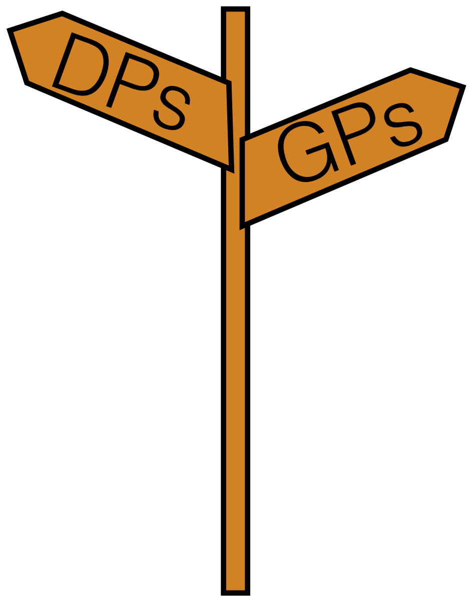
Nonparametric Bayesian Methods: Models, Algorithms, and Applications (Day 4)

Tamara Broderick
ITT Career Development Assistant Professor
Electrical Engineering & Computer Science
MIT

Roadmap



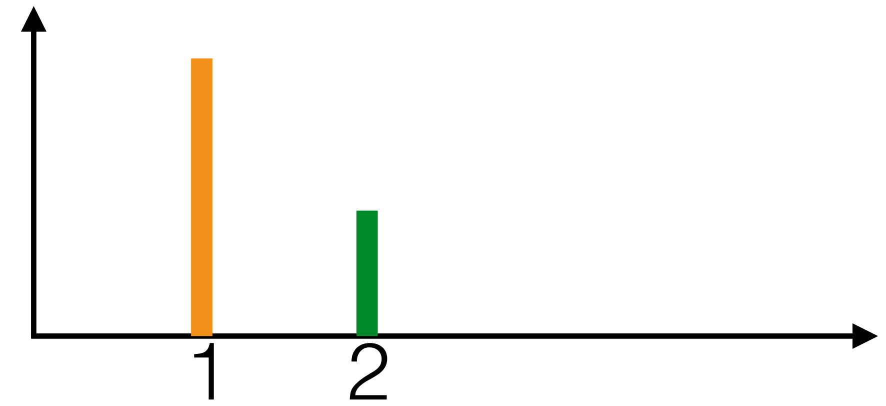
- Bayes Foundations
- Unsupervised Learning
  - Example problem: clustering
  - Example BNP model: Dirichlet process (DP)
  - Chinese restaurant process
- Supervised Learning
  - Example problem: regression
  - Example BNP model: Gaussian process (GP)
- Venture further into the wild world of Nonparametric Bayes

- Big questions
  - Why BNP?
  - What does an infinite/growing number of parameters really mean (in BNP)?
  - Why is BNP challenging but practical?

Marginal cluster assignments

Marginal cluster assignments

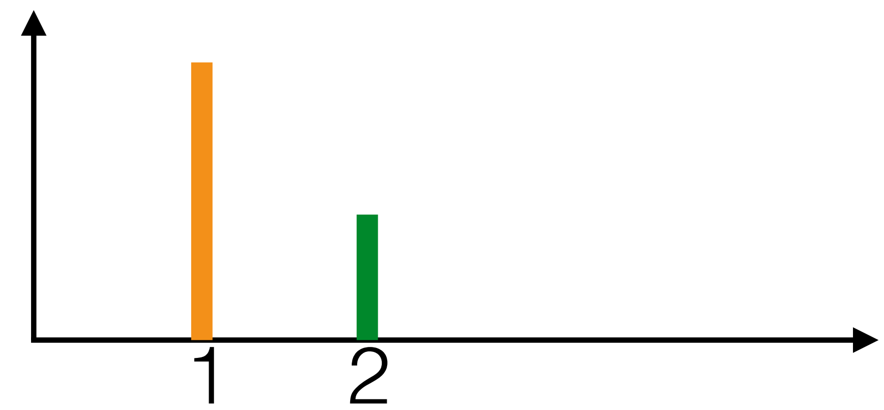
$$\rho_1 \sim \mathrm{Beta}(a_1, a_2), z_n \overset{iid}{\sim} \mathrm{Cat}(\rho_1, \rho_2)$$



Marginal cluster assignments

- Integrate out the frequencies

$$\rho_1 \sim \mathrm{Beta}(a_1, a_2), z_n \overset{iid}{\sim} \mathrm{Cat}(\rho_1, \rho_2)$$

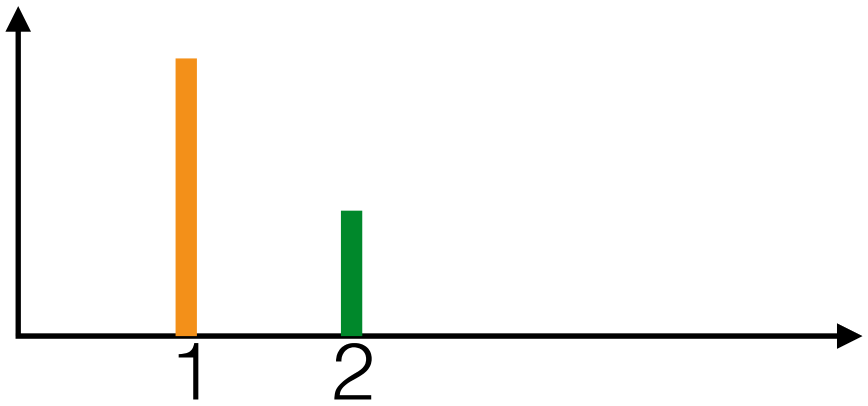


Marginal cluster assignments

- Integrate out the frequencies

$$\rho_1 \sim \mathrm{Beta}(a_1, a_2), z_n \overset{iid}{\sim} \mathrm{Cat}(\rho_1, \rho_2)$$

$$p(z_n = 1 | z_1, \ldots, z_{n-1})$$

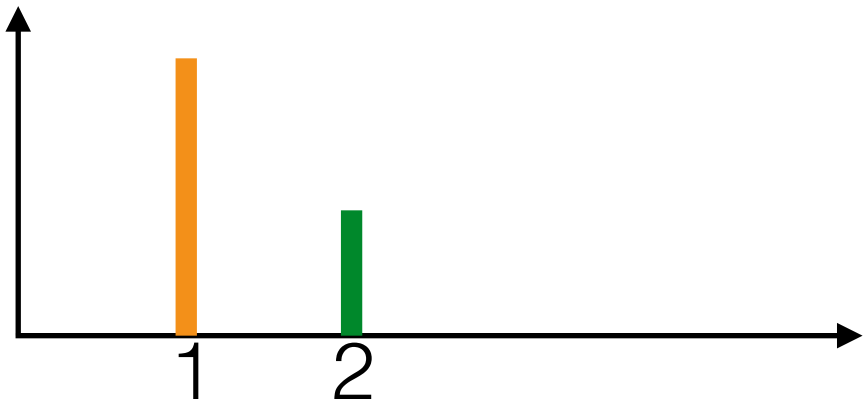


Marginal cluster assignments

- Integrate out the frequencies

$$\rho_1 \sim \mathrm{Beta}(a_1, a_2), z_n \overset{iid}{\sim} \mathrm{Cat}(\rho_1, \rho_2)$$

$$p(z_n = 1 | z_1, \ldots, z_{n-1}) = \frac{a_{1,n}}{a_{1,n} + a_{2,n}}$$

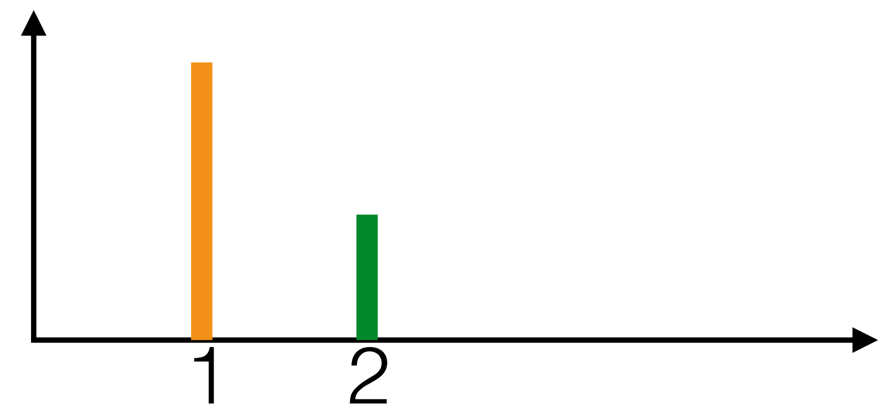


Marginal cluster assignments

- Integrate out the frequencies

$$\rho_1 \sim \mathrm{Beta}(a_1, a_2), z_n \overset{iid}{\sim} \mathrm{Cat}(\rho_1, \rho_2)$$

$$p(z_n = 1 | z_1, \ldots, z_{n-1}) = \frac{a_{1,n}}{a_{1,n} + a_{2,n}}$$



$$a_{1,n} := a_1 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 1\}, a_{2,n} = a_2 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 2\}$$

Marginal cluster assignments

- Integrate out the frequencies

$$\rho_1 \sim \text{Beta}(a_1, a_2), z_n \overset{iid}{\sim} \text{Cat}(\rho_1, \rho_2)$$

$$p(z_n = 1 | z_1, \ldots, z_{n-1}) = \frac{a_{1,n}}{a_{1,n} + a_{2,n}}$$

$$a_{1,n} := a_1 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 1\}, a_{2,n} = a_2 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 2\}$$

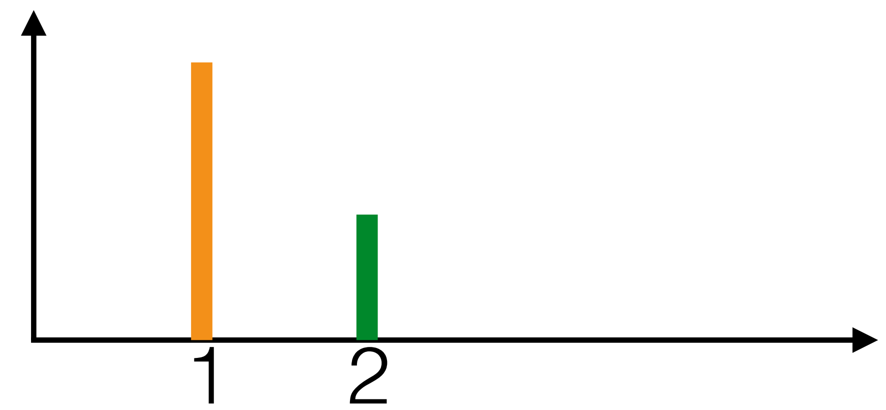
Marginal cluster assignments

- Integrate out the frequencies

$$\rho_1 \sim \mathrm{Beta}(a_1, a_2), z_n \stackrel{iid}{\sim} \mathrm{Cat}(\rho_1, \rho_2)$$

$$p(z_n = 1 | z_1, \ldots, z_{n-1}) = \frac{a_{1,n}}{a_{1,n} + a_{2,n}}$$

$$a_{1,n} := a_1 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 1\}, a_{2,n} = a_2 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 2\}$$

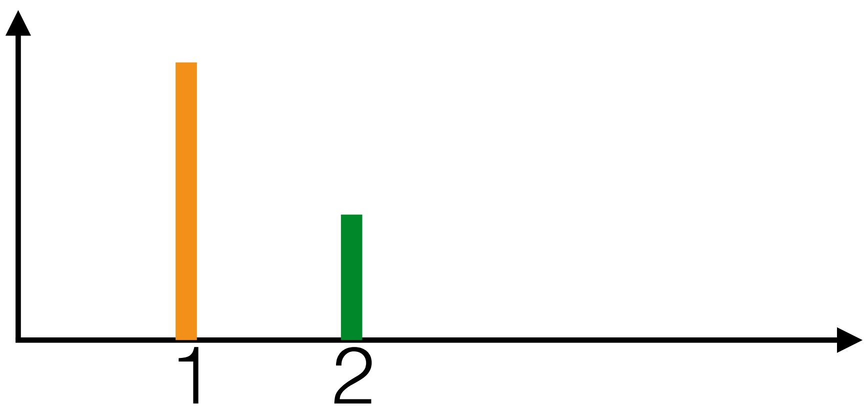


Marginal cluster assignments

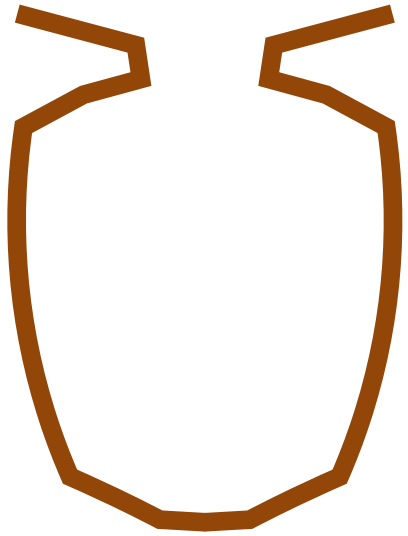
- Integrate out the frequencies

$$\rho_1 \sim \mathrm{Beta}(a_1, a_2), z_n \overset{iid}{\sim} \mathrm{Cat}(\rho_1, \rho_2)$$

$$p(z_n = 1 | z_1, \ldots, z_{n-1}) = \frac{a_{1,n}}{a_{1,n} + a_{2,n}}$$



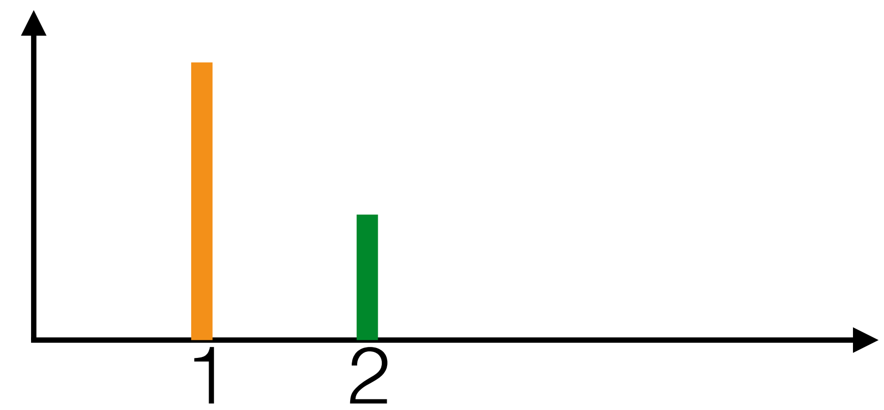
$$a_{1,n} := a_1 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 1\}, a_{2,n} = a_2 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 2\}$$



Marginal cluster assignments

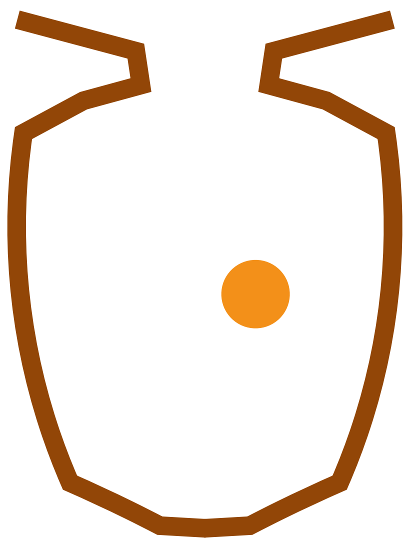
- Integrate out the frequencies

$$\rho_1 \sim \mathrm{Beta}(a_1, a_2),\, z_n \overset{iid}{\sim} \mathrm{Cat}(\rho_1, \rho_2)$$



$$p(z_n = 1 | z_1, \ldots, z_{n-1}) = \frac{a_{1,n}}{a_{1,n} + a_{2,n}}$$

$$a_{1,n} := a_1 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 1\},\, a_{2,n} = a_2 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 2\}$$

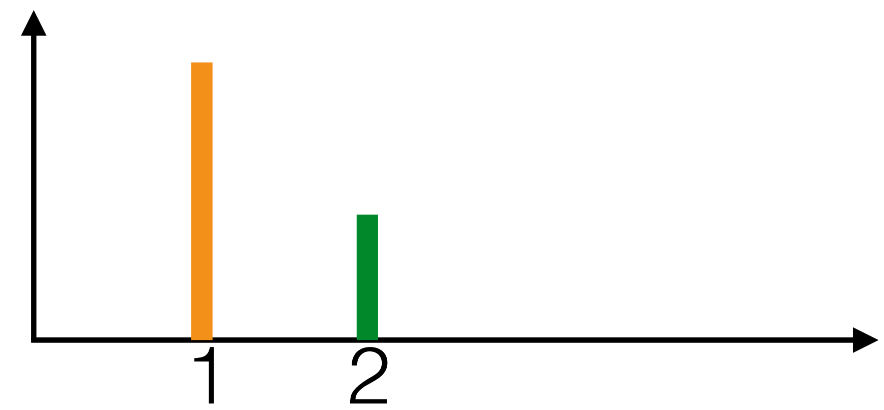


Marginal cluster assignments

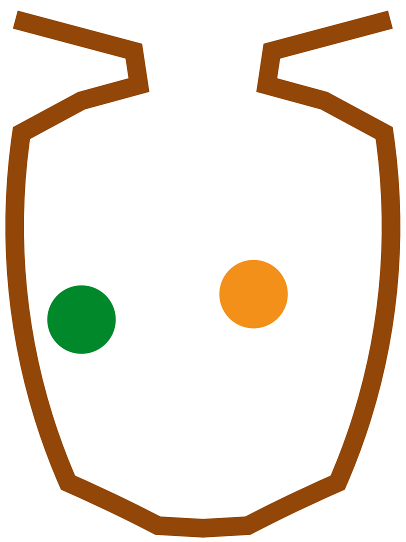
- Integrate out the frequencies

$$\rho_1 \sim \mathrm{Beta}(a_1, a_2), z_n \overset{iid}{\sim} \mathrm{Cat}(\rho_1, \rho_2)$$

$$p(z_n = 1 | z_1, \ldots, z_{n-1}) = \frac{a_{1,n}}{a_{1,n} + a_{2,n}}$$



$$a_{1,n} := a_1 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 1\}, a_{2,n} = a_2 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 2\}$$

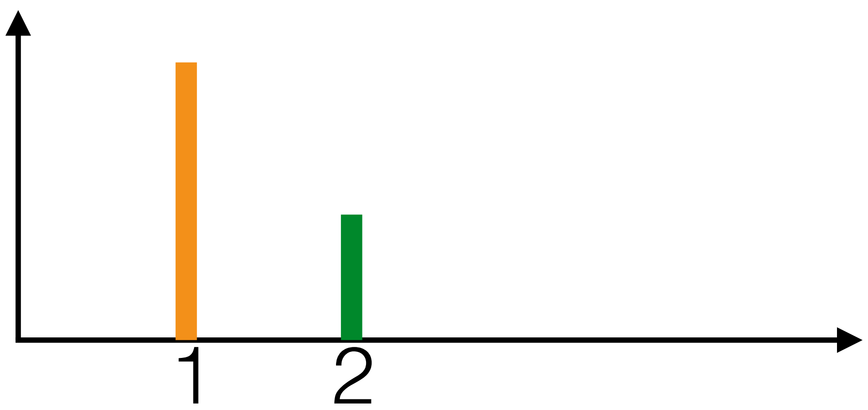


Marginal cluster assignments

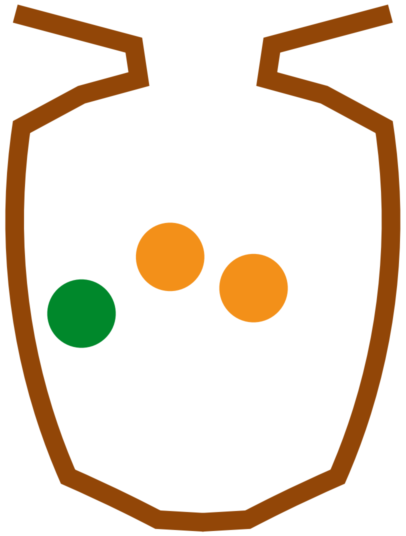
- Integrate out the frequencies

$$\rho_1 \sim \mathrm{Beta}(a_1, a_2), z_n \stackrel{iid}{\sim} \mathrm{Cat}(\rho_1, \rho_2)$$

$$p(z_n = 1 | z_1, \ldots, z_{n-1}) = \frac{a_{1,n}}{a_{1,n} + a_{2,n}}$$



$$a_{1,n} := a_1 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 1\}, a_{2,n} = a_2 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 2\}$$

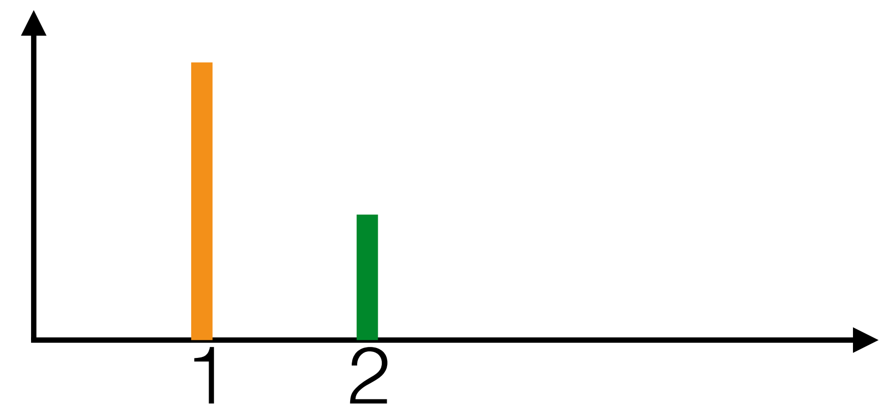


Marginal cluster assignments

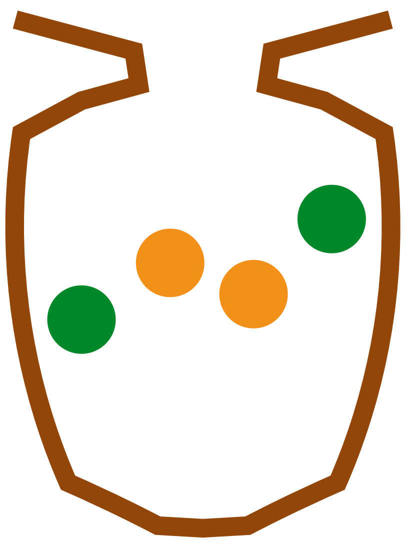
- Integrate out the frequencies

$$\rho_1 \sim \text{Beta}(a_1, a_2), z_n \overset{iid}{\sim} \text{Cat}(\rho_1, \rho_2)$$

$$p(z_n = 1 | z_1, \ldots, z_{n-1}) = \frac{a_{1,n}}{a_{1,n} + a_{2,n}}$$



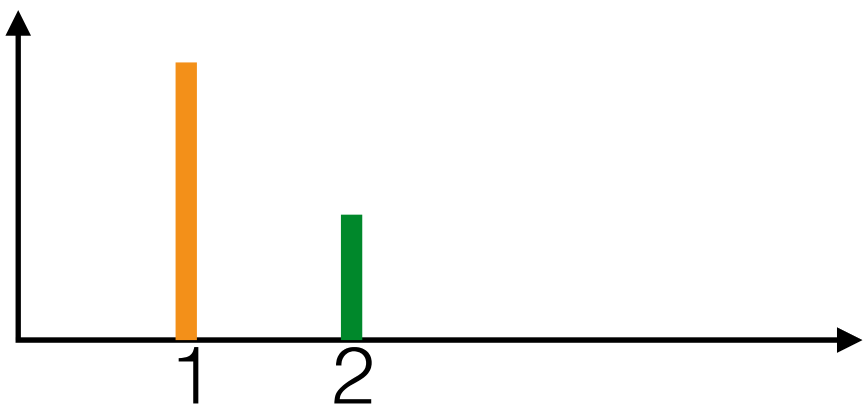
$$a_{1,n} := a_1 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 1\}, a_{2,n} = a_2 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 2\}$$



Marginal cluster assignments

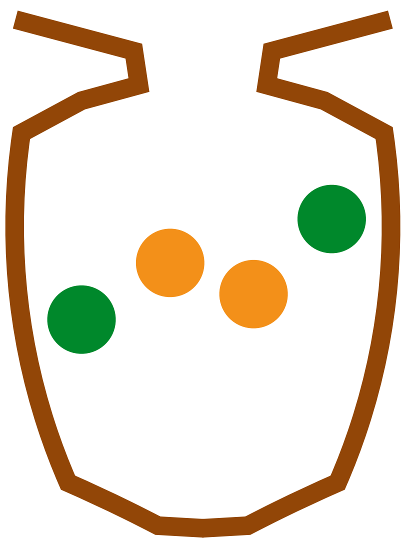
- Integrate out the frequencies

$$\rho_1 \sim \text{Beta}(a_1, a_2), z_n \stackrel{iid}{\sim} \text{Cat}(\rho_1, \rho_2)$$



$$p(z_n = 1 | z_1, \ldots, z_{n-1}) = \frac{a_{1,n}}{a_{1,n} + a_{2,n}}$$

$$a_{1,n} := a_1 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 1\}, a_{2,n} = a_2 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 2\}$$

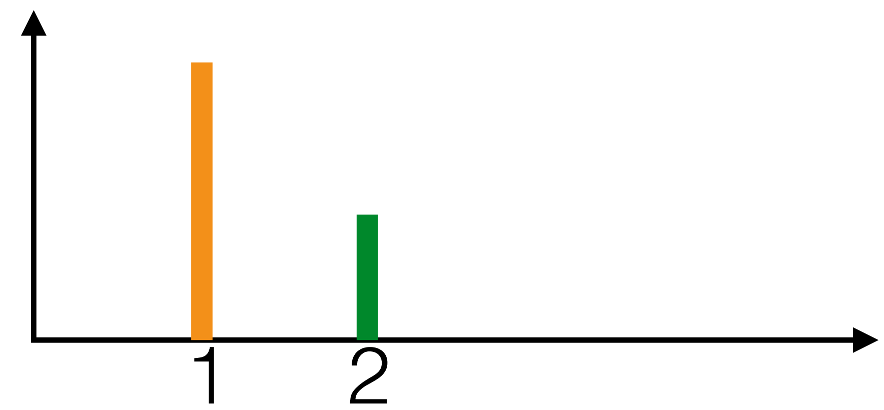


$$\lim_{n \to \infty} \frac{\#\text{ orange}}{\#\text{ total}}$$

Marginal cluster assignments

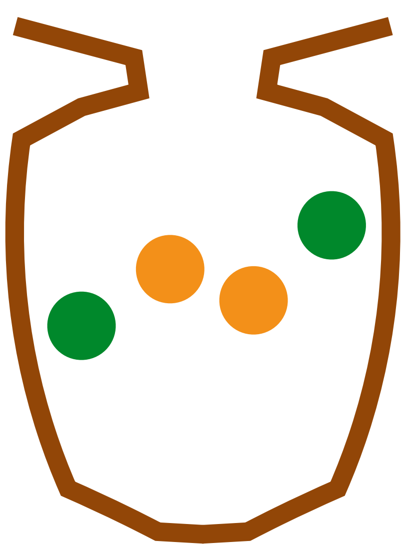
- Integrate out the frequencies

$$\rho_1 \sim \mathrm{Beta}(a_1, a_2),\, z_n \overset{iid}{\sim} \mathrm{Cat}(\rho_1, \rho_2)$$



$$p(z_n = 1 | z_1, \ldots, z_{n-1}) = \frac{a_{1,n}}{a_{1,n} + a_{2,n}}$$

$$a_{1,n} := a_1 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 1\},\, a_{2,n} = a_2 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 2\}$$

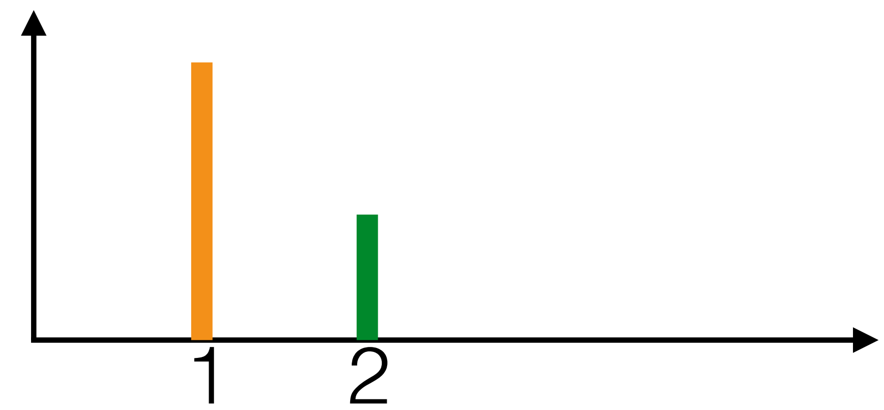


$$\lim_{n \to \infty} \frac{\#\ \mathrm{orange}}{\#\ \mathrm{total}} = \rho_{\mathrm{orange}}$$

Marginal cluster assignments

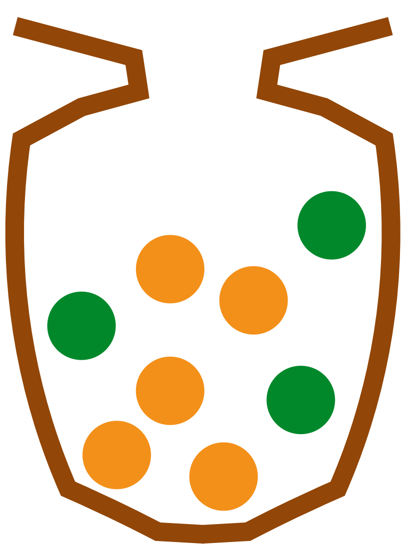
- Integrate out the frequencies

$$\rho_1 \sim \text{Beta}(a_1, a_2), z_n \stackrel{iid}{\sim} \text{Cat}(\rho_1, \rho_2)$$



$$p(z_n = 1 | z_1, \ldots, z_{n-1}) = \frac{a_{1,n}}{a_{1,n} + a_{2,n}}$$

$$a_{1,n} := a_1 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 1\}, a_{2,n} = a_2 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 2\}$$

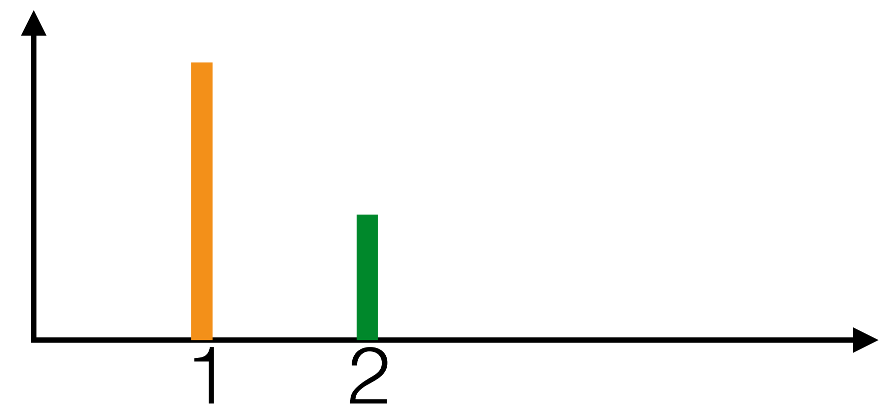


Marginal cluster assignments

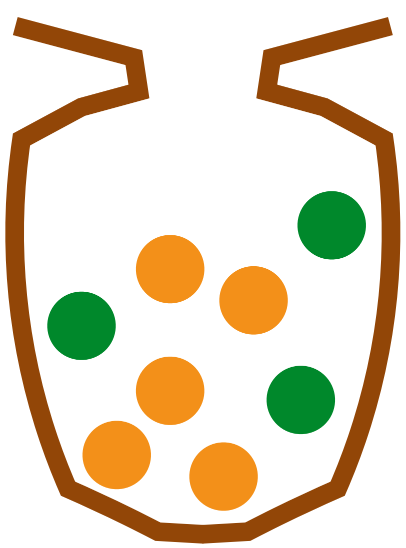
- Integrate out the frequencies

$$\rho_1 \sim \mathrm{Beta}(a_1, a_2),\ z_n \overset{iid}{\sim} \mathrm{Cat}(\rho_1, \rho_2)$$

$$p(z_n = 1 | z_1, \ldots, z_{n-1}) = \frac{a_{1,n}}{a_{1,n} + a_{2,n}}$$



$$a_{1,n} := a_1 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 1\},\ a_{2,n} = a_2 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 2\}$$



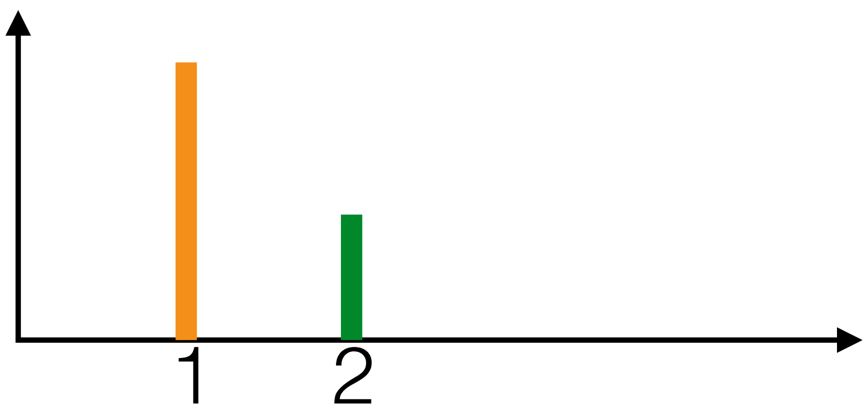
$$\lim_{n \to \infty} \frac{\#\ \text{orange}}{\#\ \text{total}}$$

Marginal cluster assignments

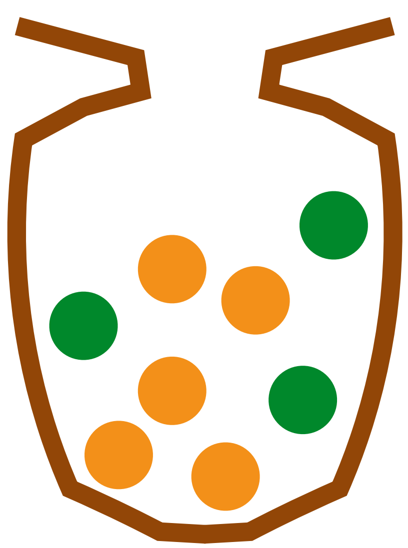
- Integrate out the frequencies

$$\rho_1 \sim \text{Beta}(a_1, a_2), z_n \overset{iid}{\sim} \text{Cat}(\rho_1, \rho_2)$$

$$p(z_n = 1 | z_1, \ldots, z_{n-1}) = \frac{a_{1,n}}{a_{1,n} + a_{2,n}}$$



$$a_{1,n} := a_1 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 1\}, a_{2,n} = a_2 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 2\}$$



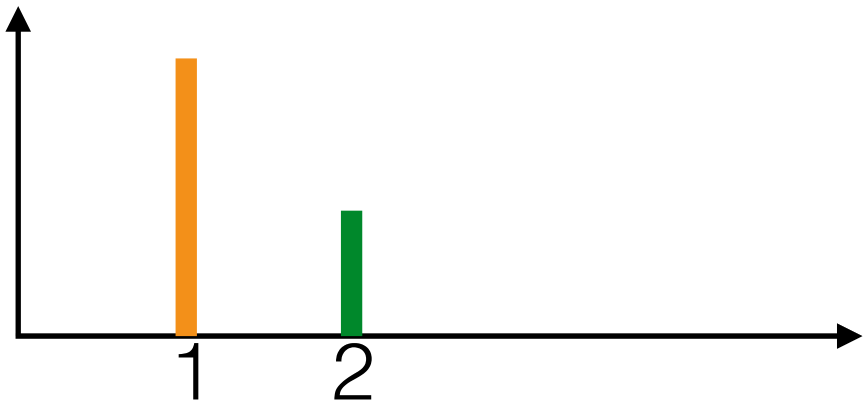
$$\lim_{n \to \infty} \frac{\#\ \text{orange}}{\#\ \text{total}} = \rho_{\text{orange}}$$

Marginal cluster assignments

- Integrate out the frequencies

$$\rho_1 \sim \mathrm{Beta}(a_1, a_2), z_n \stackrel{iid}{\sim} \mathrm{Cat}(\rho_1, \rho_2)$$

$$p(z_n = 1 | z_1, \ldots, z_{n-1}) = \frac{a_{1,n}}{a_{1,n} + a_{2,n}}$$



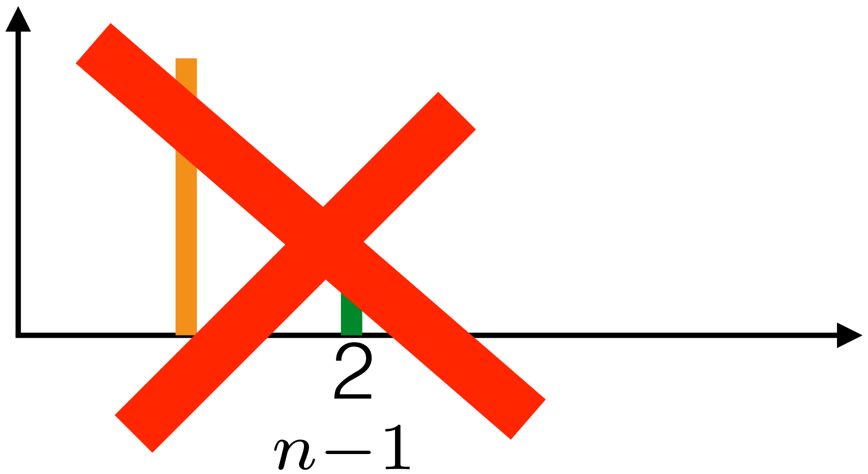
$$a_{1,n} := a_1 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 1\}, a_{2,n} = a_2 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 2\}$$

Marginal cluster assignments

- Integrate out the frequencies

$$\rho_1 \sim \mathrm{Beta}(a_1, a_2),\, z_n \overset{iid}{\sim} \mathrm{Cat}(\rho_1, \rho_2)$$

$$p(z_n = 1 | z_1, \ldots, z_{n-1}) = \frac{a_{1,n}}{a_{1,n} + a_{2,n}}$$



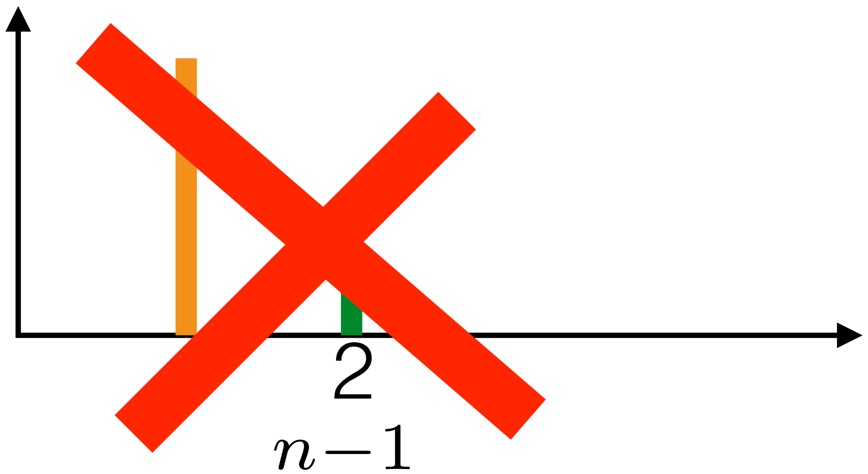
$$a_{1,n} := a_1 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 1\},\, a_{2,n} = a_2 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 2\}$$

Marginal cluster assignments

- Integrate out the frequencies

$$\rho_1 \sim \mathrm{Beta}(a_1, a_2), z_n \stackrel{iid}{\sim} \mathrm{Cat}(\rho_1, \rho_2)$$

$$p(z_n = 1 | z_1, \ldots, z_{n-1}) = \frac{a_{1,n}}{a_{1,n} + a_{2,n}}$$



$$a_{1,n} := a_1 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 1\}, a_{2,n} = a_2 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 2\}$$

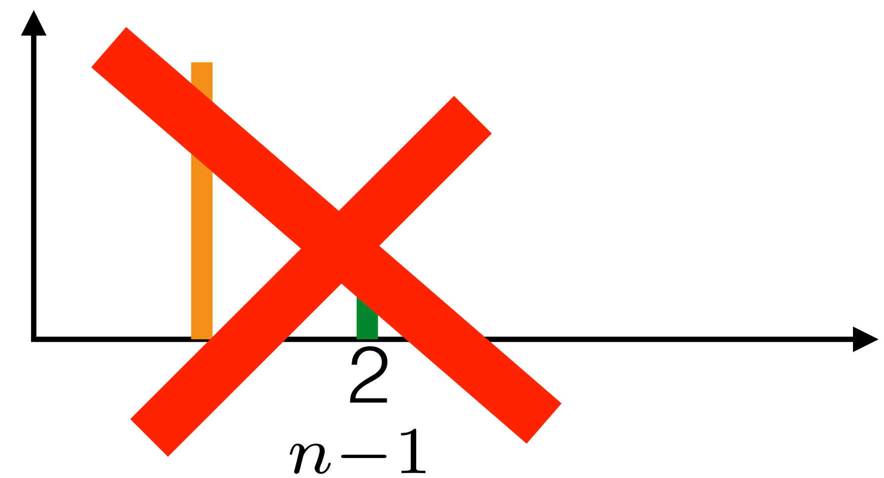
- Pólya urn

Marginal cluster assignments

- Integrate out the frequencies

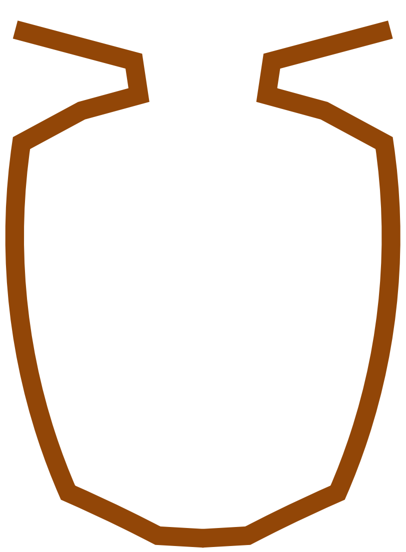
$$\rho_1 \sim \text{Beta}(a_1, a_2),\, z_n \stackrel{iid}{\sim} \text{Cat}(\rho_1, \rho_2)$$

$$p(z_n = 1 | z_1, \ldots, z_{n-1}) = \frac{a_{1,n}}{a_{1,n} + a_{2,n}}$$



$$a_{1,n} := a_1 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 1\},\, a_{2,n} = a_2 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 2\}$$

