Métodos Quantitativos para Ciência da Computação Experimental

Projeto de Experimentos

Jussara Almeida
DCC-UFMG
2017

Apresentacao derivada dos slides originais de Virgilio Almeida

Projeto de Experimentos

- Introdução: cap. 16 do texto (Jain)
- Projetos 2^k fatorial: caps. 17, 18 e 19 do texto (Jain)
- Experimentos de um-fator: cap. 20 do texto (Jain)

Terminologia do Projeto de Experimentos

- *Variável resposta*: representa o valor obtido, que é medido de acordo com as variações dos dados de entrada.
  - Exemplo: tempo de resposta, índice de precisão, utilização, outros exemplos???

- *Fatores*: as variáveis de entrada de um experimento que podem ser controladas pelo “experimentador”.
  - Exemplo: tamanho do cache, tamanho dos arquivos, tempo de seek, latência da rede, etc

- *Níveis*: os níveis de um fator são os valores específicos que podem ser atribuídos ao fator. Podem ser contínuos (ex.: tempo de seek), discretos (# de servidores) ou podem ser categóricos, como o tipo de um processador ou a classe de um certo algoritmo. Tambem chamados de *treatments*

Terminologia do Projeto de Experimentos

- *Replicação*: replicar um experimento significa re-executá-lo completamente com todos os mesmos niveis de entrada. Desde que as medidas da variável resposta são sujeitas a variações aleatórias, as replicações de um experimento são usadas para determinar o impacto do erro experimental na variável resposta.

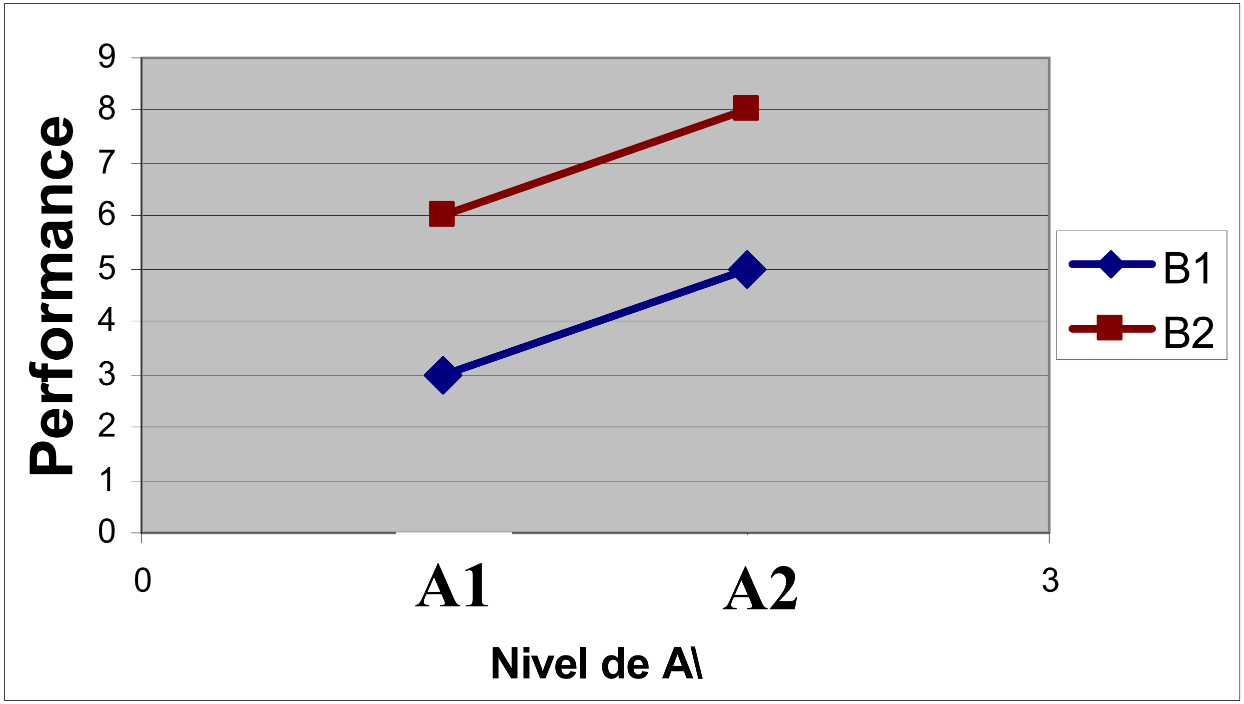
- *Interação*: uma interação entre fatores ocorre quando o efeito de um fator depende do nível de outro fator.
  - Efeito da memória na atividade de paginação.
  - Outros exemplos???

Exemplo

- Fatores que nao tem interacao:

Resposta:

	Fator A	
Fator B	A1	A2
B1	3	5
B2	6	8



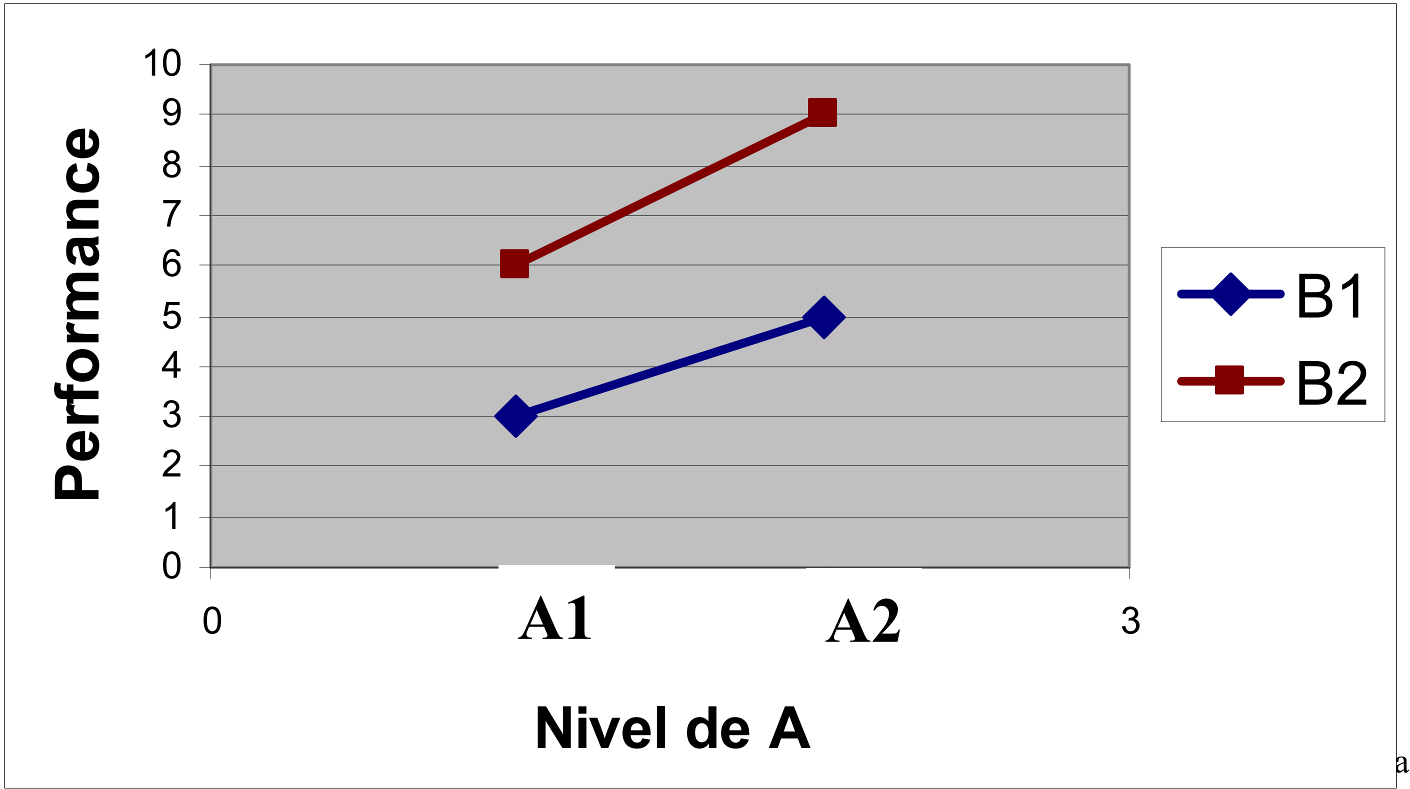
Apresentacao derivada dos slides originais de Virgilio Almeida

Exemplo

- Fatores que tem interacao:

Resposta:

	Fator A	
Fator B	A1	A2
B1	3	5
B2	6	9



Introdução ao Projeto de Experimentos: perguntas básicas

- Voce conhece as métricas?
- Voce conhece os fatores?
- Voce conhece os níveis?
- Voce tem conhecimento de como instrumentar o sistema e elaborar as cargas de teste?
- Voce sabe descrever o que fazer para realizar experimentos que suportem ou refutem as hipóteses de sua pesquisa?

