

OTIMIZAÇÃO EXPERIMENTAL COM
OPERAÇÃO EVOLUTIVA

WALMIR MONTALVÃO

ORIENTADOR

JOSÉ NORBERTO W. DACHS

Dissertação apresentada no Instituto de Matemática, Estatística e Ciência da Computação, UNICAMP, como requisito parcial para obtenção do título de Mestre em Matemática Aplicada.

Dezembro 1978

Aos meus pais .

AGRADECIMENTOS

Ao Norberto, pela natureza do problema que me propos , bem como pela considerável ajuda na etapa de redação.

Ao Carlos Vaz, pelo esforço em nos conseguir uma aplicação prática da OPEV, bem como pela ajuda na interpretação dos resultados.

Ao Prof. Ahmed pelas explicações sobre processo de Extrusão e pelas figuras fornecidas.

Ao Miguel Taube pelo incentivo para que eu viesse para o IMECC.

Ao Schneider e Mauro, pelo trabalho de revisão do texto.

À Valéria pelos auxílios computacionais.

À Isabel (FEAA), pela boa vontade e paciência em avaliar as concentrações das amostras na fase experimental deste trabalho.

À Gilda pelas explicações sobre delineamentos fatoriais, o que facilitou bastante a compreensão da OPEV.

À Angelina, pelo belo trabalho com os desenhos.

Aos colegas, professores, amigos por terem tornado gratificante a minha estadia em Campinas.

SUMÁRIO

Este trabalho tem como finalidade:

1) Revisão da Literatura sobre OPERação EVolutiva (OPEV), uma técnica de Otimização Experimental.

2) Descrição da técnica numa forma didática. A técnica é pouco empregada no Brasil e o texto poderia servir de roteiro para difusão da mesma.

3) Uma aplicação prática onde os conceitos assimilados pudessem ser empregados.

4) Simulação de um processo para treinamento de pessoas durante um curso. Isso sanaria a dificuldade de treinamento no próprio processo.

No Capítulo 1 é dada uma idéia de sistemas e modelos. Depois é caracterizada a técnica OPEV como uma Otimização Experimental.

No Capítulo 2 são apresentados os princípios estatísticos necessários à OPEV.

Nos Capítulos 3 e 4 são detalhados os procedimentos da OPEV para delineamentos fatoriais 2^2 e 2^3 .

No Capítulo 5 é apresentado um exemplo de aplicação da OPEV na Otimização de uma Centrífuga.

No Capítulo 6 é proposto o jogo da OPEV.

No apêndice 1 estão contidos alguns comentários e indicações sobre uma referência bibliográfica mais completa.

No apêndice 2 aparece o manual de uso do programa que realiza o jogo.

SUMMARY

The scope of this work is:

1) To review the literature on Evolutionary Operation, an experimental optimization technique.

2) To describe the technique in a didactical way as it is of little use in Brazil and the text could become a mean of spreading its use.

3) A practical application in which the EVOP technique is used.

4) Simulation of a process for personnel training in an EVOP course. This would avoid problems of training in the physical process itself.

In the 1st chapter an idea of systems and models is presented and then EVOP is introduced as an experimental optimization technique.

In the 2nd chapter the statistical principles on which EVOP is based are presented.

In the 3rd and 4th chapters the EVOP procedure is detailed for 2^2 and 2^3 factorial designs.

In the 5th chapter an application is made using EVOP in the optimization of a centrifugal pump.

The 6th chapter is the EVOP game.

The 1st Appendix contains comments on the bibliography.

The 2nd Appendix is a manual of the EVOP game program.

ÍNDICE

CAPÍTULO	PÁGINA
1 - SISTEMAS E MODELOS	1
2 - PRINCÍPIOS ESTATÍSTICOS	10
3 - OPERAÇÃO EVOLUTIVA	45
<i>Folha de trabalho para Programa OPEV com duas variáveis</i>	59
<i>Detalhes dos cálculos nas folhas de trabalho</i>	64
<i>Folha de trabalho para um esquema fa torial 2² sem condições adicionais de referência.</i>	68
4 - DELINEAMENTO FATORIAL 2 ³	71
<i>Folha de trabalho para programa OPEV com três variáveis</i>	89
<i>Folha de trabalho para um esquema fa torial disposto em dois blocos, com uma corrida de referência em cada blo co.</i>	90
<i>Folhas de trabalho para um modelo fa torial 2³ com disposição em dois blo cos, sem condição adicional de refe rência</i>	97

5 - OTIMIZAÇÃO DE UMA CENTRÍFUGA	100
<i>Motivação</i>	100
<i>Materiais e métodos</i>	100
<i>Medidas físicas</i>	104
<i>Procedimentos da operação</i>	106
<i>Resultados e discussão</i>	107
<i>Conclusões</i>	117
6 - O JOGO DA OPEV	119
<i>Processo de extrusão de farinha</i>	120
<i>Obtenção de pão com elevado teor proteico.</i>	123
<i>Fluxograma do jogo</i>	126
APÊNDICE I - <i>Comentários</i>	127
APÊNDICE II - <i>Manual de uso para os pro- gramas OPEVTY e OPEVWR</i>	128

