

Announcements

HW1 & Quiz 1

- Graded
- Regrade requests
- Hours, performance

HW2

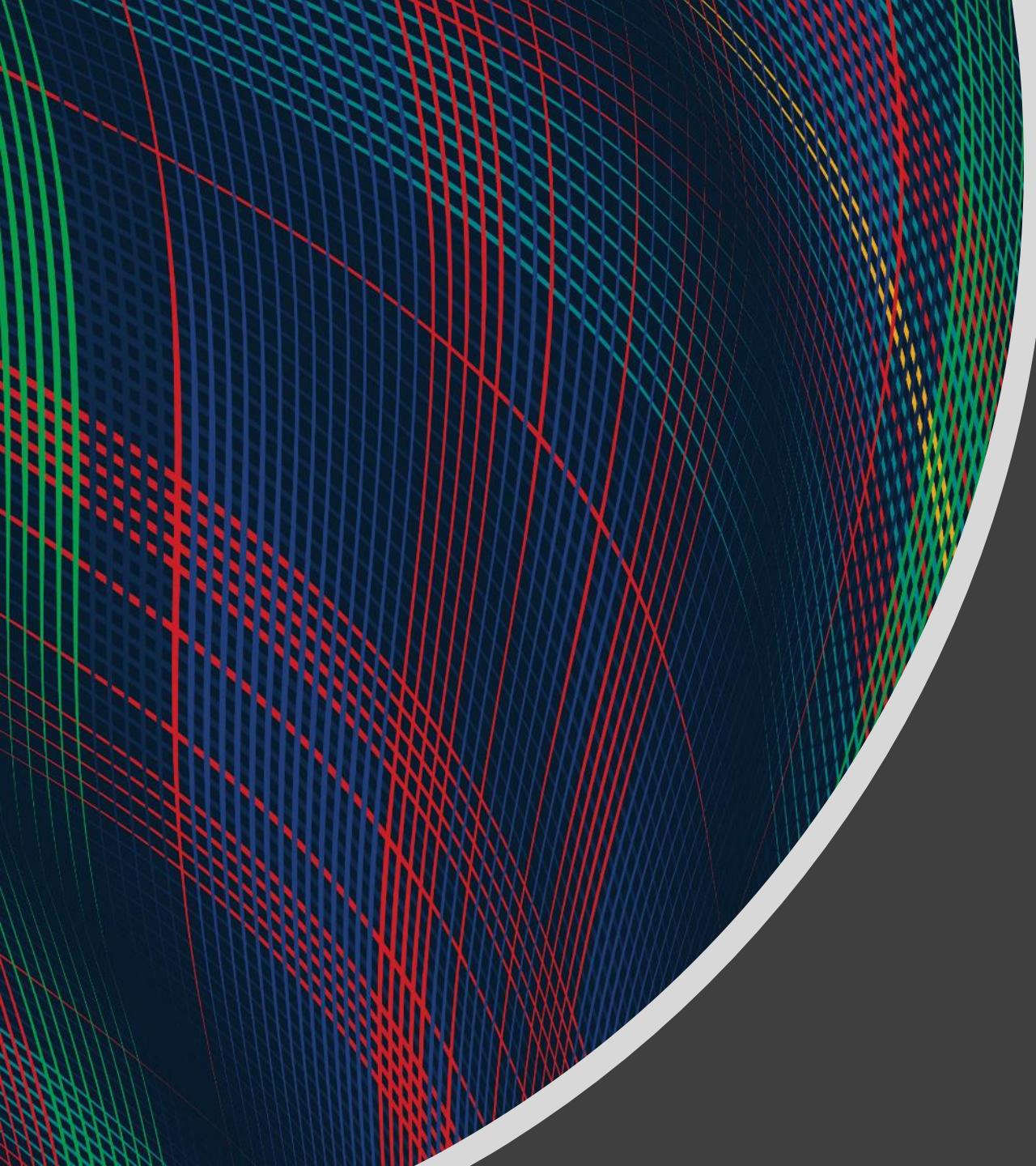
- Due Sat 10/2

Quizzes

- Mon 10/4, last 15 min. of class (calculus, optimization, Lagrange)
- Mon 10/11, last 15 min. of class (probability, statistics)

