Random partition-based inference schemes for feature allocations

Creighton Heaukulani
University of Cambridge

Joint work with

Daniel M. Roy

BNP 10, Raleigh–Durham
June 2015

Overview

1. We connect random partitions to random feature allocations.
2. Construct the Indian buffet process (IBP) from Chinese restaurant processes (CRPs).
3. Posterior inference on CRPs $\Rightarrow$ inference on IBP.
4. Generalize to any exchangeable partition, results in broad class of feature allocations.

Random partitions and feature allocations

A partition of $[n] := (1, \ldots, n)$ is a set of subsets of $[n]$, called <u>blocks</u>, whose union is $[n]$.

$$\Pi_5 = \Big\{\{1,5\},\{2,4\},\{3\}\Big\}$$

A feature allocation of $[n]$ is a (finite) set of subsets of $[n]$, called <u>features</u> (See [4, 3]).

$$F_5 = \Big\{\{1,2,5\},\{2,5\},\{3\},\{3\}\Big\}$$

Feature allocations allows the subsets to overlap.

Chinese restaurant process

$\mathrm{CRP}(\theta)$

Chinese restaurant process

$\mathrm{CRP}(\theta)$

- The first customer sits at the first table.

Chinese restaurant process

$\mathrm{CRP}(\theta)$

- The first customer sits at the first table.
- The $n+1$-st customer sits at table k with probability

$$\begin{cases} \frac{m_k}{\theta+n}, & \text{if } k \text{ old}, \\ \frac{\theta}{\theta+n}, & \text{if } k \text{ new}. \end{cases}$$

Chinese restaurant process

$\mathrm{CRP}(\theta)$

- The first customer sits at the first table.
- The $n+1$-st customer sits at table k with probability

$$\begin{cases} \frac{m_k}{\theta+n}, & \text{if } k \text{ old,} \\ \frac{\theta}{\theta+n}, & \text{if } k \text{ new.} \end{cases}$$

m_j is the # in cluster j

Chinese restaurant process
$\mathrm{CRP}(\theta)$

Indian buffet process
$\mathrm{IBP}(\gamma, \theta)$; [6, 4]

- The first customer sits at the first table.
- The $n+1$-st customer sits at table k with probability

$$\begin{cases} \frac{m_k}{\theta+n}, & \text{if } k \text{ old,} \\ \frac{\theta}{\theta+n}, & \text{if } k \text{ new.} \end{cases}$$

m_j is the # in cluster j

Chinese restaurant process
$\mathrm{CRP}(\theta)$

- The first customer sits at the first table.
- The $n+1$-st customer sits at table k with probability

$$\begin{cases} \frac{m_k}{\theta+n}, & \text{if } k \text{ old,} \\ \frac{\theta}{\theta+n}, & \text{if } k \text{ new.} \end{cases}$$

Indian buffet process
$\mathrm{IBP}(\gamma, \theta)$; [6, 4]

- The first customer takes $\mathrm{Poisson}(\gamma)$ dishes.

m_j is the # in cluster j

Chinese restaurant process

$\mathrm{CRP}(\theta)$

- The first customer sits at the first table.
- The $n+1$-st customer sits at table k with probability

$$\begin{cases} \frac{m_k}{\theta+n}, & \text{if } k \text{ old,} \\ \frac{\theta}{\theta+n}, & \text{if } k \text{ new.} \end{cases}$$

Indian buffet process

$\mathrm{IBP}(\gamma, \theta)$; [6, 4]

- The first customer takes Poisson(γ) dishes.
- The $n+1$-st customer
  - Takes previous dish k with probability $\frac{m_k}{\theta+n}$;

m_j is the # in cluster j

Chinese restaurant process

$\mathrm{CRP}(\theta)$

- The first customer sits at the first table.
- The $n+1$-st customer sits at table k with probability

$$\begin{cases} \frac{m_k}{\theta+n}, & \text{if } k \text{ old}, \\ \frac{\theta}{\theta+n}, & \text{if } k \text{ new}. \end{cases}$$

Indian buffet process

$\mathrm{IBP}(\gamma, \theta)$; [6, 4]

- The first customer takes $\mathrm{Poisson}(\gamma)$ dishes.
- The $n+1$-st customer
  - Takes previous dish k with probability $\frac{m_k}{\theta+n}$;
  - Takes $\mathrm{Poisson}(\gamma \frac{\theta}{\theta+n})$ new dishes.

m_j is the # in cluster j

Construct an IBP from i.i.d. CRPs

Construct an IBP from i.i.d. CRPs

– First customer enters buffet and takes Poisson(γ) dishes

Construct an IBP from i.i.d. CRPs

– First customer enters buffet and takes Poisson(γ) dishes

$$\overbrace{\Pi_1^{(1)} \qquad \Pi_1^{(2)}}^{\text{Poisson}(\gamma)} \qquad \stackrel{iid}{\sim} \qquad \text{CRP}(\theta)$$

Construct an IBP from i.i.d. CRPs

– First customer enters buffet and takes Poisson(γ) dishes

$$\underbrace{\overbrace{\Pi_1^{(1)} \qquad \Pi_1^{(2)}}^{\text{Poisson}(\gamma)}}_{} \quad \overset{iid}{\sim} \quad \text{CRP}(\theta)$$

$$\begin{array}{cc} \Pi_1^{(1)} & \Pi_1^{(2)} \\ \parallel & \parallel \\ \{1\} & \{1\} \end{array}$$

