Devoir – Parité démographique en classification supervisée

2025-2026
R. AZAÏS

1- $$R(g) = \mathbb{P}(g(X) \neq Y)$$
$$= \int_{\mathbb{R}} \mathbb{P}(g(X) \neq Y \mid X = x)\, f_X(dx)$$
$$= \int_{\mathbb{R}} \{1 - \mathbb{P}(Y = g(x) \mid X = x)\}\, f_X(dx)$$
$$= \int_{\mathbb{R}} \{1 - p_{g(x)}(x)\}\, f_X(dx)$$

2- On pose $g^*(x) = \underset{k}{\operatorname{argmax}}\ p_k(x)$

$$R(g) - R(g^*) = \int_{\mathbb{R}} \{p_{g^*(x)}(x) - p_{g(x)}(x)\}\, f_X(dx)$$

On $p_{g^*(x)}(x) = \max\limits_k p_k(x)$

donc $p_{g^*(x)}(x) - p_{g(x)}(x) \geq 0, \quad \forall x \in \mathbb{R}$

Ainsi $R(g) \geq R(g^*), \quad \forall g : \mathbb{R} \xrightarrow{\text{mes.}} \{1, \dots, K\}$

3- $S \in \{-1, 1\}$ $X \in \{0, 1\}$ $Y \in \{1, 2\}$

Loi de $X|S$

	$x=0$	$x=1$
$s=-1$	$\frac{3}{4}$	$\frac{1}{4}$
$s=1$	$\frac{1}{5}$	$\frac{4}{5}$

Loi de $Y|X$

	$y=1$	$y=2$
$x=0$	$\frac{2}{3}$	$\frac{1}{3}$
$x=1$	$\frac{1}{10}$	$\frac{9}{10}$

$$g^*(x) = \underset{k}{\operatorname{argmax}}\ p_k(x) \quad \text{donc} \quad \begin{cases} g^*(0) = 1 \\ g^*(1) = 2 \end{cases}$$

Ainsi, $\mathbb{P}(g^*(X) = 2 \mid S = 1) = \mathbb{P}(X = 1 \mid S = 1) = \frac{4}{5}$

et $\mathbb{P}(g^*(X) = 2 \mid S = -1) = \mathbb{P}(X = 1 \mid S = -1) = \frac{1}{4}$

De même, $\mathbb{P}(g^*(X) = 1 \mid S = 1) = \mathbb{P}(X = 0 \mid S = 1) = \frac{1}{5}$

et $\mathbb{P}(g^*(X) = 1 \mid S = -1) = \mathbb{P}(X = 0 \mid S = -1) = \frac{3}{4}$

La parité démographique n'est donc pas respectée.

4- D'après la question 1, on a

$$R(g) = \sum_{s\in\{-1,1\}} \pi_s \int_{\mathbb{R}} \{1 - p_{g(x,s)}(x,s)\} f_{X|S}(dx|s)$$

ce qu'on peut réécrire

$$R(g) = \sum_{k=1}^{K} \sum_{s\in\{-1,1\}} \pi_s \int_{\mathbb{R}} \{1 - p_k(x,s)\} \mathbb{1}_{\{g(x,s)=k\}} f_{X|S}(dx|s)$$

On a aussi, $\forall s \in \{-1,1\}$, $\forall k \in \{1,\dots,K\}$,

$$\mathbb{P}(g(X,S)=k \mid S=s) = \int_{\mathbb{R}} \underbrace{\mathbb{P}(g(X,S)=k \mid X=x, S=s)}_{= \mathbb{1}_{\{g(x,s)=k\}}} f_{X|S}(dx|s)$$

En remarquant que

$$\lambda_k \{\mathbb{P}(g(X,S)=k \mid S=1) - \mathbb{P}(g(X,S)=k \mid S=-1)]$$

$$= \lambda_k \sum_{s\in\{-1,1\}} s\, \mathbb{P}(g(X,S)=k \mid S=s),$$

On trouve

$$L(g,\lambda) = \sum_{k=1}^{K} \sum_{s\in\{-1,1\}} \pi_s \int_{\mathbb{R}} \{1 - p_k(x,s)\} \mathbb{1}_{\{g(x,s)=k\}} f_{X|S}(dx|s)$$

$$+ \sum_{k=1}^{K} \lambda_k \sum_{s\in\{-1,1\}} s \int_{\mathbb{R}} \mathbb{1}_{\{g(x,s)=k\}} f_{X|S}(dx|s)$$

$$= \sum_{k=1}^{K} \sum_{s\in\{-1,1\}} \int_{\mathbb{R}} \mathbb{1}_{\{g(x,s)=k\}} f_{X|S}(dx|s) \times \left[\pi_s \{1 - p_k(x,s)\} + \lambda_k s\right]$$

$$= \underbrace{\sum_{k=1}^{K} \sum_{s\in\{-1,1\}} \int_{\mathbb{R}} \pi_s \mathbb{1}_{\{g(x,s)=k\}} f_{X|S}(dx|s)}_{=1}$$

$$- \sum_{k=1}^{K} \sum_{s\in\{-1,1\}} \int_{\mathbb{R}} \{\pi_s p_k(x,s) - \lambda_k s\} \mathbb{1}_{\{g(x,s)=k\}} f_{X|S}(dx|s)$$