Constrained Optimization Clarification

Previous lecture slides



Mathematical Foundations for Machine Learning

Probability

Instructor: Pat Virtue

Plan

Today

Probability

- Vocab
- Notation
- Discrete distributions
- Continuous distributions
- Chain rule
- Independence

Probability

Probability Vocab

Outcomes

Sample space

Events

Probability

Random variable

Discrete random variable

Continuous random variable

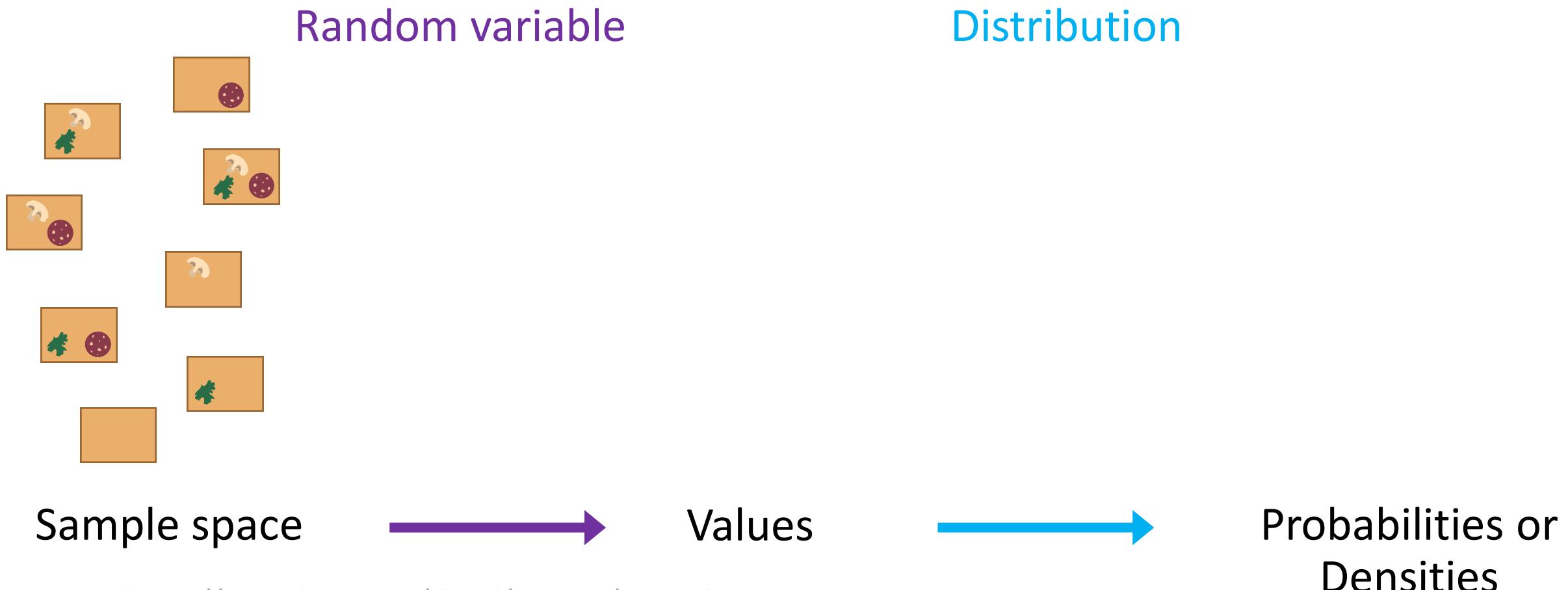
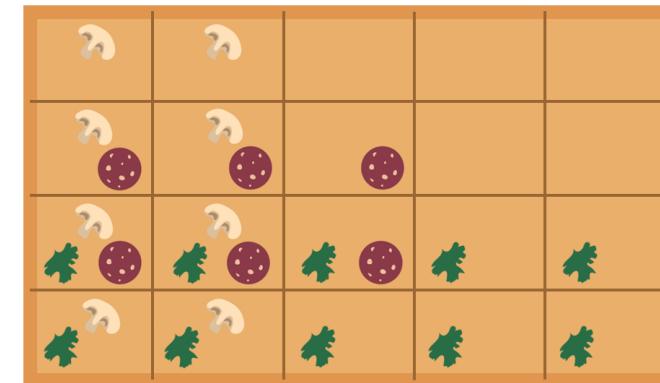
Probability mass function

