

# Mini- Curso

## Básico de Avaliação de Imóveis Com uso da Inferência Estatística

Arq. Urb. Ana Maria de Biazzí Dias de Oliveira  
[anabiazzi@uol.com.br](mailto:anabiazzi@uol.com.br)

*Arquiteta e Urbanista, Mestre em Engenharia Civil e Urbana pela Poli-USP. Certificada em Engenharia de Avaliações pelo IBAPE. Ex-Presidente do IBAPE/SP (biênio 2009-1010). Conselheira do CAU/SP. Professora da disciplina “Inferência Estatística” de cursos de pós-graduação e especialização em Engenharia de Avaliações e Perícias. Coautora do capítulo “Métodos Científicos e a Engenharia de Avaliações” do Livro - Engenharia de Avaliações do IBAPE/SP.*

Promoção:





## **Método comparativo direto de dados de mercado (ABNT 14.653-1 2019)**

***Identifica o valor de mercado do bem por meio de tratamento técnico dos atributos dos elementos comparáveis, constituintes da amostra.***

*Na aplicação deste método, a natureza dos bens, a indisponibilidade dos dados e de suas características, bem como os prazos limitados para a concepção da avaliação, podem levar à coleta de amostras que não atendem na íntegra aos pressupostos formais das amostras aleatórias simples, exigidos pelos modelos de estatística inferencial.*

*Assim, as amostras são mais bem descritas como “amostras acidentais”, que devem possuir a maior representatividade possível em relação à população, mesmo que não sejam utilizadas as técnicas tradicionais para a coleta de amostras aleatórias simples.*

*Para alcançar o máximo de representatividade da amostra, deve-se especificar claramente as características dos imóveis que compõem a população pesquisada, tomando como referência as características do imóvel avaliando, além de levar em consideração os aspectos citados em 6.4. Com a utilização desses cuidados, torna-se viável a aplicação de estatística inferencial.*



# Método Comparativo Direto

Tratamento dos dados  
da AMOSTRA para Comparação

Dados Homogêneos

Tratamento por Fatores  
HOMOGENEIZAÇÃO

Fundamentados em estudos  
de mercado, regionalizados e  
com intervalo de tempo  
definido

Dados Heterogêneos

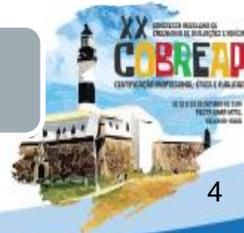
Regressão  
Linear

Redes Neurais  
(RLN)

Metodologia  
Científica

Análise  
Envoltória de  
Dados (DEA)

Regressão  
Espacial

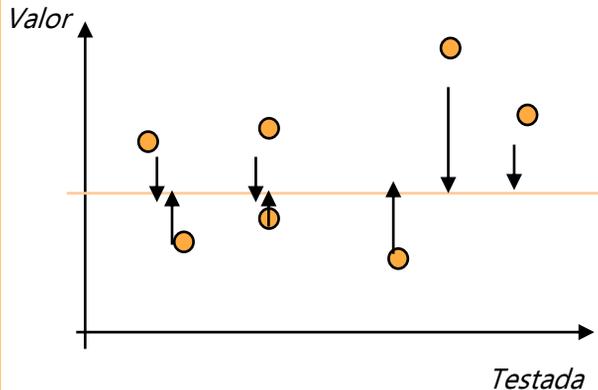


# MÉTODO COMPARATIVO

## PESQUISA DE MERCADO (AMOSTRA)

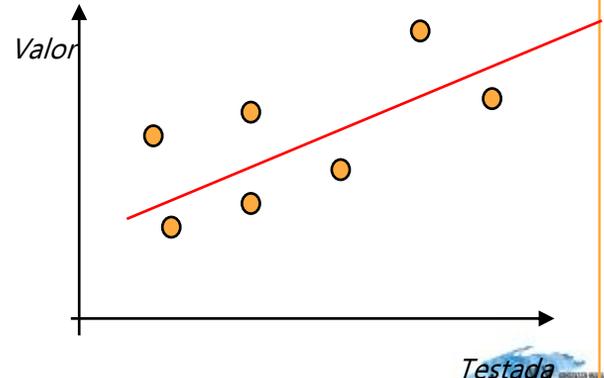
### Tratamento por Fatores X Tratamento Científico

**Tratamento por fatores:** utiliza-se "fatores" para ajustar os dados de mercado à média



**Tratamento científico**

(**análise de regressão**) procura-se encontrar a média que mais se aproxima dos dados de mercado



# Níveis de Especificação do Laudo de Avaliação

## item 9.2 ABNT 14.653-2



### Graus de FUNDAMENTAÇÃO

Função do aprofundamento do trabalho: confiabilidade, qualidade e quantidade dos dados disponíveis, podendo atingir

**Graus III, II ou I**

**Tabelas 1 (6 itens) e 2 (artigo 9.2.1 da Normas 14.653-2)**



### Graus de PRECISÃO

Função do **valor estimado** e o erro associado.  
**Amplitude do Intervalo de confiança de 80%**

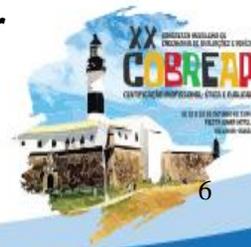
- $\leq 30\%$  - Grau III
- $\leq 40\%$  - Grau II
- $\leq 50\%$  - Grau I

Acima de 50%, não há classificação à precisão e - justificativa diagnóstico do mercado.