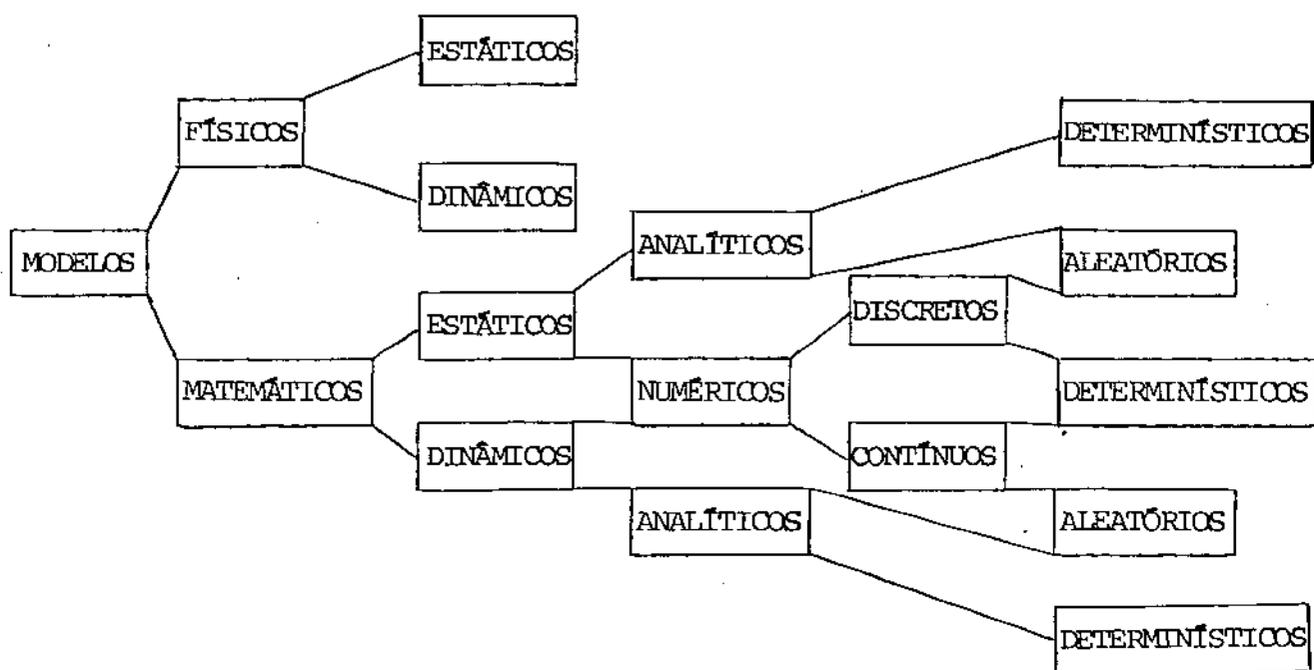
1 - SISTEMAS E MODELOS

Ao termo sistema, usado numa grande variedade de maneiras e significados, inferiremos o conceito de um conjunto organizado de objetos, estruturados de forma interativa ou interdependente.

Modelar um sistema é representá-lo convenientemente para permitir o seu estudo. Modelos podem ser utilizados na otimização ou apenas na previsão do sistema em estudo. Muitos tipos de modelos tem sido usados para o estudo de sistemas. Podem ser classificados em:

- físicos e matemáticos
- estáticos e dinâmicos
- analíticos e numéricos
- contínuos e discretos
- determinísticos e aleatórios.

Esquemáticamente teremos:



Com a evolução da PESQUISA DE OPERAÇÕES (P.O.) e o aparecimento dos computadores, a otimização conheceu um novo desenvolvimento: com a programação matemática, tem sido possível resolver problemas até então sem solução pelos métodos clássicos de cálculo. Os problemas tratados pela programação matemática são ligados à área de economia, indústria, planificação, etc.; os problemas tratados pelos métodos clássicos de otimização são mais apropriados à física e à geometria.

Resolver um modelo consiste em determinar ou estimar os valores das variáveis controladas que o otimizam segundo algum critério pré-fixado, ou que prevêm seu comportamento nas situações em estudo. Os métodos da matemática clássica podem ser usados em alguns casos para se obter soluções por dedução. Quando isso não é possível, poderão ser obtidas por indução, através de algum método numérico, por exemplo.

Seguindo o diagrama, quando o modelo é matemático, dinâmico e numérico, a técnica empregada para resolver o modelo é conhecida como simulação de sistemas. A simulação é útil no estudo dos processos onde há valores de transição, para estimar valores dos parâmetros do modelo e para analisar linhas de ação que não possam ser incorporadas ao modelo.

Ainda seguindo o diagrama, quando o modelo é matemático, estático ou dinâmico, e analítico, temos caracterizada uma função a ser otimizada segundo um critério de decisão estabelecido na formulação do problema. Essa função, conhecida como função objetiva ou função critério, pode ser otimizada com ou sem restrições ,

sendo que a região em que não existe violação das restrições é conhecida como região factível ou região admissível.

Existem, na programação matemática, vários métodos para otimizar nossa função critério na região admissível. Esses métodos podem ser classificados em 3 categorias:

- a) Os métodos usando somente valores funcionais, chamados métodos diretos.
- b) Os métodos que também fazem uso das derivadas de primeira ordem.
- c) Os métodos que também requerem conhecimento das derivadas de segunda ordem.

Em situações onde não se podem construir modelos, mas é possível realizar experiências, é aplicável a otimização experimental, desde que o sistema seja estatisticamente estável. Para uma função objetivo que seja desconhecida, e/ou tenha incorporados a ela erros aleatórios não desprezíveis, existe um conjunto de técnicas possíveis de se utilizar na melhoria da mesma.

