Topic	Question	Done?
Binomial	2	
Binomial	3	
Binomial	4	
Binomial	11	
Data analysis	12	
Data analysis	5	
Data analysis	4	
Data analysis	6	
Normal approximation of binomial	1	
Normal approximation of binomial	6	
Normal approximation of binomial	10	
Binomial hypothesis testing	3	
Binomial hypothesis testing	8	
Binomial hypothesis testing	2	
Binomial hypothesis testing	3	
Hypothesis testing of a mean	1	
Hypothesis testing of a mean	2	
Hypothesis testing of a mean	1	
Hypothesis testing of a mean	1	
Normal distribution	23	
Normal distribution	18	
Normal distribution	11	
Normal distribution	29	
Normal distribution	4	
Probability	13	
Probability	21	
Probability	47	
Probability	50	
Probability	7	
Hypothesis testing correlation	6	
Hypothesis testing correlation	7	
Exponential models	5	
Exponential models	6	

Binomial

Question 2 (***)

The random variable X has the binomial distribution $\mathrm{B}(13, 0.16)$.

a) Find $\mathrm{P}(X < 2)$.

Two independent observations of X are made.

b) Determine the probability that exactly one of these values is equal to 2

$\boxed{0.3604}$, $\boxed{0.4146}$

(a) $X \sim B(13, 0.16)$

$P(X<2) = P(X \le 1) = P(X=0,1) = \binom{13}{0}(0.16)^0(0.84)^{13} + \binom{13}{1}(0.16)^1(0.84)^{12}$

$= 0.10366 + 0.25669 = 0.3604$

(b) $P(X=2) = \binom{13}{2}(0.16)^2(0.84)^{11} = 0.29336$

$P(X \neq 2) = 1 - 0.29336 = 0.70664$

$\therefore$ REQUIRED PROBABILITY IS "2" $\times$ "NOT 2" $\times$ (2 WAYS)

$= 0.29336 \times 0.70664 \times 2$

$= 0.4146$

Question 3 (***)

The discrete random variable X has the distribution $\mathrm{B}(15, p)$.

a) Given that $p = 0.3$, determine ...

i. ... $\mathrm{P}(X < 6)$.

ii. ... $\mathrm{P}(X > 10)$.

b) Given instead that $\mathrm{P}(X = 0) = 0.04$, determine the value of p, correct to three decimal places.

c) Given instead that $\mathrm{Var}(X) = 3.15$, determine the two possible values of p.

$\boxed{0.7216}$, $\boxed{0.0007}$, $\boxed{p \approx 0.193}$, $\boxed{p = 0.3 \text{ or } 0.7}$

(a) $X \sim B(15, p)$, $p = 0.3$

(I) $P(X<6) = P(X \le 5) = \ldots$ tables $\ldots = 0.7216$

(II) $P(X>10) = P(X \ge 11) = 1 - P(X \le 10) = \ldots$ tables

$= 1 - 0.9993 = 0.0007$

(b) $X \sim B(15, p)$

$\Rightarrow P(X=0) = 0.04$

$\Rightarrow \binom{15}{0} p^0 (1-p)^{15} = 0.04$

$\Rightarrow 1 \times 1 \times (1-p)^{15} = 0.04$

$\Rightarrow (1-p)^{15} = 0.04$

$\Rightarrow 1-p = \sqrt[15]{0.04}$

$\Rightarrow 1-p = 0.80687\ldots$

$\Rightarrow p \approx 0.193$ 3 d.p.

(c) $X \sim B(15, p)$

$\text{Var}(X) = np(1-p)$

$\Rightarrow 3.15 = 15p(1-p)$

$\Rightarrow 3.15 = 15p - 15p^2$

$\Rightarrow 15p^2 - 15p + 3.15 = 0$

$\Rightarrow p^2 - p + 0.21 = 0$

$p = \frac{1 \pm \sqrt{1 - 4 \times 1 \times 0.21}}{2 \times 1}$

$p = \frac{1 \pm 0.4}{2}$

$p = \begin{cases} 0.7 \\ 0.3 \end{cases}$

Question 4 (***)

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

Given that the mean and the standard deviation of X are both 0.95, determine the value of n.

$\boxed{n = 19}$

$X \sim B(n, p)$

- $E(X) = np$

$0.95 = np$

- $\text{Var}(X) = np(1-p)$

(S.D) $\longrightarrow 0.95^2 = np(1-p)$

$0.95^2 = 0.95(1-p)$

$0.95 = 1 - p$

$\boxed{p = 0.05}$

$\therefore np = 0.95$

$0.05n = 0.95$

$n = 19$

Question 11 (****)

The records in a doctor's surgery show that 20% of the patients that make an appointment fail to turn up.

During a given weekday there are 20 appointments to see the doctor and the doctor has enough time to see all 20 patients.

a) Find the probability that all the patients **will** turn up.

b) Find the probability that more than three patients **will not** turn up.

In order to improve efficiency in the surgery, the doctor decides to make more than 20 appointments although he still has enough time to see only 20 patients.

One day the doctor has booked 21 appointments.

c) Find the probability he will be able to see all the patients that **will** turn up.

Another day the doctor has booked 25 appointments.

d) Find the probability that the doctor **will not** be able to see one of the patients that will turn up.

$\boxed{0.0115}$, $\boxed{0.5886}$, $\boxed{0.9908}$, $\boxed{0.1867}$

X = no of "no shows"

$X \sim B(20, 0.2)$

(a) $P(X=0) = \binom{20}{0}(0.2)^0(0.8)^{20} = 0.0115$

(b) $P(X>3) = P(X \geqslant 4) = 1 - P(X \leqslant 3) = \ldots$ tables...

$= 1 - 0.4114 = 0.5886$

- $X \sim B(21, 0.2)$

$P(X \geqslant 1) = 1 - P(X=0) = 1 - \binom{21}{0}(0.2)^0(0.8)^{21}$

$= 1 - 0.0092$

$= 0.9908$

- $X \sim B(25, 0.2)$

"WILL NOT BE ABLE TO SEE ONE" $\Rightarrow$ 21 TURNED UP

$\Rightarrow$ 4 NO SHOWS

$P(X=4) = \binom{25}{4}(0.2)^4(0.8)^{21} = 0.1867$

Data Analysis

Question 12

The number of bottles of red wine sold by a local supermarket over a two week period is shown below.