Construct an IBP from i.i.d. CRPs

– First customer enters buffet and takes Poisson(γ) dishes
– Mark blocks as special

$$\overbrace{\begin{matrix} \Pi_1^{(1)} & \quad & \Pi_1^{(2)} \\ \shortparallel & & \shortparallel \\ \{1\} & & \{1\} \end{matrix}}^{\text{Poisson}(\gamma)} \quad \overset{iid}{\sim} \quad \text{CRP}(\theta)$$

Construct an IBP from i.i.d. CRPs

– First customer enters buffet and takes Poisson(γ) dishes
– Mark blocks as special
– Create feature allocation from special blocks

$$\begin{array}{ccc} \Pi_1^{(1)} & & \Pi_1^{(2)} \\ \parallel & & \parallel \\ \{1\} & & \{1\} \end{array}$$

$$F_1 = \Big\{ \{1\} \quad , \quad \{1\} \Big\}$$

Construct an IBP from i.i.d. CRPs

– First customer enters buffet and takes Poisson(γ) dishes
– Mark blocks as special
– Create feature allocation from special blocks

$$\begin{array}{cc} \Pi_1^{(1)} & \Pi_1^{(2)} \\ \| & \| \\ \{1\} & \{1\} \end{array}$$

$$F_1 = \left\{ \{1\} \quad , \quad \{1\} \right\} \quad \sim \quad \mathrm{IBP}(\gamma, \theta)$$

Clearly.

Construct an IBP from i.i.d. CRPs

$$\begin{array}{ccc} \Pi_1^{(1)} & & \Pi_1^{(2)} \\ \| & & \| \\ \{1\} & & \{1\} \end{array}$$

Construct an IBP from i.i.d. CRPs

– 2nd customer enters buffet and decides on previous dishes

$$\begin{array}{cc} \Pi_1^{(1)} & \Pi_1^{(2)} \\ \| & \| \\ \{1\} & \{1\} \end{array}$$

Construct an IBP from i.i.d. CRPs

– 2nd customer enters buffet and decides on previous dishes

$$\begin{tikzcd}
2 \arrow[dr, bend left] \arrow[drr, bend left, "\sim \mathrm{CRP}(\theta)"] & & \\
 & \Pi_1^{(1)} \arrow[d, phantom, "\parallel"] & \Pi_1^{(2)} \arrow[d, phantom, "\parallel"] \\
 & \{1\} & \{1\}
\end{tikzcd}$$

Construct an IBP from i.i.d. CRPs

– 2nd customer enters buffet and decides on previous dishes

$$\begin{tikzcd}
2 \arrow[dr, bend left] \arrow[drr, bend left, "\sim \mathrm{CRP}(\theta)"] & & \\
& \Pi_2^{(1)} \arrow[d, phantom, "\parallel"] & \Pi_2^{(2)} \arrow[d, phantom, "\parallel"] \\
& \color{red}{\{1,2\}} & \color{red}{\{1\}} \\
& & \{2\}
\end{tikzcd}$$

Construct an IBP from i.i.d. CRPs

– 2nd customer enters buffet and decides on previous dishes

$$\begin{array}{cc} \Pi_2^{(1)} & \Pi_2^{(2)} \\ \| & \| \\ \{1,2\} & \{1\} \\ & \{2\} \end{array}$$

Construct an IBP from i.i.d. CRPs

– 2nd customer enters buffet and decides on previous dishes
– Samples new dishes

$$\begin{array}{cc} \Pi_2^{(1)} & \Pi_2^{(2)} \\ \| & \| \\ \{1,2\} & \{1\} \\ & \{2\} \end{array}$$

Construct an IBP from i.i.d. CRPs

– 2nd customer enters buffet and decides on previous dishes
– Samples new dishes

$$\Pi_2^{(1)} \quad \Pi_2^{(2)} \quad \Pi_2 \quad \overbrace{\quad \Pi_2 \quad}^{\text{Poisson}(\gamma)} \quad \Pi_2 \quad \overset{iid}{\sim} \quad \text{CRP}(\theta)$$

$$\begin{array}{cc} \| & \| \\ \{1,2\} & \{1\} \\ & \{2\} \end{array}$$

Construct an IBP from i.i.d. CRPs

– 2nd customer enters buffet and decides on previous dishes
– Samples new dishes

$$\begin{array}{ccccccc} & & & \overbrace{\qquad\qquad\qquad}^{\text{Poisson}(\gamma)} & & & \\ \Pi_2^{(1)} & \Pi_2^{(2)} & \Pi_2 & \Pi_2 & \Pi_2 & \overset{iid}{\sim} & \text{CRP}(\theta) \\ \parallel & \parallel & \parallel & \parallel & \parallel & & \\ {\color{red}\{1,2\}} & {\color{red}\{1\}} & \{1,2\} & \{1\} & \{1,2\} & & \\ & \{2\} & & \{2\} & & & \end{array}$$

