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: \( \{x_i\}_{i=1}^N \)
Finite Mixture Model

- Set of observation: \( \{x_i\}_{i=1}^N \)
- Constant clusters, \( K \)
Finite Mixture Model

- Set of observation: \( \{x_i\}_{i=1}^{N} \)
- Constant clusters, \( K \)
- Cluster assignment for \( x_i \) is \( c_i \in \{1, \ldots, K\} \)
Finite Mixture Model

- Set of observation: \( \{x_i\}^{N}_{i=1} \)
- Constant clusters, \( K \)
- Cluster assignment for \( x_i \) is \( c_i \in \{1,\ldots,K\} \)
- Cluster assignments vector:

\[ \alpha \]
\[ \theta \]
\[ x_i \]
Finite Mixture Model

- Set of observation: $\{x_i\}_{i=1}^{N}$
- Constant clusters, $K$
- Cluster assignment for $x_i$ is $c_i \in \{1, ..., K\}$
- Cluster assignments vector: $c = [c_1, c_2, ..., c_N]^T$
Finite Mixture Model

- Set of observation: \( \{x_i\}_{i=1}^N \)
- Constant clusters, \( K \)
- Cluster assignment for \( x_i \) is \( c_i \in \{1, \ldots, K\} \)
- Cluster assignments vector: \( c = [c_1, c_2, \ldots, c_N]^T \)
- The probability of each sample under the model:
  \[
p(x_i | \theta) = \sum_{k=1}^K p(x_i | c_i = k) p(c_i = k)
\]
- The likelihood of samples:
  \[
p(X | \theta) = \prod_{i=1}^N \sum_{k=1}^K p(x_i | c_i = k) p(c_i = k)
\]
Finite Mixture Model

- Set of observation: \( \{x_i\}_{i=1}^{N} \)
- Constant clusters, \( K \)
- Cluster assignment for \( x_i \) is \( c_i \in \{1,\ldots,K\} \)
- Cluster assignments vector: \( c = [c_1, c_2, \ldots, c_N]^T \)
- The probability of each sample under the model:
  \[
p(x_i | \theta) = \sum_{k=1}^{K} p(x_i | c_i = k) p(c_i = k)
  \]
- The likelihood of samples:
  \[
p(X | \theta) = \prod_{i=1}^{N} \sum_{k=1}^{K} p(x_i | c_i = k) p(c_i = k)
  \]
- The prior on the component probabilities (symmetric Dirichlet dits.)
  \[
  \theta | \alpha \sim \text{Dirichlet}(\frac{\alpha}{K}, \ldots, \frac{\alpha}{K}).
  \]
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: $c = [c_1, c_2, \ldots, c_N]^T$

• The model can be summarized as:

$$
\begin{align*}
\theta | \alpha &\sim \text{Dirichlet}(\frac{\alpha}{K}, \ldots, \frac{\alpha}{K}). \\
c_i | \theta &\sim \text{Discrete}(\theta)
\end{align*}
$$
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: \( c = [c_1, c_2, ..., c_N]^T \)

• The model can be summarized as:

\[
\begin{align*}
\theta &\mid \alpha \sim \text{Dirichlet}(\frac{\alpha}{K}, ..., \frac{\alpha}{K}). \\
c_i &\mid \theta \sim \text{Discrete}(\theta)
\end{align*}
\]

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

\[
p(\theta \mid c) = \frac{p(c \mid \theta).p(\theta)}{p(c)}
\]
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: \( \mathbf{c} = [c_1, c_2, \ldots, c_N]^T \)
- The model can be summarized as:
  \[
  \begin{align*}
  \theta | \alpha & \sim \text{Dirichlet}(\frac{\alpha}{K}, \ldots, \frac{\alpha}{K}). \\
  c_i | \theta & \sim \text{Discrete}(\theta)
  \end{align*}
  \]
- To have a valid model, all of the distributions must be valid!

\[
p(\theta | \mathbf{c}) = \frac{p(\mathbf{c} | \theta) \cdot p(\theta)}{p(\mathbf{c})}
\]
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: \( c = [c_1, c_2, ..., c_N]^T \)
- The model can be summarized as:

\[
\begin{align*}
\theta | \alpha &\sim \text{Dirichlet}(\frac{\alpha}{K}, ..., \frac{\alpha}{K}). \\
c_i | \theta &\sim \text{Discrete}(\theta)
\end{align*}
\]
- To have a valid model, all of the distributions must be valid!

\[
p(\theta | c) = \frac{p(c | \theta)p(\theta)}{p(c)}
\]
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: \( \mathbf{c} = [c_1, c_2, \ldots, c_N]^T \)

• The model can be summarized as:

\[
\begin{cases}
\theta | \alpha \sim \text{Dirichlet}\left(\frac{\alpha}{K}, \ldots, \frac{\alpha}{K}\right). \\
c_i | \theta \sim \text{Discrete} (\theta)
\end{cases}
\]

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

\[
p(\theta | \mathbf{c}) = \frac{p(\mathbf{c} | \theta) p(\theta)}{p(\mathbf{c})}
\]

\( p(\mathbf{c}) \)
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: \( c = [c_1, c_2, ..., c_N]^T \)

• The model can be summarized as:

\[
\begin{align*}
    \theta | \alpha & \sim \text{Dirichlet}(\frac{\alpha}{K}, ..., \frac{\alpha}{K}). \\
    c_i | \theta & \sim \text{Discrete}(\theta)
\end{align*}
\]

• To have a valid model, all of the distributions must be valid!
Finite Mixture Model contd.

\[
p(c) = \int_{\Delta_k} p(c \mid \theta) p(\theta) d\theta \quad p(\theta) = \left( D\left(\frac{\alpha}{K}, \ldots, \frac{\alpha}{K}\right)^{-1} \prod_{k=1}^{K} \theta_k^{\alpha_k-1}\right.
\]

\[\alpha \quad \theta \quad c_i \quad N\]
Finite Mixture Model contd.

\[
p(c) = \int_{\Delta_k} p(c | \theta) p(\theta) d\theta \quad p(\theta) = \left(D\left(\frac{\alpha}{K}, \ldots, \frac{\alpha}{K}\right)\right)^{-1} \prod_{k=1}^{K} \theta_k^{\alpha_k - 1}
\]

\[
m_k = \sum_{i=1}^{N} \delta(c_i = k)
\]
Finite Mixture Model contd.

\[ p(c) = \int_{\Delta_K} p(c \mid \theta) p(\theta) d\theta \quad p(\theta) = \left( \frac{\alpha}{K}, \ldots, \frac{\alpha}{K} \right)^{-1} \prod_{k=1}^{K} \theta_k^{\alpha_k-1} \]

\[ m_k \equiv \sum_{i=1}^{N} \delta(c_i = k) \quad \Rightarrow \quad p(c \mid \theta) = \prod_{k=1}^{K} \theta_k^{m_k} \]
Finite Mixture Model contd.

\[ p(c) = \int_{\Delta_k} p(c | \theta) p(\theta) d\theta \]

\[ p(\theta) = \left( D\left( \frac{\alpha}{K}, ..., \frac{\alpha}{K} \right) \right)^{-1} \prod_{k=1}^{K} \theta_k^{\alpha_k-1} \]

\[ m_k \propto \sum_{i=1}^{N} \delta(c_i = k) \Rightarrow p(c | \theta) = \prod_{k=1}^{K} \theta_k^{m_k} \]

\[ p(c) = \int_{\Delta_k} p(c | \theta) p(\theta) d\theta \]

\[ = \int_{\Delta_k} \frac{1}{D\left( \frac{\alpha}{K}, ..., \frac{\alpha}{K} \right)} \prod_{k=1}^{K} \theta_k^{m_k + \frac{\alpha}{K} - 1} \ d\theta \]
Finite Mixture Model contd.

\[
p(c) = \int_{\Delta K} p(c | \theta) p(\theta) d\theta \quad p(\theta) = \left( D\left(\frac{\alpha}{K}, \ldots, \frac{\alpha}{K}\right) \right)^{-1} \prod_{k=1}^{K} \theta_{k}^{\alpha_{k}-1} \]

\[
m_{k} = \sum_{i=1}^{N} \delta(c_{i} = k) \quad \Rightarrow p(c | \theta) = \prod_{k=1}^{K} \theta_{k}^{m_{k}} \]

\[
p(c) = \int_{\Delta K} p(c | \theta) p(\theta) d\theta \]

\[
= \int_{\Delta K} \frac{1}{D\left(\frac{\alpha}{K}, \ldots, \frac{\alpha}{K}\right)} \prod_{k=1}^{K} \theta_{k}^{m_{k}+\frac{\alpha}{K}-1} d\theta \]

\[
= \prod_{k=1}^{K} \frac{\Gamma\left(m_{k} + \frac{\alpha}{K}\right)}{\Gamma\left(\frac{\alpha}{K}\right)} \frac{\Gamma(\alpha)}{\Gamma(N + \alpha)}, \quad \text{s.t. } m_{k} = \sum_{i=1}^{N} \delta(c_{i} = k) \]
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: $K \to \infty$
- 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:
    - $K_+ \text{ number of classes for which } m_k > 0$
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:
    $K_+ \text{ number of classes for which } m_k > 0$
    $K_0 \text{ number of classes for which } m_k = 0$
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:
    - $K_+ \equiv \text{number of classes for which } m_k > 0$
    - $K_0 \equiv \text{number of classes for which } m_k = 0$
  - 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: $K \to \infty$
- Since we always have limited samples in reality, we will have limited number of clusters used; so we define two set of clusters:
  - $K_+ \rightarrow$ number of classes for which $m_k > 0$
  - $K_0 \rightarrow$ number of classes for which $m_k = 0$
- Assume a reordering, such that $\forall k > K_+ \Rightarrow m_k = 0$; and $\forall k \leq K_+ \Rightarrow m_k > 0$

