Άσκηση 21

Δείξτε ότι εάν $x_i \overset{ind}{\sim} Ga\left(\cdot|a_i, b\right)$ για $i=1,2$ τότε έχουμε ότι οι τ.μ $u = x_1 + x_2$ και $v = \dfrac{x_1}{x_1 + x_2}$ είναι ανεξάρτητες και $u \sim Ga\left(a_1 + a_2, b\right)$, $v \sim Be\left(a_1, a_2\right)$

Η από κοινού των x_1 και x_2 είναι

$$\pi\left(x_1, x_2\right) \propto x_1^{a_1-1}e^{-bx_1} \times x_2^{a_2-1}e^{-bx_2} = x_1^{a_1-1}x_2^{a_2-1}e^{-b\left(x_1+x_2\right)}.$$

Ορίζουμε τον ένα-προς-ένα μετασχηματισμό

$$T:\left\{\begin{array}{c} u = x_1 + x_2 \\ v = \dfrac{x_1}{x_1+x_2} \end{array}\right\} \Leftrightarrow T^{-1}\left\{: \begin{array}{l} x_1 = uv \\ x_2 = u\left(1-v\right) \end{array}\right\} \text{ με } \left\{\begin{array}{c} 0 < u = x_1 + x_2 < \infty \\ 0 < v = \dfrac{x_1}{x_1+x_2} < 1 \end{array}\right\}.$$

Η Ιακωβιανή του αντιστρόφου είναι $Jac\left(T^{-1}\right) = \begin{vmatrix} v & u \\ 1-v & -u \end{vmatrix} = u$,

Έτσι η από κοινού των u και v είναι

$$\pi\left(u,v\right) \propto \left(uv\right)^{a_1-1}\left(u\left(1-v\right)\right)^{a_2-1} e^{-bu} \,|\, Jac\left(T^{-1}\right)| = \left\{u^{a_1+a_2-1}e^{-bu}\right\} \times \left\{v^{a_1-1}\left(1-v\right)^{a_2-1}\right\}$$

$$\propto Ga\left(u|a_1 + a_2, b\right)Be\left(v|a_1, a_2\right)$$

Δηλαδή οι τ.μ. u και v είναι ανεξάρτητες

$$\pi\left(u,v\right) = \pi\left(u\right)\pi\left(v\right) = Ga\left(u|a_1 + a_2, b\right)Be\left(v|a_1, a_2\right)$$

Άσκηση 22

Δείξτε ότι εάν $x_i \overset{ind}{\sim} Ga\left(\cdot|a_i, b\right)$ για $i=1,2,3$ τότε

$$u = x_1 + x_2 + x_3 \sim Ga\left(\cdot|a_1 + a_2 + a_3, b\right),$$

$$\left.\begin{aligned} v_1 &= \frac{x_1}{x_1+x_2+x_3} \\ v_2 &= \frac{x_2}{x_1+x_2+x_3} \end{aligned}\right. \Rightarrow \left(v_1, v_2, \underbrace{1-v_1-v_2}_{v_3}\right) \sim D(\cdot \mid a_1, a_2, a_3)$$

και $u, (v_1, v_2)$ ανεξάρτητα.

$$T: \left\{\begin{aligned} u &= x_1+x_2+x_3 \\ v_1 &= \frac{x_1}{x_1+x_2+x_3} \\ v_2 &= \frac{x_2}{x_1+x_2+x_3} \end{aligned}\right\} \Leftrightarrow T^{-1}: \left\{\begin{aligned} x_1 &= uv_1 \\ x_2 &= uv_2 \\ x_3 &= u(1-v_1-v_2) \end{aligned}\right\} \text{ και } \left\{\begin{aligned} 0 < u = x_1+x_2+x_3 < \infty \\ 0 < v_1 = \frac{x_1}{x_1+x_2+x_3} < 1 \\ 0 < v_2 = \frac{x_2}{x_1+x_2+x_3} < 1 \end{aligned}\right\}$$

$$Jac\left(T^{-1}\right) = \begin{vmatrix} v_1 & u & 0 \\ v_2 & 0 & u \\ 1-v_1-v_2 & -u & -u \end{vmatrix} = u^2$$