Objetivos no Projeto de Experimentos

- Obter a maior quantidade de informação

- Reduzir o trabalho/esforço de experimentação
  - Tipicamente significa o menor número de experimentos

- Realizar número muito grande de experimentos não é bom (gasta-se tempo e recursos), principalmente se voce for aquele responsável pela execução dos mesmos.

- Experimentos bem projetados são mais fáceis de serem analisados.

Replicações Experimentais

- O sistema em estudo executará com vários níveis de diferentes fatores, potencialmente com diferentes cargas.
- Uma execução com um conjunto particular de níveis e dados de entrada é uma replicação.
- Em geral, é necessário realizar múltiplas replicações com um único conjunto de níveis e dados de entrada, por razões de verificação e validação estatística.

A Interação dos Fatores

- Alguns fatores tem efeitos completamente independentes um do outro.
  - Exemplo: Duplique o nível de um fator e obterá metade da resposta, independente dos outros fatores.
- Mas os efeitos de alguns fatores dependem dos valores de outros fatores
  - *Fatores interatuantes*
- A presença de fatores interatuantes complica o projeto experimental.

Problema Básico ao Projetar Experimentos

- Um determinado número de fatores foi escolhido
- Os fatores podem ou não interagir
- Como se pode projetar um experimento que captura os intervalos completos de variação dos níveis?
  - Com a menor quantidade de trabalho possível
- Qual a combinação ou combinações de níveis de fatores deseja-se medir?

Erros Comuns na Experimentação

- Ignorar o erro experimental
  - A variacao devido a um fator deve ser comparada com a variacao devido aos erros experimentais antes de se tomar uma decisao sobre o fator (ele tem impacto significativo?)
- Existência de parâmetros não controlados (nao sao fatores)
  - Somente o impacto de fatores e avaliado
- Não isolamento dos efeitos de diferentes fatores
  - Variacao de varios fatores simultaneamente
- Projetos de experimentos com um fator-de-cada-vez
  - Muito caro: nao necessariamente mais informativo
- Ignorar as interações entre os fatores
- Projetos que requerem um número excessivo de experimentos
  - Melhor considerar um subconjunto dos fatores/niveis primeiro e depois ir acrescentando fatores/niveis aos poucos.

Tipos de Projetos de Experimentos

- Projetos simples
- Projetos com fatorial completo
- Projetos com fatorial fracionado

Projeto Experimental

Fator 1

Fator 2

$(\ell_{1,0}, \ell_{1,1}, \ldots, \ell_{1,n1-1}) \times (\ell_{2,0}, \ell_{2,1}, \ldots, \ell_{2,n2-1}) \times \ldots$

$\times (\ell_{k,0}, \ell_{k,1}, \ldots, \ell_{k,nk-1})$

Fator k

k diferentes fatores, onde fator i tem n_i níveis

r replicações

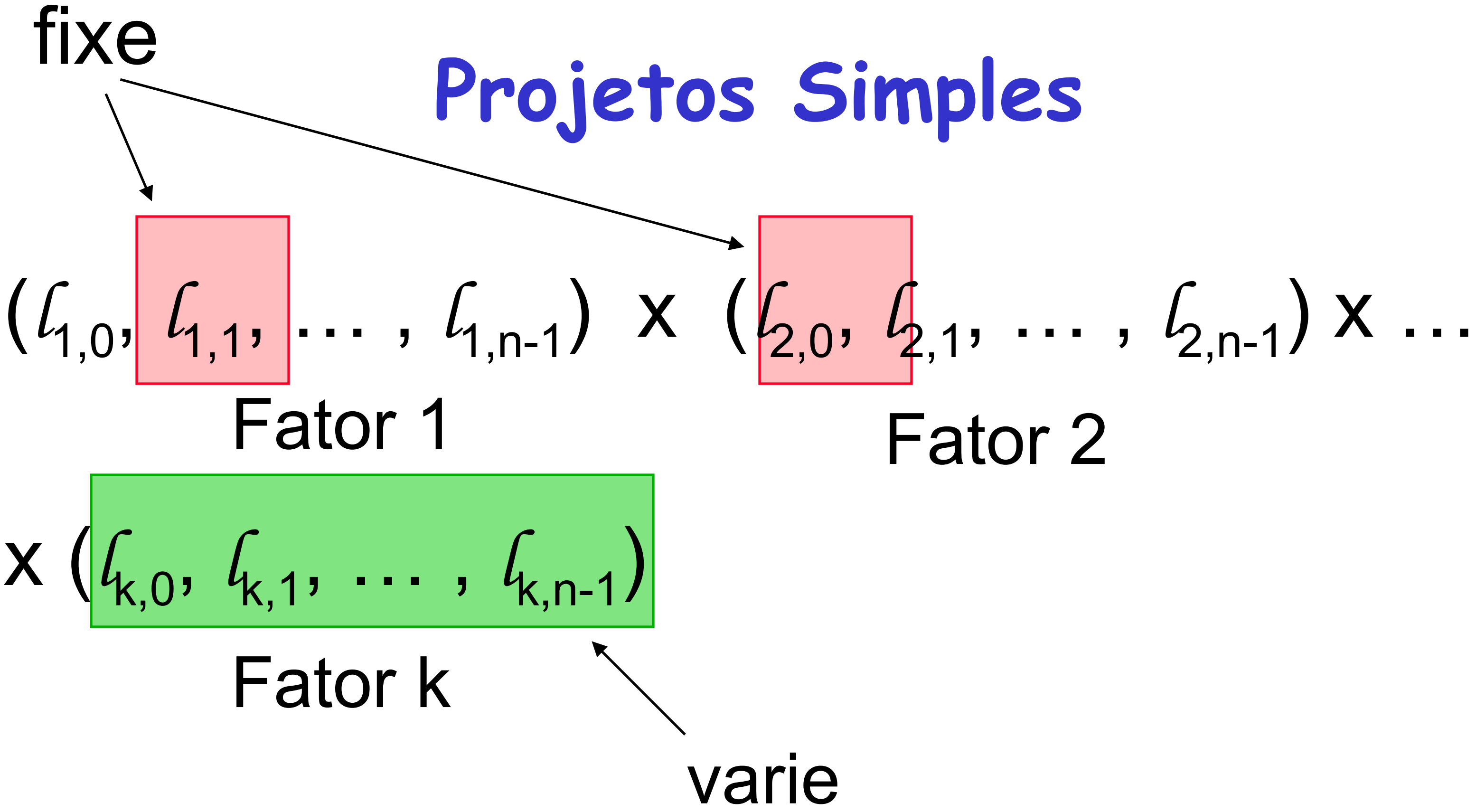
Projetos Simples

- Varie um fator de cada vez
- Para *k* fatores com o i^{esimo} fator tendo n_i níveis

$$n = 1 + \sum_{i=1}^{k} (n_i - 1)$$

- Assume que os fatores não interagem
- Usualmente requer mais esforço que se pensa
- Tente evitar esse enfoque de experimentação

Projetos Simples



fixe

$(\ell_{1,0}, \ell_{1,1}, \ldots, \ell_{1,n-1}) \times (\ell_{2,0}, \ell_{2,1}, \ldots, \ell_{2,n-1}) \times \ldots$

Fator 1

Fator 2

$\times (\ell_{k,0}, \ell_{k,1}, \ldots, \ell_{k,n-1})$

Fator k

varie

Projetos Simples

$$(\ell_{1,0}, \underset{\text{Fator 1}}{\boxed{\ell_{1,1},}} \dots , \ell_{1,n-1}) \ \times \ (\underset{\text{Fator 2}}{\boxed{\ell_{2,0}, \ell_{2,1}, \dots , \ell_{2,n-1})}} \times \dots$$

$$\times \ (\ell_{k,0}, \underset{\text{Fator k}}{\boxed{\ell_{k,1},}} \dots , \ell_{k,n-1})$$

Projetos com Fatorial Completo

- Para *k* fatores com o i^{esimo} fator tendo n_i níveis -

$$n = \prod_{i=1}^{k} n_i$$

- Teste cada combinação possível dos níveis dos fatores.
- Capture a informação completa sobre a interação de fatores
- É no entanto, um trabalho ENORME!!!
  - Principalmente se valores de n_i forem grandes