![Diagram showing the division of clusters into $K_+$ and $K_0$]
Infinite mixture model

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

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

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

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

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

$$\Gamma(x+1) = x\Gamma(x) \Rightarrow \Gamma(m_k + \frac{\alpha}{K}) = (m_k + \frac{\alpha}{K} - 1)...\left( \frac{\alpha}{K} \right)\Gamma\left( \frac{\alpha}{K} \right)$$
Infinite mixture model

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

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

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

- Now we return to the previous slides and set \( K \to \infty \) in formulas

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

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

\[
\Rightarrow p(c) = \frac{\prod_{k=1}^{K} \Gamma \left( m_k + \frac{\alpha}{K} \right)}{\left( \frac{\alpha}{K} \right)^{K}} \frac{\Gamma(\alpha)}{\Gamma(N+\alpha)} = \left( \frac{\alpha}{K} \right)^{K} \prod_{k=1}^{K} \prod_{j=1}^{m_k} \left( j + \frac{\alpha}{K} \right) \frac{\Gamma(\alpha)}{\Gamma(N+\alpha)}
\]
Infinite mixture model

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

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

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

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

If we set $K \to \infty$ the marginal likelihood will be $p(c) \to 0$. 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: $[c] = \{c_i \mid c_i \in c\}$
Infinite mixture model contd.

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

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

- Define a partition of objects;
- Want to partition \( N \) objects into \( K^+ \) classes
- Equivalence class of object partitions: \([c] = \{c_i | c_i \in c\}\)

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

\[
\Rightarrow p([c]) = \alpha^{K^+} \cdot \frac{K!}{K_0!} \frac{1}{K^+} \prod_{k=1}^{K^+} \prod_{j=1}^{m_k-1} \left( j + \frac{\alpha}{K} \right) \cdot \frac{\Gamma(\alpha)}{\Gamma(N + \alpha)}
\]
Infinite mixture model contd.

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

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

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

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

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

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

$$\Rightarrow \lim_{K \to \infty} p([c]) = \alpha^{K_+} \cdot 1 \cdot \prod_{k=1}^{K_+} (m_k - 1)! \cdot \frac{\Gamma(\alpha)}{\Gamma(N + \alpha)}$$
Infinite mixture model contd.

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

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

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

$$\Rightarrow \lim_{K \to \infty} p([c]) = \alpha^{K_+}. \frac{\Gamma(\alpha)}{\Gamma(N + \alpha)}$$

- 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: $[c] = \{ c_i \mid c_i \in c \}$

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

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

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

- 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: \( [\mathbf{c}] = \{ \mathbf{c}_i \mid \mathbf{c}_i \in \mathbf{c} \} \)

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

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

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

- 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: $[c] = \{ c_i \mid c_i \in c \}$

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

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

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

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

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

Parameter \( \alpha = 1 \)

\[ m_1 = 0 \]

\[ p(c_1 = 1) = \frac{1}{1} \]
Chinese Restaurant Process (CRP)

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

Parameter \( \alpha = 1 \)

\( m_1 = 1, \quad m_2 = 0 \)

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

\[ p(c_2 = 2 \mid c_1) = \frac{1}{1+1} \]
Chinese Restaurant Process (CRP)

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

Parameter \( \alpha = 1 \)

\( m_1 = 1, \quad m_2 = 1, \quad m_3 = 0 \)

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

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

\[ p(c_3 = 3 \mid c_{1:2}) = \frac{1}{2+1} \]
Chinese Restaurant Process (CRP)

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

Parameter \( \alpha = 1 \)

\( m_1 = 2, \quad m_2 = 1, \quad m_3 = 0 \)

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

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

\[ p(c_4 = 2 \mid c_{1:3}) = \frac{1}{3 + 1} \]
Chinese Restaurant Process (CRP)

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

Parameter \( \alpha = 1 \)

\( m_1 = 2, \quad m_2 = 2, \quad m_3 = 0 \)

\[ p(c_5 = 1 \mid c_{1:4}) = \frac{2}{4+1} \]
\[ p(c_5 = 2 \mid c_{1:4}) = \frac{2}{4+1} \]
\[ p(c_5 = 3 \mid c_{1:4}) = \frac{1}{4+1} \]
Chinese Restaurant Process (CRP)

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

Parameter \( \alpha = 1 \)

- \( m_1 = 2, \quad m_2 = 2, \quad m_3 = 1, \quad m_4 = 0 \)

- \( p(c_6 = 1 \mid c_{1:5}) = \frac{2}{5+1} \)
- \( p(c_6 = 3 \mid c_{1:5}) = \frac{1}{5+1} \)
- \( p(c_6 = 2 \mid c_{1:5}) = \frac{2}{5+1} \)
- \( p(c_6 = 4 \mid c_{1:5}) = \frac{1}{5+1} \)
Chinese Restaurant Process (CRP)

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

Parameter \( \alpha = 1 \)

\( m_1 = 2, \quad m_2 = 3, \quad m_3 = 1, \quad m_4 = 0 \)

\[
p(c_7 = 1 \mid c_{1:6}) = \frac{2}{6+1}
\]

\[
p(c_7 = 3 \mid c_{1:6}) = \frac{1}{6+1}
\]

\[
p(c_7 = 2 \mid c_{1:6}) = \frac{3}{6+1}
\]

\[
p(c_7 = 4 \mid c_{1:6}) = \frac{1}{6+1}
\]
Chinese Restaurant Process (CRP)

\[ p(c_i = k \mid c_1, \ldots, c_{i-1}) = \begin{cases} \frac{m_k}{i-1+\alpha} & k \leq K^+ \\ \alpha & k = K + 1 \end{cases} \]

Parameter \( \alpha = 1 \)

\[ m_1 = 2, \quad m_2 = 3, \quad m_3 = 2, \quad m_4 = 0 \]

\[ p(c_8 = 1 \mid c_{1:3}) = \frac{2}{7+1} \]

\[ p(c_8 = 2 \mid c_{1:3}) = \frac{3}{7+1} \]

\[ p(c_8 = 3 \mid c_{1:3}) = \frac{2}{7+1} \]

\[ p(c_8 = 4 \mid c_{1:3}) = \frac{1}{7+1} \]
Chinese Restaurant Process (CRP)

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

Parameter \( \alpha = 1 \)

\[ m_1 = 3, \quad m_2 = 3, \quad m_3 = 2, \quad m_4 = 0 \]

\[ p(c_9 = 1 \mid c_{1:3}) = \frac{3}{8+1} \]
\[ p(c_9 = 3 \mid c_{1:3}) = \frac{2}{8+1} \]
\[ p(c_9 = 2 \mid c_{1:3}) = \frac{3}{8+1} \]
\[ p(c_9 = 4 \mid c_{1:3}) = \frac{1}{8+1} \]
Chinese Restaurant Process (CRP)

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

Parameter \( \alpha = 1 \)

\[m_1 = 3, \quad m_2 = 3, \quad m_3 = 2, \quad m_4 = 1, \quad m_5 = 5\]

\[
p(c_{10} = 1 \mid c_{1:4}) = \frac{3}{9+1} \quad p(c_{10} = 3 \mid c_{1:4}) = \frac{2}{9+1} \quad p(c_{10} = 5 \mid c_{1:4}) = \frac{1}{9+1}
\]

\[
p(c_{10} = 2 \mid c_{1:4}) = \frac{3}{9+1} \quad p(c_{10} = 4 \mid c_{1:4}) = \frac{1}{9+1}
\]
CRP: Gibbs sampling

• Gibbs sampler requires full conditional

\[ p(c_i = k \mid c_{-i}, X) \propto p(X \mid c).p(c_i = k \mid c_{-i}) \]

• Finite Mixture Model:

\[ p(c_i = k \mid c_{-i}) = \frac{m_{-i,k} + \alpha}{K} \]

\[ \frac{N - 1 + \alpha}{K} \]

• Infinite Mixture Model:

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

- \( F \): latent feature matrix
- \( Z \): binary matrix
- \( V \): value matrix

\[
F = Z \otimes V
\]

- With \( p(F) = p(Z) \cdot p(V) \)
Finite Feature Model

• Generating $Z$: $(N \times K)$ binary matrix
  - For each column $k$, draw $\pi_k$ from beta distribution
  - For each object, flip a coin by $z_{ik}$

\[
\begin{align*}
\pi_k & \mid \alpha \sim \text{Beta}\left(\frac{\alpha}{K}, 1\right) & (\pi_k, 1 - \pi_k) & \sim \text{Dir}\left(\frac{\alpha}{K}, 1\right) \\
\end{align*}
\]

\[
\begin{align*}
\pi_k & \mid \alpha \sim \text{Beta}\left(\frac{\alpha}{K}, 1\right) & (\pi_k, 1 - \pi_k) & \sim \text{Dir}\left(\frac{\alpha}{K}, 1\right) \\
\end{align*}
\]
Finite Feature Model