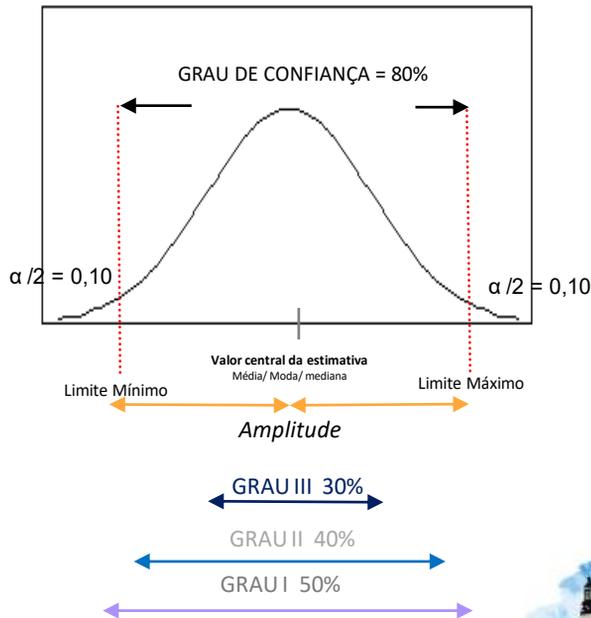
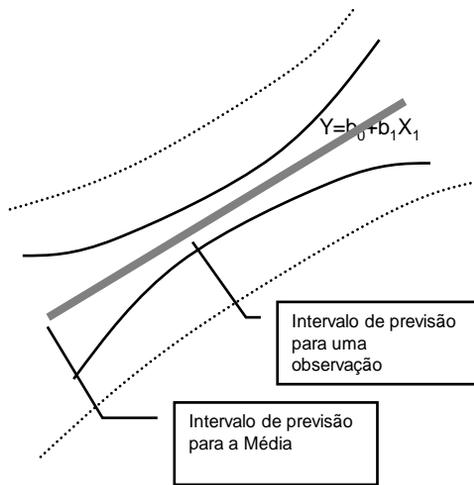
Tabela 5 (art.9.2.2 da Norma ABNT 14.653-2)

*Atender as recomendações do* **ANEXO A**

**Procedimentos para a utilização de modelos de regressão linear**



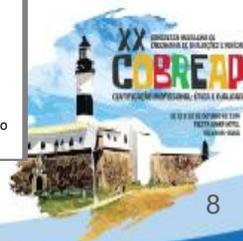
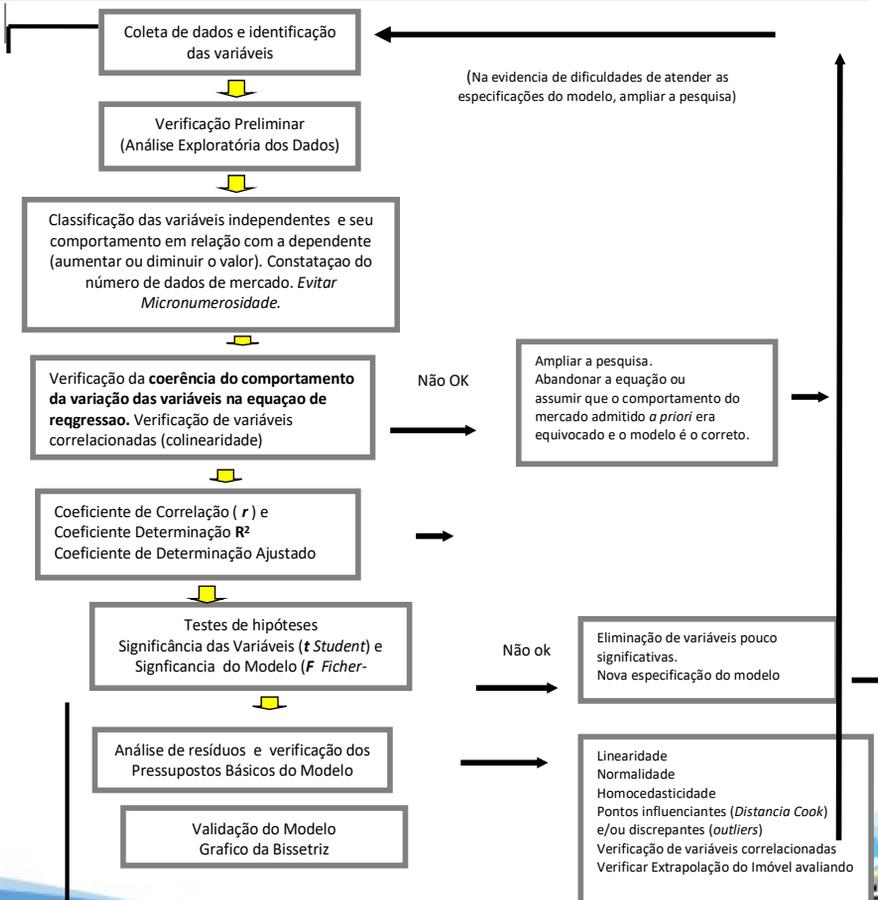
# Intervalo de confiança de 80% e Precisão do laudo



*A menor amplitude do intervalo ocorre quando as características do imóvel a avaliar estão próximas às médias das características dos dados*



# Caracterização do Imóvel Avaliando e do Contexto Imobiliário ao qual pertence





Escolha da equação de regressão para previsão do valor. Especificação do Laudo de Avaliação.