Suponhamos proposto o seguinte problema: estudar as condições de sobrevivência de uma colônia de determinada espécie num certo ambiente. Para efeito de estudo, o sistema consistirá da população e do meio em que ela se encontra. Pela maneira como o problema é proposto, faz-se necessário um modelo que possa prever o comportamento do sistema. Tal modelo deve relacionar o crescimento da população com as condições ambientais do sistema. Fixadas essas condições, haverá uma variabilidade genética (mutação) e, pela própria seleção natural, os indivíduos que melhor se adaptarem a

essas condições terão maiores chances de sobrevivência. E, à medida que ocorrem variações nas condições ambientais, acontecerão novas seleções de indivíduos.

Vejam os outros sistemas: uma cidade. Ela sofre constantes alterações no número de habitantes, no número de veículos, etc. Com sua infraestrutura afetada, precisa ser modificada. A equipe de Planejamento Urbano pondera as várias alternativas de mudança e seleciona as variações favoráveis. Isso é realizado continuamente, pois a própria execução das mudanças acarreta alteração no ambiente do sistema. Mais uma vez temos variação - seleção, sendo agora artificial essa seleção.

Vamos considerar ainda um outro sistema: uma fábrica química. Um processo industrial em implantação passa por vários estágios de desenvolvimento. Em primeiro lugar tem-se um trabalho de laboratório, frequentemente prolongado. Após isso, temos uma estimativa das proporções entre insumos utilizados e produtos obtidos. Os resultados de laboratório fornecem uma idéia preliminar de praticabilidade, permitindo objetivos realísticos a serem definidos, e conduzindo a um esboço de um processo industrial. Esse esboço pode então ser usado para construir um modelo físico dinâmico do que queremos produzir: a planta piloto. Nessa etapa ele é otimizado quase sempre experimentalmente. Agora, com as condições "ideais" de operação obtidas em escala reduzida na planta piloto, a fábrica começa a operar em escala real.

A experimentação em pequena escala fornece importantes informações acerca do processo, bem como primeiras indicações so-

bre as condições de operação em escala real. Não obstante, características como fluxo, transmissão de calor, mistura, etc., não serão mantidas com a mudança da escala de operação, e as técnicas de engenharia utilizadas na compensação se apresentam imperfeitas.

Não são esperadas, é claro, características completamente diferentes na escala real; caso isso ocorresse, os esforços empregados durante a etapa em pequena escala seriam inúteis. Na transferência do processo de pequena escala para a escala real são preservadas as principais características desse comportamento, mas existem diferenças em detalhes que são importantes economicamente. Uma ilustração é dada na fig. 1.2. Ela mostra as curvas de nível * do rendimento para uma mesma reação química conduzida em (a) escala reduzida e (b) escala real. Notamos que as curvas de nível do rendimento em uma fábrica, em geral, são distorcidas e deslocadas quando comparadas com as curvas de nível da planta piloto, muito embora as características básicas sejam muito semelhantes. A melhor combinação de tempo e concentração na escala pequena (ponto P) dará rendimento mais baixo na escala completa. Esforços para mover

* Uma função que relaciona o nosso critério com as variáveis envolvidas no processo em consideração caracteriza a conhecida superfície de resposta. Nos sistemas onde os modelos que os representam são de natureza aleatória e/ou não tenham representação analítica, somente alguns pontos são disponíveis, e não se sabe ao certo qual a forma dessa superfície. Pontos onde a função critério (pontos da superfície de resposta) tem o mesmo valor numérico constituem o curva de nível.

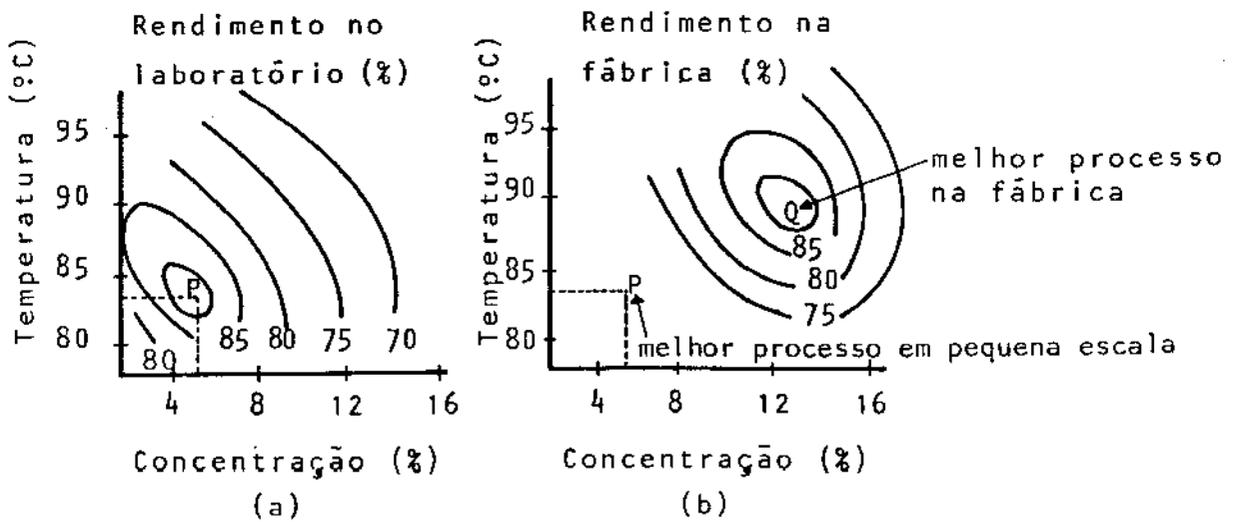


Fig. 1.1 Possível aparência da superfície de rendimento, mostrando contornos de rendimentos percentuais, para processos conduzidos.