22, 14, 11, 33, 32, 45, 4, 12, 13, 20, 27, 44, 30, 15.

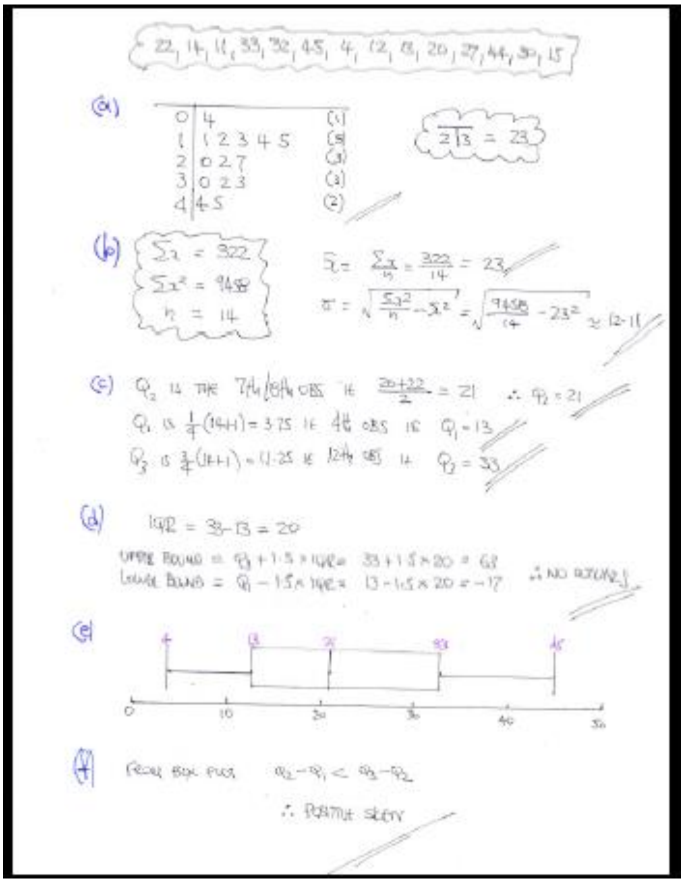
a) Display the data in a stem and leaf diagram.

b) Find the mean and the standard deviation of the data.

c) Find the lower quartile, the median and the upper quartile of the data.

d) Determine if there are any outliers.

e) Draw a suitably labelled box plot for this data.



22, 14, 11, 33, 32, 45, 4, 12, 13, 20, 27, 44, 30, 15

(a)

0	4	(1)
1	1 2 3 4 5	(5)
2	0 2 7	(3)
3	0 2 3	(3)
4	4 5	(2)

2|3 = 23

(b) $\Sigma x = 322$, $\Sigma x^2 = 9458$, $n = 14$

$\bar{x} = \frac{\Sigma x}{n} = \frac{322}{14} = 23$

$\sigma = \sqrt{\frac{\Sigma x^2}{n} - \bar{x}^2} = \sqrt{\frac{9458}{14} - 23^2} \simeq 12.11$

(c) Q_2 IS THE 7th/8th OBS IE $\frac{20+22}{2} = 21$ $\therefore Q_2 = 21$

Q_1 IS $\frac{1}{4}(14+1) = 3.75$ IE 4th OBS IE $Q_1 = 13$

Q_3 IS $\frac{3}{4}(14+1) = 11.25$ IE 12th OBS IE $Q_3 = 33$

(d) IQR = 33 − 13 = 20

UPPER BOUND = $Q_3 + 1.5 \times IQR = 33 + 1.5 \times 20 = 63$

LOWER BOUND = $Q_1 - 1.5 \times IQR = 13 - 1.5 \times 20 = -17$

$\therefore$ NO OUTLIERS

(e)

(f) FROM BOX PLOT $Q_2 - Q_1 < Q_3 - Q_2$

$\therefore$ POSITIVE SKEW

Question 5

The commuting distances of 125 individuals, rounded to the nearest mile, is summarised in the table below.

Distance (nearest mile)	Frequency
0 – 9	12
10 –19	22
20 – 29	48
30 – 39	26
40 – 49	8
50 – 59	5
60 – 69	3
70 – 79	1

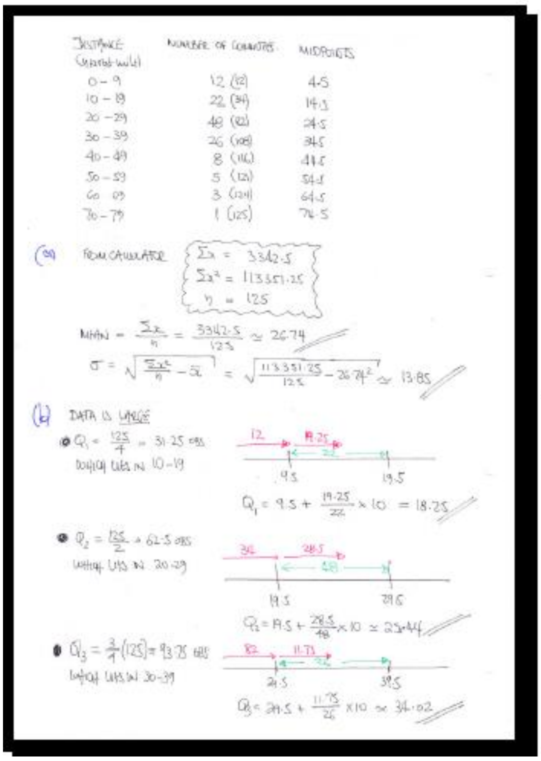
a) Estimate the mean and the standard deviation of these commuting distances.

b) Use linear interpolation to estimate the value of the median, the lower quartile and the upper quartile.

c) Determine with justification the skewness of the data.

d) Explain which out of the mean and standard deviation or the median and the interquartile range are more appropriate measures to summarize this data.

