سوال ۱ - (آ) : با روش اثبات بازگشتی به اثبات قضیه فوق می‌پردازیم :

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

$$\frac{P(x|w_1)P(w_1)}{P(x)} = \frac{P(x|w_2)P(w_2)}{P(x)} \xrightarrow{P(w_1)=P(w_2)}$$

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

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

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

$$\Longrightarrow \frac{x-a_1}{b} = \pm\left(\frac{x-a_2}{b}\right)$$

$$\Longrightarrow x-a_1 = -x+a_2$$

$$2x = a_1+a_2$$

$$x = \frac{a_1+a_2}{2}$$

باتوجه به معادله آخر ، اگر به صورت بازگشتی روابط را به سمت اولین معادله برگردیم قضیه $P(w_1|x)=P(w_2|x)$ اثبات می‌گردد .

$$P(error) = \int_{x\in R_2} p(w_1|x)\,p(x)\,dx + \int_{x\in R_1} P(w_2|x)\,P(x)\,dx$$

$\boxed{\text{سوال 1 (ب)}}$

$$= \int_{\substack{x\in R_2\\(a_2)}} P(x|w_1)P(w_1)\,dx + \int_{\substack{x\in R_1\\(a_1)}} P(x|w_2)\,P(w_2)\,dx$$

هرچه x کمتر از a_2 ← مربوط به کلاس (1) است.

هرچه x بیشتر از a_2 ← مربوط به کلاس (2) است.

$\boxed{a_1 < a_2}$

$$= \frac{1}{2\pi b}\left(\int_{-\infty}^{a_2} \frac{1}{1+\left(\frac{x-a_1}{b}\right)^2}dx + \int_{a_2}^{+\infty} \frac{1}{1+\left(\frac{x-a_2}{b}\right)^2}dx\right)$$

$$= \frac{1}{2\pi b}\left[b\,tg^{-1}\left(\frac{x-a_1}{b}\right) + c\Big|_{-\infty}^{a_2} + b\,tg^{-1}\left(\frac{x-a_2}{b}\right)\Big|_{a_2}^{+\infty}\right]$$

$$= \frac{1}{2\pi}\left[tg^{-1}\left(\frac{a_2-a_1}{b}\right) - \left(\frac{-\pi}{2}\right) + \underbrace{tg^{-1}(+\infty)}_{\frac{\pi}{2}} - \underbrace{tg^{-1}\left(\frac{a_2-a_2}{b}\right)}_{\text{صفر}}\right]$$

$$= \frac{1}{2\pi}\left[tg^{-1}\left(\frac{a_2-a_1}{b}\right) + \frac{\pi}{2} + \frac{\pi}{2} - 0\right]$$

$$= \frac{1}{2\pi}tg^{-1}\left(\frac{a_2-a_1}{b}\right) + \frac{1}{2} = \frac{1}{2} - \frac{1}{\pi}tg^{-1}\left(\frac{|a_2-a_1|}{b}\right)$$

سوال ۱ -(پ)

احتمال خطا زمانی بیشینه می شود که مرز تصمیم بین دو کلاس در نقطه‌ای قرار داده شود که دو تابع چگالی احتمال شرطی ($P(w_1|x)$) همدیگر را قطع کنند.

به عبارت هندسی تر، اگر دو تابع توزیع کوشی در نقطه تقاطع، ارتفاع برابری داشته باشند.

از لحاظ ریاضی، Classifier میزان احتمال برابری برای x های داده شده، برای واگذار کردن به کلاس ۱ یا ۲ احساس می کند.

$$P(w_1|x) = P(w_2|x) \Longrightarrow x = \frac{a_1 + a_2}{2}$$ مرز تصمیم $$\boxed{\theta = \frac{a_1 + a_2}{2}}$$

$$P(error) = \int_{x>\theta} \frac{1}{2\pi b} \frac{1}{1+\left(\frac{x-a_1}{b}\right)^2} + \int_{x<\theta} \frac{1}{2\pi b} \frac{1}{1+\left(\frac{x-a_2}{b}\right)^2}$$

$$P(error) = \frac{1}{2\pi b}\left[b\, tg^{-1}\left(\frac{x-a_1}{b}\right)\Big|_{\theta}^{\infty} + b\, tg^{-1}\left(\frac{x-a_2}{b}\right)\Big|_{-\infty}^{\theta} \right]$$

$$P(error) = \frac{1}{2\pi}\left[tg^{-1}(\infty) - \underbrace{tg^{-1}\left(\frac{\theta-a_1}{b}\right) + tg^{-1}\left(\frac{\theta-a_2}{b}\right)}_{\text{قرینه هستند}} - tg^{-1}(-\infty) \right]$$

$$P(error) = \frac{1}{2\pi}\left[\frac{\pi}{2} - \left(\frac{-\pi}{2}\right)\right] = \frac{1}{2\pi}[\pi] = \frac{1}{2}$$

