SNOW 1   JAN 13

1. A teacher asked a random sample of 10 students to record the number of hours of television, t, they watched in the week before their mock exam. She then calculated their grade, g, in their mock exam. The results are summarised as follows.

$$\sum t = 258 \qquad \sum t^2 = 8702 \qquad \sum g = 63.6 \qquad S_{gg} = 7.864 \qquad \sum gt = 1550.2$$

(a) Find S_{tt} and S_{gt}

(3)

(b) Calculate, to 3 significant figures, the product moment correlation coefficient between t and g.

(2)

The teacher also recorded the number of hours of revision, v, these 10 students completed during the week before their mock exam. The correlation coefficient between t and v was −0.753

(c) Describe, giving a reason, the nature of the correlation you would expect to find between v and g.

(2)

a) $S_{tt} = \Sigma t^2 - (\Sigma t)^2 \div n = 8702 - 258^2 \div 10 = 2045.6$

$S_{gt} = \Sigma gt - (\Sigma g)(\Sigma t) \div n = 1550.2 - \frac{258 \times 63.6}{10}$

$= -90.68$

b) $r = \frac{S_{gt}}{\sqrt{S_{gg} \times S_{tt}}} = \frac{-90.68}{\sqrt{2045.6 \times 7.864}}$

$r = -0.715$ (3sf)

c) Positive correlation. More time spent revising the higher the exam mark.

$-0.753 \Rightarrow$ the more students spent time revising, the less time they spent watching TV.

$-0.715 \Rightarrow$ the more time students spent watching TV, the lower their exam grade.

2. The discrete random variable X can take only the values 1, 2 and 3. For these values the cumulative distribution function is defined by

$$F(x) = \frac{x^3 + k}{40} \qquad x = 1, 2, 3$$

(a) Show that $k = 13$

(2)

(b) Find the probability distribution of X.

(4)

Given that $\text{Var}(X) = \dfrac{259}{320}$

(c) find the exact value of $\text{Var}(4X - 5)$.

(2)

a)

x	1	2	3
F	$\frac{1+k}{40}$	$\frac{8+k}{40}$	$\frac{27+k}{40}$

$F(3) = 1 \quad \therefore \quad \frac{27+k}{40} = 1 \Rightarrow 27 + k = 40 \Rightarrow \underline{k = 13}$ #

b)

x	1	2	3
F	$\frac{14}{40}$	$\frac{21}{40}$	$\frac{40}{40}$

$\Rightarrow$

x	1	2	3
P	$\frac{14}{40}$	$\frac{7}{40}$	$\frac{19}{40}$

c)

$E(X) = \frac{14}{40} + \frac{14}{40} + \frac{57}{40} = \frac{85}{40}$

(not needed!)

$V(4X-5) = 4^2 V(X)$

$= 16 \times \frac{259}{320} = \underline{12.95}$

3. A biologist is comparing the intervals (m seconds) between the mating calls of a certain species of tree frog and the surrounding temperature (t °C). The following results were obtained.

t °C	8	13	14	15	15	20	25	30
m secs	6.5	4.5	6	5	4	3	2	1

(You may use $\sum tm = 469.5$, $S_{tt} = 354$, $S_{mm} = 25.5$)

(a) Show that $S_{tm} = -90.5$ **(4)**

(b) Find the equation of the regression line of m on t giving your answer in the form $m = a + bt$. **(4)**

(c) Use your regression line to estimate the time interval between mating calls when the surrounding temperature is 10 °C. **(1)**

(d) Comment on the reliability of this estimate, giving a reason for your answer. **(1)**

a) $S_{tm} = \Sigma tm - (\Sigma t)(\Sigma m) \div n$ $\quad \Sigma t = 140$, $\Sigma m = 32$

$= 469.5 - 140 \times 32 \div 8 = \underline{-90.5}$ #

b) $y = a + bx \Rightarrow y \to m$, $x \to t$; $m = a + bt$ $\quad b = \frac{S_{xy}}{S_{xx}} = \frac{S_{mt}}{S_{tt}} = -0.256$

$a = \bar{y} - b\bar{x} \Rightarrow a = \bar{m} - b\bar{t} \Rightarrow a = \left(\frac{32}{8}\right) - (-0.256..)\left(\frac{140}{8}\right)$

$a = 8.47..$

$m = 8.47 - 0.256t$

c) $t = 10$ $\quad m = 8.47 - 0.256(10) \Rightarrow m \simeq 5.9$

d) reliable as this is interpolation.

