RAFFLES INSTITUTION
H2 Mathematics 9758
2023 Year 6 Term 3 Revision 14 (Summary and Tutorial)

Topic: Discrete Random Variable, Binomial Distribution, Normal Distribution

---

Summary for Discrete Random Variable

Probability Distribution

A table or formula giving the values of $\mathrm{P}(X=x)$ for every x in sample space is called the **probability distribution** of X. For the experiment of tossing a fair coin 3 times where X is the number of heads obtained, the **probability distribution** of X is as follows

x	0	1	2	3
$\mathrm{P}(X=x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

The probability distribution of X satisfy the following:

1. $0 \le \mathrm{P}(X=x) \le 1$ for all x in S.
2. $\sum_{x\in S} \mathrm{P}(X=x)=1$ where the summation is over all values of x in S.

Expectation of Discrete Random Variable

The expectation (or mean, or expected value) of a discrete random variable X taking values from a set S is given by $\mathrm{E}(X)=\sum_{x\in S} x\,\mathrm{P}(X=x)$.

Independent Discrete Random Variables

Let X and Y be two discrete random variables taking on possible values $x_1, x_2, \ldots$ and $y_1, y_2, \ldots$ respectively. The random variables X and Y are said to be **independent** if for all i and j,

$$\mathrm{P}(X=x_i \text{ and } Y=y_j)=\mathrm{P}(X=x_i)\mathrm{P}(Y=y_j).$$

Functions of a Discrete Random Variable

The expectation of $\mathrm{g}(X)$, where g is a function of X, is denoted by $\mathrm{E}(\mathrm{g}(X))=\sum_{x\in S}\mathrm{g}(x)\mathrm{P}(X=x)$.

In particular, $\mathrm{E}(X^2)=\sum_{x\in S} x^2\,\mathrm{P}(X=x)$.

Variance and Standard Deviation of a Discrete Random Variable

The variance of a discrete random variable X is given by

$$\begin{aligned}\mathrm{Var}(X) = \mathrm{E}(X-\mu)^2 &= \sum (x-\mu)^2 \, \mathrm{P}(X=x) \\ &= \mathrm{E}\left(X^2\right)-\mu^2 = \left(\sum x^2 \mathrm{P}(X=x)\right)-\mu^2\end{aligned}.$$

The standard deviation of X, denoted by σ, is defined as $\sqrt{\mathrm{Var}(X)}$.

Properties of Expectation and Variance

(*Note that these properties hold for discrete and continuous random variables*)

Let X and Y be random variables and a and b be constants. We have

Expectation	Variance
(1) $\mathrm{E}(a) = a$ (2) $\mathrm{E}(aX) = a\,\mathrm{E}(X)$ (3) $\mathrm{E}(aX \pm b) = a\mathrm{E}(X) \pm b$ (4) $\mathrm{E}(aX \pm bY) = a\mathrm{E}(X) \pm b\mathrm{E}(Y)$	(1) $\mathrm{Var}(a) = 0$ (2) $\mathrm{Var}(aX) = a^2\mathrm{Var}(X)$ (3) $\mathrm{Var}(aX \pm b) = a^2\mathrm{Var}(X)$ If X and Y are *independent* random variables, then (4) $\mathrm{Var}(aX \pm bY) = a^2\mathrm{Var}(X) + b^2\mathrm{Var}(Y)$

Important Results

In general, $\mathrm{E}(nX) = n\,\mathrm{E}(X)$

$$\mathrm{E}(X_1 + X_2 + \ldots + X_n) = \mathrm{E}\left(X_1\right) + \mathrm{E}\left(X_2\right) + \ldots + \mathrm{E}\left(X_n\right) = n\,\mathrm{E}\left(X\right)$$

but $\mathrm{Var}(nX) = n^2\,\mathrm{Var}(X)$

$$\mathrm{Var}(X_1 + X_2 + \ldots + X_n) = \mathrm{Var}(X_1) + \mathrm{Var}(X_2) + \ldots + \mathrm{Var}\left(X_n\right) = n\,\mathrm{Var}(X).$$

Summary for Binomial Distribution

Note that Binomial random variable is a special **discrete random variable**.

Conditions for an experiment to follow binomial distribution

1. It consists of n independent trials.
2. The outcome of each trial is either a success or a failure.
3. The probability of success for each trial, denoted by p, remains constant.

Note that the above must be stated **in context** of a given question.

For example, a biased coin has probability 0.56 of obtaining a head in any toss. Find the probability of getting 6 heads if the coin is tossed 8 times.

Conditions in context as follow

1. The coin is tossed independently 8 times.
2. The outcome of each toss is either a head (success) or tail (failure).
3. The probability of obtaining a head (success) remains constant at 0.56.

Binomial Distribution

If a discrete random variable X follows a **binomial distribution**, we write $X \sim \mathrm{B}(n, p)$

where the parameters of the distribution n and p refer to the number of trials and the probability of success for each trial respectively.

The probability distribution of X is given by (in MF26)

$$\mathrm{P}(X = x) = \binom{n}{x} p^x (1-p)^{n-x}, \text{ for } x = 0, 1, 2, \ldots, n, \text{ where } \binom{n}{x} = {}^nC_x = \frac{n!}{(n-x)!x!}$$

The mean and variance of X are given by

(i) $\mathrm{E}(X) = np$ (in MF26)

(ii) $\mathrm{Var}(X) = np(1-p)$ (in MF26)

Summary for Normal Distribution

Note that Normal random variable is a special continuous random variable. For continuous random variable, probability is calculated using the area under its probability density function.

