{x} = 0.785$	B1	Find sample mean
	$s^2 = (6.19 - 7.85^2/10)/9$ [= 37/12 000 or 0.003083 or 0.05553 ²]	M1	Estimate population variance (allow biased here: 0.002775 or 0.05268 ²)
	$H_0: \mu = 0.75, H_1: \mu > 0.75$ (AEF)	B1	State hypotheses (B0 for $\bar{x}$...)
	$t_{9,0.95} = 1.83[3]$	B1	State or use correct tabular t -value
	$t = (\bar{x} - 0.75)/(s/\sqrt{10}) = 1.99$ [Reject H_0 and accept H_1]	M1 A1	Find value of t (or compare $\bar{x}$ with $0.75 + 0.032 = 0.782$) Consistent conclusion
	Claim (of yield per plant increased) is justified (AEF)	B1	FT on both t -values (must be t value)
		7	
(ii)	$\bar{x} \pm t \sqrt{(s^2/10)}$	M1	Find confidence interval (must be a t value)
	$t_{9,0.975} = 2.26[2]$	B1	State or use correct tabular value of t
	0.785 ± 0.04 or [0.745, 0.825]	A1	Evaluate confidence interval (either form)
		3	

15 - (9231/21_Summer_2019_Q11) - Momentum And Impulse, Circular Motion, Inference Using Normal And T-distributions

E(i)	$\frac{1}{2}mu_P^2 = \frac{1}{2}m(2\sqrt{ag/2}) + mga$ [$u_P^2 = (25/2)ag$]	M1 A1	u_P is speed of P at lowest point, v_Q is speed of Q immediately after collision. Apply conservation of energy at lowest point (A0 if no m)
	$\frac{1}{2}4mv_Q^2 = 4mag$	M1	Find speed v_Q at lowest point by conservation of energy (A0 if no m)
	$v_Q = \sqrt{2ag}$ or $1.41\sqrt{ag}$ or $4.47\sqrt{a}$	A1	
	$mu_P = [\pm] mv_P + 4mv_Q$	M1	Find v_P using conservation of momentum (m may be omitted)
	$v_P = [\pm] (-5\sqrt{2} + 4\sqrt{2})\sqrt{ag}$	A1	
	$ v_P = (3/\sqrt{2})\sqrt{ag}$ or $2.12\sqrt{ag}$ or $0.671\sqrt{a}$ (AEF)	A1	Hence find speed of P
		7	
E(ii)	V_P is speed of P when it loses contact $\frac{1}{2}mV_P^2 = \frac{1}{2}mv_P^2 - mga(1 + \cos \alpha)$ [$V_P^2 = (9/2)ag - 2ga(1 + \cos \alpha) = (5/2 - 2 \cos \alpha)ag$]	M1 A1	Apply conservation of energy at D (A0 if no m)
	[$R_D =$] $mV_P^2/a - mg \cos \alpha = 0$ [$V_P^2 = ag \cos \alpha$]	M1 A1	Apply $F = ma$ radially at D with reaction = 0
	$(5/2 - 2 \cos \alpha)ag = ag \cos \alpha$, $\cos \alpha = 5/6$ or 0.833	A1	Combine to find $\cos \alpha$
		5	



O(i)	$ts_A / \sqrt{8} = \frac{1}{2}(16.7 - 13.5) [= 1.6]$	M1	Relate s_A to semi-width of confidence interval
	$t_{7, 0.975} = 2.365$ (to 3 s.f.)	A1	State or use correct tabular t value
	$[s_A = \sqrt{8} \times 1.6 / 2.365 = 1.9135], s_A^2 = 3.66[16]$	A1	Hence find unbiased estimate of A 's population variance
		3	
O(ii)	$H_0: \mu_A = \mu_B, H_1: \mu_A > \mu_B$ (AEF)	B1	State hypotheses (B0 for $\bar{x} \dots$)
	$[\bar{x}_A = 15.1], \bar{x}_B = 85.2 / 6 = 14.2$	B1	Find sample mean for B
	$s_B^2 = (1221.06 - 85.2^2 / 6) / 5$ $= 561 / 250$ or 2.244 or 1.498 ² (all to 3 s.f.)	M1	Estimate or imply population variance for B (allow biased here: 1.87 or 1.367 ²)
	$s^2 = (7 s_A^2 + 5 s_B^2) / 12 = 3.0709$ or 1.752 ²	M1 A1	Estimate (pooled) common variance (s_B^2 not needed explicitly)
	$t_{12, 0.95} = 1.782$	B1	State or use correct tabular t value
	$[-] t = (\bar{x}_A - \bar{x}_B) / (s \sqrt{(1/8 + 1/6)}) = 0.951$ $t < 1.78$ so [accept H_0]	M1 A1	Find value of t (or can compare $\bar{x}_A - \bar{x}_B = 0.9$ with 1.69) Correct conclusion
	mean mass of B not less than mean mass of A (AEF)	B1	
		9	
			SC1: Implicitly taking s_A^2, s_B^2 as unequal population variances (may also earn first B1 B1 M1) $z = (\bar{x}_A - \bar{x}_B) / \sqrt{(s_A^2/8 + s_B^2/6)}$ $= 0.9 / \sqrt{(0.8317)} = 0.987$ $z < 1.645$ so
			DepSC1: mean mass of B not less than mean mass of A (AEF) Comparison with $z_{0.95}$ and conclusion (FT on z) (can earn at most 5/9)

16 - (9231/23_Summer_2019_Q11) - Simple Harmonic Motion, Inference Using Normal And T-distributions

E(i)	$\frac{1}{2} m u^2 = \frac{1}{2} k m g (\frac{1}{4} a)^2 / a, k = 16 u^2 / a g$	AG	M1 A1	Verify k by using conservation of energy (M0 if ω is found and SHM formula $v = a\omega$ then used)
			2	
E(ii)	$\pm m d^2x/dt^2 = k m g x/a$		*M1	Apply Newton's law at general point (e.g. $OP = a + x$), requires m
	$d^2x/dt^2 = -(kg/a)x$ or $-(16u^2/a^2)x [\omega = \sqrt{(kg/a)} \text{ or } 4u/a]$		A1	Derive standard SHM form (requires minus sign)
	Period is $2\pi\sqrt{(a/kg)} = 2\pi\sqrt{(a^2/16u^2)} = \pi a / 2u$		DM1 A1	Find period from $2\pi/\omega$
			4	
E(iii)	$[\frac{1}{2} m v^2 = \frac{3}{4} \times \frac{1}{2} m u^2] v^2 = \frac{3}{4} u^2 \text{ or } v = (\sqrt{3}/2) u$		M1	Relate v and u using given loss in energy
	$x = (\frac{1}{4} a) \sin \omega t, v = (\frac{1}{4} a) \omega \cos \omega t \text{ or } u \cos \omega t$		M1	Relate v and t
	$\cos \omega t = (\sqrt{3}/2) u / (\frac{1}{4} a) \omega = (\sqrt{3}/2)$		M1 A1	Combine to find $\cos \omega t$
	$t = (\pi/6) / (4u/a) = \pi a / 24u \text{ or } 0.0417 \pi a / u \text{ or } 0.131 a/u$		M1 A1	and hence t
			6	



O	$\bar{x} - \bar{y} = 0.448 - 0.48 = [-] 0.032$	B1	Find difference in sample means (either sign; may be implied)
	$s_x^2 = (10.1 - 22.4^2/50) / 49 = 0.0648 / 49 \quad [= 0.001322]$	M1 A1	Estimate both population variances (may be implied)
	$s_y^2 = (16.3 - 28.8^2/60) / 59 = 2.476 / 59 \quad [= 0.04197]$	A1	(allow biased here: 0.001296 and 0.04127)
	$s_c^2 = s_x^2/50 + s_y^2/60$	M1	Estimate combined variance
	$= 0.0007259 \text{ or } 0.02694^2 \text{ (to 3 s.f. throughout)}$	A1	
	$z = 0.032 / s_c \quad \text{[or } 0.032 / s_c \sqrt{(1/50 + 1/60)}]$ $= 1.188 \quad \text{[or } 1.0895]$	M1 A1	Find value of z (either sign)
	$\Phi(z) = 0.8826 \quad \text{[or } 0.8620]$	A1	Find $\Phi(z)$
	$100 \times (1 - \Phi(z)) = 11.7 \text{ [or 13.8]}$	M1 A1	Find limiting value for α , based on one-tail test (M0 for basing on two-tail test)
	$\alpha > \text{(or } \geq) 11.7 \text{ [or 13.8]}$	A1	Find set of possible values of α Allow 11.8 if 11.74 seen (Misreading $\alpha\%$ as α loses only last A1)

O	Alternative method for question 11O		
	$\bar{x} - \bar{y} = 0.448 - 0.48 = [-] 0.032$	B1	Find difference in sample means (either sign; may be implied)
	$s_x^2 = (10.1 - 22.4^2/50) / 49 = 0.0648 / 49 \quad [= 0.001322]$	M1 A1	Estimate both population variances (may be implied)
	$s_y^2 = (16.3 - 28.8^2/60) / 59 = 2.476 / 59 \quad [= 0.04197]$	A1	(allow biased here: 0.001296 and 0.04127)
	Assume equal [population] variances	B1	State assumption
	$s_c^2 = (49 s_x^2 + 59 s_y^2) / 108 \text{ or } (0.0648 + 2.476) / 108$ $= 0.02353 \text{ or } 0.1534^2$	B1	(Find pooled estimate of common variance s_x^2 and s_y^2 not needed explicitly so may be implied by result)
	$z = 0.032 / s_c \quad \text{[or } 0.032 / s_c \sqrt{(1/50 + 1/60)}]$ $= 1.188 \quad \text{[or } 1.0895]$	M1 A1	Find value of z (either sign)
	$\Phi(z) = 0.8826 \quad \text{[or } 0.8620]$	A1	Find $\Phi(z)$
	$100 \times (1 - \Phi(z)) = 11.7 \text{ [or 13.8]}$	M1 A1	Find limiting value for α , based on one-tail test (M0 for basing on two-tail test)
	$\alpha > \text{(or } \geq) 11.7 \text{ [or 13.8]}$	A1	Find set of possible values of α Allow 11.8 if 11.74 seen (Misreading $\alpha\%$ as α loses only last A1)
		12	

17 - (9231/21_Winter_2019_Q6) - Inference Using Normal And T-distributions

(i)	$t \sqrt{(s^2/9)} = \frac{1}{2} (1.85 - 1.65) [= 0.1]$	M1	Find estimate s^2 of population variance (must be t)
	$t_{8,0.975} = 2.306 \text{ (to 3 s.f.)}$	A1	Use of correct tabular t -value
	$s^2 = 9 \times 0.04337^2 = 0.0169 \text{ or } 0.130[1]^2$	A1	
		3	
(ii)	$\bar{x} = \frac{1}{2} (1.65 + 1.85) = 1.75$ or $\Sigma x = 9 \times 1.75 = 15.75$	M1 A1	Find sample mean $\bar{x}$
	$s^2 = (\Sigma x^2 - 9 \times \bar{x}^2) / 8$ or $\{\Sigma x^2 - (\Sigma x)^2 / 9\} / 8$	M1	or Σx
	$\Sigma x^2 = 8 \times 0.0169 + 15.75^2/9 = 27.7$	A1	Find Σx^2 from s^2
		4	



18 - (9231/21_Winter_2019_Q8) - Inference Using Normal And T-distributions

$\bar{x}_A = 32.4 / 8 = 4.05$	B1	Find sample mean for A
$s_A^2 = (131.82 - 32.4^2/8) / 7$ $s_A^2 = 3/35$ (or 0.08571 or 0.2928 ² both to 3 s.f.)	M1	Estimate or imply popln. variance for A
$H_0: \mu_A = \mu_B$, $H_1: \mu_A \neq \mu_B$	B1	State hypotheses. AEF
$s^2 = (7 s_A^2 + 9 s_B^2) / 16 = 0.12497$ or 0.3535^2 or $\frac{3999}{32000}$	M1 A1	Estimate (pooled) common variance
$t_{16, 0.95} = 1.746$	B1*	State or use correct tabular t value
$[-] t = (x_A - \bar{x}_B) / s \sqrt{(1/8 + 1/10)}$	M1	
$= 0.27 / 0.1677 = 1.61$	A1	Find value of t (or can compare $\bar{x}_A - \bar{x}_B = 0.27$ with 0.293)
$t < 1.75$ so [accept H_0] mean masses are the same	DB1	Correct conclusion (FT on t, dep B1*), AEF
	9	

19 - (9231/43_Summer_2020_Q2) - Inference Using Normal And T-distributions

$s^2 = \frac{10.2}{40} + \frac{11.1}{50} = 0.477$	M1A1
$CI = (24.4 - 17.2) \pm z s$	M1
$= (24.4 - 17.2) \pm 1.645 \sqrt{0.477}$	A1
$= [6.06, 8.34]$	A1
	5

20 - (9231/41_Summer_2020_Q4) - Inference Using Normal And T-distributions

$H_0: \mu_x = \mu_y$, $H_1: \mu_x \neq \mu_y$	B1
$s_x^2 = \frac{1}{49} \left(5.1 - \frac{15.2^2}{50} \right) = 0.0097796$; $s_y^2 = \frac{1}{39} \left(4.8 - \frac{13.4^2}{40} \right) = 0.007974$	M1A1
$s^2 = \frac{0.00977959}{50} + \frac{0.007974}{40} = 0.0003949$	M1A1
$z = \frac{0.304 - 0.335}{\sqrt{0.0003949}} = (-)1.56$	M1A1
Compare with 1.645	M1
Accept H_0 : insufficient evidence to reject manager's claim	A1
	9



21 - (9231/41_Summer_2020_Q5) - Inference Using Normal And T-distributions

(a)	$\sum x = 183.6, \sum x^2 = 4249.08, \bar{x} = 22.95$	B1
	$s^2 = \frac{1}{7} \left(4249.08 - \frac{183.6^2}{8} \right) = 5.066$	M1
	$H_0: \mu = 22.0, H_1: \mu > 22.0$	B1
	$t = \frac{22.95 - 22.0}{\sqrt{\frac{s^2}{8}}} = 1.194$	M1A1
	Compare t with correct tabular value 1.895	M1
	Accept H_0 : mean distance thrown is not more than 22.0 m	A1
		7
(b)	$22.95 \pm t \sqrt{\frac{s^2}{8}}$	M1
	With $t = 2.365$	B1
	[21.1, 24.8]	A1
		3

22 - (9231/43_Summer_2020_Q5) - Inference Using Normal And T-distributions

(a)	$\bar{x} = \frac{828}{10} = 82.8, s^2 = \frac{1}{9} \left(68622 - \frac{828^2}{10} \right) = 7.0667$	B1
	$\frac{82.8 - k}{\sqrt{\frac{s^2}{10}}} \geq t$ where $t = 1.833$ (M1 if equality, A1 for inequality and correct t value)	M1 A1
	$k \leq 81.259, k \leq 81.3$	A1
		4
(b)	$H_0: \mu_A = \mu_B, H_1: \mu_A \neq \mu_B$	B1
	Pooled variance = $\frac{9 \times 7.0667 + 7 \times 9.966}{10 + 8 - 2} = s_p^2$	M1
	= 8.335	A1
	$t = \frac{82.8 - 79.8}{s_p \sqrt{\frac{1}{10} + \frac{1}{8}}} = 2.19$	M1A1
	Compare with 2.12 ($t(16, 0.975)$) and reject H_0	M1
	Population means are not the same	A1
		7



23 - (9231/41_Winter_2020_Q1) - Inference Using Normal And T-distributions

$s_x^2 = \frac{1}{39} \left(1587.0 - \frac{242.0^2}{40} \right) = 3.15128 \quad \left(= \frac{1229}{390} \right)$ $s_y^2 = \frac{1}{59} \left(2532.6 - \frac{373.2^2}{60} \right) = 3.58129 \quad \left(= \frac{26412}{7375} \right)$	M1 A1	Either unsimplified Both correct
$s^2 = \frac{3.15128}{40} + \frac{3.58129}{60} = 0.13847$	M1 A1	Pooled variance is M0A0 or $s = 0.3721$
$6.05 - 6.22 \pm zs$	M1	FT their s , must be a z value
$= 0.17 \pm 1.645\sqrt{0.13847}$	A1	With 1.645
$= [-0.442, 0.782] \text{ or } [-0.782, 0.442]$	A1	
	7	

24 - (9231/42_Winter_2020_Q1) - Inference Using Normal And T-distributions

$s^2 = \frac{313.1}{10} = 31.3$	B1	Can be implied
$t = \frac{176.2 - 172.5}{s / \sqrt{11}}$	M1	
$= 2.19$	A1	
$'2.19' < 2.228$	M1	Comparison of their t value with 2.228
accept H_0 : insufficient evidence to reject mean height is 172.5 OR sufficient evidence to accept mean height is 172.5	A1 FT	Correct conclusion in context, level of uncertainty in language used. No contradictions. CWO Do not accept: 'mean height is 172.5' Do not accept use of symbols that are undefined e.g. insufficient evidence to reject $\mu = 172.5$
	5	

25 - (9231/41_Winter_2020_Q4) - Inference Using Normal And T-distributions

Assume (population) differences are normally distributed	B1	
$H_0: \mu_X - \mu_Y = 0.3 \quad H_1: \mu_X - \mu_Y > 0.3$	B1	
Diff: 0.9 0.4 0.7 0.5 -0.2 0 -0.1 1.2 0.3	M1	Signed differences
$\Sigma d = 3.7, \quad \Sigma d^2 = 3.29$ $\bar{d} = 0.411, \quad s^2 = \frac{1}{8} \left(3.29 - \frac{3.7^2}{9} \right) = 0.2211$	M1	$s = 0.4702$
$t = \frac{0.411 - 0.3}{\sqrt{\frac{0.2211}{9}}}$ $= 0.708 \text{ or } 0.709$	M1 A1	
$'0.708' < 1.397$	M1	Compare <i>their</i> 0.708 with 1.397
Accept H_0 Insufficient evidence to support claim	A1 FT	In context, except possibly hypotheses Level of uncertainty in language used. No contradictions
	8	



26 - (9231/42_Winter_2020_Q6) - Inference Using Normal And T-distributions

(a)	$\bar{x} = \frac{1}{2}(4.22 + 3.36) = 3.79$	B1	
	$4.22 - 3.36 = \frac{2s}{\sqrt{10}}$	M1	Using z value implies M0
	With $t = 1.833$	A1	
	($s = 0.7418$) variance = 0.55(0)	A1	
		4	
(b)	$H_0: \mu_A = \mu_B \quad H_1: \mu_A > \mu_B$	B1	Not $\bar{x}$
	$s^2 = \frac{1}{7} \left(98.02 - \frac{27.86^2}{8} \right) = 0.142(507)$	M1	$= \frac{1}{7}(98.02 - 97.02)$
	Pooled estimate = $\frac{9 \times 0.550316 + 7 \times 0.142507}{16}$	M1	
	$= 0.372$	A1	
	$t = \frac{\frac{27.86}{8} - 3.79}{\sqrt{0.372} \sqrt{\frac{1}{10} + \frac{1}{8}}} = (-)1.063$	M1A1	
	'1.063' < 1.337 oe	M1	Compare <i>their</i> value with 1.337
	Accept H_0 Nassa's claim is not supported	A1ft	Correct conclusion in context, level of uncertainty in language used. No contradictions.
		8	

27 - (9231/41_Summer_2021_Q1) - Inference Using Normal And T-distributions

(a)	$\bar{x} = \frac{56.56}{7} = 8.08$	B1	Accept unsimplified.
	$s^2 = \frac{1}{6} \left(457.275 - \frac{56.56^2}{7} \right) = 0.045033 \left[= \frac{1351}{30000} \right]$	M1	Accept unsimplified.
	$t = \frac{8.08 - 8.22}{\sqrt{\frac{s^2}{7}}} = -1.745$	M1 A1	1.74 – 1.75
	Tabular value 1.440 1.745 > 1.440, so reject H_0	M1	Comparison with 1.440 and correct FT conclusion.
	There is insufficient evidence to support the hypothesis that the (population) mean is 8.22 OR sufficient evidence to suggest mean is less than 8.22 OR sufficient evidence to suggest $\mu < 8.22$ OE	A1	Correct conclusion, in context, following correct work. Level of uncertainty in language is used.
		6	
(b)	Underlying distribution is normal/ population is normal	B1	X is normal distributed.
		1	



28 - (9231/43_Summer_2021_Q1) - Inference Using Normal And T-distributions

$s^2 = \frac{1}{9} \left(2313 - \frac{152^2}{10} \right) = 0.28889$	M1	$\frac{13}{45}$, accept unsimplified.
$H_0: \mu = 14.8 \quad H_1: \mu > 14.8$	B1	If μ not used, 'population mean' required. Must see 14.8, must be = and >.
$[t] = \frac{\frac{152}{10} - 14.8}{s / \sqrt{10}}$	M1	Using unbiased estimate. p-value is 0.0215.
$[t] = 2.35$	A1	Rounds to 2.35.
Tabular value = 1.833 '2.35' > 1.833 Reject H_0	M1	Comparison with 1.833 and correct FT conclusion.
Sufficient evidence to accept Farmer B's claim oe	A1	Correct conclusion, in context, following correct work. Level of uncertainty in language is used. Allow ± 1 difference in third significant figure of their t for this mark.
	6	

29 - (9231/43_Summer_2021_Q3) - Inference Using Normal And T-distributions

$\bar{x} = \frac{89}{50} = 1.78 \quad \bar{y} = \frac{67.2}{40} = 1.68$	B1	Implied by 0.1 in the CI formula.
$\sigma_x^2 = \frac{1}{49} \left(159.4 - \frac{89^2}{50} \right) = 0.02 \left(= \frac{1}{50} \right)$	M1	One variance correct unsimplified.
$\sigma_y^2 = \frac{1}{39} \left(113.1 - \frac{67.2^2}{40} \right) = 0.0052308 \left(= \frac{17}{3250} \right)$	A1	Both correctly calculated.
$\sigma^2 = \frac{0.02}{50} + \frac{0.0052308}{40} = \left(\frac{69}{130000} \right)$	M1	Unsimplified expression must be seen, 40 and 50 in correct places.
$[\sigma^2] = 0.0005308$	A1	May be implied by correct final answer.
CI: $1.78 - 1.68 \pm z \times \sqrt{0.0005308}$	M1	Correct form for CI with a z-value.
with $z = 1.96$	A1	
0.10 ± 0.0452 or $[0.0548, 0.145]$	A1	Allow either form.
	8	SC Assuming equal variances award B1M0A0, M1A1 (for use of pooled variance formula, gives 0.0135), M1A1A0.



