

A matriz de confusão – classificação binária

		Previsão	
		<i>N</i>	<i>P</i>
<u>População total</u> <i>Pop</i>			
Condição	<i>N</i>	<u>Verdadeiro negativo</u> <i>VN</i>	<u>Falso positivo</u> (erro tipo I) <i>FP</i>
	<i>P</i>	<u>Falso negativo</u> (erro tipo II) <i>FN</i>	<u>Verdadeiro positivo</u> <i>VP</i>

Obs.: Aqui a matriz de confusão está conforme é retornada pelo `sklearn`, com a condição nas linhas e a previsão nas colunas, e classe negativa antes da classe positiva. Fique alerta pois nem todo software ou referência a ilustra da mesma maneira!

Algumas métricas derivadas da matriz de confusão

		Previsão			
	<u>População total</u> <i>Pop</i>	<i>N</i>	<i>P</i>		
Condição	<i>N</i>	<u>Verdadeiro negativo</u> <i>VN</i>	<u>Falso positivo</u> (erro tipo I) <i>FP</i>	<u>Taxa de falsos positivos</u> (<i>FPR</i>), probabilidade de alarme falso $\frac{FP}{N}$	<u>Especificidade</u> , Seletividade, <u>Taxa de verdadeiros negativos</u> (<i>TNR</i>) $\frac{VN}{N}$
	<i>P</i>	<u>Falso negativo</u> (erro tipo II) <i>FN</i>	<u>Verdadeiro positivo</u> <i>VP</i>	<u>Taxa de verdadeiros positivos</u> (<i>TPR</i>), <u>Recall</u> (Revocação), <u>Sensibilidade</u> , Probabilidade de detecção, <u>Poder</u> $\frac{VP}{P}$	<u>Taxa de falsos negativos</u> (<i>FNR</i>), Taxa de erro $\frac{FN}{P}$
	<u>Prevalência</u> $\frac{P}{Pop}$	<u>False omission rate</u> (<i>FOR</i>) $\frac{FN}{Prev.N} = \frac{FN}{VN + FN}$	<u>Valor preditivo positivo</u> , <i>Positive predictive value</i> (<i>PPV</i>), <i>Precision</i> , Precisão $\frac{VP}{Prev.P} = \frac{VP}{FP + VP}$	<u>Positive likelihood ratio</u> (<i>LR +</i>) $\frac{TPR}{FPR}$	<u>Negative likelihood ratio</u> (<i>LR -</i>) $\frac{FNR}{TNR}$
				<u>Diagnostic odds ratio</u> (<i>DOR</i>) $\frac{LR +}{LR -}$	
	<u>Acurácia</u> (<i>ACC</i>), Risco $\frac{VN + VP}{Pop}$	<u>Valor preditivo negativo</u> , <i>Negative predictive value</i> (<i>NPV</i>) $\frac{VN}{Prev.N} = \frac{VN}{VN + FN}$	<u>False discovery rate</u> (<i>FDR</i>) $\frac{FP}{Prev.P} = \frac{FP}{FP + VP}$	<u>F1 score</u> $\frac{\frac{2}{\frac{1}{PPV} + \frac{1}{TPR}}}{\frac{1}{Precisão} + \frac{1}{Revocação}} = \frac{2 \cdot PPV \cdot TPR}{PPV + TPR} = \frac{2 \cdot Precisão \cdot Revocação}{Precisão + Revocação}$	

