SML ASSIGNMENT-1

ANSWER-1.

a) Plotting $P(x|W_1)$ v/s x.

$P(x|W_2)$ v/s x. DONE in .py file

Likelihood ratio $\frac{P(x|W_1)}{P(x|W_2)}$ v/s x

b) **DECISION BOUNDARY** (Minimize error in case of :)

(i) Zero-One loss $\equiv$ Loss fxn $= \begin{bmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

$$R(\alpha_1|x) = \lambda_{11} p(W_1|x) + \lambda_{12} p(W_2|x)$$

$$= \lambda_{12} p(W_2|x) \qquad (\because \lambda_{11} = 0)$$

$$R(\alpha_2|x) = \lambda_{21} p(W_1|x) + \lambda_{22} p(W_2|x)$$

$$= \lambda_{21} p(W_1|x) \qquad (\because \lambda_{22} = 0)$$

To minimize the error,

$$P(error) = \int_{-\infty}^{\infty} P(error|x)\, P(x)\, dx.$$

we should minimize $P(error|x)$ for each x so that $P(error)$ is less.

$$P(error|x) = \begin{cases} R(\alpha_1|x) & \text{if we decide } w_1 \\ R(\alpha_2|x) & \text{if we decide } w_2 \end{cases}$$

$$\Rightarrow \quad P(error|x) = \begin{cases} P(w_2|x) & \text{if we decide } w_1 \\ P(w_1|x) & \text{if we decide } w_2 \end{cases}$$

$$(\because \lambda_{12} = \lambda_{21} = 1)$$

Hence, $P(error/x) = \min\left[P(w_1/x), P(w_2/x)\right]$

Now, Hence a good decision boundary will be x such that,

$$P(w_1|x) = P(w_2|x)$$

$$\Rightarrow \frac{P(x|w_1)\,P(w_1)}{\cancel{P(x)}} = \frac{P(x|w_2)\,P(w_2)}{\cancel{P(x)}}$$

$$\Rightarrow N(2,1) \cdot \frac{1}{\cancel{4}} = N(5,1) \cdot \frac{3}{\cancel{4}}$$

$$\Rightarrow \frac{N(2,1)}{N(5,1)} = 3$$

$$\Rightarrow \frac{e^{-1/2(x-2)^2}}{e^{-1/2(x-5)^2}} = 3$$

$$\Rightarrow e^{-1/2(6x-21)} = 3$$

$$\Rightarrow -\frac{1}{2}(6x-21) = \ln(3) \quad \{\text{Taking ln both sides}\}$$

$$\Rightarrow \boxed{x = \frac{21 - 2\ln(3)}{6}}$$

Simplify, $\bar{x} = 7/2 - \frac{\ln(3)}{3}$

$$\boxed{\bar{x} \simeq 3.137}$$ Decision Boundary

So for $x < \bar{x}$ Choose w_1.

$x > \bar{x}$ Choose w_2.