- a) pequena escala
- b) escala de fábrica

o ótimo P. da pequena escala para o ótimo Q na escala real devem, necessariamente, ser feitos na própria fábrica, já que a pesquisa em pequena escala irá sempre retornar a P.

O diagrama ilustra esse ponto, mas, naturalmente, é uma simplificação do problema. Na prática, temos distorção e deslocamento devidos ao aumento proporcional não em um espaço bi-dimensional, mas em um espaço multi-dimensional. Quase inevitavelmente o produto estará sendo industrializado a taxas de produção mais baixas, com rendimento mais baixo, e com uma qualidade inferior àquela de que a fábrica é capaz.

O processo de sintonização (refinamento) ainda está por ser feito, e com ele iremos nos preocupar. Embora o trabalho de

ajustamento seja uma tarefa importante, não é simples. Em muitas organizações industriais, as equipes de pesquisa e desenvolvimento efetuam esforço considerável para aperfeiçoar os processos, com estudos especiais efetuados no laboratório, na planta piloto e no processo em escala real. A não disponibilidade de pessoal qualificado para a tarefa limita a quantidade de investigação especial dessa natureza. Discutiremos um método adicional que poupa o uso de capacidade técnica e ajuda a completar as investigações já mencionadas.

É uma técnica simples, mas potente, chamada Operação Evolutiva - OPEV (Evolutionary Operation - EVOP). Tem sido largamente aplicada (infelizmente pouca é a sua difusão no Brasil), e com sucesso, nas indústrias químicas, mas sem dúvidas seria útil em outros tipos de indústria, como alimentos, plásticos, etc. OPEV é aplicada à fábrica em escala completa, de forma contínua e sistemática. Não requer nenhuma tarefa difícil de coleta de dados sendo que esta pode ser executada por um operário comum da fábrica após um período breve de treinamento. Operação Evolutiva não substitue a investigação básica. Entretanto, indica áreas para as quais as tentativas fundamentais poderiam ser dirigidas de forma mais proveitosa.

Já está claro que a fábrica precisará refinar as condições de operação do processo, isto é, encontrar outros valores para algumas das variáveis envolvidas de modo a obter um melhor desempenho do sistema. Algumas compensações são obtidas por ajustes empíricos na ocasião da elaboração do projeto em escala real

e sua construção. A introdução de OPEV na própria fábrica aumenta a velocidade com que ocorrem as melhorias. Mesmo onde não há dificuldade na passagem da escala reduzida para a real, poderia ser dispendioso desenvolver um processo de operação por experimentação apenas em pequena escala. Demanda laboratórios suficientemente grandes, instalações piloto e capacidade técnica para o desenvolvimento de todos os processos simultâneos, o que poderia se tornar impraticável por ser proibitivamente caro. Já que existe agora a própria planta e o pessoal para operá-la, por que não utilizá-los para a experimentação, na própria rotina de produção, objetivando melhorar as condições de operação ?.

A produção de rotina normalmente é conduzida operando a fábrica em condições rigidamente definidas, chamadas processo de trabalho. O processo de trabalho incorpora as melhores condições de operação até então conhecidas. Este procedimento de fabricação é chamado método de operação estática. Na prática estamos interessados não somente na produtividade do processo, mas também nas propriedades físicas do produto que é manufaturado. Tais propriedades poderiam cair fora dos limites de especificação caso desvios arbitrários fossem permitidos. Nosso método de operação deve incluir salvaguardas que tornem aceitavelmente pequeno o risco de obtenção de quantidades apreciáveis de produtos com qualidade insatisfatória.

O método OPEV consiste, inicialmente, em planejar um ciclo de pequenas variações nas condições de operação. Dessa maneira nós usamos a própria rotina para gerar não apenas o produto, mas

também informações necessárias para melhorá-lo*. Tendo então sido introduzida a variação controlada, os resultados devem ser submetidos ao superintendente do processo em uma forma facilmente compreensível. Isso permite a ele ver que mudanças deveriam ser feitas para melhorar a fabricação. Mais uma vez temos variação-seleção.

Espera-se alguma produção abaixo dos padrões, e deve-se eventualmente fazer previsão de custo para tal. Operação Evolutiva, entretanto, é um método quase permanente de operar a planta, e irá, portanto, exigir poucos recursos e instalações especiais. Por essa razão podem ser permitidas mudanças nos níveis das variáveis cujos efeitos sejam indetectáveis em corridas individuais, podendo ser empregadas somente técnicas simples o bastante para serem utilizadas continuamente.

* Contrastando com um sistema que encerra uma colônia e seu habitat, onde a variação e a seleção são naturais, vemos agora que a variação tem que ser introduzida artificialmente, e a seleção das variações que melhorem o processo é o nosso objetivo.

2 - PRINCÍPIOS ESTATÍSTICOS

Os cálculos requeridos na operação de um esquema OPEV são bastante simples. É altamente desejável, entretanto, que aqueles responsáveis por essa operação tenham alguma compreensão dos princípios estatísticos elementares nos quais os cálculos são baseados. Além desses conceitos estatísticos são a seguir apresentados diversos termos de uso comum em OPEV.

Corrida - Será chamada uma corrida o período de operação em que as condições do processo são mantidas em níveis fixos. Em alguns processos onde a inércia do sistema é alta, o tempo para equilibrar (transitório) pode ser demasiado longo. Entre uma corrida e outra deve ser esperado tempo suficiente para o equilíbrio ser atingido.

