



Section 7

CLT and Joint Distributions

The image features a central orange rectangular area containing the text 'CONTENT REVIEW'. This central area is framed by teal horizontal bars at the top and bottom. Scattered around the edges are several brown circles of varying sizes, creating a decorative, bubbly effect.

CONTENT REVIEW

normal distribution

this PDF is really messy :(

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

use the Phi Table to easily look up the value of $P(Z < z)$

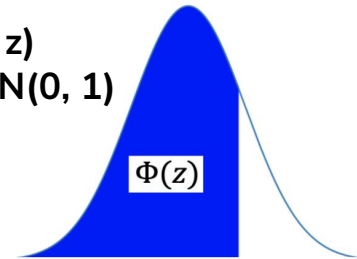
1. Find $(x-\mu)/\sigma$ of normal RV X

1. Standardize RV $X \sim N(\mu, \sigma^2)$ to get $(X-\mu)/\sigma \sim Z \sim N(0,1)$

2. Use **Phi table** to get appropriate value with $Z = (X-\mu)/\sigma$

3. Solve for X

$P(Z < z)$
if $Z \sim N(0, 1)$



normal distribution

properties of the normal distribution

Closure for Normal Distribution

Let $X \sim N(\mu, \sigma^2)$. Then $aX + b \sim N(a\mu + b, b^2\sigma^2)$

“Reproductive” Property

Let X_1, \dots, X_n be independent normal RVs with $E[X_i] = \mu$ and $\text{Var}(X_i) = \sigma_i^2$

$$X = \sum_{i=1}^n (a_i X_i + b) \sim N \left(\sum_{i=1}^n (a_i \mu_i + b), \sum_{i=1}^n a_i^2 \sigma_i^2 \right)$$

normal distribution

properties of the normal distribution

Closure for Normal Distribution

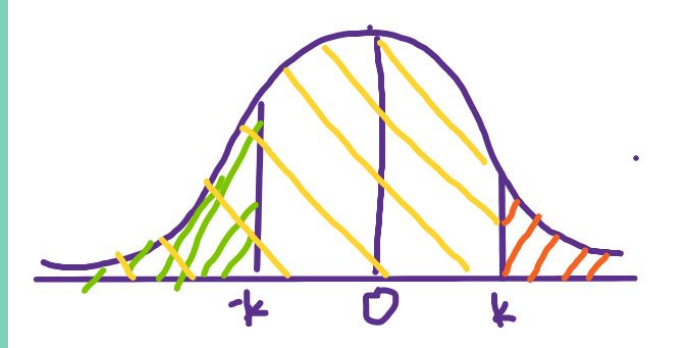
Let $X \sim N(\mu, \sigma^2)$. Then $aX + b \sim N(a\mu + b, b^2\sigma^2)$

“Reproductive” Property


Let X_1, \dots, X_n be independent normal RVs with $E[X_i] = \mu$ and $\text{Var}(X_i) = \sigma_i^2$

$$X = \sum_{i=1}^n (a_i X_i + b) \sim N \left(\sum_{i=1}^n (a_i \mu_i + b), \sum_{i=1}^n a_i^2 \sigma_i^2 \right)$$

normal distribution



- By symmetry of normal distribution, $P(X < -k) = P(X > k)$
- $\Phi(a) = P(Z < a)$ where Z (standard normal) follows $N(0, 1)$
- $\Phi(-k) = P(Z < k) = P(Z > k) = 1 - P(Z < k) = 1 - \Phi(k)$
^ this is useful because negative values aren't on the table!



*it turns out that a lot of
data collected in real
world experiments follow
a normal distribution!*



central limit theorem

if $X = X_1 + X_2 + \dots + X_n$ are iid where $E[X_i] = \mu$ and $\text{Var}(X_i) = \sigma^2$,
 X can be approximated by a normal distribution $X \sim N(n\mu, n\sigma^2)$
as n becomes really large

this also means that as n increases, $(X - \mu) / \sigma \sim Z \sim N(0,1)$

if X_i is discrete random variables, use **continuity correction**
because we'd be using a continuous distribution to approximate a discrete distribution

*To estimate probability that discrete RV lands in (integer) interval $\{a,b\}$
compute probability continuous approximation lands in interval $[a-0.5, b+0.5]$*



central limit theorem

1. set up the problem
(what do we want to solve for/what is the probability we want to be true?)