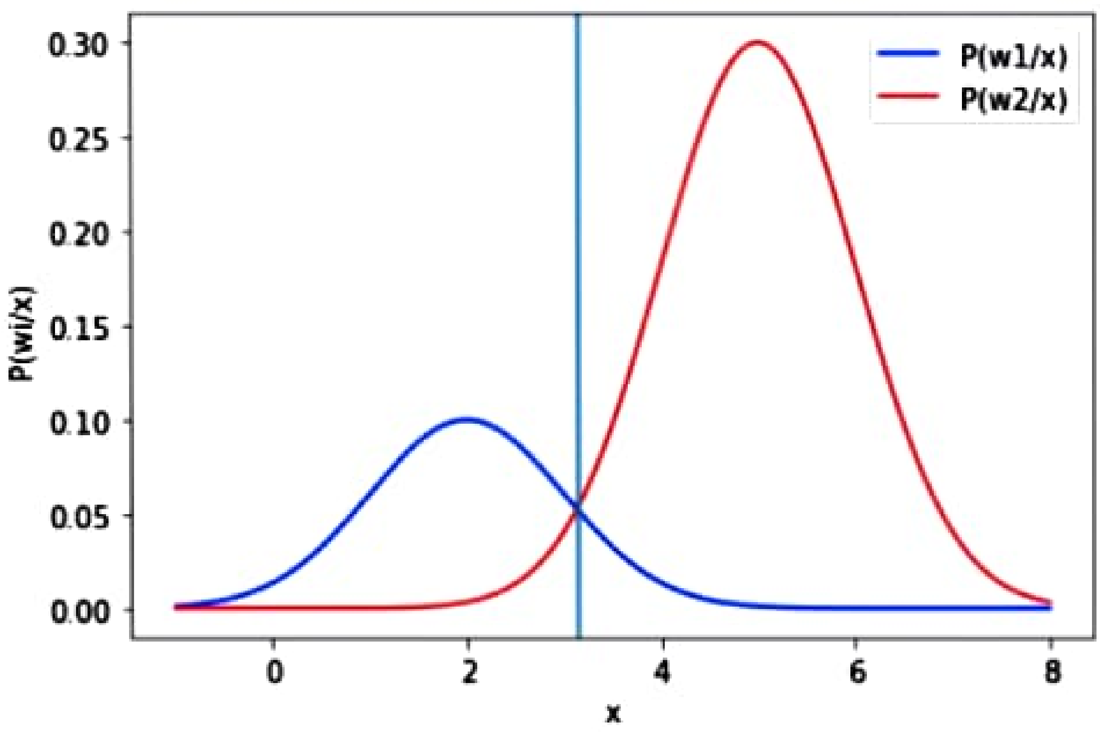
In [52]:

```python
# Decision boundary (Zero-one loss)

plt.plot(x,pw1x,'b',label = 'P(w1/x)')
plt.plot(x,pw2x,'r',label = 'P(w2/x)')

plt.axvline(x=3.137)

plt.xlabel("x")
plt.ylabel("P(wi/x)")
plt.legend()
plt.show()
```



(ii) To minimize the error for given loss fxn same decision boundary will be there, But to reduce loss / risk we will need to calculate again.

$$\lambda = \begin{bmatrix} 0 & 2 \\ 3 & 0 \end{bmatrix} = \begin{bmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22} \end{bmatrix}.$$

$\rightarrow \quad R(\alpha_1|x) = \lambda_{12} P(w_2|x) = 2P(w_2|x)$

$\rightarrow \quad R(\alpha_2|x) = \lambda_{21} P(w_1|x) = 3P(w_1|x)$

$(\because \lambda_{11} = \lambda_{22} = 0)$

$$P(\text{error}/x) = \min\{R(\alpha_1|x), R(\alpha_2|x)\}.$$

Decision boundary will be x s.t.

$$R(\alpha_1|x) = R(\alpha_2|x)$$

$$\Rightarrow \quad 2 \times P(w_2|x) = 3 \times P(w_1|x)$$

$$\Rightarrow \quad 2 \times \cancel{3}/\cancel{4} \times N(5,1) = \cancel{3} \times 1/\cancel{4} \times N(2,1)$$

$$\Rightarrow \quad \frac{N(2,1)}{N(5,1)} = 2$$

From previous calc^n we can say.

$$\Rightarrow \quad e^{-1/2(6x-21)} = 2$$

$$\Rightarrow \quad -1/2(6x-21) = \ln(2) \quad \text{(Taking ln both sides)}$$

$$\Rightarrow \quad x = \frac{21 - 2\ln(2)}{6}$$

$$\Rightarrow \quad x = 7/2 - \frac{\ln(2)}{3}$$

$$\Rightarrow \quad \boxed{x \simeq 3.268}$$

Decision boundary shifted Right because we had to pay more cost for wrong identification of one as two than vice-versa.

$\therefore$ Now less such error will be there & hence our loss/error will be minimum.

