1. Helen believes that the random variable C, representing cloud cover from the large data set, can be modelled by a discrete uniform distribution.

(a) Write down the probability distribution for C. (2)

(b) Using this model, find the probability that cloud cover is less than 50% (1)

Helen used all the data from the large data set for Hurn in 2015 and found that the proportion of days with cloud cover of less than 50% was 0.315

(c) Comment on the suitability of Helen's model in the light of this information. (1)

(d) Suggest an appropriate refinement to Helen's model. (1)

a) Daily mean cloud cover (page 12 in text book) is measured in 'oktas' or eighths of the sky covered. $\frac{0}{8}$, $\frac{1}{8}$ etc $\frac{8}{8}$

c	0	1	2	3	4	5	6	7	8
$P(C=c)$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$

b) less than 50%

$P(C<4) = \frac{4}{9}$

c) The probability 0.315 is lower than the expected value $\frac{4}{9}$ (0.4) which suggests the model is not good.

d) Cloud cover will vary from month to month and season to season, so a non-uniform distribution may be a better model.

2.

In an experiment a group of children each repeatedly throw a dart at a target.
For each child, the random variable H represents the number of times the dart hits the target in the first 10 throws.

Peta models H as $B(10, 0.1)$

(a) State two assumptions Peta needs to make to use her model.

(2)

(b) Using Peta's model, find $P(H \geqslant 4)$

(1)

For each child the random variable F represents the number of the throw on which the dart first hits the target.

Using Peta's assumptions about this experiment,

(c) find $P(F = 5)$

(2)

Thomas assumes that in this experiment no child will need more than 10 throws for the dart to hit the target for the first time. He models $P(F = n)$ as

$$P(F = n) = 0.01 + (n - 1) \times \alpha$$

where α is a constant.

(d) Find the value of α

(4)

(e) Using Thomas' model, find $P(F = 5)$

(1)

(f) Explain how Peta's and Thomas' models differ in describing the probability that a dart hits the target in this experiment.

(1)

a) $H \sim B(10, 0.1)$

1) Probability of each child hitting the target is the same

2) Dart throws are independent of each other.

b) $P(H \geqslant 4) = 1 - P(x \leq 3)$

Menu 7 → 1: Binomial CD

2: Variable
$x = 3$
$n = 10$
$p = 0.1$

$P(x \leq 3) = 0.987204$

$= 1 - 0.987204$

$= 0.012795$

c) $P(F=5) = \underbrace{0.9 \times 0.9 \times 0.9 \times 0.9}_{\text{4 misses}} \times \underbrace{0.1}_{\text{1st hit}}$

$= 0.06561$

d)

n	$P(F=n)$
1	$0.01 + (1-1) \times \alpha = 0.01$
2	$0.01 + (2-1)\alpha = 0.01 + \alpha$
3	$0.01 + (3-1)\alpha = 0.01 + 2\alpha$
...	
10	$0.01 + (10-1)\alpha = 0.01 + 9\alpha$

$\therefore (10 \times 0.01) + \alpha + 2\alpha + 3\alpha + 4\alpha + 5\alpha$
$+ 6\alpha + 7\alpha + 8\alpha + 9\alpha = 1$

$0.1 + 45\alpha = 1$

$45\alpha = 0.9$

$\alpha = \frac{1}{50} = 0.02$

e) $P(F=5) = 0.01 + (5-1) \times 0.02$

$= 0.09$

f) Peta's model assumes probability of each dart hitting is constant.
Thomas's assumes probability of hitting increases with each throw.

3. The discrete random variable D has the following probability distribution

d	10	20	30	40	50
$P(D = d)$	$\frac{k}{10}$	$\frac{k}{20}$	$\frac{k}{30}$	$\frac{k}{40}$	$\frac{k}{50}$

where k is a constant.

(a) Show that the value of k is $\frac{600}{137}$

(2)

The random variables D_1 and D_2 are independent and each have the same distribution as D.

(b) Find $P(D_1 + D_2 = 80)$
Give your answer to 3 significant figures.

(3)

A single observation of D is made.

The value obtained, d, is the common difference of an arithmetic sequence.

