

# Otimização de produtos e processos

## PARTE I

**Prof. Thais M. F. S. Vieira**  
**[tvieira@usp.br](mailto:tvieira@usp.br)**

# Objetivos

- Capacitar o aluno a
  - Reconhecer a importância do planejamento experimental
  - Planejar, executar e analisar delineamentos fatoriais
  - Ajustar modelo matemático para tomada de decisão
- Metodologia de ensino
  - Aulas expositivas e interativas contemplando:
    - Introdução ao planejamento experimental
    - Apresentação de conceitos relacionados ao tema
    - Estudos de casos
- Bibliografia básica
  - RODRIGUES, M. I.; IEMMA, A. F. **Planejamento de experimentos e otimização de processos**. Campinas : Casa do Espírito Amigo Fraternidade Fé e Amor, 2009. 357 p.
  - MYERS, R. H.; MONTGOMERY, D. C. ; ANDERSON-COOK, C. M. **Response surface methodology: process and product optimization using designed experiments**. Nova Iorque: John Wiley & Sons, 2009. 1200p

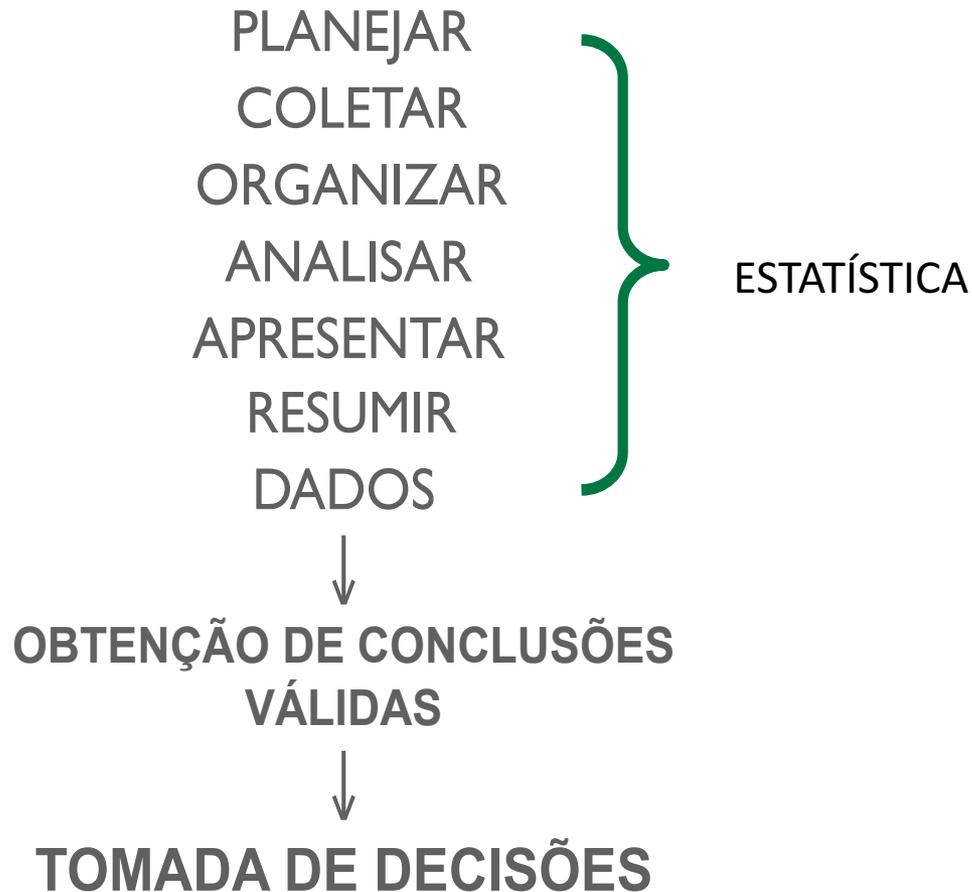
# Conteúdo

- Planejamento Experimental - Introdução
- Esquema fatorial
- Efeitos principais e interações
- ANOVA; ajuste de modelos lineares e quadráticos
- Estudos de casos

## Dicas importantes:

- A consulta ao texto será importante para que os conceitos sejam consolidados;
- A realização do trabalho em grupo também;
- Não é necessário ser especialista em estatística para começar a pensar em experimentos planejados!

# Introdução



O QUE EU GOSTARIA DE FICAR SABENDO QUANDO TERMINAR O ESTUDO/EXPERIMENTO?

# Introdução

## Estatística descritiva ou dedutiva:

- Descrever e analisar um certo grupo de dados, sem tirar conclusões ou inferências sobre um grupo maior.

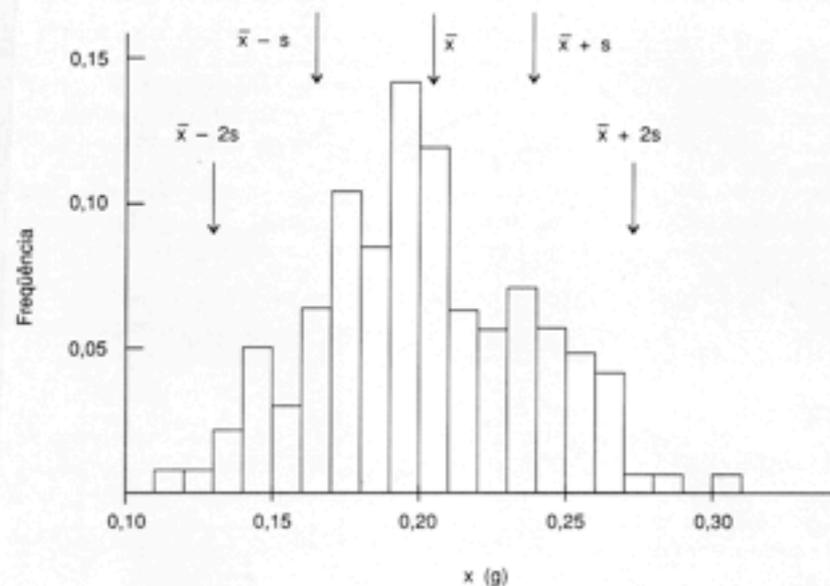
## Estatística indutiva ou inferência estatística:

- A partir de sub-conjuntos representativos (amostras) fazem-se induções sobre o conjunto estudado (população ou universo).
- Como essa inferência pode não ser absolutamente certa, a linguagem da probabilidade é sempre usada no estabelecimento das conclusões.

## Modelos Estatísticos



## Distribuição de Frequência



**Figura 2.2** Distribuição de frequências dos pesos de 140 caroços extraídos aleatoriamente de um pacote de 1 kg de feijão preto.

Probabilidade de ocorrência de erros (ou desvios) em medições

# Introdução

- ***Para  $N > 30$  (grandes amostras):***
  - as distribuições amostrais são aproximadamente normais,
  - quanto  $> N$ , melhor é esta aproximação
- ***Para  $N < 30$  (pequenas amostras):***
  - A aproximação de distribuição normal não é boa e torna-se pior com o decréscimo de  $N$ 
    - distribuição t (Student)
    - distribuição F
    - distribuição  $X^2$  (qui-quadrado)

Testes usados para determinar quão aproximadamente as distribuições empíricas (obtidas dos dados amostrais) se ajustam às distribuições teóricas

# Introdução

- Teste de Hipóteses:
  - $H_0$ : nulidade ( $\mu = \mu_0$ );
  - $H_1$ : hipótese de efeito observável  $\left\{ \begin{array}{l} \text{unilateral } (\mu < \mu_0 \text{ ou } \mu > \mu_0); \\ \text{bilateral } (\mu \neq \mu_0); \end{array} \right.$
- Dados baseados em um número de observações experimentais são também acompanhados de incertezas
- Probabilidade de rejeitar uma das hipóteses e ela ser verdadeira.
  - Aos testes de hipótese estão associados dois tipos de erros:
    - Erro tipo I e Erro tipo II

# Introdução

**Erro tipo I:** probabilidade de rejeitar  $H_0$  (concluir que amostras diferem) e  $H_0$  ser verdadeira.

**Erro tipo II:** probabilidade de não rejeitar  $H_0$  (concluir que amostras não diferem) e  $H_0$  ser falsa.

**Na ANOVA e teste F controlamos o erro tipo I:**

- Se  $F_{\text{calculado}} > F_{\text{tabelado}}$ , rejeito  $H_0$ : pelo menos uma amostra difere das demais
- **Ao nível de significância de 10%, 5%, 1%... de acordo com os prejuízos que a conclusão pode causar se ela estiver errada.**

# Planejamento Experimental

- Desenvolvimento de produtos ou processos em geral é dependente de um grande número de variáveis
- Objetivo:
  - “melhor produto”; “melhor condição de processo”
- Como definir?
  - Resposta – variáveis a serem observadas
- Quais são as variáveis que podem afetar essas respostas?
  - Variáveis explanatórias ou explicativas

# Planejamento Experimental

- Variáveis Explanatórias:
  - Fatores a serem estudados ou avaliados em um processo **(que podem ser controlados)**
  - Ex.: Condições de processo; concentrações de ingredientes e insumos
- Variáveis Dependentes:
  - Respostas desejadas (determinadas experimentalmente)
  - Ex.: Rendimento, produtividade, atributos sensoriais...

# Planejamento Experimental

- Estratégia de planejamento:
  - Usar todas as variáveis?
  - Algumas? Como selecioná-las?

Uma estratégia é selecionar o conjunto de variáveis mais importantes:  
*Screening*

*Ex: Selecionamos 2 variáveis*

Estabelecer a estratégia para encontrar o melhor produto/  
processo, com apenas 10 amostras

# Planejamento Fatorial $2^k$

“k” fatores, analisados em 2 níveis:

- Quando há vários fatores de interesse
- Os níveis dos fatores são binários
  - denotados por “inferior” e “superior”
- É realizado um experimento fatorial completamente casualizado
  - todas as combinações dos k fatores são realizadas  $2^k$  combinações

# Planejamento Fatorial $2^k$

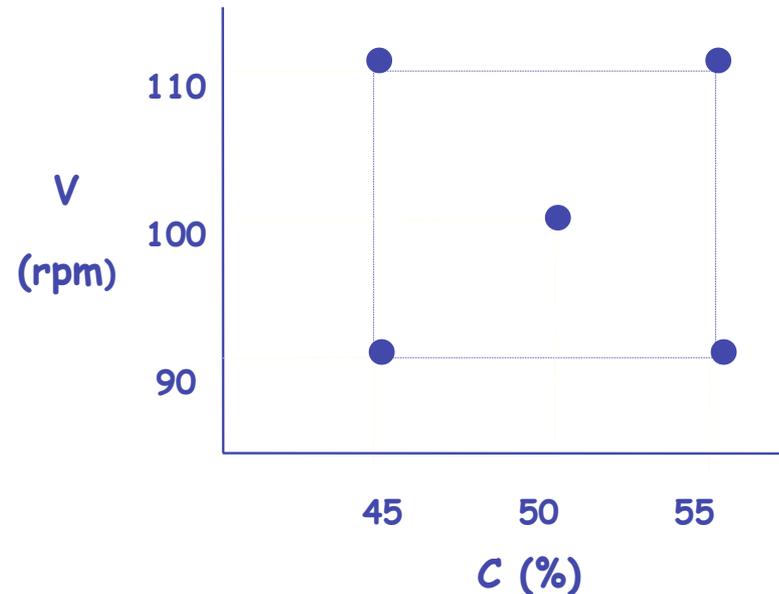
Matriz do planejamento:

- todas as possíveis combinações dos níveis inferior e superior

**Exemplo:** Maximizar o rendimento ( $y$ ) de um processo

Variáveis Explanatórias:  $x_1$  Concentração de um ingrediente (%)

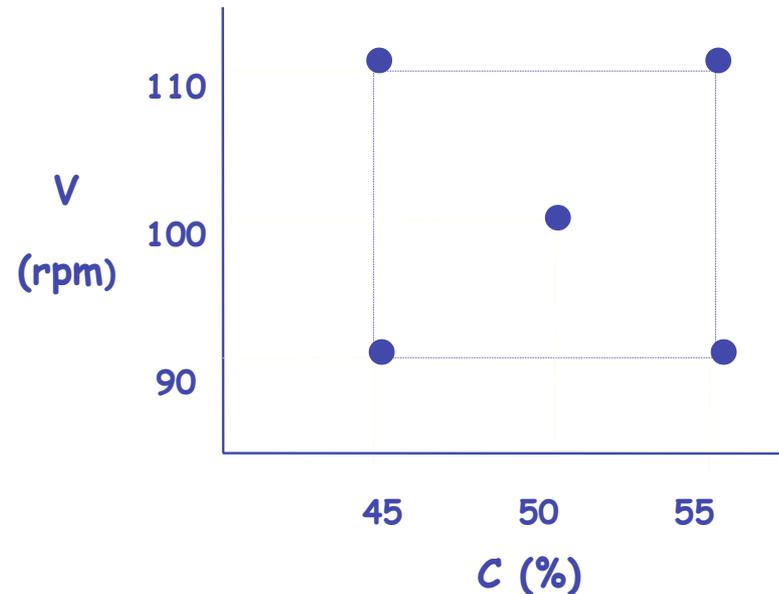
$x_2$  Velocidade de agitação (rpm)



# Planejamento Fatorial $2^k$

Faixas de estudo das variáveis explanatórias em análise no processo

Variável Explanatória	-1	0	+1
$X_1$	45	50	55
$X_2$	90	100	110



# Planejamento Fatorial $2^k$

Ensaio	Valor codificado		Valor real		Y (%)
	X <sub>1</sub>	X <sub>2</sub>	C (%)	V (rpm)	
1	- 1	- 1	45	90	69
2	+ 1	- 1	55	90	59
3	- 1	+ 1	45	110	78
4	+ 1	+ 1	55	110	67
5	0	0	50	100	68
6	0	0	50	100	66
7	0	0	50	100	69

-10

9

-11

8

Interpretação geométrica dos efeitos: Contraste entre as arestas opostas

# Efeitos principais

Ensaio	Valor codificado		Valor real		Y (%)
	X <sub>1</sub>	X <sub>2</sub>	C (%)	V (rpm)	
1	- 1	- 1	45	90	69
2	+ 1	- 1	55	90	59
3	- 1	+ 1	45	110	78
4	+ 1	+ 1	55	110	67
5	0	0	50	100	68
6	0	0	50	100	66
7	0	0	50	100	69

Efeito da Concentração:

$$C = [(y_2 - y_1) + (y_4 - y_3)] / 2$$

$$C = [(59 - 69) + (67 - 78)] / 2$$

$$C = [(-10) + (-11)] / 2$$

$$C = -10,5 \%$$



Mesma unidade da resposta

# Efeito da Interação

Ensaio	Valor codificado		Y (%)	efeitos		
	X <sub>1</sub>	X <sub>2</sub>		X <sub>1</sub>	X <sub>2</sub>	X <sub>1</sub> .X <sub>2</sub>
1	-1	-1	69	- 69	- 69	+ 69
2	1	-1	59	+59	-59	- 59
3	-1	1	78	-78	+78	-78
4	1	1	67	+67	+67	+67
				-10,5	8,5	-0,5

Se não houver interação: Efeito da Concentração seria o mesmo com qualquer velocidade

Efeito da interação = metade da diferença entre os efeitos

$$C \times V = CV = (10 - 11)/2 = - 0,5\%$$

# Significância dos efeitos

Estimativa da variância do erro experimental



Cálculo do erro padrão das respostas médias para cada tratamento



Cálculo do erro padrão dos efeitos



**Teste t de significância para os efeitos:**

$$t_{ef} = \frac{ef}{S_{ef}}$$

O efeito é significativo se  $|t_{ef}| > t_{tabelado}$

# Estimativa do erro experimental

- Ensaaios devem ser realizados de forma aleatória
- Realização de ensaios em replicata:
  - estimativa conjunta da variância de uma observação individual
  - Erro padrão
- Repetições autênticas do ponto central

Cálculo do t:

$$t = \text{Valor do Efeito} / \text{Erro Padrão}$$

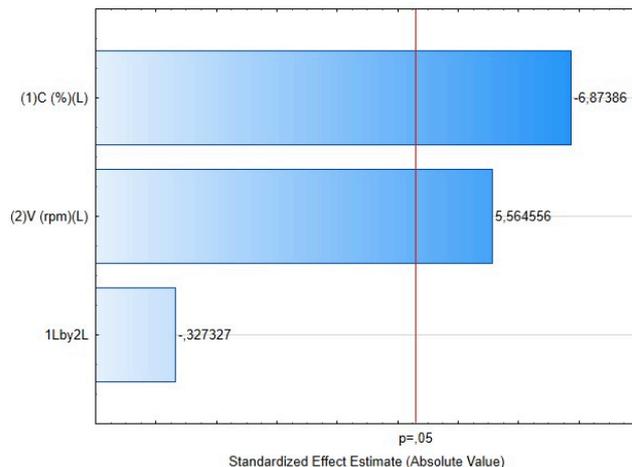
Cálculo do p-value:

- Probabilidade de significância baseado no t calculado.
- Se  $p > 0,05$  tem-se a probabilidade do efeito ser zero, ou seja Hipótese nula ( $H_0$ ) ser verdadeira

# Análise dos efeitos

- Estimativa dos efeitos:

	Efeito	Erro padrão	t calc	p
Média	68,00	0,57735	117,7795	0,000072
C (%)	-10,50	1,5275	-6,8739	0,0200515
V (rpm)	8,50	1,5275	5,5646	0,030811
1 x 2	-0,50	1,5275	0,3273	0,774506



# Ajuste do modelo

Variável Explanatória	-1	0	+1
$X_1$	45	50	55
$X_2$	90	100	110

Modelo de 1ª Ordem = modelo linear

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3$$

Onde:

$Y$  = resposta (variável dependente)

$x_1, x_2, x_3$  = níveis codificados das variáveis explanatórias

$\beta_0$  = média

$\beta$ 's = coeficientes estimados pelo método dos mínimos quadrados

# ANOVA

- **Modelo codificado:**

$$\text{Rendimento (\%)} = 68,0 - 5,25.C + 4,25.v$$

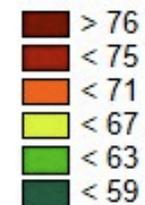
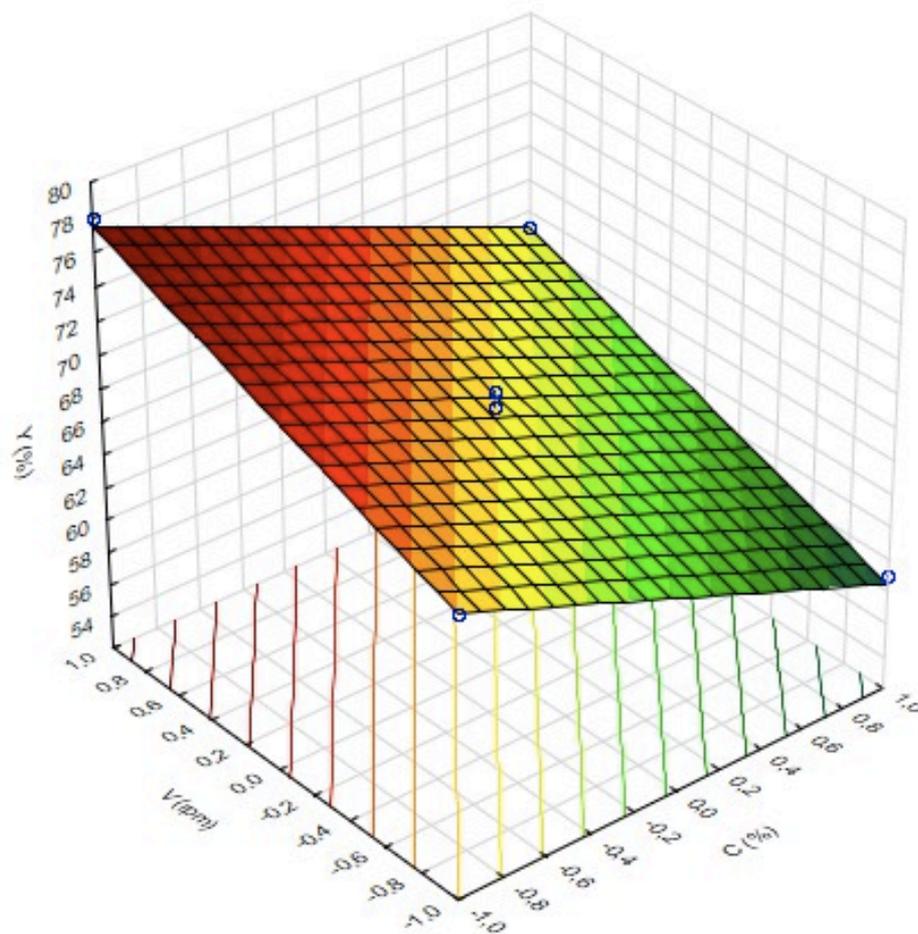
Fonte de Variação	Soma dos Quadrados	GL	Quadrado Médio	F calculado
Regressão	182,50	2	91,25	66,12
Resíduos	5,50	4	1,38	
Falta de ajuste	0,83	2	0,42	0,17
Erro puro	4,57	2	2,34	
Total	188,00	6		