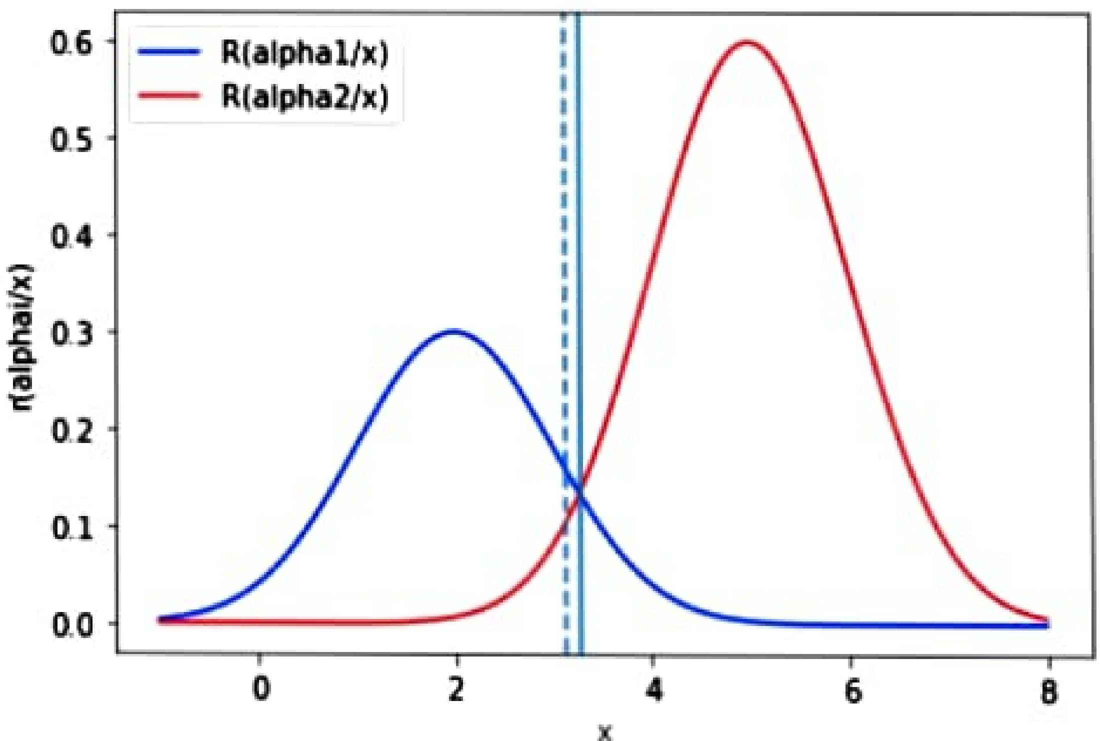
In [59]:

```python
# Decision boundary (Acc to given loss fxn)

plt.plot(x,3*pw1x,'b',label = 'R(alpha1/x)')
plt.plot(x,2*pw2x,'r',label = 'R(alpha2/x)')

plt.axvline(x=3.137,linestyle='--')
plt.axvline(x=3.268)

plt.xlabel("x")
plt.ylabel("r(alphai/x)")
plt.legend()
plt.show()
```



c) No we shouldn't prefer zero-one loss for a task like task : Cancer prediction where w_1 = Cancer predicted w_2 = No cancer, bcz there loss fxn is not symm.

ie $\lambda_{12} <<< \lambda_{21}$ ie loss incurred due to wrong guess prediction of non-cancerous patient as cancerous is less costly or loss than claiming cancerous as non cancerous. Because λ_{21} may account for someones life as well where λ_{12} may be monetary losses.

$$\left\{ \underset{\lambda_{21}}{\text{Life}} >>> \underset{\lambda_{12}}{\text{Money}} \right\}$$

Ans-2.

$$X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \qquad \mu = E[X] = \begin{bmatrix} 5 \\ -5 \\ 6 \end{bmatrix}$$

Random vector

$$Cov[X,X] = \begin{bmatrix} 5 & 2 & -1 \\ 2 & 5 & 0 \\ -1 & 0 & 4 \end{bmatrix}$$