Preencher **Tabelas 1 e 2** das Normas da ABNT e classificar o Laudo de Avaliação no

**GRAU de FUNDAMENTAÇÃO (III,II ou I)**

Verificar a amplitude de Intervalo de confiança de 80% projetado para o valor do imóvel avaliando – e enquadrar no **GRAU de PRECISÃO**

**III - para amplitude < 30%**

**II - para amplitude < 40%**

**I - para amplitude < 50%**

*Nota Se ultrapassar 50% não há classificação de Precisão – justificar com diagnóstico de mercado*

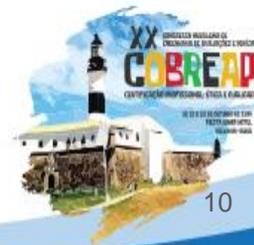


## Avaliação do bem – Conceitos iniciais NBR-146153-1 (2019)

Principais **estágios** com variações, dependendo do tipo de propriedade.

1º) **Estágio- identificação do problema** - onde são definidos a finalidade, o objetivo e identificação do bem a avaliar, conceituados no **item 6 Atividades Básicas**

- a) **finalidade:** *locação, arrendamento, comodato, aquisição, doação, alienação, dação em pagamento, permuta, garantia, fins contábeis, seguro, arrematação, adjudicação, indenização, tributação e outros;*
- b) **objetivo:** *valor de mercado de compra e venda ou de locação; outros valores, como valor em risco, valor patrimonial, valor econômico, custo de reedição, valor de liquidação forçada, valor de desmonte; indicadores de viabilidade ....*
- c) *prazo-limite para apresentação do laudo;*
- d) *expectativa em relação ao grau de fundamentação;*
- e) *forma de apresentação;*
- f) *condições a serem observadas, no caso de laudos de uso restrito.*





# Avaliação do bem – Conceitos iniciais NBR-146153-1 (2019)

## 3º) Estagio- Vistoria do bem - essencial



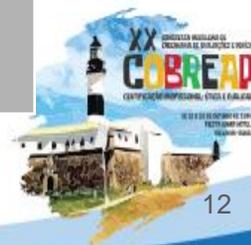
Em **casos excepcionais**, quando impossível ou inviável

### item 6.9 -14.653-1:2019 - Adoção de situação-paradigma

- **hipotética:** *avaliação de imóvel não vistoriado internamente e avaliação de terreno sem a consideração das benfeitorias existentes;*
- **virtual:** *avaliação de imóvel ou empreendimento ainda em fase de projeto ou construção, considerado na condição de pronto, e avaliação de imóvel cujas benfeitorias foram modificadas ou destruídas*

### Grau de fundamentação (Tabela 1 – item 9.2.1 –14.653-2

Item	Descrição	Grau		
		III	II	I
1	Caracterização do imóvel avaliando	Completa quanto a todas as variáveis analisadas	Completa quanto às variáveis utilizadas no modelo	Adoção de situação paradigma





# Tipos de Variáveis

**variáveis quantitativas:** variáveis que podem ser medidas ou contadas

**Contínua** – pode assumir infinitos valores dentro de seus limites • Exemplo: Área

**Discreta** – assume finitos valores • Exemplo: No. de Dormitórios

**variáveis qualitativas:** variáveis que não podem ser medidas ou contadas, mas apenas ordenadas ou hierarquizadas

**variáveis dicotômicas:** assume duas posições. .Padrão

**Baixo=0 Médio =1**

**códigos alocados:** escala lógica ordenada para diferenciar as características qualitativas

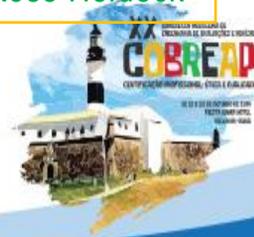
Padrão **Baixo 1; Médio, 2; Alto 3.**

**variável “Proxy”:** substitui outra de difícil mensuração e que se presume guardar relação de pertinência:

**Padrão construtivo - CUB**

**Localização - Índice Fiscal ou Renda familiar na localidade estudada**

**Estado de conservação - Ross Heideck**



# Codificação binária

$n$  situações  $n-1$  variáveis dicotômicas

Exemplo: 3 (três) bairros distintos *bairro A*; *bairro B* e *bairro C*

Variável  $X_1$ =Bairro B = 1 (imóvel localizado no Bairro B);

= 0 (zero) imóvel localizado em outro bairro;

Variável  $X_2$ =Bairro C = 1 (imóvel localizado no Bairro C);

= 0 (zero) imóvel localizado em outro bairro

	$X_1$	$X_2$	Y
A	0	0	$b_0$
B	1	0	$b_0+b_1$
C	0	1	$b_0+b_2$

Códigos Ajustados ( $x_1$ )

$$A = b_0/b_0$$

$$B = (b_0+b_1)/b_0$$

$$C = (b_0+b_2)/b_0$$

$$Y = b_0 + b_1 \cdot x_1 + b_2 \cdot x_2$$



# Variáveis qualitativas expressas por Códigos Alocados

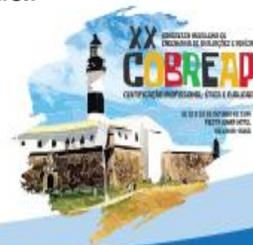
Para variáveis de tipo nominal como de tipo ordinal é possível por razões de conveniência é possível associar valores numéricos às diferentes categorias.

