Experimentos de Um Fator

- Objetivo: comparar *a* alternativas ($a \geq 2$) de uma única variavel categórica
  - Comparar impacto de diferentes processadores
  - Comparar impacto de diferentes algoritmos para um mesmo problema
  - Comparar diferentes politicas de caching
- Nao ha limites no numero de niveis para o fator analisado

Experimentos de Um Fator

- Modelo: $y_{ij} = \mu + \alpha_j + e_{ij}$
  - y_{ij}: i-esima resposta com o fator no nivel j
  - μ: resposta media
  - α_j : efeito da alternativa j
  - e_{ij}: erro experimental
- Tem-se que:

$$\sum_j \alpha_j = 0 \quad \text{e} \quad \sum_i e_{ij} = 0$$

Computando os Efeitos

- Se $y_{ij} = \mu + \alpha_j + e_{ij}$, tem-se que:

$$\sum_{i=1}^{r}\sum_{j=1}^{a} y_{ij} = ar\mu + r\sum_{j=1}^{a}\alpha_j + \sum_{i=1}^{r}\sum_{j=1}^{a} e_{ij} = ar\mu + 0 + 0$$

$$\mu = \frac{1}{ar}\sum_{i=1}^{r}\sum_{j=1}^{a} y_{ij} = \overline{\mathrm{y}_{..}} \quad : \text{media geral}$$

Computando os Efeitos

$$\overline{\mathrm{y}_{.\mathrm{j}}} = \frac{1}{r}\sum_{i=1}^{r} y_{ij} = \frac{1}{r}\sum_{i=1}^{r}(\mu + \alpha_j + e_{ij})$$

$$= \frac{1}{r}(\mathrm{r}\mu + r\alpha_j + \sum_{i=1}^{r} e_{ij}) = \mu + \alpha_j$$

$$\alpha_j = \overline{\mathrm{y}_{.\mathrm{j}}} - \mu$$

Exemplo

- Em uma comparacao de tamanho de codigo, o numero de bytes necessario para codificar uma certa carga em tres processadores diferentes, R, V e Z, foi medido, cada um 5 vezes (diferentes programadores) Os dados medidos sao mostrados na tabela abaixo:

R	144	120	176	288	144
V	101	144	211	288	72
Z	130	180	141	374	302

Assume-se que nao ha dependencia entre as codificacoes feitas para um mesmo processador. Caso contrário, deveria ter sido feito um projeto de 2 fatores

Exemplo

						Soma da Linha	Media Linha	Efeito Linha
R	144	120	176	288	144	872	174.4	-13.3
V	101	144	211	288	72	816	163.2	-24.5
Z	130	180	141	374	302	1127	225.4	37.7
					Total	2815	187.7	

Um processador medio requer 187.7 bytes de armazenamento

Em media, proc. R requer 13.3 bytes a menos que a media; V requer 24.5 bytes a menos que a media, e Z requer 37.7 bytes a mais que a media.

Estimando erros experimentais

$$\hat{y}_j = \mu + \alpha_j \qquad e_{ij} = y_{ij} - \hat{y}_j$$

- Variacao dos erros e estimada pela Suma dos Erros Quadrados (SSE)

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

Exemplo

$$\begin{bmatrix} 144 & 101 & 130 \\ 120 & 144 & 180 \\ 176 & 211 & 141 \\ 288 & 288 & 374 \\ 144 & 72 & 302 \end{bmatrix} = \begin{bmatrix} 187.7 & 187.7 & 187.7 \\ 187.7 & 187.7 & 187.7 \\ 187.7 & 187.7 & 187.7 \\ 187.7 & 187.7 & 187.7 \\ 187.7 & 187.7 & 187.7 \end{bmatrix} + \begin{bmatrix} -13.3 & -24.5 & 37.7 \\ -13.3 & -24.5 & 37.7 \\ -13.3 & -24.5 & 37.7 \\ -13.3 & -24.5 & 37.7 \\ -13.3 & -24.5 & 37.7 \end{bmatrix} +$$

$$+ \begin{bmatrix} -30.4 & -62.2 & -95.4 \\ -54.4 & -19.2 & -45.4 \\ 1.6 & 47.8 & -84.4 \\ 113.6 & 124.8 & 148.6 \\ -30.4 & -91.2 & 76.6 \end{bmatrix}$$

$$SSE = (-30.4)^2 + (-54.4)^2 + \ldots + (76.6)^2 = 94365.20$$

Alocacao de Variacao

- Se $y_{ij} = \mu + \alpha_j + e_{ij}$, tem-se que:

$$y_{ij}^2 = \mu^2 + \alpha_j^2 + e_{ij}^2 + 2\mu\alpha_j + 2\mu e_{ij} + 2\alpha_j e_{ij}$$

$$\sum_{ij} y_{ij}^2 = \sum_{ij} \mu^2 + \sum_{ij} \alpha_j^2 + \sum_{ij} e_{ij}^2$$

$$SSY = SS0 + SSA + SSE$$

$$SSO = \sum_{ij} \mu^2 = ar\mu^2 \qquad SSA = \sum_{ij} \alpha_j^2 = r\sum_{j=1}^{a} \alpha_j^2$$

$$SST = \sum_{ij} (y_{ij} - \overline{y_{..}})^2 = SSY - SS0 = SSA + SSE$$

Apresentacao derivada dos slides originais de Virgilio Almeida

Exemplo

- SSY = $144^2 + 120^2 + \ldots + 302^2 = 633639$
- SS0 = $ar\mu^2 = 3 * 5 * (187.7)^2 = 528281.7$
- $$SSA = r\sum_{j=1}^{a}\alpha_j^2 = 5\left[(-13.3)^2 + (-24.5)^2 + (37.6)^2\right] = 10992.1$$
- SST = SSY – SS0 = 633639 – 528281.7 = 105357.3
- SSE = SST – SSA = 105357.3 – 10992.1 = 94365.2

% variacao explicada pelos processadores:
10992.13 / 105357.3 = 10.4%

Resto da variacao devido a erros experimentais

Analise de Variancia

- O impacto do fator e significativo?

