Nonparametric Bayesian Statistics: Part II

Tamara Broderick
ITT Career Development Assistant Professor
Electrical Engineering & Computer Science
MIT
Recall: Part I
Recall: Part I

• Dirichlet process (DP) stick-breaking

\[ \cdots \]

\[ \cdots \]

\[ \cdots \]

\[ \cdots \]

\[ \cdots \]
Recall: Part I

- Dirichlet process (DP) stick-breaking
- Griffiths-Engen-McCloskey (GEM) distribution:

\[ \theta \sim GEM(\alpha) \]

\[ \pi_k \sim \text{Beta}(1, \alpha) \]

\[ \pi_k = \frac{2}{4k+1} - \frac{2}{4k+3} \]

\[ \sum_{j=1}^{k} \pi_j = 1 \]
Recall: Part I

- **Dirichlet process (DP) stick-breaking**
- Griffiths-Engen-McCloskey (**GEM**) distribution:
  \[ \rho = (\rho_1, \rho_2, \ldots) \sim \text{GEM}(\alpha) \]
Recall: Part I

• Dirichlet process (DP) stick-breaking

• Griffiths-Engen-McCloskey (GEM) distribution:
  \[ \rho = (\rho_1, \rho_2, \ldots) \sim \text{GEM}(\alpha) \]
Recall: Part I

- Dirichlet process (DP) stick-breaking
- Griffiths-Engen-McCloskey (GEM) distribution:
  \[ \rho = (\rho_1, \rho_2, \ldots) \sim \text{GEM}(\alpha) \]
Recall: Part I

• Dirichlet process (DP) stick-breaking

• Griffiths-Engen-McCloskey (GEM) distribution:

\[ \rho = (\rho_1, \rho_2, \ldots) \sim \text{GEM}(\alpha) \]

[McCloskey 1965; Engen 1975; Patil and Taillie 1977; Ewens 1987; Sethuraman 1994; Ishwaran, James 2001]
Recall: Part I

• Dirichlet process (DP) stick-breaking

• Griffiths-Engen-McCloskey (GEM) distribution:

\[ \rho = (\rho_1, \rho_2, \ldots) \sim \text{GEM}(\alpha) \]

1 [McCloskey 1965; Engen 1975; Patil and Taillie 1977; Ewens 1987; Sethuraman 1994; Ishwaran, James 2001]
Recall: Part I

• Dirichlet process (DP) stick-breaking

• Griffiths-Engen-McCloskey (GEM) distribution:
  \[ \rho = (\rho_1, \rho_2, \ldots) \sim \text{GEM}(\alpha) \]

[McCloskey 1965; Engen 1975; Patil and Taillie 1977; Ewens 1987; Sethuraman 1994; Ishwaran, James 2001]
Recall: Part I

- **Dirichlet process (DP) stick-breaking**
- **Griffiths-Engen-McCloskey (GEM) distribution:**

\[ \rho = (\rho_1, \rho_2, \ldots) \sim \text{GEM}(\alpha) \]

[McCloskey 1965; Engen 1975; Patil and Taillie 1977; Ewens 1987; Sethuraman 1994; Ishwaran, James 2001]
Recall: Part I

- Dirichlet process (DP) stick-breaking
- Griffiths-Engen-McCloskey (GEM) distribution:
  \[ \rho = (\rho_1, \rho_2, \ldots) \sim \text{GEM}(\alpha) \]

\[ V_k \overset{iid}{\sim} \text{Beta}(1, \alpha) \]

[McCloskey 1965; Engen 1975; Patil and Taillie 1977; Ewens 1987; Sethuraman 1994; Ishwaran, James 2001]
Recall: Part I

- **Dirichlet process (DP) stick-breaking**
- **Griffiths-Engen-McCloskey (GEM) distribution:**
  \[
  \rho = (\rho_1, \rho_2, \ldots) \sim \text{GEM}(\alpha)
  \]

\[V_k \overset{iid}{\sim} \text{Beta}(1, \alpha)\]

\[\rho_k = \left[ \prod_{j=1}^{k-1} (1 - V_j) \right] V_k\]

\[\ldots\]
Recall: Part I

• Dirichlet process (DP) stick-breaking

• Griffiths-Engen-McCloskey (GEM) distribution:
  \[ \rho = (\rho_1, \rho_2, \ldots) \sim \text{GEM}(\alpha) \]

\[ V_k \overset{iid}{\sim} \text{Beta}(1, \alpha) \quad \rho_k = \left[ \prod_{j=1}^{k-1} (1 - V_j) \right] V_k \]