Apresentacao derivada dos slides originais de Virgilio Almeida

Projetos com Fatorial Completo

$$(\underbrace{\ell_{1,0}, \ell_{1,1}, \ldots, \ell_{1,n-1}}_{\text{Fator 1}}) \times (\underbrace{\ell_{2,0}, \ell_{2,1}, \ldots, \ell_{2,n-1}}_{\text{Fator 2}}) \times \ldots$$

$$\times (\underbrace{\ell_{k,0}, \ell_{k,1}, \ldots, \ell_{k,n-1}}_{\text{Fator k}})$$

Reduzindo o trabalho em Projetos com Fatorial Completo

- Reduza o número de níveis por fator
  - Geralmente uma boa opção
  - Especialmente quando se sabe quais fatores são mais importantes
  - Para os fatores mais relevantes, use mais níveis
- Reduza o número de fatores
  - Simplifique o modelo experimental
  - Mas não retire fatores “relevantes”
- Use projetos de fatorial fracionado

Projetos com Fatorial Fracionado

- Faça a medição somente de uma combinação de niveis de fatores.
- O projeto deve ser cuidadosamente projetado para capturar melhor qualquer interação que possivelmente exista.
- Menos trabalho, porém com mais chance de imprecisões nos resultados.
  - Compromisso
- Pode ser útil quando se sabe *a priori* que alguns fatores não interagem.

Projetos Fatoriais 2^k

- Usados para determinar os efeitos de *k* fatores
  - Cada um com duas alternativas ou níveis

- Em geral, são usados de maneira preliminar, antes de estudos mais detalhados
  - Cada fator medido é representado por seu nível máximo e por seu nível mínimo.
  - Pode oferecer algum “insight” sobre as interações entre os vários fatores.

Efeitos Unidirecionais

- Efeitos que somente aumentam à medida que o nível de um fator também aumenta
  - Ou vice-versa
- Se essa característica é conhecida a priori, um projeto fatorial 2^k nos niveis máximo e mínimo pode ser útil.
- Demonstra-se quando um fator tem efeito significativo no experimento.

Projetos Fatoriais 2^2

- Dois fatores, com dois niveis cada
- Tipo mais simples de um projeto fatorial de experimentos
- Os conceitos desenvolvidos podem ajudar o entendimento dos problemas de projetos 2^k $(k > 2)$
- Exemplo simples, com finalidade pedagógica

Exemplo de um Projeto Fatorial 2^2

- Uma arquitetura de máquina de busca, composta por *N* servidores;
- Pode-se usar vários esquemas de distribuição ou escalonamento de *queries* para os servidores, por exemplo, *round-robin, gang, random, priority,* etc
- O objetivo é completar os *queries* no menor tempo possível.
- No exemplo, a métrica usada é o tempo de execução da *query* em microsegundos.

Fatores e Níveis do Exemplo

- Primeiro fator – número de servidores usados na máquina de busca experimental:
  - Varia entre 8 e 64
- Segundo fator – baseado em outros estudos, usa-se dois extremos de políticas de escalonamento: aleatorio e “round-robin”.
  - Sistema de arquivos local e global na arquitetura, que permite a distribuição de *query* para qualquer servidor.
- Outros fatores existem, mas neste exemplo, vamos ignorá-los.

Definindo as Variáveis para um Exemplo de Projeto 2² Factorial

$$x_A = \begin{cases} -1 & \text{se 8 servidores} \\ 1 & \text{se 64 servidores} \end{cases}$$

$$x_B = \begin{cases} -1 & \text{se escalonamento randomico} \\ 1 & \text{se escalonamento round robin (RR)} \end{cases}$$

Dados Amostrais para o Exemplo

- Execução única de uma carga benchmark de *queries* nas duas configurações resultou nos seguintes tempos de execucao:

	8 Serv. (-1)	64 Serv. (+1)
Rand. (-1)	820	217
RR (+1)	776	197

Modelo de Regressão Não Linear para o Exemplo

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

$820 = q_0 - q_A - q_B + q_{AB}$

$217 = q_0 + q_A - q_B - q_{AB}$

$776 = q_0 - q_A + q_B - q_{AB}$

$197 = q_0 + q_A + q_B + q_{AB}$

A = numero de servidores

B = escalonamento

Modelo de Regressão

- 4 equações e 4 variáveis
- Outra representação – tabela

Experimento	I	A	B	y
1	1	-1	-1	y_1=820
2	1	1	-1	y_2=217
3	1	-1	1	y_3=776
4	1	1	1	y_4=197

Solucionando para os q_i´s

$$q_0 = \frac{1}{4}(y_1 + y_2 + y_3 + y_4)$$

$$q_A = \frac{1}{4}(-y_1 + y_2 - y_3 + y_4)$$

$$q_B = \frac{1}{4}(-y_1 - y_2 + y_3 + y_4)$$

$$q_{AB} = \frac{1}{4}(y_1 - y_2 - y_3 + y_4)$$

Solução das Equações

q_0 = 1/4(820 + 217 + 776 + 197) = 502.5
q_A = 1/4(-820 + 217 - 776 + 197) = -295.5
q_B = 1/4(-820 - 217 + 776 + 197) = -16
q_{AB} = 1/4(820 - 217 - 776 + 197) = 6

Assim:

$y = 502.5 - 295.5x_A - 16x_B + 6x_Ax_B$

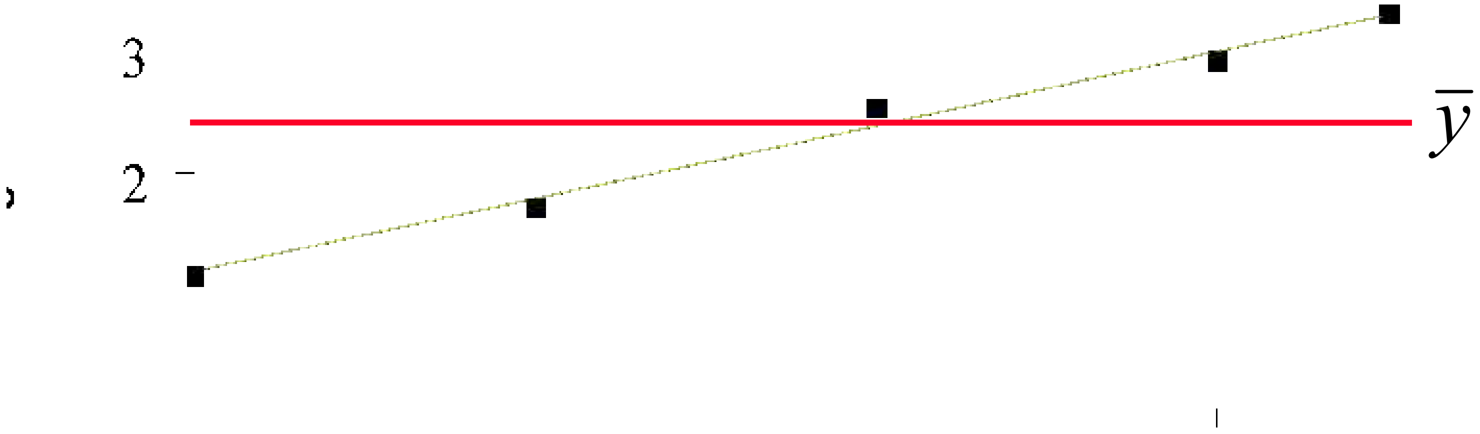
q_0 = 502.5 = tempo de execucao medio

A = numero de servidores tem maior impacto no tempo de resposta, e faz uma diferenca de +- 295,5

Alocando a Variação

- Sem regressão, a melhor estimativa de *y* é $\bar{y}$
- Valores observados de *y* diferem de $\bar{y}$ aumentando os erros (variação)
- Regressão provê uma melhor estimativa, mas ainda existem erros
- Nós podemos avaliar a qualidade da regressão pela alocação das fontes de erros.
  - Quão bom este modelo é?