• Generating $Z$ : (N*K) binary matrix
  ▫ For each column $k$, draw $\pi_k$ from beta distribution
  ▫ For each object, flip a coin by $Z_{ik}$

\[
\begin{align*}
\pi_k & \sim \text{Beta}(\frac{\alpha}{K},1) \\
Z_{ik} & \sim \text{Bernoulli}(\pi_k)
\end{align*}
\]

• Distribution of $Z$ :
**Finite Feature Model**

- **Generating $Z$:** $(N \times K)$ binary matrix
  - For each column $k$, draw $\pi_k$ from beta distribution
  - For each object, flip a coin by $z_{ik}$

  \[
  \begin{align*}
  \pi_k \mid \alpha &\sim \text{Beta}\left(\frac{\alpha}{K}, 1\right) \\
  z_{ik} \mid \pi_k &\sim \text{Bernoulli}(\pi_k)
  \end{align*}
  \]

- Distribution of $Z$:

  \[
  p(Z \mid \pi) = \prod_{k=1}^{K} \prod_{i=1}^{N} p(z_{ik} \mid \pi_k) = \prod_{k=1}^{K} \pi_k^{m_k} (1 - \pi_k)^{N-m_k}
  \]
Finite Feature Model

• Generating \( Z \): \((N \times K)\) binary matrix
  - For each column \( k \), draw \( \pi_k \) from beta distribution
  - For each object, flip a coin by \( z_{ik} \)
    \[
    \begin{aligned}
    \pi_k | \alpha &\sim \text{Beta}\left(\frac{\alpha}{K},1\right) \\
    z_{ik} | \pi_k &\sim \text{Bernoulli}(\pi_k)
    \end{aligned}
    \]

• Distribution of \( Z \):
  \[
  p(Z | \pi) = \prod_{k=1}^{K} \prod_{i=1}^{N} p(z_{ik} | \pi_k) = \prod_{k=1}^{K} \pi_k^{m_k} (1 - \pi_k)^{N - m_k}
  \]
  \[
  p(Z | \alpha) = \prod_{k} \int_{\pi_k} p(\pi_k | \alpha) p(z_{ik} | \pi_k) d\pi_k
  \]
Finite Feature Model

- **Generating $Z$ :** $(N \times K)$ binary matrix
  - For each column $k$, draw $\pi_k$ from beta distribution
  - For each object, flip a coin by $z_{ik}$

\[
\begin{align*}
\pi_k | \alpha & \sim \text{Beta}(\frac{\alpha}{K}, 1) \\
(z_{ik} | \pi_k) & \sim \text{Bernoulli}((\pi_k, 1 - \pi_k)) \\
\end{align*}
\]

- **Distribution of $Z$ :**

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

\[
p(Z | \alpha) = \prod_{k} \int_{\pi_k} p(\pi_k | \alpha) p(z_{ik} | \pi_k) d\pi_k
\]

\[
= \prod_{k=1}^{K} \frac{\alpha K}{\Gamma(N + 1 + \frac{\alpha}{K})} \frac{\Gamma(m_k + \frac{\alpha}{K})}{\Gamma(N - m_k + 1)}
\]
Finite Feature Model

- Generating $\mathbf{Z}$: $(N \times K)$ binary matrix
  - For each column $k$, draw $\pi_k$ from beta distribution
  - For each object, flip a coin by $Z_{ik}$

$$
\begin{align*}
\pi_k | \alpha &\sim \text{Beta}(\frac{\alpha}{K}, 1) \quad \leftrightarrow \quad (\pi_k, 1 - \pi_k) \sim \text{Dir}(\frac{\alpha}{K}, 1) \\
Z_{ik} | \pi_k &\sim \text{Bernoulli}(\pi_k)
\end{align*}
$$

- $\mathbf{Z}$ is sparse:

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

  - Even $K \to \infty$
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, \( \text{lof}(Z) \):
  - 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, $lof(Z)$:
  - 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, \( \text{lof}(Z) \):
  - 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, $lof(Z)$:
  - Compute history $h$ of feature (column) $k$
Indian Buffet Process

1st Representation: $K \to \infty$

- Difficulty:
  - $P(Z) \to 0$
  - Solution: define equivalence classes on random binary feature matrices.

- *left-ordered form* function of binary matrices, $lof(Z)$:
  - Compute history $h$ of feature (column) $k$

\[
h_4 = 2^3 + 2^7 + 2^9
\]
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

\[
h_4 = 2^3 + 2^7 + 2^9
\]
Indian Buffet Process

1\textsuperscript{st} Representation: $K \rightarrow \infty$

- Difficulty:
  - $P(Z) \rightarrow 0$
  - Solution: define equivalence classes on random binary feature matrices.

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

$$h_4 = 2^3 + 2^7 + 2^9$$
Indian Buffet Process

1st Representation: $K \rightarrow \infty$

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

\[ h_4 = 2^3 + 2^7 + 2^9 \]

$Z_1$ and $Z_2$ are \textit{lof} equivalent iff $lof(Z_1) = lof(Z_2) = [Z]$
Indian Buffet Process

1\textsuperscript{st} Representation: \( K \to \infty \)

- Difficulty:
  - \( P(Z) \to 0 \)
  - Solution: define equivalence classes on random binary feature matrices.

- \textit{left-ordered form} function of binary matrices, \( \text{lof}(Z) \):
  - 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 \textit{lof} equivalent iff \( \text{lof}(Z_1)=\text{lof}(Z_2) = [Z] \)

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:
  - $P(Z) \rightarrow 0$
  - Solution: define equivalence classes on random binary feature matrices.
- *left-ordered form* function of binary matrices, $\text{lof}(Z)$:
  - 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 $\text{lof}(Z_1) = \text{lof}(Z_2) = [Z]$  

Describing $[Z]$:
\[
K_h = \#\{i; h_i = h\}
\]
\[
K_+ = \sum_{i>0} K_h
\]
\[
K = K_+ + K_0
\]

Cardinality of $[Z]$:
\[
\binom{K}{K_0 \ldots K_{2^N-1}} = \frac{K!}{\prod_{h=0}^{2^N-1} K_h!}
\]
Indian Buffet Process

1st Representation: $K \to \infty$

Given:

$$\Pr(Z | \alpha) = \prod_{k=1}^{K} \frac{\alpha}{K} \frac{\Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})}$$

$$\text{card}([Z]) = \binom{K}{K_0 \ldots K_{2^{N-1}}} = \frac{K!}{\prod_{h=0}^{2^{N-1}} K_h !}$$

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

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

1\textsuperscript{st} Representation: \( K \to \infty \)

Given:

\[
\Pr(Z | \alpha) = \prod_{k=1}^{K} \frac{\alpha}{K} \frac{\Gamma(m_k + \frac{\alpha}{K})}{\Gamma(N - m_k + 1)} \frac{\Gamma(N + 1 + \frac{\alpha}{K})}{\prod_{h=0}^{2^N-1} K_h !} \\
\text{card}([Z]) = \binom{K}{K_0 \ldots K_{2^N-1}} = \frac{K!}{\prod_{h=0}^{2^N-1} K_h !}
\]

Derive when \( K \to \infty \) : \( (K_+^+ < \infty \text{ almost surely}) \)

\[
\Pr([Z] | \alpha) = \Pr(Z | \alpha) \cdot \text{card}([Z]) \\
= \frac{K!}{\prod_{h=0}^{2^N-1} K_h !} \prod_{k=1}^{K} \frac{\alpha}{K} \frac{\Gamma(m_k + \frac{\alpha}{K})}{\Gamma(N - m_k + 1)} \frac{\Gamma(N + 1 + \frac{\alpha}{K})}{\prod_{h=0}^{2^N-1} K_h !} \\
\]
Indian Buffet Process

1st Representation: $K \to \infty$

Given:

$$\Pr(Z | \alpha) = \prod_{k=1}^{K} \frac{\alpha}{K} \frac{\Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})}$$

$$\text{card}([Z]) = \left\{ \begin{array}{l} \frac{K}{K_0 ... K_{2^{N-1}}} \\ K_0 \end{array} \right\} = \frac{K!}{\prod_{h=0}^{2^{N-1}} K_h!}$$

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

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

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

$$= \frac{K!}{2^{N-1} \prod_{h=0}^{2^{N-1}} K_h!} \left( \frac{\alpha}{K} \frac{\Gamma(\frac{\alpha}{K}) \Gamma(N + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \right)^{K_0} \prod_{k=1}^{K} \frac{\alpha}{K} \frac{\Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})}$$
Indian Buffet Process

1st Representation: \( K \rightarrow \infty \)

Given:

\[
\Pr(Z \mid \alpha) = \prod_{k=1}^{K} \frac{\alpha}{K} \frac{\Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \\
\text{card}([Z]) = \binom{K}{K_0 \ldots K_{2^N - 1}} = \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} 
\]

Derive when \( K \rightarrow \infty \) : \( (K_+ < \infty \text{ almost surely}) \)

\[
\Pr([Z] \mid \alpha) = \Pr(Z \mid \alpha) \cdot \text{card}([Z]) \\
= \frac{K!}{2^{N-1}} \prod_{h=0}^{2^N-1} K_h! \left( \frac{\alpha}{K} \frac{\Gamma(\frac{\alpha}{K}) \Gamma(N + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \right)^{K_0} \prod_{k=1}^{K} \frac{\alpha}{K} \frac{\Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} 
\]
Indian Buffet Process