30 - (9231/41_Summer_2021_Q4) - Inference Using Normal And T-distributions

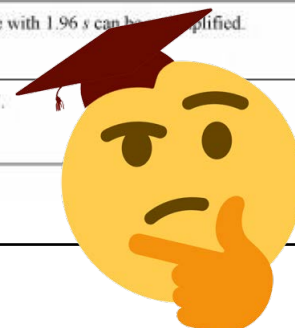
$s_x^2 = \frac{1}{49} \left(1596 - \frac{282^2}{50} \right) = 0.11265 \left[= \frac{138}{1225} \right]$	M1	One correct, unsimplified.
$s_y^2 = \frac{1}{59} \left(1808 - \frac{328^2}{60} \right) = 0.25311 \left[= \frac{224}{885} \right]$	A1	Both correct to 3sf.
$s^2 = \frac{0.11265}{50} + \frac{0.25312}{60} = 0.006471\dots$	M1 A1	May be implied.
$z = \frac{\frac{282}{50} - \frac{328}{60}}{\text{their } s}$	M1	
$ z = 2.155$	A1	Tolerance 2.15 – 2.16
Comparison with tabular value 1.96 OE $2.155 > 1.96$ Reject H_0	M1	Comparison with 1.96 OE and conclusion.
Sufficient evidence to reject population means are equal OR sufficient evidence to accept population means are not equal OR sufficient evidence to suggest $\mu_A \neq \mu_B$ OE	A1	Correct conclusion following correct work. Level of uncertainty in language is used.
	8	

31 - (9231/41_Winter_2021_Q1) - Inference Using Normal And T-distributions

(a)	$s^2 = \frac{78.5}{9} = 8.7222$	B1	$\frac{157}{18}$
	CI: $25.6 \pm t_{\alpha} \sqrt{\frac{s^2}{10}}$	M1	Correct expression with a t value.
	With $t = 1.833$	A1	With correct t value.
	$25.6 \pm 1.71(2)$ or $[23.9, 27.3]$	A1	Accept in either form, ISW. Accept inequality form.
		4	
(b)	$\frac{25.6 - k}{\sqrt{\frac{s^2}{10}}} \geq t$	M1	
	With $t = -1.383$	A1	Allow 1.383.
	26.9	A1	CAO
		3	

32 - (9231/42_Winter_2021_Q1) - Inference Using Normal And T-distributions

$s_x^2 = \frac{1}{39} \left(14320 - \frac{752^2}{40} \right) = 4.677 = \frac{304}{65}$	M1	One correct unsimplified.
$s_y^2 = \frac{1}{59} \left(40200 - \frac{1548^2}{60} \right) = 4.434 = \frac{1308}{295}$	A1	Both correct to at least 3sf.
$s^2 = \frac{4.6769}{40} + \frac{4.4339}{60}$	M1	
$0.1908 \left(= \frac{3659}{19175} \right)$	A1	
CI = $\left(\frac{752}{40} - \frac{1548}{60} \right) \pm 1.96s$	M1A1	Correct form with a z -value with 1.96 s can be simplified.
$-7.0 \pm 0.856 = [-7.86, -6.14]$	A1	Accept in either form, ISW. Accept $[6.14, 7.86]$. Accept inequality form.
	7	



33 - (9231/41_Winter_2021_Q4) - Inference Using Normal And T-distributions

$H_0: \mu_B - \mu_A = 2$ and $H_1: \mu_B - \mu_A > 2$	B1	Or use of μ_d .
Differences: 5 8 2 0 4 7 -2 1 10 -3	M1	Allow one error.
$\Sigma d = 32, \Sigma d^2 = 272$ $\bar{d} = 3.2, s^2 = \frac{1}{9} \left(272 - \frac{32^2}{10} \right) = 18.84 \left(= \frac{848}{45} \right)$	M1	Sample mean and variance.
$t = \frac{3.2 - 2}{\sqrt{\frac{s^2}{10}}} = 0.874$	M1 A1	
Compare with tabular value 1.833: $0.874 < 1.833$. Accept H_0 .	M1	Compare their value with 1.833 and conclusion.
Insufficient evidence to support claim.	A1	Correct conclusion, in context, following correct work. Level of uncertainty in language is used.
Assumption: population differences are normally distributed	B1	
	8	

34 - (9231/42_Winter_2021_Q6) - Inference Using Normal And T-distributions

$H_0: \mu_A = \mu_B$ and $H_1: \mu_A > \mu_B$	B1	
$\Sigma x = 21.2, \Sigma x^2 = 47.92$ $s_x^2 = \frac{1}{9} \left(47.92 - \frac{21.2^2}{10} \right) = 0.33067 = \frac{124}{375}$ $s_y^2 = \frac{1}{11} \left(53.75 - \frac{24.48^2}{12} \right) = 0.34644 = \frac{9527}{27500}$	M1 A1	
$s^2 = \frac{9 \times 0.33067 + 11 \times 0.34644}{10 + 12 - 2} = 0.3393$	M1 A1	
$t = \frac{2.12 - 2.04}{s \sqrt{\frac{1}{10} + \frac{1}{12}}} = 0.321$	M1 A1	Accept 0.321 – 0.322.
Critical value: 1.325 $0.321 < 1.325$ accept H_0 .	M1	Compare their value with 1.325 and correct FT conclusion.
Insufficient evidence that mean of A is greater than mean of B	A1	Correct conclusion, in context, following correct work. Level of uncertainty in language is used.
EITHER: Distributions are normal and equal variances OR: Distributions are normal	B1	Assumptions consistent with method used. Accept 'population is normal'. 'Distribution of difference of means is normal'.
Alternative method for question 6		
$H_0: \mu_A = \mu_B$ and $H_1: \mu_A > \mu_B$	B1	



$\Sigma x = 21.2 \quad \Sigma x^2 = 47.92$	M1	For one correct unsimplified.
$s_x^2 = \frac{1}{9} \left(47.92 - \frac{21.2^2}{10} \right) = 0.33067 = \frac{124}{375}$	A1	For both correct.
$s_y^2 = \frac{1}{11} \left(53.75 - \frac{24.48^2}{12} \right) = 0.34644 = \frac{9527}{27500}$		
$s^2 = \frac{0.33067}{10} + \frac{0.34644}{12} = 0.061936$	M1 A1	
$t = \frac{2.12 - 2.04}{s} = 0.321(5)$	M1 A1	Accept 0.321 – 0.322.
Critical value: 1.325 $0.321 < 1.325$ accept H_0 .	M1	Compare their value with 1.325 and correct FT conclusion.
Insufficient evidence that mean of A is greater than mean of B.	A1	Correct conclusion, in context, following correct work. Level of uncertainty in language is used.
EITHER: Distributions are normal and equal variances OR: Distributions are normal.	B1	Assumptions consistent with method used. Accept 'population is normal'. 'Distribution of difference of means is normal'.
	10	

35 - (9231/41_Summer_2022_Q1) - Inference Using Normal And T-distributions

(a)	Differences: 0.6 1.3 0.0 0.7 0.6 0.6 1.8 -0.2 -0.2 0.4	M1	Attempt at differences, allow one error.
	$\Sigma d = 5.6 \quad \Sigma d^2 = 6.74$ $s^2 = \frac{1}{9} \left(6.74 - \frac{5.6^2}{10} \right) = 0.400(4)$	M1	
	$H_0: \mu_B - \mu_A = 0.4$ and $H_1: \mu_B - \mu_A > 0.4$	B1	Allow use of μ_d
	$t = \frac{0.56 - 0.4}{\sqrt{\frac{0.4004}{10}}} (0.800)$	M1	
	$t = 0.7996 (0.800)$	A1	
	Critical value = $t_{0.90}(9) = 1.383$ Compare: ' $0.7996 < 1.383$ ' Accept H_0	M1	Compare calculated value with 1.383 and correct FT conclusion.
	Insufficient evidence to support manager's claim.	A1	Correct conclusion, in context. Level of uncertainty in language is used.
		7	
(b)	Distribution of population differences is normal	B1	
		1	

36 - (9231/43_Summer_2022_Q1) - Inference Using Normal And T-distributions

$s^2 = \frac{94.5}{7} (=13.5)$	B1	
CI: $33.8 \pm 2.365 \sqrt{\frac{s^2}{8}}$	M1	Correct formula with a t value, must be 33.8.
	B1	2.365 used.
$33.8 \pm 3.072 = [30.7, 36.9]$	A1	Either form.
	4	

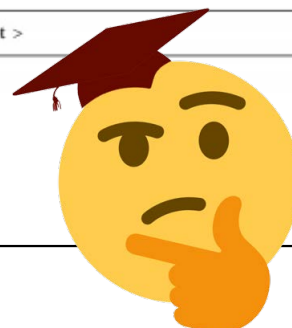


37 - (9231/41_Summer_2022_Q5) - Inference Using Normal And T-distributions

(a)	$\Sigma x = 44.4, \Sigma x^2 = 247.48$	B1	
	$s^2 = \frac{1}{7} \left(247.48 - \frac{44.4^2}{8} \right) = 0.1514 \left(= \frac{53}{350} \right)$	M1	
	CI: $\frac{44.4}{8} \pm 1.895 \sqrt{\frac{0.1514}{8}}$	M1	Correct formula with a t value.
	1.895	B1	1.895 used.
	$5.55 \pm 0.261 = [5.29, 5.81]$	A1	Either form.
		5	
(b)	$H_0: \mu = 5.9$ and $H_1: \mu < 5.9$	B1	
	$t = \frac{5.55 - 5.9}{\sqrt{\frac{s^2}{8}}} = -2.544$	M1	
	Critical value $t_{0.975}(7) = 2.365$ Compare $-2.544 < -2.365$, reject H_0 .	M1	Compare calculated value with -2.365 and correct FT conclusion.
	Sufficient evidence to support Raman's claim.	A1	Correct conclusion, in context. Level of uncertainty in language is used.
		4	

38 - (9231/43_Summer_2022_Q6) - Inference Using Normal And T-distributions

(a)	$s_x^2 = \frac{1}{39} \left(31400 - \frac{1120^2}{40} \right) = 1.0256$ $s_y^2 = \frac{1}{49} \left(37600 - \frac{1370^2}{50} \right) = 1.2653$	M1	$\frac{40}{39}$ $\frac{62}{49}$ Both required
	$s^2 = \frac{1.0256}{40} + \frac{1.2653}{50}$	M1	$\frac{2434}{47775}$ Condone $s^2 = \frac{1.03}{40} + \frac{1.27}{50}$ or 0.05115.
	$s^2 = 0.050947$	A1	May be implied by 2.66 as z value.
	$H_0: \mu_a - \mu_b = 0$ and $H_1: \mu_a - \mu_b \neq 0$	B1	
	$z = \frac{\frac{1120}{40} - \frac{1370}{50}}{s} = \frac{0.6}{\sqrt{0.050947}}$	M1	
	$z = 2.66$	A1	
	Critical value of $z = 2.326$ $2.66 > 2.326$ Reject H_0	M1	Compare their z value with 2.326 and correct FT conclusion.
	Sufficient evidence to support difference in population means oe	A1	Correct conclusion, in context. Level of uncertainty in language is used.
		8	
(b)	$z = \frac{0.6 - 0.25}{s}$	M1	Condone 1.54.
	$z = 1.55$	A1	
	$0.9392 - 0.9395$ or $0.0605 - 0.0608$	A1	
	$\alpha \geq 6.05$	A1	Condone 6.05 - 6.08, accept >
		4	



39 - (9231/41_Winter_2022_Q1) - Inference Using Normal And T-distributions

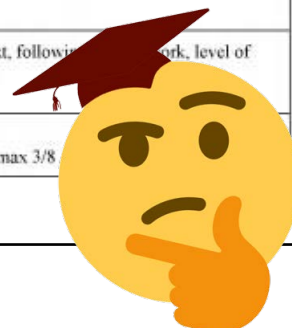
$\bar{x} = \frac{1625}{50} = 32.5$ $s_x^2 = \frac{1}{49} \left(53200 - \frac{1625^2}{50} \right) = 7.908$	M1	$\frac{775}{98}$
$\bar{y} = \frac{1854}{60} = 30.9$ $s_y^2 = \frac{1}{59} \left(57900 - \frac{1854^2}{60} \right) = 10.3627$	A1	$\frac{3057}{295}$ Both correct
$s^2 = \frac{7.908}{50} + \frac{10.3627}{60} [= 0.33087]$	M1 A1	
CI: $(32.5 - 30.9) \pm 1.96s$	M1 A1	Correct form with a z-value. Correct with 1.96.
$1.6 \pm 1.1274 = [0.473, 2.73]$	A1	At least 3sf.
	7	Pooled variance used: M1A1 M0A0 M1A1A0 max 4/7

40 - (9231/42_Winter_2022_Q1) - Inference Using Normal And T-distributions

$2.02 = \bar{x} + 1.833\sqrt{\frac{s^2}{10}}$ $1.78 = \bar{x} - 1.833\sqrt{\frac{s^2}{10}}$ Add, $\bar{x} = \frac{2.02 + 1.78}{2} = 1.90$	M1	Add.
$\Sigma x = 19$	A1	
Subtract: $2.02 - 1.78 = 2 \times 1.833\sqrt{\frac{s^2}{10}}$	M1	Allow 1.372, 1.383, 1.812 instead of 1.833.
$s^2 = 0.042859$	A1	May be implied. $s = 0.20702$
But $s^2 = \frac{1}{9} \left(\Sigma x^2 - \frac{(\Sigma x)^2}{10} \right)$	M1	
$\Sigma x^2 = 9s^2 + \frac{19^2}{10} = 36.5$	A1	36.486
		Using 1.645; maximum M1A1 M0A0 M1A0 3/6
	6	

41 - (9231/42_Winter_2022_Q3) - Inference Using Normal And T-distributions

$s_x^2 = \frac{1}{49} \left(115.2 - \frac{75.5^2}{50} \right) = 0.02439$ $s_y^2 = \frac{1}{79} \left(172.6 - \frac{116.8^2}{80} \right) = 0.02623$	M1	$\frac{239}{9800}$ $\frac{259}{9875}$
	A1	Both correct
$s^2 = \frac{0.02439}{50} + \frac{0.02623}{80} [= 0.00081565]$	M1 A1	May be implied by 1.75 for z
$z = \frac{\frac{75.5}{50} - \frac{116.8}{80}}{s}$	M1	
1.75	A1	
Compare with 1.645: $1.75 > 1.645$ Reject null hypothesis	M1	Using areas, $0.04 < 0.05$.
Sufficient evidence to suggest that population mean in country X is greater than population mean in country Y	A1	Correct conclusion in context, following correct work, level of uncertainty in language. 'Prove' gives A0.
	8	Pooled variance: $z = 1.74$ M1A1 M0A0M0A0 M1A0 max 3/8



42 - (9231/41_Winter_2022_Q6) - Inference Using Normal And T-distributions

$s_x^2 = \frac{1}{11} \left(3.26 - \frac{6.24^2}{12} \right) = 0.001382$ $s_y^2 = \frac{1}{9} \left(2.6124 - \frac{5.1^2}{10} \right) = 0.001267$	M1	$\frac{19}{13750}$ $\frac{19}{15000}$
	A1	Both correct.
$s^2 = \frac{11 \times 0.001382 + 9 \times 0.001267}{12 + 10 - 2} = 0.001330$	M1 A1	Find pooled variance.
$t = \frac{\frac{6.24}{12} - \frac{5.1}{10}}{s \sqrt{\frac{1}{12} + \frac{1}{10}}} = 0.640$	M1 A1	Use correct formula for t .
$H_0: \mu_a = \mu_b \quad H_1: \mu_a > \mu_b$	B1	
Critical value is 2.086 ($t_{0.975}(20)$) $0.640 < 2.086$ Accept H_0	M1	Compare their calculated value with 2.086 and conclusion.
There is insufficient evidence to support inspector's claim	A1	All correct except possibly B1, in context, level of uncertainty in language. 'Prove' scores A0
	9	Pooled variance not used: M1A1 M0A0 M0A0 B1 M1A0 max 4/9

43 - (9231/41_Summer_2023_Q2) - Inference Using Normal And T-distributions

$\bar{x} = \frac{99.6}{12} [= 8.3] \quad s_x^2 = \frac{21.5}{11} [= 1.9545 \text{ or } \frac{43}{22}]$	M1	Both SOL
$H_0: \mu_P = \mu_Q \quad H_1: \mu_P > \mu_Q$	B1	
Pooled estimate = $\frac{9 \times 2.16 + 11 \times 1.9545}{10 + 12 - 2}$	M1 FT	FT their 1.9545 only
2.047 (2.05)	A1	May be implied, allow 2.04
$t = \frac{9.12 - 8.3}{\sqrt{2.047} \times \sqrt{\frac{1}{10} + \frac{1}{12}}}$	M1 FT	FT their pooled estimate
1.339 (or 1.338)	A1	Accept 1.34
'1.339' < 1.725 Accept H_0 / not significant	M1	Compare with correct tabular value 1.725 and conclusion without context. Condone 'reject H_0 '.
Insufficient evidence to suggest that the (mean) time taken at P is greater than the (mean) time taken at Q.	A1	Correct conclusion in context, following correct work, level of uncertainty in language (not 'prove', not 'there is no evidence'), no contradictions. A0 if H_0 and H_1 reversed.
	8	



44 - (9231/43_Summer_2023_Q4) - Inference Using Normal And T-distributions

(a)	$H_0: \mu_x = \mu_y$ $H_1: \mu_x < \mu_y$	B1	
	$s_x^2 = \frac{1}{79} \left(338.1 - \frac{164^2}{80} \right) [= 0.02405]$ and $s_y^2 = \frac{1}{59} \left(261.1 - \frac{124.8^2}{60} \right) [= 0.02569]$	B1	Both. Implied by $\frac{19}{790}, \frac{379}{14750}$ or 3sf.
	$s^2 = \frac{0.02405}{80} + \frac{0.02569}{60}$	M1	
	$s^2 = 0.0007289$ or $s = 0.026998$ or 0.0270	A1	Implied by $z = -1.11$
	$z = \frac{\frac{164.0}{80} - \frac{124.8}{60}}{s}$	M1	FT <i>their</i> value for s .
	-1.11	A1	
	'1.11' < 1.282 Accept H_0 / not significant	M1	Compare with correct z -value 1.282 and consistent signs. Condone 'reject H_1 '. Using probabilities, $P(Z > 1.11) = 0.1333 > 0.1$.
	Insufficient evidence to support the inspector's suspicion/ Insufficient evidence that the (mean) lengths of rods from machine X are shorter than the (mean) lengths of rods from machine Y	A1	Correct conclusion in context, following correct work, level of uncertainty in language (not 'prove', not 'there is no evidence'), no contradictions. e.g. proves that the inspector is incorrect scores A0 A0 if hypotheses wrong way round.
		8	
(b)	Large sample sizes OR central limit theorem applies.	B1	
		1	

45 - (9231/41_Winter_2023_Q1) - Inference Using Normal And T-distributions

	$H_0: \mu = 4.22$ $H_1: \mu < 4.22$	B1	
	$\bar{x} = \frac{42.05}{10} [= 4.205]$, $s^2 = \frac{1}{9} \left(176.83 - \frac{42.05^2}{10} \right) [= \frac{13}{12000} = 0.0010833]$	M1	
	$t = \frac{4.205 - 4.22}{\sqrt{\frac{s^2}{10}}}$	M1	<i>Their</i> 4.205, <i>their</i> s^2 .
	$t = -1.44$	A1	Condone sign.
	Tabular value = 1.833: '1.44' < 1.833, accept H_0	M1	Compare <i>their</i> value with correct tabular value, signs consistent , allow 'not significant'.
	There is insufficient evidence to support Maya's belief.	A1	Correct conclusion in context, following correct work, level of uncertainty in language. A0 if hypotheses wrong way round or missing.
		6	



46 - (9231/43_Winter_2023_Q3) - Inference Using Normal And T-distributions

(a)	$H_0: \mu_D - \mu_E = 2 \quad H_1: \mu_D - \mu_E > 2$										B1	May use μ_d , but must be consistent with direction of differences found.
	8	8	15	-2	-3	20	-3	0	5	4	M1	Differences, at most 2 errors.
	$\Sigma d = 52, \Sigma d^2 = 816, s^2 = \frac{1}{9} \left(816 - \frac{52^2}{10} \right) \left[= \frac{2728}{45} = 60.62 \right]$										M1	<i>Their</i> values but must see 9 and 10 used correctly.
	$t = \frac{\frac{52}{10} - 2}{\sqrt{\frac{s^2}{10}}}$										M1	
	$t = 1.29967, 1.30$										A1	
	Tabular value = 2.262. '1.30' < 2.262, accept H_0										M1	Allow 'not significant'.
	Insufficient evidence to suggest that cholesterol level has reduced by more than 2.										A1	CWO. Correct conclusion in context, following correct work, level of uncertainty in language. A0 if hypotheses wrong way round or missing.
											7	
(b)	Population differences are normally distributed										B1	
											1	



1 - (9231/23_Summer_2017_Q10) - X2 Test

(i)	$\bar{x} = (1/100) \sum x f(x) = 325/100 = 3.25$ AG	B1
	Total:	1
(ii)	$E_3 = 100 \lambda^3 e^{-\lambda} / 3! \text{ with } \lambda = 3.25$	M1 A1
	$E_6 = 100 \lambda^6 e^{-\lambda} / 6! = 6.35$	A1
	$E_{\geq 7} = 100 - \sum_0^6 E_i = 4.77$	A1
	Total:	4
(iii)	H_0 : Distribution fits data (AEF)	B1
	O_i : <u>13</u> 18 26 20 16 <u>7</u> E_i : <u>16.48</u> 20.48 22.18 18.02 11.72 <u>11.12</u>	M1FT
	$\chi^2 = 0.735 + 0.300 + 0.658 + 0.218 + 1.563 + 1.527$	M1
	$= 5.00$	A1
	No. n of cells: 8 7 <u>6</u> 5 4 $\chi_{n-2, 0.95}^2$: 12.59 11.07 <u>9.488</u> 7.815 5.991	B1FT
	Accept H_0 if $\chi^2 < \text{tabular value}$ (AEF) $5.00 [\pm 0.1] < 9.49$ so distn. fits [data]	M1
	or distn. is a suitable model (AEF)	A1
	Total:	7



2 - (9231/21_Summer_2017_Q11) - Simple Harmonic Motion, X2 Test

(a)(i)	$T = 3mg \sin \alpha \quad [= 2mg]$	B1
	$T = kmg (5a/4 - a) / a \quad [= \frac{1}{4} kmg]$	B1
	$k = 8$	B1
	Total:	3
(a)(ii)	<i>EITHER:</i> $\pm 2m \, d^2 OQ / dt^2 = 2mg \sin \alpha - kmg (OQ - a) / a$	(M1 A1)
	$d^2 OQ / dt^2 = (4g/a) (7a/6 - OQ)$	A1
	$d^2 x / dt^2 = - (4g/a) x$ where $x = OQ - 7a/6$	A1)
	<i>OR:</i> $2mg \sin \alpha = kmg (e - a) / a, e = 7a/6$	(M1
	$\pm 2m \, d^2 x / dt^2 = 2mg \sin \alpha - kmg (e + x - a) / a$	M1 A1
	$d^2 x / dt^2 = - (4g/a) x$	A1)
	Centre is $7a/6$ (or $1.17 a$) from O	B1
	Period is $\pi \sqrt{a/g}$ or $0.993 \sqrt{a}$	B1
	Total:	6
(a)(iii)	$x_0 = 5a/4 - e = a/12$	B1
	$T_{min} = kmg (5a/4 - 2x_0 - a) / a = 2mg/3$	M1 A1
	$(d^2 x / dt^2)_{max} = [\pm] (4g/a) x_0 = [\pm] \frac{1}{3} g$	M1 A1
	Total:	5



(b)(i)	$x = (1/250) \sum x f(x) = 414/250 = 1.656$ AG	B1
	Total:	1
(b)(ii)	$p = x/6 = 0.276, q = 0.724$	M1 A1
	$a = 250 {}^6C_2 q^4 p^2 = 78.49 \pm 0.01$ (to 2 d.p.)	A2
	$b = 250 {}^6C_4 q^2 p^4 = 11.41 \pm 0.01$ (to 2 d.p.)	A1
	Total:	5
(b)(iii)	H_0 : Distribution fits data <i>or</i> distribution is binomial (AEF)	B1
	O_i : 48 69 78 32 23 E_i : 36.01 82.36 78.49 39.89 <u>13.26</u> (± 0.01)	M1FT A1
	$\chi^2 = 3.992 + 2.167 + 0.003 + 1.561 + 7.154$	M1
	$= 14.9$	A1
	No. n of cells: 7 6 <u>5</u> 4 3 $\chi_{n-2, 0.99}^2$: 15.09 13.28 <u>11.34</u> 9.210 6.635	B1FT
	Accept H_1 if $\chi^2 >$ tabular value (AEF) $14.9 [\pm 0.1] > 11.34$ so distrn. doesn't fit [data]	M1
	<i>or</i> manager's belief not justified (AEF)	A1
	Total:	8

3 - (9231/21_Winter_2017_Q8) - X2 Test

H_0 : No association <i>or</i> method is independent of gender (AEF)	B1	State (at least) null hypothesis
E_i : 9.0 19.8 31.2 6.0 13.2 20.8	M1 A1	Find expected values E_i (A0 if rounded to integers)
$\chi^2 = 0.111 + 3.073 + 1.482$ $+ 0.167 + 4.610 + 2.223$ $= 11.7$ (or 12.3) (to 2 d.p.)	M1 A1	Find value of χ^2 from $\sum (E_i - O_i)^2 / E_i$ [or $\sum O_i^2 / E_i - n$] (allow 12.3 if integer values of E_i used)
$\chi_{2, 0.99}^2 = 9.21$	B1	State or use correct tabular χ^2 value
Reject H_0 if $\chi^2 >$ tabular value (AEF)	M1	Valid method for reaching conclusion
$11.7 [\pm 0.1] > 9.21$ so there is an association	A1	Correct conclusion, from correct values
	8	