$$R^2 = 0,97 \text{ (\% de variação explicada)}$$

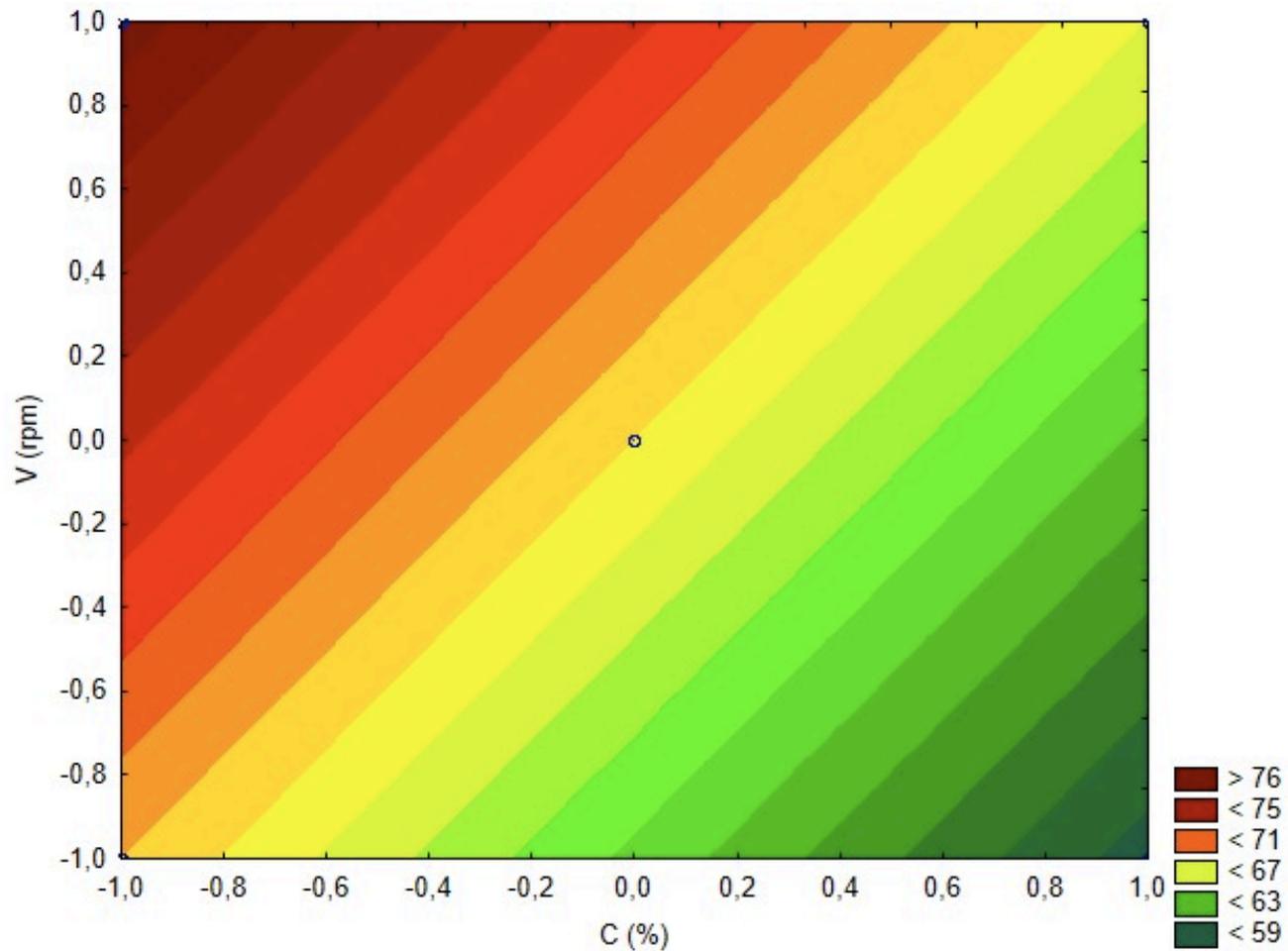
$$F_{0,95; 2; 4} = 6,94$$

$$F_{0,95; 2; 2} = 19,00$$

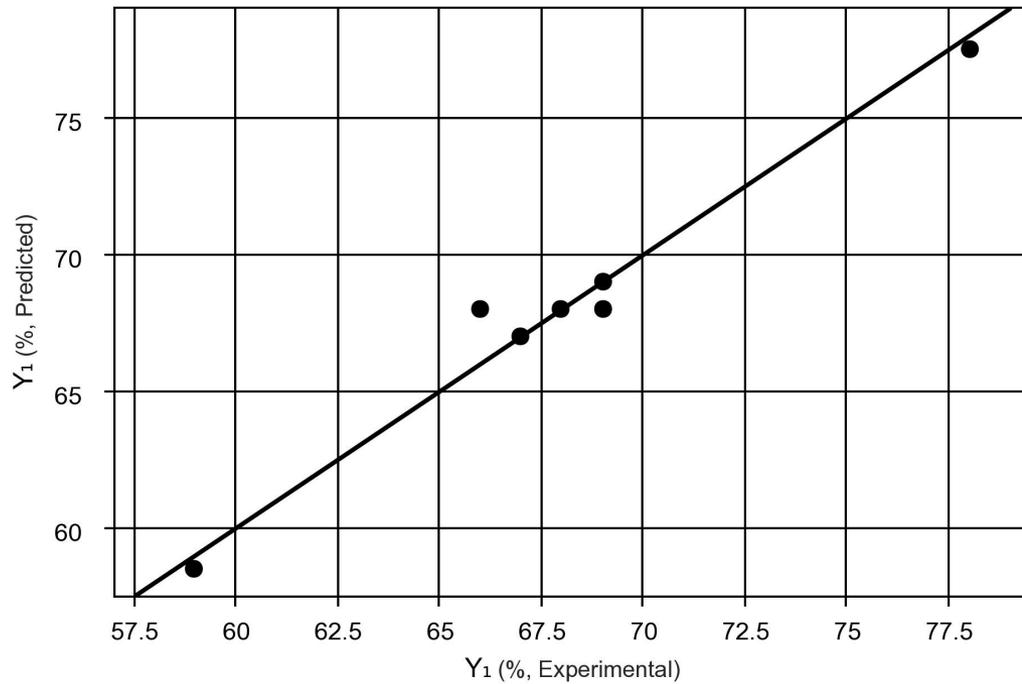
# SUPERFÍCIE DE RESPOSTA



# GRÁFICO DE CONTORNO

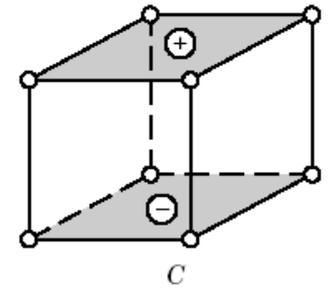
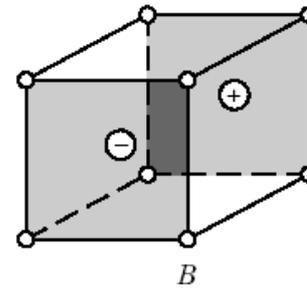
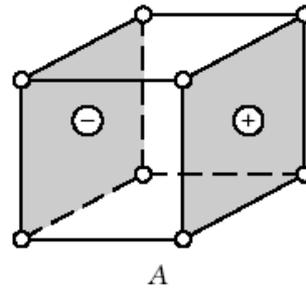
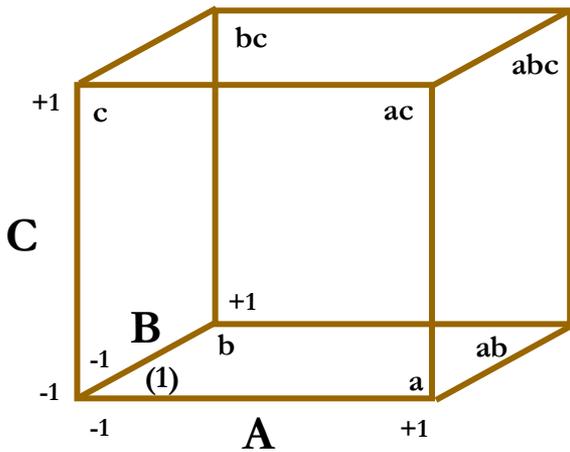


# Falta de ajuste

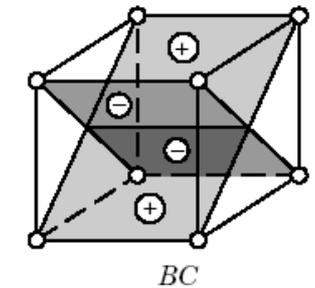
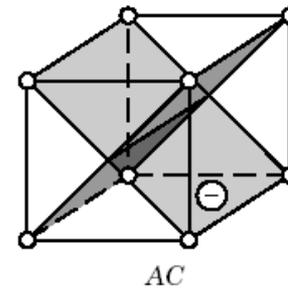
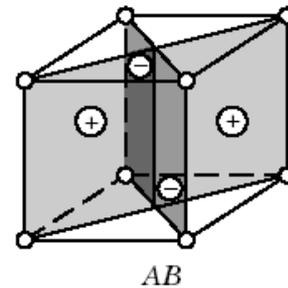


# Planejamento Fatorial $2^k$

Para 3 variáveis:  $2^3 = 8$  ensaios (níveis inferiores e superiores)