$$\pi(x_1, x_2, x_3) \propto x_1^{a_1-1} e^{-bx_1} \times x_2^{a_2-1} e^{-bx_2} \times x_3^{a_3-1} e^{-bx_3} = x_1^{a_1-1} x_2^{a_2-1} x_3^{a_3-1} e^{-b(x_1+x_2+x_3)} \Rightarrow$$

$$\pi(u, v_1, v_2) \propto (uv_1)^{a_1-1} (uv_2)^{a_2-1} \left(u(1-v_1-v_2)\right)^{a_3-1} e^{-bu} \,\left|\left|Jac\left(T^{-1}\right)\right|\right|$$

$$= \left\{u^{a_1+a_2+a_3-1} e^{-bu}\right\} \times \left\{v_1^{a_1-1} v_2^{a_2-1} (1-v_1-v_2)^{a_3-1}\right\}$$

$$\propto Ga(u \mid a_1+a_2+a_3, b)\ D(v_1, v_2 \mid a_1, a_2, a_3)$$

Η κατανομή Dirichlet είναι το multivariate ανάλογο της beta κατανομής. Για $d=2$ έχουμε:

$$Dirichlet(v_1, v_2, v_3 \mid a_1, a_2, a_3) \propto v_1^{a_1-1} v_2^{a_2-1} v_3^{a_3-1}\ I(v \in S_3),$$

$$S_3 = \left\{v = (v_1, v_2, v_3) \in \mathbb{R}^3 : v_1+v_2+v_3 = 1, 0 < v_i < 1\right\}$$

$=$ the 3 dimensional probability simplex (πολύεδρο).

Η σταθερά κανονικοποίησης C για $d=2$ είναι

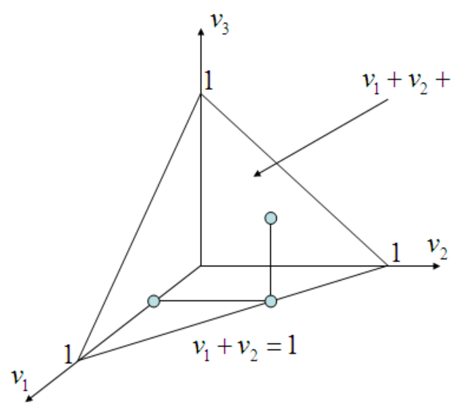
$$C^{-1} = \int_{S_3} v_1^{a_1-1} v_2^{a_2-1} v_3^{a_3-1} dv_1 dv_2 dv_3 = \int_{v_1=0}^{1} \int_{v_2=0}^{1-v_1} v_1^{a_1-1} v_2^{a_2-1} \left(1-v_1-v_2\right)^{a_3-1} dv_1 dv_2$$

Θέτοντας $v_2 = (1-v_1)t$ παίρνουμε

$$C^{-1} = \int_{v_1=0}^{1} \int_{t=0}^{1} v_1^{a_1-1} \left((1-v_1)t\right)^{a_2-1} \left(1-v_1-(1-v_1)t\right)^{a_3-1} dv_1 \left((1-v_1)dt\right) =$$

$$= \int_{v_1=0}^{1} v_1^{a_1-1} (1-v_1)^{a_2+a_3-1} dv_1 \int_{t=0}^{1} t^{a_2-1} (1-t)^{a_3-1} dt$$

$$= \left\{ \frac{\Gamma(a_1)\Gamma(a_2+a_3)}{\Gamma(a_1+a_2+a_3)} \right\} \left\{ \frac{\Gamma(a_2)\Gamma(a_3)}{\Gamma(a_2+a_3)} \right\} = \frac{\Gamma(a_1)\Gamma(a_2)\Gamma(a_3)}{\Gamma(a_1+a_2+a_3)}$$



