Distribución predictiva para procesos normalizados

Ramsés H. Mena

Complemento de curso sobre medidas aleatorias

Distribución predictiva

Teo: Si $\{X_i\}_{i=1}^{\infty}$ son intercambiables con medida de de Finetti, N-Prior(α, ν), entonces

$$\mathbb{P}(X_2 \in B \mid X_1) = \frac{\alpha(B)}{a}[a - \mathcal{I}_a] + \delta_{X_1}(B)\mathcal{I}_a \tag{1}$$

donde

$$\mathcal{I}_a := a \int_{\mathbb{R}_+} u\, e^{-a\,\psi(u)} \int_{\mathbb{R}_+} v^2 e^{-uv} \nu(dv) du$$

y

$$\psi(u) = \int_{\mathbb{R}_+} (1 - e^{-xu}) \nu(dx)$$

Caso más general

Si definimos

$$\Pi_k^{(n)}(n_1,\ldots,n_k) := \frac{a^k}{\Gamma(n)} \int_{\mathbb{R}_+} u^{n-1} e^{-a\psi(u)} \prod_{j=1}^{k} \mu_{n_j}(u)du \tag{2}$$

donde

$$\mu_n(u) := \int_{\mathbb{R}_+} v^n e^{-uv} \nu(dv)$$

Entonces

$$\mathbb{P}(X_{n+1} \in B \mid \mathbf{X}^{(n)}) = w_0^{(n)} \frac{\alpha(B)}{a} + \frac{1}{n} \sum_{j=1}^{k} w_j^n \delta_{X_j^*}(B) \tag{3}$$

donde $\{X_j^*\}_{j=1}^k$ denotan los k valores distintos en $\mathbf{X}^{(n)} := (X_1,\ldots,X_n)$ y

Caso más general

$$w_0^{(n)} = \frac{a \prod_{k+1}^{(n+1)}(n_1, \dots, n_k, 1)}{n \prod_k^{(n)}(n_1, \dots, n_k)}$$

y para $j = 1, \dots, k$

$$w_j^{(n)} = \frac{\prod_k^{(n+1)}(n_1, \dots, n_j + 1, n_k)}{\prod_k^{(n)}(n_1, \dots, n_k)}$$

Interpretación de FPPI

A $\Pi_k^{(n)}(n_1, \ldots, n_k)$ se le conoce como la Función de Probabilidad sobre Particiones Intercambiables (FPPI) y se interpreta como:

La probabilidad de obtener k valores diferentes, en una muestra intercambiable discreta de tamaño n, con la configuración de frecuencias $\mathbf{n}_k := (n_1, \ldots, n_k)$

$\Rightarrow$ Los pesos de la distribución predictiva dependen únicamente de cuantos valores distintos k hay en la muestra $\mathbf{X}^{(n)}$ y las frecuencias absolutas con esos que se repiten $\mathbf{n}_k$.

Interpretación de FPPI

A $\Pi_k^{(n)}(n_1, \ldots, n_k)$ se le conoce como la Función de Probabilidad sobre Particiones Intercambiables (FPPI) y se interpreta como:

La probabilidad de obtener k valores diferentes, en una muestra intercambiable discreta de tamaño n, con la configuración de frecuencias $\mathbf{n}_k := (n_1, \ldots, n_k)$

$\Rightarrow$ Los pesos de la distribución predictiva dependen únicamente de cuantos valores distintos k hay en la muestra $\mathbf{X}^{(n)}$ y las frecuencias absolutas con esos que se repiten $\mathbf{n}_k$.

El Proceso de Dirichlet

- $\nu(\mathrm{d}x) = \frac{e^{-x}}{x}\,\mathrm{d}x$

$$\mathcal{I}_a \quad = \quad \text{TAREA} \; = \frac{1}{a+1} \tag{4}$$

El Proceso de Dirichlet

$$\mu_n(u) = \int_{\mathbb{R}_+} v^n e^{-uv} \nu(dv) = \int_0^\infty v^{n-1} e^{-(u+1)v} \mathrm{d}v = \frac{\Gamma(n)}{(1+u)^n} \tag{5}$$

Entonces simplificando

$$\Pi_k^{(n)}(n_1, \ldots, n_k) \;=\; \text{TAREA} \;=\; \frac{a^k}{(a)_n} \prod_{i=1}^{k} (n_i - 1)! \tag{6}$$

El Proceso de Dirichlet

Entonces,

$$w_0^{(n)} = \frac{a}{n}\frac{\Gamma(n+1)\Gamma(n+a)}{\Gamma(n+a+1)\Gamma(n)} = \frac{a}{a+n}$$

y

$$w_j^{(n)} = n_j\frac{\Gamma(n+1)\Gamma(n+a)}{\Gamma(n+a+1)\Gamma(n)} = n_j\frac{n}{n+a}$$

para $j = 1, \ldots, k$. Finalmente la distribución predictiva se simplifica como

$$\mathbb{P}(X_{n+1} \in B \mid \mathbf{X}^{(n)}) = \frac{a}{a+n}\frac{\alpha(B)}{a} + \sum_{j=1}^{k}\frac{n_j}{a+n}\delta_{X_j^*}(B) \tag{7}$$

