

RADEGAZ NASSER JUNIOR
LUIZ ALBERTO PRETTI

**ANÁLISE DOS RESÍDUOS NA REGRESSÃO LINEAR MÚLTIPLA
PELO MÉTODO DOS MÍNIMOS QUADRADOS ORDINÁRIOS**

Trabalho de Avaliação

Goiânia/GO
2021

Análise dos Resíduos na Regressão Linear Múltipla Pelo Método dos Mínimos Quadrados Ordinários

Resumo

A Regressão Linear Múltipla pelo Método dos Mínimos Quadrados Ordinários é uma técnica utilizada intensivamente pelos profissionais da área de avaliações imobiliárias para obter os valores de mercados de imóveis de tipologias diversas. Para alcançar seus objetivos seguem recomendações normativas ditadas pela ABNT NBR 14.653/1-2019 e ABNT NBR 14.653/2 – 2011 e invariavelmente concluem estabelecendo Graus de Precisão e Fundamentação como recomendados por esta norma. Entretanto, mesmo procedendo assim, baseado nas citadas ABNT NBR, não há qualquer garantia de que a Regressão Linear Múltipla tenha alcançado de fato seus objetivos técnicos no que se refere ao atendimento dos pressupostos estatísticos mínimos e indispensáveis, ou seja, a possibilidade de violação destes pressupostos continua presente. Este artigo se propõe a explicitar o passo a passo das análises informais e direcionar onde necessário às análises técnicas formais teóricas, possibilitando análises mais consistentes que conduzam a resultados mais justificáveis e consistentes.

Palavras chave: ***Análise; Regressão Linear; Resíduos; Pressupostos.***

Introdução

O modelo de regressão linear simples (RLS) ou múltipla (RLM), aplicado a avaliação de bens imobiliários, tem por objetivo estimar os estimadores do modelo estatístico inferencial e testar hipóteses sobre eles a fim de comprovar se realmente representam os parâmetros da população examinada. Esses estimadores em regra, são os coeficientes, a média e o desvio padrão que após testados podem representar os parâmetros da população ou não.

Em caso afirmativo devem ainda atender a algumas premissas quanto ao comportamento da variável em relação à colaboração na formação da variável dependente, isto é, se contribuem positivamente ou negativamente.

O MMQO permite a obtenção do modelo inicial informando à priori que os resíduos ou erros da estimação mostram-se normalmente distribuídos ou não.

Para que a análise seja completa avança-se ao método clássico da regressão linear (MCRL) onde são analisados os pressupostos estatísticos que conferirão confiabilidade o modelo, dentre os quais se destacam:

- 1 – Preparação da base dados ou amostra
- 2 – Minimização dos resíduos (erros)
- 3 – Normalidade dos resíduos.
- 4 – Outliers e Pontos Influenciadores dos resíduos
- 5 – Independência dos resíduos
- 6 – Linearidade dos resíduos
- 7 – Homocedastidade dos resíduos
- 8 – Multicolinearidade das variáveis

A elaboração eficiente e consistente da análise de uma Regressão Linear Múltipla (RLM) pelo Método dos Mínimos Quadrados Ordinários (MMQO) exige que os erros residuais sejam minimizados por métodos algébricos em relação a uma tendência média, com o objetivo de obter estimativas tanto quanto precisas necessárias para o atendimento de regras de precisão e especificações de confiabilidade e erro.

Preparação da base dados ou amostra

Inicialmente, conhecidas as características físicas do bem imobiliário a avaliar, identifica-se os atributos correspondentes que os agentes de mercado normalmente valorizam e divulgam nos meios de comunicação locais além dos comumente encontrados em imóveis da mesma tipologia.

Com essas informações, constrói-se uma matriz de dados numéricos, também chamada de “data frame” porque identifica nominalmente as variáveis e seus atributos, onde as variáveis serão armazenadas conforme suas características, quantitativas e qualitativas, para o preenchimento com as informações coletadas na pesquisa imobiliária.

Esses procedimentos devem seguir as orientações constantes da ABNT NBR 14.653-2 de 2011.

Minimização dos resíduos (erros)

A minimização dos resíduos dos dados pertencentes a uma amostra em análise estatística consiste em aproximar tanto quanto possível os pontos dispersos em um espaço xy de uma reta de tendência como mostrado na figura 1 seguinte:

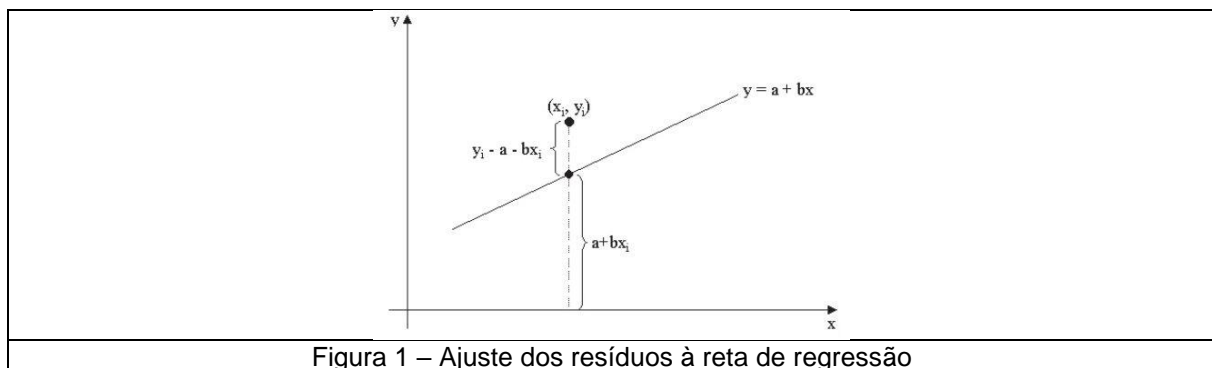


Figura 1 – Ajuste dos resíduos à reta de regressão

Examinando-se a figura 1 vê-se que quanto mais próximo os pontos (x_i, y_i) estiverem da reta regressora $(y = a + bx)$ mais preciso será o retorno obtido por esta expressão. O que o método dos mínimos quadrados ordinários faz, é permitir que pela minimização dos erros obtenham-se os estimadores (a) e (b) da expressão da reta, que nada mais seria do que reta de tendência para estimar (y) quando os resíduos estiverem minimizados.