- Pólya urn

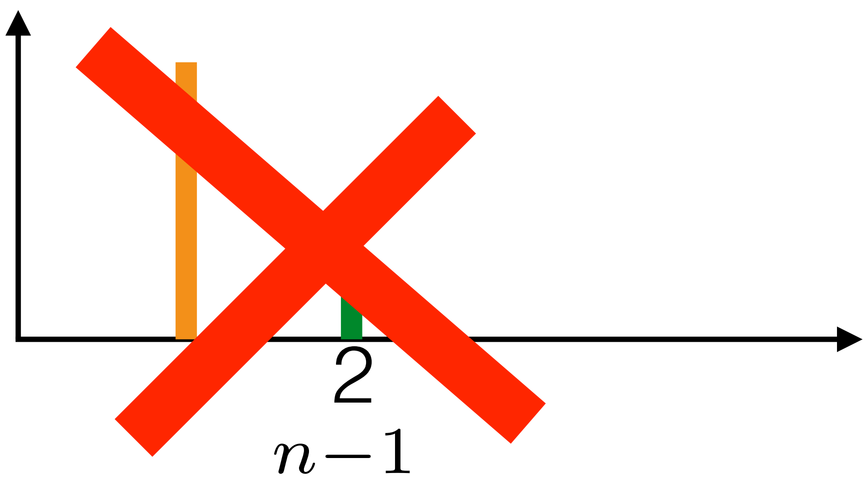


Marginal cluster assignments

- Integrate out the frequencies

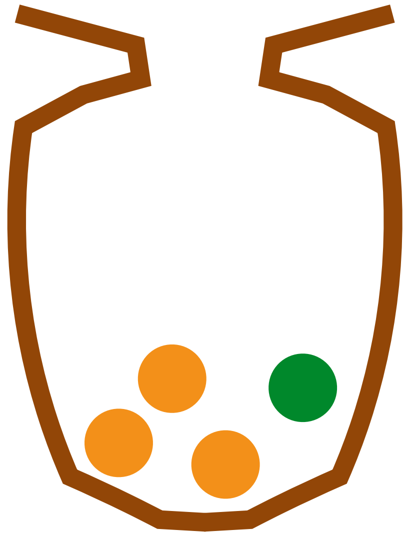
$$\rho_1 \sim \mathrm{Beta}(a_1, a_2),\, z_n \overset{iid}{\sim} \mathrm{Cat}(\rho_1, \rho_2)$$

$$p(z_n = 1 | z_1, \ldots, z_{n-1}) = \frac{a_{1,n}}{a_{1,n} + a_{2,n}}$$



$$a_{1,n} := a_1 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 1\},\, a_{2,n} = a_2 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 2\}$$

- Pólya urn

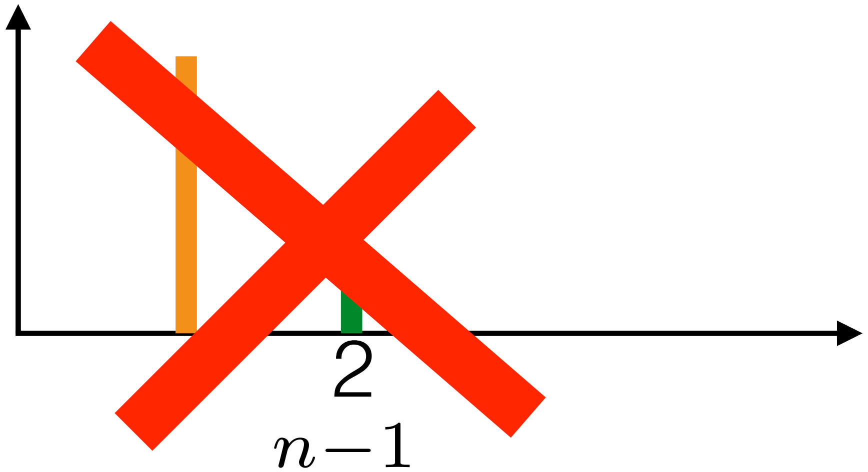


Marginal cluster assignments

- Integrate out the frequencies

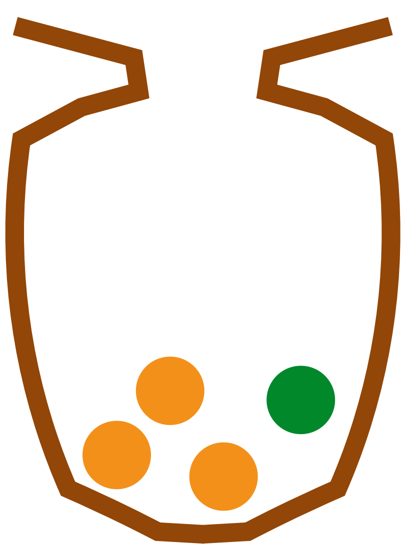
$$\rho_1 \sim \mathrm{Beta}(a_1, a_2),\, z_n \overset{iid}{\sim} \mathrm{Cat}(\rho_1, \rho_2)$$

$$p(z_n = 1 | z_1, \ldots, z_{n-1}) = \frac{a_{1,n}}{a_{1,n} + a_{2,n}}$$



$$a_{1,n} := a_1 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 1\},\, a_{2,n} = a_2 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 2\}$$

- Pólya urn
  - Choose any ball with equal probability

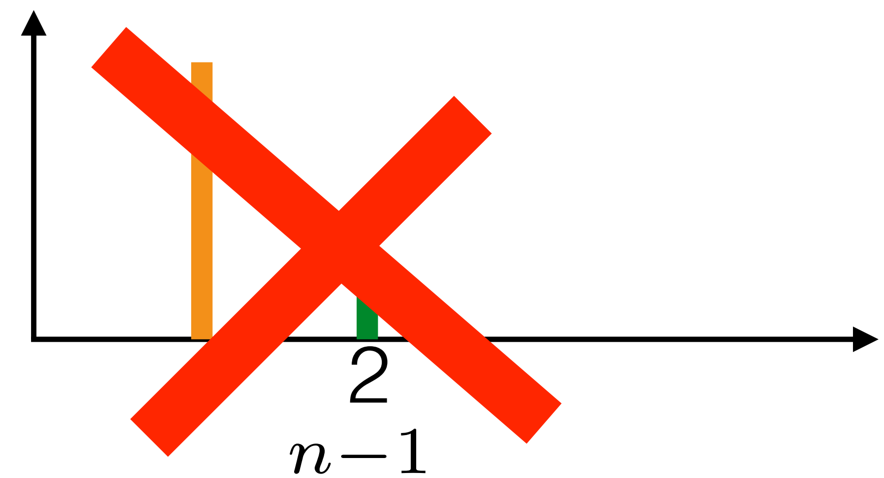


Marginal cluster assignments

- Integrate out the frequencies

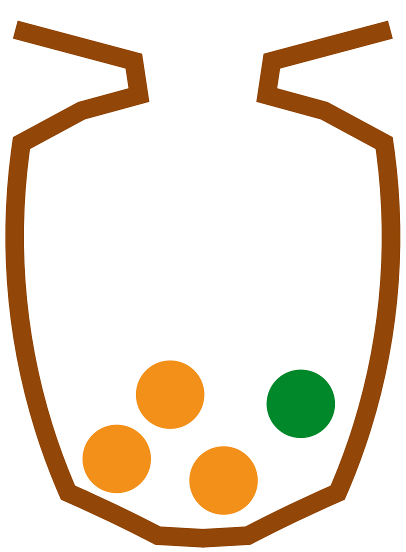
$$\rho_1 \sim \mathrm{Beta}(a_1, a_2),\, z_n \overset{iid}{\sim} \mathrm{Cat}(\rho_1, \rho_2)$$

$$p(z_n = 1 | z_1, \ldots, z_{n-1}) = \frac{a_{1,n}}{a_{1,n} + a_{2,n}}$$



$$a_{1,n} := a_1 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 1\},\, a_{2,n} = a_2 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 2\}$$

- Pólya urn
    - Choose any ball with equal probability
    - Replace and add ball of same color

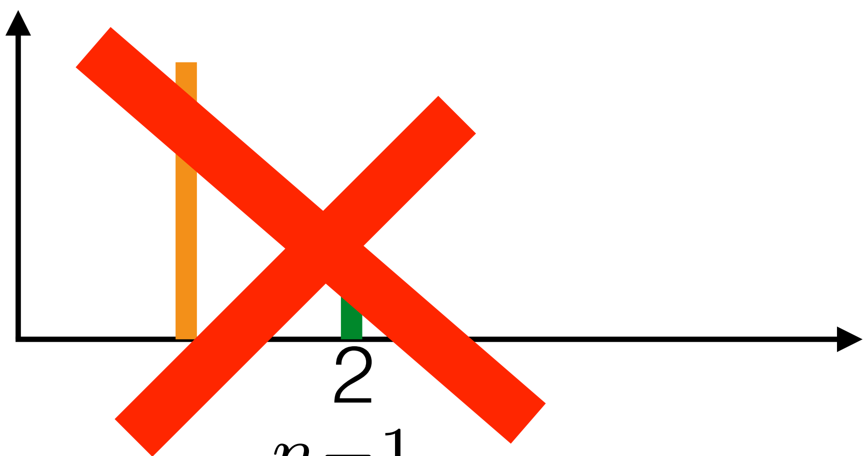


Marginal cluster assignments

- Integrate out the frequencies

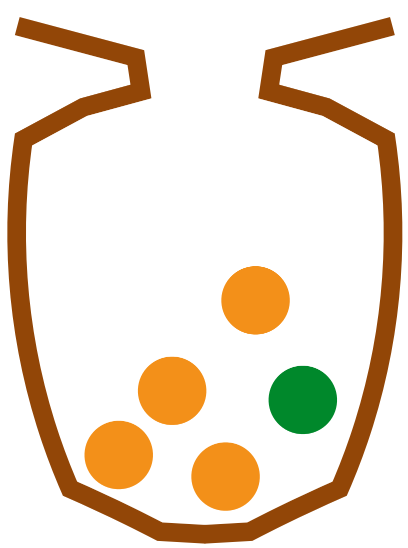
$$\rho_1 \sim \mathrm{Beta}(a_1, a_2),\, z_n \overset{iid}{\sim} \mathrm{Cat}(\rho_1, \rho_2)$$

$$p(z_n = 1 | z_1, \ldots, z_{n-1}) = \frac{a_{1,n}}{a_{1,n} + a_{2,n}}$$



$$a_{1,n} := a_1 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 1\},\, a_{2,n} = a_2 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 2\}$$

- Pólya urn
  - Choose any ball with equal probability
  - Replace and add ball of same color



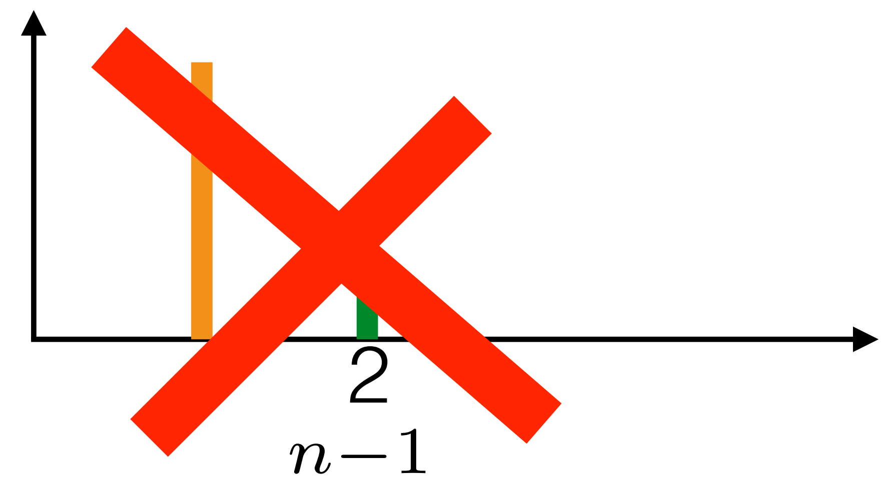
Marginal cluster assignments

- Integrate out the frequencies

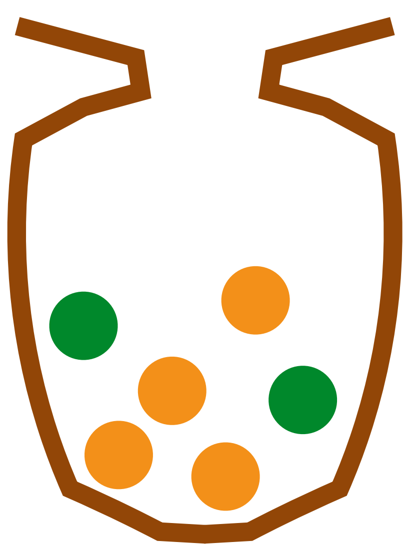
$$\rho_1 \sim \text{Beta}(a_1, a_2),\ z_n \stackrel{iid}{\sim} \text{Cat}(\rho_1, \rho_2)$$

$$p(z_n = 1 | z_1, \ldots, z_{n-1}) = \frac{a_{1,n}}{a_{1,n} + a_{2,n}}$$

$$a_{1,n} := a_1 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 1\},\ a_{2,n} = a_2 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 2\}$$



- Pólya urn
  - Choose any ball with equal probability
  - Replace and add ball of same color

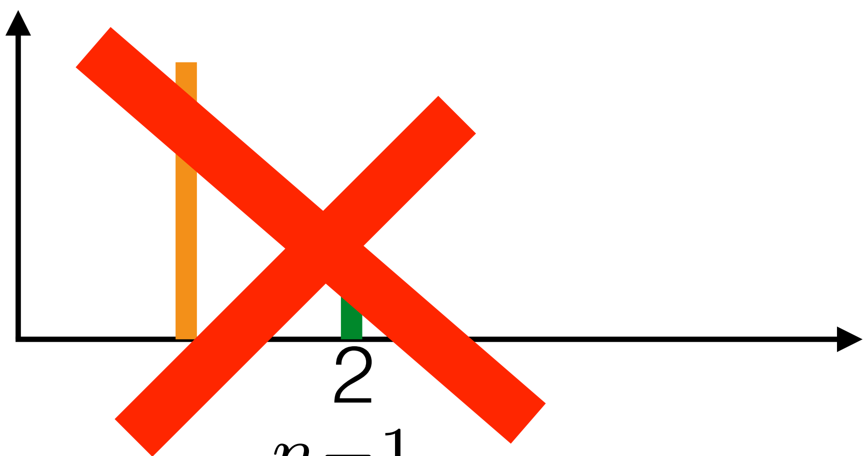


Marginal cluster assignments

- Integrate out the frequencies

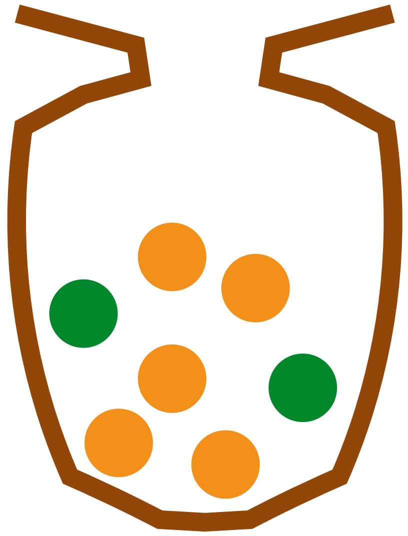
$$\rho_1 \sim \mathrm{Beta}(a_1, a_2),\ z_n \overset{iid}{\sim} \mathrm{Cat}(\rho_1, \rho_2)$$

$$p(z_n = 1 | z_1, \ldots, z_{n-1}) = \frac{a_{1,n}}{a_{1,n} + a_{2,n}}$$



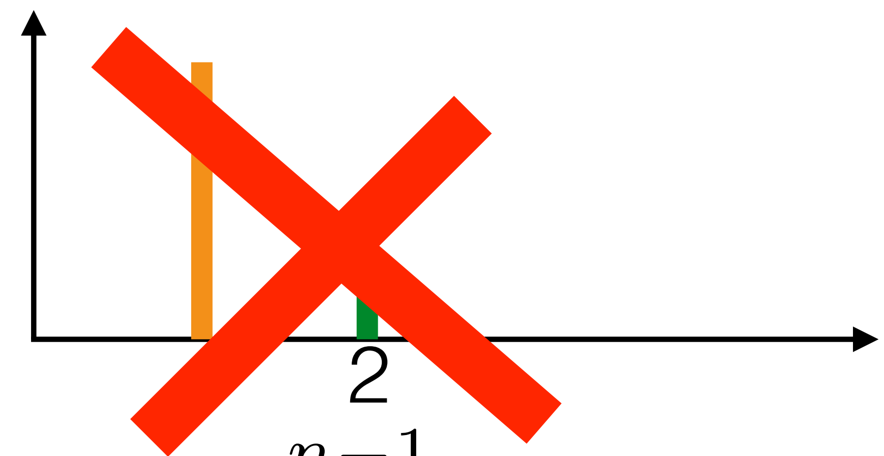
$$a_{1,n} := a_1 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 1\},\ a_{2,n} = a_2 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 2\}$$

- Pólya urn
  - Choose any ball with equal probability
  - Replace and add ball of same color



Marginal cluster assignments

- Integrate out the frequencies

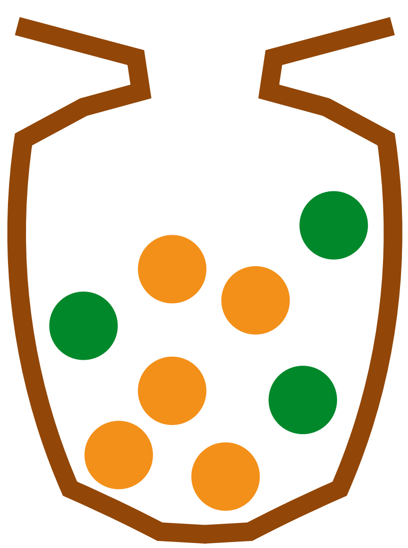


$$\rho_1 \sim \text{Beta}(a_1, a_2), z_n \overset{iid}{\sim} \text{Cat}(\rho_1, \rho_2)$$

$$p(z_n = 1 | z_1, \ldots, z_{n-1}) = \frac{a_{1,n}}{a_{1,n} + a_{2,n}}$$

$$a_{1,n} := a_1 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 1\}, a_{2,n} = a_2 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 2\}$$

- Pólya urn
  - Choose any ball with equal probability
  - Replace and add ball of same color

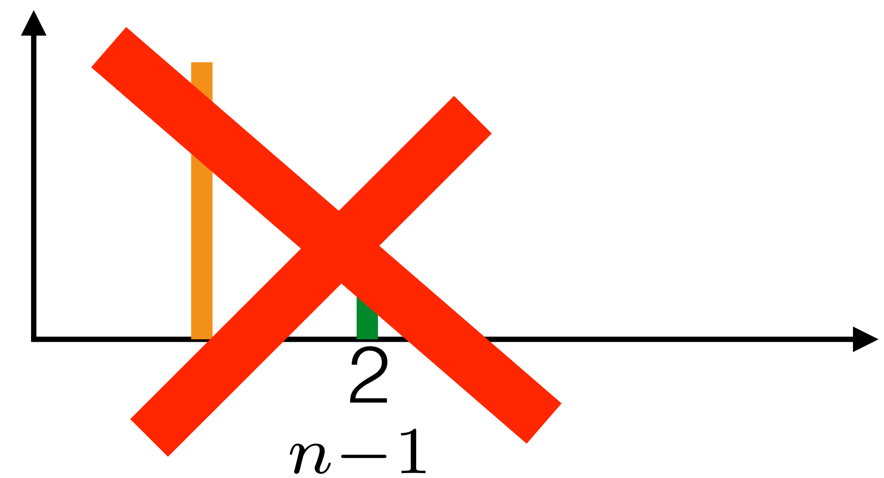


Marginal cluster assignments

- Integrate out the frequencies

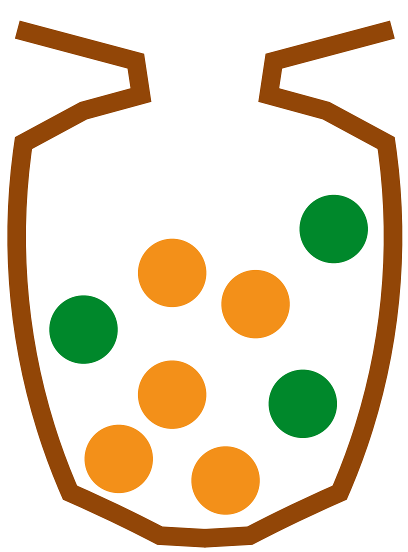
$$\rho_1 \sim \mathrm{Beta}(a_1, a_2),\, z_n \overset{iid}{\sim} \mathrm{Cat}(\rho_1, \rho_2)$$

$$p(z_n = 1 | z_1, \ldots, z_{n-1}) = \frac{a_{1,n}}{a_{1,n} + a_{2,n}}$$



$$a_{1,n} := a_1 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 1\},\, a_{2,n} = a_2 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 2\}$$

- Pólya urn
  - Choose any ball with equal probability
  - Replace and add ball of same color



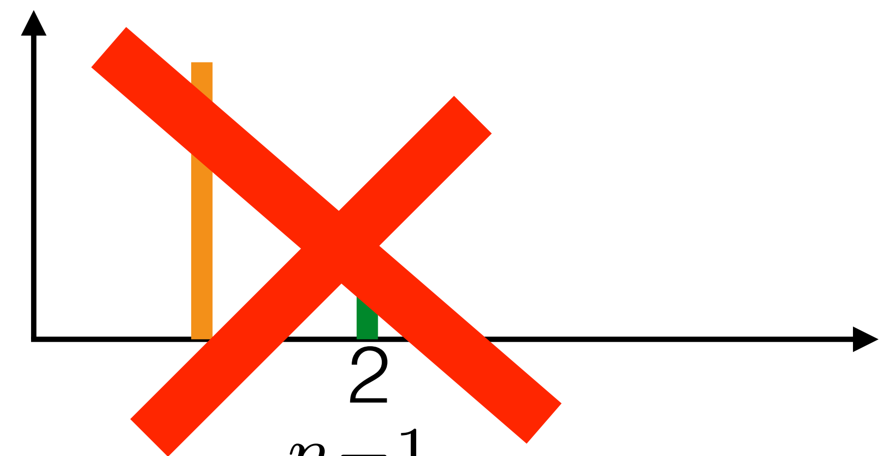
$$\lim_{n \to \infty} \frac{\#\ \text{orange}}{\#\ \text{total}}$$

Marginal cluster assignments

- Integrate out the frequencies

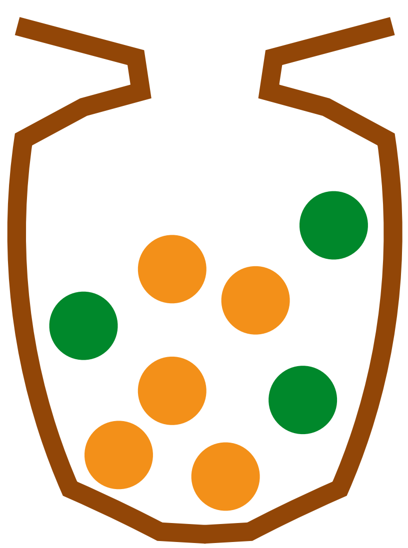
$$\rho_1 \sim \mathrm{Beta}(a_1, a_2),\ z_n \overset{iid}{\sim} \mathrm{Cat}(\rho_1, \rho_2)$$

$$p(z_n = 1 | z_1, \ldots, z_{n-1}) = \frac{a_{1,n}}{a_{1,n} + a_{2,n}}$$



$$a_{1,n} := a_1 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 1\},\ a_{2,n} = a_2 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 2\}$$

- Pólya urn
  - Choose any ball with equal probability
  - Replace and add ball of same color



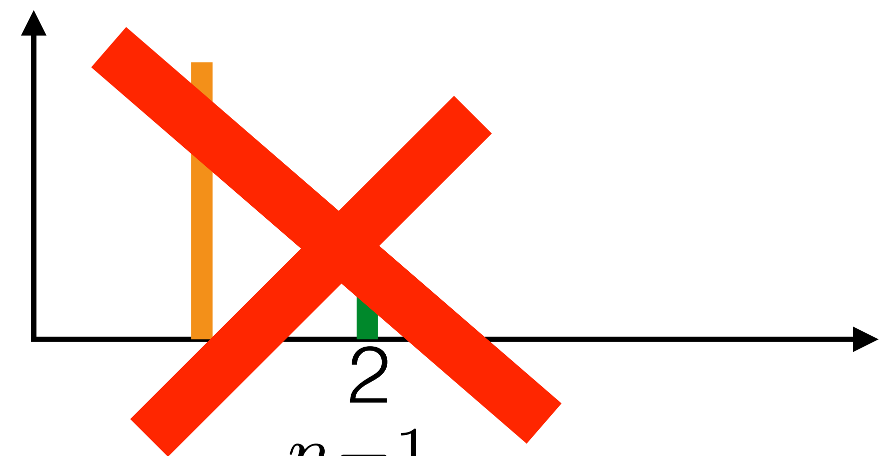
$$\lim_{n \to \infty} \frac{\#\ \text{orange}}{\#\ \text{total}} = \rho_{\text{orange}}$$

Marginal cluster assignments

- Integrate out the frequencies

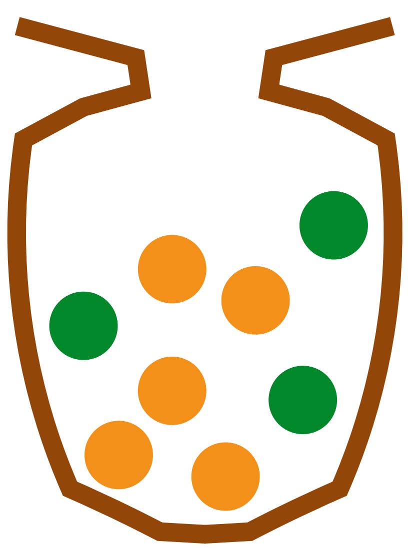
$$\rho_1 \sim \text{Beta}(a_1, a_2),\ z_n \overset{iid}{\sim} \text{Cat}(\rho_1, \rho_2)$$

$$p(z_n = 1 | z_1, \ldots, z_{n-1}) = \frac{a_{1,n}}{a_{1,n} + a_{2,n}}$$



$$a_{1,n} := a_1 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 1\},\ a_{2,n} = a_2 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 2\}$$

- Pólya urn
  - Choose any ball with equal probability
  - Replace and add ball of same color



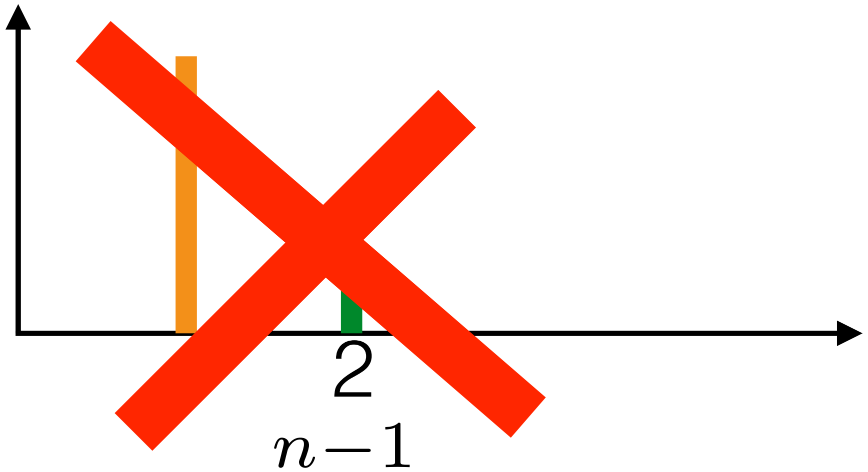
$$\lim_{n \to \infty} \frac{\#\ \text{orange}}{\#\ \text{total}} = \rho_{\text{orange}} \overset{d}{=} \text{Beta}(a_{\text{orange}}, a_{\text{green}})$$

Marginal cluster assignments

- Integrate out the frequencies

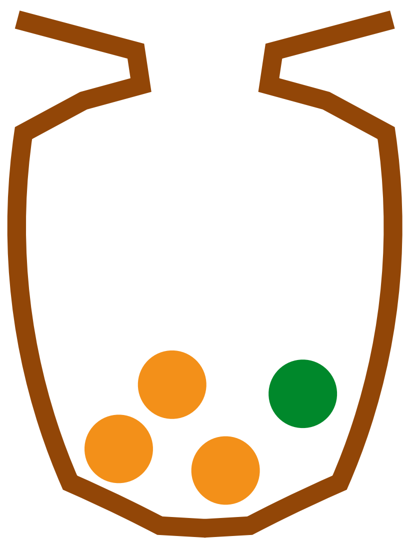
$$\rho_1 \sim \text{Beta}(a_1, a_2),\ z_n \overset{iid}{\sim} \text{Cat}(\rho_1, \rho_2)$$

$$p(z_n = 1 | z_1, \ldots, z_{n-1}) = \frac{a_{1,n}}{a_{1,n} + a_{2,n}}$$



$$a_{1,n} := a_1 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 1\},\ a_{2,n} = a_2 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 2\}$$

- Pólya urn
  - Choose any ball with equal probability
  - Replace and add ball of same color



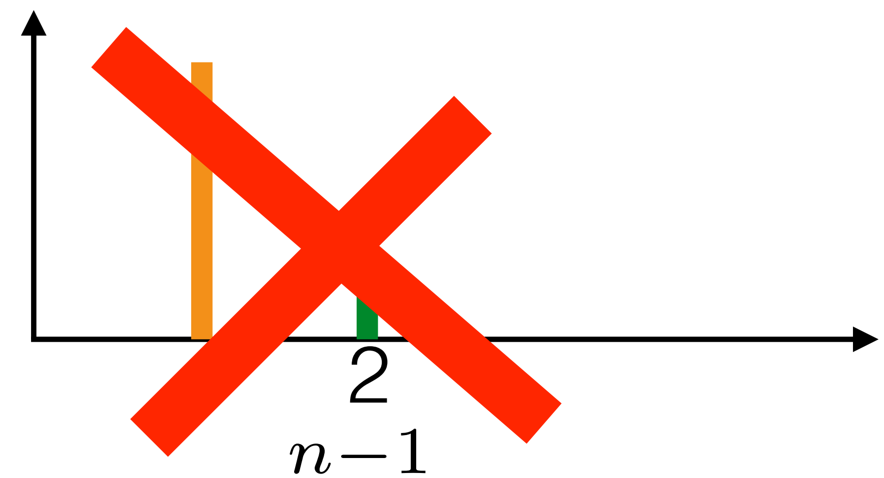
$$\lim_{n \to \infty} \frac{\#\ \text{orange}}{\#\ \text{total}} = \rho_{\text{orange}} \overset{d}{=} \text{Beta}(a_{\text{orange}}, a_{\text{green}})$$

Marginal cluster assignments

- Integrate out the frequencies

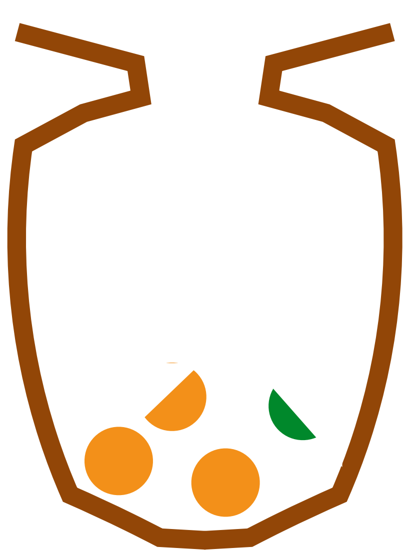
$$\rho_1 \sim \text{Beta}(a_1, a_2),\, z_n \overset{iid}{\sim} \text{Cat}(\rho_1, \rho_2)$$

$$p(z_n = 1 | z_1, \ldots, z_{n-1}) = \frac{a_{1,n}}{a_{1,n} + a_{2,n}}$$



$$a_{1,n} := a_1 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 1\},\, a_{2,n} = a_2 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 2\}$$

- Pólya urn
  - Choose any ball with equal probability
  - Replace and add ball of same color



$$\lim_{n \to \infty} \frac{\#\ \text{orange}}{\#\ \text{total}} = \rho_{\text{orange}} \overset{d}{=} \text{Beta}(a_{\text{orange}}, a_{\text{green}})$$

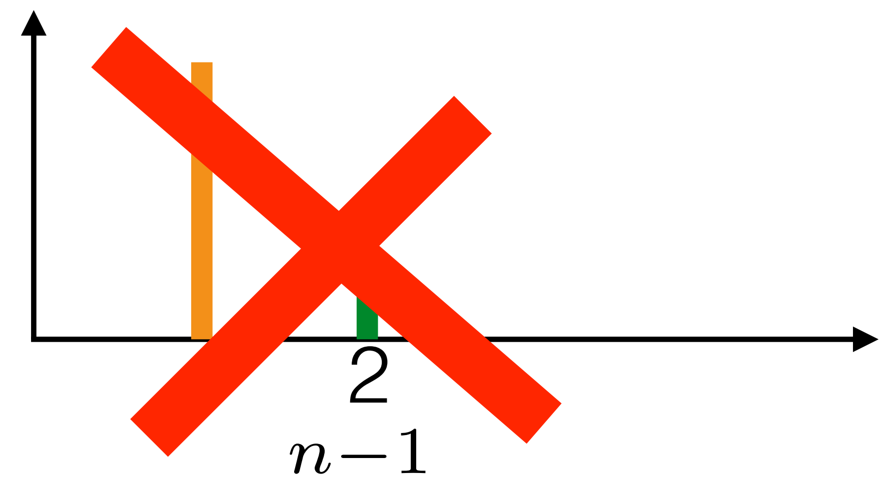
Marginal cluster assignments

- Integrate out the frequencies

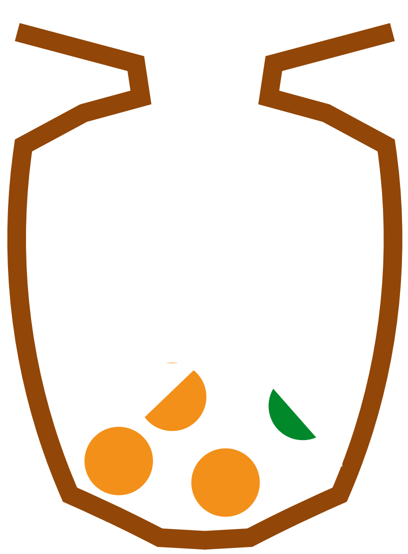
$$\rho_1 \sim \text{Beta}(a_1, a_2),\ z_n \overset{iid}{\sim} \text{Cat}(\rho_1, \rho_2)$$

$$p(z_n = 1 | z_1, \ldots, z_{n-1}) = \frac{a_{1,n}}{a_{1,n} + a_{2,n}}$$

$$a_{1,n} := a_1 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 1\},\ a_{2,n} = a_2 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 2\}$$



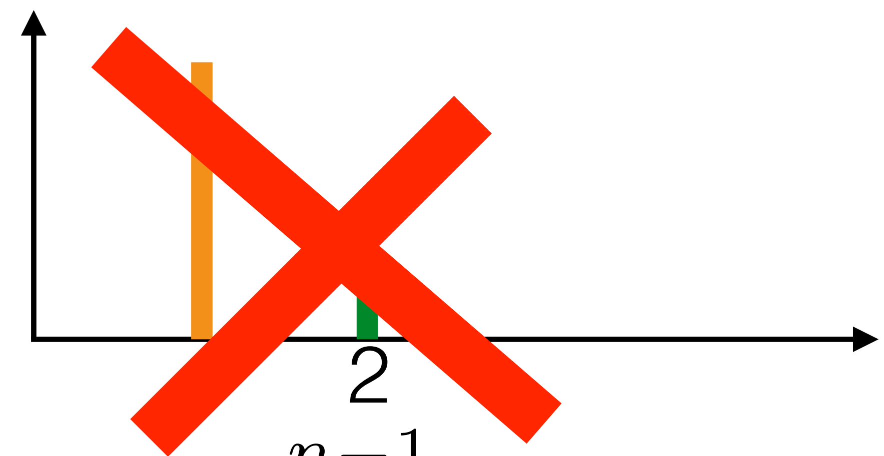
- Pólya urn
  - Choose any ball with prob proportional to its mass
  - Replace and add ball of same color



$$\lim_{n \to \infty} \frac{\#\ \text{orange}}{\#\ \text{total}} = \rho_{\text{orange}} \overset{d}{=} \text{Beta}(a_{\text{orange}}, a_{\text{green}})$$

Marginal cluster assignments

- Integrate out the frequencies

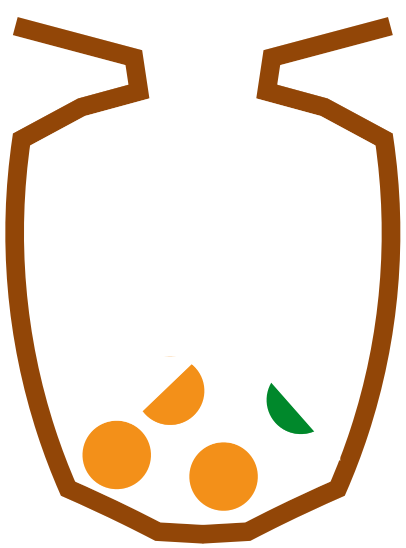


$$\rho_1 \sim \mathrm{Beta}(a_1, a_2), z_n \overset{iid}{\sim} \mathrm{Cat}(\rho_1, \rho_2)$$

$$p(z_n = 1 | z_1, \ldots, z_{n-1}) = \frac{a_{1,n}}{a_{1,n} + a_{2,n}}$$

$$a_{1,n} := a_1 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 1\}, a_{2,n} = a_2 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 2\}$$

- Pólya urn
  - Choose any ball with prob proportional to its mass
  - Replace and add ball of same color

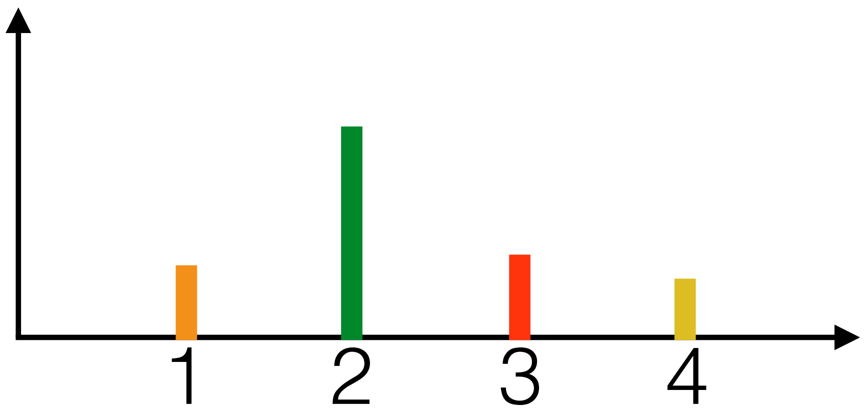


$$\lim_{n \to \infty} \frac{\#\ \text{orange}}{\#\ \text{total}} = \rho_{\text{orange}} \overset{d}{=} \mathrm{Beta}(a_{\text{orange}}, a_{\text{green}})$$

$$\mathrm{PolyaUrn}(a_{\text{orange}}, a_{\text{green}})$$

Marginal cluster assignments

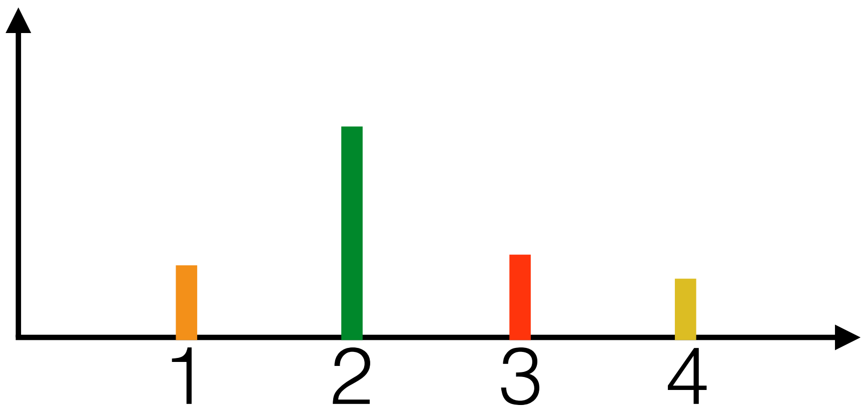
- Integrate out the frequencies



Marginal cluster assignments

- Integrate out the frequencies

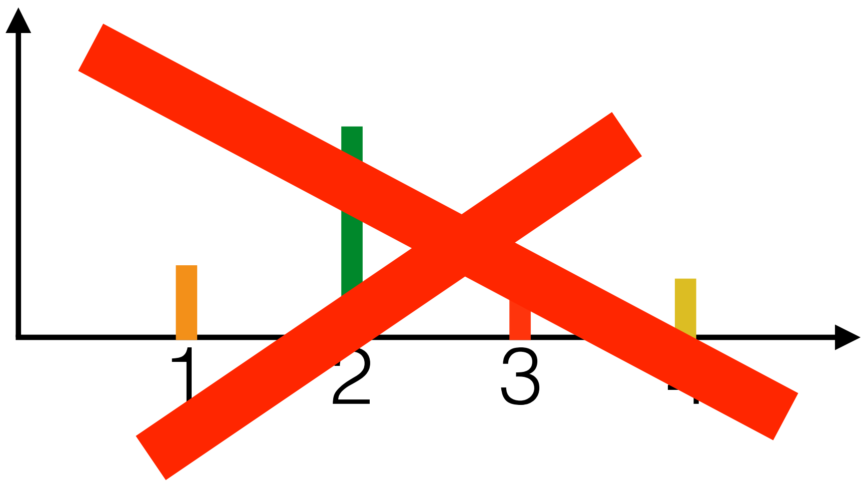
$$\rho_{1:K} \sim \mathrm{Dirichlet}(a_{1:K}), z_n \overset{iid}{\sim} \mathrm{Cat}(\rho_{1:K})$$



Marginal cluster assignments

- Integrate out the frequencies

$$\rho_{1:K} \sim \mathrm{Dirichlet}(a_{1:K}), z_n \stackrel{iid}{\sim} \mathrm{Cat}(\rho_{1:K})$$

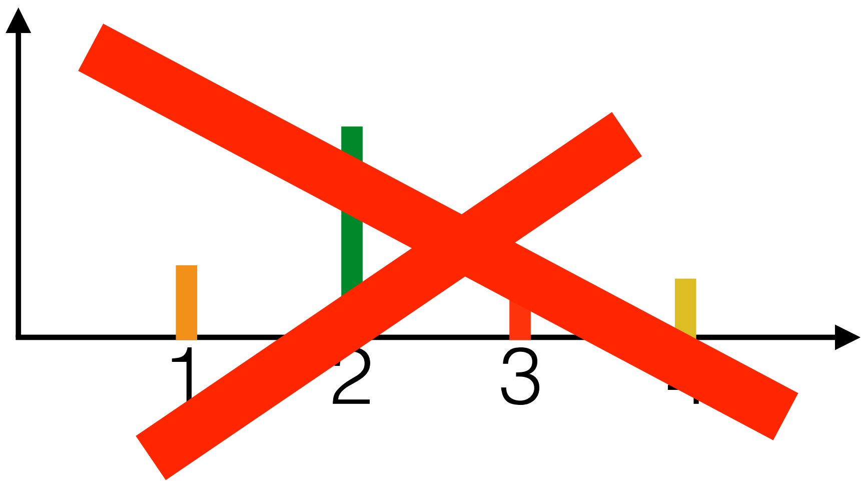


Marginal cluster assignments

- Integrate out the frequencies

$$\rho_{1:K} \sim \text{Dirichlet}(a_{1:K}), z_n \overset{iid}{\sim} \text{Cat}(\rho_{1:K})$$