El Proceso de estable normalizado

- $\nu(\mathrm{d}x) = c\,x^{-(1+\gamma)}\,\mathrm{d}x, \quad \gamma \in (0,1)$

$$
\begin{aligned}
\mathcal{I}_a &= a\int_0^\infty u\,e^{-a\psi(u)}\int_0^\infty v^2 e^{-uv}\nu\,(dv)\,du \\
&= a\int_0^\infty u\,e^{-a\,\frac{c\Gamma(1-\gamma)u^\gamma}{\gamma}}\int_0^\infty v^{1-\gamma}e^{-uv}dv\,du \\
&= a\,c\,\Gamma\,(2-\gamma)\int_0^\infty u^{\gamma-1}\,e^{-a\,\frac{c\Gamma(1-\gamma)u^\gamma}{\gamma}}du \\
&= 1-\gamma
\end{aligned}
$$

El Proceso de estable normalizado

$$\begin{aligned}\mu_n(u) &= \int_0^\infty v^n e^{-uv} \nu(dv) \\ &= c\int_0^\infty v^{n-1-\gamma} e^{-uv} dv \\ &= cu^{-n+\gamma}\Gamma(n-\gamma).\end{aligned}$$

También

$$\Pi_{k+1}^{(n+1)}(n_1,\ldots,n_k,1) = c^{k+1}\Gamma(1-\gamma)\prod_{i=1}^{k}\Gamma(n_i-\gamma)\int_0^\infty u^{(k+1)\gamma-1} e^{-a\,\frac{c\Gamma(1-\gamma)u^\gamma}{\gamma}} du$$

$$\Pi_k^{(n)}(n_1,\ldots,n_k) = c^k\prod_{i=1}^{k}\Gamma(n_i-\gamma)\int_0^\infty u^{k\gamma-1} e^{-a\,\frac{c\Gamma(1-\gamma)u^\gamma}{\gamma}} du.$$

El Proceso de estable normalizado

Entonces

$$w_0^{(n)} = \frac{a}{n} \frac{c^{k+1}\Gamma\left(1-\gamma\right)\prod_{i=1}^{k}\Gamma\left(n_i-\gamma\right)\int_0^\infty u^{(k+1)\gamma-1}e^{-a\frac{c\Gamma(1-\gamma)u^\gamma}{\gamma}}du}{c^k\prod_{i=1}^{k}\Gamma\left(n_i-\gamma\right)\int_0^\infty u^{k\gamma-1}e^{-a\frac{c\Gamma(1-\gamma)u^\gamma}{\gamma}}du} = \frac{k}{n}\gamma$$

y

$$\begin{aligned}\Pi_k^{(n+1)}\left(n_1,\ldots,n_j+1,\ldots,n_k\right) &= c^k\left(n_j-1\right)\prod_{i=1}^{k}\Gamma\left(n_i-\gamma\right)\times \\ &\quad \int_0^\infty u^{k\gamma-1}e^{-a\frac{c\Gamma(1-\gamma)u^\gamma}{\gamma}}du.\end{aligned}$$

El Proceso de estable normalizado

Resultando en

$$w_j^{(n)} = \frac{c^k \left(n_j - 1\right) \prod_{i=1}^{k} \Gamma\left(n_i - \gamma\right) \int_0^\infty u^{k\gamma-1} e^{-a \frac{c\Gamma(1-\gamma)u^\gamma}{\gamma}} du}{c^k \prod_{i=1}^{k} \Gamma\left(n_i - \gamma\right) \int_0^\infty u^{k\gamma-1} e^{-a \frac{c\Gamma(1-\gamma)u^\gamma}{\gamma}} du} = \left(n_j - \gamma\right).$$

y por lo tanto

$$P\left(X_{n+1} \in B \mid \mathbf{X}^{(n)}\right) = \gamma \frac{k}{n} \frac{\alpha\left(B\right)}{a} + \frac{1}{n} \sum_{j=1}^{k} \left(n_j - \gamma\right) \delta_{X_j^*}\left(B\right).$$

Entendiendo la FPPI

Dado un conjunto finito F, una partición de F en k grupos es una colección de conjuntos disjuntos no-vacíos $\{A_1, A_2, \ldots, A_k\}$ cuya unión es F, i.e. $F = \cup_{i=1}^{k} A_j$, con $A_j \neq \emptyset\ \forall j$ y $A_i \cap A_j = \emptyset$ para $i \neq j$

- $\mathcal{P}_{[n]}^{k}$ Conjunto de particiones del conjunto $[n] := \{1, \ldots, n\}$ en k grupos, e.g. $\mathcal{P}_{[4]}^{2} = \{\{123, 4\}, \{124, 3\}, \{12, 34\}, \{134, 2\}, \{13, 24\}, \{14, 23\}, \{1, 234\}\}$

$S_{n,k} = \frac{1}{k!} \sum_{i=0}^{k} (-1)^i \binom{k}{i} (k-i)^n$ $\rightarrow$Números de Stirling del 2^{do} tipo, indican el número de formas de particionar un conjunto $[n]$ en k conjuntos no-vacíos $S_{n,k} = \#\mathcal{P}_{[n]}^{k}$,e.g. $S(4,2) = 7$

- $\mathcal{P}_{[n]} := \cup_{k=1}^{n} \mathcal{P}_{[n]}^{k}$ Conjunto de todas las particiones de $[n]$

