Q1: a. In this case, the loss for making error is λ_s, and the loss of rejection is λ_r. So the optimal discriminant function will base on the Bayes classifier, which means:

$$g_i(x) = -R(\alpha_i|x)$$

When $i = 1, \dots, c$,

$$-R(\alpha_i|x) = -\left[\sum_{j=1}^{c} \lambda(\alpha_i|w_j) P(w_j|x)\right]$$, Since loss of making error is λ_s, $R(\alpha_i|x) = \sum_{j=1}^{c} \lambda_s P(w_j|x) - \lambda_s P(w_i|x)$

$$-R(\alpha_i|x) = -\left[\lambda_s - \lambda_s P(w_i|x)\right] = \lambda_s P(w_i|x) - \lambda_s$$

When $i = c+1$,

$$-R(\alpha_{c+1}|x) = -\left[\sum_{j=1}^{c} \lambda_r P(w_j|x)\right] = -\lambda_r$$

To compare $g_i(x)$ and $g_j(x)$, there are two situation:

①. i and j both in classes, $g_i(x) > g_j(x)$ means:

$$\lambda_s P(w_i|x) - \lambda_s > \lambda_s P(w_j|x) - \lambda_s$$

$$\Rightarrow P(w_i|x) > P(w_j|x)$$

$$\Rightarrow P(x|w_i)P(w_i) > P(x|w_j)P(w_j)$$, decide i

otherwise, j

②. $i = c+1$ and j in classes, $g_i(x) > g_j(x)$ means:

$$-\lambda_r > \lambda_s P(w_j|x) - \lambda_s$$

$$\Rightarrow \lambda_s - \lambda_r > \lambda_s P(w_j|x)$$

$$\Rightarrow \cancel{\frac{\lambda_s-\lambda_r}{\lambda_s} > P(x|w_j)P(w_j) / \sum_{i=1}^{c}}$$

$$\Rightarrow \frac{\lambda_s - \lambda_r}{\lambda_s} > \frac{P(x|w_j)P(w_j)}{P(x)}$$

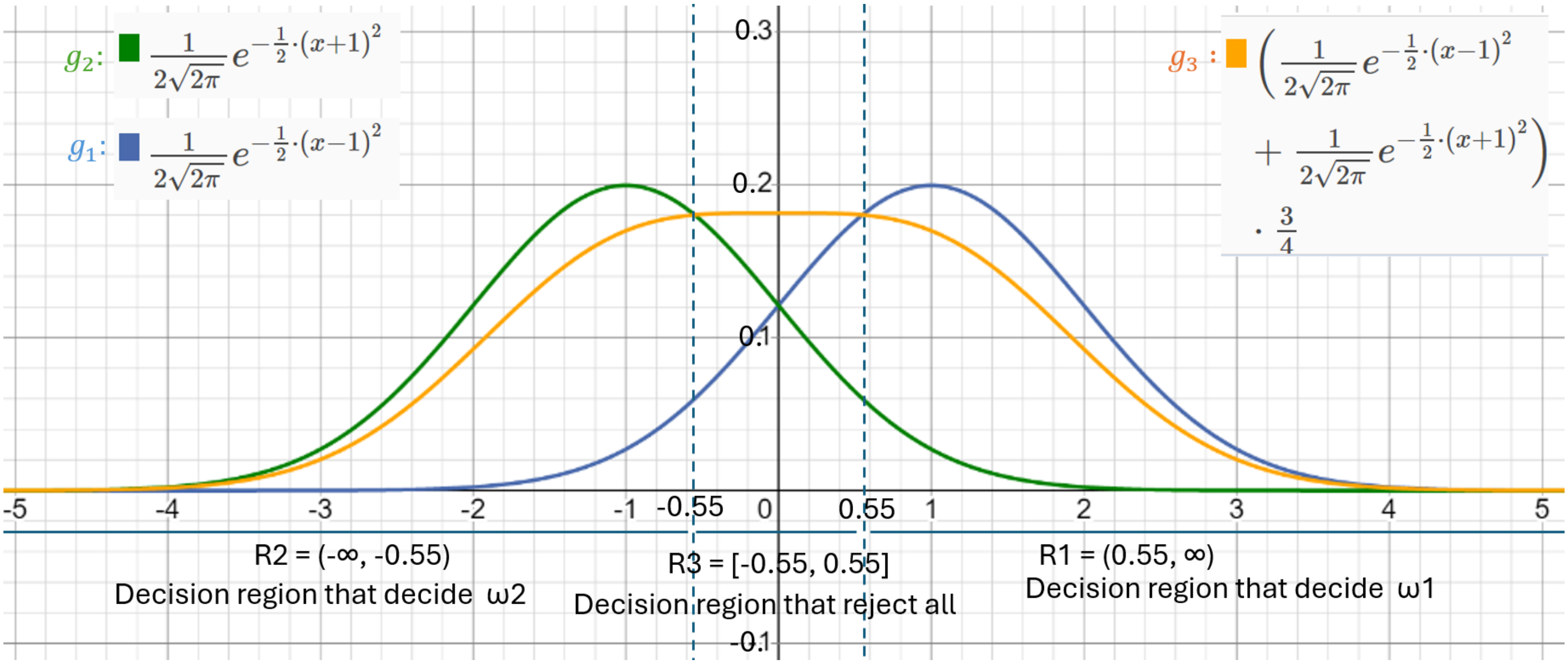
$$\Rightarrow \frac{\lambda_s - \lambda_r}{\lambda_s} \sum_{j=1}^{c} p(x|w_j)p(w_j) > p(x|w_j)p(w_j)$$