Construct an IBP from i.i.d. CRPs

– 2nd customer enters buffet and decides on previous dishes
– Samples new dishes
– Keep partition if 2 forms a singleton; marked as special

$$\begin{array}{ccccccc} & & & \overbrace{\hspace{4em}}^{\text{Poisson}(\gamma)} & & & \\ \Pi_2^{(1)} & \Pi_2^{(2)} & \Pi_2 & \Pi_2 & \Pi_2 & \overset{iid}{\sim} & \text{CRP}(\theta) \\ \| & \| & \| & \| & \| & & \\ {\color{red}\{1,2\}} & {\color{red}\{1\}} & \{1,2\} & \{1\} & \{1,2\} & & \\ & \{2\} & & \{2\} & & & \end{array}$$

Construct an IBP from i.i.d. CRPs

– 2nd customer enters buffet and decides on previous dishes
– Samples new dishes
– Keep partition if 2 forms a singleton; marked as special

$$
\begin{array}{ccc}
\Pi_2^{(1)} & \Pi_2^{(2)} & \Pi_2^{(3)} \\
\| & \| & \| \\
\{1,2\} & \{1\} & \{1\} \\
 & \{2\} & \{2\}
\end{array}
$$

Construct an IBP from i.i.d. CRPs

– 2nd customer enters buffet and decides on previous dishes
– Samples new dishes
– Keep partition if 2 forms a singleton; marked as special
– Form a feature allocation from the special blocks

$$\Pi_2^{(1)} \quad \Pi_2^{(2)} \quad \Pi_2^{(3)}$$

$$\Pi_2^{(1)} = \{1,2\} \qquad \Pi_2^{(2)} = \{1\} \;\; \{2\} \qquad \Pi_2^{(3)} = \{1\} \;\; \{2\}$$

$$F_2 = \Big\{ \{1,2\} \;,\; \{1\} \;,\; \{2\} \Big\}$$

Construct an IBP from i.i.d. CRPs

– 2nd customer enters buffet and decides on previous dishes
– Samples new dishes
– Keep partition if 2 forms a singleton; marked as special
– Form a feature allocation from the special blocks

$$\Pi_2^{(1)} = \{\{1,2\}\} \qquad \Pi_2^{(2)} = \{\{1\}, \{2\}\} \qquad \Pi_2^{(3)} = \{\{1\}, \{2\}\}$$

$$F_2 = \left\{ \{1,2\}, \{1\}, \{2\} \right\} \sim \mathrm{IBP}(\gamma, \theta)$$

Easy to show

Construct an IBP from i.i.d. CRPs

– $n+1$-st customer enters buffet and decides on previous dishes

$\Pi_n^{(1)} \quad \cdots \quad \Pi_n^{(K_n)}$

Construct an IBP from i.i.d. CRPs

– $n+1$-st customer enters buffet and decides on previous dishes

$$\begin{tikzcd} n+1 \arrow[dr, bend left] \arrow[drrr, bend left] & & & & \sim \mathrm{CRP}(\theta) \\ & \Pi_n^{(1)} & \cdots & \Pi_n^{(K_n)} & \end{tikzcd}$$

Construct an IBP from i.i.d. CRPs

– $n+1$-st customer enters buffet and decides on previous dishes

$$
\begin{tikzcd}[column sep=small]
n+1 \arrow[dr, bend left] \arrow[drrr, bend left] & & & & \sim \mathrm{CRP}(\theta) \\
& \Pi_{n+1}^{(1)} & \cdots & \Pi_{n+1}^{(K_n)} &
\end{tikzcd}
$$

Construct an IBP from i.i.d. CRPs

– $n + 1$-st customer enters buffet and decides on previous dishes
– Samples new dishes

$$\Pi^{(1)}_{n+1} \quad \cdots \quad \Pi^{(K_n)}_{n+1} \qquad \overbrace{\Pi_{n+1} \quad \cdots \quad \Pi_{n+1}}^{\text{Poisson}(\gamma)} \quad \overset{iid}{\sim} \quad \text{CRP}(\theta)$$

Construct an IBP from i.i.d. CRPs

– $n+1$-st customer enters buffet and decides on previous dishes
– Samples new dishes

$$\Pi_{n+1}^{(1)} \quad \cdots \quad \Pi_{n+1}^{(K_n)} \qquad \underbrace{\{\Pi_n^{(K_n+1)}, \{n+1\}\} \quad \cdots \quad \{\Pi_n^{(K_{n+1})}, \{n+1\}\}}^{\text{Kept Singletons}}$$

Construct an IBP from i.i.d. CRPs

– $n+1$-st customer enters buffet and decides on previous dishes
– Samples new dishes
– Form a feature allocation from the special blocks

$$\Pi_{n+1}^{(1)} \quad \cdots \quad \Pi_{n+1}^{(K_n)} \qquad \{\Pi_n^{(K_n+1)}, \{n+1\}\} \quad \cdots \quad \{\Pi_n^{(K_{n+1})}, \{n+1\}\}$$

$$F_{n+1} \quad = \quad \left\{ \; \Pi_{n+1,\star}^{(1)} \quad , \quad \cdots \quad , \quad \Pi_{n+1,\star}^{(K_{n+1})} \right\}$$