Código Alocado
A = 1
B = 2
C = 3

$$Y = b_0 + b_1 * x_1$$



Importante - estes valores numéricos não tem significado como tal. Ou seja, ao assumir os valores 1, 2 e 3 para as categorias A,B,C, os números 1, 2 e 3 são simbólicos para representar as categorias. Não pode significar por exemplo que significa padrão fino tenha valor igual ao dobro do valor do superior.



## 5º) Estagio - Quantidade de dados de mercado ( $n$ )

Atender o **item 2** dos Graus de Fundamentação da **ABNT-NBR 14653-2**:

<b>2</b>	Quantidade mínima de dados de mercado, efetivamente utilizados	<b>III</b>	<b>II</b>	<b>I</b>
		<b>6 (k + 1)</b> , onde k é o número de variáveis independentes	<b>4 (k + 1)</b> , onde k é o número de variáveis independentes	<b>3 (k + 1)</b> , onde k é o número de variáveis independentes

Evitar a **micronumerosidade** - o número mínimo de dados efetivamente utilizados ( $n$ ) no modelo deve obedecer aos seguintes critérios, com respeito ao número de variáveis independentes ( $k$ ):

$$n \geq 3 (k + 1)$$

para  $n \leq 30$ ,  $n_i \geq 3$

para  $30 < n \leq 100$ ,  $n_i \geq 10\% n$

para  $n > 100$ ,  $n_i \geq 10$

$n_i$  é o número de dados de mesma característica, no caso de utilização de variáveis **dicotômicas e variáveis qualitativas expressas por códigos alocados** ou códigos ajustados;



## 4º) Estágio – Montagem do Modelo de Regressão Linear

Variável Dependente

Intercepto

Parâmetros

Variáveis Independentes,

Erro Aleatório

$$Y_i = b_0 + b_1 X_{i1} + \dots + b_k X_{ik} (+ \epsilon)$$

$Y =$  (Valor)  $\left[ \begin{array}{l} \text{Predito por } X_1, X_2, X_3, \dots, \\ \text{segundo uma função} \end{array} \right] + \left[ \begin{array}{l} \text{Efeito aleatório} \end{array} \right]$

A finalidade de uma equação de regressão é estimar valores de uma **variável (y- valor de um imóvel)**, com base em **valores conhecidos das outras (x1, x2, x3, área, localização, padrão....)** Ou seja, é a **média de Y, condicional às informações de X (E[Y|X])**.

onde  $\epsilon$  é uma variável aleatória, que produz desvios devido a :

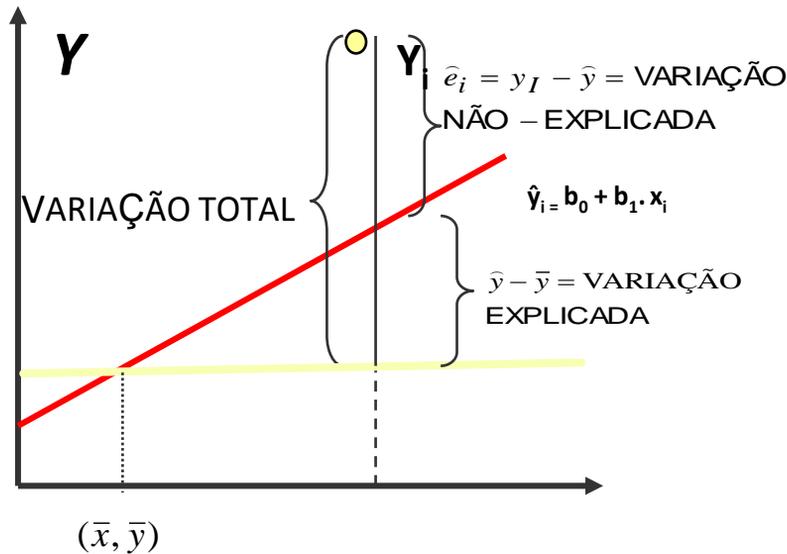
- variáveis de significância menor, não incluídas no modelo;
- imprevisibilidade do comportamento humano;
- erros de observação ou medidas





# Regressão Múltipla - Significância do Modelo

Teste  $F$  – calcula-se a razão entre as variâncias explicada e não explicada e compara-se com  $F$  tabelado por Fisher



# INTERPRETAÇÃO DOS RESULTADOS

A primeira estatística de interesse é o **coeficiente de determinação ( $R^2$ )** usualmente interpretado como a proporção da variância na variável dependente explicada pela variação das variáveis independentes.