For discriminant function, we will try to make decision base on a given x. So in this question it is not only consider the comparison between classes.

According to result in Problem 13, decide w_i if $P(w_i|x) \geq P(w_j|x)$ for all classes and $P(w_i|x) \geq 1 - \frac{\lambda_r}{\lambda_s}$ rejection otherwise.

And $P(w_i|x) = \frac{P(x|w_i)P(w_i)}{\sum_{j=1}^{c} P(x|w_j)P(w_j)}$. So we used $P(x|w_i)P(w_i)$ for $g_i(x)$ when $i = 1, \dots, c$ and $\frac{\lambda_s - \lambda_r}{\lambda_s}\sum_{j=1}^{c} P(x|w_j)P(w_j)$ when $i = c+1$.

Q1. B



Q1. C

As λr/λs is increased from 0 to 1, the decision region of rejection (in this case is R3) decreases from infinite to zero, and the decision region for classes will expand. In a two-classes category, it will show following:

When λr/λs is equal to 0, the decision region of R3 is from -∞ to ∞,

and the decision for other choices (in this case is R1 and R2) do not exist.

When 0 < λr/λs < ½ , R3 exists and decreases,

and the decision for other choices (in this case is R1 and R2) will appear and expand.

When ½ ≤ λr/λs < 1, R3 do not exist, and the decision for other choices (in this case is R1 and R2) will continuously expand.

When λr/λs = 1, the discriminant function of rejection is equal to 0.

Q2. A

Q2. a

~~Th~~ For minimum-error-rate case, the discriminant function $g_i(x) = P(w_i|x)$

In this case, $\begin{cases} g_1 = P(w_1|x) \\ g_2 = P(w_2|x) \end{cases}$

To find decision rule and decide w_1, ~~wh~~ we should find $g(x) = g_1(x) - g_2(x) > 0$

Which means: $P(w_1|x) - P(w_2|x) > 0$

$$p(x|w_1)P(w_1) - p(x|w_2)P(w_2) > 0 \quad , x \in [0, 1]$$

$$\frac{2}{5} - 2x \cdot \frac{3}{5} > 0$$

$$x < \frac{1}{3}$$

So the decision rule is:

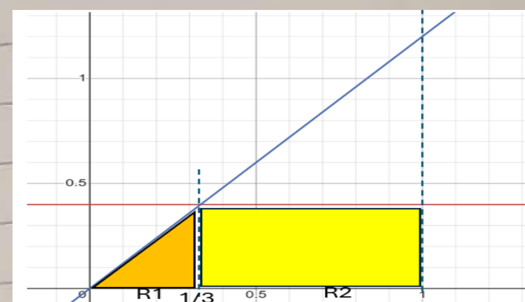
Decide w_1 if $x \in [0, \frac{1}{3})$; Otherwise decide w_2

The decision region $R_1 = [0, \frac{1}{3})$, $R_2 = [\frac{1}{3}, 1]$

Q2.B

Q2.b.

