

UNIVERSIDADE FEDERAL DE PERNAMBUCO
CURSO DE GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO

Análise de regressão aplicada à estimação dos tempos de fabricação
de painéis elétricos

Daniel Cavalcanti de Souza
Orientador: Prof. Luciano Lins

RECIFE, Junho/2009



UNIVERSIDADE FEDERAL DE PERNAMBUCO

DEPARTAMENTO DE ENGENHARIA DE PRODUÇÃO

CURSO DE GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO

**ANÁLISE DE REGRESSÃO APLICADA À ESTIMAÇÃO DOS
TEMPOS DE FABRICAÇÃO DE PAINÉIS ELÉTRICOS**

TRABALHO DE CONCLUSÃO DE CURSO DE GRADUAÇÃO

POR

DANIEL CAVALCANTI DE SOUZA

Orientador: Prof. Luciano Lins

RECIFE, JUNHO / 2009

S731a

Souza, Daniel Cavalcanti de.

Análise de regressão aplicada à estimação dos tempos de fabricação de painéis elétricos / Daniel Cavalcanti de Souza. - Recife: O Autor, 2009.

vi, 52 folhas, il : grafs., tabs., figs.

TCC (Graduação) – Universidade Federal de Pernambuco. CTG. Curso de Graduação em Engenharia de Produção, 2009.

Inclui bibliografia e Apêndice.

1. Engenharia de Produção. 2. Análise de Regressão. 3. Painéis Elétricos. 4. Tempo de Fabricação. I. Título.

UFPE

658.5

CDD (22. ed.)

BCTG/2009-113

AGRADECIMENTOS

Em cada etapa de nossas vidas, pessoas importantes dão contribuição ao nosso desenvolvimento pessoal, intelectual e profissional. Muitas vezes, nos distanciamos de algumas destas pessoas pelo efeito natural do cotidiano. Porém, mesmo com esta distância que há de ocorrer, vou sempre lembrar e me sentir agradecido por terem certas pessoas, passado por minha trajetória acadêmica.

Agradeço especialmente a meu orientador neste trabalho, professor Luciano Lins, que até os últimos instantes do desenvolvimento desta monografia motivou-me e transmitiu-me conhecimento.

Agradeço aos meus pais por me motivarem em todas as fases da minha vida.

Agradeço a minha namorada Clarissa por ter dado apoio e segurança durante este trabalho.

Agradeço a Deus, por possibilitar mais uma vez o êxito em meus trabalhos e estar presente em minha vida.

RESUMO

A estimação de tempos de manufatura pode ser tarefa difícil mediante critérios subjetivos, a depender, por exemplo, da força utilizada no processo, se são máquinas ou pessoas. Quando são pessoas, a demanda de tempo costuma variar bastante, o mesmo ocorrendo quando se trata de produtos customizados, sendo por vezes de grande valor fazer aplicações de ferramentas estatísticas para fazer estimativas. O presente trabalho objetiva aplicar a análise de regressão linear para melhorar a qualidade de estimativas de tempos de fabricação de painéis elétricos em uma empresa do ramo de energias renováveis. O trabalho apresenta a forma como ocorrem os processos de manufatura de painéis elétricos. Uma amostra de dados do processo de fabricação foi colhida ao longo do desenvolvimento do trabalho. Estes dados são apresentados graficamente de forma a fornecer melhor compreensão e análise, e em cima desta base foram aplicados os conceitos da análise de regressão para construir um modelo matemático estimador de tempos e da análise de variância para medir a qualidade do modelo desenvolvido. Estudo posterior pode revelar outro modelo, mais ajustado. Porém, para os dados observados, o trabalho traz um modelo adequado e com desempenho bastante razoável para se realizar estimativas de tempo de manufatura de painéis elétricos estudada.

Palavras-chave: Análise de Regressão, Painéis Elétricos, Tempo de Fabricação.

LISTA DE FIGURAS

Figura 2.1: Diagrama de dispersão.....	7
Figura 2.2: reta de regressão com inclinação.....	8
Figura 2.3: representação gráfica das variabilidades da variável resposta.....	9
Figura 2.4: bom padrão de resíduos	14
Figura 2.5: variância não constante de resíduos.....	14
Figura 2.6: exemplo de erros de modelo inadequado	15
Figura 3.1: fluxograma de produção de painéis.....	22
Figura 4.1: Distribuição da variável lead time	29
Figura 4.2: gráfico de probabilidade normal para lead time	30
Figura 4.3: distribuição da variável raiz de lead time.....	30
Figura 4.4: gráfico de probabilidade normal para raiz de lead time.....	31
Figura 4.5: distribuição da variável homem-hora	31
Figura 4.6: gráfico de probabilidade normal para homem-hora	32
Figura 4.7: distribuição da variável ln(homem-hora).....	32
Figura 4.8: gráfico de probabilidade normal para ln(homem-hora)	33
Figura 4.9: histograma da variável tipo.....	34
Figura 4.10: histograma da variável cubículos	34
Figura 4.11: histograma da variável cubículo01.....	35
Figura 4.12: histograma da variável cubículos	35
Figura 4.13: histograma da variável barramentos	36
Figura 4.14: histograma da variável automação	36
Figura 4.15: histograma da variável pontos de ligação.....	37
Figura 4.16: histograma da variável tensão	37
Figura 4.17: histograma da variável terceiros.....	38
Figura 4.18: <i>box-plot</i> de ln(homem-hora) versus barramentos.....	39
Figura 4.19: <i>box-plot</i> de ln(homem-hora) versus automação	39
Figura 4.20: <i>box-plot</i> de ln(homem-hora) versus tensão.....	40
Figura 4.21: <i>box-plot</i> de ln(homem-hora) versus terceiros	40
Figura 4.22: <i>box-plot</i> de ln(homem-hora) versus tipo.....	42
Figura 4.23: <i>box-plot</i> de ln(homem-hora) versus tipo, com outliers e pontos extremos.....	42
Figura 4.24: Normalidade dos resíduos	44
Figura 4.25: Resíduos x valores preditos.....	44

LISTA DE TABELAS

Tabela 3.1: agrupamento de tipos de painéis semelhantes e codificação da variável tipo.....	28
Tabela 4.1: Resultados do modelo	41
Tabela 4.2: resultados do modelo sem outliers e extremos da variável tipo	43

SUMÁRIO

1	INTRODUÇÃO	1
1.1	Problemática	1
1.2	Justificativa e relevância.....	2
1.3	Objetivos	3
1.3.1	Objetivo Geral	3
1.3.2	Objetivos Específicos.....	3
1.4	Metodologia	3
1.5	Organização do Trabalho	4
2	ANÁLISE DE REGRESSÃO.....	5
2.1	Introdução	5
2.2	Regressão linear simples	6
2.2.1	Modelo geral de Regressão Linear Simples	7
2.2.2	Coefficiente de Correlação	8
2.2.3	Coefficiente de determinação	9
2.2.4	Estimadores de dispersão	10
2.2.5	Suposições do modelo	11
2.2.6	Testes de significância.....	11
2.2.7	Análise de resíduos	13
2.3	Regressão linear múltipla.....	15
2.3.1	A estimação dos parâmetros do modelo	16
2.3.2	Coefficiente de determinação múltiplo	16
2.3.3	Coefficiente de determinação múltiplo ajustado.....	16
2.3.4	Suposições do modelo	17
2.3.5	Testes de significância.....	17
2.3.6	Multicolinearidade	19
2.3.7	Variáveis qualitativas independentes	19
3	MANUFATURA DE PAINÉIS ELÉTRICOS.....	21
3.1	A empresa	21
3.1.1	Contexto de mercado.....	21
3.1.2	O processo de manufatura de painéis	21
4	RESULTADOS	27
4.1	Levantamento das variáveis.....	27

4.2	Elaboração do modelo	28
4.2.1	Testes de normalidade e Identificação da variável resposta	29
4.2.2	Histogramas de candidatas a variáveis explicativas.....	33
4.2.3	Impacto das candidatas a variáveis independentes na média de ln(homem-hora)	38
4.3	Aplicação do modelo de regressão linear múltipla	41
5	CONCLUSÕES.....	46
5.1	Aspectos positivos e limitações do trabalho.....	46
5.2	Sugestões para futuros trabalhos	47
	REFERÊNCIAS BIBLIOGRÁFICAS.....	48
	APÊNDICE A	49
	APÊNDICE B	51

1 INTRODUÇÃO

Este capítulo introduz o leitor à monografia, apresentando a problemática, justificativa e relevância do trabalho, os objetivos, a metodologia e a organização dos capítulos.

1.1 Problemática

A crescente demanda mundial por fontes renováveis de energia elétrica tem acentuado o crescimento das empresas do segmento de geração de energia. Algumas empresas do setor são focadas em integração de sistemas de energia: planejam, integram e executam a montagem completa de sistemas de alta eficiência. Fornecem estudos de viabilidade, projetos, construção civil da obra, instalações elétricas, integração de sistemas de energia, *Start-up* e comissionamento, assistência técnica, etc. Existe algo que parece comum a várias empresas no setor: elas têm constantemente buscado evolução e têm adotado conceitos cada vez mais avançados em engenharia de sistemas.

Desta forma, para obter competitividade é desejável ter capacidade compatível com a demanda, produzir com qualidade, de forma mais eficiente, ter planejamento adequado, e atender o cliente em tempo acordado. Esses objetivos muitas vezes são interdependentes. De tal maneira, por exemplo, que para atender o cliente em tempo é necessário planejar adequadamente. Por isso é de grande importância ser capaz de estimar tempos de fabricação de produtos.

Um segmento do mercado de geração de energia consiste em empresas fabricantes de painéis elétricos. São elas que produzem os vários tipos de painéis necessários ao controle dos sistemas de energia. Além de serem vários os tipos, dificilmente dois sistemas de energia elétrica terão configurações iguais de painéis. Por este motivo é grande a dificuldade de se estimar o tempo de fabricação desse produto.

O tempo de fabricação dos produtos é uma fração do *lead time* - lapso de tempo entre pedido e entrega (SLACK, 2002). Essa fração do *lead time* é conhecida também como “*lead time* de Fabricação”, ou “*lead time* interno”, e será acerca deste tempo que investigações serão feitas ao longo deste trabalho. Alguns trabalhos relatam o uso de técnicas de estimação de *lead time* utilizando *data mining* (mineração de dados), análise de regressão linear, método indutivo, construção da árvore de decisão e regressão, seleção e discretização de atributos, etc. (OZTURK; KAYALIGIL; OZDEMIREL, 2005).

São várias as potenciais fontes de variação de tempo no processo fabril e, nasce então neste contexto o interesse em aplicar ferramentas estatísticas com o intuito de prever o tempo estimado necessário à fabricação de painéis.

1.2 Justificativa e relevância

Atualmente os gerentes de uma empresa fabricante de painéis elétricos no Brasil estimam os tempos de fabricação de seus produtos através da intuição, da vivência prática adquirida em seu processo de fabricação, que há um bom tempo observam, e tomam apenas essas observações como base para estimativas, o que pode acarretar imprecisões.

Uma grande fonte de variação nos tempos de manufatura de painéis concentra-se no fato de a produção de painéis ser bem mais dependente dos funcionários do que das máquinas, são eles que executam praticamente todas as atividades, não as máquinas. Em um processo que depende quase que exclusivamente de máquinas, e não pessoas, a estimação de tempos pode ser bem mais fácil e direta, apenas avaliando-se capacidades de produção das máquinas e fazendo cálculos de proporções. Além disso, quando os produtos são padronizados, também existe facilidade maior em estimar os tempos de fabricação, mesmo que esta seja executada por pessoas, o produto estará sempre se repetindo, e até mesmo o cálculo da média pode ser de grande valia. Porém no processo em estudo, os produtos não são padronizados (a empresa customiza seus produtos aos clientes) e os dois fatores mencionados dificultam a estimação de tempos.

O ato de estimar, dentro da engenharia, deve ocorrer da forma mais racional possível, traduzindo observações em grandezas mensuráveis e passíveis de análise e interpretação numérica. É importante observar que decisões tomadas com critérios matemáticos de avaliação costumam ter resultados mais positivos que decisões tomadas à luz de critérios subjetivos.

Atualmente é grande a facilidade de manipular dados em ferramentas computacionais, o que permite que diferentes modelos sejam rapidamente testados e escolhidos com segurança especificada.

Além disso, a empresa torna-se bastante dependente da experiência dos gestores na estimação dos tempos, quando um modelo matemático dispensaria grande parte dessa experiência – qualquer pessoa com o mínimo conhecimento sobre os produtos seria capaz de introduzir variáveis no modelo e extrair o tempo de fabricação estimado.

1.3 Objetivos

1.3.1 Objetivo Geral

Aliar conceitos de Análise de Variância e de Regressão para tentar explicar de forma estimada como variam os tempos de fabricação dos diversos tipos de painéis elétricos de acordo com variações nas variáveis explicativas que serão propostas.

1.3.2 Objetivos Específicos

- Explorar os tópicos conceituais relevantes à elaboração deste estudo
- Investigar as variáveis que potencialmente influenciam no tempo de fabricação
- Realizar uma modelagem matemática, transformando as informações do problema em uma função de várias variáveis
- Realizar testes estatísticos que comprovem a provável dependência ou não-dependência das variáveis estudadas sobre o tempo de fabricação com nível de segurança probabilística estabelecido previamente
- Avaliar e melhorar o modelo caso necessário e possível

1.4 Metodologia

Segundo Selltiz et al (1965), a finalidade da pesquisa é “descobrir respostas para questões, mediante a aplicação de métodos científicos”. As questões que estão sendo levantadas, assim como suas respostas, são de objetivo prático e dessa forma a pesquisa classifica-se como aplicada (ANDER-EGG, 1978).

Inicialmente foram realizadas entrevistas com um gerente de produção que tem conhecimento dos dados necessários a execução do estudo. Estes dados foram classificados em dois tipos: fatores potencialmente influenciáveis no tempo de fabricação e tempo de fabricação. Uma técnica de produção de idéias em grupo chamada *brainstorming* consiste em listar opiniões de todas as pessoas do grupo, sejam estas aparentemente plausíveis ou não. Em seguida, estes pensamentos são selecionados, com o intuito de se apresentar a solução mais eficaz para o problema. (ANZIEU; MARTIN, 1968). Foi feito um *brainstorming* para tentar identificar variáveis que supostamente explicam o comportamento da variabilidade dos tempos de fabricação dos painéis.

Foi feita, em paralelo às entrevistas, uma pesquisa bibliográfica para identificar trabalhos já desenvolvidos relacionados ao tema estudado. Dessa forma evita-se realizar

pesquisa que já foi feita. Além disso, essa forma de pesquisa permite a aquisição de conhecimentos teóricos acerca das ferramentas que serão utilizadas. Em estância seguinte, foi feita a aplicação dos conhecimentos obtidos na revisão bibliográfica aos dados observados. Estes receberam um pré-tratamento, buscando evitar dados discrepantes que possam potencializar problemas ao modelo.

1.5 Organização do Trabalho

O trabalho foi desenvolvido em cinco capítulos, obedecendo à seqüência a seguir.

O primeiro capítulo introduz o leitor ao trabalho, descrevendo a problemática, justificativa, objetivos e metodologia sob a qual foi desenvolvida a monografia. O segundo capítulo consiste em revisão bibliográfica sobre análise de regressão linear e análise de variância, trazendo os casos de regressão simples e regressão múltipla. O capítulo 3 descreve de maneira sucinta o contexto a nível mercadológico e fabril em que se encontram a empresa e o processo onde será aplicado o modelo de regressão para estimação de tempo necessário à manufatura. O capítulo 4 traz o levantamento de dados e o desenvolvimento de um modelo estimador para o referido tempo, contemplando análises por gráficos acerca das observações e de verificação de alguns pressupostos do modelo. O quinto e último capítulo conclui o trabalho apresentando críticas ao modelo obtido e sugere estudos futuros sobre o tema abordado na obra.

2 ANÁLISE DE REGRESSÃO

Este capítulo apresenta a base conceitual necessária à construção do modelo estimador de tempos de manufatura de painéis elétricos.

2.1 Introdução

A análise de Regressão teve origem em trabalhos de astronomia desenvolvidos pelo matemático, astrônomo e físico Johann Carl Friedrich Gauss no início do século XIX (1809 a 1821). É uma ferramenta da inferência estatística que possui grande utilidade em trabalhos científicos e tem sido utilizada em vários trabalhos devido à sua aplicabilidade e facilidade de operação. (DANTAS, 1998)

É uma técnica que fornece informações acerca do comportamento de uma variável (dependente) sobre outras (independentes). Para aplicação da ferramenta, parte-se de observação e registro das ocorrências das variáveis dependente e independentes, formando vetores de variáveis a cada observação. O método irá processar os dados e conduzir a uma expressão matemática que irá estimar a variável dependente ao se arbitrar valores das variáveis independentes.

Para se chegar à expressão (curva) que melhor se ajuste aos dados observados, utiliza-se o método dos mínimos quadrados. Consiste em minimizar os erros cometidos ao se usar a curva de predição da variável dependente ao invés dos valores reais observados. Como teremos desvios positivos e negativos, a soma de desvios é mais ideal quanto mais próxima estiver de zero. Por isso, quanto menor a soma de quadrados de erros, mais ajustada aos pontos estará a curva.

