



Indian Buffet Process

Xiaolong Wang and Daniel Khashabi

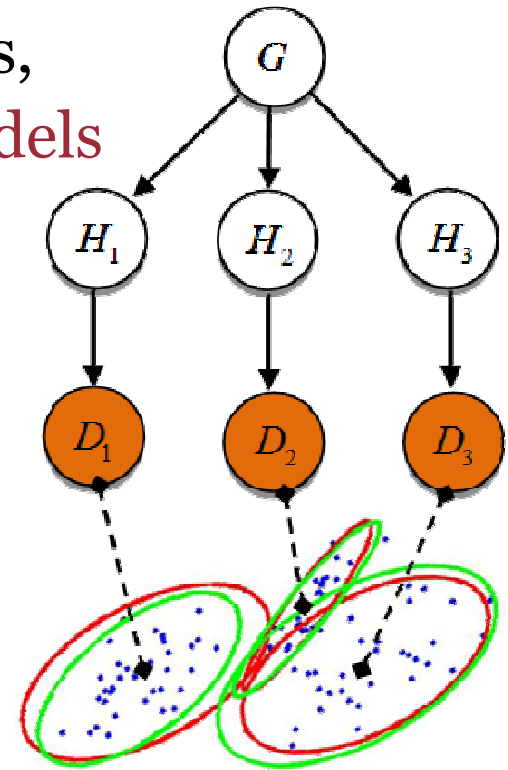
UIUC, 2013

Contents

- Mixture Model
 - Finite mixture
 - Infinite mixture
- Matrix Feature Model
 - Finite features
 - Infinite features(Indian Buffet Process)

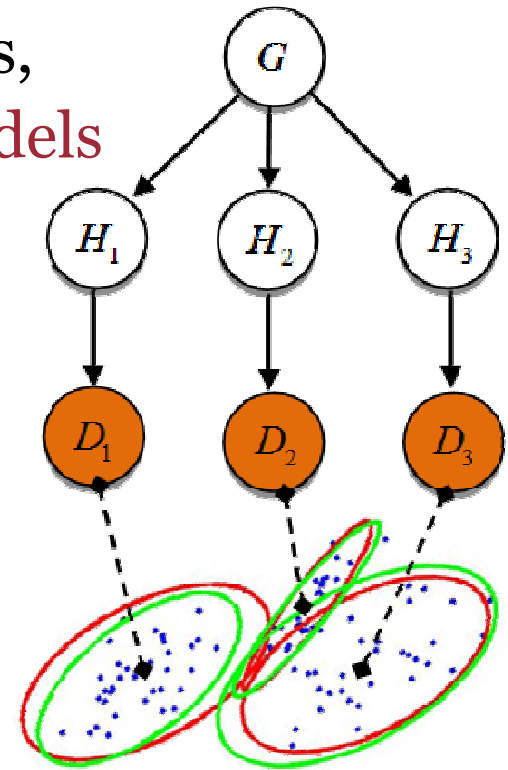
Bayesian Nonparametrics

- Models with undefined number of elements,
 - Dirichlet Process for infinite mixture models
 - With various applications
 - Hierarchies
 - Topics and syntactic classes
 - Objects appearing in one image



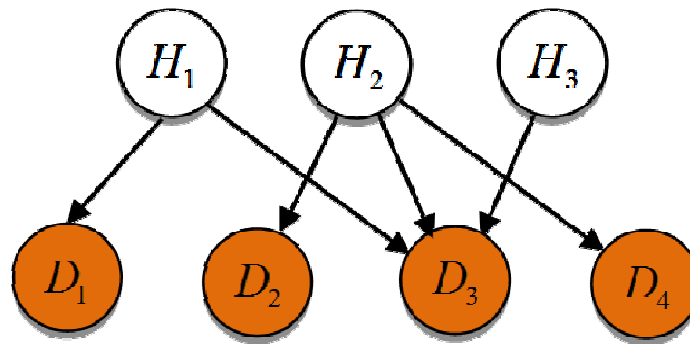
Bayesian Nonparametrics

- Models with undefined number of elements,
 - Dirichlet Process for infinite mixture models
 - With various applications
 - Hierarchies
 - Topics and syntactic classes
 - Objects appearing in one image
- Cons
 - The models are limited to the case that could be modeled using DP.
 - i.e. set of observations are generated by only one latent component



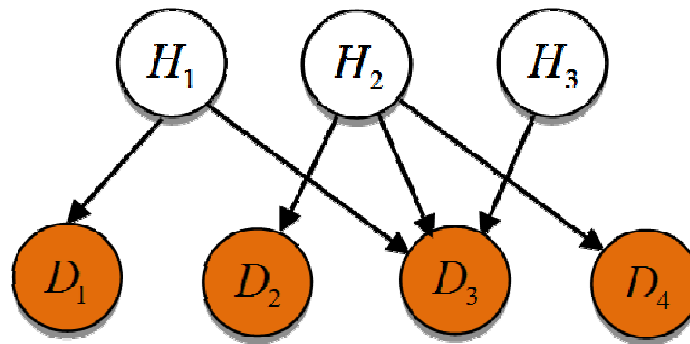
Bayesian Nonparametrics contd.

- In practice there might be more complicated interaction between latent variables and observations



Bayesian Nonparametrics contd.

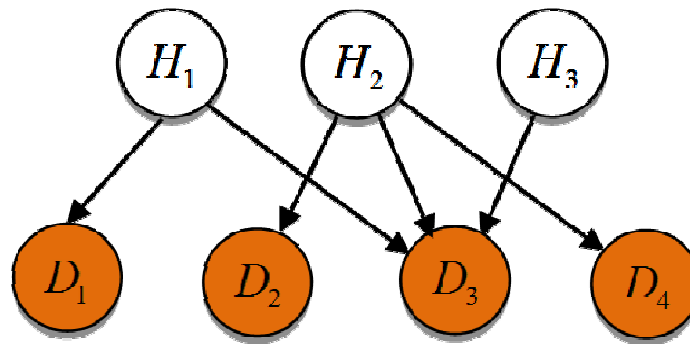
- In practice there might be more complicated interaction between latent variables and observations



- Solution
 - Looking for more flexible nonparametric models

Bayesian Nonparametrics contd.

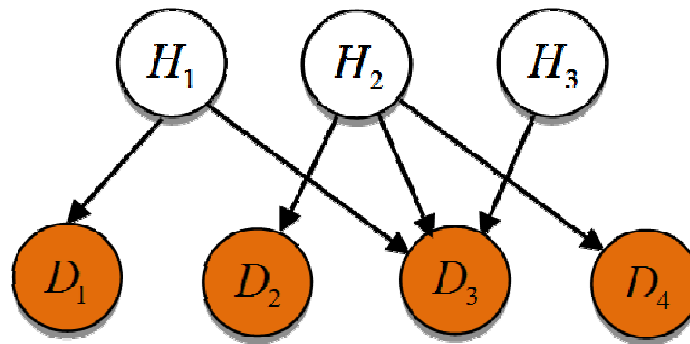
- In practice there might be more complicated interaction between latent variables and observations



- Solution
 - Looking for more flexible nonparametric models
 - Such interaction could be captured via a **binary matrix**
 - Infinite features means infinite number of columns

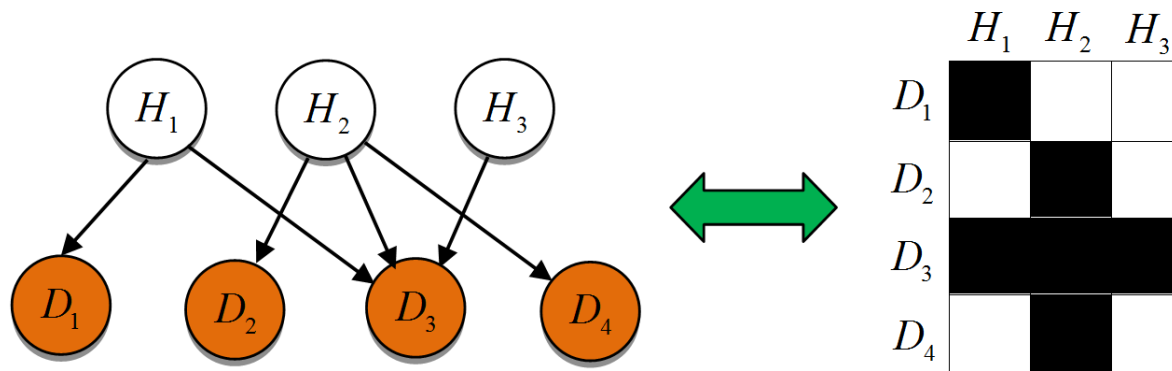
Bayesian Nonparametrics contd.

- In practice there might be more complicated interaction between latent variables and observations



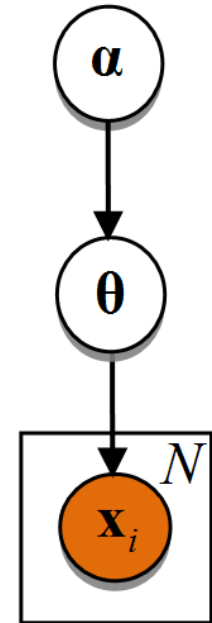
- Solution

- Looking for more flexible nonparametric models
- Such interaction could be captured via a **binary matrix**
- Infinite features means infinite number of columns



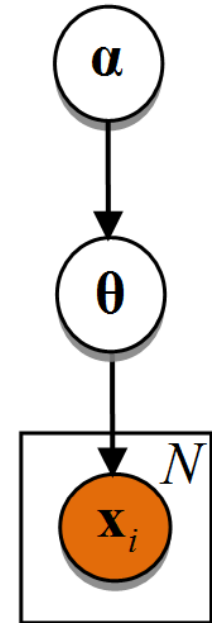
Finite Mixture Model

- Set of observation: $\{\mathbf{x}_i\}_{i=1}^N$



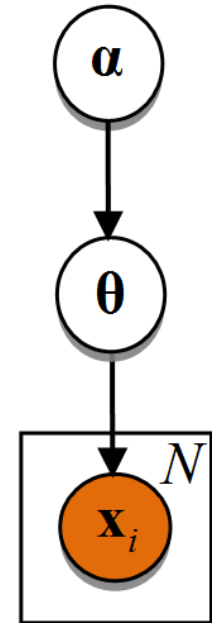
Finite Mixture Model

- Set of observation: $\{\mathbf{x}_i\}_{i=1}^N$
- Constant clusters, K



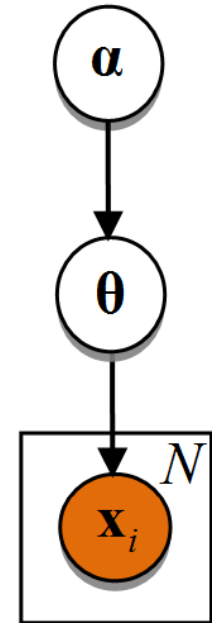
Finite Mixture Model

- Set of observation: $\{\mathbf{x}_i\}_{i=1}^N$
- Constant clusters, K
- Cluster assignment for \mathbf{x}_i is $c_i \in \{1, \dots, K\}$



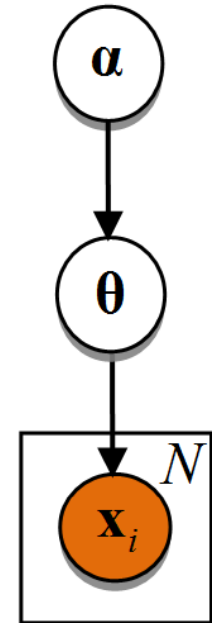
Finite Mixture Model

- Set of observation: $\{\mathbf{x}_i\}_{i=1}^N$
- Constant clusters, K
- Cluster assignment for \mathbf{x}_i is $c_i \in \{1, \dots, K\}$
- Cluster assignments vector :



Finite Mixture Model

- Set of observation: $\{\mathbf{x}_i\}_{i=1}^N$
- Constant clusters, K
- Cluster assignment for \mathbf{x}_i is $c_i \in \{1, \dots, K\}$
- Cluster assignments vector : $\mathbf{c} = [c_1, c_2, \dots, c_N]^T$



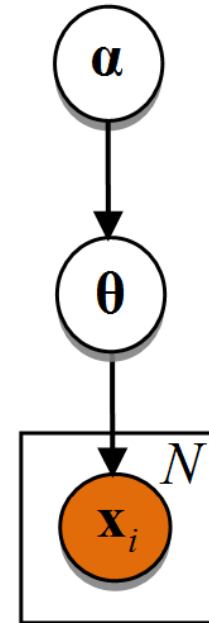
Finite Mixture Model

- Set of observation: $\{\mathbf{x}_i\}_{i=1}^N$
- Constant clusters, K
- Cluster assignment for \mathbf{x}_i is $c_i \in \{1, \dots, K\}$
- Cluster assignments vector : $\mathbf{c} = [c_1, c_2, \dots, c_N]^T$
- The probability of each sample under the model:

$$p(\mathbf{x}_i | \theta) = \sum_{k=1}^K p(\mathbf{x}_i | c_i = k) p(c_i = k)$$

- The likelihood of samples:

$$p(\mathbf{X} | \theta) = \prod_{i=1}^N \sum_{k=1}^K p(\mathbf{x}_i | c_i = k) p(c_i = k)$$



Finite Mixture Model

- Set of observation: $\{\mathbf{x}_i\}_{i=1}^N$
- Constant clusters, K
- Cluster assignment for \mathbf{x}_i is $c_i \in \{1, \dots, K\}$
- Cluster assignments vector : $\mathbf{c} = [c_1, c_2, \dots, c_N]^T$
- The probability of each sample under the model:

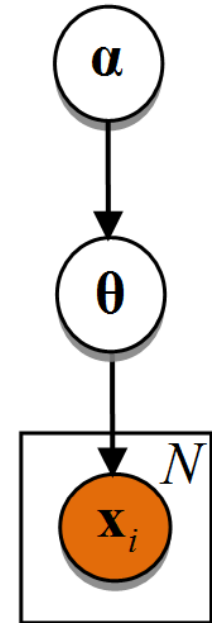
$$p(\mathbf{x}_i | \theta) = \sum_{k=1}^K p(\mathbf{x}_i | c_i = k) p(c_i = k)$$

- The likelihood of samples:

$$p(\mathbf{X} | \theta) = \prod_{i=1}^N \sum_{k=1}^K p(\mathbf{x}_i | c_i = k) p(c_i = k)$$

- The prior on the component probabilities (symmetric Dirichlet dits.)

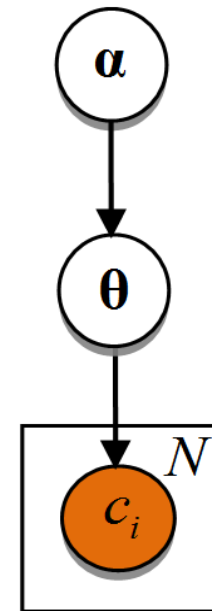
$$\theta | \alpha \sim \text{Dirichlet}\left(\frac{\alpha}{K}, \dots, \frac{\alpha}{K}\right).$$



Finite Mixture Model

- Since we want the mixture model to be valid for **any general component** $p(x_j | c_j = i)$ we only assume the number of **cluster assignments** to be the goal of learning this mixture model !
- Cluster assignments: $\mathbf{c} = [c_1, c_2, \dots, c_N]^T$
- The model can be summarized as :

$$\begin{cases} \boldsymbol{\theta} | \boldsymbol{\alpha} \sim \text{Dirichlet}(\frac{\alpha}{K}, \dots, \frac{\alpha}{K}). \\ c_i | \boldsymbol{\theta} \sim \text{Discrete}(\boldsymbol{\theta}) \end{cases}$$



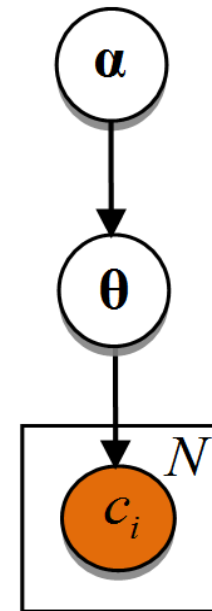
Finite Mixture Model

- Since we want the mixture model to be valid for **any general component** $p(x_j | c_j = i)$ we only assume the number of **cluster assignments** to be the goal of learning this mixture model !
- Cluster assignments: $\mathbf{c} = [c_1, c_2, \dots, c_N]^T$
- The model can be summarized as :

$$\begin{cases} \boldsymbol{\theta} | \boldsymbol{\alpha} \sim \text{Dirichlet}(\frac{\alpha}{K}, \dots, \frac{\alpha}{K}). \\ c_i | \boldsymbol{\theta} \sim \text{Discrete}(\boldsymbol{\theta}) \end{cases}$$

- To have a valid model, all of the distributions must be valid !

$$p(\boldsymbol{\theta} | \mathbf{c}) = \frac{p(\mathbf{c} | \boldsymbol{\theta}).p(\boldsymbol{\theta})}{p(\mathbf{c})}$$



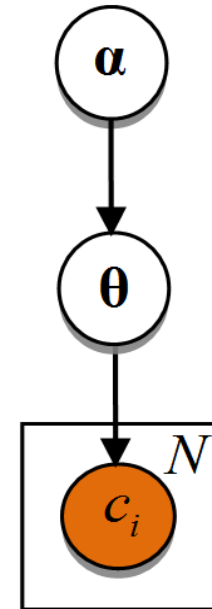
Finite Mixture Model

- Since we want the mixture model to be valid for **any general component** $p(x_j | c_j = i)$ we only assume the number of **cluster assignments** to be the goal of learning this mixture model !
- Cluster assignments: $\mathbf{c} = [c_1, c_2, \dots, c_N]^T$
- The model can be summarized as :

$$\begin{cases} \boldsymbol{\theta} | \boldsymbol{\alpha} \sim \text{Dirichlet}(\frac{\alpha}{K}, \dots, \frac{\alpha}{K}). \\ c_i | \boldsymbol{\theta} \sim \text{Discrete}(\boldsymbol{\theta}) \end{cases}$$

- To have a valid model, all of the distributions must be valid !

$$p(\boldsymbol{\theta} | \mathbf{c}) = \frac{p(\mathbf{c} | \boldsymbol{\theta}) \cdot \boxed{p(\boldsymbol{\theta})} \Rightarrow \text{prior}}{p(\mathbf{c})}$$



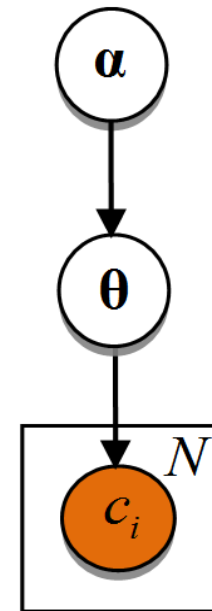
Finite Mixture Model

- Since we want the mixture model to be valid for **any general component** $p(x_j | c_j = i)$ we only assume the number of **cluster assignments** to be the goal of learning this mixture model !
- Cluster assignments: $\mathbf{c} = [c_1, c_2, \dots, c_N]^T$
- The model can be summarized as :

$$\begin{cases} \boldsymbol{\theta} | \boldsymbol{\alpha} \sim \text{Dirichlet}(\frac{\alpha}{K}, \dots, \frac{\alpha}{K}). \\ c_i | \boldsymbol{\theta} \sim \text{Discrete}(\boldsymbol{\theta}) \end{cases}$$

- To have a valid model, all of the distributions must be valid !

$$p(\boldsymbol{\theta} | \mathbf{c}) = \frac{\overset{\text{likelihood}}{\boxed{p(\mathbf{c} | \boldsymbol{\theta})}} \cdot \boxed{p(\boldsymbol{\theta})} \overset{\text{prior}}{\longrightarrow}}{p(\mathbf{c})}$$



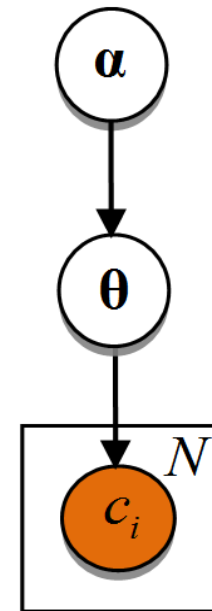
Finite Mixture Model

- Since we want the mixture model to be valid for **any general component** $p(x_j | c_j = i)$ we only assume the number of **cluster assignments** to be the goal of learning this mixture model !
- Cluster assignments: $\mathbf{c} = [c_1, c_2, \dots, c_N]^T$
- The model can be summarized as :

$$\begin{cases} \boldsymbol{\theta} | \boldsymbol{\alpha} \sim \text{Dirichlet}(\frac{\alpha}{K}, \dots, \frac{\alpha}{K}). \\ c_i | \boldsymbol{\theta} \sim \text{Discrete}(\boldsymbol{\theta}) \end{cases}$$

- To have a valid model, all of the distributions must be valid !

$$\text{posterior} \leftarrow p(\boldsymbol{\theta} | \mathbf{c}) = \frac{\overset{\text{likelihood}}{p(\mathbf{c} | \boldsymbol{\theta})} \cdot p(\boldsymbol{\theta}) \rightarrow \text{prior}}{p(\mathbf{c})}$$



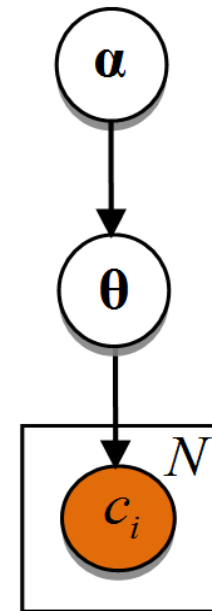
Finite Mixture Model

- Since we want the mixture model to be valid for **any general component** $p(x_j | c_j = i)$ we only assume the number of **cluster assignments** to be the goal of learning this mixture model !
- Cluster assignments: $\mathbf{c} = [c_1, c_2, \dots, c_N]^T$
- The model can be summarized as :

$$\begin{cases} \boldsymbol{\theta} | \boldsymbol{\alpha} \sim \text{Dirichlet}(\frac{\alpha}{K}, \dots, \frac{\alpha}{K}). \\ c_i | \boldsymbol{\theta} \sim \text{Discrete}(\boldsymbol{\theta}) \end{cases}$$

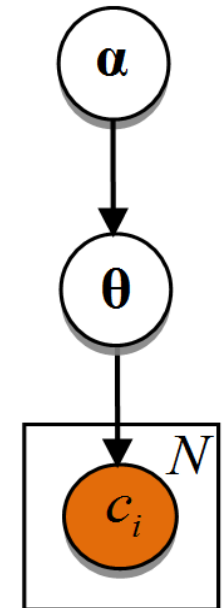
- To have a valid model, all of the distributions must be valid !

$$\text{posterior} \leftarrow p(\boldsymbol{\theta} | \mathbf{c}) = \frac{\overset{\text{likelihood}}{p(\mathbf{c} | \boldsymbol{\theta})} \cdot \overset{\text{prior}}{p(\boldsymbol{\theta})}}{\underset{\text{Marginal likelihood (Evidence)}}{p(\mathbf{c})}}$$



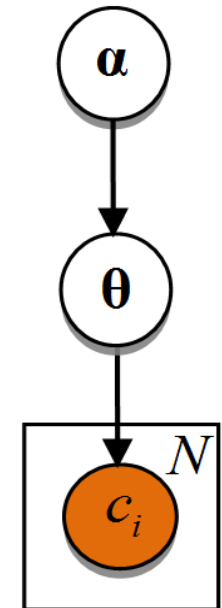
Finite Mixture Model contd.

$$p(\mathbf{c}) = \int_{\Delta_K} p(\mathbf{c} | \boldsymbol{\theta}) p(\boldsymbol{\theta}) d\boldsymbol{\theta} \quad p(\boldsymbol{\theta}) = \left(D\left(\frac{\alpha}{K}, \dots, \frac{\alpha}{K}\right) \right)^{-1} \prod_{k=1}^K \theta_k^{\alpha_k - 1}$$



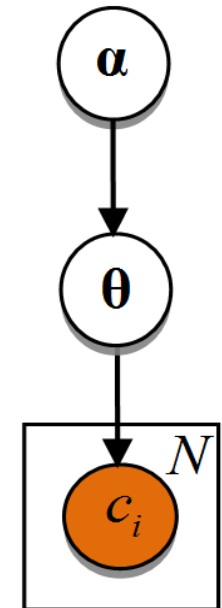
Finite Mixture Model contd.

$$p(\mathbf{c}) = \int_{\Delta_K} p(\mathbf{c} | \boldsymbol{\theta}) p(\boldsymbol{\theta}) d\boldsymbol{\theta} \quad p(\boldsymbol{\theta}) = \left(D\left(\frac{\alpha}{K}, \dots, \frac{\alpha}{K}\right) \right)^{-1} \prod_{k=1}^K \theta_k^{\alpha_k - 1}$$
$$m_k \propto \sum_{i=1}^N \delta(c_i = k)$$



Finite Mixture Model contd.

$$p(\mathbf{c}) = \int_{\Delta_K} p(\mathbf{c} | \boldsymbol{\theta}) p(\boldsymbol{\theta}) d\boldsymbol{\theta} \quad p(\boldsymbol{\theta}) = \left(D\left(\frac{\alpha}{K}, \dots, \frac{\alpha}{K}\right) \right)^{-1} \prod_{k=1}^K \theta_k^{\alpha_k - 1}$$
$$m_k \triangleq \sum_{i=1}^N \delta(c_i = k) \quad \Rightarrow \quad p(\mathbf{c} | \boldsymbol{\theta}) = \prod_{k=1}^K \theta_k^{m_k}$$

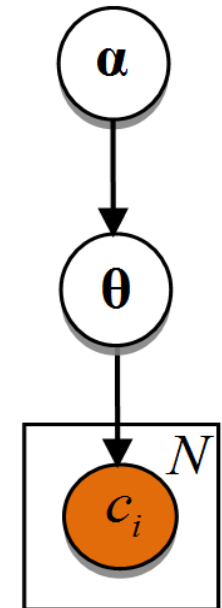


Finite Mixture Model contd.

$$p(\mathbf{c}) = \int_{\Delta_K} p(\mathbf{c} | \boldsymbol{\theta}) p(\boldsymbol{\theta}) d\boldsymbol{\theta} \quad p(\boldsymbol{\theta}) = \left(D\left(\frac{\alpha}{K}, \dots, \frac{\alpha}{K}\right) \right)^{-1} \prod_{k=1}^K \theta_k^{\alpha_k - 1}$$

$$m_k \triangleq \sum_{i=1}^N \delta(c_i = k) \Rightarrow p(\mathbf{c} | \boldsymbol{\theta}) = \prod_{k=1}^K \theta_k^{m_k}$$

$$\begin{aligned} p(\mathbf{c}) &= \int_{\Delta_K} p(\mathbf{c} | \boldsymbol{\theta}) p(\boldsymbol{\theta}) d\boldsymbol{\theta} \\ &= \int_{\Delta_K} \frac{1}{D\left(\frac{\alpha}{K}, \dots, \frac{\alpha}{K}\right)} \prod_{k=1}^K \theta_k^{m_k + \frac{\alpha}{K} - 1} d\boldsymbol{\theta} \end{aligned}$$



Finite Mixture Model contd.

