Normal μοντέλο με γνωστή διασπορά, και άγνωστο μέσο. Εδώ έχουμε να εκτιμήσουμε τη location παράμετρο του normal μοντέλου για γνωστό precision $\tau = \sigma^{-2}$ δηλαδή η κατανομή δειγματοληψίας είναι $\pi\left(x_i|\vartheta\right) = N\left(x_i \mid \vartheta, \tau^{-1}\right)$. Θα δείξουμε ότι η συζυγής κατανομή είναι Normal. Αναπαριστούμε πρώτα τη $\pi\left(x_i|\vartheta\right)$ σαν μέλος της EF

$$\pi\left(x_i|\vartheta\right) = \sqrt{\frac{\tau}{2\pi}} \exp\left[-\frac{\tau}{2}\left(x_i - \vartheta\right)^2\right] = \sqrt{\frac{\tau}{2\pi}} \exp\left(-\frac{\tau}{2}x_i^2\right)\exp\left(-\frac{\tau}{2}\vartheta^2\right)\exp\left(\tau\vartheta x_i\right)$$

έτσι έχουμε

$$h\left(x_i\right) = \sqrt{\frac{\tau}{2\pi}} \exp\left[-\frac{\tau}{2}x_i^2\right],\ g\left(\vartheta\right) = \exp\left(-\frac{\tau}{2}\vartheta^2\right),\ c\left(\vartheta\right) = \tau\vartheta,\ t\left(x_i\right) = x_i .$$

Η πιθανοφάνεια είναι

$$\pi\left(x|\vartheta\right) = \prod_{i=1}^{n} \sqrt{\frac{\tau}{2\pi}} \exp\left[-\frac{\tau}{2}\left(x_i - \vartheta\right)^2\right],\ \text{με}\ \sigma^2 = Var\left(x_i|\vartheta\right) = \tau^{-1},$$

$$= \left(\frac{\tau}{2\pi}\right)^{n/2} \exp\left[-\frac{\tau}{2}\left(\sum_{i=1}^{n} x_i^2 - 2\vartheta\sum_{i=1}^{n} x_i + n\vartheta^2\right)\right]$$

$$= \left(\frac{\tau}{2\pi}\right)^{n/2} \exp\left(-\frac{\tau}{2}\sum_{i=1}^{n} x_i^2\right)\exp\left(-\frac{1}{2}n\tau\vartheta^2\right)\exp\left(\tau\vartheta n\overline{x}\right)$$

$$\overset{\vartheta}{\propto} \exp\left(-\frac{\tau n}{2}\vartheta^2\right)\exp\left(\tau\vartheta n\,\overline{x}\right) \overset{\vartheta}{\propto} N\left(\vartheta \mid \overline{x}, \frac{1}{n\tau}\right) = N\left(\vartheta \mid \overline{x}, \frac{\sigma^2}{n}\right),$$

με $t_{suff}\left(x\right) = \sum_{i=1}^{n} t\left(x_i\right) = n\overline{x}$.

$$\pi_{NCP}\left(\vartheta\right) \propto \exp\left(-\frac{\tau d}{2}\vartheta^2\right)\exp\left(b\tau\vartheta\right) \propto N\left(\vartheta \mid \frac{b}{d}, \frac{1}{\tau d}\right).$$

Θέτουμε λοιπόν $\pi\left(\vartheta\right) = N\left(\vartheta \mid m, c^{-1}\right) \propto \exp\left(-\frac{c}{2}\vartheta^2 + cm\vartheta\right)$ και η posterior γίνεται:

$$\pi\left(\vartheta|x\right) \propto \exp\left(-\frac{c}{2}\vartheta^2 + cm\vartheta\right) \times \exp\left(-\frac{\tau n}{2}\vartheta^2 + \tau\vartheta n\,\overline{x}\right)$$

$$= \exp\left(-\frac{1}{2}(c+\tau n)\vartheta^2 + (cm+\tau n\overline{x})\vartheta \right) \propto N\left(\vartheta | \frac{mc+n\tau\overline{x}}{c+n\tau}, \frac{1}{c+n\tau} \right)$$

Ο εκτιμητής ως προς τετραγωνική συνάρτηση απώλειας είναι:

$$\hat{\vartheta}_{BAYES} = \mathbb{E}(\vartheta|x) = \frac{mc+n\tau\overline{x}}{c+n\tau} = \left(\frac{c}{c+n\tau} \right) m + \left(1 - \frac{c}{c+n\tau} \right) \overline{x}$$

Normal model με συνολική απουσία a-priori πεποιθήσεων

Το posterior precision για μεγάλο n είναι

$$n\tau + c = \frac{1}{Var(\vartheta|x)} = \frac{n}{Var(x_i|\vartheta)} + \frac{1}{Var(\vartheta)} \overset{n>>1}{\approx} \frac{n}{Var(x_i|\vartheta)} = \frac{n}{\sigma^2}$$

Το posterior mean για μεγάλο n γίνεται

$\mathbb{E}(\vartheta|x) = \frac{mc+n\tau\overline{x}}{c+n\tau} \overset{n>>1}{\approx} \overline{x}$, δηλαδή $\hat{\vartheta}_{Bayes} \overset{n>>1}{\approx} \hat{\vartheta}_{MLE}$.

Έτσι για $n>>1$ έχουμε $[\vartheta|x] \sim N\left(\cdot | \overline{x}, \frac{\sigma^2}{n} \right)$. Δηλαδή οριακά η prior δεν επηρεάζει τη posterior.

Γνωρίζουμε άλλωστε ότι εφόσον $x_i|\vartheta \sim N(\cdot | \vartheta, \sigma^2)$ θα έχουμε και

$$\frac{1}{n}\sum_{i=1}^{n} x_i | \vartheta \sim N\left(\cdot | \vartheta, \frac{\sigma^2}{n} \right)$$

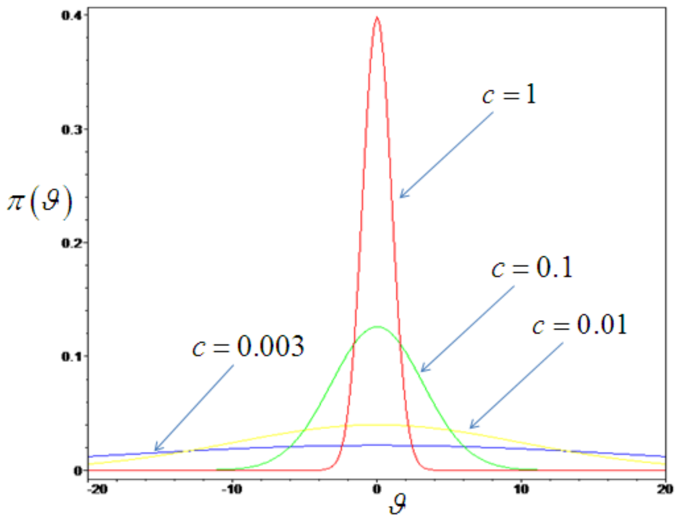
Θυμηθείτε ότι:

$$\left[x_i | \mu_i, \sigma_i^2 \right] \overset{ind}{\sim} N(\mu_i, \sigma_i^2) \Rightarrow \sum_{i=1}^{n} \alpha_i x_i \sim N\left(\sum_{i=1}^{n} \alpha_i \mu_i, \sum_{i=1}^{n} \alpha_i^2 \sigma_i^2 \right).$$