$B_n := \sum_{k=1}^{n} S_{n,k} = \sum_{k=1}^{n} \frac{k^n}{k!} \sum_{j=0}^{n-k} \frac{(-1^j)}{j!}$ $\rightarrow$Números de Bell, e.g. $B_1 = 1, B_3 = 5, B_4 = 15, B_5 = 52, \ldots, B_{10} = 115975$

Si $|A_j| = \{$el número de elementos en $A_j\}$ entonces la sucesión $(|A_1|, \ldots, |A_k|)$ de tamaños de grupos de una partición de $[n]$ define una

Entendiendo la FPPI

Dado un conjunto finito F, una partición de F en k grupos es una colección de conjuntos disjuntos no-vacíos $\{A_1, A_2, \ldots, A_k\}$ cuya unión es F, i.e. $F = \cup_{i=1}^{k} A_j$, con $A_j \neq \emptyset \; \forall j$ y $A_i \cap A_j = \emptyset$ para $i \neq j$

- $\mathcal{P}_{[n]}^{k}$ Conjunto de particiones del conjunto $[n] := \{1, \ldots, n\}$ en k grupos, e.g. $\mathcal{P}_{[4]}^{2} = \{\{123, 4\}, \{124, 3\}, \{12, 34\}, \{134, 2\}, \{13, 24\}, \{14, 23\}, \{1, 234\}\}$

$S_{n,k} = \frac{1}{k!}\sum_{i=0}^{k}(-1)^i\binom{k}{i}(k-i)^n$ →Números de Stirling del 2^{do} tipo, indican el número de formas de particionar un conjunto $[n]$ en k conjuntos no-vacíos $S_{n,k} = \#\mathcal{P}_{[n]}^{k}$,e.g. $S(4,2) = 7$

- $\mathcal{P}_{[n]} := \cup_{k=1}^{n}\mathcal{P}_{[n]}^{k}$ Conjunto de todas las particiones de $[n]$

$B_n := \sum_{k=1}^{n} S_{n,k} = \sum_{k=1}^{n}\frac{k^n}{k!}\sum_{j=0}^{n-k}\frac{(-1^j)}{j!}$ →Números de Bell, e.g. $B_1 = 1, B_3 = 5, B_4 = 15, B_5 = 52, \ldots, B_{10} = 115975$

Si $|A_j| = \{$el número de elementos en $A_j\}$ entonces la sucesión $(|A_1|, \ldots, |A_k|)$ de tamaños de grupos de una partición de $[n]$ define una

Entendiendo la FPPI

Dado un conjunto finito F, una partición de F en k grupos es una colección de conjuntos disjuntos no-vacíos $\{A_1, A_2, \ldots, A_k\}$ cuya unión es F, i.e. $F = \cup_{i=1}^{k} A_j$, con $A_j \neq \emptyset \; \forall j$ y $A_i \cap A_j = \emptyset$ para $i \neq j$

- $\mathcal{P}_{[n]}^{k}$ Conjunto de particiones del conjunto $[n] := \{1, \ldots, n\}$ en k grupos, e.g. $\mathcal{P}_{[4]}^{2} = \{\{123, 4\}, \{124, 3\}, \{12, 34\}, \{134, 2\}, \{13, 24\}, \{14, 23\}, \{1, 234\}\}$

$S_{n,k} = \frac{1}{k!}\sum_{i=0}^{k}(-1)^i\binom{k}{i}(k-i)^n$ →Números de Stirling del 2^{do} tipo, indican el número de formas de particionar un conjunto $[n]$ en k conjuntos no-vacíos $S_{n,k} = \#\mathcal{P}_{[n]}^{k}$,e.g. $S(4, 2) = 7$

- $\mathcal{P}_{[n]} := \cup_{k=1}^{n}\mathcal{P}_{[n]}^{k}$ Conjunto de todas las particiones de $[n]$

$B_n := \sum_{k=1}^{n} S_{n,k} = \sum_{k=1}^{n}\frac{k^n}{k!}\sum_{j=0}^{n-k}\frac{(-1^j)}{j!}$ →Números de Bell, e.g. $B_1 = 1, B_3 = 5, B_4 = 15, B_5 = 52, \ldots, B_{10} = 115975$

Si $|A_j| = \{$el número de elementos en $A_j\}$ entonces la sucesión $(|A_1|, \ldots, |A_k|)$ de tamaños de grupos de una partición de $[n]$ define una

Entendiendo la FPPI

Dado un conjunto finito F, una partición de F en k grupos es una colección de conjuntos disjuntos no-vacíos $\{A_1, A_2, \dots, A_k\}$ cuya unión es F, i.e. $F = \cup_{i=1}^{k} A_j$, con $A_j \neq \emptyset \ \forall j$ y $A_i \cap A_j = \emptyset$ para $i \neq j$

- $\mathcal{P}_{[n]}^{k}$ Conjunto de particiones del conjunto $[n] := \{1, \dots, n\}$ en k grupos, e.g. $\mathcal{P}_{[4]}^{2} = \{\{123, 4\}, \{124, 3\}, \{12, 34\}, \{134, 2\}, \{13, 24\}, \{14, 23\}, \{1, 234\}\}$

$S_{n,k} = \frac{1}{k!}\sum_{i=0}^{k}(-1)^i\binom{k}{i}(k-i)^n$ →Números de Stirling del 2^{do} tipo, indican el número de formas de particionar un conjunto $[n]$ en k conjuntos no-vacíos $S_{n,k} = \#\mathcal{P}_{[n]}^{k}$,e.g. $S(4,2) = 7$