Variação nos Resultados - Mesmo quando as condições do processo são mantidas constantes, os resultados observados variam de corrida para corrida. Essa variação provém de muitas fontes diferentes. Existem erros de medidas e erros analíticos, mas a maior causa da variação é a impossibilidade de reproduzir precisamente as condições pretendidas durante qualquer corrida em particular.

Ciclo e Fase - Uma única execução de um conjunto pré-determinado de corridas é denominada ciclo. Por causa da variabilidade dos resultados, é conveniente repetir o mesmo ciclo um número suficiente de

vezes que permita indicar qual mudança, (se houver), nas condições de operação, é aconselhável para melhorar o processo. Temos então caracterizada uma fase.

Dispondo os Dados em Ordem - Suponhamos que para algum processo químico particular, as observações de rendimento para 10 corridas sucessivas tenham sido as seguintes: 66.1, 63.7, 65.8, 64.8, 65.7, 66.3, 64.8, 65.2, 64.2, 64.3, 65.3. Esses resultados são mapeados na fig. 2.1 na ordem em que foram obtidos, o que às vezes permite detetar tendências ou outros fenômenos importantes.

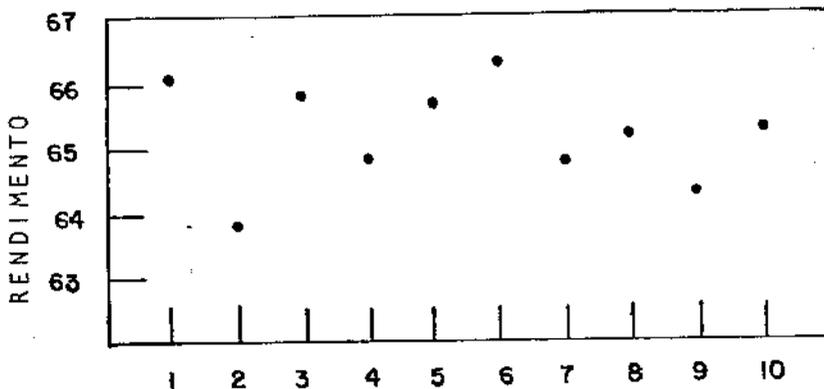


Fig. 2.1 Dez observações de rendimento dispostos na ordem em que foram obtidos.

Suposição de Aleatoriedade e Independência - Doravante vamos assumir situações onde os resultados variam aleatoriamente em torno de algum valor médio. Suporemos também que a probabilidade do desvio do valor médio exceder um tamanho especificado não é afetado pelo valor de outra observação. Quando isso é verdade, dizemos que os

desvios são estatisticamente independentes. A suposição de aleatoriedade, embora nem sempre possa ser assumida com rigor, fornecerá uma aproximação adequada na maioria das vezes.

O Diagrama de Pontos - Quando pode ser assumida a independência entre as observações, a sequência no tempo é irrelevante, nada se perdendo por ignorá-la. Com essa suposição, o mapa da fig. 2.1 pode ser refeito conforme a fig. 2.2 - chamada de diagrama de pontos.

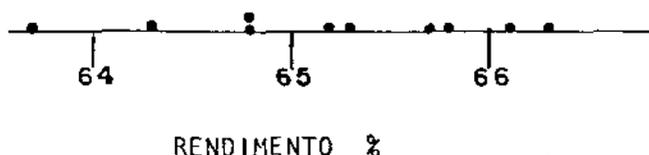


Fig. 2.2 Diagrama de pontos para rendimento percentual.

O diagrama de pontos é útil para representar a distribuição de um corpo de dados com até 20 observações aproximadamente. Em particular, ele nos chama a atenção para:

- 1 - Localização geral das observações (no exemplo podemos ver que os rendimentos estão mais próximos de 65% do que 85% ou 35%).
- 2 - A dispersão das observações (no exemplo sua amplitude cobre cerca de 3 unidades percentuais).

Distribuição de Frequência - Quando um grande corpo de dados está sendo estudado, os pontos individuais em um diagrama de pontos se apresentam de forma bastante confusa. É mais fácil apreciar os dados construindo uma distribuição de frequência. Para isso, podemos

subdividir o eixo das abcissas em intervalos iguais, com algum tamanho escolhido, e cravamos os pontos que caem em cada intervalo

INTERVALO DA CLASSE	NÚMERO DE OBSERVAÇÕES POR INTERVALO=FREQUÊNCIA
ABAIXO - 62.5	0
62.5 - 63.0	2
63.0 - 64.5	4
63.5 - 64.0	6
64.0 - 64.5	11
64.5 - 65.0	18
65.0 - 65.5	21
65.5 - 66.0	17
66.0 - 66.5	12
66.5 - 67.0	8
67.0 - 67.5	1
ACIMA DE 67.5	0
TOTAL	100

TABELA 2.1 - DISTRIBUIÇÃO DE FREQUÊNCIA DE
100 OBSERVAÇÕES

no respectivo valor central. Em cada um deles podemos construir um retângulo cuja altura é proporcional ao número de observações naquele intervalo. Geralmente, é adequado o uso de 10 a 20 deles, chamados classes, cobrindo uma amplitude relevante. A fig. 2.3 mos

tra a distribuição de frequência para 100 observações de rendimento em um processo de produção. Os 100 valores são distribuídos em

