Metodologia de Superfície de Resposta e Otimização

LOPP
UFRGS

O que é Modelagem?

- As equações de regressão podem ser usadas para analisar a relação entre uma variável dependente (variável de resposta) e várias variáveis independentes (fatores controláveis)
- A ANOVA permite identificar os fatores controláveis significativos sobre a VR e os gráficos permitem identificar o ajuste ótimo dos fatores controláveis
- A regressão ainda permite prever o valor da VR para uma determinada combinação dos fatores controláveis através da equação de regressão

Quando utilizar?

- A ANOVA permite trabalhar com fatores controláveis qualitativos (não-métricas) ou quantitativos (métricas), desde que a VR seja quantitativa
- A regressão pode ser usada quando a VR e os fatores controláveis são quantitativos
- Sob certas circunstâncias é possível incluir fatores controláveis não-métricos (transformando os FC do tipo ordinais ou nominais em variável com codificação dicotômica)
- Quando a variável de resposta for não-métrica utiliza-se a regressão logística ou loglinear

Quando utilizar?

- **A definição do ajuste ótimo da ANOVA é realizado graficamente e fica restrito aos ensaios realizados**
- **A definição do ajuste ótimo da regressão é realizado por Pesquisa operacional e permite interpolação dos fatores controláveis pela equação de regressão**

Introdução à MSR

- A Metodologia de Superfície de Resposta (MSR) envolve uma série de técnicas orientadas à análise de experimentos planejados de modo a gerar informações suficientes para a modelagem das respostas de interesse através de superfícies n-dimensionais.
- Após a construção de modelos para a resposta, o interesse recai na busca do ajuste ótimo, ou seja, na busca de regiões que conduzam a um valor mínimo, máximo ou nominal, conforme a característica da resposta em questão.

LOPP
UFRGS

Aplicações da MSR

- **A MSR tem ampla aplicação dentro da engenharia, contribuindo para a otimização de produtos ou processos, principalmente quando os fatores controláveis são a níveis contínuos .**

- **Apesar do potencial da MSR no que se refere a otimização de produtos e processos, essa metodologia é pouco empregada no Brasil, pois exige o domínio dos conceitos básicos de projeto de experimentos, regressão múltipla e otimização, e poucas escolas de engenharia mantém cursos que contemplem todas essas áreas.**

Etapas no uso da MSR

- A proposta da MSR é responder questões gerais referente ao comportamento da resposta dentro do intervalo de interesse e, em particular, mapear regiões de alto desempenho.

- Os estudos envolvem três etapas principais:

- Planejar o experimento, distribuindo adequadamente os pontos experimentais

- Estimar os coeficientes da equação da superfície de resposta

- Explorar a superfície de resposta encontrando o ajuste dos fatores que maximiza a resposta

LOPP
UFRGS

Estratégia de análise

- A estratégia de análise supõe que a resposta Y possa ser representada por uma função polinomial dos fatores controláveis $X_1, X_2, ..., X_k$.
- Entre os modelos possíveis, estão o modelo linear,

$$Y = b_0 + b_1X_1 + b_2X_2 + ... + b_kX_k$$

- o modelo quadrático,

$$Y = b_0 + b_1X_1 + b_2X_2 + ... + b_{11}X_1^2 + b_2X_2^2 + ... + b_{12}X_1X_2 + ...$$

- e também modelos não lineares.

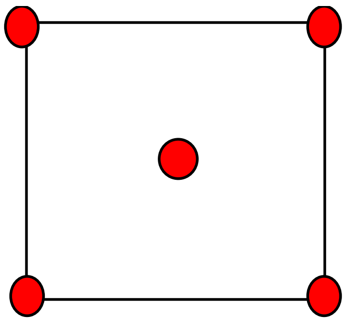
LOPP
UFRGS

Projetos de superfície de resposta

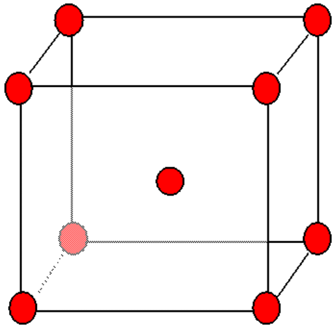
- Os coeficientes dos modelos podem ser estimados mais eficientemente se for usado um projeto experimental adequado para a coleta de dados.
- Projetos para o ajuste de superfícies de resposta são chamados de projetos de superfície de resposta.
- Por exemplo, para ajustar modelos lineares, toda a classe de experimentos 2^k são particularmente eficientes.
- Eles permitem fracionamento, blocagem e a suposição de linearidade pode ser testada acrescentando-se alguns pontos centrais.

Exemplos de projetos para modelos lineares

- **Um projeto 2^2 com um ponto central**



- **Um projeto 2^3 com um ponto central**

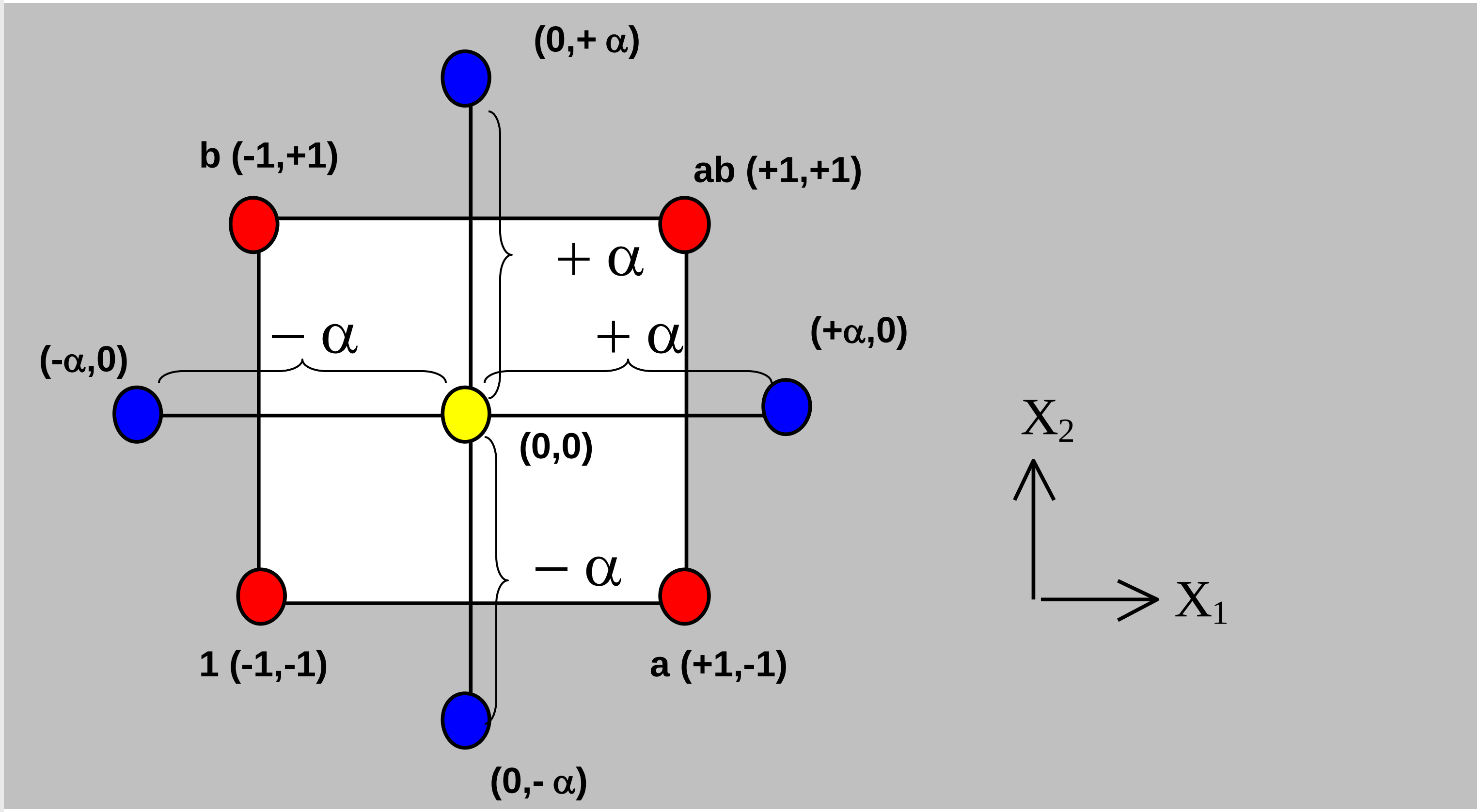


LOPP
UFRGS

Projetos para modelos quadráticos

- **Para o ajuste de modelos quadráticos, o Projeto Composto de Segunda Ordem (PCSO) é recomendado.**
- **Esse projeto, que será visto a seguir tem inúmeras vantagens.**
- **Ele tem como base um projeto 2^k, exige um número pequeno de ensaios, pode contemplar blocagem, rotacionalidade e ortogonalidade.**
- **O que aparece a seguir é um exemplo de um PCSO para um experimento de dois fatores controláveis:**

Exemplo de um PCSO



LOPP
UFRGS

Matriz experimental

- **A matriz experimental para esse experimento seria a seguinte:**

Rodada	X1	X2		Y
1 (1)	-1	-1	Fatorial	
2 (a)	+1	-1	Fatorial	
3 (b)	-1	+1	Fatorial	
4 (ab)	+1	+1	Fatorial	
5	α	0	Estrela	
6	$-\alpha$	0	Estrela	
7	0	α	Estrela	
8	0	$-\alpha$	Estrela	
9	0	0	Central	

LOPP
UFRGS

Construção dos PCSO

- Como pode ser visto, o PCSO é a soma de um experimento 2^k, mais uma estrela, mais pontos centrais.
- Por isso o nome projeto composto.
- Os pontos da parte fatorial (2^k) permitem a estimativa de termos lineares e interações.
- Os pontos da estrela, permitem a estimativa de efeitos quadráticos puros.

LOPP
UFRGS

Construção dos PCSO

- De forma geral, os PCSO consistem de três partes:

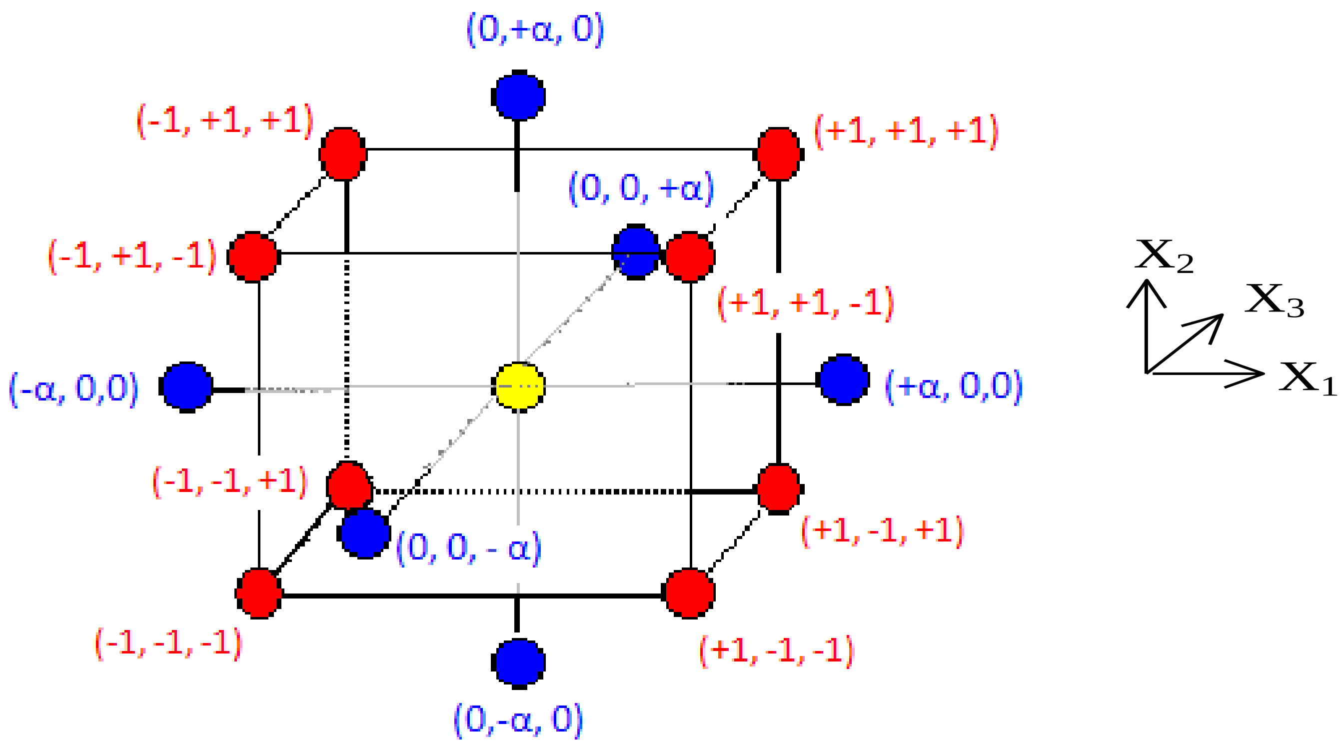
a) A parte fatorial, ou seja 2^k vértices de um cubo k dimensional (ou uma fração desses vértices) com coordenadas $\pm1, \pm1, ..., \pm1$.

b) A parte em estrela, 2xK vértices com coordenadas 0, ..., $\pm\alpha$, ...,0.

c) n_o pontos centrais, com coordenadas 0,0,...

Exemplo de um PCSO - 3 fatores

- **A figura a seguir apresenta um PCSO para um experimento de três fatores:**



LOPP
UFRGS

PCSO para 3 fatores

Rodada	X1 (A)	X2 (B)	X3 (C)		Y
1 (1)	-1	-1	-1	Fatorial	
2 (a)	1	-1	-1	Fatorial	
3 (b)	-1	1	-1	Fatorial	
4 (ab)	1	1	-1	Fatorial	
5 (c)	-1	-1	1	Fatorial	
6 (ac)	1	-1	1	Fatorial	
7 (bc)	-1	1	1	Fatorial	
8 (abc)	1	1	1	Fatorial	
9	α	0	0	Estrela	
10	$-\alpha$	0	0	Estrela	
11	0	α	0	Estrela	
12	0	$-\alpha$	0	Estrela	
13	0	0	α	Estrela	
14	0	0	$-\alpha$	Estrela	
15	0	0	0	Central	

PCSO para 4 fatores

	X1 (A)	X2 (B)	X3 (C)	X4 (D)
1	-1	-1	-1	-1
a	1	-1	-1	-1
b	-1	1	-1	-1
ab	1	1	-1	-1
c	-1	-1	1	-1
ac	1	-1	1	-1
bc	-1	1	1	-1
abc	1	1	1	-1
d	-1	-1	-1	1
ad	1	-1	-1	1
bd	-1	1	-1	1
abd	1	1	-1	1
cd	-1	-1	1	1
acd	1	-1	1	1
bcd	-1	1	1	1
abcd	1	1	1	1
estrela X1	alfa	0	0	0
estrela X1	-alfa	0	0	0
estrela X2	0	alfa	0	0
estrela X2	0	-alfa	0	0
estrela X3	0	0	alfa	0
estrela X3	0	0	-alfa	0
estrela X4	0	0	0	alfa
estrela X4	0	0	0	-alfa
ponto central	0	0	0	0
ponto central	0	0	0	0

LOPP
UFRGS

Características dos PCSO

- Caso necessário, o projeto pode contemplar repetições do ponto central, aumentando os graus de liberdade do termo de erro, ou seja, permitindo uma avaliação mais precisa da variância experimental.
- O valor de alfa pode ser definido de modo que o projeto tenha algumas propriedades interessantes.
- Por exemplo, alfa pode ser calculado para atribuir rotacionalidade ou ortogonalidade ao projeto.

Rotacionalidade

- **Um projeto rotacional assegura a mesma precisão nas estimativas de Y para todos os pontos do espaço amostral.**
- **Para atribuir rotacionalidade ao projeto, o valor de alfa deve ser definido usando:**

$$\alpha = F^{\frac{1}{4}}$$

- **onde F se refere ao número de pontos da parte fatorial**

LOPP
UFRGS

Ortogonalidade

- Outra possibilidade é atribuir ao projeto a condição de ortogonalidade.

- Nesse caso, a estimativa dos coeficientes de termos lineares e quadráticos resultam independente, ou seja, essas estimativas não se alteram quando algum termo é eliminado do modelo.

Ortogonalidade

- **Para atribuir ortogonalidade ao projeto, o valor de alfa deve ser definido usando:**

$$\alpha = \left\{ \frac{\left[(F+T)^{1/2} - F^{1/2} \right]^2}{4 \times n^2} \times F \right\}^{\frac{1}{4}}$$

- **onde F se refere ao número de pontos da parte fatorial**
- **T é o número de pontos adicionais (estrela mais pontos centrais), multiplicado pelo número de repetições n.**

LOPP
UFRGS

Blocos ortogonais

- **Por fim, os PCSO são particularmente eficientes quando existe a necessidade de blocagem.**
- **Nesse caso, o projeto é normalmente dividido em dois:**

- **um bloco contendo a parte fatorial e**
- **o outro bloco contendo a parte em estrela.**

- **Os pontos centrais são utilizados para assegurar o mesmo número de ensaios em cada bloco.**

LOPP
UFRGS

Blocos ortogonais

- Para assegurar que os blocos serão ortogonais entre si, o que irá permitir extrair o efeito entre blocos, caso ele exista, basta ter o mesmo número de ensaios em cada bloco e definir o valor de alfa usando:

$$\alpha = \sqrt{\frac{F}{2}}$$

LOPP
UFRGS

Blocos ortogonais

- Quando o experimento está dividido em blocos ortogonais, realiza-se a parte fatorial mais os pontos centrais
- Se os pontos centrais detectarem a falta de ajuste do modelo linear, roda-se a estrela para investigar os efeitos quadráticos
- Ou seja, a parte da estrela só é realizada caso necessário

LOPP
UFRGS

Modelagem das VR

- Inicialmente, a partir dos resultados do experimento planejado, chega-se a modelos para as VR:
- Esses modelos podem contemplar média e variabilidade:

$Y_1 = f1\ (X_1, X_2, ..., X_k)$ $\qquad$ $\sigma_{Y1} = g1\ (X_1, X_2, ..., X_k)$

$Y_2 = f2\ (X_1, X_2, ..., X_k)$ $\qquad$ $\sigma_{Y2} = g2\ (X_1, X_2, ..., X_k)$

: :

$Y_p = fp\ (X_1, X_2, ..., X_k)$ $\qquad$ $\sigma_{Yp} = gp\ (X_1, X_2, ..., X_k)$

- Para todas as variáveis de resposta Y_j, já se conhece de antemão o seu valor alvo, os limites de especificação e a importância relativa (IR_j).

LOPP
UFRGS

Regressão Linear Simples

- A regressão linear simples se aplica àquelas situações onde há duas variáveis (digamos, X e Y) que podem possuir uma relação de causa e efeito.
- A variável X é chamada de variável independente ou fator controlável (causa) e a variável Y é a variável dependente ou variável de resposta (efeito, que depende de X).
- É dito relação linear simples, pois supõe-se tendência linear entre as variáveis e simples por ser uma única variável independente (X).

LOPP
UFRGS

Regressão Linear Simples

Seja que existam dados coletados (pares de valores X e Y) associando uma variável de resposta Y (variável dependente) com uma variável regressora X (variável independente).

E suponha que a relação entre Y e X seja aproximadamente linear.