1 [McCloskey 1965; Engen 1975; Patil and Taillie 1977; Ewens 1987; Sethuraman 1994; Ishwaran, James 2001]
Recall: Part I

- Dirichlet process (DP) stick-breaking
- Griffiths-Engen-McCloskey (GEM) distribution:

$$\rho = (\rho_1, \rho_2, \ldots) \sim \text{GEM}(\alpha)$$

$$V_k \overset{iid}{\sim} \text{Beta}(1, \alpha) \quad \rho_k = \left[ \prod_{j=1}^{k-1} (1 - V_j) \right] V_k$$

- Part of: DP mixture model

---

1 [McCloskey 1965; Engen 1975; Patil and Taillie 1977; Ewens 1987; Sethuraman 1994; Ishwaran, James 2001]
DP or not DP, that is the question
DP or not DP, that is the question

- GEM: 

- Compare to:
  - Finite (small $K$) mixture model
  - Finite (large $K$) mixture model

Time series …
DP or not DP, that is the question

- GEM: 
  ![Bar chart with multiple colors]
- Compare to:

![Scatter plot with points]
DP or not DP, that is the question

• GEM:  

• Compare to:
  • Finite (small $K$) mixture model
DP or not DP, that is the question

- GEM: 

- Compare to:
  - Finite (small $K$) mixture model
  - Finite (large $K$) mixture model
DP or not DP, that is the question

• GEM:

• Compare to:
  • Finite (small $K$) mixture model
  • Finite (large $K$) mixture model
  • Time series
Nonparametric Bayes: Part II

• Last time:

  • Understand what it means to have an infinite/growing number of parameters
  • Finite representation allows use of infinite model
  • www.tamarabroderick.com/tutorials.html

• This time:

  • Avoid the infinity of parameters for inference
  • e.g. Chinese restaurant process
Nonparametric Bayes: Part II

• Last time:
  • Understand what it means to have an infinite/growing number of parameters

• Avoid the infinity of parameters for inference
  • e.g. Chinese restaurant process
Nonparametric Bayes: Part II

- Last time:
  - Understand what it means to have an infinite/growing number of parameters
  - Finite representation allows use of infinite model
Nonparametric Bayes: Part II

• Last time:
  • Understand what it means to have an infinite/growing number of parameters
  • Finite representation allows use of infinite model
  • www.tamarabroderick.com/tutorials.html
Nonparametric Bayes: Part II

• Last time:
  • Understand what it means to have an infinite/growing number of parameters
  • Finite representation allows use of infinite model
  • www.tamarabroderick.com/tutorials.html

• This time:
Nonparametric Bayes: Part II

• Last time:
  • Understand what it means to have an infinite/growing number of parameters
  • Finite representation allows use of infinite model
  • www.tamarabroderick.com/tutorials.html

• This time:
  • Avoid the infinity of parameters for inference
Nonparametric Bayes: Part II

• Last time:
  • Understand what it means to have an infinite/growing number of parameters
  • Finite representation allows use of infinite model
  • [Website](www.tamarabroderick.com/tutorials.html)

• This time:
  • Avoid the infinity of parameters for inference
  • e.g. Chinese restaurant process
Marginal cluster assignments

\[ \rho_1 \sim \text{Beta}(a_1, a_2), \quad z_n \overset{iid}{\sim} \text{Cat}(\rho_1, \rho_2) \]
Marginal cluster assignments

- Integrate out the frequencies

\[ \rho_1 \sim \text{Beta}(a_1, a_2), \quad z_n \overset{iid}{\sim} \text{Cat}(\rho_1, \rho_2) \]
Marginal cluster assignments

