演習問題 2.10

ガンマ関数の性質 $\Gamma(\ x\ +\ 1\ )\ =\ x\,\Gamma(\ x\ )$ を用いて、式 (2.38) のディクリレ分布の平均、分散、および共分散の結果を導出せよ。

$$\mathrm{E}[\ \mu_j\ ]\ =\ \frac{\alpha_j}{\alpha_0}$$

… (2.273)

$$\mathrm{var}[\ \mu_j\ ]\ =\ \frac{\alpha_j\ (\ \alpha_0\ -\ \alpha_j\ )}{\alpha_0^2\ (\ \alpha_0\ +\ 1\ )}$$

… (2.274)

$$\mathrm{cov}[\ \mu_j\,\mu_l\ ]\ =\ -\ \frac{\alpha_j\,\alpha_l}{\alpha_0^2\ (\ \alpha_0\ +\ 1\ )}\qquad \text{(ただし、}j\ \neq\ l\text{ である。)}$$

… (2.275)

ただし、α_0 は、式 (2.39) で定義されている。

[期待値]

ある関数 $f\ (\ x\ )$ の確率分布 $p\ (\ x\ )$ の下での平均値のこと。

$$\mathrm{E}[\ f\ ] = \int p\ (\ x\ )\ f\ (\ x\ )\ dx$$

… (1.34)

[分散]

$f\ (\ x\ )$ がその平均値 $\mathrm{E}[\ f\ (\ x\ )\ ]$ の周りでどれぐらいバラつくかを表す尺度のこと。

$$\mathrm{var}[\ f\ ] =\ \mathrm{E}[\ (\ f\ (\ x\ )\ -\ \mathrm{E}[\ f\ (\ x\ )\ ]\ )^2\ ]$$

… (1.38)

$$=\ \mathrm{E}[\ f\ (\ x\ )^2\ ] - \mathrm{E}[\ f\ (\ x\ )\ ]^2$$

… (1.39)

[共分散]

$$\mathrm{cov}[\ x\ ,y\ ] = \mathrm{E}_{x,y}[\ \{\ x - \mathrm{E}[\ x\ ]\ \}\ \{\ y - \mathrm{E}[\ y\ ]\ \}\ ]$$
$$= \mathrm{E}_{x,y}[\ xy\ ] -\ \mathrm{E}[\ x\ ]\ \mathrm{E}[\ y\ ]$$

… (1.41)

[ガンマ関数]

階乗の概念を一般化した特殊関数である。ここで、x は任意の正実数である。

$$\Gamma(\ x\ )\ \equiv\ \int_0^\infty \mu^{x\ -\ 1}\,e^{\ -\ u}\,du$$

… (1.141)

基本的性質として、ガンマ関数は、自然数 n について、

$$\Gamma(\ n\ +\ 1\ )\ =\ n\,!$$

… (1.141)′

が成立する。また、非整数でのガンマ関数の値のうち、最も有名であるのは、ガウス積分になる $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$ である。

[ディクリレ分布]

○ 共役性をもつ多項分布

$$\mathrm{Dir}(\ \boldsymbol{\mu}\mid\boldsymbol{\alpha}\ ) = \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1)\cdots\Gamma(\alpha_M)}\prod_{k=1}^{M}\mu_k^{\alpha_k-1}$$

$$\left(= \underbrace{\frac{\Gamma(\ \alpha_1 + \alpha_2 + \cdots + \alpha_M\ )}{\Gamma(\alpha_1)\Gamma(\alpha_2)\cdots\Gamma(\alpha_M)}}_{\text{正規化係数}}\prod_{k=1}^{M}\mu_k^{\alpha_k-1}\right)$$

… (2.38)

ただし、$0 \le \mu_k \le 1$ かつ $\sum_k \mu_k = 1$ である。ここで、$\alpha_1, \cdots, \alpha_M$ は、この分布のパラメータで、$\boldsymbol{\alpha} = (\ \alpha_1, \cdots, \alpha_M\ )^{\mathrm{T}}$ を表す。また、α_0 は、

$$\alpha_0 = \sum_{k=1}^{K}\alpha_k$$

… (2.39)

