

Avaliação de Classificação

Prof. Dr. Rogério Galante Negri

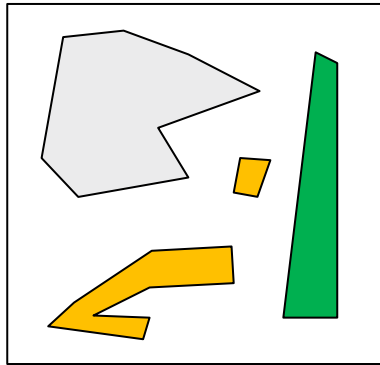
Avaliação de Classificação

- Verificar a precisão de um resultado de classificação
 - As classes do resultado precisam ser conhecidas
 - São necessários dados de referência (amostras de teste)
 - Dados pré-existentes (levantamentos, mapas, literatura, etc)
 - Baixo custo, mas as informações podem estar desatualizadas
 - Erros pré-existentes raramente conhecidos
 - Pode haver incompatibilidade entre classes (diferença semântica)
 - Dados de campo
 - Custo elevado (logística, localização precisa, etc)
 - Defasagem temporal entre obtenção do dado usado na classificação e a checagem
 - A amostragem é enviesada (pontos que se tem acesso)
 - Fotointerpretação (com imagens de maiores resoluções)
 - Baixo custo, mas depende a experiência do fotointérprete
 - Cuidado para não induzir um resultado positivo falso
 - Resultado da classificação: imagem classificada ou mapa temático
-

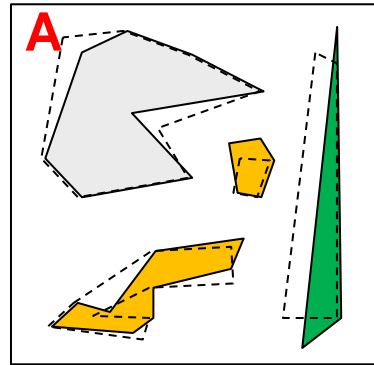
Unidade Amostral

- Comparação entre um mapa temático e uma referência pode ser feita:
 - Pontos simples (ou pixels numa imagem)
 - Grupos de pontos ou pixels
 - Polígonos (ou objetos)
 - Grupos de polígonos
-

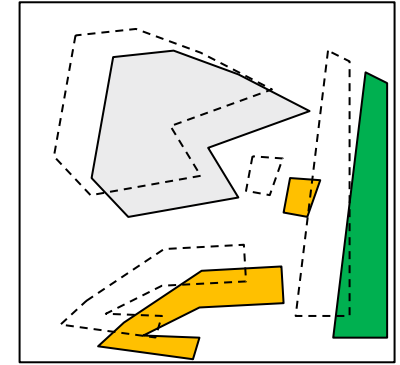
Exatidão Temática



mapa de referência
(verdade)



problemas de
deslocamento
e diferenças
de resolução

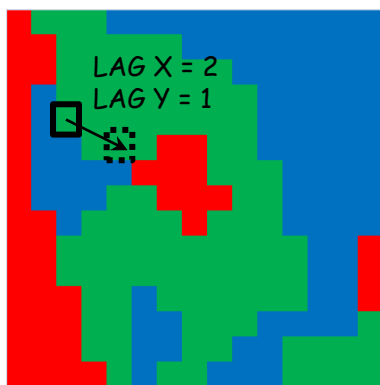


erros de
inclusão e
exclusão se
compensam

- Qual dos mapas A ou B é mais parecido com o mapa de referência?
- Com relação ao posicionamento → **A**
- Com relação à área total de cada classe → **B**

Dependência Espacial

- A utilização de pontos amostrais próximos cria a falsa impressão de que o conjunto amostral é grande quando, na realidade, há muita redundância de informação
- Prejudica principalmente a avaliação da incerteza (variância) das estimativas obtidas a partir da amostra
- Para minimizar os problemas, determina-se qual a distância (Lag em x e em y) a partir do qual os pontos escolhidos podem ser considerados independentes



Teste de Independência (χ^2)