To find error rate in (a), we will decide ω_1 in R_1 and decide ω_2 in R_2.



The error rate will be:

$$\int_{R_1} R(\alpha_1|x)\,p(x)\,dx + \int_{R_2} R(\alpha_2|x)\,p(x)\,dx$$

$$= \int_{R_1} \left[\lambda_{11} P(\omega_1) p(x|\omega_1) + \lambda_{12} P(\omega_2) p(x|\omega_2)\right] dx$$

$$+ \int_{R_2} \left[\lambda_{21} P(\omega_1) p(x|\omega_1) + \lambda_{22} P(\omega_2) p(x|\omega_2)\right] dx, \qquad \lambda_{11} = \lambda_{22} = 0, \quad \lambda_{21} = \lambda_{12} = 1$$

$$= \int_{R_1} P(\omega_2) P(x|\omega_2)\,dx + \int_{R_2} P(\omega_1) p(x|\omega_1)\,dx$$

$$= \int_0^{\frac{1}{3}} 2x \cdot \frac{3}{5}\,dx + \cancel{\int_{\frac{1}{3}}^{1} \frac{5}{9}\,dx} + \int_{\frac{1}{3}}^{1} \frac{2}{5}\,dx = \frac{1}{3}$$

The error rate is $\frac{1}{3}$

Q3.

To find decision rule, we should ~~find~~ decide

ω_1 when $g_1(x) > g_2(x)$. And $g_i(x) = -R(\alpha_i|x)$

So that means: $R(\alpha_1|x) < R(\alpha_2|x)$

$$\lambda_{11} P(\omega_1|x) + \lambda_{12} P(\omega_2|x) < \lambda_{21} P(\omega_1|x) + \lambda_{22} P(\omega_2|x)$$

And $\lambda_{11} = \lambda_{22} = 0$,

$$\lambda_{12} P(x|\omega_2) P(\omega_2) < \lambda_{21} P(x|\omega_1) P(\omega_1)$$

$$\frac{2\lambda_{12}}{\lambda_{21}} < \frac{P(x|\omega_1)}{P(x|\omega_2)}$$

$$\frac{2\lambda_{12}}{\lambda_{21}} < EXP\left[-\frac{(x-\mu_1)^2 - (x-\mu_2)^2}{2\sigma^2}\right]$$

$$\ln(2) + \ln(\lambda_{12}) - \ln(\lambda_{21}) < \frac{2x(\mu_1-\mu_2) + \mu_2^2 - \mu_1^2}{2\sigma^2}$$

$$x(\mu_1-\mu_2) > \frac{1}{2} \cdot \left[2\sigma^2(\ln(2) + \ln(\lambda_{12}) - \ln(\lambda_{21})) + \mu_1^2 - \mu_2^2\right]$$

So the decision boundary is

$$x_0 = \frac{1}{2(\mu_1-\mu_2)}\left[2\sigma^2 \cdot \left(\ln\left(2 \cdot \frac{\lambda_{12}}{\lambda_{21}}\right)\right) + \mu_1^2 - \mu_2^2\right]$$

~~If $(\mu_1 - \mu_2) > 0$~~

If $\mu_1 > \mu_2$, then $R_1 = (x_0, \infty)$

decide w_1 if $x \in R_1$, otherwise decide w_2

If $\mu_1 < \mu_2$, then $R_1 = (-\infty, x_0)$

decide w_1 if $x \in R_0$, otherwise decide w_2

Q4.A

Q4. A

$$P(W_0 \mid X_1) = \frac{\sum_{j=0}^{1}\sum_{i=0}^{1} P(X_1, y_i, z_j, W_0)}{P(X_1)}$$