である。

[ディクリレ分布の正規化条件]

$$\int_0^1\prod_{k=1}^{M}\mu_k^{\alpha_k-1}\,d\boldsymbol{\mu} = \frac{\Gamma(\alpha_1)\cdots\Gamma(\alpha_M)}{\Gamma(\alpha_0)}$$

… (2.38)′

[解]

まず、ディクリレ分布の平均 (2.273) を示す。式 (1.34) より、平均 $\mathrm{E}[\mu_j]$ は、

$$\mathrm{E}[\mu_j] = \int_0^1 \mu_j\,\mathrm{Dir}(\ \boldsymbol{\mu}\mid\boldsymbol{\alpha}\ )\,d\boldsymbol{\mu}$$

$$= \int_0^1 \mu_j\,\frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1)\cdots\Gamma(\alpha_M)}\prod_{j=1}^{M}\mu_j^{\alpha_j-1}\ d\boldsymbol{\mu}$$

$$= \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1)\cdots\Gamma(\alpha_M)}\int_0^1 \mu_j\prod_{k=1}^{M}\mu_k^{\alpha_k-1}\ d\boldsymbol{\mu}$$

… ※

となる。ここで、上記式中の $\int_0^1 \mu_j \prod_{k=1}^{M}\mu_k^{\alpha_k-1}\ d\boldsymbol{\mu}$ は、

$$\int_0^1 \mu_j\prod_{k=1}^{M}\mu_k^{\alpha_k-1}\ d\boldsymbol{\mu}$$

$$= \int_0^1 \mu_j\left(\ \mu_1^{\alpha_1-1}\cdot\mu_2^{\alpha_2-1}\cdot\cdots\cdot\mu_j^{\alpha_j-1}\cdot\cdots\cdot\mu_{M-1}^{\alpha_{M-1}-1}\cdot\mu_M^{\alpha_M-1}\ \right)d\boldsymbol{\mu}$$

$$= \int_0^1 \mu_1^{\alpha_1 - 1} \cdot \mu_2^{\alpha_2 - 1} \cdot \cdots \cdot \left(\mu_j^{\alpha_j - 1} \cdot \mu_j \right) \cdot \cdots \cdot \mu_{M-1}^{\alpha_{M-1} - 1} \cdot \mu_M^{\alpha_M - 1} \, d\boldsymbol{\mu}$$

$$= \int_0^1 \mu_1^{\alpha_1 - 1} \cdot \mu_2^{\alpha_2 - 1} \cdot \cdots \cdot \mu_j^{(\alpha_j + 1) - 1} \cdot \cdots \cdot \mu_{M-1}^{\alpha_{M-1} - 1} \cdot \mu_M^{\alpha_M - 1} \, d\boldsymbol{\mu}$$

と展開できるので、正規化条件 (2.38)′ を用いて、

$$= \frac{\Gamma(\alpha_1) \cdots \Gamma\left(\alpha_j + 1\right) \cdots \Gamma(\alpha_M)}{\Gamma(\alpha_0 + 1)}$$

と書き表せる。このことから、式 ※ は、

$$= \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1) \cdots \Gamma\left(\alpha_j\right) \cdots \Gamma(\alpha_M)} \cdot \frac{\Gamma(\alpha_1) \cdots \Gamma\left(\alpha_j + 1\right) \cdots \Gamma(\alpha_M)}{\Gamma(\alpha_0 + 1)}$$

と書き直せ、ガンマ関数の性質 (1.141)′ より、

$$= \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1) \cdots \Gamma\left(\alpha_j\right) \cdots \Gamma(\alpha_M)} \cdot \frac{\Gamma(\alpha_1) \cdots \alpha_j \Gamma\left(\alpha_j\right) \cdots \Gamma(\alpha_M)}{\alpha_0 \Gamma(\alpha_0)}$$

$$= \frac{\alpha_j}{\alpha_0}$$

となる。よって、ディクリレ分布の平均 (2.273) が示せた。

次に、ディクリレ分布の分散 (2.274) を示す。式 (1.39) より、分散 $\mathrm{var}\left[\mu_j\right]$ は、

$$\mathrm{var}\left[\mu_j\right] = \mathrm{E}\left[\mu_j^2\right] - \mathrm{E}\left[\mu_j\right]^2$$

