A Level H2 Math

Normal Distribution Test 1

Q1

(a) An examination taken by a large number of students is marked out of a total score of 100. It is found that the mean is 73 marks and that the standard deviation is 15 marks.

(i) Give a reason why the normal distribution is not a good model for the distribution of marks for the examination. [1]

(ii) The marks for a random sample of 50 students is recorded. Find the probability that the mean mark of this sample lies between 70 and 75. [2]

(b) The interquartile range of a distribution is the difference between the upper and lower quartile values for the distribution. The lower quartile value, l, of a distribution X, is such that $\mathrm{P}(X < l) = 0.25$. The upper quartile value, u, of the same distribution is such that $\mathrm{P}(X < u) = 0.75$.

The marks of another examination is known to follow a normal distribution. If a student who scores 51 marks is at the 80th percentile, and the interquartile range is found to be 10.8 marks, find the mean mark and the standard deviation of the marks scored by students who took the examination. [5]

(c) In a third examination, the marks scored by students are normally distributed with a mean of 52 marks and a standard deviation of 13 marks.

(i) If 50 is the passing mark and 289 students are expected to pass, how many candidates are there? [2]

(ii) Find the smallest integer value of m such that more than 90% of the candidates will score within m marks of the mean. [3]

Q2

A manufacturing plant processes raw material for a supplier. An order placed with the plant is considered to be a bulk order when a worker is expected to process more than 300 kg (kilograms) of raw material.

Albert uses a machine to process X kg of raw material and Bob uses a separate machine to process Y kg of raw material on a working day. X and Y are independent random variables with the distributions $\mathrm{N}(296,\ 8^2)$ and $\mathrm{N}(290,\ 12^2)$ respectively.

(i) Find the probability that Albert processes more than 300 kg of raw material on a randomly selected working day. [2]

(ii) Find the probability that, over a period of 15 independent working days, there are exactly four working days on which Albert processes more than 300 kg of raw material. [2]

(iii) Find the probability that the total amount of raw material Bob processes over two working days exceeds twice the amount of raw material Albert processes on one working day. [4]

The plant receives a bulk order and Albert wants to have a probability of at least 0.95 of meeting the order.

(iv) This can be done by changing the value of $\propto$, the mean amount of raw material Albert processes using the machine, but the standard deviation remains unchanged. Find the least value of $\propto$. [3]

Q3

Each month the amount of electricity, X measured in kilowatt-hours (kWh), used by a household in a particular city may be assumed to follow a normal distribution with mean 950 and standard deviation σ. The charge for electricity used per month is fixed at $0.22 per kWh.

(i) Given that 65% of the households uses less than 960 kWh of electricity in a month, find the value of σ, correct to 1 decimal place. [2]

For the rest of the question, σ is the value found in part (i).

(ii) Find the probability that the difference in the amount of electricity used among 2 randomly chosen households in a particular month is not more than 30 kWh. [3]

(iii) In the month of August, the mayor of the city decides to provide 50% and 30% subsidies for the electricity bills of households in the North and South distincts of the city respectively. Find the probability that the total electricity bill of 2 randomly chosen North district households and 1 South district household is less than $360. [4]

(iv) In December, a random sample of n households is chosen to study the mean monthly electricity usage per household in the city. Find the least value of n if the probability of the sample mean being less than 955 kWh is at least 0.9. [3]

Answers

Normal Distribution Test 1

Q1

(a) (i)	Let X be the random variable 'marks of an examination'. By GC, $P(X > 100) = 0.0359$ if $X \sim N(73, 15^2)$ i.e., there are 3.59% of the students scoring more than the maximum mark of 100, which is impossible.	Common wrong answers are: The students marks are not independent of one another / the mean should be around 50 / mark is a discrete random variable / mark cannot take negative values or values above 100. Students need to understand that normal distribution is a model to help analyze the data and can be applied as long the population is large and the values that it cannot take have negligible probabilities.
(a) (ii)	Since $n = 50 \geq 20$ is large, by Central Limit Theorem, $\overline{X} \sim N(73, \frac{15^2}{50})$ approximately. $\therefore\ P(70 < \overline{X} < 75) = 0.748$	Majority assumed X is normal and then applied CLT for $\overline{X}$. This questions shows that most people do not understand the meaning of $\overline{X}$.