- Precisa comparar a variacao explicada pelo fator com a variacao explicada pelos erros experimentais

Analise de Variancia: Executando Teste F

- Graus de liberdade (df):

$$SSY = SS0 + SSA + SSE$$
$$ar = 1 + (a-1) + a(r-1)$$

- Quadrados Medios de A e de Erros
  - MSA = SSA / df(A) = SSA / (a-1)
  - MSE = SSE / df(e) = SSE / a(r-1)
- Computar razao MSA / MSE: esta razão segue uma distribuição F
  - Se maior que valor da tabela F com df=a-1 no numerador e df =a(r-1) no denominador e 1-α % de confianca:
    - Fator e significativo com confianca de 1-α %
    - Efeitos explicam fração significativa da variação na resposta (SSA é significativamente maior que SSE)
  - Caso contrario: fator nao e significativo

Exemplo: Tabela ANOVA

Componente	Soma Quadrados	% Var.	DF	MS	F Computado	F Tabela
Y	SSY = 633639.00					
y..	SS0 = 528281.69					
y – y..	SST = 105357.31	100	14			
A	SSA = 10992.13	10.4	2	5496.1	0.7	3.89
Erros	SSE = 94365.20	89.6	12	7863.8		

95% confiança

$$s_e = \sqrt{MSE} = \sqrt{7863.77} = 88.68$$

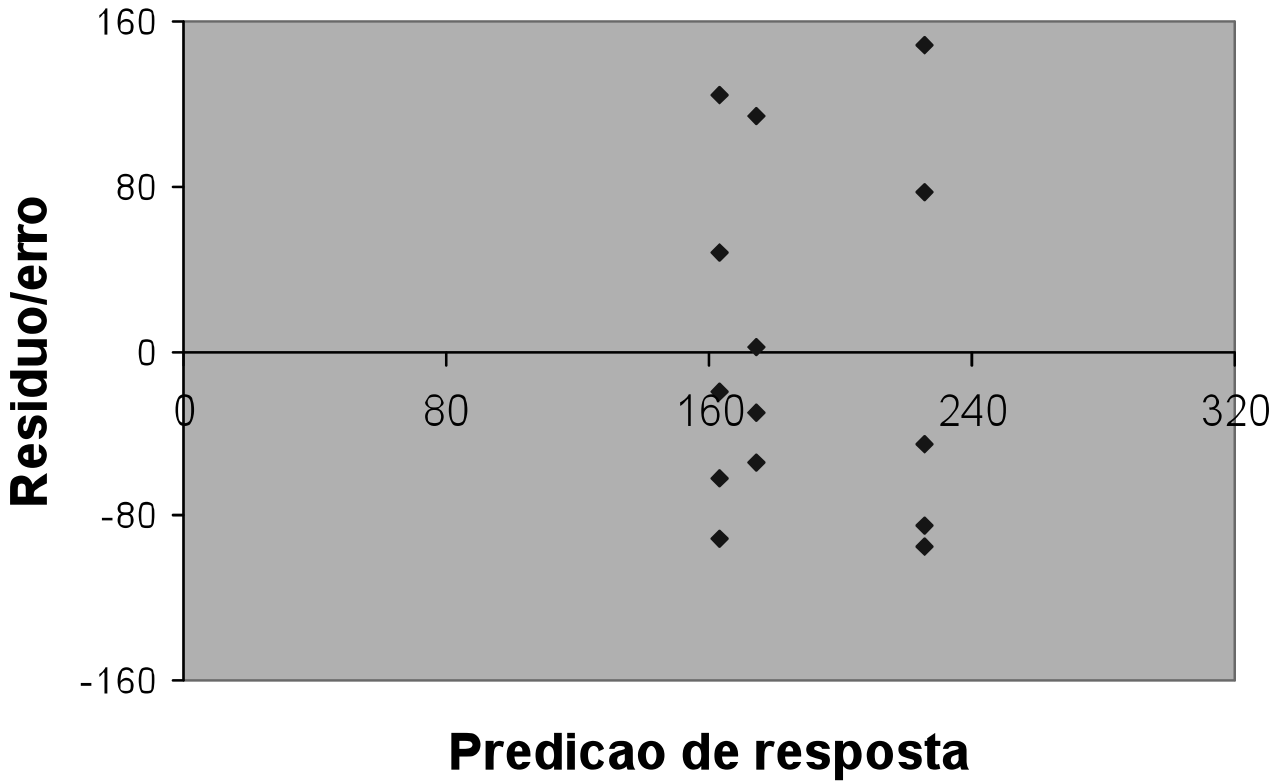
$F_{computado} < F_{tabela}$: diferencas em tamanhos de codigo sao principalmente devido aos "erros experimentais"

Diagnostico Visual

- Premissas do modelo
  - Efeitos dos fatores sao aditivos
  - Erros sao aditivos
  - Erros sao independentes dos niveis do fator
  - Erros sao normalmente distribuidos
  - Erros tem mesma variancia para todos niveis de fator
- Mesmos testes visuais aplicados a experimentos fatoriais podem ser aplicados aqui para verificar a aplicabilidade do modelo ou entender resultados “estranhos”

Exemplo

- Grafico de residuos (ou erros) X previsao de resposta



- Residuos nao sao pequenos se comparados com respostas
- Espalhamento no eixo y > espalhamento no eixo x
  - Variacao devido aos erros > variacao devido ao fator
- Espalhamento dos erros e homogeneo e nao ha tendencias claras

Intervalos de Confianca para Efeitos

Parametro	*Estimativa*	*Variancia*
μ	$\overline{y_{..}}$	s_e^2 / ar
α_j	$\overline{y_{.j}} - \overline{y_{..}}$	$s_e^2(a-1) / ar$
$\mu + \alpha_j$	$\overline{y_{.j}}$	s_e^2 / r
$\sum_{j=1}^{a} h_j \alpha_j, \sum_{j=1}^{a} h_j = 0$	$\sum_{j=1}^{a} h_j \overline{y_{.j}}$	$\sum_{j=1}^{a} s_e^2 h_j^2 / r$
s_e^2	$\sum \frac{e_{ij}^2}{a(r-1)}$	

Estimar IC para resposta media com fator em certo nivel

Estimar IC para diferenca dos efeitos de niveis do fator (Ex: α_1 - α_2)

Graus de liberdade dos erros : a (r - 1)

Apresentacao derivada dos slides originais de Virgilio Almeida

Exemplo

- Desvio padrao dos parametros

$$s_e^2 = \frac{94365.2}{12} = 7863.8$$

$$\text{Desvio padrao dos erros} = s_e = \sqrt{7863.8} = 88.7$$

$$\text{Desvio padrao de } \mu = s_e \Big/ \sqrt{ar} = 88.7 \Big/ \sqrt{15} = 22.9$$

$$\text{Desvio padrao de } \alpha_j = s_e \sqrt{(a-1)/(ar)} = 32.4$$

Exemplo

- 90% IC para os parametros ($t_{0.95,12}$ = 1.782)

$$\mu = 187.7 \pm (1.782)(22.9) = (146.9, 228.5)$$

$$\alpha_1 = -13.3 \pm (1.782)(32.4) = (-71.0, 44.4)$$

$$\alpha_2 = -24.5 \pm (1.782)(32.4) = (-82.2, 33.2)$$

$$\alpha_3 = 37.6 \pm (1.782)(32.4) = (-20.0, 95.4)$$

Tamanho do codigo em media e significativamente diferente de 0 (faz sentido)

Porem nenhum dos efeitos dos tres processadores sao significativos. Logo, nao podemos dizer com 90% de confianca que os processadores tem um efeito significativo no tamanho do codigo