Nothing Noteworthy obsⁿ can be made

$Y = A^T X + B$ where $A = (2, -1, 2)^T$

$\Downarrow \qquad \Rightarrow A^T = [2 \ -1 \ 2]$

Taking Expⁿ both sides $\quad$ & $B = 5$

$$E[Y] = E[A^T X + B]$$

Since A^T & B are constants $\therefore E[A^T] = A^T$ & $E[B] = B$

$$\Rightarrow E[Y] = \mu_Y = A^T . E[X] + B$$

$$= [2 \ -1 \ 2] \begin{bmatrix} 5 \\ -5 \\ 6 \end{bmatrix} + 5$$

$$\mu_Y = [(2 \times 5) + (-1 \times -5) + (2 \times 6)] + 5$$

$$\mu_Y = 10 + 5 + 12 + 5$$

$$\boxed{\mu_Y = 32}$$

$\therefore$ Mean of $Y = A^T X + B$ is 32.

Ans-3 Cauchy's pdf

$$P(x/w_i) = \frac{1}{\pi b} \cdot \frac{1}{1 + \left(\frac{x-a_i}{b}\right)^2}, \quad i = 1, 2.$$

(i) for zero-one loss $\lambda = \begin{bmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

Minimum error rate is when,

$$P(\text{error}/x) = \min [P(w_1/x), P(w_2/x)].$$

& Decision boundary is x such that,

$$P(w_1/x) = P(w_2/x)$$

Assuming $P(w_1) = P(w_2) = 1/2$.

$$\frac{P(x/w_1)\,\cancel{P(w_1)}}{\cancel{P(x)}} = \frac{P(x/w_2)\,\cancel{P(w_2)}}{\cancel{P(x)}}$$

$$\Rightarrow P(x/w_1) = P(x/w_2)$$

$$\Rightarrow \frac{\cancel{1}}{\cancel{\pi b}} \cdot \frac{1}{1 + \left(\frac{x-a_1}{b}\right)^2} = \frac{\cancel{1}}{\cancel{\pi b}} \cdot \frac{1}{1 + \left(\frac{x-a_2}{b}\right)^2}$$

$$\Rightarrow \cancel{1} + \left(\frac{x-a_1}{\cancel{b}}\right)^2 = \cancel{1} + \left(\frac{x-a_2}{\cancel{b}}\right)^2$$

$$\Rightarrow \cancel{x^2} - 2a_1 x + a_1^2 = \cancel{x^2} - 2a_2 x + a_2^2$$

$$\Rightarrow 2x(a_2 - a_1) = a_2^2 - a_1^2$$

$$\Rightarrow \boxed{\bar{x} = \frac{a_1 + a_2}{2}.}$$

Decision Boundary.

(Independent of b)

