J. Appl. Prob. **62,** 1–35 (2025)
Printed in England

SUPPLEMENTARY MATERIAL: THE DUTCH DRAW: CONSTRUCTING A UNIVERSAL BASELINE FOR BINARY CLASSIFICATION PROBLEMS

SANDJAI BHULAI,* *** *Vrije Universiteit Amsterdam*

ETIENNE VAN DE BIJL,** **** *Centrum Wiskunde & Informatica*

MARK HOOGENDOORN,***** *Vrije Universiteit Amsterdam*

JAN KLEIN,** ****** *Centrum Wiskunde & Informatica*

JORIS PRIES,** ******* *Centrum Wiskunde & Informatica*

This Supplementary Material contains the complete theoretical analysis used to gather the information presented in Sec. 2 and 3 of the Dutch Draw article, and more specifically, Tables 2, 3, and 4. Each section is dedicated to one of the evaluation measures.

Submitted 31 March 2022; accepted 22 July 2024.
* Postal address: Vrije Universiteit Amsterdam, Department of Mathematics, De Boelelaan 1111, 1081 HV, Amsterdam, Netherlands.
** Postal address: Centrum Wiskunde & Informatica, Department of Stochastics, Science Park 123, 1098 XG, Amsterdam, Netherlands.
*** Email address: s.bhulai@vu.nl
**** Email address: evdb@cwi.nl
***** Postal address: Department of Computer Science, De Boelelaan 1111, 1081 HV, Amsterdam, Netherlands. Email: m.hoogendoorn@vu.nl
****** Email address: jan_g_klein@outlook.com
******* Email address: jorispries@gmail.com
******** Email address: mei@cwi.nl

The following definitions are frequently used in the Supplementary Material:

$$\mathrm{TP}_\theta = \mathrm{TP}_\theta, \tag{B1}$$
$$\mathrm{FP}_\theta = \hat{P}_\theta - \mathrm{TP}_\theta, \tag{B2}$$
$$\mathrm{FN}_\theta = P - \mathrm{TP}_\theta, \tag{B3}$$
$$\mathrm{TN}_\theta = N - \hat{P}_\theta + \mathrm{TP}_\theta. \tag{B4}$$

$$X_\theta(a,b) := a \cdot \mathrm{TP}_\theta + b \text{ with } a, b \in \mathbb{R},$$
$$f_{X_\theta}(a,b) := \text{probability distribution of } X_\theta(a,b).$$

$$\mathbb{E}[X_\theta(a,b)] = a \cdot \mathbb{E}[\mathrm{TP}_\theta] + b = a \cdot \frac{\lfloor M \cdot \theta \rceil}{M} \cdot P + b. \tag{1}$$

$$\mathcal{D}(\mathrm{TP}_\theta) := \{i \in \mathbb{N}_0 : \max\{0, \lfloor M \cdot \theta \rceil - (M - P)\} \leq i \leq \min\{P, \lfloor M \cdot \theta \rceil\}\},$$
$$\mathcal{R}(X_\theta(a,b)) := \{a \cdot i + b\}_{i \in \mathcal{D}(\mathrm{TP}_\theta)}. \tag{R}$$

An overview of the entire Supplementary Material can be viewed in Table 1.

TABLE 1: ***Overview of the Supplementary Material:*** Each measure is discussed in the corresponding section in the Supplementary Material

Measure	TP	TN	FN	FP	TPR	TNR	FNR	FPR	PPV	NPV	FDR	FOR
Section	1	2	3	4	5	6	7	8	9	10	11	12
Measure	F_β	J	MK	Acc	BAcc	MCC	κ	FM	$\mathrm{G}^{(2)}$	PT	TS	
Section	13	14	15	16	17	18	19	20	21	22	23	

1. Number of True Positives

The *number of True Positives* TP_θ is one of the four base measures. This measure indicates how many of the predicted positive observations are actually positive. Under the DD methodology, each evaluation measure can be written in terms of TP_θ.

1.1. Definition and distribution

Since we want to formulate each measure in terms of TP_θ, we have for TP_θ:

$$\mathrm{TP}_\theta \overset{(B1)}{=} X_\theta(1,0) \sim f_{X_\theta}(1,0).$$

The range of this base measure depends on θ. Therefore, Eq. (R) yields the range of this measure:

$$\mathrm{TP}_\theta \in \mathcal{R}(X_\theta(1,0)).$$

1.2. Expectation

The expectation of TP_θ using the DD is given by

$$\mathbb{E}[\text{TP}_\theta] = \mathbb{E}[X_\theta(1,0)] \overset{(1)}{=} \frac{\lfloor M \cdot \theta \rceil}{M} \cdot P = \theta^* \cdot P. \tag{2}$$

1.3. Optimal baselines

The optimal expectation gives the DD baseline. Eq. (2) shows that the expected value depends on the parameter θ. Therefore, either the minimum or maximum of the expectation yields the baseline. They are given by

$$\min_{\theta \in [0,1]} (\mathbb{E}[\text{TP}_\theta]) = P \cdot \min_{\theta \in [0,1]} \left(\frac{\lfloor M \cdot \theta \rceil}{M} \right) = 0,$$
$$\max_{\theta \in [0,1]} (\mathbb{E}[\text{TP}_\theta]) = P \cdot \max_{\theta \in [0,1]} \left(\frac{\lfloor M \cdot \theta \rceil}{M} \right) = P.$$

The values of $\theta \in [0,1]$ that minimize or maximize the expected value are θ_{min} and θ_{max}, respectively, and are defined as

$$\theta_{\text{min}} \in \arg\min_{\theta \in [0,1]} (\mathbb{E}[\text{TP}_\theta]) = \arg\min_{\theta \in [0,1]} \left(\frac{\lfloor M \cdot \theta \rceil}{M} \right) = \left[0, \frac{1}{2M}\right),$$
$$\theta_{\text{max}} \in \arg\max_{\theta \in [0,1]} (\mathbb{E}[\text{TP}_\theta]) = \arg\max_{\theta \in [0,1]} \left(\frac{\lfloor M \cdot \theta \rceil}{M} \right) = \left[1 - \frac{1}{2M}, 1\right].$$

Equivalently, the discrete optimizers $\theta^*_{\text{min}} \in \Theta^*$ and $\theta^*_{\text{max}} \in \Theta^*$ are determined by

$$\theta^*_{\text{min}} \in \arg\min_{\theta^* \in \Theta^*} \{\mathbb{E}[\text{TP}_{\theta^*}]\} = \arg\min_{\theta^* \in \Theta^*} \{\theta^*\} = \{0\},$$
$$\theta^*_{\text{max}} \in \arg\max_{\theta^* \in \Theta^*} \{\mathbb{E}[\text{TP}_{\theta^*}]\} = \arg\max_{\theta^* \in \Theta^*} \{\theta^*\} = \{1\}.$$

2. Number of True Negatives

The *number of True Negatives* TN_θ is also one of the four base measures. This base measure counts the number of negative predicted instances that are actually negative.

2.1. Definition and distribution

Since we want to formulate each measure in terms of TP_θ, we have for TN_θ:

$$\text{TN}_\theta = M - P - \lfloor M \cdot \theta \rceil + \text{TP}_\theta,$$

which corresponds to Eq. (B4). Furthermore,

$$\text{TN}_\theta \overset{(B4)}{=} X_\theta(1, M - P - \lfloor M \cdot \theta \rceil) \sim f_{X_\theta}(1, M - P - \lfloor M \cdot \theta \rceil),$$

and for its range

$$\text{TN}_\theta \overset{(R)}{\in} \mathcal{R}(X_\theta(1, M - P - \lfloor M \cdot \theta \rceil)).$$

2.2. Expectation

TN_θ is linear in TP_θ with slope $a = 1$ and intercept $b = M - P - \lfloor M \cdot \theta \rceil$, so its expectation is given by

$$\mathbb{E}[TN_\theta] = \mathbb{E}[X_\theta(1, M - P - \lfloor M \cdot \theta \rceil)] \overset{(1)}{=} 1 \cdot \mathbb{E}[TP_\theta] + M - P - \lfloor M \cdot \theta \rceil$$
$$= \left(1 - \frac{\lfloor M \cdot \theta \rceil}{M}\right)(M - P) = (1 - \theta^*)(M - P).$$

2.3. Optimal baselines

To determine the range of the expectation of TN_θ and obtain baselines, its extreme values are calculated:

$$\min_{\theta \in [0,1]} (\mathbb{E}[TN_\theta]) = (M - P) \min_{\theta \in [0,1]} \left(1 - \frac{\lfloor M \cdot \theta \rceil}{M}\right) = 0,$$
$$\max_{\theta \in [0,1]} (\mathbb{E}[TN_\theta]) = (M - P) \max_{\theta \in [0,1]} \left(1 - \frac{\lfloor M \cdot \theta \rceil}{M}\right) = M - P.$$

The associated optimization values $\theta_{\min} \in [0, 1]$ and $\theta_{\max} \in [0, 1]$ are

$$\theta_{\min} \in \underset{\theta \in [0,1]}{\arg\min} (\mathbb{E}[TN_\theta]) = \underset{\theta \in [0,1]}{\arg\min} \left(1 - \frac{\lfloor M \cdot \theta \rceil}{M}\right) = \left[1 - \frac{1}{2M}, 1\right],$$
$$\theta_{\max} \in \underset{\theta \in [0,1]}{\arg\max} (\mathbb{E}[TN_\theta]) = \underset{\theta \in [0,1]}{\arg\max} \left(1 - \frac{\lfloor M \cdot \theta \rceil}{M}\right) = \left[0, \frac{1}{2M}\right).$$

The discrete equivalents $\theta^*_{\min} \in \Theta^*$ and $\theta^*_{\max} \in \Theta^*$ are then determined by

$$\theta^*_{\min} \in \underset{\theta^* \in \Theta^*}{\arg\min} \{\mathbb{E}[TN_{\theta^*}]\} = \underset{\theta^* \in \Theta^*}{\arg\min} \{1 - \theta^*\} = \{1\},$$
$$\theta^*_{\max} \in \underset{\theta^* \in \Theta^*}{\arg\max} \{\mathbb{E}[TN_{\theta^*}]\} = \underset{\theta^* \in \Theta^*}{\arg\max} \{1 - \theta^*\} = \{0\}.$$

3. Number of False Negatives

The *number of False Negative* FN_θ is one of the four base measures. This base measure counts the number of mistakes made by predicting instances negative while the actual labels are positive.

3.1. Definition and distribution

Eq. (B3) shows that FN_θ can be expressed in terms of TP_θ:

$$FN_\theta \overset{(B3)}{=} P - TP_\theta = X_\theta(-1, P) \sim f_{X_\theta}(-1, P),$$

and for its range:

$$FN_\theta \overset{(R)}{\in} \mathcal{R}(X_\theta(-1, P)).$$

3.2. Expectation

As Eq. (B3) shows, FN_θ is linear in TP_θ with slope $a = -1$ and intercept $b = P$. Hence, the expectation of FN_θ is given by

$$\mathbb{E}[\mathrm{FN}_\theta] = \mathbb{E}[X_\theta(-1, P)] \overset{(1)}{=} -1 \cdot \mathbb{E}[\mathrm{TP}_\theta] + P = \left(1 - \frac{\lfloor M \cdot \theta \rceil}{M}\right) \cdot P = \left(1 - \theta^*\right) \cdot P.$$

3.3. Optimal baselines

The range of the expectation of FN_θ determines the baselines. The extreme values are given by

$$\min_{\theta \in [0,1]} (\mathbb{E}[\mathrm{FN}_\theta]) = P \cdot \min_{\theta \in [0,1]} \left(1 - \frac{\lfloor M \cdot \theta \rceil}{M}\right) = 0,$$
$$\max_{\theta \in [0,1]} (\mathbb{E}[\mathrm{FN}_\theta]) = P \cdot \max_{\theta \in [0,1]} \left(1 - \frac{\lfloor M \cdot \theta \rceil}{M}\right) = P.$$

The associated optimization values $\theta_{\min} \in [0, 1]$ and $\theta_{\max} \in [0, 1]$ are then

$$\theta_{\min} \in \underset{\theta \in [0,1]}{\arg\min} (\mathbb{E}[\mathrm{FN}_\theta]) = \underset{\theta \in [0,1]}{\arg\min} \left(1 - \frac{\lfloor M \cdot \theta \rceil}{M}\right) = \left[1 - \frac{1}{2M}, 1\right],$$
$$\theta_{\max} \in \underset{\theta \in [0,1]}{\arg\max} (\mathbb{E}[\mathrm{FN}_\theta]) = \underset{\theta \in [0,1]}{\arg\max} \left(1 - \frac{\lfloor M \cdot \theta \rceil}{M}\right) = \left[0, \frac{1}{2M}\right),$$

respectively. The discrete versions $\theta^*_{\min} \in \Theta^*$ and $\theta^*_{\max} \in \Theta^*$ of the optimizers are as follows:

$$\theta^*_{\min} \in \underset{\theta^* \in \Theta^*}{\arg\min} \{\mathbb{E}[\mathrm{FN}_{\theta^*}]\} = \underset{\theta^* \in \Theta^*}{\arg\min} \{1 - \theta^*\} = \{1\},$$
$$\theta^*_{\max} \in \underset{\theta^* \in \Theta^*}{\arg\max} \{\mathbb{E}[\mathrm{FN}_{\theta^*}]\} = \underset{\theta^* \in \Theta^*}{\arg\max} \{1 - \theta^*\} = \{0\}.$$

4. Number of False Positives

The *number of False Positives* FP_θ is one of the four base measures. This base measure counts the number of mistakes made by predicting instances as positive while the actual labels are negative.

