Tarea 4, Reconocimiento de Patrones

Francisco Javier Peralta Ramírez

1. Calcula el clasificador bayesiano optimo para Y. Podemos calcular $P(X), P(Y)$ y usamos los datos de la tabla, los cuales muestran los valores para $P(X \cap Y)$

$$
\begin{aligned}
P(X=0) &= \cup_{y=0}^{2} P(X=0 \cap Y=y) \\
P(X=0) &= P(X=0 \cap Y=0) \\
&\quad + P(X=0 \cap Y=1) \\
P(X=0) &= 0.2 + 0.15 = 0.35
\end{aligned}
$$

$$
\begin{aligned}
P(X=1) &= \cup_{y=0}^{2} P(X=1 \cap Y=y) \\
P(X=1) &= P(X=1 \cap Y=0) \\
&\quad + P(X=1 \cap Y=1) \\
P(X=1) &= 0.15 + 0.1 = 0.25
\end{aligned}
$$

$$
\begin{aligned}
P(X=2) &= \cup_{y=0}^{2} P(X=2 \cap Y=y) \\
P(X=2) &= P(X=2 \cap Y=0) \\
&\quad + P(X=2 \cap Y=1) \\
P(X=2) &= 0.3 + 0.1 = 0.4
\end{aligned}
$$

Usando la definición de las propiedades condicionales tenemos que $P(Y=y|X=x) = P(X=x \cap Y=x)/P(X)$ con lo que obtenemos

$$
\begin{aligned}
P(Y=1|X=0) &= \frac{P(X=0 \cap Y=1)}{P(X=0)} \\
&= \frac{0.2}{0.35} = 0.57 \\
P(Y=2|X=0) &= \frac{P(X=0 \cap Y=2)}{P(X=0)} \\
&= \frac{0.15}{0.35} = 0.43 \\
&\rightarrow \text{Tomamos } Y=1
\end{aligned}
$$

$$
\begin{aligned}
P(Y=1|X=1) &= \frac{P(X=1 \cap Y=1)}{P(X=1)} \\
&= \frac{0.15}{0.25} = 0.6 \\
P(Y=2|X=1) &= \frac{P(X=1 \cap Y=2)}{P(X=1)} \\
&= \frac{0.1}{0.25} = 0.4 \\
&\rightarrow \text{Tomamos } Y=1
\end{aligned}
$$

$$
\begin{aligned}
P(Y=1|X=2) &= \frac{P(X=2 \cap Y=1)}{P(X=2)} \\
&= \frac{0.3}{0.4} = 0.75 \\
P(Y=2|X=2) &= \frac{P(X=2 \cap Y=2)}{P(X=2)} \\
&= \frac{0.1}{0.4} = 0.25 \\
&\rightarrow \text{Tomamos } Y=1
\end{aligned}
$$

La probabilidad de cometer un error está dada por

$$P(Y=2|X=0)P(X=0) + P(Y=2|X=1)P(X=1) + P(Y=2|X=2)P(X=2)$$

en este caso, como sólo escojemos $Y=1$ el error es $P(Y=2) = 0.15 + 0.1 + 0.1 = 0.35$

2. Deriva el clasificador Bayesiano Óptimo para el caso de dos clases y una funció de costo simétrica cuando:

$$P(X=x|Y=1) \sim \mathcal{N}(\mu_1, \Sigma) \qquad P(X=x|Y=2) \sim \mathcal{N}(\mu_2, \Sigma)$$

y

$$P(Y=1) = 2P(Y=2)$$

La regla de Bayes nos dice:

$$P(Y=y|X=x) = \frac{P(X=x|Y=y)}{P(X)} P(Y=y)$$

Ya que la función de costo es simétrica, tomamos $Y=1$ si $P(Y=1|X=x) > P(Y=2|X=x)$ y $Y=2$ de lo contrario. Con esto tenemos la desigualdad

$$
\begin{aligned}
P(Y=1|X=x) &> P(Y=2|X=x) && (1) \\
\frac{P(X=x|Y=1)}{P(X)} P(Y=1) &> \frac{P(X=x|Y=2)}{P(X)} P(Y=2) && (2) \\
P(X=x|Y=1)P(Y=1) &> P(X=x|Y=2)P(Y=2) && (3) \\
2P(X=x|Y=1) &> P(X=x|Y=2) && (4) \\
2\mathcal{N}(\mu_1, \Sigma) &> \mathcal{N}(\mu_2, \Sigma) && (5) \\
2\mu_1 + 2\mathcal{N}(0, \Sigma) &> \mathcal{N}(0, \Sigma) + \mu_2 && (6) \\
\mathcal{N}(0, \Sigma) &> \mu_2 - 2\mu_1 && (7)
\end{aligned}
$$