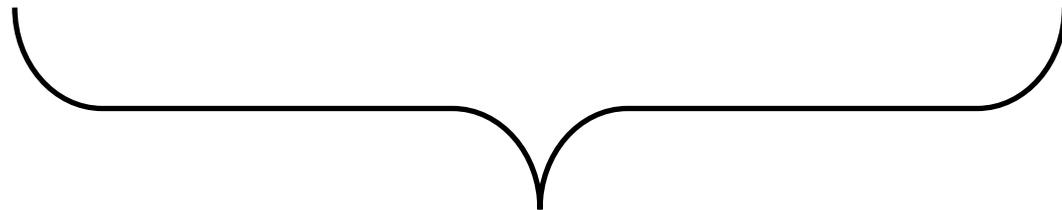
(a) Main effects



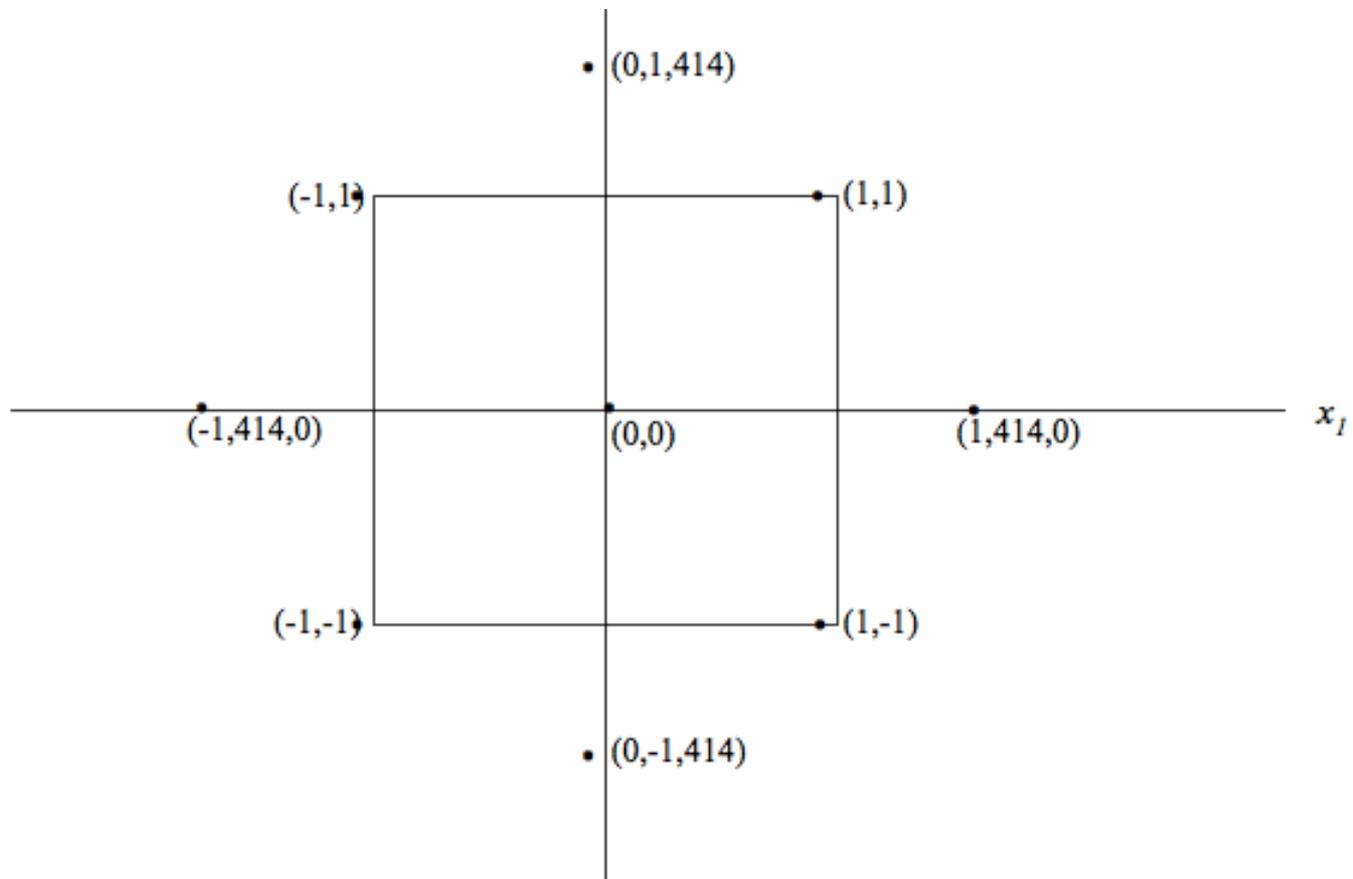
(b) Two-factor interaction

# FATORIAL COMPLETO COM PONTO CENTRAL

Variável	Nível de variação				
exploratória	$-\alpha$	<b>-1</b>	<b>0</b>	<b>+1</b>	$+\alpha$
$X_1$		30	35	40	
$X_2$		115	125	135	



Modelo de 2ª Ordem = modelo quadrático



# Modelo de 2ª Ordem

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{33} x_3^2 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3$$

Onde:

Y = resposta (variável dependente)

$x_1, x_2, x_3$  = níveis codificados das variáveis explanatórias

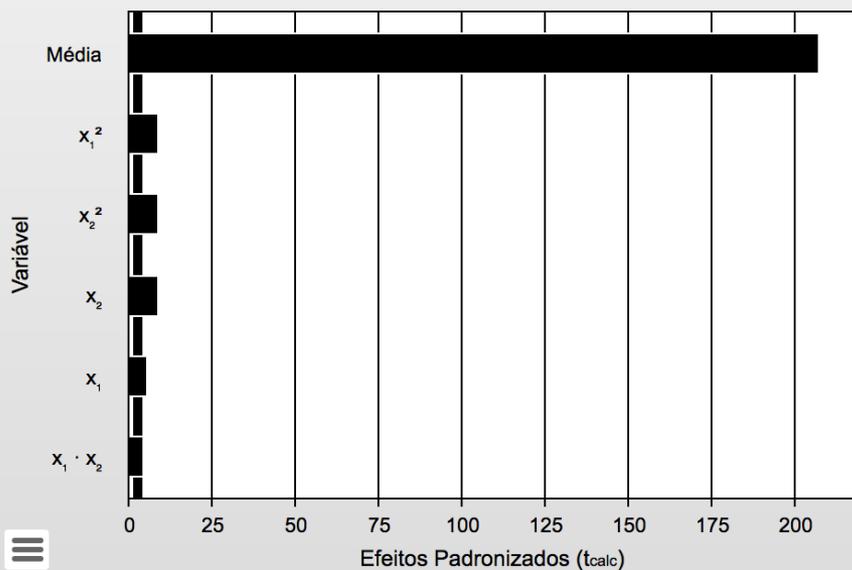
$\beta_0$  = média

$\beta$ 's = coeficientes estimados pelo método dos mínimos quadrados

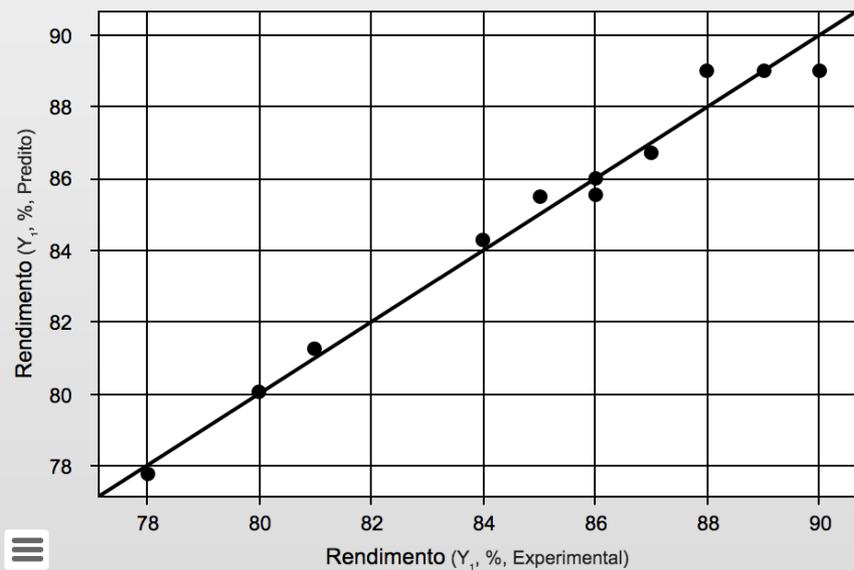
## Resultados do Planejamento com adiç3o dos pontos axiais

Ensaio	X <sub>1</sub>	X <sub>2</sub>	C (%)	V (rpm)	Y (%)
1	- 1	- 1	30	115	86
2	+ 1	- 1	40	115	85
3	- 1	+ 1	30	135	78
4	+ 1	+ 1	40	135	84
5	0	0	35	125	90
6	0	0	35	125	88
7	0	0	35	125	89
8	-1,41	0	28	125	81
9	0	+1,41	35	139	80
10	+ 1,41	0	42	125	86
11	0	-1,41	35	119	87

