Ορίζουμε την πληροφορία κατά Fisher $I_x(\vartheta)$ σαν το ποσό της πληροφορίας που περιέχει η παρατήρηση x για την παράμετρο ϑ. Συμβολίζοντας με $S(x|\vartheta)$ την λογαριθμική παράγωγο της πιθανοφάνειας ως προς την παράμετρο ϑ (score function), η πληροφορία κατά Fisher για μια μονοδιάστατη παράμετρο ϑ ορίζεται ως

$$I_x(\vartheta) = \mathbb{E}\left[S(x|\vartheta)^2 \mid \vartheta \right],$$

ισοδύναμα

$$I_x(\vartheta) = \int_X S(x|\vartheta)^2 \, \Pi(dx|\vartheta) = \int_X S(x|\vartheta)^2 \, \pi(x|\vartheta) dx \geq 0 .$$

Η πληροφορία κατά Fisher είναι το variance του score function $S(x|\vartheta)$.
Πράγματι

$$Var\left[S(x|\vartheta) \mid \vartheta \right] = \mathbb{E}\left[S(x|\vartheta)^2 \mid \vartheta \right] - \mathbb{E}\left[S(x|\vartheta) \mid \vartheta \right]^2 = I_x(\vartheta) - \mathbb{E}\left[S(x|\vartheta) \mid \vartheta \right]^2$$

ενώ

$$\mathbb{E}\left[S(x|\vartheta) \mid \vartheta \right] = \int_X S(x|\vartheta) \pi(x|\vartheta) dx = \int_X \frac{\partial}{\partial \vartheta} \pi(x|\vartheta) dx = \frac{\partial}{\partial \vartheta} \int_X \pi(x|\vartheta) dx = 0 .$$

Τώρα μπορούμε να δείξουμε ότι **η πληροφορία κατά Fisher είναι αθροιστική ως προς της παρατηρήσεις**. Δηλαδή εάν έχουμε n παρατηρήσεις, $x_i \mid \vartheta \overset{iid}{\sim} \pi(\cdot \mid \vartheta)$ για $1 \leq i \leq n$,θα έχουμε ότι

$$I_x(\vartheta) = \sum_{i=1}^{n} I_{x_i}(\vartheta) .$$

Πράγματι

$$I_x(\vartheta) = \mathbb{E}\left[S(x|\vartheta)^2 \mid \vartheta \right] = \mathbb{E}\left\{ \left[\frac{\partial}{\partial \vartheta} \log \prod\nolimits_{i=1}^{n} \pi(x_i|\vartheta) \right]^2 \mid \vartheta \right\}$$

$$= \mathbb{E}\left\{ \left[\sum\nolimits_{i=1}^{n} \frac{\partial}{\partial \vartheta} \log \pi(x_i|\vartheta) \right]^2 \mid \vartheta \right\}$$

$$= \mathbb{E}\left\{ \sum\nolimits_{i=1}^{n} \left[\frac{\partial}{\partial \vartheta} \log \pi(x_i|\vartheta) \right]^2 + 2 \sum_{1 \leq i < j \leq n} \frac{\partial}{\partial \vartheta} \log \pi(x_i|\vartheta) \frac{\partial}{\partial \vartheta} \log \pi(x_j|\vartheta) \mid \vartheta \right\}$$

$$= \sum\nolimits_{i=1}^{n} \mathbb{E}\left\{ \left[\frac{\partial}{\partial \vartheta} \log \pi(x_i|\vartheta) \right]^2 \mid \vartheta \right\} + 2 \sum_{1 \leq i < j \leq n} \mathbb{E}\left\{ \frac{\partial}{\partial \vartheta} \log \pi(x_i|\vartheta) \mid \vartheta \right\} \mathbb{E}\left\{ \frac{\partial}{\partial \vartheta} \log \pi(x_j|\vartheta) \mid \vartheta \right\}$$

$$= \sum_{i=1}^{n} \mathbb{E}\left\{ \left[\frac{\partial}{\partial \vartheta} \log \pi\left(x_i | \vartheta\right) \right]^2 | \vartheta \right\} = \sum_{i=1}^{n} \mathbb{E}\left[S\left(x_i | \vartheta\right)^2 | \vartheta \right] = \sum_{i=1}^{n} I_{x_i}\left(\vartheta\right).$$

<u>Πρόταση</u>: Η πληροφορία κατά Fisher (για μία παρατήρηση x) δίνεται ισοδύναμα από την σχέση:

$$I_x\left(\vartheta\right) = -\mathbb{E}\left[\frac{\partial^2}{\partial \vartheta^2} \log \pi\left(x | \vartheta\right) | \vartheta \right].$$

Παραγωγίζοντας την $\mathbb{E}\left[S\left(x|\vartheta\right) | \vartheta \right] = 0$ έχουμε:

$$0 = \frac{\partial}{\partial \vartheta} \mathbb{E}\left[\frac{\partial}{\partial \vartheta} \log \pi\left(x|\vartheta\right) | \vartheta \right] = \int_X \frac{\partial}{\partial \vartheta} \left\{ \left(\frac{\partial}{\partial \vartheta} \log \pi\left(x|\vartheta\right) \right) \pi\left(x|\vartheta\right) \right\} dx$$

$$= \int_X \left\{ \frac{\partial^2}{\partial \vartheta^2} \log \pi\left(x|\vartheta\right) \right\} \pi\left(x|\vartheta\right) dx + \int_X \frac{\partial}{\partial \vartheta} \log \pi\left(x|\vartheta\right) \frac{\partial}{\partial \vartheta} \pi\left(x|\vartheta\right) dx$$

$$= \int_X \left\{ \frac{\partial^2}{\partial \vartheta^2} \log \pi\left(x|\vartheta\right) \right\} \pi\left(x|\vartheta\right) dx + \int_X \frac{\partial}{\partial \vartheta} \log \pi\left(x|\vartheta\right) \left\{ \frac{\frac{\partial}{\partial \vartheta} \pi\left(x|\vartheta\right)}{\pi\left(x|\vartheta\right)} \right\} \pi\left(x|\vartheta\right) dx$$

$$= \int_X \left(\frac{\partial^2}{\partial \vartheta^2} \log \pi\left(x|\vartheta\right) \right) \pi\left(x|\vartheta\right) dx + \int_X \left(\frac{\partial}{\partial \vartheta} \log \pi\left(x|\vartheta\right) \right)^2 \pi\left(x|\vartheta\right) dx$$

$$= E\left[\frac{\partial^2}{\partial \vartheta^2} \log \pi\left(x|\vartheta\right) \right] + E\left[\left(\frac{\partial}{\partial \vartheta} \log \pi\left(x|\vartheta\right) \right)^2 \right].$$

Παρατήρηση: Η ισότητα

$$I_x\left(\vartheta\right) = E\left[\left(\frac{\partial}{\partial \vartheta} \log \pi\left(x|\vartheta\right) \right)^2 \right] = \mathbb{E}\left[-\frac{\partial^2}{\partial \vartheta^2} \log \pi\left(x|\vartheta\right) \right]$$

