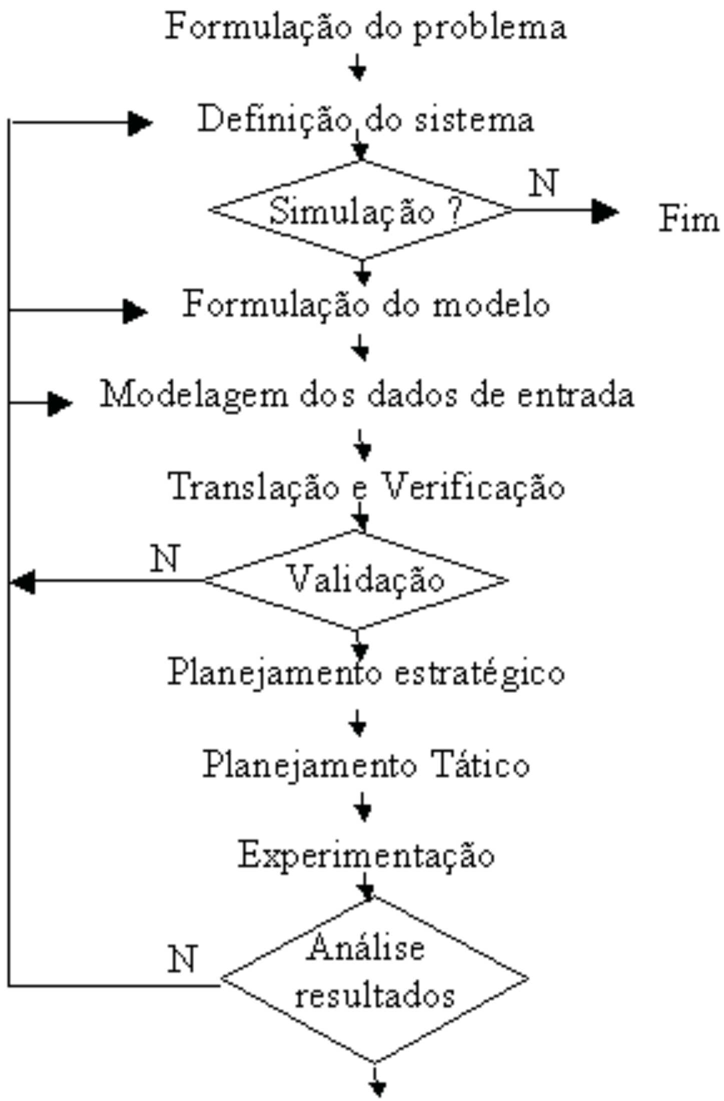
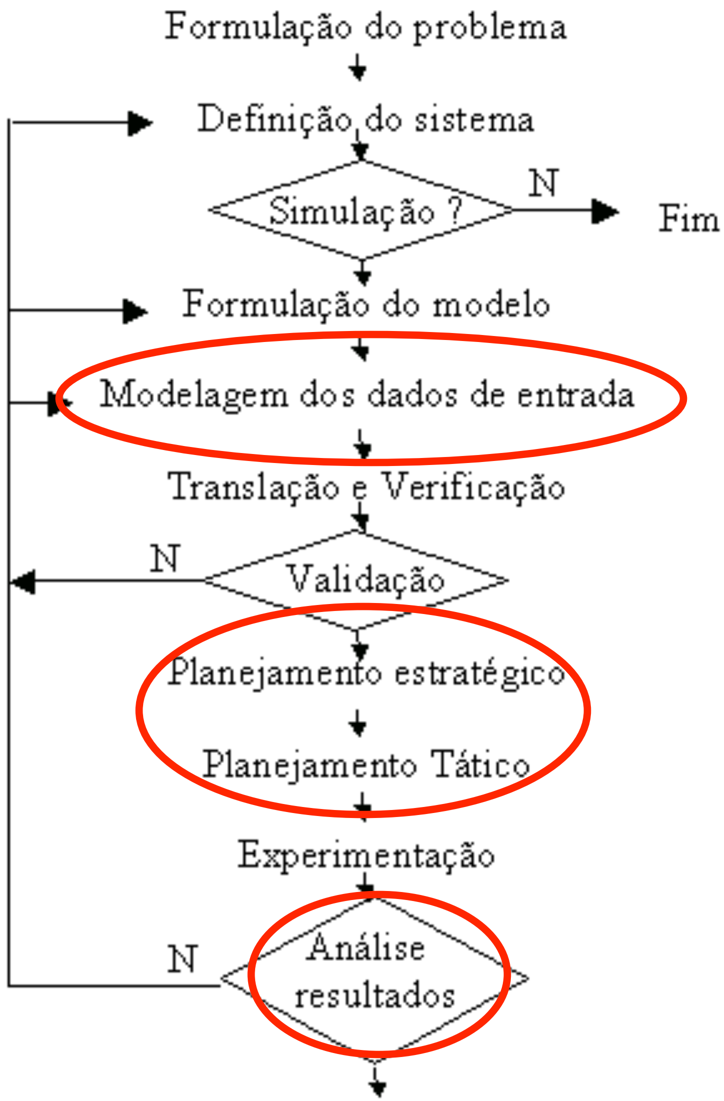
A preparação de experimentos