Construct an IBP from i.i.d. CRPs

– $n+1$-st customer enters buffet and decides on previous dishes
– Samples new dishes
– Form a feature allocation from the special blocks

$$\Pi_{n+1}^{(1)} \quad \cdots \quad \Pi_{n+1}^{(K_n)} \qquad \{\Pi_n^{(K_n+1)}, \{n+1\}\} \quad \cdots \quad \{\Pi_n^{(K_{n+1})}, \{n+1\}\}$$

$$\bullet \quad \bullet \quad \bullet$$

$$F_{n+1} \quad = \quad \left\{ \Pi_{n+1,\star}^{(1)} \quad , \quad \cdots \quad , \quad \Pi_{n+1,\star}^{(K_{n+1})} \right\} \quad \sim \quad \mathrm{IBP}(\gamma, \theta)$$

Posterior Inference

- Maintain random partitions as latent variables
  - Gibbs sample assignments to $\Pi_{n+1}^{(k)}$ from posterior distribution
  - $\Rightarrow$ resamples assignments to F_{n+1}

$$\Pi_{n+1}^{(1)} \quad \cdots \quad \Pi_{n+1}^{(K_n)} \qquad \{\Pi_n^{(K_n+1)}, \{n+1\}\} \quad \cdots \quad \{\Pi_n^{(K_{n+1})}, \{n+1\}\}$$

$$\bullet \quad \bullet \quad \bullet$$

$$F_{n+1} \quad = \quad \left\{ \Pi_{n+1,\star}^{(1)} \quad , \quad \cdots \quad , \quad \Pi_{n+1,\star}^{(K_{n+1})} \right\}$$

Posterior Inference

$$
\begin{array}{ccc}
\Pi_5^{(1)} & \Pi_5^{(2)} & \Pi_5^{(3)} \\
\shortparallel & \shortparallel & \shortparallel \\
\{1,3,5\} & \{1,2\} & \{1,4\} \\
\{2,4\} & \{3\} & \{2,5\} \\
 & \{4,5\} & \{3\}
\end{array}
$$

$$
F_5 = \Big\{ \{1,3,5\} \;,\; \{1,2\} \;,\; \{3\} \Big\}
$$

Posterior Inference

– Resample assignment of 3 in $\Pi_5^{(k)} \Rightarrow$ Resample allocation in F_5

$$\begin{array}{ccc}
\Pi_5^{(1)} & \Pi_5^{(2)} & \Pi_5^{(3)} \\
\| & \| & \| \\
\{1,3,5\} & \{1,2\} & \{1,4\} \\
\{2,4\} & \{3\} & \{2,5\} \\
 & \{4,5\} & \{3\}
\end{array}$$

$$F_5 = \Big\{ \ \{1,3,5\} \ , \ \{1,2\} \ , \ \{3\} \ \Big\}$$

Posterior Inference

– Resample assignment of 3 in $\Pi_5^{(k)} \Rightarrow$ Resample allocation in F_5

$$\begin{array}{ccc} \Pi_5^{(1)} & \Pi_5^{(2)} & \Pi_5^{(3)} \\ \parallel & \parallel & \parallel \\ \{1,5\} & \{1,2\} & \{1,4\} \\ \{2,4\} & \{3\} & \{2,5\} \\ & \{4,5\} & \{3\} \end{array}$$

Remove 3 from $\Pi_5^{(1)}$

$$F_5 = \Big\{ \{1,5\}, \{1,2\}, \{3\} \Big\}$$

Posterior Inference

– Resample assignment of 3 in $\Pi_5^{(k)}$ $\Rightarrow$ Resample allocation in F_5

3

$$\begin{array}{ccc} \Pi_5^{(1)} & \Pi_5^{(2)} & \Pi_5^{(3)} \\ \shortparallel & \shortparallel & \shortparallel \\ \{1,5\} & \{1,2\} & \{1,4\} \\ \{2,4\} & \{3\} & \{2,5\} \\ \{\quad\} & \{4,5\} & \{3\} \end{array}$$

$\sim \mathrm{CRP}(\theta)$

Resample 3 in $\Pi_5^{(1)}$
from posterior

$$F_5 = \Big\{ \{1,5\} , \{1,2\} , \{3\} \Big\}$$

Posterior Inference

– Resample assignment of 3 in $\Pi_5^{(k)} \Rightarrow$ Resample allocation in F_5

3

$\Pi_5^{(1)}$	$\Pi_5^{(2)}$	$\Pi_5^{(3)}$
$\parallel$	$\parallel$	$\parallel$
$\{1,5\}$	$\{1,2\}$	$\{1,4\}$
$\{2,4\}$	$\{3\}$	$\{2,5\}$
$\{\quad\}$	$\{4,5\}$	$\{3\}$

$\sim \mathrm{CRP}(\theta)$

Resample 3 in $\Pi_5^{(1)}$
from posterior

Typically only 2
likelihood terms!