Algumas métricas derivadas da matriz de confusão – Exemplo 1

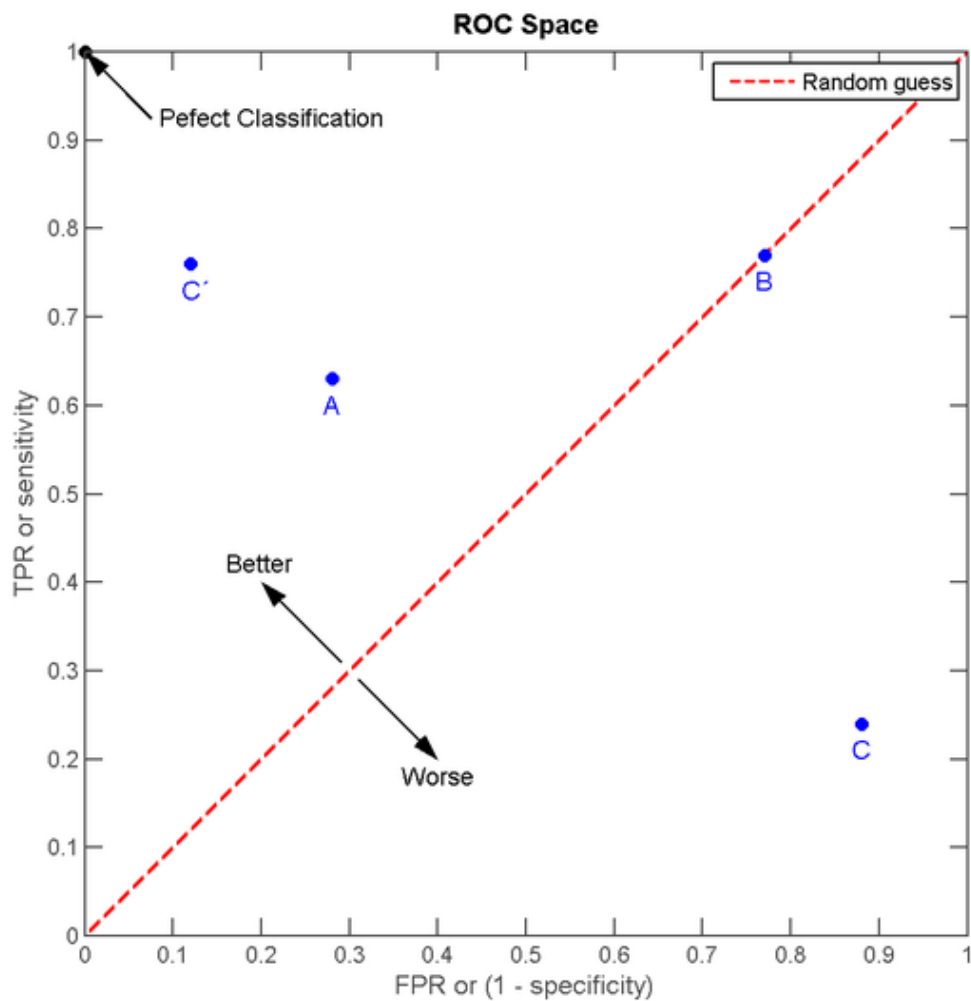
		Previsão			
		<u>População total</u> <i>Pop</i> = 100	<i>N</i> = 90	<i>P</i> = 10	
Condição	<i>N</i> = 90	<u>Verdadeiro negativo</u> <i>VN</i> = 85	<u>Falso positivo</u> (<u>erro tipo I</u>) <i>FP</i> = 5	<u>Taxa de falsos positivos</u> (<i>FPR</i>), probabilidade de alarme falso $\frac{FP}{N} = 0,055 \dots$	<u>Especificidade</u> , Seletividade, <u>Taxa de verdadeiros negativos</u> (<i>TNR</i>) $\frac{VN}{N} = 0,944 \dots$
	<i>P</i> = 10	<u>Falso negativo</u> (<u>erro tipo II</u>) <i>FN</i> = 5	<u>Verdadeiro positivo</u> <i>VP</i> = 5	<u>Taxa de verdadeiros positivos</u> (<i>TPR</i>), <u>Recall</u> (Revocação), <u>Sensibilidade</u> , Probabilidade de detecção, <u>Poder</u> $\frac{VP}{P} = 0,5$	<u>Taxa de falsos negativos</u> (<i>FNR</i>), Taxa de erro $\frac{FN}{P} = 0,5$
	<u>Prevalência</u> $\frac{P}{Pop} = 0,1$	<u>False omission rate</u> (<i>FOR</i>) $\frac{FN}{Prev.N} = \frac{FN}{VN + FN} = 0,055 \dots$	<u>Valor preditivo positivo</u> , <i>Positive predictive value</i> (<i>PPV</i>), <i>Precision</i> , Precisão $\frac{VP}{Prev.P} = \frac{VP}{FP + VP} = 0,5$	<u>Positive likelihood ratio</u> (<i>LR +</i>) $\frac{TPR}{FPR} = 9$	<u>Negative likelihood ratio</u> (<i>LR -</i>) $\frac{FNR}{TNR} = 0,529 \dots$
	<u>Acurácia</u> (<i>ACC</i>), Risco $\frac{VN + VP}{Pop} = 0,9$	<u>Valor preditivo negativo</u> , <i>Negative predictive value</i> (<i>NPV</i>) $\frac{VN}{Prev.N} = \frac{VN}{VN + FN} = 0,944 \dots$	<u>False discovery rate</u> (<i>FDR</i>) $\frac{FP}{Prev.P} = \frac{FP}{FP + VP} = 0,5$	<u>Diagnostic odds ratio</u> (<i>DOR</i>) $\frac{LR +}{LR -} = 17$	
				<u>F1 score</u> $\frac{\frac{2}{\frac{1}{PPV} + \frac{1}{TPR}}}{\frac{1}{Precisão} + \frac{1}{Revocação}} = 0,5$ $\frac{2 \cdot PPV \cdot TPR}{PPV + TPR} = \frac{2 \cdot Precisão \cdot Revocação}{Precisão + Revocação}$	