Probability density function

Parameters

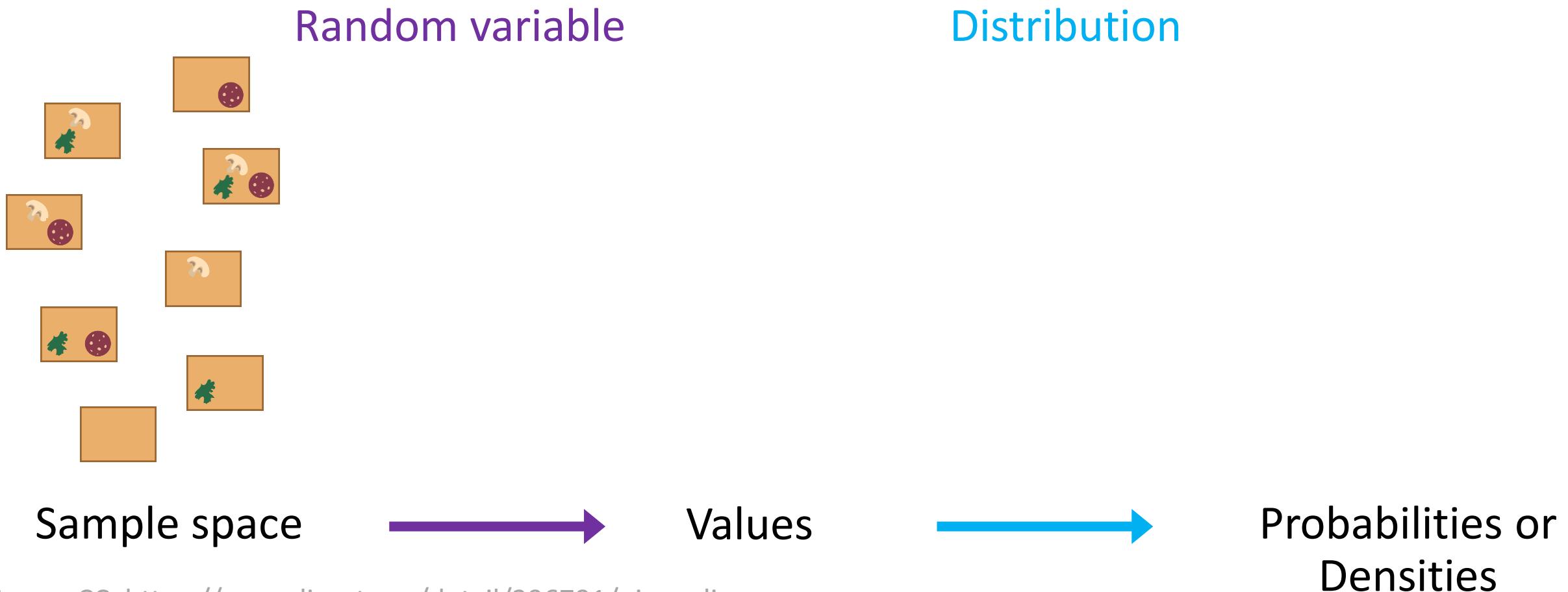
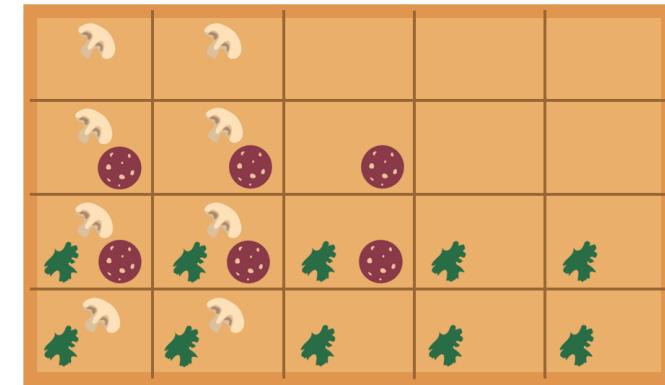
Example