In H2 Math syllabus, we just need to know how to use GC to evaluate these probabilities.

We also need to be aware that

$\mathrm{P}(X<x) \neq \mathrm{P}(X \leq x)$ for discrete random variable (Binomial included)

BUT

$\mathrm{P}(X<x)=\mathrm{P}(X \leq x)$ for continuous random variable

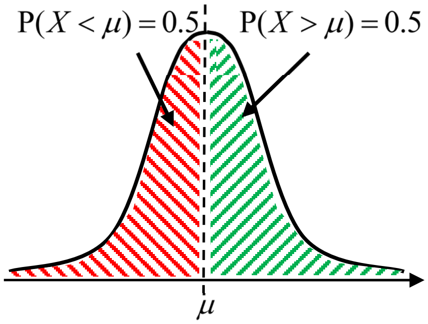
Normal Distribution

If a continuous random variable X follows a normal distribution, we write $X \sim \mathrm{N}(\mu,\sigma^2)$,

where $\mathrm{E}(X)=\mu$ and $\mathrm{Var}(X)=\sigma^2$.

Properties of a Normal Curve

Let $X \sim \mathrm{N}(\mu,\sigma^2)$.



(1) It is symmetrical about the line $x=\mu$.

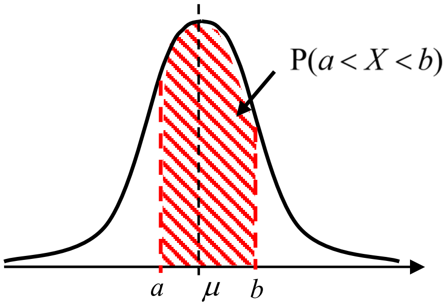
(2) The mean, median and mode are all equal to μ.

(3) It approaches the $x-$axis as $x \to \pm\infty$.

(4) Area under the graph gives the probabilities,

i.e., $\mathrm{P}(a<X<b)=\int_a^b \mathrm{f}(x)\,\mathrm{d}x$

where $y=\mathrm{f}(x)$ represents the probability density function of the normal curve.



Hence $\mathrm{P}(a<X<b)$ is given by the area under the graph from $x=a$ to $x=b$.

(5) Total area under the curve is 1.

(6) $\mathrm{P}(\mu-\sigma < X < \mu+\sigma) \approx 0.68$

$\mathrm{P}(\mu-2\sigma < X < \mu+2\sigma) \approx 0.95$

$\mathrm{P}(\mu-3\sigma < X < \mu+3\sigma) \approx 0.997$

i.e., approximately 68%, 95% and 99.7% of the values drawn from a normal distribution lies within 1, 2 and 3 standard deviations of the mean respectively.

Standard Normal Distribution

Let $X \sim \mathrm{N}(\mu, \sigma^2)$. The random variable Z, which is the standard normal variable, is defined by $Z = \dfrac{X-\mu}{\sigma}$. The standard normal distribution is $Z \sim \mathrm{N}(0,1)$.

The process of converting $X \sim \mathrm{N}(\mu, \sigma^2)$ into $Z \sim \mathrm{N}(0,1)$ is known as **standardization**.

$$\mathrm{P}(X \le x) = \mathrm{P}\left(\frac{X-\mu}{\sigma} \le \frac{x-\mu}{\sigma}\right) = \mathrm{P}\left(Z \le \frac{x-\mu}{\sigma}\right)$$

Using the Properties of Expectation and Variance of Random Variables, we have the following results for independent Normal Random Variables

Let $X \sim \mathrm{N}(\mu_1, \sigma_1{}^2)$ and $Y \sim \mathrm{N}(\mu_2, \sigma_2{}^2)$ be **independent** random variables and *a* and *b* be constants. We have

(1) $X \pm Y \sim \mathrm{N}(\mu_1 \pm \mu_2, \sigma_1{}^2 + \sigma_2{}^2)$

(2) $aX \pm b \sim \mathrm{N}(a\mu_1 \pm b, a^2\sigma_1{}^2)$

(3) $aX \pm bY \sim \mathrm{N}(a\mu_1 \pm b\mu_2, a^2\sigma_1{}^2 + b^2\sigma_2{}^2)$

Revision Tutorial Questions

Source of Question: ACJC Prelim 9758/2017/02/Q6

1. Alex and his friend stand randomly in a queue with 3 other people. The random variable X is the number of people standing between Alex and his friend.

(i) Show that P(X = 2) = 0.2. [2]

(ii) Tabulate the probability distribution of X. [2]

(iii) Find E(X) and E(X – 1)^2. Hence find Var (X). [3]

Solution

1

(i) $\mathrm{P}(X=2)=\mathrm{P}(A{*}{*}F{*},{*}A{*}{*}F)=2\left(\dfrac{2\times 3!}{5!}\right)=\dfrac{1}{5}=0.2$ (shown)

