Outline

I. Clusters

- Overview
- Distribution
- Proportions
- Random probability measure

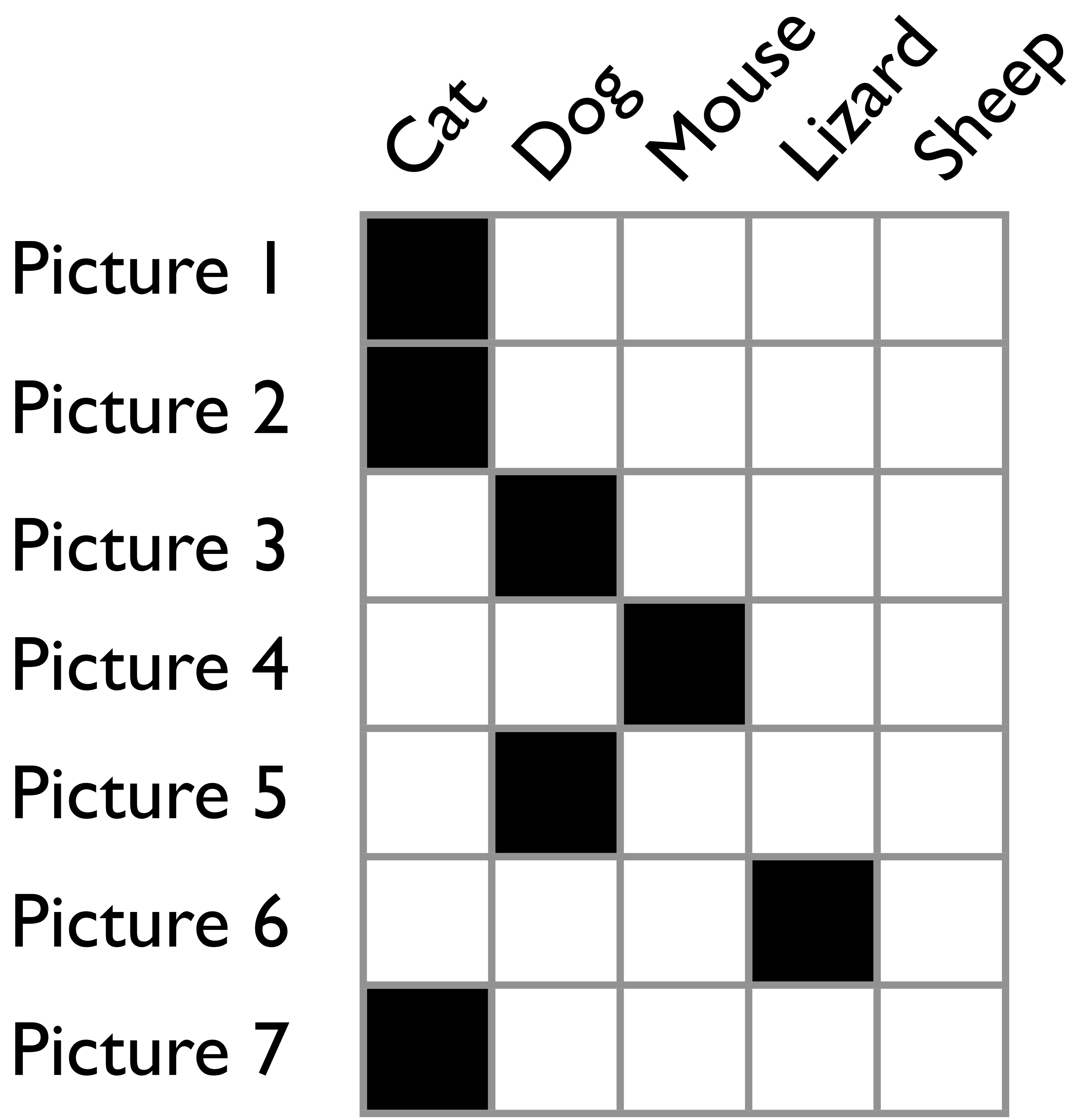
II. Features

Outline

I. Clusters

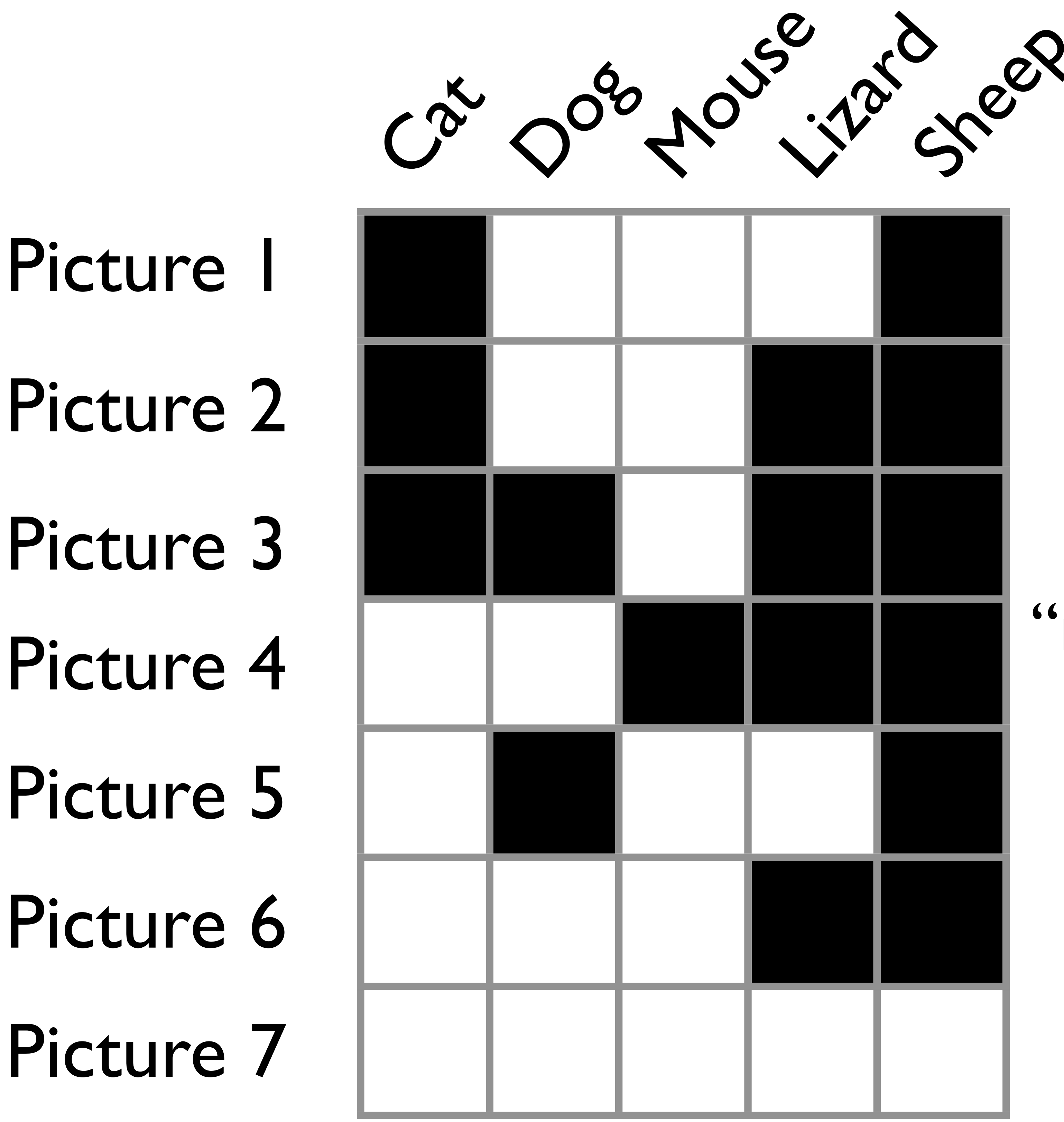
II. Features

Clustering



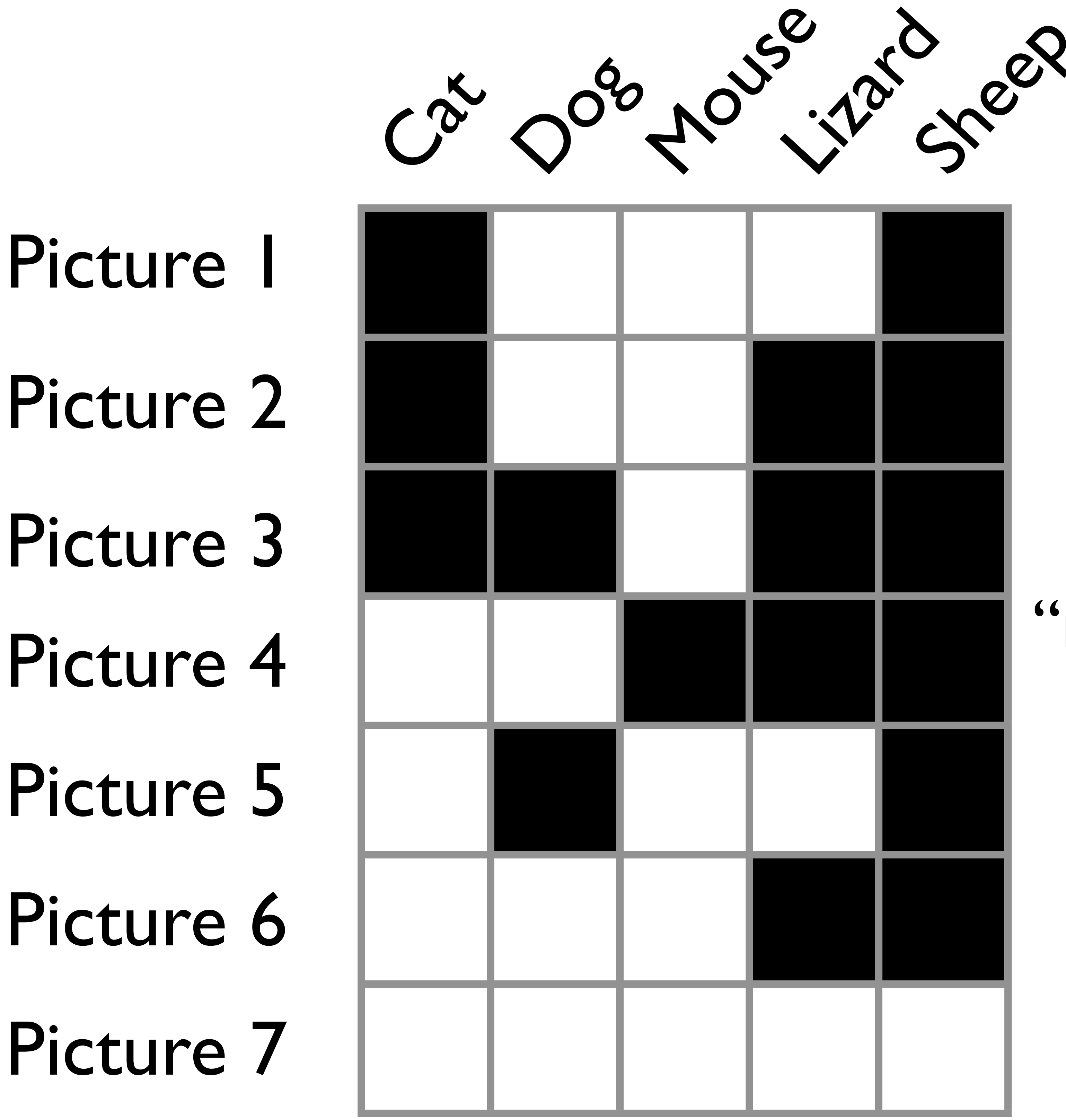
Cat
Dog
Mouse
Lizard
Sheep
Picture 1
Picture 2
Picture 3
Picture 4
Picture 5
Picture 6
Picture 7

Feature allocation



"feature model",
"admixture",
"mixed membership",
"topic model"

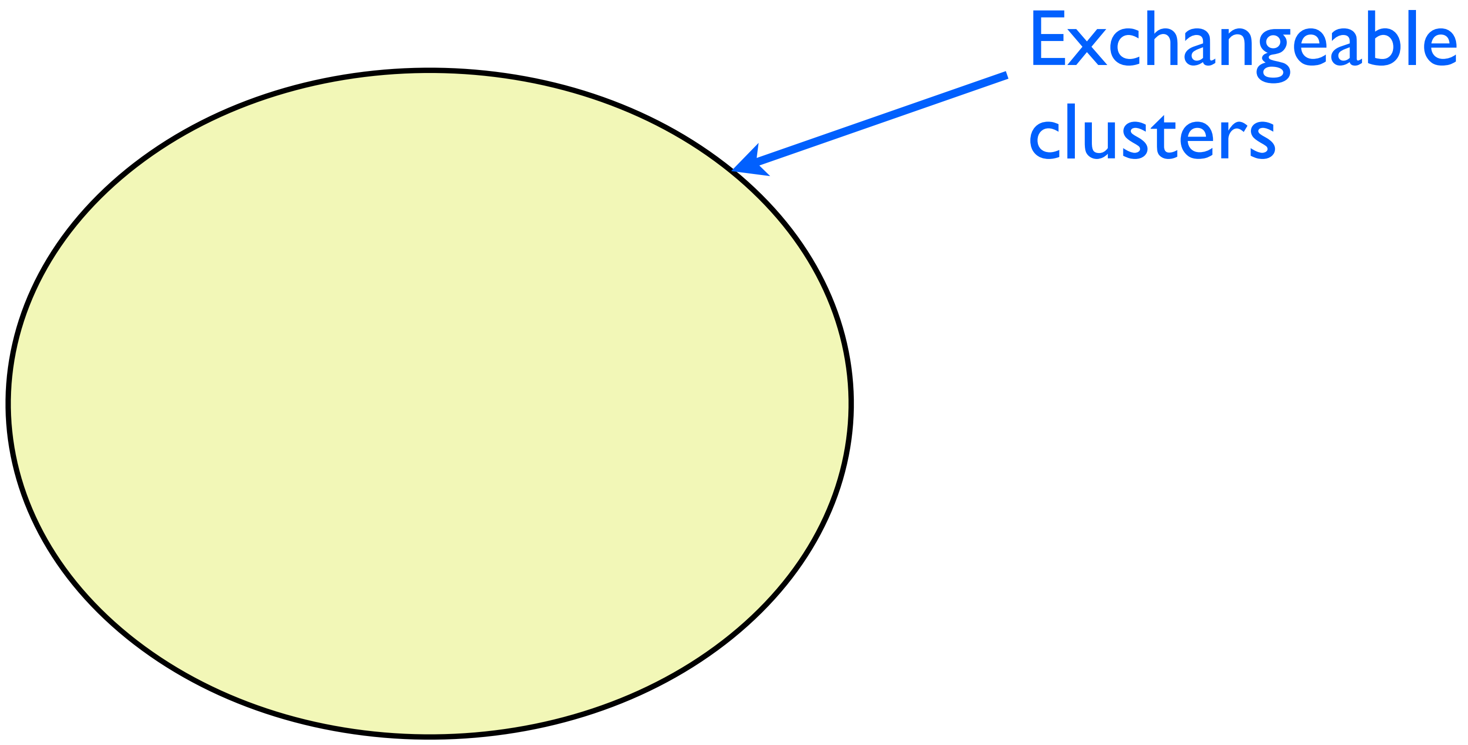
Feature allocation



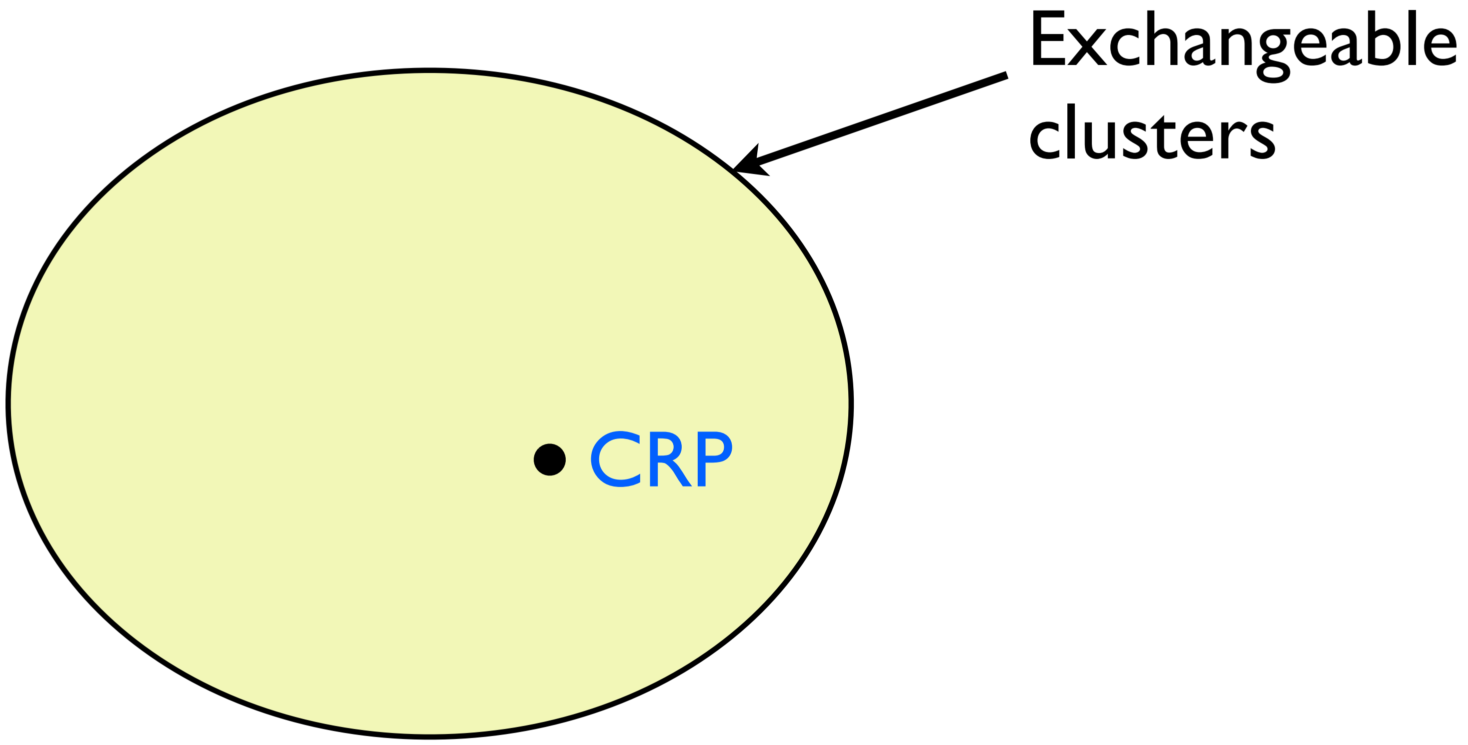
"feature model",
"admixture",
"mixed membership",
"topic model"

- Finite number of features per data point!

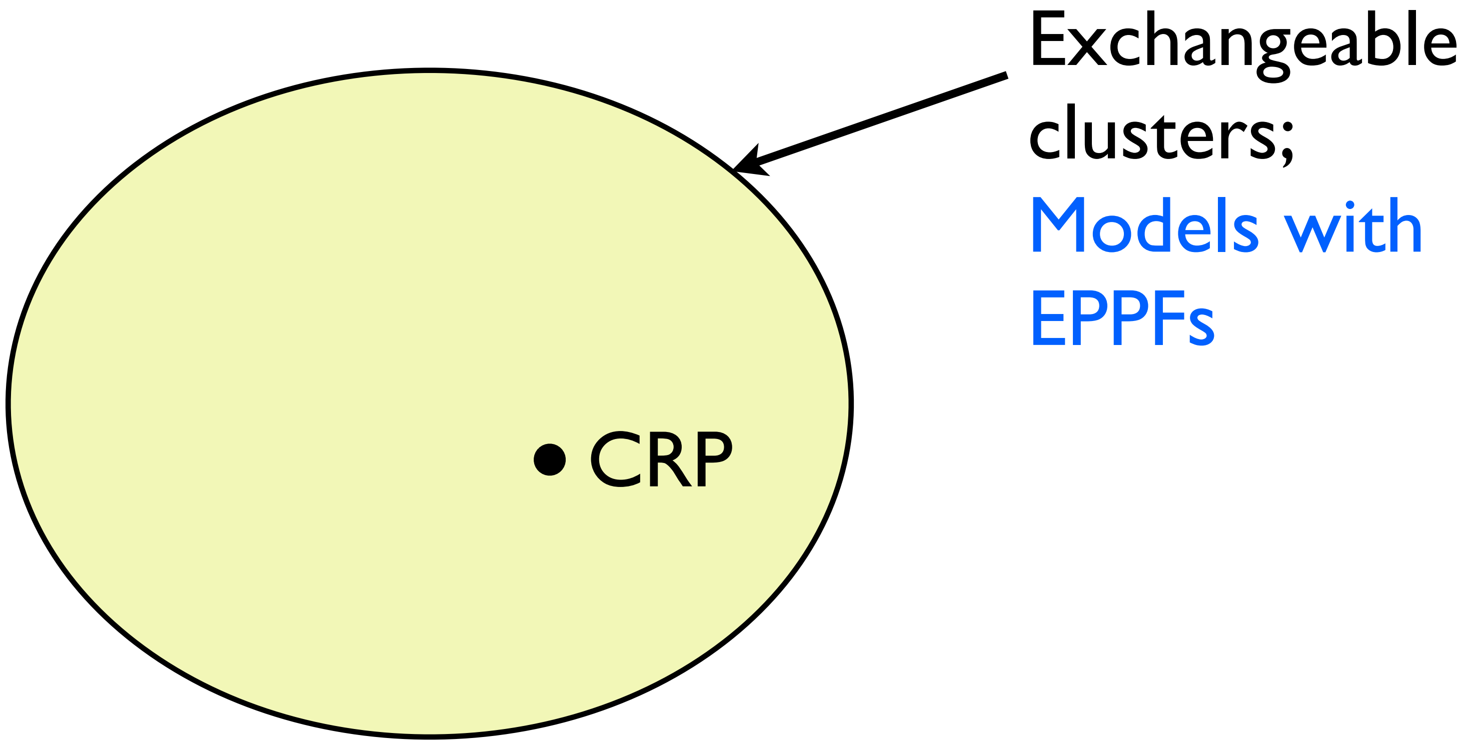
Clusters and features



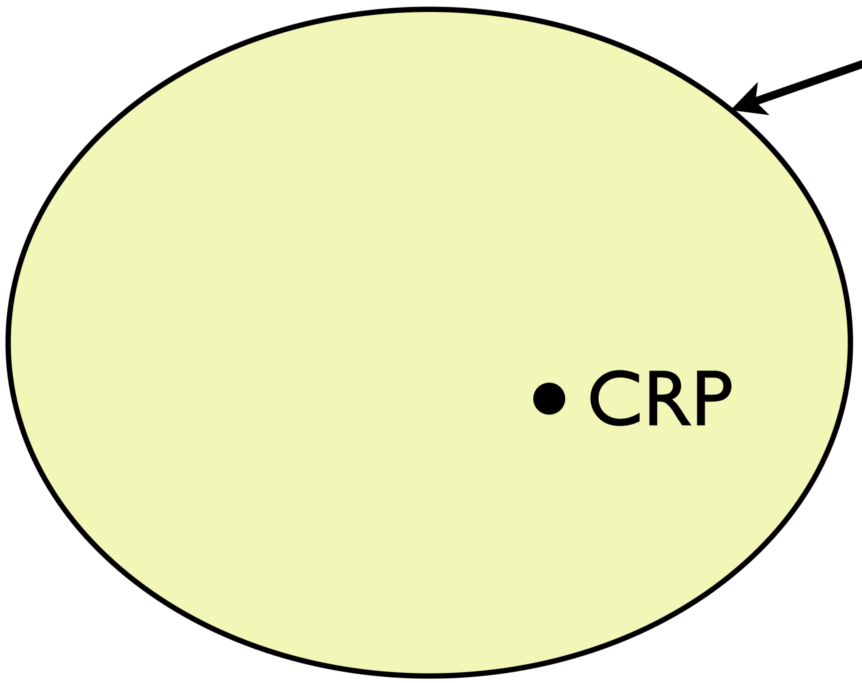
Clusters and features



Clusters and features

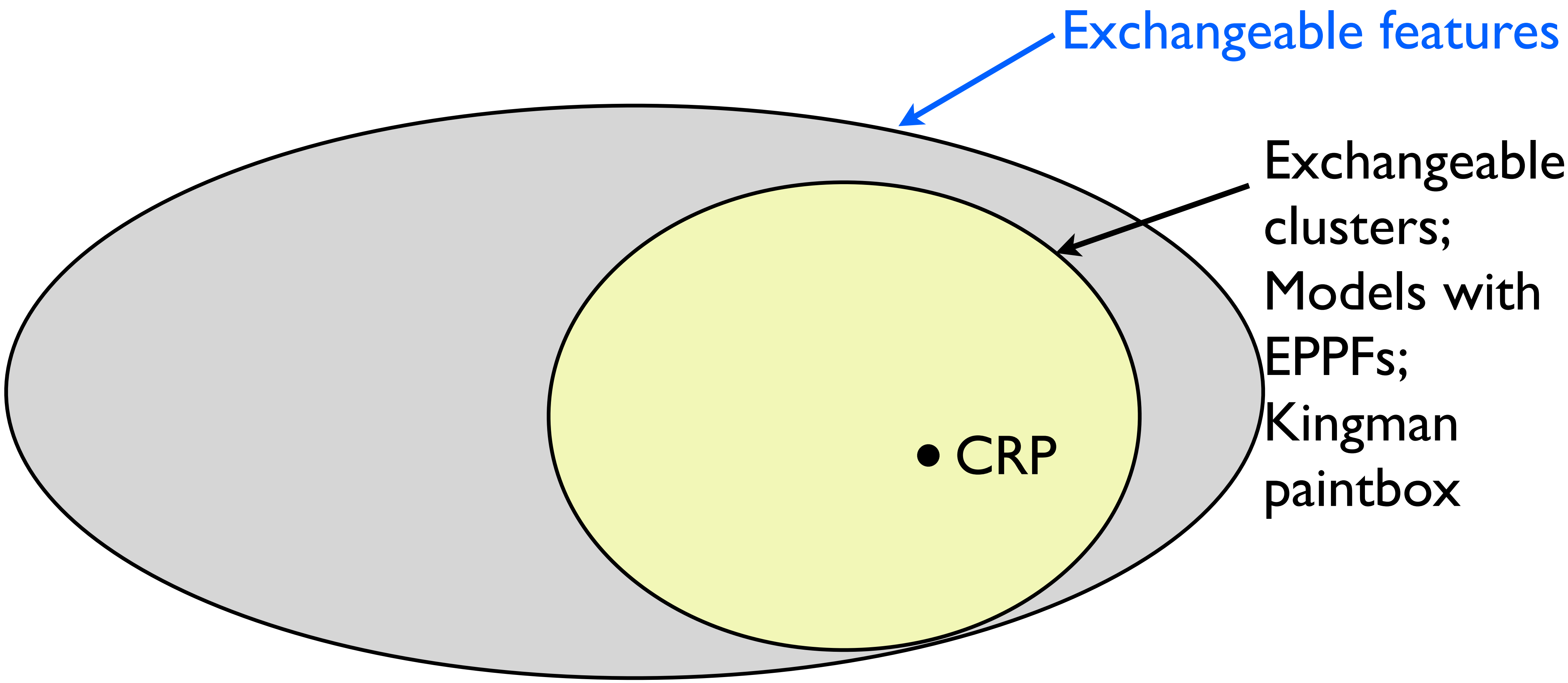


Clusters and features



Exchangeable clusters;
Models with EPPFs;
Kingman paintbox

Clusters and features



Outline

I. Clusters

II. Features

Outline

I. Clusters

II. Features

- Overview
- Distribution
- Frequencies
- Random measure
- Paintbox

Outline

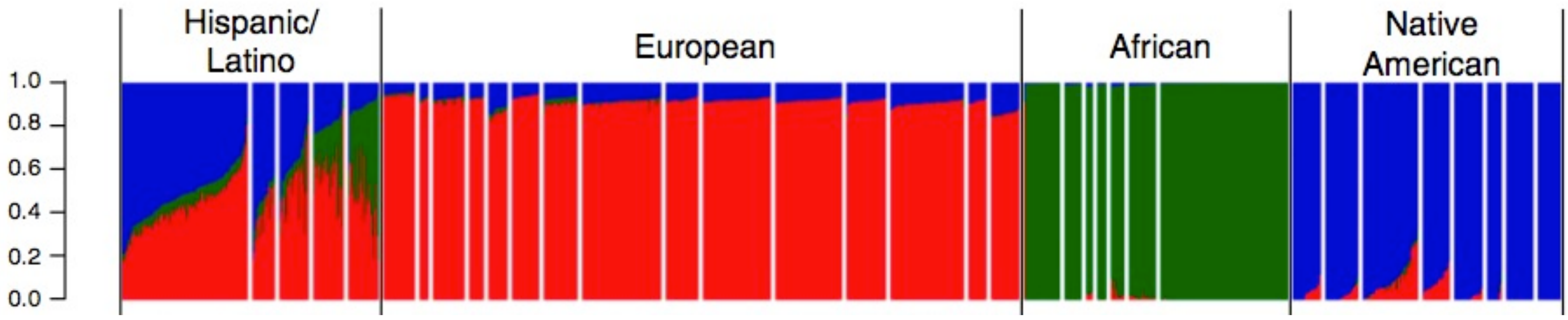
I. Clusters

II. Features

- Overview
- Distribution
- Frequencies
- Random measure
- Paintbox

Feature allocation

Genetic admixture



[Bryc et al 2010]

Feature allocation

Topic analysis

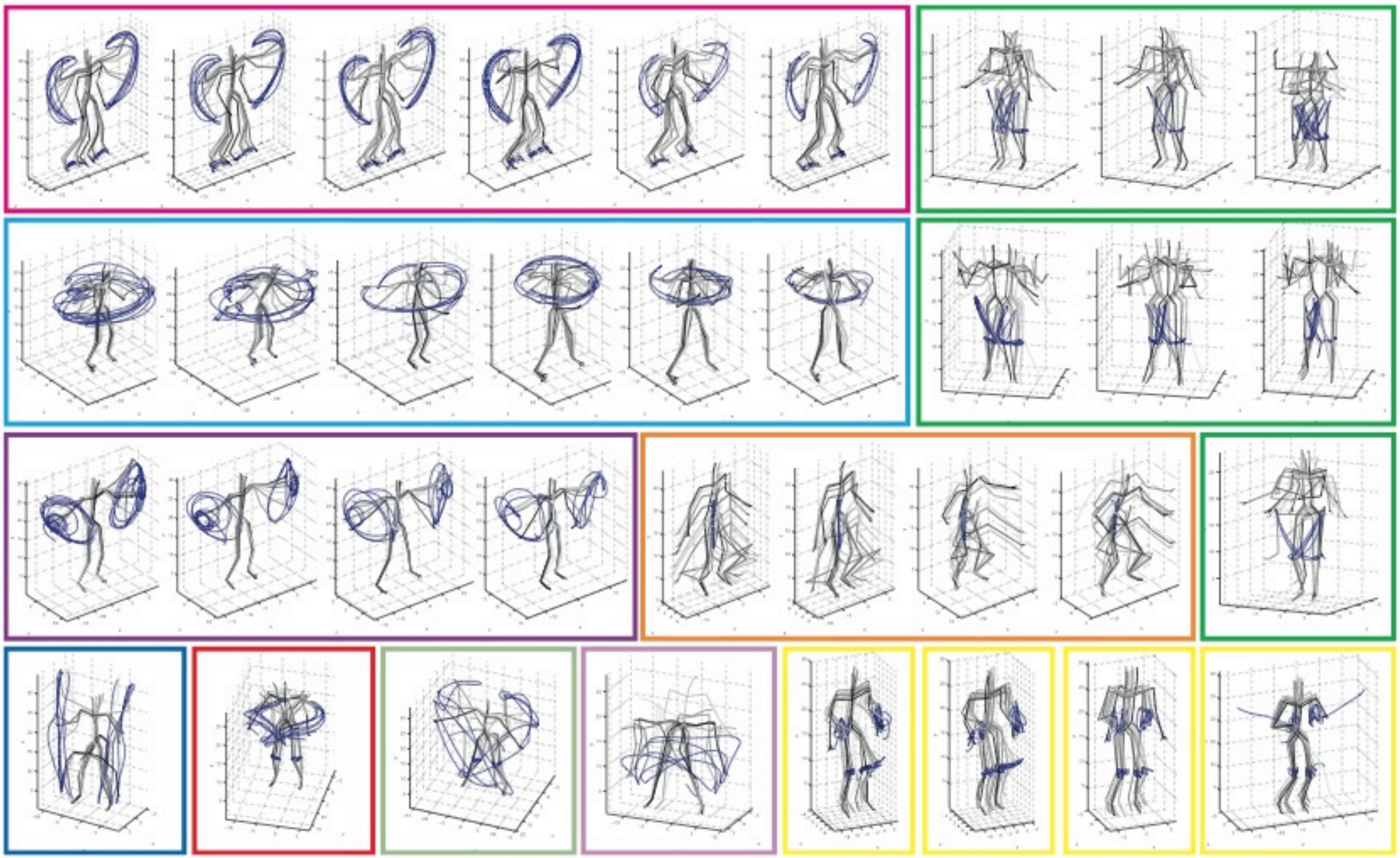
NEW	MILLION	CHILDREN	SCHOOL
FILM	TAX	WOMEN	STUDENTS
SHOW	PROGRAM	PEOPLE	SCHOOLS
MUSIC	BUDGET	CHILD	EDUCATION
MOVIE	BILLION	YEARS	TEACHERS
PLAY	FEDERAL	FAMILIES	HIGH
MUSICAL	YEAR	WORK	PUBLIC
BEST	SPENDING	PARENTS	TEACHER
ACTOR	NEW	SAYS	BENNETT
FIRST	STATE	FAMILY	MANIGAT
YORK	PLAN	WELFARE	NAMPHY
OPERA	MONEY	MEN	STATE
THEATER	PROGRAMS	PERCENT	PRESIDENT
ACTRESS	GOVERNMENT	CARE	ELEMENTARY
LOVE	CONGRESS	LIFE	HAITI

arst Foundation will give $1.25 million to Lincoln Center, Metropoli-
k Philharmonic and Juilliard School. "Our board felt that we had a
a mark on the future of the performing arts with these grants an act
our traditional areas of support in health, medical research, education
Hearst Foundation President Randolph A. Hearst said Monday in
incoln Center's share will be $200,000 for its new building, which
and provide new public facilities. The Metropolitan Opera Co. and
will receive $400,000 each. The Juilliard School, where music and
the performing arts are taught, will get $250,000. The Hearst Foundation, a leading supporter of the Lincoln Center Consolidated Corporate Fund, will make its usual annual $100,000 donation, too.

[Blei et al 2003]

Feature allocation

Switching behaviors



[Fox et al 2009]

Outline

I. Clusters

II. Features
- Overview
- Distribution
- Frequencies
- Random measure
- Paintbox

Outline

I. Clusters

II. Features

- Overview
- Distribution
- Frequencies
- Random measure
- Paintbox

Feature allocation

- Clustering: Exchangeable partition probability function (EPPF)

Feature allocation

- Clustering: Exchangeable partition probability function (EPPF)
- What about feature allocations?

Feature allocation

- Clustering: Exchangeable partition probability function (EPPF)
- What about feature allocations?

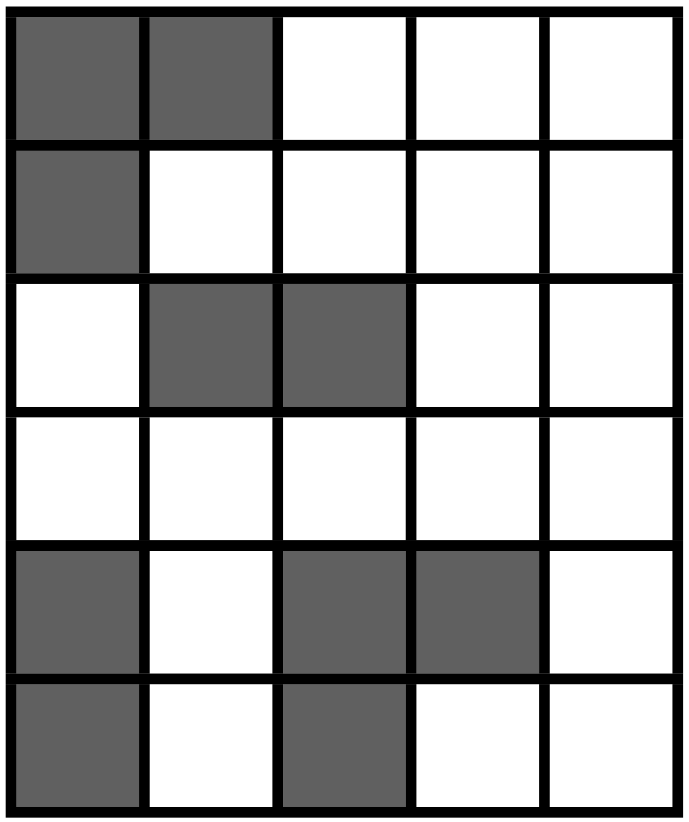
“Exchangeable feature probability function” (EFPF)?

EFPF Example

Indian buffet process

[Griffiths, Ghahramani 2006]

EFPF Example

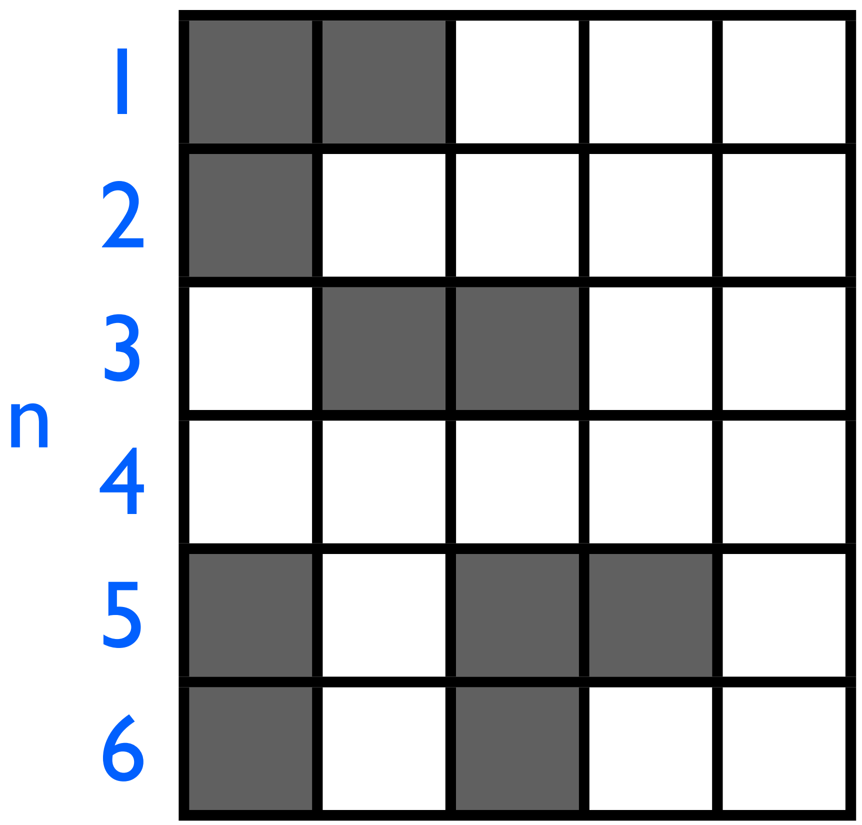


Indian buffet process

- Restaurant ⇔ feature allocation

[Griffiths, Ghahramani 2006]

EFPF Example

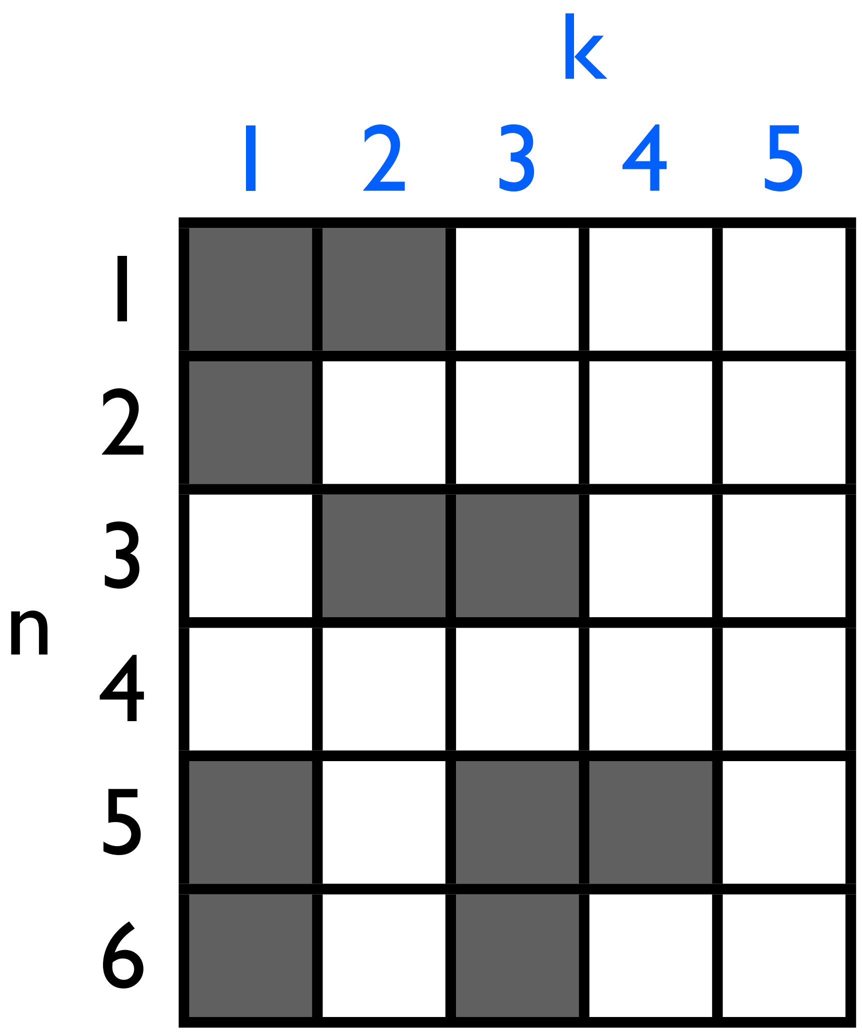


Indian buffet process

- Restaurant ⇔ feature allocation
  - Customer ⇔ index

[Griffiths, Ghahramani 2006]

EFPF Example



Indian buffet process

- Restaurant ⇔ feature allocation
  - Customer ⇔ index
  - Dish in the buffet ⇔ feature

[Griffiths, Ghahramani 2006]

EFPF Example

k

n	1	2	3	4	5
1					
2					
3					
4					
5					
6					

Indian buffet process

[Griffiths, Ghahramani 2006]

EFPF Example

k

n	1	2	3	4	5
1					
2					
3					
4					
5					
6					

Indian buffet process

For n = 1, 2, ..., N

1. Customer *n* prefers popular dishes

2. Customer *n* might try something new

[Griffiths, Ghahramani 2006]

EFPF Example

k

n	1	2	3	4	5
1					
2					
3					
4					
5					
6					

Indian buffet process

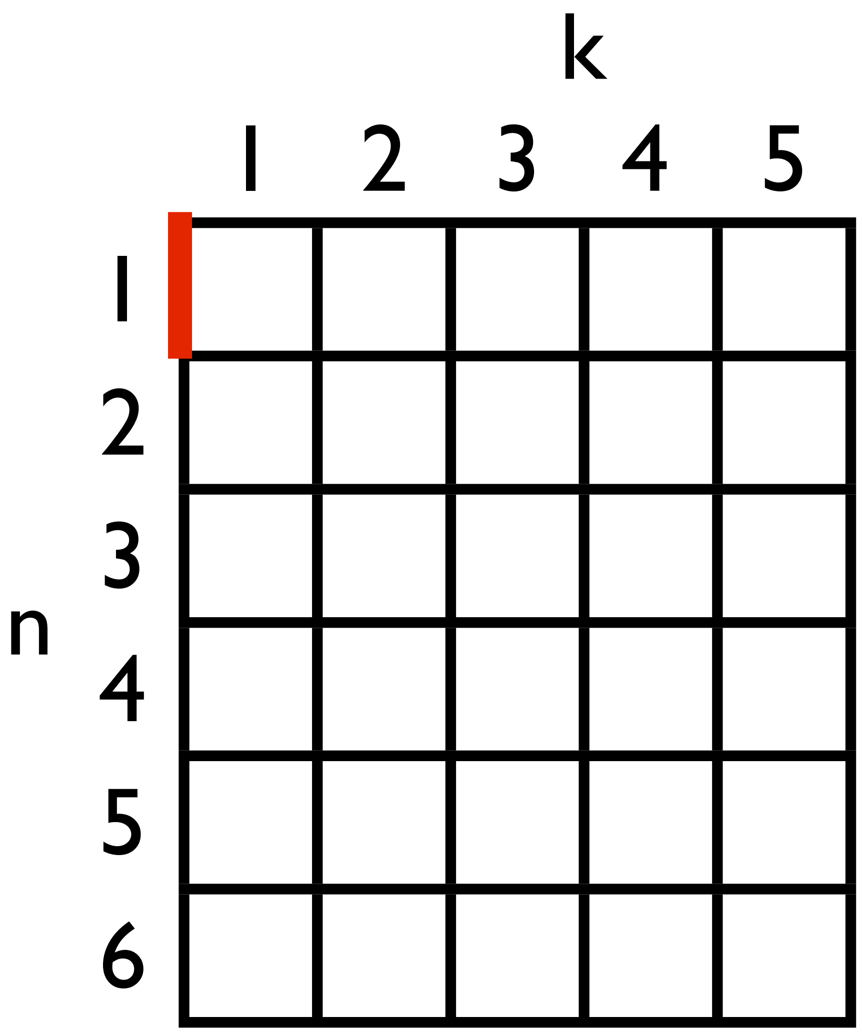
For n = 1, 2, ..., N

1. Customer *n* prefers popular dishes

2. Customer *n* might try something new

[Griffiths, Ghahramani 2006]

EFPF Example



Indian buffet process

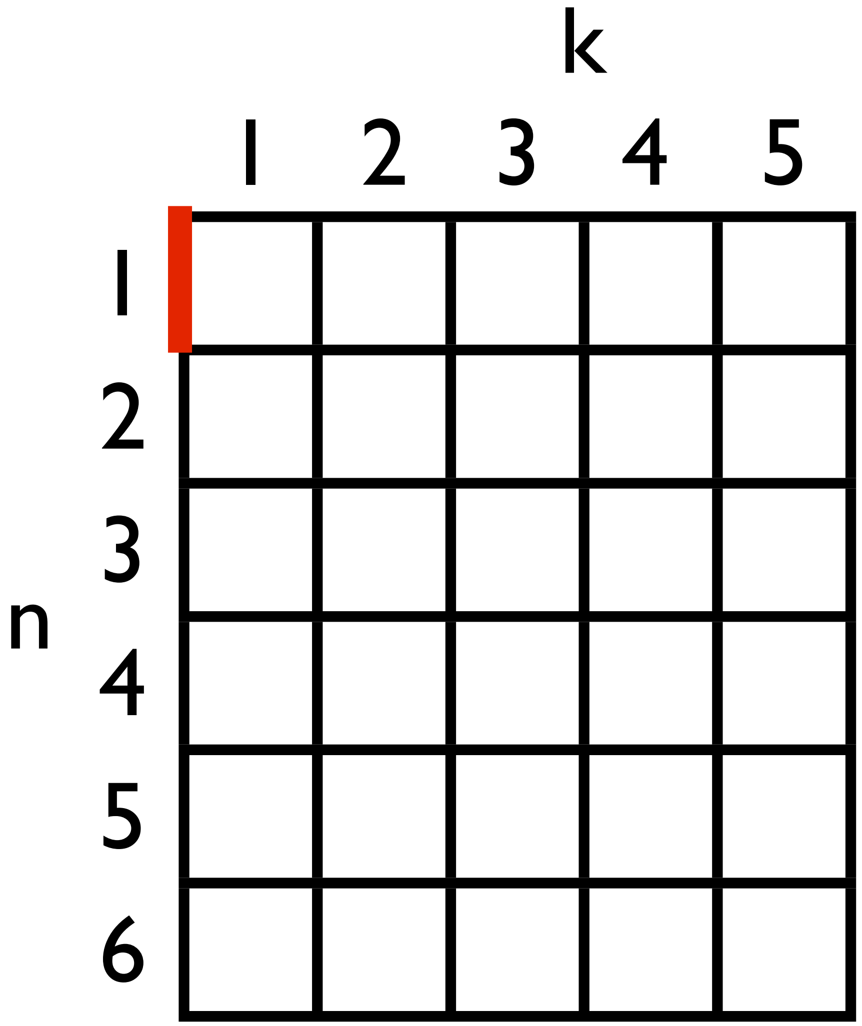
For n = 1, 2, ..., N

1. Customer *n* samples an existing dish *k* that has already been sampled $M_{n-1,k}$ times with probability $\dfrac{M_{n-1,k}}{\theta + n - 1}$

2. Customer *n* might try something new

[Griffiths, Ghahramani 2006]

EFPF Example



concentration parameter

Indian buffet process

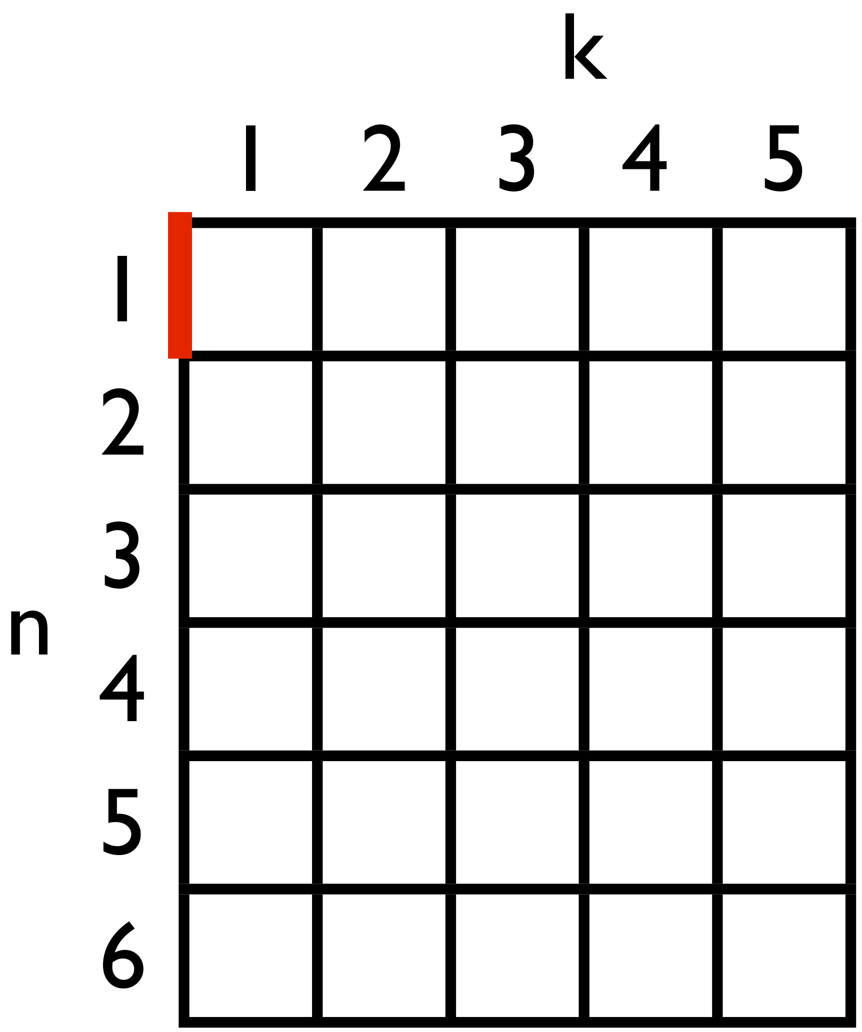
For n = 1, 2, ..., N

1. Customer *n* samples an existing dish *k* that has already been sampled $M_{n-1,k}$ times with probability $\dfrac{M_{n-1,k}}{\theta + n - 1}$

2. Customer *n* might try something new

[Griffiths, Ghahramani 2006]

EFPF Example



Indian buffet process

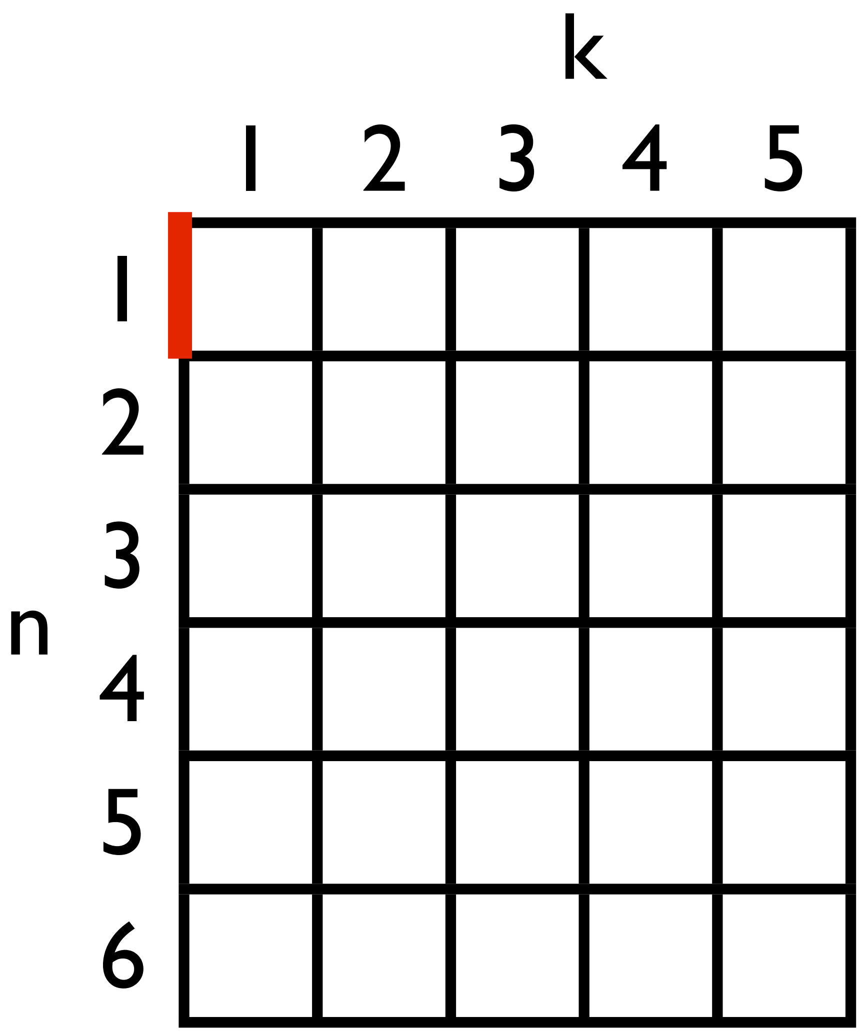
For n = 1, 2, ..., N

1. Customer *n* samples an existing dish *k* that has already been sampled $M_{n-1,k}$ times with probability $\dfrac{M_{n-1,k}}{\theta + n - 1}$

2. Customer *n* might try something new

[Griffiths, Ghahramani 2006]

EFPF Example



Indian buffet process

For n = 1, 2, ..., N

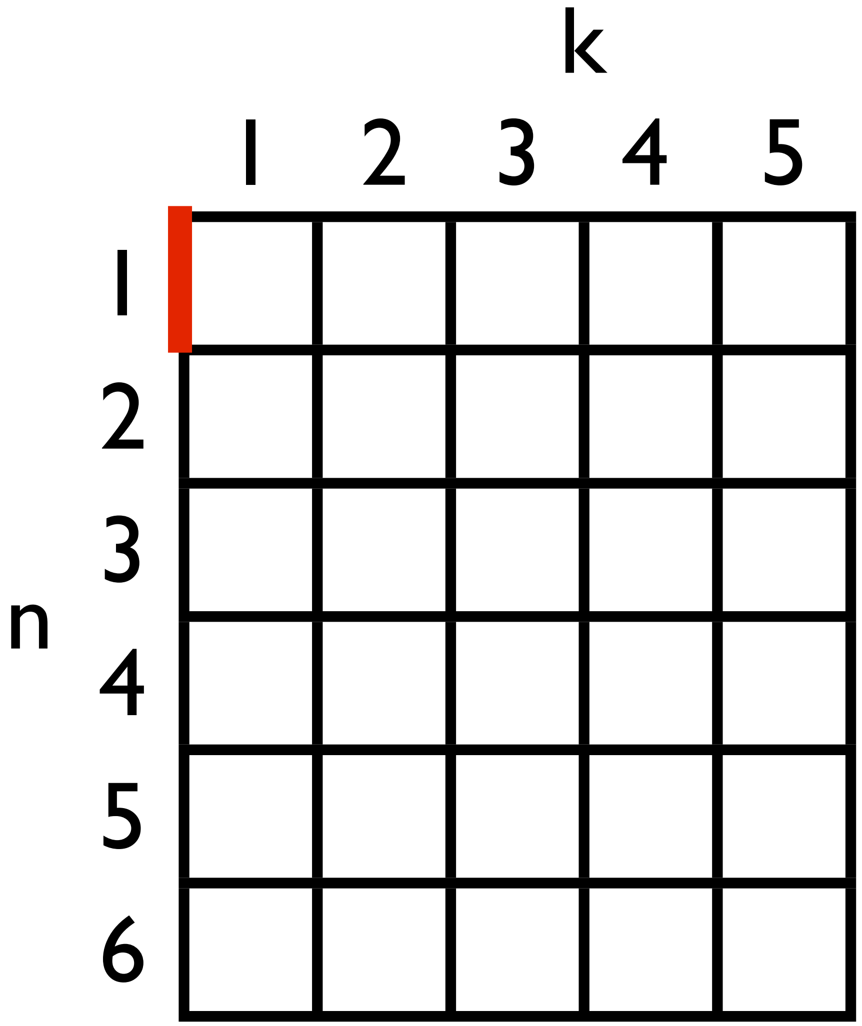
1. Customer *n* samples an existing dish *k* that has already been sampled $M_{n-1,k}$ times with probability $\dfrac{M_{n-1,k}}{\theta + n - 1}$

2. Number of new dishes for index *n*:

$$K_n^+ = \text{Poisson}\left(\gamma \frac{\theta}{\theta + n - 1}\right)$$

[Griffiths, Ghahramani 2006]

EFPF Example



Indian buffet process

For n = 1, 2, ..., N

1. Customer *n* samples an existing dish *k* that has already been sampled $M_{n-1,k}$ times with probability $\dfrac{M_{n-1,k}}{\theta + n - 1}$

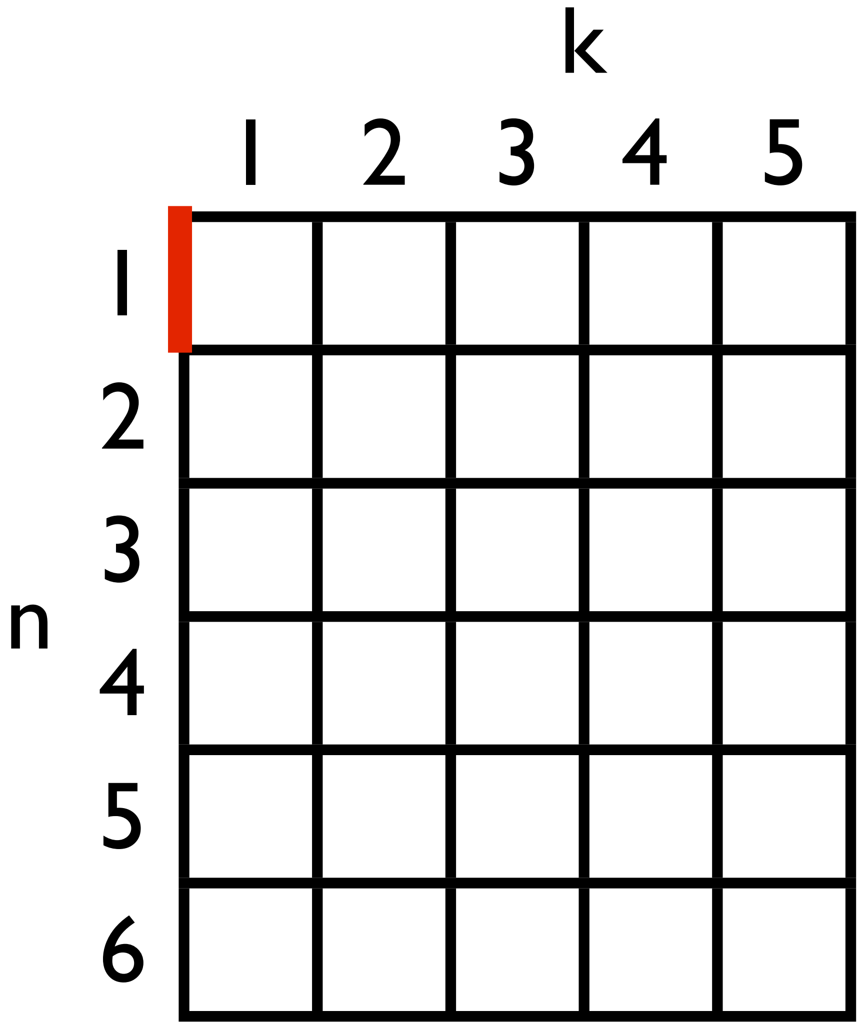
2. Number of new dishes for index *n*:

$$K_n^+ = \text{Poisson}\left(\gamma \frac{\theta}{\theta + n - 1}\right)$$

mass parameter

[Griffiths, Ghahramani 2006]

EFPF Example



Indian buffet process

For n = 1, 2, ..., N

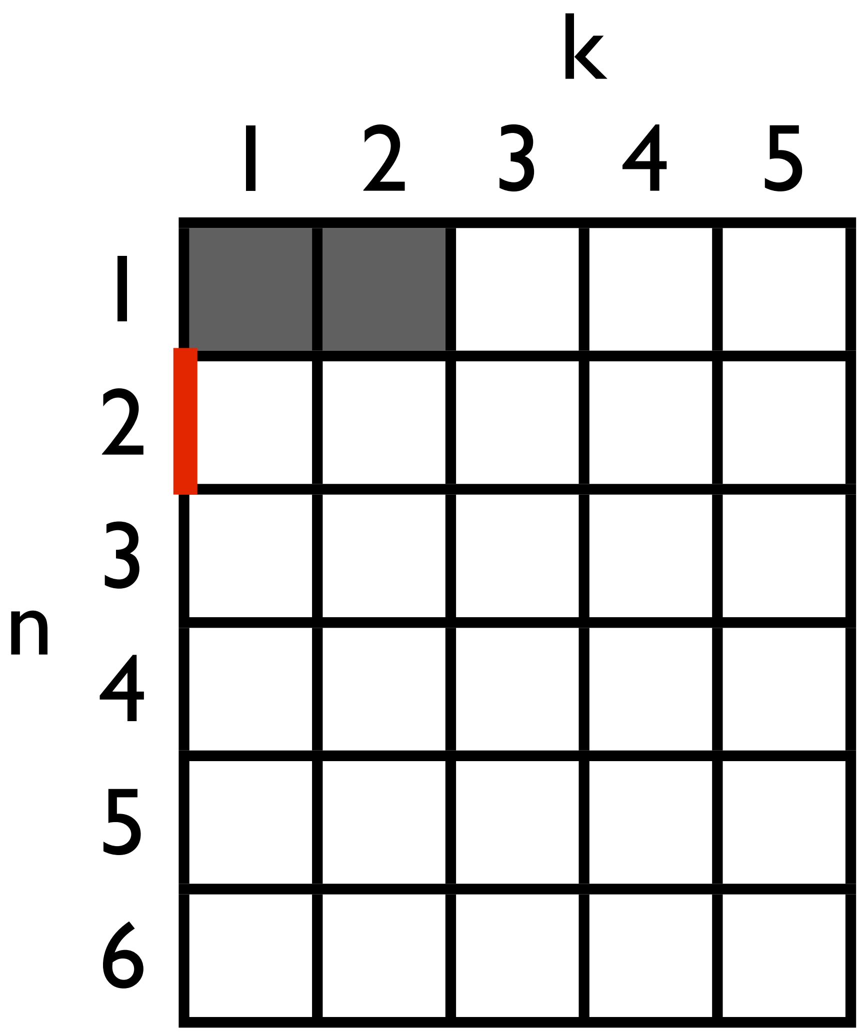
1. Customer *n* samples an existing dish *k* that has already been sampled $M_{n-1,k}$ times with probability $\frac{M_{n-1,k}}{\theta + n - 1}$

2. Number of new dishes for index *n*:

$$K_n^+ = \mathrm{Poisson}\left(\gamma \frac{\theta}{\theta + n - 1}\right)$$

$$\mathrm{Poisson}\left(\gamma\right)$$

EFPF Example



Indian buffet process

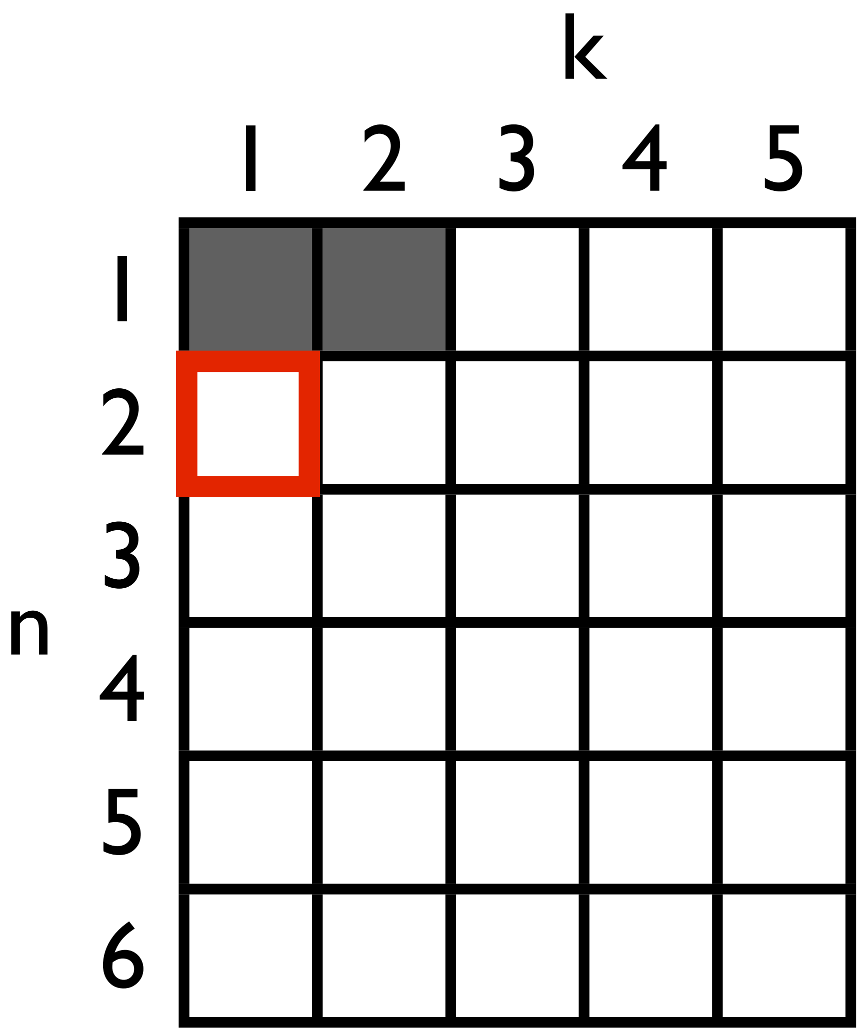
For n = 1, 2, ..., N

1. Customer *n* samples an existing dish *k* that has already been sampled $M_{n-1,k}$ times with probability $\dfrac{M_{n-1,k}}{\theta + n - 1}$

2. Number of new dishes for index *n*:

$$K_n^+ = \text{Poisson}\left(\gamma \frac{\theta}{\theta + n - 1}\right)$$

EFPF Example



Indian buffet process

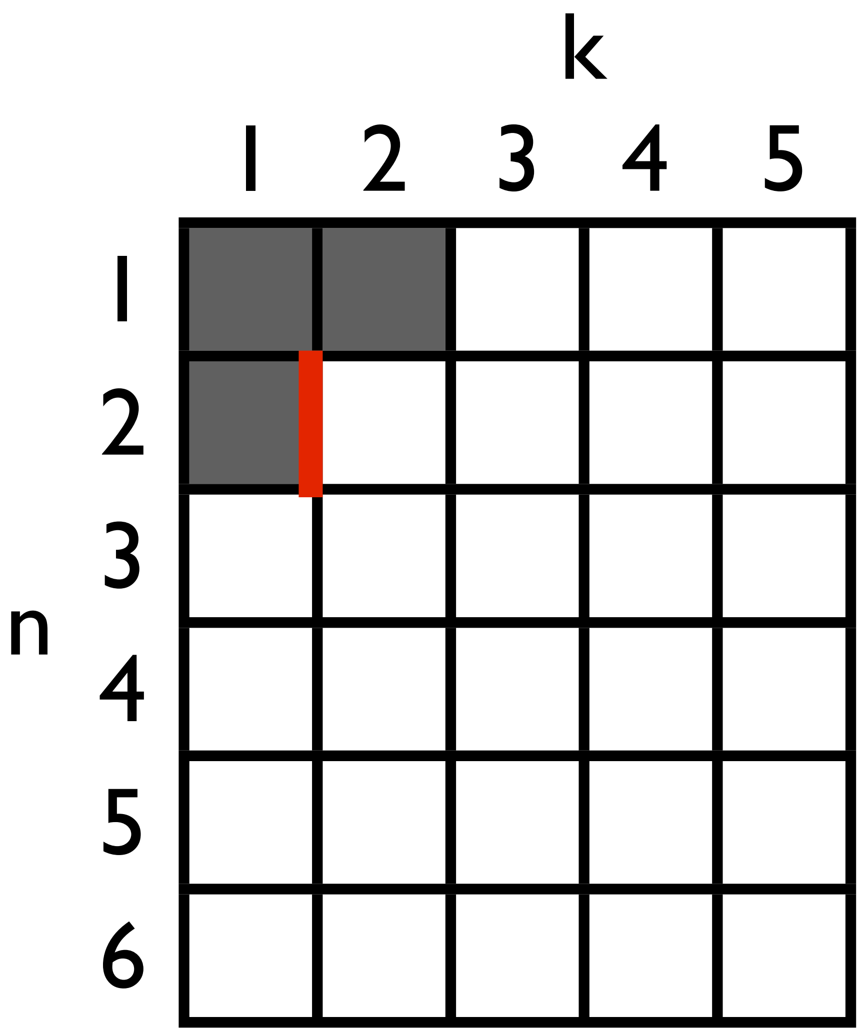
For n = 1, 2, ..., N

1. Customer *n* samples an existing dish *k* that has already been sampled $M_{n-1,k}$ times with probability $\frac{M_{n-1,k}}{\theta + n - 1}$ $\frac{1}{\theta + 1}$

