TA2 Actividades de decisión

Índice

1 Introducción

☐ En un problema de clasificación en C clases, tanto los posibles estados de la naturaleza como las posibles acciones son etiquetas de clase, $\mathcal{H} = \mathcal{A} = \mathcal{Y} = \{1, \dots, C\}$. Si la pérdida asociada a cada par estado-acción es la 0-1, $\ell(y^*, \hat{y}) = \mathbb{I}(y^* \neq \hat{y})$, la pérdida esperada de una acción $\hat{y}$, a posteriori de una observación $\boldsymbol{x}$, es:

1. $R(\hat{y} \mid \boldsymbol{x}) = p(y = \hat{y} \mid \boldsymbol{x})$
2. $R(\hat{y} \mid \boldsymbol{x}) = 1 - p(y = \hat{y} \mid \boldsymbol{x})$
3. $R(\hat{y} \mid \boldsymbol{x}) = 1 - p(y \neq \hat{y} \mid \boldsymbol{x})$
4. Ninguna de las anteriores

Solución: La 2.

---

2 Ejercicio

☐ Supóngase que tenemos dos cajas con 40 naranjas cada una. La primera caja contiene 9 naranjas Navelina y 31 Caracara. La segunda caja contiene 20 naranjas de cada tipo. Ahora supóngase que se escoge una caja al azar, y luego una naranja al azar de la caja escogida. Si la naranja escogida es Navelina, la probabilidad P de que proceda de la primera caja es:

1. $0/4 \leq P < 1/4$
2. $1/4 \leq P < 2/4$
3. $2/4 \leq P < 3/4$
4. $3/4 \leq P \leq 4/4$

Solución: La 2; $P = 0.31$

---

3 Clasificación sensible al coste

☐ Sea un problema de clasificación en dos clases, 0 (negativo) y 1 (positivo), tal que el coste de acertar es nulo, mientras que el coste de falso negativo es dos veces mayor que el de falso positivo. El estimador de Bayes para este problema consiste en decidirse por positivo si la probabilidad de positivo, p_1, es:

1. $p_1 \geq 1/2$
2. $p_1 \geq 1/3$
3. $p_1 \geq 1/4$
4. Ninguna de las anteiores.

Solución: La 2; $p_1 \geq \frac{1}{1+c}$, donde c es el coste relativo de falso negativo frente al de falso positivo, esto es, $\ell_{10} = c\ell_{01}$.

4 Clasificación con opción de rechazo

☐ Sea un problema de clasificación de pérdida con opción de rechazo (acción 0): $\mathcal{H} = \mathcal{Y} = \{1, \dots, C\}$ y $\mathcal{A} = \mathcal{Y} \cup \{0\}$

$$\ell(y^*, a) = \begin{cases} 0 & \text{si } a = y^* \quad \text{(acertar no tiene coste)} \\ \lambda_r & \text{si } a = 0 \quad \text{(coste del rechazo)} \\ \lambda_e & \text{otro caso} \quad \text{(mismo coste para todo error, mayor que el de rechazo, } \lambda_r < \lambda_e) \end{cases}$$

La pérdida esperada de rechazo, a posteriori de una observación $\boldsymbol{x}$, $R(a \mid \boldsymbol{x})$, $a = 0$, es:

1. λ_r
2. λ_e
3. $\max(\lambda_r, \lambda_e)$
4. Ninguna de las anteriores.

Solución: La 1

$$R(0 \mid \boldsymbol{x}) = \sum_{y^* \in \mathcal{Y}} \ell(y^*, 0)\, p(y^* \mid \boldsymbol{x}) = \sum_{y^* \in \mathcal{Y}} \lambda_r\, p(y^* \mid \boldsymbol{x}) = \lambda_r \sum_{y^* \in \mathcal{Y}} p(y^* \mid \boldsymbol{x}) = \lambda_r$$

☐ Sea un problema de clasificación de pérdida con opción de rechazo (acción 0): $\mathcal{H} = \mathcal{Y} = \{1, \dots, C\}$ y $\mathcal{A} = \mathcal{Y} \cup \{0\}$

$$\ell(y^*, a) = \begin{cases} 0 & \text{si } a = y^* \quad \text{(acertar no tiene coste)} \\ \lambda_r & \text{si } a = 0 \quad \text{(coste del rechazo)} \\ \lambda_e & \text{otro caso} \quad \text{(mismo coste para todo error, mayor que el de rechazo, } \lambda_r < \lambda_e) \end{cases}$$

La pérdida esperada de un no rechazo, a posteriori de una observación $\boldsymbol{x}$, $R(a \mid \boldsymbol{x})$, $a = \hat{y} \in \mathcal{Y}$, es:

1. λ_e
2. $\lambda_e\, p(\hat{y} \mid \boldsymbol{x})$
3. $\lambda_e\, (1 - p(\hat{y} \mid \boldsymbol{x}))$
4. Ninguna de las anteriores.