4.1. Definition and distribution

Each base measure can be expressed in terms of TP_θ, thus we have for FP_θ:

$$\mathrm{FP}_\theta \overset{(B2)}{=} \lfloor M \cdot \theta \rceil - \mathrm{TP}_\theta = X_\theta(-1, \lfloor M \cdot \theta \rceil) \sim f_{X_\theta}(-1, \lfloor M \cdot \theta \rceil),$$

and for its range:

$$\mathrm{FP}_\theta \overset{(R)}{\in} \mathcal{R}(X_\theta(-1, \lfloor M \cdot \theta \rceil)).$$

4.2. Expectation

As Eq. (B2) shows, FP_θ is linear in TP_θ with slope $a = -1$ and intercept $b = \lfloor M \cdot \theta \rceil$, thus the expectation of FP_θ is defined as

$$\begin{aligned}\mathbb{E}[\mathrm{FP}_\theta] &= \mathbb{E}[X_\theta(-1, \lfloor M \cdot \theta \rceil)] \overset{(1)}{=} -1 \cdot \mathbb{E}[\mathrm{TP}_\theta] + \lfloor M \cdot \theta \rceil = \frac{\lfloor M \cdot \theta \rceil}{M} \cdot (M - P) \\ &= \theta^* \cdot (M - P)\,.\end{aligned}$$

4.3. Optimal baselines

The extreme values of its expectation give the baselines of FP_θ. Hence:

$$\begin{aligned}\min_{\theta \in [0,1]} (\mathbb{E}[\mathrm{FP}_\theta]) &= (M - P) \min_{\theta \in [0,1]} \left(\frac{\lfloor M \cdot \theta \rceil}{M} \right) = 0, \\ \max_{\theta \in [0,1]} (\mathbb{E}[\mathrm{FP}_\theta]) &= (M - P) \max_{\theta \in [0,1]} \left(\frac{\lfloor M \cdot \theta \rceil}{M} \right) = M - P.\end{aligned}$$

The corresponding optimization values $\theta_{\min} \in [0, 1]$ and $\theta_{\max} \in [0, 1]$ are

$$\begin{aligned}\theta_{\min} &\in \underset{\theta \in [0,1]}{\arg\min} (\mathbb{E}[\mathrm{FP}_\theta]) = \underset{\theta \in [0,1]}{\arg\min} \left(\frac{\lfloor M \cdot \theta \rceil}{M} \right) = \left[0, \frac{1}{2M}\right), \\ \theta_{\max} &\in \underset{\theta \in [0,1]}{\arg\max} (\mathbb{E}[\mathrm{FP}_\theta]) = \underset{\theta \in [0,1]}{\arg\max} \left(\frac{\lfloor M \cdot \theta \rceil}{M} \right) = \left[1 - \frac{1}{2M}, 1\right].\end{aligned}$$

The discrete versions $\theta^*_{\min} \in \Theta^*$ and $\theta^*_{\max} \in \Theta^*$ of the optimization values are determined by

$$\begin{aligned}\theta^*_{\min} &\in \underset{\theta^* \in \Theta^*}{\arg\min} \{\mathbb{E}[\mathrm{FP}_{\theta^*}]\} = \underset{\theta^* \in \Theta^*}{\arg\min} \{\theta^*\} = \{0\}, \\ \theta^*_{\max} &\in \underset{\theta^* \in \Theta^*}{\arg\max} \{\mathbb{E}[\mathrm{FP}_{\theta^*}]\} = \underset{\theta^* \in \Theta^*}{\arg\max} \{\theta^*\} = \{1\}.\end{aligned}$$

5. True Positive Rate

The *True Positive Rate* TPR_θ, *Recall*, or *Sensitivity* is the performance measure that presents the fraction of positive observations that are correctly predicted. This makes it a fundamental performance measure in binary classification.

5.1. Definition and distribution

The True Positive Rate is commonly defined as

$$\mathrm{TPR}_\theta = \frac{\mathrm{TP}_\theta}{P}. \tag{3}$$

Hence, $P > 0$ should hold, otherwise, the denominator is zero. Now, TPR_θ is linear in TP_θ and can therefore be written as

$$\mathrm{TPR}_\theta = X_\theta \left(\frac{1}{P}, 0 \right) \sim f_{X_\theta} \left(\frac{1}{P}, 0 \right), \tag{4}$$

and for its range:

$$\text{TPR}_\theta \overset{(R)}{\in} \mathcal{R}\left(X_\theta\left(\frac{1}{P}, 0\right)\right).$$

5.2. Expectation

Since TPR_θ is linear in TP_θ with slope $a = 1/P$ and intercept $b = 0$, its expectation is

$$\mathbb{E}[\text{TPR}_\theta] = \mathbb{E}\left[X_\theta\left(\frac{1}{P}, 0\right)\right] \overset{(1)}{=} \frac{1}{P} \cdot \mathbb{E}[\text{TP}_\theta] + 0 = \frac{\lfloor M \cdot \theta \rceil}{M} = \theta^*.$$

5.3. Optimal baselines

The range of the expectation of TPR_θ directly determines the baselines. The extreme values are given by

$$\min_{\theta \in [0,1]} (\mathbb{E}[\text{TPR}_\theta]) = \min_{\theta \in [0,1]} \left(\frac{\lfloor M \cdot \theta \rceil}{M}\right) = 0,$$
$$\max_{\theta \in [0,1]} (\mathbb{E}[\text{TPR}_\theta]) = \max_{\theta \in [0,1]} \left(\frac{\lfloor M \cdot \theta \rceil}{M}\right) = 1.$$

Furthermore, the corresponding optimization values $\theta_{\min} \in [0, 1]$ and $\theta_{\max} \in [0, 1]$ are given by

$$\theta_{\min} \in \underset{\theta \in [0,1]}{\arg\min} (\mathbb{E}[\text{TPR}_\theta]) = \underset{\theta \in [0,1]}{\arg\min} \left(\frac{\lfloor M \cdot \theta \rceil}{M}\right) = \left[0, \frac{1}{2M}\right),$$
$$\theta_{\max} \in \underset{\theta \in [0,1]}{\arg\max} (\mathbb{E}[\text{TPR}_\theta]) = \underset{\theta \in [0,1]}{\arg\max} \left(\frac{\lfloor M \cdot \theta \rceil}{M}\right) = \left[1 - \frac{1}{2M}, 1\right].$$

The discrete versions $\theta^*_{\min} \in \Theta^*$ and $\theta^*_{\max} \in \Theta^*$ of the optimizers are then

$$\theta^*_{\min} \in \underset{\theta^* \in \Theta^*}{\arg\min} \{\mathbb{E}[\text{TPR}_{\theta^*}]\} = \underset{\theta^* \in \Theta^*}{\arg\min} \{\theta^*\} = \{0\},$$
$$\theta^*_{\max} \in \underset{\theta^* \in \Theta^*}{\arg\max} \{\mathbb{E}[\text{TPR}_{\theta^*}]\} = \underset{\theta^* \in \Theta^*}{\arg\max} \{\theta^*\} = \{1\},$$

respectively.

6. True Negative Rate

The *True Negative Rate* TNR_θ, *Specificity*, or *Selectivity* is the measure that shows how relatively well the negative observations are correctly predicted. Hence, this performance measure is a fundamental measure in binary classification.

61 ### 6.1. Definition and distribution

The True Negative Rate is commonly defined as

$$\mathrm{TNR}_\theta = \frac{\mathrm{TN}_\theta}{N}.$$

Hence, $N := M - P > 0$ should hold, otherwise, the denominator is zero. By using Eq. (B4), TNR_θ can be rewritten as

$$\mathrm{TNR}_\theta = \frac{M - P - \lfloor M \cdot \theta \rceil + \mathrm{TP}_\theta}{M - P} = 1 - \frac{\lfloor M \cdot \theta \rceil - \mathrm{TP}_\theta}{M - P}.$$

62 Hence, it is linear in TP_θ and can therefore be written as

$$\mathrm{TNR}_\theta = X_\theta \left(\frac{1}{M - P}, 1 - \frac{\lfloor M \cdot \theta \rceil}{M - P} \right) \sim f_{X_\theta} \left(\frac{1}{M - P}, 1 - \frac{\lfloor M \cdot \theta \rceil}{M - P} \right), \tag{5}$$

and for its range:

$$\mathrm{TNR}_\theta \overset{(R)}{\in} \mathcal{R} \left(X_\theta \left(\frac{1}{M - P}, 1 - \frac{\lfloor M \cdot \theta \rceil}{M - P} \right) \right).$$

63 ### 6.2. Expectation

Since TNR_θ is linear in TP_θ in terms of $X_\theta(a, b)$ with slope $a = 1/(M - P)$ and intercept $b = 1 - \lfloor M \cdot \theta \rceil/(M - P)$, its expectation is

$$\begin{aligned}\mathbb{E}[\mathrm{TNR}_\theta] &= \mathbb{E}\left[X_\theta \left(\frac{1}{M - P}, 1 - \frac{\lfloor M \cdot \theta \rceil}{M - P} \right) \right] \overset{(1)}{=} \frac{1}{M - P} \cdot \mathbb{E}[\mathrm{TP}_\theta] + 1 - \frac{\lfloor M \cdot \theta \rceil}{M - P} \\ &= 1 - \frac{\lfloor M \cdot \theta \rceil}{M} = 1 - \theta^*.\end{aligned}$$

64 ### 6.3. Optimal baselines

The extreme values of the expectation of TNR_θ determine the baselines. The range is given by

$$\begin{aligned}\min_{\theta \in [0,1]} (\mathbb{E}[\mathrm{TNR}_\theta]) &= \min_{\theta \in [0,1]} \left(1 - \frac{\lfloor M \cdot \theta \rceil}{M} \right) = 0, \\ \max_{\theta \in [0,1]} (\mathbb{E}[\mathrm{TNR}_\theta]) &= \max_{\theta \in [0,1]} \left(1 - \frac{\lfloor M \cdot \theta \rceil}{M} \right) = 1.\end{aligned}$$

Moreover, the optimization values $\theta_{\min} \in [0, 1]$ and $\theta_{\max} \in [0, 1]$ corresponding to the extreme values are defined as

$$\begin{aligned}\theta_{\min} &\in \operatorname*{arg\,min}_{\theta \in [0,1]} (\mathbb{E}[\mathrm{TNR}_\theta]) = \operatorname*{arg\,min}_{\theta \in [0,1]} \left(1 - \frac{\lfloor M \cdot \theta \rceil}{M} \right) = \left[1 - \frac{1}{2M}, 1 \right], \\ \theta_{\max} &\in \operatorname*{arg\,max}_{\theta \in [0,1]} (\mathbb{E}[\mathrm{TNR}_\theta]) = \operatorname*{arg\,max}_{\theta \in [0,1]} \left(1 - \frac{\lfloor M \cdot \theta \rceil}{M} \right) = \left[0, \frac{1}{2M} \right),\end{aligned}$$

respectively. The discrete versions $\theta^*_{\min} \in \Theta^*$ and $\theta^*_{\max} \in \Theta^*$ of the optimizers are given by

$$\theta^*_{\min} \in \underset{\theta^* \in \Theta^*}{\arg\min} \{\mathbb{E}[\mathrm{TNR}_{\theta^*}]\} = \underset{\theta^* \in \Theta^*}{\arg\min} \{1 - \theta^*\} = \{1\},$$
$$\theta^*_{\max} \in \underset{\theta^* \in \Theta^*}{\arg\max} \{\mathbb{E}[\mathrm{TNR}_{\theta^*}]\} = \underset{\theta^* \in \Theta^*}{\arg\max} \{1 - \theta^*\} = \{0\}.$$

7. False Negative Rate

The *False Negative Rate* FNR_θ or *Miss Rate* is the performance measure that indicates the relative number of incorrectly predicted positive observations. Therefore, it can be seen as the counterpart to the True Positive Rate discussed in Sec. 5.

7.1. Definition and distribution

The False Negative Rate is commonly defined as

$$\mathrm{FNR}_\theta = \frac{\mathrm{FN}_\theta}{P}.$$

Hence, $P > 0$ should hold, otherwise, the denominator is zero. With the aid of Eq. (B3), FNR_θ can be reformulated to

$$\mathrm{FNR}_\theta = \frac{P - \mathrm{TP}_\theta}{P} = 1 - \frac{\mathrm{TP}_\theta}{P}.$$

Thus, it is linear in TP_θ and can therefore be written as

$$\mathrm{FNR}_\theta = X_\theta\left(-\frac{1}{P}, 1\right) \sim f_{X_\theta}\left(-\frac{1}{P}, 1\right),$$

and for its range:

$$\mathrm{FNR}_\theta \overset{(R)}{\in} \mathcal{R}\left(X_\theta\left(-\frac{1}{P}, 1\right)\right).$$

7.2. Expectation

Because FNR_θ is linear in TP_θ with slope $a = -1/P$ and intercept $b = 1$, its expectation is

$$\mathbb{E}[\mathrm{FNR}_\theta] = \mathbb{E}\left[X_\theta\left(-\frac{1}{P}, 1\right)\right] \overset{(1)}{=} -\frac{1}{P} \cdot \mathbb{E}[\mathrm{TP}_\theta] + 1 = 1 - \frac{\lfloor M \cdot \theta \rceil}{M} = 1 - \theta^*.$$

7.3. Optimal baselines

The range of the expectation of FNR_θ determines the baselines. The extreme values are given by:

$$\min_{\theta \in [0,1]} (\mathbb{E}[\mathrm{FNR}_\theta]) = \min_{\theta \in [0,1]} \left(1 - \frac{\lfloor M \cdot \theta \rceil}{M}\right) = 0,$$
$$\max_{\theta \in [0,1]} (\mathbb{E}[\mathrm{FNR}_\theta]) = \max_{\theta \in [0,1]} \left(1 - \frac{\lfloor M \cdot \theta \rceil}{M}\right) = 1.$$

Furthermore, the optimizers $\theta_{\min} \in [0,1]$ and $\theta_{\max} \in [0,1]$ for the extreme values are as follows:

$$\theta_{\min} \in \operatorname*{arg\,min}_{\theta\in[0,1]} (\mathbb{E}[\mathrm{FNR}_\theta]) = \operatorname*{arg\,min}_{\theta\in[0,1]} \left(1 - \frac{\lfloor M\cdot\theta\rceil}{M}\right) = \left[1-\frac{1}{2M}, 1\right],$$

$$\theta_{\max} \in \operatorname*{arg\,max}_{\theta\in[0,1]} (\mathbb{E}[\mathrm{FNR}_\theta]) = \operatorname*{arg\,max}_{\theta\in[0,1]} \left(1 - \frac{\lfloor M\cdot\theta\rceil}{M}\right) = \left[0, \frac{1}{2M}\right),$$

respectively. The discrete versions $\theta^*_{\min} \in \Theta^*$ and $\theta^*_{\max} \in \Theta^*$ of the optimization values are then:

$$\theta^*_{\min} \in \operatorname*{arg\,min}_{\theta^*\in\Theta^*} \{\mathbb{E}[\mathrm{FNR}_{\theta^*}]\} = \operatorname*{arg\,min}_{\theta^*\in\Theta^*} \{1-\theta^*\} = \{1\},$$

$$\theta^*_{\max} \in \operatorname*{arg\,max}_{\theta^*\in\Theta^*} \{\mathbb{E}[\mathrm{FNR}_{\theta^*}]\} = \operatorname*{arg\,max}_{\theta^*\in\Theta^*} \{1-\theta^*\} = \{0\}.$$

8. False Positive Rate

The *False Positive Rate* FPR_θ or *Fall-out* is the performance measure that shows the fraction of incorrectly predicted negative observations. Hence, it can be seen as the counterpart to the True Negative Rate that is introduced in Sec. 6.