Random variable for spinach or no spinach



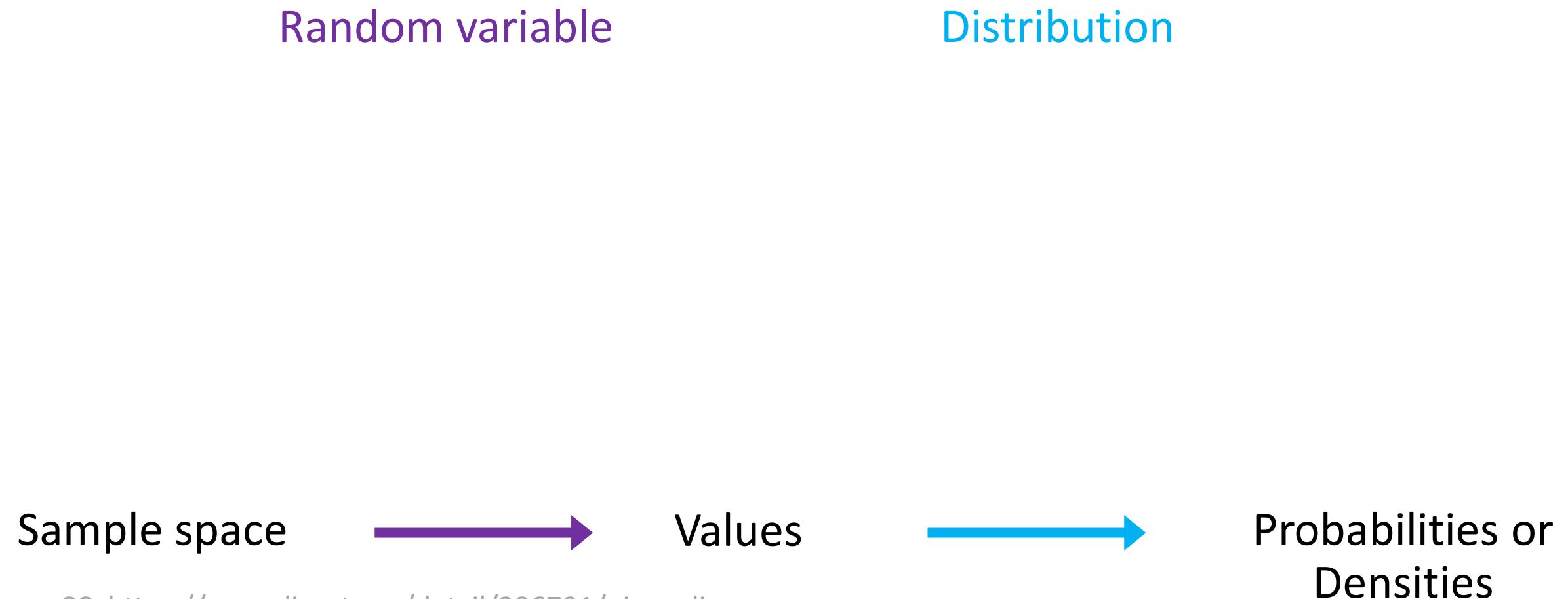
Example

Random variable for topping type with three categories: none, non-meat, meat



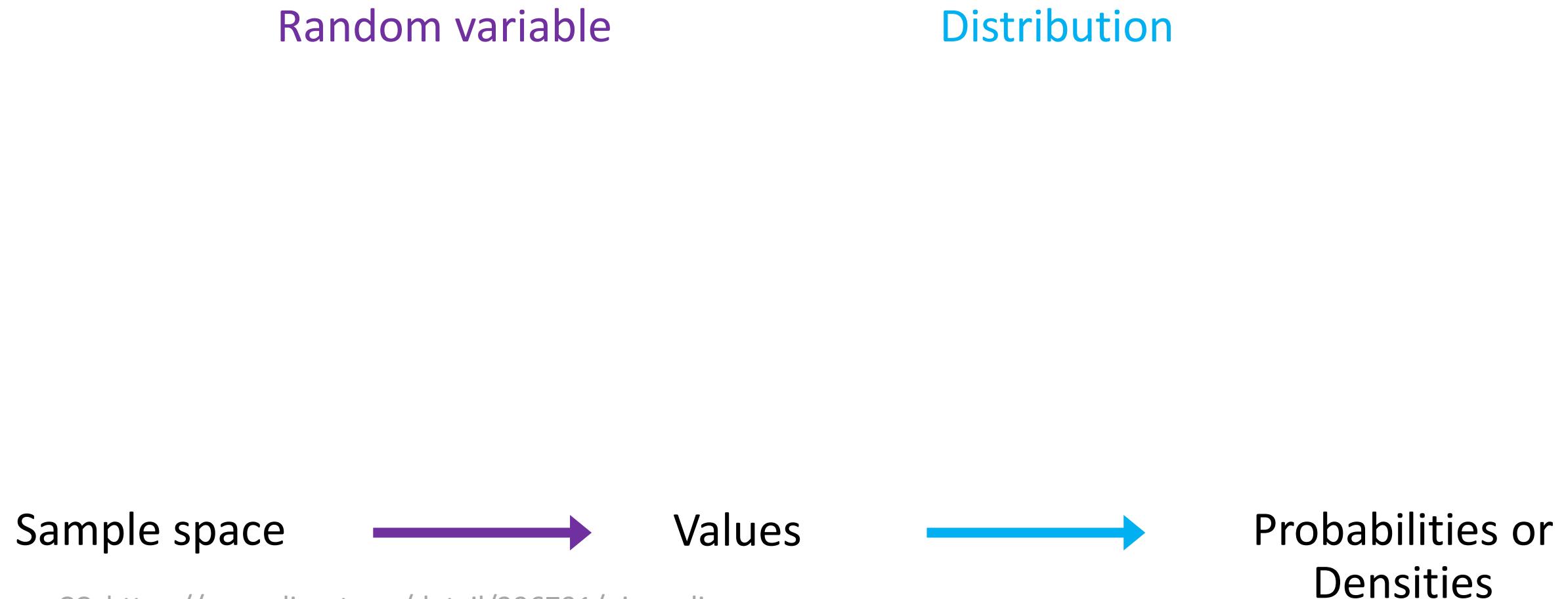