São vários os modelos de regressão já desenvolvidos. O modelo mais elementar é o de regressão linear simples, com apenas uma variável independente, e suposição de comportamento linear da variável dependente em relação a esta. Semelhante ao modelo de regressão linear simples, o modelo de regressão linear múltipla, também considera linear, essa relação, porém admite mais de uma variável independente no modelo. Outros modelos de regressão também podem ser usados para predição de uma variável dependente, a depender da natureza e do comportamento das variáveis. Segundo Lachos (2009), a regressão polinomial busca aproximar as observações à curva gerada por um polinômio de grau escolhido. O modelo de regressão logística é usado quando se deseja estimar a probabilidade de ocorrência de uma variável qualitativa, arbitrando-se a variável independente (OGLIARI, 2009).

Serão discutidos neste trabalho, os modelos de regressão linear simples e o modelo de regressão linear múltipla.

2.2 Regressão linear simples

Devido a seu caráter elementar e facilidade de entendimento, o primeiro modelo de regressão a ser apresentado é o de regressão linear simples. Muitos conceitos que forem apresentados no modelo de Regressão linear simples podem ser estendidos a outros modelos de regressão.

Uma função $f(x)$ é dita linear quando, ao se variar a variável independente x , o valor da função $f(x)$ sofre variação proporcional a x . Uma função linear assume a forma geral (STEWART, 2006):

$$f(x) = ax + b \quad (2.1)$$

ou,

$$y = ax + b \quad (2.2)$$

O modelo de regressão linear simples tentará aproximar os pontos observados a uma função linear, cujos pontos formam uma reta quando se representa as variáveis em plano cartesiano. Depois de construída a função que se aproxima dos pontos observados, esta poderá ser usada para fazer previsões da variável dependente ao se arbitrar a variável independente dentro de um campo de arbítrio permitido.

Suponha que se dispõe de dados de duas variáveis, e deseja-se investigar o grau de interação entre elas, ou seja, como uma variável responde a uma perturbação (variação) na outra. Para utilizar o modelo de regressão linear simples, é interessante que se observe tendência de linearidade nos pontos ao se analisar o diagrama de dispersão, como na figura 2.1.

Fica evidente então que se deseja estimar, com menor engano possível, os parâmetros a e b da equação da reta que representa (explica) o comportamento de y em relação à variável x , como na figura 2.2.

Seja Y o valor da variável dependente observado na amostra e seja \hat{Y} o valor correspondente estimado. Ao construir uma reta de regressão e utilizá-la para fazer previsões da variável dependente entrando-se com a dependente, comete-se um erro, há uma diferença entre o valor estimado e o valor real observado. Esse erro será denotado por “ e ” e tem valor:

$$e = Y - \hat{Y} \quad (2.3)$$

Não é difícil concluir que se a reta estiver bem ajustada aos pontos observados, deverá ser nula, a soma de erros, já que existirão valores acima e abaixo da reta. Mas isso não é o bastante para garantir que a reta está bem ajustada, é preciso garantir que sejam mínimos estes erros. Não é operacional matematicamente, fazer considerações diretamente aos erros para se obter a reta de mínimos erros. Mas é bastante prático considerar os quadrados dos erros – se os erros forem mínimos, também o serão seus quadrados, e então se deseja minimizar a soma de quadrados de erros (DANTAS, 1998).

2.2.1 Modelo geral de Regressão Linear Simples

O modelo geral de regressão linear simples busca explicar como perturbações na variável independente podem explicar perturbações na variável dependente (ALLARD, 1980).

O modelo de regressão linear simples pode ser representado na forma:

$$\hat{Y}_i = b_0 + b_1 \cdot X_i \text{ (amostra)} \quad (2.4)$$

$$\hat{Y}_i = \beta_0 + \beta_1 \cdot X_i \text{ (população)} \quad (2.5)$$

$$Y_i = \hat{Y}_i + e_i \text{ (amostra)} \quad (2.6)$$

$$Y_i = \hat{Y}_i + \varepsilon_i \text{ (população)} \quad (2.7)$$

Para que o modelo de regressão linear simples tenha bom desempenho, é interessante que o diagrama de dispersão demonstre certa tendência de linearidade nos pontos observados. Caso demonstre outra tendência, é necessário realizar transformações matemáticas na variável independente, ou na dependente, ou em ambas a fim de garantir tendência de linearidade, como será tratado em tópico específico (DANTAS, 1998).

Suponha que foram observados n pontos de uma amostra, como apresentado no diagrama de dispersão na figura 2.1.

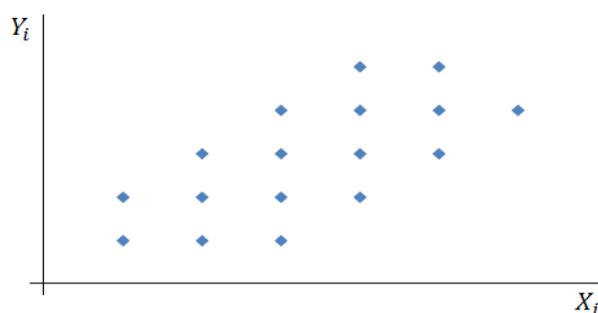


Figura 2.1: Diagrama de dispersão

Os valores observados da variável resposta parecem seguir um comportamento linear.

Deseja-se encontrar a reta que esteja mais próxima possível de todos os pontos da amostra, como na figura 2.2:

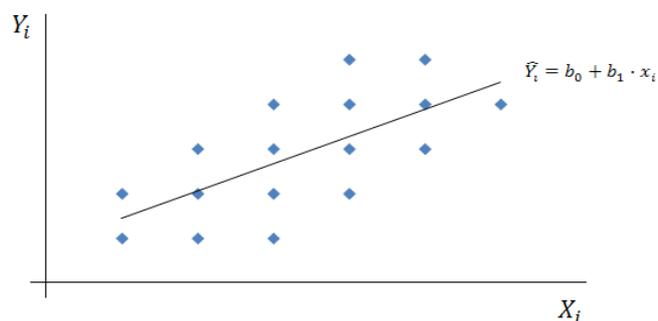


Figura 2.2: reta de regressão com inclinação

Como o comportamento da reta será descrito por $\hat{Y}_i = b_0 + b_1 \cdot X_i$ e os erros são dados por $e_i = Y_i - \hat{Y}_i$, então a soma de quadrados de erros pode ser escrita na forma:

$$U = \sum_i^n (Y_i - b_0 - b_1 \cdot X_i)^2 \quad (2.8)$$

Agora, percebe-se que dois parâmetros podem influenciar a soma de quadrados de erros, quais sejam, b_0 e b_1 . Quanto maior for b_0 , mais deslocada para cima será a reta, e quanto maior for b_1 , mais inclinada será a reta de regressão. Assim, deseja-se ajustar a reta de modo a obter a menor soma de quadrados de erros possível. O método dos mínimos quadrados fornece as estimativas dos parâmetros:

$$b_0 = \bar{Y} - b_1 \cdot \bar{X} \quad (2.9)$$

$$b_1 = \frac{\sum_{i=1}^n X_i Y_i - n \bar{X} \bar{Y}}{\sum_{i=1}^n X_i^2 - n \bar{X}^2} \quad (2.10)$$

Assim chega-se a reta de regressão linear simples por mínimos quadrados:

$$\hat{Y}_i = b_0 + b_1 \cdot X_i \quad (2.11)$$

2.2.2 Coeficiente de Correlação

O coeficiente de correlação é definido por

$$r = \frac{\sum_i^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_i^n (X_i - \bar{X})^2 \cdot \sum_i^n (Y_i - \bar{Y})^2}} \quad (2.12)$$

É uma medida da dependência linear entre as variáveis dependente e independente. O valor de r deve estar contido no intervalo $-1 \leq r \leq 1$ e quanto mais próximo de 1 estiver o módulo de r , mais forte a característica de linearidade entre as variáveis (ANDERSON *et al*, 2007).

Quando for negativo o valor de r , a reta de regressão será decrescente, ou seja, os valores de Y decrescem com o crescimento de X , ou crescem com o decréscimo de X

(relação inversa). Quando for positivo o valor de r , a reta de regressão será uma reta crescente, ou seja, os valores de Y crescem com o crescimento de X , ou decrescem com o decrescimento de X (relação direta). Quando o valor de r for nulo, não há correlação entre as variáveis e se estará diante de uma reta de regressão horizontal (DANTAS, 1998).

2.2.3 Coeficiente de determinação

O coeficiente de determinação é uma medida do poder de explicação do modelo. Representa a proporção de variabilidade da variável dependente que pode ser explicada por variações nas variáveis independentes (DANTAS, 1998).

Seu valor é definido por:

$$R = r^2 \quad (2.13)$$

Substituindo o valor de r , encontra-se:

$$R = \frac{(\sum_i^n (X_i - \bar{X})(Y_i - \bar{Y}))^2}{\sum_i^n (X_i - \bar{X})^2 \cdot \sum_i^n (Y_i - \bar{Y})^2} \quad (2.14)$$

Como o valor de R é o quadrado de r , R poderá assumir valores contidos no intervalo $0 \leq R \leq 1$. Quanto mais próximo da unidade estiver o valor de R , maior o poder de explicação do modelo com as variáveis consideradas, quanto mais próximo de 0 estiver o valor de R , menor será o poder de explicação do modelo com as variáveis consideradas. Pode-se afirmar que R é uma medida da exatidão do ajuste da linha de regressão.

É interessante comentar que ao realizar a regressão de Y sobre X ou de X sobre Y , teremos equações completamente simétricas para R . Logo, os cálculos de R para as duas regressões deverão levar ao mesmo valor de R . (ALLARD, 1980)

Outra maneira de chegar ao coeficiente de determinação é observando graficamente as variabilidades da variável resposta, como indica a figura 2.3:

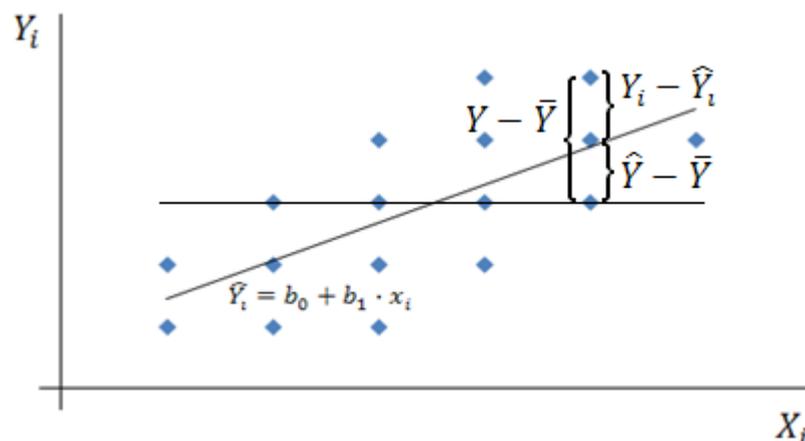


Figura 2.3: representação gráfica das variabilidades da variável resposta

$$\text{Soma dos quadrados dos erros: } SSE = \sum(Y_i - \hat{Y}_i)^2 \quad (2.15)$$

$$\text{Soma dos quadrados da regressão: } SSR = \sum(\hat{Y}_i - \bar{Y})^2 \quad (2.16)$$

$$\text{Soma total dos quadrados: } SST = \sum(Y_i - \bar{Y})^2$$

$$R = r^2 = \frac{SSR}{SST} \quad (2.17)$$

2.2.4 Estimadores de dispersão

2.2.5.1 Variância (σ^2)

A variância é calculada pela média dos desvios quadráticos, isto é, a média das diferenças entre cada elemento observado (X_i) e a média de X_i , elevados ao quadrado (DANTAS, 1998, p. 75). Porém esta média possui um caráter tendencioso, e é necessário corrigi-la por um fator multiplicativo $\frac{n}{(n-1)}$. Desta forma, tem-se:

$$\sigma^2 = \frac{\sum_i^n (X_i - \bar{X})^2}{n-1} \quad (2.18)$$

Para o modelo geral de regressão, a estimativa de σ^2 é obtida calculando-se o erro médio quadrático, que consiste no quociente entre a soma de quadrados de resíduos e o número de graus de liberdade desta soma:

$$\hat{\sigma}^2 = \frac{\sum_i^n (Y_i - \hat{Y}_i)^2}{n-p} \quad (2.19)$$

Com p parâmetros, n observações e $n-p$ graus de liberdade (cada parâmetro corresponde à perda de um grau de liberdade).

No caso da regressão linear simples (reta), tem-se:

$$\hat{\sigma}^2 = s^2 = \frac{\sum_i^n (Y_i - \hat{Y}_i)^2}{n-2} = MSE = \frac{SSE}{n-2} \quad (2.20)$$

2.2.5.2 Desvio-padrão

A medida de dispersão mais usada, que pode ser considerada como uma medida de variabilidade dos dados de uma distribuição de frequências. Isto é, o desvio padrão mede a dispersão dos valores individuais em torno da média. Para seu cálculo, deve-se obter a média da distribuição e, a seguir, determinar os desvios para mais e para menos a partir da mesma. Assim, o desvio padrão é a raiz quadrada da média aritmética dos quadrados dos desvios, esses tomados a partir da média aritmética (SIAS, 2009).

O desvio-padrão é definido como a raiz quadrada positiva da variância. Para o modelo geral de regressão o desvio-padrão é estimado por (DANTAS, 1998):

$$\hat{\sigma} = \sqrt{\frac{\sum_i^n (Y_i - \hat{Y}_i)^2}{n-p}} \quad (2.21)$$

No caso da regressão linear simples (reta), tem-se:

$$\hat{\sigma} = s = \sqrt{\frac{\sum_i^n (Y_i - \hat{Y}_i)^2}{n-2}} \quad (2.22)$$

2.2.5 Suposições do modelo

- Modelo de regressão suposto: $Y = \beta_0 + \beta_1 \cdot X + \varepsilon$
- Equação de regressão estimada por método de mínimos quadrados: $\hat{Y}_i = b_0 + b_1 \cdot X_i$.
- O termo de erro ε é uma variável aleatória com uma média, ou valor esperado, igual a zero.
- A variância de ε , σ^2 , é a mesma para todos os valores de X .
- Os valores de ε são independentes.
- O termo de erro é uma variável aleatória normalmente distribuída.

Implicações:

- $E(\beta_0) = \beta_0$; $E(\beta_1) = \beta_1$; $E(Y) = \beta_0 + \beta_1 \cdot X$
- Variância de Y nas proximidades da reta de regressão é igual a σ^2 e é a mesma para todos os valores de X .
- O valor de ε para um valor esperado em particular de X não está relacionado ao valor de ε para qualquer outro valor de X . Assim, o valor de Y para um valor em particular de X não está relacionado com o valor de Y para qualquer outro valor de X .
- Uma vez que Y é uma função linear de ε , Y também é uma variável aleatória normalmente distribuída. (ANDERSON et al, 2007)

2.2.6 Testes de significância

O modelo de regressão linear simples calcula a média, ou o valor esperado, de Y através de uma função linear de X , $\hat{Y}_i = b_0 + b_1 \cdot X_i$. Para investigar se uma relação linear é significativa, deve-se realizar um teste de hipóteses (H_0 e H_1) no estimador b_1 , do coeficiente de inclinação da reta (DANTAS, 1998):

$$H_0: b_1 = 0$$

$$H_1: b_1 \neq 0$$

Se H_0 for rejeitada, significa que é possível concluir que há relação linear entre X e Y . Se H_0 não for rejeitada não haverá evidências suficientes para concluir que existe relação linear significativa.

Dois testes comumente são usados, e ambos precisam de uma estimativa de σ^2 para serem usados. São o teste t e o teste F .

2.2.7.1 O teste t

O estimador b_1 possui as seguintes propriedades:

$$b_1 \sim N(\beta_1, \sigma_{b_1}) \quad (2.23)$$

$$\text{Valor esperado } E(b_1) = \beta_1 \quad (2.24)$$

$$\text{Desvio Padrão } \sigma_{b_1} = \frac{\sigma}{\sqrt{\sum(X_i - \bar{X})^2}} \quad (2.25)$$

$$\text{Estimador do desvio-padrão de } b_1: \hat{\sigma}_{b_1} = s_{b_1} = \frac{s}{\sqrt{\sum(X_i - \bar{X})^2}} \quad (2.26)$$

A estatística de teste $\frac{b_1 - \beta_1}{s_{b_1}}$ segue uma distribuição t de Student com $n - 2$ graus de liberdade. Como se deseja testar a hipótese nula, então a estatística resulta em $t = \frac{b_1}{s_{b_1}}$. A estatística de teste deve ser comparada com o valor, encontrado na distribuição de probabilidade, correspondente ao número de graus de liberdade e nível de significância (α) desejado. A hipótese nula poderá ser rejeitada se a estatística de teste, t , for maior do que o valor de t encontrado na distribuição, o que equivale a rejeitar a hipótese nula se o valor- p for menor do que o nível de significância desejado, α (ANDERSON *et al*, 2007).

2.2.7.2 O teste F

O teste F avalia a significância da relação entre as variáveis explicativas de um modelo e a variável resposta. Para o caso da regressão linear simples, o teste t possui a mesma aplicabilidade do teste F . Porém, o teste t é limitado a apenas uma variável explicativa, ao contrário do teste F (DANTAS, 1998).

Para realizar o teste F , é necessário calcular MSR e MSE .

$$MSR = \frac{SSR}{GLR} \quad (2.27)$$

$$MSE = \frac{SSE}{n-2} \quad (2.28)$$

$GLR =$ número de graus de liberdade da regressão

Para os modelos apresentados nesse trabalho, o valor de GLR é sempre igual ao número de variáveis do modelo.