ισχύει για τις περιπτώσεις όπου

$$\mathbb{E}\left[S\left(x|\vartheta\right) | \vartheta \right] = \int_X \frac{\partial}{\partial \vartheta} \pi\left(x|\vartheta\right) dx = \frac{\partial}{\partial \vartheta} \int_X \pi\left(x|\vartheta\right) dx = 0,$$

δηλαδή **όταν ισχύει** $\int_X \frac{\partial}{\partial\vartheta}\pi(x|\vartheta)dx = \frac{\partial}{\partial\vartheta}\int_X \pi(x|\vartheta)dx$. Για παράδειγμα εάν το στήριγμα X της πυκνότητας του μοντέλου $\pi(x|\vartheta)$ εξαρτάται από την παράμετρο ϑ, τότε γενικά έχουμε ότι $\int_{X(\vartheta)} \frac{\partial}{\partial\vartheta}\pi(x|\vartheta)dx \neq \frac{\partial}{\partial\vartheta}\int_{X(\vartheta)} \pi(x|\vartheta)dx$ με αποτέλεσμα

$$I_x(\vartheta) = E\left[\left(\frac{\partial}{\partial\vartheta}\log\pi(x|\vartheta)\right)^2\right] \neq -\mathbb{E}\left[\frac{\partial^2}{\partial\vartheta^2}\log\pi(x|\vartheta)\right].$$

Για παράδειγμα εάν $[x_i|\vartheta] \overset{iid}{\sim} \mathcal{U}(\cdot|0,\vartheta),\ i=1,\ldots,n$ τότε $X = X(\vartheta) = (0,\vartheta)$ και για μία παρατήρηση x έχουμε

$$\int_{X(\vartheta)} \frac{\partial}{\partial\vartheta}\pi(x|\vartheta)dx = \int_{x=0}^{\vartheta} \frac{\partial}{\partial\vartheta}\vartheta^{-1}dx = -\vartheta^{-1} \neq \frac{\partial}{\partial\vartheta}\int_{X(\vartheta)} \pi(x|\vartheta)dx = \frac{\partial}{\partial\vartheta}\int_{x=0}^{\vartheta} \frac{1}{\vartheta}dx = 0.$$

Όμως η πληροφορία κατά Fisher για μία παρατήρηση x_i είναι:

$$I_{x_i}(\vartheta) = \mathbb{E}\left[S(x_i|\vartheta)^2 \,|\, \vartheta\right] = \mathbb{E}\left[\left(-\vartheta^{-1}\right)^2 \,|\, \vartheta\right] = \vartheta^{-2} \Rightarrow I_x(\vartheta) = n\vartheta^{-2}.$$

<u>Άσκηση</u>: Εάν $\pi(\cdot|\vartheta) \in EF$ τότε για παρατήρηση x τέτοια ώστε $x\,|\,\vartheta \sim \pi(\cdot|\vartheta)$, ισχύει $\mathbb{E}\left[t_{suff}(x)\,|\,\vartheta\right] = -\frac{g'(\vartheta)}{g(\vartheta)c'(\vartheta)}$. Επαληθεύστε την προηγούμενη σχέση όταν $x\,|\,\vartheta \sim Ga(a,\vartheta)$.

Έχουμε ότι

$$\pi(\cdot|\vartheta) \in EF \Leftrightarrow \pi(x|\vartheta) = h(x)g(\vartheta)e^{c(\vartheta)t(x)},\ g(\vartheta) = \left\{\int_X h(u)e^{c(\vartheta)t(u)}du\right\}^{-1},\ X \neq X(\vartheta),$$

που δίνει $\int_X \frac{\partial}{\partial\vartheta}\pi(x|\vartheta)dx = \frac{\partial}{\partial\vartheta}\int_X \pi(x|\vartheta)dx = 0$, τότε για $t = t_{suff}$ έχουμε

$$\begin{aligned}
0 &= \int_X \frac{\partial}{\partial\vartheta}\pi(x|\vartheta)dx = \int_X \frac{\partial}{\partial\vartheta}\left\{h(x)g(\vartheta)e^{c(\vartheta)t(x)}\right\}dx \\
&= g'(\vartheta)\int_X h(x)e^{c(\vartheta)t(x)}dx + g(\vartheta)c'(\vartheta)\int_X h(x)t(x)e^{c(\vartheta)t(x)}dx \\
&= \frac{g'(\vartheta)}{g(\vartheta)}\int_X h(x)g(\vartheta)e^{c(\vartheta)t(x)}dx + c'(\vartheta)\int_X t(x)h(x)g(\vartheta)e^{c(\vartheta)t(x)}dx
\end{aligned}$$

$$=\frac{g'(\vartheta)}{g(\vartheta)}\int_X \pi(x|\vartheta)dx + c'(\vartheta)\int_X t(x)\pi(x|\vartheta)dx$$

$=\frac{g'(\vartheta)}{g(\vartheta)} + c'(\vartheta)\mathbb{E}\left[t(x)\,|\,\vartheta\right]$, από όπου, $\mathbb{E}\left[t(x)\,|\,\vartheta\right] = -\frac{g'(\vartheta)}{g(\vartheta)c'(\vartheta)}$.

Έχουμε ότι $Ga(a,\vartheta)\in EF$, εφόσον $\pi(x|\vartheta)=\frac{\vartheta^a}{\Gamma(a)}x^{a-1}e^{-\vartheta x} = \left(\frac{x^{a-1}}{\Gamma(a)}\right)\left(\vartheta^a\right)\left(e^{-\vartheta x}\right)$

δηλαδή $g(\vartheta)=\vartheta^a$ και $c(\vartheta)=-\vartheta$, και $t(x)=x$, που δίνουν $\frac{g'(\vartheta)}{g(\vartheta)c'(\vartheta)}=-\frac{a}{\vartheta}$ και

$\mathbb{E}\left[t(x)\,|\,\vartheta\right]=\mathbb{E}\left[x\,|\,\vartheta\right]=\frac{a}{\vartheta}$.

Για μια οποιαδήποτε στατιστική συνάρτηση $t=t(x)$, ισχύει η ανισότητα $I_{t(x)}(\vartheta)\le I_x(\vartheta)$. Εάν όμως η $t(x)$ είναι επαρκής ως προς την ϑ τότε έχουμε ότι $I_{t(x)}(\vartheta)=I_x(\vartheta)$.