2. if the distribution is discrete, apply continuity correction here

2. apply CLT to $X = X_1 + \dots + X_n$

3. convert to standard normal


4. solve!

joint distribution

shows the distribution for two (or more) random variables

$$p_{X,Y}(x, y) = P(X=x, Y=y)$$


| | Discrete | Continuous |
|--|--|---|
| Joint PMF/PDF | $p_{X,Y}(x, y) = \mathbb{P}(X = x, Y = y)$ | $f_{X,Y}(x, y) \neq \mathbb{P}(X = x, Y = y)$ |
| Joint range/support $\Omega_{X,Y}$ | $\{(x, y) \in \Omega_X \times \Omega_Y : p_{X,Y}(x, y) > 0\}$ | $\{(x, y) \in \Omega_X \times \Omega_Y : f_{X,Y}(x, y) > 0\}$ |
| Joint CDF | $F_{X,Y}(x, y) = \sum_{t \leq x, s \leq y} p_{X,Y}(t, s)$ | $F_{X,Y}(x, y) = \int_{-\infty}^x \int_{-\infty}^y f_{X,Y}(t, s) ds dt$ |
| Normalization | $\sum_{x,y} p_{X,Y}(x, y) = 1$ | $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx dy = 1$ |
| Marginal PMF/PDF | $p_X(x) = \sum_y p_{X,Y}(x, y)$ | $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dy$ |
| Expectation | $\mathbb{E}[g(X, Y)] = \sum_{x,y} g(x, y) p_{X,Y}(x, y)$ | $\mathbb{E}[g(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f_{X,Y}(x, y) dx dy$ |
| Independence must have | $\forall x, y, p_{X,Y}(x, y) = p_X(x) p_Y(y)$ $\Omega_{X,Y} = \Omega_X \times \Omega_Y$ | $\forall x, y, f_{X,Y}(x, y) = f_X(x) f_Y(y)$ $\Omega_{X,Y} = \Omega_X \times \Omega_Y$ |
| Conditional PMF/PDF | $p_{X Y}(x y) = \frac{p_{X,Y}(x,y)}{p_Y(y)}$ | $f_{X Y}(x y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$ |
| Conditional Expectation | $\mathbb{E}[X Y = y] = \sum_x x \cdot p_{X Y}(x y)$ | $\mathbb{E}[X Y = y] = \int_{-\infty}^{\infty} x f_{X Y}(x y) dx$ |



Suppose X and Y have the following joint PMF:

| X/Y | 1 | 2 | 3 |
|-----|-----|-----|-----|
| 0 | 0 | 0.2 | 0.1 |
| 1 | 0.3 | 0 | 0.4 |

these are
the values
 X takes on



Suppose X and Y have the following joint PMF:

| X/Y | 1 | 2 | 3 |
|-----|-----|-----|-----|
| 0 | 0 | 0.2 | 0.1 |
| 1 | 0.3 | 0 | 0.4 |

these are
the values
 Y takes on

these are
the values
 X takes on

Suppose X and Y have the following joint PMF:

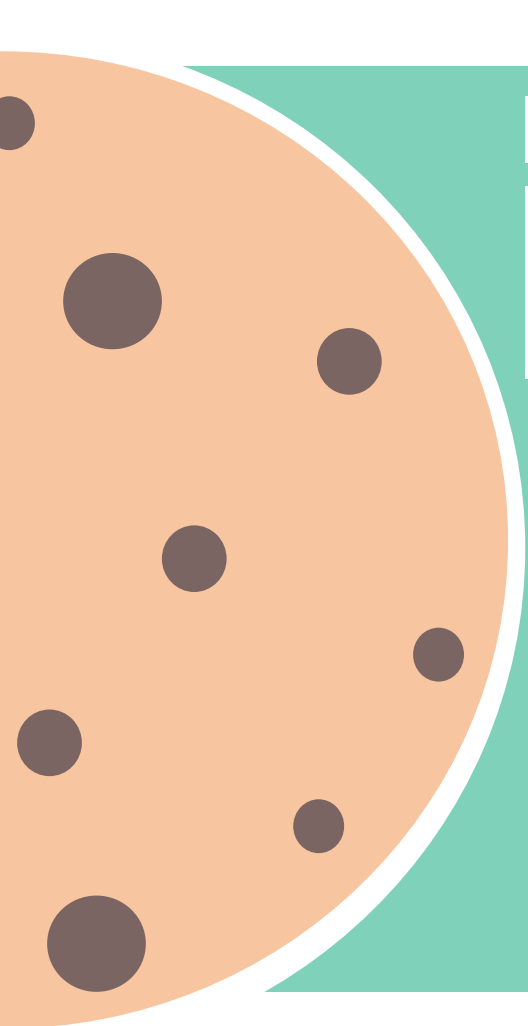
| X/Y | 1 | 2 | 3 |
|-----|-----|-----|-----|
| 0 | 0 | 0.2 | 0.1 |
| 1 | 0.3 | 0 | 0.4 |