Não se pode avaliar a capacidade explicativa de um modelo de regressão a partir do  $R^2$ .

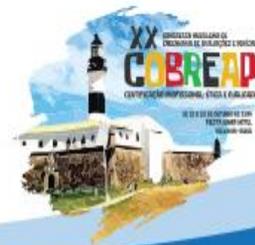
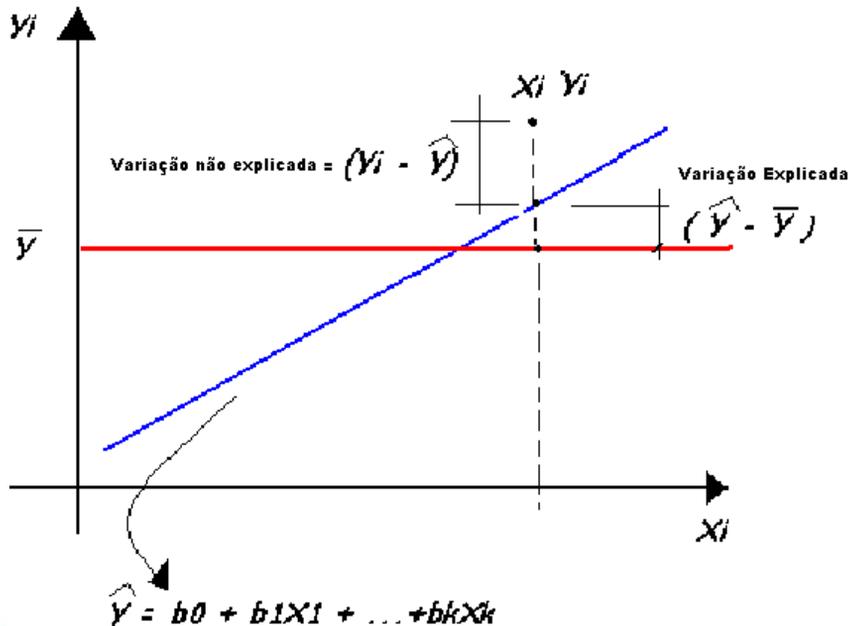
O  $R^2$  **ajustado** é uma medida similar ao  $R^2$ , mas que controla pelo número de observações e variáveis incluídas no modelo.

Embora a adição de variáveis independentes irá sempre fazer com que o coeficiente de determinação aumente, o **coeficiente de determinação ajustado** pode DIMINUIR se as variáveis independentes acrescentadas têm pouco poder explicativo



# Regressão Múltipla Significância do Modelo

## Decomposição da variável



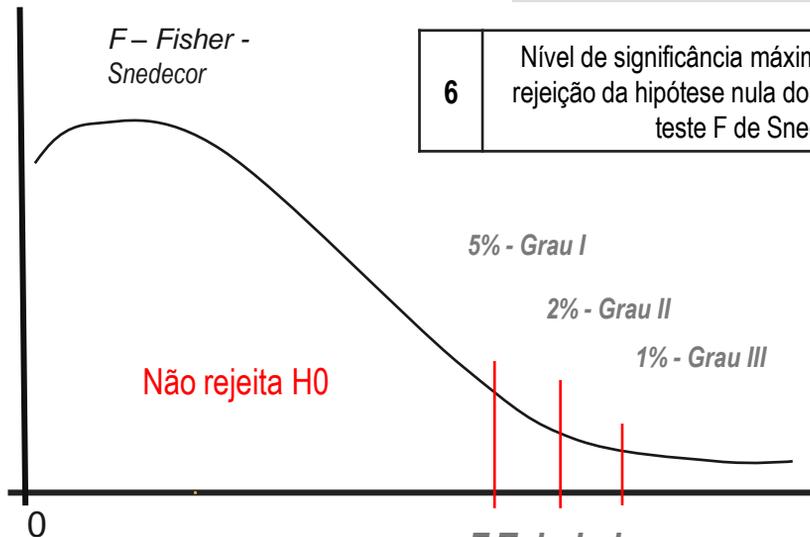
# Testes Significância do Modelo

**H0:**  $b_1, b_2, b_3, b_4, b_5 = 0$

**H1:**  $b_1, b_2, b_3, b_4, b_5 \neq 0$

**Hipótese Nula** *nenhuma variável escolhida é importante para explicar a variabilidade dos preços.*

**Hipótese Alternativa** – *peelo menos uma variável é importante*



<b>6</b>	Nível de significância máximo admitido para a rejeição da hipótese nula do modelo através do teste F de Snedecor	III 1 %	II 2 %	I 5 %
----------	--	------------	-----------	----------



Fundamentação – item 6  
Tabela ABNT

**F Tabelado**  
Graus de Liberdade  
Numerador  $K$   
Denominador  $(n-k-1)$

F Calculado do modelo



# Testes de Hipóteses das variáveis isoladas

$$Y = + b_0 - b_1 * X_1 + b_2 * X_2 - b_3 * X_3$$



Portanto, em função da variação de  $x$ , são obtidos  $Y$ , (ou valores), diferentes