Simulação de 60 pares de pontos aleatórios (10000 iterações)

Porcentagem das simulações cuja independência foi comprovada

		LAG X				
		0	1	2	3	4
LAG Y	0	---	0,0%	38,2%	76,4%	78,0%
	1	0,0%	0,0%	47,5%	74,2%	55,3%
	2	0,0%	13,5%	62,9%	72,0%	80,5%
	3	4,4%	39,7%	67,0%	49,7%	63,6%
	4	30,2%	55,5%	60,4%	56,8%	67,9%

Metodologia de Avaliação

- A partir dos pares de pontos (polígonos ou grupos) avaliados, é construída a Matriz de Confusão
- Antes de mais nada:
 - As classes são excludentes (cada ponto em apenas uma classe)
 - Todos os pontos avaliados devem pertencer a alguma classe (não pode considerar a classe “não classificado”)
 - O uso de classes muito semelhantes (e.g., subtipos) pode induzir a um excesso de erros (ou confusões) que certamente prejudicarão a avaliação global da classificação
 - A interpretação dos resultados depende diretamente da unidade amostral adotada

		Referência				Total
		1	2	...	c	
Classificação	1	x_{11}	x_{12}	...	x_{1c}	x_{1+}
	2	x_{21}	x_{22}	...	x_{2c}	x_{2+}
	⋮	⋮	⋮	⋮	⋮	⋮
	c	x_{c1}	x_{c2}	...	x_{cc}	x_{c+}
Total	x_{+1}	x_{+2}	...	x_{+c}	n	

Matriz de Confusão

		Referência				Total
		1	2	...	c	
Classificação	1	x_{11}	x_{12}	...	x_{1c}	x_{1+}
	2	x_{21}	x_{22}	...	x_{2c}	x_{2+}
	⋮	⋮	⋮	⋮	⋮	⋮
	c	x_{c1}	x_{c2}	...	x_{cc}	x_{c+}
	Total	x_{+1}	x_{+2}	...	x_{+c}	n

x_{ij} – número de pontos da classe i (referência), classificados na classe j (classificação)

x_{kk} – número total de pontos corretamente classificados da classe k

x_{+j} – número total de pontos avaliados da classe j na referência

x_{i+} – número total de pontos avaliados da classe i na classificação

Matriz de Confusão

		Referência				Total
		1	2	3	4	
Classificação	1	13	0	24	6	43
	2	8	10	5	6	29
	3	8	4	27	4	43
	4	0	0	1	34	35
	Total	29	14	57	50	150

$$\text{Exatidão Total (ou Global)} = \frac{\sum_{k=1}^c x_{kk}}{n} \quad \left\{ \begin{array}{l} \text{mínimo} = 0 \\ \text{máximo} = 1 \text{ (ou 100\%)} \end{array} \right.$$

$$= \frac{13 + 10 + 27 + 34}{150} = \frac{84}{150} = 0,56$$

Matriz de Confusão

Ponto de vista do
“Produtor”

Quanto da classe k
foi “vista” pelo
classificador?

		Referência				Total
		1	2	3	4	
Classificação	1	13	0	24	6	43
	2	8	10	5	6	29
	3	8	4	27	4	43
	4	0	0	1	34	35
	Total	29	14	57	50	150

$$\text{Exatidão do Produtor da classe } k = \frac{x_{kk}}{x_{+k}}$$

$$\text{Exatidão do Produtor da classe } 2 = \frac{10}{14} = 0,7143$$

$$\text{Erro de omissão da classe } 2 = \frac{0 + 4 + 0}{14} = \frac{4}{14} = 0,2857$$

Matriz de Confusão

Ponto de vista do Consumidor (Usuário)

Quanto do que foi classificado como k é realmente da classe k ?