## Gráfico de Pareto



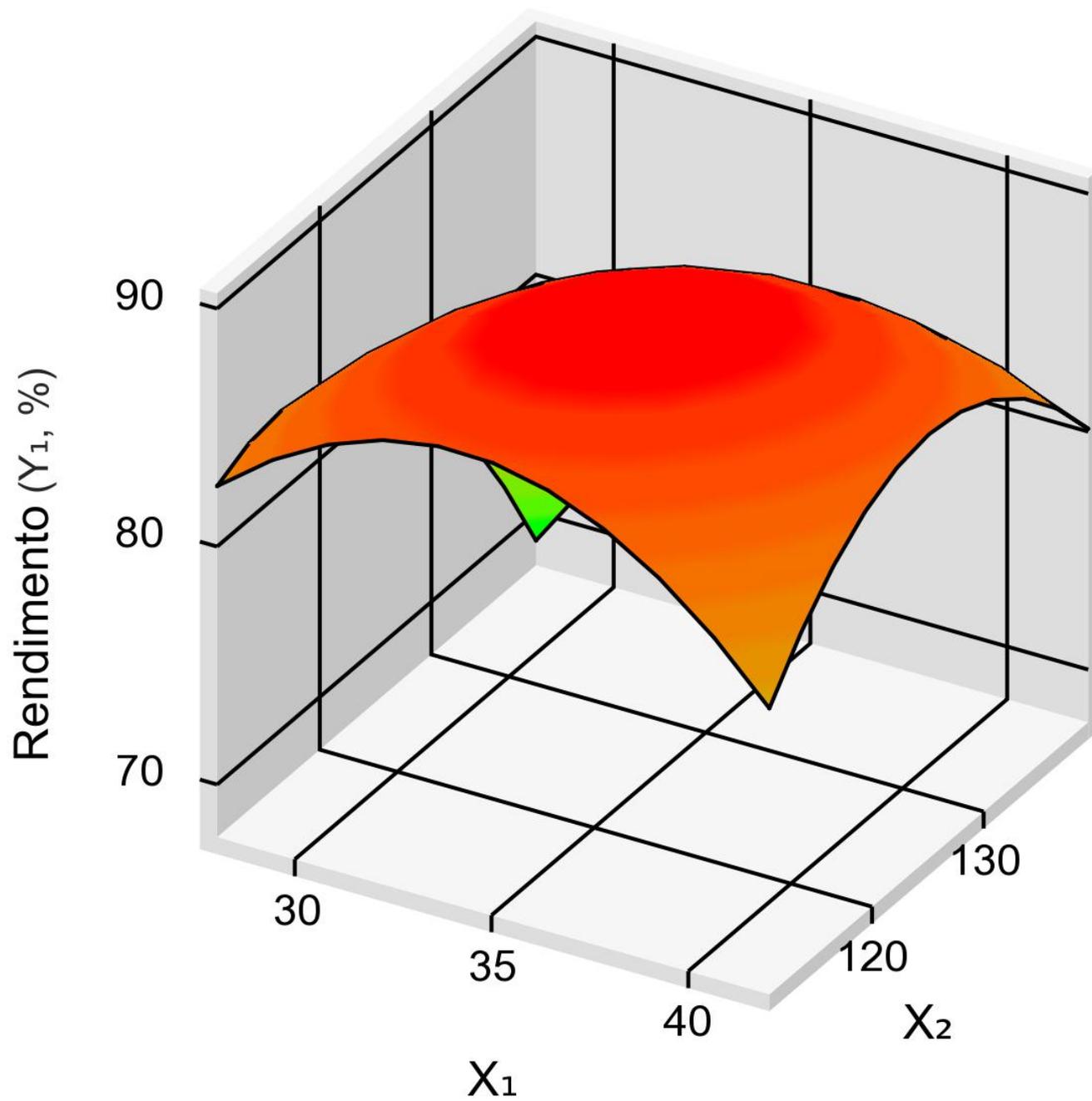
## Valores Experimentais × Preditos

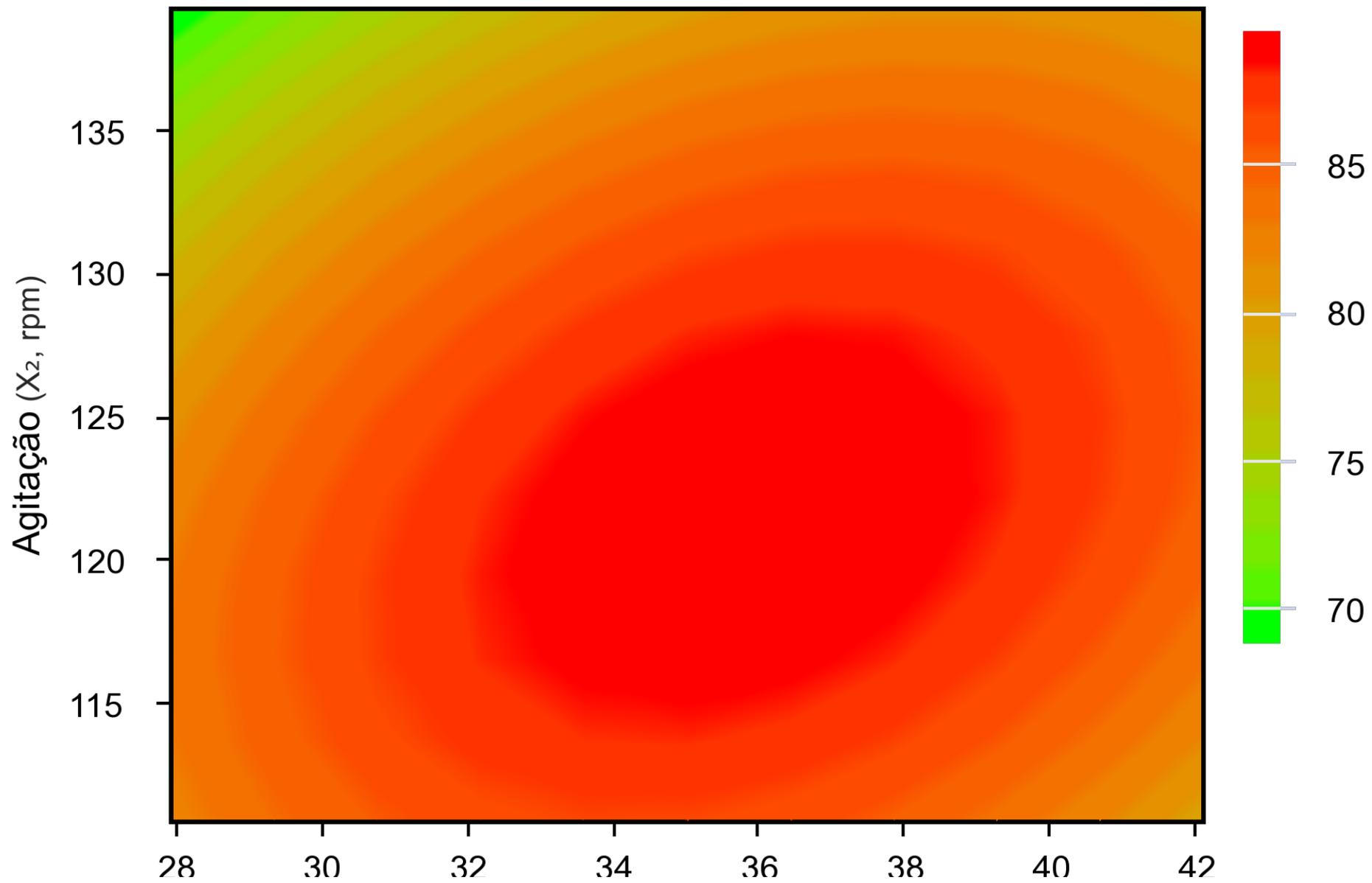


## ANOVA

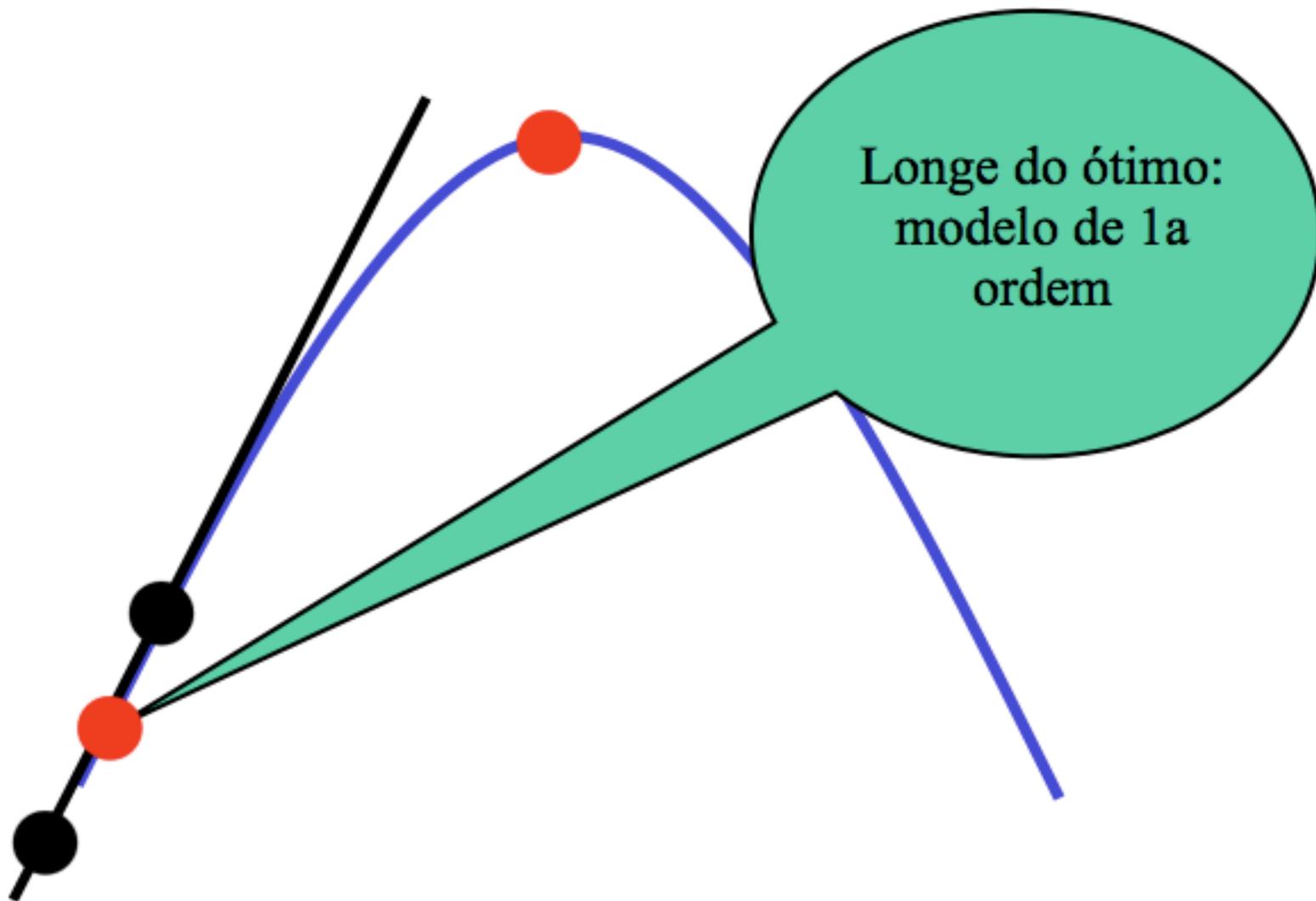
Fonte de Variação	Soma dos Quadrados	Graus de Liberdade	Quadrado Médio	F <sub>calc</sub>	p-valor
Regressão	144,1	5	28,8	52,2	0,00026
Resíduos	2,8	5	0,6		
Falta de Ajuste	0,8	3	0,3	0,3	0,85503
Erro Puro	2,0	2	1,0		
Total	146,9	10			

