**演習問題 2.11**

ディクリレ分布の下での $\ln \mu_j$ の期待値を、$\alpha_j$ についての導関数として表すと、

$$\mathrm{E}\left[\ \ln \mu_j\ \right] = \Psi\left( \alpha_j \right) - \Psi\left( \alpha_0 \right)$$

… ( 2.276 )

になることを示せ。ただし、$\alpha_0$ は、式 ( 2.39 ) で定義され、$\Psi(\cdot)$ は「 **ディガンマ関数** ( digamma function ) 」

$$\Psi( a ) \equiv \frac{d}{da} \ln \Gamma( a )$$

… ( 2.277 )

である。

**[ 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 )

**[ 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 )′

**[ 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、$\alpha_0$ 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 )′

**[ ディガンマ関数 ]**

ガンマ関数に対し、その対数微分をとった関数のこと

$$\Psi(x) \equiv \frac{d}{dx} \ln \Gamma(x) = \frac{\Gamma'(x)}{\Gamma(x)}$$

… ( 2.277 )′

[ ディガンマ関数の性質 ]

ディガンマ関数は、$x = 0, -1, -2, \cdots$ で漸近線をもち、以下の漸化式を満たす。

$$\Psi(x + 1) = \Psi(x) + \frac{1}{x}$$

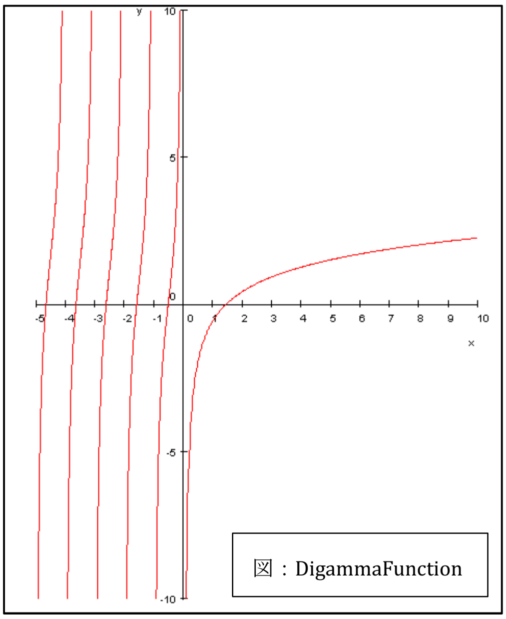
$$\Psi(x + n) = \Psi(x) + \sum_{k=1}^{n} \frac{1}{x + k - 1}$$

( ただし、$n$ は自然数である。)

また、$\Psi(1) = -\gamma$ であり、$\gamma = -0.5772\cdots$ は「 **オイラー定数** 」である。$x = 1$ のとき、

$$\Psi(1 + n) = -\gamma + \sum_{k=1}^{n} \frac{1}{k}$$

が成り立つ。



図：DigammaFunction

[ 解 ]

ディクリレ分布の下での $\ln \mu_j$ の期待値を、$\alpha_j$ についての導関数として表すと、式 ( 2.276 ) になることを示す。式 ( 1.34 ) より、 $\mathrm{E}[\ln \mu_j]$ は、

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

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

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

… ※

と書き表せる。ここで、

$$\frac{\partial}{\partial \alpha_j} \prod_{k=1}^{M} \mu_k^{\alpha_k - 1} = \frac{\partial}{\partial \alpha_j} \left( \mu_1^{\alpha_1 - 1} \cdot \mu_2^{\alpha_2 - 1} \cdot \cdots \cdot \mu_j^{\alpha_j - 1} \cdot \cdots \cdot \mu_{M-1}^{\alpha_{M-1} - 1} \cdot \mu_M^{\alpha_M - 1} \right)$$

$$= \mu_1^{\alpha_1 - 1} \cdot \mu_2^{\alpha_2 - 1} \cdot \cdots \cdot \mu_j^{\alpha_j - 1} \ln \mu_j \cdot \cdots \cdot \mu_{M-1}^{\alpha_{M-1} - 1} \cdot \mu_M^{\alpha_M - 1}$$

$$= ( \ln \mu_j ) \prod_{k=1}^{M} \mu_k^{\alpha_k - 1}$$

[ 指数関数の導関数 ]