Exemplo

- Comparando processador R com processador V, computamos IC para α_1 - α_2 (note que $\sum_j h_j = 0$)

$$\text{Valor medio para } \alpha_1 - \alpha_2 = \overline{y_{.1}} - \overline{y_{.2}} = 174.4 - 163.2 = 11.2$$

$$\text{Desvio padrao para } \alpha_1 - \alpha_2 = s_e \sqrt{\left(\sum h_j^2\right) / r} = 88.7\sqrt{2/5} = 56.1$$

$$90\%\ CI \text{ para } \alpha_1 - \alpha_2 = 11.2 \pm (1.782)(56.1) = (-88.7, 111.1)$$

$$90\%\ CI \text{ para } \alpha_1 - \alpha_3 = (174.4 - 225.4) \pm (1.782)(56.1)$$
$$= (-140.9, 48.9)$$

$$90\%\ CI \text{ para } \alpha_2 - \alpha_3 = (163.2 - 225.4) \pm (1.782)(56.1)$$
$$= (-162.1, 37.7)$$

Todos ICs incluem 0: nenhum proc. é superior

Amostras de tamanhos diferentes para cada nivel

- Modelo ainda e o mesmo: $y_{ij} = \mu + \alpha_j + e_{ij}$

- Alem disto: $\sum_j r_j \alpha_j = 0$

dado que r_j : numero de replicacoes no nivel j

- Definimos $N = \sum_j r_j$

Intervalos de Confianca para Efeitos

Parametro	*Estimativa*	*Variancia*
μ	$\overline{y_{..}}$	s_e^2 / N
α_j	$\overline{y_{.j}} - \overline{y_{..}}$	$s_e^2 (N - r_j) / N r_j$
$\mu + \alpha_j$	$\overline{y_{.j}}$	s_e^2 / r_j
$\sum_{j=1}^{a} h_j \alpha_j, \sum_{j=1}^{a} h_j = 0$	$\sum_{j=1}^{a} h_j \overline{y_{.j}}$	$s_e^2 \sum_{j=1}^{a} h_j^2 / r_j$
s_e^2	$\frac{\sum e_{ij}^2}{N - a}$	

Graus de liberdade dos erros : N - a

Apresentacao derivada dos slides originais de Virgilio Almeida

Tabela ANOVA

Componente	Soma Quadrados	% Var.	DF	MS
Y	SSY = $\sum y_{ij}^2$		N	
y..	SS0 = $N\mu^2$		1	
y – y..	SST = SSY – SSO	100	N-1	
A	SSA= $\sum_j r_j \alpha_j^2$	$100\dfrac{SSA}{SST}$	a-1	$MSA = \dfrac{SSA}{a-1}$
Erros	SSE = SST-SSA	$100\dfrac{SSE}{SST}$	N-a	$MSE = \dfrac{SSE}{N-a}$

$$s_e = \sqrt{MSE}$$

$$F_{computado} = \frac{MSA}{MSE} \qquad F_{tabela} = F_{[1-\alpha, a-1, N-a]}$$

Exemplo 2

- Suponha que voce descubra que 3 das observacoes no exemplo anterior nao tinham sido obtidas de forma correta. Logo elas nao deveriam ser utilizadas na analise. Das tres observacoes incorretas, suponha que uma seja do sistema V e duas do sistema Z. Agora voce tem entao um projeto de um fator unico com amostras de tamanhos diferentes. Refaca a analise:

Exemplo 2

						Soma da Linha	Media Linha	Efeito Linha
R	144	120	176	288	144	872	174.4	2.15
V	101	144	211	288		744	186	13.75
Z	130	180	141			451	150.33	-21.92
					Total	2067	172.25	

Um processador medio requer 172.25 bytes de armazenamento

Em media, proc. R requer 2.15 bytes a mais que a media; V requer 13.75 bytes a mais que a media, e Z requer 21.92 bytes a menos que a media.

Exemplo 2

$$\begin{bmatrix} 144 & 101 & 130 \\ 120 & 144 & 180 \\ 176 & 211 & 141 \\ 288 & 288 & \\ 144 & & \end{bmatrix} = \begin{bmatrix} 172.25 & 172.25 & 172.25 \\ 172.25 & 172.25 & 172.25 \\ 172.25 & 172.25 & 172.25 \\ 172.25 & 172.25 & \\ 172.25 & & \end{bmatrix} + \begin{bmatrix} 2.15 & 13.75 & -21.92 \\ 2.15 & 13.75 & -21.92 \\ 2.15 & 13.75 & -21.92 \\ 2.15 & 13.75 & \\ 2.15 & & \end{bmatrix} +$$

$$+ \begin{bmatrix} -30.4 & -85.0 & -20.33 \\ -54.4 & -42.0 & 29.67 \\ 1.6 & 25.0 & -9.33 \\ 113.6 & 102.0 & \\ -30.40 & & \end{bmatrix}$$

$$SSE = (-30.4)^2 + (-54.4)^2 + \ldots + (-9.33)^2 = 39113.87$$

Exemplo 2

- $SSY = 144^2 + 120^2 + \ldots + 141^2 = 397375$
- $SS0 = N\mu^2 = 12*(172.25)^2 = 356040.75$
- $SSA = 5\alpha_1^2 + 4\alpha_2^2 + 3\alpha_3^2 = 2220.38$
- SSE = 39113.87
- SST = SSY – SS0 = 41334.25

% variacao explicada pelos processadores:

2220.38 / 41334.25 = 5.4%

Resto da variacao devido a erros experimentais

Tabela ANOVA

Componente	Soma Quadrados	% Var.	DF	MS
Y	397375.0			
y..	356040.75			
y – y..	41334.25	100	11	
A	2220.38	5.37	2	1110.19
Erros	39113.87	94.63	9	4345.99

$$s_e = \sqrt{MSE} = \sqrt{4345.99} = 65.92$$

$$F_{computado} = \frac{MSA}{MSE} = 0.26 \qquad F_{tabela} = F_{[1-\alpha, a-1, N-a]} = 3.01$$

Sumario

- Ver tabela 20.1 (pagina 341)

Exercício 1

- Seu melhor amigo acabou de desenvolver um algoritmo revolucionário, chamado Xulambis, para o reconhecimento de ações humanas. Visando uma posterior comparação com o algoritmo estado-da-arte Zambis, ele fez um estudo para estimar a sensibilidade do algoritmo Xulambis, em termos de taxa de reconhecimento (%), aos seus dois parâmetros principais, α e β . Com base nos resultados de um projeto fatorial (com uma única replicação), mostrado nas 3 primeiras colunas da tabela abaixo, ele concluiu que o algoritmo Xulambis é muito mais sensível ao parâmetro α que ao parâmetro β , obtendo um desempenho melhor (maior taxa de reconhecimento) para valores mais altos de α.