		Referência				Total
		1	2	3	4	
Classificação	1	13	0	24	6	43
	2	8	10	5	6	29
	3	8	4	27	4	43
	4	0	0	1	34	35
	Total	29	14	57	50	150

$$\text{Exatidão do Consumidor da classe } k = \frac{x_{kk}}{x_{k+}}$$

$$\text{Exatidão do Consumidor da classe } 2 = \frac{10}{29} = 0,3448$$

$$\text{Erro de comissão (inclusão) da classe } 2 = \frac{8 + 5 + 6}{29} = \frac{19}{29} = 0,6552$$

Avaliação da Exatidão

		Referência				Total
		1	2	3	4	
Classificação	1	13	0	24	6	43
	2	8	10	5	6	29
	3	8	4	27	4	43
	4	0	0	1	34	35
Total		29	14	57	50	150

Exatidão Total = 0,56

E se a classificação fosse realizada de modo totalmente aleatória?

		Referência				Total
		1	2	3	4	
Classificação	1	8,31	4,01	16,34	14,33	43
	2	5,61	2,71	11,02	9,67	29
	3	8,31	4,01	16,34	14,33	43
	4	6,77	3,27	13,30	11,67	35
Total		29	14	57	50	150

$$\text{Exatidão Total} = \frac{8,31 + 2,71 + 16,34 + 11,67}{150} = \frac{39,03}{150} = 0,26$$

Apenas 26%... Será que este acerto foi conseguido de modo casual?

Medida de Concordância Kappa

		Referência				Total
		1	2	3	4	
Classificação	1	13	0	24	6	43
	2	8	10	5	6	29
	3	8	4	27	4	43
	4	0	0	1	34	35
	Total	29	14	57	50	150

Kappa (\mathbf{K}) – medida de concordância $\hat{\kappa} = \frac{\theta_1 - \theta_2}{1 - \theta_2}$

$$\theta_1 = \frac{\sum_{k=1}^c x_{kk}}{n} \quad \text{exatidão total (observada)}$$

$$\theta_2 = \frac{\sum_{k=1}^c x_{k+} x_{+k}}{n^2} \quad \begin{array}{l} \text{exatidão total} \\ \text{(se classificação fosse aleatória)} \end{array}$$

$$\mathbf{K} \begin{cases} \text{mínimo} = < 0 & (\theta_1 < \theta_2) \\ \text{máximo} = 1 & (\theta_1 = 1) \end{cases}$$

Kappa

		Referência				Total
		1	2	3	4	
Classificação	1	13	0	24	6	43
	2	8	10	5	6	29
	3	8	4	27	4	43
	4	0	0	1	34	35
	Total	29	14	57	50	150

$$\theta_1 = \sum_{k=1}^c x_{kk}$$

$$\hat{\kappa} = \frac{\theta_1 - \theta_2}{1 - \theta_2} = \frac{0,56 - 0,26}{1 - 0,26} = 0,4053$$

$$\theta_2 = \sum_{k=1}^c x_{k+} x_{+k}$$

$$\hat{\kappa} = \frac{\theta_1 - \theta_2}{1 - \theta_2}$$

Se a classificação fosse totalmente aleatória, o valor esperado para o Kappa seria ZERO

Será que este valor é significativamente superior a ZERO? [Teste de Hipóteses!]

Variância do Kappa

$$\hat{\kappa} = \frac{\theta_1 - \theta_2}{1 - \theta_2} \quad \theta_1 = \sum_{k=1}^c x_{kk} \quad \theta_2 = \sum_{k=1}^c x_{k+} x_{+k}$$

$$Var(\hat{\kappa}) = \frac{1}{n} \left[\frac{\theta_1(1 - \theta_1)}{(1 - \theta_2)^2} + \frac{2(1 - \theta_1)(2\theta_1\theta_2 - \theta_3)}{(1 - \theta_2)^3} + \frac{(1 - \theta_1)^2(\theta_4 - 4\theta_2^2)}{(1 - \theta_2)^4} \right]$$

$$\theta_3 = \sum_{k=1}^c x_{kk} (x_{k+} + x_{+k}) / n^2 \quad \theta_4 = \sum_{i=1}^c \sum_{j=1}^c x_{ij} (x_{i+} + x_{+j})^2 / n^3$$