For $a_1 = 3$ & $a_2 = 5$

$$\bar{x} = \frac{3+5}{2}$$

$$\Rightarrow \boxed{\bar{x} = 4}$$

If $x < \bar{x}$ choose w_1

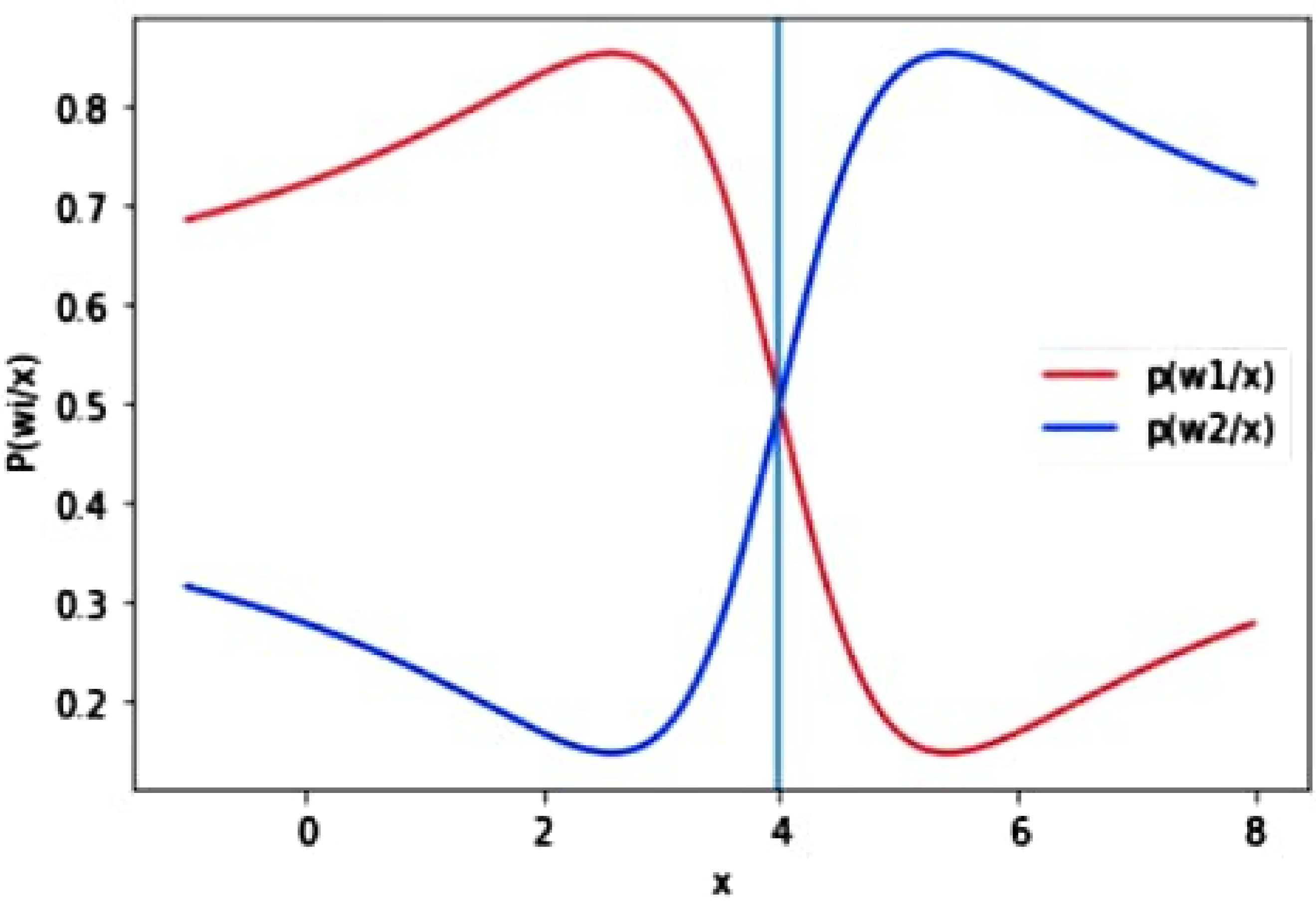
If $x > \bar{x}$ choose w_2

```python
# DECISION BOUNDARY

plt.plot(x,pw1x,'r',label = 'p(w1/x)')
plt.plot(x,pw2x,'b',label = 'p(w2/x)')

plt.axvline(x = 4)

plt.xlabel("x")
plt.ylabel("P(wi/x)")
plt.legend()
plt.show()
```



(iii):

$$P(w_1|x) = \frac{P(x|w_1)\,P(w_1)}{P(x)}$$

where $P(x) = P(x|w_1)P(w_1) + P(x|w_2)P(w_2)$

Done in .py file

3) Overall error rate $P(\text{Error})$

let R_1 be the region where classifier predicts w_1.

& R_2 " " " w_2.

$\therefore$ R_1 is $\rightarrow$ $-\infty$ to $4\,(\bar{x})$ $(-\infty, 4)\ R_1$

R_2 is $\Rightarrow$ $\bar{x}\,(4)$ to ∞ $(4, \infty)\ R_2$

$$P(\text{error}/x) = \begin{cases} P(w_1/x) & \text{if } x \in R_2 \\ P(w_2/x) & \text{if } x \in R_1 \end{cases}.$$

$$\therefore\ P(\text{error}/x) = \min\{P(w_1/x), P(w_2/x)\}.$$

Total overall Error rate $= P(\text{error}) = \int_{-\infty}^{\infty} P(\text{error}/x)\, P(x)\, dx.$

$$= \int P(w_2/x)\, P(x)\, dx + \int_{R_2} P(w_1/x)\, P(x)\, dx.$$

$$= \int_{-\infty}^{4} P(x/w_2)\, P(w_2)\, dx + \int_{4}^{\infty} P(x/w_1)\, P(w_1)\, dx.$$