2. Number of new dishes for index *n*:

$$K_n^+ = \mathrm{Poisson}\left(\gamma \frac{\theta}{\theta + n - 1}\right)$$

EFPF Example



Indian buffet process

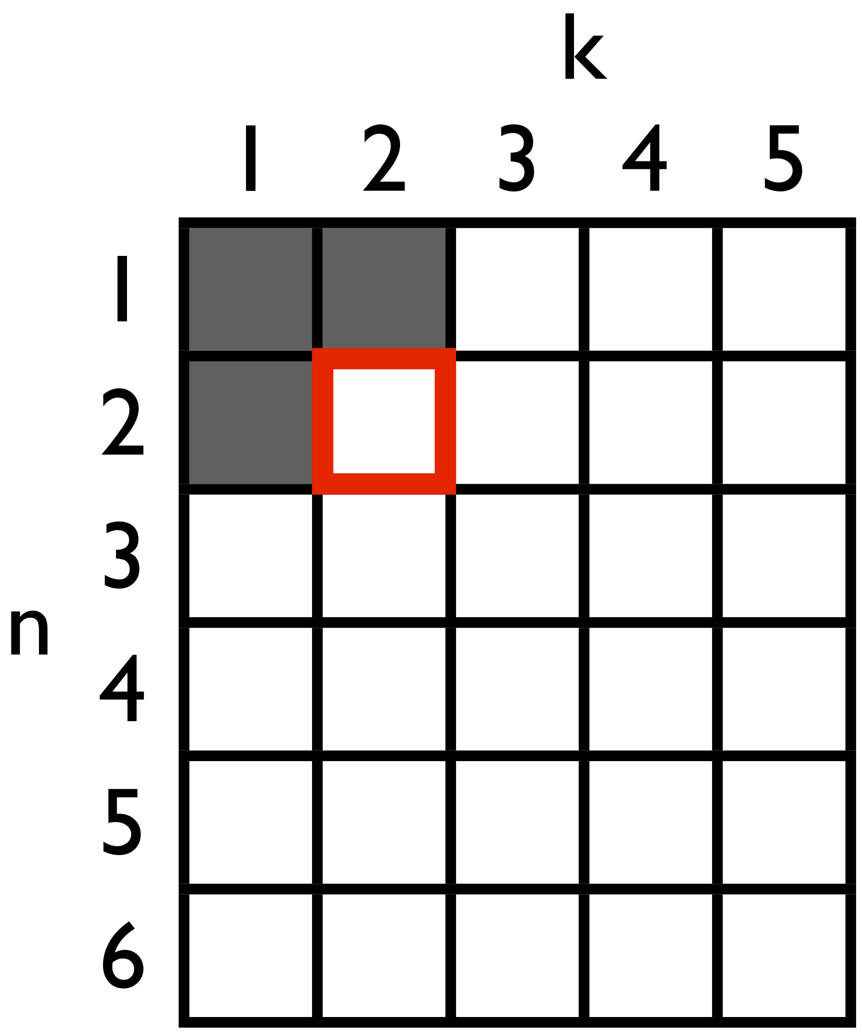
For n = 1, 2, ..., N

1. Customer *n* samples an existing dish *k* that has already been sampled $M_{n-1,k}$ times with probability $\dfrac{M_{n-1,k}}{\theta + n - 1}$

2. Number of new dishes for index *n*:

$$K_n^+ = \mathrm{Poisson}\left(\gamma \frac{\theta}{\theta + n - 1}\right)$$

EFPF Example



Indian buffet process

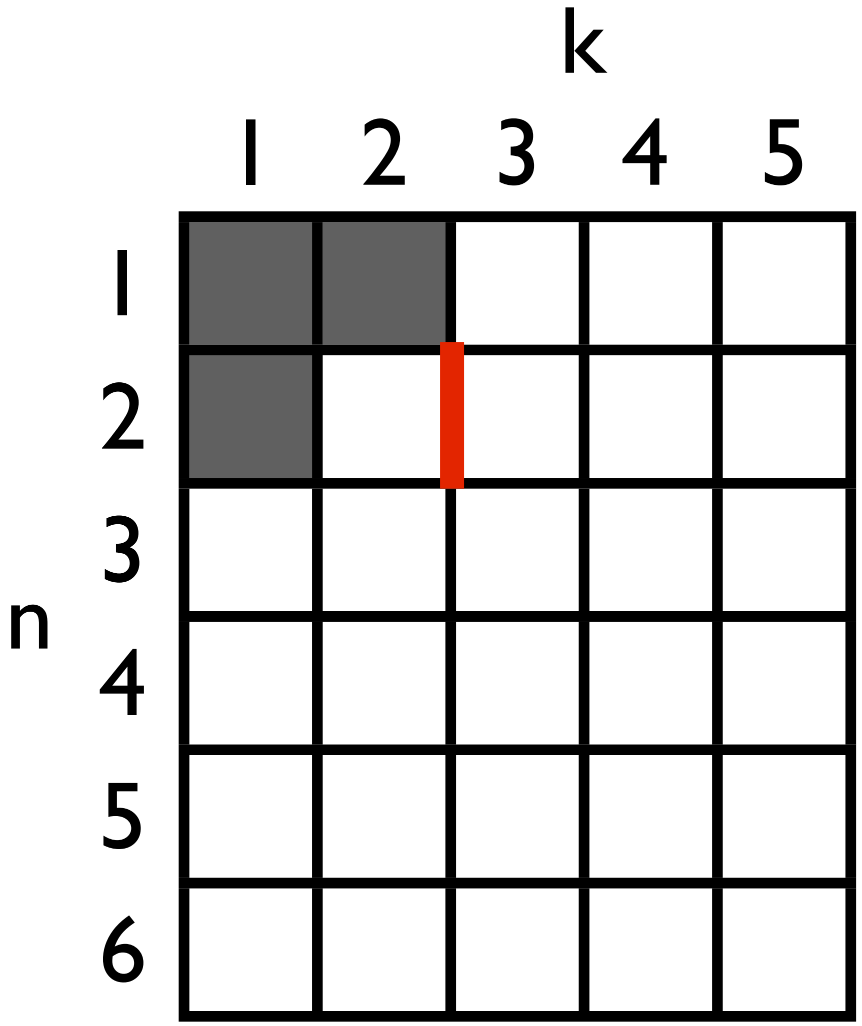
For n = 1, 2, ..., N

1. Customer *n* samples an existing dish *k* that has already been sampled $M_{n-1,k}$ times with probability $\dfrac{M_{n-1,k}}{\theta + n - 1}$ $\dfrac{1}{\theta + 1}$

2. Number of new dishes for index *n*:

$$K_n^+ = \text{Poisson}\left(\gamma \frac{\theta}{\theta + n - 1}\right)$$

EFPF Example



Indian buffet process

For n = 1, 2, ..., N

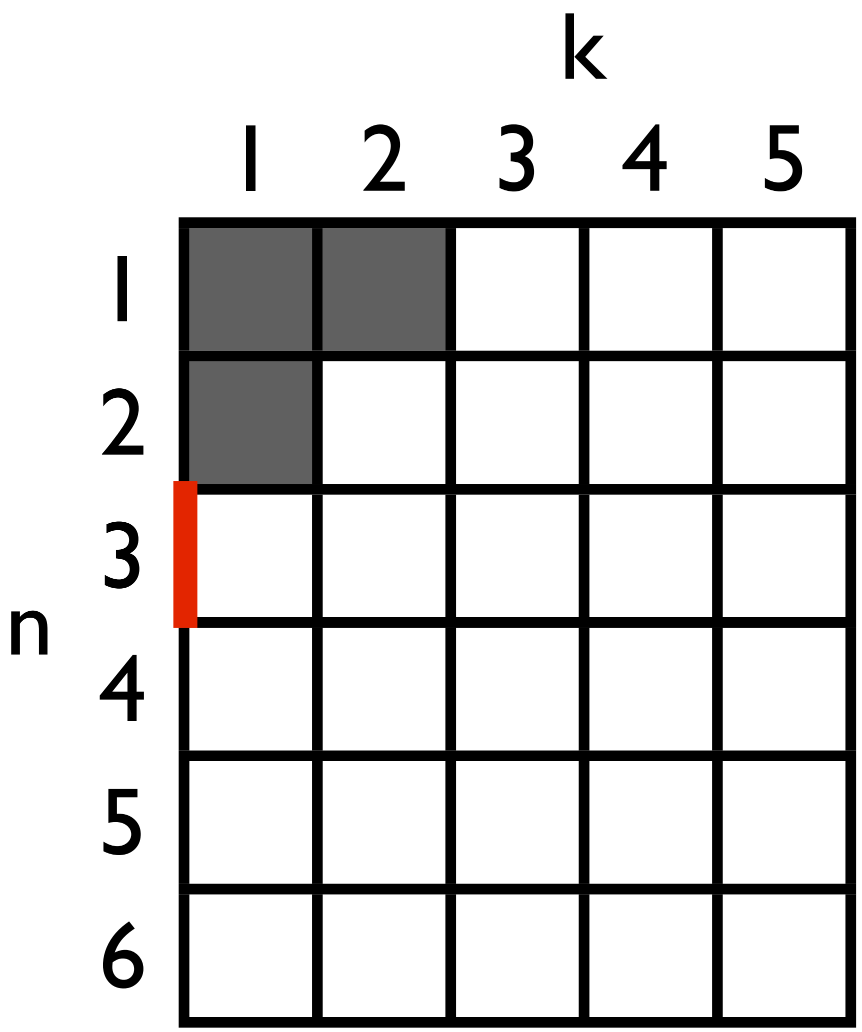
1. Customer *n* samples an existing dish *k* that has already been sampled $M_{n-1,k}$ times with probability $\dfrac{M_{n-1,k}}{\theta + n - 1}$

2. Number of new dishes for index *n*:

$$K_n^+ = \text{Poisson}\left(\gamma \frac{\theta}{\theta + n - 1}\right)$$

$$\text{Poisson}\left(\gamma \frac{\theta}{\theta + 1}\right)$$

EFPF Example



Indian buffet process

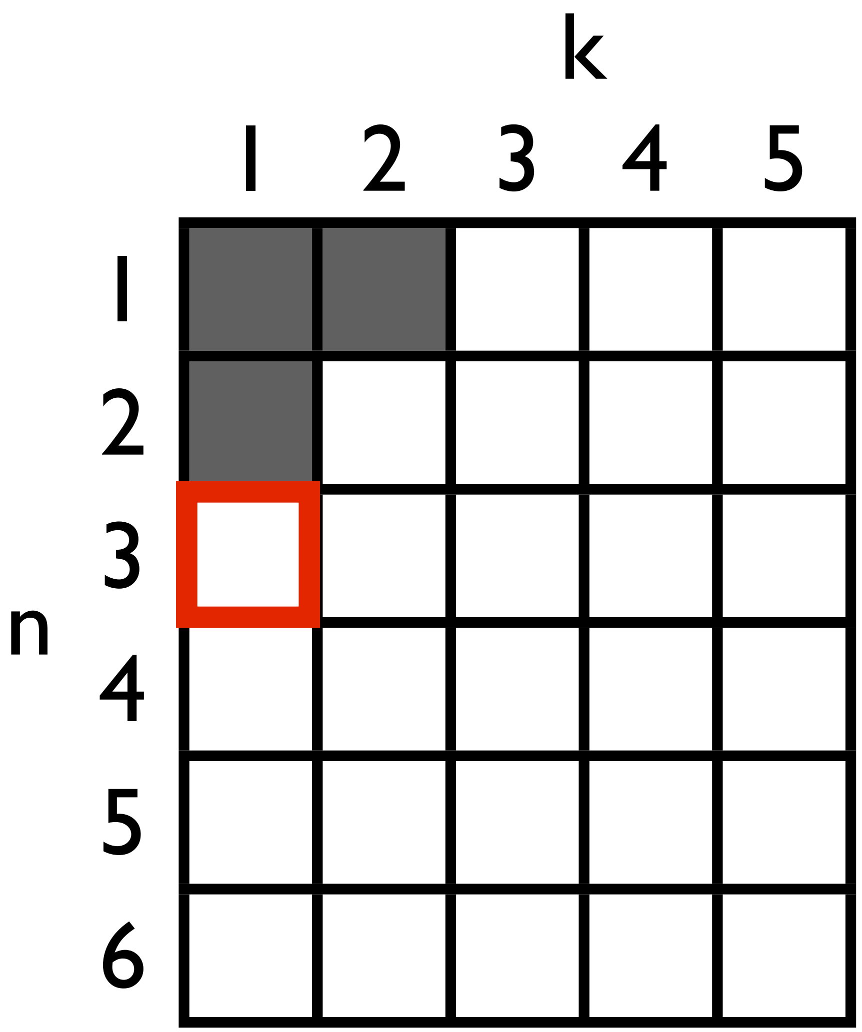
For n = 1, 2, ..., N

1. Customer *n* samples an existing dish *k* that has already been sampled $M_{n-1,k}$ times with probability $\dfrac{M_{n-1,k}}{\theta + n - 1}$

2. Number of new dishes for index *n*:

$$K_n^+ = \text{Poisson}\left(\gamma \frac{\theta}{\theta + n - 1}\right)$$

EFPF Example



Indian buffet process

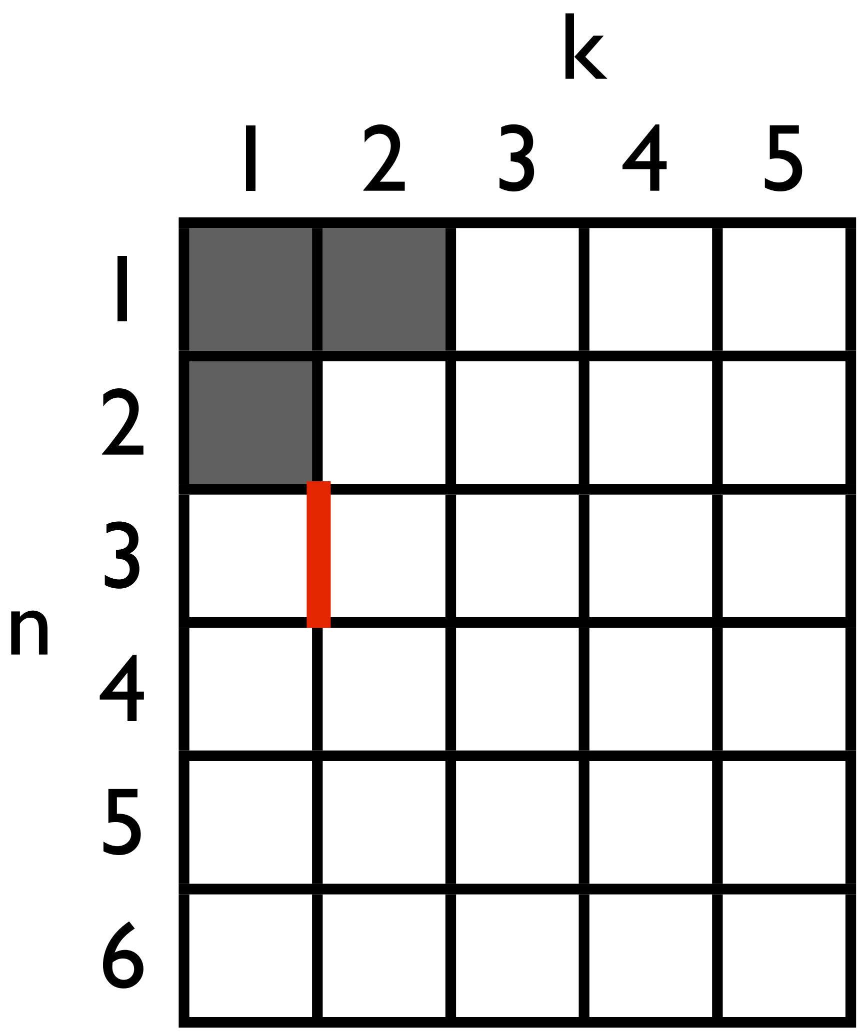
For n = 1, 2, ..., N

1. Customer *n* samples an existing dish *k* that has already been sampled $M_{n-1,k}$ times with probability $\dfrac{M_{n-1,k}}{\theta + n - 1}$ $\dfrac{2}{\theta + 2}$

2. Number of new dishes for index *n*:

$$K_n^+ = \mathrm{Poisson}\left(\gamma \frac{\theta}{\theta + n - 1}\right)$$

EFPF Example



Indian buffet process

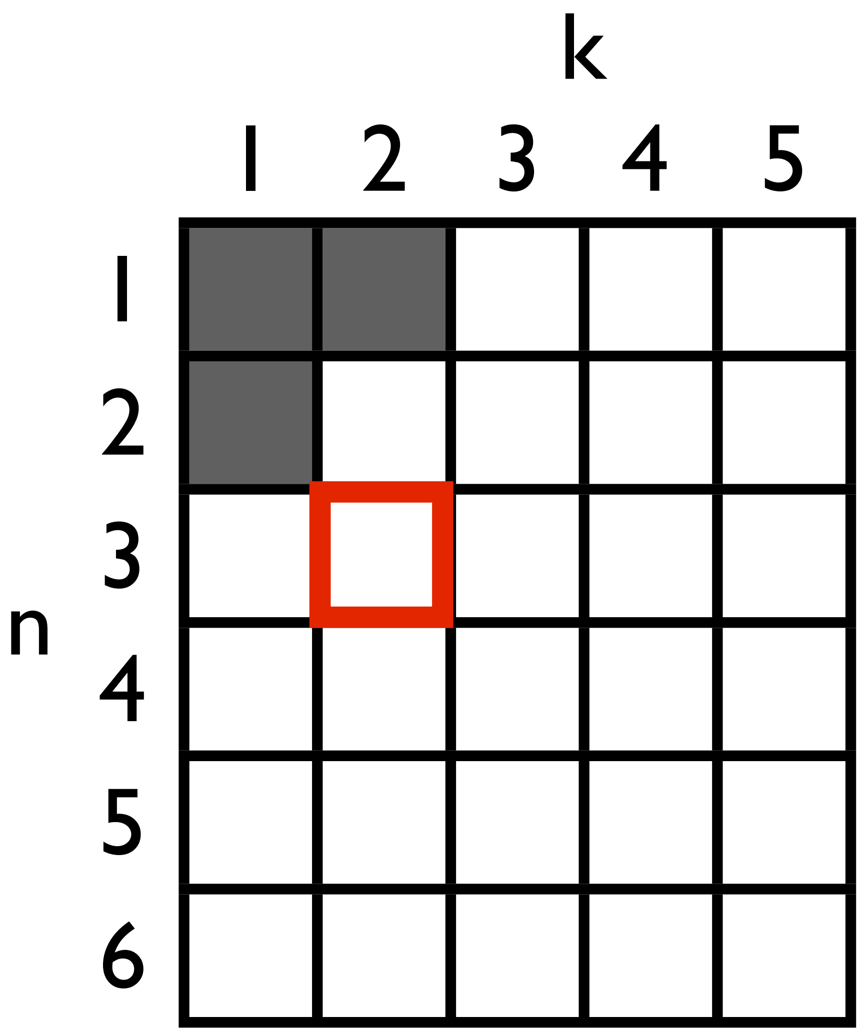
For n = 1, 2, ..., N

1. Customer *n* samples an existing dish *k* that has already been sampled $M_{n-1,k}$ times with probability $\frac{M_{n-1,k}}{\theta + n - 1}$

2. Number of new dishes for index *n*:

$$K_n^+ = \text{Poisson}\left(\gamma \frac{\theta}{\theta + n - 1}\right)$$

EFPF Example



Indian buffet process

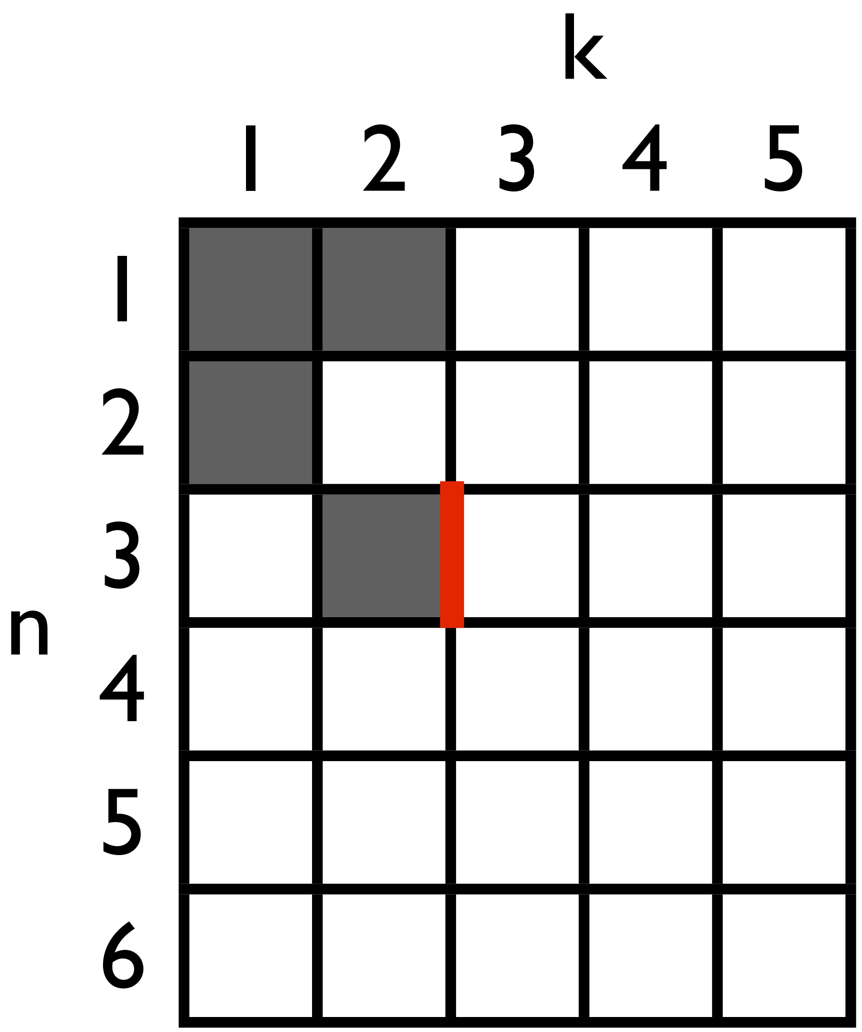
For n = 1, 2, ..., N

1. Customer *n* samples an existing dish *k* that has already been sampled $M_{n-1,k}$ times with probability $\frac{M_{n-1,k}}{\theta + n - 1}$ $\frac{1}{\theta + 2}$

2. Number of new dishes for index *n*:

$$K_n^+ = \text{Poisson}\left(\gamma \frac{\theta}{\theta + n - 1}\right)$$

EFPF Example



Indian buffet process

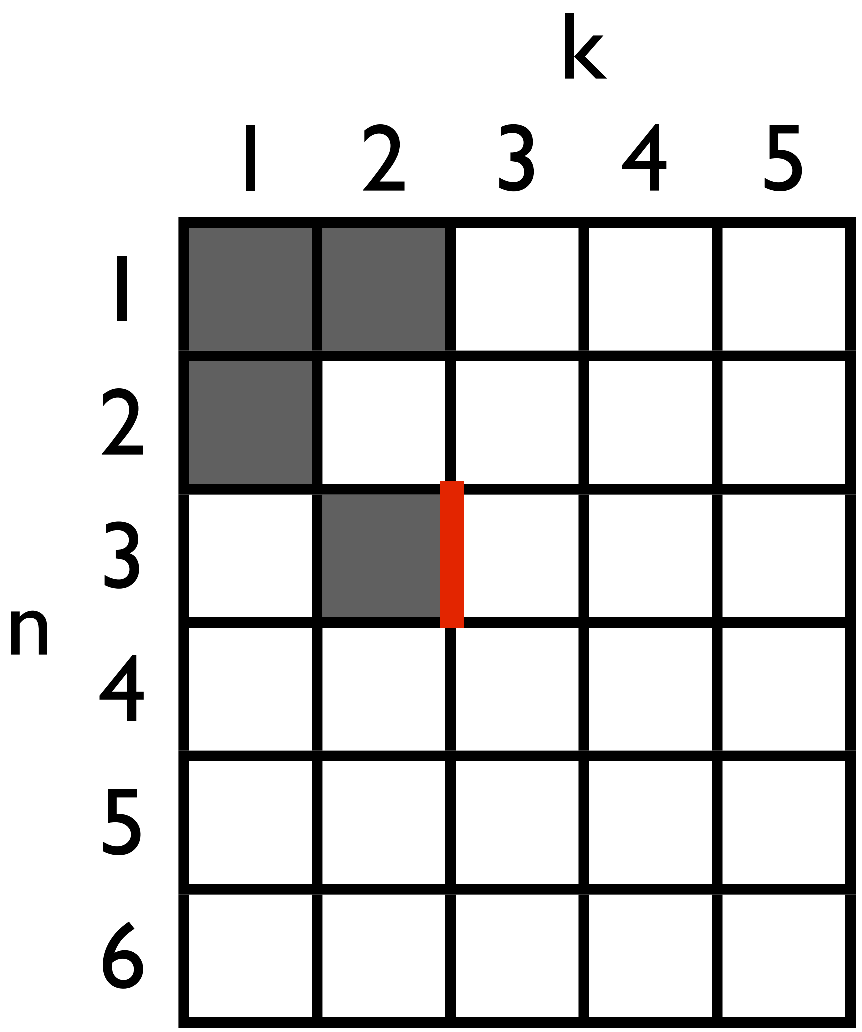
For n = 1, 2, ..., N

1. Customer *n* samples an existing dish *k* that has already been sampled $M_{n-1,k}$ times with probability $\frac{M_{n-1,k}}{\theta+n-1}$

2. Number of new dishes for index *n*:

$$K_n^+ = \text{Poisson}\left(\gamma \frac{\theta}{\theta+n-1}\right)$$

EFPF Example



Indian buffet process

For n = 1, 2, ..., N

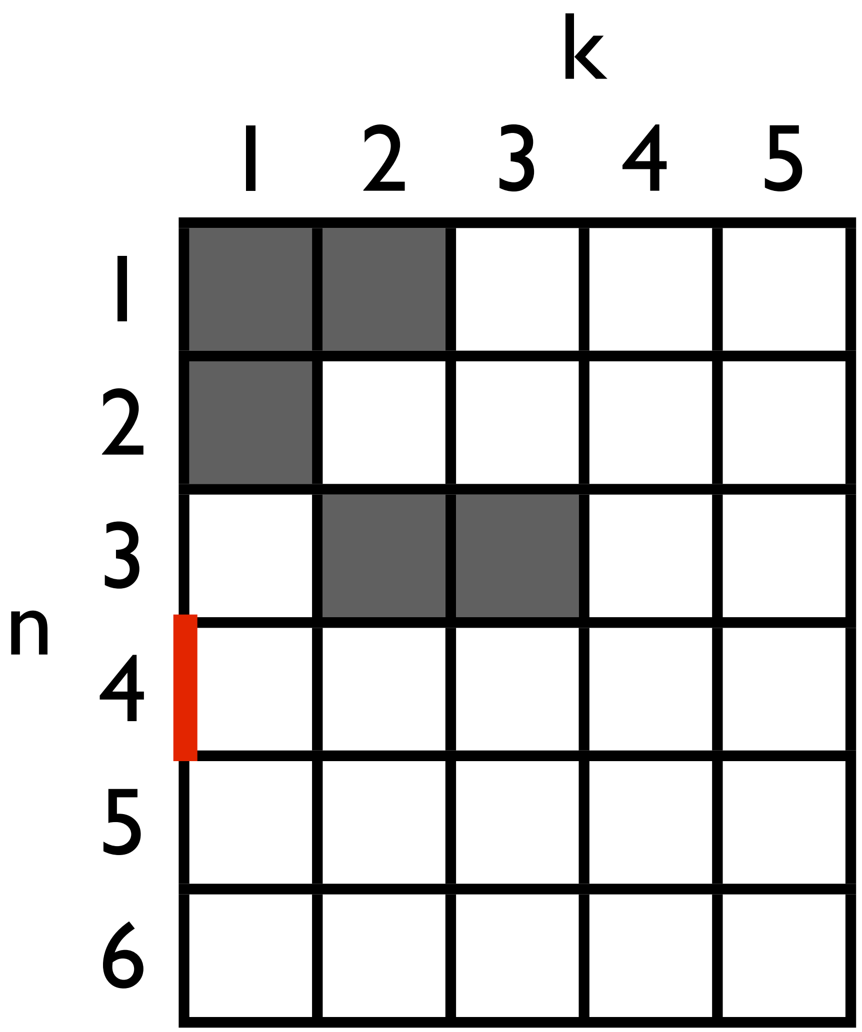
1. Customer *n* samples an existing dish *k* that has already been sampled $M_{n-1,k}$ times with probability $\frac{M_{n-1,k}}{\theta + n - 1}$

2. Number of new dishes for index *n*:

$$K_n^+ = \text{Poisson}\left(\gamma \frac{\theta}{\theta + n - 1}\right)$$

$$\text{Poisson}\left(\gamma \frac{\theta}{\theta + 2}\right)$$

EFPF Example



Indian buffet process

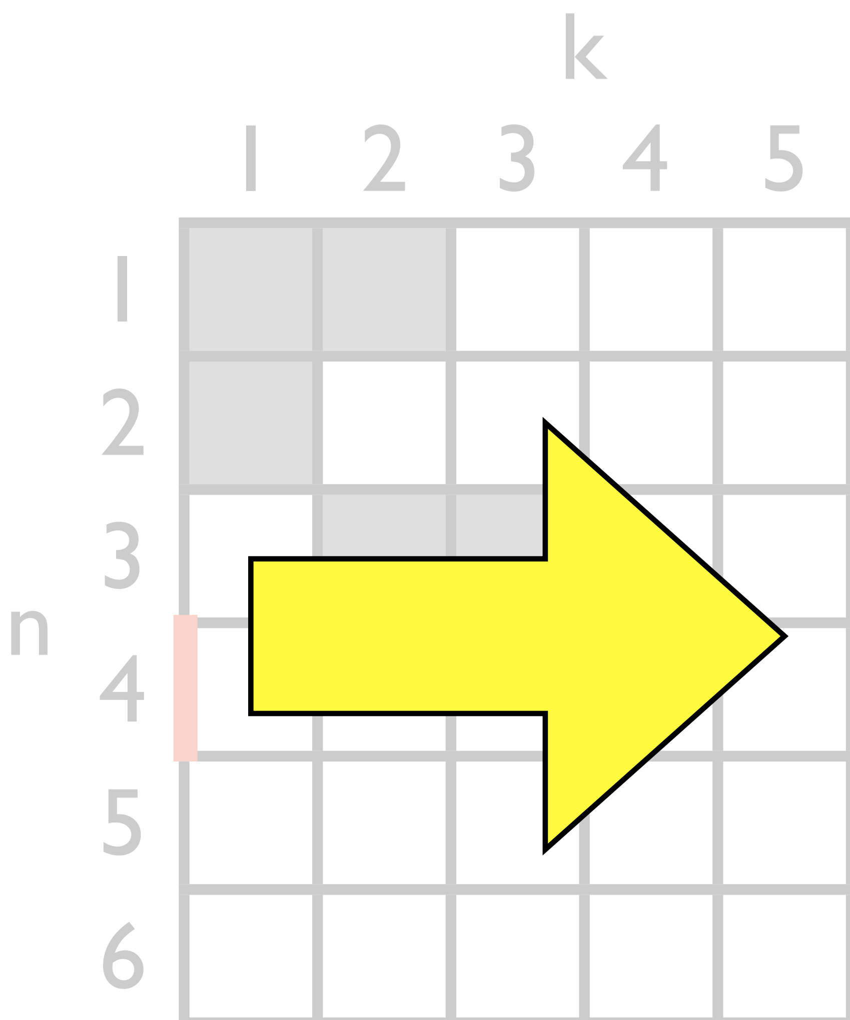
For n = 1, 2, ..., N

1. Customer *n* samples an existing dish *k* that has already been sampled $M_{n-1,k}$ times with probability $\dfrac{M_{n-1,k}}{\theta + n - 1}$

2. Number of new dishes for index *n*:

$$K_n^+ = \text{Poisson}\left(\gamma \frac{\theta}{\theta + n - 1}\right)$$

EFPF Example



Indian buffet process

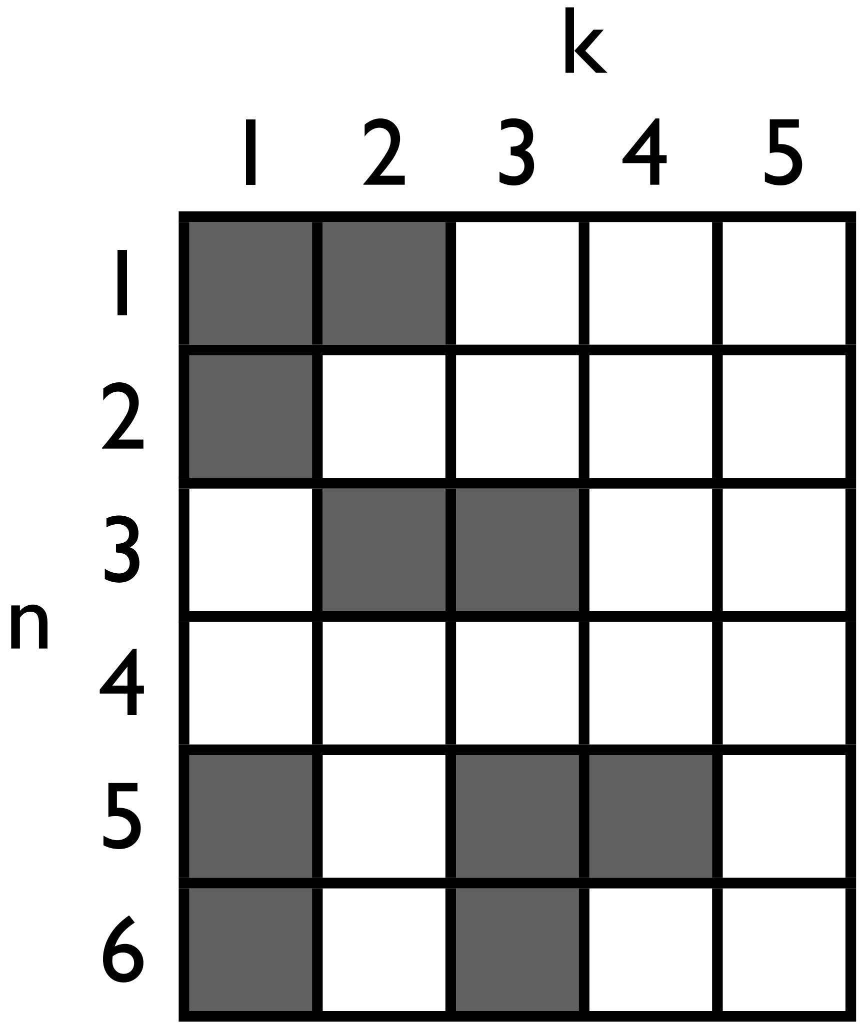
For n = 1, 2, ..., N

1. Customer *n* samples an existing dish *k* that has already been sampled $M_{n-1,k}$ times with probability $\dfrac{M_{n-1,k}}{\theta + n - 1}$

2. Number of new dishes for index *n*:

$$K_n^+ = \text{Poisson}\left(\gamma \frac{\theta}{\theta + n - 1}\right)$$

EFPF Example



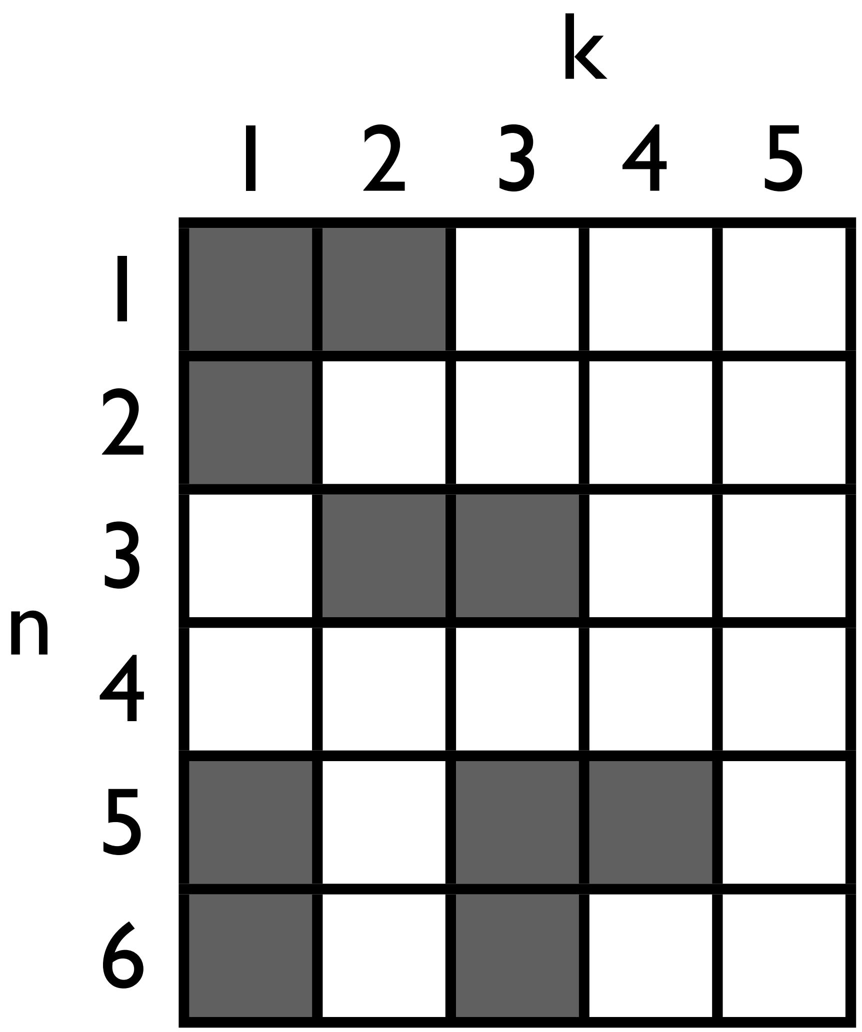
Indian buffet process

For n = 1, 2, ..., N

1. Customer *n* samples an existing dish *k* that has already been sampled $M_{n-1,k}$ times with probability $\dfrac{M_{n-1,k}}{\theta + n - 1}$
2. Number of new dishes for index *n*:

$$K_n^+ = \mathrm{Poisson}\left(\gamma \frac{\theta}{\theta + n - 1}\right)$$

EFPF Example



concentration parameter

Indian buffet process

For n = 1, 2, ..., N

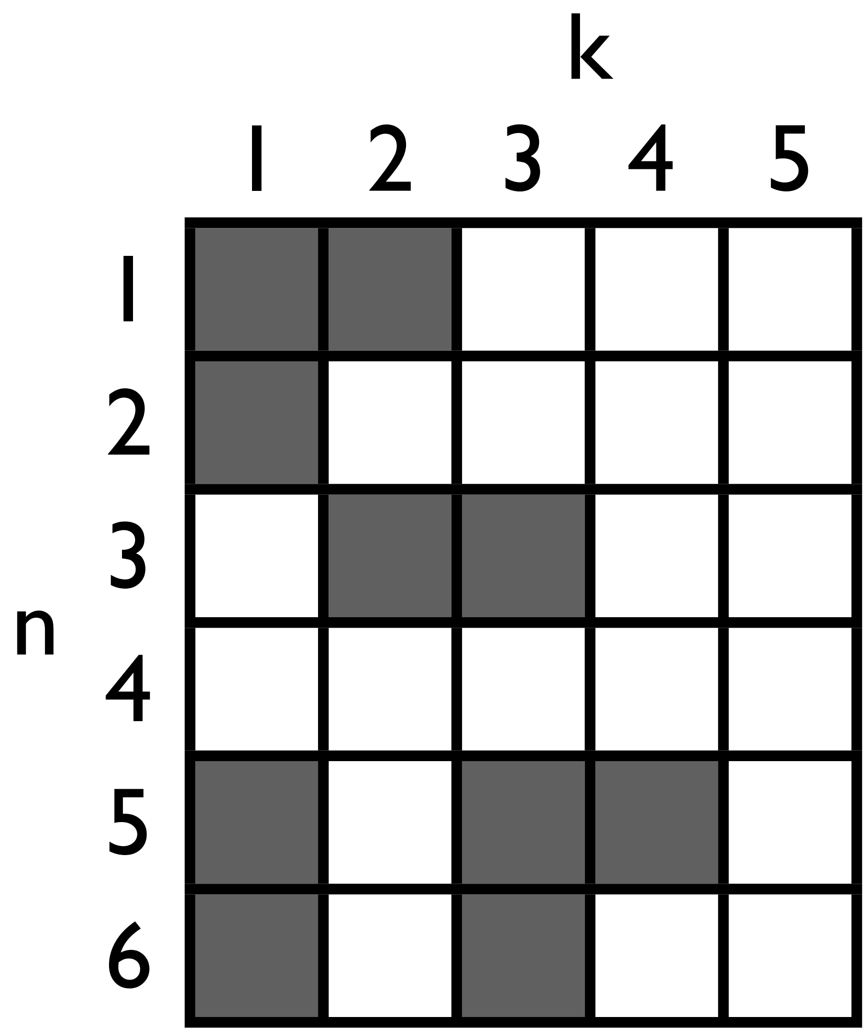
1. Customer *n* samples an existing dish *k* that has already been sampled $M_{n-1,k}$ times with probability $\frac{M_{n-1,k}}{\theta + n - 1}$

2. Number of new dishes for index *n*:

$$K_n^+ = \text{Poisson}\left(\gamma \frac{\theta}{\theta + n - 1}\right)$$

mass parameter

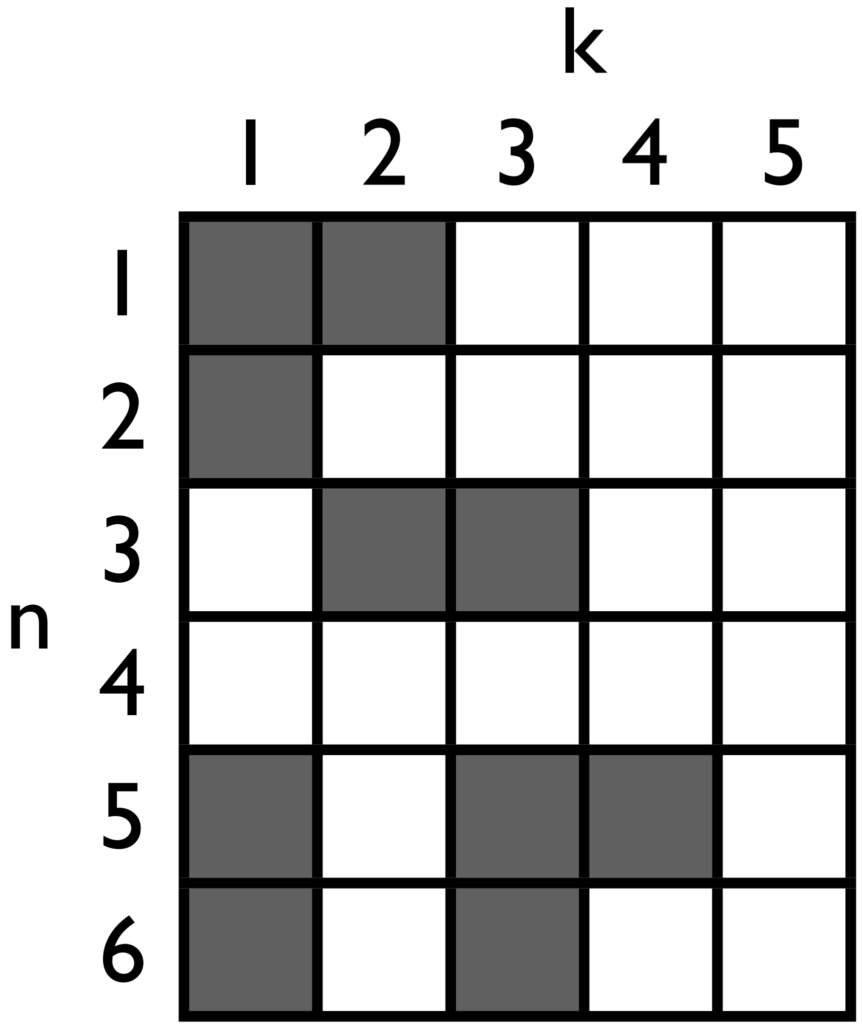
EFPF Example



Indian buffet process

$\mathbb{P}(F'_N = f'_N)$

EFPF Example

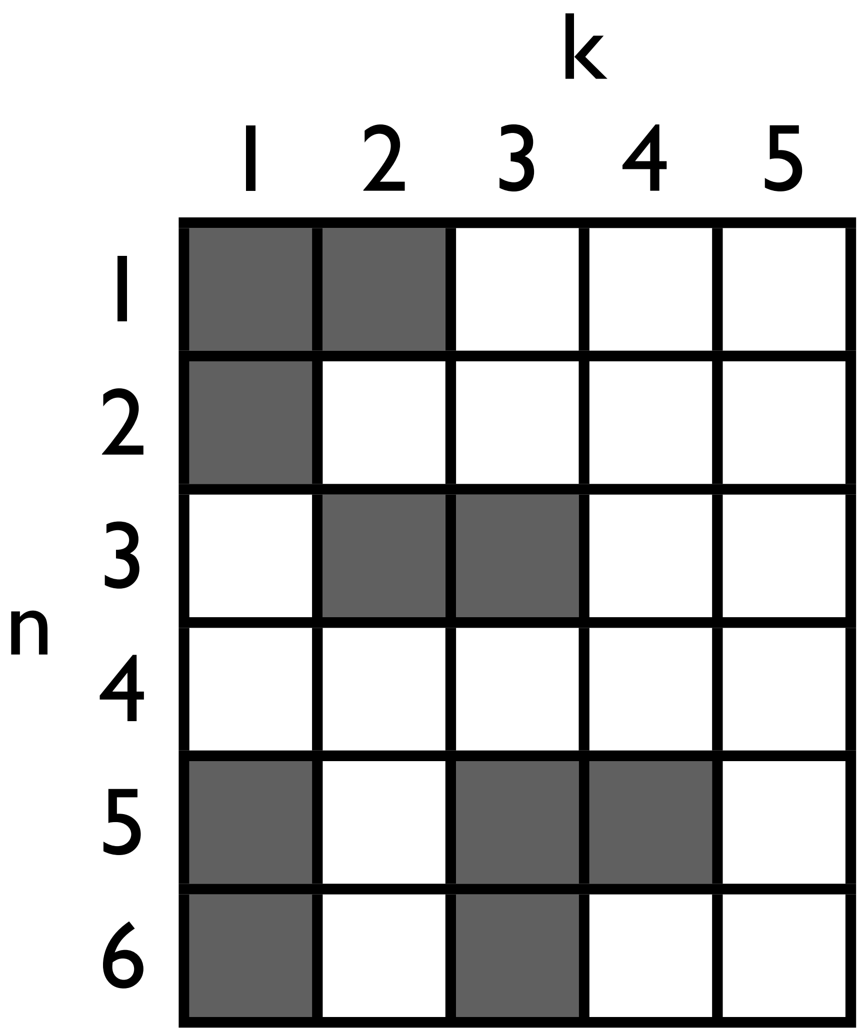


Indian buffet process

$$\mathbb{P}(F'_N = f'_N)$$

$$= \frac{1}{K_N!}(\theta\gamma)^{K_N} \exp\left(-\theta\gamma \sum_{n=1}^{N} (\theta + n - 1)^{-1}\right) \cdot \prod_{k=1}^{K_N} \frac{\Gamma(|A_k|)\Gamma(N - |A_k| + \theta)}{\Gamma(N+\theta)}$$

EFPF Example



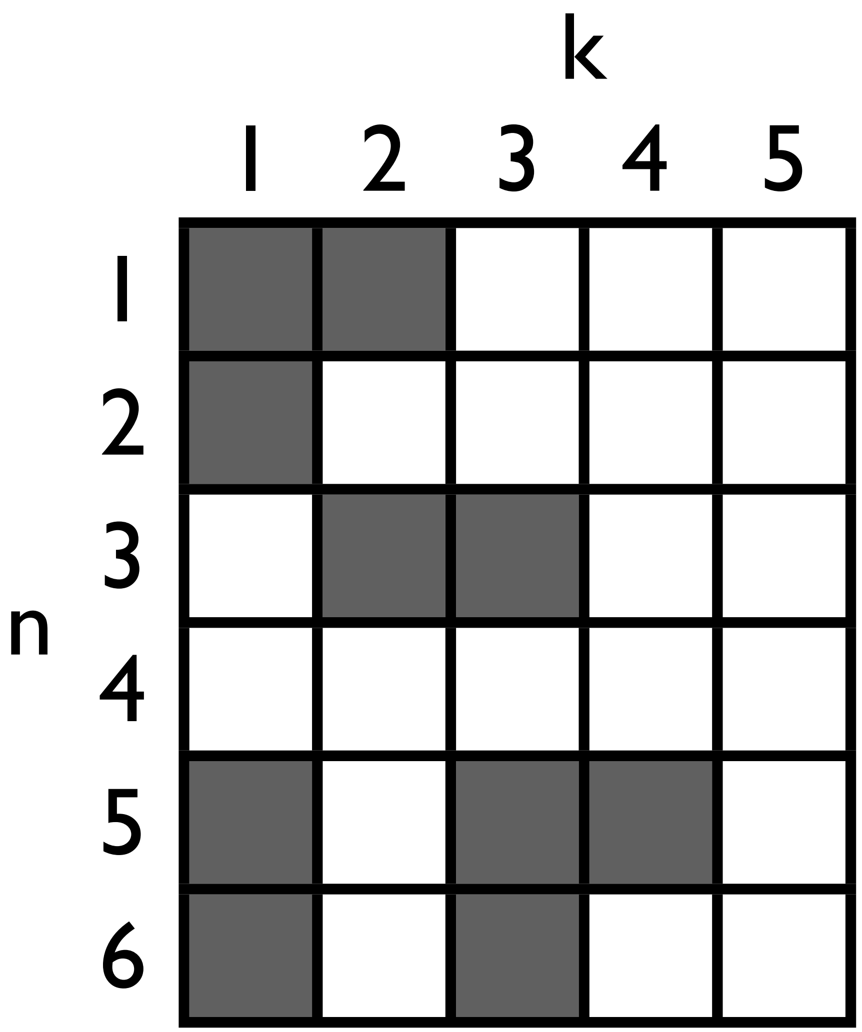
Indian buffet process

Size of kth feature

$$\mathbb{P}(F'_N = f'_N)$$

$$= \frac{1}{K_N!}(\theta\gamma)^{K_N} \exp\left(-\theta\gamma \sum_{n=1}^{N} (\theta + n - 1)^{-1}\right) \cdot \prod_{k=1}^{K_N} \frac{\Gamma(|A_k|)\Gamma(N - |A_k| + \theta)}{\Gamma(N + \theta)}$$

EFPF Example



Indian buffet process

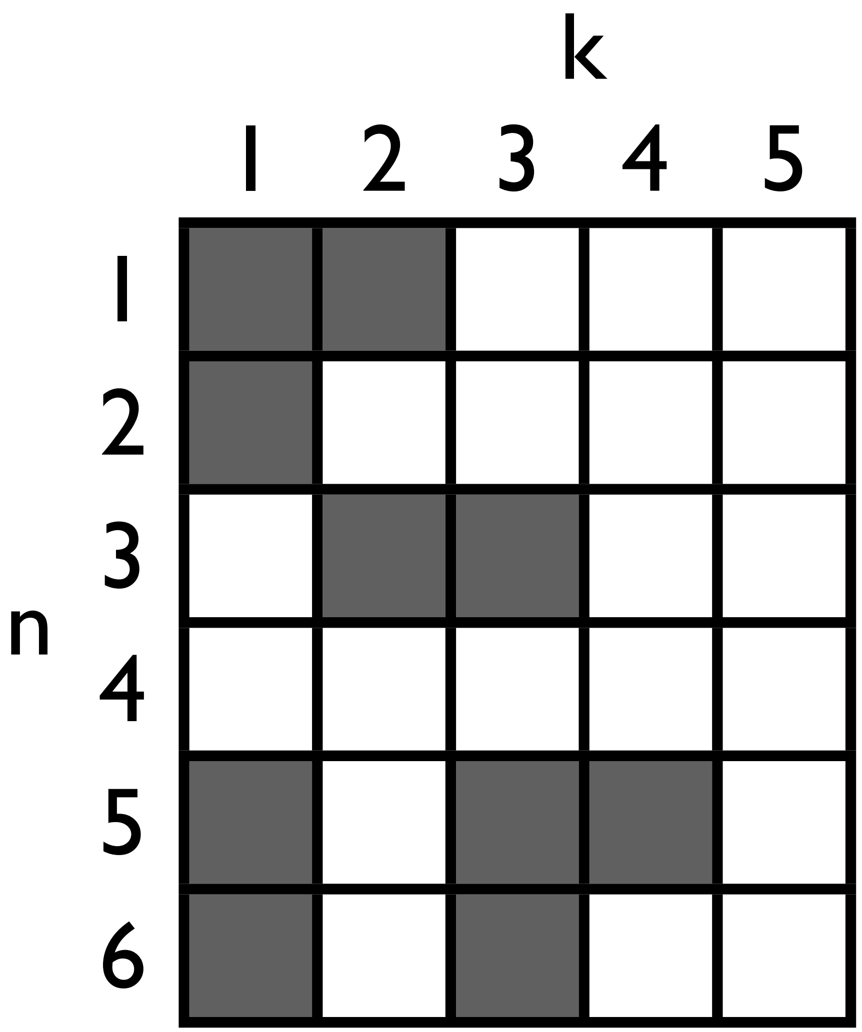
Size of *k*th feature

Number of features

$$\mathbb{P}(F'_N = f'_N)$$

$$= \frac{1}{K_N!}(\theta\gamma)^{K_N} \exp\left(-\theta\gamma \sum_{n=1}^{N} (\theta + n - 1)^{-1}\right) \cdot \prod_{k=1}^{K_N} \frac{\Gamma(|A_k|)\Gamma(N - |A_k| + \theta)}{\Gamma(N+\theta)}$$

EFPF Example



Indian buffet process

Number of data points

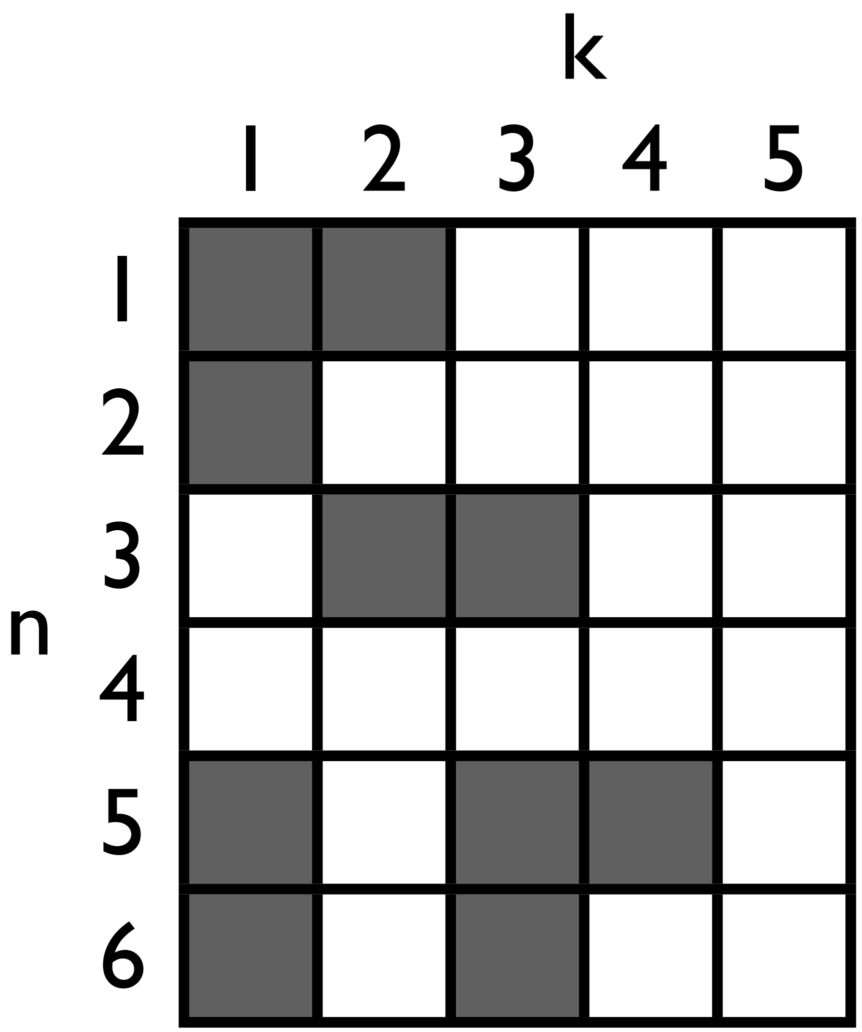
Size of kth feature

Number of features

$$\mathbb{P}(F'_N = f'_N)$$

$$= \frac{1}{K_N!}(\theta\gamma)^{K_N} \exp\left(-\theta\gamma \sum_{n=1}^{N} (\theta + n - 1)^{-1}\right) \cdot \prod_{k=1}^{K_N} \frac{\Gamma(|A_k|)\Gamma(N - |A_k| + \theta)}{\Gamma(N + \theta)}$$

EFPF Example



Indian buffet process

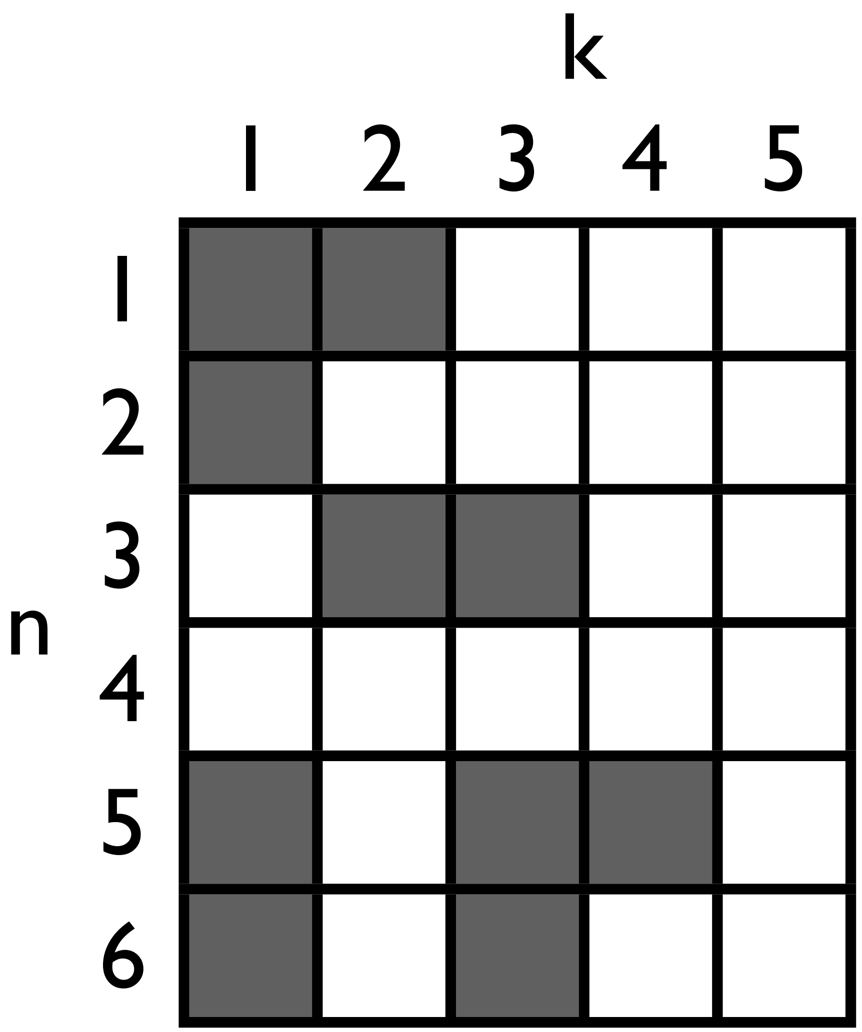
Number of data points

Size of *k*th feature

Number of features

$$
\begin{aligned}
&\mathbb{P}(F'_N = f'_N) \\
&= \frac{1}{K_N!}(\theta\gamma)^{K_N} \exp\left(-\theta\gamma \sum_{n=1}^{N} (\theta + n - 1)^{-1}\right) \cdot \prod_{k=1}^{K_N} \frac{\Gamma(|A_k|)\Gamma(N - |A_k| + \theta)}{\Gamma(N + \theta)} \\
&= p(N, |A_1|, \ldots, |A_K|)
\end{aligned}
$$

EFPF Example



Indian buffet process

Number of data points

Size of *k*th feature

Number of features

$$
\begin{aligned}
&\mathbb{P}(F'_N = f'_N) \\
&= \frac{1}{K_N!}(\theta\gamma)^{K_N} \exp\left(-\theta\gamma \sum_{n=1}^{N} (\theta + n - 1)^{-1}\right) \cdot \prod_{k=1}^{K_N} \frac{\Gamma(|A_k|)\Gamma(N - |A_k| + \theta)}{\Gamma(N+\theta)} \\
&= p(N, |A_1|, \ldots, |A_K|)
\end{aligned}
$$

p: symmetric in feature sizes

EFPF Example

Indian buffet process

Number of data points

Size of kth feature

Number of features

$$\mathbb{P}(F'_N = f'_N)$$

$$= \frac{1}{K_N!}(\theta\gamma)^{K_N} \exp\left(-\theta\gamma \sum_{n=1}^{N} (\theta + n - 1)^{-1}\right) \cdot \prod_{k=1}^{K_N} \frac{\Gamma(|A_k|)\Gamma(N - |A_k| + \theta)}{\Gamma(N + \theta)}$$

$$= p(N, |A_1|, \ldots, |A_K|)$$

p: symmetric in feature sizes

EFPF Example

Indian buffet process

- Exchangeable

Number of data points

Size of kth feature

Number of features

$$\mathbb{P}(F'_N = f'_N)$$

$$= \frac{1}{K_N!}(\theta\gamma)^{K_N} \exp\left(-\theta\gamma \sum_{n=1}^{N} (\theta + n - 1)^{-1}\right) \cdot \prod_{k=1}^{K_N} \frac{\Gamma(|A_k|)\Gamma(N - |A_k| + \theta)}{\Gamma(N + \theta)}$$

$$= p(N, |A_1|, \ldots, |A_K|)$$

p: symmetric in feature sizes

EFPF Example

Indian buffet process

- Exchangeable
- Consistent

Number of data points

Size of kth feature

Number of features

$$\mathbb{P}(F'_N = f'_N)$$

$$= \frac{1}{K_N!}(\theta\gamma)^{K_N} \exp\left(-\theta\gamma \sum_{n=1}^{N} (\theta + n - 1)^{-1}\right) \cdot \prod_{k=1}^{K_N} \frac{\Gamma(|A_k|)\Gamma(N - |A_k| + \theta)}{\Gamma(N + \theta)}$$

$$= p(N, |A_1|, \ldots, |A_K|)$$

p: symmetric in feature sizes

EFPF Example

Indian buffet process

- Exchangeable
- Consistent
- Unbounded number of features

Number of data points

Size of kth feature

Number of features

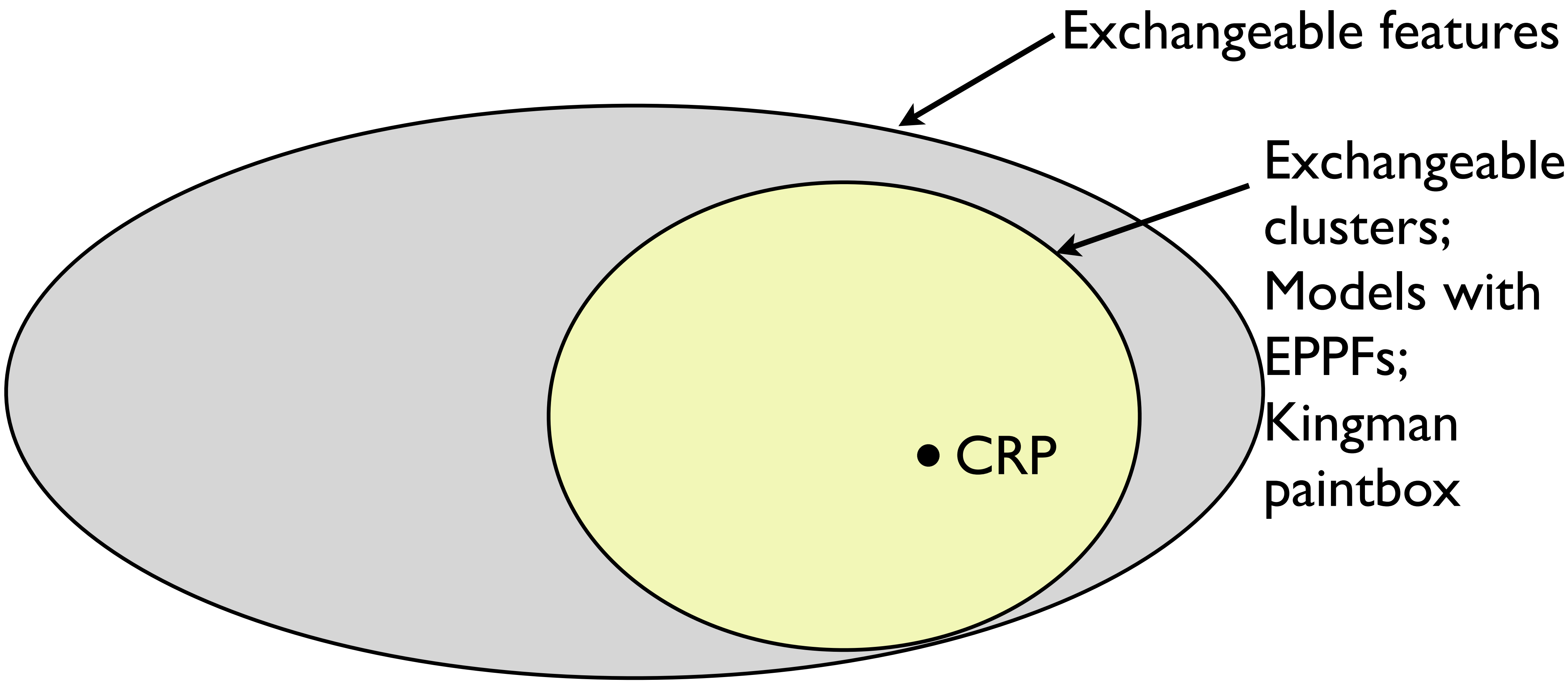
$$\mathbb{P}(F'_N = f'_N)$$

$$= \frac{1}{K_N!}(\theta\gamma)^{K_N} \exp\left(-\theta\gamma \sum_{n=1}^{N} (\theta + n - 1)^{-1}\right) \cdot \prod_{k=1}^{K_N} \frac{\Gamma(|A_k|)\Gamma(N - |A_k| + \theta)}{\Gamma(N + \theta)}$$

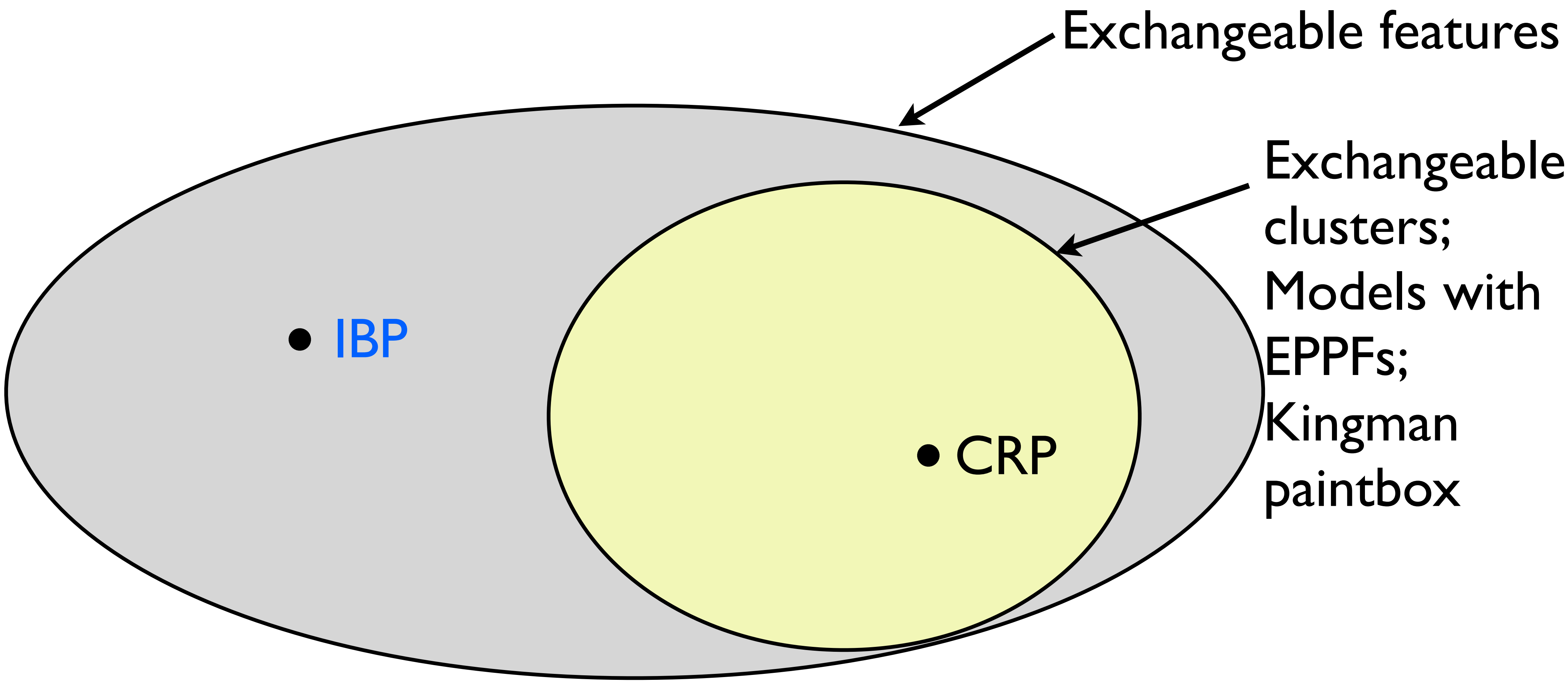
$$= p(N, |A_1|, \ldots, |A_K|)$$

p: symmetric in feature sizes

Clusters and features



Clusters and features



EFPF Example

Indian buffet process

Number of data points

Size of kth feature

Number of features

$$\mathbb{P}(F'_N = f'_N)$$

$$= \frac{1}{K_N!}(\theta\gamma)^{K_N} \exp\left(-\theta\gamma \sum_{n=1}^{N} (\theta + n - 1)^{-1}\right) \cdot \prod_{k=1}^{K_N} \frac{\Gamma(|A_k|)\Gamma(N - |A_k| + \theta)}{\Gamma(N + \theta)}$$

$$= p(N, |A_1|, \ldots, |A_K|)$$

p: symmetric in feature sizes

EFPF Example

Indian buffet process

Number of data points

Size of *k*th feature

Number of features

$$\mathbb{P}(F'_N = f'_N)$$

$$= \frac{1}{K_N!} (\theta\gamma)^{K_N} \exp\left(-\theta\gamma \sum_{n=1}^{N} (\theta + n - 1)^{-1}\right) \cdot \prod_{k=1}^{K_N} \frac{\Gamma(|A_k|)\Gamma(N - |A_k| + \theta)}{\Gamma(N + \theta)}$$