$$\sum_{j=0}^{1}\sum_{i=0}^{1} P(X_1, y_i, z_j, W_0)$$

$$= P(X_1, y_0, z_0, W_0) + P(X_1, y_0, z_1, W_0)$$

$$+ P(X_1, y_1, z_0, W_0) + P(X_1, y_1, z_1, W_0)$$

In Bayes network, since there are conditional independent in variables,

$$\sum_{j=0}^{1}\sum_{i=0}^{1} P(X_1, y_i, z_j, W_0)$$

$$= \sum_{j=0}^{1}\sum_{i=0}^{1} P(X_1) P(y_i \mid X_1) P(z_j \mid y_i) P(W_0 \mid z_j)$$

$$= P(X_1) P(y_0 \mid X_1) P(z_0 \mid y_0) P(W_0 \mid z_0) + P(X_1) P(y_0 \mid X_1) P(z_1 \mid y_0) P(W_0 \mid z_1)$$

$$+ P(X_1) P(y_1 \mid X_1) P(z_0 \mid y_1) P(W_0 \mid z_0) + P(X_1) P(y_1 \mid X_1) P(z_1 \mid y_1) P(W_0 \mid z_1)$$

$$\cancel{=0.6(} \quad \begin{cases} P(y_0 \mid X_1) = 1 - P(y_1 \mid X_1) = 0.6 \\ P(z_0 \mid y_0) = 1 - P(z_1 \mid y_0) = 0.4 \end{cases}$$

? ... Auording to the provided conditional probability,

$$\sum_{j=0}^{1}\sum_{i=0}^{1} P(X_1, y_i, z_j, W_0) = 0.3786.$$

$$P(W_0 \mid X_1) = \frac{0.3786}{P(X_1)} = 0.631$$

Q4.B

Q4. b

$$p(x_0|w_1) = \frac{p(w_1|x_0)\,p(x_0)}{\sum_{i=0}^{1} p(w_1|x_i)\,p(x_i)}$$

$$p(x_0)p(w_1|x_0) = \sum_{i=0}^{1}\sum_{j=0}^{1} p(x_0, y_i, z_j, w_1)$$

$$= \sum_{i=0}^{1}\sum_{j=0}^{1} p(x_0)\,p(y_i|x_0)\,p(z_j|y_i)\,p(w_1|z_j)$$

$$= 0.1497$$

$$p(x_1)p(w_1|x_1) = \sum_{j=0}^{1}\sum_{i=0}^{1} \cancel{p(x_1, w)}\; p(x_1, y_i, z_j, w_1)$$

$$= \sum_{j=0}^{1}\sum_{i=0}^{1} p(x_1)\,p(y_i|x_1)\,p(z_j|y_i)\,p(w_1|z_j)$$

$$= 0.2214$$

$$p(x_0|w_1) = \frac{0.1497}{0.1497+0.2214} = 0.403$$

Q5.

