The negative binomial Indian buffet process

Creighton Heaukulani
University of Cambridge

Joint work with **Daniel M. Roy** (Cambridge)

BNP 9, Amsterdam
June 2013

Outline

1. The Indian buffet process (IBP; [GG06, GGS07]) induces a distribution on *allocations of features* to individuals. We are interested in **count extensions of the IBP**, i.e., each individual may have *multiple copies* of each feature.

2. ([TJ07]). The IBP is the combinatorial structure of an exchangeable sequence of Bernoulli processes directed by a beta process base measure.

3. We are interested in the combinatorial structure of an exchangeable sequence of negative binomial processes ([BMPJ11, ZHDC12]), directed by a beta process, which we describe as the negative binomial Indian buffet process, a *count* extension of the IBP.

Plan

1. Review the IBP and connect it to the theory of completely random measures.

2. Develop a different continuum of Pólya urn schemes perspective for sampling the IBP.

3. Develop a method to **sample** a beta negative binomial process using the continuum of Pólya urn schemes intuition.

4. Present the corresponding negative binomial IBP.

Indian buffet process

Let $\alpha, c > 0$. A sequence of customers go through an Indian buffet with an infinite number of dishes.

Indian buffet process

Let $\alpha, c > 0$. A sequence of customers go through an Indian buffet with an infinite number of dishes.

- The first customer $n = 1$ tastes Poisson(α) dishes;

Indian buffet process

Let $\alpha, c > 0$. A sequence of customers go through an Indian buffet with an infinite number of dishes.

- The first customer $n = 1$ tastes Poisson(α) dishes;
- The $n + 1$-st customer:

Indian buffet process

Let $\alpha, c > 0$. A sequence of customers go through an Indian buffet with an infinite number of dishes.

- The first customer $n = 1$ tastes Poisson(α) dishes;
- The $n + 1$-st customer:
  - takes Bernoulli$\left(\frac{m_k}{c+n}\right)$ servings of every previously sampled dish k, where m_k is the total number of servings taken of dish k by the previous n customers;

Indian buffet process

Let $\alpha, c > 0$. A sequence of customers go through an Indian buffet with an infinite number of dishes.

- The first customer $n = 1$ tastes $\text{Poisson}(\alpha)$ dishes;
- The $n + 1$-st customer:
  - takes $\text{Bernoulli}\left(\frac{m_k}{c+n}\right)$ servings of every previously sampled dish k, where m_k is the total number of servings taken of dish k by the previous n customers;
  - tastes $\text{Poisson}\left(\alpha \frac{c}{c+n}\right)$ new dishes.

IBP: matrix perspective

rows = individuals/customers
columns = features/dishes

$n = 1$	1	1				
$n = 2$		1	1	1		
$\vdots$	1		1		1	
		1		1	1	1

IBP: matrix perspective

rows = individuals/customers
columns = features/dishes

$n = 1$	1	1				
$n = 2$		1	1	1		
$\vdots$	1		1		1	
		1		1	1	1

[TJ07] connected the IBP with the theory of completely random measures.

IBP: matrix perspective

rows = ~~individuals/customers~~ Bernoulli processes
columns = ~~features/dishes~~ locations in Ω

	ω_1	ω_2	ω_3	ω_4	$\cdots$	
X_1	1	1				
X_2		1	1	1		
X_3	1		1		1	
X_4		1		1	1	1

[TJ07] connected the IBP with the theory of completely random measures.

IBP: matrix perspective

rows = ~~individuals/customers~~ Bernoulli processes
columns = ~~features/dishes~~ locations in Ω

	ω_1	ω_2	ω_3	ω_4	$\cdots$	
X_1	1	1				
X_2		1	1	1		
X_3	1		1		1	
X_4		1		1	1	1

ω_k i.i.d.

[TJ07] connected the IBP with the theory of completely random measures.

IBP: matrix perspective

rows = ~~individuals/customers~~ Bernoulli processes
columns = ~~features/dishes~~ locations in Ω

$$
\begin{array}{c|c|c|c|c|c|c|}
 & \omega_1 & \omega_2 & \omega_3 & \omega_4 & \cdots & \\
\hline
X_1 & 1 & 1 & & & & \\
\hline
X_2 & & 1 & 1 & 1 & & \\
\hline
X_3 & 1 & & 1 & & 1 & \\
\hline
X_4 & & 1 & & 1 & 1 & 1 \\
\hline
\end{array}
$$

ω_k i.i.d.

[TJ07] connected the IBP with the theory of completely random measures.

IBP: matrix perspective

rows = ~~individuals/customers~~ Bernoulli processes
columns = ~~features/dishes~~ locations in Ω

	ω_1	ω_2	ω_3	ω_4	$\cdots$	
X_1	1	1				
X_2		1	1	1		
X_3	1		1		1	
X_4		1		1	1	1

ω_k i.i.d.

$X_2 = \delta_{\omega_2} + \delta_{\omega_3} + \delta_{\omega_4}$

[TJ07] connected the IBP with the theory of completely random measures.

IBP: matrix perspective

rows = ~~individuals/customers~~ Bernoulli processes
columns = ~~features/dishes~~ locations in Ω

$$X_{n+1} \mid X_1, \ldots, X_n \;\sim\; \mathrm{BeP}\Big(\frac{c}{c+n}\widetilde{B}_0 + \frac{1}{c+n}\sum_{i=1}^{n} X_i\Big)$$

	ω_1	ω_2	ω_3	ω_4	$\cdots$	
X_1	1	1				
X_2		1	1	1		
X_3	1		1		1	
X_4		1		1	1	1

ω_k i.i.d.

$X_2 = \delta_{\omega_2} + \delta_{\omega_3} + \delta_{\omega_4}$

[TJ07] connected the IBP with the theory of completely random measures.

IBP: completely random measures perspective

Let $(X_n)_{n\in\mathbb{N}}$ satisfy

$$X_{n+1} \mid X_1, \ldots, X_n \;\sim\; \mathrm{BeP}\Big(\frac{c}{c+n}\widetilde{B}_0 + \frac{1}{c+n}\sum_{i=1}^{n} X_i\Big).$$

IBP: completely random measures perspective

Let $(X_n)_{n\in\mathbb{N}}$ satisfy

$$X_{n+1} \mid X_1, \ldots, X_n \;\sim\; \mathrm{BeP}\Big(\frac{c}{c+n}\widetilde{B}_0 + \frac{1}{c+n}\sum_{i=1}^{n} X_i\Big).$$

Thm ([TJ07]): Then there exists a beta process

$$B \;\sim\; \mathrm{BP}(c, \widetilde{B}_0)$$

such that

$$(X_n)_{n\in\mathbb{N}} \mid B \;\overset{iid}{\sim}\; \mathrm{BeP}(B).$$

How to get counts

$$(X_n)_{n\in\mathbb{N}} \mid B \overset{iid}{\sim} \mathrm{BeP}(B)$$

$n=1$	1	1				
$n=2$		1	1	1		
$\vdots$	1		1		1	
		1		1	1	1

How to get counts

$$(X_n)_{n\in\mathbb{N}} \mid B \overset{iid}{\sim} \mathrm{BeP}(B)$$

Bernoulli(p) $\rightarrow$ geometric(p): # successes before the **first failure**