$a > 0$，$a \neq 0$ において、

$$\frac{\partial a^x}{\partial x} = a^x \ln a$$

が成り立つ。特に、$a = e$ のとき、

$$\frac{\partial e^x}{\partial x} = e^x$$

となる。

であることから、式 ※ は、$\alpha_j$ についての導関数として、

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

と書き表すことができる。さらに、ディクリレ分布の正規化条件 ( 2.38 )′ より、上記の式は、

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

$$= \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1) \cdots \Gamma\left(\alpha_j\right) \cdots \Gamma(\alpha_M)} \cdot \frac{\partial}{\partial \alpha_j} [ \{ \Gamma(\alpha_1) \cdots \Gamma(\alpha_M) \} \{ \Gamma(\alpha_0) \}^{-1} ]$$

… ※※

と書き直せる。ここで、上記の式の微分項は、

$$\frac{\partial}{\partial \alpha_j} [ \{ \Gamma(\alpha_1) \cdots \Gamma(\alpha_M) \} \{ \Gamma(\alpha_0) \}^{-1} ]$$

$$= \{ \Gamma(\alpha_1) \cdots \Gamma(\alpha_M) \}' \{ \Gamma(\alpha_0) \}^{-1} + \{ \Gamma(\alpha_1) \cdots \Gamma(\alpha_M) \} [ \{ \Gamma(\alpha_0) \}^{-1} ]'$$

$$= \left\{ \Gamma(\alpha_1) \cdots \Gamma'\left(\alpha_j\right) \cdots \Gamma(\alpha_M) \right\} \{ \Gamma(\alpha_0) \}^{-1}$$

$$+ \{ \Gamma(\alpha_1) \cdots \Gamma(\alpha_M) \} [ - \Gamma(\alpha_0)^{-2} \{ \Gamma(\alpha_0) \}' ]$$

$$= \frac{\Gamma(\alpha_1) \cdots \Gamma'\left(\alpha_j\right) \cdots \Gamma(\alpha_M)}{\Gamma(\alpha_0)} - \{ \Gamma(\alpha_1) \cdots \Gamma(\alpha_M) \} \frac{\Gamma'(\alpha_0)}{\Gamma(\alpha_0)^2}$$

$$= \frac{\Gamma(\alpha_1) \cdots \Gamma'\left(\alpha_j\right) \cdots \Gamma(\alpha_M)}{\Gamma(\alpha_0)} - \frac{\{ \Gamma(\alpha_1) \cdots \Gamma(\alpha_M) \} \Gamma'(\alpha_0)}{\Gamma(\alpha_0)^2}$$

[ 積の微分 ]

2 つの関数 $f(x)$ と $g(x)$ があったとき、

$$( f(x) g(x) )' = f'(x) g(x) + f(x) g'(x)$$

が成り立つ。

と計算できるので、式 ※※ は、

$$= \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1)\cdots\Gamma(\alpha_M)} \cdot \left[ \frac{\Gamma(\alpha_1)\cdots\Gamma'\left(\alpha_j\right)\cdots\Gamma(\alpha_M)}{\Gamma(\alpha_0)} - \frac{\{\Gamma(\alpha_1)\cdots\Gamma(\alpha_M)\}\Gamma'(\alpha_0)}{\Gamma(\alpha_0)^2} \right]$$

$$= \frac{\Gamma'\left(\alpha_j\right)}{\Gamma\left(\alpha_j\right)} - \frac{\Gamma'(\alpha_0)}{\Gamma(\alpha_0)}$$

と整理でき、ディガンマ関数 ( 2.277 )′ より、上記の式は、

$$= \Psi\left(\alpha_j\right) - \Psi(\alpha_0)$$

と書き直せる。以上より、ディクリレ分布の下での $\ln \mu_j$ の期待値を、$\alpha_j$ についての導関数として表すと、式 ( 2.276 ) になることを示せた。