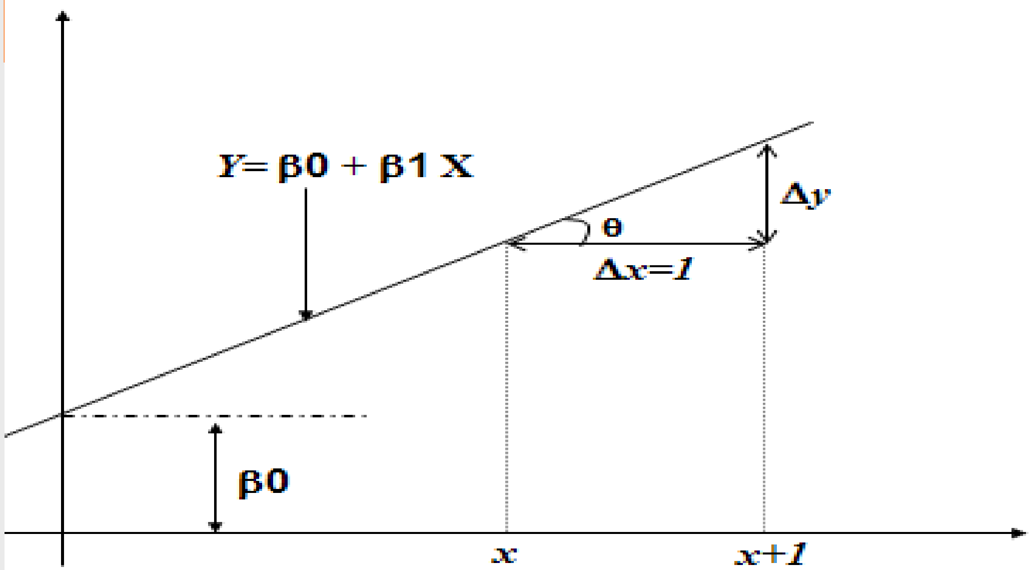
Então o valor esperado de Y para cada valor de X virá dado por

$$E\,(Y/X) = \beta_0 + \beta_1 X$$

onde os parâmetros da relação linear, β_0 e β_1, são desconhecidos e estimados através de estimativas amostrais

$$Y = b_o + b_1 X$$

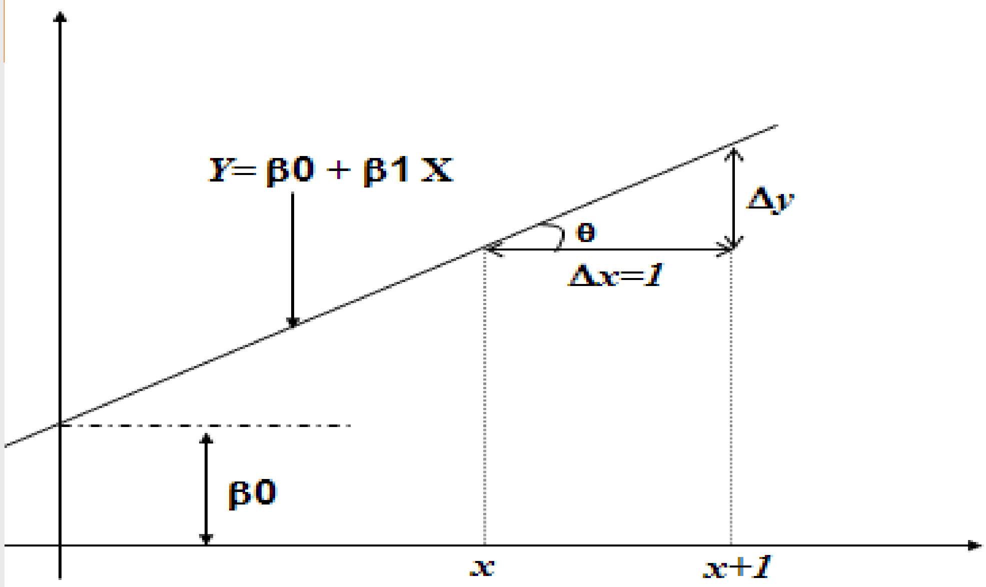
Equação de regressão



- Cada observação Y pode ser descrita pelo modelo:
- $Y = \beta_0 + \beta_1 X + \varepsilon$
- ε é o erro aleatório, com média 0 e variância σ^2.
- β_0 é a interseção (valor de Y para X = 0)
- β_1 é a inclinação da reta, que descreve a variação de Y para cada unidade de X.

LOPP
UFRGS

Equação de regressão



O coeficiente angular β_1 é a tangente do angulo e quanto maior o valor mais inclinada é a reta.

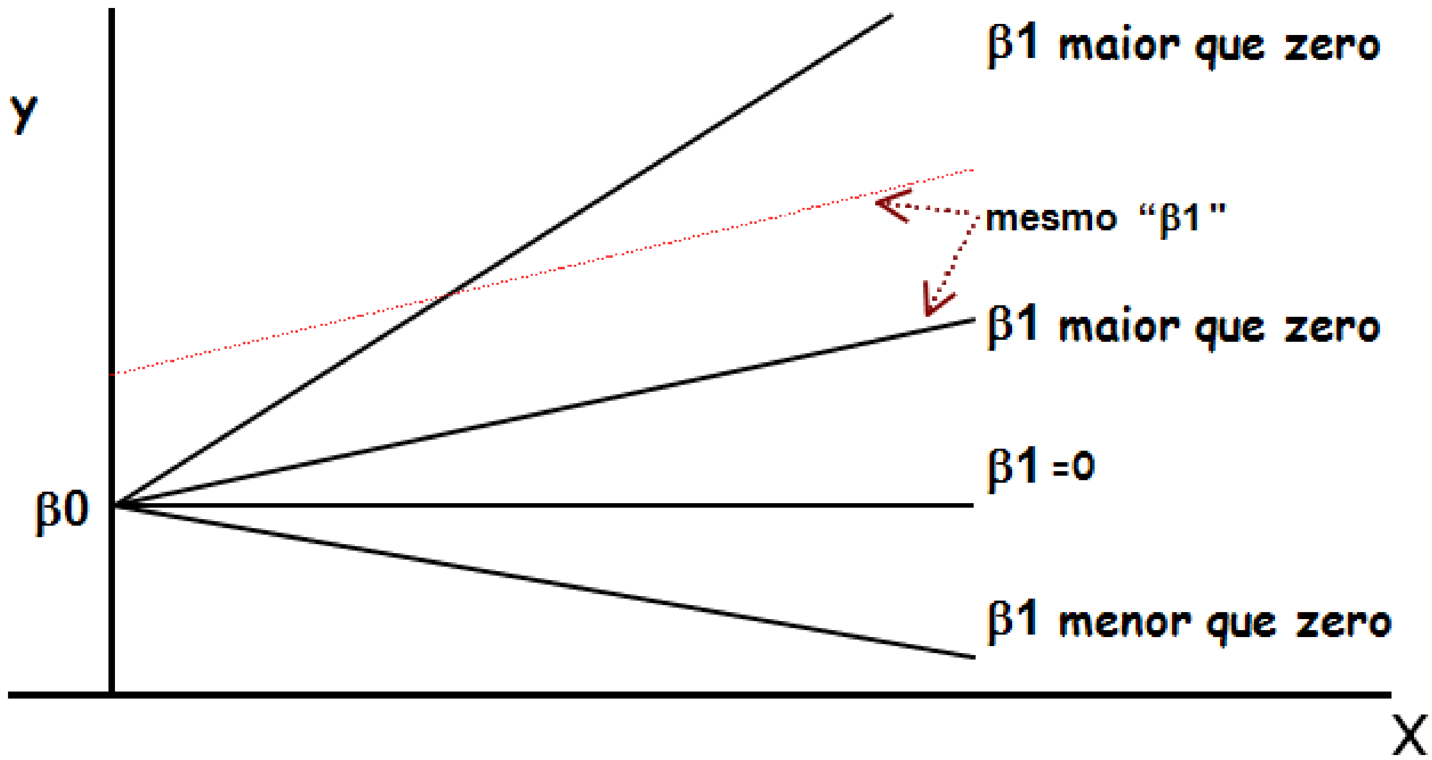
Se "β_1" é positivo, a reta é crescente

Se "β_1" é negativo, a reta é decrescente

Se "β_1" é zero, Y não depende de X e a reta é paralela ao eixo X na altura do valor β_0

Coeficiente angular

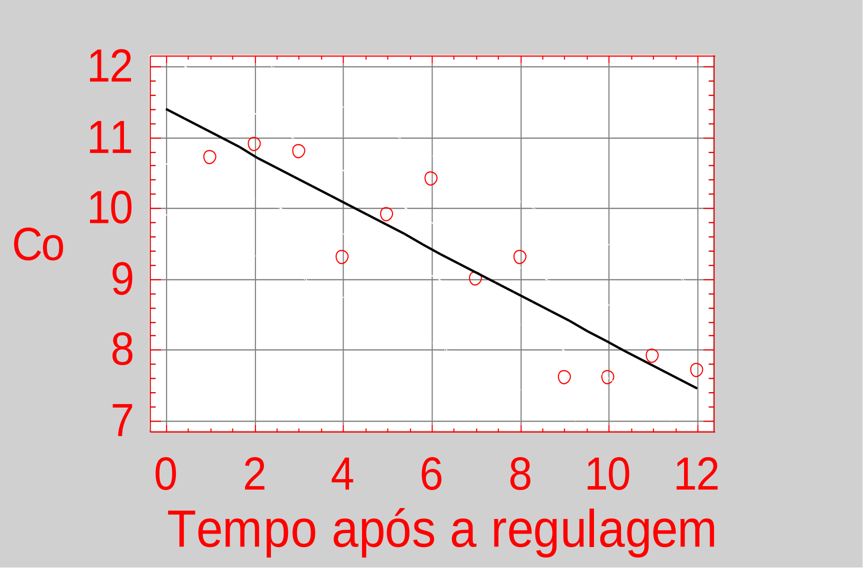
Alguns exemplos de coeficientes β_1



LOPP
UFRGS

- **Se há n pares de dados (Y_1, X_1), ..., (Yn, Yn) é possível estimar os parâmetros β_0 e β_1 através de estimativas amostrais (b_o e b_1) usando o método dos Mínimos Quadrados Ordinários (MQO), o qual busca minimizar os resíduos:** $R_i = Y_{obs} - Y_{est}$

$$L = \sum_{i=1}^{n} (Y_{obs} - Y_{est})^2 = \sum_{i=1}^{n} (Y_{obs} - (b_0 + b_1 X_i))^2$$

- 

LOPP
UFRGS

Regressão Linear Simples

O uso do método MQO conduz as seguintes estimativas:

$$b1 = S_{XY} / S_{XX}$$

$$bo = \overline{Y} - b_1 \overline{X}$$

Onde:

$$S_{XX} = \sum X_i^2 - \left(\sum X_i\right)^2 / n$$

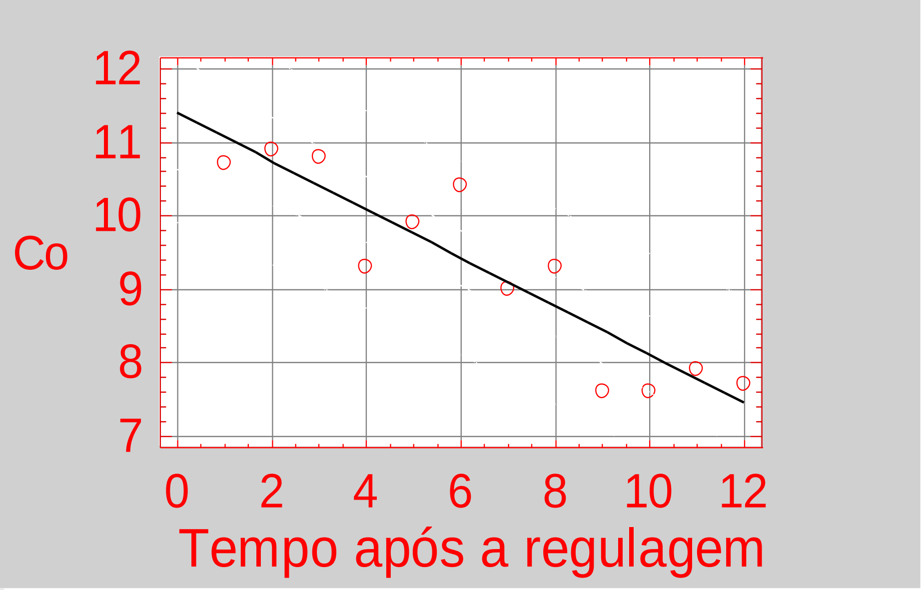
$$S_{YY} = \sum Y_i^2 - \left(\sum Y_i\right)^2 / n$$

$$S_{XY} = \sum X_i Y_i - \left(\sum X_i\right)\left(\sum Y_i\right) / n$$

LOPP
UFRGS

Exemplo

- **Após uma regulagem eletrônica um veículo apresenta um rendimento ideal no que tange a rendimento de combustível.**
- **Com o passar do tempo esse rendimento vai se degradando. Os dados da tabela representam o rendimento (Y) medido a cada mês (X) após a regulagem do veículo.**



Meses(X)	Rendimento(Y)
1	10.7
2	10.9
3	10.8
4	9.3
5	9.5
6	10.4
7	9
8	9.3
9	7.6
10	7.6
11	7.9
12	7.7

LOPP
UFRGS

Exemplo

Meses(X)	Rendimento(Y)	X^2	Y^2	X*Y
1	10,7	1	114,49	10,7
2	10,9	4	118,81	21,8
3	10,8	9	116,64	32,4
4	9,3	16	86,49	37,2
5	9,5	25	90,25	47,5
6	10,4	36	108,16	62,4
7	9	49	81	63
8	9,3	64	86,49	74,4
9	7,6	81	57,76	68,4
10	7,6	100	57,76	76
11	7,9	121	62,41	86,9
12	7,7	144	59,29	92,4
78	**110,7**	**650**	**1039,55**	**673,1**
6,5	**9,225**			

- $\Sigma X_i = 78{,}00;$ $\Sigma X_i^2 = 650{,}00\ ;$ $\bar{X} = 6{,}50$
- $\Sigma Y_i = 110{,}70;$ $\Sigma Y_i^2 = 1039{,}55\ ;$ $\bar{Y} = 9{,}225$

LOPP
UFRGS

Exemplo

Desvio-padrão de X:

$$S_{XX} = \sum X_i^2 - \left(\sum X_i\right)^2 \Big/ n = 650 - (78)^2 / 12 = 143{,}00$$

Desvio-padrão de Y:

$$S_{YY} = \sum Y_i^2 - \left(\sum Y_i\right)^2 \Big/ n = 1039{,}55 - (110{,}70)^2 / 12 = 18{,}34$$

Covariância de X,Y:

$$S_{XY} = \sum X_i Y_i - \left(\sum X_i\right)\left(\sum Y_i\right) \Big/ n = 673{,}1 - (78 \times 110{,}70) / 12 = -46{,}45$$

Coeficientes

$$b1 = S_{XY} / S_{XX} = -46{,}25 / 143 = -0{,}325$$

$$bo = \overline{Y} - b_1 \overline{X} = 9{,}225 - (-0{,}325) \times 6{,}50 = 11{,}34$$

Equação de regressão

$$Y = b_0 + b_1 X_i$$

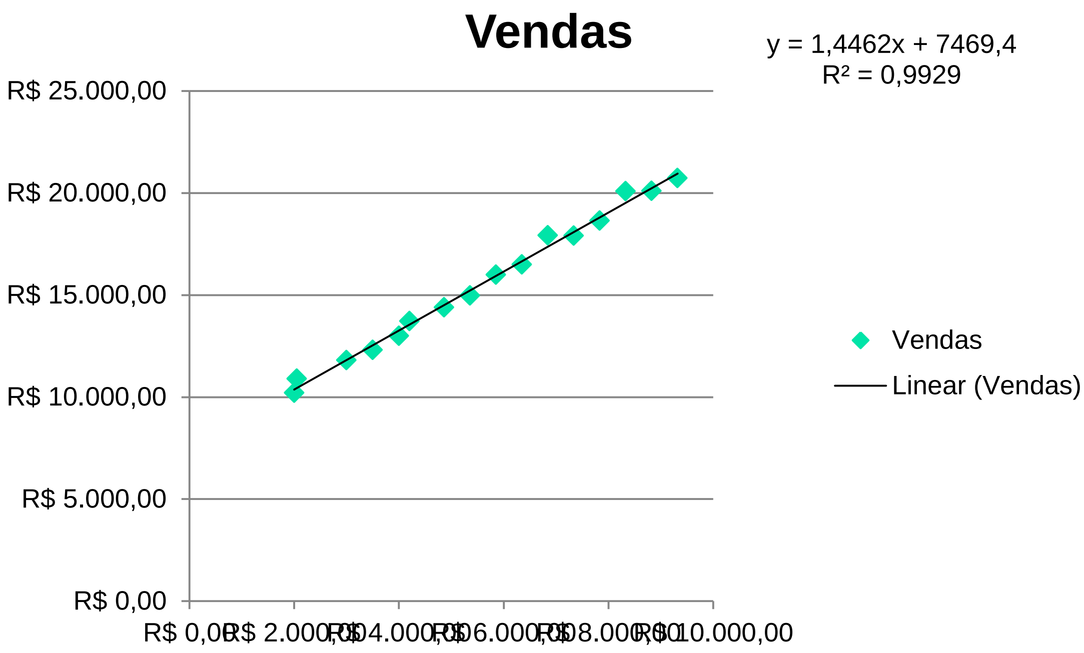
$$Y = 11{,}34 + (-0{,}325) \times X$$

LOPP
UFRGS

Exercício

- **Um gerente deseja investigar o efeito do investimento em marketing sobre o volume de vendas.**

Invest	Vendas
R$ 2.000,00	R$ 10.200,00
R$ 2.050,00	R$ 10.900,00
R$ 3.000,00	R$ 11.800,00
R$ 3.500,00	R$ 12.300,00
R$ 4.000,00	R$ 12.990,00
R$ 4.200,00	R$ 13.720,00
R$ 4.860,00	R$ 14.400,00
R$ 5.355,71	R$ 14.980,00
R$ 5.851,43	R$ 15.990,00
R$ 6.347,14	R$ 16.500,00
R$ 6.842,86	R$ 17.920,00
R$ 7.338,57	R$ 17.900,00
R$ 7.834,29	R$ 18.650,00
R$ 8.330,00	R$ 20.089,00
R$ 8.825,71	R$ 20.100,00
R$ 9.321,43	R$ 20.730,00



LOPP
UFRGS

Exercício

- **Um gerente deseja investigar o efeito do investimento em marketing sobre o volume de vendas.**

RESUMO DOS RESULTADOS

Estatística de regressão	
R múltiplo	0,996464396
R-Quadrado	0,992941292
R-quadrado ajustado	0,992437099
Erro padrão	298,2228194
Observações	16

ANOVA

	gl	SQ	MQ	F	F de significação
Regressão	1	175149195	1,75E+08	1969,3658	1,83463E-16
Resíduo	14	1245115,9	88936,85		
Total	15	176394310,9			

	Coeficientes	Erro padrão	Stat t	valor-P	95% inferiores	95% superiores	Inferior 95,0%	Superior 95,0%
Interseção	7469,369498	197,2415564	37,86915	1,661E-15	7046,328434	7892,411	7046,328	7892,411
Invest	1,446165736	0,032587787	44,37754	1,835E-16	1,376271885	1,51606	1,376272	1,51606

LOPP
UFRGS

Exercício- Interpretação do Resultado

A equação apresenta um valor de b_1= 1,44
Ou seja, para cada R$1,00 de investimento em marketing, o volume de vendas aumenta em R$ 1,44

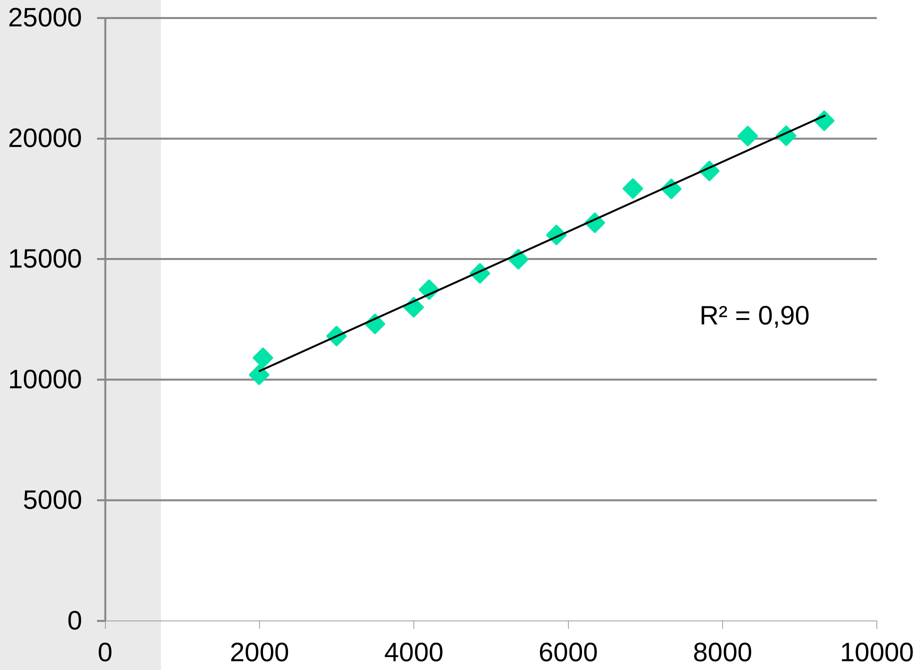
A equação só é válida para o intervalo estudado (ou seja entre R$2000,00 e R$9310,00)

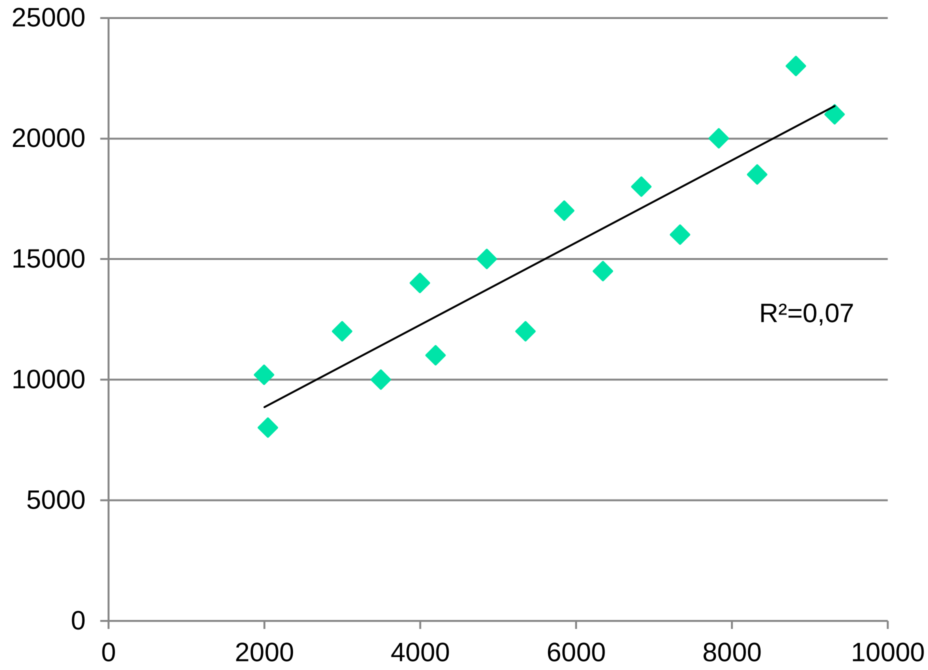
Não é recomendável aplicar a equação para previsão de valores fora do intervalo investigado