8.1. Definition and distribution

The False Positive Rate is commonly defined as

$$\mathrm{FPR}_\theta = \frac{\mathrm{FP}_\theta}{N}.$$

Hence, $N := M - P$ should hold, otherwise, the denominator is zero. By using Eq. (B2), FPR_θ can be restated as

$$\mathrm{FPR}_\theta = \frac{\lfloor M\cdot\theta\rceil - \mathrm{TP}_\theta}{M-P}. \tag{6}$$

Note that it is linear in TP_θ and can therefore be written as

$$\mathrm{FPR}_\theta = X_\theta\left(-\frac{1}{M-P}, \frac{\lfloor M\cdot\theta\rceil}{M-P}\right) \sim f_{X_\theta}\left(-\frac{1}{M-P}, \frac{\lfloor M\cdot\theta\rceil}{M-P}\right),$$

with range:

$$\mathrm{FPR}_\theta \overset{(R)}{\in} \mathcal{R}\left(X_\theta\left(-\frac{1}{M-P}, \frac{\lfloor M\cdot\theta\rceil}{M-P}\right)\right).$$

8.2. Expectation

Since FPR_θ is linear in TP_θ with slope $a = -1/(M-P)$ and intercept $b = \lfloor M\cdot\theta\rceil/(M-P)$, its expectation is given by

$$\mathbb{E}[\mathrm{FPR}_\theta] = \mathbb{E}\left[X_\theta\left(-\frac{1}{M-P}, \frac{\lfloor M\cdot\theta\rceil}{M-P}\right)\right] \overset{(1)}{=} -\frac{1}{M-P}\cdot\mathbb{E}[\mathrm{TP}_\theta] + \frac{\lfloor M\cdot\theta\rceil}{M-P} = \frac{\lfloor M\cdot\theta\rceil}{M} = \theta^*.$$

8.3. Optimal baselines

The extreme values of the expectation of FPR_θ determine the baselines. The range is given by

$$\min_{\theta \in [0,1]} (\mathbb{E}[\text{FPR}_\theta]) = \min_{\theta \in [0,1]} \left(\frac{\lfloor M \cdot \theta \rceil}{M} \right) = 0,$$
$$\max_{\theta \in [0,1]} (\mathbb{E}[\text{FPR}_\theta]) = \max_{\theta \in [0,1]} \left(\frac{\lfloor M \cdot \theta \rceil}{M} \right) = 1.$$

Moreover, the optimizers $\theta_{\text{min}} \in [0, 1]$ and $\theta_{\text{max}} \in [0, 1]$ for the extreme values are determined by

$$\theta_{\text{min}} \in \underset{\theta \in [0,1]}{\arg\min} (\mathbb{E}[\text{FPR}_\theta]) = \underset{\theta \in [0,1]}{\arg\min} \left(\frac{\lfloor M \cdot \theta \rceil}{M} \right) = \left[0, \frac{1}{2M} \right),$$
$$\theta_{\text{max}} \in \underset{\theta \in [0,1]}{\arg\max} (\mathbb{E}[\text{FPR}_\theta]) = \underset{\theta \in [0,1]}{\arg\max} \left(\frac{\lfloor M \cdot \theta \rceil}{M} \right) = \left[1 - \frac{1}{2M}, 1 \right],$$

respectively. The discrete forms $\theta^*_{\text{min}} \in \Theta^*$ and $\theta^*_{\text{max}} \in \Theta^*$ of these are then

$$\theta^*_{\text{min}} \in \underset{\theta^* \in \Theta^*}{\arg\min} \{\mathbb{E}[\text{FNR}_{\theta^*}]\} = \underset{\theta^* \in \Theta^*}{\arg\min} \{\theta^*\} = \{0\},$$
$$\theta^*_{\text{max}} \in \underset{\theta^* \in \Theta^*}{\arg\max} \{\mathbb{E}[\text{FNR}_{\theta^*}]\} = \underset{\theta^* \in \Theta^*}{\arg\max} \{\theta^*\} = \{1\}.$$

9. Positive Predictive Value

The *Positive Predictive Value* PPV_θ or *Precision* is the performance measure that considers the fraction of all positively predicted observations that are, in fact, positive. Therefore, it provides an indication of how cautious the model is in assigning positive predictions. A large value means the model is cautious in predicting observations as positive, while a small value means the opposite.

9.1. Definition and distribution

The Positive Predictive Value is commonly defined as

$$\text{PPV}_\theta = \frac{\text{TP}_\theta}{\text{TP}_\theta + \text{FP}_\theta}. \tag{7}$$

By using Eq. (B1) and (B2), this definition can be reformulated to

$$\text{PPV}_\theta = \frac{\text{TP}_\theta}{\lfloor M \cdot \theta \rceil}.$$

Note that this performance measure is only defined whenever $\lfloor M \cdot \theta \rceil > 0$, otherwise the denominator is zero. Therefore, we assume specifically for PPV_θ that $\theta \geq \frac{1}{2M}$. The definition of PPV_θ is linear in TP_θ and can thus be formulated as

$$\text{PPV}_\theta = X_\theta \left(\frac{1}{\lfloor M \cdot \theta \rceil}, 0 \right) \sim f_{X_\theta} \left(\frac{1}{\lfloor M \cdot \theta \rceil}, 0 \right), \tag{8}$$

with range:

$$\mathrm{PPV}_\theta \overset{(R)}{\in} \mathcal{R}\left(X_\theta\left(\frac{1}{\lfloor M \cdot \theta \rceil}, 0\right)\right).$$

9.2. Expectation

Because PPV_θ is linear in TP_θ with slope $a = 1/\lfloor M \cdot \theta \rceil$ and intercept $b = 0$, its expectation is

$$\mathbb{E}[\mathrm{PPV}_\theta] = \mathbb{E}\left[X_\theta\left(\frac{1}{\lfloor M \cdot \theta \rceil}, 0\right)\right] \overset{(1)}{=} \frac{1}{\lfloor M \cdot \theta \rceil} \cdot \mathbb{E}[\mathrm{TP}_\theta] + 0 = \frac{P}{M}.$$

9.3. Optimal baselines

The baselines are determined by the extreme values of the expectation of PPV_θ:

$$\min_{\theta \in [1/(2M), 1]} (\mathbb{E}[\mathrm{PPV}_\theta]) = \frac{P}{M},$$
$$\max_{\theta \in [1/(2M), 1]} (\mathbb{E}[\mathrm{PPV}_\theta]) = \frac{P}{M},$$

because the expectation does not depend on θ. Hence, the optimization values $\theta_{\min}$ and $\theta_{\max}$ are simply all allowed values for θ:

$$\theta_{\min} = \theta_{\max} \in \left[\frac{1}{2M}, 1\right].$$

Consequently, the discrete versions $\theta^*_{\min}$ and $\theta^*_{\max}$ of these optimizers are in the set of all allowed discrete values:

$$\theta^*_{\min} = \theta^*_{\max} \in \Theta^* \setminus \{0\}.$$

10. Negative Predictive Value

The *Negative Predictive Value* NPV_θ is the performance measure that indicates the fraction of all negatively predicted observations that are, in fact, negative. Hence, it shows how cautious the model is in assigning negative predictions. A large value means the model is cautious in predicting observations negatively, while a small value means the opposite.

10.1. Definition and distribution

The *Negative Predictive Value* is commonly defined as

$$\mathrm{NPV}_\theta = \frac{\mathrm{TN}_\theta}{\mathrm{TN}_\theta + \mathrm{FN}_\theta}.$$

With the help of Eq. (B3) and (B4), this definition can be rewritten as

$$\mathrm{NPV}_\theta = 1 - \frac{P - \mathrm{TP}_\theta}{M - \lfloor M \cdot \theta \rceil}.$$

Note that this performance measure is only defined whenever $\lfloor M \cdot \theta \rceil < M$, otherwise the denominator is zero. Therefore, we assume specifically for NPV_θ that $\theta < 1 - \frac{1}{2M}$. The definition of NPV_θ is linear in TP_θ and can thus be formulated as

$$\mathrm{NPV}_\theta = X_\theta \left(\frac{1}{M - \lfloor M \cdot \theta \rceil}, 1 - \frac{P}{M - \lfloor M \cdot \theta \rceil} \right) \sim f_{X_\theta} \left(\frac{1}{M - \lfloor M \cdot \theta \rceil}, 1 - \frac{P}{M - \lfloor M \cdot \theta \rceil} \right), \tag{9}$$

with range:

$$\mathrm{NPV}_\theta \overset{(R)}{\in} \mathcal{R} \left(X_\theta \left(\frac{1}{M - \lfloor M \cdot \theta \rceil}, 1 - \frac{P}{M - \lfloor M \cdot \theta \rceil} \right) \right).$$

10.2. Expectation

Since NPV_θ is linear in TP_θ with slope $a = 1/(M - \lfloor M \cdot \theta \rceil)$ and intercept $b = 1 - P/(M - \lfloor M \cdot \theta \rceil)$, its expectation is given by

$$\begin{aligned} \mathbb{E}[\mathrm{NPV}_\theta] &= \mathbb{E}\left[X_\theta \left(\frac{1}{M - \lfloor M \cdot \theta \rceil}, 1 - \frac{P}{M - \lfloor M \cdot \theta \rceil} \right) \right] \\ &\overset{(1)}{=} \frac{1}{M - \lfloor M \cdot \theta \rceil} \cdot \mathbb{E}[\mathrm{TP}_\theta] + 1 - \frac{P}{M - \lfloor M \cdot \theta \rceil} = 1 - \frac{P}{M}. \end{aligned}$$

10.3. Optimal baselines

The extreme values of the expectation of NPV_θ determine the baselines. They are given by

$$\begin{aligned} \min_{\theta \in [0, 1 - 1/(2M))} (\mathbb{E}[\mathrm{NPV}_\theta]) &= 1 - \frac{P}{M}, \\ \max_{\theta \in [0, 1 - 1/(2M))} (\mathbb{E}[\mathrm{NPV}_\theta]) &= 1 - \frac{P}{M}, \end{aligned}$$

because the expectation does not depend on θ. Consequently, the optimization values θ_{min} and θ_{max} are all allowed values for θ:

$$\theta_{\mathrm{min}} = \theta_{\mathrm{max}} \in \left[0, 1 - \frac{1}{2M} \right).$$

This also means the discrete forms θ^*_{min} and θ^*_{max} of the optimizers are in the set of all allowed discrete values:

$$\theta^*_{\mathrm{min}} = \theta^*_{\mathrm{max}} \in \Theta^* \setminus \{1\}.$$

11. False Discovery Rate

The *False Discovery Rate* FDR_θ is the performance measure that looks at the fraction of positively predicted observations that are actually negative. Therefore, it can be seen as the counterpart to the Positive Predictive Value that we discuss in Sec. 9. Consequently, a small value means the model is cautious in predicting observations as positive, while a large value means the opposite.

107 **11.1. Definition and distribution**

The *False Discovery Rate* is commonly defined as

$$\mathrm{FDR}_\theta = \frac{\mathrm{FP}_\theta}{\mathrm{TP}_\theta + \mathrm{FP}_\theta} = 1 - \mathrm{PPV}_\theta.$$

With the help of Eq. (8), this definition can be rewritten as

$$\mathrm{FDR}_\theta = 1 - \frac{\mathrm{TP}_\theta}{\lfloor M \cdot \theta \rceil}.$$

Note that this performance measure is only defined whenever $\lfloor M \cdot \theta \rceil > 0$, otherwise the denominator is zero. Therefore, we assume specifically for FDR_θ that $\theta > \frac{1}{2M}$. The definition of FDR_θ is linear in TP_θ and can thus be formulated as

$$\mathrm{FDR}_\theta = X_\theta\left(-\frac{1}{\lfloor M \cdot \theta \rceil}, 1\right) \sim f_{X_\theta}\left(-\frac{1}{\lfloor M \cdot \theta \rceil}, 1\right),$$

with range:

$$\mathrm{FDR}_\theta \overset{(R)}{\in} \mathcal{R}\left(X_\theta\left(-\frac{1}{\lfloor M \cdot \theta \rceil}, 1\right)\right).$$

108 **11.2. Expectation**

Since FDR_θ is linear in TP_θ with slope $a = -1/\lfloor M \cdot \theta \rceil$ and intercept $b = 1$, its expectation is given by

$$\mathbb{E}[\mathrm{FDR}_\theta] = \mathbb{E}\left[X_\theta\left(-\frac{1}{\lfloor M \cdot \theta \rceil}, 1\right)\right] \overset{(1)}{=} -\frac{1}{\lfloor M \cdot \theta \rceil} \cdot \mathbb{E}[\mathrm{TP}_\theta] + 1 = 1 - \frac{P}{M}.$$

109 **11.3. Optimal baselines**

The extreme values of the expectation of FDR_θ determine the baselines. Its range is given by

$$\min_{\theta \in (1/(2M), 1]} (\mathbb{E}[\mathrm{FDR}_\theta]) = 1 - \frac{P}{M},$$
$$\max_{\theta \in (1/(2M), 1]} (\mathbb{E}[\mathrm{FDR}_\theta]) = 1 - \frac{P}{M},$$

because the expectation does not depend on θ. Consequently, the optimization values $\theta_{\min}$ and $\theta_{\max}$ are all allowed values for θ:

$$\theta_{\min} = \theta_{\max} \in \left(\frac{1}{2M}, 1\right].$$

This also means the discrete forms $\theta^*_{\min}$ and $\theta^*_{\max}$ of the optimizers are in the set of all allowed discrete values:

$$\theta^*_{\min} = \theta^*_{\max} \in \Theta^* \setminus \{0\}.$$

12. False Omission Rate

The *False Omission Rate* FOR_θ is the performance measure that considers the fraction of observations that are predicted negative but are in fact positive. Hence, it can be seen as the counterpart to the Negative Predictive Value introduced in Sec. 10. Consequently, a small value means the model is cautious in negatively predicting observations, while a large value means the opposite.