- $\mathcal{P}_{[n]} := \cup_{k=1}^{n}\mathcal{P}_{[n]}^{k}$ Conjunto de todas las particiones de $[n]$

$B_n := \sum_{k=1}^{n} S_{n,k} = \sum_{k=1}^{n}\frac{k^n}{k!}\sum_{j=0}^{n-k}\frac{(-1^j)}{j!}$ →Números de Bell, e.g. $B_1 = 1, B_3 = 5, B_4 = 15, B_5 = 52, \dots, B_{10} = 115975$

Si $|A_j| = \{$el número de elementos en $A_j\}$ entonces la sucesión $(|A_1|, \dots, |A_k|)$ de tamaños de grupos de una partición de $[n]$ define una

Entendiendo la FPPI

Dado un conjunto finito F, una partición de F en k grupos es una colección de conjuntos disjuntos no-vacíos $\{A_1, A_2, \ldots, A_k\}$ cuya unión es F, i.e. $F = \cup_{i=1}^{k} A_j$, con $A_j \neq \emptyset \; \forall j$ y $A_i \cap A_j = \emptyset$ para $i \neq j$

- $\mathcal{P}_{[n]}^{k}$ Conjunto de particiones del conjunto $[n] := \{1, \ldots, n\}$ en k grupos, e.g. $\mathcal{P}_{[4]}^{2} = \{\{123, 4\}, \{124, 3\}, \{12, 34\}, \{134, 2\}, \{13, 24\}, \{14, 23\}, \{1, 234\}\}$

$S_{n,k} = \frac{1}{k!} \sum_{i=0}^{k} (-1)^i \binom{k}{i} (k-i)^n$ →Números de Stirling del 2^{do} tipo, indican el número de formas de particionar un conjunto $[n]$ en k conjuntos no-vacíos $S_{n,k} = \#\mathcal{P}_{[n]}^{k}$,e.g. $S(4, 2) = 7$

- $\mathcal{P}_{[n]} := \cup_{k=1}^{n} \mathcal{P}_{[n]}^{k}$ Conjunto de todas las particiones de $[n]$

$B_n := \sum_{k=1}^{n} S_{n,k} = \sum_{k=1}^{n} \frac{k^n}{k!} \sum_{j=0}^{n-k} \frac{(-1^j)}{j!}$ →Números de Bell, e.g. $B_1 = 1, B_3 = 5, B_4 = 15, B_5 = 52, \ldots, B_{10} = 115975$

Si $|A_j| = \{$el número de elementos en $A_j\}$ entonces la sucesión $(|A_1|, \ldots, |A_k|)$ de tamaños de grupos de una partición de $[n]$ define una

Estructuras combinatóricas

- $\mathcal{C}_n$: conjunto de composiciones de n e.g. (→orden importa!)
  $\mathcal{C}_4 = \{[4],[1,3],[1,1,1,1],[3,1],[2,2],[1,2,1],[2,1,1],[1,1,2]\}$
- ρ_n: partición de un entero n (no $[n]$!!) e.g.(→orden no importa!)
  $\rho_4 = \{[4],[3,1],[2,2],[2,1,1],[1,1,1,1]\}$

Ex: $\mathscr{P}_{[3]} = \{\{\{1\},\{2\},\{3\}\},\{\{2,3\},\{1\}\},\{\{1,3\},\{2\}\},\{\{3\},\{1,2\}\},\{\{1,2,3\}\}\}$

$$\mathcal{C}_3 = \{(1,1,1),(2,1),(1,2),(3)\}$$

Entonces se necesita cumplir

$$\Pi_3^{(3)}(1,1,1) + 3 * \Pi_2^{(3)}(2,1) + \Pi_1^{(3)}(3) = 1$$

en general

$$\sum_{k=1}^{n} \sum_{(n_1,\ldots,n_k)} \binom{n}{n_1,\ldots,n_k} \frac{1}{k!} \Pi_k^{(n)}(n_1,\ldots,n_k) = 1,$$

donde la segunda suma corre sobre todas las composiciones de n con k partes

$\Rightarrow$ Se reduce de B_n a $\#\mathcal{C}_n = 2^{n-1}$, todavía muy grande!!

Estructuras combinatóricas

- $\mathcal{C}_n$: conjunto de composiciones de n e.g. ($\rightarrow$orden importa!)
  $\mathcal{C}_4 = \{[4],[1,3],[1,1,1,1],[3,1],[2,2],[1,2,1],[2,1,1],[1,1,2]\}$
- ρ_n: partición de un entero n (no $[n]$!!) e.g.($\rightarrow$orden no importa!)
  $\rho_4 = \{[4],[3,1],[2,2],[2,1,1],[1,1,1,1]\}$

Ex: $\mathscr{P}_{[3]} = \{\{\{1\},\{2\},\{3\}\},\{\{2,3\},\{1\}\},\{\{1,3\},\{2\}\},\{\{3\},\{1,2\}\},\{\{1,2,3\}\}\}$

$$\mathcal{C}_3 = \{(1,1,1),(2,1),(1,2),(3)\}$$

Entonces se necesita cumplir

$$\Pi_3^{(3)}(1,1,1) + 3 * \Pi_2^{(3)}(2,1) + \Pi_1^{(3)}(3) = 1$$

en general

$$\sum_{k=1}^{n} \sum_{(n_1,\ldots,n_k)} \binom{n}{n_1,\ldots,n_k} \frac{1}{k!} \Pi_k^{(n)}(n_1,\ldots,n_k) = 1,$$

donde la segunda suma corre sobre todas las composiciones de n con k partes

$\Rightarrow$ Se reduce de B_n a $\#\mathcal{C}_n = 2^{n-1}$, todavía muy grande!!