Algebricamente esses estimadores para a regressão linear simples, que pode ser estendido para a regressão linear múltipla, são obtidos pelo sistema linear a seguir mostrado na figura 2.

$$\begin{cases} na + \left(\sum_{i=1}^n x_i \right) b = \sum_{i=1}^n y_i \\ \left(\sum_{i=1}^n x_i \right) a + \left(\sum_{i=1}^n x_i^2 \right) b = \sum_{i=1}^n x_i y_i \end{cases}$$

Figura 2 - Ajuste algébrico dos dados à reta de regressão simples pelo MQO

Alternativamente pode ser obtido por cálculos matriciais, sem maiores dificuldades, até mesmo com maior facilidade com o uso de softwares dedicados à elaboração de planilhas, por exemplo, o EXCEL, pela expressão genérica da figura 3.

$$\hat{\beta} = (X'X)^{-1} X'y = X^{-1}(X')^{-1} X'y = X^{-1}y.$$

Figura 3 – Ajuste matricial dos dados à reta de regressão simples pelo MQO

Observe-se que estes são os passos iniciais para obtenção dos resíduos mínimos proporcionados pelo MQO que permite determinar a estatística referente ao Coeficiente de Determinação, dado por (R^2) , indicador apenas qualitativo do poder de explicação do modelo estatístico, mas não suficiente por si só.

A partir desse ponto, a análise passa para o MCRL, método clássico da regressão linear que tem por objetivo, a partir do MMQO examinar mais detalhadamente os resíduos produzidos e testá-los convenientemente.

Normalidade dos resíduos

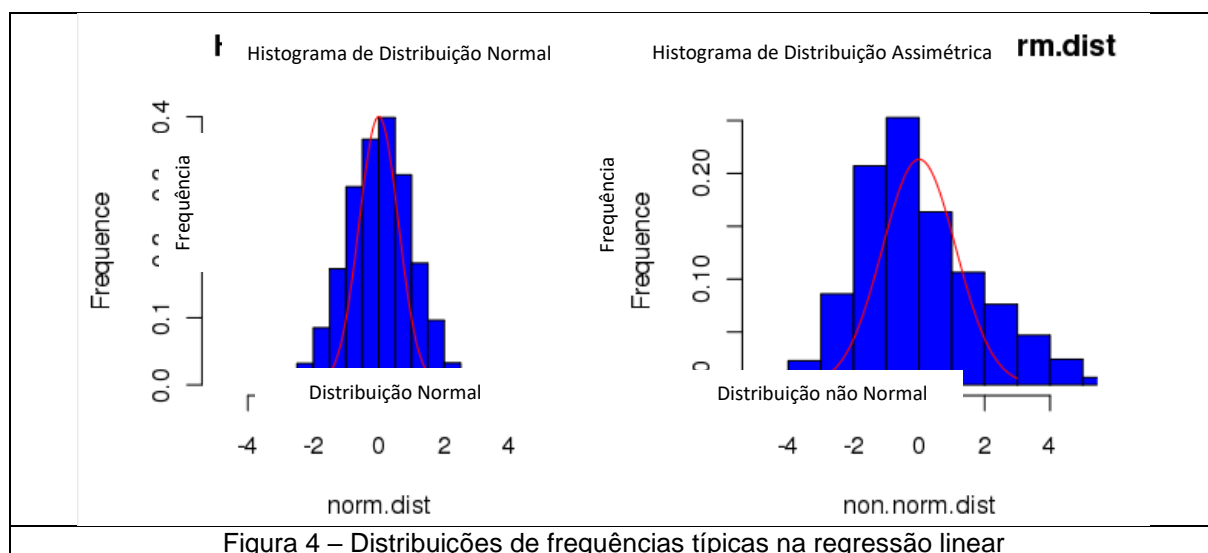
A normalidade dos resíduos é o pressuposto fundamental para a análise correta da regressão linear simples (RLS) ou múltipla (RLM) pelo Método dos Mínimos Quadrados Ordinários (MMQO).

Além de ser um pressuposto da própria RLS ou RLM, ele também é um pressuposto para a aplicação para a maioria dos testes formais em complemento às diversas verificações possíveis informais via gráficos de dispersão disponíveis nos softwares customizados, regularmente utilizados.

Este pressuposto reveste-se de importância pelo fato de, ocorrendo a situação em que os resíduos não sigam uma distribuição normal, irremediavelmente os intervalos de confiança da regressão não serão confiáveis e os valores-p dos coeficientes tendem a ser inexatos tornando-os ineficientes. São situações em que aparecem, inadvertidamente, os erros tipo I ou tipo II dos testes de hipóteses.

Na prática das avaliações o avaliador verifica a normalidade comparando os percentuais distribuídos entre ± 1 desvio padrão, ± 2 desvios padrão e ± 3 desvios padrão e aceita quando as variações entre os percentuais calculados e os percentuais 68%, 90% e 99% estão próximos até algo em torno de $\pm 5\%$ em cada intervalo.

Ocorre que mesmo isso acontecendo, nem sempre o histograma de frequências dos dados aparenta normalidade como se observa na figura seguinte em que esses histogramas apresentam a forma normal e não normal de frequências.