Έτσι για $d=2$ παίρνουμε την non-minimal αναπαράσταση της Dirichlet κατανομής σαν μια singular (ιδιάζουσα) κατανομή με support το πολύεδρο πιθανότητας $S_3 \subset \mathbb{R}^3$

$$D(v_1, v_2, v_3 \mid a_1, a_2, a_3) = \frac{\Gamma(a_1+a_2+a_3)}{\Gamma(a_1)\Gamma(a_2)\Gamma(a_3)} v_1^{a_1-1} v_2^{a_2-1} v_3^{a_3-1} \, I(v \in S_3)$$

$$\propto v_1^{a_1-1} v_2^{a_2-1} v_3^{a_3-1} \, I(v \in S_3)$$

Η minimal αναπαράσταση της Dirichlet είναι:

$$D(v_1, v_2 \mid a_1, a_2, a_3) = \frac{\Gamma(a_1+a_2+a_3)}{\Gamma(a_1)\Gamma(a_2)\Gamma(a_3)} v_1^{a_1-1} v_2^{a_2-1} (1-v_1-v_2)^{a_3-1}$$

$$\times I\left(0 < v_1+v_2 < 1, 0 < v_1 < 1, 0 < v_2 < 1\right)$$

$$\propto v_1^{a_1-1} v_2^{a_2-1} \left(1-v_1-v_2\right)^{a_3-1} \times I\left(0<v_1+v_2<1, 0<v_1<1, 0<v_2<1\right)$$

Οι περιθώριες κατανομές της Dirichlet αρχίζοντας από την minimal αναπαράσταση είναι

$$\pi\left(v_1\right)=C\int_{v_2=0}^{1-v_1} v_1^{a_1-1} v_2^{a_2-1}\left(1-v_1-v_2\right)^{a_3-1} dv_2 = C\int_{t=0}^{1} v_1^{a_1-1}\left(t\left(1-v_1\right)\right)^{a_2-1}\left(1-v_1-t\left(1-v_1\right)\right)^{a_3-1}\left(1-v_1\right)dt$$

$$= C\, v_1^{a_1-1}\left(1-v_1\right)^{a_2+a_3-1}\int_{t=0}^{1} t^{a_2-1}\left(1-t\right)^{a_3-1} d\,t = C\, v_1^{a_1-1}\left(1-v_1\right)^{a_2+a_3-1}\frac{\Gamma\left(a_2\right)\Gamma\left(a_3\right)}{\Gamma\left(a_2+a_3\right)}$$

$$=\frac{\Gamma\left(a_1+a_2+a_3\right)}{\Gamma\left(a_1\right)\Gamma\left(a_2\right)\Gamma\left(a_3\right)}\frac{\Gamma\left(a_2\right)\Gamma\left(a_3\right)}{\Gamma\left(a_2+a_3\right)} v_1^{a_1-1}\left(1-v_1\right)^{a_2+a_3-1}$$

$$=\frac{\Gamma\left(a_1+\left(a_2+a_3\right)\right)}{\Gamma\left(a_1\right)\Gamma\left(a_2+a_3\right)} v_1^{a_1-1}\left(1-v_1\right)^{a_2+a_3-1} = Be\left(v_1 \mid a_1, a_2+a_3\right),$$

και λόγω συμμετρίας

$$\pi\left(v_2\right)=C\int_{v_1=0}^{1-v_2} v_1^{a_1-1} v_2^{a_2-1}\left(1-v_1-v_2\right)^{a_3-1} dv_1 = Be\left(v_2 \mid a_2, a_1+a_3\right)$$

Για $d=1$ η non-minimal αναπαράσταση της Dirichlet είναι

$$D\left(v_1, v_2 \mid a_1, a_2\right) = \frac{\Gamma\left(a_1\right)\Gamma\left(a_2\right)}{\Gamma\left(a_1+a_2\right)} v_1^{a_1-1} v_2^{a_2-1} I\left(\left(v_1, v_2\right)\in S_2\right)$$

και η minimal αναπαράσταση

$$D\left(v_1 \mid a_1, a_2\right) = Be\left(v_1 \mid a_1, a_2\right) = \frac{\Gamma\left(a_1+a_2\right)}{\Gamma\left(a_1\right)\Gamma\left(a_2\right)} v_1^{a_1-1}\left(1-v_1\right)^{a_2-1} I\left(0<v_1<1\right)$$

που είναι η Beta κατανομή.