4 - (9231/21_Summer_2018_Q8) - X2 Test

H_0 : Colour is independent of type or no association between colour and type (AEF)	B1	State (at least) null hypothesis in full
E_i : 47.67 43.33 39 36.67 33.33 30 25.67 23.33 21 (to 3 s.f.)	M1 A1	Find expected values E_i (A0 if rounded to integers)
$\chi^2 = 0.597 + 1.241 + 0.103$ $+ 1.603 + 1.333 + 0.033$ $+ 0.212 + 0.019 + 0.429 = 5.57$ (to 3 s.f.)	M1 A1	Find value of χ^2 from $\Sigma (E_i - O_i)^2 / E_i$ [or $\Sigma O_i^2 / E_i - n$] (allow 5.64 for this A1 if integer values of E_i used)
$\chi_{4,0.05}^2 = 7.779$ or 7.78	B1	State or use correct tabular χ^2 value
Accept H_0 if $\chi^2 <$ tabular value (AEF)	M1	Valid method for reaching conclusion
$5.57 [\pm 0.01] < 7.78$ so independent or no association (AEF)	A1	Correct (abbreviated) conclusion, from approx. correct values
	8	

5 - (9231/22_Winter_2018_Q10) - X2 Test

(i)	$\bar{x} = \Sigma x f(x) = 118/40 = 2.95$ AG	B1	Verify mean of sample data (B0 for $\bar{x} = 118/40 = 2.95$)
	$\sigma_n^2 = (454 - 118^2/40) / 40 = 2.65$ or $\sigma_{n-1}^2 = (454 - 118^2/40) / 39 = 2.72$	B1	Find variance of sample data (accept either σ_n^2 or σ_{n-1}^2)
	$2.95 \approx 2.65$ or 2.72	B1	Valid comment on correct values
		3	
(ii)	$E_4 = 40 \lambda^4 e^{-\lambda} / 4!$ with $\lambda = 2.95$ [= 40×0.1652] (allow $(\lambda/4) \times E_3 = 0.7375 \times 8.96$)	M1A1	State expression for reqd. expected value E_4
		2	
(iii)	H_0 : [Poisson] distribution fits data (AEF)	B1	State (at least) null hypothesis in full
	O_i : E_i : 8 8 10 5 9 8.27 9.11 8.96 6.61 7.05	M1	Combine values consistent with all exp. values ≥ 5
	$\chi^2 = 0.0088 + 0.1352 + 0.1207 + 0.3921 + 0.5394$ $= 1.20$ (to 3 s.f.)	M1	Find value of χ^2 from $\Sigma (E_i - O_i)^2 / E_i$ [or $\Sigma O_i^2 / E_i - n$]
		A1	
	No. of cells: 8 7 6 5 $\chi_{0.2,0.95}^2$: 12.59 11.07 9.488 7.815	B1	State or use consistent tabular value $\chi_{0.2,0.95}^2$ (to 3 s.f.) [FT on number, n , of cells used to find χ^2]
	Accept H_0 if $\chi^2 <$ tabular value (AEF)	M1	State or imply valid method for conclusion
	$1.20 [\pm 0.1] < 7.81[5]$ so [Poisson] distn. fits [data] or distn. is a suitable model (AEF)	A1	Conclusion (requires both values correct)
		7	



6 - (9231/21_Winter_2018_Q11) - Circular Motion, X2 Test

A(i)	$\frac{1}{2}mv_C^2 = \frac{1}{2}mu^2 + mga(1 - \cos \alpha)$ [$v_C^2 = u^2 + (2/5)ag$]	M1A1	Find speed v_C at C by conservation of energy (A0 if no m)
	$R_A = mg \cos \alpha + ma^2/a$	B1	Find reaction R_A at A by using $F = ma$
	$R_C = mg + mv_C^2/a$	B1	Find reaction R_C at C by using $F = ma$
	$8\{ag + u^2 + (2/5)ag\} = 9\{(4/5)ag + u^2\}$	M1	Verify u^2 using $R_C/R_A = 9/8$
	$u^2 = (8 + 16/5 - 36/5)ag = 4ag$ AG	A1	
		6	
A(ii)	$\frac{1}{2}mv_B^2 = \frac{1}{2}mu^2 - 2mga \cos \alpha$ or $\frac{1}{2}mv_C^2 = mga(1 + \cos \alpha)$ [$v_C^2 = (22/5)ag$]	M1	Find speed v_B at B by conservation of energy (A0 if no m)
	$v_B^2 = (4 - 16/5)ag$ or $(22/5 - 18/5)ag = (4/5)ag$	*A1	
	≥ 0 so P reaches height of B (AEF; dep *A1)	A1	(or valid argument based on R_A)
	$R_A = mv_B^2/a - mg \cos \alpha = 4mg/5 - 4mg/5 = 0$	M1**A1	Find reaction R_B at B by using $F = ma$
	≥ 0 so P still [just] in contact at B (AEF; dep **A1)	A1	
		6	
B(i)	$\bar{x} = (1/200) \sum x f(x) = 30\,064 / 200 = 150.3[2]$	M1A1	Find sample mean to 4 s.f. using mid-interval values (B1 for $29\,964 / 200 = 149.8[2]$ or $30\,164 / 200 = 150.8[2]$)
		2	

B(ii)	$P(151 \leq X < 152)$ $= P((151 - 150)/1.2 \leq Z < (152 - 150)/1.2)$ $= P(0.833 \leq Z < 1.667)$ $= 0.9522 - 0.7975 [= 0.1547] \text{ (to 4 s.f.)}$	M1A1	Find $P(151 \leq X < 152)$
	$E_{151} = 190.44 - 159.5$ or $200 \times 0.1547 = 30.94$ AG	A1	and hence expected frequency for $151 \leq x < 152$
		3	
B(iii)	H_0 : [Normal] distribution fits data (AEF)	B1	State (at least) null hypothesis in full
	O_i : 3 23 52 69 36 17	M1	Combine values consistent with all exp. values ≥ 5
	E_i : <u>9.56</u> 30.94 59.50 59.50 30.94 <u>9.56</u>		
	$\chi^2 = 4.501 + 2.038 + 0.945 + 1.517 + 0.828 + 5.790$	M1	Find value of χ^2 from $\sum (E_i - O_i)^2 / E_i$ [or $\sum O_i^2 / E_i - n$]
	$= 15.6$ (to 3 s.f.)	A1	
	No. n of cells: 8 7 6 5 $\chi_{0.05}^2$: 14.07 12.59 <u>11.07</u> 9.488	B1	State or use consistent tabular value $\chi_{0.05}^2$ (to 3 s.f.) [FT on number, n , of cells used to find χ^2]
	Accept H_1 if $\chi^2 >$ tabular value (AEF)	M1	State or imply valid method for conclusion
	$15.6 [\pm 0.1] > 11.1$ so distr. doesn't fit [data] (AEF)	A1	Conclusion (requires both values correct)
		7	



7 - (9231/23_Summer_2019_Q8) - X2 Test

H_0 : Holidays are independent of salesman or no association between holidays and salesman (AEF)	B1	State (at least) null hypothesis Find expected values (lose A1 if rounded to integers) Find value of χ^2 from $\Sigma (E_i - O_i)^2 / E_i$ [or $\Sigma O_i^2 / E_i - n$] State or use correct tabular χ^2 value
E_i : 29.68 33.04 21.28 23.32 25.96 16.72 (to 1 d.p.)	M1 A1	
χ^2 = 0.7380 + 0.7446 + 0.0037 + 0.9392 + 0.9477 + 0.0047	M1	
= 3.38 (to 3 s.f.)	A1	
$\chi_{2,0.95}^2$ = 4.605 (to 3 s.f.)	B1	
Accept H_0 if $\chi^2 <$ tabular value	M1	Compare their calculated value with their χ^2 value and appropriate conclusion
Type of holidays is independent of salesman (AEF)	A1	Correct conclusion, from correct values (3.37 – 3.39)
	8	

8 - (9231/21_Summer_2019_Q9) - X2 Test

(i)	$E_3 = (3/16) \int_{1/8}^{2/4} (4-x)^{1/2} dx$ $= (3/16) [-(2/3)(4-x)^{3/2}]_{1/8}^{2/4}$	M1	State or imply expression for required expected value E_3 of X
	$= (2 \cdot 4^{3/2} - 1 \cdot 6^{3/2}) / 8 = 1.694/8$ or 0.2118	A1	Find expected value E_3 (may be implied in finding 50 E_3) (M1 A1 requires adequate explicit working)
	50 $E_3 = 10.59$ AG	A1	Hence verify corresponding expected frequency
		3	
(ii)	H_0 : Distribution fits data (AEF)	B1	State (at least) null hypothesis in full
	O_i 18 16 8 8 E_i : 14.22 12.54 10.59 12.65	M1	Combine values consistent with all exp. values ≥ 5
	χ^2 = 1.005 + 0.955 + 0.633 + 1.709 = 4.30	M1 A1	Find value of χ^2 from $\Sigma (E_i - O_i)^2 / E_i$ [or $\Sigma O_i^2 / E_i - n$]
	No. n of cells: 5 4 3 $\chi_{n-1,0.95}^2$: 9.488 7.815 5.991 (to 3 s.f.)	B1	FT on number, n , of cells used to find χ^2 State or use consistent tabular value $\chi_{n-1,0.95}^2$
	Accept H_0 if $\chi^2 <$ tabular value (AEF)	M1	State or imply valid method for conclusion
	4.30 [± 0.01] < 7.81[5] so distribution fits [data] or distribution is a suitable model (AEF)	A1	Conclusion (requires both values approx. correct)
		7	



9 - (9231/21_Winter_2019_Q11) - Simple Harmonic Motion, X2 Test

A(i)	$10 (AP - 0.6) / 0.6 = 20 (1.2 - AP - 0.4) / 0.4$	M1 A1	Verify AP by equating equilibrium tensions. AEF
	$4 AP - 2.4 = 9.6 - 12 AP$ $AP = 0.75 \text{ [m]}$	A1	AG
		3	
A(ii)	$m \frac{d^2x}{dt^2} = -10 (0.15 + x) / 0.6 + 20 (0.05 - x) / 0.4$ or $m \frac{d^2x}{dt^2} = +10 (0.15 - x) / 0.6 - 20 (0.05 + x) / 0.4$	M1 A1 A1	Apply Newton's law at $0.75 + x$ or $0.75 - x$ from A (M1 requires LHS and 2 tensions: A1 for each correct tension)
	$\frac{1}{2} \frac{d^2x}{dt^2} = -(80 / 1.2) x$, $\frac{d^2x}{dt^2} = -100 x$	M1 A1	Simplify to give SHM eqn. in standard form
	$T = 2\pi/\omega = 2\pi/10 = \pi/5$ or 0.628 [s]	DB1	State the period T with FT on ω from SHM eqn.
		6	
A(iii)	$(a = 0.75 - 0.7 = 0.05)$ $v_{\max} = \omega \times a$	M1	Find speed at equilibrium position from ωa
	$v_{\max} = 0.5 \text{ [m s}^{-1}\text{]}$	A1	
		2	
A(iv)	$x = a/2 = 0.025$	M1	Find value of x giving half max. acceln.
	$v = \omega \sqrt{(a^2 - x^2)} = 10 \sqrt{(0.05^2 - 0.025^2)}$	M1	
	$v = 0.433 \text{ [m s}^{-1}\text{]}$	A1	Find corresponding speed
		3	
B(i)	$\bar{x} = (1/40) \sum x f(x) = 68/40 = 1.7$	B1	Find mean of sample
	$(1/40) \sum x^2 f(x) = 178/40 = 4.45$, $\text{Var} = 4.45 - 1.7^2 = 1.56$	B1	Find variance of sample
	Mean and variance are similar so Poisson may be suitable	B1	State valid comment
		3	
B(ii)	$a = 40 \times 1.6^3 e^{-1.6} / 5! = 40 \times 0.01764$	M1	AG
	$a = 0.706$	A1	Verify a from Poisson term
	$b = 40 - 39.758 = 0.242$	B1*	Find b
		3	
B(iii)	H_0 : Distribution fits/models data	B1	State (at least) null hypothesis in full
	O_i : 6 15 9 10 E_i : 8.076 12.921 10.337 <u>8.666</u>	DM1 A1	Combine values consistent with all exp. values ≥ 5 (FT on b , dep B1*)
	$\chi^2 = 0.5337 + 0.3345 + 0.1729 + 0.2053 = 1.25$	M1	Find value of χ^2 from $\sum (E_i - O_i)^2 / E_i$ [or $\sum O_i^2 / E_i - n$]
	$\chi^2 = 1.25$	A1	
	No. n of cells: 7 6 5 4 3 $\chi_{0.1, 0.9}^2$: 10.64 9.236 7.779 <u>6.251</u> 4.605	DB1	State or use consistent tabular value $\chi_{0.1, 0.9}^2$ (to 3 s.f.) [FT on number, n , of cells used to find χ^2]
	Accept H_0 if $\chi^2 < \text{tabular value}$ (using their values)	M1	AEF
	$1.25 [\pm 0.1] < 6.25$ so distn. fits [data] or distn. is a suitable model	A1	Conclusion (requires both values approx. correct). AEF
		8	



10 - (9231/41_Summer_2020_Q1) - X2 Test

	<table> <tr> <td>49 E = 56.8</td><td>51 E = 47.6</td><td>40 E = 35.6</td></tr> <tr> <td>93 E = 85.2</td><td>68 E = 71.4</td><td>49 E = 53.4</td></tr> </table>	49 E = 56.8	51 E = 47.6	40 E = 35.6	93 E = 85.2	68 E = 71.4	49 E = 53.4	M1A1
49 E = 56.8	51 E = 47.6	40 E = 35.6						
93 E = 85.2	68 E = 71.4	49 E = 53.4						
	$\frac{(49-56.8)^2}{56.8} + \frac{(51-47.6)^2}{47.6} + \frac{(40-35.6)^2}{35.6} + \frac{(93-85.2)^2}{85.2} + \frac{(68-71.4)^2}{71.4} + \frac{(49-53.4)^2}{53.4}$	M1						
	= 3.096 (3.10)	A1						
	Use appropriate tabular value = 4.605	M1						
	3.096 < 4.605 so the grades awarded are independent of the background	A1						
		6						

11 - (9231/43_Summer_2020_Q1) - X2 Test

	<table> <tr> <td>31 E = 22</td><td>35 E = 37.4</td><td>22 E = 28.6</td></tr> <tr> <td>19 E = 28</td><td>50 E = 47.6</td><td>43 E = 36.4</td></tr> </table>	31 E = 22	35 E = 37.4	22 E = 28.6	19 E = 28	50 E = 47.6	43 E = 36.4	M1A1
31 E = 22	35 E = 37.4	22 E = 28.6						
19 E = 28	50 E = 47.6	43 E = 36.4						
	$\frac{(31-22)^2}{22} + \frac{(35-37.4)^2}{37.4} + \frac{(22-28.6)^2}{28.6} + \frac{(19-28)^2}{28} + \frac{(50-47.6)^2}{47.6} + \frac{(43-36.4)^2}{36.4}$	M1						
	9.569 (9.57)	A1						
	Tabular value = 5.991 and comparing 9.569 > 5.991	M1						
	So standard achieved is dependent on reading scheme used	A1						
		6						

12 - (9231/41_Winter_2020_Q3) - X2 Test

<table><tr><td>x</td><td>0</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td></tr><tr><td>Observed freq</td><td>12</td><td>39</td><td>46</td><td>37</td><td>12</td><td>4</td></tr><tr><td>Bin prob</td><td>0.11603</td><td>0.31236</td><td>0.33642</td><td>0.18115</td><td>0.04877</td><td>0.00525</td></tr><tr><td>Expected freq</td><td>17.404</td><td>46.858</td><td>50.462</td><td>27.172</td><td>7.316</td><td>0.7878</td></tr></table>	x	0	1	2	3	4	5	Observed freq	12	39	46	37	12	4	Bin prob	0.11603	0.31236	0.33642	0.18115	0.04877	0.00525	Expected freq	17.404	46.858	50.462	27.172	7.316	0.7878	M1 A1	Attempt at E values (at least 4 correct) All correct, to 2 dp or better
x	0	1	2	3	4	5																								
Observed freq	12	39	46	37	12	4																								
Bin prob	0.11603	0.31236	0.33642	0.18115	0.04877	0.00525																								
Expected freq	17.404	46.858	50.462	27.172	7.316	0.7878																								
Add last two columns: 16, 8.104	M1																													
$\frac{(17.404 - 12)^2}{17.404} + \frac{(46.858 - 39)^2}{46.858} + \frac{(50.462 - 46)^2}{50.462} + \frac{(27.172 - 37)^2}{27.172} + \frac{(8.104 - 16)^2}{8.104} = 14.64 \text{ or } 14.65$	M1 A1	Accept 14.6 – 14.7																												
'14.64' > 9.49	M1	Compare <i>their</i> value with 9.49																												
Freya's claim is not supported or Data does not fit the distribution	A1 FT	Correct conclusion in context, FT only <i>their</i> 14.64 SC for Poisson M0M1M1M0 max 2																												
	7																													



13 - (9231/42_Winter_2020_Q3) - X2 Test

(a)	$\int_0^{0.5} \left(\frac{4}{9}x - \frac{1}{9}x^2 \right) dx \quad I = \frac{1}{9} \left[2 \times 0.5^2 - \frac{1}{3} 0.5^3 \right]$ $(= 0.0509(26) \text{ or } \frac{11}{216})$	M1	Statement of integration with correct limits														
	Freq = $p - 0.0509(26) \times 200 = 10.19$ AG	A1	Requires sight of 0.0509 or $\frac{11}{216}$														
	By addition to 200, $q = 26.85 \quad \left(\frac{725}{27} \right)$	B1	Or by integration														
		3															
(b)	<table border="1"><tr><td>O</td><td>5</td><td>23</td><td>40</td><td>41</td><td>46</td><td>45</td></tr><tr><td>E</td><td>10.19</td><td>26.85</td><td>37.96</td><td>43.52</td><td>43.52</td><td>37.96</td></tr></table> Test statistic = $2.6433 + 0.5520 + 0.1096 + 0.1459 + 0.1413 + 1.3056$	O	5	23	40	41	46	45	E	10.19	26.85	37.96	43.52	43.52	37.96	M1	At least 4 correct
	O	5	23	40	41	46	45										
	E	10.19	26.85	37.96	43.52	43.52	37.96										
	4.89 or 4.90	A1															
	'4.90' < 11.07:	M1	Compare their value with 11.07														
	PDF is a satisfactory model for the data.	A1 FT	Correct conclusion in context, ft only their 4.90														
	4																

14 - (9231/41_Summer_2021_Q2) - X2 Test

<table><tr><td>72 (61.425)</td><td>42 (44.46)</td><td>52 (54.405)</td><td>68 (73.71)</td></tr><tr><td>33 (43.575)</td><td>34 (31.54)</td><td>41 (38.595)</td><td>58 (52.29)</td></tr></table>	72 (61.425)	42 (44.46)	52 (54.405)	68 (73.71)	33 (43.575)	34 (31.54)	41 (38.595)	58 (52.29)	M1	At least 3 correct to 4sf.
72 (61.425)	42 (44.46)	52 (54.405)	68 (73.71)							
33 (43.575)	34 (31.54)	41 (38.595)	58 (52.29)							
Calculates E values (in brackets above)	A1	All correct to 4sf (if 6.037 not found).								
Test value: $\sum \frac{(O-E)^2}{E} = 1.8206 + 0.1361 + 0.1063 + 0.4423 + 2.5664 + 0.1919 + 0.1499 + 0.6235$	M1	Correct formula used.								
= 6.037	A1	Correct to 3 sf.								
H ₀ : (driving) test success is independent of instructor	B1									
Tabular value; 3 df 10% = 6.251 6.04 < 6.251 Accept H ₀	M1	Compare with correct tabular value and conclusion.								
Sufficient evidence to suggest that (driving) test success is independent of instructor	A1	Correct conclusion, in context, following correct work. Level of uncertainty in language is used. Allow ±1 in the third significant figure of test value.								
	7									



15 - (9231/43_Summer_2021_Q5) - X2 Test

(a)	(Mean of distribution = 2.4) Variance = 2.25	M1	Allow 1.5 ² .
	Mean approx. equal to variance (so Poisson might be suitable)	A1	Must have 'approximately' oe, not equality. Must have 2.25.
		2	
(b)	Po(2.5) leads to frequencies:	M1	At least 3 correct to 3sf.
	14.77[5], 36.938, 46.173, 38.477, 24.048, 12.024, 5.010, 2.554	A1	All correct to 4sf.
	Combine last two values: 7.564	M1	FT their table values if final figure is less than 5.
	$\sum \frac{(O-E)^2}{E} = 2.6227 + 0.4198 + 0.00065 + 1.4529 + 0.04570$ $+ 1.3148 + 0.7845$	M1	Apply correct formula to their frequencies.
	6.64	A1	Correct to 3sf.
(b)	H_0 : distribution fits the data	B1	Must mention data and distribution.
	Tabular value, 5 degrees of freedom, is 9.236 '6.64' < 9.236, so accept H_0	M1	Compare with correct tabular value and conclusion.
	Insufficient evidence to show that data does not follow a Poisson distribution	A1	Correct conclusion, in context, following correct work. Level of uncertainty in language is used. Allow ± 1 difference in third significant figure.
		8	

16 - (9231/42_Winter_2021_Q2) - X2 Test

(a)	$P(130 \leq x < 140) = P\left(\frac{130-125}{12} \leq Z < \frac{140-125}{12}\right)$ $P(0.4167 < Z < 1.25)$	M1	Must be seen. Must see 0.4167 and 1.25, may be implied by 0.8944 and 0.6616 Accept 0.417.												
	$0.8944 - 0.6616 = 0.2328$ Multiply by 60 leads to 13.97	A1	AG Accept 0.6615. Need to see 0.2328 or 0.2329 or 13.968 or (0.8944 - 0.6616) × 60.												
		2													
(b)	<table border="1"><tr><td>0</td><td>3</td><td>15</td><td>23</td><td>11</td><td>8</td></tr><tr><td>1.12</td><td>5.22</td><td>13.97</td><td>19.38</td><td>13.97</td><td>6.34</td></tr></table>	0	3	15	23	11	8	1.12	5.22	13.97	19.38	13.97	6.34	M1	Combine first two values.
	0	3	15	23	11	8									
	1.12	5.22	13.97	19.38	13.97	6.34									
	Test stat = 1.7588 + 0.0761 + 0.6743 + 0.6309 + 0.4352	M1													
	3.58	A1	Accept 3.57 – 3.58 . SC: 3.88, if values not combined, scores M1A0.												
	H_0 : distribution fits data $N(125, 12^2)$ is a good model for the data oe	B1	Must mention distribution and data.												
	4 degrees of freedom, so tabular value = 9.488 $3.58 < 9.488$ Accept H_0 .	M1	Compare their value with 9.488 (9.49) and correct FT conclusion Or, if values not combined, compare their value with 11.07 and correct FT conclusion for M1A0.												
	There is sufficient evidence that the normal distribution fits the data There is sufficient evidence that the claim is supported by the data	A1	Correct conclusion, in context, following correct work. Level of uncertainty in language is used.												
	6														



17 - (9231/41_Winter_2021_Q3) - X2 Test

(a)	<table><tr><td>x</td><td>0</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6 or more</td></tr><tr><td>O</td><td>35</td><td>48</td><td>43</td><td>15</td><td>6</td><td>3</td><td>0</td></tr><tr><td>E</td><td>40.874</td><td>57.704</td><td>35.641</td><td>12.579</td><td>2.775</td><td>0.392</td><td>0.0352</td></tr></table>	x	0	1	2	3	4	5	6 or more	O	35	48	43	15	6	3	0	E	40.874	57.704	35.641	12.579	2.775	0.392	0.0352	B1	3 dp or better.
	x	0	1	2	3	4	5	6 or more																			
	O	35	48	43	15	6	3	0																			
	E	40.874	57.704	35.641	12.579	2.775	0.392	0.0352																			
$p = 57.704$																											
	$q = 0.035$ or 0.036	B1	Accept numbers which round to 0.035 or 0.036.																								
		2																									
(b)	Combine frequencies less than 5: last 4 columns give 24 / 15.7812	M1	Allow last 2 or 3 columns combined for this M1.																								
	$\sum \frac{(O-E)^2}{E} = 0.8444 + 1.6319 + 1.5199 + 4.2803$	M1																									
	8.28	A1	Accept 8.27 – 8.28. SC: 25.60, if values not combined, scores M1A0. SC: 14.96 if last 3 combined, scores M1A0. SC: if last 2 combined, scores M1A0.																								
	H_0 : B(8, 0.15) fits the data	B1	Must mention distribution and data.																								
	3 degrees of freedom, tabular value = 7.815. 8.28 > 7.815 Reject H_0 .	M1	Compare their value with 7.815 and correct FT conclusion 3 columns combined compared with 9.488 and correct conclusion																								
	Insufficient evidence to support manager's claim.	A1	Correct conclusion, in context, following correct work. Level of uncertainty in language is used.																								
		6																									

18 - (9231/42_Summer_2022_Q4) - X2 Test

(a)	$p = 16.03$	B1	
	$q = 1.70$	B1	Condone 1.7
		2	



(b)

0	1	2	3	4	5	6	≥ 7
12	20	36	32	13	6	1	0
9.85	24.63	30.78	25.65	16.03	8.02	3.34	1.70

Combine last two columns: O value = 1, E value = 5.04

Calculate values of $\frac{(O-E)^2}{E}$:

0.4693 0.8704 0.8853 1.5720
 0.5727 0.5088 3.2384

Test statistic = 8.12

 H_0 : data fits a Poisson distribution with mean 2.5 H_1 : data does not fit a Poisson distribution with mean 2.5

Critical value of chi-squared = 10.64

Compare: '8.12' < 10.64, accept H_0

There is sufficient evidence to support the scientist's claim / there
 is sufficient evidence to suggest that the data fits a Poisson
 distribution with mean 2.5

M1

Add last 2 columns (or 3 columns: 13.06).