Example

Random variable for number of heads after
two flips of a fair coin



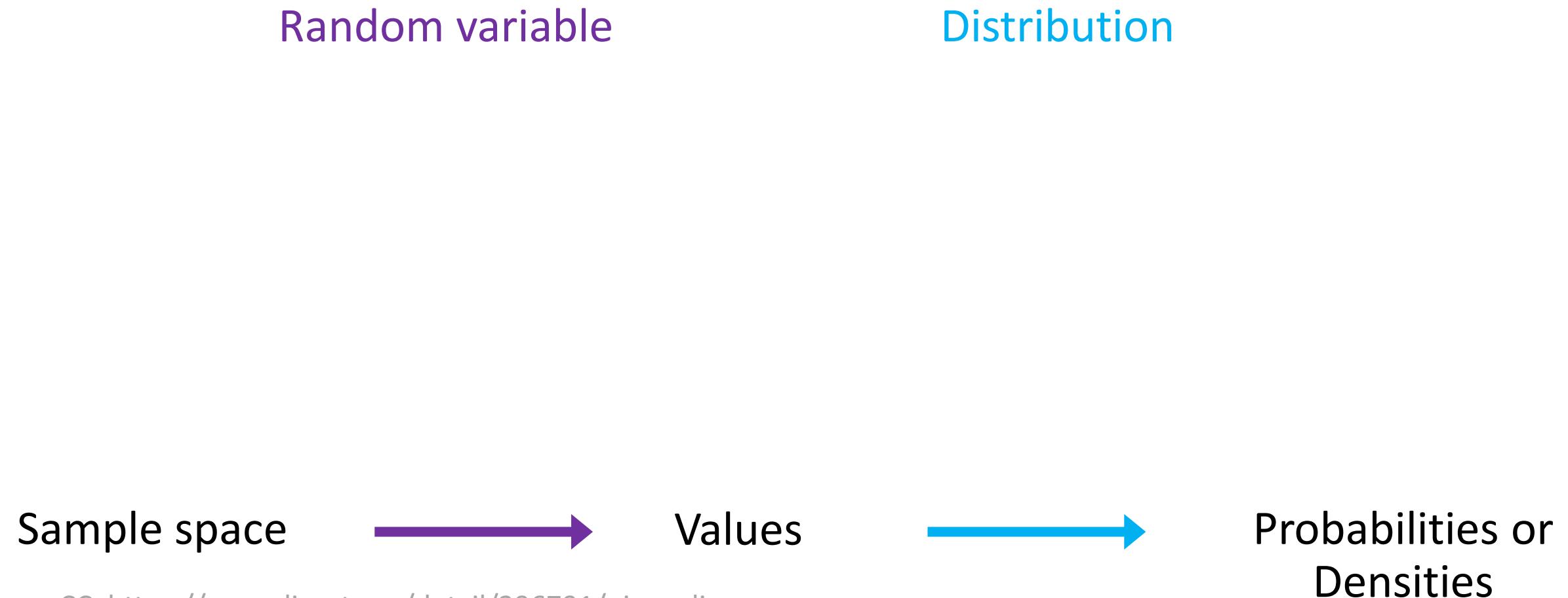
Example

Random variable for number of heads after
two flips of a *biased* coin that lands heads 75%



