1. Review of Main Concepts

(a) **Multivariate: Discrete to Continuous:**

<table>
<thead>
<tr>
<th>Joint PMF/PDF</th>
<th>Discrete</th>
<th>Continuous</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p_{X,Y}(x, y) )</td>
<td>( \mathbb{P}(X = x, Y = y) )</td>
<td>( f_{X,Y}(x, y) \neq \mathbb{P}(X = x, Y = y) )</td>
</tr>
<tr>
<td>Joint range/support</td>
<td>( {(x, y) \in \Omega_X \times \Omega_Y : p_{X,Y}(x, y) &gt; 0} )</td>
<td>( {(x, y) \in \Omega_X \times \Omega_Y : f_{X,Y}(x, y) &gt; 0} )</td>
</tr>
<tr>
<td>Joint CDF</td>
<td>( F_{X,Y}(x, y) = \sum_{t \leq x, s \leq y} p_{X,Y}(t, s) )</td>
<td>( F_{X,Y}(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(t, s) , ds , dt )</td>
</tr>
<tr>
<td>Normalization</td>
<td>( \sum_{x,y} p_{X,Y}(x,y) = 1 )</td>
<td>( \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x,y) , dx , dy = 1 )</td>
</tr>
<tr>
<td>Marginal PMF/PDF</td>
<td>( p_X(x) = \sum_{y} p_{X,Y}(x,y) )</td>
<td>( f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) , dy )</td>
</tr>
<tr>
<td>Expectation</td>
<td>( \mathbb{E}[g(X,Y)] = \sum_{x,y} g(x,y)p_{X,Y}(x,y) )</td>
<td>( \mathbb{E}[g(X,Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y)f_{X,Y}(x,y) , dx , dy )</td>
</tr>
<tr>
<td>Independence</td>
<td>( \forall x, y, p_{X,Y}(x,y) = p_X(x)p_Y(y) )</td>
<td>( \forall x, y, f_{X,Y}(x,y) = f_X(x)f_Y(y) )</td>
</tr>
<tr>
<td></td>
<td>( \Omega_{X,Y} = \Omega_X \times \Omega_Y )</td>
<td>( \Omega_{X,Y} = \Omega_X \times \Omega_Y )</td>
</tr>
</tbody>
</table>

(b) **Law of Total Probability (r.v. version):** If \( X \) is a discrete random variable, then
\[
\mathbb{P}(A) = \sum_{x \in \Omega_X} \mathbb{P}(A | X = x)p_X(x) \quad \text{discrete } X
\]

(c) **Law of Total Expectation (Event Version):** Let \( X \) be a discrete random variable, and let events \( A_1, ..., A_n \) partition the sample space. Then,
\[
\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X | A_i] \mathbb{P}(A_i)
\]

(d) **Conditional Expectation:** See table. Note that linearity of expectation still applies to conditional expectation: \( \mathbb{E}[X + Y | A] = \mathbb{E}[X | A] + \mathbb{E}[Y | A] \)

(e) **Law of Total Expectation (RV Version):** Suppose \( X \) and \( Y \) are random variables. Then,
\[
\mathbb{E}[X] = \sum_{y} \mathbb{E}[X | Y = y] p_Y(y) \quad \text{discrete version.}
\]

(f) **Conditional distributions (not covered in class)**

<table>
<thead>
<tr>
<th>Conditional PMF/PDF</th>
<th>Discrete</th>
<th>Continuous</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p_{X,Y}(x,y) )</td>
<td>( p_{X,Y}(x,y) )</td>
<td>( f_{X,Y}(x,y) )</td>
</tr>
<tr>
<td>( p_X(x) )</td>
<td>( p_X(x) )</td>
<td>( f_X(x) )</td>
</tr>
<tr>
<td>( p_Y(y) )</td>
<td>( p_Y(y) )</td>
<td>( f_Y(y) )</td>
</tr>
</tbody>
</table>

(g) The following have not been covered as of 11/19:

- Law of Total Probability (continuous)
\[
\mathbb{P}(A) = \int_{x \in \Omega_X} \mathbb{P}(A | X = x) f_X(x) \, dx
\]