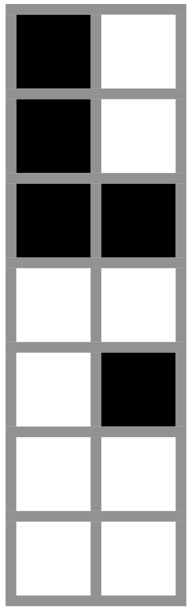
$$= p(N, |A_1|, \ldots, |A_K|)$$ “EFPF”

p: symmetric in feature sizes

[Broderick, Jordan, Pitman 2012]

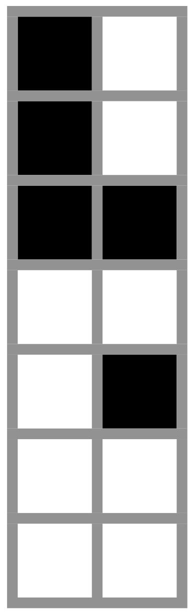
No EFPF

Counterexample

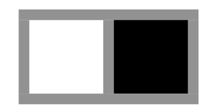


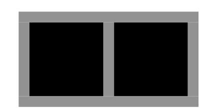
No EFPF

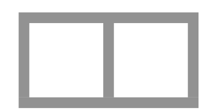
Counterexample



$\mathbb{P}(\text{row} =$  $) = p_1$

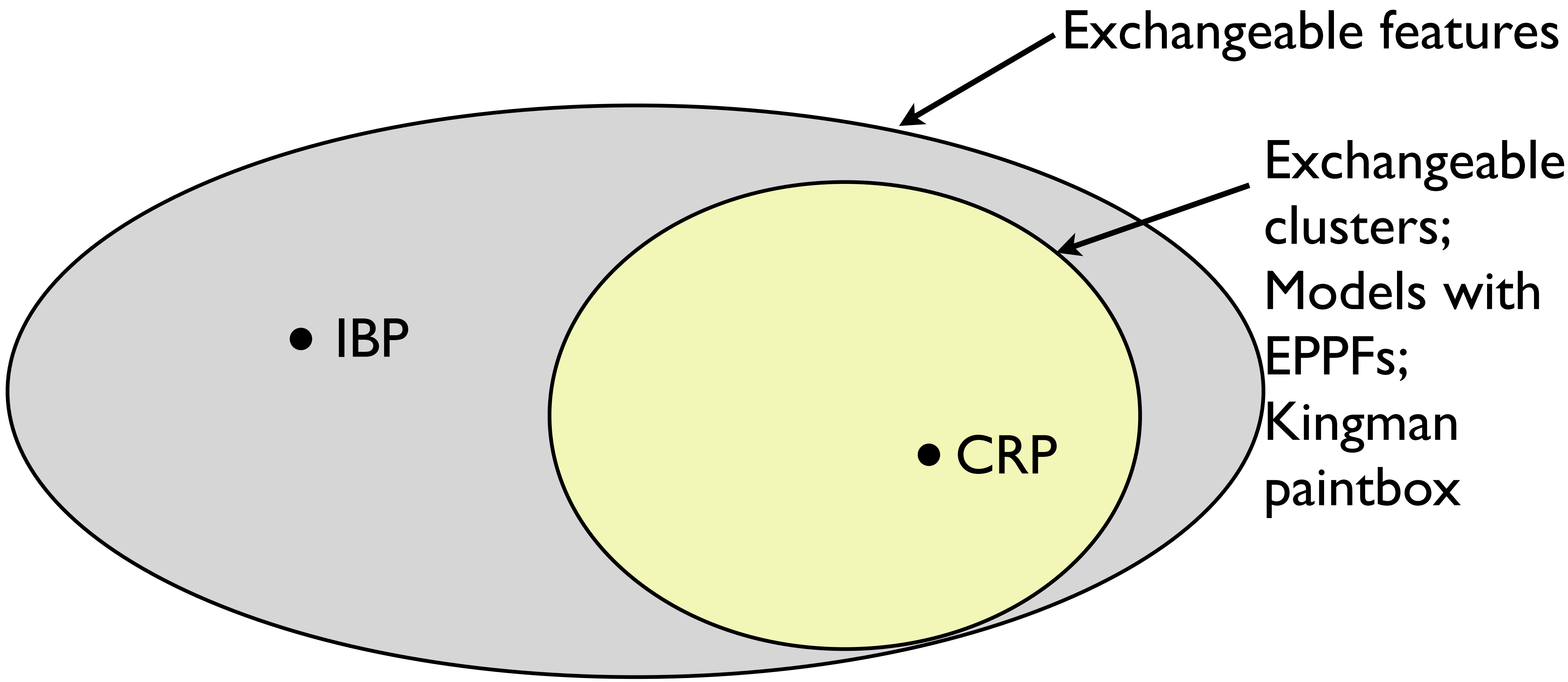
$\mathbb{P}(\text{row} =$  $) = p_2$

$\mathbb{P}(\text{row} =$  $) = p_3$

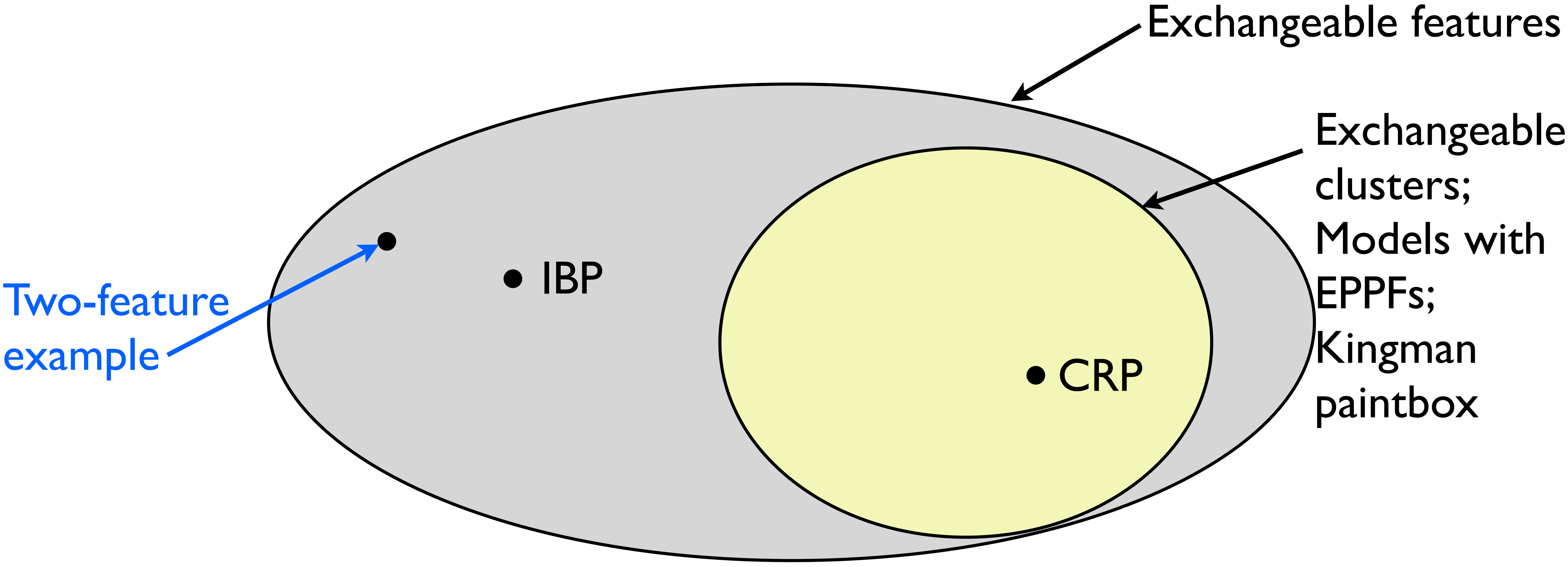
$\mathbb{P}(\text{row} =$  $) = p_4$

[Broderick, Jordan, Pitman 2012]

Clusters and features

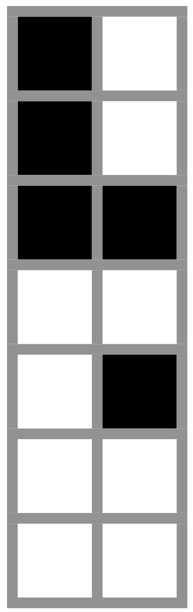


Clusters and features



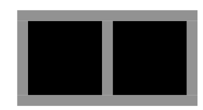
No EFPF

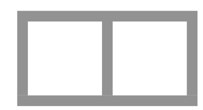
Counterexample

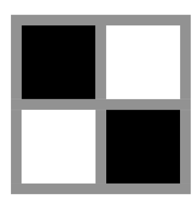
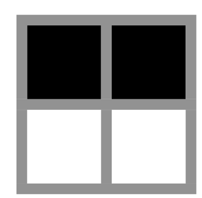


$\mathbb{P}(\text{row} =$  $) = p_1$

$\mathbb{P}(\text{row} =$  $) = p_2$

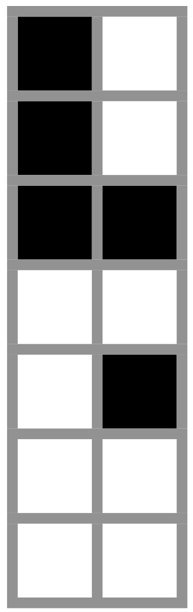
$\mathbb{P}(\text{row} =$  $) = p_3$

$\mathbb{P}(\text{row} =$  $) = p_4$

$\mathbb{P}($  $)$ $\mathbb{P}($  $)$

No EFPF

Counterexample



$$\mathbb{P}(\text{row} = \blacksquare\square) = p_1$$

$$\mathbb{P}(\text{row} = \square\blacksquare) = p_2$$

$$\mathbb{P}(\text{row} = \blacksquare\blacksquare) = p_3$$

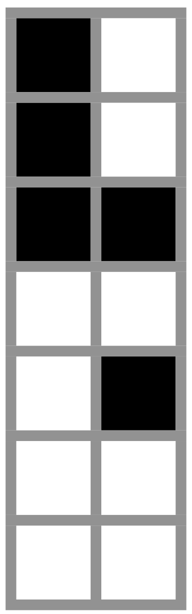
$$\mathbb{P}(\text{row} = \square\square) = p_4$$

$$\mathbb{P}\left(\begin{matrix}\blacksquare & \square \\ \square & \blacksquare\end{matrix}\right) \qquad \mathbb{P}\left(\begin{matrix}\blacksquare & \blacksquare \\ \square & \square\end{matrix}\right)$$

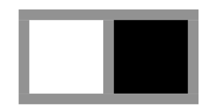
$$p_1 p_2 \qquad p_3 p_4$$

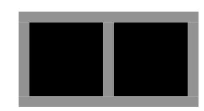
No EFPF

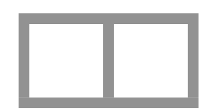
Counterexample

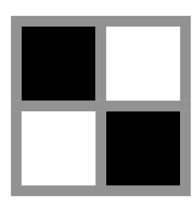
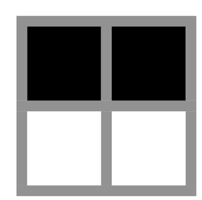


$\mathbb{P}(\text{row} =$  $) = p_1$

$\mathbb{P}(\text{row} =$  $) = p_2$

$\mathbb{P}(\text{row} =$  $) = p_3$

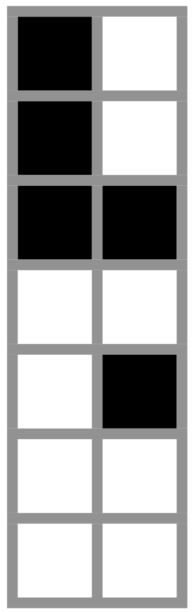
$\mathbb{P}(\text{row} =$  $) = p_4$

$\mathbb{P}($  $) \neq \mathbb{P}($  $)$

$$p_1 p_2 \neq p_3 p_4$$

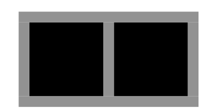
No EFPF

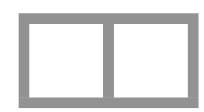
Counterexample

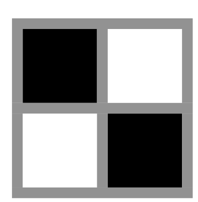
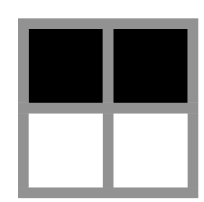


$\mathbb{P}(\text{row} =$  $) = p_1$

$\mathbb{P}(\text{row} =$  $) = p_2$

$\mathbb{P}(\text{row} =$  $) = p_3$

$\mathbb{P}(\text{row} =$  $) = p_4$

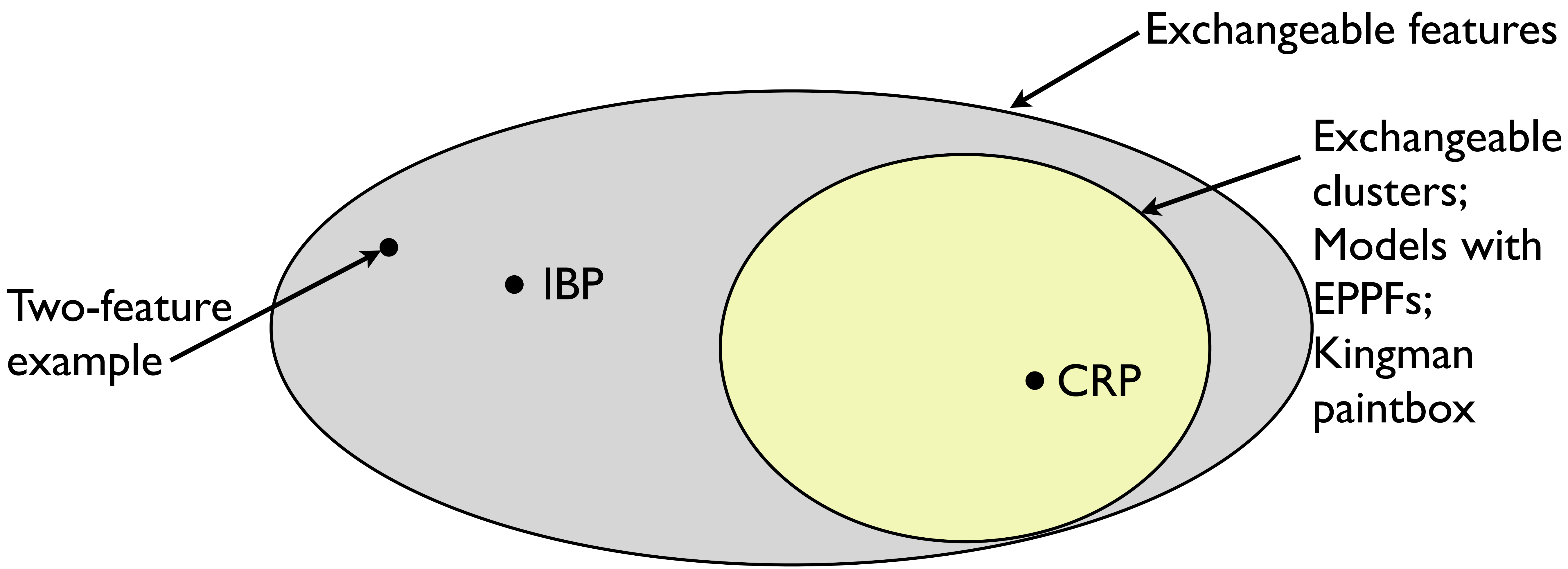
$\mathbb{P}($  $) \neq \mathbb{P}($  $)$

$$p_1 p_2 \neq p_3 p_4$$

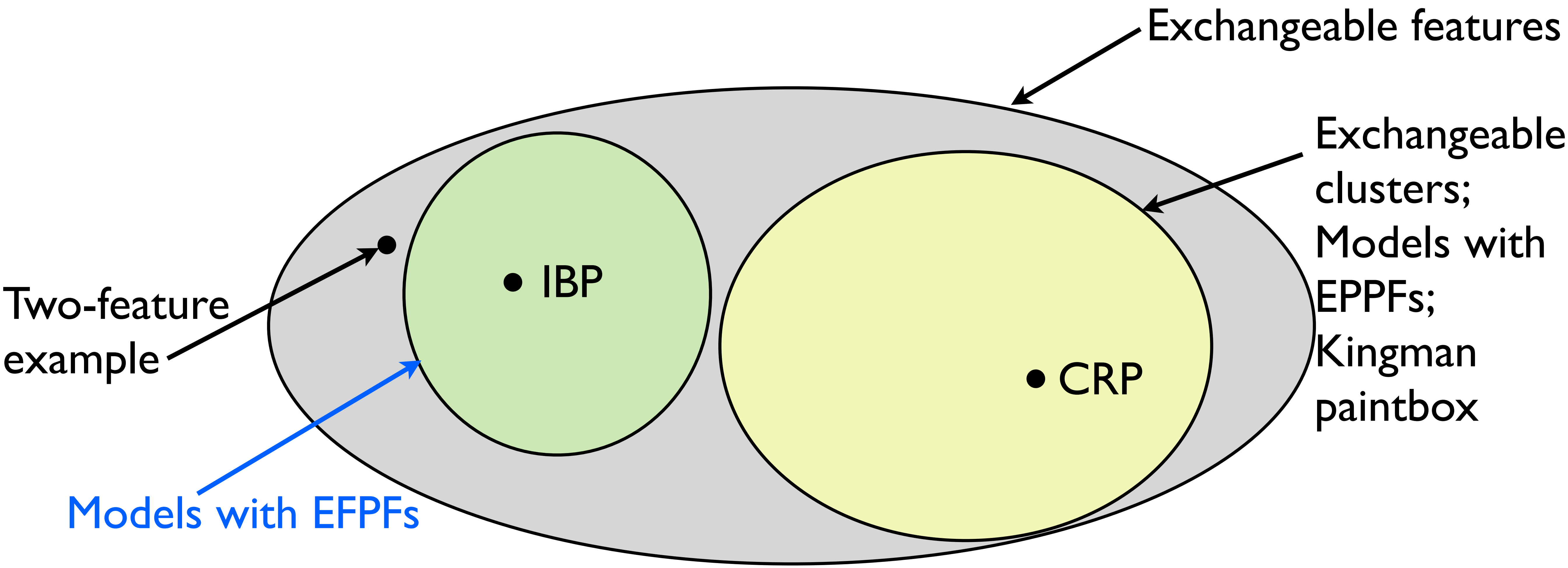
(EFPF)

$$\mathbb{P}(F'_N = f'_N) = p(N, |A_1|, |A_2|, \ldots, |A_K|)$$

Clusters and features



Clusters and features



Outline

I. Clusters

II. Features

- Overview
- Distribution
- Frequencies
- Random measure
- Paintbox

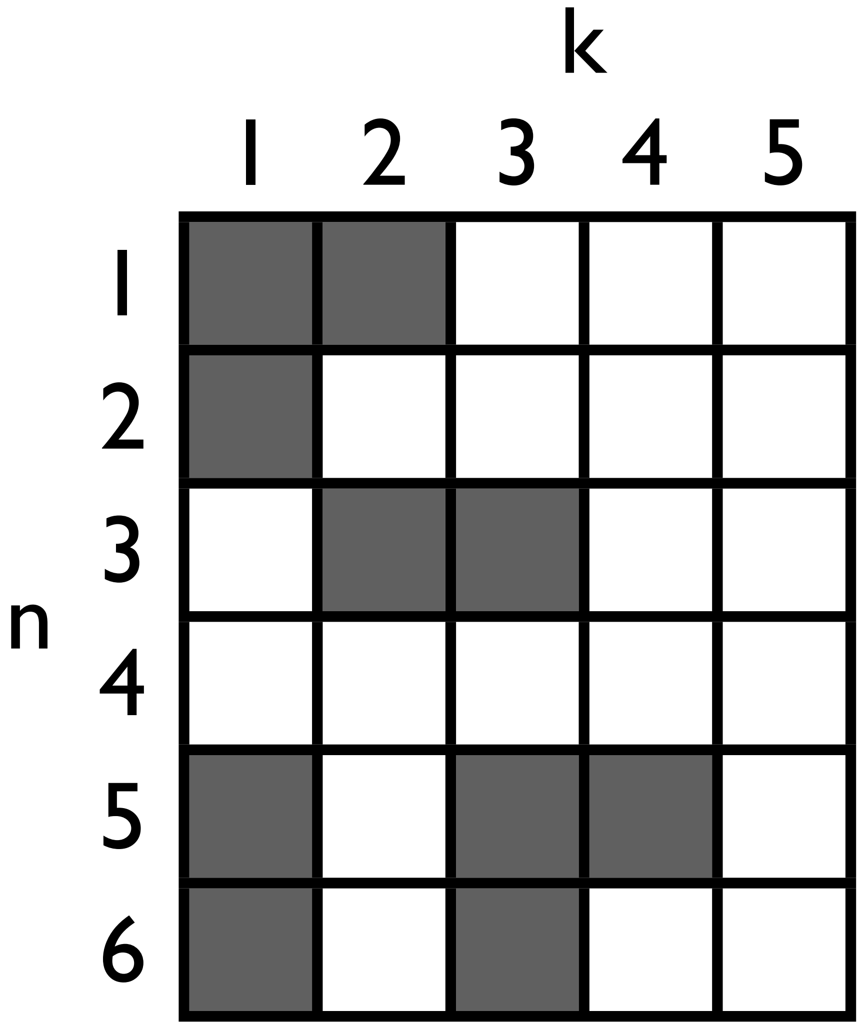
Outline

I. Clusters

II. Features

- Overview
- Distribution
- Frequencies
- Random measure
- Paintbox

IBP as Polya urns



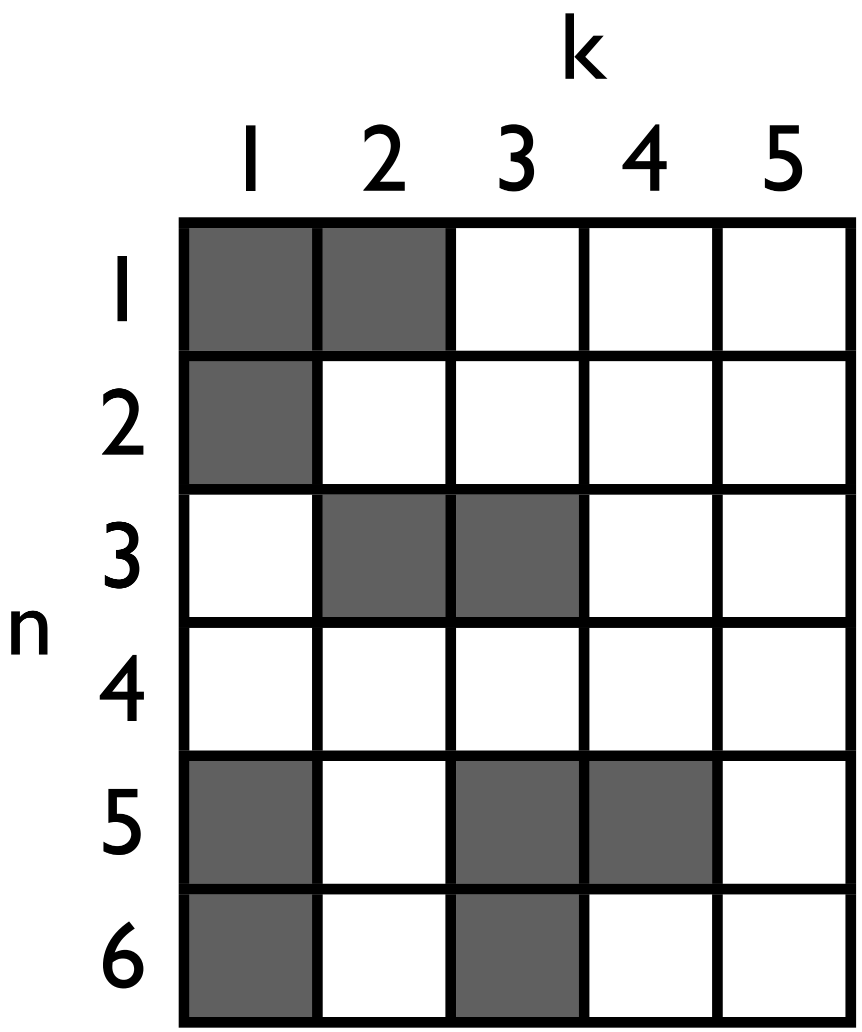
Indian buffet process

For n = 1, 2, ..., N

1. Customer *n* samples an existing dish *k* that has already been sampled $M_{n-1,k}$ times with probability $\dfrac{M_{n-1,k}}{\theta + n - 1}$
2. Number of new dishes for index *n*:

$$K_n^+ = \mathrm{Poisson}\left(\gamma \frac{\theta}{\theta + n - 1}\right)$$

IBP as Polya urns



Indian buffet process

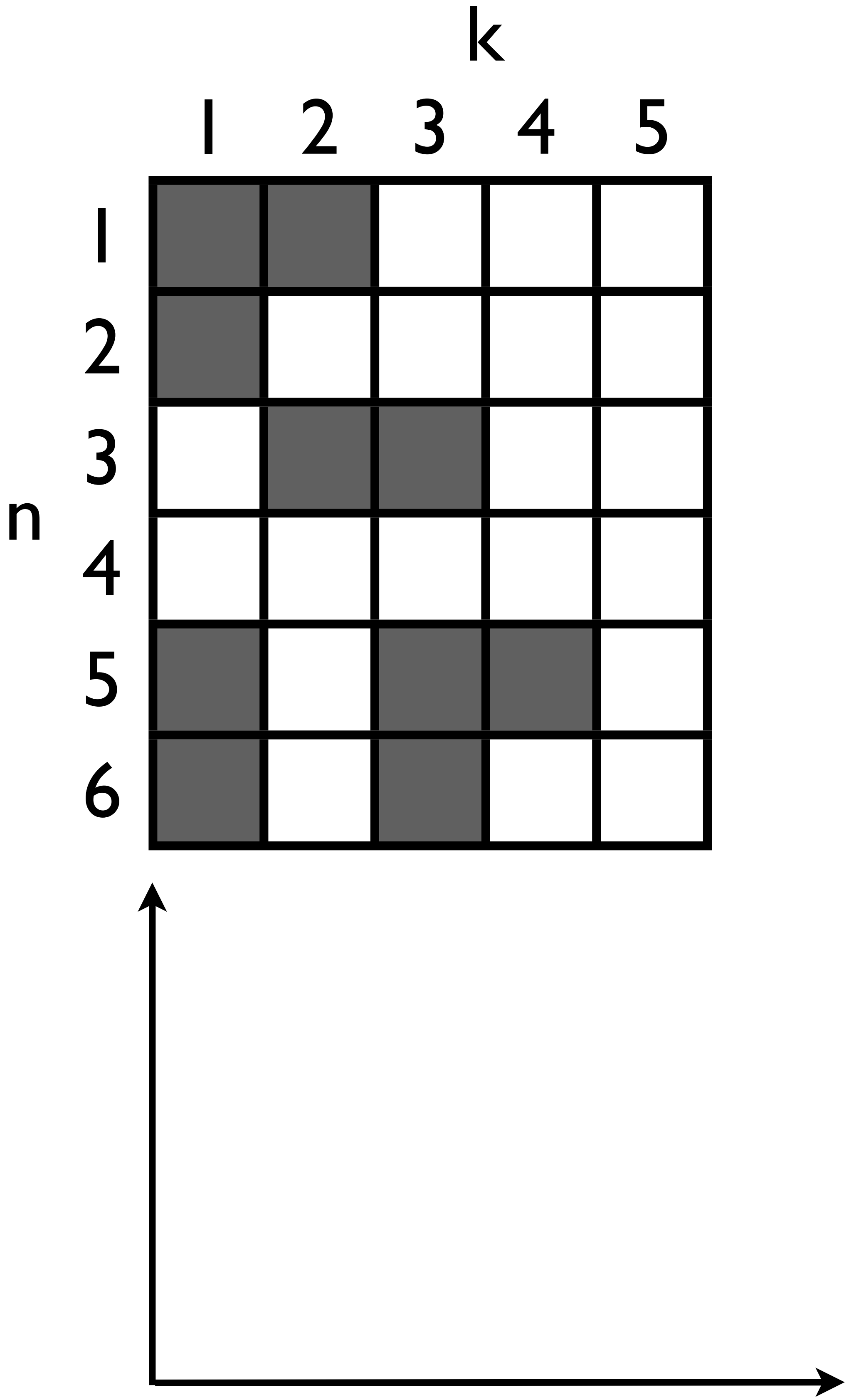
For n = 1, 2, ..., N

1. Customer *n* samples an existing dish *k* that has already been sampled $M_{n-1,k}$ times with probability $\dfrac{M_{n-1,k}}{\theta + n - 1}$

2. Number of new dishes for index *n*:

$$K_n^+ = \mathrm{Poisson}\left(\gamma \frac{\theta}{\theta + n - 1}\right)$$

IBP as Polya urns



Indian buffet process

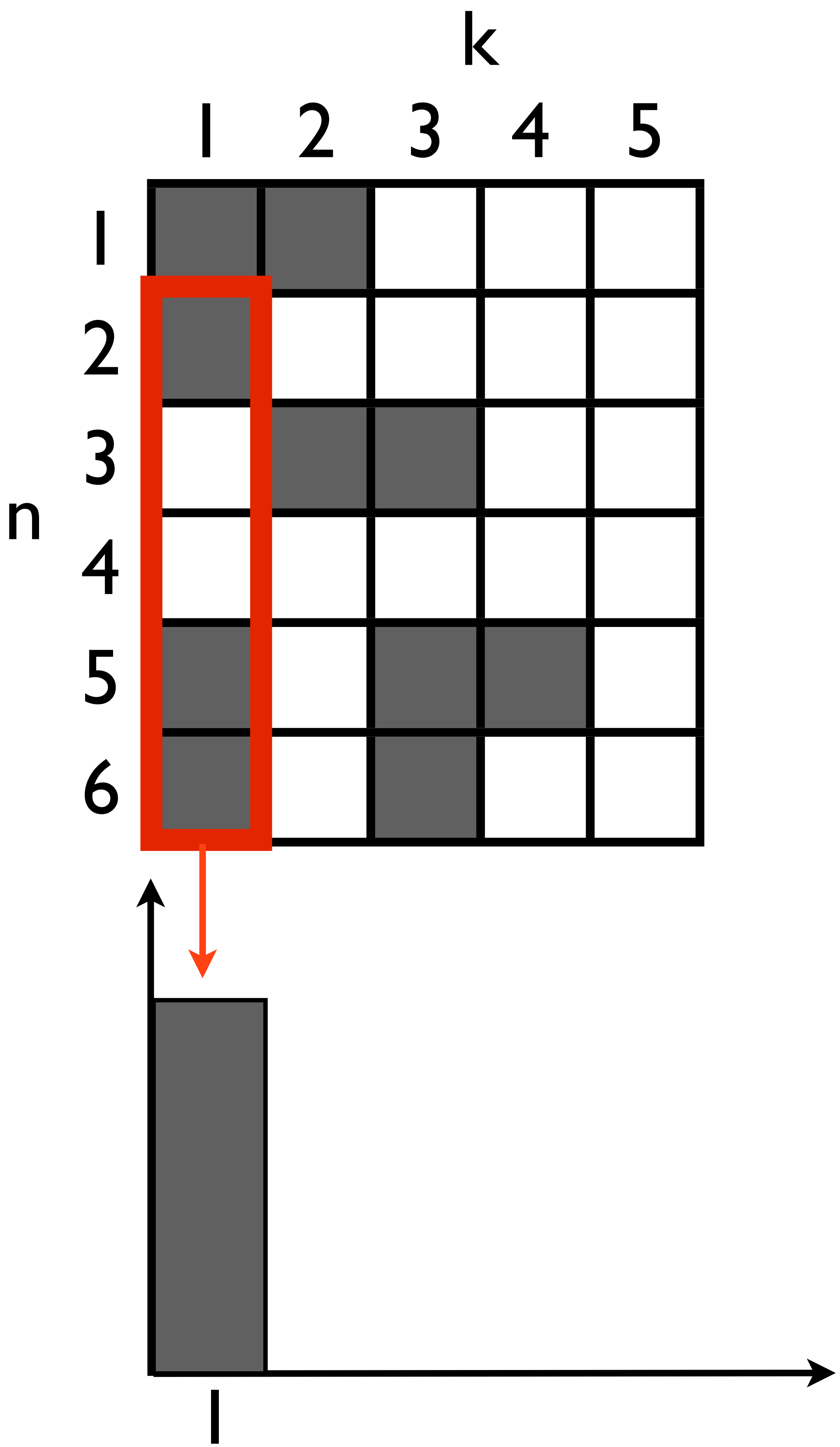
For n = 1, 2, ..., N

1. Customer *n* samples an existing dish *k* that has already been sampled $M_{n-1,k}$ times with probability $\dfrac{M_{n-1,k}}{\theta + n - 1}$

2. Number of new dishes for index *n*:

$$K_n^+ = \mathrm{Poisson}\left(\gamma \frac{\theta}{\theta + n - 1}\right)$$

IBP as Polya urns



Indian buffet process

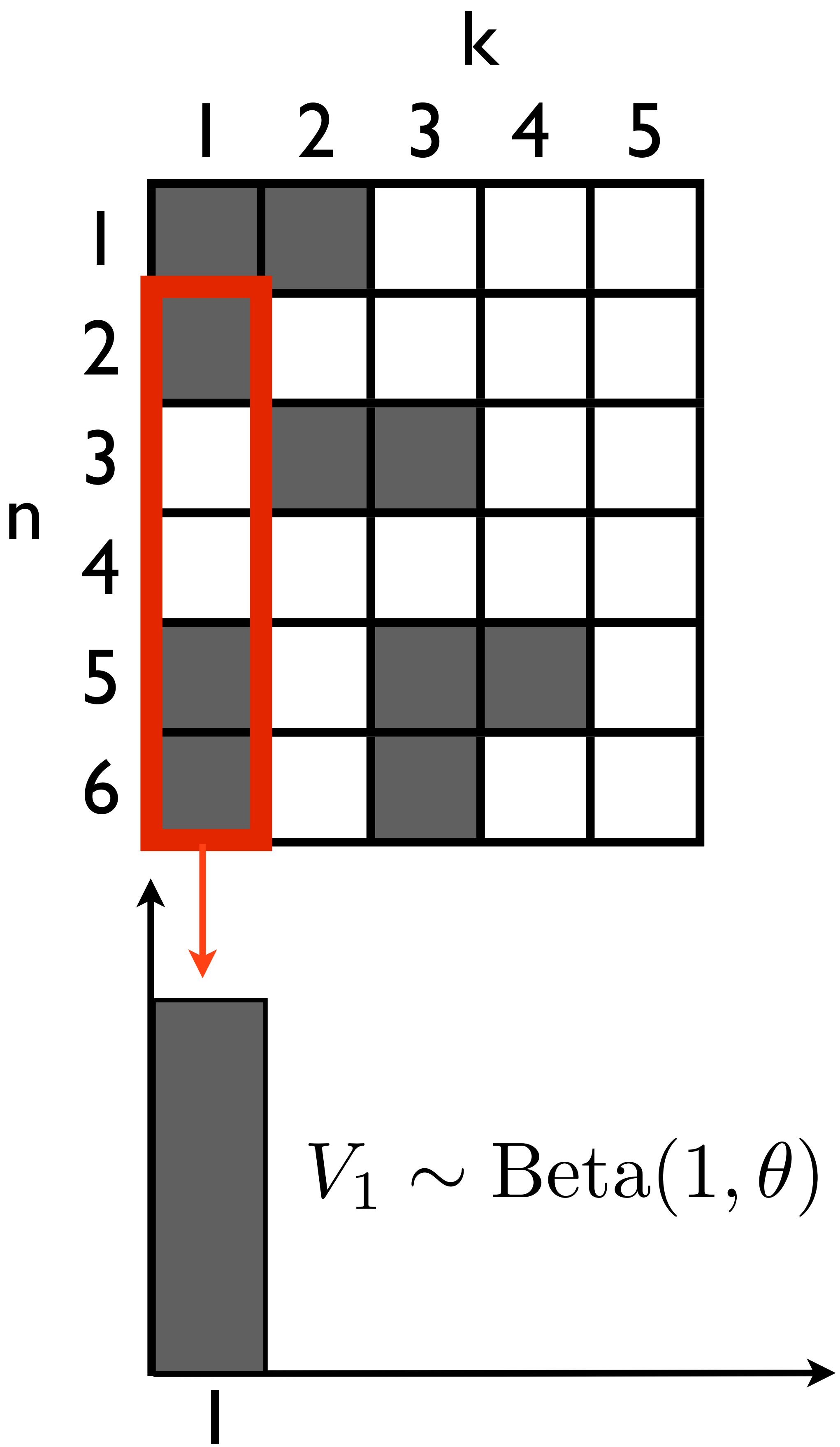
For n = 1, 2, ..., N

1. Customer *n* samples an existing dish *k* that has already been sampled $M_{n-1,k}$ times with probability $\dfrac{M_{n-1,k}}{\theta + n - 1}$

2. Number of new dishes for index *n*:

$$K_n^+ = \text{Poisson}\left(\gamma \frac{\theta}{\theta + n - 1}\right)$$

IBP as Polya urns



Indian buffet process

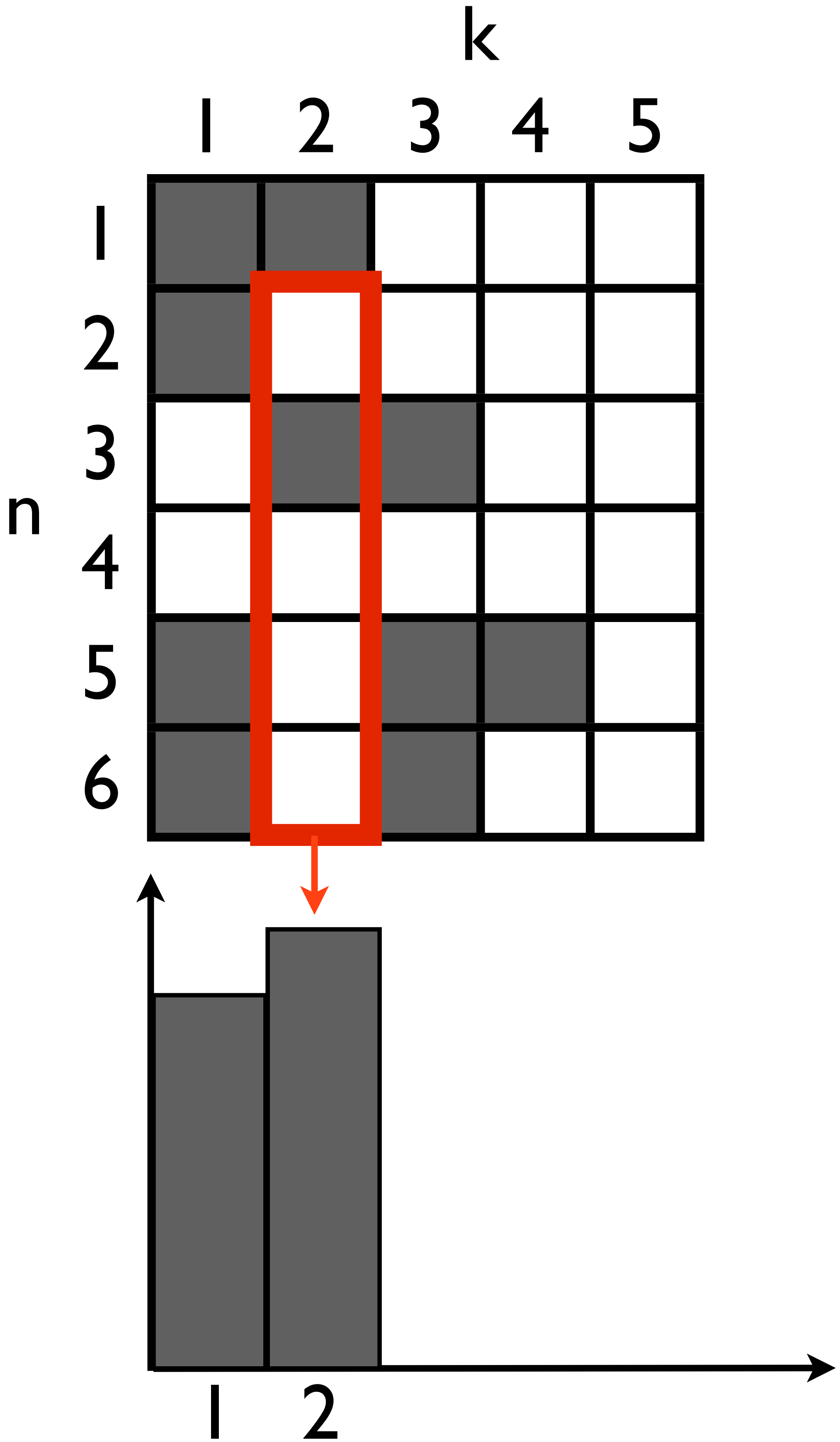
For n = 1, 2, ..., N

1. Customer *n* samples an existing dish *k* that has already been sampled $M_{n-1,k}$ times with probability $\dfrac{M_{n-1,k}}{\theta + n - 1}$

2. Number of new dishes for index *n*:

$$K_n^+ = \mathrm{Poisson}\left(\gamma \frac{\theta}{\theta + n - 1}\right)$$

IBP as Polya urns



Indian buffet process

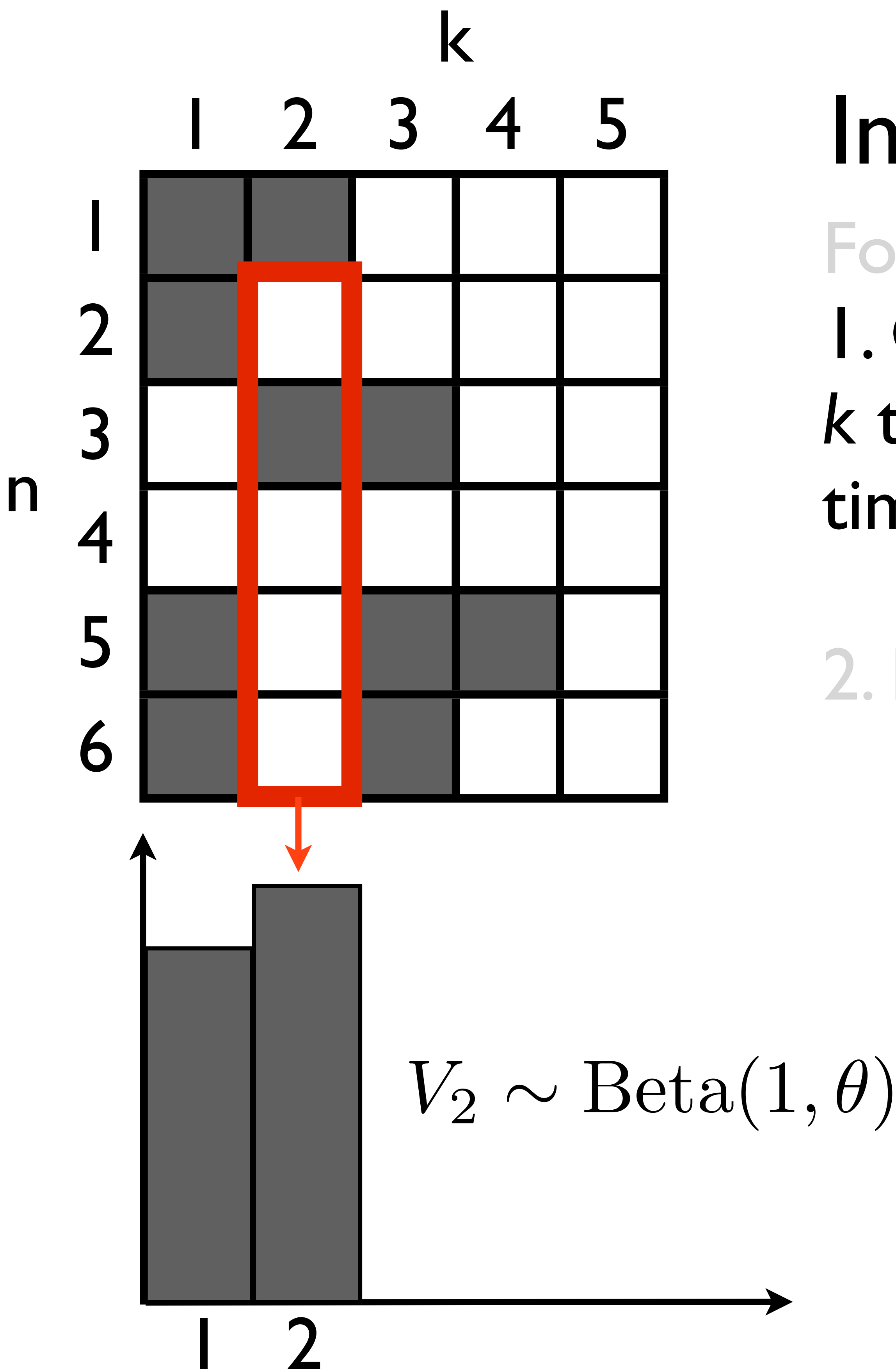
For n = 1, 2, ..., N

1. Customer *n* samples an existing dish *k* that has already been sampled $M_{n-1,k}$ times with probability $\dfrac{M_{n-1,k}}{\theta + n - 1}$

2. Number of new dishes for index *n*:

$$K_n^+ = \text{Poisson}\left(\gamma \frac{\theta}{\theta + n - 1}\right)$$

IBP as Polya urns



Indian buffet process

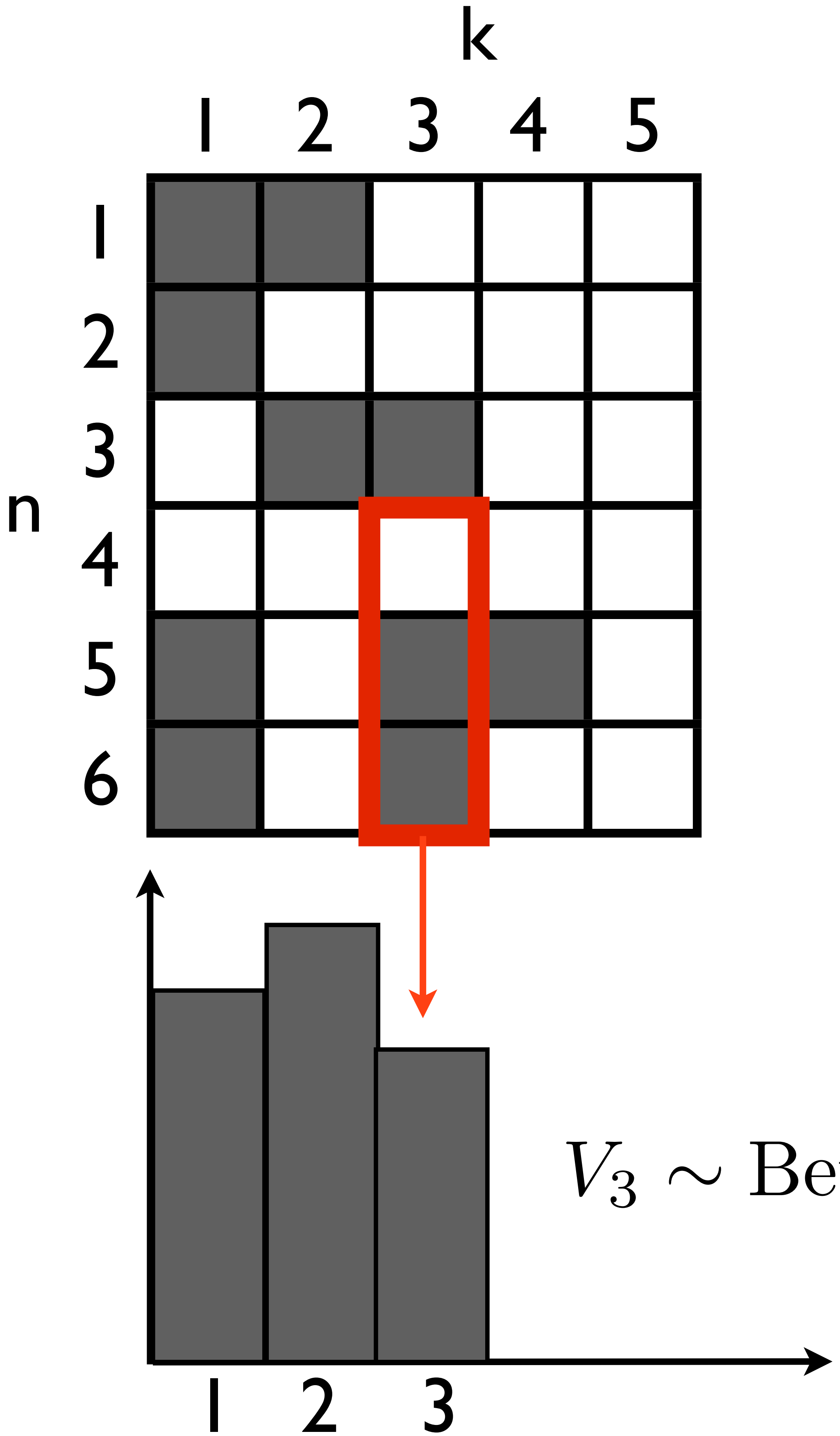
For n = 1, 2, ..., N

1. Customer *n* samples an existing dish *k* that has already been sampled $M_{n-1,k}$ times with probability $\dfrac{M_{n-1,k}}{\theta + n - 1}$

2. Number of new dishes for index *n*:

$$K_n^+ = \mathrm{Poisson}\left(\gamma \frac{\theta}{\theta + n - 1}\right)$$

IBP as Polya urns



Indian buffet process

For n = 1, 2, ..., N

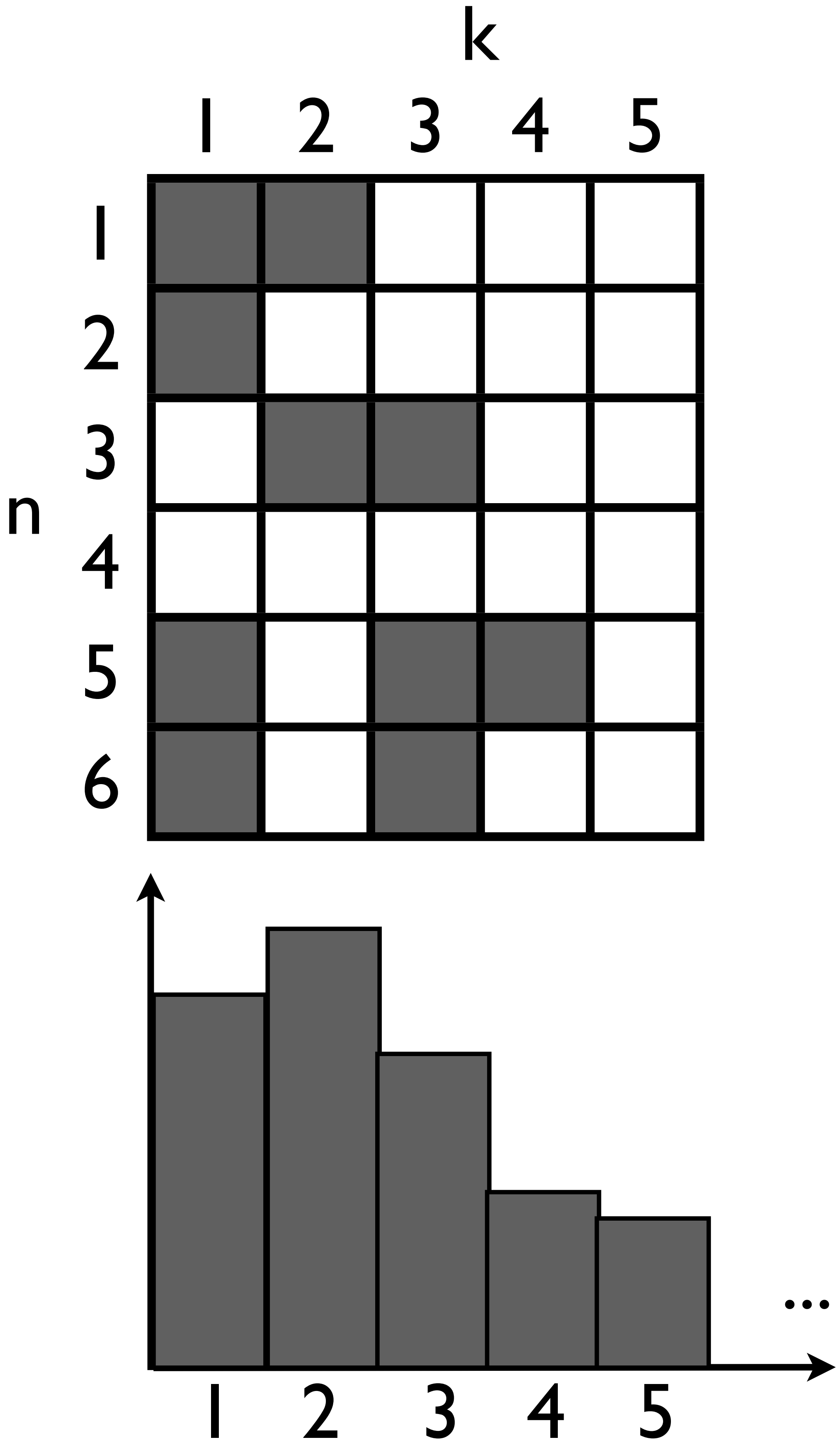
1. Customer *n* samples an existing dish *k* that has already been sampled $M_{n-1,k}$ times with probability $\dfrac{M_{n-1,k}}{\theta + n - 1}$

2. Number of new dishes for index *n*:

$$K_n^+ = \text{Poisson}\left(\gamma \frac{\theta}{\theta + n - 1}\right)$$

$$V_3 \sim \text{Beta}(1, 2 + \theta)$$

IBP as Polya urns



Indian buffet process

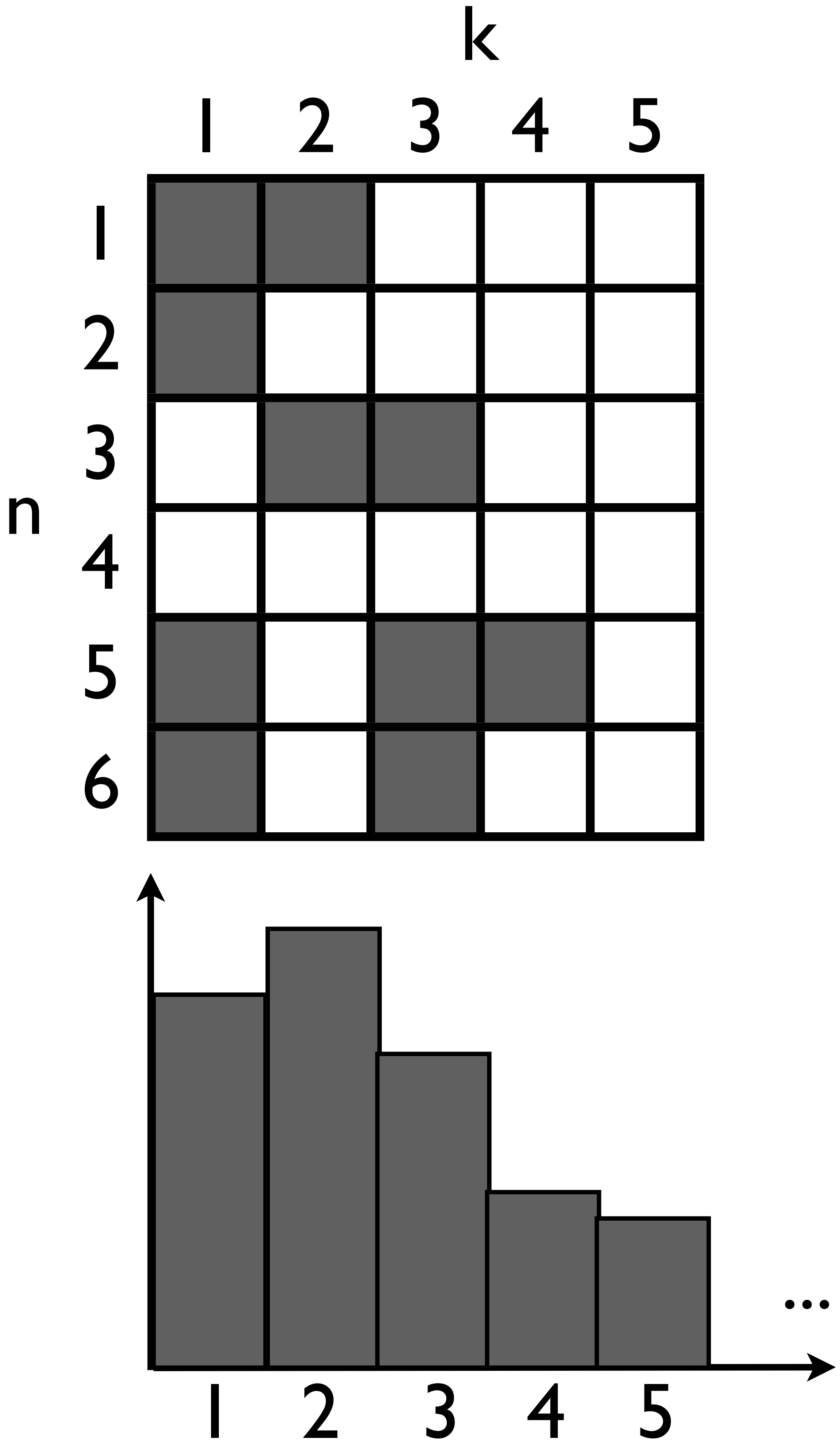
For n = 1, 2, ..., N

1. Customer *n* samples an existing dish *k* that has already been sampled $M_{n-1,k}$ times with probability $\dfrac{M_{n-1,k}}{\theta + n - 1}$

2. Number of new dishes for index *n*:

$$K_n^+ = \mathrm{Poisson}\left(\gamma \frac{\theta}{\theta + n - 1}\right)$$

IBP as Polya urns



Indian buffet process

For n = 1, 2, ..., N

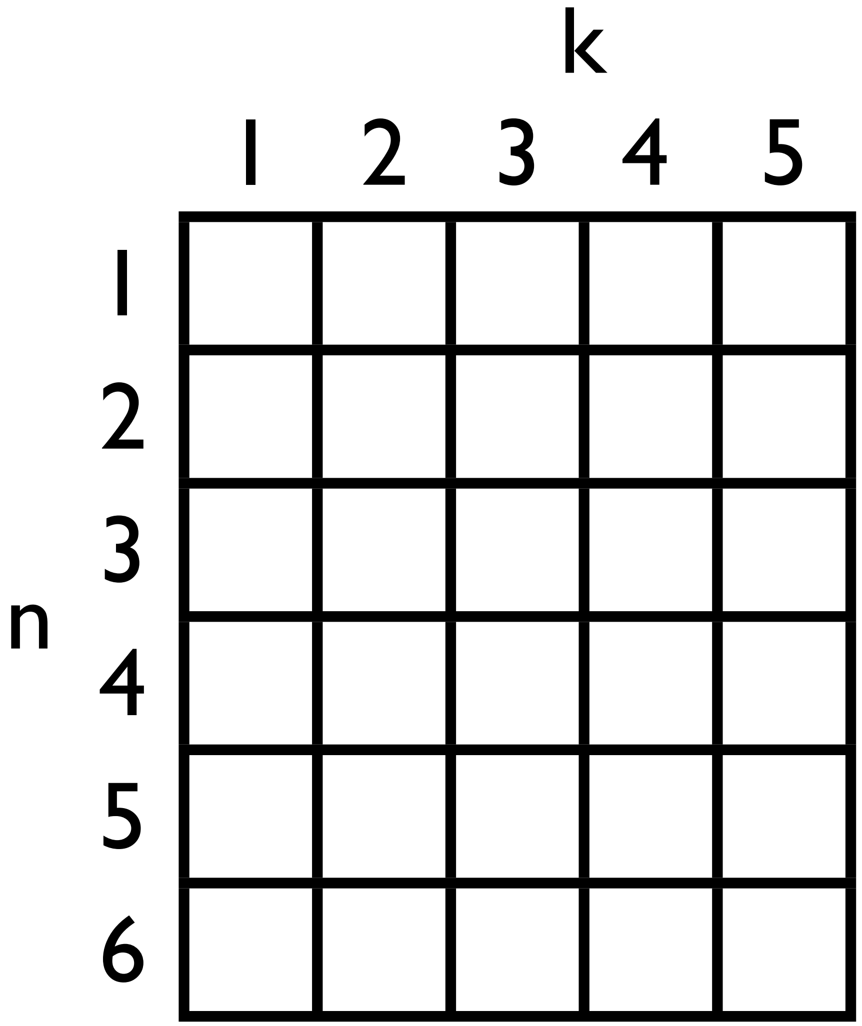
1. Customer *n* samples an existing dish *k* that has already been sampled $M_{n-1,k}$ times with probability $\dfrac{M_{n-1,k}}{\theta + n - 1}$

2. Number of new dishes for index *n*:

$$K_n^+ = \mathrm{Poisson}\left(\gamma \frac{\theta}{\theta + n - 1}\right)$$

Not a distribution

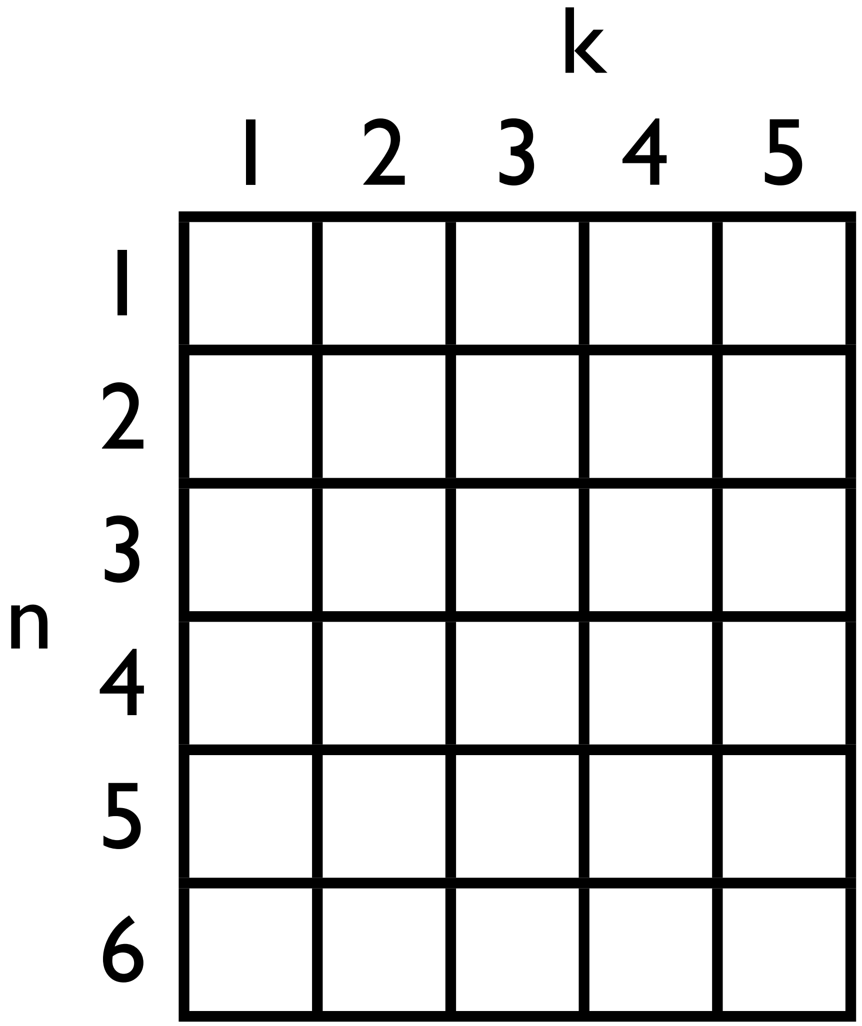
IBP as Polya urns



Another way to generate the IBP:

[Hjort 1990; Thibaux, Jordan 2007]

IBP as Polya urns

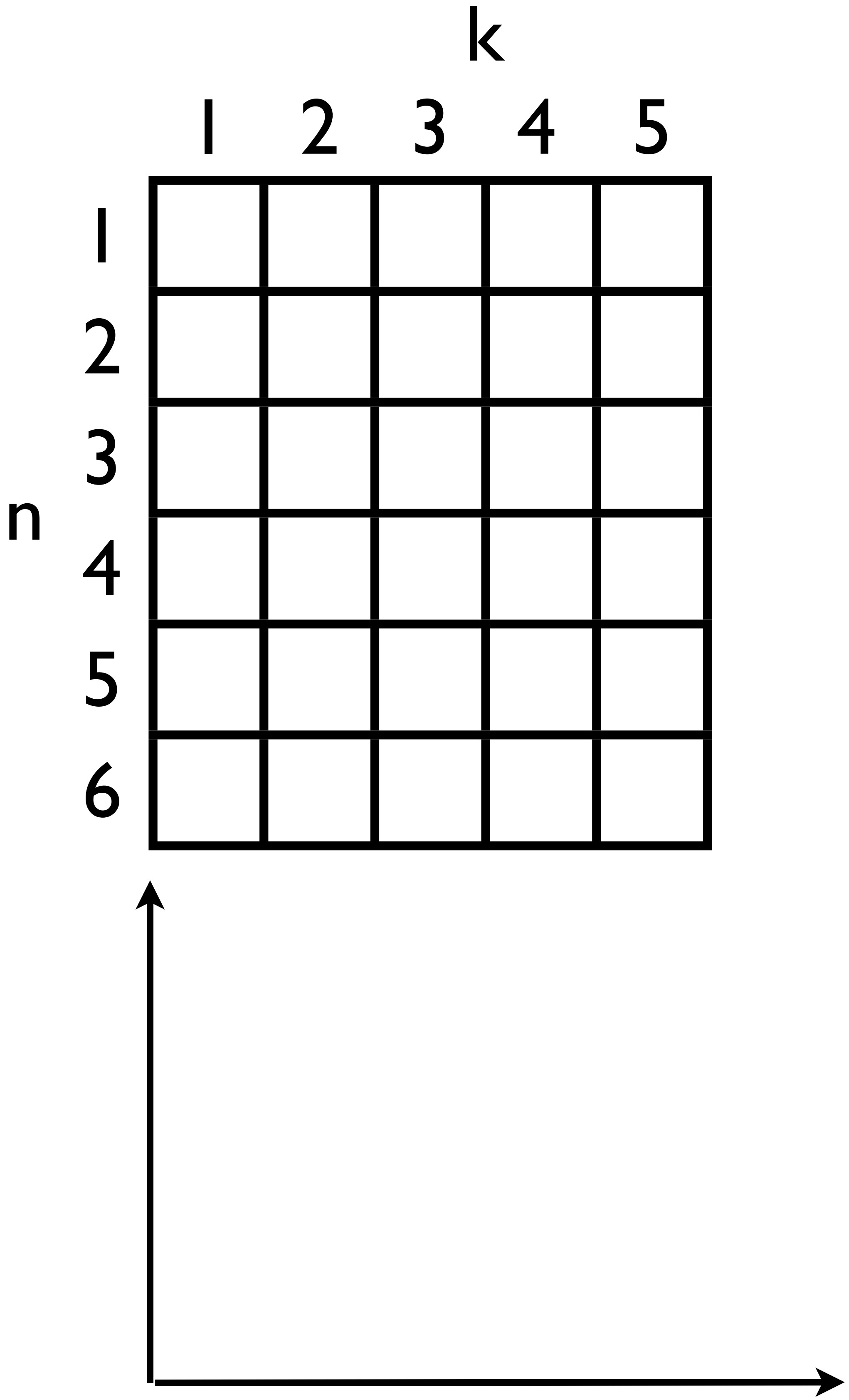


Another way to generate the IBP:

1. Generate feature frequencies

[Hjort 1990; Thibaux, Jordan 2007]