LOPP
UFRGS

Coeficiente de Determinação R^2

- **Coeficiente de determinação R^2 é uma medida de quão bem a equação de regressão se ajusta as dados amostrais**

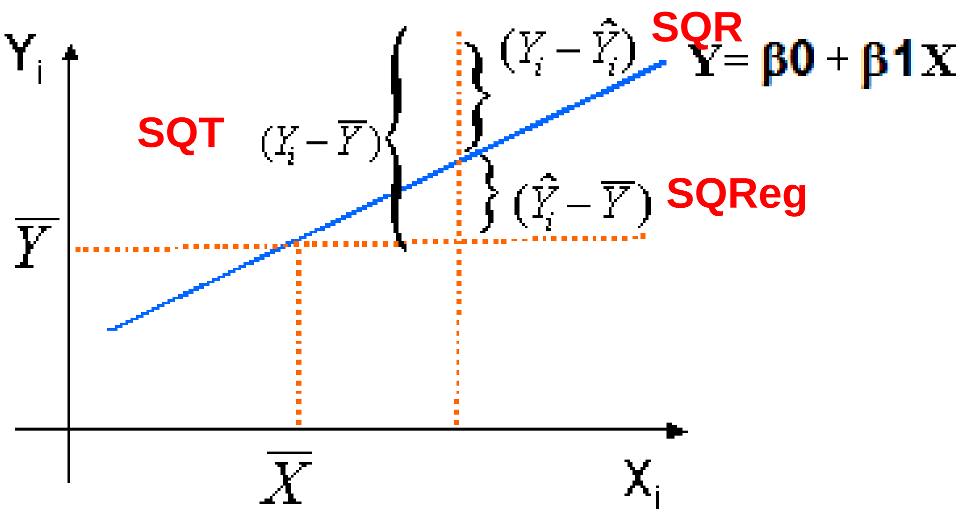




- **O Coeficiente de determinação R^2 equivale a proporção da variância dos valores de Y que pode ser atribuída à regressão com a variável X, logo** $0 \leq R^2 \leq 1$

LOPP
UFRGS

Coeficiente de Determinação R^2



- O R^2 indica o percentual da variabilidade de Y que é explicado pela modelo de regressão em função de X.
- Se $R^2 = 1$, todas as observações estarão sobre a reta definida pelo modelo (caso não houvesse variabilidade residual)
- Se $R^2 = 0$, não há nenhuma relação entre a variável dependente e as variáveis independentes

LOPP
UFRGS

Coeficiente de Determinação R^2

- O coeficiente de determinação R^2 é calculado segundo:

$$R^2 = \frac{S_{xy}^2}{S_{xx} \times S_{yy}} = \frac{-46{,}45^2}{143{,}00 \times 18{,}34} = 0{,}82$$

$$R^2 = \frac{SQ\,\mathrm{Re}\,g}{SQT} = \frac{SQT - SQR}{SQT} = 1 - \frac{SQR}{SQT}$$

- Para o exemplo analisado resultou $R^2 = 0{,}82$

• 82% da variabilidade nos resultados de rendimento de combustível pode ser devida ao tempo decorrido após a regulagem

• 18% da variabilidade total é devido a outros fatores que não foram investigados ou fatores de ruído

LOPP
UFRGS

Adequação do modelo de regressão

- A adequação do ajuste e as suposições do modelo podem ser verificadas através da análise dos resíduos.
- Os resíduos são a diferença entre os valor observados e o valores previstos pela equação

$$R_i = Y_{obs} - Y_{est}$$

$$R_i = Y_{obs} - (b_0 + b_1 X_i)$$

- Verifica-se as seguintes hipóteses
  - A distribuição do erro possui média zero
  - A variância do erro é constante (homogeneidade da variância)
  - A distribuição do erro é normal
  - Os valores do erro são independentes dos y observados

LOPP
UFRGS

Análise de resíduos

- **A Figura (a) representa uma situação onde o ajuste é adequado pois os resíduos são aleatórios, enquanto que a Figura (b) representa uma situação onde o modelo linear não se ajusta bem aos dados pois os resíduos apresentam um padrão não-aleatório.**

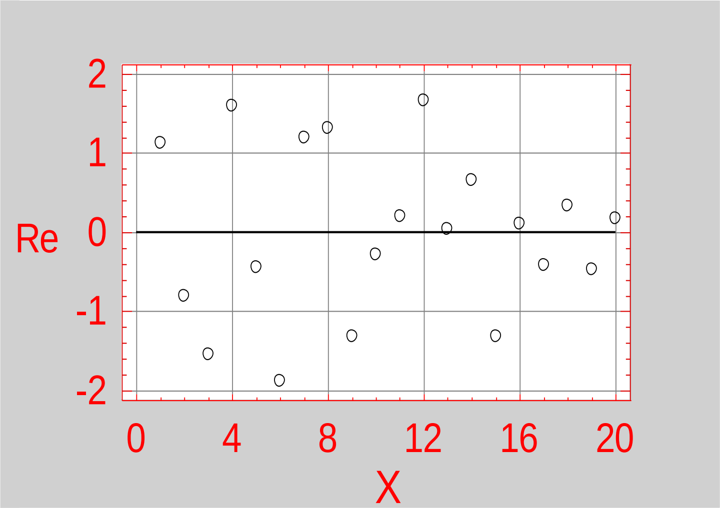


Fig. (a)

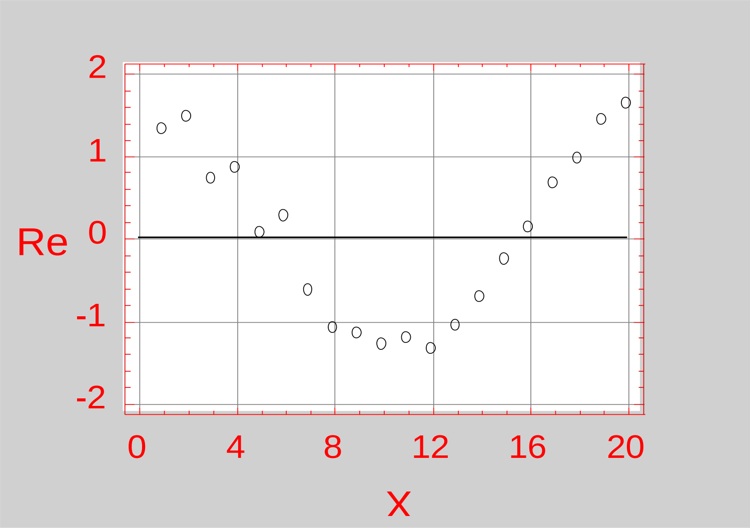


Fig. (b)

LOPP
UFRGS

Análise de resíduos

• Caso os resíduos não estejam distribuídos aleatoriamente é indício de falta de ajuste do modelo linear, ou seja, faltou algum termo quadrático, um logarítmico ou outros (de uma ou mais variáveis independentes)

• Quando isso acontecer deve-se acrescentar novos termos na equação.

• Se o modelo linear não fornece um bom ajuste, as vezes o problema pode ser contornado trabalhando-se com valores transformados de X ou Y, por exemplo:

$$Y = b_0 + b_1\sqrt{X}$$

$$Y = b_0 + b_1 X* \qquad \text{onde } X* = \sqrt{X}$$

LOPP
UFRGS

Análise de resíduos

- A suposição de homogeneidade da variância σ^2 ao longo do intervalo de X pode ser verificada analisando o gráfico de Resíduos $\times$ X.
- Na Figura (a) verifica-se a suposição de homogeneidade da variância e na Figura (b) essa suposição é violada.

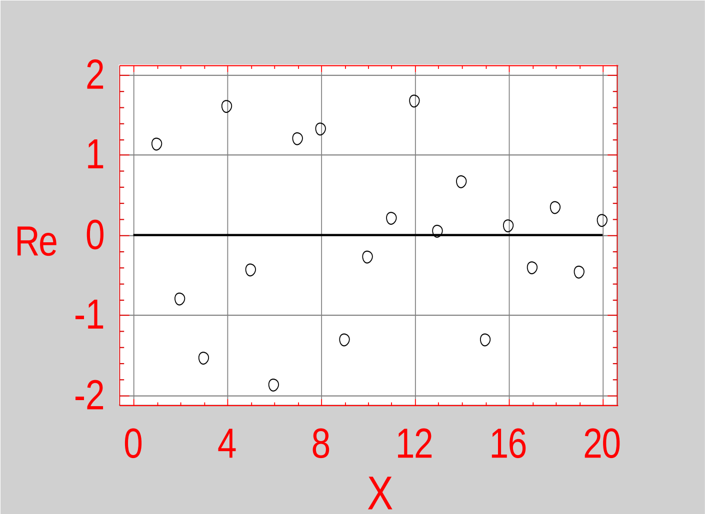


Fig. (a)

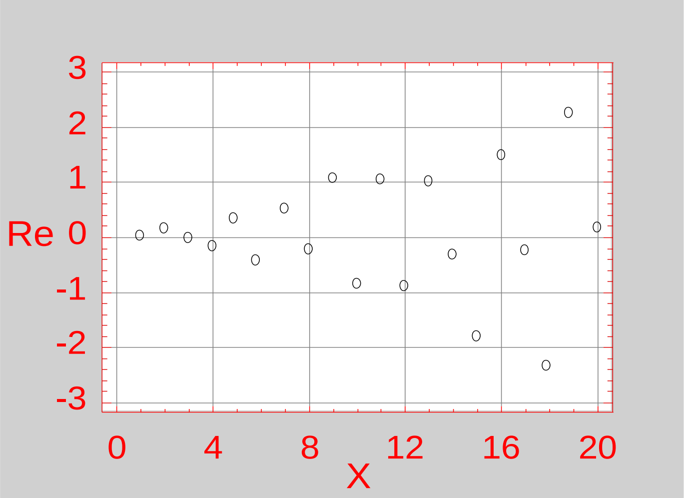


Fig. (b)

LOPP
UFRGS

Análise de resíduos

Se a suposição de homogeneidade da variância é rejeitada, pode-se usar

o método da regressão linear ponderada, onde se busca os valores de β_0 e β_1 que minimizam

$$L = \sum (Y_{obs} - Y_{est})^2 = \sum K_i (Y_i - (b_0 + b_1 X_i))^2$$

Nesse caso, os pesos k_i são inversamente proporcionais à variância.

Análise dos resíduos padronizados

Se o resíduo de algum ponto tem comportamento diferente dos resíduos dos outros pontos este ponto pode não pertencer a esse grupo de dados.

Deve-se observar os valores dos resíduos padronizados para verificar a existência de valores atípicos.

Se houver registro de alguma causa especial que tenha afetado essa coleta, essa observação pode ser eliminada do conjunto e a análise de regressão deve ser rodada novamente, possivelmente fornecendo um modelo mais preciso.

Análise dos resíduos padronizados

- **Os resíduos padronizados são calculados dividindo-se o resíduo pelo desvio-padrão**

$$Z_i = \frac{Y_{obs} - Y_{est}}{S} = \frac{Y_{obs} - (b_0 + b_1 X_i)}{S}$$

$$S = \sqrt{SQR / n - 2}$$

$$SQR = S_{YY} - b_1 S_{XY}$$

- **Considera-se um valor atípico quando o resíduo padronizado (*Z_i*) for maior do que:**

Z_i > 3,00 (quando o intervalo de confiança adotado for 99,73%)

Z_i > 1,96 (quando o intervalo de confiança adotado for 95,00 %)

LOPP
UFRGS

Regressão múltipla

Embora haja muitos problemas em que uma variável pode ser predita com bastante precisão em termos de outra, é claro que as predições devem melhorar se for levado em conta informações adicionais importantes. Por exemplo:

- Pode-se fazer melhores predições sobre o desempenho de funcionários recém contratados se for levado em consideração não somente sua formação, mas também seu tempo de experiência e sua personalidade;
- Pode-se fazer melhor predição do sucesso de um novo produto se for considerado não somente sua qualidade, mas o potencial de procura e a concorrência.
- A qualidade de um processo químico pode depender da temperatura, pressão e taxa de agitação.

LOPP
UFRGS

O modelo geral da regressão múltipla

Uma equação de regressão linear múltipla expressa uma relação entre uma variável dependente (Y) e duas ou mais variáveis independentes (X's)

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k + \varepsilon$$

Onde:

Y = valor da variável dependente

β_0 = coeficiente de intersecção

k = número de variáveis independentes

$X_1, X_2, \ldots, X_k$, = variáveis independentes

$\beta_1, \beta_2, \ldots, \beta_k$ = coeficientes das variáveis independentes

ε = termo de erro

LOPP
UFRGS

O modelo geral da regressão múltipla

- **O problema então é estimar o valor dos coeficientes β_i a partir de um conjunto de dados do tipo:**

Y	X_1	X_2	...	X_k
y_1	x_{11}	x_{12}	...	x_{1k}
y_2	x_{21}	x_{22}	...	x_{2k}
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
y_n	x_{n1}	x_{n2}	...	x_{nk}

- **Ou seja, estimar b_o, b_1, b_2 ...**

LOPP
UFRGS

Exemplo

Um distribuidor de cerveja está analisando seu sistema de distribuição. Especificamente ele está interessado em prever o tempo requerido para atender um ponto de venda.

O engenheiro industrial acredita que os dois fatores mais importantes são o número de caixas de cerveja fornecidas e a distância do depósito ao posto de venda.

Coletar dados pareados

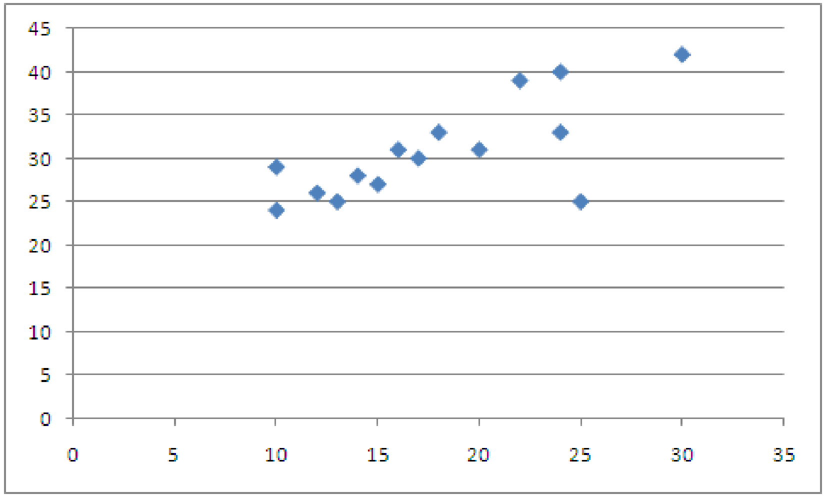
Os dados devem ser coletados pareados. As linhas representam as observações coletadas e as colunas os fatores controláveis (X's) e a Variável de resposta (Y)

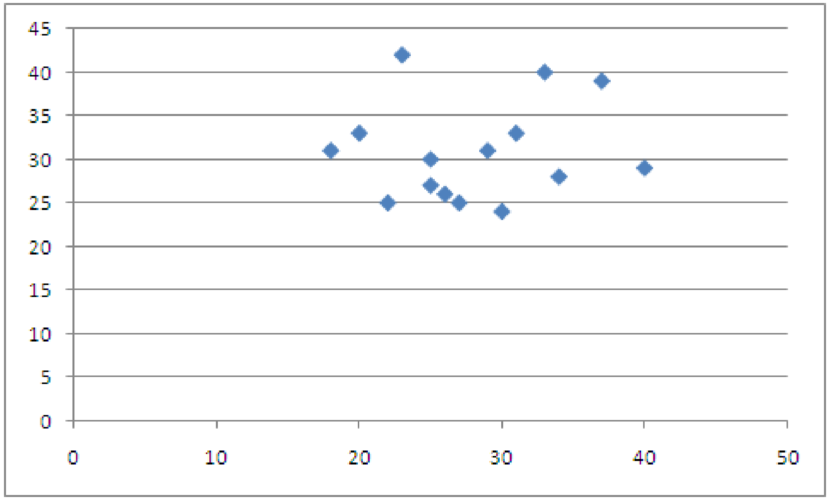
X_1: número de caixas	X_2: distância	Y: tempo
10	30	24
15	25	27
10	40	29
20	18	31
25	22	25
18	31	33
12	26	26
14	34	28
16	29	31
22	37	39
24	20	33
17	25	30
13	27	25
30	23	12
24	33	40

LOPP
UFRGS

Análise preliminar dos dados

- **Fazer um gráfico de dispersão das variáveis independentes *versus* variável dependente**





X_1=num de caixas **X_2 =*distância***

LOPP
UFRGS

Análise do modelo

Teste da significância do modelo de regressão

- Realiza-se o teste de hipótese *F* para confirmar a "inexistência de relação entre X e Y ".

H_0: $\beta_1 ; \beta_2 ; ...; \beta_k = 0$

H_1: $\beta_{ji} \neq 0$ para pelo menos uma variável independente

- Calcula-se o $F_{calculado}$ fazendo F=*MQReg/ MQR*
- A hipótese nula será rejeitada quando:

$F_{calculado} > F_{\alpha/2,\, k,\, n-k-1}$ *ou* *valor-p<0,05 (5%)*

ANOVA

	gl	*SQ*	*MQ*	*F*	*F de significação*
Regressão	2	331,3585994	165,6792997	$\boxed{16,7954}$	$\boxed{0,000332524}$
Resíduo	12	118,3747339	9,864561157		
Total	14	449,7333333			

LOPP
UFRGS

Identificação dos fatores significativos

Teste individual sobre a significância de cada parâmetro b_i

Se os resíduos seguem o modelo normal, os parâmetros b_i também irão seguir esse modelo, ou seja:

$$b_i \rightarrow N\left(\beta_i, \sigma_{bi}^2\right)$$

De modo que para testar as hipóteses

$$H_0: \beta_i = 0$$

$$H_1: \beta_i \neq 0$$

Usa-se a distribuição de *Student*, calculando:

$$t_i = b_i / Sb_i$$

A hipótese nula será rejeitada quando:

$$\left|t_i\right| > t_{\alpha/2\ ,\ n-k-1}$$ ***ou valor-p < 0,05 (5%)***

O valor-p representa a probabilidade de errar na afirmação de que o fator tem efeito significativo quando não tem

LOPP
UFRGS

Identificação dos fatores significativos

	Coeficientes	*Erro padrão*	*Stat t*	*valor-P*
Interseção	2,31120209	5,857302725	0,394584709	0,70007
Variável X 1	0,87720461	0,153034597	5,732067324	9,4E-05
Variável X 2	0,45592077	0,146762335	3,106524374	0,00908

Todos os termos são significativos pois os *valores-p* são menores que 0,05 (5%), resultando na equação:

$$Y = 2,31 + 0,877X_1 + 0,459X_2$$

Os termos que não são significativos não devem permanecer no modelo

Retirar um termo por vez e rodar novamente a rotina de regressão até definir uma equação com apenas termos significativos

LOPP
UFRGS

Análise dos coeficiente de determinação R_2

R^2 é uma medida de quão bem a equação de regressão múltipla se ajusta as dados amostrais;

R^2 indica a percentagem da variabilidade total que é explicada pelo modelo de regressão.