- Fases do processo de experimentação (simulação)

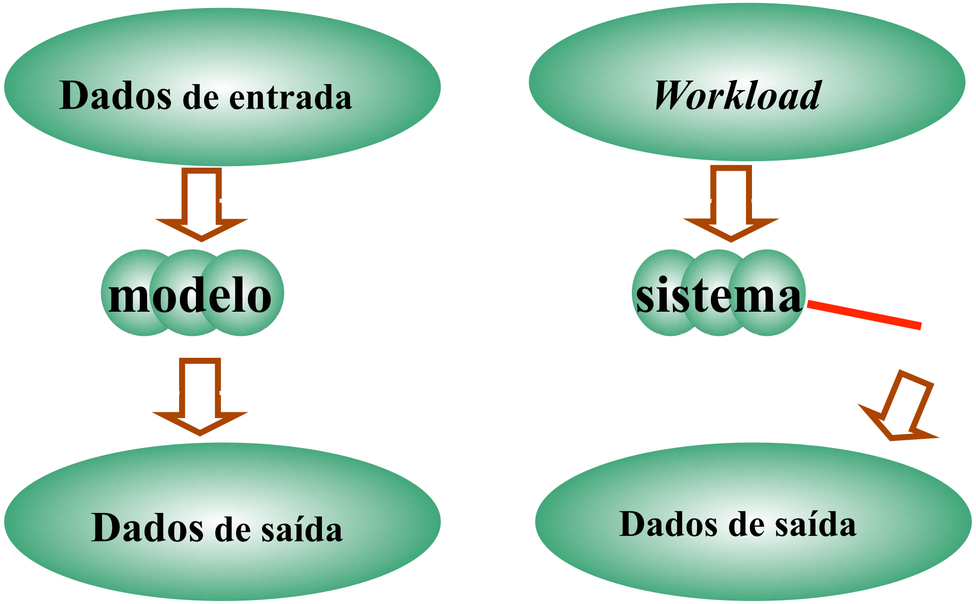


A preparação de experimentos

- Fases do processo de experimentação (simulação)



Preparação de experimentos de Simulação e Monitoração



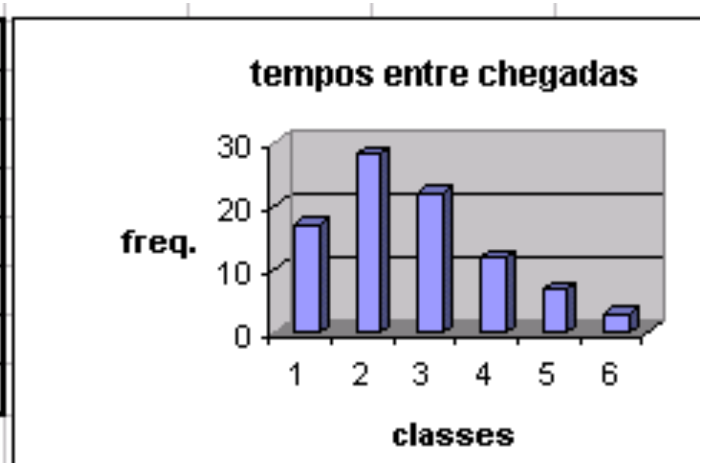
A preparação de experimentos

- Modelagem de dados de entrada
  - Identificação de distribuições de probabilidade
  - Estimar parâmetros
  - Fit- Tests
- Planejamento de experimentos
- Análise de dados de saída

Modelagem de Dados de Entrada

- Identificação de distribuições de probabilidade
  - A partir dos dados coletados:
    - Elaborar histograma ⇨ distribuições de freqüências
    - Selecionar uma distribuição de probabilidade
  - Ex.: tempos entre-chegadas de clientes em um caixa de banco