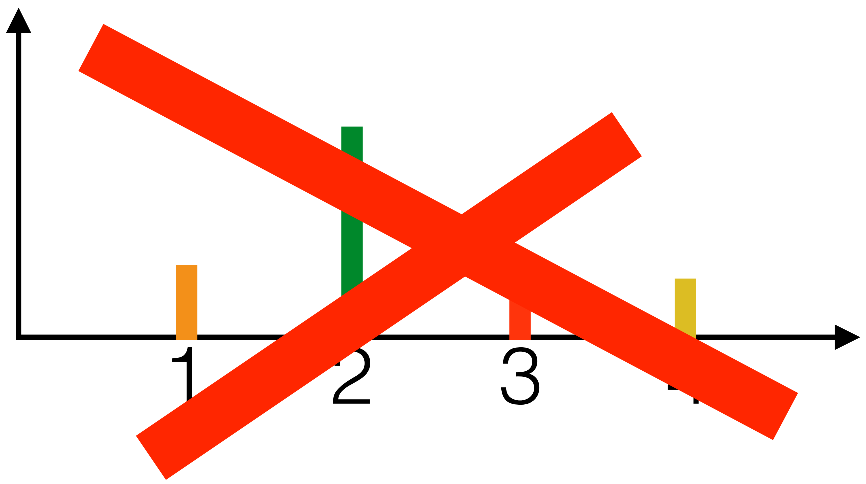
$$p(z_n = k | z_1, \ldots, z_{n-1}) = \frac{a_{k,n}}{\sum_{j=1}^{K} a_{j,n}}$$



Marginal cluster assignments

- Integrate out the frequencies

$$\rho_{1:K} \sim \mathrm{Dirichlet}(a_{1:K}), z_n \overset{iid}{\sim} \mathrm{Cat}(\rho_{1:K})$$



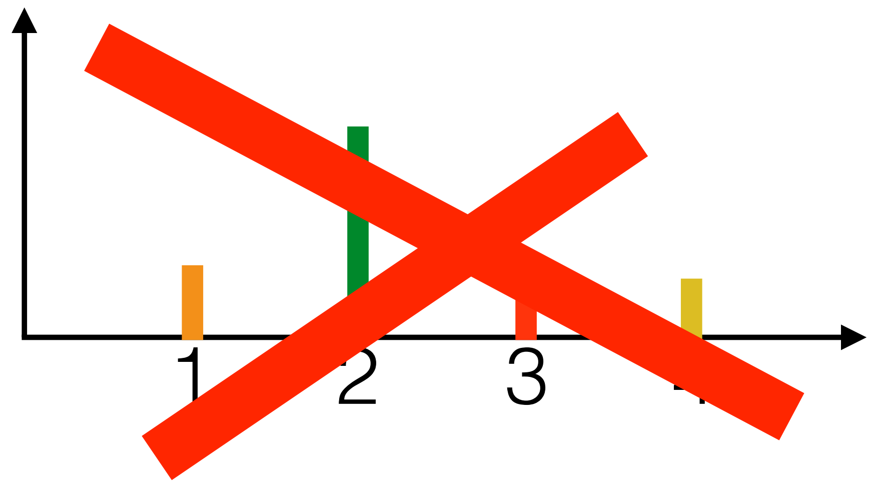
$$p(z_n = k | z_1, \ldots, z_{n-1}) = \frac{a_{k,n}}{\sum_{j=1}^{K} a_{j,n}}$$

$$a_{k,n} := a_k + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = k\}$$

Marginal cluster assignments

- Integrate out the frequencies

$$\rho_{1:K} \sim \mathrm{Dirichlet}(a_{1:K}), z_n \stackrel{iid}{\sim} \mathrm{Cat}(\rho_{1:K})$$



$$p(z_n = k | z_1, \ldots, z_{n-1}) = \frac{a_{k,n}}{\sum_{j=1}^{K} a_{j,n}}$$

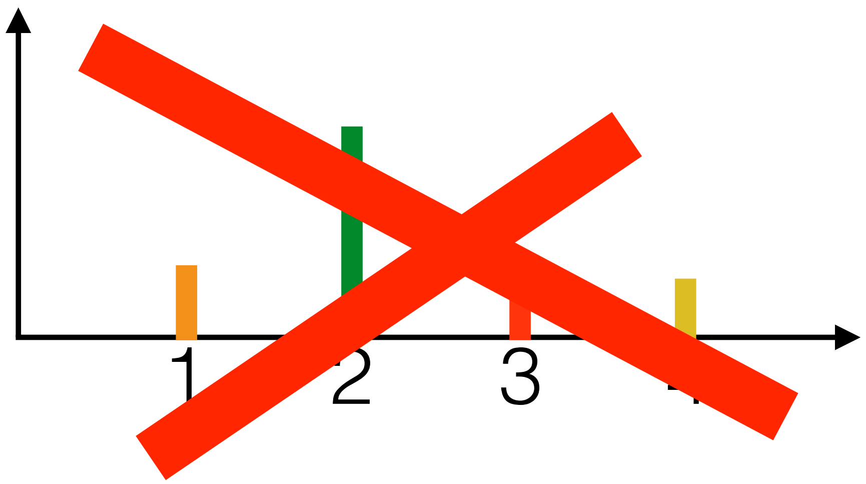
$$a_{k,n} := a_k + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = k\}$$

- multivariate Pólya urn

Marginal cluster assignments

- Integrate out the frequencies

$$\rho_{1:K} \sim \mathrm{Dirichlet}(a_{1:K}), z_n \overset{iid}{\sim} \mathrm{Cat}(\rho_{1:K})$$



$$p(z_n = k | z_1, \ldots, z_{n-1}) = \frac{a_{k,n}}{\sum_{j=1}^{K} a_{j,n}}$$

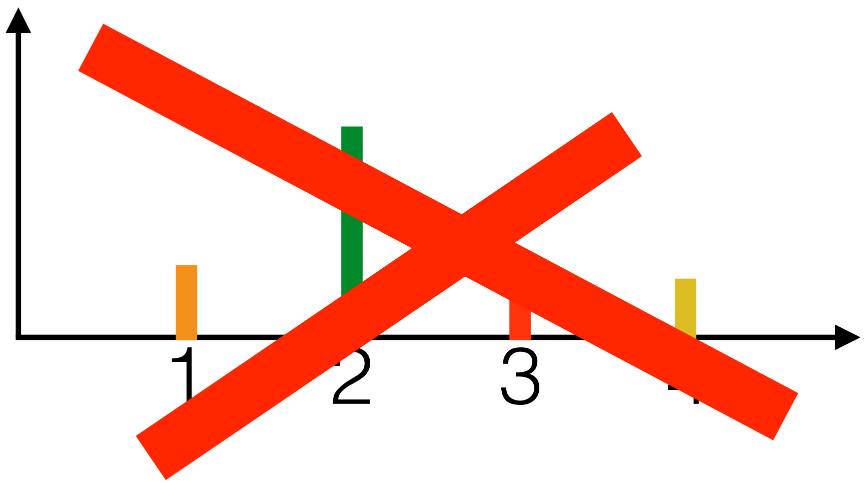
$$a_{k,n} := a_k + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = k\}$$

- multivariate Pólya urn

Marginal cluster assignments

- Integrate out the frequencies

$$\rho_{1:K} \sim \mathrm{Dirichlet}(a_{1:K}),\ z_n \overset{iid}{\sim} \mathrm{Cat}(\rho_{1:K})$$



$$p(z_n = k | z_1, \ldots, z_{n-1}) = \frac{a_{k,n}}{\sum_{j=1}^{K} a_{j,n}}$$

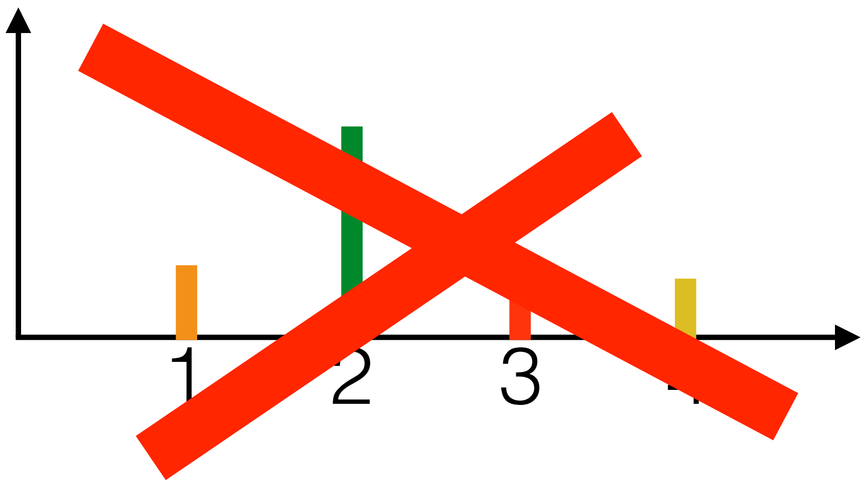
$$a_{k,n} := a_k + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = k\}$$

- multivariate Pólya urn
  - Choose any ball with prob proportional to its mass

Marginal cluster assignments

- Integrate out the frequencies

$$\rho_{1:K} \sim \mathrm{Dirichlet}(a_{1:K}),\, z_n \overset{iid}{\sim} \mathrm{Cat}(\rho_{1:K})$$



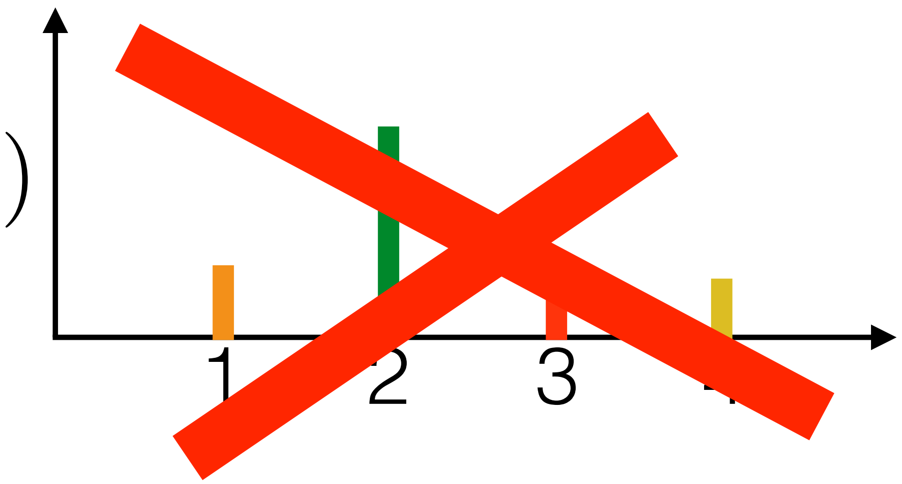
$$p(z_n = k | z_1, \ldots, z_{n-1}) = \frac{a_{k,n}}{\sum_{j=1}^{K} a_{j,n}}$$

$$a_{k,n} := a_k + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = k\}$$

- multivariate Pólya urn
  - Choose any ball with prob proportional to its mass
  - Replace and add ball of same color

Marginal cluster assignments

- Integrate out the frequencies



$$\rho_{1:K} \sim \text{Dirichlet}(a_{1:K}), z_n \stackrel{iid}{\sim} \text{Cat}(\rho_{1:K})$$

$$p(z_n = k | z_1, \ldots, z_{n-1}) = \frac{a_{k,n}}{\sum_{j=1}^{K} a_{j,n}}$$

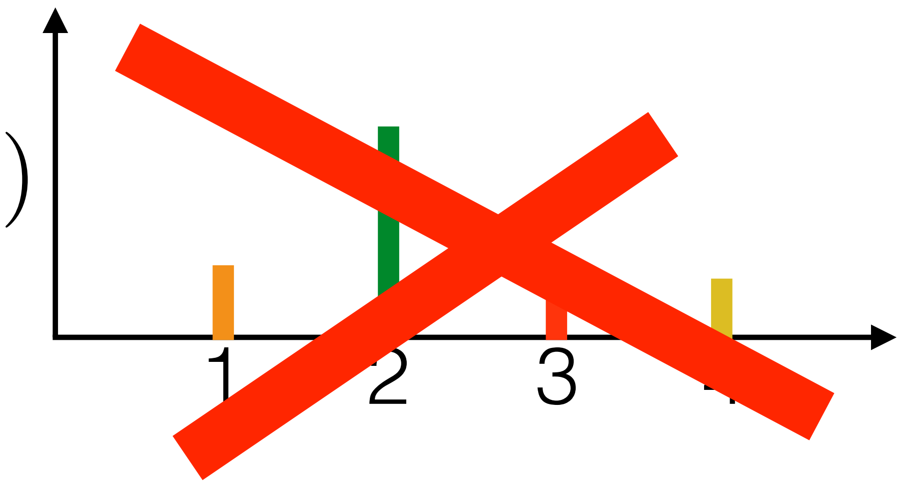
$$a_{k,n} := a_k + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = k\}$$

- multivariate Pólya urn
  - Choose any ball with prob proportional to its mass
  - Replace and add ball of same color

$$\lim_{n \to \infty} \frac{(\#\ \text{orange},\ \#\ \text{green},\ \#\ \text{red},\ \#\ \text{yellow})}{\#\ \text{total}}$$

Marginal cluster assignments

- Integrate out the frequencies



$$\rho_{1:K} \sim \text{Dirichlet}(a_{1:K}),\, z_n \stackrel{iid}{\sim} \text{Cat}(\rho_{1:K})$$

$$p(z_n = k | z_1, \ldots, z_{n-1}) = \frac{a_{k,n}}{\sum_{j=1}^{K} a_{j,n}}$$

$$a_{k,n} := a_k + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = k\}$$

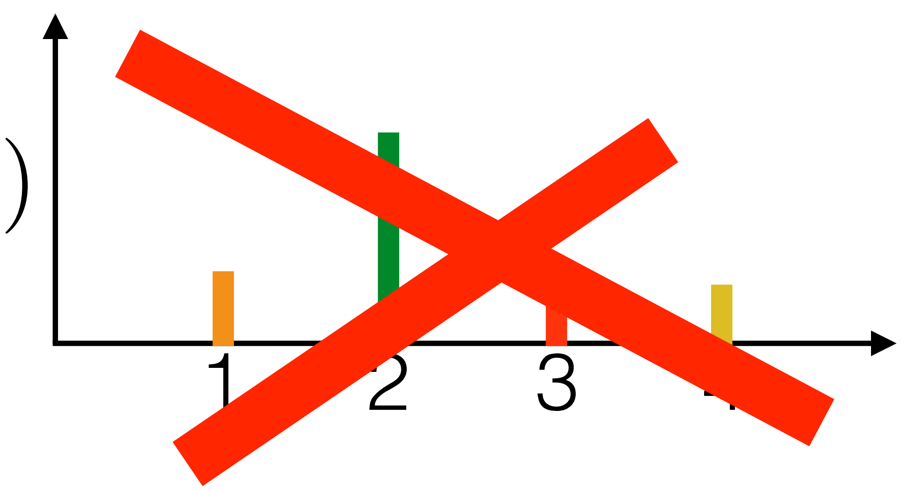
- multivariate Pólya urn
    - Choose any ball with prob proportional to its mass
    - Replace and add ball of same color

$$\lim_{n \to \infty} \frac{(\#\text{ orange}, \#\text{ green}, \#\text{ red}, \#\text{ yellow})}{\#\text{ total}}$$

$$\to (\rho_{\text{orange}}, \rho_{\text{green}}, \rho_{\text{red}}, \rho_{\text{yellow}})$$

Marginal cluster assignments

- Integrate out the frequencies

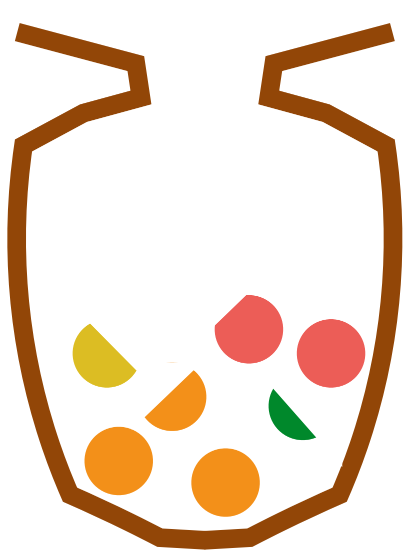


$$\rho_{1:K} \sim \text{Dirichlet}(a_{1:K}),\, z_n \overset{iid}{\sim} \text{Cat}(\rho_{1:K})$$

$$p(z_n = k | z_1, \ldots, z_{n-1}) = \frac{a_{k,n}}{\sum_{j=1}^{K} a_{j,n}}$$

$$a_{k,n} := a_k + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = k\}$$

- multivariate Pólya urn
  - Choose any ball with prob proportional to its mass
  - Replace and add ball of same color



$$\lim_{n\to\infty} \frac{(\#\text{ orange}, \#\text{ green}, \#\text{ red}, \#\text{ yellow})}{\#\text{ total}}$$

$$\rightarrow (\rho_{\text{orange}}, \rho_{\text{green}}, \rho_{\text{red}}, \rho_{\text{yellow}})$$

$$\overset{d}{=} \text{Dirichlet}(a_{\text{orange}}, a_{\text{green}}, a_{\text{red}}, a_{\text{yellow}})$$

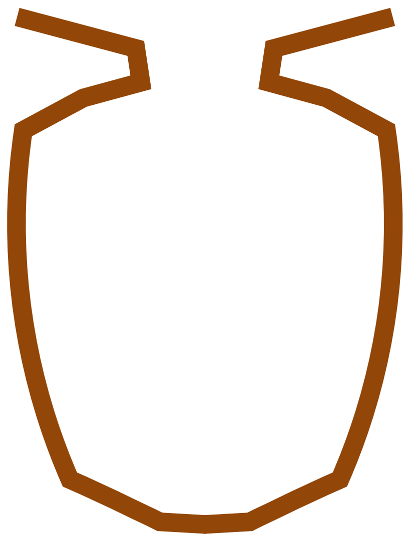
Marginal cluster assignments

- Hoppe urn / Blackwell-MacQueen urn

[Blackwell, MacQueen 1973; Hoppe 1984]

Marginal cluster assignments

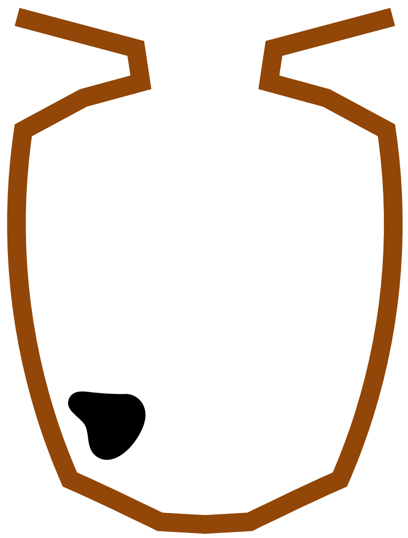
- Hoppe urn / Blackwell-MacQueen urn



[Blackwell, MacQueen 1973; Hoppe 1984]

Marginal cluster assignments

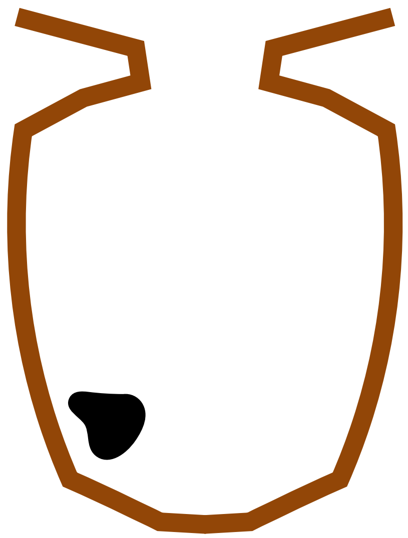
- Hoppe urn / Blackwell-MacQueen urn



[Blackwell, MacQueen 1973; Hoppe 1984]

Marginal cluster assignments

- Hoppe urn / Blackwell-MacQueen urn

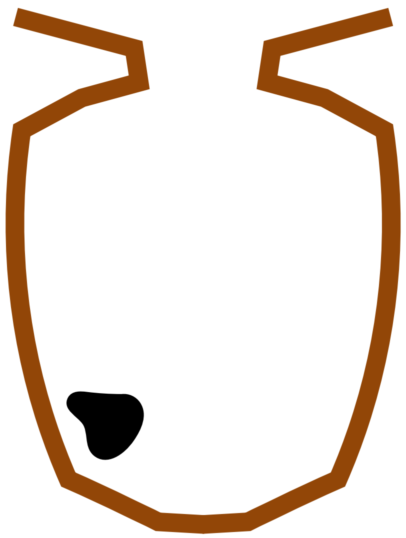


  - Choose ball with prob proportional to its mass

[Blackwell, MacQueen 1973; Hoppe 1984]

Marginal cluster assignments

- Hoppe urn / Blackwell-MacQueen urn

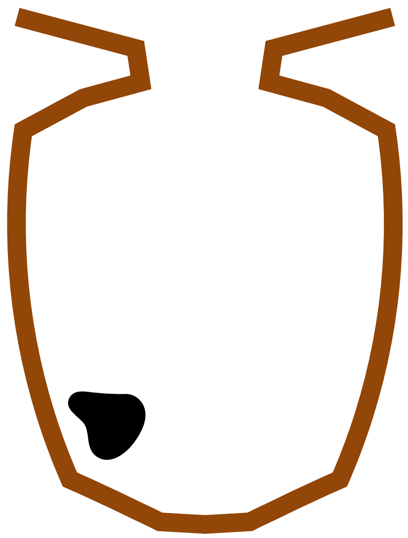


- Choose ball with prob proportional to its mass
  - If black, replace and add ball of new color

[Blackwell, MacQueen 1973; Hoppe 1984]

Marginal cluster assignments

- Hoppe urn / Blackwell-MacQueen urn

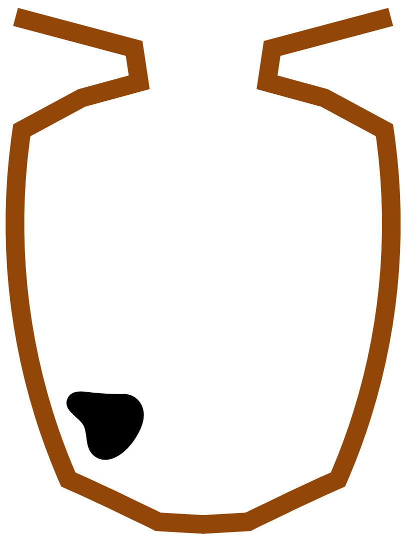


- Choose ball with prob proportional to its mass
    - If black, replace and add ball of new color
    - Else, replace and add ball of same color

[Blackwell, MacQueen 1973; Hoppe 1984]

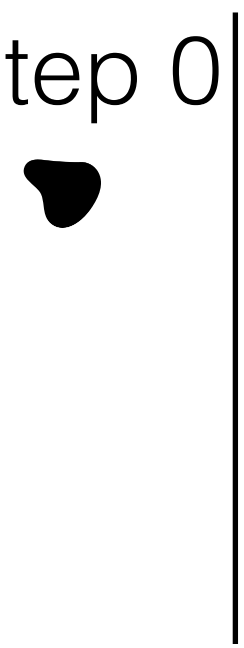
Marginal cluster assignments

- Hoppe urn / Blackwell-MacQueen urn



- Choose ball with prob proportional to its mass
  - If black, replace and add ball of new color
  - Else, replace and add ball of same color

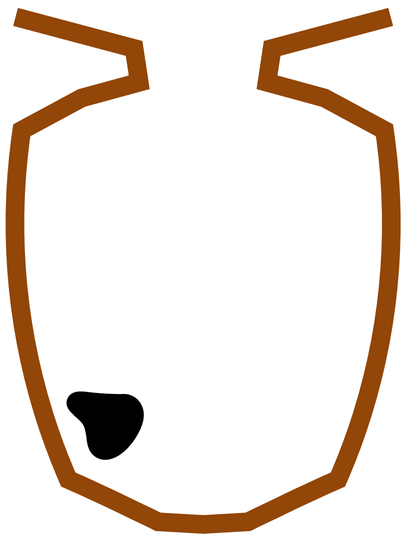
Step 0



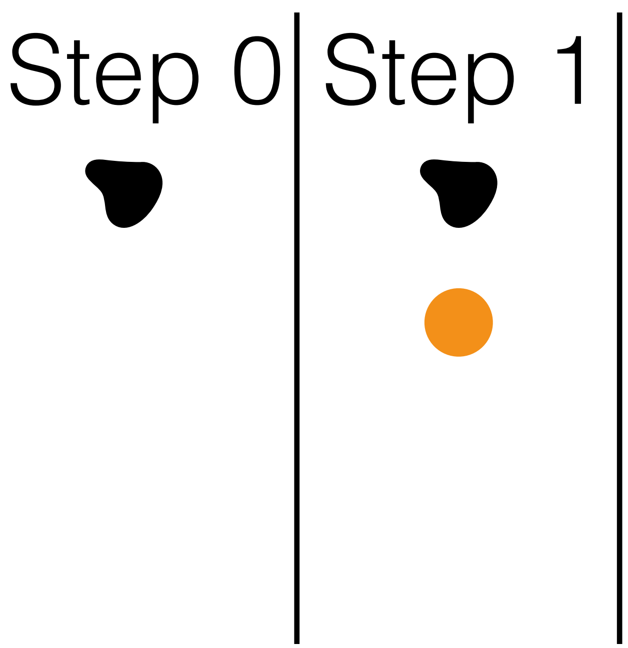
[Blackwell, MacQueen 1973; Hoppe 1984]

Marginal cluster assignments

- Hoppe urn / Blackwell-MacQueen urn



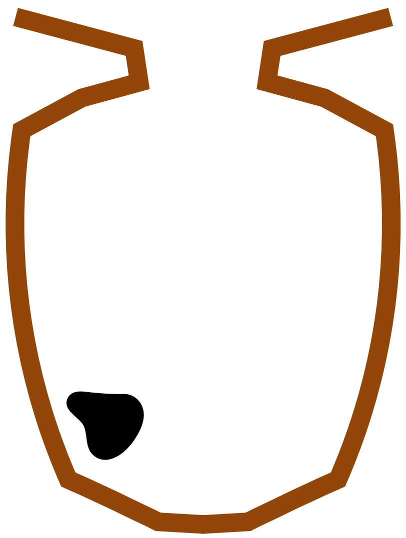
  - Choose ball with prob proportional to its mass
    - If black, replace and add ball of new color
    - Else, replace and add ball of same color



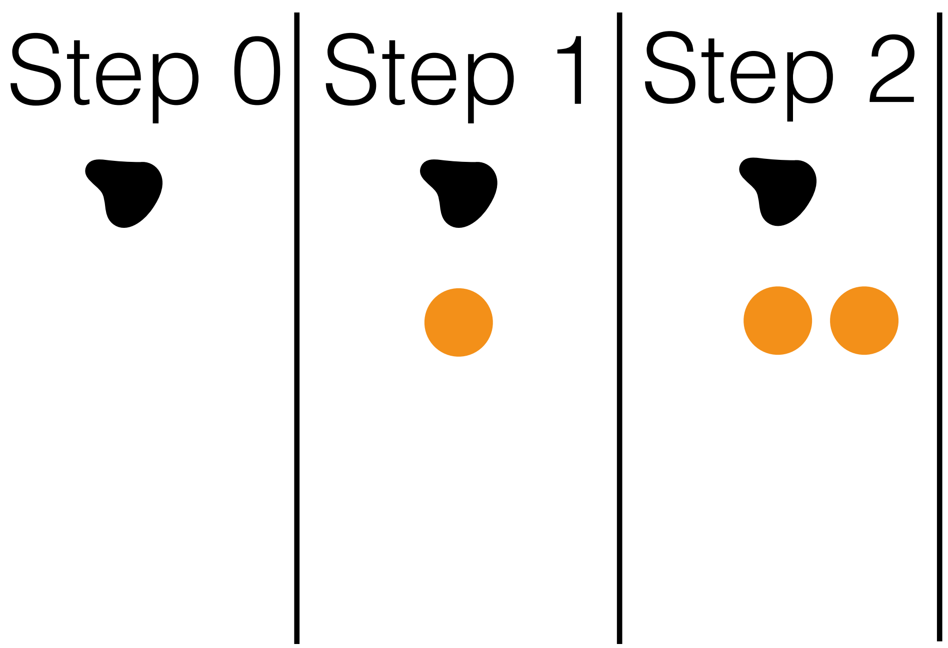
[Blackwell, MacQueen 1973; Hoppe 1984]

Marginal cluster assignments

- Hoppe urn / Blackwell-MacQueen urn



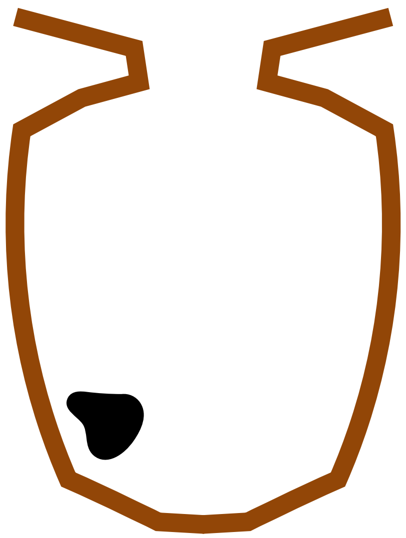
- Choose ball with prob proportional to its mass
  - If black, replace and add ball of new color
  - Else, replace and add ball of same color



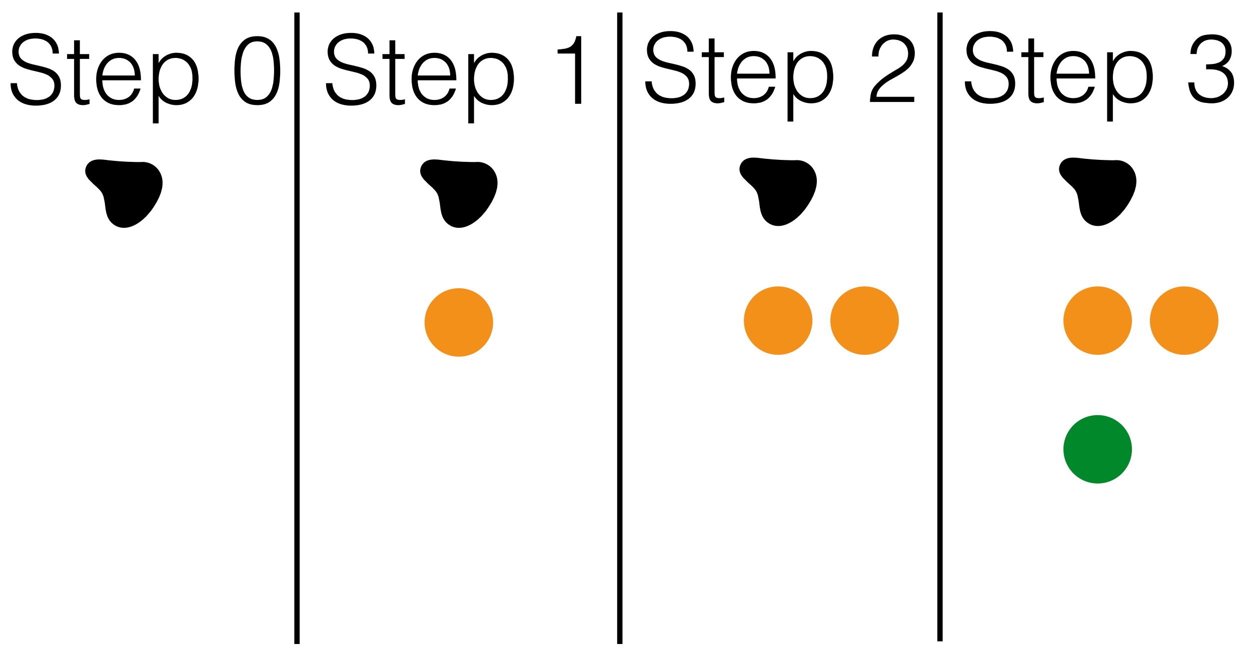
[Blackwell, MacQueen 1973; Hoppe 1984]

Marginal cluster assignments

- Hoppe urn / Blackwell-MacQueen urn



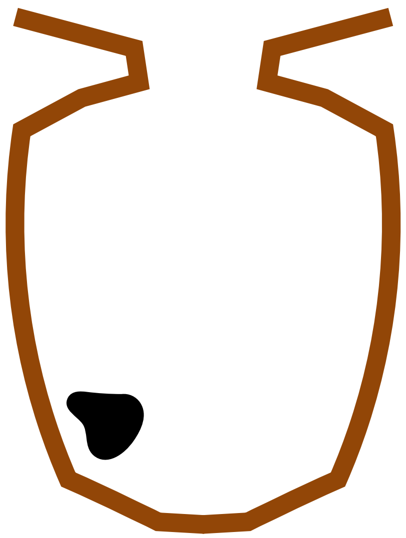
- Choose ball with prob proportional to its mass
  - If black, replace and add ball of new color
  - Else, replace and add ball of same color



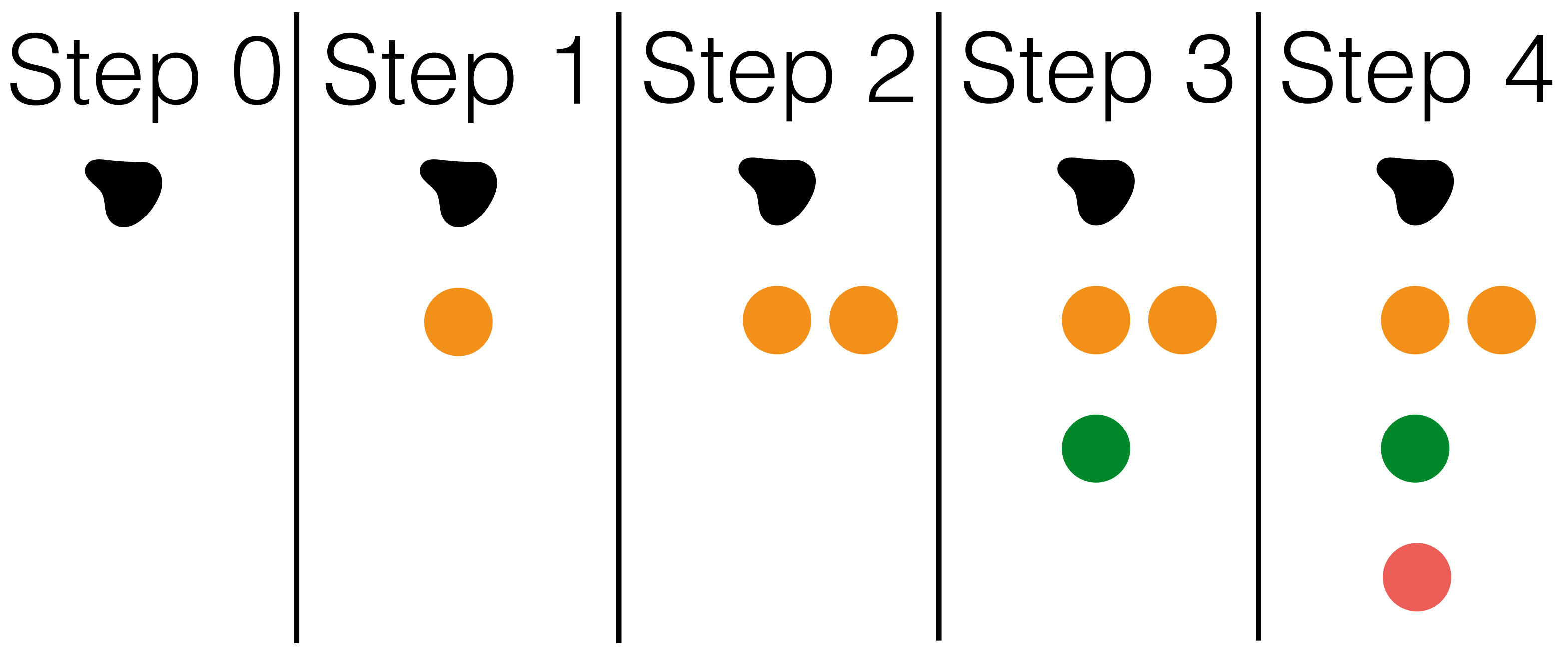
[Blackwell, MacQueen 1973; Hoppe 1984]

Marginal cluster assignments

- Hoppe urn / Blackwell-MacQueen urn



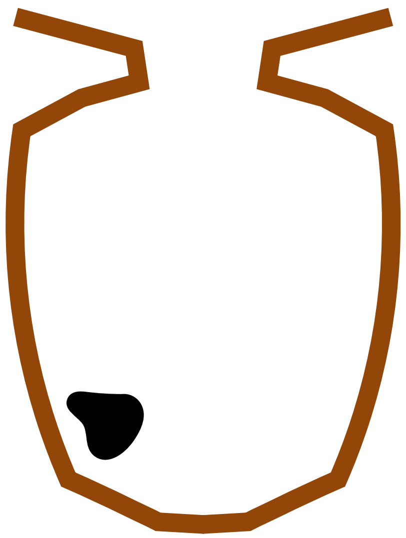
  - Choose ball with prob proportional to its mass
    - If black, replace and add ball of new color
    - Else, replace and add ball of same color



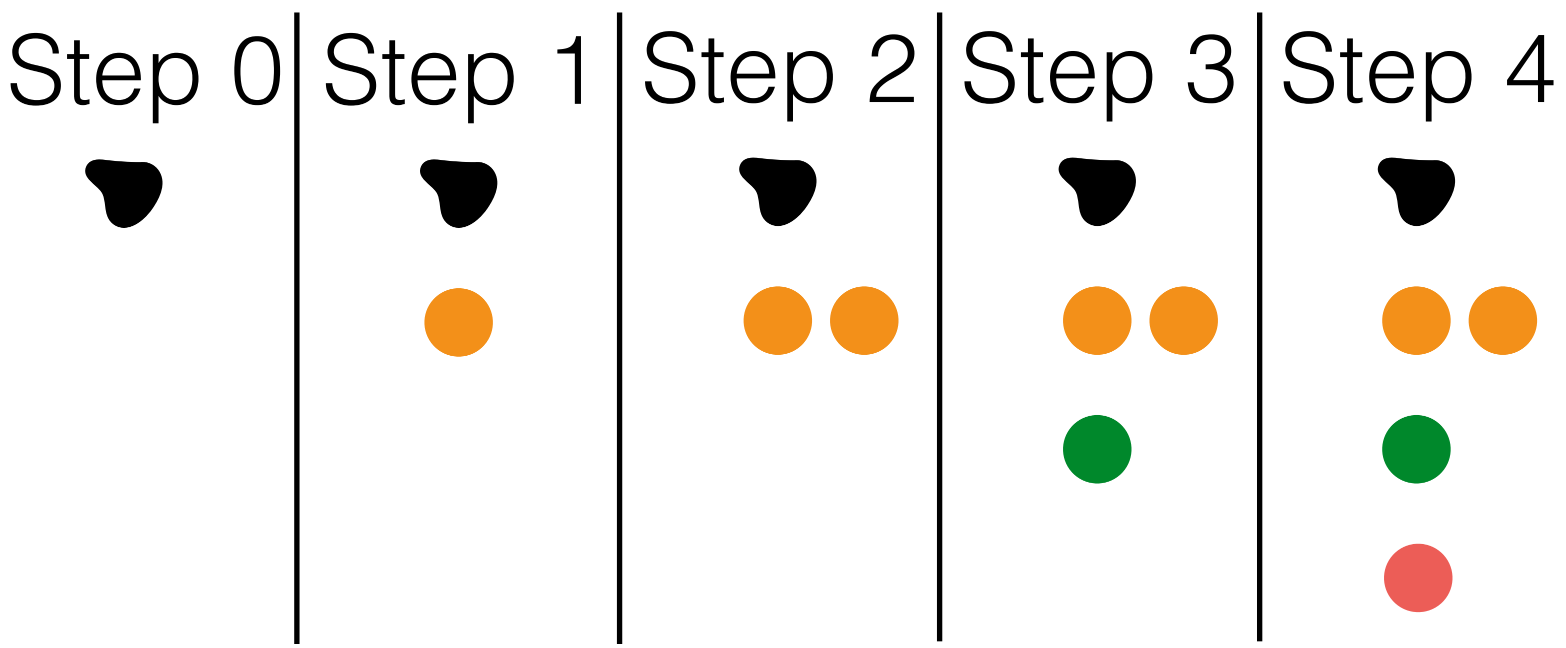
[Blackwell, MacQueen 1973; Hoppe 1984]

Marginal cluster assignments

- Hoppe urn / Blackwell-MacQueen urn



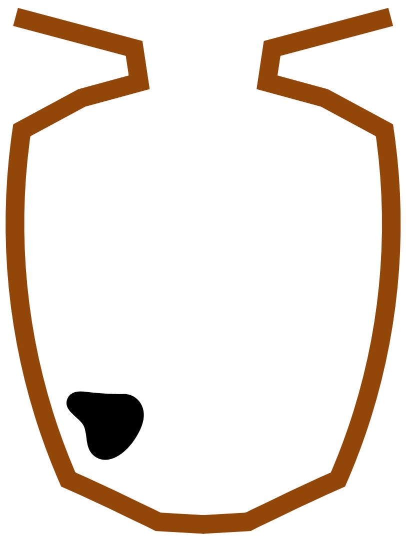
  - Choose ball with prob proportional to its mass
    - If black, replace and add ball of new color
    - Else, replace and add ball of same color



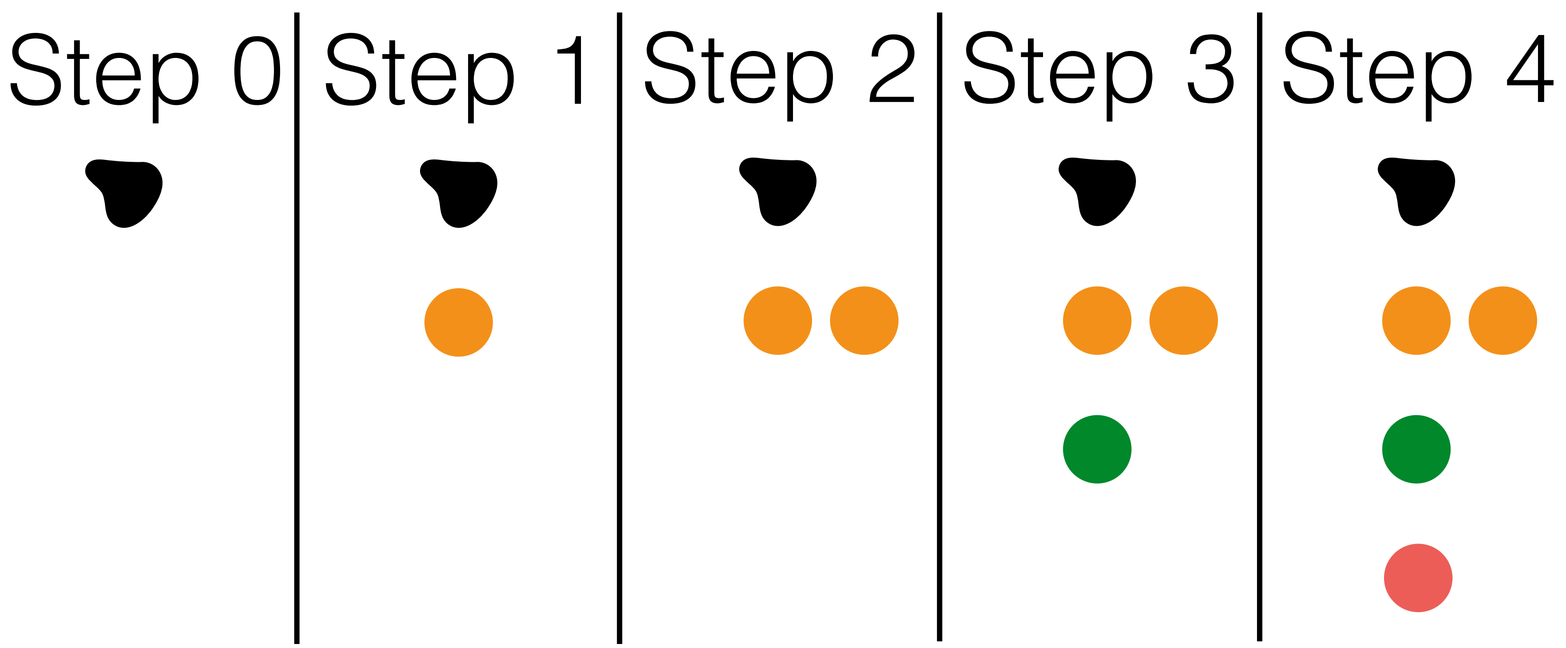
$$(\#\text{orange}, \#\text{other}) =$$

Marginal cluster assignments

- Hoppe urn / Blackwell-MacQueen urn



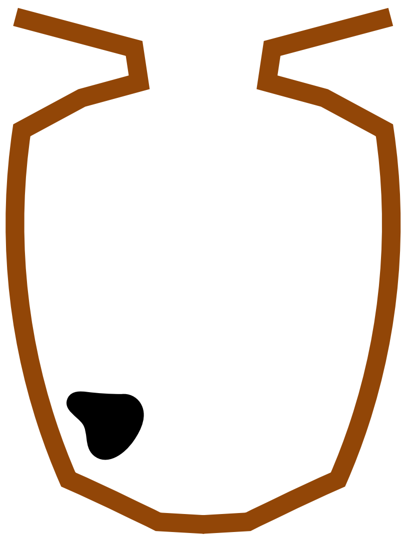
  - Choose ball with prob proportional to its mass
    - If black, replace and add ball of new color
    - Else, replace and add ball of same color



