1. Two male ladybugs are sitting at opposite corners of a large window composed of individual square glass panes, as shown. At the same instant, they start walking at the same rate along the pane edges, one edge at a time, in the hope of eventually reaching the female sitting at the center of the window. At the start of every “move” along a single edge, each male randomly – and independently – selects only either the horizontal or vertical direction shown, each with the probability indicated. However, the female is only willing to wait for each male to complete a total of $n = 5$ horizontal and vertical moves along the edges, before she becomes impatient and flies away. Given this information, answer each of the following.

[Hint: Let the random variable $X_1 =$ “number of ‘→’ moves in $n = 5$ independent moves by Male 1,” and $X_2 =$ “number of ‘←’ moves in $n = 5$ independent moves by Male 2.”]

In order for either male to reach the female in $n = 5$ moves, he must make $X = 3$ horizontal moves and $5 − X = 2$ vertical moves, in any combination.

Let event $M_1 =$ “$X_1 = 3$” = “Male 1 reaches the female in $n = 5$ moves,” and likewise, let event $M_2 =$ “$X_2 = 3$” = “Male 2 reaches the female in $n = 5$ moves.”

(a) Calculate the probability that Male 1 meets the female in $n = 5$ moves. (5 pts)

$$P(M_1) = P(X_1 = 3) = \binom{5}{3} 0.6^3 0.4^2 = 0.3456$$

(b) Calculate the probability that Male 2 meets the female in $n = 5$ moves. (5 pts)

$$P(M_2) = P(X_2 = 3) = \binom{5}{3} 0.4^3 0.6^2 = 0.2304$$

(c) Calculate the probability that both males meet the female in $n = 5$ moves. (5 pts)

Via independence, $P(M_1 \cap M_2) = P(M_1)P(M_2) = 0.3456 \times 0.2304 = 0.0796$.

(d) Calculate the probability that neither male meets the female in $n = 5$ moves. (5 pts)

Via independence,

$$P(M_1^c \cap M_2^c) = P(M_1^c)P(M_2^c) = 1 − 0.3456 \times 1 − 0.2304 = 0.6544 \times 0.7696 = 0.5036.$$
(e) Calculate the probability that only Male 1 meets the female in $n = 5$ moves. (5 pts)

\[ P(M_1 \cap M_2^c) = P(M_1^c) P(M_2^c) = 0.3456 \times 1 - 0.2304 = 0.3456 	imes 0.7696 = 0.2660 \]

(f) Calculate the probability that only Male 2 meets the female in $n = 5$ moves. (5 pts)

\[ P(M_1^c \cap M_2) = P(M_1^c) P(M_2) = 1 - 0.3456 \times 0.2304 = 0.6544 \times 0.2304 = 0.1508 \]
2. (a) An experiment involving a fair deck of 52 cards is to be conducted. Cards will be randomly selected from the deck \textit{with replacement}. The random variable of interest is $X =$ “Number of trials until a ‘face card’ (i.e., Jack, Queen, or King) first appears.”

Can the resulting outcomes be considered a sequence of Bernoulli trials? Be specific!

\textbf{Yes!} Each outcome is independent of the others; since this is conducted \textit{with replacement}, the probability of “Success” (i.e., “face card”) does not change!

Which probability distribution is the most appropriate to model this experiment? Determine all parameter values. That is, $X \sim ?$

\textbf{Geometric distribution, with $P(\text{Success}) = \frac{12}{52}$}.

Calculate the probability that a face card will first appear on the fourth trial. Show all work…

\begin{align*}
p(4) &= P(X = 4) = \left( \frac{40}{52} \right)^3 \left( \frac{12}{52} \right) = 0.10504
\end{align*}

(b) A second experiment involving the same deck is to be conducted. Cards will be randomly selected from the deck, but \textit{without replacement}, until a “face card” (i.e., Jack, Queen, or King) first appears.

Can the resulting outcomes be considered a sequence of Bernoulli trials? Be specific!

\textbf{No!} Each outcome is not independent of the others; since this is conducted \textit{without replacement}, the probability of “Success” (i.e., “face card”) on each trial affects that of the next trial.

Calculate the probability that a face card will first appear on the fourth trial. Show all work…

\begin{align*}
\text{Now, } P(X = 4) &= \left( \frac{40}{52} \right) \left( \frac{39}{51} \right) \left( \frac{38}{50} \right) \left( \frac{12}{49} \right) = 0.10948
\end{align*}

(c) Suppose that in general, a population of finite size $N$ units contains $s$ “Successes.” Units are to randomly drawn from the population \textit{without replacement}.