Η προηγούμενη posterior μπορεί να προκύψει και όταν $c \to 0^+$. Σε αυτήν την περίπτωση οι a-priori πεποιθήσεις έχουν πολύ μικρή βαρύτητα εφόσον $Var(\vartheta) \to +\infty$. Πιο αναλυτικά όταν $c \to 0^+$ έχουμε $\pi(\vartheta) \propto \exp\left(-\frac{c}{2}(\vartheta - m)^2 \right) \to 1$,

δηλαδή η prior γίνεται $\pi(\vartheta) \propto 1$ (flat prior δες επόμενο κεφάλαιο) και $\pi(\vartheta) = const$

που δίνει $\int_{\mathbb{R}} \pi(\vartheta)d\vartheta = \infty$ και η prior είναι μη ολοκληρώσιμη. **Λέμε τότε ότι έχουμε μια improper prior. Οι εκτίμηση με την χρήση *improper prior* είναι αποδεκτή εάν η αντίστοιχη posterior ολοκληρώσιμη (proper posterior)**, δηλαδή $\int_{\mathbb{R}} \pi(\vartheta|x)d\vartheta = 1$.



Στο προηγούμενο σχήμα η $\pi(\vartheta) = N\left(\vartheta|\, 0,\ (0.003)^{-1}\right)$ είναι «κοντά» σε flat prior, δηλαδή **ισομήκη διαστήματα του ϑ να είναι περίπου ισοπίθανα.**

Για $\pi_{imp}(\vartheta) \propto 1$ η posterior γίνεται

$$\pi(\vartheta|x) \propto 1 \times \exp\left(-\frac{\tau n}{2}\vartheta^2\right)\exp(\tau\vartheta n\,\bar{x}) = \exp\left\{-\frac{\tau n}{2}\left(\vartheta^2 - 2\bar{x}\vartheta\right)\right\}$$

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

Το ίδιο αποτέλεσμα παίρνουμε κατ' ευθείαν από την παρατήρηση ότι

$$\pi(x|\vartheta) \overset{\vartheta}{\propto} N\left(\vartheta|\,\bar{x}, \frac{\sigma^2}{n}\right) \Rightarrow \pi(\vartheta|x) \propto 1 \times N\left(\vartheta|\,\bar{x}, \frac{\sigma^2}{n}\right)$$

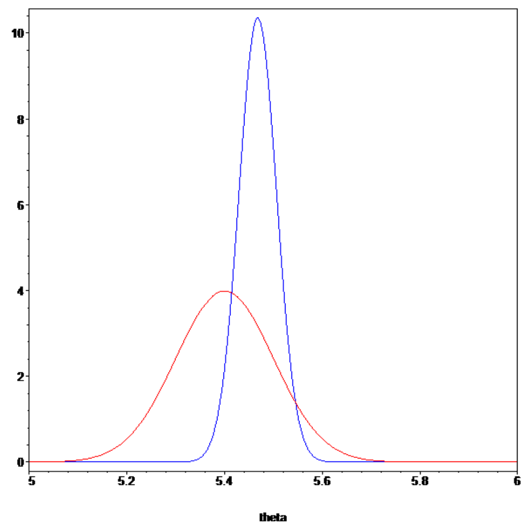
$$\begin{array}{ll} prior & \pi_{imp}(\vartheta) \propto 1 \\ likelihood & \pi(x|\vartheta) = \prod_{i=1}^{n} N(x_i|\vartheta, \sigma^2) \\ posterior & \pi(\vartheta|x) = N\left(\vartheta|\,\overline{x}, \dfrac{\sigma^2}{n}\right) \end{array}$$

Παρατηρούμε ότι στην περίπτωση του flat prior $\pi(\vartheta) \propto 1$ για $\vartheta \in \mathbb{R}$, θα έχουμε $E(\vartheta|x) = \lim_{c \to 0^+} \dfrac{mc + n\tau\overline{x}}{n\tau + c} = \overline{x}$ και η posterior δίνει την ίδια εκτίμηση όπως και η κλασσική στατιστική.

Αριθμητικό παράδειγμα

Ο H. Cavendish τον 18^o αιώνα έκανε $n = 23$ μετρήσεις για την πυκνότητα ϑ της γης με $\overline{x} = 5.48$, ενώ θεώρησε τη σ^2 γνωστή $Var(x_i|\vartheta) = \dfrac{1}{25} = 0.04$. Δηλαδή είχε δειγματοληπτική κατανομή $x_i|\vartheta \overset{iid}{\sim} N\left(\cdot|\vartheta, \dfrac{1}{25}\right)$, $1 \le i \le 23$ Από προηγούμενα πειράματα η a-priori πληροφορία ήταν $\vartheta \sim N\left(\cdot\,|\,5.40, \dfrac{1}{100}\right)$. Σύμφωνα με τα προηγούμενα έχουμε

$$m = 5.40,\ c = 100,\ \tau = 25,\ n = 23 \Rightarrow [\vartheta|x] \sim N\left(5.468148, \frac{1}{675}\right)$$



Prior και posterior predictives

Για μια παρατήρηση $x|\vartheta \sim N\left(\vartheta,\tau^{-1}\right)$ και prior $\vartheta \sim N\left(m,c^{-1}\right)$ έχουμε

$$\pi(x)=\int_{\mathbb{R}} N\left(x|\ \vartheta,\tau^{-1}\right)N\left(\vartheta|m,c^{-1}\right)d\vartheta$$

$$=\frac{\sqrt{c\tau}}{2\pi}\int_{\mathbb{R}}\exp\left\{-\frac{c}{2}(\vartheta-m)^2\right\}\exp\left\{-\frac{\tau}{2}(x-\vartheta)^2\right\}d\vartheta$$

$$=\frac{\sqrt{c\tau}}{2\pi}\exp\left\{-\frac{c}{2}m^2-\frac{\tau}{2}x^2\right\}\int_{\mathbb{R}}\exp\left\{-\frac{c+\tau}{2}\vartheta^2+\left(mc+x\tau\right)\vartheta\right\}d\vartheta$$

$$=\frac{\sqrt{c\tau}}{2\pi}\exp\left\{-\frac{c}{2}m^2-\frac{\tau}{2}x^2\right\}\exp\left\{\frac{\left(mc+x\tau\right)^2}{2\left(c+\tau\right)}\right\}\int_{\mathbb{R}}\exp\left\{-\frac{c+\tau}{2}\left(\vartheta-\frac{mc+x\tau}{c+\tau}\right)^2\right\}d\vartheta$$

$$=\frac{\sqrt{c\tau}}{2\pi}\exp\left\{-\frac{c}{2}m^2-\frac{\tau}{2}x^2\right\}\exp\left\{\frac{\left(mc+x\tau\right)^2}{2\left(c+\tau\right)}\right\}\sqrt{\frac{2\pi}{c+\tau}}$$

$$=\sqrt{\frac{c\,\tau}{2\pi\left(c+\tau\right)}}\exp\left\{-\frac{c\tau}{c+\tau}\left(x-m\right)^2\right\}=N\left(x|\ m,\frac{c+\tau}{c\,\tau}\right)$$

$$=N\left(x|\ m,\ c^{-1}+\tau^{-1}\right).$$

Έχουμε λοιπόν ότι:

$$\pi(x)=\int_{\mathbb{R}} N\left(x|\ \vartheta,\tau^{-1}\right)N\left(\vartheta|m,c^{-1}\right)d\vartheta=N\left(x|\ m,\ c^{-1}+\tau^{-1}\right),$$