Estatística de teste:

$$F = \frac{MSR}{MSE} \quad (2.29)$$

Observando-se a figura 2.3, verifica-se que quanto maior for o valor de β_1 , maior será a inclinação da reta e, portanto mais explicativa será a variável independente. Com isso, maior será o valor de MSR comparado a MSE , e maior será o valor da estatística F . Dessa forma quanto menor for o valor de F , mais próximo se estará da aceitação da hipótese nula.

Se a hipótese nula for verdadeira, MSR e MSE são duas estimativas independentes de σ^2 e MSR/MSE segue uma distribuição F com 1 grau de liberdade no numerador e $n - 2$ graus no denominador (ANDERSON *et al*, 2007).

Regra de rejeição:

Critério do valor p : Rejeitar H_0 se o valor $p \leq \alpha$

Critério do valor crítico: Rejeitar H_0 se o $F \geq F_\alpha$

Onde F_α é obtido dos pontos críticos (F) da distribuição de Snedecor tabelados por Fischer com um grau de liberdade no numerador e $n - 2$ no denominador (DANTAS, 1998).

2.2.7 Análise de resíduos

O modelo de regressão linear simples apresentado contempla a seguinte forma:

$$\hat{Y}_i = \beta_0 + \beta_1 \cdot X_i \quad (2.30)$$

$$Y_i = \hat{Y}_i + \varepsilon_i \text{ ou,} \quad (2.31)$$

$$Y_i = \beta_0 + \beta_1 \cdot X_i + \varepsilon_i \quad (2.32)$$

Onde o resíduo, ou erro, ε_i é a diferença entre o valor observado da variável dependente e seu valor estimado pela regressão.

Os testes t e F baseiam-se em suposições que foram feitas acerca dos resíduos ε_i na seção 2.2.6. Assim a rejeição ou aceitação da hipótese nula, H_0 , tem validade se forem verificadas essas suposições. Por isso é importante realizar uma análise dos resíduos para se

concluir se existem evidências de significância da relação linear entre as variáveis estudadas. É interessante dispor os valores da variável independente, da dependente, da estimativa, e dos resíduos referentes a cada observação i em um plano cartesiano, como indicado na figura 2.4. Para o caso da regressão linear simples, o padrão de comportamento de resíduos, será o mesmo ao se plotar ε_i contra X_i , ou ε_i contra \hat{Y}_i . Para o caso da regressão múltipla, a plotagem de \hat{Y}_i contra ε_i é mais amplamente utilizada em virtude da presença de mais de uma variável independente. O comportamento dos resíduos possui um bom padrão quando a variabilidade destes não tende a aumentar ou diminuir ao se variar o valor de X_i ou \hat{Y}_i , como mostrado na figura 2.4. As figuras 2.5 e 2.6 indicam um mau padrão de resíduos.

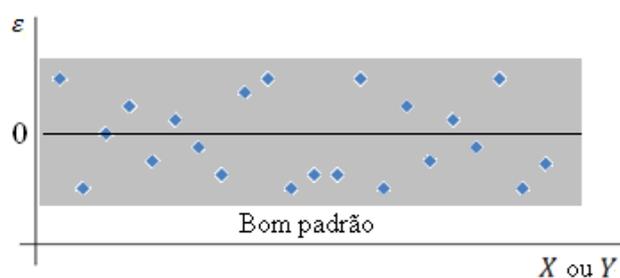


Figura 2.4: bom padrão de resíduos

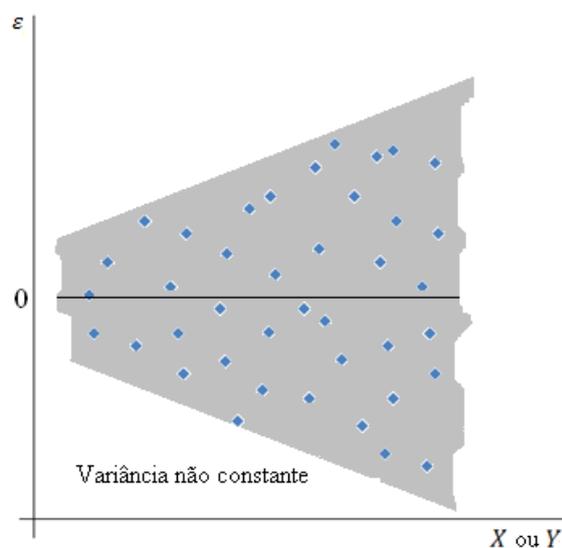


Figura 2.5: variância não constante de resíduos

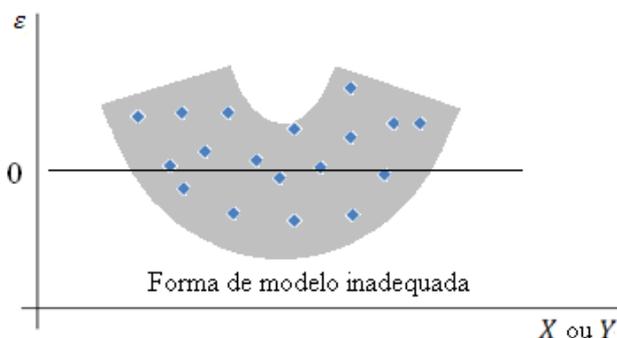


Figura 2.6: exemplo de erros de modelo inadequado

2.3 Regressão linear múltipla

O modelo de regressão simples é usado para explicar o comportamento de uma variável dependente através de uma variável independente. Nas mais freqüentes situações do mundo real, uma variável pode ser influenciada por vários fatores. Dessa forma um modelo que seja capaz de agrupar mais de uma variável explicativa pode se mostrar, em geral, mais eficiente que um modelo com apenas uma. Tais modelos são conhecidos como modelos múltiplos e seguem desenvolvimento semelhante ao da regressão linear simples.

O modelo de regressão linear múltipla considera linear, a relação entre a variável dependente e as variáveis independentes. Deste modo, o modelo segue a seguinte forma:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p + \varepsilon \quad (2.33)$$

Onde o índice p denota o número de variáveis no modelo e ε , o termo de erro da previsão pelo modelo de regressão. Deseja-se encontrar a equação de previsão do valor esperado de y em função das variáveis explicativas, que contemple a melhor aproximação possível dos reais valores de y . Tal equação é chamada equação de regressão múltipla e é da forma:

$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p \quad (2.34)$$

Como em geral não se dispõe de todos os dados da população, é feita uma estimativa com base em dados amostrais observados, e com esses dados estima-se os valores dos parâmetros $\beta_0, \beta_1, \beta_2, \dots, \beta_p$ por $b_0, b_1, b_2, \dots, b_p$. Utilizando os parâmetros estimados para compor a equação de regressão múltipla, tem-se a equação de regressão múltipla estimada:

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + \dots + b_px_p \quad (2.35)$$

2.3.1 A estimação dos parâmetros do modelo

Tal como ocorrido para a regressão linear simples, deseja-se minimizar os erros cometidos ao se usar a equação de regressão linear múltipla. Da mesma maneira, recorre-se ao método dos mínimos quadrados, procurando os parâmetros que minimizem a soma (DANTAS, 1998):

$$U = \sum_{i=1}^n (Y - \hat{Y})^2 \quad (2.36)$$

A dedução dos valores de estimativas de parâmetros não é simples como no caso da regressão linear simples, por envolver álgebra matricial e operações mais complexas e estar além do escopo deste trabalho. Ferramentas computacionais são bastante utilizadas para se chegar às estimativas de parâmetros e, as saídas desses softwares, além dessas estimativas, fornecem diversas informações acerca do desempenho do modelo, como as que foram demonstradas para o modelo de regressão linear simples.

2.3.2 Coeficiente de determinação múltiplo

Assim como na regressão linear simples, este coeficiente representa a proporção de variabilidade da variável dependente que pode ser explicada pela variabilidade das variáveis independentes e é calculado da mesma maneira na regressão múltipla (Berenson, 2004). Deste modo teremos a mesma expressão:

$$R = r^2 = \frac{SSR}{SST} \quad (2.37)$$

$$\text{Onde } SST = SSR + SSE. \text{ Com:} \quad (2.38)$$

$$\text{Soma dos quadrados dos erros: } SSE = \sum (Y_i - \hat{Y}_i)^2 \quad (2.39)$$

$$\text{Soma dos quadrados da regressão: } SSR = \sum (\hat{Y}_i - \bar{Y})^2 \quad (2.40)$$

$$\text{Soma total dos quadrados: } SST = \sum (Y_i - \bar{Y}_i)^2 \quad (2.41)$$

2.3.3 Coeficiente de determinação múltiplo ajustado

Na análise de regressão múltipla, toda vez que uma variável independente é adicionada, o coeficiente de determinação aumenta, mesmo que esta nova variável não contribua

significativamente para a explicação da variância residual. Desta forma, é útil analisar o coeficiente ajustado, que, ao contrário do coeficiente de determinação, diminui se a variável adicionada possuir um poder de explicação baixo, podendo ter até um valor negativo (Élin Duxus, 2009). Este coeficiente leva em conta o número de variáveis que está no modelo.

O valor do coeficiente de determinação múltiplo ajustado é calculado por:

$$R_a^2 = 1 - (1 - R^2) \frac{n-1}{n-p-1} \quad (2.42)$$

2.3.4 Suposições do modelo

- Modelo de regressão linear múltipla: $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p + \varepsilon$
- O termo de erro ε é uma variável aleatória com uma média, ou valor esperado, igual a zero.
- A variância de ε , σ^2 , é a mesma para todos os valores das variáveis independentes.
- Os valores de ε são independentes.
- O termo de erro é uma variável aleatória normalmente distribuída

Implicações:

- $E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p$ (2.43)
- Variância de y nas proximidades da linha de regressão é igual a σ^2 e é a mesma para todos os valores das variáveis independentes x_1, x_2, \dots, x_p .
- O valor de ε para um conjunto de valores em particular das variáveis independentes não guarda relação com o valor de erro para qualquer outro conjunto de valores.
- A variável dependente y também é uma variável aleatória normalmente distribuída. (ANDERSON et al, 2007)

2.3.5 Testes de significância

Na regressão simples, o teste t e o teste F eram utilizados com o mesmo propósito, testar se a variável independente x é significativa para explicar variabilidade da variável dependente y . Na análise de regressão múltipla, esses testes possuem propósitos diferentes. O teste F é utilizado para testar se existe uma relação significativa entre a variável dependente e todas as variáveis independentes, por isso é chamado teste de significância global. O teste t é

utilizado para testar se cada variável, individualmente, é significativa, por isso é chamado teste de significância individual (ANDERSON *et al*, 2007).

2.3.5.1 Teste F

Hipóteses do teste F :

$$H_0: \beta_1 = \beta_2 = \dots = \beta_p = 0 \text{ (hipótese a ser testada)}$$

H_1 : pelo menos um dos parâmetros é diferente de zero

Estatística de teste:

$$F = \frac{MSR}{MSE} \tag{2.44}$$

$$\text{Onde, } MSR = \frac{SSR}{p} = \text{Média de quadrados da regressão} \tag{2.45}$$

$$\text{E, } MSE = \frac{SSE}{n-p-1} = \text{Erro médio quadrático.} \tag{2.46}$$

Regras de rejeição da hipótese nula:

Critério do valor p : Rejeitar H_0 se valor $p \leq \alpha$

Critério do valor crítico: Rejeitar H_0 se $F \geq F_\alpha$

Onde F_α é obtido dos pontos críticos (F) da distribuição de Snedecor tabelados por Fischer com p graus de liberdade no numerador e $n - p - 1$ no denominador.

2.3.5.2 Teste t

O teste t é utilizado para se testar a significância individual dos parâmetros, sendo este aplicado separadamente a cada parâmetro estimado (ANDERSON *et al*, 2007). Segundo Dantas (1998), o teste deve ocorrer da forma apresentada abaixo.

Hipóteses do teste t :

$$H_0: \beta_i = 0 \text{ (hipótese a ser testada)}$$

$$H_1: \beta_i \neq 0$$

Estatística de teste:

$$t = \frac{b_i}{s_{b_i}} \quad (2.47)$$

Onde s_{b_i} é o desvio-padrão da estimativa b_i .

Regras de rejeição da hipótese nula:

Critério do valor p : Rejeitar H_0 se valor $p \leq \alpha$

Critério do valor crítico: Rejeitar H_0 se $t \leq -t_{\alpha/2}$ ou $t \geq t_{\alpha/2}$

Onde $t_{\alpha/2}$ baseia-se em uma distribuição t com $n - p - 1$ graus de liberdade.

2.3.6 Multicolinearidade

Em um problema de regressão múltipla, diversas vezes depara-se com uma situação onde algumas variáveis independentes guardam certa dependência linear. Para medir o quanto estas variáveis estão relacionadas deve-se medir a correlação entre elas, duas a duas, da mesma maneira que foi apresentado para a regressão simples. Quando estas variáveis estão muito correlacionadas, diz-se que existe um problema de multicolinearidade.

Quando existe grande multicolinearidade entre variáveis independentes, ao se fazer um teste t em cada variável é possível que não se possa descartar a hipótese nula caso a caso. Mesmo assim, o teste F pode indicar que existe relação significativa entre a variável dependente e as variáveis independentes (significância global). Este é um problema causado freqüentemente pela multicolinearidade.

Os práticos estatísticos chegaram experimentalmente à conclusão de que a multicolinearidade pode causar problemas quando a correlação entre variáveis explicativas for maior que 0,7 (ANDERSON *et al*, 2007).

2.3.7 Variáveis qualitativas independentes

Quando se deseja incluir no modelo uma característica quantitativa, por exemplo, sexo masculino ou feminino, reparo mecânico ou elétrico, noite ou dia, ausência ou presença de alguma característica, etc. deve se usar esse tipo de variável, chamada variável simulada. Deve-se incorporar uma variável da seguinte forma (as opções devem ser exclusivas):

0 se possuir uma característica
1 se possuir a outra

A presença de uma variável simulada no modelo admite a interpretação de que na verdade poderá haver dois modelos de regressão múltipla. Se a variável simulada mostrar-se significativa para explicar variações na variável resposta e fixando-se as outras variáveis independentes, quando ela assumir valor zero, o modelo de regressão fornecerá a estimativa da variável resposta quando estiver presente uma das características qualitativas na observação, e quando ela assumir valor 1, o modelo fornecerá a estimativa da variável resposta quando estiver presente a outra característica. No caso de um modelo com uma variável quantitativa e uma qualitativa, a interpretação de haver 2 modelos, fornecerá um gráfico com 2 retas de regressão paralelas, em que a distância entre elas representa a diferença entre as médias, ponto a ponto, da variável resposta ao se fazer presente uma característica ou outra (ANDERSON *et al*, 2007).

3 MANUFATURA DE PAINÉIS ELÉTRICOS

Este capítulo situa o leitor no contexto de mercado da empresa e no processo fabril onde são produzidos os painéis elétricos.

3.1 A empresa

3.1.1 Contexto de mercado

A AREVA KOBLITZ atua no mercado de integração de sistemas de geração e cogeração de energia, ou seja, o sistema que produz a energia é o mesmo que a consome. A empresa fornece serviços para a construção de centrais termoeletricas e hidrelétricas, desde projetos até o comissionamento e entrega do sistema.

A empresa empenha-se em desenvolver e implementar em seus projetos, tecnologias mais livres de impactos ambientais de forma a promover o desenvolvimento sustentável. A empresa foi fundada em 1975 na cidade do Recife com o nome KOBLITZ, município que sedia até hoje a AREVA KOBLITZ. Desde então tem crescido fisicamente e financeiramente, tendo a sede, em aproximadamente trinta anos, quadruplicado seu tamanho físico.

A elaboração e fornecimento dos painéis elétricos necessários à construção do sistema de geração de energia são feitos pela própria empresa, na unidade Recife, que abriga o processo de fabricação destes equipamentos. Além da fabricação dos painéis, a AREVA KOBLITZ executa a montagem do sistema e comissionamento, as demais atividades da construção da obra são executadas, atualmente, por terceiros.

A antiga KOBLITZ tornou-se multinacional no início de 2008, quando se uniu à divisão brasileira de bioenergias do grupo AREVA, empresa do ramo de energia que atua no desenvolvimento de geração de energias limpas, livres de CO₂, e de soluções para transporte e distribuição de eletricidade. A posição de multinacional vem dando espaço à expansão da rede de negócios da empresa, beneficiando seu desenvolvimento e fortalecendo-a.

2.3.8 O processo de manufatura de painéis

O Processo de manufatura de painéis pode ser representado pelo fluxograma apresentado na figura 3.1.