- Law of total expectation (continuous)
\[
\mathbb{E}[X] = \int_{y \in \Omega_Y} \mathbb{E}[X | Y = y] f_Y(y) \, dy
\]
2. Joint PMF's
Suppose X and Y have the following joint PMF:

<table>
<thead>
<tr>
<th>X/Y</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0.2</td>
<td>0.1</td>
</tr>
<tr>
<td>1</td>
<td>0.3</td>
<td>0</td>
<td>0.4</td>
</tr>
</tbody>
</table>

(a) Identify the range of X ($\Omega_X$), the range of Y ($\Omega_Y$), and their joint range ($\Omega_{X,Y}$).

Solution:
$\Omega_X = \{0, 1\}$, $\Omega_Y = \{1, 2, 3\}$, and $\Omega_{X,Y} = \{(0, 2), (0, 3), (1, 1), (1, 3)\}$

(b) Find the marginal PMF for X, $p_X(x)$ for $x \in \Omega_X$.

Solution:

$p_X(0) = \sum_y p_{X,Y}(0, y) = 0 + 0.2 + 0.1 = 0.3$

$p_X(1) = 1 - p_X(0) = 0.7$

(c) Find the marginal PMF for Y, $p_Y(y)$ for $y \in \Omega_Y$.

Solution:

$p_Y(1) = \sum_x p_{X,Y}(x, 1) = 0 + 0.3 = 0.3$

$p_Y(2) = \sum_x p_{X,Y}(x, 2) = 0.2 + 0 = 0.2$

$p_Y(3) = \sum_x p_{X,Y}(x, 3) = 0.1 + 0.4 = 0.5$

(d) Are X and Y independent? Why or why not?

Solution:

No, since a necessary condition is that $\Omega_{X,Y} = \Omega_X \times \Omega_Y$.

(e) Find $\mathbb{E}[X^3Y]$.

Solution:

Note that $X^3 = X$ since $X$ takes values in $\{0, 1\}$.

$\mathbb{E}[X^3Y] = \mathbb{E}[XY] = \sum_{(x,y) \in \Omega_{X,Y}} xy p_{X,Y}(x, y) = 1 \cdot 1 \cdot 0.3 + 1 \cdot 3 \cdot 0.4 = 1.5$
3. Trinomial Distribution
A generalization of the Binomial model is when there is a sequence of \( n \) independent trials, but with three outcomes, where \( \mathbb{P}(\text{outcome } i) = p_i \) for \( i = 1, 2, 3 \) and of course \( p_1 + p_2 + p_3 = 1 \). Let \( X_i \) be the number of times outcome \( i \) occurred for \( i = 1, 2, 3 \), where \( X_1 + X_2 + X_3 = n \). Find the joint PMF \( p_{X_1,X_2,X_3}(x_1, x_2, x_3) \) and specify its value for all \( x_1, x_2, x_3 \in \mathbb{R} \).

**Solution:**

Same argument as for the binomial PMF:

\[
p_{X_1,X_2,X_3}(x_1, x_2, x_3) = \begin{pmatrix} n \\ x_1, x_2, x_3 \end{pmatrix} \prod_{i=1}^{3} p_i^{x_i} = \frac{n!}{x_1!x_2!x_3!} p_1^{x_1} p_2^{x_2} p_3^{x_3}
\]

where \( x_1 + x_2 + x_3 = n \) and are nonnegative integers.

4. Do You “Urn” to Learn More About Probability?
Suppose that 3 balls are chosen without replacement from an urn consisting of 5 white and 8 red balls. Let \( X_i = 1 \) if the \( i \)-th ball selected is white and let it be equal to 0 otherwise. Give the joint probability mass function of

(a) \( X_1, X_2 \)