- **Se $R^2 = 1$, todas as observações estarão sobre o hiperplano definido pelo modelo.**
- **Se $R^2 = 0$, não há nenhuma relação entre a variável de resposta e as variáveis independentes.**
- **No exemplo, R^2 =*0,7367,ou seja,* 73,67% da variabilidade do fenômeno pode ser explicado pelo modelo de regressão**

Estatística de regressão	
R múltiplo	0,85836418
R-Quadrado	0,73678906
R-quadrado ajustado	0,69292057
Erro padrão	3,14078989
Observações	15

LOPP
UFRGS

Análise do R^2 ajustado

***R*²** **ajustado faz um ajuste no *R*² considerando o número de variáveis e o tamanho amostral.**

$$R^2_{ajustado} = 1 - \frac{(n-1)}{[n-(k+1)]}(1-R^2)$$

Onde: n é o tamanho da amostra e k é o número de variáveis independentes

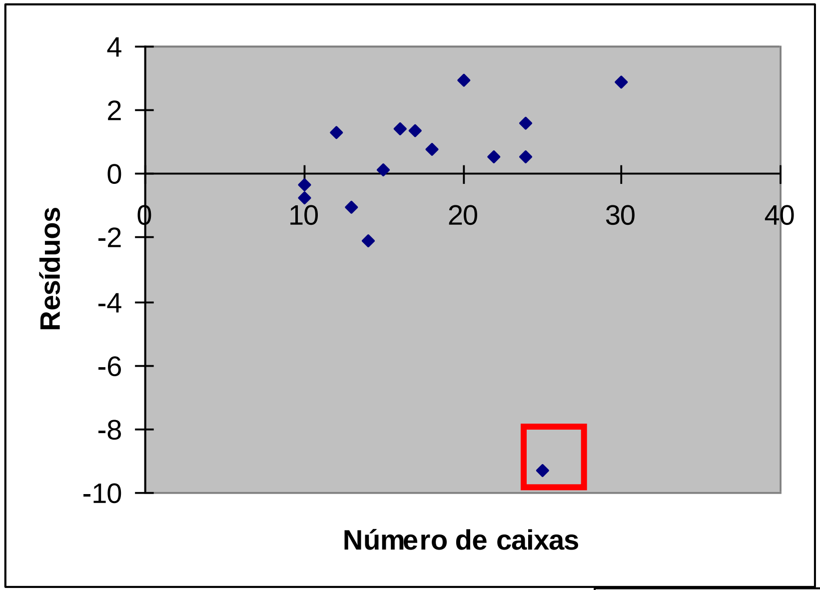
Verificar o valor do *R*² ajustado e compará-lo com o valor do *R*²

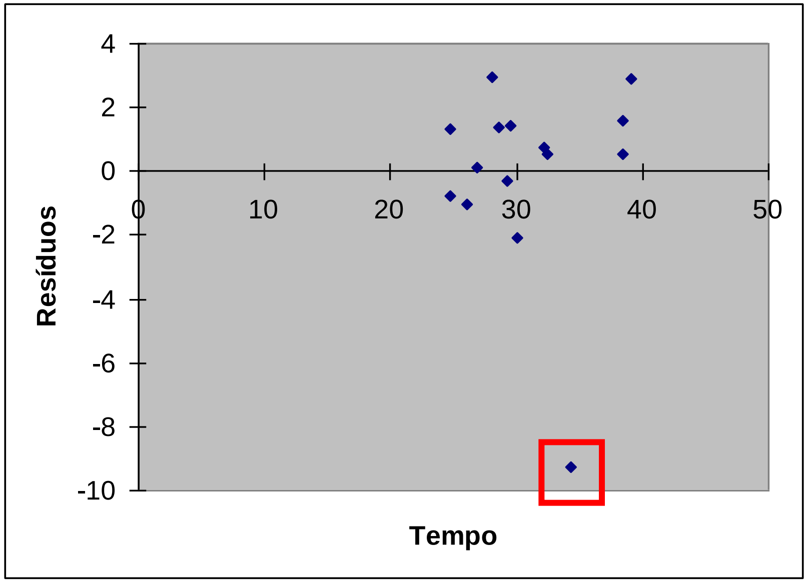
Se esses valores forem muito diferentes, pode-se afirmar que há um excesso de variáveis no modelo.

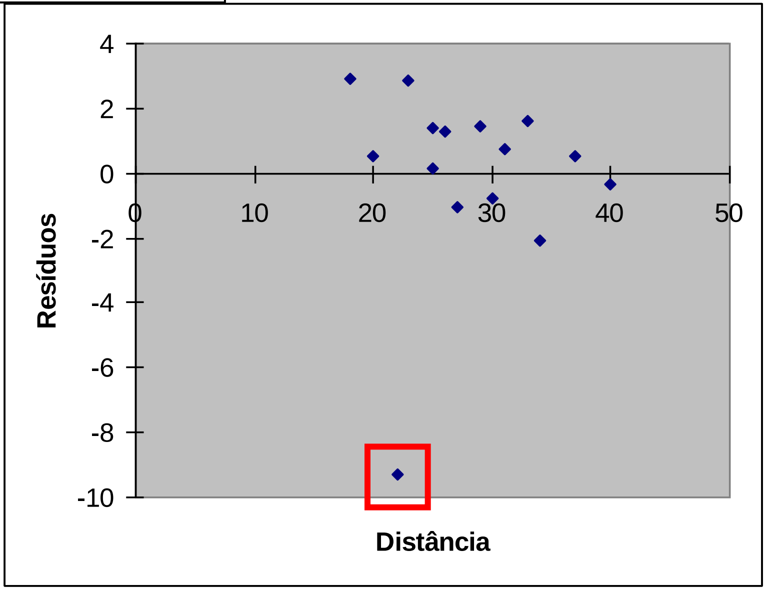
Estatística de regressão	
R múltiplo	0,85836418
R-Quadrado	0,73678906
R-quadrado ajustado	0,69292057
Erro padrão	3,14078989
Observações	15

LOPP
UFRGS

Análise gráfica dos resíduos







LOPP
UFRGS

Análise dos gráficos dos resíduos

Os resíduos do ponto 5 tem comportamento diferente dos resíduos dos outros pontos

Este fato pode indicar que este ponto não pertence a esse grupo de dados

Se houver registro de alguma causa especial que tenha afetado esta entrega em particular, essa observação pode ser eliminada do conjunto e a análise poderia ser refeita, possivelmente fornecendo um modelo mais preciso.

Dados atípicos

Seguindo na análise, observar a coluna dos resíduos padronizados para verificar a existência de valores atípicos.

Se isso acontecer o dado atípico deve der eliminado e a rotina de regressão deve ser rodada novamente

Observação	*Y previsto*	*Resíduos*	*Resíduos padrão*
1	24,7608713	-0,760871317	-0,261665052
2	26,8672905	0,132709479	0,045639035
3	29,320079	-0,320079022	-0,110075767
4	28,0618682	2,938131815	1,010428962
5	34,2715743	-9,271574323	-3,188511547
6	32,234429	0,765571022	0,263281289
7	24,6915975	1,308402542	0,449962053
8	30,0933728	-2,093372844	-0,71991479
9	29,5681782	1,431821786	0,492406159
10	38,478772	0,521227954	0,179251257
11	32,4825282	0,517471829	0,177959518
12	28,6216997	1,378300256	0,474000006
13	26,0247228	-1,02472284	-0,352404079
14	39,1135182	2,88648185	0,992666443
15	38,4094982	1,590501813	0,546976513

Codificação dos níveis

- **Nas equações de regressão múltipla, pode ser interessante comparar os efeitos dos diferentes fatores controláveis (termos) da equação;**
- **Neste caso, é necessário padronizar o intervalo de variação dos diferentes termos da equação, para que os coeficientes sejam diretamente comparáveis entre si;**
- **É necessário converter os níveis reais do intervalo de investigação em níveis codificados do intervalo**
- **O nível baixo será o nível –1 e o nível alto será o nível +1**

Fórmula para codificação dos níveis

- **Fórmula para converter os níveis reais (NR) em níveis codificados (NC):**

$$NC = \frac{NR - VC}{((LSI - LII)/2)} = \frac{NR - ((LSI - LII)/2 + LII)}{((LSI - LII)/2)}$$

- ***VC* representa o valor central do intervalo investigado;**

- ***LSI* representam o limite superior do intervalo investigado;**

- ***LII* representam o limite inferior do intervalo investigado.**

LOPP
UFRGS

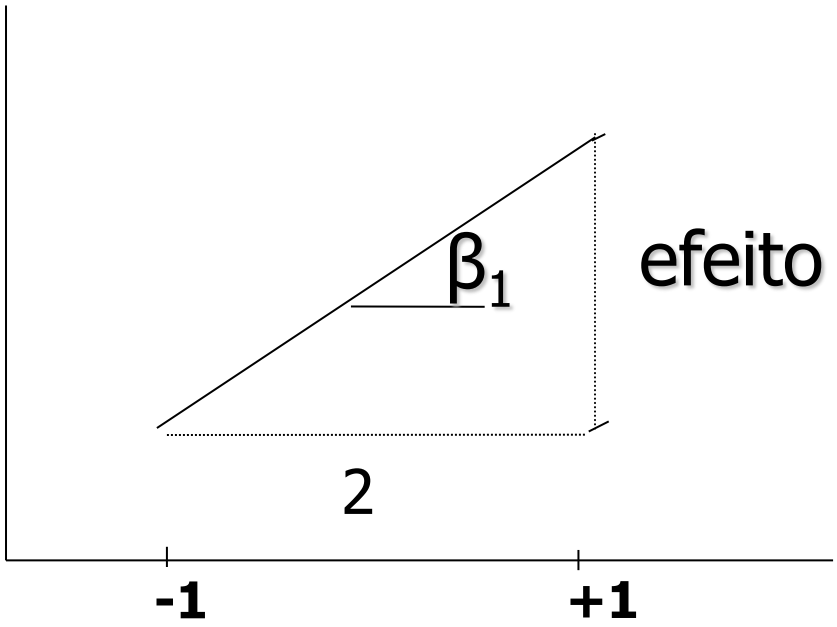
Codificação dos níveis

Por exemplo, o intervalo de investigação da temperatura é de 100 ºC a 120 ºC

- **LII = 100 ºC e o LSI = 120 ºC.**
- **O valor central é calculado como:** $VC = \dfrac{120-100}{2} + 100 = 110$
- **O nível 100ºC é calculado como:** $NC = \dfrac{100-110}{(120-100)/2} = -1$
- O **nível 110ºC é calculado como** $NC = \dfrac{110-110}{(120-100)/2} = 0$
- O **nível 120ºC é calculado como** $NC = \dfrac{120-110}{(120-100)/2} = +1$

Coeficiente x efeito

- **Como os níveis variam de –1 a +1 o cálculo do coeficiente β_1 em uma equação de regressão equivale ao efeito do fator calculado na ANOVA dividido por dois (distancia entre -1 e +1).**



$$\beta_1 = \frac{efeitoANOVA}{2}$$

LOPP
UFRGS

Seleção de variáveis do modelo

Quando existir um número muito grande de variáveis candidatas ao modelo (n>30 ou 40), a análise do modelo de regressão pode tornar-se muito complexa por parte do usuário.

Dessa forma, sugere-se a utilização de algum método que pré-selecione as variáveis mais importantes para caracterizar a variável resposta.

Existem 3 categorias diferentes do método passo-a-passo:

(i) seleção progressiva,

(ii) eliminação regressiva e

(iii) regressão por etapas.

A estatística F parcial, que é o valor do teste F para a variável xi dado que as variáveis xj (j≠i) já estão no modelo:

$$F_i = \frac{SQ\,\mathrm{Re}\,g(\beta_i / \beta_0, \beta_j)}{QMR(x_j, x_i)}$$

Seleção de variáveis do modelo

A seleção progressiva consiste na adição gradual de variáveis controláveis ao modelo. Esse método inicia com o modelo com apenas o parâmetro de intercepto e, a seguir, seleciona-se a variável de controle que apresenta maior coeficiente de correlação simples com a resposta.

Caso o F parcial para essa variável for maior que fentra, deve-se adicionar essa variável ao modelo.

A cada etapa deste método, seleciona-se a variável que apresente maior correlação com a resposta, essa é adicionada desde que o valor do teste parcial F seja maior que fentra.

O procedimento finaliza assim que nenhuma variável puder ser adicionada ao modelo

Seleção de variáveis do modelo

A eliminação regressiva consiste na remoção gradual de variáveis controláveis do modelo que contenha todas as variáveis candidatas.

Como primeiro passo deve-se calcular o valor F parcial para todas as variáveis como se cada uma delas fosse a última a entrar no modelo.

A seguir, seleciona-se a variável que possua o menor valor parcial F, se esse valor for menor que fsai, deve-se retirar essa variável do modelo.

A cada etapa desse método, seleciona-se a variável que apresentar menor valor parcial F, essas são removidas sempre que o valor do teste parcial F for menor que fsai.

Quando não existirem variáveis a serem excluídas do modelo, o método se finaliza

LOPP
UFRGS

Seleção de variáveis do modelo

A regressão por etapas é a combinação da seleção progressiva e da eliminação regressiva.

Inicia-se com o modelo com nenhuma variável e a cada etapa escolhe-se a variável com maior correlação com a resposta, que deverá ser adicionada ao modelo se $F_i > f_{entra}$.

Caso a variável x_i seja inserida no modelo, analisa-se a possibilidade de retirada das outras variáveis já selecionadas dado que x_i está presente.

O procedimento termina assim que não existir nenhuma variável para ser incluída ou removida do modelo.

Seleção do melhor modelo

O primeiro critério é a escolha do modelo que maximiza o coeficiente de determinação.

Esse critério deve ser utilizado com cuidado, pois a adição de novas variáveis ao modelo acarreta um aumento do R2, sem de fato melhorar o modelo.

Um outro critério lógico é a escolha do modelo que apresente menor estimativa para os erros, ou seja, deve-se escolher o modelo que minimize MQR

Seleção do melhor modelo

Também é possível utilizar o coeficiente de determinação ajustado. Essa estatística é uma modificação do coeficiente de determinação, que considera também o número de variáveis do modelo, penalizando modelos que possuam mais variáveis.

O segundo critério seria a escolha do modelo que apresente esses dois coeficientes aproximadamente iguais.

Dessa forma, os critérios para escolha de modelos devem balancear a minimização do SQRp com o aumento da complexidade da equação.

FUNÇÃO DE PERDA MULTIVARIADA

Introdução a Função de perda Multivariada

- Na maioria dos estudos experimentais, existe mais de uma variável de resposta de interesse, exigindo o uso de algum procedimento multivariado na busca do ajuste ótimo dos fatores controláveis.
- O procedimento que será mostrado a seguir baseia-se na utilização da Função de Perda Multivariada como função objetivo a ser otimizada.
- Trata-se de um procedimento bastante genérico que irá fornecer resultados consistentes na maioria das aplicações práticas.

Função de Perda

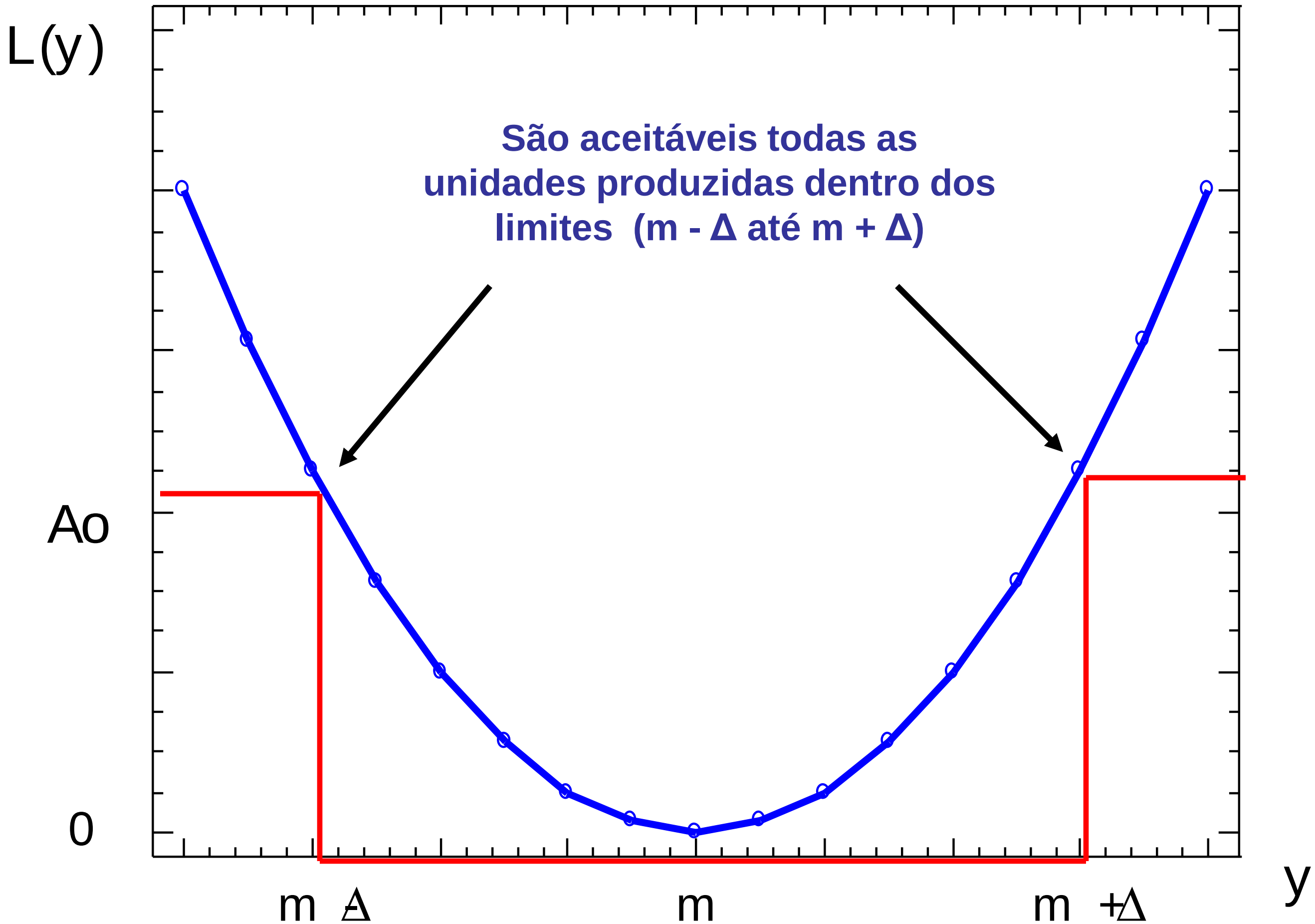
- **A função de perda é empregada para quantificar a perda que um produto impõem à sociedade pela falta de qualidade.**
- **Em muitos caso, essa perda resulta aproximadamente proporcional ao quadrado do desvio da meta estabelecida para uma certa característica de qualidade, ou seja:**

$$Z_i = K_j \left[(Y_j - T_j)^2 \right]$$

- Z é o valor que a função de perda “Z” assume para um dado ajuste “i” do conjunto dos fatores controláveis;
- Y é o valor observado da VRj no ajuste i
- T ou m é o valor alvo da VRj

LOPP
UFRGS

Gráfico da Função de Perda Quadrática



LOPP
UFRGS

Ponderação das VR

- **Na otimização, é preciso atribuir pesos a cada VR. Esses pesos K, têm duas funções:**

a) normalizar os valores que representam os desvios do alvo, obtidos nas unidades de grandeza da VR, para que os desvios de todas as VR possam ser diretamente comparáveis;

b) considerar a importância relativa (IR_j) de cada VR.

LOPP
UFRGS

Ponderação das VR

- **Para todas as variáveis de resposta Y_j, deve-se conhecer o valor alvo, os limites de especificação e a importância relativa (IR_j)**

- **Existem três tipos de variáveis de resposta**

Nominal-é-melhor

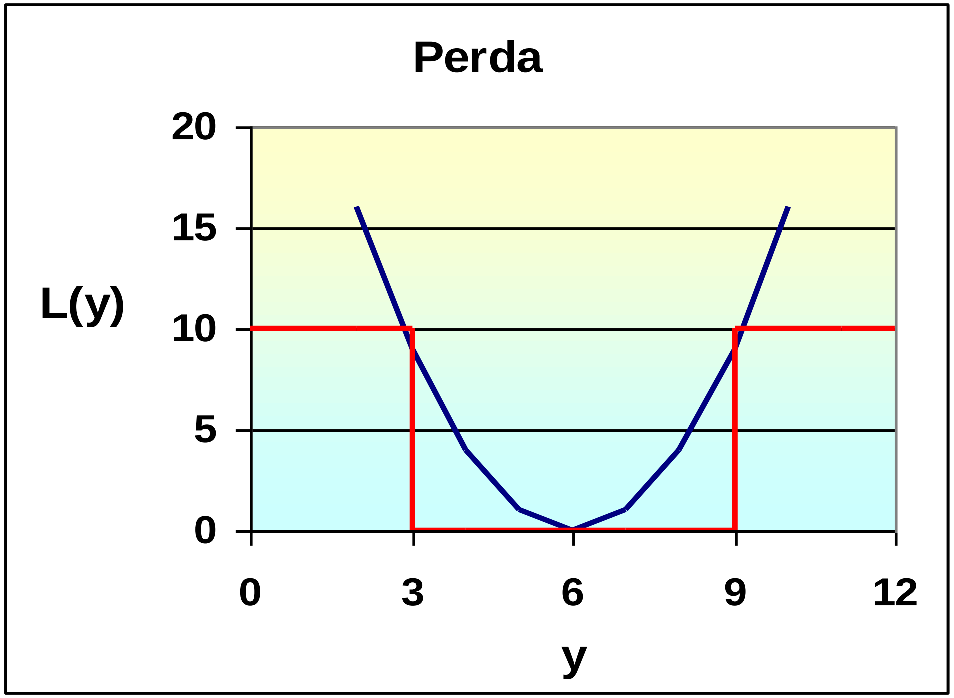
Maior-é-melhor

Menor-é-melhor

Função de Perda Nominal-é-melhor

- **Nominal-é-melhor** se refere às características que têm um valor alvo e qualquer desvio deste valor (para cima ou para baixo) incorre em uma perda de qualidade

T = 6,0, Especificações: 6,0 +/- 3,0



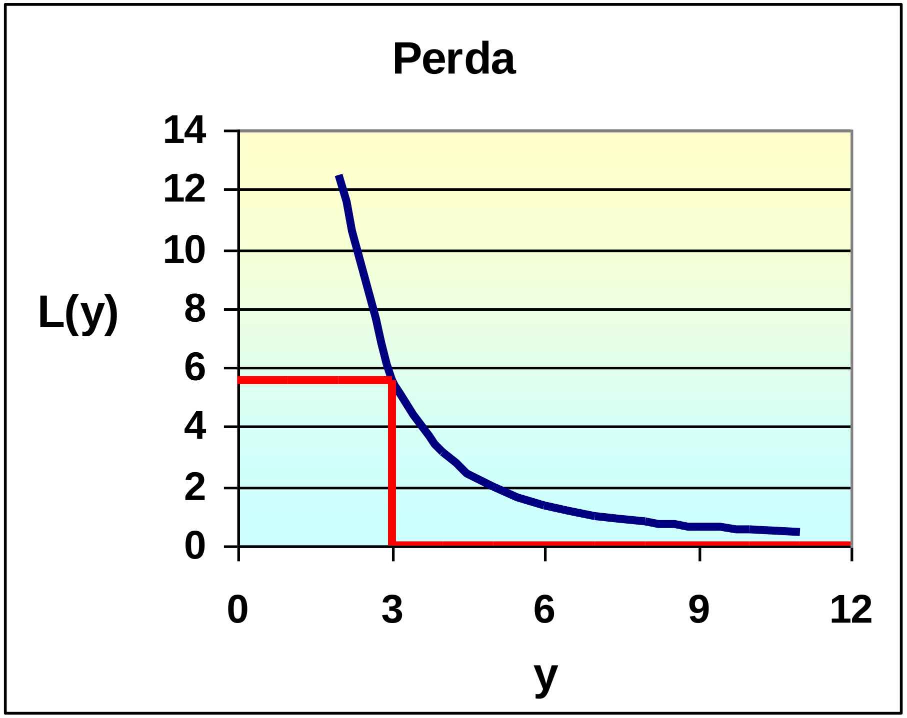
$$K_j = \frac{IR_j}{((LSE - LIE)/2)^2}$$

LOPP
UFRGS

Função de Perda Maior-é-melhor

- **Maior-é-melhor** se refere às características que têm um valor mínimo estabelecido e, se esse valor for superado tanto melhor. Não existe limite superior de especificação.

