آرین شُکر - ۴۰۰۲۳۴۹۴

2.19

likelihood function: $p(x) \sim N(\mu, \Sigma) = \dfrac{1}{(2\pi)^{l/2}\sqrt{|\Sigma|}} e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)}$

log - likelihood

$$\Rightarrow L(\theta) = \ln\left(\prod_{k=1}^{n} p(x_k|\mu, \Sigma)\right) = \sum_{k=1}^{n} \ln(p(x_k|\mu,\Sigma))$$

$$\sqrt{\frac{1}{(2\pi)^l \Sigma}}$$

$$\ln(p(x_k|\mu,\Sigma)) = -\frac{1}{2}(x_k-\mu)^T\Sigma^{-1}(x_k-\mu) + \ln\frac{1}{(2\pi)^{l/2}\sqrt{|\Sigma|}} =$$

$$-\frac{1}{2}(x_k-\mu)^T\Sigma^{-1}(x_k-\mu) - \frac{1}{2}\ln(2\pi)^l|\Sigma|$$

independent from index

$$\Rightarrow L(\theta) = \sum_{k=1}^{n} \ln(p(x_k|\mu,\Sigma)) = \sum_{k=1}^{n}\left[-\frac{1}{2}(x_k-\mu)^T\Sigma^{-1}(x_k-\mu) - \underbrace{\frac{1}{2}\ln(2\pi)^l|\Sigma|}_{}\right]$$

$$= -\frac{1}{2}\sum_{k=1}^{n}(x_k-\mu)^T\Sigma^{-1}(x_k-\mu) - \frac{n}{2}\ln(2\pi)^l|\Sigma|$$

$$\frac{\partial L(\theta)}{\partial\theta} = 0 \Rightarrow \begin{cases} \dfrac{\partial L(\theta)}{\partial\mu} = 0 & \text{I} \\[2ex] \dfrac{\partial L(\theta)}{\partial\Sigma} = 0 & \text{II} \end{cases}$$

I. $$\frac{\partial L(\theta)}{\partial\mu} = \partial\left(-\frac{1}{2}\sum_{k=1}^{n}(x_k-\mu)^T\Sigma^{-1}(x_k-\mu) - \frac{n}{2}\ln(2\pi)^l|\Sigma|\right)\Big/\partial(\mu)$$

we know from vector derivation that: $\dfrac{\partial x^TAx}{\partial x} = x^T(A+A^T)$, $\partial(x_k-\mu) = -\partial(\mu)$

$$\Rightarrow \frac{\partial L(\theta)}{\partial\mu} = -\frac{1}{2}\sum_{k=1}^{n} -(x_k-\mu)(\Sigma^{-1}+\Sigma^{-1T}) = -\frac{1}{2}\sum_{k=1}^{n} -2(x_k-\mu)(\Sigma^{-1}) = 0$$

$(\Sigma^{-1})^T = (\Sigma^T)^{-1} = \Sigma^{-1}$

covariance matrix is symmetrical

$$\Rightarrow \sum_{k=1}^{n}(x_k-\mu)(\Sigma^{-1}) = 0$$

$$\Rightarrow (x_1-\mu)(\Sigma^{-1}) + (x_2-\mu)(\Sigma^{-1}) + \cdots + (x_n-\mu)(\Sigma^{-1}) = 0 \xrightarrow[\text{by } \Sigma]{\text{right multiply}}$$

$$\Rightarrow (x_1-\mu) + (x_2-\mu) + \cdots + (x_n-\mu) = 0 \Rightarrow \sum_{k=1}^{n} x_k = n\mu \Rightarrow \boxed{\hat{\mu} = \frac{1}{n}\sum_{k=1}^{n} x_k}$$

(II). $\frac{\partial l(\theta)}{\partial (\Sigma)} = \partial\left(-\frac{1}{2}\sum_{k=1}^{n} (x_k-\mu)^T \Sigma^{-1} (x_k-\mu) - \frac{n}{2}\ln\left((2\pi)^d|\Sigma|\right)\right) \Big/ \partial(\Sigma)$

we need two derivatives:

1. $\frac{\partial a^t x b}{\partial x}$: $\frac{\partial a^t x b}{\partial x} = \partial\left(\sum_{i=1}^{n}\sum_{j=1}^{m} a_i x_{ij} b_j\right) \Big/ \partial(x_{ij}) = \left[a_i b_j\right] = a b^t$

2. $\frac{\partial \ln|\Sigma|}{\partial \Sigma} = (\Sigma^{-1})^t$ → found online

$= (\Sigma^{-1})^t = (\Sigma^t)^{-1} = \boxed{\Sigma^{-1}}$

Additionally we know that $|\Sigma| = \frac{1}{|\Sigma^{-1}|} \Rightarrow$ extrema of Σ and Σ^{-1} coincide so we derive for Σ^{-1}

$$l(\theta) = -\frac{1}{2}\sum_{k=1}^{n}(x_k-\mu)^T \Sigma^{-1}(x_k-\mu) - \frac{n}{2}\ln(2\pi)^d + \frac{n}{2}\ln|\Sigma^{-1}|$$

$$\rightarrow \frac{\partial l(\theta)}{\partial \Sigma^{-1}} = -\frac{1}{2}\sum(x_k-\mu)(x_k-\mu)^T + \frac{n}{2}\Sigma \Rightarrow \boxed{\hat{\Sigma} = \frac{1}{n}\sum(x_k-\mu)(x_k-\mu)^T}$$

(e.20)

$$E(\hat{\Sigma}) = E\left(\frac{1}{N-1}\sum_{k=1}^{N}(x_k-\hat{\mu})(x_k-\hat{\mu})^T\right) = \frac{1}{N-1} E\left(\sum_{k=1}^{N}\left[x_k x_k^T - x_k\hat{\mu}^T - \hat{\mu}x_k^T + \hat{\mu}\hat{\mu}^T\right]\right)$$

we know $\sum_{i=1}^{N} x_i = N\hat{\mu}$ and $\sum_{i=1}^{N} x_i^T = N\hat{\mu}^T$

$$\rightarrow E(\hat{\Sigma}) = \frac{1}{N-1} E\left(\left[\sum_{k=1}^{N}\|x_k\|^2\right] - N\hat{\mu}\hat{\mu}^T - N\hat{\mu}\hat{\mu}^T + N\hat{\mu}\hat{\mu}^T\right) = \frac{1}{N-1} E\left(\left[\sum_{k=1}^{N}\|x_k\|^2\right] - N\|\hat{\mu}\|^2\right)$$

$$= \frac{1}{N-1} E\left(\sum_{k=1}^{N}\|x_k\|^2\right) - \frac{N}{N-1} E(\|\hat{\mu}\|^2) = \frac{N}{N-1} E(\|x\|^2) - \frac{N}{N-1} E(\|\hat{\mu}\|^2)$$

$$N E\left(\frac{1}{N}\sum_{k=1}^{N}\|x_k\|^2\right) = N E(\|x\|^2)$$ (I)

from a second definition of covariance: $\mathrm{cov}(x,y) = E(xy) - E(x)E(y)$

$$\xrightarrow{x=y^T} \Sigma = E(\|x\|^2) - E(x)^2$$

$$\hat{\Sigma} = E(\|\hat{\mu}\|^2) - \underbrace{E(\hat{\mu})^2}_{\hat{\mu}^2} \rightarrow \frac{N}{N-1}\left(E(\|x\|^2) - E(\|\hat{\mu}\|^2)\right)$$

$$= \frac{N}{N-1}\left(\Sigma + \cancel{E(x)^2} - (\hat{\Sigma} + \cancel{E(\hat{\mu})^2})\right) = \frac{N}{N-1}(\Sigma - \hat{\Sigma})$$

$$= \frac{N}{N-1}\underbrace{\left(\frac{1}{N}\sum_{k=1}^{N}(x_k-\mu)(x_k-\mu)^T - \frac{1}{N-1}\sum_{k=1}^{N}(x_k-\hat{\mu})(x_k-\hat{\mu})\right)}_{\rightarrow \frac{N-1}{N}\Sigma?}$$