Solución: La 3

$$\begin{aligned} R(\hat{y} \mid \boldsymbol{x}) &= \sum_{y^* \in \mathcal{Y}} \ell(y^*, \hat{y})\, p(y^* \mid \boldsymbol{x}) \\ &= \ell(\hat{y}, \hat{y})\, p(\hat{y} \mid \boldsymbol{x}) + \sum_{y^* \in \mathcal{Y} \setminus \{\hat{y}\}} \ell(y^*, \hat{y})\, p(y^* \mid \boldsymbol{x}) \\ &= 0\, p(\hat{y} \mid \boldsymbol{x}) + \sum_{y^* \in \mathcal{Y} \setminus \{\hat{y}\}} \lambda_e\, p(y^* \mid \boldsymbol{x}) \\ &= \lambda_e \sum_{y^* \in \mathcal{Y} \setminus \{\hat{y}\}} p(y^* \mid \boldsymbol{x}) \\ &= \lambda_e (1 - p(\hat{y} \mid \boldsymbol{x})) \end{aligned}$$

☐ Sea un problema de clasificación de pérdida con opción de rechazo (acción 0): $\mathcal{H} = \mathcal{Y} = \{1, \dots, C\}$ y $\mathcal{A} = \mathcal{Y} \cup \{0\}$

$$\ell(y^*, a) = \begin{cases} 0 & \text{si } a = y^* \quad (\text{acertar no tiene coste}) \\ \lambda_r & \text{si } a = 0 \quad (\text{coste del rechazo}) \\ \lambda_e & \text{otro caso} \quad (\text{mismo coste para todo error, mayor que el de rechazo, } \lambda_r < \lambda_e) \end{cases}$$

La pérdida esperada de una acción $a \in \mathcal{A}$, a posteriori de una observación $\boldsymbol{x}$, es:

$$R(a \mid \boldsymbol{x}) = \begin{cases} \lambda_r & \text{si } \ a = 0 \\ \lambda_e(1 - p(\hat{y} \mid \boldsymbol{x})) & \text{si } \ a = \hat{y} \in \mathcal{Y} \end{cases}$$

Sea y^* una clase de máxima probabilidad a posteriori, $y^* = \operatorname{argmax}_{y \in \mathcal{Y}} p(y \mid \boldsymbol{x})$, y sean $p^* = p(y^* \mid \boldsymbol{x})$ y $\lambda^* = 1 - \frac{\lambda_r}{\lambda_e}$. El estimador de Bayes, $\pi^*(\boldsymbol{x})$, es:

1. y^* si $p^* > \lambda^*$; 0 en otro caso
2. y^* si $p^* \leq \lambda^*$; 0 en otro caso
3. y^* en cualquier caso
4. Ninguna de las anteriores.

Solución:

$$\begin{aligned} \pi^*(\boldsymbol{x}) &= \operatorname*{argmin}_{a \in \mathcal{A}} \ R(a \mid \boldsymbol{x}) \\ &= \operatorname*{argmin}_{a \in \mathcal{A}} \begin{cases} \lambda_r & \text{si } a = 0 \\ \lambda_e(1 - p(\hat{y} \mid \boldsymbol{x})) & \text{si } a = \hat{y} \in \mathcal{Y} \end{cases} \\ &= \operatorname*{argmin}_{a \in \{0, y^*\}} \begin{cases} \lambda_r & \text{si } a = 0 \\ \lambda_e(1 - p^*) & \text{si } a = y^* \end{cases} \\ &= \begin{cases} y^* & \text{si } \lambda_e(1 - p^*) < \lambda_r \\ 0 & \text{en otro caso} \end{cases} \\ &= \begin{cases} y^* & \text{si } p^* > \lambda^* \\ 0 & \text{en otro caso} \end{cases} \end{aligned}$$

---

5 Matrices de confusión binarias

☐ Sea $\hat{y}_\tau(\boldsymbol{x}) = \mathbb{I}(p(y=1 \mid \boldsymbol{x}) \geq 1-\tau)$ un clasificador basado en un umbral $\tau \in [0,1]$ para un un problema de clasificación binaria. Tras clasificar 12 muestras (de test) con un cierto $\tau,$ se ha obtenido la siguiente matriz de confusión:

y	$\hat{0}$	$\hat{1}$
0	3	1
1	2	6

La true negative rate (o specificity) TNR_τ y la true positive rate (hit, recall o sensitivity) TPR_τ son:

1. $\mathrm{TNR}_\tau < 0.5$ y $\mathrm{TPR}_\tau < 0.5$
2. $\mathrm{TNR}_\tau < 0.5$ y $\mathrm{TPR}_\tau \geq 0.5$
3. $\mathrm{TNR}_\tau \geq 0.5$ y $\mathrm{TPR}_\tau < 0.5$
4. $\mathrm{TNR}_\tau \geq 0.5$ y $\mathrm{TPR}_\tau \geq 0.5$

Solución: La 4:

$$\mathrm{TNR}_\tau = \mathrm{TN}_\tau / N = 3/4 = 0.75$$
$$\mathrm{TPR}_\tau = \mathrm{TP}_\tau / P = 6/8 = 0.75$$

☐ Sea $\hat{y}_\tau(\boldsymbol{x}) = \mathbb{I}(p(y=1 \mid \boldsymbol{x}) \geq 1-\tau)$ un clasificador basado en un umbral $\tau \in [0,1]$ para un un problema de clasificación binaria. Tras clasificar 12 muestras (de test) con un cierto $\tau,$ se ha obtenido la siguiente matriz de confusión:

y	$\hat{0}$	$\hat{1}$
0	3	1
1	2	6

La false discovery rate FDR_τ y la precisión (o positive predictive value) PPV_τ son:

1. $\mathrm{FDR}_\tau < 0.5$ y $\mathrm{PPV}_\tau < 0.5$
2. $\mathrm{FDR}_\tau < 0.5$ y $\mathrm{PPV}_\tau \geq 0.5$
3. $\mathrm{FDR}_\tau \geq 0.5$ y $\mathrm{PPV}_\tau < 0.5$
4. $\mathrm{FDR}_\tau \geq 0.5$ y $\mathrm{PPV}_\tau \geq 0.5$

Solución: La 2:

$$\mathrm{FDR}_\tau = \mathrm{FP}_\tau / \hat{P}_\tau = 1/7 = 0.14$$
$$\mathrm{PPV}_\tau = \mathrm{TP}_\tau / \hat{P}_\tau = 6/7 = 0.86$$

---

6 Curvas ROC

☐ Sea $\hat{y}_\tau(\boldsymbol{x}) = \mathbb{I}(p(y=1 \mid \boldsymbol{x}) \geq 1-\tau)$ un clasificador basado en un umbral $\tau \in [0,1]$ para un un problema de clasificación binaria. La curva ROC de $\hat{y}_\tau$ se obtiene calculando, para todo τ:

1. La TPR_τ en función de la FPR_τ
2. La FPR_τ en función de la TPR_τ
3. La PPV_τ en función de la TPR_τ
4. Ninguna de las anteriores

Solución: La 1

☐ Sea $\hat{y}_\tau(\boldsymbol{x}) = \mathbb{I}(p(y=1 \mid \boldsymbol{x}) \geq 1-\tau)$ un clasificador basado en un umbral $\tau \in [0,1]$ para un un problema de clasificación binaria. En relación con la curva ROC de $\hat{y}_\tau,$ indica la opción incorrecta o la última opción si las tres primeras son correctas:

1. Si $p(y=1 \mid \boldsymbol{x}) = 1$ para las muestras positivas y 0 para las negativas, la curva ROC se reduce a un punto en la esquina superior izquierda y otro en la superior derecha
2. Si $p(y=1 \mid \boldsymbol{x}) = \mathrm{Unif}(0,1),$ la curva ROC es una recta diagonal
3. En el caso de un clasificador realista, la curva ROC es monótona creciente por encima de la diagonal; mejor cuanto más arriba
4. Las tres opciones anteriores son correctas

Solución: La 4

☐ Para facilitar la comparación de curvas ROC de distintos clasificadores, se suelen utilizar medidas que resumen una curva ROC mediante un escalar. En relación con estas medidas, indica la opción incorrecta o la última opción si las tres primeras son correctas:

1. El área bajo la curva (AUC, area under curve) es una medida muy popular
2. La equal error rate (EER, o cross-over rate) es igual a la abscisa del punto en el que la curva ROC cruza la diagonal que va de la esquina superior izquierda a la inferior derecha
3. Tanto AUC como EER son medidas de bondad, esto es, mejor cuanto mayor sean
4. Las tres opciones anteriores son correctas

Solución: La 3; la AUC sí, pero la EER no; mejor cuanto menor EER

Problema. Sea $p(y \mid \boldsymbol{x})$ un modelo probabilístico para clasificar $\boldsymbol{x}$ como positivo ($y=1$) o negativo ($y=0$). Se desea estudiar su comportamiento en función de un umbral $\tau \in [0,1]$ que actúe a modo de margen de positividad; esto es, tal que prediga positivo si $p(y=1 \mid \boldsymbol{x}) \geq 1-\tau;$ negativo si no. Así pues, dado $\tau,$ la clase predicha de $\boldsymbol{x}$ es:

$$\hat{y}_\tau(\boldsymbol{x}) = \mathbb{I}(p(y=1 \mid \boldsymbol{x}) \geq 1-\tau)$$