IBP as Polya urns

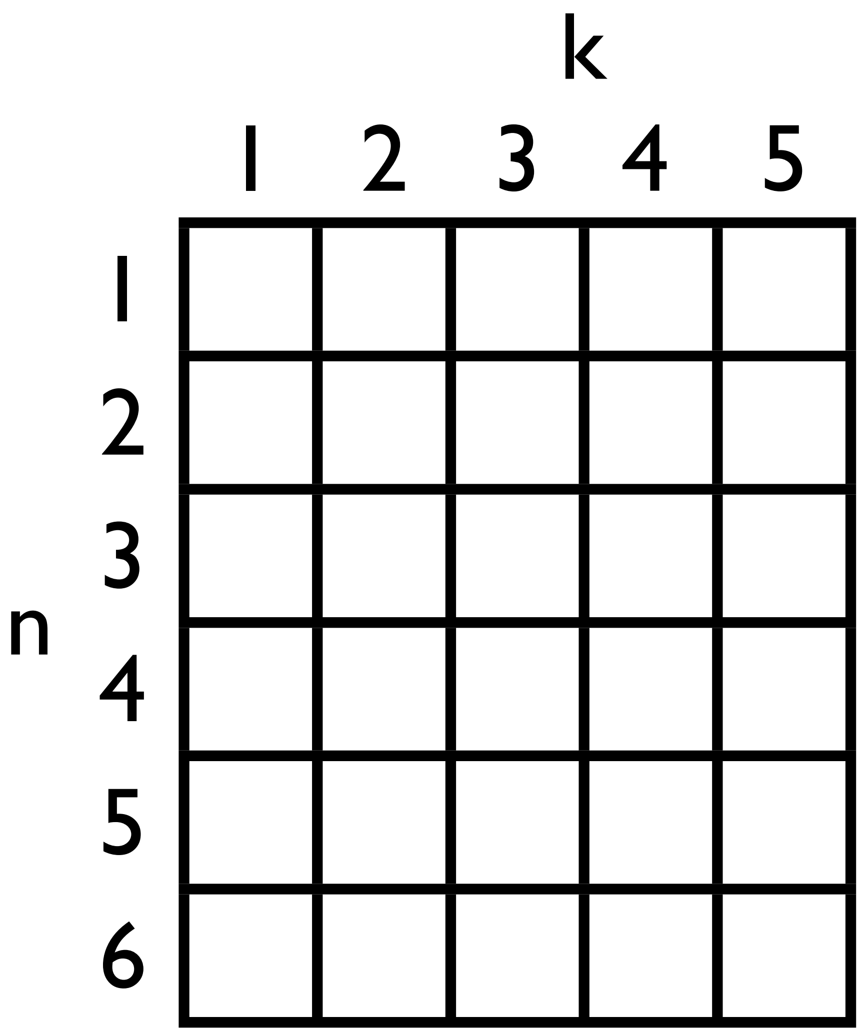


Another way to generate the IBP:

1. Generate feature frequencies

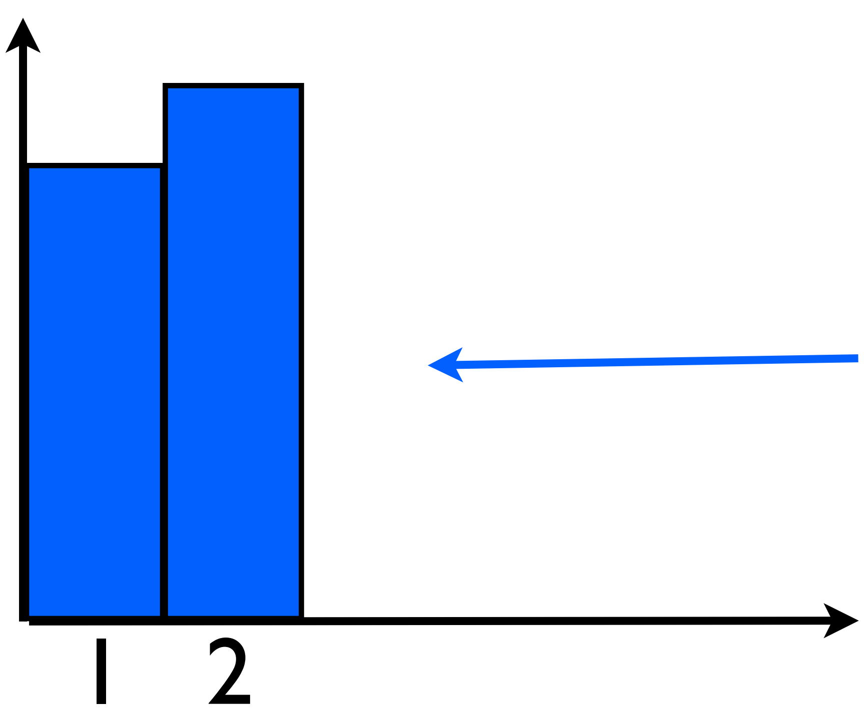
[Hjort 1990; Thibaux, Jordan 2007]

IBP as Polya urns



Another way to generate the IBP:

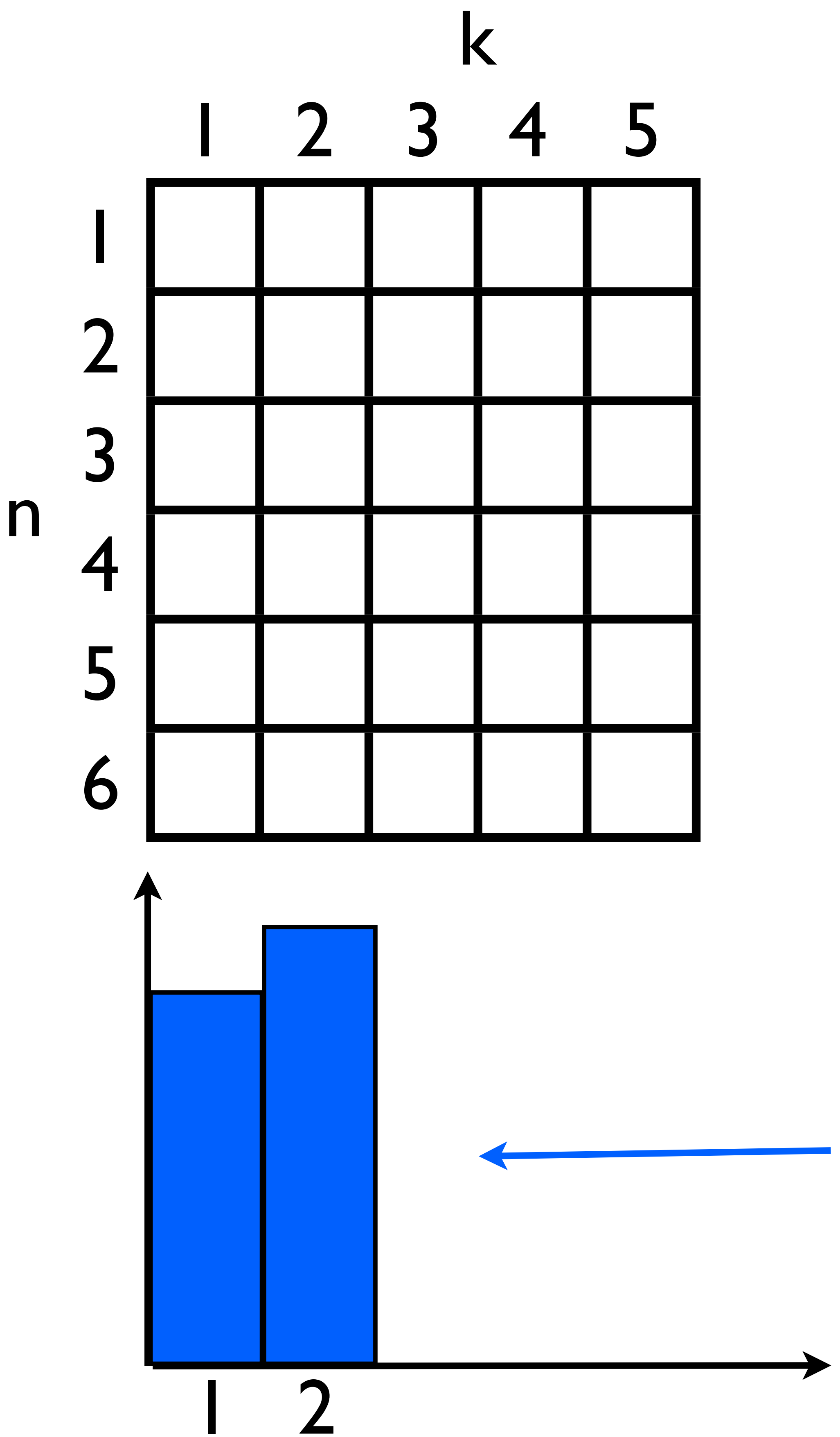
1. Generate feature frequencies



Poisson number of frequencies

[Hjort 1990; Thibaux, Jordan 2007]

IBP as Polya urns



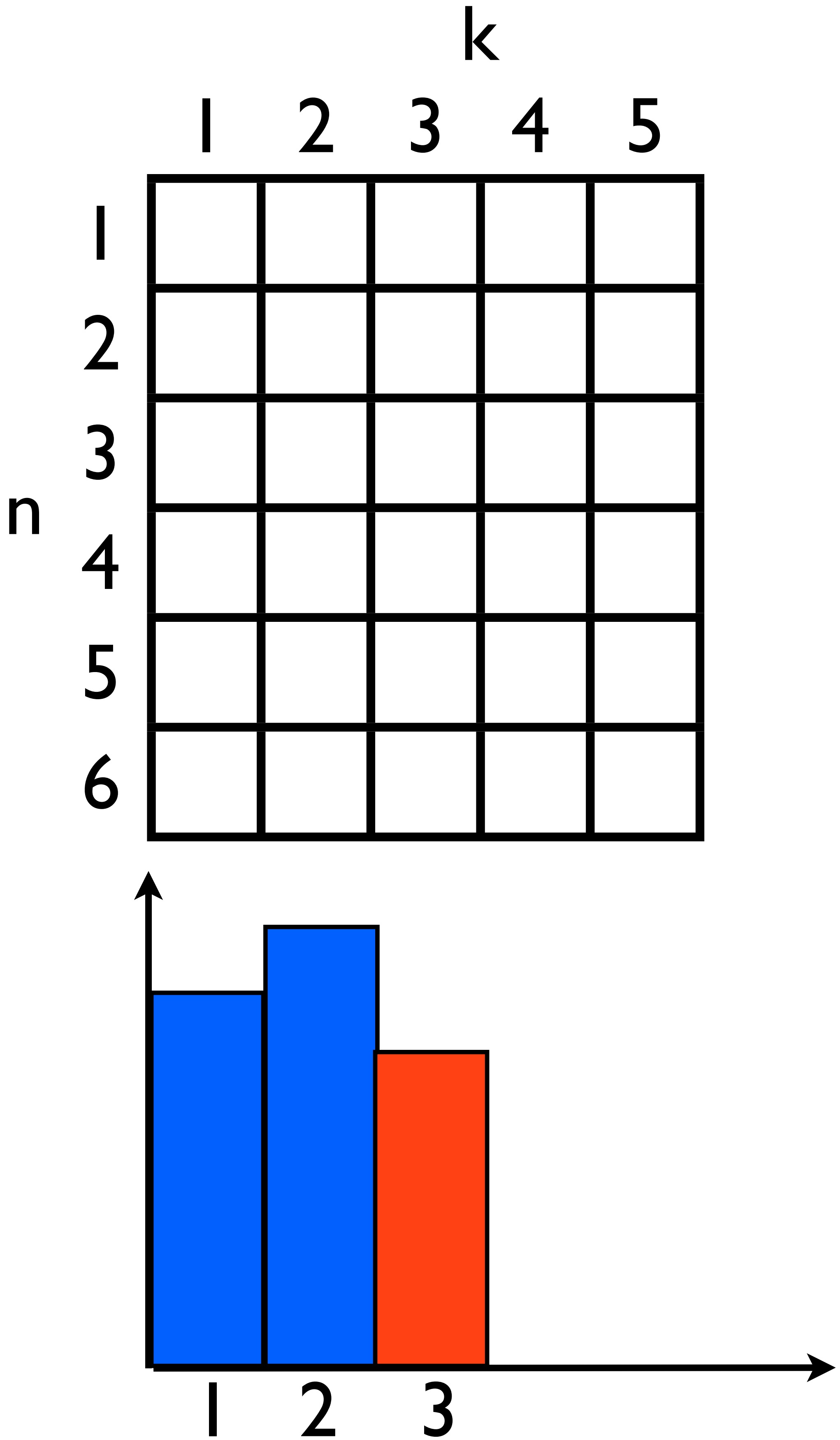
Another way to generate the IBP:

1. Generate feature frequencies

Poisson number of frequencies
Each has a beta distribution

[Hjort 1990; Thibaux, Jordan 2007]

IBP as Polya urns

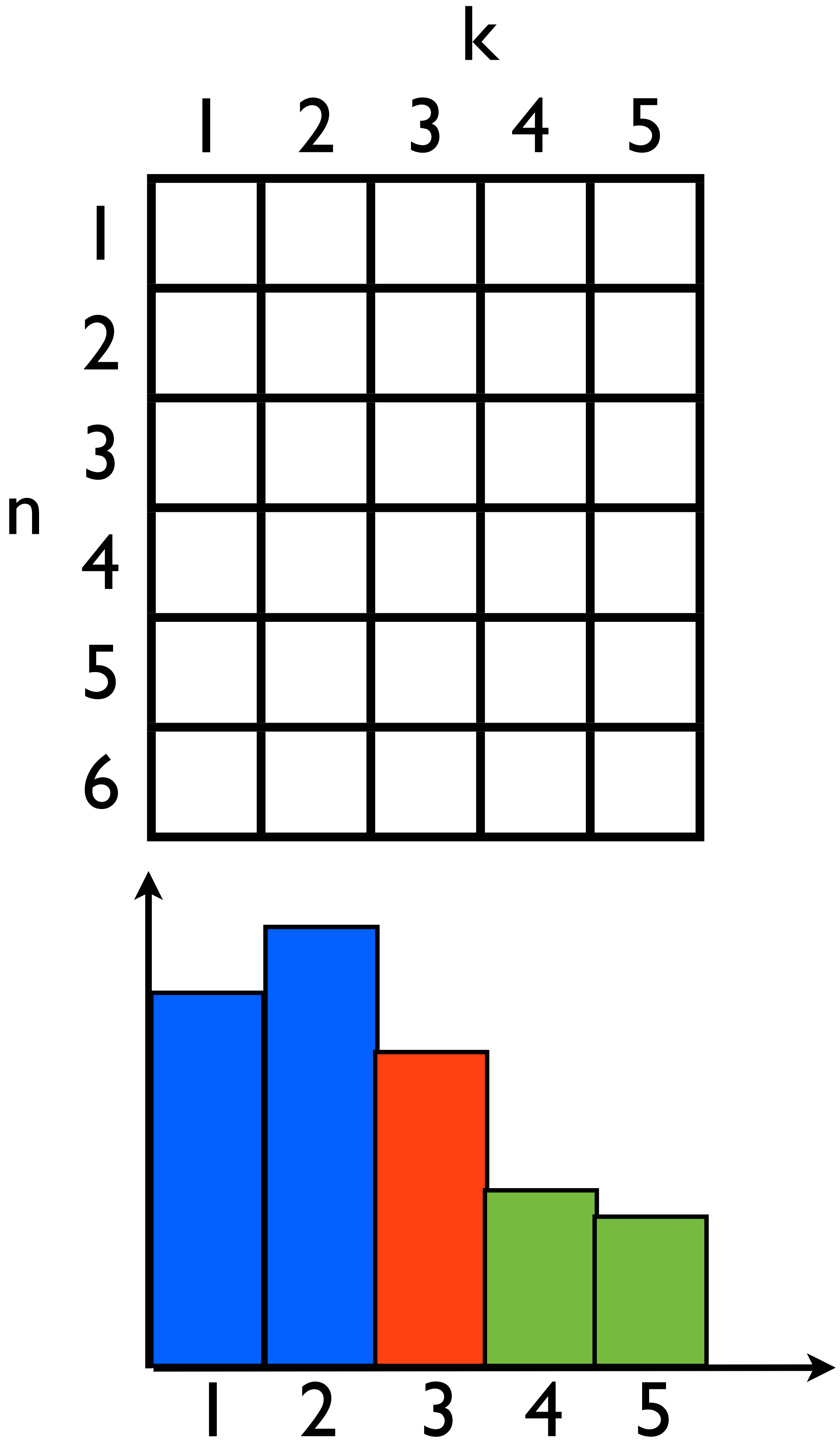


Another way to generate the IBP:

1. Generate feature frequencies

[Hjort 1990; Thibaux, Jordan 2007]

IBP as Polya urns

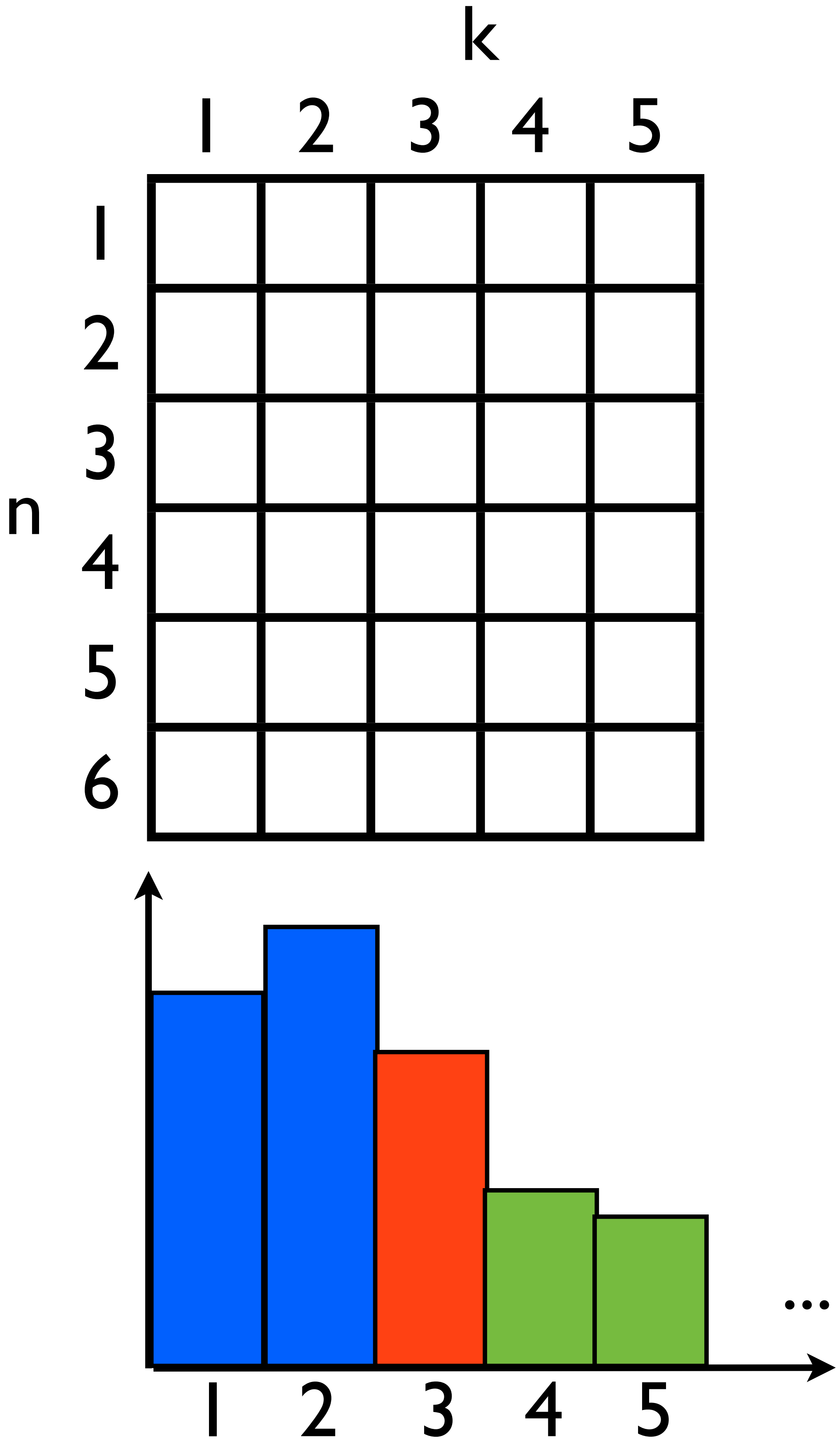


Another way to generate the IBP:

1. Generate feature frequencies

[Hjort 1990; Thibaux, Jordan 2007]

IBP as Polya urns

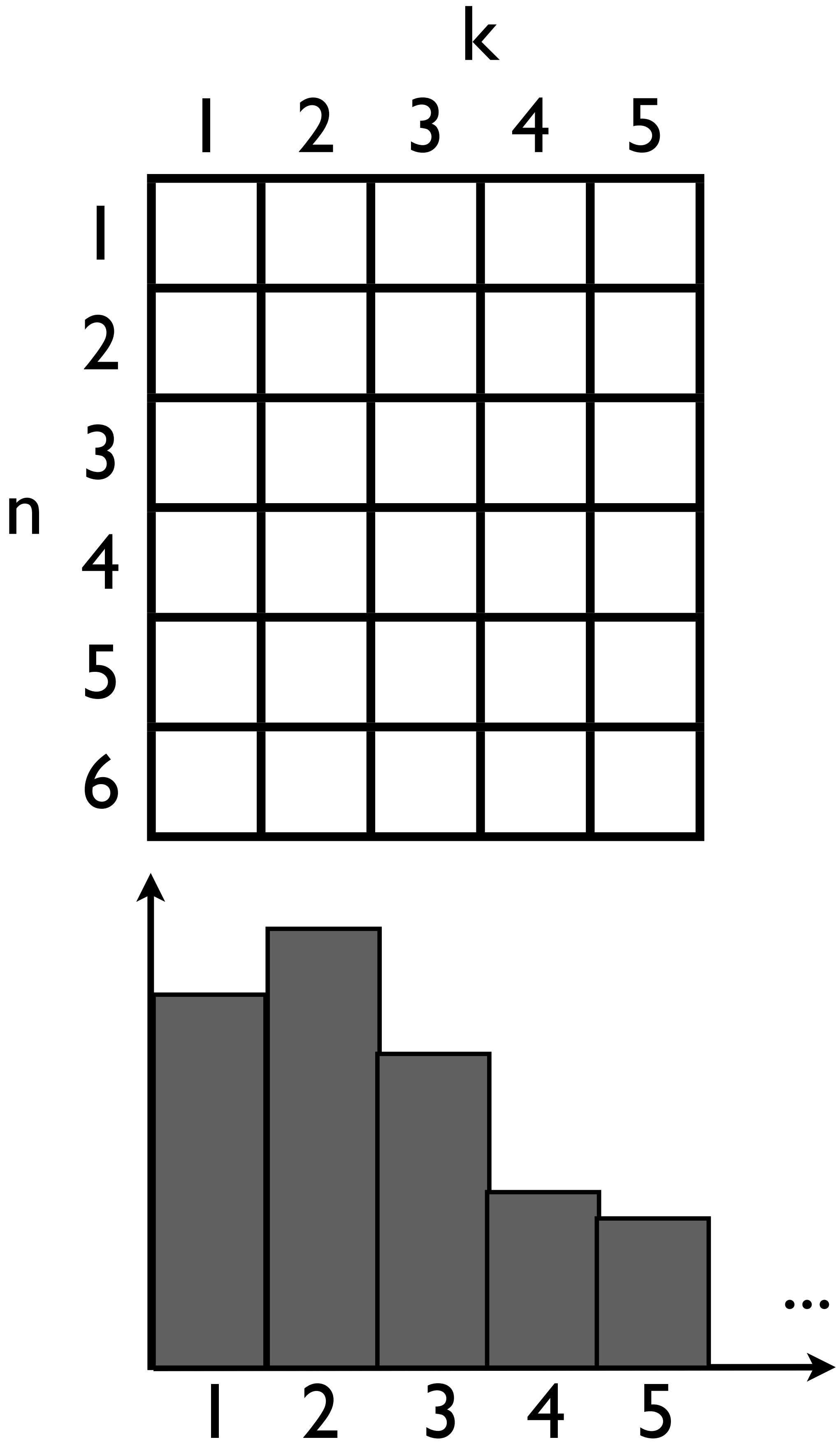


Another way to generate the IBP:

1. Generate feature frequencies

[Hjort 1990; Thibaux, Jordan 2007]

IBP as Polya urns

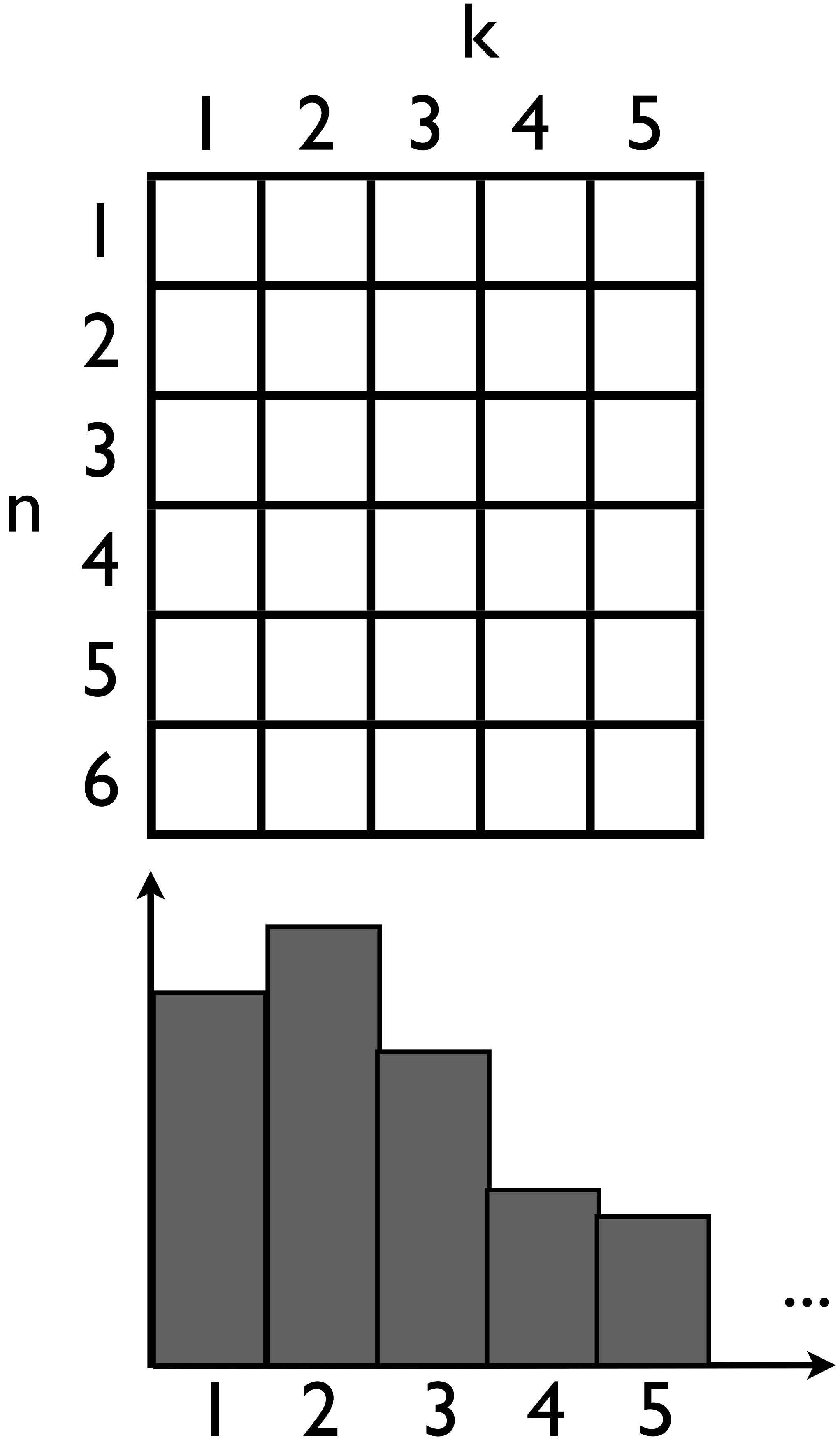


Another way to generate the IBP:

1. Generate feature frequencies

[Hjort 1990; Thibaux, Jordan 2007]

IBP as Polya urns

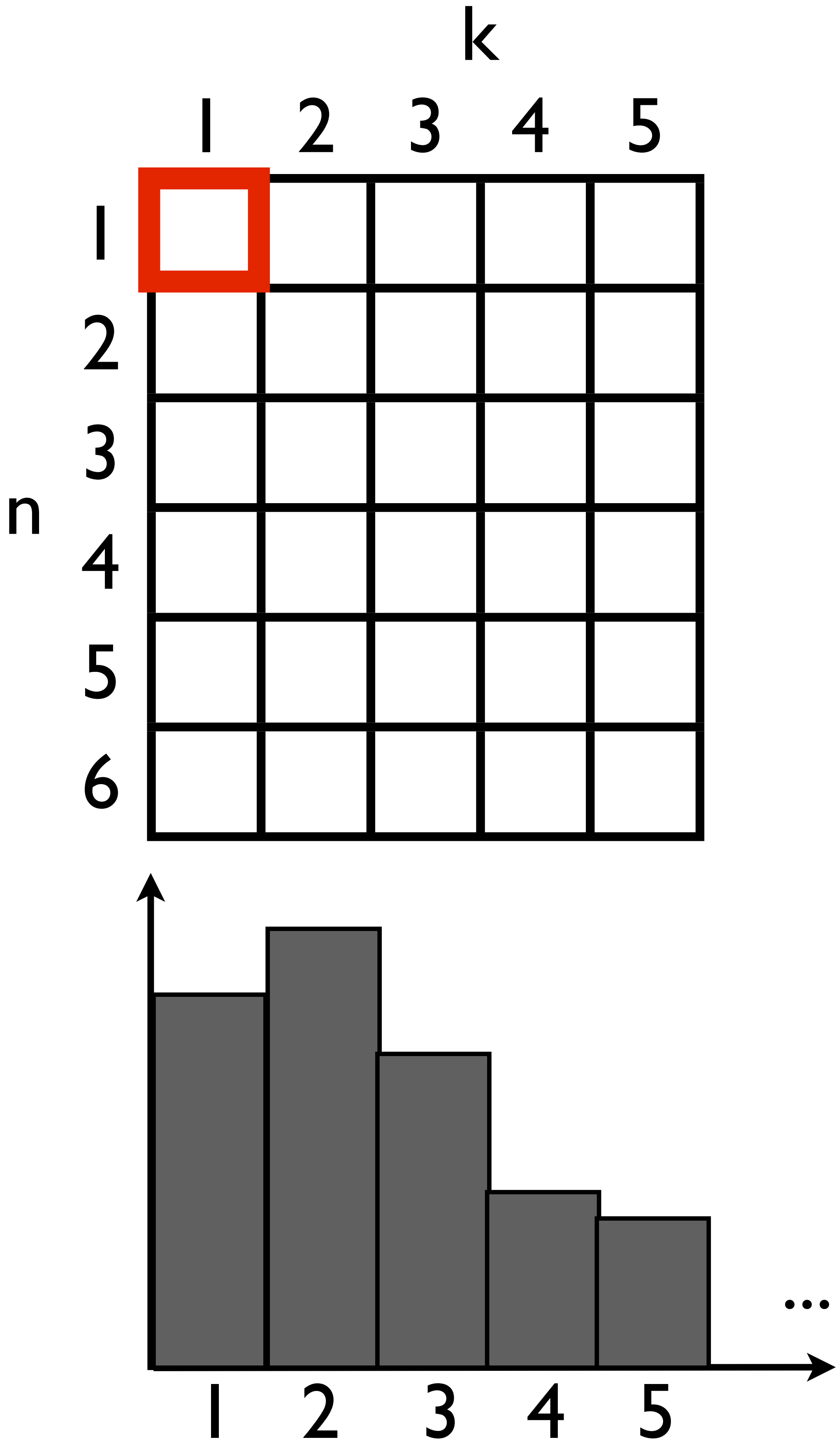


Another way to generate the IBP:

1. Generate feature frequencies
2. For each data point and for each feature, make independent Bernoulli draws

[Hjort 1990; Thibaux, Jordan 2007]

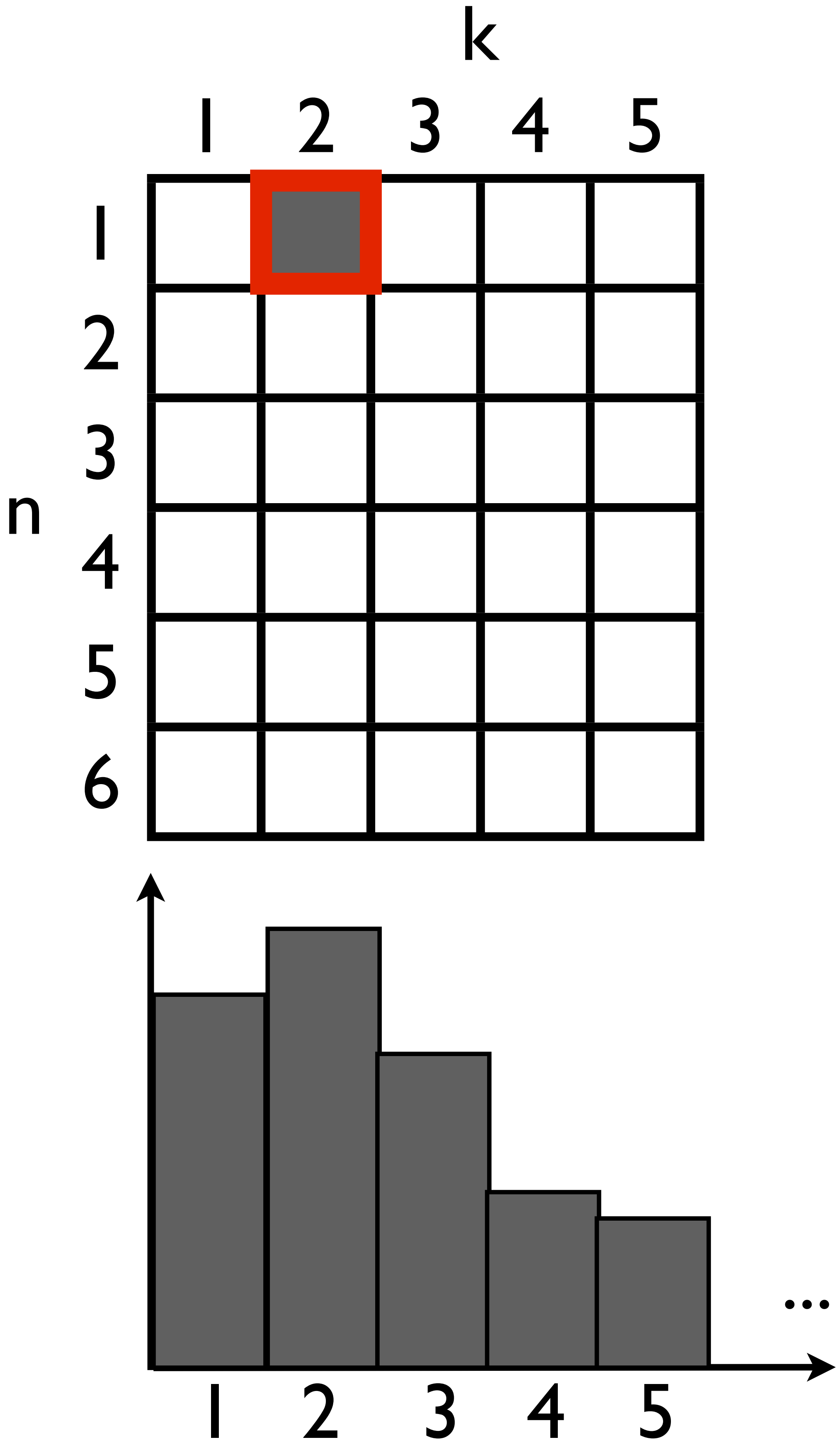
IBP as Polya urns



Another way to generate the IBP:

1. Generate feature frequencies
2. For each data point and for each feature, make independent Bernoulli draws

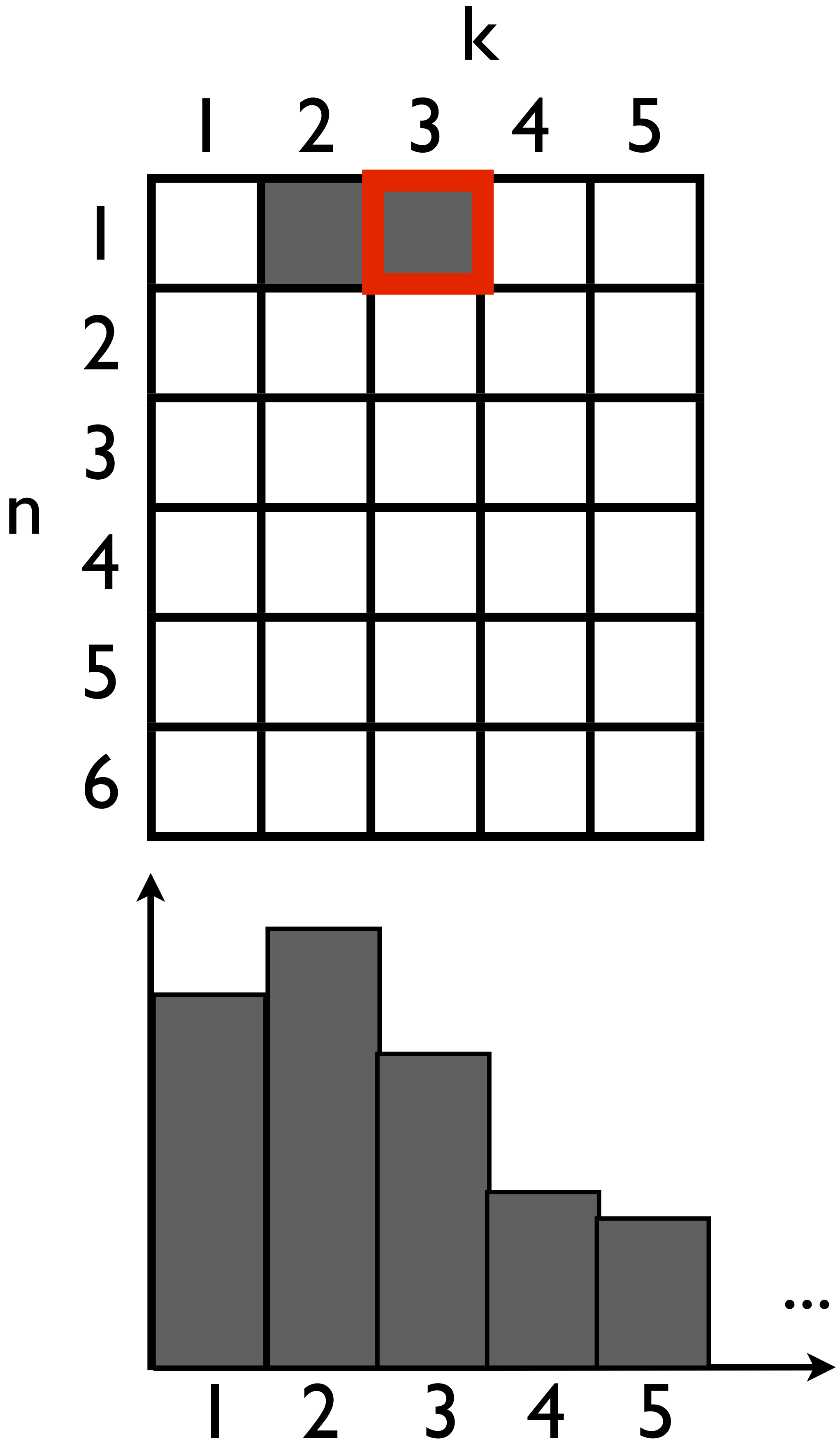
IBP as Polya urns



Another way to generate the IBP:

1. Generate feature frequencies
2. For each data point and for each feature, make independent Bernoulli draws

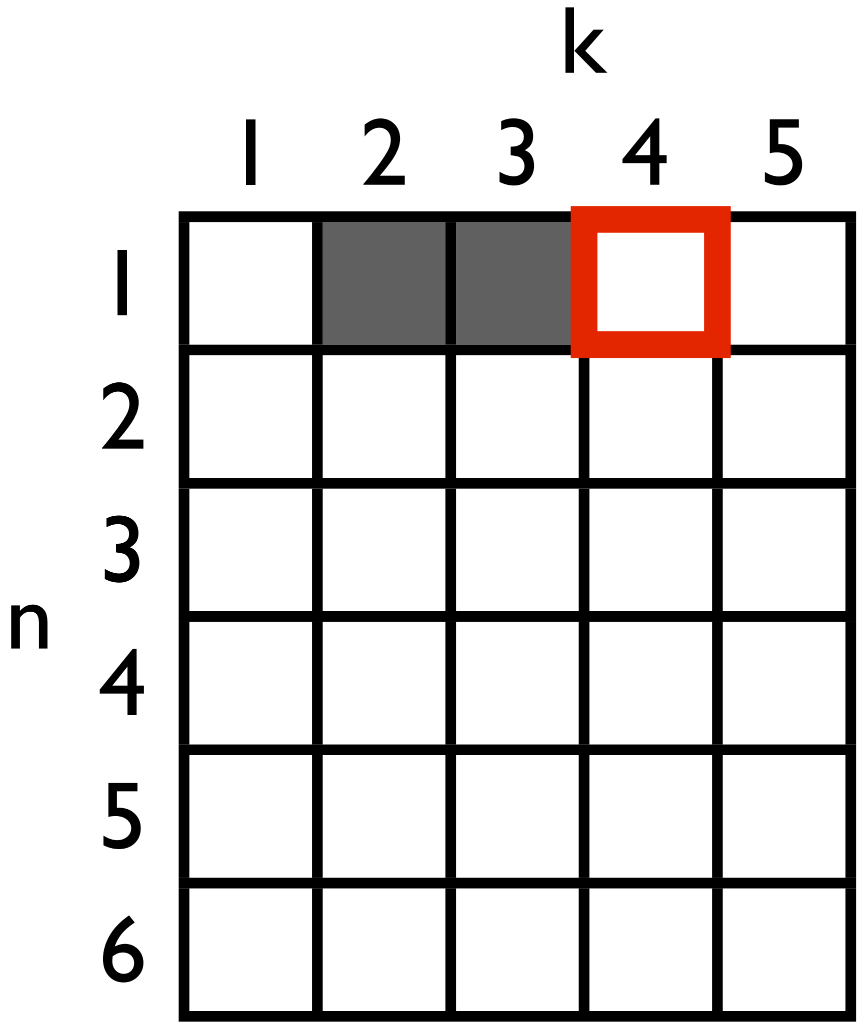
IBP as Polya urns

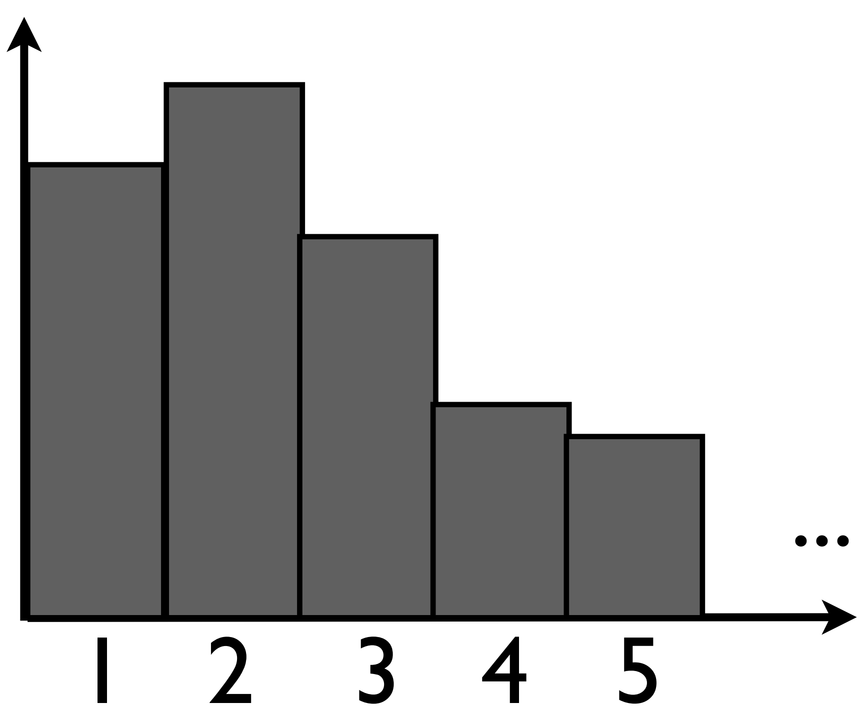


Another way to generate the IBP:

1. Generate feature frequencies
2. For each data point and for each feature, make independent Bernoulli draws

IBP as Polya urns

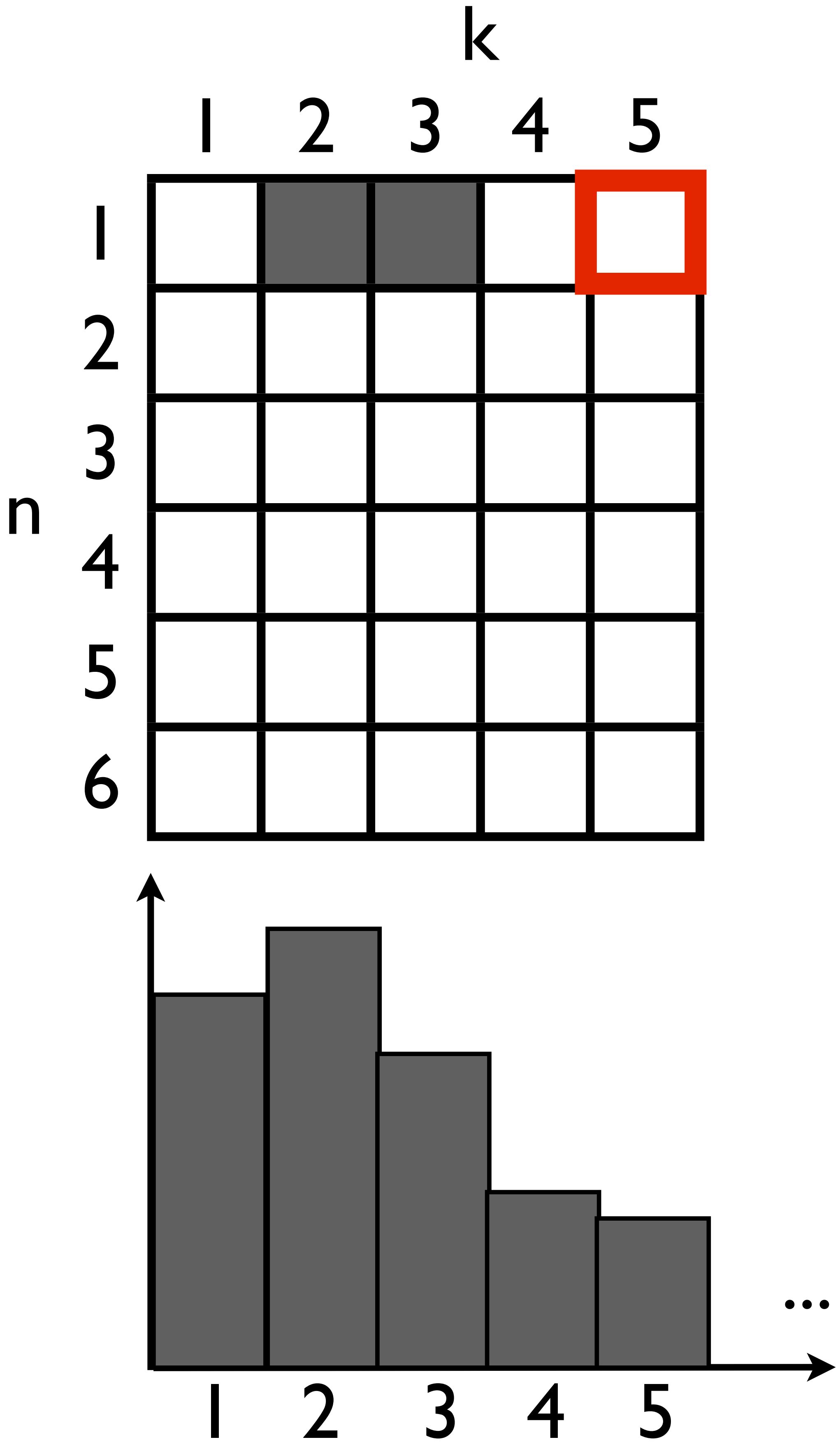




Another way to generate the IBP:

1. Generate feature frequencies
2. For each data point and for each feature, make independent Bernoulli draws

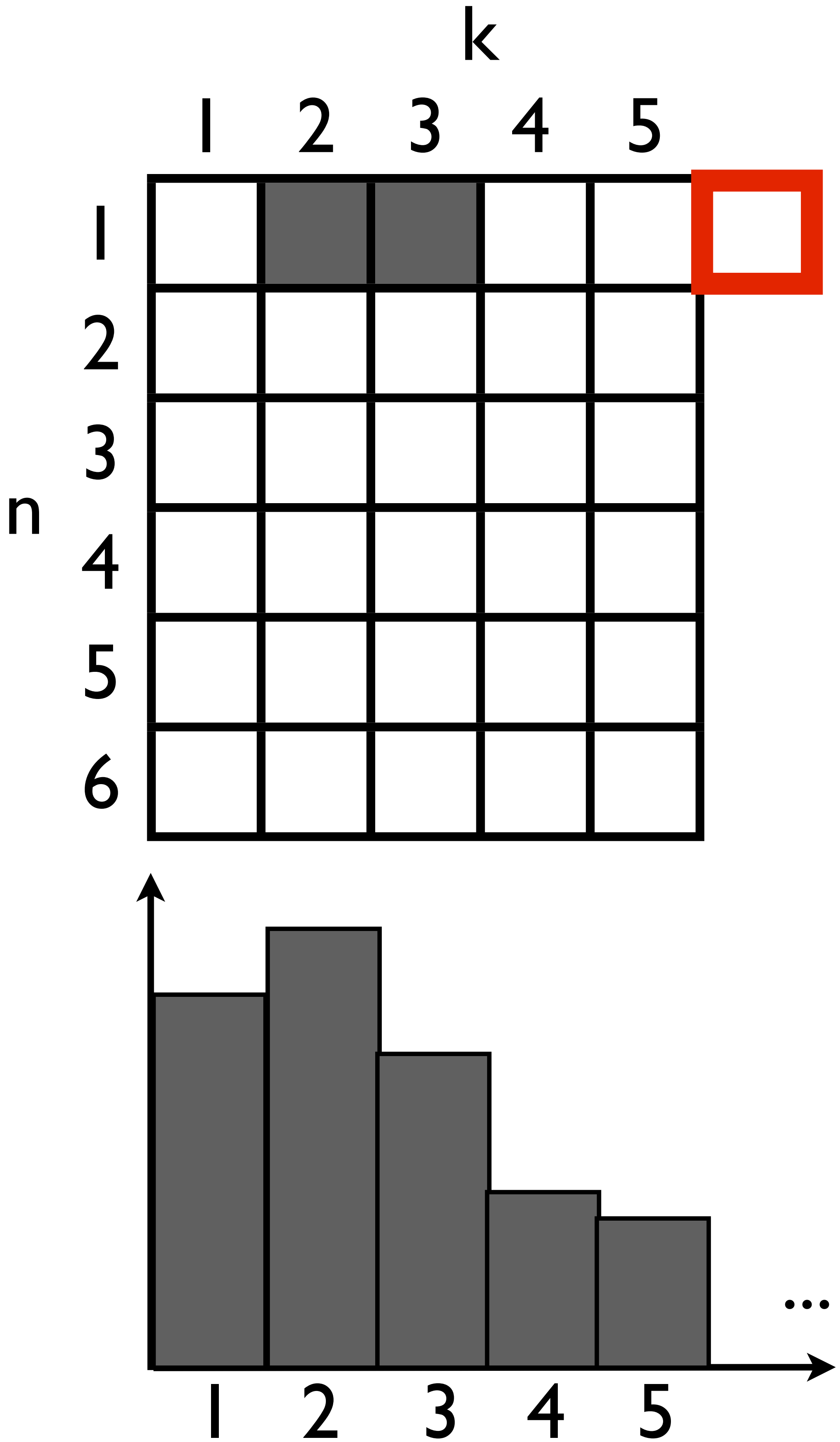
IBP as Polya urns



Another way to generate the IBP:

1. Generate feature frequencies
2. For each data point and for each feature, make independent Bernoulli draws

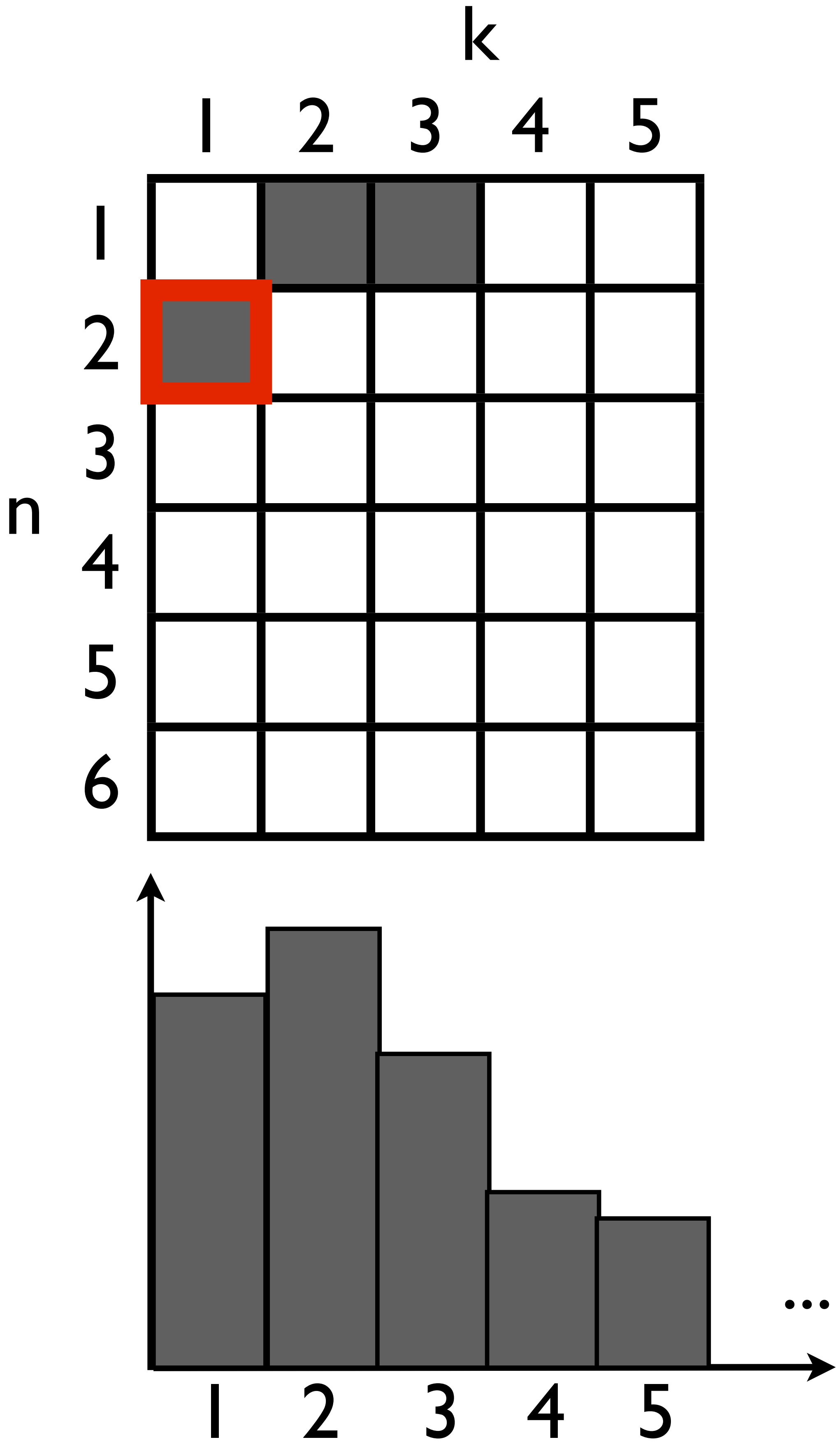
IBP as Polya urns



Another way to generate the IBP:

1. Generate feature frequencies
2. For each data point and for each feature, make independent Bernoulli draws

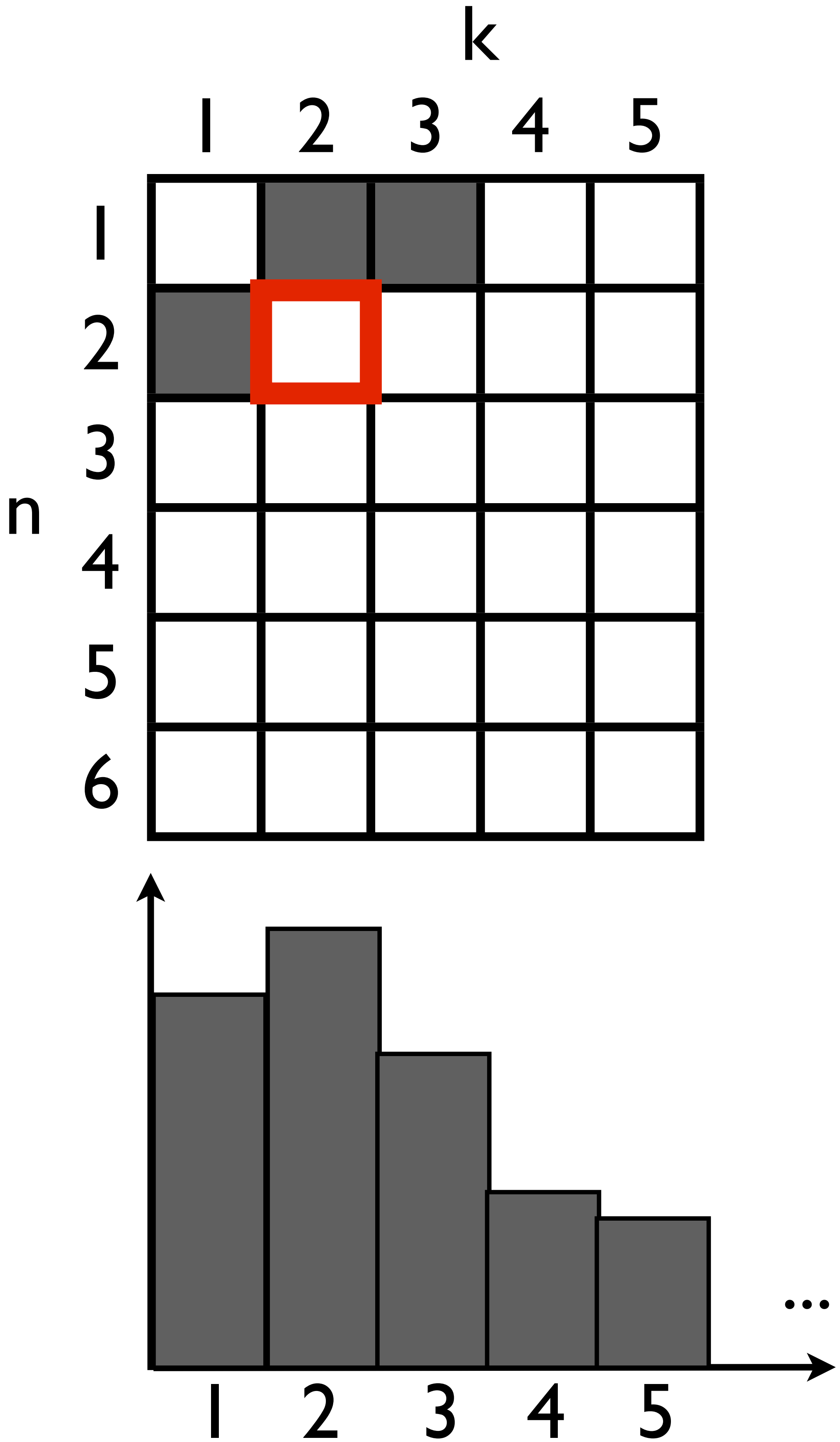
IBP as Polya urns



Another way to generate the IBP:

1. Generate feature frequencies
2. For each data point and for each feature, make independent Bernoulli draws

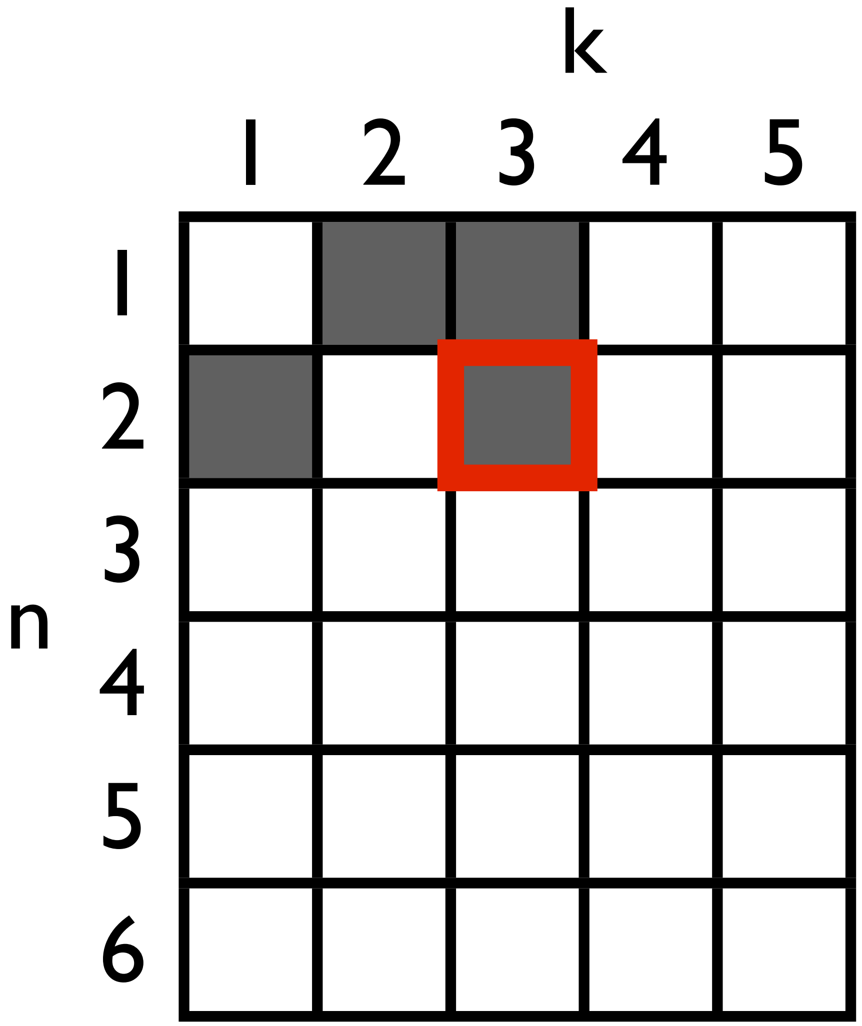
IBP as Polya urns

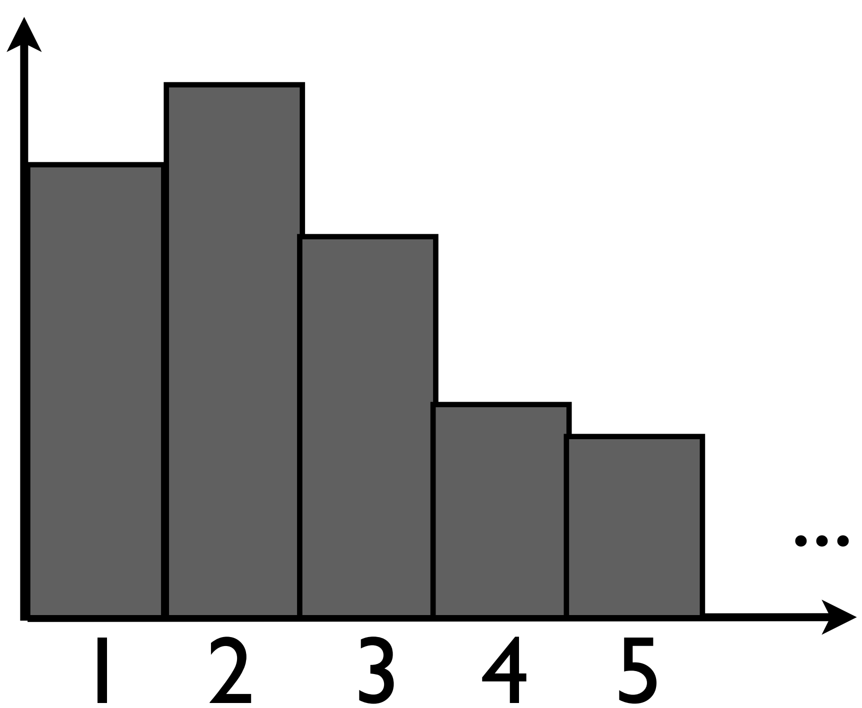


Another way to generate the IBP:

1. Generate feature frequencies
2. For each data point and for each feature, make independent Bernoulli draws

IBP as Polya urns

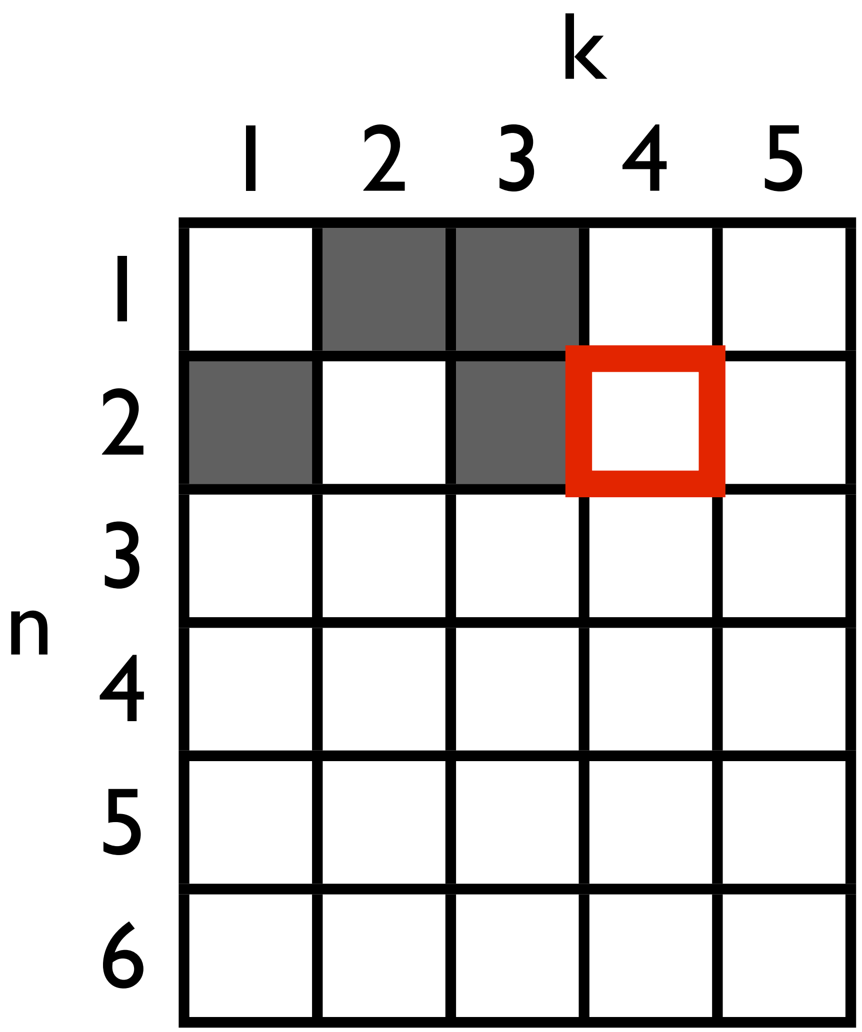


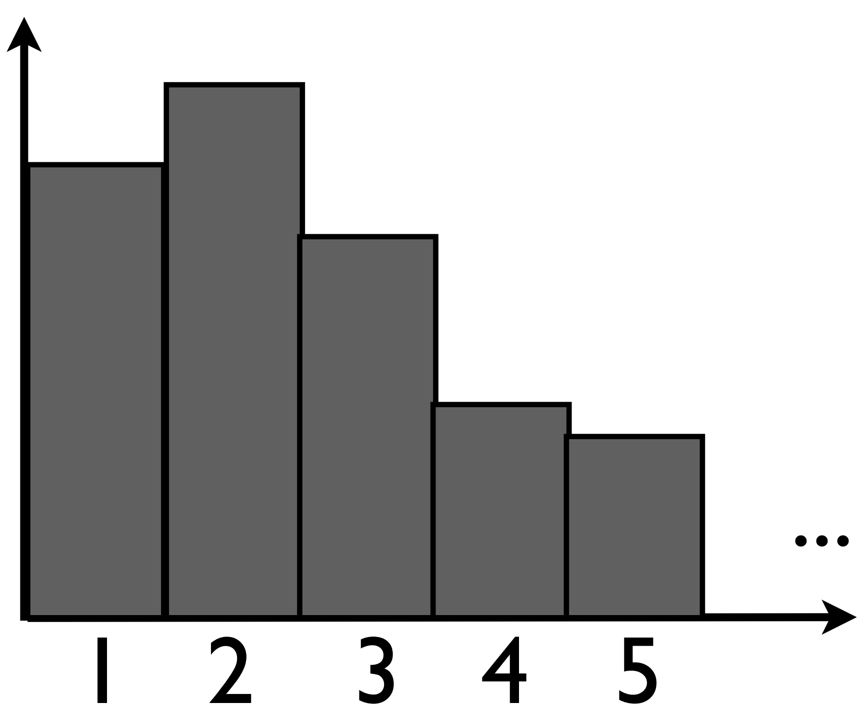


Another way to generate the IBP:

1. Generate feature frequencies
2. For each data point and for each feature, make independent Bernoulli draws

IBP as Polya urns

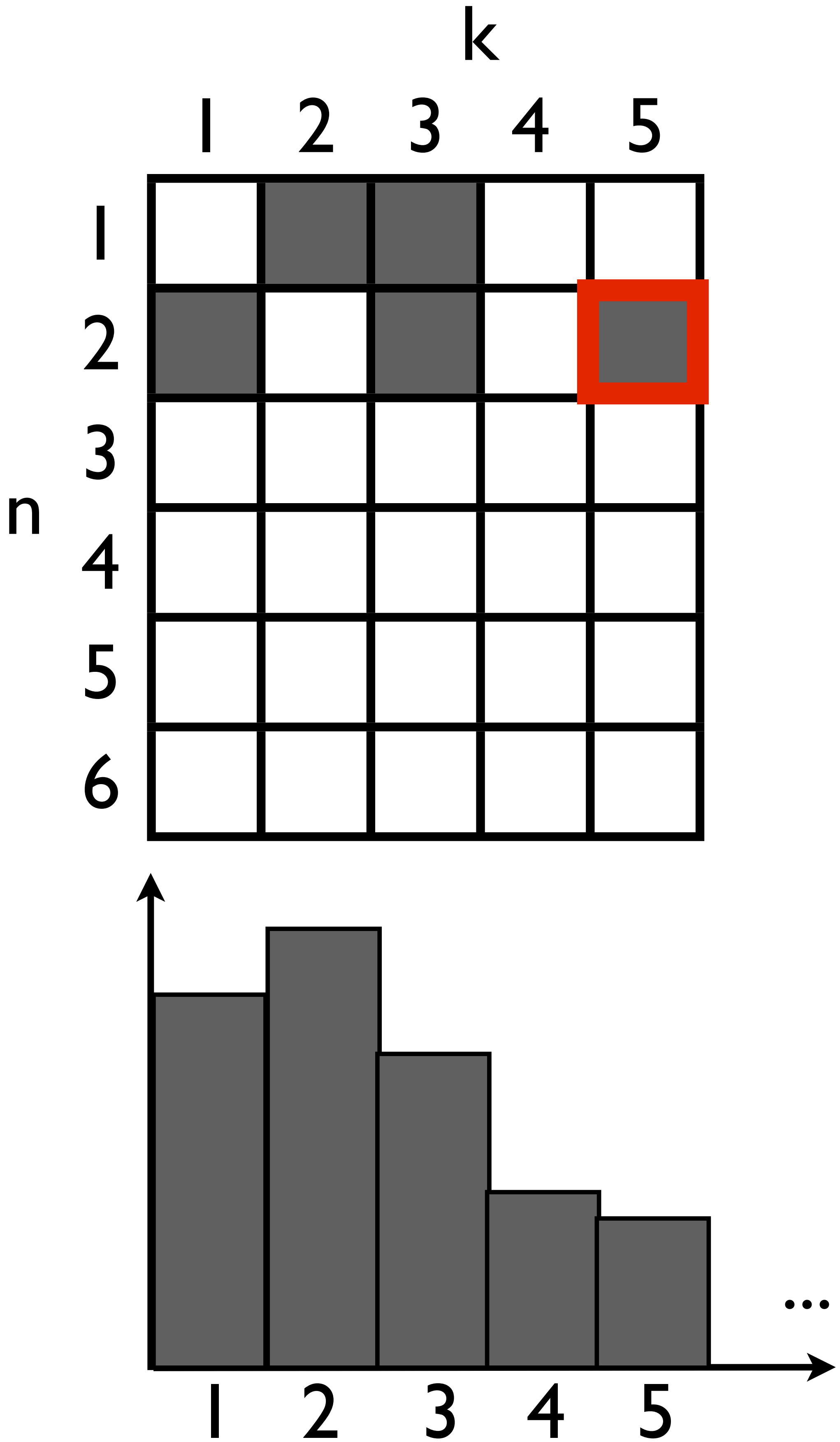




Another way to generate the IBP:

1. Generate feature frequencies
2. For each data point and for each feature, make independent Bernoulli draws

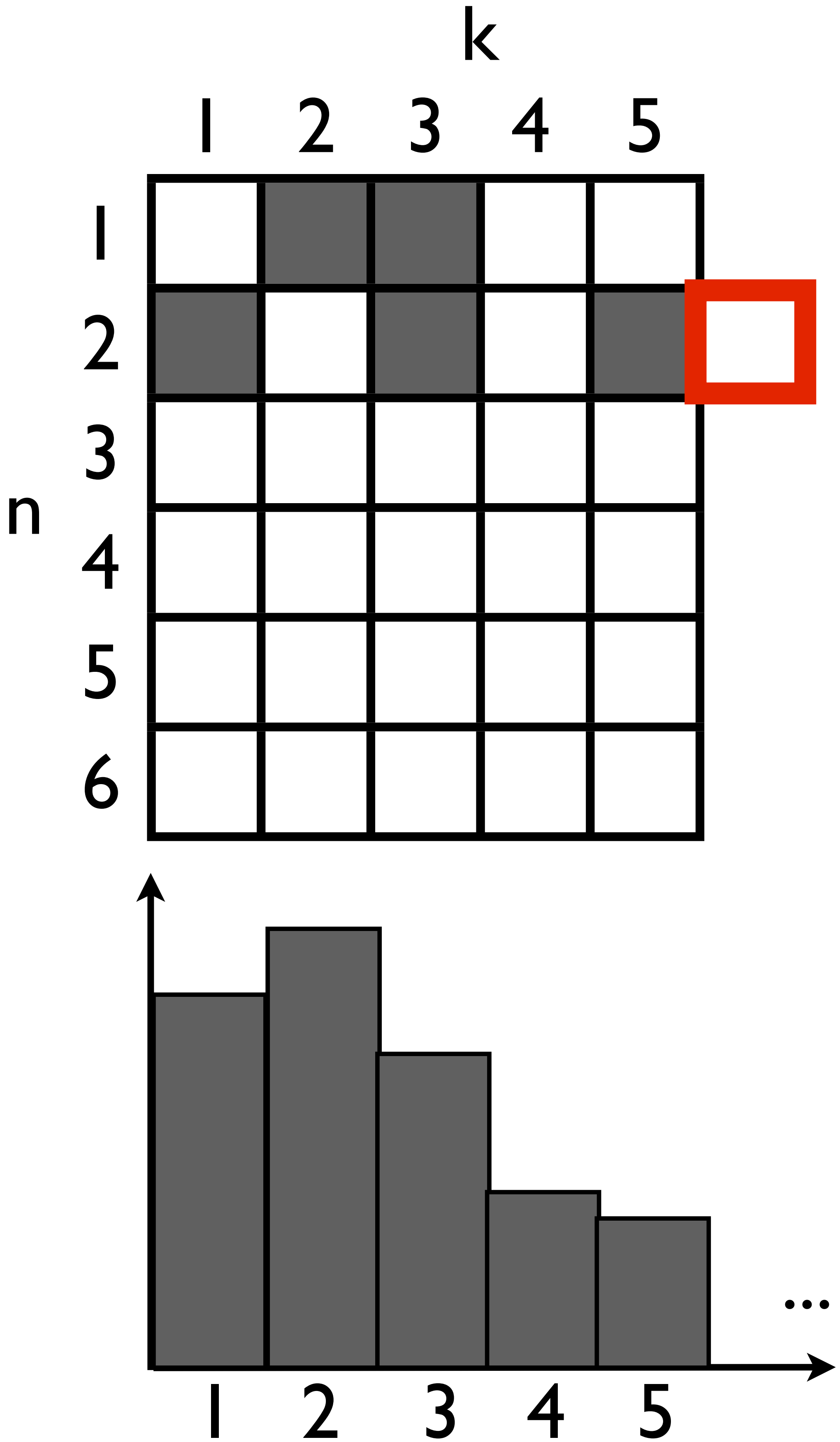
IBP as Polya urns



Another way to generate the IBP:

1. Generate feature frequencies
2. For each data point and for each feature, make independent Bernoulli draws

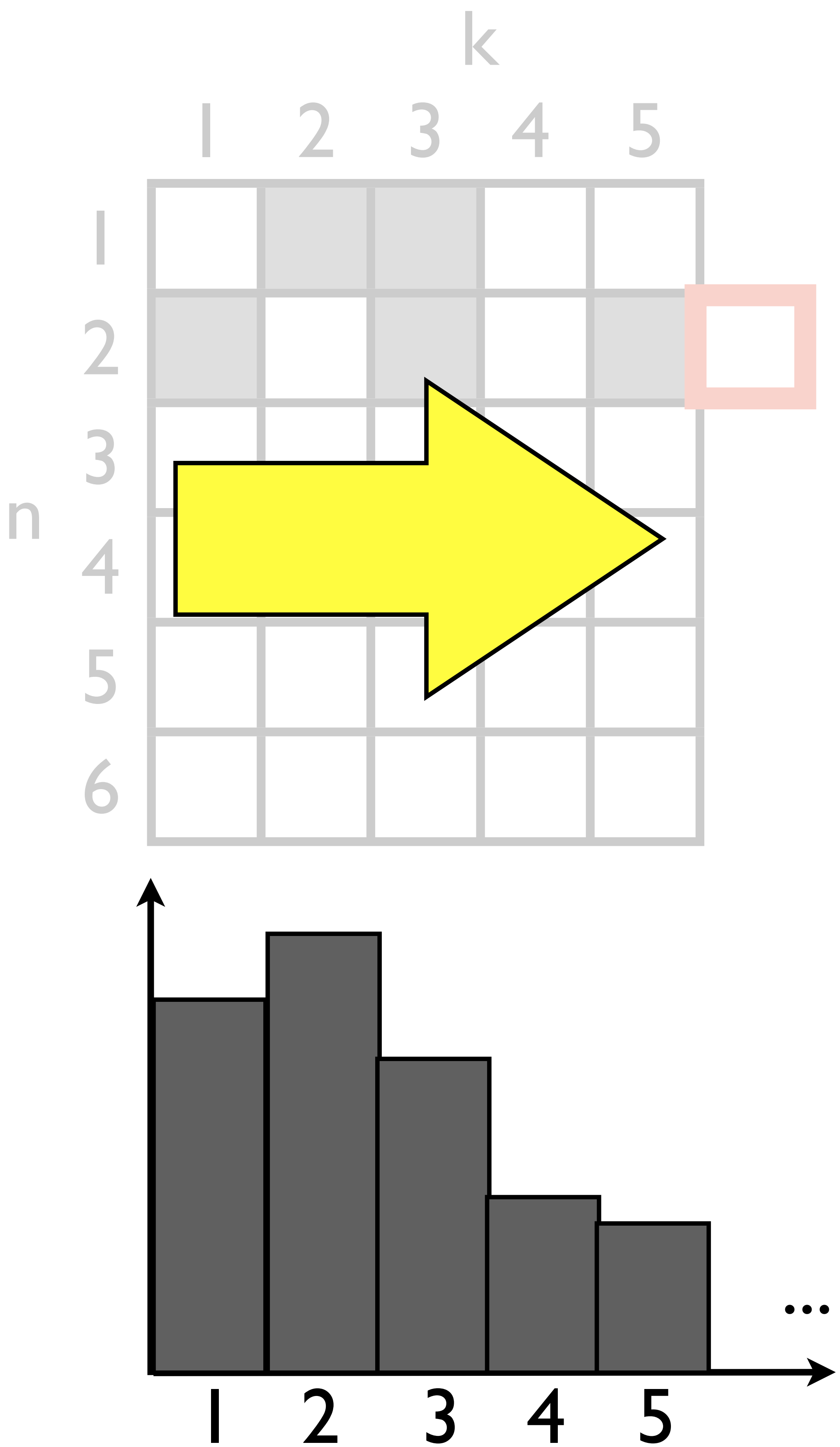
IBP as Polya urns



Another way to generate the IBP:

1. Generate feature frequencies
2. For each data point and for each feature, make independent Bernoulli draws

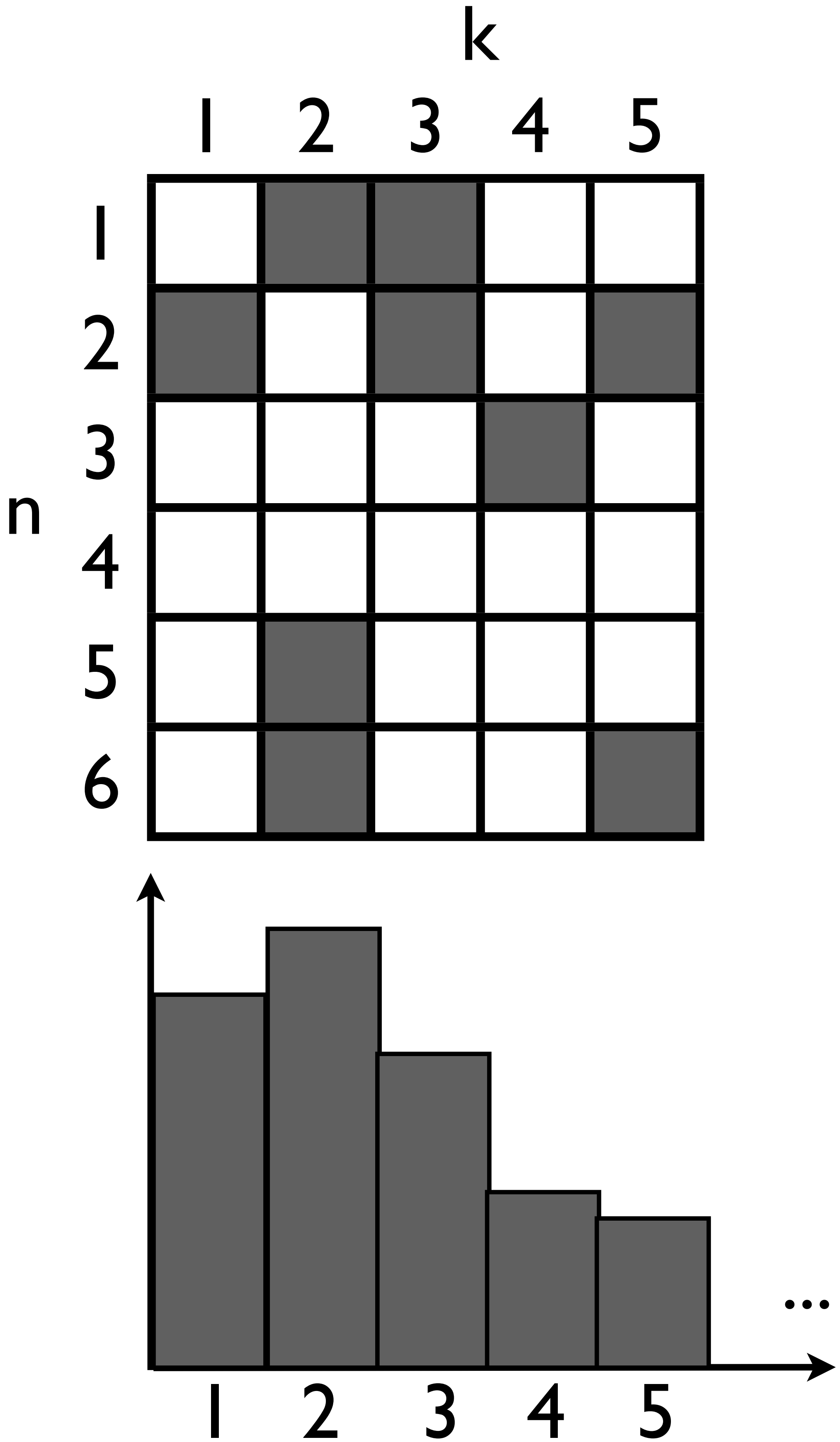
IBP as Polya urns



Another way to generate the IBP:

1. Generate feature frequencies
2. For each data point and for each feature, make independent Bernoulli draws

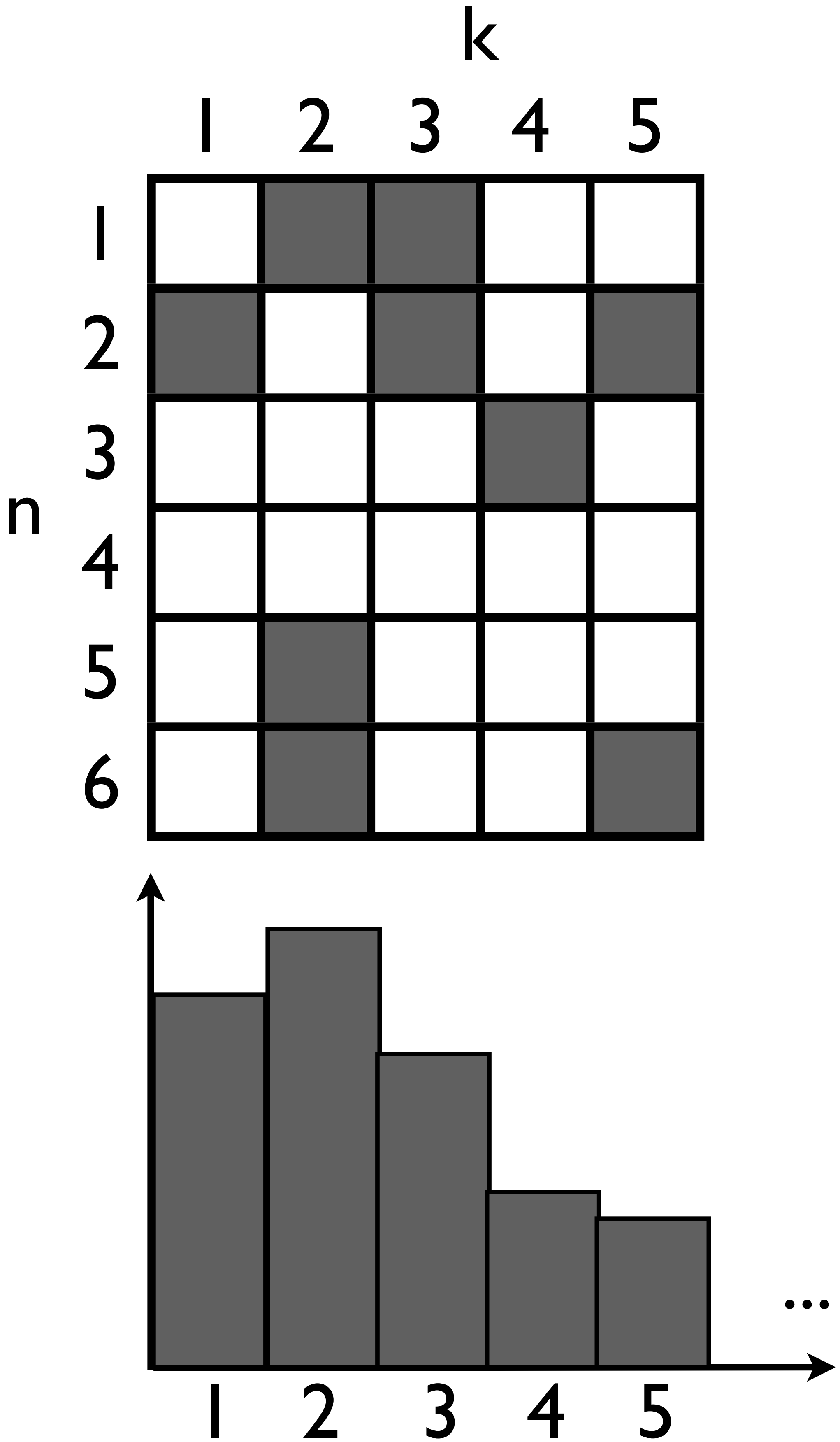
IBP as Polya urns



Another way to generate the IBP:

1. Generate feature frequencies
2. For each data point and for each feature, make independent Bernoulli draws

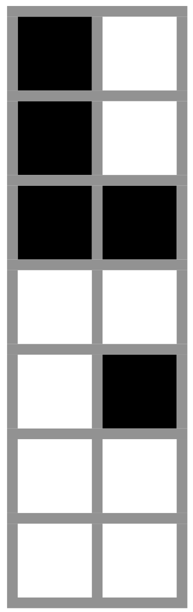
Frequency model



1. Generate feature frequencies
2. For each data point and for each feature, make independent Bernoulli draws

Not a frequency model

Counterexample



$$\mathbb{P}(\text{row} = \blacksquare\square) = p_1$$

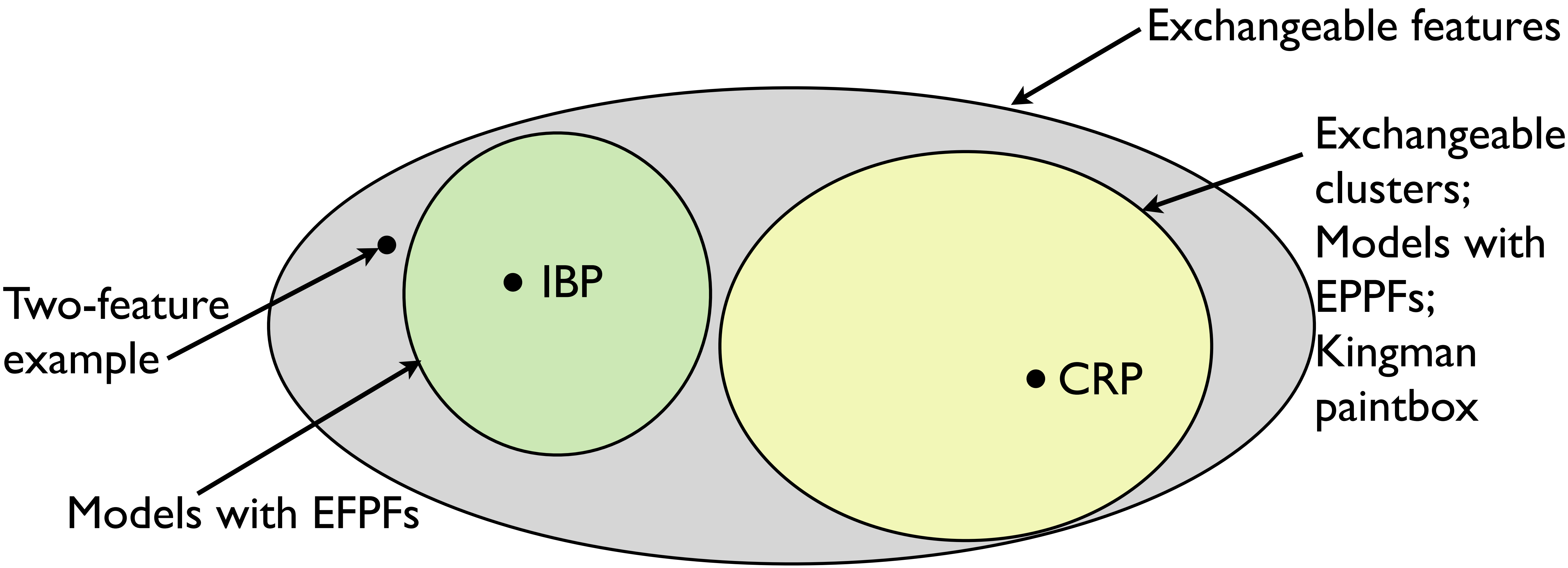
$$\mathbb{P}(\text{row} = \square\blacksquare) = p_2$$

$$\mathbb{P}(\text{row} = \blacksquare\blacksquare) = p_3$$

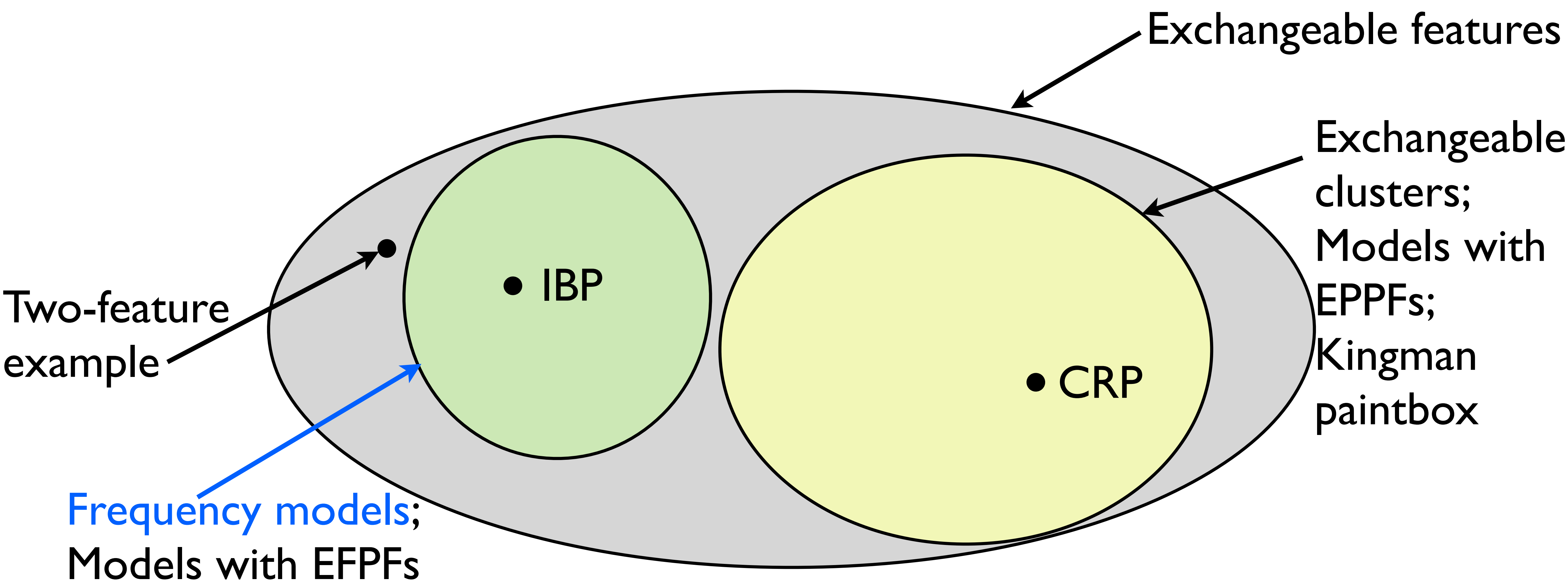
$$\mathbb{P}(\text{row} = \square\square) = p_4$$

$$p_1 \neq (p_1 + p_3)(p_2 + p_3)$$

Clusters and features



Clusters and features



[Broderick, Pitman, Jordan (submitted)]

Outline

I. Clusters

II. Features

- Overview
- Distribution
- Frequencies
- Random measure
- Paintbox

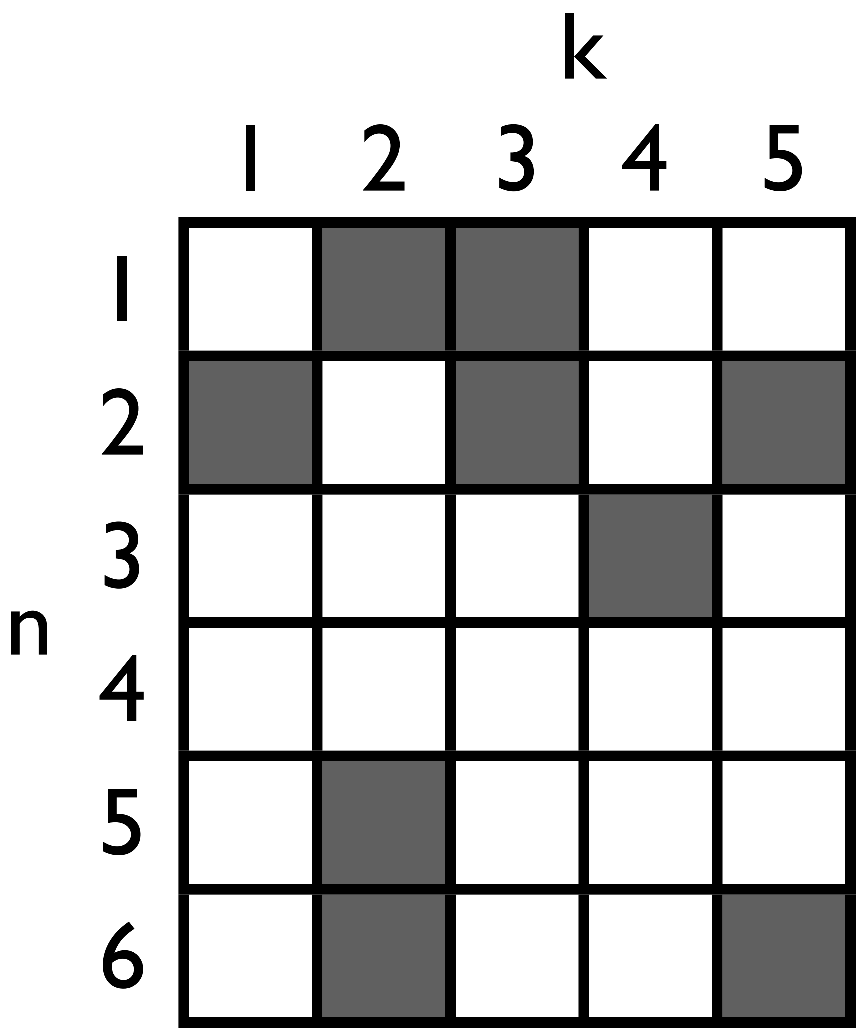
Outline

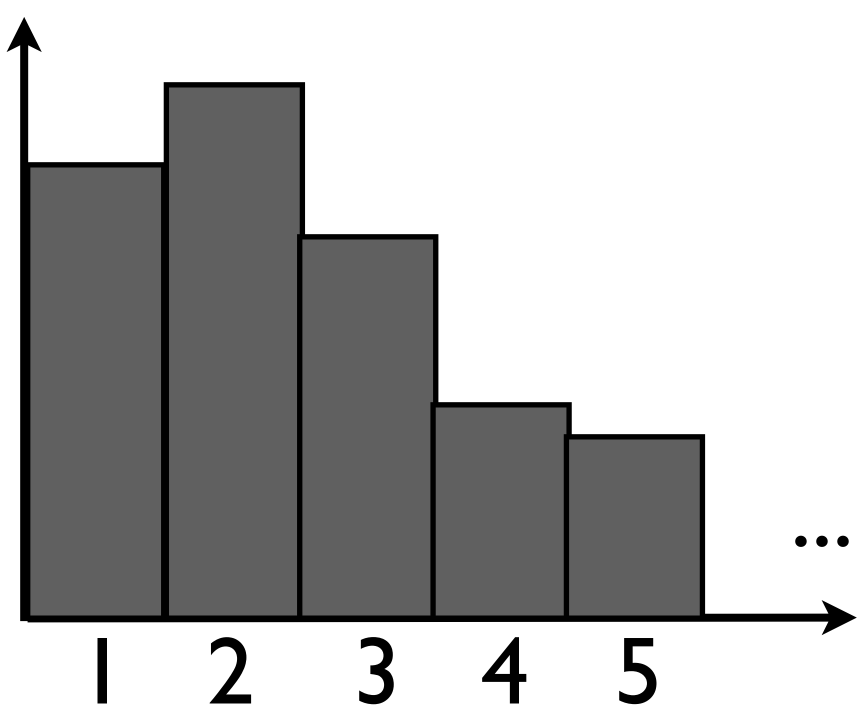
I. Clusters

II. Features

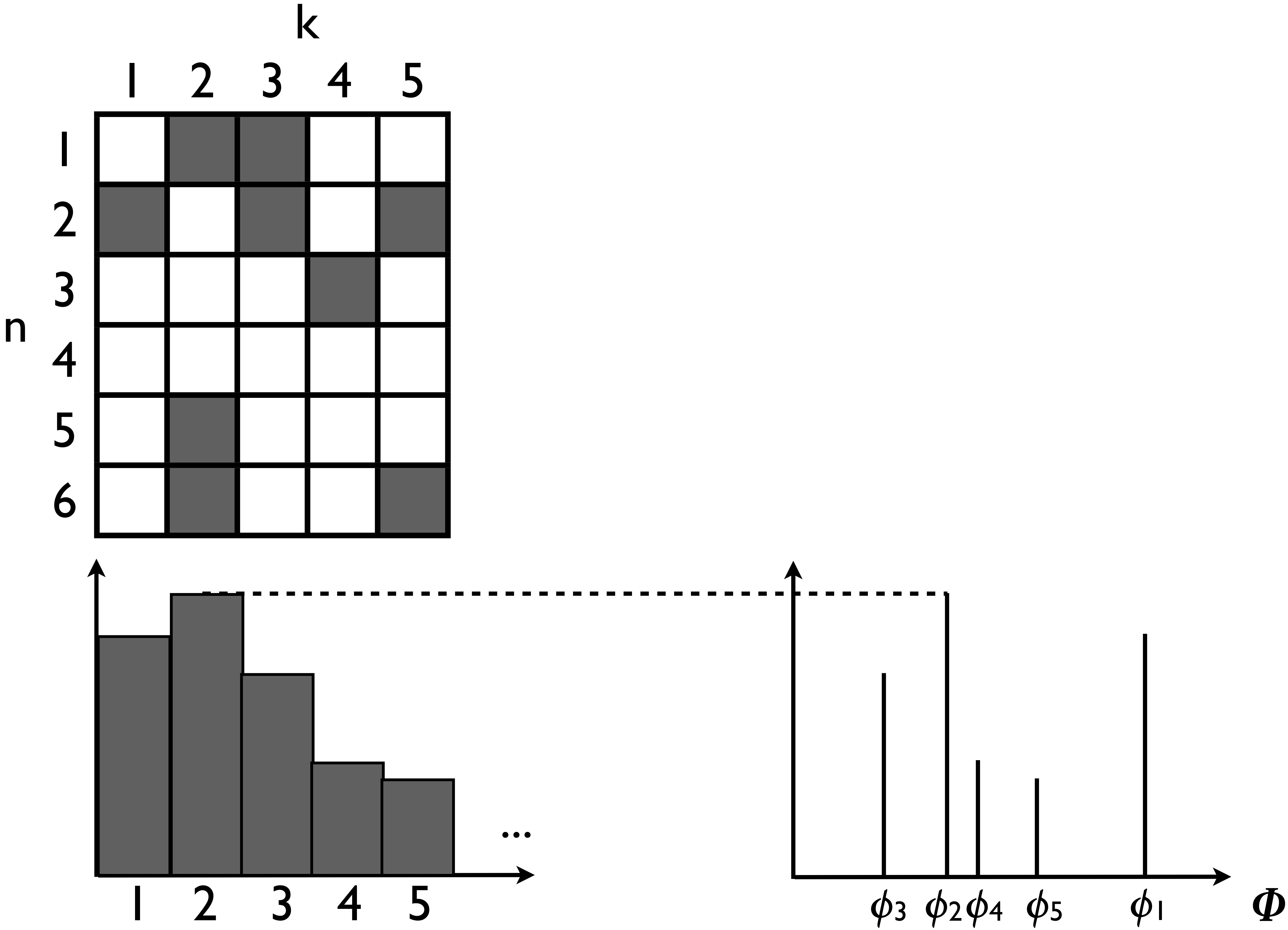
- Overview
- Distribution
- Frequencies
- Random measure
- Paintbox

Frequency model

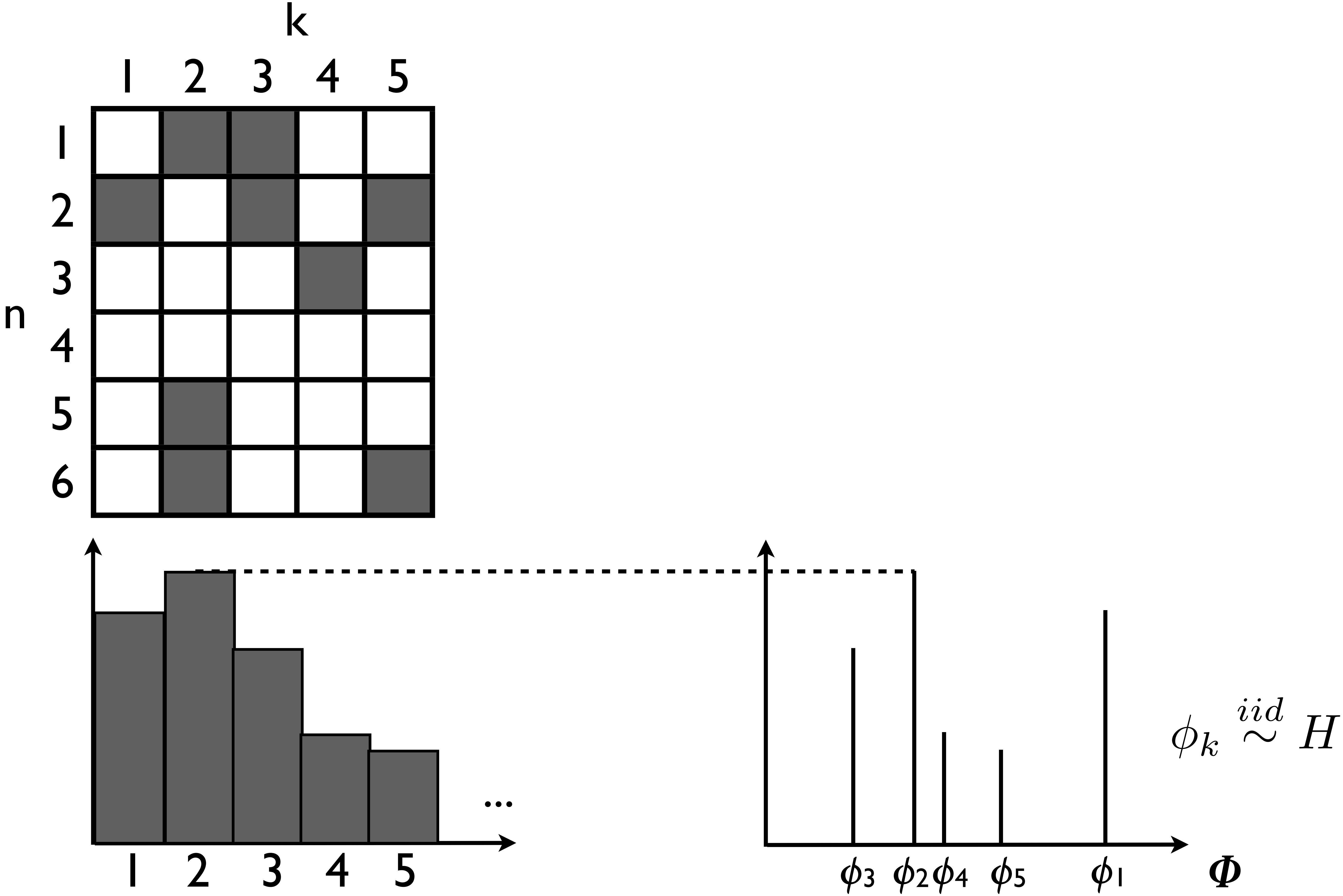




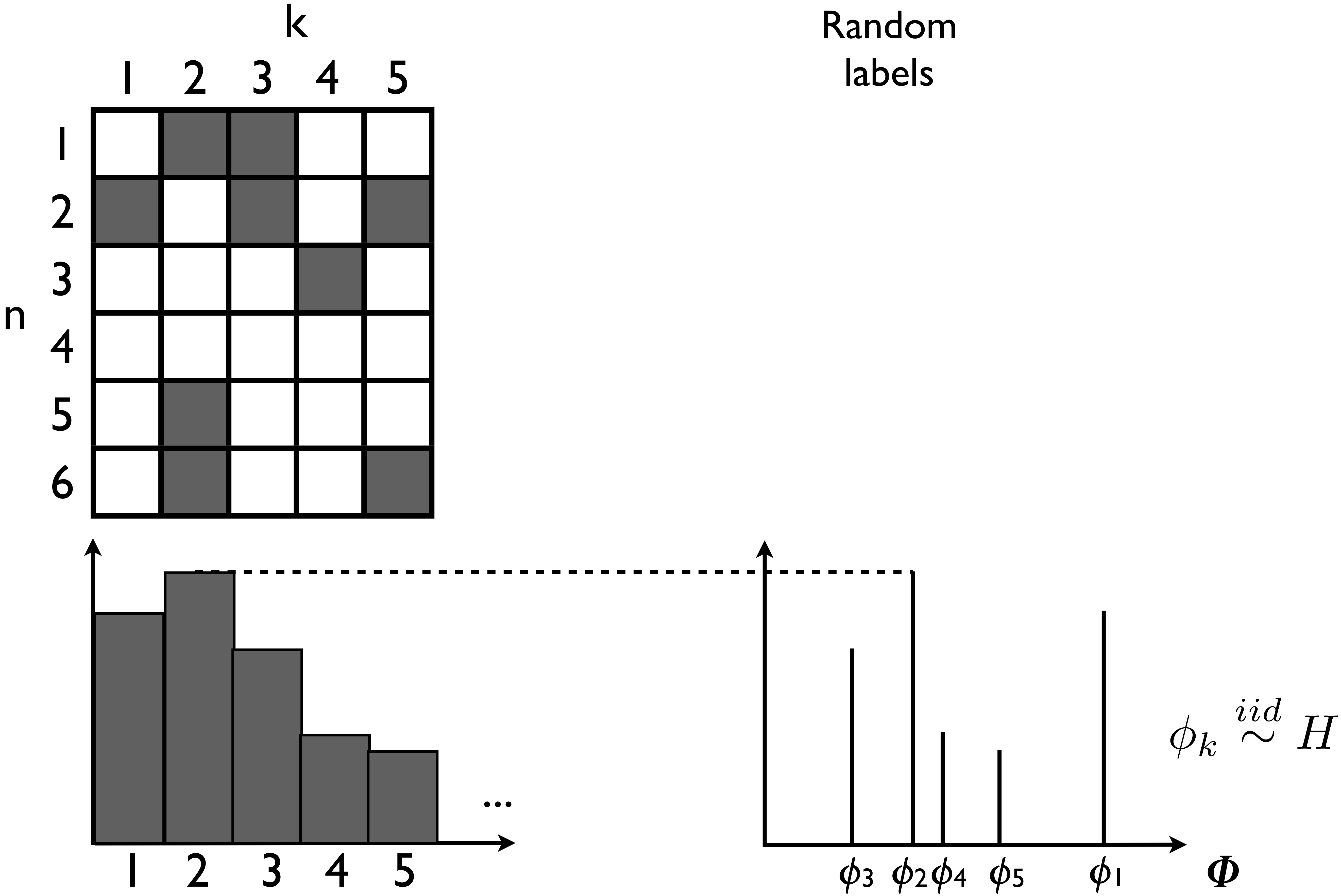
Frequency model



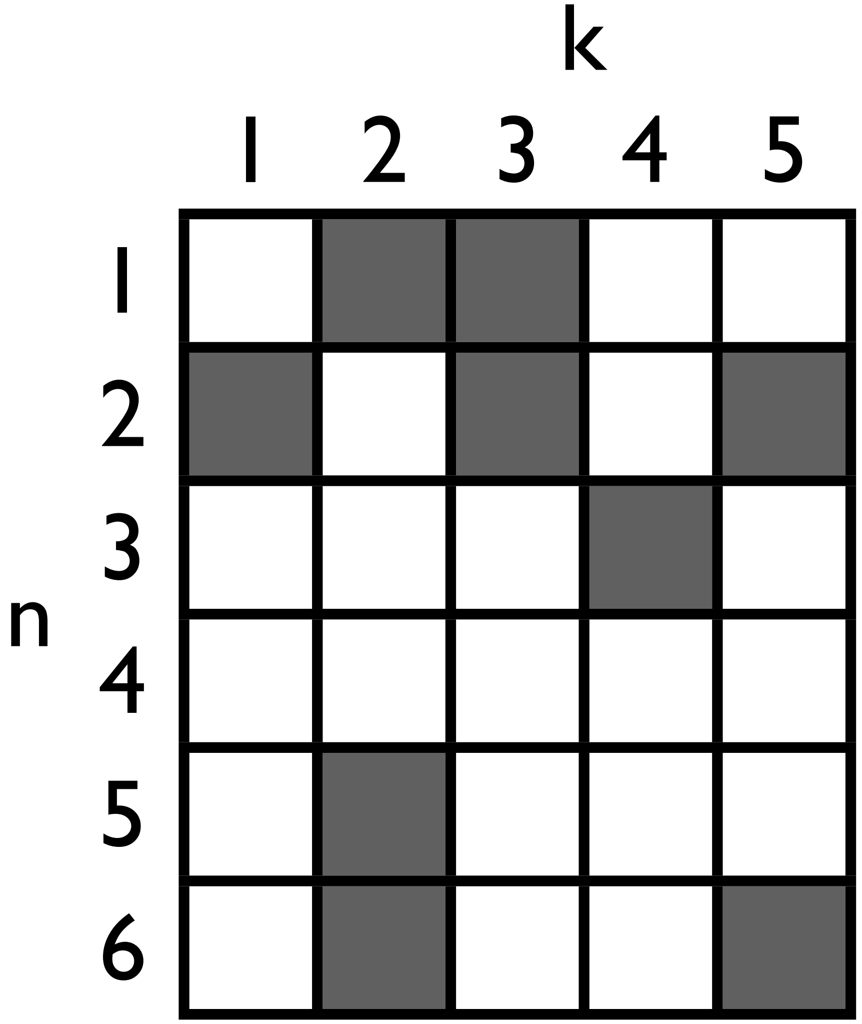
Frequency model



Frequency model



Frequency model



Random labels

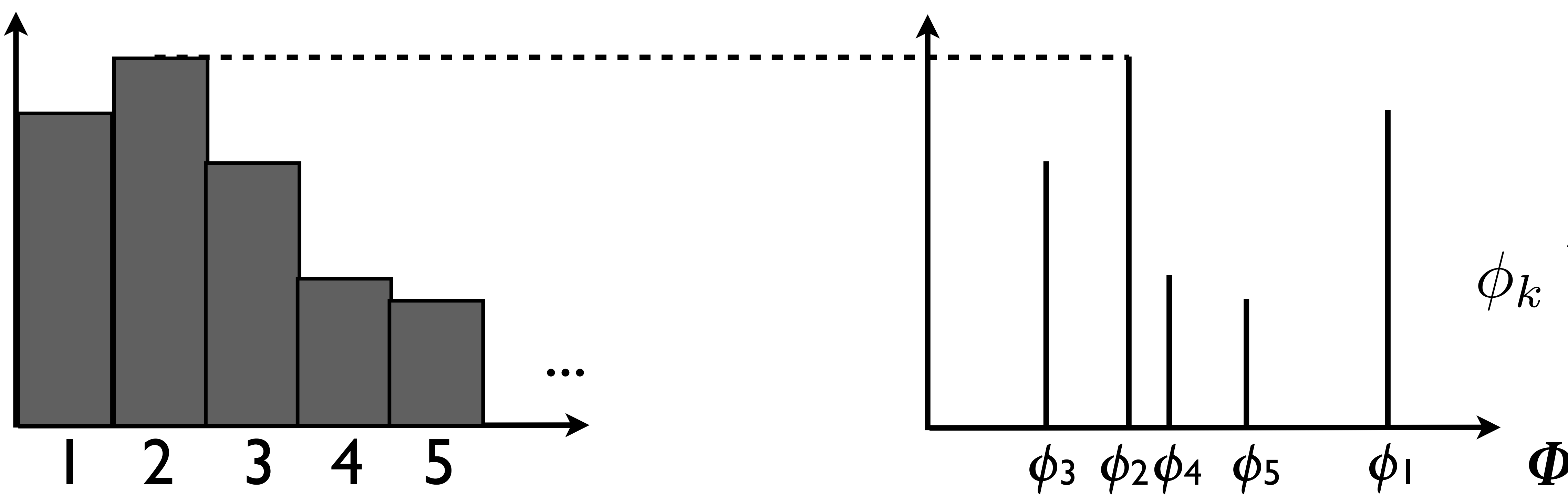
$Z'_1 = \{\phi_2, \phi_3\}$

$Z'_2 = \{\phi_1, \phi_3, \phi_5\}$

$Z'_3 = \{\phi_4\}$

$Z'_4 = \{\}$

$Z'_5 = \{\phi_2\}$



$$\phi_k \stackrel{iid}{\sim} H$$

Frequency model

ϕ_2

ϕ_3

X_1

[Griffiths, Ghahramani 2011]

Random labels

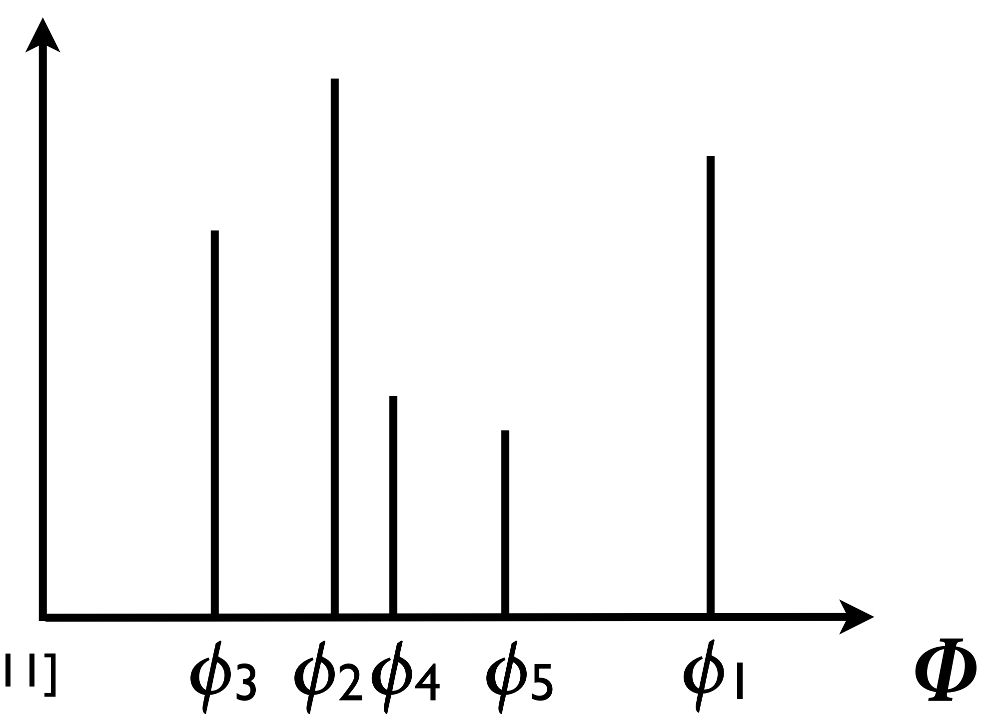
$Z'_1 = \{\phi_2, \phi_3\}$

$Z'_2 = \{\phi_1, \phi_3, \phi_5\}$

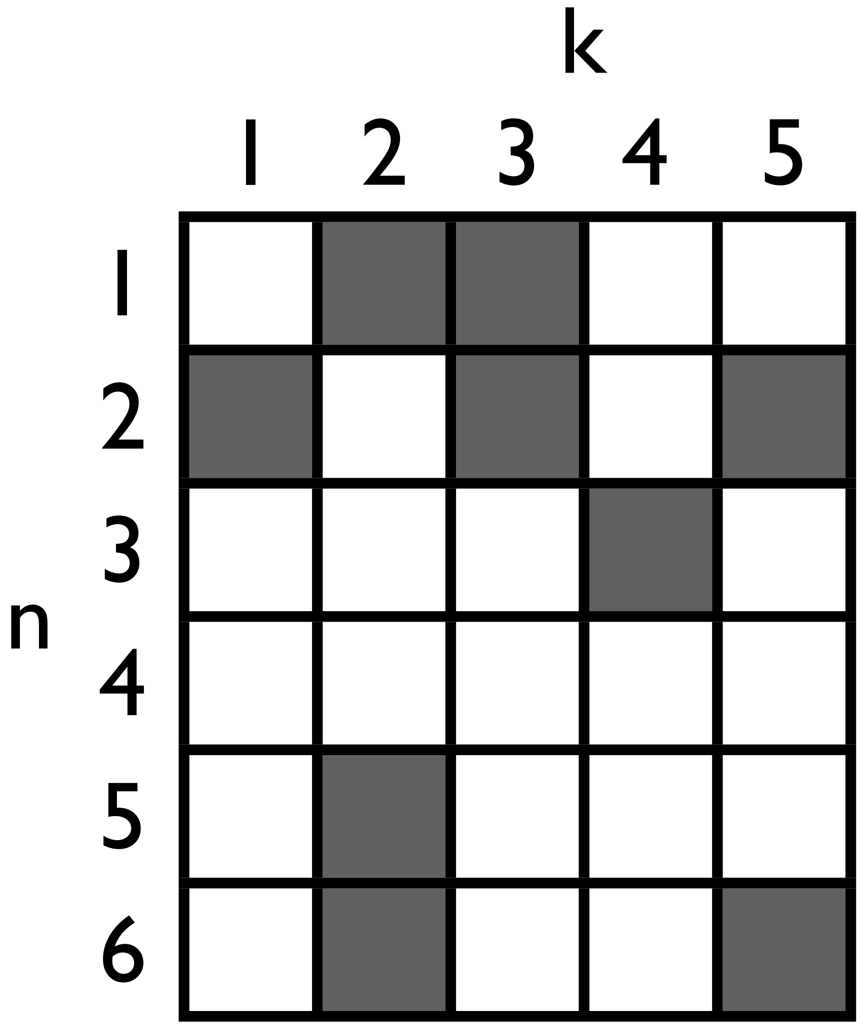
$Z'_3 = \{\phi_4\}$

$Z'_4 = \{\}$

$Z'_5 = \{\phi_2\}$



Frequency model



Random labels

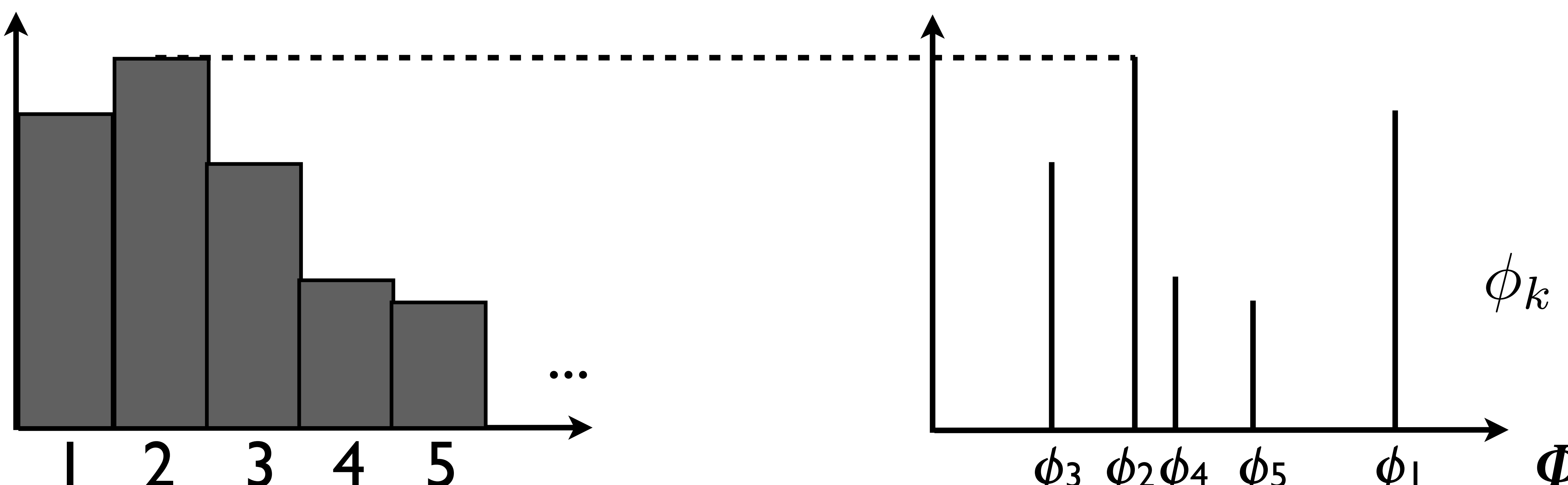
$Z'_1 = \{\phi_2, \phi_3\}$

$Z'_2 = \{\phi_1, \phi_3, \phi_5\}$

$Z'_3 = \{\phi_4\}$

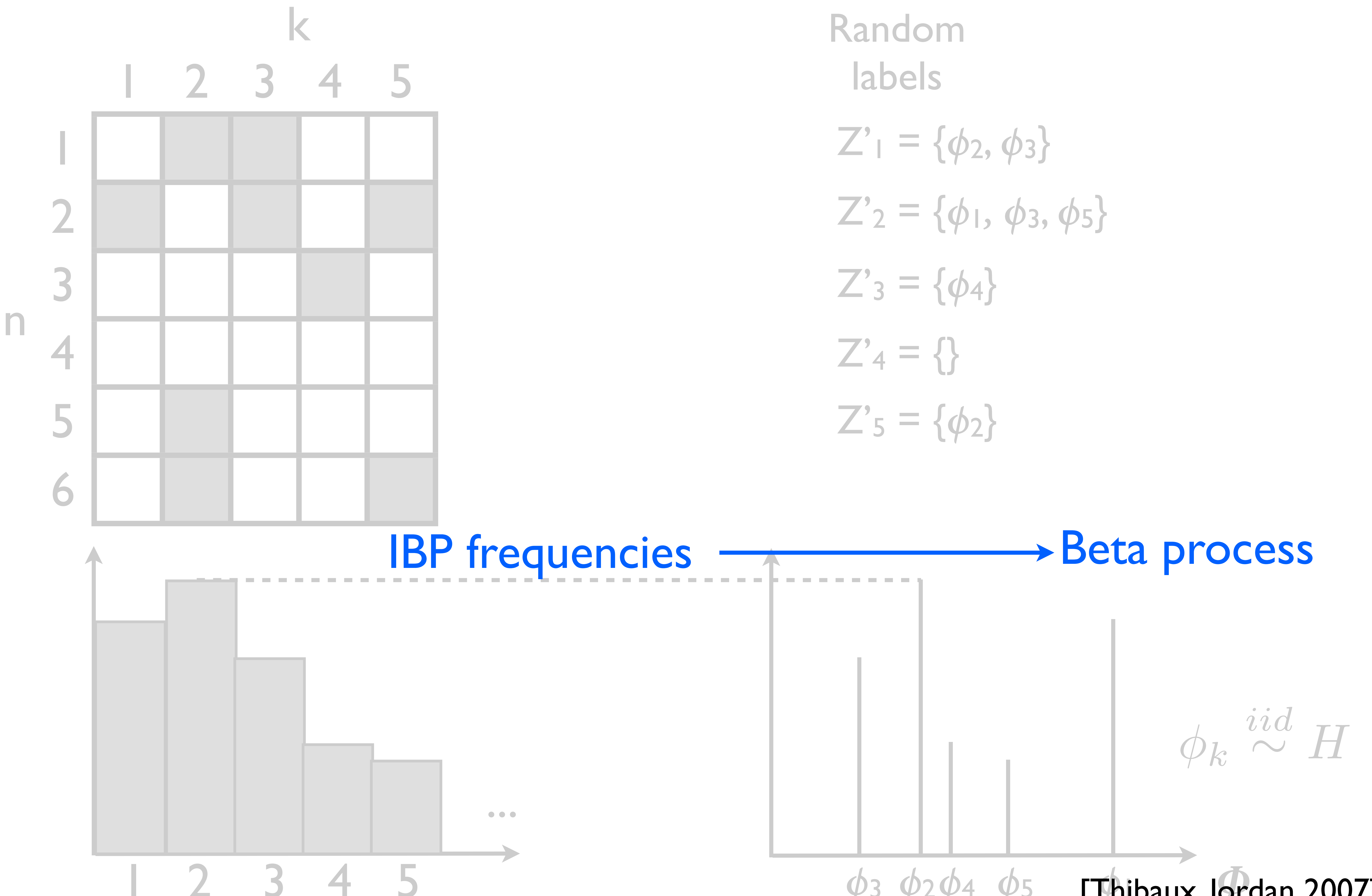
$Z'_4 = \{\}$

$Z'_5 = \{\phi_2\}$



$$\phi_k \stackrel{iid}{\sim} H$$

Frequency model



[Thibaux, Jordan 2007]

(Some) features: augmentation

(Some) features: augmentation

random feature
allocation & EFPF

(Some) features: augmentation

random feature allocation & EFPF

IBP

(Some) features: augmentation

$$f_9 = \{\{1,5,6\},\{8\},\{8\},\{8,5\},\{1,6\}\}$$

random feature allocation & EFPF

IBP

(Some) features: augmentation

$$f_9 = \{\{1,5,6\},\{8\},\{8\},\\ \{8,5\},\{1,6\}\}$$

random feature allocation & EFPF

IBP

(continuous-valued) random feature labels

(Some) features: augmentation

$$f_9 = \{\{1,5,6\},\{8\},\{8\},\\ \{8,5\},\{1,6\}\}$$

random feature allocation & EFPF

IBP

(continuous-valued) random feature labels

IBP with feature means

(Some) features: augmentation

$$f_9 = \{\{1,5,6\},\{8\},\{8\},\{8,5\},\{1,6\}\}$$

$Z'_1 = \{\phi_1, \phi_2\}, \ldots, Z'_9 = \{\}$

random feature allocation & EFPF

IBP

(continuous-valued) random feature labels

IBP with feature means

(Some) features: augmentation

$$f_9 = \{\{1,5,6\},\{8\},\{8\},\{8,5\},\{1,6\}\}$$

$Z'_1 = \{\phi_1, \phi_2\}, \ldots, Z'_9 = \{\}$

random feature allocation & EFPF

IBP

(continuous-valued) random feature labels

IBP with feature means

feature frequencies

(Some) features: augmentation

$$f_9 = \{\{1,5,6\},\{8\},\{8\},$$
$$\{8,5\},\{1,6\}\}$$

Z'$_1$ = $\{\phi_1, \phi_2\}$, ..., Z'$_9$ = {}

random feature allocation & EFPF

IBP

(continuous-valued) random feature labels

IBP with feature means

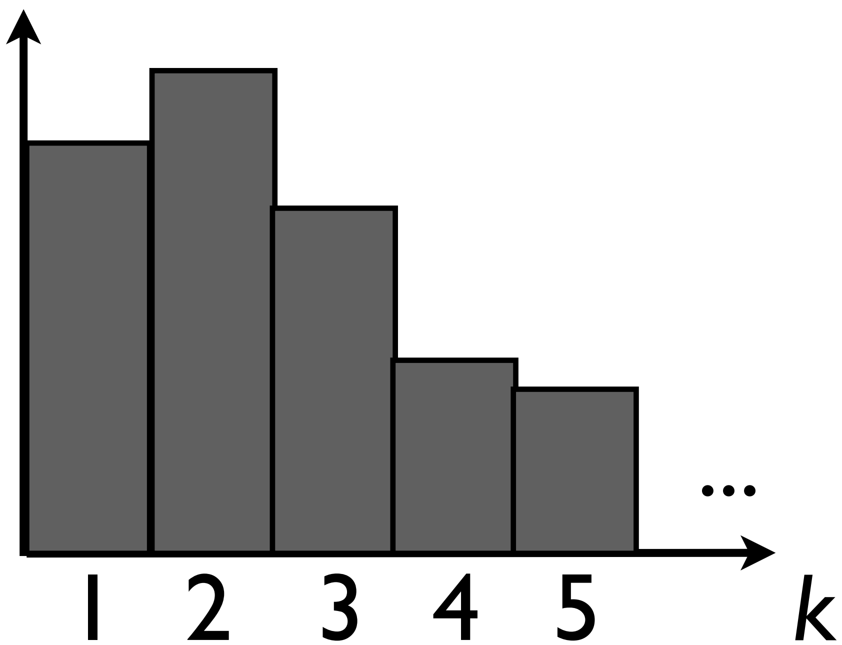
feature frequencies

IBP frequencies

(Some) features: augmentation

$$f_9 = \{\{1,5,6\},\{8\},\{8\},\{8,5\},\{1,6\}\}$$

$Z'_1 = \{\phi_1, \phi_2\}, \ldots, Z'_9 = \{\}$



random feature allocation & EFPF

IBP

(continuous-valued) random feature labels

IBP with feature means

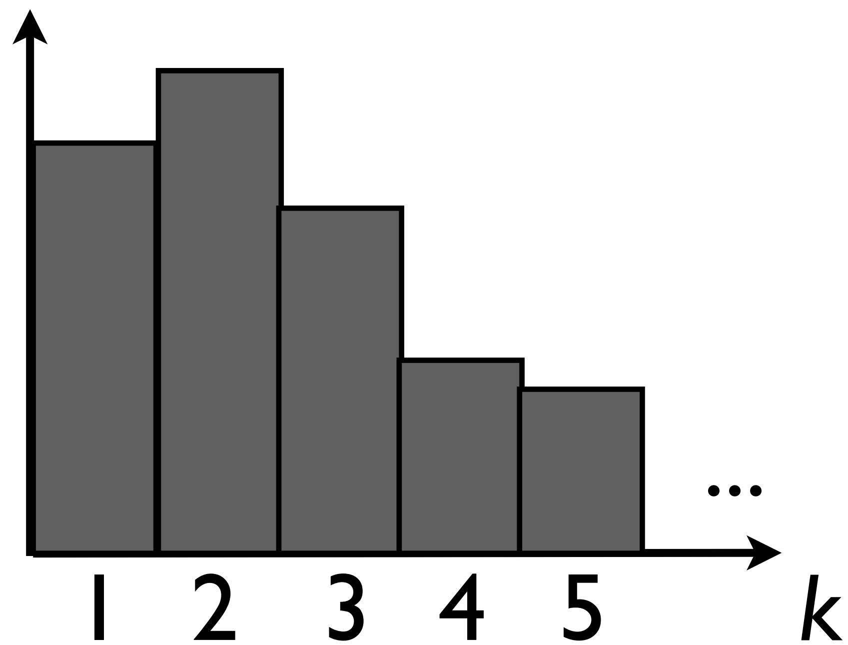
feature frequencies

IBP frequencies

(Some) features: augmentation

$$f_9 = \{\{1,5,6\},\{8\},\{8\},\{8,5\},\{1,6\}\}$$

Z'[1] = {ϕ_1, ϕ_2}, ..., Z'[9] = {}



random feature allocation & EFPF

IBP

(continuous-valued) random feature labels

IBP with feature means

feature frequencies

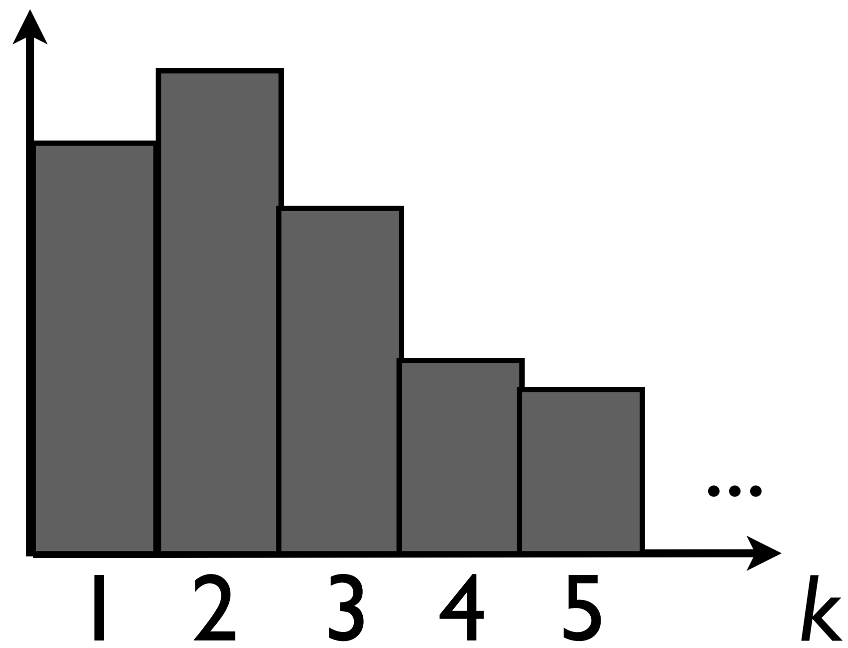
IBP frequencies

random, discrete measure with atoms in [0,1]

(Some) features: augmentation

$$f_9 = \{\{1,5,6\},\{8\},\{8\},\{8,5\},\{1,6\}\}$$

$Z'_1 = \{\phi_1, \phi_2\}, \ldots, Z'_9 = \{\}$



random feature allocation & EFPF

IBP

(continuous-valued) random feature labels

IBP with feature means

feature frequencies

IBP frequencies

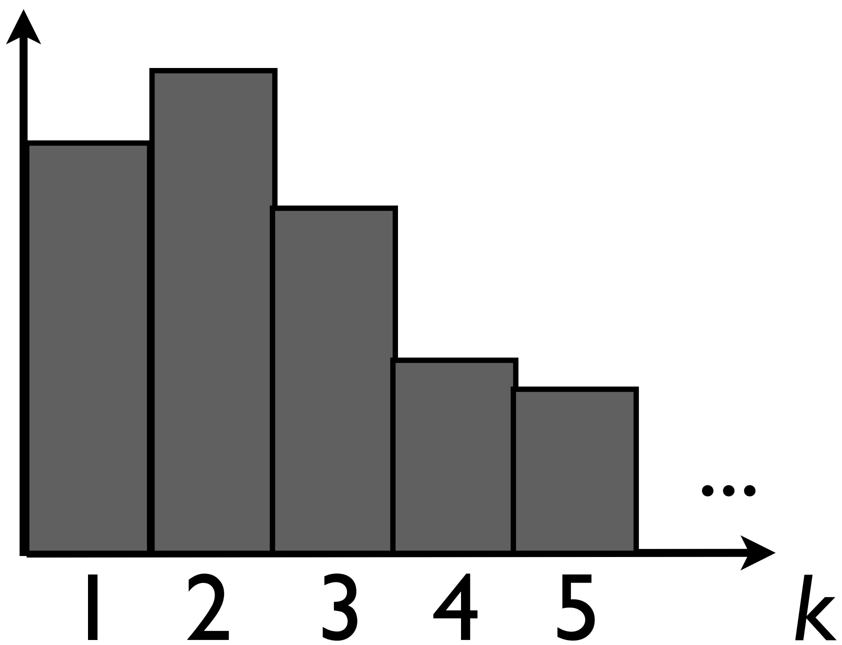
random, discrete measure with atoms in [0,1]