Alvo = 12,0, Limite inferior: 3,0



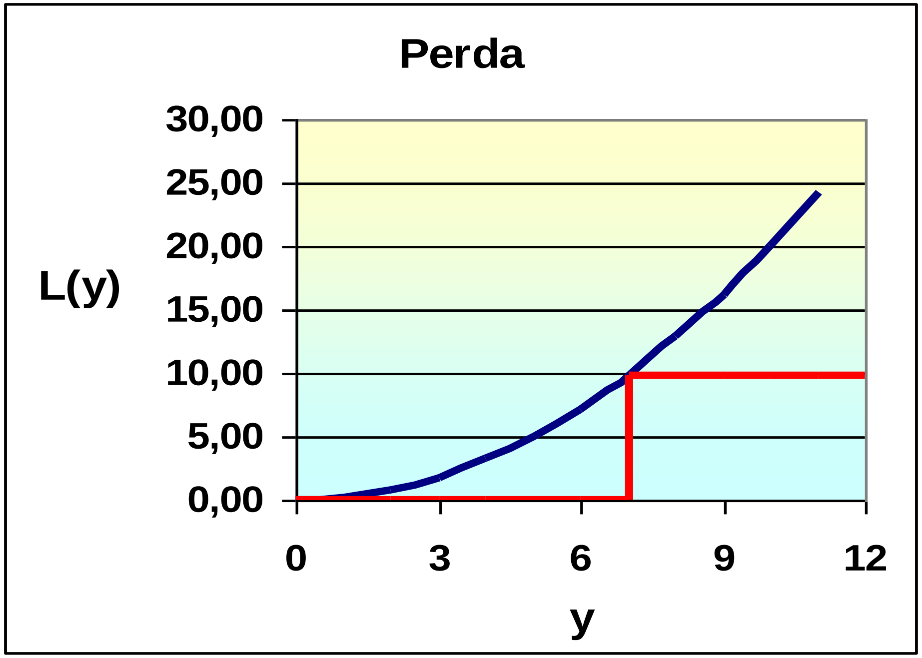
$$K_j = \frac{IR_j}{(\text{alvo} - LIE)^2}$$

LOPP
UFRGS

Função de Perda Menor-é-melhor

Menor-é-melhor se refere às características que têm um valor máximo estabelecido e, se esse valor for menor tanto melhor. Não existe limite inferior de especificação.

Alvo = 1,0, Limite superior: 7,0



$$K_j = \frac{IR_j}{(LSE - alvo)^2}$$

Função Perda Multivariada

- **A expressão da função de perda multivariada para múltiplas VR é a seguinte:**

$$\hat{Z}_i = \sum_{J=1}^{J} K_j \left[(\hat{Y}_j - T_j)^2 \right]$$

- $\hat{Z}_i$ é o valor que a função de perda "Z" assume para um dado ajuste "i" do conjunto dos fatores controláveis;

Função Perda Multivariada

- K_j é a ponderação atribuída a VR "j";
- $\hat{Y}_j$ é o modelo matemático que fornece uma estimativa da média da VR "j" em função do ajuste i dos fatores controláveis;
- Tj é o valor alvo para a VR "j";
- Nota: para VR do tipo maior é melhor ou menor é melhor, quando o valor de $\hat{Y}_j$ supera o alvo, atribui-se zero para o correspondente desvio do alvo;

LOPP
UFRGS

Exemplo de Função Perda Multivariada

- **Exemplo da função de perda para duas VR:**
- **primeira VR=Produtividade do tipo Maior-é-melhor**
- **segunda VR=Qualidade do tipo Nominal-é-melhor**

$$\hat{Z}(i) = \frac{IR_1}{(alvo_1 - LIE_1)^2} \times \left(\hat{Y}_1 - T_1\right)^2 + \frac{IR_2}{\left((LSE_2 - LIE_2)/2\right)^2} \times \left(\hat{Y}_2 - T_2\right)^2$$

LOPP
UFRGS

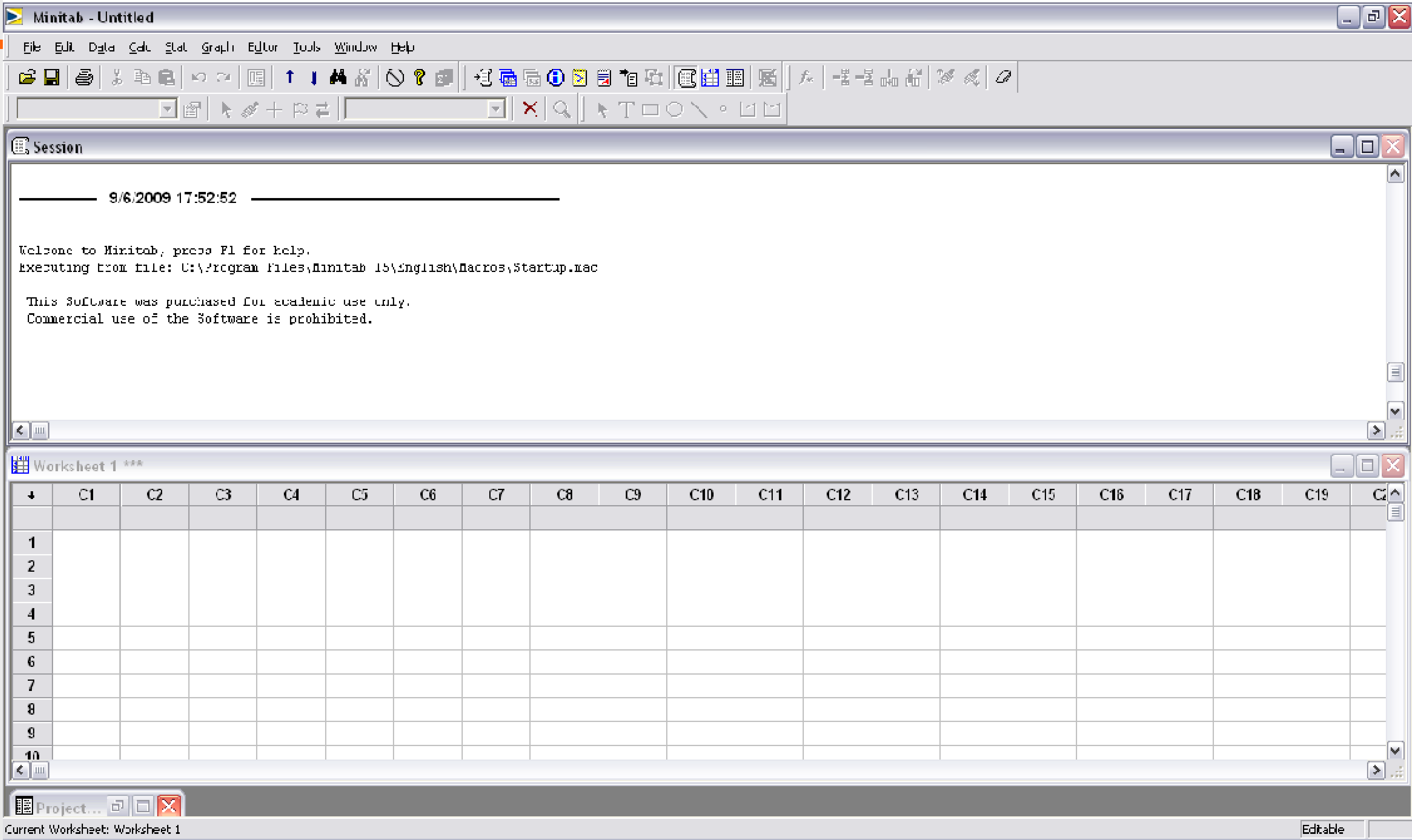
Notas sobre a Função perda Multivariada

- Vale observar que a perda é função das equações das VR , que por sua vez são função dos fatores controláveis X.
- Logo, em última análise tem-se que a perda é função de X, ou seja, é função do ajuste dos fatores controláveis.
- Observa-se também que a perda cresce quadraticamente quando qualquer VR afasta-se do alvo (ou em regiões onde aumenta a variabilidade das VR).
- Assim, o objetivo é encontrar o ajuste dos fatores controláveis que minimiza a função perda.

Função perda Multivariada

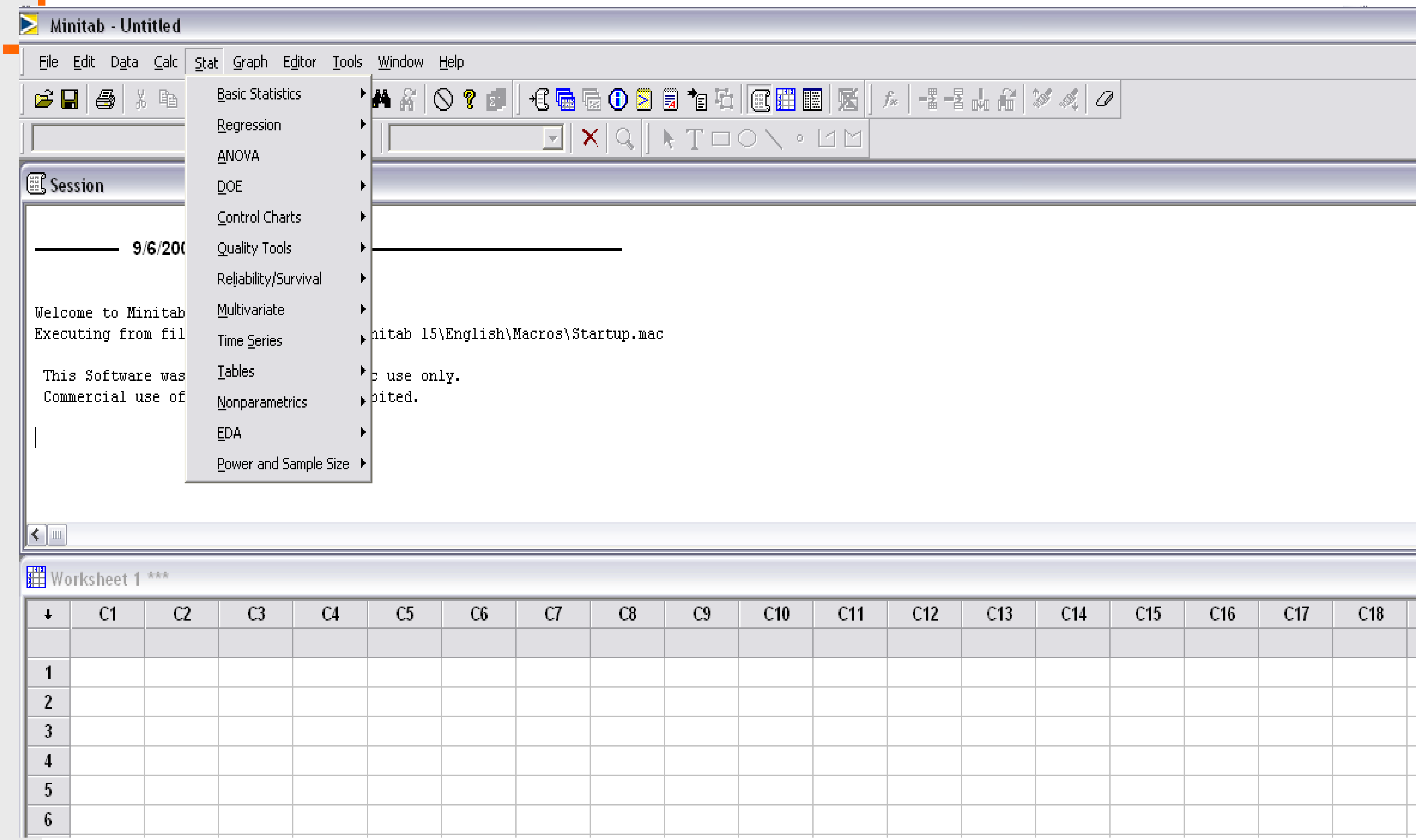
- Este ajuste ótimo estará associado a uma região onde as VR estão próximas de seus respectivos alvos (ou em regiões onde a variabilidade é pequena).
- Em projetos com muitos fatores, a busca do ponto ótimo exige suporte computacional.
- O Solver do Excel pode ser usado para identificar a melhor combinação dos fatores controláveis que otimiza o conjunto das variáveis de resposta simultaneamente.

Software Minitab 15



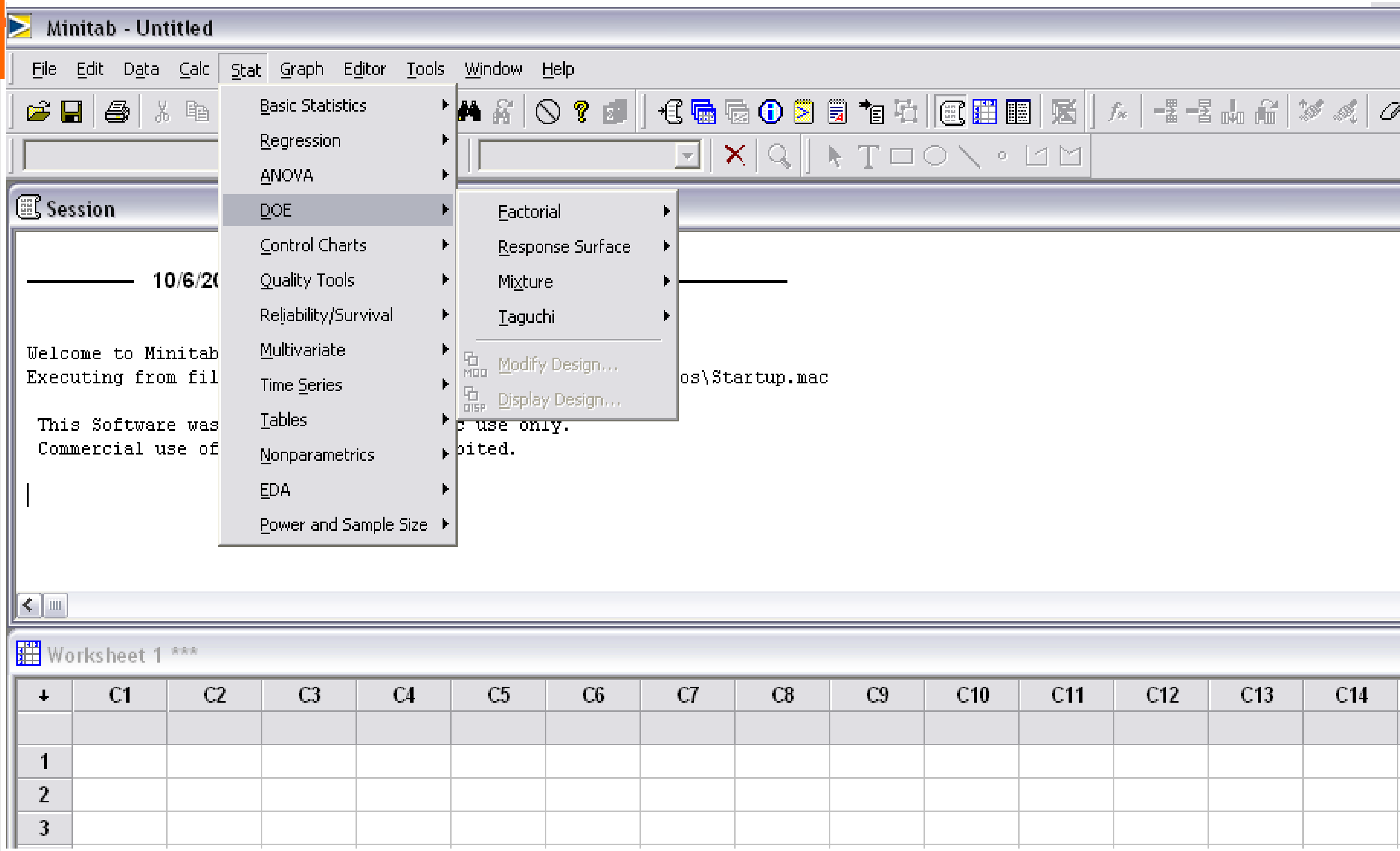
- **Desenvolvido pelo MIT (Massachussets Institute of Technology) em 1972, tem o principal foco em aplicações em engenharia.**

Menu Stat



- Onde se localiza as ferramentas estatísticas do software, tais como: Estat. básica; Regressão; Análise de Variância; dentre outras.