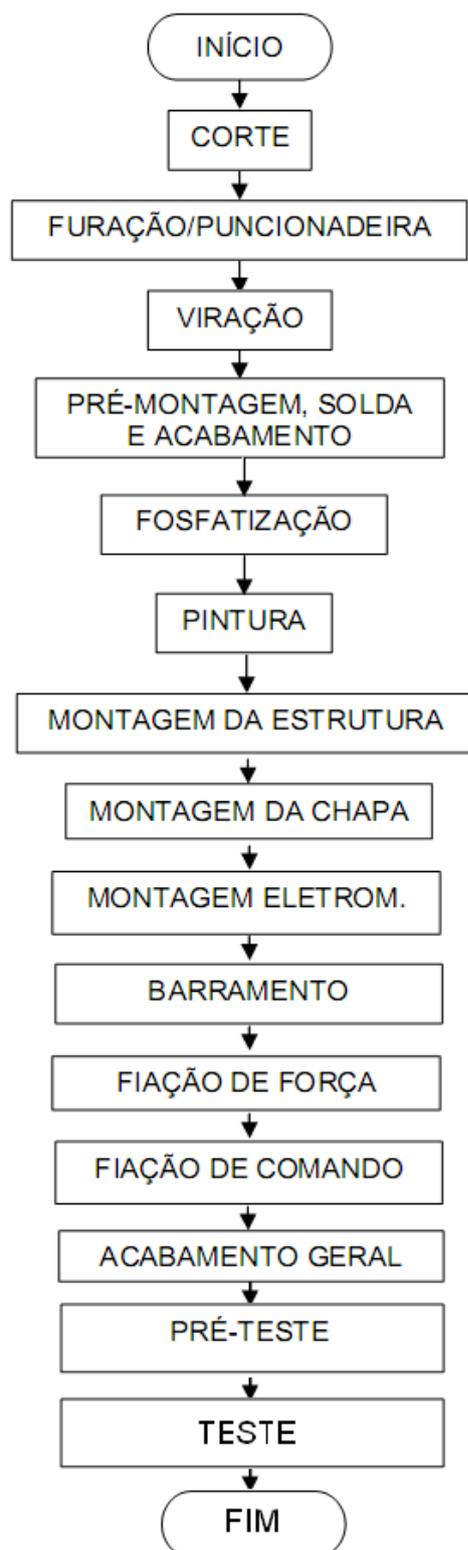


Figura 3.1: fluxograma de produção de painéis

É a própria AREVA KOBBLITZ quem confecciona seus painéis. A estrutura mecânica é toda construída na fábrica, e os componentes elétricos e eletrônicos são montados na estrutura. Entre os setores da empresa, encontra-se a divisão de painéis, responsável pela produção de painéis elétricos.

O processo de fabricação de painéis se inicia quando é enviado ao DPI (departamento de produção industrial), o projeto da estrutura mecânica do painel. Após receber o projeto, o DPI o envia ao setor de serralharia, que executa a ferramentação do projeto (programação da puncionadeira para executar furos e cortes mais precisos). Após a ferramentação, dá-se início o processo de manufatura dos painéis.

O primeiro ponto no processo de transformação da matéria prima é o corte de chapas, que podem ser de aço carbono ou de zinco (as de zinco não passarão por solda, acabamento, fosfatização, nem pintura). As chapas são retiradas de seus respectivos fardos através de ponte rolante e sistema de carrinho sobre trilhos e são transportadas até a primeira máquina envolvida no processo, a guilhotina. Nela são feitos cortes maiores, para deixar a chapa em um tamanho ideal para ser processada na puncionadeira, de acordo com a ferramentação.

Em seguida, as chapas seguem para a puncionadeira, máquina de suma importância no processo da serralharia. É ela quem realiza cortes mais precisos, detalhes, venezianas, quinas, furações, etc. Tudo com bastante agilidade e precisão.

A etapa seguinte é a viração de chapas. É onde são feitas todas as dobras nas chapas e as folhas começam a ganhar geometrias tridimensionais. A etapa de viração de chapas conta com duas viradeiras hidráulicas, a primeira está sempre configurada com todas as matrizes de dobra (formando uma única matriz contínua) e a segunda sofre setups conforme necessidade. A primeira executa dobras em uma direção e suas dobras paralelas e as demais dobras são feitas na segunda viradeira.

Em seguida, as chapas em aço carbono viradas seguem para o setor de solda e acabamento, onde as chapas já viradas são ponteadas com solda MIG, esquadrejadas e soldadas para formar peças de painéis. Após a solda, as peças devem passar por acabamento com lixadeiras elétricas para remover excessos de solda e as eventuais falhas de soldagem são corrigidas com solda elétrica.

Neste ponto do processo as peças já possuem forma, mas antes de começar a formar o painel, as peças em aço carbono, devem passar pelos processos de fosfatização e pintura. O processo de fosfatização tem como objetivo remover qualquer resíduo de gordura, óleo, sujeiras, oxidações e aplicar uma camada de fosfato sobre a chapa. Para isso as peças passam

por banho em oito tanques com soluções aquosas e água, entre elas uma solução desengraxante, uma ácida, uma básica, lavagem em água, passivador, refinador de camada, e fosfato de zinco, de onde vem o nome do processo.

Agora as peças estão prontas para receberem a pintura. As peças são penduradas em um trilho suspenso que as conduzem à cabine de pintura. Os trilhos eletrizados encarregam-se de eletrizar as peças, e a tinta em pó, é eletrizada com carga oposta, fazendo a tinta aderir eletrostaticamente à peça. Porém a tinta até agora nada mais é do que pó sendo atraído pela chapa. Por isso as peças devem passar por uma estufa para que haja fusão da tinta e completa adesão às peças.

Depois de pintadas, as peças serão agregadas umas as outras, em células de montagem, num processo chamado de montagem mecânica. As peças são montadas de forma a liberar por cada ciclo de montagem, um “cubículo” ou “coluna” (frações de painéis e principal unidade de medida da produção). Aqui são feitos furos que não foram feitos na punctionadeira, corrigidos furos errados, e as peças são fixadas umas às outras por porcas, parafusos, rebites, etc. É de grande importância que os processos da serralharia evitem falhas, pois neste momento do processo de montagem de estrutura, suas não-conformidades geram um grande impacto, ocasionando dificuldade na montagem devido a furos não coincidentes, problemas de viração, peças fora de esquadro impossibilitando o devido encaixe, etc.

Após o fim desta etapa do processo, a estrutura já possui forma de painel, ou pelo menos de coluna e está pronto para sofrer o processo de montagem elétrica. É neste momento que pode haver a interferência mais significativa no processo, quando a estrutura está montada, mas não foram cumpridos alguns prazos de chegada de componentes, ocasionando a espera de painéis semi-acabados por estes materiais.

Agora se inicia a montagem elétrica, podendo esta etapa ser dividida em Fiação de força, Barramento, Montagem da chapa, Montagem Eletromecânica, Acabamento Geral, Fiação de Comando e Pré-Teste.

A montagem da chapa de montagem inicia selecionando-se as chapas conforme projeto e posicionando-as sobre cavaletes. Em seguida os materiais e equipamentos são distribuídos conforme lay out da chapa apresentado no projeto do painel. Caso o montador, líder de produção ou o supervisor verifique que o lay out apresentado não é a melhor opção, o projetista envolvido é consultado para discutir a alteração do projeto. Em seguida, as posições dos componentes são marcadas na chapa, furos são feitos e os componentes são fixados. Nesta etapa também são colocadas as calhas plásticas que irão abrigar a fiação.

Após a montagem da chapa, é feita a montagem eletromecânica. É feita a montagem das travessas e suportes, obedecendo às cotas especificadas no projeto elétrico. São colocados os TC's, TP's, Chaves Seccionadoras, disjuntores, bases e fusíveis, nos seus devidos suportes e fixados com parafusos, arruelas e porcas.

A confecção do barramento passa por operações semelhantes à serralharia, ocorrendo inicialmente o corte das barras de cobre e virações. As barras (100%) são posicionadas no painel para checar o alinhamento das mesmas, devendo ser desempenadas quando empenos forem observados. Em seguida são retiradas do painel e recebem as marcações de contato elétrico, de furos e são feitas as devidas furações. São novamente posicionadas no painel para checar os furos. Depois são identificadas conforme sua respectiva fase (R, S, T). As barras são novamente retiradas do painel e as áreas de contato são lixadas e recebem líquido de pratear. São encapsuladas quando especificado em projeto, pintadas conforme cor da fase especificada em projeto, e retornam ao painel, onde são posicionadas pela última vez e são apertadas as conexões do barramento.

A próxima etapa é a fiação de força. Os cabos são posicionados realizando-se seu percurso no painel, em seguida são colocados os terminais de compressão nas duas extremidades do cabo. Quando o terminal não for pré-isolado, isola-se a junção do cabo com o terminal com fita isolante. O cabo é fixado no ponto de origem e no ponto de destino. Caso o componente que sedia o ponto de origem ou destino não estiver no painel, o cabo deve ser marcado com indicação de sua ligação. Os cabos recebem acabamento sendo amarrados com abraçadeiras de nylon.

A fiação de comando é feita de acordo com a planilha de montagem que consta no projeto do painel. Cada cabo deve respeitar bitolas e cores especificadas em projeto. Cada cabo recebe anilha de identificação em suas duas extremidades, em seguida recebem terminais na extremidade da chapa de montagem, são feitas as ligações desta mesma extremidade em seus pontos de origem e os cabos são dispostos nas calhas plásticas. Em seguida, a chapa de montagem é transportada até o painel, onde é fixada. A fiação é organizada, são colocados terminais nas extremidades dos cabos que irão para a porta, e os cabos são fixados em seus pontos de destino. Em caso de painel de Média, fazer o chicote do cabeamento de comando do disjuntor ou outro componente que seja separado do conjunto, no posto de trabalho da fiação. É feito acabamento na fiação com abraçadeiras, fazendo amarração e fechando as calhas plásticas.

Em seguida é feito o acabamento final do painel, colocando-se etiquetas de identificação de PVC na porta do painel, a etiqueta de identificação do fabricante, é feito o fechamento do painel com as laterais, colocadas borrachas nos aros de portas, é feita a montagem do sinótico, limpeza e polimento com cera.

Antes de ser levado ao setor de inspeção e controle de qualidade, é feito o pré-teste, que consiste num teste ponto a ponto verificando se há continuidade entre o ponto de origem e o ponto de destino dos cabos. Caso todas as ligações estejam corretas, o painel segue para o teste de qualidade. Caso contrário, o montador deve identificar e corrigir o erro de montagem.

Por último, o painel passará pelo setor de inspeção e controle de qualidade onde são feitos vários testes, entre eles inspeção visual, checando cor, medidas, lay out, fixação de componentes, acabamento em geral, vedação, trancas, etiquetas, aterramento, conferência de identificação de anilhas, etc. É feito o teste funcional, onde são acionados os equipamentos do painel, testando disjuntores, chaves, contadores, etc. Caso seja verificada alguma não conformidade, o departamento de produção industrial é acionado e são feitos os ajustes necessários, dando fim ao processo de produção de painéis.

4 RESULTADOS

Este capítulo apresenta a identificação e distribuições de variáveis, a aplicação da análise de regressão e o modelo desenvolvido com a ferramenta estatística, bem como a avaliação de seu desempenho.

4.1 Levantamento das variáveis

Após um *brainstorming* com coordenadores e o autor, chegou-se a um conjunto de características candidatas a variáveis explicativas, quais sejam:

- Cubículos – representa a quantidade de colunas, ou cubículos, que possui o painel observado.
- Requisições – indica o número de itens requisitados para a ordem de produção (OP) do painel.
- Barramentos – indica a presença ou ausência de barramento, se a observação for sem barramentos, a variável assume valor 0, caso contrário, 1.
- Funcionários – é o número médio de funcionários a trabalhar na montagem do painel, ao longo de seu processo de manufatura.
- Automação – indica se o painel será com ou sem automação. Em caso afirmativo, a variável assume valor 1, caso contrário, 0.
- Ligações – representa a quantidade de pontos de ligação da fiação do painel.
- Tensão – indica se a tensão de operação do painel é baixa ou média. Para baixa, a variável assume valor 0, caso contrário, 1.
- Terceiros – indica se a montagem elétrica do painel será executada pela própria AREVA KOBLITZ – valor 0 – ou por empresas terceirizadas – valor 1.
- Pintura – indica o tipo de pintura da estrutura, se é pó ou poliéster. Esta característica de projeto está relacionada à exposição do painel, caso seja abrigado, recebe pintura pó e a variável assume valor 0, caso o painel seja exposto ao tempo, recebe pintura poliéster e a variável assume valor 1.
- Tipo – Os diversos painéis fabricados pela AREVA KOBLITZ (conforme apêndice A) podem ser agrupados em classes por semelhança em dificuldade de montagem, tipos e quantidades de equipamentos utilizados, função a que se destinam, etc. Tipo é uma variável que pode assumir valores inteiros no intervalo de 1 a 14, onde cada valor representa uma classe, ou tipo de painel

conforme tabela 3.1. A ordenação numérica foi feita de acordo com a quantidade de homens-hora necessária a manufatura de cada tipo. Não houve observação dos tipos PNGBT-PE e MP/MC, por isso foram colocados no fim da lista.

Tabela 3.1: agrupamento de tipos de painéis semelhantes e codificação da variável tipo

Tipo	Painéis
1	MSS
2	QDF / QDCA / QDCC / QDFL / QDL / QDC / QDR
3	PNCLP / PNMED / PNR
4	PNSAT / PNSEC
5	PNPE / PNC / PNP / PNPA / PNE / PNPC / PNTU / PNCS
6	PNDBT
7	PNSA
8	CCMBT
9	PNDMT / PNGMT / PNPMT / CCMMT
10	CCMBT (BACK TO BACK)
11	PNAT / PNST
12	PKS30
13	PNGBT-PE
14	MP / MC

É interessante destacar que todas as variáveis investigadas são discretas. A seguir é feita uma breve descrição destas.

Após as sugestões de variáveis, e posterior avaliação, chegou-se a conclusão de que a variável tipo de pintura não pode ser considerada como impactante no tempo de produção de um painel pelo fato de o processo de pintura ser exatamente o mesmo ao se usar qualquer dos dois tipos de pintura. Também não pode ser levada em consideração a variável número médio de funcionários a trabalhar na montagem do painel, por ser este valor calculado a partir do *lead time* e da quantidade de homens-hora necessária. Somente as demais variáveis devem ser investigadas quanto a seus respectivos poderes de explicação.

4.2 Elaboração do modelo

Todo o desenvolvimento da aplicação prática deste estudo foi feita com referência à base de dados constante no apêndice B.

Neste trabalho, se imaginou dois caminhos para se construir um modelo de regressão estimador de tempos de fabricação. Um deles usando *lead time*, medido em dias, como

variável resposta e outro, usando a quantidade de homens-hora necessária à manufatura do painel como variável resposta.

4.2.1 Testes de normalidade e Identificação da variável resposta

Uma das suposições do modelo de regressão linear é que a variável resposta segue uma distribuição normal de probabilidade. Desta forma, é necessário investigar se os dados seguem esta exigência do modelo, através de testes de aderência. Três testes são freqüentemente utilizados: o de Kolmogorov-Smirnov, Shapiro-Wilk e Lilliefors. Neste trabalho foram utilizados os testes de Kolmogorov-Smirnov e Lilliefors como critérios de aprovação da aproximação dos dados à distribuição gaussiana. Para ambos, o indicador do teste, p , que é apresentado acima dos histogramas, deve ser maior que 0,2 para uma aderência aceitável. Não necessariamente os três testes chegam ao mesmo resultado.

Deseja-se, então, investigar qual candidata a variável resposta, ou sua transformação, deverá ser adotada na análise de acordo com seu desempenho nos testes de aderência. A primeira a ser testada, foi a variável *lead time*, conforme figura 4.1.

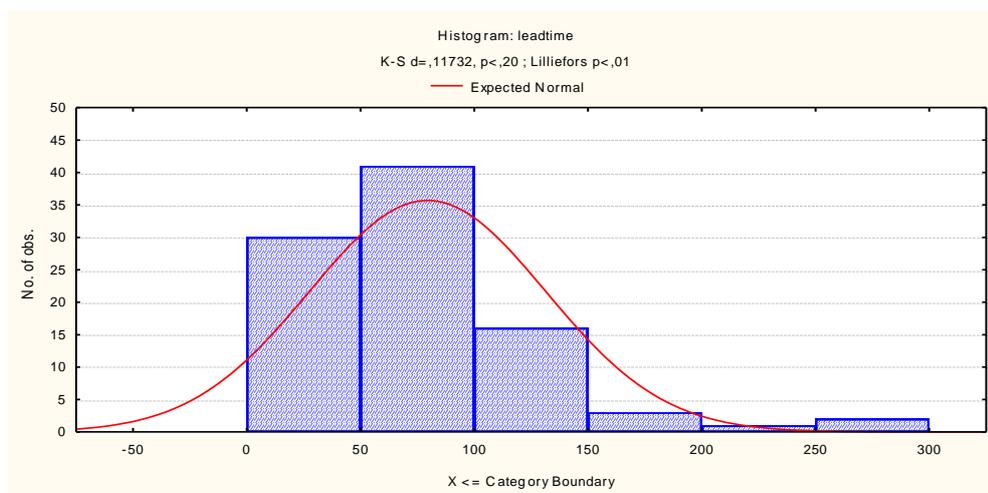


Figura 4.1: Distribuição da variável *lead time*

Percebe-se que a distribuição dos valores do *lead time* não passou nos testes de aderência e, portanto, não se pode assumir que a variável *lead time* segue distribuição normal.

Outra maneira de se verificar normalidade é utilizando o gráfico de probabilidade normal (normal-plot) que apresenta os valores esperados da distribuição normal, contra as observações da variável investigada. Quanto mais próximos estiverem os pontos da reta apresentada, mais próximo de seguir uma distribuição normal estará a variável analisada. Um gráfico de distribuição normal para *lead time* é apresentado na figura 4.2.