(also PMCC, $r = -0.95$; so there is strong evidence that negative correlation exists)

4. The length of time, L hours, that a phone will work before it needs charging is normally distributed with a mean of 100 hours and a standard deviation of 15 hours.

(a) Find $P(L > 127)$.

(3)

(b) Find the value of d such that $P(L < d) = 0.10$

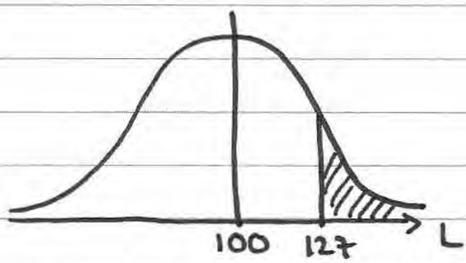
(3)

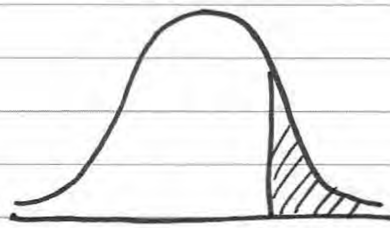
Alice is about to go on a 6 hour journey.
Given that it is 127 hours since Alice last charged her phone,

(c) find the probability that her phone will not need charging before her journey is completed.

(4)

a)

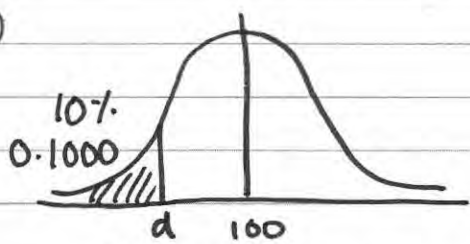




$\frac{127-100}{15}$
$= 1.8$

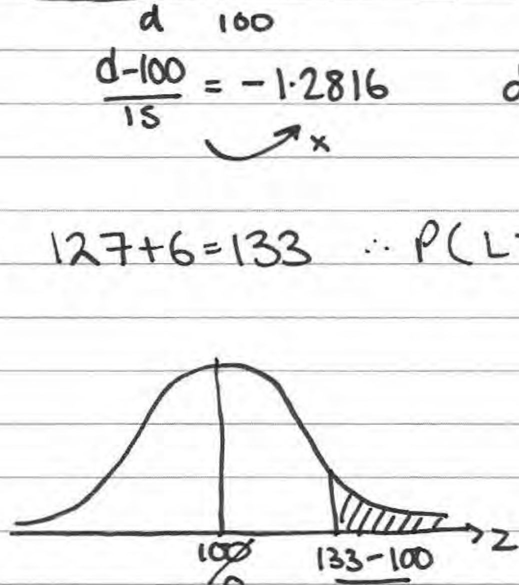
$\Phi(1.8) = 0.9641 \quad \therefore P(L>127) = \underline{0.0359}$

b)



$\frac{d-100}{15} = -1.2816 \xrightarrow{\times}$ $\quad d - 100 = -19.224 \xrightarrow{+100}$ $\therefore d = \underline{80.78}$

c) $127+6=133 \quad \therefore P(L>133)$



$2.2 \rightarrow 0.9861$

$\therefore P(L>133) = \underline{0.0139}$

$\therefore P(L>133 \mid L>127) = \frac{0.0139}{0.0359}$

$= \underline{0.39}$

5. A survey of 100 households gave the following results for weekly income £y.

Income y (£)	Mid-point	Frequency f	
$0 \leqslant y < 200$	100	12	12
$200 \leqslant y < 240$	220	28	40
$240 \leqslant y < 320$	280	22	62
$320 \leqslant y < 400$	360	18	80
$400 \leqslant y < 600$	500	12	92
$600 \leqslant y < 800$	700	8	100

(You may use $\sum fy^2 = 12\ 452\ 800$)

A histogram was drawn and the class $200 \leqslant y < 240$ was represented by a rectangle of width 2 cm and height 7 cm.