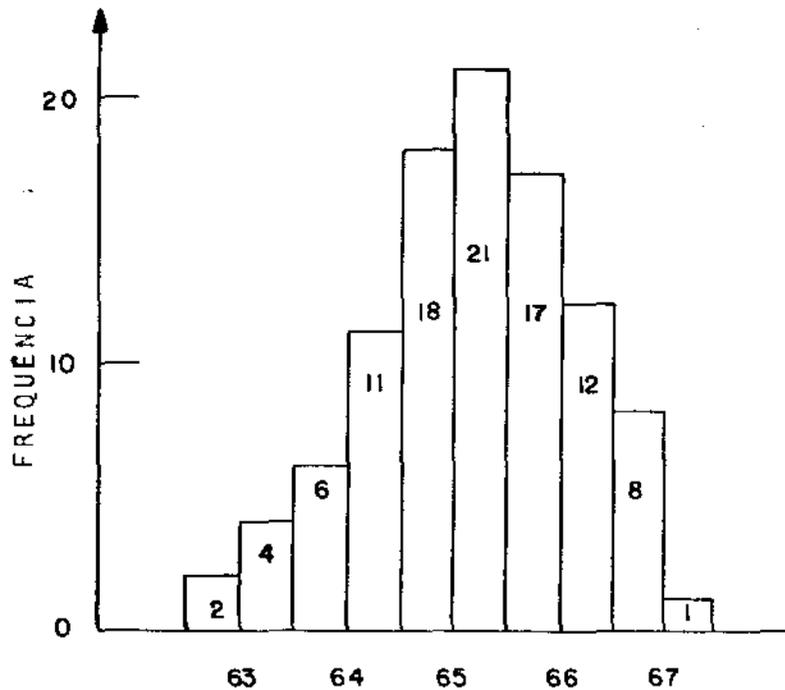


Fig. 2.3 Distribuição de frequência para 100 observações de rendimento.

intervalos selecionados conforme mostrado na Tabela 2.1. O número de observações em cada classe é chamado frequência da classe. Então, para esse exemplo, duas observações se situam na primeira delas que tem intervalo 62.5-63.0, tal que sua frequência é 2.

A distribuição de frequência da fig. 2.3 mostra a localização e dispersão dos valores observados. Se fôssemos trabalhar com um número muito grande de observações, poderíamos nos dar ao luxo de tornar as classes de frequência mais estreitas e numerosas, e ainda manteríamos uma quantidade razoável de observações em cada classe. Se o número total de observações fosse muito grande, os intervalos poderiam ser extremamente pequenos e a forma do diagrama

tenderia a uma curva suave conforme mostrada na fig. 2.4. Se nós

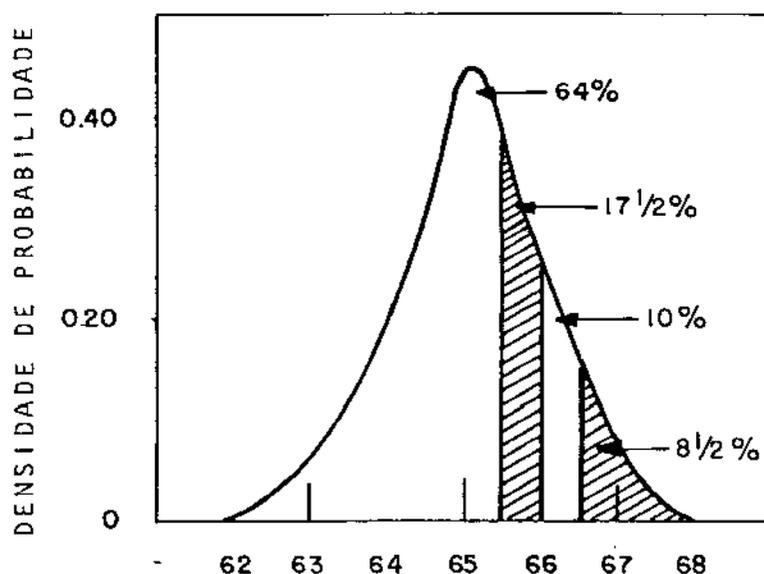


Fig. 2.4. Densidade de probabilidade para rendimento percentual.

arranjarmos uma escala vertical tal que a área sob essa curva seja unitária, a curva é chamada função densidade de probabilidade ou mais simplesmente densidade. A densidade, com algumas características que veremos depois, identifica a variável aleatória associada às observações.

População e Amostra - Uma grande coleção (conceitualmente infinita) de todos os resultados possíveis de ocorrer ao efetuarmos alguma operação particular é chamada população. Uma variável aleatória, com densidade como a da fig. 2.4, descreve uma população dizendo as frequências relativas com que os resultados de diferentes tipos podem ocorrer.