The first 4 terms of this arithmetic sequence are the angles, measured in degrees, of quadrilateral Q

(c) Find the exact probability that the smallest angle of Q is more than 50°

(5)

4a) $k\left(\frac{1}{10} + \frac{1}{20} + \frac{1}{30} + \frac{1}{40} + \frac{1}{50}\right) = 1$

$\frac{137}{600}k = 1$

$k = \frac{600}{137}$ (as required)

b) For $D_1 + D_2 = 80$

$D_1 = 30 \quad D_2 = 50$

or $D_1 = 50 \quad D_2 = 30$

or $D_1 = 40 \quad D_2 = 40$

$= \frac{k}{30} \times \frac{k}{50} + \frac{k}{50} \times \frac{k}{30} + \frac{k}{40} \times \frac{k}{40}$

$= k^2\left(\frac{1}{1500} + \frac{1}{1500} + \frac{1}{1600}\right)$

$= \left(\frac{600}{137}\right)^2\left(\frac{1}{1500} + \frac{1}{1500} + \frac{1}{1600}\right)$

$= \frac{705}{18769} = 0.03756 = 0.0376$ (3 sf)

1st	2nd	3rd	4th
c) a	$a+d$	$a+2d$	$a+3d$

$$S_n = \frac{n}{2}(2a + (n-1)d)$$

$n = 4$, $S_n = 360$ (angles in quadrilateral)

$$360 = 2(2a + 3d)$$

$$\therefore 2a + 3d = 180$$

From table

if $d = 10$, $a = \dfrac{180 - 30}{2} = 75$ ✓

$d = 20$, $a = \dfrac{180 - 60}{2} = 60$ ✓

$d = 30$, $a = \dfrac{180 - 90}{2} = 45$ ✗

a must be > 50

$\therefore$ $d = 10$ or $d = 20$

$$= \frac{k}{10} + \frac{k}{20}$$

$$= \frac{600}{137}\left(\frac{1}{10} + \frac{1}{20}\right)$$

$$= \frac{90}{137}$$ exact probability

4. A machine puts liquid into bottles of perfume. The amount of liquid put into each bottle, D ml, follows a normal distribution with mean 25 ml $\mu = 25$

Given that 15% of bottles contain less than 24.63 ml

(a) find, to 2 decimal places, the value of k such that $P(24.63 < D < k) = 0.45$ **(5)**

A random sample of 200 bottles is taken.

(b) Using a normal approximation, find the probability that fewer than half of these bottles contain between 24.63 ml and k ml **(3)**

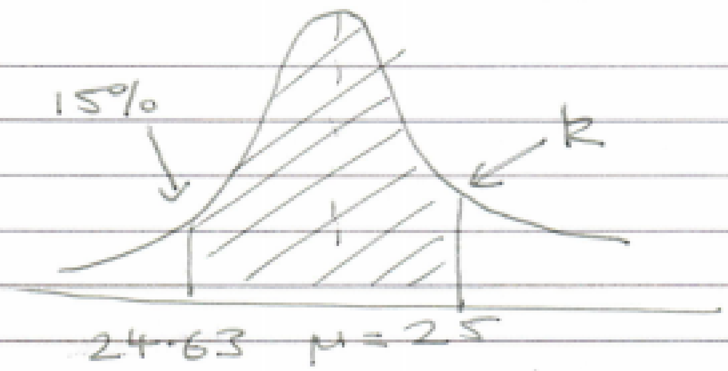
The machine is adjusted so that the standard deviation of the liquid put in the bottles is now 0.16 ml

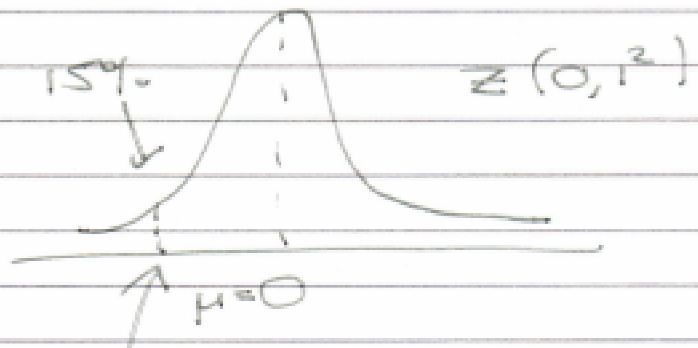
Following the adjustments, Hannah believes that the mean amount of liquid put in each bottle is less than 25 ml

She takes a random sample of 20 bottles and finds the mean amount of liquid to be 24.94 ml

(c) Test Hannah's belief at the 5% level of significance.
You should state your hypotheses clearly. **(5)**

a)





Inverse normal
Area = 0.15
$\sigma = 1$
$\mu = 0$

$Z = -1.0364$ $\quad$ $Z = -1.0364$

$$Z = \frac{X - \mu}{\sigma}$$

$$-1.0364 = \frac{24.63 - 25}{\sigma}$$

$$\Rightarrow \sigma = \frac{24.63 - 25}{-1.0364} = 0.35699$$

$X \sim N(25, 0.35699^2)$

% shaded $= 45\%$

$15\% + 45\%$

Inverse normal
Area $= 0.6$
$\sigma = 0.35699$
$\mu = 25$

$k = 25.09$ (2dp)

b) $n = 200$ $p = 0.45$
(between 24.63 and k)