(using bayes)

$$= \frac{1}{2}\left[\int_{-\infty}^{4} P(x/w_2)\, dx + \int_{4}^{\infty} P(x/w_1)\, dx\right]$$

$P(w_1) = P(w_2)$

$$\Rightarrow \frac{1}{2}\left[\int_{-\infty}^{4} \frac{1}{\pi}\cdot\frac{dx}{1+(x-5)^2} + \int_{4}^{\infty} \frac{1}{\pi}\cdot\frac{dx}{1+(x-3)^2}\right].$$

using $a_1 = 3$, $a_2 = 5$ \& $b = 1$

$$\Rightarrow \frac{1}{2\pi}\left[\int_{-\infty}^{4} \frac{dx}{1+(x-5)^2} + \int_{4}^{\infty} \frac{dx.}{1+(x-3)^2}\right].$$

$$\Rightarrow \frac{1}{2\pi}\left[\pi/4 + \pi/4\right].$$

$\Rightarrow \underline{\underline{1/4}}$ Answer. $\Rightarrow$ $\boxed{P(\text{error}) = 1/4}$

Overall error rate

for general a_1, a_2 \& b

$$P(\text{error}) = \frac{1}{2} - \frac{1}{\pi}\tan^{-1}\left|\frac{a_2 - a_1}{2b}\right|$$

$\therefore$ Substituting $a_2 = 5$, $a_3 = 3$, $b = 1$

$$\Rightarrow P(\text{error}) = \frac{1}{2} - \frac{1}{\pi}\tan^{-1}(1)$$

$$= \frac{1}{2} - \frac{1}{4}$$

$P(\text{error}) = 1/4$ overall error rate.

Answer-4

$$x = \begin{bmatrix} a \\ b \end{bmatrix}_{2\times 1}$$

a : Bernoulli R.V : $P_a(\theta) = \left\{ \begin{array}{ll} 1-\theta & a=0 \\ \theta & a=1 \\ 0 & \text{otherwise} \end{array} \right\}$

b : Gaussian R.V : $P_b(m, \sigma^2) = N(m, \sigma^2)$

$$= \frac{1}{\sqrt{2\pi\sigma}} \exp\left\{-\tfrac{1}{2}\frac{(x-m)^2}{\sigma^2}\right\}$$

Covariance of x : $\begin{bmatrix} \theta(1-\theta) & 0 \\ 0 & \sigma^2 \end{bmatrix}$

$\therefore$ Var(a) = $\theta(1-\theta)$. Var(b) = σ^2

Cov(a,b) = Cov(b,a) = 0

$\therefore$ Since Cov(a,b) = Cov(b,a) = 0

a $ b are INDEPENT RANDOM VARIABLES.

OR UNCORRELATED R.VS.

$\therefore$ Joint Pdf of a $ b,

$$\boxed{P(x) = P(a = x_1 \mid b = x_2) = P_a(x_1) . P_b(x_2)}$$

$$P(x) = \left((1-\theta)^{(1-x_1)} . \theta^{(x_1)}\right) . \left(\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\tfrac{1}{2}\frac{(x_2-m)^2}{\sigma^2}\right\}\right)$$

where, $x_1 \in \{0, 1\}$.

$x_2 \in \mathbb{R}$

Ans-4

(b) We have found pdf $p(x)$ from part (a).