(a) Calculate the width and the height of the rectangle representing the class $320 \leqslant y < 400$

(3)

(b) Use linear interpolation to estimate the median weekly income to the nearest pound.

(2)

(c) Estimate the mean and the standard deviation of the weekly income for these data.

(4)

One measure of skewness is $\dfrac{3(\text{mean} - \text{median})}{\text{standard deviation}}$.

(d) Use this measure to calculate the skewness for these data and describe its value.

(2)

Katie suggests using the random variable X which has a normal distribution with mean 320 and standard deviation 150 to model the weekly income for these data.

(e) Find $P(240 < X < 400)$.

(2)

(f) With reference to your calculations in parts (d) and (e) and the data in the table, comment on Katie's suggestion.

(2)

a)

200 - 240

$\boxed{\begin{array}{l} A = 14\,cm^2 \\ F = 28 \end{array}}$ $\uparrow \div 2$ $h = 7cm$

$CW = 40$ $\downarrow \div 20$

$w = 2cm$

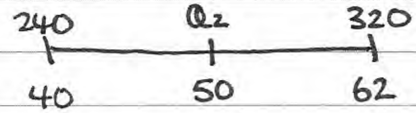
320 - 400

$\boxed{\begin{array}{l} F = 18 \\ \Rightarrow A = 9cm^2 \end{array}}$ $h = \frac{9}{4} = 2.25cm$

$CW = 80$

$\therefore \underline{w = 4cm}, \ \underline{h = 2.25cm}$

b)



$\frac{Q_2 - 240}{80} = \frac{10}{22} \Rightarrow \underline{Q_2 = 276.4}$

$\simeq £276$

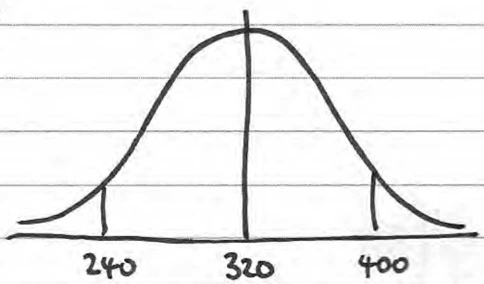
c) $\Sigma fy = 31600$ $\quad \bar{y} = \frac{\Sigma fy}{n} = \frac{31600}{100} = \underline{316}$

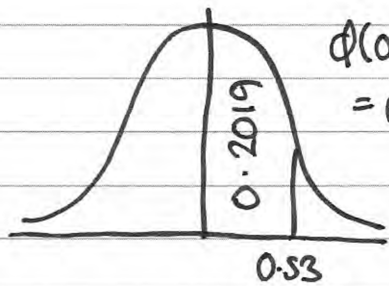
$S_{yy} = \Sigma fy - (\Sigma fy)^2 \div n = 12\,452\,800 - 31600^2 \div 100$

$S_{yy} = 2467200 \quad \div n \Rightarrow \sqrt{\ } \Rightarrow sd = \underline{157.1}$

d) $\text{Skew} = \frac{3(316 - 276.4)}{157.1} = +0.76 \Rightarrow$ Positive skew

e)





$\Phi(0.53)$

$= 0.7019$

$\frac{240-320}{150} = -0.53 \qquad \frac{400-320}{150} = +0.53$

$\therefore P(240 < X < 400) = 2 \times 0.2019 = 0.4038$

f) from table $P(240 < X < 400) = \frac{22+18}{100} = 0.40$

So this seems reasonable but since data is skewed :O normal distribution is not sensible.

6. A fair blue die has faces numbered 1, 1, 3, 3, 5 and 5. The random variable B represents the score when the blue die is rolled.

(a) Write down the probability distribution for B.

(2)

(b) State the name of this probability distribution.

(1)

(c) Write down the value of $\mathrm{E}(B)$.

(1)

A second die is red and the random variable R represents the score when the red die is rolled.

The probability distribution of R is

r	2	4	6
$\mathrm{P}(R = r)$	$\frac{2}{3}$	$\frac{1}{6}$	$\frac{1}{6}$

(d) Find $\mathrm{E}(R)$.

(2)

(e) Find $\mathrm{Var}(R)$.

(3)

Tom invites Avisha to play a game with these dice.

