

03_dirichlet_distribution

September 29, 2021

1 Latent Dirichlet Allocation - Interactive Simulation

1.1 Imports

```
[1]: import warnings
warnings.filterwarnings('ignore')
```

```
[2]: %matplotlib inline

import numpy as np
import pandas as pd

# Visualization
from ipywidgets import interact, FloatSlider
import matplotlib.pyplot as plt
import seaborn as sns
```

```
[3]: sns.set_style('white')
```

1.2 Simulate Dirichlet Distribution

```
[4]: f = FloatSlider(value=1,
                    min=1e-2,
                    max=10,
                    step=1e-2,
                    continuous_update=False,
                    description='Alpha')

@interact(alpha=f)
def sample_dirichlet(alpha):
    topics = 10
    draws = 9
    alphas = np.full(shape=topics, fill_value=alpha)
    samples = np.random.dirichlet(alpha=alphas, size=draws)

    fig, axes = plt.subplots(nrows=3,
```

```

        ncols=3,
        sharex=True,
        sharey=True,
        figsize=(14, 8))

axes = axes.flatten()
plt.setp(axes, ylim=(0, 1))
for i, sample in enumerate(samples):
    axes[i].bar(x=list(range(10)), height=sample,
               color=sns.color_palette("Set2", 10))
fig.suptitle('Dirichlet Allocation | 10 Topics, 9 Samples')
sns.despine()
fig.tight_layout()
plt.subplots_adjust(top=.95)

```

```

interactive(children=(FloatSlider(value=1.0, continuous_update=False,
    ↪description='Alpha', max=10.0, min=0.01,...

```