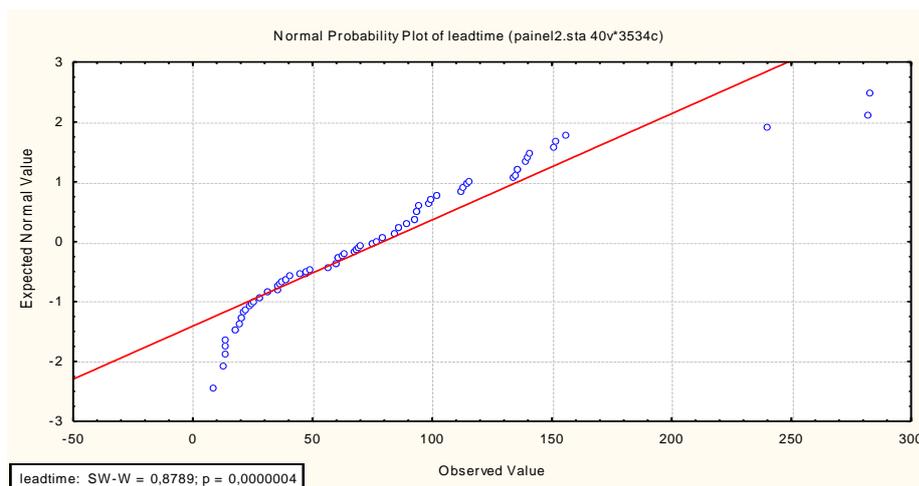


Figura 4.2: gráfico de probabilidade normal para lead time

Verifica-se que não há uma boa aproximação da distribuição à normal.

Testando-se a transformação raiz quadrada no *lead time*, obtêm-se os resultados indicados nas figuras 4.3 e 4.4.

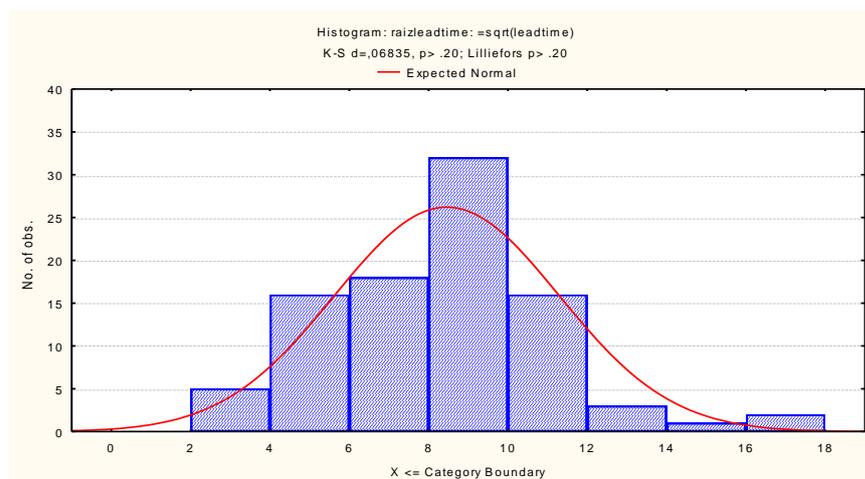


Figura 4.3: distribuição da variável raiz de lead time

Neste caso, a distribuição está mais próxima à normal, e a hipótese de a variável raiz de *lead time* seguir uma distribuição normal foi aceita nos dois testes. Verifica-se no gráfico da figura 4.4 que os pontos estão mais alinhados à reta.

Testando-se a variável homem-hora, como indicam as figuras 4.5 e 4.6, verifica-se que está distante a condição de normalidade para esta candidata, não há boa aproximação à distribuição normal.

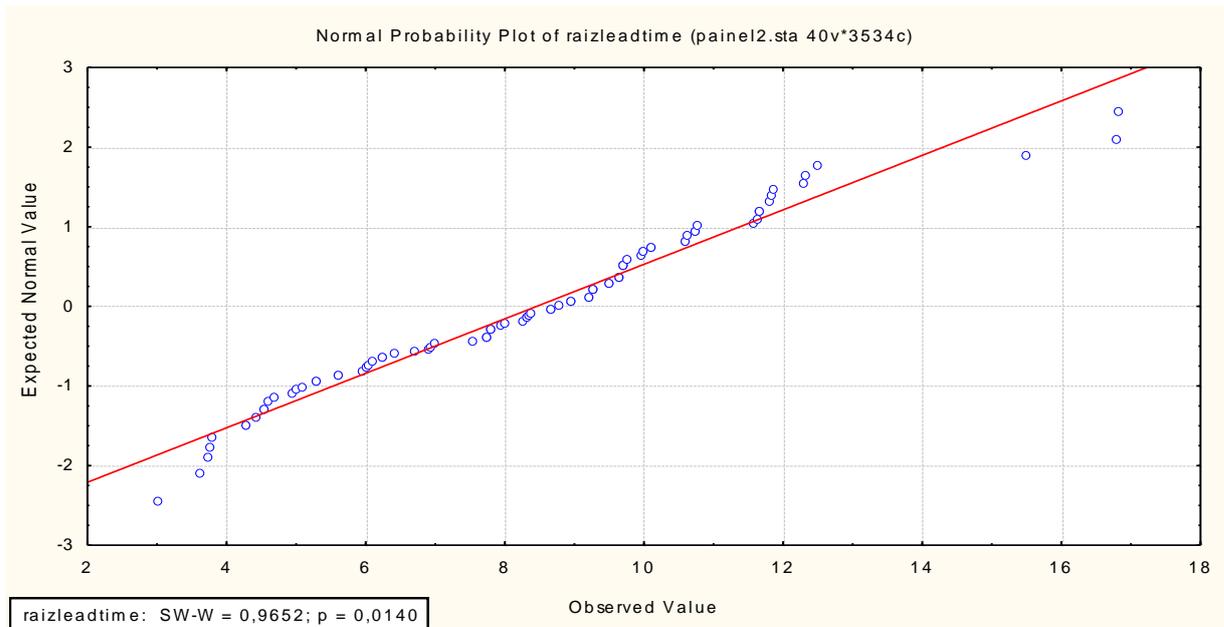


Figura 4.4: gráfico de probabilidade normal para raiz de lead time

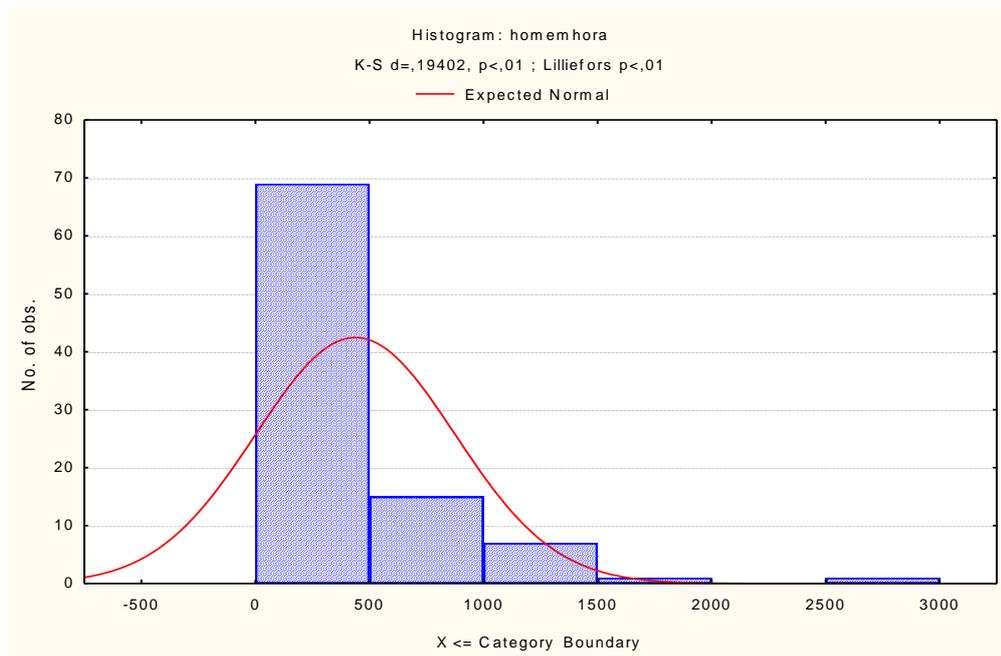


Figura 4.5: distribuição da variável homem-hora

Porém, ao tomar-se o logaritmo natural da variável homem hora, verifica-se um melhor resultado, como mostrado nas figuras 4.7 e 4.8.

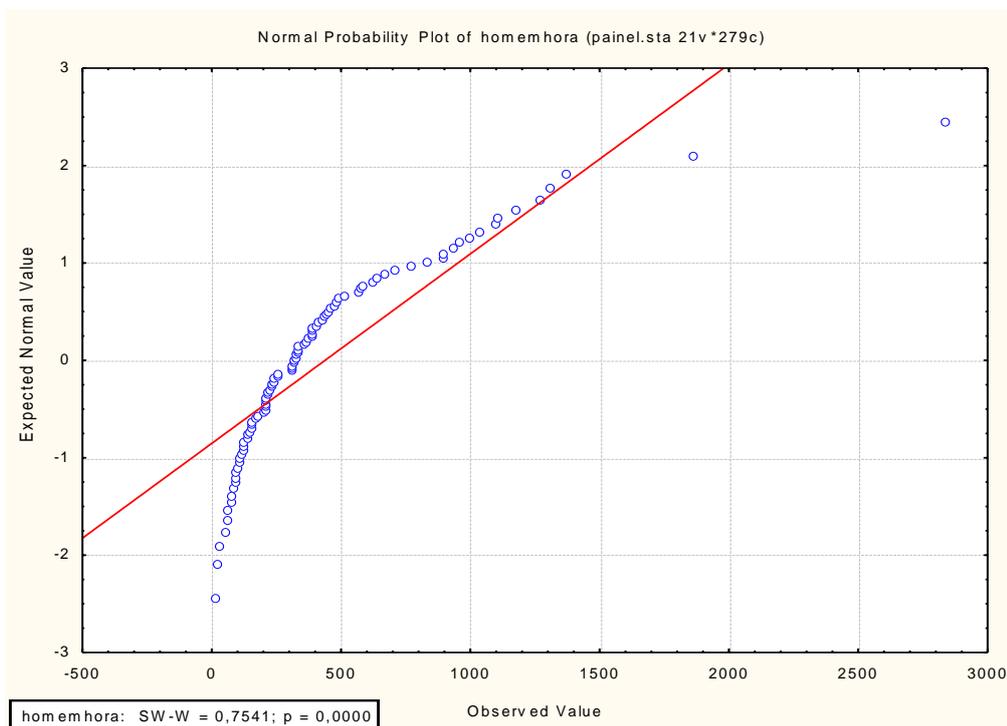


Figura 4.6: gráfico de probabilidade normal para homem-hora

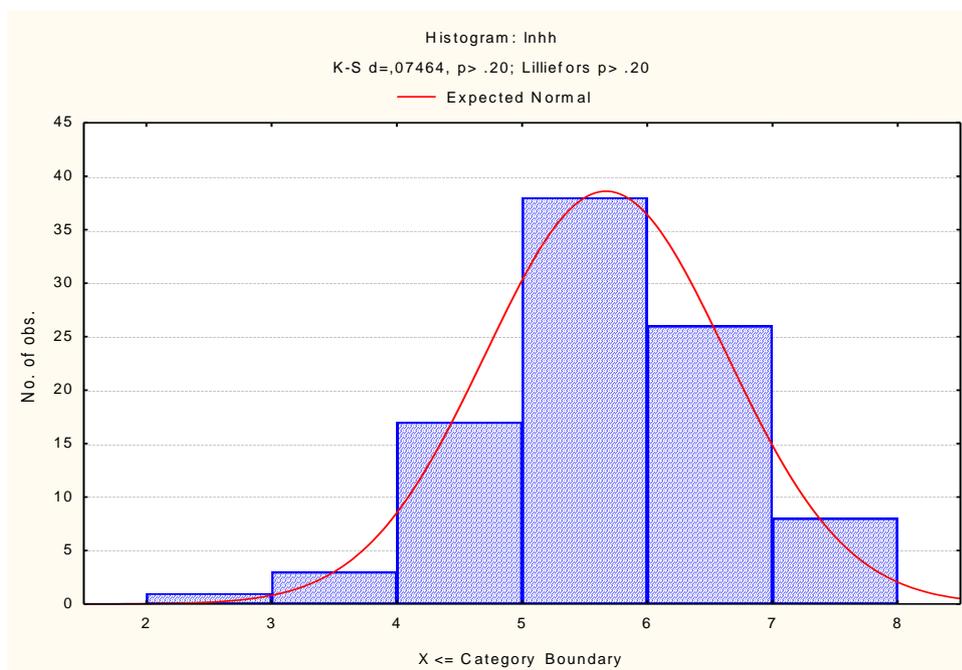


Figura 4.7: distribuição da variável $\ln(\text{homem-hora})$

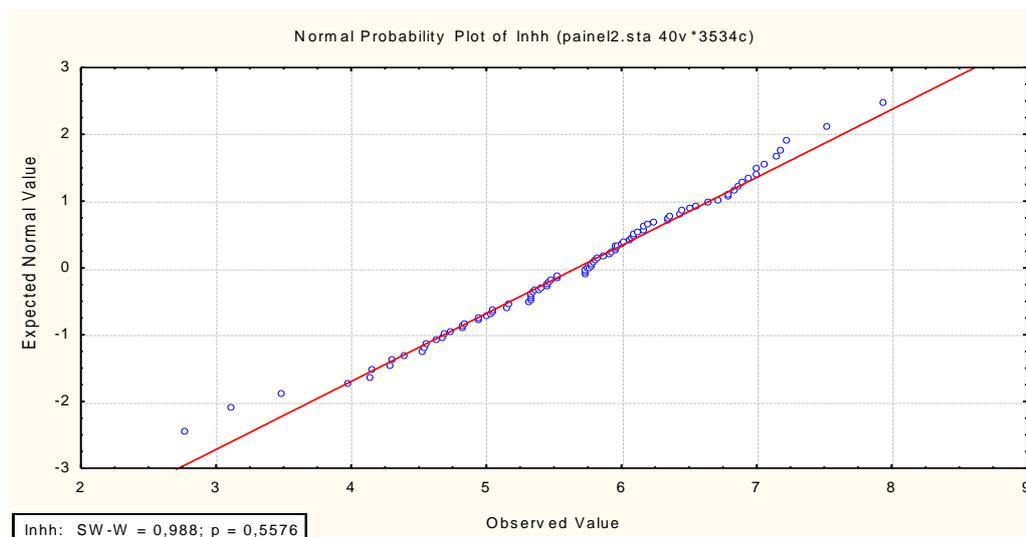


Figura 4.8: gráfico de probabilidade normal para $\ln(\text{homem-hora})$

Há uma boa aproximação à condição de normalidade, ainda melhor que na variável raiz de *lead time*. Além disso, é mais interessante escolher como variável resposta a quantidade de homens-hora necessária à produção do painel, ou suas transformações, por ser um tempo apenas de produção. Parte do *lead time* em algumas observações pode representar tempo ocioso, onde o painel encontra-se parado na fábrica, sem sofrer montagem, devido a atrasos de material, urgência em outros painéis, etc., o que pode causar mais variabilidade nos dados. Tendo uma estimativa da quantidade de homem-hora necessária a montagem, pode-se manipular a quantidade de mão-de-obra adotada e estimar o *lead time*, supondo que não haverá paradas no processo.

4.2.2 Histogramas de candidatas a variáveis explicativas

É interessante montar um histograma para investigar a quantidade de observações de cada nível para cada variável.

Investigando-se a variável tipo, percebe-se que há uma grande concentração de observações de painéis dos tipos 3, 4, 5 e 8, como indica o histograma da figura 4.9. Desta maneira o modelo terá grandes chances de ser mais representativo para observações destes tipos.

Construindo-se um histograma da variável cubículos, verifica-se uma grande concentração de observações de painéis com apenas 1 cubículo, conforme figura 4.10.

É possível tentar contornar este problema utilizando-se uma variável chamada dicotômica, dunning, ou ainda, simulada. Neste caso criou-se a variável cubiculo01. Quando

a observação for painel com 1 cubículo, a variável assumirá valor 1, caso contrário assumirá valor 0. Desta forma, quando o painel possuir 1 cubículo, a média da variável resposta será estimada por um modelo, caso contrário, será estimada por outro, tendo o coeficiente da variável, a função de deslocar a média estimada pelo modelo, para mais, ou para menos.