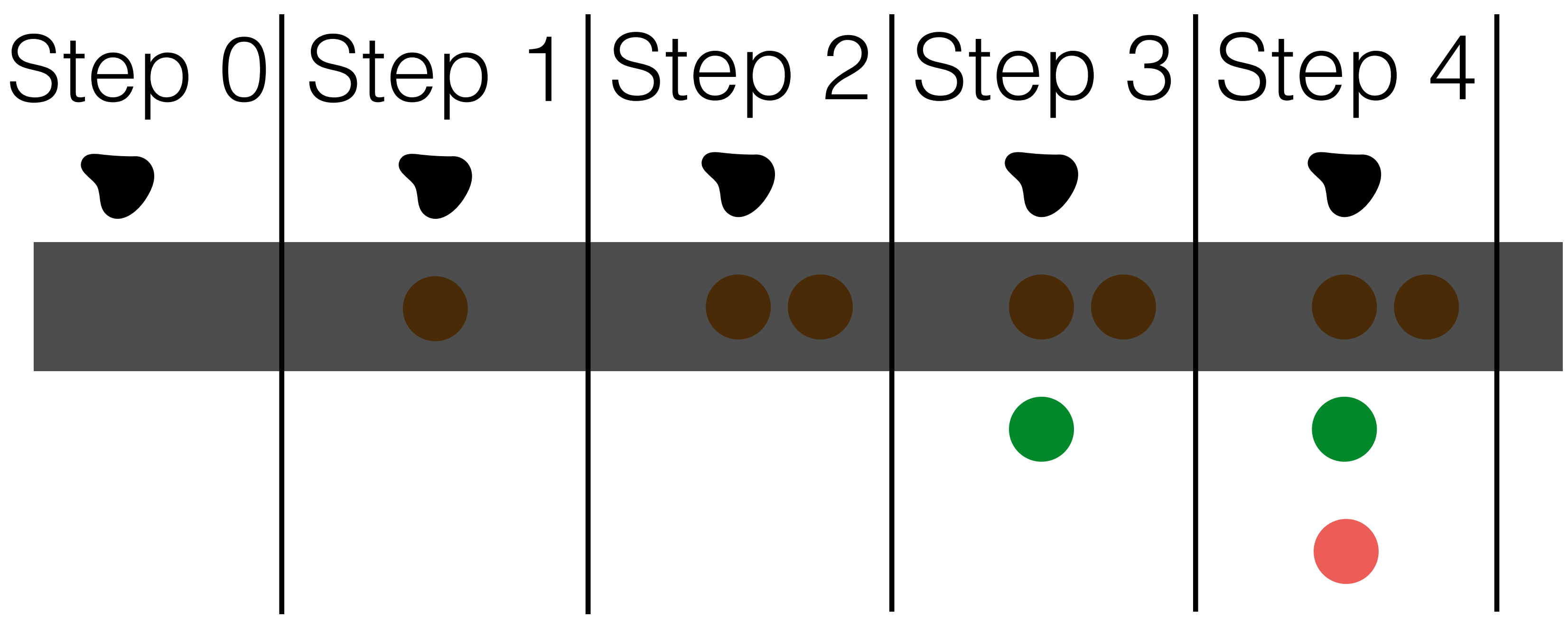
$$(\#\text{orange}, \#\text{other}) = \text{PolyaUrn}(1, \alpha)$$

Marginal cluster assignments

- Hoppe urn / Blackwell-MacQueen urn



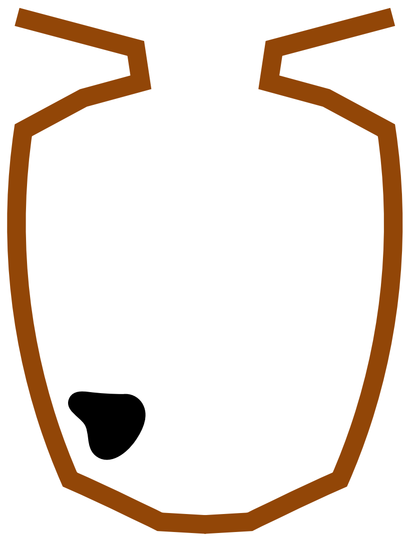
  - Choose ball with prob proportional to its mass
    - If black, replace and add ball of new color
    - Else, replace and add ball of same color



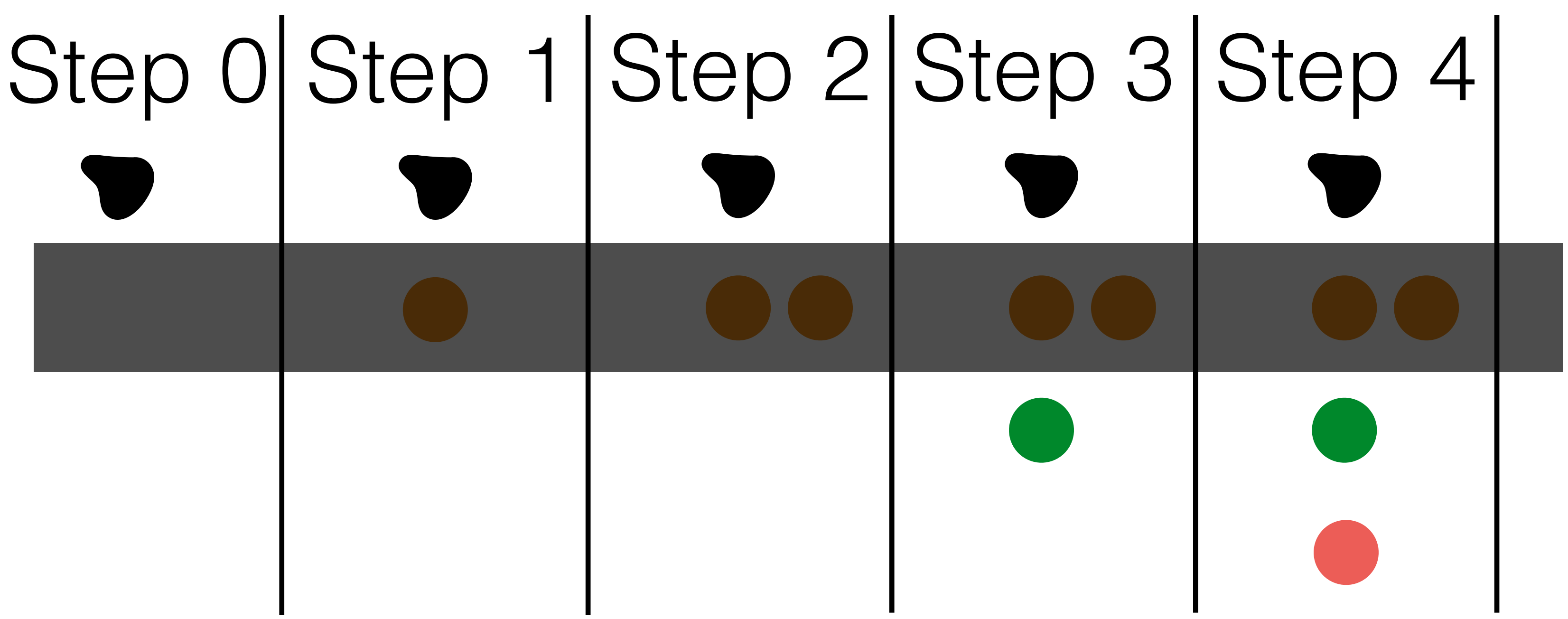
$$(\#\text{orange}, \#\text{other}) = \text{PolyaUrn}(1, \alpha)$$

Marginal cluster assignments

- Hoppe urn / Blackwell-MacQueen urn



- Choose ball with prob proportional to its mass
  - If black, replace and add ball of new color
  - Else, replace and add ball of same color

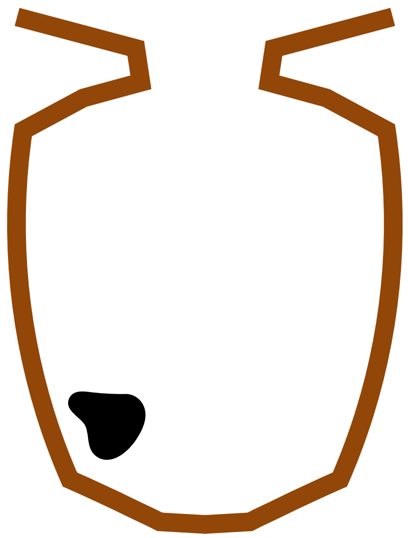


$(\#\text{orange}, \#\text{other}) = \text{PolyaUrn}(1, \alpha)$

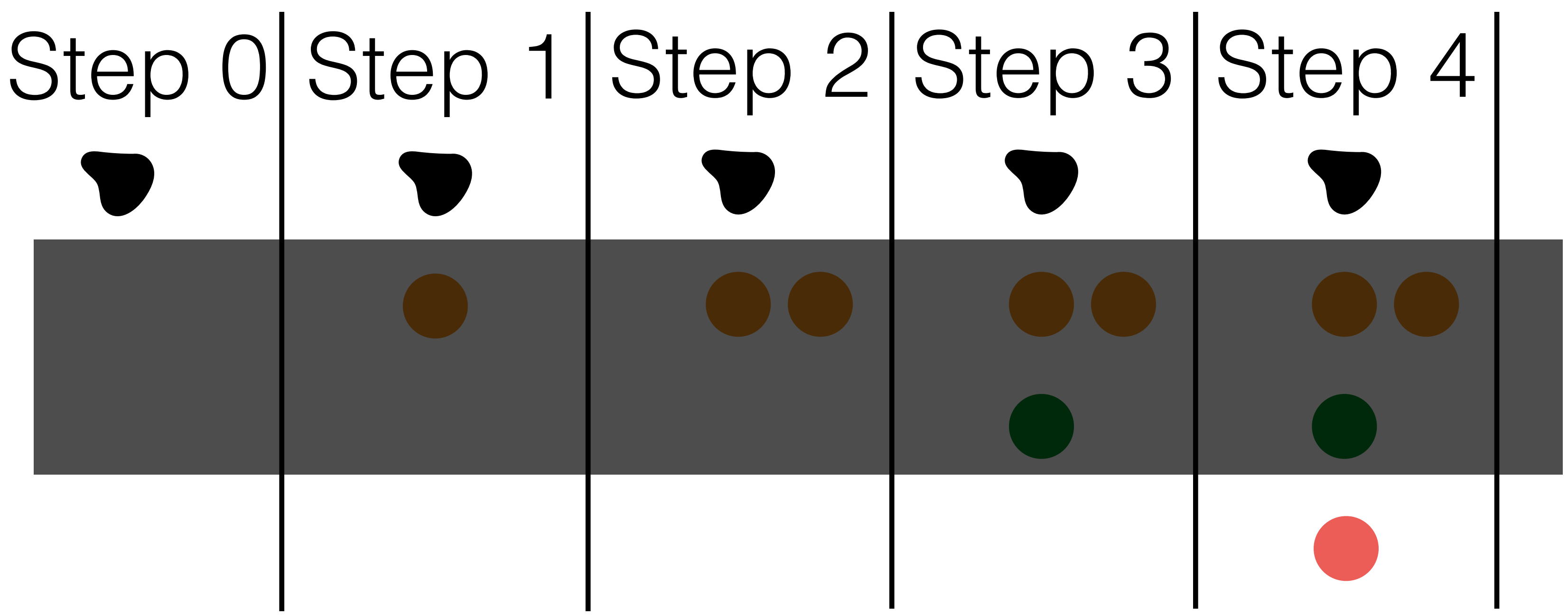
- not orange: $(\#\text{green}, \#\text{other}) = \text{PolyaUrn}(1, \alpha)$

Marginal cluster assignments

- Hoppe urn / Blackwell-MacQueen urn



- Choose ball with prob proportional to its mass
  - If black, replace and add ball of new color
  - Else, replace and add ball of same color

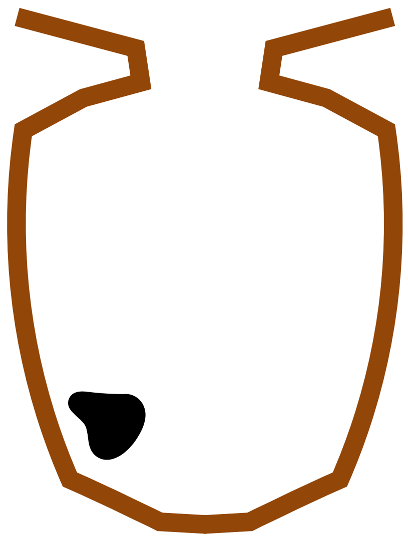


$$(\#\mathrm{orange}, \#\mathrm{other}) = \mathrm{PolyaUrn}(1, \alpha)$$

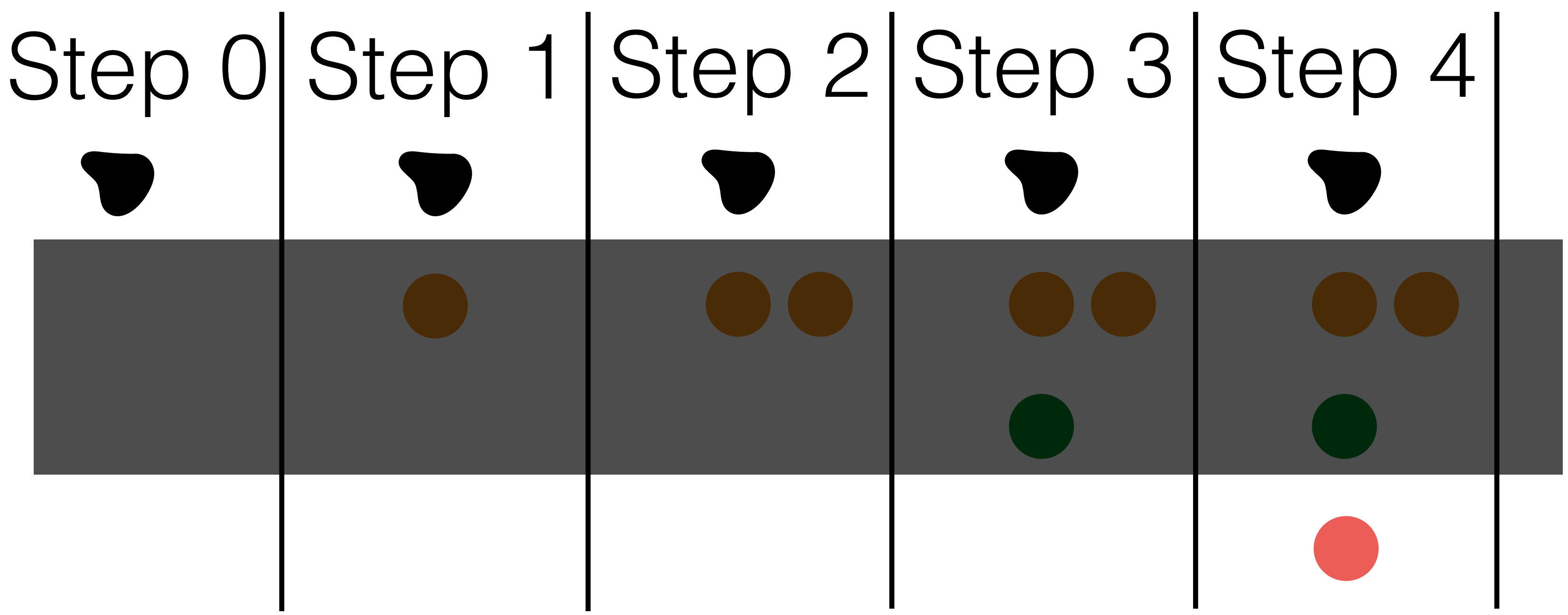
- not orange: $(\#\mathrm{green}, \#\mathrm{other}) = \mathrm{PolyaUrn}(1, \alpha)$

Marginal cluster assignments

- Hoppe urn / Blackwell-MacQueen urn



  - Choose ball with prob proportional to its mass
    - If black, replace and add ball of new color
    - Else, replace and add ball of same color

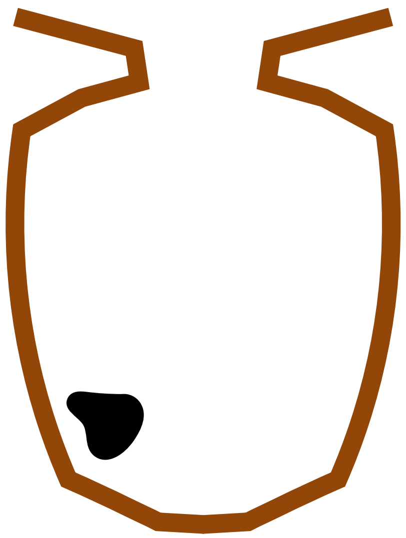


$$(\#\text{orange}, \#\text{other}) = \text{PolyaUrn}(1, \alpha)$$

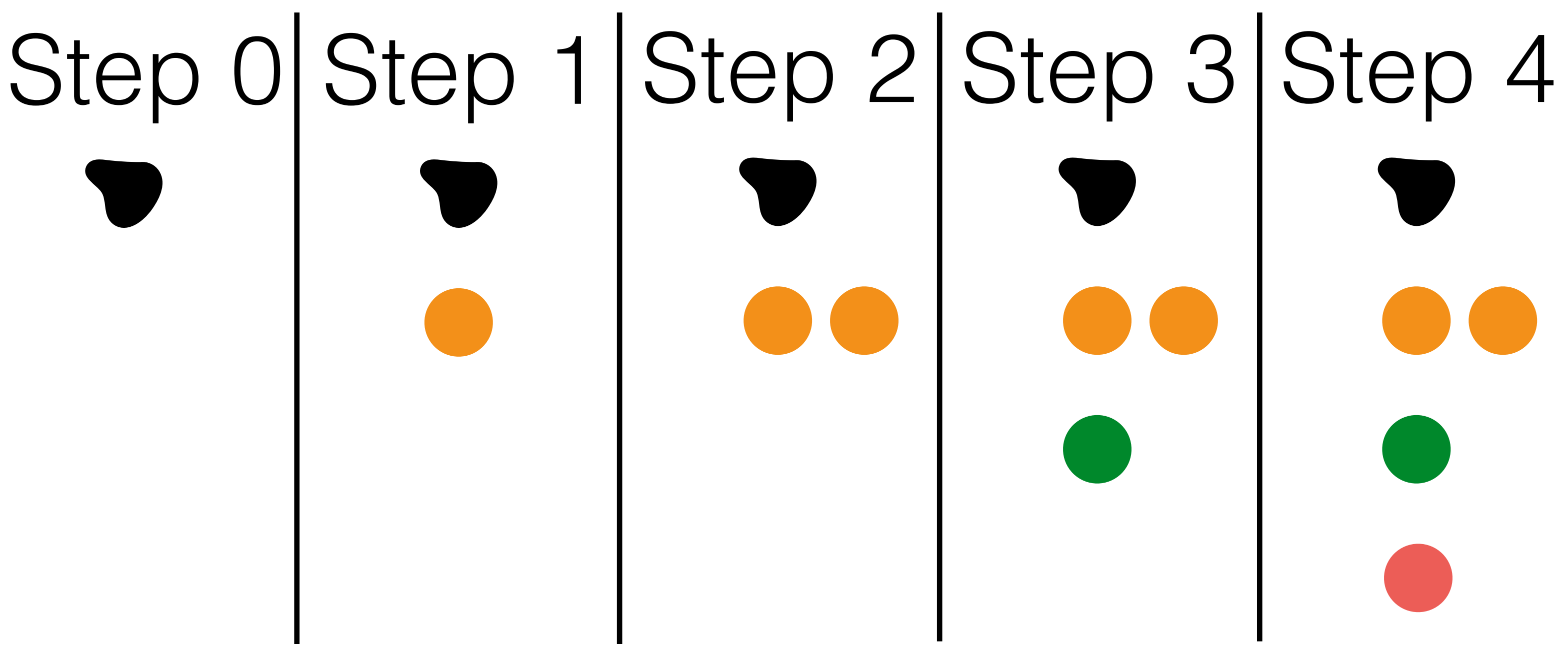
- not orange: $(\#\text{green}, \#\text{other}) = \text{PolyaUrn}(1, \alpha)$
- not orange, green: $(\#\text{red}, \#\text{other}) = \text{PolyaUrn}(1, \alpha)$

Marginal cluster assignments

- Hoppe urn / Blackwell-MacQueen urn



  - Choose ball with prob proportional to its mass
    - If black, replace and add ball of new color
    - Else, replace and add ball of same color

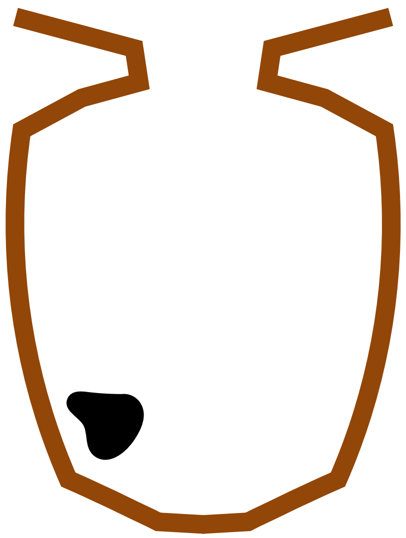


$(\#\text{orange}, \#\text{other}) = \text{PolyaUrn}(1, \alpha)$

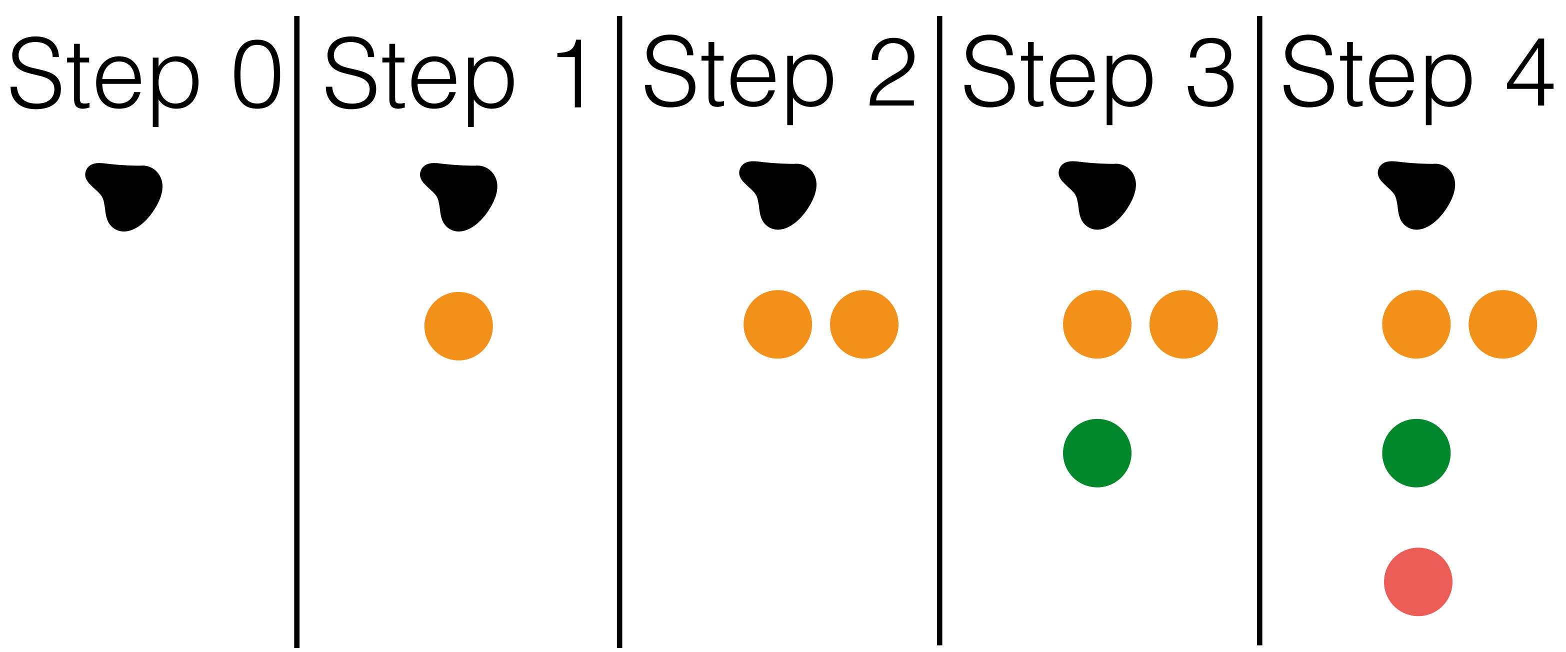
- not orange: $(\#\text{green}, \#\text{other}) = \text{PolyaUrn}(1, \alpha)$
- not orange, green: $(\#\text{red}, \#\text{other}) = \text{PolyaUrn}(1, \alpha)$

Marginal cluster assignments

- Hoppe urn / Blackwell-MacQueen urn



- Choose ball with prob proportional to its mass
  - If black, replace and add ball of new color
  - Else, replace and add ball of same color



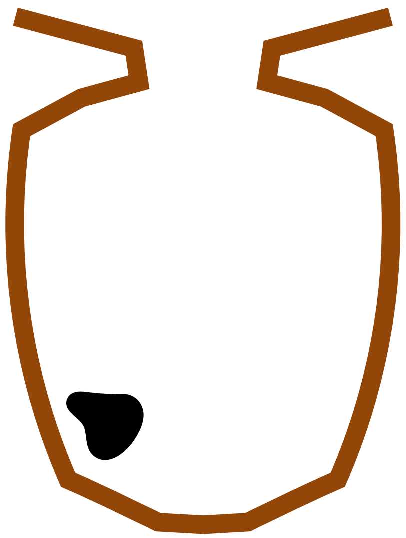
$$V_k \overset{iid}{\sim} \mathrm{Beta}(1, \alpha)$$

$$(\#\mathrm{orange}, \#\mathrm{other}) = \mathrm{PolyaUrn}(1, \alpha)$$

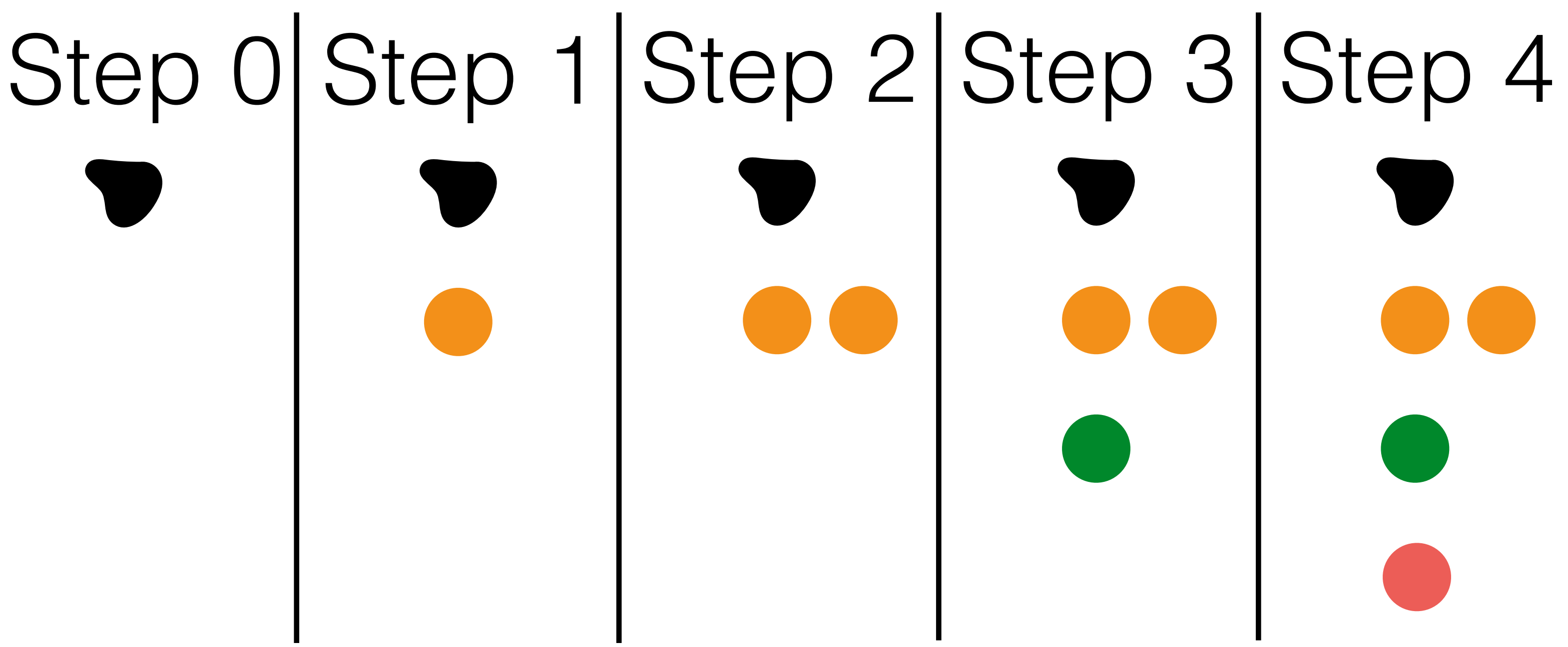
- not orange: $(\#\mathrm{green}, \#\mathrm{other}) = \mathrm{PolyaUrn}(1, \alpha)$
- not orange, green: $(\#\mathrm{red}, \#\mathrm{other}) = \mathrm{PolyaUrn}(1, \alpha)$

Marginal cluster assignments

- Hoppe urn / Blackwell-MacQueen urn



  - Choose ball with prob proportional to its mass
    - If black, replace and add ball of new color
    - Else, replace and add ball of same color



$$V_k \overset{iid}{\sim} \mathrm{Beta}(1, \alpha)$$

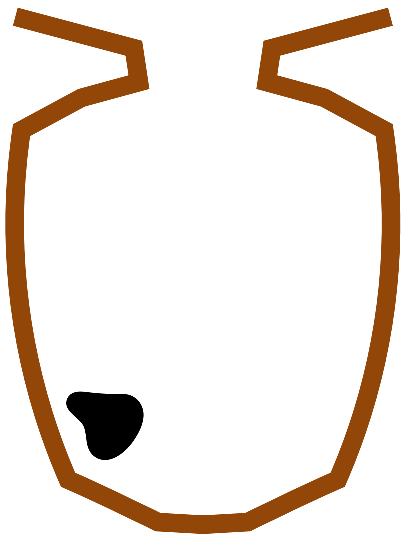
$$\rho_1 = V_1$$

$$(\#\text{orange}, \#\text{other}) = \mathrm{PolyaUrn}(1, \alpha)$$

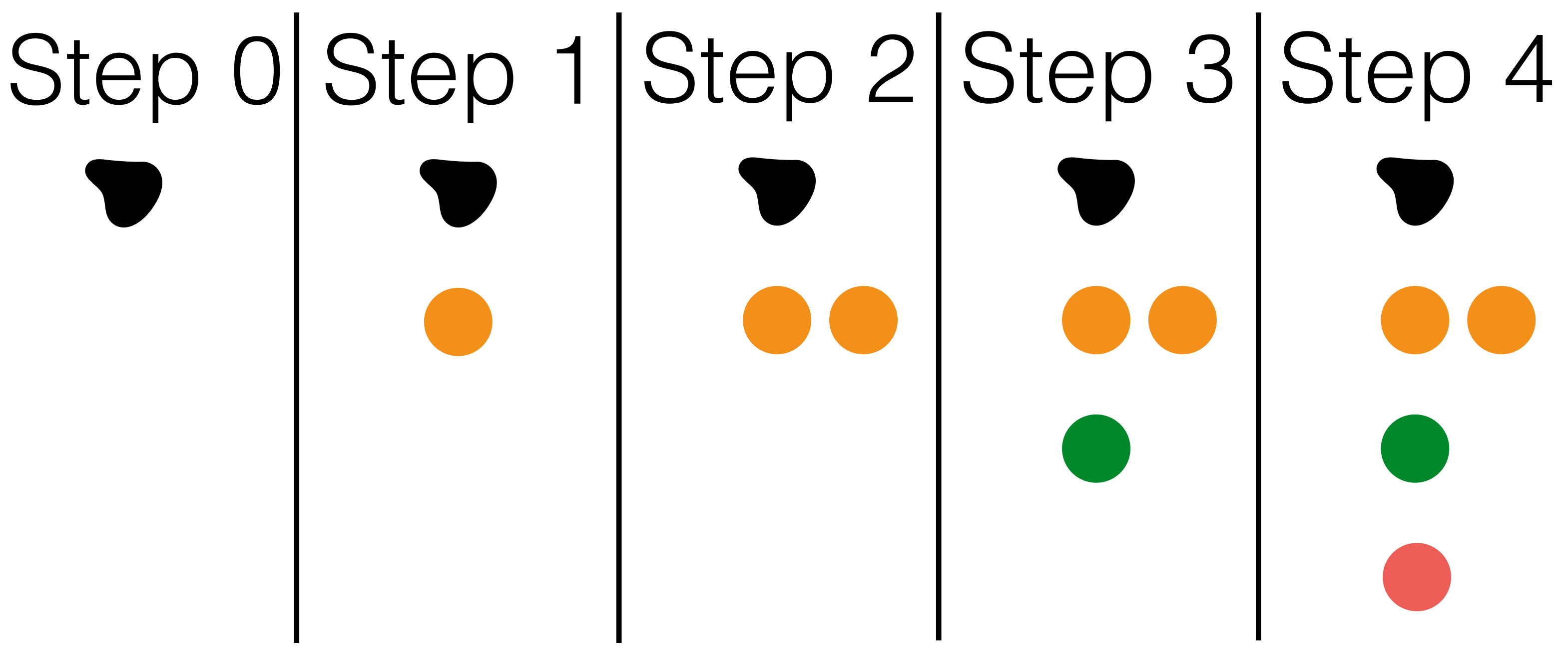
- not orange: $(\#\text{green}, \#\text{other}) = \mathrm{PolyaUrn}(1, \alpha)$
- not orange, green: $(\#\text{red}, \#\text{other}) = \mathrm{PolyaUrn}(1, \alpha)$

Marginal cluster assignments

- Hoppe urn / Blackwell-MacQueen urn



- Choose ball with prob proportional to its mass
  - If black, replace and add ball of new color
  - Else, replace and add ball of same color



$$V_k \overset{iid}{\sim} \mathrm{Beta}(1, \alpha)$$

$$\rho_1 = V_1$$

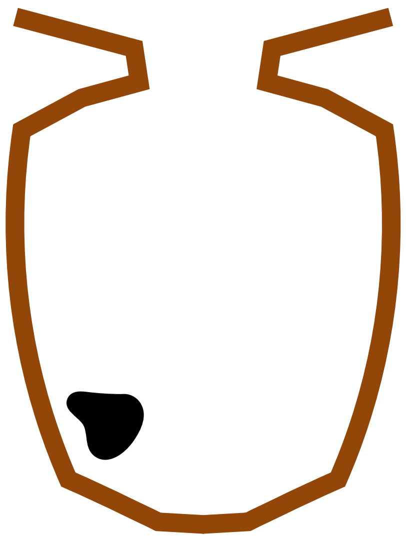
$$\rho_2 = (1 - V_1)V_2$$

$$(\#\text{orange}, \#\text{other}) = \mathrm{PolyaUrn}(1, \alpha)$$

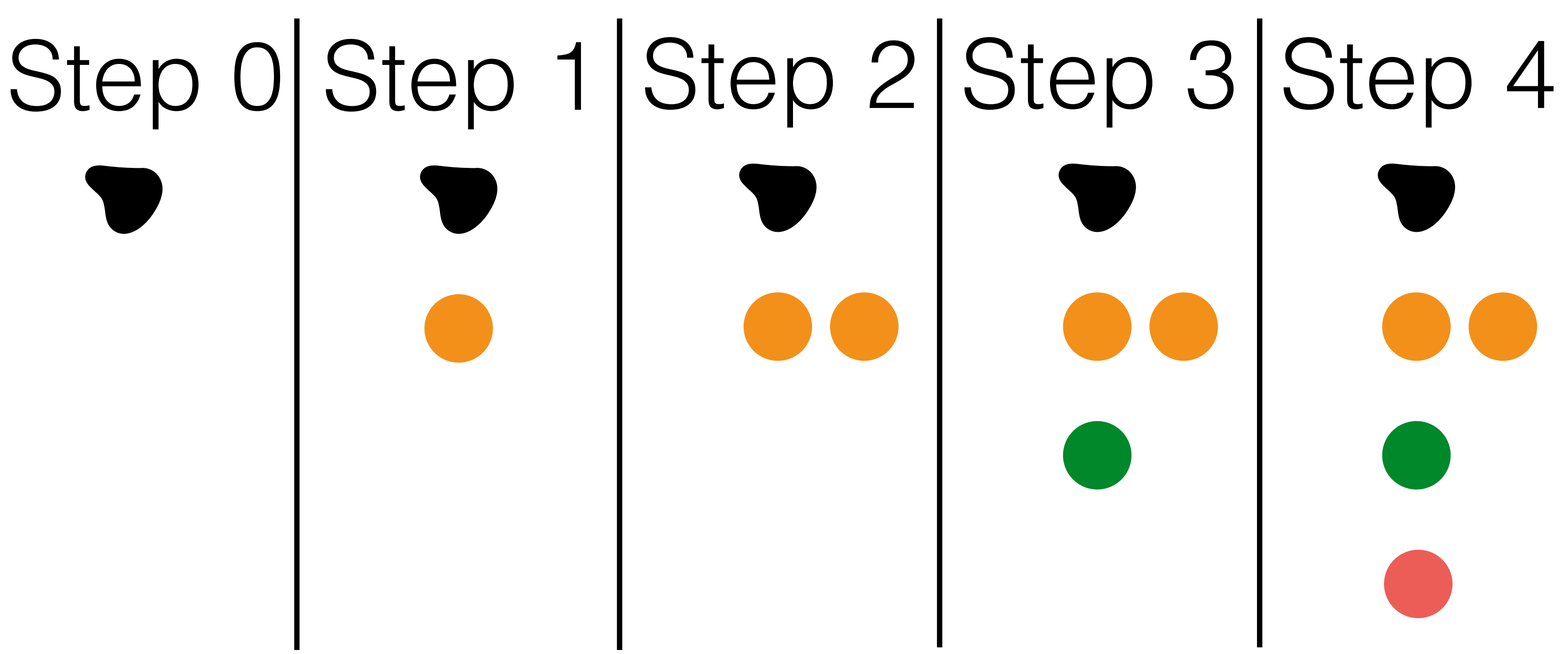
- not orange: $(\#\text{green}, \#\text{other}) = \mathrm{PolyaUrn}(1, \alpha)$
- not orange, green: $(\#\text{red}, \#\text{other}) = \mathrm{PolyaUrn}(1, \alpha)$

Marginal cluster assignments

- Hoppe urn / Blackwell-MacQueen urn



- Choose ball with prob proportional to its mass
  - If black, replace and add ball of new color
  - Else, replace and add ball of same color



$$V_k \stackrel{iid}{\sim} \mathrm{Beta}(1, \alpha)$$

$$\rho_1 = V_1$$

$$\rho_2 = (1 - V_1) V_2$$

$$\rho_3 = [\prod_{k=1}^{2} (1 - V_k)] V_3$$

$$(\#\text{orange}, \#\text{other}) = \text{PolyaUrn}(1, \alpha)$$

- not orange: $(\#\text{green}, \#\text{other}) = \text{PolyaUrn}(1, \alpha)$
- not orange, green: $(\#\text{red}, \#\text{other}) = \text{PolyaUrn}(1, \alpha)$

Chinese restaurant process

Chinese restaurant process

- Same thing we just did

Chinese restaurant process

- Same thing we just did
- Each customer walks into the restaurant

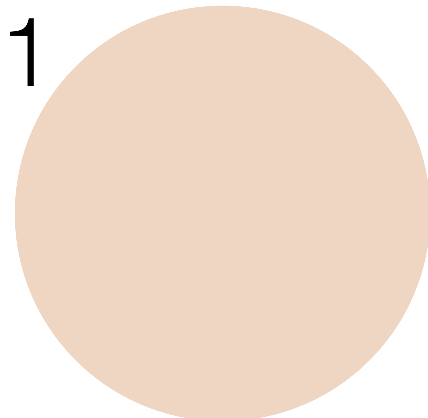
Chinese restaurant process

- Same thing we just did
- Each customer walks into the restaurant
  - Sits at existing table with prob proportional to # people there

Chinese restaurant process

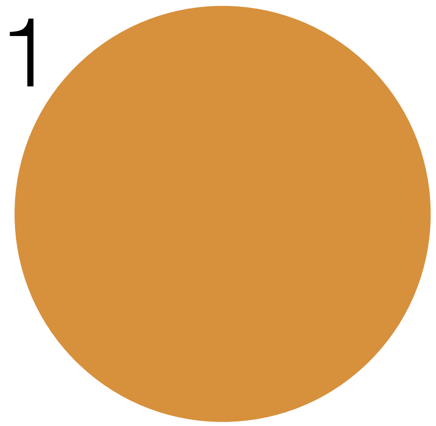
- Same thing we just did
- Each customer walks into the restaurant
  - Sits at existing table with prob proportional to # people there
  - Forms new table with prob proportional to α

Chinese restaurant process



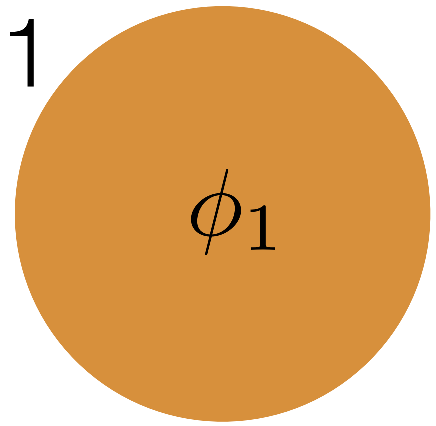
- Same thing we just did
- Each customer walks into the restaurant
  - Sits at existing table with prob proportional to # people there
  - Forms new table with prob proportional to α