$\mu = np = 200 \times 0.45 = 90$
$\sigma = \sqrt{np(1-p)} = \sqrt{90(0.55)} = 7.0356$

$X \sim N(90, 7.0356^2)$

μ σ

Fewer than half of 200 is 99

$\therefore P(X \leq 99)$ needs continuity correction
$P(X \leq 99.5)$

Normal CD
Lower -100
Upper 99.5
$\sigma = 7.0356$
$\mu = 90$

$= 0.911536$

$= 0.912$ (3sf)

5. A health centre claims that the time a doctor spends with a patient can be modelled by a normal distribution with a mean of 10 minutes and a standard deviation of 4 minutes.

(a) Using this model, find the probability that the time spent with a randomly selected patient is more than 15 minutes.

(1)

Some patients complain that the mean time the doctor spends with a patient is more than 10 minutes.

The receptionist takes a random sample of 20 patients and finds that the mean time the doctor spends with a patient is 11.5 minutes.

(b) Stating your hypotheses clearly and using a 5% significance level, test whether or not there is evidence to support the patients' complaint.

(4)

The health centre also claims that the time a dentist spends with a patient during a routine appointment, T minutes, can be modelled by the normal distribution where $T \sim N(5, 3.5^2)$

(c) Using this model,

(i) find the probability that a routine appointment with the dentist takes less than 2 minutes

(1)

(ii) find $P(T < 2 \mid T > 0)$

(3)

(iii) hence explain why this normal distribution may not be a good model for T.

(1)

The dentist believes that she cannot complete a routine appointment in less than 2 minutes.

She suggests that the health centre should use a refined model only including values of $T > 2$

(d) Find the median time for a routine appointment using this new model, giving your answer correct to one decimal place.

(5)

a) $X \sim N(10, 4^2)$ Normal CD

$P(X > 15)$ lower $= 15$

upper $= 1000$

$\sigma = 4$

$\mu = 10$

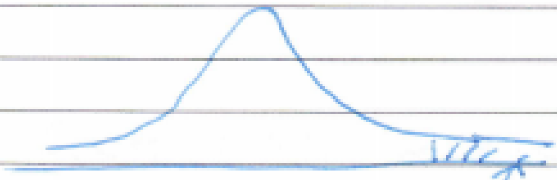
$p = 0.1056$

Question 5 continued

b) $n = 20$ $\qquad \sigma = \frac{\sigma}{\sqrt{n}} = \frac{4}{\sqrt{20}} = 0.8944$

$\bar{X} = 11.5$ $\qquad H_0: \mu = 10 \quad H_1: \mu > 10$

5% 1 tail test



0.95

Inverse normal
Area = 0.95
$\sigma = 0.8944$
$\mu = 10$

test statistic
$\bar{X} = 11.5$ minutes
in tail

critical value

$x = 11.4711$

As 11.5 is in tail, reject H_0, accept H_1, there is evidence to support complaint that $\mu > 10$ minutes for appointments now.

c) (i) $T \sim N(5, 3.5^2)$ $\qquad$ Normal CD
$P(T < 2)$
lower = −1000
upper = 2
$\sigma = 3.5$
$\mu = 5$

$p = 0.1957$ (4dp)

(ii) $P(0 < T < 2) = 0.1191$

$P(T < 2 \mid T > 0)$

$= \frac{0.1191}{0.9234}$

$P(T<2)$

lower 0
upper 1000

$= 0.1290$ (4dp)

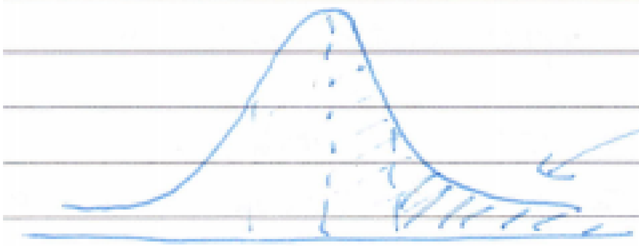
5a) median $p = 0.5$

$$\therefore \frac{P(T > t)}{P(T > 2)} = 0.5$$

$$P(T > t) = 0.5 \times P(T > 2)$$

Lower = 2
upper = 1000
$\sigma = 3.5$
$\mu = 5$
$p = 0.8043$

$$P(T > t) = 0.5 \times 0.8043 = 0.4022$$



t $p = 0.4022$

$1 - 0.4022 = 0.5978$

0.5978

Inverse normal
Area = 0.5978
$\sigma = 3.5$
$\mu = 5$

$t = 5.8667$
$t = 5.9$ minutes (1dp)