**Solution:**

Here is one way of defining the joint pmf of \( X_1, X_2 \)

\[
\mathbb{P}(X_1 = 1, X_2 = 1) = \mathbb{P}(X_1 = 1)\mathbb{P}(X_2 = 1 \mid X_1 = 1) = \frac{5}{13} \cdot \frac{4}{12} = \frac{20}{156}
\]

\[
\mathbb{P}(X_1 = 1, X_2 = 0) = \mathbb{P}(X_1 = 1)\mathbb{P}(X_2 = 0 \mid X_1 = 1) = \frac{5}{13} \cdot \frac{8}{12} = \frac{40}{156}
\]

\[
\mathbb{P}(X_1 = 0, X_2 = 1) = \mathbb{P}(X_1 = 0)\mathbb{P}(X_2 = 1 \mid X_1 = 0) = \frac{8}{13} \cdot \frac{5}{12} = \frac{40}{156}
\]

\[
\mathbb{P}(X_1 = 0, X_2 = 0) = \mathbb{P}(X_1 = 0)\mathbb{P}(X_2 = 0 \mid X_1 = 0) = \frac{8}{13} \cdot \frac{7}{12} = \frac{56}{156}
\]

(b) \( X_1, X_2, X_3 \)

**Solution:**

Instead of listing out all the individual probabilities, we could write a more compact formula for the pmf. In this problem, the denominator is always \( P(13, k) \), where \( k \) is the number of random variables in the joint pmf. And the numerator is \( P(5, i) \) times \( P(8, j) \) where \( i \) and \( j \) are the number of 1s and 0s, respectively. In this case the number of 1s is \( x_1 + x_2 + x_3 \), and the number of 0s is \((1 - x_1) + (1 - x_2) + (1 - x_3)\). Then, we can write the pmf as follows:

\[
p_{X_1,X_2,X_3}(x_1, x_2, x_3) = \frac{10!}{13!} \cdot \frac{5!}{(5 - x_1 - x_2 - x_3)!} \cdot \frac{8!}{(5 + x_1 + x_2 + x_3)!}
\]

5. Successes
Consider a sequence of independent Bernoulli trials, each of which is a success with probability \( p \). Let \( X_1 \) be the number of failures preceding the first success, and let \( X_2 \) be the number of failures between the first 2 successes. Find the joint pmf of \( X_1 \) and \( X_2 \). Write an expression for \( E[\sqrt{X_1X_2}] \). You can leave your answer in the form of a sum.
Solution:

$X_1$ and $X_2$ take on two particular values $x_1$ and $x_2$, when there are $x_1$ failures followed by one success, and then $x_2$ failures followed by one success. Since the Bernoulli trials are independent the joint pmf is

$$
p_{X_1,X_2}(x_1,x_2) = (1-p)^{x_1}p \cdot (1-p)^{x_2}p = (1-p)^{x_1+x_2}p^2
$$

By the definition of expectation

$$
E[\sqrt{X_1X_2}] = \sum_{(x_1,x_2)\in\Omega_{X_1,X_2}} \sqrt{x_1x_2} \cdot (1-p)^{x_1+x_2}p^2
$$

6. Continuous joint density I

The joint probability density function of $X$ and $Y$ is given by

$$
f_{X,Y}(x,y) = \begin{cases} 
\frac{6}{7} \left( x^2 + \frac{xy}{2} \right) & 0 < x < 1, \ 0 < y < 2 \\
0 & \text{otherwise.}
\end{cases}
$$

(a) Verify that this is indeed a joint density function.

(b) Compute the marginal density function of $X$.

(c) Find $Pr(X > Y)$. (Uses the continuous law of total probability which we have not covered in class as of 11/19.)

(d) Find $P(Y > \frac{1}{2}|X < \frac{1}{2})$.

(e) Find $E(X)$.

(f) Find $E(Y)$

Solution:

(a) A joint density function will integrate to 1 over all possible values. Since it $f_{X,Y}$ is 0 outside of the specified range, we can just integrate over that range using Wolfram Alpha, to get:

$$
\int_0^2 \int_0^1 \frac{6}{7} \left( x^2 + \frac{xy}{2} \right) dx dy = 1
$$

(b) By definition, the marginal density function of $X$ is the integration over all values of $y$, which again we can only go over the range of $y$ to get:

$$
f_X(x) = \int_0^2 \frac{6}{7} \left( x^2 + \frac{xy}{2} \right) dy = \frac{6}{7}x(2x+1)
$$

Note that this is only the case when $0 < x < 1$. Otherwise, the PDF is equal to 0.