$\boxed{\bar{x} \approx 26.74}$, $\boxed{\sigma \approx 13.85}$, $\boxed{Q_1 = 18.1 - 18.4}$, $\boxed{Q_2 = 25.3 - 25.5}$, $\boxed{Q_3 = 33.9 - 34.1}$, $\boxed{\text{positive skew}}$, $\boxed{\text{median \& IQR}}$



DISTANCE (nearest mile)	NUMBER OF COUNTRIES	MIDPOINTS
0 – 9	12 (12)	4.5
10 – 19	22 (34)	14.5
20 – 29	48 (82)	24.5
30 – 39	26 (108)	34.5
40 – 49	8 (116)	44.5
50 – 59	5 (121)	54.5
60 – 69	3 (124)	64.5
70 – 79	1 (125)	74.5

(a) FROM CALCULATOR $\Sigma x = 3342.5$, $\Sigma x^2 = 113351.25$, $n = 125$

$$\text{MEAN} = \frac{\Sigma x}{n} = \frac{3342.5}{125} \simeq 26.74$$

$$\sigma = \sqrt{\frac{\Sigma x^2}{n} - \bar{x}^2} = \sqrt{\frac{113351.25}{125} - 26.74^2} \simeq 13.85$$

(b) DATA IS LARGE

- $Q_1 = \frac{125}{4} = 31.25$ OBS, WHICH LIES IN 10–19

12, 19.25, 22, 9.5, 19.5

$$Q_1 = 9.5 + \frac{19.25}{22} \times 10 = 18.25$$

- $Q_2 = \frac{125}{2} = 62.5$ OBS, WHICH LIES IN 20–29

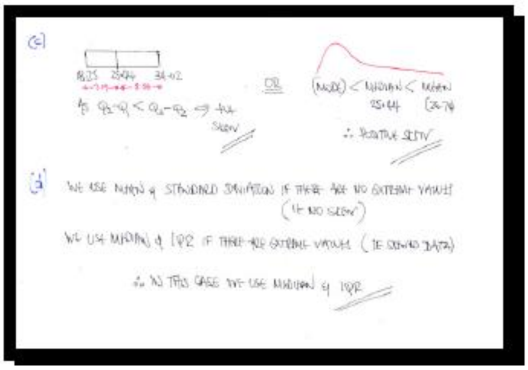
34, 28.5, 48, 19.5, 29.5

$$Q_2 = 19.5 + \frac{28.5}{48} \times 10 \simeq 25.44$$

- $Q_3 = \frac{3}{4}(125) = 93.75$ OBS, WHICH LIES IN 30–39

82, 11.75, 26, 29.5, 39.5

$$Q_3 = 29.5 + \frac{11.75}{26} \times 10 \simeq 34.02$$



(c) 18.25, 25.44, 34.02

AS $Q_2 - Q_1 < Q_3 - Q_2 \Rightarrow$ +VE SKEW

OR

(MODE) < MEDIAN < MEAN
25.44 (26.74)

∴ POSITIVE SKEW

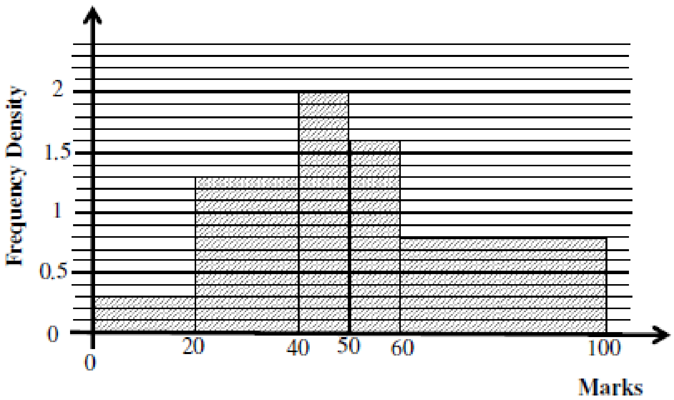
(d) WE USE MEAN & STANDARD DEVIATION IF THERE ARE NO EXTREME VALUES (IE NO SKEW)

WE USE MEDIAN & IQR IF THERE ARE EXTREME VALUES (IE SKEWED DATA)

∴ IN THIS CASE WE USE MEDIAN & IQR

Question 4

The histogram below shows the distribution of the marks of 250 students.

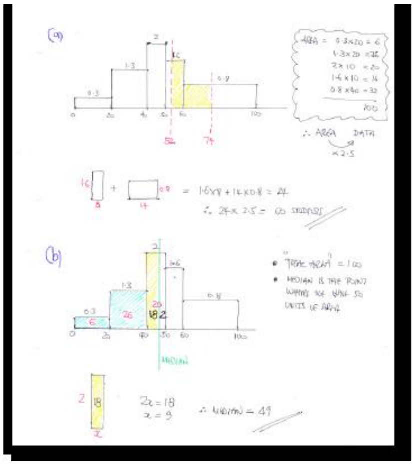


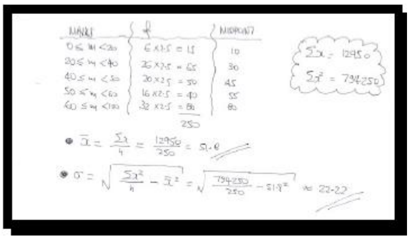
a) Estimate how many students scored between 52 and 74 marks.

b) Use the histogram estimate the median.

c) Find an estimate of the mean and standard deviation of the marks of these students.

$\boxed{60}$, $\boxed{49}$, $\boxed{\bar{x} \simeq 51.8}$, $\boxed{\sigma \simeq 22.22}$





Question 6

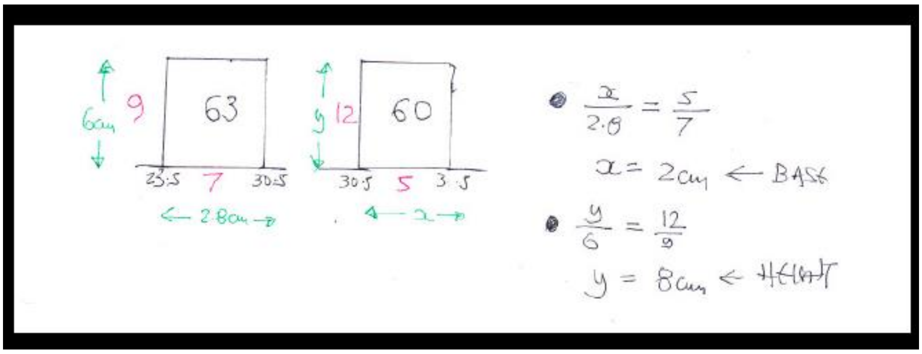
In a histogram the weights of hamsters, correct to the nearest gram, are plotted on the x axis.