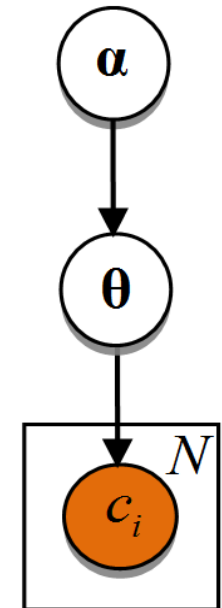
$$p(\mathbf{c}) = \int_{\Delta_K} p(\mathbf{c} | \boldsymbol{\theta}) p(\boldsymbol{\theta}) d\boldsymbol{\theta} \quad p(\boldsymbol{\theta}) = \left(D\left(\frac{\alpha}{K}, \dots, \frac{\alpha}{K}\right) \right)^{-1} \prod_{k=1}^K \theta_k^{\alpha_k - 1}$$

$$m_k \triangleq \sum_{i=1}^N \delta(c_i = k) \quad \Rightarrow \quad p(\mathbf{c} | \boldsymbol{\theta}) = \prod_{k=1}^K \theta_k^{m_k}$$

$$p(\mathbf{c}) = \int_{\Delta_K} p(\mathbf{c} | \boldsymbol{\theta}) p(\boldsymbol{\theta}) d\boldsymbol{\theta}$$

$$= \int_{\Delta_K} \frac{1}{D\left(\frac{\alpha}{K}, \dots, \frac{\alpha}{K}\right)} \prod_{k=1}^K \theta_k^{m_k + \frac{\alpha}{K} - 1} d\boldsymbol{\theta}$$

$$= \frac{\prod_{k=1}^K \Gamma\left(m_k + \frac{\alpha}{K}\right)}{\left(\Gamma\left(\frac{\alpha}{K}\right)\right)^K} \frac{\Gamma(\alpha)}{\Gamma(N + \alpha)}, \quad \text{s.t.} \quad m_k = \sum_{i=1}^N \delta(c_i = k)$$





Infinite mixture model

Infinite mixture model

- Infinite clusters likelihood
 - It is like saying that we have : $K \rightarrow \infty$

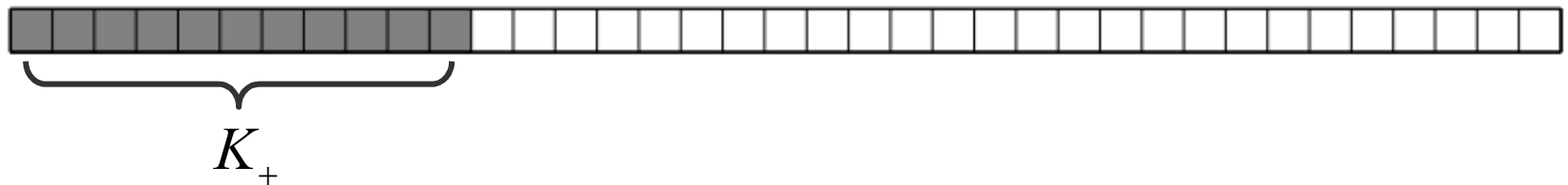
Infinite mixture model

- Infinite clusters likelihood
 - It is like saying that we have : $K \rightarrow \infty$
- Since we always have limited samples in reality, we will have limited number of clusters used; so we define two set of clusters:



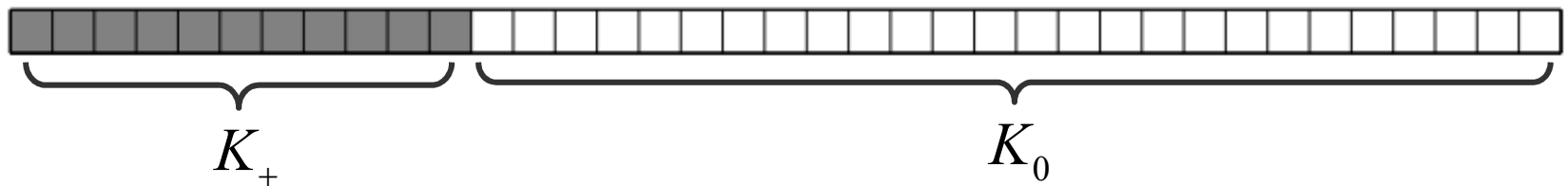
Infinite mixture model

- Infinite clusters likelihood
 - It is like saying that we have : $K \rightarrow \infty$
- Since we always have limited samples in reality, we will have limited number of clusters used; so we define two set of clusters:
 K_+ number of classes for which $m_k > 0$



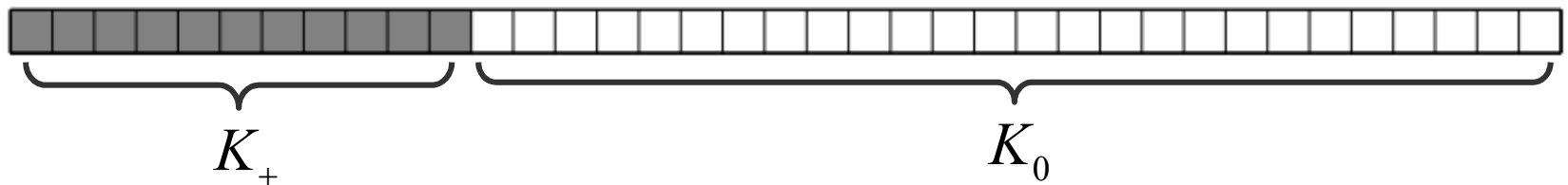
Infinite mixture model

- Infinite clusters likelihood
 - It is like saying that we have : $K \rightarrow \infty$
- Since we always have limited samples in reality, we will have limited number of clusters used; so we define two set of clusters:
 - K_+ number of classes for which $m_k > 0$
 - K_0 number of classes for which $m_k = 0$



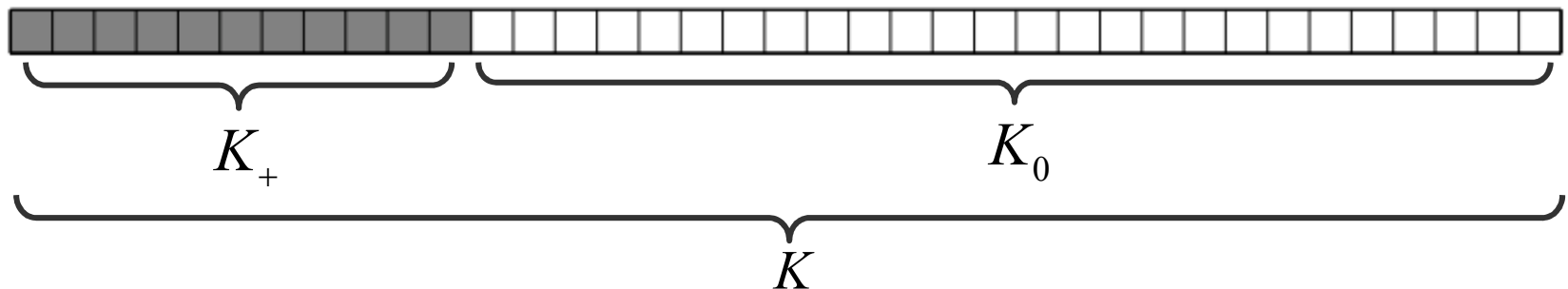
Infinite mixture model

- Infinite clusters likelihood
 - It is like saying that we have : $K \rightarrow \infty$
- Since we always have limited samples in reality, we will have limited number of clusters used; so we define two set of clusters:
 - K_+ number of classes for which $m_k > 0$
 - K_0 number of classes for which $m_k = 0$
- Assume a reordering, such that $\forall k > K_+ \Rightarrow m_k = 0$; and $\forall k \leq K_+ \Rightarrow m_k > 0$



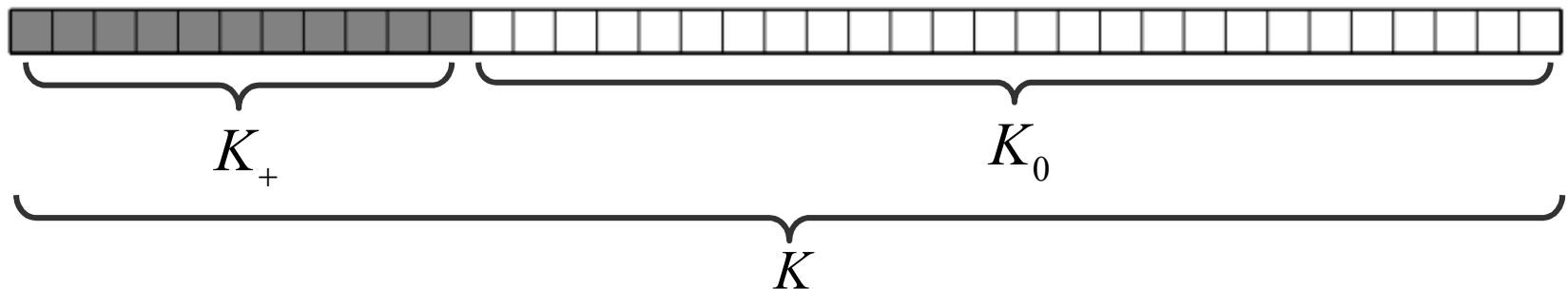
Infinite mixture model

- Infinite clusters likelihood
 - It is like saying that we have : $K \rightarrow \infty$
- Since we always have limited samples in reality, we will have limited number of clusters used; so we define two set of clusters:
 - K_+ number of classes for which $m_k > 0$
 - K_0 number of classes for which $m_k = 0$
- Assume a reordering, such that $\forall k > K_+ \Rightarrow m_k = 0$; and $\forall k \leq K_+ \Rightarrow m_k > 0$



Infinite mixture model

- Infinite clusters likelihood
 - It is like saying that we have : $K \rightarrow \infty$
- Since we always have limited samples in reality, we will have limited number of clusters used; so we define two set of clusters:
 - K_+ number of classes for which $m_k > 0$
 - K_0 number of classes for which $m_k = 0$
- Assume a reordering, such that $\forall k > K_+ \Rightarrow m_k = 0$; and $\forall k \leq K_+ \Rightarrow m_k > 0$



- Infinite clusters likelihood
 - It is like saying that we have : $K \rightarrow \infty$
 - Infinite dimensional multinomial cluster distribution.

Infinite mixture model

- Now we return to the **previous slides** and set $K \rightarrow \infty$ in formulas

$$p(\mathbf{c}) = \frac{\prod_{k=1}^K \Gamma\left(m_k + \frac{\alpha}{K}\right)}{\left(\Gamma\left(\frac{\alpha}{K}\right)\right)^K} \frac{\Gamma(\alpha)}{\Gamma(N + \alpha)}, \quad \text{s.t.} \quad m_k = \sum_{i=1}^N \delta(c_i = k)$$

Infinite mixture model

- Now we return to the **previous slides** and set $K \rightarrow \infty$ in formulas

$$p(\mathbf{c}) = \frac{\prod_{k=1}^K \Gamma\left(m_k + \frac{\alpha}{K}\right)}{\left(\Gamma\left(\frac{\alpha}{K}\right)\right)^K} \frac{\Gamma(\alpha)}{\Gamma(N + \alpha)}, \quad \text{s.t.} \quad m_k = \sum_{i=1}^N \delta(c_i = k)$$

$$\Gamma(x+1) = x\Gamma(x) \Rightarrow \Gamma\left(m_k + \frac{\alpha}{K}\right) = \left(m_k + \frac{\alpha}{K} - 1\right) \dots \left(\frac{\alpha}{K}\right) \Gamma\left(\frac{\alpha}{K}\right)$$

Infinite mixture model

- Now we return to the **previous slides** and set $K \rightarrow \infty$ in formulas

$$p(\mathbf{c}) = \frac{\prod_{k=1}^K \Gamma\left(m_k + \frac{\alpha}{K}\right)}{\left(\Gamma\left(\frac{\alpha}{K}\right)\right)^K} \frac{\Gamma(\alpha)}{\Gamma(N + \alpha)}, \quad \text{s.t.} \quad m_k = \sum_{i=1}^N \delta(c_i = k)$$

$$\Gamma(x+1) = x\Gamma(x) \Rightarrow \Gamma\left(m_k + \frac{\alpha}{K}\right) = \left(m_k + \frac{\alpha}{K} - 1\right) \dots \left(\frac{\alpha}{K}\right) \Gamma\left(\frac{\alpha}{K}\right) \Rightarrow \frac{\Gamma\left(m_k + \frac{\alpha}{K}\right)}{\Gamma\left(\frac{\alpha}{K}\right)} = \left(\frac{\alpha}{K}\right) \prod_{j=1}^{m_k-1} \left(j + \frac{\alpha}{K}\right)$$

Infinite mixture model

- Now we return to the **previous slides** and set $K \rightarrow \infty$ in formulas

$$p(\mathbf{c}) = \frac{\prod_{k=1}^K \Gamma\left(m_k + \frac{\alpha}{K}\right)}{\left(\Gamma\left(\frac{\alpha}{K}\right)\right)^K} \frac{\Gamma(\alpha)}{\Gamma(N + \alpha)}, \quad \text{s.t.} \quad m_k = \sum_{i=1}^N \delta(c_i = k)$$

$$\Gamma(x+1) = x\Gamma(x) \Rightarrow \Gamma\left(m_k + \frac{\alpha}{K}\right) = \left(m_k + \frac{\alpha}{K} - 1\right) \dots \left(\frac{\alpha}{K}\right) \Gamma\left(\frac{\alpha}{K}\right) \Rightarrow \frac{\Gamma\left(m_k + \frac{\alpha}{K}\right)}{\Gamma\left(\frac{\alpha}{K}\right)} = \left(\frac{\alpha}{K}\right) \prod_{j=1}^{m_k-1} \left(j + \frac{\alpha}{K}\right)$$

$$\Rightarrow p(\mathbf{c}) = \frac{\prod_{k=1}^K \Gamma\left(m_k + \frac{\alpha}{K}\right)}{\left(\Gamma\left(\frac{\alpha}{K}\right)\right)^K} \frac{\Gamma(\alpha)}{\Gamma(N + \alpha)} = \left(\frac{\alpha}{K}\right)^{K_+} \prod_{k=1}^{K_+} \prod_{j=1}^{m_k-1} \left(j + \frac{\alpha}{K}\right) \frac{\Gamma(\alpha)}{\Gamma(N + \alpha)}$$

Infinite mixture model

- Now we return to the **previous slides** and set $K \rightarrow \infty$ in formulas

$$p(\mathbf{c}) = \frac{\prod_{k=1}^K \Gamma\left(m_k + \frac{\alpha}{K}\right)}{\left(\Gamma\left(\frac{\alpha}{K}\right)\right)^K} \frac{\Gamma(\alpha)}{\Gamma(N + \alpha)}, \quad \text{s.t.} \quad m_k = \sum_{i=1}^N \delta(c_i = k)$$

$$\Gamma(x+1) = x\Gamma(x) \Rightarrow \Gamma\left(m_k + \frac{\alpha}{K}\right) = \left(m_k + \frac{\alpha}{K} - 1\right) \dots \left(\frac{\alpha}{K}\right) \Gamma\left(\frac{\alpha}{K}\right) \Rightarrow \frac{\Gamma\left(m_k + \frac{\alpha}{K}\right)}{\Gamma\left(\frac{\alpha}{K}\right)} = \left(\frac{\alpha}{K}\right)^{m_k-1} \prod_{j=1}^{m_k-1} \left(j + \frac{\alpha}{K}\right)$$

$$\Rightarrow p(\mathbf{c}) = \frac{\prod_{k=1}^K \Gamma\left(m_k + \frac{\alpha}{K}\right)}{\left(\Gamma\left(\frac{\alpha}{K}\right)\right)^K} \frac{\Gamma(\alpha)}{\Gamma(N + \alpha)} = \left(\frac{\alpha}{K}\right)^{K_+} \prod_{k=1}^{K_+} \prod_{j=1}^{m_k-1} \left(j + \frac{\alpha}{K}\right) \frac{\Gamma(\alpha)}{\Gamma(N + \alpha)}$$

If we set $K \rightarrow \infty$ the marginal likelihood will be $p(\mathbf{c}) \rightarrow 0$.
 Instead we can model this problem, by defining probabilities on **partitions of samples**, instead of **class labels for each sample**.

Infinite mixture model contd.

- Define a partition of objects;
- Want to partition N objects into K_+ classes

Infinite mixture model contd.

- Define a partition of objects;
- Want to partition N objects into K classes
- Equivalence class of object partitions: $[\mathbf{c}] = \{\mathbf{c}_i \mid \mathbf{c}_i \in \mathbf{c}\}$

Infinite mixture model contd.

- Define a partition of objects;
- Want to partition N objects into K_+ classes
- Equivalence class of object partitions: $[\mathbf{c}] = \{\mathbf{c}_i \mid \mathbf{c}_i \in \mathbf{c}\}$

$$p([\mathbf{c}]) = \sum_{\mathbf{c} \in [\mathbf{c}]} p(\mathbf{c}) = \frac{K!}{K_0!} \left(\frac{\alpha}{K} \right)^{K_+} \prod_{k=1}^{K_+} \prod_{j=1}^{m_k-1} \left(j + \frac{\alpha}{K} \right) \frac{\Gamma(\alpha)}{\Gamma(N + \alpha)}$$

Infinite mixture model contd.

- Define a partition of objects;
- Want to partition N objects into K_+ classes
- Equivalence class of object partitions: $[\mathbf{c}] = \{\mathbf{c}_i \mid \mathbf{c}_i \in \mathbf{c}\}$

$$p([\mathbf{c}]) = \sum_{\mathbf{c} \in [\mathbf{c}]} p(\mathbf{c}) = \frac{K!}{K_0!} \left(\frac{\alpha}{K} \right)^{K_+} \prod_{k=1}^{K_+} \prod_{j=1}^{m_k-1} \left(j + \frac{\alpha}{K} \right) \frac{\Gamma(\alpha)}{\Gamma(N + \alpha)}$$

$$\Rightarrow p([\mathbf{c}]) = \alpha^{K_+} \cdot \frac{K!}{K_0! K^{K_+}} \cdot \prod_{k=1}^{K_+} \prod_{j=1}^{m_k-1} \left(j + \frac{\alpha}{K} \right) \cdot \frac{\Gamma(\alpha)}{\Gamma(N + \alpha)}$$

Infinite mixture model contd.

- Define a partition of objects;
- Want to partition N objects into K_+ classes
- Equivalence class of object partitions: $[\mathbf{c}] = \{\mathbf{c}_i \mid \mathbf{c}_i \in \mathbf{c}\}$

$$p([\mathbf{c}]) = \sum_{\mathbf{c} \in [\mathbf{c}]} p(\mathbf{c}) = \frac{K!}{K_0!} \left(\frac{\alpha}{K}\right)^{K_+} \prod_{k=1}^{K_+} \prod_{j=1}^{m_k-1} \left(j + \frac{\alpha}{K}\right) \frac{\Gamma(\alpha)}{\Gamma(N + \alpha)}$$

$$\Rightarrow p([\mathbf{c}]) = \alpha^{K_+} \cdot \frac{K!}{K_0! K^{K_+}} \cdot \prod_{k=1}^{K_+} \prod_{j=1}^{m_k-1} \left(j + \frac{\alpha}{K}\right) \cdot \frac{\Gamma(\alpha)}{\Gamma(N + \alpha)}$$

Infinite mixture model contd.

- Define a partition of objects;
- Want to partition N objects into K_+ classes
- Equivalence class of object partitions: $[\mathbf{c}] = \{\mathbf{c}_i \mid \mathbf{c}_i \in \mathbf{c}\}$

$$p([\mathbf{c}]) = \sum_{\mathbf{c} \in [\mathbf{c}]} p(\mathbf{c}) = \frac{K!}{K_0!} \left(\frac{\alpha}{K}\right)^{K_+} \prod_{k=1}^{K_+} \prod_{j=1}^{m_k-1} \left(j + \frac{\alpha}{K}\right) \frac{\Gamma(\alpha)}{\Gamma(N + \alpha)}$$

$$\Rightarrow p([\mathbf{c}]) = \alpha^{K_+} \cdot \frac{K!}{K_0! K^{K_+}} \cdot \prod_{k=1}^{K_+} \prod_{j=1}^{m_k-1} \left(j + \frac{\alpha}{K}\right) \cdot \frac{\Gamma(\alpha)}{\Gamma(N + \alpha)}$$

$$\Rightarrow \lim_{K \rightarrow \infty} p([\mathbf{c}]) = \alpha^{K_+} \cdot 1 \cdot \prod_{k=1}^{K_+} (m_k - 1)! \cdot \frac{\Gamma(\alpha)}{\Gamma(N + \alpha)}$$

Infinite mixture model contd.

- Define a partition of objects;
- Want to partition N objects into K_+ classes
- Equivalence class of object partitions: $[\mathbf{c}] = \{\mathbf{c}_i \mid \mathbf{c}_i \in \mathbf{c}\}$

$$p([\mathbf{c}]) = \sum_{\mathbf{c} \in [\mathbf{c}]} p(\mathbf{c}) = \frac{K!}{K_0!} \left(\frac{\alpha}{K}\right)^{K_+} \prod_{k=1}^{K_+} \prod_{j=1}^{m_k-1} \left(j + \frac{\alpha}{K}\right) \frac{\Gamma(\alpha)}{\Gamma(N + \alpha)}$$

$$\Rightarrow p([\mathbf{c}]) = \alpha^{K_+} \cdot \frac{K!}{K_0! K^{K_+}} \cdot \prod_{k=1}^{K_+} \prod_{j=1}^{m_k-1} \left(j + \frac{\alpha}{K}\right) \cdot \frac{\Gamma(\alpha)}{\Gamma(N + \alpha)}$$

$$\Rightarrow \lim_{K \rightarrow \infty} p([\mathbf{c}]) = \alpha^{K_+} \cdot 1 \cdot \prod_{k=1}^{K_+} (m_k - 1)! \cdot \frac{\Gamma(\alpha)}{\Gamma(N + \alpha)}$$

- **Valid probability** distribution for an infinite mixture model

Infinite mixture model contd.

- Define a partition of objects;
- Want to partition N objects into K_+ classes
- Equivalence class of object partitions: $[\mathbf{c}] = \{\mathbf{c}_i \mid \mathbf{c}_i \in \mathbf{c}\}$

$$p([\mathbf{c}]) = \sum_{\mathbf{c} \in [\mathbf{c}]} p(\mathbf{c}) = \frac{K!}{K_0!} \left(\frac{\alpha}{K}\right)^{K_+} \prod_{k=1}^{K_+} \prod_{j=1}^{m_k-1} \left(j + \frac{\alpha}{K}\right) \frac{\Gamma(\alpha)}{\Gamma(N + \alpha)}$$

$$\Rightarrow p([\mathbf{c}]) = \alpha^{K_+} \cdot \frac{K!}{K_0! K^{K_+}} \cdot \prod_{k=1}^{K_+} \prod_{j=1}^{m_k-1} \left(j + \frac{\alpha}{K}\right) \cdot \frac{\Gamma(\alpha)}{\Gamma(N + \alpha)}$$

$$\Rightarrow \lim_{K \rightarrow \infty} p([\mathbf{c}]) = \alpha^{K_+} \cdot 1 \cdot \prod_{k=1}^{K_+} (m_k - 1)! \cdot \frac{\Gamma(\alpha)}{\Gamma(N + \alpha)}$$

- **Valid probability** distribution for an infinite mixture model
- Exchangeable with respect to clusters assignments !

Infinite mixture model contd.

- Define a partition of objects;
- Want to partition N objects into K_+ classes
- Equivalence class of object partitions: $[\mathbf{c}] = \{\mathbf{c}_i \mid \mathbf{c}_i \in \mathbf{c}\}$

$$p([\mathbf{c}]) = \sum_{\mathbf{c} \in [\mathbf{c}]} p(\mathbf{c}) = \frac{K!}{K_0!} \left(\frac{\alpha}{K} \right)^{K_+} \prod_{k=1}^{K_+} \prod_{j=1}^{m_k-1} \left(j + \frac{\alpha}{K} \right) \frac{\Gamma(\alpha)}{\Gamma(N + \alpha)}$$

$$\Rightarrow p([\mathbf{c}]) = \alpha^{K_+} \cdot \frac{K!}{K_0! K^{K_+}} \cdot \prod_{k=1}^{K_+} \prod_{j=1}^{m_k-1} \left(j + \frac{\alpha}{K} \right) \cdot \frac{\Gamma(\alpha)}{\Gamma(N + \alpha)}$$

$$\Rightarrow \lim_{K \rightarrow \infty} p([\mathbf{c}]) = \alpha^{K_+} \cdot 1 \cdot \prod_{k=1}^{K_+} (m_k - 1)! \cdot \frac{\Gamma(\alpha)}{\Gamma(N + \alpha)}$$

- **Valid probability** distribution for an infinite mixture model
- Exchangeable with respect to clusters assignments !
 - Important for Gibbs sampling (and Chinese restaurant process)

Infinite mixture model contd.

- Define a partition of objects;
- Want to partition N objects into K_+ classes
- Equivalence class of object partitions: $[\mathbf{c}] = \{\mathbf{c}_i \mid \mathbf{c}_i \in \mathbf{c}\}$

$$p([\mathbf{c}]) = \sum_{\mathbf{c} \in [\mathbf{c}]} p(\mathbf{c}) = \frac{K!}{K_0!} \left(\frac{\alpha}{K}\right)^{K_+} \prod_{k=1}^{K_+} \prod_{j=1}^{m_k-1} \left(j + \frac{\alpha}{K}\right) \frac{\Gamma(\alpha)}{\Gamma(N + \alpha)}$$

$$\Rightarrow p([\mathbf{c}]) = \alpha^{K_+} \cdot \frac{K!}{K_0! K^{K_+}} \cdot \prod_{k=1}^{K_+} \prod_{j=1}^{m_k-1} \left(j + \frac{\alpha}{K}\right) \cdot \frac{\Gamma(\alpha)}{\Gamma(N + \alpha)}$$

$$\Rightarrow \lim_{K \rightarrow \infty} p([\mathbf{c}]) = \alpha^{K_+} \cdot 1 \cdot \prod_{k=1}^{K_+} (m_k - 1)! \cdot \frac{\Gamma(\alpha)}{\Gamma(N + \alpha)}$$

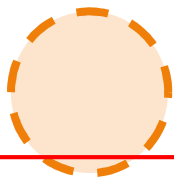
- **Valid probability** distribution for an infinite mixture model
- Exchangeable with respect to clusters assignments !
 - Important for Gibbs sampling (and Chinese restaurant process)
 - De Finetti's theorem: explains why **exchangeable** observations are **conditionally independent** given some **probability distribution**

Chinese Restaurant Process (CRP)

$$p(c_i = k \mid c_1, \dots, c_{i-1}) = \begin{cases} \frac{m_k}{i-1+\alpha} & k \leq K_+ \\ \frac{\alpha}{i-1+\alpha} & k = K+1 \end{cases}$$

Parameter $\alpha = 1$

$$m_1 = 0$$



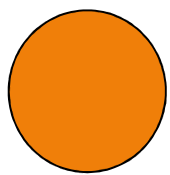
$$p(c_1 = 1) = \frac{1}{1}$$

Chinese Restaurant Process (CRP)

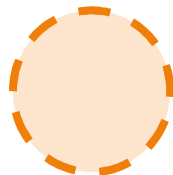
$$p(c_i = k \mid c_1, \dots, c_{i-1}) = \begin{cases} \frac{m_k}{i-1+\alpha} & k \leq K_+ \\ \frac{\alpha}{i-1+\alpha} & k = K+1 \end{cases}$$

Parameter $\alpha = 1$

$$m_1 = 1,$$



$$m_2 = 0$$



$$p(c_2 = 1 \mid c_1) = \frac{1}{1+1}$$

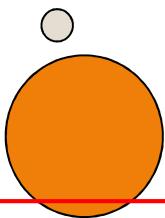
$$p(c_2 = 2 \mid c_1) = \frac{1}{1+1}$$

Chinese Restaurant Process (CRP)

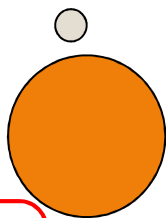
$$p(c_i = k \mid c_1, \dots, c_{i-1}) = \begin{cases} \frac{m_k}{i-1+\alpha} & k \leq K_+ \\ \frac{\alpha}{i-1+\alpha} & k = K+1 \end{cases}$$

Parameter $\alpha = 1$

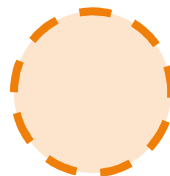
$m_1 = 1,$



$m_2 = 1,$



$m_3 = 0$



$$p(c_3 = 1 \mid c_{1:2}) = \frac{1}{2+1}$$

$$p(c_3 = 3 \mid c_{1:2}) = \frac{1}{2+1}$$

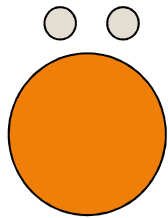
$$p(c_3 = 2 \mid c_{1:2}) = \frac{1}{2+1}$$

Chinese Restaurant Process (CRP)

$$p(c_i = k \mid c_1, \dots, c_{i-1}) = \begin{cases} \frac{m_k}{i-1+\alpha} & k \leq K_+ \\ \frac{\alpha}{i-1+\alpha} & k = K+1 \end{cases}$$

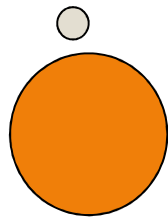
Parameter $\alpha = 1$

$$m_1 = 2,$$



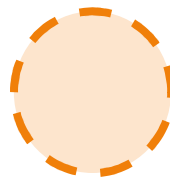
$$p(c_4 = 1 \mid c_{1:3}) = \frac{2}{3+1}$$

$$m_2 = 1,$$



$$p(c_4 = 2 \mid c_{1:3}) = \frac{1}{3+1}$$

$$m_3 = 0$$

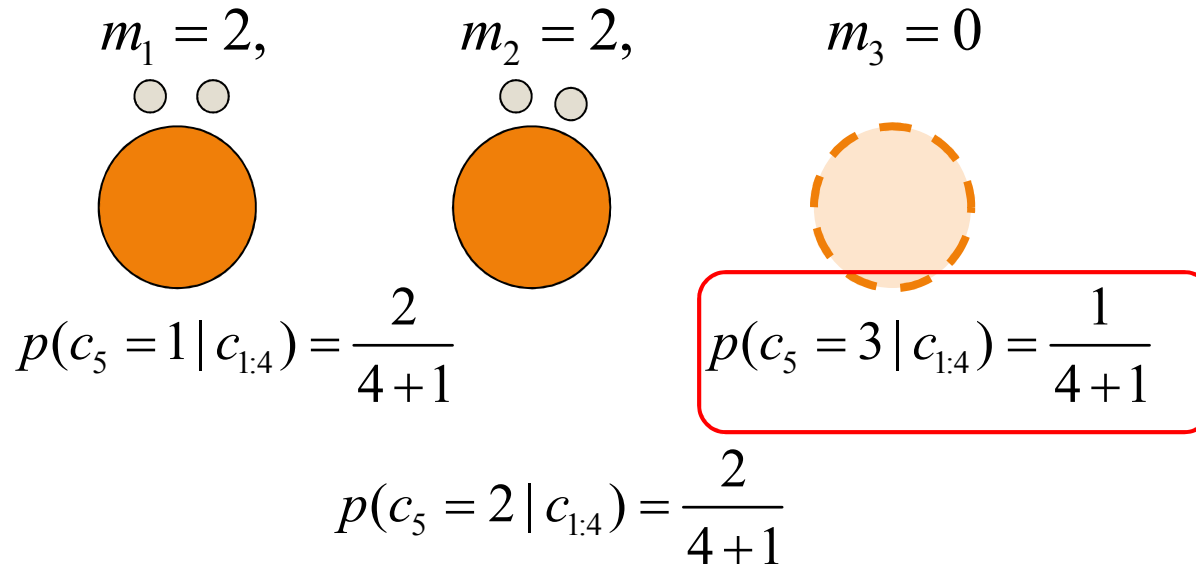


$$p(c_4 = 3 \mid c_{1:3}) = \frac{1}{3+1}$$

Chinese Restaurant Process (CRP)

$$p(c_i = k | c_1, \dots, c_{i-1}) = \begin{cases} \frac{m_k}{i-1+\alpha} & k \leq K_+ \\ \frac{\alpha}{i-1+\alpha} & k = K+1 \end{cases}$$

Parameter $\alpha = 1$

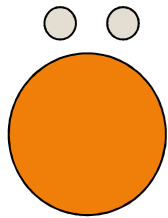


Chinese Restaurant Process (CRP)

$$p(c_i = k | c_1, \dots, c_{i-1}) = \begin{cases} \frac{m_k}{i-1+\alpha} & k \leq K_+ \\ \frac{\alpha}{i-1+\alpha} & k = K+1 \end{cases}$$

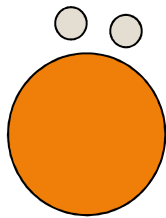
Parameter $\alpha = 1$

$$m_1 = 2,$$



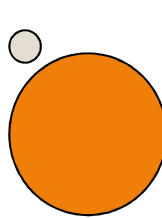
$$p(c_6 = 1 | c_{1:5}) = \frac{2}{5+1}$$

$$m_2 = 2,$$



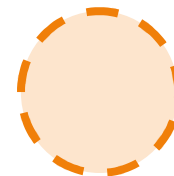
$$p(c_6 = 2 | c_{1:5}) = \frac{2}{5+1}$$

$$m_3 = 1,$$



$$p(c_6 = 3 | c_{1:5}) = \frac{1}{5+1}$$

$$m_4 = 0$$

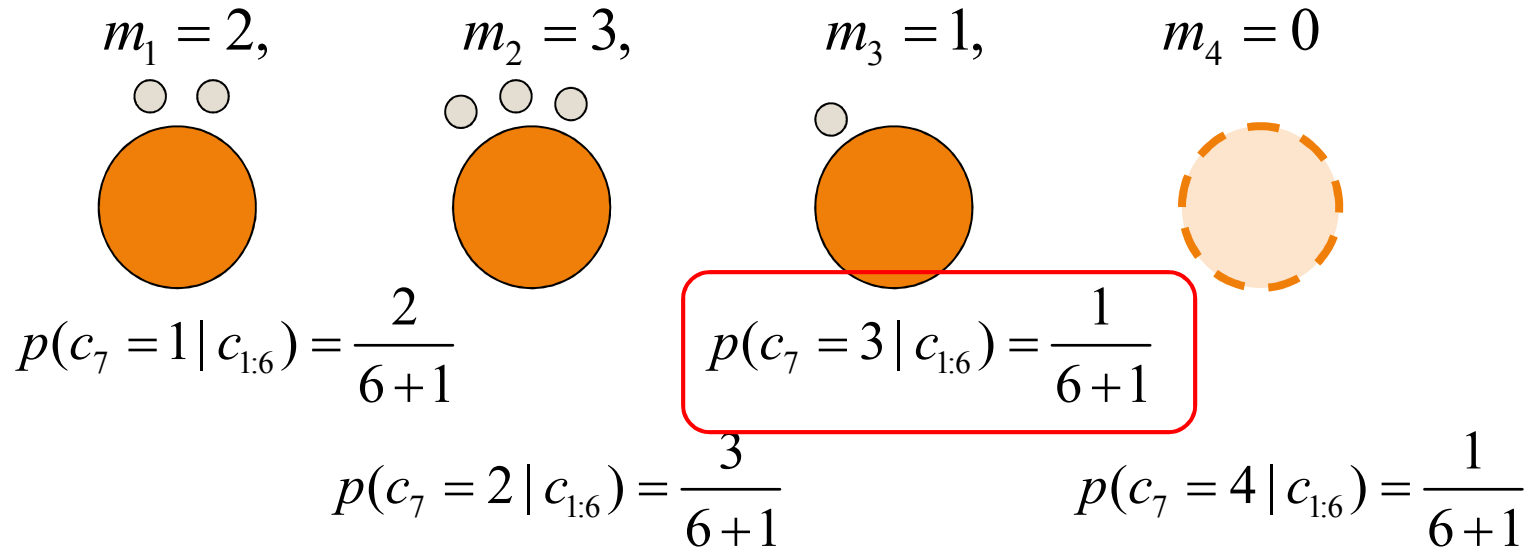


$$p(c_6 = 4 | c_{1:5}) = \frac{1}{5+1}$$

Chinese Restaurant Process (CRP)

$$p(c_i = k | c_1, \dots, c_{i-1}) = \begin{cases} \frac{m_k}{i-1+\alpha} & k \leq K_+ \\ \frac{\alpha}{i-1+\alpha} & k = K+1 \end{cases}$$