Fator α	Fator β	Projeto Fatorial (uma replicação)	Replicações extras	
			2ª replicação	3ª replicação
-1	-1	70%	74%	64%
1	-1	80%	90%	84%
-1	1	45%	34%	43%
1	1	65%	55%	64%

Exercício 1

a) Com base apenas na única replicação inicialmente feita (3ª coluna da tabela), você concorda com a conclusão do seu amigo quanto à sensibilidade de Xulambis aos dois parâmetros considerados?

b) Com base apenas na única replicação inicialmente feita (3ª coluna da tabela), você concorda com a conclusão do seu amigo de que o desempenho de Xulambis é melhor para valores mais altos de α?

c) Conhecendo bem a variabilidade inerente ao problema de reconhecimento de ações humanas, você pede ao seu amigo que execute pelo menos mais 2 replicações para cada cenário. Os resultados são mostrados na 4ª e na 5ª colunas da tabela. Suas conclusões em (a) e (b) se mantêm?

d) Qual a porcentagem total da variação nos dados, observada no projeto com 3 replicações, você consegue explicar com a variação dos fatores?

e) Ainda considerando o projeto com 3 replicações, o efeito devido à interação entre os fatores α e β é significativo com 95% de confiança? Qual a maior confiança que você pode atribuir à afirmativa de que a interação entre os dois fatores é significativa?

Exercício 2

- Um sistema foi avaliado quanto ao impacto de 4 parâmetros – A, B, C e D. Para tanto, um projeto fatorial 2^{4-1} foi realizado e os resultados são mostrados abaixo. O subscrito representa o nível de cada fator (1 para inferior, e 2 para superior).

	C_1D_1	C_1D_2	C_2D_1	C_2D_2
A_1B_1	-	40	15	-
A_1B_2	-	20	10	-
A_2B_1	100	-	-	30
A_2B_2	120	-	-	50

- Qual o polinômio gerador deste projeto? Liste todos os *confoundings* , caso existam, e discuta o que eles representam em termos do impacto na precisão do modelo.
- Quantifique todos os efeitos e as porcentagens de variação explicada por cada um deles.
- Você sugeriria um projeto melhor? Qual e por quê?

Exercício 3

- O tempo de execução de um algoritmo foi avaliado em função de 3 fatores, a CPU (fator A), o tamanho da memória (fator B) e o número de discos (fator C). Foram executadas três replicações para cada configuração considerada, conforme mostrado na tabela abaixo. A tabela também mostra alguns resultados agregados da análise das observações. Responda:

Experimento	I	A	B	C	Tempo de execução(ms)			
					t_1	t_2	t_3	$\overline{t}$
1	1	-1	-1	1	98	100	102	100
2	1	1	-1	-1	245	249	256	250
3	1	-1	1	-1	45	54	52	50
4	1	1	1	1	300	301	299	300
Total	700	400	0	100				
Total/4	175	100	0	25				

$$\text{SSE} = 117$$

Exercício 3

(a) Qual o polinômio gerador deste projeto? Liste todos os confoundings, caso existam, e discuta o que eles representam em termos do impacto na precisão do modelo?

(b) Quais são as estimativas e desvios padrões para cada efeito? Apresente o seu modelo para estimativa do tempo de execuçao do algoritmo em questão.

(c) Qual a porcentagem da variação explicada por cada efeito e pelo modelo? Você está satisfeito com seu modelo? Justifique. Qual o seu próximo passo na avaliação do impacto dos fatores CPU, memória e disco no desempenho do algoritmo em questão?

(d) Com uma confiança de 90%, você poderia dizer que o efeito A é significativamente diferente do efeito C? Justifique.

(e) Você pode propor um melhor projeto de experimentos, mantendo o mesmo custo, que o proposto acima? Justifique

Projeto Fatorial Completo com Dois Fatores Sem Replicação

- Dois fatores A e B com a e b níveis respectivamente
  - Fatores são categóricos (ou tratados como tal)
    - Fatores quantitativos: usar regressão
  - Cada experimento é executado somente uma vez
    - Não irá conseguir separar interação de erros experimentais

$$y_{ij} = \mu + \alpha_j + \beta_i + e_{ij}$$

$$\sum_j \alpha_j = 0 \qquad \sum_i \beta_i = 0$$

Efeitos dos fatores e erros são aditivos

Cálculo dos efeitos

- Similar ao projeto de 1 fator

$$y_{ij} = \mu + \alpha_j + \beta_i + e_{ij}$$

$$\mu = \overline{y_{..}}$$

$$\alpha_j = \overline{y_{.j}} - \overline{y_{..}}$$

$$\beta_i = \overline{y_{i.}} - \overline{y_{..}}$$

$$\sum_j \alpha_j = 0 \qquad \sum_i \beta_i = 0$$

Exemplo

Tempo de processamento de diferentes cargas em três configurações de cache

Workloads	Two Caches	One Cache	No Cache
ASM	54.0	55.0	106.0
TECO	60.0	60.0	123.0
SIEVE	43.0	43.0	120.0
DHRYSTONE	49.0	52.0	111.0
SORT	49.0	50.0	108.0

Exemplo

Workloads	Two Caches	One Cache	No Cache	Row Sum	Row Mean	Row Effect
ASM	54.0	55.0	106.0	215.0	71.7	-0.5
TECO	60.0	60.0	123.0	243.0	81.0	8.8
SIEVE	43.0	43.0	120.0	206.0	68.7	-3.5
DHRYSTONE	49.0	52.0	111.0	212.0	70.7	-1.5
SORT	49.0	50.0	108.0	207.0	69.0	-3.2
Column Sum	255.0	260.0	568.0	1083.0		
Column Mean	51.0	52.0	113.6		72.2	
Column effect	-21.2	-20.2	41.4			

Em média: tempo = 72.2 ms (μ)

- Tempo com dois caches é em média 21.2 ms menor que a média
- Tempo com um cache é em média 20.2 ms menor que média
- Tempo sem cache é em média 41.44 ms maior que a média
- Carga também afeta: carga ASM gasta 0.5 ms a menos que a média enquanto TECO gasta 8.8 ms a mais que a média

Estimando "Erros Experimentais"

$$\hat{y}_{ij} = \mu + \alpha_j + \beta_i$$

$$e_{ij} = y_{ij} - \hat{y}_{ij} = y_{ij} - \mu - \alpha_j - \beta_i$$

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

Para o exemplo: $\hat{y}_{11} = \mu + \alpha_1 + \beta_1 = 72.2 - 21.2 - 0.5 = 50.5$

$y_{11} = 54$

$e_{11} = 3.5$

Estimando "Erros Experimentais"

Para o exemplo:

Workloads	Two Caches	One Cache	No Cache
ASM	3.5	3.5	-7.1
TECO	0.2	-0.8	0.6
SIEVE	-4.5	-5.5	9.9
DHRYSTONE	-0.5	1.5	-1.1
SORT	1.2	1.2	-2.4