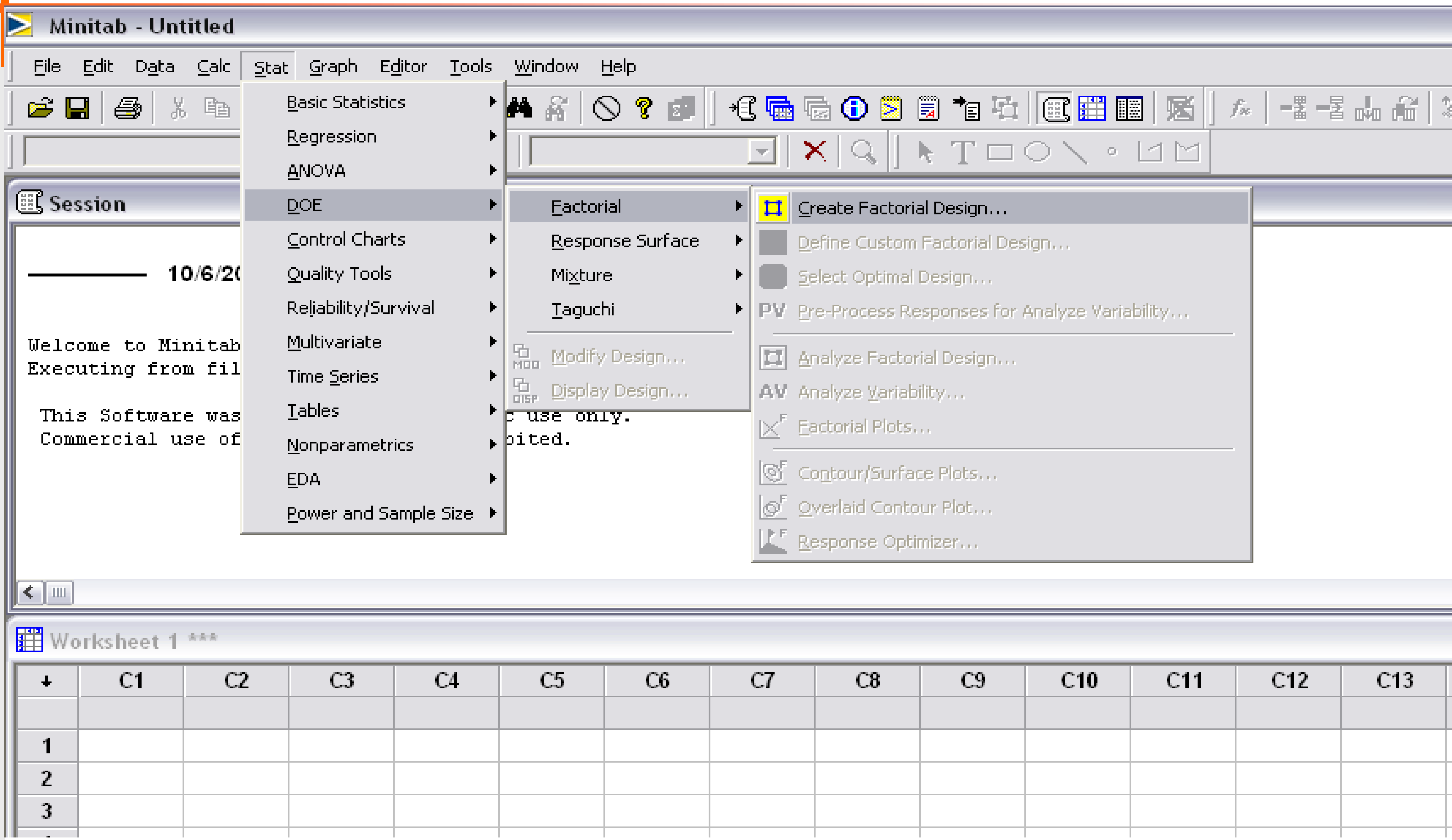
Projeto de Experimentos



- **Design que o Software permite;**
- **Caminho: Stat > DOE.**

LOPP
UFRGS

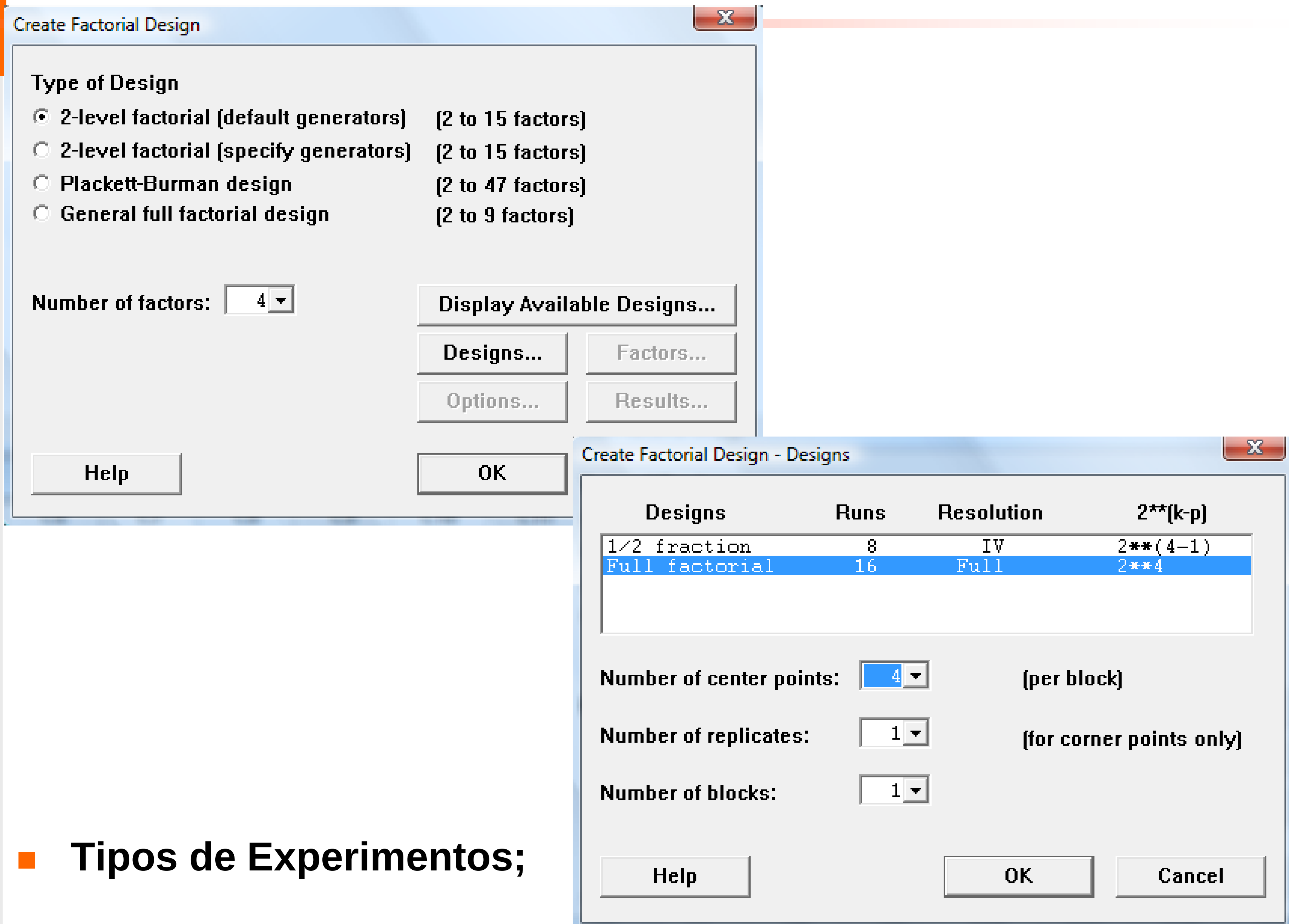
Planejar um Experimento



- Ex. Projeto de experimentos 2^k;
- Caminho: Stat > DOE > Factorial > Create.

LOPP
UFRGS

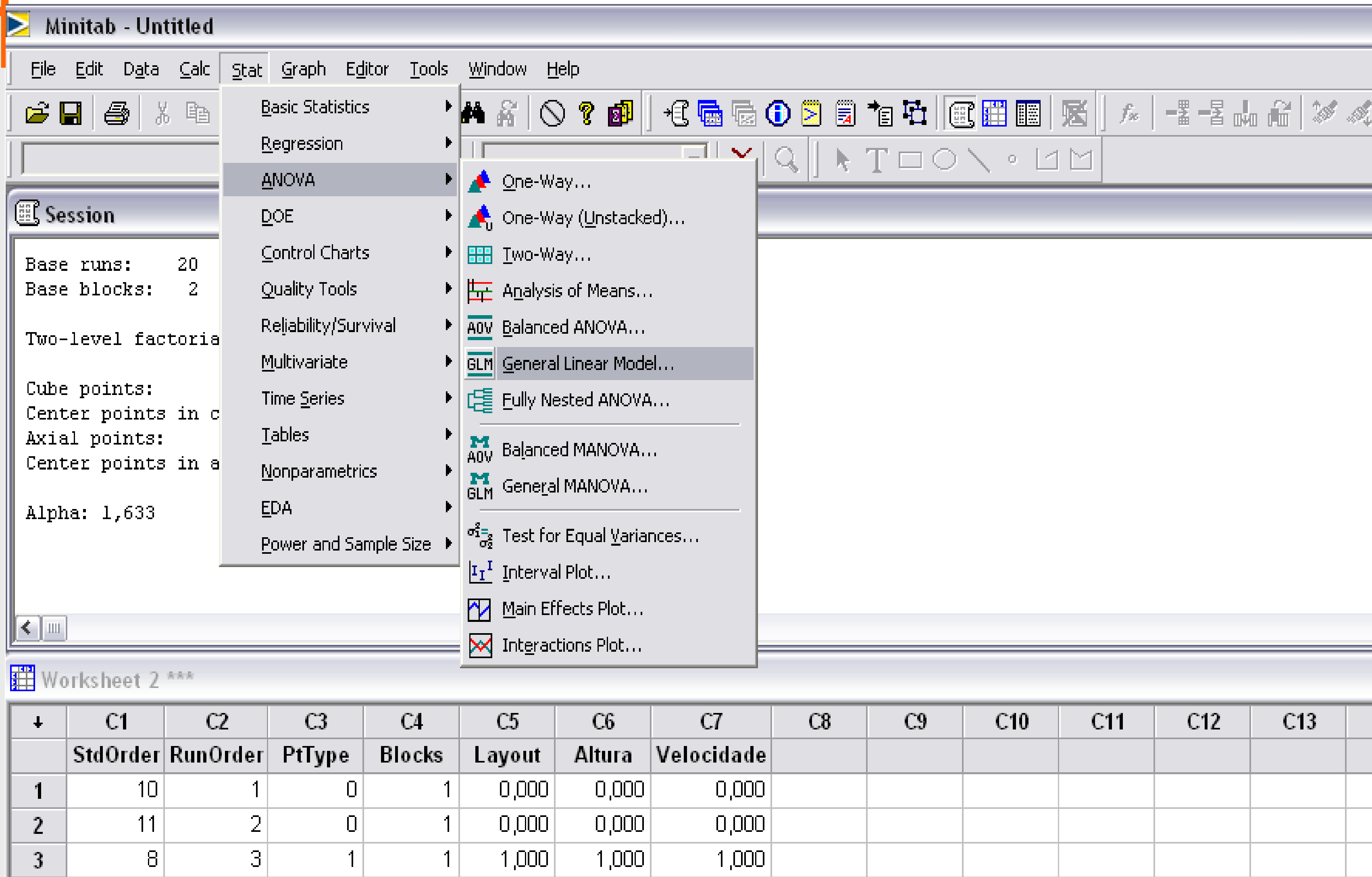
Planejar um Experimento



- **Tipos de Experimentos;**

LOPP
UFRGS

Analisar o Experimento



- Caminho: Stat > ANOVA > General Linear Model.

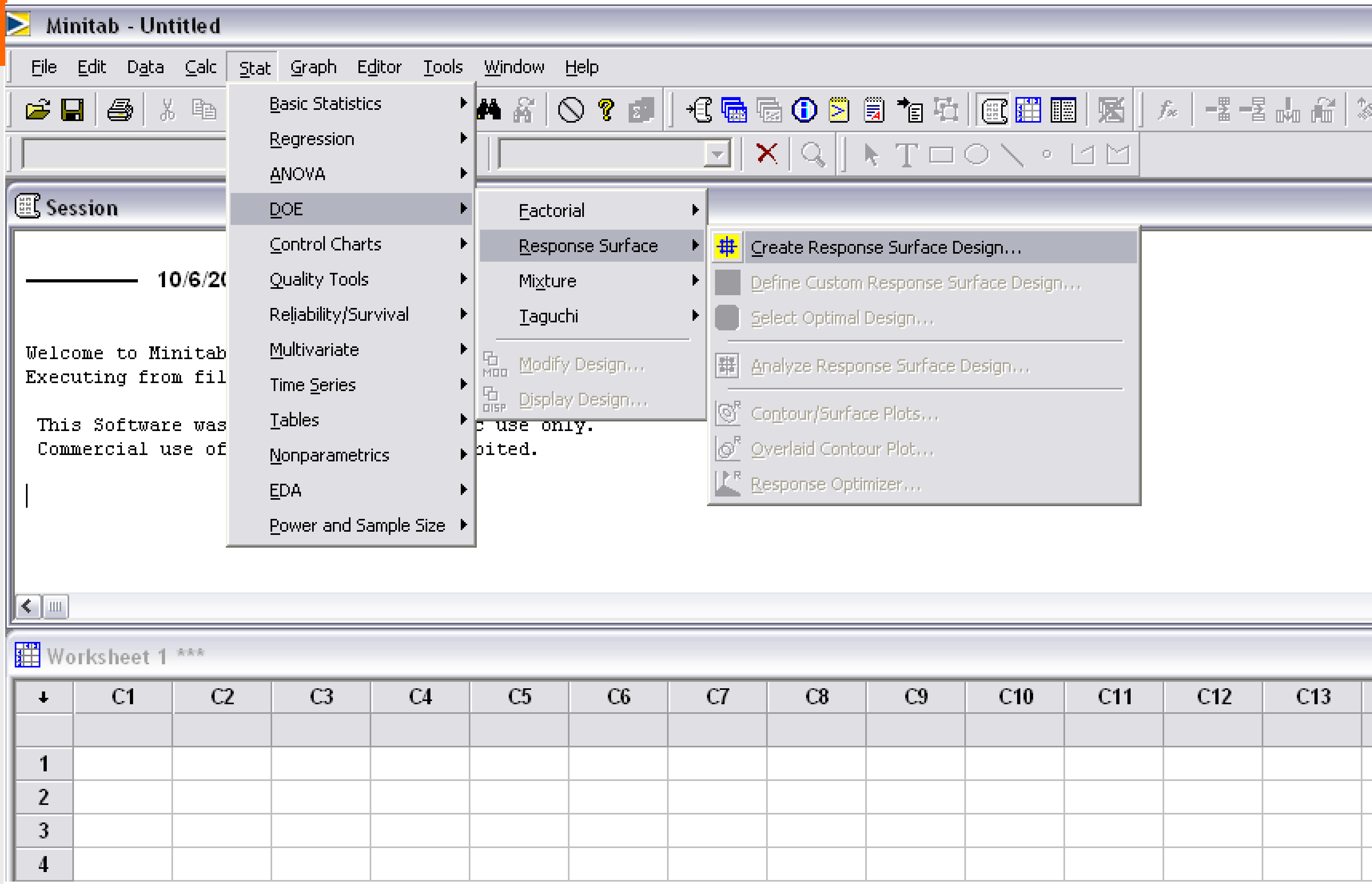
LOPP
UFRGS

Estudo Experimental

- Utilizar o software Minitab para Planejar, Analisar e Otimizar o estudo experimental;
- Exemplo: Excel 10.2
- Característica de Qualidade:
  - Y1: Produtividade;
  - Y2: N. defeituosos.
- Fatores Controláveis:
  - X1: Layout;
  - X2: Altura da Bancada;
  - X3: Velocidade da Esteira;

LOPP
UFRGS

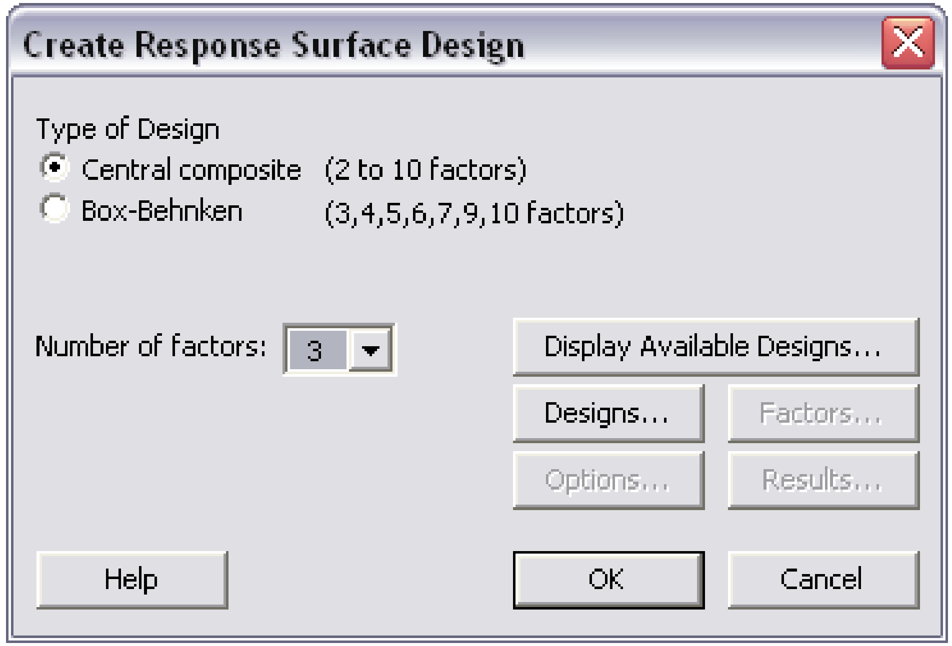
Planejar uma Superfície de Resposta



LOPP
UFRGS

- **Caminho: Stat > DOE > Response Surface > Create RSD.**

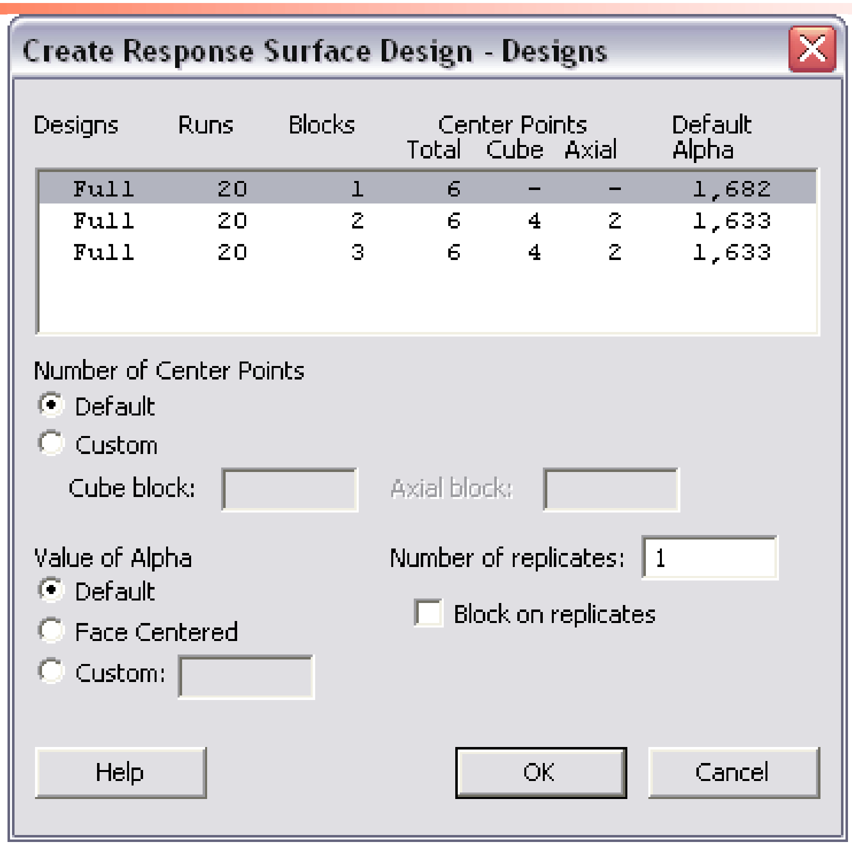
Planejar uma MSR



- **Selecione a opção Central Composite;**
- **Escolha o número de fatores (ex. 3);**
- **Clique em Designs.**
- **O restante das opções é semelhante ao Experimento 2^k**

LOPP
UFRGS

Planejar uma MSR



- Escolha o número de pontos centrais;
- Escolha do valor de Alpha;

LOPP
UFRGS

Resultado do Design

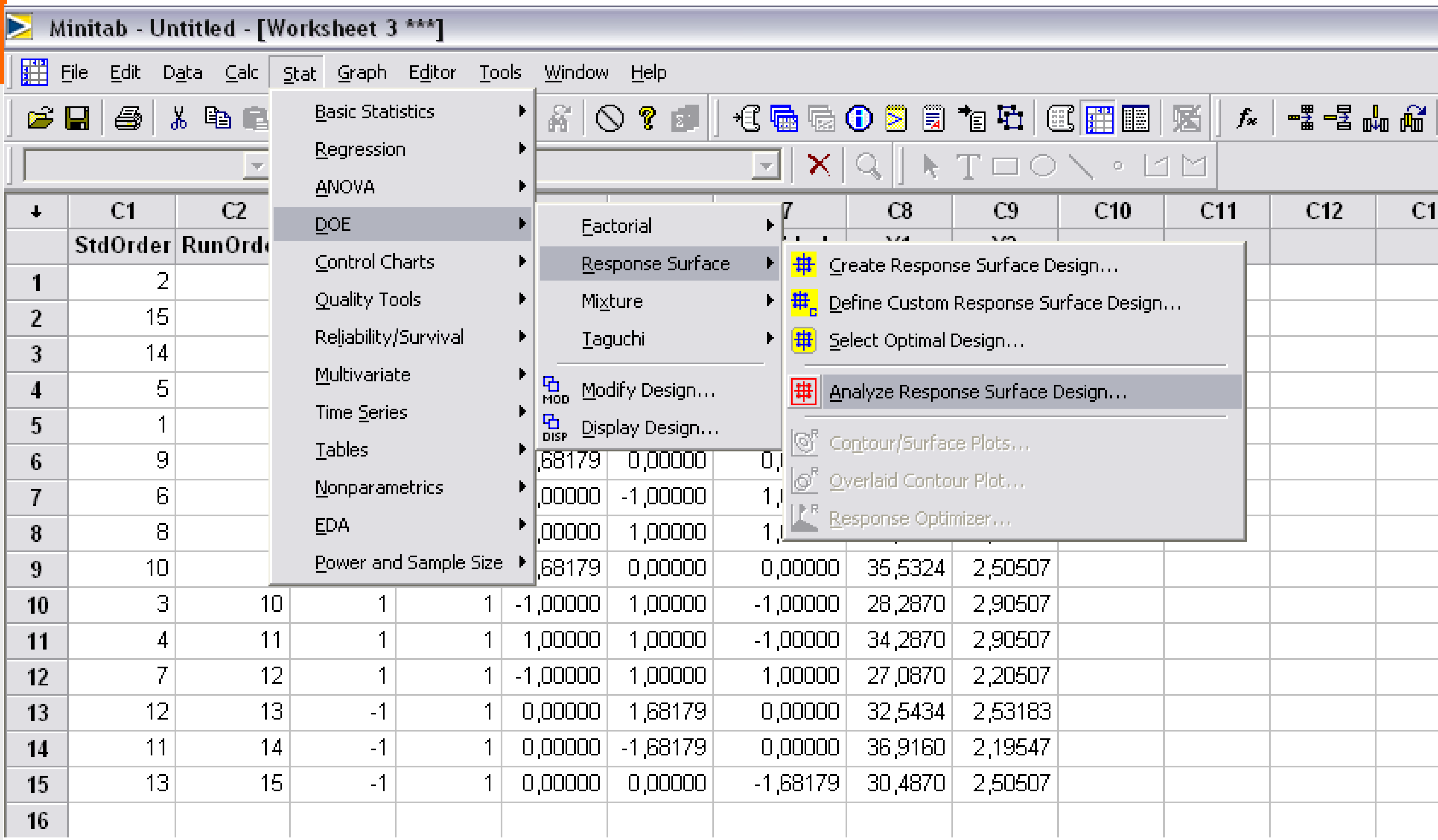
Minitab - Untitled - [Worksheet 3 ***]