$n=1$	1	1				
$n=2$		1	1	1		
$\vdots$	1		1		1	
		1		1	1	1

How to get counts

$$(X_n)_{n\in\mathbb{N}} \mid B \overset{iid}{\sim} \mathrm{BeP}(B)$$

Bernoulli(p) $\to$ geometric(p): # successes before the **first failure**

	ω_1	ω_2	ω_3	ω_4	$\cdots$	
X_1	1	1				
X_2		1	1	1		
X_3	1		1		1	
X_4		1		1	1	1

$\delta_{\omega_1} + \ldots$

How to get counts

$$(X_n)_{n\in\mathbb{N}} \mid B \overset{iid}{\sim} \mathrm{BeP}(B)$$

Bernoulli(p) $\to$ geometric(p): # successes before the **first failure**

	ω_1	ω_2	ω_3	ω_4	$\cdots$	
X_1	1	1				
X_2		1	1	1		
X_3	1		1		1	
X_4		1		1	1	1

$$\delta_{\omega_1} + 2\,\delta_{\omega_2} + \ldots$$

How to get counts

$$(X_n)_{n\in\mathbb{N}} \mid B \overset{iid}{\sim} \mathrm{BeP}(B)$$

Bernoulli(p) $\to$ geometric(p): # successes before the **first failure**

	ω_1	ω_2	ω_3	ω_4	$\cdots$	
X_1	1	1				
X_2		1	1	1		
X_3	1		1		1	
X_4		1		1	1	1

$$\delta_{\omega_1} + 2\,\delta_{\omega_2}$$

How to get counts

$$(X_n)_{n\in\mathbb{N}} \mid B \overset{iid}{\sim} \mathrm{BeP}(B)$$

Bernoulli(p) $\rightarrow$ geometric(p): # successes before the **first failure**

	ω_1	ω_2	ω_3	ω_4	$\cdots$	
X_1	1	1				
X_2		1	1	1		
X_3	1		1		1	
X_4		1		1	1	1

$$\delta_{\omega_1} + 2\,\delta_{\omega_2} \;\sim\; \mathrm{BGP}(\widetilde{B}_0)$$

How to get counts

$$(X_n)_{n\in\mathbb{N}} \mid B \overset{iid}{\sim} \mathrm{BeP}(B)$$

Bernoulli(p) $\to$ geometric(p): # successes before the **first failure**

	ω_1	ω_2	ω_3	ω_4	$\cdots$	
X_1	1	1				
X_2		1	1	1		
X_3	1		1		1	
X_4		1		1	1	1

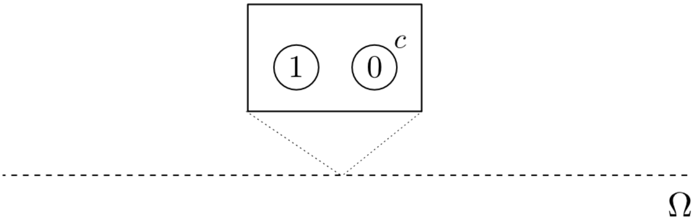
$$\delta_{\omega_1} + 2\,\delta_{\omega_2} \;\sim\; \mathrm{BGP}(\widetilde{B}_0)$$

NB(r, p): # successes before the r-th failure

Urn scheme perspective

Urn scheme perspective

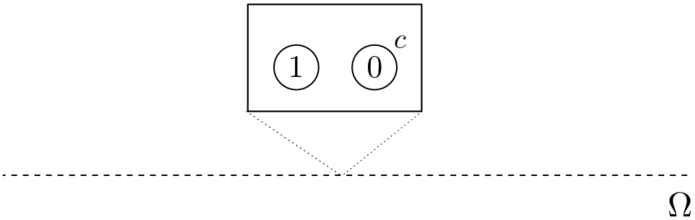
Imagine an infinite number of independent Pólya urn schemes, each with two tables labelled 1 and 0



Urn scheme perspective

Imagine an infinite number of independent Pólya urn schemes, each with two tables labelled 1 and 0

i.e., a **continuum of Pólya urn schemes**



Urn scheme perspective

How to sample $X_1 \sim \mathrm{BeP}(\widetilde{B}_0)$

Urn scheme perspective

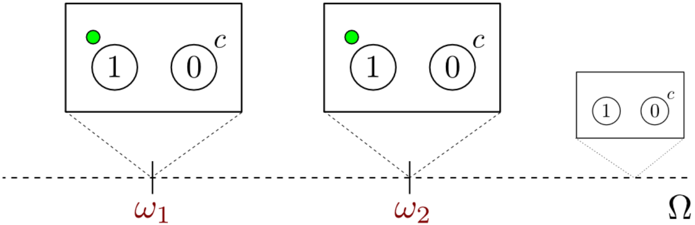
How to sample $X_1 \sim \text{BeP}(\widetilde{B}_0)$

- A Poisson(α) number of urn schemes succeed.

Urn scheme perspective

How to sample $X_1 \sim \mathrm{BeP}(\widetilde{B}_0)$

- A Poisson(α) number of urn schemes succeed.

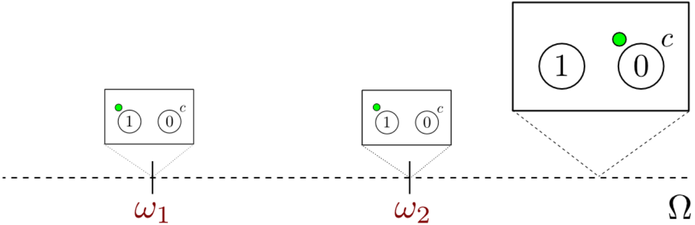


$\widetilde{B}_0(\Omega) = \alpha$

Urn scheme perspective

How to sample $X_1 \sim \mathrm{BeP}(\widetilde{B}_0)$

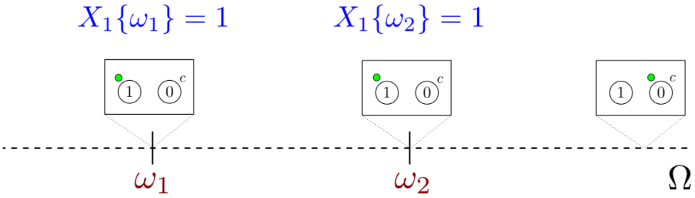
- A Poisson(α) number of urn schemes succeed.
- The remaining urn schemes on the continuum fail.



Urn scheme perspective

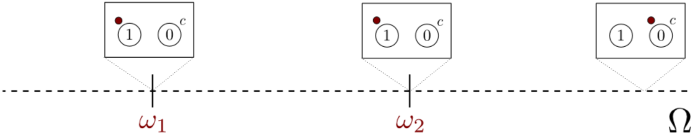
How to sample $X_1 \sim \mathrm{BeP}(\widetilde{B}_0)$

- A Poisson(α) number of urn schemes succeed.
- The remaining urn schemes on the continuum fail.



X_1 done.