δηλαδή η μίξη της οικογένειας $N\left(x|\ \vartheta,\tau^{-1}\right)$ με μέτρο μίξης $\Pi\left(d\vartheta\right)=N\left(\vartheta|m,c^{-1}\right)d\vartheta$ είναι και πάλι **normal με μέσο, το μέσο του μέτρου μίξης και διασπορά το άθροισμα των διασπορών**.

Έστω τώρα ότι έχουμε δει τις πρώτες n παρατηρήσεις x_i και θεωρούμε μελλοντική (unobserved) παρατήρηση $x_{n+1}|\vartheta \sim N\left(\vartheta,\tau^{-1}\right)$. Το posterior predictive είναι:

$$\pi\left(x_{n+1}|x\right)=\int_{\Theta}\pi\left(x_{n+1}|\vartheta\right)\pi\left(\vartheta|x\right)d\vartheta=\int_{\mathbb{R}}N\left(x_{n+1}|\vartheta,\frac{1}{\tau}\right)N\left(\vartheta\,|\,\frac{mc+n\tau\overline{x}}{c+n\tau},\frac{1}{c+n\tau}\right)d\vartheta$$

$$=N\left(x_{n+1}|\,\frac{mc+n\tau\overline{x}}{c+n\tau},\frac{1}{\tau}+\frac{1}{c+n\tau}\right).$$

Normal – normal μοντέλο με άγνωστή διασπορά, και γνωστό μέσο. Εδώ έχουμε να εκτιμήσουμε τη scale παράμετρο του normal μοντέλου για γνωστό location. Η κατανομή δειγματοληψίας για n παρατηρήσεις είναι:

(i). $x_i|\vartheta \overset{iid}{\sim} N\left(\cdot|\mu,\vartheta^{-1}\right)$ (για εκτίμηση του precision)

(ii). $x_i|\vartheta \overset{iid}{\sim} N\left(\cdot|\mu,\vartheta\right)$ (για εκτίμηση της διασποράς).

<u>**Για το μοντέλο (i):**</u> έχουμε δειγματοληπτική κατανομή

$$\pi\left(x_i|\vartheta\right)=\sqrt{\frac{\vartheta}{2\pi}}\exp\left[-\frac{\vartheta}{2}\left(x_i-\mu\right)^2\right],$$

και EF παραμέτρους: $h\left(x_i\right)=\sqrt{\dfrac{1}{2\pi}},\ g\left(\vartheta\right)=\vartheta^{1/2},\ c\left(\vartheta\right)=-\dfrac{\vartheta}{2},\ t\left(x_i\right)=\left(x_i-\mu\right)^2$.

Η πιθανοφάνεια θα είναι

$$\pi\left(x|\vartheta\right)\overset{\vartheta}{\propto}\vartheta^{n/2}\exp\left(-\frac{\vartheta}{2}nS^2\right)\propto Ga\left(\vartheta|\frac{n}{2}+1,\frac{nS^2}{2}\right),$$

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

Ο συζυγής prior είναι

$$\pi_{NCP}\left(\vartheta\right)\propto g\left(\vartheta\right)^d\exp\left\{bc\left(\vartheta\right)\right\}=\vartheta^{d/2}\exp\left(-\frac{b\vartheta}{2}\right)\propto Ga\left(\vartheta|\frac{d}{2}+1,\frac{b}{2}\right).$$

Θέτοντας λοιπόν prior $\pi\left(\vartheta\right)=Ga\left(\vartheta\,|\,a,b\right)$, η posterior γίνεται

$$\pi\left(\vartheta|x\right)\propto\vartheta^{a-1}\exp\left(-b\vartheta\right)\times\vartheta^{n/2}\exp\left(-\frac{\vartheta}{2}nS^{2}\right)\propto Ga\left(\vartheta\,|\,a+\frac{n}{2},b+\frac{nS^{2}}{2}\right).$$

Ο εκτιμητής κατά Bayes ως προς τετραγωνική συνάρτηση απώλειας είναι

$$\hat{\vartheta}_{Bayes}=\mathbb{E}\left(\vartheta|x\right)=\frac{a+\frac{n}{2}}{b+\frac{nS^{2}}{2}}=\left(\frac{a}{b}\right)\left(\frac{b}{b+\frac{nS^{2}}{2}}\right)+\left(\frac{1}{S^{2}}\right)\left(1-\frac{b}{b+\frac{nS^{2}}{2}}\right)$$

$$=\mathbb{E}\left(\vartheta\right)\left(\frac{b}{b+\frac{nS^{2}}{2}}\right)+\hat{\vartheta}_{MLE}\left(1-\frac{b}{b+\frac{nS^{2}}{2}}\right).$$

Παρατηρούμε ότι σε αυτήν την περίπτωση, το ποσοστό της κλασικής εκτίμησης $\hat{\vartheta}_{MLE}$, μέσα στην εκτίμηση κατά Bayes $\hat{\vartheta}_{Bayes}$, είναι συνάρτηση της $t_{Suff}\left(x\right)$. Εξαρτάται δηλαδή από τις παρατηρήσεις x.