(c) First, we rearrange our initial probability. Then, by the continuous law of total probability:

$$
Pr(X > Y) = 1 - Pr(X \leq Y) = 1 - \int_{-\infty}^{\infty} Pr(X \leq Y | Y = y) f_Y(y) dy = 1 - \int_{-\infty}^{\infty} Pr(X \leq y) f_Y(y) dy
$$

Once again, we can instead integrate over just the range of $y$, getting:

$$
1 - \int_0^2 Pr(X \leq y) f_Y(y) dy
$$
We have to remember that \( x \) operates on the range \( 0 < x < 1 \). Thus, the CDF is 1 past the top of that range, so we have:

\[
1 - \int_0^1 \mathbb{P}(X \leq y) f_Y(y) dy - \int_1^2 f_Y(y) dy
\]

So, now we just need to find the CDF of \( X \), and the marginal PDF of \( Y \). For the former:

\[
F_X(x) = \int_0^x \frac{6}{7} u (2u + 1) du = \frac{1}{7} x^2 (4x + 3)
\]

For the latter:

\[
f_Y(y) = \int_0^1 \frac{6}{7} (x^2 + \frac{xy}{2}) dx = \frac{1}{14} (3y + 4)
\]

Putting these together, we get that:

\[
\mathbb{P}(X > Y) = 1 - \int_0^1 \frac{1}{7} y^2 (4y + 3) \frac{1}{14} (3y + 4) dy - \int_1^2 \frac{1}{14} (3y + 4) dy = 1 - \frac{253}{1960} - \frac{17}{28} = \frac{517}{1960} = 0.2638
\]

(d) By the definition of conditional probability:

\[
\mathbb{P}(Y > \frac{1}{2} \mid X < \frac{1}{2}) = \frac{\mathbb{P}(Y > \frac{1}{2}, X < \frac{1}{2})}{\mathbb{P}(X < \frac{1}{2})}
\]

We can solve for the numerator by integrating over the given range for the joint distribution. We know that the joint PDF is 0 outside of its range, so we can truncate the integral to that:

\[
\mathbb{P}(Y > \frac{1}{2}, X < \frac{1}{2}) = \int_{1/2}^\infty \int_{-\infty}^{1/2} f_{X,Y}(x,y) dxdy = \int_{1/2}^\infty \int_{-\infty}^{1/2} \frac{6}{7} \left( x^2 + \frac{xy}{2} \right) dxdy
\]

\[
= \int_{1/2}^1 \int_0^{1/2} \frac{6}{7} \left( x^2 + \frac{xy}{2} \right) dxdy = \frac{69}{448}
\]

For the denominator, we can integrate over the marginal distribution that we found before:

\[
\int_{-\infty}^{1/2} \frac{6}{7} x (2x + 1) dx = \frac{5}{28}
\]

Putting these together, we get:

\[
\mathbb{P}(Y > \frac{1}{2} \mid X < \frac{1}{2}) = \frac{69}{448} = 0.8625
\]

(e) By definition, and because of the aforementioned range situation:

\[
\mathbb{E}[X] = \int_0^1 f_X(x) x dx = \int_0^1 \frac{6}{7} x (2x + 1) dx = \frac{5}{7}
\]

(f) By definition, and because of the aforementioned range situation:

\[
\mathbb{E}[Y] = \int_0^2 f_Y(y) y dy = \int_0^2 \frac{1}{14} (3y + 4) y dy = \frac{8}{7}
\]
7. Continuous joint density II

The joint density of \( X \) and \( Y \) is given by

\[
f_{X,Y}(x, y) = \begin{cases} \ xe^{-(x+y)} & x > 0, y > 0 \\ 0 & \text{otherwise.} \end{cases}
\]

and the joint density of \( W \) and \( V \) is given by

\[
f_{W,V}(w, v) = \begin{cases} \ 2 & 0 < w < v, 0 < v < 1 \\ 0 & \text{otherwise.} \end{cases}
\]

Are \( X \) and \( Y \) independent? Are \( W \) and \( V \) independent?