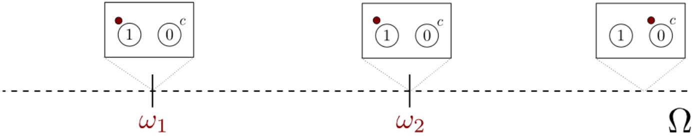
How to sample $X_2 \mid X_1$



$$X_2 \mid X_1 \;\sim\; \mathrm{BeP}\Big(\tfrac{c}{c+1}\widetilde{B}_0 + \tfrac{1}{c+1}X_1\Big)$$

How to sample $X_2 \mid X_1$

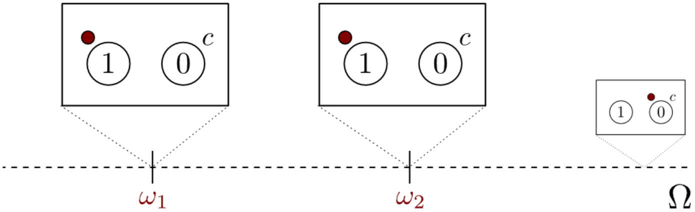
Run the urn schemes forward one more step



$$X_2 \mid X_1 \;\sim\; \mathrm{BeP}\left(\frac{c}{c+1}\widetilde{B}_0 + \frac{1}{c+1}X_1\right)$$

How to sample $X_2 \mid X_1$

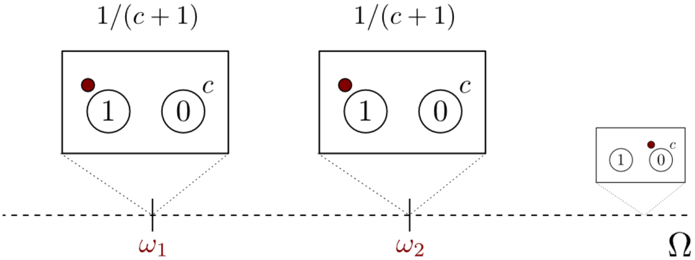
Run the urn schemes forward one more step
$\Rightarrow$ Consider the previous atoms in X_1



$$X_2 \mid X_1 \sim \mathrm{BeP}\left(\frac{c}{c+1}\widetilde{B}_0 + \frac{1}{c+1}X_1\right)$$

How to sample $X_2 \mid X_1$

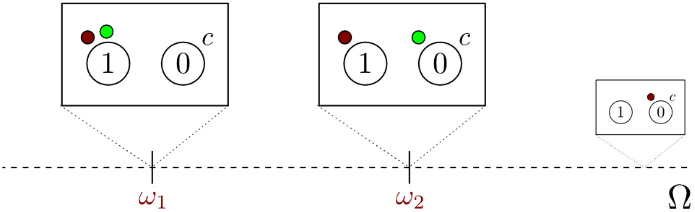
Run the urn schemes forward one more step
$\Rightarrow$ Consider the previous atoms in X_1



$$X_2 \mid X_1 \sim \mathrm{BeP}\left(\frac{c}{c+1}\widetilde{B}_0 + \frac{1}{c+1}X_1\right)$$

How to sample $X_2 \mid X_1$

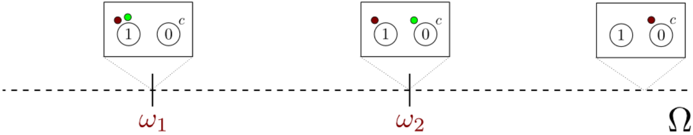
Run the urn schemes forward one more step
$\Rightarrow$ Consider the previous atoms in X_1



$$X_2 \mid X_1 \sim \mathrm{BeP}\left(\frac{c}{c+1}\widetilde{B}_0 + \frac{1}{c+1}X_1\right)$$

How to sample $X_2 \mid X_1$

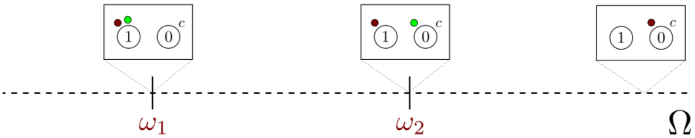
Run the urn schemes forward one more step
$\Rightarrow$ Consider the previous atoms in X_1



$$X_2 \mid X_1 \sim \mathrm{BeP}\left(\frac{c}{c+1}\widetilde{B}_0 + \frac{1}{c+1}X_1\right)$$

How to sample $X_2 \mid X_1$

Run the urn schemes forward one more step
$\Rightarrow$ Consider the previous atoms in X_1
$\Rightarrow$ Consider urns **not in** X_1



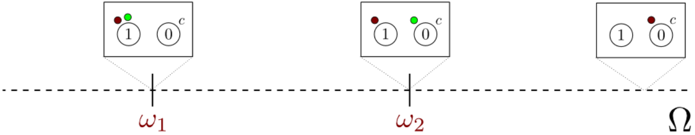
$$X_2 \mid X_1 \;\sim\; \mathrm{BeP}\Big(\tfrac{c}{c+1}\widetilde{B}_0 + \tfrac{1}{c+1}X_1\Big)$$

How to sample $X_2 \mid X_1$

Run the urn schemes forward one more step

$\Rightarrow$ Consider the previous atoms in X_1

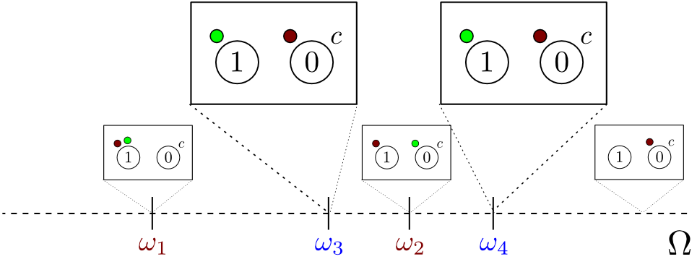
$\Rightarrow$ Consider urns **not in** X_1; a Poisson$\left(\alpha\frac{c}{c+1}\right)$ new urns succeed



$$X_2 \mid X_1 \ \sim\ \mathrm{BeP}\left(\tfrac{c}{c+1}\widetilde{B}_0 + \tfrac{1}{c+1}X_1\right)$$

How to sample $X_2 \mid X_1$

Run the urn schemes forward one more step
$\Rightarrow$ Consider the previous atoms in X_1
$\Rightarrow$ Consider urns **not in** X_1; a Poisson$\left(\alpha \frac{c}{c+1}\right)$ new urns succeed



$$X_2 \mid X_1 \;\sim\; \mathrm{BeP}\left(\frac{c}{c+1}\widetilde{B}_0 + \frac{1}{c+1}X_1\right)$$

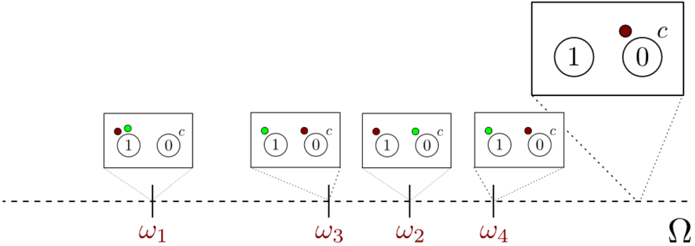
How to sample $X_2 \mid X_1$

Run the urn schemes forward one more step

$\Rightarrow$ Consider the previous atoms in X_1

$\Rightarrow$ Consider urns **not in** X_1; a Poisson$\left(\alpha \frac{c}{c+1}\right)$ new urns succeed

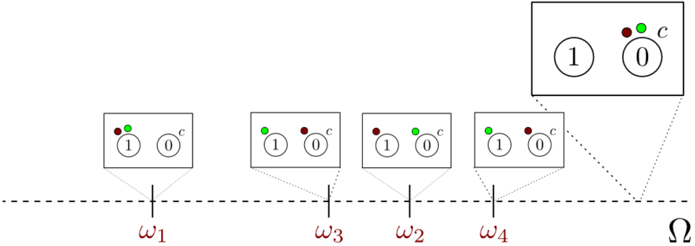
$\Rightarrow$ The remaining urns on the continuum fail



$$X_2 \mid X_1 \;\sim\; \mathrm{BeP}\left(\tfrac{c}{c+1}\widetilde{B}_0 + \tfrac{1}{c+1} X_1\right)$$