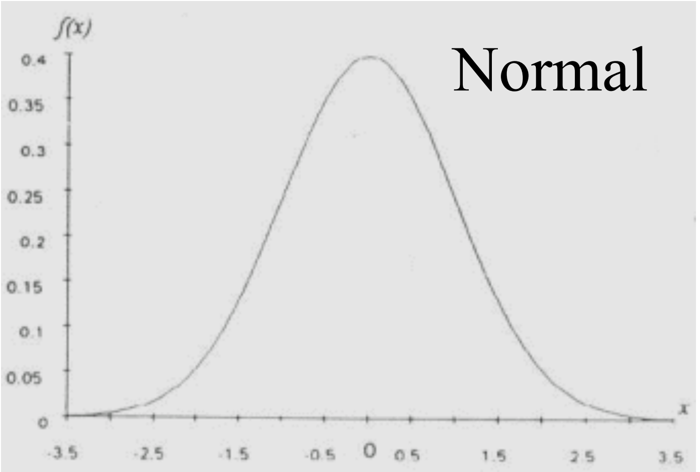
classes	freqüência
1	17
2	28
3	22
4	12
5	7
6	3
n =	89

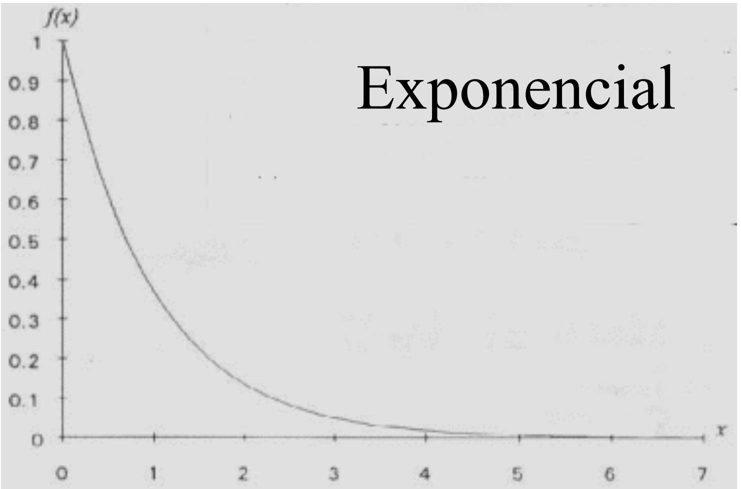


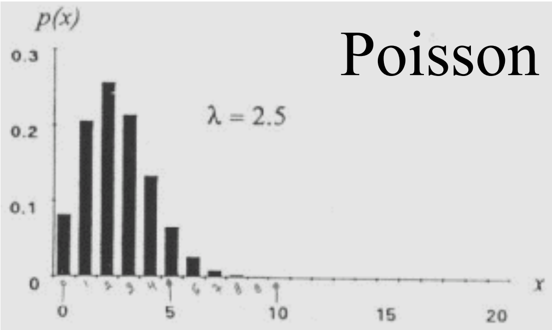
Modelagem de Dados de Entrada

- Critérios para seleção da distribuição
  - Natureza do processo sendo modelado
  - Comparação visual das curvas

- Exemplos







Modelagem de Dados de Entrada

- Exponencial
  - Tempos entre eventos (entre chegadas, por ex.)
  - Eventos independentes (*memoryless*)
- Poisson
  - Modela o número de eventos independentes que ocorrem em um intervalo de tempo
- Normal
  - Duração de tarefas onde a variabilidade é baixa; curva simétrica;
  - 68% da sua área está entre média ± desvio-padrão

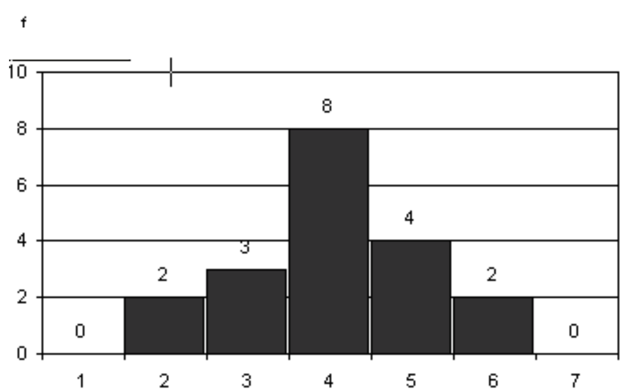
Modelagem de Dados de Entrada

- Estimar parâmetros da distribuição escolhida
  - Se foi selecionada uma Normal, então calcula-se:
    - Média
    - Desvio-padrão

$$\overline{X} = \left(\sum_{i=1}^{k} f_i X_i \right) / n$$

$$S = \left(\left(\sum f_i X_i^{\,2} - n\overline{X}^2 \right) / (n-1) \right)^{1/2}$$