(ii)

$$\mathrm{P}(X=0)=\mathrm{P}(AF{*}{*}{*},{*}AF{*}{*},{*}{*}AF{*},{*}{*}{*}AF)=4\left(\frac{2\times 3!}{5!}\right)=\frac{2}{5}$$
$$=0.4$$

$$\mathrm{P}(X=1)=\mathrm{P}(A{*}F{*}{*},{*}A{*}F{*},{*}{*}A{*}F)=3\left(\frac{2\times 3!}{5!}\right)=\frac{3}{10}=0.3$$

$$\mathrm{P}(X=3)=\mathrm{P}(A{*}{*}{*}F)=\left(\frac{2\times 3!}{5!}\right)=\frac{1}{10}=0.1$$

x	0	1	2	3
$\mathrm{P}(X=x)$	0.4	0.3	0.2	0.1

(iii)

$$\mathrm{E}(X)=\sum_{all\ x}x\mathrm{P}(X=x)=0(0.4)+1(0.3)+2(0.2)+3(0.1)=1$$

$$\mathrm{E}(X-1)^2=\sum_{all\ x}(x-1)^2\mathrm{P}(X=x)=1(0.4)+0(0.3)+1(0.2)+4(0.1)=1$$

$$\mathrm{Var}(X)=\mathrm{E}(X-\mu)^2=\mathrm{E}(X-1)^2=1$$

Source of Question: JJC Prelim 9758/2017/02/Q5

2 The probability distribution of a discrete random variable, X, is shown below.

x	1	2
$\mathrm{P}(X=x)$	a	b

Find $\mathrm{E}(X)$ and $\mathrm{Var}(X)$ in terms of a. [5]

Solution:

2 $b = 1-a$

x	1	2
$\mathrm{P}(X=x)$	a	$1-a$

$$\begin{aligned}
\mathrm{E}(X) &= 1(a)+2(1-a) \\
&= \underline{\underline{2-a}} \\
\mathrm{E}(X^2) &= 1^2(a)+2^2(1-a) \\
&= 4-3a \\
\mathrm{Var}(X) &= \mathrm{E}(X^2)-[\mathrm{E}(X)]^2 \\
&= 4-3a-(2-a)^2 \\
&= 4-3a-(4-4a+a^2) \\
&= \underline{\underline{a-a^2}}
\end{aligned}$$

Source of Question: MJC Prelim 9758/2017/02/Q6

3 The probability function of X is given by

$$\mathrm{P}(X=x)=\begin{cases}(2x-1)\theta & \text{if } x=1,2,3\\ k & \text{if } x=4\\ 0 & \text{otherwise}\end{cases}$$

where $0<\theta<\frac{1}{9}$.

(i) Show that $k=1-9\theta$. Find, in terms of θ , the probability distribution of X. [2]

(ii) Find $\mathrm{E}(X)$ in terms of θ and hence show that $\mathrm{Var}(X)=26\theta-196\theta^2$. [3]

(iii) The random variable Y is related to X by the formula $Y=a+bX$, where a and b are non-zero constants. Given that $\mathrm{Var}(Y)=\frac{1}{3}b^2$, find the value of θ . [3]

Solution:

Solution:

(i)

x	1	2	3	4
$\mathrm{P}(X=x)$	θ	3θ	5θ	k

Since $\sum_{\text{all } x}\mathrm{P}(X=x)=1$,

$\theta+3\theta+5\theta+k=1$

$\therefore k=1-9\theta$

Probability distribution of X is

x	1	2	3	4
$\mathrm{P}(X=x)$	θ	3θ	5θ	$1-9\theta$

(ii)

$$\begin{aligned}\mathrm{E}(X)&=1(\theta)+2(3\theta)+3(5\theta)+4(1-9\theta)\\&=\theta+6\theta+15\theta+4-36\theta\\&=4-14\theta\end{aligned}$$

$$\begin{aligned}\mathrm{E}(X^2)&=1^2(\theta)+2^2(3\theta)+3^2(5\theta)+4^2(1-9\theta)\\&=\theta+12\theta+45\theta+16-144\theta\\&=16-86\theta\end{aligned}$$

$$\begin{aligned}\mathrm{Var}(X) &= \mathrm{E}(X^2)-[\mathrm{E}(X)]^2\\ &=16-86\theta-(4-14\theta)^2\\ &=16-86\theta-(16-112\theta+196\theta^2)\\ &=26\theta-196\theta^2\end{aligned}$$

(iii)

$Y = a+bX$

$\mathrm{Var}(Y)=\mathrm{Var}(a+bX)$

$\mathrm{Var}(Y)=b^2\mathrm{Var}(X)$

$\frac{1}{3}b^2 = b^2(26\theta-196\theta^2)$

$196\theta^2-26\theta+\frac{1}{3}=0 \quad (\because \mathrm{b}\neq 0)$

Using GC,

$\theta = 0.0144$ or $\theta = 0.118$ (rejected $\because 0<\theta<\frac{1}{9}$)

Source of Question: AJC Prelim 9758/2017/02/Q10

4 Males and females visiting an amusement park have heights, in centimetres, which are normally distributed with means and standard deviations as shown in the following table:

	Mean (cm)	Standard deviation (cm)
Male	165	12
Female	155	σ

It is found that 38.29% of the females have heights between 150 cm and 160 cm.

(i) Show that $\sigma = 10.0$ cm, correct to 3 significant figures. [2]

(ii) Find the probability that the height of a randomly chosen female is within 20 cm of three-quarter the height of a randomly chosen male. State an assumption that is necessary for the calculation to be valid. [4]

The amount, $*X*, a visitor has to pay for a popular ride in the park is $10 if the visitor's height is at least 120 cm but less than 150 cm, and $*m* if the visitor's height is 150 cm and above. If the visitor's height is less than 120 cm, he/she does not need to pay for the ride.