2.22.

$$p(x;\theta) = \theta^2 x \exp(-\theta x) u(x)$$

$$\rightarrow \text{log-likelihood, } L(\theta) = \ln \prod_{k=1}^{N} p(x_k;\theta) = \sum_{k=1}^{N} \ln(p(x_k;\theta))$$

$$= \sum_{k=1}^{N}\left[2\ln\theta + \ln x_k - \theta x_k + \ln u(x)\right] \Rightarrow \frac{\partial L(\theta)}{\partial(\theta)} = 0$$

$$\Rightarrow \sum_{k=1}^{N}\left[\frac{2}{\theta} + 0 - x_k + 0\right] = 0 \Rightarrow N\cdot\frac{2}{\theta} - \sum_{k=1}^{N} x_k = 0 \Rightarrow \boxed{\hat{\theta} = \frac{2N}{\sum_{k=1}^{N} x_k}}$$

2.30. Maximize $H = -\int p(x)\ln p(x)dx$ subjected to:

$$\begin{cases} \int p(x)dx = 1 \\ \int x p(x)dx = \mu \\ \int (x-\mu)^2 p(x)dx = \sigma^2 \end{cases}$$

$$\xrightarrow[\text{lagrangian terms}]{\text{form the}} H_L = -\int p(x)\ln p(x)dx + \lambda_1\left(\int p(x)dx - 1\right) + \lambda_2\left(\int x p(x)dx - \mu\right) + \lambda_3\left(\int (x-\mu)^2 p(x)dx - \sigma^2\right)$$

$$\Rightarrow \frac{\partial H_L}{\partial p(x)} = -\int\left([\ln p(x)] + p(x)\frac{1}{p(x)}\right)dx + \lambda_1\int 1dx + \lambda_2\int x dx + \lambda_3\int (x-\mu)^2 dx = 0$$

$$\Rightarrow -\int \ln p(x)dx - \int 1 dx + \lambda_1\int dx + \lambda_2\int x dx + \lambda_3\int (x-\mu)^2 dx = 0$$

Scanned with CamScanner

$$\Rightarrow \int \ln p(x)\,dx = \int \left[\lambda_1 + \lambda_2 x + \lambda_3 (x-\mu)^2 - 1\right] dx \xrightarrow[\text{from both sides}]{d/dx}$$

$$\rightarrow \ln p(x) = \lambda_1 + \lambda_2 x + \lambda_3 (x-\mu)^2 - 1 \Rightarrow p(x) = e^{\lambda_1 + \lambda_2 x + \lambda_3 (x-\mu)^2 - 1}$$

from the constraints we know:

$$\left.\begin{aligned}
&\int e^{\lambda_1 + \lambda_2 x + \lambda_3 (x-\mu)^2 - 1}\,dx = 1 \\
&\int x e^{\lambda_1 + \lambda_2 x + \lambda_3 (x-\mu)^2 - 1}\,dx = \mu \\
&\int (x-\mu)^2 e^{\lambda_1 + \lambda_2 x + \lambda_3 (x-\mu)^2 - 1}\,dx = \sigma^2
\end{aligned}\right\} \longrightarrow \text{solve for } \lambda_1, \lambda_2, \lambda_3 \text{ using wolfram mathematica}$$

$$\lambda_1 = \frac{1}{4} \ln\left(\frac{e^4}{4\pi^2\sigma^4}\right) = 1 - \frac{1}{2}\ln(2\pi\sigma^2) \;\Big|\; \lambda_2 = 0 \;\Big|\; \lambda_3 = -\frac{1}{2\sigma^2}$$

$$\rightarrow p(x) = e^{\cancel{1} - \frac{1}{2}\ln(2\pi\sigma^2) + 0 + \left(-\frac{1}{2\sigma^2}\right)(x-\mu)^2 \cancel{-1}} = e^{-\frac{1}{2}\ln(2\pi\sigma^2)}\, e^{\left(-\frac{1}{2\sigma^2}\right)(x-\mu)^2}$$

$$= \frac{1}{\sqrt{2\pi}\,\sigma} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \sim N(\mu, \sigma^2)$$

ست سوم

مسئله اول.