- Integrate out the frequencies

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

\[ p(z_n = 1|z_1, \ldots, z_{n-1}) \]
Marginal cluster assignments

- Integrate out the frequencies

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

\[ p(z_n = 1|z_1, \ldots, z_{n-1}) = \int p(z_n = 1|\rho_1)p(\rho_1|z_1, \ldots, z_{n-1})d\rho_1 \]
Integrate out the frequencies

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

\[ p(z_n = 1|z_1, \ldots, z_{n-1}) = \int p(z_n = 1|\rho_1)p(\rho_1|z_1, \ldots, z_{n-1})d\rho_1 \]

\[ = \int \rho_1 \text{Beta}(\rho_1|a_{1,n}, a_{2,n})d\rho_1 \]
Marginal cluster assignments

• Integrate out the frequencies

\[ \rho_1 \sim \text{Beta}(a_1, a_2), \ z_n \overset{iid}{\sim} \text{Cat}(\rho_1, \rho_2) \]
\[ p(z_n = 1|z_1, \ldots, z_{n-1}) \]
\[ = \int p(z_n = 1|\rho_1)p(\rho_1|z_1, \ldots, z_{n-1})d\rho_1 \]
\[ = \int \rho_1 \text{Beta}(\rho_1|a_{1,n}, a_{2,n})d\rho_1 \]
\[ a_{1,n} := a_1 + \sum_{m=1}^{n-1} 1\{z_m = 1\}, \ a_{2,n} = a_2 + \sum_{m=1}^{n-1} 1\{z_m = 2\} \]
Marginal cluster assignments

- Integrate out the frequencies

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

\[ p(z_n = 1|z_1, \ldots, z_{n-1}) = \int p(z_n = 1|\rho_1)p(\rho_1|z_1, \ldots, z_{n-1})d\rho_1 \]

\[ = \int \rho_1 \text{Beta}(\rho_1|a_{1,n}, a_{2,n})d\rho_1 \]

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

\[ = \int \rho_1 \frac{\Gamma(a_{1,n} + a_{2,n})}{\Gamma(a_{1,n})\Gamma(a_{2,n})} \rho_1^{a_{1,n}-1}(1 - \rho_1)^{a_{2,n}-1}d\rho_1 \]
Marginal cluster assignments

- Integrate out the frequencies

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

\[
p(z_n = 1|z_1, \ldots, z_{n-1})
= \int p(z_n = 1|\rho_1)p(\rho_1|z_1, \ldots, z_{n-1}) d\rho_1
= \int \rho_1 \text{Beta}(\rho_1|a_{1,n}, a_{2,n}) d\rho_1
\]

\[
a_{1,n} := a_1 + \sum_{m=1}^{n-1} 1\{z_m = 1\}, \quad a_{2,n} = a_2 + \sum_{m=1}^{n-1} 1\{z_m = 2\}
\]

\[
= \int \rho_1 \frac{\Gamma(a_{1,n} + a_{2,n})}{\Gamma(a_{1,n})\Gamma(a_{2,n})} \rho_1^{a_{1,n}-1}(1 - \rho_1)^{a_{2,n}-1} d\rho_1
\]

\[
= \frac{\Gamma(a_{1,n} + a_{2,n})}{\Gamma(a_{1,n})\Gamma(a_{2,n})} \frac{\Gamma(a_{1,n} + 1)\Gamma(a_{2,n})}{\Gamma(a_{1,n} + a_{2,n} + 1)}
\]
Marginal cluster assignments

- Integrate out the frequencies
  \[ \rho_1 \sim \text{Beta}(a_1, a_2), \quad z_n \overset{iid}{\sim} \text{Cat}(\rho_1, \rho_2) \]
  \[ p(z_n = 1|z_1, \ldots, z_{n-1}) = \int p(z_n = 1|\rho_1)p(\rho_1|z_1, \ldots, z_{n-1})d\rho_1 \]
  \[ = \int \rho_1 \text{Beta}(\rho_1|a_{1,n}, a_{2,n})d\rho_1 \]
  \[ a_{1,n} := a_1 + \sum_{m=1}^{n-1} 1\{z_m = 1\}, \quad a_{2,n} = a_2 + \sum_{m=1}^{n-1} 1\{z_m = 2\} \]
  \[ = \frac{\Gamma(a_{1,n} + a_{2,n})}{\Gamma(a_{1,n})\Gamma(a_{2,n})} \frac{\Gamma(a_{1,n} + 1)\Gamma(a_{2,n})}{\Gamma(a_{1,n} + a_{2,n} + 1)} \frac{a_{1,n}}{a_{1,n} + a_{2,n}} \]
Marginal cluster assignments