(iii) Assuming that a visitor purchasing a ticket for the ride is equally likely to be a male or female, find in terms of *m*, the probability distribution of *X*. [3]

Given that the expected amount a visitor will pay for a ride is $17.93, show that $m = 20.00$, correct to 2 decimal places. [1]

(iv) Three visitors were randomly chosen. Find the probability that the total amount they paid for a ride together is more than $40. [3]

Solution:

Let M denote the random variable: Height of a male visitor in cm. $M \sim \mathrm{N}(165, 12^2)$

Let F denote the random variable: Height of a female visitor in cm. $F \sim \mathrm{N}(155, \sigma^2)$

$$\mathrm{P}(150 < F < 160) = 0.3829$$

$$\mathrm{P}(-\frac{5}{\sigma} < Z < \frac{5}{\sigma}) = 0.3829$$

$$\mathrm{P}(Z < -\frac{5}{\sigma}) = \frac{1-0.3829}{2} = 0.30855$$

From G.C. $-\frac{5}{\sigma} = -\ 0.4999646$

$\Rightarrow \sigma = 10.0$ cm (3 sig. fig.)

<table>
<tr><td>(ii)</td><td>

$\frac{3}{4}M - F \sim N(\frac{3}{4}\times 165 - 155, \frac{9}{16}\times 12^2 + 10^2) = N(-31.25, 181)$

$P(\left|\frac{3}{4}M - F\right| \le 20) = P(-20 \le \frac{3}{4}M - F \le 20)$

$= 0.201 \quad (3 \text{ sig. fig.})$

Assumption: The heights of all male and female visitors are independent of one another.

</td></tr>
<tr><td>(iii)</td><td>

Probability Distribution of X:

<table>
<tr><th>x (in \$)</th><th>P(X = x)</th></tr>
<tr><td>0</td><td>$\frac{1}{2}P(M < 120) + \frac{1}{2}P(F < 120) = 0.00016056$</td></tr>
<tr><td>10</td><td>$\frac{1}{2}P(120 \le M < 150) + \frac{1}{2}P(120 \le F < 150) = 0.20693$</td></tr>
<tr><td>m</td><td>$\frac{1}{2}P(M \ge 150) + \frac{1}{2}P(F \ge 150) = 0.79291$</td></tr>
</table>

Given E(X) = 17.93 = 0 (0.00016056)+10(0.20693)+m(0.79291)
$\Rightarrow$ m = 20.00 (shown)

</td></tr>
<tr><td>(iv)</td><td>

P(X_1+ X_2 + X_3 > 40) = P(20, 20, 20) + 3. P(20, 20, 10)
= 0.889

</td></tr>
</table>

Source of Question: HCI Prelim 9758/2017/02/Q5

5 The independent random variables X and Y are normally distributed with the same mean 7 but different variances $\text{Var}(X)$ and $\text{Var}(Y)$, respectively. It is given that $\text{P}(X < 10) = \text{P}(Y > 6)$.

(i) Show that $\text{Var}\left(X\right) = 9\text{Var}\left(Y\right)$. [3]

(ii) If $\text{Var}\left(Y\right) = 1$, find $\text{P}(X < 9)$. [2]

Solution:

(i)

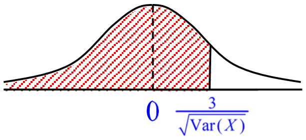
$X \sim \text{N}\left(7, \text{Var}(X)\right)$

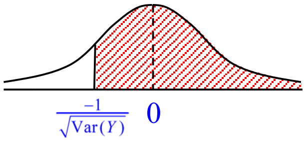
$Y \sim \text{N}\left(7, \text{Var}(Y)\right)$

$\text{P}(X < 10) = \text{P}(Y > 6)$

$$\text{P}\left(Z < \frac{10-7}{\sqrt{\text{Var}(X)}}\right) = \text{P}\left(Z > \frac{6-7}{\sqrt{\text{Var}(Y)}}\right)$$

$$\text{P}\left(Z < \frac{3}{\sqrt{\text{Var}(X)}}\right) = \text{P}\left(Z > \frac{-1}{\sqrt{\text{Var}(Y)}}\right)$$





$$\therefore \frac{3}{\sqrt{\text{Var}(X)}} = -\left(\frac{-1}{\sqrt{\text{Var}(Y)}}\right)$$

$$3\sqrt{\text{Var}(Y)} = \sqrt{\text{Var}(X)}$$

Hence $\text{Var}(X) = 9\text{Var}(Y)$ (shown)

(ii)

$\text{Var}(X) = 9(1) = 9$

$X \sim \text{N}\left(7, 9\right)$

$\therefore \text{P}(X < 9) = 0.748$

Source of Question: AJC Prelim 9758/2017/02/Q10

6 The number of days of gestation for a Dutch Belted cow is normally distributed, with a mean of μ days and a standard deviation of σ days. 8.08% of this cattle breed has a gestation period shorter than 278 days whereas 21.2% has a gestation period longer than 289 days. Find the values of μ and σ, giving your answers correct to 3 significant figures. [3]

(i) Find the probability that the mean gestation period for thirty-two randomly chosen Dutch Belted cows is more than 287 days. State a necessary assumption for your calculation to be valid. [3]

For another cattle breed, the Jersey cow, the number of days of gestation is normally distributed with a mean of 278 days and a standard deviation of 2.5 days. During gestation, a randomly chosen pregnant Dutch Belted cow eats 29 kg of feed daily while a randomly chosen pregnant Jersey cow eats 26 kg of feed daily.

