

本の P.173 にある、IBP について考える。

論文の (15)(16) を見ると、CRP に関してはわかりやすい。

(16) を考慮して、

$$P(c) = \prod_n P(c_n | c_1, \dots, c_{n-1}) = \frac{1}{(N-1-\alpha)!} \alpha^{K_+} \prod_k^{K_+} (m_k - 1)! \quad (1)$$

となる。ここで、otherwise の方は 0 でないクラスの数なので、 K_+ 回選択される。それ以外のときは、(16) で m_k がそれぞれカウントアップされるので、階乗の項が出てくる。(しかし、追加される前の個数をかけるので、-1 がつく。)

逆に (15) を考慮すると、

$$P(c) = \alpha^{K_+} \left(\prod_k^{K_+} (m_k - 1)! \right) \frac{\Gamma(\alpha)}{\Gamma(N + \alpha)} = \alpha^{K_+} \left(\prod_k^{K_+} (m_k - 1)! \right) \frac{\Gamma(\alpha)}{(N - 1 + \alpha)! \Gamma(\alpha)} = \alpha^{K_+} \left(\prod_k^{K_+} (m_k - 1)! \right) \frac{1}{(N - 1 + \alpha)!} \quad (2)$$

よって、(16) の手続きで、(15) の確率分布が出てくる。

なお、 c_i はすでに選ばれている k 個の料理と新しい料理 $k+1$ のいずれかになる。確率として、 c_1, \dots, c_{i-1} に条件付けられると、 $\sum_k P(c_i | c_1, \dots, c_{i-1}) = 1$ となる。 i 番目のサンプリング前で、 $\sum m_k = i - 1$ になることを考慮すると、

$$P(c_i | c_1, \dots, c_{i-1}) = \sum_k \frac{m_k}{i - 1 + \alpha} + \frac{\alpha}{i - 1 + \alpha} = \frac{i - 1}{i - 1 + \alpha} + \frac{\alpha}{i - 1 + \alpha} = 1 \quad (3)$$

(6.51) に関して考える。

$$\begin{aligned} p([M]) &= \frac{(\alpha\beta)^{H_+}}{\prod_{i \geq 1} H_i!} \exp\left(-\left(\alpha \sum_{j=1}^N \frac{\beta}{j + \beta - 1}\right)\right) \prod_{h=1}^{H_+} \frac{\Gamma(N_h) \Gamma(N - N_h + \beta)}{\Gamma(N + \beta)} \\ &= \frac{(\alpha\beta)^{H_+}}{\prod_{i \geq 1} H_i!} \exp\left(-\left(\alpha \sum_{j=1}^N \frac{\beta}{j + \beta - 1}\right)\right) \prod_{h=1}^{H_+} \frac{(N_h - 1)! (N - N_h + \beta - 1)! \Gamma(\beta)}{(N + \beta - 1)! \Gamma(\beta)} \\ &= \frac{(\alpha\beta)^{H_+}}{\prod_{i \geq 1} H_i!} \exp\left(-\left(\alpha \sum_{j=1}^N \frac{\beta}{j + \beta - 1}\right)\right) \prod_{h=1}^{H_+} \frac{(N_h - 1)! (N - N_h + \beta - 1)!}{(N + \beta - 1)!} \end{aligned} \quad (4)$$

なお、 $(N + \beta)! = (N + \beta)(N - 1 + \beta) \cdots (1 + \beta)\beta$ であり、

$$H_+ = \sum_{i \geq 1} H_i \quad (5)$$

IBP の手続きは、P.173 の 1,2 のように行う。ただ、 $n=1$ のときに $\frac{\alpha\beta}{n+\beta-1} = \alpha$ なので、記載方法を変えると、

$$\begin{aligned} p(z_{nh} = 1) &= \frac{N_{nh}}{n + \beta - 1} \text{ (すでに選択されている料理 } h \text{ が選ばれる確率。)} \\ n \text{ はこのときまでに選んだ人の数, } N_{nh} \text{ はその時まで } h \text{ を選んでいる人の数。)} \\ p_n(x_n | \frac{\alpha\beta}{n + \beta - 1}) &= Poi(x_n | \frac{\alpha\beta}{n + \beta - 1}) = \frac{\frac{\alpha\beta}{n + \beta - 1}^{x_n}}{x_n!} e^{-\frac{\alpha\beta}{n + \beta - 1}} \\ &\text{(} n \text{ 番目の人がこれまで取られていない新しい料理の数。)} \end{aligned} \quad (6)$$