File Edit Data Calc Stat Graph Editor Tools Window Help

↓	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13
	StdOrder	RunOrder	PtType	Blocks	Layout	Altura	Velocidade	Y1	Y2				
1	2	1	1	1	1,00000	-1,00000	-1,00000	35,6870	2,00507				
2	15	2	0	1	0,00000	0,00000	0,00000	30,4870	2,50507				
3	14	3	-1	1	0,00000	0,00000	1,68179	30,4870	2,50507				
4	5	4	1	1	-1,00000	-1,00000	1,00000	30,8870	2,70507				
5	1	5	1	1	-1,00000	-1,00000	-1,00000	29,6870	2,00507				
6	9	6	-1	1	-1,68179	0,00000	0,00000	25,4417	2,50507				
7	6	7	1	1	1,00000	-1,00000	1,00000	36,8870	2,70507				
8	8	8	1	1	1,00000	1,00000	1,00000	33,0870	2,20507				
9	10	9	-1	1	1,68179	0,00000	0,00000	35,5324	2,50507				
10	3	10	1	1	-1,00000	1,00000	-1,00000	28,2870	2,90507				
11	4	11	1	1	1,00000	1,00000	-1,00000	34,2870	2,90507				
12	7	12	1	1	-1,00000	1,00000	1,00000	27,0870	2,20507				
13	12	13	-1	1	0,00000	1,68179	0,00000	32,5434	2,53183				
14	11	14	-1	1	0,00000	-1,68179	0,00000	36,9160	2,19547				
15	13	15	-1	1	0,00000	0,00000	-1,68179	30,4870	2,50507				
16													

- Inserir Variáveis Resposta na Worksheet;

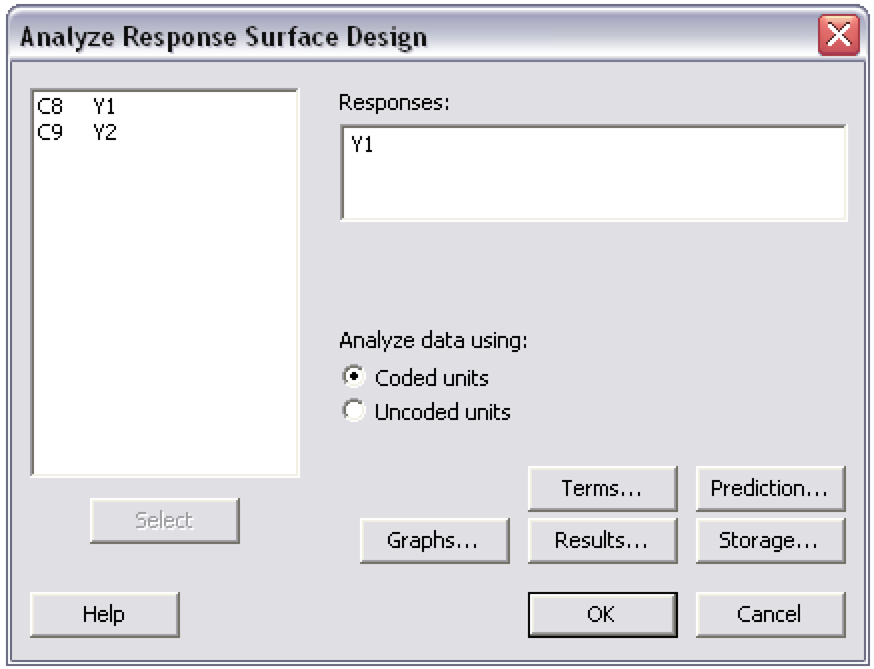
Analisar uma MSR



LOPP
UFRGS

- Caminho: DOE > Response Surface > Analyse RSD;
- É possível analisar por ANOVA ou Regression;

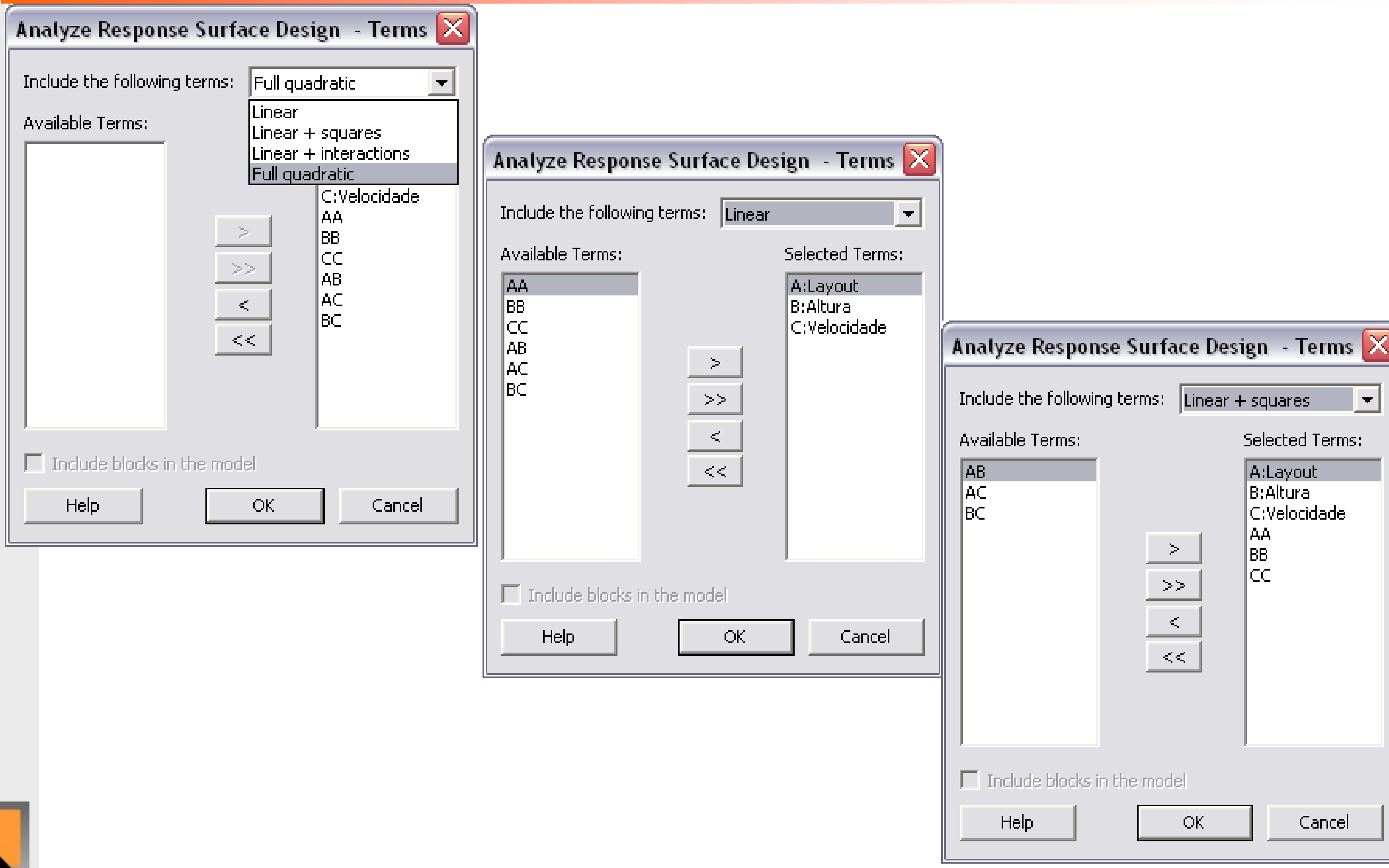
Analisar uma MSR



LOPP
UFRGS

- Insira a Variável Y desejada;
- Clique em “Terms” para selecionar o tipo de relação entre as Variáveis;

Analisar uma MSR



LOPP
UFRGS

- **Escolher o tipo de relação entre as Variáveis;**

Resultado da Análise “Y1”

Response Surface Regression: Y1 versus Layout; Altura; Velocidade

The analysis was done using coded units.

Estimated Regression Coefficients for Y1

Term	Coef	SE Coef	T	P
Constant	31,5767	0,3349	94,278	0,000
Layout	3,0000	0,4053	7,402	0,000
Altura	-1,3000	0,4053	-3,207	0,005
Velocidade	-0,0000	0,4053	-0,000	1,000

S = 1,49787   PRESS = 60,9089
R-Sq = 80,26%   R-Sq(pred) = 66,51%   R-Sq(adj) = 76,56%

Analysis of Variance for Y1

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Regression	3	145,992	145,992	48,6639	21,69	0,000
Linear	3	145,992	145,992	48,6639	21,69	0,000
Residual Error	16	35,898	35,898	2,2436		
Lack-of-Fit	11	35,898	35,898	3,2634	*	*
Pure Error	5	0,000	0,000	0,0000		
Total	19	181,889				

Unusual Observations for Y1

Obs	StdOrder	Y1	Fit	SE Fit	Residual	St Resid	
15	12	32,609	29,390	0,760	3,218	2,49	R
20	11	36,981	33,763	0,760	3,218	2,49	R

R denotes an observation with a large standardized residual.

Resultado da Análise “Y1”

Response Surface Regression: Y1 versus Layout; Altura

The analysis was done using coded units.

Estimated Regression Coefficients for Y1

Term	Coef	SE Coef	T	P
Constant	30,552	0,1213	251,826	0,000
Layout	3,000	0,1148	26,131	0,000
Altura	-1,300	0,1148	-11,324	0,000
Altura*Altura	1,500	0,1108	13,544	0,000

S = 0,424264   PRESS = 4,53737
R-Sq = 98,42%   R-Sq(pred) = 97,51%   R-Sq(adj) = 98,12%

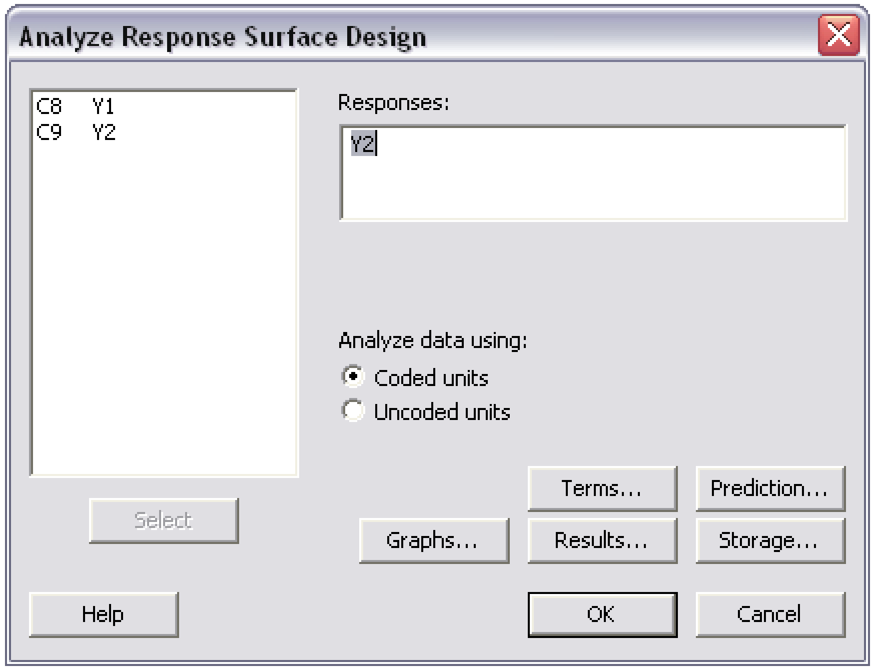
Analysis of Variance for Y1

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Regression	3	179,009	179,009	59,6698	331,50	0,000
Linear	2	145,992	145,992	72,9959	405,53	0,000
Square	1	33,018	33,018	33,0177	183,43	0,000
Residual Error	16	2,880	2,880	0,1800		
Lack-of-Fit	5	0,000	0,000	0,0000	0,00	1,000
Pure Error	11	2,880	2,880	0,2618		
Total	19	181,889				

- **Modelo Final de Y1: X1 + X2 + $X2^2$**
- **Interpretar os Resultados;**

LOPP
UFRGS

Analisar uma MSR



- Insira a Variável Y2;
- Clique em “Terms” para selecionar o tipo de relação entre as Variáveis;

LOPP
UFRGS

Resultado da Análise "Y2"

Response Surface Regression: Y2 versus Layout; Altura; Velocidade

The analysis was done using coded units.

Estimated Regression Coefficients for Y2

Term	Coef	SE Coef	T	P
Constant	10,6600	0,2778	38,379	0,000
Layout	0,0000	0,2628	0,000	1,000
Altura	4,0000	0,2628	15,219	0,000
Velocidade	-0,0000	0,2628	-0,000	1,000
Altura*Altura	-1,0000	0,2536	-3,944	0,001

S = 0,971295   PRESS = 27,0362
R-Sq = 94,28%   R-Sq(pred) = 89,07%   R-Sq(adj) = 92,75%

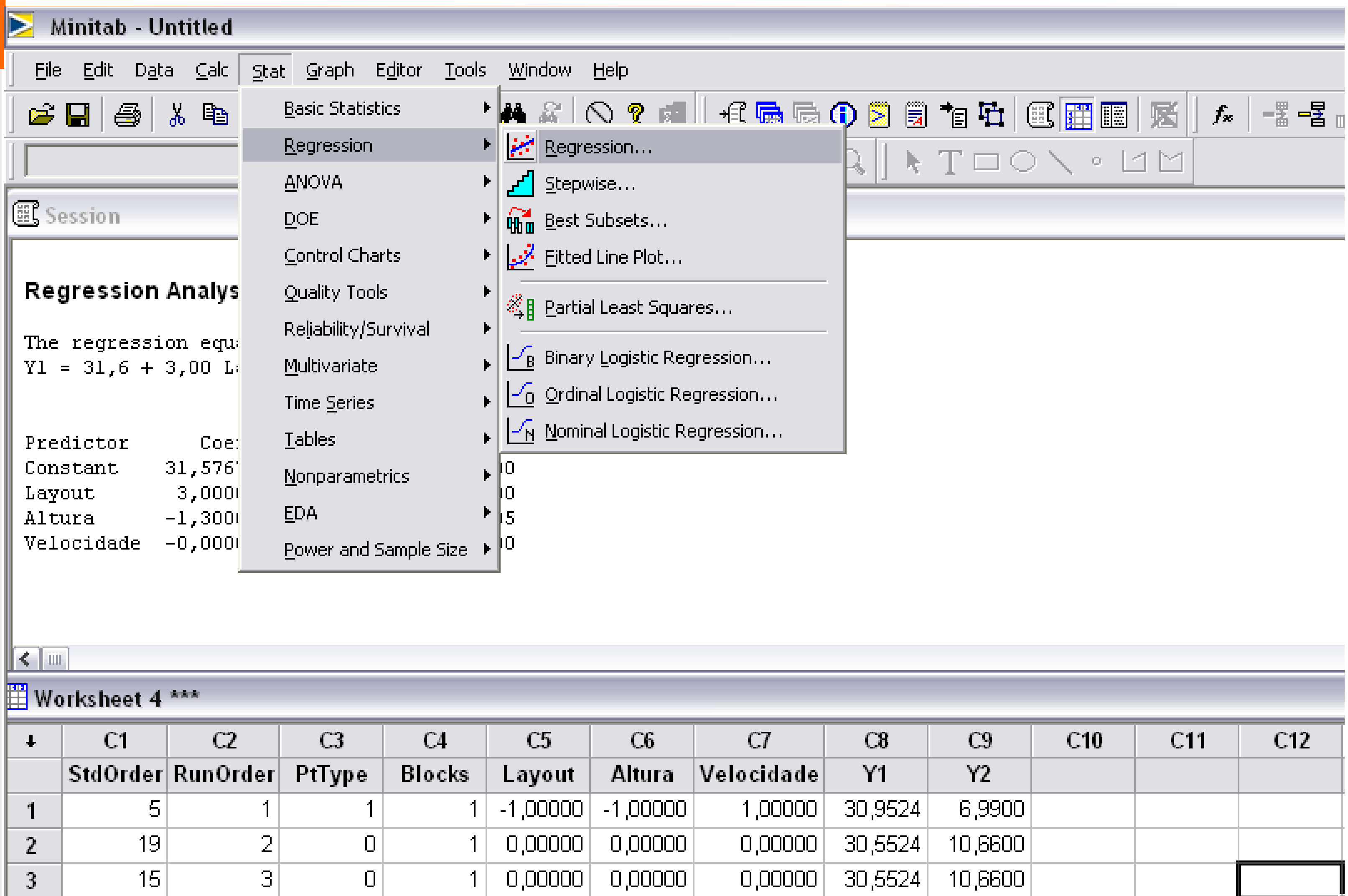
Analysis of Variance for Y2

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Regression	4	233,184	233,184	58,2960	61,79	0,000
Linear	3	218,510	218,510	72,8366	77,21	0,000
Square	1	14,675	14,675	14,6745	15,55	0,001
Residual Error	15	14,151	14,151	0,9434		
Lack-of-Fit	10	14,151	14,151	1,4151	*	*
Pure Error	5	0,000	0,000	0,0000		
Total	19	247,335				

- **Modelo Final de Y2: X2 + X2^2**

LOPP
UFRGS

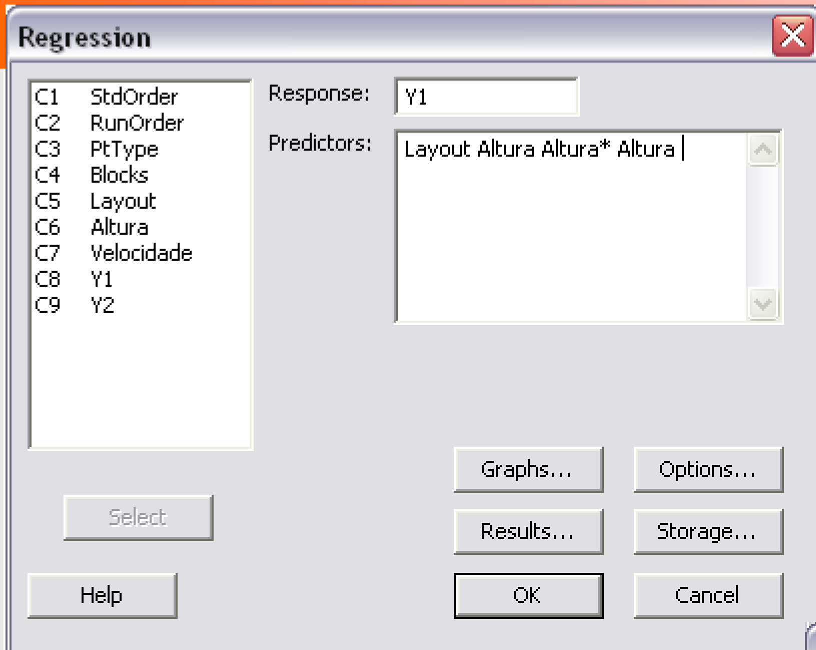
Analisar utilizando Regression

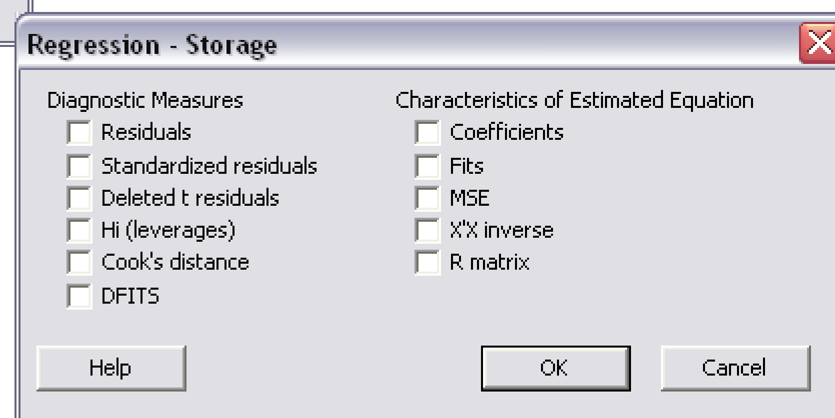


- Caminho: Regression > Regression;

LOPP
UFRGS

Regression





LOPP
UFRGS

Resultado da Regression “Y1”