Θα συμβολίζουμε την non-minimal αναπαράσταση της Dirichlet με

$$D\left(v_1,\ldots,v_n | a_1,\ldots,a_n\right) \propto v_1^{a_1-1}\cdots v_n^{a_n-1} I\left(v\in S_n\right),$$

και την minimal με

$$Dm\left(v_1,\ldots,v_{n-1} | a_1,\ldots,a_n\right) \propto v_1^{a_1-1}\cdots v_{n-1}^{a_{n-1}-1}\left(1-v_1-\cdots-v_n\right)^{a_n-1} I\left(v\in S_n\right)$$

Άσκηση 23

Εάν $v \sim D_3(\cdot | a)$ δείξτε ότι για $i = 1,2,3$ έχουμε $\mathbb{E}(v_i^k) = \frac{(a_i)_{(k)}}{(a_1 + a_2 + a_3)_{(k)}}$

$$E(v_1^k) = \int_{S_3} v_1^k D_3(v | a) dv_1 dv_2 dv_3 = B(a_1, a_2, a_3)^{-1} \int_{v_1=0}^{1} \int_{v_2=0}^{1-v_1} v_1^{k+a_1-1} v_2^{a_2-1} (1 - v_1 - v_2)^{a_3-1} dv_1 dv_2$$

Θέτοντας $v_2 = (1-v_1)t$ παίρνουμε

$$E(v_1^k) = B(a_1, a_2, a_3)^{-1} \int_{v_1=0}^{1} \int_{t=0}^{1} v_1^{k+a_1-1} ((1-v_1)t)^{a_2-1} (1 - v_1 - (1-v_1)t)^{a_3-1} dv_1 ((1-v_1) dt) =$$

$$= B(a_1, a_2, a_3)^{-1} \int_{t=0}^{1} t^{a_2-1} (1-t)^{a_3-1} dt \int_{v_1=0}^{1} v_1^{k+a_1-1} (1-v_1)^{a_2+a_3-1} dv_1$$

$$= \left(\frac{\Gamma(a_1 + a_2 + a_3)}{\Gamma(a_1)\Gamma(a_2)\Gamma(a_3)} \right) \left(\frac{\Gamma(a_2)\Gamma(a_3)}{\Gamma(a_2 + a_3)} \right) \left(\frac{\Gamma(k + a_1)\Gamma(a_2 + a_3)}{\Gamma(k + a_1 + a_2 + a_3)} \right)$$

$$= \left(\frac{\Gamma(a_1 + a_2 + a_3)}{\Gamma(a_1)} \right) \left(\frac{(k + a_1 - 1)\cdots(a_1)\Gamma(a_1)}{(k + a_1 + a_2 + a_3 - 1)\cdots(a_1 + a_2 + a_3)\Gamma(a_1 + a_2 + a_3)} \right)$$

$$= \frac{(a_1)_{(k)}}{(a_1 + a_2 + a_3)_{(k)}}$$

Λόγω συμμετρίας παίρνουμε το ζητούμενο αποτέλεσμα και για $i = 2,3$.

Άσκηση 24

Δείξτε ότι εάν

$x_i \overset{ind}{\sim} Ga(\cdot | a_i, b)$ για $i = 1, \ldots, d+1$

τότε $u = \sum_{i=1}^{d+1} x_i \sim Ga\left(\cdot | \sum_{i=1}^{d+1} a_i, b \right)$,

$$v_i = \frac{x_i}{\sum_{i=1}^{d+1} x_i},\ 1 \le i \le d \quad \Rightarrow \quad \left(v_1, \ldots, v_d, \underbrace{1 - v_1 - \ldots - v_d}_{v_{d+1}} \right) \sim Dirichlet(\cdot | a_1, \ldots, a_{d+1})$$

και $u, (v_1, \ldots, v_d)$ ανεξάρτητα.

