Joint & Conditional Entropy

Problem:

Grant, Coulton, and Shengwei are researchers studying the behavior of a unique species of creatures inhabiting an island. They record two characteristics of these creatures: their **size (X)** and their **color (Y)**. The joint distribution of these characteristics is as follows:

X/Y	1	2	3	4
1	1/8	1/16	1/32	1/32
2	1/16	1/8	1/32	1/32
3	1/16	1/16	1/16	1/16
4	1/4	0	0	0

Questions:

(a) Grant, Coulton, and Shengwei are analyzing the data they collected on the creatures' size X and color Y on the island. What is the entropy of the size X and the entropy of the color Y of these creatures? Solve for $H(X)$ and $H(Y)$. Solve by hand then validate with code.

(b) Considering the creatures on the island, if one knows the color of a creature, how much uncertainty remains about its size, and vice versa? Determine the conditional entropy of the size given the color and the conditional entropy of the color given the size. Solve for $H(X|Y)$ and $H(Y|X)$. Solve by hand then validate with code.

(c) Grant, Coulton, and Shengwei have computed the joint entropy of the size and color of the creatures on the island. What is the joint entropy of these characteristics? Solve for $H(X, Y)$. Solve by hand then validate with code.

(d) After discovering the size of the creatures on the island, how much uncertainty is left about their color? Calculate the reduction in uncertainty about the color given the knowledge of the size. Solve for $H(Y)$ - $H(Y|X)$. Solve by hand then validate with code.

(e) Grant, Coulton, and Shengwei want to quantify the amount of information shared between the size and color of the creatures on the island. Compute $I(X;Y)$. Solve by hand then validate with code.

In [42]:
```python
import numpy as np
```

Solutions:

(a) Solve for $H(X)$ and $H(Y)$. Solve by hand then validate with code.

Marginal Distribution of Size (X):

$P(X=x) = \sum_y P(X=x, Y=y)$

$P(X=1) = [P(X=1,Y=1) + P(X=1,Y=2) + P(X=1,Y=3)$
$+P(X=1,Y=4)]$ = $\frac{1}{2}$

$P(X=2) = [P(X=2,Y=1) + P(X=2,Y=2) + P(X=2,Y=3)$
$+P(X=2,Y=4)]$ = $\frac{1}{4}$

$P(X=3) = [P(X=3,Y=1) + P(X=3,Y=2) + P(X=3,Y=3)$
$+P(X=3,Y=4)]$ = $\frac{1}{8}$

$P(X=4) = [P(X=4,Y=1) + P(X=4,Y=2) + P(X=4,Y=3)$
$+P(X=4,Y=4)]$ = $\frac{1}{8}$

Marginal Distribution of Color (Y):

$P(Y=y) = \sum_x P(X=x, Y=y)$

$P(Y=1) = [P(X=1,Y=1) + P(X=2,Y=1) + P(X=3,Y=1)$
$+P(X=4,Y=1)]$ = $\frac{1}{4}$

$P(Y=2) = [P(X=1,Y=2) + P(X=2,Y=2) + P(X=3,Y=2)$
$+P(X=4,Y=2)]$ = $\frac{1}{4}$

$P(Y=3) = [P(X=1,Y=3) + P(X=2,Y=3) + P(X=3,Y=3)$
$+P(X=4,Y=3)]$ = $\frac{1}{4}$

$P(Y=4) = [P(X=1,Y=4) + P(X=2,Y=4) + P(X=3,Y=4)$
$+P(X=4,Y=4)]$ = $\frac{1}{4}$

Entropy of Size (X):

$$H(X) = [-\sum_x P(X=x)\log_2 P(X=x)]$$

$$H(X) = [-\left(\frac{1}{2}\log_2\frac{1}{2} + \frac{1}{4}\log_2\frac{1}{4} + \frac{1}{8}\log_2\frac{1}{8} + \frac{1}{8}\log_2\frac{1}{8}\right)]$$

$$H(X) = \frac{7}{4}\text{ bits}$$

Entropy of Color (Y):

$$H(Y) = [-\sum_y P(Y=y)\log_2 P(Y=y)]$$

$$H(Y) = [-\left(\frac{1}{4}\log_2\frac{1}{4} + \frac{1}{4}\log_2\frac{1}{4} + \frac{1}{4}\log_2\frac{1}{4} + \frac{1}{4}\log_2\frac{1}{4}\right)]$$

$$H(Y) = 2\text{ bits}$$

In [43]:

```
# Define the joint distribution as a numpy array
joint_distribution = np.array([
    [1/8, 1/16, 1/16, 1/4],
    [1/16, 1/8, 1/16, 0],
    [1/32, 1/32, 1/16, 0],
    [1/32, 1/32, 1/16, 0]
])

# Marginal Distribution of Size (X)
px = np.sum(joint_distribution, axis=1)

# Marginal Distribution of Color (Y)
py = np.sum(joint_distribution, axis=0)

# Entropy of Size (X)
hx = -np.sum(px * np.log2(px[px > 0]))

# Entropy of Color (Y)
hy = -np.sum(py * np.log2(py[py > 0]))

print(f"Entropy of Size (H(X)): {hx:.2f} bits")
print(f"Entropy of Color (H(Y)): {hy:.2f} bits")
```

```
Entropy of Size (H(X)): 1.75 bits
Entropy of Color (H(Y)): 2.00 bits
```

Grant, Coulton, and Shengwei determined that the entropy of the creatures' size ($H(X)$) is $\frac{7}{4}$ bits, while the entropy of the color ($H(Y)$) is 2 bits.