Estructuras combinatóricas

- $\mathcal{C}_n$: conjunto de composiciones de n e.g. ($\rightarrow$orden importa!)
  $\mathcal{C}_4 = \{[4],[1,3],[1,1,1,1],[3,1],[2,2],[1,2,1],[2,1,1],[1,1,2]\}$
- ρ_n: partición de un entero n (no $[n]$!!) e.g.($\rightarrow$orden no importa!)
  $\rho_4 = \{[4],[3,1],[2,2],[2,1,1],[1,1,1,1]\}$

Ex: $\mathscr{P}_{[3]} = \{\{\{1\},\{2\},\{3\}\},\{\{2,3\},\{1\}\},\{\{1,3\},\{2\}\},\{\{3\},\{1,2\}\},\{\{1,2,3\}\}\}$

$$\mathcal{C}_3 = \{(1,1,1),(2,1),(1,2),(3)\}$$

Entonces se necesita cumplir

$$\Pi_3^{(3)}(1,1,1) + 3 * \Pi_2^{(3)}(2,1) + \Pi_1^{(3)}(3) = 1$$

en general

$$\sum_{k=1}^{n} \sum_{(n_1,\ldots,n_k)} \binom{n}{n_1,\ldots,n_k} \frac{1}{k!} \Pi_k^{(n)}(n_1,\ldots,n_k) = 1,$$

donde la segunda suma corre sobre todas las composiciones de n con k partes

$\Rightarrow$ Se reduce de B_n a $\#\mathcal{C}_n = 2^{n-1}$, todavía muy grande!!

Estructuras combinatóricas

- $\mathcal{C}_n$: conjunto de composiciones de n e.g. ($\rightarrow$orden importa!)
  $\mathcal{C}_4 = \{[4], [1,3], [1,1,1,1], [3,1], [2,2], [1,2,1], [2,1,1], [1,1,2]\}$
- ρ_n: partición de un entero n (no $[n]$!!) e.g.($\rightarrow$orden no importa!)
  $\rho_4 = \{[4], [3,1], [2,2], [2,1,1], [1,1,1,1]\}$

Ex: $\mathscr{P}_{[3]} = \{\{\{1\},\{2\},\{3\}\}, \{\{2,3\},\{1\}\}, \{\{1,3\},\{2\}\}, \{\{3\},\{1,2\}\}, \{\{1,2,3\}\}\}$

$$\mathcal{C}_3 = \{(1,1,1), (2,1), (1,2), (3)\}$$

Entonces se necesita cumplir

$$\Pi_3^{(3)}(1,1,1) + 3 * \Pi_2^{(3)}(2,1) + \Pi_1^{(3)}(3) = 1$$

en general

$$\sum_{k=1}^{n} \sum_{(n_1,\ldots,n_k)} \binom{n}{n_1,\ldots,n_k} \frac{1}{k!} \Pi_k^{(n)}(n_1,\ldots,n_k) = 1,$$

donde la segunda suma corre sobre todas las composiciones de n con k partes

$\Rightarrow$ Se reduce de B_n a $\#\mathcal{C}_n = 2^{n-1}$, todavía muy grande!!

Estructuras combinatóricas

- $\mathcal{C}_n$: conjunto de composiciones de n e.g. ($\rightarrow$orden importa!)
  $\mathcal{C}_4 = \{[4],[1,3],[1,1,1,1],[3,1],[2,2],[1,2,1],[2,1,1],[1,1,2]\}$
- ρ_n: partición de un entero n (no $[n]$!!) e.g.($\rightarrow$orden no importa!)
  $\rho_4 = \{[4],[3,1],[2,2],[2,1,1],[1,1,1,1]\}$

Ex: $\mathscr{P}_{[3]} = \{\{\{1\},\{2\},\{3\}\},\{\{2,3\},\{1\}\},\{\{1,3\},\{2\}\},\{\{3\},\{1,2\}\},\{\{1,2,3\}\}\}$

$$\mathcal{C}_3 = \{(1,1,1),(2,1),(1,2),(3)\}$$

Entonces se necesita cumplir

$$\Pi_3^{(3)}(1,1,1) + 3 * \Pi_2^{(3)}(2,1) + \Pi_1^{(3)}(3) = 1$$

en general

$$\sum_{k=1}^{n} \sum_{(n_1,\ldots,n_k)} \binom{n}{n_1,\ldots,n_k} \frac{1}{k!} \Pi_k^{(n)}(n_1,\ldots,n_k) = 1,$$

donde la segunda suma corre sobre todas las composiciones de n con k partes

$\Rightarrow$ Se reduce de B_n a $\#\mathcal{C}_n = 2^{n-1}$, todavía muy grande!!

