Third Subtask:-

Name: Amit Kumar
Roll: 20CS30003

Q. $$p(y) = \begin{cases} \phi & \text{if } y=1 \\ 1-\phi & \text{if } y=0 \end{cases} \quad -(1)$$

$$p(x/y=0) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x-\mu_0)^T \Sigma^{-1} (x-\mu_0)\right) \quad -(2)$$

$$p(x/y=1) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x-\mu_1)^T \Sigma^{-1} (x-\mu_1)\right) \quad -(3)$$

ϕ, μ_0, μ_1 & Σ being parameters.

Show that :—

$$p(y=1 \mid x; \phi, \mu_0, \mu_1, \Sigma) = \frac{1}{1+\exp(-(\theta^T x + \theta_0))}$$

where $\theta \in \mathbb{R}^n$ & $\theta_0 \in \mathbb{R}$ are f^n of ϕ, μ_0, μ_1 & Σ

Ans:- $P(y|x) = \dfrac{P(x/y) \cdot P(y)}{P(x)}$ using Bayes Rule.

when $y=1$,

$$\Rightarrow P(y=1/x) = \frac{P(x/y=1) \cdot p(y=1)}{P(x)}$$

Since $P(x) = P(x/y=0)\,P(y=0) + P(x/y=1)\,P(y=1)$

$$\Rightarrow P(y=1/x) = \frac{P(x/y=1) \cdot p(y=1)}{P(x/y=0)\,P(y=0) + P(x/y=1)\,P(y=1)}$$

Dividing Numerator & Denominator by $P(x/y=1)\,P(y=1)$

we get,

$\Rightarrow P(y=1/x) = \dfrac{1}{\dfrac{P(x/y=0)P(y=0)}{P(x/y=1)P(y=1)} + 1}$

$\therefore$ ~~Rdp~~ Now, $P(y=0) = 1-\phi$ & $P(y=1) = \phi$
using eqn (1)

Now, using eqn (2), (3) & eqn (1) we get,

$$\Rightarrow P(y=1/x) = \dfrac{1}{\dfrac{\cancel{\frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}}}\exp\left(-\frac{1}{2}(x-\mu_0)^T\Sigma^{-1}(x-\mu_0)\right)\cdot(1-\phi)}{\cancel{\frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}}}\exp\left(-\frac{1}{2}(x-\mu_1)^T\Sigma^{-1}(x-\mu_1)\right)\cdot\phi} + 1}$$

Now, $\frac{1-\phi}{\phi}$ can be written as $\exp\left(\log\left(\frac{1-\phi}{\phi}\right)\right)$

$$\Rightarrow P(y=1/x) = \dfrac{1}{\exp\left[-\frac{1}{2}(x-\mu_0)^T\Sigma^{-1}(x-\mu_0) + \frac{1}{2}(x-\mu_1)^T\Sigma^{-1}(x-\mu_1) + \log\left(\frac{1-\phi}{\phi}\right)\right] + 1}$$

$$\Rightarrow P(y=1/x) = \frac{1}{\exp(\lambda)+1} \quad -(4)$$

Now let us shorten λ :-

$$\lambda = -\frac{1}{2}(x-\mu_0)^T\Sigma^{-1}(x-\mu_0) + \frac{1}{2}(x-\mu_1)^T\Sigma^{-1}(x-\mu_1) + \log\left(\frac{1-\phi}{\phi}\right)$$

$$\Rightarrow \lambda = -\frac{1}{2}\left[\cancel{x^T\Sigma^{-1}x} - 2x^T\Sigma^{-1}\mu_0 + \mu_0^T\Sigma^{-1}\mu_0\right] + \frac{1}{2}\left[\cancel{x^T\Sigma^{-1}x} - 2x^T\Sigma^{-1}\mu_1 + \mu_1^T\Sigma^{-1}\mu_1\right] + \log\left(\frac{1-\phi}{\phi}\right)$$

$\because (A+B)^T = A^T + B^T$ & $x^T\Sigma^{-1}\mu_0 = \mu_0\Sigma^{-1}x^T$

$$\Rightarrow \lambda = x^T \Sigma^{-1} \mu_0 - x^T \Sigma^{-1} \mu_1 - \frac{1}{2} \mu_0^T \Sigma \mu_0 + \frac{1}{2} \mu_1^T \Sigma^{-1} \mu_1 + \log\left(\frac{1-\phi}{\phi}\right)$$

$$\Rightarrow \lambda = x^T \Sigma^{-1} (\mu_0 - \mu_1) - \frac{1}{2} \mu_0^T \Sigma^{-1} \mu_0 + \frac{1}{2} \mu_1^T \Sigma^{-1} \mu_1 + \log\left(\frac{1-\phi}{\phi}\right)$$

θ_0

$$\Rightarrow \lambda = -\left[x^T \Sigma^{-1} (\mu_1 - \mu_0) + \frac{1}{2} \mu_0^T \Sigma^{-1} \mu_0 - \frac{1}{2} \mu_1^T \Sigma^{-1} \mu_1 - \log\left(\frac{1-\phi}{\phi}\right)\right]$$

let $\theta_0 = \frac{1}{2} \mu_0 \Sigma^{-1} \mu_0 - \frac{1}{2} \mu_1^T \Sigma^{-1} \mu_1 - \log\left(\frac{1-\phi}{\phi}\right)$

& $\theta = \Sigma^{-1} (\mu_1 - \mu_0)$

then, $\lambda = -\left[x^T \theta + \theta_0 \right)$

$$\Rightarrow \lambda = -\left[\theta^T x + \theta_0 \right] \quad -(5)$$

$\because$ $x^T \theta = \theta^T x$ which is a 1×1 matrix.
i.e. $(x^T \theta)^T = (x^T \theta)$
$= \theta^T x = x^T \theta$.

Now, put eqⁿ (5) in eqⁿ (4)

$$\Rightarrow P(y=1/x) = \frac{1}{1 + \exp(-(\theta^T x + \theta_0))}, \quad \theta = \Sigma^{-1} (\mu_1 - \mu_0)$$

$$\theta_0 = \frac{1}{2} \mu_0 \Sigma^{-1} \mu_0 - \frac{1}{2} \mu_1^T \Sigma^{-1} \mu_1 - \log\left(\frac{1-\phi}{\phi}\right)$$

Hence proved.