(ii) Find the value of *a* such that during their respective gestation periods, there is a probability of 0.35 that the amount of feed consumed by a randomly chosen pregnant Jersey cow exceeds half of the amount consumed by a randomly chosen pregnant Dutch Belted cow by less than *a* kg. Express your answer to the nearest kg. [2]

(iii) Calculate the probability that during their respective gestation periods, the difference between the amount of feed consumed by three randomly chosen pregnant Dutch Belted cows and four randomly chosen pregnant Jersey cows is more than 4000 kg. State clearly the parameters of the distribution used in the calculation. [3]

Solution:

	Let *D* be the random variable denoting the number of days of gestation for a Dutch Belted cow. $\text{P}(D < 278) = 0.0808$ $\text{P}(Z < \frac{278-\mu}{\sigma}) = 0.0808$ $\frac{278-\mu}{\sigma} = -1.39971 - - - (1)$ $\text{P}(D > 289) = 0.212$ $\text{P}(Z < \frac{289-\mu}{\sigma}) = 0.788$ $\frac{289-\mu}{\sigma} = 0.799501 - - - (2)$ Solving (1)&(2), $\mu = 285.001064$ and $\sigma = 5.0017961$ $\mu = 285$ (3 s.f.) and $\sigma = 5.00$ (3 s.f.) . **(i)**

$\overline{D} \sim \mathrm{N}(285.001064, \frac{5.0017961^2}{32})$.

$\mathrm{P}(\overline{D} > 287) = 0.0118629 = 0.0119$ (3 s.f.)

The number of days of gestation for a Dutch Belted cow is independent of the number of days of gestation of another Dutch Belted cow.

(ii)

$J \sim \mathrm{N}(278, 2.50^2)$

$D \sim \mathrm{N}(285.001064, 5.0017961^2)$

Let $X = 26J - \frac{1}{2}29D$

$X \sim \mathrm{N}(3095.48457\ ,\ 9485.02698)$

$\mathrm{P}(0 < X < a) = 0.35$

$\therefore a = 3057.95778 \approx 3058$

(iii)

$D \sim \mathrm{N}(285.001064, 5.0017961^2)$

$J \sim \mathrm{N}(278, 2.50^2)$

Let C_1 denote the random variable of the amount of feed cosumed by 3 pregnant Dutch belted cows.

Let C_2 denote the random variable of the amount of feed cosumed by 4 pregnant Jersey cows.

$C_1 = 29(D_1 + D_2 + D_3) \sim \mathrm{N}(24795.09257, 63120.32374)$

$C_2 = 26(\mathrm{J}_1 + J_2 + J_3 + J_4) \sim \mathrm{N}(28912, 16900)$

$C_1 - C_2 \sim \mathrm{N}(-4117, 80020.3237)$

$$\begin{aligned}
&\mathrm{P}(|C_1 - C_2| > 4000) \\
&\quad = \mathrm{P}(C_1 - C_2 < -4000) + \mathrm{P}(C_1 - C_2 > 4000) \\
&\quad = 0.6604182314 \\
&\quad = 0.660 \text{ (3 s.f.)}
\end{aligned}$$

Or

$D \sim \mathrm{N}(285.00, 5.0018^2)$

$J \sim \mathrm{N}(278, 2.50^2)$

Let C_1 denote the random variable of the amount of feed cosumed by 3 pregnant Dutch belted cows.

Let C_2 denote the random variable of the amount of feed cosumed by 4 pregnant Jersey cows.

$C_1 = 29(D_1 + D_2 + D_3) \sim \mathrm{N}(24795, 63120.42217)$

$C_2 = 26(\mathrm{J}_1 + J_2 + J_3 + J_4) \sim \mathrm{N}(28912, 16900)$

$C_1 - C_2 \sim \mathrm{N}(-4117, 80020.422)$

$\mathrm{P}(|C_1 - C_2| > 4000)$

$= \mathrm{P}(C_1 - C_2 < -4000) + \mathrm{P}(C_1 - C_2 > 4000)$

$= 0.66041814$

$= 0.660$ (3 s.f.)

Source of Question: PJC Prelim 9758/2017/02/Q7

7 The masses, in kilograms, of black sea bass fish and red tilapia fish sold in a supermarket are normally distributed. The means and standard deviations of these distributions, and the selling prices, in \$ per kilogram, are given in the following table.

	Mean Mass (kg)	Standard Deviation (kg)	Selling Price (\$ per kg)
Black sea bass fish	1.10	0.20	12
Red tilapia fish	0.55	0.05	9

(i) Ayden bought 2 black sea bass fish and 3 red tilapia fish. Find the probability that he pays more than \$40. State an assumption needed in your calculation. [4]

(ii) Five red tilapia fish are randomly chosen. Find the probability that the fifth red tilapia fish is the third red tilapia fish weighing less than half a kilogram. [3]

Solution:

(i)

$X \sim$ mass of a black sea bass fish. $X \sim \mathrm{N}\left(1.1, 0.2^2\right)$

$Y \sim$ mass of a red tilapia fish. $Y \sim \mathrm{N}\left(0.55, 0.05^2\right)$

Let T be the total cost of 2 black sea bass and 3 red tilapia. Then

$T = 12\left(X_1 + X_2\right) + 9\left(Y_1 + Y_2 + Y_3\right)$

$\mathrm{E}\left(T\right) = (12)(2)\mathrm{E}(X) + (9)(3)\mathrm{E}(Y)$

$= 26.4 + 14.85$

$= 41.25$

$\mathrm{Var}\left(T\right) = (12)^2(2)\mathrm{Var}(X) + (9)^2(3)\mathrm{Var}(Y)$

$= 11.52 + 0.6075$

$= 12.1275$

Thus $T \sim \mathrm{N}\left(41.25\ ,\ 12.1275\right)$.