How to sample $X_2 \mid X_1$

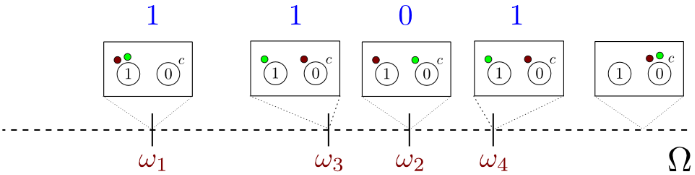
Run the urn schemes forward one more step
$\Rightarrow$ Consider the previous atoms in X_1
$\Rightarrow$ Consider urns **not in** X_1; a $\text{Poisson}\left(\alpha \frac{c}{c+1}\right)$ new urns succeed
$\Rightarrow$ The remaining urns on the continuum fail



$$X_2 \mid X_1 \;\sim\; \text{BeP}\left(\frac{c}{c+1}\widetilde{B}_0 + \frac{1}{c+1}X_1\right)$$

How to sample $X_2 \mid X_1$

X_2 done.



$$X_2 \mid X_1 \;\sim\; \mathrm{BeP}\left(\frac{c}{c+1}\widetilde{B}_0 + \frac{1}{c+1}X_1\right)$$

A count extension

Bernoulli(p): distribution of successful trial with success probability p;

NB(r, p): distribution of # successes until r failures.

A count extension

Bernoulli(p): distribution of successful trial with success probability p;

NB(r, p): distribution of # successes until r failures.

Intuition: The Bernoulli process runs one trial at every urn scheme. So the *negative binomial process* continues running trials until r failures at each urn scheme.

Key point: Instead of binary indicators at each urn scheme, we get integer-valued counts.

Negative binomial process

Negative binomial process

Negative binomial processes are parameterized by a “base measure”

$$B_0 = \widetilde{B}_0 + \sum_{k=1}^{\kappa} b_k \delta_{\omega_k}. \tag{1}$$

Negative binomial process

Negative binomial processes are parameterized by a "base measure"

$$B_0 = \widetilde{B}_0 + \sum_{k=1}^{\kappa} b_k \delta_{\omega_k}. \tag{1}$$

Defn. We call $X \sim \mathrm{NBP}(r, B_0)$ a <u>negative binomial process</u>, when it is a completely random measure with

- fixed component (as defined by [BMPJ11, ZHDC12]):

$$\sum_{k=1}^{\kappa} \zeta_k \delta_{\omega_k}, \qquad \zeta_k \overset{ind}{\sim} \mathrm{NB}(r, b_k), \quad \kappa \in \mathbb{N} \cup \{\infty\}, \tag{2}$$

Negative binomial process

Negative binomial processes are parameterized by a "base measure"

$$B_0 = \widetilde{B}_0 + \sum_{k=1}^{\kappa} b_k \delta_{\omega_k}. \qquad (1)$$

Defn. We call $X \sim \mathrm{NBP}(r, B_0)$ a <u>negative binomial process</u>, when it is a completely random measure with

- fixed component (as defined by [BMPJ11, ZHDC12]):

$$\sum_{k=1}^{\kappa} \zeta_k \delta_{\omega_k}, \qquad \zeta_k \overset{ind}{\sim} \mathrm{NB}(r, b_k), \quad \kappa \in \mathbb{N} \cup \{\infty\}, \qquad (2)$$

- ordinary component (our definition): a Poisson process with intensity $r\widetilde{B}_0$.

Beta negative binomial process

Beta negative binomial process

Defn. We call $X \sim \text{BNBP}(r, c, B_0)$ a beta negative binomial process when there exists a beta process

$$B \sim \text{BP}(c, B_0)$$

such that

$$X \mid B \sim \text{NBP}(r, B). \tag{3}$$

Beta negative binomial process

Defn. We call $X \sim \mathrm{BNBP}(r, c, B_0)$ a beta negative binomial process when there exists a beta process

$$B \sim \mathrm{BP}(c, B_0)$$

such that

$$X \mid B \sim \mathrm{NBP}(r, B). \tag{3}$$

- The beta process has an infinite number of atoms w.p. one

Beta negative binomial process

Defn. We call $X \sim \mathrm{BNBP}(r, c, B_0)$ a beta negative binomial process when there exists a beta process

$$B \sim \mathrm{BP}(c, B_0)$$

such that

$$X \mid B \sim \mathrm{NBP}(r, B). \tag{3}$$

- The beta process has an infinite number of atoms w.p. one
- [BMPJ11, ZHDC12] use finite approximations to the beta process in order to produce samples of X

Beta negative binomial process

Defn. We call $X \sim \mathrm{BNBP}(r, c, B_0)$ a beta negative binomial process when there exists a beta process

$$B \sim \mathrm{BP}(c, B_0)$$

such that

$$X \mid B \sim \mathrm{NBP}(r, B). \tag{3}$$

- The beta process has an infinite number of atoms w.p. one
- [BMPJ11, ZHDC12] use finite approximations to the beta process in order to produce samples of X

Our goal: give a method to directly sample X without representing the underlying beta process.

Start simple

Consider a $\mathrm{NB}(r, p)$ distribution when $r = 1$, i.e., a geometric distribution

$$\mathrm{NB}(1, p) = \mathrm{geometric}(p), \tag{4}$$

which counts the number of successful trials (with success probability p) before the **first failure**.

Start simple

Consider a $\mathrm{NB}(r, p)$ distribution when $r = 1$, i.e., a geometric distribution

$$\mathrm{NB}(1, p) = \mathrm{geometric}(p), \tag{4}$$

which counts the number of successful trials (with success probability p) before the **first failure**.

Intuition: Run the urn scheme until the first failure.

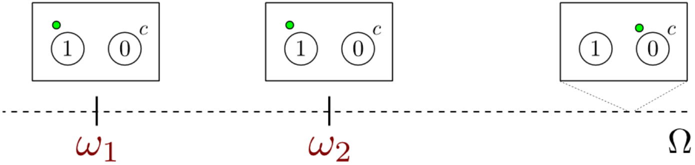
How to sample $X \sim \mathrm{BGP}(c, \widetilde{B}_0)$

How to sample $X \sim \mathrm{BGP}(c, \widetilde{B}_0)$

Run all urn schemes at once

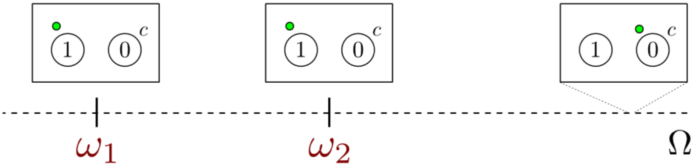
How to sample $X \sim \mathrm{BGP}(c, \widetilde{B}_0)$

Run all urn schemes at once



How to sample $X \sim \mathrm{BGP}(c, \widetilde{B}_0)$

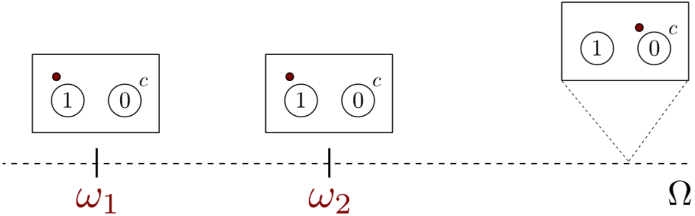
We continue until all urn schemes have failed once!



How to sample $X \sim \mathrm{BGP}(c, \widetilde{B}_0)$

We continue until all urn schemes have failed once!

