Assg. 1 Q1 →

Given: $X \sim N(\mu_1, \Sigma_1)$ for class ω_1
$X \sim N(\mu_2, \Sigma_2)$ for class ω_2

Linear Trans. → $Y = AX + b$ → k-dim. vector
$\downarrow$
$k \times d$ mat.

Now, if $X \sim N(\mu, \Sigma)$ then Y is also Gaussian with,
$Y \sim N(A\mu_1 + b, A\Sigma_1 A^T)$ for ω_1
and, $Y \sim N(A\mu_2 + b, A\Sigma_2 A^T)$ for ω_2

Now, the discriminant functions $g_i(Y)$ are proportional to log ~~likelihood~~ of posterior prob.,

$$\Rightarrow g_i(Y) = \ln(P(\omega_i)) - \frac{1}{2}\ln|(2\pi)^{d/2}\Sigma_i'| - \frac{1}{2}(Y-\mu_i')^T \Sigma_i'^{-1}(Y-\mu_i')$$

where $\mu_i' = A\mu_i + B$
$\Sigma_i' = A\Sigma_i A^T$ } Acc. to our transformation.

Now, to find the decision boundary we set $g_1(Y) = g_2(Y)$ in the above eqn, which simplifies to,

.5

$$\Rightarrow \ln P(\omega_1) - \ln P(\omega_2) = \frac{1}{2}\ln\frac{|\Sigma_2'|}{|\Sigma_1'|} \cdot \frac{1}{2}\left[(Y-\mu_1')^T\Sigma_1'^{-1}(Y-\mu_1') + (Y-\mu_2')^T\Sigma_2'^{-1}(Y-\mu_2')\right]$$

The eqn above defines a quadratic surface in Y on the RHS and the LHS is the log of the prior ratio.

Specifically, Y: $Y^T(\Sigma_1'^{-1} - \Sigma_2'^{-1})Y + 2[\Sigma_2'^{-1}\mu_2' - \Sigma_1'^{-1}\mu_1]^T Y$

~~+ constant = 0~~

.5

$+ (\mu_2'^T\Sigma_2'^{-1}\mu_2' - \mu_1'^T\Sigma_1'^{-1}\mu_1') + 2\ln\frac{P(\omega_2)}{P(\omega_1)} = 0$ → quadratic RHS in terms of Y (decision boundary)

Now, in the special case of equal covariances,

$$\Sigma_1 = \Sigma_2 = \Sigma$$

$$\Rightarrow \Sigma_1' = \Sigma_2' = A \Sigma A^T$$

$\therefore$ The boundary simplifies to a linear eqn in

$$Y: (\mu_1' - \mu_2')^T \Sigma'^{-1} Y = \frac{\ln P(\omega_2)}{\ln P(\omega_1)} \cancel{\text{constant}}$$

$$+ \underbrace{\frac{1}{2}\left(\mu_1'^T \Sigma_1'^{-1} \mu_1 - \mu_2'^T \Sigma_2'^{-1} \mu_2'\right)}_{\text{constant}}$$

Q2 Poisson Distribution: Properties & Bayes classification

$$P(x|\lambda) = \frac{e^{-\lambda}\lambda^{x}}{x!}$$

Discrete distribution
$x = 0, 1, 2, - -$ $\lambda \in \mathbb{R}$

(a) To prove: $E[x] = \lambda$

$$E[x] = \sum_{x} x P(x)$$

$$= \sum_{x=0}^{\infty} x . \frac{e^{-\lambda}\lambda^{x}}{x!} = \sum_{x=1}^{\infty} x . \frac{e^{-\lambda}\lambda^{x}}{x!}$$

$$= e^{-\lambda} \sum_{x=1}^{\infty} \cancel{e} \frac{\lambda^{x}}{x!} . x$$

$$e^{-\lambda}\left[\frac{\lambda}{1!} + \frac{\lambda^{2}.2}{2!} + \frac{\lambda^{3}.3}{3!} + \frac{\lambda^{4}.4}{4!} + - - - \right]$$