Formular a hipótese nula  $H_0$  e a hipótese alternativa  $H_1$

$H_0 : b_1 = 0$  *a variável não é importante para explicar a variabilidade dos preços.*

$H_1 : b_1 \neq 0$  

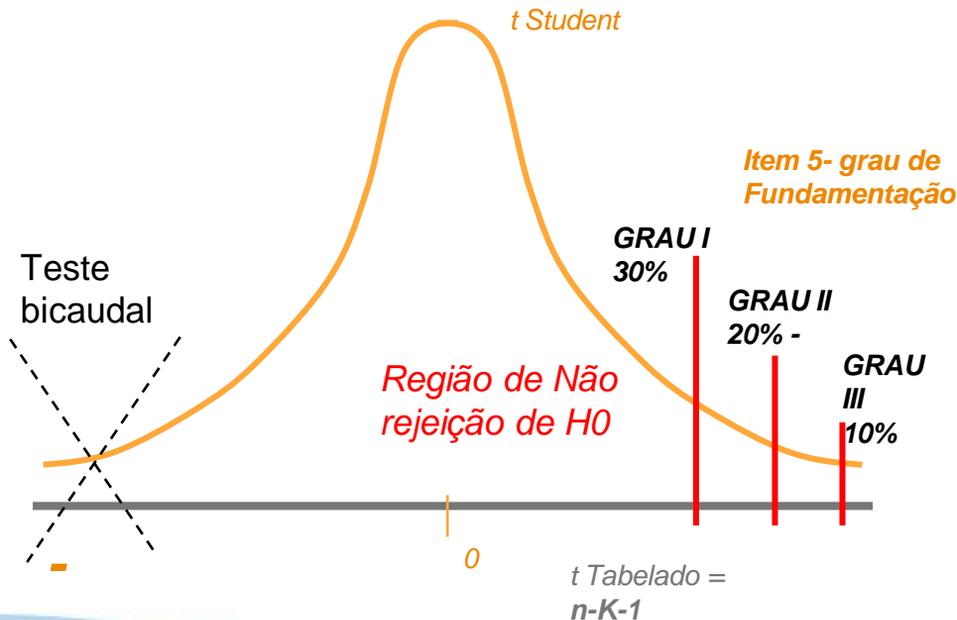
Se o coeficiente  $b_1$  for Provável de ser zero, então  $y$  não depende do valor de  $x$ . Isso significa dizer que eles não são linearmente relacionados. Alternativamente, se o valor  $b_1$  não é igual a zero, as duas variáveis estão relacionadas.

Importante que os *coeficientes* ( $b_1, b_2, b_3, b_4, b_5$ ) da sejam significativos na explicação da variação do valor ( $Y$ )



# Teste de Significância Individual das variáveis

Item	Descrição	Grau		
		III	II	I
5	Nível de significância (somatório do valor das duas caudas) máximo para a rejeição da hipótese nula de cada regressor (teste bicaudal)	10%	20%	30%





**Distribuição t de Student**

Coeficiente de Confiança						
Duas caudas	0,80	0,90	0,95	0,98	0,990	0,9990
Uma cauda	0,90	0,95	0,98	0,99	0,995	0,9995
Testes de Hipóteses						
Duas caudas	0,20	0,10	0,05	0,02	0,010	0,0010
Uma cauda	0,10	0,05	0,03	0,01	0,005	0,0005
1	3,078	6,314	12,706	31,821	63,656	636,578
2	1,886	2,920	4,303	6,965	9,925	31,600
3	1,638	2,353	3,182	4,541	5,841	12,924
4	1,533	2,132	2,776	3,747	4,604	8,610
5	1,476	2,015	2,571	3,365	4,032	6,869
6	1,440	1,943	2,447	3,143	3,707	5,959
7	1,415	1,895	2,365	2,998	3,499	5,408
8	1,397	1,860	2,306	2,896	3,355	5,041
9	1,383	1,833	2,262	2,821	3,250	4,781
10	1,372	1,812	2,228	2,764	3,169	4,587
11	1,363	1,796	2,201	2,718	3,106	4,437
12	1,356	1,782	2,179	2,681	3,055	4,318
13	1,350	1,771	2,160	2,650	3,012	4,221
14	1,345	1,761	2,145	2,624	2,977	4,140
15	1,341	1,753	2,131	2,602	2,947	4,073
16	1,337	1,746	2,120	2,583	2,921	4,015
17	1,333	1,740	2,110	2,567	2,898	3,965
18	1,330	1,734	2,101	2,552	2,878	3,922
19	1,328	1,729	2,093	2,539	2,861	3,883
20	1,325	1,725	2,086	2,528	2,845	3,850
21	1,323	1,721	2,080	2,518	2,831	3,819
22	1,321	1,717	2,074	2,508	2,819	3,792
23	1,319	1,714	2,069	2,500	2,807	3,768
24	1,318	1,711	2,064	2,492	2,797	3,745
25	1,316	1,708	2,060	2,485	2,787	3,725
26	1,315	1,706	2,056	2,479	2,779	3,707