In this histogram the class $24-30$ has a frequency of 63 and is represented by a rectangle of base 2.8 cm and height 6 cm.

In the same histogram the class $31-35$ has a frequency of 60.

Determine the measurements, in cm, of the rectangle that represents the class $31-35$.

$\boxed{\text{base} = 2 \text{ cm}}$, $\boxed{\text{height} = 8 \text{ cm}}$



Normal approximations of the binomial distribution

Question 1

The discrete random variable X has probability distribution

$$X \sim B(160, 0.125).$$

Use a distributional approximation, to find $P(18 \le X < 25)$.

$$\boxed{P(18 \le X < 25) = 0.5840}$$

Question 6

A popular bag of confectionary contains 20 sweets, of which $\frac{1}{5}$ are expected to be orange in flavour.

a) Find the probability that once such bag selected at random will contain at least 3 but no more than 7 orange flavoured sweets.

A family size bag of the same confectionary contains 90 sweets. The proportion of the orange flavoured sweets in these bags is also expected to be $\frac{1}{5}$.

b) Use a distributional approximation, to find the probability that a randomly selected family size bag, will contain less than 25 orange flavoured sweets.

$\boxed{0.7818}$, $\boxed{0.9567}$

(a) X = orange sweet

$X \sim B(20, 0.2)$

$P(3 \le X \le 7) = P(X \le 7) - P(X \le 2) = 0.9679 - 0.2061 = 0.7818$

(b) $X \sim B(90, 0.2)$

$\left.\begin{array}{l} \text{MEAN} = np = 90 \times 0.2 = 18 > 5 \\ \text{VARIANCE} = np(1-p) = 18 \times 0.8 > 5 \end{array}\right\}$ APPROXIMATE BY NORMAL

$Y \sim N(18, 14.4)$

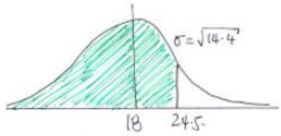
$P(X < 25) = P(X \le 24)$

$= P(Y < 24.5)$

$= P\left(Z < \frac{24.5 - 18}{\sqrt{14.4}}\right)$

$= \Phi(1.7129\ldots)$

$= 0.9567$ INTERPOLATED FIGURE



Question 10

The discrete random variable $X \sim \mathrm{B}(n, p)$.

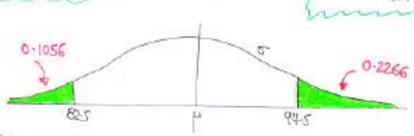
The value of n and the value of p are such so that X can be approximated by a Normal distribution.

Using a Normal approximation the probability that X is at most 82 is 0.1056 and the probability that X is less than 95 is 0.7734.

Determine the value of n and the value of p.

$\boxed{n = 150}$, $\boxed{p = 0.6}$

$X \sim B(n, p)$ APPROXIMATED BY $Y \sim N(np, np(1-p))$



- $P(X \le 82) = 0.1056$

$P(Y < 82.5) = 0.1056$

$P(Y > 82.5) = 0.8944$

$P\left(Z > \frac{82.5 - \mu}{\sigma}\right) = 0.8944$

$\frac{82.5 - \mu}{\sigma} = \Phi^{-1}(0.8944)$

$\frac{82.5 - \mu}{\sigma} = -1.25$

$82.5 - \mu = -1.25\sigma$

$\boxed{82.5 + 1.25\sigma = \mu}$

- $P(X < 95) = 0.7734$

$P(X \le 94) = 0.7734$

$P(Y < 94.5) = 0.7734$

$P\left(Z < \frac{94.5 - \mu}{\sigma}\right) = 0.7734$

$\frac{94.5 - \mu}{\sigma} = \Phi^{-1}(0.7734)$

$\frac{94.5 - \mu}{\sigma} = 0.75$

$94.5 - \mu = 0.75\sigma$

$\boxed{94.5 - 0.75\sigma = \mu}$

$82.5 + 1.25\sigma = 94.5 - 0.75\sigma$

$2\sigma = 12$

$\boxed{\sigma = 6}$ & $\boxed{\mu = 90}$

Thus $\left.\begin{array}{l} np = 90 \\ \sqrt{np(1-p)} = 6 \end{array}\right\} \Rightarrow \left.\begin{array}{l} np = 90 \\ np(1-p) = 36 \end{array}\right\} \Rightarrow 90(1-p) = 36$

$1 - p = 0.4$

$p = 0.6$

Binomial Hypothesis testing

Question 3

The owner of a corner shop believes that 25% of the customers who buy crisps will buy the "cheese and onion" variety.

He finds that in the last 30 customers who bought crisps, only 4 customers bought the "cheese and onion" variety.

Test, at the 5% level of significance, whether there is evidence to suggest that the proportion of customers who choose the "cheese and onion" variety is lower than 25%.

$\boxed{\text{not significant evidence, } 9.79\% > 5\%}$

X = a "cheese & onion" customer
$X \sim B(30, 0.25)$

$H_0: p = 0.25$
$H_1: p < 0.25$
where p is the proportion of people who buy cheese & onion crisps in general

Testing at the 5% significance on the basis that $x = 4$

$P(X \leq 4) = 0.0979$
$= 9.79\%$
$> 5\%$

NOT SIGNIFICANT EVIDENCE THAT FEWER THAN 25% CUSTOMERS BUY CHEESE & ONION CRISPS
IE REJECT H_1, ACCEPT H_0

Question 8

A pub manager feels that since the introduction of the "smoking ban" in his pub, the proportion of the non smoking customers visiting his pub has increased.

It had been established that before the "smoking ban" 15% of the customers visiting his pub were non smokers.

He now finds 8 non smokers in a random sample of 20 customers.

Test, at the 1% level of significance, whether there is evidence to suggest that the proportion of the non smoking customers visiting his pub has increased.