$$T: \begin{array}{l} u=\sum_{i=1}^{d+1} x_i \\ v_i=\dfrac{x_1}{\sum_{i=1}^{d+1} x_i} \end{array} \Leftrightarrow \quad T^{-1}: \begin{array}{l} x_i=uv_i,\ 1\le i\le d \\ x_{d+1}=u\left(1-\sum_{i=1}^{d} x_i\right) \end{array} \text{ και } \begin{array}{l} 0<u=\sum_{i=1}^{d+1} x_i<\infty \\ 0<v_i=\dfrac{x_i}{\sum_{j=1}^{d+1} x_j}<1,\ 1\le i\le d \end{array}$$

$$Jac\left(T^{-1}\right)=\begin{vmatrix} v_1 & u & \cdots & 0 \\ v_2 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 1-v_1-\cdots-v_d & -u & \cdots & -u \end{vmatrix}=(-1)^{d+1}u^d$$

$$\pi\left(x_1,\ldots,x_{d+1}\right) \propto \prod_{i=1}^{d+1} x_i^{a_i-1}e^{-bx_i} = \left\{\prod_{i=1}^{d+1} x_i^{a_i-1}\right\}\exp\left(-b\sum_{i=1}^{d+1} x_i\right) \Rightarrow$$

$$\pi\left(u,v_1,\ldots,v_d\right) \propto \left\{\prod_{i=1}^{d}\left(uv_i\right)^{a_i-1}\right\}\left(u\left(1-v_1-\cdots-v_d\right)\right)^{a_{d+1}-1} e^{-bu}\left|Jac\left(T^{-1}\right)\right|$$

$$=\left\{u^{\sum_{i=1}^{d+1} a_i-1} e^{-bu}\right\}\left\{\left(\prod_{i=1}^{d} v_i^{a_i-1}\right)\left(\underbrace{1-v_1-\cdots-v_d}_{v_{d+1}}\right)^{a_{d+1}-1}\right\}$$

$$\propto Ga\left(u\,|\sum_{i=1}^{d+1} a_i, b\right)\ Dirichlet\left(v_1,\ldots,v_d,v_{d+1}\,|\,a_1,\ldots,a_d,a_{d+1}\right)$$

Έχουμε δηλαδή ότι

$$Dirichlet\left(v_1,\ldots,v_d\,|\,a_1,\ldots,a_d,a_{d+1}\right)=C\,v_1^{a_1-1}\cdots v_d^{a_d-1}\left(1-v_1-\cdots-v_d\right)^{a_{d+1}-1}$$
$$\times I\left(0<v_1+\cdots+v_d<1, 0<v_i<1\right).$$

Δείχνουμε τώρα με επαγωγή ότι το normalizing constant C για γενικό d είναι

$$C^{-1}=\int_{S_{d+1}} \prod_{i=1}^{d+1} v_i^{a_i-1}dv=\frac{\prod_{i=1}^{d+1}\Gamma\left(a_i\right)}{\Gamma\left(\sum_{i=1}^{d+1} a_i\right)}\equiv B\left(a_1,\ldots,a_{d+1}\right),$$

με $S_{d+1}=\left\{v=\left(v_1,\ldots,v_{d+1}\right)\in\mathbb{R}^{d+1}: v_1+\cdots+v_{d+1}=1, 0<v_i<1\right\}$ το $(d+1)$-probability simplex.

Στην άσκηση 22 το δείξαμε για $d=2$. Έστω ότι η προς απόδειξη σχέση ισχύει για $d-1$, δηλαδή ότι $\int_{S_d}\prod_{i=1}^{d} v_i^{a_i-1} dv_1\cdots dv_d = B(a_1,\dots,a_d)$, τότε

$$\int_{S_{d+1}}\prod_{i=1}^{d+1} v_i^{a_i-1} dv = \int_{v_1=0}^{1}\int_{v_2=0}^{1-v_1}\cdots\int_{v_{d-1}=0}^{1-v_1-\cdots-v_{d-2}}\int_{v_d=0}^{1-v_1-\cdots-v_{d-1}}\left\{\prod_{i=1}^{d} v_i^{a_i-1}\right\}(1-v_1-\cdots-v_{d-1}-v_d)^{a_{d+1}-1} dv_1\cdots dv_d$$