$$\lambda e^{-\lambda}\left[1 + \frac{2\lambda}{2!} + \frac{3\lambda^{2}}{3!} + - - \right]$$

$$\lambda e^{-\lambda}\left[\underbrace{1 + \frac{\lambda}{1!} + \frac{\lambda^{2}}{2!} + \frac{\lambda^{3}}{3!} + -}_{\text{Taylor expansion of } e^{\lambda}}\right] \qquad \frac{n}{n!} = \frac{1}{(n-1)!}$$

$$= \lambda \cancel{e^{-\lambda}} . \cancel{e^{\lambda}}$$

$$= \lambda$$

(b) To prove $Var[x] = \lambda$

$$Var[x] = E[(x - E[x])^2] = E[x^2] - (E[x])^2$$

from part (a) $E[x] = \lambda$

$$\Rightarrow Var[x] = E[x^2] - (\lambda)^2$$

$$= E[x^2] - \lambda^2$$

$E[x^2] = ?$

→

$$E[x^2] = \sum_{x} x^2 P(x) \qquad \left\{ E[f(x)] = \sum_{x} f(x)P(x) \right\}$$

$$= \sum_{x=0}^{\infty} x^2 \frac{e^{-\lambda}\lambda^x}{x!} = e^{-\lambda}\sum_{x=1}^{\infty} \frac{x^2\lambda^x}{x!}$$

$$\sum_{x=1}^{\infty} \frac{x^2\lambda^x}{x!} = \sum_{x=1}^{\infty} \frac{x\lambda^x}{(x-1)!}$$

$$= \frac{\lambda}{0!} + \frac{2\lambda^2}{1!} + \frac{3\lambda^3}{2!} + \frac{4\lambda^4}{3!} + \cdots$$

$$= (\lambda) + \left(\frac{\lambda^2}{1!} + \frac{\lambda^2}{1!}\right) + \left(\frac{\lambda^3}{2!} + \frac{\lambda^3}{2!} + \frac{\lambda^3}{2!}\right) + \cdots$$

$$= \lambda\left[1 + \frac{\lambda^2}{1!} + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \cdots\right] + \left(\frac{\lambda^2}{1!} + \frac{2\lambda^3}{2!} + \frac{3\lambda^4}{3!} + \cdots\right)$$

$$\lambda\left[1 + \frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \cdots\right] + \lambda^2\left[1 + \frac{2\lambda}{2!} + \frac{3\lambda^2}{3!} + \cdots\right]$$

$$\lambda\left[\underbrace{1 + \frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \cdots}\right] + \lambda^2\left[1 + \underbrace{\frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \cdots}\right]$$

Taylor expansion of e^{λ}

$$(\lambda e^{\lambda} + \lambda^2 e^{\lambda})$$

$$E[x^2] = e^{-\lambda}\sum_{x=1}^{\infty} \frac{x^2\lambda^x}{x!}$$

$$= \cancel{e^{-\lambda}}(\lambda + \lambda^2)\cancel{e^{\lambda}}$$

$$= \lambda + \lambda^2$$

$$Var[x] = E[x^2] - \lambda^2$$

$$= \lambda + \cancel{\lambda^2} - \cancel{\lambda^2}$$

$$= \lambda \quad ✓$$

(c) 2 equiprobable categories — ω_1, ω_2 $P(\omega_1) = P(\omega_2) = 1/2$

let parameter of $\omega_1 \to \lambda_1$

$\omega_2 \to \lambda_2$ given $\underline{\lambda_1 > \lambda_2}$

Bayes classification rule?

we will choose/classify a sample as class ω_1 if

$P(\omega_1|x) > P(\omega_2|x)$

$$P(x|\omega_1) = \frac{\lambda_1^x e^{-\lambda_1}}{x!}$$

By Baye's rule

$$P(x|\omega_2) = \frac{\lambda_2^x e^{-\lambda_2}}{x!}$$

classify ω_1 if $P(\omega_1|x) > P(\omega_2|x)$

$(x = 0, 1, 2 \ldots)$ (DISCRETE CASE)