Na situação em que o histograma se mostra assimétrico, contrariando informalmente o pressuposto de normalidade dos resíduos, convém que se faça a análise formal da distribuição dos resíduos. Neste caso é necessário verificar se a assimetria compromete a normalidade dos resíduos a um nível de confiança, pelo menos aceitável, e que assim seja possível assumir essa tolerância no resultado da regressão e justificar o resultado.

Essa análise informal pode ser realizada utilizando-se a técnica do Papel (ou gráfico) de Probabilidade Normal em que é possível obter um conjunto de dados de entrada com duas series, uma com os resíduos em análise ordenados e outra com as probabilidades acumuladas padronizadas desses resíduos ordenados pela ordem de coleta. A figura 4 demonstra as apresentações desse gráfico em algumas situações.

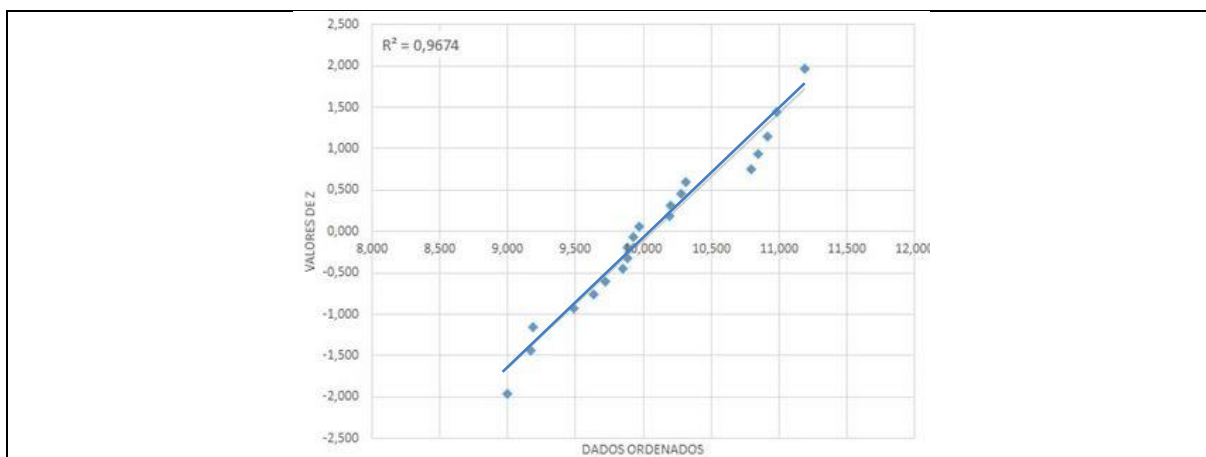
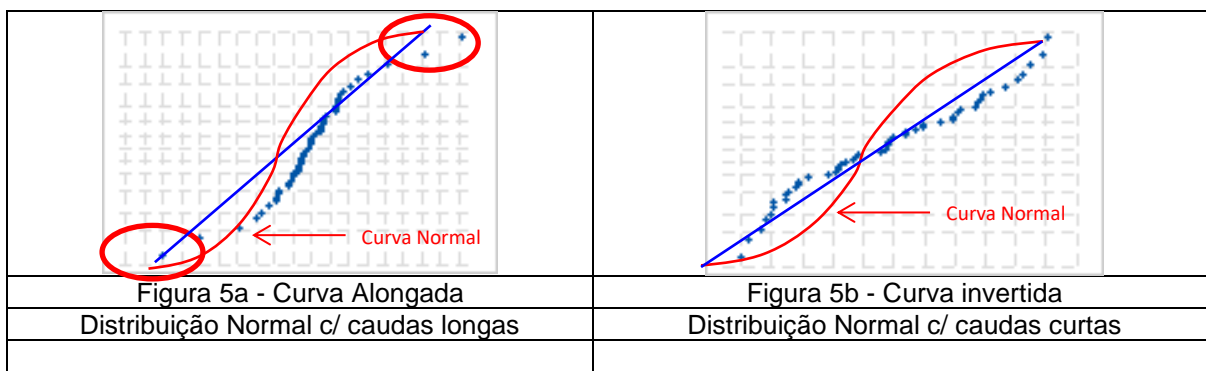


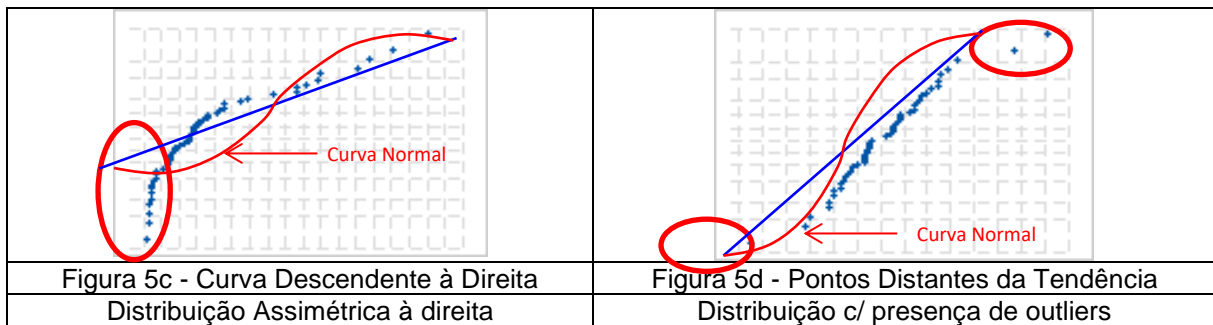
Figura 4 – Gráfico Papel de Probabilidade – Distribuição Normal

Não obstante, outros padrões podem se apresentar na análise e sugerir que mais investigações devem ser implementadas para a busca da normalidade dos resíduos, fundamental para a análise da regressão linear com sucesso.

