Universidade Federal de Santa Catarina  
Programa de Pós-Graduação em Engenharia de Transportes e Gestão Territorial

Tópicos Especiais em Gestão Territorial

**Trabalho 2 – Trimestre 2017.2**

**Regressão linear**

Professor: Norberto Hochheim

Acadêmicos: Gustavo Andreas Hochheim

Ivan Merencio

Micael Desengrini

Ricardo Cesar dos Passos

Florianópolis, 09 de agosto de 2017.

Sumário

[**1.** **EQUAÇÃO** 3](#_Toc489880172)

[**1.1** **Equação de regressão e de estimação** 3](#_Toc489880173)

[**1.2** **Equação de regressão e de estimação** 3](#_Toc489880175)

[**2.** **ESTATÍSTICAS DO MODELO** 5](#_Toc489880176)

[**2.1** **Estatísticas da regressão** 5](#_Toc489880177)

[**2.2** **Significância dos bi’s** 5](#_Toc489880178)

[**2.3** **Análise de variância** 6](#_Toc489880179)

[**3.** **PODER DE PREDIÇÃO DO MODELO** 6](#_Toc489880180)

[**4.** **PRESSUPOSTOS BÁSICOS** 7](#_Toc489880181)

[**4.1** **Multicolinearidade** 7](#_Toc489880182)

[**4.2** **Normalidade dos resíduos** 8](#_Toc489880183)

[**4.3** **Homocedasticidade** 11](#_Toc489880184)

[**4.4** **Micronumerosidade** 11](#_Toc489880185)

# **EQUAÇÃO**

# **Equação de regressão e de estimação**

As variáveis independentes analisadas no modelo foram: Área do terreno, distância ao polo valorizador (no caso o mar) e número de pavimentos que podem ser construídos no terreno. As variáveis “Esquina” e “Testada” foram analisadas, mas não possuíram relação com o modelo. Abaixo estão listadas as equações de melhor ajuste para a variável dependente, “Valor Total”:

**Equação de regressão:**

****

**Equação de estimação:**

****



# **Equação de regressão e de estimação**

Abaixo podem ser vistos os coeficientes de cada variável separadamente:



Observa-se que os sinais dos coeficientes do modelo estão de acordo com o mercado, pois o valor do terreno aumenta com o aumento da área, possuindo o coeficiente sinal positivo; o valor do terreno sofre um decréscimo com o afastamento do mar, possuindo o seu coeficiente um sinal negativo, porém, o mesmo está em sua função inversa, estando assim de acordo com o esperado; o número de pavimentos valoriza o terreno, quanto mais se puder construir sobre o mesmo. Seu coeficiente apresenta-se com sinal negativo, mas o mesmo está em sua função inversa, mostrando que o modelo está de acordo com o esperado.

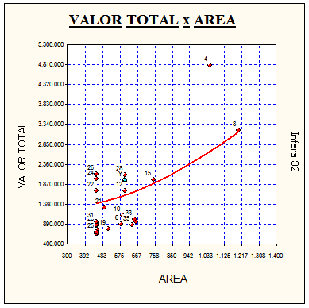
1. **retirar**

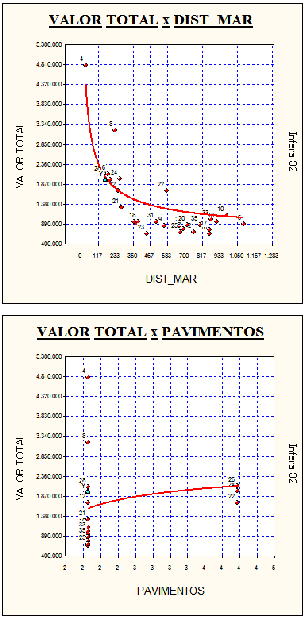
Abaixo, a tabela resumo das variáveis:

|  |  |  |
| --- | --- | --- |
| Análise das variáveis dependentes | | |
| Variável | Descrição | Crescimento esperado |
| Área do terreno | Quantitativa contínua | Crescimento positivo. Espera-se que aumento da área do terreno agregue valor ao mesmo. |
| Distância ao mar | Quantitativa contínua | Crescimento negativo. O valor deve diminuir com o afastamento da orla do mar. |
| Pavimentos construíveis | Quantitativa discreta | Crescimento positivo. O valor do terreno deve aumentar de acordo com uma maior possibilidade construtiva. |

1. retirar
   1. retirar

Abaixo podem ser visto também os gráficos que mostram como o valor de um imóvel tomado como exemplo (JUR\_01) se altera ao mudarem-se as variáveis independentes uma a uma. Pode-se ver que as variações de valor apresentam-se novamente coerentes:

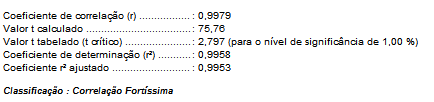




# **ESTATÍSTICAS DO MODELO**

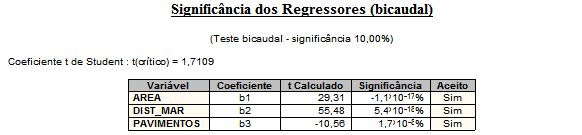
# **Estatísticas da regressão**

Por meio do *software* INFER32 calculou-se a correlação do modelo. Na figura abaixo, percebe-se que a correlação entre as variáveis independentes e a variável dependente é fortíssima, pois seu r é de 0,9979, indicando que o modelo prevê bem o comportamento da variável dependente. Além disso, o r² apresenta-se também muito alto, com 0,9958, ou seja, 99,58% da variação dos valores em relação à média é explicada pela equação de regressão, sendo que o restante, 0,42% são atribuídos à perturbações aleatórias e variáveis que não estão incluídas no modelo.



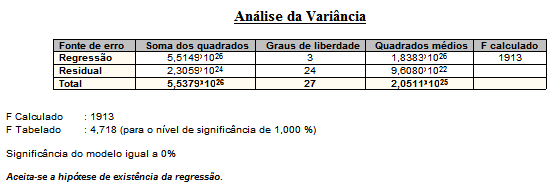
# **Significância dos bi’s**

Na figura abaixo pode ser analisada a significância de cada coeficiente do modelo:

****

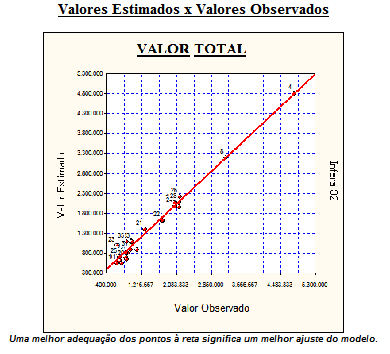
Todas variáveis foram aceitas na significância dos regressores (bicaudal), sendo que o maior valor encontrado foi 1,7 x 10-8%, o que permite, inclusive, enquadrar o modelo como tendo o grau de fundamentação III nesse quesito.

# **Análise de variância**

****

A hipótese de existência da regressão foi aceita pelo modelo (houve rejeição da hipótese nula de não existir regressão através do teste F de Snedecor. Além disso, como a significância do modelo é igual a 0%, o modelo pode ser enquadrado como de grau III nesse quesito, pois é inferior a 1%.

# **PODER DE PREDIÇÃO DO MODELO**

****

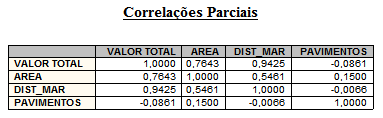
O modelo apresenta um bom poder de predição, com os dados amostrais apresentando valores estimados e observados muito próximos. Além disso, a distribuição dos valores se faz de forma bem distribuída e aleatória, denotando um bom poder de predição para o mesmo.

# **PRESSUPOSTOS BÁSICOS**

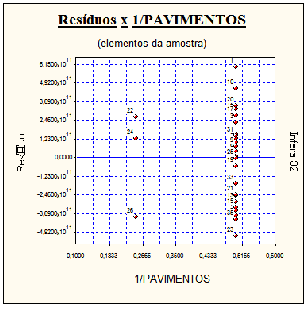
1. **retirar**

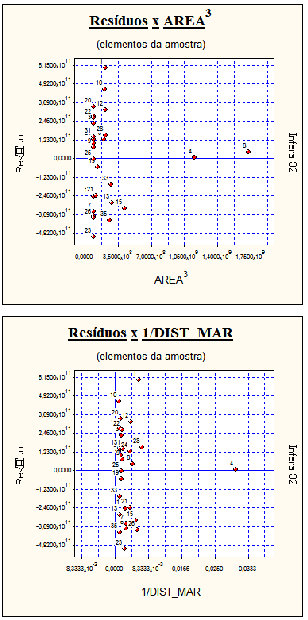
# **Multicolinearidade**

Deve-se observar se o modelo possui multicolinearidade, pois o mesmo não pode apresentar relações exatas, ou ainda muito fortes, entre quaisquer variáveis independentes. Abaixo pode ser acompanhada a matriz de correlações do modelo adotado:

****

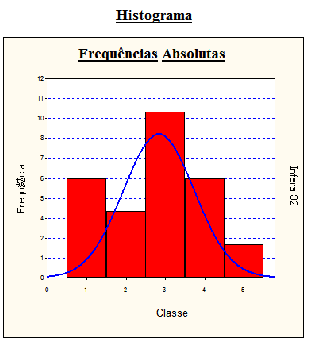
Observa-se na matriz acima que todos os valores estão abaixo de 0,8000, não indicando uma relação forte demais entre as variáveis independentes. ~~Apenas as variáveis~~ Além disso, os gráficos de resíduos apresentam-se de forma aleatória, concluindo que o modelo não apresenta multicolinearidade, como pode ser acompanhado abaixo:

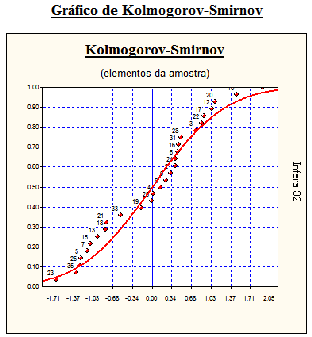




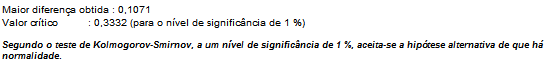
# **Normalidade dos resíduos**

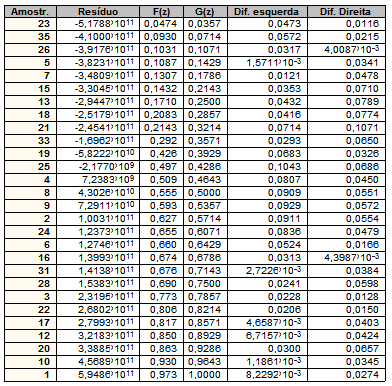
É importante verificar se os resíduos apresentam distribuição normal. Para tanto, pode-se realizar análises gráficas por diferentes métodos, ou ainda testes de normalidade para análise da hipótese H0 (resíduos apresentam normalidade), ou H1 (resíduos não apresentam normalidade). Neste trabalho a normalidade dos resíduos foi averiguada pelo histograma de frequências absolutas, gráfico de Kolmogorov-Smirnov e pelo teste de Kolmogorov-Smirnov, como pode ser acompanhado abaixo:

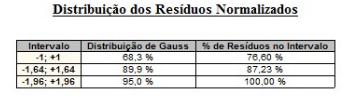
****

****

****

****

****





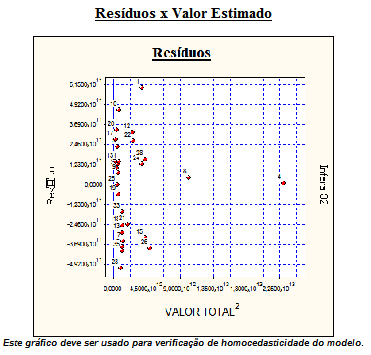
O histograma acima apresentou a distribuição de frequências absolutas que se assemelham com a curva de distribuição normal (com exceção apenas da primeira classe) e, além disso, o gráfico de Kolmogorov-Smirnov apresentou pontos que se aproximam da curva.

O teste de Kolmogorov-Smirnov aceitou a hipótese H0 de que há normalidade para a significância de 1%. Já a distribuição de resíduos normalizados ficou bastante próxima da esperada, com exceção do intervalo -1;+1.

Conclui-se, pelos dados acima e principalmente pelo resultado do teste de Komogorov-Smirnov, que os resíduos do modelo apresentam distribuição normal.

# **Homocedasticidade**

Para verificar a homocedasticidade, podemos realizar análise gráfica dos resíduos versus valores ajustados, que devem apresentar pontos sem padrão definido, dispostos aleatoriamente:

****

O gráfico de resíduos não apresenta mancha definida, indicando que o modelo é homocedástico.

# **Micronumerosidade**

Para evitar micronumerosidade, a NBR 14653-2 (2011) exige que o número mínimo de dados efetivamente utilizados no modelo (n) deve obedecer ao seguinte critério com relação ao número de variáveis independentes (k):

e para

onde *ni* é o número de dados de mesma característica, no caso de utilização de variáveis dicotômicas e variáveis qualitativas expressas por códigos alocados ou códigos ajustados.

Após a escolha do modelo e retirada dos *outliers*, obtivemos:

Nº de elementos da amostra: 28

Nº de variáveis independentes: 3

substituindo os valores na inequação acima:

Como o modelo não apresenta variáveis qualitativas expressas por códigos alocados ou ainda variáveis dicotômicas, a segunda verificação não precisa ser realizada.

Com os valores mostrados acima respeitando as exigências da norma, podemos concluir que não existe micronumerosidade no modelo.