- Integrate out the frequencies

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

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

\[
a_{1,n} := a_1 + \sum_{m=1}^{n-1} 1\{z_m = 1\}, \quad a_{2,n} = a_2 + \sum_{m=1}^{n-1} 1\{z_m = 2\}
\]
Marginal cluster assignments

- Integrate out the frequencies
  \[ \rho_1 \sim \text{Beta}(a_1, a_2), \ z_n \overset{iid}{\sim} \text{Cat}(\rho_1, \rho_2) \]
  \[ p(z_n = 1|z_1, \ldots, z_{n-1}) = \frac{a_{1,n}}{a_{1,n} + a_{2,n}} \]
  \[ a_{1,n} := a_1 + \sum_{m=1}^{n-1} 1\{z_m = 1\}, \ a_{2,n} = a_2 + \sum_{m=1}^{n-1} 1\{z_m = 2\} \]

- Pólya urn
Marginal cluster assignments

- Integrate out the frequencies
  \[ \rho_1 \sim \text{Beta}(a_1, a_2), \quad z_n \overset{iid}{\sim} \text{Cat}(\rho_1, \rho_2) \]
  \[ p(z_n = 1|z_1, \ldots, z_{n-1}) = \frac{a_{1,n}}{a_{1,n} + a_{2,n}} \]
  \[ a_{1,n} := a_1 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 1\}, \quad a_{2,n} = a_2 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 2\} \]

- Pólya urn
Marginal cluster assignments

- Integrate out the frequencies

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

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

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

- Pólya urn
Marginal cluster assignments

- Integrate out the frequencies
  \[ \rho_1 \sim \text{Beta}(a_1, a_2), \quad z_n \overset{iid}{\sim} \text{Cat}(\rho_1, \rho_2) \]
  \[ p(z_n = 1|z_1, \ldots, z_{n-1}) = \frac{a_1,n}{a_1,n + a_2,n} \]
  \[ a_{1,n} := a_1 + \sum_{m=1}^{n-1} 1\{z_m = 1\}, \quad a_{2,n} = a_2 + \sum_{m=1}^{n-1} 1\{z_m = 2\} \]

- Pólya urn
  - Choose any ball with equal probability
Marginal cluster assignments

- Integrate out the frequencies
  \[ \rho_1 \sim \text{Beta}(a_1, a_2), \quad z_n \overset{iid}{\sim} \text{Cat}(\rho_1, \rho_2) \]
  \[ p(z_n = 1|z_1, \ldots, z_{n-1}) = \frac{a_{1,n}}{a_{1,n} + a_{2,n}} \]
  \[ a_{1,n} := a_1 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 1\}, \quad a_{2,n} = a_2 + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = 2\} \]

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

• Integrate out the frequencies

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

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

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

• Pólya urn

  • Choose any ball with equal probability
  • Replace and add ball of same color

\[ \lim_{n \to \infty} \frac{\# \text{ orange}}{\# \text{ total}} = \text{Beta}(a_{\text{orange}}, a_{\text{green}}) \]
Marginal cluster assignments

- Integrate out the frequencies

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

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

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

- 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
  \[ \rho_1 \sim \text{Beta}(a_1, a_2), \ z_n \overset{iid}{\sim} \text{Cat}(\rho_1, \rho_2) \]
  \[ p(z_n = 1|z_1, \ldots, z_{n-1}) = \frac{\alpha_1}{\alpha_1 + \alpha_2} \]
  \[ a_{1,n} := a_1 + \sum_{m=1}^{n-1} 1\{z_m = 1\}, \ a_{2,n} = a_2 + \sum_{m=1}^{n-1} 1\{z_m = 2\} \]