$\Rightarrow$ The remaining urns on the continuum have already failed

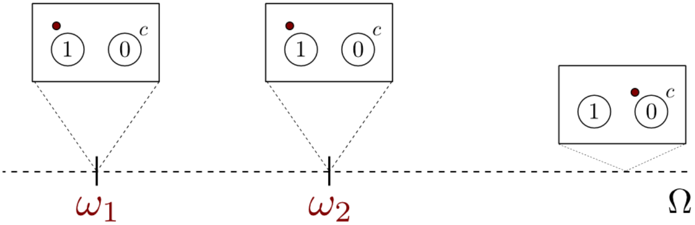


How to sample $X \sim \mathrm{BGP}(c, \widetilde{B}_0)$

We continue until all urn schemes have failed once!

$\Rightarrow$ The remaining urns on the continuum have already failed

$\Rightarrow$ We **complete** the atoms

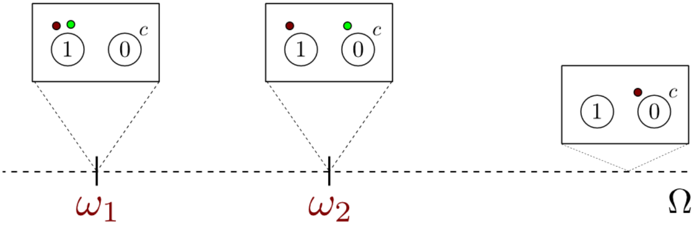


How to sample $X \sim \mathrm{BGP}(c, \widetilde{B}_0)$

We continue until all urn schemes have failed once!

⇒ The remaining urns on the continuum have already failed

⇒ We **complete** the atoms

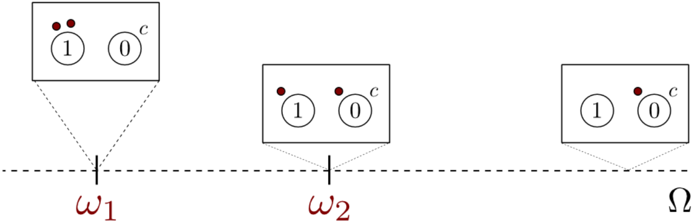


How to sample $X \sim \mathrm{BGP}(c, \widetilde{B}_0)$

We continue until all urn schemes have failed once!

$\Rightarrow$ The remaining urns on the continuum have already failed

$\Rightarrow$ We **complete** the atoms

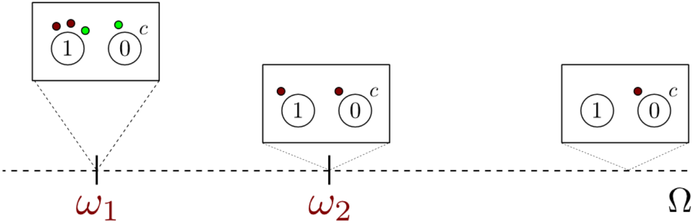


How to sample $X \sim \mathrm{BGP}(c, \widetilde{B}_0)$

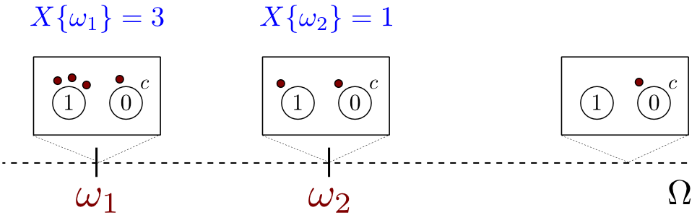
We continue until all urn schemes have failed once!

$\Rightarrow$ The remaining urns on the continuum have already failed

$\Rightarrow$ We **complete** the atoms



How to sample $X \sim \mathrm{BGP}(c, \widetilde{B}_0)$



$$X \sim \mathrm{BGP}(c, \widetilde{B}_0)$$

Ordinary component: BGP

Thm ([HR13]). Let $Y \sim \mathrm{PP}(\widetilde{B}_0)$ and let

$$\zeta_s \overset{ind}{\sim} \text{beta-geometric}(1, c), \quad s \in \Omega,$$

be independent from Y. Then

$$X = \sum_{s \in Y} (1 + \zeta_s)\delta_s \sim \mathrm{BGP}(c, \widetilde{B}_0).$$

Ordinary component: BGP

Thm ([HR13]). Let $Y \sim \mathrm{PP}(\widetilde{B}_0)$ and let

$$\zeta_s \overset{ind}{\sim} \underbrace{\text{beta-geometric}(1, c),}_{\textbf{completing the atom}} \quad s \in \Omega,$$

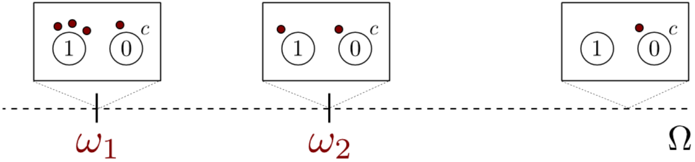
be independent from Y. Then

$$X = \sum_{s \in Y} (1 + \zeta_s)\delta_s \sim \mathrm{BGP}(c, \widetilde{B}_0).$$

How to sample $X \sim \mathrm{BNBP}(r, c, \widetilde{B}_0)$

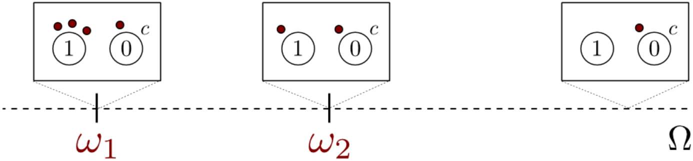
How to sample $X \sim \mathrm{BNBP}(r, c, \widetilde{B}_0)$

Start where $X \sim \mathrm{BGP}(c, \widetilde{B}_0)$ left off



How to sample $X \sim \mathrm{BNBP}(r, c, \widetilde{B}_0)$

Start where $X \sim \mathrm{BGP}(c, \widetilde{B}_0)$ left off

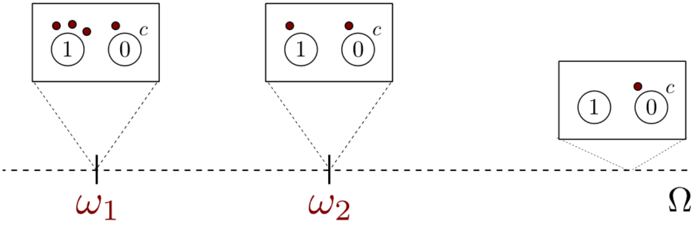


Continue running the urn schemes untill r failures!

How to sample $X \sim \mathrm{BNBP}(r, c, \widetilde{B}_0)$

Start where $X \sim \mathrm{BGP}(c, \widetilde{B}_0)$ left off

$\Rightarrow$ complete the atoms



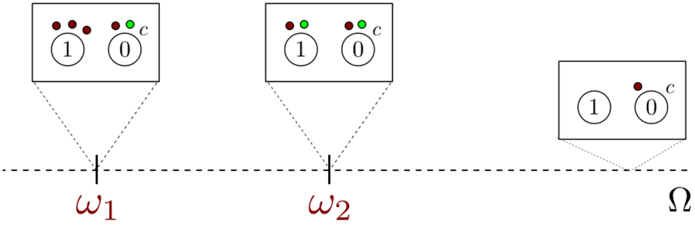
Continue running the urn schemes untill r failures!