(b) Solve for $H(X|Y)$ and $H(Y|X)$. Solve by hand then validate with code.

Calculation of $H(X|Y)$:

$$H(X|Y) = \sum_y P(Y=y)H(X|Y=y)$$

$$H(X|Y) = [P(Y=1)\cdot H(X|Y=1) + P(Y=2)\cdot H(X|Y=2) + P(Y=3)\cdot H(X|Y=3) + P(Y=4)\cdot H(X|Y=4)]$$

$$H(X|Y) = [\tfrac{1}{4}\cdot H(\tfrac{1}{2},\tfrac{1}{4},\tfrac{1}{8},\tfrac{1}{8}) + \tfrac{1}{4}\cdot H(\tfrac{1}{4},\tfrac{1}{2},\tfrac{1}{8},\tfrac{1}{8}) + \tfrac{1}{4}\cdot H(\tfrac{1}{4},\tfrac{1}{4},\tfrac{1}{4},\tfrac{1}{4}) + \tfrac{1}{4}\cdot H(1,0,0,0)]$$

$$H(X|Y) = [\tfrac{1}{4}\cdot\tfrac{7}{4} + \tfrac{1}{4}\cdot\tfrac{7}{4} + \tfrac{1}{4}\cdot 2 + \tfrac{1}{4}\cdot 0] = \tfrac{11}{8} \text{ bits}$$

Calculation of $H(Y|X)$:

$$H(Y|X) = \sum_x P(X=x)H(Y|X=x)$$

$$H(Y|X) = [P(X=1)\cdot H(Y|X=1) + P(X=2)\cdot H(Y|X=2) + P(X=3)\cdot H(Y|X=3) + P(X=4)\cdot H(Y|X=4)]$$

$$H(Y|X) = [\tfrac{1}{2}\cdot H(\tfrac{1}{4},\tfrac{1}{8},\tfrac{1}{8},\tfrac{1}{2}) + \tfrac{1}{4}\cdot H(\tfrac{1}{4},\tfrac{1}{2},\tfrac{1}{4},0) + \tfrac{1}{8}\cdot H(\tfrac{1}{4},\tfrac{1}{4},\tfrac{1}{2},0) + \tfrac{1}{8}\cdot H(\tfrac{1}{4},\tfrac{1}{4},\tfrac{1}{2},0)]$$

$$H(Y|X) = [\tfrac{1}{2}\cdot\tfrac{7}{4} + \tfrac{1}{4}\cdot\tfrac{3}{2} + \tfrac{1}{8}\cdot\tfrac{3}{2} + \tfrac{1}{8}\cdot\tfrac{3}{2}] = \tfrac{13}{8} \text{ bits}$$

In [44]:

```python
# Define the joint distribution as a numpy array
joint_distribution = np.array([
    [1/8, 1/16, 1/16, 1/4],
    [1/16, 1/8, 1/16, 0],
    [1/32, 1/32, 1/16, 0],
    [1/32, 1/32, 1/16, 0]
])

# Calculation of H(X|Y)
hxy = (1/4) * (7/4) + (1/4) * (7/4) + (1/4) * 2 + (1/4) * 0

# Calculation of H(Y|X)
hyx = (1/2) * (7/4) + (1/4) * (3/2) + (1/8) * (3/2) + (1/8) * (3/2)

print(f"Conditional Entropy of Size given Color (H(X|Y)): {hxy:.2f} bi
print(f"Conditional Entropy of Color given Size (H(Y|X)): {hyx:.2f} bi
```

```
Conditional Entropy of Size given Color (H(X|Y)): 1.38 bits
Conditional Entropy of Color given Size (H(Y|X)): 1.62 bits
```

In their analysis of the island creatures' data, Grant, Coulton, and Shengwei determined the conditional entropy of size given color ($H(X|Y)$) to be $\frac{11}{8}$ bits and the conditional entropy of color given size ($H(Y|X)$) to be $\frac{13}{8}$ bits. These values quantify the remaining uncertainty about a creature's size given its color, and vice versa.

(c) Solve for $H(X,Y)$. Solve by hand then validate with code.