Algumas métricas derivadas da matriz de confusão – Exemplo 1

		Previsão			
		<u>População total</u> <i>Pop</i> = 100	<i>N</i> = 83	<i>P</i> = 17	
Condição	<i>N</i> = 90	<u>Verdadeiro negativo</u> <i>VN</i> = 82	<u>Falso positivo</u> (<u>erro tipo I</u>) <i>FP</i> = 8	<u>Taxa de falsos positivos</u> (<i>FPR</i>), probabilidade de alarme falso $\frac{FP}{N} = 0,088 \dots$	<u>Especificidade</u> , Seletividade, <u>Taxa de verdadeiros negativos</u> (<i>TNR</i>) $\frac{VN}{N} = 0,911 \dots$
	<i>P</i> = 10	<u>Falso negativo</u> (<u>erro tipo II</u>) <i>FN</i> = 1	<u>Verdadeiro positivo</u> <i>VP</i> = 9	<u>Taxa de verdadeiros positivos</u> (<i>TPR</i>), <u>Recall</u> (Revocação), <u>Sensibilidade</u> , Probabilidade de detecção, <u>Poder</u> $\frac{VP}{P} = 0,9$	<u>Taxa de falsos negativos</u> (<i>FNR</i>), Taxa de erro $\frac{FN}{P} = 0,1$
	<u>Prevalência</u> $\frac{P}{Pop} = 0,1$	<u>False omission rate</u> (<i>FOR</i>) $\frac{FN}{Prev.N} = \frac{FN}{VN + FN} = 0,012 \dots$	<u>Valor preditivo positivo</u> , <i>Positive predictive value</i> (<i>PPV</i>), <i>Precision</i> , Precisão $\frac{VP}{Prev.P} = \frac{VP}{FP + VP} = 0,529 \dots$	<u>Positive likelihood ratio</u> (<i>LR +</i>) $\frac{TPR}{FPR} = 10,125$	<u>Negative likelihood ratio</u> (<i>LR -</i>) $\frac{FNR}{TNR} = 0,109 \dots$
				<u>Diagnostic odds ratio</u> (<i>DOR</i>) $\frac{LR +}{LR -} = 92,25$	
	<u>Acurácia</u> (<i>ACC</i>), Risco $\frac{VN + VP}{Pop} = 0,91$	<u>Valor preditivo negativo</u> , <i>Negative predictive value</i> (<i>NPV</i>) $\frac{VN}{Prev.N} = \frac{VN}{VN + FN} = 0,987 \dots$	<u>False discovery rate</u> (<i>FDR</i>) $\frac{FP}{Prev.P} = \frac{FP}{FP + VP} = 0,470 \dots$	<u>F1 score</u> $\frac{2}{\frac{1}{PPV} + \frac{1}{TPR}} = \frac{2}{\frac{1}{Precisão} + \frac{1}{Revocação}} = 0,666 \dots$ $\frac{2 \cdot PPV \cdot TPR}{PPV + TPR} = \frac{2 \cdot Precisão \cdot Revocação}{Precisão + Revocação}$	