Dois caminhos podem ser seguidos:

1 – Utilizando-se de um software de planilhas, por exemplo, o EXCEL, preparar um gráfico com as probabilidades acumuladas padronizadas (eixo y) x resíduos ordenados (eixo x) e a partir daí obter a reta de tendência dos resíduos plotados e o seu coeficiente de determinação R^2 . Esse coeficiente calculado estando bastante próximo de 1, significa que os dados analisados têm uma distribuição normal. A Figura 5 demonstra algumas configurações das violações da normalidade possíveis de ser identificadas pela análise informal dos Gráficos de Papel de Probabilidade:





Até este ponto, é possível saber se os resíduos são normalmente distribuídos qualitativamente, ou seja, o quanto os resíduos se aproximam da curva de distribuição normal padronizada, (t) ou (z) conforme o caso, o que é possível avaliando-se o Coeficiente de Determinação (R^2) da curva de tendência, mas não se sabe quanto quantitativamente e nem com que confiabilidade.

Isso, porque o Coeficiente de Determinação informa percentualmente quão próximo os resíduos estão perto da média, mas não quanto numericamente, o que é dado pelo erro padrão e pelo valor- p.

Assim, essa avaliação pode ser feita qualitativamente como segue, através do grau de relacionamento entre as variáveis independentes com a variável dependente, pelo exame do Coeficiente de Correlação da regressão, como no Quadro 1:

$R = 0$	Relação nula
$0 < R \leq 0,30$	Relação fraca
$0,30 < R \leq 0,60$	Relação média
$0,60 < R \leq 0,90$	Relação forte
$0,90 < R \leq 0,99$	Relação fortíssima
$R = 1$	Relação perfeita
Quadro 1 - Relacionamento das variáveis independentes com a variável dependente	

2 - Regredindo as Probabilidades Acumuladas Padronizadas (eixo y) x Resíduos Ordenados (eixo x), ainda usando o EXCEL, e analisando-se os resultados do valor-p dos coeficientes, define-se quantitativamente a confiabilidade da normalidade dos resíduos para uma significância especificada previamente, finalizando a análise.

Uma análise formal, alternativa à análise informal anterior, pode ser feita utilizando-se do software R / R-Studio.

Por esse caminho procede-se como indicado no quadro a seguir, aplicando-se os comandos seguintes finalizando com a conclusão e análise:

```

## Normalidade dos resíduos:

## Teste de Shapiro-Wilk

## Teste p-value para a distribuição normal dos resíduos
shapiro.test(mod$residuals)

## Nível de significância = 0,05

## Hipóteses

## H0: Distribuição dos resíduos = Normal => p > 0,05
## H0: Distribuição dos resíduos ? Normal => p = 0,05

## No Console do R-Studio deve aparecer a seguinte resposta:
>
≤

## data: mod$residuals
## w = 0,99555, p-value = 0,8269

## Análise e conclusão:

## Como p-value = 0,08269 (calculado) > p-value 0,05 (crítico)
## A distribuição dos resíduos é normal

```

Teste de Normalidade no R / R-Studio

Observa-se que neste exemplo didático o p-value calculado foi superior ao p-value crítico, sugerindo que há evidências que a distribuição dos resíduos siga uma distribuição normal.

Outliers e Pontos Influenciantes dos resíduos

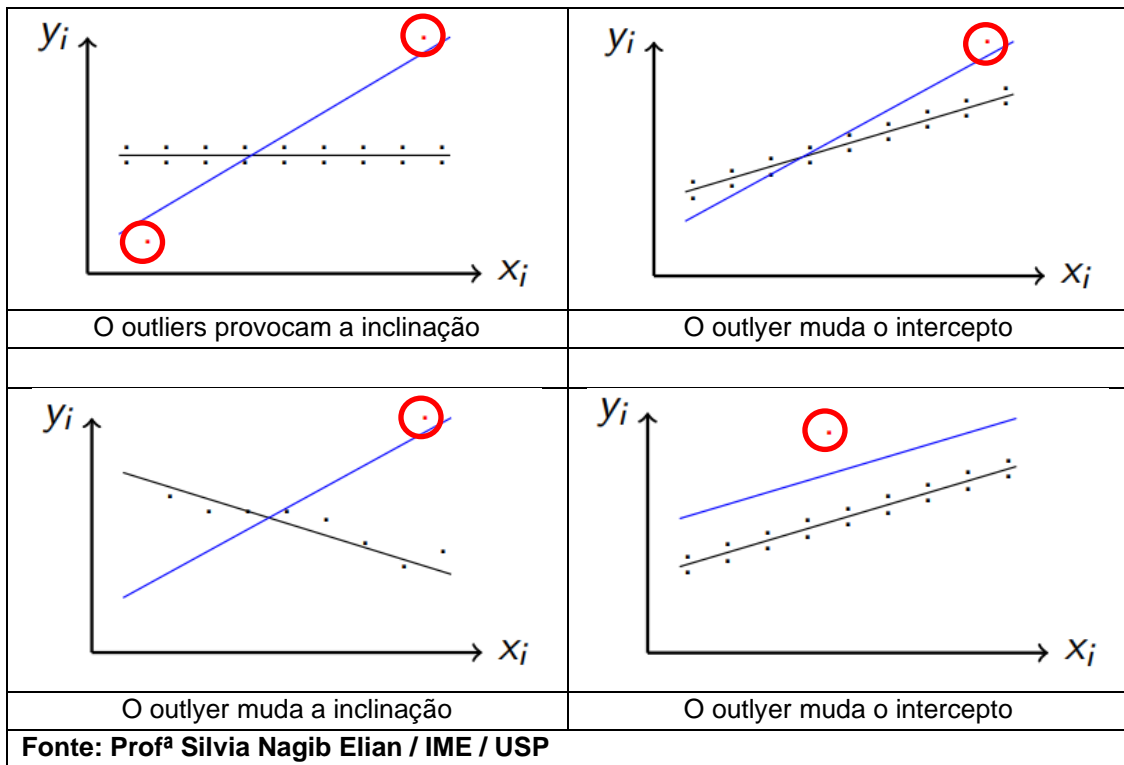
Os outliers, definidos na ABNT NBR 14.653-2 / 2011, item 3.48, são pontos atípicos e como afirma a norma, são estranhos à massa de dados formadora da amostra que pretende sintetizar o mercado colhida para análise e conclusão quanto ao valor de avaliação dentro de um intervalo de confiança.