Gráfico dos Parametros de Estimativa
exemplo: regressão e a média



Alocação de Variação para o Modelo 2^2

- Calcule a variância amostral de y

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

- Numerador é o SST (variação total) $SST = \sum_{i=1}^{2^2} \left(y_i - \overline{y}\right)^2$

  (nao confundir variacao com variancia)

  Outra formula para SST e: SST = $2^2 q_A^2 + 2^2 q_B^2 + 2^2 q_{AB}^2$

  (derivacao no livro)

- Podemos usar isso para entender as causas da variação de y

Termos no SST

- $2^2 q_A^2$ é parte da variação explicada pelo efeito de A (SSA)
- $2^2 q_B^2$ é parte da variação explicada pelo efeito de B (SSB)
- $2^2 q_{AB}^2$ é parte da variação explicada pelo efeito da interação de A e B (SSAB)

$$SST = SSA + SSB + SSAB$$

Variações no Exemplo

- SST = 350449
- SSA = 349281
- SSB = 1024
- SSAB = 144
- Pode-se agora calcular e entender a fração da variação total causada por cada efeito.

Frações de Variação no Exemplo

- Fração explicada por A: 99.67%
- Fração explicada por B: 0.29%
- Fração explicada pela interação de A e B: 0.04%

- Assim, quase toda variação vem do número de servidores da arquitetura e o esquema de escalonamento tem um efeito desprezível na performance da máquina de busca em estudo.

- Se o objetivo é diminuir o tempo de resposta de *queries*, deve-se então concentrar no número de servidores e não no esquema de distribuição-escalonamento (exemplo hipotético!)

Projetos com Fatorial 2^k

- Usado para analisar os efeitos de *k* fatores, cada um com niveis de duas alternativas
- Projetos 2^2 fatorial são um caso especial

Exemplo

- No projeto de um sistema, os tres fatores de maior impacto e que precisam ser estudados sao : tamanho do cache, tamanho da memoria, e se 1 ou 2 processadores serao usados.

Os tres fatores e os seus niveis sao apresentados abaixo:

Fator	Nivel –1	Nivel 1
Tamanho da Memoria A	1 GB	4 GB
Tamanho do Cache B	128 KB	256 KB
# Processadores C	1	2

Exemplo

- O projeto 2^3 e o desempenho medido, em MIPS, e mostrado na tabela abaixo:

	1GB		4GB	
Cache(KB)	1 Proc.	2 Proc.	1 Proc	2 Proc
128	14	46	22	58
256	10	50	34	86

Solucao

A	B	C	Y
-1	-1	-1	14
1	-1	-1	22
-1	1	-1	10
1	1	-1	34
-1	-1	1	46
1	-1	1	58
-1	1	1	50
1	1	1	86

Solucao

I	A	B	C	Y
1	-1	-1	-1	14
1	1	-1	-1	22
1	-1	1	-1	10
1	1	1	-1	34
1	-1	-1	1	46
1	1	-1	1	58
1	-1	1	1	50
1	1	1	1	86

Solucao

I	A	B	C	AB	AC	BC	ABC	Y
1	-1	-1	-1	1	1	1	-1	14
1	1	-1	-1	-1	-1	1	1	22
1	-1	1	-1	-1	1	-1	1	10
1	1	1	-1	1	-1	-1	-1	34
1	-1	-1	1	1	-1	-1	1	46
1	1	-1	1	-1	1	-1	-1	58
1	-1	1	1	-1	-1	1	-1	50
1	1	1	1	1	1	1	1	86

Solucao

I	A	B	C	AB	AC	BC	ABC	Y
1	-1	-1	-1	1	1	1	-1	14
1	1	-1	-1	-1	-1	1	1	22
1	-1	1	-1	-1	1	-1	1	10
1	1	1	-1	1	-1	-1	-1	34
1	-1	-1	1	1	-1	-1	1	46
1	1	-1	1	-1	1	-1	-1	58
1	-1	1	1	-1	-1	1	-1	50
1	1	1	1	1	1	1	1	86
320	80	40	160	40	16	24	9	Total
40	10	5	20	5	2	3	1	Total/8
Media ou q_0	q_A	q_B	q_C	q_{AB}	q_{AC}	q_{BC}	q_{ABC}	

Solucao

I	A	B	C	AB	AC	BC	ABC	Y
1	-1	-1	-1	1	1	1	-1	14
1	1	-1	-1	-1	-1	1	1	22
1	-1	1	-1	-1	1	-1	1	10
1	1	1	-1	1	-1	-1	-1	34
1	-1	-1	1	1	-1	-1	1	46
1	1	-1	1	-1	1	-1	-1	58
1	-1	1	1	-1	-1	1	-1	50
1	1	1	1	1	1	1	1	86
320	80	40	160	40	16	24	9	Total
40	10	5	20	5	2	3	1	Total/8

$$SST = 2^3(q_A^2 + q_B^2 + q_C^2 + q_{AB}^2 + q_{AC}^2 + q_{BC}^2 + q_{ABC}^2)$$

$$= 8(10^2 + 5^2 + 20^2 + 5^2 + 2^2 + 3^2 + 1^2)$$

$$= 800 + 200 + 3200 + 200 + 32 + 72 + 8 = 4512$$

Solucao

- A porcao da variacao explicada por cada fator e suas interacoes sao:
  - A : 800/4512 = 18%
  - B: 200/4512 = 4%
  - C: 3200/4512 = 71%
  - AB: 200/4512 = 4%
  - AC: 32/4512 = 1%
  - BC: 72/4512 = 2%
  - ABC: 8/4512 = 0% -> pode ignorar

Projetos Fatoriais $2^k r$

- Projetos fatoriais 2^k nao permitem estimar os erros experimentais já que nenhum experimento e repetido
- Se cada um dos 2^k forem replicados r vezes, teremos $2^k r$ observacoes
  - Projetos fatoriais $2^k r$
  - Poderemos estimar os erros experimentais
  - Poderemos comparar a % da variacao devido a cada fator ou interacao com a % da variacao devido aos erros experimentais
    - Fator/interação tem impacto significativo?

Projetos Fatoriais 2^2r

- Assume o modelo generico:

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

- Computar os efeitos (coeficientes) de forma similar aos projetos 2^k

Exemplo

- Um sistema foi avaliado considerando 2 fatores (A e B) e variando cada fator em dois niveis. Cada experimento foi repetido 3 vezes e os resultados sao mostrados na tabela abaixo:

I	A	B	AB	y	Media(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

$q_0 = 41$ $q_A = 21.5$ $q_B = 9.5$ $q_{AB} = 5$

Estimando erros experimentais

$$\widehat{y}_i = q_0 + q_A x_{Ai} + q_B x_{Bi} + q_{AB} x_{Ai} x_{Bi}$$

$$e_{ij} = y_{ij} - \widehat{y}_i$$

$$\sum_{i,j} e_{ij} = 0$$

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

Apresentacao derivada dos slides originais de Virgilio Almeida

Exemplo

Calculando os erros experimentais:

I	A	B	AB	y_{ij}	$\hat{y}$	e_{ij}
1	-1	-1	1	(15,18,12)	15	(0,3,-3)
1	1	-1	-1	(45,48,51)	48	(-3,0,3)
1	-1	1	-1	(25,28,19)	24	(1,4,-5)
1	1	1	1	(75,75,81)	77	(-2,-2,4)

$$SSE = 0 + 9 + 9 + 9 + 0 + 9 + 1 + 16 + 25 + 4 + 4 + 16 = 102$$

Alocacao de Variação

Variacao total ou Soma Total dos Quadrados SST e dada por

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

$$= \sum_{i,j} y_{ij}{}^2 - \sum_{i,j} q_0{}^2 = \sum_{i,j} y_{ij}{}^2 - 2^2 r q_0{}^2$$

$$= \sum_{i,j} y_{ij}{}^2 - 2^2 r q_0{}^2$$

$$SST = SSY - SS0$$

Apresentacao derivada dos slides originais de Virgilio Almeida

Alocacao de Variação

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

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

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

$$SSA = 2^2 r q_A^2 \qquad SSB = 2^2 r q_B^2$$

$$SSAB = 2^2 r q_{AB}^2$$

$$SSE = SSY - 2^2 r (q_0^2 + q_A^2 + q_B^2 + q_{AB}^2)$$

Exemplo

- Alocando as variacoes:

$SSY = 15^2 + 18^2 + 12^2 + \ldots\ldots + 81^2 = 27204$

$SS0 = 2^2 r q_0^2 = 4*3*41^2 = 20172$

$SSA = 2^2 r q_A^2 = 4*3*(21.5)^2 = 5547$

$SSB = 2^2 r q_B^2 = 4*3*(9.5)^2 = 1083$

$SSAB = 2^2 r q_{AB}^2 = 4*3*(5)^2 = 300$

$SSE = SSY - 2^2 r(q_0^2 + q_A^2 + q_B^2 + q_{AB}^2)$

$= 27204 – 12(41^2 + 21.5^2 + 9.5^2 + 5^2) = 102$

(igual anterior)

$SST = SSY – SS0 = 27204 – 20172 = 7032$

Exemplo

Fator A explica 5547/7032 = 78.88% da variacao.