Algumas métricas derivadas da matriz de confusão – Exemplo 1

		Previsão			
		<u>População total</u> <i>Pop</i> = 100	<i>N</i> = 90	<i>P</i> = 10	
Condição	<i>N</i> = 90	<u>Verdadeiro negativo</u> <i>VN</i> = 90	<u>Falso positivo</u> (<u>erro tipo I</u>) <i>FP</i> = 0	<u>Taxa de falsos positivos</u> (<i>FPR</i>), probabilidade de alarme falso $\frac{FP}{N} = 0$	<u>Especificidade</u> , Seletividade, <u>Taxa de verdadeiros negativos</u> (<i>TNR</i>) $\frac{VN}{N} = 1$
	<i>P</i> = 10	<u>Falso negativo</u> (<u>erro tipo II</u>) <i>FN</i> = 0	<u>Verdadeiro positivo</u> <i>VP</i> = 10	<u>Taxa de verdadeiros positivos</u> (<i>TPR</i>), <u>Recall</u> (Revocação), <u>Sensibilidade</u> , Probabilidade de detecção, <u>Poder</u> $\frac{VP}{P} = 1$	<u>Taxa de falsos negativos</u> (<i>FNR</i>), Taxa de erro $\frac{FN}{P} = 0$
				<u>Positive likelihood ratio</u> (<i>LR +</i>) $\frac{TPR}{FPR} = \infty$	<u>Negative likelihood ratio</u> (<i>LR -</i>) $\frac{FNR}{TNR} = 0$
				<u>Diagnostic odds ratio</u> (<i>DOR</i>) $\frac{LR +}{LR -} = \infty$	
				<u>F1 score</u> $\frac{\frac{2}{\frac{1}{PPV} + \frac{1}{TPR}}}{\frac{2}{\frac{1}{Precisão} + \frac{1}{Revocação}}} = 1$ $\frac{2 \cdot PPV \cdot TPR}{PPV + TPR} = \frac{2 \cdot Precisão \cdot Revocação}{Precisão + Revocação}$	
		<u>Prevalência</u> $\frac{P}{Pop} = 0,1$	<u>False omission rate</u> (<i>FOR</i>) $\frac{FN}{Prev.N} = \frac{FN}{VN + FN} = 0$	<u>Valor preditivo positivo</u> , <i>Positive predictive value</i> (<i>PPV</i>), <i>Precision</i> , <i>Precisão</i> $\frac{VP}{Prev.P} = \frac{VP}{FP + VP} = 1$	
		<u>Acurácia</u> (<i>ACC</i>), Risco $\frac{VN + VP}{Pop} = 1$	<u>Valor preditivo negativo</u> , <i>Negative predictive value</i> (<i>NPV</i>) $\frac{VN}{Prev.N} = \frac{VN}{VN + FN} = 1$	<u>False discovery rate</u> (<i>FDR</i>) $\frac{FP}{Prev.P} = \frac{FP}{FP + VP} = 0$	



[Wikipedia - https://en.wikipedia.org/wiki/Receiver_operating_characteristic]

A		Previsão	
	<u>População total</u> <i>Pop</i> = 200	<i>N</i> = 109	<i>P</i> = 91
Condição	<i>N</i> = 100	<u>Verdadeiro negativo</u> <i>VN</i> = 72	<u>Falso positivo (erro tipo I)</u> <i>FP</i> = 28
	<i>P</i> = 100	<u>Falso negativo (erro tipo II)</u> <i>FN</i> = 37	<u>Verdadeiro positivo</u> <i>VP</i> = 63
<i>TPR</i> = 0,63		<i>FPR</i> = 0,28	
<i>F1</i> = 0,66		<i>ACC</i> = 0,68	

B		Previsão	
	<u>População total</u> <i>Pop</i> = 200	<i>N</i> = 46	<i>P</i> = 154
Condição	<i>N</i> = 100	<u>Verdadeiro negativo</u> <i>VN</i> = 23	<u>Falso positivo (erro tipo I)</u> <i>FP</i> = 77
	<i>P</i> = 100	<u>Falso negativo (erro tipo II)</u> <i>FN</i> = 23	<u>Verdadeiro positivo</u> <i>VP</i> = 77
<i>TPR</i> = 0,77		<i>FPR</i> = 0,77	
<i>F1</i> = 0,61		<i>ACC</i> = 0,50	

C		Previsão	
	<u>População total</u> <i>Pop</i> = 200	<i>N</i> = 88	<i>P</i> = 112
Condição	<i>N</i> = 100	<u>Verdadeiro negativo</u> <i>VN</i> = 12	<u>Falso positivo (erro tipo I)</u> <i>FP</i> = 88
	<i>P</i> = 100	<u>Falso negativo (erro tipo II)</u> <i>FN</i> = 76	<u>Verdadeiro positivo</u> <i>VP</i> = 24
<i>TPR</i> = 0,24		<i>FPR</i> = 0,88	
<i>F1</i> = 0,23		<i>ACC</i> = 0,18	

C'		Previsão	
	<u>População total</u> <i>Pop</i> = 200	<i>N</i> = 112	<i>P</i> = 88
Condição	<i>N</i> = 100	<u>Verdadeiro negativo</u> <i>VN</i> = 88	<u>Falso positivo (erro tipo I)</u> <i>FP</i> = 12
	<i>P</i> = 100	<u>Falso negativo (erro tipo II)</u> <i>FN</i> = 24	<u>Verdadeiro positivo</u> <i>VP</i> = 76
<i>TPR</i> = 0,76		<i>FPR</i> = 0,12	
<i>F1</i> = 0,81		<i>ACC</i> = 0,82	