Como tomar decisões?

Dani Gamerman
Instituto de Matemática – UFRJ

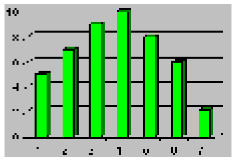
http://www.dme.ufrj.br/~dani

Estatística: o que é ?

Estatística pode ser pensada como a ciência de aprendizagem a partir de dados.

No nosso cotidiano, precisamos tomar decisões, muitas vezes decisões rápidas.

INFORMAÇÃO

()

→ ANÁLISE → DECISÕES

Em linhas gerais, a Estatística fornece métodos que auxiliam o processo de tomada de decisão.

Exemplo:
Problema de decisão médica

Paciente chega ao consultório com uma queixa.
Médico suspeita que ele tem doença D.

Suspeita é quantificada através de probabilidade
Prob (D+) = 0,6, especificada subjetivamente.

Doença (D+) → Tratar (T+)

Tratamento é doloroso e caro.

Sem doença (D-) → Não tratar (T-)

Que decisão tomar?

Tratar (T+) ou mandar embora (T-)

Outro exemplo:

Na área financeira, temos que tomar a decisão de investir ou não em um ativo.

Suspeitamos que o ativo vai valorizar nos próximos dias $\rightarrow$ Probabilidade(valorização) = 0,6.

Problema é totalmente diferente mas ...
a estrutura de decisão é idêntica.

Isso acontece em várias outras áreas:

1) **Linha de produção industrial**: controle de qualidade por amostragem;
2) **Marketing**: lançamento de produto;
3) **Agronomia**: utilização de variedades de uma planta;
4) **Psicologia**: reação a um estímulo.

Problema de uma decisão:

ela pode estar errada!

Não se quer tratar pessoas sadias . ⟷ Não se quer deixar de obter lucro.

Não se quer deixar de tratar pessoas doentes. ⟷ Não se quer investir para ter prejuízo.

Elementos de uma decisão:

1) Possíveis estados da natureza: D+ e D-

2)Probabilidades dos estados:
Prob(D+)=0,6 e Prob(D-)=0,4

3) Possíveis ações: T+ e T-

4) Utilidades (ou perdas):

	TRATAR	MANDAR EMBORA
DOENTE (D+)	500	-1000
SADIO (D-)	-500	100

Utilidades esperadas:

$E[U(T+)] = 0,6 \times 500 + 0,4 \times (-500) = 100$

$E[U(T–)] = 0,6 \times (-1000) + 0,4 \times 100 = -560$

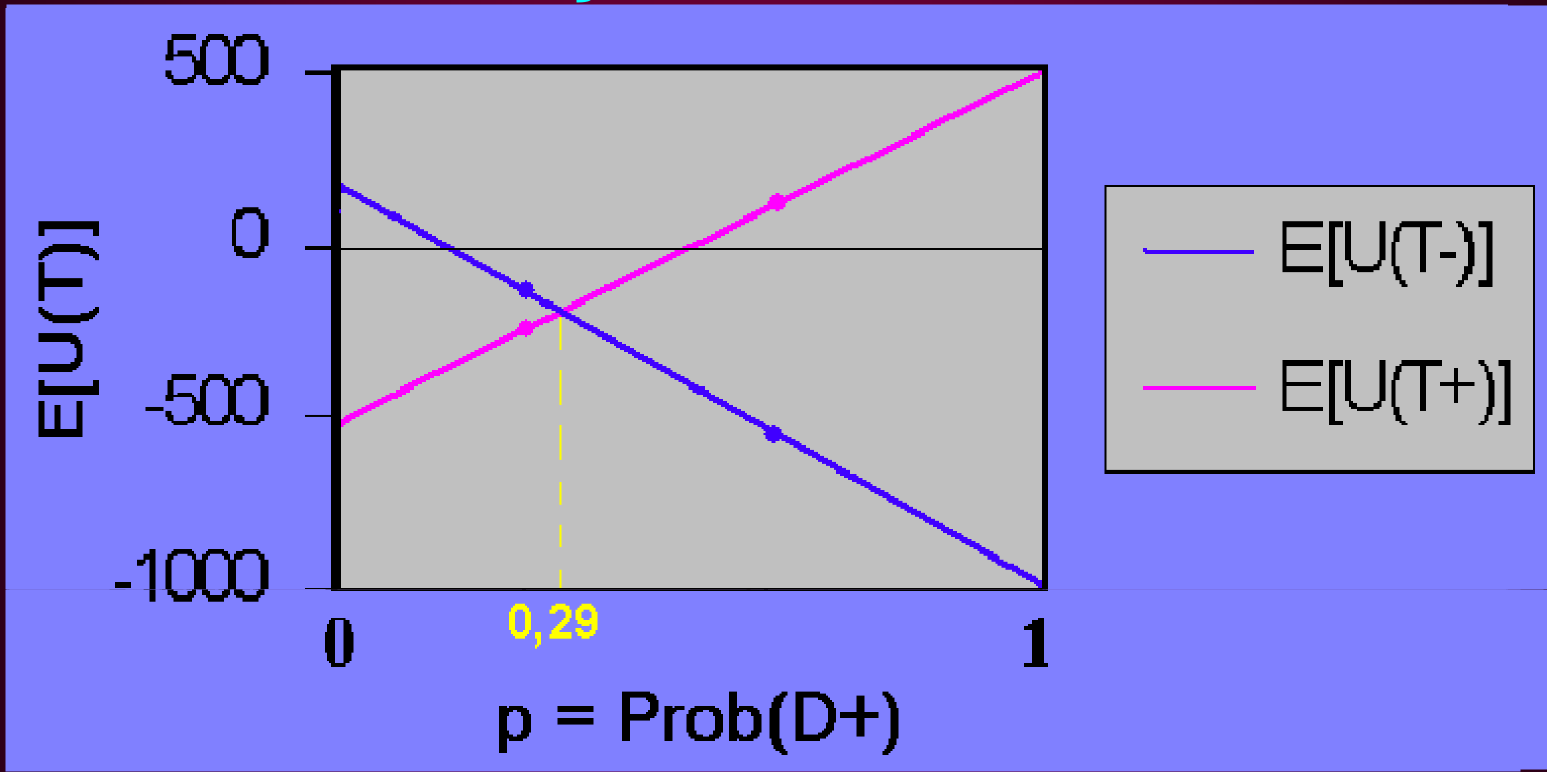
$E[U(T+)] > E[U(T–)]$ → decisão correta é **tratar o paciente**.