# classe	inf	sup	f	X	f.X	X-x	f.x2
1		2	0				0
2	2	4	2	3	6	-4,11	33,706
3	4	6	3	5	15	-2,11	13,296
4	6	8	8	7	56	-0,11	0,0886
5	8	10	4	9	36	1,89	14,36
6	10	12	2	11	22	3,89	30,338
7	12		0				0
			Σ		Σ		Σ
			19		135		91,789
média=	7,105		n				
desvio=	2,258						



Modelagem de Dados de Entrada

- Goodness-of-Fit Tests (GOD) ou simplesmente "Fit Tests"
  - **Testes Paramétricos: se baseiam no Teorema do Limite Central (TLC)**
    - *Qui-Quadrado* (χ^2)
  - **Testes não-paramétricos:** não supõe nada em relação à forma da distribuição sendo testada; podem ser usados quando o tamanho da amostra é pequeno ($n \leq 30$), o que exclui a aplicação do TLC
    - *Kolmogorov-Smirnov* (KS) ; *Mann-Whitney*

Modelagem de Dados de Entrada

- Exemplo

	Limites	L-x	z=(L-x)/s	Normal	Normal * n	freq teo esp	freq obs	O-E	(0-E)2	(O-E)2 / E
1	-100	-107	-47,43	0,500	9,500					
	2	-5,1	-2,26	0,488	9,274	0,226	0,000	-0,226	0,051	0,226
2	2	-5,1	-2,26	0,488	9,274					
	4	-3,1	-1,38	0,415	7,894	1,381	2,000	0,619	0,384	0,278
3	4	-3,1	-1,38	0,415	7,894					
	6	-1,1	-0,49	0,188	3,567	4,327	3,000	-1,327	1,760	0,407
4	6	-1,1	-0,49	0,188	3,567					
	8	0,89	0,40	-0,154	-2,927	6,494	8,000	1,506	2,269	0,349
5	8	0,89	0,40	-0,154	-2,927					
	10	2,89	1,28	-0,400	-7,601	4,675	4,000	-0,675	0,455	0,097
6	10	2,89	1,28	-0,400	-7,601					
	12	4,89	2,17	-0,485	-9,213	1,612	2,000	0,388	0,150	0,093
7	12	4,89	2,17	-0,485	-9,213					
	100	92,9	41,14	-0,500	-9,500	0,287	0,000	-0,287	0,082	0,287