$$F_5 = \Big\{ \{1,5\} , \{1,2\} , \{3\} \Big\}$$

Posterior Inference

– Resample assignment of 3 in $\Pi_5^{(k)} \Rightarrow$ Resample allocation in F_5

$$
\begin{array}{ccc}
\Pi_5^{(1)} & \Pi_5^{(2)} & \Pi_5^{(3)} \\
\| & \| & \| \\
\color{red}{\{1,5\}} & \color{red}{\{1,2\}} & \{1,4\} \\
\{2,4\} & \{3\} & \{2,5\} \\
\{3\} & \{4,5\} & \color{red}{\{3\}}
\end{array}
$$

$$F_5 = \Big\{ \{1,5\} , \{1,2\} , \{3\} \Big\}$$

Posterior Inference

– Resample assignment of 3 in $\Pi_5^{(k)} \Rightarrow$ Resample allocation in F_5
– Resample singletons

$$
\begin{array}{ccc}
\Pi_5^{(1)} & \Pi_5^{(2)} & \Pi_5^{(3)} \\
\| & \| & \| \\
\color{red}{\{1,5\}} & \color{red}{\{1,2\}} & \{1,4\} \\
\{2,4\} & \{3\} & \{2,5\} \\
\{3\} & \{4,5\} & \color{red}{\{3\}}
\end{array}
$$

$$F_5 = \Big\{ \{1,5\} , \{1,2\} , \{3\} \Big\}$$

Posterior Inference

– Resample assignment of 3 in $\Pi_5^{(k)}$ $\Rightarrow$ Resample allocation in F_5
– Resample singletons

$$\begin{array}{ccc} \Pi_5^{(1)} & \Pi_5^{(2)} & \Pi_5^{(3)} \\ \| & \| & \| \\ \{1,5\} & \{1,2\} & \{1,4\} \\ \{2,4\} & \{3\} & \{2,5\} \\ \{3\} & \{4,5\} & \{3\} \end{array}$$

Propose removing singletons

$$F_5 = \Big\{ \{1,5\} , \{1,2\} , \{3\} \Big\}$$

Posterior Inference

– Resample assignment of 3 in $\Pi_5^{(k)} \Rightarrow$ Resample allocation in F_5
– Resample singletons

$$
\begin{array}{cc}
\Pi_5^{(1)} & \Pi_5^{(2)} \\
\parallel & \parallel \\
\{1,5\} & \{1,2\} \\
\{2,4\} & \{3\} \\
\{3\} & \{4,5\}
\end{array}
$$

$$F_5 = \left\{ \{1,5\} , \{1,2\} \right\}$$

Posterior Inference

– Resample assignment of 3 in $\Pi_5^{(k)} \Rightarrow$ Resample allocation in F_5
– Resample singletons

$$
\begin{array}{ccccccc}
\Pi_5^{(1)} & \Pi_5^{(2)} & \Pi_5 & \overbrace{\Pi_5}^{\text{Poisson}(\gamma)} & \Pi_5 & \overset{iid}{\sim} & \text{CRP}(\theta) \\
\| & \| & & & & & \\
\{1,5\} & \{1,2\} & & & & & \\
\{2,4\} & \{3\} & & & & & \\
\{3\} & \{4,5\} & & & & &
\end{array}
$$

$$F_5 = \left\{ \{1,5\} , \{1,2\} \right\}$$

Posterior Inference

– Resample assignment of 3 in $\Pi_5^{(k)}$ $\Rightarrow$ Resample allocation in F_5
– Resample singletons

$$
\begin{array}{ccccccc}
 & & & \overbrace{\hspace{6em}}^{\text{Poisson}(\gamma)} & & & \\
\Pi_5^{(1)} & \Pi_5^{(2)} & \Pi_5 & \Pi_5 & \Pi_5 & \overset{iid}{\sim} & \text{CRP}(\theta) \\
\| & \| & \| & \| & \| & & \\
\{1,5\} & \{1,2\} & \{1,2,3,5\} & \{1\} & \{1,2\} & & \\
\{2,4\} & \{3\} & \{4\} & \{2\} & \{3\} & & \\
\{3\} & \{4,5\} & & \{3\} & \{4,5\} & & \\
 & & & \{4,5\} & & &
\end{array}
$$

$$
F_5 \quad = \quad \Big\{ \quad \{1,5\} \quad , \quad \{1,2\} \quad \Big\}
$$

Posterior Inference

– Resample assignment of 3 in $\Pi_5^{(k)}$ $\Rightarrow$ Resample allocation in F_5
– Resample singletons

$$\begin{array}{ccccccc}
 & & & \overbrace{\hphantom{\Pi_5 \quad \Pi_5}}^{\text{Poisson}(\gamma)} & & & \\
\Pi_5^{(1)} & \Pi_5^{(2)} & \Pi_5 & \Pi_5 & \Pi_5 & \overset{iid}{\sim} & \text{CRP}(\theta) \\
\| & \| & \| & \| & \| & & \\
\{1,5\} & \{1,2\} & \{1,2,3,5\} & \{1\} & \{1,2\} & & \\
\{2,4\} & \{3\} & \{4\} & \{2\} & \{3\} & & \\
\{3\} & \{4,5\} & & \{3\} & \{4,5\} & & \\
 & & & \{4,5\} & & &
\end{array}$$

$$F_5 = \Big\{ \{1,5\} , \{1,2\} \Big\}$$

Posterior Inference

– Resample assignment of 3 in $\Pi_5^{(k)} \Rightarrow$ Resample allocation in F_5
– Resample singletons