1st Representation: \( K \rightarrow \infty \)

Given:

\[
\Pr(Z | \alpha) = \prod_{k=1}^{K} \frac{\alpha}{K} \frac{\Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})}
\]

\[
card([Z]) = \left\{ \begin{array}{c} K \\ K_0 \ldots K_{2^{N-1}} \end{array} \right\} = \frac{K!}{\prod_{h=0}^{2^{N-1}} K_h!}.
\]

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

\[
\Pr([Z] | \alpha) = \Pr(Z | \alpha) \cdot card([Z])
\]

\[
= \frac{K!}{2^{N-1}} \left( \prod_{h=0}^{2^{N-1}} K_h! \right) \left( \frac{\alpha}{K} \frac{\Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \right)^{K_0} \left( \prod_{k=1}^{K} \frac{\alpha}{K} \frac{\Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \right)^{K_+}
\]

\[
= \frac{K!}{2^{N-1}} \left( \prod_{h=0}^{2^{N-1}} K_h! \right) \left( \frac{\alpha}{K} \frac{\Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \right)^{K_0} \left( \prod_{k=1}^{K} \frac{\alpha}{K} \frac{\Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \right)^{K_+}
\]

\[
\Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)
\]

\[
\Gamma(N + 1 + \frac{\alpha}{K})
\]

\[
\Gamma(N + 1 + \frac{\alpha}{K})
\]

\[
\Gamma(N + 1 + \frac{\alpha}{K})
\]

\[
\Gamma(\frac{\alpha}{K}) \Gamma(N + 1)
\]
Indian Buffet Process

1st Representation: \( K \rightarrow \infty \)

Given:

\[
\Pr(Z | \alpha) = \prod_{k=1}^{K} \frac{\alpha}{K} \frac{\Gamma(m_k + \alpha/K)}{\Gamma(N) \Gamma(\frac{\alpha}{K})} \]

\[
\text{card}([Z]) = \binom{K}{K_0 \ldots K_{2^N-1}} = \frac{K!}{\prod_{h=0}^{2^N-1} K_h !}
\]

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

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

\[
= \frac{K!}{2^N-1} \left( \prod_{h=0}^{2^N-1} K_h ! \right) \frac{\Gamma(m_k + \alpha/K) \Gamma(N-m_k + 1)}{\Gamma(N+1+\alpha/K) \Gamma(N+1)} \prod_{k=1}^{K_+} \frac{\Gamma(\alpha/K) \Gamma(N+1)}{\Gamma(\frac{\alpha}{K}) \Gamma(N+1)}
\]
Indian Buffet Process

$1^{st}$ Representation: $K \rightarrow \infty$

Given:

$$\Pr(Z | \alpha) = \prod_{k=1}^{K} \frac{\alpha \Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})}$$

$$\text{card}([Z]) = \left\{ \begin{array}{c} K \\ K_0 \ldots K_{2^{N-1}} \end{array} \right\} = \frac{K!}{\prod_{h=0}^{2^{N-1}} K_h !}$$

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

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

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

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

1\textsuperscript{st} Representation: \( K \rightarrow \infty \)

\begin{align*}
\text{Given:} \\
\Pr(Z | \alpha) &= \prod_{k=1}^{K} \frac{\alpha}{K} \frac{\Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \\
\text{card}([Z]) &= \binom{K}{K_0 \ldots K_{2^{N-1}}} = \frac{K!}{\prod_{h=0}^{2^{N-1}-1} K_h !} \\
\Pr([Z] | \alpha) &= \Pr(Z | \alpha) \cdot \text{card}([Z]) \\
&= \frac{K!}{2^{N-1}} \left( \frac{\alpha}{K} \frac{\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{align*}

Derive when \( K \rightarrow \infty \) : \( (K_+ < \infty \) almost surely)
Indian Buffet Process

1st Representation: \( K \to \infty \)

Given:
\[
\Pr(Z | \alpha) = \prod_{k=1}^{K} \frac{\alpha}{K} \frac{\Gamma(m_k + \alpha)}{K} \frac{\Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})}
\]

\[
\text{card}([Z]) = \binom{K}{K_0 \ldots K_{2^N - 1}} = \frac{K!}{\prod_{h=0}^{2^N-1} K_h !}
\]

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

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

\[
= \frac{K!}{2^{N-1} \prod_{h=0}^{2^N-1} K_h !} \left( \frac{\alpha}{K} \frac{\Gamma(\frac{\alpha}{K})}{\Gamma(N + 1 + \frac{\alpha}{K})} \right)^K
\]

\[
= \frac{\Gamma(m_k + \frac{\alpha}{K})}{\Gamma(N - m_k + 1)} \frac{\Gamma(N + 1 + \frac{\alpha}{K})}{\Gamma(N + 1)}
\]

\[
(m_k - 1)! \cdot \frac{\alpha}{K}
\]
Indian Buffet Process

1st Representation: $K \to \infty$

Given:

$$\Pr(Z | \alpha) = \prod_{k=1}^{K} \frac{\alpha K}{K} \frac{\Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})}$$

$$\text{card}([Z]) = \left\{ \begin{array}{c} K \\ K_0 \ldots K_{2^{N-1}} \end{array} \right\} = \frac{K!}{\prod_{h=0}^{2^{N-1} - 1} K_h !}$$

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

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

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

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

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

1\textsuperscript{st} Representation: \( K \rightarrow \infty \)

Given:

\[
\Pr(Z | \alpha) = \prod_{k=1}^{K} \frac{\alpha}{K} \frac{\Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})}
\]

\[
\text{card}(\{Z\}) = \left\{ \begin{array}{c} K \\ K_0 \ldots K_{2^{N-1}} \end{array} \right\} = \frac{K!}{\prod_{h=0}^{2^{N-1}-1} K_h!}
\]

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

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

\[
= \frac{K!}{2^{N-1} \prod_{h=0}^{2^{N-1}-1} K_h!} \left( \frac{\alpha}{K} \frac{\Gamma(\frac{\alpha}{K}) \Gamma(N + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \right)^K
\]

\[
\prod_{k=1}^{K_+} \frac{\Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K}) \Gamma(N + 1)}
\]

\[
\prod_{k=1}^{K_+} (m_k - 1)! \cdot \frac{\alpha}{K} \frac{(N - m_k)!}{N!}
\]
Indian Buffet Process

1st Representation: \( K \rightarrow \infty \)

Given:

\[
\Pr(Z | \alpha) = \prod_{k=1}^{K} \frac{\alpha K}{K} \frac{\Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + \frac{\alpha}{K})}
\]

\[
\text{card}([Z]) = \binom{K}{K_0 \ldots K_{2^N-1}} = \frac{K!}{\prod_{h=0}^{2^N-1} K_h !}
\]

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

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

\[
= \frac{K!}{2^{N-1}} \prod_{h=0}^{2^N-1} K_h ! \left( \frac{\alpha K}{K} \frac{\Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + \frac{\alpha}{K})} \right)^K \prod_{k=1}^{K_+} \frac{\Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + \frac{\alpha}{K})}
\]

\[
\prod_{k=1}^{K_+} (m_k - 1)! \cdot \frac{\alpha}{K} \frac{(N - m_k)!}{N!}
\]
Indian Buffet Process

1st Representation: \( K \to \infty \)

Given:

\[
\Pr(Z | \alpha) = \prod_{k=1}^{K} \frac{\alpha}{K} \Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1) \frac{\Gamma(N + 1 + \frac{\alpha}{K})}{\Gamma(N + 1)} \]

\[
card([Z]) = \frac{K}{K_0 \ldots K_{2^{N-1}} + 1} = \frac{K!}{\prod_{h=0}^{2^{N-1}} K_h} \Gamma(N + 1) \Gamma(N + 1 + \frac{\alpha}{K})
\]

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

\[
\Pr([Z] | \alpha) = \Pr(Z | \alpha) \cdot card([Z])
\]

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

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

\[
\Pr(Z | \alpha) = \prod_{k=1}^{K} \frac{\alpha K (m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} 
\]

\[
\text{card}([Z]) = \begin{pmatrix} K \\ K_0 \ldots K_{2^{N-1}} \end{pmatrix} = \frac{K!}{\prod_{h=0}^{2^{N-1}} K_h!} 
\]

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

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

\[
= \frac{K!}{2^{N-1} \prod_{h=0}^{2^{N-1}} K_h!} \left( \frac{\alpha K \Gamma(\frac{\alpha}{K}) \Gamma(N + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \right)^K 
\]

\[
= \prod_{k=1}^{K_+} \left( \frac{\Gamma(\frac{\alpha}{K} + 1) \Gamma(N + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \right)^{m_k - 1} \frac{\alpha K (N - m_k)}{K} \frac{(N - m_k)!}{N!} 
\]
Indian Buffet Process

1st Representation: \( K \rightarrow \infty \)

Given:

\[
\Pr(Z | \alpha) = \prod_{k=1}^{K} \frac{\alpha}{K} \frac{\Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})}
\]

\[
card([Z]) = \sum_{K_0 \ldots K_2^{N-1}} = \frac{K!}{\prod_{h=0}^{K_h} K_h !} \cdot \frac{\Gamma(N + 1 + \frac{\alpha}{K})}{\Gamma(N + 1)}
\]

Derive when \( K \rightarrow \infty \) : \( (K_+ < \infty \) almost surely)
Indian Buffet Process

1st Representation: \( K \rightarrow \infty \)