$\boxed{\text{significant evidence, } 0.59\% < 1\%}$

$X =$ no of non smokers
$X \sim B(20, 0.15)$

$H_0 : p = 0.15$
$H_1 : p > 0.15$

WHERE p IS THE PROPORTION OF NON SMOKING CUSTOMERS IN GENERAL

TESTING AT 1% SIGNIFICANCE ON THE BASIS AT $x = 8$

$$P(X \geq 8) = 1 - P(X \leq 7) = 1 - 0.9941 = 0.0059 = 0.59\% < 1\%$$

SIGNIFICANT EVIDENCE THAT THE PROPORTION OF NON SMOKERS VISITING THE PUB HAS INCREASED, IE REJECT H_0 & ACCEPT H_1

Question 2

The proportion of tiles with minor faults produced in a factory is thought to be 10% .

The factory manager believes that the proportion is higher due to the old machinery.

He inspects a random sample of 20 tiles.

a) Find the critical region to test the manager's belief, at the 5% level of significance.

b) State the actual significance level for a test using the critical region of part **(a)**.

Four faulty tiles were found in the sample.

c) Complete the test.

$\boxed{\text{C.R.} = \{5,6,7,...,20\}}$, $\boxed{4.32\%}$, $\boxed{\text{not significant evidence, } 4 \text{ not in C.R.}}$

(a) $X =$ faulty tile
$X \sim B(20, 0.1)$

$H_0 : p = 0.1$
$H_1 : p > 0.1$

$p =$ is the proportion of faulty tiles in the entire production

CRITICAL REGION REQUIRED AT 5% SIGNIFICANCE

- $P(X \geq 4) = 1 - P(X \leq 3) = 1 - 0.8670 = 0.1330 = 13.3\% > 5\%$ NOT SIGNIFICANT
- $P(X \geq 5) = 1 - P(X \leq 4) = 1 - 0.9568 = 0.0432 = 4.32\% < 5\%$ SIGNIFICANT

$\therefore$ CRITICAL REGION $5 \leq x \leq 20$

(b) ACTUAL SIGNIFICANCE IS 4.32%

(c) 4 IS NOT IN THE CRITICAL REGION; NOT SIGNIFICANT EVIDENCE TO SUGGEST THAT MACHINERY PRODUCES MORE FAULTY TILES (ACCEPT H_0, REJECT H_1)

Question 3

The recruitment director of a large accounting firm believes that maths graduates are more successful when applying for positions in his firm compared with graduates of other subjects.

One in five job applicants to this firm is successful.

The recruitment director selects a random sample of 25 maths graduate applicants.

a) Find the critical region to test at the 5% level of significance the director's belief.

b) State the actual significance level for a test using the critical region of part (a).

Ten successful maths graduate applicants were found in the sample.

c) Complete the test.

$\boxed{\text{C.R.} = \{9, 10, 11, \ldots, 25\}}$, $\boxed{4.68\%}$, $\boxed{\text{significant evidence, 10 is in C.R.}}$

(a) X = SUCCESSFUL "MATHS" APPLICANT

$X \sim B(25, 0.2)$

$H_0 : p = 0.2$

$H_1 : p > 0.2$

p is proportion of successful "maths" applicants in general

CRITICAL REGION REQUIRED AT 5% LEVEL OF SIGNIFICANCE

- $P(X \geq 8) = 1 - P(X \leq 7) = 1 - 0.8909$
  $= 0.1091 = 10.91\% > 5\%$
  NOT SIGNIFICANT
- $P(X \geq 9) = 1 - P(X \leq 8) = 1 - 0.9532$
  $= 0.0468 = 4.68\% < 5\%$
  SIGNIFICANT

C.R $\Rightarrow 9 \leq x \leq 25$

(b) ACTUAL SIGNIFICANCE IS 4.68%

(c) 10 IS IN THE CRITICAL REGION. THERE IS SIGNIFICANT EVIDENCE TO SUPPORT THE DIRECTOR'S BELIEF (ACCEPT H_1, REJECT H_0)

Hypothesis testing of the mean of a normal distribution

Question 1

The weights of Madeira cakes are Normally distributed with a mean of 450 grams and a standard deviation of 25.

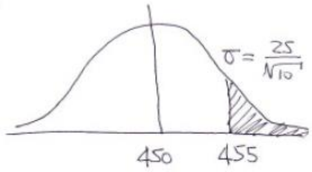
Find the probability that the sample mean of 10 randomly selected Madeira cakes will exceed 455 grams.

$\boxed{0.264}$

$X =$ weight of Madeira cake

$X \sim N(450, 25^2)$

$\bar{X} \sim N\left(450, \left(\frac{25}{\sqrt{10}}\right)^2\right)$



$$
\begin{aligned}
P(\bar{X} > 455) &= 1 - P(\bar{X} < 455) \\
&= 1 - P\left(Z < \frac{455 - 450}{25/\sqrt{10}}\right) \\
&= 1 - \Phi(0.6325) \\
&= 1 - 0.7357 \\
&= 0.264
\end{aligned}
$$

Question 2

David's commuting times are Normally distributed with a mean of 42 minutes and a standard deviation of 15.

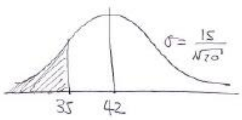
Find the probability that 20 of David's commuting times will have a sample mean of less than 35 minutes.

$\boxed{0.019}$

$T =$ David's commuting time

$T \sim N(42, 15^2)$

$\bar{T} \sim N\left(42, \left(\frac{15}{\sqrt{20}}\right)^2\right)$



$$
\begin{aligned}
P(\bar{T} < 35) &= 1 - P(\bar{T} > 35) \\
&= 1 - P\left(Z > \frac{35 - 42}{15/\sqrt{20}}\right) \\
&= 1 - \Phi(-2.0869\ldots) \\
&= 1 - 0.9812 \\
&= 0.019
\end{aligned}
$$

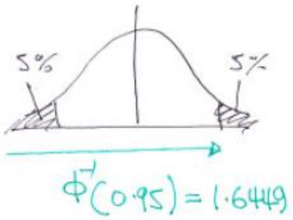
Question 1

The weights of bags of popcorn are Normally distributed with unknown mean μ and standard deviation of 10 grams.