Formally prove that the random variable $X =$ “Number of trials until the first Success appears” has pmf

\begin{align*}
p(x) &= P(X = x) = \binom{N - x}{s - 1} \binom{s}{x} \quad \text{for } x = 1, 2, 3, 4, \ldots
\end{align*}

\textbf{Hint:} Start by generalizing the procedure in (b), then apply some algebra.
Proof: For $x = 1, 2, 3, 4, \ldots$ the probability of achieving the first Success after $(x-1)$ Failures without replacement would be

\[ p(x) = \frac{(N-s)!}{N!/(N-s)!} \times \frac{(N-s-x+1)!}{(N-x)!} \times \frac{s}{(s-1)!/N!/(N-s)!} \]

\[ = \frac{(N-x)!}{(N-s-x+1)!} \times \frac{s}{N!/(N-s)!} \]

\[ = \frac{(N-x)!}{(N-s-x+1)!} \times \frac{s!}{(s-1)!/N!/(N-s)!} \]

\[ = \frac{(N-x)!}{(N-s-x+1)!} \times \frac{s!}{(s-1)!/N!/(N-s)!} \]

\[ = \left( \frac{N-x}{s-1} \right) \frac{1}{\binom{N}{s}} \times \frac{n!}{k!(n-k)!} \]

QED

(d) Use this formula to recalculate the probability in (b), and show agreement in your answers.

\[ N = 52, s = 12, x = 4: \quad p(4) = P(X = 4) = \frac{\binom{48}{11}}{\binom{52}{12}} = \frac{22595200368}{206379406870} = 0.10948, \text{ which agrees with (b).} \]
3.

(a) Starting from 0, a flea randomly jumps either one unit or three units to the right, with fixed probability $p$ or $1 - p$ respectively, as shown in the probability histogram below.* Also shown is the accompanying pmf chart, for the random variable $X = \text{“Number of units jumped.”}$

\[
\begin{array}{c|c}
 x & p(x) \\
\hline
1 & p \\
3 & 1 - p \\
\end{array}
\]

\[ E[X] = \sum x p(x) = 1(p) + 3(1 - p) = 3 - 2p \]

(b) Starting from 0, an ant crawls horizontally, and can randomly stop anywhere ($Y$) in the interval $[0, 4]$, with corresponding probability given by the piecewise uniform pdf $f(y) = \begin{cases} 
\frac{p}{2}, & 0 \leq y < 2 \\
1 - \frac{p}{2}, & 2 \leq y \leq 4 
\end{cases}$ (and 0 elsewhere), as illustrated below.*

* As usual, $0 < p < 1$. The illustration shown is for $0 < p < 0.5$ without loss of generality, for if $0.5 < p < 1$, the heights are just reversed.
Calculate the expected distance crawled, i.e., the mean of \( Y \), in terms of \( p \). Verify that this mean is the same as in (a), and give a brief explanation for why this should be. (It does not have to be mathematically rigorous.) Show all work! (5 pts)

\[
\mu_y = E[Y] = \int_{-\infty}^{\infty} y f(y) \, dy = \int_{0}^{2} y \left( \frac{p}{2} \right) \, dy + \int_{2}^{4} y \left( \frac{1-p}{2} \right) \, dy
\]

\[
= \left( \frac{p}{2} \right) \int_{0}^{2} y \, dy + \left( \frac{1-p}{2} \right) \int_{2}^{4} y \, dy
\]

\[
= \left( \frac{p}{2} \right) \left[ \frac{y^2}{2} \right]_{0}^{2} + \left( \frac{1-p}{2} \right) \left[ \frac{y^2}{2} \right]_{2}^{4}
\]

\[
= \frac{p}{2} \left[ \frac{2^3}{3} \right] + \frac{1-p}{2} \left[ \frac{4^3}{2} \right] = p + 3(1-p) = \frac{3}{2} - 2p, \text{ which agrees with } \mu_x \text{ in (a}).
\]

This is not surprising, since from a physical perspective, both systems have their total mass \( 1 \) distributed in exactly the same way. Hence the “balance point” must be the same for both.

Calculate the variance of \( Y \), in terms of \( p \). Express in simplest possible form. Show all work! (5 pts)

\[
\sigma_y^2 = E[Y^2] - (E[Y])^2 = \int_{-\infty}^{\infty} y^2 f(y) \, dy - \mu_y^2
\]

\[
= \int_{0}^{2} y^2 \left( \frac{p}{2} \right) \, dy + \int_{2}^{4} y^2 \left( \frac{1-p}{2} \right) \, dy - (3-2p)^2
\]

\[
= \frac{p}{2} \int_{0}^{2} y^2 \, dy + \frac{1-p}{2} \int_{2}^{4} y^2 \, dy - (3-2p)^2
\]

\[
= \frac{p}{2} \left[ \frac{y^3}{3} \right]_{0}^{2} + \frac{1-p}{2} \left[ \frac{y^3}{3} \right]_{2}^{4} - (3-2p)^2
\]

\[
= \frac{8}{3} \frac{p}{2} + \frac{28}{3} \frac{1-p}{2} - (3-2p)^2
\]

\[
= \frac{4}{3} p + \frac{28}{3} (1-p) - (9 - 12p + 4p^2)
\]

\[
= \frac{1}{3} + 4p - 4p^2 = \frac{1}{3} + 4p(1-p) \text{ (compare with } \sigma_x^2 \text{)}.
\]
Determine the cdf \( F(y) = P(Y \leq y) \), for all \( y \) in [0, 4], in terms of \( p \). Show all work! (10 pts)