Podemos calcular $P(X)$ usando

$$P(X = x) = P(X = x|Y = 1)P(Y = 1) + P(X = x|Y = 2)P(Y = 2)$$
$$P(X = x) = \mathcal{N}(\mu_1, \Sigma)P(Y = 1) + \mathcal{N}(\mu_2, \Sigma)P(Y = 2)$$
$$P(X = x) = \mathcal{N}(\mu_1, \Sigma)\frac{2}{3} + \mathcal{N}(\mu_2, \Sigma)\frac{1}{3}$$
$$P(X = x) = \frac{2\mu_1}{3} + \mathcal{N}(0, \Sigma)\frac{2}{3} + \frac{\mu_2}{3} + \mathcal{N}(0, \Sigma)\frac{1}{3}$$
$$P(X = x) = \frac{1}{3}(2\mu_1 + \mu_2) + \mathcal{N}(0, \Sigma)$$
$$X \sim \mathcal{N}(\frac{1}{3}(2\mu_1 + \mu_2), \Sigma)$$

En la ecuación *7* podemos sumar $\frac{1}{3}(2\mu_1 + \mu_2)$ a ambos lados

$$\mathcal{N}(0, \Sigma) + \frac{1}{3}(2\mu_1 + \mu_2) > \mu_2 - 2\mu_1 + \frac{1}{3}(2\mu_1 + \mu_2)$$
$$X > \mu_2 - 2\mu_1 + \frac{1}{3}(2\mu_1 + \mu_2)$$
$$X > \frac{1}{3}(3\mu_2 - 6\mu_1 + 2\mu_1 + \mu_2)$$
$$X > \frac{4}{3}(\mu_2 - \mu_1)$$

En otras palabras, escogemos $Y = 1$ cuando $X > \frac{4}{3}(\mu_2 - \mu_1)$

3. Considera un problema de clasificación de X en dos categorías, $Y = 0$ y $Y = 1$. Supongamos que $P(Y = 0) = P(Y = 1)$ y que $P(X|Y = i)$ sigue una distribución Poisson con parámetro λ_i.

   Derive el clasificador Bayesiano óptimo si el costo de un falso positivo es dos veces el costo de un falso negativo.

   Recordamos que la distribución Poisson está dada por

$$f(x, \lambda) = \frac{e^{-\lambda}\lambda^x}{x!}$$

   Bajo el clasificador Bayesiano óptimo escogeremos $Y = 1$ cuando

$$P(Y = 1|X = x) > 2P(Y = 0|X = x)$$
$$P(X = x|Y = 1)\frac{P(Y = 1)}{P(X = x)} > 2P(X = x|Y = 0)\frac{P(Y = 0)}{P(X = x)}$$
$$P(X = x|Y = 1) > 2P(X = x|Y = 0)$$
$$\frac{e^{-\lambda_1}\lambda_1^x}{x!} > 2\frac{e^{-\lambda_0}\lambda_0^x}{x!}$$
$$e^{-\lambda_1}\lambda_1^x > 2e^{-\lambda_0}\lambda_0^x$$
$$\frac{e^{-\lambda_1}}{2e^{-\lambda_0}} > \frac{\lambda_0^x}{\lambda_1^x}$$
$$\frac{e^{\lambda_0 - \lambda_1}}{2} > \left(\frac{\lambda_0}{\lambda_1}\right)^x$$
$$\log\left(\frac{e^{\lambda_0 - \lambda_1}}{2}\right) > \log\left[\left(\frac{\lambda_0}{\lambda_1}\right)^x\right]$$
$$\log(e^{\lambda_0 - \lambda_1}) - \log(2) > x(\log\lambda_0 - \log\lambda_1)$$
$$\frac{\lambda_0 - \lambda_1 - \log(2)}{\log\lambda_0 - \log\lambda_1} > x$$

   Cuando $\frac{\lambda_0 - \lambda_1 - \log(2)}{\log\lambda_0 - \log\lambda_1} < x$ escogemos $Y = 1$ y cuando $\frac{\lambda_0 - \lambda_1 - \log(2)}{\log\lambda_0 - \log\lambda_1} < x$ escogemos $Y = 0$