Não menos danosos, são os pontos influenciantes, definidos na ABNT NBR 14.653-2 / 2011, item 3.56 também como pontos atípicos que quando retirados da amostra, alteram significativamente os parâmetros estimados ou a estrutura do modelo. A ABNT NBR 14.653-2 / 2011 não explicita o motivo, mas cabe aqui informar que alteram em favor da qualidade e melhora da precisão do modelo.

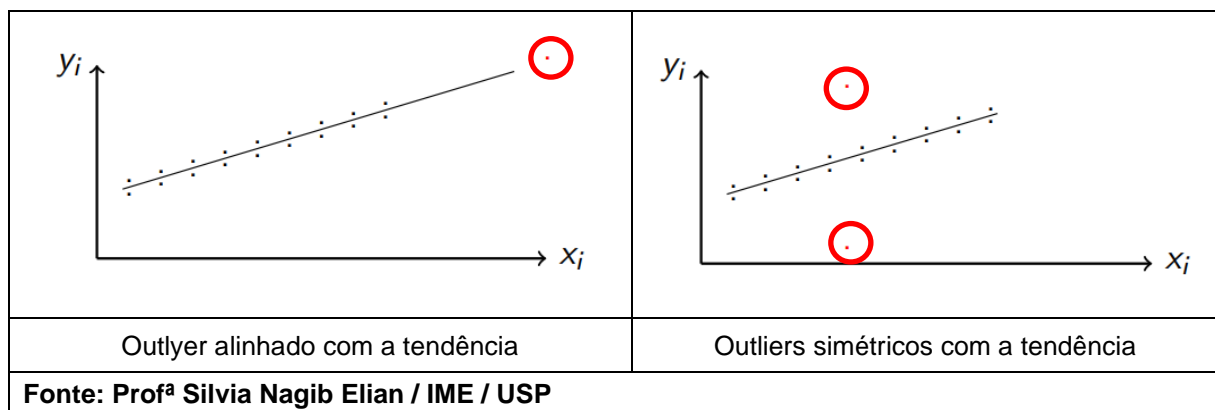
Os outliers, ainda que muitas vezes tolerados em modelos de regressão simples como os pontos influenciantes não devem ser admitidos conforme especifica a 14.653-2 / 2011 no item A.2 do Anexo 2.

Esses pontos comprometem seriamente os resultados estimados pela da regressão linear, seja, RLS ou RLM pelo Método dos Mínimos Quadrados (MQO).

As figuras com os gráficos dos resíduos a seguir demonstram graficamente como ocorre esse comprometimento pela mudança da posição da reta de tendência da regressão:



Não obstante, desde que cuidados adicionais sejam empreendidos existem pontos atípicos que não alteram a reta de regressão e após análise circunstanciada do modelo resultante podendo até ser tolerados, como demonstrado a seguir, desde que não se constituam em pontos influenciantes no resultado, normalmente identificados pela Distância de COOK



Uma análise formal, alternativa à análise informal anterior, pode ser feita utilizando-se do software R / R-Studio. Por esse caminho procede-se como indicado no quadro a seguir, aplicando-se os comandos seguintes finalizando com a conclusão e análise:

```
## Outliers nos resíduos:

## Cálculo dos resíduos padronizados

summary(rstandard(mod))

## No Console do R-Studio deve aparecer a seguinte resposta:

##   Min.      1º Qu.      Median      Mean      3º Qu.      Máx.
## -2,9693035 -0,7338720 -0,0079448 -0,0009045 -0,6220131  2,6209702

## Nesta tabela estão informados:
##   Os valores mínimo (Min.) e máximo (Máx.) dos resíduos padronizados.
##   Os primeiro (1ºQu.) e terceiro quartis (3ed Qu.) da distribuição padroniozada
##   Os valores da mediana (Median) e media (Mean) da distribuição padronizada

## Neste exemplo, existem resíduos padronizados acima de 2 desvios padrões
```

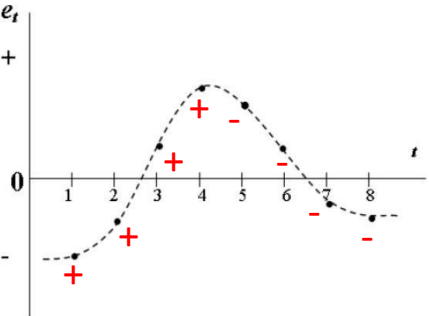
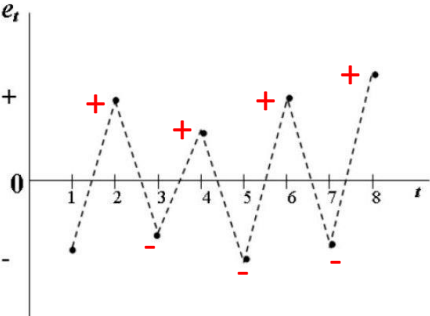
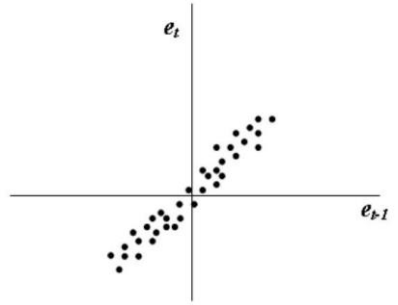
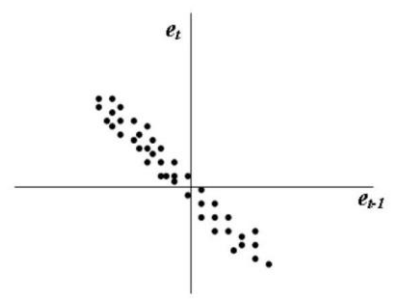
Resíduos padronizados no R-Studio

Observa-se que neste exemplo didático existem resíduos padronizados acima de 2 desvios padrões, que devem ser identificados e removidos para o perfeito enquadramento de sua distribuição normal conforme recomendação normativa.