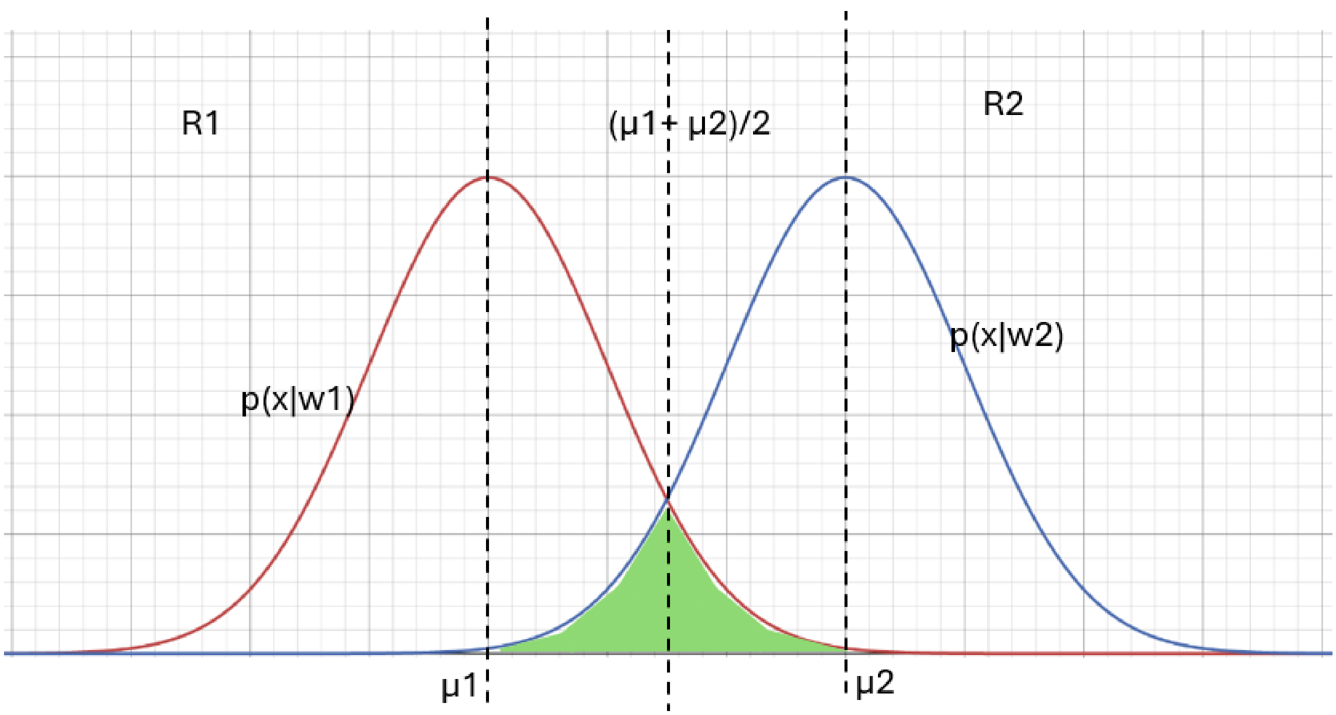
The answer to this question is False.

The decision boundary for a two-class classification which class conditionals are Gaussian density will have multiple situations based on their covariance matrix.

When the covariance matrixes for both Gaussian densities are equal, their decision boundary is linear (hyperplanes). Especially, when the covariance matrix is diagonal, this hyperplane is orthogonal to the line linking the means.

When the covariance matrixes for Gaussian densities are not equal, the decision boundary is hyperquadratics. It can be a surface other than linear (hyperplanes). So the answer is False.

Q6.A



To find the minimum probability of error in a two-category problem means finding the error of decision ω_1 in region R_2 and the error of decision ω_2 in decision region R_1.

In this problem, the decision boundary is:

$$p(x|\omega_1)\,p(\omega_1) = p(x|\omega_2)\,p(\omega_2)$$

$$p(x|\omega_1)\,p(\omega_1) - p(x|\omega_2)\,p(\omega_2) = 0$$

$$\frac{1}{\sqrt{2\pi}\sigma} EXP\left[-\frac{(x-\mu_1)^2}{2\sigma^2}\right]\cdot\frac{1}{2} - \frac{1}{\sqrt{2\pi}\sigma} EXP\left[-\frac{(x-\mu_2)^2}{2\sigma^2}\right]\cdot\frac{1}{2} = 0,$$ ~~since they~~

Since they has the same variance and same piror probability, decision boundary is $(x-\mu_1)^2 - (x-\mu_2)^2 = 0$

$$x = \frac{\mu_1+\mu_2}{2}$$

calculate the minimum probability of error:

there will have two situation: ①. $\mu_1 > \mu_2$

$$P_e = \int_{-\infty}^{\frac{\mu_1+\mu_2}{2}} p(x|\omega_1)\,p(\omega_1)\,dx + \int_{\frac{\mu_2+\mu_1}{2}}^{\infty} p(x|\omega_2)\,p(\omega_2)\,dx$$