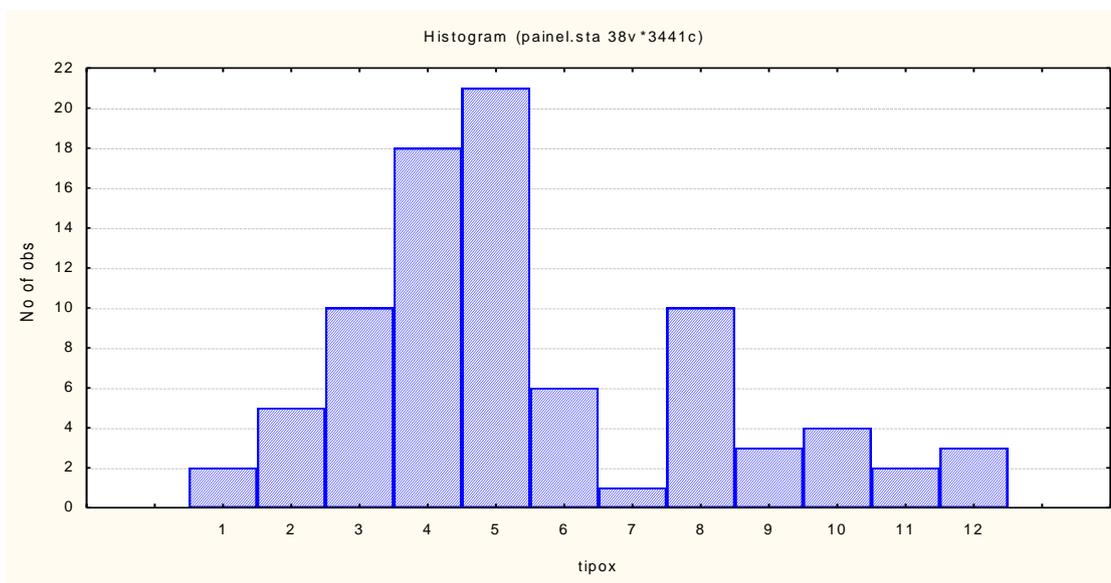


Figura 4.9: histograma da variável tipo

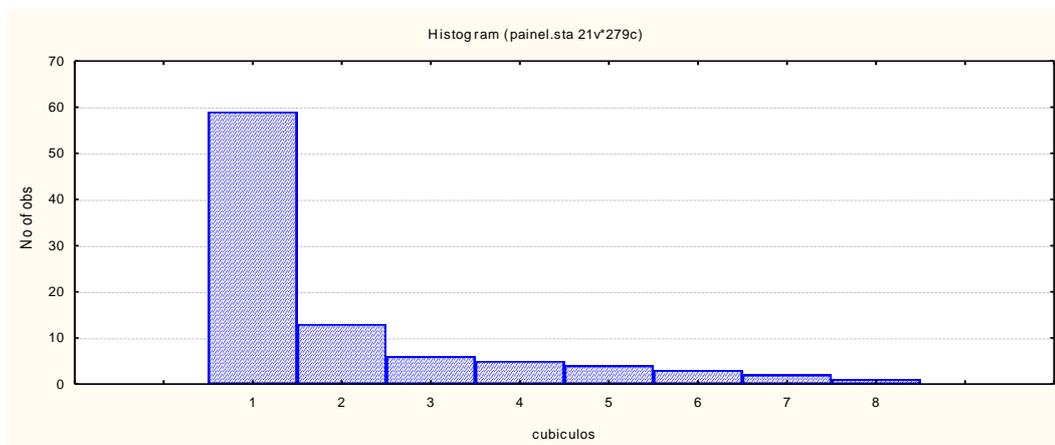


Figura 4.10: histograma da variável cubículos

Com a transformação as observações de número de cubículos ficam mais equilibradas, como indica a figura 4.11.

Conforme ilustra o histograma da figura 4.12, há uma grande concentração de observações de quantidade de itens requisitados entre 37 e 769,6. Possivelmente justifica-se pela grande quantidade de observações de painéis com apenas 1 cubículo (talvez haja

correlação entre número cubículos e requisições, mas nada se pode afirmar com relação a isto apenas observando-se o histograma).



Figura 4.11: histograma da variável cubículo01

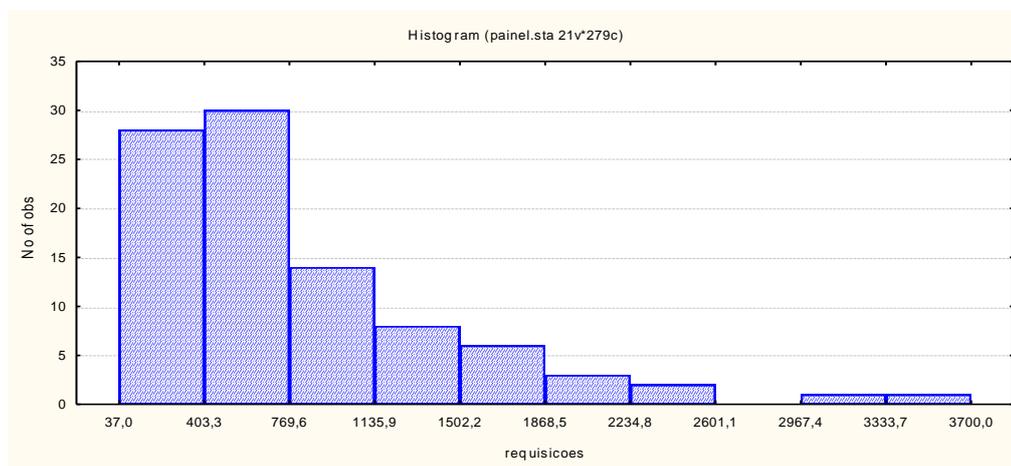


Figura 4.12: histograma da variável cubículos

É praticamente insignificante a quantidade de observações de painéis (ao menos nesta amostra) que não possuem barramento, como apresentado na figura 4.13. Por isso o modelo, a priori, deverá ser montado sem considerar esta variável.

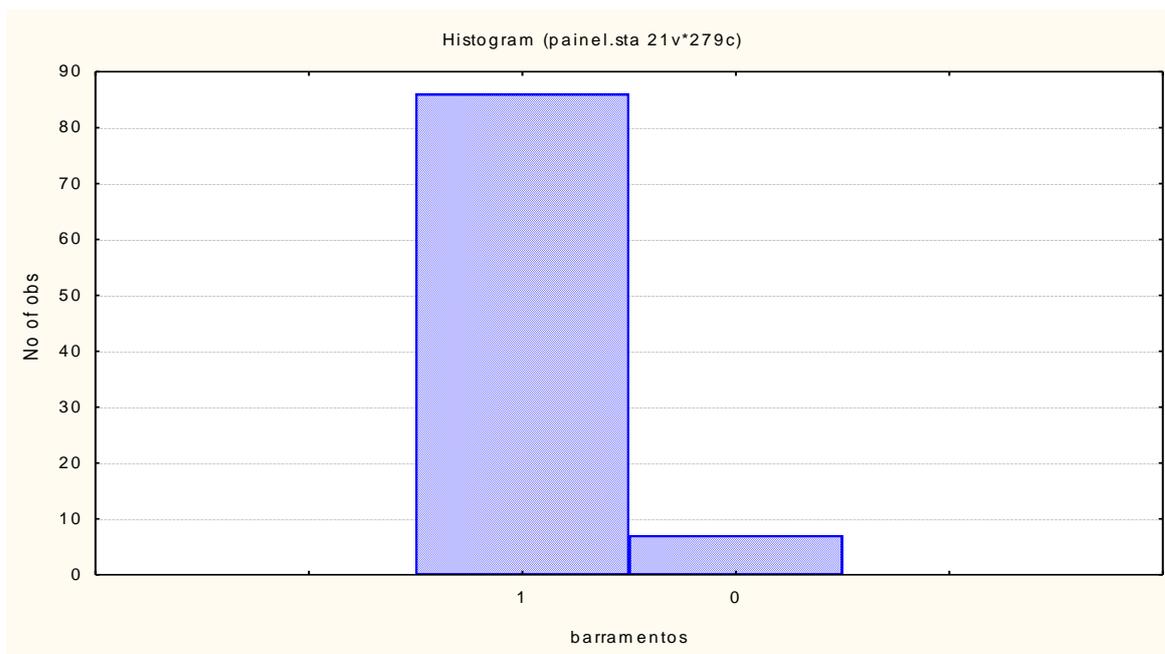


Figura 4.13: histograma da variável barramentos

Verifica-se na figura 4.14 que a quantidade observada de painéis com automação está bem equilibrada com a de painéis sem automação.

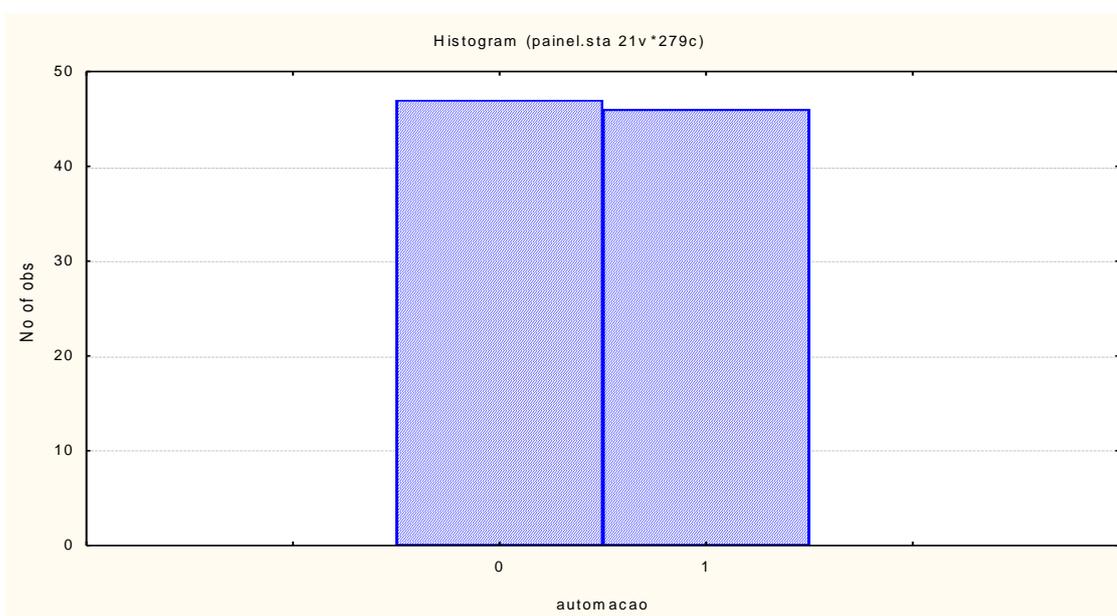


Figura 4.14: histograma da variável automação

O histograma da figura 4.15 indica uma grande concentração de observações de quantidade de pontos de ligação entre 0 e 500. Possivelmente justifica-se pela grande quantidade de observações de painéis com apenas um cubículo, mas não se pode assegurar verdadeira, esta hipótese.

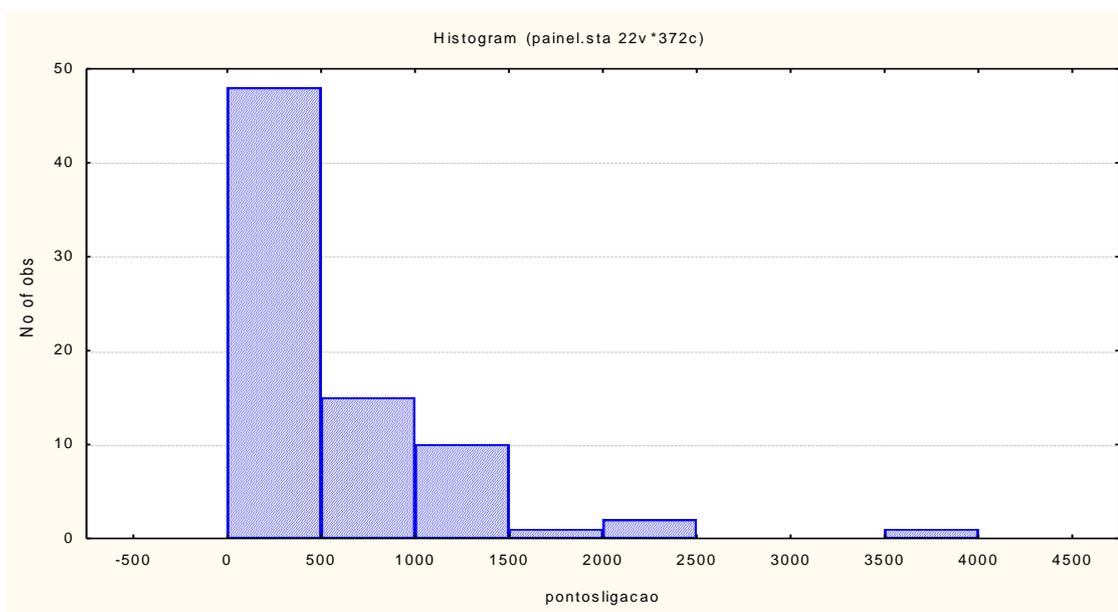


Figura 4.15: histograma da variável pontos de ligação

A maior parte dos painéis observados pertence à classe de baixa tensão, como ilustra a figura 4.16.

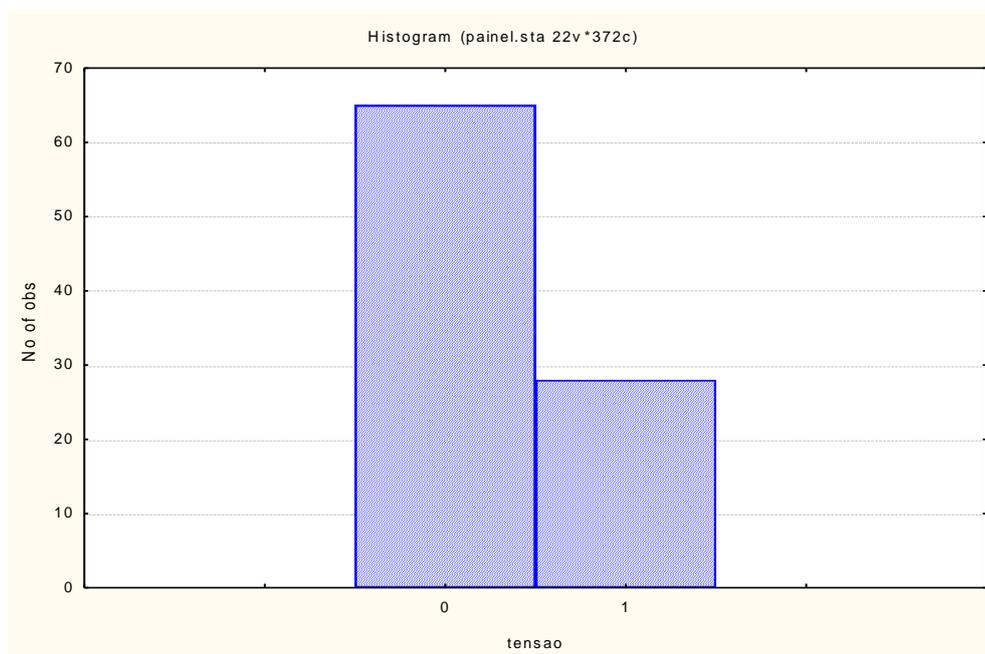


Figura 4.16: histograma da variável tensão

A observação do histograma 4.17 está de acordo com a realidade, a quantidade de painéis elétricos montados em terceirizadas é menor que a de painéis montados internamente, é natural que se colhendo aleatoriamente uma amostra, esta contenha mais observações de montagem interna do que em terceiros.



Figura 4.17: histograma da variável terceiros

4.2.3 Impacto das candidatas a variáveis independentes na média de $\ln(\text{homem-hora})$

Antes de se tentar montar um modelo, é possível investigar se as candidatas a variáveis explicativas parecem causar efeito significativo na média da variável resposta. É útil para tal análise, dispor os dados em um gráfico do tipo *box-plot* com representação da média, primeiro quartil, terceiro quartil, mínimo e máximo da amostra, de forma a ilustrar graficamente a dispersão dos dados.

O *box-plot* da figura 4.18 representa a média de $\ln(\text{hh})$ para a amostra de painéis que não possuem barramentos, e para a amostra dos que possui. Percebe-se que as médias são diferentes (a região de dispersão dos dados com barramentos, não invade a dos sem barramentos), porém caso outras variáveis explicativas se mostrem muito mais impactantes na média, esta variável perde força no poder de explicação do modelo.

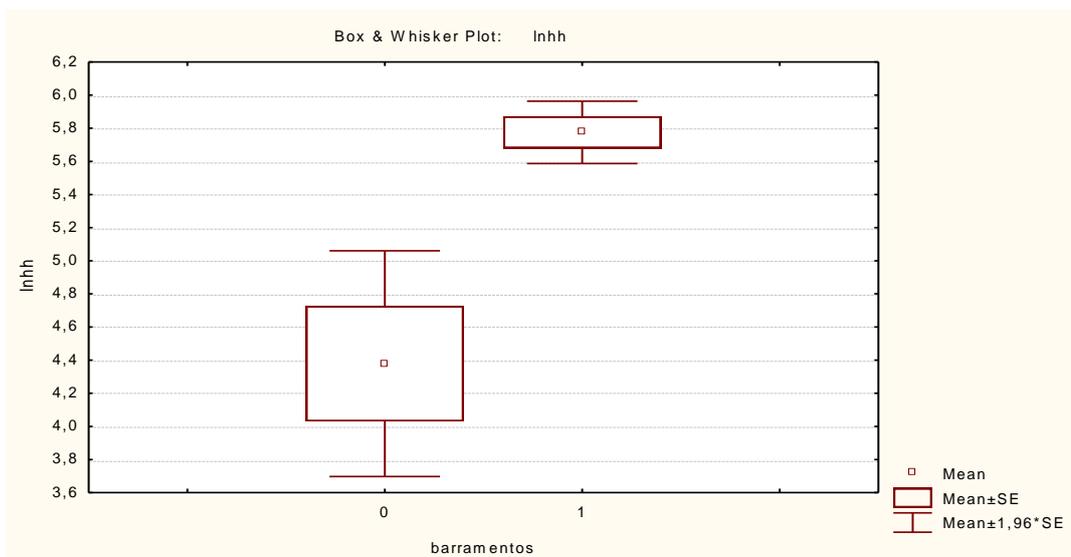


Figura 4.18: box-plot de ln(homem-hora) versus barramentos

Conforme indica a figura 4.19, a média amostral de ln(hh) para painéis que possuem automação é maior que a média dos que não possuem, porém não se pode assegurar que as médias populacionais são diferentes, devido à dispersão dos dois tipos de observações possuírem região em comum, inclusive percebe-se que o terceiro quartil dos sem automação é maior que o primeiro quartil dos com automação.