ω_1 if $$\frac{P(x|\omega_1)P(\omega_1)}{\cancel{P(x)}} > \frac{P(x|\omega_2)P(\omega_2)}{\cancel{P(x)}}$$

$P(\omega_1) = P(\omega_2) = 1/2$ (given equiprobable)

ω_1 if $P(x|\omega_1) > P(x|\omega_2)$

$$\frac{e^{-\lambda_1}\lambda_1^x}{\cancel{x!}} > \frac{e^{-\lambda_2}\lambda_2^x}{\cancel{x!}}$$

$$e^{-\lambda_1}\lambda_1^x > e^{-\lambda_2}\lambda_2^x$$

take log → (since log is monotonically increasing)

$$-\lambda_1 + x\ln\lambda_1 > -\lambda_2 + x\ln\lambda_2$$

$$x\ln\lambda_1 - x\ln\lambda_2 > \lambda_1 - \lambda_2$$

$$x(\ln\lambda_1/\lambda_2) > \lambda_1 - \lambda_2$$

Classify x as ω_1 if $x > \dfrac{\lambda_1 - \lambda_2}{\ln(\lambda_1/\lambda_2)}$ 0.5

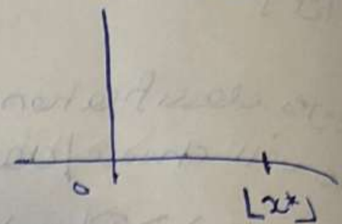
x as class ω_1 if $x > \dfrac{\lambda_1 - \lambda_2}{\ln(\lambda_1/\lambda_2)}$ Ans

otherwise class ω_2

(d) Baye's error rate

↳ probability of misclassification

Discrete Case



Decision boundary from part(c)

$$x^* = \frac{\lambda_1 - \lambda_2}{\ln(\lambda_1/\lambda_2)} \quad \text{(let this be } x^*)$$

$\lfloor x^* \rfloor$ = Smallest integer $\leq x^*$

we classify w_2 if $x \leq x^*$

↳ misclassification probability or probability of error $= P(x \leq x^* | w_1) P(w_1)$

(~~given~~ we classified w_2, true class w_1)

and we classify w_1 if $x > x^*$ i.e

error → $P(x > x^* | w_2) P(w_2)$

$P(w_1) = P(w_2) = 0.5$

Total error $= P(x \leq x^* | w_1) P(w_1) + P(x > x^* | w_2) P(w_2)$

$$= \frac{1}{2} \Big[\underbrace{P(x \leq x^* | w_1)}_{a} + \underbrace{P(x > x^* | w_2)}_{b} \Big]$$

$$P(x \leq x^* | \lambda_1) = \sum_{x=0}^{\lfloor x^* \rfloor} \frac{e^{-\lambda_1} \lambda_1^x}{x!} \quad \text{— a}$$

$$P(x > x^* | \lambda_2) = \sum_{\lceil x^* \rceil}^{\infty} \frac{e^{-\lambda_2} \lambda_2^x}{x!}$$

0.5

we know that $\displaystyle\sum_0^\infty \frac{e^{-\lambda_2} \lambda_2^x}{x!} = 1$

$$\sum_{\lceil x^* \rceil}^{\infty} \frac{e^{-\lambda_2} \lambda_2^x}{x!} = 1 - \sum_{x=0}^{\lfloor x^* \rfloor} \frac{e^{-\lambda_2} \lambda_2^x}{x!} \quad \text{— b}$$