Πράγματι εάν η t είναι επαρκής θα υπάρχουν συναρτήσεις h_{FN} και g_{FN} τέτοιες ώστε $\pi(x|\vartheta)=h_{FN}(x)\,g_{FN}(t(x);\vartheta)$. Έτσι για μια παρατήρηση x και μετασχηματισμό $y=t(x)$ που είναι ένα προς ένα, έχουμε

$$\pi(y|\vartheta)=h_{FN}\left(t^{-1}(y)\right)g_{FN}(y;\vartheta)\left|\frac{dt^{-1}(y)}{dy}\right|,$$

και

$$I_y(\vartheta)=\mathbb{E}\left[-\frac{\partial^2}{\partial\vartheta^2}\log\pi(y|\vartheta)|\vartheta\right]=\mathbb{E}\left\{-\frac{\partial^2}{\partial\vartheta^2}\log\left[h_{FN}\left(t^{-1}(y)\right)g_{FN}(y;\vartheta)\left|\frac{dt^{-1}(y)}{dy}\right|\right]|\vartheta\right\}$$

$$=\mathbb{E}\left\{-\frac{\partial^2}{\partial\vartheta^2}\log g_{FN}(y;\vartheta)|\vartheta\right\}=\mathbb{E}\left\{-\frac{\partial^2}{\partial\vartheta^2}\log\frac{\pi(x|\vartheta)}{h_{FN}(x)}|\vartheta\right\}=\mathbb{E}\left\{-\frac{\partial^2}{\partial\vartheta^2}\log\pi(x|\vartheta)|\vartheta\right\}=I_x(\vartheta).$$

<u>Πρόταση</u>

Δείξτε ότι για κάθε αμερόληπτο εκτιμητή $t=t(x)$ του ϑ η συνάρτηση τετραγωνικού σφάλματος $R\left(t(x)|\vartheta\right)=E\left\{\left[t(x)-\vartheta\right]^2|\vartheta\right\}$ (risk function) ικανοποιεί την ανισότητα $R\left(t(x)|\vartheta\right)\ge I_x(\vartheta)^{-1}$

Για $t=t(x)$ με $\mathbb{E}\left|t\right|<\infty$, θέτουμε $\psi(\vartheta)=\mathbb{E}\left[t|\vartheta\right]$ τότε

$$\psi'(\vartheta)=\frac{d}{d\vartheta}\left\{\int_X t(x)\pi(x|\vartheta)dx\right\}=\int_X t(x)\frac{\partial}{\partial\vartheta}\pi(x|\vartheta)dx$$

$$= \int_X t(x)\left(\frac{\partial}{\partial \vartheta}\log \pi(x|\vartheta)\right)\pi(x|\vartheta)\,dx = \int_X t(x) S(x|\vartheta)\pi(x|\vartheta)\,dx$$
$$= \mathbb{E}\left[t(x) S(x|\vartheta)|\vartheta\right] \Rightarrow \psi'(\vartheta) = Cov\left[t(x), S(x|\vartheta)|\vartheta\right]$$

Δηλαδή, η στατιστική $S(x|\vartheta)$, έχει μέσο 0, διασπορά $I_x(\vartheta)$ και συνδιασπορά ως προς $t(x)$ ίση με $\psi'(\vartheta)$.

Χρησιμοποιώντας την ανισότητα συνδιασποράς (covariance inequality) παίρνουμε το φράγμα Cramer Rao

$$\psi'(\vartheta)^2 = Cov\left[t(x), S(x|\vartheta)\,|\vartheta\right]^2 \le Var\left[t(x)|\vartheta\right] Var\left[S(x|\vartheta)\right] = Var\left[t(x)|\vartheta\right] I_x(\vartheta)$$

και $Var\left[t(x)|\vartheta\right] \ge \dfrac{\psi'(\vartheta)^2}{I_x(\vartheta)}$.

Εάν η στατιστική t είναι αμερόληπτος εκτιμητής (unbiased estimator) για το ϑ, δηλαδή $\mathbb{E}[t|\vartheta] = \vartheta$, θα έχουμε $\psi(\vartheta) = \vartheta$ και

$$R\big(t(x)|\vartheta\big) = \mathbb{E}\left\{\left[t(x) - \vartheta\right]^2 |\vartheta\right\} = \mathbb{E}\left\{\left[t(x) - \mathbb{E}\left[t(x)|\vartheta\right]\right]^2 |\vartheta\right\} = Var\left[t(x)|\vartheta\right].$$

Έτσι για αμερόληπτο εκτιμητή $t(x)$ του ϑ ισχύει η ανισότητα $R\big(t(x)|\vartheta\big) \ge I_x(\vartheta)^{-1}$.

Το Jeffrey's prior χρησιμοποιείται για τον ορισμό noninformative priors (reference priors), και είναι της μορφής

$$\pi^J(\vartheta) \propto \sqrt{I_x(\vartheta)}.$$

Εάν $\eta = \varphi(\vartheta)$ με $\varphi: \Theta \to \mathrm{H}$, αντιστρέψιμος μετασχηματισμός[1] της παραμέτρου ϑ, τότε από το αρχικό μοντέλο $\mathrm{M}_1 = \left\{ f_1(\,\cdot|\vartheta) : \vartheta \in \Theta \right\}$ ορίζουμε την επαναπαραμετροποίηση $\mathrm{M}_2 = \left\{ f_2(\,\cdot|\eta) : \eta \in \mathrm{H} \right\}$ όπου $f_2(\,\cdot|\eta) = f_1\left(\,\cdot|\varphi^{-1}(\eta)\right)$. Εάν η prior πυκνότητα $\pi_1(\vartheta)$ με στήριγμα το Θ είναι non – informative για το μοντέλο M_1, αυτό δεν σημαίνει ότι και η μετασχηματισμένη πυκνότητα

$\pi_2(\eta) = \pi_1\left(\varphi^{-1}(\eta)\right)\left|\dfrac{d\varphi^{-1}(\eta)}{d\eta}\right|$ με στήριγμα H,

είναι non – informative για το μοντέλο M_2. Εάν όμως θέσουμε $\pi_1(\vartheta) = \pi^J(\vartheta)$ τότε

[1] ένα – προς – ένα και επί (bijection).

$$\pi_2(\eta)=\pi_1\left(\varphi^{-1}(\eta)\right)\left|\frac{d\varphi^{-1}(\eta)}{d\eta}\right|=\pi^J(\eta),$$

και τότε μετασχηματίζουμε τον ϑ – noninformative prior σε η – noninformative prior.