Parameter $\alpha = 1$

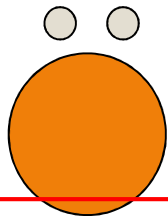


Chinese Restaurant Process (CRP)

$$p(c_i = k | c_1, \dots, c_{i-1}) = \begin{cases} \frac{m_k}{i-1+\alpha} & k \leq K_+ \\ \frac{\alpha}{i-1+\alpha} & k = K+1 \end{cases}$$

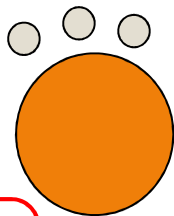
Parameter $\alpha = 1$

$$m_1 = 2,$$



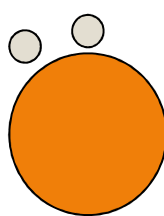
$$p(c_8 = 1 | c_{1:3}) = \frac{2}{7+1}$$

$$m_2 = 3,$$



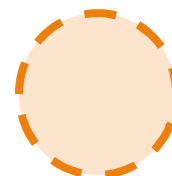
$$p(c_8 = 2 | c_{1:3}) = \frac{3}{7+1}$$

$$m_3 = 2,$$



$$p(c_8 = 3 | c_{1:3}) = \frac{2}{7+1}$$

$$m_4 = 0$$

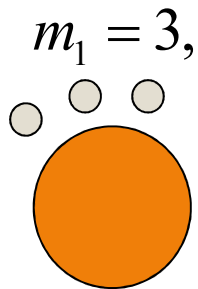


$$p(c_8 = 4 | c_{1:3}) = \frac{1}{7+1}$$

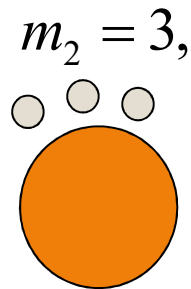
Chinese Restaurant Process (CRP)

$$p(c_i = k | c_1, \dots, c_{i-1}) = \begin{cases} \frac{m_k}{i-1+\alpha} & k \leq K_+ \\ \frac{\alpha}{i-1+\alpha} & k = K+1 \end{cases}$$

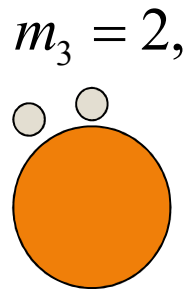
Parameter $\alpha = 1$



$$p(c_9 = 1 | c_{1:3}) = \frac{3}{8+1}$$



$$p(c_9 = 2 | c_{1:3}) = \frac{3}{8+1}$$



$$p(c_9 = 3 | c_{1:3}) = \frac{2}{8+1}$$

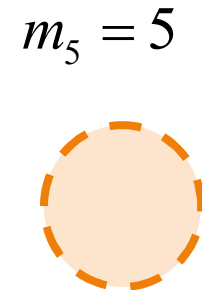
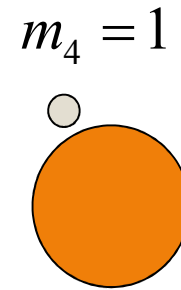
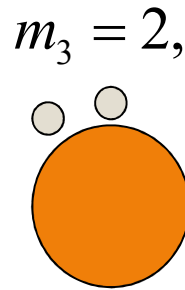
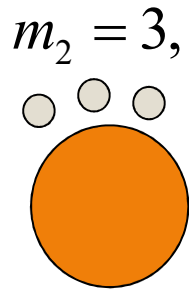
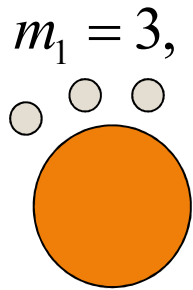


$$p(c_9 = 4 | c_{1:3}) = \frac{1}{8+1}$$

Chinese Restaurant Process (CRP)

$$p(c_i = k | c_1, \dots, c_{i-1}) = \begin{cases} \frac{m_k}{i-1+\alpha} & k \leq K_+ \\ \frac{\alpha}{i-1+\alpha} & k = K+1 \end{cases}$$

Parameter $\alpha = 1$



$$p(c_{10} = 1 | c_{1:4}) = \frac{3}{9+1}$$

$$p(c_{10} = 3 | c_{1:4}) = \frac{2}{9+1}$$

$$p(c_{10} = 5 | c_{1:4}) = \frac{1}{9+1}$$

$$p(c_{10} = 2 | c_{1:4}) = \frac{3}{9+1}$$

$$p(c_{10} = 4 | c_{1:4}) = \frac{1}{9+1}$$

CRP: Gibbs sampling

- Gibbs sampler requires full conditional

$$p(c_i = k \mid \mathbf{c}_{-i}, \mathbf{X}) \propto p(\mathbf{X} \mid \mathbf{c}).p(c_i = k \mid \mathbf{c}_{-i})$$

- Finite Mixture Model:

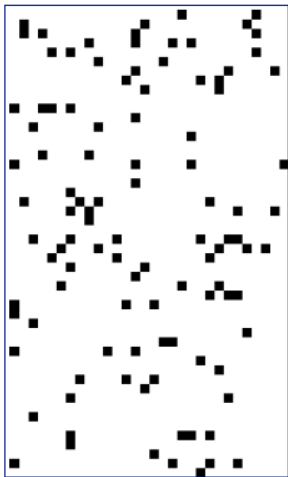
$$p(c_i = k \mid \mathbf{c}_{-i}) = \frac{m_{-i,k} + \frac{\alpha}{K}}{N - 1 + \alpha}$$

- Infinite Mixture Model:

$$p(c_i = k \mid \mathbf{c}_{-i}) = \begin{cases} \frac{m_{-i,k}}{N - 1 + \alpha} & m_{-i,k} > 0 \\ \frac{\alpha}{N - 1 + \alpha} & k = K_{-i} + 1 \\ 0 & \text{otherwise} \end{cases}$$

Beyond the limit of single label

- In **Latent Class Models**:
 - Each object (word) has only one latent label (topic)
 - Finite number of latent labels: **LDA**
 - Infinite number of latent labels: **DPM**
- In **Latent Feature (latent structure) Models**:
 - Each object (graph) has multiple latent features (entities)
 - Finite number of latent features: **Finite Feature Model (FFM)**
 - Infinite number of latent features: **Indian Buffet Process (IBP)**



- Rows are data points
- Columns are latent features
- Movie Preference Example:
 - Rows are movies: *Rise of the Planet of the Apes*
 - Columns are latent features:
 - Made in U.S.
 - Is Science fiction
 - Has apes in it ...



Latent Feature Model

K features

N objects

■	■	□	□	■
■	□	■	□	
□	■	■		
■	□			
■				

\mathbf{Z}

K features

N objects

0.9	1.4	0	0	-0.3
-3.2	0	0.9	0	
0	0.2	-2.8		
1.8	0			
-0.1				

continuous \mathbf{F}

K features

N objects

1	3	0	0	4
5	0	3	0	
0	1	4		
2	0			
5				

discrete \mathbf{F}

- \mathbf{F} : latent feature matrix
- \mathbf{Z} : binary matrix
- \mathbf{V} : value matrix

$$\mathbf{F} = \mathbf{Z} \otimes \mathbf{V}$$

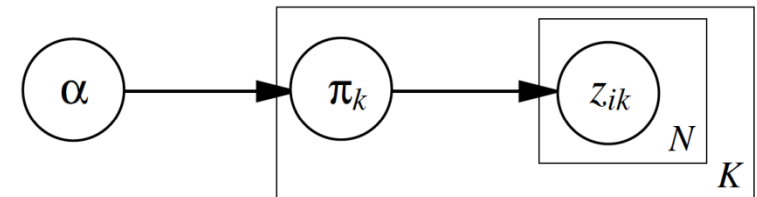
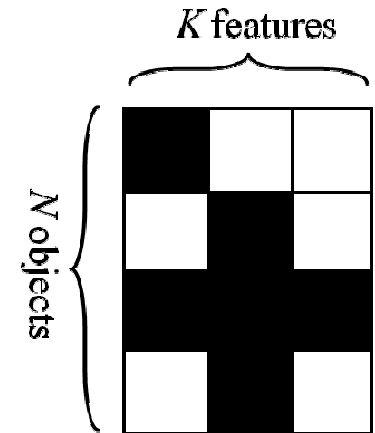
- With $p(\mathbf{F}) = p(\mathbf{Z}) \cdot p(\mathbf{V})$

Finite Feature Model

- Generating \mathbf{Z} : ($N \times K$) binary matrix

- For each column k , draw π_k from beta distribution
- For each object, flip a coin by z_{ik}

$$\begin{cases} \pi_k \mid \alpha \sim \text{Beta}(\frac{\alpha}{K}, 1) & \longleftrightarrow & (\pi_k, 1 - \pi_k) \sim \text{Dir}(\frac{\alpha}{K}, 1) \\ z_{ik} \mid \pi_k \sim \text{Bernoulli}(\pi_k) \end{cases}$$



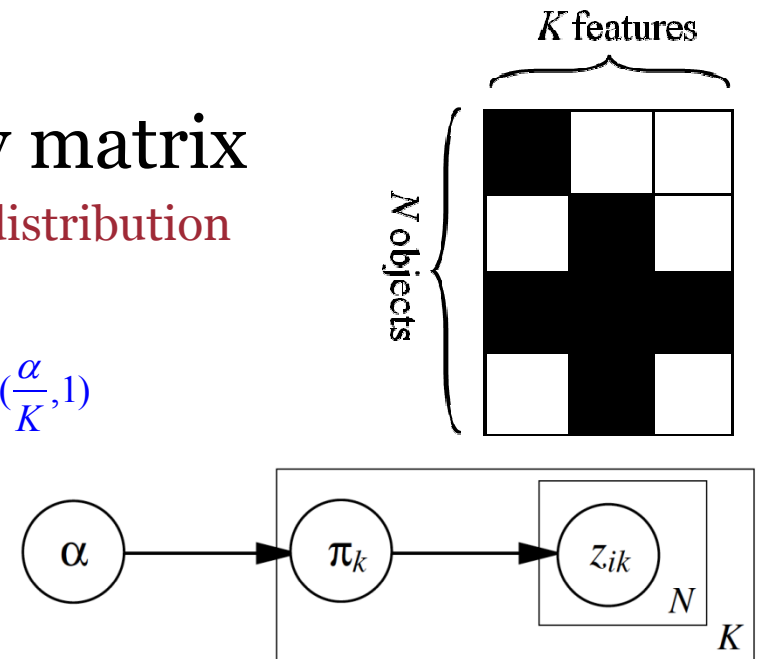
Finite Feature Model

- Generating \mathbf{Z} : ($N \times K$) binary matrix

- For each column k , draw π_k from beta distribution
- For each object, flip a coin by z_{ik}

$$\begin{cases} \pi_k \mid \alpha \sim \text{Beta}(\frac{\alpha}{K}, 1) & \longleftrightarrow & (\pi_k, 1 - \pi_k) \sim \text{Dir}(\frac{\alpha}{K}, 1) \\ z_{ik} \mid \pi_k \sim \text{Bernoulli}(\pi_k) \end{cases}$$

- Distribution of \mathbf{Z} :

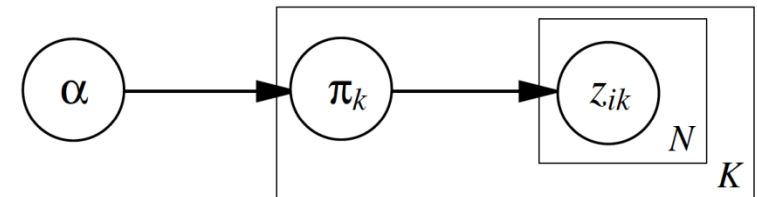
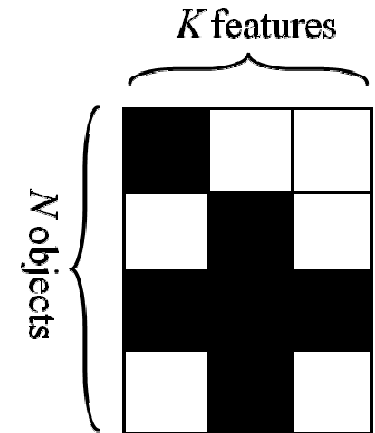


Finite Feature Model

- Generating \mathbf{Z} : ($N \times K$) binary matrix

- For each column k , draw π_k from beta distribution
- For each object, flip a coin by z_{ik}

$$\begin{cases} \pi_k \mid \alpha \sim \text{Beta}(\frac{\alpha}{K}, 1) & \longleftrightarrow & (\pi_k, 1 - \pi_k) \sim \text{Dir}(\frac{\alpha}{K}, 1) \\ z_{ik} \mid \pi_k \sim \text{Bernoulli}(\pi_k) \end{cases}$$



- Distribution of \mathbf{Z} :

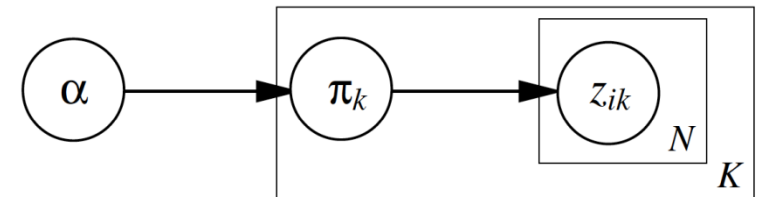
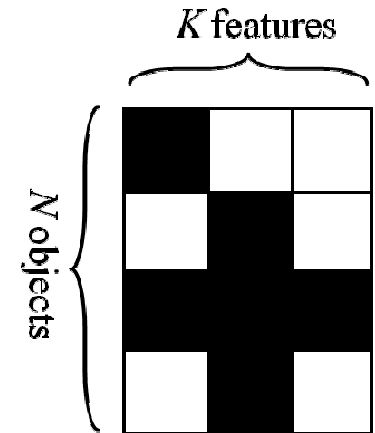
$$p(\mathbf{Z} \mid \boldsymbol{\pi}) = \prod_{k=1}^K \prod_{i=1}^N p(z_{ik} \mid \pi_k) = \prod_{k=1}^K \pi_k^{m_k} (1 - \pi_k)^{N - m_k}$$

Finite Feature Model

- Generating \mathbf{Z} : ($N \times K$) binary matrix

- For each column k , draw π_k from beta distribution
- For each object, flip a coin by z_{ik}

$$\begin{cases} \pi_k | \alpha \sim \text{Beta}(\frac{\alpha}{K}, 1) & \longleftrightarrow & (\pi_k, 1 - \pi_k) \sim \text{Dir}(\frac{\alpha}{K}, 1) \\ z_{ik} | \pi_k \sim \text{Bernoulli}(\pi_k) \end{cases}$$



- Distribution of \mathbf{Z} :

$$p(\mathbf{Z} | \boldsymbol{\pi}) = \prod_{k=1}^K \prod_{i=1}^N p(z_{ik} | \pi_k) = \prod_{k=1}^K \pi_k^{m_k} (1 - \pi_k)^{N - m_k}$$

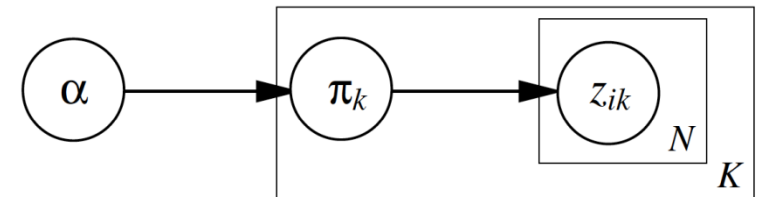
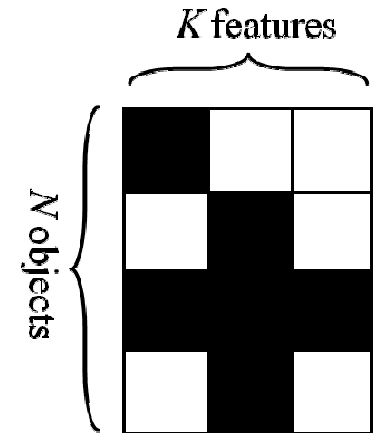
$$p(\mathbf{Z} | \alpha) = \prod_k^K \int_{\pi_k} p(\pi_k | \alpha) p(z_{\cdot k} | \pi_k) d\pi_k$$

Finite Feature Model

- Generating \mathbf{Z} : ($N \times K$) binary matrix

- For each column k , draw π_k from beta distribution
- For each object, flip a coin by z_{ik}

$$\begin{cases} \pi_k | \alpha \sim \text{Beta}(\frac{\alpha}{K}, 1) & \longleftrightarrow & (\pi_k, 1 - \pi_k) \sim \text{Dir}(\frac{\alpha}{K}, 1) \\ z_{ik} | \pi_k \sim \text{Bernoulli}(\pi_k) \end{cases}$$



- Distribution of \mathbf{Z} :

$$p(\mathbf{Z} | \boldsymbol{\pi}) = \prod_{k=1}^K \prod_{i=1}^N p(z_{ik} | \pi_k) = \prod_{k=1}^K \pi_k^{m_k} (1 - \pi_k)^{N - m_k}$$

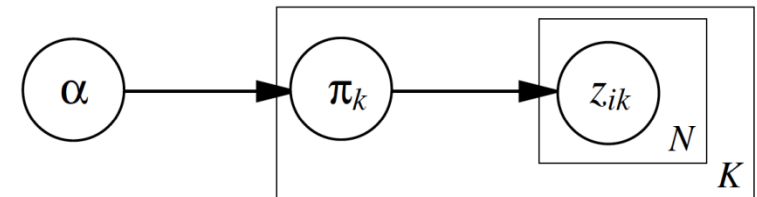
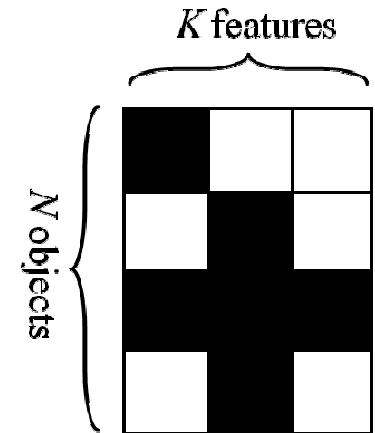
$$\begin{aligned} p(\mathbf{Z} | \alpha) &= \prod_k \int_{\pi_k} p(\pi_k | \alpha) p(z_{\cdot k} | \pi_k) d\pi_k \\ &= \prod_{k=1}^K \frac{\frac{\alpha}{K} \Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \end{aligned}$$

Finite Feature Model

- Generating \mathbf{Z} : ($N \times K$) binary matrix

- For each column k , draw π_k from beta distribution
- For each object, flip a coin by z_{ik}

$$\begin{cases} \pi_k \mid \alpha \sim \text{Beta}(\frac{\alpha}{K}, 1) & \longleftrightarrow & (\pi_k, 1 - \pi_k) \sim \text{Dir}(\frac{\alpha}{K}, 1) \\ z_{ik} \mid \pi_k \sim \text{Bernoulli}(\pi_k) \end{cases}$$



- \mathbf{Z} is sparse:

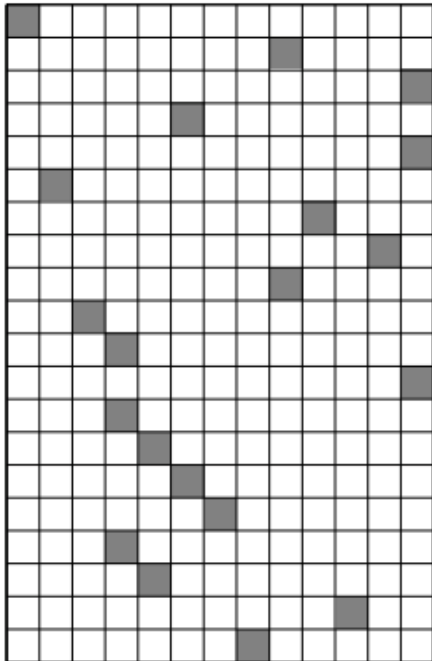
$$\mathbb{E}[\mathbf{1}^T \mathbf{Z} \mathbf{1}] = K \mathbb{E}[\mathbf{1}^T \mathbf{Z}] = K \sum_{i=1}^N \mathbb{E}[z_{ik}] = KN \mathbb{E}[\pi_k] = N \frac{\frac{\alpha}{K}}{1 + \frac{\alpha}{K}} \leq N\alpha$$

- Even $K \rightarrow \infty$

Indian Buffet Process

1st Representation: $K \rightarrow \infty$

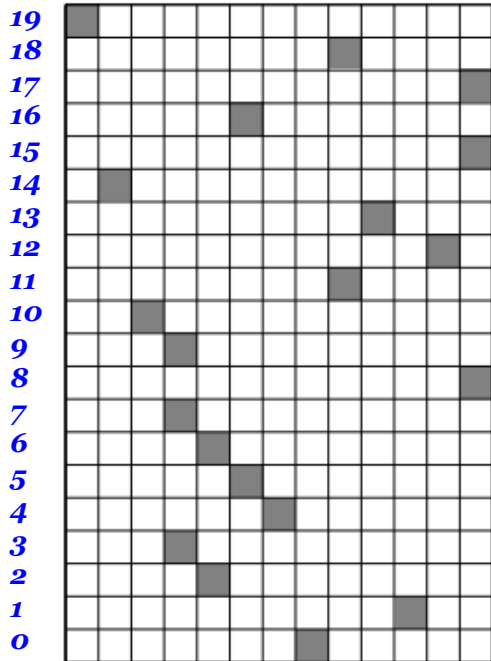
- Difficulty:
 - $P(Z) \rightarrow 0$
 - Solution: define equivalence classes on random binary feature matrices.
- *left-ordered form* function of binary matrices, $lof(\mathbf{Z})$:
 - Compute history \mathbf{h} of feature (column) \mathbf{k}



Indian Buffet Process

1st Representation: $K \rightarrow \infty$

- Difficulty:
 - $P(\mathbf{Z}) \rightarrow 0$
 - Solution: define equivalence classes on random binary feature matrices.
- *left-ordered form* function of binary matrices, $lof(\mathbf{Z})$:
 - Compute history \mathbf{h} of feature (column) k



1st Representation: $K \rightarrow \infty$

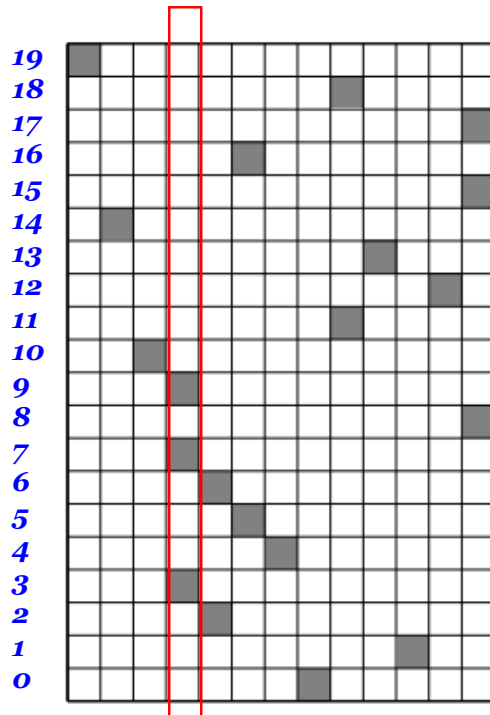
1st Representation: $K \rightarrow \infty$

- $P(Z) \rightarrow 0$

- **Solution:** define equivalence classes on random binary feature matrices.

- *left-ordered form* function of binary matrices, $lof(\mathbf{Z})$:

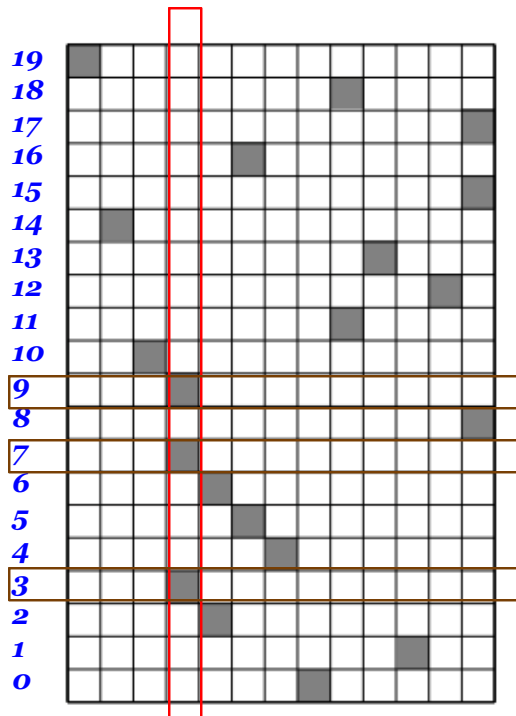
- Compute history **h** of feature (column) **k**



Indian Buffet Process

1st Representation: $K \rightarrow \infty$

- Difficulty:
 - $P(Z) \rightarrow 0$
 - Solution: define equivalence classes on random binary feature matrices.
- *left-ordered form* function of binary matrices, $lof(\mathbf{Z})$:
 - Compute history h of feature (column) k

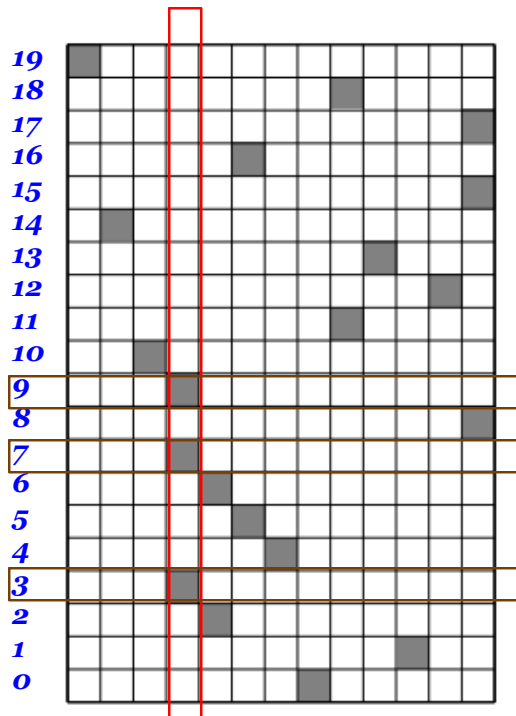


Indian Buffet Process

1st Representation: $K \rightarrow \infty$

- Difficulty:
 - $P(Z) \rightarrow 0$
 - Solution: define equivalence classes on random binary feature matrices.
- *left-ordered form* function of binary matrices, $lof(\mathbf{Z})$:
 - Compute history h of feature (column) k

$$h_4 = 2^3 + 2^7 + 2^9$$

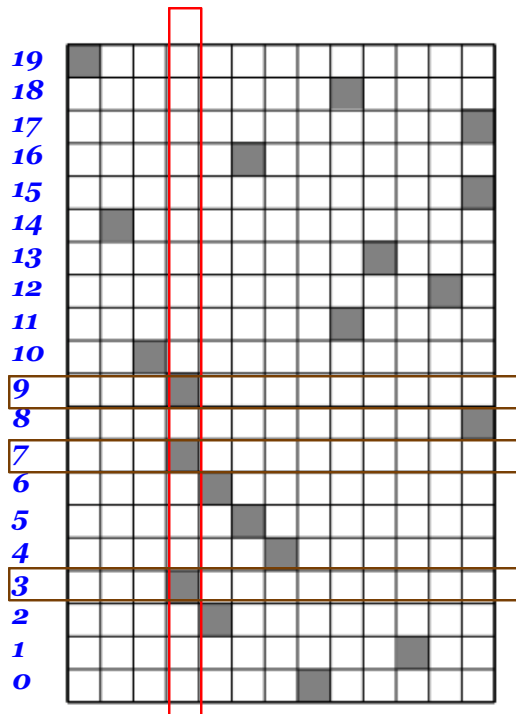


Indian Buffet Process

1st Representation: $K \rightarrow \infty$

- Difficulty:
 - $P(Z) \rightarrow 0$
 - Solution: define equivalence classes on random binary feature matrices.
- *left-ordered form* function of binary matrices, $lof(\mathbf{Z})$:
 - Compute history h of feature (column) k
- Order features by h decreasingly

$$h_4 = 2^3 + 2^7 + 2^9$$

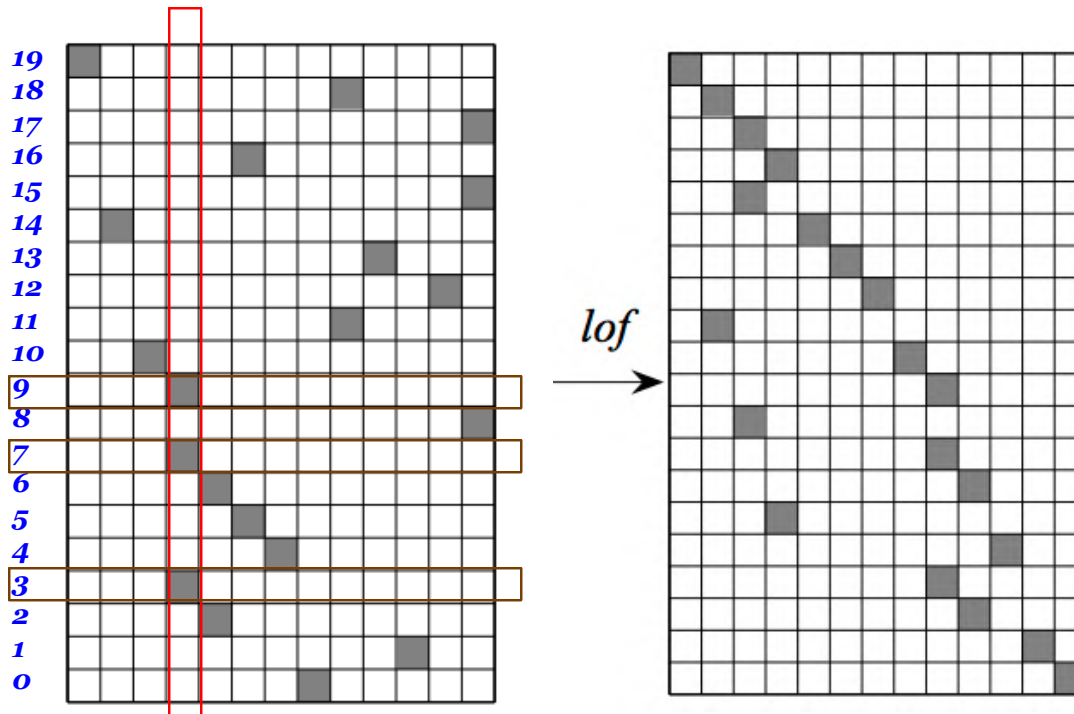


Indian Buffet Process

1st Representation: $K \rightarrow \infty$

- Difficulty:
 - $P(Z) \rightarrow 0$
 - Solution: define equivalence classes on random binary feature matrices.
- *left-ordered form* function of binary matrices, $lof(\mathbf{Z})$:
 - Compute history h of feature (column) k
 - Order features by h decreasingly