Se dispone de $M=10$ muestras de test; 5 negativas y 5 positivas. Sus clases reales y probabilidades de positivo según el modelo son:

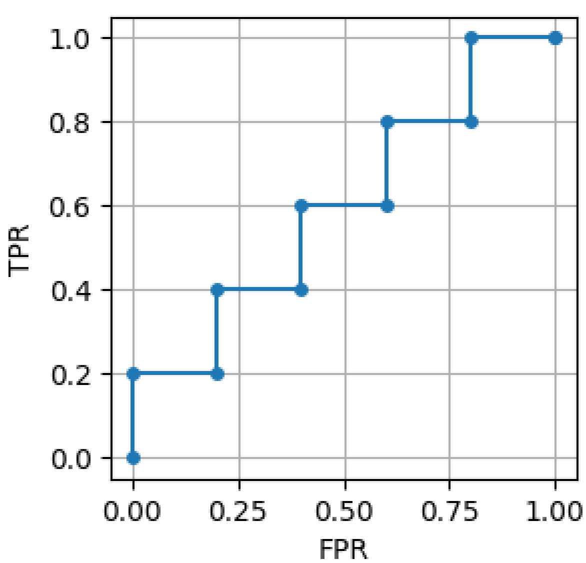
$\boldsymbol{x}_n$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_3$	$\boldsymbol{x}_4$	$\boldsymbol{x}_5$	$\boldsymbol{x}_6$	$\boldsymbol{x}_7$	$\boldsymbol{x}_8$	$\boldsymbol{x}_9$	$\boldsymbol{x}_{10}$
y_n	1	0	1	0	1	0	1	0	1	0
$p(y_n=1 \mid \boldsymbol{x})$	0.81	0.76	0.58	0.47	0.39	0.28	0.23	0.12	0.11	0.05

Construye una tabla con $\mathrm{TP}_\tau,$ $\mathrm{FP}_\tau,$ $\mathrm{FN}_\tau,$ $\mathrm{TN}_\tau,$ TPR_τ y FPR_τ en función de $\tau;$ y dibuja la curva ROC.

Solución: variar τ de 0 a 1 equivale a recorrer las muestras en orden decreciente de probabilidad

τ	$1-\tau$	TP_τ	FP_τ	FN_τ	TN_τ	TPR_τ	FPR_τ
0.00	1.00	0	0	5	5	0.0	0.0
0.19	0.81	1	0	4	5	0.2	0.0
0.24	0.76	1	1	4	4	0.2	0.2
0.42	0.58	2	1	3	4	0.4	0.2
0.53	0.47	2	2	3	3	0.4	0.4
0.61	0.39	3	2	2	3	0.6	0.4
0.72	0.28	3	3	2	2	0.6	0.6
0.77	0.23	4	3	1	2	0.8	0.6
0.88	0.12	4	4	1	1	0.8	0.8
0.89	0.11	5	4	0	1	1.0	0.8
0.95	0.05	5	5	0	0	1.0	1.0
1.00	0.00	5	5	0	0	1.0	1.0