سوال ① -(ت)

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

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

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

$$|x-a_1| = |x-a_2|$$

$$-x+a_1 = x-a_2$$

$$a_1+a_2 = 2x$$

$$\boxed{x = \frac{a_1+a_2}{2}}$$ مرز تصمیم

طبق سوال ① -(پ) برای محاسبه احتمال خطا خواهیم داشت :

$$P(error) = \frac{1}{2}$$

سوال ① - (ث)

$$\begin{pmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix}$$

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

$$\underbrace{P(\omega_1|x)\lambda_{11}}_{\text{صفر}} + P(\omega_2|x)\lambda_{12} = P(\omega_1|x)\lambda_{21} + \underbrace{P(\omega_2|x)\lambda_{22}}_{\text{صفر}}$$

$$P(\omega_2|x) = 2P(\omega_1|x)$$

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

$$|x-a_2| = \sqrt{2}\,|x-a_1|$$

$$x-a_2 = \sqrt{2}x - a_1\sqrt{2} \longrightarrow x = \frac{\sqrt{2}a_1 - a_2}{\sqrt{2}-1}$$

$$-x+a_2 = \sqrt{2}x - a_1\sqrt{2} \longrightarrow x = \frac{a_2 + a_1\sqrt{2}}{\sqrt{2}+1}$$

سوال ۲

$$\boxed{\lambda_{ii} \neq 0}$$

$$R(i|x) = \sum_{j=1}^{c} P(y=j|x)\,\lambda_{ij}$$

$$R(w_1|x) < R(w_2|x) :$$

ناحیه مربوط به کلاس اول

$$P(w_1|x)\lambda_{11} + P(w_2|x)\lambda_{12} < P(w_1|x)\lambda_{21} + P(w_2|x)\lambda_{22}$$

$$\frac{P(x|w_1)P(w_1)}{P(x)}\lambda_{11} + \frac{P(x|w_2)P(w_2)}{P(x)}\lambda_{12} < \frac{P(x|w_1)P(w_1)}{P(x)}\lambda_{21} + \frac{P(x|w_2)P(w_2)}{P(x)}\lambda_{22}$$

$$\underline{P(x|w_1)P(w_1)}\lambda_{11} + \underline{\underline{P(x|w_2)P(w_2)}}\lambda_{12} < \underline{P(x|w_1)P(w_1)}\lambda_{21} + \underline{\underline{P(x|w_2)P(w_2)}}\lambda_{22}$$

$$P(x|w_1)P(w_1)\left[\lambda_{11} - \lambda_{21}\right] < P(x|w_2)P(w_2)\left[\lambda_{22} - \lambda_{12}\right]$$

$$\frac{P(x|w_1)}{P(x|w_2)} < \frac{\lambda_{22} - \lambda_{12}}{\lambda_{11} - \lambda_{21}}\;\frac{P(w_2)}{P(w_1)}$$

بخش (آ)

سوال (۲)

ناحیه تصمیم کلاس (۱)

$$\frac{f(x|w_2)}{f(x|w_1)} = \frac{P(x|w_2)}{P(x|w_1)} \quad : \quad R(w_2|x) > R(w_1|x)$$

ناحیه تصمیم کلاس (۲)

$$P(w_1|x)\lambda_{21} + P(w_2|x)\lambda_{22} > P(w_1|x)\lambda_{11} + P(w_2|x)\lambda_{12}$$

$$\underline{P(x|w_1)P(w_1)}\lambda_{21} + \underline{\underline{P(x|w_2)P(w_2)}}\lambda_{22} > \underline{P(x|w_1)P(w_1)}\lambda_{11} + \underline{\underline{P(x|w_2)P(w_2)\lambda_{12}}}$$

$$P(x|w_1)P(w_1)\boxed{\lambda_{21} - \lambda_{11}} > P(x|w_2)P(w_2)\left[\lambda_{12} - \lambda_{22}\right]$$

$$\frac{P(w_1)}{P(w_2)}\left[\frac{\lambda_{21} - \lambda_{11}}{\lambda_{12} - \lambda_{22}}\right] > \frac{P(x|w_2)}{P(x|w_1)}$$