Σ

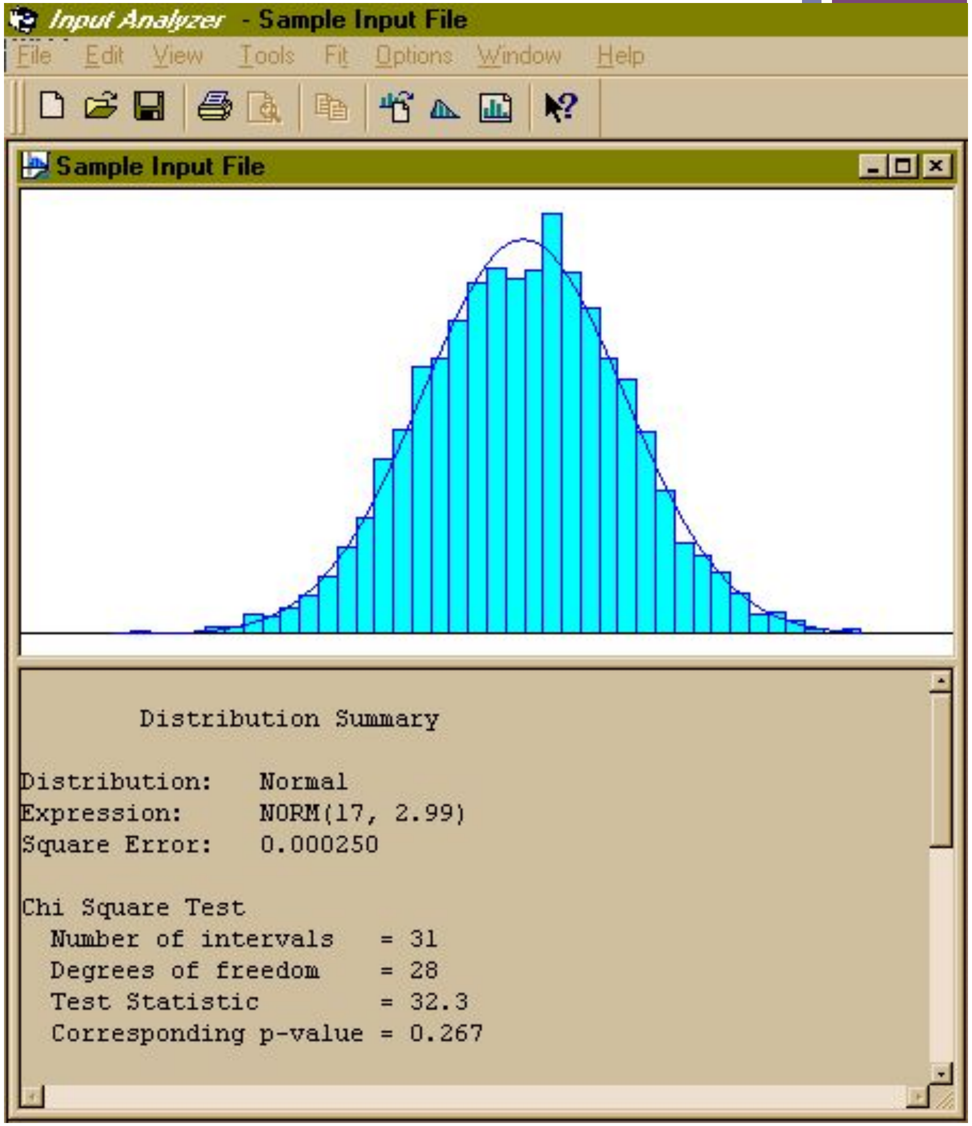
D=classes-2-1
D=7-2-1=4

19,000	19		**chi calc=**	**1,7375**
nível signif:	0.05	0.5	0.90	0.95
chi tabela:	9,49	3,35	1,064	0,711

Modelagem de Dados de Entrada

- Systems Modeling Corp. – Arena 3.5

Input Analyzer



Planejamento estratégico

- Objetivos:
  - Encontrar o conjunto de experimentos que melhor representa a resposta desejada
  - Este conjunto deve ter o menor tamanho possível = minimizar o número de experimentos (*runs*)
  - Obter o máximo de informação com o menor esforço possível
  - Análise de sensibilidade do modelo aos dados de entrada

Planejamento estratégico

- Terminologia
  - Variável de resposta
  - Fatores
    - Fatores primários
    - Fatores secundários
  - Níveis (por fator)
  - Interação entre fatores
  - Replicação
  - Experimento simétrico: k fatores, todos com q níveis

Planejamento estratégico

- Erros comuns no projeto de experimentos
  - Influência do erro é ignorada
  - Parâmetros importantes não são controlados
  - Efeitos de fatores diferentes não são isolados
  - Interações são ignoradas
  - Número de *runs* é excessivo

Planejamento estratégico

- Tipos de projetos (*designs*)
  - simples
  - 2^k fatorial (*full factorial*)
  - $2^k r$ fatorial (*full factorial*)
  - 2^{k-p} fracionários

Projeto Simples

- A partir de uma configuração típica, varia um fator por vez
  ⇨como este fator influencia o desempenho
- Não é estatisticamente eficiente
- Se os fatores possuírem interações, pode levar a conclusões errôneas
- Se forem k fatores cada um com n_i níveis, então o número total de experimentos n será $n=1+\sum (n_i-1)$ onde i=1 a k

Projeto 2^k fatorial

Cache Size (kbytes)	Memory Size 4 Mbytes	Memory Size 16 Mbytes
1	15	45
2	25	75

Projeto 2^k fatorial

❖

$$x_A = \begin{cases} -1 & \text{if 4 Mbytes memory} \\ 1 & \text{if 16 Mbytes memory} \end{cases}$$

$$x_B = \begin{cases} -1 & \text{if 1 kbyte cache} \\ 1 & \text{if 2 kbytes cache} \end{cases}$$