Given:

\[
\Pr(Z | \alpha) = \prod_{k=1}^{K} \frac{\alpha}{K} \frac{\Gamma(m_k + \alpha) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \]

\[
\text{card}([Z]) = \left\{ \frac{K}{K_0 \ldots K_{2^N - 1}} \right\} = \frac{K!}{\prod_{h=0}^{2^N-1} K_h !} \]

Derive when \( K \rightarrow \infty \) : \( (K_+ < \infty \text{ almost surely}) \)

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

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

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

\[
\prod_{k=1}^{K_+} \frac{(m_k - 1)! \frac{\alpha (N - m_k)}{K N!}}{N!}
\]
Indian Buffet Process

1st Representation: \( K \rightarrow \infty \)

Given:

\[
\Pr(Z|\alpha) = \prod_{k=1}^{K} \frac{\alpha}{K} \frac{\Gamma(m_k + \frac{\alpha}{K})\Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})}
\]

\[
\text{card}([Z]) = \left\{ \frac{K}{K_0 \ldots K_{2^{N-1}}} \right\} = \frac{K!}{\prod_{h=0}^{2^{N-1}} K_h !}
\]

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

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

\[
= \frac{K!}{2^{N-1}} \prod_{h=0}^{2^{N-1}} K_h ! \left\{ \frac{\alpha}{K} \frac{\Gamma(\alpha)\Gamma(N+1)}{\Gamma(N+1+\frac{\alpha}{K})} \right\}^{K} \prod_{k=1}^{K_+} \frac{\Gamma(m_k + \frac{\alpha}{K})\Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})}
\]

\[
= \prod_{k=1}^{K_+} (m_k - 1)! \cdot \frac{\alpha}{K} \frac{(N - m_k)!}{N!}
\]
Indian Buffet Process

1st Representation: \( K \to \infty \)

Given:

\[
\Pr(Z | \alpha) = \prod_{k=1}^{K} \frac{\alpha}{K} \frac{\Gamma(m_k + \frac{\alpha}{K})\Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})}
\]

\[
\text{card}([Z]) = \left\{\frac{K}{K_0 \ldots K_{2^n - 1}}\right\} = \frac{K!}{\prod_{h=0}^{2^n - 1} K_h!}
\]

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

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

\[
= \frac{K!}{2^{2^n - 1} \prod_{h=0}^{2^n - 1} K_h!} \left(\frac{\alpha}{K} \frac{\Gamma(\frac{\alpha}{K})\Gamma(N + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})}\right)^{K} \prod_{k=1}^{K_+} \frac{\Gamma(m_k + \frac{\alpha}{K})\Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \prod_{j=1}^{N} \left(1 + \frac{\alpha}{j}\right)^K \prod_{k=1}^{K_+} (m_k - 1)! \cdot \frac{\alpha}{K} \frac{(N - m_k)!}{N!}
\]
Given:

\[
\Pr(Z \mid \alpha) = \prod_{k=1}^{K} \frac{\alpha}{K} \frac{\Gamma(m_k + \alpha) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})}
\]

\[
\text{card}([Z]) = \left\{ \begin{array}{c}
K \\
K_0 \cdots K_{2^{N-1}}
\end{array} \right\} = \frac{K!}{\prod_{h=0}^{2^{N-1}} K_h !}
\]

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

\[
\Pr([Z] \mid \alpha) = \Pr(Z \mid \alpha) \cdot \text{card}([Z])
\]

\[
= \frac{K!}{2^{N-1} \prod_{h=0}^{2^{N-1}} K_h !} \left( \frac{\alpha}{K} \frac{\Gamma(\frac{\alpha}{K}) \Gamma(N + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \right)^{K_+} \frac{\Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1)}
\]

\[
\prod_{k=1}^{K_+} (m_k - 1)! \cdot \frac{\alpha}{K} \frac{(N - m_k)!}{N!}
\]
Indian Buffet Process

1st Representation: \( K \rightarrow \infty \)

Given:

\[
\Pr(Z | \alpha) = \prod_{k=1}^{K} \frac{\alpha}{K} \frac{\Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \\
\text{card}(\{Z\}) = \frac{K}{K_0 \ldots K_{2^{N-1}}} = \frac{K!}{\prod_{h=0}^{2^N-1} K_h !} 
\]

Derive when \( K \rightarrow \infty : (K_+ < \infty \text{ almost surely}) \)

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

1st Representation: \( K \rightarrow \infty \)

Given:

\[
\Pr(Z \mid \alpha) = \prod_{k=1}^{K} \frac{\alpha}{K} \frac{\Gamma(m_k + \frac{\alpha}{K})}{\Gamma(N - m_k + 1)} \frac{\Gamma(N + 1 + \frac{\alpha}{K})}{\Gamma(N + 1)}
\]

\[
\text{card}([Z]) = \left\{ \frac{K}{K_0 \ldots K_{2^{N-1}}} \right\} = \frac{K!}{\prod_{h=0}^{2^{N-1}} K_h!}
\]

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

\[
\Pr([Z] \mid \alpha) = \Pr(Z \mid \alpha) \cdot \text{card}([Z])
\]

\[
= \frac{K!}{2^{N-1} \prod_{h=0}^{2^{N-1}} K_h!} \left( \frac{\alpha}{K} \frac{\Gamma(\frac{\alpha}{K})}{\Gamma(N + 1 + \frac{\alpha}{K})} \right)^N
\]

\[
= \frac{K!}{2^{N-1} \prod_{h=0}^{2^{N-1}} K_h!} \prod_{j=1}^{N} \left(1 + \frac{\alpha / j}{K} \right)^{-K} \prod_{k=1}^{K_+} \frac{\alpha (N - m_k)!}{K (N - m_k)!}
\]
Indian Buffet Process

1\textsuperscript{st} Representation: \( K \rightarrow \infty \)

Given:

\[
\Pr(Z | \alpha) = \prod_{k=1}^{K} \frac{\alpha}{K} \frac{\Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \\
\text{card}([Z]) = \binom{K}{K_0 \ldots K_{2^{n-1}}} = \frac{K!}{\prod_{h=0}^{2^{n-1}} K_h !} 
\]

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

\[
\Pr([Z] | \alpha) = \Pr(Z | \alpha) \cdot \text{card}([Z]) \\
= \frac{K!}{2^{N-1} \prod_{h=0}^{2^{n-1}} K_h !} \left( \frac{\alpha}{K} \frac{\Gamma(\frac{\alpha}{K}) \Gamma(N + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \right)^K \\
= \frac{K!}{\prod_{k=1}^{K_+} (m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)} \frac{\Gamma(\frac{\alpha}{K}) \Gamma(N + 1)}{\Gamma(N + 1)} \\
\prod_{j=1}^{N} (1 + \frac{\alpha / j}{K})^{-K} \prod_{k=1}^{K_+} (m_k - 1)! \cdot \frac{\alpha (N - m_k)!}{K \text{N!}} 
\]
Indian Buffet Process

1st Representation: $K \rightarrow \infty$

Given:

$$\Pr(Z | \alpha) = \frac{\frac{\alpha}{K} \frac{\Gamma(m_k + \frac{\alpha}{K})}{\Gamma(N-m_k+1)}}{\prod_{k=1}^{K} \frac{\Gamma(N+1+\frac{\alpha}{K})}{\Gamma(N+1+\frac{\alpha}{K})}}$$

$$\text{card}(\{Z\}) = \left\{ \frac{K}{K_0 \ldots K_{2^N-1}} \right\} = \frac{K!}{\prod_{h=0}^{2^N-1} K_h !}$$

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

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

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

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

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

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

1st Representation:  $K \rightarrow \infty$

Given:

$$\Pr(Z | \alpha) = \prod_{k=1}^{K} \frac{\alpha}{K} \frac{\Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})}$$

$$card([Z]) = \binom{K}{K_0 \ldots K_{2^{N-1}}} = \frac{K!}{\prod_{h=0}^{N} K_h !}$$

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

$$\Pr([Z] | \alpha) = \Pr(Z | \alpha) \cdot card([Z])$$

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

1st Representation: \( K \rightarrow \infty \)

Given:

\[
\Pr(Z \mid \alpha) = \prod_{k=1}^{K} \frac{\alpha}{K} \frac{\Gamma(m_k + \frac{\alpha}{K})}{\Gamma(N - m_k + 1)} \frac{\Gamma(N + 1 + \frac{\alpha}{K})}{\Gamma(N + 1 + \frac{\alpha}{K})} \\
\text{card}([Z]) = \left\{ \frac{K}{K_0 \ldots K_{2^N-1}} \right\} = \frac{K!}{\prod_{h=0}^{2^N-1} K_h!} \\
\]

Derive when \( K \rightarrow \infty \): \((K_+ < \infty \text{ almost surely})\)

\[
\Pr([Z] \mid \alpha) = \Pr(Z \mid \alpha) \cdot \text{card}([Z]) \\
= \frac{K!}{2^N - 1} \frac{K!}{K_0 \prod_{h=1}^{K} K_h!} \prod_{j=1}^{N} \left(1 + \frac{\alpha / j}{K}\right)^{-K} \\
\left(\prod_{k=1}^{K} (m_k - 1)! \cdot \frac{\alpha}{K} \frac{(N - m_k)!}{N!}\right)
\]
Indian Buffet Process

1st Representation: \( K \rightarrow \infty \)

Given:

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

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

\[
\Pr([Z] | \alpha) = \Pr(Z | \alpha) \cdot \text{card}([Z]) \\
= \frac{K!}{2^{N-1} K_0! \prod_{h=1}^{2^{N-1}} K_h!} \prod_{j=1}^{N} (1 + \frac{\alpha / j}{K})^{-K} \\
\prod_{k=1}^{K} \frac{(m_k - 1)! \alpha}{K} \frac{(N - m_k)!}{N!} \\
\frac{\alpha^{K_+}}{K^{K_+}} \prod_{k=1}^{K_+} (m_k - 1)! \frac{(N - m_k)!}{N!} \\
\]
Indian Buffet Process

1st Representation: $K \rightarrow \infty$

Given:

$$\Pr(Z | \alpha) = \prod_{k=1}^{K} \frac{\alpha}{K} \frac{\Gamma(m_k + \alpha)}{\Gamma(m_k + 1)} \frac{\Gamma(N - m_k + 1)}{\Gamma(N + \frac{\alpha}{K})}$$

$$\text{card}([Z]) = \left\{ \begin{array}{c} K \\ K_0 \ldots K_{2^{N-1}} \end{array} \right\} = \frac{K!}{\prod_{h=0}^{K} K_h !}$$

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

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

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

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

1st Representation: \( K \rightarrow \infty \)

Given:

\[
\Pr(Z \mid \alpha) = \prod_{k=1}^{K} \frac{\alpha K}{K} \frac{\Gamma(m_k + \frac{\alpha}{K})}{\Gamma(N - m_k + 1)} \frac{\Gamma(N + 1 + \frac{\alpha}{K})}{\prod_{h=0}^{K_h} K_h} \\
\text{card}([Z]) = \begin{pmatrix} \mathcal{K} \\ K_0 \ldots K_{2^N-1} \end{pmatrix} = \frac{K!}{\prod_{h=0}^{K_h} K_h!}
\]

Derive when \( K \rightarrow \infty \) : \( (K_+ < \infty \text{ almost surely}) \)

\[
\Pr([Z] \mid \alpha) = \Pr(Z \mid \alpha) \cdot \text{card}([Z]) \\
= \frac{K!}{2^N - 1} \prod_{j=1}^{N} \left(1 + \frac{\alpha / j}{K}\right)^{-K} \prod_{k=1}^{K} (m_k - 1)! \frac{\alpha}{K} \frac{(N - m_k)!}{N!} \prod_{k=1}^{K_+} (m_k - 1)! \frac{(N - m_k)!}{N!}
\]
Indian Buffet Process

1st Representation: \( K \to \infty \)

Given:

\[
\Pr(Z | \alpha) = \prod_{k=1}^{K} \frac{\alpha}{K} \frac{\Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})}
\]

\[
\text{card}([Z]) = \binom{K}{K_0 \ldots K_{2^{N-1}}} = \frac{K!}{\prod_{h=0}^{2^{N-1}} K_h!}
\]

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

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

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

\[
= \frac{K!}{K_0 ! K^{K_+} \prod_{h=1}^{2^{N-1}} K_h !} \left( \frac{\alpha^{K_+}}{\prod_{k=1}^{K_+} (m_k - 1)! (N - m_k)! / N!} \right)
\]
Indian Buffet Process

1st Representation: \( K \rightarrow \infty \)

Given:

\[
\begin{align*}
\Pr(Z | \alpha) &= \prod_{k=1}^{K} \frac{\alpha}{K} \frac{\Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \\
\text{card}(\{Z\}) &= \binom{K}{K_0 \ldots K_{2^{N-1}}} = \frac{K!}{\prod_{h=0}^{N-1} K_h !} \\
\end{align*}
\]

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

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

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

\[
\prod_{k=1}^{K_+} (m_k - 1)! \frac{(N - m_k)!}{N!}
\]
Indian Buffet Process

1st Representation: \( K \to \infty \)

Given:

\[
\Pr(Z | \alpha) = \prod_{k=1}^{K} \frac{\alpha}{K} \frac{\Gamma(m_k + \alpha)}{\Gamma(N - m_k + 1)} \frac{\Gamma(N + 1 + \frac{\alpha}{K})}{\Gamma(N + 1)}
\]

\[
\text{card}([Z]) = \begin{pmatrix} K \\ K_0 \ldots K_{2^{N-1}} \end{pmatrix} = \frac{K!}{\prod_{h=0}^{2^{N-1}} K_h!}
\]

Derive when \( K \to \infty \) : \( (K_+ < \infty \text{ almost surely}) \)

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

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

\[
\frac{\alpha^{K_+}}{2^{N-1} \prod_{h=1}^{K} K_h!} \prod_{k=1}^{K_+} (m_k - 1)! \frac{(N - m_k)!}{N!}
\]
Indian Buffet Process

1st Representation: \( K \rightarrow \infty \)

Given:
\[
\Pr(Z|\alpha) = \prod_{k=1}^{K} \frac{\alpha}{K} \frac{\Gamma(m_k + \frac{\alpha}{K})}{\Gamma(N - m_k + 1)} \frac{\Gamma(N + 1 + \frac{\alpha}{K})}{\prod_{h=0}^{2^{N-1}} K_h!}
\]
\[
\text{card}([Z]) = \left( \sum_{k=0}^{K} \frac{K}{K_0 \ldots K_{2^{N-1}}} \right) = \frac{K!}{\prod_{h=0}^{2^{N-1}} K_h!}
\]

Derive when \( K \rightarrow \infty \) : \( (K_+ < \infty \text{ almost surely}) \)

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

\[
= \frac{K!}{2^{N-1}} \prod_{j=1}^{N} \left( 1 + \frac{\alpha}{j} \right)^{-K} \prod_{j=1}^{N} e^{-\frac{\alpha}{j}} \prod_{h=1}^{K} \frac{\alpha^{K_+}}{m_k - 1} \frac{(N - m_k)!}{K! N!}
\]
Indian Buffet Process

1st Representation: \( K \rightarrow \infty \)

Given:

\[
\Pr(Z | \alpha) = \prod_{k=1}^{K} \frac{\alpha}{K} \frac{\Gamma(m_k + \alpha)}{\Gamma(N-m_k+1)} \frac{\Gamma(N+1+\frac{\alpha}{K})}{\Gamma(N+1)}
\]

\[
card([Z]) = \binom{K}{K_0...K_{2^{N-1}}} = \frac{K!}{\prod_{h=0}^{2^{N-1}} K_h!}
\]

Derive when \( K \rightarrow \infty : (K_+ < \infty \text{ almost surely}) \)

\[
\Pr([Z] | \alpha) = \Pr(Z | \alpha) \cdot card([Z])
\]

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

\[
\frac{\alpha^{K_+}}{\prod_{h=1}^{2^{N-1}} K_h!} \prod_{k=1}^{K_+} (m_k - 1)! \frac{(N-m_k)!}{N!}
\]
Indian Buffet Process

1st Representation: $K \rightarrow \infty$

Given:

$$\Pr(Z | \alpha) = \prod_{k=1}^{\infty} \frac{\alpha}{K} \frac{\Gamma(m_k + \alpha)\Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})}$$

$$\text{card}([Z]) = \left\{ \begin{array}{c} K \\ K_0 \ldots K_{2^N-1} \end{array} \right\} = \frac{K!}{\prod_{h=0}^{2^N-1} K_h!}$$

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

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

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

$$= e^{-\alpha H_N}$$

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

\[
\Pr(Z | \alpha) = \prod_{k=1}^{K} \frac{\alpha}{K} \frac{\Gamma(m_k + \frac{\alpha}{K}) \Gamma(N - m_k + 1)}{\Gamma(N + 1 + \frac{\alpha}{K})} \Gamma(N + 1 + \frac{\alpha}{K})
\]

\[
card([Z]) = \begin{cases} K & K_0 \ldots K_{2^{N-1}} \\ K_0 & \ldots & K_{2^{N-1}}! \end{cases} = \frac{K!}{\prod_{h=0}^{h=N} K_h!}
\]

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

\[
\Pr([Z] | \alpha) = \Pr(Z | \alpha) \cdot card([Z])
\]

\[
= e^{-\alpha H} \prod_{h=1}^{K} \frac{\alpha^{K_+}}{K_h!} \prod_{k=1}^{K} \frac{(m_k - 1)! (N - m_k)!}{N!}
\]

Note:

- \( \Pr([Z] | \alpha) \) is well defined because \( K_+ < \infty \) a.s.
- \( \Pr([Z] | \alpha) \) depends on \( K_h \):
  - The number of features (columns) with history \( h \)
  - Permute the rows (data points) does not change \( K_h \) (exchangeability)
Indian Buffet Process

2nd Representation: Customers & Dishes

- Indian restaurant with infinitely many infinite dishes (columns)

Poission Distribution: \( p(k \mid \lambda) = \frac{\lambda^k e^{-\lambda}}{k!} \)
Indian Buffet Process

2\textsuperscript{st} Representation: Customers & Dishes

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

Poission Distribution: $p(k \mid \lambda) = \frac{\lambda^k e^{-\lambda}}{k!}$
Indian Buffet Process

2\textsuperscript{st} Representation: Customers & Dishes

- Indian restaurant with infinitely many infinite dishes (columns)
  - The first customer tastes \textbf{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 $k$