θέτοντας $v_d=(1-v_1-\cdots-v_{d-1})t$ το ολοκλήρωμα γίνεται

$$\int_{v_1=0}^{1}\int_{v_2=0}^{1-v_1}\cdots\int_{v_{d-1}=0}^{1-v_1-\cdots-v_{d-2}}\left\{\prod_{i=1}^{d-1} v_i^{a_i-1}\right\}(1-v_1-\cdots-v_{d-1})^{a_d+a_{d+1}-1} dv_1\cdots dv_{d-1}\int_{t=0}^{1} t^{a_d-1}(1-t)^{a_{d+1}-1}dt$$

$$=B(a_1,\dots,a_{d-1},a_d+a_{d+1})B(a_d,a_{d+1})=\frac{\Gamma(a_1)\cdots\Gamma(a_{d-1})\Gamma(a_d+a_{d+1})}{\Gamma(a_1+\ldots+a_{d+1})}\frac{\Gamma(a_d)\Gamma(a_{d+1})}{\Gamma(a_d+a_{d+1})}$$

$$=\frac{\Gamma(a_1)\cdots\Gamma(a_{d+1})}{\Gamma(a_1+\ldots+a_{d+1})}=B(a_1,\dots,a_{d+1})$$

Άσκηση 25

Δίνεται το μοντέλο $x_i|a_i,\beta \overset{ind}{\sim} Po(\cdot|a_i\beta)$ για $i=1,\dots n$. Να γίνει εκτίμηση του $\vartheta=(a,\beta)$, όπου $a=(a_1,\dots,a_n)$ και $\sum_{i=1}^{n} a_i=1$, $0<a_i<1$, $i=1,\dots n$. Ποία τα posterior means των a_i, $i=1,\dots n$ και β?

Η συνάρτηση πιθανοφάνειας είναι

$$\pi(x|a,\beta)=\prod_{i=1}^{n}\pi(x_i|a_i,\beta)=\prod_{i=1}^{n}Po(x_i|a_i\beta)=\prod_{i=1}^{n}\frac{e^{-a_i\beta}(a_i\beta)^{x_i}}{x_i!}$$

$$\overset{a,\beta}{\propto}\prod_{i=1}^{n}\left\{e^{-a_i\beta}(a_i\beta)^{x_i}\right\}=e^{-\beta}\beta^{n\bar{x}}\prod_{i=1}^{n}a_i^{x_i}$$

θέτοντας a-priori

$$\pi(a,\beta)=\pi(a)\pi(\beta)=Dirichlet_n(a|q)Ga(\beta|\gamma,\delta)\propto\left\{\prod_{i=1}^{n}a_i^{q_i-1}\right\}\left\{\beta^{\gamma-1}e^{-\delta\beta}\right\},$$

όπου $q=(q_1,\dots,q_n)$, γ,δ hyperparameters, θα έχουμε την posterior

$$\pi\left(a,\beta|x\right)\propto\pi\left(a\right)\pi\left(\beta\right)\pi\left(x|a,\beta\right)\propto\left\{\prod_{i=1}^{n}a_i^{q_i-1}\right\}\times\left\{\beta^{\gamma-1}e^{-\delta\beta}\right\}\times\left\{e^{-\beta}\beta^{n\overline{x}}\prod_{i=1}^{n}a_i^{x_i}\right\}$$

$$=e^{-(1+\delta)\beta}\beta^{\gamma+n\overline{x}-1}\prod_{i=1}^{n}a_i^{x_i+q_i-1}\propto Dirichlet_n\left(a|q+x\right)Ga\left(\beta|\gamma+n\overline{x},\delta+1\right)$$

Δηλαδή $\pi\left(a,\beta|x\right)=\pi\left(a|x\right)\pi\left(\beta|x\right)$ με

$$\pi\left(a|x\right)=Dirichlet_n\left(a|q+x\right)$$

$$\pi\left(\beta|x\right)=Ga\left(\beta|\gamma+n\overline{x},\delta+1\right).$$

Επειδή $a_i|x\sim Be\left(\cdot|q_i+x_i,\sum_{j\neq i}\left(q_j+x_j\right)\right)$ παίρνουμε

$$E\left(a_i|x\right)=\frac{q_i+x_i}{\left(q_i+x_i\right)+\sum_{j\neq i}\left(q_j+x_j\right)}=\frac{q_i+x_i}{\sum_{j=1}^{n}\left(q_j+x_j\right)},\ i=1,\dots n$$

$$E\left(\beta|x\right)=\frac{\gamma+n\overline{x}}{\delta+1}$$

<u>Παρατηρήστε ότι</u> το scale parameter της Gamma κάνει update από δ σε $\delta+1$ που σημαίνει ότι δεν εξαρτάται από τις παρατηρήσεις. Δηλαδή θα μπορούσαμε εξ' αρχής να θέσουμε, ας πούμε, $\delta=1$.