these are the values Y takes on

these are the values X takes on

the values in the table show the joint probabilities - for example this highlighted value is $P(X=1, Y=3)$



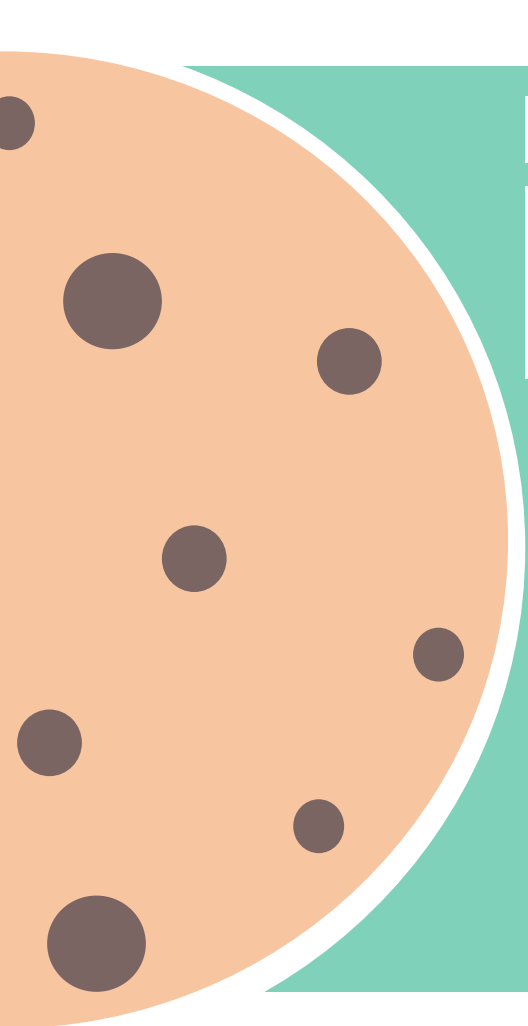


Suppose X and Y have the following joint PMF:

| X/Y | 1 | 2 | 3 |
|-----|-----|-----|-----|
| 0 | 0 | 0.2 | 0.1 |
| 1 | 0.3 | 0 | 0.4 |

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

- Based on the table (specifically what's highlighted in yellow and pink), we know that the range of X is $\{0, 1\}$, and the range of Y is $\{1, 2, 3\}$




Suppose X and Y have the following joint PMF:

| X/Y | 1 | 2 | 3 |
|-----|-----|-----|-----|
| 0 | 0 | 0.2 | 0.1 |
| 1 | 0.3 | 0 | 0.4 |

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

- Based on the table (specifically what's highlighted in yellow and pink), we know that the range of X is $\{0, 1\}$, and the range of Y is $\{1, 2, 3\}$
- The joint range is the set of *pairs of x and y* that both X and Y can take on at the same time. In other words, the pairs of values x and y such that the joint PMF $P(X=x, Y=y)$ is greater than 0.