Em muitos casos, uma população completa de possíveis resultados não é disponível. Não obstante, a idéia de uma população subjacente em uma dada operação é um conceito bastante importante. Quando nós efetuamos uma corrida e determinamos o rendimento em algum conjunto de condições, agimos como se fôssemos participar de um grande sorteio. Podemos imaginar uma enorme urna de loteria contendo um número infinito de bilhetes. Cada bilhete é marcado com um valor de percentagem de rendimento que poderia ser obtido quando nós tentássemos operar a aparelhagem nas condições escolhidas. A função densidade de probabilidade descreve o número relativo de bilhetes na urna, inscritos com os vários resultados de rendimento. Caso suponhamos que a operação com que estejamos envolvidos tenha os resultados estatisticamente independentes, nada se pode dizer sobre o valor particular de resultado que será obtido, mesmo conhecendo valores observados anteriormente. Não obstante, se nós conhecermos a função densidade de probabilidade para a população em questão, nós poderemos fazer afirmações probabilísticas acerca dos possíveis resultados. Por exemplo, a escala vertical da fig. 2.4 foi preparada de forma que a área total sob a curva é igual a um. A área sob a curva para a direita de qualquer valor de rendimento representa a proporção de rendimentos de uma população que exceda esse valor; A área à direita de 66.5 (sombreada na fig. 2.4) é 8.5% da área total sob a curva. Isso poderia implicar, em nossa analogia, que do número total de bilhetes, 8.5% teriam números mais altos que 66.5. Supondo que os valores dos rendimentos variam aleatoria-

mente em torno da média, então a probabilidade de obter um valor de rendimento maior que 66.5, naquele conjunto fixado de condições, é 8.5%. De forma equivalente, a área sombreada entre 65.5 e 66 é 17.5% da área total sob a curva na fig. 2.4, de forma que a probabilidade de obter um rendimento entre 65.5 e 66 é 17.5%.

Agora suponhamos que tudo o que nós temos são as observações dispostas na fig. 2.2. Do ponto de vista estatístico, nós olharíamos essas observações como uma amostra aleatória de uma (hipotética) população infinita descrita por uma (desconhecida) função densidade de probabilidade; nós poderíamos imaginar que nossos 10 resultados constituem 10 bilhetes tirados da urna. Embora tudo que nós tenhamos seja uma amostra de 10 observações do processo, nós gostaríamos de poder fazer afirmações acerca seu comportamento em geral, isto é, de poder fazer afirmações acerca da população. Qualquer afirmação dessa natureza certamente será indutiva. O principal objetivo da Estatística é raciocinar do particular para o geral, isto é, de uma amostra que seja disponível para uma população não disponível.

Medidas Características: média e variância - Para discutir a população hipotética e sua função densidade de probabilidade, nós precisamos ter maneiras de medir suas características. Uma característica muito importante é aquela que possa medir sua posição geral ou locação. Na figura 2.5, nós vemos duas densidades idênticas na forma, mas com locação diferente.

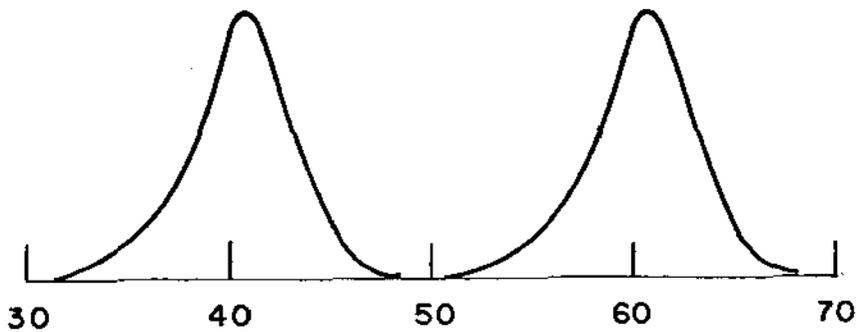


Fig. 2.5 Duas densidades tendo diferentes
locações.

A *média* μ : uma medida de locação - a medida mais útil para o nosso propósito é a média da função densidade. Para uma população que contenha um número infinito de observações, a média geralmente é definida em termos de uma integral. Nós evitamos essa sofisticação supondo que a população contém N observações e que N é muito grande. Supomos que y_1, y_2, \dots, y_N são as observações abrangendo a população e nós denotamos a média por μ (a letra grega μ). Então

$$\mu = \frac{y_1 + y_2 + \dots + y_N}{N} = \frac{\sum_{i=1}^N y_i}{N}$$

A média μ localiza fisicamente aquele eixo vertical em torno do qual a função de densidade balançaria (o centro de massa).

O desvio padrão σ : uma medida de dispersão - A média μ informa acerca da locação, mas outros tipos de informação também são importantes; por exemplo, se soubéssemos apenas que o rendimento médio do processo é 50%, o que poderia ser afirmado acerca de o rendimento ser 50% durante todo o processo ou de ser 100% em metade do processo e zero na outra metade? Certamente será importante para compreender o que é a população em questão se, além de sua média, soubermos alguma medida de dispersão. Na fig. 2.6, nós vemos duas variáveis aleatórias que tem a mesma locação, mas uma é mais dispersa que a outra.

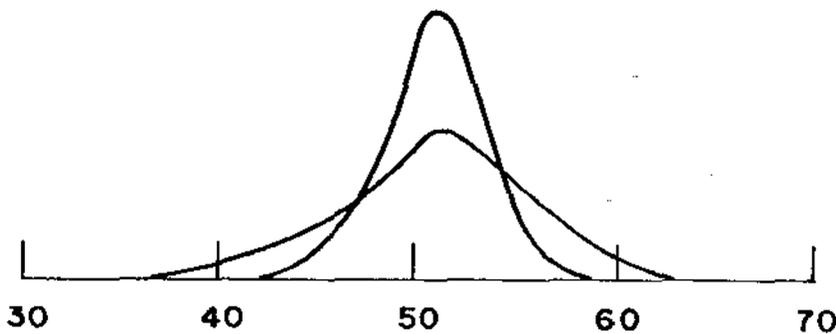


Fig. 2.6 Duas densidades tendo diferentes dispersões

Como uma medida de dispersão, empregamos a média dos quadrados dos desvios em torno da média:

$$\sigma^2 = \frac{(y_1 - \mu)^2 + (y_2 - \mu)^2 + \dots + (y_N - \mu)^2}{N} = \frac{\sum_{i=1}^N (y_i - \mu)^2}{N}$$