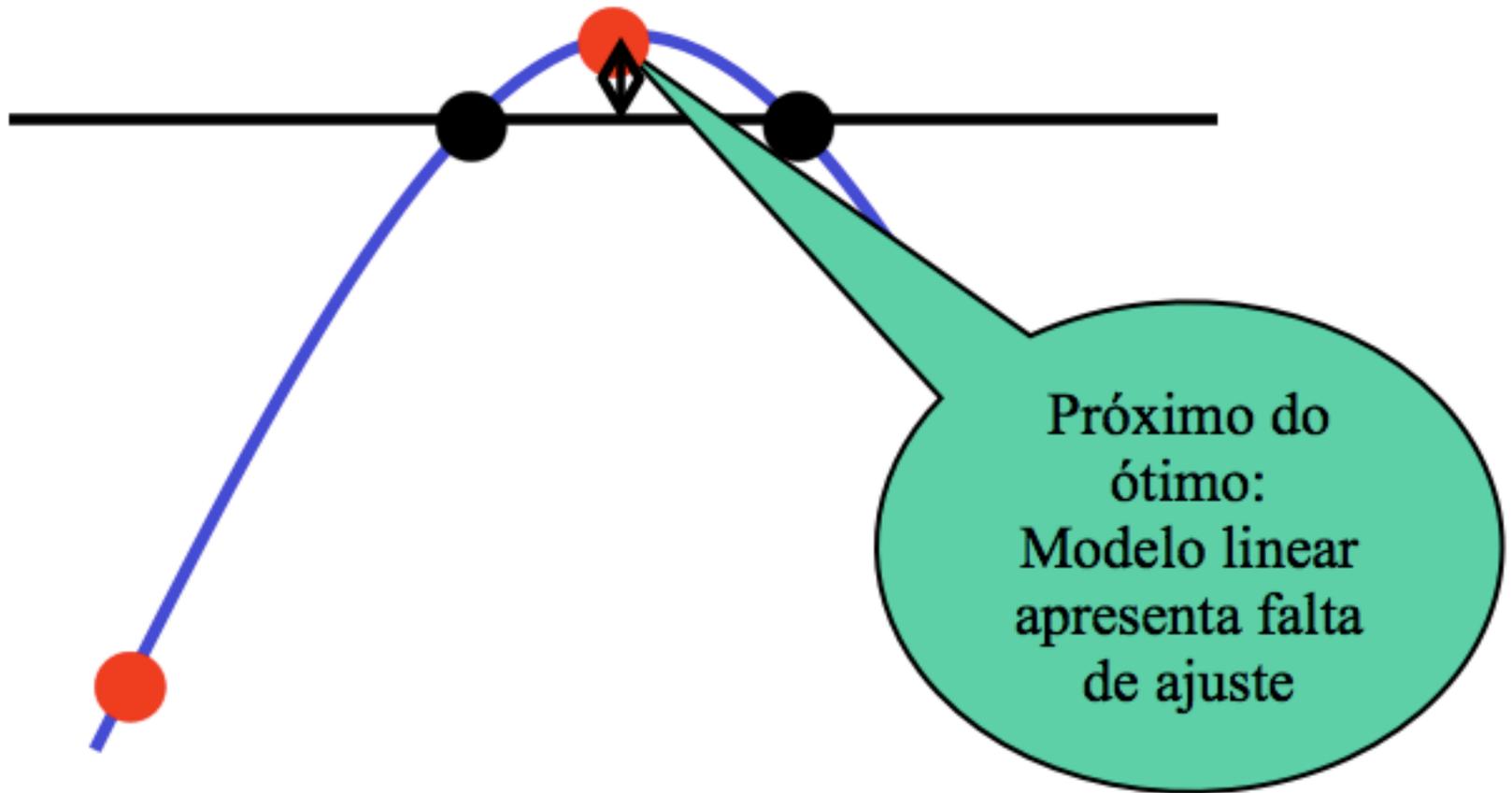
$R^2 = 98,12\%$





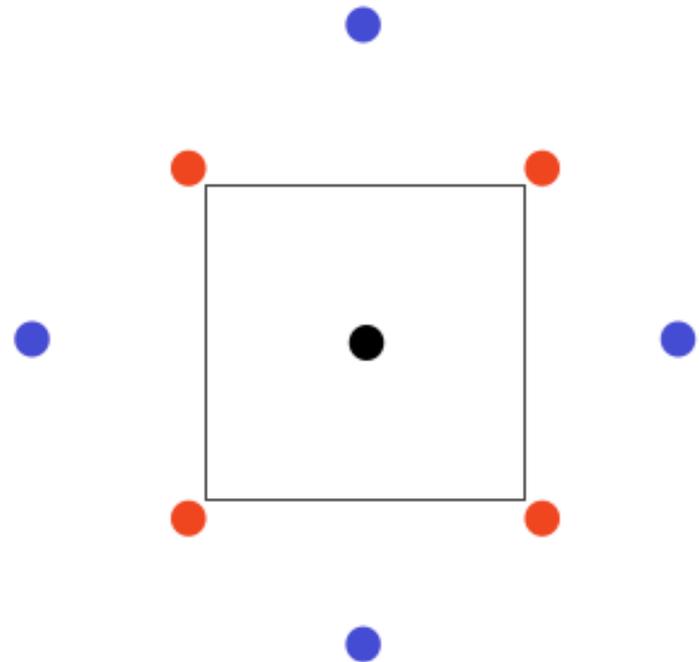
# Modelos

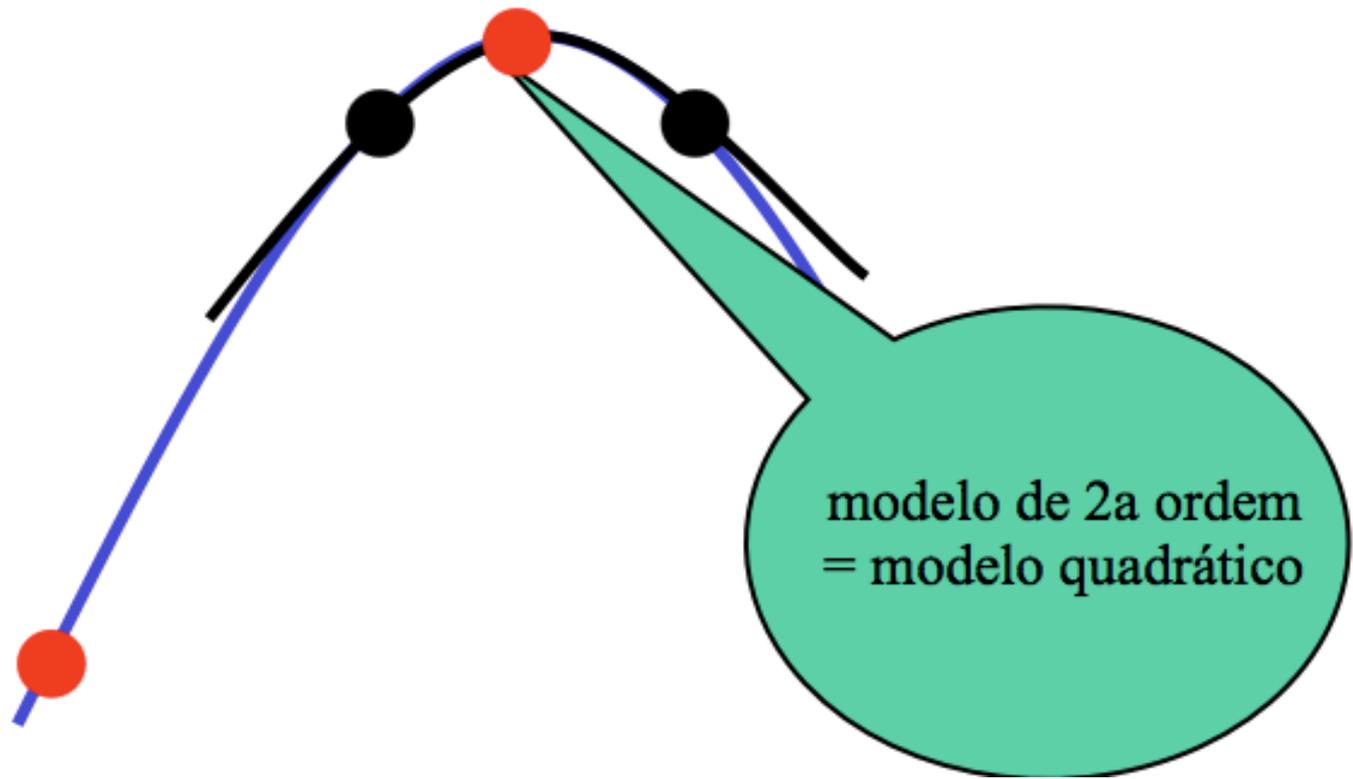




# Central Composite Design

- 3 partes:
  - Pontos fatoriais
  - Pontos centrais
  - Pontos axiais





# Estratégia de otimização

1. Screening

2. Modelos de 1a ordem

3. Modelos de 2a ordem:

- Central Composite Design (CCD)
  - precisão para estimar os efeitos
- [Box-Behnken Design](#)
  - menos preciso na estimativa dos efeitos

***Pesquisador***

Otimizar o rendimento de um processo



**Temperatura e Pressão**

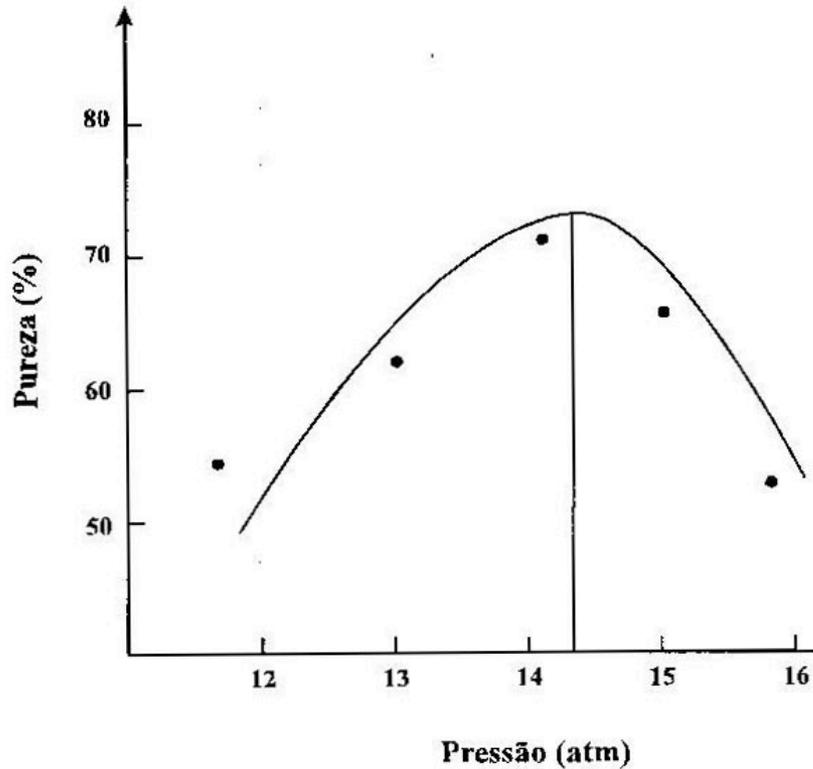
**Quais os valores das 2 variáveis que produzem o maior rendimento?**



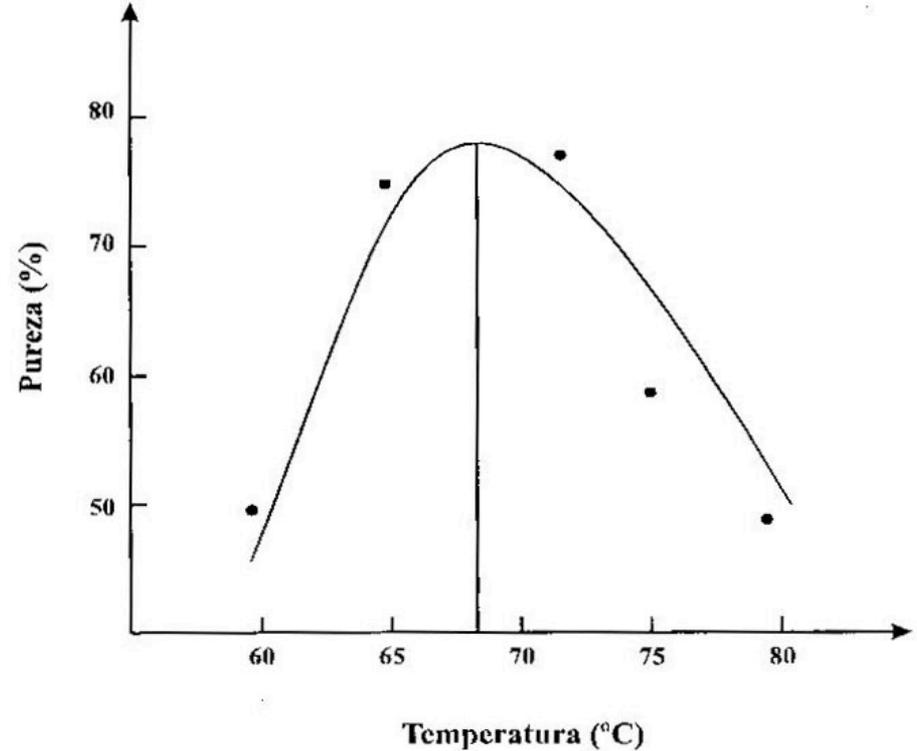
**Sugestão: manter o processo sob controle**