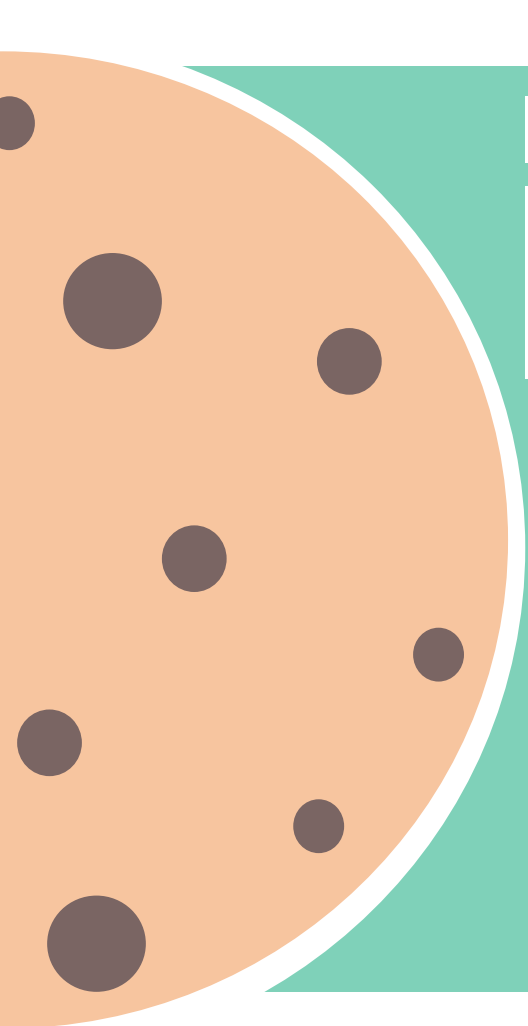
A large, stylized illustration of a chocolate chip cookie is positioned on the left side of the slide. The cookie is light brown with several dark brown chocolate chips scattered across its surface. The right edge of the cookie is curved, and it overlaps the green background of the slide.

Suppose X and Y have the following joint PMF:

| X/Y | 1 | 2 | 3 |
|-----|-----|-----|-----|
| 0 | 0 | 0.2 | 0.1 |
| 1 | 0.3 | 0 | 0.4 |

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

- Based on the table (specifically what's highlighted in yellow and pink), we know that the range of X is $\{0, 1\}$, and the range of Y is $\{1, 2, 3\}$
- The joint range is the set of *pairs of x and y* that both X and Y can take on at the same time. In other words, the pairs of values x and y such that the joint PMF $P(X=x, Y=y)$ is greater than 0.
- In this case, that would be the set of pairs:
$$\Omega_{X,Y} = \{(0, 2), (0, 3), (1, 1), (1, 3)\}$$
- We don't include the pairs $(0, 1)$ and $(1, 2)$ because the joint pmf is 0 for those pairs

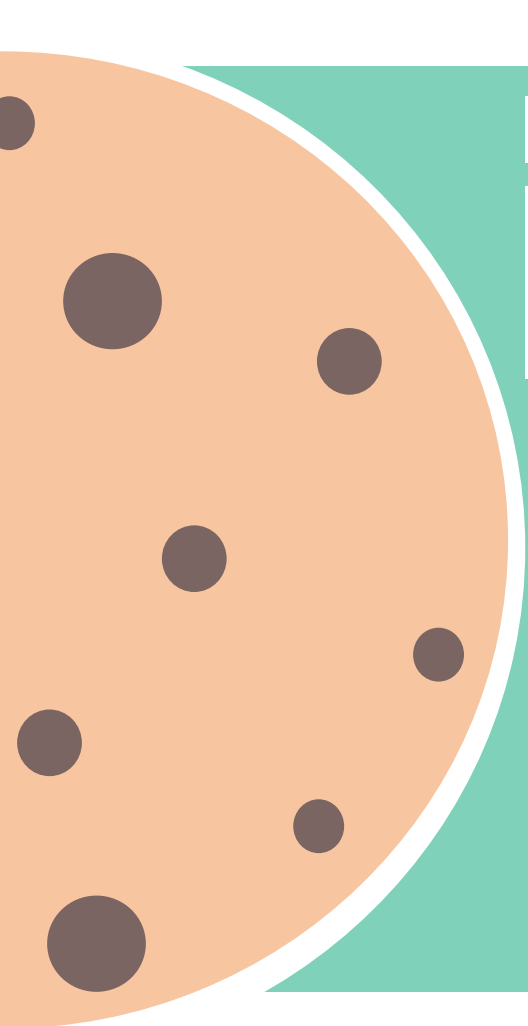


Suppose X and Y have the following joint PMF:

| X/Y | 1 | 2 | 3 |
|-----|-----|-----|-----|
| 0 | 0 | 0.2 | 0.1 |
| 1 | 0.3 | 0 | 0.4 |

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

- The marginal PMF for X is just like asking what the PMF for X is based on the joint PMF
- There are only two values in the range of X so we can try defining this PMF by looking at the probabilities for each value separately!
- First, we want to find $P(X=0)$ -