Left side of $\frac{\mu_1+\mu_2}{2}$ will be R_2, otherwise is R_1

$$P_e = \frac{1}{2}\left[\int_{-\infty}^{\frac{\mu_1+\mu_2}{2}} \frac{1}{\sqrt{2\pi}} Exp\left(-\frac{(x-\mu_1)^2}{2\sigma^2}\right)dx + \int_{\frac{\mu_1+\mu_2}{2}}^{\infty} \frac{1}{\sqrt{2\pi}} EXP\left(-\frac{(x-\mu_2)^2}{2\sigma^2}\right)dx\right]$$

then standardized :

$$P_e = \frac{1}{2}\left[P\left(x \le \frac{\mu_1+\mu_2}{2} \middle| \omega_1\right) + P\left(x > \frac{\mu_1+\mu_2}{2} \middle| \omega_2\right) \right]$$

$$= \frac{1}{2}\left[\varphi_0\left(x \le \frac{\frac{\mu_1+\mu_2}{2}-\mu_1}{\delta}\right) + 1 - \varphi_0\left(x < \frac{\frac{\mu_1+\mu_2}{2}-\mu_2}{\delta}\right) \right]$$

$$= \frac{1}{2}\left[\Phi\left(\frac{\mu_2-\mu_1}{2\delta}\right) + \left(1-\Phi\left(\frac{\mu_1-\mu_2}{2\delta}\right)\right) \right]$$

Since $\frac{\mu_2-\mu_1}{2\delta} = -\frac{\mu_1-\mu_2}{2\delta}$, $\Phi\left(\frac{\mu_2-\mu_1}{2\delta}\right) = 1-\Phi\left(\frac{\mu_1-\mu_2}{2\delta}\right)$

$$P_e = \Phi\left(\frac{\mu_2-\mu_1}{2\delta}\right) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{a} e^{-\frac{t^2}{2}}dt, \quad a = \frac{\mu_2-\mu_1}{2\delta}$$

$$= \frac{1}{\sqrt{2\pi}}\int_{-a}^{\infty} e^{-\frac{t^2}{2}}dt, \quad -a = \frac{\mu_1-\mu_2}{2\delta}$$

② $\mu_1 < \mu_2$

Left side ($x < \frac{\mu_1+\mu_2}{2}$) will be R_1, $x > \frac{\mu_1+\mu_2}{2}$ will be R_2.

$$P_e = \frac{1}{2}\left[\int_{-\infty}^{\frac{\mu_1+\mu_2}{2}} \frac{1}{\sqrt{2\pi}} EXP\left(-\frac{(x-\mu_2)^2}{2\delta^2}\right)dx + \int_{\frac{\mu_1+\mu_2}{2}}^{\infty} \frac{1}{\sqrt{2\pi}} EXP\left(-\frac{(x-\mu_1)^2}{2\delta^2}\right)dx \right]$$

then ~~new~~ standardized:

$$P_e = \frac{1}{2}\left(\Phi\left(\frac{\mu_1-\mu_2}{2\delta}\right) + \left(1-\Phi\left(\frac{\mu_2-\mu_1}{2\delta}\right)\right) \right) = \Phi\left(\frac{\mu_1-\mu_2}{2\delta}\right)$$

$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{-a} e^{-\frac{t^2}{2}}dt, \quad -a = \frac{\mu_1-\mu_2}{2\delta}$$

$$= \frac{1}{\sqrt{2\pi}}\int_{a}^{\infty} e^{-\frac{t^2}{2}}dt, \quad a = \frac{\mu_2-\mu_1}{2\delta}$$

In both cases, a will be ~~the~~ a positive number.

when $\mu_1 > \mu_2$, $\frac{\mu_1 - \mu_2}{2\delta} > 0$

when $\mu_2 > \mu_1$, $\frac{\mu_2 - \mu_1}{2\delta} > 0$.

So $a = \frac{|\mu_1 - \mu_2|}{2\delta}$

Q6.B

As $\frac{|\mu_2 - \mu_1|}{\delta}$ goes to infinity,

$$\frac{1}{\sqrt{2\pi}\frac{|\mu_2 - \mu_1|}{2\cdot\delta}} Exp \frac{-\left(\frac{|\mu_2 - \mu_1|}{2\delta}\right)^2}{2}.$$ goes to zero.

Since $P_e \leq \frac{1}{\sqrt{2\pi}a} Exp \frac{-a^2}{2}$, P_e goes to zero