CS181: Computer Vision

Discrete Probability Distribution

Joint probability distribution
Marginal probability distribution
Conditional probability distribution



Md Alimoor Reza
Associate Professor of Computer Science

Probability Basics

- **Probability Distribution**
- **Joint Probability Distribution**
- **Marginal Probability Distribution**
- **Conditional Probability Distribution**

Probability Basics

- **Probability Distribution**
  - Discrete Probabilities
  - Continuous Probabilities

In this lecture, we are going to cover only distributions for discrete events;

- **Probability Distribution and Random Variable**
  - Probability distribution is a function which will depend on a random variable eg, X
  - If the random variable X takes discrete values then you get discrete probabilities

Random Number

- Random numbers are useful many applications:
  - Simulating a coin toss — random flipping of head or tail
  - Simulating a dice roll — random roll of one of six sides
  - Simulating a card shuffling - randomly selecting cards (out of 52)

- Python provides library to generate random numbers
  - You can import random module to get access to random number generating functions

```
import random

rand_number = random.randint(1, 10)
print(rand_number)
```

10 coin toss events — random flipping of head or tail

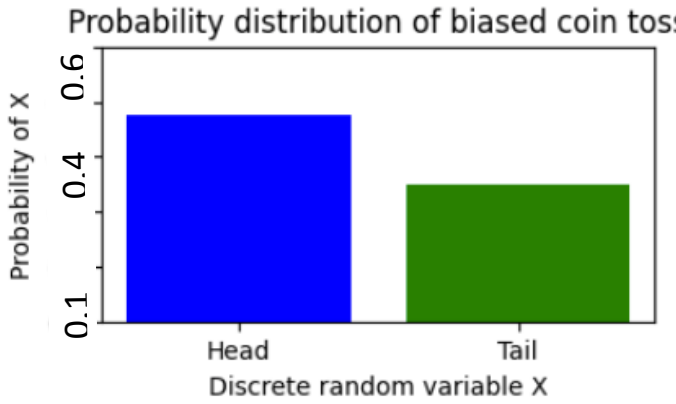
- For example, let's say there are 10 coins tossing events as follows. The **sample space** is the set of *all possible value that* ***X*** *can take*. Then *P(X=head)=?*, *P(X=tail)=?*



10 coin toss event — random flipping of head or tail

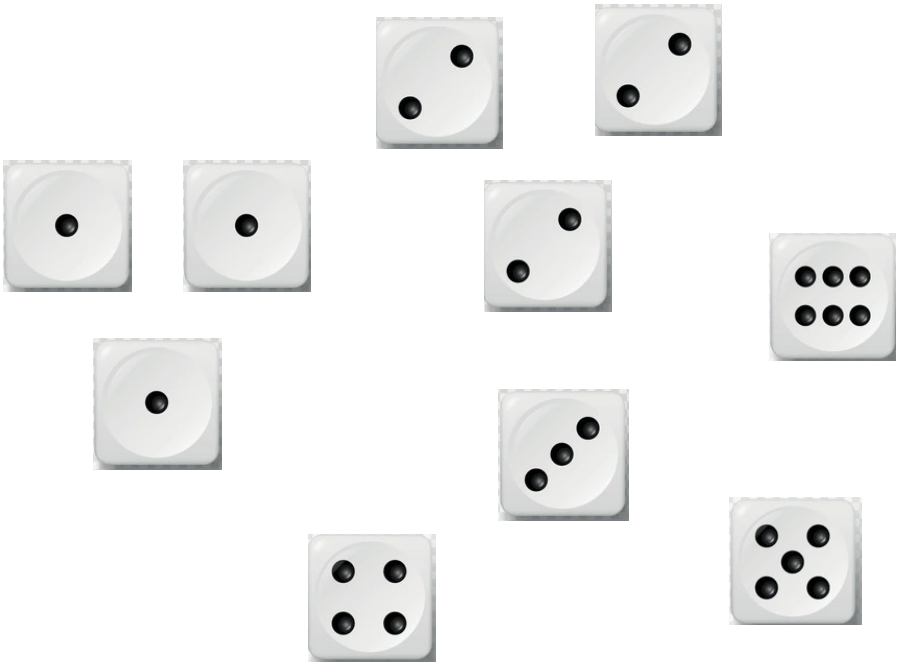
- **Discrete Probability Distribution (1 dimensional event):**
  - Let's assume X is random variable which can take discrete values
  - P(X) is a function that maps from all possible values of X to the probability of the corresponding event. For example:
    - X is a random variable
    - The **sample space** is the set of *all possible value that X can take*.
    - A probability is assigned to the random variable for each value from the sample space.
    - For example, X is a random variable for 10-coin toss events. X can take one of the 2 values from the **sample space** *{Head, Tail}*. Its probability distribution is:

$$P(X = Head) = 0.60$$
$$P(X = Tail) = 0.40$$



10 dice rolling events — random roll of one of six sides

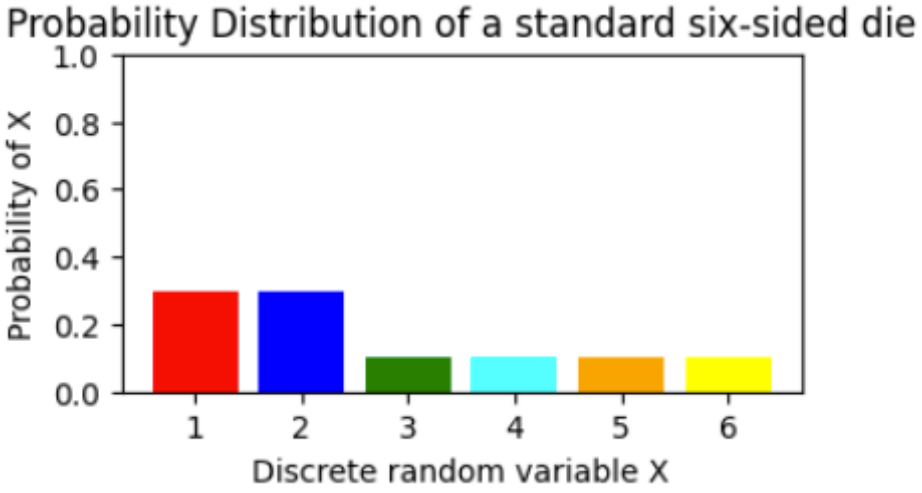
- For example, let's say there are 10-dice rolling events as follows. The **sample space** is the set of *all possible value that* ***X*** *can take*. Then *P(X=1)=?, P(X=2)=?, P(X=3)=?, P(X=4)=?, P(X=5)=?, and P(X=6)=?*.



10 dice rolling events — random roll of one of six sides

- **Discrete Probability Distribution (1 dimensional event):**
  - Let's assume *X* is random variable which can take discrete values
  - P(*X*) is a function that maps from all possible values of X to the probability of the corresponding event.
    - X is a random variable which can take one of the 6 values for a standard six-sided die, ie, the **sample space** of X: *{1, 2, 3, 4, 5, 6}*. Its probability distribution is:

$$P(X = 1) = 0.3$$
$$P(X = 2) = 0.3$$
$$P(X = 3) = 0.1$$
$$P(X = 4) = 0.1$$
$$P(X = 5) = 0.1$$
$$P(X = 6) = 0.1$$



Probability Basics

- **Discrete Probability Distribution (multi-dimensional event):**
  - What about when we have more than two random variables?
    - both random variables may take discrete values.
  - Let's assume there are two random variables, X and Y, each of which can take on discrete values.
  - Essentially, we are transitioning from a one-dimensional probability distribution to a higher-dimensional space. Let's start with a 2-dimensional probability distribution.

Probability Basics

- Probability Distribution
- **Joint Probability Distribution**
- Marginal Probability Distribution
- Conditional Probability Distribution

Probability Basics

- **Discrete Probability Distribution (2 dimensional event):**
  - Let's assume there are two random variables, X and Y, each of which can take on discrete values. So the **sample space** can be shown as table as follows:

1st dimension

2nd dimension

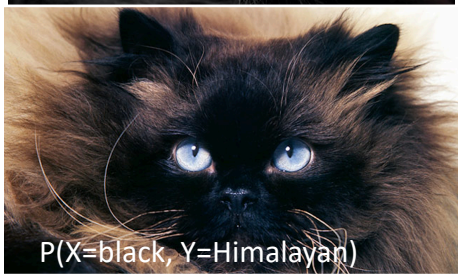
	Discrete value of 1st random variable X	**Discrete value of 1st random variable X**
Discrete value of 2nd random variable Y	$P(X=a_1, Y=b_1)$	$P(X=a_2, Y=b_1)$
Discrete value of 2nd random variable Y	$P(X=a_1, Y=b_2)$	$P(X=a_2, Y=b_2)$

Probability Basics

- **Discrete Probability Distribution (2 dimensional event):**
  - X = random variable indicating color
  - Y = random variable indicating breed
  - What is P(X=white, Y=Persian)?