$SSE = (3.5)^2 + (0.2)^2 + \ldots (-2.4)^2 = 236.81$

Alocação de Variação

SST = SSY – SS0 = SSA + SSB + SSE

$$SSY = \sum_{ij} y_{ij}^2 = 91595$$

$$SSO = ab\mu^2 = 3 \times 5 \times (72.2)^2 = 78192.59$$

$$SSA = b\sum_{j} \alpha_j^2 = 5 \times ((-21.2)^2 + (-20.2)^2 + (41.4)^2) = 12857.2$$

$$SSB = a\sum_{i} \beta_i^2 = 3 \times ((-0.5)^2 + (8.8)^2 + (-3.5)^2 + (-1.5)^2 + (3.2)^2) = 308.40$$

$$SST = SSY - SS0 = 91595 - 78192.59 = 13402.41$$

$$SSE = SST - SSA - SSB = 236.81$$

$$\%VarA = \frac{SSA}{SST} = 95.9\%$$

$$\%VarB = \frac{SSB}{SST} = 2.3\%$$

$$\%\text{var inexplicada} = \frac{\text{SSE}}{\text{SST}} = 1.8\%$$

ANOVA

- Objetivo: testar significância de um fator usando teste-F

- Graus de liberdade

$$\begin{array}{ccccccccc} \mathrm{SSY} & = & \mathrm{SS0} & + & \mathrm{SSA} & + & \mathrm{SSB} & + & \mathrm{SSE} \\ ab & = & 1 & + & (a-1) & + & (b-1) & + & (a-1)(b-1) \end{array}$$

- Means Squares

$$\mathrm{MSA} = \frac{\mathrm{SSA}}{a-1}$$

$$\mathrm{MSB} = \frac{\mathrm{SSB}}{b-1}$$

$$\mathrm{MSE} = \frac{\mathrm{SSE}}{(a-1)(b-1)}$$

$$\frac{MSA}{MSE} \sim F_{[a-1,(a-1)(b-1)]}$$

$Se\ \frac{\mathrm{MSA}}{\mathrm{MSE}} > F_{[1-\alpha, a-1, (a-1)(b-1)]}$ então fator A é significativo com $(1-\alpha)$% de confiança

$Se\ \frac{\mathrm{MSB}}{\mathrm{MSE}} > F_{[1-\alpha, b-1, (a-1)(b-1)]}$ então fator *B* é significativo com $(1-\alpha)$% de confiança

Tabela ANOVA

Compo-nent	Sum of Squares	%Variation	DF	Mean Square	F-Comp.	F-Table
y	$\mathrm{SSY} = \sum y_{ij}^2$		ab			
$\bar{y}_{...}$	$\mathrm{SS0} = ab\mu^2$		1			
$y - \bar{y}_{...}$	SST=SSY-SS0	100	$ab - 1$			
A	$\mathrm{SSA} = b\Sigma\alpha_j^2$	$100\left(\frac{\mathrm{SSA}}{\mathrm{SST}}\right)$	$a - 1$	$\mathrm{MSA} = \frac{\mathrm{SSA}}{a-1}$	$\frac{\mathrm{MSA}}{\mathrm{MSE}}$	$F_{[1-\alpha, a-1, (a-1)(b-1)]}$
B	$\mathrm{SSB} = a\Sigma\beta_i^2$	$100\left(\frac{\mathrm{SSB}}{\mathrm{SST}}\right)$	$b - 1$	$\mathrm{MSB} = \frac{\mathrm{SSB}}{b-1}$	$\frac{\mathrm{MSB}}{\mathrm{MSE}}$	$F_{[1-\alpha, b-1, (a-1)(b-1)]}$
e	$\mathrm{SSE} = \mathrm{SST} - (\mathrm{SSA} + \mathrm{SSB})$	$100\left(\frac{\mathrm{SSE}}{\mathrm{SST}}\right)$	$(a-1)(b-1)$	$\mathrm{MSE} = \frac{\mathrm{SSE}}{(a-1)(b-1)}$		

Exemplo: Tabela ANOVA

Component	Sum of Squares	%Variation	DF	Mean Square	F-Comp.	F-Table
y	91595.00					
$y_{..}$	78192.59					
y-$y_{..}$	13402.41	100.0%	14			
Caches	12857.20	95.9%	2	6428.60	217.2	3.1
Workloads	308.40	2.3%	4	77.10	2.6	2.8
Errors	236.80	1.8%	8	29.60		

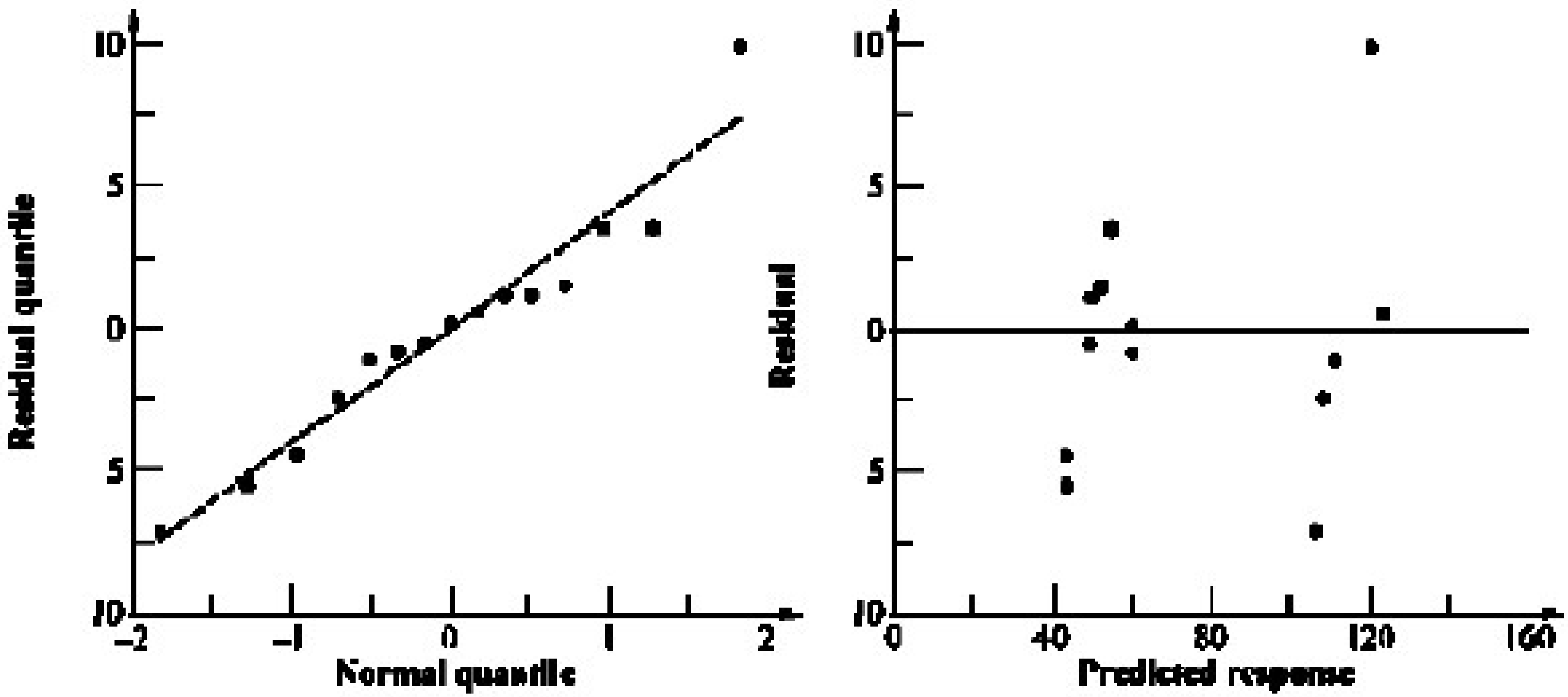
$$s_e = \sqrt{\text{MSE}} = \sqrt{29.60} = 5.44$$