$\Pi_5^{(1)}$	$\Pi_5^{(2)}$	$\Pi_5^{(3)}$	$\Pi_5^{(4)}$
$\shortparallel$	$\shortparallel$	$\shortparallel$	$\shortparallel$
$\{1,5\}$	$\{1,2\}$	$\{1\}$	$\{1,2\}$
$\{2,4\}$	$\{3\}$	$\{2\}$	$\{3\}$
$\{3\}$	$\{4,5\}$	$\{3\}$	$\{4,5\}$
		$\{4,5\}$	

$$F_5 = \Big\{ \ \{1,5\} \ , \ \{1,2\} \ \Big\}$$

Posterior Inference

– Resample assignment of 3 in $\Pi_5^{(k)} \Rightarrow$ Resample allocation in F_5
– Resample singletons

$$
\begin{array}{cccc}
\Pi_5^{(1)} & \Pi_5^{(2)} & \Pi_5^{(3)} & \Pi_5^{(4)} \\
\parallel & \parallel & \parallel & \parallel \\
\color{red}{\{1,5\}} & \color{red}{\{1,2\}} & \{1\} & \{1,2\} \\
\{2,4\} & \{3\} & \{2\} & \color{red}{\{3\}} \\
\{3\} & \{4,5\} & \color{red}{\{3\}} & \{4,5\} \\
 & & \{4,5\} &
\end{array}
$$

$$
F_5 = \Big\{ \{1,5\}, \{1,2\}, \color{red}{\{3\}}, \color{red}{\{3\}} \Big\}
$$

Posterior Inference

– Resample assignment of 3 in $\Pi_5^{(k)} \Rightarrow$ Resample allocation in F_5
– Resample singletons

$$\begin{array}{cccc}
\Pi_5^{(1)} & \Pi_5^{(2)} & \Pi_5^{(3)} & \Pi_5^{(4)} \\
\parallel & \parallel & \parallel & \parallel \\
\{1,5\} & \{1,2\} & \{1\} & \{1,2\} \\
\{2,4\} & \{3\} & \{2\} & \{3\} \\
\{3\} & \{4,5\} & \{3\} & \{4,5\} \\
 & & \{4,5\} &
\end{array}$$

$$F_5 = \left\{ \{1,5\} , \{1,2\} , \{3\} , \{3\} \right\}$$

Accept/Reject with MH proposal

Posterior Inference

Key Points:

- Same complexity as typical Gibbs sampling for the IBP
- Only requires a posterior inference procedure for the CRP

Exchangeable partitions (more generally)

A random partition $\Pi_n = \{A_1, \ldots, A_{B_n}\}$ of $[n]$ is exchangeable iff

$$\mathbb{P}\{B_n = k, |A_1| = n_1, \ldots, |A_k| = n_k\} = f_\Pi(n_1, \ldots, n_k)$$

for some symmetric function f_Π, called the exchangeable partition probability function, or **EPPF**.

The EPPF f_Π characterizes the distribution of Π_n.

Exchangeable partitions (more generally)

To sample f_Π-CRP:

Customer sits at table j with probability

$$\begin{cases} \frac{f_\Pi(N_1,\ldots,N_j+1,\ldots,N_k)}{f_\Pi(N_1,\ldots,N_k)}, & \text{if } j \text{ old,} \\ \\ \frac{f_\Pi(N_1,\ldots,N_k,1)}{f_\Pi(N_1,\ldots,N_k)}, & \text{if } j \text{ new.} \end{cases}$$

N_j is number assigned to j-th block
k is number of blocks

Generalized IBPs

– First customer samples Poisson(γ) i.i.d. CRP partitions

Generalized IBPs

– First customer samples Poisson(γ) i.i.d. ~~CRP~~ f_Π partitions

Generalized IBPs

– First customer samples Poisson(γ) i.i.d. ~~CRP~~ f_Π partitions
– $n + 1$-st customer samples previous dishes

$\Pi_n^{(1)} \quad \cdots \quad \Pi_n^{(K_n)}$

Generalized IBPs

– First customer samples Poisson(γ) i.i.d. ~~CRP~~ f_Π partitions
– $n+1$-st customer samples previous dishes

$$\begin{tikzcd}
n+1 \arrow[dr, bend left] \arrow[drrr, bend left] & & & \sim \mathrm{CRP}(\theta) \\
& \Pi^{(1)}_{n+1} & \cdots & \Pi^{(K_n)}_{n+1}
\end{tikzcd}$$

Generalized IBPs

– First customer samples Poisson(γ) i.i.d. ~~CRP~~ f_Π partitions
– $n+1$-st customer samples previous dishes

$$\begin{tikzcd} n+1 \arrow[dr, bend left] \arrow[drrr, bend left] & & & & \sim \xcancel{\mathrm{CRP}(\theta)} \quad f_\Pi \\ & \Pi^{(1)}_{n+1} & \cdots & \Pi^{(K_n)}_{n+1} & \end{tikzcd}$$

Generalized IBPs

– First customer samples Poisson(γ) i.i.d. ~~CRP~~ f_Π partitions
– $n+1$-st customer samples previous dishes
– Samples new dishes

$$\Pi_{n+1}^{(1)} \quad \cdots \quad \Pi_{n+1}^{(K_n)} \qquad \overbrace{\Pi_{n+1} \quad \cdots \quad \Pi_{n+1}}^{\text{Poisson}(\gamma)} \quad \overset{iid}{\sim} \quad \text{CRP}(\theta)$$