Pressupondo amostras independentes e TLC válido
(amostra grande)

$$Z \sim N(0,1)$$

Índice Kappa – T.H.

classificação 1

		referência				
		A	B	C	D	E
A	13	0	3	0	0	
B	8	10	5	0	0	
C	8	4	27	0	0	
D	2	0	1	25	0	
E	0	0	0	0	44	

$$\hat{\kappa}_1 = 0,7364$$

$$\text{Var}(\hat{\kappa}_1) = 0,001664$$

classificação 2

		A	B	C	D	E
A	25	0	2	0	0	
B	2	28	0	0	4	
C	0	2	30	2	0	
D	0	0	1	33	0	
E	0	0	0	0	21	

$$\hat{\kappa}_2 = 0,8911$$

$$\text{Var}(\hat{\kappa}_2) = 0,000831$$

Perceba que foram usadas amostras diferentes para avaliar cada classificação

$$H_0: \kappa_1 - \kappa_2 = 0$$

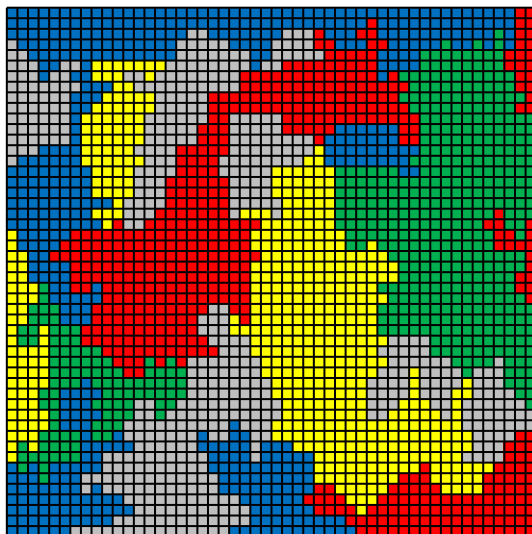
$$H_1: \kappa_1 - \kappa_2 < 0$$

$$Z = \frac{\kappa_1 - \kappa_2}{\sqrt{\text{Var}(\kappa_1) + \text{Var}(\kappa_2)}} \sim N(0,1)$$

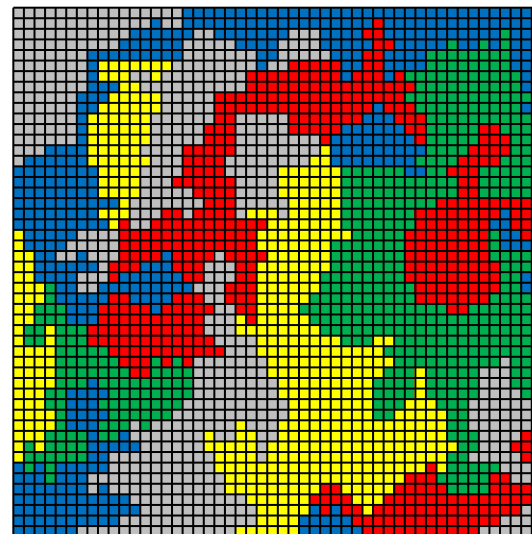
$$z = \frac{\hat{\kappa}_1 - \hat{\kappa}_2}{\sqrt{\text{Var}(\hat{\kappa}_1) + \text{Var}(\hat{\kappa}_2)}} = \frac{0,7364 - 0,8911}{\sqrt{0,001664 + 0,000831}} = -3,10 \quad \text{valor-P} = 0,0010$$

Conclusão: rejeita-se H_0 com 5% de significância, ou seja, a concordância entre a classificação e a referência é maior para o classificador 2