Chinese restaurant process



- Same thing we just did
- Each customer walks into the restaurant
    - Sits at existing table with prob proportional to # people there
    - Forms new table with prob proportional to α

Chinese restaurant process

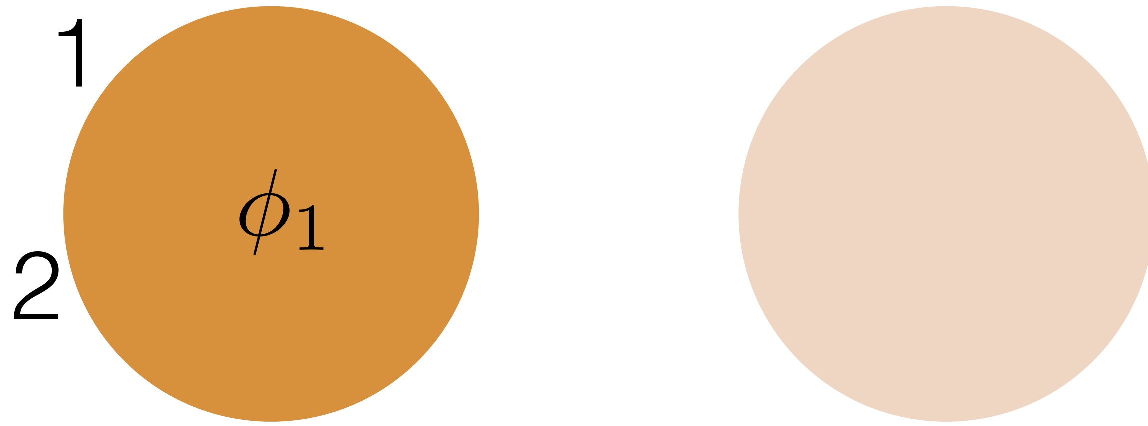


- Same thing we just did
- Each customer walks into the restaurant
  - Sits at existing table with prob proportional to # people there
  - Forms new table with prob proportional to α

Chinese restaurant process

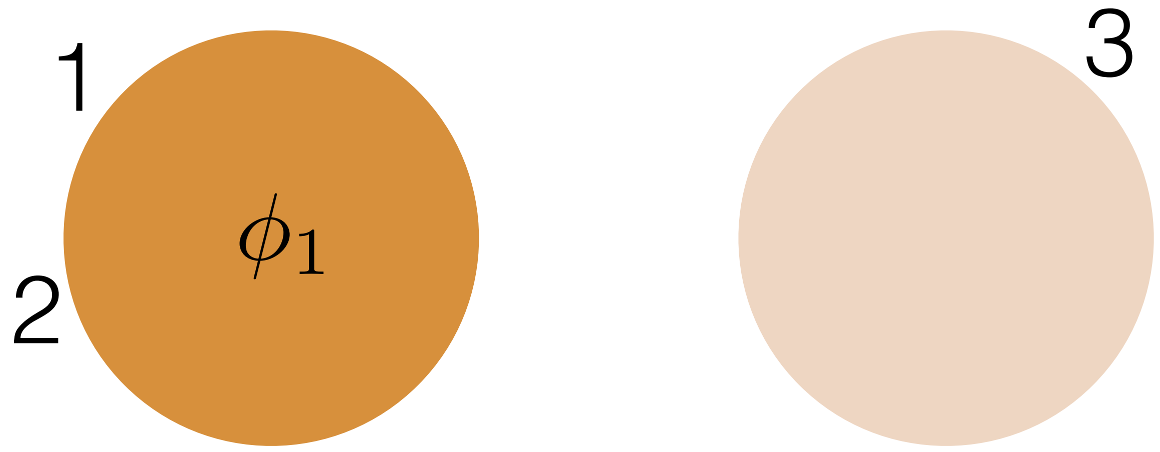
- Same thing we just did
- Each customer walks into the restaurant
  - Sits at existing table with prob proportional to # people there
  - Forms new table with prob proportional to α

Chinese restaurant process



- Same thing we just did
- Each customer walks into the restaurant
  - Sits at existing table with prob proportional to # people there
  - Forms new table with prob proportional to α

Chinese restaurant process



- Same thing we just did
- Each customer walks into the restaurant
  - Sits at existing table with prob proportional to # people there
  - Forms new table with prob proportional to α

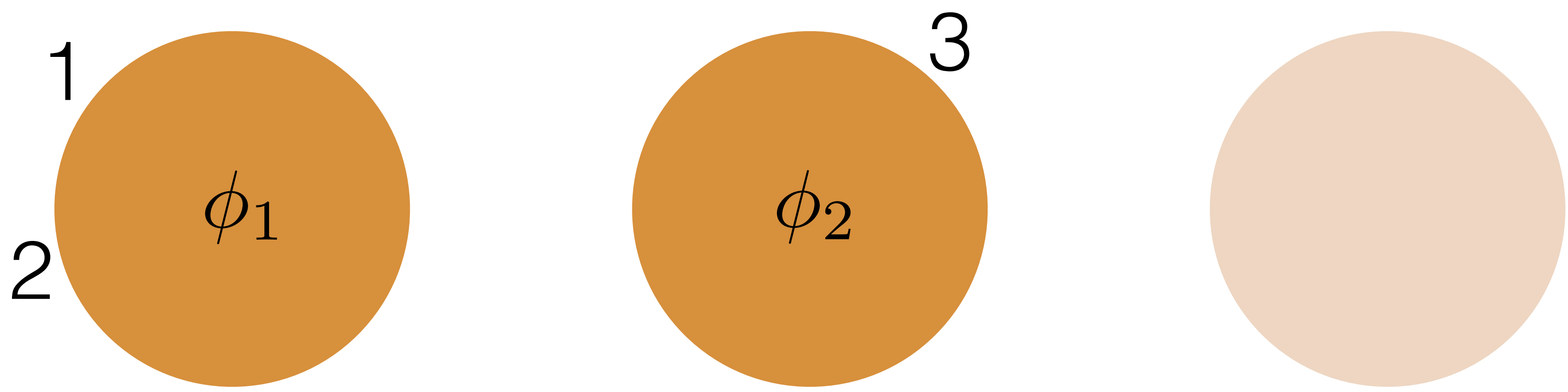
Chinese restaurant process

- Same thing we just did
- Each customer walks into the restaurant
  - Sits at existing table with prob proportional to # people there
  - Forms new table with prob proportional to α

Chinese restaurant process

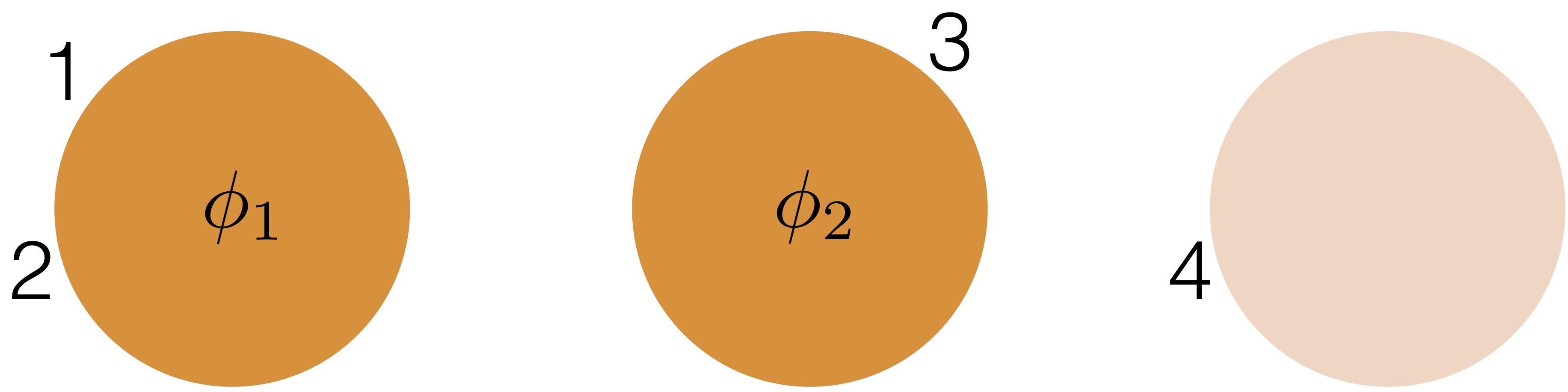
- Same thing we just did
- Each customer walks into the restaurant
  - Sits at existing table with prob proportional to # people there
  - Forms new table with prob proportional to α

Chinese restaurant process



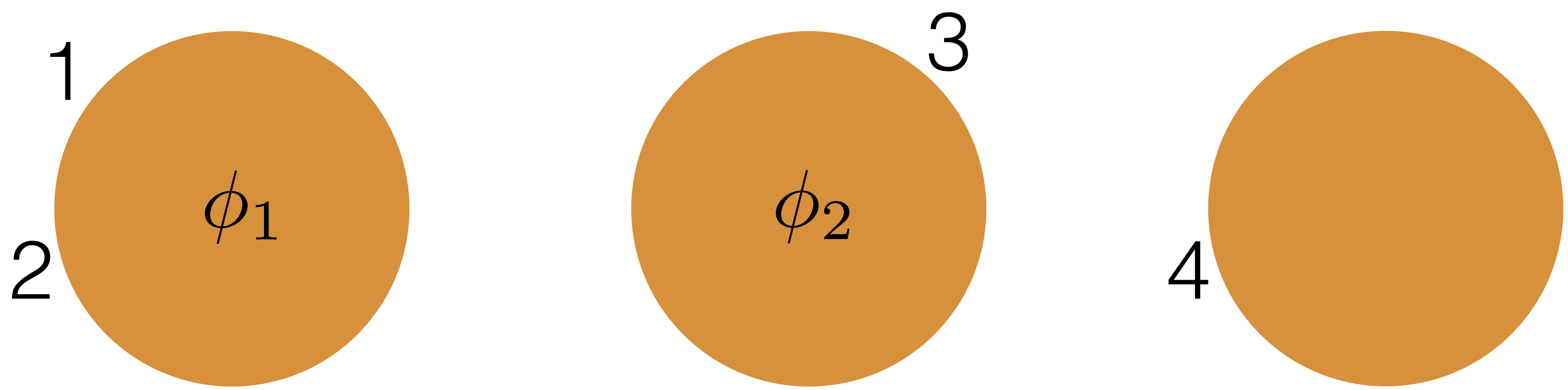
- Same thing we just did
- Each customer walks into the restaurant
  - Sits at existing table with prob proportional to # people there
  - Forms new table with prob proportional to α

Chinese restaurant process



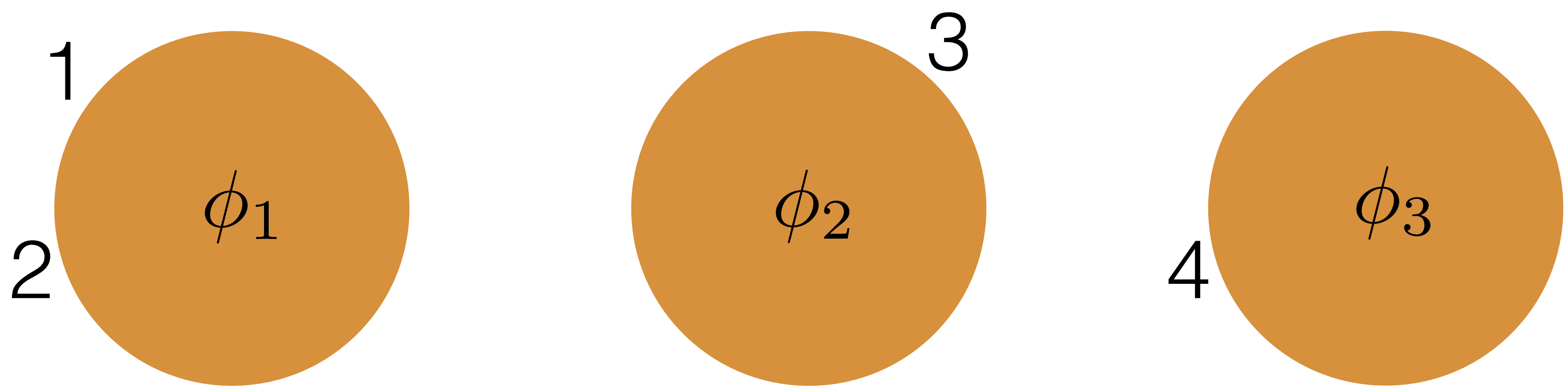
- Same thing we just did
- Each customer walks into the restaurant
  - Sits at existing table with prob proportional to # people there
  - Forms new table with prob proportional to α

Chinese restaurant process



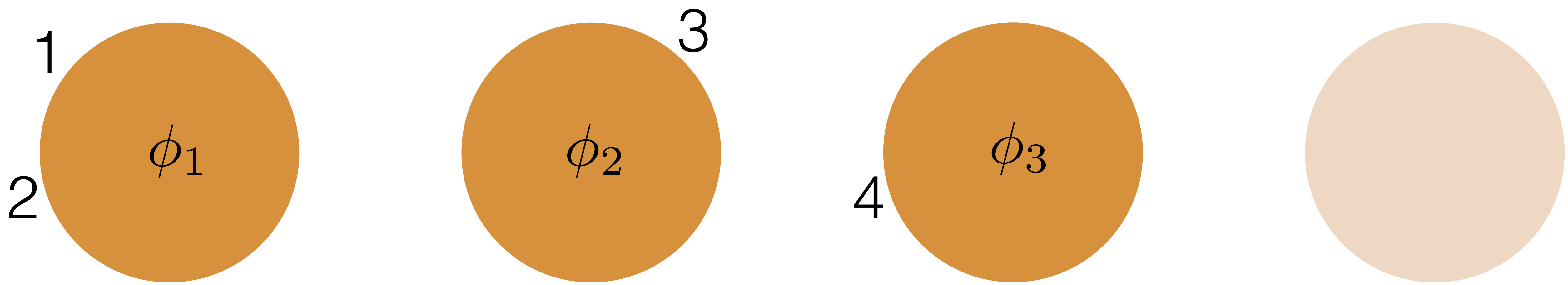
- Same thing we just did
- Each customer walks into the restaurant
  - Sits at existing table with prob proportional to # people there
  - Forms new table with prob proportional to α

Chinese restaurant process



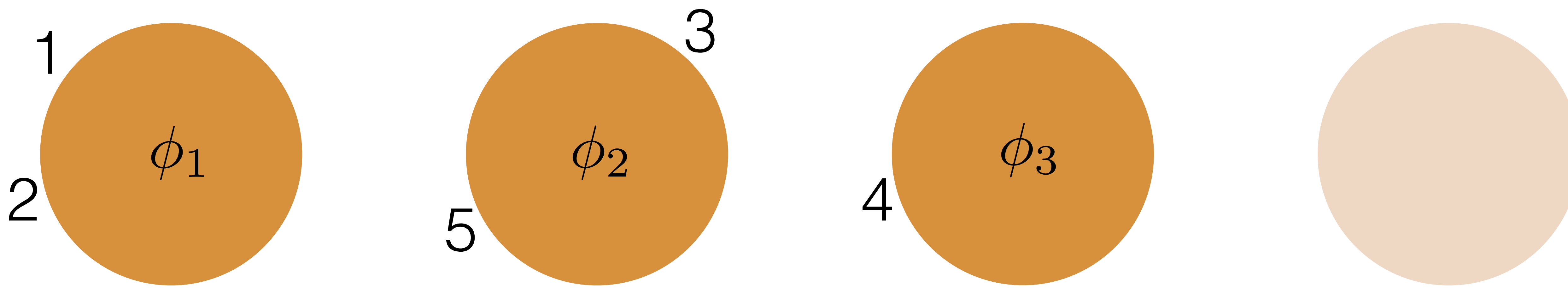
- Same thing we just did
- Each customer walks into the restaurant
  - Sits at existing table with prob proportional to # people there
  - Forms new table with prob proportional to α

Chinese restaurant process



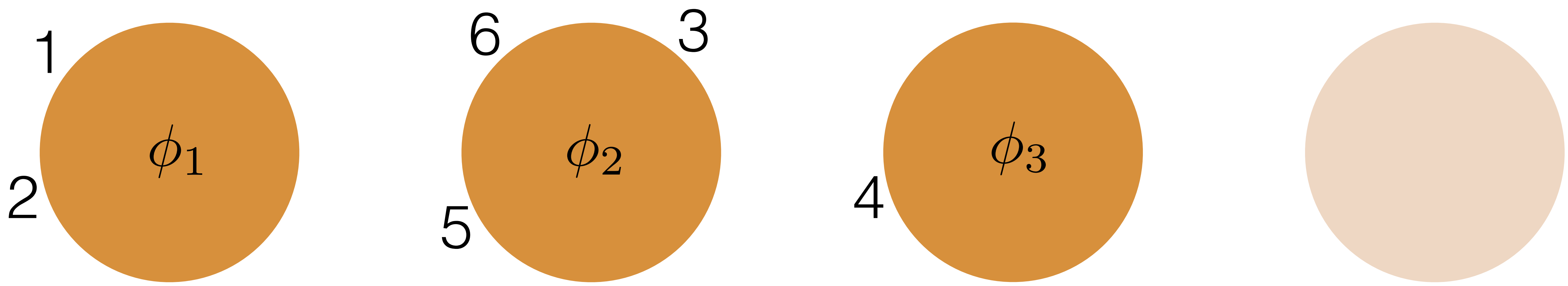
- Same thing we just did
- Each customer walks into the restaurant
  - Sits at existing table with prob proportional to # people there
  - Forms new table with prob proportional to α

Chinese restaurant process



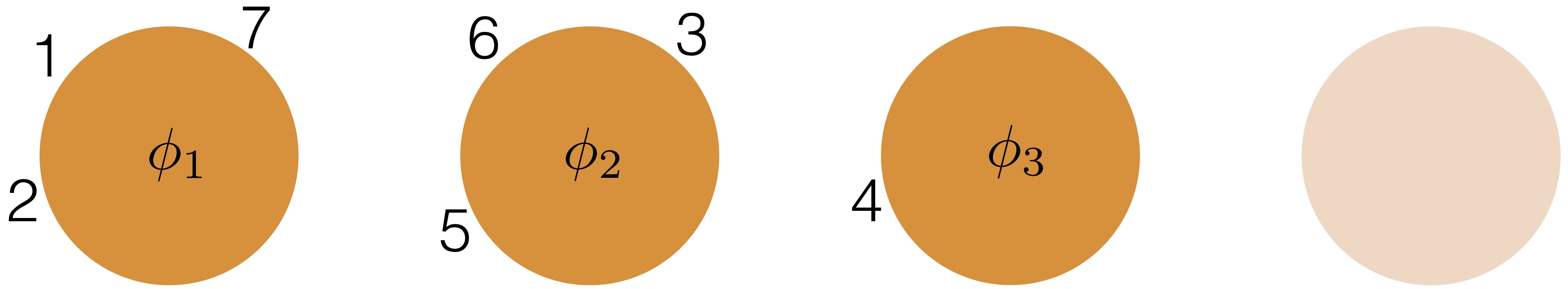
- Same thing we just did
- Each customer walks into the restaurant
    - Sits at existing table with prob proportional to # people there
    - Forms new table with prob proportional to α

Chinese restaurant process



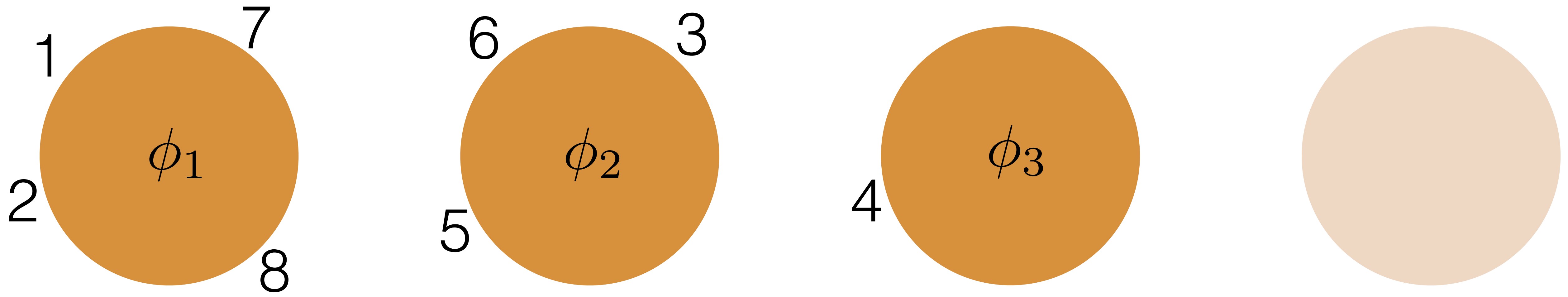
- Same thing we just did
- Each customer walks into the restaurant
  - Sits at existing table with prob proportional to # people there
  - Forms new table with prob proportional to α

Chinese restaurant process



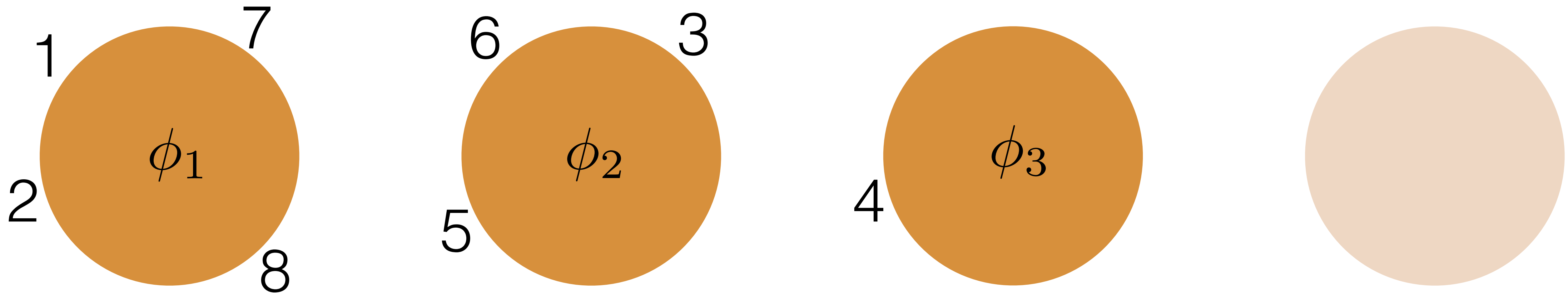
- Same thing we just did
- Each customer walks into the restaurant
  - Sits at existing table with prob proportional to # people there
  - Forms new table with prob proportional to α

Chinese restaurant process



- Same thing we just did
- Each customer walks into the restaurant
  - Sits at existing table with prob proportional to # people there
  - Forms new table with prob proportional to α

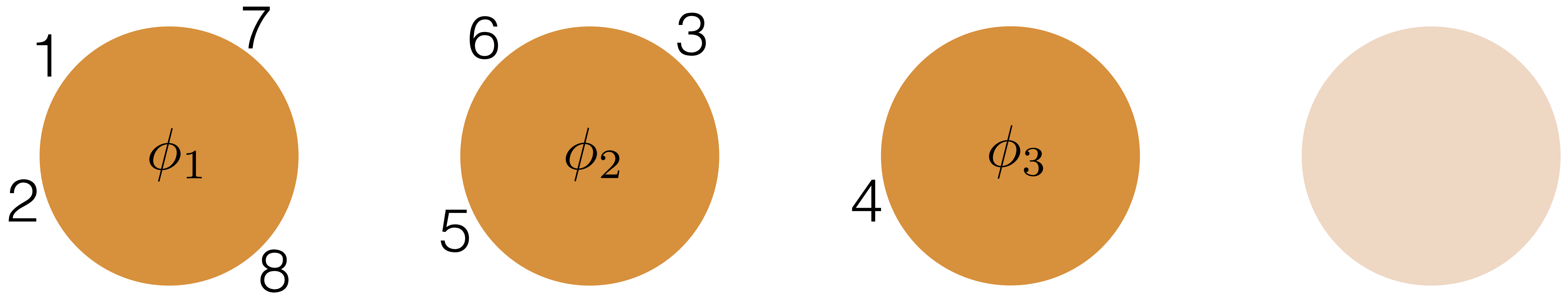
Chinese restaurant process



- Same thing we just did
- Each customer walks into the restaurant
  - Sits at existing table with prob proportional to # people there
  - Forms new table with prob proportional to α

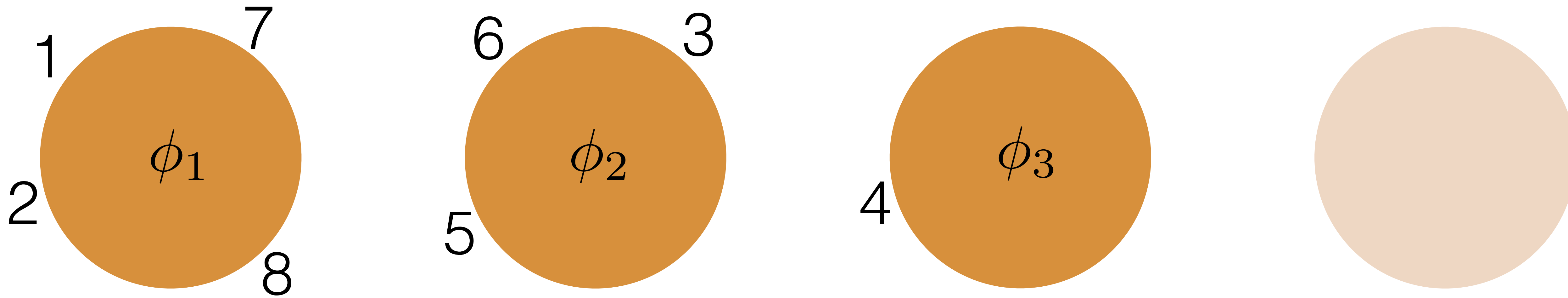
[Aldous 1983]

Chinese restaurant process



- Same thing we just did
- Each customer walks into the restaurant
  - Sits at existing table with prob proportional to # people there
  - Forms new table with prob proportional to α
- Marginal for the Categorical likelihood with GEM prior

Chinese restaurant process

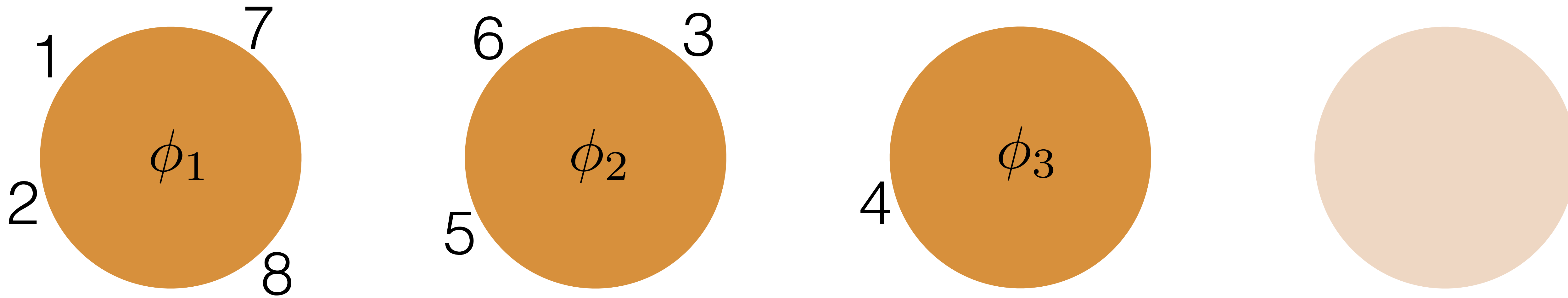


- Same thing we just did
- Each customer walks into the restaurant
  - Sits at existing table with prob proportional to # people there
  - Forms new table with prob proportional to α
- Marginal for the Categorical likelihood with GEM prior

So far: Dirichlet process, Chinese restaurant process

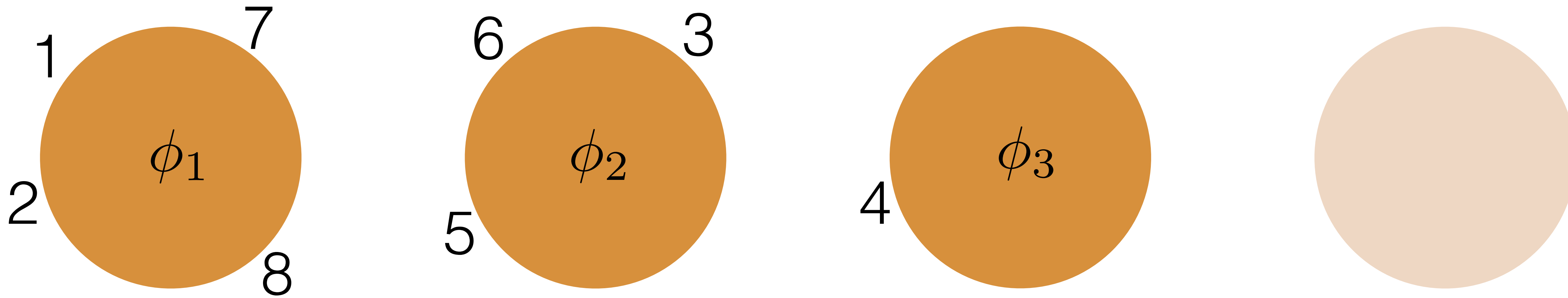
- Infinity of parameters, growing number of parameters

Chinese restaurant process



- Same thing we just did
- Each customer walks into the restaurant
  - Sits at existing table with prob proportional to # people there
  - Forms new table with prob proportional to α
- Marginal for the Categorical likelihood with GEM prior

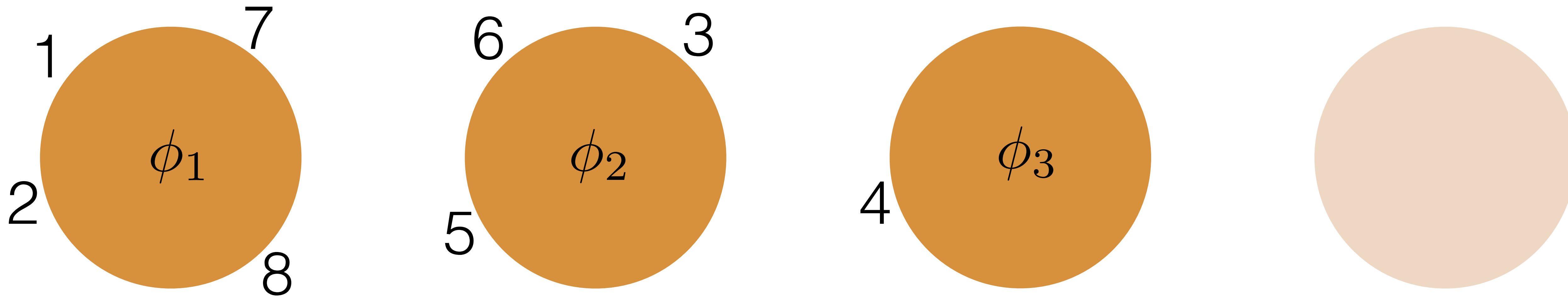
Chinese restaurant process



- Same thing we just did
- Each customer walks into the restaurant
  - Sits at existing table with prob proportional to # people there
  - Forms new table with prob proportional to α
- Marginal for the Categorical likelihood with GEM prior

$$z_1 = z_2 = z_7 = z_8 = 1, z_3 = z_5 = z_6 = 2, z_4 = 3$$

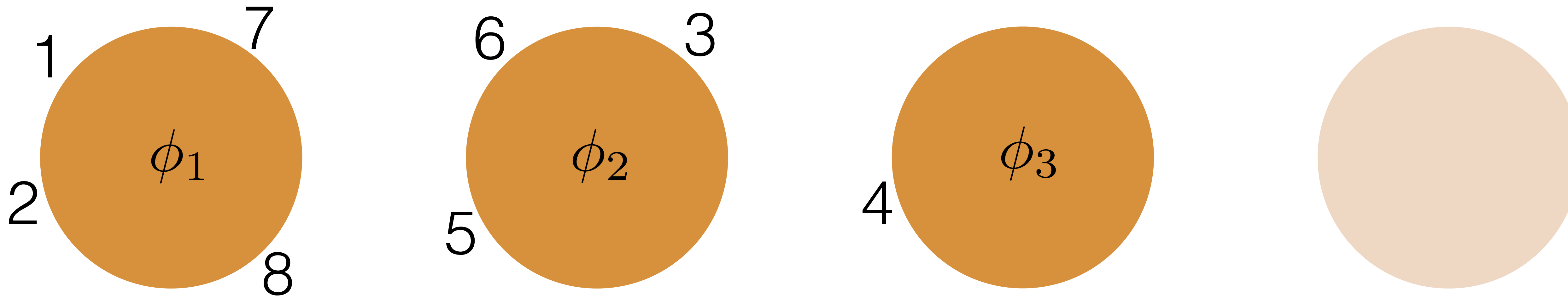
Chinese restaurant process



- Same thing we just did
- Each customer walks into the restaurant
  - Sits at existing table with prob proportional to # people there
  - Forms new table with prob proportional to α
- Marginal for the Categorical likelihood with GEM prior

$$z_1 = z_2 = z_7 = z_8 = 1, z_3 = z_5 = z_6 = 2, z_4 = 3$$
$$\Rightarrow \Pi_8 = \{\{1, 2, 7, 8\}, \{3, 5, 6\}, \{4\}\}$$

Chinese restaurant process

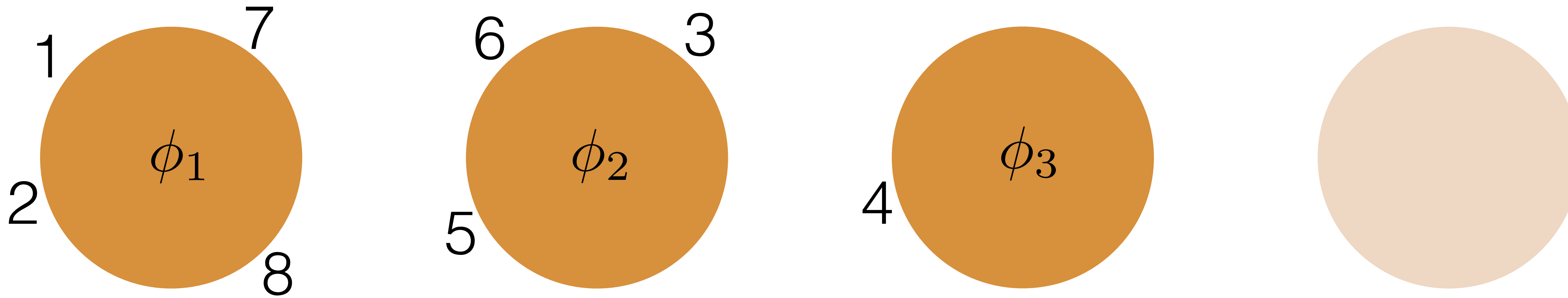


- Same thing we just did
- Each customer walks into the restaurant
  - Sits at existing table with prob proportional to # people there
  - Forms new table with prob proportional to α
- Marginal for the Categorical likelihood with GEM prior

$$z_1 = z_2 = z_7 = z_8 = 1, z_3 = z_5 = z_6 = 2, z_4 = 3$$
$$\Rightarrow \Pi_8 = \{\{1, 2, 7, 8\}, \{3, 5, 6\}, \{4\}\}$$

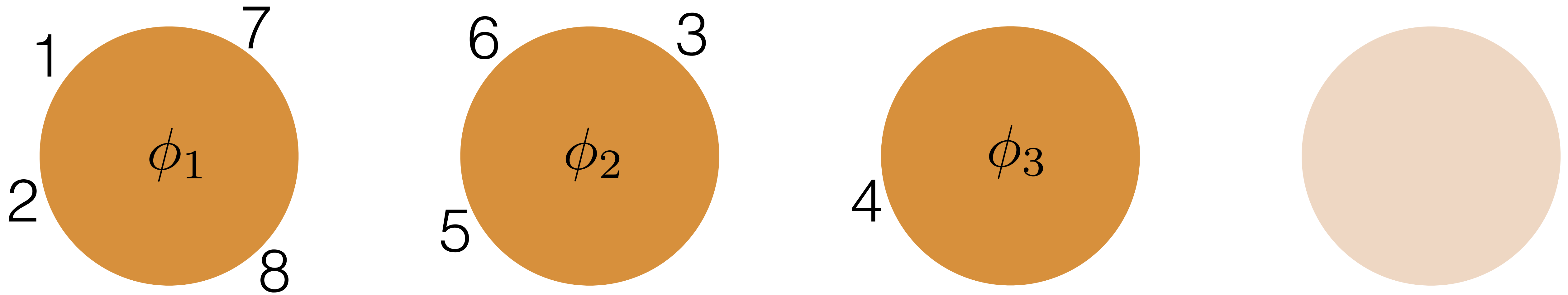
- *Partition of [8]*: set of mutually exclusive & exhaustive sets of $[8] := \{1, \ldots, 8\}$

Chinese restaurant process



- Probability of this seating:

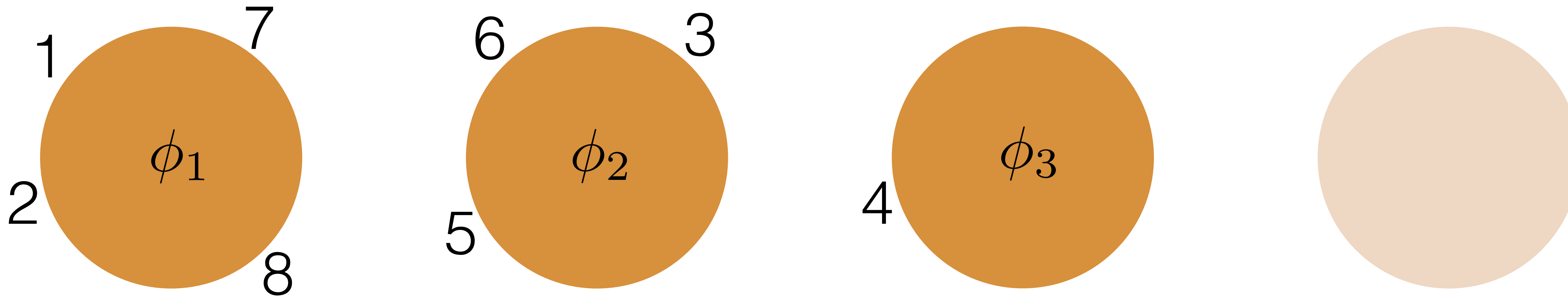
Chinese restaurant process



- Probability of this seating:

  $$\frac{\alpha}{\alpha}$$

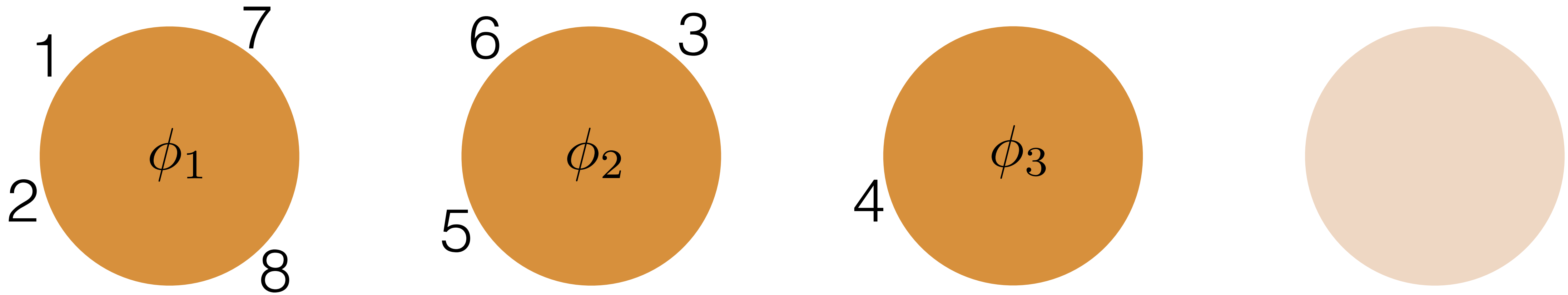
Chinese restaurant process



- Probability of this seating:

$$\frac{\alpha}{\alpha} \cdot \frac{1}{\alpha + 1}$$

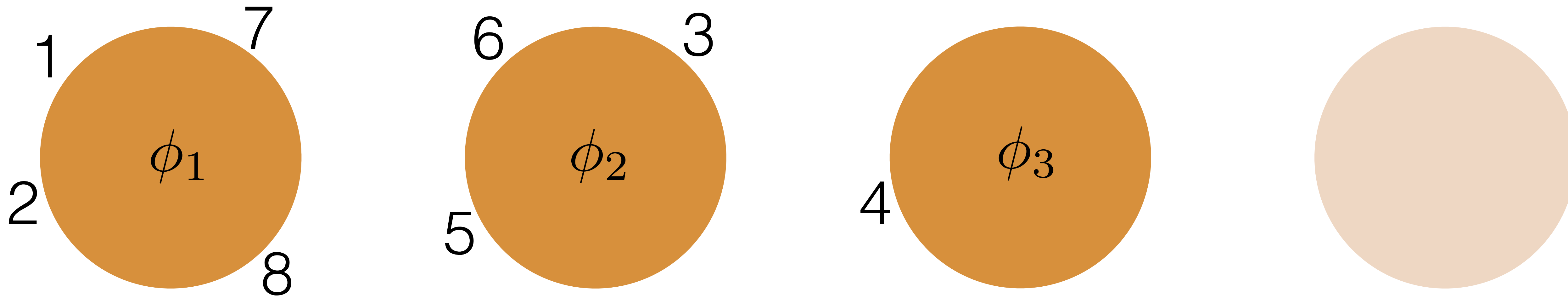
Chinese restaurant process



- Probability of this seating:

$$\frac{\alpha}{\alpha} \cdot \frac{1}{\alpha+1} \cdot \frac{\alpha}{\alpha+2}$$

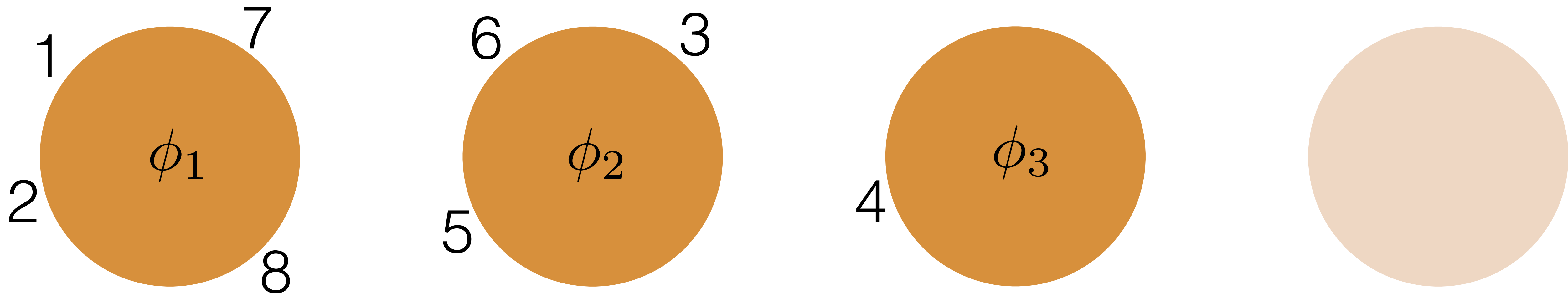
Chinese restaurant process



- Probability of this seating:

$$\frac{\alpha}{\alpha} \cdot \frac{1}{\alpha+1} \cdot \frac{\alpha}{\alpha+2} \cdot \frac{\alpha}{\alpha+3}$$