このとき、

$$H_+ = \sum_n^N x_n \quad (7)$$

$$p_n(z_{nh} = 0) = \frac{n + \beta - 1 - N_{nh}}{n + \beta - 1} \quad (8)$$

論文の Fig.5 に関して、(35) が成り立つことを確認する。なお、本とは異なり、 $\beta = 1$ となっている。

まず、すでに選ばれている料理と新しい料理に関する部分に分ける。これらは独立なので、

$$P(Z) = P_{z\text{chosen}}(Z)P_{z\text{new}}(Z) \quad (9)$$

fig.5 で新しい料理は各列で一番上に出てくる 1 のところに対応し、それより下の部分が、すでに選ばれている部分に相当する。逆に新しいところよりも上の 0 のところは必ず 0 になるので確率 1 になる。

$P_{z\text{new}}(Z)$ を考えると、ポアソン分布の積になる。

$$\begin{aligned} P_{z\text{new}}(Z) &= \prod_{i=1}^N \left(\frac{(\frac{\alpha}{i})^{K_1^{(i)}}}{K_1^{(i)}!} \exp(-\frac{\alpha}{i}) \right) = \exp(-\alpha \sum_{i=1}^N \frac{1}{i}) \prod_{i=1}^N \left((\frac{1}{i})^{K_1^{(i)}} \frac{\alpha^{K_1^{(i)}}}{K_1^{(i)}!} \right) \\ &= \exp(-\alpha H_N) \alpha^{K_+} \prod_{i=1}^N \left((\frac{1}{i})^{K_1^{(i)}} \frac{1}{K_1^{(i)}!} \right) = \frac{\exp(-\alpha H_N) \alpha^{K_+}}{\prod_{i=1}^N K_1^{(i)}!} \prod_{i=1}^N (\frac{1}{i})^{K_1^{(i)}} \end{aligned} \quad (10)$$

次に、 $P_{z\text{chosen}}(Z)$ を考える。

$$P_{z\text{chosen}}(Z) = \prod_k^{K_+} \left(\prod_{i=a(k)+1, z_{i,k}=1} \frac{N_{nk}}{i} \prod_{i=a(k)+1, z_{i,k}=0} \frac{i - N_{nk}}{i} \right) \quad (11)$$

$a(k)$ は k 列で初めて、1 になる列の番号。

これに関して、論文の fig.5 で具体的にしてみる。

例えば、1 列目だと、 $a(k) = 1$ になり、

$$\begin{aligned} P_{z\text{chosen}}(k=1) &= \prod_{i=2, z_{i,1}=1} \frac{N_{n1}}{i} \prod_{i=2, z_{i,1}=0} \frac{i - N_{n1}}{i} = \prod_{i=2}^6 \frac{i-1}{i} \frac{1}{7} \prod_{i=8}^{13} \frac{i-2}{i} \frac{2}{14} \frac{3}{15} \frac{12}{16} \frac{4}{17} \prod_{i=18}^{20} \frac{i-5}{i} \\ &= 1 \frac{(1 \cdots (m_{k=1}-1))(1 \cdots (N - m_{k=1}))}{20!} = a(k=1) \frac{(m_{k=1}-1)!(N - m_{k=1})!}{N!} \end{aligned} \quad (12)$$

となる。 $a(k)=1$ となるところに関してはこれが、成り立っていることがわかりやすい。それ以外のところも同様に確認できる。

例えば 22 列目を考える。 $a(k=1) = 3$ になる。

$$\begin{aligned} P_{z\text{chosen}}(k=22) &= \prod_{i=4, z_{i,22}=1} \frac{N_{n22}}{i} \prod_{i=4, z_{i,22}=0} \frac{i - N_{n22}}{i} = \prod_{i=4}^{10} \frac{i-1}{i} \frac{1}{11} \prod_{i=12}^{19} \frac{i-2}{i} \frac{2}{20} \\ &= \frac{(1 \cdots (m_{k=22}-1))(3 \cdots (N - m_{k=22}))}{(4 \cdots 20)} = 3 \frac{(1 \cdots (m_{k=22}-1))(1 \cdot 2)(3 \cdots (N - m_{k=22}))}{(1 \cdot 2)3(4 \cdots 20)} \\ &= a(k) \frac{(m_{k=22}-1)!(N - m_{k=22})!}{N!} \end{aligned} \quad (13)$$

この様になるので、すべての k に対して、

$$P_{z\text{chosen}}(k) = a(k) \frac{(m_k - 1)!(N - m_k)!}{N!} \quad (14)$$

よって、

$$P_{z\text{chosen}}(Z) = \prod_{k=1}^{K_+} a(k) \frac{(m_k - 1)!(N - m_k)!}{N!} \quad (15)$$

これから、

$$P(Z) = \prod_{k=1}^{K_+} a(k) \frac{(m_k - 1)!(N - m_k)!}{N!} \frac{\exp(-\alpha H_N) \alpha^{K_+}}{\prod_{i=1}^N K_1^{(i)}!} \prod_{i=1}^N \left(\frac{1}{i}\right)^{K_1^{(i)}} = \frac{\alpha^{K_+}}{\prod_{i=1}^N K_1^{(i)}!} \exp(-\alpha H_N) \prod_{k=1}^{K_+} \frac{(m_k - 1)!(N - m_k)!}{N!} \quad (16)$$