M1

Calculate values.

A1

B1

Need data and distribution.
 e.g Data fits the given distribution.
 Number of butterflies per square metre fits Po(2.5).

M1

Compare their test statistic with 10.64.
 Note: allow 9.236 if 3 columns combined
 or 12.02 if none combined.

A1

Correct conclusion, in context.
 Level of uncertainty in language is used.

6

19 - (9231/41_Winter_2022_Q2) - X2 Test

(a)	Using B(8, 0.6) $p = 0.098$, $q = 41.804$, $r = 31.353$	B1	One correct.
		B1	Other two correct. If B0B0 scored, SCB1 if all 3 correct but not rounded to 3dp.
		2	
(b)	H_0 : the data fits the binomial distribution B(8, 0.6) H_1 : the data does not fit the binomial distribution B(8, 0.6)	B1	
	Add first 3 columns: O value 8, E value 7.471 AND last two columns: O value 11, E value 15.956	M1	Both.
	Chi-squared sum: $0.0374 + 1.5817 + 2.9655 + 0.8058 + 0.9139 + 1.5394$	M1	Method must be seen.
	7.84	A1	If M0 awarded SCB1 for 7.84.
	Critical value, 5df = 9.236 $7.84 < 9.236$ Accept H_0	M1	Compare calculated value with 9.236 OR 13.36 if no columns added. 12.02 adding only last 2 columns. 10.64 adding only first 3 columns.
	There is insufficient evidence to reject the director's claim.	A1	Correct work only except possibly B1 for hypotheses, in context, level of uncertainty in language. 'prove' scores A0
		6	



20 - (9231/42_Winter_2022_Q2) - X2 Test

H_0 : performance is independent of region H_1 : performance is not independent of region	B1										
<table border="1"> <tr> <td>62 (53.9)</td><td>41 (44.1)</td><td>44 (49)</td></tr> <tr> <td>102 (106.7)</td><td>94 (87.3)</td><td>95 (97)</td></tr> <tr> <td>56 (59.4)</td><td>45 (48.6)</td><td>61 (54)</td></tr> </table>	62 (53.9)	41 (44.1)	44 (49)	102 (106.7)	94 (87.3)	95 (97)	56 (59.4)	45 (48.6)	61 (54)	M1 A1	Calculate expected frequencies, allow 1 error.
62 (53.9)	41 (44.1)	44 (49)									
102 (106.7)	94 (87.3)	95 (97)									
56 (59.4)	45 (48.6)	61 (54)									
Test stat = $1.217 + 0.218 + 0.510 + 0.207 + 0.514 + 0.041 + 0.195 + 0.267 + 0.907$	M1	Calculate test statistic.									
4.08	A1	SCB2 for 4.08 if totally unsupported. SCM1A1 B1 for 4.08 supported only by expected frequencies.									
Tabular value, 4df = 7.779 $1.08 < 7.779$ Accept H_0	M1	Compare with 7.779 and conclusion.									
Insufficient evidence that performance is not independent of region	A1	Correct conclusion in context, following correct work, level of uncertainty in language. 'Prove' gives A0.									
	7										

21 - (9231/43_Summer_2023_Q6) - X2 Test

(a)	H_0 : grade is independent of age H_1 : grade is not independent of age	B1	Condone 'Ability to remember' instead of 'grade'.																					
	Calculate expected values, shown in table in bold		M1	At least 2 correct values or expressions.																				
	<table><tr><td>25</td><td>21.2</td><td>16</td><td>24.4</td><td>19</td><td>14.4</td><td>60</td></tr><tr><td>28</td><td>31.8</td><td>45</td><td>36.6</td><td>17</td><td>21.6</td><td>90</td></tr><tr><td>53</td><td></td><td>61</td><td></td><td>36</td><td></td><td>150</td></tr></table>	25	21.2	16	24.4	19	14.4	60	28	31.8	45	36.6	17	21.6	90	53		61		36		150	A1	6 correct values or expressions.
	25	21.2	16	24.4	19	14.4	60																	
	28	31.8	45	36.6	17	21.6	90																	
	53		61		36		150																	
	Calculate chi-squared values: $0.6811 + 2.8918 + 1.4694 + 0.4541 + 1.9279 + 0.9796$		M1	At least 2 correct values (at least 3sf) or expressions seen.																				
	8.40[4]		A1																					
	Tabular value, 2 degrees of freedom = 7.378 '8.404' > 7.378 and reject H_0 /significant		M1	Compare their value with 7.378 and conclusion without context. Condone 'accept H_1 '.																				
	Sufficient evidence to suggest that grade is not independent of age.		A1	Correct conclusion in context, following correct work, level of uncertainty in language (not 'prove', not 'there is no evidence'), no contradictions. A0 if hypotheses wrong way round.																				
		7																						
(b) Value, 4 degrees of freedom = 11.14 $10.91 < 11.14$ Accept H_0 /not significant		M1	Condone 'reject H_1 '.																					
Insufficient evidence to suggest that grade is not independent of age.		A1	CAO																					
		2																						
(c)	For example, result in part (b) because the table contains more information e.g. More degrees of freedom, more groups, more detail	B1	Any appropriate comment to support part (a) or part (b). Allow 'more specific'. Not 'more data' or 'more accurate' on its own.																					
		1																						



22 - (9231/41_Winter_2023_Q2) - X2 Test

H_0 : opinion is independent of street	H_1 : opinion is not independent of street	B1	
33 32.24 54 57.66 42 43.4 26 21.7 155		M1	At least 2 correct values or expressions.
19 19.76 39 35.34 28 26.6 9 13.3 95		A1	6 correct values or expressions.
52 93 70 35 250			
Calculate chi-squared values: $0.0179 + 0.2323 + 0.0452 + 0.8521 + 0.0292 + 0.3790 + 0.0737 + 1.3902$		M1	Correct to 3 decimal places. At least 2 correct values or expressions.
3.02		A1	SC B1 3.02 following M1M0 SC B2 3.02 following M0M0
Tabular value, 3 degrees of freedom 7.815: '3.02' < 7.815, accept H_0		M1	Allow 'not significant'
Insufficient evidence to suggest that opinion depends on street.		A1	CWO. Correct conclusion in context, following correct work, level of uncertainty in language. A0 if hypotheses wrong way round or missing.
		7	

23 - (9231/42_Winter_2023_Q2) - X2 Test

(a)	0 5 13 17 21 16 9 5 4	B1	Each.
	2.718 9.512 16.646 19.421 16.993 11.895 6.939 3.469 2.407	B1	
		2	
(b)	H_0 : Po(3.5) is a good fit to the data H_1 : Po(3.5) is not a good fit to the data	B1	
	Combine first 2 columns: 5, 12.23 And last 2 columns: 9, 5.876	M1	Both.
	Chi-squared values: $4.274 + 0.799 + 0.302 + 0.945 + 1.417 + 0.612 + 1.661$	M1	Allow if no or incorrect number of columns added. At least two 'correct' values (3 sf) or expressions seen from their grouping (or lack of). $2.718 + 2.140 + \dots + 0.6757 + 1.054$
	10.0	A1	AWRT 10.0 If M0 awarded then SC B1 for 10.0.
	Tabular value: 10.64 '10.0' < 10.64, accept H_0 /not significant	M1	Allow equivalent to 10.64 if columns not combined or only one pair combined (12.02 one pair combined, 13.36 none combined).
	Insufficient evidence to suggest that Po(3.5) is not a good fit to the data	A1	Correct conclusion in context, following correct work, level of uncertainty in language. A0 if hypotheses the wrong way round or missing.
		6	



1 - (9231/23_Summer_2017_Q7) - Bivariate Data

(i)	$r = \sqrt{(0.46 \times 0.93)} = 0.654$	M1 A1
	Total:	2
(ii)	$H_0: \rho = 0, H_1: \rho \neq 0$	B1
	$r_{12, 5\%} = 0.576$	B1
	Accept H_1 if $ r > \text{tab. value}$ (AEF)	M1
	0.654 [or 0.65] > 0.576 so there is non-zero correlation	A1
	Total:	4

2 - (9231/21_Summer_2017_Q10) - Bivariate Data

(i)	$\Sigma x = 20, \Sigma y = 30, \Sigma xy = 111, \Sigma x^2 = 110, \Sigma y^2 = 190$ $S_{xy} = 111 - 20 \times 30/5 = -9$ or -1.8 $S_{xx} = 110 - 20^2/5 = 30$ or 6 $[S_{yy} = 190 - 30^2/5 = 10$ or $2]$ $b = S_{xy} / S_{xx} = -9/30 = -3/10$ or -0.3	M1 A1
	$(y - 6) = b(x - 4), y = -0.3x + 7.2$	M1 A1
	Total:	4
(ii)	$r = S_{xy} / \sqrt{(S_{xx} S_{yy})} = -9 / \sqrt{(30 \times 10)}$	M1 A1
	$= -0.520$	*A1
	Total:	3
(iii)	$H_0: \rho = 0, H_1: \rho \neq 0$	B1
	$r_{5, 10\%} = 0.805$	*B1
	Accept H_0 if $ r < \text{tab. value}$ (AEF)	M1
	No [non-zero] correlation (AEF)	DA1
	Total:	4



3 - (9231/23_Summer_2017_Q11) - Rotation Of A Rigid Body, Bivariate Data

(a)(i)	$I_{AB} = \frac{1}{3} 8m (3a/2)^2 + 8m (a/2)^2$ or $8m (3a)^2/12 + 8m (a/2)^2$ [= $8ma^2$]	M1 A1
	$I_{\text{shell}} = \frac{2}{3} ma^2 + m (2a)^2$ [= $(14/3) ma^2$]	M1
	$I_{\text{sphere}} = (2/5) (3m/2)a^2 + (3m/2) (2a)^2$ [= $(33/5) ma^2$]	M1 A1
	$I = (8 + 14/3 + 33/5) ma^2 = (289/15) ma^2$ AG	A1
	Total:	6
(a)(ii)	$\frac{1}{2} I \omega^2 = (5mg/2) 2a (1 - \cos \alpha) - 8mg (a/2) (1 - \cos \alpha)$ or $((21/2)mg \times 2a/21) (1 - \cos \alpha)$	M1 A1
	$= mga (1 - \cos \alpha) = 5mga/6$	A1
	$\omega^2 = 25g/289a$ or $0.0865 g/a$ or $\omega = (5/17) \sqrt{(g/a)}$ or $0.294 \sqrt{(g/a)}$	A1
	$v_{\text{max}} = 2a\omega = 2a \sqrt{(25g/289a)}$ $= \sqrt{(100ag/289)}$ or $(10/17)\sqrt{(ag)}$	M1
	or $0.588\sqrt{(ag)}$ or $1.86\sqrt{a}$ (AEF)	A1
	Total:	6
(b)(i)	e.g. $S_{xy} = 4510.99 - 191 \times 188.8/8 = 3.39$ or 0.424 $S_{xx} = 4564.46 - 191^2/8 = 4.335$ or 0.542 $b = S_{xy} / S_{xx} = 3.39/4.335 = 0.782$	M1 A1
	$(y - 188.8/8) = b (x - 191/8)$ $(y - 23.6) = 0.782 (x - 23.875), y = 0.782x + 4.93$	M1 A1
	Total:	4



(b)(ii)	$H_0: \mu_x - \mu_y = 0.2, H_1: \mu_x - \mu_y > 0.2$ (AEF)	B1
	$d_i: 0.3 \quad 0.2 \quad 0 \quad 0.6 \quad 0.3 \quad 0.5 \quad 0.2 \quad 0.1$	M1
	$\bar{d} = 2.2 / 8 = 0.275$	B1
	$s^2 = (0.88 - 2.2^2/8) / 7$	M1
	$[= 11/280 \text{ or } 0.0393 \text{ or } 0.198^2]$ $t_{7, 0.9} = 1.41[5]$	B1
	$t = (\bar{d} - 0.2) / (s/\sqrt{8}) = 1.07$	M1 A1
	[Accept H_0 .] No evidence for coach's belief (AEF)	B1 FT
Total:		8

4 - (9231/21_Winter_2017_Q9) - Bivariate Data

(i)	$r = S_{xy} / \sqrt{(S_{xx} S_{yy})}$ with e.g. $S_{xy} = 212.62 - 35.9 \times 36.8/8 = 47.48$ (or 5.935) $S_{xx} = 216.47 - 35.9^2/8 = 55.37$ (or 6.921) $S_{yy} = 244.96 - 36.8^2/8 = 75.68$ (or 9.46) (all to 3 s.f.)	M1 A1	Find correlation coefficient r
	$r = 0.733$	*A1	(Insufficient working loses first A1)
		3	
(ii)	$H_0: \rho = 0, H_1: \rho > 0$	B1	State both hypotheses (B0 for $r \dots$)
	EITHER: $r_{8, 1\%} = 0.789$	(*B1	State or use correct tabular one-tail r -value
	Accept H_0 if $ r < \text{tab. } r\text{-value}$ (AEF)	M1)	State or imply valid method for conclusion
	OR: $t_r = r\sqrt{(n-2) / (1 - r^2)} = 2.64, t_{6, 0.95} = 3.143$	(*B1	
	Accept H_0 if $ t_r < \text{tab. } t\text{-value}$ (AEF)	M1)	
	No positive correlation (AEF)	DA1	
(iii)	$r_{14, 1\%} = 0.661, r_{15, 1\%} = 0.641$ so $n_{\text{obs}} = 15$	M1 A1	Find n_{obs} from relevant two-tail tabular value[s] SC: Award B1 for stating 15 without any justification
		2	



5 - (9231/21_Winter_2017_Q11) - Circular Motion, Bivariate Data

A(i)	$\frac{1}{2}mv^2 = \frac{1}{2}mu^2 - mga(\cos \alpha + \cos \beta)$	M1 A1	Find v^2 at A' from conservation of energy (A0 if no m)
	$mv^2/a = mg \cos \beta$ $u^2 = ag \cos \beta + 2ag(1/16 + \cos \beta)$	B1	Use $F = ma$ radially at A' with $R_{A'} = 0$ Use $\cos \alpha = 1/16$ and eliminate v^2 to verify u^2
	$= (1/8) ag(1 + 24 \cos \beta)$ AG	M1 A1	
		5	

A(ii)	$\frac{1}{2}mv^2 = \frac{1}{2}mu^2 - mga(1 + \cos 2\beta)$	B1	Find w^2 at B' from conservation of energy (B0 if no m)
	$mw^2/a = mg \cos 2\beta$ $u^2 = ag \cos 2\beta + 2ag(1 + \cos 2\beta)$	B1	Use $F = ma$ radially at B' with $R_{B'} = 0$ Eliminate w^2 to find u^2
	$= ag(2 + 3 \cos 2\beta)$	M1	
	$1 + 24 \cos \beta = 8(2 + 3 \cos 2\beta)$ $= 16 + 48 \cos^2 \beta - 24$ $16 \cos^2 \beta - 8 \cos \beta - 3 = 0$	M1	Combine eqns for u^2 Formulate and solve quadratic to find $\cos \beta$
	$\cos \beta = 3/4$ [rejecting $-1/4$]	M1 A1	
		6	
A(iii)	$u^2 = (19/8) ag$ or $2.375 ag$ $R = mu^2/a + mg \cos \alpha$ $= (19/8 + 1/16) mg$	B1	Find u^2 using value of $\cos \beta$ Use $F = ma$ radially at A to find reaction R at A
	$= (39/16) mg$ or $2.44 mg$	M1 A1	
		3	

B(i)	e.g. $S_{xy} = 218.72 - 45.3 \times 40.5/9 = 14.87$ or 1.652 $S_{xx} = 245.59 - 45.3^2/9 = 17.58$ or 1.953 or $S_{yy} = 195.11 - 40.5^2/9 = 12.86$ or 1.429		Find reqd. values
	EITHER: $b_1 = S_{xy} / S_{xx} = 14.87 / 17.58 = 0.84585 = 0.846$ $(y - 40.5/9) = b_1(x - 45.3/9)$ $(y - 4.5) = 0.846(x - 5.033)$	(M1 A1)	Find gradient in $y - \bar{y} = b_1(x - \bar{x})$ to 3 s.f. Find eqn. of regression line to 3 s.f.
	$y = 0.846x + 0.24257 = 0.846x + 0.243$	M1 A1	
	$x = 4.68$	A1	Find x when $y = 4.2$
	OR: $b_2 = S_{xy} / S_{yy} = 14.87 / 12.86 = 1.1563 = 1.16$ $(x - 45.3/9) = b_2(y - 40.5/9)$ $(x - 5.033) = 1.16(y - 4.5)$	(M1 A1)	Find gradient in $x - \bar{x} = b_2(y - \bar{y})$ to 3 s.f. Find eqn. of regression line to 3 s.f.
	$x = 1.16y - 0.170[04]$	M1A1	
	$x = 4.69$	A1	Find x when $y = 4.2$ (A0 for $x = 4.70$)
		5	



B(ii)	$H_0: \mu_x - \mu_y = 0.3, H_1: \mu_x - \mu_y > 0.3$ (AEF)	B1	State hypotheses (B0 for $\bar{x}, \dots$)
	$\bar{d} = 4.8 / 9$ or $8/15$ or 0.533 (where $d = x - y$)	B1	Find sample mean
	$s^2 = (3.26 - 4.8^2/9) / 8$ $= 7/80$ or 0.0875 or 0.296^2	M1 A1	Estimate population variance (allow biased here: $7/90$ or 0.0778 or 0.279^2)
	$t = (\bar{d} - 0.3) / (s/\sqrt{9}) = 2.37$	M1 A1	Find value of t
	$t_{8, 0.975} = 2.306$ or 2.31	B1	State or use correct tabular t -value (or can compare $\bar{d}$ with $0.3 + t_{8, 0.975} s/\sqrt{9} = 0.527$)
	Reject H_0 if $t >$ tabular value $2.37 [\pm 0.1] > 2.31$ so accept belief (AEF)	M1	Valid method for reaching conclusion Correct conclusion, from correct values
	[of increase of more than 0.3] (AEF)	A1	SC: Wrong (hypothesis) test can earn only B1 for hypotheses
		9	

6 - (9231/23_Summer_2018_Q6) - Bivariate Data

(i)	$H_0: \rho = 0, H_1: \rho \neq 0$	B1	State both hypotheses (B0 for $r, \dots$)
	EITHER: $r_{15, 10\%} = 0.441$	*B1	State or use correct tabular two-tail r -value
	Accept H_0 if $0.430 <$ tab. r -value (AEF)	M1	State or imply valid method for conclusion
	OR: $t_r = r\sqrt{(n-2)/(1-r^2)} = 1.72, t_{13, 0.95} = 1.771$	(*B1)	(Rarely seen)
	Accept H_0 if $ t_r <$ tab. T -value (AEF)	(M1)	
	No [non-zero] correlation (AEF)	A1	Correct conclusion (dep *B1)
(ii)	$r_{9, 5\%} = 0.621, r_{9, 5\%} = 0.582$ so $N_{\text{crit}} = 9$ (or equivalent argument)	M1 A1	Find N_{crit} from relevant [one-tail] tabular values SC: Award B1 for stating 9 with inadequate justification
		2	

7 - (9231/23_Summer_2018_Q8) - Bivariate Data

(i)	$b \times 0.6331 = 0.9797^2, b = 1.516$ (3 d.p.)	M1 A1	Find b from given gradients and coefficient
		2	
(ii)	$46.5/6 [= 7.75] = b \times \bar{x} + 1.306$ or $b \times (\Sigma x)/6 + 1.306$	M1	Find p from means and regression line of y on x
	$\bar{x} = 4.25$ or $\Sigma x = 25.5$ (3 s.f.)	A1	(intermediate values may be implied)
	$6 \bar{x}$ or $\Sigma x = 25.5$ or $p, p = 4.2$ (1 d.p.)	M1 A1	
		4	
(iii)	EITHER: $(\Sigma x)/6 = 0.6331 (\Sigma y)/6 + d [d = -0.656]$	M1	Find d from means and regression line of x on y
	$x = 0.6331 \times 8.5 + d = 4.725[3]$ or 4.73	M1 A1	Estimate x when $y = 8.5$ using regression line of x on y
	OR: $x - (\Sigma x)/6 = 0.6331 \{8.5 - (\Sigma y)/6\},$	(M2)	Combine above into single step
	$x = 4.725[3]$ or 4.73	(A1)	
	SC: $x = (8.5 - 1.306) / b [= 4.7452] = 4.75$	(B1)	SC: Estimate x when $y = 8.5$ using regression line of y on x
		3	



8 - (9231/21_Summer_2018_Q11) - Circular Motion, Bivariate Data

A(i)	$\frac{1}{2}mv^2 = \frac{1}{2}mu^2 + mga \cos \theta$ and	M1	Use conservation of energy (A0 if no m) and
	$T = mv^2/a + mg \cos \theta$ $T = mu^2/a + 3mg \cos \theta$	M1 A1	find tension T at P_1 by using $F = ma$ (A1 for both correct) Combine to verify $\cos \theta$
	$(11/5)mg = 2mg/5 + 3mg \cos \theta, \cos \theta = 3/5$	AG	M1 A1 by substituting $T = (11/5)mg$ and $u = \sqrt{(2ag/5)}$
	$v^2 = u^2 + 2ag \cos \theta = (2/5 + 6/5)ag = 8ag/5$ $v = \sqrt{(8ag/5)}$ or $\sqrt{(1.6ag)}$ or $4\sqrt{a}$	B1 (AEF)	Find v at P_1 (can assume $\cos \theta = 3/5$)
		6	
A(ii)	$t = (a \sin \theta) / (v \cos \theta) = 4a/3v \quad [= \frac{1}{3}\sqrt{a}]$	M1 A1	Find time t to P_2 by considering horizontal motion
	$h = (v \sin \theta) t + \frac{1}{2}gt^2$ $= 16a/15 + 5a/9 = 73a/45$ or $1.62a$	M1 A1	Find ht. fallen at P_2 by considering vertical motion
	$OP_2 = h + a \cos \theta = 20a/9$ or $2.22a$	M1 A1	Find OP_2
		6	
B(i)	$\Sigma x = 48, \Sigma y = 29 + 2p, \Sigma xy = 242 + 17p,$ $\Sigma x^2 = 450, [\Sigma y^2 = 219 + 2p^2]$ $S_{xy} = 242 + 17p - 48 \times (29 + 2p)/6 = 10 + p$ $S_{xx} = 450 - 48^2/6 = 66$ $[S_{yy} = 219 + 2p^2 - (29 + 2p)^2/6 = (8p^2 - 116p + 473)/6]$	M1 A1	Find required values
	$0.25 = S_{xy} / S_{xx} = (10 + p)/66, \quad p = 6.5$	M1 A1	Find p from gradient in eqn. of regression line
	$(29 + 2p)/6 = 0.25 \times 48/6 + k, \quad k = 7 - 2 = 5$	M1 A1	Find k from means and regression line
		6	