Πράγματι

$$\pi_2(\eta)=\pi_1\left(\varphi^{-1}(\eta)\right)\left|\frac{d\varphi^{-1}(\eta)}{d\eta}\right|=\pi_1(\vartheta)\left|\frac{d\vartheta}{d\eta}\right|=\pi^J(\vartheta)\left|\frac{d\vartheta}{d\eta}\right|$$

$$\propto\sqrt{I_x(\vartheta)}\left|\frac{d\vartheta}{d\eta}\right|=\sqrt{\mathbb{E}\left[\left(\frac{\partial}{\partial\vartheta}\log f_1(x|\vartheta)\right)^2|\vartheta\right]}\left|\frac{d\vartheta}{d\eta}\right|$$

$$=\sqrt{\mathbb{E}\left[\left(\frac{\partial}{\partial\vartheta}\log f_1(x|\vartheta)\left|\frac{d\vartheta}{d\eta}\right|\right)^2|\vartheta\right]}=\sqrt{\mathbb{E}\left[\left(\frac{\partial}{\partial\eta}\log f_2(x|\eta)\right)^2|\eta\right]}$$

$$=\sqrt{I_x(\eta)}\propto\pi^J(\eta)\Rightarrow\pi_2(\eta)=\pi^J(\eta).$$

Δείξαμε λοιπόν την εξής πρόταση:

Πρόταση **Το Jeffrey's prior είναι αναλλοίωτο κάτω από μετασχηματισμούς (reparametrization invariant) αυτό το γεγονός αναφέρεται σαν Jeffrey's principle.**

Σημειώστε ότι $\pi(\varphi)=\pi(\vartheta)\left|\frac{d\vartheta}{d\varphi}\right|$ σε standard συμβολισμό σημαίνει ότι

$$f_{\mathrm{H}}(\eta)=f_{\Theta}\left(\varphi^{-1}(\eta)\right)\left|\frac{d\varphi^{-1}(\eta)}{d\eta}\right|\Leftrightarrow f_{\mathrm{H}}(\eta)|d\eta|=f_{\Theta}(\vartheta)|d\vartheta|.$$

Πράγματι εάν $\varphi\uparrow$ τότε και $\varphi^{-1}\uparrow\Rightarrow\frac{d\vartheta}{d\varphi}=\left|\frac{d\vartheta}{d\varphi}\right|$

$$f_{\mathrm{H}}(\eta)d\eta=P\{\eta<\mathrm{H}\le\eta+d\eta\}=P\{\eta<\varphi(\Theta)\le\eta+d\eta\}=P\left\{\varphi^{-1}(\eta)<\Theta\le\varphi^{-1}(\eta+d\eta)\right\}$$
$$=P\{\vartheta<\Theta\le\vartheta+d\vartheta\}=f_{\Theta}(\vartheta)d\vartheta\Leftrightarrow f_{\mathrm{H}}(\eta)|d\eta|=f_{\Theta}(\vartheta)|d\vartheta|$$

Ενώ εάν $\varphi\downarrow$ τότε και $\varphi^{-1}\downarrow\Rightarrow\frac{d\vartheta}{d\varphi}=-\left|\frac{d\vartheta}{d\varphi}\right|$

$$f_{\mathrm{H}}(\eta)d\eta=P\{\eta<\mathrm{H}\le\eta+d\eta\}=P\{\eta<\varphi(\Theta)\le\eta+d\eta\}=P\left\{\varphi^{-1}(\eta+d\eta)\le\Theta<\varphi^{-1}(\eta)\right\}$$

$$= P\{\vartheta + d\vartheta \le \Theta < \vartheta\} = \int_{\vartheta+d\vartheta}^{\vartheta} f_\Theta(u)\,du = -f_\Theta(\vartheta)\,d\vartheta \Leftrightarrow f_{\mathrm{H}}(\eta)|d\eta| = f_\Theta(\vartheta)|d\vartheta| .$$

Jeffrey's prior για το beta – binomial μοντέλο: Έστω Bernoulli παρατηρήσεις $x_i \mid \vartheta \overset{iid}{\sim} Bernoulli(\cdot \mid \vartheta), 1 \le i \le n$. Θεωρώντας μία παρατήρηση x_i έχουμε $\pi(x_i \mid \vartheta) = \vartheta^{x_i}(1-\vartheta)^{1-x_i}$, $x_i \in \{0,1\}$

$$I_{x_i}(\vartheta) = \mathbb{E}\left\{-\frac{\partial^2}{\partial\vartheta^2}\log\left[\vartheta^{x_i}(1-\vartheta)^{1-x_i}\right] \mid \vartheta\right\} = \mathbb{E}\left\{\frac{x_i}{\vartheta^2} + \frac{1-x_i}{(1-\vartheta)^2} \mid \vartheta\right\}$$
$$= \frac{\mathbb{E}(x_i \mid \vartheta)}{\vartheta^2} + \frac{1-\mathbb{E}(x_i \mid \vartheta)}{(1-\vartheta)^2} = \frac{\vartheta}{\vartheta^2} + \frac{1-\vartheta}{(1-\vartheta)^2} = \frac{1}{\vartheta(1-\vartheta)} ,$$

που δίνει

$$I_x(\vartheta) = \sum_{i=1}^{n} I_{x_i}(\vartheta) = \frac{n}{\vartheta(1-\vartheta)} \propto \frac{1}{\vartheta(1-\vartheta)} .$$

Το αντίστοιχο Jeffrey's prior είναι

$$\pi^J(\vartheta) \propto \sqrt{I_x(\vartheta)} \propto g^J(\vartheta) = \vartheta^{-1/2}(1-\vartheta)^{-1/2} ,$$

που είναι proper, επειδή η σταθερά κανονικοποίησης του $g^J(\vartheta)$ είναι

$$C^{-1} = \int_0^1 g^J(\vartheta)\,d\vartheta = \int_0^1 \vartheta^{1/2-1}(1-\vartheta)^{1/2-1}\,d\vartheta = \frac{\Gamma(1/2)^2}{\Gamma(1)} = \pi .$$

Έτσι $\pi^J(\vartheta) = C\,g^J(\vartheta) = \dfrac{1}{\pi}\dfrac{1}{\sqrt{\vartheta(1-\vartheta)}} = Be(\vartheta \mid 1/2, 1/2) \propto \dfrac{1}{\sqrt{\mathrm{var}(x \mid \vartheta)}}$,

που είναι η κατανομή arcsine. Η αντίστοιχη posterior για n παρατηρήσεις θα είναι

$$\pi^J(\vartheta \mid x) \propto \pi^J(\vartheta)\pi(x \mid \vartheta)$$
$$\propto \left\{\vartheta^{-1/2}(1-\vartheta)^{-1/2}\right\} \times \left\{\vartheta^{n\bar{x}}(1-\vartheta)^{n-n\bar{x}}\right\} \propto Be\left(\vartheta \mid n\bar{x} + \frac{1}{2}, n - n\bar{x} + \frac{1}{2}\right).$$

Παρατηρούμε ότι $\mathbb{E}^J(\vartheta \mid x) = \dfrac{n\bar{x} + 1/2}{n+1} \approx \hat{\vartheta}_{MLE} = \bar{x}$.

Ας αναζητήσουμε noninformative συζυγείς priors για το binomial μοντέλο.