- 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
  \[ \rho_1 \sim \text{Beta}(a_1, a_2), \quad z_n \overset{iid}{\sim} \text{Cat}(\rho_1, \rho_2) \]
  \[ p(z_n = 1|z_1, \ldots, z_{n-1}) = \frac{a_{1,n}}{a_{1,n} + a_{2,n}} \]

  \[ a_{1,n} := a_1 + \sum_{m=1}^{n-1} 1\{z_m = 1\}, \quad a_{2,n} = a_2 + \sum_{m=1}^{n-1} 1\{z_m = 2\} \]

• 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

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

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

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

• 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
Marginal cluster assignments

- Integrate out the frequencies

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

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

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

\[ a_{k,n} := a_k + \sum_{m=1}^{n-1} 1\{z_m = k\} \]
Marginal cluster assignments

- Integrate out the frequencies
  \[
  \rho_{1:K} \sim \text{Dirichlet}(a_{1:K}), \ z_n \sim \text{Cat}(\rho_{1:K})
  \]
  \[
  p(z_n = k | z_1, \ldots, z_{n-1}) = \frac{a_{k,n}}{\sum_{j=1}^{K} a_{j,n}}
  \]
  \[
  a_{k,n} := a_k + \sum_{m=1}^{n-1} \mathbf{1}\{z_m = k\}
  \]

- Multivariate Pólya urn
Marginal cluster assignments

• Integrate out the frequencies
  \( \rho_{1:K} \sim \text{Dirichlet}(a_{1:K}) \), \( z_n \overset{iid}{\sim} \text{Cat}(\rho_{1:K}) \)
  \[
p(z_n = k | z_1, \ldots, z_{n-1}) = \frac{a_{k,n}}{\sum_{j=1}^{K} a_{j,n}}
  \]
  \[
a_{k,n} := a_k + \sum_{m=1}^{n-1} 1\{z_m = k\}
  \]

• multivariate Pólya urn
Marginal cluster assignments

- Integrate out the frequencies
  \( \rho_{1:K} \sim \text{Dirichlet}(a_{1:K}) \), \( z_n \overset{iid}{\sim} \text{Cat}(\rho_{1:K}) \)

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

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

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

- Integrate out the frequencies
  \[ \rho_{1:K} \sim \text{Dirichlet}(a_{1:K}), \quad z_n \overset{iid}{\sim} \text{Cat}(\rho_{1:K}) \]
  
  \[ p(z_n = k | z_1, \ldots, z_{n-1}) = \frac{a_{k,n}}{\sum_{j=1}^{K} a_{j,n}} \]

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

- 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

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

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

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

• 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

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

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

\[ a_{k,n} := \alpha_k + \sum_{m=1}^{n-1} 1\{z_m = k\} \]

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

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

\[ \to (\rho_{\text{orange}}, \rho_{\text{green}}, \rho_{\text{red}}, \rho_{\text{yellow}}) \]
Marginal cluster assignments

- Integrate out the frequencies

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

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

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

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

\[ \lim_{n \to \infty} \left( \frac{\# \text{ orange}, \# \text{ green}, \# \text{ red}, \# \text{ yellow}}{\# \text{ total}} \right) \]

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

\[ d = \text{Dirichlet}(a_{\text{orange}}, a_{\text{green}}, a_{\text{red}}, a_{\text{yellow}}) \]
Marginal cluster assignments

• Hoppe urn / Blackwell-MacQueen urn
Marginal cluster assignments

- Hoppe urn / Blackwell-MacQueen urn

[Blackwell, MacQueen 1973; Hoppe 1984]
Marginal cluster assignments

- Hoppe urn / Blackwell-MacQueen urn
Marginal cluster assignments

• Hoppe urn / Blackwell-MacQueen urn

  • Choose ball with prob proportional to its mass
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

Step 0 | Step 1
---|---
• | •
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 | Step 1 | Step 2
--- | --- | ---
\[ \text{\begin{center}
\includegraphics[width=0.2\textwidth]{urn}
\end{center}} \] | \[ \text{\begin{center}
\includegraphics[width=0.2\textwidth]{black}
\end{center}} \] | \[ \text{\begin{center}
\includegraphics[width=0.2\textwidth]{black}
\end{center}} \]