$$h_4 = 2^3 + 2^7 + 2^9$$

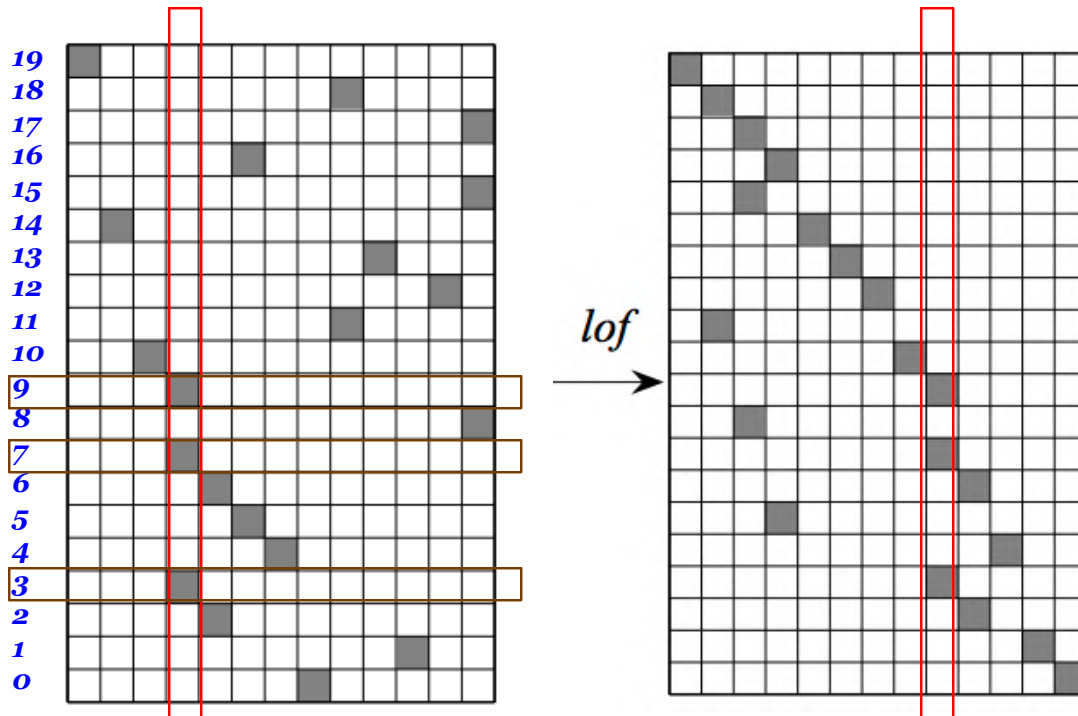


Indian Buffet Process

1st Representation: $K \rightarrow \infty$

- Difficulty:
 - $P(Z) \rightarrow 0$
 - Solution: define equivalence classes on random binary feature matrices.
- *left-ordered form* function of binary matrices, $lof(\mathbf{Z})$:
 - Compute history h of feature (column) k
- Order features by h decreasingly

$$h_4 = 2^3 + 2^7 + 2^9$$



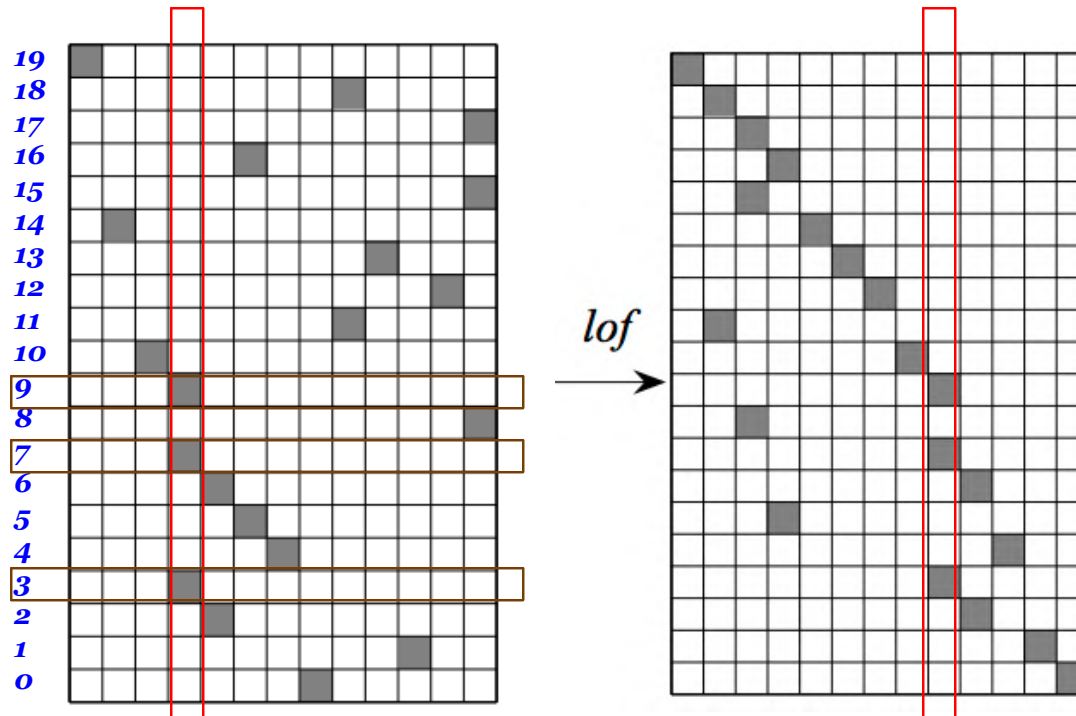
Indian Buffet Process

1st Representation: $K \rightarrow \infty$

- Difficulty:
 - $P(Z) \rightarrow 0$
 - Solution: define equivalence classes on random binary feature matrices.
- *left-ordered form* function of binary matrices, $lof(Z)$:
 - Compute history h of feature (column) k
- Order features by h decreasingly

$$h_4 = 2^3 + 2^7 + 2^9$$

Z_1 Z_2 are ***lof*** equivalent
iff $lof(\mathbf{Z}_1) = lof(\mathbf{Z}_2) = [\mathbf{Z}]$



Indian Buffet Process

1st Representation: $K \rightarrow \infty$

- Difficulty:
 - $P(Z) \rightarrow 0$
 - Solution: define equivalence classes on random binary feature matrices.
- *left-ordered form* function of binary matrices, $lof(Z)$:
 - Compute history h of feature (column) k
 - Order features by h decreasingly

$$h_4 = 2^3 + 2^7 + 2^9$$

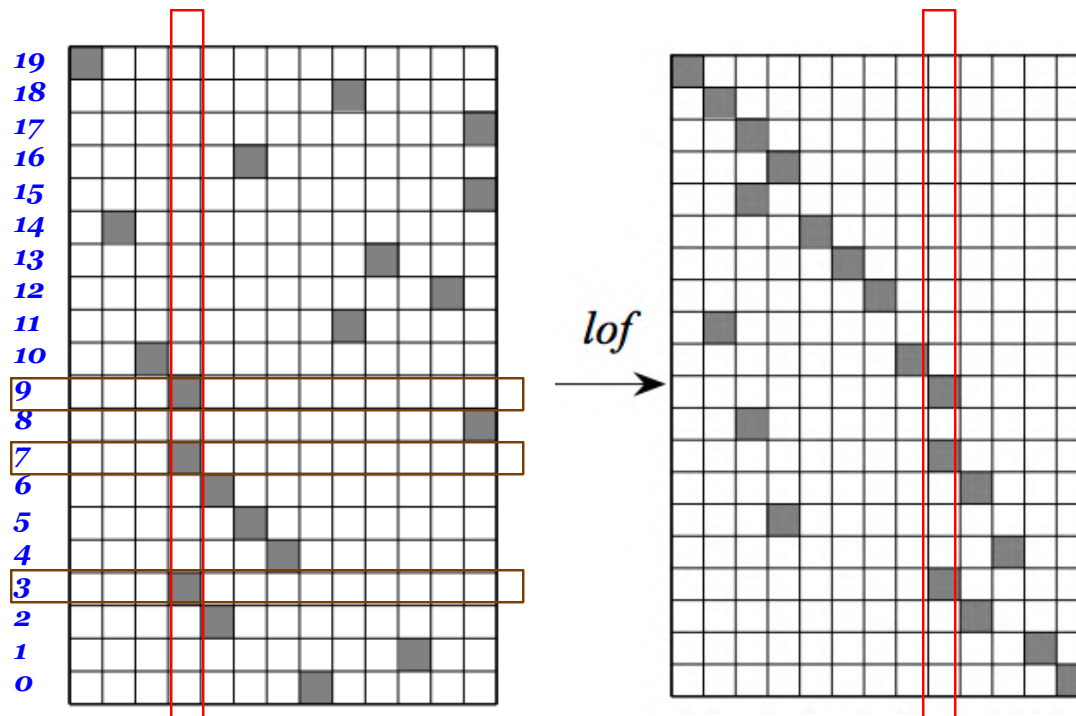
Z_1 Z_2 are *lof* equivalent
iff $lof(Z_1) = lof(Z_2) = [Z]$

Describing $[Z]$:

$$K_h = \#\{i; h_i = h\}$$

$$K_+ = \sum_{i>0} K_h$$

$$K = K_+ + K_0$$



Indian Buffet Process

1st Representation: $K \rightarrow \infty$

- Difficulty:
 - $P(\mathbf{Z}) \rightarrow 0$
 - Solution: define equivalence classes on random binary feature matrices.
- *left-ordered form* function of binary matrices, $\text{lof}(\mathbf{Z})$:
 - Compute history h of feature (column) k
- Order features by h decreasingly

$$h_4 = 2^3 + 2^7 + 2^9$$

\mathbf{Z}_1 \mathbf{Z}_2 are **lof equivalent**
iff $\text{lof}(\mathbf{Z}_1) = \text{lof}(\mathbf{Z}_2) = [\mathbf{Z}]$

Describing $[\mathbf{Z}]$:

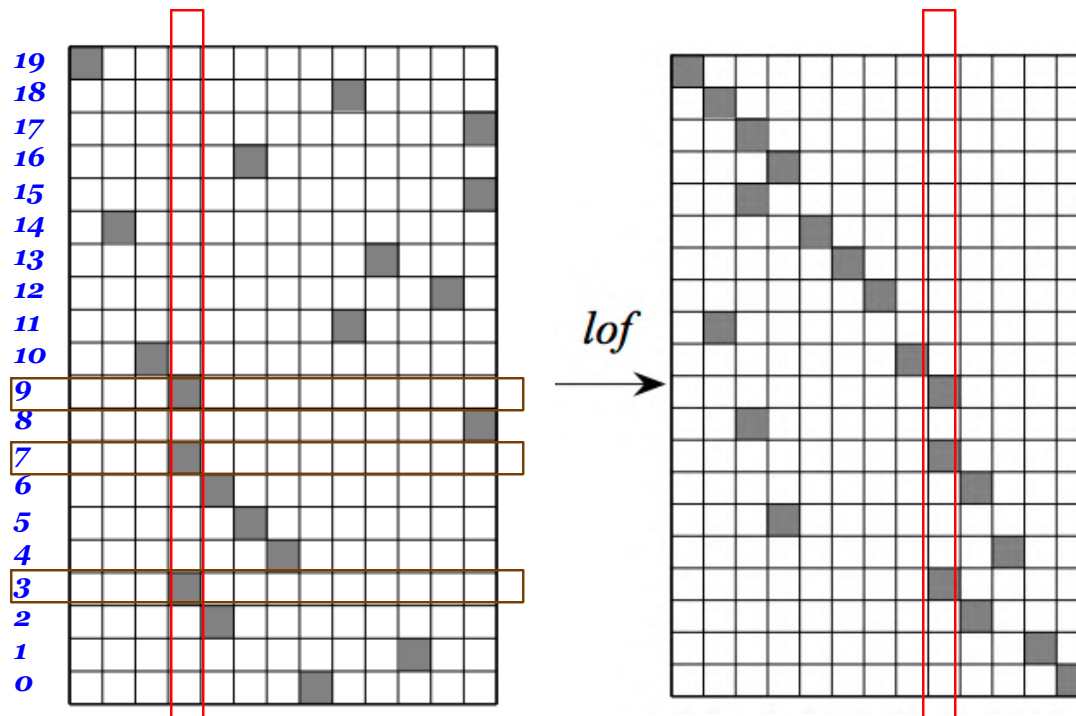
$$K_h = \#\{i; h_i = h\}$$

$$K_+ = \sum_{i>0} K_h$$

$$K = K_+ + K_0$$

Cardinality of $[\mathbf{Z}]$:

$$\binom{K}{K_0 \dots K_{2^N-1}} = \frac{K!}{\prod_{h=0}^{2^N-1} K_h!}$$



Indian Buffet Process

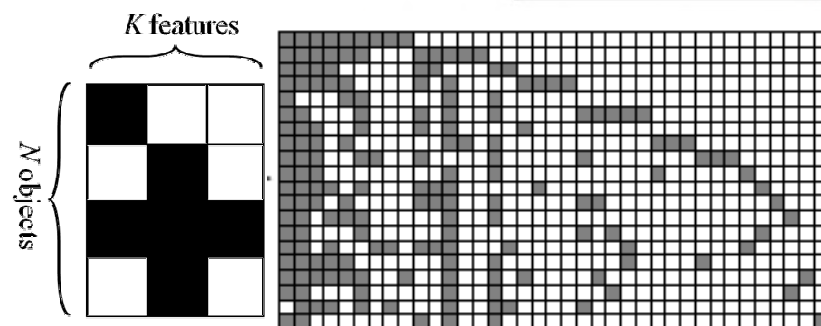
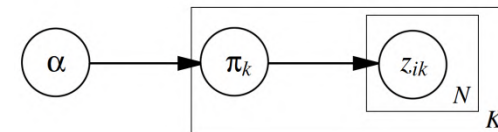
1st Representation: $K \rightarrow \infty$

Given:

$$\left\{ \begin{array}{l} \Pr(\mathbf{Z} | \alpha) = \prod_{k=1}^K \frac{\frac{\alpha}{K} \Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \\ \text{card}([Z]) = \binom{K}{K_0 \dots K_{2^N-1}} = \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \end{array} \right.$$

Derive when $K \rightarrow \infty$: ($K_+ < \infty$ almost surely)

$$\Pr([Z] | \alpha) = \Pr(\mathbf{Z} | \alpha) \cdot \text{card}([Z])$$

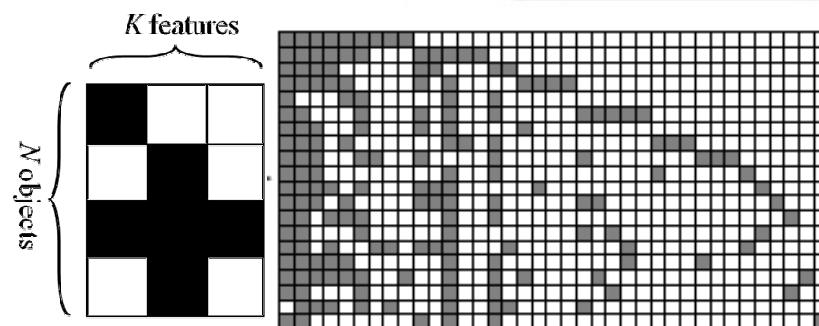


Indian Buffet Process

1st Representation: $K \rightarrow \infty$

Given:

$$\left\{ \begin{array}{l} \Pr(\mathbf{Z} | \alpha) = \prod_{k=1}^K \frac{\frac{\alpha}{K} \Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \\ \text{card}([Z]) = \binom{K}{K_0 \dots K_{2^N-1}} = \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \end{array} \right.$$



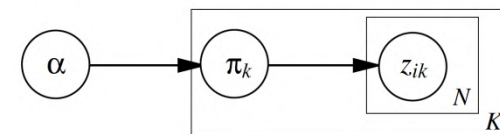
Derive when $K \rightarrow \infty$: ($K_+ < \infty$ almost surely)

$$\Pr([Z] | \alpha) = \Pr(\mathbf{Z} | \alpha) \cdot \text{card}([Z])$$

$$= \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \prod_{k=1}^K \frac{\frac{\alpha}{K} \Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})}$$

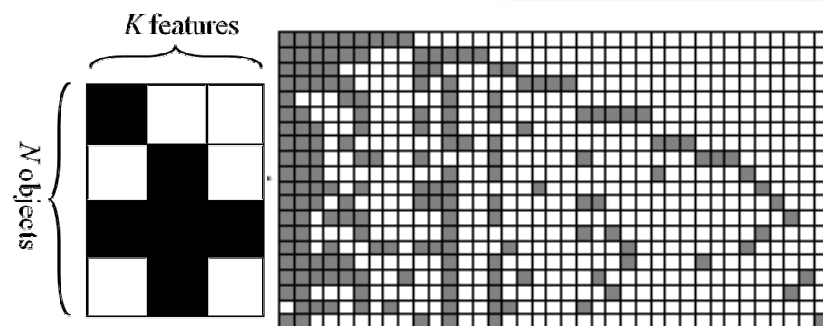
Indian Buffet Process

1st Representation: $K \rightarrow \infty$



Given:

$$\left\{ \begin{array}{l} \Pr(\mathbf{Z} | \alpha) = \prod_{k=1}^K \frac{\frac{\alpha}{K} \Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \\ \text{card}([Z]) = \binom{K}{K_0 \dots K_{2^N-1}} = \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \end{array} \right.$$



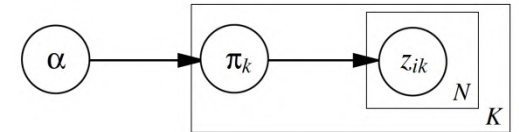
Derive when $K \rightarrow \infty$: ($K_+ < \infty$ almost surely)

$$\Pr([Z] | \alpha) = \Pr(\mathbf{Z} | \alpha) \cdot \text{card}([Z])$$

$$\begin{aligned} &= \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \prod_{k=1}^K \frac{\frac{\alpha}{K} \Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \\ &= \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \left(\frac{\frac{\alpha}{K} \Gamma(\frac{\alpha}{K}) \Gamma(N + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \right)^{K_0} \prod_{k=1}^K \frac{\frac{\alpha}{K} \Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \end{aligned}$$

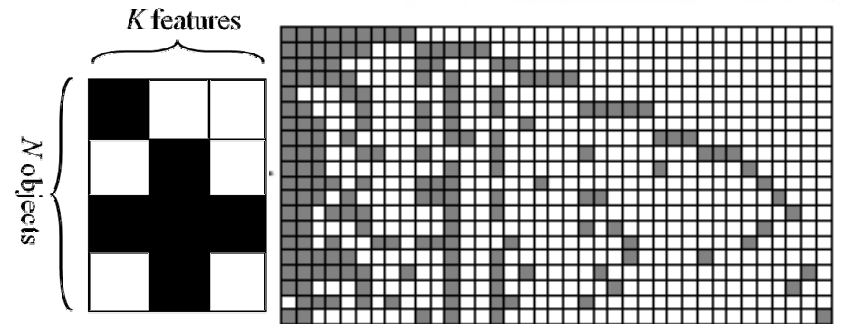
Indian Buffet Process

1st Representation: $K \rightarrow \infty$



Given:

$$\left\{ \begin{array}{l} \Pr(\mathbf{Z} | \alpha) = \prod_{k=1}^K \frac{\frac{\alpha}{K} \Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \\ \text{card}([Z]) = \binom{K}{K_0 \dots K_{2^N-1}} = \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \end{array} \right.$$

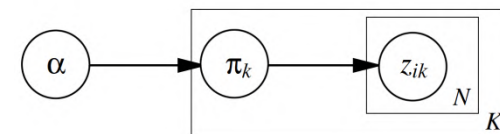


Derive when $K \rightarrow \infty$: ($K_+ < \infty$ almost surely)

$$\begin{aligned} \Pr([Z] | \alpha) &= \Pr(\mathbf{Z} | \alpha) \cdot \text{card}([Z]) \\ &= \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \left(\frac{\frac{\alpha}{K} \Gamma(\frac{\alpha}{K}) \Gamma(N + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \right)^{K_0} \prod_{k=1}^K \frac{\frac{\alpha}{K} \Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \end{aligned}$$

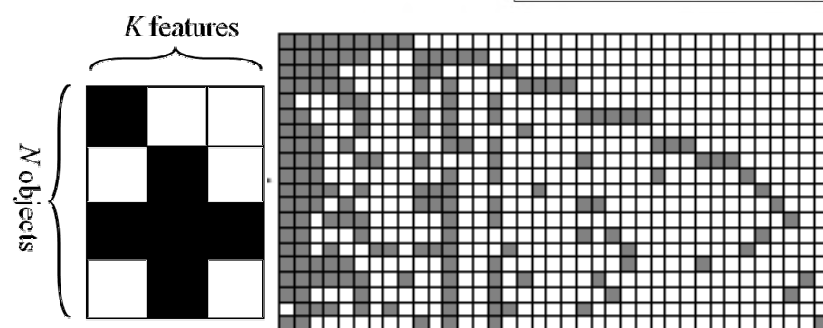
Indian Buffet Process

1st Representation: $K \rightarrow \infty$



Given:

$$\left\{ \begin{array}{l} \Pr(\mathbf{Z} | \alpha) = \prod_{k=1}^K \frac{\frac{\alpha}{K} \Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \\ \text{card}([Z]) = \binom{K}{K_0 \dots K_{2^N-1}} = \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \end{array} \right.$$



Derive when $K \rightarrow \infty$: ($K_+ < \infty$ almost surely)

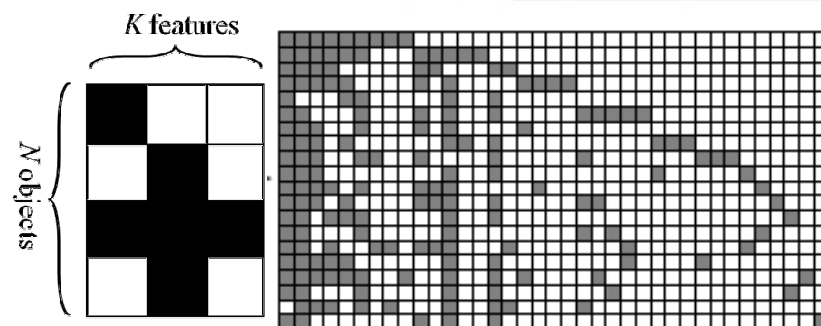
$$\begin{aligned} \Pr([Z] | \alpha) &= \Pr(\mathbf{Z} | \alpha) \cdot \text{card}([Z]) \\ &= \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \left(\frac{\frac{\alpha}{K} \Gamma(\frac{\alpha}{K}) \Gamma(N + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \right)^{K_0} \prod_{k=1}^K \frac{\frac{\alpha}{K} \Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \\ &= \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \left(\frac{\frac{\alpha}{K} \Gamma(\frac{\alpha}{K}) \Gamma(N + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \right)^K \prod_{k=1}^{K_+} \frac{\Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(\frac{\alpha}{K}) \Gamma(N + 1)} \end{aligned}$$

Indian Buffet Process

1st Representation: $K \rightarrow \infty$

Given:

$$\left\{ \begin{array}{l} \Pr(\mathbf{Z} | \alpha) = \prod_{k=1}^K \frac{\frac{\alpha}{K} \Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \\ \text{card}([Z]) = \binom{K}{K_0 \dots K_{2^N-1}} = \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \end{array} \right.$$

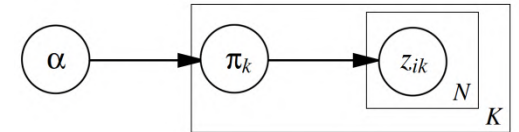


Derive when $K \rightarrow \infty$: ($K_+ < \infty$ almost surely)

$$\begin{aligned} \Pr([Z] | \alpha) &= \Pr(\mathbf{Z} | \alpha) \cdot \text{card}([Z]) \\ &= \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \left(\frac{\frac{\alpha}{K} \Gamma(\frac{\alpha}{K}) \Gamma(N + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \right)^K \prod_{k=1}^{K_+} \frac{\Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(\frac{\alpha}{K}) \Gamma(N + 1)} \end{aligned}$$

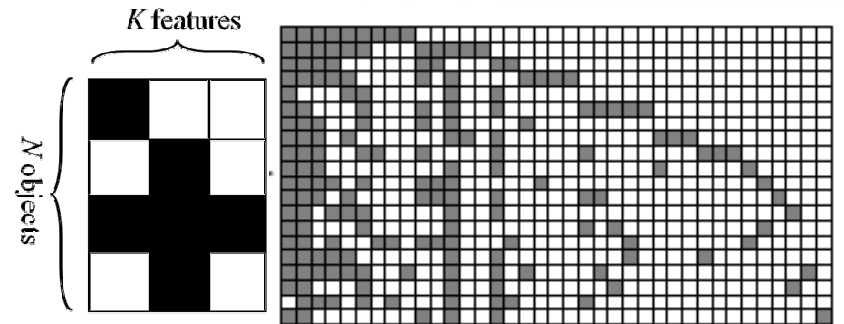
Indian Buffet Process

1st Representation: $K \rightarrow \infty$



Given:

$$\left\{ \begin{array}{l} \Pr(\mathbf{Z} | \alpha) = \prod_{k=1}^K \frac{\frac{\alpha}{K} \Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \\ \text{card}([Z]) = \binom{K}{K_0 \dots K_{2^N-1}} = \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \end{array} \right.$$



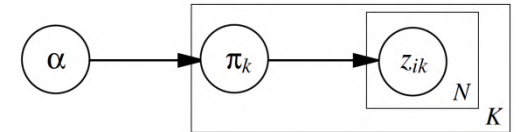
Derive when $K \rightarrow \infty$: ($K_+ < \infty$ almost surely)

$$\begin{aligned} \Pr([Z] | \alpha) &= \Pr(\mathbf{Z} | \alpha) \cdot \text{card}([Z]) \\ &= \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \left(\frac{\frac{\alpha}{K} \Gamma(\frac{\alpha}{K}) \Gamma(N + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \right)^K \prod_{k=1}^{K_+} \frac{\Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(\frac{\alpha}{K}) \Gamma(N + 1)} \end{aligned}$$

$$\prod_{j=0}^{m_k-1} (j + \frac{\alpha}{K})$$

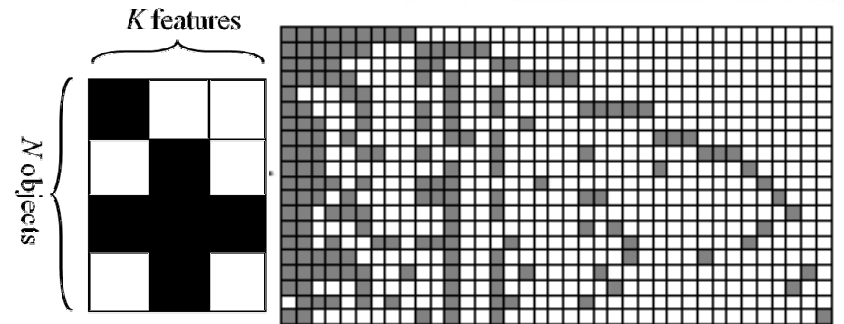
Indian Buffet Process

1st Representation: $K \rightarrow \infty$



Given:

$$\left\{ \begin{array}{l} \Pr(\mathbf{Z} | \alpha) = \prod_{k=1}^K \frac{\frac{\alpha}{K} \Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \\ \text{card}([Z]) = \binom{K}{K_0 \dots K_{2^N-1}} = \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \end{array} \right.$$

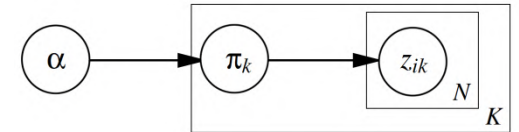


Derive when $K \rightarrow \infty$: ($K_+ < \infty$ almost surely)

$$\begin{aligned} \Pr([Z] | \alpha) &= \Pr(\mathbf{Z} | \alpha) \cdot \text{card}([Z]) \\ &= \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \left(\frac{\frac{\alpha}{K} \Gamma(\frac{\alpha}{K}) \Gamma(N + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \right)^K \prod_{k=1}^{K_+} \frac{\Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(\frac{\alpha}{K}) \Gamma(N + 1)} \end{aligned}$$

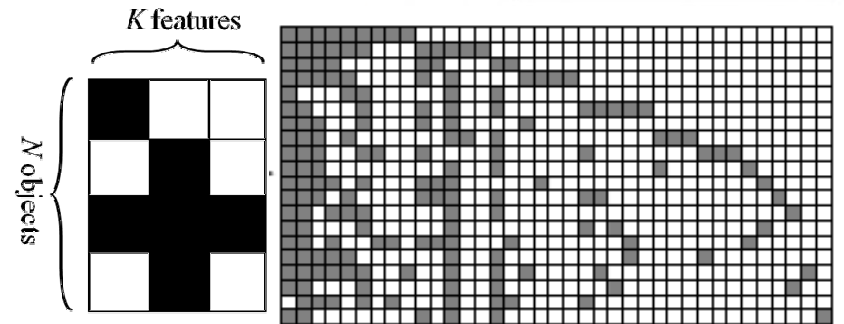
Indian Buffet Process

1st Representation: $K \rightarrow \infty$



Given:

$$\left\{ \begin{aligned} \Pr(\mathbf{Z} | \alpha) &= \prod_{k=1}^K \frac{\frac{\alpha}{K} \Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \\ \text{card}([Z]) &= \binom{K}{K_0 \dots K_{2^N-1}} = \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \end{aligned} \right.$$



Derive when $K \rightarrow \infty$: ($K_+ < \infty$ almost surely)

$$\begin{aligned} \Pr([Z] | \alpha) &= \Pr(\mathbf{Z} | \alpha) \cdot \text{card}([Z]) \\ &= \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \left(\frac{\frac{\alpha}{K} \Gamma(\frac{\alpha}{K}) \Gamma(N + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \right)^K \prod_{k=1}^{K_+} \frac{\Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(\frac{\alpha}{K}) \Gamma(N + 1)} \end{aligned}$$

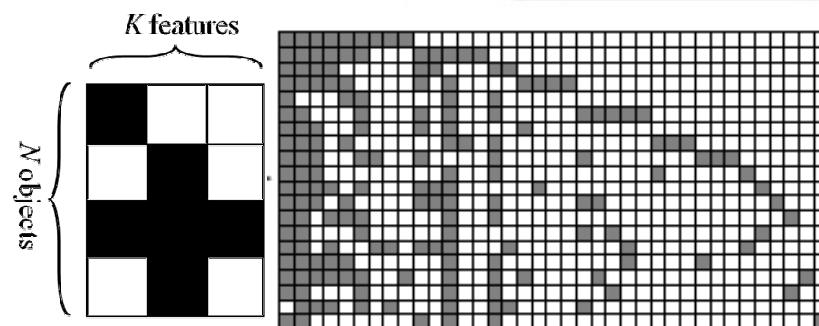
$$(m_k - 1)! \cdot \frac{\alpha}{K}$$

Indian Buffet Process

1st Representation: $K \rightarrow \infty$

Given:

$$\left\{ \begin{array}{l} \Pr(\mathbf{Z} | \alpha) = \prod_{k=1}^K \frac{\frac{\alpha}{K} \Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \\ \text{card}([Z]) = \binom{K}{K_0 \dots K_{2^N-1}} = \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \end{array} \right.$$



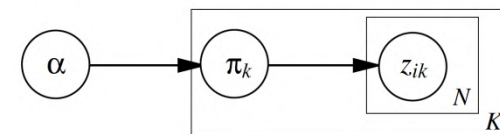
Derive when $K \rightarrow \infty$: ($K_+ < \infty$ almost surely)

$$\begin{aligned} \Pr([Z] | \alpha) &= \Pr(\mathbf{Z} | \alpha) \cdot \text{card}([Z]) \\ &= \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \left(\frac{\frac{\alpha}{K} \Gamma(\frac{\alpha}{K}) \Gamma(N + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \right)^K \prod_{k=1}^{K_+} \frac{\Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(\frac{\alpha}{K}) \Gamma(N + 1)} \end{aligned}$$

$$(m_k - 1)! \cdot \frac{\alpha}{K} \frac{(N - m_k)!}{N!}$$

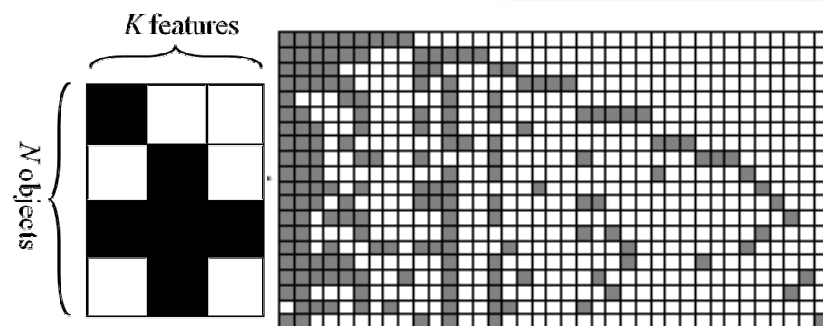
Indian Buffet Process

1st Representation: $K \rightarrow \infty$



Given:

$$\left\{ \begin{array}{l} \Pr(\mathbf{Z} | \alpha) = \prod_{k=1}^K \frac{\frac{\alpha}{K} \Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \\ \text{card}([Z]) = \binom{K}{K_0 \dots K_{2^N-1}} = \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \end{array} \right.$$



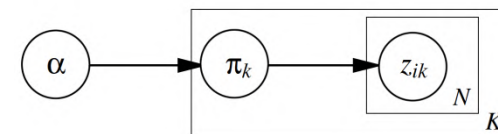
Derive when $K \rightarrow \infty$: ($K_+ < \infty$ almost surely)

$$\begin{aligned} \Pr([Z] | \alpha) &= \Pr(\mathbf{Z} | \alpha) \cdot \text{card}([Z]) \\ &= \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \left(\frac{\frac{\alpha}{K} \Gamma(\frac{\alpha}{K}) \Gamma(N + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \right)^K \prod_{k=1}^{K_+} \frac{\Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(\frac{\alpha}{K}) \Gamma(N + 1)} \end{aligned}$$

$$\prod_{k=1}^{K_+} (m_k - 1)! \cdot \frac{\alpha}{K} \frac{(N - m_k)!}{N!}$$

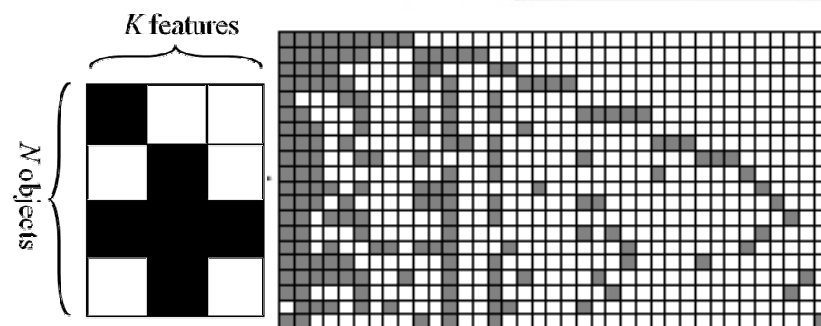
Indian Buffet Process

1st Representation: $K \rightarrow \infty$



Given:

$$\left\{ \begin{array}{l} \Pr(\mathbf{Z} | \alpha) = \prod_{k=1}^K \frac{\frac{\alpha}{K} \Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \\ \text{card}([Z]) = \binom{K}{K_0 \dots K_{2^N-1}} = \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \end{array} \right.$$

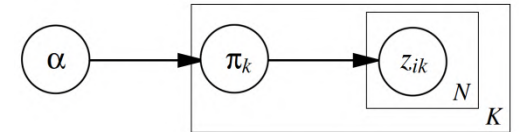


Derive when $K \rightarrow \infty$: ($K_+ < \infty$ almost surely)

$$\begin{aligned} \Pr([Z] | \alpha) &= \Pr(\mathbf{Z} | \alpha) \cdot \text{card}([Z]) \\ &= \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \left(\frac{\frac{\alpha}{K} \Gamma(\frac{\alpha}{K}) \Gamma(N+1)}{\Gamma(N+1+\frac{\alpha}{K})} \right)^K \prod_{k=1}^{K_+} \frac{\Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(\frac{\alpha}{K}) \Gamma(N+1)} \\ &= \left(\frac{\frac{\alpha}{K} \Gamma(\frac{\alpha}{K}) \Gamma(N+1)}{\Gamma(N+1+\frac{\alpha}{K})} \right)^K \prod_{k=1}^{K_+} (m_k - 1)! \cdot \frac{\alpha}{K} \frac{(N - m_k)!}{N!} \end{aligned}$$

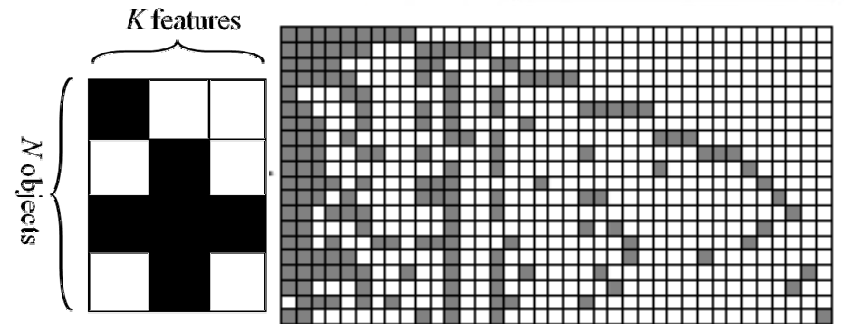
Indian Buffet Process

1st Representation: $K \rightarrow \infty$



Given:

$$\left\{ \begin{array}{l} \Pr(\mathbf{Z} | \alpha) = \prod_{k=1}^K \frac{\frac{\alpha}{K} \Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \\ \text{card}([Z]) = \binom{K}{K_0 \dots K_{2^N-1}} = \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \end{array} \right.$$



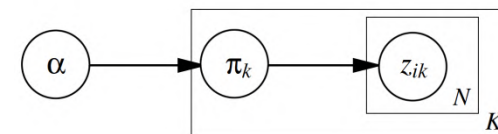
Derive when $K \rightarrow \infty$: ($K_+ < \infty$ almost surely)

$$\begin{aligned} \Pr([Z] | \alpha) &= \Pr(\mathbf{Z} | \alpha) \cdot \text{card}([Z]) \\ &= \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \left(\frac{\frac{\alpha}{K} \Gamma(\frac{\alpha}{K}) \Gamma(N + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \right)^K \prod_{k=1}^{K_+} \frac{\Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(\frac{\alpha}{K}) \Gamma(N + 1)} \end{aligned}$$

$$\prod_{k=1}^{K_+} (m_k - 1)! \cdot \frac{\alpha}{K} \frac{(N - m_k)!}{N!}$$

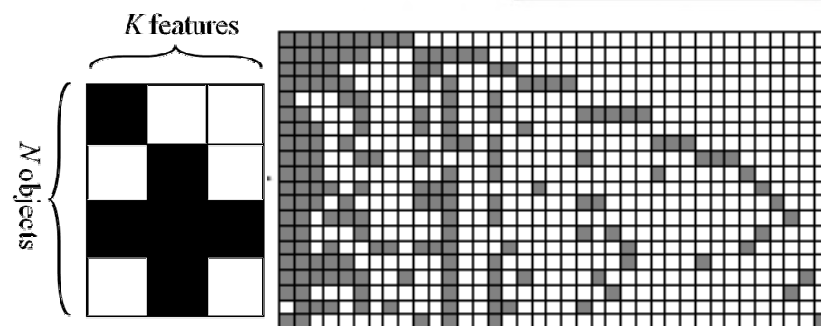
Indian Buffet Process

1st Representation: $K \rightarrow \infty$



Given:

$$\left\{ \begin{array}{l} \Pr(\mathbf{Z} | \alpha) = \prod_{k=1}^K \frac{\frac{\alpha}{K} \Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \\ \text{card}([Z]) = \binom{K}{K_0 \dots K_{2^N-1}} = \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \end{array} \right.$$

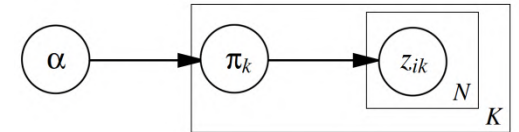


Derive when $K \rightarrow \infty$: ($K_+ < \infty$ almost surely)

$$\begin{aligned} \Pr([Z] | \alpha) &= \Pr(\mathbf{Z} | \alpha) \cdot \text{card}([Z]) \\ &= \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \left(\frac{\frac{\alpha}{K} \Gamma(\frac{\alpha}{K}) \Gamma(N+1)}{\Gamma(N+1+\frac{\alpha}{K})} \right)^K \prod_{k=1}^{K_+} \frac{\Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(\frac{\alpha}{K}) \Gamma(N+1)} \\ &\quad \left(\frac{\Gamma(\frac{\alpha}{K} + 1) \Gamma(N+1)}{\Gamma(N+1+\frac{\alpha}{K})} \right)^K \prod_{k=1}^{K_+} (m_k - 1)! \cdot \frac{\alpha}{K} \frac{(N - m_k)!}{N!} \end{aligned}$$

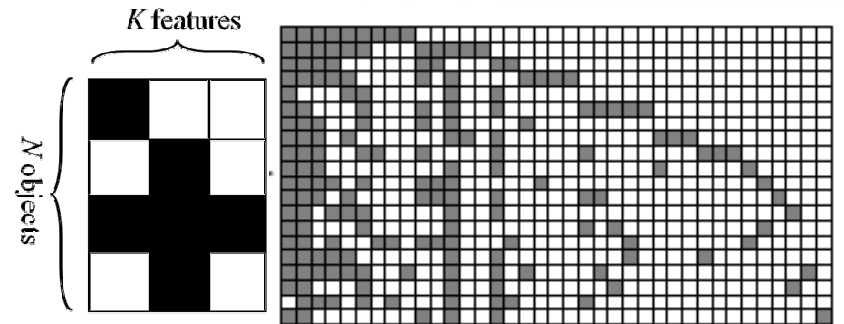
Indian Buffet Process

1st Representation: $K \rightarrow \infty$



Given:

$$\left\{ \begin{array}{l} \Pr(\mathbf{Z} | \alpha) = \prod_{k=1}^K \frac{\frac{\alpha}{K} \Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \\ \text{card}([Z]) = \binom{K}{K_0 \dots K_{2^N-1}} = \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \end{array} \right.$$



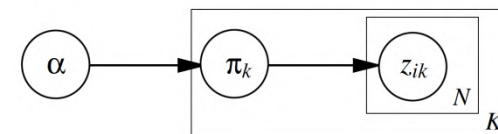
Derive when $K \rightarrow \infty$: ($K_+ < \infty$ almost surely)

$$\begin{aligned} \Pr([Z] | \alpha) &= \Pr(\mathbf{Z} | \alpha) \cdot \text{card}([Z]) \\ &= \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \left(\frac{\frac{\alpha}{K} \Gamma(\frac{\alpha}{K}) \Gamma(N + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \right)^K \prod_{k=1}^{K_+} \frac{\Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(\frac{\alpha}{K}) \Gamma(N + 1)} \end{aligned}$$

$$\prod_{k=1}^{K_+} (m_k - 1)! \cdot \frac{\alpha}{K} \frac{(N - m_k)!}{N!}$$

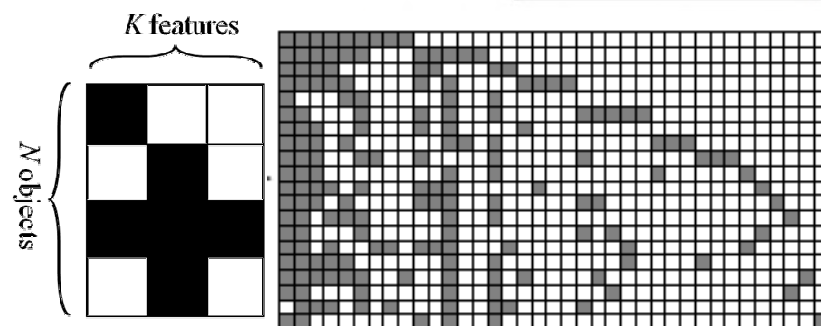
Indian Buffet Process

1st Representation: $K \rightarrow \infty$



Given:

$$\left\{ \begin{array}{l} \Pr(\mathbf{Z} | \alpha) = \prod_{k=1}^K \frac{\frac{\alpha}{K} \Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \\ \text{card}([Z]) = \binom{K}{K_0 \dots K_{2^N-1}} = \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \end{array} \right.$$

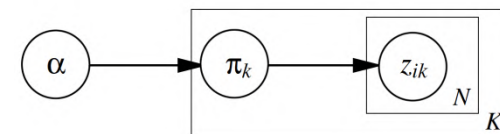


Derive when $K \rightarrow \infty$: ($K_+ < \infty$ almost surely)

$$\begin{aligned} \Pr([Z] | \alpha) &= \Pr(\mathbf{Z} | \alpha) \cdot \text{card}([Z]) \\ &= \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \left(\frac{\frac{\alpha}{K} \Gamma(\frac{\alpha}{K}) \Gamma(N+1)}{\Gamma(N+1+\frac{\alpha}{K})} \right)^K \prod_{k=1}^{K_+} \frac{\Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(\frac{\alpha}{K}) \Gamma(N+1)} \\ &= \left(\frac{N!}{\prod_{j=1}^N (j + \frac{\alpha}{K})} \right)^K \prod_{k=1}^{K_+} (m_k - 1)! \cdot \frac{\alpha}{K} \frac{(N - m_k)!}{N!} \end{aligned}$$

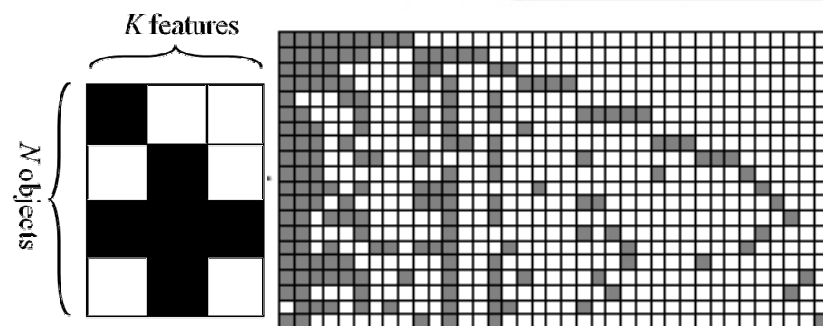
Indian Buffet Process

1st Representation: $K \rightarrow \infty$



Given:

$$\left\{ \begin{array}{l} \Pr(\mathbf{Z} | \alpha) = \prod_{k=1}^K \frac{\frac{\alpha}{K} \Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \\ \text{card}([Z]) = \binom{K}{K_0 \dots K_{2^N-1}} = \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \end{array} \right.$$



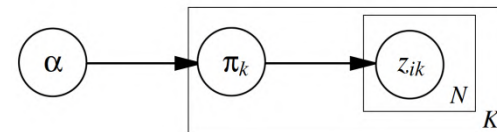
Derive when $K \rightarrow \infty$: ($K_+ < \infty$ almost surely)

$$\begin{aligned} \Pr([Z] | \alpha) &= \Pr(\mathbf{Z} | \alpha) \cdot \text{card}([Z]) \\ &= \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \left(\frac{\frac{\alpha}{K} \Gamma(\frac{\alpha}{K}) \Gamma(N + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \right)^K \prod_{k=1}^{K_+} \frac{\Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(\frac{\alpha}{K}) \Gamma(N + 1)} \end{aligned}$$

$$\prod_{k=1}^{K_+} (m_k - 1)! \cdot \frac{\alpha}{K} \frac{(N - m_k)!}{N!}$$

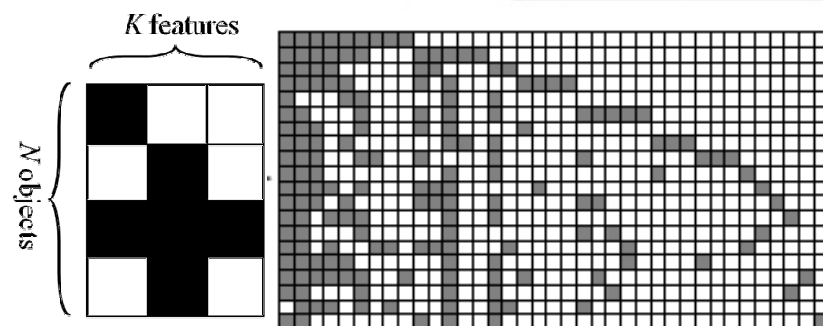
Indian Buffet Process

1st Representation: $K \rightarrow \infty$



Given:

$$\left\{ \begin{array}{l} \Pr(\mathbf{Z} | \alpha) = \prod_{k=1}^K \frac{\frac{\alpha}{K} \Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \\ \text{card}([Z]) = \binom{K}{K_0 \dots K_{2^N-1}} = \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \end{array} \right.$$

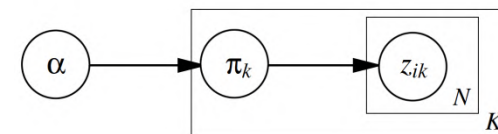


Derive when $K \rightarrow \infty$: ($K_+ < \infty$ almost surely)

$$\begin{aligned} \Pr([Z] | \alpha) &= \Pr(\mathbf{Z} | \alpha) \cdot \text{card}([Z]) \\ &= \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \left(\frac{\frac{\alpha}{K} \Gamma(\frac{\alpha}{K}) \Gamma(N + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \right)^K \prod_{k=1}^{K_+} \frac{\Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(\frac{\alpha}{K}) \Gamma(N + 1)} \\ &\quad \left(\prod_{j=1}^N \frac{1}{(1 + \frac{\alpha}{K} \cdot j)} \right)^K \prod_{k=1}^{K_+} (m_k - 1)! \cdot \frac{\alpha}{K} \frac{(N - m_k)!}{N!} \end{aligned}$$

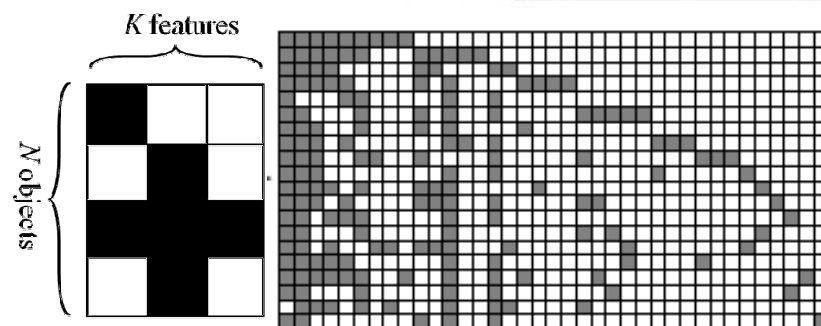
Indian Buffet Process

1st Representation: $K \rightarrow \infty$



Given:

$$\left\{ \begin{array}{l} \Pr(\mathbf{Z} | \alpha) = \prod_{k=1}^K \frac{\frac{\alpha}{K} \Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \\ \text{card}([Z]) = \binom{K}{K_0 \dots K_{2^N-1}} = \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \end{array} \right.$$



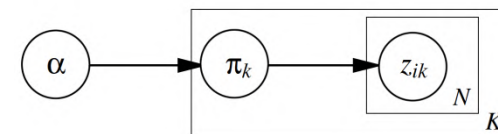
Derive when $K \rightarrow \infty$: ($K_+ < \infty$ almost surely)

$$\begin{aligned} \Pr([Z] | \alpha) &= \Pr(\mathbf{Z} | \alpha) \cdot \text{card}([Z]) \\ &= \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \left(\frac{\frac{\alpha}{K} \Gamma(\frac{\alpha}{K}) \Gamma(N + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \right)^K \prod_{k=1}^{K_+} \frac{\Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(\frac{\alpha}{K}) \Gamma(N + 1)} \end{aligned}$$

$$\prod_{k=1}^{K_+} (m_k - 1)! \cdot \frac{\alpha}{K} \frac{(N - m_k)!}{N!}$$

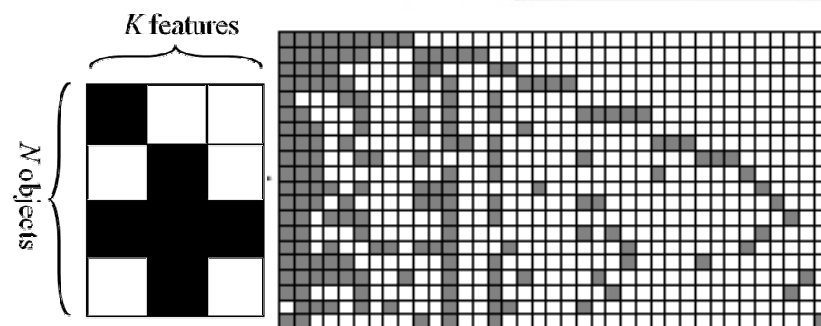
Indian Buffet Process

1st Representation: $K \rightarrow \infty$



Given:

$$\left\{ \begin{array}{l} \Pr(\mathbf{Z} | \alpha) = \prod_{k=1}^K \frac{\frac{\alpha}{K} \Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \\ \text{card}([Z]) = \binom{K}{K_0 \dots K_{2^N-1}} = \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \end{array} \right.$$

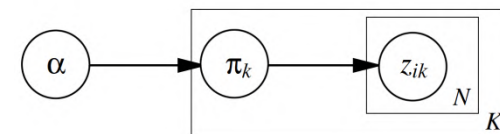


Derive when $K \rightarrow \infty$: ($K_+ < \infty$ almost surely)

$$\begin{aligned} \Pr([Z] | \alpha) &= \Pr(\mathbf{Z} | \alpha) \cdot \text{card}([Z]) \\ &= \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \left[\frac{\frac{\alpha}{K} \Gamma(\frac{\alpha}{K}) \Gamma(N + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \right]^K \prod_{k=1}^{K_+} \frac{\Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(\frac{\alpha}{K}) \Gamma(N + 1)} \\ &= \prod_{j=1}^N \left(1 + \frac{\alpha}{K j}\right)^{-K} \prod_{k=1}^{K_+} (m_k - 1)! \cdot \frac{\alpha}{K} \frac{(N - m_k)!}{N!} \end{aligned}$$

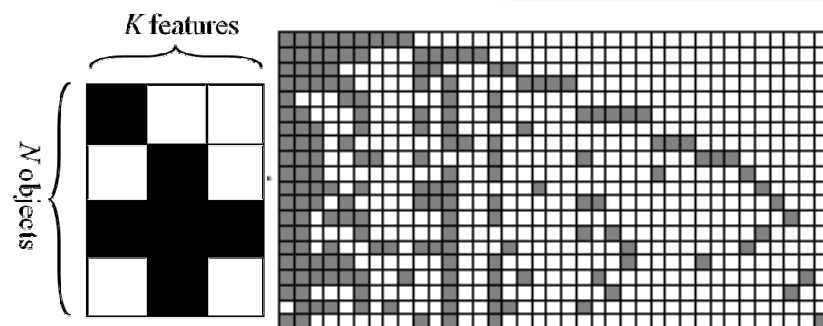
Indian Buffet Process

1st Representation: $K \rightarrow \infty$



Given:

$$\left\{ \begin{array}{l} \Pr(\mathbf{Z} | \alpha) = \prod_{k=1}^K \frac{\frac{\alpha}{K} \Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \\ \text{card}([Z]) = \binom{K}{K_0 \dots K_{2^N-1}} = \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \end{array} \right.$$

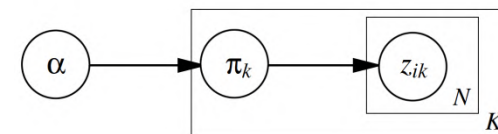


Derive when $K \rightarrow \infty$: ($K_+ < \infty$ almost surely)

$$\begin{aligned} \Pr([Z] | \alpha) &= \Pr(\mathbf{Z} | \alpha) \cdot \text{card}([Z]) \\ &= \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \left[\frac{\frac{\alpha}{K} \Gamma(\frac{\alpha}{K}) \Gamma(N + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \right]^K \prod_{k=1}^{K_+} \frac{\Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(\frac{\alpha}{K}) \Gamma(N + 1)} \\ &= \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \prod_{j=1}^N \left(1 + \frac{\alpha / j}{K} \right)^{-K} \prod_{k=1}^{K_+} (m_k - 1)! \cdot \frac{\alpha}{K} \frac{(N - m_k)!}{N!} \end{aligned}$$

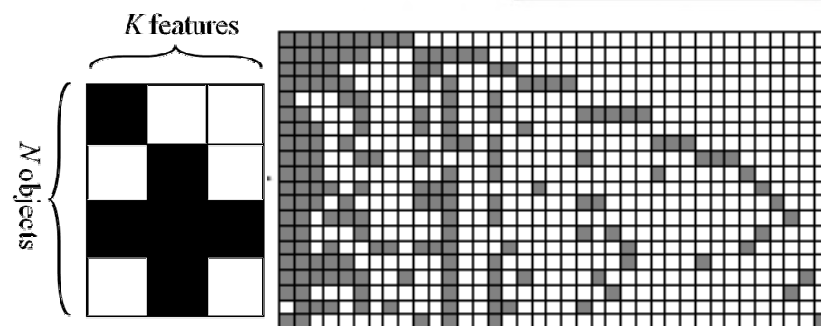
Indian Buffet Process

1st Representation: $K \rightarrow \infty$



Given:

$$\left\{ \begin{array}{l} \Pr(\mathbf{Z} | \alpha) = \prod_{k=1}^K \frac{\frac{\alpha}{K} \Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \\ \text{card}([Z]) = \binom{K}{K_0 \dots K_{2^N-1}} = \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \end{array} \right.$$

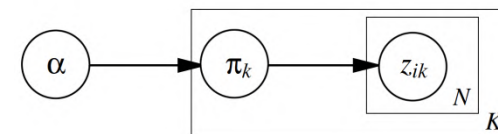


Derive when $K \rightarrow \infty$: ($K_+ < \infty$ almost surely)

$$\begin{aligned} \Pr([Z] | \alpha) &= \Pr(\mathbf{Z} | \alpha) \cdot \text{card}([Z]) \\ &= \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \left[\frac{\frac{\alpha}{K} \Gamma(\frac{\alpha}{K}) \Gamma(N + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \right]^K \prod_{k=1}^{K_+} \frac{\Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(\frac{\alpha}{K}) \Gamma(N + 1)} \\ &= \prod_{j=1}^N \left(1 + \frac{\alpha / j}{K} \right)^{-K} \prod_{k=1}^{K_+} (m_k - 1)! \cdot \frac{\alpha}{K} \frac{(N - m_k)!}{N!} \end{aligned}$$

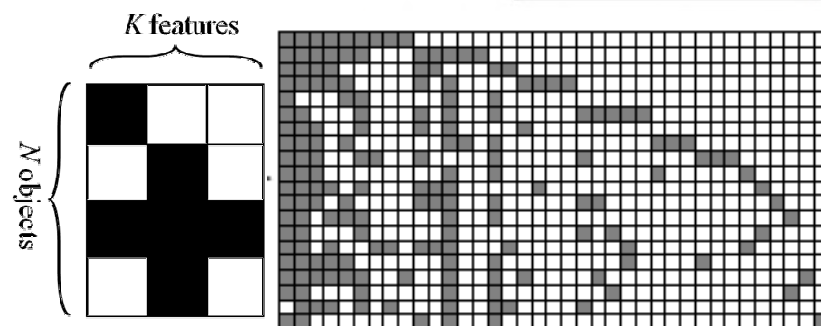
Indian Buffet Process

1st Representation: $K \rightarrow \infty$



Given:

$$\left\{ \begin{array}{l} \Pr(\mathbf{Z} | \alpha) = \prod_{k=1}^K \frac{\frac{\alpha}{K} \Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \\ \text{card}([Z]) = \binom{K}{K_0 \dots K_{2^N-1}} = \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \end{array} \right.$$



Derive when $K \rightarrow \infty$: ($K_+ < \infty$ almost surely)

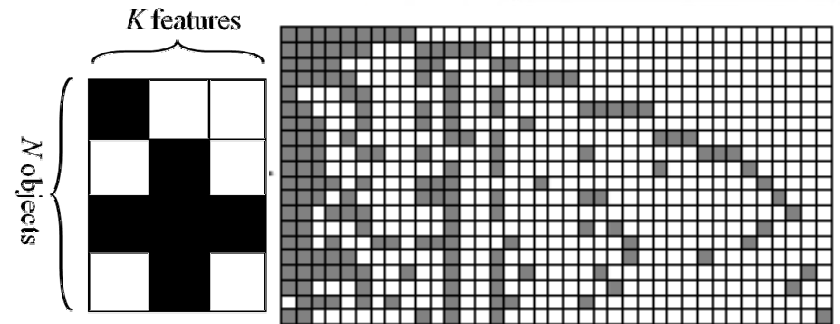
$$\begin{aligned} \Pr([Z] | \alpha) &= \Pr(\mathbf{Z} | \alpha) \cdot \text{card}([Z]) \\ &= \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \left[\frac{\frac{\alpha}{K} \Gamma(\frac{\alpha}{K}) \Gamma(N + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \right]^K \prod_{k=1}^{K_+} \frac{\Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(\frac{\alpha}{K}) \Gamma(N + 1)} \\ &= \frac{K!}{K_0! \prod_{h=1}^{2^N-1} K_h!} \prod_{j=1}^N \left(1 + \frac{\alpha / j}{K} \right)^{-K} \prod_{k=1}^{K_+} (m_k - 1)! \cdot \frac{\alpha}{K} \frac{(N - m_k)!}{N!} \end{aligned}$$

Indian Buffet Process

1st Representation: $K \rightarrow \infty$

Given:

$$\left\{ \begin{array}{l} \Pr(\mathbf{Z} | \alpha) = \prod_{k=1}^K \frac{\frac{\alpha}{K} \Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \\ \text{card}([Z]) = \binom{K}{K_0 \dots K_{2^N-1}} = \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \end{array} \right.$$