ناحیه تصمیم کلاس (۲)

سوال ۳

$$P(w_1|x) = P(w_2|x) \longrightarrow x = ?$$

$$\frac{P(x|w_1)\cancel{P(w_1)}}{\cancel{P(x)}} = \frac{P(x|w_2)\cancel{P(w_2)}}{\cancel{P(x)}} \quad ; P(w_1)=P(w_2)$$

$$\frac{x}{\sigma_1^2}\exp\left(-\frac{x^2}{2\sigma_1^2}\right)\cancel{P(w_1)} = \frac{x}{\sigma_2^2}\exp\left(-\frac{x^2}{2\sigma_2^2}\right)\cancel{P(w_2)}$$

$$\frac{\sigma_2^2}{\sigma_1^2} = \frac{\exp\left(\frac{-x^2}{2\sigma_2^2}\right)}{\exp\left(\frac{-x^2}{2\sigma_1^2}\right)}$$

$$\left(\frac{\sigma_2}{\sigma_1}\right)^2 = \exp\left(\frac{-x^2}{2\sigma_2^2} + \frac{x^2}{2\sigma_1^2}\right)$$

$$\mathrm{Ln}\left(\frac{\sigma_2}{\sigma_1}\right)^2 = x^2\left(\frac{1}{2\sigma_1^2} - \frac{1}{2\sigma_2^2}\right)$$

$$2\,\mathrm{Ln}\left(\frac{\sigma_2}{\sigma_1}\right) = \frac{x^2}{2}\left(\frac{\sigma_2^2 - \sigma_1^2}{\sigma_1^2\sigma_2^2}\right)$$

$$x = \sqrt{\frac{4\,\mathrm{Ln}\left(\frac{\sigma_2}{\sigma_1}\right)}{\frac{\sigma_2^2 - \sigma_1^2}{\sigma_1^2\sigma_2^2}}}$$

$N(x \mid \mu_1, \sigma_1)$ $\quad \sigma_1 = \sigma_2$ $\quad$ سوال ٤ $\quad$ مطلوب

$N(x \mid \mu_2, \sigma_2)$ $\quad \mu_1 \neq \mu_2 \longrightarrow \mu_1 < \mu_2 \longrightarrow \begin{cases} \theta < \mu_1 \\ \theta > \mu_2 \end{cases}$

$\max(\mu_1, \mu_2) < \theta$

$\theta < \min(\mu_1, \mu_2)$

~~$P(\ldots)$~~

$$P(\omega_1 \mid x) = P(\omega_2 \mid x)$$

$$\frac{P(x \mid \omega_1)\, P(\omega_1)}{\cancel{P(x)}} = \frac{P(x \mid \omega_2)\, P(\omega_2)}{\cancel{P(x)}}$$

$$\frac{1}{\sqrt{2\pi}\,\sigma_1} \exp\left(\frac{-1}{2} \frac{(x-\mu_1)^2}{\sigma_1^2}\right) = \frac{1}{\sqrt{2\pi}\,\sigma_2} \exp\left(\frac{-1}{2} \frac{(x-\mu_2)^2}{\sigma_2^2}\right)$$

$$\frac{-1}{2} (x-\mu_1)^2 = \frac{-1}{2} (x-\mu_2)^2$$

$$|x-\mu_1| = x-\mu_2$$

$$\begin{cases} x-\mu_1 = x-\mu_2 \quad \times \text{ تناقض} \\ -x+\mu_1 = x-\mu_2 \end{cases}$$

$$\mu_1 + \mu_2 = 2x$$

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

با هیچ شرایطی اتفاق مطلوب سوال رخ نخواهد داد.

$$\left.\begin{aligned} \frac{\mu_1+\mu_2}{2} &< \mu_1 \\ \frac{\mu_1+\mu_2}{2} &> \mu_2 \end{aligned}\right\} \longrightarrow \mu_1+\mu_2 < 2\mu_1 \longrightarrow \boxed{\mu_2 < \mu_1 \quad \times}$$