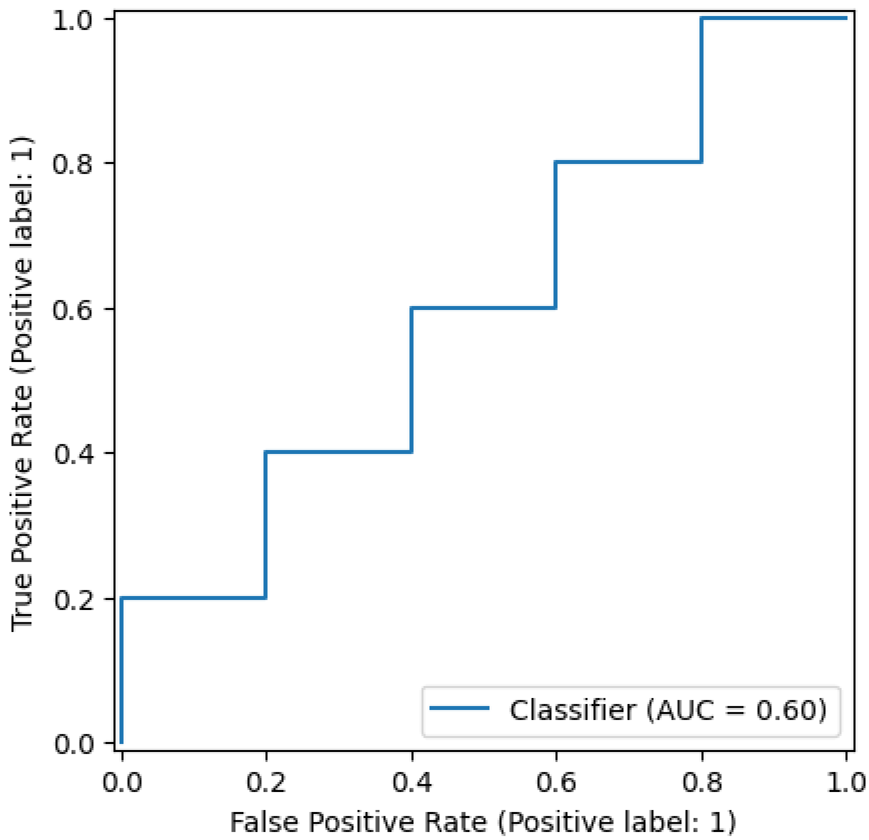
In [ ]:
```
import numpy as np; import matplotlib.pyplot as plt
TPR = np.array([0.0, 0.2, 0.2, 0.4, 0.4, 0.6, 0.6, 0.8, 0.8, 1.0, 1.0, 1.0])
FPR = np.array([0.0, 0.0, 0.2, 0.2, 0.4, 0.4, 0.6, 0.6, 0.8, 0.8, 1.0, 1.0])
plt.figure(figsize=(3, 3)); plt.plot(FPR, TPR, marker='o', markersize=4)
plt.xlabel("FPR"); plt.ylabel("TPR"); plt.grid(); plt.show()
```



In [ ]:

```python
import numpy as np; from sklearn import metrics
y_true = np.array([1, 0, 1, 0, 1, 0, 1, 0, 1, 0])
y_score = np.array([0.81, 0.76, 0.58, 0.47, 0.39, 0.28, 0.23, 0.12, 0.11, 0.05])
FPR, TPR, thresholds = metrics.roc_curve(y_true, y_score)
print(np.c_[(FPR, TPR, thresholds)])
metrics.RocCurveDisplay.from_predictions(y_true, y_score);
```

```
[[0.   0.    inf]
 [0.   0.2  0.81]
 [0.2  0.2  0.76]
 [0.2  0.4  0.58]
 [0.4  0.4  0.47]
 [0.4  0.6  0.39]
 [0.6  0.6  0.28]
 [0.6  0.8  0.23]
 [0.8  0.8  0.12]
 [0.8  1.   0.11]
 [1.   1.   0.05]]
```



7 Curvas PR

☐ Sea $\hat{y}_\tau(\boldsymbol{x}) = \mathbb{I}(p(y=1 \mid \boldsymbol{x}) \geq 1-\tau)$ un clasificador basado en un umbral $\tau \in [0,1]$ para un un problema de clasificación binaria. La curva PR de $\hat{y}_\tau$ se obtiene calculando, para todo τ:

1. La PPV_τ en función de la TPR_τ
2. La TPR_τ en función de la PPV_τ
3. La PPV_τ en función de la FPR_τ
4. La FPR_τ en función de la PPV_τ

Solución: La 1

☐ Sea $\hat{y}_\tau(\boldsymbol{x}) = \mathbb{I}(p(y=1 \mid \boldsymbol{x}) \geq 1-\tau)$ un clasificador basado en un umbral $\tau \in [0,1]$ para un un problema de clasificación binaria. En relación con la curva PR de $\hat{y}_\tau,$ indica la opción incorrecta o la última opción si las tres primeras son correctas:

1. Si $p(y=1 \mid \boldsymbol{x}) = 1$ para las muestras positivas y 0 para las negativas, la curva PR se reduce a un punto en $(1,1)$ y otro en $(1, P/M)$
2. Si $p(y=1 \mid \boldsymbol{x}) = \mathrm{Unif}(0,1),$ la curva PR es una recta horizontal a altura P/M
3. En el caso de un clasificador realista, la curva PR es aproximadamente monótona decreciente, por encima de P/M; mejor cuanto más arriba
4. Las tres opciones anteriores son correctas