X = random variable indicating **color**

Y=random variable indicating **breed**

	white	**black**
Persian	 P(X=white, Y=Persian)	 P(X=black, Y=Persian)
Himalayan	 P(X=white, Y=Himalayan)	 P(X=black, Y=Himalayan)

10 events of cat selection with 2-dimensions (color and breed)


Persian White


Himalayan White


Persian Black


Himalayan Black


Persian White


Himalayan White


Himalayan Black


Himalayan Black


Persian White


Himalayan Black

P(X=white, Y=Persian) = ?

P(X=white, Y=Himalayan) = ?

P(X=black, Y=Persian) = ?

P(X=black, Y=Himalayan) = ?

Joint Probability Distribution

- Sample space is set of all possible outcomes of the random variable
  - X = random variable indicating color of the cat
  - Y = random variable indicating breed of the cat

- The full joint probability distribution assigns a probability to each element of the **sample space**

X = random variable indicating **color**

Y=random variable indicating **breed**

	white	black
Persian	 P(X=white, Y=Persian) **0.3**	 P(X=black, Y=Persian) **0.1**
Himalayan	 P(X=white, Y=Himalayan) **0.2**	 P(X=black, Y=Himalayan) **0.4**

Joint Probability Distribution

- Sample space is set of all possible outcomes of the random variable
  - X = random variable indicating color of the cat
  - Y = random variable indicating breed of the cat

- The full joint probability distribution assigns a probability to each element of the sample space. You can also list the probabilities of the sample space in the format below:

X	Y	P(X,Y)
White	Persian	0.3
White	Himalayan	0.2
Black	Persian	0.1
Black	Himalayan	0.4

Joint Probability Distribution

- The joint probability distribution expresses probability distribution of observing varied instance of (x, y) where some paired outcome occurs more frequently than others
  - X = random variable with three discrete values {a, b, c}
  - Y = random variable with three discrete values { α , β , γ }
- The joint probability distribution P(X, Y) is expressed using the table

	X=a	X=b	X=c
Y= α	0.1	0.05	0.1
Y= β	0.1	0.3	0.1
Y= γ	0.05	0.05	0.15

Probability Basics

- Probability Distribution
- Joint Probability Distribution
- **Marginal Probability Distribution**
- Conditional Probability Distribution

Marginal Probability Distribution

- It expresses **probability distribution of any single random variable** from a joint probability distribution
  - What is the **probability distribution of the breed** of the cat?
    - You need to find the probability of the cat being a Persian
    - You need to find the probability of the cat being a Himalayan

X = random variable indicating **color**

Y=random variable indicating **breed**

	white	black
Persian	 0.3	 0.1
Himalayan	 0.2	 0.4

finds the probability of X happening regardless of Y taking any value.

Marginal Probability Distribution

- It expresses **probability distribution of any single random variable** from a joint probability distribution
  - What is the **probability distribution of the color** of the cat?
    - You need to find the probability of the cat being a white
    - You need to find the probability of the cat being a black

X = random variable indicating **color**

Y=random variable indicating **breed**

	white	black
Persian	 0.3	 0.1
Himalayan	 0.2	 0.4

Marginal Probability Distribution

finds the probability of X happening regardless of Y taking any value.

- The marginal probability distribution expresses probability distribution of any single random variable from a joint probability distribution by summing over all other variables

$$P(X) = \sum_{y} P(X, Y = y)$$

	X=a	X=b	X=c
Y= α	0.1	0.05	0.1
Y= β	0.1	0.3	0.1
Y= γ	0.05	0.05	0.15

finds the probability of X happening regardless of Y taking any value.

Marginal Probability Distribution

	X=a	X=b	X=c
Y=α	0.1	0.05	0.1
Y=β	0.1	0.3	0.1
Y=γ	0.05	0.05	0.15

$$P(X) = P(X, Y = \alpha) + P(X, Y = \beta) + P(X, Y = \gamma)$$

- Aggregating individual marginal probability values into a histogram to make the probability distribution as follows:

  [P(X=a), P(X=b), P(X=c)]

- Marginal distribution has the following interpretation — "It finds the probability of X happening regardless of Y taking any value of α β γ

finds the probability of X happening regardless of Y taking any value.