How to sample $X \sim \mathrm{BNBP}(r, c, \widetilde{B}_0)$

Start where $X \sim \mathrm{BGP}(c, \widetilde{B}_0)$ left off

$\Rightarrow$ complete the atoms

E.g., if $r = 2$, we stop.

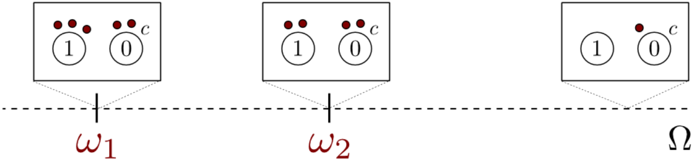


Continue running the urn schemes untill r failures!

How to sample $X \sim \mathrm{BNBP}(r, c, \widetilde{B}_0)$

Start where $X \sim \mathrm{BGP}(c, \widetilde{B}_0)$ left off

$\Rightarrow$ complete the atoms

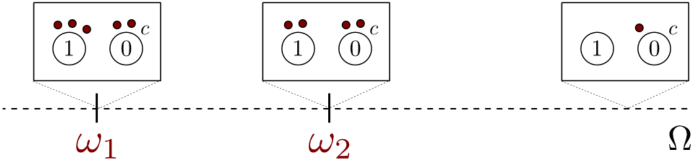


How to sample $X \sim \text{BNBP}(r, c, \widetilde{B}_0)$

Start where $X \sim \text{BGP}(c, \widetilde{B}_0)$ left off

$\Rightarrow$ complete the atoms

$\Rightarrow$ advance the remaining urns on the continuum

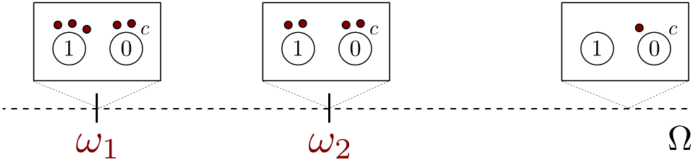


How to sample $X \sim \mathrm{BNBP}(r, c, \widetilde{B}_0)$

Start where $X \sim \mathrm{BGP}(c, \widetilde{B}_0)$ left off

$\Rightarrow$ complete the atoms

$\Rightarrow$ advance the remaining urns on the continuum



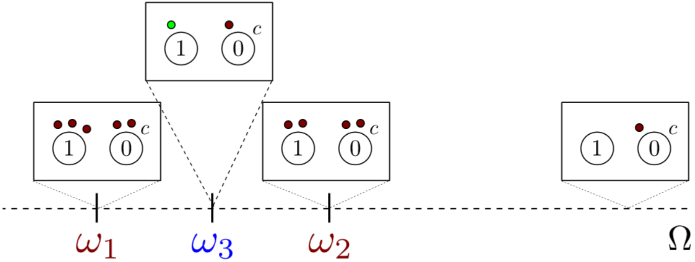
- A $\mathrm{Poisson}\left(\alpha \frac{c}{c+1}\right)$ number of new urns succeed

How to sample $X \sim \mathrm{BNBP}(r, c, \widetilde{B}_0)$

Start where $X \sim \mathrm{BGP}(c, \widetilde{B}_0)$ left off

$\Rightarrow$ complete the atoms

$\Rightarrow$ advance the remaining urns on the continuum



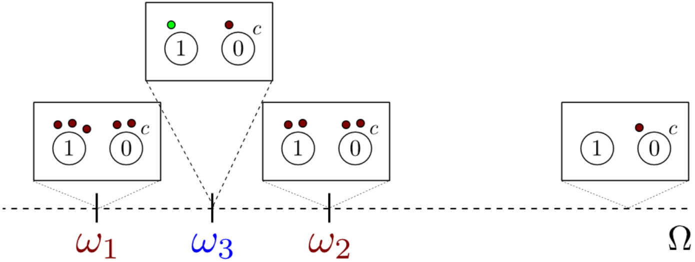
- A $\mathrm{Poisson}\left(\alpha \frac{c}{c+1}\right)$ number of new urns succeed

How to sample $X \sim \mathrm{BNBP}(r, c, \widetilde{B}_0)$

Start where $X \sim \mathrm{BGP}(c, \widetilde{B}_0)$ left off

$\Rightarrow$ complete the atoms

$\Rightarrow$ advance the remaining urns on the continuum



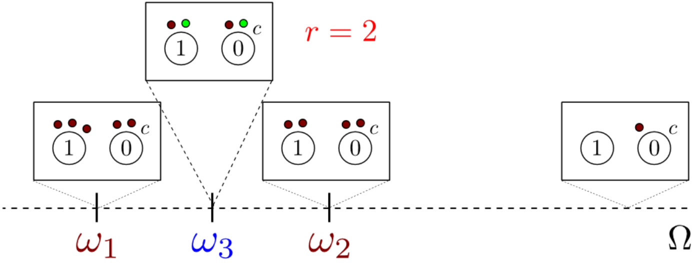
- A $\mathrm{Poisson}\left(\alpha \frac{c}{c+1}\right)$ number of new urns succeed
- **Complete** the new atom

How to sample $X \sim \text{BNBP}(r, c, \widetilde{B}_0)$

Start where $X \sim \text{BGP}(c, \widetilde{B}_0)$ left off

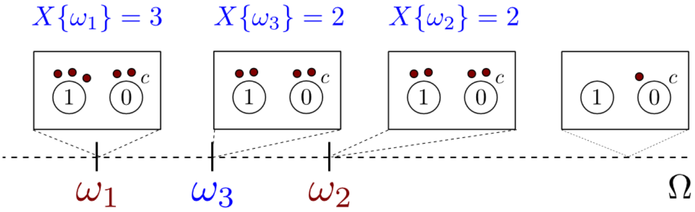
$\Rightarrow$ complete the atoms

$\Rightarrow$ advance the remaining urns on the continuum



- A $\text{Poisson}\left(\alpha \frac{c}{c+1}\right)$ number of new urns succeed
- **Complete** the new atom

How to sample $X \sim \text{BNBP}(r, c, \widetilde{B}_0)$



$X \sim \text{BNBP}(r, c, \widetilde{B}_0)$

Ordinary component: BNBP

Thm ([HR13]). Let

$$Y_\ell \overset{ind}{\sim} \mathrm{PP}\Big(\frac{c}{c+\ell-1}\widetilde{B}_0\Big), \qquad \ell \in [r] = (1,\dots,r), \tag{5}$$

and let

$$\zeta_{\ell,s} \overset{ind}{\sim} \text{beta-NB}(r-\ell+1, 1, c+\ell-1), \qquad \ell \in [r],\ s \in \Omega, \tag{6}$$

independent also from (Y_r). Then

$$X = \sum_{\ell=1}^{r} \sum_{s \in Y_\ell} (1+\zeta_{\ell,s})\delta_s \ \sim\ \mathrm{BNBP}(r, c, \widetilde{B}_0). \tag{7}$$

Ordinary component: BNBP

Thm ([HR13]). Let

$$Y_\ell \overset{ind}{\sim} \mathrm{PP}\Big(\frac{c}{c+\ell-1}\widetilde{B}_0\Big), \qquad \ell \in [r] = (1,\dots,r), \tag{5}$$

and let

$$\zeta_{\ell,s} \overset{ind}{\sim} \text{beta-NB}(r-\ell+1, 1, c+\ell-1), \qquad \ell \in [r],\ s \in \Omega, \tag{6}$$

independent also from (Y_r). Then

$$X = \sum_{\ell=1}^{r} \sum_{s \in Y_\ell} (1+\zeta_{\ell,s})\delta_s \ \sim\ \mathrm{BNBP}(r, c, \widetilde{B}_0). \tag{7}$$

Key point: atoms arise in a sequence of r rounds.