[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 | Step 1 | Step 2 | Step 3
[black] | [black] | [black] | [black]
[orange] | [orange] | [orange] | [green]
```

[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  Step 1  Step 2  Step 3  Step 4

[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

\[
\begin{align*}
\text{Step 0} & \quad \text{Step 1} & \quad \text{Step 2} & \quad \text{Step 3} & \quad \text{Step 4} \\
\begin{array}{c}
\black \quad & \black \quad & \black \quad & \black \quad & \black \\
\end{array} & \begin{array}{c}
\black \quad \black \quad & \black \quad \black \quad & \black \quad \black \quad & \black \quad \black \quad & \black \quad \black \quad \\
\end{array} & \begin{array}{c}
\black \quad \black \quad \black \quad & \black \quad \black \quad \black \quad & \black \quad \black \quad \black \quad & \black \quad \black \quad \black \quad & \black \quad \black \quad \black \quad \\
\end{array} & \begin{array}{c}
\black \quad \black \quad \black \quad \black \quad & \black \quad \black \quad \black \quad \black \quad & \black \quad \black \quad \black \quad \black \quad & \black \quad \black \quad \black \quad \black \quad & \black \quad \black \quad \black \quad \black \quad \\
\end{array} & \begin{array}{c}
\black \quad \black \quad \black \quad \black \quad \black \quad & \black \quad \black \quad \black \quad \black \quad \black \quad & \black \quad \black \quad \black \quad \black \quad \black \quad & \black \quad \black \quad \black \quad \black \quad \black \quad & \black \quad \black \quad \black \quad \black \quad \black \quad \\
\end{array} \\
(\#\text{orange}, \#\text{other}) = \text{PolyaUrn}(1, \alpha)
\end{align*}
\]
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

\[
\begin{array}{c|c|c|c|c}
\text{Step 0} & \text{Step 1} & \text{Step 2} & \text{Step 3} & \text{Step 4} \\
\black & \black & \black & \black & \black \\
\end{array}
\]

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

- not orange
  - not orange, green:

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

\[
\begin{align*}
\text{Step 0} & : \black\black \\
\text{Step 1} & : \black\black\orange \\
\text{Step 2} & : \black\black\black\orange\orange \\
\text{Step 3} & : \black\black\black\orange\orange\green\green \\
\text{Step 4} & : \black\black\black\orange\orange\green\green\red\red \\
\end{align*}
\]

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

- 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

\[
V_k \overset{iid}{\sim} \text{Beta}(1, \alpha)
\]

<table>
<thead>
<tr>
<th>Step 0</th>
<th>Step 1</th>
<th>Step 2</th>
<th>Step 3</th>
<th>Step 4</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="#" alt="Ball" /></td>
<td><img src="#" alt="Ball" /></td>
<td><img src="#" alt="Ball" /></td>
<td><img src="#" alt="Ball" /></td>
<td><img src="#" alt="Ball" /></td>
</tr>
<tr>
<td><img src="#" alt="Ball" /></td>
<td><img src="#" alt="Ball" /></td>
<td><img src="#" alt="Ball" /></td>
<td><img src="#" alt="Ball" /></td>
<td><img src="#" alt="Ball" /></td>
</tr>
<tr>
<td><img src="#" alt="Ball" /></td>
<td><img src="#" alt="Ball" /></td>
<td><img src="#" alt="Ball" /></td>
<td><img src="#" alt="Ball" /></td>
<td><img src="#" alt="Ball" /></td>
</tr>
<tr>
<td><img src="#" alt="Ball" /></td>
<td><img src="#" alt="Ball" /></td>
<td><img src="#" alt="Ball" /></td>
<td><img src="#" alt="Ball" /></td>
<td><img src="#" alt="Ball" /></td>
</tr>
</tbody>
</table>

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

\[ V_k \overset{iid}{\sim} \text{Beta}(1, \alpha) \]
\[ \rho_1 = V_1 \]

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

Step 0

Step 1

Step 2

Step 3

Step 4

\[ V_k \overset{iid}{\sim} \text{Beta}(1, \alpha) \]
\[ \rho_1 = V_1 \]
\[ \rho_2 = (1 - V_1)V_2 \]

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

\[
\begin{align*}