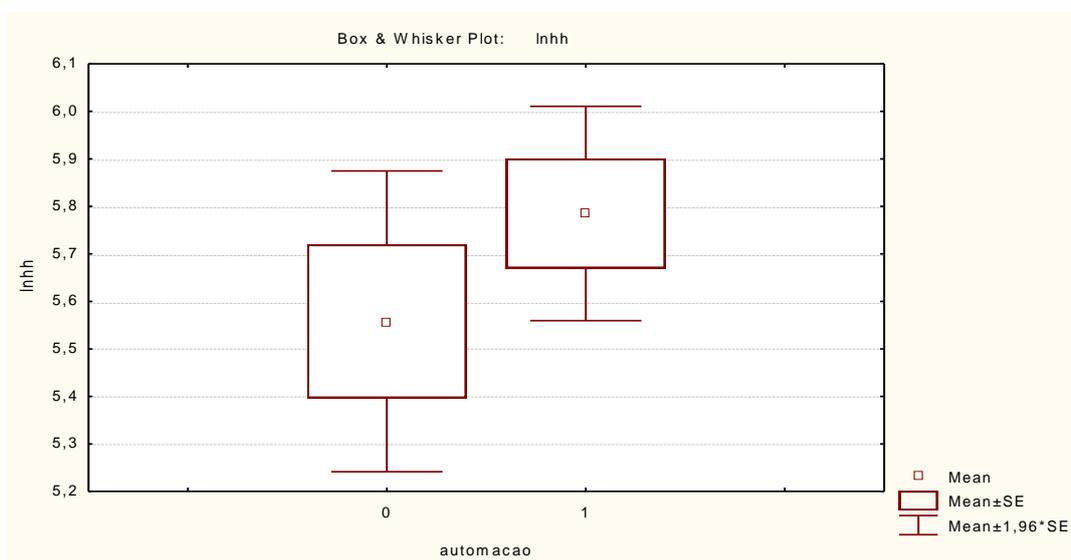


Figura 4.19: box-plot de ln(homem-hora) versus automação

No caso da tensão de operação do painel, alta ou baixa, há mais evidências de que as médias de ln(hh) para os dois tipos de observação são diferentes, como mostrados na figura 4.20.

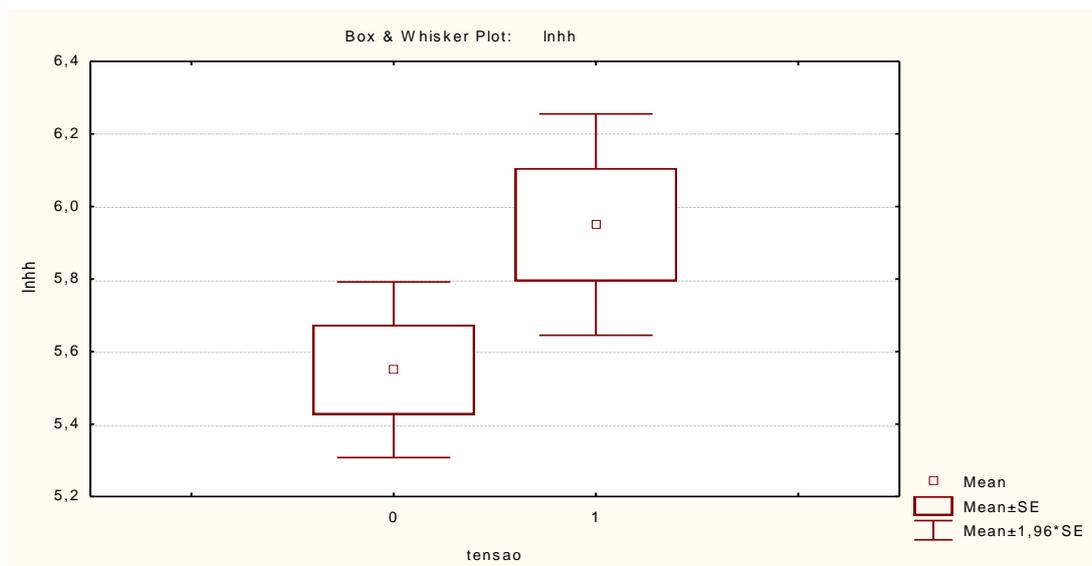


Figura 4.20: box-plot de ln(homem-hora) versus tensão

É importante lembrar que está se investigando se as variáveis independentes são responsáveis por variações na média de ln(homem-hora). Caso a montagem elétrica seja terceirizada, é natural a necessidade de menos mão-de-obra interna, pois o painel não passará em um dos setores, e se estará estimando a mão de obra necessária excluindo-se a montagem elétrica, que é feita nos terceiros, mas não pode ser feita em paralelo a outras etapas. Por isso no caso de montagem elétrica em terceiros será necessário questionar aos mesmos a quantidade de mão-de-obra necessária em homem-hora (caso a variável seja incluída no modelo), para somá-la a variável resposta, e só então se estará diante da estimativa de homens-hora total necessários à manufatura do painel. Um *box-plot* de ln(homem-hora) é mostrado na figura 4.21.

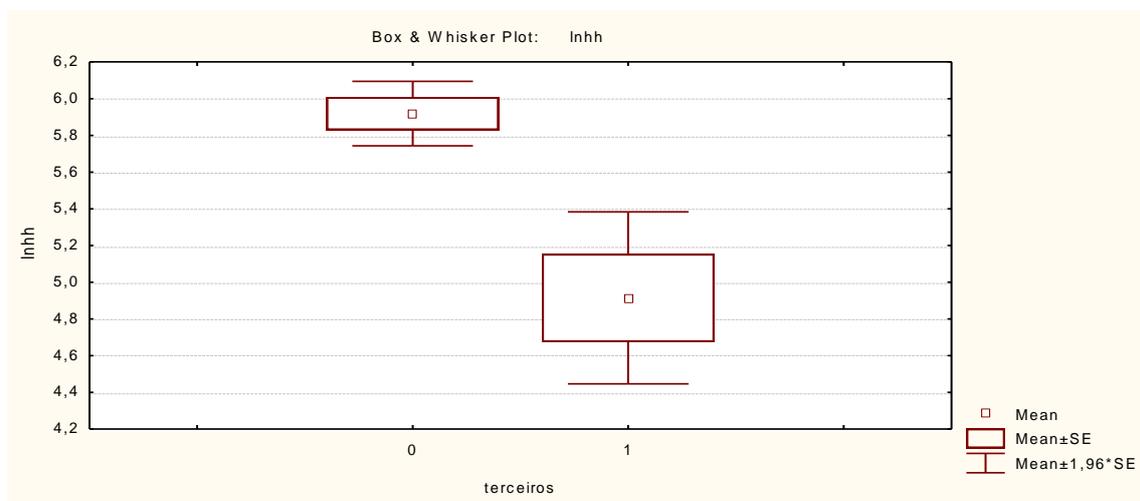


Figura 4.21: box-plot de ln(homem-hora) versus terceiros

Percebe-se que a média de ln(hh) pode variar conforme se observem tipos diferentes de painéis (figura 4.22).

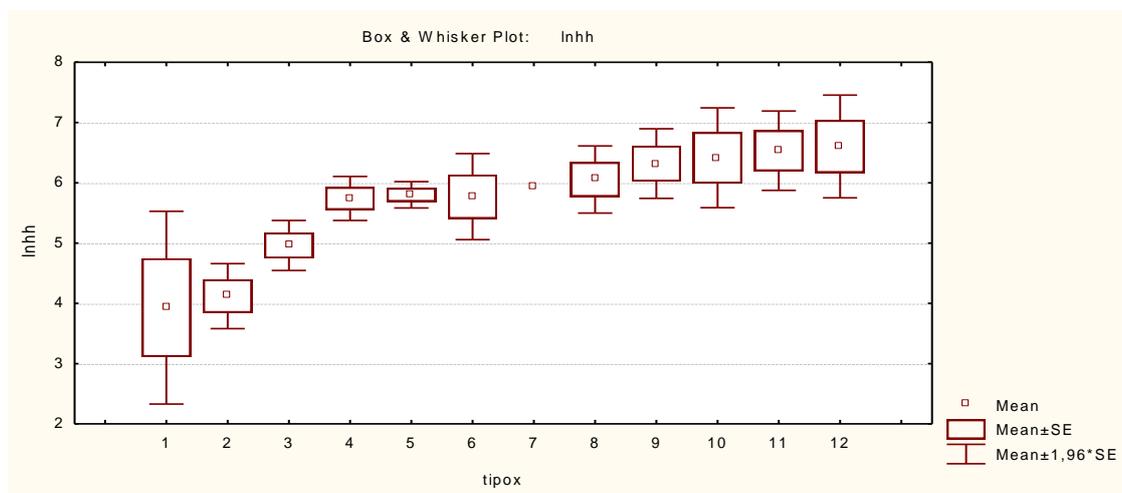


Figura 4.22: box-plot de ln(homem-hora) versus tipo

4.3 Aplicação do modelo de regressão linear múltipla

Vários modelos de regressão linear foram construídos e testados após as investigações feitas nos itens anteriores e entre estes, o que apresentou melhores resultados, como poder de explicação e significância de parâmetros, foi o modelo com variável resposta ln(hh) e variáveis explicativas: cubículos, terceiros e tipo, como mostrado na tabela 4.1.

Tabela 4.1: Resultados do modelo

Regression Summary for Dependent Variable: ln(hh) (painel3.sta)						
R= ,78336774 R²= ,61366502 Adjusted R²= ,60064249						
F(3,89)=47,123 p<,00000 Std.Error of estimate: ,60732						
N=93	Beta	Std.Err. of Beta	B	Std.Err. of B	t(89)	p-level
Intercept			4,89549	0,156963	31,18886	0,000000
cubiculos	0,480549	0,078409	0,27745	0,045270	6,12871	0,000000
terceiros	-0,497427	0,068148	-1,10201	0,150977	-7,29924	0,000000
tipox	0,245737	0,078010	0,09074	0,028805	3,15007	0,002223

Verifica-se que com modelo:

A expressão matemática que estima a quantidade de homens-hora necessária é:

$$\ln(hh) = 4,89549 + 0,27745 \cdot X_1 - 1,10201 \cdot X_2 + 0,09074 \cdot X_3 \tag{4.1}$$

$$\text{ou: } hh = \exp(4,89549 + 0,27745 \cdot X_1 - 1,10201 \cdot X_2 + 0,09074 \cdot X_3). \tag{4.2}$$

Com X_1 sendo quantidade de cubículos, X_2 sendo tipo de montagem elétrica e X_3 sendo o tipo de painel, conforme descrição anterior.

O coeficiente de X_1 indica que, fixando-se as outras variáveis, para o aumento de uma unidade no número de cubículos de um painel, há um aumento no logaritmo natural da mão-de-obra necessária de 0,277745.

O coeficiente de X_2 representa a redução na média do logaritmo natural da mão-de-obra necessária quando o painel sofre montagem elétrica terceirizada, fixando-se as demais variáveis.

O coeficiente de X_3 representa o aumento na média do logaritmo natural da mão-de-obra necessária, ao se variar de um tipo para seu consecutivo (do menor para o maior índice de tipo), fixando-se as demais variáveis.

O valor da variável resposta quando forem nulas todas as variáveis explicativas é 4,89549 – neste caso este valor não possui interpretação, uma vez que deve ser nula, a mão-de-obra necessária para este tipo de observação (quando não se está fabricando painel algum).

De acordo com os testes de significância realizados pelo software, as três variáveis explicativas mostram-se significantes para o modelo ao nível de 5% de significância.

Encontrou-se $R^2 = 0,61366502$, o que indica que aproximadamente 61,37% da variabilidade do logaritmo natural de homem-hora, pode ser explicada por variações nas variáveis cubículos, terceiros e tipo.

Seu valor ajustado, $R_a^2 = 0,60064249$, tem o mesmo significado de R^2 , porém levando em consideração o tamanho da amostra. De acordo com esse dado, 60,06% da variabilidade do logaritmo natural de homem-hora, pode ser explicada por variações nas variáveis cubículos, terceiros e tipo.

Um gráfico *box-plot* de $\ln(hh)$ contra tipo de painel, como mostrado na figura 4.23, revela a existência de alguns pontos discrepantes que podem interferir negativamente na análise. Estes pontos são os chamados *outliers* e extremos.

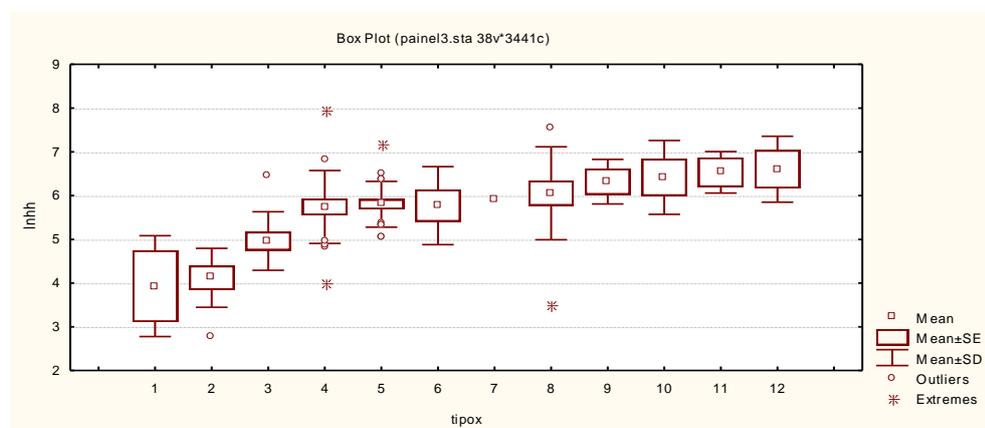


Figura 4.23: *box-plot* de $\ln(\text{homem-hora})$ versus *tipo*, com *outliers* e pontos extremos

Depois da retirada destas observações discrepantes, uma regressão de $\ln(hh)$ sobre as mesmas variáveis explicativas fornece os resultados indicados na tabela 4.2.

Tabela 4.2: resultados do modelo sem outliers e extremos da variável tipo

Regression Summary for Dependent Variable: lnhh (painel.sta)						
R= ,83365085 R ² = ,69497374 Adjusted R ² = ,68367647						
F(3,81)=61,517 p<0,0000 Std.Error of estimate: ,48658						
N=85	Beta	Std.Err. of Beta	B	Std.Err. of B	t(81)	p-level
Intercept			4,862307	0,130318	37,31098	0,000000
cubiculos	0,483071	0,073581	0,256508	0,039071	6,56520	0,000000
terceiros	-0,457493	0,063193	-0,927584	0,128126	-7,23960	0,000000
tipox	0,318682	0,073748	0,104811	0,024255	4,32123	0,000044

Para este novo modelo, a expressão matemática que estima a quantidade necessária de homens-hora é:

$$\ln(hh) = 4,862307 + 0,256508 \cdot X_1 - 0,927584 \cdot X_2 + 0,104811 \cdot X_3 \quad (4.3)$$

$$\text{ou: } hh = \exp(4,862307 + 0,256508 \cdot X_1 - 0,927584 \cdot X_2 + 0,104811 \cdot X_3) \quad (4.4)$$

O coeficiente de X_1 indica que, fixando-se as outras variáveis, para o aumento de uma unidade no número de cubículos de um painel, há um aumento no logaritmo natural da mão-de-obra necessária de 0,256508.

O valor da variável resposta quando forem nulas todas as variáveis explicativas é 4,862307 – neste caso este valor não possui interpretação, uma vez que deve ser nula, a mão-de-obra necessária para este tipo de observação (quando não se está fabricando painel algum).

De acordo com os testes de significância realizados pelo software, as três variáveis explicativas mostram-se significantes para o modelo ao nível de 5% de significância.

Encontrou-se $R^2 = 0,69497374$, o que indica que aproximadamente 69,50% da variabilidade do logaritmo natural de homem-hora, pode ser explicada por variações nas variáveis cubículos, terceiros e tipo.

Seu valor ajustado, $R_a^2 = 0,68367647$, tem o mesmo significado de R^2 , porém levando em consideração o tamanho da amostra. De acordo com esse dado, 68,37% da variabilidade do logaritmo natural de homem-hora, pode ser explicada por variações nas variáveis cubículos, terceiros e tipo.

Pela comparação descrita acima, o segundo modelo se mostra mais eficiente na estimação de mão-de-obra necessária a manufatura de painéis.

Uma das suposições do modelo de regressão linear múltipla é que os resíduos são normalmente distribuídos. Um teste de normalidade como foi feito na identificação da

variável resposta foi feito para os resíduos do modelo final, comparando a distribuição dos resíduos com uma distribuição normal, como na figura 4.24, afirma a normalidade dos resíduos. O gráfico de resíduos contra valores preditos, não demonstra padrão tão bom quanto desejado, como mostrado na figura 4.25, porém as outras análises feitas dão força ao modelo desenvolvido.