Estructura de agrupamiento

- Una muestra de tamaño n, $X_i \overset{\text{iid}}{\sim} P$, $i = 1, \ldots, n$ con $P \sim \text{N-Prior}(\alpha, \nu)$
- $\Rightarrow P[X_i = X_j] > 0, \forall i \neq j$

Entonces los n datos pueden particionarse en en:

- $K_n = k$ clases, tal que dos observaciones pertenecen a la misma clase si éstas coinciden.
- Las K_n clases tienen frecuencias $N_{1,n}, N_{2,n}, \ldots, N_{K_n,n}$ tal que $\sum_{i=1}^{K_n} N_{i,n} = n$

Estructura de agrupamiento

- Una muestra de tamaño n, $X_i \overset{\text{iid}}{\sim} P$, $i = 1, \ldots, n$ con $P \sim \text{N-Prior}(\alpha, \nu)$
- $\Rightarrow P[X_i = X_j] > 0, \forall i \neq j$

Entonces los n datos pueden particionarse en en:

- $K_n = k$ clases, tal que dos observaciones pertenecen a la misma clase si éstas coinciden.
- Las K_n clases tienen frecuencias $N_{1,n}, N_{2,n}, \ldots, N_{K_n,n}$ tal que $\sum_{i=1}^{K_n} N_{i,n} = n$

Estructura de agrupamiento

- Una muestra de tamaño n, $X_i \overset{\text{iid}}{\sim} P$, $i = 1, \ldots, n$ con $P \sim \text{N-Prior}(\alpha, \nu)$
- $\Rightarrow P[X_i = X_j] > 0, \forall i \neq j$

Entonces los n datos pueden particionarse en en:

- $K_n = k$ clases, tal que dos observaciones pertenecen a la misma clase si éstas coinciden.
- Las K_n clases tienen frecuencias $N_{1,n}, N_{2,n}, \ldots, N_{K_n,n}$ tal que $\sum_{i=1}^{K_n} N_{i,n} = n$

Estructura de agrupamiento

Para cada $n \geq 1$ y $1 \leq k \leq n$ sea

$$\Pi_k^{(n)}(n_1, \ldots, n_k) = \mathrm{P}(\text{datos en } k \text{ clases con frecuencias } n_1, \ldots, n_k)$$

$\Rightarrow$ Una FPPI bien definida!

(1) $\Pi_k^{(n)}(n_1, \ldots, n_k)$ es invariante con respecto a permutaciones de $n_1, \ldots, n_k$ (intercambiable)

(2) Se tiene la siguiente regla de adición

$$\Pi_k^{(n)}(n_1, \ldots, n_k) = \Pi_{k+1}^{(n+1)}(n_1, \ldots, n_k, 1) + \sum_{j=1}^{k} \Pi_{k+1}^{(n+1)}(n_1, \ldots, n_j + 1, \ldots, n_k)$$

La colección $\{\Pi_k^{(n)} : 1 \leq k \leq n, \ k \geq 1\}$ define una FPPI

Estructura de agrupamiento

Para cada $n \geq 1$ y $1 \leq k \leq n$ sea

$$\Pi_k^{(n)}(n_1, \ldots, n_k) = \mathrm{P}(\text{datos en } k \text{ clases con frecuencias } n_1, \ldots, n_k)$$

$\Rightarrow$ Una FPPI bien definida!

(1) $\Pi_k^{(n)}(n_1, \ldots, n_k)$ es invariante con respecto a permutaciones de $n_1, \ldots, n_k$ (intercambiable)

(2) Se tiene la siguiente regla de adición

$$\Pi_k^{(n)}(n_1, \ldots, n_k) = \Pi_{k+1}^{(n+1)}(n_1, \ldots, n_k, 1) + \sum_{j=1}^{k} \Pi_{k+1}^{(n+1)}(n_1, \ldots, n_j + 1, \ldots, n_k)$$

La colección $\{\Pi_k^{(n)} : 1 \leq k \leq n,\ k \geq 1\}$ define una FPPI

Estructura de agrupamiento

Para cada $n \geq 1$ y $1 \leq k \leq n$ sea

$$\Pi_k^{(n)}(n_1, \ldots, n_k) = \mathrm{P}(\text{datos en } k \text{ clases con frecuencias } n_1, \ldots, n_k)$$

$\Rightarrow$ Una FPPI bien definida!

(1) $\Pi_k^{(n)}(n_1, \ldots, n_k)$ es invariante con respecto a permutaciones de $n_1, \ldots, n_k$ (intercambiable)

(2) Se tiene la siguiente regla de adición

$$\Pi_k^{(n)}(n_1, \ldots, n_k) = \Pi_{k+1}^{(n+1)}(n_1, \ldots, n_k, 1) + \sum_{j=1}^{k} \Pi_{k+1}^{(n+1)}(n_1, \ldots, n_j + 1, \ldots, n_k)$$

La colección $\{\Pi_k^{(n)} : 1 \leq k \leq n, \ k \geq 1\}$ define una FPPI

Estructura de agrupamiento

Para cada $n \geq 1$ y $1 \leq k \leq n$ sea

$$\Pi_k^{(n)}(n_1, \ldots, n_k) = \mathrm{P}(\text{datos en } k \text{ clases con frecuencias } n_1, \ldots, n_k)$$

$\Rightarrow$ Una FPPI bien definida!