Exercise 1: Marginal Probability Distribut

	X=a	X=b	X=c
Y=α	0.1	0.05	0.1
Y=β	0.1	0.3	0.1
Y=γ	0.05	0.05	0.15

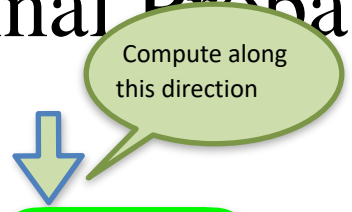
$$P(X) = P(X, Y = \alpha) + P(X, Y = \beta) + P(X, Y = \gamma)$$

- Find the marginal distribution P(X)?

$$P(X = a) = ?$$
$$P(X = b) = ?$$
$$P(X = c) = ?$$

finds the probability of X happening regardless of Y taking any value.

Solution: Marginal Probability Distributi



	X=a	X=b	X=c
Y=α	0.1	0.05	0.1
Y=β	0.1	0.3	0.1
Y=γ	0.05	0.05	0.15

$$P(X) = P(X, Y = \alpha) + P(X, Y = \beta) + P(X, Y = \gamma)$$

- Find the marginal distribution P(X=a)?

$$P(X = a) = P(X = a, Y = \alpha) + (X = a, Y = \beta) + (X = a, Y = \gamma)$$

$$= ?$$

finds the probability of X happening regardless of Y taking any value.

Solution: Marginal Probability Distributi

	X=a	X=b	X=c
Y=α	0.1	0.05	0.1
Y=β	0.1	0.3	0.1
Y=γ	0.05	0.05	0.15

$$P(X) = P(X, Y = \alpha) + P(X, Y = \beta) + P(X, Y = \gamma)$$

- Find the marginal distribution P(X=a)?

$$\begin{aligned} P(X = a) &= P(X = a, Y = \alpha) + (X = a, Y = \beta) + (X = a, Y = \gamma) \\ &= 0.1 + (X = a, Y = \beta) + (X = a, Y = \gamma) \\ &= 0.1 + 0.1 + (X = a, Y = \gamma) \\ &= 0.1 + 0.1 + 0.05 \\ &= 0.25 \end{aligned}$$

finds the probability of X happening regardless of Y taking any value.

Solution: Marginal Probability Distribution

Compute along this direction

	X=a	X=b	X=c
Y=α	0.1	0.05	0.1
Y=β	0.1	0.3	0.1
Y=γ	0.05	0.05	0.15

$$P(X) = P(X, Y = \alpha) + P(X, Y = \beta) + P(X, Y = \gamma)$$

- Find the marginal distribution P(X=b)?

$$P(X = b) = P(X = b, Y = \alpha) + (X = b, Y = \beta) + (X = b, Y = \gamma)$$

$$= ?$$

finds the probability of X happening regardless of Y taking any value.

Solution: Marginal Probability Distributi

	X=a	X=b	X=c
Y=α	0.1	0.05	0.1
Y=β	0.1	0.3	0.1
Y=γ	0.05	0.05	0.15

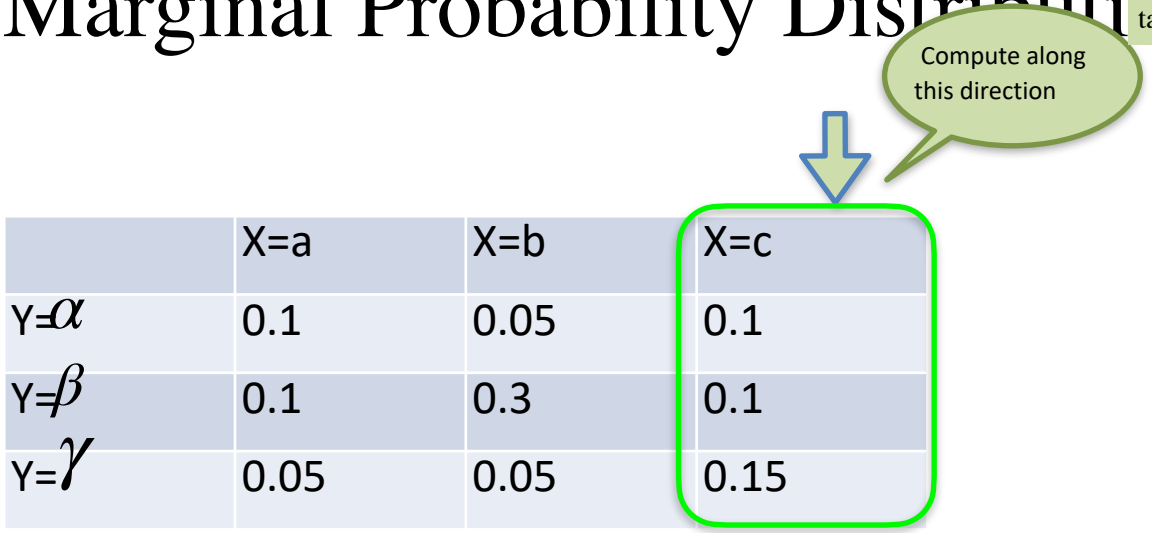
$$P(X) = P(X, Y = \alpha) + P(X, Y = \beta) + P(X, Y = \gamma)$$