12.1. Definition and distribution

The *False Omission Rate* is commonly defined as

$$\mathrm{FOR}_\theta = \frac{\mathrm{FN}_\theta}{\mathrm{TN}_\theta + \mathrm{FN}_\theta}.$$

With the aid of Eq. (B3), this can be reformulated to

$$\mathrm{FOR}_\theta = \frac{P - \mathrm{TP}_\theta}{M - \lfloor M \cdot \theta \rceil}.$$

Note that this performance measure is only defined whenever $\lfloor M \cdot \theta \rceil < M$, otherwise the denominator is zero. Therefore, we assume specifically for FOR_θ that $\theta < 1 - \frac{1}{2M}$. Now, FOR_θ is linear in TP_θ and can therefore be written as

$$\mathrm{FOR}_\theta = X_\theta \left(-\frac{1}{M - \lfloor M \cdot \theta \rceil}, \frac{P}{M - \lfloor M \cdot \theta \rceil} \right) \sim f_{X_\theta} \left(-\frac{1}{M - \lfloor M \cdot \theta \rceil}, \frac{P}{M - \lfloor M \cdot \theta \rceil} \right),$$

with range:

$$\mathrm{FOR}_\theta \overset{(R)}{\in} \mathcal{R} \left(X_\theta \left(-\frac{1}{M - \lfloor M \cdot \theta \rceil}, \frac{P}{M - \lfloor M \cdot \theta \rceil} \right) \right).$$

12.2. Expectation

Because FOR_θ is linear in TP_θ with slope $a = -1/(M - \lfloor M \cdot \theta \rceil)$ and intercept $b = P/(M - \lfloor M \cdot \theta \rceil)$, its expectation is

$$\begin{aligned}\mathbb{E}[\mathrm{FOR}_\theta] &= \mathbb{E} \left[X_\theta \left(-\frac{1}{M - \lfloor M \cdot \theta \rceil}, \frac{P}{M - \lfloor M \cdot \theta \rceil} \right) \right] \\ &\overset{(1)}{=} -\frac{1}{M - \lfloor M \cdot \theta \rceil} \cdot \mathbb{E}[\mathrm{TP}_\theta] + \frac{P}{M - \lfloor M \cdot \theta \rceil} = \frac{P}{M}.\end{aligned}$$

12.3. Optimal baselines

The range of the expectation of FOR_θ determines the baselines. The extreme values are defined as

$$\min_{\theta \in [0, 1-1/(2M))} (\mathbb{E}[\mathrm{FOR}_\theta]) = \frac{P}{M},$$

$$\max_{\theta \in [0, 1-1/(2M))} (\mathbb{E}[\mathrm{FOR}_\theta]) = \frac{P}{M},$$

because the expectation does not depend on θ. Consequently, the optimization values $\theta_{\min}$ and $\theta_{\max}$ are all allowed values for θ:

$$\theta_{\min} = \theta_{\max} \in \left[0, 1 - \frac{1}{2M}\right).$$

This also means the discrete forms $\theta^*_{\min}$ and $\theta^*_{\max}$ of the optimizers are in the set of all allowed discrete values:

$$\theta^*_{\min} = \theta^*_{\max} \in \Theta^* \setminus \{1\}.$$

13. F_β score

The F_β *score* $\mathrm{F}_\theta^{(\beta)}$ was introduced by Chinchor (1992). It is the weighted harmonic average between the True Positive Rate (TPR_θ) and the Positive Predictive Value (PPV_θ). These two performance measures are discussed extensively in Sec. 5 and 9. The F_β score balances predicting the actual positive observations correctly (TPR_θ) and being cautious in predicting observations as positive (PPV_θ). The factor $\beta > 0$ indicates how much more TPR_θ is weighted compared to PPV_θ.

13.1. Definition and distribution

The F_β score is commonly defined as

$$\mathrm{F}_\theta^{(\beta)} = \frac{1+\beta^2}{\frac{1}{\mathrm{PPV}_\theta} + \frac{\beta^2}{\mathrm{TPR}_\theta}}.$$

By using the definitions of TPR_θ and PPV_θ in Eq. (3) and (7), $\mathrm{F}_\theta^{(\beta)}$ can be formulated in terms of the base measures:

$$\mathrm{F}_\theta^{(\beta)} = \frac{(1+\beta^2)\cdot \mathrm{TP}_\theta}{\beta^2 \cdot P + \mathrm{TP}_\theta + \mathrm{FP}_\theta}$$

Eq. (B1) and (B2) allow us to write the formulation above in terms of only TP_θ:

$$\mathrm{F}_\theta^{(\beta)} = \frac{(1+\beta^2)\cdot \mathrm{TP}_\theta}{\beta^2 \cdot P + \lfloor M \cdot \theta \rceil}.$$

Note that $P > 0$ and $\lfloor M \cdot \theta \rceil > 0$, otherwise TPR_θ or PPV_θ is not defined, and hence, $\mathrm{F}_\theta^{(\beta)}$ is not defined. Now, $\mathrm{F}_\theta^{(\beta)}$ is linear in TP_θ and can be formulated as

$$\mathrm{F}_\theta^{(\beta)} = X_\theta\left(\frac{1+\beta^2}{\beta^2 \cdot P + \lfloor M \cdot \theta \rceil}, 0\right),$$

with range:

$$\mathrm{F}_\theta^{(\beta)} \overset{(R)}{\in} \mathcal{R}\left(X_\theta\left(\frac{1+\beta^2}{\beta^2 \cdot P + \lfloor M \cdot \theta \rceil}, 0\right)\right).$$

127 **13.2. Expectation**

Because $\mathrm{F}_\theta^{(\beta)}$ is linear in TP_θ with slope $a = (1+\beta^2)/(\beta^2 P + \lfloor M\cdot\theta\rceil)$ and intercept $b = 0$, its expectation is given by

$$\begin{aligned}\mathbb{E}[\mathrm{F}_\theta^{(\beta)}] &= \mathbb{E}\left[X_\theta\left(\frac{1+\beta^2}{\beta^2\cdot P + \lfloor M\cdot\theta\rceil}, 0\right)\right] \overset{(1)}{=} \frac{1+\beta^2}{\beta^2\cdot P + \lfloor M\cdot\theta\rceil}\cdot\mathbb{E}[\mathrm{TP}_\theta] + 0\\ &= \frac{\lfloor M\cdot\theta\rceil\cdot P\cdot(1+\beta^2)}{M\cdot(\beta^2\cdot P + \lfloor M\cdot\theta\rceil)}\\ &= \frac{(1+\beta^2)\cdot P\cdot\theta^*}{\beta^2\cdot P + M\cdot\theta^*}. \end{aligned} \tag{10}$$

128 **13.3. Optimal baselines**

To determine the extreme values of the expectation of $\mathrm{F}_\theta^{(\beta)}$, and therefore the baselines, the derivative of the function $f : [0,1] \to [0,1]$ defined as

$$f(t) = \frac{(1+\beta^2)\cdot P\cdot t}{\beta^2\cdot P + M\cdot t}$$

is calculated. First note that $\mathbb{E}[\mathrm{F}_\theta^{(\beta)}] = f(\lfloor M\cdot\theta\rceil/M)$. The derivative is given by

$$\frac{\mathrm{d}f(t)}{\mathrm{d}t} = \frac{\beta^2(1+\beta^2)\cdot P^2}{(\beta^2\cdot P + M\cdot t)^2}.$$

It is strictly positive for all t in its domain; thus, f is strictly increasing in t. This means $\mathbb{E}[\mathrm{F}_\theta^{(\beta)}]$ given in Eq. (10) is non-decreasing in both θ and θ^*. This is because the term $\lfloor M\cdot\theta\rceil/M$ is non-decreasing in θ. Hence, the extreme values of the expectation of $\mathrm{F}_\theta^{(\beta)}$ are its border values:

$$\begin{aligned}\min_{\theta\in[1/(2M),1]}\left(\mathbb{E}[\mathrm{F}_\theta^{(\beta)}]\right) &= \min_{\theta\in[1/(2M),1]}\left(\frac{(1+\beta^2)\cdot P\cdot\lfloor M\cdot\theta\rceil}{M(\beta^2\cdot P + \lfloor M\cdot\theta\rceil)}\right) = \frac{(1+\beta^2)\cdot P}{M(\beta^2\cdot P + 1)},\\ \max_{\theta\in[1/(2M),1]}\left(\mathbb{E}[\mathrm{F}_\theta^{(\beta)}]\right) &= \max_{\theta\in[1/(2M),1]}\left(\frac{(1+\beta^2)\cdot P\cdot\lfloor M\cdot\theta\rceil}{M(\beta^2\cdot P + \lfloor M\cdot\theta\rceil)}\right) = \frac{(1+\beta^2)\cdot P}{\beta^2\cdot P + M}.\end{aligned}$$

Consequently, the optimization values $\theta_{\min}$ and $\theta_{\max}$ for the extreme values are given by

$$\theta_{\min} \in \operatorname*{arg\,min}_{\theta\in[1/(2M),1]}\left(\mathbb{E}[\mathrm{F}_\theta^{(\beta)}]\right) = \operatorname*{arg\,min}_{\theta\in[1/(2M),1]}\left(\frac{\lfloor M\cdot\theta\rceil}{\beta^2\cdot P + \lfloor M\cdot\theta\rceil}\right) = \begin{cases}[\frac{1}{2},1] & \text{if } M = 1\\ \left[\frac{1}{2M},\frac{3}{2M}\right) & \text{if } M > 1,\end{cases}$$

$$\theta_{\max} \in \operatorname*{arg\,max}_{\theta\in[1/(2M),1]}\left(\mathbb{E}[\mathrm{F}_\theta^{(\beta)}]\right) = \operatorname*{arg\,max}_{\theta\in[1/(2M),1]}\left(\frac{\lfloor M\cdot\theta\rceil}{\beta^2\cdot P + \lfloor M\cdot\theta\rceil}\right) = \begin{cases}[\frac{1}{2},1] & \text{if } M = 1\\ \left[1-\frac{1}{2M},1\right] & \text{if } M > 1,\end{cases}$$

respectively. Following this reasoning, the discrete forms θ^*_{min} and θ^*_{max} are given by

$$\theta^*_{\text{min}} \in \underset{\theta^* \in \Theta^* \setminus \{0\}}{\arg\min} \left\{ \mathbb{E}[\mathrm{F}^{(\beta)}_{\theta^*}] \right\} = \underset{\theta^* \in \Theta^* \setminus \{0\}}{\arg\min} \left\{ \frac{\theta^*}{\beta^2 \cdot P + M \cdot \theta^*} \right\} = \left\{ \frac{1}{M} \right\},$$
$$\theta^*_{\text{max}} \in \underset{\theta^* \in \Theta^* \setminus \{0\}}{\arg\max} \left\{ \mathbb{E}[\mathrm{F}^{(\beta)}_{\theta^*}] \right\} = \underset{\theta^* \in \Theta^* \setminus \{0\}}{\arg\max} \left\{ \frac{\theta^*}{\beta^2 \cdot P + M \cdot \theta^*} \right\} = \{1\}.$$

14. Youden's J Statistic

The *Youden's J Statistic* J_θ, *Youden's Index*, or *(Bookmaker) Informedness* was introduced by Youden (1950) to capture the performance of a diagnostic test as a single statistic. It incorporates both the True Positive and True Negative rates, discussed in Sec. 5 and 6, respectively. Youden's J Statistic shows how well the model can correctly predict both the positive as well as the negative observations.

14.1. Definition and distribution

The Youden's J Statistic is commonly defined as

$$\mathrm{J}_\theta = \mathrm{TPR}_\theta + \mathrm{TNR}_\theta - 1.$$

By using Eq. (4) and (5), which provide the definitions of TPR_θ and TNR_θ in terms of TP_θ, the definition of J_θ can be reformulated as

$$\mathrm{J}_\theta = \frac{M \cdot \mathrm{TP}_\theta - P \cdot \lfloor M \cdot \theta \rceil}{P\,(M - P)}.$$

Because TPR_θ needs $P > 0$, and TNR_θ needs $N > 0$, we have both these assumptions for J_θ. Consequently, $M > 1$. Now, J_θ is linear in TP_θ and can therefore be written as

$$\mathrm{J}_\theta = X_\theta \left(\frac{M}{P\,(M-P)}, -\frac{\lfloor M \cdot \theta \rceil}{M - P} \right) \sim f_{X_\theta} \left(\frac{M}{P\,(M-P)}, -\frac{\lfloor M \cdot \theta \rceil}{M - P} \right),$$

with range:

$$\mathrm{J}_\theta \overset{(R)}{\in} \mathcal{R} \left(X_\theta \left(\frac{M}{P\,(M-P)}, -\frac{\lfloor M \cdot \theta \rceil}{M - P} \right) \right).$$

14.2. Expectation

Since J_θ is linear in TP_θ with slope $a = M/(P\,(M-P))$ and intercept $b = -\lfloor M \cdot \theta \rceil/(M-P)$, its expectation is given by

$$\mathbb{E}[\mathrm{J}_\theta] = \mathbb{E}\left[X_\theta \left(\frac{M}{P\,(M-P)}, -\frac{\lfloor M \cdot \theta \rceil}{M - P} \right) \right] \overset{(1)}{=} \frac{M}{P\,(M-P)} \cdot \mathbb{E}[\mathrm{TP}_\theta] - \frac{\lfloor M \cdot \theta \rceil}{M - P} = 0.$$

14.3. Optimal baselines

The extreme values of the expectation of J_θ determine the baselines. They are given by

$$\min_{\theta\in[0,1]} (\mathbb{E}[J_\theta]) = 0,$$
$$\max_{\theta\in[0,1]} (\mathbb{E}[J_\theta]) = 0,$$

because the expected value does not depend on θ. Consequently, the optimization values $\theta_{\min}$ and $\theta_{\max}$ can be any value in the domain of θ:

$$\theta_{\min} = \theta_{\max} \in [0, 1].$$

This also holds for the discrete forms $\theta^*_{\min}$ and $\theta^*_{\max}$ of the optimizers:

$$\theta^*_{\min} = \theta^*_{\max} \in \Theta^*.$$

15. Markedness

The *Markedness* MK_θ or *deltaP* is a performance measure mostly used in linguistics and social sciences. It combines both the Positive Predictive Value and the Negative Predictive Value. These two measures are discussed in Sec. 9 and 10. The Markedness indicates how cautious the model is in predicting observations as positive and also how cautious it is in predicting them as negative.