Suponha agora que Prob(D+)=0,25.
Refazendo as contas, temos:

$E[U(T+)] = -250$ $E[U(T–)] = -175$

$E[U(T+)] < E[U(T–)]$ → Agora, a decisão correta é **não tratar o paciente.**

Variação da decisão



Portanto, se p < 0,29 $\rightarrow$ T-, se p > 0,29 $\rightarrow$ T+

Esse processo de decisão é **puramente subjetivo** (nas probabilidades e utilidades).

Essa variação da decisão com *p* vai se repetir adiante numa aplicação de poluição ambiental.

Métodos *objetivos* de decisão

Envolvem necessariamente só medições.

Exemplo médico: exames clínicos.

Suponha que existe exame (E) para doença (D) com a seguinte performance:

$\text{Prob}(E+|\ D+) = 0{,}9$ ← alta sensibilidade

$\text{Prob}(E+|\ D-) = 0{,}2$ ← alta especificidade

Decisão agora é baseada em <u>estratégias</u> de ação

1) EA1 - agir de acordo com exame

$$E = E+ \rightarrow D = D+ \rightarrow T+$$

$$E = E- \rightarrow D = D- \rightarrow T-$$

☑ pois o exame é muito bom!

Mas e se Prob(E+| D-) = 0,5 ?
(exame tem baixa especificidade)

2) EA2 - sempre tratar o paciente

$$E = E+ \rightarrow T+$$

$$E = E- \rightarrow T+$$

Revendo, tabela de utilidades:

	Tratar	Mandar embora
Doente (D+)	500	-1000
Sadio (D-)	-500	100

e tabela de probabilidades:

	E+	E-
Doente (D+)	0,9	0,1
Sadio (D-)	0,2	0,8

Só se pode calcular utilidades esperadas supondo conhecido o estado de saúde do paciente.

$E[U(EA1)|\, D+] = 0{,}9 \times 500 + 0{,}1 \times (-1000) = 350$

$E[U(EA1)|\, D-] = 0{,}2 \times (-500) + 0{,}8 \times (100) = -20$

Estratégias de ação:

EA1 - agir de acordo com o exame

EA2 - sempre tratar paciente

	E[U(EA1)]	E[U(EA2)]
Doente (D+)	350	500
Sadio (D-)	-20	-500
Mínimo	-20	-500

Como comparar diferentes estratégias sem especificar probabilidades de doença?

Critério minimax: maximizar a menor utilidade.

Como -20 > -500 → EA1 é melhor.

Critério maximax → EA2 é melhor.