B(ii)	$[\Sigma y = 42, \Sigma y^2 = 303.5, S_{yy} = 303.5 - 42^2/6 = 9.5]$ $r = S_{xy} / \sqrt{(S_{xx}S_{yy})} = 16.5 / \sqrt{(66 \times 9.5)} = 0.659$	M1 *A1	Find correlation coefficient r
		2	
B(iii)	$H_0: \rho = 0, H_1: \rho > 0$	B1	State both hypotheses (B0 for $r \dots$)
	<i>EITHER:</i> $r_s, s_b = 0.729$	*B1	State or use correct tabular one-tail r -value
	Accept H_0 if $ r < \text{tab. } r\text{-value}$ (AEF)	M1	State or imply valid method for conclusion
	<i>OR:</i> $t_r = r\sqrt{(n-2)/(1-r^2)} = 1.75, t_{4,0.95} = 2.132$	(*B1)	(Rarely seen)
	Accept H_0 if $ t_r < \text{tab. } t\text{-value}$ (AEF)	(M1)	
	No positive correlation (AEF)	A1	Correct conclusion (dep *A1, *B1)
		4	

9 - (9231/22_Winter_2018_Q9) - Bivariate Data

(i)	$H_0: \rho = 0, H_1: \rho \neq 0$	B1	State both hypotheses (B0 for $r \dots$)
	<i>EITHER:</i> $r_s, s_b = 0.878$	*B1	State or use correct tabular two-tail r -value
	Accept H_0 if $ r < \text{tab. value}$ (AEF)	M1	State or imply valid method for conclusion
	<i>OR:</i> $t_r = r\sqrt{(n-2)/(1-r^2)} = -2.08, t_{4,0.975} = 2.776$ or 2.78	(*B1)	(Rarely seen)
	Accept H_0 if $ t_r < \text{tab. } t\text{-value}$ (AEF)	(M1)	
	No evidence of [non-zero] correlation (AEF)	A1	Correct conclusion (dep *B1)
		4	



(ii)	$cd = r^2 = (-0.7214)^2 = 0.52042$	M1	Find cd
	$\bar{x} = 10.8 + d(4.2 + c\bar{x})$ [$\bar{y} = 2.137$]	M1	Find 2nd eqn. for c, d using <u>means</u> in eqns. of regression lines Combine to find c (or d)
	$c = 4.2 r^2 / \{(1 - r^2)\bar{x} - 10.8\} = -0.294$ or $d = \{(1 - r^2)\bar{x} - 10.8\} / 4.2 = -1.77$	M1A1	
	$d = 0.7214^2 / c = -1.77$ or $c = 0.7214^2 / d = -0.294$	A1	and hence d (or c)
		5	
(iii)	$x = 10.8 - 1.77 \times 3.5 = 4.60$ [5] [y on x gives 2.38]	B1	Find y from eqn. of regression line of x on y
	e.g. not reliable since no evidence of correlation or reasonably reliable since 0.7214 close to 1 or not reliable since 0.7214 not close to 1 or reliability unclear as degree of extrapolation unknown	B1	Valid comment on reliability (AEF)
		2	

10 - (9231/21_Winter_2018_Q10) - Bivariate Data

(i)	$\bar{x} = (\bar{y} - 1.1664) / 0.4604 = 11.02$ or $\Sigma x = 110.2$	M1	Find $\bar{x}$ or Σx using eqn. of regression line of y on x
	$S_{xx} = 1419.98 - 110.2^2 / 10 = 205.576$ $S_{yy} = 439.68 - 62.4^2 / 10 = 50.304$ (AEF)	M1	Find S_{xx} and S_{yy} (may be earned in part (ii))
	EITHER: $b = S_{xy} / S_{xx} = (S_{xy} / S_{yy}) \times (S_{yy} / S_{xx})$ $= 0.4604 \times 205.576 / 50.304 = 1.88$ [15]	M1A1	Find gradient b to 3 s.f. in $x - \bar{x} = b(y - \bar{y})$
	OR: $S_{xy} = 0.4604 \times S_{xx} = 94.647$ $b = S_{xy} / S_{xx} = 94.647 / 50.304 = 1.88$ [15]	(M1A1)	
	$(x - 11.02) = b(y - 6.24), x = 1.88y - 0.721$	M1A1	and hence eqn. of regression line of x on y
		6	
(ii)	$r = S_{xy} / \sqrt{(S_{xx} S_{yy})} = 94.647 / \sqrt{(205.576 \times 50.304)}$ or $\sqrt{(0.4604 \times b)} = \sqrt{(0.4604 \times 1.8815)}$ (AEF)	M1	Find correlation coefficient r
	$= 0.931$	*A1	
		2	
(iii)	$H_0: \rho = 0, H_1: \rho > 0$	B1	State both hypotheses ($B0$ for $r, \dots$)
	EITHER: $r_{10, 5\%} = 0.549$	*B1	State or use correct tabular one-tail r -value
	Reject H_0 if $ r > \text{tab. } r\text{-value}$ (AEF)	M1	State or imply valid method for conclusion
	OR: $t_r = r\sqrt{(n-2) / (1 - r^2)} = 7.21, t_{8, 0.05} = 1.86$	(*B1)	(Rarely seen)
	Reject H_0 if $ t_r > \text{tab. } t\text{-value}$ (AEF)	M1	
	Evidence of positive correlation (AEF)	A1	Correct conclusion (dep *A1, *B1)
		4	



11 - (9231/21_Summer_2019_Q10) - Bivariate Data

(i)	$\Sigma x = 25, \Sigma y = 25 + q, \Sigma xy = 135 + 4q,$ $\Sigma x^2 = 141, \Sigma y^2 = 159 + q^2$	B1	Find required values
	$S_{xy} = 135 + 4q - 25(25 + q)/5 = [10 - q]$ (AEF) [$S_{xx} = 141 - 25^2/5 = 16$] $S_{yy} = (159 + q^2) - (25 + q)^2/5 = [34 - 10q + 4q^2/5]$	M1 A1	(S_{xx}, S_{xx}, S_{yy} may be scaled by the same constant)
	$40 - 4q = 170 - 50q + 4q^2, 2q^2 - 23q + 65 = 0$ $(2q - 13)(q - 5) = 0, q = 5$	M1 A1	Equate gradient 5/4 in line of x on y to S_{xy} / S_{xx} and solve quadratic to find integer value of q
		5	
(ii)	$c = 25/5 - (5/4)(25 + q)/5 = -(5 + q) / 4$	M1 A1	Find c from $\bar{x} - (5/4) \bar{y}$
	$= -5/2$ or -2.5	A1	
		3	
(iii)	$r = S_{xy} / \sqrt{(S_{xx} S_{yy})} = 5 / \sqrt{(16 \times 4)}$	M1 A1	Find correlation coefficient r
	$= 5/8$ or 0.625	A1	
		3	

12 - (9231/23_Summer_2019_Q10) - Bivariate Data

(i)	$b = S_{xy} / S_{xx}$ and $r = S_{xy} / \sqrt{(S_{xx} S_{yy})}$ so	M1	Relate gradient b in $y = bx + c$ to r, S_{xx}, S_{yy}
	$b = r \sqrt{(S_{yy} / S_{xx})}$	A1	
	$= 0.5815 \sqrt{(7.9473 / 3.3086)} = 0.901[2]$	M1 A1	Find S_{xx} and S_{yy} and hence find b to 3 s.f.
	$(y - 6.7375) = b(x - 3.3125), y = 0.901x + 3.75$	M1 A1	Find equation of regression line of y on x
	Alternative method for question 10(i)		
	$S_{xy} = r \sqrt{(S_{xx} S_{yy})} = r \sqrt{(8x_1^2)(8y_1^2)}$ $= 8 \times 0.5815 \sqrt{(7.9473 \times 3.3086)}$	M1	Find S_{xy} (allow values consistently scaled by factor 8)
	$= 23.855$ (allow 2.9818) (to 4 s.f.)	A1	
	$b = S_{xy} / S_{xx} = 23.85 / (8 \times 3.3086) = 0.901[2]$	M1 A1	Hence find b to 3 s.f.
	$(y - 6.7375) = b(x - 3.3125), y = 0.901x + 3.75$	M1 A1	Find equation of regression line of y on x
		6	
(ii)	$H_0: \rho = 0, H_1: \rho > 0$	B1	State both hypotheses (B0 for $r \dots$)
	$r_{8, 95} = 0.621$	B1	State or use correct tabular one-tail r -value
	Accept H_0 if $0.5815 < \text{tab. } r\text{-value}$ (AEF)	M1	State or imply valid method for conclusion
	No evidence of positive correlation (AEF)	A1	Correct conclusion
	Alternative method for question 10(ii)		
	$H_0: \rho = 0, H_1: \rho > 0$	B1	State both hypotheses (B0 for $r \dots$)
	$t_r = r \sqrt{(n-2) / (1 - r^2)} = 1.75, t_{0.05} = 1.943$	B1	
	Accept H_0 if $ t_r < \text{tab. } t\text{-value}$ (AEF)	M1	State or imply valid method for conclusion
	No evidence of positive correlation (AEF)	A1	Correct conclusion
		4	
(iii)	$y = 9.16$	B1	Find y when $x = 6.0$
	Unreliable since r is small or r is not close to 1 or no correlation or 6.0 is not close to mean of x (AEF)	B1	Reasonable comment on reliability
		2	



13 - (9231/21_Winter_2019_Q9) - Bivariate Data

(i)	$\Sigma x = 15, \Sigma y = 7 + p + q, \Sigma xy = 17 + 2p + 3q$ $\Sigma x^2 = 55, [\Sigma y^2 = 21 + p^2 + q^2]$	M1	Find required summations
	$S_{xx} = 55 - 15^2/5 = 10$ and $S_{yy} = 17 + 2p + 3q - 15 \times (7 + p + q)/5 = -4 - p$	M1 A1	
	$-0.5 = S_{xy} / S_{xx} = (-4 - p)/10 \quad p = 1$	M1 A1	Find p from gradient in eqn. of regression line
	$(7 + p + q)/5 = -0.5 \times 15/5 + 3.5 \quad q = 2$	M1 A1	Find q from means and regression line
		7	
(ii)	$\Sigma y = 10, \Sigma y^2 = 26, S_{yy} = 26 - 10^2/5 = 6$	M1	Find S_{yy}
	$r = S_{xy} / \sqrt{(S_{xx}S_{yy})} = -5 / \sqrt{(10 \times 6)}$	M1	Find correlation coefficient r
	$r = -0.645[5]$ [allow -0.646]	A1	
		3	



1 - (9231/41_Summer_2020_Q2) - Non Parametric Test

(a)	H_0 : population median is 6.40, H_1 : population median $\neq$ 6.40	BL																				
	Calculate differences and signed ranks	MIA																				
	<table><tr><td>0.04</td><td>-0.24</td><td>-0.78</td><td>-0.58</td><td>0.11</td><td>0.22</td><td>-0.21</td><td>0.02</td><td>-0.06</td><td>-0.12</td></tr><tr><td>2</td><td>-8</td><td>-10</td><td>-9</td><td>4</td><td>7</td><td>-6</td><td>1</td><td>-3</td><td>-5</td></tr></table>	0.04	-0.24	-0.78	-0.58	0.11	0.22	-0.21	0.02	-0.06	-0.12	2	-8	-10	-9	4	7	-6	1	-3	-5	
	0.04	-0.24	-0.78	-0.58	0.11	0.22	-0.21	0.02	-0.06	-0.12												
	2	-8	-10	-9	4	7	-6	1	-3	-5												
	Test statistic = $4 + 7 + 2 + 1 = 14$	AL																				
	Compare with correct critical value 8	ML																				
Accept H_0 : insufficient evidence to reject claim (FT their test statistic)	AL FT																					
	6																					
(b)	Symmetrically distributed about the median	BL																				
		1																				

2 - (9231/43_Summer_2020_Q6) - Non Parametric Test

(a)	Difference of location test for populations not known to be normal				B1
					1

(b)	<table><tr><td>10.6</td><td>2</td><td>10.5</td><td>1</td></tr><tr><td>11.1</td><td>7</td><td>10.7</td><td>3</td></tr><tr><td>11.8</td><td>12</td><td>10.8</td><td>4</td></tr><tr><td>12.0</td><td>14</td><td>10.9</td><td>5</td></tr><tr><td>12.1</td><td>15</td><td>11.0</td><td>6</td></tr><tr><td>12.2</td><td>16</td><td>11.2</td><td>8</td></tr><tr><td>12.3</td><td>17</td><td>11.3</td><td>9</td></tr><tr><td>12.4</td><td>18</td><td>11.6</td><td>10</td></tr><tr><td>13.2</td><td>21</td><td>11.7</td><td>11</td></tr><tr><td>13.5</td><td>22</td><td>11.9</td><td>13</td></tr><tr><td>13.8</td><td>23</td><td>12.5</td><td>19</td></tr><tr><td>13.9</td><td>24</td><td>12.6</td><td>20</td></tr></table>				10.6	2	10.5	1	11.1	7	10.7	3	11.8	12	10.8	4	12.0	14	10.9	5	12.1	15	11.0	6	12.2	16	11.2	8	12.3	17	11.3	9	12.4	18	11.6	10	13.2	21	11.7	11	13.5	22	11.9	13	13.8	23	12.5	19	13.9	24	12.6	20	M1
	10.6	2	10.5	1																																																	
	11.1	7	10.7	3																																																	
	11.8	12	10.8	4																																																	
	12.0	14	10.9	5																																																	
	12.1	15	11.0	6																																																	
	12.2	16	11.2	8																																																	
	12.3	17	11.3	9																																																	
	12.4	18	11.6	10																																																	
	13.2	21	11.7	11																																																	
	13.5	22	11.9	13																																																	
	13.8	23	12.5	19																																																	
	13.9	24	12.6	20																																																	
	Total ranks: 109 or 191				A1																																																
$H_0: m_x = m_y$ $H_1: m_x > m_y$				B1																																																	
Use normal approximation with attempts at mean and variance				M1																																																	
mean = 150, variance = 300				A1																																																	

(b)	$\frac{109.5 - 150}{\sqrt{300}} = -2.338$	M1A1
	Probability 0.0097	A1
	This is less than 0.05, so reject H_0	M1
	There is evidence of an effect on growth	A1
		10



3 - (9231/41_Winter_2020_Q2) - Non Parametric Test

(a)	$H_0: m = 200$ $H_1: m < 200$	B1	Allow 'median' in words, allow m not defined
	10 -14 -12 8 -16 -9 15 -2 -4	M1	Signed differences (at most 3 errors)
	Ranks: 5 -7 -6 3 -9 -4 8 -1 -2	A1	Award for correct rank order, ignore signs
	Sum ranks $T = 16$	A1	CWO
	Critical value 8 and compare ' 16 ' > 8	M1	Compare their T with 8
	Accept H_0 Insufficient evidence to support scientist's belief	A1	In context, all correct, except possibly hypotheses. Level of uncertainty in language used. No contradictions
		6	
(b)	Magnitude of differences from median are taken into account	B1	Must mention magnitude and differences
		1	

4 - (9231/42_Winter_2020_Q2) - Non Parametric Test

(a)	H_0 : difference in (population) medians = 0 H_1 : difference in (population) medians < 0 (or > 0)	B1 B1	Correct hypotheses (median or m) All notation identified e.g. m = population median if used
	Differences: -3 4 5 -1 -8 -19 -11 10 -13 -7 9 -16	M1	At most 3 errors
	Ranks: -2 3 4 -1 -6 -12 -9 8 -10 -5 7 -11	A1	Award for correct rank order, ignore signs
	Sum $T = 3 + 4 + 8 + 7 = 22$	A1	cwo
	Compare with critical value 17: ' 22 ' > 17	M1	Compare <i>their</i> T with 17
	Accept H_0 Data does not support student's claim	A1	In context, all correct, except possibly second B1 Level of uncertainty in language used. No contradictions.
		7	
(b)	Rank for A becomes + 2,	M1	$T = 24$
	Changing sign of difference can only reduce evidence in favour of the claim.	A1	still > 17 and test result unchanged
		2	

5 - (9231/43_Summer_2021_Q2) - Non Parametric Test

H_0 : difference in population medians = 0 ($m_a = m_b$)	B1	Correct hypotheses, allow m_d . Do not allow μ or mean.
H_1 : difference in population medians $\neq 0$ ($m_a \neq m_b$)	B1	'population' included.
Diff: 12 -14 -1 3 -13 8 -5 15 6	M1	Allow one error.
Rank: 6 -8 -1 2 -7 5 -3 9 4	M1	Allow one error.
Smaller sum 19	A1	(Other sum is 26.)
Critical tabular value = 8 '19' > 8 so accept H_0	M1	Comparison with 8 and correct FT conclusion.
There is insufficient evidence that marks differ	A1	Correct conclusion, in context, following correct work, except possibly the hypotheses. Level of uncertainty in language is used.
	7	



6 - (9231/41_Summer_2021_Q5) - Non Parametric Test

(a)	Underlying distribution or population of differences is unknown. Underlying distributions or population of scores for both X and Y are unknown	B1	Not known to be normal. Condone 'scores for X and Y cannot be assumed to be normally distributed'.
		1	
(b)	H_0 : difference of population medians = 0	B1	Correct hypotheses, allow m but not μ or mean.
	H_1 : difference of population median $\neq 0$	B1	'population' included.
	Diff: 4 -1 -21 3 -7 -9 -5 8 -10 -12 -11	M1	Differences, allow one error.
	Rank: 3 -1 -11 2 -5 -7 -4 6 -8 -10 -9	M1	Signed ranks, allow one sign error.
	$[Q = -55,] P = 11$	A1	$P = 11$
	Critical tabular value = 13 '11' < 13 so reject H_0	M1	Comparison with 13 and correct ft conclusion.
	Sufficient evidence of a preference for one of the uniforms OE	A1	Correct conclusion, in context, following correct work, except possibly second B1. Level of uncertainty in language is used.
		7	

7 - (9231/42_Winter_2021_Q4) - Non Parametric Test

(a)	<table border="1"> <tr><td>12</td><td>1</td><td>28</td><td>4</td></tr> <tr><td>18</td><td>2</td><td>31</td><td>6</td></tr> <tr><td>25</td><td>3</td><td>35</td><td>9</td></tr> <tr><td>29</td><td>5</td><td>36</td><td>10</td></tr> <tr><td>32</td><td>7</td><td>37</td><td>11</td></tr> <tr><td>33</td><td>8</td><td>41</td><td>13</td></tr> <tr><td>40</td><td>12</td><td>48</td><td>15</td></tr> <tr><td>46</td><td>14</td><td>49</td><td>16</td></tr> </table>	12	1	28	4	18	2	31	6	25	3	35	9	29	5	36	10	32	7	37	11	33	8	41	13	40	12	48	15	46	14	49	16	M1	Attempt at ranking.
12	1	28	4																																
18	2	31	6																																
25	3	35	9																																
29	5	36	10																																
32	7	37	11																																
33	8	41	13																																
40	12	48	15																																
46	14	49	16																																
	Test statistic: 52	A1																																	
	$H_0: m_x = m_y$ and $H_1: m_x \neq m_y$	B1	Allow in words but 'population' must be included.																																
	Critical value for (8, 8) is 49.	*B1	Allow 51 if clearly one-tail test in hypotheses.																																
	52 > 49 Accept H_0	DM1	Compare their calculated value with 49 and correct FT conclusion.																																
	Insufficient evidence of difference in medians.	A1	Correct conclusion, in context, following correct work. Level of uncertainty in language is used.																																
		6																																	
(b)	Not a paired sample.	B1																																	
	Underlying distribution/population not (known to be) normal Underlying distribution/population unknown	B1	B0 for 'data is not normally distributed' B0 for 'marks are not normally distributed'																																
		2																																	



8 - (9231/41_Winter_2021_Q6) - Non Parametric Test

<table><tr><td>51</td><td>1</td><td>72</td><td>3</td></tr><tr><td>55</td><td>2</td><td>86</td><td>5</td></tr><tr><td>75</td><td>4</td><td>170</td><td>10</td></tr><tr><td>98</td><td>6</td><td>175</td><td>11</td></tr><tr><td>137</td><td>7</td><td>220</td><td>12</td></tr><tr><td>152</td><td>8</td><td>260</td><td>17</td></tr><tr><td>167</td><td>9</td><td>262</td><td>18</td></tr><tr><td>235</td><td>13</td><td>302</td><td>19</td></tr><tr><td>238</td><td>14</td><td>311</td><td>20</td></tr><tr><td>242</td><td>15</td><td>320</td><td>21</td></tr><tr><td>256</td><td>16</td><td>333</td><td>22</td></tr><tr><td></td><td></td><td>351</td><td>23</td></tr></table>	51	1	72	3	55	2	86	5	75	4	170	10	98	6	175	11	137	7	220	12	152	8	260	17	167	9	262	18	235	13	302	19	238	14	311	20	242	15	320	21	256	16	333	22			351	23	M1	Attempt at ranking.
51	1	72	3																																															
55	2	86	5																																															
75	4	170	10																																															
98	6	175	11																																															
137	7	220	12																																															
152	8	260	17																																															
167	9	262	18																																															
235	13	302	19																																															
238	14	311	20																																															
242	15	320	21																																															
256	16	333	22																																															
		351	23																																															
Total ranks: 95	A1																																																	
$H_0 : m_x = m_y$ and $H_1 : m_x \neq m_y$	B1	Allow words, must include 'population' median.																																																
Use normal approximation with attempts at mean and variance.	M1																																																	
Mean =132, variance = 264	A1																																																	
$\frac{95.5 - 132}{\sqrt{264}}$	M1	Allow no or wrong continuity correction for M1 only.																																																
- 2.246	A1	CAO																																																
Critical value is - 1.96. - 2.246 < - 1.96 reject H_0 .	M1	Compare their value with - 1.96. Or area comparison 0.0123 or 0.0124 with 0.025 and FT conclusion.																																																
There is sufficient evidence of a difference in levels.	A1	Correct conclusion, in context, following correct work. Level of uncertainty in language is used.																																																
	9																																																	

9 - (9231/43_Summer_2022_Q5) - Non Parametric Test

Signs: + - - + + + - + + + and use Binomial	M1	11 signs (values not needed) allow one error.
Consider $B(11, 0.5)$ with $P(X \geq 8)$ or $P(X \leq 3)$	M1	Use of correct binomial.
Compare 0.227 with 0.1 or 0.113 with 0.05	A1	
Differences: 0.06 -0.05 -0.03 0.08 0.04 0.02 0.10 -0.15 0.09 0.01 0.12	M1	Allow one error.
Signed Ranks: 6 -5 -3 7 4 2 9 -11 8 1 10	M1	Allow one swap.
$(P = 47) Q = 19$	A1	
Critical value from table = 13	B1	
$19 > 13$ Accept H_0	M1	FT <i>their</i> (19) compared with 13 and correct FT conclusion
Insufficient evidence to support median not being 5.50	A1	BOTH conclusions correct from correct working, in context. Level of uncertainty in language.
	9	