Calculation of $H(X,Y)$:

$$H(X,Y) = H(X) + H(Y|X)$$

$$H(X,Y) = [\frac{7}{4} + \frac{13}{8}]$$

$$H(X,Y) = \frac{27}{8} \text{ bits}$$

Alternatively,

$$H(X,Y) = H(Y) + H(X|Y)$$

$$H(X,Y) = [2 + \frac{11}{8}]$$

$$H(X,Y) = \frac{27}{8} \text{ bits}$$

In [45]:

```
# Define the joint distribution as a numpy array
joint_distribution = np.array([
    [1/8, 1/16, 1/16, 1/4],
    [1/16, 1/8, 1/16, 0],
    [1/32, 1/32, 1/16, 0],
    [1/32, 1/32, 1/16, 0]
])

# Calculation of H(X) and H(Y)
px = np.sum(joint_distribution, axis=1)
hx = -np.sum(px * np.log2(px[px > 0]))

py = np.sum(joint_distribution, axis=0)
hy = -np.sum(py * np.log2(py[py > 0]))

# Calculation of H(Y|X)
hyx = (1/2) * (7/4) + (1/4) * (3/2) + (1/8) * (3/2) + (1/8) * (3/2)

# Calculation of H(X,Y)
hxy = hx + hyx

print(f"Joint Entropy of Size and Color (H(X,Y)): {hxy:.2f} bits")
```

```
Joint Entropy of Size and Color (H(X,Y)): 3.38 bits
```

Grant, Coulton, and Shengwei have determined the joint entropy of the size and color of the creatures on the island to be $\frac{27}{8}$ bits. This quantifies the total uncertainty associated with both characteristics considered together.

---

(4) Solve for $H(Y)$ - $H(Y|X)$. Solve by hand then validate with code.

Calculation of $H(Y)$ - $H(Y|X)$:

$$H(Y) - H(Y|X) = [2 - \frac{13}{8}]$$

$$H(Y) - H(Y|X) = [\frac{16}{8} - \frac{13}{8}]$$

$$H(Y) - H(Y|X) = \frac{3}{8} \text{ bits}$$

In [46]:

```
# Define the joint distribution as a numpy array
joint_distribution = np.array([
    [1/8, 1/16, 1/16, 1/4],
    [1/16, 1/8, 1/16, 0],
    [1/32, 1/32, 1/16, 0],
    [1/32, 1/32, 1/16, 0]
])

# Calculation of H(Y)
py = np.sum(joint_distribution, axis=0)
hy = -np.sum(py * np.log2(py[py > 0]))

# Calculation of H(Y|X)
hyx = (1/2) * (7/4) + (1/4) * (3/2) + (1/8) * (3/2) + (1/8) * (3/2)

# Calculation of H(Y) - H(Y|X)
hy_minus_hyx = hy - hyx

print(f"Reduction in uncertainty about color given knowledge of size: 
    {hy_minus_hyx:.2f} bits")
```

```
Reduction in uncertainty about color given knowledge of size: 0.38 bits
```

After discovering the size of the creatures on the island, there remains $\frac{3}{8}$ bits of uncertainty about their color. This represents the reduction in uncertainty about color given knowledge of size, as computed by the difference between the entropy of color
($H(Y)$) and the conditional entropy of color given size ($H(Y|X)$).

---

(5) Compute $I(X;Y)$. Solve by hand then validate with code.

Calculation of $I(X;Y)$:

$$I(X;Y) = H(X) - H(X|Y)$$

$$I(X;Y) = [\frac{7}{4} - \frac{11}{8}]$$

$$I(X;Y) = [\frac{14}{8} - \frac{11}{8}]$$

$$I(X;Y) = \frac{3}{8} \text{ bits}$$

In [47]:

```
# Define the joint distribution as a numpy array
joint_distribution = np.array([
    [1/8, 1/16, 1/16, 1/4],
    [1/16, 1/8, 1/16, 0],
    [1/32, 1/32, 1/16, 0],
    [1/32, 1/32, 1/16, 0]
])

# Calculation of H(X)
px = np.sum(joint_distribution, axis=1)
hx = -np.sum(px * np.log2(px[px > 0]))


# Calculation of H(X|Y)
hxy = (1/4) * (7/4) + (1/4) * (7/4) + (1/4) * 2 + (1/4) * 0

# Calculation of I(X;Y)
ixy = hx - hxy

print(f"Mutual Information between Size and Color (I(X;Y)): {ixy:.2f}
```

```
Mutual Information between Size and Color (I(X;Y)): 0.38 bits
```

Grant, Coulton, and Shengwei aimed to quantify the information shared between the size and color of the island creatures. The mutual information ($I(X;Y)$) between size and color is $\frac{3}{8}$ bits, representing the amount of information about one variable that can be inferred from the other.

💡 **Insightful Discovery:** *Affirmation of Bayes' Rule for Conditional Entropy*

In exploring the equations derived from the problem set, particularly $I(X;Y) = H(Y) - H(Y|X) = H(X) - H(X|Y)$, a fascinating observation emerges. It is hereby made evident that Bayes' rule for conditional entropy, which is given by $H(Y|X) = H(X|Y) - H(X) + H(Y)$, is validated. This finding underscores the profound interplay between mutual information and conditional entropy, shedding light on the inherent relationships within the dataset.