Derive when $K \rightarrow \infty$: ($K_+ < \infty$ almost surely)

$$\Pr([Z] | \alpha) = \Pr(\mathbf{Z} | \alpha) \cdot \text{card}([Z])$$

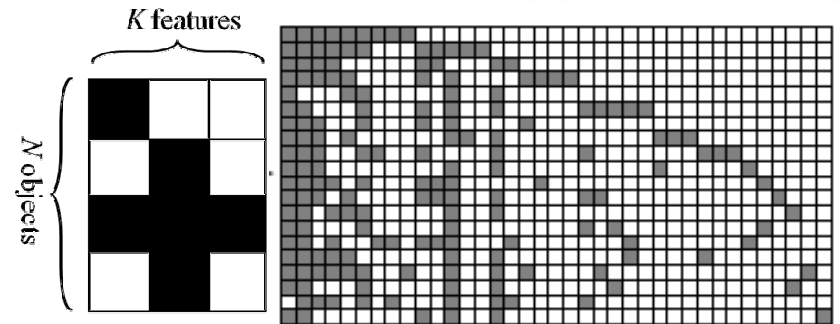
$$= \frac{K!}{K_0! \prod_{h=1}^{2^N-1} K_h!} \prod_{j=1}^N \left(1 + \frac{\alpha / j}{K}\right)^{-K} \prod_{k=1}^{K_+} (m_k - 1)! \cdot \frac{\alpha}{K} \frac{(N - m_k)!}{N!}$$

Indian Buffet Process

1st Representation: $K \rightarrow \infty$

Given:

$$\left\{ \begin{array}{l} \Pr(\mathbf{Z} | \alpha) = \prod_{k=1}^K \frac{\frac{\alpha}{K} \Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \\ \text{card}([Z]) = \binom{K}{K_0 \dots K_{2^N-1}} = \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \end{array} \right.$$



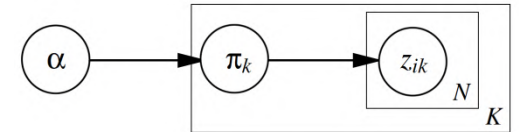
Derive when $K \rightarrow \infty$: ($K_+ < \infty$ almost surely)

$$\Pr([Z] | \alpha) = \Pr(\mathbf{Z} | \alpha) \cdot \text{card}([Z])$$

$$= \frac{K!}{K_0! \prod_{h=1}^{2^N-1} K_h!} \prod_{j=1}^N \left(1 + \frac{\alpha / j}{K}\right)^{-K} \prod_{k=1}^{K_+} (m_k - 1)! \cdot \frac{\alpha}{K} \frac{(N - m_k)!}{N!}$$

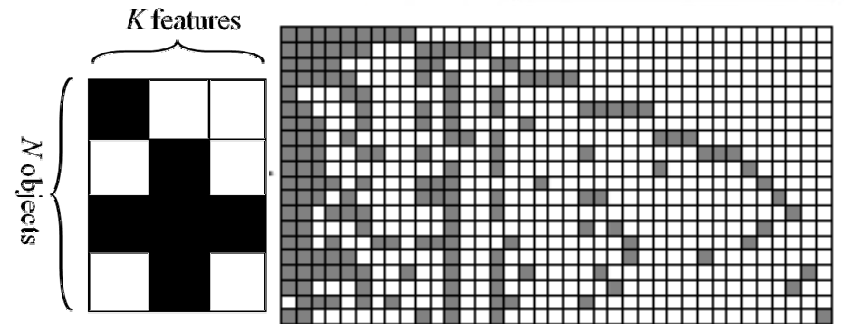
Indian Buffet Process

1st Representation: $K \rightarrow \infty$



Given:

$$\left\{ \begin{array}{l} \Pr(\mathbf{Z} | \alpha) = \prod_{k=1}^K \frac{\frac{\alpha}{K} \Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \\ \text{card}([Z]) = \binom{K}{K_0 \dots K_{2^N-1}} = \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \end{array} \right.$$



Derive when $K \rightarrow \infty$: ($K_+ < \infty$ almost surely)

$$\Pr([Z] | \alpha) = \Pr(\mathbf{Z} | \alpha) \cdot \text{card}([Z])$$

$$= \frac{K!}{K_0! \prod_{h=1}^{2^N-1} K_h!} \prod_{j=1}^N \left(1 + \frac{\alpha / j}{K}\right)^{-K}$$

$$\prod_{k=1}^{K_+} (m_k - 1)! \cdot \frac{\alpha}{K} \frac{(N - m_k)!}{N!}$$

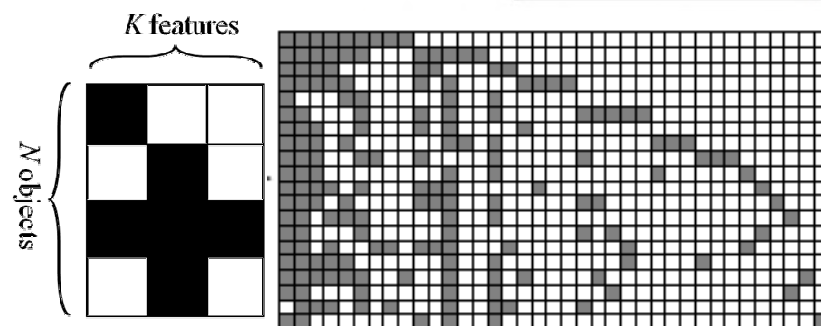
$$\frac{\alpha^{K_+}}{K^{K_+}} \prod_{k=1}^{K_+} (m_k - 1)! \frac{(N - m_k)!}{N!}$$

Indian Buffet Process

1st Representation: $K \rightarrow \infty$

Given:

$$\left\{ \begin{array}{l} \Pr(\mathbf{Z} | \alpha) = \prod_{k=1}^K \frac{\frac{\alpha}{K} \Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \\ \text{card}([Z]) = \binom{K}{K_0 \dots K_{2^N-1}} = \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \end{array} \right.$$



Derive when $K \rightarrow \infty$: ($K_+ < \infty$ almost surely)

$$\Pr([Z] | \alpha) = \Pr(\mathbf{Z} | \alpha) \cdot \text{card}([Z])$$

$$= \frac{K!}{K_0! \prod_{h=1}^{2^N-1} K_h!} \prod_{j=1}^N \left(1 + \frac{\alpha / j}{K}\right)^{-K} \prod_{k=1}^{K_+} (m_k - 1)! \frac{\alpha}{K} \frac{(N - m_k)!}{N!}$$

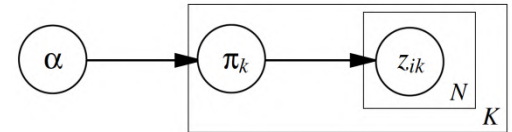
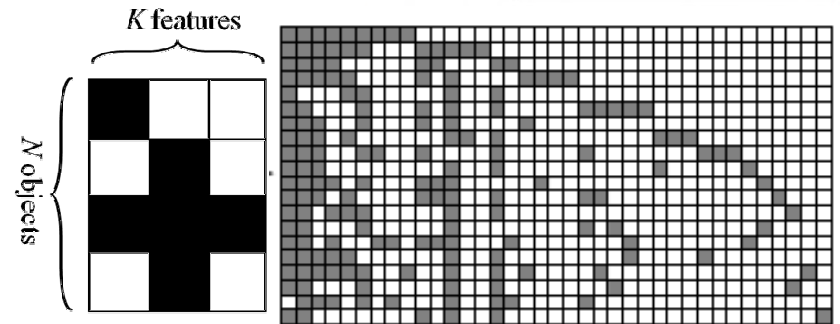
$$\frac{K!}{K_0! \prod_{h=1}^{2^N-1} K_h!} \frac{\alpha^{K_+}}{K^{K_+}} \prod_{k=1}^{K_+} (m_k - 1)! \frac{(N - m_k)!}{N!}$$

Indian Buffet Process

1st Representation: $K \rightarrow \infty$

Given:

$$\left\{ \begin{array}{l} \Pr(\mathbf{Z} | \alpha) = \prod_{k=1}^K \frac{\frac{\alpha}{K} \Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \\ \text{card}([Z]) = \binom{K}{K_0 \dots K_{2^N-1}} = \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \end{array} \right.$$



Derive when $K \rightarrow \infty$: ($K_+ < \infty$ almost surely)

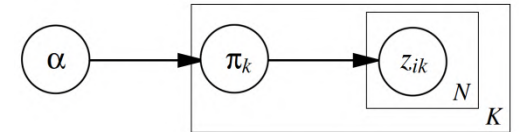
$$\Pr([Z] | \alpha) = \Pr(\mathbf{Z} | \alpha) \cdot \text{card}([Z])$$

$$= \frac{K!}{K_0! \prod_{h=1}^{2^N-1} K_h!} \prod_{j=1}^N \left(1 + \frac{\alpha / j}{K}\right)^{-K} \prod_{k=1}^{K_+} (m_k - 1)! \frac{\alpha}{K} \frac{(N - m_k)!}{N!}$$

$$\prod_{k=1}^{K_+} (m_k - 1)! \frac{(N - m_k)!}{N!}$$

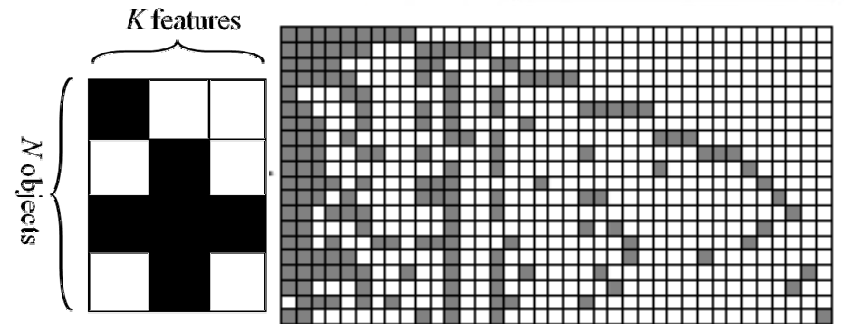
Indian Buffet Process

1st Representation: $K \rightarrow \infty$



Given:

$$\left\{ \begin{array}{l} \Pr(\mathbf{Z} | \alpha) = \prod_{k=1}^K \frac{\frac{\alpha}{K} \Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \\ \text{card}([Z]) = \binom{K}{K_0 \dots K_{2^N-1}} = \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \end{array} \right.$$



Derive when $K \rightarrow \infty$: ($K_+ < \infty$ almost surely)

$$\Pr([Z] | \alpha) = \Pr(\mathbf{Z} | \alpha) \cdot \text{card}([Z])$$

$$= \frac{K!}{K_0! \prod_{h=1}^{2^N-1} K_h!} \prod_{j=1}^N \left(1 + \frac{\alpha / j}{K}\right)^{-K} \prod_{k=1}^{K_+} (m_k - 1)! \frac{\alpha}{K} \frac{(N - m_k)!}{N!}$$

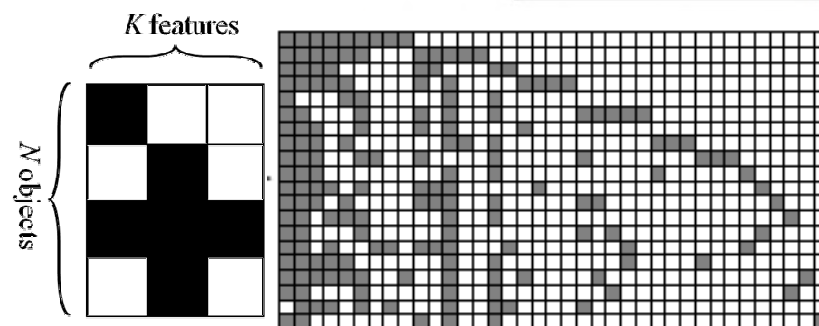
$$\frac{K!}{K_0! K^{K_+}} \frac{\alpha^{K_+}}{\prod_{h=1}^{2^N-1} K_h!} \prod_{k=1}^{K_+} (m_k - 1)! \frac{(N - m_k)!}{N!}$$

Indian Buffet Process

1st Representation: $K \rightarrow \infty$

Given:

$$\left\{ \begin{array}{l} \Pr(\mathbf{Z} | \alpha) = \prod_{k=1}^K \frac{\frac{\alpha}{K} \Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \\ \text{card}([Z]) = \binom{K}{K_0 \dots K_{2^N-1}} = \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \end{array} \right.$$



Derive when $K \rightarrow \infty$: ($K_+ < \infty$ almost surely)

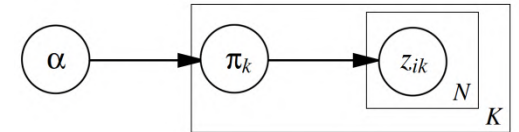
$$\Pr([Z] | \alpha) = \Pr(\mathbf{Z} | \alpha) \cdot \text{card}([Z])$$

$$= \frac{K!}{K_0! \prod_{h=1}^{2^N-1} K_h!} \prod_{j=1}^N \left(1 + \frac{\alpha / j}{K}\right)^{-K} \prod_{k=1}^{K_+} (m_k - 1)! \frac{\alpha}{K} \frac{(N - m_k)!}{N!}$$

$$\prod_{k=1}^{K_+} (m_k - 1)! \frac{(N - m_k)!}{N!}$$

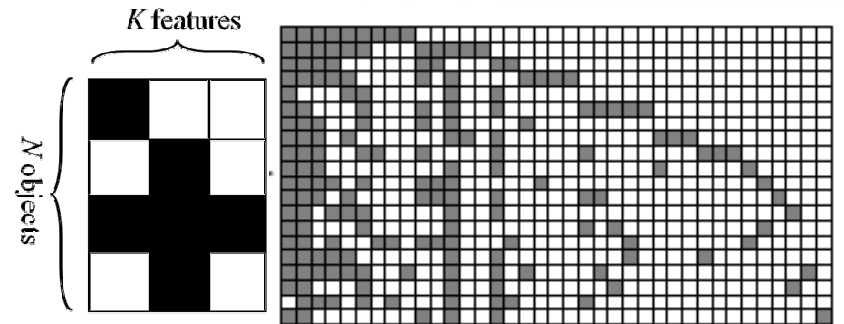
Indian Buffet Process

1st Representation: $K \rightarrow \infty$



Given:

$$\left\{ \begin{array}{l} \Pr(\mathbf{Z} | \alpha) = \prod_{k=1}^K \frac{\frac{\alpha}{K} \Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \\ \text{card}([Z]) = \binom{K}{K_0 \dots K_{2^N-1}} = \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \end{array} \right.$$



Derive when $K \rightarrow \infty$: ($K_+ < \infty$ almost surely)

$$\Pr([Z] | \alpha) = \Pr(\mathbf{Z} | \alpha) \cdot \text{card}([Z])$$

$$= \frac{K!}{K_0! \prod_{h=1}^{2^N-1} K_h!} \prod_{j=1}^N \left(1 + \frac{\alpha / j}{K}\right)^{-K}$$

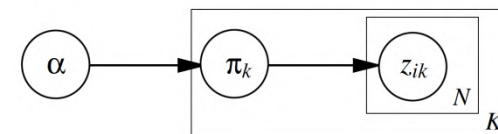
$$\prod_{k=1}^{K_+} (m_k - 1)! \cdot \frac{\alpha}{K} \frac{(N - m_k)!}{N!}$$

$$\frac{\alpha^{K_+}}{\prod_{h=1}^{2^N-1} K_h!}$$

$$\prod_{k=1}^{K_+} (m_k - 1)! \frac{(N - m_k)!}{N!}$$

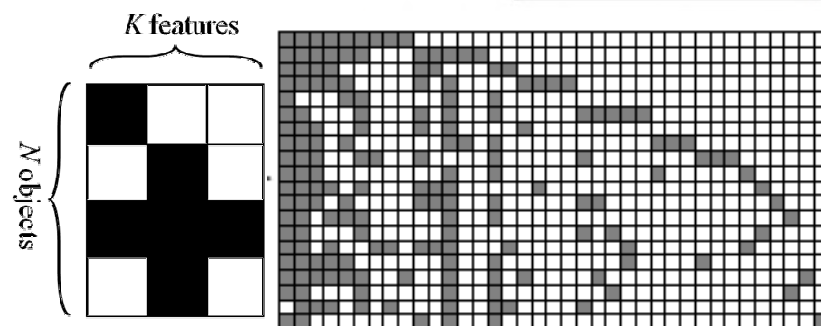
Indian Buffet Process

1st Representation: $K \rightarrow \infty$



Given:

$$\left\{ \begin{aligned} \Pr(\mathbf{Z} | \alpha) &= \prod_{k=1}^K \frac{\frac{\alpha}{K} \Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \\ \text{card}([Z]) &= \binom{K}{K_0 \dots K_{2^N-1}} = \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \end{aligned} \right.$$



Derive when $K \rightarrow \infty$: ($K_+ < \infty$ almost surely)

$$\Pr([Z] | \alpha) = \Pr(\mathbf{Z} | \alpha) \cdot \text{card}([Z])$$

$$= \frac{K!}{K_0! \prod_{h=1}^{2^N-1} K_h!} \prod_{j=1}^N \left(1 + \frac{\alpha / j}{K}\right)^{-K}$$

$$\prod_{k=1}^{K_+} (m_k - 1)! \cdot \frac{\alpha}{K} \frac{(N - m_k)!}{N!}$$

$$\prod_{j=1}^N e^{-\frac{\alpha / j}{K}}$$

$$\frac{\alpha^{K_+}}{\prod_{h=1}^{2^N-1} K_h!}$$

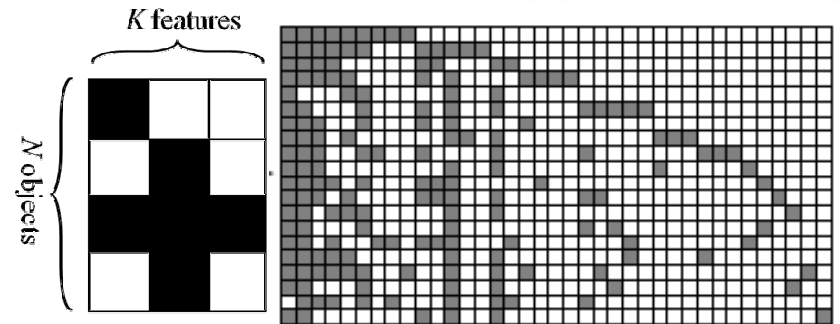
$$\prod_{k=1}^{K_+} (m_k - 1)! \frac{(N - m_k)!}{N!}$$

Indian Buffet Process

1st Representation: $K \rightarrow \infty$

Given:

$$\left\{ \begin{array}{l} \Pr(\mathbf{Z} | \alpha) = \prod_{k=1}^K \frac{\frac{\alpha}{K} \Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \\ \text{card}([Z]) = \binom{K}{K_0 \dots K_{2^N-1}} = \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \end{array} \right.$$



Derive when $K \rightarrow \infty$: ($K_+ < \infty$ almost surely)

$$\Pr([Z] | \alpha) = \Pr(\mathbf{Z} | \alpha) \cdot \text{card}([Z])$$

$$= \frac{K!}{K_0! \prod_{h=1}^{2^N-1} K_h!} \prod_{j=1}^N \left(1 + \frac{\alpha / j}{K}\right)^{-K} \prod_{k=1}^{K_+} (m_k - 1)! \frac{\alpha}{K} \frac{(N - m_k)!}{N!}$$

$$\frac{\alpha^{K_+}}{\prod_{h=1}^{2^N-1} K_h!}$$

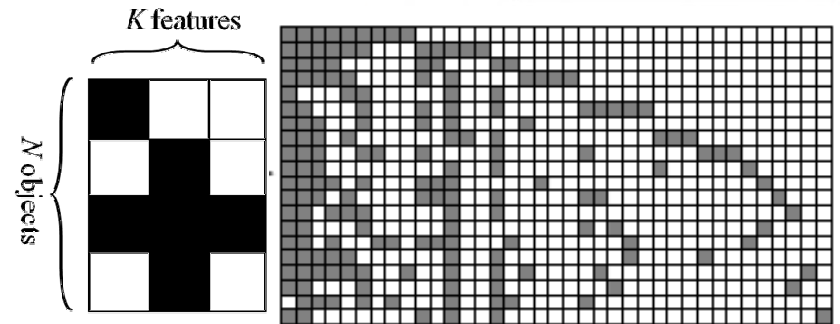
$$\prod_{k=1}^{K_+} (m_k - 1)! \frac{(N - m_k)!}{N!}$$

Indian Buffet Process

1st Representation: $K \rightarrow \infty$

Given:

$$\left\{ \begin{array}{l} \Pr(\mathbf{Z} | \alpha) = \prod_{k=1}^K \frac{\frac{\alpha}{K} \Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \\ \text{card}([Z]) = \binom{K}{K_0 \dots K_{2^N-1}} = \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \end{array} \right.$$



Derive when $K \rightarrow \infty$: ($K_+ < \infty$ almost surely)

$$\Pr([Z] | \alpha) = \Pr(\mathbf{Z} | \alpha) \cdot \text{card}([Z])$$

$$= \frac{K!}{K_0! \prod_{h=1}^{2^N-1} K_h!} \prod_{j=1}^N \left(1 + \frac{\alpha/j}{K}\right)^{-K}$$

$$\prod_{k=1}^{K_+} (m_k - 1)! \cdot \frac{\alpha}{K} \frac{(N - m_k)!}{N!}$$

$$= e^{-\alpha H_N} \frac{\alpha^{K_+}}{\prod_{h=1}^{2^N-1} K_h!}$$

$$\prod_{k=1}^{K_+} (m_k - 1)! \frac{(N - m_k)!}{N!}$$

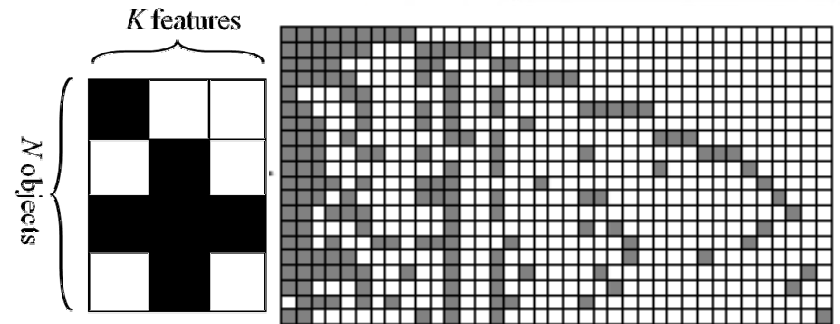
$$H_N = \sum_{j=1}^N \frac{1}{j} \quad \text{Harmonic sequence sum}$$

Indian Buffet Process

1st Representation: $K \rightarrow \infty$

Given:

$$\left\{ \begin{array}{l} \Pr(\mathbf{Z} | \alpha) = \prod_{k=1}^K \frac{\frac{\alpha}{K} \Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \\ \text{card}([Z]) = \binom{K}{K_0 \dots K_{2^N-1}} = \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \end{array} \right.$$



Derive when $K \rightarrow \infty$: ($K_+ < \infty$ almost surely)

$$\begin{aligned} \Pr([Z] | \alpha) &= \Pr(\mathbf{Z} | \alpha) \cdot \text{card}([Z]) \\ &= e^{-\alpha H_N} \frac{\alpha^{K_+}}{\prod_{h=1}^{2^N-1} K_h!} \prod_{k=1}^{K_+} \frac{(m_k - 1)!(N - m_k)!}{N!} \end{aligned}$$

Note:

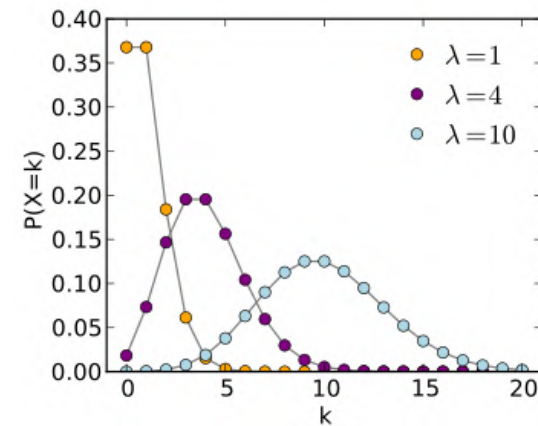
- $\Pr([Z] | \alpha)$ is well defined because $K_+ < \infty$ a.s.
- $\Pr([Z] | \alpha)$ depends on \mathbf{K}_h :
 - The **number of features** (columns) with history h
 - Permute the rows (data points) does not change \mathbf{K}_h (exchangeability)

Indian Buffet Process

2st Representation: Customers & Dishes

- Indian restaurant with infinitely many infinite dishes (columns)

Poisson Distribution: $p(k | \lambda) = \frac{\lambda^k e^{-\lambda}}{k!}$



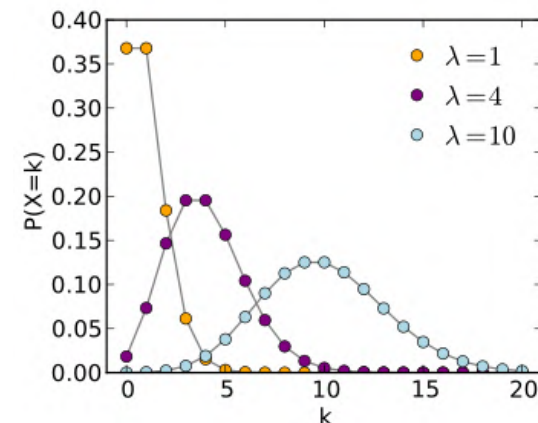
Indian Buffet Process

2st Representation: Customers & Dishes

- Indian restaurant with infinitely many infinite dishes (columns)
 - The first customer tastes **first** $K_1^{(1)}$ dishes, sample $K_1^{(1)} \sim \text{Poisson}(\alpha)$



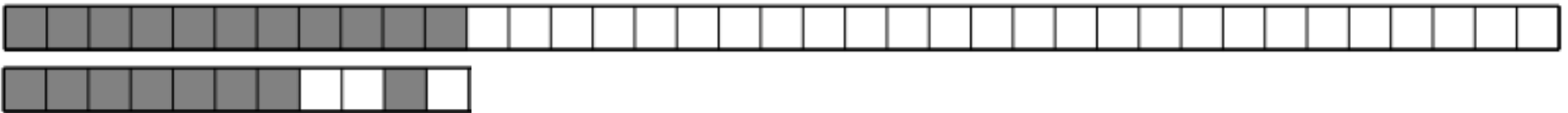
Poisson Distribution: $p(k | \lambda) = \frac{\lambda^k e^{-\lambda}}{k!}$



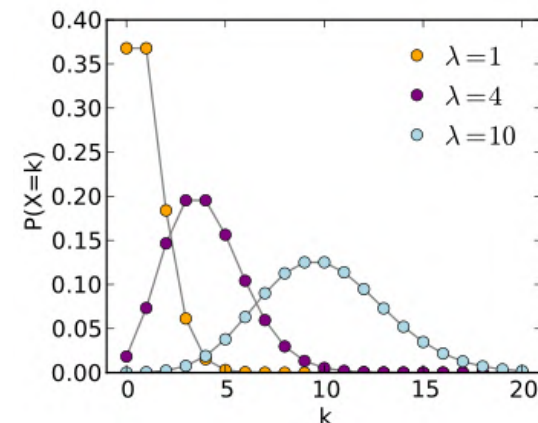
Indian Buffet Process

2st Representation: Customers & Dishes

- Indian restaurant with infinitely many infinite dishes (columns)
 - The first customer tastes **first** $K_1^{(1)}$ dishes, sample $K_1^{(1)} \sim \text{Poisson}(\alpha)$
 - The i -th customer:
 - Taste a previously sampled dish with probability $\frac{m_k}{i+1}$
 - m_k : number of previously customers taking dish k



Poisson Distribution: $p(k | \lambda) = \frac{\lambda^k e^{-\lambda}}{k!}$



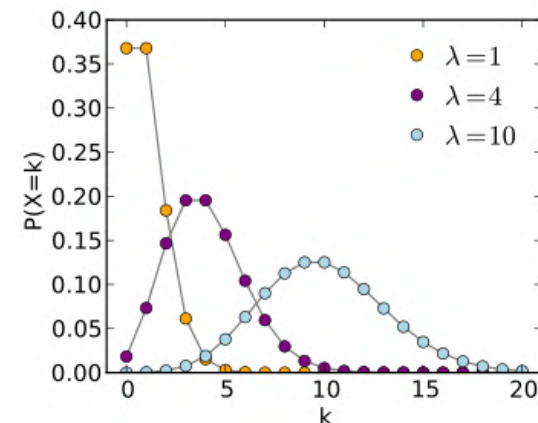
Indian Buffet Process

2st Representation: Customers & Dishes

- Indian restaurant with infinitely many infinite dishes (columns)
 - The first customer tastes **first** $K_1^{(1)}$ dishes, sample $K_1^{(1)} \sim \text{Poisson}(\alpha)$
 - The i -th customer:
 - Taste a previously sampled dish with probability $\frac{m_k}{i+1}$
 - m_k : number of previously customers taking dish k
 - Taste **following** $K_1^{(i)}$ new dishes, sample $K_1^{(i)} \sim \text{Poisson}(\frac{\alpha}{i})$



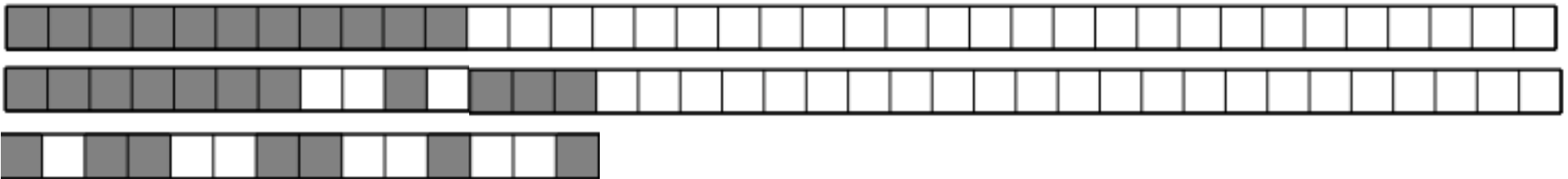
Poisson Distribution: $p(k | \lambda) = \frac{\lambda^k e^{-\lambda}}{k!}$



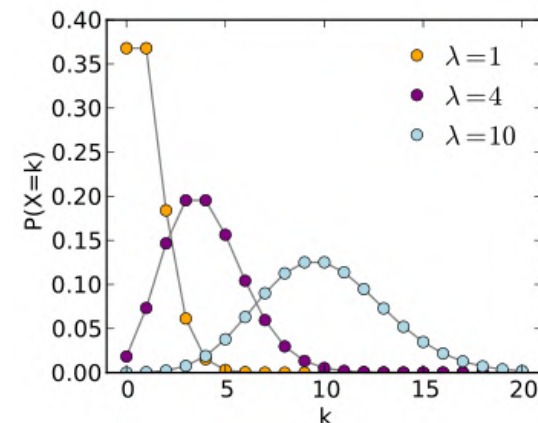
Indian Buffet Process

2st Representation: Customers & Dishes

- Indian restaurant with infinitely many infinite dishes (columns)
 - The first customer tastes **first** $K_1^{(1)}$ dishes, sample $K_1^{(1)} \sim \text{Poisson}(\alpha)$
 - The i -th customer:
 - Taste a previously sampled dish with probability $\frac{m_k}{i+1}$
 - m_k : number of previously customers taking dish k
 - Taste **following** $K_1^{(i)}$ new dishes, sample $K_1^{(i)} \sim \text{Poisson}(\frac{\alpha}{i})$