Solución: La 4

☐ Sea $\hat{y}_\tau(\boldsymbol{x}) = \mathbb{I}(p(y=1 \mid \boldsymbol{x}) \geq 1-\tau)$ un clasificador basado en un umbral $\tau \in [0,1]$ para un un problema de clasificación binaria. La curva PR de $\hat{y}_\tau,$ $\mathcal{P}(\tau)$ en función de $\mathcal{R}(\tau)$ para todo $\tau,$ es aproximadamente monótona decreciente, por encima de $P/M.$ Ahora bien, el decrecimiento monótono no está garantizado ya que, si $\tau' \leq \tau,$ se cumple $\mathcal{R}(\tau') \leq \mathcal{R}(\tau),$ pero puede que no se cumpla $\mathcal{P}(\tau') \geq \mathcal{P}(\tau);$ de hecho, no se cumple cada vez que un aumento de τ (margen de positividad) conlleva un nuevo verdadero positivo. Con el fin de suavizar la curva PR y garantizar su decrecimiento monótono, en su lugar suele usarse la curva PR interpolada. Esta curva sustituye $\mathcal{P}(\tau)$ por:

1. $\min_{\tilde{\tau} < \tau} \mathcal{P}(\tilde{\tau})$
2. $\max_{\tilde{\tau} < \tau} \mathcal{P}(\tilde{\tau})$
3. $\min_{\tilde{\tau} \geq \tau} \mathcal{P}(\tilde{\tau})$
4. $\max_{\tilde{\tau} \geq \tau} \mathcal{P}(\tilde{\tau})$

Solución: La 4

Problema. Sea $p(y \mid \boldsymbol{x})$ un modelo probabilístico para clasificar $\boldsymbol{x}$ como positivo ($y = 1$) o negativo ($y = 0$). Se desea estudiar su comportamiento en función de un umbral $\tau \in [0, 1]$ que actúe a modo de margen de positividad; esto es, tal que prediga positivo si $p(y = 1 \mid \boldsymbol{x}) \geq 1 - \tau$; negativo si no. Así pues, dado τ, la clase predicha de $\boldsymbol{x}$ es:

$$\hat{y}_\tau(\boldsymbol{x}) = \mathbb{I}(p(y = 1 \mid \boldsymbol{x}) \geq 1 - \tau)$$

Se dispone de $M = 10$ muestras de test; 5 negativas y 5 positivas. Sus clases reales y probabilidades de positivo según el modelo son:

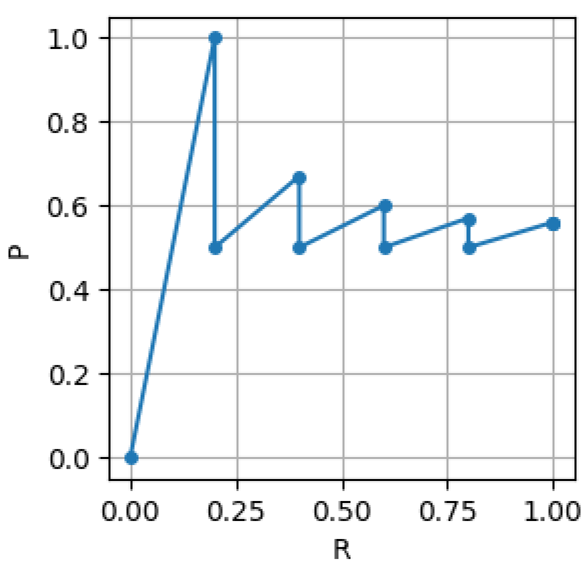
$\boldsymbol{x}_n$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_3$	$\boldsymbol{x}_4$	$\boldsymbol{x}_5$	$\boldsymbol{x}_6$	$\boldsymbol{x}_7$	$\boldsymbol{x}_8$	$\boldsymbol{x}_9$	$\boldsymbol{x}_{10}$
y_n	1	0	1	0	1	0	1	0	1	0
$p(y_n = 1 \mid \boldsymbol{x})$	0.81	0.76	0.58	0.47	0.39	0.28	0.23	0.12	0.11	0.05

Construye una tabla con TP_τ, FP_τ, FN_τ, TN_τ, $\mathcal{P}_\tau$ y $\mathcal{R}_\tau$ en función de τ; y dibuja la curva PR.

Solución: variar τ de 0 a 1 equivale a recorrer las muestras en orden decreciente de probabilidad