Fator B explica 1083/7032 = 15.4%

A interacao AB explica 4.27% da variacao.

Os restantes 1.45% sao inexplicados e sao atribuidos aos erros experimentais

Intervalos de Confianca para os Efeitos

A variancia dos erros ou Mean Square Errors (MSE) pode ser estimada a partir do SSE como a seguir:

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

Graus de liberdade do calculo do SSE iguais a $2^2(r-1)$ pois os r erros correspondentes as replicacoes de cada experimento devem somar 0.

Isto e valido para todos os 2^2 experimentos.

Intervalos de Confianca para os Efeitos

Desvio padrao dos efeitos:

$$s_{q_0} = s_{q_A} = s_{q_B} = s_{q_{AB}} = \frac{s_e}{\sqrt{2^2 r}}$$

Intervalos de Confianca:

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

Apresentacao derivada dos slides originais de Virgilio Almeida

Exemplo

- Calculando os Intervalos de Confianca:

Lembrando que SSE = 102

$$s_e = \sqrt{\frac{SSE}{2^2(r-1)}} = \sqrt{\frac{102}{2^2(3-1)}} = \sqrt{12.75} = 3.57$$

$$s_{q_i} = \frac{s_e}{\sqrt{2^2 r}} = \frac{3.57}{\sqrt{12}} = 1.03$$

Exemplo

- Calculando os Intervalos de Confianca:

  95% quantil da variavel t com

  $2^2(r-1) = 8$ graus de liberdade $= 1.86$

  Intervalos de confianca de 90% : $q_i \pm t s_{q_i}$

  $q_0 : (39.08, 42.91)$ $q_A : (19.58, 23.41)$

  $q_B : (7.58, 11.41)$ $q_{AB} : (3.08, 6.91)$

- Todos os coeficientes sao significativos com 90% de confianca

Intervalos de Confianca para Contrastes de Efeitos

Um contraste e uma combinacao linear cujos coeficientes somam zero.

Queremos calcular a variancia de:

$$\sum h_i q_i \qquad \text{dado} \quad \text{que} \quad \sum \mathrm{h}_i = 0$$

$$s^2_{\sum h_i q_i} = \frac{s_e^2 \sum h_i^2}{2^2 r}$$

Intervalos de Confianca para Contrastes de Efeitos

Um intervalo de confianca para o contraste pode ser calculado usando a variavel t com $1-\alpha/2$ e $2^2(r-1)$ graus de liberdade

Este procedimento pode ser usado para computar ICs para a resposta estimada para configuracoes especificas (niveis dos fatores em valores especificos)

Exemplo

Seja o contraste u = $q_A + q_B - 2q_{AB}$ com coeficientes 0, 1, 1 e –2

(soma dos coeficientes igual a 0)

u = 21.5 + 9.5 – 2*5 = 11

Variancia de u

$$s_u^2 = \frac{s_e^2 \times (0+1+1+4)}{2^2 \times 3} = 6.375$$

$$s_u = \sqrt{6.375} = 2.52$$

$$t_{[0.95,8]} = 1.86$$

$$u \pm ts_u = 11 \pm 1.86 \times 2.52 = (6.31,15.69)$$

Intervalos de Confianca para Estimativas de Respostas

Queremos estimar a media das respostas obtidas em *m* repeticoes (futuras) do mesmo experimento (mesma combinacao de niveis de fatores).

A estimativa da media e:

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

Intervalos de Confianca para Estimativas de Respostas

O desvio padrao da estimativa e

$$s_{\hat{y}_m} = s_e \sqrt{\frac{1}{n_{eff}} + \frac{1}{m}}$$

$$n_{eff} = \frac{\#\ \text{total de observacoes}}{1\ +\ \text{soma dos DFs dos parametros usados em } \hat{y}}$$

$$n_{eff} = \frac{2^2 r}{1+2^k} = \frac{2^2 r}{1+4} \quad \text{onde k e o numero de parametros}$$

$$(q_0, q_A, q_B, q_{AB})$$

Intervalos de Confianca para Estimativas de Respostas

Intervalo de Confianca:

$$\hat{y} \pm t_{[1-\alpha/2, 2^2(r-1)]} s_{\hat{y}_m}$$

Estimativa da resposta para uma unica execucao (m=1)

$$s_{\hat{y}_m} = s_e \sqrt{\frac{5}{2^2 r} + 1}$$

Estimativa da resposta para a media da populacao (m=∞)

$$s_{\hat{y}_m} = s_e \sqrt{\frac{5}{2^2 r}}$$

Exemplo

No exemplo, vamos computar o IC para a resposta media com $x_A = -1$ e $x_B = -1$

Quatro ICs podem ser calculados:

1) Estimativa da resposta media para UM experimento futuro de confirmacao

$$\hat{y}_1 = q_0 - q_A - q_B + q_{AB} = 41 - 21.5 - 9.5 + 5 = 15$$

$$s_{\hat{y}_1} = s_e\sqrt{\frac{5}{2^2 r} + 1} = 3.57\sqrt{\frac{5}{12} + 1} = 4.25$$

$$t_{[0.95,8]} = 1.86$$

$$90\%\ \mathrm{IC}: 15 \pm 1.86 \times 4.25 = (8.09, 22.91)$$

Exemplo

2) Estimativa da resposta media para 5 experimentos no futuro:

$$\hat{y}_1 = q_0 - q_A - q_B + q_{AB} = 41 - 21.5 - 9.5 + 5 = 15$$

$$s_{\hat{y}_1} = s_e \sqrt{\frac{5}{2^2 r} + \frac{1}{m}} = 3.57 \sqrt{\frac{5}{12} + \frac{1}{5}} = 2.80$$

$$t_{[0.95,8]} = 1.86$$

$$90\%\ \mathrm{IC}: 15 \pm 1.86 \times 2.80 = (9.79, 20.29)$$

Exemplo

3) Estimativa da resposta media para um grande numero de experimentos no futuro:

$$\hat{y}_1 = q_0 - q_A - q_B + q_{AB} = 41 - 21.5 - 9.5 + 5 = 15$$

$$s_{\hat{y}_1} = s_e \sqrt{\frac{5}{2^2 r}} = 3.57 \sqrt{\frac{5}{12}} = 2.30$$

$$t_{[0.95,8]} = 1.86$$

$$90\%\ \mathrm{IC}: 15 \pm 1.86 \times 2.30 = (10.72, 19.28)$$

Exemplo

4) Resposta media atual (nao e previsao): queremos um IC para um contraste

$$\text{Contraste} \quad \hat{y}_1 = q_0 - q_A - q_B + q_{AB} = 41 - 21.5 - 9.5 + 5 = 15$$

$$s_{\hat{y}_1} = \sqrt{\frac{s_e^2 \sum h_i^2}{2^2 r}} = \sqrt{\frac{12.75(1+1+1+1)}{12}} = 2.06$$

$$t_{[0.95,8]} = 1.86$$

$$90\%\ \mathrm{IC}: 15 \pm 1.86 \times 2.06 = (11.17, 18.83)$$

As estimativas futuras sempre tem CIs mais largos (ou maior variancia por causa dos erros nos experimentos futuros, que devem ser adicionados aos correntes

Premissas

As seguintes premissas foram feitas nas derivacoes anteriores:

- Erros estatisticamente independentes
- Erros sao aditivos
- Erros sao normalmente distribuidos
- Erros tem desvio padrao constante
- Efeitos dos fatores sao aditivos

- Necessidade de testes visuais
  (os mesmos aplicados para regressao linear,
  veremos isto depois)

Modelos Multiplicativos para Experimentos 2^2r

- O modelo usado anteriormente assume que os efeitos, suas interacoes e os erros experimentais sao aditivos

- Analista precisa validar que efeitos sao mesmo aditivos

- Contra-exemplo: desempenho de processadores para diferentes cargas

Exemplo

- Seja y_{ij} o tempo necessario para executar uma carga de w_j instrucoes em um processador que executa $1/v_i$ instrucoes por segundo:

  $y_{ij} = v_i w_j$

- Os efeitos dos dois fatores nao sao aditivos, mas sim multiplicativos: transformacao

  $$\log(y_{ij}) = \log(v_i) + \log(w_j)$$

- Modelo: $y_{ij}' = q_0 + q_A x_A + q_B x_B + q_{AB} x_A x_B + e_{ij}$

onde $y_{ij}' = \log(y_{ij})$

Exemplo

- Considere o caso de dois processadores A1 e A2 nos quais foram testados dois benchmarks B1 e B2. Cada experimento foi repetido 3 vezes e os tempos de execucao medidos (segundos) sao dados na tabela abaixo. Uma analise direta usando o modelo aditivo tambem e mostrado

I	A	B	AB	y	Media(y)
1	-1	-1	1	(85.10, 79.50, 147.90)	104.17
1	1	-1	-1	(0.891, 1.047, 1.072)	1.003
1	-1	1	-1	(0.955, 0.933, 1.122)	1.003
1	1	1	1	(0.0148, 0.0126, 0.0118)	0.013
106.19	–104.15	–104.15	102.17		Total
26.55	-26.04	-26.04	25.54		Total/4

- Conclusao: ha uma grande interacao entre processadores e benchmarks, o que leva a concluir que a selecao dos processadores deveria depender do benchmark (???)