15.1. Definition and distribution

Markedness is commonly defined as

$$\text{MK}_\theta = \text{PPV}_\theta + \text{NPV}_\theta - 1.$$

This definition of MK_θ can be reformulated in terms of TP_θ by using Eq. (8) and (9):

$$\text{MK}_\theta = \frac{M \cdot \text{TP}_\theta - P \cdot \lfloor M \cdot \theta \rceil}{\lfloor M \cdot \theta \rceil (M - \lfloor M \cdot \theta \rceil)}.$$

Note that MK_θ is only defined for $M > 1$ and $\theta \in [1/(2M), 1 - 1/(2M))$, otherwise the denominator becomes zero. The assumption $M > 1$ automatically follows from the assumptions $\hat{P} > 0$ and $\hat{N} > 0$, which hold for PPV_θ and NPV_θ, respectively. In other words, at least one observation predicted positive and at least one predicted negative; thus, $M > 1$. Now, MK_θ is linear in TP_θ and can therefore be written as

$$\begin{aligned}\text{MK}_\theta &= X_\theta \left(\frac{M}{\lfloor M \cdot \theta \rceil (M - \lfloor M \cdot \theta \rceil)}, -\frac{P}{M - \lfloor M \cdot \theta \rceil} \right) \\ &\sim f_{X_\theta} \left(\frac{M}{\lfloor M \cdot \theta \rceil (M - \lfloor M \cdot \theta \rceil)}, -\frac{P}{M - \lfloor M \cdot \theta \rceil} \right),\end{aligned}$$

with range:

$$\text{MK}_\theta \overset{(R)}{\in} \mathcal{R} \left(X_\theta \left(\frac{M}{\lfloor M \cdot \theta \rceil (M - \lfloor M \cdot \theta \rceil)}, -\frac{P}{M - \lfloor M \cdot \theta \rceil} \right) \right).$$

15.2. Expectation

By using slope $a = M/(\lfloor M \cdot \theta \rceil (M - \lfloor M \cdot \theta \rceil))$ and intercept $b = -P/(M - \lfloor M \cdot \theta \rceil)$, the expectation of MK_θ can be calculated:

$$\begin{aligned}\mathbb{E}[\mathrm{MK}_\theta] &= \mathbb{E}\left[X_\theta\left(\frac{M}{\lfloor M \cdot \theta \rceil (M - \lfloor M \cdot \theta \rceil)}, -\frac{P}{M - \lfloor M \cdot \theta \rceil}\right)\right] \\ &\overset{(1)}{=} \frac{M}{\lfloor M \cdot \theta \rceil (M - \lfloor M \cdot \theta \rceil)} \cdot \mathbb{E}[\mathrm{TP}_\theta] - \frac{P}{M - \lfloor M \cdot \theta \rceil} = 0.\end{aligned}$$

15.3. Optimal baselines

The extreme values of the expectation of MK_θ determine the baselines. Its range is given by:

$$\min_{\theta \in [1/(2M), 1-1/(2M))} (\mathbb{E}[\mathrm{MK}_\theta]) = 0,$$
$$\max_{\theta \in [1/(2M), 1-1/(2M))} (\mathbb{E}[\mathrm{MK}_\theta]) = 0,$$

since the expected value does not depend on θ. Therefore, the optimization values $\theta_{\min}$ and $\theta_{\max}$ are in the set of allowed values for θ:

$$\theta_{\min} = \theta_{\max} \in \left[\frac{1}{2M}, 1 - \frac{1}{2M}\right).$$

This also means the discrete forms $\theta^*_{\min}$ and $\theta^*_{\max}$ of the optimizers are in the set of the allowed discrete values:

$$\theta^*_{\min} = \theta^*_{\max} \in \Theta^* \setminus \{0, 1\}.$$

16. Accuracy

The *Accuracy* Acc_θ is the performance measure that assesses how good the model is in correctly predicting the observations without distinguishing between positive or negative observations.

16.1. Definition and distribution

The Accuracy is commonly defined as

$$\mathrm{Acc}_\theta = \frac{\mathrm{TP}_\theta + \mathrm{TN}_\theta}{M}.$$

By using Eq. (B4), this can be restated as

$$\mathrm{Acc}_\theta = \frac{2 \cdot \mathrm{TP}_\theta + M - P - \lfloor M \cdot \theta \rceil}{M}.$$

Note that it is linear in TP_θ and can therefore be written as

$$\mathrm{Acc}_\theta = X_\theta\left(\frac{2}{M}, \frac{M - P - \lfloor M \cdot \theta \rceil}{M}\right) \sim f_{X_\theta}\left(\frac{2}{M}, \frac{M - P - \lfloor M \cdot \theta \rceil}{M}\right), \qquad (11)$$

with range:

$$\mathrm{Acc}_\theta \overset{(R)}{\in} \mathcal{R}\left(X_\theta\left(\frac{2}{M}, \frac{M - P - \lfloor M \cdot \theta \rceil}{M}\right)\right).$$

153 **16.2. Expectation**

Since Acc_θ is linear in TP_θ with slope $a = 2/M$ and intercept $b = (M - P - \lfloor M \cdot \theta \rceil)/M$, its expectation can be derived:

$$\begin{aligned}\mathbb{E}[\mathrm{Acc}_\theta] &= \mathbb{E}\left[X_\theta\left(\frac{2}{M}, \frac{M - P - \lfloor M \cdot \theta \rceil}{M}\right)\right] \overset{(1)}{=} \frac{2}{M} \cdot \mathbb{E}[\mathrm{TP}_\theta] + \frac{M - P - \lfloor M \cdot \theta \rceil}{M} \\ &= \frac{(M - \lfloor M \cdot \theta \rceil)(M - P) + \lfloor M \cdot \theta \rceil \cdot P}{M^2} = \frac{(1 - \theta^*)(M - P) + \theta^* \cdot P}{M}. \quad (12)\end{aligned}$$

154 **16.3. Optimal baselines**

The range of the expectation of Acc_θ directly determines the baselines. To determine the extreme values of Acc_θ, the derivative of the function $f : [0, 1] \to [0, 1]$ defined as

$$f(t) = \frac{(1 - t)(M - P) + P \cdot t}{M}$$

is calculated. First, note that $\mathbb{E}[\mathrm{Acc}_\theta] = f(\lfloor M \cdot \theta \rceil / M)$. The derivative is given by

$$\frac{\mathrm{d}f(t)}{\mathrm{d}t} = \frac{2P - M}{M}.$$

It does not depend on t, but whether the derivative is positive or negative depends on P and M. Whenever $P > \frac{M}{2}$, then f is strictly increasing for all t in its domain. If $P < \frac{M}{2}$, then f is strictly decreasing. When $P = \frac{M}{2}$, f is constant. Consequently, the same holds for $\mathbb{E}[\mathrm{Acc}_\theta]$ given in Eq. (12). This is because the term $\lfloor M \cdot \theta \rceil / M$ is non-decreasing in θ. Thus, the extreme values of the expectation of Acc_θ are given by

$$\begin{aligned}\min_{\theta \in [0,1]} (\mathbb{E}[\mathrm{Acc}_\theta]) &= \begin{cases} \frac{P}{M} & \text{if } P < \frac{M}{2} \\ 1 - \frac{P}{M} & \text{if } P \geq \frac{M}{2} \end{cases} = \min\left\{\frac{P}{M}, 1 - \frac{P}{M}\right\}, \\ \max_{\theta \in [0,1]} (\mathbb{E}[\mathrm{Acc}_\theta]) &= \begin{cases} 1 - \frac{P}{M} & \text{if } P < \frac{M}{2} \\ \frac{P}{M} & \text{if } P \geq \frac{M}{2} \end{cases} = \max\left\{\frac{P}{M}, 1 - \frac{P}{M}\right\}.\end{aligned}$$

This means that the optimization values $\theta_{\min} \in [0,1]$ and $\theta_{\max} \in [0,1]$ for these extreme values are given by

$$\theta_{\min} \in \underset{\theta\in[0,1]}{\arg\min}\left(\mathbb{E}[\mathrm{Acc}_\theta]\right) = \begin{cases} \left[1-\frac{1}{2M}, 1\right] & \text{if } P < \frac{M}{2} \\ [0,1] & \text{if } P = \frac{M}{2} \\ \left[0, \frac{1}{2M}\right) & \text{if } P > \frac{M}{2}, \end{cases} \tag{13}$$

$$\theta_{\max} \in \underset{\theta\in[0,1]}{\arg\max}\left(\mathbb{E}[\mathrm{Acc}_\theta]\right) = \begin{cases} \left[0, \frac{1}{2M}\right) & \text{if } P < \frac{M}{2} \\ [0,1] & \text{if } P = \frac{M}{2} \\ \left[1-\frac{1}{2M}, 1\right] & \text{if } P > \frac{M}{2}, \end{cases} \tag{14}$$

respectively. Consequently, the discrete versions $\theta^*_{\min} \in \Theta^*$ and $\theta^*_{\max} \in \Theta^*$ of the optimizers are given by

$$\theta^*_{\min} \in \underset{\theta^*\in\Theta^*}{\arg\min}\left\{\mathbb{E}[\mathrm{Acc}_{\theta^*}]\right\} = \begin{cases} \{1\} & \text{if } P < \frac{M}{2} \\ \Theta^* & \text{if } P = \frac{M}{2} \\ \{0\} & \text{if } P > \frac{M}{2}, \end{cases} \tag{15}$$

$$\theta^*_{\max} \in \underset{\theta^*\in\Theta^*}{\arg\max}\left\{\mathbb{E}[\mathrm{Acc}_{\theta^*}]\right\} = \begin{cases} \{0\} & \text{if } P < \frac{M}{2} \\ \Theta^* & \text{if } P = \frac{M}{2} \\ \{1\} & \text{if } P > \frac{M}{2}, \end{cases} \tag{16}$$

respectively.

17. Balanced Accuracy

The *Balanced Accuracy* BAcc_θ is the mean of the True Positive Rate and True Negative Rate, which are discussed in Sec. 5 and 6. It determines how good the model is in correctly predicting the positive observations and in correctly predicting the negative observations on average.

17.1. Definition and distribution

The Balanced Accuracy is commonly defined as

$$\mathrm{BAcc}_\theta = \frac{1}{2} \cdot (\mathrm{TPR}_\theta + \mathrm{TNR}_\theta).$$

By using Eq. (4) and (5), this can be reformulated as

$$\mathrm{BAcc}_\theta = \frac{1}{2}\left(\frac{\mathrm{TP}_\theta}{P} + 1 - \frac{\lfloor M\cdot\theta\rceil - \mathrm{TP}_\theta}{M-P}\right) = \frac{M\cdot\mathrm{TP}_\theta}{2P\,(M-P)} + \frac{M-P-\lfloor M\cdot\theta\rceil}{2\,(M-P)}.$$

Note that $P > 0$ and $N > 0$ should hold, otherwise TPR_θ or TNR_θ is not defined. Consequently, $M > 1$. Note that BAcc_θ is linear in TP_θ and can therefore be written as

$$\mathrm{BAcc}_\theta = X_\theta\left(\frac{M}{2P\,(M-P)}, \frac{M-P-\lfloor M\cdot\theta\rceil}{2\,(M-P)}\right) \sim f_{X_\theta}\left(\frac{M}{2P\,(M-P)}, \frac{M-P-\lfloor M\cdot\theta\rceil}{2\,(M-P)}\right),$$

with range:

$$\mathrm{BAcc}_\theta \overset{(R)}{\in} \mathcal{R}\left(X_\theta\left(\frac{M}{2P(M-P)}, \frac{M-P-\lfloor M\cdot\theta\rceil}{2(M-P)}\right)\right).$$

17.2. Expectation

BAcc_θ is linear in TP_θ with slope $a = M/(2P(M-P))$ and intercept $b = (M - P - \lfloor M\cdot\theta\rceil)/(2(M-P))$, so its expectation can be derived:

$$\begin{aligned}\mathbb{E}[\mathrm{BAcc}_\theta] &= \mathbb{E}\left[X_\theta\left(\frac{M}{2P(M-P)}, \frac{M-P-\lfloor M\cdot\theta\rceil}{2(M-P)}\right)\right] \\ &\overset{(1)}{=} \frac{M}{2P(M-P)}\cdot\mathbb{E}[\mathrm{TP}_\theta] + \frac{M-P-\lfloor M\cdot\theta\rceil}{2(M-P)} = \frac{1}{2}.\end{aligned}$$

17.3. Optimal baselines

The baselines are directly determined by the ranges of the expectation of BAcc_θ. Since the expectation is constant, its extreme values are the same:

$$\begin{aligned}\min_{\theta\in[0,1]}\left(\mathbb{E}[\mathrm{BAcc}_\theta]\right) &= \frac{1}{2}, \\ \max_{\theta\in[0,1]}\left(\mathbb{E}[\mathrm{BAcc}_\theta]\right) &= \frac{1}{2}.\end{aligned}$$

This means that the optimization values $\theta_{\min} \in [0,1]$ and $\theta_{\max} \in [0,1]$ for these extreme values are simply

$$\begin{aligned}\theta_{\min} &\in \arg\min_{\theta\in[0,1]}\left(\mathbb{E}[\mathrm{BAcc}_\theta]\right) = [0,1], \\ \theta_{\max} &\in \arg\max_{\theta\in[0,1]}\left(\mathbb{E}[\mathrm{BAcc}_\theta]\right) = [0,1],\end{aligned}$$

respectively. Consequently, the discrete versions $\theta^*_{\min} \in \Theta^*$ and $\theta^*_{\max} \in \Theta^*$ of the optimizers are given by

$$\begin{aligned}\theta^*_{\min} &\in \arg\min_{\theta^*\in\Theta^*}\left\{\mathbb{E}[\mathrm{Acc}_{\theta^*}]\right\} = \Theta^*, \\ \theta^*_{\max} &\in \arg\max_{\theta^*\in\Theta^*}\left\{\mathbb{E}[\mathrm{Acc}_{\theta^*}]\right\} = \Theta^*,\end{aligned}$$

respectively.