Efeito do cache é significativo

Efeito da carga não é significativo

Testes Visuais

Mesmo testes usados em projeto de 1 fator



Apresentacao derivada dos slides originais de Virgilio Almeida

Intervalos de Confiança

Computar IC para diferentes parâmetros usando graus de liberdade dos erros (a-1)(b-1)

Parameter	Estimate	Variance
μ	$\bar{y}_{..}$	s_e^2/ab
α_j	$\bar{y}_{.j}$-$\bar{y}_{..}$	$s_e^2(a-1)/ab$
$\mu+\alpha_j$	$\bar{y}_{.j}$	s_e^2/b
β_i	$\bar{y}_{i.}$-$\bar{y}_{..}$	$s_e^2(b-1)/ab$
$\mu+\alpha_j+\beta_i$	$\bar{y}_{.j}+\bar{y}_{i.}$-$\bar{y}_{..}$	$s_e^2(a+b-1)/(ab)$
$\sum_{j=1}^a h_j\ \alpha_j,\ \sum_{j=1}^a h_j = 0$	$\sum_{j=1}^a h_j\ \bar{y}_{.j}$	$s_e^2 \sum_{j=1}^a h_j^2/b$
$\sum_{i=1}^b h_i\ \beta_i,\ \sum_{i=1}^b h_i = 0$	$\sum_{i=1}^b h_i\ \bar{y}_{i.}$	$s_e^2 \sum_{i=1}^b h_i^2/a$
s_e^2	$\{\sum_{j=1}^a \sum_{i=1}^b e_{ij}^2\}/\{(a-1)(b-1)\}$	

Degrees of freedom for errors = (a-1)(b-1)

Intervalos de Confiança: Exemplo

Graus de liberdade = (a-1)(b-1) = 2*4 = 8

$t_{0.95,8}$= 1.86 s_e = 5.44

Parameter	Mean Effect	Std. Dev.	Confidence Interval
μ	72.2	1.4	(69.6, 74.8)
Caches			
Two Caches	-21.2	1.99	(-24.9, -17.5)
One Cache	-20.2	1.99	(-23.9, -16.5)
No Cache	41.4	1.99	(37.7, 45.1)

Intervalos de Confiança: Exemplo

Graus de liberdade = (a-1)(b-1) = 2*4 = 8

$t_{0.95,8} = 1.86 \quad s_e = 5.44$

Parameter	Mean Effect	Std. Dev.	Confidence Interval
μ	72.2	1.4	(69.6, 74.8)
Caches			
Two Caches	-21.2	1.99	(-24.9, -17.5)
One Cache	-20.2	1.99	(-23.9, -16.5)
No Cache	41.4	1.99	(37.7, 45.1)

$$s_\mu = \frac{s_e}{\sqrt{ab}} = \frac{5.44}{\sqrt{3 \times 5}}$$

$$s_{\alpha_j} = s_e \sqrt{\frac{a-1}{ab}} = 5.44 \sqrt{\frac{3-1}{3 \times 5}}$$

Intervalos de Confiança: Exemplo

Graus de liberdade = (a-1)(b-1) = 2*4 = 8

$t_{0.95,8}$= 1.86 s_e = 5.44

Parameter	Mean Effect	Std. Dev.	Confidence Interval
ASM	-0.5	2.81	(-5.8, 4.7)†
TECO	8.8	2.81	(3.6, 14.0)
SIEVE	-3.5	2.81	(-8.8, 1.7)†
DHRYSTONE	-1.5	2.81	(-6.8, 3.7)†
SORT	-3.2	2.81	(-8.4, 2.0)†

† ⇒ Not significant

$$s_{\beta_i} = s_e\sqrt{\frac{b-1}{ab}} = 5.44\sqrt{\frac{5-1}{3\times 5}}$$

Intervalos de Confiança: Exemplo

Graus de liberdade = (a-1)(b-1) = 2*4 = 8

$t_{0.95,8} = 1.86$

IC para diferença dos efeitos de duas configurações de cache alternativas

	Two Caches	One Cache	No Cache
Two Caches		(-7.4, 5.4)†	(-69.0, -56.2)
One Cache			(-68.0, -55.2)

† ⇒ Not significant

$$s_{\alpha_i - \alpha_j} = s_e \sqrt{\frac{\sum_{j=1}^{a} h_j^2}{b}} = 5.44\sqrt{\frac{1^2 + 1^2 + 0^2}{5}}$$

Outros aspectos

Modelos aditivos X modelos multiplicativos

ver exemplos no livro

Observações faltantes: generalização das fórmulas quando não tem resultados de experimentos para todas as ab configurações

Projeto Fatorial Completo com Dois Fatores Com Replicação

- Dois fatores A e B com a e b níveis respectivamente
  - Fatores são categóricos ou tratados como tal
  - Cada experimento é executado múltiplas vezes
  - Permite separar interação de erros experimentais

$$y_{ijk} = \mu + \alpha_j + \beta_i + \gamma_{ij} + e_{ijk}$$

Efeitos dos fatores e erros são aditivos

$$\sum_{j=1}^{a} \alpha_j = 0 \qquad \sum_{i=1}^{b} \beta_i = 0$$

$$\sum_{j=1}^{a} \gamma_{1j} = \sum_{j=1}^{a} \gamma_{2j} = \cdots = \sum_{j=1}^{a} \gamma_{bj} = 0$$

$$\sum_{i=1}^{b} \gamma_{i1} = \sum_{i=1}^{b} \gamma_{i2} = \cdots = \sum_{i=1}^{b} \gamma_{ia} = 0$$

$$\sum_{k=1}^{r} e_{ijk} = 0 \quad \forall i, j$$

Cálculo dos Efeitos

$$y_{ijk} = \mu + \alpha_j + \beta_i + \gamma_{ij} + e_{ijk}$$

$$\overline{y_{ij.}} = \mu + \alpha_j + \beta_i + \gamma_{ij}$$

$$\overline{y_{i..}} = \mu + \beta_i \qquad \beta_i = \overline{y_{i..}} - \mu$$

$$\overline{y_{.j.}} = \mu + \alpha_j \longrightarrow \alpha_j = \overline{y_{.j.}} - \mu$$

$$\overline{y_{...}} = \mu \qquad \gamma_{ij} = \overline{y_{ij.}} - (\mu + \alpha_j + \beta_i)$$

$$\gamma_{ij} = \overline{y_{ij.}} - (\mu + \overline{y_{.j.}} - \mu + \overline{y_{i..}} - \mu)$$

$$\gamma_{ij} = \overline{y_{ij.}} - \overline{y_{.j.}} - \overline{y_{i..}} + \overline{y_{...}}$$