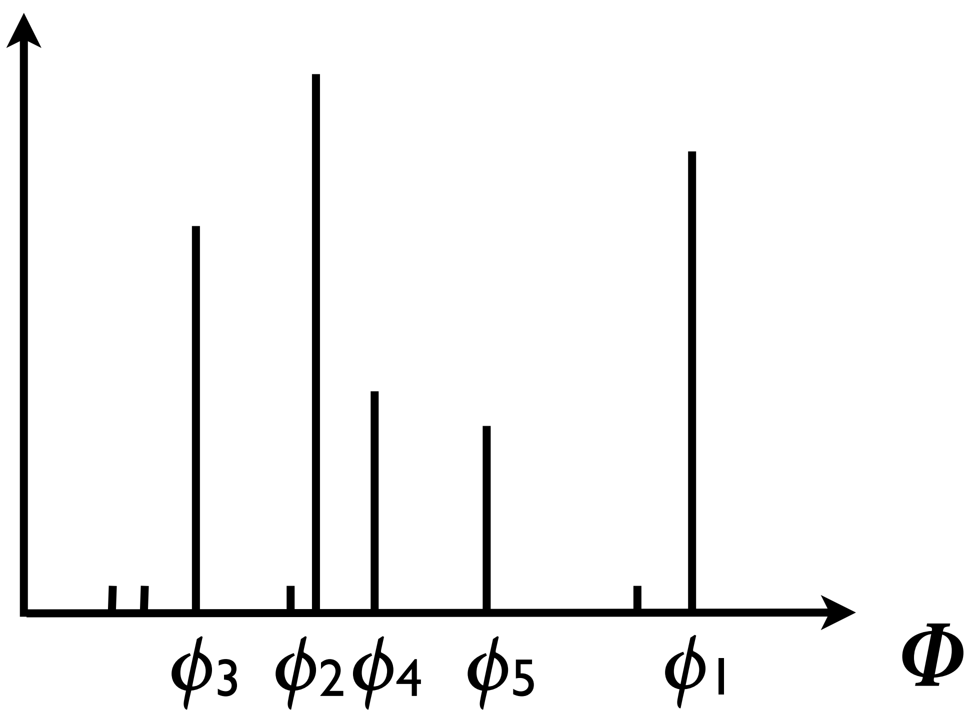
beta process

(Some) features: augmentation

$$f_9 = \{\{1,5,6\},\{8\},\{8\},\{8,5\},\{1,6\}\}$$

$Z'_1 = \{\phi_1, \phi_2\}, \ldots, Z'_9 = \{\}$





random feature allocation & EFPF

IBP

(continuous-valued) random feature labels

IBP with feature means

feature frequencies

IBP frequencies

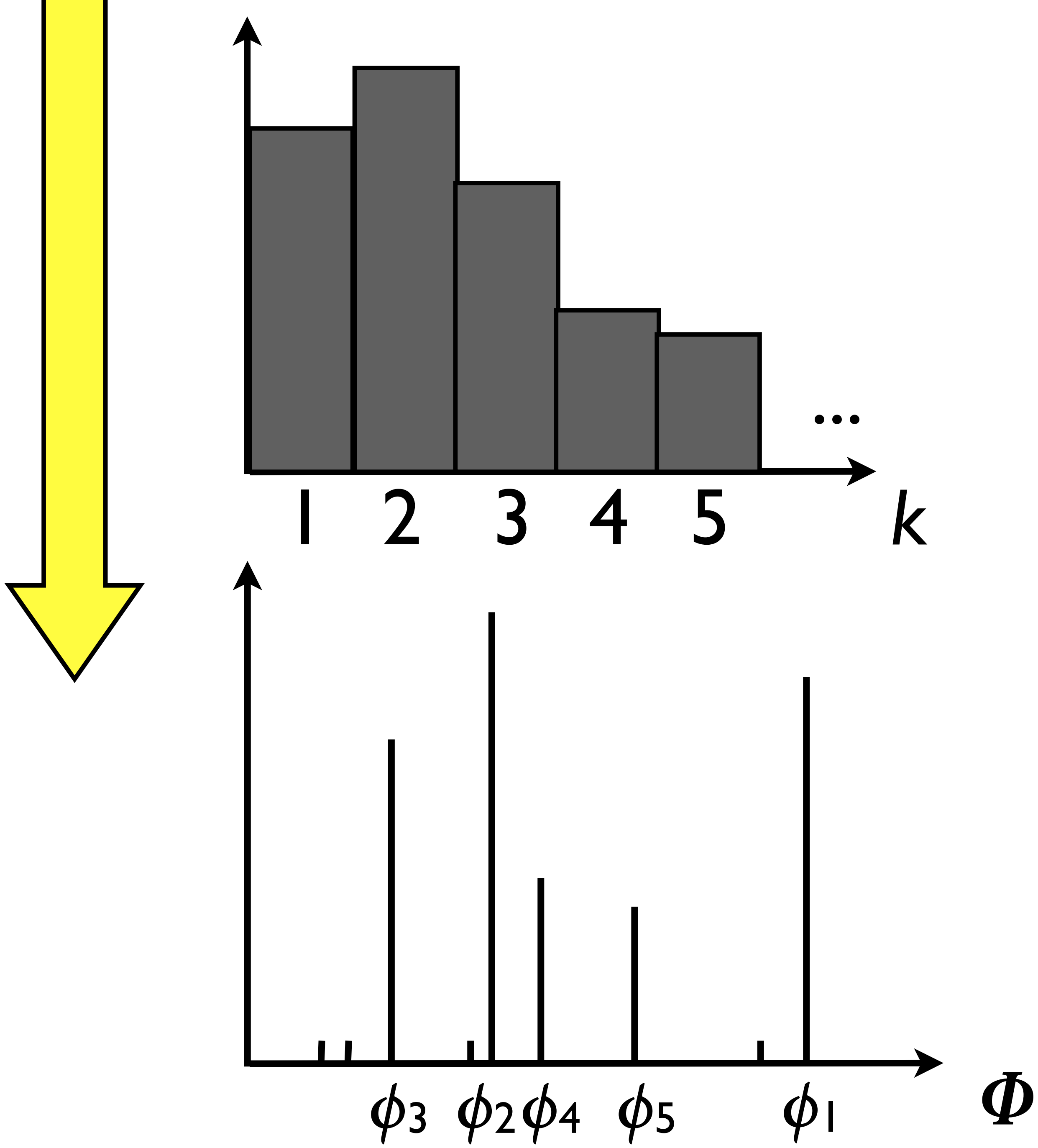
random, discrete measure with atoms in [0,1]

beta process

(Some) features: augmentation

$$f_9 = \{\{1,5,6\},\{8\},\{8\},\{8,5\},\{1,6\}\}$$

Z'₁ = {ϕ_1, ϕ_2}, ..., Z'₉ = {}



random feature allocation & EFPF

IBP

(continuous-valued) random feature labels

IBP with feature means

feature frequencies

IBP frequencies

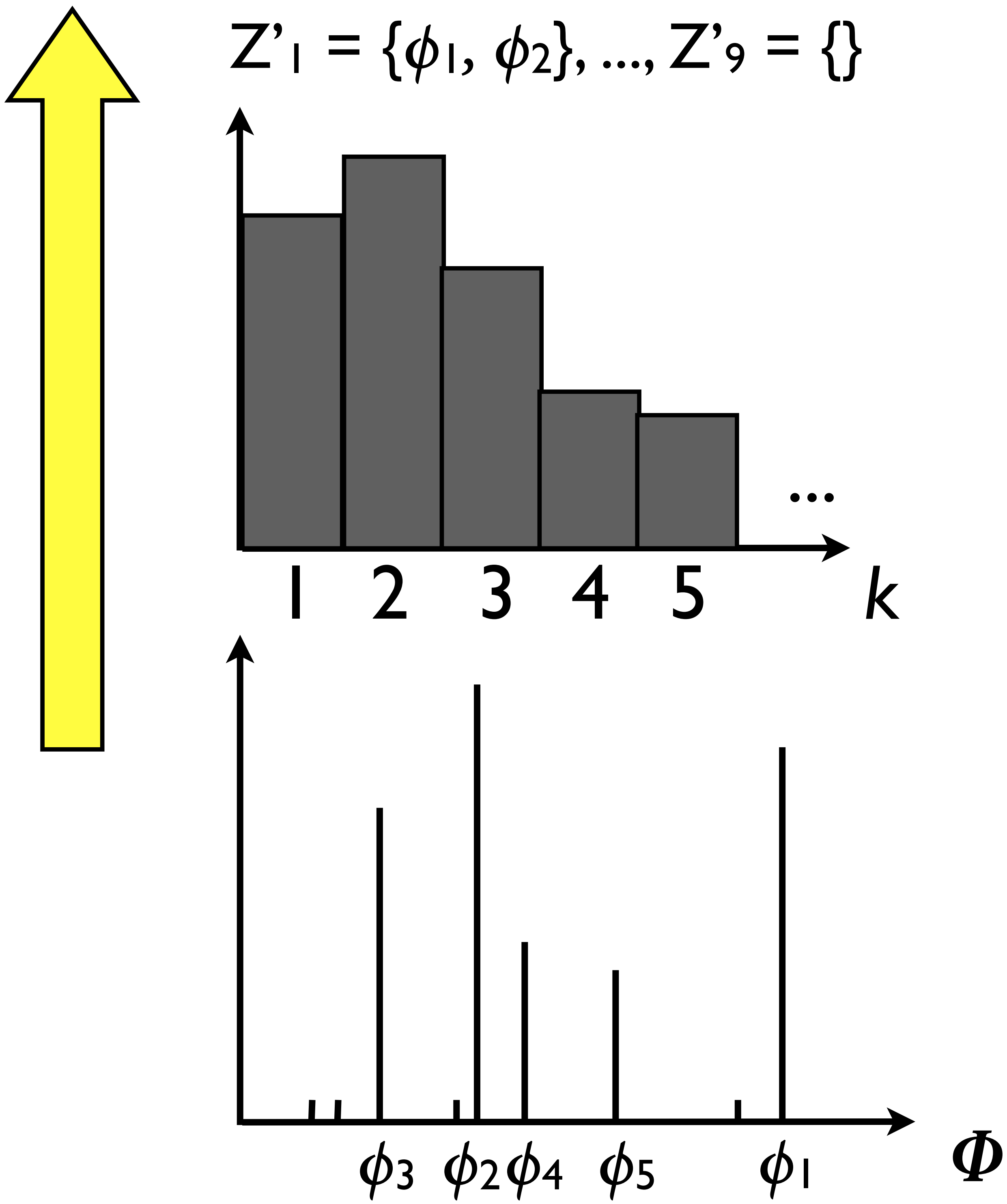
random, discrete measure with atoms in [0,1]

beta process

(Some) features: integrating out

$$f_9 = \{\{1,5,6\},\{8\},\{8\},\{8,5\},\{1,6\}\}$$

Z'₁ = {ϕ_1, ϕ_2}, ..., Z'₉ = {}



random feature allocation & EFPF

IBP

(continuous-valued) random feature labels

IBP with feature means

feature frequencies

IBP frequencies

random, discrete measure with atoms in [0,1]

beta process

Outline

I. Clusters

II. Features

- Overview
- Distribution
- Frequencies
- Random measure
- Paintbox

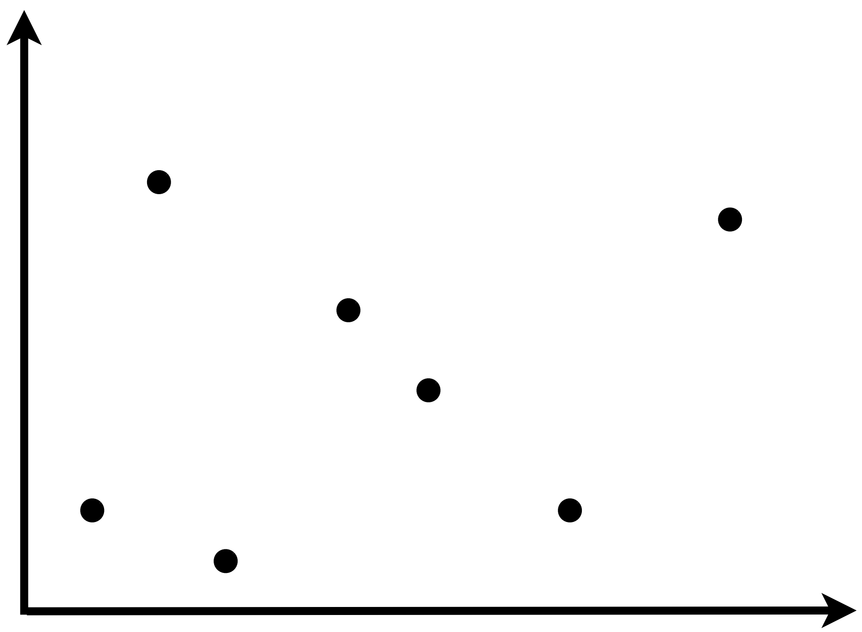
Outline

I. Clusters

II. Features

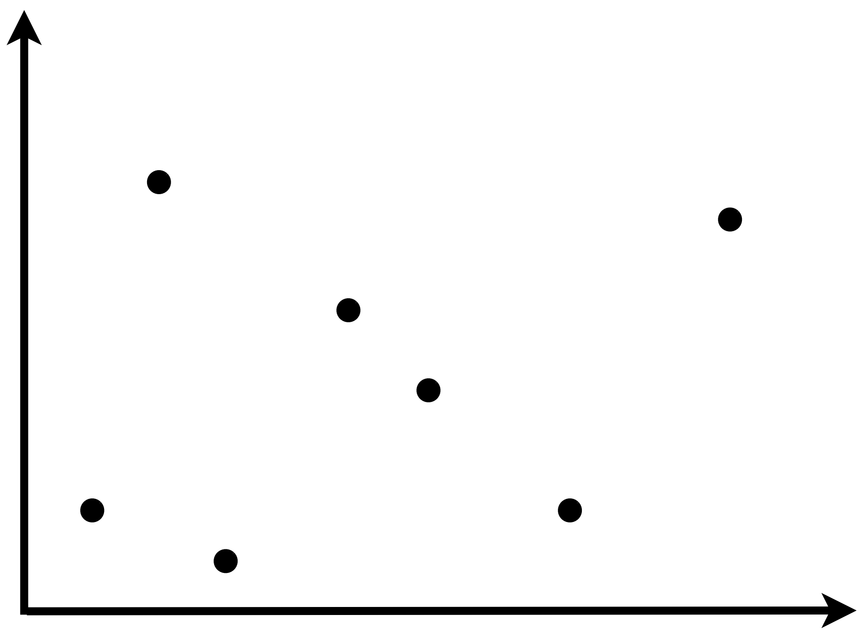
- Overview
- Distribution
- Frequencies
- Random measure
  - ◇ Poisson process
- Paintbox

Aside: Poisson process



[Kingman 1993]

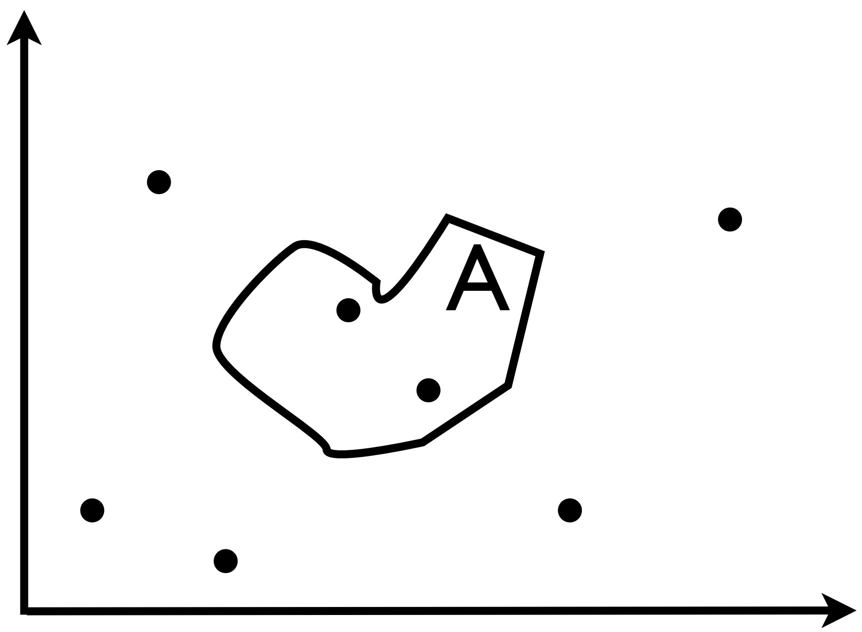
Aside: Poisson process



$$M \sim \mathrm{PP}(\nu)$$

[Kingman 1993]

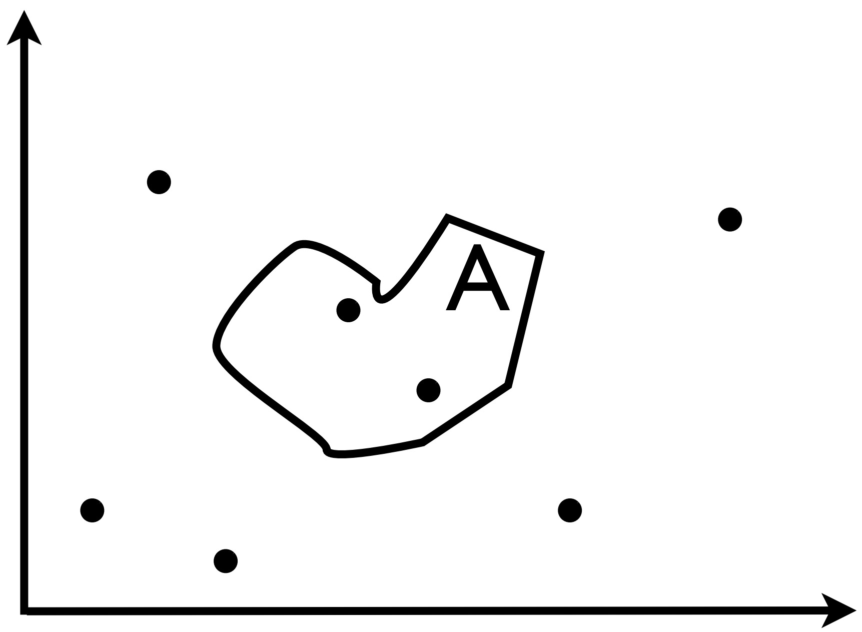
Aside: Poisson process



$$M \sim \mathrm{PP}(\nu)$$

[Kingman 1993]

Aside: Poisson process

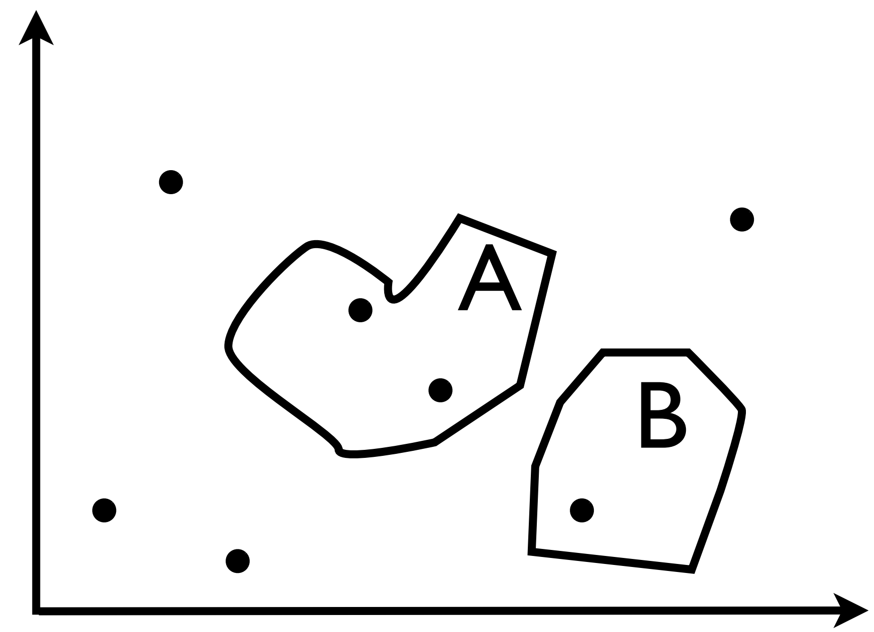


$$M \sim \mathrm{PP}(\nu)$$

$$M(A) \sim \mathrm{Poisson}(\nu(A))$$

[Kingman 1993]

Aside: Poisson process

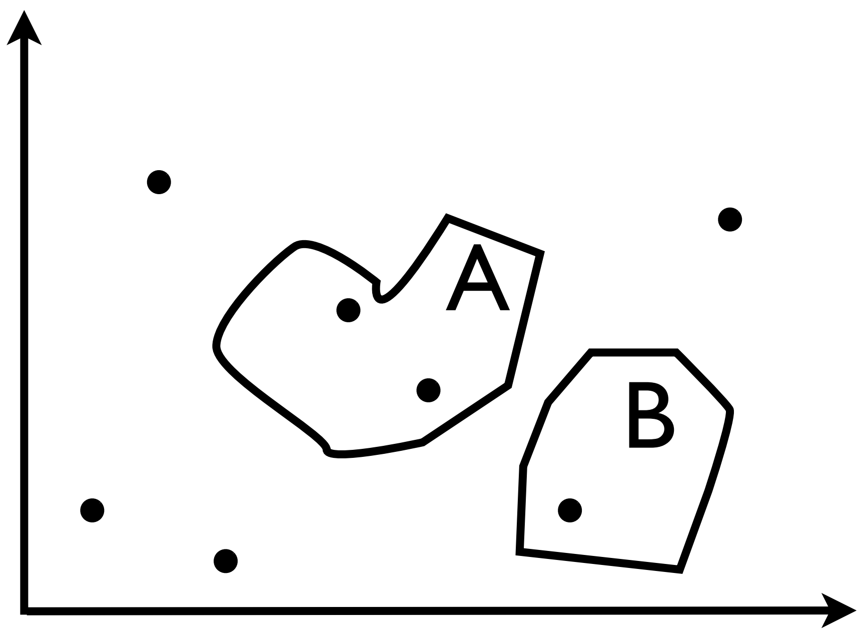


$$M \sim \mathrm{PP}(\nu)$$

$$M(A) \sim \mathrm{Poisson}(\nu(A))$$

[Kingman 1993]

Aside: Poisson process



$$M \sim \mathrm{PP}(\nu)$$

$$M(A) \sim \mathrm{Poisson}(\nu(A))$$

$$M(A) \perp M(B)$$

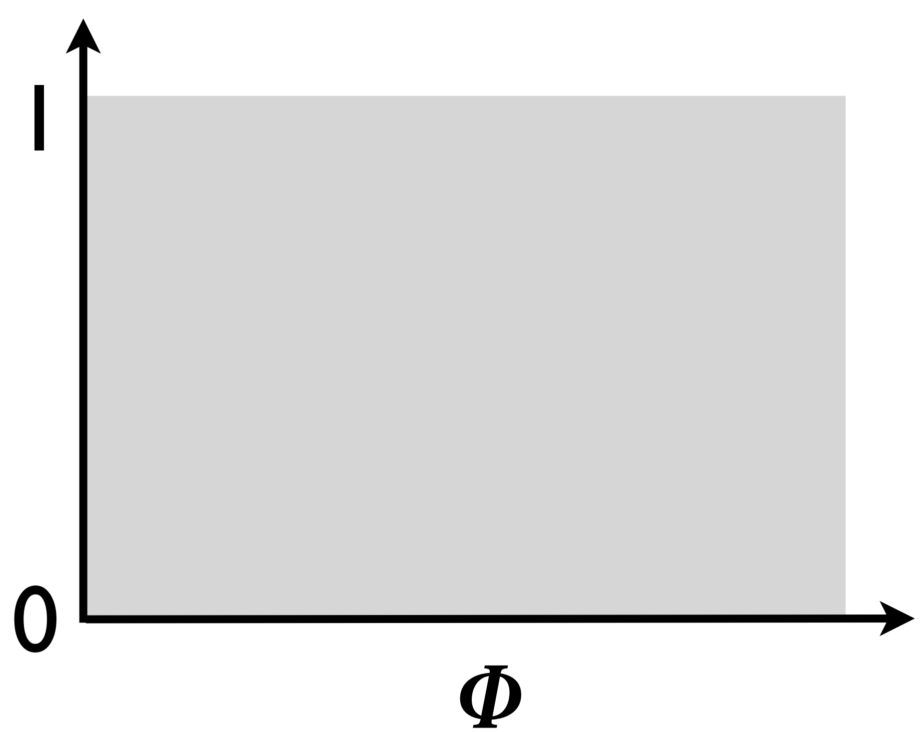
[Kingman 1993]

Aside: Poisson process

$$M \sim \mathrm{PP}(\nu)$$

$$M(A) \sim \mathrm{Poisson}(\nu(A))$$

$$M(A) \perp M(B)$$

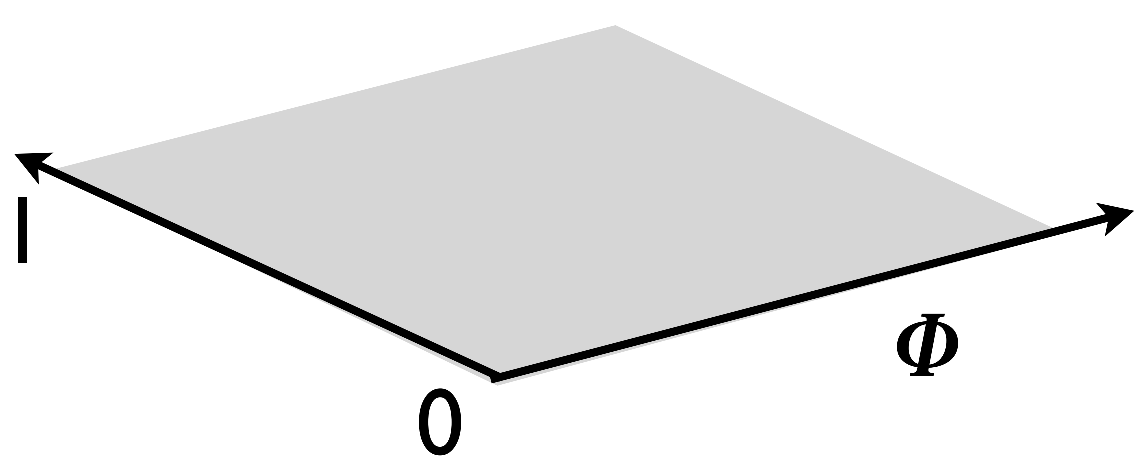


Aside: Poisson process

$$M \sim \mathrm{PP}(\nu)$$

$$M(A) \sim \mathrm{Poisson}(\nu(A))$$

$$M(A) \perp M(B)$$

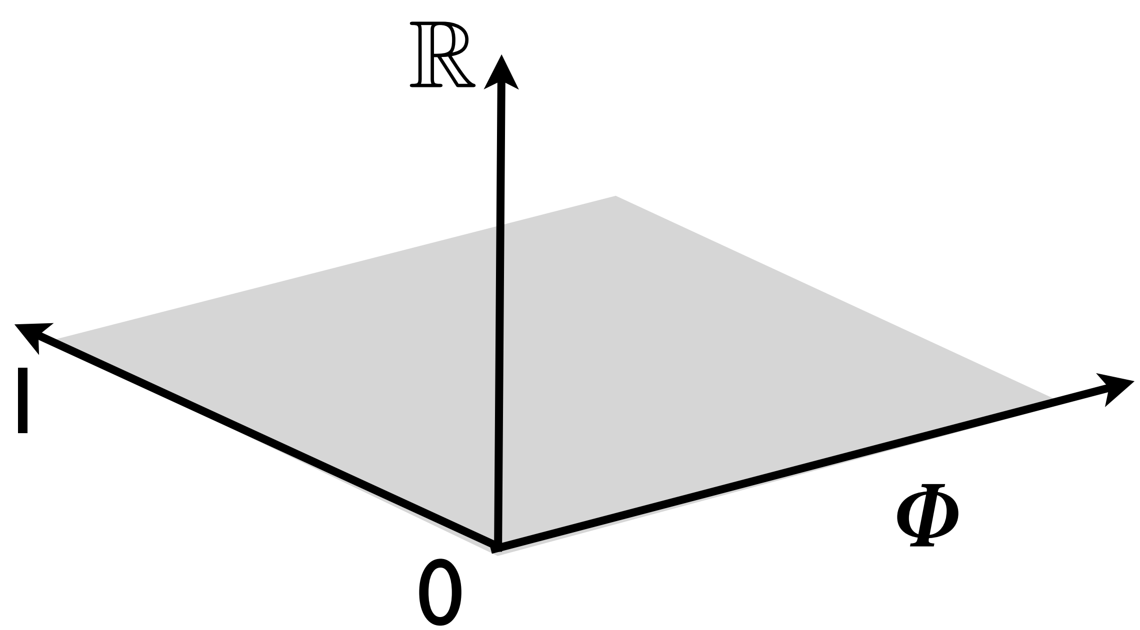


Aside: Poisson process

$$M \sim \mathrm{PP}(\nu)$$

$$M(A) \sim \mathrm{Poisson}(\nu(A))$$

$$M(A) \perp M(B)$$

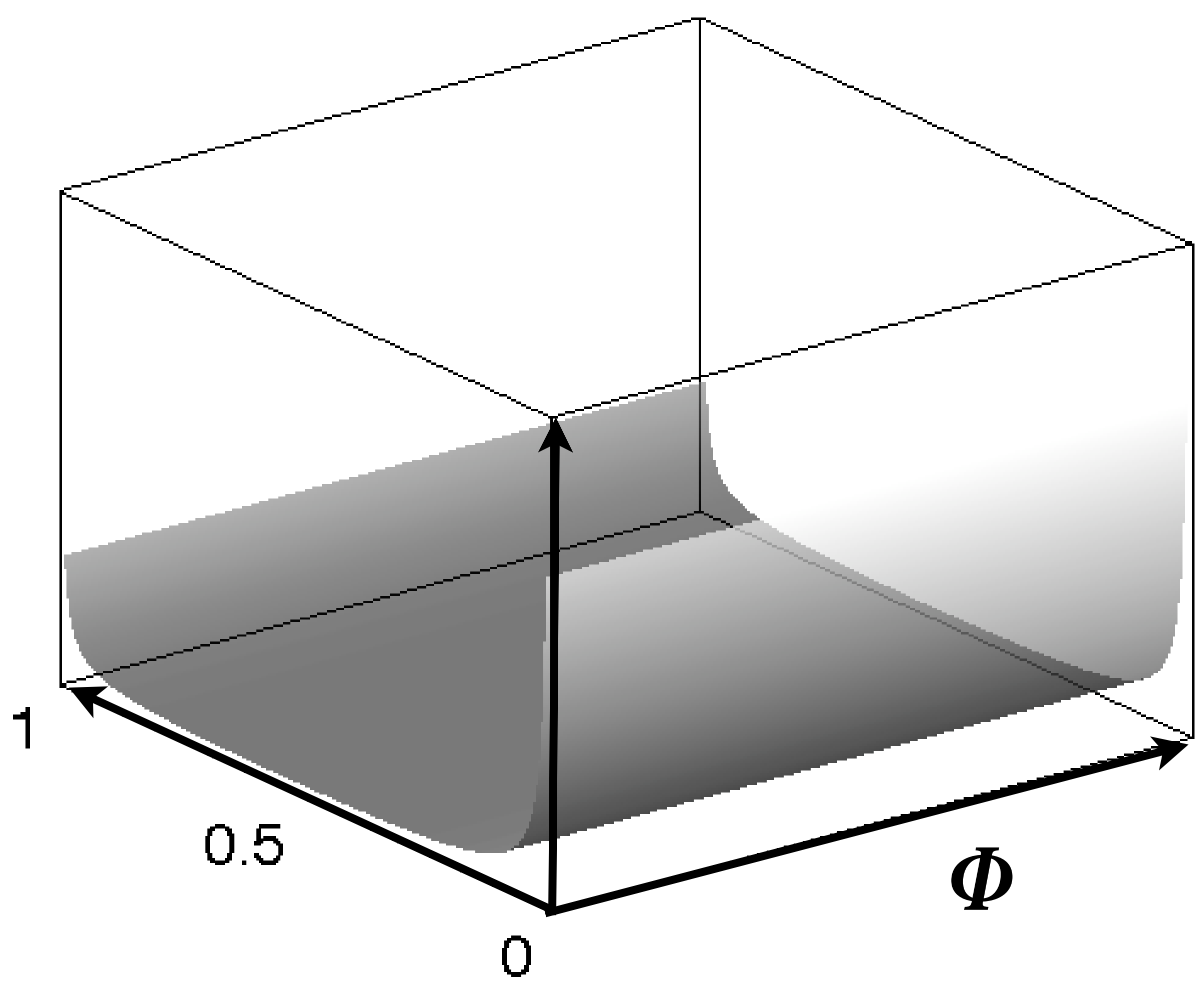


Aside: Poisson process

$$M \sim \mathrm{PP}(\nu)$$

$$M(A) \sim \mathrm{Poisson}(\nu(A))$$

$$M(A) \perp M(B)$$

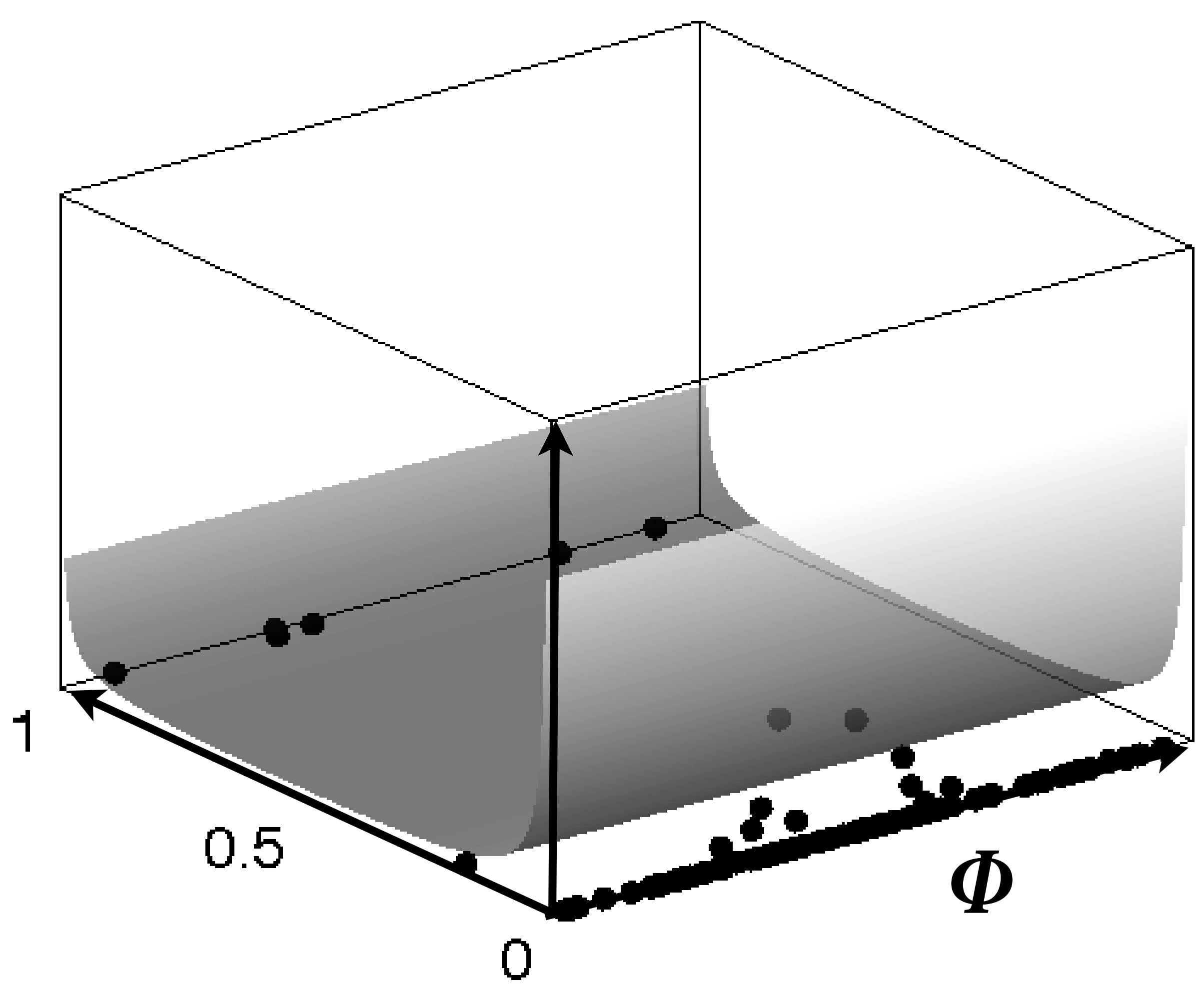


Aside: Poisson process

$$M \sim \mathrm{PP}(\nu)$$

$$M(A) \sim \mathrm{Poisson}(\nu(A))$$

$$M(A) \perp M(B)$$



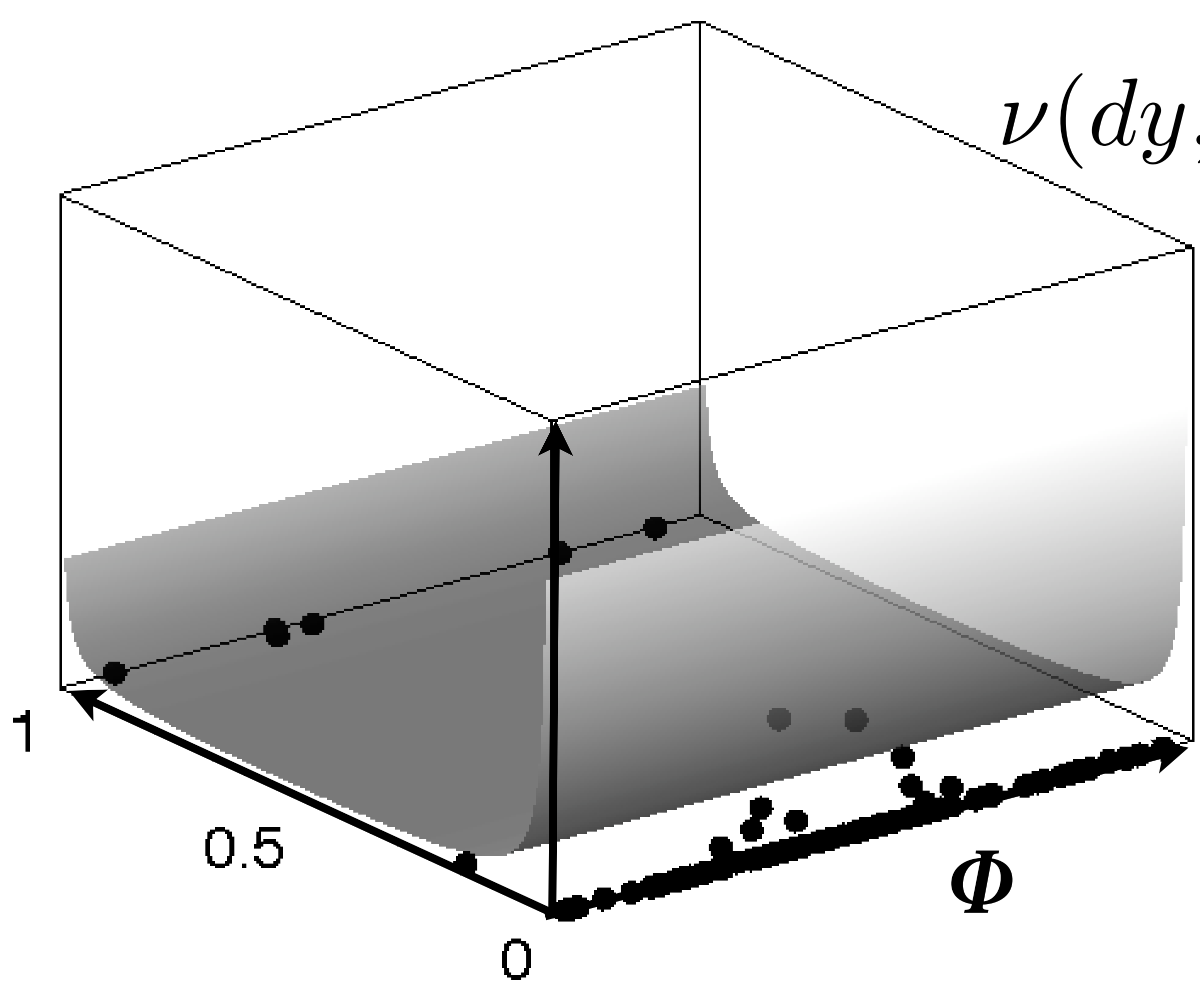
Aside: Poisson process

$$M \sim \mathrm{PP}(\nu)$$

$$M(A) \sim \mathrm{Poisson}(\nu(A))$$

$$M(A) \perp M(B)$$

$$\nu(dy, d\phi) = \gamma\theta y^{-1}(1-y)^{\theta-1}\ dy\ H(d\phi)$$



[Hjort 1990; Thibaux, Jordan 2007]

Aside: Poisson process

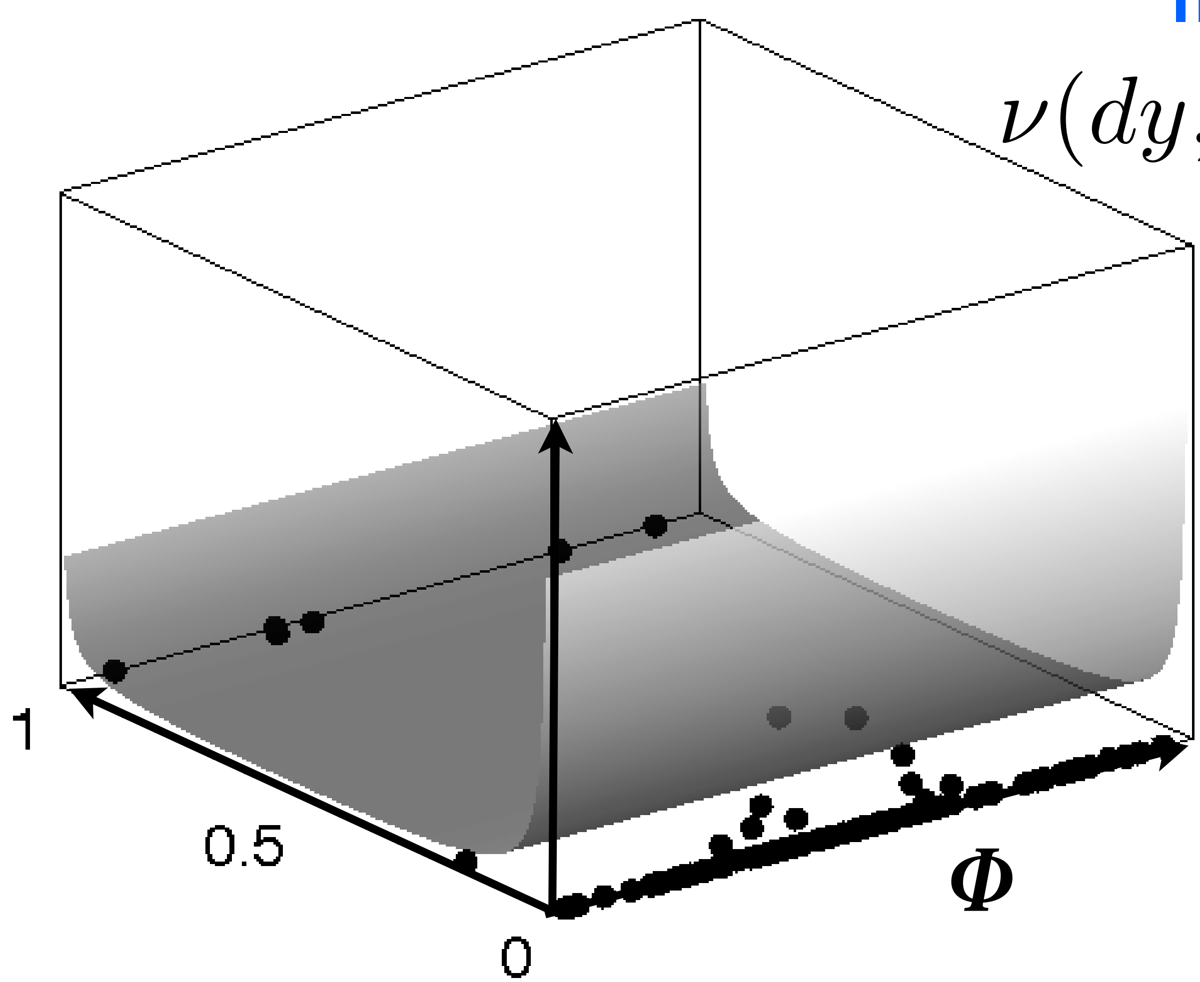
$$M \sim \mathrm{PP}(\nu)$$

$$M(A) \sim \mathrm{Poisson}(\nu(A))$$

$$M(A) \perp M(B)$$

Improper beta distribution

$$\nu(dy, d\phi) = \gamma\theta y^{-1}(1-y)^{\theta-1}\, dy\; H(d\phi)$$



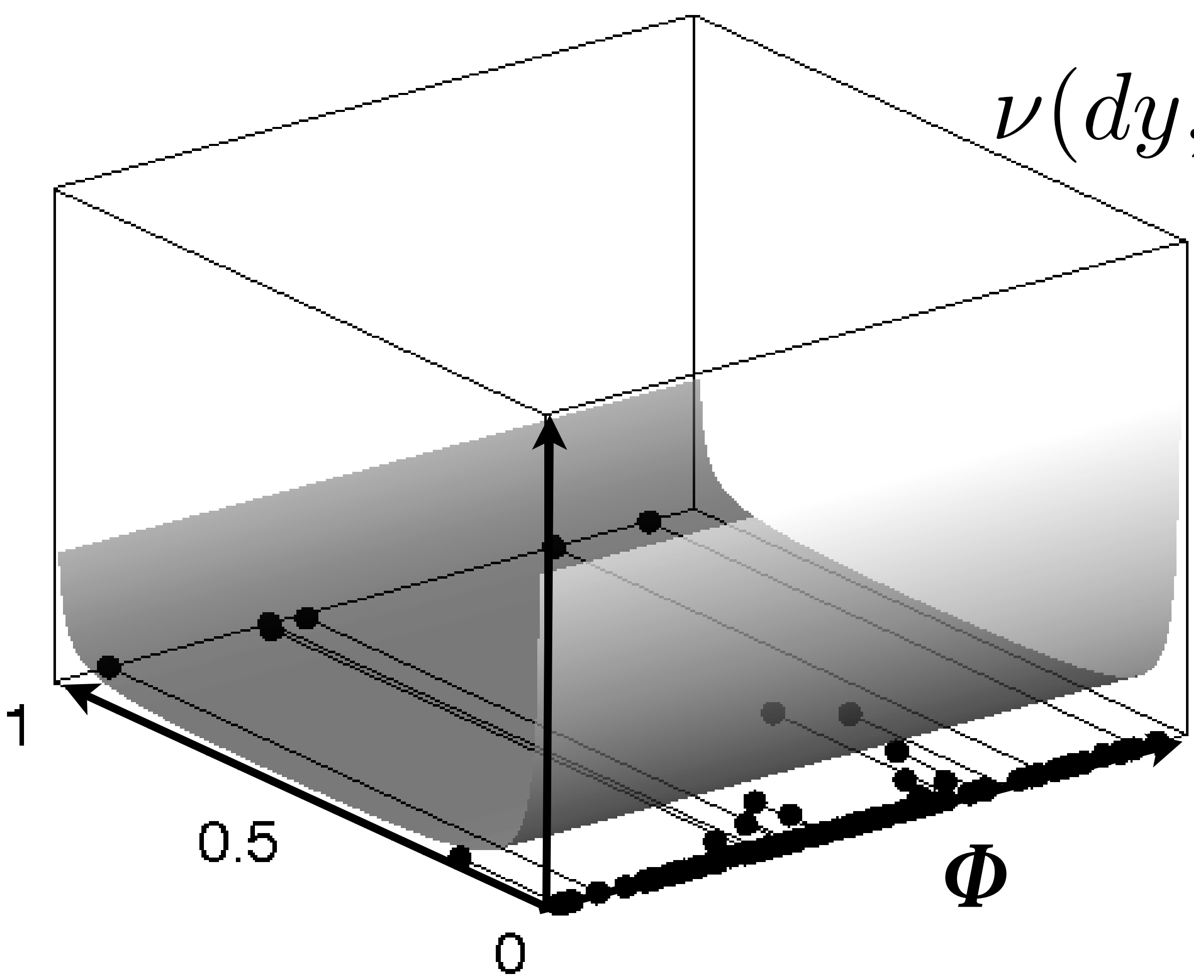
[Hjort 1990; Thibaux, Jordan 2007]

Aside: Poisson process

$$M \sim \mathrm{PP}(\nu)$$

$$M(A) \sim \mathrm{Poisson}(\nu(A))$$

$$M(A) \perp M(B)$$



$$\nu(dy, d\phi) = \gamma\theta y^{-1}(1-y)^{\theta-1}\; dy\; H(d\phi)$$

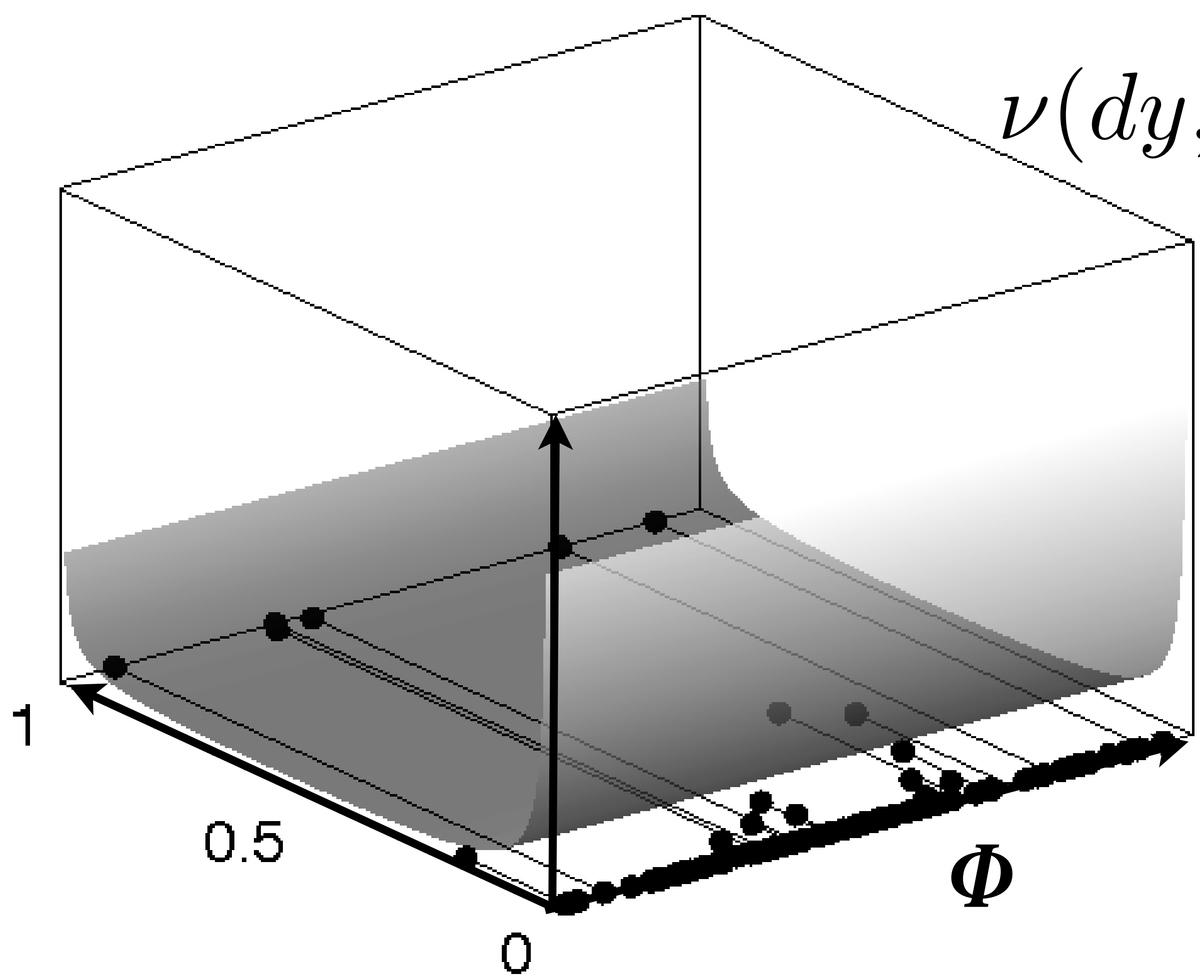
Aside: Poisson process

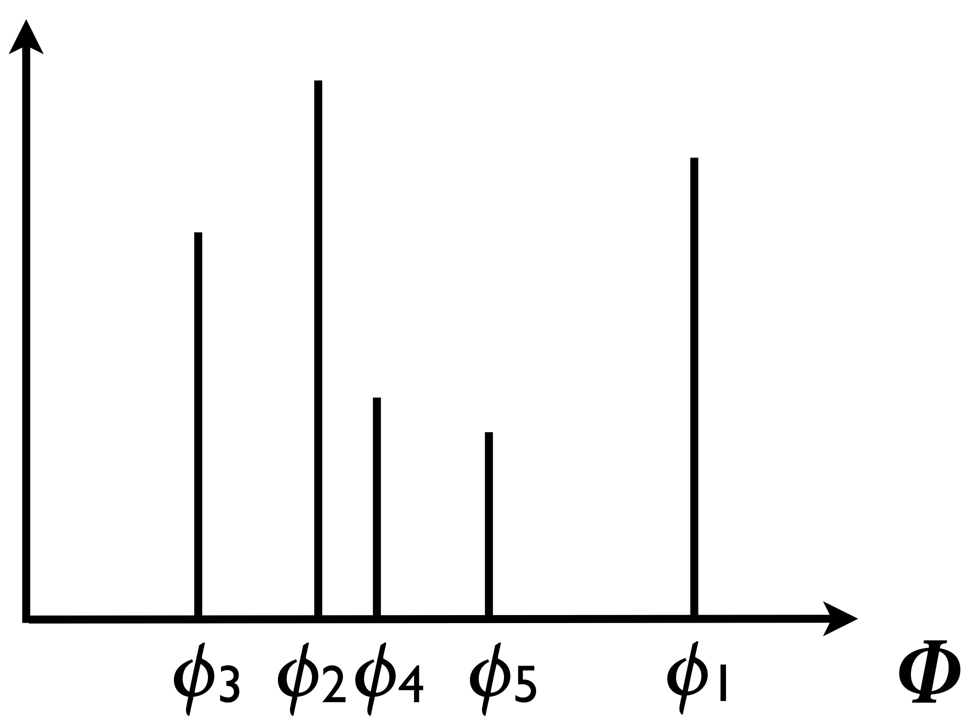
$$M \sim \mathrm{PP}(\nu)$$

$$M(A) \sim \mathrm{Poisson}(\nu(A))$$

$$M(A) \perp M(B)$$

$$\nu(dy, d\phi) = \gamma\theta y^{-1}(1-y)^{\theta-1}\ dy\ H(d\phi)$$





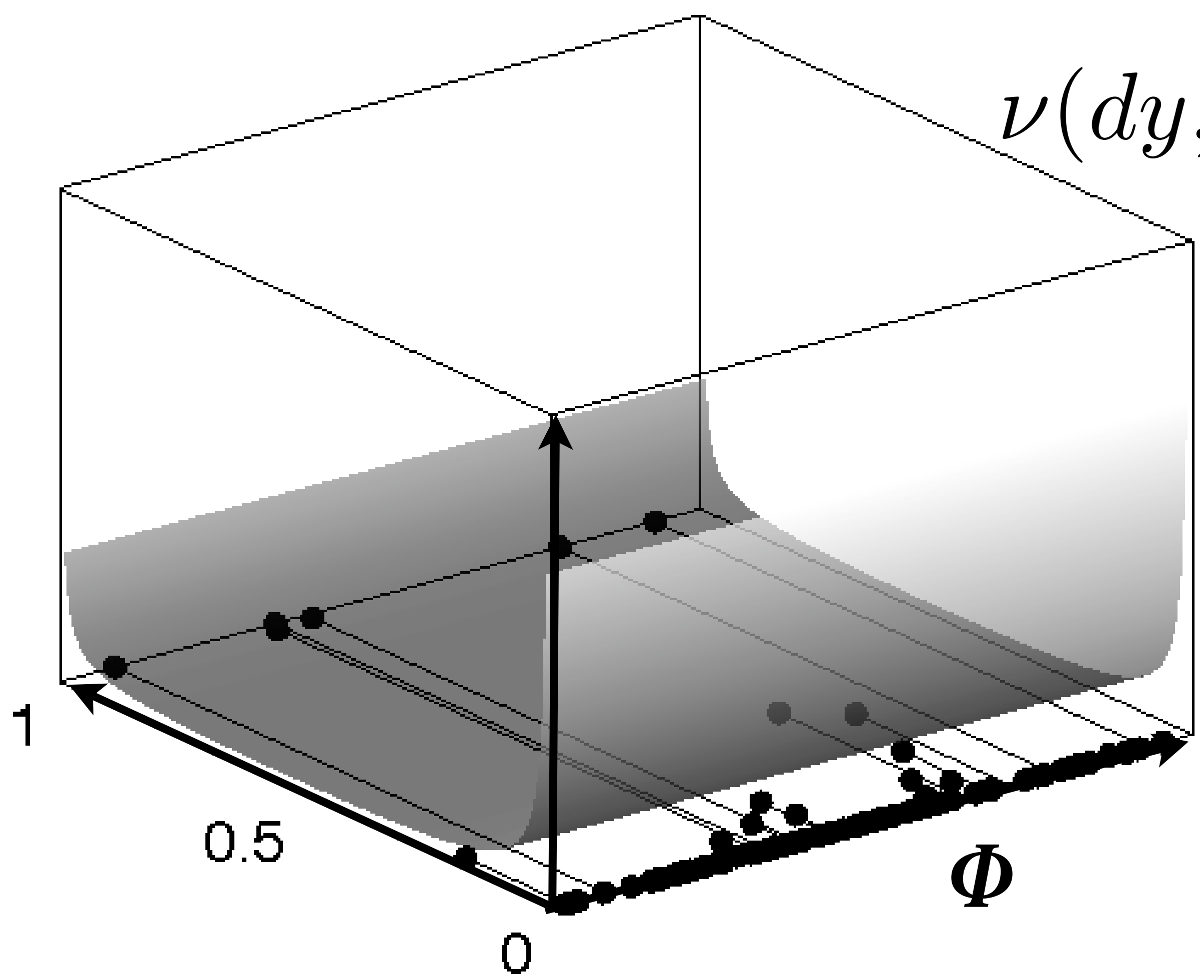
Aside: Poisson process

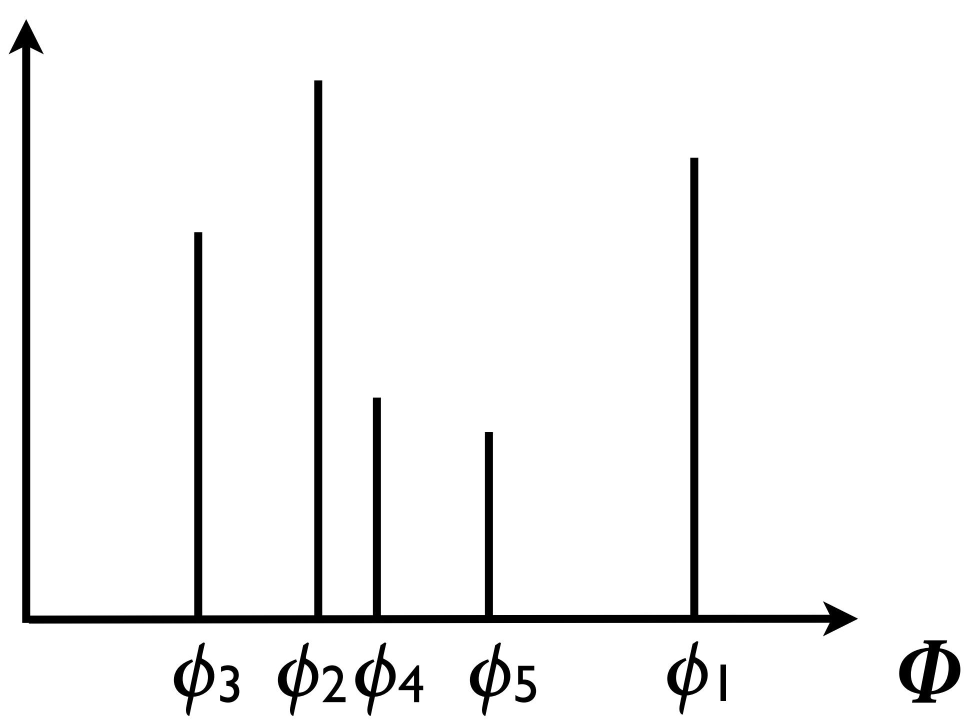
$M \sim \mathrm{PP}(\nu)$

$M(A) \sim \mathrm{Poisson}(\nu(A))$

$M(A) \perp M(B)$

$$\nu(dy, d\phi) = \gamma\theta y^{-1}(1-y)^{\theta-1}\ dy\ H(d\phi)$$



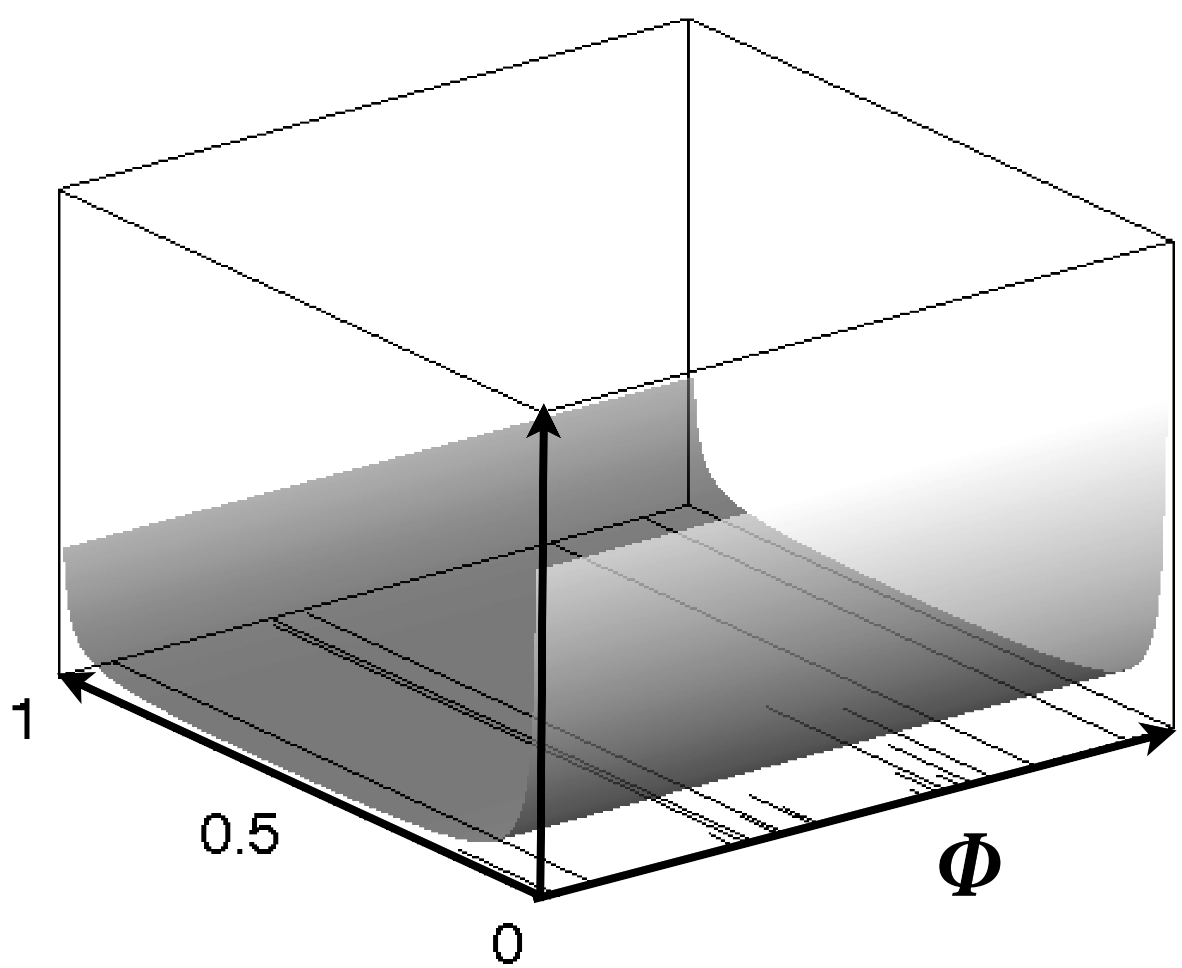


Aside: Poisson process

$$M \sim \mathrm{PP}(\nu)$$

$$M(A) \sim \mathrm{Poisson}(\nu(A))$$

$$M(A) \perp M(B)$$



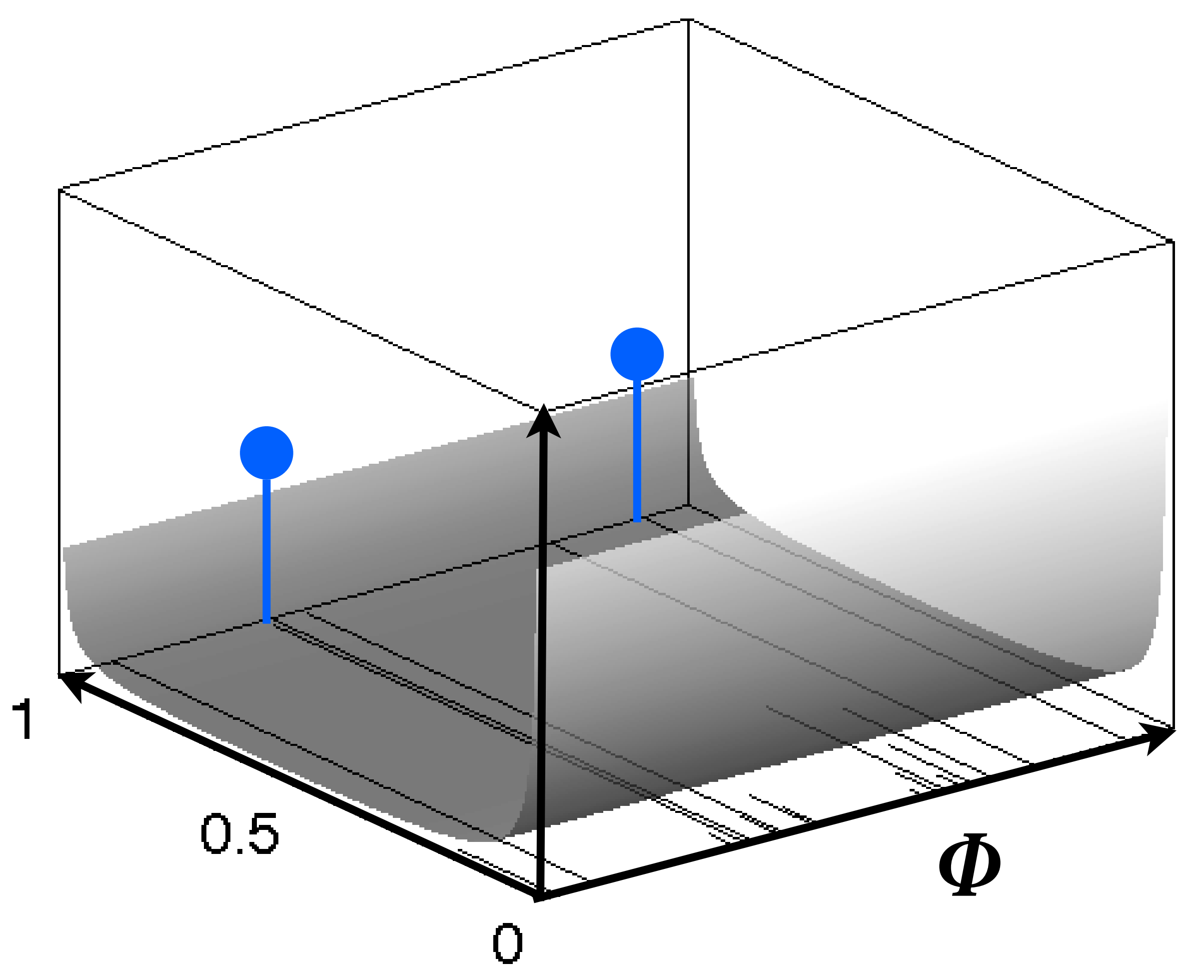
[Kingman 1993]

Aside: Poisson process

$$M \sim \mathrm{PP}(\nu)$$

$$M(A) \sim \mathrm{Poisson}(\nu(A))$$

$$M(A) \perp M(B)$$



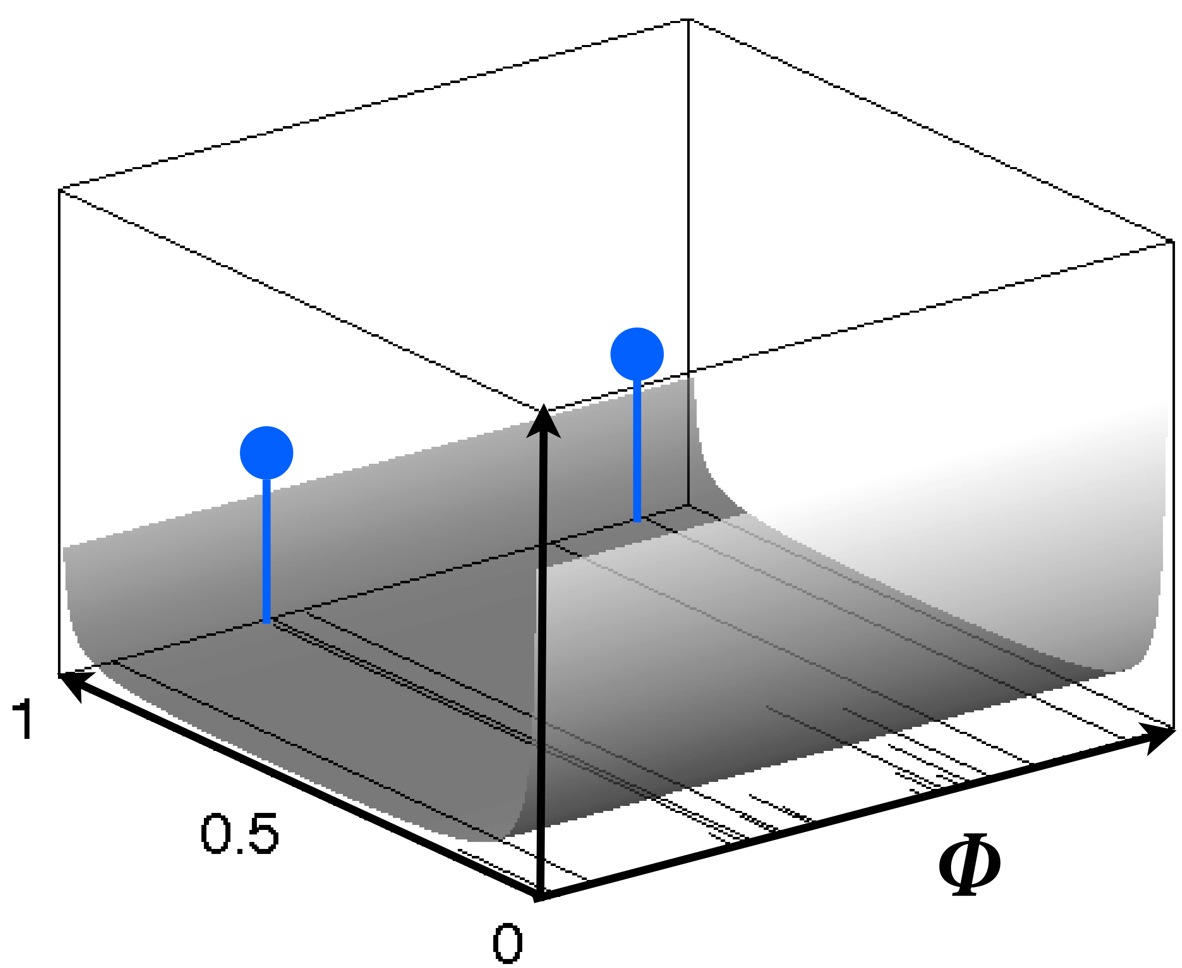
[Kingman 1993]