5- Pour minimiser $R_\lambda(g) = L(g,\lambda)$ en g, d'après l'expression au-dessus, il suffit, pour tout (x,s), de défini $g(x,s) = k$ où k maximise $\pi_s\, p_k(x,s) - \lambda_k s$, ce qui est équivalent à maximiser

$$p_k(x,s) - \frac{\lambda_k s}{\pi_s}$$

donc $\quad g^*_\lambda(x,s) \in \underset{k}{\operatorname{argmax}} \; p_k(x,s) - \frac{\lambda_k s}{\pi_s}$

6 - Par définition de g^*_λ, $\quad \forall (x,s) \in \mathbb{R}\times\{-1,1\}$,

$$\{\pi_s p_k(x,s) - s\lambda_k\} \mathbb{1}_{\{g^*_\lambda(x,s)=k\}}$$

$$= \begin{cases} 0 & \text{si } k \notin \underset{j}{\operatorname{argmax}} \; p_j(x,s) - \frac{\lambda_j s}{\pi_s} \\ \max\limits_j \; p_j(x,s) - \frac{\lambda_j s}{\pi_s} & \text{sinon} \end{cases}$$

Donc $\quad \sum_{k=1}^{K} \{\pi_s p_k(x,s) - s\lambda_k\} \mathbb{1}_{\{g^*_\lambda(x,s)=k\}}$

$$= \max_k \{\pi_s p_k(x,s) - s\lambda_k\}$$

Et donc λ^* maximise

$$1 - \sum_{s\in\{-1,1\}} \int_{\mathbb{R}} \max_k \{\pi_s p_k(x,s) - s\lambda_k\} f_{X|S}(dx|s)$$

7- * La règle $g^*_{\lambda^*}$ fait intervenir des paramètres inconnus : π_s, $p_k(x,s)$, $f_{X|S}$

qu'on commence par estimer : $\hat{\pi}_s$, $\hat{p}_k(x,s)$ et $\hat{f}_{X|S}$

inutile a posteriori

* Pour estimer λ^*, il faut évaluer

$$A(\lambda) = \sum_{s\in\{-1,1\}} \int_{\mathbb{R}} \max_k \{\pi_s p_k(x,s) - s\lambda_k\} f_{X|S}(dx|s)$$

Pour cela, on peut appliquer une méthode de Monte-Carlo :

$$A(\lambda) = \sum_{s\in\{-1,1\}} \int_{\mathbb{R}} \underbrace{\frac{\max_k \{\pi_s p_k(x,s) - s\lambda_k\}}{\pi_s}}_{= F_\lambda(x,s)} \underbrace{f_{X,S}(dx,s)}_{= \pi_s f_{X|S}(dx|s)}$$

$$= \mathbb{E}[F_\lambda(X,S)]$$

La fonctionnelle $\mathcal{F}_\lambda$ est inconnue mais peut être estimée :

$$\hat{\mathcal{F}}_\lambda(x,s) = \frac{\max_k \{\hat{\pi}_s \hat{p}_k(x,s) - s\lambda_k\}}{\hat{\pi}_s}$$

Et donc $\mathcal{A}(\lambda)$ peut être estimé par

$$\hat{\mathcal{A}}(\lambda) = \mathbb{E}\left[\hat{\mathcal{F}}_\lambda(X,S) \mid \hat{\pi}_s, \hat{p}_k(x,s)\right]$$

C'est cette espérance qu'on évalue par Monte-Carlo, à partir de données qui N'ont PAS servi à construire $\hat{\pi}_s$ et $\hat{p}_k(x,s)$:

$$\hat{\hat{\mathcal{A}}}(\lambda) = \frac{1}{n}\sum_{i=1}^{n} \hat{\mathcal{F}}_\lambda(X_i,S_i)$$

$$\xrightarrow[n\to\infty]{p.s.} \hat{\mathcal{A}}(\lambda)$$

Alors, on peut poser $\hat{\lambda}^* = \operatorname{argmax} \hat{\hat{\mathcal{A}}}$

* On peut maintenant définir $\hat{g}_{\hat{\lambda}^*}$:

$$\hat{g}_{\hat{\lambda}^*}(x,s) = \operatorname*{argmax}_k \; \hat{p}_k(x,s) - \frac{s\hat{\lambda}_k^*}{\hat{\pi}_s}$$