Generalized IBPs

– First customer samples Poisson(γ) i.i.d. ~~CRP~~ f_Π partitions
– $n+1$-st customer samples previous dishes
– Samples new dishes

$$\Pi_{n+1}^{(1)} \quad \cdots \quad \Pi_{n+1}^{(K_n)} \qquad \overbrace{\Pi_{n+1} \quad \cdots \quad \Pi_{n+1}}^{\text{Poisson}(\gamma)} \quad \overset{iid}{\sim} \quad \xcancel{\text{CRP}(\theta)}$$

$$f_\Pi$$

Generalized IBPs

– First customer samples Poisson(γ) i.i.d. ~~CRP~~ f_Π partitions
– $n+1$-st customer samples previous dishes
– Samples new dishes

$$\Pi_{n+1}^{(1)} \quad \cdots \quad \Pi_{n+1}^{(K_n)} \qquad \overbrace{\{\Pi_n^{(K_n+1)}, \{n+1\}\} \quad \cdots \quad \{\Pi_n^{(K_{n+1})}, \{n+1\}\}}^{\text{Kept Singletons}}$$

Generalized IBPs

– First customer samples Poisson(γ) i.i.d. ~~CRP~~ f_Π partitions
– $n+1$-st customer samples previous dishes
– Samples new dishes

$$\Pi_{n+1}^{(1)} \quad \cdots \quad \Pi_{n+1}^{(K_n)} \qquad \{\Pi_n^{(K_n+1)}, \{n+1\}\} \quad \cdots \quad \{\Pi_n^{(K_{n+1})}, \{n+1\}\}$$

$$\bullet \quad \bullet \quad \bullet$$

$$F_{n+1} \quad = \quad \left\{ \Pi_{n+1,\star}^{(1)} \quad , \quad \cdots \quad , \quad \Pi_{n+1,\star}^{(K_{n+1})} \right\}$$

Generalized IBPs

– First customer samples Poisson(γ) i.i.d. ~~CRP~~ f_Π partitions
– $n+1$-st customer samples previous dishes
– Samples new dishes

$$\Pi_{n+1}^{(1)} \quad \cdots \quad \Pi_{n+1}^{(K_n)} \qquad \{\Pi_n^{(K_n+1)}, \{n+1\}\} \quad \cdots \quad \{\Pi_n^{(K_{n+1})}, \{n+1\}\}$$

$$\bullet \quad \bullet \quad \bullet$$

$$F_{n+1} = \left\{ \Pi_{n+1,\star}^{(1)}, \cdots, \Pi_{n+1,\star}^{(K_{n+1})} \right\} \sim \mathrm{IBP}(\gamma, \theta)$$

Generalized IBPs

– First customer samples Poisson(γ) i.i.d. ~~CRP~~ f_Π partitions
– $n+1$-st customer samples previous dishes
– Samples new dishes

$$\Pi_{n+1}^{(1)} \quad \cdots \quad \Pi_{n+1}^{(K_n)} \qquad \{\Pi_n^{(K_n+1)}, \{n+1\}\} \quad \cdots \quad \{\Pi_n^{(K_{n+1})}, \{n+1\}\}$$

$$F_{n+1} = \left\{ \Pi_{n+1,\star}^{(1)}, \cdots, \Pi_{n+1,\star}^{(K_{n+1})} \right\} \sim \cancel{\mathrm{IBP}(\gamma,\theta)}$$

$\mathrm{IBP}(\gamma, f_\Pi)$

Generalized IBP; Roy (2014) [9]

Generalized IBPs

Generalized IBPs are parameterized by an EPPF

EPPF		Feature allocation model
Dirichlet process	$\rightarrow$	IBP (G&G, 2006; GG&S, 2007) [6, 4]
Pitman–Yor process	$\rightarrow$	stable IBP (T&G, 2009; BJ&P, 2012) [11, 1]
Gibbs-type prior (G&P, 2006) [5]	$\rightarrow$	Gibbs-type IBP (H&R, 2015) [7]

Inference for an f_Π-IBP

Inference algorithm looks the same.

Inference for an f_{Π}-IBP

Inference algorithm looks the same.

$$\begin{array}{cccc} 3 & & & \\ & \Pi_5^{(1)} & \Pi_5^{(2)} & \Pi_5^{(3)} \\ & \| & \| & \| \\ & \color{red}{\{1,5\}} & \color{red}{\{1,2\}} & \{1,4\} \\ & \{2,4\} & \{3\} & \{2,5\} \\ & \{\quad\} & \{4,5\} & \color{red}{\{3\}} \end{array}$$

$$\sim \xcancel{\mathrm{CRP}(\theta)}\ \color{blue}{f_{\Pi}}$$

$$F_5 = \Big\{ \{1,5\}, \{1,2\}, \{3\} \Big\}$$

Inference for an f_Π-IBP

Inference algorithm looks the same.

$$
\begin{array}{cccccc}
\Pi_5^{(1)} & \Pi_5^{(2)} & \Pi_5 & \underbrace{\Pi_5}^{\text{Poisson}(\gamma)} & \Pi_5 & \overset{iid}{\sim} \; \underset{f_\Pi}{\xcancel{\text{CRP}(\theta)}} \\
\| & \| & & & & \\
\{1,5\} & \{1,2\} & & & & \\
\{2,4\} & \{3\} & & & & \\
\{3\} & \{4,5\} & & & &
\end{array}
$$