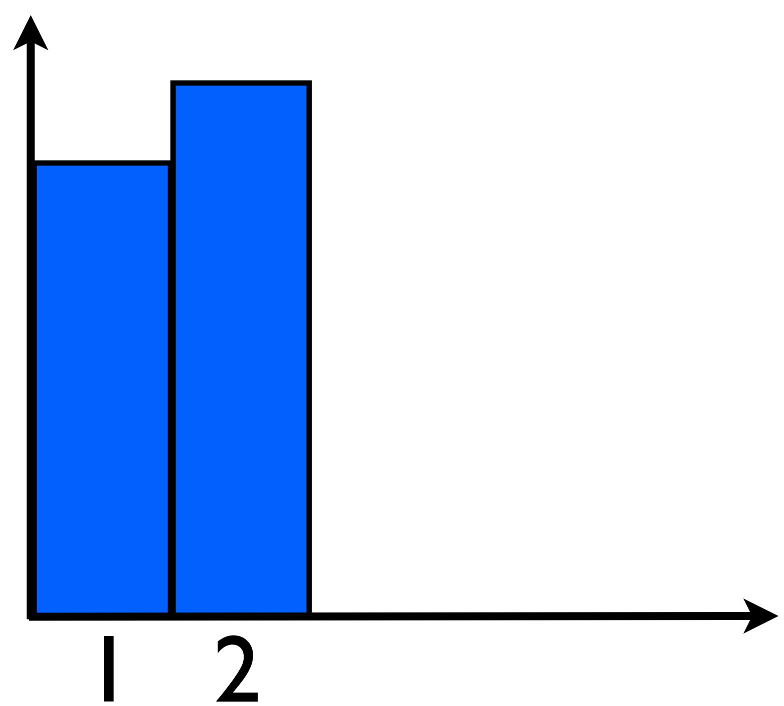
Aside: Poisson process

$$M \sim \mathrm{PP}(\nu)$$

$$M(A) \sim \mathrm{Poisson}(\nu(A))$$

$$M(A) \perp M(B)$$





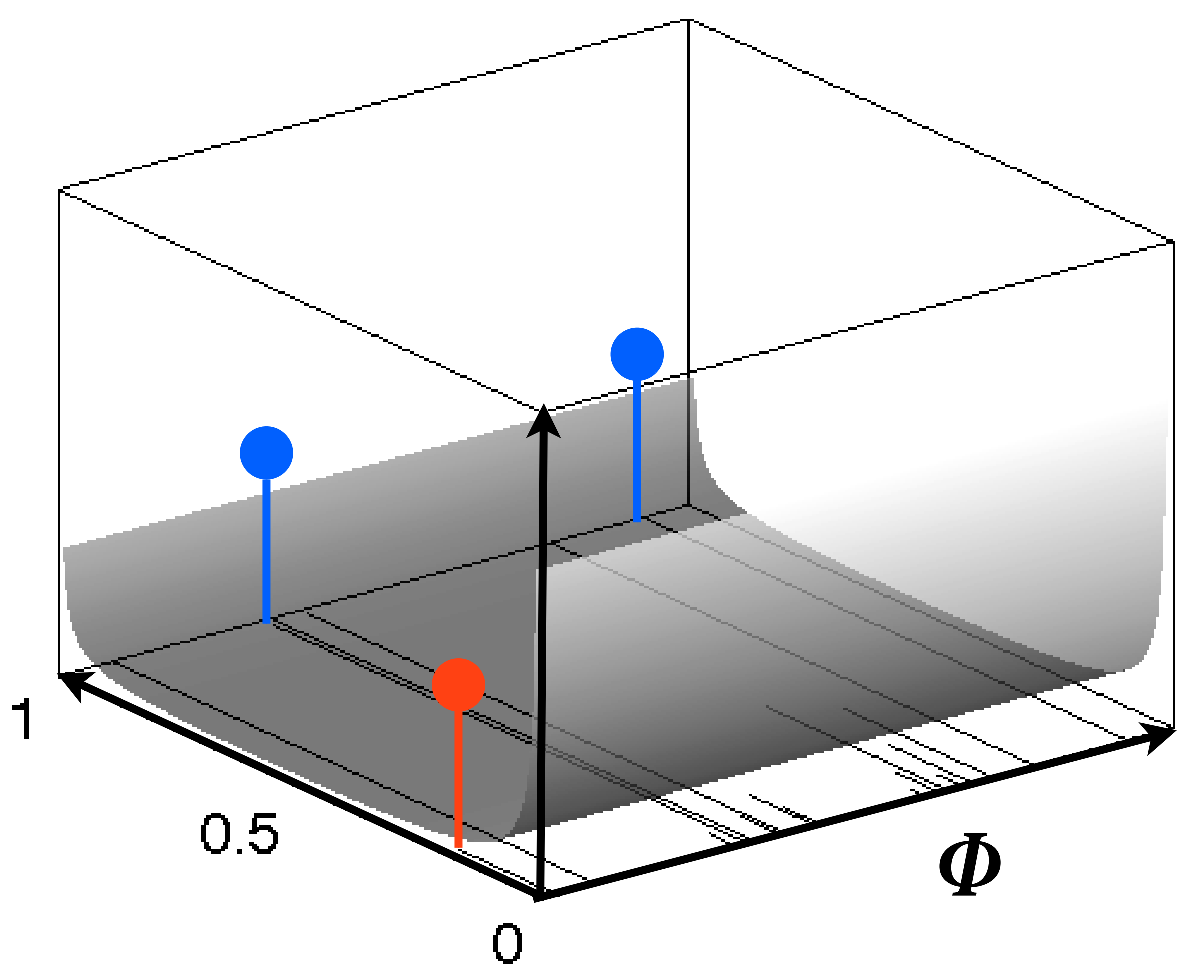
[Kingman 1993]

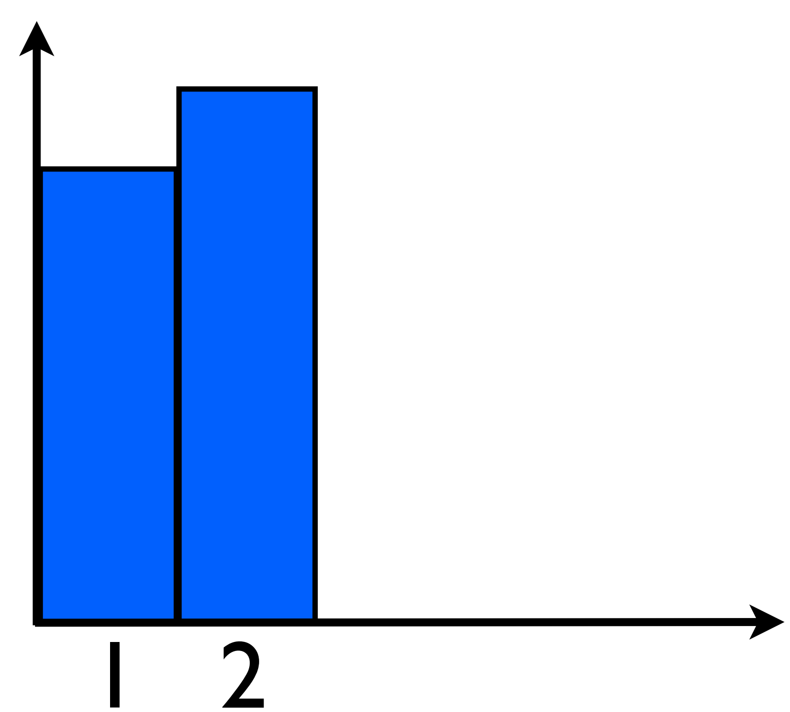
Aside: Poisson process

$$M \sim \mathrm{PP}(\nu)$$

$$M(A) \sim \mathrm{Poisson}(\nu(A))$$

$$M(A) \perp M(B)$$





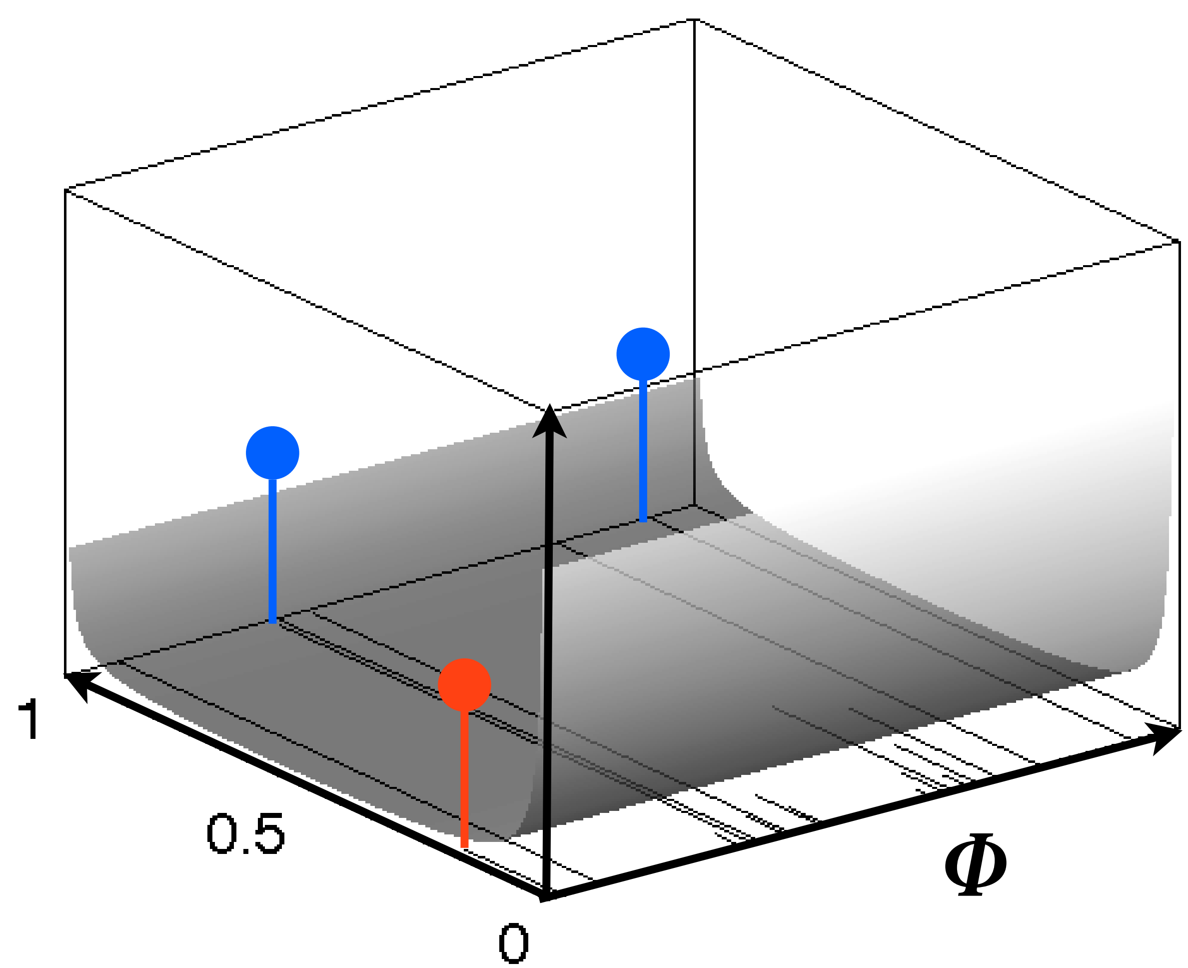
[Kingman 1993]

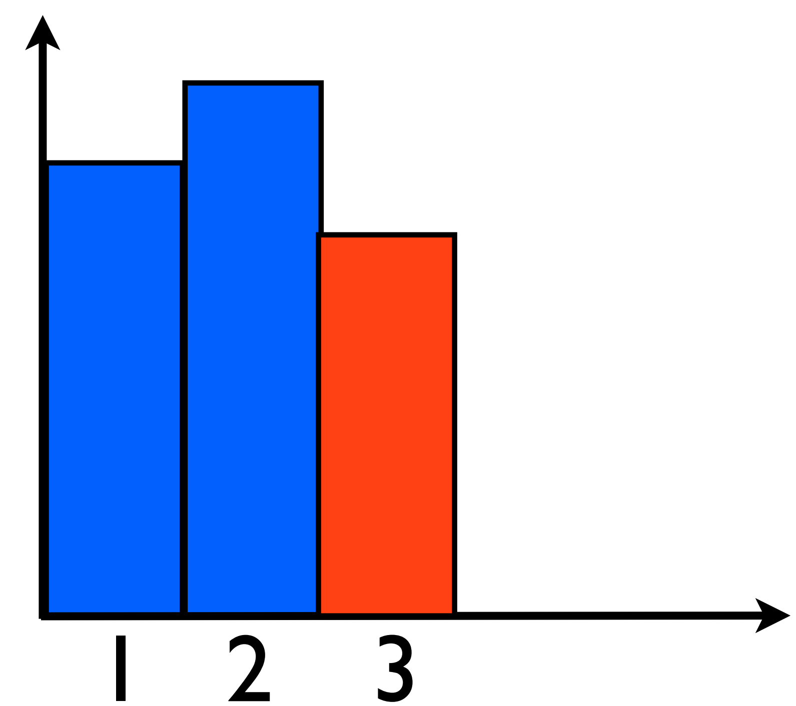
Aside: Poisson process

$$M \sim \mathrm{PP}(\nu)$$

$$M(A) \sim \mathrm{Poisson}(\nu(A))$$

$$M(A) \perp M(B)$$





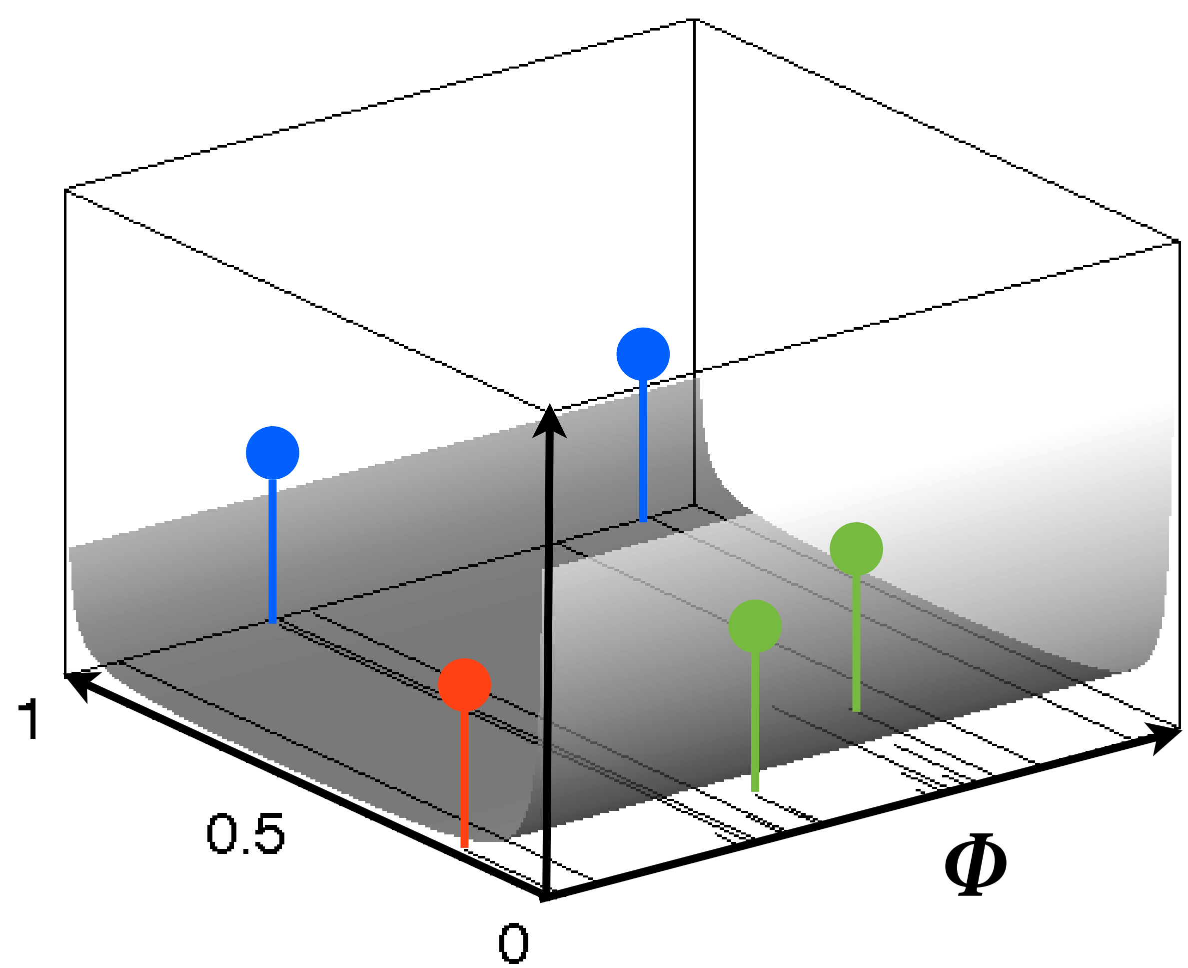
[Kingman 1993]

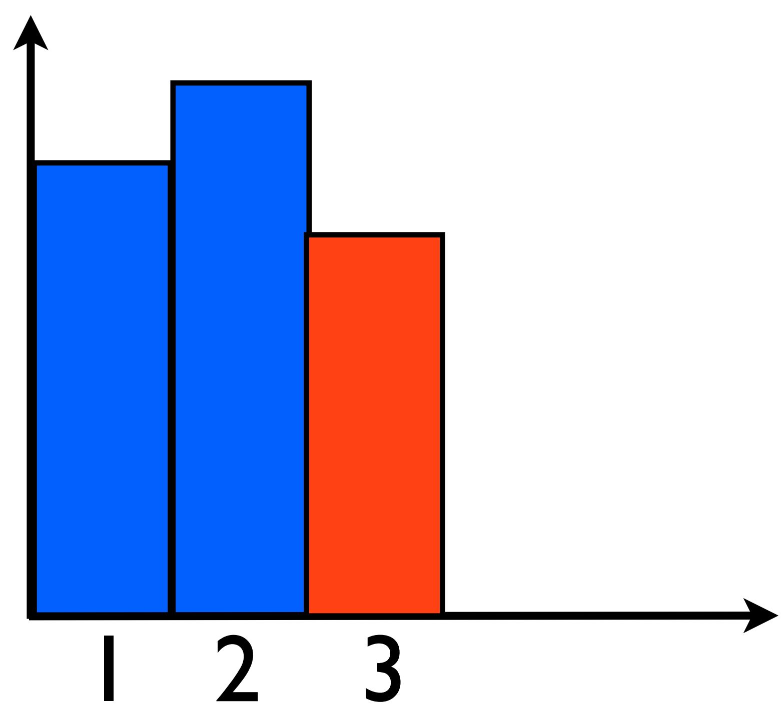
Aside: Poisson process

$$M \sim \mathrm{PP}(\nu)$$

$$M(A) \sim \mathrm{Poisson}(\nu(A))$$

$$M(A) \perp M(B)$$





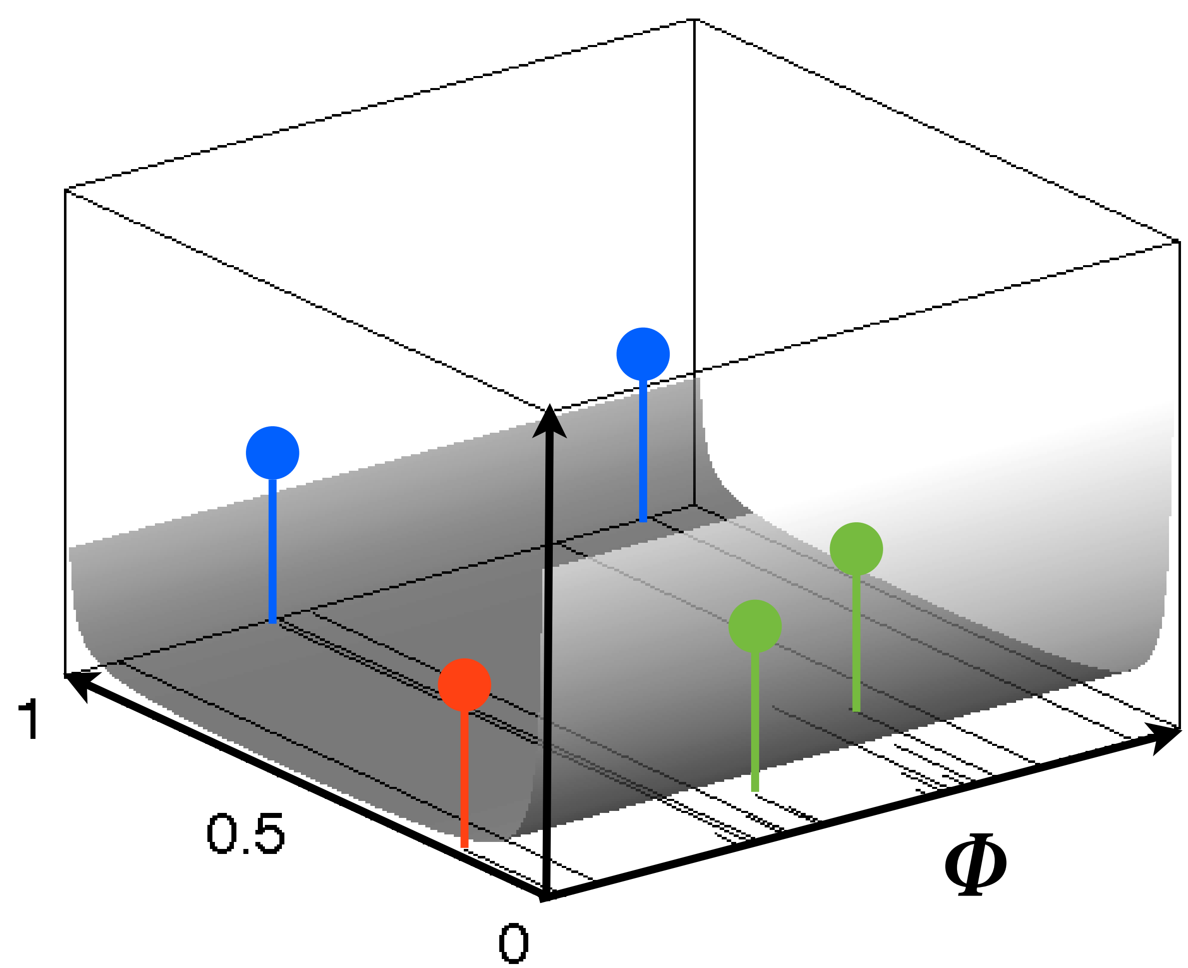
[Kingman 1993]

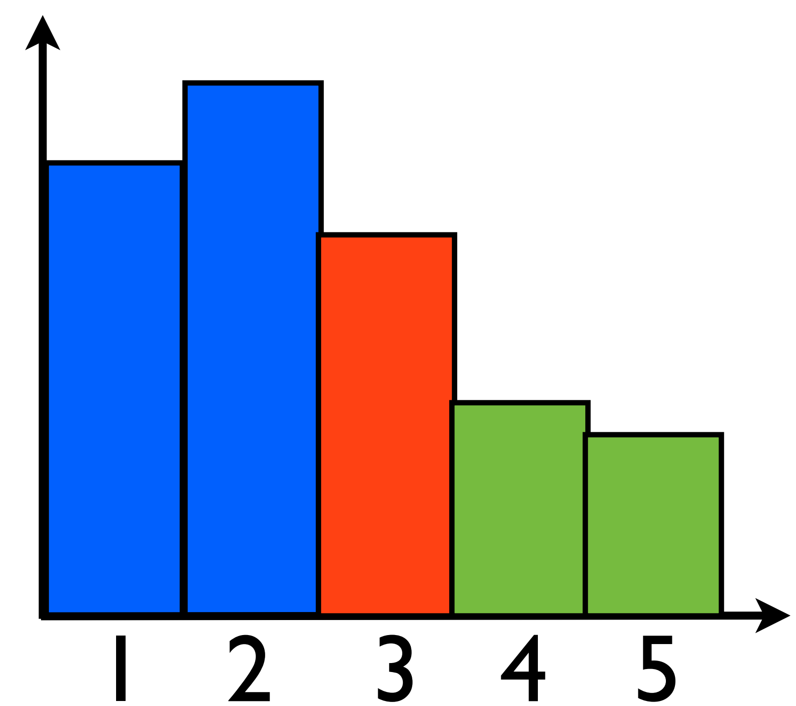
Aside: Poisson process

$$M \sim \mathrm{PP}(\nu)$$

$$M(A) \sim \mathrm{Poisson}(\nu(A))$$

$$M(A) \perp M(B)$$





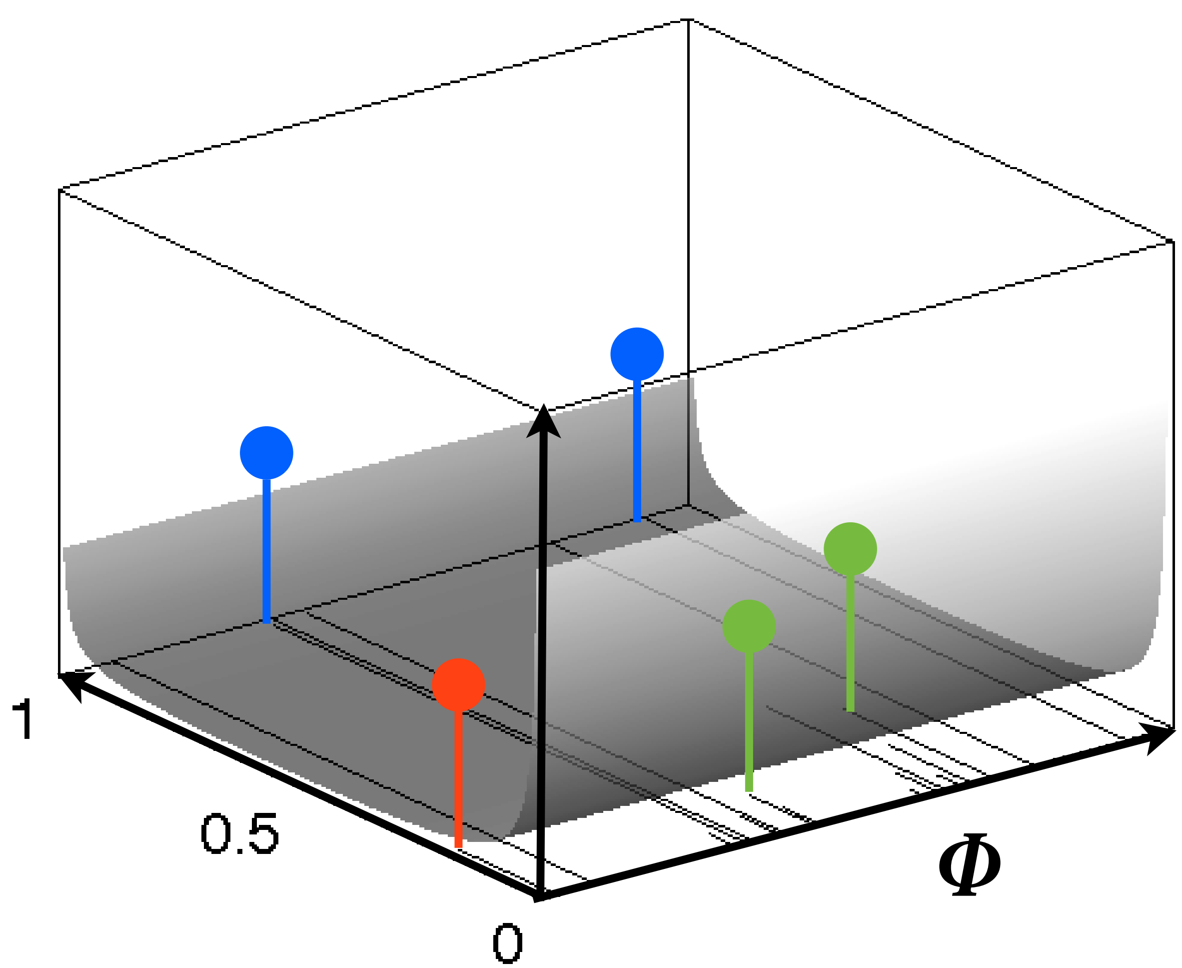
[Kingman 1993]

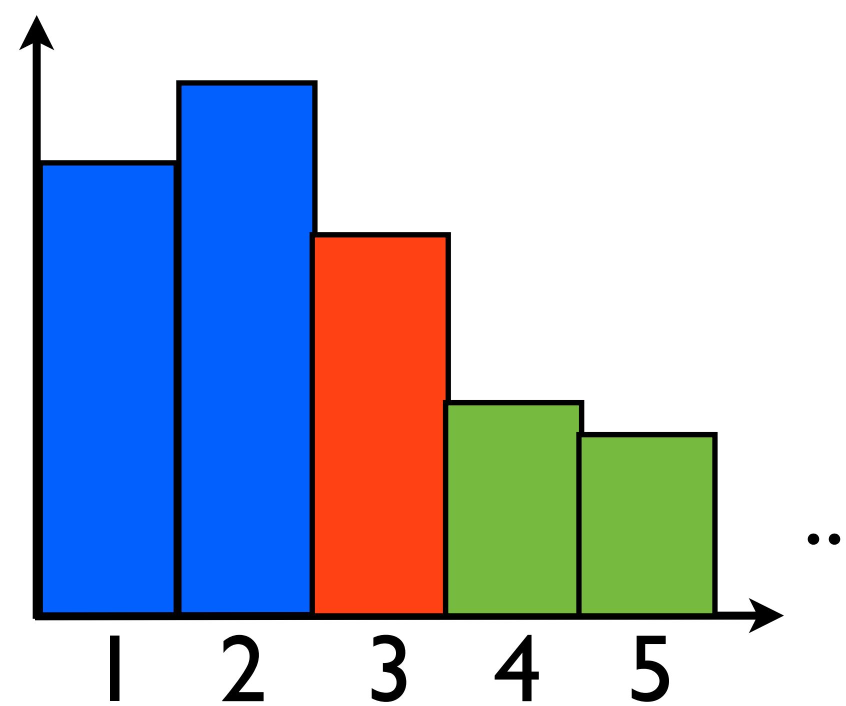
Aside: Poisson process

$$M \sim \mathrm{PP}(\nu)$$

$$M(A) \sim \mathrm{Poisson}(\nu(A))$$

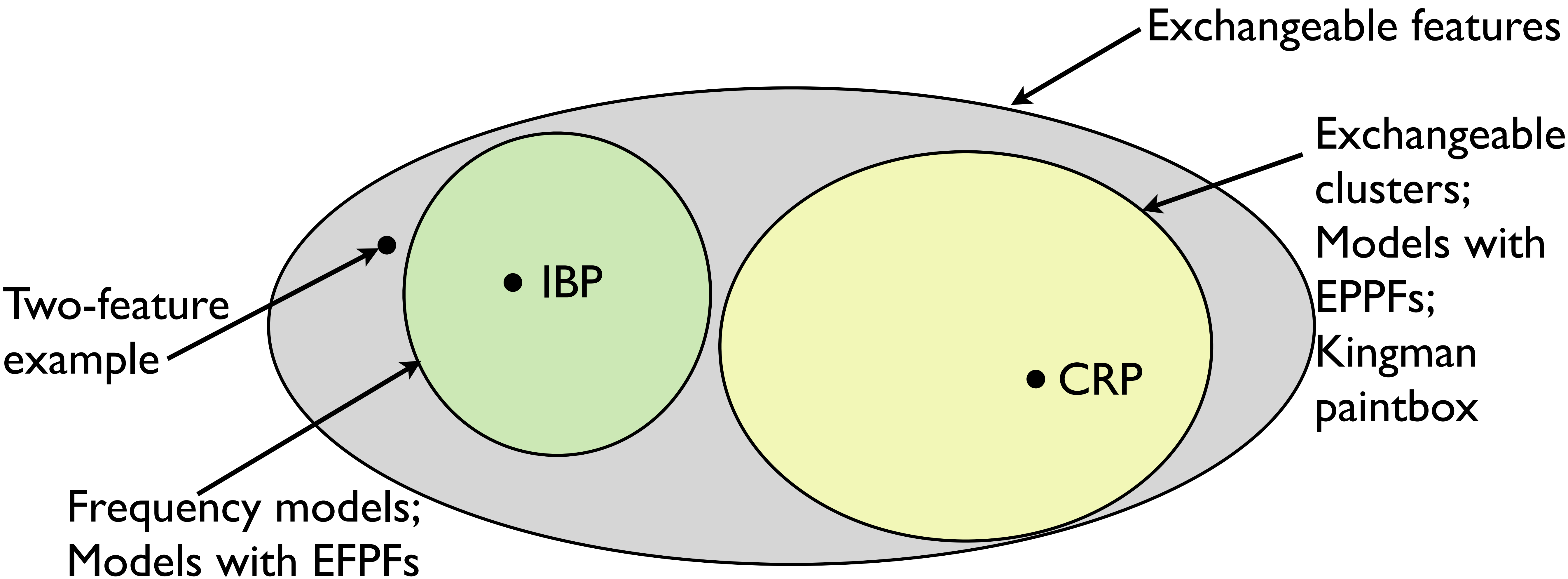
$$M(A) \perp M(B)$$



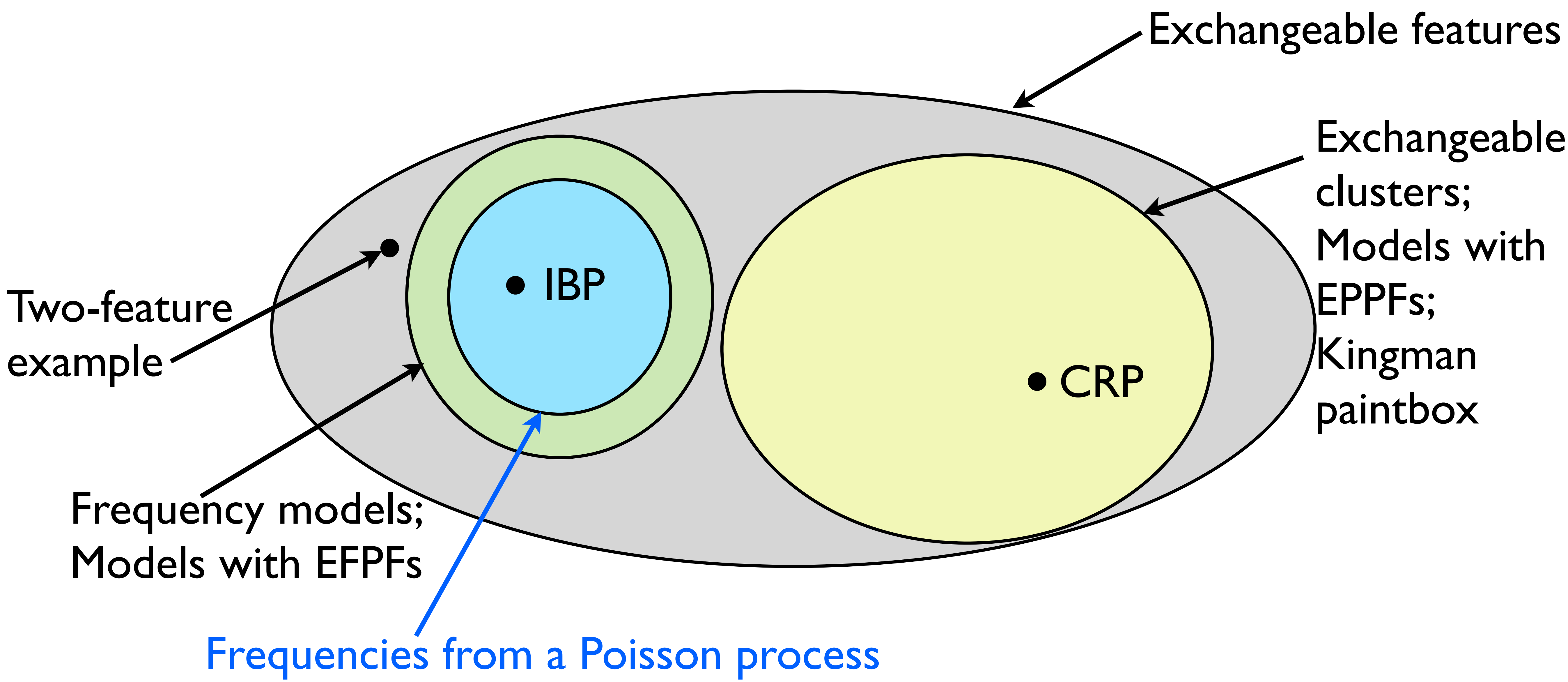


[Kingman 1993]

Clusters and features



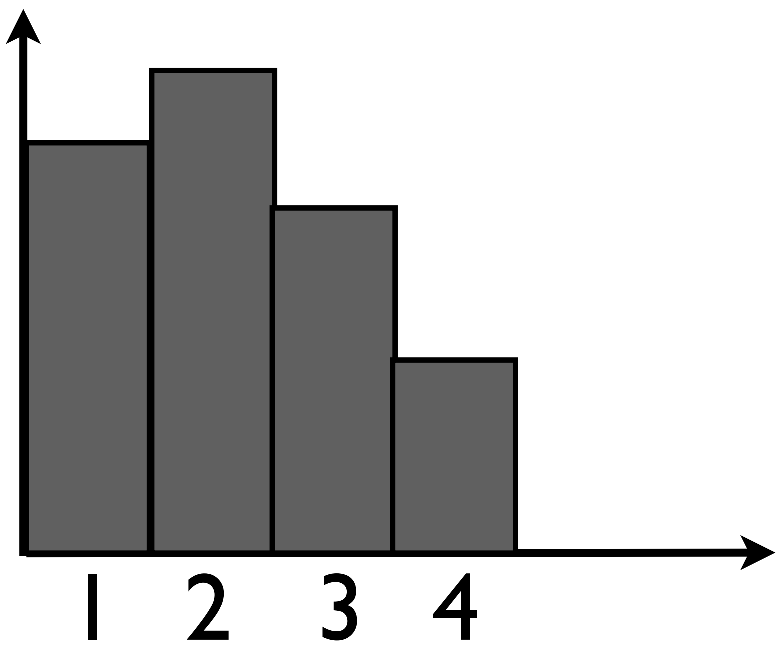
Clusters and features

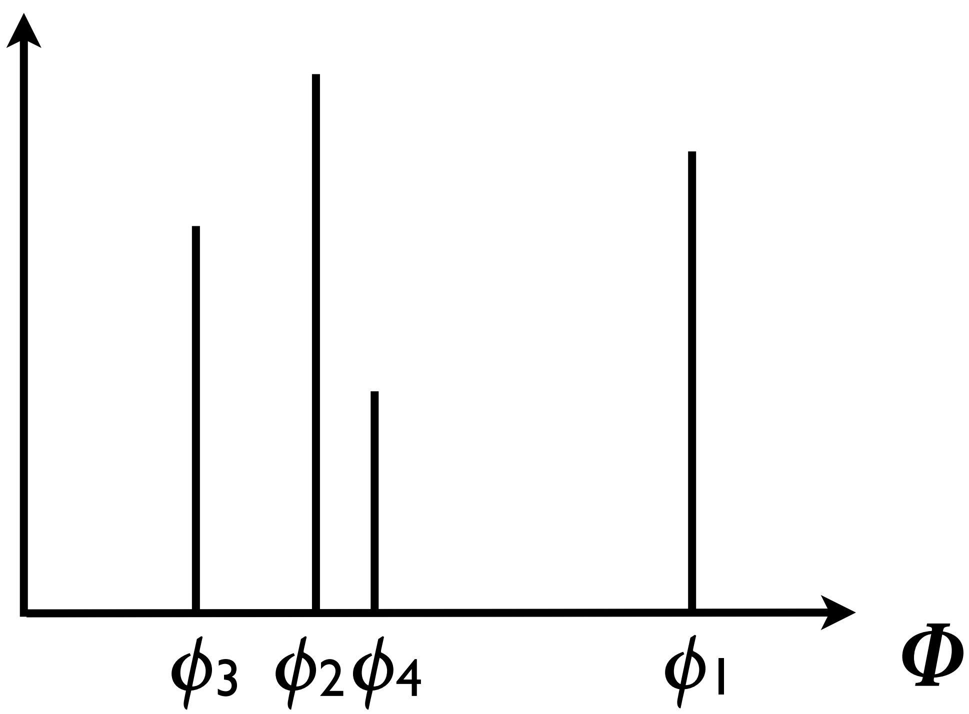


Why the IBP?

Finite, fixed number of features

$$f_9 = \{\{1,5,6\},\{8\},\{8\},\{8,5\},\{1,6\}\}$$





$$Z_{nk} \stackrel{iid}{\sim} \mathrm{Bern}(q_k) \quad k = 1, \ldots, K$$

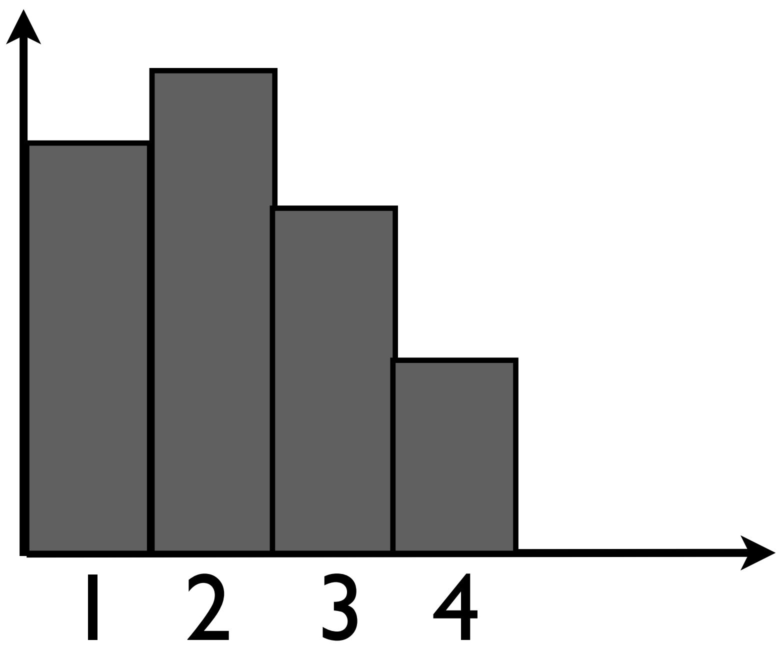
$$q_k \stackrel{iid}{\sim} \mathrm{Beta}(1, \theta) \quad k = 1, \ldots, K$$

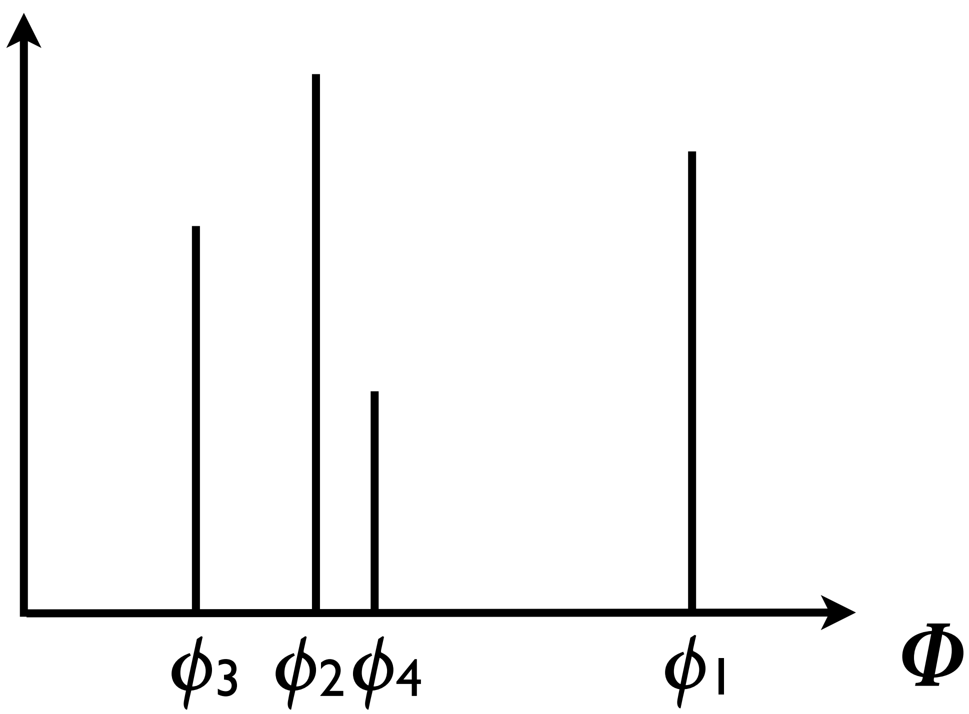
$$\phi_k \stackrel{iid}{\sim} H \quad k = 1, \ldots, K$$

Why the IBP?

Unbounded number of features

$$f_9 = \{\{1,5,6\},\{8\},\{8\},\\ \{8,5\},\{1,6\}\}$$





$$Z_{nk} \stackrel{iid}{\sim} \text{Bern}(q_k)$$
$$k = 1, \ldots, K$$

$$q_k \stackrel{iid}{\sim} \text{Beta}(1, \theta)$$
$$k = 1, \ldots, K$$

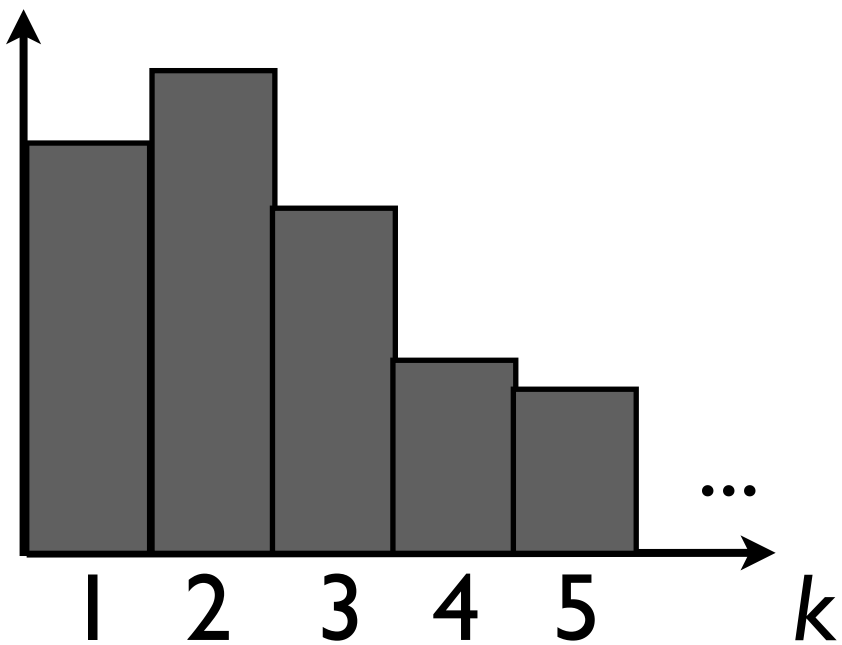
$$\phi_k \stackrel{iid}{\sim} H$$
$$k = 1, \ldots, K$$

Why the IBP?

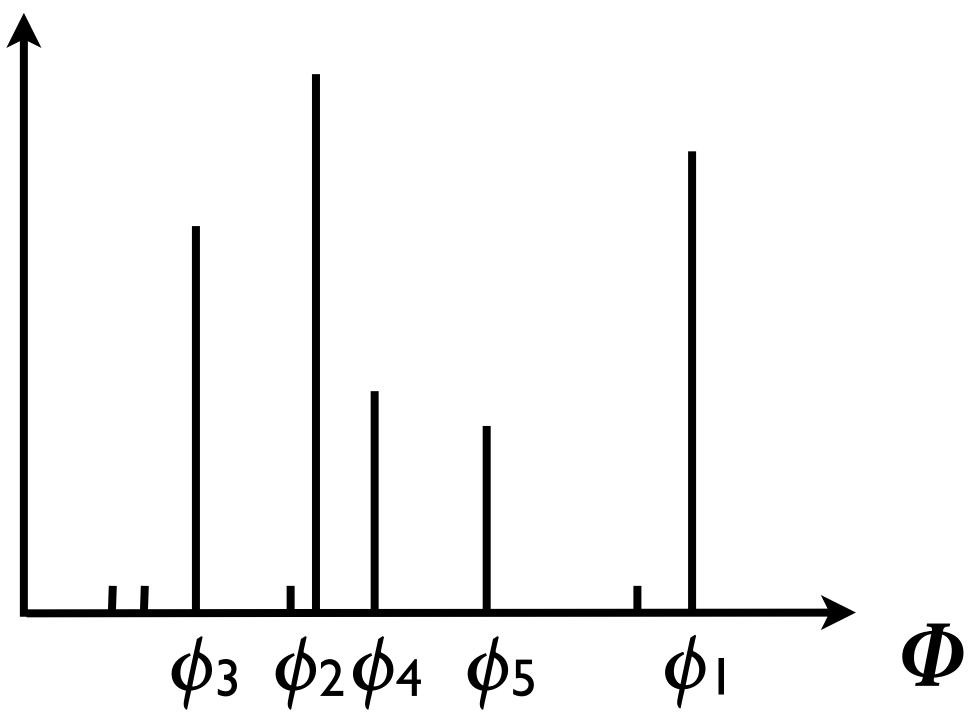
Unbounded number of clusters

$$f_9 = \{\{1,5,6\},\{8\},\{8\},\{8,5\},\{1,6\}\}$$

$$Z_{nk} \overset{iid}{\sim} \mathrm{Bern}(q_k)$$
$$k = 1, 2, \ldots$$

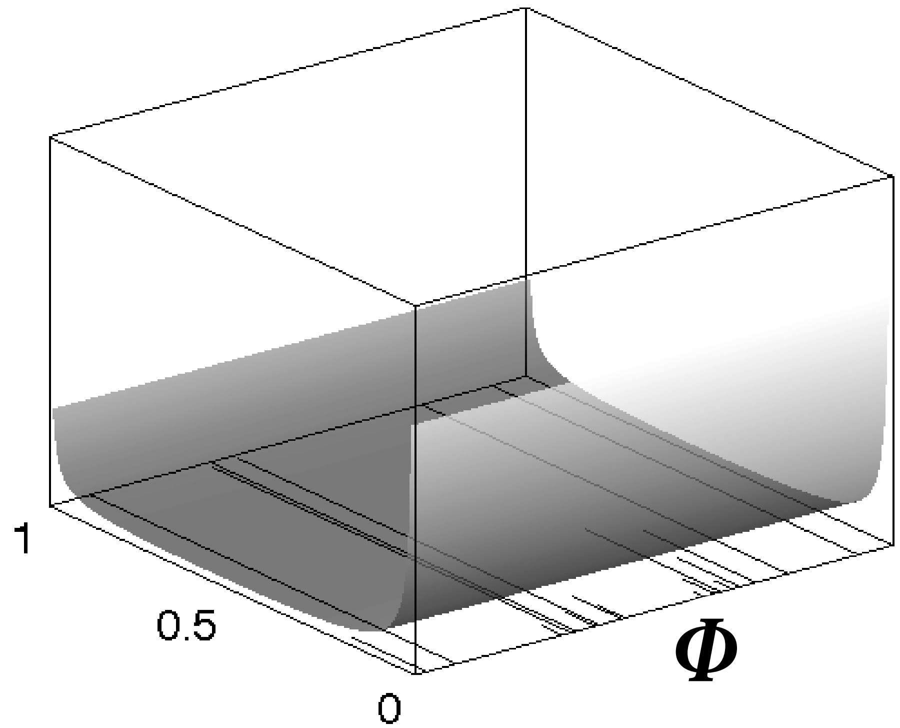


$(q_k)_{k=1}^{\infty}$ ~atom weights of beta process(γ, θ)



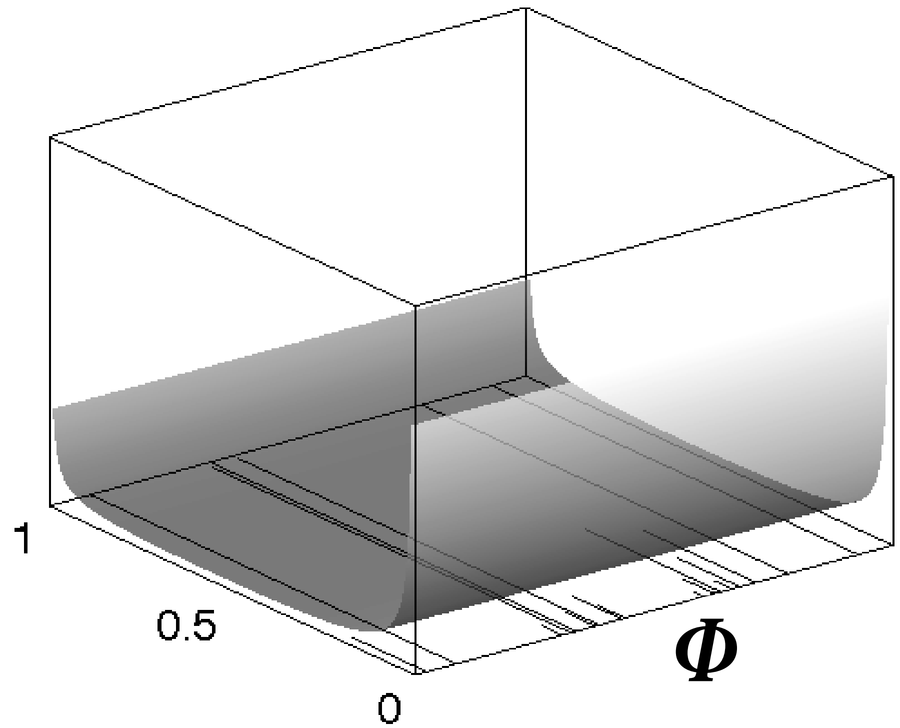
$$\phi_k \overset{iid}{\sim} H$$
$$k = 1, 2, \ldots$$

Poisson processes



• improper beta density for Poisson process

Poisson processes

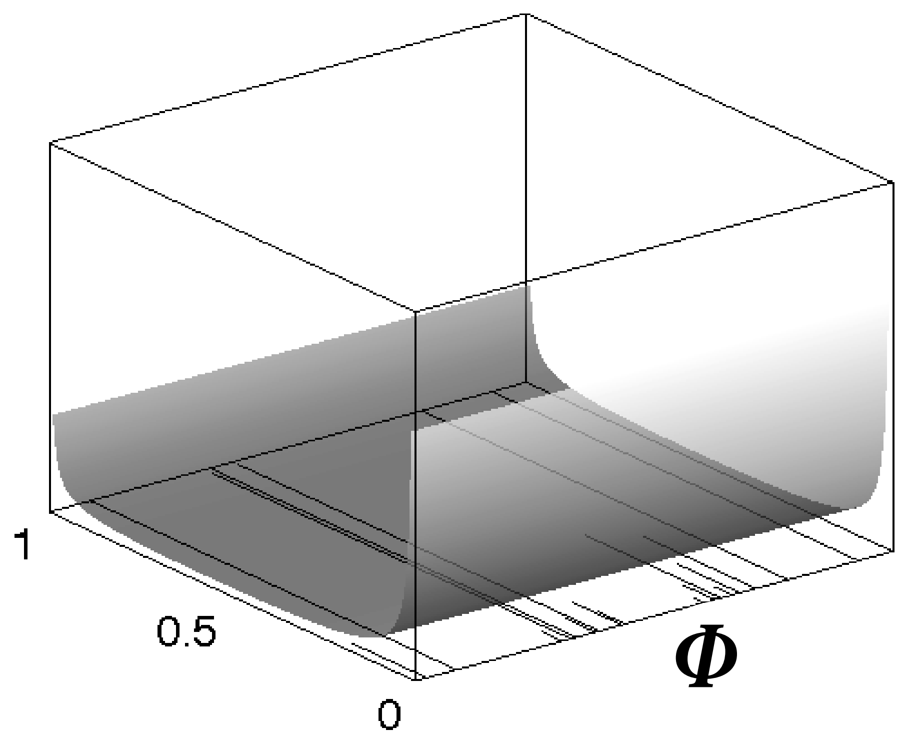


• improper beta density for Poisson process

↓

• beta process

Poisson processes



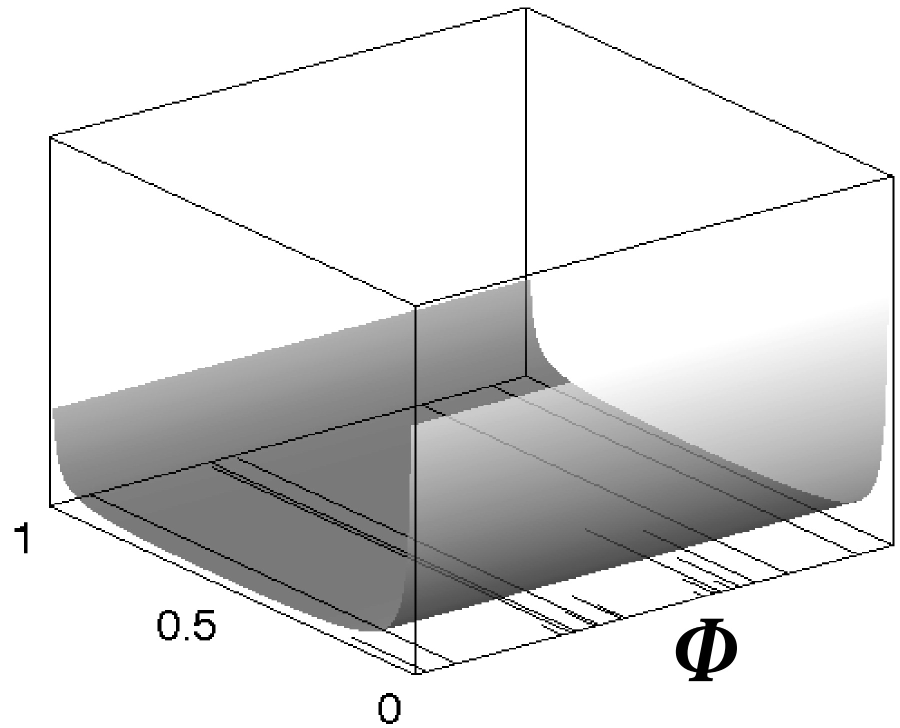
• improper beta density for Poisson process

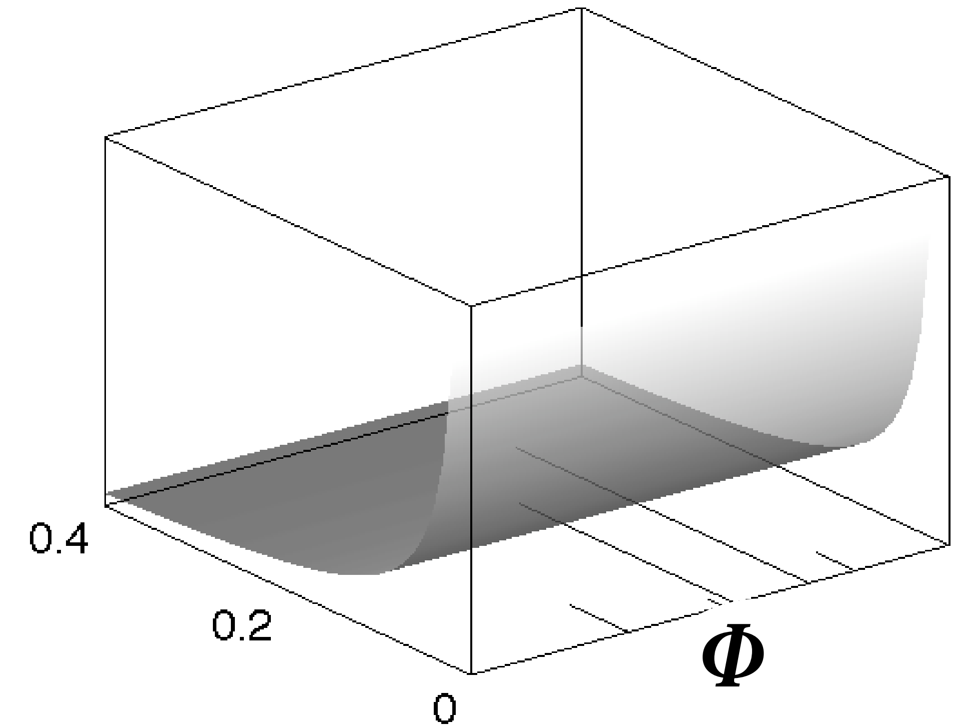
↓

• beta process

• ~ independent beta random variables

Poisson processes





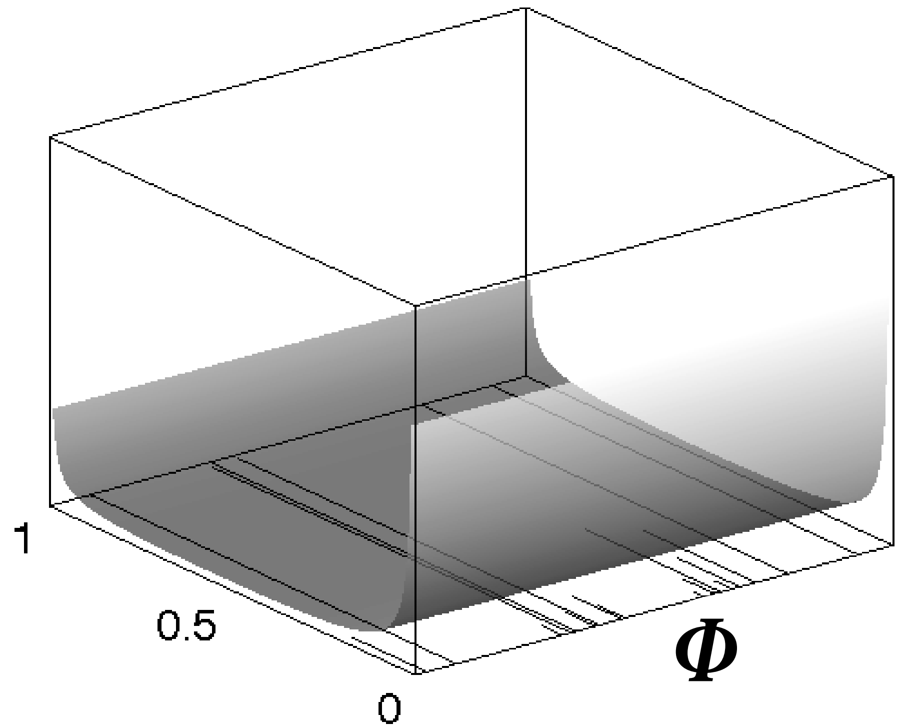
- improper beta density for Poisson process

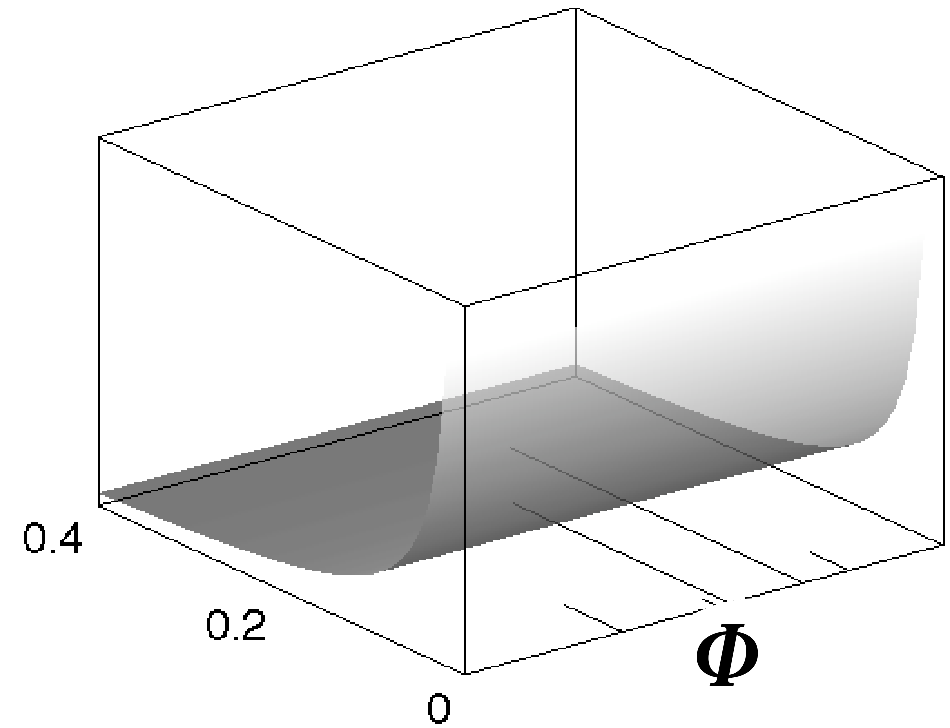
↓

- beta process

- ~ independent beta random variables

Poisson processes





- improper beta density for Poisson process

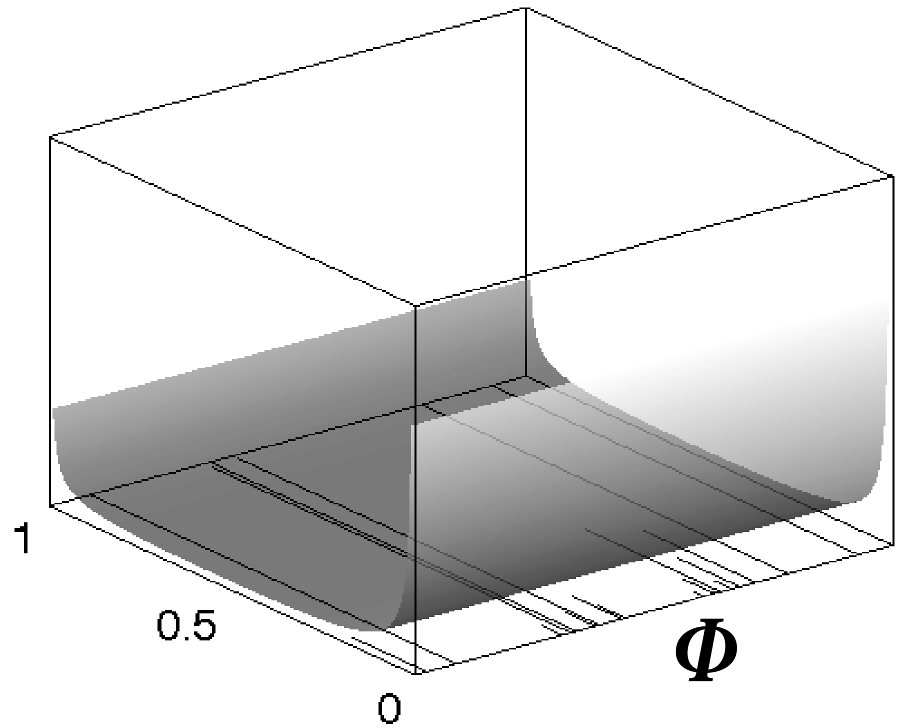
↓

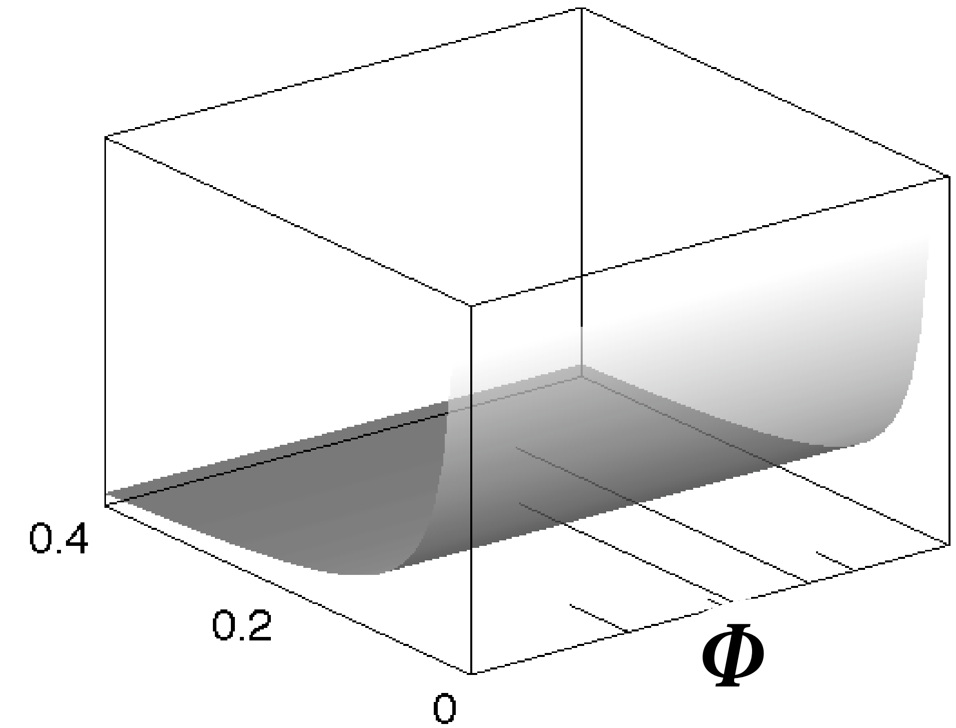
- beta process

- ~ independent beta random variables

- improper gamma density for Poisson process

Poisson processes





• improper beta density for Poisson process

↓

• beta process

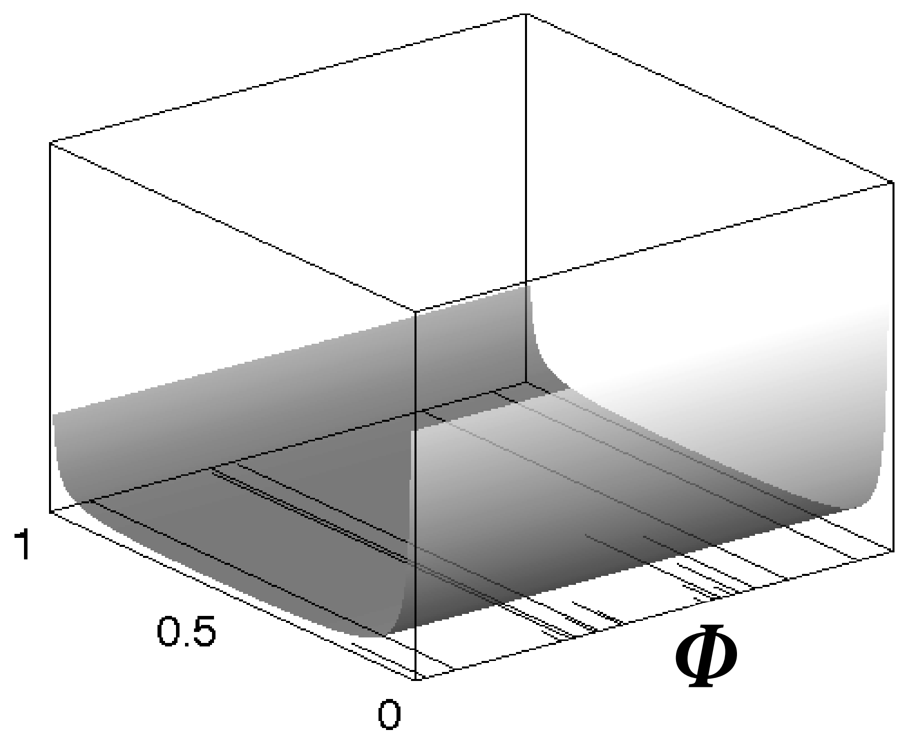
• ~ independent beta random variables

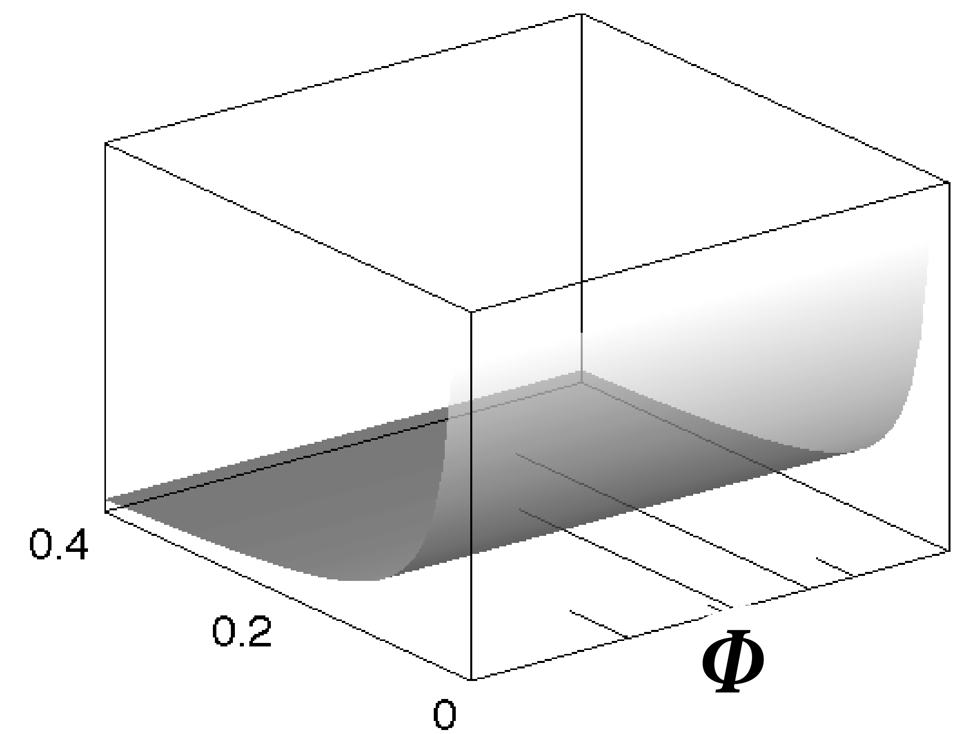
• improper gamma density for Poisson process