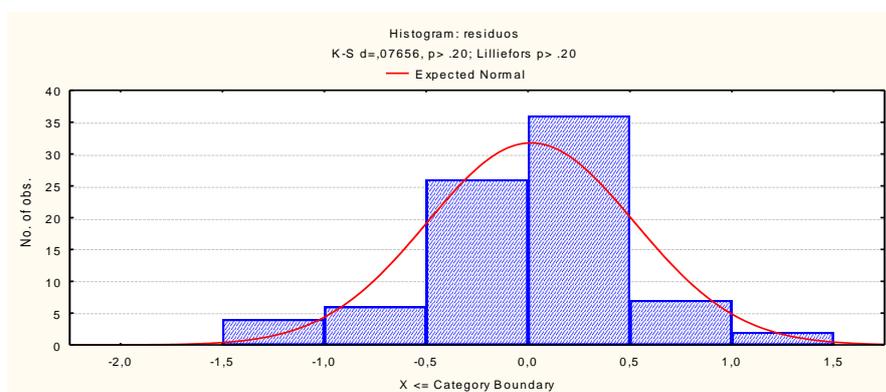


Figura 4.24: Normalidade dos resíduos

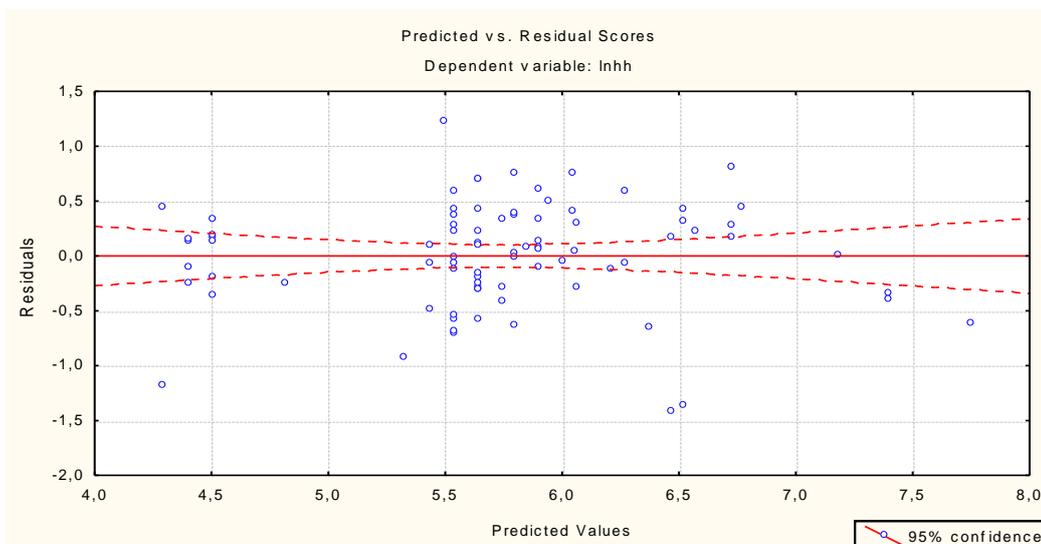


Figura 4.25: Resíduos x valores preditos

Apesar dos bons resultados, o modelo não pode ser utilizado para estimar tempos de fabricação de painéis com quantidade de cubículos superior a oito, nem de painéis dos tipos 13 e 14, uma vez que nenhuma observação destes casos foi colhida. Pelo que apontam os histogramas apresentados nas figuras 4.9, 4.10 e 1.17, O modelo é mais representativo para

painéis com número de cubículos próximo de 1, dos tipos 3, 4, 5 e 8 e com montagem elétrica interna.

5 CONCLUSÕES

Este capítulo encerra o trabalho, identificando os pontos positivos, as limitações do modelo encontrado, as dificuldades encontradas no desenvolvimento do trabalho e sugerindo propostas para estudos futuros.

5.1 Aspectos positivos e limitações do trabalho

A idéia inicial deste trabalho foi utilizar um modelo de regressão para estimar diretamente o tempo de fabricação de painéis. Porém ao se tentar utilizar o *lead time* como variável resposta, esta e suas transformações se mostraram inadequadas, pelo fato de não satisfazerem o pressuposto de serem normalmente distribuídas. Provavelmente, esta não adequação se deve ao fato de alguns painéis passarem algum tempo aguardando chegada de material e outros não – o que pode gerar muita variabilidade.

O trabalho revela que pode ser adequada a estimação de tempos de fabricação, para esta amostra, através do valor fornecido pelo modelo de regressão com variável resposta $\ln(\text{homem-hora})$, tomando-se a exponencial deste valor e dividido-a pelo número médio de funcionários disponíveis para o painel cujo tempo de fabricação se está estimando. É importante lembrar que diversas vezes não é possível aumentar muito a quantidade de mão-de-obra aplicada à manufatura do painel – na maioria das etapas do processo fabril em estudo, é impraticável operar com mais de dois funcionários por cubículo, à exceção apenas da pintura, que ocorre com quatro funcionários.

O modelo apresentado neste trabalho foi desenvolvido com a base de dados disponível, extraíndo-a do sistema de informação da empresa e de planilhas de controle operacional de atividades da fábrica, colhendo observações em um intervalo de aproximadamente três meses, e filtrando-as por painéis que haviam sido completamente embarcados ao cliente (o envio dos painéis pode ocorrer de forma fracionada). Alguns dados referentes ao número de pontos de ligação não foram colhidos pela sua não disponibilidade.

Haveria maior facilidade de se coletar dados para gerar modelos de regressão, caso o sistema de informação fosse mais direcionado à análise, fornecendo informações resumidas em um relatório com observações das variáveis candidatas no modelo. Houve demora na coleta de alguns dados referentes a certas variáveis, em especial automação e pontos de ligação, pelo fato da visualização só ser possível observando os projetos dos painéis. Tal problema seria contornado se ao elaborar cada projeto, os responsáveis alimentassem uma

planilha com a ordem de produção do painel e os respectivos valores das variáveis ligadas a suas atividades.

5.2 Sugestões para futuros trabalhos

Para uma análise mais precisa dos dados, seria interessante, em estudo futuro, realizar uma nova coleta de dados, a fim de tentar equilibrar as quantidades observadas em cada nível, de cada variável independente, de forma a se ter maior segurança na análise dos dados e maior capacidade, do modelo a se definir, em retratar o comportamento da população dos dados. Além disto, outros modelos de regressão, a exemplo do polinomial podem ser testados quanto à suas capacidades de estimar o comportamento da variável resposta.

REFERÊNCIAS BIBLIOGRÁFICAS

- ALLARD, R. J. *Introdução à Econometria*. Rio de Janeiro: Editora Guanabara Dois S.A., 1980.
- ANDER-EGG, E; IDÁÑEZ, M. J. A. *Cómo elaborar un proyecto – guía para diseñar proyectos sociales y culturales*. Madrid: Instituto de Ciências Sociales Aplicadas, 1997.
- ANDERSON, D. R.; SWEENEY, D. J.; WILLIAMS, T. A. *Estatística Aplicada à administração e Economia*. 2. ed. São Paulo: Cengage Learning, 2007.
- ANZIEU, D.; MARTIN, J. *La Dynamique des Groupes Restreints*, Paris, P.U.F., 1968.
- BERENSON, M. L. et al. *Basic Business Statistics – concepts and applications*. Prentice Hall, 2004.
- BUSSAB, W. O. *Análise de Variância e de Regressão*. 2. ed. São Paulo: Editora Atual, 1940.
- DANTAS, R. A. *Engenharia de Avaliações – uma introdução à metodologia científica*. 1. ed. São Paulo: Pini, 1998.
- Élin Duxus – Soluções inteligentes para suas idéias. Disponível em: < http://www.duxus.com.br/produtos/tecnologia/addin/metrixus/manual/index2_2_12.php > Acesso em: 17 de maio de 2009.
- HOSMER, D. W.; LEMESHOW, S. *Basic Applied Logistic Regression*. Nova Iorque: John Wiley & Sons, 1989.
- LACHOS, V. H. *Modelos Log - Linear e Polinomial*. Disponível em: < www.ime.unicamp.br/~hlachos/RegreLogLog.pdf > Acesso em: 04 de maio de 2009.
- OGLIARI, P. J. *Regressão Logística*. Disponível em: < www.inf.ufsc.br/~ogliari/arquivos/regressao_logistica.ppt > Acesso em: 04 de maio de 2009.
- OZTURK, A.; KAYALIGIL, S.; OZDEMIREL, N. E. Manufacturing *lead time* estimation using data mining. *European Journal of Operational Research*, 2006. Volume 173.
- SELLTIZ, C.; JAHODA, M.; DEUTSCH, M.; COOK, S. *Métodos de pesquisa nas relações sociais*. São Paulo: EPU, 1975.
- SIAS - Sistema de informação aplicada à saúde. Disponível em: < http://www2.inf.furb.br/sias/saude/Textos/desvio_padrao.htm > Acesso em: 04 de maio de 2009.
- SLACK, N. *Administração da Produção*. 2. ed. São Paulo: Editora Atlas, 2002.
- STEWART, J. *Cálculo*. 5. ed. São Paulo: Cengage Learning, 2006. Volume I.
- STEWART, J. *Cálculo*. 5. ed. São Paulo: Cengage Learning, 2006. Volume II.

APÊNDICE A

O mix de Painéis

Os painéis elétricos da AREVA KOBLITZ dividem-se em:

- Painéis de média tensão
 - Painel de Força em Média Tensão (principal)
 - PNDMT
 - Distribuição Geral em MT
 - Entrada de Energia MT
 - Despachos MT
 - PNGMT – Painel de Força em Média Tensão
 - PNPMT – Partida para Motor
 - CCMMT – Centro de Controle de Motores em Média Tensão
 - Painel de Força em média tensão (auxiliar)
 - PNST – Surto de Tensão para motor / gerador
 - PNAT – Aterramento para Gerador / Transformador
 - PNSEC – Seccionamento MT
 - PNSAT – Surto de Tensão e Aterramento para Gerador
 - Painel de Força em média tensão (PKS-30)
 - PNDMT
 - Distribuição Geral em MT
 - Entrada de Energia MT
 - Despachos MT
 - PNGMT – Painel de Força em Média Tensão
 - PNPMT – Partida para Motor
 - CCMMT – Centro de Controle de Motores em Média Tensão
- Painéis de Baixa Tensão
 - Painel de força em Baixa Tensão
 - PNDBT
 - Distribuição Geral em BT
 - Entrada de Energia BT
 - Despacho(s) BT

- PNGBT – Disjunção de Força para Gerador
- PNGBTPE – Disjunção de Força para Gerador – Proteção e excitação
- Centro de Controle de Motores em Baixa Tensão
 - CCMBT – Centro de controle de Motores em BT – Fabricação
 - PNFP – Correção de Fator de Potência
 - PNPBT – Partida para motor
 - CCMBT – Centro de controle de Motores em BT – Fabricação back-to-back
 - CCMBT – Centro de controle de Motores em BT – Fabricação gaveta
- Painel de Comando
 - PNCLP
 - PNMED – Medição para faturamento
 - PNR – Painel de Remota
 - PNC – Comando
 - PNP – Proteção
 - PNPA – Proteção e acoplamento
 - PNE – Excitação de Gerador
 - PNPC – Proteção e Comando
 - PNPE – Proteção e Excitação
 - PNTU – Controle de Turbina de Gerador
 - PNCS – Controle do Sincronismo
- Quadro
 - QDCA – Distribuição de Serviços Auxiliares CA
 - QDCC – Distribuição de Serviços Auxiliares CC
 - QDF – Distribuição de Força
 - QDFL – Distribuição de Força e Iluminação
 - QDL – Quadro de iluminação
 - QDC – Quadro de Comando
 - QDR – Remota
 - PNSA – Distribuição de Serviços Auxiliares CA/CC
 - MSS – Mesa para o Sistema de Supervisão
 - MP – Mesa de Paralelismo para Geradores
 - Mesa de Comando

APÊNDICE B

Coleta de dados

Observação	Lead time em dias	Homens-hora	Cubículos	Requisições	Barramento	Automação	Ligações	Tensão	Terceiros	Tipo
1	156,00	140,40	1	164	1	1	345	0	0	5
2	45,00	96,40	1	66	1	0		0	1	6
3	69,50	391,98	2	709	1	1	1001	0	0	5
4	28,00	388,70	2	371	1	0	930	0	0	6
5	94,00	93,60	1	77	0	0	14	0	1	2
6	112,00	81,80	1	87	0	0	47	0	0	2
7	94,00	897,90	3	1201	1	1	1202	1	0	9
8	94,00	309,10	1	436	1	1	392	1	0	12
9	94,00	491,00	1	684	1	0	144	1	0	11
10	18,25	1861,00	4	1584	1	0		0	0	8
11	9,13	1269,60	8	1225	1	0		0	0	8
12	24,33	156,10	3	641	1	0	67	0	0	8
13	18,25	1106,80	4	763	1	0		0	0	8
14	49,00	32,60	1	92	1	0	101	0	1	8
15	80,00	669,00	2	582	1	1	1069	0	0	5
16	21,25	707,70	4	954	1	0	1234	0	1	8
17	100,00	771,90	3	368	1	0	1160	0	0	8
18	22,00	994,40	4	556	1	0		0	0	8
19	60,00	325,00	1	460	1	1	569	1	0	9
20	63,00	583,10	1	447	1	1	383	1	0	9
21	283,00	447,14	5	1362	1	1	1136	0	1	8
22	134,00	2834,40	1	49	1	0	46	1	0	4
23	68,00	16,10	1	63	0	0	13	0	1	2
24	70,00	64,00	1	45	0	0		0	1	2
25	282,00	406,08	7	2241	1	1	2460	0	1	8
26	37,25	175,45	4	798	1	1	688	0	1	8
27	77,00	210,60	1	507	1	1	694	0	0	5
28	31,50	388,00	2	671	1	1	961	0	0	5
29	31,50	415,20	2	607	1	1	961	0	0	5
30	100,00	206,90	1	612	1	1	968	0	0	5
31	93,00	436,00	1	1599	1	0	23	0	0	6

Observação	Lead time	Homens-hora	Cubículos	Requisições	Barramento	Automação	Ligações	Tensão	Terceiros	Tipo
32	240,00	960,00	1	67	1	0	175	1	0	11
33	93,00	206,50	1	1599	1	0	23	0	0	6
34	93,00	309,50	1	230	1	0	309	0	0	8
35	93,00	203,50	1	230	1	0	309	0	0	8
36	90,00	376,65	1	55	0	0	105	0	0	7
37	140,00	1311,80	7	984	1	1	1401	0	0	5
38	152,00	173,28	6	1650	1	1	2389	0	1	10
39	75,00	477,00	2	121	1	0	105	1	0	4
40	61,00	125,05	1	37	1	0	23	1	0	4
41	61,00	126,88	1	37	1	0	23	1	0	4
42	61,00	142,13	1	37	1	0	23	1	0	4
43	85,00	107,95	1	772	1	1	413	0	1	3
44	39,00	572,00	1	1015	1	1		0	0	5
45	99,00	94,05	1	43	0	0	11	0	1	2
46	39,00	232,50	1	967	1	1		0	0	5
47	36,00	570,10	1	307	1	1		0	0	5
48	60,00	318,40	1	1011	1	1		0	0	5
49	116,00	368,88	1	417	1	0	97	1	0	4
50	115,00	391,00	1	417	1	0	93	1	0	4
51	113,00	316,40	1	416	1	0	97	1	0	4
52	112,00	239,68	1	417	1	0	97	1	0	4
53	20,67	894,87	3	702	1	0	95	1	0	4
54	20,67	637,90	3	708	1	0	95	1	0	4
55	60,00	238,90	1	967	1	1		0	0	5
56	57,00	430,30	1	307	1	1		0	0	5
57	115,00	224,60	1	410	1	0	65	1	0	4
58	64,00	208,80	1	968	1	1	889	0	0	5
59	48,00	155,40	1	2001	1	1	880	0	0	5
60	80,00	208,32	1	620	1	1	739	0	0	5
61	136,00	252,96	1	1099	1	1	878	0	0	3
62	26,00	623,84	3	3140	1	1	3920	0	0	3

Observação	Lead time	Homens-hora	Cubículos	Requisições	Barramento	Automação	Ligações	Tensão	Terceiros	Tipo
63	80,00	74,40	1	633	1	1	125	0	1	3
64	136,00	63,24	1	2270	1	1	1722	0	1	3
65	136,00	215,02	1	1249	1	1	925	0	0	3
66	135,00	125,55	1	767	1	1	799	0	1	3
67	13,83	1034,35	6	1324	1	1	1148	0	1	10
68	36,50	831,65	2	1743	1	1		0	1	10
69	25,00	932,33	6	1384	1	1	1360	0	1	10
70	69,00	336,70	2	140	1	0	170	1	0	4
71	69,00	324,50	2	142	1	0	174	1	0	4
72	141,00	110,00	1	1154	1	1	472	0	1	3
73	102,00	157,40	1	463	1	1	140	0	0	5
74	28,00	73,20	1	165	0	0	3	0	1	2
75	14,20	1097,42	5	2112	1	1	20	1	0	12
76	75,00	331,20	2	874	1	1	806	0	0	5
77	95,00	103,80	1	536	1	1	343	0	1	3
78	14,40	1177,30	5	2122	1	1	806	1	0	12
79	139,00	354,45	1	896	1	1	1072	0	0	5
80	19,60	1368,86	5	998	1	1		0	0	6
81	47,50	481,00	2	1575	1	0	102	1	0	4
82	94,00	54,00	1	72	1	0	102	1	0	4
83	85,00	219,10	1	1110	1	1	192	0	0	5
84	35,50	512,80	2	3700	1	1	199	0	0	5
85	151,00	311,00	1	1463	1	1	291	0	0	5
86	90,00	22,50	1	74	1	0		0	1	1
87	86,00	458,50	1	509	1	0	50	1	0	4
88	86,00	337,50	1	506	1	0	46	1	0	4
89	86,00	149,00	1	506	1	0	50	1	0	4
90	86,00	252,80	1	506	1	0	46	1	0	4
91	102,00	234,60	1	54	1	0	82	0	0	6
92	13,00	444,40	2	450	1	0	469	0	0	8
93	41,00	115,00	1	95	1	0		0	1	1