$\mathrm{P}(T > 40) = 0.64018 \approx 0.640$ (3 s.f.)

An assumption needed is the price / mass of all fish are independent of one another.

(ii)

Probability required $= \dfrac{4!}{2!2!}\ [\mathrm{P}\ (Y > 0.5)]^2\ [\mathrm{P}\ (Y < 0.5)]^3 \approx 0.0170$

Source of Question: DHS Prelim 9758/2018/02/Q6

8 A test consists of 15 multiple choice questions, where each question has n possible options, of which only one is correct. A student took the test by randomly choosing an answer to each question. It is known that the probability of answering exactly 3 questions correctly is the same as the probability of answering exactly 4 questions correctly.

(i) By forming an equation in terms of *n*, find the value of *n*. [3]

Each correct answer is awarded 3 marks and each incorrect answer carries a penalty of 1 mark. The score is the total marks awarded based on the number of correct and incorrect answers.

(ii) Find the expected score, *s*, obtained by the student. [3]

(iii) Find the probability that the score obtained by the student is within 4 marks of *s*. [2]

Solution:

8(i)

Let *X* be the random variable denoting the number of questions answered correctly out of 15.

$X \sim \mathrm{B}\left(15, \frac{1}{n}\right)$

$\mathrm{P}(X = 3) = \mathrm{P}(X = 4)$

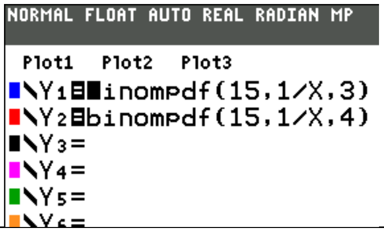
$\binom{15}{3}\left(\frac{1}{n}\right)^3\left(\frac{n-1}{n}\right)^{12} = \binom{15}{4}\left(\frac{1}{n}\right)^4\left(\frac{n-1}{n}\right)^{11}$

$\left(\frac{n-1}{n}\right) = 3\left(\frac{1}{n}\right)$

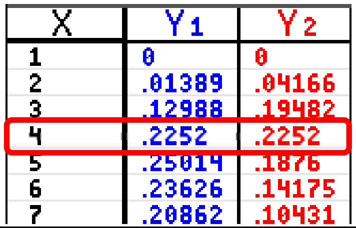
$n = 4$

<u>Using your GC to check the answer</u>

From GC, $n = 4$



```
NORMAL FLOAT AUTO REAL RADIAN MP
 Plot1  Plot2  Plot3
■\Y1■binompdf(15,1/X,3)
■\Y2■binompdf(15,1/X,4)
■\Y3=
■\Y4=
■\Y5=
■\Y6=
```



X	Y1	Y2
1	0	0
2	.01389	.04166
3	.12988	.19482
4	.2252	.2252
5	.25014	.1876
6	.23626	.14175
7	.20862	.10431

(ii)

$X \sim \mathrm{B}\left(15, \frac{1}{4}\right)$

<u>**Method 1**</u>

Let *T* be the score for 15 questions.

Then $T = 3X - (15 - X) = 4X - 15$

$$s = \mathrm{E}(T) = 4\mathrm{E}(X) - 15$$
$$= 4\left[15\left(\frac{1}{4}\right)\right] - 15 = 0$$

Method 2

Let A be the score for 1 question. $\mathrm{E}(A) = 3\left(\frac{1}{4}\right) - 1\left(\frac{3}{4}\right) = 0$

$$s = \mathrm{E}(T) = \mathrm{E}(A_1 + A_2 + \ldots + A_{15}) = 15\mathrm{E}(A) = 0$$

Method 3

Let Y be the random variable denoting the number of questions answered incorrectly out of 15.

$Y \sim \mathrm{B}\left(15, \frac{3}{4}\right)$.

$$s = 3\mathrm{E}(X) - \mathrm{E}(Y)$$
$$= 3\left[15\left(\frac{1}{4}\right)\right] - \left[15\left(\frac{3}{4}\right)\right] = 0$$

(iii) $\mathrm{P}(-4 < T < 4) = \mathrm{P}(-4 < 4X - 15 < 4)$

$= \mathrm{P}(2.75 < X < 4.75)$

$= \mathrm{P}(3 \le X \le 4)$

$= \mathrm{P}(X \le 4) - \mathrm{P}(X \le 2)$ or $2\mathrm{P}(X = 4)$ or $2\mathrm{P}(X = 3)$

$= 0.450$

Source of Question: EJC Prelim 9758/2018/02/Q7

9 When Mr. Lee sends a text message to any of his students over the weekend, he gets a reply, on average, about 6 out of 10 times. On a particular weekend, Mr. Lee sends a text message to 25 students.

(i) State, in the context of this question, two assumptions needed to model the number of students that reply by a binomial distribution. [2]

(ii) Explain why one of the assumptions stated in part (i) may not hold in this context. [1]

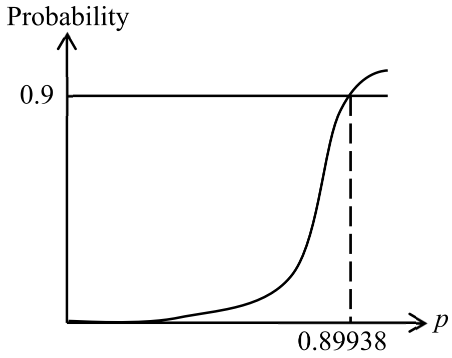
Assume now that these assumptions do in fact hold. Stating the distribution clearly,

(iii) find the probability that at least half of the students reply. [3]

It is given instead that the probability of a randomly chosen student replying is p. Find the least value of p such that there is at least a 90% chance of more than 20 students replying to the text message. [3]

Solution:

9	(i) A student replying has a constant probability of 0.6. The event of a student replying is independent of another student. OR Whether a student replies is independent of another student.
	(ii) Some students are more likely to reply a message. Hence the assumption of equal probability might not hold.
	(iii) Let X be the number of students that reply, out of 25. $X \sim \mathrm{B}(25, 0.6)$ Required probability $= \mathrm{P}(X \geq 13)$ $= 1 - \mathrm{P}(X \leq 12)$ $= 0.84623 \quad (5\mathrm{sf})$ $= 0.846 \quad (3\mathrm{sf})$
	Let Y be the number of students that reply, out of 25. Let p be the new probability of a student replying. $Y \sim \mathrm{B}(25, p)$ Given that $\mathrm{P}(Y > 20) \geq 0.9 \quad \left(\text{or } \mathrm{P}(Y \leq 20) \leq 0.1\right)$

	 Therefore, the least probability of a student replying is 0.899 [accept 0.900].

Source of Question: IJC Prelim 9758/2018/02/Q5

10 The random variable X has the distribution $\mathrm{B}(25, p)$, where $0 < p < 1$. Given that $\mathrm{P}(X \le 1) = 0.15$, write down an equation for the value of p and find this value numerically. Hence find $\mathrm{Var}(X)$. [4]

Solution:

10	$\mathrm{P}(X \le 1) = 0.15$ $\binom{25}{0}(1-p)^{25} + \binom{25}{1}p(1-p)^{24} = 0.15$ $(1-p)^{25} + 25p(1-p)^{24} = 0.15$ $\left[\text{or } (1-p)^{24}(1+24p) = 0.15\right]$ Using GC, $p = 0.12865$ $= 0.129$ (to 3 s.f.) $\mathrm{Var}(X) = np(1-p)$ $= 25(0.12865)(1-0.12865)$ $= 2.80$ (to 3 s.f.)

Source of Question: NYJC Prelim 9758/2018/02/Q6

11 A game is played using a fair six-sided die, a pawn and a simple board as shown below.

S	1	2	3	4	5	**E**

Initially, the pawn is placed on square **S**. The game is played by throwing the die and moving the pawn in the following manner:

S 1 2 3 4 5 **E** 5 4 3 2 1 2 3 4 5 **E**………

Thus, for example, if the first and second throw of the die gives a "5" and "4" respectively, the final position of the pawn will be on square "3".

The game will stop when the pawn stops at square **E**.

Let X be the random variable denoting the number of throws of the die required to move the pawn such that it stops at square **E**.

(i) Show $\mathrm{P}(X=2)=\frac{5}{36}$. [1]

(ii) Find the probability that more than two throws of the die are needed for the pawn to stop at square **E** given that the first throw of the die gives an even number. [3]

It is now given that for each game, a player has a maximum of 3 throws of the die and a special prize is given to any player who uses not more than two throws for the pawn to stop at square **E**.

(iii) Find the probability of a player winning a special prize in at least three but not more than eight games out of ten games. [3]

(iv) Find the least number of games needed so that the probability of winning at least a special prize is at least 0.998. [3]

Solution:

11(i)	Let Di be the random variable denoting the score of the ith throw of the die for $i = 1, 2$ $\mathrm{P}(X=2)$ $= \mathrm{P}(D_1=1,D_2=5)+\mathrm{P}(D_1=2,D_2=4)+\mathrm{P}(D_1=3,D_2=3)$ $+\mathrm{P}(D_1=4,D_2=2)+\mathrm{P}(D_1=5,D_2=1)$ $=\frac{1}{6}\left(\frac{1}{6}\right)5=\frac{5}{36}$
11(ii)	$\mathrm{P}(X>2 \mid D_1 \text{ is even})$ $=\dfrac{\mathrm{P}(X>2 \text{ and } D_1 \text{ is even})}{\mathrm{P}(D_1 \text{ is even})}$ $=\dfrac{\mathrm{P}(D_1=2,D_2\neq 4)+\mathrm{P}(D_1=4,D_2\neq 2)}{\mathrm{P}(D_1 \text{ is even})}$ $=\dfrac{\frac{1}{6}\left(\frac{5}{6}\right)+\frac{1}{6}\left(\frac{5}{6}\right)}{\frac{3}{6}}=\frac{5}{9}$
11(iii)	Let W be the random variable denoting number of games, out of 10, that a special prize is won. P(winning a special prize)

$= \mathrm{P}(X=1)+\mathrm{P}(X=2)=\frac{1}{6}+\frac{5}{36}=\frac{11}{36}$

$W \sim \mathrm{B}\left(10, \frac{11}{36}\right)$

Probability required

$= \mathrm{P}(3 \le W \le 8) = \mathrm{P}(W \le 8) - \mathrm{P}(W \le 2)$

$= 0.63173 = 0.632$

11(iv)

Let n be the number of games needed.

Let Y be the random variable denoting number of games, out of n, that a special prize is won.

$Y \sim \mathrm{B}\left(n, \frac{11}{36}\right)$

$\mathrm{P}(Y \ge 1) \ge 0.998 \Rightarrow 1-\mathrm{P}(Y=0) \ge 0.998$

Using GC,

n	$1-\mathrm{P}(Y=0)$	$1-\mathrm{P}(Y=0)-0.998$
17	0.99797	-3×10^{-5}
18	0.99859	5.9×10^{-4}

Least n = 18

12 **A graphing calculator is not to be used for this question.**

Joel the student lives at Holland Village and he is considered late if he reaches the school parade square after 0740. He catches the 0630 bus to school with probability p, or the 0650 train with probability $1 - p$. Measured in minutes, the journey time for the bus follows a normal distribution with mean 40 and standard deviation 12. The journey time for the train follows another normal distribution with mean 30 and standard deviation 4.

Journey times from the bus-stop where he alights to the parade square in school follows a normal distribution with mean 18 and standard deviation 5. The journey time from the train station where he alights to the parade square in school also follows a normal distribution with mean 10 and standard deviation 3.

Show that Joel is more likely to be late for school if he takes the bus, writing down any assumptions necessary in your calculations.

[6]

Let A be the probability Joel is late if he takes the bus, and B the probability he is late if the takes the train. Over a period of M days, where M is large, Joel noted that in 60% of the occasions when he was late, he had taken the bus. Obtain an estimate of p in terms of A and B. [3]

Solution

12	Let X and Y be the time Joel takes to reach the parade square in school if he takes the bus and the train respectively, in minutes. Assuming independence of all travel times by bus, train or foot, we have $X \sim \mathrm{N}(58,13^2)$ and $Y \sim \mathrm{N}(40,5^2)$. Thus probability Joel is late if he takes the bus is $\mathrm{P}(X>70)=\mathrm{P}\left(Z>\frac{70-58}{13}\right)=\mathrm{P}\left(Z>\frac{12}{13}\right)$. The probability that Joel is late if he takes the train is $\mathrm{P}(X>50)=\mathrm{P}\left(Z>\frac{50-40}{5}\right)=\mathrm{P}\left(Z>2\right)$. Now since $2>\frac{12}{13}$, $\mathrm{P}\left(Z>\frac{12}{13}\right)>\mathrm{P}\left(Z>2\right)$ so Joel is more likely to be late if he takes the bus.
	$\mathrm{P}(\text{taken bus} \mid \text{late})=\frac{\mathrm{P}(\text{taken bus and late})}{\mathrm{P}(\text{late})}$ $=\frac{Ap}{Ap+B(1-p)}=\frac{3}{5}$ $\Rightarrow 5Ap=3Ap+3B(1-p)$ $\Rightarrow (2A+3B)p=3B$ $\Rightarrow p=\frac{3B}{2A+3B}$

13 3 bar magnets are placed randomly end-to-end in a straight line. If adjacent magnets have ends of opposite polarities facing each other, they join together to form a single unit. If they have ends of the same polarity facing each other, they stand apart. Let L be the number of separate units.

(i) Show that $\mathrm{P}(L=1)=\mathrm{P}(L=3)=\frac{1}{4}$. [2]

(ii) Hence find the exact values of E(L) and Var(L). [4]

Instead of just 3 bar magnets, there are now 10 bar magnets placed randomly end-to-end in a straight line. We say there is a change in polarity when ends of opposite polarities face each other. Let S be the total number of changes in polarities.

(iii) State a necessary assumption for S to follow a binomial distribution. Assuming the condition holds, state the parameters of S. [3]

(iv) Hence find the expected number and variance of the number of separate units. [3]

Solution

13 **(i)**	For there to be 1 unit, all 3 magnets must be joined together. Thus we must have either all 3 magnets face NS or SN. Thus $\mathrm{P}(L=1)=\frac{2}{2^3}=\frac{1}{4}.$ For there to be 3 units, all 3 magnets have to be separate. Thus either we have (NS)(SN)(NS) or (SN)(NS)(SN). Thus $\mathrm{P}(L=3)=\frac{2}{2^3}=\frac{1}{4}.$
(ii)	So we have $\mathrm{P}(L=2)=1-\frac{1}{4}-\frac{1}{4}=\frac{1}{2}.$ So $\mathrm{E}(L)=(1)\frac{1}{4}+(2)\frac{1}{2}+(3)\frac{1}{4}=2.$ Similarly, $\mathrm{E}(L^2)=(1^2)\frac{1}{4}+(2^2)\frac{1}{2}+(3^2)\frac{1}{4}=4.5$ $\Rightarrow \mathrm{Var}(L)=\mathrm{E}(L^2)-\mathrm{E}(L)=0.5$
(iii)	$S\sim \mathrm{B}(9,0.5)$ Here we assume that the event there is a change in polarity between 2 magnets is independent of the event there is a change in polarity between another 2 magnets. Note that assuming that either 1) There are only 2 possible outcomes (change or no change) OR 2) there is a constant probability of change of polarity is not valid here as they are inherent in the question. For the latter, due to the magnets being placed randomly, between any 2 magnets, there are only 4 possible outcomes: NN, SS, SN, NS, of which 2 have a change in polarity.
(iv)	If we consider T to be the number of non-changes in polarity, we have $T+1=L$, but we have clearly $S+T=9$, which implies $L=9-S+1=10-S$ Thus $\mathrm{E}(L)=10-\mathrm{E}(S)=10-9(0.5)=5.5$ and $\mathrm{Var}(L)=\mathrm{Var}(S)=9(0.5)(0.5)=2.25$.