(1) $\Pi_k^{(n)}(n_1, \ldots, n_k)$ es invariante con respecto a permutaciones de $n_1, \ldots, n_k$ (intercambiable)

(2) Se tiene la siguiente regla de adición

$$\Pi_k^{(n)}(n_1, \ldots, n_k) = \Pi_{k+1}^{(n+1)}(n_1, \ldots, n_k, 1) + \sum_{j=1}^{k} \Pi_{k+1}^{(n+1)}(n_1, \ldots, n_j + 1, \ldots, n_k)$$

La colección $\{\Pi_k^{(n)} : 1 \leq k \leq n, \ k \geq 1\}$ define una FPPI

Distribución predictiva

Supongamos que $X_1, \dots, X_n$ toman k valores distintos $X_1^*, \dots, X_k^*$ y n_j de éstos son iguales a X_j^*. Entonces

$$\mathbb{P}[X_{n+1} = \text{ nuevo} \,|\, X_1, \dots, X_n] = \frac{\Pi_{k+1}^{(n+1)}(n_1, \dots, n_k, 1)}{\Pi_k^{(n)}(n_1, \dots, n_k)}$$

y

$$\mathbb{P}[X_{n+1} = X_j^* \,|\, X_1, \dots, X_n] = \frac{\Pi_k^{(n+1)}(n_1, \dots, n_j + 1, \dots, n_k)}{\Pi_k^{(n)}(n_1, \dots, n_k)}$$

Distribución predictiva

Supongamos que $X_1, \ldots, X_n$ toman k valores distintos $X_1^*, \ldots, X_k^*$ y n_j de éstos son iguales a X_j^*. Entonces

$$\mathbb{P}[X_{n+1} = \text{ nuevo} \mid X_1, \ldots, X_n] = \frac{\Pi_{k+1}^{(n+1)}(n_1, \ldots, n_k, 1)}{\Pi_k^{(n)}(n_1, \ldots, n_k)}$$

y

$$\mathbb{P}[X_{n+1} = X_j^* \mid X_1, \ldots, X_n] = \frac{\Pi_k^{(n+1)}(n_1, \ldots, n_j + 1, \ldots, n_k)}{\Pi_k^{(n)}(n_1, \ldots, n_k)}$$

Distribución predictiva

Supongamos que $X_1, \ldots, X_n$ toman k valores distintos $X_1^*, \ldots, X_k^*$ y n_j de éstos son iguales a X_j^*. Entonces

$$\mathbb{P}[X_{n+1} = \text{ nuevo} \,|\, X_1, \ldots, X_n] = \frac{\Pi_{k+1}^{(n+1)}(n_1, \ldots, n_k, 1)}{\Pi_k^{(n)}(n_1, \ldots, n_k)}$$

y

$$\mathbb{P}[X_{n+1} = X_j^* \,|\, X_1, \ldots, X_n] = \frac{\Pi_k^{(n+1)}(n_1, \ldots, n_j + 1, \ldots, n_k)}{\Pi_k^{(n)}(n_1, \ldots, n_k)}$$

FPPI derivadas del proceso de Dirichlet

Proceso de Dirichlet

$$\Pi_k^{(n)}(n_1, \ldots, n_k) = \frac{a^k}{(a)_n} \prod_{j=1}^{k} (n_j - 1)!$$

La distribución (inicial) para K_n (No. de obs. distintas) está dada por

$$\mathbb{P}[K_n = k] = \frac{a^k}{(a)_n} \, |s(n, k)| \tag{8}$$

donde $s(n, k)$ para $n \geq k \geq 1$ denota los números de Stirling del primer tipo. (Número de permutaciones de n elementos en k grupos)

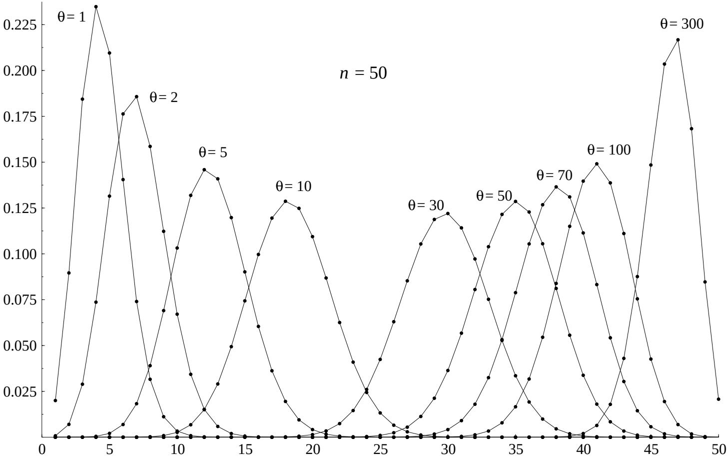


Figure: Distribución de K_n, cuando $\tilde{P} \sim \mathscr{D}_{(\theta,P_0)}$ y θ varía.