(a) $p(x_k|\theta) = \theta e^{-\theta x_k}$

$$\rightarrow l(\theta) = \ln \prod_{k=1}^{n} p(x_k|\theta) = \sum_{k=1}^{n} \ln(\theta e^{-\theta x_k}) = \sum_{k=1}^{n} [\ln\theta - \theta x_k]$$

$$\rightarrow \frac{\partial l(\theta)}{\partial\theta} = 0 \Rightarrow \sum_{k=1}^{n}\left[\frac{1}{\theta} - x_k\right] = 0 \Rightarrow \frac{n}{\theta} - \sum_{k=1}^{n} x_k = 0 \Rightarrow \theta = \frac{n}{\sum_{k=1}^{n} x_k} > 0 \rightarrow$$ مطابق محدودیت $\theta > 0$

(b) $p(x_k|\theta) = \theta c^{\theta} x_k^{-(\theta+1)}$

$$\rightarrow l(\theta) = \ln \prod_{k=1}^{n} p(x_k|\theta) = \sum_{k=1}^{n} \ln(\theta c^{\theta} x_k^{-(\theta+1)}) = \sum_{k=1}^{n}\left[\ln\theta + \theta\ln c + \underbrace{(-\theta+1)\ln x_k}_{-\theta\ln x_k + \ln x_k}\right]$$

$$\Rightarrow \frac{\partial l(\theta)}{\partial(\theta)} = 0 \Rightarrow \sum_{k=1}^{n}\left[\frac{1}{\theta} + \ln c - \ln x_k\right] = 0 \Rightarrow \frac{n}{\theta} + n\ln c - \sum_{k=1}^{n} \ln x_k = 0$$

$$\Rightarrow \theta = \frac{n}{\sum_{k=1}^{n} \ln x_k - n\ln c} \rightarrow$$ صورت کسر همواره مثبت، مخرج کسر نیز $x_k \geq c$ همواره مثبت است. بنابراین محدودیت $\theta > 0$ رعایت می‌شود. $(\sum \ln x_k \geq \sum \ln c = n\ln c)$

(c) $p(x_k|\theta) = c\theta^{c} x_k^{-(c+1)}$

$$\rightarrow l(\theta) = \ln \prod_{k=1}^{n} p(x_k|\theta) = \sum_{k=1}^{n} \ln(c\theta^{c} x_k^{-(c+1)}) = \sum_{k=1}^{n}\left[\ln c + c\ln\theta - (c+1)\ln x_k\right]$$

$$\Rightarrow \frac{\partial l(\theta)}{\partial(\theta)} = 0 \Rightarrow \sum_{k=1}^{n} \frac{c}{\theta} = 0 \Rightarrow \frac{nc}{\theta} = 0 \rightarrow \begin{cases} n = 0 \rightarrow \text{بی معنی} \\ \text{or} \\ c = 0 \rightarrow \text{امکان ندارد } c > 0 \end{cases}$$

بنابراین مشخص است که مشتق این تابع نمی‌تواند صفر باشد و با افزایش مقدار θ مقدار $l(\theta)$ نیز به صورت بی‌کران افزایش می‌یابد. اما اگر به یکی از محدودیت‌ها دقت کنیم: $x_k \geq \theta$ ⇐ بهترین انتخاب ممکن برای θ یعنی بزرگترین θ ممکن برابر با $\underline{\min_k\{x_k\}}$ خواهد بود.