Για την prior predictive και μία παρατήρηση x, τέτοια ώστε $\pi\left(x\,|\,\vartheta\right)=N\left(x\,|\,\mu,\vartheta^{-1}\right)$ και $\pi\left(\vartheta\right)=Ga\left(\vartheta\,|\,a,b\right)$ έχουμε:

$$\pi\left(x\right)=\int_{\vartheta>0}N\left(x\,|\,\mu,\vartheta^{-1}\right)Ga\left(\vartheta\,|\,a,b\right)d\vartheta=\int_{\vartheta=0}^{\infty}\sqrt{\frac{\vartheta}{2\pi}}\exp\left\{-\frac{\vartheta}{2}\left(x-\mu\right)^{2}\right\}\frac{b^{a}}{\Gamma\left(a\right)}\vartheta^{a-1}e^{-b\vartheta}d\vartheta$$

$$=\frac{b^{a}}{\Gamma\left(a\right)}\frac{1}{\sqrt{2\pi}}\int_{\vartheta=0}^{\infty}\vartheta^{a+1/2-1}\exp\left\{-\left(b+\frac{1}{2}\left(x-\mu\right)^{2}\right)\vartheta\right\}d\vartheta$$

$$=\frac{\Gamma\left(a+1/2\right)}{\Gamma\left(a\right)\Gamma\left(1/2\right)}\frac{b^{a}}{\sqrt{2}}\left(b+\frac{1}{2}\left(x-\mu\right)^{2}\right)^{-\left(a+1/2\right)}=\frac{\Gamma\left(a+1/2\right)}{\Gamma\left(a\right)\Gamma\left(1/2\right)}\frac{b^{a}}{\sqrt{2}}b^{-\left(a+1/2\right)}\left(1+\frac{1}{2b}\left(x-\mu\right)^{2}\right)^{-\left(a+1/2\right)}$$

$$=\frac{\Gamma\left(a+1/2\right)}{\Gamma\left(a\right)\Gamma\left(1/2\right)}\left(\frac{1}{2b}\right)^{1/2}\left(1+\frac{1}{2b}\left(x-\mu\right)^{2}\right)^{-\left(a+1/2\right)}$$, όπου $\Gamma\left(1/2\right)=\sqrt{\pi}$.

Λέμε ότι η τ.μ x ακολουθεί την student με μέσο μ, scale λ που είναι ανάλογο του precision και ν βαθμούς ελευθερίας, όταν $x\sim St\left(\mu,\lambda^{-1},\nu\right)$ με πυκνότητα:

$$St\left(x \mid \mu, \lambda^{-1}, \nu\right)=\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\Gamma\left(\frac{1}{2}\right)}\left(\frac{\lambda}{\nu}\right)^{1/2}\left(1+\frac{\lambda}{\nu}(x-\mu)^2\right)^{-\frac{\nu+1}{2}}.$$

Συγκρίνοντας τις δύο τελευταίες κατανομές, έχουμε ότι $\frac{\lambda}{\nu}=\frac{1}{2b}$ και $a+\frac{1}{2}=\frac{\nu+1}{2}$.

<u>Αναγνωρίζουμε λοιπόν ότι</u> η prior predictive $\pi(x)$ είναι student με μέσο μ, $\lambda=\frac{a}{b}$ και $\nu=2a$, ή ότι

$$\pi(x)=St\left(x \mid \mu, \frac{b}{a}, 2a\right).$$

Για την μέση τιμή και διασπορά της τ.μ. x έχουμε:

$$\mathbb{E}(x)=\mathbb{E}\left(\mathbb{E}(x \mid \vartheta)\right)=\mathbb{E}(\mu)=\mu$$

$$\mathbb{V}(x)=\mathbb{V}\left(\mathbb{E}(x \mid \vartheta)\right)+\mathbb{E}\left(\mathbb{V}(x \mid \vartheta)\right)=\mathbb{V}(\mu)+\mathbb{E}\left(\vartheta^{-1}\right)=\int_{\vartheta>0}\vartheta^{-1}Ga(\vartheta \mid a,b)\,d\vartheta$$

$$=\frac{b\Gamma(a-1)}{\Gamma(a)}=\frac{b}{a-1},\ a>1,$$

ή ως προς τις αρχικές παραμέτρους λ και ν

$$\mathbb{V}(x)=\frac{b}{a-1}=\frac{2a}{\frac{a}{b}(2a-2)}=\frac{\nu}{\lambda(\nu-2)},\ \nu=2a>2 \text{ και } \lambda \propto \frac{1}{\mathbb{V}(x)}.$$

Η posterior predictive για future (unobserved) observation y μετά την παρατήρηση της x είναι

$$\pi(y \mid x)=\int_{\vartheta>0}N\left(y \mid \mu, \vartheta^{-1}\right)Ga\left(\vartheta \mid a+\frac{n}{2}, b+\frac{nS^2}{2}\right)d\vartheta=St\left(y \middle| \mu, \frac{b+\frac{nS^2}{2}}{a+\frac{n}{2}}, 2\left(a+\frac{n}{2}\right)\right)$$

με $\mathbb{E}(y \mid x)=\mathbb{E}\left(\mathbb{E}(y \mid \vartheta) \mid x\right)=\mathbb{E}(\mu \mid x)=\mu$ και $\mathbb{V}(y \mid x)=\dfrac{b+\frac{nS^2}{2}}{a+\frac{n}{2}-1}$, $a+\frac{n}{2}>1$.

Θυμηθείτε ότι επειδή η τ.μ. $(y\,|\,x)$ προέρχεται από μίξη έχουμε

$$\mathbb{E}(y\,|\,x)=\int_{y\in\mathbb{R}} y\pi(y\,|\,x)dy=\int_{y\in\mathbb{R}} y\int_{\vartheta\in\mathbb{R}^+}\pi(y\,|\,\vartheta)\pi(\vartheta\,|\,x)d\vartheta dy$$

$$=\int_{\vartheta\in\mathbb{R}^+}\left(\int_{y\in\mathbb{R}} y\pi(y\,|\,\vartheta)d\,y\right)\pi(\vartheta\,|\,x)d\vartheta=\int_{\vartheta\in\mathbb{R}^+}\mathbb{E}(y\,|\,\vartheta)\pi(\vartheta\,|\,x)d\vartheta=\mathbb{E}\big(\mathbb{E}(y\,|\,\vartheta)\,|\,x\big).$$

<u>Άσκηση</u>

Λέμε ότι η τ.μ $x\sim St\left(\mu,\lambda^{-1},\nu\right)$ ακολουθεί την student με μέσο μ, λ (ανάλογο του precision) και ν βαθμούς ελευθερίας, όταν

$$St\left(x\,|\,\mu,\lambda^{-1},\nu\right)=\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\Gamma\left(\frac{1}{2}\right)}\left(\frac{\lambda}{\nu}\right)^{1/2}\left(1+\frac{\lambda}{\nu}(x-\mu)^2\right)^{-\frac{\nu+1}{2}}.$$

Να αποδειχτούν τα παρακάτω:

1. (student characterization I) Για γνωστά μ και λ έχουμε:
   $\left\{\begin{array}{c}\vartheta\sim Ga\left(\frac{\nu}{2},\frac{\nu}{2}\right)\\ x\,|\,\vartheta\sim N\left(\mu,\frac{1}{\lambda\vartheta}\right)\end{array}\right\}\Rightarrow$ Marginally $x\sim St\left(\mu,\lambda^{-1},\nu\right)$, δηλαδή η student γεννιέται

   από τη μίξη της οικογένειας κατανομών $N\left(x\,|\,\mu,\frac{1}{\lambda\vartheta}\right)$ ως προς το gamma

   μέτρο $\Pi(d\vartheta)=Ga\left(\vartheta\,|\,\frac{\nu}{2},\frac{\nu}{2}\right)d\vartheta$.

2. H normal $N\left(x\,|\,\mu,\lambda^{-1}\right)$ είναι student με ∞ βαθμούς ελευθερίας και η Cauchy είναι student με 1 βαθμό ελευθερίας.