Exemplo: Tamanho de código em 5 cargas e 4 computadores

Workloads	Processors W	X	Y	Z
I	7006	12042	29061	9903
	6593	11794	27045	9206
	7302	13074	30057	10035
J	3207	5123	8960	4153
	2883	5632	8064	4257
	3523	4608	9677	4065
K	4707	9407	19740	7089
	4935	8933	19345	6982
	4465	9964	21122	6678
L	5107	5613	22340	5356
	5508	5947	23102	5734
	4743	5161	21446	4965
W	6807	12243	28560	9803
	6392	11995	26846	9306
	7208	12974	30559	10233

Exemplo: Tamanho de código em 5 cargas e 4 computadores

	Processors			
Workloads	W	X	Y	Z
I	3.8455	4.0807	4.4633	3.9958
	3.8191	4.0717	4.4321	3.9641
	3.8634	4.1164	4.4779	4.0015
J	3.5061	3.7095	3.9523	3.6184
	3.4598	3.7507	3.9066	3.6291
	3.5469	3.6635	3.9857	3.6091
K	3.6727	3.9735	4.2953	3.8506
	3.6933	3.9510	4.2866	3.8440
	3.6498	3.9984	4.3247	3.8246
L	3.7082	3.7492	4.3491	3.7288
	3.7410	3.7743	4.3636	3.7585
	3.6761	3.7127	4.3313	3.6959
M	3.8330	4.0879	4.4558	3.9914
	3.8056	4.0790	4.4289	3.9688
	3.8578	4.1131	4.4851	4.0100

Modelo Multiplicativo: Transformação logarítmica

Exemplo: Tamanho de código em 5 cargas e 4 computadores

Cálculo das médias das observações por célula

Workloads	Processors W	X	Y	Z	Row Sum	Row Mean	Row Effect
I	3.8427	4.0896	4.4578	3.9871	16.3772	4.0943	0.1520
J	3.5043	3.7079	3.9482	3.6188	14.7792	3.6948	-0.2475
K	3.6720	3.9743	4.3022	3.8397	15.7882	3.9470	0.0047
L	3.7084	3.7454	4.3480	3.7277	15.5295	3.8824	-0.0599
M	3.8321	4.0933	4.4566	3.9900	16.3720	4.0930	0.1507
Col Sum	18.5594	19.6105	21.5128	19.1635	78.8463		
Col Mean	3.7119	3.9221	4.3026	3.8327		3.9423	
Col effect	-0.2304	-0.0202	0.3603	-0.1096			

Modelo multiplicativo: $\hat{y}_{ij} = 10^{\mu} 10^{\alpha_j} 10^{\beta_i} 10^{\gamma_{ij}}$

μ=3.9423: tamanho médio = $10^{3.9423}$=8710

Processador W requer em média um fator de $10^{-0.2304}$ a menos

Carga I requer em média um fator de $10^{0.1520}$ a mais

Exemplo: Tamanho de código em 5 cargas e 4 computadores

$$\overline{y_{ij.}} = \mu + \alpha_j + \beta_i + \gamma_{ij}$$

Cálculo dos efeitos das interações:

$$\gamma_{ij} = \overline{y_{ij.}} - (\mu + \alpha_j + \beta_i)$$

Workloads	W	X	Y	Z
I	-0.0212	0.0155	0.0032	0.0024
J	0.0399	0.0333	-0.1069	0.0337
K	-0.0447	0.0475	-0.0051	0.0023
L	0.0564	-0.1168	0.1054	-0.0450
M	-0.0305	0.0205	0.0033	0.0066

Soma das linhas/colunas tem que dar 0

Exemplo: Tamanho de código em 5 cargas e 4 computadores

Cálculo dos efeitos das interações:

$$\overline{y_{ij.}} = \mu + \alpha_j + \beta_i + \gamma_{ij}$$

$$\gamma_{ij} = \overline{y_{ij.}} - (\mu + \alpha_j + \beta_i)$$

Workloads	W	X	Y	Z
I	-0.0212	0.0155	0.0032	0.0024
J	0.0399	0.0333	-0.1069	0.0337
K	-0.0447	0.0475	-0.0051	0.0023
L	0.0564	-0.1168	0.1054	-0.0450
M	-0.0305	0.0205	0.0033	0.0066

Interpretação: carga I no processador W requer um fator de $10^{(-0.0212)}$ a menos que a média das cargas no processador W. Mesma coisa quando comparado à média dos tamanhos da carga I em todos os processadores

la

Alocação de Variação

$$\hat{y}_{ij} = \mu + \alpha_j + \beta_i + \gamma_{ij} = \bar{y}_{ij.} \qquad e_{ijk} = y_{ijk} - \bar{y}_{ij.}$$

$$\begin{array}{ccccccccccc} \sum_{ijk} y_{ijk}^2 & = & abr\mu^2 & + & br\sum_j \alpha_j^2 & + & ar\sum_i \beta_i^2 & + & r\sum_{ij} \gamma_{ij}^2 & + & \sum_{ijk} e_{ijk}^2 \\ \mathrm{SSY} & = & \mathrm{SS0} & + & \mathrm{SSA} & + & \mathrm{SSB} & + & \mathrm{SSAB} & + & \mathrm{SSE} \end{array}$$

$$\begin{array}{ccccccccccccc} \mathrm{SST} & = & \mathrm{SSY} & - & \mathrm{SS0} & = & \mathrm{SSA} & + & \mathrm{SSB} & + & \mathrm{SSAB} & + & \mathrm{SSE} \\ 4.44 & = & 936.95 & - & 932.51 & = & 2.93 & + & 1.33 & + & 0.15 & + & 0.03 \\ 100\% & = & & & & = & 65.96\% & + & 29.9\% & + & 3.48\% & + & 0.66\% \end{array}$$