18. Matthews Correlation Coefficient

The *Matthews Correlation Coefficient* MCC_θ was established by Matthews (1975). However, its definition is identical to that of the Yule phi coefficient, which was introduced by Yule (1912). The performance measure can be seen as the correlation coefficient between the actual and predicted classes. Hence, it is one of the few measures that lies in $[-1,1]$ instead of $[0,1]$.

171 18.1. Definition and distribution

The Matthews Correlation Coefficient is commonly defined as

$$\mathrm{MCC}_\theta = \frac{\mathrm{TP}_\theta \cdot \mathrm{TN}_\theta - \mathrm{FN}_\theta \cdot \mathrm{FP}_\theta}{\sqrt{(\mathrm{TP}_\theta + \mathrm{FP}_\theta)(\mathrm{TP}_\theta + \mathrm{FN}_\theta)(\mathrm{TN}_\theta + \mathrm{FP}_\theta)(\mathrm{TN}_\theta + \mathrm{FN}_\theta)}}.$$

172 By using Eq. (B2) and (B4), this definition can be reformulated as

$$\mathrm{MCC}_\theta = \frac{M \cdot \mathrm{TP}_\theta - P \cdot \lfloor M \cdot \theta \rceil}{\sqrt{\lfloor M \cdot \theta \rceil \cdot P\,(M-P)\,(M - \lfloor M \cdot \theta \rceil)}}. \tag{17}$$

To ensure the denominator is non-zero, the assumptions $P > 0$, $N > 0$, $\hat{P} := \lfloor M \cdot \theta \rceil > 0$, and $\hat{N} := M - \lfloor M \cdot \theta \rceil > 0$ must hold. If one of these assumptions is violated, then the denominator in Eq. (17) is zero, and MCC_θ is not defined. Therefore, we have for MCC_θ that $\frac{1}{2M} \leq \theta < 1 - \frac{1}{2M}$ and $M > 1$. Next, to improve readability we introduce the variable $C(M, P, \theta)$ to replace the denominator in Eq. (17):

$$C(M, P, \theta) := \sqrt{\lfloor M \cdot \theta \rceil \cdot P\,(M-P)\,(M - \lfloor M \cdot \theta \rceil)}.$$

The definition of MCC_θ is linear in TP_θ and can thus be formulated as

$$\mathrm{MCC}_\theta = X_\theta\left(\frac{M}{C(M,P,\theta)}, \frac{-P \cdot \lfloor M \cdot \theta \rceil}{C(M,P,\theta)}\right) \sim f_{X_\theta}\left(\frac{M}{C(M,P,\theta)}, \frac{-P \cdot \lfloor M \cdot \theta \rceil}{C(M,P,\theta)}\right),$$

with range:

$$\mathrm{MCC}_\theta \overset{(R)}{\in} \mathcal{R}\left(X_\theta\left(\frac{M}{C(M,P,\theta)}, \frac{-P \cdot \lfloor M \cdot \theta \rceil}{C(M,P,\theta)}\right)\right).$$

173 18.2. Expectation

MCC_θ is linear in TP_θ with slope $a = M/C(M,P,\theta)$ and intercept $b = -P \cdot \lfloor M \cdot \theta \rceil / C(M,P,\theta)$, so its expectation can be derived from Eq. (1):

$$\begin{aligned}\mathbb{E}[\mathrm{MCC}_\theta] &= \mathbb{E}\left[X_\theta\left(\frac{M}{C(M,P,\theta)}, \frac{-P \cdot \lfloor M \cdot \theta \rceil}{C(M,P,\theta)}\right)\right] \overset{(1)}{=} \frac{M}{C(M,P,\theta)} \cdot \mathbb{E}[\mathrm{TP}_\theta] - \frac{P \cdot \lfloor M \cdot \theta \rceil}{C(M,P,\theta)} \\ &= 0.\end{aligned}$$

174 18.3. Optimal baselines

The baselines are directly determined by the ranges of the expectation of MCC_θ. Since the expectation is constant, its extreme values are the same:

$$\min_{\theta \in [1/(2M), 1-1/(2M))} (\mathbb{E}[\mathrm{MCC}_\theta]) = 0,$$
$$\max_{\theta \in [1/(2M), 1-1/(2M))} (\mathbb{E}[\mathrm{MCC}_\theta]) = 0.$$

This means that the optimization values $\theta_{\min}$ and $\theta_{\max}$ for these extreme values are simply:

$$\theta_{\min} = \theta_{\max} \in \left[\frac{1}{2M}, 1 - \frac{1}{2M}\right),$$

respectively. Consequently, the discrete versions $\theta^*_{\min}$ and $\theta^*_{\max}$ of the optimizers are given by:

$$\theta^*_{\min} = \theta^*_{\max} \in \Theta^* \setminus \{0, 1\}.$$

19. Cohen's Kappa

Cohen's kappa κ_θ is a less straightforward performance measure than the other measures discussed in this research. It is used to quantify the inter-rater reliability for two raters of categorical observations (Kvålseth, 1989). In our case, we compare the first rater, which is the DD classifier, with the perfect rater, which assigns the true label to each observation.

19.1. Definition and distribution

Although there are several definitions for Cohen's kappa, here we choose the following:

$$\kappa_\theta = \frac{P_o^\theta - P_e^\theta}{1 - P_e^\theta},$$

with P_o^θ the Accuracy Acc_θ as defined in Sec. 16 and P_e^θ the probability that the shuffle approach assigns the true label by chance. These two values can be expressed in terms of the base measures as follows:

$$P_o^\theta = \mathrm{Acc}_\theta = \frac{\mathrm{TP}_\theta + \mathrm{TN}_\theta}{M},$$
$$P_e^\theta = \frac{(\mathrm{TP}_\theta + \mathrm{FP}_\theta) \cdot P + (\mathrm{TN}_\theta + \mathrm{FN}_\theta)\,(M - P)}{M^2}.$$

By using Eq. (11), (B1), (B2), (B3) and (B4) the above can be rewritten as

$$P_o^\theta = \frac{2 \cdot \mathrm{TP}_\theta + M - P - \lfloor M \cdot \theta \rceil}{M},$$
$$P_e^\theta = \frac{\lfloor M \cdot \theta \rceil \cdot P + (M - \lfloor M \cdot \theta \rceil)\,(M - P)}{M^2}.$$

Note that for κ_θ to be well-defined, we need $1 - P_e^\theta \neq 0$. In other words,

$$\lfloor M \cdot \theta \rceil \cdot P + (M - \lfloor M \cdot \theta \rceil)\,(M - P) \neq M^2.$$

This simplifies to

$$\frac{\lfloor M \cdot \theta \rceil}{M} \neq \frac{P}{2P - M}. \tag{18}$$

The left-hand side is by definition in the interval $[0, 1]$. For the right-hand side to be in that interval, we first need $P/(2P - M) \geq 0$. Since $P \geq 0$, that means $2P - M > 0$; hence, $P > \frac{M}{2}$. Secondly, $P/(2P - M) \leq 1$. Since we know $P > \frac{M}{2}$, we obtain $P \geq M$. This inequality reduces to $P = M$, because P is always at most M. Whenever $P = M$, then Eq. (18) becomes

$$\frac{\lfloor M \cdot \theta \rceil}{M} \neq 1.$$

To summarize, when $P < M$, then all $\theta \in [0, 1]$ are allowed in κ_θ, but when $P = M$, then $\theta < 1 - 1/(2M)$.

Now, by using P_o^θ and P_e^θ in the definition of Cohen's kappa, we obtain:

$$\kappa_\theta = \frac{2 \cdot M \cdot \mathrm{TP}_\theta - 2 \cdot \lfloor M \cdot \theta \rceil \cdot P}{P\left(M - \lfloor M \cdot \theta \rceil\right) + (M - P)\lfloor M \cdot \theta \rceil}.$$

To improve readability, we introduce the variables a_{κ_θ} and b_{κ_θ} defined as

$$\begin{aligned} a_{\kappa_\theta} &= \frac{2M}{P\left(M - \lfloor M \cdot \theta \rceil\right) + (M - P)\lfloor M \cdot \theta \rceil} \\ b_{\kappa_\theta} &= -\frac{2 \cdot \lfloor M \cdot \theta \rceil \cdot P}{P\left(M - \lfloor M \cdot \theta \rceil\right) + (M - P)\lfloor M \cdot \theta \rceil}. \end{aligned}$$

Hence, κ_θ is linear in TP_θ and can be written as

$$\kappa_\theta = X_\theta\left(a_{\kappa_\theta}, b_{\kappa_\theta}\right) \sim f_{X_\theta}\left(a_{\kappa_\theta}, b_{\kappa_\theta}\right),$$

with range:

$$\kappa_\theta \overset{(R)}{\in} \mathcal{R}\left(X_\theta\left(a_{\kappa_\theta}, b_{\kappa_\theta}\right)\right).$$

19.2. Expectation

As Cohen's kappa is linear in TP_θ, its expectation can be derived:

$$\begin{aligned} \mathbb{E}[\kappa_\theta] &= \mathbb{E}\left[X_\theta\left(a_{\kappa_\theta}, b_{\kappa_\theta}\right)\right] \overset{(1)}{=} a_{\kappa_\theta} \cdot \mathbb{E}[\mathrm{TP}_\theta] + b_{\kappa_\theta} \\ &= \frac{2 \cdot \lfloor M \cdot \theta \rceil \cdot P}{P\left(M - \lfloor M \cdot \theta \rceil\right) + (M - P)\lfloor M \cdot \theta \rceil} - \frac{2 \cdot \lfloor M \cdot \theta \rceil \cdot P}{P\left(M - \lfloor M \cdot \theta \rceil\right) + (M - P)\lfloor M \cdot \theta \rceil} \\ &= 0. \end{aligned}$$

19.3. Optimal baselines

The baselines are directly determined by the ranges of the expectation of κ_θ. Since the expectation is constant, its extreme values are the same:

$$\begin{cases} \min_{\theta\in[0,1]} (\mathbb{E}[\kappa_\theta]) = 0 & \text{if } P < M \\ \min_{\theta\in[0,1-1/(2M))} (\mathbb{E}[\kappa_\theta]) = 0 & \text{if } P = M, \end{cases}$$

$$\begin{cases} \max_{\theta\in[0,1]} (\mathbb{E}[\kappa_\theta]) = 0 & \text{if } P < M \\ \max_{\theta\in[0,1-1/(2M))} (\mathbb{E}[\kappa_\theta]) = 0 & \text{if } P = M. \end{cases}$$

This means that the optimization values θ_{min} and θ_{max} for these extreme values are simply all allowed values:

$$\begin{cases} \theta_{\text{min}} = \theta_{\text{max}} \in [0, 1] & \text{if } P < M \\ \theta_{\text{min}} = \theta_{\text{max}} \in \left[0, 1 - \frac{1}{2M}\right] & \text{if } P = M, \end{cases}$$

respectively. Consequently, the discrete versions θ^*_{min} and θ^*_{max} of the optimizers are given by

$$\begin{cases} \theta^*_{\text{min}} = \theta^*_{\text{max}} \in \Theta^* & \text{if } P < M \\ \theta^*_{\text{min}} = \theta^*_{\text{max}} \in \Theta^* \setminus \{1\} & \text{if } P = M. \end{cases}$$

20. Fowlkes-Mallows Index

The *Fowlkes-Mallows Index* FM_θ or *G-mean 1* was introduced by (Fowlkes and Mallows, 1983) as a way to calculate the similarity between two clusterings. It is the geometric average between the True Positive Rate (TPR_θ) and Positive Predictive Value (PPV_θ), which are discussed in Sec. 5 and 9, respectively. It balances correctly predicting the actual positive observations (TPR_θ) and being cautious in predicting observations as positive (PPV_θ).