Kappa



referência



classificação

$$H_0: \kappa = 0,7$$

$$H_1: \kappa > 0,7$$

$$\hat{\kappa} = 0,7400$$

$$\text{Var}(\hat{\kappa}) = 0,000103$$

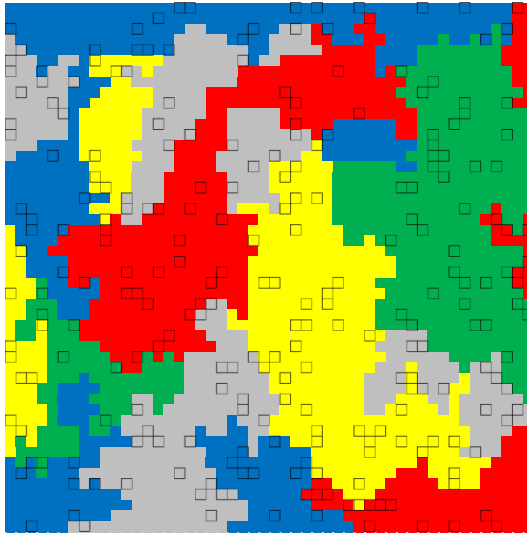
$$z = \frac{\hat{\kappa} - 0,7}{\sqrt{\text{Var}(\hat{\kappa})}} = 3,9475$$

$$\text{Valor-P} = 3,95 \cdot 10^{-5}$$

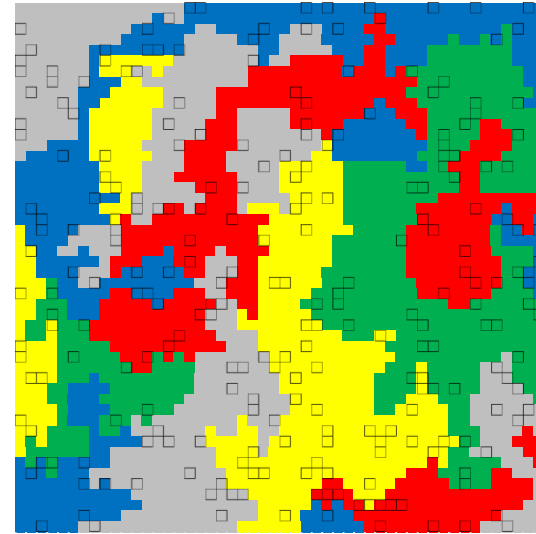
		REFERÊNCIA					$n_{i\cdot}$
		Red	Yellow	Green	Blue	Grey	
CLASSIF.	Red	360	0	110	0	0	470
	Yellow	0	410	0	70	0	480
	Green	0	90	390	0	50	530
	Blue	80	0	0	370	0	450
	Grey	60	0	0	60	450	570
$n_{\cdot j}$		500	500	500	500	500	2500

Conclusão: rejeita-se H_0 com 5% de significância, ou seja, o kappa dessa classificação é maior que 0,7 (esta conclusão pode estar equivocada pois desconsiderou a autocorrelação espacial)

Kappa



referência



classificação

$$H_0: \kappa = 0,7$$

$$H_1: \kappa > 0,7$$

$$\hat{\kappa} = 0,7336$$

$$\text{Var}(\hat{\kappa}) = 0,00103$$

$$z = \frac{\hat{\kappa} - 0,7}{\sqrt{\text{Var}(\hat{\kappa})}} = 1,0447$$

$$\text{Valor-P} = 0,1481$$

		REFERÊNCIA					$n_{i\cdot}$
		Red	Yellow	Green	Blue	Grey	
CLASSIF.	Red	26	0	8	0	0	34
	Yellow	0	50	0	8	0	58
	Green	0	8	35	0	6	49
	Blue	14	0	0	33	0	47
	Grey	6	0	0	3	53	62
$n_{\cdot j}$		46	58	43	44	59	250

Conclusão: aceita-se H_0

Bibliografia da aula

- Camilo Daleles Rennó. **Notas de Aula – Estatística Aplicada ao Sensoriamento Remoto**, INPE, São José dos Campos, 2012. Disponível em:
www.dpi.inpe.br/~camilo/estatistica2012/ppt/aula17.ppt

Estatística: Aplicação ao Sensoriamento Remoto

SER 202

Aula 17 - ANO 2012

Camilo Daleles Rennó

<http://www.dpi.inpe.br/~camilo/estatistica/>