Combinando subjetividade com medições (Teorema de Bayes)

E traz informação sobre o estado do paciente:

diz quão verossímil cada estado é.

Exemplo:

1) Como $Prob(E+|D+)/Prob(E+|D-) = 0,9/0,2 = 4,5 \rightarrow$

sob E+, D+ é 4,5 vezes mais verossímil que D-

2)Como $Prob(E-|D+)/Prob(E-|D-) = 0,1/0,8 = 1/8 \rightarrow$

sob E-, D- é 8 vezes mais verossímil que D+

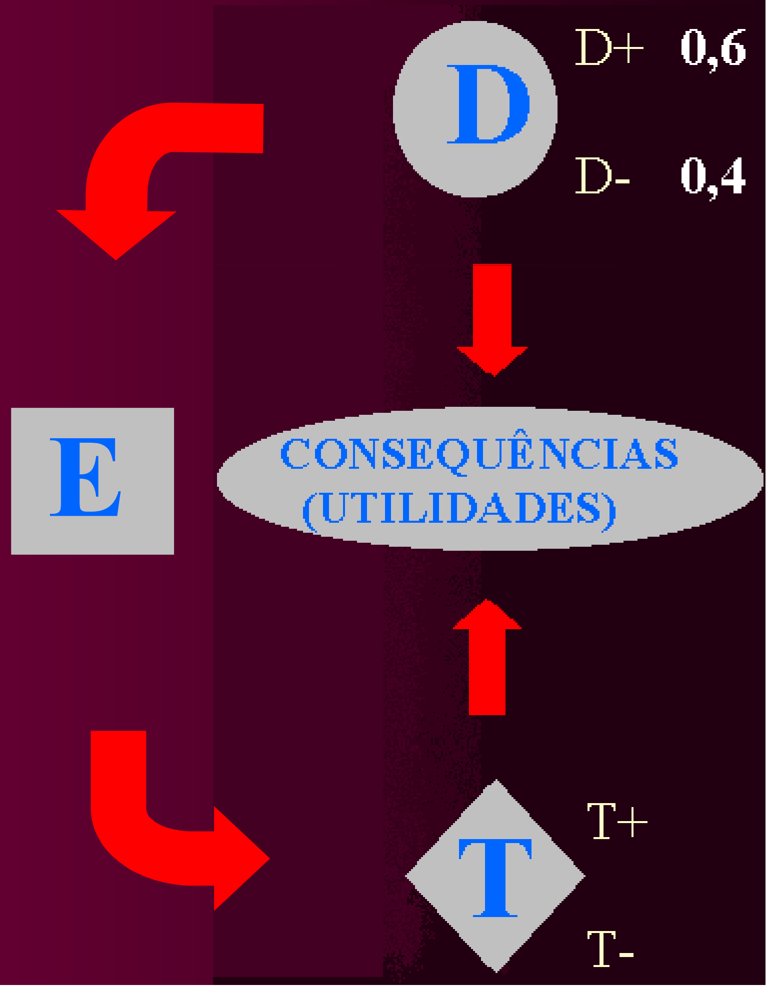
Verossimilhança: $V_E\,(D) = Prob(E|D)$

Teorema de Bayes:

$\text{Prob}(D|\,E) \propto V_E\,(D) \times \text{Prob}(D)$

Regra de atualização de probabilidades

Regra de decisão não muda: basta introduzir medições e usar as probabilidades atualizadas.



Operando o teorema de Bayes

Suponha que o Exame deu E+. Pelo teorema de Bayes:

$$\text{Prob}(D+|E+) \propto 0{,}9 \times 0{,}6 = 0{,}54$$

$$\text{Prob}(D-|E+) \propto 0{,}2 \times 0{,}4 = 0{,}08$$

Normalizando, $\text{Prob}(D+|E+) = 0{,}87$ e $\text{Prob}(D-|E+) = 0{,}13$

Se $\text{Prob}(D+) = 0{,}60 \xrightarrow[\text{de Bayes}]{\text{Teorema}} \text{Prob}(D+|E+) = 0{,}87$

→ decisão permanece igual (tratar o paciente).

Se $\text{Prob}(D+) = 0{,}25 \xrightarrow[\text{de Bayes}]{\text{Teorema}} \text{Prob}(D+|E+) = 0{,}66$

→ decisão do médico muda após exame.

Intervalo filosófico

Todo aprendizado na vida envolve combinar informação prévias com experiências que vamos acumulando.

Coerência → Teorema de Bayes é A regra de aprendizado que devemos usar.

Ele fornece a distribuição da incerteza sobre aquilo que não conhecemos dado aquilo que conhecemos

Incoerência → você se dispõe a entrar em jogos onde perde dinheiro com certeza.

Metodologia objetiva é questionável:
além dos problemas operacionais,
também envolve especificações subjetivas de
probabilidades e utilidades.

Melhor é reconhecer caráter subjetivo
e, de quebra, resolver problemas
adequadamente.

Generalização

No exemplo médico:

Estado da natureza : 2 possibilidades (D+/D-)

Medições: 2 possibilidades (E+/E-)

No caso geral, estado da natureza e/ou medições podem ter vários valores possíveis.

Exemplo: determinação de uma grandeza física (dureza de um material)
Estado da natureza (desconhecido) – X
Medição (a ser conhecida) - Y
O que podemos dizer sobre X (após observar Y)?

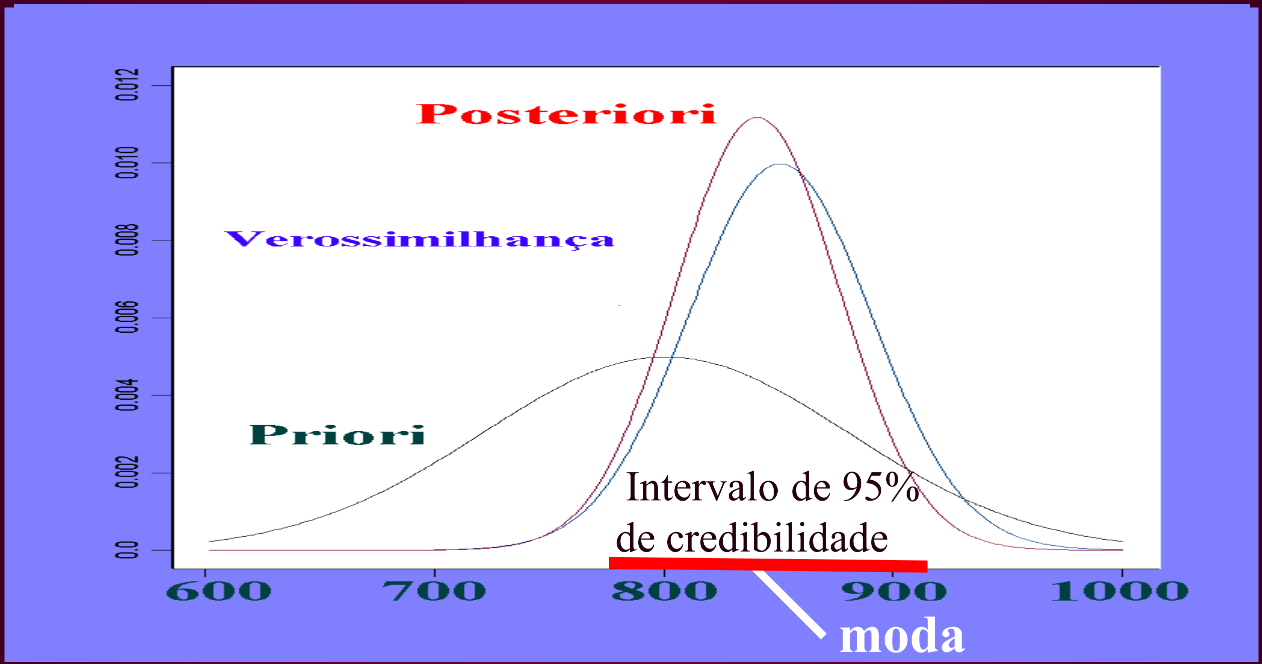
Pela teoria dos materiais ⟶ Prob(X = x).

Características da medição ⟶ Prob (Y = y |X = x).

Usualmente, Y = X + erro de medicão

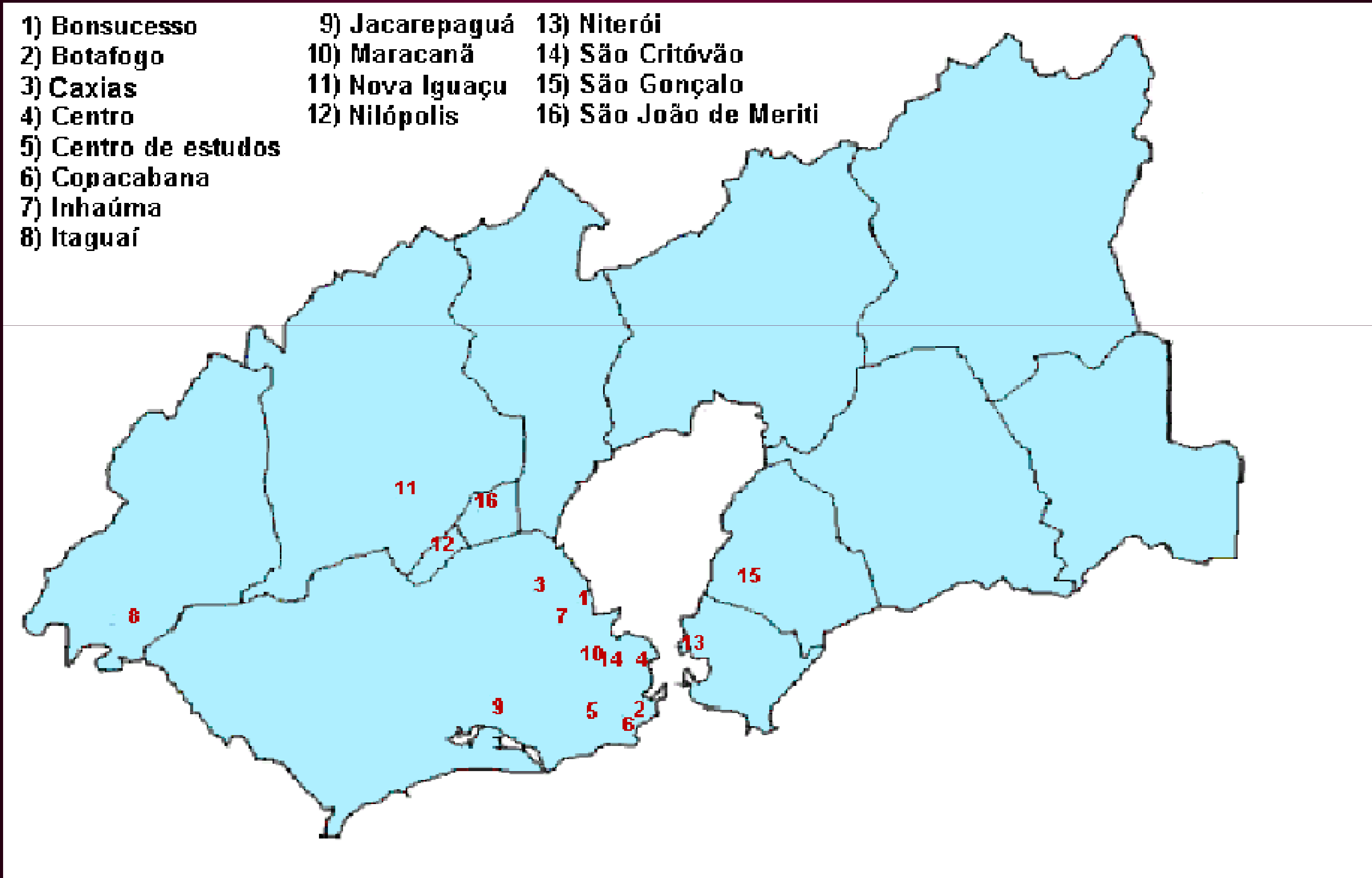
Após observar Y=y, temos V_y (X=x)

Teorema de Bayes ⟶ Prob (X= x |Y = y)



Aplicação (com Marina S. Paez)

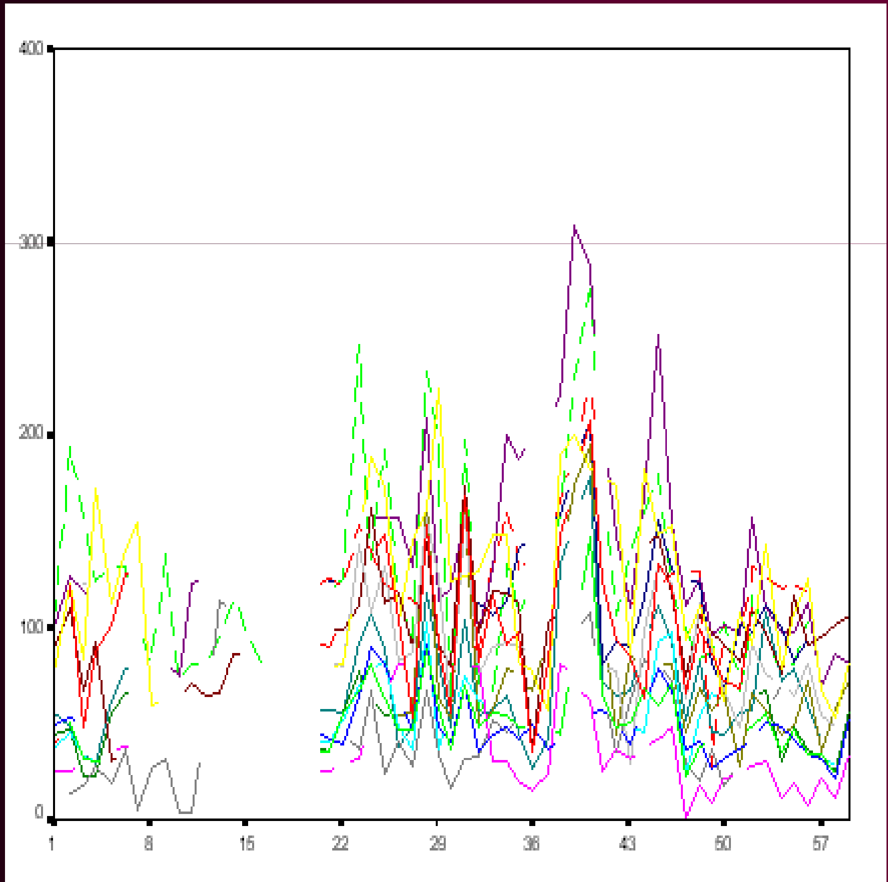
Monitoramento da poluição no Rio de Janeiro.



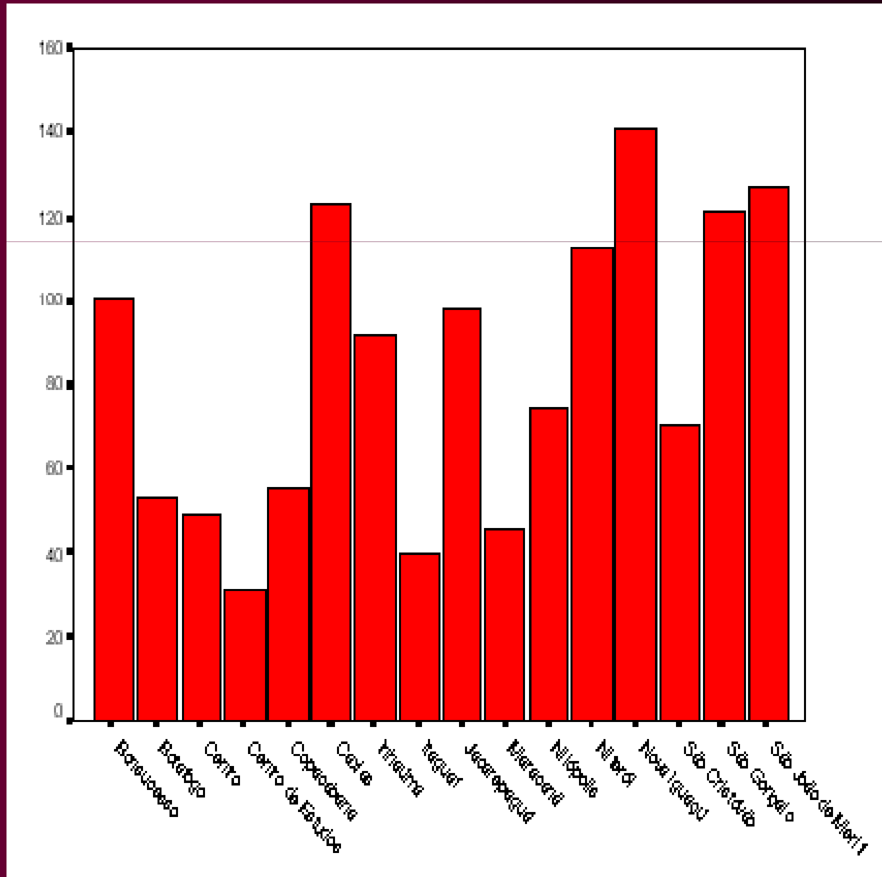
Medições: concentração de PM10

(material particulado, medido em $\mu g/m^3$) dados observados a cada 6 dias em 1999.

níveis ao longo do tempo



médias por estação



Elementos de uma decisão:

1) Órgãos fiscalizadores do governo:

nível de atenção - 100 $\mu g/m^3$.

Para cada local, a decisão será tomada em função de

$p = Prob(PM10 > 100 | Dados)$.

2) Possíveis ações:

T_1 : Não se preocupar

T_2 : Alerta (monitorar com maior frequência)

T_3 : Emergência (fechar fábricas, multar, ...)

Que decisão tomar em cada local?

Problema estatístico:

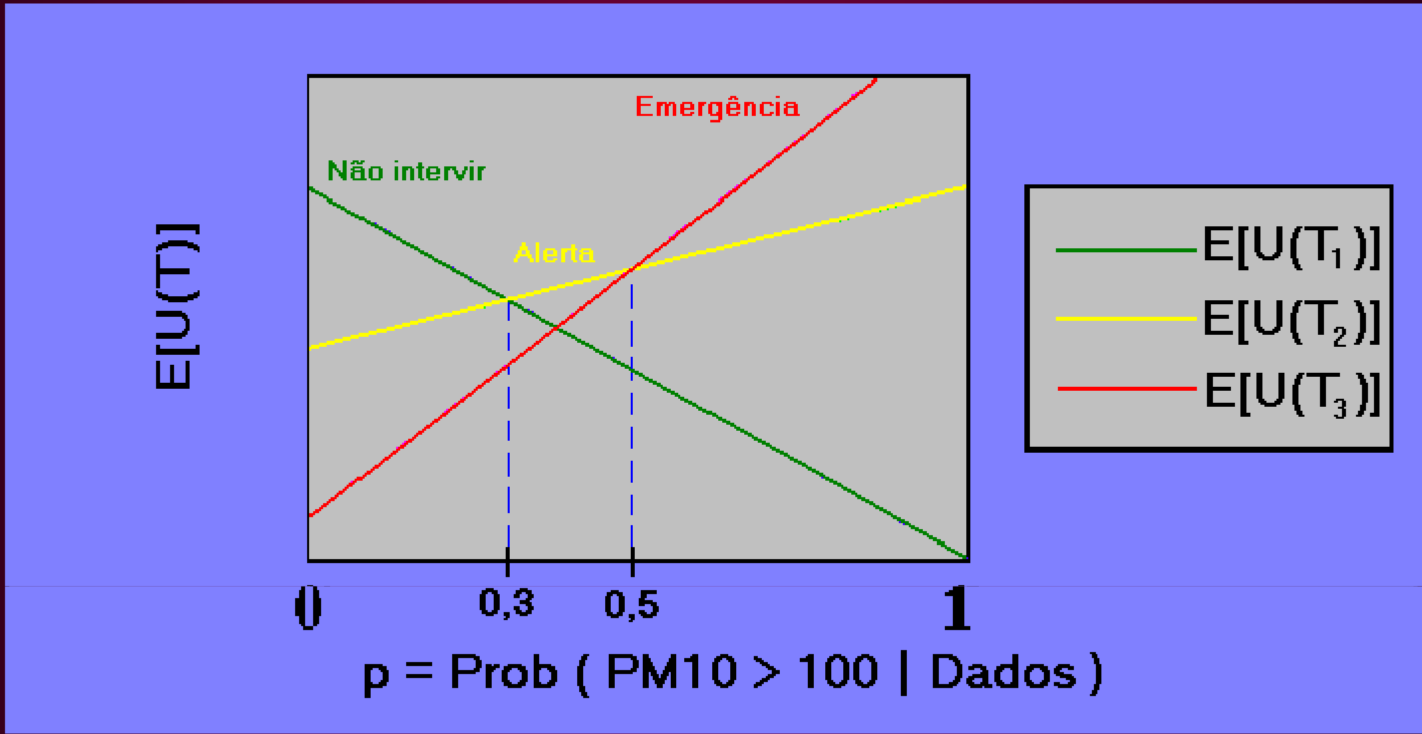
Expandir os resultados das medições dos 16 postos de monitoramento para todos os outros locais (interpolação espacial).

Problemas de uma decisão:

Não se quer intervir em áreas não poluídas. ⟷ Não se quer gastar dinheiro sem necessidade.

Não se quer deixar de intervir em áreas poluídas. ⟷ Não se quer pôr em risco a saúde da população.

Variação da decisão

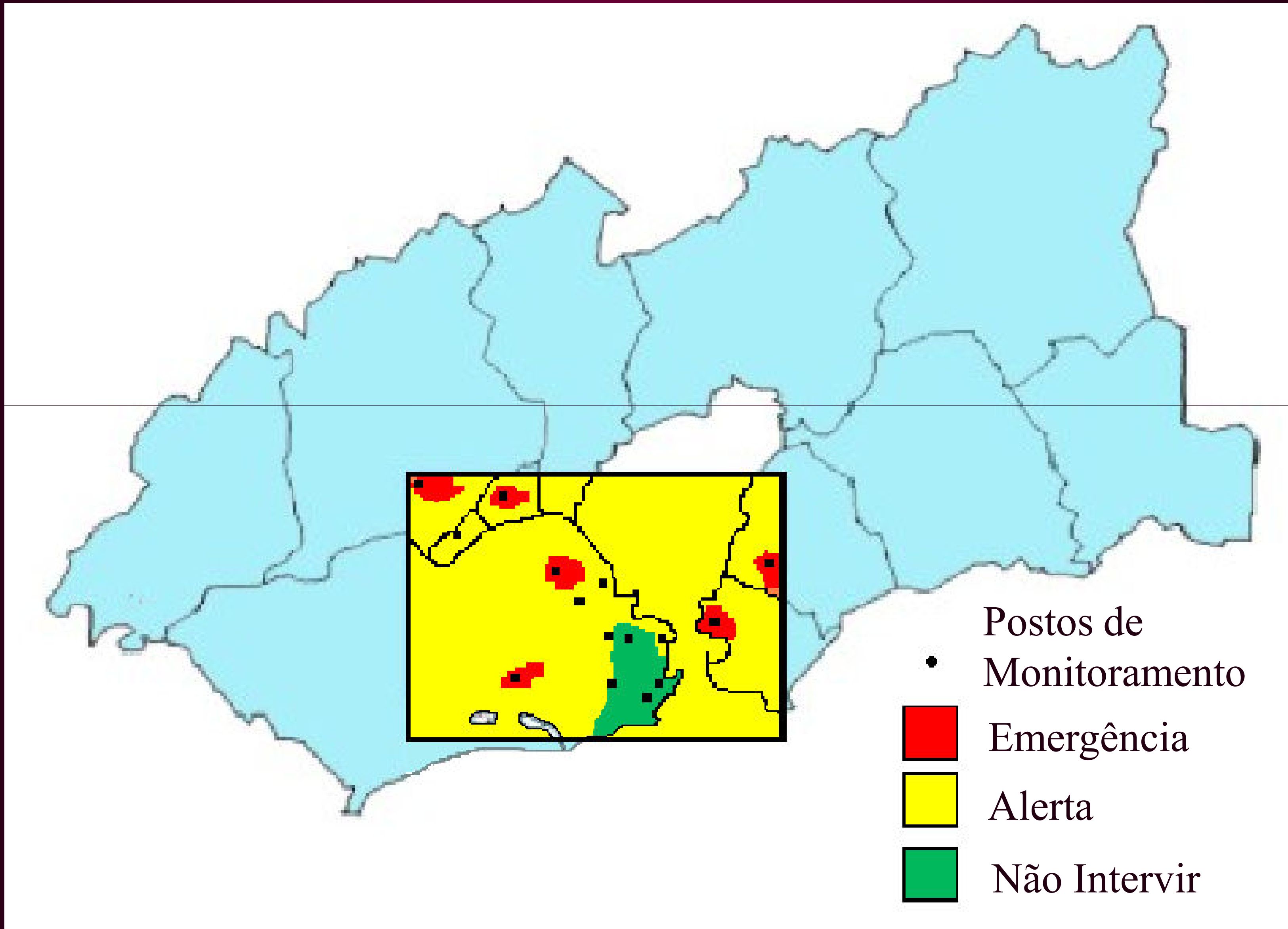


$0 < p < 0,3 \rightarrow$ **$E[U(T_1)] > E[U(T_2)]$** **$E[U(T_1)] > E[U(T_3)]$** ⇒ Não se preocupar

$0,3 < p < 0,5 \rightarrow$ **$E[U(T_2)] > E[U(T_1)]$** **$E[U(T_2)] > E[U(T_3)]$** ⇒ Alerta

$0,5 < p < 1 \rightarrow$ **$E[U(T_3)] > E[U(T_1)]$** **$E[U(T_3)] > E[U(T_2)]$** ⇒ Emergência

Interpolação espacial:



Conclusões

A Estatística utiliza métodos matemáticos para solucionar problemas reais de tomada de decisão.

Assim, em situações onde poderíamos contar unicamente com a sorte, temos um instrumento que nos possibilita aumentar as chances de tomar a decisão certa.

Como tomar decisões?

Dani Gamerman

Instituto de Matemática – UFRJ

Tel: 562 7911

Fax: 562 7374

dani@im.ufrj.br

http://acd.ufrj.br/~dani

Preparado por Thaís Cristina O Fonseca e Oswaldo G Souza Jr. (alunos de IC)