(d) $p(x_k|\theta) = \sqrt{\theta}\, x_k^{\sqrt{\theta}-1}$

$$\rightarrow l(\theta) = \ln \prod_{k=1}^{n} p(x_k|\theta) = \sum_{k=1}^{n} \ln p(x_k|\theta) = \sum_{k=1}^{n} \left[\ln\sqrt{\theta} + (\sqrt{\theta}-1)\ln x_k\right]$$

$$\frac{\partial l(\theta)}{\partial\theta} = \sum_{k=1}^{n}\left[\frac{\left(\frac{1}{2\sqrt{\theta}}\right)}{\sqrt{\theta}} + \frac{\ln x_k}{2\sqrt{\theta}}\right] = \sum_{k=1}^{n}\left[\frac{1}{2\theta} + \frac{\ln x_k}{2\sqrt{\theta}}\right] = 0$$

$$\Rightarrow \sum_{k=1}^{n} \frac{1+\sqrt{\theta}\ln x_k}{2\theta} = 0 \rightarrow \frac{n}{2\theta} + \frac{1}{2\sqrt{\theta}}\underbrace{\boxed{\sum_{k=1}^{n}\ln x_k}}_{A} = 0 \Rightarrow \frac{n}{2\theta} = \frac{-A}{2\sqrt{\theta}}$$

$$\Rightarrow 2n\sqrt{\theta} = -2A\theta \Rightarrow n\sqrt{\theta} + A\theta = 0 \xrightarrow{\sqrt{\theta}=u} Au^2 + nu = 0 \Rightarrow u(Au+n) = 0$$

$$\rightarrow \begin{cases} u=0 \rightarrow \sqrt{\theta}=0 \longrightarrow \text{غیرممکن چون } \theta>0 \\ Au+n=0 \Rightarrow \sqrt{\theta}\left(\sum_{k=1}^{n}\ln x_k\right) + n = 0 \Rightarrow \sqrt{\theta} = \dfrac{-n}{\sum_{k=1}^{n}\ln x_k} \Rightarrow \theta = \dfrac{n^2}{\left[\sum_{k=1}^{n}\ln x_k\right]^2} \end{cases}$$

این عبارت باید مثبت باشد چون برابر حاصل یک رادیکال است.

از آنجایی که x_k ها بین صفر و یک هستند بنابراین جمع لگاریتم آن ها یک عدد منفی خواهد بود که با منفی صورت کسر ساده می شود.

(e) $p(x_k|\theta) = \left(\frac{x_k}{\theta^2}\right) e^{-\frac{x_k^2}{2\theta^2}}$

$$\rightarrow l(\theta) = \ln\prod_{k=1}^{n} p(x_k|\theta) = \sum_{k=1}^{n}\ln p(x_k|\theta) = \sum_{k=1}^{n}\left[\ln x_k - 2\ln\theta - \frac{x_k^2}{2\theta^2}\right]$$

$$\frac{\partial l(\theta)}{\partial\theta} = \sum_{k=1}^{n}\left[\frac{-2}{\theta} - \frac{x_k^2}{2}(-2\theta^{-3})\right] = \frac{-2n}{\theta} + \sum_{k=1}^{n} x_k^2\theta^{-3} = \frac{-2n}{\theta} + \frac{1}{\theta^3}\underbrace{\sum_{k=1}^{n} x_k^2}_{A} = 0$$

$$\Rightarrow \frac{-2n}{\theta} + \frac{A}{\theta^3} = 0 \Rightarrow \frac{-2n\theta^2 + A}{\theta^3} = 0 \Rightarrow A = 2n\theta^2 \Rightarrow \theta^2 = \frac{A}{2n} \Rightarrow \theta = \pm\sqrt{\frac{A}{2n}}$$

همواره مثبت چون A مثبت و 2n مثبت

جواب منفی غیرممکن چون $\theta>0$

$$\theta = \sqrt{\left(\sum_{k=1}^{n} x_k^2\right)/2n}$$

Scanned with CamScanner

(f) $p(x_k|\theta) = \theta c x_k^{c-1} e^{-\theta x_k^c}$

$$\rightarrow l(\theta) = \ln \prod_{k=1}^{n} p(x_k|\theta) = \sum_{k=1}^{n} \ln p(x_k|\theta) = \sum_{k=1}^{n} \left[\ln\theta + \ln c + (c-1)\ln x_k - \theta x_k^c\right]$$

$$\rightarrow \frac{\partial l(\theta)}{\partial \theta} = \sum_{k=1}^{n}\left[\frac{1}{\theta} - x_k^c\right] = 0 \rightarrow \frac{n}{\theta} - \sum_{k=1}^{n} x_k^c = 0 \rightarrow \theta = \frac{n}{\sum_{k=1}^{n} x_k^c} \rightarrow$$