Bayes error rate

$$= \boxed{\frac{1}{2} \left[\sum_{x=0}^{\lfloor x^* \rfloor} \frac{e^{-\lambda_1} \lambda_1^x}{x!} + 1 - \sum_{x=0}^{\lfloor x^* \rfloor} \frac{e^{-\lambda_2} \lambda_2^x}{x!} \right] \quad x^* = \frac{\lambda_1 - \lambda_2}{\ln(\lambda_1/\lambda_2)}}$$

Ans

Assigment - 1

Solution 3) The decision rule minimizes the expected risk:

$$R(\omega_i|x) = \sum_{j=1}^{2} \lambda_{ji}\, P(\omega_j|x)$$

A sample x is assigned to ω_1 if -

$$R(\omega_1|x) < R(\omega_2|x)$$

From eq. (17), we define risk based discriminant function as -

.5

$$g(x) = \ln P(x|\omega_1) - \ln P(x|\omega_2) + \ln\frac{P(\omega_1)}{P(\omega_2)} + \ln\frac{\lambda_{21} - \lambda_{11}}{\lambda_{21} - \lambda_{22}}$$

For independent features, we express likelihood ratio -

$$\ln P(x|\omega_1) - \ln P(x|\omega_2) = \sum_{i=1}^{d} \ln\frac{P(x_i|\omega_1)}{P(x_i|\omega_2)}$$

If we assume Bernoulli - distributed features where -

$p_i = P(x_i = 1|\omega_1)$

$q_i = P(x_i = 1|\omega_2)$

likelihood ratio simplifies to - $\sum_{i=1} \ln\frac{1-p_i}{1-q_i}$

final discriminent function is

$$g(x) = \omega^T x + \omega_0$$

where, $\omega_0 = \sum_{i=1}^{d} \ln\frac{1-p_i}{1-q_i} + \frac{\ln P(\omega_1)}{\ln P(\omega_2)} + \ln\frac{\lambda_{21} - \lambda_{11}}{\lambda_{12} - \lambda_{22}}$

.5

Q.4

- prior probabilities $P(\omega_1)$ and $P(\omega_2)$
- error matrix $\Lambda = \begin{pmatrix} 0 & \lambda_{12} \\ \lambda_{21} & 0 \end{pmatrix}$

- $R_1 = \{x \mid x \le x^*\}$ ... where we decide ω_1.
- $R_2 = \{x \mid x > x^*\}$ ... where we decide ω_2.
- we want to determine the rule (x^*) that minimizes the expected Risk.

- Bayes decision Rule chooses the class that minimizes the expected loss

$$R(\text{decide } \omega_1 \mid x) = \lambda_{12} \, P(\omega_2 \mid x)$$
$$R(\text{decide } \omega_2 \mid x) = \lambda_{21} \, P(\omega_1 \mid x)$$

At x, we compare

$$\lambda_{12} \, P(\omega_2 \mid x) \quad \text{vs} \quad \lambda_{21} \, P(\omega_1 \mid x)$$

if $\lambda_{12} P(\omega_2 \mid x) < \lambda_{21} P(\omega_1 \mid x)$ ... we choose ω_1

and if $\lambda_{12} P(\omega_2 \mid x) > \lambda_{21} P(\omega_1 \mid x)$ ... we choose ω_2

Boyes Rule :

$$p(\omega_j \mid x) = \frac{P(x \mid \omega_j) \, P(\omega_j)}{P(x)}$$

$$\therefore \quad \lambda_{12} P(\omega_2 \mid x) = \lambda_{12} \frac{P(x \mid \omega_2) \, P(\omega_2)}{P(x)} \quad \ldots (1)$$

$$\therefore \quad \lambda_{21} P(\omega_1 \mid x) = \lambda_{21} \frac{P(x \mid \omega_1) \, P(\omega_1)}{P(x)} \quad \ldots (2)$$

from 1 and 2

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

$$\frac{P(x \mid \omega_1)}{P(x \mid \omega_2)} > \frac{\lambda_{12} \times P(\omega_2)}{\lambda_{21} \times P(\omega_1)} \qquad \text{Decides } \omega_1$$

1

Otherwise decide ω_2.