Solution:

For two random variables \( X, Y \) to be independent, \( f_{X,Y}(x, y) = f_X(x)f_Y(y) \). Let’s start with \( X \) and \( Y \) by finding their marginal PDFs. By definition, the marginal density function of \( X \) is the integration over all values of \( Y \). Since the joint PDF is 0 outside of \( y > 0 \), we can integrate only over that range to get:

\[
f_X(x) = \int_0^\infty xe^{-(x+y)} \, dy = e^{-x} x
\]

We do the same to get the PDF of \( Y \), with the same range of \( x \): \( x > 0 \):

\[
f_Y(y) = \int_0^\infty xe^{-(x+y)} \, dx = e^{-y}
\]

Since \( e^{-x} x \cdot e^{-y} = xe^{-(x+y)} = xe^{-(x+y)} \), \( X \) and \( Y \) are independent.

We can identify that \( W \) and \( V \) are not independent simply by looking at their range of \( f_{W,V}(w, v) \). Graphing it with \( w \) as the "x-axis" and \( v \) as the "y-axis", we get:
The triangular space between the lines is the range. Looking at it, we can see that it is not rectangular, and therefore it is impossible that $\Omega_{W,V} = \Omega_W \times \Omega_V$. Remember, the joint range being the cartesian product of the marginal ranges is not sufficient for independence, but it is necessary. Therefore, this is enough to show that they are not independent.

We could have also shown the same process for $W$ and $V$. We can start with the marginal PDF of $W$. Since the joint PDF is 0 outside of the range $0 < v < 1$, we can integrate over that range to get:

$$f_W(w) = \int_0^1 2 dv = 2$$

We do the same to get the PDF of $V$, integrating with respect to $w$ over the range $0 < w < v$:

$$f_V(v) = \int_0^v 2 dw = 2v$$

Since $2 \cdot 2v \neq 2$, $W$ and $V$ are NOT independent.

### 8. Trapped Miner

A miner is trapped in a mine containing 3 doors.

- $D_1$: The 1st door leads to a tunnel that will take him to safety after 3 hours.
- $D_2$: The 2nd door leads to a tunnel that returns him to the mine after 5 hours.
- $D_3$: The 3rd door leads to a tunnel that returns him to the mine after a number of hours that is Binomial with parameters $(12, \frac{1}{3})$.

At all times, he is equally likely to choose any one of the doors. What is the expected number of hours for this miner to reach safety?

**Solution:**

Let $T = \text{number of hours for the miner to reach safety}$. ($T$ is a random variable)

Let $D_i$ be the event the $i^{th}$ door is chosen. $i \in \{1, 2, 3\}$. Finally, let $T_3$ be the time it takes to return to the mine in the third case only (a random variable). Note that the expectation of $T_3$ is $12 \cdot \frac{1}{3}$ because it is binomially distributed with parameters $n = 12, p = \frac{1}{3}$. By Law of Total Expectation, linearity of expectation, and by applying the conditional expectations given by the problem statement:

$$E[T] = E[T \mid D_1] P(D_1) + E[T \mid D_2] P(D_2) + E[T \mid D_3] P(D_3)$$

$$= 3 \cdot \frac{1}{3} + (5 + E[T]) \cdot \frac{1}{3} + (E[T_3 + T]) \cdot \frac{1}{3}$$

$$= 3 \cdot \frac{1}{3} + (5 + E[T]) \cdot \frac{1}{3} + (E[T_3] + E[T]) \cdot \frac{1}{3}$$

$$= 3 \cdot \frac{1}{3} + (5 + E[T]) \cdot \frac{1}{3} + (4 + E[T]) \cdot \frac{1}{3}$$

Solving this equation for $E[T]$, we get

$$E[T] = 12$$

Therefore, the expected number of hours for this miner to reach safety is 12.
9. Lemonade Stand
Suppose I run a lemonade stand, which costs me $100 a day to operate. I sell a drink of lemonade for $20. Every person who walks by my stand either buys a drink or doesn’t (no one buys more than one). If it is raining, \( n_1 \) people walk by my stand, and each buys a drink independently with probability \( p_1 \). If it isn’t raining, \( n_2 \) people walk by my stand, and each buys a drink independently with probability \( p_2 \). It rains each day with probability \( p_3 \), independently of every other day. Let \( X \) be my profit over the next week. In terms of \( n_1, n_2, p_1, p_2 \) and \( p_3 \), what is \( \mathbb{E}[X] \)?