Άσκηση 26

Να δειχθεί ότι το conjugate prior της Multinomial (πολυωνυμικής) κατανομής είναι η κατανομή Dirichlet

Ας θεωρήσουμε μια πολυωνυμική παρατήρηση x τέτοια ώστε $x=\left(x_1,\dots,x_m\right)$ με $\sum_{j=1}^{m}x_j=M$, όπου x_j ο αριθμός των αποτελεσμάτων τύπου j για $1\leq j\leq m$ σε M πολυωνυμικές δοκιμές. Τότε $x|p\sim Multinomial_m\left(\cdot|M,p\right)$ για διάνυσμα πιθανότητας $\vartheta=p=\left(p_1,\dots,p_m\right)$ και M γνωστό, δηλαδή

$$\pi\left(x|p\right)=Multinomial\left(x_1,\dots,x_m|M,\left(p_1,\dots,p_m\right)\right)$$

$$=\binom{M}{x_1,\dots,x_m}p_1^{x_1}\cdots p_m^{x_m}\propto p_1^{x_1}\cdots p_m^{x_m}$$

όπου $\binom{M}{x_1,\ldots,x_m} \equiv \frac{M!}{\prod_{j=1}^{m} x_j!}$ ο πολυωνυμικός συντελεστής για τον οποίο ξέρουμε

ότι ισχύει $\sum_{x_1+\ldots+x_m=M} \binom{M}{x_1,\ldots,x_m} y_1^{x_1} \cdots y_m^{x_m} = \left(y_1+\ldots+y_m\right)^M$. Σημειώστε ότι η

παραπάνω έκφραση για την multinomial είναι ιδιάζουσα.

Θέτοντας σαν prior

$$\pi(p) = Dirichlet_m(p|a) \propto \prod_{j=1}^{m} p_i^{a_i-1},$$

η posterior γίνεται

$$\pi(p|x) \propto \left\{\prod_{j=1}^{m} p_j^{a_j-1}\right\}\left\{\prod_{j=1}^{m} p_j^{x_j}\right\} = \prod_{j=1}^{m} p_j^{a_j+x_j-1} \propto Dirichlet_m(p|a+x)$$

Άσκηση 27

Δείξτε ότι στην προηγούμενη άσκηση, το posterior mean $E(p_i|x)$ είναι κυρτός γραμμικός συνδυασμός του prior mean $E(p_i)$ και maximum likelihood estimator $(p_i)_{MLE}$ για κάθε $1 \le i \le m$

Γνωρίζουμε ότι

$$E(p_i) = \frac{a_i}{\sum_{j=1}^{m} a_j} \text{ και } E(p_i|x) = \frac{a_i + x_i}{\sum_{j=1}^{m}\left(a_j + x_j\right)} = \frac{a_i + x_i}{\sum_{j=1}^{m} a_j + M} \text{ για } 1 \le i \le m.$$

Το $(p_i)_{MLE}$ μπορούμε να το υπολογίσουμε από τις σχέσεις

$$\frac{\partial}{\partial p_j}\log\left\{\pi(x|p)\right\} = 0,\ 1 \le j \le m-1,$$

$$x_m = M - \sum_{r=1}^{m-1} x_r,\ p_m = 1 - \sum_{r=1}^{m-1} p_r \iff$$

$$\frac{\partial}{\partial p_j}\left\{x_1 \log(p_1) + \cdots + x_{m-1}\log(p_{m-1}) + \left(M - \sum_{r=1}^{m-1} x_r\right)\log\left(1 - \sum_{r=1}^{m-1} p_r\right)\right\} = 0 \iff$$

$$\frac{x_j}{p_j} = \frac{M - \sum_{r=1}^{m-1} x_r}{1 - \sum_{r=1}^{m-1} p_r}\,,\ 1 \le j \le m-1 \ \Leftrightarrow\ \left(p_j\right)_{MLE} = \frac{x_j}{M}\,,\ 1 \le j \le m$$