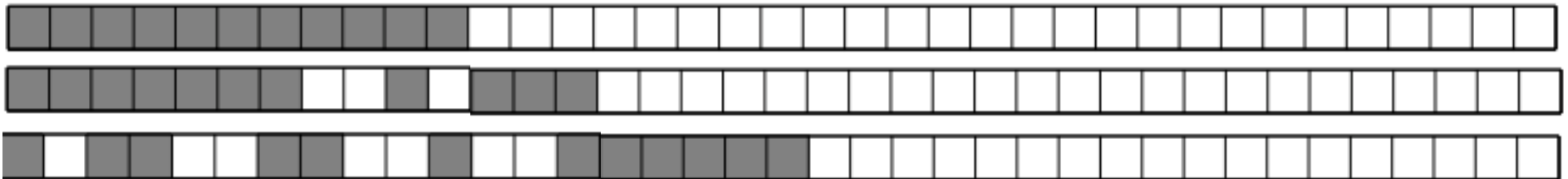
Poisson Distribution: $p(k | \lambda) = \frac{\lambda^k e^{-\lambda}}{k!}$



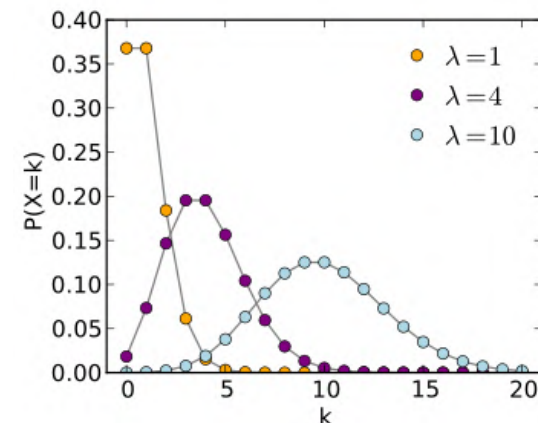
Indian Buffet Process

2st Representation: Customers & Dishes

- Indian restaurant with infinitely many infinite dishes (columns)
 - The first customer tastes **first** $K_1^{(1)}$ dishes, sample $K_1^{(1)} \sim \text{Poisson}(\alpha)$
 - The i -th customer:
 - Taste a previously sampled dish with probability $\frac{m_k}{i+1}$
 - m_k : number of previously customers taking dish k
 - Taste **following** $K_1^{(i)}$ new dishes, sample $K_1^{(i)} \sim \text{Poisson}(\frac{\alpha}{i})$



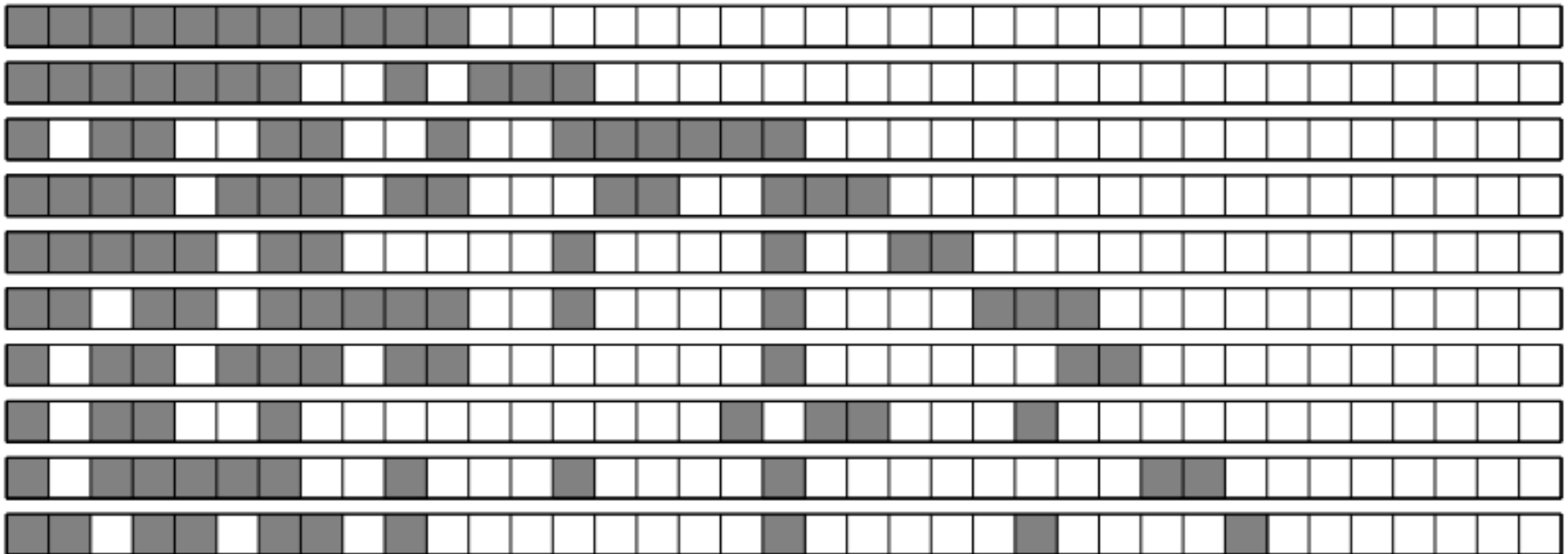
Poisson Distribution: $p(k | \lambda) = \frac{\lambda^k e^{-\lambda}}{k!}$



Indian Buffet Process

2st Representation: Customers & Dishes

- Indian restaurant with infinitely many infinite dishes (columns)
 - The first customer tastes **first** $K_1^{(1)}$ dishes, sample $K_1^{(1)} \sim \text{Poisson}(\alpha)$
 - The i -th customer:
 - Taste a previously sampled dish with probability $\frac{m_k}{i+1}$
 - m_k : number of previously customers taking dish k
 - Taste **following** $K_1^{(i)}$ new dishes, sample $K_1^{(i)} \sim \text{Poisson}(\frac{\alpha}{i})$

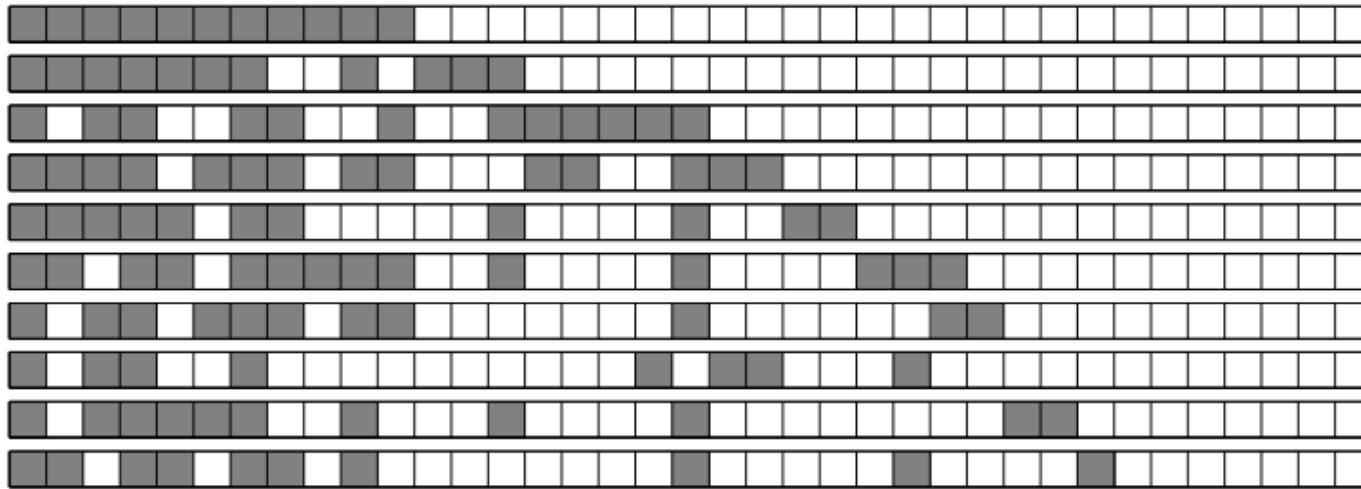


Indian Buffet Process

2st Representation: Customers & Dishes

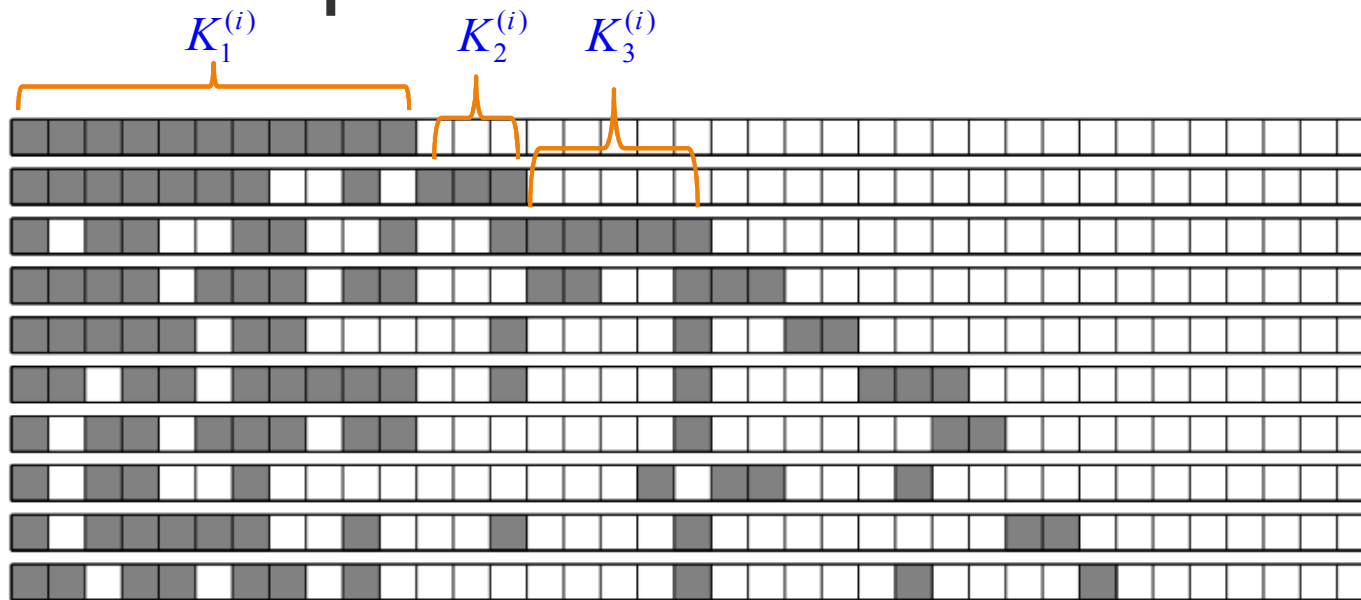
Indian Buffet Process

2st Representation: Customers & Dishes



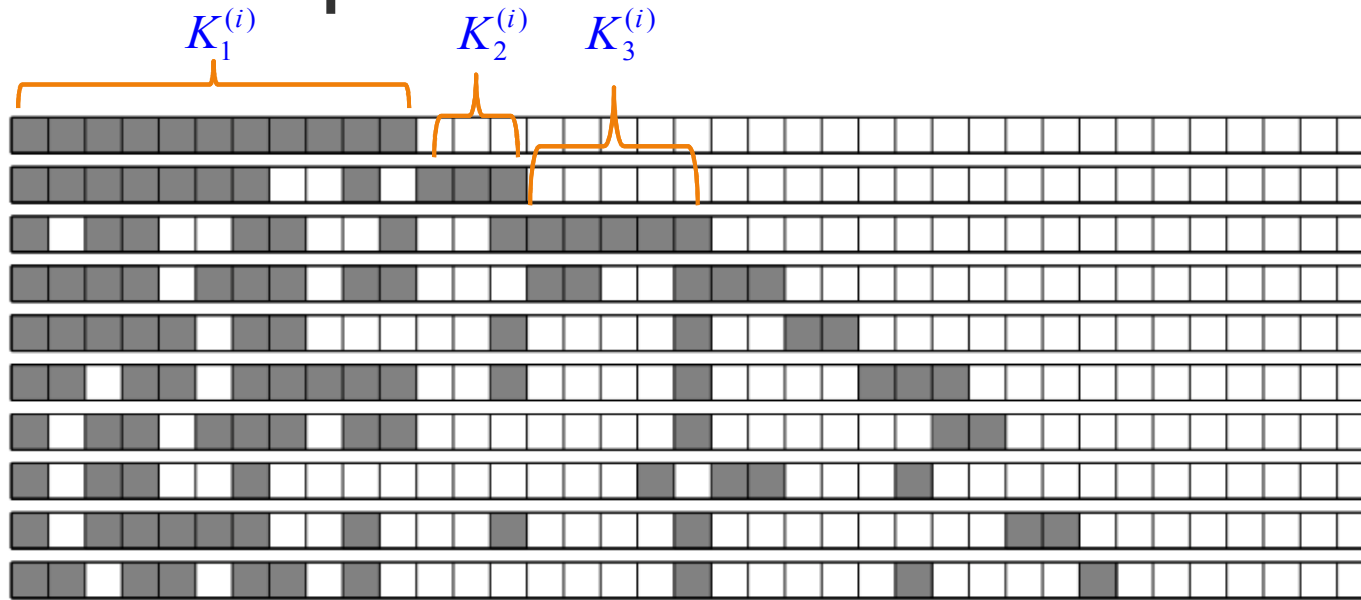
Indian Buffet Process

2st Representation: Customers & Dishes



Indian Buffet Process

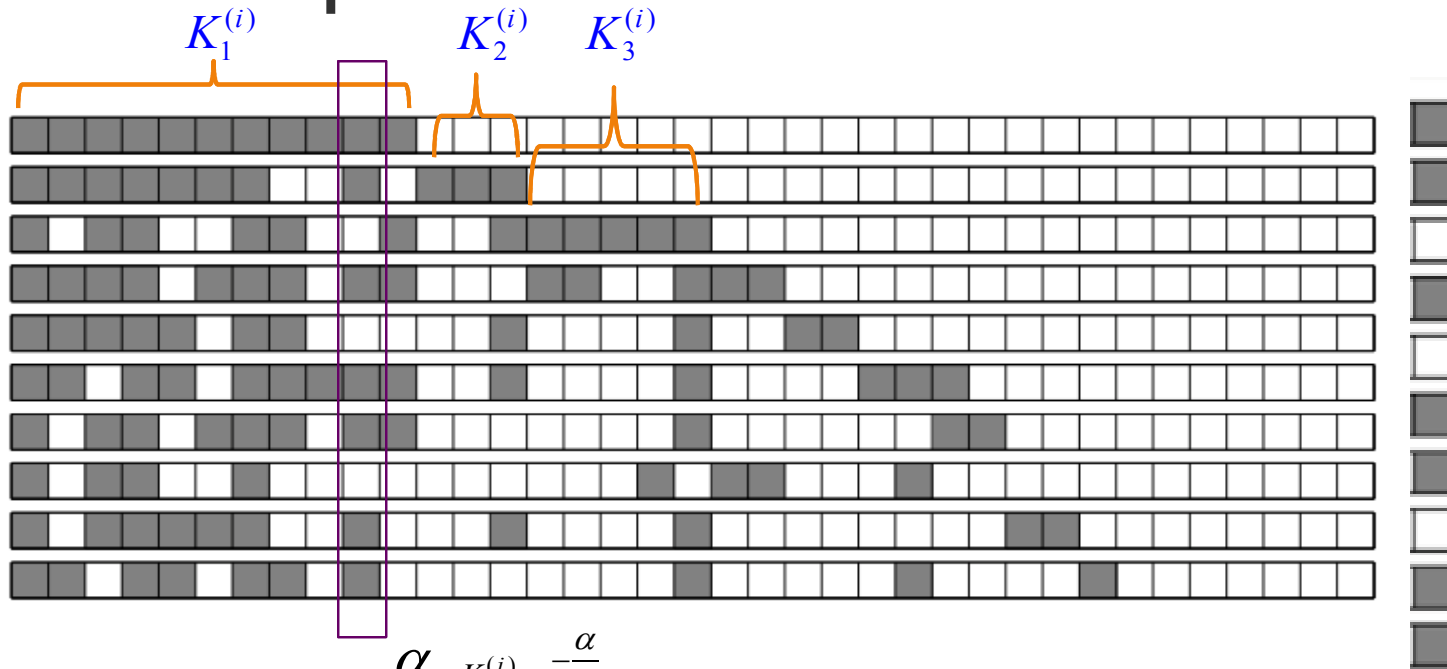
2st Representation: Customers & Dishes



$$P_{new_dish} = \prod_i^N \frac{\left(\frac{\alpha}{i}\right)^{K_1^{(i)}} e^{-\frac{\alpha}{i}}}{K_1^{(i)}!} = \frac{\alpha^{K_+} e^{-\alpha H_N}}{\prod_i^N K_1^{(i)}} \prod_i^N \left(\frac{1}{i}\right)^{K_1^{(i)}}$$

Indian Buffet Process

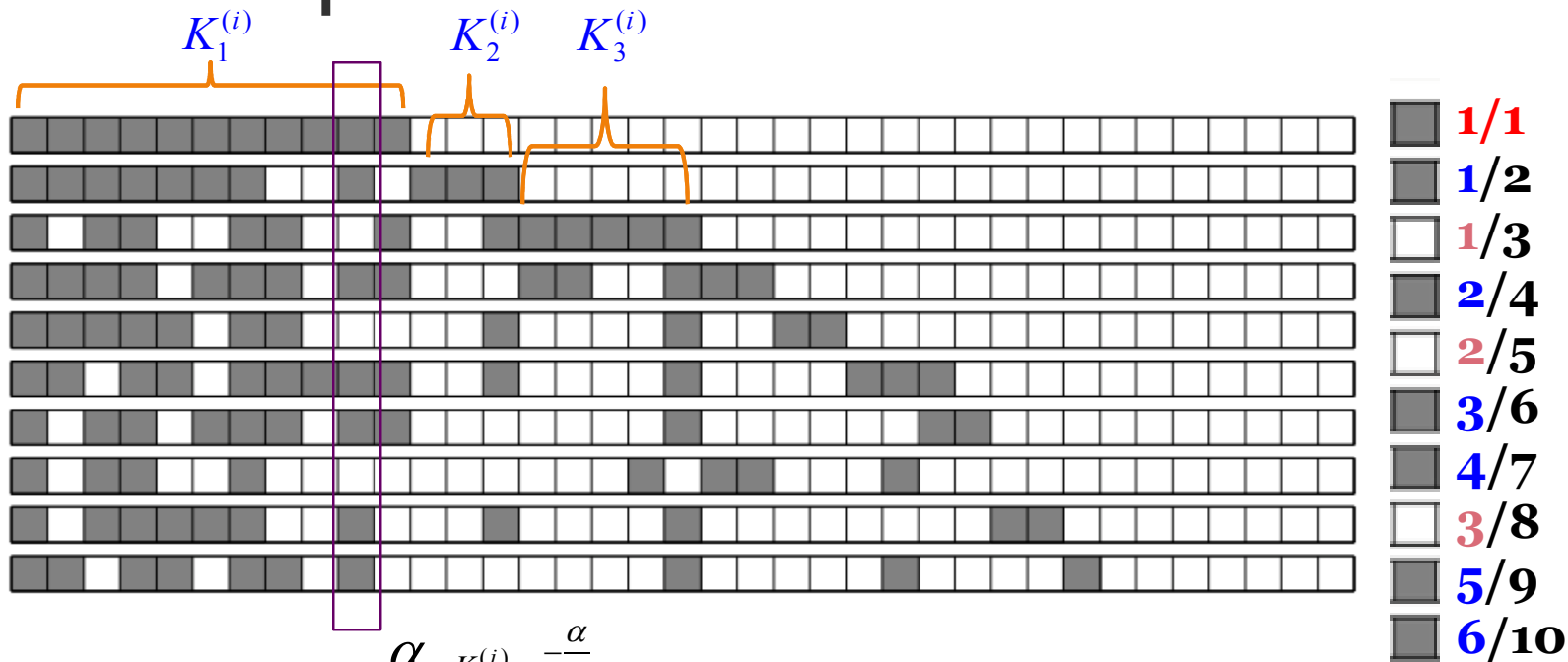
2st Representation: Customers & Dishes



$$P_{new_dish} = \prod_i^N \frac{\left(\frac{\alpha}{i}\right)^{K_1^{(i)}} e^{-\frac{\alpha}{i}}}{K_1^{(i)}!} = \frac{\alpha^{K_+} e^{-\alpha H_N}}{\prod_i^N K_1^{(i)}} \prod_i^N \left(\frac{1}{i}\right)^{K_1^{(i)}}$$

Indian Buffet Process

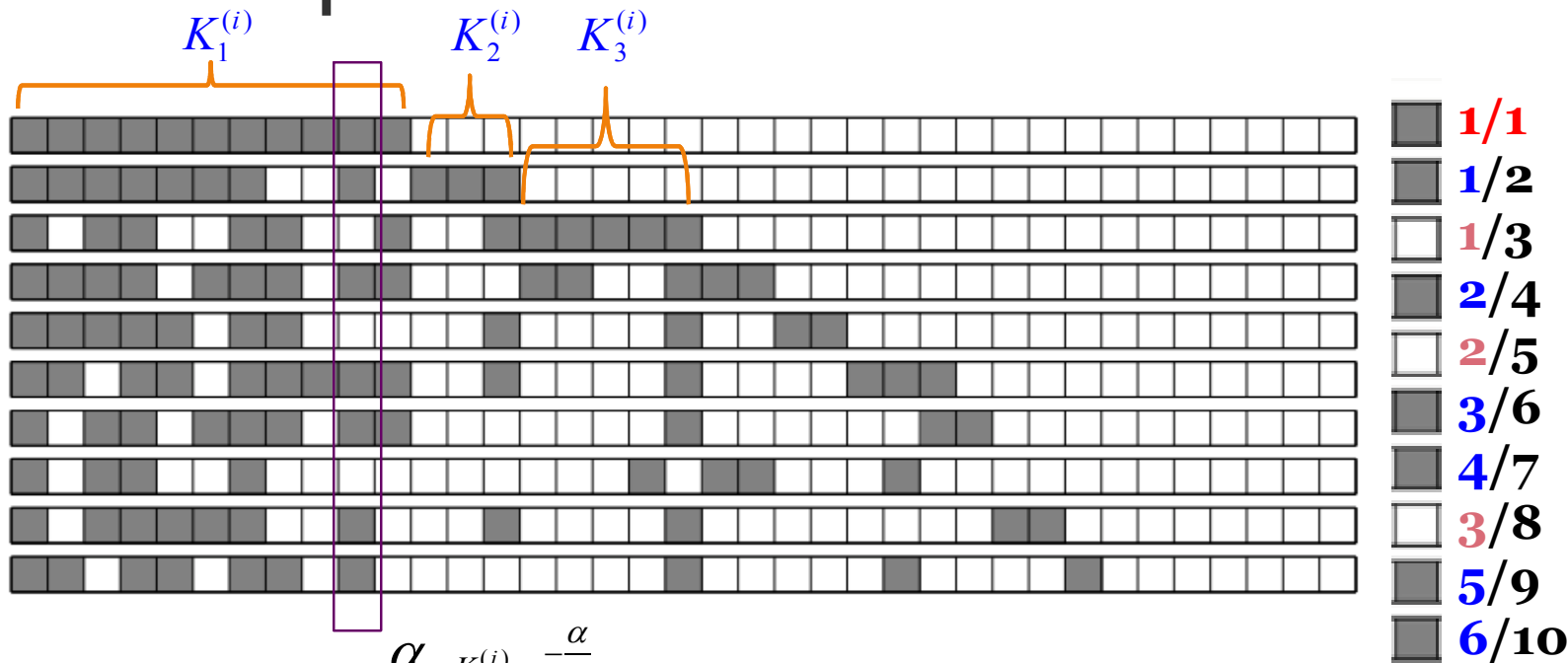
2st Representation: Customers & Dishes



$$P_{new_dish} = \prod_i^N \frac{\left(\frac{\alpha}{i}\right)^{K_1^{(i)}} e^{-\frac{\alpha}{i}}}{K_1^{(i)}!} = \frac{\alpha^{K_+} e^{-\alpha H_N}}{\prod_i^N K_1^{(i)}} \prod_i^N \left(\frac{1}{i}\right)^{K_1^{(i)}}$$

Indian Buffet Process

2st Representation: Customers & Dishes

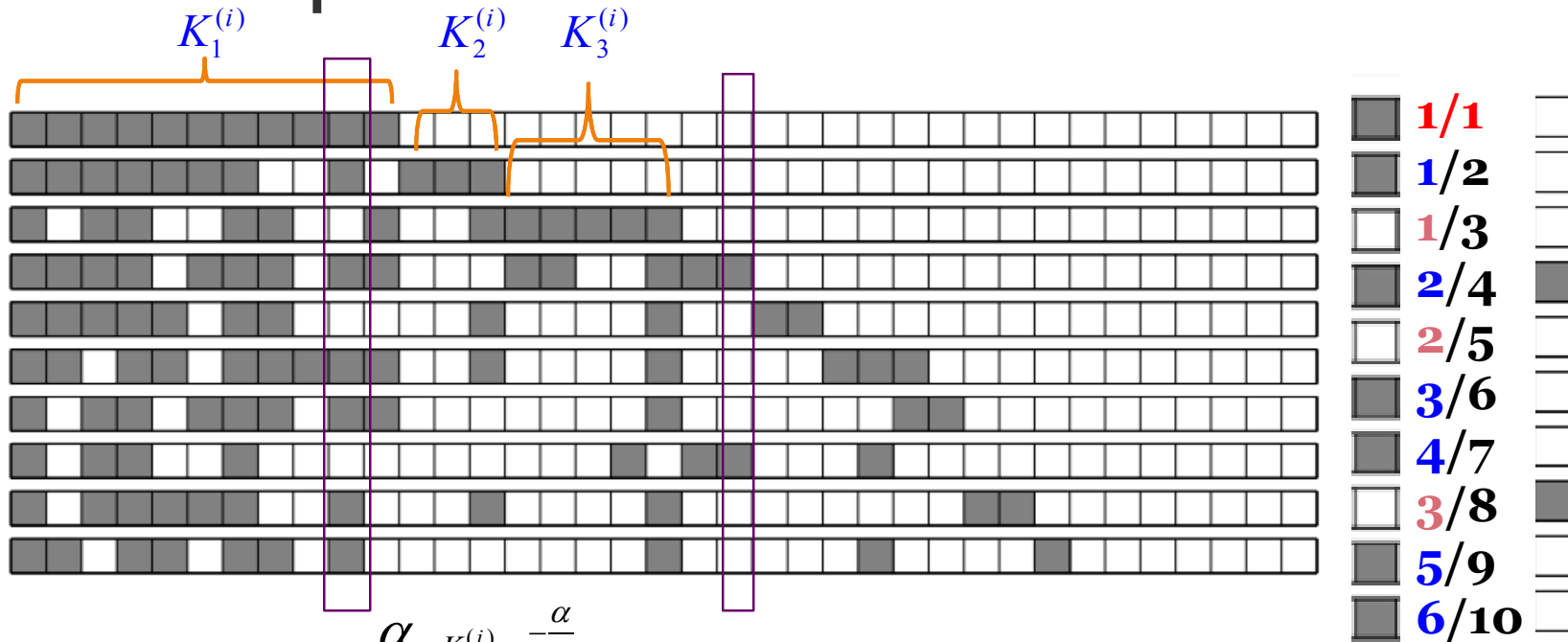


$$P_{new_dish} = \prod_i \frac{\left(\frac{\alpha}{i}\right)^{K_1^{(i)}} e^{-\frac{\alpha}{i}}}{K_1^{(i)}!} = \frac{\alpha^{K_+} e^{-\alpha H_N}}{\prod_i K_1^{(i)}} \prod_i \left(\frac{1}{i}\right)^{K_1^{(i)}}$$

$$\frac{3!6!}{10!}$$

Indian Buffet Process

2st Representation: Customers & Dishes

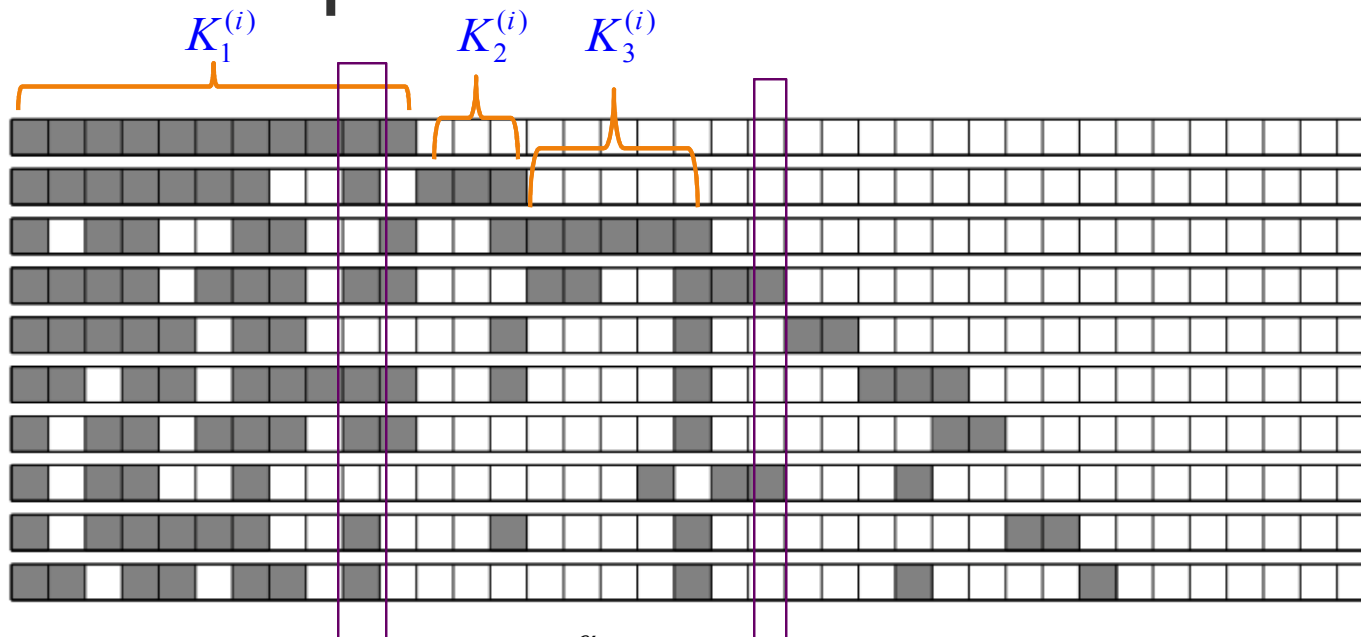


$$P_{new_dish} = \prod_i \frac{\left(\frac{\alpha}{i}\right)^{K_1^{(i)}} e^{-\frac{\alpha}{i}}}{K_1^{(i)}!} = \frac{\alpha^{K_+} e^{-\alpha H_N}}{\prod_i K_1^{(i)}} \prod_i \left(\frac{1}{i}\right)^{K_1^{(i)}}$$

$$\frac{3!6!}{10!}$$

Indian Buffet Process

2st Representation: Customers & Dishes



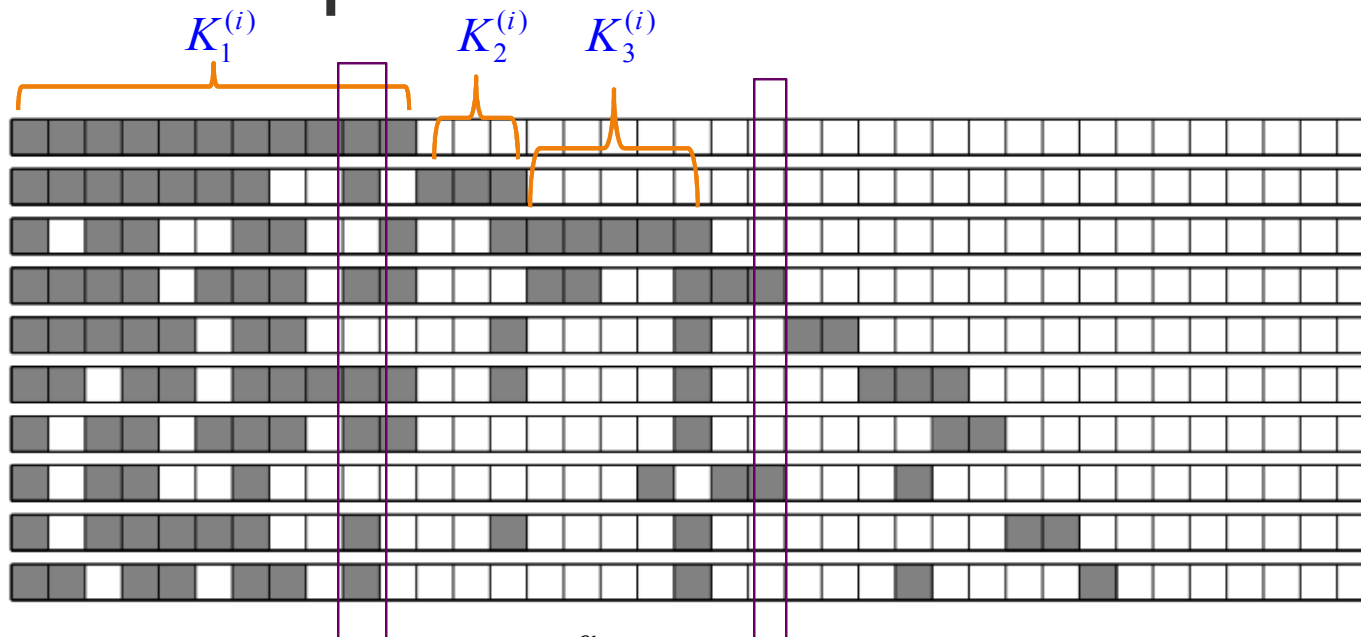
	1/1		1/1
	1/2		2/2
	1/3		3/3
	2/4		1/4
	2/5		4/5
	3/6		5/6
	4/7		6/7
	3/8		1/8
	5/9		7/9
	6/10		8/10

$$\frac{3!6!}{10!}$$

$$P_{new_dish} = \prod_i \frac{\left(\frac{\alpha}{i}\right)^{K_1^{(i)}} e^{-\frac{\alpha}{i}}}{K_1^{(i)}!} = \frac{\alpha^{K_+} e^{-\alpha H_N}}{\prod_i K_1^{(i)}} \prod_i \left(\frac{1}{i}\right)^{K_1^{(i)}}$$