***Experimentos do tipo Um-Fator-de-Cada-Vez***

(1) Fixar T e variar P



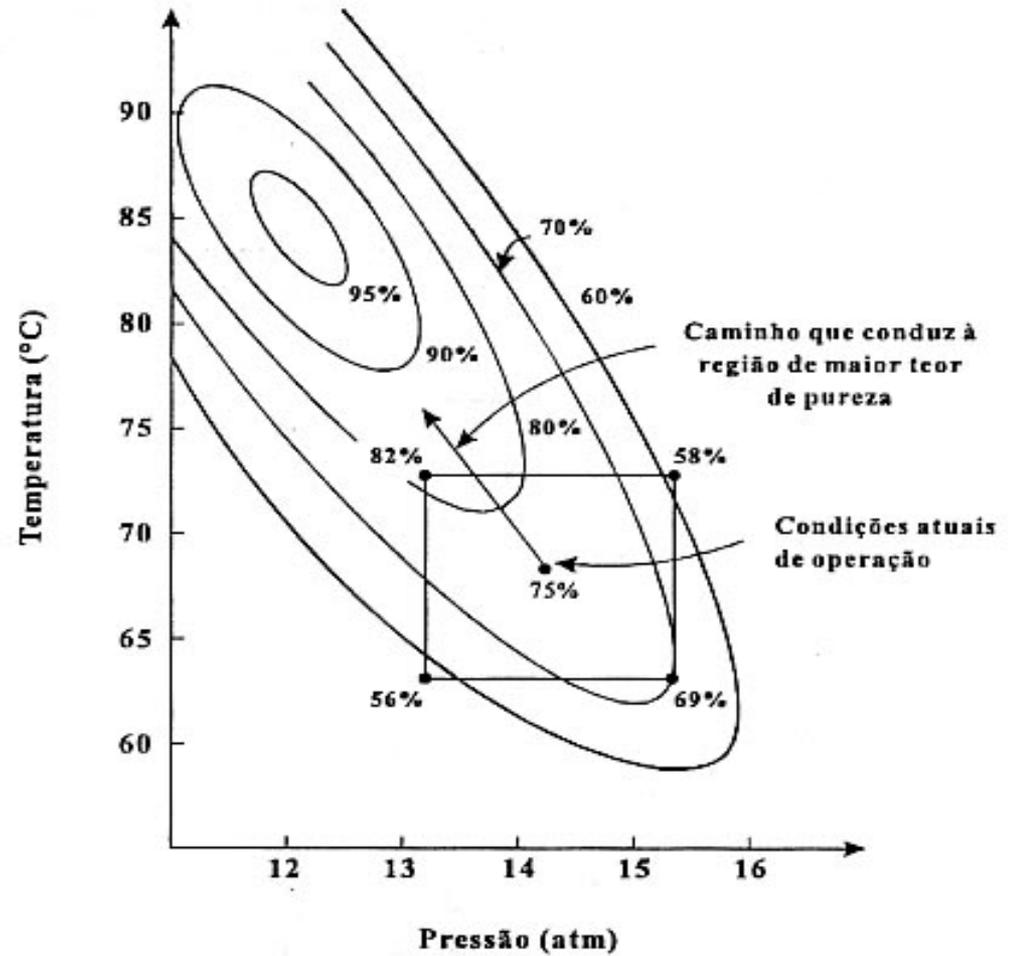
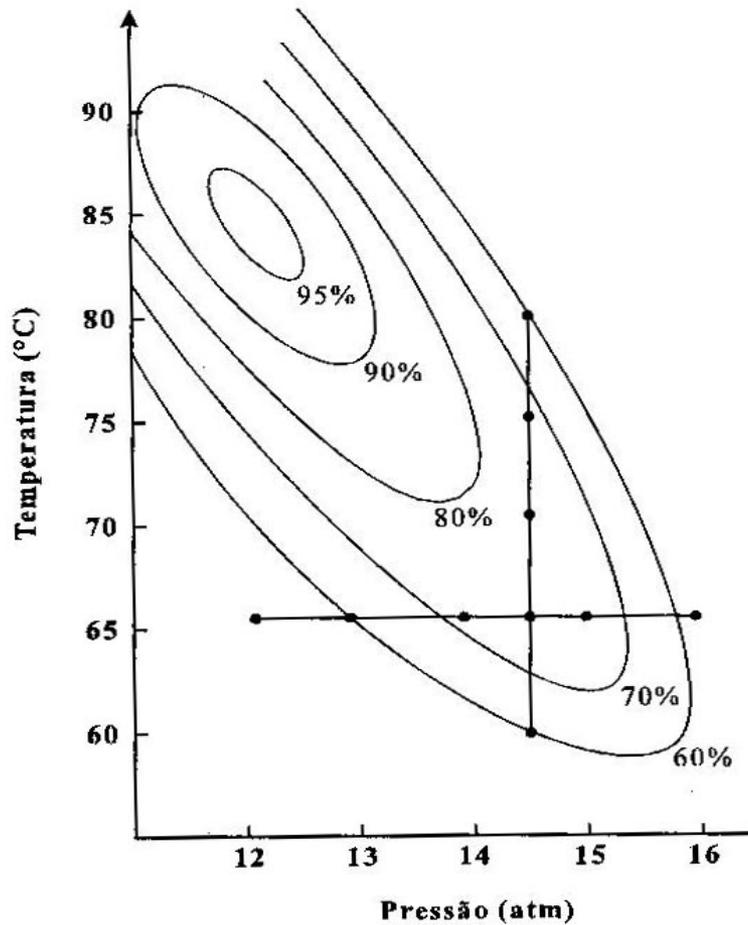
(2) Manter P no valor ótimo e variar T até > Pureza



***Experimento acabou e descobrimos os valores ótimos para maximização da Pureza.***

***Certo ou Errado?***

## Experimentos Fatoriais x “um-fator-de-cada-vez”



# Vantagens do Planejamento Experimental

- Reduz o número de experiências, com **melhor qualidade** de informação nos resultados
- Análise multivariável: permite verificar e quantificar efeitos sinérgicos e antagônicos entre as variáveis explanatórias estudadas
- É possível otimizar mais de uma resposta ao mesmo tempo

Exemplo do uso da Metodologia Univariável  
(Um-Fator-Por-Vez)  
versus

Planejamento Experimental Fatorial

Estudo de Caso:

**Avaliação do pH e temperatura na atividade da inulinase**

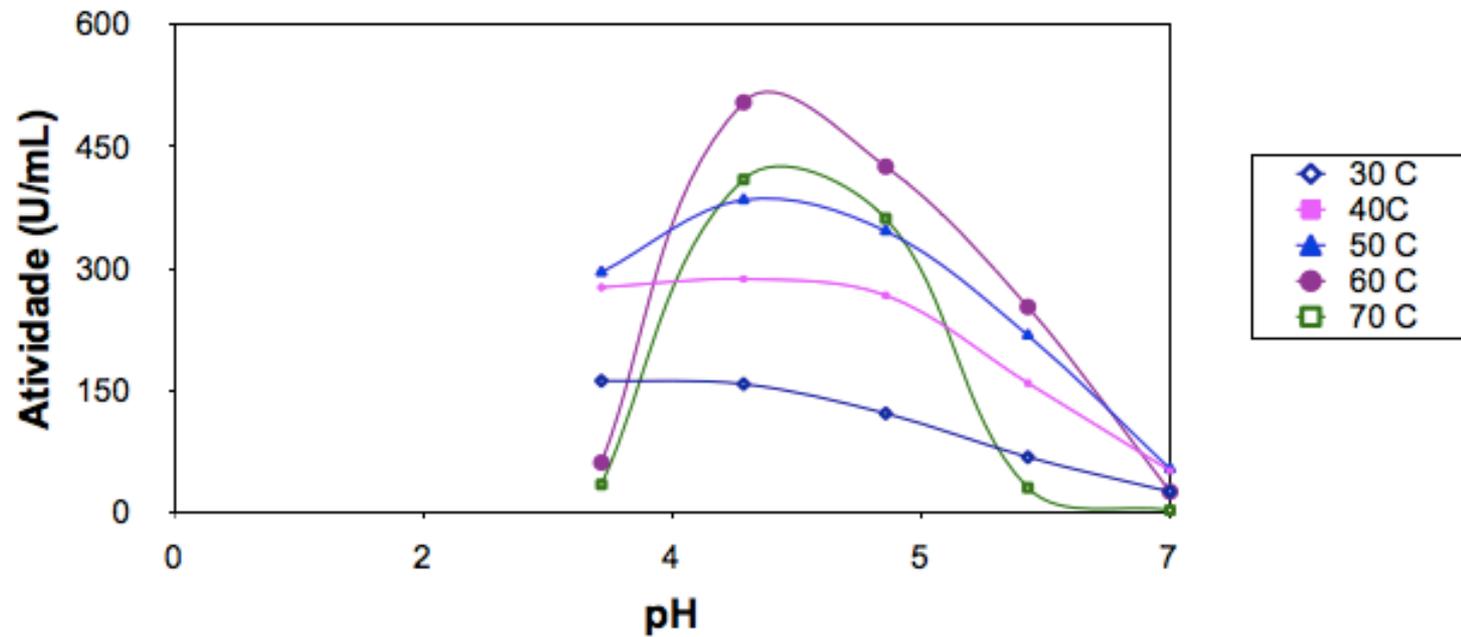
***Estudo Fatorial (25 ensaios x repetições):***

Resultados da atividade enzimática para diversas combinações de pH e temperatura

Temp.(°C) pH	30	40	50	60	70
3	162	277	295	61	34
4	158	287	384	504	409
5	122	267	346	425	361
6	68	159	218	253	30
7	25	51	53	26	3

## Comportamento da atividade no estudo

### Temperatura x pH



## Planejamento Experimental Fatorial

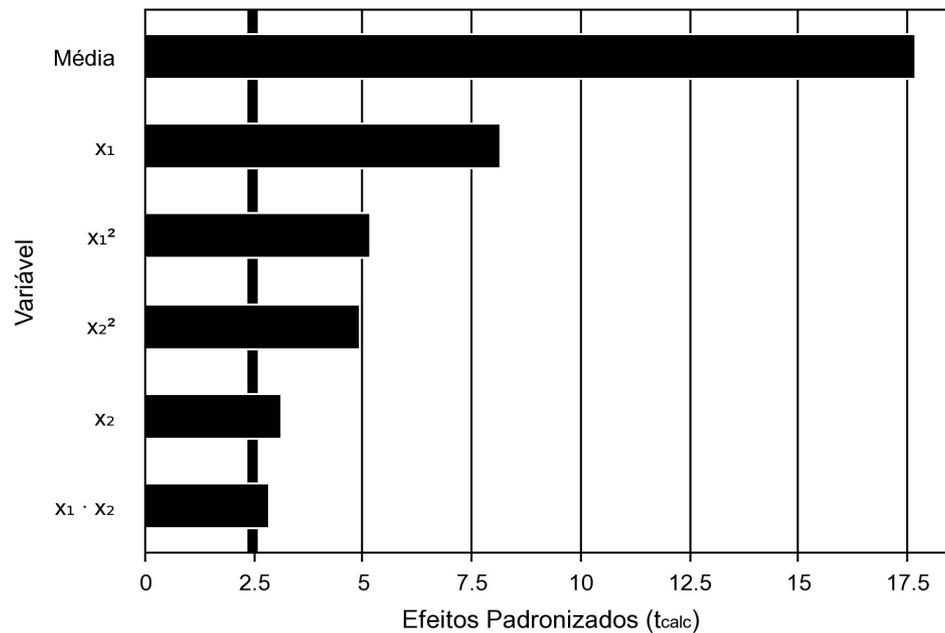
Faixa de estudo das variáveis explanatórias: valores codificados e reais

	-1,41	-1	0	+1	+1,41
pH	3,0				7,0
Temperatura (°C)	30				70



## Matriz do Planejamento Experimental Fatorial

Ensaio	X1	X2	pH	Temperatura	Resposta
1	-1	-1	3,6	36	272
2	+1	-1	6,4	36	83
3	-1	+1	3,6	64	457
4	+1	+1	6,4	64	16
5	-1,41	0	3,0	50	360
6	+1,41	0	7,0	50	83
7	0	-1,41	5,0	30	132
8	0	+1,41	5,0	70	328
9	0	0	5,0	50	396
10	0	0	5,0	50	412
11	0	0	5,0	50	393
12	0	0	5,0	50	371