**Solution:**
Let \( R \) be the event it rains. Let \( X_i \) be how many drinks I sell on day \( i \) for \( i = 1, \ldots, 7 \). We are interested in \( X = \sum_{i=1}^{7} (20X_i - 100) \). We have \( X_i | R \sim \text{Binomial}(n_1, p_1) \), so \( \mathbb{E}[X_i | R] = n_1p_1 \). Similarly, \( X_i | R^C \sim \text{Binomial}(n_2, p_2) \), so \( \mathbb{E}[X_i | R^C] = n_2p_2 \). By the law of total expectation,

\[
\mu = \mathbb{E}[X] = \mathbb{E}[X_i | R] \mathbb{P}(R) + \mathbb{E}[X_i | R^C] \mathbb{P}(R^C) = n_1p_1p_3 + n_2p_2(1 - p_3)
\]

Hence, by linearity of expectation,

\[
\mathbb{E}[X] = \mathbb{E}\left[ \sum_{i=1}^{7} (20X_i - 100) \right] = 20 \sum_{i=1}^{7} \mathbb{E}[X_i] - 700 = 140\mu - 700
\]

\[
= 140 \cdot (n_1p_1p_3 + n_2p_2(1 - p_3)) - 700.
\]

10. Particle Emissions
Suppose we are measuring particle emissions, and the number of particles emitted follows a Poisson distribution with parameter \( \lambda \), \( X \sim \text{Poisson}(\lambda) \). Suppose our device to measure emissions is not always entirely accurate sometimes we fail to observe particles that actually emitted. So for each particle actually emitted, say we have with parameter \( \lambda \), \( Y \) be the number of particles we observed. What distribution does \( Y \) follow with what parameters, and what is \( \mathbb{E}[Y] \)?

**Solution:**
Let \( p \) be the probability of actually recording it, independently of other particles. Let \( R \) be the event it rains. Let \( X \sim \text{Poisson}(\lambda) \) and \( R \sim \text{Bernoulli}(p) \). Suppose we are measuring particle emissions, and the number of particles emitted follows a Poisson distribution with parameter \( \lambda \), \( X \sim \text{Poisson}(\lambda) \). Suppose our device to measure emissions is not always entirely accurate sometimes we fail to observe particles that actually emitted. So for each particle actually emitted, say we have with parameter \( \lambda \), \( Y \) be the number of particles we observed. What distribution does \( Y \) follow with what parameters, and what is \( \mathbb{E}[Y] \)?

\[
p_Y(y) = \mathbb{P}(Y = y)
\]

\[
= \sum_{x=y}^{\infty} \mathbb{P}(Y = y | X = x) \mathbb{P}(X = x) \quad \text{(Law of Total Probability)}
\]

\[
= \sum_{x=y}^{\infty} \binom{x}{y} p^y (1-p)^{x-y} \cdot e^{-\lambda} \frac{\lambda^x}{x!}
\]

\[
= e^{-\lambda} \sum_{x=y}^{\infty} \frac{x!}{y!(x-y)!} \frac{\lambda^x}{x!} p^y (1-p)^{x-y}
\]

\[
= e^{-\lambda} \frac{p^y}{y!} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} (1-p)^k (let \ k = x-y)
\]

\[
= e^{-\lambda} \frac{(\lambda p)^y}{y!} \sum_{k=0}^{\infty} \frac{(\lambda (1-p))^k}{k!}
\]

\[
= e^{-\lambda (1-p)} \cdot e^{\lambda (1-p)}
\]

\[
= e^{-p\lambda (1-p)^y}
\]
So \( Y \sim \text{Poisson}(\mu) \) and \( \mathbb{E}[Y] = \mu \).

11. In between
(Covers ideas that have not been covered in class.) Suppose that \( X_1 \) and \( X_2 \) are discrete uniform random variables in \( \{1, \ldots, 2n\} \) (i.e., \( X_1 \) and \( X_2 \) are equally likely to take any of the values \( 1, \ldots, 2n \)) and let \( Y = \min(X_1, X_2) \). What is the conditional pmf \( p_{Y|X_1}(y \mid x_1) \) and conditional CDF \( F_{Y|X_1}(y \mid x_1) \). What is \( E[Y \mid X_1 = x_1] \)? (For the definitions of conditional pmf, conditional CDF, see the review at the top of this worksheet.)