ANOVA

Graus de liberdade:

$$
\begin{array}{ccccccccccc}
\text{SSY} & = & \text{SS0} & + & \text{SSA} & + & \text{SSB} & + & \text{SSAB} & + & \text{SSE} \\
abr & = & 1 & + & (a-1) & + & (b-1) & + & (a-1)(b-1) & + & ab(r-1)
\end{array}
$$

- $$
\begin{aligned}
\frac{\text{MSA}}{\text{MSE}} &\sim F[a-1, ab(r-1)] \\
\frac{\text{MSB}}{\text{MSE}} &\sim F[b-1, ab(r-1)] \\
\frac{\text{MSAB}}{\text{MSE}} &\sim F[(a-1)(b-1), ab(r-1)]
\end{aligned}
$$

Tabela ANOVA

Compo-nent	Sum of Squares	%Variation	DF	Mean Square	F-Comp.	F-Table
y	$\mathrm{SSY} = \sum y_{ij}^2$		abr			
$\bar{y}_{\ldots}$	$\mathrm{SS0} = abr\mu^2$		1			
$y - \bar{y}_{\ldots}$	$\mathrm{SST} = \mathrm{SSY} - \mathrm{SS0}$	100	$abr - 1$			
A	$\mathrm{SSA} = br\Sigma\alpha_j^2$	$100\left(\frac{\mathrm{SSA}}{\mathrm{SST}}\right)$	$a - 1$	$\mathrm{MSA} = \frac{\mathrm{SSA}}{a-1}$	$\frac{\mathrm{MSA}}{\mathrm{MSE}}$	$F_{[1-\alpha; a-1, ab(r-1)]}$
B	$\mathrm{SSB} = ar\Sigma\beta_i^2$	$100\left(\frac{\mathrm{SSB}}{\mathrm{SST}}\right)$	$b - 1$	$\mathrm{MSB} = \frac{\mathrm{SSB}}{b-1}$	$\frac{\mathrm{MSB}}{\mathrm{MSE}}$	$F_{[1-\alpha; b-1, ab(r-1)]}$
AB	$\mathrm{SSAB} = r\Sigma\gamma_{ij}^2$	$100\left(\frac{\mathrm{SSAB}}{\mathrm{SST}}\right)$	$(a-1)(b-1)$	$\mathrm{MSAB} = \frac{\mathrm{SSAB}}{(a-1)(b-1)}$	$\frac{\mathrm{MSA}}{\mathrm{MSE}}$	$F_{[1-\alpha,(a-1)(b-1), ab(r-1)]}$
e	$\mathrm{SSE} = \mathrm{SST} - (\mathrm{SSA} + \mathrm{SSB} + \mathrm{SSAB})$	$100\left(\frac{\mathrm{SSE}}{\mathrm{SST}}\right)$	$ab(r-1)$	$\mathrm{MSE} = \frac{\mathrm{SSE}}{ab(r-1)}$		

Tabela ANOVA: exemplo

Component	Sum of Squares	%Variation	DF	Mean Square	F-Comp.	F-Table
y	936.95					
$\bar{y}_{...}$	932.51					
$y - \bar{y}_{...}$	4.44	100.00%	59			
Processors	2.93	65.96%	3	0.9765	1340.01	2.23
Workloads	1.33	29.90%	4	0.3320	455.65	2.09
Interactions	0.15	3.48%	12	0.0129	17.70	1.71
Errors	0.03	0.66%	40	0.0007		

$$s_e = \sqrt{\text{MSE}} = \sqrt{0.0008} = 0.03$$

Os dois fatores e a interação são significativos com 90% de confiança

Intervalos de Confiança

Parameter	Estimate	Variance
μ	$\bar{y}_{...}$	s_e^2/abr
α_j	$\bar{y}_{i..}$-$\bar{y}_{...}$	$s_e^2(a-1)/abr$
β_i	$\bar{y}_{.j.}$-$\bar{y}_{...}$	$s_e^2(b-1)/abr$
γ_{ij}	$\bar{y}_{ij.}$-$\bar{y}_{i..}$-$\bar{y}_{.j.}$+$\bar{y}_{...}$	$s_e^2(a-1)(b-1)/abr$
$\Sigma\ h_j\ \alpha_j,\ \Sigma\ h_j{=}0$	$\Sigma\ h_j\ \bar{y}_{.j.}$	$\Sigma\ h_j^2 s_e^2/br$
$\Sigma\ h_i\ \beta_i,\ \Sigma\ h_i{=}0$	$\Sigma\ h_i\ \bar{y}_{i..}$	$\Sigma\ h_i^2 s_e^2/ar$
s_e^2	$\Sigma\ e_{ijk}^2/\{ab(r-1)\}$	

Degrees of freedom for errors = ab(r-1)

Intervalos de Confiança: exemplo

$$s_{\alpha_j} = s_e\sqrt{\frac{a-1}{abr}} = 0.03\sqrt{\frac{4-1}{4\times5\times3}} = 0.0060$$

Graus de liberdade: ab(r-1) = 40 (use distr. Normal)
Para 90% de confiança = $z_{0.95} = 1.645$

$$\begin{aligned}\alpha_1 \mp t\, s_{\alpha_1} &= -0.2304 \mp 1.645 \times 0.0060\\ &= -0.2304 \mp \quad 0.00987\\ &= (-0.2406, -0.2203)\end{aligned}$$

Significativo!

Intervalos de Confiança: exemplo

Parameter	Mean Effect	Std. Dev.	Confidence Interval
μ	3.9423	0.0035	(3.9364, 3.9482)
Processors			
W	-0.2304	0.0060	(-0.2406, -0.2203)
X	-0.0202	0.0060	(-0.0304, -0.0100)
Y	0.3603	0.0060	(0.3501, 0.3704)
Z	-0.1096	0.0060	(-0.1198, -0.0995)
Workloads			
I	0.1520	0.0070	(0.1402, 0.1637)
J	-0.2475	0.0070	(-0.2592, -0.2358)
K	0.0047	0.0070	(-0.0070, 0.0165)†
L	-0.0599	0.0070	(-0.0717, -0.0482)
M	0.1507	0.0070	(0.1390, 0.1624)

† ⇒ Not significant

Intervalos de Confiança: exemplo

Workloads	W	X	Y	Z
I	(-0.0415, -0.0009)	(-0.0048, 0.0358)†	(-0.0171, 0.0236)†	(-0.0179, 0.0228)†
J	(0.0196, 0.0602)	(0.0130, 0.0536)	(-0.1272, -0.0865)	(0.0133, 0.0540)
K	(-0.0650, -0.0243)	(0.0271, 0.0678)	(-0.0254, 0.0152)†	(-0.0180, 0.0226)†
L	(0.0361, 0.0768)	(-0.1371, -0.0964)	(0.0850, 0.1257)	(-0.0654, -0.0247)
M	(-0.0508, -0.0101)	(0.0002, 0.0408)	(-0.0170, 0.0236)†	(-0.0137, 0.0270)†

† ⇒ Not significant

Mesmos testes visuais!!!