3. Για τις ροπές της student έχουμε: $\mathbb{E}\left(x^k\right)=\left\{\begin{array}{cc} defined & k<\nu\\ \infty-\infty & k=odd\geq\nu\\ \infty & k=even\geq\nu\end{array}\right\}$.

4. (Student Characterization II) εάν x και y ανεξάρτητες τ.μ. με $x \sim N(0,1)$ και $y \sim \mathcal{X}_\nu^2 = Ga\left(\frac{\nu}{2}, \frac{1}{2}\right)$, τότε $z = \frac{x}{\sqrt{y/\nu}} \sim St(0,1,\nu)$ δηλαδή standard student με ν βαθμούς ελευθερίας. Θέτοντας $u = \frac{z}{\sqrt{\lambda}} + \mu$ παίρνουμε $u \sim St\left(\mu, \lambda^{-1}, \nu\right)$.

[Θα δείξουμε μόνο το 2.]

Θέλουμε να δείξουμε ότι $\lim_{\nu\to\infty} St\left(x \mid \mu, \lambda^{-1}, \nu\right) = N\left(x \mid \mu, \lambda^{-1}\right)$.

Αρχίζοντας από την ασυμπτοτική προσέγγιση του Stirling που ισχύει για κάθε θετικό «μεγάλο» πραγματικό ν έχουμε

$$\Gamma(\nu+1) \approx \sqrt{2\pi\nu}\,\nu^\nu e^{-\nu} \Leftrightarrow \Gamma(\nu) \approx \sqrt{2\pi}\,(\nu-1)^{\nu-1/2} e^{-\nu+1} \approx \sqrt{2\pi}\,\nu^{\nu-1/2} e^{-\nu+1},$$

που δίνει

$$\frac{\Gamma(\nu+1/2)}{\nu^{1/2}\,\Gamma(\nu)} = \left(1+\frac{1}{2\nu}\right)^\nu e^{-1/2} \to 1 \text{ όταν } \nu \to \infty .$$

Αντικαθιστώντας όπου ν το $\nu/2$ παίρνουμε $\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\left(\frac{\nu}{2}\right)^{1/2}\Gamma\left(\frac{\nu}{2}\right)} \to 1$ όταν $\nu \to \infty$, έτσι

$$\lim_{\nu\to\infty} St(x \mid \mu, \lambda, \nu) = \lim_{\nu\to\infty} \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \left(\frac{\lambda}{\pi\nu}\right)^{1/2} \left(1+\frac{\lambda(x-\mu)^2}{\nu}\right)^{-\frac{\nu+1}{2}}$$

$$= \sqrt{\frac{\lambda}{2\pi}} \left\{ \lim_{\nu\to\infty} \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\left(\frac{\nu}{2}\right)^{1/2}\Gamma\left(\frac{\nu}{2}\right)} \right\} \left\{ \lim_{\nu\to\infty} \left(1+\frac{\lambda(x-\mu)^2}{\nu}\right)^{\nu+1} \right\}^{-1/2} = N\left(x \mid \mu, \lambda^{-1}\right).$$

Για την Cauchy έχουμε:

$$St\left(x \mid \mu, \lambda^{-1}, 1\right) = \frac{\Gamma(1)}{\Gamma\left(\frac{1}{2}\right)^2}(\lambda)^{1/2}\left(1+\lambda(x-\mu)^2\right)^{-1} = \frac{1}{\pi}\frac{\sqrt{\lambda}}{1+\lambda(x-\mu)^2}$$

Θέτοντας $\beta^{-1} = \sqrt{\lambda}$ παίρνουμε $St\left(x \mid \mu, \lambda^{-1}, 1\right) = \frac{1}{\pi}\frac{\beta}{\beta^2+(x-\mu)^2} = Ca(x \mid \mu, \beta)$.

<u>**Για το μοντέλο (ii)**</u> έχουμε δειγματοληπτική κατανομή:

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

και EF παραμέτρους: $h(x_i) = \sqrt{\frac{1}{2\pi}},\ g(\vartheta) = \vartheta^{-1/2},\ c(\vartheta) = -\frac{1}{2\vartheta},\ t(x_i) = (x_i-\mu)^2$.

Λέμε ότι το $\vartheta \sim Ig(a,b)$ ακολουθεί την inverse gamma με παραμέτρους a και b, όταν $\vartheta^{-1} \sim Ga(a,b)$. Εύκολα δείχνεται ότι ισχύει $Ig(\vartheta \mid a,b) = \frac{b^a}{\Gamma(a)}\vartheta^{-(a+1)}e^{-\frac{b}{\vartheta}},\ \vartheta > 0$.

Η πιθανοφάνεια είναι $\pi(x \mid \vartheta) \overset{\vartheta}{\propto} \vartheta^{-n/2}\exp\left(-\frac{1}{2\vartheta}nS^2\right) \propto Ig\left(\vartheta \mid \frac{n}{2}-1, \frac{nS^2}{2}\right)$,

με επαρκή στατιστική $t_{Suff}(x) = nS^2 = \sum_{i=1}^{n}(x_i-\mu)^2$.

Ο συζυγής prior είναι

$$\pi_{NCP}(\vartheta) \propto g(\vartheta)^d \exp\left[bc(\vartheta)\right] = \vartheta^{-d/2}\exp\left(-\frac{b}{2\vartheta}\right) \propto Ig\left(\vartheta \mid \frac{d}{2}-1, \frac{b}{2}\right).$$

Θέτοντας prior $\pi(\vartheta) = Ig(\vartheta \mid a,b) \propto \vartheta^{-(a+1)}e^{-\frac{b}{\vartheta}}$, η posterior είναι:

$$\pi(\vartheta \mid x) \propto \vartheta^{-(a+1)}e^{-\frac{b}{\vartheta}} \times \vartheta^{-n/2}\exp\left(-\frac{nS^2}{2\vartheta}\right) \propto Ig\left(\vartheta \mid a+\frac{n}{2}, b+\frac{nS^2}{2}\right).$$

Δεν είναι δύσκολα να δούμε ότι

$\mathbb{E}(\vartheta) = \frac{b}{a-1},\ a > 1$ και $\hat{\vartheta}_{MLE} = S^2$ έτσι

$$\hat{\vartheta}_{BAYES}=\mathbb{E}\left(\vartheta|x\right)=\frac{b+\frac{nS^2}{2}}{a+\frac{n}{2}-1}=\left(\frac{b}{a-1}\right)\left(\frac{a-1}{a+\frac{n}{2}-1}\right)+S^2\left(1-\frac{a-1}{a+\frac{n}{2}-1}\right).$$

Δηλαδή με την επαναπαραμετροποίηση $\vartheta\mapsto\varphi=\vartheta^{-1}=$ variance το ποσοστό της κλασικής εκτίμησης $\hat{\vartheta}_{MLE}$ στην εκτίμηση κατά Bayes $\hat{\vartheta}_{BAYES}$ είναι ανεξάρτητο του δείγματος x.