The sample mean $\bar{x}$ of 8 such bags was found to be 130 grams.

Find a 90% confidence interval for the mean weight of a bag of popcorn.

$\boxed{(124.2, 135.8)}$



$\mu = \bar{x} \pm \frac{\sigma}{\sqrt{n}} \Phi^{-1}(0.95)$

$\mu = 130 \pm \frac{10}{\sqrt{8}} \times 1.6449$

$\mu = 130 \pm 5.82$

$CI = (124.2, 135.8)$

Question 1

The continuous random variable X is Normally distributed with mean μ and standard deviation 10.

A test is to be carried out for the hypotheses

$$H_0 : \mu = 40 \text{ versus } H_1 : \mu > 40 .$$

A random sample of 5 observations of X produced a sample mean of 44.

Carry out the test at the 1% significance level.

$\boxed{\text{not significant}}$

$X \sim N(40, 10^2)$

$H_0 : \mu = 40$

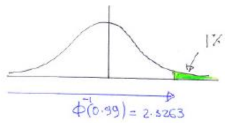
$H_1 : \mu > 40$

$\bar{x} = 44$

$n = 5$

$\sigma = 10$

1% SIGNIFICANCE



- Z-STAT $= \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{44 - 40}{10/\sqrt{5}}$

Z-STAT $= 0.8944\ldots$

- AS $0.8944 < 2.3263$ THERE IS NO SIGNIFICANT EVIDENCE THAT $\mu > 40$. REJECT H_1

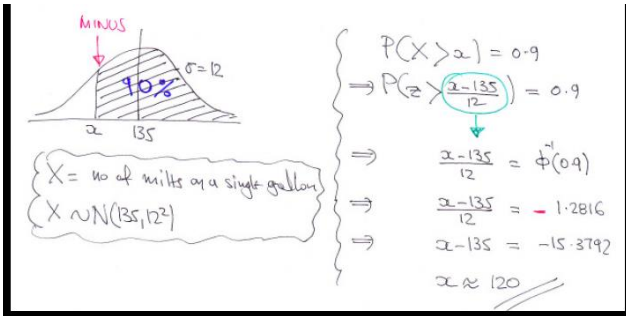
Normal distribution

Question 23

The number of miles Mark's motorbike can travel on a single gallon of petrol can be modelled by a Normal distribution with a mean of 135 miles and a standard deviation of 12.

Find, to the nearest mile, the longest journey that Mark can make, if he is to have at least a 90% chance of completing it on a single gallon of petrol.

$\boxed{120}$



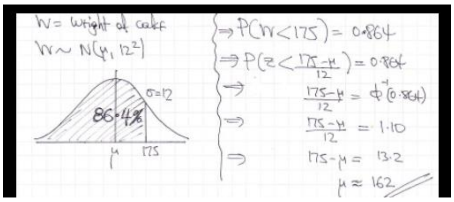
Question 18

The weights of fairy cakes produced by a baker are Normally distributed with a standard deviation of 12 grams.

Find their mean weight if 86.4% of these cakes weigh less than 175 grams.

$\boxed{162}$

#

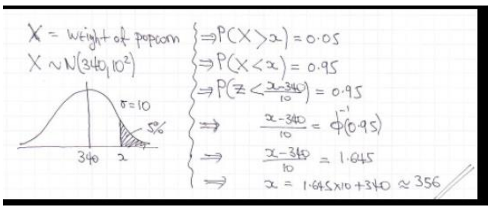


Question 11

The weights of popcorn bags sold in a cinema are assumed to be Normally distributed with a mean of 340 grams and a standard deviation of 10.

Find to the nearest gram, the weight exceeded by 5% of these popcorn bags.

$\boxed{356}$



Question 29

The lengths of pine needles are Normally distributed.

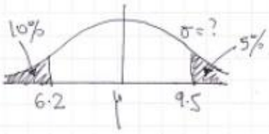
10% of these needles are shorter than 6.2 cm and 5% are longer than 9.5 cm.

Find the mean and the standard deviation of the length of a pine needle.

$$\boxed{\mu = 7.64 - 7.65,\ \ \sigma = 1.13}$$

L = length of needle

$L \sim N(\mu, \sigma^2)$



$\Rightarrow P(L<6.2) = 0.1$

$\Rightarrow P(L>6.2) = 0.9$

$\Rightarrow P\left(Z > \frac{6.2-\mu}{\sigma}\right) = 0.9$

$\Rightarrow \frac{6.2-\mu}{\sigma} = -\Phi^{-1}(0.9)$

$\Rightarrow \frac{6.2-\mu}{\sigma} = -1.28$

$\Rightarrow 6.2-\mu = -1.28\sigma$

$\Rightarrow \boxed{6.2 + 1.28\sigma = \mu}$

$\Rightarrow P(L>9.5) = 0.05$

$\Rightarrow P(L<9.5) = 0.95$

$\Rightarrow P\left(Z < \frac{9.5-\mu}{\sigma}\right) = 0.95$

$\Rightarrow \frac{9.5-\mu}{\sigma} = \Phi^{-1}(0.95)$

$\Rightarrow \frac{9.5-\mu}{\sigma} = 1.645$

$\Rightarrow 9.5-\mu = 1.645\sigma$

$\Rightarrow \boxed{9.5 - 1.645\sigma = \mu}$

HENCE $6.2 + 1.28\sigma = 9.5 - 1.645\sigma$

$2.925\sigma = 3.3$

$\sigma \approx 1.13$

AND $\mu = 6.2 + 1.28\sigma$

$\mu \approx 7.64$

4. A teacher wants to investigate the sports played by students at her school in their free time. She decides to ask a random sample of 120 pupils to complete a short questionnaire.

(a) Give two reasons why the teacher might choose to use a sample survey rather than a census.

(2 marks)

(b) Suggest a suitable sampling frame that she could use. **(1 mark)**

The teacher believes that 1 in 20 of the students play tennis in their free time. She uses the data collected from her sample to test if the proportion is different from this.

(c) Using a suitable approximation and stating the hypotheses that she should use, find the critical region for this test. The probability for each tail of the region should be as close as possible to 5%.