10 - (9231/41_Summer_2022_Q6) - Non Parametric Test

Differences/rank:	M1 A1	Attempt at differences, allow 4 errors.																																																
<table><tr><td>0.2</td><td>-0.5</td><td>-1.6</td><td>-2.9</td><td>2.5</td><td>8.5</td><td>1.2</td><td>5.2</td></tr><tr><td>2</td><td>-5</td><td>-11</td><td>-18</td><td>17</td><td>24</td><td>9</td><td>21</td></tr><tr><td>-0.1</td><td>0.8</td><td>7.8</td><td>1.9</td><td>-0.3</td><td>-1.8</td><td>-0.7</td><td>-1.4</td></tr><tr><td>-1</td><td>7</td><td>23</td><td>13</td><td>-3</td><td>-12</td><td>-6</td><td>-10</td></tr><tr><td>-0.9</td><td>2.1</td><td>2.3</td><td>-5.4</td><td>-2.0</td><td>3.4</td><td>4.7</td><td>0.4</td></tr><tr><td>-8</td><td>15</td><td>16</td><td>-22</td><td>-14</td><td>19</td><td>20</td><td>4</td></tr></table>	0.2	-0.5	-1.6	-2.9	2.5	8.5	1.2	5.2	2	-5	-11	-18	17	24	9	21	-0.1	0.8	7.8	1.9	-0.3	-1.8	-0.7	-1.4	-1	7	23	13	-3	-12	-6	-10	-0.9	2.1	2.3	-5.4	-2.0	3.4	4.7	0.4	-8	15	16	-22	-14	19	20	4		
0.2	-0.5	-1.6	-2.9	2.5	8.5	1.2	5.2																																											
2	-5	-11	-18	17	24	9	21																																											
-0.1	0.8	7.8	1.9	-0.3	-1.8	-0.7	-1.4																																											
-1	7	23	13	-3	-12	-6	-10																																											
-0.9	2.1	2.3	-5.4	-2.0	3.4	4.7	0.4																																											
-8	15	16	-22	-14	19	20	4																																											
Q = 110 P = 190	A1	Either sum correct.																																																
H ₀ : population median = 18.0 and H ₁ : population median ≠ 18.0	B1	Must be 'population', allow <i>m</i> .																																																
Mean = $\frac{1}{4}n(n+1) = \frac{1}{4} \times 24 \times 25 = 150$	B1	Normal approximation: mean.																																																
Variance = $\frac{1}{24}n(n+1)(2n+1) = \frac{1}{24} \times 24 \times 25 \times 49 = 1225$	B1																																																	
$\frac{110.5 - 150}{\sqrt{1225}}$	M1	Attempt at test statistic, allow incorrect or missing cc, FT <i>their</i> 110.																																																
-1.129	A1																																																	
10% 2-tail, critical value is 1.645 -1.129 > -1.645 or 0.1296 > 0.05, accept H ₀ .	M1	Compare with -1.645, correct FT conclusion.																																																
Insufficient evidence to suggest that median is not 18.0/ Insufficient evidence to reject teacher's claim.	A1	Correct conclusion, in context. Level of uncertainty in language is used.																																																
	10																																																	

11 - (9231/41_Winter_2022_Q3) - Non Parametric Test

(a)	Differences: 4 -8 3 1 -7 11 19 6 10 -13 14 5 -9 12	M1	Attempt at signed differences, allow 2 errors.
	Signed Ranks: 3 -7 2 1 -6 10 14 5 9 -12 13 4 -8 11	M1	Ranks, allow 4 errors.
	P = 72, Q = 33 T = 33	A1	33 clearly identified.
	H_0 : difference in population medians = 0 H_1 : difference in population medians > 0	B1	Correct, allow m for this mark.
	Critical tabular value = 25	B1	
	33 > 25 accept H_1	M1	Compares their value of sum of ranks with their 25 and conclusion.
	Insufficient evidence to support claim	A1	All correct except possibly first B1, in context, level of uncertainty in language. 'Prove' scores A0.
		7	
(b)	Underlying/population distribution of differences is unknown (not known to be normal)	B1	
		1	



12 - (9231/42_Winter_2022_Q6) - Non Parametric Test

(a)	359	3	342	1	M1	Wrong test: max B1 for hypotheses, B1B1 for correct mean and variance Attempt at rankings (allow up to 4 errors)
	374	5	355	2		
	398	7	361	4		
	415	10	384	6		
	427	11	401	8		
	452	14	414	9		
	461	16	437	12		
	482	18	443	13		
	502	20	454	15		
	512	22	472	17		
	545	23	491	19		
	612	24	506	21		
	Test statistic = 127				A1	Clearly identified
	H_0 : population medians are equal H_1 : population median for X is greater than population median for Y				B1	
	Mean = $\frac{1}{2} \times 12 \times 25 = 150$				B1	
	Variance = $\frac{1}{12} \times 12 \times 12 \times 25 = 300$				B1	
(a)	$\frac{127.5 - 150}{\sqrt{300}}$				M1	Allow incorrect or no continuity correction.
	-1.299				A1	
	Compare with -1.645: -1.299 > -1.645, or 0.097 > 0.05 Accept H_0				M1	Valid comparison with 1.645 or 0.05 and reach correct fit conclusion.
	Insufficient evidence to support manager's claim				A1	Correct conclusion in context, following correct work, level of uncertainty in language. 'Prove' is A0.
					9	
(b)	Not appropriate/no, not the same people				B1	OE No and reason needed, e.g. individuals in the samples cannot be paired up.
					1	



13 - (9231/41_Summer_2023_Q4) - Non Parametric Test

(a)	<table><tr><td>2</td><td>5</td><td>1</td><td>-6</td><td>42</td><td>11</td><td>-3</td><td>7</td><td>20</td><td>12</td><td>-8</td><td>-10</td><td>50</td></tr><tr><td>2</td><td>4</td><td>1</td><td>-5</td><td>12</td><td>9</td><td>-3</td><td>6</td><td>11</td><td>10</td><td>-7</td><td>-8</td><td>13</td></tr></table>	2	5	1	-6	42	11	-3	7	20	12	-8	-10	50	2	4	1	-5	12	9	-3	6	11	10	-7	-8	13	M1	Differences (allow up to 3 errors)
	2	5	1	-6	42	11	-3	7	20	12	-8	-10	50																
	2	4	1	-5	12	9	-3	6	11	10	-7	-8	13																
		A1	Correct rank order, ignore signs																										
	$[P = 68,] \quad Q = 23$	A1	CWO																										
	H_0 : difference in population medians = 0 H_1 : population median in 2022 > population median in 2018	B1	'Population' required. Accept use of m , not μ . Do not accept 'difference between population medians > 0' without 2022, 2018 OE specified																										
	Critical value 21	B1																											
	'23' > '21' and accept H_0	M1	Compare their 23 with their 21 and FT conclusion. Their 23 must be less than 46.																										
	Insufficient evidence to support researcher's belief / insufficient evidence that the (median) number of employees in 2022 is greater than the (median) number of employees in 2018	A1	Correct conclusion in context, following correct work, level of uncertainty in language (not 'prove', not 'there is no evidence'), no contradictions. A0 if hypotheses reversed																										
		7																											
(b)	Rank for G will now be +3, giving $Q = 20$ which is < 21	M1	Must include numbers, ft their 21 from part (a) and ft their 20 (23 – 3)																										
	Change in conclusion	A1	CWO Condone 'reject H_0 ' OE.																										
		2																											

14 - (9231/43_Summer_2023_Q3) - Non Parametric Test

(a)	<table><tr><td>5</td><td>7</td><td>-6</td><td>1</td><td>-4</td><td>2</td><td>8</td><td>26</td><td>-10</td></tr><tr><td>4</td><td>6</td><td>-5</td><td>1</td><td>-3</td><td>2</td><td>7</td><td>9</td><td>-8</td></tr></table>	5	7	-6	1	-4	2	8	26	-10	4	6	-5	1	-3	2	7	9	-8	M1	Differences, allow at most 2 errors.
	5	7	-6	1	-4	2	8	26	-10												
	4	6	-5	1	-3	2	7	9	-8												
		A1	Correct rank order, ignore signs.																		
	$[P = 29] \quad Q = 16$	A1	Condone P not excluded.																		
	H_0 : population medians equal or $m_1 = m_2$ H_1 : population median $X >$ population median Y or $m_1 > m_2$	B1	'Population' required. Accept use of m , not μ . Do not accept 'difference between population medians > 0 ' without X , Y OE specified.																		
	Critical value = 8	B1																			
	$16 > 8$ and accept H_0 / not significant	M1	Compare their '16' with their '8' and conclusion. Their '16' must be less than 23. Ignore <i>their</i> hypotheses. Condone 'reject H_1 '.																		
Insufficient evidence to support teacher's belief or insufficient evidence that the marks/median in Paper 1 are/is higher than the marks/median in Paper 2	A1	Correct conclusion in context, following correct work, level of uncertainty in language (not 'prove', not 'there is no evidence'), no contradictions. e.g. Proves that the teacher is incorrect scores A0, A0 if hypotheses wrong way round.																			
	7																				
(b)	The population differences are symmetrical (about the median difference)	B1	Words in bold, or their equivalent, are required.																		
		1																			



15 - (9231/41_Winter_2023_Q6) - Non Parametric Test

H_0 : population medians are equal H_1 : population median after > population median before	B1 Do not accept 'differences between population medians < 0' unless difference defined. Allow $H_0: m = 22, H_1: m > 22.0$																																																		
<table><tr><td>-0.8</td><td>-7</td><td>1.5</td><td>12</td><td>0.9</td><td>8</td><td>-3.4</td><td>-19</td><td>-2.6</td><td>-18</td></tr><tr><td>0.1</td><td>1</td><td>4.5</td><td>21</td><td>-1.8</td><td>-14</td><td>3.7</td><td>20</td><td>-1.4</td><td>-11</td></tr><tr><td>0.3</td><td>3</td><td>-4.6</td><td>-22</td><td>0.2</td><td>2</td><td>5</td><td>23</td><td>1.9</td><td>15</td></tr><tr><td>6.2</td><td>25</td><td>0.6</td><td>5</td><td>5.2</td><td>24</td><td>1</td><td>9</td><td>1.7</td><td>13</td></tr><tr><td>-2.2</td><td>-16</td><td>0.7</td><td>6</td><td>1.3</td><td>10</td><td>-0.5</td><td>-4</td><td>2.3</td><td>17</td></tr></table>	-0.8	-7	1.5	12	0.9	8	-3.4	-19	-2.6	-18	0.1	1	4.5	21	-1.8	-14	3.7	20	-1.4	-11	0.3	3	-4.6	-22	0.2	2	5	23	1.9	15	6.2	25	0.6	5	5.2	24	1	9	1.7	13	-2.2	-16	0.7	6	1.3	10	-0.5	-4	2.3	17	M1 Signed differences, allow at most 4 errors M1 Attempt at ranks (ignore signs).
-0.8	-7	1.5	12	0.9	8	-3.4	-19	-2.6	-18																																										
0.1	1	4.5	21	-1.8	-14	3.7	20	-1.4	-11																																										
0.3	3	-4.6	-22	0.2	2	5	23	1.9	15																																										
6.2	25	0.6	5	5.2	24	1	9	1.7	13																																										
-2.2	-16	0.7	6	1.3	10	-0.5	-4	2.3	17																																										
$(W_+ = 214) \quad W_- = 111$	A1 Cao identified, or used, as test statistic.																																																		
Normal: mean = $\frac{1}{4}n(n+1) = \frac{1}{4} \times 25 \times 26 = 162.5$	B1																																																		
Variance = $\frac{1}{24}n(n+1)(2n+1) = \frac{1}{24} \times 25 \times 26 \times 51 = 1381.25$	B1																																																		
z- value: $\frac{111.5 - 162.5}{\sqrt{1381.25}}$	M1 Allow missing or incorrect continuity correction. Their 111 must come from ranks.																																																		
-1.37	A1 CAO																																																		
Tabular value is -1.645 ; $-1.37 > -1.645$, or $0.915 < 0.95$ oe, accept H_0 .	M1 Consistent signs. Allow 'not significant'.																																																		
Insufficient evidence to support the teacher's claim Insufficient evidence to suggest that the distances thrown are further when it is hot	A1 All correct. Correct conclusion in context, following correct work, level of uncertainty in language. A0 if hypotheses wrong way round or missing																																																		
	10																																																		

16 - (9231/42_Winter_2023_Q5) - Non Parametric Test

(a)	H_0 : population medians are equal or $m_x = m_y$ H_1 : population median for X < population median for Y or $m_x < m_y$	B1	Do not accept 'difference between population medians < 0' without X or Y oe specified.																																								
	<table><tr><td>X</td><td></td><td>Y</td><td></td></tr><tr><td>4.0</td><td>1</td><td>4.5</td><td>2</td></tr><tr><td>4.6</td><td>3</td><td>4.9</td><td>6</td></tr><tr><td>4.7</td><td>4</td><td>5.1</td><td>8</td></tr><tr><td>4.8</td><td>5</td><td>5.3</td><td>10</td></tr><tr><td>5.0</td><td>7</td><td>5.4</td><td>11</td></tr><tr><td>5.2</td><td>9</td><td>5.7</td><td>13</td></tr><tr><td>5.6</td><td>12</td><td>5.9</td><td>15</td></tr><tr><td>5.8</td><td>14</td><td>6.3</td><td>16</td></tr><tr><td></td><td></td><td>6.4</td><td>17</td></tr></table> Sum: 55 98	X		Y		4.0	1	4.5	2	4.6	3	4.9	6	4.7	4	5.1	8	4.8	5	5.3	10	5.0	7	5.4	11	5.2	9	5.7	13	5.6	12	5.9	15	5.8	14	6.3	16			6.4	17	M1	Rankings, allow at most 3 errors.
	X		Y																																								
	4.0	1	4.5	2																																							
	4.6	3	4.9	6																																							
	4.7	4	5.1	8																																							
	4.8	5	5.3	10																																							
	5.0	7	5.4	11																																							
5.2	9	5.7	13																																								
5.6	12	5.9	15																																								
5.8	14	6.3	16																																								
		6.4	17																																								
Test statistic = 55	A1																																										
Tabular value for $m=8, n=9$ is 54	B1																																										
'55' > '54', accept H_0 /not significant	M1	Ft <i>their</i> '55'. Must come from ranks. Ft <i>their</i> '54', must come from table.																																									
Insufficient evidence to support manager's claim. Insufficient evidence to suggest that the median time of machine X is less than the median time of machine Y .	A1	Correct conclusion in context, following correct work, level of uncertainty in language. A0 if hypothesises the wrong way round or missing.																																									
	6																																										



(b)	$H_0: \mu_x = \mu_y$ $H_1: \mu_x < \mu_y$	B1	
	$\sum x = 39.7$ $\sum x^2 = 199.33$ $\sum y = 49.5$ $\sum y^2 = 275.47$ $s_x^2 = \frac{1}{7} \left(199.33 - \frac{39.7^2}{8} \right) = 0.33125$	B1	
	$s_y^2 = \frac{1}{8} \left(275.47 - \frac{49.5^2}{9} \right) = 0.4025$	B1	
	Pooled variance $s^2 = \frac{7 \times 0.33125 + 8 \times 0.4025}{8 + 9 - 2}$	M1	
	0.36925	A1	
	$t = \frac{\frac{39.7}{8} - \frac{49.5}{9}}{s \sqrt{\frac{1}{8} + \frac{1}{9}}}$	M1	
	$t = -1.82$	A1	
	Tabular value = 1.753: $1.82 > 1.753$	M1	
	Reject H_0 , sufficient evidence that mean for machine X is less than mean for machine Y.	A1	CWO Correct conclusion in context, following correct work, level of uncertainty in language.
(c)	t -test is assuming a normal distribution, and with equal variances. This may not be true. So, no reason to expect results to be the same. Outliers affect part (b) but not part (a).	B1	Not specific to data in question. Mention of normal distribution is not enough.
		1	



1 - (9231/41_Summer_2020_Q3) - Continuous Random Variable

(a)	$E(X) = \int_0^1 \frac{3}{16} x(2 - \sqrt{x}) dx + \int_1^9 \frac{3}{16} \frac{x}{\sqrt{x}} dx$	M1
	$\frac{3}{16} \left[x^2 - \frac{2}{5} x^{5/2} \right] + \frac{3}{16} \left[\frac{2}{3} x^{3/2} \right]$	A1
	$\frac{269}{80}$	A1
		3
(b)	$F(x) = \begin{cases} \frac{3}{8} \left(x - \frac{1}{3} x^{3/2} \right), & 0 \leq x < 1 \\ \frac{3}{8} x^{1/2} - \frac{1}{8}, & 1 \leq x \leq 9 \end{cases}$	M1A1
	$G(y) = \begin{cases} \frac{3}{8} \left(y^2 - \frac{1}{3} y^3 \right), & 0 \leq y < 1 \\ \frac{3}{8} y - \frac{1}{8}, & 1 \leq y \leq 3 \end{cases}$	M1A1
	$(G(y) = 0 \text{ for } y \leq 0 \text{ and } = 1 \text{ for } y \geq 3)$ $g(y) = \begin{cases} \frac{3}{8} (2y - y^2), & 0 \leq y < 1 \\ \frac{3}{8}, & 1 \leq y \leq 3 \\ = 0 \text{ otherwise} \end{cases}$	A1
		5

2 - (9231/43_Summer_2020_Q3) - Continuous Random Variable

(a)	$F(x) = \begin{cases} \frac{x^2}{10}, & 0 \leq x < 2 \\ \frac{1}{15} (10x - x^2) - \frac{2}{3}, & 2 \leq x \leq 5 \end{cases}$	M1A1
	$= 0 \text{ for } x < 0 \text{ and } = 1 \text{ for } x > 5$	A1
		3
(b)	$F(m) = \frac{1}{2} \text{ so } 2m^2 - 20m + 35 = 0$	M1
	$m = 5 - \frac{1}{2} \sqrt{30} = 2.26$	A1
		2
(c)	$E(X^2) = \int_0^2 x^2 \cdot \frac{x}{5} dx + \int_2^5 \frac{2}{15} (5-x) \cdot x^2 dx = \left[\frac{x^4}{20} \right] + \left[\frac{2}{15} \left(\frac{5}{3} x^3 - \frac{x^4}{4} \right) \right]$	M1
	$= 6.5$	A1
		2



(d)	$F(3) - F(1)$	M1
	$= \frac{11}{15} - \frac{1}{10} = \frac{19}{30}$	A1
Alternative method for 3(d)		
	$\int_1^3 \frac{x}{5} dx + \int_2^3 \frac{2}{15}(5-x) dx = \left[\frac{x^2}{10} \right] + \left[\frac{2}{15} \left(5x - \frac{x^2}{2} \right) \right]$	M1
	$= \frac{3}{10} + \frac{1}{3} = \frac{19}{30}$	A1
		2

3 - (9231/42_Winter_2020_Q4) - Continuous Random Variable

(a)	$F(6) - F(3) = \frac{9}{20}$ or 0.45	B1	
		1	
(b)	$f(x) = \begin{cases} \frac{1}{30}x, & 2 \leq x \leq 8 \\ 0, & \text{otherwise} \end{cases}$	B1	May be implied. Only need to see $f(x) = \frac{1}{30}x$
	$E(\sqrt{X}) = \int_2^8 \frac{1}{30}x^{\frac{3}{2}} dx = \left[\frac{1}{75}x^{\frac{5}{2}} \right]$	M1	Integrated
	$= 2.34$	A1	2.338... Answer of 2.34 without $\frac{1}{75}x^{\frac{5}{2}}$ scores B1B1 (2/3)
		3	
(c)	$\text{Var}(\sqrt{X}) = \int_2^8 \frac{1}{30}x^2 dx = \left[\frac{1}{90}x^3 \right] - E(\sqrt{X})^2$ $\text{Var}(\sqrt{X}) = \left[\frac{1}{30}x^2 \right] - E(\sqrt{X})^2$	M1	
	$= 5.6 - 2.338^2$ $= 0.133$ or 0.134	A1	
		2	
(d)	$G(y) = \frac{1}{60}y^{\frac{2}{3}} - \frac{1}{15}$	M1	CDF for Y
	$g(y) = \frac{1}{90}y^{-\frac{1}{3}}$ for $8 \leq y \leq 512$ (0 otherwise)	M1 A1	Differentiate to find PDF for Y Correct $g(y)$ and correct range seen anywhere
		3	



4 - (9231/41_Winter_2020_Q6) - Continuous Random Variable

(a)	UQ: $F(U) = 0.75: u^2 - 16u + 45 = 0$	M1	Obtain quadratic equation for u or l
	$u = 8 - \sqrt{19} \quad (= 3.64)$	A1	Value of UQ
	LQ: $l^2 - 16l + 15 = 0, \quad l = 1$	A1	Value of LQ
	IQR = UQ - LQ = $7 - \sqrt{19} = 2.64$	A1 FT	
		4	
(b)	$f(x) = \begin{cases} \frac{1}{30}(8-x), & 0 \leq x \leq 6 \\ 0, & \text{otherwise} \end{cases}$	B1	May be implied, only need to see $f(x) = \frac{1}{30}(8-x)$
	Find $E(X^3)$ by integration: $\frac{1}{30} \int_0^6 (8x^3 - x^4) dx$ $= \left[2x^4 - \frac{1}{5}x^5 \right]_0^6$	M1 A1	Attempt integration (with PDF not CDF) limits not required Correct integrated expression, with limits
	$= 34.56$	A1	
		4	
(c)	$G(y) = \frac{1}{60}(16y^2 - y^4)$	M1	CDF for Y
	$g(y) = \frac{1}{60}(32y - 4y^3) \quad \text{for } 0 \leq y \leq \sqrt{6}$	M1A1	Differentiate to find PDF for Y Correct $g(y)$ and correct range seen anywhere
		3	

5 - (9231/41_Summer_2021_Q3) - Continuous Random Variable

(a)	$\int_0^9 \frac{2}{81} x^{1.5} dx$	B1	$f(x) = \frac{2}{81}x$, correct expression as integrand.
	$\frac{2}{81} \times \frac{2}{5} x^{2.5}$	M1	Integrate, FT only on their PDF expression.
	$\frac{12}{5} = 2.4$	A1	
		3	
(b)	$\int_0^9 \frac{2}{81} x^2 dx - (their a)^2$	M1	
	$\frac{2}{81} \times \frac{9^3}{3} - 2.4^2 = \frac{6}{25}$	A1	
		2	
(c)	$G(y) = \frac{1}{81} y^6$	M1	
	$g(y) = \frac{2}{27} y^5$	M1	
	Fully correct including 'for $0 \leq y \leq \sqrt[3]{9}$ and 0 otherwise'	A1	Condone 2.08
		3	



6 - (9231/43_Summer_2021_Q6) - Continuous Random Variable

(a)	$F(x) = \begin{cases} \frac{1}{8}x & 0 \leq x < 1, \\ \frac{1}{28}\left(8x - \frac{1}{2}x^2\right) - \frac{1}{7} & 1 \leq x \leq 8. \end{cases}$	M1 Integrate both parts (all powers of x increase by 1), allow M1 if no constant term. A1 Both parts correct, with domains. Any equivalent form, for example: $\frac{2}{7}x - \frac{1}{56}x^2 - \frac{1}{7}$.
	$F(x) = 0$ for $x < 0$ and $F(x) = 1$ for $x > 8$.	A1
Alternative method for question 6(a)		
	$F(x) = \begin{cases} \frac{1}{8}x & 0 \leq x < 1, \\ 1 - \frac{1}{56}(8-x)^2 & 1 \leq x \leq 8. \end{cases}$	M1 Integrate both parts (all powers of x increase by 1), allow M1 if no constant term. A1 Both parts correct, with domains. Any equivalent form, for example: $\frac{1}{56}(16x - x^2 - 8)$.
	$F(x) = 0$ for $x < 0$ and $F(x) = 1$ for $x > 8$.	A1
		3
(b)	$F(a) = \frac{5}{7}$ leading to $\frac{1}{28}\left(8a - \frac{1}{2}a^2\right) - \frac{1}{7} = \frac{5}{7}$	*M1 Correct method and attempt to simplify, ft their part (a).
	$a^2 - 16a + 48 = 0$ $a = 4$ [$a = 12$]	DM1 Obtain quadratic and solve.
	$a = 4$	A1 Correct single answer, WWW.
		3
(c)	$G(y) = \begin{cases} \frac{1}{8}y^3 & 0 \leq y < 1, \\ \frac{1}{56}(16y^3 - y^6) - \frac{1}{7} & 1 \leq y \leq 2. \end{cases}$	M1 Substitute $x = y^3$ into both parts of their $F(x)$ from part (a). A1 Correct functions in terms of y , any equivalent form.
	$0 \leq y < 1$ and $1 \leq y \leq 2$ seen in CDF or PDF	A1 Condone $\leq$ instead of $<$.
	$g(y) = \begin{cases} \frac{3}{8}y^2 & 0 \leq y < 1, \\ \frac{3}{28}(8y^2 - y^5) & 1 \leq y < 2, \\ 0 & \text{otherwise.} \end{cases}$	M1 Differentiate their $G(y)$, all powers of y decrease by 1. A1 Must include 'otherwise'.
		5
(c)	Alternative method for question 6(c)	
	$y = x^{\frac{1}{3}}, \quad dy = \frac{1}{3}x^{-\frac{2}{3}}dx, \quad dx = 3y^2dy$	M1
	$0 \leq x < 1, \quad f(x)dx = \frac{1}{8}dx = \frac{3}{8}y^2dy$	M1
	$g(y) = \frac{3}{8}y^2$	A1
	$1 \leq x \leq 8, \quad f(x)dx = \frac{1}{28}(8-x)dx = \frac{1}{28}(8-y^3)3y^2dy$ $g(y) = \frac{3}{28}(8y^2 - y^5)$	A1
	$0 \leq y < 1$ and $1 \leq y \leq 2$ seen in CDF or PDF	A1
		5



7 - (9231/41_Winter_2021_Q2) - Continuous Random Variable

(a)	$f(x) = \begin{cases} 1+x & -1 \leq x \leq 0, \\ 1-x & 0 < x \leq 1, \\ 0 & \text{otherwise.} \end{cases}$	M1	Differentiation attempted.
		A1	All correct, including 0 otherwise.
		2	
(b)	$P\left(-\frac{1}{2} \leq X \leq \frac{1}{2}\right) = F\left(\frac{1}{2}\right) - F\left(-\frac{1}{2}\right) = \frac{7}{8} - \frac{1}{8}$	M1	Can do by integration with correct limits.
		A1	
		2	
(c)	$\int_{-1}^0 x^2(1+x)dx + \int_0^1 x^2(1-x)dx$ $\left[\frac{1}{3}x^3 + \frac{1}{4}x^4\right] + \left[\frac{1}{3}x^3 - \frac{1}{4}x^4\right]$	M1	
		A1	
		2	
(d)	$\int_{-1}^0 x^4(1+x)dx + \int_0^1 x^4(1-x)dx - \left(\frac{1}{6}\right)^2$ $\left[\frac{1}{5}x^5 + \frac{1}{6}x^6\right] + \left[\frac{1}{5}x^5 - \frac{1}{6}x^6\right] - \frac{1}{36}$	M1	Complete method with their mean squared explicit.
		A1	
		2	

8 - (9231/42_Winter_2021_Q3) - Continuous Random Variable

(a)	$\text{Total prob} = 1, \text{ so } \int_0^1 \left(a + \frac{1}{5}x\right)dx + \int_1^2 \left(2a - \frac{1}{5}x\right)dx = 1$ $\left[ax + \frac{1}{10}x^2\right] + \left[2ax - \frac{1}{10}x^2\right] = 1$	*M1	Correct expression integrated and equated to 1. Powers of x correct.
		DM1	Limits used and attempt to solve.
		A1	
		3	
(b)	$\int_0^1 \frac{1}{5}(2+x)x^2 dx + \int_1^2 \frac{1}{5}(4-x)x^2 dx$ $\frac{1}{5}\left[\frac{2}{3}x^3 + \frac{1}{4}x^4\right] + \frac{1}{5}\left[\frac{4}{3}x^3 - \frac{1}{4}x^4\right]$	M1	Correct expressions integrated, FT their a .
		A1	$\frac{11}{60} + \frac{67}{60}$
		2	



(c)	$F(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{5} \left(2x + \frac{1}{2}x^2 \right) & 0 \leq x < 1, \\ \frac{1}{5} \left(4x - \frac{1}{2}x^2 - 1 \right) & 1 \leq x \leq 2, \\ 1 & x > 2. \end{cases}$	M1	Integration of their PDF.
		A1	Middle 2 parts correct.
	$= 0 \quad (x < 0) \text{ and } = 1 \quad (x > 2)$	A1	First and last parts correct and all domains correct, with no gaps.
		3	

9 - (9231/41_Summer_2022_Q3) - Continuous Random Variable

(a)	$k \left[2x^2 - \frac{1}{3}x^3 + k \left[6x - \frac{1}{2}x^2 \right] \right] = 1$ $k \left[8 - \frac{8}{3} + 36 - 18 - 12 + 2 \right] = \frac{40}{3} \quad k = 1$ $k = \frac{3}{40}$	B1	Integrate and equate to 1, no errors seen.
		1	
(b)	$E(X^2) = k \int_0^2 (4x^3 - x^4) dx + k \int_2^6 (6x^2 - x^3) dx =$ $k \left[x^4 - \frac{1}{5}x^5 \right] + k \left[2x^3 - \frac{1}{4}x^4 \right] = k[9.6 + 96] = 7.92$	M1*	Attempt at $E(X^2)$, correct limits, integrated.
	$\text{Var}(X) = E(X^2) - E^2(X) = \text{their } (7.92) - 2.5^2$	depM1	Use of correct formula, with their $E(X^2)$.
	1.67	A1	
		3	
(c)	$\int_0^2 kx(4-x) dx + \int_2^m k(6-x) = \frac{1}{2}$	M1*	Correct expression and attempt to integrate.
	$\int_0^2 kx(4-x) dx = \frac{2}{5}$	B1	May be implied
	$\frac{2}{5} + k \left[6x - \frac{1}{2}x^2 \right] = \frac{1}{2}$ $6m - \frac{1}{2}m^2 - 12 + 2 = \frac{1}{10k}$ $m^2 - 12m + \frac{68}{3} = 0$	depM1	First integral and $\frac{1}{2}$ may be seen combined as $\frac{1}{10}$. Obtain quadratic equation.
	$m = 6 \pm \frac{2}{3}\sqrt{30}$ $m = 6 - \frac{2}{3}\sqrt{30} (= 2.35)$	A1	Single answer.
		4	



10 - (9231/43_Summer_2022_Q4) - Continuous Random Variable

(a)	$E(\sqrt{X}) = \int \sqrt{x} \cdot \frac{3}{8} \left(1 + \frac{1}{x^2}\right) dx = \int_1^3 \frac{3}{8} (x^{0.5} + x^{-1.5}) dx$	M1	
	$\frac{3}{8} \left[\frac{2}{3} x^{1.5} - 2x^{-0.5} \right]$	A1	
	$\frac{3}{8} \left(2\sqrt{3} - \frac{2}{\sqrt{3}} - \frac{2}{3} + 2 \right) = \frac{1}{2} (1 + \sqrt{3}) (= 1.37)$	A1	
		3	
(b)	$F(x) = \frac{3}{8} \left(x - \frac{1}{x} \right)$	M1	Integrate.
	$G(y) = \frac{3}{8} \left(\sqrt{y} - \frac{1}{\sqrt{y}} \right)$	M1	Change variable throughout in their CDF.
	$g(y) = \frac{3}{16} (y^{-0.5} + y^{-1.5}), \quad (1 \leq y \leq 9)$	A1	Differentiate to find g(y).
	0 otherwise	A1	$1 \leq y \leq 9$ and 0 otherwise, dependent on second M1.
		4	
(c)	$F(p) = 0.4$ so $\frac{3}{8} \left(p - \frac{1}{p} \right) = 0.4$	M1*	Correct method.
	$15p^2 - 16p - 15 = 0$ $(5p+3)(3p-5) = 0, \quad p = \frac{5}{3}$	M1 dep	Obtain quadratic and solve.
	40th percentile of Y is $\frac{25}{9}$	A1	
		3	

11 - (9231/42_Winter_2022_Q4) - Continuous Random Variable

(a)	$\int_0^1 kx dx + \int_1^2 kx dx = [kx] + \left[\frac{k}{2} x^2 \right], \quad k + \frac{3}{2}k = 1, \quad k = \frac{2}{5}$	B1	Integration and $k + 3k/2 = 1$ or seen.
		1	
(b)	$F(x) = \begin{cases} kx & 0 \leq x < 1, \\ \frac{k}{2} x^2 + \frac{1}{5} & 1 \leq x \leq 2 \end{cases}$	M1	CDF or integration with $\frac{1}{5}x^2 + \frac{1}{5}$ seen.
	UQ: $\frac{k}{2} x^2 + \frac{1}{5} = \frac{3}{4}$	M1	Method for UQ
	$\left[x^2 + 1 = 5 \times \frac{3}{4}, \quad x^2 = \frac{11}{4} \right] \quad x = \frac{1}{2} \sqrt{11}$	A1	UQ
	LQ: $kx = \frac{1}{4}, \quad x = \frac{5}{8}$	B1	LQ
	$IQR = \frac{1}{2} \sqrt{11} - \frac{5}{8} = 1.03(3)$	A1	CAO
		5	



(c)	$E(X) = \int_0^1 kx \, dx + \int_1^2 kx^2 \, dx = \left[\frac{k}{2} x^2 \right] + \left[\frac{k}{3} x^3 \right] = \frac{17}{15}$	M1	With correct limits.
	$E(X^2) = \int_0^1 kx^2 \, dx + \int_1^2 kx^3 \, dx = \left[\frac{k}{3} x^3 \right] + \left[\frac{k}{4} x^4 \right] = \frac{49}{30}$	M1	With correct limits.
	$\text{Var}(X) = \frac{49}{30} - \left(\frac{17}{15} \right)^2$	M1	Using correct formula with numerical values.
	$\frac{157}{450} \quad (= 0.349)$	A1	
		4	

12 - (9231/41_Winter_2022_Q5) - Continuous Random Variable

(a)	$F(U) = 0.75$, so $1 - \frac{1}{144}(12-U)^2 = \frac{3}{4}$	M1	
	$(12-U)^2 = 36$, $12-U = \pm 6$, $U = 6$ only	A1	
		2	
(b)	$f(x) = \frac{1}{72}(12-x)$	B1	Differentiate given PDF.
	$E(X^2) = \int_0^{12} \frac{1}{72}(12x^2 - x^3) \, dx = \frac{1}{72} \left(4x^3 - \frac{1}{4}x^4 \right) [= 24]$	M1	Integrate $x^2 f(x)$ with their $f(x)$.
	$E(X^4) = \int_0^{12} \frac{1}{72}(12x^4 - x^5) \, dx = \frac{1}{72} \left(\frac{12}{5}x^5 - \frac{1}{6}x^6 \right) [= 1382.4]$	M1	Integrate $x^4 f(x)$ with their $f(x)$.
	$\text{Var}(X^2) = E(X^4) - E^2(X^2) [= 1382.4 - 24^2]$	M1	Using correct formula with numerical values possibly unsimplified.
	$1382.4 - 576 = 806.4$	A1	
		5	
(c)	$G(y) = \begin{cases} 0 & y < 0, \\ 1 - \frac{1}{144}(12-y^2)^2 & 0 \leq y \leq \sqrt{12}, \\ 1 & y > \sqrt{12}. \end{cases}$	M1	Change the variable: $1 - \frac{1}{144}(12-y^2)^2$ Allow domain unchanged at this stage.
	$\frac{1}{36}(12y - y^3)$	M1	Differentiate their $G(y)$.
	$g(y) = \begin{cases} \frac{1}{36}(12y - y^3) & 0 \leq y \leq \sqrt{12} \\ 0 & \text{otherwise} \end{cases}$	A1	Fully correct.

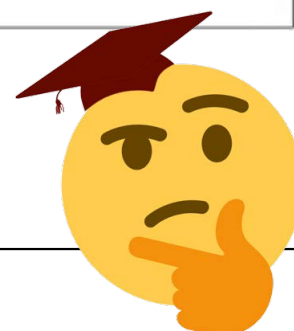


13 - (9231/41_Summer_2023_Q3) - Continuous Random Variable

(a)	$\int_{1.5}^2 \frac{1}{24} \left(\frac{4}{x^2} + x^2 \right) dx = \frac{1}{24} \left[-\frac{4}{x} + \frac{x^3}{3} \right]$	M1	Integration with correct powers and correct limits seen.
	$a = 50 \times 0.092014 = 4.6007$	A1	AG $\frac{53}{576}$ or $\frac{1325}{288}$ or 0.092014 or 4.60069 seen
	$b = 11.4211$	B1	Or 11.421(0)
		3	
(b)	Combine first two columns: 6, 9.0278	M1	Must be seen, or implied by 1.0155.
	Calculate value of chi-squared: $1.0155 + 0.5715 + 0.7661 + 0.2183 + 0.5885$	M1	At least 2 correct values (at least 3sf) or expressions seen. Allow columns not combined or three columns combined.
	3.16	A1	CWO. Dependent on M1M1 scored. SCB1 for 3.16 with no working
	H_0 : f is a good fit for the data H_1 : f is not a good fit for the data	B1	
	Tabular value of chi-squared: 7.78 "3.16" < 7.78 and accept H_0	M1	Correct tabular value; allow correct FT value if columns not combined (9.236) or three columns combined (6.251).
	Insufficient evidence to suggest that f is not a good fit for the data Condone: sufficient evidence to suggest that f is a good fit for the data.	A1	Correct conclusion in context, following correct work, level of uncertainty in language (not 'prove', not 'there is no evidence'), no contradictions. A0 if hypotheses reversed
		6	

14 - (9231/41_Summer_2023_Q6) - Continuous Random Variable

(a)	$F(x) = \frac{3}{28} \left(2e^{\frac{1}{2}x} - 8e^{-\frac{1}{2}x} + c \right)$	M1	+ c not required.
	$F(x) = \frac{3}{28} \left(2e^{\frac{1}{2}x} - 8e^{-\frac{1}{2}x} \right) + \frac{9}{14}$	A1	
	Correct ranges including $0 \leq x \leq 2\ln 3$ associated with their $F(x)$ $F(x) = 0$ for $x < 0$ and 1 for $x > 2\ln 3$	B1	No gaps in range.
		3	
(b)	$G(y) = \frac{3}{28} \left(2y - \frac{8}{y} + 6 \right)$	M1	y substituted into their F .
	$1 \leq y \leq 3$	B1	Seen anywhere, condone $1 \leq y \leq e^{\ln 3}$.
	$g(y) = \frac{3}{28} \left(2 + \frac{8}{y^2} \right)$	A1	Correct expression, not containing logs.
		3	
(c)	$\frac{3}{28} \left(2y - \frac{8}{y} + 6 \right) = \frac{3}{10}$	M1	Their $G(y) = \frac{3}{10}$
	$5y^2 + 8y - 20 = 0$	M1	Obtain from $G(y)$ and solve quadratic equation to find y .
	$y = 1.35$	A1	Single correct answer.
		3	



(d)	$E(Y^4) = \int_1^3 \frac{3}{28} \left(2 + \frac{8}{y^2} \right) y^4 dy = \int_1^3 \frac{3}{28} (2y^4 + 8y^2) dy = \frac{3}{28} \left[\frac{2}{5} y^5 + \frac{8}{3} y^3 \right]$	1	Integrate $y^4 \times \text{their } g(y)$. Limits must be 1 and 3.
	17.8	A1	$\frac{89}{5}$
		2	

15 - (9231/43_Summer_2023_Q1) - Continuous Random Variable

(a)	$F(x) = \frac{1}{6} \left[\frac{3}{2} x^{\frac{2}{3}} - 3x^{\frac{1}{3}} \right] + c$	M1	Integrate $f(x)$, + c not required, at least one power correct.
	$\frac{1}{4} x^{\frac{2}{3}} - \frac{1}{2} x^{\frac{1}{3}} + \frac{1}{4}$	A1	AEF
	Correct ranges including $1 \leq x \leq 27$ associated with their $F(x)$ $F(x) = 0$ for $x < 1$ and $F(x) = 1$ for $x > 27$	A1	No gaps.
		3	
(b)	$G(y) = \frac{1}{4} (y^2 - 2y + 1)$	M1	Using $y = x^{\frac{1}{3}}$ in their $F(x)$.
	$g(y) = \frac{1}{2} (y - 1)$	A1	Correct, AEF, 0 otherwise not required.
	For $1 \leq y \leq 3$	B1	Seen anywhere, correct variable.
		3	
(c)	$\frac{1}{4} (y^2 - 2y + 1) = \frac{1}{2}, \quad (y - 1)^2 = 2$ OE	M1	Equate their $G(y)$ to $\frac{1}{2}$ and solve to find y .
	$[y =] \quad 1 + \sqrt{2}$	A1	Only.
		2	

16 - (9231/41_Winter_2023_Q4) - Continuous Random Variable

(a)	$P(X \leq 2) = \frac{1}{3}, \text{ so } \left[\frac{mx^3}{2} \right]_0^2 = \frac{1}{3}$ $m \left(\frac{4}{2} \right) = \frac{1}{3}, \quad m = \frac{1}{6}$	B1	AG Reasoning required.
	$\int_2^6 \left(\frac{k}{x^2} + c \right) dx = \frac{2}{3} \text{ so } \left[-\frac{k}{x} + cx \right]_2^6 = 4c + \frac{1}{3}k = \frac{2}{3} \text{ or } k + 12c = 2$	M1	Attempt to integrate and correct use of correct limits to form linear equation in k and c .
	Also, $2m = \frac{k}{4} + c$ or $\frac{1}{3} = \frac{k}{4} + c$ or $3k + 12c = 4$	M1	Matching at $x = 2$.
	Solve, $k = 1, \quad c = \frac{1}{12}$	A1	
		4	



(b)	LQ: $\int_0^L \frac{1}{6} x dx = \frac{1}{4}, \quad L = \sqrt{3}$	B1	
	UQ: $\frac{1}{3} + \int_{\frac{1}{2}}^U \left(\frac{k}{x^2} + c \right) dx = \frac{3}{4}$	M1	Or find and use CDF. May be in terms of m .
	$-\frac{k}{U} + Uc + \frac{k}{2} - 2c = \frac{5}{12}$ $U^2 - U - 12 = 0$	M1	Attempt at integral and correct use of correct limits. Simplify to quadratic equation, may be in terms of k and c .
	$U = 4$	A1	
	$IQR = 4 - \sqrt{3}$	A1 FT	FT their UQ – their LQ, exact values only.
		5	

17 - (9231/42_Winter_2023_Q4) - Continuous Random Variable

(a)	UQ = 4 so $\int_4^6 c dx = \frac{1}{4}; \quad c[6-4] = \frac{1}{4}, \quad c = \frac{1}{8}$	B1	AG. Some indication of reasoning required.
	$\int_0^4 \frac{1}{128} (4ax - bx^3) dx = \frac{3}{4}, \quad \frac{1}{128} \left[2ax^2 - \frac{b}{4} x^4 \right] = \frac{3}{4}, \quad 32a - 64b = 96$	M1	Integrate and use correct limits to form equation in a and b .
	Equate at $x = 4; \quad a - 4b = 1$	M1	
	Solve: $a = 5, b = 1$	A1	
		4	
(b)	$\int_0^m \frac{1}{128} (20x - x^3) dx = \frac{1}{2}$ OR $F(x) = \frac{1}{128} \left(20x^2 - \frac{b}{4} x^4 \right) = \frac{5}{64} x^2 - \frac{1}{512} x^4$	M1*	Attempt at integral and use of correct limits. Could be in terms of a, b, c . Or attempt at relevant part of CDF and set $F(m) = \frac{1}{2}$. Could be in terms of a, b, c .
	$m^4 - 40m^2 + 256 = 0$	M1	Simplify to quartic. Could be in terms of a, b, c .
	$m = 2\sqrt{2}$	A1	Positive answer only.
		3	
(c)	$E(\sqrt{X}) = \frac{1}{128} \int_0^4 20x^{\frac{3}{2}} - x^{\frac{7}{2}} dx + \frac{1}{4} \int_4^6 \frac{1}{8} x^{\frac{1}{2}} dx$	M1	Their PDF multiplied by $\sqrt{x}$ with correct limits. Could be in terms of a, b, c .
	$\frac{1}{128} \left[8x^{\frac{5}{2}} - \frac{2}{9} x^{\frac{9}{2}} \right]_0^4 + \frac{1}{12} \left[\frac{2}{3} x^{\frac{3}{2}} \right]_4^6$	M1	Correct use of correct limits in their integral. Could be in terms of a, b, c . May be implied by correct final answer.
	$\frac{1}{128} \left(2^8 - \frac{1}{9} 2^{10} \right) + \frac{1}{12} (6\sqrt{6} - 8) = \frac{4}{9} + \frac{1}{2} \sqrt{6} = 1.67$	A1	CAO 1.67 following (first) M0 SC B2
		3	



1 - (9231/43_Summer_2020_Q4) - Probability Generating Function

(a)	$G_X(t) = 0.2t + 0.5t^2 + 0.3t^3$	B1
	$G_Y(t) = (0.2t + 0.5t^2 + 0.3t^3)^2$	M1
	$= 0.04t^2 + 0.2t^3 + 0.37t^4 + 0.3t^5 + 0.09t^6$	A1
		3
(b)	$G''_Y(t) = 0.08t + 0.6t^2 + 1.48t^3 + 1.5t^4 + 0.54t^5$	M1
	$E(Y) = 4.2$	A1
	$G''_{YY}(t) = 0.08 + 1.2t + 4.44t^2 + 6t^3 + 2.7t^4$	M1
	Use $G''_{YY}(1) + G'_Y(1) - (G'_Y(1))^2$	M1
	$\text{Var}(Y) = 14.42 + 4.2 - 4.2^2 = 0.98$	A1
		5

2 - (9231/41_Summer_2020_Q6) - Probability Generating Function

(a)	$P(\text{BB}) = \frac{6}{10} \times \frac{5}{9} = \frac{1}{3}$ $P(\text{RB}, \text{BR}) = \frac{6}{10} \times \frac{4}{9} = \frac{8}{15}$ $P(\text{RR}) = \frac{4}{10} \times \frac{3}{9} = \frac{2}{15}$	M1
	$\frac{1}{3} + \frac{8}{15}t + \frac{2}{15}t^2$ (FT their probabilities)	A1 FT
		2
(b)	$P(1 \text{ H}) = \frac{2}{3}(1-p) + \frac{1}{3}p = \frac{7}{12}$	M1
	$p = \frac{1}{4}$	A1
	PGF of $Y = \frac{1}{4} + \frac{7}{12}t + \frac{1}{6}t^2$	A1
		3
(c)	PGF of $Z = \left(\frac{1}{3} + \frac{8}{15}t + \frac{2}{15}t^2\right)\left(\frac{1}{4} + \frac{7}{12}t + \frac{1}{6}t^2\right)$	B1M1
	$= \frac{1}{180}(15 + 59t + 72t^2 + 30t^3 + 4t^4)$ AEF	A1
		3
(d)	Attempt to differentiate their $G_X(t)$ and evaluate $G'_X(1)$	M1
	$E(Z) = (59 + 144 + 90 + 16)/180 = \frac{309}{180} = 1.72$	A1
		2



3 - (9231/41_Winter_2020_Q5) - Probability Generating Function

(a)	$G_X(t) = \frac{1}{4} + \frac{1}{2}t + \frac{1}{4}t^2$	B1	Accept $(0.5+0.5t)^2$
		1	
(b)	$P(0H) = \frac{12}{36}$ $P(1H) = \frac{16}{36}$ $P(2H) = \frac{7}{36}$ $P(3H) = \frac{1}{36}$	M1 A1	Attempt at probs, at least 2 correct All correct
	$G_Y(t) = \frac{12}{36} + \frac{16}{36}t + \frac{7}{36}t^2 + \frac{1}{36}t^3$	B1 FT	FT <i>their</i> probabilities, must be cubic with 4 non-zero terms
		3	
(c)	$G_Z(t) = \left(\frac{1}{4} + \frac{1}{2}t + \frac{1}{4}t^2\right)\left(\frac{12}{36} + \frac{16}{36}t + \frac{7}{36}t^2 + \frac{1}{36}t^3\right)$	M1	Attempt to multiply their two PGF
	$= \frac{1}{144}(12 + 40t + 51t^2 + 31t^3 + 9t^4 + t^5)$	M1 A1	Obtain quintic expression and collect terms
		3	
(d)	$G'_Z(t) = \frac{1}{144}(40 + 102t + 93t^2 + 36t^3 + 5t^4)$	M1	Differentiate
	$E(Z) = G'_Z(1) = \frac{23}{12} = (= 1.92)$	A1	
		2	
(e)	2	B1 FT	FT power of term with largest coefficient in <i>their</i> $G_Z(t)$
		1	

4 - (9231/42_Winter_2020_Q5) - Probability Generating Function

(a)	$P(X=r) = {}^nC_r p^r (1-p)^{n-r}$ or ${}^nC_r p^r q^{n-r}$	B1	
	$G_X(t) = \sum_0^n nCr p^r (1-p)^{n-r} t^r$	M1	Accept minimum of 4 terms including the last. Shown with specific value of n is M0
	$\sum_0^n nCr (pt)^r (1-p)^{n-r}$ $= (q + pt)^n$	A1	At least one intermediate step to be shown, with p and t grouped AG
		3	
(b)	$G'_{X'}(t) = n(q + pt)^{n-1} \times p$	M1	
	So $E(X) = G'_{X'}(1) = np(q + p)^{n-1}$ and $q + p = 1$ so $E(X) = np$	A1	
	$G''_{X'}(t) = n(n-1)(q + pt)^{n-2} \times p \times p$	M1	
	$\text{Var}(X) = n(n-1)p^2 + np - (np)^2$	M1	
	$np(1-p)$	A1	
		5	



5 - (9231/43_Summer_2021_Q4) - Probability Generating Function

(a)	$G'(t) = \frac{4t}{(3-2t^2)^2}$	M1	Differentiate to obtain $\frac{kt}{(3-2t^2)^2}$ OE.
	so $E(X) = G'(1) = 4$	A1	CAO WWW
	$G''(t) = \frac{12+24t^2}{(3-2t^2)^3}$ or $4(3-2t^2)^{-2} + 32t^2(3-2t^2)^{-3}$	M1	OE. Differentiate, allow only numerical or sign slips.
	$\text{Var}(X) = G''(1) + G'(1) - (G'(1))^2 = 36 + 4 - 16$	M1	Substitute their values into correct formula, dependent on attempt at $G''(t)$.
	$[\text{Var}(X)] = 24$	A1	CAO WWW
		5	
(b)	$G_X(t) = \frac{1}{3-2t^2} = (3-2t^2)^{-1}$ $= \frac{1}{3} \left(1 + \frac{2}{3}t^2 + \frac{4}{9}t^4 + \dots \right)$	M1	Expand given expression or give expression for the term in t^4 .
	$P(X=4) = \text{their coefficient of } t^4$	M1	Found from legitimate method.
	$[P(X=4)] = \frac{4}{27}$	A1	WWW
		3	

6 - (9231/41_Summer_2021_Q6) - Probability Generating Function

(a)	$P(3R, 2R, 1R, 0R) = \frac{1}{5}, \frac{3}{5}, \frac{1}{5}, 0$	M1	At least 2 correct probabilities used in a polynomial.
	$G_X(t) = \frac{1}{5}t + \frac{3}{5}t^2 + \frac{1}{5}t^3$	A1	
		2	
(b)	$P(2H, 1H, 0H) = \frac{1}{16}, \frac{6}{16}, \frac{9}{16}$	M1	One correct probability and all three adding to 1 used in a polynomial.
	$G_T(t) = \frac{9}{16} + \frac{6}{16}t + \frac{1}{16}t^2$	A1	
		2	
(c)	Attempt to multiply results from part (a) and part (b)	M1	
	Obtain polynomial	M1	
	$\frac{9}{80}t + \frac{33}{80}t^2 + \frac{28}{80}t^3 + \frac{9}{80}t^4 + \frac{1}{80}t^5$	A1	Accept exact equivalent decimals.
		3	
(d)	$E(Z) = G'_Z(1) = \frac{1}{80}(9+66+84+36+5)$	M1	Differentiate and put $t=1$.
	$= \frac{200}{80} = 2.5$	A1	
	$G''_Z(1) = \frac{1}{80}(66+168+108+20) = 4.525 = \left[\frac{362}{80} \right]$	M1	
(d)	$\text{Var}(Z) = G''(1) + G'(1) - (G'(1))^2$	M1	Use result.
	$\frac{362}{80} + \frac{200}{80} - \left(\frac{200}{80} \right)^2 = 0.775 = \left[\frac{31}{40} \right]$	A1	
		5	



7 - (9231/41_Winter_2021_Q5) - Probability Generating Function

(a)	$P(3, 6, 9) = \frac{3}{9} \times \frac{2}{8} \times \frac{1}{7} = \frac{6}{504}$	B1	At least 2 probabilities correct.
	$P(\text{Two of } 3, 6, 9) = \frac{3}{9} \times \frac{2}{8} \times \frac{6}{7} \times 3 = \frac{108}{504}$		
	$P(\text{one of } 3, 6, 9) = \frac{3}{9} \times \frac{6}{8} \times \frac{5}{7} \times 3 = \frac{270}{504}$		
	$P(\text{none of } 3, 6, 9) = \frac{6}{9} \times \frac{5}{8} \times \frac{4}{7} = \frac{120}{504}$		
	$G_x(t) = \frac{20}{84} + \frac{45}{84}t + \frac{18}{84}t^2 + \frac{1}{84}t^3$	M1 A1	Attempt with at least 3 probabilities in a polynomial, CAO.
		3	
(b)	$P(\text{both even}) = \frac{12}{72}$	M1	
	$P(\text{one even}) = \frac{40}{72}$		
	$P(\text{no even}) = \frac{20}{72}$		
	$G_y(t) = \frac{5}{18} + \frac{10}{18}t + \frac{3}{18}t^2$	A1	CAO
		2	
(c)	$\frac{1}{1512} \left(\frac{20}{84} + \frac{45}{84}t + \frac{18}{84}t^2 + \frac{1}{84}t^3 \right) \left(\frac{5}{18} + \frac{10}{18}t + \frac{3}{18}t^2 \right)$	M1	Method and attempt to multiply.
		M1	Multiplication to obtain single quintic polynomial.
	$\frac{1}{1512} (100 + 425t + 600t^2 + 320t^3 + 64t^4 + 3t^5)$	A1	CAO
		3	
(d)	$G'(1) = \frac{1}{1512} (425 + 1200 + 960 + 256 + 15)$	M1	
	$\frac{17}{9}$	A1	CWO
		2	



8 - (9231/42_Winter_2021_Q5) - Probability Generating Function

(a)	$(1+4+9+16)k = 1$ so $k = \frac{1}{30}$	B1	
		1	
(b)	$G_X(t) = \frac{1}{30}t + \frac{4}{30}t^2 + \frac{9}{30}t^3 + \frac{16}{30}t^4$	M1A1	Using their k in a polynomial, at least two terms correct for their k .
		2	
(c)	$\left(\frac{1}{30}t + \frac{4}{30}t^2 + \frac{9}{30}t^3 + \frac{16}{30}t^4\right)\left(\frac{1}{4} + \frac{1}{2}t + \frac{1}{4}t^2\right)$	M1	Method and attempt to multiply.
	$\frac{1}{120}(t + 6t^2 + 18t^3 + 38t^4 + 41t^5 + 16t^6)$	M1A1	Multiplication to obtain single polynomial of order 6.
		3	
(d)	Given: $G'(1) = 13/3$ $G''(t) = \frac{1}{120}(12 + 108t + 456t^2 + 820t^3 + 480t^4)$	M1	Differentiate twice.
	$\text{Var}(X) = G''(1) + \frac{13}{3} - \left(\frac{13}{3}\right)^2$	M1	Use correct formula.
	$\frac{1876}{120} + \frac{13}{3} - \frac{169}{9} = \frac{107}{90}$ or 1.19	A1	CAO
		3	

9 - (9231/41_Summer_2022_Q2) - Probability Generating Function

(a)	$G_Y(t) = 0.2 + 0.74t + 0.9t^2 + 0.36t^3$ $G_Y'(t) = 0.74 + 1.8t + 1.08t^2$	M1	PGF differentiated twice.
	$G_Y'(1) = 2.2$ $G_Y''(1) = 3.62$	A1	
	$\text{Var}(Y) = G_Y''(1) + G_Y'(1) - (G_Y'(1))^2$ $3.62 + 2.2 - 2.2^2$	M1	Correct formula used and attempt at $\text{Var}(Y)$.
	0.98	A1	
		4	
(b)	$G_X(t) = (a + bt + ct^2)^2$	M1*	Correct method.
	So $a^2 = 0.04$, $a = 0.2$ $c^2 = 0.09$, $c = 0.3$ $2ab = 0.2$, $b = \frac{0.1}{0.2} = 0.5$	depM1	Attempt to find all coefficients.
	$G_X(t) = 0.2 + 0.5t + 0.3t^2$	A1	Note: $G_Y = \frac{1}{100}(1+t)^2(2+3t)^2$ M1, accept without $\frac{1}{100}$. So $G_X = \frac{1}{10}(1+t)(2+3t)$ dep M1, must have $\frac{1}{10}$. Final answer A1.
		3	

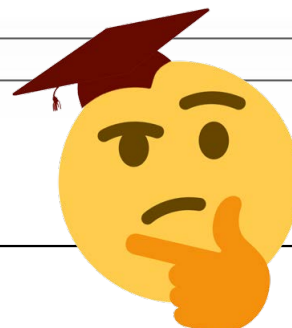


10 - (9231/43_Summer_2022_Q2) - Probability Generating Function

H_0 : the size of shell is independent of the location H_1 : the size of shell is not independent of the location					B1											
<table><tr><td>O</td><td>68</td><td>69</td><td>96</td><td>81</td></tr><tr><td>E</td><td>60.29</td><td>77.87</td><td>100.48</td><td>75.36</td></tr></table>					O	68	69	96	81	E	60.29	77.87	100.48	75.36	M1 A1	Attempt at calculation of E values.
O	68	69	96	81												
E	60.29	77.87	100.48	75.36												
<table><tr><td>O</td><td>28</td><td>55</td><td>64</td><td>39</td></tr><tr><td>E</td><td>35.71</td><td>46.13</td><td>59.52</td><td>44.64</td></tr></table>					O	28	55	64	39	E	35.71	46.13	59.52	44.64		
O	28	55	64	39												
E	35.71	46.13	59.52	44.64												
Calculate values of $\frac{(O-E)^2}{E}$: 0.9859 1.0104 0.1997 0.4221 1.6646 1.7055 0.3372 0.7126					M1	Attempt at Chi-squared contributions. Accuracy of ± 0.1 .										
Test statistic = 7.04					A1	Correct total to 3sf.										
Tabular value for 3df = 6.251 7.04 > 6.251 Reject H_0 .					M1	Compare with 6.251 and correct FT conclusion. May be implied by FT correct conclusion in words.										
Insufficient evidence to suggest size of shell is independent of location/Evidence suggests that size of shell is dependent on (beach) location					A1	Correct conclusion, in context. Level of uncertainty in language is used.										
					7											

11 - (9231/43_Summer_2022_Q3) - Probability Generating Function

(a)		$G_X(t) = \frac{6}{12} + \frac{5}{12}t + \frac{1}{12}t^2$ $P(0 \text{ heads}) = (1-a)(1-b) = \frac{6}{12}$ $P(1 \text{ head}) = a(1-b) + (1-a)b = \frac{5}{12}$ $P(2 \text{ heads}) = ab = \frac{1}{12}$ $\left(a+b = \frac{7}{12}, ab = \frac{1}{12}\right)$	M1	Obtain two correct equations and attempt to solve.
Solve to give $a = \frac{1}{3} \quad \left(b = \frac{1}{4}\right)$			A1	Correct value for a .
			2	
(b)		$G_Y(t) = \left(\frac{6}{12} + \frac{5}{12}t + \frac{1}{12}t^2\right)^2$ $\frac{1}{144}(36 + 60t + 37t^2 + 10t^3 + t^4)$	M1	Square their $G_X(t)$.
			M1	Obtain quartic polynomial from 3-term $G_X(t)$.
			A1	
			3	
(c)		$G_Y'(t) = \frac{1}{144}(60 + 74t + 30t^2 + 4t^3)$ $G_Y''(t) = \frac{1}{144}(74 + 60t + 12t^2)$	M1	Differentiate their $G_Y(t)$ twice.
$\text{Var}(Y) = G_Y''(1) + G_Y'(1) - (G_Y'(1))^2$ $\frac{1}{144}(146) + \frac{7}{6} - \frac{49}{36} = \frac{59}{72} = 0.819$			M1	Use correct formula and attempt at $\text{Var}(Y)$.
0.819			A1	
			3	



12 - (9231/41_Summer_2022_Q4) - Probability Generating Function

(a)	$p = 16.03$	B1																									
	$q = 1.70$	B1	Condone 1.7																								
		2																									
(b)	<table><tr><td>0</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>≥ 7</td></tr><tr><td>12</td><td>20</td><td>36</td><td>32</td><td>13</td><td>6</td><td>1</td><td>0</td></tr><tr><td>9.85</td><td>24.63</td><td>30.78</td><td>25.65</td><td>16.03</td><td>8.02</td><td>3.34</td><td>1.70</td></tr></table>	0	1	2	3	4	5	6	≥ 7	12	20	36	32	13	6	1	0	9.85	24.63	30.78	25.65	16.03	8.02	3.34	1.70	M1	Add last 2 columns (or 3 columns: 13.06).
	0	1	2	3	4	5	6	≥ 7																			
	12	20	36	32	13	6	1	0																			
	9.85	24.63	30.78	25.65	16.03	8.02	3.34	1.70																			
	Combine last two columns: O value = 1, E value = 5.04																										
	Calculate values of $\frac{(O-E)^2}{E}$; 0.4693 0.8704 0.8853 1.5720 0.5727 0.5088 3.2384	M1	Calculate values.																								
	Test statistic = 8.12	A1																									
	H_0 : data fits a Poisson distribution with mean 2.5 H_1 : data does not fit a Poisson distribution with mean 2.5	B1	Need data and distribution. e.g Data fits the given distribution. Number of butterflies per square metre fits Po(2.5).																								
	Critical value of chi-squared = 10.64 Compare: '8.12' < 10.64, accept H_0	M1	Compare their test statistic with 10.64. Note: allow 9.236 if 3 columns combined or 12.02 if none combined.																								
	There is sufficient evidence to support the scientist's claim / there is sufficient evidence to suggest that the data fits a Poisson distribution with mean 2.5	A1	Correct conclusion, in context. Level of uncertainty in language is used.																								
	6																										

13 - (9231/41_Winter_2022_Q4) - Probability Generating Function

(a)	$P(0H) = \frac{1}{3} \times \frac{1}{3} \times \frac{1}{3} = \frac{1}{27}$ $P(1H) = \frac{2}{3} \times \frac{1}{3} \times \frac{1}{3} \times 3 = \frac{6}{27}$ $P(2H) = \frac{1}{3} \times \frac{2}{3} \times \frac{2}{3} \times 3 = \frac{12}{27}$ $P(3H) = \frac{2}{3} \times \frac{2}{3} \times \frac{2}{3} = \frac{8}{27}$	M1	At least 2 probabilities correct.
	$G_X(t) = \frac{1}{27} + \frac{6}{27}t + \frac{12}{27}t^2 + \frac{8}{27}t^3$	M1A1	Correct form, fit their probabilities.
	OR: PGF for one coin is $\frac{1}{3} + \frac{2}{3}t$	M1	
	For 3 coins, PGF is $\left(\frac{1}{3} + \frac{2}{3}t\right)^3$	M1A1	
		3	
(b)	$G_Z(t) = \frac{1}{108}(1+2t+t^2)(1+6t+12t^2+8t^3)$	M1	Correct method.
	$\frac{1}{108}(1+8t+25t^2+38t^3+28t^4+8t^5)$	M1	Obtain quintic polynomial.
	$\frac{1}{108} + \frac{2}{27}t + \frac{25}{108}t^2 + \frac{19}{54}t^3 + \frac{7}{27}t^4 + \frac{2}{27}t^5$	A1	
		3	



(c)	$G_Z'(t) = \frac{1}{108}(8 + 50t + 114t^2 + 112t^3 + 40t^4)$ $E(Z) = \frac{1}{108}(8 + 50 + 114 + 112 + 40)$	M1	Differentiate and evaluate at $t = 1$
	$\frac{324}{108} = 3$	A1	
	Alternative method for questions 4(c)		
	$\frac{2}{3} + \frac{2}{3} + \frac{2}{3} + \frac{1}{2} + \frac{1}{2}$	M1	Summing expected values for each coin
	3	A1	
		2	

14 - (9231/42_Winter_2022_Q5) - Probability Generating Function

(a)	$G_X(t) = \frac{9}{16} + \frac{6}{16}t + \frac{1}{16}t^2$	M1 A1	2 probabilities correct, in a polynomial
		2	
(b)	$G_Z(t) = \left(\frac{9}{16} + \frac{6}{16}t + \frac{1}{16}t^2\right)\left(\frac{25}{36} + \frac{10}{36}t + \frac{1}{36}t^2\right)$	M1	Second PGF correct and multiplied by part (a).
	$\frac{1}{576}(225 + 240t + 94t^2 + 16t^3 + t^4)$	M1	Obtains a quartic polynomial.
	Or $\frac{25}{64} + \frac{5}{12}t + \frac{47}{288}t^2 + \frac{1}{36}t^3 + \frac{1}{576}t^4$	A1	Note: $\frac{1}{576}(t+3)^2(t+5)^2$ scores M1M1A0.
		3	
(c)	$G_Z'(t) = \frac{1}{576}(240 + 188t + 48t^2 + 4t^3)$ $G_Z''(t) = \frac{1}{576}(188 + 96t + 12t^2)$	M1	Differentiate twice.
	$\text{Var}(Z) = \frac{1}{576}(188 + 96 + 12) + \frac{5}{6} - \frac{25}{36}$	M1	Use correct formula.
	$\frac{47}{72}$	A1	
	OR: $\text{Var}(Z) = 2 \times \frac{1}{4} \times \frac{3}{4} + 2 \times \frac{1}{6} \times \frac{5}{6}$	M1 M1	One term correct. Two terms present and added.
	$\frac{47}{72}$	A1	
		3	
(d)	1	B1 FT	FT their power with greatest coefficient in part (b).
		1	

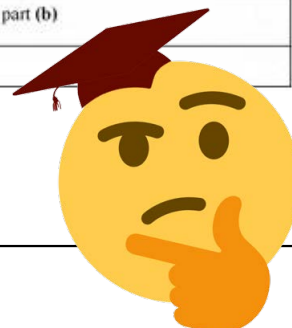


15 - (9231/41_Summer_2023_Q5) - Probability Generating Function

(a)	$P(3H) = \frac{1}{60} \quad P(2H) = \frac{9}{60} \quad P(1H) = \frac{26}{60} \quad P(0H) = \frac{24}{60}$	B1	Two probabilities correct, seen anywhere.
	$G_X(t) = \frac{24}{60} + \frac{26}{60}t + \frac{9}{60}t^2 + \frac{1}{60}t^3$	M1	Cubic polynomial with 4 probabilities as coefficients with the correct powers of t from their working. Equivalent forms are acceptable.
		A1	Correct, AEF.
		3	
(b)	$G_Y(t) = \frac{1}{4} + \frac{1}{2}t + \frac{1}{4}t^2$	B1	
	$G_Z(t) = \left(\frac{24}{60} + \frac{26}{60}t + \frac{9}{60}t^2 + \frac{1}{60}t^3 \right) \left(\frac{1}{4} + \frac{1}{2}t + \frac{1}{4}t^2 \right)$	M1	Attempt to multiply their two PGFs.
	$\frac{1}{240} (24 + 74t + 85t^2 + 45t^3 + 11t^4 + t^5)$	M1	Obtain quintic expression and collect terms
		A1	Correct, AEF.
(c)		4	
	Differentiate: $G'_Z(t) = \frac{1}{240} (74 + 170t + 135t^2 + 44t^3 + 5t^4)$	M1	Differentiate <i>their</i> G .
	$G'_Z(1) = \frac{428}{240} = \frac{107}{60} = 1.78$	A1	Any correct form.
		2	

16 - (9231/43_Summer_2023_Q5) - Probability Generating Function

(a)	$[G_X(t) = k(1 + 3t + 4t^2)] \quad k(1 + 3 + 4) = 1, k = \frac{1}{8}$	B1	Working required.
	$G_X(t) = k(3 + 8t)$	M1	Or $\sum px = 0 \times \frac{1}{8} + 1 \times \frac{3}{8} + 2 \times \frac{4}{8} = \frac{11}{8}$
	$E(X) = G'_X(1) = \frac{11}{8}$	A1	AG Evidence of using $t = 1$ or $\sum px$ required. CWO
		3	
(b)	$G_A(t) = \frac{1}{8}(1 + 3t + 4t^2) \times \frac{1}{3}t^2(1 + 2t)$	M1	Multiply the two PGFs to obtain a single polynomial of degree 5.
	$\frac{1}{24}(t^2 + 5t^3 + 10t^4 + 8t^5)$	A1	May have t^2 as a factor.
		2	
(c)	$G'_A(t) = \frac{1}{24}(2t + 15t^2 + 40t^3 + 40t^4)$ $G''_A(t) = \frac{1}{24}(2 + 30t + 120t^2 + 160t^3)$	M1	Differentiate twice, allow one slip.
	$\text{Var}(X) = G''_A(1) + G'_A(1) - (G'_A(1))^2 = \frac{1}{24}(312) + \frac{97}{24} - \left(\frac{97}{24}\right)^2$	M1	Use correct formula.
	0.707	A1	$\frac{407}{576}$
		3	
(d)	$(Z =) 4$	B1 FT	FT <i>their</i> final polynomial in part (b)
		1	



17 - (9231/41_Winter_2023_Q5) - Probability Generating Function

(a)	$P(X=r) = p(1-p)^{r-1} = pq^{r-1}$	B1	Implied by $G(t) = \sum pq^{r-1} t^r$.
	$G(t) = \sum pq^r (qt)^{r-1}$	M1	(Infinite summation) with common ratio (qt) and first term or common factor p indicated or GP identified.
	$G(t) = pq \left(\frac{1}{1-qt} \right) = \frac{pq}{1-qt}$	A1	AG Convincingly obtained.
		3	
(b)	$G'(t) = pq(1-qt)^{-2} + p(1-qt)^{-1} [= p(1-qt)^{-2}]$	M1	Attempt to differentiate as a product/quotient.
	$G''(t) = 2pq^2t(1-qt)^{-3} + 2pq(1-qt)^{-2} [= 2qp(1-qt)^{-3}]$	M1	Attempt to differentiate <i>their</i> $G'(t)$.
	$G'(1) = p(1-q)^{-2} = \frac{1}{p}, \quad G''(1) = \frac{2q}{p^2}$	M1	Substitute $t=1$ into their $G'(t)$ and $G''(t)$ and use formula for $\text{Var}(X)$ (at some point).
	$\text{Var}(X) = \frac{2q}{p^2} + \frac{1}{p} - \left(\frac{1}{p}\right)^2$	M1	$(1-q)$ replaced by p in denominator throughout, and some cancellation seen (at some point).
	$\frac{q}{p^2}$	A1	AG Convincingly obtained.
		5	
(c)	$G_X(t) = \frac{\frac{1}{6}t}{1-\frac{5}{6}t} \left[= \frac{t}{6-5t} \right]$	B1	Implied by correct final answer.
	$G_X(t) = (G_X(t))^2 = \left(\frac{t}{6-5t} \right)^2$	B1	Allow $\left(\frac{\frac{1}{6}t}{1-\frac{5}{6}t} \right)^2$.
		2	

18 - (9231/42_Winter_2023_Q3) - Probability Generating Function

(a)	$P(3R) = \frac{120}{504} \quad P(2R) = \frac{270}{504} \quad P(1R) = \frac{108}{504} \quad P(0R) = \frac{6}{504}$	B1	2 correct probabilities.
	$G_X(t) = \frac{1}{504} (6 + 108t + 270t^2 + 120t^3) = \frac{1}{84} (1 + 18t + 45t^2 + 20t^3)$	M1	Cubic polynomial with their probabilities.
	$\frac{1}{84} + \frac{3}{14}t + \frac{15}{28}t^2 + \frac{5}{21}t^3$	A1	Correct.
		3	
(b)	$G_X(t) = \frac{1}{84} (1 + 18t + 45t^2 + 20t^3) \times \frac{1}{12} (1 + 6t + 5t^2)$	M1	Attempt to multiply out the brackets.
	$\frac{1}{1008} (1 + 24t + 158t^2 + 380t^3 + 345t^4 + 100t^5)$	M1	Obtain quintic polynomial (may not be simplified).
	$\frac{1}{1008} + \frac{1}{42}t + \frac{79}{504}t^2 + \frac{95}{252}t^3 + \frac{115}{336}t^4 + \frac{25}{252}t^5$	A1	Correct.
		3	



(c)	$G' = \frac{1}{1008} (24 + 316t + 1140t^2 + 1380t^3 + 500t^4)$ $G'' = \frac{1}{1008} (316 + 2280t + 4140t^2 + 2000t^3)$	M1	Differentiate twice. May not see derivatives in terms of t .
	$\left[E(Z) = \frac{3360}{1008} \right] \frac{10}{3}$	B1	Seen or implied, NFWW (e.g. sampling with replacement).
	$\text{Var}(Z) = \frac{8736}{1008} + \frac{10}{3} - \left(\frac{10}{3} \right)^2 = \frac{26}{3} - \frac{10}{3} - \frac{100}{9}$	M1	Use formula with their $G''(1)$ and $G'(1)$.
	$\frac{8}{9}$	A1	Or 0.889
		4	