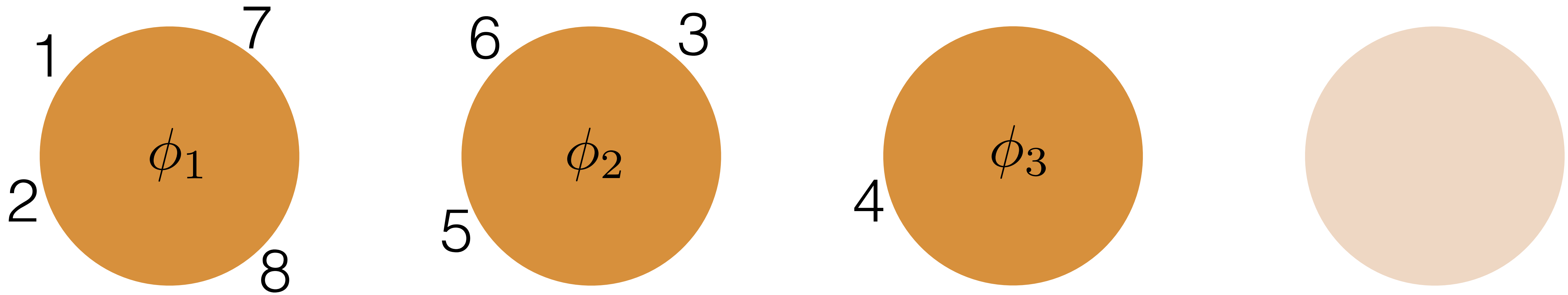
Chinese restaurant process



- Probability of this seating:

$$\frac{\alpha}{\alpha} \cdot \frac{1}{\alpha+1} \cdot \frac{\alpha}{\alpha+2} \cdot \frac{\alpha}{\alpha+3} \cdot \frac{1}{\alpha+4}$$

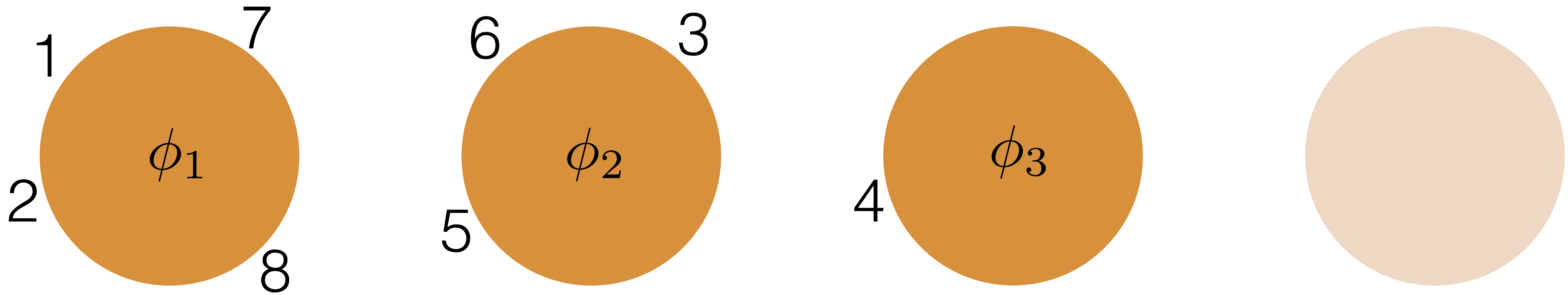
Chinese restaurant process



- Probability of this seating:

$$\frac{\alpha}{\alpha} \cdot \frac{1}{\alpha+1} \cdot \frac{\alpha}{\alpha+2} \cdot \frac{\alpha}{\alpha+3} \cdot \frac{1}{\alpha+4} \cdot \frac{2}{\alpha+5}$$

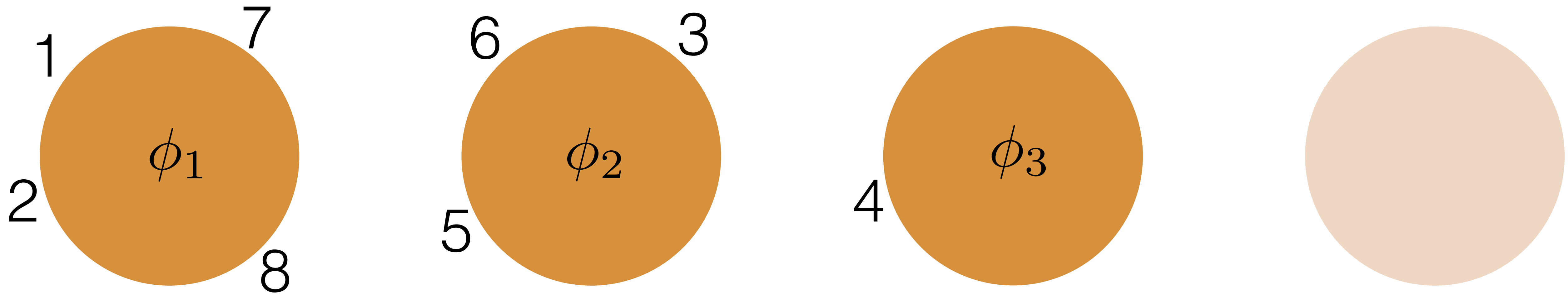
Chinese restaurant process



- Probability of this seating:

$$\frac{\alpha}{\alpha} \cdot \frac{1}{\alpha+1} \cdot \frac{\alpha}{\alpha+2} \cdot \frac{\alpha}{\alpha+3} \cdot \frac{1}{\alpha+4} \cdot \frac{2}{\alpha+5} \cdot \frac{2}{\alpha+6}$$

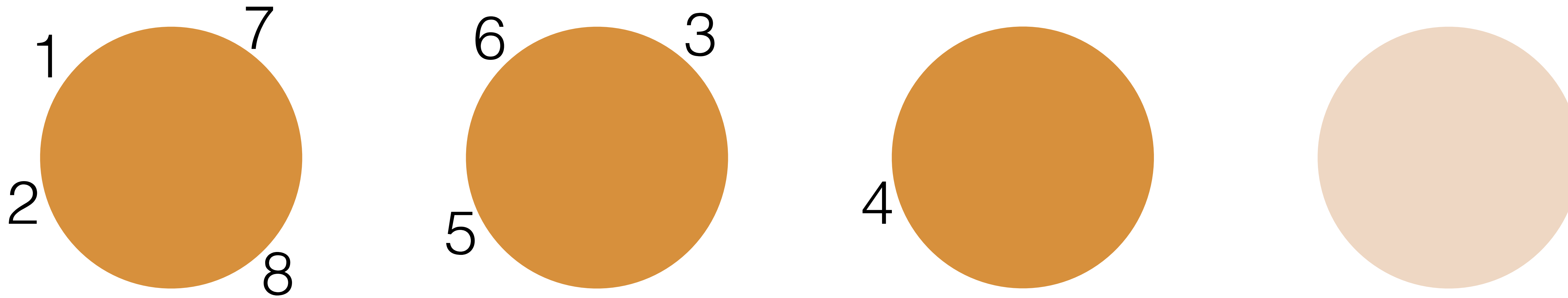
Chinese restaurant process



- Probability of this seating:

$$\frac{\alpha}{\alpha} \cdot \frac{1}{\alpha+1} \cdot \frac{\alpha}{\alpha+2} \cdot \frac{\alpha}{\alpha+3} \cdot \frac{1}{\alpha+4} \cdot \frac{2}{\alpha+5} \cdot \frac{2}{\alpha+6} \cdot \frac{3}{\alpha+7}$$

Chinese restaurant process

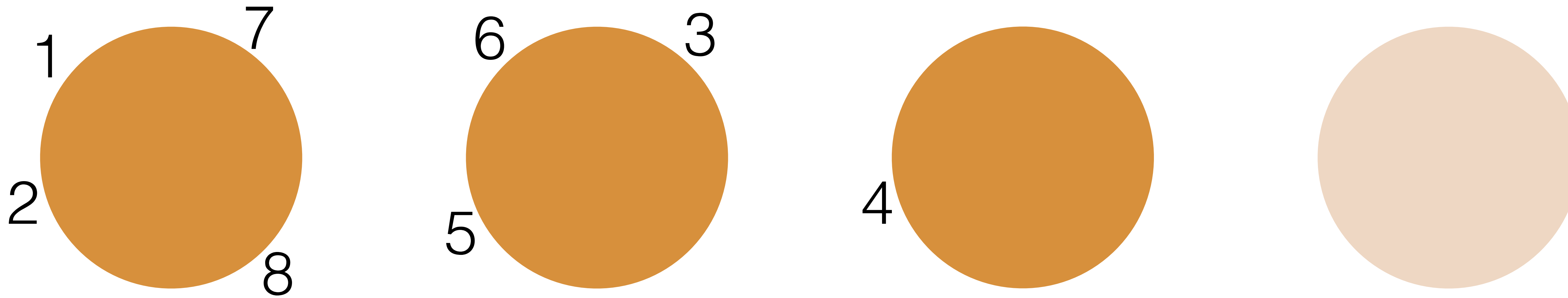


- Probability of this seating:

$$\frac{\alpha}{\alpha} \cdot \frac{1}{\alpha+1} \cdot \frac{\alpha}{\alpha+2} \cdot \frac{\alpha}{\alpha+3} \cdot \frac{1}{\alpha+4} \cdot \frac{2}{\alpha+5} \cdot \frac{2}{\alpha+6} \cdot \frac{3}{\alpha+7}$$

- Probability of *N* customers (K_N tables, n_k at table *k*):

Chinese restaurant process

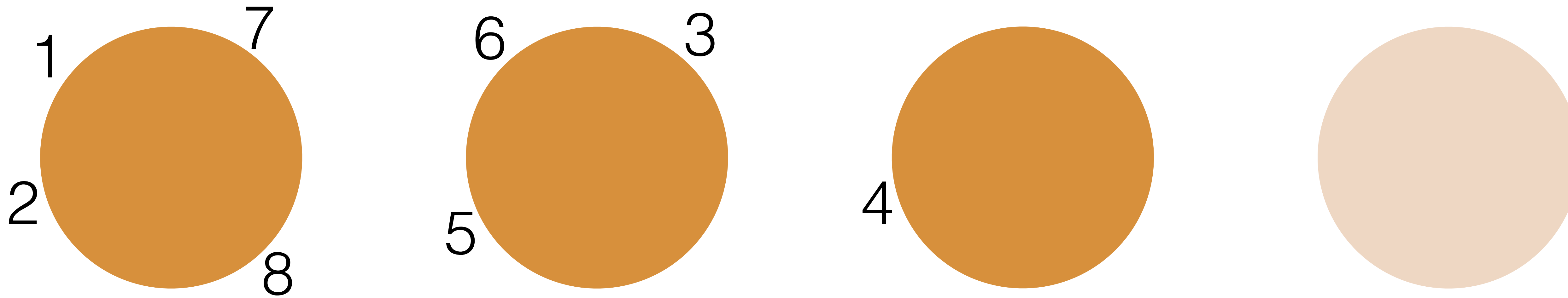


- Probability of this seating:

$$\frac{\alpha}{\alpha} \cdot \frac{1}{\alpha+1} \cdot \frac{\alpha}{\alpha+2} \cdot \frac{\alpha}{\alpha+3} \cdot \frac{1}{\alpha+4} \cdot \frac{2}{\alpha+5} \cdot \frac{2}{\alpha+6} \cdot \frac{3}{\alpha+7}$$

- Probability of *N* customers (K_N tables, n_k at table *k*):

Chinese restaurant process



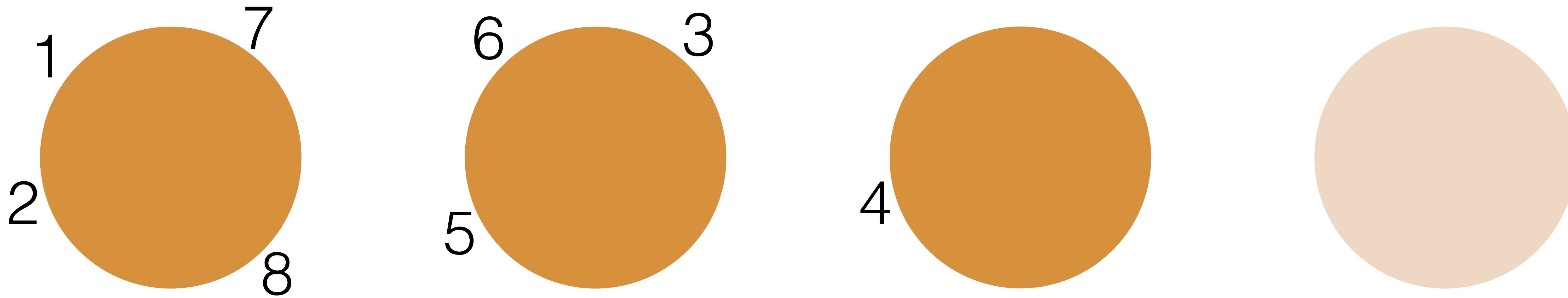
- Probability of this seating:

$$\frac{\alpha}{\alpha} \cdot \frac{1}{\alpha+1} \cdot \frac{\alpha}{\alpha+2} \cdot \frac{\alpha}{\alpha+3} \cdot \frac{1}{\alpha+4} \cdot \frac{2}{\alpha+5} \cdot \frac{2}{\alpha+6} \cdot \frac{3}{\alpha+7}$$

- Probability of *N* customers (K_N tables, n_k at table *k*):

$$\frac{\quad}{\alpha \cdots (\alpha + N - 1)}$$

Chinese restaurant process



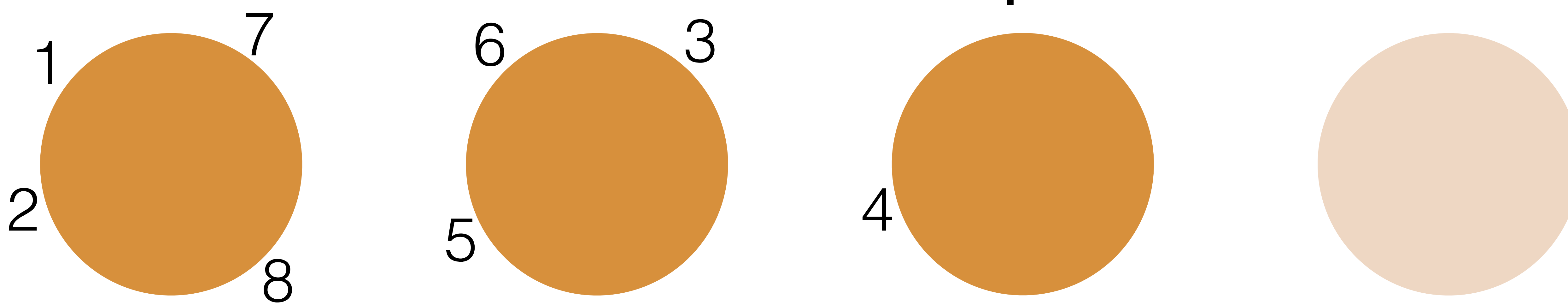
- Probability of this seating:

$$\frac{\alpha}{\alpha} \cdot \frac{1}{\alpha+1} \cdot \frac{\alpha}{\alpha+2} \cdot \frac{\alpha}{\alpha+3} \cdot \frac{1}{\alpha+4} \cdot \frac{2}{\alpha+5} \cdot \frac{2}{\alpha+6} \cdot \frac{3}{\alpha+7}$$

- Probability of *N* customers (K_N tables, n_k at table *k*):

$$\frac{\alpha^{K_N}}{\alpha \cdots (\alpha + N - 1)}$$

Chinese restaurant process



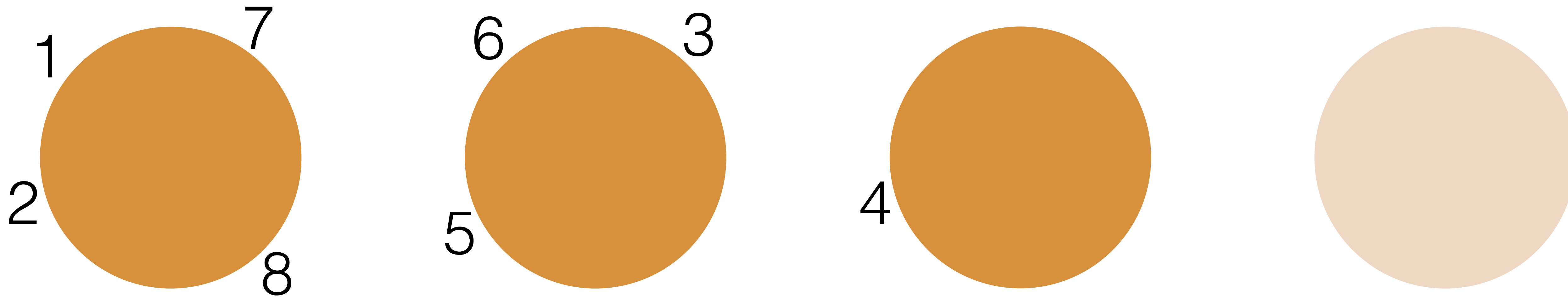
- Probability of this seating:

$$\frac{\alpha}{\alpha} \cdot \frac{1}{\alpha+1} \cdot \frac{\alpha}{\alpha+2} \cdot \frac{\alpha}{\alpha+3} \cdot \frac{1}{\alpha+4} \cdot \frac{2}{\alpha+5} \cdot \frac{2}{\alpha+6} \cdot \frac{3}{\alpha+7}$$

- Probability of *N* customers (K_N tables, n_k at table *k*):

$$\frac{\alpha^{K_N}}{\alpha \cdots (\alpha + N - 1)}$$

Chinese restaurant process



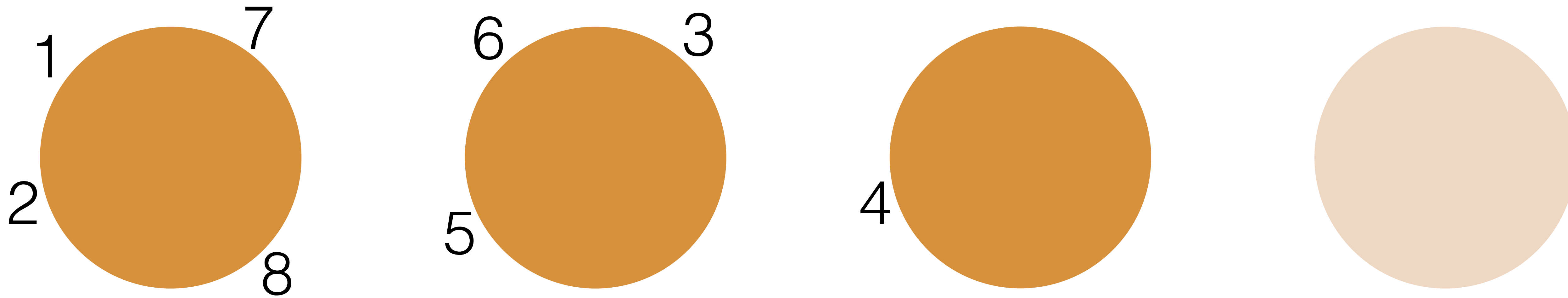
- Probability of this seating:

$$\frac{\alpha}{\alpha} \cdot \frac{1}{\alpha+1} \cdot \frac{\alpha}{\alpha+2} \cdot \frac{\alpha}{\alpha+3} \cdot \frac{1}{\alpha+4} \cdot \frac{2}{\alpha+5} \cdot \frac{2}{\alpha+6} \cdot \frac{3}{\alpha+7}$$

- Probability of *N* customers (K_N tables, n_k at table *k*):

$$\frac{\alpha^{K_N}}{\alpha \cdots (\alpha + N - 1)}$$

Chinese restaurant process



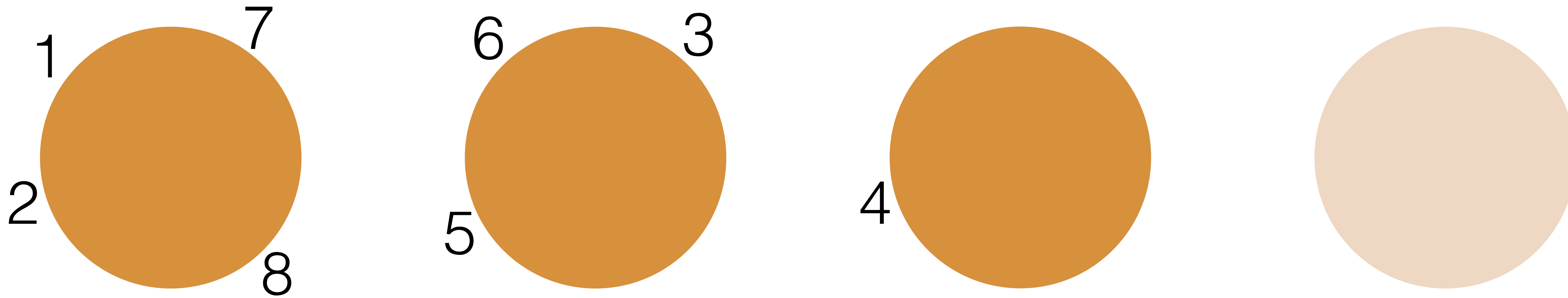
- Probability of this seating:

$$\frac{\alpha}{\alpha} \cdot \frac{1}{\alpha+1} \cdot \frac{\alpha}{\alpha+2} \cdot \frac{\alpha}{\alpha+3} \cdot \frac{1}{\alpha+4} \cdot \frac{2}{\alpha+5} \cdot \frac{2}{\alpha+6} \cdot \frac{3}{\alpha+7}$$

- Probability of *N* customers (K_N tables, n_k at table *k*):

$$\frac{\alpha^{K_N}}{\alpha \cdots (\alpha + N - 1)}$$

Chinese restaurant process



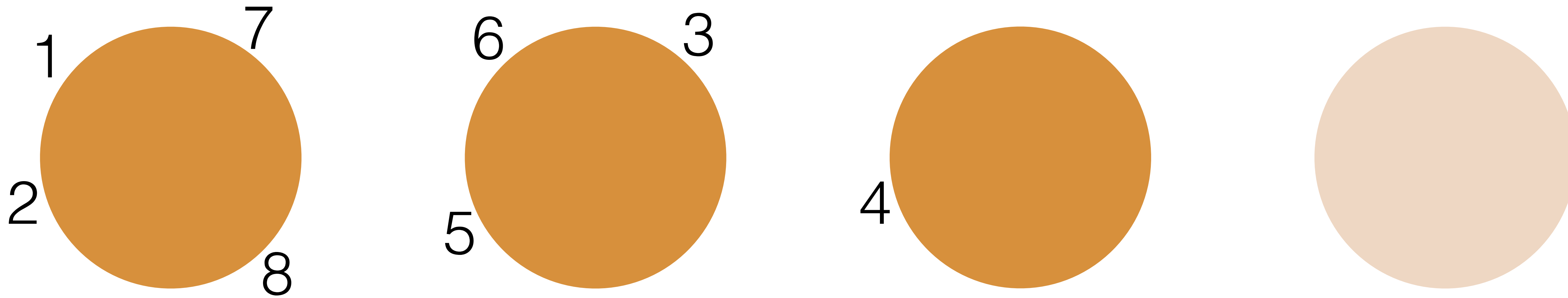
- Probability of this seating:

$$\frac{\alpha}{\alpha} \cdot \frac{1}{\alpha+1} \cdot \frac{\alpha}{\alpha+2} \cdot \frac{\alpha}{\alpha+3} \cdot \frac{1}{\alpha+4} \cdot \frac{2}{\alpha+5} \cdot \frac{2}{\alpha+6} \cdot \frac{3}{\alpha+7}$$

- Probability of *N* customers (K_N tables, n_k at table *k*):

$$\frac{\alpha^{K_N}}{\alpha \cdots (\alpha+N-1)}$$

Chinese restaurant process



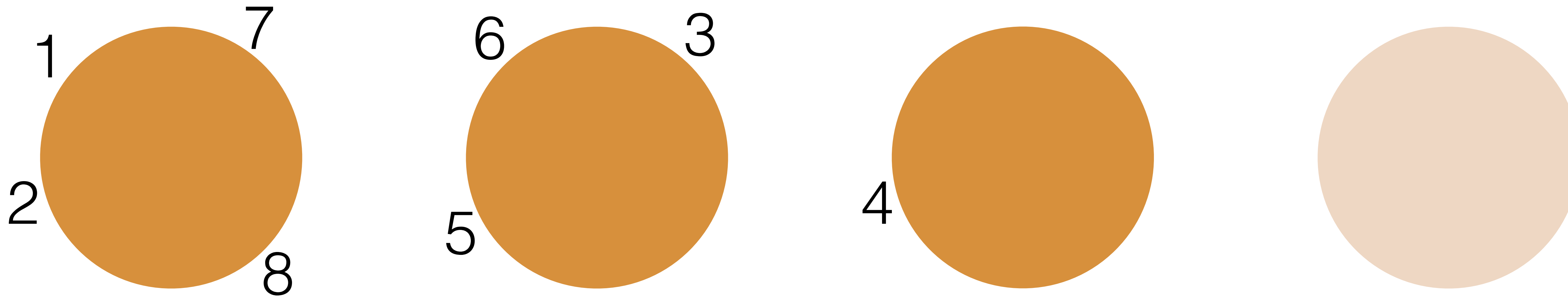
- Probability of this seating:

$$\frac{\alpha}{\alpha} \cdot \frac{1}{\alpha+1} \cdot \frac{\alpha}{\alpha+2} \cdot \frac{\alpha}{\alpha+3} \cdot \frac{1}{\alpha+4} \cdot \frac{2}{\alpha+5} \cdot \frac{2}{\alpha+6} \cdot \frac{3}{\alpha+7}$$

- Probability of *N* customers (K_N tables, n_k at table *k*):

$$\frac{\alpha^{K_N} \prod_{k=1}^{K_N} (n_k - 1)!}{\alpha \cdots (\alpha + N - 1)}$$

Chinese restaurant process



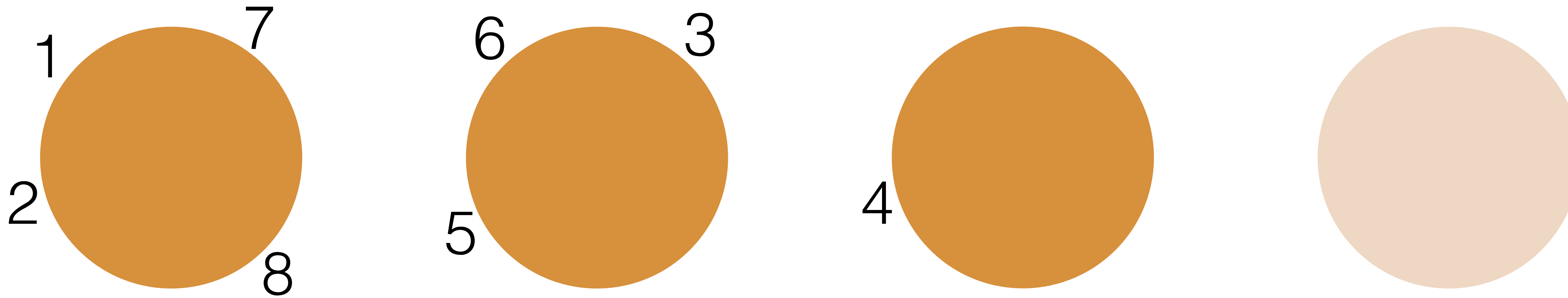
- Probability of this seating:

$$\frac{\alpha}{\alpha} \cdot \frac{1}{\alpha+1} \cdot \frac{\alpha}{\alpha+2} \cdot \frac{\alpha}{\alpha+3} \cdot \frac{1}{\alpha+4} \cdot \frac{2}{\alpha+5} \cdot \frac{2}{\alpha+6} \cdot \frac{3}{\alpha+7}$$

- Probability of N customers (K_N tables, $\#C$ at table C):

$$\frac{\alpha^{K_N} \prod_{C \in \Pi_N} (\#C - 1)!}{\alpha \cdots (\alpha + N - 1)}$$

Chinese restaurant process



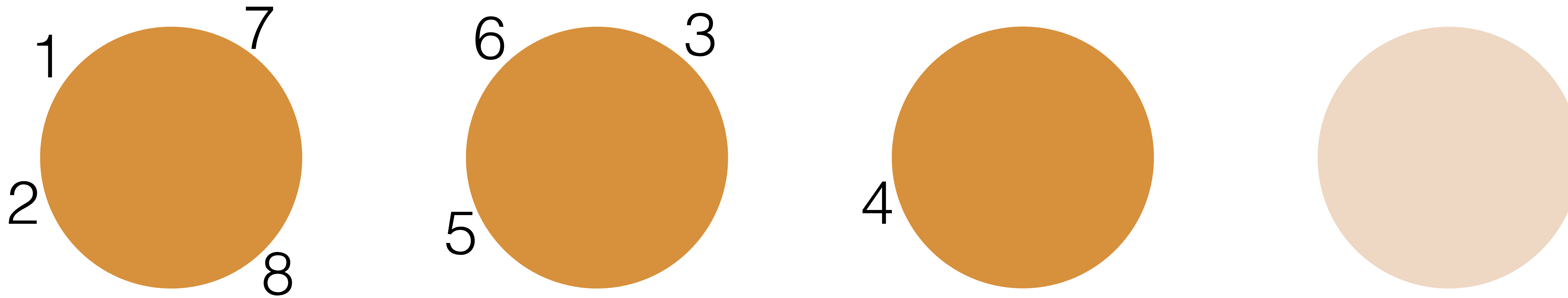
- Probability of this seating:

$$\frac{\alpha}{\alpha} \cdot \frac{1}{\alpha+1} \cdot \frac{\alpha}{\alpha+2} \cdot \frac{\alpha}{\alpha+3} \cdot \frac{1}{\alpha+4} \cdot \frac{2}{\alpha+5} \cdot \frac{2}{\alpha+6} \cdot \frac{3}{\alpha+7}$$

- Probability of N customers (K_N tables, $\#C$ at table C):

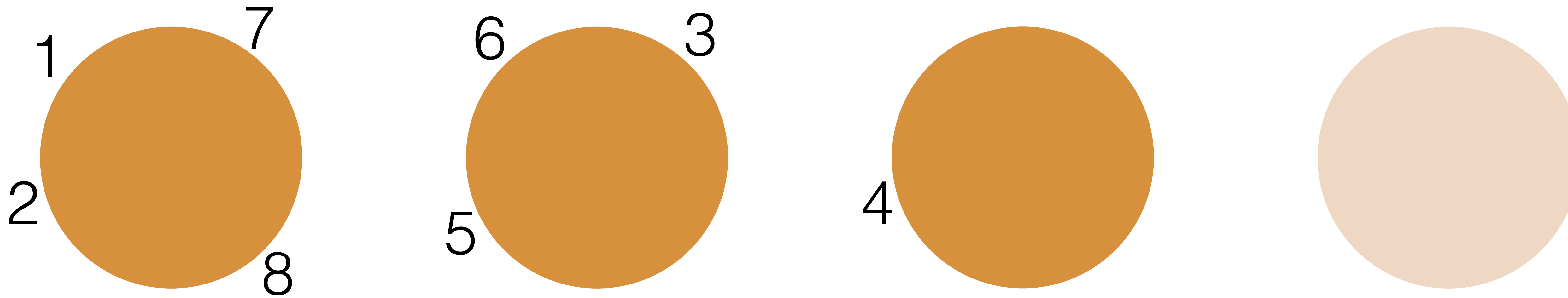
$$\frac{\alpha^{K_N} \prod_{C \in \Pi_N} (\#C - 1)!}{\alpha \cdots (\alpha + N - 1)} = \mathbb{P}(\Pi_N = \pi_N)$$

Chinese restaurant process



- Probability of this seating:
  $$\frac{\alpha}{\alpha} \cdot \frac{1}{\alpha+1} \cdot \frac{\alpha}{\alpha+2} \cdot \frac{\alpha}{\alpha+3} \cdot \frac{1}{\alpha+4} \cdot \frac{2}{\alpha+5} \cdot \frac{2}{\alpha+6} \cdot \frac{3}{\alpha+7}$$
- Probability of N customers (K_N tables, $\#C$ at table C):
  $$\frac{\alpha^{K_N} \prod_{C \in \Pi_N} (\#C - 1)!}{\alpha \cdots (\alpha + N - 1)} = \mathbb{P}(\Pi_N = \pi_N)$$
- Prob doesn't depend on customer order: *exchangeable*

Chinese restaurant process



- Probability of this seating:

$$\frac{\alpha}{\alpha} \cdot \frac{1}{\alpha+1} \cdot \frac{\alpha}{\alpha+2} \cdot \frac{\alpha}{\alpha+3} \cdot \frac{1}{\alpha+4} \cdot \frac{2}{\alpha+5} \cdot \frac{2}{\alpha+6} \cdot \frac{3}{\alpha+7}$$

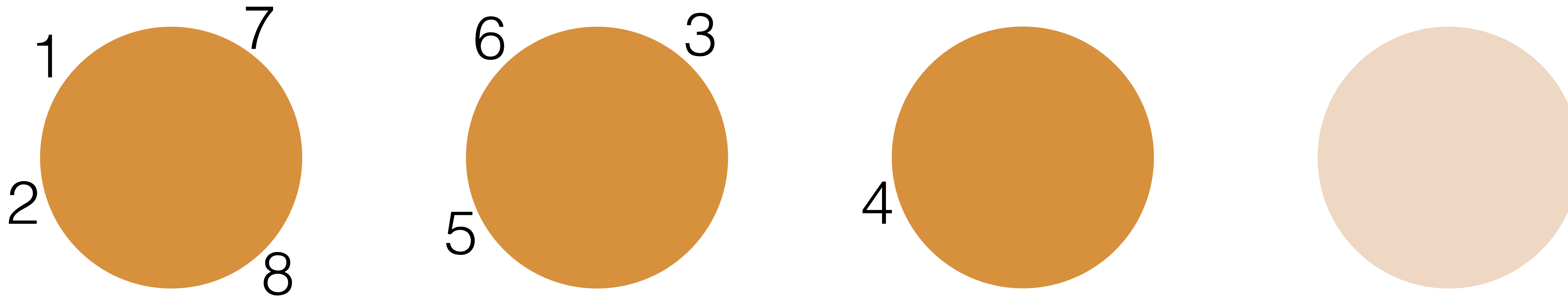
- Probability of *N* customers (K_N tables, #*C* at table *C*):

$$\frac{\alpha^{K_N} \prod_{C \in \Pi_N} (\#C - 1)!}{\alpha \cdots (\alpha + N - 1)} = \mathbb{P}(\Pi_N = \pi_N)$$

- Prob doesn't depend on customer order: *exchangeable*

$$\mathbb{P}(\Pi_8 = \{\{1,2,7,8\},\{3,5,6\},\{4\}\}) = \mathbb{P}(\Pi_8 = \{\{2,3,8,1\},\{4,6,7\},\{5\}\})$$

Chinese restaurant process



- Probability of this seating:

$$\frac{\alpha}{\alpha} \cdot \frac{1}{\alpha+1} \cdot \frac{\alpha}{\alpha+2} \cdot \frac{\alpha}{\alpha+3} \cdot \frac{1}{\alpha+4} \cdot \frac{2}{\alpha+5} \cdot \frac{2}{\alpha+6} \cdot \frac{3}{\alpha+7}$$

- Probability of N customers (K_N tables, $\#C$ at table C):

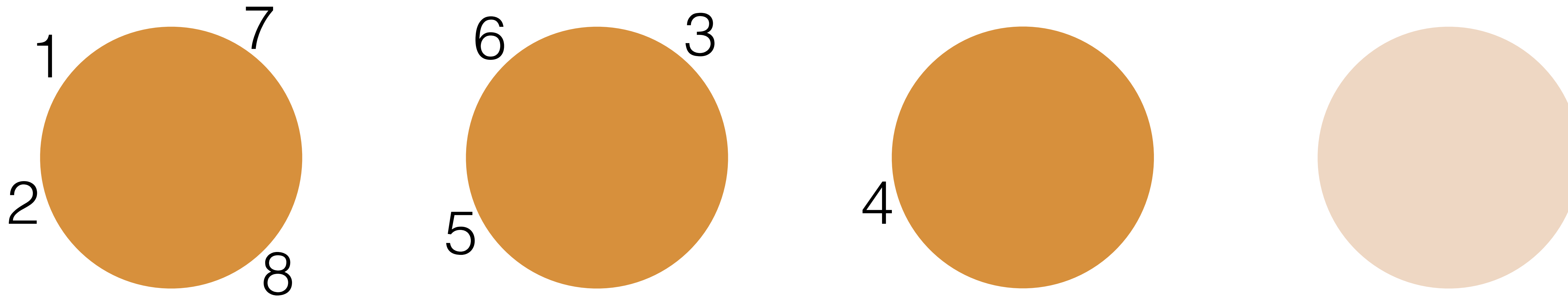
$$\frac{\alpha^{K_N} \prod_{C \in \Pi_N} (\#C - 1)!}{\alpha \cdots (\alpha + N - 1)} = \mathbb{P}(\Pi_N = \pi_N)$$

- Prob doesn't depend on customer order: *exchangeable*

$$\mathbb{P}(\Pi_8 = \{\{1,2,7,8\},\{3,5,6\},\{4\}\}) = \mathbb{P}(\Pi_8 = \{\{2,3,8,1\},\{4,6,7\},\{5\}\})$$

- Can always pretend n is the last customer and calculate $p(\Pi_N | \Pi_{N,\text{-}n})$

Chinese restaurant process



- Probability of this seating:

$$\frac{\alpha}{\alpha} \cdot \frac{1}{\alpha+1} \cdot \frac{\alpha}{\alpha+2} \cdot \frac{\alpha}{\alpha+3} \cdot \frac{1}{\alpha+4} \cdot \frac{2}{\alpha+5} \cdot \frac{2}{\alpha+6} \cdot \frac{3}{\alpha+7}$$

- Probability of *N* customers (K_N tables, #*C* at table *C*):

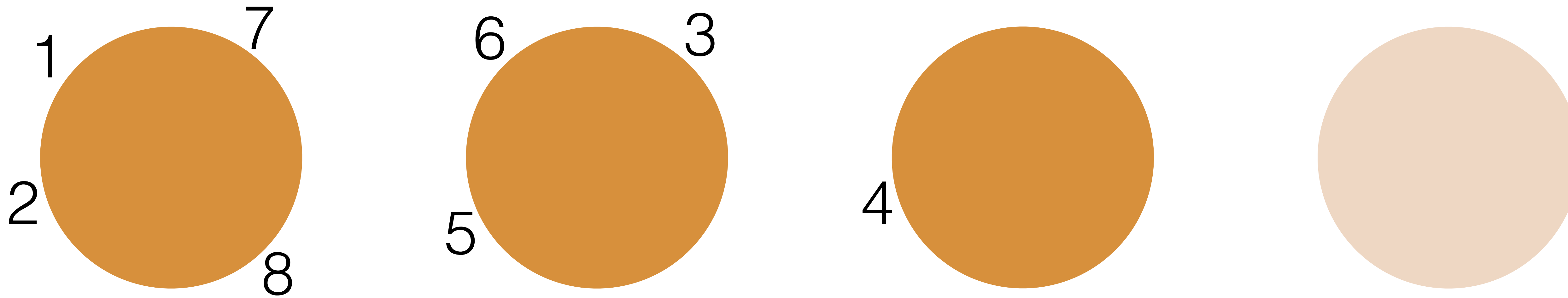
$$\frac{\alpha^{K_N} \prod_{C \in \Pi_N} (\#C - 1)!}{\alpha \cdots (\alpha + N - 1)} = \mathbb{P}(\Pi_N = \pi_N)$$

- Prob doesn't depend on customer order: *exchangeable*

$$\mathbb{P}(\Pi_8 = \{\{1,2,7,8\},\{3,5,6\},\{4\}\}) = \mathbb{P}(\Pi_8 = \{\{2,3,8,1\},\{4,6,7\},\{5\}\})$$

- Can always pretend *n* is the last customer and calculate $p(\Pi_N | \Pi_{N,\text{-}n})$
    - e.g. $\Pi_{8,\text{-}5} = \{\{1,2,7,8\},\{3,6\},\{4\}\}$

Chinese restaurant process



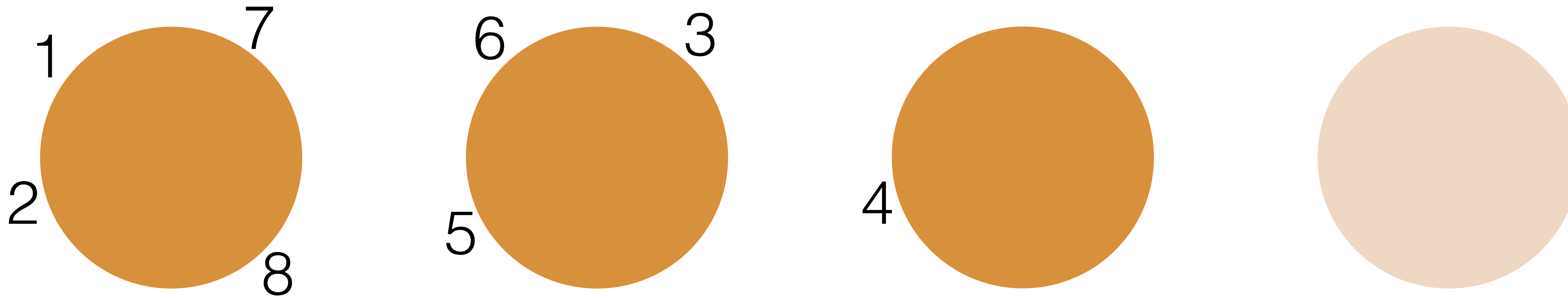
- Probability of *N* customers (K_N tables, #*C* at table *C*):

$$\frac{\alpha^{K_N} \prod_{C \in \Pi_N} (\#C - 1)!}{\alpha \cdots (\alpha + N - 1)} = \mathbb{P}(\Pi_N = \pi_N)$$

- So:

$$p(\Pi_N | \Pi_{N,-n}) =$$

Chinese restaurant process

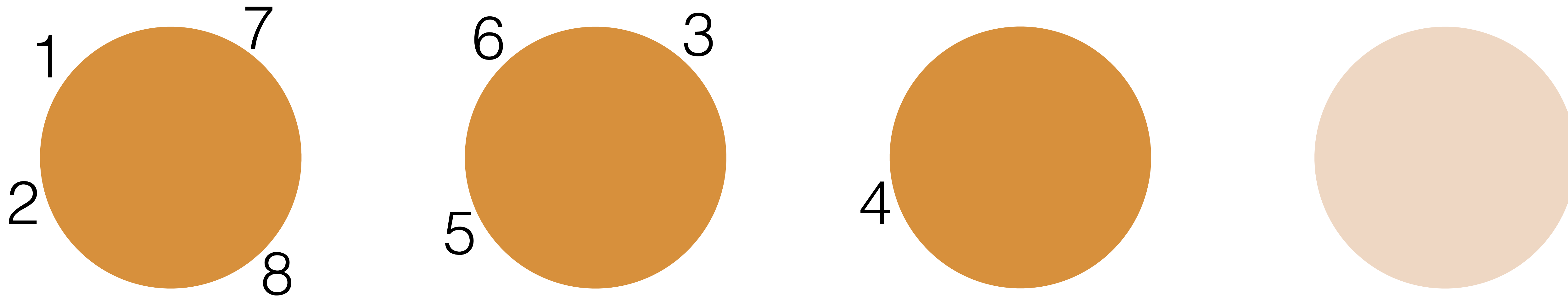


- Probability of N customers (K_N tables, $\#C$ at table C):

$$\frac{\alpha^{K_N} \prod_{C \in \Pi_N} (\#C - 1)!}{\alpha \cdots (\alpha + N - 1)} = \mathbb{P}(\Pi_N = \pi_N)$$