Exemplo

- Pontos que levam ao questionamento desta analise
  - Mais importante e a explicacao fisica: sabemos que os efeitos de processador e benchmark nao se somam, mas multiplicam
    - Modelo errado desde o inicio
  - O intervalo de valores medidos e muito grande (0.0118 a 147.90). A razao entre max e min 147.90/0.0118 = 12534 e muito grande. Tirar a media aritmetica de valores tao dispersos nao e apropriado. Este intervalo grande pede uma transformacao logaritmica

Exemplo

- Pontos que levam ao questionamento desta analise
  - Um grafico dos residuos X respostas estimadas mostra que:
    - Residuos (erros) nao sao pequenos se comparados com respostas
    - Variabilidade dos residuos aumenta com resposta : transformacao log
  - Um grafico quantil-quantil para a distribuicao normal mostra que residuos extremos nao seguem a mesma reta dos valores intermediarios: residuos tem cauda mais pesada que normal
    - Residuos nao seguem distribuicao normal.

Exemplo

- Refazendo a analise com o modelo multiplicativo leva aos seguintes resultados (note que tiramos o logaritmo das respostas y)

I	A	B	AB	y	Media(y)
1	-1	-1	1	(1.93, 1.90, 2.17)	2.00
1	1	-1	-1	(-0.05, 0.02, 0.03)	0.00
1	-1	1	-1	(-0.02, -0.03, 0.05)	0.00
1	1	1	1	(-1.83, -1.90, -1.93)	-1.89
0.11	-3.89	-3.89	0.11		
0.03	-0.97	-0.97	0.03		

Exemplo

- Porcentagem da variacao explicada nos dois modelos

Modelo Aditivo

Fator	I	A	B	AB	e
Efeito	26.55	-26.04	-26.04	25.54	
%Variacao		30.1	30.1	29.00	10.8
IC	(16.35, 36.74)	(-36.23, -15.84)	(-36.23, -15.84)	(15.35,35.74)	

Todos coeficientes sao significativos

Modelo Multiplicativo

Fator	I	A	B	AB	e
Efeito	0.03	-0.97	-0.97	0.03	
%Variacao		49.9	49.9	0	0.2
IC	(-0.02,0.07)	(-1.02,-0.93)	(-1.02, -0.93)	(-0.02, 0.07)	

Media e interacao nao sao significativos

Exemplo

- Graficos de residuo e quantil-quantil mostram resultados muito mais apropriados (ver livro): modelo multiplicativo e o adequado

$\log(y) = q_0 + q_A x_A + q_B x_B + q_{AB} x_A x_B + e$ ou

$$y = 10^{q_0} 10^{q_A x_A} 10^{q_B x_B} 10^{q_{AB} x_A x_B} 10^{e}$$

$$y = 10^{0.03} 10^{-0.97 x_A} 10^{-0.97 x_B} 10^{0.03 x_A x_B} 10^{e}$$

$$y = 1.07 \times 0.107^{x_A} \times 0.107^{x_B} \times 1.07^{x_A x_B} \times 10^{e}$$

- *Conhecimento sobre o sistema deve sempre tomar precedencia sobre consideracoes estatisticas!!*

Projetos Fatoriais $2^k r$

- *Ver Tabela 18.1 no livro com sumario de todos os passos para realizar o projeto $2^k r$!!*

Projetos Fatoriais Fracionarios 2^{k-p}

- Quando o numero de fatores *k* e muito grande, o custo de um projeto fatorial de experimentos pode ser muito caro
  - Requer 2^k experimentos (com/sem replicacao)
- Projetos fracionarios podem ser usados:
  - Permite analisar o impacto de k fatores (2 niveis) com um numero menor (2^{k-p}) de experimentos
  - Porem nao consegue isolar os efeitos de cada fator e de suas interacoes
  - Para minimizar este problema (*confounding*), e crucial um projeto cuidadoso dos experimentos a serem realizados

Exemplo: 2^{7-4} projeto

# Experimento	I	A	B	C	D	E	F	G
1	1	-1	-1	-1	1	1	1	-1
2	1	1	-1	-1	-1	-1	1	1
3	1	-1	1	-1	-1	1	-1	1
4	1	1	1	-1	1	-1	-1	-1
5	1	-1	-1	1	1	-1	-1	1
6	1	1	-1	1	-1	1	-1	-1
7	1	-1	1	1	-1	-1	1	-1
8	1	1	1	1	1	1	1	1

Niveis dos fatores em cada experimento cuidadosamente escolhidos

Apresentacao derivada dos slides originais de Virgilio Almeida

Projetos Fatoriais Fracionarios 2^{k-p}

- Colunas da tabela de experimentos precisam ser mutuamente exclusivas:
  - A soma de cada coluna e zero

$$\sum_i x_{i,j} = 0 \qquad \forall j$$

$$\text{onde } x_{i,j} \text{ e o nivel do fator j no experimento i}$$

  - A soma dos produtos de quaisquer duas colunas e zero $\sum_i x_{i,j} x_{i,l} = 0 \qquad \forall j \neq l$
  - A soma dos quadrados de cada coluna e 2^{k-p}

$$\sum_i x_{i,j}{}^2 = 2^{k-p}$$

Projetos Fatoriais Fracionarios 2^{k-p}

- A ortogonalidade permite calcular os efeitos e as suas contribuicoes para a variacao dos valores de y
- No exemplo anterior, estamos assumindo um modelo

$$y = q_0 + q_A x_A + q_B X_B + q_C x_C + q_D x_D + q_E x_E + q_F x_F + q_G x_G$$

- Podemos calcular cada efeito como anteriormente. Exemplo:

$$q_A = \frac{\sum_i y_i x_{Ai}}{8} = \frac{-y_1 + y_2 - y_3 + y_4 - y_5 + y_6 - y_7 + y_8}{8}$$

- Da mesma maneira, as formulas para desvio padrao dos efeitos e ICs podem ser computadas como para projetos fatoriais completos (com/sem replicacao) substituindo 2^k

Preparando a tabela de sinais para um experimento 2^{k-p}

- Escolha k-p fatores e prepare uma tabela de sinais completa para o projeto fatorial completo com k-p fatores.
  - Isto resultara em um tabela com 2^{k-p} linhas e 2^{k-p} colunas
  - A primeira coluna, marcada como I, contem somente 1's
  - As proximas k – p colunas serao marcadas com os k-p fatores escolhidos. As colunas restantes sao produtos destes fatores
- Das 2^{k-p} – (k-p) –1 colunas que restaram a direita da tabela, escolha p colunas e marque-as como os p fatores que nao foram escolhidos no passo 1.