20.1. Definition and distribution

The Fowlkes-Mallows Index is commonly defined as

$$\text{FM}_\theta = \sqrt{\text{TPR}_\theta \cdot \text{PPV}_\theta}.$$

By using the definitions of TPR_θ and PPV_θ in terms of TP_θ in Eq. (4) and (8), respectively, we obtain:

$$\text{FM}_\theta = \frac{\text{TP}_\theta}{\sqrt{P \cdot \lfloor M \cdot \theta \rceil}}.$$

Since TPR_θ is only defined when $P > 0$ and PPV_θ only when $\hat{P} := \lfloor M \cdot \theta \rceil > 0$, also FM_θ has these assumptions. Therefore, $\theta \geq \frac{1}{2M}$. The definition of FM_θ is linear in

TP_θ and can thus be formulated as

$$\mathrm{FM}_\theta = X_\theta\left(\frac{1}{\sqrt{P\cdot\lfloor M\cdot\theta\rceil}}, 0\right) \sim f_{X_\theta}\left(\frac{1}{\sqrt{P\cdot\lfloor M\cdot\theta\rceil}}, 0\right),$$

with range:

$$\mathrm{FM}_\theta \overset{(R)}{\in} \mathcal{R}\left(X_\theta\left(\frac{1}{\sqrt{P\cdot\lfloor M\cdot\theta\rceil}}, 0\right)\right).$$

194 **20.2. Expectation**

Because FM_θ is linear in TP_θ with slope $a = \frac{1}{\sqrt{P\cdot\lfloor M\cdot\theta\rceil}}$ and intercept $b = 0$, its expectation is

$$\mathbb{E}[\mathrm{FM}_\theta] = \mathbb{E}\left[X_\theta\left(\frac{1}{\sqrt{P\cdot\lfloor M\cdot\theta\rceil}}, 0\right)\right] \overset{(1)}{=} \frac{1}{\sqrt{P\cdot\lfloor M\cdot\theta\rceil}}\cdot\mathbb{E}[\mathrm{TP}_\theta] + 0 = \frac{\sqrt{P\cdot\lfloor M\cdot\theta\rceil}}{M}$$
$$= \sqrt{\frac{\theta^*\cdot P}{M}}.$$

195 **20.3. Optimal baselines**

The extreme values of the expectation of FM_θ determine the baselines. They are given by:

$$\min_{\theta\in[1/(2M),1]}(\mathbb{E}[\mathrm{FM}_\theta]) = \min_{\theta\in[1/(2M),1]}\left(\frac{\sqrt{P\cdot\lfloor M\cdot\theta\rceil}}{M}\right) = \frac{\sqrt{P}}{M},$$
$$\max_{\theta\in[1/(2M),1]}(\mathbb{E}[\mathrm{FM}_\theta]) = \max_{\theta\in[1/(2M),1]}\left(\frac{\sqrt{P\cdot\lfloor M\cdot\theta\rceil}}{M}\right) = \sqrt{\frac{P}{M}},$$

because the expectation is a non-decreasing function in θ. Note that the minimum and maximum are equal when $M = 1$. Consequently, the optimizers $\theta_{\min}$ and $\theta_{\max}$ for the extreme values are determined by:

$$\theta_{\min} \in \underset{\theta\in[1/(2M),1]}{\arg\min}(\mathbb{E}[\mathrm{FM}_\theta]) = \underset{\theta\in[1/(2M),1]}{\arg\min}\left(\frac{\sqrt{P\cdot\lfloor M\cdot\theta\rceil}}{M}\right) = \begin{cases}\left[\frac{1}{2M}, 1\right] & \text{if } M = 1\\ \left[\frac{1}{2M}, \frac{3}{2M}\right) & \text{if } M > 1,\end{cases}$$

$$\theta_{\max} \in \underset{\theta\in[1/(2M),1]}{\arg\max}(\mathbb{E}[\mathrm{FM}_\theta]) = \underset{\theta\in[1/(2M),1]}{\arg\max}\left(\frac{\sqrt{P\cdot\lfloor M\cdot\theta\rceil}}{M}\right) = \begin{cases}\left[\frac{1}{2M}, 1\right] & \text{if } M = 1\\ \left[1-\frac{1}{2M}, 1\right] & \text{if } M > 1,\end{cases}$$

respectively. The discrete forms $\theta^*_{\min}$ and $\theta^*_{\max}$ of these are given by:

$$\theta^*_{\min} \in \underset{\theta^* \in \Theta^* \setminus \{0\}}{\arg\min} \{\mathbb{E}[\mathrm{FM}_{\theta^*}]\} = \underset{\theta^* \in \Theta^* \setminus \{0\}}{\arg\min} \left\{ \sqrt{\frac{\theta^* \cdot P}{M}} \right\} = \left\{ \frac{1}{M} \right\},$$

$$\theta^*_{\max} \in \underset{\theta^* \in \Theta^* \setminus \{0\}}{\arg\max} \{\mathbb{E}[\mathrm{FM}_{\theta^*}]\} = \underset{\theta^* \in \Theta^* \setminus \{0\}}{\arg\max} \left\{ \sqrt{\frac{\theta^* \cdot P}{M}} \right\} = \{1\}.$$

21. G-mean 2

The *G-mean 2* $\mathrm{G}^{(2)}_\theta$ was established by (Kubat et al., 1998). This performance measure is the geometric average between the True Positive Rate (TPR_θ) and True Negative Rate (TNR_θ), which we discuss in Sec. 5 and 6, respectively. Hence, it balances correctly predicting the positive observations and correctly predicting the negative observations.

21.1. Definition and distribution

The G-mean 2 is defined as

$$\mathrm{G}^{(2)}_\theta = \sqrt{\mathrm{TPR}_\theta \cdot \mathrm{TNR}_\theta}.$$

Since TPR_θ needs the assumption $P > 0$ and TNR_θ needs $N := M - P > 0$, we have these restrictions also for $\mathrm{G}^{(2)}_\theta$. Consequently, $M > 1$. Now, by using the definitions of TPR_θ and TNR_θ in terms of TP_θ in, respectively, Eq. (4) and (5), we obtain:

$$\mathrm{G}^{(2)}_\theta = \sqrt{\frac{\mathrm{TP}_\theta \cdot (M - P - \lfloor M \cdot \theta \rceil) + \mathrm{TP}^2_\theta}{P\,(M-P)}}.$$

This function is not a linear function of TP_θ, and hence, we cannot write it in the form $X_\theta\,(a, b) = a \cdot \mathrm{TP}_\theta + b$ for some variables $a, b \in \mathbb{R}$.

21.2. Expectation

Since $\mathrm{G}^{(2)}_\theta$ is not linear in TP_θ, we cannot easily use the expectation of TP_θ to determine that for $\mathrm{G}^{(2)}_\theta$. However, we can determine the second moment of $\mathrm{G}^{(2)}_\theta$:

$$\begin{aligned}
\mathbb{E}\left[\left(\mathrm{G}^{(2)}_\theta\right)^2\right] &= \frac{M - P - \lfloor M \cdot \theta \rceil}{P\,(M-P)} \cdot \mathbb{E}[\mathrm{TP}_\theta] + \frac{1}{P\,(M-P)} \cdot \mathbb{E}[\mathrm{TP}^2_\theta] \\
&= \frac{M - P - \lfloor M \cdot \theta \rceil}{P\,(M-P)} \cdot \frac{\lfloor M \cdot \theta \rceil}{M} \cdot P + \frac{1}{P\,(M-P)} \cdot \left(\mathbf{Var}[\mathrm{TP}_\theta] + \mathbb{E}[\mathrm{TP}_\theta]^2\right) \\
&= \frac{(M - P - \lfloor M \cdot \theta \rceil) \cdot \lfloor M \cdot \theta \rceil}{M\,(M-P)} + \frac{\frac{\lfloor M\cdot\theta \rceil (M - \lfloor M\cdot\theta \rceil) P (M-P)}{M^2 (M-1)} + \left(\frac{\lfloor M\cdot\theta \rceil}{M} \cdot P\right)^2}{P\,(M-P)} \\
&= \frac{\lfloor M \cdot \theta \rceil \cdot (M - \lfloor M \cdot \theta \rceil)}{M(M-1)} = \theta^* \cdot \left(1 - \theta^*\right) \cdot \frac{M}{M-1}.
\end{aligned}$$

Of course, since the distribution of TP_θ is known, the expectation of $\mathrm{G}^{(2)}_\theta$ can always be numerically calculated.

21.3. Optimal baselines

We can show for the $G_\theta^{(2)}$ that the performance upper bound for the DD baseline is given by 0.5 by taking the inequality of the geometric mean and the arithmetic mean:

$$\mathbb{E}[G_\theta^{(2)}] = \mathbb{E}\left[\sqrt{\frac{TP_\theta}{P}\cdot\frac{TN_\theta}{N}}\right] \leq \frac{1}{2}\cdot\mathbb{E}\left[\frac{TP_\theta}{P}+\frac{TN_\theta}{N}\right]$$
$$= \frac{1}{2}\cdot(\mathbb{E}\left[\frac{TP_\theta}{P}\right]+\mathbb{E}\left[\frac{TN_\theta}{N}\right]) = \frac{1}{2}\cdot(\theta^*+(1-\theta^*)) = \frac{1}{2}$$

Another helpful lower bound on the performance of the DD can be derived when labeling randomly $M-P$ observations positive. If a θ is selected s.t. $\theta^* = \frac{M-P}{M}$, then:

$$\mathbb{E}[G^{(2)}_{\theta^*=\frac{M-P}{M}}] = \mathbb{E}\left[\sqrt{\frac{TP_\theta\cdot 0+TP_\theta^2}{P\,(M-P)}}\right] = \frac{1}{\sqrt{P\cdot(M-P)}}\cdot\mathbb{E}[TP_\theta] = \frac{\sqrt{P\cdot(M-P)}}{M}$$

It can be observed that when $P = M-P$, the lower and upper bounds are 0.5. This implies that the maximum expectation can be achieved by randomly predicting 50

Since the function $\varphi:\mathbb{R}\to\mathbb{R}_{\geq 0}$ given by $\varphi(x)=x^2$ is a convex function, we have by Jensen's inequality that:

$$\mathbb{E}[G_\theta^{(2)}]^2 \leq \mathbb{E}\left[\left(G_\theta^{(2)}\right)^2\right] = \theta^*\left(1-\theta^*\right)\frac{M}{M-1}.$$

This means that

$$\mathbb{E}[G_\theta^{(2)}] \leq \sqrt{\theta^*\,(1-\theta^*)\,\frac{M}{M-1}}.$$

Therefore, whenever $\theta^*\in\{0,1\}$, then $\mathbb{E}[G_\theta^{(2)}]\leq 0$. Since $G_\theta^{(2)}\geq 0$, it must hold that $\mathbb{E}[G_\theta^{(2)}]=0$. Hence, the set $\{0,1\}$ contains minimizers for $\mathbb{E}[G_\theta^{(2)}]$. The continuous version of this set is the interval $[0,1/(2M))\cup[1-1/(2M),1]$. To show that this interval contains the only possible values for the minimizers, consider the definition for the expectation of $G_\theta^{(2)}$:

$$\mathbb{E}\left[G_\theta^{(2)}\right] = \sum_{k\in\mathcal{D}(TP_\theta)}\sqrt{\frac{k\cdot((M-P)-(\lfloor M\cdot\theta\rceil-k))}{P\,(M-P)}}\cdot\mathbb{P}(TP_\theta=k),$$

where $\mathcal{D}(TP_\theta)$ is the domain of TP_θ, i.e. the set of values k such that $\mathbb{P}(TP_\theta=k)>0$. Now, let θ be such that $1/(2M)\leq\theta<1-1/(2M)$. Furthermore, consider the summand $S_k^{(\theta)}$ corresponding to $k=\min\{P,\lfloor M\cdot\theta\rceil\}\in\mathcal{D}(TP_\theta)$:

$$S^{(\theta)}_{k=\min\{P,\lfloor M\cdot\theta\rceil\}} = \begin{cases}\sqrt{\frac{M-\lfloor M\cdot\theta\rceil}{M-P}}\cdot\mathbb{P}(TP_\theta=P) & \text{if } P\leq\lfloor M\cdot\theta\rceil\\ \sqrt{\frac{\lfloor M\cdot\theta\rceil}{P}}\cdot\mathbb{P}(TP_\theta=\lfloor M\cdot\theta\rceil) & \text{if } P>\lfloor M\cdot\theta\rceil,\end{cases}$$

which is strictly positive in both cases. Hence, there is at least one term in the summation in the definition of $\mathbb{E}\left[\mathrm{G}_\theta^{(2)}\right]$ that is larger than 0; thus, the expectation is strictly positive for $1/(2M) \leq \theta < 1 - 1/(2M)$. Consequently, the minimization values $\theta_{\min} \in [0, 1]$ are:

$$\theta_{\min} \in \underset{\theta \in [0,1]}{\arg\min} \left(\mathbb{E}[\mathrm{G}_\theta^{(2)}]\right) = \left[0, \frac{1}{2M}\right) \cup \left[1 - \frac{1}{2M}, 1\right].$$

Following this reasoning, the discrete form $\theta^*_{\min} \in \Theta^*$ is given by:

$$\theta^*_{\min} \in \underset{\theta^* \in \Theta^*}{\arg\min} \left\{\mathbb{E}[\mathrm{G}_\theta^{(2)}]\right\} = \{0, 1\}.$$

22. Prevalence Threshold (PT)

A relatively new performance measure named *Prevalence Threshold* (PT_θ) was introduced by (Balayla, 2020). We could not find many articles that use this measure, but it is included for completeness. However, this performance measure has an inherent problem that eliminates the possibility of determining all statistics.

22.1. Definition and distribution

The *Prevalence Threshold* PT_θ is commonly defined as

$$\mathrm{PT}_\theta = \frac{\sqrt{\mathrm{TPR}_\theta \cdot \mathrm{FPR}_\theta} - \mathrm{FPR}_\theta}{\mathrm{TPR}_\theta - \mathrm{FPR}_\theta}.$$

By using the definitions of TPR_θ and FPR_θ in terms of TP_θ (see Equations (4) and (6)), we obtain:

$$\mathrm{PT}_\theta = \frac{\sqrt{P \cdot (M-P) \cdot \mathrm{TP}_\theta \cdot (\lfloor M \cdot \theta \rceil - \mathrm{TP}_\theta)} - P(\lfloor M \cdot \theta \rceil - \mathrm{TP}_\theta)}{M \cdot \mathrm{TP}_\theta - P \cdot \lfloor M \cdot \theta \rceil}. \tag{19}$$

It is clear that this performance measure is not a linear function of TP_θ, therefore we cannot easily calculate its expectation. However, there are more fundamental problems with PT_θ.

22.2. Division by zero

Eq. (19) shows that PT_θ is a problematic measure. When is the denominator zero? This happens when $\mathrm{TP}_\theta = (\lfloor M \cdot \theta \rceil / M) \cdot P$. In this case, the fraction is undefined, as the denominator is zero. Furthermore, also the numerator is zero in that case. The number of true positives TP_θ can attain the value $(\lfloor M \cdot \theta \rceil / M) \cdot P = \theta^* \cdot P$ whenever the latter is also an integer. For example, this always happens for $\theta^* \in \{0, 1\}$. But even when $\theta^* \in \Theta^* \setminus \{0, 1\}$, PT_θ is still only safe to use when M and P are coprime, i.e. when the only positive integer that is a divisor of both of them is 1. Otherwise, there are always values of $\theta^* \in \Theta^* \setminus \{0, 1\}$ that cause $\theta^* \cdot P$ to be an integer and therefore PT_θ to be undefined when TP_θ attains that value.