) (b)	Let Y be the random variable 'marks of a school examination'. $Y \sim N(\mu, \sigma^2)$ $P(Y < 51) = 0.8$ $P(Z < \frac{51-\mu}{\sigma}) = 0.8$ $\frac{51-\mu}{\sigma} = 0.84162$ $\mu + 0.84162\sigma = 51$ $P(\mu - 5.4 < Y < \mu + 5.4) = 0.5$ $P(\frac{-5.4}{\sigma} < Z < \frac{5.4}{\sigma}) = 0.5$ $P(Z < -\frac{5.4}{\sigma}) = 0.25$ $-\frac{5.4}{\sigma} = -0.67449$ $\therefore \quad \sigma = 8.01$ $\therefore \quad \mu = 51 - 0.84162(8.0061) = 44.3$	Quite a number had problem with 80th percentile: $P(Y > 51) = 0.8$ & $P(Y = 51)$ are WRONG! Standardisation should be $\boxed{Z = \frac{X-\mu}{\sigma}}$ Note that InvNorm $(0.8) = 0.84162$ InvNorm $(0.8) \neq 0.8$ Note the interquartile range and its related probability: $P(Y < u) - P(Y < l) = 0.5$ where $u - l = 10.8$ $P(Y < u) - P(Y < l) = 10.8$ is WRONG!
) (c) (i)	Let M be the random variable 'marks of another school examination'. $M \sim N(52, 13^2)$ $P(50 < M) = 0.56113$ Number of passes = (total candidature) $\times 0.56113 = 289$ $\therefore$ total candidature $= 289 \div 0.56113 = 515$	
0 (c) (ii)	$P(\lvert M - 52 \rvert < m) > 0.9 \Rightarrow P(52 - m < M < 52 + m) > 0.9$ where $M \sim N(52, 13^2)$ $\Rightarrow P(M < 52 - m) < 0.05$ $\Rightarrow \quad 52 - m < 30.6$ $\Rightarrow \quad m > 21.4$ $\therefore$ Smallest integral value of $m = 22$	**P**: Missing first step **R**: *m* marks from mean, 90%, more than, etc. As 52 & 13 are given, there is no need for standardisation. The preferred method is InvNorm(0.05, 52, 13). Trial and error using GC table is not advisable.

Q2

(i)	$X \sim \text{N}(296,\ 8^2)$ $\quad$ $Y \sim \text{N}(290,\ 12^2)$ Required probability = [P(X > 300)] = 0.30854 (5 s.f.) (0.3085375322) = 0.309 (3 s.f.)
(ii)	Let W be the number of days in which Albert processes more than 300 kg of raw material on that day out of 15 days. $W \sim \text{B}(15, 0.30854)$ $\text{P}(W = 4) = 0.214$ (3 s.f.)
(iii)	Let $S = Y_1 + Y_2 - 2X$ $\text{E}(S) = \text{E}(Y_1) + \text{E}(Y_2) - 2\text{E}(X) = 2\times 290 - 2\times 296 = -12$ $\text{Var}(S) = 2\text{Var}(Y) + 2^2\,\text{Var}(X) = 2\times 12^2 + 2^2\times 8^2 = 544$ Hence, $S \sim \text{N}(-12,\ 544)$ $\text{P}(S > 0) = 0.303$ (3 s.f.) (0.3034526994)
(iv)	$X \sim \text{N}(\mu,\ 8^2)$ $\text{P}(X > 300) = \text{P}\left(Z > \frac{300-\mu}{8}\right) \geq 0.95$ $\text{P}\left(Z \leq \frac{300-\mu}{8}\right) \leq 0.05$ $\frac{300-\mu}{8} \leq -1.6449$ $\mu \geq 313.1592$ Least value of $\mu = 314$ kg (3 s.f.)

Q3

(i)

$X \sim \text{N}(950,\ \sigma^2)$

Given that $\text{P}(X < 960) = 0.65$,

then $\text{P}(Z < \dfrac{960-950}{\sigma}) = 0.65$

$\Rightarrow \dfrac{960-950}{\sigma} = 0.3853204726$

$\Rightarrow \sigma = 25.95242327 = 26.0$ (1 decimal place)

(ii)

Let X_1 and X_2 be the amount of electricity used by the 2 randomly chosen household in a particular month.

Then $X_1 - X_2 \sim \text{N}(0,\ 26.0^2 + 26.0^2)$

Thus, $\text{P}(|X_1 - X_2| \le 30)$

$= \text{P}(-30 \le X_1 - X_2 \le 30)$

$= 0.585$

(iii)

Let N_1, N_2 and S be the amount of electricity used by the 2 randomly chosen households in the North District and household in the South district respectively in August.

Then their total electricity bill = \$ T

= \$ $0.5 \times 0.22 \times (N_1 + N_2) + 0.7 \times 0.22 \times S$

= \$ $0.11N_1 + 0.11N_2 + 0.154S$

Then

$\text{E}(T) = 0.11 \times 950 \times 2 + 0.154 \times 950 = 355.3$

$\text{Var}(T) = 0.11^2 \times 26.0^2 \times 2 + 0.154^2 \times 26.0^2 = 32.391216$

So, $T \sim \text{N}(355.3,\ 32.391216)$

Hence, $\text{P}(T < 360) = 0.796$

(iv)

Let $\overline{X} = \dfrac{X_1 + X_2 + X_3 + \ldots\ldots + X_n}{n} \sim \mathrm{N}\left(950, \dfrac{26.0^2}{n}\right)$

where X_i :electricity usage for each of the randomly selected household in the month of December

Then, we have

$$\mathrm{P}\left(\overline{X} < 955\right) \geq 0.9$$

$$\Rightarrow \mathrm{P}\left(Z < \frac{955-950}{26.0/\sqrt{n}}\right) \geq 0.9$$

$$\Rightarrow \frac{955-950}{26.0/\sqrt{n}} \geq 1.281551567$$

$$\Rightarrow \frac{26.0}{\sqrt{n}} \leq \frac{5}{1.281551567} = 3.901520726$$

$$\Rightarrow \sqrt{n} \geq \frac{26}{3.901520726}$$

$$\Rightarrow n \geq 44.40980429$$

Thus, the least value of n is 45.