- Find the marginal distribution P(X=b)?

$$\begin{aligned} P(X = b) &= P(X = b, Y = \alpha) + (X = b, Y = \beta) + (X = b, Y = \gamma) \\ &= 0.05 + (X = b, Y = \beta) + (X = b, Y = \gamma) \\ &= 0.05 + 0.3 + (X = b, Y = \gamma) \\ &= 0.05 + 0.3 + 0.05 \\ &= 0.40 \end{aligned}$$

finds the probability of X happening regardless of Y taking any value.

Solution: Marginal Probability Distributi



$$P(X) = P(X, Y = \alpha) + P(X, Y = \beta) + P(X, Y = \gamma)$$

- Find the marginal distribution P(X=c)?

$$P(X = c) = P(X = c, Y = \alpha) + (X = c, Y = \beta) + (X = c, Y = \gamma)$$

$$= ?$$

finds the probability of X happening regardless of Y taking any value.

Solution: Marginal Probability Distributi

	X=a	X=b	X=c
Y=α	0.1	0.05	0.1
Y=β	0.1	0.3	0.1
Y=γ	0.05	0.05	0.15

$$P(X) = P(X, Y = \alpha) + P(X, Y = \beta) + P(X, Y = \gamma)$$

- Find the marginal distribution P(X=c)?

$$\begin{aligned} P(X = c) &= P(X = c, Y = \alpha) + (X = c, Y = \beta) + (X = c, Y = \gamma) \\ &= 0.10 + (X = c, Y = \beta) + (X = c, Y = \gamma) \\ &= 0.10 + 0.10 + (X = c, Y = \gamma) \\ &= 0.10 + 0.10 + 0.15 \\ &= 0.35 \end{aligned}$$

finds the probability of X happening regardless of Y taking any value.

Marginal Probability Distribution

	X=a	X=b	X=c
Y=α	0.1	0.05	0.1
Y=β	0.1	0.3	0.1
Y=γ	0.05	0.05	0.15

$$P(X) = P(X, Y = \alpha) + P(X, Y = \beta) + P(X, Y = \gamma)$$

- Aggregating individual marginal probability values into a histogram to make the probability distribution as follows:

[P(X=a), P(X=b), P(X=c)]

$[0.25, 0.40, 0.35]$

It sums up to 1.0, which satisfies the property of a probability distribution.

- Marginal distribution has the following interpretation — "It finds the probability of X happening regardless of Y taking any value of α β γ

Exercise 2: Marginal Probability Distribut

finds the probability of Y happening regardless of X taking any value.

Compute along these directions one after another

	X=a	X=b	X=c
Y=α	0.1	0.05	0.1
Y=β	0.1	0.3	0.1
Y=γ	0.05	0.05	0.15

$$P(Y) = P(X = a, Y) + (X = b, Y) + (X = c, Y)$$

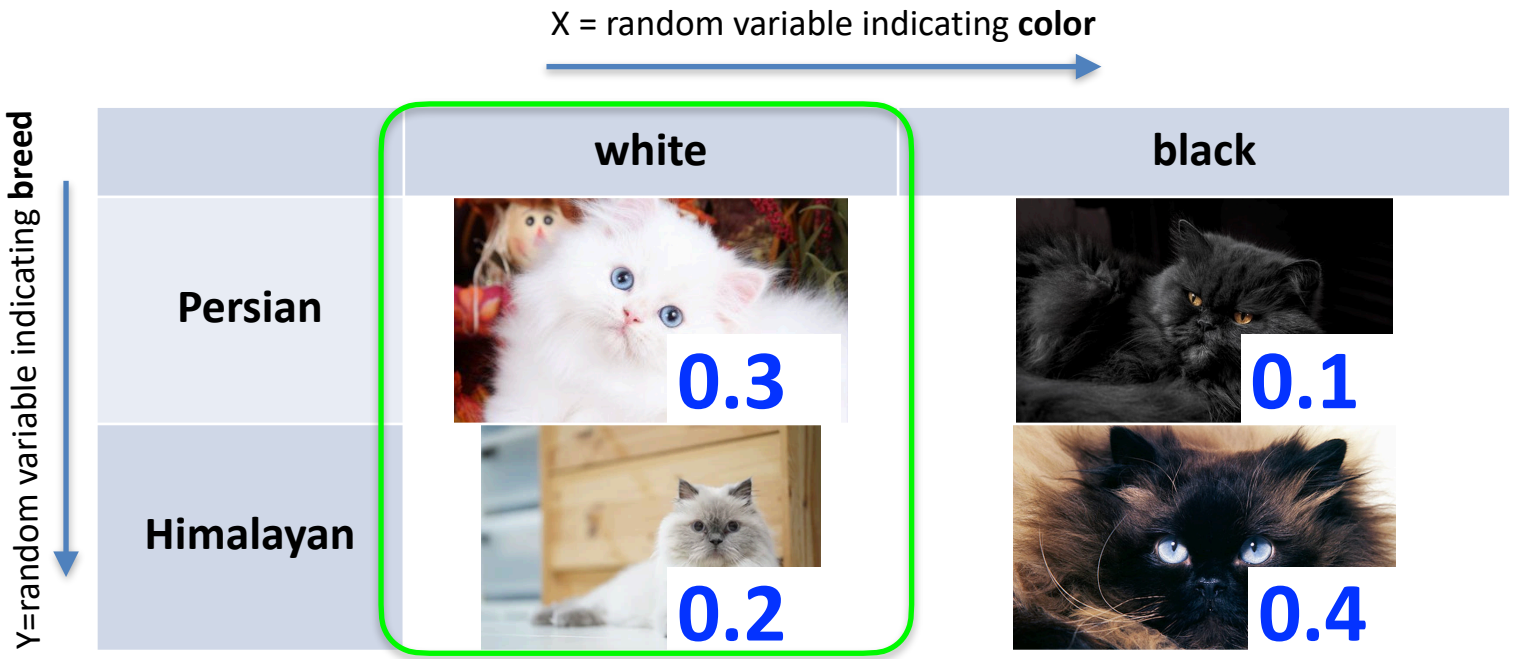
- Similarly, can you find the marginal distribution P(Y)?

$$P(Y = \alpha) = ?$$
$$P(Y = \beta) = ?$$
$$P(Y = \gamma) = ?$$

finds the probability of **color** happening regardless of **breed** taking any value.

Group Exercise

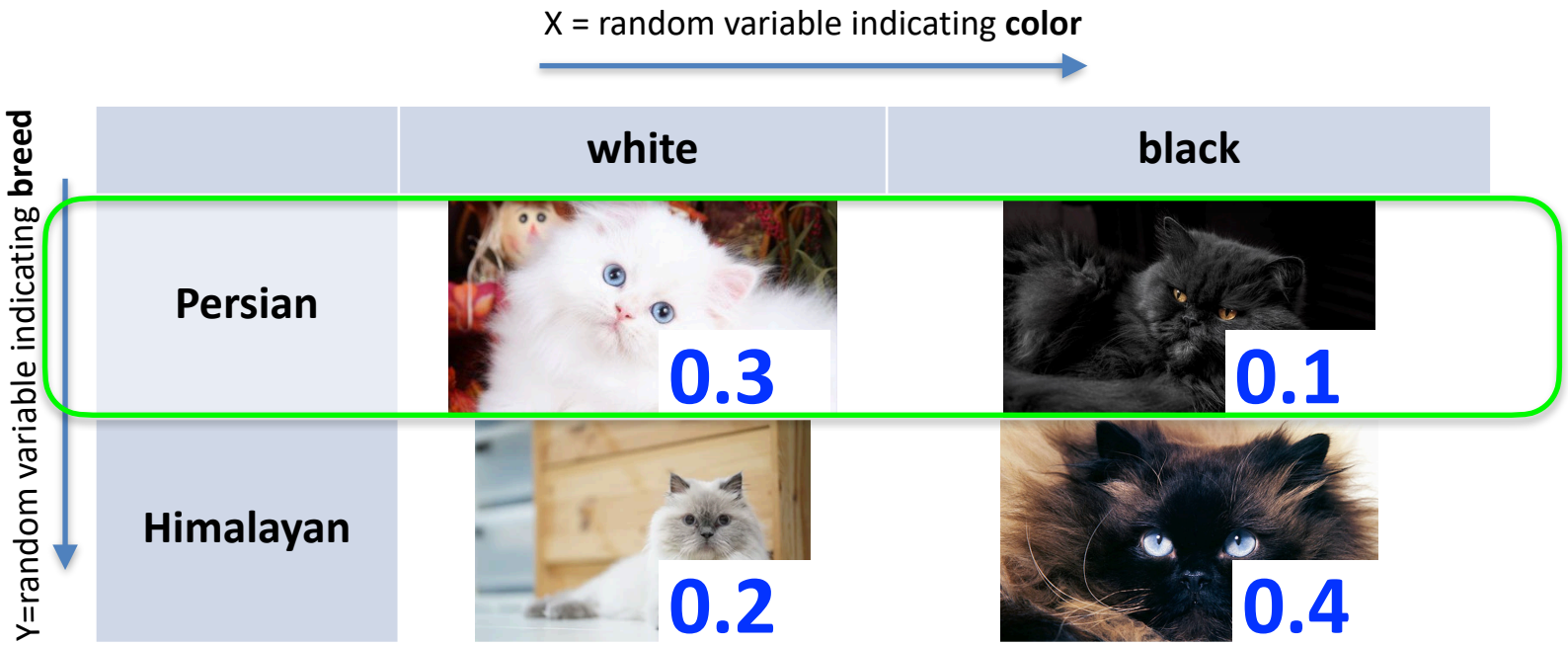
- What is the **probability distribution of the color** of a cat?
- P(X=color)=?
  - You need to find the probability of the cat being a white
  - You need to find the probability of the cat being a black



finds the probability of **breed** happening regardless of **color** taking any value.

Group Exercise

- What is the **probability distribution of the breed** of a cat?
- P(X=breed)=?
  - You need to find the probability of the cat being a Persian
  - You need to find the probability of the cat being a Himalayan



Probability Basics

- Probability Distribution
- Joint Probability Distribution
- Marginal Probability Distribution
- **Conditional Probability Distribution**

Conditional Probability Distribution

finds the probability of X happening after observing other event Y=alpha normalized by the marginal probability of that other event P(Y=alpha)

- **It tells us the relative propensity**
- Conditional probability distribution of any single random variable *(eg, X)* conditioned on other variable's *(eg, Y)* value fixed to a particular value
- Conditional distribution has the following form:

$$P(X \mid Y = \alpha)$$

	X=a	X=b	X=c
Y=α	0.1	0.05	0.1
Y=β	0.1	0.3	0.1
Y=γ	0.05	0.05	0.15

- Individual terms of the conditional distribution has the following form:

$$P(X = a \mid Y = \alpha) = \frac{P(X = a, Y = \alpha)}{P(Y = \alpha)}$$

We computed this marginal probability term earlier

$$= \frac{P(X = a, Y = \alpha)}{P(X = a, Y = \alpha) + P(X = b, Y = \alpha) + P(X = c, Y = \alpha)}$$

If you don't recall, here is how it can be recomputed

Conditional Probability Distribution

finds the probability of X happening after observing other event Y=alpha normalized by the marginal probability of that other event P(Y=alpha)

- It tells us the relative propensity
- Conditional probability distribution of any single random variable *(eg, X)* conditioned on other variable's *(eg, Y)* value fixed to a particular value

	X=a	X=b	X=c
Y=α	0.1	0.05	0.1
Y=β	0.1	0.3	0.1
Y=γ	0.05	0.05	0.15

- In this example, you need to compute three conditional probabilities to form a valid distribution

$$P(X = a \mid Y = \alpha) = ?$$

$$P(X = b \mid Y = \alpha) = ?$$

$$P(X = c \mid Y = \alpha) = ?$$

Conditional Probability Distribution

finds the probability of X happening after observing other event Y=alpha normalized by the marginal probability of that other event P(Y=alpha)

- It tells us the relative propensity
- Conditional probability distribution of any single random variable *(eg, X)* conditioned on other variable's *(eg, Y)* value fixed to a particular value

	X=a	X=b	X=c
Y=α	0.1	0.05	0.1
Y=β	0.1	0.3	0.1
Y=γ	0.05	0.05	0.15

We computed this marginal probability term earlier

$$P(Y = \alpha) = 0.25$$

- In this example, you need to compute three conditional terms to make it a distribution

$$P(X = a \mid Y = \alpha) = \frac{P(X = a, Y = \alpha)}{P(Y = \alpha)}$$

$$P(X = b \mid Y = \alpha) = \frac{P(X = b, Y = \alpha)}{P(Y = \alpha)}$$

$$P(X = c \mid Y = \alpha) = \frac{P(X = c, Y = \alpha)}{P(Y = \alpha)}$$

Exercise 3: Conditional Probability Distrib

finds the probability of X happening after observing other event Y=alpha normalized by the marginal probability of that other event P(Y=alpha)

	X=a	X=b	X=c
Y=α	0.1	0.05	0.1
Y=β	0.1	0.3	0.1
Y=γ	0.05	0.05	0.15

$$P(X|Y=\alpha)$$

- Aggregating individual conditional probability values into a histogram to make the probability distribution as follows:

$$[P(X=a\,|\,Y=\alpha), P(X=b\,|\,Y=\alpha), P(X=c\,|\,Y=\alpha)]$$

Exercise 3: Conditional Probability Distrib

finds the probability of X happening after observing other event Y=alpha normalized by the marginal probability of that other event P(Y=alpha)

- It tells us the relative propensity
- Conditional probability distribution of any single random variable *(eg, X)* conditioned on other variable's *(eg, Y)* **value fixed to a particular value**

$$P(X = a \mid Y = \alpha) = ?$$

$$= \frac{P(X = a, Y = \alpha)}{P(Y = \alpha)}$$

$$= \frac{0.1}{P(Y = \alpha)}$$

$$= \frac{0.1}{0.25}$$

$$= 0.4$$

We computed this marginal probability term earlier

$$P(Y = \alpha) = 0.25$$

	X=a	X=b	X=c
Y=α	0.1	0.05	0.1
Y=β	0.1	0.3	0.1
Y=γ	0.05	0.05	0.15

Exercise 3: Conditional Probability Distrib

finds the probability of X happening after observing other event Y=alpha normalized by the marginal probability of that other event P(Y=alpha)

- It tells us the relative propensity
- Conditional probability distribution of any single random variable *(eg, X)* conditioned on other variable's *(eg, Y)* **value fixed to a particular value**

$$P(X = b \mid Y = \alpha) = ?$$

$$= \frac{P(X = b, Y = \alpha)}{P(Y = \alpha)}$$

$$= \frac{0.05}{P(Y = \alpha)}$$

$$= \frac{0.05}{0.25}$$

$$= 0.2$$

We computed this marginal probability term earlier

$$P(Y = \alpha) = 0.25$$

	X=a	X=b	X=c
Y=α	0.1	0.05	0.1
Y=β	0.1	0.3	0.1
Y=γ	0.05	0.05	0.15

Exercise 3: Conditional Probability Distrib

finds the probability of X happening after observing other event Y=alpha normalized by the marginal probability of that other event P(Y=alpha)

- It tells us the relative propensity
- Conditional probability distribution of any single random variable *(eg, X)* conditioned on other variable's *(eg, Y)* **value fixed to a particular value**

$$P(X = c \mid Y = \alpha) = ?$$

$$= \frac{P(X = c, Y = \alpha)}{P(Y = \alpha)}$$

$$= \frac{0.1}{P(Y = \alpha)}$$

$$= \frac{0.1}{0.25}$$

$$= 0.4$$

We computed this marginal probability term earlier

$$P(Y = \alpha) = 0.25$$

	X=a	X=b	X=c
Y=α	0.1	0.05	0.1
Y=β	0.1	0.3	0.1
Y=γ	0.05	0.05	0.15

Exercise 3: Conditional Probability Distrib

finds the probability of X happening after observing other event Y=alpha normalized by the marginal probability of that other event P(Y=alpha)

	X=a	X=b	X=c
Y=α	0.1	0.05	0.1
Y=β	0.1	0.3	0.1
Y=γ	0.05	0.05	0.15

$$P(X \mid Y = \alpha)$$

- Aggregating individual conditional probability values into a histogram to make the probability distribution as follows:

$$[P(X = a \mid Y = \alpha), P(X = b \mid Y = \alpha), P(X = c \mid Y = \alpha)]$$

[0.4, 0.2, 0.4]

Conditional Probability Distribution

- It expresses **probability distribution of any single random variable conditioned** on other random variables
- **It tells us the relative propensity**
  - What is the probability of it's **color** when you picked **a Persian cat**?
    - What is the probability it is a White: P(Y=White| X=Persian) = ?
    - What is the probability it is a Black: P(Y=Black| X=Persian) = ?

	white	black
Persian	 0.3	 0.1
Himalayan	 0.2	 0.4

Probability Basics

- **Probability Distribution**
- **Joint Probability Distribution**
- **Marginal Probability Distribution**
- **Conditional Probability Distribution**