$\pi\left(x_i \mid \vartheta\right)=\vartheta^{x_i}\left(1-\vartheta\right)^{1-x_i}=g\left(\vartheta\right)\exp\left(x_i\, c\left(\vartheta\right)\right)$ με $g\left(\vartheta\right)=1-\vartheta$ και $c\left(\vartheta\right)=\log\left(\dfrac{\vartheta}{1-\vartheta}\right)$

$\pi\left(x|\vartheta\right)=\vartheta^{n\overline{x}}\left(1-\vartheta\right)^{n-n\overline{x}}=g\left(\vartheta\right)^{n}\exp\left(n\overline{x}\, c\left(\vartheta\right)\right)$ και θεωρώντας $\pi\left(\vartheta\right)\propto g\left(\vartheta\right)^{a}\exp\left(bc\left(\vartheta\right)\right)$ παίρνουμε $\pi\left(\vartheta \mid x\right)\propto g\left(\vartheta\right)^{a+n\overline{x}}\exp\left(\left(b+n\overline{x}\right)c\left(\vartheta\right)\right)\propto Be\left(\vartheta \mid a+n\overline{x}, b+n-n\overline{x}\right)$.

Όλες οι prior $\pi\left(\vartheta\right)=Be\left(\vartheta \mid a,b\right)$ με $a=b<1$ είναι noninformative Θα μπορούσαμε μάλιστα να συμπεριλάβουμε και την περίπτωση της uniform $a=b=1$, αλλά και την περίπτωση όπου $a=b=0$. Στην τελευταία περίπτωση, έχουμε $\pi\left(\vartheta\right)\propto\vartheta^{-1}\left(1-\vartheta\right)^{-1}\propto\dfrac{1}{\operatorname{var}\left(x \mid \vartheta\right)}$ (improper συζυγής) και $Var\left(\vartheta\right)=\infty$.

Ορίζουμε τις οικογένειες πυκνοτήτων $\mathcal{P}_{BB}\left(\vartheta\right)=\left\{\vartheta\sim Be\left(a,b\right): b=a,\, 0\le a\le 1\right\}$, και $\mathcal{C}_{Bin}\left(\vartheta\right)=\left\{\vartheta\sim Be\left(a,b\right): a\ge 0, b\ge 0\right\}$. Η οικογένεια $\mathcal{P}_{BB}\left(\vartheta\right)$ είναι η οικογένεια των noninformative prior για το beta – binomial μοντέλο που είναι υποσύνολο της οικογένειας $\mathcal{C}_{Bin}\left(\vartheta\right)$ των συζυγών prior για το binomial μοντέλο δειγματοληψίας.

Παρατηρούμε ότι $\pi^{J}\left(\vartheta\right)\in\mathcal{P}_{BB}\left(\vartheta\right)$, αλλά η $\pi^{J}\left(\vartheta\right)$ ξεχωρίζει από τα υπόλοιπα στοιχεία του $\mathcal{P}_{BB}\left(\vartheta\right)$ γιατί είναι η μόνη noninformative prior που είναι αναλλοίωτη κάτω από μετασχηματισμούς (δηλαδή και η μετασχηματισμένη της είναι noninformative). Φυσικά για κάθε $\pi\left(\vartheta\right)\in\mathcal{P}_{BB}\left(\vartheta\right)$ έχουμε
$E\left(\vartheta|x\right)=\dfrac{a+n\overline{x}}{2a+n}$ και επειδή $0\le a\le 1$ έχουμε $E\left(\vartheta|x\right)\cong\hat{\vartheta}_{MLE}=\overline{x}$.

<u>Περίπτωση Beta – Negative Binomial</u>

Έστω $x|\vartheta\sim NB\left(\cdot|n,\vartheta\right)$,.έχουμε $\pi\left(x|\vartheta\right)=\dbinom{x+n-1}{n-1}\vartheta^{n}\left(1-\vartheta\right)^{x}$ και

$$I_x\left(\vartheta\right)=-E\left\{\frac{\partial^2}{\partial\vartheta^2}\log\left[\vartheta^{n}\left(1-\vartheta\right)^{x}\right]|\vartheta\right\}=-E\left\{-\frac{n}{\vartheta^2}-\frac{x}{\left(1-\vartheta\right)^2}|\vartheta\right\}=\frac{n}{\vartheta^2}+\frac{E\left(x|\vartheta\right)}{\left(1-\vartheta\right)^2}$$

$$\Pi_x\left(s\right)=\left\{\frac{\vartheta}{1-s\left(1-\vartheta\right)}\right\}^{n},\ \Pi'_x\left(1\right)=\mathbb{E}\left(x|\vartheta\right)=\frac{n\left(1-\vartheta\right)}{\vartheta}$$

που δίνει

$$I_x\left(\vartheta\right)=\frac{n}{\vartheta^2}+\frac{n}{\vartheta\left(1-\vartheta\right)}=\frac{n}{\vartheta^2\left(1-\vartheta\right)}\propto\frac{1}{\vartheta^2\left(1-\vartheta\right)},$$

και το αντίστοιχο Jeffrey’s prior είναι

$$\pi^{J}\left(\vartheta\right)\propto\sqrt{I_x\left(\vartheta\right)}\propto\vartheta^{-1}\left(1-\vartheta\right)^{-1/2},$$

που είναι improper διότι $\int_0^1 \vartheta^{-1}(1-\vartheta)^{-1/2}\,d\vartheta = \infty$.
Η αντίστοιχη posterior είναι

$$\pi(\vartheta|x) \propto \pi^J(\vartheta)\pi(x|\vartheta) \propto \left\{\vartheta^{-1}(1-\vartheta)^{-1/2}\right\} \times \left\{\vartheta^n (1-\vartheta)^x\right\} \propto Be\left(\vartheta|\ n,\ x+\frac{1}{2}\right).$$

Παρατηρούμε ότι $E(\vartheta|x) = \dfrac{n}{n+x+1/2} \approx \dfrac{n}{n+x} = \hat{\vartheta}_{MLE}$

Περίπτωση Gamma – Poisson
Έστω δεδομένα Poisson $x_i|\vartheta \overset{iid}{\sim} Po(\cdot|\vartheta)$, $1 \le i \le n$. Θεωρώντας μία παρατήρηση x_i έχουμε $\pi(x_i|\vartheta) = \dfrac{e^{-\vartheta}\vartheta^{x_i}}{x_i!}$ και

$$I_{x_i}(\vartheta) = -\mathbb{E}\left\{\frac{\partial^2}{\partial\vartheta^2}\log\left[\frac{e^{-\vartheta}\vartheta^{x_i}}{x_i!}\right]|\vartheta\right\} = -\mathbb{E}\left\{-\frac{x_i}{\vartheta^2}|\vartheta\right\} = \frac{\mathbb{E}(x_i|\vartheta)}{\vartheta^2} = \frac{1}{\vartheta},$$

που δίνει

$$I_x(\vartheta) = \sum_{i=1}^{n} I_{x_i}(\vartheta) = \frac{n}{\vartheta} \propto \frac{1}{\vartheta}.$$