… ※※

と書き表せる。ここで、上記の式の第一項 $\mathrm{E}\left[\mu_j^2\right]$ を求める。式 (1.34) より、$\mathrm{E}\left[\mu_j^2\right]$ は、

$$\mathrm{E}\left[\mu_j^2\right] = \int_0^1 \mu_j^2 \, \mathrm{Dir}(\boldsymbol{\mu} \mid \boldsymbol{\alpha}) \, d\boldsymbol{\mu}$$

$$= \int_0^1 \mu_j^2 \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1) \cdots \Gamma(\alpha_M)} \prod_{j=1}^{M} \mu_j^{\alpha_j - 1} \, d\boldsymbol{\mu}$$

$$= \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1) \cdots \Gamma(\alpha_M)} \int_0^1 \mu_j^2 \prod_{k=1}^{M} \mu_k^{\alpha_k - 1} \, d\boldsymbol{\mu}$$

… ※※※

となる。ここで、上記式中の $\int_0^1 \mu_j^2 \prod_{k=1}^{M} \mu_k^{\alpha_k - 1} \, d\boldsymbol{\mu}$ は、

$$\int_0^1 \mu_j^2 \prod_{k=1}^{M} \mu_k^{\alpha_k - 1} \, d\boldsymbol{\mu}$$

$$= \int_0^1 \mu_j^2 \left(\mu_1^{\alpha_1 - 1} \cdot \mu_2^{\alpha_2 - 1} \cdot \cdots \cdot \mu_j^{\alpha_j - 1} \cdot \cdots \cdot \mu_{M-1}^{\alpha_{M-1} - 1} \cdot \mu_M^{\alpha_M - 1} \right) d\boldsymbol{\mu}$$

$$= \int_0^1 \mu_1^{\alpha_1 - 1} \cdot \mu_2^{\alpha_2 - 1} \cdot \cdots \cdot \left(\mu_j^{\alpha_j - 1} \cdot \mu_j^2 \right) \cdot \cdots \cdot \mu_{M-1}^{\alpha_{M-1} - 1} \cdot \mu_M^{\alpha_M - 1} \, d\boldsymbol{\mu}$$

$$= \int_0^1 \mu_1^{\alpha_1 - 1} \cdot \mu_2^{\alpha_2 - 1} \cdot \cdots \cdot \mu_j^{(\alpha_j + 2) - 1} \cdot \cdots \cdot \mu_{M-1}^{\alpha_{M-1} - 1} \cdot \mu_M^{\alpha_M - 1} \, d\boldsymbol{\mu}$$

と展開できるので、正規化条件 (2.38)′ を用いて、

$$= \frac{\Gamma(\alpha_1) \cdots \Gamma(\alpha_j + 2) \cdots \Gamma(\alpha_M)}{\Gamma(\alpha_0 + 2)}$$

と書き表せる。このことから、式 ※※※ は、

$$= \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1) \cdots \Gamma(\alpha_j) \cdots \Gamma(\alpha_M)} \cdot \frac{\Gamma(\alpha_1) \cdots \Gamma(\alpha_j + 2) \cdots \Gamma(\alpha_M)}{\Gamma(\alpha_0 + 2)}$$

と書き直せ、ガンマ関数の性質 (1.141)′ より、

$$= \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1) \cdots \Gamma(\alpha_M)} \cdot \frac{\Gamma(\alpha_1) \cdots \alpha_j(\alpha_j + 1)\Gamma(\alpha_j) \cdots \Gamma(\alpha_M)}{\alpha_0(\alpha_0 + 1)\Gamma(\alpha_0)}$$

$$= \frac{\alpha_j(\alpha_j + 1)}{\alpha_0(\alpha_0 + 1)}$$

となる。よって、上記の結果と式 (2.273) を用いることにより、式 ※※ は、

$$= \frac{\alpha_j(\alpha_j + 1)}{\alpha_0(\alpha_0 + 1)} - \left(\frac{\alpha_j}{\alpha_0}\right)^2 = \frac{\alpha_j(\alpha_j + 1) - \alpha_j^2(\alpha_0 + 1)}{\alpha_0^2(\alpha_0 + 1)} = \frac{\alpha_j(\alpha_0 - \alpha_j)}{\alpha_0^2(\alpha_0 + 1)}$$