A negative binomial IBP

Each customer $n \in \mathbb{N}$ goes through $r \in \mathbb{N}$ rounds at the buffet.

A negative binomial IBP

Each customer $n \in \mathbb{N}$ goes through $r \in \mathbb{N}$ rounds at the buffet.

- $n = 1$: Independently for each of r rounds $\ell = 1, \ldots, r$,
  - tastes $\mathrm{Poisson}\left(\alpha \frac{c}{c+\ell-1}\right)$ dishes;

A negative binomial IBP

Each customer $n \in \mathbb{N}$ goes through $r \in \mathbb{N}$ rounds at the buffet.

- $n = 1$: Independently for each of r rounds $\ell = 1, \ldots, r$,
  - tastes $\text{Poisson}\left(\alpha \frac{c}{c+\ell-1}\right)$ dishes;
  - returns to each dish for $\text{beta-NB}(r - \ell + 1, 1, c + \ell - 1)$ additional servings;

A negative binomial IBP

Each customer $n \in \mathbb{N}$ goes through $r \in \mathbb{N}$ rounds at the buffet.

- $n = 1$: Independently for each of r rounds $\ell = 1, \ldots, r$,
  - tastes Poisson$\left(\alpha \frac{c}{c+\ell-1}\right)$ dishes;
  - returns to each dish for beta-NB$(r - \ell + 1, 1, c + \ell - 1)$ additional servings;
- subsequent customers $n \geq 2$:
  - takes beta-NB$(r, m_k, c + (n-1)r)$ servings of each previously tasted dish k, where m_k is the total number of servings taken by first n customers;

A negative binomial IBP

Each customer $n \in \mathbb{N}$ goes through $r \in \mathbb{N}$ rounds at the buffet.

- $n = 1$: Independently for each of r rounds $\ell = 1, \ldots, r$,
  - tastes $\mathrm{Poisson}\big(\alpha \, \frac{c}{c+\ell-1}\big)$ dishes;
  - returns to each dish for $\text{beta-NB}(r-\ell+1, 1, c+\ell-1)$ additional servings;
- subsequent customers $n \geq 2$:
  - takes $\text{beta-NB}(r, m_k, c+(n-1)r)$ servings of each previously tasted dish k, where m_k is the total number of servings taken by first n customers;
  - independently for each of r rounds $\ell = 1, \ldots, r$,
    - tries $\mathrm{Poisson}\left(\alpha \, \frac{c}{c+(n-1)r+\ell-1}\right)$ new dishes;

A negative binomial IBP

Each customer $n \in \mathbb{N}$ goes through $r \in \mathbb{N}$ rounds at the buffet.

- $n = 1$: Independently for each of r rounds $\ell = 1, \ldots, r$,
  - tastes $\text{Poisson}\big(\alpha \frac{c}{c+\ell-1}\big)$ dishes;
  - returns to each dish for $\text{beta-NB}(r - \ell + 1, 1, c + \ell - 1)$ additional servings;
- subsequent customers $n \geq 2$:
  - takes $\text{beta-NB}(r, m_k, c + (n-1)r)$ servings of each previously tasted dish k, where m_k is the total number of servings taken by first n customers;
  - independently for each of r rounds $\ell = 1, \ldots, r$,
    - tries $\text{Poisson}\left(\alpha \frac{c}{c+(n-1)r+\ell-1}\right)$ new dishes;
    - returns to each new dish for $\text{beta-NB}(r - \ell + 1, 1, c + (n-1)r + \ell - 1)$ additional servings.

Probability function

Let $\mathcal{H}_n \equiv \mathbb{Z}_+^n \backslash \{0^n\}$, and for $h \in \mathcal{H}_n$, let M_h count the number of features k where every customer n has $h(n)$ copies of feature k.

Probability function

Let $\mathcal{H}_n \equiv \mathbb{Z}_+^n \backslash \{0^n\}$, and for $h \in \mathcal{H}_n$, let M_h count the number of features k where every customer n has $h(n)$ copies of feature k.

Claim. The probability function for the NB-IBP is given by

$$\frac{\alpha^{K_N}}{\prod_{h\in\mathcal{H}_N} M_h!} \exp\left(-\alpha \sum_{n=1}^{N}\sum_{\ell=1}^{r} \frac{c}{c+nr+\ell-1}\right) \prod_{h\in\mathcal{H}_N}\left[\frac{c}{c+Nr}\frac{\Gamma(s(h))\Gamma(c+Nr+1)}{\Gamma(c+Nr+s(h))} \prod_{n\in\mathbb{N}}\binom{r+h(n)-1}{r-1}\right]^{M_h}$$

where $s(h) \equiv \sum_{n=1}^{N} h(n)$.

General $r > 0$

- The previous urn scheme was only valid for $r \in \mathbb{N}$.
- Urn schemes for general $r > 0$ reduces to the case of **fractional** $r \in (0, 1)$, which is desirable for some applications.
  - We use Poisson process calculus to reduce the r rounds to one slightly more clever round.
  - Completions are no longer beta-NB variables, but what we call harmonic mixtures.

General $r > 0$

There is an analytical extension of the NB-IBP to general values of $r > 0$. The probability function looks similar:

Claim ([HR13]). Let $r > 0$. The probability function for the NB-IBP is given by

$$\begin{aligned} &\frac{(c\,\alpha)^{K_N}}{\prod_{h \in \mathcal{H}_N} M_h!} \exp\Big(-c\,\alpha \sum_{n=1}^{N} [\psi(c+(n+1)r) - \psi(c+nr)]\Big) \\ &\qquad \times \prod_{h \in \mathcal{H}_N} \left[\frac{\Gamma(s(h))\Gamma(c+Nr)}{\Gamma(c+Nr+s(h))} \prod_{n \in \mathbb{N}} \frac{\Gamma(r+h(n))}{h(n)!\,\Gamma(r)}\right]^{M_h}. \end{aligned} \tag{8}$$

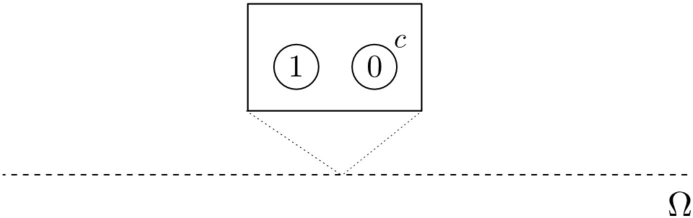
Key difference: No longer a concept of *rounds*.

Extensions

We can also characterize the combinatorial structure of an exchangeable sequence of negative binomial processes, directed by a generalized beta process [Roy13], parametrized by a measurable family of EPPFs.

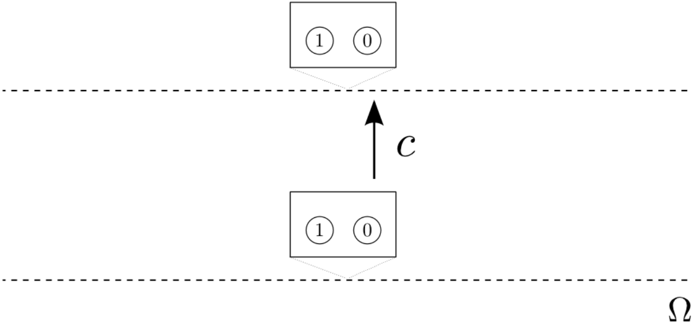
Extensions

We can also characterize the combinatorial structure of an exchangeable sequence of negative binomial processes, directed by a generalized beta process [Roy13], parametrized by a measurable family of EPPFs.