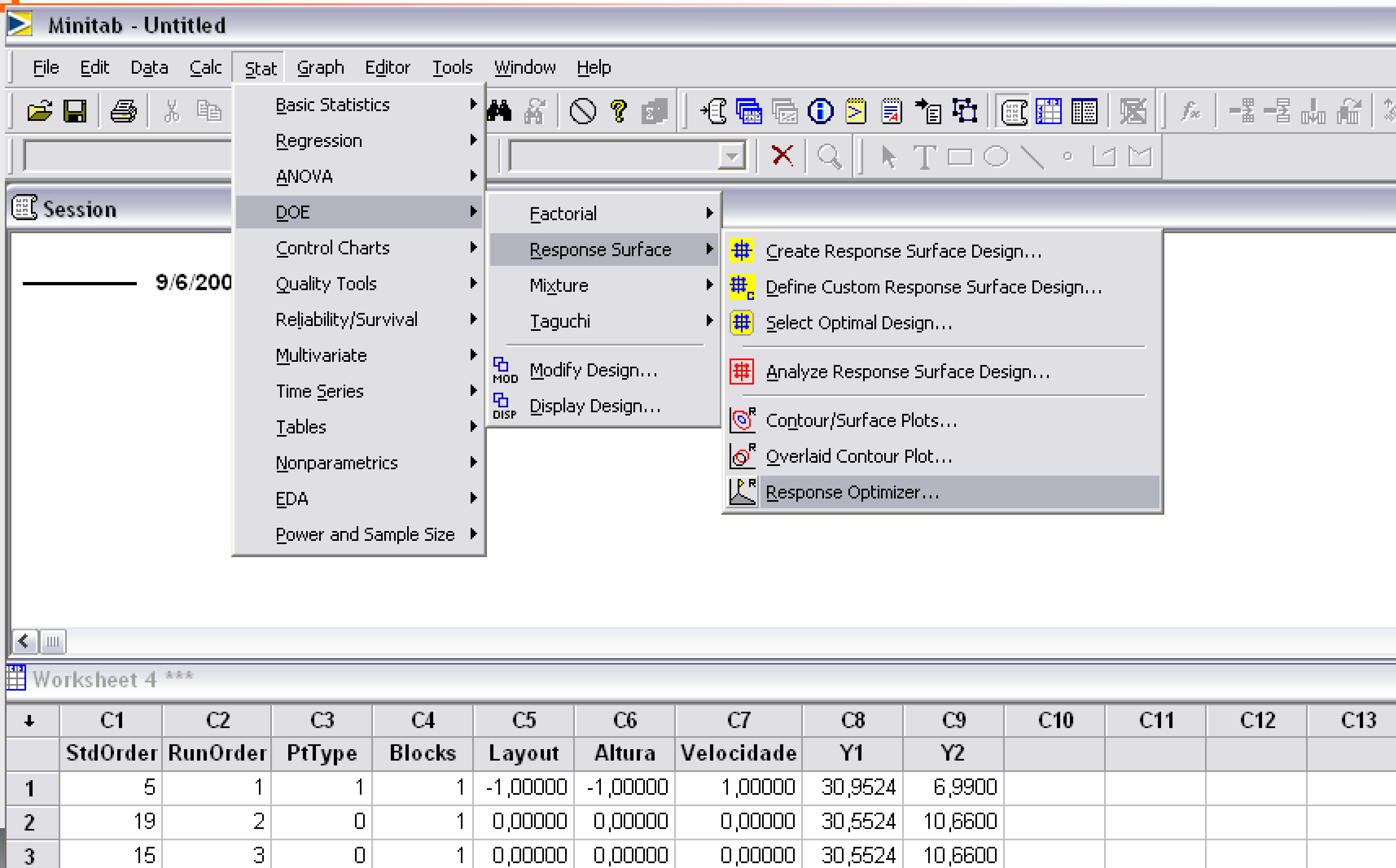
Regression Analysis: Y1 versus Layout; Altura; Velocidade

```
The regression equation is
Y1 = 31,6 + 3,00 Layout - 1,30 Altura - 0,000 Velocidade


Predictor       Coef  SE Coef      T      P
Constant     31,5767   0,3349  94,28  0,000
Layout        3,0000   0,4053   7,40  0,000
Altura       -1,3000   0,4053  -3,21  0,005
Velocidade   -0,0000   0,4053  -0,00  1,000


S = 1,49787   R-Sq = 80,3%   R-Sq(adj) = 76,6%


Analysis of Variance

Source          DF       SS      MS      F      P
Regression       3  145,992  48,664  21,69  0,000
Residual Error  16   35,898   2,244
Total           19  181,889


Source      DF   Seq SS
Layout       1  122,912
Altura       1   23,080
Velocidade   1    0,000


Unusual Observations

Obs  Layout      Y1     Fit  SE Fit  Residual  St Resid
 15    0,00  32,609  29,390   0,760     3,218     2,49R
 20    0,00  36,981  33,763   0,760     3,218     2,49R
```

Response Surface Regression: Y1 versus Layout; Altura; Velo

```
The analysis was done using coded units.

Estimated Regression Coefficients for Y1

Term           Coef  SE Coef       T      P
Constant    31,5767   0,3349  94,278  0,000
Layout       3,0000   0,4053   7,402  0,000
Altura      -1,3000   0,4053  -3,207  0,005
Velocidade  -0,0000   0,4053  -0,000  1,000


S = 1,49787    PRESS = 60,9089
R-Sq = 80,26%  R-Sq(pred) = 66,51%  R-Sq(adj) = 76,56%


Analysis of Variance for Y1

Source          DF   Seq SS   Adj SS   Adj MS      F      P
Regression       3  145,992  145,992  48,6639  21,69  0,000
  Linear         3  145,992  145,992  48,6639  21,69  0,000
Residual Error  16   35,898   35,898   2,2436
  Lack-of-Fit   11   35,898   35,898   3,2634      *      *
  Pure Error     5    0,000    0,000   0,0000
Total           19  181,889


Unusual Observations for Y1

Obs  StdOrder      Y1     Fit  SE Fit  Residual  St Resid
 15        12  32,609  29,390   0,760     3,218    2,49 R
 20        11  36,981  33,763   0,760     3,218    2,49 R

R denotes an observation with a large standardized residual.
```

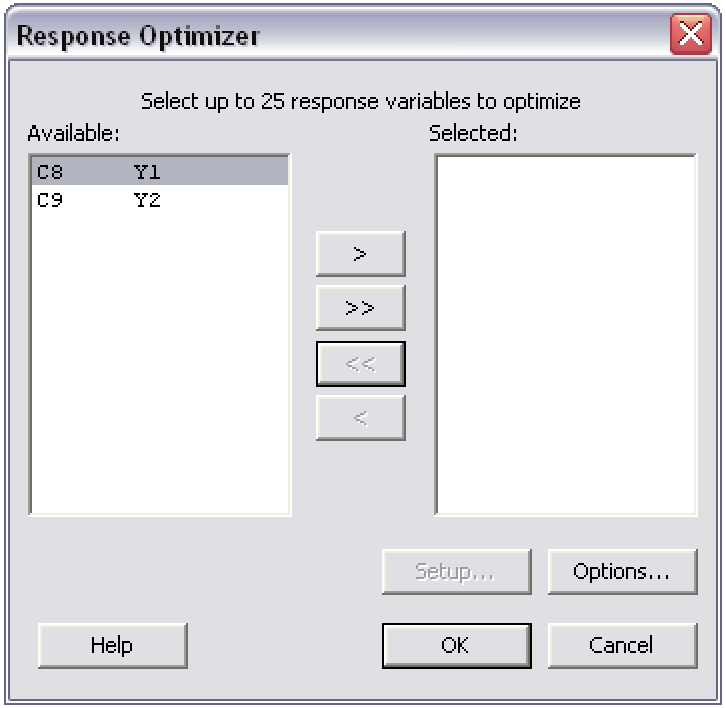
Analisar uma Otimização



- **Caminho: DOE > Response Surface > Resp. Optimizer;**

LOPP
UFRGS

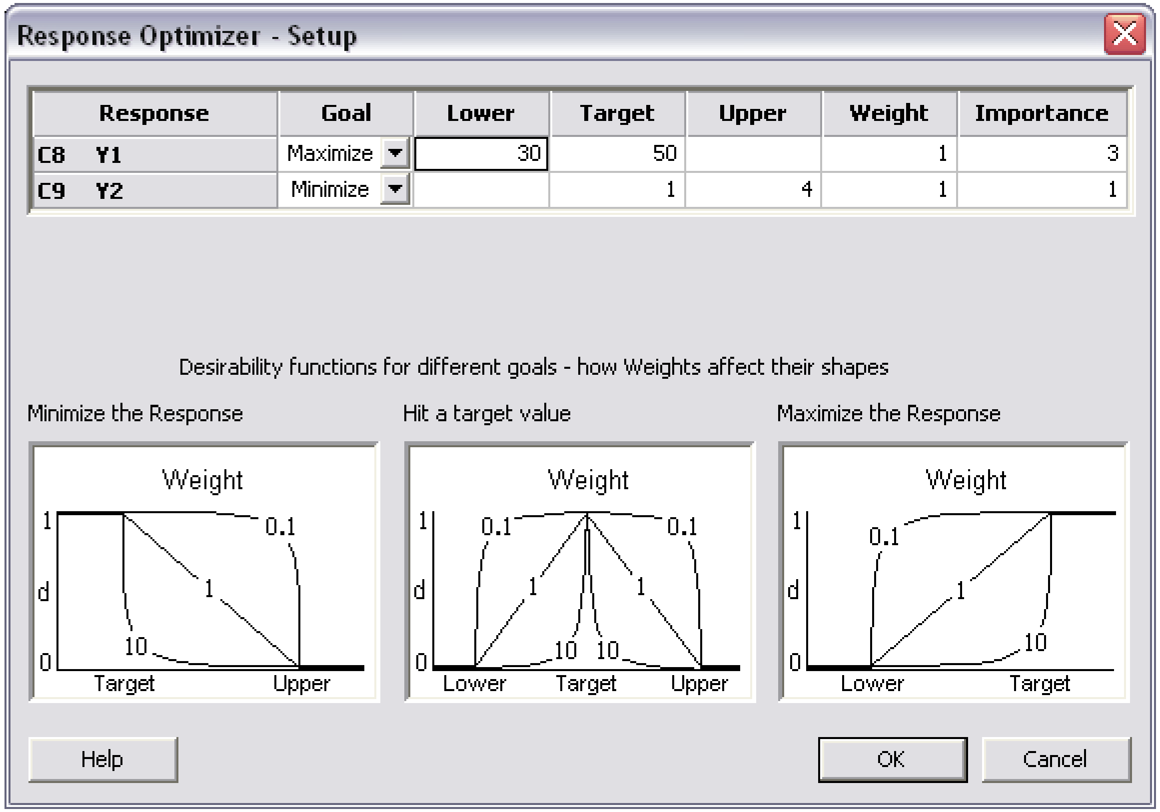
Analisar uma Otimização



LOPP
UFRGS

- Insira as Variáveis para Otimizar;
- Clique em “Setup” para definir o valor Alvo e Limites;
- Clique em “Options”para definir valores iniciais;

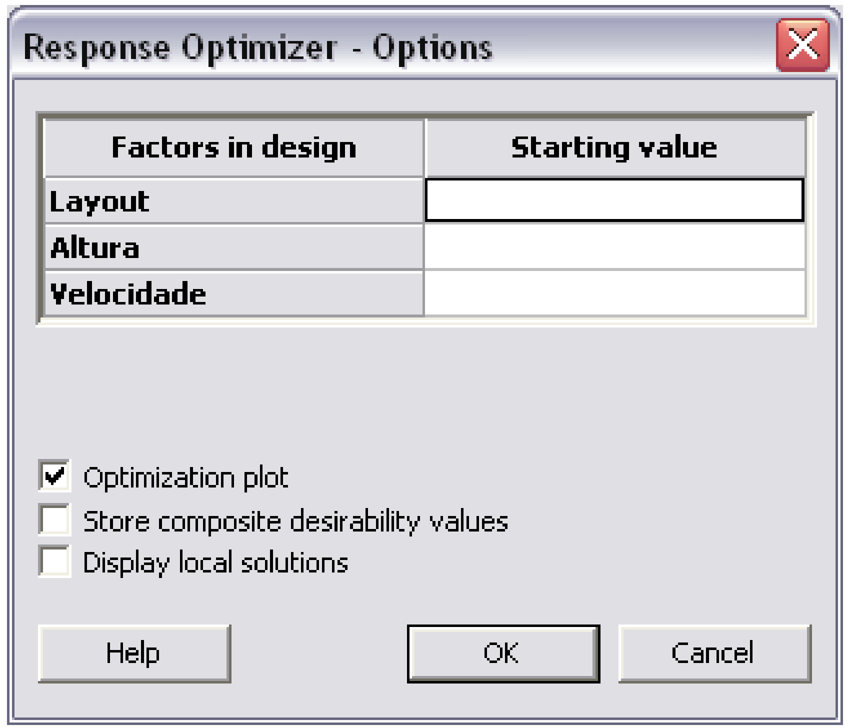
Analisar uma Otimização



LOPP
UFRGS

- **Na opção "Setup";**
- **Inserir as Especificações das Variáveis Respostas;**

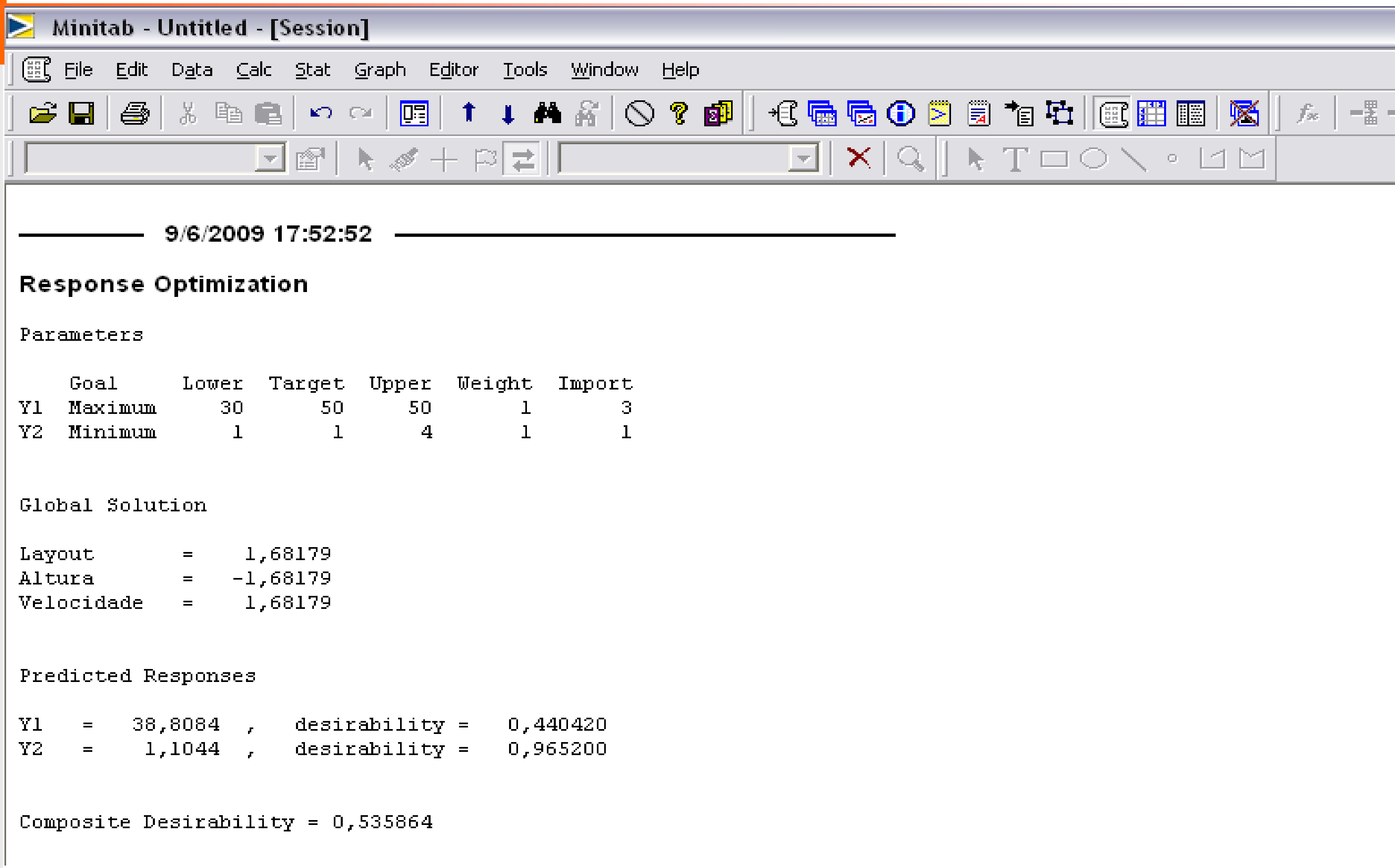
Analisar uma Otimização



LOPP
UFRGS

- **Na opção “Options”;**
- **Inserir valores iniciais para os Fatores Controláveis;**

Analisar uma Otimização



9/6/2009 17:52:52

Response Optimization

Parameters

	Goal	Lower	Target	Upper	Weight	Import
Y1	Maximum	30	50	50	1	3
Y2	Minimum	1	1	4	1	1

Global Solution

```
Layout      =    1,68179
Altura      =   -1,68179
Velocidade  =    1,68179
```

Predicted Responses

```
Y1   =   38,8084  ,   desirability =   0,440420
Y2   =    1,1044  ,   desirability =   0,965200
```

Composite Desirability = 0,535864

- **Inserir valores iniciais para os Fatores Controláveis;**

LOPP
UFRGS

Analisar uma Otimização

Minitab - Untitled - [Session]

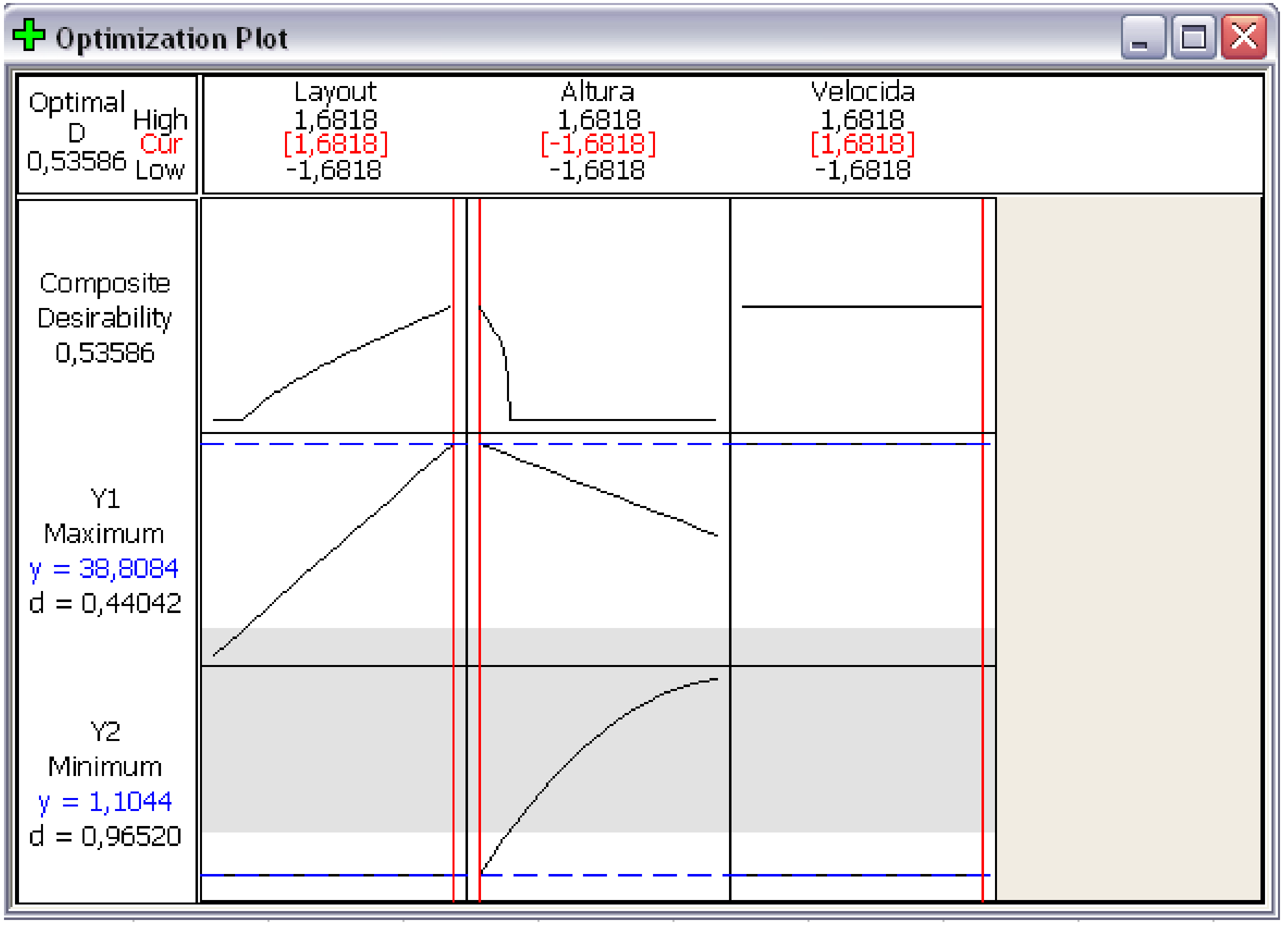
File Edit Data Calc Stat Graph Editor Tools Window Help

Response Optimization

```
Parameters

    Goal      Lower  Target  Upper  Weight  Import
Y1  Maximum      30      50     50       1       3
Y2  Minimum       1       1      4       1       1


Starting Point

Layout      =   -1
Altura      =   -1
Velocidade  =   -1


Global Solution

Layout      =    1,68179
Altura      =   -1,68179
Velocidade  =    1,68179


Predicted Responses

Y1   =    38,8084  ,   desirability =   0,440420
Y2   =     1,1044  ,   desirability =   0,965200


Composite Desirability = 0,535864
```

LOPP
UFRGS

Analisar uma Otimização



LOPP
UFRGS