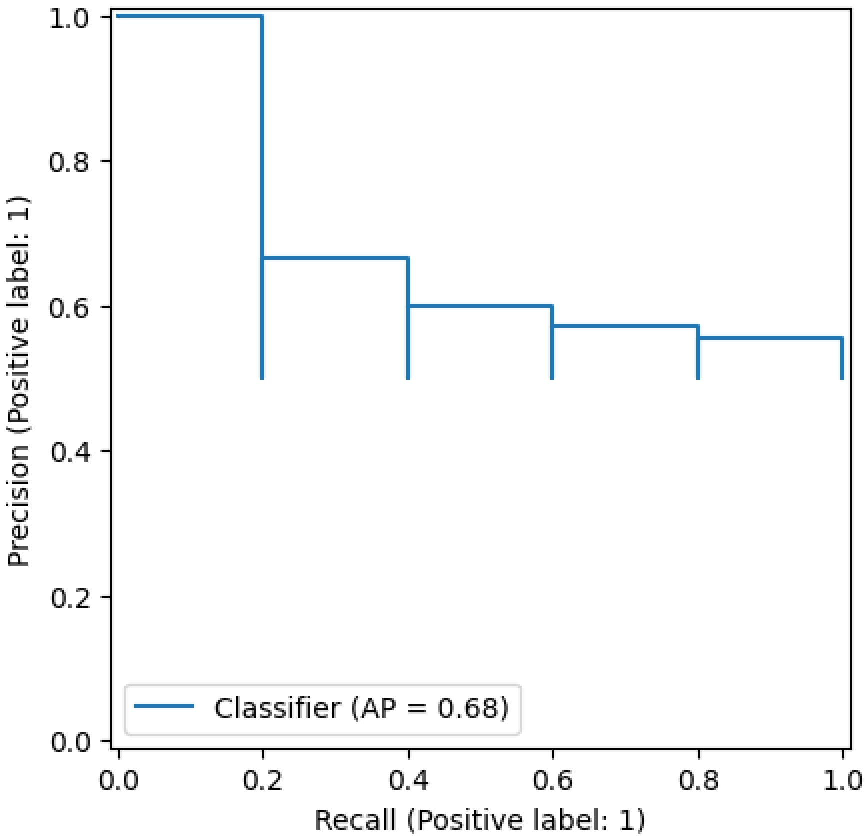
τ	$1-\tau$	TP_τ	FP_τ	FN_τ	TN_τ	$\mathcal{P}_\tau$	$\mathcal{R}_\tau$
0.00	1.00	0	0	5	5	0.00	0.0
0.19	0.81	1	0	4	5	1.00	0.2
0.24	0.76	1	1	4	4	0.50	0.2
0.42	0.58	2	1	3	4	0.67	0.4
0.53	0.47	2	2	3	3	0.50	0.4
0.61	0.39	3	2	2	3	0.60	0.6
0.72	0.28	3	3	2	2	0.50	0.6
0.77	0.23	4	3	1	2	0.57	0.8
0.88	0.12	4	4	1	1	0.50	0.8
0.89	0.11	5	4	0	1	0.56	1.0
0.95	0.05	5	5	0	0	0.56	1.0
1.00	0.00	5	5	0	0	0.50	1.0

In [ ]:
```
import numpy as np; import matplotlib.pyplot as plt
P = np.array([0.00, 1.00, 0.50, 0.67, 0.50, 0.60, 0.50, 0.57, 0.50, 0.56, 0.56, 0.56])
R = np.array([0.0, 0.2, 0.2, 0.4, 0.4, 0.6, 0.6, 0.8, 0.8, 1.0, 1.0, 1.0])
plt.figure(figsize=(3, 3)); plt.plot(R, P, marker='o', markersize=4)
plt.xlabel("R"); plt.ylabel("P"); plt.grid(); plt.show()
```



In [ ]:

```
import numpy as np; from sklearn import metrics
y_true = np.array([1, 0, 1, 0, 1, 0, 1, 0, 1, 0])
y_score = np.array([0.81, 0.76, 0.58, 0.47, 0.39, 0.28, 0.23, 0.12, 0.11, 0.05])
metrics.PrecisionRecallDisplay.from_predictions(y_true, y_score);
```



☐ Para facilitar la comparación de curvas PR de distintos clasificadores, se suelen utilizar medidas que resumen una curva PR mediante un escalar. En relación con estas medidas, indica la opción incorrecta o la última opción si las tres primeras son correctas:

1. Una medida muy popular es el área bajo la curva PR o average precision (AP)
2. La mean average precision (mAP) es la AP media de varias curvas PR, p. ej. de cada clase en clasificación multiclase
3. Tanto la AP como la mAP pueden limitarse a un cierto rango de umbrales relevantes en lugar de todo τ
4. Las tres opciones anteriores son correctas

Solución: La 4

8 F-scores

☐ Los F-scores se introducen para establecer una única medida que evalúe adecuadamente la calidad de cualquier par precisión-cobertura. Dado $\beta \geq 0,\ F_\beta$ se define como:

1. $F_\beta = \dfrac{1}{2}(\mathcal{P} + \mathcal{R})$
2. $F_\beta = \dfrac{1}{1+\beta}(\mathcal{P} + \beta\mathcal{R})$
3. $F_\beta = (1+\beta^2)\,\dfrac{\mathcal{P}\,\mathcal{R}}{\beta^2\mathcal{P} + \mathcal{R}}$
4. Ninguna de las anteriores