One solution would be to say $\mathrm{PT}_\theta := c$, $c \in [0,1]$, whenever both the numerator and denominator are zero. However, this c is arbitrary and directly influences the optimization of the expectation. This makes the optimal parameter values dependent on c, which is beyond the scope of this chapter. Thus, no statistics are derived for the Prevalence Threshold PT_θ.

23. Threat Score (TS) / Critical Success Index (CSI)

The *Threat Score* (Palmer and Allen, 1949) TS_θ or *Critical Success Index* (Schaefer, 1990) is a performance measure that is used for evaluation of forecasting binary weather events: it either happens in a specific location or it does not. It was already used in 1884 to evaluate the prediction of tornadoes (Schaefer, 1990). The Threat Score is the ratio of successful event forecasts (TP_θ) to the total number of positive predictions ($\mathrm{TP}_\theta + \mathrm{FP}_\theta$) and the number of events that were missed (FN_θ).

23.1. Definition and distribution

The Threat Score is thus defined as

$$\mathrm{TS}_\theta = \frac{\mathrm{TP}_\theta}{\mathrm{TP}_\theta + \mathrm{FP}_\theta + \mathrm{FN}_\theta}.$$

By using Eq. (B2) and (B3), this definition can be reformulated as

$$\mathrm{TS}_\theta = \frac{\mathrm{TP}_\theta}{P + \lfloor M \cdot \theta \rceil - \mathrm{TP}_\theta}.$$

Note that TS_θ is well-defined whenever $P > 0$. The definition of TS_θ is not linear in TP_θ, and so there are no $a, b \in \mathbb{R}$ such that we can write the definition as $X_\theta\,(a,b)$.

23.2. Expectation

Because TS_θ is not linear in TP_θ, determining the expectation is less straightforward than for other performance measures. The definition of the expectation is

$$\mathbb{E}[\mathrm{TS}_\theta] = \sum_{k \in \mathcal{D}(\mathrm{TP}_\theta)} \frac{k}{P + \lfloor M \cdot \theta \rceil - k} \cdot \mathbb{P}(\mathrm{TP}_\theta = k).$$

Unfortunately, we cannot explicitly solve this sum, but it can be calculated numerically.

23.3. Optimal baselines

Although no explicit formula can be given for the expectation, we can calculate its extreme values and the corresponding optimizers.

Minimal Baseline Firstly, we show that $\theta_{\min} \in [0, \frac{1}{2M})$ constitutes a minimum and that there are no θ outside this interval also yielding this minimum. To this end,

$$\mathbb{E}[\mathrm{TS}_{\theta_{\min}}] = \sum_{k \in \mathcal{D}(\mathrm{TS}_{\theta_{\min}})} \frac{k}{P + 0 - k} \cdot \mathbb{P}(\mathrm{TS}_{\theta_{\min}} = k) = 0,$$

because $\mathcal{D}(\mathrm{TS}_{\theta_{\min}}) = \{0\}$. This is the lowest possible value since TS_θ is a non-negative performance measure; hence, $\mathbb{E}[\mathrm{TS}_\theta] \geq 0$ for any $\theta \in [0, 1]$. Now, let $\theta' \geq \frac{1}{2M}$, then there exists a $k' > 0$ such that $\mathbb{P}(\mathrm{TP}_{\theta'} = k') > 0$. Consequently, $\mathbb{E}[\mathrm{TS}_{\theta'}] > 0$ and this means the interval $[0, \frac{1}{2M})$ contains the only values that constitute the minimum. In summary,

$$\min_{\theta \in [0,1]} (\mathbb{E}[\mathrm{TS}_\theta]) = 0,$$
$$\theta_{\min} \in \arg\min_{\theta \in [0,1]} (\mathbb{E}[\mathrm{TS}_\theta]) = \left[0, \frac{1}{2M}\right).$$

Since $\theta^*_{\min}$ is the discretization of $\theta_{\min}$, it corresponds to 0. More precisely:

$$\theta^*_{\min} \in \arg\min_{\theta^* \in \Theta^*} \{\mathbb{E}[\mathrm{TS}_{\theta^*}]\} = \{0\}.$$

Maximal baseline Secondly, to determine the maximum of $\mathbb{E}[\mathrm{TS}_\theta]$ and the corresponding parameter $\theta_{\max}$, we determine an upper bound for the expectation, show that this value is attained for a specific interval, and that there is no θ outside this interval also yielding this value. To do this, assume that $\lfloor M \cdot \theta \rceil > 0$. This makes sense, because $\lfloor M \cdot \theta \rceil = 0$ implies $\theta < 1/(2M)$ and such a θ would yield the minimum 0. Now,

$$\begin{aligned}
\mathbb{E}[\mathrm{TS}_\theta] &= \sum_{k \in \mathcal{D}(\mathrm{TP}_\theta)} \frac{k}{P + \lfloor M \cdot \theta \rceil - k} \cdot \mathbb{P}(\mathrm{TP}_\theta = k) \\
&\leq \sum_{k \in \mathcal{D}(\mathrm{TP}_\theta)} \frac{k}{P + \lfloor M \cdot \theta \rceil - P} \cdot \mathbb{P}(\mathrm{TP}_\theta = k) = \frac{1}{\lfloor M \cdot \theta \rceil} \sum_{k \in \mathcal{D}(\mathrm{TP}_\theta)} k \cdot \mathbb{P}(\mathrm{TP}_\theta = k) \\
&= \frac{\mathbb{E}[\mathrm{TP}_\theta]}{\lfloor M \cdot \theta \rceil} \overset{(1)}{=} \frac{P}{M}.
\end{aligned}$$

Next, let $\theta_{\max} \in [1 - 1/(2M), 1]$, then

$$\mathbb{E}[\mathrm{TS}_{\theta_{\max}}] = \sum_{k=M-(M-P)}^{P} \frac{k}{P + M - k} \cdot \mathbb{P}(\mathrm{TP}_{\theta_{\max}} = k) = \frac{P}{P + M - P} \cdot \mathbb{P}(\mathrm{TP}_{\theta_{\max}} = P) = \frac{P}{M},$$

because $\mathbb{P}(\mathrm{TP}_{\theta_{\max}} = P) = 1$. Hence, the upper bound is attained for $\theta_{\max} \in [1 - 1/(2M), 1]$, and thus, $\theta_{\max}$ is a maximizer.

Now, specifically for $P = 1$, we show that the interval of maximizers is actually

$[1/(2M), 1]$. Thus, let $\theta \in [1/(2M), 1 - 1/(2M))$, then $0 < \lfloor M \cdot \theta \rceil < M$ and

$$\begin{aligned}
\mathbb{E}[\mathrm{TS}_\theta] &= \sum_{k=\max\{0,\lfloor M\cdot\theta\rceil-(M-1)\}}^{\min\{1,\lfloor M\cdot\theta\rceil\}} \frac{k}{1+\lfloor M\cdot\theta\rceil - k} \cdot \mathbb{P}(\mathrm{TP}_\theta = k) \\
&= \frac{0}{1+\lfloor M\cdot\theta\rceil - 0} \cdot \mathbb{P}(\mathrm{TP}_\theta = 0) + \frac{1}{1+\lfloor M\cdot\theta\rceil - 1} \cdot \mathbb{P}(\mathrm{TP}_\theta = 1) \\
&= \frac{1}{\lfloor M\cdot\theta\rceil} \cdot \mathbb{P}(\mathrm{TP}_\theta = 1) = \frac{1}{\lfloor M\cdot\theta\rceil} \cdot \left(\frac{\binom{1}{1}\binom{M-1}{\lfloor M\cdot\theta\rceil - 1}}{\binom{M}{\lfloor M\cdot\theta\rceil}} \right) = \frac{1}{M},
\end{aligned}$$

which is exactly the upper bound $\mathbb{E}[\mathrm{TS}_{\theta_{\max}}] = P/M$ for $P = 1$.

Next, to show that the maximizers are only in $[1 - 1/(2M), 1]$ for $P > 1$, assume there is a $\theta' < 1 - \frac{1}{2M}$ that also yields the maximum. Hence, there is a $k' \in \mathcal{D}(\mathrm{TP}_{\theta'})$ with $0 < k' < P$ such that $\mathbb{P}(\mathrm{TP}_{\theta'} = k')$. This means

$$\begin{aligned}
\mathbb{E}[\mathrm{TS}_{\theta'}] &= \sum_{k\in\mathcal{D}(\mathrm{TP}_{\theta'})} \frac{k}{P+\lfloor M\cdot\theta'\rceil - k} \cdot \mathbb{P}(\mathrm{TP}_{\theta'} = k) \\
&= \frac{k'}{P+\lfloor M\cdot\theta'\rceil - k'} \cdot \mathbb{P}(\mathrm{TP}_{\theta'} = k') + \sum_{k\in\mathcal{D}(\mathrm{TP}_{\theta'})\setminus\{k'\}} \frac{k}{P+\lfloor M\cdot\theta'\rceil - k} \cdot \mathbb{P}(\mathrm{TP}_{\theta'} = k) \\
&\leq \frac{k'}{P+\lfloor M\cdot\theta'\rceil - (P-1)} \cdot \mathbb{P}(\mathrm{TP}_{\theta'} = k') + \\
&\quad \sum_{k\in\mathcal{D}(\mathrm{TP}_{\theta'})\setminus\{k'\}} \frac{k}{P+\lfloor M\cdot\theta'\rceil - P} \cdot \mathbb{P}(\mathrm{TP}_{\theta'} = k) \\
&= \frac{k'}{\lfloor M\cdot\theta'\rceil + 1} \mathbb{P}(\mathrm{TP}_{\theta'} = k') + \sum_{k\in\mathcal{D}(\mathrm{TP}_{\theta'})\setminus\{k'\}} \frac{k}{\lfloor M\cdot\theta'\rceil} \mathbb{P}(\mathrm{TP}_{\theta'} = k) \\
&< \frac{k'}{\lfloor M\cdot\theta'\rceil} \cdot \mathbb{P}(\mathrm{TP}_{\theta'} = k') + \sum_{k\in\mathcal{D}(\mathrm{TP}_{\theta'})\setminus\{k'\}} \frac{k}{\lfloor M\cdot\theta'\rceil} \cdot \mathbb{P}(\mathrm{TP}_{\theta'} = k) \\
&= \frac{1}{\lfloor M\cdot\theta'\rceil} \sum_{k\in\mathcal{D}(\mathrm{TP}_{\theta'})} k \cdot \mathbb{P}(\mathrm{TP}_{\theta'} = k) = \frac{P}{M}.
\end{aligned}$$

Hence, there is a strict inequality $\mathbb{E}[\mathrm{TS}_{\theta'}] < \frac{P}{M}$ and this means θ' is not a maximizer of the expectation. Consequently, the maximizers are only in the interval $[1 - 1/(2M), 1]$ for $P > 1$. In summary,

$$\begin{aligned}
\max_{\theta\in[0,1]} (\mathbb{E}[\mathrm{TS}_\theta]) &= \frac{P}{M}, \\
\theta_{\max} \in \arg\max_{\theta\in[0,1]} (\mathbb{E}[\mathrm{TS}_\theta]) &= \begin{cases} \left[\frac{1}{2M}, 1\right] & \text{if } P = 1 \\ \left[1 - \frac{1}{2M}, 1\right] & \text{if } P > 1. \end{cases}
\end{aligned}$$

Since $\theta^*_{\max}$ is the discretization of $\theta_{\max}$, we obtain:

$$\theta^*_{\max} \in \underset{\theta^* \in \Theta^*}{\arg\max} \{\mathbb{E}[\mathrm{TS}_{\theta^*}]\} = \begin{cases} \Theta^* \setminus \{0\} & \text{if } P = 1 \\ \{1\} & \text{if } P > 1. \end{cases}$$

References

Chinchor N. (1992) Muc-4 evaluation metrics. In: Proceedings of the 4th Conference on Message Understanding, Association for Computational Linguistics, USA, MUC4 ’92, p. 22–29, DOI https://doi.org/10.3115/1072064.1072067 16

Balayla J. (2020) Prevalence threshold (ϕe) and the geometry of screening curves. *PLoS ONE* 15(10):e0240215, DOI https://doi.org/10.1371/journal.pone.0240215 31

Fowlkes E. B., Mallows C. L. (1983) A method for comparing two hierarchical clusterings. *Journal of the American Statistical Association* 78(383):553–569, DOI https://doi.org/10.1080/01621459.1983.10478008 27

Kubat M., Holte R. C., Matwin S. (1998) Machine learning for the detection of oil spills in satellite radar images. *Machine Learning* 30(2/3):195–215, DOI https://doi.org/10.1023/a:1007452223027 29

Kvålseth T. O. (1989) Note on cohen’s kappa. *Psychological Reports* 65(1):223–226, DOI https://doi.org/10.2466/pr0.1989.65.1.223 25

Matthews B. (1975) Comparison of the predicted and observed secondary structure of t4 phage lysozyme. *Biochimica et Biophysica Acta (BBA) - Protein Structure* 405(2):442–451, DOI https://doi.org/10.1016/0005-2795(75)90109-9 23

Palmer W., Allen R. (1949) Note on the accuracy of forecasts concerning the rain problem. *US Weather Bureau manuscript* 32

Schaefer J. T. (1990) The critical success index as an indicator of warning skill. *Weather and Forecasting* 5(4):570–575, DOI https://doi.org/10.1175/1520-0434(1990)005<0570:tcsiaa>2.0.co;2 32

Yule G. U. (1912) On the methods of measuring association between two attributes. *Journal of the Royal Statistical Society* 75(6):579, DOI https://doi.org/10.2307/2340126 23

Youden W. J. (1950) Index for rating diagnostic tests. *Cancer* 3(1):32–35, DOI https://doi.org/10.1002/1097-0142(1950)3:1<32::aid-cncr2820030106>3.0.co;2-3 18