$\theta > 0$ برقرار است

چون $x_k > 0$ و $c > 0$

$\rightarrow x_k^c > 0$

(g) $$p(x_k|\theta) = \begin{cases} 1 & \theta - \frac{1}{2} \leq x_k \leq \theta + \frac{1}{2} \\ 0 & \text{o.w.} \end{cases}$$

اگر حداقل یکی از نقاط خارج از بازه‌ی $[\theta - \frac{1}{2}, \theta + \frac{1}{2}]$ انتخاب شوند $\prod_{k=1}^{n} p(x_k|\theta)$ برابر صفر می‌شود و در غیر این صورت برابر 1 می‌شود. در هر دو حالت مشتق این عبارت نسبت به θ برابر صفر است.
بنابراین انتخاب θ هیچ تاثیری ندارد و مقدار likelihood تابع به ازای هر θ ثابت است.

مسئله‌ی دوم.

نمونه‌ها به صورت مستقل انتخاب شده‌اند بنابراین: joint likelihood: $\prod_{j=1}^{m}(p(x_j;\mu_1,\sigma^2))\prod_{j=1}^{n}(p(y_j;\mu_2,\sigma^2))$

$$\rightarrow L = \prod_{i=1}^{m}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x_i-\mu_1)^2}{2\sigma^2}}\prod_{j=1}^{n}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(y_j-\mu_2)^2}{2\sigma^2}}$$

$$= \left[\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^m e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{m}(x_i-\mu_1)^2}\right]\left[\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n e^{-\frac{1}{2\sigma^2}\sum_{j=1}^{n}(y_j-\mu_2)^2}\right]$$

$$= \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^{m+n} e^{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{m}(x_i-\mu_1)^2 + \sum_{j=1}^{n}(y_j-\mu_2)^2\right]}$$

$$\text{log-likelihood} = \ln(L) = -(m+n)\ln\sqrt{2\pi}\sigma - \frac{1}{2\sigma^2}\left[\sum_{i=1}^{m}(x_i-\mu_1)^2 + \sum_{j=1}^{n}(y_j-\mu_2)^2\right]$$

$$\frac{\partial \ln(L)}{\partial \mu_1} = -\frac{1}{2\sigma^2}\left(-2\sum_{i=1}^{m}(x_i-\mu_1)\right) = 0 \Rightarrow \sum_{i=1}^{m}(x_i-\mu_1) = 0 \rightarrow \mu_1 = \frac{\sum_{i=1}^{m} x_i}{m}$$

بنابر تقارن مسئله نسبت به μ_1 و μ_2: $\mu_2 = \frac{\sum_{j=1}^{n} y_j}{n}$

اکنون نسبت به σ مشتق می‌گیریم.

1

اگر قرار دهیم: $A = \sum (x_i-\mu_1)^2 + \sum (y_j-\mu_2)^2$

$$\frac{\partial \ln(L)}{\partial \sigma} = \frac{-(m+n)\sqrt{2\pi}}{\sigma\sqrt{2\pi}} + \frac{A}{\sigma^3} = 0 \Rightarrow -(m+n)\sigma^2 + A = 0$$

$$\Rightarrow \sigma = \sqrt{\frac{A}{(m+n)}} = \boxed{\sqrt{\frac{\sum_{i=1}^{m}(x_i-\mu_1)^2 + \sum_{j=1}^{n}(y_j-\mu_2)^2}{(m+n)}}}$$