Solución: La 3

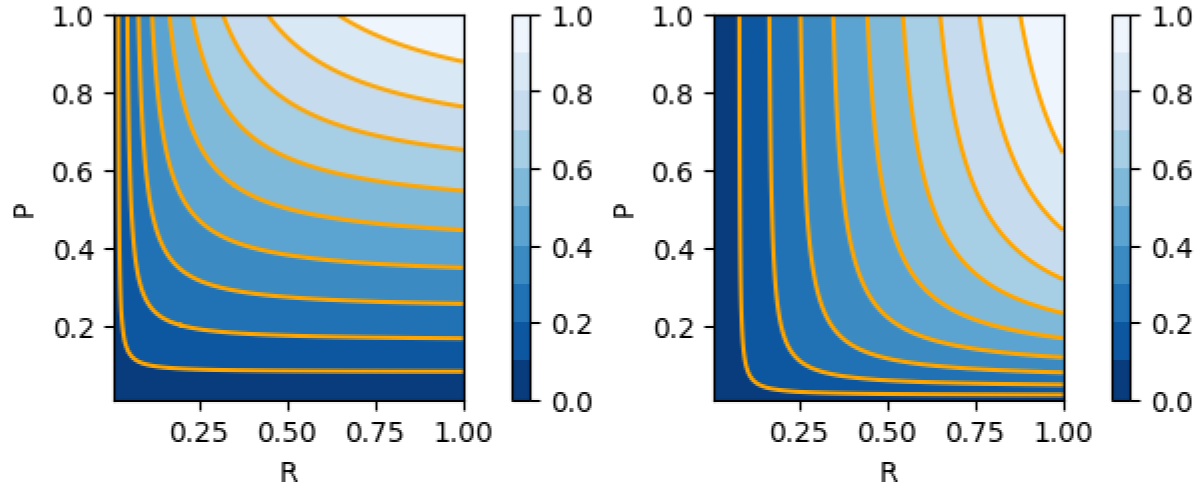
☐ Los F-scores se introducen para establecer una única medida que evalúe adecuadamente la calidad de cualquier par precisión-cobertura. Ahora bien, la elección de un F-score concreto, $F_\beta,$ depende del valor de $\beta \geq 0$ que se considere adecuado en la tarea específica objeto de evaluación. En relación con los valores de $\beta,$ indica la opción incorrecta o la última opción si las tres primeras son correctas:

1. En general, β se interpreta como la importancia relativa de la cobertura frente a la precisión
2. F_1 coincide con la media aromónica de precisión y cobertura
3. F_0 coincide con la precisión y F_∞ con la cobertura
4. Las tres opciones anteriores son correctas

Solución: La 4

☐ La figura siguiente muestra F_β en función de $\mathcal{P}$ y $\mathcal{R}$ para dos valores de β : β_{izq} en la gráfica de la izquierda y β_{dcha} en la de la derecha; ambas gráficas también muestran curvas de nivel para facilitar su interpretación. A la vista de estas gráficas, podemos afirmar que:

1. $\beta_{\text{izq}} < 1 < \beta_{\text{dcha}}$
2. $\beta_{\text{dcha}} < 1 < \beta_{\text{izq}}$
3. $\beta_{\text{izq}} < \beta_{\text{dcha}}$, y $\beta_{\text{izq}} \geq 1$ o $\beta_{\text{dcha}} \leq 1$
4. $\beta_{\text{dcha}} < \beta_{\text{izq}}$, y $\beta_{\text{dcha}} \geq 1$ o $\beta_{\text{izq}} \leq 1$



Solución: La 1

In [ ]:
```python
import numpy as np; import matplotlib.pyplot as plt
R, P = np.meshgrid(np.linspace(0.01, 1, 100), np.linspace(0.01, 1, 100)); RP = np.c_[np.ravel(R), np.ravel(P)]
B = (.5, 2); nrows = 1; ncols = len(B) # len(B) plots de F + plot de media
_, axs = plt.subplots(nrows=nrows, ncols=ncols, figsize=(6, 5*nrows/ncols), constrained_layout=True)
for ax, b in zip(axs.flat, B):
    bsquared = np.square(b); F = lambda rp: (1 + bsquared) * rp.prod() / ( bsquared * rp[1] + rp[0] )
    FF = np.apply_along_axis(F, 1, RP); ax.set_xlabel('R'); ax.set_ylabel('P')
    ax.contour(R, P, FF.reshape(R.shape), 10, colors='orange', linestyles='solid')
    cp = ax.contourf(R, P, FF.reshape(R.shape), 10, cmap='Blues_r'); plt.colorbar(cp, ax=ax)
```