- So: $$p(\Pi_N | \Pi_{N,-n}) = \bigg\{$$

Chinese restaurant process



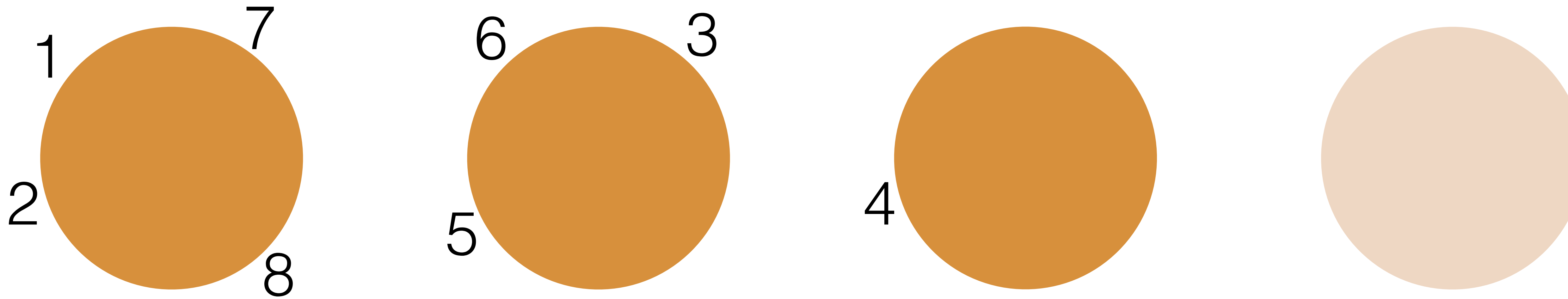
- Probability of N customers (K_N tables, $\#C$ at table C):

$$\frac{\alpha^{K_N} \prod_{C \in \Pi_N} (\#C - 1)!}{\alpha \cdots (\alpha + N - 1)} = \mathbb{P}(\Pi_N = \pi_N)$$

- So:

$$p(\Pi_N | \Pi_{N,-n}) = \begin{cases} & \text{if } n \text{ joins cluster } C \\ & \text{if } n \text{ starts a new cluster} \end{cases}$$

Chinese restaurant process



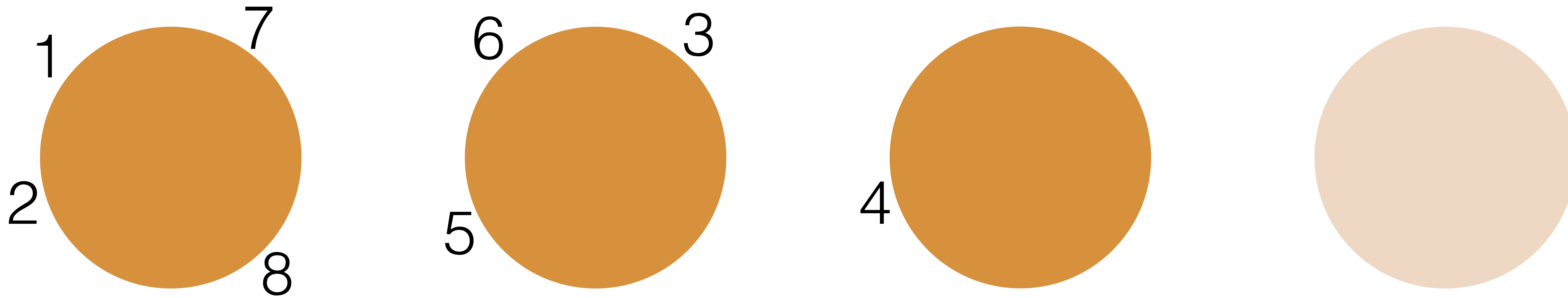
- Probability of N customers (K_N tables, $\#C$ at table C):

$$\frac{\alpha^{K_N} \prod_{C \in \Pi_N} (\#C - 1)!}{\alpha \cdots (\alpha + N - 1)} = \mathbb{P}(\Pi_N = \pi_N)$$

- So:

$$p(\Pi_N | \Pi_{N,-n}) = \begin{cases} \frac{\#C}{\alpha+N-1} & \text{if } n \text{ joins cluster } C \\ & \text{if } n \text{ starts a new cluster} \end{cases}$$

Chinese restaurant process



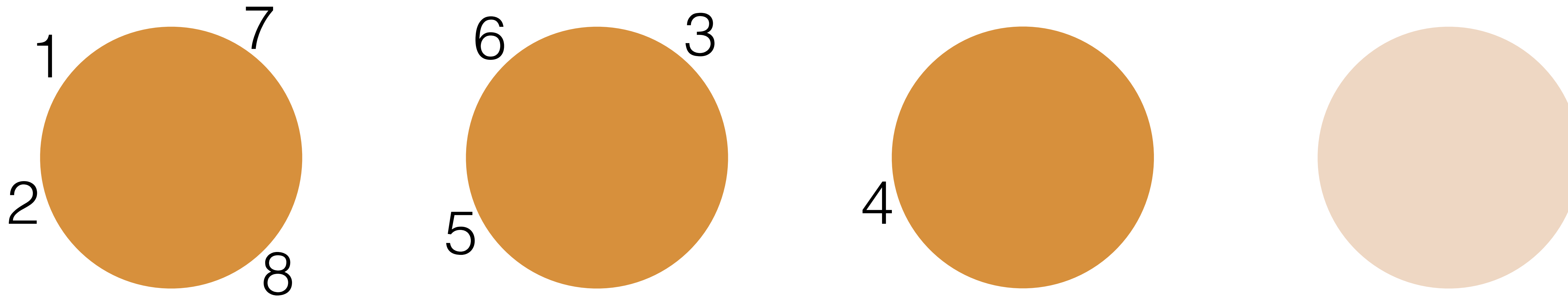
- Probability of N customers (K_N tables, $\#C$ at table C):

$$\frac{\alpha^{K_N} \prod_{C \in \Pi_N} (\#C - 1)!}{\alpha \cdots (\alpha + N - 1)} = \mathbb{P}(\Pi_N = \pi_N)$$

- So:

$$p(\Pi_N | \Pi_{N,-n}) = \begin{cases} \frac{\#C}{\alpha + N - 1} & \text{if } n \text{ joins cluster } C \\ \frac{\alpha}{\alpha + N - 1} & \text{if } n \text{ starts a new cluster} \end{cases}$$

Chinese restaurant process



- Probability of N customers (K_N tables, $\#C$ at table C):

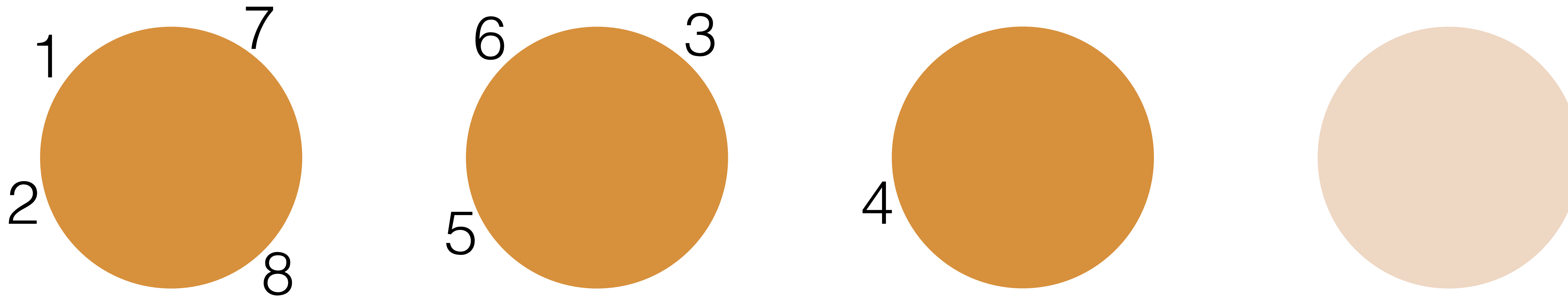
$$\frac{\alpha^{K_N} \prod_{C \in \Pi_N} (\#C - 1)!}{\alpha \cdots (\alpha + N - 1)} = \mathbb{P}(\Pi_N = \pi_N)$$

- So:

$$p(\Pi_N | \Pi_{N,-n}) = \begin{cases} \frac{\#C}{\alpha+N-1} & \text{if } n \text{ joins cluster } C \\ \frac{\alpha}{\alpha+N-1} & \text{if } n \text{ starts a new cluster} \end{cases}$$

- Gibbs sampling review:

Chinese restaurant process



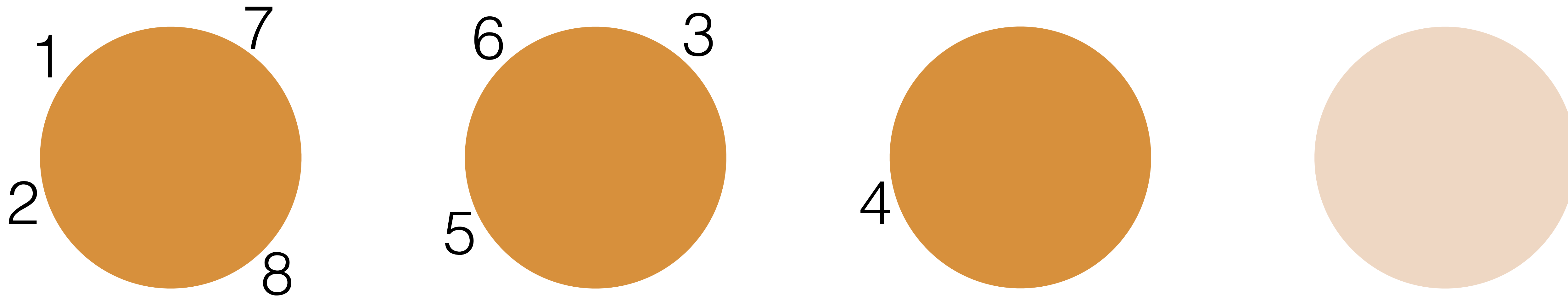
- Probability of N customers (K_N tables, #C at table C):

$$\frac{\alpha^{K_N} \prod_{C \in \Pi_N} (\#C - 1)!}{\alpha \cdots (\alpha + N - 1)} = \mathbb{P}(\Pi_N = \pi_N)$$

- So: $$p(\Pi_N | \Pi_{N,-n}) = \begin{cases} \frac{\#C}{\alpha + N - 1} & \text{if } n \text{ joins cluster } C \\ \frac{\alpha}{\alpha + N - 1} & \text{if } n \text{ starts a new cluster} \end{cases}$$

- Gibbs sampling review: target distribution $p(v_1, v_2, v_3)$

Chinese restaurant process



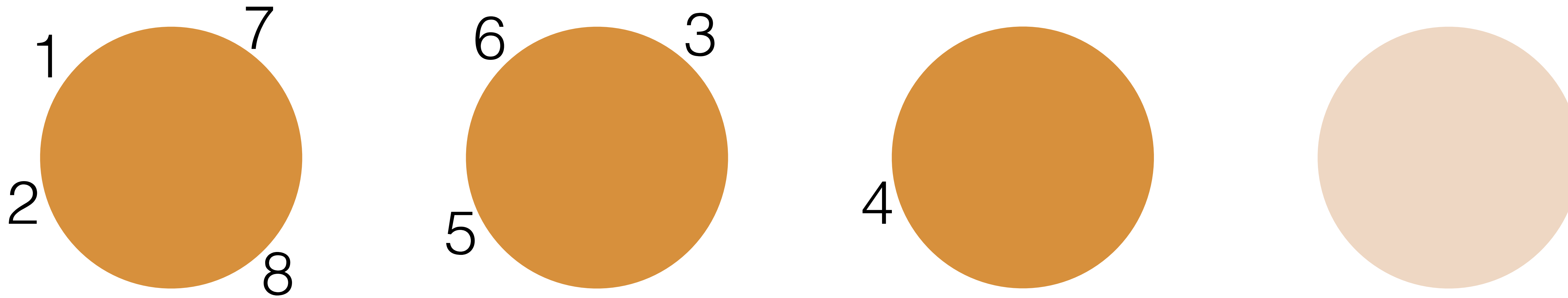
- Probability of *N* customers (K_N tables, #*C* at table *C*):

$$\frac{\alpha^{K_N} \prod_{C \in \Pi_N} (\#C - 1)!}{\alpha \cdots (\alpha + N - 1)} = \mathbb{P}(\Pi_N = \pi_N)$$

- So: $$p(\Pi_N | \Pi_{N,-n}) = \begin{cases} \frac{\#C}{\alpha+N-1} & \text{if } n \text{ joins cluster } C \\ \frac{\alpha}{\alpha+N-1} & \text{if } n \text{ starts a new cluster} \end{cases}$$

- Gibbs sampling review: target distribution $p(v_1, v_2, v_3)$
  - Start: $v_1^{(0)}, v_2^{(0)}, v_3^{(0)}$

Chinese restaurant process



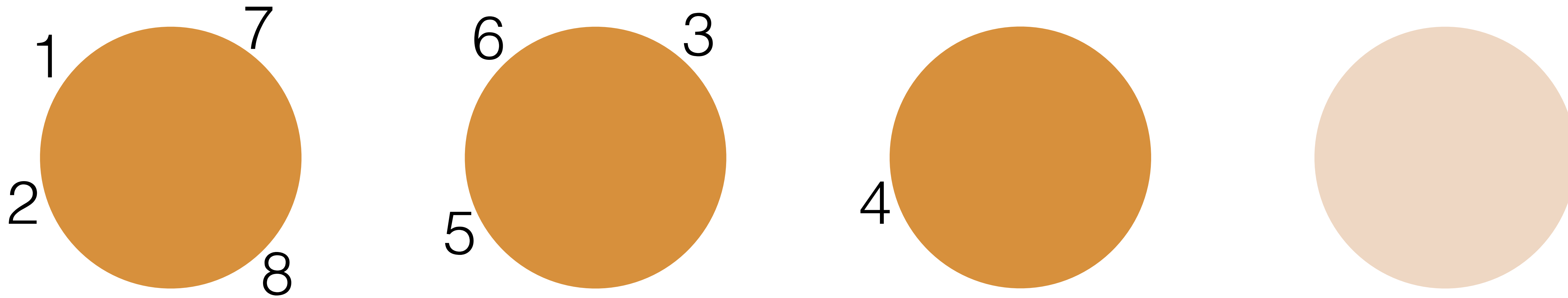
- Probability of N customers (K_N tables, $\#C$ at table C):

$$\frac{\alpha^{K_N} \prod_{C \in \Pi_N} (\#C - 1)!}{\alpha \cdots (\alpha + N - 1)} = \mathbb{P}(\Pi_N = \pi_N)$$

- So: $$p(\Pi_N | \Pi_{N,-n}) = \begin{cases} \frac{\#C}{\alpha+N-1} & \text{if } n \text{ joins cluster } C \\ \frac{\alpha}{\alpha+N-1} & \text{if } n \text{ starts a new cluster} \end{cases}$$

- Gibbs sampling review: target distribution $p(v_1, v_2, v_3)$
  - Start: $v_1^{(0)}, v_2^{(0)}, v_3^{(0)}$
  - t th step: $v_1^{(t)} \sim p(v_1 | v_2^{(t-1)}, v_3^{(t-1)})$

Chinese restaurant process



- Probability of N customers (K_N tables, $\#C$ at table C):

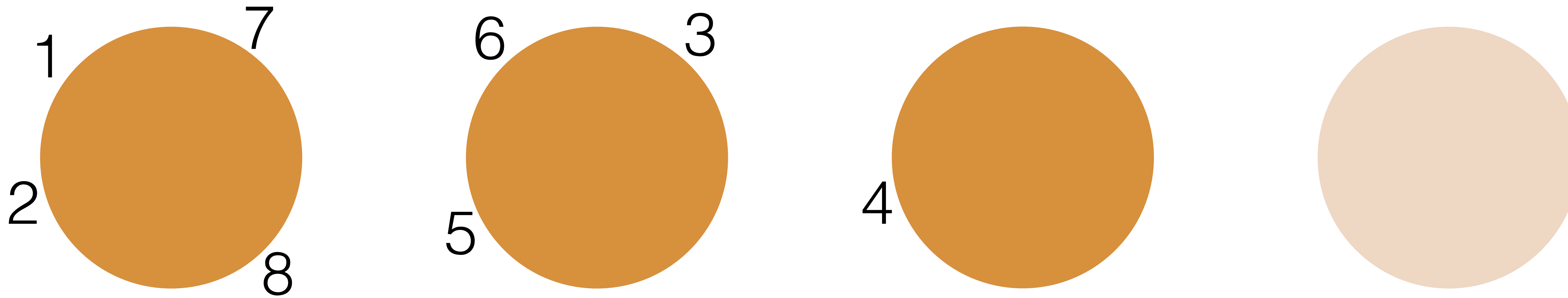
$$\frac{\alpha^{K_N} \prod_{C \in \Pi_N} (\#C - 1)!}{\alpha \cdots (\alpha + N - 1)} = \mathbb{P}(\Pi_N = \pi_N)$$

- So: $$p(\Pi_N | \Pi_{N,-n}) = \begin{cases} \frac{\#C}{\alpha+N-1} & \text{if } n \text{ joins cluster } C \\ \frac{\alpha}{\alpha+N-1} & \text{if } n \text{ starts a new cluster} \end{cases}$$

- Gibbs sampling review: target distribution $p(v_1, v_2, v_3)$
  - Start: $v_1^{(0)}, v_2^{(0)}, v_3^{(0)}$
  - t th step: $v_1^{(t)} \sim p(v_1 | v_2^{(t-1)}, v_3^{(t-1)})$

$$v_2^{(t)} \sim p(v_2 | v_1^{(t)}, v_3^{(t-1)})$$

Chinese restaurant process



- Probability of N customers (K_N tables, #C at table C):
  $$\frac{\alpha^{K_N} \prod_{C \in \Pi_N} (\#C - 1)!}{\alpha \cdots (\alpha + N - 1)} = \mathbb{P}(\Pi_N = \pi_N)$$
- So: $$p(\Pi_N | \Pi_{N,-n}) = \begin{cases} \frac{\#C}{\alpha+N-1} & \text{if } n \text{ joins cluster } C \\ \frac{\alpha}{\alpha+N-1} & \text{if } n \text{ starts a new cluster} \end{cases}$$
- Gibbs sampling review: target distribution $p(v_1, v_2, v_3)$
  - Start: $v_1^{(0)}, v_2^{(0)}, v_3^{(0)}$
  - t th step: $v_1^{(t)} \sim p(v_1 | v_2^{(t-1)}, v_3^{(t-1)})$

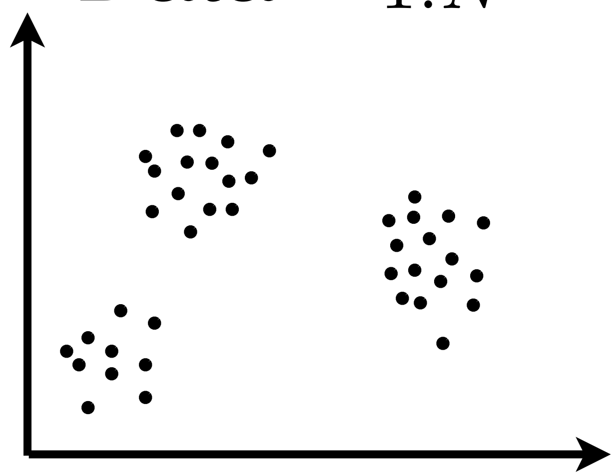
    $v_2^{(t)} \sim p(v_2 | v_1^{(t)}, v_3^{(t-1)})$

    $v_3^{(t)} \sim p(v_3 | v_1^{(t)}, v_2^{(t)})$

CRP mixture model: inference

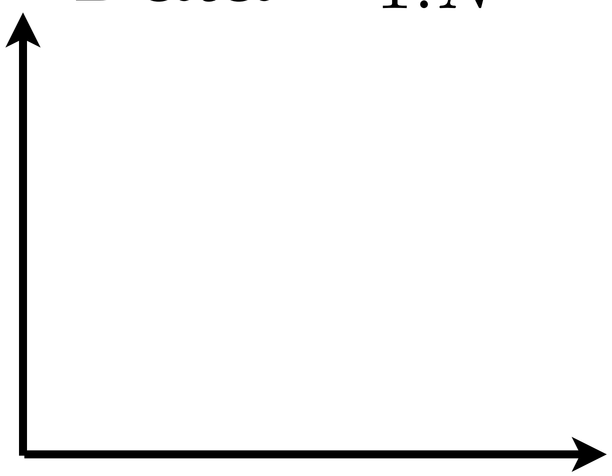
CRP mixture model: inference

- Data $x_{1:N}$



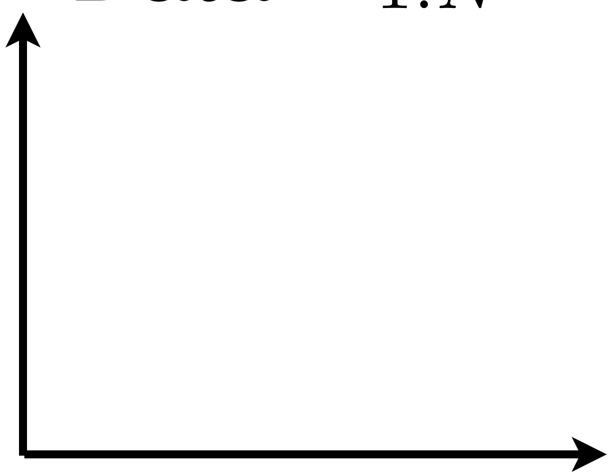
CRP mixture model: inference

- Data $x_{1:N}$



CRP mixture model: inference

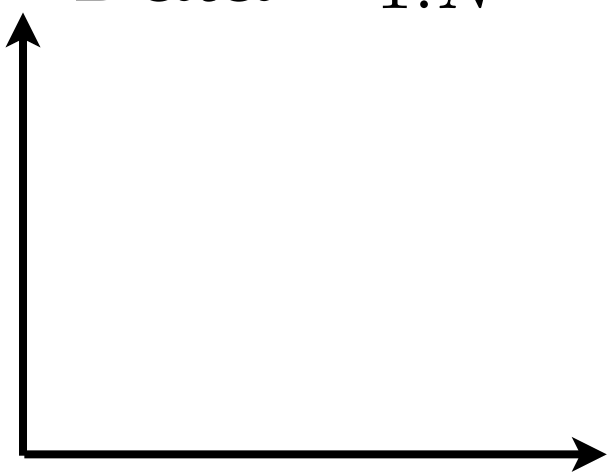
- Data $x_{1:N}$



- Generative model

CRP mixture model: inference

- Data $x_{1:N}$

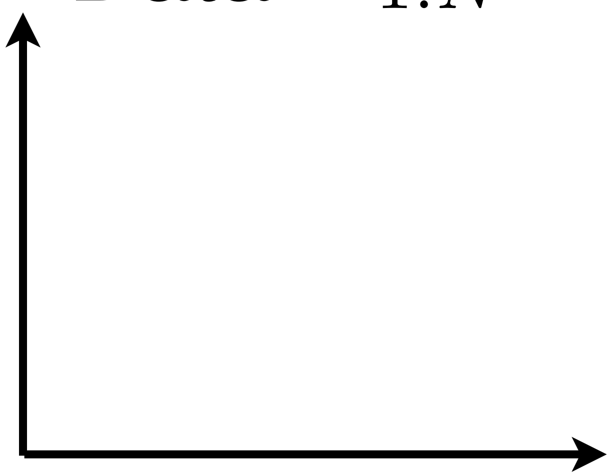


- Generative model

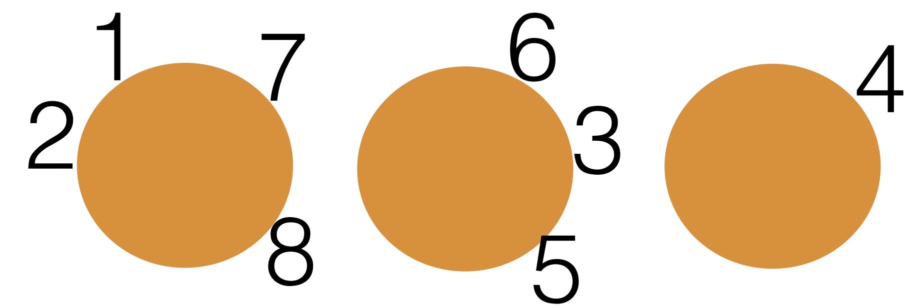
$$\Pi_N \sim \mathrm{CRP}(N, \alpha)$$

CRP mixture model: inference

- Data $x_{1:N}$



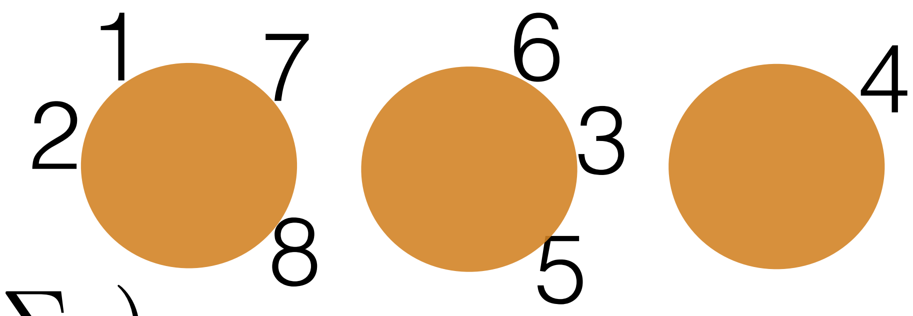
- Generative model
  $\Pi_N \sim \mathrm{CRP}(N, \alpha)$



CRP mixture model: inference

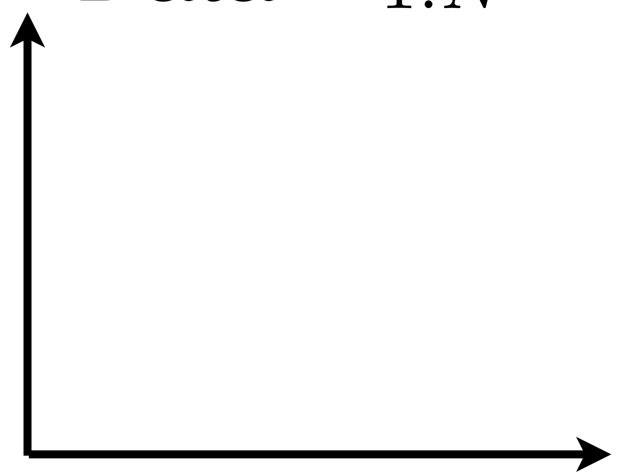
- Data $x_{1:N}$

- Generative model
$\Pi_N \sim \mathrm{CRP}(N, \alpha)$
$\forall C \in \Pi_N, \mu_C \stackrel{iid}{\sim} \mathcal{N}(\mu_0, \Sigma_0)$



CRP mixture model: inference

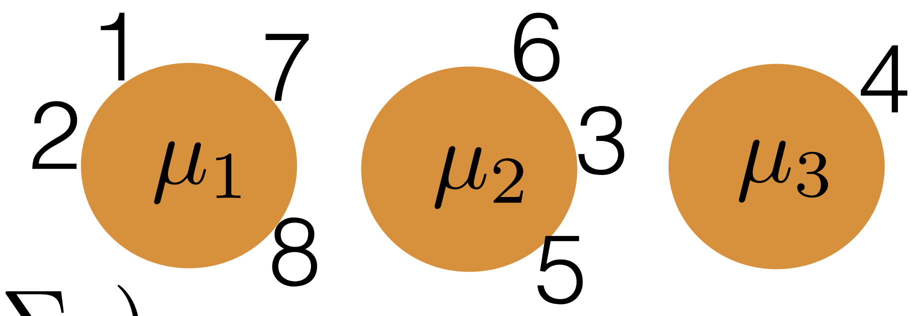
- Data $x_{1:N}$



- Generative model

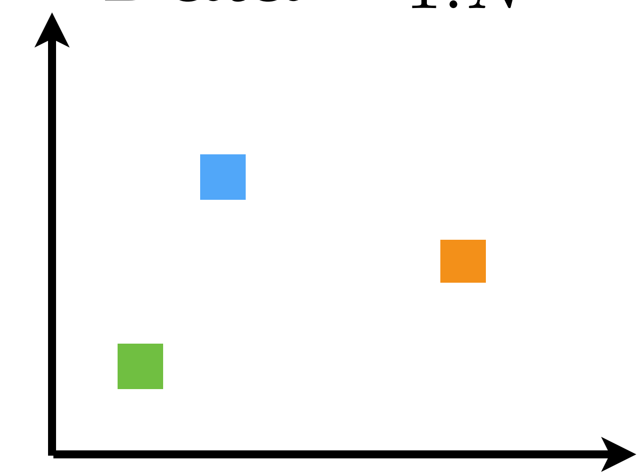
$$\Pi_N \sim \mathrm{CRP}(N, \alpha)$$

$$\forall C \in \Pi_N, \mu_C \stackrel{iid}{\sim} \mathcal{N}(\mu_0, \Sigma_0)$$



CRP mixture model: inference

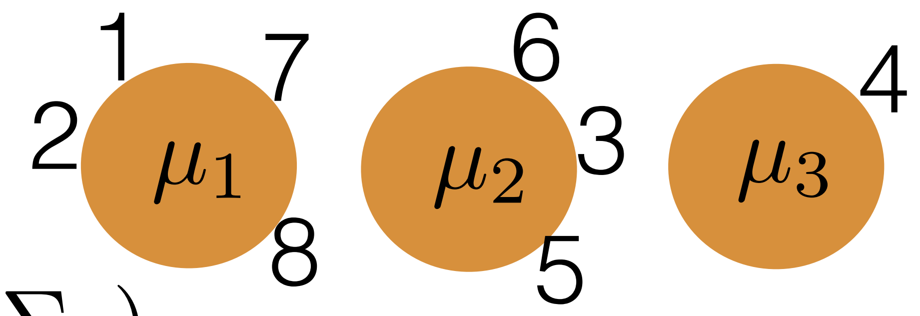
- Data $x_{1:N}$



- Generative model

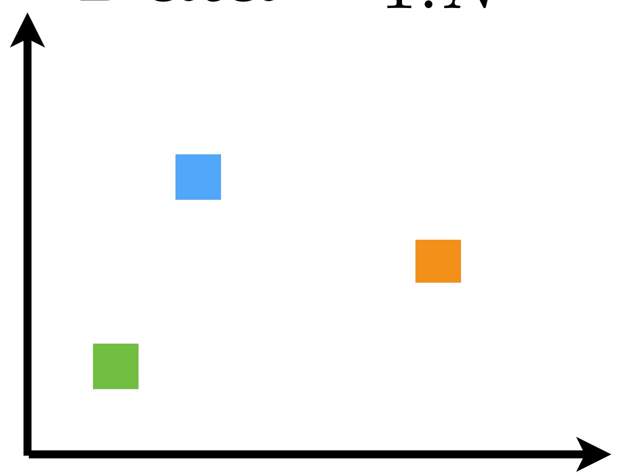
$$\Pi_N \sim \mathrm{CRP}(N, \alpha)$$

$$\forall C \in \Pi_N, \mu_C \overset{iid}{\sim} \mathcal{N}(\mu_0, \Sigma_0)$$



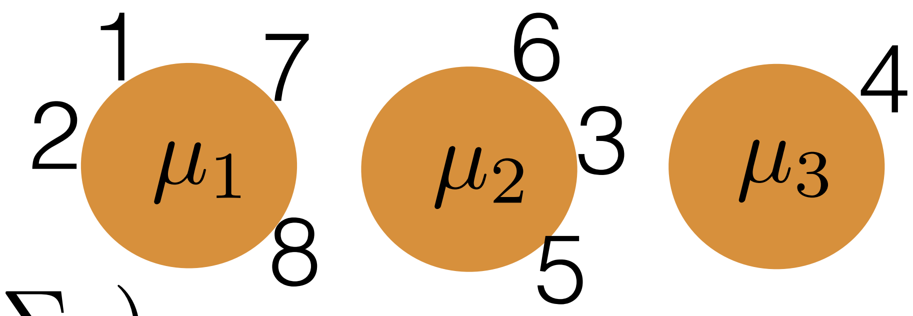
CRP mixture model: inference

- Data $x_{1:N}$



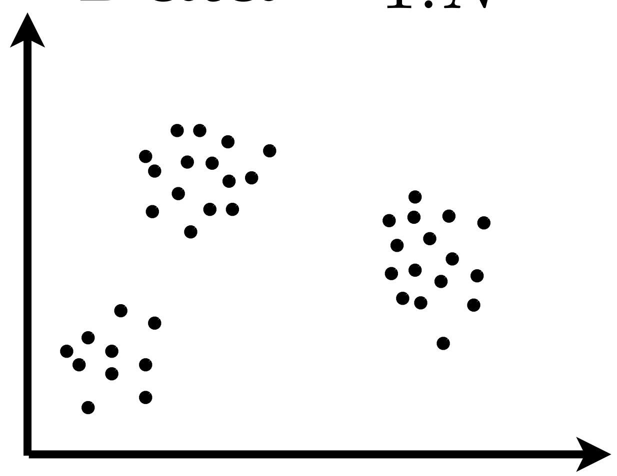
- Generative model

$$\Pi_N \sim \mathrm{CRP}(N, \alpha)$$
$$\forall C \in \Pi_N, \mu_C \overset{iid}{\sim} \mathcal{N}(\mu_0, \Sigma_0)$$
$$\forall C \in \Pi_N, \forall n \in C, x_n \overset{indep}{\sim} \mathcal{N}(\mu_C, \Sigma)$$



CRP mixture model: inference

- Data $x_{1:N}$

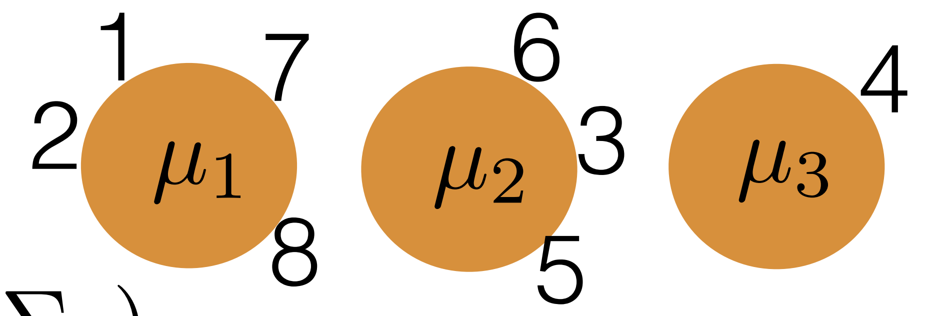


- Generative model

$$\Pi_N \sim \mathrm{CRP}(N, \alpha)$$

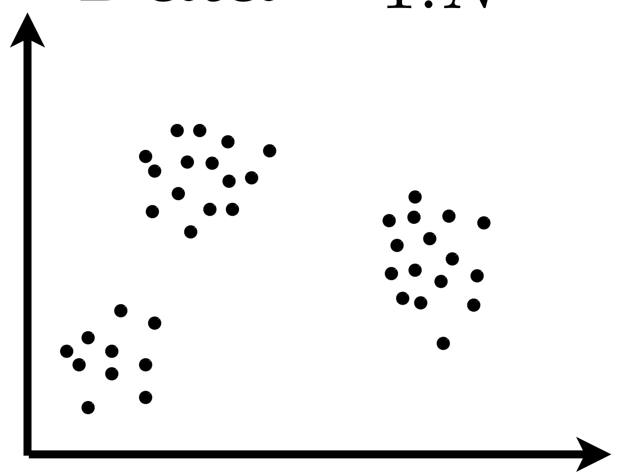
$$\forall C \in \Pi_N, \mu_C \overset{iid}{\sim} \mathcal{N}(\mu_0, \Sigma_0)$$

$$\forall C \in \Pi_N, \forall n \in C, x_n \overset{indep}{\sim} \mathcal{N}(\mu_C, \Sigma)$$



CRP mixture model: inference

- Data $x_{1:N}$

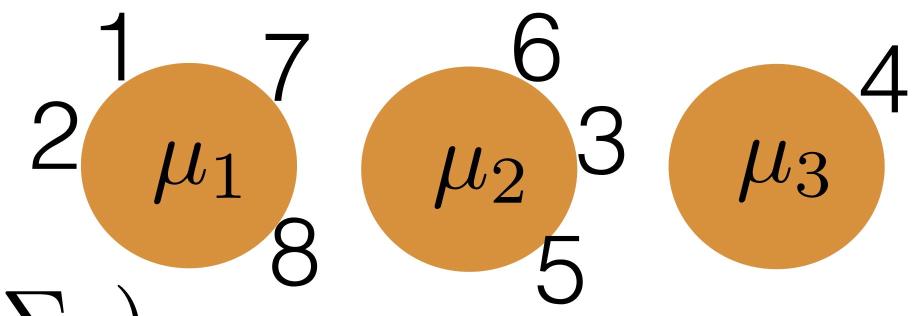


- Generative model

$$\Pi_N \sim \mathrm{CRP}(N, \alpha)$$

$$\forall C \in \Pi_N, \mu_C \overset{iid}{\sim} \mathcal{N}(\mu_0, \Sigma_0)$$

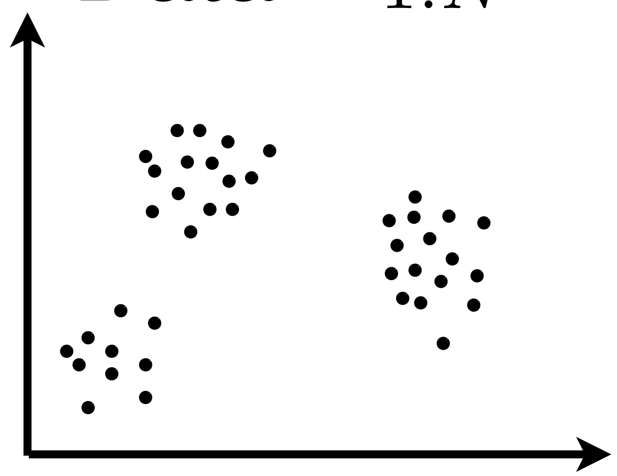
$$\forall C \in \Pi_N, \forall n \in C, x_n \overset{indep}{\sim} \mathcal{N}(\mu_C, \Sigma)$$



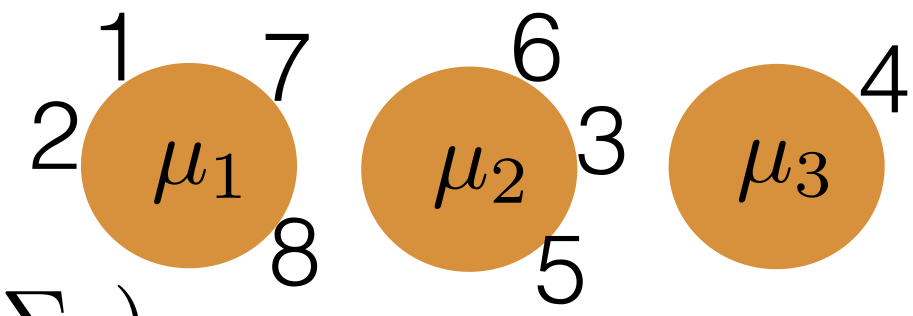
- Want: posterior

CRP mixture model: inference

- Data $x_{1:N}$



- Generative model



$$\Pi_N \sim \mathrm{CRP}(N, \alpha)$$

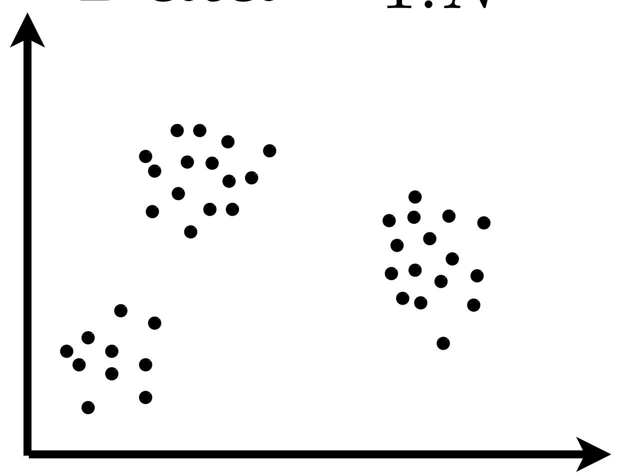
$$\forall C \in \Pi_N, \mu_C \overset{iid}{\sim} \mathcal{N}(\mu_0, \Sigma_0)$$

$$\forall C \in \Pi_N, \forall n \in C, x_n \overset{indep}{\sim} \mathcal{N}(\mu_C, \Sigma)$$

- Want: posterior $p(\Pi_N | x_{1:N})$

CRP mixture model: inference

- Data $x_{1:N}$

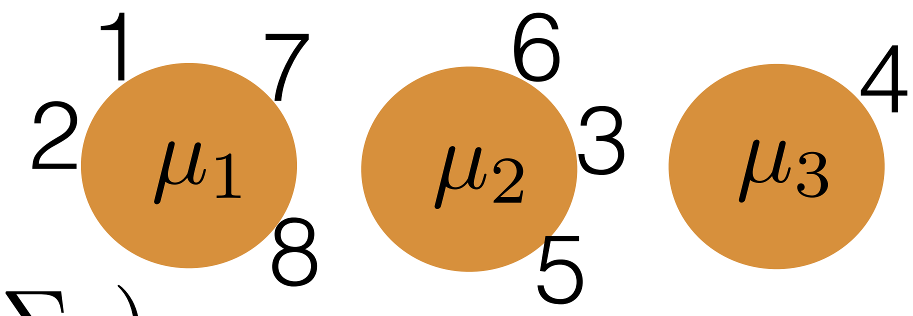


- Generative model

$$\Pi_N \sim \mathrm{CRP}(N, \alpha)$$

$$\forall C \in \Pi_N, \mu_C \overset{iid}{\sim} \mathcal{N}(\mu_0, \Sigma_0)$$

$$\forall C \in \Pi_N, \forall n \in C, x_n \overset{indep}{\sim} \mathcal{N}(\mu_C, \Sigma)$$

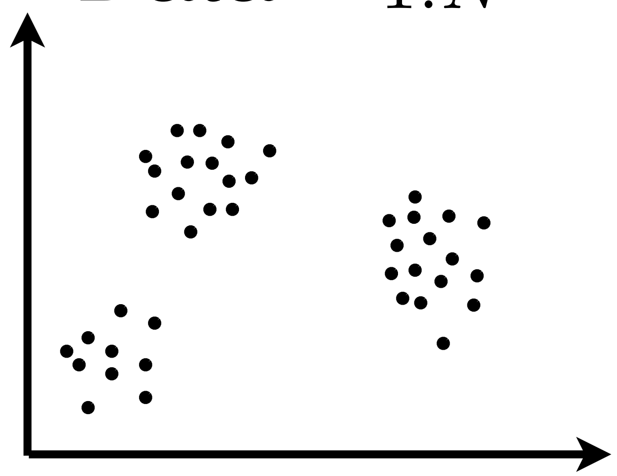


- Want: posterior $p(\Pi_N | x_{1:N})$

- Gibbs sampler:

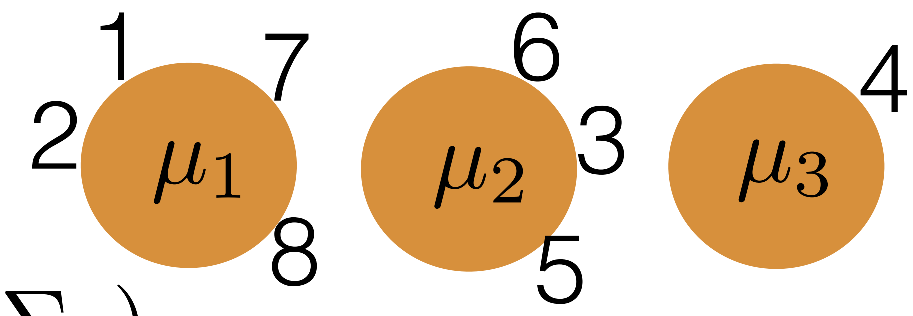
CRP mixture model: inference

- Data $x_{1:N}$



- Generative model

$$\Pi_N \sim \mathrm{CRP}(N, \alpha)$$
$$\forall C \in \Pi_N, \mu_C \overset{iid}{\sim} \mathcal{N}(\mu_0, \Sigma_0)$$
$$\forall C \in \Pi_N, \forall n \in C, x_n \overset{indep}{\sim} \mathcal{N}(\mu_C, \Sigma)$$



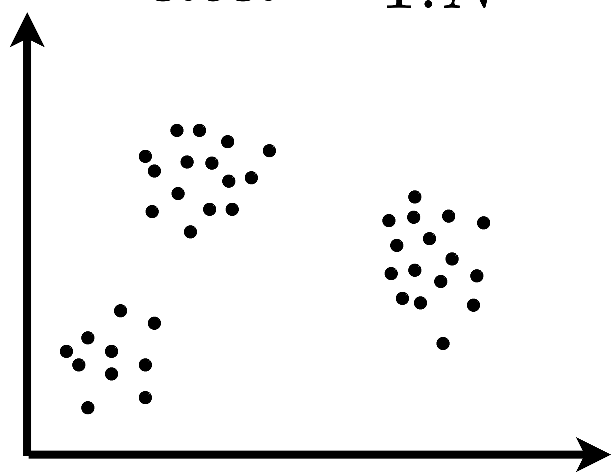
- Want: posterior $p(\Pi_N | x_{1:N})$

- Gibbs sampler:

$$p(\Pi_N | \Pi_{N,\text{-}n}, x)$$

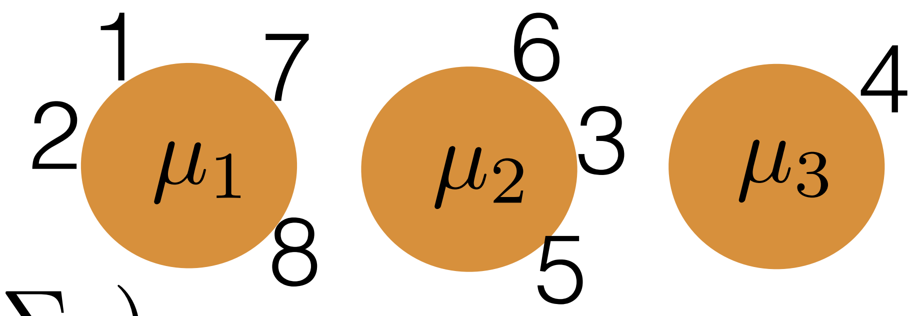
CRP mixture model: inference

- Data $x_{1:N}$



- Generative model

$$\Pi_N \sim \mathrm{CRP}(N, \alpha)$$
$$\forall C \in \Pi_N, \mu_C \stackrel{iid}{\sim} \mathcal{N}(\mu_0, \Sigma_0)$$
$$\forall C \in \Pi_N, \forall n \in C, x_n \stackrel{indep}{\sim} \mathcal{N}(\mu_C, \Sigma)$$



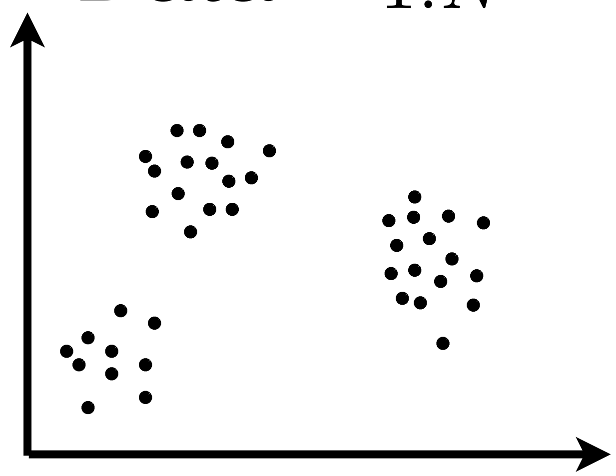
- Want: posterior $p(\Pi_N | x_{1:N})$

- Gibbs sampler:

$$p(\Pi_N | \Pi_{N,\text{-}n}, x) = \Bigg\{$$

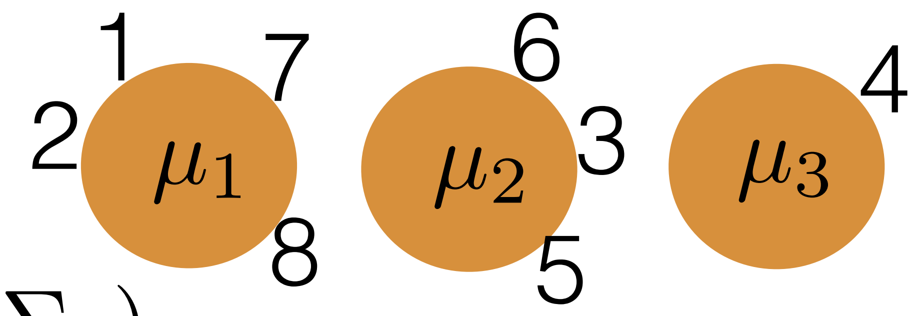
CRP mixture model: inference

- Data $x_{1:N}$



- Generative model

$$\Pi_N \sim \mathrm{CRP}(N, \alpha)$$
$$\forall C \in \Pi_N, \mu_C \stackrel{iid}{\sim} \mathcal{N}(\mu_0, \Sigma_0)$$
$$\forall C \in \Pi_N, \forall n \in C, x_n \stackrel{indep}{\sim} \mathcal{N}(\mu_C, \Sigma)$$



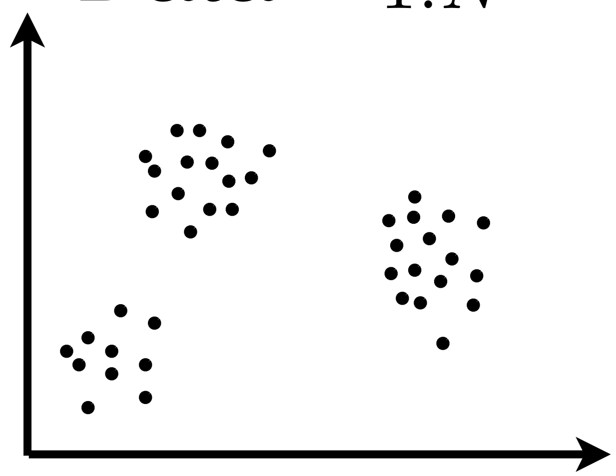
- Want: posterior $p(\Pi_N | x_{1:N})$

- Gibbs sampler:

$$p(\Pi_N | \Pi_{N,\text{-}n}, x) = \Bigg\{ \qquad \text{if } n \text{ joins cluster } C$$

CRP mixture model: inference

- Data $x_{1:N}$

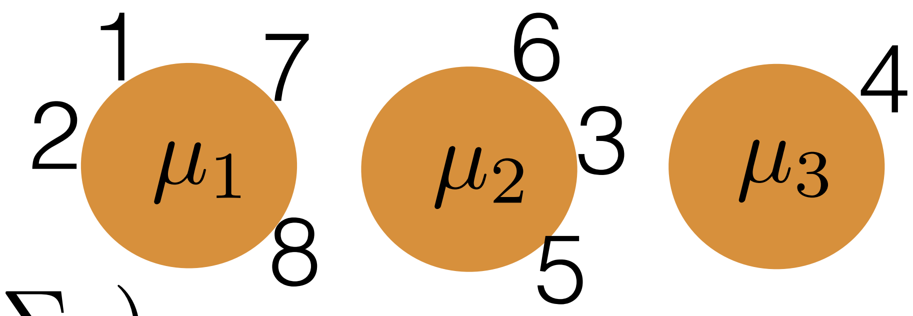


- Generative model

$$\Pi_N \sim \mathrm{CRP}(N, \alpha)$$

$$\forall C \in \Pi_N, \mu_C \overset{iid}{\sim} \mathcal{N}(\mu_0, \Sigma_0)$$

$$\forall C \in \Pi_N, \forall n \in C, x_n \overset{indep}{\sim} \mathcal{N}(\mu_C, \Sigma)$$



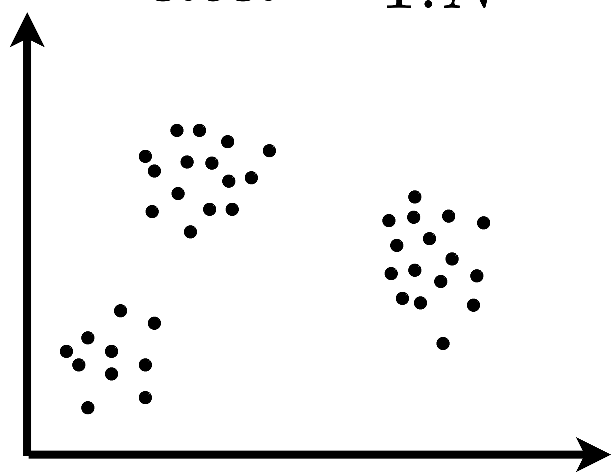
- Want: posterior $p(\Pi_N | x_{1:N})$

- Gibbs sampler:

$$p(\Pi_N | \Pi_{N,\text{-}n}, x) = \begin{cases} & \text{if } n \text{ joins cluster } C \\ & \text{if } n \text{ starts a new cluster} \end{cases}$$

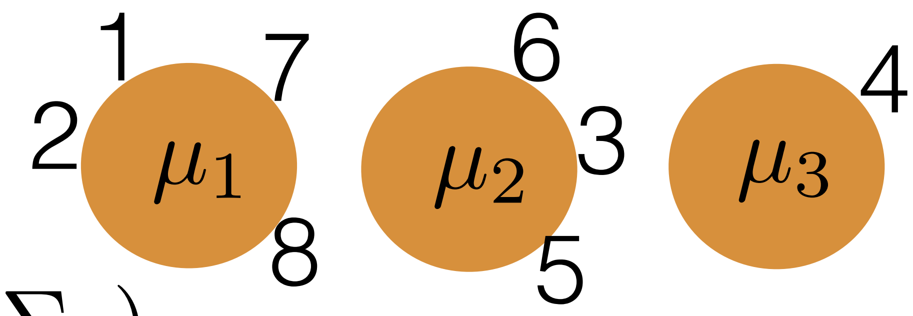
CRP mixture model: inference

- Data $x_{1:N}$



- Generative model

$$\Pi_N \sim \mathrm{CRP}(N, \alpha)$$
$$\forall C \in \Pi_N, \mu_C \overset{iid}{\sim} \mathcal{N}(\mu_0, \Sigma_0)$$
$$\forall C \in \Pi_N, \forall n \in C, x_n \overset{indep}{\sim} \mathcal{N}(\mu_C, \Sigma)$$

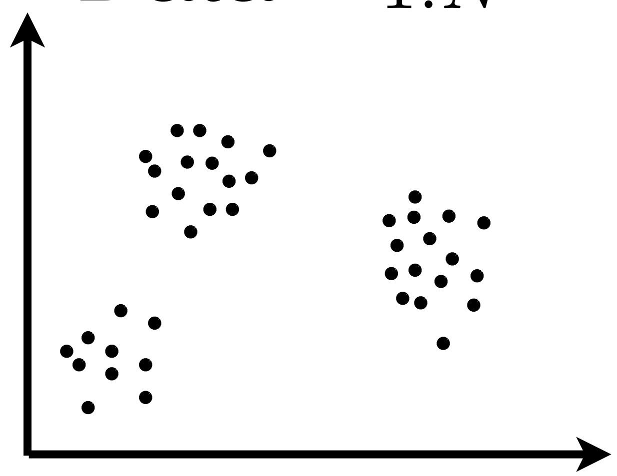


- Want: posterior $p(\Pi_N | x_{1:N})$
- Gibbs sampler:

$$p(\Pi_N | \Pi_{N,\text{-}n}, x) = \begin{cases} \frac{\#C}{\alpha+N-1} p(x_{C \cup \{n\}} | x_C) & \text{if } n \text{ joins cluster } C \\ & \text{if } n \text{ starts a new cluster} \end{cases}$$

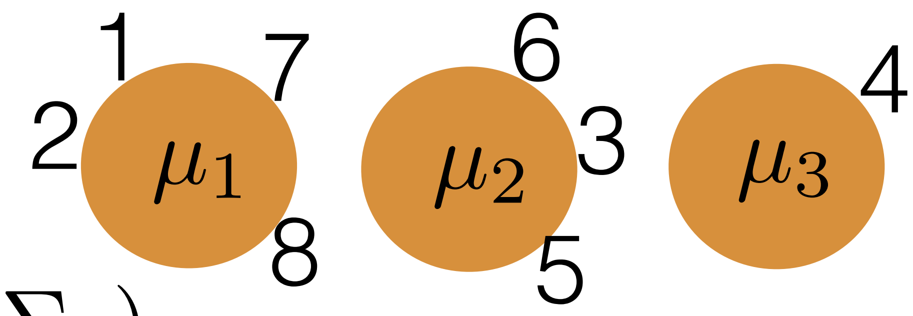
CRP mixture model: inference

- Data $x_{1:N}$



- Generative model

$$\Pi_N \sim \mathrm{CRP}(N, \alpha)$$
$$\forall C \in \Pi_N, \mu_C \stackrel{iid}{\sim} \mathcal{N}(\mu_0, \Sigma_0)$$
$$\forall C \in \Pi_N, \forall n \in C, x_n \stackrel{indep}{\sim} \mathcal{N}(\mu_C, \Sigma)$$



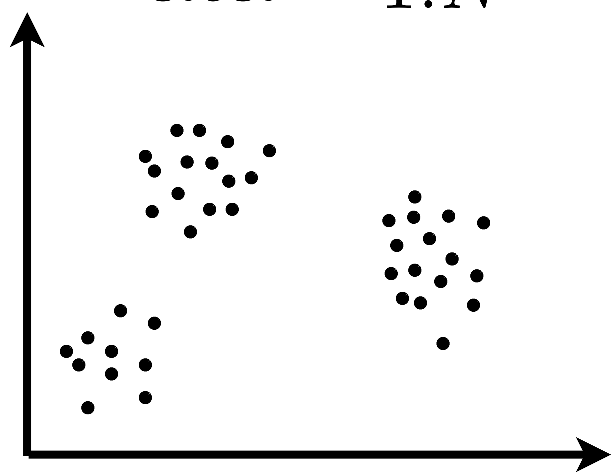
- Want: posterior $p(\Pi_N | x_{1:N})$

- Gibbs sampler:

$$p(\Pi_N | \Pi_{N,\text{-}n}, x) = \begin{cases} \frac{\#C}{\alpha+N-1} p(x_{C\cup\{n\}} | x_C) & \text{if } n \text{ joins cluster } C \\ \frac{\alpha}{\alpha+N-1} p(x_{\{n\}}) & \text{if } n \text{ starts a new cluster} \end{cases}$$

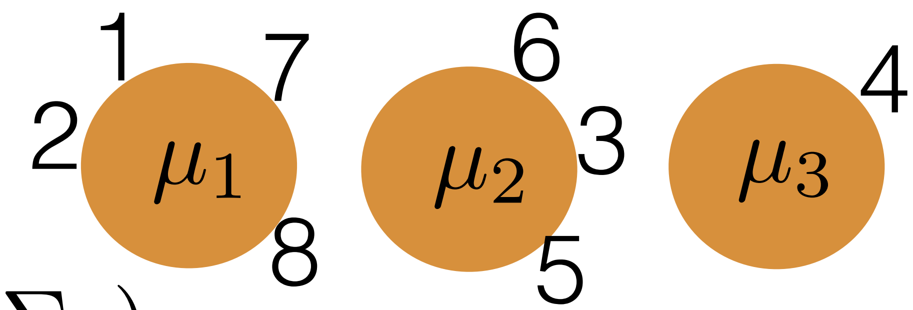
CRP mixture model: inference

- Data $x_{1:N}$



- Generative model

$$\Pi_N \sim \mathrm{CRP}(N, \alpha)$$
$$\forall C \in \Pi_N, \mu_C \overset{iid}{\sim} \mathcal{N}(\mu_0, \Sigma_0)$$
$$\forall C \in \Pi_N, \forall n \in C, x_n \overset{indep}{\sim} \mathcal{N}(\mu_C, \Sigma)$$



- Want: posterior $p(\Pi_N | x_{1:N})$

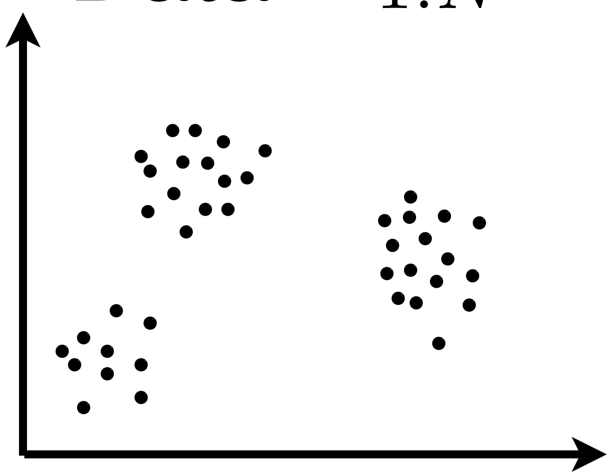
- Gibbs sampler:

$$p(\Pi_N | \Pi_{N,\text{-}n}, x) = \begin{cases} \frac{\#C}{\alpha + N - 1} p(x_{C \cup \{n\}} | x_C) & \text{if } n \text{ joins cluster } C \\ \frac{\alpha}{\alpha + N - 1} p(x_{\{n\}}) & \text{if } n \text{ starts a new cluster} \end{cases}$$

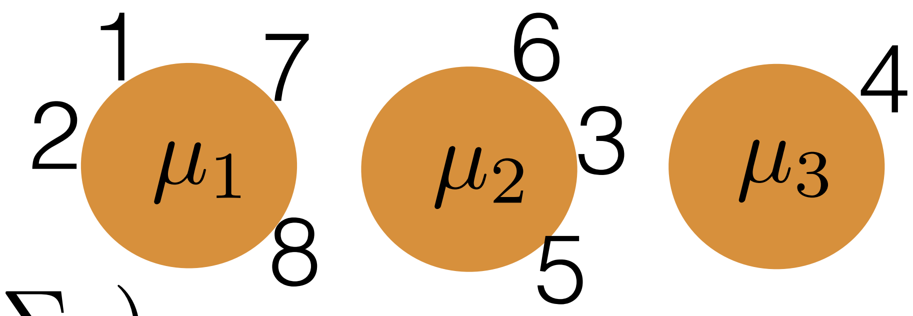
- For completeness: $p(x_{C \cup \{n\}} | x_C) =$

CRP mixture model: inference

- Data $x_{1:N}$



- Generative model



$$\Pi_N \sim \mathrm{CRP}(N, \alpha)$$
$$\forall C \in \Pi_N, \mu_C \overset{iid}{\sim} \mathcal{N}(\mu_0, \Sigma_0)$$
$$\forall C \in \Pi_N, \forall n \in C, x_n \overset{indep}{\sim} \mathcal{N}(\mu_C, \Sigma)$$

- Want: posterior $p(\Pi_N | x_{1:N})$

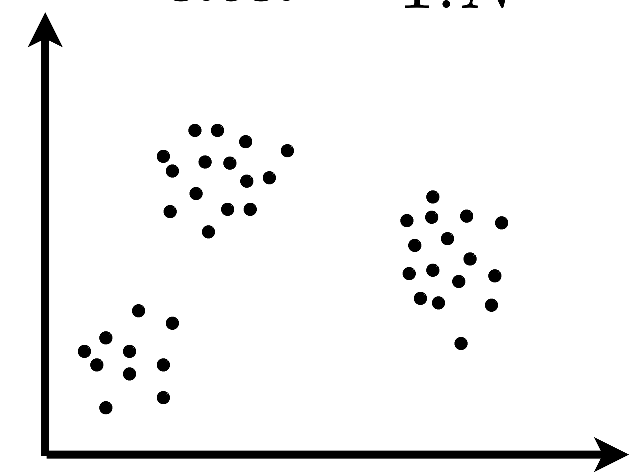
- Gibbs sampler:

$$p(\Pi_N | \Pi_{N,\text{-}n}, x) = \begin{cases} \frac{\#C}{\alpha + N - 1} p(x_{C \cup \{n\}} | x_C) & \text{if } n \text{ joins cluster } C \\ \frac{\alpha}{\alpha + N - 1} p(x_{\{n\}}) & \text{if } n \text{ starts a new cluster} \end{cases}$$

- For completeness: $p(x_{C \cup \{n\}} | x_C) = \mathcal{N}(\tilde{m}, \tilde{\Sigma} + \Sigma)$

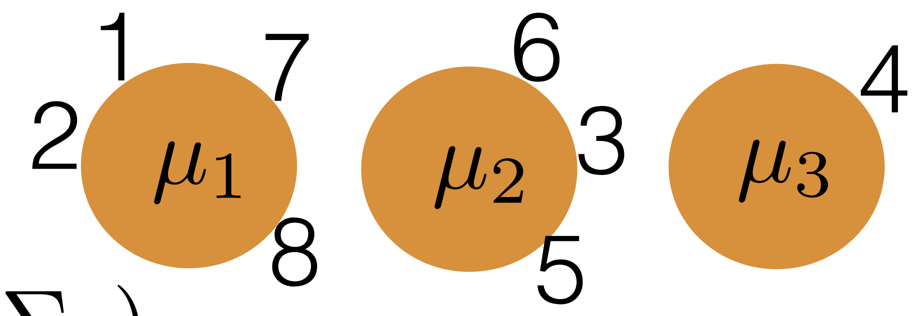
CRP mixture model: inference

- Data $x_{1:N}$



- Generative model

$$\Pi_N \sim \mathrm{CRP}(N, \alpha)$$
$$\forall C \in \Pi_N, \mu_C \overset{iid}{\sim} \mathcal{N}(\mu_0, \Sigma_0)$$
$$\forall C \in \Pi_N, \forall n \in C, x_n \overset{indep}{\sim} \mathcal{N}(\mu_C, \Sigma)$$



- Want: posterior $p(\Pi_N | x_{1:N})$

- Gibbs sampler:

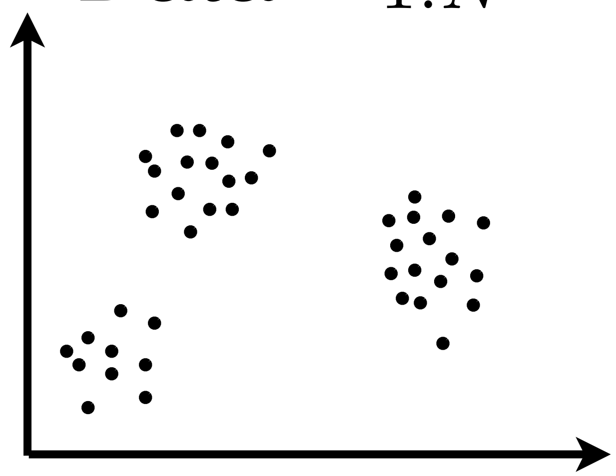
$$p(\Pi_N | \Pi_{N,\text{-}n}, x) = \begin{cases} \frac{\#C}{\alpha + N - 1} p(x_{C \cup \{n\}} | x_C) & \text{if } n \text{ joins cluster } C \\ \frac{\alpha}{\alpha + N - 1} p(x_{\{n\}}) & \text{if } n \text{ starts a new cluster} \end{cases}$$

- For completeness: $p(x_{C \cup \{n\}} | x_C) = \mathcal{N}(\tilde{m}, \tilde{\Sigma} + \Sigma)$

$$\tilde{\Sigma}^{-1} := \Sigma_0^{-1} + (\#C)\Sigma^{-1}$$
$$\tilde{m} := \tilde{\Sigma}\left(\Sigma^{-1} \sum_{m \in C} x_m + \Sigma_0^{-1} \mu_0\right)$$

CRP mixture model: inference

- Data $x_{1:N}$

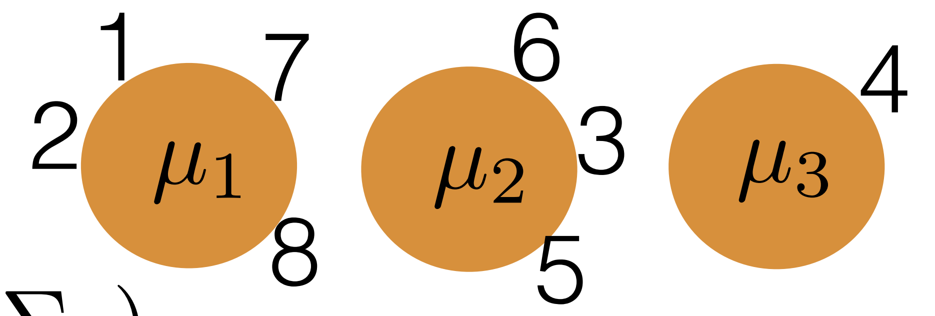


- Generative model

$$\Pi_N \sim \mathrm{CRP}(N, \alpha)$$

$$\forall C \in \Pi_N, \mu_C \overset{iid}{\sim} \mathcal{N}(\mu_0, \Sigma_0)$$

$$\forall C \in \Pi_N, \forall n \in C, x_n \overset{indep}{\sim} \mathcal{N}(\mu_C, \Sigma)$$



- Want: posterior $p(\Pi_N | x_{1:N})$

- Gibbs sampler:

$$p(\Pi_N | \Pi_{N,\text{-}n}, x) = \begin{cases} \frac{\#C}{\alpha + N - 1} p(x_{C \cup \{n\}} | x_C) & \text{if } n \text{ joins cluster } C \\ \frac{\alpha}{\alpha + N - 1} p(x_{\{n\}}) & \text{if } n \text{ starts a new cluster} \end{cases}$$

- For completeness: $p(x_{C \cup \{n\}} | x_C) = \mathcal{N}(\tilde{m}, \tilde{\Sigma} + \Sigma)$

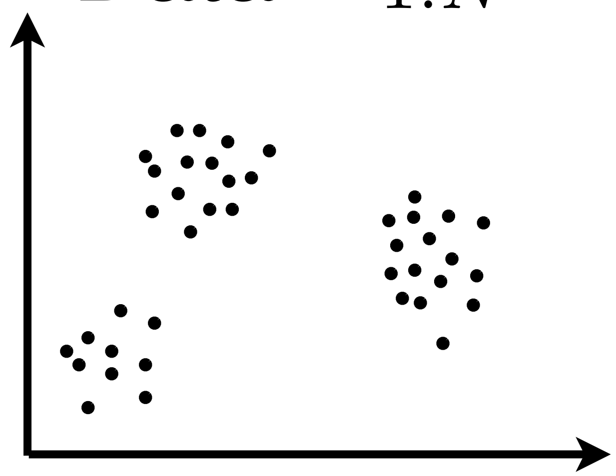
$$\tilde{\Sigma}^{-1} := \Sigma_0^{-1} + (\#C)\Sigma^{-1}$$

$$\tilde{m} := \tilde{\Sigma}\left(\Sigma^{-1} \sum_{m \in C} x_m + \Sigma_0^{-1}\mu_0\right)$$

[MacEachern 1994; Neal 1992; Neal 2000]

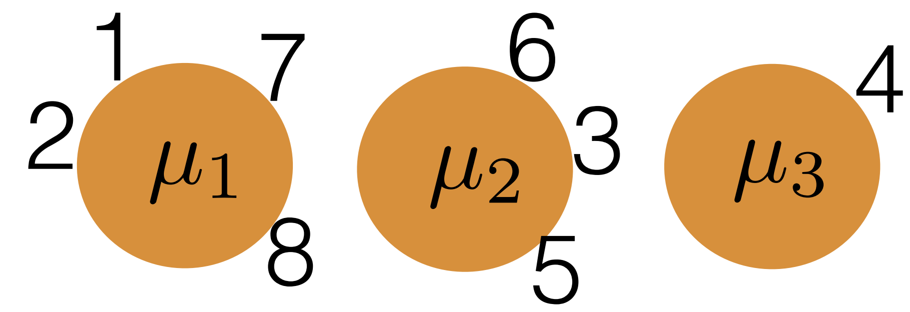
CRP mixture model: inference

- Data $x_{1:N}$



- Generative model

$$\Pi_N \sim \mathrm{CRP}(N, \alpha)$$
$$\forall C \in \Pi_N, \phi_C \overset{iid}{\sim} G_0$$
$$\forall C \in \Pi_N, \forall n \in C, x_n \overset{indep}{\sim} \mathcal{N}(\mu_C, \Sigma)$$



- Want: posterior $p(\Pi_N | x_{1:N})$

- Gibbs sampler:

$$p(\Pi_N | \Pi_{N,\text{-}n}, x) = \begin{cases} \frac{\#C}{\alpha+N-1} p(x_{C \cup \{n\}} | x_C) & \text{if } n \text{ joins cluster } C \\ \frac{\alpha}{\alpha+N-1} p(x_{\{n\}}) & \text{if } n \text{ starts a new cluster} \end{cases}$$

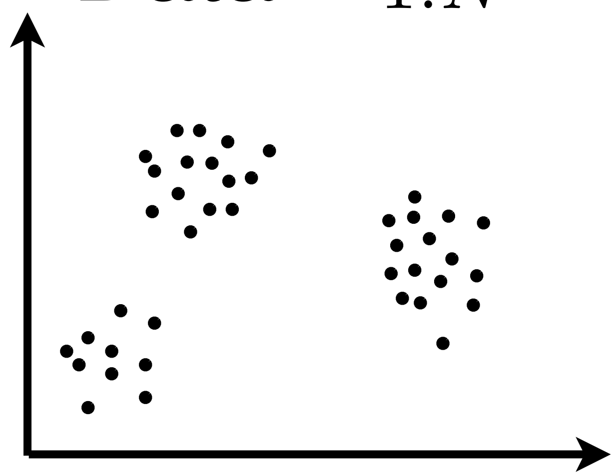
- For completeness: $p(x_{C \cup \{n\}} | x_C) = \mathcal{N}(\tilde{m}, \tilde{\Sigma} + \Sigma)$

$$\tilde{\Sigma}^{-1} := \Sigma_0^{-1} + (\#C)\Sigma^{-1}$$
$$\tilde{m} := \tilde{\Sigma}\left(\Sigma^{-1} \sum_{m \in C} x_m + \Sigma_0^{-1} \mu_0\right)$$

[MacEachern 1994; Neal 1992; Neal 2000]

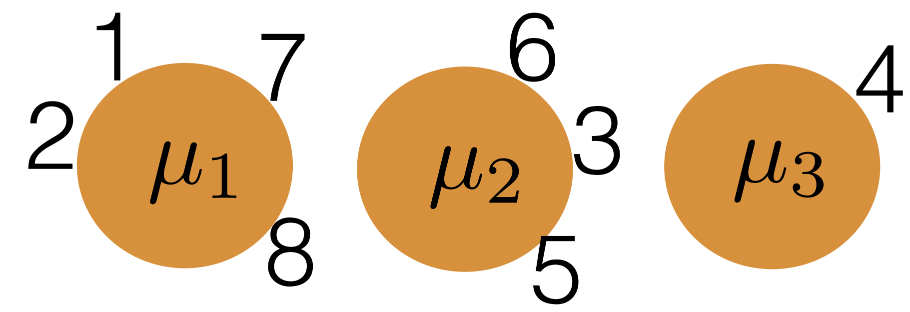
CRP mixture model: inference

- Data $x_{1:N}$



- Generative model

$$\Pi_N \sim \mathrm{CRP}(N, \alpha)$$
$$\forall C \in \Pi_N, \boxed{\phi_C \overset{iid}{\sim} G_0}$$
$$\forall C \in \Pi_N, \forall n \in C, x_n \overset{indep}{\sim} \boxed{F(\phi_C)}$$



- Want: posterior $p(\Pi_N | x_{1:N})$

- Gibbs sampler:

$$p(\Pi_N | \Pi_{N,\text{-}n}, x) = \begin{cases} \frac{\#C}{\alpha+N-1} p(x_{C \cup \{n\}} | x_C) & \text{if } n \text{ joins cluster } C \\ \frac{\alpha}{\alpha+N-1} p(x_{\{n\}}) & \text{if } n \text{ starts a new cluster} \end{cases}$$

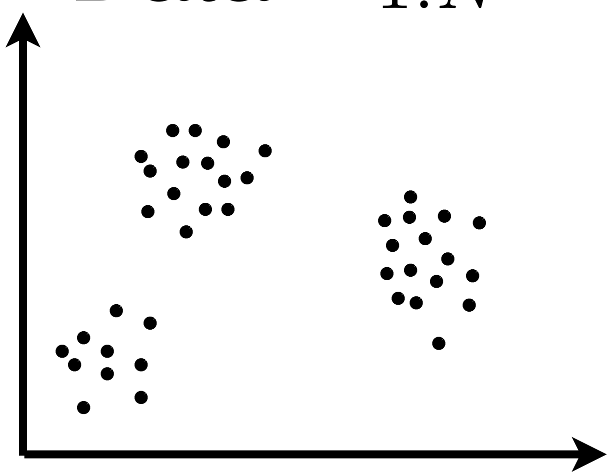
- For completeness: $p(x_{C \cup \{n\}} | x_C) = \mathcal{N}(\tilde{m}, \tilde{\Sigma} + \Sigma)$

$$\tilde{\Sigma}^{-1} := \Sigma_0^{-1} + (\#C)\Sigma^{-1}$$
$$\tilde{m} := \tilde{\Sigma}\left(\Sigma^{-1} \sum_{m \in C} x_m + \Sigma_0^{-1}\mu_0\right)$$

[MacEachern 1994; Neal 1992; Neal 2000]

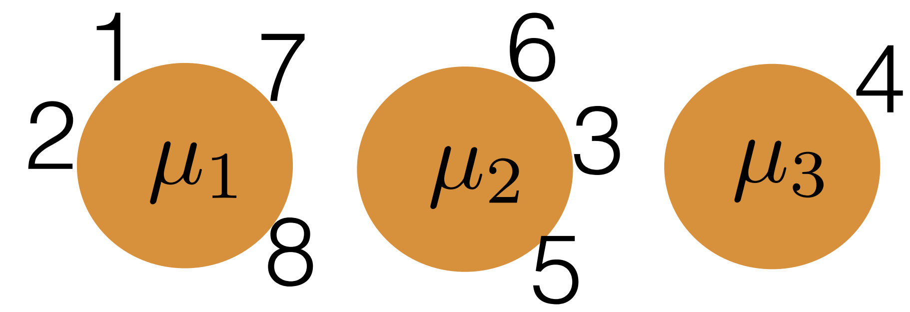
CRP mixture model: inference

- Data $x_{1:N}$



- Generative model

$$\Pi_N \sim \mathrm{CRP}(N, \alpha)$$
$$\forall C \in \Pi_N, \boxed{\phi_C \overset{iid}{\sim} G_0}$$
$$\forall C \in \Pi_N, \forall n \in C, x_n \overset{indep}{\sim} \boxed{F(\phi_C)}$$



- Want: posterior $p(\Pi_N | x_{1:N})$

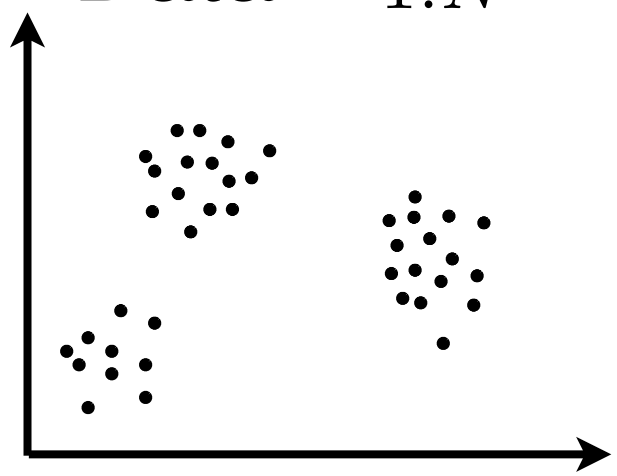
- Gibbs sampler:

$$p(\Pi_N | \Pi_{N,\text{-}n}, x) = \begin{cases} \frac{\#C}{\alpha+N-1} p(x_{C\cup\{n\}} | x_C) & \text{if } n \text{ joins cluster } C \\ \frac{\alpha}{\alpha+N-1} p(x_{\{n\}}) & \text{if } n \text{ starts a new cluster} \end{cases}$$

[MacEachern 1994; Neal 1992; Neal 2000]

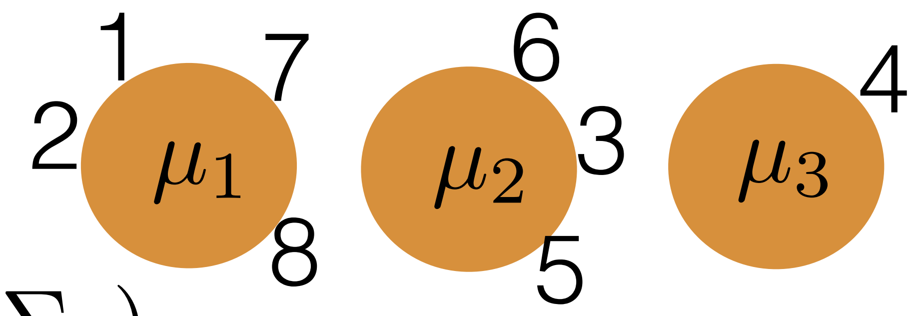
CRP mixture model: inference

- Data $x_{1:N}$



- Generative model

$$\Pi_N \sim \mathrm{CRP}(N, \alpha)$$
$$\forall C \in \Pi_N, \mu_C \stackrel{iid}{\sim} \mathcal{N}(\mu_0, \Sigma_0)$$
$$\forall C \in \Pi_N, \forall n \in C, x_n \stackrel{indep}{\sim} \mathcal{N}(\mu_C, \Sigma)$$



- Want: posterior $p(\Pi_N | x_{1:N})$
- Gibbs sampler:

$$p(\Pi_N | \Pi_{N,\text{-}n}, x) = \begin{cases} \frac{\#C}{\alpha + N - 1} p(x_{C \cup \{n\}} | x_C) & \text{if } n \text{ joins cluster } C \\ \frac{\alpha}{\alpha + N - 1} p(x_{\{n\}}) & \text{if } n \text{ starts a new cluster} \end{cases}$$

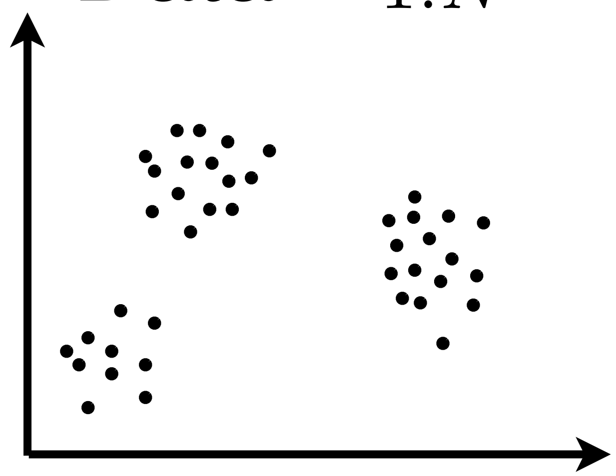
- For completeness: $p(x_{C \cup \{n\}} | x_C) = \mathcal{N}(\tilde{m}, \tilde{\Sigma} + \Sigma)$

$$\tilde{\Sigma}^{-1} := \Sigma_0^{-1} + (\#C)\Sigma^{-1}$$
$$\tilde{m} := \tilde{\Sigma} \left(\Sigma^{-1} \sum_{m \in C} x_m + \Sigma_0^{-1} \mu_0 \right)$$

[MacEachern 1994; Neal 1992; Neal 2000]

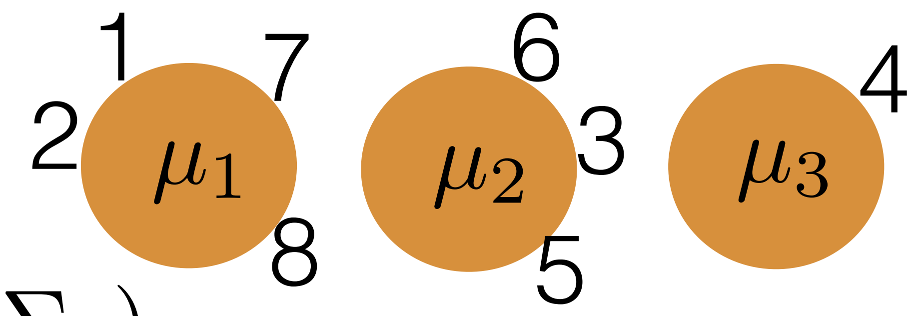
CRP mixture model: inference

- Data $x_{1:N}$



- Generative model

$$\Pi_N \sim \mathrm{CRP}(N, \alpha)$$
$$\forall C \in \Pi_N, \mu_C \overset{iid}{\sim} \mathcal{N}(\mu_0, \Sigma_0)$$
$$\forall C \in \Pi_N, \forall n \in C, x_n \overset{indep}{\sim} \mathcal{N}(\mu_C, \Sigma)$$



- Want: posterior $p(\Pi_N | x_{1:N})$

- Gibbs sampler:

$$p(\Pi_N | \Pi_{N,\text{-}n}, x) = \begin{cases} \frac{\#C}{\alpha+N-1} p(x_{C\cup\{n\}} | x_C) & \text{if } n \text{ joins cluster } C \\ \frac{\alpha}{\alpha+N-1} p(x_{\{n\}}) & \text{if } n \text{ starts a new cluster} \end{cases}$$

- For completeness: $p(x_{C\cup\{n\}} | x_C) = \mathcal{N}(\tilde{m}, \tilde{\Sigma} + \Sigma)$

$$\tilde{\Sigma}^{-1} := \Sigma_0^{-1} + (\#C)\Sigma^{-1}$$
$$\tilde{m} := \tilde{\Sigma}\left(\Sigma^{-1} \sum_{m\in C} x_m + \Sigma_0^{-1}\mu_0\right)$$

[demo]

[MacEachern 1994; Neal 1992; Neal 2000]