Exemplo: $2^{7-4} = 2^3$ projeto

# Experimento	I	A	B	C	AB	AC	BC	ABC
1	1	-1	-1	-1	1	1	1	-1
2	1	1	-1	-1	-1	-1	1	1
3	1	-1	1	-1	-1	1	-1	1
4	1	1	1	-1	1	-1	-1	-1
5	1	-1	-1	1	1	-1	-1	1
6	1	1	-1	1	-1	1	-1	-1
7	1	-1	1	1	-1	-1	1	-1
8	1	1	1	1	1	1	1	1

Crie a tabela para um experimento 2^3 completo.
Escolha fatores A, B e C

Exemplo: $2^{7-4} = 2^3$ projeto

# Experimento	I	A	B	C	D	E	F	G
1	1	-1	-1	-1	1	1	1	-1
2	1	1	-1	-1	-1	-1	1	1
3	1	-1	1	-1	-1	1	-1	1
4	1	1	1	-1	1	-1	-1	-1
5	1	-1	-1	1	1	-1	-1	1
6	1	1	-1	1	-1	1	-1	-1
7	1	-1	1	1	-1	-1	1	-1
8	1	1	1	1	1	1	1	1

Marque as 4 colunas mais a direita com os 4 fatores faltantes (D-G)

Exemplo 2 : 2^{4-1} = 2^3 projeto

# Experimento	I	A	B	C	AB	AC	BC	ABC
1	1	-1	-1	-1	1	1	1	-1
2	1	1	-1	-1	-1	-1	1	1
3	1	-1	1	-1	-1	1	-1	1
4	1	1	1	-1	1	-1	-1	-1
5	1	-1	-1	1	1	-1	-1	1
6	1	1	-1	1	-1	1	-1	-1
7	1	-1	1	1	-1	-1	1	-1
8	1	1	1	1	1	1	1	1

Crie a tabela para um experimento 2^3 completo.
Escolha fatores A, B e C

Exemplo 2 : $2^{4-1} = 2^3$ projeto

# Experimento	I	A	B	C	AB	AC	BC	D
1	1	-1	-1	-1	1	1	1	-1
2	1	1	-1	-1	-1	-1	1	1
3	1	-1	1	-1	-1	1	-1	1
4	1	1	1	-1	1	-1	-1	-1
5	1	-1	-1	1	1	-1	-1	1
6	1	1	-1	1	-1	1	-1	-1
7	1	-1	1	1	-1	-1	1	-1
8	1	1	1	1	1	1	1	1

Arbitrariamente escolha a ultima coluna para o fator D

Este projeto permitira analisar o impacto dos efeitos principais bem como das interacoes AB, AC e BC

Confounding

- Um problema com projetos fracionarios e que nem todos os efeitos podem ser determinados
- Somente a influencia combinada de dois ou mais efeitos pode ser computada
- Este problema e conhecido como confounding
- Exemplo: no projeto 2^{4-1} anterior:

$$q_D = \sum_i y_i x_{Di} = \frac{-y_1 + y_2 + y_3 - y_4 + y_5 - y_6 - y_7 + y_8}{8} =$$

$$= \sum_i y_i x_{Ai} x_{Bi} x_{Ci} = q_{ABC}$$

Confounding

- Na verdade o somatorio dado nao e nem q_D nem q_{ABC} mas sim a soma dos dois efeitos

$$q_D + q_{ABC} = \sum_i y_i x_{Ai} x_{Bi} x_{Ci} =$$

$$= \frac{-y_1 + y_2 + y_3 - y_4 + y_5 - y_6 - y_7 + y_8}{8}$$

- Sem o projeto fatorial completo, nao e possivel obter estimativas separadas para os efeitos D e ABC
- Isto nao e um problema serio se e sabido, a priori, que o impacto da interacao entre A, B e C e pequena se comparada como o efeito D.

Confounding

- Notacao: D = ABC
- Note que outros efeitos tambem sao “confounded”
  - Ex: A = BCD
- Em um projeto 2^{4-1}, somente 8 dos 16 efeitos podem ser computados.
- Logo cada quantidade computada e na verdade a soma de dois efeitos.
- No exemplo anterior:

A = BCD B = ACD C = ABD AB = CD
AC = BD BC = AD ABC = D I = ABCD

Apresentacao derivada dos slides originais de Virgilio Almeida

Projeto 2^{4-1} Alternativo

- Para o exemplo 2^{4-1} um projeto alternativo e:

# Experimento	I	A	B	C	D	AC	BC	ABC
1	1	-1	-1	-1	1	1	1	-1
2	1	1	-1	-1	-1	-1	1	1
3	1	-1	1	-1	-1	1	-1	1
4	1	1	1	-1	1	-1	-1	-1
5	1	-1	-1	1	1	-1	-1	1
6	1	1	-1	1	-1	1	-1	-1
7	1	-1	1	1	-1	-1	1	-1
8	1	1	1	1	1	1	1	1

Projeto 2^{4-1} Alternativo

- Neste caso, o projeto contem os seguintes confoundings:

I = ABD, A = BD, B = AD, C = ABCD

D = AB, AC = BCD, BC = ACD, ABC = CD

- Este projeto pode nao ser tao bom quanto o anterior. Por que?

Algebra de Confounding

- O primeiro projeto 2^{4-1} mostrado e chamado projeto I = ABCD. O alternativo e chamado projeto I = ABD
  - Polinomio gerador
- E possivel obter todos os outros confoundings multiplicando os dois lados do polinomio gerador por termos diferentes e usando as seguintes regras:
  - A media I e tratada como unidade. Ex: IA = A
  - Qualquer termo elevado a potencia de 2 e apagado.

    Ex: $AB^2C = AC$
- Logo, se I = ABCD:

$AI = A^2BCD \Rightarrow A = BCD$

$BI = AB^2CD \Rightarrow B = ACD$

Algebra de Confounding

- Em um projeto 2^{k-p} , 2^p efeitos sao confounded. O polinomio gerador tem 2^p termos
- No exemplo 2^{7-4} , o projeto foi obtido substituindo as colunas AB, AC , BC e ABC por D, E , F e G. Logo:

  D = AB   E = AC   F = BC   G = ABC
- Multiplicando cada equacao pelo termo a esquerda:

  I = ABD   I = ACE   I = BCF   I = ABCG
- O produto de qualquer subconjunto dos termos acima tambem e igual a I. Logo, o polinomio gerador completo e:

  I = ABD = ACE = BCF = ABCG= BCDE = ACDF = CDG =

  = ABEF = BEG= AFG = DEF = ADEG = BDFG = ABDG =

  = CEFG = ABCDEFG

Exemplo

- Realizacao de experimentos para projeto de um escalonador. O sistema alvo permite processamento de texto, processamento de dados e processamento batch. O objetivo era descobrir qual tipo de escalonador usar para cada tipo de carga. Seis fatores foram considerados:
  - O escalonador e preemptivo: nao (-1) ou sim (+1)
  - O time slice e pequeno (-1) ou grande (+1)
  - Numero de filas: uma (-1) ou duas (+1) (2a fila tem time slide menor)
  - Requeueing: duas filas (-1) ou cinco filas (+1)
  - Dar preferencia a processos esperando por muito tempo para melhorar a justica: nao (-1) ou sim (+1)

Um projeto de experimentos 2^{5-1} com gerador I = ABCDE foi utilizado . Os niveis dos fatores e os throughputs para cada tipo de carga obtidos sao mostrados na tabela a seguir

Exemplo

Experimento	A	B	C	D	E	T_{text}	T_{data}	T_{batch}
1	-1	-1	-1	-1	1	15	25	15.2
2	1	-1	-1	-1	-1	11	41	3
3	-1	1	-1	-1	-1	25	36	21
4	1	1	-1	-1	1	10	15.7	8.6
5	-1	-1	1	-1	-1	14	63.9	7.5
6	1	-1	1	-1	1	10	13.2	7.5
7	-1	1	1	-1	1	28	36.3	20.2
8	1	1	1	-1	-1	11	23	3
9	-1	-1	-1	1	-1	14	66.1	6.4
10	1	-1	-1	1	1	10	9.1	8.4
11	-1	1	-1	1	1	27	34.6	15.7
12	1	1	-1	1	-1	11	23	3
13	-1	-1	1	1	1	14	26	12
14	1	-1	1	1	-1	11	38	2
15	-1	1	1	1	-1	25	35	17.2
16	1	1	1	1	1	11	22	2

Exemplo

Efeitos medios e % variacao explicada para cada tipo de carga:

Confounded Effects		Ttext		Tdata		Tbatch	
1	2	Estim.	%var	Estim.	%var	Estim.	%var
I	ABCDE	15.44		31.74		9.54	
A	BCDE	**-4.81**	**55.5**	**-8.62**	**31**	**-4.86**	**58.,8**
B	ACDE	**3.06**	**22.5**	**-3.54**	**5.2**	**1.79**	**8**
C	ABDE	0.06	0	0.43	0.1	-0.62	1
D	ABCE	-0.06	0	-0.02	0	-1.21	3.6
AB	CDE	**-2.94**	**20.7**	1.34	0.8	**-2.33**	**13.5**
AC	BDE	0.06	0	0.49	0.1	-0.44	0.5
AD	BCE	0.19	0.1	-0.08	0	0.37	0.3
BC	ADE	0.19	0.1	0.44	0.1	-0.12	0
BD	ACE	0.06	0	0.47	0.1	-0.66	1.1
CD	ABE	-0.19	0.1	-1.91	1.5	0.58	0.8
DE	ABC	-0.06	0	0.21	0	-0.47	0.5
CE	ABD	0.06	0	1.21	0.6	-0.16	0.1
BE	ACD	0.31	0.2	**7.96**	**26.4**	-1.37	4.7