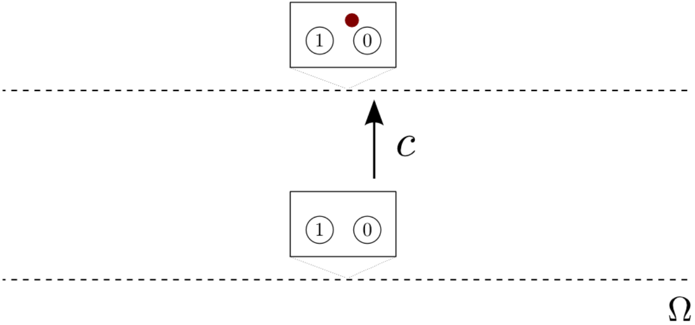
Extensions

We can also characterize the combinatorial structure of an exchangeable sequence of negative binomial processes, directed by a generalized beta process [Roy13], parametrized by a measurable family of EPPFs.



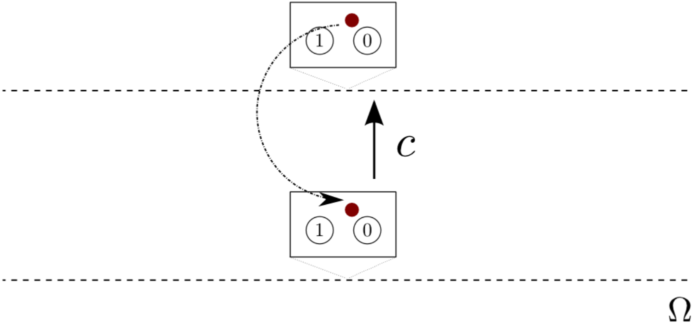
Extensions

We can also characterize the combinatorial structure of an exchangeable sequence of negative binomial processes, directed by a generalized beta process [Roy13], parametrized by a measurable family of EPPFs.



Extensions

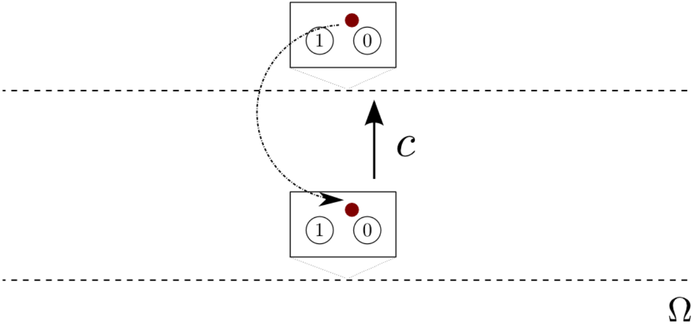
We can also characterize the combinatorial structure of an exchangeable sequence of negative binomial processes, directed by a generalized beta process [Roy13], parametrized by a measurable family of EPPFs.



Extensions

We can also characterize the combinatorial structure of an exchangeable sequence of negative binomial processes, directed by a generalized beta process [Roy13], parametrized by a measurable family of EPPFs.

A continuum of generalized Blackwell-MacQueen urn schemes



Extensions

Special cases include:

- hierarchies of beta processes [TJ07]
- stable beta processes [TG09, BJP12]
  - corresponding power-law NB-IBP
- hierarchies of stable beta processes

Probability function

Claim. The probability function for the NB-IBP is given by

$$\frac{\alpha^{K_N}}{\prod_{h\in\mathcal{H}_N} M_h!}\exp\left(-\alpha\sum_{n=1}^{N}\sum_{\ell=1}^{r}\frac{c}{c+nr+\ell-1}\right)$$
$$\prod_{h\in\mathcal{H}_N}\left[\frac{c}{c+Nr}\frac{\Gamma(s(h))\Gamma(c+Nr+1)}{\Gamma(c+Nr+s(h))}\right.$$
$$\left.\prod_{n\in\mathbb{N}}\binom{r+h(n)-1}{r-1}\right]^{M_h}$$

Probability function from a generalized beta

Claim. The probability function for the generalized NB-IBP is given by

$$\frac{\alpha^{K_N}}{\prod_{h\in\mathcal{H}_N} M_h!}\exp\left(-\alpha\sum_{n=1}^{N}\sum_{\ell=1}^{r}\mathbb{P}(K_{nr+\ell} > K_{nr+\ell-1})\right)$$
$$\prod_{h\in\mathcal{H}_N}\Bigg[\mathbb{P}(K_{Nr+1} > K_{Nr}, Z_{Nr+s(h)} = \cdots = Z_{Nr+1})$$
$$\prod_{n\in\mathbb{N}}\binom{r+h(n)-1}{r-1}\Bigg]^{M_h}$$

Probability function from a generalized beta

Claim. The probability function for the generalized NB-IBP is given by

$$\frac{\alpha^{K_N}}{\prod_{h\in\mathcal{H}_N} M_h!}\exp\Bigg(-\alpha\sum_{n=1}^{N}\sum_{\ell=1}^{r}\mathbb{P}(K_{nr+\ell}>K_{nr+\ell-1})\Bigg)$$

$$\prod_{h\in\mathcal{H}_N}\Bigg[\mathbb{P}(K_{Nr+1}>K_{Nr}, Z_{Nr+s(h)}=\cdots=Z_{Nr+1})$$

$$\prod_{n\in\mathbb{N}}\binom{r+h(n)-1}{r-1}\Bigg]^{M_h}$$

$\mathbb{P}(K_{n+1}>K_n)$: probability of a new table at step $n+1$ ...
$\mathbb{P}(K_{n+1}>K_n, Z_{n+m}=\cdots=Z_{n+1})$: probability of also staying there for next m steps

T. Broderick, M. I. Jordan, and J. Pitman.
Beta processes, stick-breaking, and power laws.
Bayesian Analysis, 7(1):1–38, 2012.

T. Broderick, L. Mackey, J. Paisley, and M. Jordan.
Combinatorial clustering and the beta-negative binomial process, 2011, 1111.1802v2.

T. L. Griffiths and Z. Ghahramani.
Infinite latent feature models and the Indian buffet process.
In Adv. in Neural Inform. Processing Syst. 18, pages 475–482. MIT Press, Cambridge, MA, 2006.

Z. Ghahramani, T. Griffiths, and P. Sollich.
Bayesian nonparametric latent feature models.
Bayesian Statistics, 8:201–226, 2007.
See also the discussion and rejoinder.

C. Heaukulani and D. M. Roy.
Urn schemes for the beta negative binomial process and the negative binomial Indian buffet process, 2013, Preprint.

D. M. Roy.
The combinatorial structure underlying the beta process is described by a continuum of blackwell-macqueen urn schemes, 2013, Preprint.

Y. W. Teh and D. Görür.
Indian buffet processes with power-law behavior.
In Adv. in Neural Inform. Processing Syst. 23. MIT Press, Cambridge, MA, 2009.

R. Thibaux and M. I. Jordan.
Hierarchical beta processes and the Indian buffet process.
In Proc. of the 11th Conf. on A.I. and Stat., 2007.

M. Zhou, L. Hannah, D. Dunson, and L. Carin.
Beta-negative binomial process and Poisson factor analysis.
In Proc. AISTATS, 2012.