FPPI derivadas del proceso de gamma-generalizado

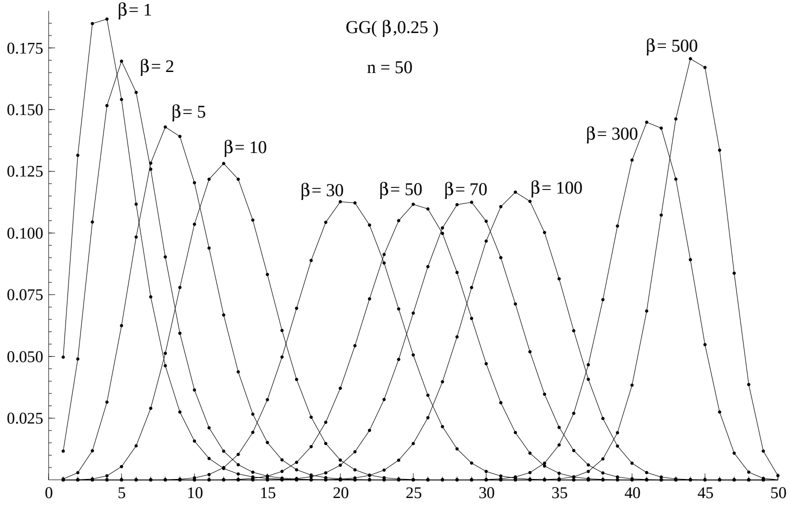
- $\nu(\mathrm{d}x) = \frac{\exp(-\tau x)\, x^{-(1+\sigma)}}{\Gamma(1-\sigma)}\mathrm{d}x$, con $\sigma \in (0,1)$ y $\tau \geq 0$

$$\mathbb{P}[K_n = k] = \frac{\mathrm{e}^{\,\beta}\, \mathscr{C}(n,k,\sigma)}{\sigma\, \Gamma(n)} \sum_{i=0}^{n-1} \binom{n-1}{i} (-1)^i \beta^{i/\sigma}\, \Gamma\left(k - \frac{i}{\sigma};\, \beta\right),$$

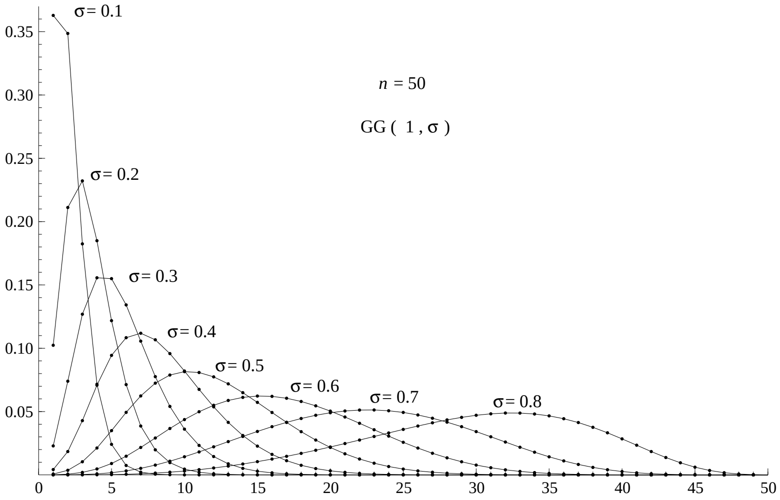
donde $\beta := \tau^\sigma/\sigma$ y

$$\mathscr{C}(n,k;\sigma) := \frac{1}{k!}\sum_{j=0}^{k} (-1)^j \binom{k}{j} (-j\sigma)_n$$

denota el coeficiente factorial generalizado.



- β controls the location of the distribution ($\beta \uparrow \implies \mathbb{E}(K_n) \uparrow$);



- σ controls the flatness ($\sigma \uparrow \implies$ flatness $\uparrow$).

Estadística Bayesiana NP

Si a $\{X_i\}_{i=1}^{\infty}$, un conjunto de v.a.'s con valores en $\mathbb{X}$, se les modela mediante intercambiabilidad con medida de de Finetti no-paramétrica resultante de un subordinador de Lévy normalizado entonces:

- Si $\mathbb{X}$ es discreto:

$$X_i \overset{\text{ind}}{\sim} P, \quad P \overset{\text{iid}}{\sim} \mathscr{P}(aP_0)$$

- Si $\mathbb{X}$ es continuo:

$$X_i \overset{\text{ind}}{\sim} P, \quad P(B) = \int K_\theta(B) G(d\theta), \quad G \overset{\text{iid}}{\sim} \mathscr{P}(aP_0)$$

  donde $K_\theta(\cdot)$ es un kernel absolutamente continuo, y G es un proceso normalizado con medida $\mathscr{P}$ inducida por la selección de ν.

  NOTA: agrupamiento a nivel de los parámetros θ's $\Rightarrow$ agrupamiento a nivel de las observaciones X_i's

Estadística Bayesiana NP

- ¿ Cómo elegir $\mathscr{P}$ $(\nu(\mathrm{d}x)$ dentro del mundo de procesos normalizados?
- Hasta el momento la elección es escoger el modelo tan general como sea posible/operacional/adecuado
- Una vez hecho esto el proceso dependerá de parámetro(s), e.g. a en el Dirichlet, γ en el estable o (σ, β) en el gamma generalizado.
- ¿ Cómo elegir dichos parámetros?
  - Informativo: Si se tiene información acerca del número de grupos subyacente a los datos X_i's, entonces se puede(n) escoger el/los parámetro(s) de tal forma que la $\mathbb{P}[K_n = k]$ correspondiente tenga moda en dicho número.
  - No informativo: Escoger el/los parámetro(s) de tal forma que la $\mathbb{P}[K_n = k]$ sea lo más uniforme posible