Για την prior predictive έχουμε: $\pi\left(x\right)=\int\limits_{\vartheta>0}N\left(x\,|\,\mu,\vartheta\right)Ig\left(\vartheta\,|\,a,b\right)d\vartheta$. Κάνοντας τον μετασχηματισμό $\varphi=\vartheta^{-1}$ παίρνουμε:

$$\pi\left(x\right)=-\int\limits_{\varphi=\infty}^{0}N\left(x\,|\,\mu,\varphi^{-1}\right)\left\{Ig\left(\varphi^{-1}\,|\,a,b\right)\varphi^{-2}\right\}d\varphi=\int\limits_{\varphi>0}N\left(x\,|\,\mu,\varphi^{-1}\right)Ga\left(\varphi\,|\,a,b\right)d\varphi$$

$$=\frac{\Gamma\left(a+1/2\right)}{\Gamma\left(a\right)\Gamma\left(1/2\right)}\left(\frac{1}{2b}\right)^{1/2}\left(1+\frac{1}{2b}\left(x-\mu\right)^2\right)^{-\left(a+1/2\right)}=St\left(x\,|\,\mu,\frac{b}{a},2a\right).$$

Ομοίως για την posterior predictive έχουμε:

$$\pi\left(y\,|\,x\right)=\int\limits_{\vartheta>0}N\left(y\,|\,\mu,\vartheta\right)Ig\left(\vartheta\,|\,a+\frac{n}{2},b+\frac{nS^2}{2}\right)d\vartheta$$

$$=\int\limits_{\varphi>0}N\left(x\,|\,\mu,\varphi^{-1}\right)Ga\left(\varphi\,|\,a+\frac{n}{2},b+\frac{nS^2}{2}\right)d\varphi=St\left(y\middle|\mu,\frac{b+\frac{nS^2}{2}}{a+\frac{n}{2}},2\left(a+\frac{n}{2}\right)\right).$$

<u>Άσκηση</u>

Να αποδειχτούν τα παρακάτω:

1. Εάν $z_i\overset{iid}{\sim}N\left(0,1\right)\ 1\le i\le\nu$ τότε $z=\sum_{i=1}^{\nu}z_i^2\sim\mathcal{X}_\nu^2=Ga\left(\frac{\nu}{2},\frac{1}{2}\right)$ δηλαδή το z ακολουθεί την chi – squared με ν βαθμούς ελευθερίας.
2. Εάν $x\sim Ga\left(a,b\right)$ τότε το $y=x^{-1}\sim Ig\left(a,b\right)$ ακολουθεί inverse gamma με

$$Ig\left(y\,|\,a,b\right)=\frac{b^a}{\Gamma(a)}y^{-(a+1)}e^{-\frac{b}{y}},\ y>0$$

$$\mathbb{E}\left(y^k\right)=\frac{b^k}{(a-k)_{(k)}}\ \text{όταν}\ a>k\ \text{και}\ Var(y)=\frac{b^2}{(a-1)^2(a-2)}\ \text{όταν}\ a>2\,.$$

3. Εάν $y\sim Ig(a,b)$ τότε $z=sy\sim Ig(a,sb)$.

4. Η scaled inverse chi – squared κατανομή με v_0 βαθμούς ελευθερίας και rate σ_0^2. ορίζεται σαν

   $$Inv\mathcal{X}^2\left(v_0,\sigma_0^2\right)=\mathcal{X}^{-2}\left(v_0,\sigma_0^2\right)=Ig\left(\frac{v_0}{2},\frac{v_0\sigma_0^2}{2}\right),$$

   δηλαδή είναι η $s=v_0\sigma_0^2$ – scaled inverse gamma κατανομή για $a=v_0/2$, και $b=1/2$. Δείξτε ότι $\mathcal{X}^{-2}\left(\vartheta|v_0,\sigma_0^2\right)\propto\vartheta^{-\left(\frac{v_0}{2}+1\right)}\exp\left(-\frac{v_0\sigma_0^2}{2\vartheta}\right)$.

1. Εάν $z_i\overset{iid}{\sim}N(0,1)$ και θέσουμε $y_i=T(z_i)=z_i^2$

$$\pi_{Y_i}(y_i)=\sum_{y_i=z_i^2}\pi_{Z_i}(z_i)\left|Jac\left(T^{-1}(z_i)\right)\right|$$

$$=N\left(\sqrt{y_i}\,|\,0,1\right)\left|\frac{d}{dy_i}\left(\sqrt{y_i}\right)\right|+N\left(-\sqrt{y_i}\,|\,0,1\right)\left|\frac{d}{dy_i}\left(-\sqrt{y_i}\right)\right|$$

$$=N\left(\sqrt{y_i}\,|\,0,1\right)\frac{1}{\sqrt{y_i}}=\frac{1}{\sqrt{2\pi y_i}}e^{-y_i/2}=\frac{(1/2)^{1/2}}{\Gamma(1/2)}y^{1/2-1}e^{-y_i/2}=Ga\left(y_i\,|\,\frac{1}{2},\frac{1}{2}\right),\ y>0\,.$$

Όπου $\Gamma(1/2)=\sqrt{\pi}$ επειδή

$$\Gamma\left(\frac{1}{2}\right)=\int_{\mathbb{R}^+}x^{-1/2}e^{-x}dx=\int_{\mathbb{R}^+}\left(\frac{u^2}{2}\right)^{-1/2}e^{-\frac{u^2}{2}}(udu)=\sqrt{2}\int_{\mathbb{R}^+}e^{-\frac{u^2}{2}}du$$

$$=\frac{1}{\sqrt{2}}\int_{\mathbb{R}}e^{-\frac{u^2}{2}}du=\sqrt{\pi}\int_{\mathbb{R}}\frac{1}{\sqrt{2\pi}}e^{-\frac{u^2}{2}}du=\sqrt{\pi}\int_{\mathbb{R}}N(u\,|\,0,1)du=\sqrt{\pi}\ .$$

Όμως $z_i^2 \overset{iid}{\sim} Ga\left(\frac{1}{2},\frac{1}{2}\right)$ $1 \le i \le \nu$, δίνει $\sum_{i=1}^{\nu} z_i^2 \sim \chi_\nu^2 = Ga\left(\frac{\nu}{2},\frac{1}{2}\right)$.