(6 marks)

(d) State the significance level of this test. **(1 mark)**

4.	*(a)*	e.g. quicker; may not be able to get all pupils to respond	B2	
	(b)	school roll	B1	
	(c)	let X = no. of students who play tennis $\therefore X \sim \mathrm{B}(120, \frac{1}{20})$	M1	
		$\mathrm{H}_0 : p = \frac{1}{20}$ $\quad \mathrm{H}_1 : p \neq \frac{1}{20}$	B1	
		Using Po approx. $X \approx\sim \mathrm{Po}(6)$	M1	
		$\mathrm{P}(X \leq 2) = 0.0620$; $\mathrm{P}(X \leq 10) = 0.9574$	M1 A1	
		$\therefore$ C.R. is $X \leq 2$ or $X \geq 11$	A1	
	(d)	$0.0620 + 0.0426 = 0.1046$	A1	**(10)**

Probability

Question 13 (**+)

The quality controller in a factory examined 250 components for three types of minor faults, known as fault A, fault B and fault C. His results are summarized as follows.

- 4 components with type C fault only.
- 8 components with type A and C but no B fault.
- 9 components with type A fault only.
- 7 components with type B and C but no A fault.
- 1 component with type B fault only.
- 5 components with type A and B but no C fault.
- 2 components with all three types of fault.

a) Draw a fully completed Venn diagram to represent this information.

A component is selected at random.

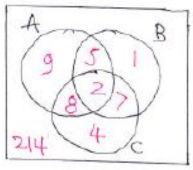
b) Find the probability that the selected component has a type C fault.

c) Given that the selected component has a type B fault, find the probability that the component has all three types of fault.

d) Given that the selected component has a fault, find the probability it has two faults.

$\boxed{0.084}$, $\boxed{\frac{2}{15}}$, $\boxed{\frac{5}{9}}$

(a)



(b) $P(C) = \frac{8+2+7+4}{250} = \frac{21}{250}$

(c) $P(\text{all 3 faults} \mid \text{fault } B) = \frac{2}{5+1+2+7}$

$= \frac{2}{15}$

(d) $P(\text{2 faults} \mid \text{faulty}) = \frac{5+8+7}{36}$

$= \frac{20}{36} = \frac{5}{9}$

Question 21 (***)

The events A and B satisfy

$$P(A) = 0.5 \text{, } P(B) = 0.2 \text{ and } P(A|B) = 0.3 \text{.}$$

Determine …

a) … $P(A \cap B)$.

b) … $P(A \cup B)$.

c) … $P(B|A)$.

$\boxed{P(A \cap B) = 0.06}$, $\boxed{P(A \cup B) = 0.64}$, $\boxed{P(B|A) = 0.12}$

(a) $P(A) = 0.5$, $P(B) = 0.2$, $P(A|B) = 0.3$

$P(A|B) = \frac{P(A \cap B)}{P(B)}$

$0.3 = \frac{P(A \cap B)}{0.2}$

$P(A \cap B) = 0.06$

(b) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$P(A \cup B) = 0.5 + 0.2 - 0.06$

$P(A \cup B) = 0.64$

(c) $P(B|A) = \frac{P(B \cap A)}{P(A)}$

$P(B|A) = \frac{0.06}{0.5}$

$P(B|A) = 0.12$

Question 47 (****)

The events A and B satisfy

$$P(A) = 0.45, \ P(A \cap B') = 0.25 \ \text{ and } \ P(A \cup B) = 0.8.$$

a) Illustrate the above information in a fully completed Venn diagram.

b) Determine …

i. … $P(A|B')$.

ii. … $P(B'|A')$

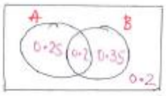
c) Find $P(A \cap B' | A' \cup B')$.

$$\boxed{P(A|B') = \tfrac{5}{9}}, \ \boxed{P(B'|A') = \tfrac{4}{11}}, \ \boxed{P(A \cap B' | A' \cup B') = \tfrac{5}{16}}$$

(a) $P(A) = 0.45$

$P(A \cap B') = 0.25$

$P(A \cup B) = 0.8$



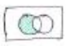
(b) (I) $P(A|B') = \frac{P(A \cap B')}{P(B')} = \frac{0.25}{0.25+0.2} = \frac{0.25}{0.45} = \frac{5}{9}$

(II) $P(B'|A') = \frac{P(B' \cap A')}{P(A')} = \frac{0.2}{0.55} = \frac{4}{11}$

(c) $P(A \cap B' | A' \cup B') = \frac{0.25}{0.8} = \frac{5}{16}$

$\downarrow$ 0.25 (from Venn DIAGRAM)

$\downarrow$ $P(A' \cup B') = P(A') + P(B') - P(A' \cap B')$
$= 0.55 + 0.45 - 0.2$
$= 0.8$



Question 50 (****)

The events A and B are such so that

$$P(A \cup B) = \tfrac{3}{5}, \ P(A) = \tfrac{2}{5}.$$

Determine $P(B)$ in each of the following three cases.

a) If A and B are mutually exclusive.

b) If A and B are independent.

c) If $P(B|A) = \frac{1}{5}$

$$\boxed{P(B) = \tfrac{1}{5}}, \ \boxed{P(B) = \tfrac{1}{3}}, \ \boxed{P(B) = \tfrac{7}{25}}$$

$$\left\{ P(A \cup B) = \frac{3}{5} \quad P(A) = \frac{2}{5} \right\}$$

(a) $P(A \cup B) = P(A) + P(B) \leftarrow$ MUTUALLY EXCLUSIVE

$$\frac{3}{5} = \frac{2}{5} + P(B)$$

$$\therefore P(B) = \frac{1}{5}$$

(b) $P(A \cup B) = P(A) + P(B) - P(A)P(B) \longleftarrow$ INDEPENDENT $P(A \cap B) = P(A) \times P(B)$

$$\frac{3}{5} = \frac{2}{5} + P(B) - \frac{2}{5}P(B)$$

$$\frac{1}{5} = P(B) - \frac{2}{5}P(B)$$

$$\frac{1}{5} = \frac{3}{5}P(B)$$

$$P(B) = \frac{1}{3}$$

(c) $P(B|A) = \dfrac{P(B \cap A)}{P(A)}$

$$\frac{1}{5} = \frac{P(B \cap A)}{\frac{2}{5}}$$

$$\frac{2}{25} = P(B \cap A)$$

- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$$\frac{3}{5} = \frac{2}{5} + P(B) - \frac{2}{25}$$

$$\frac{7}{25} = P(B)$$

$$P(B) = \frac{7}{25}$$

Question 7 (***+)

Bag X contains 3 one pound coins and 2 two pound coins.