Το αντίστοιχο Jeffrey's prior είναι

$$\pi^J(\vartheta) \propto \sqrt{I_x(\vartheta)} \propto \vartheta^{-1/2},$$

που δεν είναι proper εφόσον $\int_{\mathbb{R}^+} \pi^J(\vartheta)\,d\vartheta = \infty$. Όμως η αντίστοιχη posterior για n παρατηρήσεις είναι proper

$$\pi(\vartheta|x) \propto \pi^J(\vartheta)\pi(x|\vartheta) \propto \left\{\vartheta^{-1/2}\right\} \times \left\{e^{-n\vartheta}\vartheta^{n\bar{x}}\right\} \propto Ga\left(\vartheta|\ n\bar{x}+\frac{1}{2},\ n\right).$$

Παρατηρούμε ότι $E(\vartheta|x) = \dfrac{n\bar{x}+1/2}{n} \approx \hat{\vartheta}_{MLE} = \bar{x}$

Περίπτωση Gamma – Gamma
1. Έστω δεδομένα Gamma $x_i|\vartheta \overset{iid}{\sim} Ga(\cdot|a,\vartheta)$, $1 \le i \le n$. Έχουμε $\pi(x_i|\vartheta) \propto \vartheta^a e^{-\vartheta x_i}$ και $\pi(x|\vartheta) \propto \vartheta^{na}e^{-\vartheta n\bar{x}}$

$$I_x(\vartheta) = -\mathbb{E}\left\{\frac{\partial^2}{\partial\vartheta^2}\log\left[\vartheta^{na-n}e^{-\vartheta n\bar{x}}\right]|\vartheta\right\} = -\mathbb{E}\left\{-\frac{na}{\vartheta^2}|\vartheta\right\} \propto \frac{1}{\vartheta^2},$$

Το αντίστοιχο Jeffrey's prior είναι $\pi^J(\vartheta) \propto \vartheta^{-1}$, που δεν είναι proper εφόσον $\int_{\mathbb{R}^+} \vartheta^{-1}d\vartheta = \infty$. Όμως η αντίστοιχη posterior για n παρατηρήσεις είναι proper

$$\pi^J(\vartheta|x) \propto \pi^J(\vartheta)\pi(x|\vartheta) \propto \{\vartheta^{-1}\}\times\{\vartheta^{na}e^{-\vartheta n\bar{x}}\} = \vartheta^{na-1}e^{-\vartheta n\bar{x}} \propto Ga(\vartheta|\, na, n\bar{x}).$$

Παρατηρούμε ότι $\mathbb{E}(\vartheta|x) = \frac{a}{\bar{x}} = \hat{\vartheta}_{MLE}$.

2. . Έστω δεδομένα Gamma $x_i|\vartheta \overset{iid}{\sim} Ga(\cdot|\vartheta, b)$, $1\le i\le n$. Τότε $\pi(x_i|\vartheta) = \frac{b^\vartheta}{\Gamma(\vartheta)}x_i^{\vartheta-1}e^{-bx_i}$ και γνωρίζουμε ότι ο συζυγής prior είναι nonstandard. Η πληροφορία κατά Fisher είναι

$$I_{x_i}(\vartheta) = -\mathbb{E}\left\{\frac{\partial^2}{\partial\vartheta^2}\log\left[\frac{b^\vartheta}{\Gamma(\vartheta)}x_i^{\vartheta-1}e^{-bx_i}\right]|\vartheta\right\} = -\mathbb{E}\{-\Psi'(\vartheta)|\vartheta\} = \Psi'(\vartheta) \Rightarrow I_x(\vartheta) = n\Psi'(\vartheta),$$

με $\Psi'(\vartheta) \triangleq \frac{d^2}{d\vartheta^2}\log\Gamma(\vartheta) = \sum_{k=0}^{\infty}\frac{1}{(\vartheta+k)^2}, \forall\vartheta\in\mathbb{R}\setminus\mathbb{Z}^{\le 0}$.

Το Jeffrey’s prior είναι $\pi^J(\vartheta) \propto \sqrt{\Psi'(\vartheta)}$.

Περίπτωση Normal – Normal με άγνωστο location και γνωστό scale

Εδώ έχουμε δεδομένα της μορφής $x_i|\vartheta \overset{iid}{\sim} N(\cdot|\vartheta,\sigma^2)$, $1\le i\le n$. Έχουμε

$$\pi(x_i|\vartheta) \propto \exp\left\{-\frac{1}{2\sigma^2}(x_i-\vartheta)^2\right\}.$$

$$I_{x_i}(\vartheta) = -\mathbb{E}\left\{\frac{\partial^2}{\partial\vartheta^2}\left[-\frac{1}{2\sigma^2}(x_i-\vartheta)^2\right]|\vartheta\right\} = -\mathbb{E}\left\{-\frac{1}{\sigma^2}|\vartheta\right\} = \frac{1}{\sigma^2},\ I_x(\vartheta) = \frac{n}{\sigma^2}$$

Το αντίστοιχο Jeffrey’s prior είναι $\pi^J(\vartheta) \propto 1$, που δεν είναι proper εφόσον $\int_{\mathbb{R}}\pi^J(\vartheta)d\vartheta = \infty$. Όμως η αντίστοιχη posterior για n παρατηρήσεις είναι proper

$$\pi(\vartheta|x) \propto \pi^J(\vartheta)\pi(x|\vartheta) \propto 1\times\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^n(x_i-\vartheta)^2\right\}$$

$$= \exp\left\{-\frac{\tau}{2}\left(\sum_{i=1}^n x_i^2 - 2n\bar{x}\vartheta + n\vartheta^2\right)\right\} \propto \exp\left\{-\frac{\tau}{2}\left(n\vartheta^2 - 2n\bar{x}\vartheta\right)\right\} = \exp\left\{-\frac{\tau n}{2}\left(\vartheta^2 - 2\bar{x}\vartheta\right)\right\}$$

$$\propto \exp\left\{-\frac{\tau n}{2}(\vartheta-\bar{x})^2\right\} \propto N\left(\vartheta|\,\bar{x}, \frac{\sigma^2}{n}\right)$$

Παρατηρούμε ότι $E(\vartheta|x) = \hat{\vartheta}_{MLE}$

Περίπτωση Normal – Normal με γνωστό location και άγνωστο scale

Εδώ έχουμε δεδομένα της μορφής $x_i|\vartheta \overset{iid}{\sim} N(\cdot|\mu,\vartheta)$, $1\le i\le n$. Για μία παρατήρηση έχουμε

$$\pi\left(x_i \mid \vartheta\right)=\frac{1}{\sqrt{2\pi\vartheta}}\exp\left[-\frac{1}{2\vartheta}\left(x_i-\mu\right)^2\right]\propto\frac{1}{\sqrt{\vartheta}}\exp\left[-\frac{1}{2\vartheta}\left(x_i-\mu\right)^2\right].$$

$$I_{x_i}\left(\vartheta\right)=-\mathbb{E}\left\{\frac{\partial^2}{\partial\vartheta^2}\left[-\frac{1}{2}\log\left(\vartheta\right)-\frac{1}{2\vartheta}\left(x_i-\mu\right)^2\right]\mid\vartheta\right\}=-\mathbb{E}\left\{\frac{1}{2\vartheta^2}-\frac{1}{\vartheta^3}\left(x_i-\mu\right)^2\mid\vartheta\right\}$$
$$=-\frac{1}{2\vartheta^2}+\frac{1}{\vartheta^3}\mathbb{E}\left[\left(x_i-\mu\right)^2\mid\vartheta\right]=-\frac{1}{2\vartheta^2}+\frac{1}{\vartheta^2}=\frac{1}{2\vartheta^2},\ I_x\left(\vartheta\right)=\frac{n}{2\vartheta^2}$$