There are N iid samples drawn. $(x_i, y_i) \Rightarrow (z_i)$

Joint probability of these N samples is $q(\bar{X})$. $\bar{X} = (z_1, z_2 \dots z_n)$

$$\Rightarrow \quad q(\bar{x}) = P(z_1, z_2, z_3 \dots \dots z_n) = P(z_1).P(z_2)\dots P(z_n)$$

where $z_i = (x_i, y_i)$

$$\Rightarrow \quad q(\bar{x}) = \prod_{i=1}^{N} P(z_i)$$

where $P(z_i) = P(a = x_i \mid b = y_i) = (1-\theta)^{(1-x_i)} \theta^{x_i} \left(\frac{1}{\sqrt{2\pi\sigma}} \exp\left\{ -\frac{1}{2} \left(\frac{y-m}{\sigma} \right)^2 \right\} \right)$

$$\therefore \quad q(\bar{x}) = \prod_{i=1}^{N} (1-\theta)^{(1-x_i)} \cdot \theta^{x_i} \cdot \frac{1}{\sqrt{2\pi\sigma}} \exp\left\{ -\frac{1}{2} \left(\frac{y_i - m}{\sigma} \right)^2 \right\}$$

Take ln both sides

$$\ln(q(\bar{x})) = \sum_{i=1}^{N} \ln\left((1-\theta)^{(1-x_i)} \cdot \theta^{x_i} \cdot \frac{1}{\sqrt{2\pi\sigma}} \exp\left\{ -\frac{1}{2} \left(\frac{y_i - m}{\sigma} \right)^2 \right\} \right)$$

$$\rightarrow \quad \ln(q(\bar{x})) = \sum_{i=1}^{N} \Big\{ (1-x_i)\ln(1-\theta) + x_i \cdot \ln(\theta) + \ln\left(\frac{1}{\sqrt{2\pi\sigma}}\right) + \ln\left(\exp\left(-\tfrac{1}{2}\left(\frac{y_i-m}{b}\right)^2\right)\right) \Big\}$$

$$\ln(q(\bar{x})) = \sum_{i=1}^{N} \Big\{ (1-x_i)\ln(1-\theta) + x_i \ln(\theta) + \ln\left(1/\sqrt{2\pi\sigma}\right) + \left(-\tfrac{1}{2}\frac{(y_i-m)^2}{b^2}\right) \Big\}$$

$$\ln(q(\bar{x})) = \ln(1-\theta)\sum_{i=1}^{N}(1-x_i) + \ln\theta \sum_{i=1}^{N} x_i + \underbrace{\sum_{i=1}^{N}\left(\ln\left(\frac{1}{\sqrt{2\pi\sigma}}\right) - \tfrac{1}{2}\frac{(y_i-m)^2}{b^2}\right)}$$

Independent of θ

To find value of θ for which $q(\bar{x})$ is Max^m

Differentiate both sides & $\frac{d(\ln q(\bar{x}))}{d\theta} = 0$

wrt θ.

$$\rightarrow \quad \frac{d(\ln(q(\bar{x})))}{d\theta} = -\left(\frac{1}{1-\theta}\right)\sum_{i=1}^{N}(1-x_i) + \frac{1}{\theta}\sum_{i=1}^{N} x_i + 0$$

Partial differentiation

Independent of θ

$$\left\{ \frac{d(\ln(x))}{dx} = 1/x \right\}.$$

$$\theta = \frac{1}{\theta-1}\sum_{i=1}^{N}(1-x_i) + \frac{1}{\theta}\sum_{i=1}^{N}x_i$$

$$\Rightarrow \frac{1}{1-\theta}\cdot\sum_{i=1}^{N}(1-x_i) = \frac{1}{\theta}\sum_{i=1}^{N}x_i$$

$$\Rightarrow \frac{N - \sum_{i=1}^{N}x_i}{1-\theta} = \frac{\sum_{i=1}^{N}x_i}{\theta}$$

$$\Rightarrow N\theta - \cancel{\theta\sum_{i=1}^{N}x_i} = \sum_{i=1}^{N}x_i - \cancel{\theta\sum_{i=1}^{N}x_i}$$

$$\Rightarrow \boxed{\theta_{max} = \frac{\sum_{i=1}^{N}x_i}{N}}$$

Answer

<u>Value of θ</u> for which $q(X)$ is Maximum

is <u>Mean of x_i</u> for all $i = 1,\ldots N$

where x_i is the values corresponding to

(a) <u>Bernaulli random values only</u>.