Poission Distribution: $p(k \mid \lambda) = \frac{\lambda^k e^{-\lambda}}{k!}$
Indian Buffet Process
2nd Representation: Customers & Dishes

- Indian restaurant with infinitely many infinite dishes (columns)
  - The first customer tastes \textit{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 \textit{following} $K_1^{(i)}$ new dishes, sample $K_1^{(i)} \sim \text{Poisson}\left(\frac{\alpha}{i}\right)$

Poission Distribution: $p(k | \lambda) = \frac{\lambda^k e^{-\lambda}}{k!}$
Indian Buffet Process

2nd 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 $K_1^{(i)} \sim \text{Poisson}(\frac{\alpha}{i})$

Poission Distribution:

\[ p(k \mid \lambda) = \frac{\lambda^k e^{-\lambda}}{k!} \]
Indian Buffet Process

2nd 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 \( K_1^{(i)} \sim \text{Poisson}(\frac{\alpha}{i}) \)

Poission Distribution: \( p(k \mid \lambda) = \frac{\lambda^k e^{-\lambda}}{k!} \)
Indian Buffet Process

2\textsuperscript{st} Representation: Customers \& Dishes

- Indian restaurant with infinitely many infinite dishes (columns)
  - The first customer tastes \textbf{first} $K_{1}^{(1)}$ dishes, sample $K_{1}^{(1)} \sim \text{Poission}(\alpha)$
  - 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 \textbf{following} $K_{1}^{(i)}$ new dishes, sample $K_{1}^{(i)} \sim \text{Poission}\left(\frac{\alpha}{i}\right)$
Indian Buffet Process
2\textsuperscript{st} Representation: Customers & Dishes
Indian Buffet Process
2^{st} Representation: Customers & Dishes
Indian Buffet Process
2\textsuperscript{nd} Representation: Customers & Dishes

$K_1^{(i)}$ $K_2^{(i)}$ $K_3^{(i)}$
Indian Buffet Process

2nd Representation: Customers & Dishes

\[ P_{\text{new\_dish}} = \prod_{i}^{N} \left( \frac{\alpha}{i} \right)^{K_{1}^{(i)}} \frac{\alpha}{K_{1}^{(i)}} e^{-\alpha H_{N}} \right) \]

\[ \cong \frac{\alpha^{K_{+}} e^{-\alpha H_{N}}}{\prod_{i}^{N} K_{1}^{(i)}} \prod_{i}^{N} \frac{1}{i^{K_{1}^{(i)}}} \]
Indian Buffet Process

2nd Representation: Customers & Dishes

\[ P_{\text{new dish}} = \prod_{i}^{N} \frac{\left(\frac{\alpha}{i}\right)^{K_{1}^{(i)}} e^{-\frac{\alpha}{i}}}{K_{1}^{(i)}!} = \frac{\alpha^{K_{1}^{(i)}} e^{-\alpha H_{N}}}{\prod_{i}^{N} K_{1}^{(i)}} \prod_{i}^{N} \left(\frac{1}{i}\right)^{K_{1}^{(i)}} \]
Indian Buffet Process

2\textsuperscript{st} Representation: Customers & Dishes

\[
P_{\text{new\_dish}} = \prod_{i}^{N} \frac{\left(\frac{\alpha}{i}\right)^{K_1^{(i)}} e^{\frac{-\alpha}{i}}}{K_1^{(i)}!} = \frac{\alpha^{K_+} e^{-\alpha H_N}}{\prod_{i}^{N} K_1^{(i)}} \prod_{i}^{N} \frac{1}{i^{K_1^{(i)}}}
\]
Indian Buffet Process

2\textsuperscript{st} Representation: Customers & Dishes

\[ P_{\text{new\_dish}} = \prod_{i=1}^{N} \left( \frac{\alpha}{i} \right)^{K_1^{(i)}} \frac{\alpha}{K_1^{(i)}} e^{-\alpha H_N} \frac{\alpha^{K_1^{(i)}} e^{-\alpha H_N} \prod_{i=1}^{N} \left( \frac{1}{i} \right)^{K_1^{(i)}}}{\prod_{i=1}^{N} K_1^{(i)}} \]
Indian Buffet Process

2\textsuperscript{nd} Representation: Customers \& Dishes

\[ P_{\text{new\_dish}} = \prod_{i}^{N} \frac{\left(\frac{\alpha}{i}\right)^{K_{i}^{(i)}} e^{\frac{\alpha}{i}}}{K_{1}^{(i)}!} = \frac{\alpha^{K_{+}} e^{-\alpha H_{N}}}{\prod_{i}^{N} K_{1}^{(i)}} \prod_{i}^{N} \frac{1}{i^{K_{1}^{(i)}}} \]

\[ \frac{3!6!}{10!} \]
Indian Buffet Process

2\textsuperscript{st} Representation: Customers & Dishes

\[ P_{\text{new\_dish}} = \prod_{i}^{N} \left( \frac{\alpha}{i} \right)^{K_{1}^{(i)}} e^{-\frac{\alpha}{i}} \frac{K_{1}^{(i)}}{K_{1}^{(i)}}! = \alpha^{K_{1}^{(i)}} e^{-\alpha H_{N}} \prod_{i}^{N} \left( \frac{1}{i} \right)^{K_{1}^{(i)}} \prod_{i}^{N} \frac{1}{i} \right) \]

\[
\frac{3!6!}{10!}
\]
Indian Buffet Process

2\textsuperscript{st} Representation: Customers & Dishes

\[ P_{new\_dish} = \prod_{i}^{N} \left( \frac{\alpha}{i} \right)^{K^{(i)}} \frac{e^{-\frac{\alpha}{i}}}{K^{(i)}!} = \frac{\alpha^{K_{+}} e^{-\alpha H_{N}}}{\prod_{i}^{N} K^{(i)} \prod_{i}^{N} \left( \frac{1}{i} \right)^{K^{(i)}}} \]

\[
\begin{array}{c|c|c|c|c|c|c|c|c|c|c|c|c}
\hline
\hline
\end{array}
\]
Indian Buffet Process

2\textsuperscript{nd} Representation: Customers & Dishes

\[ P_{\text{new\_dish}} = \prod_{i=1}^{N} \frac{\left(\alpha\right)^{K_1^{(i)}}}{i^{K_1^{(i)}}} = \frac{\alpha^{K_+} e^{-\alpha H_N}}{K_1^{(N)}} \prod_{i=1}^{N} \frac{1}{i^{K_1^{(i)}}} \]

\[ \prod_{i=1}^{N} \frac{1}{i^{K_1^{(i)}}} P_{\text{old\_dishes}} = \prod_{k=1}^{K_+} \frac{(N-m_k)! (m_k-1)!}{N!} \]
Indian Buffet Process

2\textsuperscript{st} Representation: Customers & Dishes

From:
\[ P_{\text{new\_dish}} = \prod_i^N \left( \frac{\alpha}{K_1^{(i)}} \right)^{\frac{\alpha}{K_1^{(i)}}} e^{-\frac{\alpha}{K_1^{(i)}}} = \frac{\alpha^{K_+} e^{-\alpha H_N}}{\prod_i^N K_1^{(i)}} \prod_i^N \left( \frac{1}{i} \right)^{K_1^{(i)}} \]

\[ \prod_i^N \left( \frac{1}{i} \right)^{K_1^{(i)}} P_{\text{old\_dishes}} = \prod_k^N \frac{(N - m_k)! (m_k - 1)!}{N!} \]

We have:
\[ P_{\text{IBP}}(Z | \alpha) = \frac{\alpha^{K_+} e^{-\alpha H_N}}{\prod_i^N K_1^{(i)}} \prod_k^N \frac{(N - m_k)! (m_k - 1)!}{N!} \]
Indian Buffet Process
2nd Representation: Customers & Dishes

From:
\[ P_{\text{new dish}} = \prod_i^n \left( \frac{\alpha}{K_i} \right)^{K_i(i)} \frac{1}{K_i(i)!} = \frac{\alpha^K e^{-\alpha N}}{\prod K_i(i)} \prod_i^n \frac{1}{K_i(i)} \]
\[ \prod_i^n \left( \frac{1}{i} \right)^{K_i(i)} P_{\text{old dishes}} = \prod_k^{K} \frac{(N - m_k)!}{N!} \]

We have:
\[ P_{\text{IBP}}(Z \mid \alpha) = \frac{\alpha^K e^{-\alpha N}}{\prod K_i(i)} \prod_k^{K} \frac{(N - m_k)!}{N!} \]

Note:
Indian Buffet Process
2\textsuperscript{st} Representation: Customers & Dishes

From:
\[ P_{\text{new \_dish}} = \prod_i^n \left( \frac{\alpha}{i} \right)^{K_1^{(i)}} \frac{\alpha_i}{K_1^{(i)}} = \frac{\alpha^K e^{-\alpha H_N}}{\prod K_1^{(i)}} \prod_i^N \left( \frac{1}{i} \right)^{K_1^{(i)}} \]
\[
\prod_i^N \left( \frac{1}{i} \right)^{K_1^{(i)}} P_{\text{old \_dishes}} = \prod_k \frac{(N - m_k)! (m_k - 1)!}{N!}
\]

We have:
\[ P_{\text{IBP}}(Z \mid \alpha) = \frac{\alpha^K e^{-\alpha H_N}}{\prod_i^N K_1^{(i)}} \prod_k \frac{(N - m_k)! (m_k - 1)!}{N!} \]