- Clearly, since \( f(y) = 0 \) for all \( y < 0 \), it follows that \( F(y) = 0 \) there as well.
- Therefore, for any fixed but arbitrary \( 0 \leq y < 2 \), it follows that \( F(y) = 0 + \) the area under \( f(y) = \frac{p}{2} \) from 0 to \( y \). Since this is a simple rectangular area, it can be found directly, without the use of calculus: base \( \times \) height = \( (y-0)\left(\frac{p}{2}\right) \). Alternatively, \( F(y) = \int_0^y f(t) \, dt = \int_0^y \left(\frac{p}{2}\right) \, dt = \left[\frac{p}{2}t\right]_0^y = \left[\frac{p}{2}y\right] \). Note that \( F(0) = 0 \) and \( F(2) = p \), hence this is the equation of a straight line from (0, 0) to \((2, p)\), as shown below.

- Now for any fixed but arbitrary \( 2 \leq y \leq 4 \), it follows that \( F(y) = F(2) + \int_2^y f(t) \, dt = p + \left(1 - \frac{p}{2}\right)(y-2) \) or written more conventionally, \( \left(1 - \frac{p}{2}\right)y + (2p - 1) \). Note that \( F(2) = p \) and \( F(4) = 1 \), hence this is the equation of a straight line from \((2, p)\) to \((4, 1)\), as shown below.

- Finally, \( f(y) = 0 \) for all \( y > 4 \), therefore \( F(y) = 1 \) there.

Combining these into a single formal statement, the complete description appears below:

\[
F(y) = \begin{cases} 
0, & y < 0 \\
\frac{p}{2}y, & 0 \leq y < 2 \\
p + \left(1 - \frac{p}{2}\right)(y-2), & 2 \leq y \leq 4 \\
1, & y > 4 
\end{cases}
\]

Sketch a labeled graph of the cdf below, for \( 0 < p < 0.5 \). (5 pts)
(d) For all $0 < h < 2$, consider the interval $[2-h, 2+h]$, symmetric about $Y = 2$. Using the cdf found in (c), compute the probability $P(2-h \leq Y \leq 2+h)$ in terms of $h$, and confirm that it does not depend on the value of $p$. Show all work!

$$P(2-h \leq Y \leq 2+h) = P(Y \leq 2+h) - P(Y \leq 2-h)$$
$$= F(2+h) - F(2-h)$$
$$= \left[p + \left(\frac{1-p}{2}\right)h\right] - \left[\frac{p}{2}(2-h)\right]$$
$$= \frac{h}{2} \text{ for all } p.$$
4. The relation between the *continuous* variable “$X = \text{Foot Length (inches)}$” and the *discrete* variable “$Y = \text{Men’s Shoe Size}$” is shown in the two horizontal scales below. (For example, men’s shoe size 9 corresponds to foot length between $10\frac{2}{6}$ and $10\frac{3}{6}$ inches.) Suppose that in a certain population of men, $X$ is normally distributed with mean $\mu = 11$ inches, and standard deviation $\sigma = \frac{2}{3}$ inches, i.e., $X \sim N(11, \frac{2}{3})$, as illustrated.

(a) Calculate the probability that a randomly selected man has a shoe size of 12 or larger. Show all work!

$Y = \text{Men’s Shoe Size (discrete)}$; $X = \text{Foot Length (continuous)} \sim N(11, \frac{2}{3})$.

$P(Y \geq 12) = P(X \geq 11\frac{2}{6}) = P\left(Z \geq \frac{11\frac{2}{6} - 11}{\frac{2}{3}}\right) = P(Z \geq 0.5) = 1 - 0.69146 = \boxed{0.30854}$

(b) Calculate the probability that a randomly selected man has a shoe size larger than 12. Show all work!

$P(Y > 12) = P(Y \geq 12\frac{1}{2}) = P(X \geq 11\frac{1}{2}) = P\left(Z \geq \frac{11\frac{1}{2} - 11}{\frac{2}{3}}\right) = P(Z \geq 0.75) = 1 - 0.77337 = \boxed{0.22663}$

(c) Calculate the probability that a randomly selected man has a shoe size of exactly 12. Show all work!

$P(Y = 12) = P(Y \geq 12) - P(Y > 12) = 0.30854 - 0.22663 = \boxed{0.08191}$

(d) Calculate the probability that a randomly selected man has a shoe size of exactly 12, given that his shoe size is 12 or larger. Show all work!

$P(Y = 12 \mid Y \geq 12) = \frac{P(Y = 12 \cap Y \geq 12)}{P(Y \geq 12)} = \frac{0.08191}{0.30854} = \boxed{0.2658}$