2. $\pi(y) = Ga\left(y^{-1} \mid p, q\right)\left(y^{-2}\right) = \frac{q^p}{\Gamma(p)} y^{-(p+1)} e^{-\frac{q}{y}}, \; y > 0$

$$\mathbb{E}\left(y^k\right) = \int_{\mathbb{R}^+} y^k Ig\left(y \mid p, q\right) dy = \int_{\mathbb{R}^+} y^k \left\{ \frac{q^p}{\Gamma(p)} \frac{1}{y^{p+1}} e^{-\frac{q}{y}} dy \right\}$$

$$= \frac{q^p}{\Gamma(p)} \int_{\mathbb{R}^+} \left(\frac{1}{y}\right)^{p-k+1} e^{-\frac{q}{y}} dy = \frac{q^p}{\Gamma(p)} \int_{\mathbb{R}^+} u^{p-k+1} e^{-qu} \frac{du}{u^2} = \frac{q^p}{\Gamma(p)} \int_{\mathbb{R}^+} u^{p-k-1} e^{-qu} du$$

$$= \frac{q^p}{\Gamma(p)} \frac{1}{q^{p-k}} \int_{\mathbb{R}^+} v^{(p-k)-1} e^{-v} dv = \frac{q^k \, \Gamma(p-k)}{\Gamma(p)}$$, όταν $p > k$,

δηλαδή για όλα τα $k \in \mathbb{N}$ και $p > 0$ για τα οποία ισχύει $\left|\Gamma(p-k)\right| < \infty$. Εάν $p > k$, και επειδή

$$\Gamma(p+\tau) = (p+\tau-1)\Gamma(p+\tau-1) = \cdots = (p)(p+1)\cdots(p+\tau-1)\Gamma(p) = (p)_{(\tau)}\,\Gamma(p),$$

έχουμε

$$\mathbb{E}\left(y^k\right) = \frac{q^k \Gamma(p-k)}{\Gamma(p)} = \frac{q^k \Gamma(p-k)}{\Gamma\left((p-k)+k\right)} = \frac{q^k \Gamma(p-k)}{(p-k)_{(k)}\,\Gamma(p-k)} = \frac{q^k}{(p-k)_{(k)}}.$$

Έτσι

$$\mathbb{E}(y) = \frac{q\Gamma(p-1)}{\Gamma(p)} = \frac{q}{p-1}, \; p > 1 \text{ και } Var(y) = \frac{q^2}{(p-1)^2 (p-2)}, \; p > 2.$$

3. $z = sy, \; s > 0$ και $y \sim IG(p,q)$, τότε έχουμε:

$$f\left(z \mid p, q, s\right) = s^{-1} IG\left(s^{-1} z \mid p, q\right) = s^{-1} \frac{q^p}{\Gamma(p)} \frac{s^{p+1}}{z^{p+1}} e^{-\frac{sq}{z}}, \; z > 0$$

$= \frac{(sq)^p}{\Gamma(p)} z^{-(p+1)} e^{-\frac{(sq)}{z}}$, που δίνει $z = sy \sim Ig(p, sq)$.

4. Παίρνοντας υπ’ όψιν ότι $Ig\left(z|\,p,sq\right)=\dfrac{\left(sq\right)^{p}}{\Gamma\left(p\right)}z^{-\left(p+1\right)}e^{-\frac{\left(sq\right)}{z}}$ και ότι

$\mathcal{X}^{-2}\left(v_0,\sigma_0^2\right)=Ig\left(\dfrac{v_0}{2},\dfrac{v_0\sigma_0^2}{2}\right)$, έχουμε

$$\mathcal{X}^{-2}\left(\vartheta\,|\,v_0,\sigma_0^2\right)=\frac{1}{\Gamma\left(v_0/2\right)}\left(\frac{v_0\sigma_0^2}{2}\right)^{\frac{v_0}{2}}\vartheta^{-\left(\frac{v_0}{2}+1\right)}\exp\left(-\frac{v_0\sigma_0^2}{2\vartheta}\right)\propto\vartheta^{-\left(\frac{v_0}{2}+1\right)}\exp\left(-\frac{v_0\sigma_0^2}{2\vartheta}\right).$$

Με τον νέο συμβολισμό το μοντέλο (ii)

$$\pi\left(x_i\,|\,\vartheta\right)=N\left(x_i\,|\,\mu,\vartheta\right)=\frac{1}{\sqrt{2\pi\vartheta}}\exp\left[-\frac{1}{2\vartheta}\left(x_i-\mu\right)^2\right]$$

$$\pi\left(x|\vartheta\right)\overset{\vartheta}{\propto}\vartheta^{-n/2}\exp\left(-\frac{1}{2\vartheta}nS^2\right)\propto Ig\left(\vartheta\,|\,\frac{n}{2}-1,\frac{nS^2}{2}\right)$$

$$\hat{\vartheta}_{BAYES}=\mathbb{E}\left(\vartheta|x\right)=\left(\frac{b}{a-1}\right)\left(\frac{a-1}{a+\frac{n}{2}-1}\right)+S^2\left(1-\frac{a-1}{a+\frac{n}{2}-1}\right)$$

$$\pi\left(\vartheta\right)\propto\vartheta^{-\left(a+1\right)}e^{-\frac{b}{\vartheta}}\propto Ig\left(\vartheta\,|\,a,b\right),\;\pi\left(\vartheta|x\right)=Ig\left(\vartheta\,|\,a+\frac{n}{2},b+\frac{nS^2}{2}\right)$$

$$\pi\left(x\right)=St\left(x\,|\,\mu,\frac{a}{b},2a\right),\;\pi\left(y\,|\,x\right)=St\left(y|\,\mu,\frac{a+\frac{n}{2}}{b+\frac{nS^2}{2}},2\left(a+\frac{n}{2}\right)\right)$$

επαναδιατυπώνεται στην πιο standard μορφή για $a=\dfrac{v_0}{2},b=\dfrac{v_0\sigma_0^2}{2}$

$$\pi\left(\vartheta\right)=\mathcal{X}^{-2}\left(\vartheta\,|\,v_0,\sigma_0^2\right)=Ig\left(\vartheta\,|\,\frac{v_0}{2},\frac{v_0\sigma_0^2}{2}\right)$$

$$\pi\left(\vartheta|x\right)=\mathcal{X}^{-2}\left(\vartheta \mid \underbrace{v_0+n}_{v_n}, \underbrace{\frac{v_0\sigma_0^2+nS^2}{v_0+n}}_{\sigma_n^2}\right)=Ig\left(\vartheta \mid \frac{v_0+n}{2}, \frac{v_0\sigma_0^2+nS^2}{2}\right)$$

$$\hat{\vartheta}_{BAYES}=\mathbb{E}\left(\vartheta|x\right)=\left(\frac{v_0\sigma_0^2}{v_0-2}\right)\left(\frac{v_0-2}{v_0+n-2}\right)+\left(S^2\right)\left(1-\frac{v_0-2}{v_0+n-2}\right)$$

$$\pi\left(x\right)=St\left(x \mid \mu,\sigma_0^2,v_0\right),\ \pi\left(y \mid x\right)=St\left(y|\mu,\sigma_n^2,v_n\right).$$

Επίσης η αναπαράσταση (characterization 1)της student σαν μίξη, τώρα γίνεται προφανείς, εφόσον από την εξίσωση

$$\pi\left(x\right)=\int_{\vartheta>0} N\left(x \mid \mu,\vartheta^{-1}\right)Ga\left(\vartheta \mid a,b\right)d\vartheta=St\left(x \mid \mu,\frac{b}{a},2a\right),$$