Ζητάμε $0 \le \gamma_i \le 1$ τέτοιο ώστε

$$E\left(p_i | x\right) = \left(1-\gamma_i\right) E\left(p_i\right) + \gamma_i \left(p_i\right)_{MLE} \ \Leftrightarrow\ \gamma_i = \frac{M}{\sum_{j=1}^{m} a_j + M} \text{ για κάθε } 1 \le i \le m$$

Άσκηση 28

Χρησιμοποιώντας τον ορισμό της συνάρτησης gamma $\Gamma\left(a\right) = \int_{\mathbb{R}^+ \ni u} u^{a-1} e^{-u} du$ δείξτε

ότι beta ολοκλήρωμα $B\left(p,q\right) = \int_{x=0}^{1} x^{p-1} \left(1-x\right)^{q-1} dx$, για $p > 0$ και $q > 0$, έχει την

αναπαράσταση $B\left(p,q\right) = \dfrac{\Gamma\left(p\right)\Gamma\left(q\right)}{\Gamma\left(p+q\right)}$.

Λύση

$$\Gamma\left(p\right)\Gamma\left(q\right) = \int_{\mathbb{R}^+ \ni u} u^{p-1} e^{-u} du \int_{\mathbb{R}^+ \ni v} v^{q-1} e^{-v} dv$$

θέτοντας $u = x^2$ και $v = y^2$ παίρνουμε

$$\Gamma\left(p\right)\Gamma\left(q\right) = 4 \int_{\mathbb{R}^+ \ni x} x^{2p-1} e^{-x^2} dx \int_{\mathbb{R}^+ \ni y} y^{2q-1} e^{-y^2} dy = 4 \int_{\mathbb{R}^+ \ni x} \int_{\mathbb{R}^+ \ni y} e^{-\left(x^2+y^2\right)} x^{2p-1} y^{2q-1} dx dy$$

θέτοντας $x = r\cos\left(\vartheta\right)$ και $y = r\sin\left(\vartheta\right)$ παίρνουμε

$$\Gamma\left(p\right)\Gamma\left(q\right) = 4 \int_{\mathbb{R}^+ \ni r} \int_{(0,\pi/2) \ni \vartheta} e^{-r^2} r^{2p-1} \cos\left(\vartheta\right)^{2p-1} r^{2q-1} \sin\left(\vartheta\right)^{2q-1} \left|Jac\left(T^{-1}\right)\right| dr d\vartheta$$

$$= 4 \int_{\mathbb{R}^+ \ni r} \int_{(0,\pi/2) \ni \vartheta} e^{-r^2} r^{2(p+q)-1} \cos\left(\vartheta\right)^{2p-1} \sin\left(\vartheta\right)^{2q-1} dr d\vartheta$$

$$= 4 \int_{\mathbb{R}^+ \ni r} e^{-r^2} r^{2(p+q)-1} dr \int_{(0,\pi/2) \ni \vartheta} \cos\left(\vartheta\right)^{2p-1} \sin\left(\vartheta\right)^{2q-1} d\vartheta$$

Θέτοντας $r = x^{1/2}$ παίρνουμε

$$\int_{\mathbb{R}^+ \ni r} e^{-r^2} r^{2(p+q)-1} dr = \frac{1}{2} \int_{\mathbb{R}^+ \ni x} e^{-x} r^{p+q-1} dr = \frac{1}{2} \Gamma(p+q)$$

Θέτοντας $\cos(\vartheta) = \sqrt{x}$ παίρνουμε $\sin(\vartheta) = \sqrt{1-x}$ και $d\vartheta = \frac{-dx}{2\sqrt{x(1-x)}}$

$$\int_{(0,\pi/2) \ni \vartheta} \cos(\vartheta)^{2p-1} \sin(\vartheta)^{2q-1} \, d\vartheta = \frac{1}{2} \int_{(0,1) \ni x} x^{p-1} (1-x)^{q-1} \, dx$$

Έτσι η εξίσωση για το $\Gamma(p)\Gamma(q)$ γίνεται

$$\Gamma(p)\Gamma(q) = 4 \left\{ \frac{1}{2} \Gamma(p+q) \right\} \left\{ \frac{1}{2} \int_{(0,1) \ni x} x^{p-1} (1-x)^{q-1} \, dx \right\}$$