Independência dos resíduos

A independência dos resíduos não é comentada na ABNT NBR 14.653-2 / 2011 e constitui-se em fator importante para a obtenção de modelos confiáveis.

Deve ser verificada sempre, pois que resíduos se auto correlacionando dentro da estrutura de dados da amostra não contribui para a normalidade dos mesmos, invalidando ou comprometendo seriamente a obtenção de resultados confiáveis.

	
<p style="text-align: center;">Sequências resíduos longas do tipo (++++-----)</p>	<p style="text-align: center;">Sequências resíduos alternadas do tipo (+ - - + - -)</p>
	
<p style="text-align: center;">Auto correlação positiva dos resíduos</p>	<p style="text-align: center;">Auto correlação negativa dos resíduos</p>

Fonte: Prof. Ilton G. Soares / FGV

Uma análise formal, alternativa à análise informal anterior, pode ser feita utilizando-se do software R / R-Studio.

Por esse caminho procede-se como indicado no quadro a seguir, aplicando-se os comandos seguintes finalizando com a conclusão e análise:

```
## Auto-correlação dos resíduos padronizados do modelo
durbinWatsonTest(mod)

## Hipóteses
## H0: Não há autocorrelação dos resíduos => p > 0,05
## H1: Há autocorrelação dos resíduos l => p = 0,05

## No Console do R-Studio deve aparecer a seguinte resposta:

## lag      Aucorrelacion D-W      Statistic      p-value
## 1         0,08871924         1,811357         0,218
## Alternative hipotesis: rho != 0

## a análise destes resultados consiste em:
##          1,5 > D-W < 2,5 => A estatística D-W neste intervalo sinaliza que
##                               não há autorcorrelação dos resíduos padronizados.
##          p-value > 0,05 => Esta estatística confirma a estatística de D-W, ou
##                               seja, não há autocorrelação dos resíduos

## Conclusão: Não havendo autorcorrelação, esta análise sinalizada que os resíduos
##                               são independentes entre si.
```

Independência dos resíduos no R-Studio

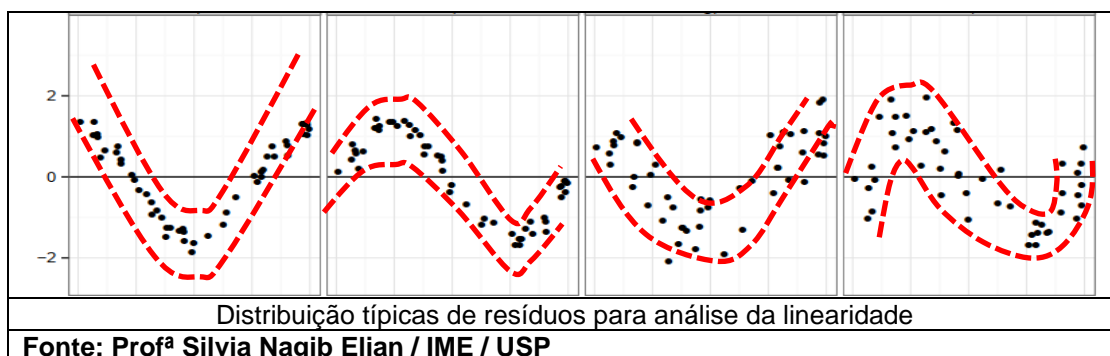
Em síntese, a violação do pressuposto da autocorrelação entre os resíduos examina a independência dos mesmos sendo um erro que resulta na falta de confiabilidade dos testes de significância e intervalos de confiança, quando não independentes, ainda que os coeficientes do modelo permaneçam não viesados ou não tendenciosos e consistentes.

Linearidade dos resíduos

A linearidade dos resíduos é um pressuposto importante e indispensável até mesmo pelos próprios objetivos dos Método RLS ou RLM, lineares por definição.

Quando esse pressuposto é violado, isso implica que a estimativa do erro aumentará, isto é, o intervalo de confiança será maior, porque nesta situação os valores observados na amostra não se aproximarão dos valores calculados ou preditos pela RLS ou RLM (local onde passará a reta).

Portanto trata-se de um pressuposto fundamental já que as equações de regressão RLS ou RLM são modelos lineares por definição.



A não linearidade dos resíduos sugere que os resíduos tendem a seguir um padrão diferente da distribuição linear representativa de uma distribuição aleatória.

Deve ser observada junto com a variância dos resíduos que devem se manter constante sem tendências ascendentes ou descendentes com as variáveis independentes.