**Solution:**
The conditional pmf is

\[
p_{Y|X_1}(y \mid x_1) = \mathbb{P}(\min(X_1, X_2) = y \mid X_1 = x_1) = \frac{\mathbb{P}(\min(X_1, X_2) = y, X_1 = x_1)}{\mathbb{P}(X_1 = x_1)} = \begin{cases} 0 & \text{if } y > x_1 \\ 1 - \frac{x_1 - 1}{2n} & \text{if } y = x_1 \\ \frac{1}{2n} & \text{if } 1 \leq y < x_1 \end{cases}
\]

Explanation for the final equality:

- Since \( \min(X_1, X_2) \leq X_1 \), if \( y > x_1 \), then we have
  \( \mathbb{P}(\min(X_1, X_2) = y, X_1 = x_1) = 0 \)

- When \( y = x_1 \), we have
  \[
  \mathbb{P}(\min(X_1, X_2) = x_1 \mid X_1 = x_1) = \mathbb{P}(X_2 \geq x_1 \mid X_1 = x_1) = \mathbb{P}(X_2 \geq x_1) = \frac{2n - x_1 + 1}{2n} = 1 - \frac{x_1 - 1}{2n}
  \]

The second to last equality holds because \( X_1 \) and \( X_2 \) are independent.

- In the case where \( y < x_1 \), \( \min(X_1, X_2) = X_2 \). So for \( y < x_1 \),
  \[
  \mathbb{P}(Y = y \mid X_1 = x_1) = \mathbb{P}(X_2 = y \mid X_1 = x_1) = \mathbb{P}(X_2 = y) = \frac{1}{2n}.
  \]

Again, we used the independence of \( X_1 \) and \( X_2 \).

The conditional CDF can be computed as follows

\[
F_{Y|X_1}(y \mid x_1) = \sum_{i=1}^{y} p_{Y|X_1}(i \mid x_1) = \begin{cases} 0 & \text{if } y < 0 \\ \frac{y}{2n} & \text{if } y < x_1 \\ 1 & \text{if } y \geq x_1 \end{cases}
\]

The expected value can be computed as follows:

\[
E[Y \mid X_1 = x_1] = \sum_{y=1}^{2n} y p_{Y|X_1}(y \mid x_1) = \sum_{y=1}^{x_1-1} y \frac{1}{2n} + \sum_{y=x_1}^{x_1} y \left(1 - \frac{y-1}{2n}\right) = \frac{1}{2n} \sum_{y=1}^{x_1-1} y + x_1 \left(1 - \frac{x_1 - 1}{2n}\right)
\]

12. 3 points on a line
(This problem uses the continuous law of total probability which has not yet be covered in class.) Three points \( X_1, X_2, X_3 \) are selected at random on a line \( L \) (continuous independent uniform distributions). What is the probability that \( X_2 \) lies between \( X_1 \) and \( X_3 \)?
Solution:
Let $X_1, X_2, X_3 \sim Unif(0, 1)$.

\[
P(X_1 < X_2 < X_3) = \int_{-\infty}^{\infty} P(X_1 < X_2 < X_3 \mid X_2 = x) f_{X_2}(x) \, dx
\]

Continuous LoTP

\[
= \int_{-\infty}^{\infty} P(X_1 < x, X_3 > x) f_{X_2}(x) \, dx
\]

independence of $X_1, X_2, X_3$

\[
= \int_{-\infty}^{\infty} P(X_1 < x) P(x < X_3) f_{X_2}(x) \, dx
\]

Independence

\[
= \int_{-\infty}^{\infty} F_{X_1}(x) (1 - F_{X_3}(x)) f_{X_2}(x) \, dx
\]

\[
= \int_{0}^{1} x (1 - x) 1 \, dx
\]

\[
= \frac{x^2}{2} - \frac{x^3}{3} \bigg|_{0}^{1} = \frac{1}{6}
\]