και θέτοντας $a=\frac{v}{2}, b=\frac{v\sigma^2}{2}$ έχουμε

$$\pi\left(x\right)=\int_{\vartheta>0} N\left(x \mid \mu,\vartheta^{-1}\right)Ga\left(\vartheta \mid \frac{v}{2},\frac{v\sigma^2}{2}\right)d\vartheta=St\left(x \mid \mu,\sigma^2,v\right)$$

αλλά $Ga\left(\vartheta \mid \alpha,\sigma^2\beta\right)d\vartheta=Ga\left(\sigma^2\vartheta \mid \alpha,\beta\right)d\left(\sigma^2\vartheta\right)$, που δίνει για $\varphi=\sigma^2\vartheta$,

$$\pi\left(x\right)=\int_{\varphi>0} N\left(x \mid \mu,\sigma^2\varphi^{-1}\right)Ga\left(\varphi \mid \frac{v}{2},\frac{v}{2}\right)d\varphi=St\left(x \mid \mu,\sigma^2,v\right).$$

Απουσία a-priori πεποιθήσεων

Εάν θεωρήσουμε ότι η prior δίνει πληροφορία ισοδύναμη με v_0 «prior παρατηρήσεις» η συνολική απουσία a-priori πεποιθήσεων ισοδυναμεί με αποδοχή της conjugate prior $\pi\left(\vartheta\right)=Inv\chi^2\left(\vartheta|0,\sigma_0^2\right)\ \propto\ \frac{1}{\vartheta}$ *ή αλλιώς την «zero prior observation» a-priori κατανομή, που είναι improper εφόσον* $\int_{\mathbb{R}^+}\frac{d\vartheta}{\vartheta}=\infty$. Τότε

$$\pi\left(\vartheta|x\right)\ \propto\ \left\{\vartheta^{-1}\right\}\times\left\{\vartheta^{-n/2}\exp\left(-\frac{nS^2}{2\vartheta}\right)\right\}=\vartheta^{-\left(\frac{n}{2}+1\right)}\exp\left(-\frac{nS^2}{2\vartheta}\right)\propto\mathcal{X}^{-2}\left(\vartheta|\frac{n}{2},\frac{nS^2}{2}\right),$$

που δίνει

$$\mathbb{E}\left(\vartheta|x\right)=\frac{\frac{nS^2}{2}}{\frac{n}{2}-1}\approx S^2=\hat{\vartheta}_{MLE}$$

<u>Άσκηση</u>

Δείξτε ότι $\pi\left(x\right)=\int\limits_{\mathbb{R}^+} N\left(x|\ \mu,\vartheta\right)\mathcal{X}^{-2}\left(\vartheta|v_0,\sigma_0^2\right)d\vartheta=St\left(x|\mu,\sigma_0^2,v_0\right)$.

Πράγματι

$$\pi\left(x\right)=\int\limits_{\mathbb{R}^+} N\left(x|\ \mu,\vartheta\right)\mathcal{X}^{-2}\left(\vartheta|v_0,\sigma_0^2\right)d\vartheta=\int\limits_{\mathbb{R}^+} N\left(x|\ \mu,\vartheta\right)Ig\left(\vartheta|\frac{v_0}{2},\frac{v_0\sigma_0^2}{2}\right)d\vartheta$$

$$=\frac{1}{\Gamma\left(v_0/2\right)}\left(\frac{v_0\sigma_0^2}{2}\right)^{\frac{v_0}{2}}\left(\frac{1}{2\pi}\right)^{\frac{1}{2}}\int_{\mathbb{R}^+}\left(\frac{1}{\vartheta}\right)^{\frac{v_0+1}{2}+1}\exp\left(-\frac{v_0\sigma_0^2+\left(x-\mu\right)^2}{2\vartheta}\right)d\vartheta$$

θέτοντας $u=\frac{1}{\vartheta}$ παίρνουμε

$$\pi\left(x\right)=\frac{1}{\Gamma\left(v_0/2\right)}\left(\frac{v_0\sigma_0^2}{2}\right)^{\frac{v_0}{2}}\left(\frac{1}{2\pi}\right)^{\frac{1}{2}}\int_{\mathbb{R}^+}u^{\frac{v_0+1}{2}+1}\exp\left(-\frac{v_0\sigma_0^2+\left(x-\mu\right)^2}{2}u\right)u^{-2}du$$

$$=\frac{1}{\Gamma\left(v_0/2\right)}\left(\frac{v_0\sigma_0^2}{2}\right)^{\frac{v_0}{2}}\left(\frac{1}{2\pi}\right)^{\frac{1}{2}}\int_{\mathbb{R}^+}u^{\frac{v_0+1}{2}-1}\exp\left(-\frac{v_0\sigma_0^2+\left(x-\mu\right)^2}{2}u\right)du$$

θέτοντας $\tau=\frac{v_0\sigma_0^2+\left(x-\mu\right)^2}{2}u$ παίρνουμε

$$\pi\left(x\right)=\frac{1}{\Gamma\left(v_0/2\right)}\left(\frac{v_0\sigma_0^2}{2}\right)^{\frac{v_0}{2}}\left(\frac{1}{2\pi}\right)^{\frac{1}{2}}\int_{\mathbb{R}^+}\left(\frac{2\tau}{v_0\sigma_0^2+\left(x-\mu\right)^2}\right)^{\frac{v_0+1}{2}-1}e^{-\tau}\frac{2d\tau}{v_0\sigma_0^2+\left(x-\mu\right)^2}$$

$$=\frac{1}{\Gamma\left(v_0/2\right)}\left(\frac{v_0\sigma_0^2}{2}\right)^{\frac{v_0}{2}}\left(\frac{1}{2\pi}\right)^{\frac{1}{2}}\left(\frac{2}{v_0\sigma_0^2+\left(x-\mu\right)^2}\right)^{\frac{v_0+1}{2}}\int_{\mathbb{R}^+}\tau^{\frac{v_0+1}{2}-1}e^{-\tau}d\tau$$

$$= \frac{1}{\Gamma(v_0/2)} \left(v_0\sigma_0^2\right)^{\frac{v_0}{2}} \left(\frac{1}{\pi}\right)^{\frac{1}{2}} \left(\frac{1}{v_0\sigma_0^2 + (x-\mu)^2}\right)^{\frac{v_0+1}{2}} \Gamma\left(\frac{v_0+1}{2}\right)$$

$$= \frac{\Gamma\left(\frac{v_0+1}{2}\right)}{\Gamma\left(\frac{v_0}{2}\right)\sqrt{\pi}} \left(v_0\sigma_0^2\right)^{\frac{v_0}{2}} \left(v_0\sigma_0^2 + (x-\mu)^2\right)^{-\frac{v_0+1}{2}} = \frac{\Gamma\left(\frac{v_0+1}{2}\right)}{\Gamma\left(\frac{v_0}{2}\right)\sigma_0\sqrt{\pi v_0}} \left\{1 + \frac{1}{v_0}\left(\frac{x-\mu}{\sigma_0}\right)^2\right\}^{-\frac{v_0+1}{2}}$$

$$= St\left(x|\mu, \sigma_0^2, v_0\right).$$