❖

$$15 = q_0 - q_A - q_B + q_{AB}$$

$$45 = q_0 + q_A - q_B - q_{AB}$$

$$25 = q_0 - q_A + q_B - q_{AB}$$

$$75 = q_0 + q_A + q_B + q_{AB}$$

Projeto 2^k fatorial

Experiment	A	B	y
1	-1	-1	y_1
2	1	-1	y_2
3	-1	1	y_3
4	1	1	y_4

Projeto 2^k fatorial

❖

I	A	B	AB	y
1	−1	−1	1	15
1	1	−1	−1	45
1	−1	1	−1	25
1	1	1	1	75
160	80	40	20	Total
40	20	10	5	Total/4

Projeto 2^k fatorial

$$s_y^2 = \frac{\sum_{i=1}^{2^2}(y_i - \bar{y})^2}{2^2 - 1}$$

-
-
-
-
- $$SST = \sum_{i=1}^{2^2}(y_i - \bar{y})^2$$

  $$SST = 2^2 q_A^2 + 2^2 q_B^2 + 2^2 q_{AB}^2$$
-
-

Projeto 2^k fatorial

Projeto $2^k r$ fatorial

❖

$$y = q_0 + q_A x_A + q_B x_B + q_{AB} x_A x_B + e$$

❖

❖

❖

I	A	B	AB	y	Mean $\bar{y}$
1	−1	−1	1	(15, 18, 12)	15
1	1	−1	−1	(45, 48, 51)	48
1	−1	1	−1	(25, 28, 19)	24
1	1	1	1	(75, 75, 81)	77
164	86	38	20		Total
41	21.5	9.5	5		Total/4

Projeto $2^k r$ fatorial

❖

	Effect				Estimated Response,	Measured Responses			Errors		
	I	A	B	AB							
i	41	21.5	9.5	5	$\hat{y}_i$	y_{i1}	y_{i2}	y_{i3}	e_{i1}	e_{i2}	e_{i3}
1	1	−1	−1	1	15	15	18	12	0	3	−3
2	1	1	−1	−1	48	45	48	51	−3	0	3
3	1	−1	1	−1	24	25	28	19	1	4	−5
4	1	1	1	1	77	75	75	81	−2	−2	4

❖

$$SSE = \sum_{i=1}^{2^2}\sum_{j=1}^{r} e_{ij}^2$$

$$\begin{aligned} SSE &= 0^2 + 3^2 + (-3)^2 + (-3)^2 + 0^2 + 3^2 + 1^2 \\ &\quad + 4^2 + (-5)^2 + (-2)^2 + (-2)^2 + 4^2 \\ &= 102 \end{aligned}$$

Projeto $2^k r$ fatorial

$$SST = \sum_{i,j} (y_{ij} - \bar{y}_{..})^2$$

$$\sum_{i,j} (y_{ij} - \bar{y}_{..})^2 = 2^2 r q_A^2 + 2^2 r q_B^2 + 2^2 r q_{AB}^2 + \sum_{i,j} e_{ij}^2$$

$$SST = SSA + SSB + SSAB + SSE$$

Projeto $2^k r$ fatorial

$$SSY = SS0 + SSA + SSB + SSAB + SSE$$

$$SST = SSY - SS0 = SSA + SSB + SSAB + SSE$$

$$SSY = 15^2 + 18^2 + 12^2 + 45^2 + \cdots + 75^2 + 75^2 + 81^2 = 27{,}204$$

$$SS0 = 2^2 r q_0^2 = 12 \times 41^2 = 20{,}172$$

$$SSA = 2^2 r q_A^2 = 12 \times (21.5)^2 = 5547$$

$$SSB = 2^2 r q_B^2 = 12 \times (9.5)^2 = 1083$$

$$SSAB = 2^2 r q_{AB}^2 = 12 \times 5^2 = 300$$

$$SSE = 27{,}204 - 2^2 \times 3(41^2 + 21.5^2 + 9.5^2 + 5^2) = 102$$

$$SST = SSY - SS0 = 27{,}204 - 20{,}172 = 7032$$

Projeto $2^k r$ fatorial

$$s_e^2 = \frac{\text{SSE}}{2^2(r-1)}$$

$$s_e = \sqrt{\frac{\text{SSE}}{2^2(r-1)}} = \sqrt{\frac{102}{8}} = \sqrt{12.75} = 3.57$$

$$s_{q_i} = s_e / \sqrt{(2^2 r)} = 3.57/\sqrt{12} = 1.03$$

Projeto $2^k r$ fatorial

-

$$q_i \mp t_{[1-\alpha/2;2^2(r-1)]} s_{q_i}$$

-
-
-

Projeto 2^{k-p} fracionário

ntos

Planejamento tático

Condições iniciais

Condições iniciais

Análise de regime permanente

❖

o

•

o

•