となり、(35) が成立することがわかる。

論文の (35) の下のパラグラフのように、そのまま列を入れ替えたものは同じ確率で、同じ同値類になるし、そもそも、同じ列であれば、その分、数が減る。

そのため、論文にあるように、 $\frac{\prod_{i=1}^N K_1^{(i)}!}{\prod_{i=1}^{2^N-1} K_h!}$ をかけると、 $P([Z])$ になる。

よって、

$$P([Z]) = \frac{\prod_{i=1}^N K_1^{(i)}!}{\prod_{i=1}^{2^N-1} K_h!} P(Z) = \frac{\alpha^{K_+}}{\prod_{i=1}^{2^N-1} K_h!} \exp(-\alpha H_N) \prod_{k=1}^{K_+} \frac{(m_k - 1)!(N - m_k)!}{N!} \quad (17)$$

となり、この手続で、(34) のサンプリングができる。

論文では $\beta = 1$ だが、これをもとに戻す。このとき、(10) が以下のようにになる。

$$\begin{aligned} P_{z\text{new}}(Z) &= \prod_{i=1}^N \left(\frac{\left(\frac{\alpha\beta}{i+\beta-1}\right)^{H_1^{(i)}}}{H_1^{(i)}!} \exp\left(-\frac{\alpha\beta}{i+\beta-1}\right) \right) = \exp\left(-\alpha \sum_{i=1}^N \frac{\beta}{i+\beta-1}\right) \prod_{i=1}^N \left(\left(\frac{1}{i+\beta-1}\right)^{H_1^{(i)}} \frac{(\alpha\beta)^{H_1^{(i)}}}{H_1^{(i)}!} \right) \\ &= \exp(-\alpha \bar{H}_+) (\alpha\beta)^{H_+} \prod_{i=1}^N \left(\left(\frac{1}{i}\right)^{H_1^{(i)}} \frac{1}{H_1^{(i)}!} \right) = \frac{\exp(-\alpha \bar{H}_+) (\alpha\beta)^{H_+}}{\prod_{i=1}^N H_1^{(i)}!} \prod_{i=1}^N \left(\frac{1}{i+\beta-1} \right)^{H_1^{(i)}} \end{aligned} \quad (18)$$

(11) は、

$$P_{z\text{chosen}}(Z) = \prod_h^{H_+} \left(\prod_{i=a(h)+1, z_{i,h}=1} \frac{N_{nh}}{i+\beta-1} \prod_{i=a(h)+1, z_{i,h}=0} \frac{i+\beta-1-N_{nh}}{i+\beta-1} \right) \quad (19)$$

もちろん途中も変更になるが、(15) は、

$$P_{z\text{chosen}}(Z) = \prod_{h=1}^{H_+} (a(h) + \beta - 1) \frac{(N_h - 1)!(N - N_h + \beta - 1)!}{(N + \beta - 1)!} \quad (20)$$

(16) は

$$\begin{aligned} P(Z) &= \prod_{h=1}^{H_+} (a(h) + \beta - 1) \frac{(N_h - 1)!(N - N_h + \beta - 1)!}{(N + \beta - 1)!} \frac{\exp(-\alpha \bar{H}_+) (\alpha\beta)^{H_+}}{\prod_{i=1}^N H_1^{(i)}!} \prod_{i=1}^N \left(\frac{1}{i+\beta-1} \right)^{H_1^{(i)}} \\ &= \frac{(\alpha\beta)^{H_+}}{\prod_{i=1}^N H_1^{(i)}!} \exp(-\alpha \bar{H}_+) \prod_{h=1}^{H_+} \frac{(N_h - 1)!(N - N_h + \beta - 1)!}{(N + \beta - 1)!} \end{aligned} \quad (21)$$

(17) は

$$P([Z]) = \frac{(\alpha\beta)^{H_+}}{\prod_{i=1}^{2^N-1} H_i!} \exp(-\alpha\overline{H}_+) \prod_{h=1}^{H_+} \frac{(N_h - 1)!(N - N_h + \beta - 1)!}{(N + \beta - 1)!} = \frac{(\alpha\beta)^{H_+}}{\prod_{i=1}^{2^N-1} H_i!} \exp(-\alpha\overline{H}_+) \prod_{h=1}^{H_+} \frac{\Gamma(N_h)\Gamma(N - N_h + \beta)}{\Gamma(N + \beta)} \quad (22)$$

となり、本の (6.51) のようにサンプリングできていることがわかる。

P.173 の下にある、生成されるバイナリの特徴は論文の 4.6 に記載がある。客一人あたりの料理の数は $Poi(\alpha)$ に従うとあり、論文でも交換可能であることを理由に成り立つとしている。おそらく成り立つが、頑張ったら、数式でも示せそう。他の特性はポアソン分布の特性などを利用すると比較的わかりやすい。