となるので、ディクリレ分布の分散 (2.274) が示せた。

最後に、ディクリレ分布の共分散 (2.275) を示す。式 (1.41) より、2つの確率変数 μ_j と μ_l の共分散 $\mathrm{cov}[\mu_j \mu_l]$ は、

$$\mathrm{cov}[\mu_j \mu_l] = \mathrm{E}[\mu_j \mu_l] - \mathrm{E}[\mu_j]\mathrm{E}[\mu_l]$$

… ※※※※

と書き表せる。ここで、上記の式の第一項 $\mathrm{E}[\mu_j \mu_l]$ を求める。式 (1.34) より、$\mathrm{E}[\mu_j \mu_l]$ は、

$$\mathrm{E}[\mu_j \mu_l] = \int_0^1 \mu_j \mu_l \, \mathrm{Dir}(\boldsymbol{\mu} \mid \boldsymbol{\alpha}) \, d\boldsymbol{\mu}$$

$$= \int_0^1 \mu_j \mu_l \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1) \cdots \Gamma(\alpha_M)} \prod_{j=1}^{M} \mu_j^{\alpha_j - 1} \, d\boldsymbol{\mu}$$

$$= \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1) \cdots \Gamma(\alpha_M)} \int_0^1 \mu_j \mu_l \prod_{k=1}^{M} \mu_k^{\alpha_k - 1} \, d\boldsymbol{\mu}$$

… ※※※※※

となる。ここで、上記式中の $\int_0^1 \mu_j \prod_{k=1}^{M} \mu_k^{\alpha_k - 1} \, d\boldsymbol{\mu}$ は、

$$\int_0^1 \mu_j \mu_l \prod_{k=1}^{M} \mu_k^{\alpha_k - 1} \, d\boldsymbol{\mu}$$

$$= \int_0^1 \mu_1^{\alpha_1 - 1} \cdot \mu_2^{\alpha_2 - 1} \cdot \cdots \cdot \mu_j^{(\alpha_j + 1) - 1} \cdot \cdots \cdot \mu_l^{(\alpha_l + 1) - 1} \cdot \cdots \cdot \mu_{M-1}^{\alpha_{M-1} - 1} \cdot \mu_M^{\alpha_M - 1} \, d\boldsymbol{\mu}$$

と展開できるので、正規化条件 (2.38)′ を用いて、

$$= \frac{\Gamma(\alpha_1) \cdots \Gamma(\alpha_j + 1) \cdots \Gamma(\alpha_l + 1) \cdots \Gamma(\alpha_M)}{\Gamma(\alpha_0 + 2)}$$

と書き表せる。このことから、式 ※※※※※ は、

$$= \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1) \cdots \Gamma(\alpha_j) \cdots \Gamma(\alpha_l) \cdots \Gamma(\alpha_M)} \cdot \frac{\Gamma(\alpha_1) \cdots \Gamma(\alpha_j + 1) \cdots \Gamma(\alpha_l + 1) \cdots \Gamma(\alpha_M)}{\Gamma(\alpha_0 + 2)}$$

と書き直せ、ガンマ関数の性質 (1.141)′ より、

$$= \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1) \cdots \Gamma(\alpha_j) \cdots \Gamma(\alpha_l) \cdots \Gamma(\alpha_M)} \cdot \frac{\Gamma(\alpha_1) \cdots \alpha_j \Gamma(\alpha_j) \cdots \alpha_l \Gamma(\alpha_l) \cdots \Gamma(\alpha_M)}{\alpha_0 (\alpha_0 + 1) \Gamma(\alpha_0)}$$

$$= \frac{\alpha_j \alpha_l}{\alpha_0 (\alpha_0 + 1)}$$

となる。よって、上記の結果と式 (2.273) を用いることにより、式 ※※※※ は、

$$= \frac{\alpha_j \alpha_l}{\alpha_0 (\alpha_0 + 1)} - \frac{\alpha_j}{\alpha_0} \cdot \frac{\alpha_l}{\alpha_0} = \frac{\alpha_0 \alpha_j \alpha_l - (\alpha_0 + 1) \alpha_j \alpha_l}{\alpha_0^2 (\alpha_0 + 1)} = - \frac{\alpha_j \alpha_l}{\alpha_0^2 (\alpha_0 + 1)}$$

となるので、ディクリレ分布の共分散 (2.275) が示せた。