$$F_5 = \Big\{ \{1,5\} , \{1,2\} \Big\}$$

Probability of a new species

EPPF	# singletons/new features
$\mathrm{CRP}(\theta)$ (Dirichlet process)	$\mathrm{Poisson}\left(\gamma \times \frac{\theta}{\theta+n}\right)$
$\mathrm{CRP}(\theta, \alpha)$ (Pitman–Yor process)	$\mathrm{Poisson}\left(\gamma \times \frac{\Gamma(\theta+1)\Gamma(\theta+\alpha+n)}{\Gamma(\theta+n+1)\Gamma(\theta+\alpha)}\right)$
f_Π	$\mathrm{Poisson}\left(\gamma \times \mathbb{E}\left[\frac{f_\Pi(N_1,\ldots,N_k,1)}{f_\Pi(N_1,\ldots,N_k)}\right]\right)$

singletons/features is determined by probability of a new species

$$\frac{f_\Pi(N_1,\ldots,N_k,1)}{f_\Pi(N_1,\ldots,N_k)}$$

Generalizations: Hierarchical beta processes

– Feature allocations induced by hierarchical beta processes (T&J, 2009) [12]

Generalizations: Hierarchical beta processes

– Feature allocations induced by <u>hierarchical beta processes</u> (T&J, 2009) [12]

– Multiple blocks may be special (Roy, 2015) [9]

$\Pi_5^{(1)}$	$\Pi_5^{(2)}$	$\Pi_5^{(3)}$	$\Pi_5^{(4)}$
$\parallel$	$\parallel$	$\parallel$	$\parallel$
$\{1,5\}$	$\{1,2\}$	$\{1\}$	$\{1,2\}$
$\{2,4\}$	$\{3\}$	$\{2\}$	$\{3\}$
$\{3\}$	$\{4,5\}$	$\{3\}$	$\{4,5\}$
		$\{4,5\}$	

$$F_5 = \Big\{ \{1,3,5\}, \{1,2\}, \{3,4,5\}, \{3\} \Big\}$$

Generalizations: Dynamic IBPs

Replace with a Markov process of partitions (Fragmentation and Coagulation processes (Teh et al., 2010) [10]).

Generalizations: Dynamic IBPs

Replace with a Markov process of partitions (Fragmentation and Coagulation processes (Teh et al., 2010) [10]).

Produces a Markov process of feature allocations (current work with D. Roy and Z. Ghahramani).

$$\Pi_n^{(1)}(t) \qquad \cdots \qquad \Pi_n^{(K_n)}(t)$$

$$F_n(t) = \left\{ \Pi_{n,\star}^{(1)}(t) \, , \, \cdots \, , \, \Pi_{n,\star}^{(K_n)}(t) \right\} \sim \mathrm{IBP}(\gamma, f_\Pi)$$

Generalizations: Multisets

– The urn scheme intuition extends to multisets (see my talk from BNP9).

– Negative binomial processes (Broderick et al., 2014) [2]; (Zhou et al., 2014) [13]; (H& R, 2015) [8]

– Partition perspective results in generalizations and inference procedures.

Summary

1. Construct feature allocations (FAs) from random partitions
2. Different random partitions result in different classes of FAs
3. Posterior inference on the partitions results in posterior inference in the FAs

[1] T. Broderick, M. I. Jordan, and J. Pitman. Beta processes, stick-breaking and power laws. Bayesian Analysis, 7(2):439–476, 2012.

[2] T. Broderick, L. Mackey, J. Paisley, and M. I. Jordan. Combinatorial clustering and the beta-negative binomial process. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2014. Special issue on Bayesian nonparametrics.

[3] T. Broderick, J. Pitman, and M. I. Jordan. Feature allocations, probability functions, and paintboxes. Bayesian Analysis, 8(4):801–836, 2013.

[4] Z. Ghahramani, T. L. Griffiths, and P. Sollich. Bayesian nonparametric latent feature models. In Bayesian Statistics, volume 8, pages 201–226. Oxford University Press, 2007. See also the discussion and rejoinder.

[5] A. Gnedin and J. Pitman. Exchangeable Gibbs partitions and Stirling triangles. Journal of Mathematical Sciences, 138(3):5674–5685, 2006.

[6] T. L. Griffiths and Z. Ghahramani. Infinite latent feature models and the Indian buffet process. In Advances in Neural Information Processing Systems 18, 2006.

[7] C. Heaukulani and D. M. Roy. Gibbs-type Indian buffet processes. Preprint.

[8] C. Heaukulani and D. M. Roy. The combinatorial structure of beta negative binomial processes. arXiv Preprint: 1401.0062 [math.ST] (version 3), 2015. To appear in Bernoulli.

[9] Daniel M. Roy. The continuum-of-urns scheme, generalized beta and Indian buffet processes, and hierarchies thereof. arXiv Preprint: 1501.00208 [math.PR] (version 1), 2014.

[10] Y. W. Teh, C. Blundell, and L. T. Elliott. Modelling genetic variations with fragmentation-coagulation processes. In Advances in Neural Information Processing Systems 24, 2011.

[11] Y. W. Teh and D. Görür. Indian buffet processes with power-law behavior. In Advances in Neural Information Processing Systems 22, 2009.

[12] R. Thibaux and M. I. Jordan. Hierarchical beta processes and the Indian buffet process. In Proceedings of the 24th International Conference on Machine Learning, 2007.

[13] M. Zhou, L. Hannah, D. Dunson, and L. Carin. Beta-negative binomial process and Poisson factor analysis. In Proceedings of the 15th International Conference on Artificial Intelligence and Statistics, 2012.