V_k & \overset{iid}{\sim} \text{Beta}(1, \alpha) \\
\rho_1 &= V_1 \\
\rho_2 &= (1 - V_1)V_2 \\
\rho_3 &= \left[\prod_{k=1}^{2}(1 - V_k)\right]V_3
\end{align*}
\]

\begin{align*}
(#\text{orange}, #\text{other}) &= \text{PolyaUrn}(1, \alpha) \\
\text{not orange: } (#\text{green}, #\text{other}) &= \text{PolyaUrn}(1, \alpha) \\
\text{not orange, green: } (#\text{red}, #\text{other}) &= \text{PolyaUrn}(1, \alpha)
\end{align*}
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 \( \alpha \)

Marginal for the Categorical likelihood with GEM prior

\( z_1 = z_2 = z_7 = z_8 = 1, z_3 = z_5 = z_6 = 2, z_4 = 3 \)

\( \mathcal{S}_8 = \{\{1, 2, 7, 8\}, \{3, 5, 6\}, \{4\}\} \)

\( \mathcal{S}_{[8]} = \{1, \ldots, 8\} \)
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 $\alpha$
Chinese restaurant process

1. Same thing we just did
2. Each customer walks into the restaurant
   - Sits at existing table with prob proportional to # people there
   - Forms new table with prob proportional to $\alpha$
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 $\alpha$
Chinese restaurant process

1. Same thing we just did
2. Each customer walks into the restaurant
   1. Sits at existing table with prob proportional to # people there
   2. Forms new table with prob proportional to $\alpha$
Chinese restaurant process

1. \( \phi_1 \)

- 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 \( \alpha \)
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 $\alpha$
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 $\alpha$

\[ \phi_1 \]

\[ 1, 2, 3 \]
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 $\alpha$

$\phi_1$
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 $\alpha$
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 $\alpha$

Diagram:

1. $\phi_1$
2. $\phi_2$
3. (blank circle)

Partition of $[8]$: set of mutually exclusive & exhaustive sets

$[8] = \{1, \ldots, 8\}$

$
\phi_1 = \{1, 2, 7, 8\},
\phi_2 = \{3, 5, 6\},
\phi_3 = \{4\}.
$

$marginal$ for the Categorical likelihood with GEM prior

$\phi_1$
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 $\alpha$
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 $\alpha$
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 $\alpha$
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 $\alpha$
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 $\alpha$
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 $\alpha$
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 $\alpha$
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 $\alpha$

\[ \{1, 2, 7, 8 \}, \{3, 5, 6 \}, \{4 \} \]
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 $\alpha$

\[\phi_1, \phi_2, \phi_3\]

\[\{1, 2, 7, 8\}, \{3, 5, 6\}, \{4\}\]
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 $\alpha$
• 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 $\alpha$
- Marginal for the Categorical likelihood with GEM prior

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

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 $\alpha$
- 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:
  \[
  \frac{1}{\alpha^2}
  \]
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:
  \[
  \alpha \cdot \frac{1}{\alpha + 1} \cdot \alpha \cdot \frac{\alpha}{\alpha + 2} \cdot \alpha \cdot \frac{\alpha}{\alpha + 3} \cdot \frac{1}{\alpha + 4} \cdot \frac{2}{\alpha + 5} \cdot \frac{2}{\alpha + 6} \cdot \frac{3}{\alpha + 7}
  \]
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 \( k \)):

\[
\frac{\alpha \cdots (\alpha + N - 1)}{\alpha \cdots (\alpha + 1)}
\]
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 \( k \)):
\[
\frac{\alpha^{K_N}}{\alpha \cdots (\alpha + N - 1)}
\]
• 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 \( k \)):
\[ \frac{\alpha^{K_N} \prod_{k=1}^{K_N} (n_k - 1)!}{\alpha \cdots (\alpha + N - 1)} \]
References (Part II)