\textbf{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

2\textsuperscript{st} Representation: Customers & Dishes

From:
\[
P_{\text{new\_dish}} = \prod_i \frac{\alpha^{K_1^{(i)}} e^{\frac{\alpha}{K_1^{(i)}}}}{K_1^{(i)}!} = \frac{\alpha^K e^{-\alpha H_N}}{\prod K_1^{(i)}} \prod_i \frac{1}{i^{K_1^{(i)}}}
\]

\[
\prod_i \frac{1}{i^{K_1^{(i)}}} P_{\text{old\_dishes}} = \prod_k \frac{(N - m_k)! (m_k - 1)!}{N!}
\]

We have:
\[
P_{IBP}(Z \mid \alpha) = \frac{\alpha^K e^{-\alpha H_N}}{\prod K_1^{(i)}} \prod_k \frac{(N - m_k)! (m_k - 1)!}{N!}
\]

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 \textit{lof} equivalent class $[Z]$
Indian Buffet Process

2\textsuperscript{st} Representation: Customers & Dishes

From:

\[ P_{\text{new\_dish}} = \prod_i \frac{\alpha^K_i}{K_1^{(i)}!} \frac{e^{\alpha_i}}{K_1^{(i)}!} = \frac{\alpha^{K_+} e^{-\alpha H_N} \prod_i \frac{1}{i^{K_1^{(i)}}}} {\prod_i K_1^{(i)} \prod_i \left( \frac{1}{i} \right)^{K_1^{(i)}}} \]

\[ \prod_i \frac{1}{i^{K_1^{(i)}}} P_{\text{old\_dishes}} = \prod_k \frac{(N - m_k)! (m_k - 1)!}{N!} \]

We have:

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

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

2nd Representation: Customers & Dishes

From:

\[ P_{\text{new\_dish}} = \prod_{i}^{N} \left( \frac{\alpha^{K_{1}^{(i)}}}{i^{K_{1}^{(i)}}} \right) = \frac{\alpha^{K_{1}^{(i)}} e^{-\alpha H_{N}}}{\prod_{i}^{N} K_{1}^{(i)}} \prod_{i}^{N} \frac{1}{i^{K_{1}^{(i)}}} \]

\[ \prod_{i}^{N} \left( \frac{1}{i^{K_{1}^{(i)}}} \right) P_{\text{old\_dishes}} = \prod_{k}^{k_{+}} \frac{N-m_{k}! (m_{k}-1)!}{N!} \]

We have:

\[ P_{IBP}(Z \mid \alpha) = \frac{\alpha^{K_{1}^{(i)}} e^{-\alpha H_{N}}}{\prod_{i}^{N} K_{1}^{(i)}} \prod_{k}^{K_{+}} \frac{(N-m_{k}! (m_{k}-1)!}{N!} \]

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:
  \[ \prod_{i=1}^{N} K_{1}^{(i)} \]
  \[ \frac{2^{N-1}}{\prod_{h=1}^{2^{N-1}} K_{h}!} \]
Indian Buffet Process

2nd Representation: Customers & Dishes

From:

\[ P_{\text{new\_dish}} = \prod_i^N \frac{\alpha_i^{K_1(i)} e^{\alpha_i}}{K_1(i)!} = \frac{\alpha^K e^{-\alpha H_N}}{\prod_i^N K_1(i)} \prod_i^N \left( \frac{1}{i} \right)^{K_1(i)} \]

\[ \prod_i^N \left( \frac{1}{i} \right)^{K_1(i)} P_{\text{old\_dishes}} = \prod_k^K \frac{(N - m_k)! (m_k - 1)!}{N!} \]

We have:

\[ P_{\text{IBP}}(Z \mid \alpha) = \frac{\alpha^K e^{-\alpha H_N}}{\prod_i^N K_1(i)} \prod_k^K \frac{(N - m_k)! (m_k - 1)!}{N!} \]

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:
  \[\prod_i^N K_1(i) \quad \prod_{h=1}^{2^{N-1}} K_h!\]
Indian Buffet Process

2nd Representation: Customers & Dishes

From:

\[ P_{\text{new
dish}} = \prod_i^n \left( \frac{\alpha}{i} \right)^{K_i^{(i)}} \frac{\alpha}{K_1^{(i)}}! = \frac{\alpha^{K+} e^{-\alpha H_N}}{\prod_i^n K_1^{(i)}} \prod_i^n \left( \frac{1}{i} \right)^{K_i^{(i)}} \]

\[ \prod_i^n \left( \frac{1}{i} \right)^{K_i^{(i)}} P_{\text{old
dish}} = \prod_k \frac{(N - m_k)!(m_k - 1)!}{N!} \]

We have:

\[ P_{\text{IBP}}(Z \mid \alpha) = \frac{\alpha^{K+} e^{-\alpha H_N}}{\prod_i^n K_1^{(i)}} \prod_k \frac{(N - m_k)!(m_k - 1)!}{N!} \]

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:

\[ \prod_i^{2^{N-1}} K_1^{(i)} \prod_{h=1}^{K} K_h! \]

\[ P_{\text{IBP}}([Z] \mid \alpha) = \frac{\alpha^{K+} e^{-\alpha H_N}}{\prod_{h=1}^{2^{N-1}} K_h!} \prod_k \frac{(N - m_k)!(m_k - 1)!}{N!} \]
Indian Buffet Process

2\textsuperscript{st} Representation: Customers & Dishes

From:

\[ P_{\text{new\_dish}} = \prod_i^n \left( \frac{\alpha_i}{i} \right)^{K_i(i)} \frac{\alpha}{K_1(i)!} = \frac{\alpha_{N}^{K_{+}} e^{-\alpha H_{N}}}{\prod_i^n K_i(i)} \prod_i^n \frac{1}{i}^{K_i(i)} \]

\[ \prod_i^n \left( \frac{1}{i} \right)^{K_i(i)} P_{\text{old\_dishes}} = \prod_k^{K_+} \frac{(N - m_k)! (m_k - 1)!}{N!} \]

We have:

\[ P_{\text{IBP}}(Z \mid \alpha) = \frac{\alpha_{N}^{K_{+}} e^{-\alpha H_{N}}}{\prod_i^n K_i(i)} \prod_k^{K_+} \frac{(N - m_k)! (m_k - 1)!}{N!} \]

**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:
  \[ \frac{\prod_i^n K_i(i)}{2^{N-1}} \prod_{h=1}^{K_{+}} K_h! \]

2\textsuperscript{nd} representation is equivalent to 1\textsuperscript{st} representation
Indian Buffet Process

3rd Representation: Distribution over Collections of Histories

- Directly generating the left ordered form \((lof)\) matrix \(Z\)
- For each history \(h\):
  - \(m_h\): number of non-zero elements in \(h\)
  - Generate \(K_h\) columns of history \(h\)
    
    \[K_h \sim \text{Poisson}(\alpha \frac{(m_h - 1)! (N - m_h)!}{N!})\]

- The distribution over collections of histories

\[
P(K) = \prod_{h=1}^{2^{N-1} - 1} \left( \frac{\alpha (m_h - 1)! (N - m_h)!}{N! K_h!} \right) \exp \left\{ -\alpha \frac{(m_h - 1)! (N - m_h)!}{N!} \right\}
\]

\[
= \frac{\alpha \Sigma_{h=1}^{2^{N-1} - 1} K_h}{\prod_{h=1}^{2^{N-1} - 1} K_h!} \exp\{\alpha H_N\} \prod_{h=1}^{2^{N-1} - 1} \left( \frac{(m_h - 1)! (N - m_h)!}{N!} \right)^{K_h}
\]

- Note:
  - Permute digits in \(h\) does not change \(m_h\) (nor \(P(K)\))
  - Permute rows means customers are exchangeable
Indian Buffet Process

- **Effective dimension of the model $K_+$:**
  - Follow Poission distribution: $K_+ \sim \text{Poisson}(\alpha H_N)$
  - Derives from 2$^{nd}$ representation by summing Poission components

- **Number of features possessed by each object:**
  - Follow Possion distribution: $\text{Poisson}(\alpha)$
  - Derives from 2$^{nd}$ representation:
    - The first customer chooses $\text{Poisson}(\alpha)$ dishes
    - The customers are exchangeable and thus can be permuted

- **$Z$ is sparse:**
  - Non-zero element
  - Derives from the 2$^{nd}$ (or 1$^{st}$) representation: $\alpha$
    - Expected number of non-zeros for each row is $N\alpha$
    - Expected entries in $Z$ is
IBP: Gibbs sampling

- Need to have the full conditional
  \[ p(z_{ik} = 1 \mid Z_{-(ik)}, X) \propto p(X \mid Z) p(z_{ik} = 1 \mid Z_{-(ik)}) \]
  - \( Z_{-(n,k)} \) denotes the entries of \( Z \) other than \( z_{nk} \).
  - \( p(X \mid Z) \) depends on the model chosen for the observed data.
- By exchangeability, consider generating row as the last customer:

- By IBP, in which
  \[ p(z_{ik} = 1 \mid z_{-ik}) = \frac{m_{-i,k}}{N} \]
  - If sample \( z_{ik}=0 \), and \( m_{k}=0 \): delete the row
  - At the end, draw new dishes from \( \text{Pois} \left( \frac{\alpha}{n} \right) \) with considering \( p(X \mid Z) \)
    - 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

- Slides: Tom Griffiths, “The Indian buffet process”.