## Regressão

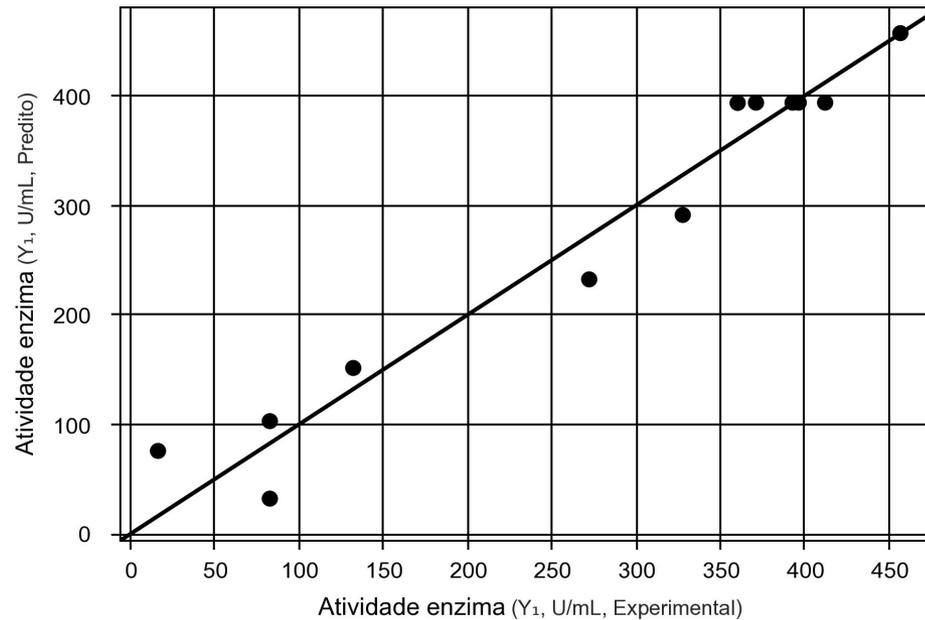
	Nome	Coeficiente	Erro Padrão	t calculado	p-valor
<input type="checkbox"/>	Média	393,00	22,19	17,71	0,0000
<input checked="" type="checkbox"/>	X <sub>1</sub>	-127,72	15,69	-8,14	0,0002
<input checked="" type="checkbox"/>	X <sub>1</sub> <sup>2</sup>	-90,44	17,55	-5,15	0,0021
<input checked="" type="checkbox"/>	X <sub>2</sub>	49,40	15,69	3,15	0,0199
<input checked="" type="checkbox"/>	X <sub>2</sub> <sup>2</sup>	-86,19	17,55	-4,91	0,0027
<input checked="" type="checkbox"/>	X <sub>1</sub> · X <sub>2</sub>	-63,00	22,19	-2,84	0,0296

$$Y_1 = 393 - 127,72 x_1 - 90,44 x_1^2 + 49,40 x_2 - 86,19 x_2^2 - 63 x_1 x_2$$

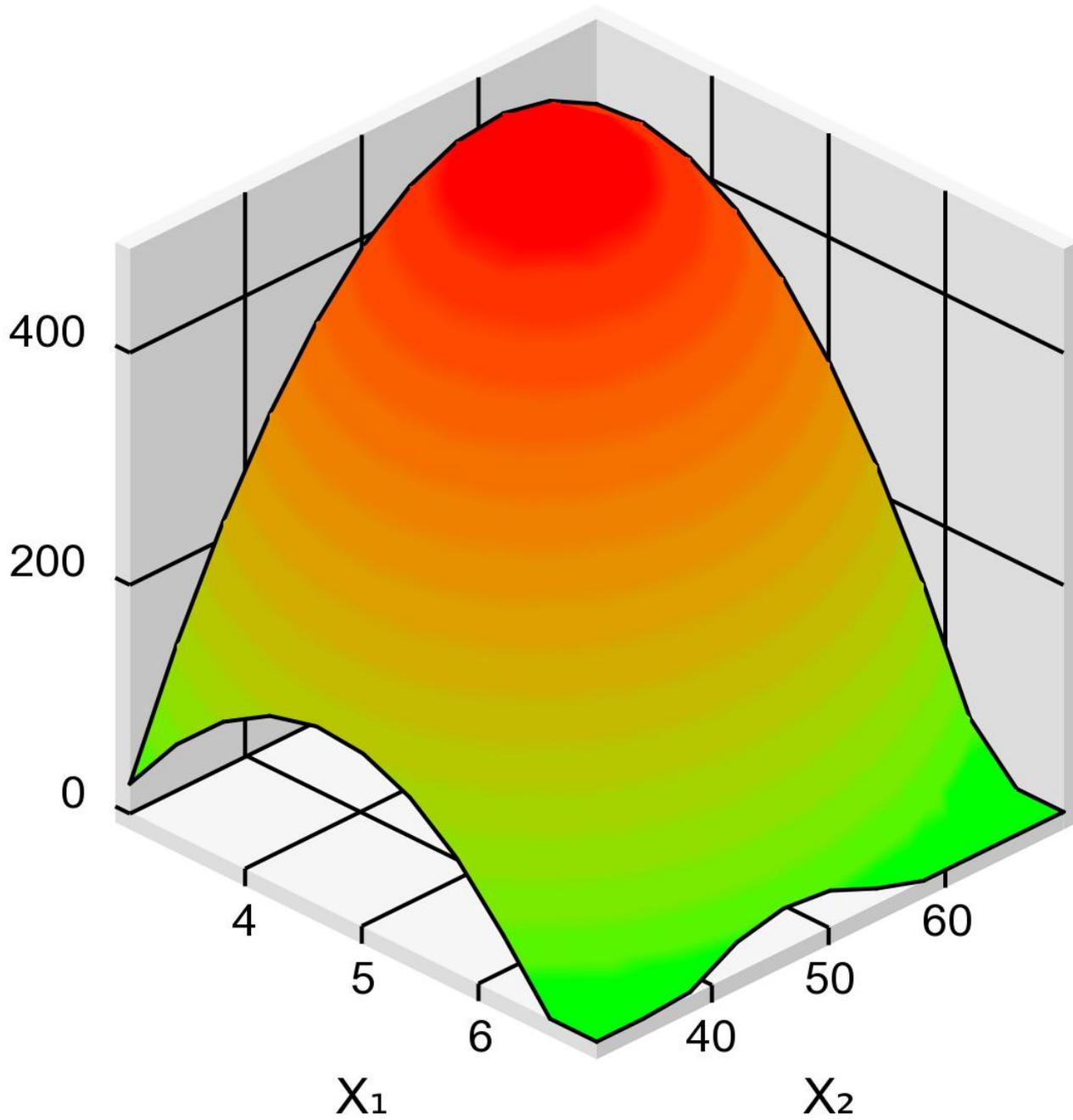
## ANOVA

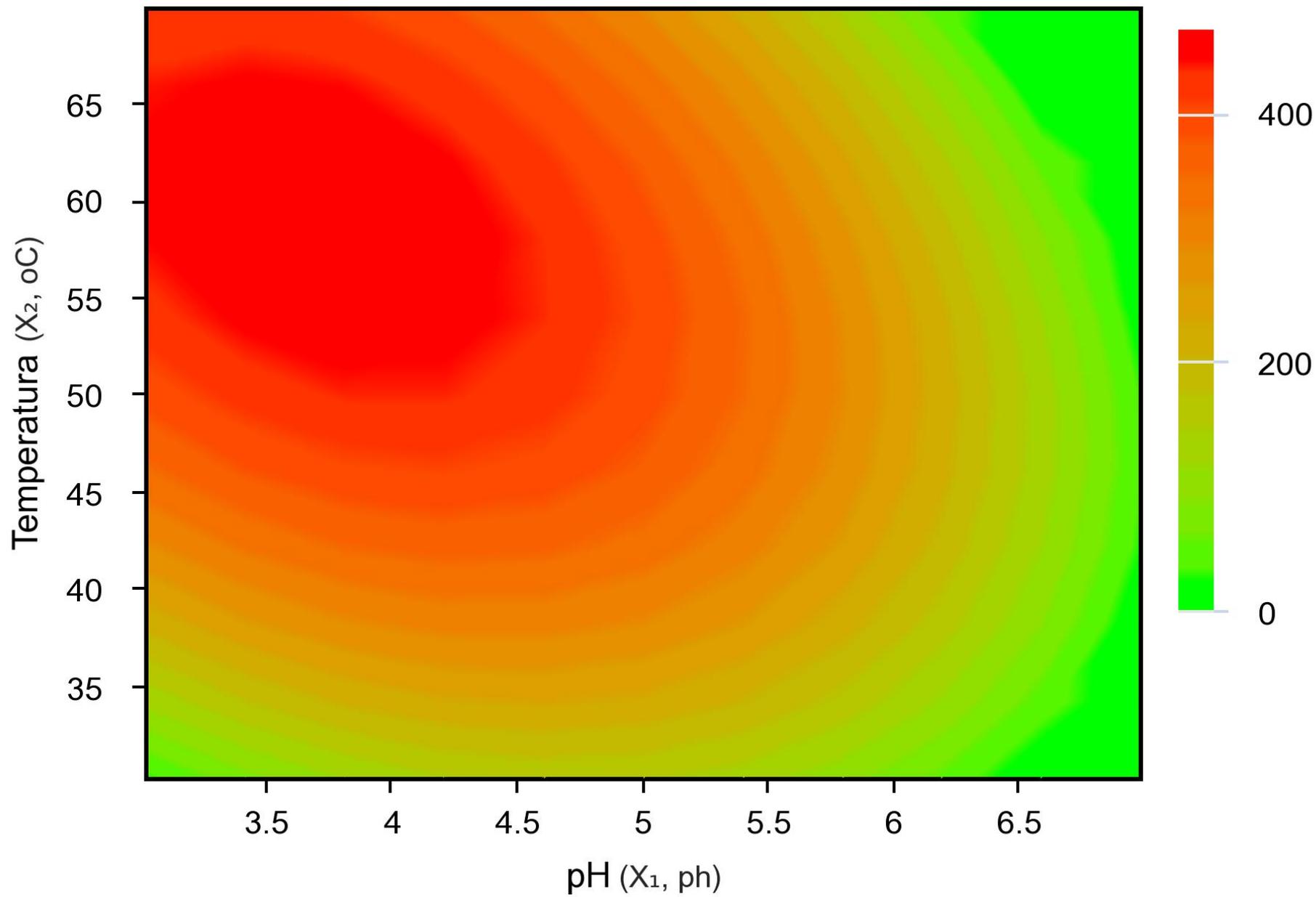
Fonte de Variação	Soma dos Quadrados	Graus de Liberdade	Quadrado Médio	F <sub>calc</sub>	p-valor
Regressão	249153,5	5	49830,7	25,3	0,00058
Resíduos	11820,8	6	1970,1		
Falta de Ajuste	10966,8	3	3655,6	12,8	0,03224
Erro Puro	854,0	3	284,7		
Total	260974,3	11			

R<sup>2</sup> = 95,47%



Atividade enzima ( $Y_1$ , U/mL)





## Conclusões

- ✓ uso da metodologia do planejamento experimental:
  - ✓ possível a realização de 12 ensaios, ao invés dos 25 para chegar as condições ótimas de temperatura e pH
- ✓ Quanto maior o número de variáveis a estudar, mais a técnica do planejamento experimental reduz o número de ensaios necessários, quando comparado a metodologia univariável