سوال ۶ الف

$$P(X;\theta) = P(x_1, x_2, \dots, x_N;\theta) = \prod_{k=1}^{N} P(x_k;\theta)$$

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

$$\hat{\theta}_{ML} = \arg\max_{\theta} \prod_{k=1}^{N} P(x_k;\theta)$$

$$L = \arg\max_{\theta} \prod_{k=1}^{N} P(x_k;\theta)$$

$$L = Ln\left(\prod_{j=1}^{N} \frac{\lambda^{x_j} e^{-\lambda}}{x_j!}\right)$$

$$L(\lambda; x_1, x_2, \dots, x_n) = \sum_{j=1}^{N} Ln\left(\frac{\lambda^{x_j} e^{-\lambda}}{x_j!}\right)$$

$$L(\lambda; x_1, x_2, \dots, x_n) = \sum_{j=1}^{N} Ln(\lambda^{x_j}) + Ln(e^{-\lambda_1}) - Ln(x_j!)$$

سوال ۶ - (ب) با توجه به قضیه زیر خواهیم داشت :

Gamma Prior $\longrightarrow$ Conjugate to Poisson distribution

$$P(\lambda|D) \propto \underbrace{P(D|\lambda)}_{\text{Poisson distribution}} \; \underbrace{P(\lambda)}_{\text{Gamma distribution}}$$

طبق قضیه بالا احتمال پسین را به صورت زیر خواهیم داشت :

$$P(\lambda|D) = \text{Gamma}(\lambda, \alpha', \beta')$$

$$\alpha' = \alpha + \sum_{i=1}^{N} x_i$$

$$\beta' = \beta + N$$

N : تعداد داده‌ها در D

سوال ۶ - (پ) بله ، یک احتمال پیشین درصورتی Conjugate خواهد بود که هم خانواده با احتمال پسین باشند .

درمثال فوق (۶-ب) احتمال پیشین از توزیع گاما ، Likelihood Function از نوع پواسون واحتمال پسین نیز از توزیع گاما می باشد که نتیجه می گیریم که احتمال پسین از نوع conjugate prior می باشد.

سوال ۶ - (ت) :

$$\underbrace{\text{Posterior}(\lambda|x)}_{\text{Gamma}(\lambda|\alpha', \beta')} \propto \text{Likelihood}(x|\lambda) \times \underbrace{\text{Prior}(\lambda)}_{\text{Gamma}(\lambda|\alpha, \beta)}$$

با توجه به این که می دانیم مقدار بیشینه در توزیع گاما به ازای $\lambda = \frac{\alpha - 1}{\beta}$ رخ می دهد ، به ازای α' و β' به روزرسانی شده نیز به ازای $\lambda' = \frac{\alpha' - 1}{\beta'}$ به بیشینه مقدار توزیع گاما خواهیم رسید .

$$\lambda' = \frac{\alpha' - 1}{\beta'} = \frac{\alpha + \sum_{i=1}^{N} x_i - 1}{\beta + N}$$

سوال ۶ -(ث): تحت شرایط خاصی جواب مثبت است.

با افزایش حجم نمونه‌ها، تأثیر توزیع پیشین کاهش می‌یابد و تابع احتمال پسین به تابع likelihood میل خواهد کرد.

تحت شرایطی مثل سازگاری تابع احتمال پیشین و پسین، یا همچنین Smooth بودن تابع توزیع احتمال، اتفاق مد نظر سوال رخ می‌دهد.

سوال ۶ - (ج): مواردی که MLE برتری دارد:
(Maximum Likelihood Estimator)

- تابع likelihood برای زمانی که مناسب رفتار می‌کند و مطابق توزیع داده عمل می‌کند مناسب است.
همچنین زمانی که حجم داده‌های بسیار زیادی در اختیار است، این روش مناسب است.

- زمانی که اطلاعات شفافی درباره توزیع پیشین نداریم و یا توزیع داده‌ها را نمی‌دانیم این روش مناسب‌تر است.

مواردی که MAP برتری دارد:
(Maximum a Posteriori)

- زمانی که داده‌های قابل اطمینان که بتوان توزیع پیشین آن را تشکیل داد، داشته باشیم، آن‌گاه ایجاد MAP گزینه بهتری خواهد بود که منجر به دقت بهتری خواهد شد.
- زمانی که میزان داده‌ها کم است و تابع likelihood به صورت شفاف ایجاد نشده باشد.

(*) در کل انتخاب بین این دو روش به قابل اتکا بودن داده‌ها از نظر توزیع، تعداد داده‌ها، میزان پیچیدگی تابع likelihood و ... وابسته خواهد بود.