**G.L =Graus de Liberdade -Amostra com 25 dados e 4 variáveis (n-k-1) =20**



## A.2 Pressupostos básicos (anexo A – ABNT 14.653-2)

“Ressalta-se a necessidade, quando se usam modelos de regressão, de observar os seus pressupostos básicos, principalmente no que concerne à sua especificação, **normalidade**, **homocedasticidade**, **não-multicolinearidade**, **não-autocorrelação**, **independência** e **inexistência de pontos atípicos**, com o objetivo de obter avaliações não tendenciosas, eficientes e consistentes, em especial as seguintes:



# Pressupostos Básicos – *Modelo de Regressão Linear*

## *Primeira hipótese*

- O número de dados ( $n$ ) deve ser superior ao número de parâmetros estimados ( $K$ )

*Definido nas formulas do **item 2** do Grau de Fundamentação do Laudo*

GRAU III =  $6(k+1)$

GRAU II =  $4(k+1)$

GRAU I =  $3(k+1)$



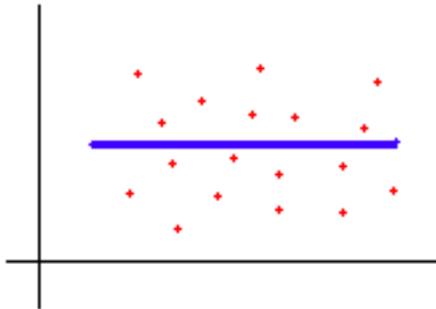


# Pressupostos Básicos – Modelo de Regressão Linear

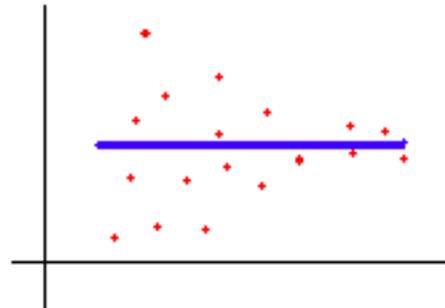
## Terceira hipótese

- Os erros são variáveis aleatórias com variância constante - Princípio da Homocedasticidade.

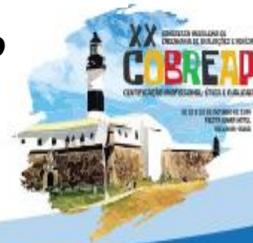
Verificada pela análise gráfica dos *resíduos* versus *valores ajustados*, que devem apresentar pontos dispostos aleatoriamente, sem nenhum padrão definido;



**Homocedástico**



**Heterocedástico**



## Pressupostos Básicos – Modelo de Regressão Linear

### Quarta hipótese

- **Ausência de Colinearidade** (ou multicolinearidade) **entre as variáveis independentes.**

não é permitida a existência de relação linear entre qualquer das variáveis explicativas, ou seja, elas têm de ser linearmente independentes.

Para verificação da multicolinearidade deve-se, em primeiro lugar, analisar a **Matriz das Correlações**, que espelha as dependências lineares de primeira ordem entre as variáveis independentes, com atenção especial para resultados superiores a 0,80

efeitos da colinearidade - *imprecisão no cálculo dos parâmetros da regressão*

- variâncias das estimativas dos parâmetros podem ser muito grandes e acarretar a aceitação da hipótese nula e a eliminação de variáveis fundamentais
- inversão de sinais em relação ao esperado
- impossibilidade de separar dos efeitos nos preços entre as variáveis independentes



# Pressupostos Básicos – Modelo de Regressão Linear

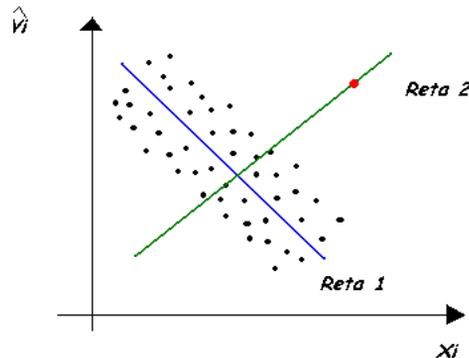
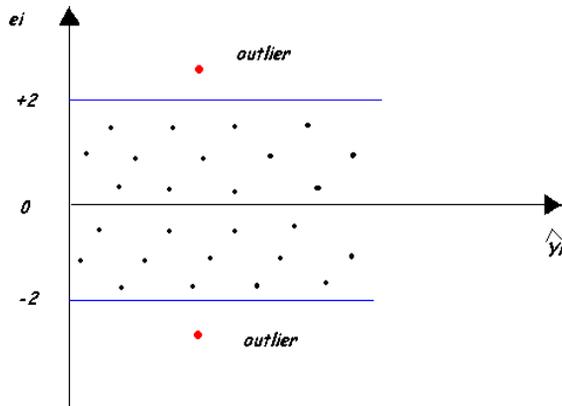
## Quinta hipótese Observações Atípicas

### Outlier

Elemento que apresenta grande resíduo em relação aos dados da amostra, ou resíduo superior a dois desvios padrão.

### Pontos influenciantes

Aqueles que embora apresentem pequenos resíduos se distanciam da massa de dados, alterando a tendência do mercado.