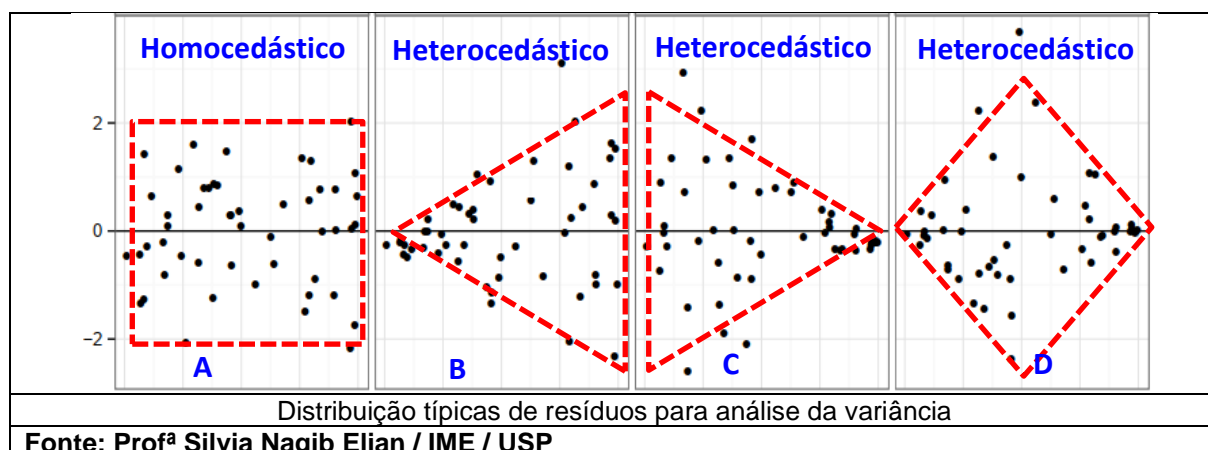
Homocedasticidade dos resíduos

A homocedasticidade é a consideração exigência para aceitação do modelo que a variância dos resíduos seja constante. Os resíduos não se comportando assim, são ditos heterocedásticos ou com variâncias variáveis.

Essa variância não constante pode ser atribuída à diversas fontes, todas elas já examinadas nos tópicos anteriores quais sejam, a presença da não normalidade dos resíduos, a existência de outliers e/ou pontos influenciantes nos resíduos, a não independência dos resíduos e a não linearidade dos resíduos.

Observe-se que não está listada a auto correlação das variáveis que será abordada a seguir e não influencia na variância do modelo.

Graficamente, por uma análise informal, os resíduos da equação de regressão RLS ou RLM pelo método MQO podem ser analisados conforme as distribuições seguintes e daí tirar as conclusões necessárias para aceitação ou não do modelo estatístico.



Nas distribuições acima fica notável que a variância dos resíduos somente é constante no caso A, sendo variável nos casos B, C e D.

Uma análise formal, alternativa à análise informal anterior, pode ser feita utilizando-se do software R / R-Studio.

Por esse caminho procede-se como indicado no quadro a seguir, aplicando-se os comandos seguintes finalizando com a conclusão e análise:

```
## Homocedasticidade (Breusch-Pagan):  
  
## Variância constante dos resíduos padronizados do modelo == 0,80  
  
## Cálculo da correlação entre as variáveis independentes:  
  
## Hipóteses  
  
## H0: Há homocedasticidade entre as variáveis independentes      => p > 0,05  
## H0: Não há homocedasticidade entre as variáveis independentes => p = 0,05  
  
bptest(mod)  
  
## No Console do R-Studio deve aparecer a seguinte resposta:  
  
##           Studentized Breusch-Pagan test  
## data: mod  
## BP = 1,9771    df = 2    p-value = 0,3771  
  
## Como p-value = 0,3771 (calculado) > p-value 0,05 (crítico)  
## Há homocedasticidade entre as variáveis independentes
```

Teste de Homocedasticidade no R / R-Studio

Em síntese, a violação do pressuposto da homocedasticidade é um erro muito grave e intolerável porque resulta que os testes de significância e intervalos de confiança obtidos resultam errados. Isso acontece em razão dos modelos obtidos pelo método dos Mínimos Quadrados Ordinários (MQO) perderem a propriedade de produzirem a melhor estimativa dos parâmetros populacionais. Em geral originam-se nas medidas dos atributos das variáveis independentes.

Multicolinearidade das variáveis independentes

A multicolinearidade das variáveis independentes é um fenômeno que surge quando essas variáveis se correlacionam mais fortemente entre si e não apenas à variável dependente para a qual pretende inferir uma resposta. mas também uns aos outros. Em resumo, constituem-se em variáveis que se comportam de forma redundante.

Para fins de constatação dessa redundância, admite-se que duas ou mais variáveis independentes são auto correlacionadas ou redundantes quando a correlação entre elas ultrapassa a 80%, visto que essa auto-correlação é inevitável em regressões lineares, sendo toleráveis até este limite, além do qual deve ser evitada.

As consequências da presença de multicolinearidade em regressões lineares reflete-se na estimação do intervalo de confiança com erros-padrão elevados no caso de multicolinearidade moderada ou severa e até mesmo a impossibilidade de qualquer

estimação se a multicolinearidade for perfeita com coeficientes de determinação iguais ou acima de 90%.

Analisando um exemplo disponível no site do professor Alexandre Gori com os recursos disponíveis no EXCEL obtém-se a seguinte conclusão:

Multicolinearidade das Variáveis Independentes

PIB	Pop	CO2
13,20	3,20	1,50
197,00	35,50	8,70
128,60	19,10	2,80
286,40	40,40	9,40
72,60	3,10	4,40
167,80	22,30	8,40
114,40	8,40	3,20
58,00	9,00	0,90

RESUMO DOS RESULTADOS

Estatística de regressão	
R múltiplo	0,882450661
R-Quadrado	0,778719168 (*) obs
R-Quad ajust	0,690206836
Erro padrão	1,905561105
Observações	8

Pela ANOVA o modelo é significativo

	gl	SQ	MQ	F	F de signif
Regressão	2	63,89	31,95	8,80	0,02
Resíduo	5	18,16	3,63		
Total	7	82,05			

CORRELAÇÃO

	PIB	Pop	CO2
PIB	1,00	0,94	0,94
Pop	0,94	1,00	0,84
CO2	0,88	0,84	1,00

Obs: Correlações superiores à 80% sinalizando a multicolinearidade entre variáveis

ANOVA

	Coefficientes	Erro padrão	Stat t	valor-P	Inferior 95,0%	Superior 95,0%
Interseção	0,47234	1,32844	0,35556	0,73668	-2,94254	3,88721
PIB	0,03048	0,02486	1,22567	0,27490	-0,03344	0,09439
Pop	0,02757	0,15044	0,18329	0,86177	-0,35915	0,41429

O intervalo dos coeficientes passa por ZERO.
Pelo exame dos coeficientes há indícios de multicolinearidade. Os coeficientes p com alpha > 5% não são confiáveis.