Exemplo

- Conclusoes:
  - Os valores ideais para os parametros sao diferentes para as tres cargas. Observando as % de variacao explicadas, nota-se que os fatores com maior impacto sao:
    - Processamento de texto: A (preempcao), B (time slice) e a interacao AB
    - Processamento de dados: A (preempcao), B (time slice), E (fairness) e BE
    - Batch: A (preempcao), B (time slice), E (fairness) e AB
  - O fator C (# filas) ou qualquer um de suas interacoes parece nao ter impacto significativo no throughput. O mesmo e valido para o fator D (requeuing)
  - O fator B tem menor impacto que o fator A, que por sua vez tem impacto significativo para os 3 tipos de cargas
  - O fator E e importante para jobs interativos e tambem (embora com menor impacto) para processamento batch

Exemplo 2

- O tempo de CPU gasto por dois formatadores de texto, LaTex e troff, foi medido usando arquivos sinteticos de diferentes tamanhos e niveis de complexidade. Seis fatores, cada um com dois niveis, foram escolhidos para o estudo:

Simbolo	Fator	Nivel –1	Nivel +1
A	formatador	latex	troff
B	tamanho arquivo	2100 Bytes	25000 Bytes
C	# equacoes	0	10
D	# floats	0	10
E	# tabelas	0	10
F	# footnotes	0	10

Apresentacao derivada dos slides originais de Virgilio Almeida

Exemplo 2

Um projeto fatorial fracionario 2^{6-1} foi realizado com polinomio gerador I = BCDEF. Os maiores efeitos e interacoes, computados a partir da tabela de sinais sao mostrados abaixo:

Simbolo	Fator	Efeito	%Variacao
B	Tamanho	12	39.4
A	Formatador	9.4	24.4
C	# equacoes	7.5	15.6
AC	Formatador X #Equacoes	7.2	14.4
E	# tabelas	3.5	3.4
F	# footnotes	1.6	0.7

Exemplo 2

Conclusoes:

- Mais de 90% da variacao pode ser explicada pelos tres fatores: tamanho, formatador e # equacoes e uma interacao de segunda ordem
- A variação no tamanho dos arquivos de entrada foi muito grande, tornando o efeito maior que o efeito dos formatadores sendo analisados
- A interacao “Formatador X Tamanho” e baixa. Isto indica que mudar o tamanho do arquivo afeta os dois programas (latex e troff) de maneira similar.

Exemplo 2

- A alta % de variacao explicada pela interacao “Formatador X # equacoes” indica que a escolha do formatador depende do número de equacoes no texto. Se considerarmos apenas o formatador e o # de equacoes, o tempo de CPU relativo gasto nas 4 combinacoes e mostrado abaixo:

	Numero de Equacoes	
Formatador	-1 (0)	+1 (10)
-1 (Latex)	-9.7	-9.1
+1 (troff)	-5.3	24.1

Isto mostra que troff gasta muito tempo de CPU se ha equacoes no texto.

Exemplo 2

- Se possivel, os experimentos deveriam ser refeitos considerando uma variacao menor no tamanho dos arquivos, de forma que os programas (formatadores) e nao a carga aparecam como o fator mais significativo. Alternativamente, o número de niveis de tamanhos de arquivos deveria ser aumentado.

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:

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

Apresentacao derivada dos slides originais de Virgilio Almeida

Computando os Efeitos

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

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

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

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.

Apresentacao derivada dos slides originais de Virgilio Almeida

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)

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

Apresentacao derivada dos slides originais de Virgilio Almeida

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:

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

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

$$SSY = SS0 + SSA + SSE$$

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

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

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

Apresentacao derivada dos slides originais de Virgilio Almeida

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

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

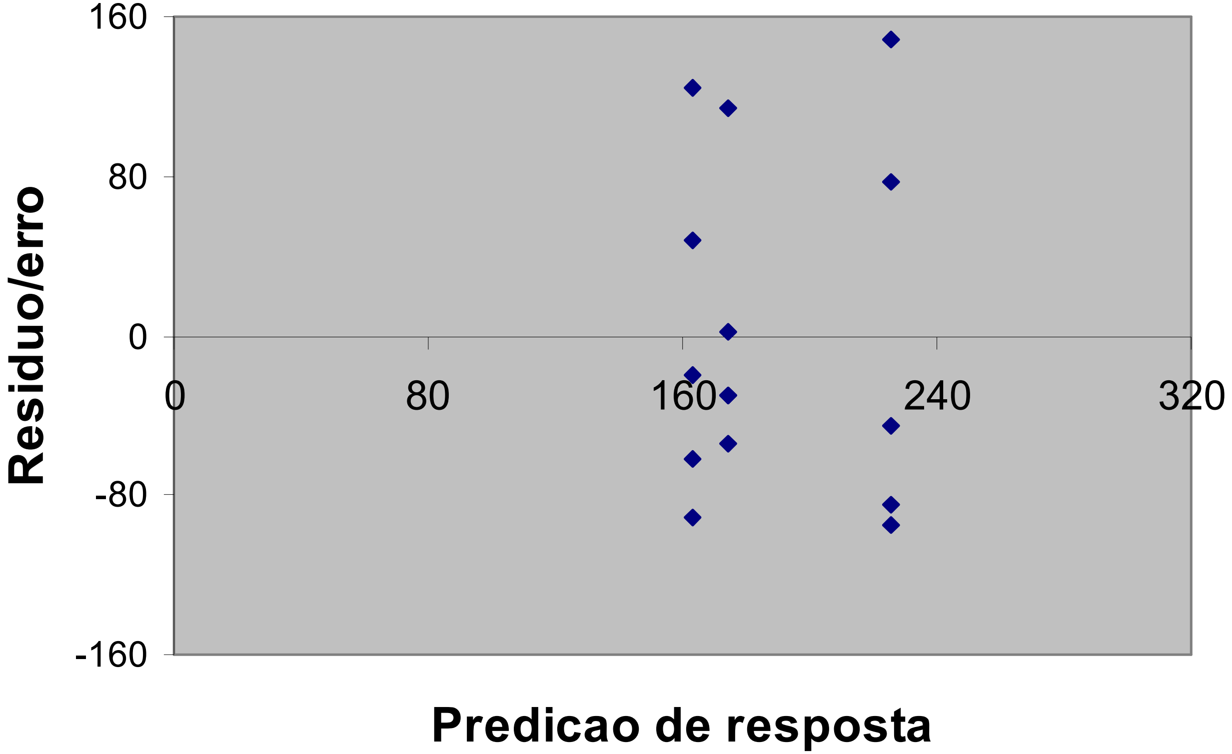
$F_{computado} < F_{tabela}$: diferencas em tamanhos de codigo sao principalmente devido aos “erros experimentais”

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	$\overline{y}_{..}$	$s_e^2 \big/ ar$
α_j	$\overline{y}_{.j} - \overline{y}_{..}$	$s_e^2(a-1) \big/ ar$
$\mu + \alpha_j$	$\overline{y}_{.j}$	$s_e^2 \big/ r$
$\sum_{j=1}^{a} h_j \alpha_j, \sum_{j=1}^{a} h_j = 0$	$\sum_{j=1}^{a} h_j \overline{y}_{.j}$	$\sum_{j=1}^{a} s_e^2 h_j^2 \big/ r$
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

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

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

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

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

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 $\sum_j h_j = 0$)

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

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

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

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

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

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 $N = \sum_j r_j$

Apresentacao derivada dos slides originais de Virgilio Almeida

Intervalos de Confianca para Efeitos

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

Graus de liberdade dos erros : N - a

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= $\sum_j r_j \alpha_j^2$	$100\frac{SSA}{SST}$	a-1	$MSA = \frac{SSA}{a-1}$
Erros	SSE = SST-SSA	$100\frac{SSE}{SST}$	N-a	$MSE = \frac{SSE}{N-a}$

$$s_e = \sqrt{MSE}$$

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

Apresentacao derivada dos slides originais de Virgilio Almeida

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.

Apresentacao derivada dos slides originais de Virgilio Almeida

Exemplo 2

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

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

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

Apresentacao derivada dos slides originais de Virgilio Almeida

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

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

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

Apresentacao derivada dos slides originais de Virgilio Almeida

Sumario

- Ver tabela 20.1 (pagina 341)

Apresentacao derivada dos slides originais de Virgilio Almeida