Indian Buffet Process

2st Representation: Customers & Dishes



	1/1		1/1
	1/2		2/2
	1/3		3/3
	2/4		1/4
	2/5		4/5
	3/6		5/6
	4/7		6/7
	3/8		1/8
	5/9		7/9
	6/10		8/10

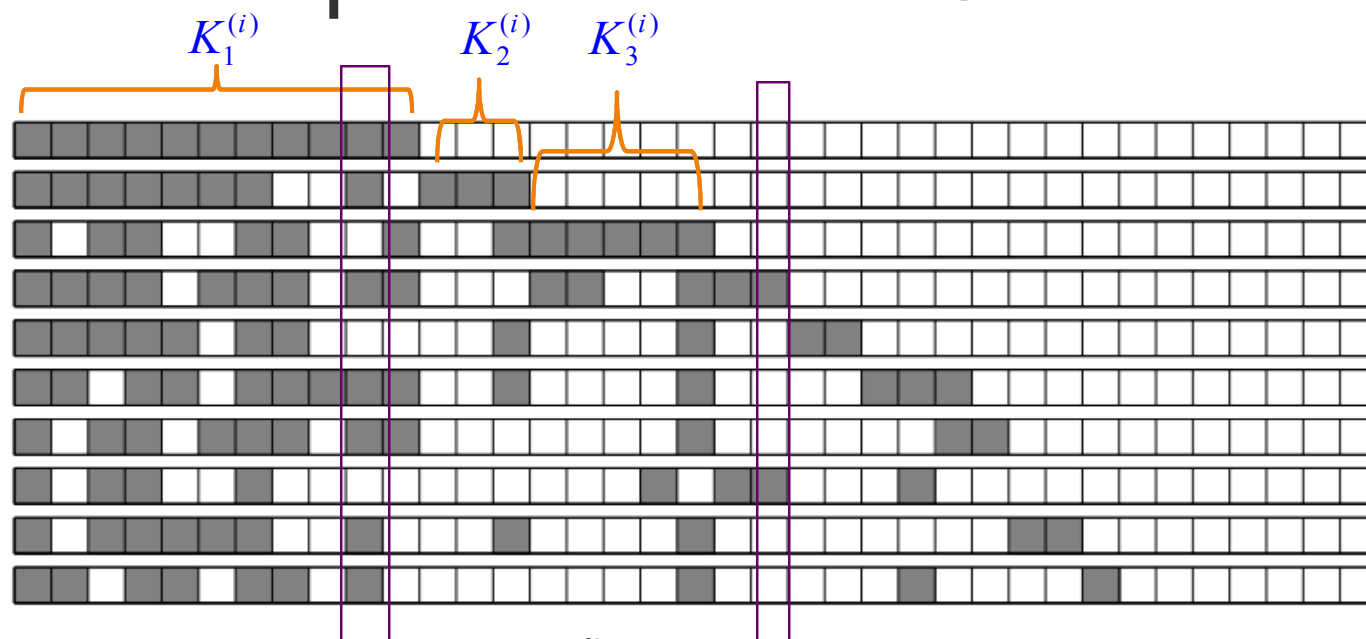
$$\frac{3!6!}{10!}$$

$$\frac{8!1!}{10!}$$

$$P_{new_dish} = \prod_i \frac{\left(\frac{\alpha}{i}\right)^{K_1^{(i)}} e^{-\frac{\alpha}{i}}}{K_1^{(i)}!} = \frac{\alpha^{K_+} e^{-\alpha H_N}}{\prod_i K_1^{(i)}} \prod_i \left(\frac{1}{i}\right)^{K_1^{(i)}}$$

Indian Buffet Process

2st Representation: Customers & Dishes



	1/1		1/1
	1/2		2/2
	1/3		3/3
	2/4		1/4
	2/5		4/5
	3/6		5/6
	4/7		6/7
	3/8		1/8
	5/9		7/9
	6/10		8/10

$$\frac{3!6!}{10!}$$

$$\frac{8!1!}{10!}$$

$$P_{new_dish} = \prod_i \frac{\left(\frac{\alpha}{i}\right)^{K_1^{(i)}} e^{-\frac{\alpha}{i}}}{K_1^{(i)}!} = \frac{\alpha^{K_+} e^{-\alpha H_N}}{\prod_i K_1^{(i)}} \prod_i \left(\frac{1}{i}\right)^{K_1^{(i)}}$$

$$\prod_i \left(\frac{1}{i}\right)^{K_1^{(i)}} P_{old_dishes} = \prod_k \frac{(N - m_k)!(m_k - 1)!}{N!}$$

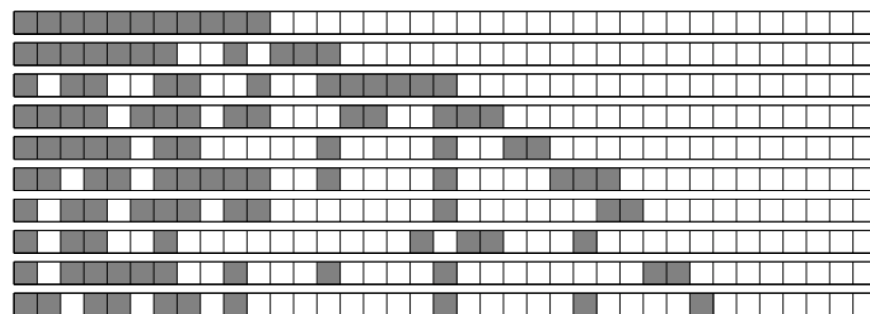
Indian Buffet Process

2st Representation: Customers & Dishes

From:

$$P_{new_dish} = \prod_i^N \frac{(\frac{\alpha}{i})^{K_1^{(i)}} e^{-\frac{\alpha}{i}}}{K_1^{(i)}!} = \frac{\alpha^{K_+} e^{-\alpha H_N}}{\prod_i^N K_1^{(i)}} \prod_i^N (\frac{1}{i})^{K_1^{(i)}}$$

$$\prod_i^N (\frac{1}{i})^{K_1^{(i)}} P_{old_dishes} = \prod_k^{K_+} \frac{(N - m_k)!(m_k - 1)!}{N!}$$



We have:

$$P_{IBP}(Z | \alpha) = \frac{\alpha^{K_+} e^{-\alpha H_N}}{\prod_i^N K_1^{(i)}} \prod_k^{K_+} \frac{(N - m_k)!(m_k - 1)!}{N!}$$

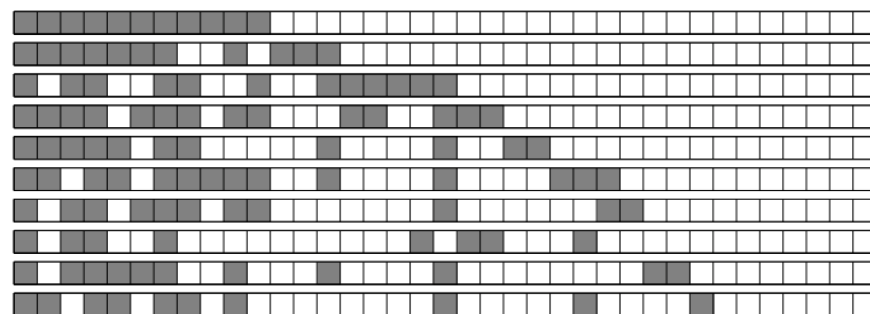
Indian Buffet Process

2st Representation: Customers & Dishes

From:

$$P_{new_dish} = \prod_i^N \frac{(\frac{\alpha}{i})^{K_1^{(i)}} e^{-\frac{\alpha}{i}}}{K_1^{(i)}!} = \frac{\alpha^{K_+} e^{-\alpha H_N}}{\prod_i^N K_1^{(i)}} \prod_i^N (\frac{1}{i})^{K_1^{(i)}}$$

$$\prod_i^N (\frac{1}{i})^{K_1^{(i)}} P_{old_dishes} = \prod_k^{K_+} \frac{(N - m_k)!(m_k - 1)!}{N!}$$



We have:

$$P_{IBP}(Z | \alpha) = \frac{\alpha^{K_+} e^{-\alpha H_N}}{\prod_i^N K_1^{(i)}} \prod_k^{K_+} \frac{(N - m_k)!(m_k - 1)!}{N!}$$

Note:

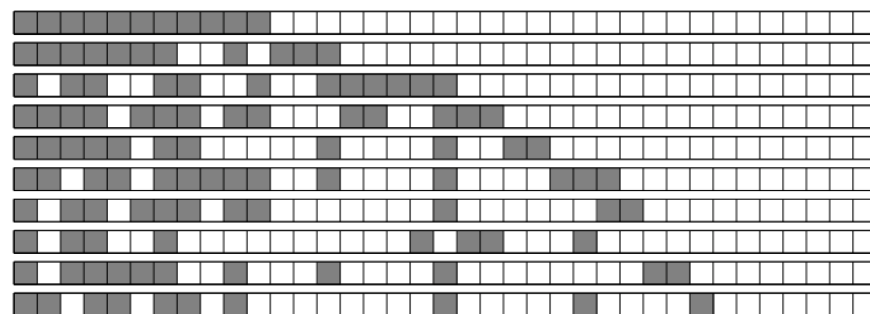
Indian Buffet Process

2st Representation: Customers & Dishes

From:

$$P_{new_dish} = \prod_i^N \frac{(\frac{\alpha}{i})^{K_1^{(i)}} e^{-\frac{\alpha}{i}}}{K_1^{(i)}!} = \frac{\alpha^{K_+} e^{-\alpha H_N}}{\prod_i^N K_1^{(i)}} \prod_i^N (\frac{1}{i})^{K_1^{(i)}}$$

$$\prod_i^N (\frac{1}{i})^{K_1^{(i)}} P_{old_dishes} = \prod_k^{K_+} \frac{(N - m_k)!(m_k - 1)!}{N!}$$



We have:

$$P_{IBP}(Z | \alpha) = \frac{\alpha^{K_+} e^{-\alpha H_N}}{\prod_i^N K_1^{(i)}} \prod_k^{K_+} \frac{(N - m_k)!(m_k - 1)!}{N!}$$

Note:

- Permute $K_1^{(1)}$, next $K_1^{(2)}$, next $K_1^{(3)}$,... next $K_1^{(N)}$ dishes (columns) does not change $P(Z)$

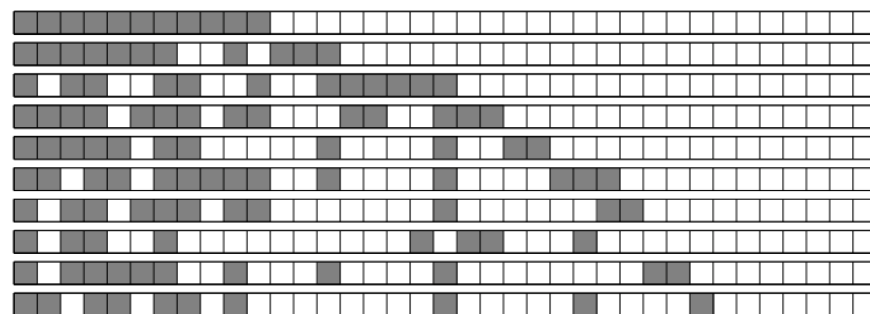
Indian Buffet Process

2st Representation: Customers & Dishes

From:

$$P_{new_dish} = \prod_i^N \frac{(\frac{\alpha}{i})^{K_1^{(i)}} e^{-\frac{\alpha}{i}}}{K_1^{(i)}!} = \frac{\alpha^{K_+} e^{-\alpha H_N}}{\prod_i^N K_1^{(i)}} \prod_i^N (\frac{1}{i})^{K_1^{(i)}}$$

$$\prod_i^N (\frac{1}{i})^{K_1^{(i)}} P_{old_dishes} = \prod_k^{K_+} \frac{(N - m_k)!(m_k - 1)!}{N!}$$



We have:

$$P_{IBP}(Z | \alpha) = \frac{\alpha^{K_+} e^{-\alpha H_N}}{\prod_i^N K_1^{(i)}} \prod_k^{K_+} \frac{(N - m_k)!(m_k - 1)!}{N!}$$

Note:

- Permute $K_1^{(1)}$, next $K_1^{(2)}$, next $K_1^{(3)}$,... next $K_1^{(N)}$ dishes (columns) does not change $P(Z)$
- The permuted matrices are all of same *lof* equivalent class $[Z]$

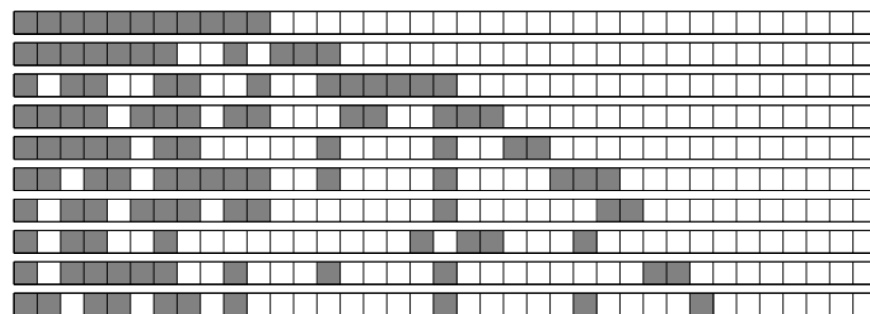
Indian Buffet Process

2st Representation: Customers & Dishes

From:

$$P_{new_dish} = \prod_i^N \frac{(\frac{\alpha}{i})^{K_1^{(i)}} e^{-\frac{\alpha}{i}}}{K_1^{(i)}!} = \frac{\alpha^{K_+} e^{-\alpha H_N}}{\prod_i^N K_1^{(i)}} \prod_i^N (\frac{1}{i})^{K_1^{(i)}}$$

$$\prod_i^N (\frac{1}{i})^{K_1^{(i)}} P_{old_dishes} = \prod_k^{K_+} \frac{(N - m_k)!(m_k - 1)!}{N!}$$



We have:

$$P_{IBP}(Z | \alpha) = \frac{\alpha^{K_+} e^{-\alpha H_N}}{\prod_i^N K_1^{(i)}} \prod_k^{K_+} \frac{(N - m_k)!(m_k - 1)!}{N!}$$

Note:

- Permute $K_1^{(1)}$, next $K_1^{(2)}$, next $K_1^{(3)}$,... next $K_1^{(N)}$ dishes (columns) does not change $P(Z)$
- The permuted matrices are all of same *lof* equivalent class $[Z]$
- Number of permutation:

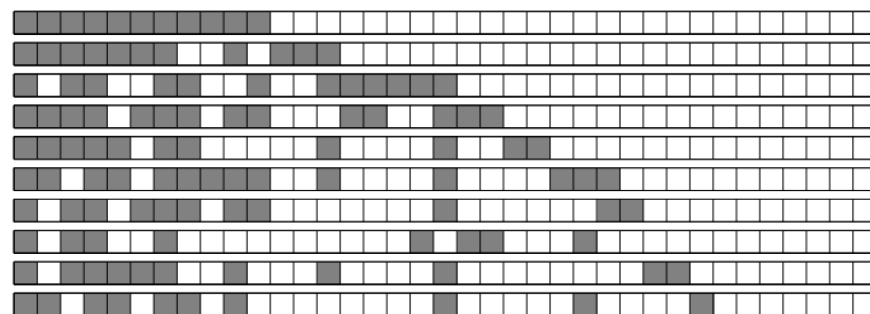
Indian Buffet Process

2st Representation: Customers & Dishes

From:

$$P_{new_dish} = \prod_i^N \frac{(\frac{\alpha}{i})^{K_1^{(i)}} e^{-\frac{\alpha}{i}}}{K_1^{(i)}!} = \frac{\alpha^{K_+} e^{-\alpha H_N}}{\prod_i^N K_1^{(i)}} \prod_i^N (\frac{1}{i})^{K_1^{(i)}}$$

$$\prod_i^N (\frac{1}{i})^{K_1^{(i)}} P_{old_dishes} = \prod_k^{K_+} \frac{(N - m_k)!(m_k - 1)!}{N!}$$



We have:

$$P_{IBP}(Z | \alpha) = \frac{\alpha^{K_+} e^{-\alpha H_N}}{\prod_i^N K_1^{(i)}} \prod_k^{K_+} \frac{(N - m_k)!(m_k - 1)!}{N!}$$

Note:

- Permute $K_1^{(1)}$, next $K_1^{(2)}$, next $K_1^{(3)}$,... next $K_1^{(N)}$ dishes (columns) does not change $P(Z)$
- The permuted matrices are all of same *lof* equivalent class $[Z]$
- Number of permutation: $\frac{\prod_i^N K_1^{(i)}}{2^N - 1}$

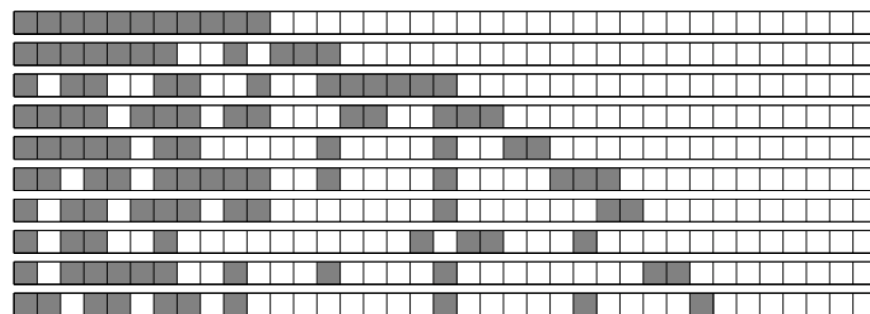
Indian Buffet Process

2st Representation: Customers & Dishes

From:

$$P_{new_dish} = \prod_i^N \frac{(\frac{\alpha}{i})^{K_1^{(i)}} e^{-\frac{\alpha}{i}}}{K_1^{(i)}!} = \frac{\alpha^{K_+} e^{-\alpha H_N}}{\prod_i^N K_1^{(i)}} \prod_i^N (\frac{1}{i})^{K_1^{(i)}}$$

$$\prod_i^N (\frac{1}{i})^{K_1^{(i)}} P_{old_dishes} = \prod_k^{K_+} \frac{(N - m_k)!(m_k - 1)!}{N!}$$



We have:

$$P_{IBP}(Z | \alpha) = \frac{\alpha^{K_+} e^{-\alpha H_N}}{\prod_i^N K_1^{(i)}} \prod_k^{K_+} \frac{(N - m_k)!(m_k - 1)!}{N!}$$

Note:

- Permute $K_1^{(1)}$, next $K_1^{(2)}$, next $K_1^{(3)}$,... next $K_1^{(N)}$ dishes (columns) does not change $P(Z)$
- The permuted matrices are all of same *lof* equivalent class $[Z]$
- Number of permutation: $\frac{\prod_i^N K_1^{(i)}}{2^N - 1}$

$$\frac{\prod_i^N K_1^{(i)}}{2^N - 1} \longrightarrow \prod_{h=1}^{K_+} K_h!$$

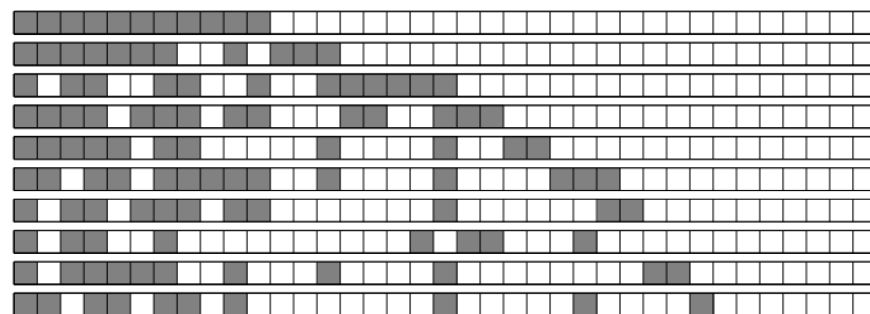
Indian Buffet Process

2st Representation: Customers & Dishes

From:

$$P_{new_dish} = \prod_i^N \frac{(\frac{\alpha}{i})^{K_1^{(i)}} e^{-\frac{\alpha}{i}}}{K_1^{(i)}!} = \frac{\alpha^{K_+} e^{-\alpha H_N}}{\prod_i^N K_1^{(i)}} \prod_i^N (\frac{1}{i})^{K_1^{(i)}}$$

$$\prod_i^N (\frac{1}{i})^{K_1^{(i)}} P_{old_dishes} = \prod_k^{K_+} \frac{(N - m_k)!(m_k - 1)!}{N!}$$



We have:

$$P_{IBP}(Z | \alpha) = \frac{\alpha^{K_+} e^{-\alpha H_N}}{\prod_i^N K_1^{(i)}} \prod_k^{K_+} \frac{(N - m_k)!(m_k - 1)!}{N!}$$

Note:

- Permute $K_1^{(1)}$, next $K_1^{(2)}$, next $K_1^{(3)}$,... next $K_1^{(N)}$ dishes (columns) does not change $P(Z)$
- The permuted matrices are all of same *lof* equivalent class $[Z]$
- Number of permutation: $\frac{\prod_i^N K_1^{(i)}}{2^{N-1} \prod_{h=1}^{K_+} K_h!}$

$$\frac{\prod_i^N K_1^{(i)}}{2^{N-1} \prod_{h=1}^{K_+} K_h!} \longrightarrow P_{IBP}([Z] | \alpha) = \frac{\alpha^{K_+} e^{-\alpha H_N}}{\prod_{h=1}^{K_+} K_h!} \prod_k^{K_+} \frac{(N - m_k)!(m_k - 1)!}{N!}$$

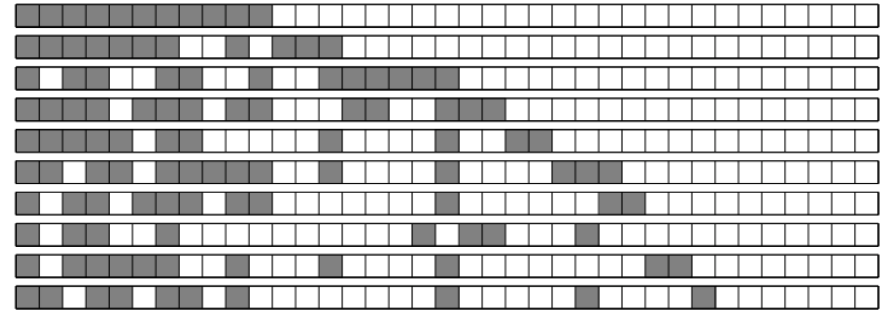
Indian Buffet Process

2st Representation: Customers & Dishes

From:

$$P_{new_dish} = \prod_i^N \frac{(\frac{\alpha}{i})^{K_1^{(i)}} e^{-\frac{\alpha}{i}}}{K_1^{(i)}!} = \frac{\alpha^{K_+} e^{-\alpha H_N}}{\prod_i^N K_1^{(i)}} \prod_i^N (\frac{1}{i})^{K_1^{(i)}}$$

$$\prod_i^N (\frac{1}{i})^{K_1^{(i)}} P_{old_dishes} = \prod_k^{K_+} \frac{(N - m_k)!(m_k - 1)!}{N!}$$



We have:

$$P_{IBP}(Z | \alpha) = \frac{\alpha^{K_+} e^{-\alpha H_N}}{\prod_i^N K_1^{(i)}} \prod_k^{K_+} \frac{(N - m_k)!(m_k - 1)!}{N!}$$

Note:

- Permute $K_1^{(1)}$, next $K_1^{(2)}$, next $K_1^{(3)}$,... next $K_1^{(N)}$ dishes (columns) does not change $P(Z)$
- The permuted matrices are all of same *lof* equivalent class $[Z]$
- Number of permutation: $\frac{\prod_i^N K_1^{(i)}}{2^{N-1} \prod_{h=1}^{K_+} K_h!}$

**2nd representation
is equivalent to
1st representation**

$$\frac{\prod_i^N K_1^{(i)}}{2^{N-1} \prod_{h=1}^{K_+} K_h!} \longrightarrow P_{IBP}([Z] | \alpha) = \frac{\alpha^{K_+} e^{-\alpha H_N}}{\prod_{h=1}^{K_+} K_h!} \prod_k^{K_+} \frac{(N - m_k)!(m_k - 1)!}{N!}$$

Indian Buffet Process

3rd Representation: Distribution over Collections of Histories

- Directly generating the left ordered form (*lof*) matrix \mathbf{Z}
- For each history h :
 - m_h : number of non-zero elements in h
 - Generate K_h columns of history h

$$K_h \sim \text{Poisson}\left(\alpha \frac{(m_h - 1)!(N - m_h)!}{N!}\right)$$

- The distribution over collections of histories

$$\begin{aligned} P(\mathbf{K}) &= \prod_{h=1}^{2^N-1} \frac{\left(\alpha \frac{(m_h-1)!(N-m_h)!}{N!}\right)^{K_h}}{K_h!} \exp\left\{-\alpha \frac{(m_h-1)!(N-m_h)!}{N!}\right\} \\ &= \frac{\alpha^{\sum_{h=1}^{2^N-1} K_h}}{\prod_{h=1}^{2^N-1} K_h!} \exp\{-\alpha H_N\} \prod_{h=1}^{2^N-1} \left(\frac{(m_h-1)!(N-m_h)!}{N!}\right)^{K_h}, \end{aligned}$$

- **Note:**

- **Permute digits in h does not change m_h (nor $P(\mathbf{K})$)**
- **Permute rows means customers are exchangeable**

Indian Buffet Process

- Effective dimension of the model K_+ :
 - Follow Poission distribution: $K_+ \sim \text{Poission}(\alpha H_N)$
 - Derives from 2nd representation by summing Poission components
- Number of features possessed by each object:
 - Follow Possion distribution: $\text{Poission}(\alpha)$
 - Derives from 2nd representation:
 - The first customer chooses $\text{Poission}(\alpha)$ dishes
 - The customers are exchangeable and thus can be purmuted
- \mathbf{Z} is sparse:
 - Non-zero element
 - Derives from the 2nd (or 1st) representation: α
 - Expected number of non-zeros for each row is $N\alpha$
 - Expected entries in \mathbf{Z} is

IBP: Gibbs sampling

- Need to have the full conditional

$$p(z_{ik} = 1 \mid \mathbf{Z}_{-(ik)}, \mathbf{X}) \propto p(\mathbf{X} \mid \mathbf{Z}) p(z_{ik} = 1 \mid \mathbf{Z}_{-(ik)})$$

- $\mathbf{Z}_{-(n,k)}$ denotes the entries of \mathbf{Z} other than z_{nk} .
- $p(\mathbf{X} \mid \mathbf{Z})$ depends on the model chosen for the observed data.
- By exchangeability, consider generating row as the last customer:

- By IBP, in which $p(z_{ik} = 1 \mid \mathbf{z}_{-ik}) = \frac{m_{-i,k}}{N}$

- If sample $z_{ik}=0$, and $m_k=0$: delete the row
- At the end, draw new dishes from $\text{Pois}(\frac{\alpha}{n})$ with considering $p(\mathbf{X} \mid \mathbf{Z})$

- Approximated by truncation, computing probabilities for a range of values of new dishes up to a upper bound

More talk on applications

- Applications

- As prior distribution in models with infinite number of features.
- Modeling Protein Interactions
- Models of bipartite graph consisting of one side with undefined number of elements.
- Binary Matrix Factorization for Modeling Dyadic Data
- Extracting Features from Similarity Judgments
- Latent Features in Link Prediction
- Independent Components Analysis and Sparse Factor Analysis

- More on inference

- Stick-breaking representation (Yee Whye Teh et al., 2007)
- Variational inference (Finale Doshi-Velez et al., 2009)
- Accelerated Inference (Finale Doshi-Velez et al., 2009)

References

- Griffiths, T. L., & Ghahramani, Z. (2011). The indian buffet process: An introduction and review. *Journal of Machine Learning Research*, 12, 1185-1224.
- Slides: Tom Griffiths, “*The Indian buffet process*”.