Modelo de regressão:
CO2 = 0,47234 + PIB x 0,03048 + Pop x 0,02757

(*) obs = O R2 elevado não garante a boa performance do modelo.

Fonte: Prof. Alexandre Gori / UNICAMP

Utilizando-se do mesmo exemplo, apenas ilustrativo desta análise, desenvolvido com o software customizado SisDea, obtém-se os mesmos resultados, para as correlações simples no SisDea, o que, a depender do analista, podem ser equivocadamente negligenciadas e conduzir a resultados equivocados, principalmente por não encontrar a disponibilidade imediata do valor da probabilidade p, o fator determinante para a aceitação dos coeficientes do modelo.

```
## Multicolinearidade
## Correlação máxima entre variáveis independentes == 0,80

Premissas

## Há colinearidade entre as variáveis independentes      => r > 0,80
## Não há colinearidade entre as variáveis independentes  => r = 0,80
## Complementarmente => VIF = 10,00 sinaliza que há multicolinearidade

pairs.panels(dados)
### Multicolinearidade: r > 0.9 (ou 0.8)

vif(mod)
### Multicolinearidade: VIF > 10

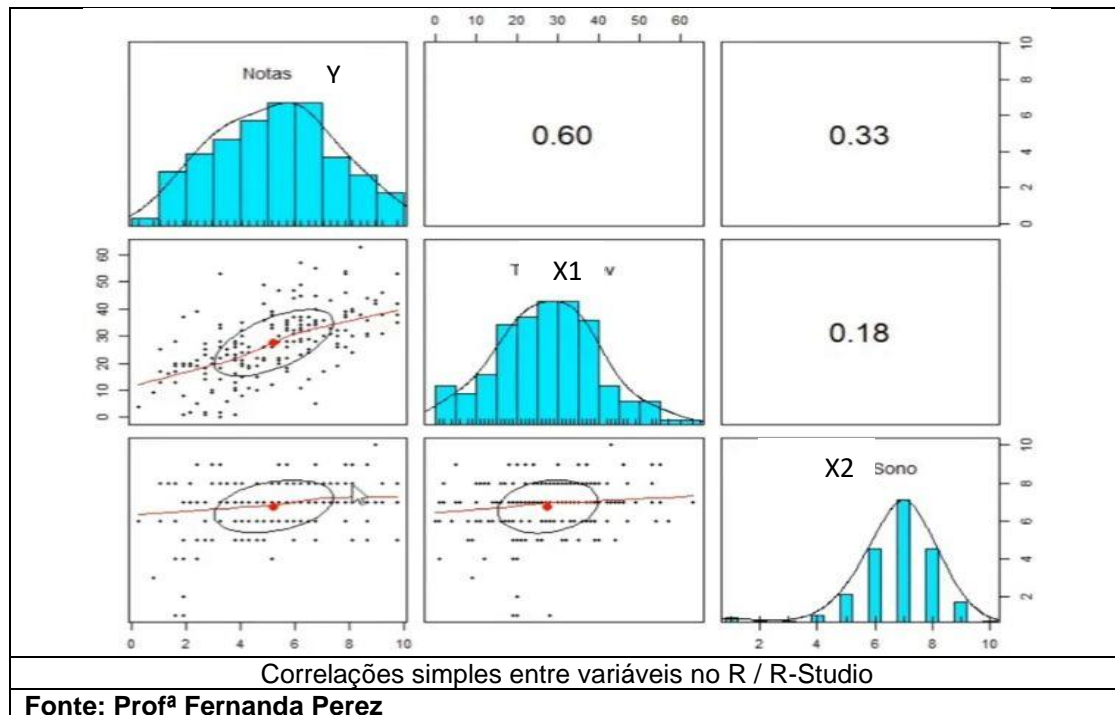
## No Console do R-Studio deve aparecer a seguinte resposta:
##   pairs.panels(dados)
##   vif(mod)

##           X1           X2
##   1,034234     1,934234

## Como os VIF entre as variáveis independentes calculados retornaram
VIF X1X2 = VIF X2X1 < VIF = 10, não há evidências que exista
multicolinearidade no modelo.
```

Teste de Multicolinearidade no R / R-Studio

Complementarmente, o software R disponibiliza a tabela de correlações onde se pode observar a correlação entre as variáveis independentes X1 e X2, igual a 0,18, como no quando a seguir, está abaixo de 0,80.



Em resumo, um modelo com a presença de variáveis independentes multicolineares mudará drasticamente se uma delas for removida, revelando-se, portanto, um modelo impreciso.

Considerações finais

A ABNT NBR 14.653-2 / 2011, textualmente no Anexo A2 alerta que “Ressalta-se a necessidade, quando se usam modelos de regressão, de observar os seus pressupostos básicos, apresentados a seguir, 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...”

Importante notar nesta nota da ABNT NBR 14.653-2 / 2011 é a referência taxativa a “... inexistência de pontos atípicos...”, que pelas definições da própria norma em seus itens 3.48 e 3.56 são intoleráveis no modelo de RLS ou RLM pelo Método dos Mínimos Quadrados.

Finalizando, o gráfico dos resíduos studentizados x quantis com as bandas do envelope dos erros padrões dos resíduos simulados para uma significância especificada, sempre que possível, deve ser examinado por possibilitar a identificação daqueles pontos da distribuição que eventualmente extrapolem esses limites.

Esse exame final permite a retirada de eventuais resíduos atípicos ou influenciantes, não identificados em outras análises.

Esse gráfico, tem um aspecto como o da figura a seguir:

