**INE 7001** - Procedimentos de Análise Bidimensional de variáveis QUANTITATIVAS utilizando o Microsoft Excel 2007.

Professor Marcelo Menezes Reis

O objetivo deste texto é apresentar os principais procedimentos de Análise Bidimensional de variáveis quantitativas, tal como apresentados em sala, mas utilizando a planilha eletrônica Excel. Os dados estão na planilha "Temperatura e vendas", do arquivo Bidimensional.xls, disponível nas páginas das disciplinas: contém as informações sobre 250 pares de observações temperatura (em graus Celsius) e quantidade vendida de refrigerantes.

Os procedimentos foram preparados utilizando a versão 2007 do Excel. Há algumas diferenças em relação às versões mais modernas (2010), mas a essência permanece a mesma.

## 1. Construção de diagrama de dispersão para as variáveis.

No presente caso, em que há apenas 2 variáveis, é possível construir um diagrama de dispersão, relacionando temperatura e vendas. O objetivo é avaliar a força, a direção e a forma de uma eventual correlação entre elas: com isso será possível avaliar qual modelo de regressão aplicar para prever os valores de uma variável em função dos da outra. Os dados de interesse estão mostrados na figura 1:

	A	В
1	Temperatura	Vendas
2	31.19	1321
3	31.28	1492
4	29.85	1495
5	32.41	1386
6	32.17	1672
7	31.87	1498
8	34.89	2702
9	30.84	1413
10	36.21	3252
11	31.36	1502
12	33.83	1937
13	33.45	2136
14	29.82	1480
15	33.27	2014

Na coluna A encontram-se os valores de Temperatura, e na coluna B os das Vendas. É preciso identificar corretamente qual variável é a independente e qual é a dependente: caso contrário o diagrama estará completamente errado, o modelo eventualmente ajustado também, e as decisões tomadas com base neles pouca validade terão. É razoável imaginar que a Temperatura possa influenciar as Vendas de refrigerante: maiores valores de Temperatura poderiam causar maiores valores de Vendas. Sendo assim, Temperatura será a variável independente, sendo então representada no eixo X, e Vendas a variável dependente, ocupando o eixo Y.

Passamos agora a construção do diagrama de dispersão propriamente dito. Recomenda-se colocar o cursor em uma célula vazia da planilha, para evitar que o Excel selecione automaticamente dados que não sejam do nosso interesse.

## Figura 1 - Temperatura e vendas

Em seguida, no menu Inserir, procure por Gráficos, depois por Dispersão, e depois selecione "Dispersão Somente com Marcadores", como na Figura 2:

	<b>.</b> • •	P × 🛕 🗋	) 📂 🖨 🎔   🚾	∋ Q↓ X↓ 🕻	<u>∎ ⊠</u> ) =			Exer	mplobidimen	isional - N
	Início	Inserir	Layout da Página	Fórmulas	Dados	Revisão	Exibiçã	ão Desenvolved	or Suple	mentos
Ĩ	•				/X ٵ			🔅 🜔		Α
Tab Dinâr	ela Tabela nica *	Imagem	Clip-art Formas Sm	artArt Colunas	Linhas Pizza	a Barras	Área [	Dispersão Gráficos *	Hiperlink	Caixa de Texto
	Tabelas		Ilustrações			Gráfico	Dispersão			
	E1	- (	f <sub>x</sub>					· · ·	bd	
	A		В	С	D		E			
1	Tempe	ratura	Vendas					Dispersão Some	nte com Mar	cadores
2		31.19	1321			_		Comparar pare	es de valores.	
-		24.20	1400					valores não es	tiverem em or	rdem
3		51,28	1492				Ļ	no eixo X ou q	uando medidas	
4		29,85	1495					separadas.		

Figura 2 - Menu Inserir - Gráficos - Dispersão - Dispersão somente com Marcadores

Após pressionar "Dispersão Somente com Marcadores" surgirá um gráfico em branco. Precisamos, então, entrar com os dados. Basta selecionar o gráfico e pressionar o botão direito do mouse para surgirem as várias opções disponíveis, como na Figura 3:

Calit	ri (( + 10 → A* A* 💁 + 🗹 + 🟈 Z 📰 🗐 A + 🗄 😤 🗄 +
×	Recortar
1 1 1 1	<u>C</u> olar Redefinir para <u>C</u> oincidir Estilo
A nu B	Eonte Alterar Tipo de Gráfico Selecionar Dados
	<u>Mover Gráfico</u> <u>R</u> otação 3D
	<u>Ag</u> rupar ► Trazer para a F <u>r</u> ente ► Enviar para <u>T</u> rás ►
	At <u>r</u> ibuir macro <u>F</u> ormatar Área do Gráfico

Figura 3 - Opções do gráfico de dispersão

A opção que nos interessa aqui é "Selecionar Dados". Pressionando-a, chegamos à Figura 4:

Selecionar Fonte de Dados	
Intervalo de dados do gráfico:	=Plan11\$E\$1
Entradas de Leoenda (Série)	Róhulos do Eivo Horizontal (Cateopriae)
Adicionar Editar	Remover         Remover <t< th=""></t<>
Células Ocultas e Vazias	OK Cancelar

No campo "Intervalo de dados do gráfico" está a célula vazia que selecionamos no início. Para podermos entrar com dados precisamos adicionar uma nova série de dados: pressionando "Adicionar" surgirá a **Erro! Fonte de referência não encontrada.** Basta selecionar as células onde estão os dados de X, no nosso caso as células A2 a A251, e as de Y (B2 a B251). AVISO IMPORTANTE: selecione as células na planilha, não digite os nomes, não sei por que, não funciona... Veja a



Editar Série	? <mark>×</mark>
Nome da série:	
Valores de <u>X</u> da série:	Selecione um Intervalo
Valana da V da africa	Selecione um Intervalo
={1}	= 1
	OK Cancelar

Figura 5 - Caixa de seleção de dados

Nome da série:	
	Selecione um Interval
Valores de <u>X</u> da série:	
=Plan1!\$A\$2:\$A\$251	= 31, 19; 31, 28;
Valores de <u>Y</u> da série:	
=Plan1!\$8\$2:\$8\$251	= 1321; 1492; 14



Pressionando OK volta-se à Figura 4 e o resultado é o gráfico da Figura 7.



Reis

Figura 7 - Diagrama de dispersão de Temperatura x Vendas de refrigerante - 1ª tentativa

Obviamente o gráfico da Figura 7 não está pronto: aparece uma legenda desnecessária (Série 1, que apenas faria sentido se fossemos acrescentar mais outros conjuntos de dados ao mesmo gráfico, o que não é o caso), não há informação sobre os nomes das variáveis (e nem título do gráfico), e a escala horizontal não permite uma adequada visualização dos pontos.

Para remover a legenda basta selecioná-la com o mouse e pressionar "Del". Os outros aspectos exigem a seleção do gráfico. Ao fazer isso o Excel habilita as "Ferramentas de gráfico" que incluem "Design", "Layout" e "Formatar", conforme visto na Figura 8.

	) 🖬 🤊 - Cl - I	4 🗋 🞽 🖨 爷 🖻	😔 21 KI 🛍	• 🔊 • 📲	Exemplobidin	nensional - Micros	oft Excel	Ferramentas d	e Gráfico				
9	Início Inser	ir Layout da Página	Fórmulas	Dados Revisã	) Exibição	Desenvolvedor	Suplementos	Design	Layout Format	ar			
Altera de Gr	r Tipo Salvar como ráfico Modelo	Alternar Seleciona Linha/Coluna Dados						^	^		^	•	•
	Tipo	Dados	Lay	out de Gráfico				Es	tilos de Gráfico				
	Gráfico 1	$ f_x$											
	А	В	С	D	E	F	G	Н	1	J	K	L	

Figura 8 - Ferramentas de Gráfico - Design

É possível alterar o tipo de gráfico e fazer várias alterações no estilo (mudando as cores dos pontos e o fundo). Não vamos mudar nada neste menu. Selecione "Layout", para chegar na Figura 9.

	🎽 🖨 🌱 I 🔂 😔 💵	- 🔁 🔣 ) 🔹 👘 👘	Exemplobidimensional - Mi	crosoft Excel	Ferramentas de G	Fráfico				
Início Inserir Li	ayout da Página Fórmula	s Dados Revisão	Exibição Desenvolve	edor Suplemento	os Design Lay	rout Formatar				
Área do Gráfico 👻					do.				III.	Nome do Gráfico:
🥸 Formatar Seleção				S	U.U.U.U.U.U.U.U.U.U.U.U.U.U.U.U.U.U.U.					Gráfico 1
🚰 Redefinir para Coincidir Estilo	Imagem Formas Caixa de Texto	Gráfico * dos Eixos *	Dados      Dados      Dados	Eixos Linhas de Grade *	Area de Parede do Plotagem * Gráfico *	Base do Rotação Gráfico * 3D	Linha de Tendência *	<ul> <li>Linhas Barr</li> <li>Superiores/I</li> </ul>	ras Barras de Inferiores * Erros *	
Seleção Atual	Inserir	Rót	tulos	Eixos	Plano de l	Fundo		Análise		Propriedades
Gráfico 1 🔹 💿	$f_{\mathbf{x}}$									
A	B C	D	E F	G	Н	I	J	К	L	Μ
·			1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1		1.27					

Figura 9 - Ferramentas de Gráfico - Layout

Neste menu podemos acrescentar o título do gráfico, os títulos dos eixos, e a própria formatação dos eixos (incluindo as suas escalas). Selecionando "Título do Gráfico" podemos escolher seu posicionamento, Acima do Gráfico (Figura 10). Depois basta selecionar o título e mudá-lo, resultando na Figura 11.

	) 🛃 🤊 - ୧୯ - 🛕 🗋	) 📁 🛱 🖤 🔽	5 😔 2↓ X↓	H 🔣 )	Ŧ	Exemplobidimensio		
9	Início Inserir	Layout da Página	Fórmulas	s Dad	os Revisã	ão Exibição D		
EixoH	orizontal (Valor)		A		l da.			
🏼 🥎 Fo	ormatar Seleção							
ar Re	edefinir para Coincidir Est	ilo Imagem Forn	nas Caixa de Texto	Título do Títulos Legenda Rótulos do Gráfico y dos Eixos y Dados y				
	Seleção Atual	Ins	erir		Nenhum			
	Gráfico 1 🛛 👻 🤇	$f_x$		Não exibir um título de gráfico				
	А	В	С		Título Sobre Sobrepor títe	posto Centralizado ulo centralizado no		
1	Temperatura	Vendas			gráfico sem i	redimensionar gráfico		
-	remperatara	Vendus			Acima do Gr	áfico		
2	31,19	1321		dn	Exibir título a gráfico e red	acima da área do limensionar gráfico		
3	31,28	1492		Mais Opções de Título				

Figura 10 - Posicionamento do título do gráfico



Figura 11 - Diagrama de dispersão de Temperatura x Vendas de refrigerante - 2ª tentativa

Para os nomes dos eixos basta selecionar "Título do Eixo Horizontal Principal", que será Temperatura em graus Celsius (variável X, independente, que PODE influenciar a outra), e "Título do Eixo Vertical Principal", que será Vendas em R\$ 1000 (variável Y, dependente). Ver Figura 12.

<b>C</b>	) 🚽 🍠 🔹 (	원 🔨 🚨 🗋	) 🗁 🖨 💙 🛛	🖻 😔 24 X4	后 🔝 ) 🔻	E	xemplobidim	nensional - Mi	icrosof	ft Excel	Ferram	ientas de Grá	fico	
	Início	Inserir	Layout da Página	a Fórmulas	Dados	Revisão	Exibição	Desenvolve	edor	Suplement	os Desig	jn Layou	ut Forma	atar
Área o 🏷 Fo	do Gráfico ormatar Seleçã edefinir para C	o Coincidir Est	Imagem For	mas Caixa	Título do Tí Gráfico * dos	ítulos s Eixos •	da Rótulos o Dados v	de Tabela de	Eixo	s Linhas de Grade *	Área de Plotagem •	Parede do B Gráfico T G	ase do Rota ráfico • 30	) Ição D
	Seleção A	tual	In	serir	<b></b>	] Título do E	ixo <u>H</u> orizonta	l Principal 🔸		Eixos	-	Plano de Fu	ndo	
	Gráfico 1	• (	$f_x$		Ith	Título do E	ixo <u>V</u> ertical Pr	incipal 🔸	Ida	Nenhum		-1		
	A		В	С	D	)	E	F	Ľ		r um titulo de	eixo		
1	Temper	ratura	Vendas		*				0 <sub>dh</sub>	Título Gira Exibir títu	ado lo do eixo gira	ado e redime	nsionar	F
2		31,19	1321			Ve	ndas	de re		Título Ver	tical			h
3		31,28	1492			•••				Exibir títu redimensi	lo do eixo con onar gráfico	n texto vertica	ale	Г
4		29,85	1495		4000				-da	Título Ho Exibir títu	<b>rizontal</b> lo do eixo hor	izontalmente	e	E
5		32,41	1386	;	3500					redimensi	onar gráfico de Título de F	ivo Vertical P	rincinal	Ŀ
										mais opções	de intuio de c	ixo vertical P	mapain	

Figura 12 - Seleção dos títulos dos eixos principais

No eixo vertical escolhemos a posição "Título Girado" e no horizontal "Título Abaixo do Eixo". Novamente, precisamos selecionar e mudar os títulos, resultando na Figura 13. É crucial que haja referência à escala das variáveis.



Figura 13 - Diagrama de dispersão de Temperatura x Vendas de refrigerante - 3ª tentativa

Agora precisamos modificar apenas a escala do eixo horizontal, que está variando de 0 a 40 graus Celsius. Mas, o menor valor de temperatura está acima de 25 graus Celsius, portanto devemos mudar o mínimo da escala para começar em 25, mantendo o máximo em 40, e podemos também modificar o passo da escala de 5 para 1 grau. Veja a Figura 14.

<b>C</b>	) 🖬 🤊 - Ci - 🛕 🗋	) 🎽 🖨 💙   🖻		🗄 🔣 ) 🕫	Exemplobidir	nensional - Mic	rosoft	Excel	Ferrar	nentas de	Gráfico			
	Início Inserir	Layout da Página	Fórmulas	Dados Rev	visão Exibição	Desenvolved	dor	Suplement	os Desi	gn 🗍 La	ayout	Formatar		
Área 🍌 F 欲 R	do Gráfico ormatar Seleção edefinir para Coincidir Est	Imagem Form	A Caixa de Texto	Título do Títulos Gráfico y dos Eixos	Legenda Rótulos	de Tabela de Dados *	Eixos	Linhas de Grade 🕶	Área de Plotagem *	Parede de Gráfico *	Base do Gráfico	Rotação 3D	Linha de Tendência	Linhas • • Sup
	Seleção Atual	Inse	rir		Rótulos		<u>a</u> ,	Eixo <u>H</u> orizon	tal Principal		Nenhu	um vibir oixo		
	Gráfico 1 🔫 🤇	f <sub>x</sub>						Eixo <u>V</u> ertical	Principal	<u>است</u>		abii eixo		
	А	В	С	D	E	F		G	H	t din	Mostra Exibir	ar Eixo Pad eixo com a	rão ordem e os ró	itulos
1	Temperatura	Vendas		1							padrão Mortr	) ar Eivo am l	Milbarar	
2	31,19	1321			Vendas	de re	fri	igera	nte r	) 🏙	Exibir em mil	eixo com nu hares	úmeros repre	entados
3	31,28	1492						0		300000	Mostr	ar Eixo em	Milhões	antadas
4	29.85	1495		400	0						em mil	hões	meros repre	entauos
5	32,41	1386		350	0						Mostra Exibir em bil	ar Eixo em I eixo com nu hões	Bilhões úmeros repre	entados
6	32,17	1672		ຣ <sup>300</sup>	0					Log	Mostr	ar Eixo com eixo usando	<b>Escala Loga</b> o uma escala	<b>ítmica</b> Iogarítmica
7	31,87	1498		<b>8</b> 250	0						de bas	e 10		-
0	24.00	2702		- <del></del>	-						<u>M</u> ais Opço	oes de Eixo	Horizontal Pi	incipal

Figura 14 - Opções do Eixo Horizontal Principal

Selecionando "Mais Opções do Eixo Horizontal Principal" chegamos à Figura 15.



Os dois aspectos mais importantes são "Opções de Eixo", que incluem a escala, e "Número", que permite mudar o formato dos dados, incluindo o número de casas decimais.

Em "Opções de Eixo" vamos mudar o mínimo de "Automático" para "Fixo", com o valor mudado para 25. E "Unidade principal" de "Automático" para "Fixo", com o valor mudado para 1. O resultado pode ser visto na Figura 16.

Figura 15 - Opções de formatação de eixo



Figura 16 - Diagrama de dispersão de Temperatura x Vendas de refrigerante - final

Agora podemos fazer uma análise do diagrama de dispersão:

- as variáveis parecem estar fortemente correlacionadas, porque os pontos encontram-se bastante próximos.

- a correlação entre elas parece ser positiva, pois se observa que a nuvem de pontos tem um comportamento crescente, ou seja, maiores valores de temperatura, maiores valores de vendas (e é razoável imaginar que realmente um aumento na temperatura cause um aumento nas vendas).

- quanto à forma do relacionamento, isto é, que tipo de curva poderia ser ajustada aos dados para realização de previsões, talvez seja interessante pensar em um polinômio de segundo grau, ou uma exponencial; a utilização de uma reta talvez não seja uma boa ideia.

Se colocarmos o mouse sobre o gráfico (na parte branca) e pressionarmos o botão esquerdo, teremos uma situação semelhante à mostrada na Figura 17.



Observe que ao selecionar o gráfico, as células que contém os dados de origem têm suas bordas coloridas, o que pode ser útil para avaliar se não houve erro ou falta de alguns valores.

Figura 17 - Diagrama de dispersão: gráfico e dados

## 2. Ajuste de uma tendência a um diagrama de dispersão.

Imagine que quiséssemos ajustar uma reta ao diagrama de dispersão mostrado na Figura 16, não obstante a análise feita. Como proceder? O Excel permite ajustar uma variedade de curvas aos dados mostrados em um diagrama de dispersão, e ainda calcula os coeficientes das equações das curvas, pelo método dos mínimos quadrados (ou seja, obtém os coeficientes minimizam a soma dos quadrados dos desvios entre os valores observados e os previstos por cada curva).

Para fazer o ajuste de qualquer curva, que no Excel significa adicionar uma linha de tendência, o primeiro passo é colocar o cursor sobre os pontos do gráfico e pressionar o botão esquerdo do mouse. Alguns pontos ficarão salientados, tal como mostrado na Figura 18.



Figura 18 - Seleção de pontos no gráfico



Figura 19 - Opções de modificação dos dados

Em seguida, mantendo o cursor sobre pontos, precisamos os pressionar botão direito do 0 mouse, e surgirão as opções possíveis para os dados, entre elas "Adicionar linha de tendência", tal como mostrado na Figura 19.

"Adicionar Se pressionarmos linha de tendência" na Figura 19 chegamos à tela mostrada na Figura 20. O tipo padrão de linha é a linear (reta), mas podemos selecionar outras. No nosso problema vamos manter a curva linear, mas queremos que o Excel exiba a equação e o valor de Rquadrado (coeficiente de determinação) no gráfico. Então, em "Opções" (Figura 21) selecionamos ambos. Pressionando "OK" gráfico 0 ficará como o da Figura 22.



Figura 20 - Tipos de curva Figura 21 - Opções para os tipos de curva



Observe no canto superior direito da figura a equação da reta, com um coeficiente angular positivo (reta crescente), e o coeficiente de determinação, que vale 0,8565. Este valor significa que cerca de 85.65% da variabilidade média das vendas pode ser explicada pela variabilidade média da temperatura, através do modelo de regressão.

Figura 22 - Diagrama de dispersão com reta

Embora o valor de  $R^2$  sugira que a reta é um bom modelo de regressão, devemos observar com cuidado o gráfico, e lembrar a análise feita na Figura 16. Realmente a reta passa "entre" a maioria dos pontos, mas talvez outra curva apresente um melhor ajuste aos dados (polinômio de segundo grau ou exponencial, conforme sugerido anteriormente). Para realmente saber se o modelo ajustado é bom precisamos analisar seus resíduos.

## 3. Análise de resíduos

Uma vez tendo construído o diagrama de dispersão para as duas variáveis, e adicionada a linha de tendência a ele, pode ser interessante realizar a análise dos resíduos do modelo. Se o modelo for apropriado os resíduos deverão ter um comportamento aleatório, sem nenhum padrão identificável, mostrando que a variação residual, que não pode ser explicada pelo modelo é realmente casual, e ele poderá ser utilizado para realizar previsões e seus resultados serão úteis na tomada de decisão. Se, porém, algum padrão for detectado nos resíduos a variância residual não é aleatória, o que significa que o modelo não está conseguindo "explicar" de maneira consistente o

relacionamento entre as variáveis, e, portanto, as previsões feitas pelo modelo são questionáveis. Isso pode acontecer mesmo que o  $R^2$  assuma um valor elevado. Sendo assim a análise de resíduos é indispensável para avaliar a adequação de qualquer modelo de regressão, sendo especialmente importante nos casos de regressão múltipla, onde muitas vezes não é possível plotar um gráfico dos dados.

Pensando nos dados de Vendas e Temperatura, estudados nos itens 1 e 2, que culminaram no gráfico mostrado na Figura 22, queremos analisar os resíduos do modelo linear (reta). O primeiro passo é calcular os valores de vendas previstos pelo modelo linear: na célula C2 da planilha inserimos a fórmula com a equação da reta obtida pelo Excel, tal como na Figura 23.

	SE	▼ X √	= =(255.17	7*A2)-6451.7				
	A	В	C	D	E			
1	Temperatura	Vendas	Y predito					
2	31.19	1321	=(255.17*A	2)-6451.7				
3	31.28	1492						
4	29.85	1495						
_								

Observe que a fórmula é construída em função da temperatura (cujo primeiro valor está na célula A2). Após digitar a fórmula e pressionar "Enter" (ou "Return", dependendo do computador), podemos colocar o cursor sobre a célula C2, selecionando-a.

Figura 23 - Fórmula de previsão de vendas (reta)

Para estender os cálculos a todos os valores de temperatura basta "arrastar" a fórmula até a última linha do arquivo. As previsões de vendas através do modelo linear estarão então completas.

Para calcular os resíduos devemos obter a diferença entre os valores observados de Vendas e os valores previstos através do modelo linear. A Figura 24 mostra isso.

	SE	<u>- × √</u>	=   =B2-C2		
	A	В	С	D	E
1	Temperatura	Vendas	Y predito	Resíduos	
2	31.19	1321	1507.052	=B2-C2	
3	31.28	1492	1530.018		
4	29.85	1495	1165.125		

Novamente, basta construir a fórmula para o primeiro valor e "arrastá-la" até a última linha para obter todos os resíduos do modelo.

Figura 24 - Cálculo dos resíduos

A obtenção dos resíduos é muito importante, mas dependendo da unidade das variáveis os resíduos poderão ser consideravelmente grandes em valores absolutos, embora em termos relativos sejam pequenos, ou o contrário. Podemos ter resíduos pequenos em termos absolutos, mas substancialmente grandes em termos relativos. Para que a análise seja feita objetivamente é preciso *padronizar* os resíduos: subtraí-los de sua média esperada (que deve ser igual a zero se o modelo for bom) e dividir pelo seu desvio padrão. O cálculo do desvio padrão dos resíduos está mostrado na Figura 25.

SE 💽 🗙 🗸 = =DESVPAD(D2:D251)					51)
	A	В	С	D	E
1	Temperatura	Vendas	Y predito	Resíduos	Desvio padrão dos resíduos
2	31.19	1321	1507.052	-186.052	=DESVPAD(D2:D251)
3	31.28	1492	1530.018	-38.0176	
4	29.85	1495	1165 125	329 8755	

Inserimos a fórmula do desvio padrão amostral, com os dados das células D2 a D251, que contêm os resíduos calculados anteriormente. O resultado está mostrado na Figura 26.

Figura 25 - Cálculo do desvio padrão dos resíduos

Para obter os resíduos padronizados basta dividir cada resíduo pelo desvio padrão. Para que não haja problemas ao "arrastar" a fórmula é preciso dar uma referência absoluta ao denominador da fórmula: acrescentar \$ antes da letra que designa a coluna e antes do número que designa linha, tal como na Figura 26.

	SE	<u>-</u> X √	= =D2/\$E	\$2		
	A	В	С	D	E	F
1	Temperatura	Vendas	Y predito	Resíduos	Desvio padrão dos resíduos	Resíduos padronizados
2	31.19	1321	1507.052	-186.052	207.4415413	=D2/\$E\$2
3	31.28	1492	1530.018	-38.0176		
4	29.85	1495	1165.125	329.8755		

Figura 26 - Cálculo dos desvios padronizados

Para obter todos os resíduos basta "arrastar" a fórmula até a última linha do arquivo.

Uma vez obtidos os resíduos padronizados podemos fazer a sua análise propriamente dita. Precisamos construir dois diagramas de dispersão dos resíduos: resíduos padronizados em função de X (Temperatura), e resíduos padronizados em função dos valores preditos. O procedimento é semelhante ao visto no item 1, mudando apenas os valores de X e de Y, escrevendo os títulos adequados, e modificando as escalas horizontal e vertical, se necessário. IMPORTANTE: A ESCALA VERTICAL DEVE SER SIMÉTRICA EM RELAÇÃO A ZERO; se, por exemplo, o Excel apresentar os resíduos de -2 a +5, devemos mudar a escala para que fique de -5 a +5. Ao fazer o diagrama dos resíduos padronizados da reta para os dados da Figura 26 o Excel apresentará uma escala vertical de -3 a +4, devemos mudá-la para de -4 a +4. No diagrama de resíduos padronizados em função da temperatura o Excel apresentará a escala de 0 a 25, tal como no diagrama de dispersão da Figura 13, exigindo a mudança. Os diagramas resultantes estão na Figura 27 e Figura 28.



Fazendo a análise dos resíduos mostrados na Figura 27.

1) Número de resíduos positivos é semelhante ao dos negativos.

2) As distâncias dos resíduos positivos a zero são maiores do que as dos negativos.

3) Há um padrão nos resíduos, parece uma parábola.

Figura 27 - Resíduos padronizados por temperatura - Modelo linear



Fazendo a análise dos resíduos mostrados na Figura 28.

1) Número de resíduos positivos é semelhante ao dos negativos.

2) As distâncias dos resíduos positivos a zero são maiores do que as dos negativos.

3) Há um padrão nos resíduos, parece uma parábola.

Juntando a análise dos dois diagramas chegamos à conclusão que o modelo linear NÃO é apropriado para o problema, pois seus resíduos não se comportam de forma aleatória.

**Figura 28 - Resíduos padronizados por valores previstos - Modelo linear** Sugerimos a utilização de outro modelo.

Repetindo o procedimento da Figura 18 à Figura 21, podemos escolher o modelo Polinômio do 2º grau. O resultado pode ser visto na Figura 29, superposto ao resultado da Figura 22.



Figura 29 - Diagrama de dispersão com reta e polinômio do 2º grau

Percebe-se que o coeficiente de determinação do polinômio de 2º grau é maior do que o da reta. E, também, o ajuste da curva do polinômio de 2º grau aos pontos é bem melhor. Provavelmente os resíduos serão melhores do que os da reta. Outros modelos poderiam ser ajustados, resultando na Figura 30.



Figura 30 - Diagrama de dispersão com cinco modelos de regressão

Todos os cinco modelos aplicáveis estão no gráfico da Figura 30: reta, polinômio de 2º grau, logarítmico, exponencial e potência. Mas, observe o formato dos coeficientes no modelo potência: está científico, 1,81E-04x<sup>4,62E+00</sup>. Isso significa 0,000181x<sup>4,62</sup>, que é o formato que devemos usar nas previsões. Às vezes o Excel automaticamente apresenta as equações de um modelo em formato

científico, e com um número insuficiente de casas decimais, o que pode prejudicar nossas previsões. Para mudar o formato e as casas decimais veja o procedimento a seguir.

Selecione a equação do modelo potência na Figura 30 e pressione o botão direito do mouse: surgirá a Figura 31: dentre as opções possíveis pressione "Formatar Rótulo de Linha de Tendência", resultando na Figura 32.





			-	- 77ES
Formatar Rótulo de I	Linha de Tendê	ncia		? ×
Númer				
Numero	Número			
Preenchimento	Categoria:	Casas decimais:	2	
Cor da Borda	Geral Número			
Estilos de Borda	Moeda			
Sombra	Data			
Formato 3D	Porcentagem			
Alinhamento	Fração Científico			
	Texto			
	Personalizado			
	Código de Forn	natação:		-
	0,00E+00		Adicionar	
	Para criar um f	ormato personaliza	do, preencha a caixa	Código de
	Vincular à f	onte		
				Fechar

Figura 32 - Formatação de rótulo de dados: Número

Às vezes o Excel apresenta os dados em formato científico, mas na categoria "Geral". Se quisermos que os números sejam apresentados da forma usual devemos escolher "Número" e escolher quantas casas decimais forem necessárias: no nosso caso, como o Excel usou E-04, deve-se escolher no mínimo 4, mas o ideal é um pouco mais para ganhar precisão nas previsões, 6, por exemplo. O resultado pode ser visto na Figura 33.



Figura 33 - Diagrama de dispersão com cinco modelos de regressão - modificado

Na Figura 23 fizemos a previsão usando o modelo de Reta, agora apresentaremos as previsões pelos outros modelos disponíveis:

DISTT 🔻		=23,039*A2^2-1220,1*A2+17			
	R		S	Т	
1	Ypred	Pol2	ResPol2	SPol2	F
2	=23,0	39* <mark>A</mark>	2^2-1220,1	* <mark>A2</mark> +170	74

Figura 34 - Modelo polinômio de 20 grau (para equação da Figura 38)

Na Figura 34 é possível observar que no lugar de X colocamos a primeira célula do intervalo que contém os valores de temperatura (célula A2). Observe que o ^ é o símbolo de potenciação no Excel (e no Calc também). Basta arrastar até a célula R251 para completar a previsão pelo modelo polinômio de 2º grau. O cálculo dos resíduos, desvio padrão dos resíduos e resíduos padronizados é análogo ao caso da reta (para este e para os próximos modelos).

D	ISTT 👻	=8021,5*LN(A2)-26071		
	V	W	Х	
1	YpredLN	ResLN	SLN	
2	=8021,5*L	N(A2)-2607	71	

Figura 35 - Modelo logarítmico (para equação da Figura 33)

Na Figura 35 é possível observar que no lugar de X colocamos a primeira célula do intervalo que contém os valores de temperatura (célula A2). Observe que LN() é uma função do Excel (e do Calc também) que permite calcular o logaritmo neperiano (com base igual a e, a constante de Neper, igual a 2, 71828...). Basta arrastar até a célula V251 para completar a previsão pelo modelo logarítmico.

DISTT 👻			=0,000181*A2^4,624102			
	Z		AA	AB		
1	YpredP	ot F	lesPot	Spot		
2	=0,000181*A2^4,624102					

Figura 36 - Modelo potência (para equação da Figura 33)

Na Figura 36 é possível observar que no lugar de X colocamos a primeira célula do intervalo que contém os valores de temperatura (célula A2). Observe que X (no caso o conteúdo da célula A2) é elevado (^) a 4,624102, que é expoente do modelo potência (ver Figura 33). Basta arrastar até a célula Z251 para completar a previsão pelo modelo potência.

DISTT 👻		-	=15,408*EXP(0,1458*A2)		
	AD		AE	AF	
1	YpredExp		ResExp	Sexp	
2	=15,4	*A2)			

Figura 37 - Modelo exponencial (para equação da Figura 38)

Na Figura 37 é possível observar que no lugar de X colocamos a primeira célula do intervalo que contém os valores de temperatura (célula A2). Observe que EXP() é uma função do Excel (e do Calc também) que permite calcular o valor da constante de Neper (e = 2, 71828...) elevada ao produto de 0,1458 pelo conteúdo da célula A2). Basta arrastar até a célula AD251 para completar a previsão pelo modelo exponencial.

Vejam os resíduos padronizados do modelo polinômio do 2º grau:



Fazendo a análise dos resíduos mostrados na Figura 38.

1) Número de resíduos positivos é semelhante ao dos negativos.

2) As distâncias dos resíduos positivos e negativos a zero são semelhantes.

3) Os resíduos distribuem-se aleatoriamente, sem padrão.

Figura 38 - Resíduos do polinômio de 2º grau por temperatura





Fazendo a análise dos resíduos mostrados na Figura 39.

1) Número de resíduos positivos é semelhante ao dos negativos.

2) As distâncias dos resíduos positivos e negativos a zero são semelhantes.

3) Os resíduos distribuem-se aleatoriamente, sem padrão.

Juntando a análise dos dois diagramas chegamos à conclusão que o modelo de polinômio de 2º grau é apropriado para o problema, pois seus resíduos se comportam de forma aleatória.