# Pontos influenciantes

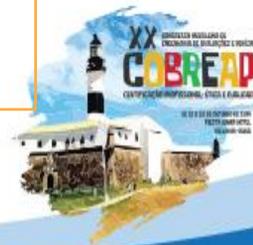
- Quando usamos regressão múltipla, ocasionalmente se encontra que algum subconjunto de observações influentes.
- Algumas vezes, essas observações que influenciam estão relativamente longe da vizinhança onde o resto dos dados foi coletado



Se esses pontos que influenciam forem pontos "ruins", ou errôneos de algum modo, então eles devem ser eliminados.

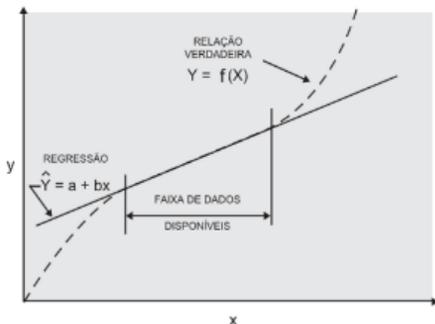
Distância de Cook (se  $D > 1$ , o ponto é influente na regressão – pode também ser um *outlier*)

Para  $D < 1$ , a observação é a que tem maior valor de influência na regressão, mas está dentro da gama de valores

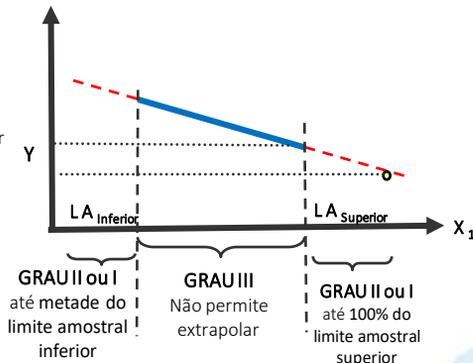


# EXTRAPOLAÇÕES ( item 5 da tabela Fundamentação ABNT 14.653-2

Item	Grau		
	III	II	I
4	Não admitida	Admitida para apenas uma variável, desde que: a) as medidas das características do imóvel avaliando não sejam superiores a 100% do limite amostral superior, nem inferiores à metade do limite amostral inferior. b) o valor estimado não ultrapasse 15% do valor calculado no limite da fronteira amostral, para a referida variável.	Admitida, desde que: a) as medidas das características do imóvel avaliando não sejam superiores a 100% do limite amostral superior, nem inferiores à metade do limite amostral inferior. b) o valor estimado não ultrapasse 20% do valor calculado no limite da fronteira amostral, para as referidas variáveis, simultaneamente.



Diferença inferior a 15% GRAU II; a 20% GRAU I











## Grau de precisão - Tabela 5

Descrição	Grau		
	III	II	I
Amplitude do intervalo de confiança de 80 % em torno da estimativa de tendência central	≤ 30 %	≤ 40 %	≤ 50 %

*Nota: Quando a amplitude do intervalo de confiança ultrapassar 50%, não há classificação do resultado quanto à precisão e é necessária justificativa com base no diagnóstico do mercado.*

### **Campo de arbítrio** NBR 14653-2

**8.2.1.5.1** O **campo de arbítrio** definido em 3.8 da ABNT NBR 14653-1:2001 é o intervalo com **amplitude de 15%, para mais e para menos**, em torno da estimativa de tendência central utilizada na avaliação.

**8.2.1.5.2** O campo de arbítrio pode ser utilizado quando variáveis relevantes para a avaliação do imóvel não tiverem sido contempladas no modelo, por escassez de dados de mercado, por inexistência de fatores de homogeneização aplicáveis ou porque essas variáveis não se apresentaram estatisticamente significantes em modelos de regressão, desde que a amplitude de até mais ou menos 15 % seja suficiente para absorver as influências não consideradas e que os ajustes sejam justificados.





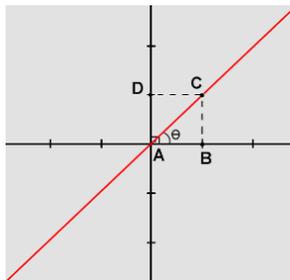
## validação do modelo – 14653-2, :

“O poder de predição do modelo **deve** ser verificado a partir do gráfico de preços observados na abscissa versus valores estimados pelo modelo na ordenada, que deve apresentar pontos próximos da bissetriz do primeiro quadrante. Alternativamente, podem ser utilizados procedimentos de validação.”

**Nota explicativa:** *Bissetriz do primeiro quadrante*

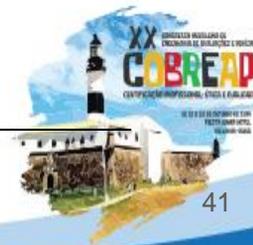
Sabendo que  $y=x$ , o gráfico dessa função passa pelos pontos (0,0) e (1,1).

Sejam  $A=(0,0)$ ,  $B=(1,0)$ ,  $C=(1,1)$ ,  $D=(0,1)$  e  $\theta$  o ângulo que a reta  $y=x$  forma com o eixo  $x$ , conforme gráfico:



Tem-se: 
$$\operatorname{tg}\theta = \frac{BC}{AB} = \frac{1}{1} = 1 \Rightarrow \theta = 45^\circ$$

Logo a reta  $y=x$  contém as bissetrizes do  $1^\circ$  (e do  $3^\circ$ ) quadrantes.



Obrigada