Bag Y contains 1 one pound coin and 3 two pound coins.

A statistical experiment consists of

- a coin being picked **at random** from bag X and placed into bag Y.
- then a coin being picked **at random** from bag Y and placed back into bag X.

Find the probability …

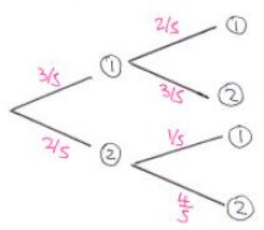
a) … that a one pound coin will be picked on both occasions in this experiment.

b) … that at the end of the experiment both bags contain £7.

A third coin is picked **at random** from bag X at the end of the experiment.

c) Determine the probability that it will be a two pound coin.

$\boxed{\frac{6}{25}}$, $\boxed{\frac{14}{25}}$, $\boxed{\frac{57}{125}}$



(a) $P(1,1) = \frac{3}{5} \times \frac{2}{5} = \frac{6}{25}$

(b) $P(1,1 \text{ or } 2,2) = \left(\frac{3}{5} \times \frac{2}{5}\right) + \left(\frac{2}{5} \times \frac{4}{5}\right)$

$$= \frac{6}{25} + \frac{8}{25} = \frac{14}{25}$$

(c)

$$\left.\begin{array}{ll} 11{:}2 & \frac{3}{5} \times \frac{2}{5} \times \frac{2}{5} = \frac{12}{125} \\ 12{:}2 & \frac{3}{5} \times \frac{3}{5} \times \frac{3}{5} = \frac{27}{125} \\ 21{:}2 & \frac{2}{5} \times \frac{1}{5} \times \frac{1}{5} = \frac{2}{125} \\ 22{:}2 & \frac{2}{5} \times \frac{4}{5} \times \frac{2}{5} = \frac{16}{125} \end{array}\right\} \text{Add } \frac{57}{125}$$

Hypothesis testing of correlation

6 The correlation coefficient between the amount of water used in a town on 30 summer days and the temperature is 0.85.

a Jane thinks that there is a correlation between water usage and temperature on a summer day. Write down the null and alternative hypotheses Jane should use to test her suspicion.

b Conduct the test at the 5% significance level.

c Karl says that if people use more water, the days will be warmer. Give two reasons why your hypothesis test does not support Karl's statement.

7 The level of antibodies in blood is thought to go down as the dose of a medical drug is increased.

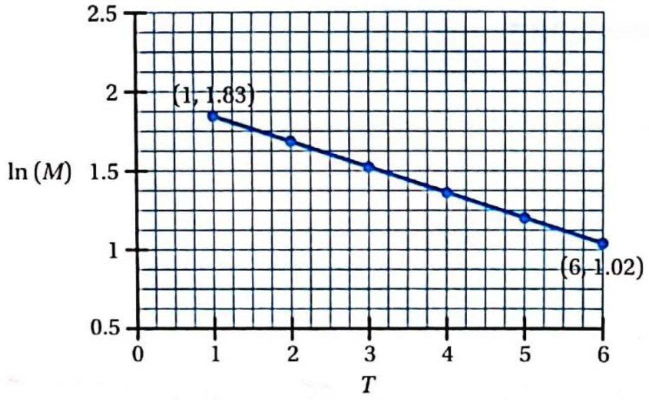
a State appropriate null and alternative hypotheses to test this statement.

b In a sample of 65 patients the correlation coefficient between these two variables is found to be -0.34. Conduct an appropriate hypothesis test at the 1% significance level.

c What are the advantages of using a 1% significance level rather than a 5% significance level for medical tests?

Modelling with exponential functions

5 A scientist is modelling exponential decay of the amount of substance in a chemical reaction. She proposes a model of the form $M = Kc^t$ where M is the mass of the substance in grams, t is the time in seconds since the start of the reaction, and K and c are constants. The mass of the substance is recorded for the first six seconds of the reaction. The graph of $\ln(M)$ against t is shown.

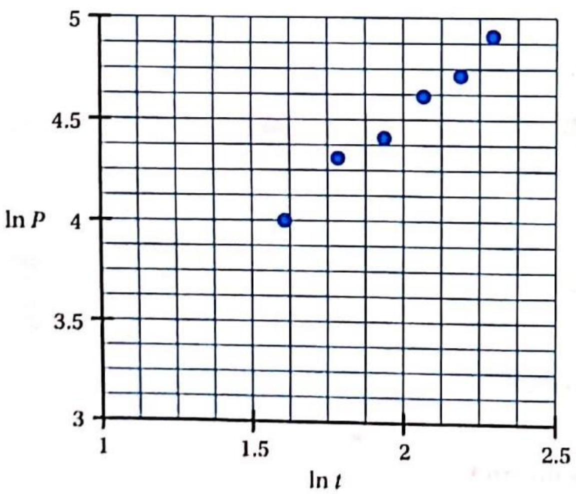


a The points are found to lie on a straight line. Find its equation, giving parameters to 2 significant figures.

b Hence find the values of K and c.

c How long, to the nearest second, will it take for the mass of the substance to fall below 1 gram?

6 A model for the size of the population of a city predicts that the population will grow according to the equation $P = Ct^n$ where P thousand is the number of people and t is the number of years since the measurements began. The graph shows $y = \ln(P)$ plotted against $x = \ln(t)$.



a Draw a line of best fit on the graph and find its equation in the form $y = mx + c$.

b Hence estimate the values of C and n.

c According to this model, after how many years will the population first exceed 200 000?

5 a $\ln(M) = -0.16t + 2.0$ b $K = 7.4$, $c = 0.85$

c 13 seconds

6 a $y = 1.2x + 2$ b $C = 7.4$, $n = 1.2$

c 5.2 years