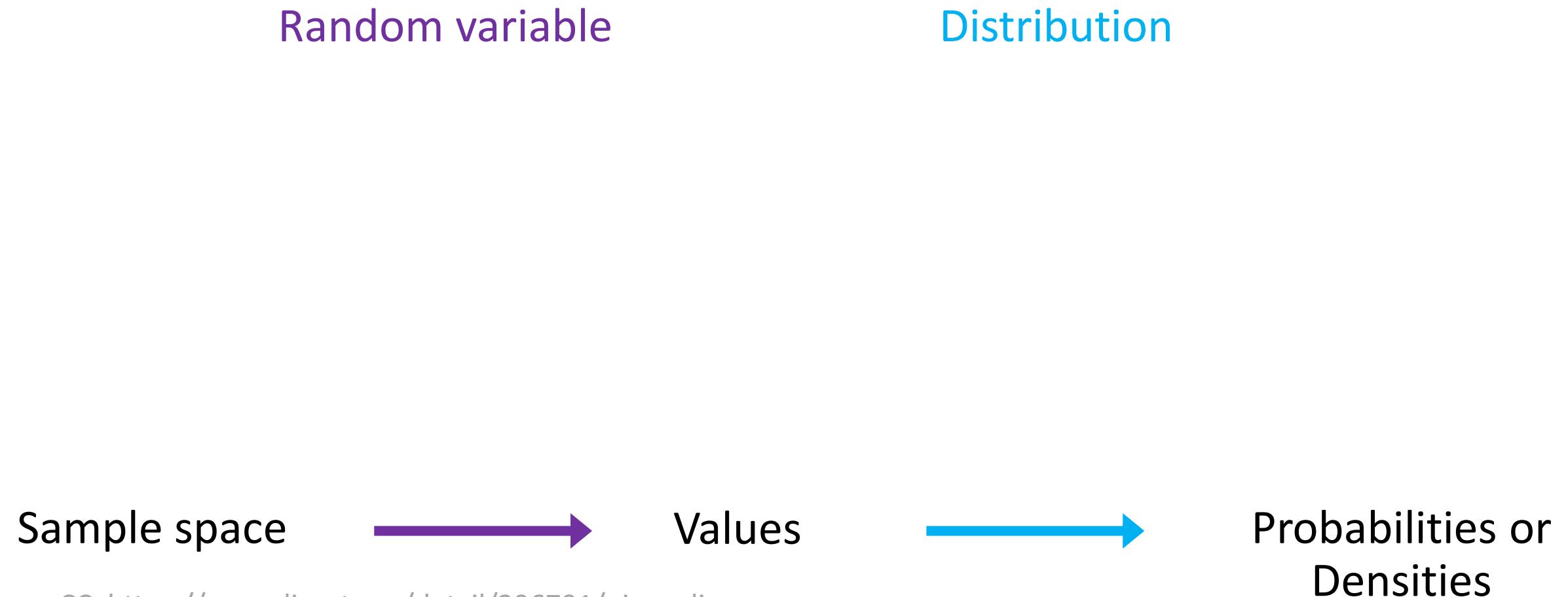
Example

Random variable for cat in picture or not



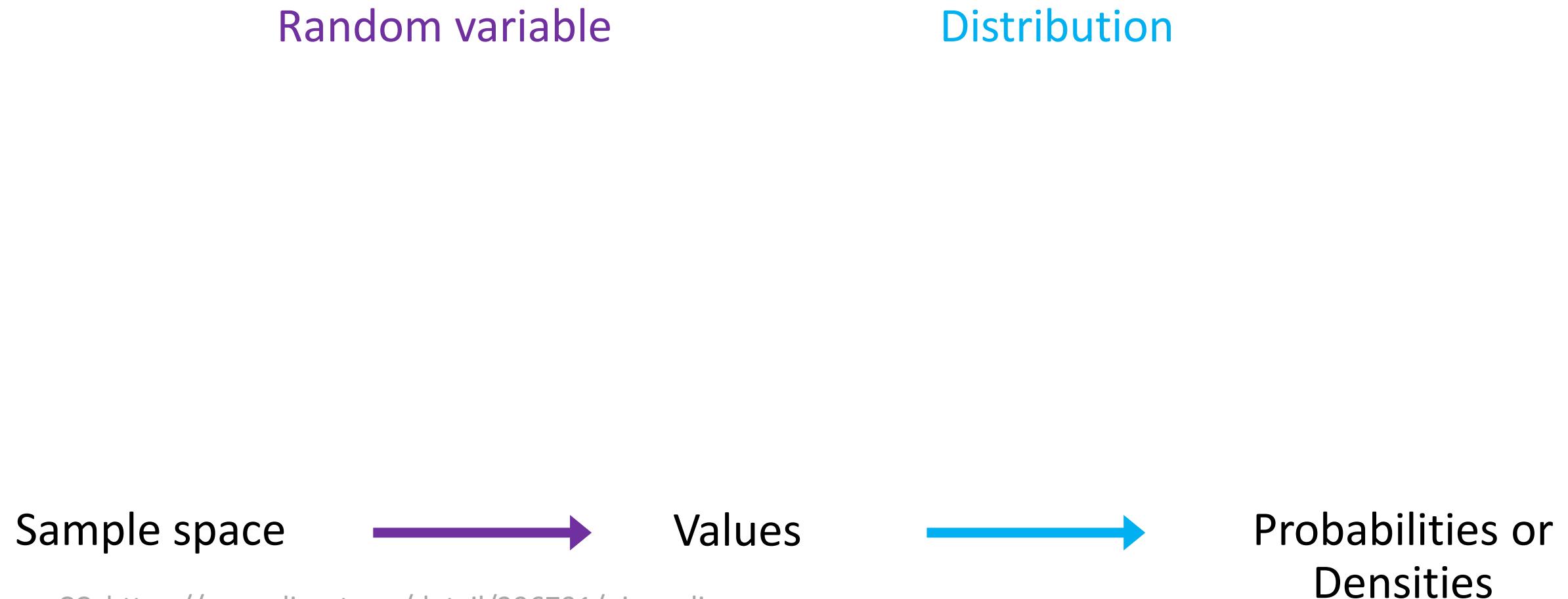
Example

Random variable for animal species in picture assuming one animal picture and available species: dog, cat, pig



Example

Random variable for height of student



Probability Vocab

Outcomes

Sample space

Events

Probability

Random variable

Discrete random variable

Continuous random variable

Probability mass function

Probability density function

Parameters

Example Discrete Distributions

Bernoulli

Categorical

Binomial

Multinomial

Uniform

Example Continuous Distributions

Gaussian

Beta

Laplace

Probability Vocab

Marginal

Joint

Conditional

Notation

[Notation doc](#)

Probability Toolbox

- Algebra
- Three axioms of probability
- Theorem of total probability
- Definition of conditional probability
- Product rule
- Bayes' theorem
- Chain rule
- Independence
- Conditional independence

Tools Summary

Adding to our toolbox

1. Definition of conditional probability

$$P(A|B) = \frac{P(A, B)}{P(B)}$$

2. Product Rule

$$P(A, B) = P(A|B)P(B)$$

3. Bayes' theorem

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}$$

4. Chain Rule...