مسئله‌ی سوم:

a) با توجه به تمرین های حل شده‌ی قبلی و مطالب بررسی شده در کلاس می دانیم که تخمین ML برای توزیع نرمال به صورت مقابل است. ←

$$N(\mu, \sigma^2): \begin{cases} \hat{\mu}_{ML} = \frac{1}{N}\sum_{i=1}^{N} x_i \\ \hat{\sigma}^2_{ML} = \frac{1}{N}\sum_{i=1}^{N}(x_i - \hat{\mu}_{ML})^2 \end{cases}$$

این مقادیر را برای نمونه های داده شده در صورت سوال محاسبه کنیم:

$$\hat{\mu}_{ML,\omega_1} = 1 \qquad \hat{\sigma}^2_{ML,\omega_1} = 0.094$$

$$\Rightarrow p(x|\omega_1) = \frac{1}{0.094\sqrt{2\pi}} e^{-\frac{(x-1)^2}{2(0.094)^2}}$$

b) طبق اسلاید های مطرح شده در کلاس می دانیم که برای یک توزیع نرمال یادگیری بیزی به نتایج زیر منجر می شود:

$$\mu_N = \frac{N\sigma_0^2\bar{x} + \sigma^2\mu_0}{N\sigma_0^2 + \sigma^2}, \quad \sigma_N^2 = \frac{\sigma^2\sigma_0^2}{N\sigma_0^2 + \sigma^2}, \quad \bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$$

$$p(x|\omega_2) = \frac{1}{\sqrt{2\pi(\sigma^2+\sigma_N^2)}} e^{-\frac{1}{2}\frac{(x-\mu_N)^2}{(\sigma^2+\sigma_N^2)}}$$

اکنون به محاسبه‌ی این مقادیر می پردازیم:

$\mu_0 = 0.5 \; , \; \sigma_0^2 = 0.5 \; , \; \bar{x} = 0 \; , \; \sigma_N^2 = \dfrac{1 \times 0.5}{10 \times 0.5 + 1} \approx 0.083$

$\sigma_c^2 = 1$

$$\mu_N = \frac{10 \times 0.5 \times 0 + 1 \times 0.5}{10 \times 0.5 + 1} \approx 0.083$$

$$\Rightarrow P(x|w_2) = \frac{1}{\sqrt{2\pi(1+0.083)}} e^{-\frac{1}{2}\frac{(x-0.083)^2}{(1+0.083)}}$$

ic

$$l_{12} \equiv \frac{P(\underline{x}|w_1)}{P(\underline{x}|w_2)} >(<) \cancel{\frac{P(w_2)}{P(w_1)}}^{\,1}$$

$$\Rightarrow l_{12} \equiv \frac{P(\underline{x}|w_1)}{P(\underline{x}|w_2)} >(<) 1 \longrightarrow l_{12} \equiv P(\underline{x}|w_1) >(<) P(\underline{x}|w_2)$$

برای محاسبه‌ی مرز تصمیم $P(\underline{x}|w_1) = P(\underline{x}|w_2)$ را حل می‌کنیم:

$$P(\underline{x}_0|w_1) = \frac{1}{0.094\cancel{\sqrt{2\pi}}} e^{\frac{-(x_0-1)^2}{0.017672}} = P(\underline{x}_0|w_2) = \frac{1}{1.04\cancel{\sqrt{2\pi}}} e^{\frac{-(x_0-0.083)^2}{2.166}}$$

$$\Rightarrow e^{\frac{-(x_0-1)^2}{0.0177}} = \underbrace{\cancel{\frac{0.094}{1.04}}}_{\approx 0.09} e^{\frac{-(x_0-0.083)^2}{2.166}} \xrightarrow{\ln} \frac{-(x_0-1)^2}{0.0177} = \ln(0.09) - \frac{(x_0-0.083)^2}{2.166}$$

$$\xrightarrow{\text{mathematica}} \begin{cases} x_1 = 0.784 \\ x_2 = 1.23 \end{cases} \longrightarrow \begin{cases} c = w_1 \; , \; x_1 < x < x_2 \\ c = w_2 \; , \; \text{o.w.} \end{cases}$$