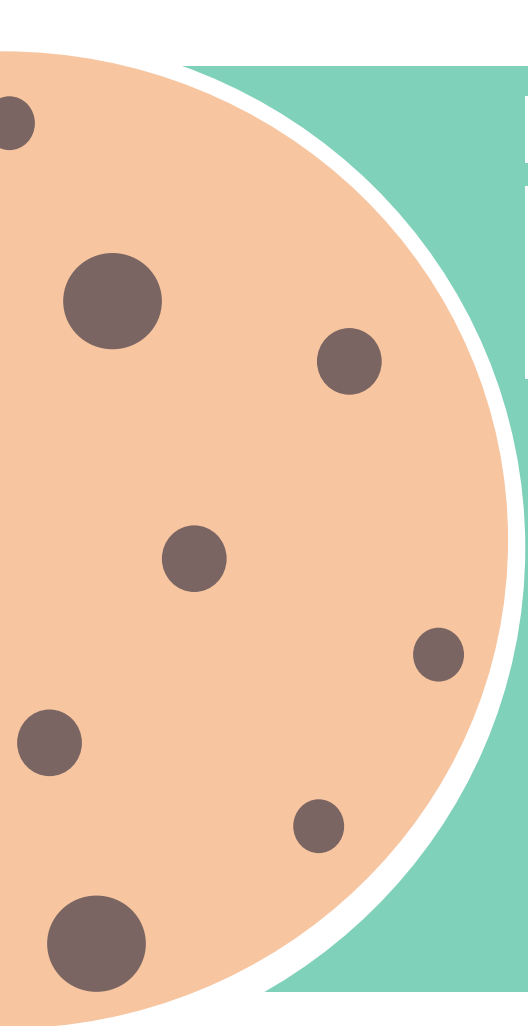


Suppose X and Y have the following joint PMF:

| X/Y | 1 | 2 | 3 |
|-----|-----|-----|-----|
| 0 | 0 | 0.2 | 0.1 |
| 1 | 0.3 | 0 | 0.4 |

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

- The marginal PMF for X is just like asking what the PMF for X is based on the joint PMF
- There are only two values in the range of X so we can try defining this PMF by looking at the probabilities for each value separately!
- First, we want to find $P(X=0)$ - Based on the joint PMF, and using LOTP
$$P(X=0) = P(X=0, Y=1) + P(X=0, Y=2) + P(X=0, Y=3) = 0 + 0.2 + 0.1 = 0.3$$

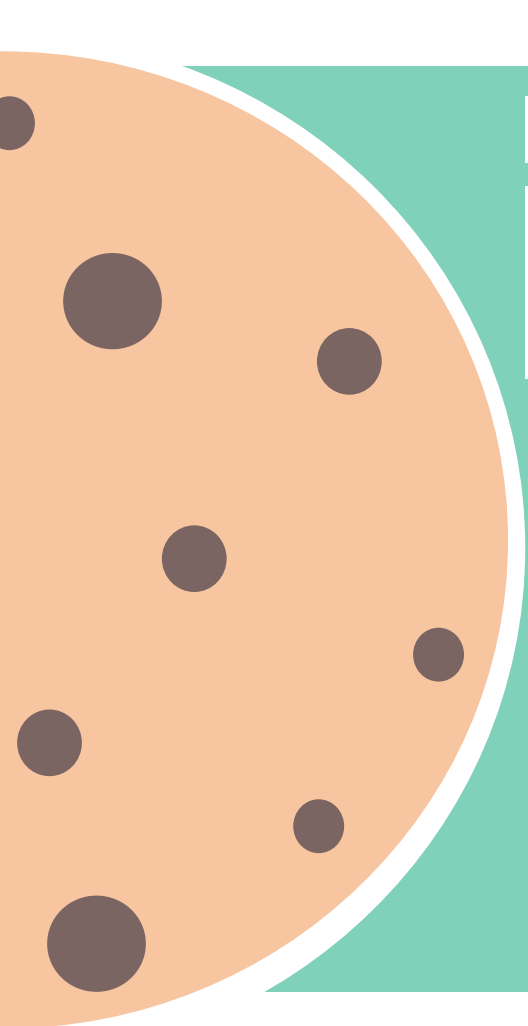


Suppose X and Y have the following joint PMF:

| X/Y | 1 | 2 | 3 |
|-----|-----|-----|-----|
| 0 | 0 | 0.2 | 0.1 |
| 1 | 0.3 | 0 | 0.4 |

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

- The marginal PMF for X is just like asking what the PMF for X is based on the joint PMF
- There are only two values in the range of X so we can try defining this PMF by looking at the probabilities for each value separately!
- First, we want to find $P(X=0)$ - Based on the joint PMF, and using LOTP
$$P(X=0) = P(X=0, Y=1) + P(X=0, Y=2) + P(X=0, Y=3) = 0 + 0.2 + 0.1 = 0.3$$
- Similarly, $P(X=1) = 0.3 + 0 + 0.4$

A large, stylized illustration of a cookie with several dark brown chocolate chips, positioned on the left side of the slide. The cookie is light brown and has a white border.

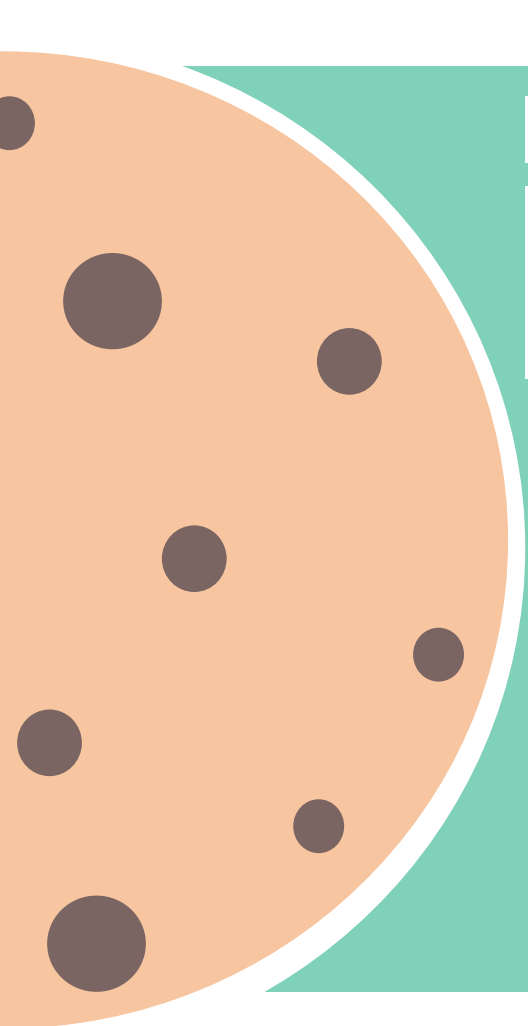
Suppose X and Y have the following joint PMF:

| X/Y | 1 | 2 | 3 |
|-----|-----|-----|-----|
| 0 | 0 | 0.2 | 0.1 |
| 1 | 0.3 | 0 | 0.4 |

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

- The marginal PMF for X is just like asking what the PMF for X is based on the joint PMF
- There are only two values in the range of X so we can try defining this PMF by looking at the probabilities for each value separately!
- First, we want to find $P(X=0)$ - Based on the joint PMF, and using LOTP
 $P(X=0) = P(X=0, Y=1) + P(X=0, Y=2) + P(X=0, Y=3) = 0 + 0.2 + 0.1 = 0.3$
- Similarly, $P(X=1) = 0.3 + 0 + 0.4$
- In general, the marginal pmf looks like:

$$p_X(x) = \sum_{y \in \Omega_Y} p_{X,Y}(x, y)$$



Suppose X and Y have the following joint PMF:

| X/Y | 1 | 2 | 3 |
|-----|-----|-----|-----|
| 0 | 0 | 0.2 | 0.1 |
| 1 | 0.3 | 0 | 0.4 |

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

- If we want to prove that X and Y are independent, we would need to show both of the conditions to be true:

$$\forall x, y, p_{X,Y}(x, y) = p_X(x)p_Y(y)$$
$$\Omega_{X,Y} = \Omega_X \times \Omega_Y$$

- However, note that the second condition is not true in this case! (take a look at the ranges we found in part a). So, X and Y are not independent.

The slide features a teal horizontal bar at the top and bottom. A large orange rectangle occupies the center. The text "Thanks for coming today!" is centered within the orange rectangle. Brown circles of various sizes are scattered around the slide, primarily on the left and bottom right sides.

Thanks for coming today!