↓

• "gamma process"

Poisson processes





- improper beta density for Poisson process

↓

- beta process

- ~ independent beta random variables

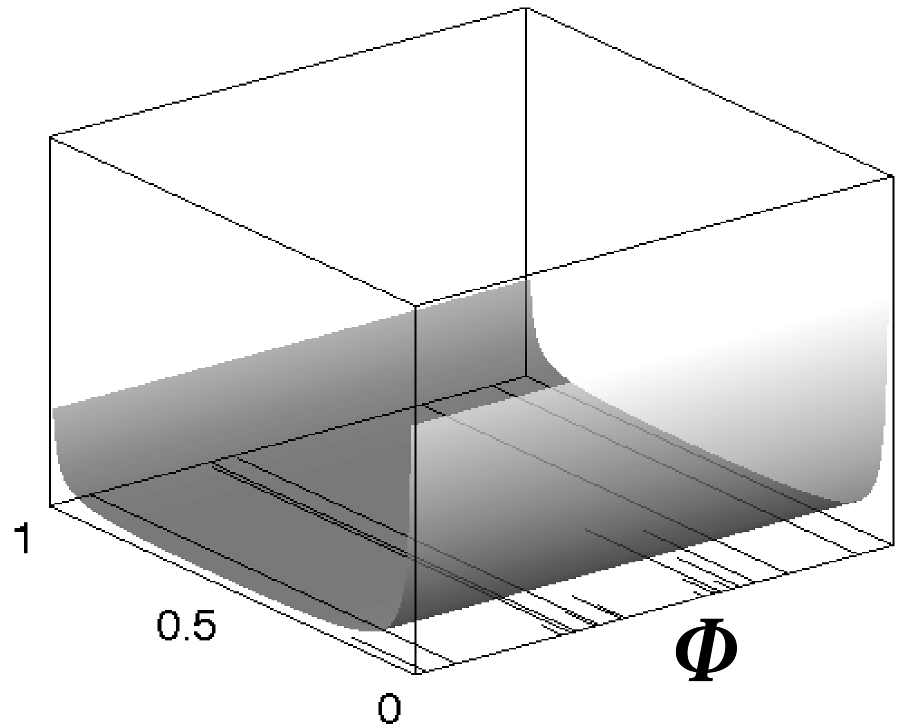
- improper gamma density for Poisson process

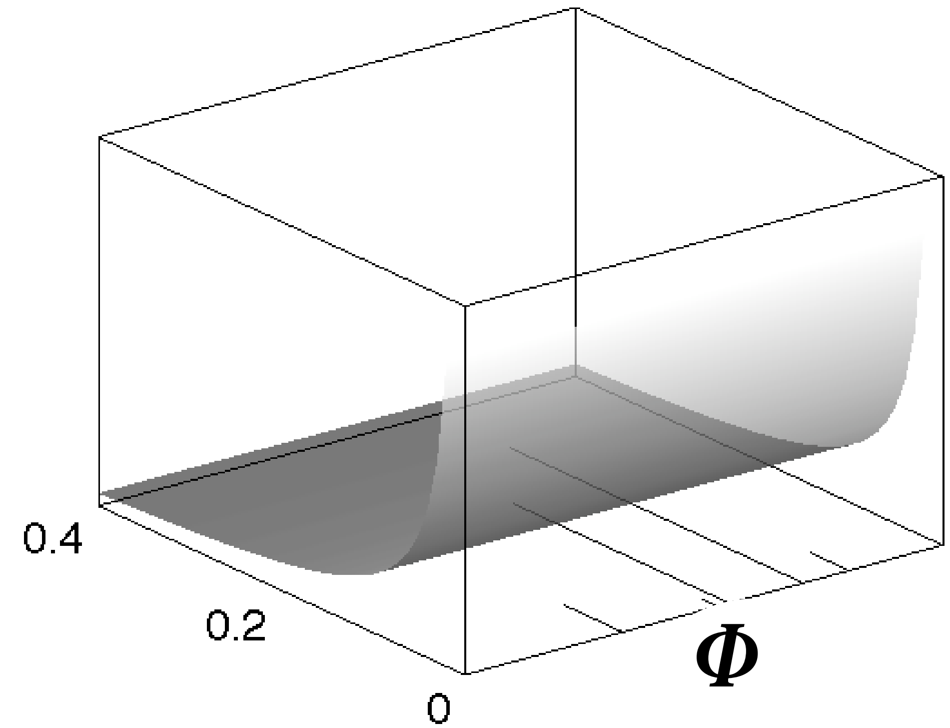
↓

- "gamma process"

- ~ independent gamma random variables

Poisson processes





- improper beta density for Poisson process

↓

- beta process

- ~ independent beta random variables

- improper gamma density for Poisson process

↓

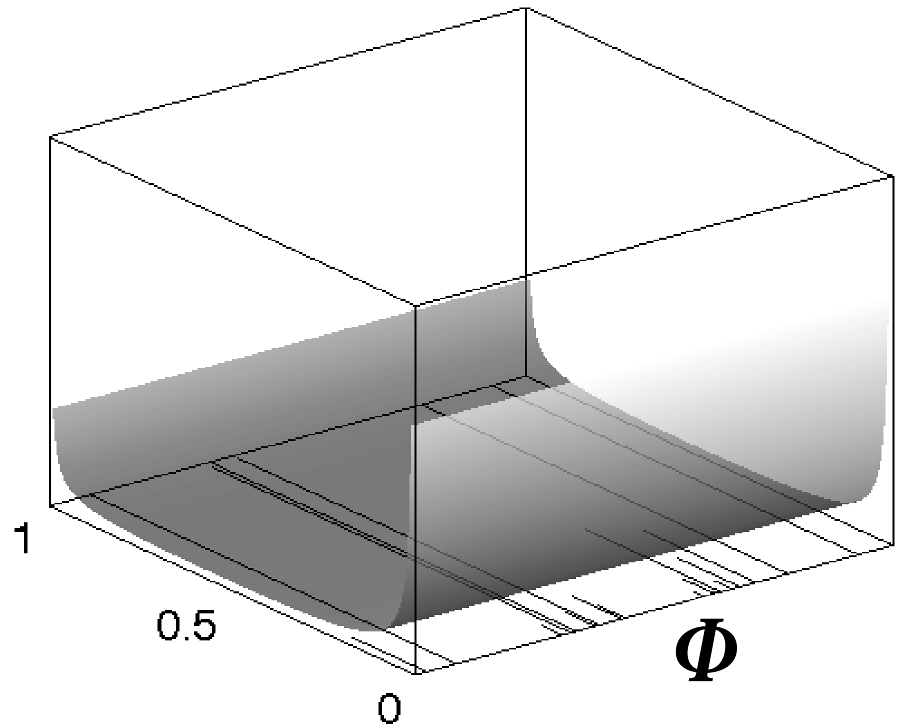
- "gamma process"

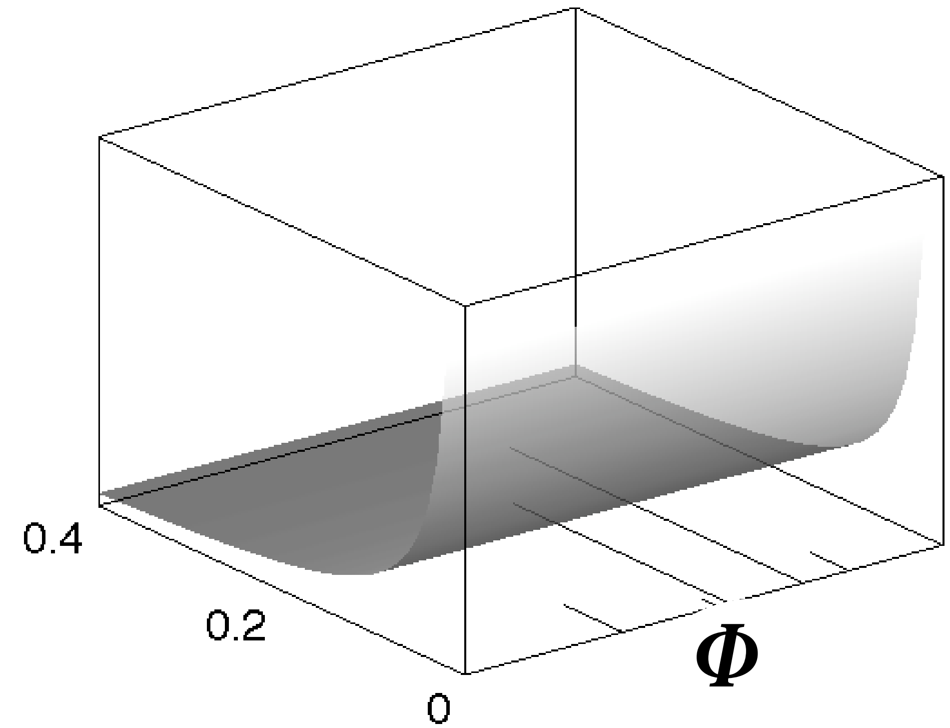
- ~ independent gamma random variables

↓

- normalize: Dirichlet random draw

Poisson processes





- improper beta density for Poisson process

↓

- beta process

- ~ independent beta random variables

- improper gamma density for Poisson process

↓

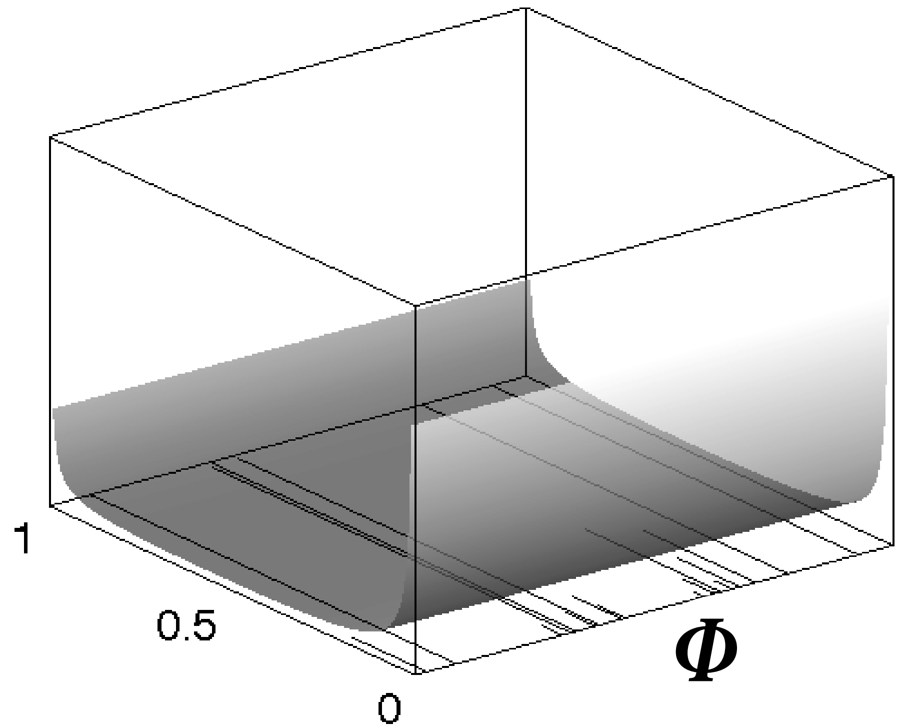
- "gamma process"

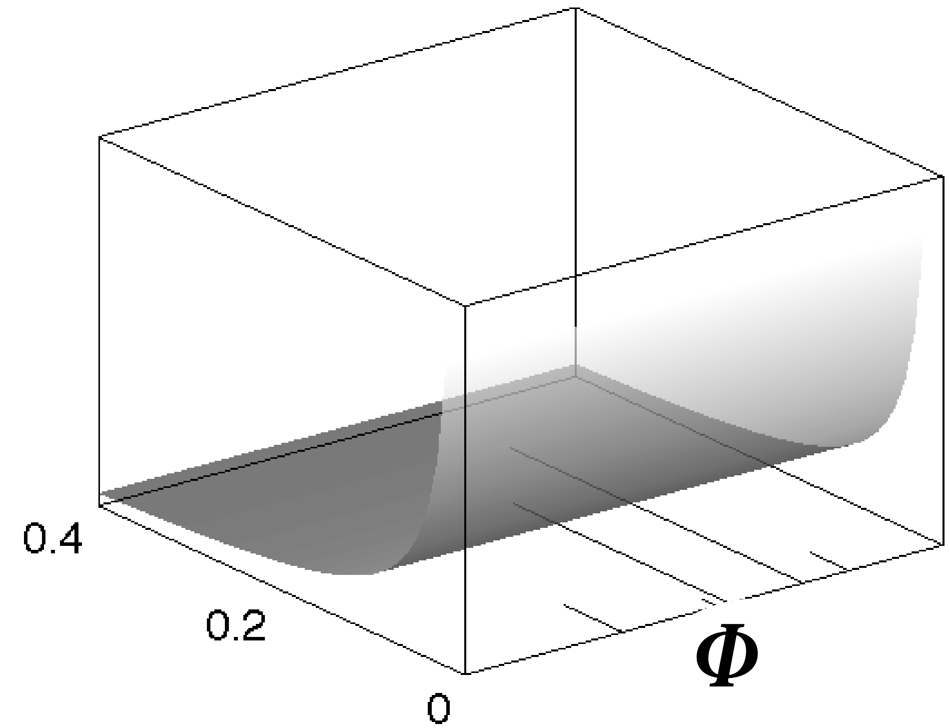
- ~ independent gamma random variables

↓

- normalize: Dirichlet random draw

Poisson processes





- improper beta density for Poisson process

↓

- beta p

- normalize: Dirichlet process

- ~ independent beta random variables

- improper gamma density for Poisson process

↓

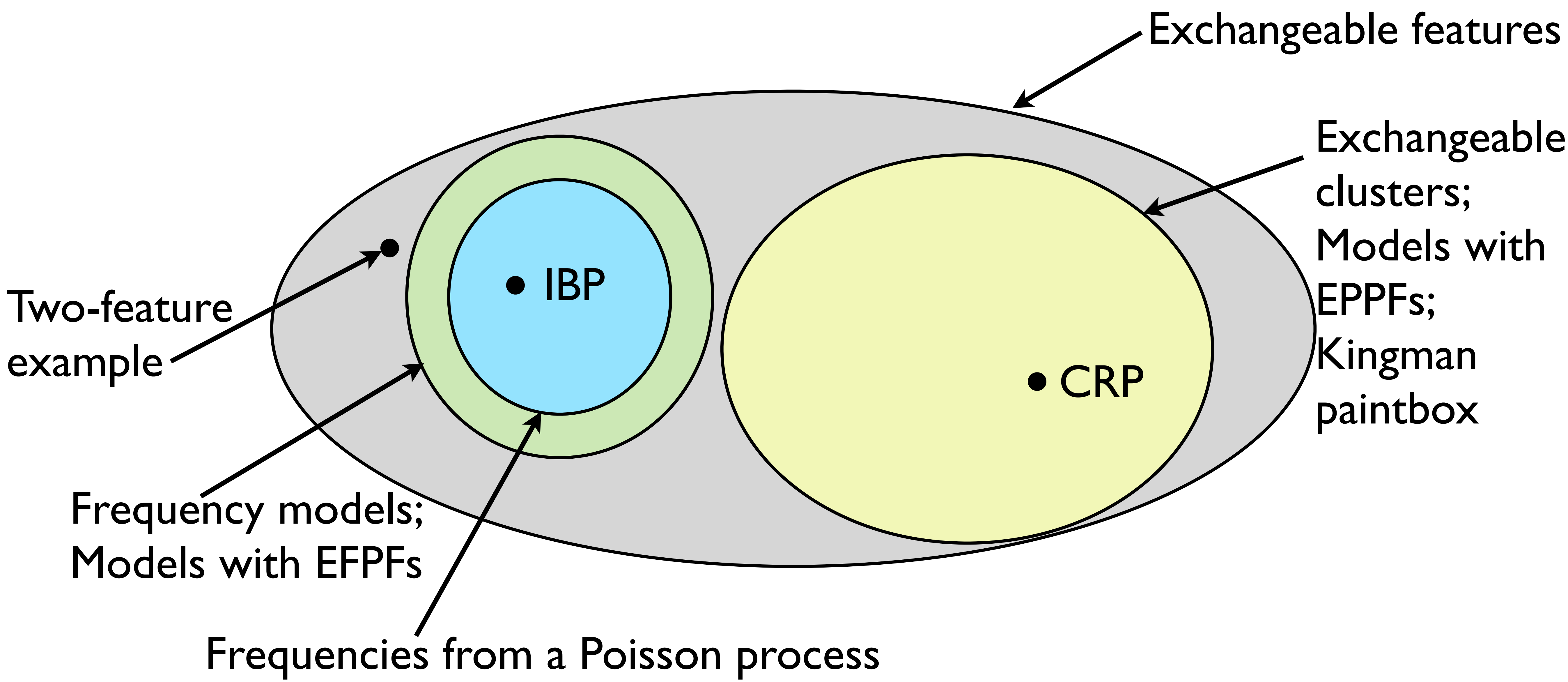
- “gamma process”

- ~ independent gamma random variables

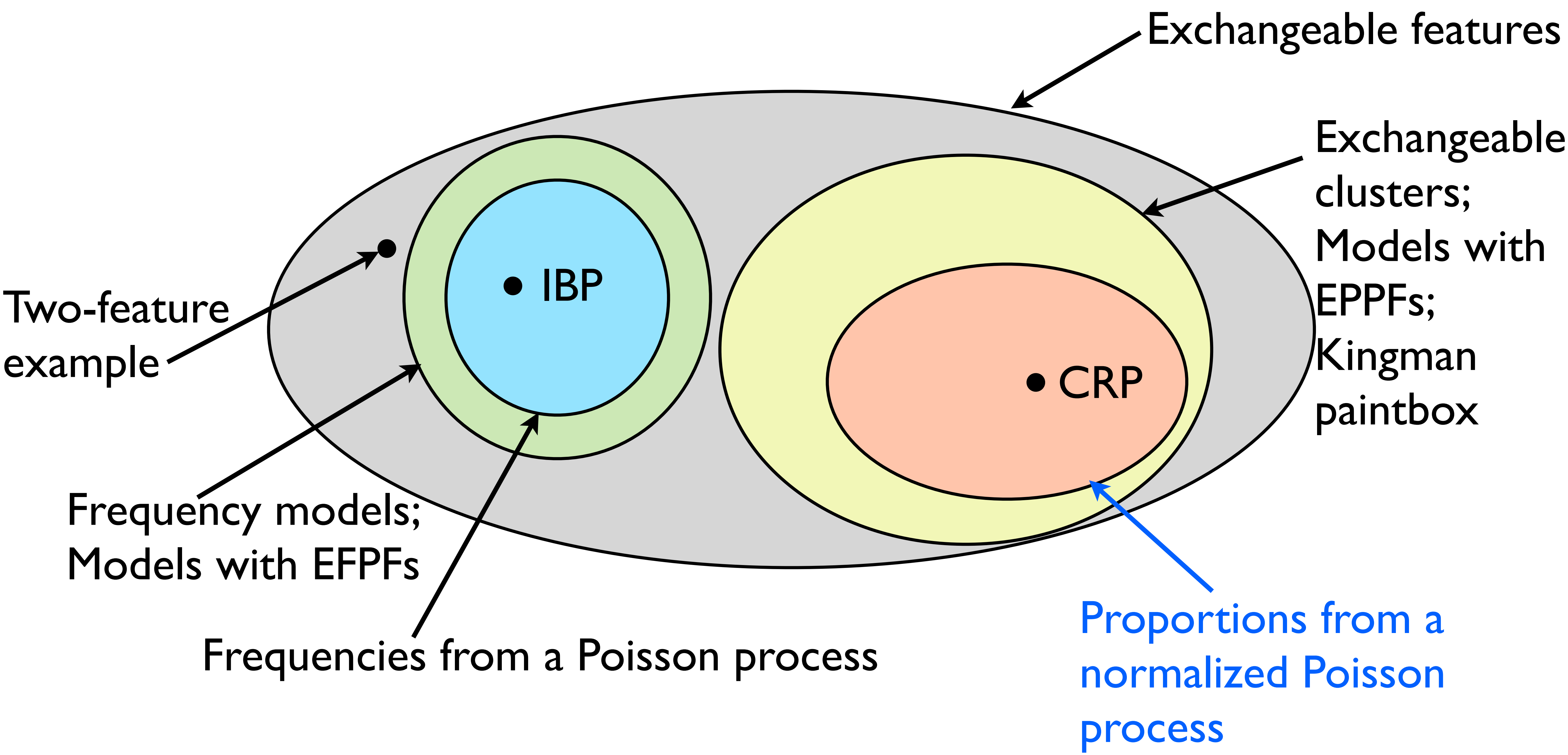
↓

- normalize: Dirichlet random draw

Clusters and features



Clusters and features



Outline

I. Clusters

II. Features

- Overview
- Distribution
- Frequencies
- Random measure
- Paintbox

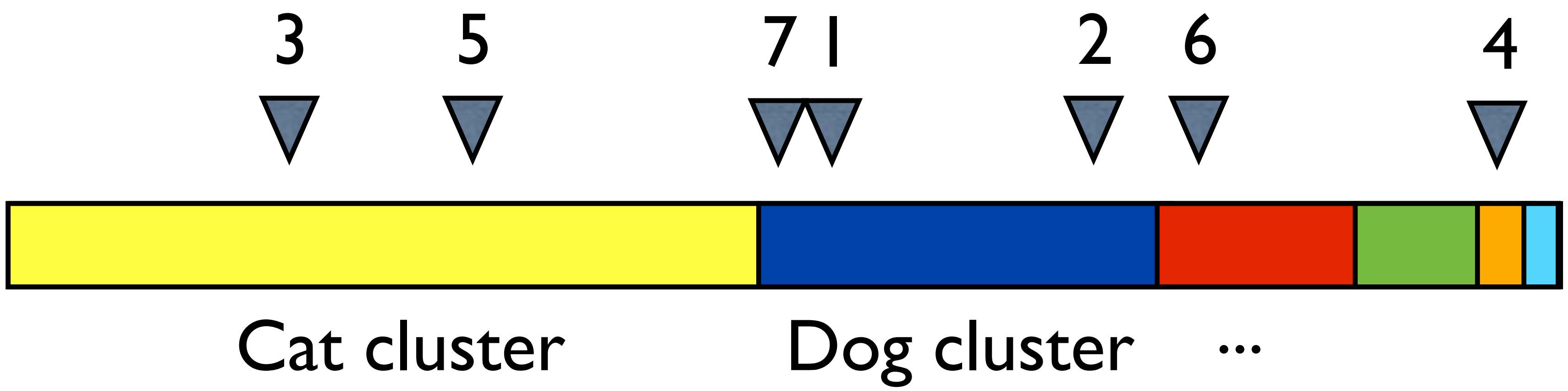
Outline

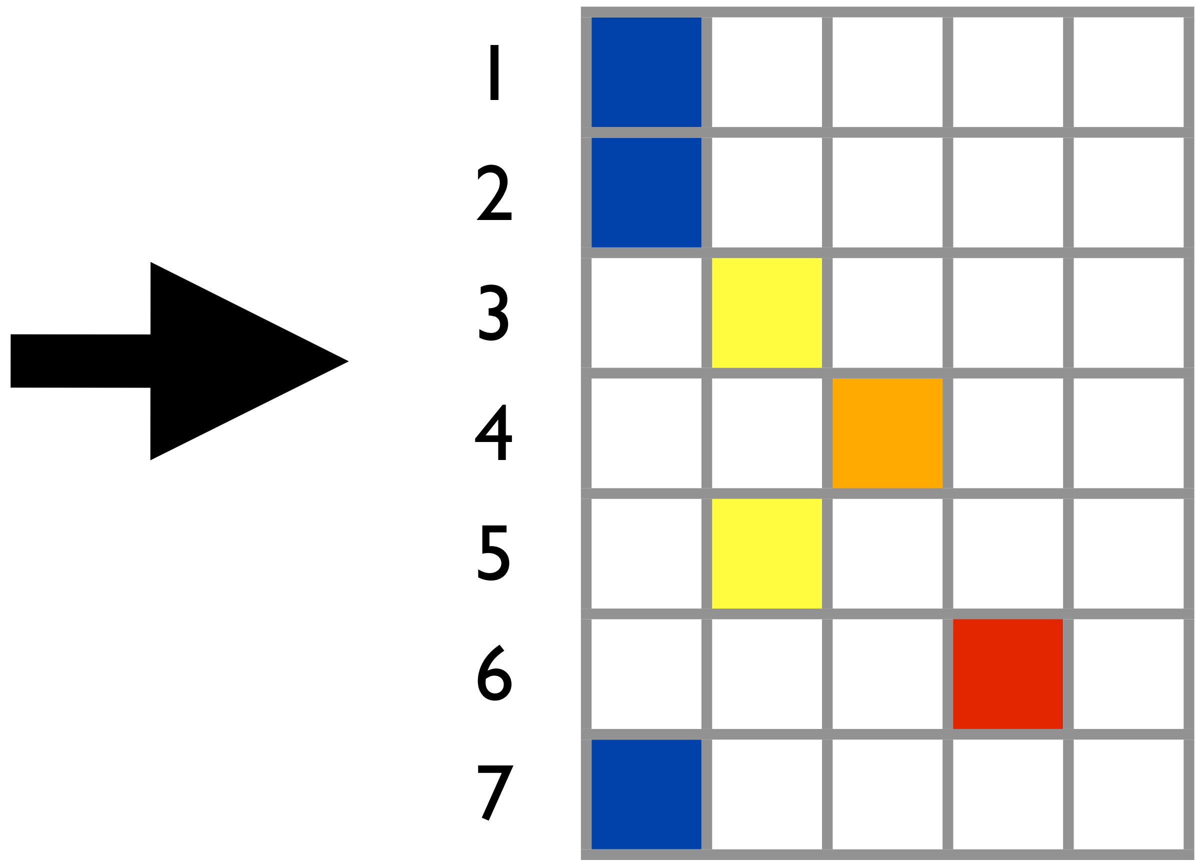
I. Clusters

II. Features

- Overview
- Distribution
- Frequencies
- Random measure
- Paintbox

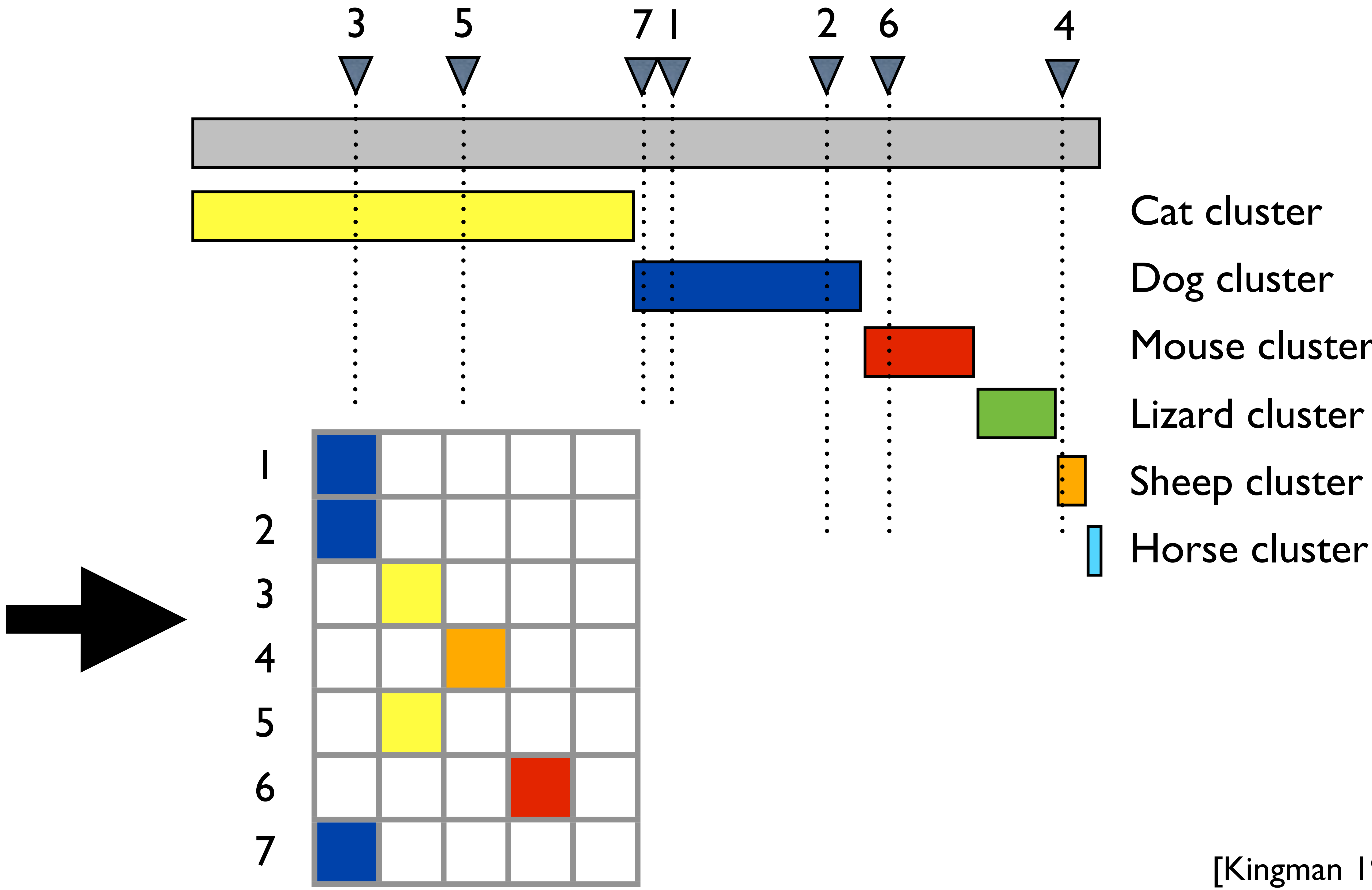
Kingman paintbox





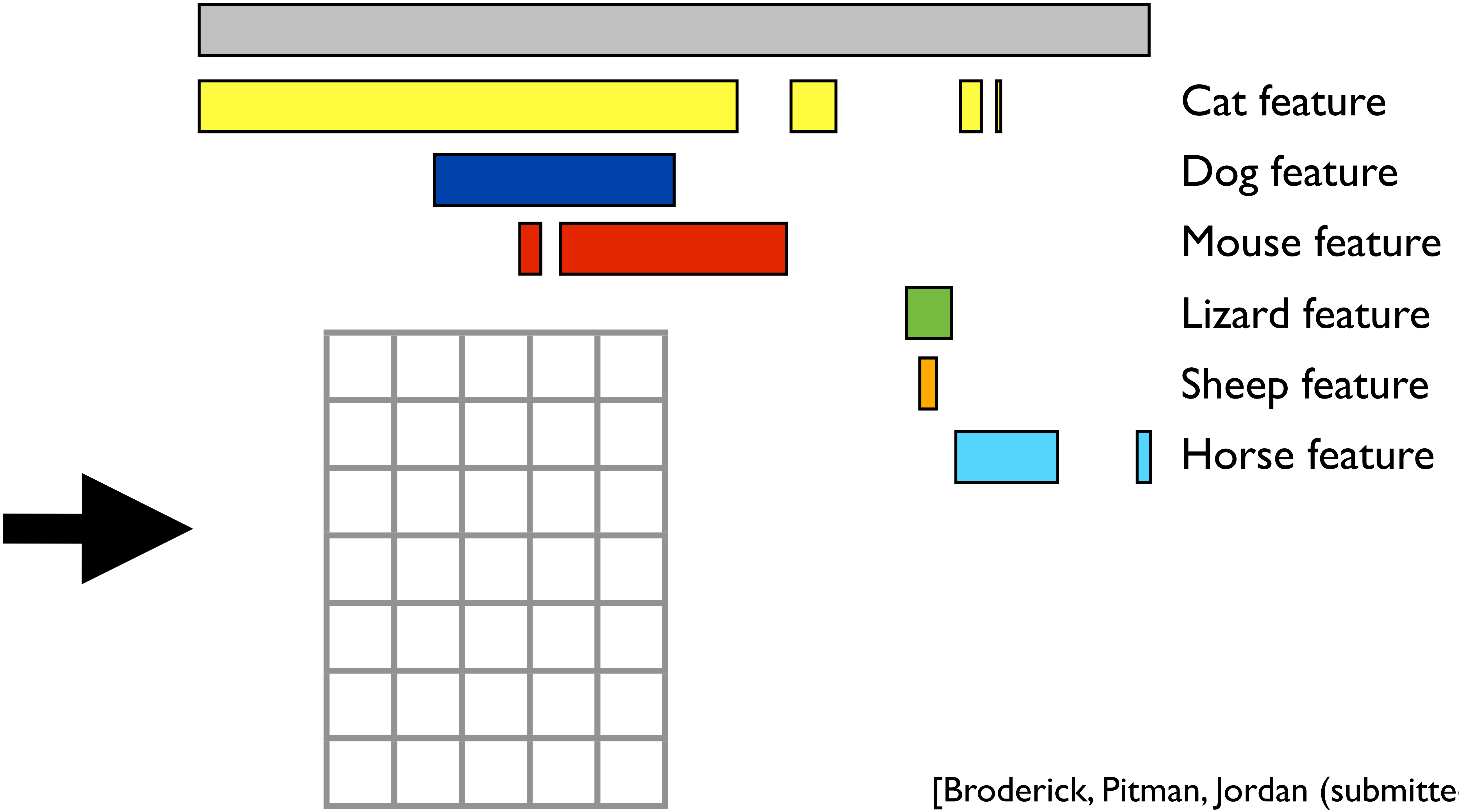
[Kingman 1978]

Kingman paintbox



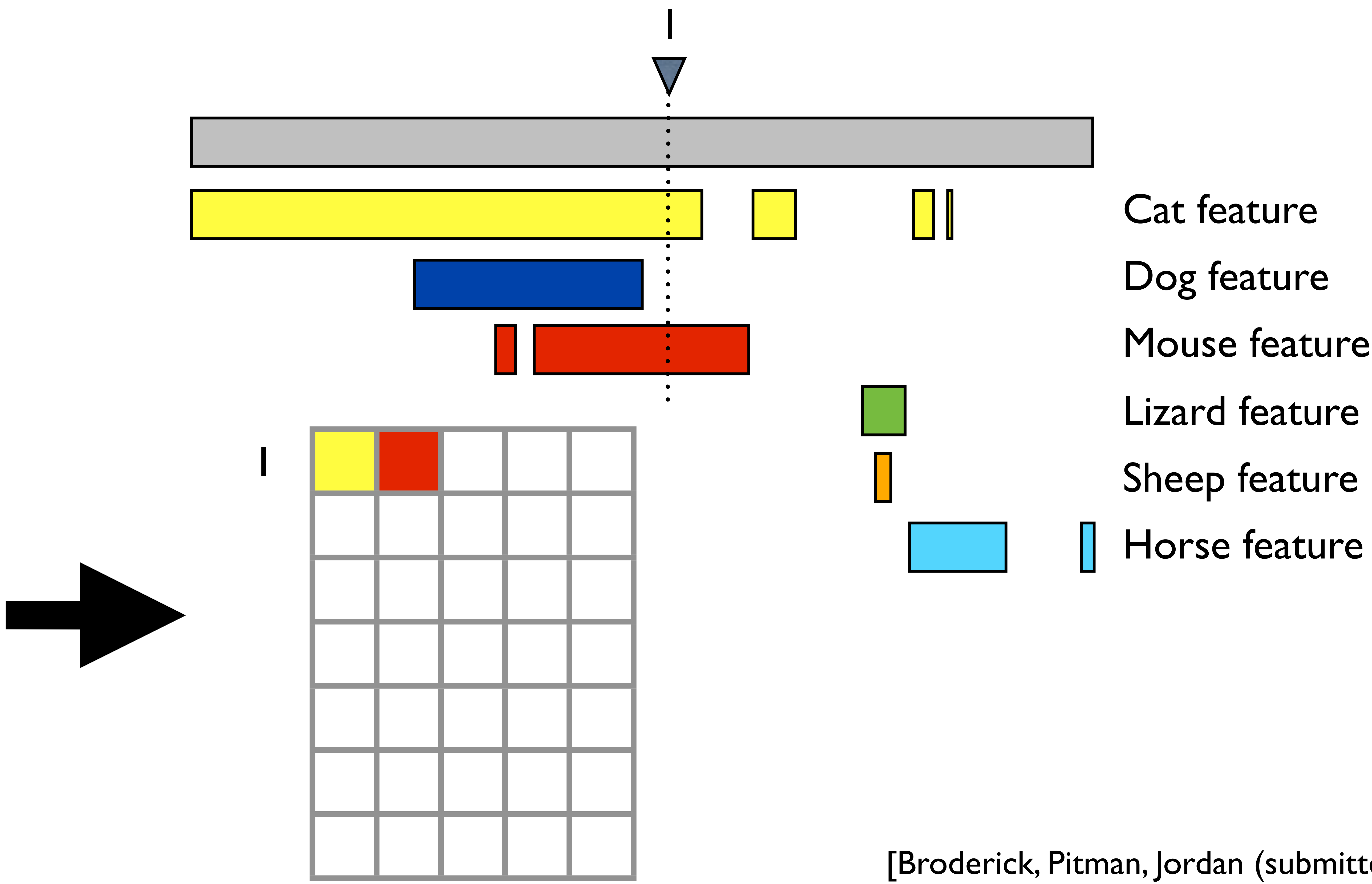
[Kingman 1978]

Feature paintbox



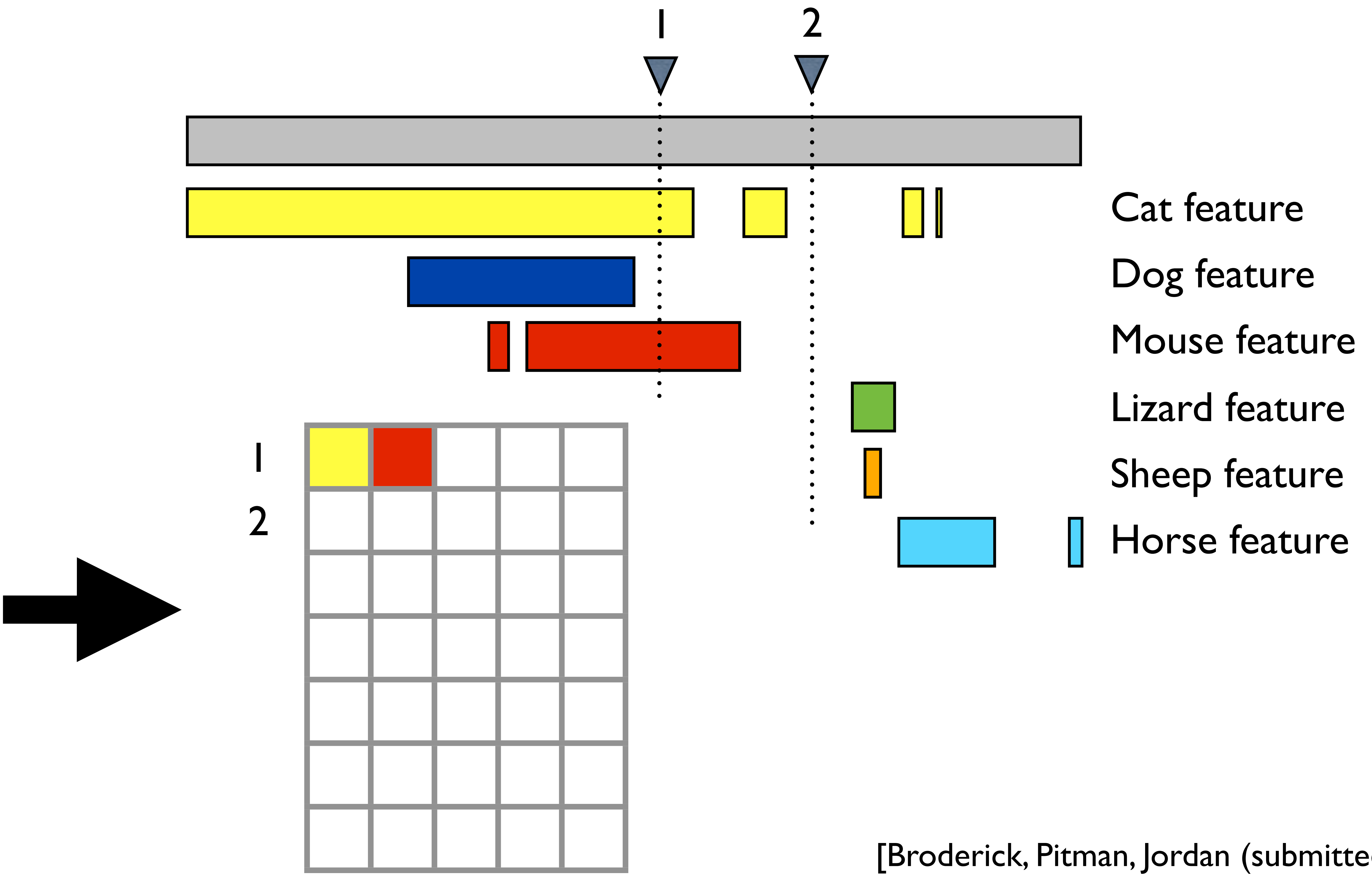
[Broderick, Pitman, Jordan (submitted)]

Feature paintbox



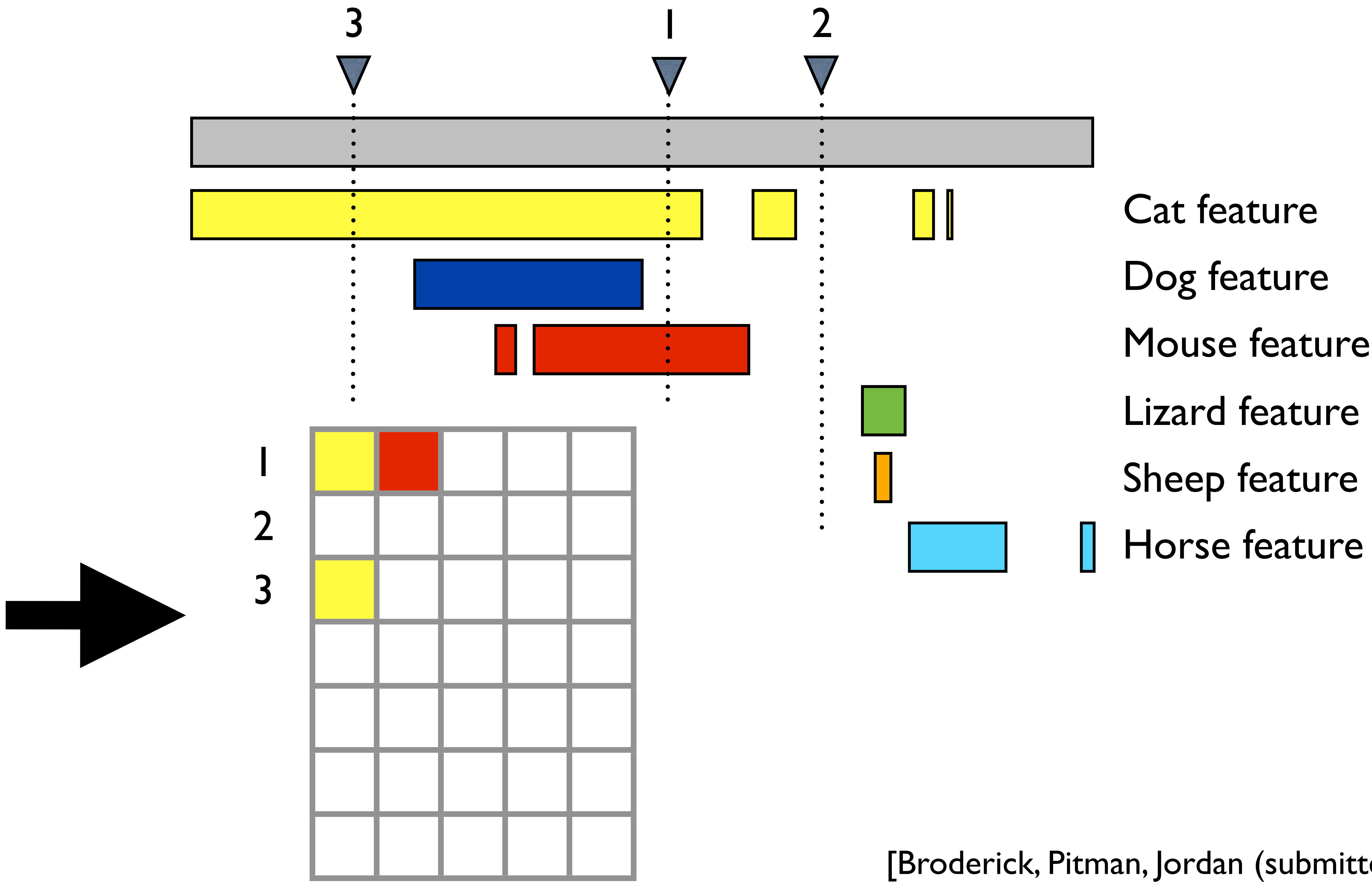
[Broderick, Pitman, Jordan (submitted)]

Feature paintbox



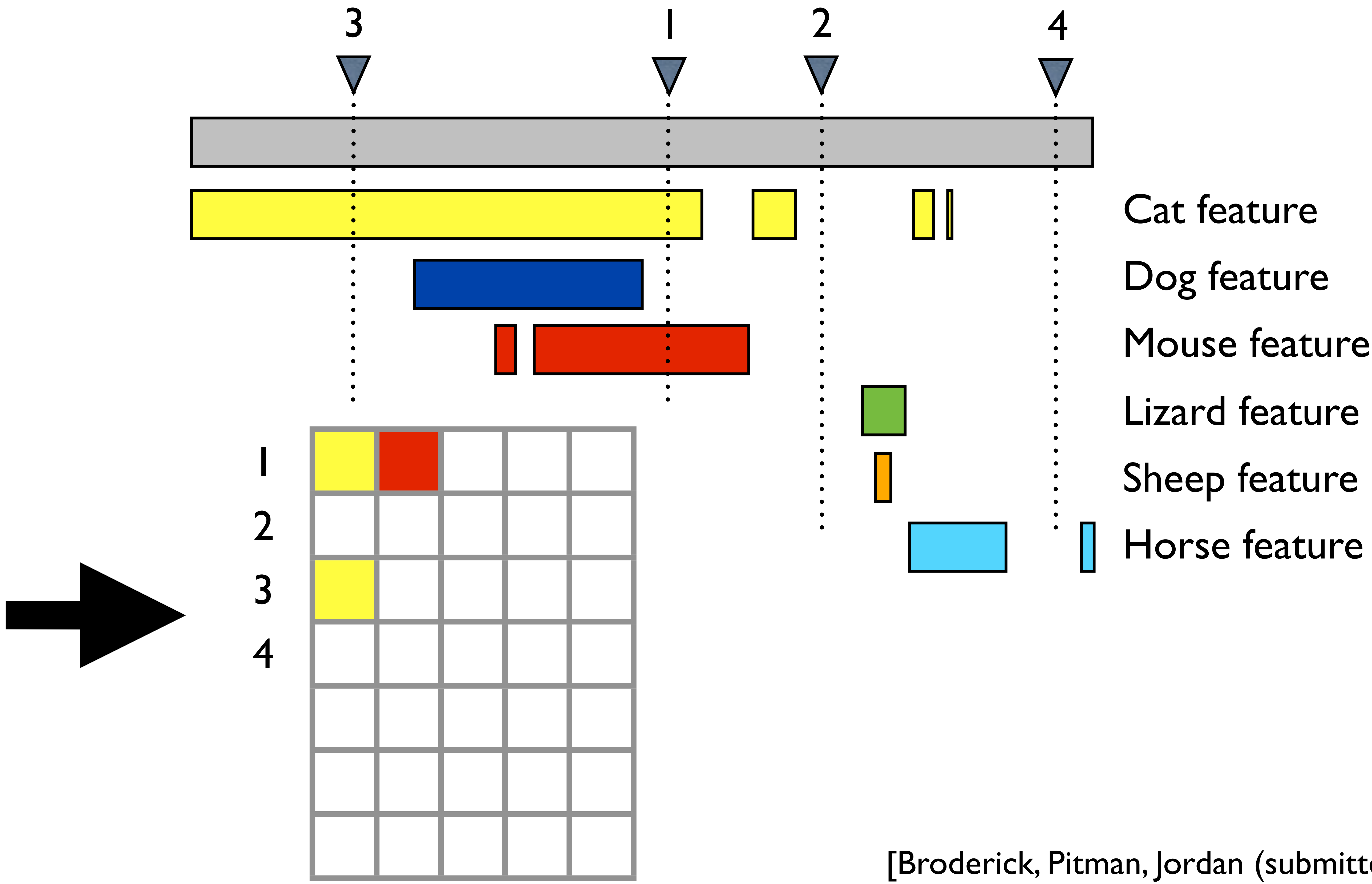
[Broderick, Pitman, Jordan (submitted)]

Feature paintbox



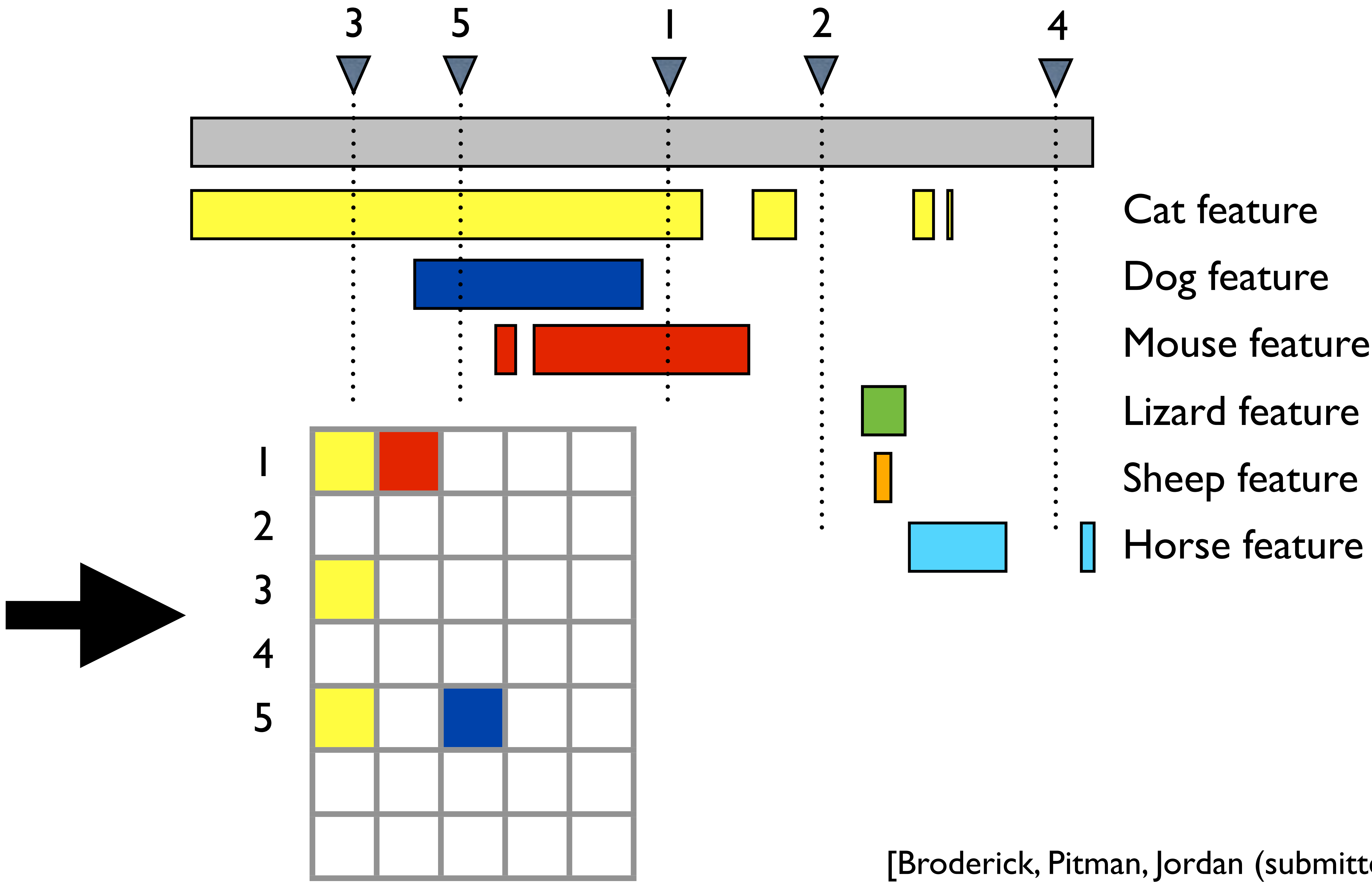
[Broderick, Pitman, Jordan (submitted)]

Feature paintbox



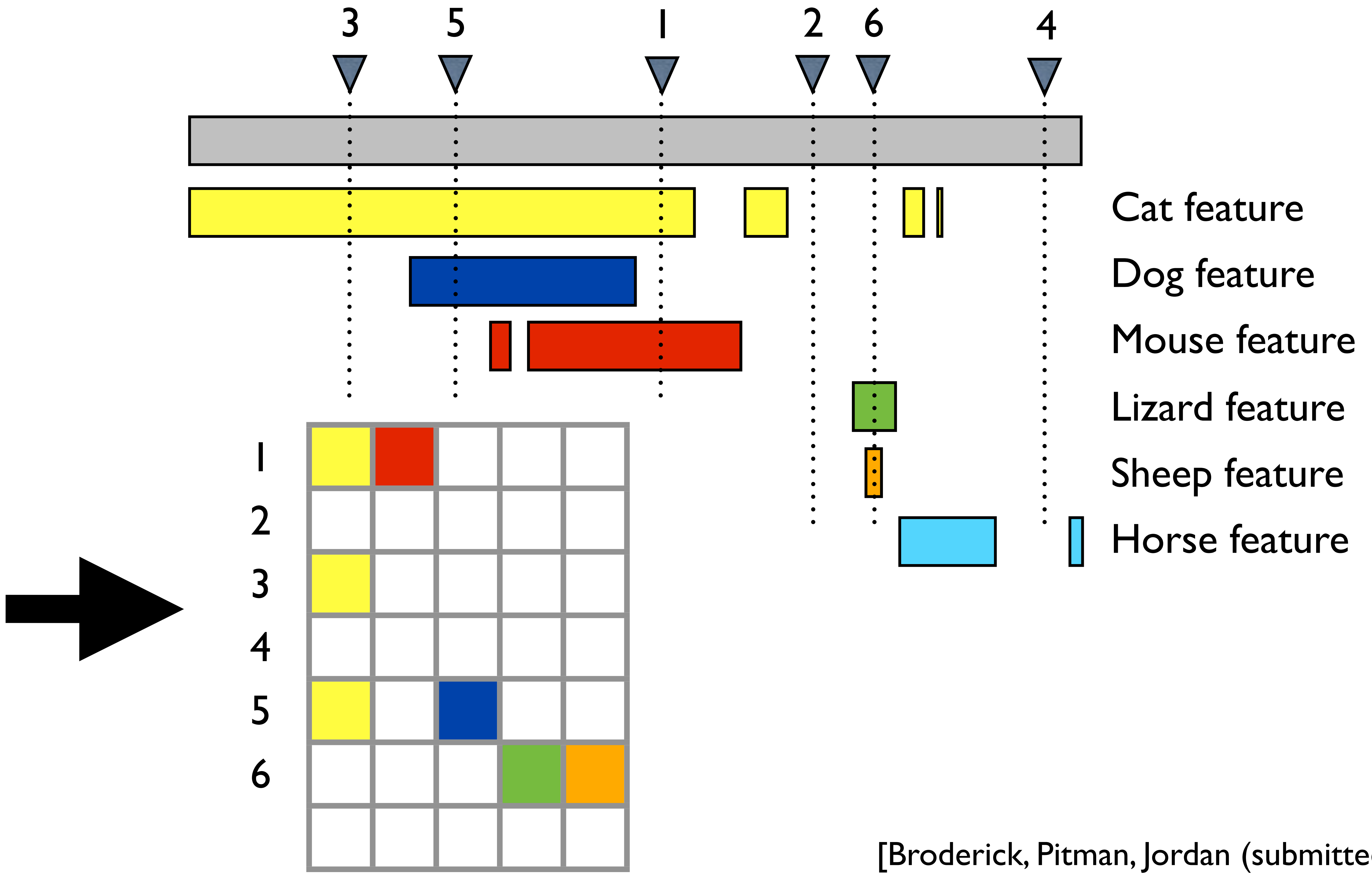
[Broderick, Pitman, Jordan (submitted)]

Feature paintbox



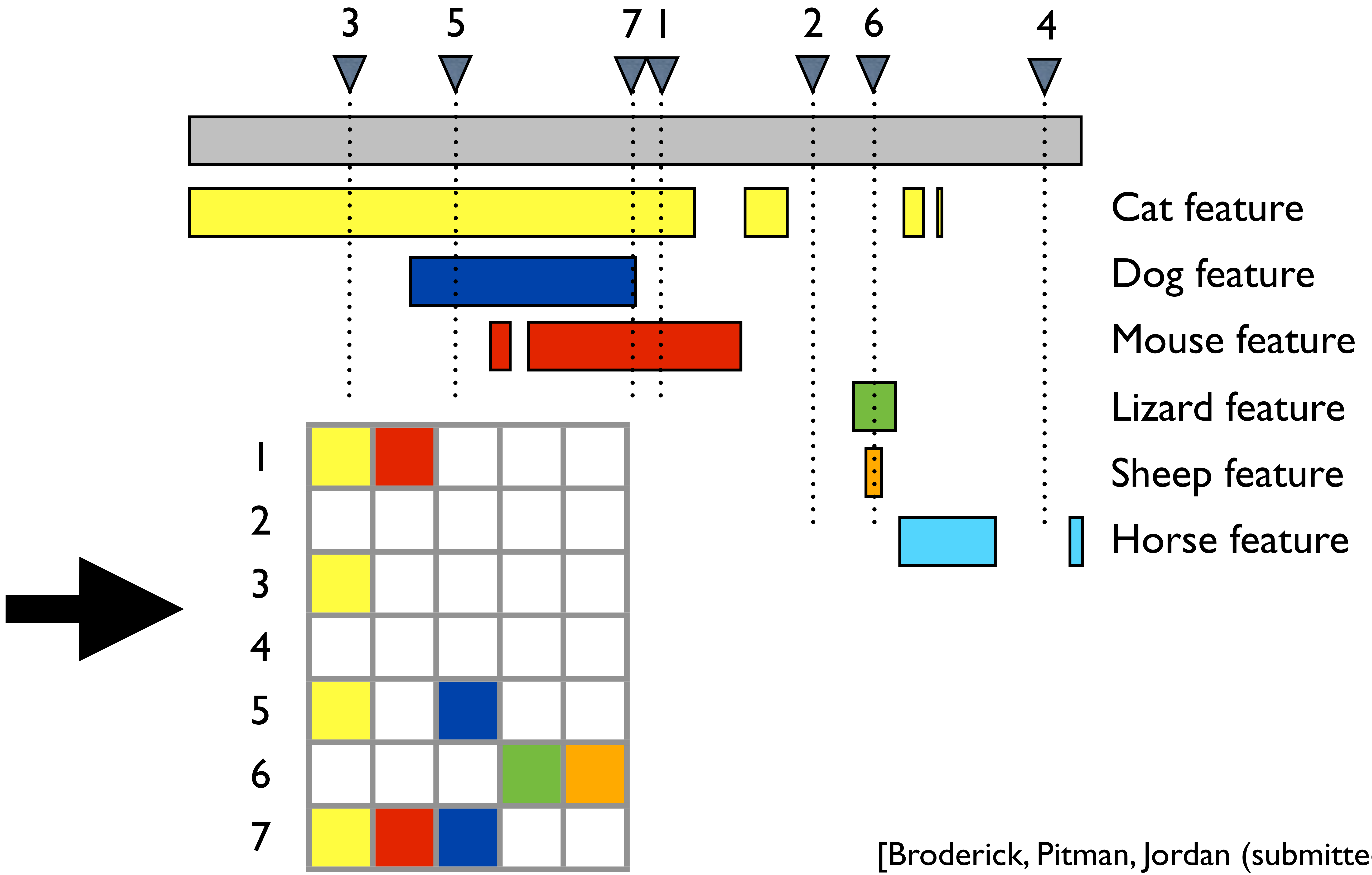
[Broderick, Pitman, Jordan (submitted)]

Feature paintbox



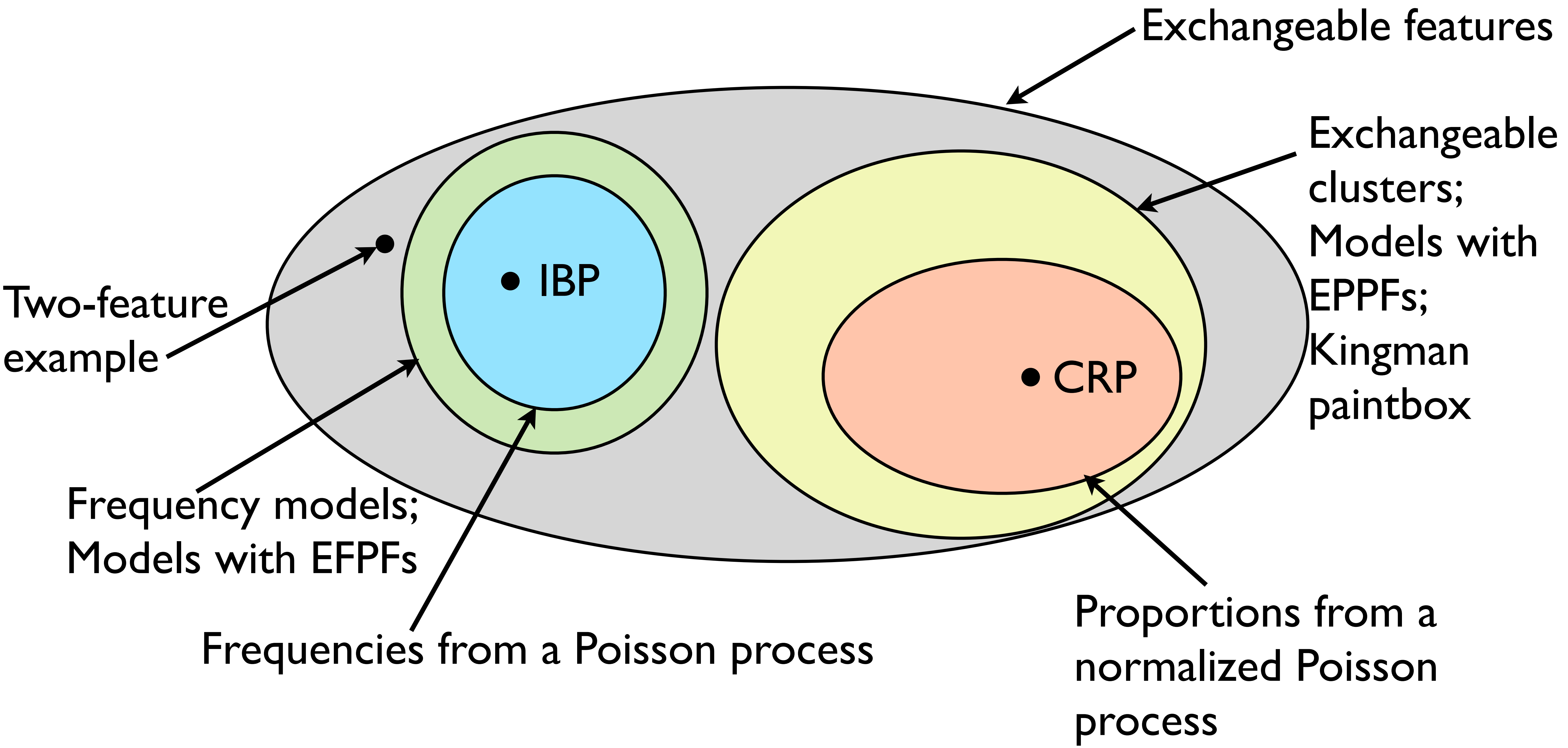
[Broderick, Pitman, Jordan (submitted)]

Feature paintbox

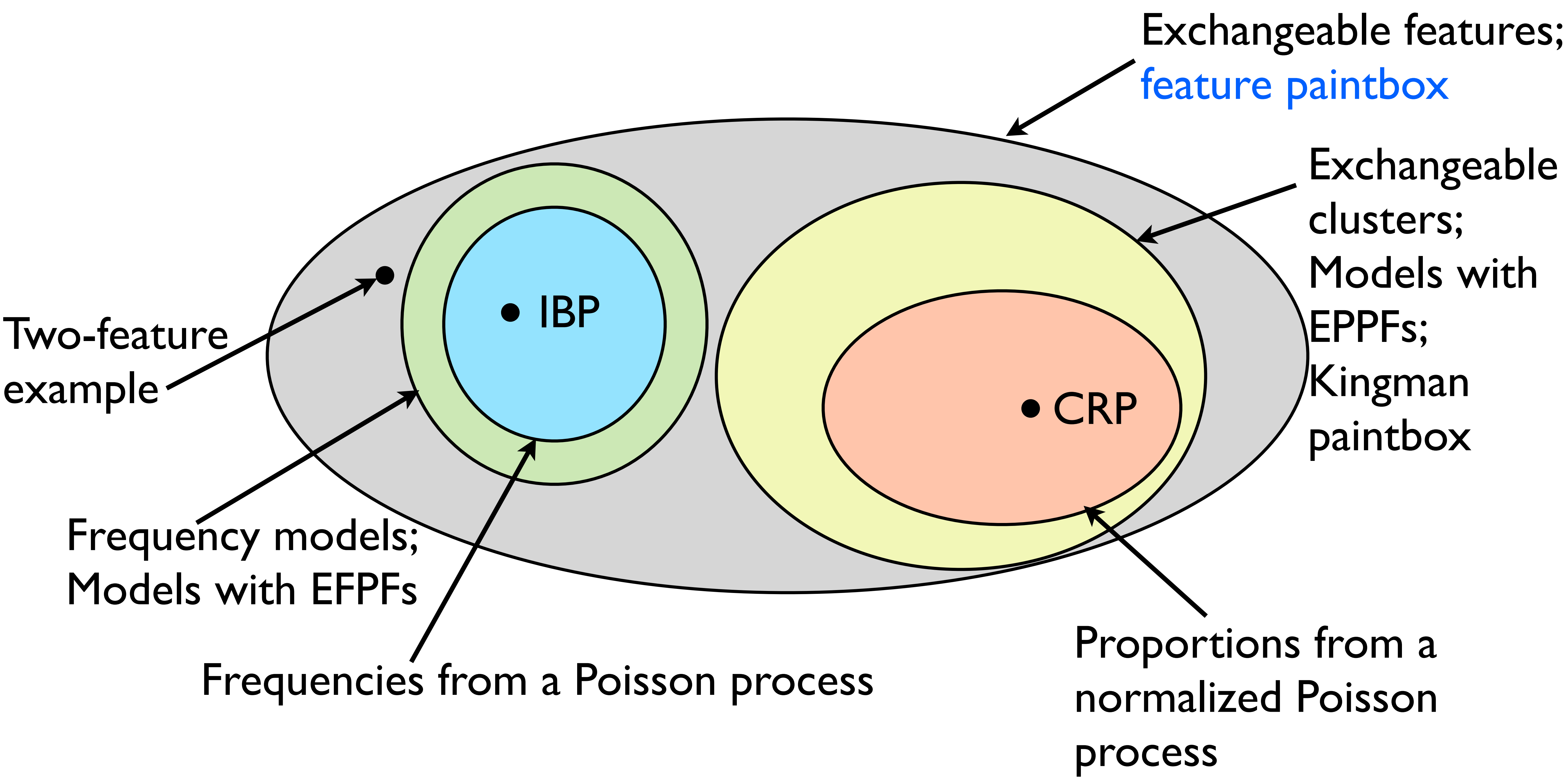


[Broderick, Pitman, Jordan (submitted)]

Clusters and features



Clusters and features



Clusters and features