Το αντίστοιχο Jeffrey's prior είναι $\pi^J\left(\vartheta\right)\propto\frac{1}{\vartheta}$, που δεν είναι proper εφόσον $\int_{\mathbb{R}}\pi^J\left(\vartheta\right)d\vartheta=\infty$. Όμως η αντίστοιχη posterior για n παρατηρήσεις είναι proper

$$\pi\left(x\mid\vartheta\right)=\frac{\vartheta^{-n/2}}{\left(2\pi\right)^{n/2}}\exp\left[-\frac{1}{2\vartheta}\sum_{i=1}^{n}\left(x_i-\mu\right)^2\right]\propto\vartheta^{-n/2}\exp\left[-\frac{1}{2\vartheta}t\left(x\right)\right],$$

όπου $t\left(x\right)=\sum_{i=1}^{n}\left(x_i-\mu\right)^2$

$$\pi\left(\vartheta\mid x\right)\ \propto\ \left\{\vartheta^{-1}\right\}\times\left\{\vartheta^{-n/2}\exp\left(-\frac{1}{2\vartheta}t\left(x\right)\right)\right\}=\vartheta^{-\left(\frac{n}{2}+1\right)}\exp\left(-\frac{1}{2\vartheta}t\left(x\right)\right)\propto In\ \chi^2\left(\vartheta\mid n,\,t\right),$$

όπου $Inv\chi^2\left(n,\,t\right)=IG\left(\frac{n}{2},\,\frac{t\left(x\right)}{2}\right)$, που δίνει

$$\mathbb{E}\left(\vartheta\mid x\right)=\frac{\frac{nt\left(x\right)}{2}}{\frac{n}{2}-1}\approx t\left(x\right)=\widehat{\vartheta}_{MLE}$$

Παρατήρηση

Εάν $X\sim Ga\left(a,b\right)$ τότε $Y=X^{-1}\sim IG\left(a,b\right)$ με $IG\left(y\mid a,b\right)=\frac{b^a}{\Gamma\left(a\right)}y^{-\left(a+1\right)}\exp\left(-\frac{b}{y}\right),\ y>0$

Περίπτωση Gamma – Lognormal

Εδώ έχουμε δεδομένα της μορφής $x_i\mid\vartheta\overset{iid}{\sim}LN\left(\cdot\mid\mu,\vartheta^{-1}\right)$, $1\leq i\leq n$. Έχουμε

$$\pi\left(x_i\mid\vartheta\right)=\frac{\sqrt{\vartheta}}{x_i\sqrt{2\pi}}\exp\left[-\frac{\vartheta}{2}\left(\log\left(x_i\right)-\mu\right)^2\right]\propto\sqrt{\vartheta}\exp\left[-\frac{\vartheta}{2}\left(\log\left(x_i\right)-\mu\right)^2\right].$$

$$I_{x_i}(\vartheta) = -\mathbb{E}\left\{\frac{\partial^2}{\partial\vartheta^2}\left[\frac{1}{2}\log(\vartheta) - \frac{\vartheta}{2}\left(\log(x_i)-\mu\right)^2\right]|\vartheta\right\} = -\mathbb{E}\left\{-\frac{1}{2\vartheta^2}|\vartheta\right\} = \frac{1}{2\vartheta^2}.$$

Και έτσι παίρνουμε $I_x(\vartheta) = \frac{n}{2\vartheta^2}$.

Το αντίστοιχο Jeffrey's prior είναι $\pi^J(\vartheta) \propto \frac{1}{\vartheta}$, που δεν είναι proper εφόσον $\int_{\mathbb{R}} \pi^J(\vartheta)\, d\vartheta = \infty$. Όμως η αντίστοιχη posterior για n παρατηρήσεις $x = (x_1, \ldots, x_n)$ είναι proper:

$$\pi(\vartheta|x) \;\propto\; \left\{\vartheta^{-1}\right\} \times \left\{\vartheta^{n/2}\exp\left(-\frac{\vartheta}{2}\sum_{i=1}^{n}\left(\log(x_i)-\mu\right)^2\right)\right\} \propto Ga\left(\vartheta|\frac{n}{2}, \frac{1}{2}\sum_{i=1}^{n}\left(\log(x_i)-\mu\right)^2\right),$$

που δίνει

$$\mathbb{E}(\vartheta|x) = \frac{\frac{n}{2}}{\frac{1}{2}\sum_{i=1}^{n}\left(\log(x_i)-\mu\right)^2} = \frac{1}{\frac{1}{n}\sum_{i=1}^{n}\left(\log(x_i)-\mu\right)^2} = \hat{\vartheta}_{MLE}$$

Για τον υπολογισμό του ΕΜΠ έχουμε

$$\frac{\partial}{\partial\vartheta}\log\pi(x|\vartheta) = \frac{\partial}{\partial\vartheta}\left[\frac{n}{2}\log(\vartheta) - \frac{\vartheta}{2}\sum_{i=1}^{n}\left(\log(x_i)-\mu\right)^2\right]$$

$$= \frac{n}{2\vartheta} - \frac{1}{2}\sum_{i=1}^{n}\left(\log(x_i)-\mu\right)^2 = 0 \Rightarrow \vartheta = \left\{\frac{1}{n}\sum_{i=1}^{n}\left(\log(x_i)-\mu\right)^2\right\}^{-1}$$

$$\frac{\partial^2}{\partial\vartheta^2}\log\pi(x|\vartheta) = \frac{\partial}{\partial\vartheta}\left[\frac{n}{2\vartheta} - \frac{1}{2}\sum_{i=1}^{n}\left(\log(x_i)-\mu\right)^2\right] = -\frac{n}{2\vartheta^2} < 0 \Rightarrow \log\pi(x|\vartheta)$$ είναι convex,

αλλά και $\frac{\partial^2}{\partial\vartheta^2}\pi(x|\vartheta)$ είναι convex εφόσον

$$\frac{\partial^2}{\partial\vartheta^2}\log\pi(x|\vartheta) = \frac{\partial}{\partial\vartheta}\left[\frac{\pi(x|\vartheta)}{\frac{\partial}{\partial\vartheta}\pi(x|\vartheta)}\right] = \frac{\left(\frac{\partial}{\partial\vartheta}\pi(x|\vartheta)\right)^2 - \pi(x|\vartheta)\frac{\partial^2}{\partial\vartheta^2}\pi(x|\vartheta)}{\left(\frac{\partial}{\partial\vartheta}\pi(x|\vartheta)\right)^2}$$

$$\Rightarrow \frac{\partial^2}{\partial \vartheta^2}\pi(x|\vartheta) > \frac{\left(\frac{\partial}{\partial \vartheta}\pi(x|\vartheta)\right)^2}{\pi(x|\vartheta)} > 0 .$$

Οπότε $\widehat{\vartheta}_{MLE} = \left\{\frac{1}{n}\sum_{i=1}^{n}\left(\log(x_i)-\mu\right)^2\right\}^{-1} = \mathbb{E}(\vartheta|x)$.