Tom spins a fair coin with one side labelled 2 and the other side labelled 5. When Avisha sees the number showing on the coin she then chooses one of the dice and rolls it. If the number showing on the die is greater than the number showing on the coin, Avisha wins, otherwise Tom wins.

Avisha chooses the die which gives her the best chance of winning each time Tom spins the coin.

(f) Find the probability that Avisha wins the game, stating clearly which die she should use in each case.

(4)

a)

B	1	3	5
P	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

b) Uniform distribution

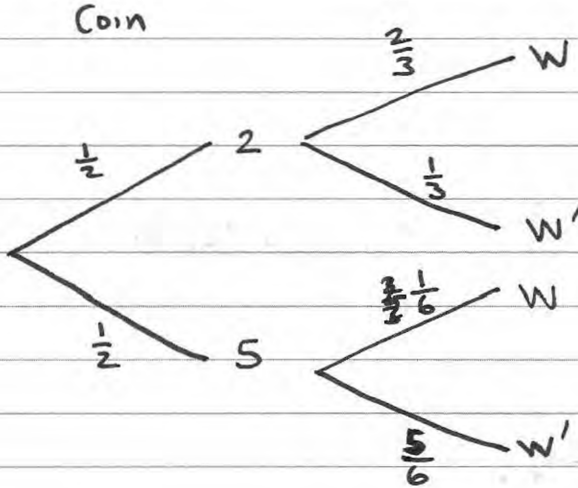
c) $E(B) = 3$

d) $E(R) = \frac{4}{3} + \frac{4}{6} + \frac{6}{6} = \frac{18}{6} = 3$

e) $E(R^2) = \frac{16}{6} + \frac{16}{6} + \frac{36}{6} = \frac{68}{6}$

$V(R) = E(R^2) - E(R)^2 = \frac{68}{6} - 3^2 = \frac{7}{3}$.

f) If Coin $= 2$, Auisha chooses blue dice to roll $P(W) = \frac{2}{3}$
If Coin $= 5$, Auisha chooses red dice to roll $P(W) = \frac{1}{6}$



$2, W = \frac{2}{6}$

$5, W = \frac{1}{12}$

$\therefore$ Prob Auisha wins

$= \frac{2}{6} + \frac{1}{12} = \frac{5}{12}$

7. Given that

$$P(A) = 0.35\,, \quad P(B) = 0.45 \quad \text{and} \quad P(A \cap B) = 0.13$$

find

(a) $P(A \cup B)$ **(2)**

(b) $P(A' \mid B')$ **(2)**

The event C has $P(C) = 0.20$

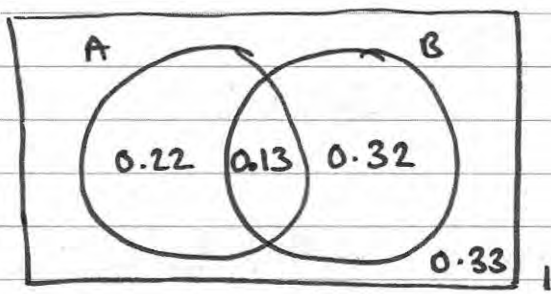
The events A and C are mutually exclusive and the events B and C are independent.

(c) Find $P(B \cap C)$ **(2)**

(d) Draw a Venn diagram to illustrate the events A, B and C and the probabilities for each region. **(4)**

(e) Find $P([B \cup C]')$ **(2)**

a)

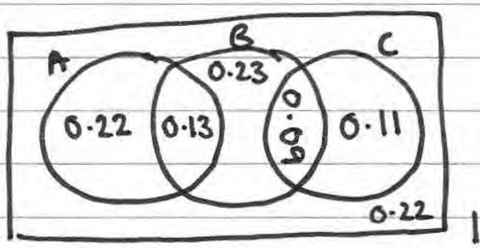


$P(A \cup B) = 0.67$

b) $P(A' \mid B') = \dfrac{0.33}{0.55} = 0.6$

c) $P(B \cap C) = P(B) \times P(C) = 0.45 \times 0.20 = 0.09$

d)



$P(B \cup C) = 0.56$

$\Rightarrow P([B \cup C]') = \underline{0.44}$