Package "crp": Chinese Restaurant Process

Ruben Abbout & Jack Potrykus

Mar 14, 2021

CONTENTS:

class chinese_restaurant_process.**ChineseRestaurantProcess**(*alpha*)

Initializes and provides utilities to draw from Chinese Restaurant Process

animate()

Animate the progress of the CRT by looping through self.history and producing a bar plot at each interval

get_table_dict()

Get a dictionary of pairs from get_table_names() and get_table_sizes()

Returns keys = table ids (int), values = # of customers at each table (int)

Return type dict

get_table_names()

Get an array of the names of each table

Returns array of the table names

Return type np.array

get_table_sizes()

Get an array of the number of people at each table

Returns array of number of people at each table

Return type np.array

iter(*niter*)

Advance the process niter times.

Parameters **niter** (*int*) – number of draws to make from the Process

Returns Returns self, so you can call it inline with assignment, i.e. crp = ChineseRestaurantProcess(alpha=1).iter(100)

Return type *ChineseRestaurantProcess*

to_pandas()

Produce a pd.DataFrame object summarizing results of simulation

visualize()

Visualize the final state of the Process as a bar plot

Returns Bar plot as described

Return type plt.figure

class chinese_restaurant_mixture.**ChineseRestaurantMixture**(*alpha*, *param_prior*, *sampler*)

Inherits from ChineseRestaurantProcess class, with the extra utility provided simply being a wrapper around the Process.

animate(*clear=False*)

Animate the progress of the CRT by looping through self.history and producing a bar plot at each interval

reset()

Reset the ChineseRestaurantMixture by clearing self.datapoints

sample(*sample_size*, *reset=False*)

Sample sample_size points from a Chinese Restaurant Mixture process.

Parameters

- **sample_size** (*int*) – number of points to sample
- **reset** (*bool*) – whether or not to reset self.datapoints before proceeding

visualize(*first_n=None*, *clear=False*)

Visualize the final state of the Mixture Process as a kernel density estimate

Parameters **first_n** (*int*) – (optional) visualize the first first_n datapoints (to be used in animate() method). Default behavior is to use all datapoints

Returns Kernel density estimate of datapoints

Return type plt.figure

CHAPTER

ONE

INDICES AND TABLES

- genindex
- modindex
- search

PYTHON MODULE INDEX

C

INDEX