<u>Περίπτωση Pareto – Uniform</u>

Γνωρίζουμε ότι ο συζυγής prior για το μοντέλο $[x_i|\vartheta] \overset{iid}{\sim} \mathcal{U}(\cdot|0,\vartheta),\ i=1,\ldots,n$ έχει κατανομή Pareto, $\pi(\vartheta) = Pa(\vartheta\,|\,a,b) \propto \vartheta^{-(a+1)}1(\vartheta > b)$, και ότι η πληροφορία κατά Fisher είναι $I_x(\vartheta) = n\vartheta^{-2}$, έτσι $\pi^J(\vartheta) \propto \frac{1}{\vartheta}$

$$\pi^J(\vartheta\,|\,x) \propto \vartheta^{-1}\times\vartheta^{-n}1\left(\vartheta \ge x_{(n)}\right) = \vartheta^{-(n+1)}1\left(\vartheta \ge x_{(n)}\right) \propto Pa\left(\vartheta\,|\,n, x_{(n)}\right),$$

που δίνει $\mathbb{E}^J(\vartheta|x) = \frac{n\,x_{(n)}}{n-1} \approx x_{(n)},\ n > 1$.

Πολυδιάστατη πληροφορία κατά Fisher

Όταν η παράμετρος ϑ είναι ένα d-διάστατο διάνυσμα, η πληροφορία κατά Fisher παίρνει την μορφή ενός $d\times d$ πίνακα, τον Fisher Information Matrix $I_x(\vartheta)$, με $(i,j)-$στοιχείο

$$\left(I_x(\vartheta)\right)_{ij} = \mathbb{E}\left\{S_i(x|\vartheta)S_j(x|\vartheta)\right\}$$

όπου $S_i(x|\vartheta) = \frac{\partial}{\partial\vartheta_i}\ln\pi(x|\vartheta),\ 1\le i\le d$. Επειδή $\mathbb{E}\left[S_i(x|\vartheta)\right]=0$ για κάθε i, θα έχουμε $\left(I_x(\vartheta)\right)_{ij} = Cov\left[S_i(x|\vartheta), S_j(x|\vartheta)\right]$. Αποδεικνύεται ότι όπως και στην μονοδιάστατη περίπτωση θα έχουμε

$$\left(I_x(\vartheta)\right)_{ij} = -\mathbb{E}\left[\frac{\partial^2}{\partial\vartheta_i\partial\vartheta_j}\log\pi(x|\vartheta)\,|\,\vartheta\right].$$

Το πολυδιάστατο Jeffrey's prior ορίζεται σαν $\pi^J(\vartheta) \propto \sqrt{\det\left(I_x(\vartheta)\right)}$

Παράδειγμα

Έστω κανονικές παρατηρήσεις $\left[x_i|\mu,\sigma^2\right] \overset{iid}{\sim} N\left(\cdot|\mu,\sigma^2\right)$ με $\vartheta=\left(\mu,\sigma^2\right)$. Θέλουμε να εκτιμήσουμε και τις δύο παραμέτρους μ και σ^2. Θα υπολογίσουμε το Jeffrey's prior χρησιμοποιώντας μόνο μια παρατήρηση x.

$$\pi\left(x|\mu,\sigma^2\right)\propto\frac{1}{\sigma}\exp\left[-\frac{1}{2\sigma^2}(x-\mu)^2\right],\ \log\left[\pi\left(x|\vartheta\right)\right]\propto-\log(\sigma)-\frac{1}{2\sigma^2}(x-\mu)^2,$$

$$\frac{\partial\log\left[\pi\left(x|\vartheta\right)\right]}{\partial\mu}=\frac{1}{\sigma^2}(x-\mu),\ \frac{\partial^2\log\left[\pi\left(x|\vartheta\right)\right]}{\partial\mu^2}=-\frac{1}{\sigma^2},\ -\mathbb{E}\left\{\frac{\partial^2\log\left[\pi\left(x|\vartheta\right)\right]}{\partial\mu^2}|\vartheta\right\}=\frac{1}{\sigma^2},$$

$$\frac{\partial^2\log\left[\pi\left(x|\vartheta\right)\right]}{\partial\mu\,\partial\sigma}=-\frac{2}{\sigma^3}(x-\mu),\ -\mathbb{E}\left\{\frac{\partial^2\log\left[\pi\left(x|\vartheta\right)\right]}{\partial\mu\,\partial\sigma}|\vartheta\right\}=\frac{2}{\sigma^3}\left(\mathbb{E}\left(x|\vartheta\right)-\mu\right)=0,$$

$$\frac{\partial\log\left[\pi\left(x|\vartheta\right)\right]}{\partial\sigma}=-\frac{1}{\sigma}+\frac{1}{\sigma^3}(x-\mu)^2,\ \frac{\partial^2\log\left[\pi\left(x|\vartheta\right)\right]}{\partial\sigma^2}=\frac{1}{\sigma^2}-\frac{3}{\sigma^4}(x-\mu)^2,$$

$$-\mathbb{E}\left\{\frac{\partial^2\log\left[\pi\left(x|\vartheta\right)\right]}{\partial\sigma^2}|\vartheta\right\}=-\frac{1}{\sigma^2}+\frac{3}{\sigma^4}\mathbb{E}\left\{(x-\mu)^2|\vartheta\right\}=\frac{2}{\sigma^2},$$

με αποτέλεσμα

$$I_x\left(\mu,\sigma^2\right)=\begin{bmatrix}\frac{1}{\sigma^2} & 0\\ 0 & \frac{2}{\sigma^2}\end{bmatrix},$$

και

$$\pi^J\left(\mu,\sigma^2\right)\propto\sqrt{\det\left(I_x\left(\mu,\sigma^2\right)\right)}\propto\frac{1}{\sigma^2}.$$

Ισχύει ότι $\pi^J\left(\mu,\sigma^2\right)=\pi^J(\mu)\pi^J\left(\sigma^2\right)\propto1\times\frac{1}{\sigma^2}$