Ela é chamada a variância da população. A raiz quadrada positiva dessa quantidade, que tem as mesmas unidades que as observações, é

chamada desvio padrão, e, como nós já indicamos, é denotado pela letra grega sigma. Então,

$$\sigma = \sqrt{\frac{\sum_{i=1}^N (y_i - \mu)^2}{N}}$$

A média e a variância de uma variável aleatória nos informam as características importantes de locação e dispersão, respectivamente. Entretanto elas não caracterizam de todo a função densidade dessa variável aleatória. É necessário saber algo a respeito de sua forma, caso queiramos fazer afirmações probabilísticas precisas. Temos sorte de podermos predizer, aproximadamente, a forma que, via de regra, ocorrerá nas situações com que quase sempre nos depararemos em experimentos industriais.

Variável aleatória Normal - Conforme mencionamos, as observações em um processo industrial variam inevitavelmente. Os desvios $e = y - \mu$ da média verdadeira μ podem ser provenientes de uma grande variedade de fontes, incluindo, por exemplo, erros analíticos, variações na qualidade da matéria prima e aditivos, erros em medida de quantidades do produto, erros nas medidas de taxas de fluxo, vazamento no encanamento, efeito de mudança de temperatura no ambiente, erros de medida de peso do produto final, etc. O erro global e que ocorre em uma observação é, necessariamente, alguma função

$$e = f(e_1, e_2, \dots, e_m)$$

dos erros componentes, que são denotados por e_1, e_2, \dots, e_m . Se

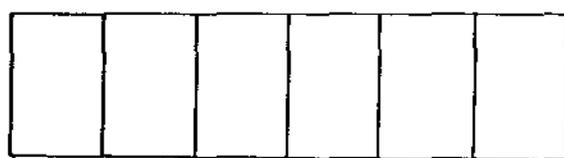
os erros componentes variam regularmente em pequenas amplitudes, isso pode ser escrito aproximadamente por

$$e = K_1 e_1 + K_2 e_2 + \dots + K_m e_m ,$$

onde os K_i são constantes. Há um teorema da estatística matemática, chamado teorema central do limite onde, grosseiramente falando, se diz que: sob a influência proveniente de várias fontes de erro de magnitude compável , a função de densidade de e tende à da Normal quando o número de componentes é aumentado, independente de qual densidade os e_i individuais possam ter. Para ilustração, a função densidade de probabilidade de uma variável aleatória Normal é mostrada na fig. 2.7 c, e tem um número de propriedades que discutiremos resumidamente.

Um exemplo do efeito do teorema central do limite é ilustrado na fig. 2.7. Suponha que uma única componente de erro e_i pudesse tomar os valores $-5, -3, -1, 1, 3, 5$ com igual probabilidade. Então a ocorrência de erros aleatórios dessa fonte poderia ser simulada pelo arremessmo de um dado com $-5, -3, -1, 1, 3, 5$ inscritos nas seis faces. A função de probabilidade do erro dessa única fonte poderia ser representada pela fig. 2.7 a. A função pro babilidade dos erros agregados $e = e_1 + e_2 + e_3 + e_4$ para quatro componentes igualmente importantes e_1, e_2, e_3, e_4 , cada um tendo esse mesmo tipo de distribuição, poderia ser simulado pela soma total das quatro faces de dados do tipo mencionado, e é mostrado na fig. 2.7 b. A função de probabilidade do erro agregado é muito se-

melhante à forma da densidade da Normal mostrada por comparação na fig. 2.7 c. Quanto mais erros componentes são incluídos, mais a densidade de sua soma se aproxima da Normal. Na prática, a densidade do erro não será exatamente Normal. Não obstante, tal suposição fornece uma aproximação que é bem adequada para nossos propósitos.



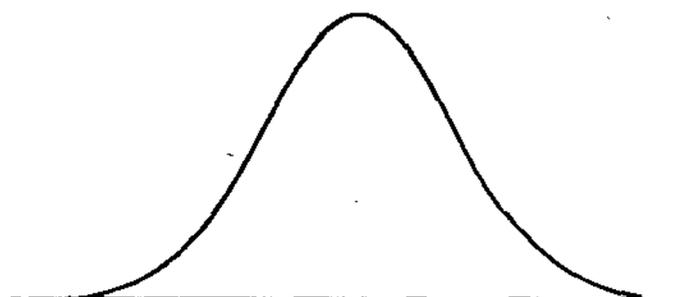
-5 -3 -1 1 3 5

(a)



-20 -10 0 10 20

(b)



(c)

Fig. 2.7 O efeito do teorema central do limite e a densidade Normal:

- a) função de probabilidade de erro e_1 proveniente de uma única fonte.
- b) função de probabilidade do erro $e = e_1 + e_2 + e_3 + e_4$ proveniente de quatro fontes independentes.
- c) função de densidade Normal.

Qualquer variável aleatória com função densidade Normal é completamente especificada por duas medidas: a média μ e o desvio padrão σ . A fig. 2.8 ilustra esse ponto, onde vemos que a Normal é simétrica em torno de sua média μ . Também veremos como o desvio padrão σ mede a sua dispersão. O valor σ mede a distância do seu centro até aquele ponto da curva chamado ponto de inflexão. O

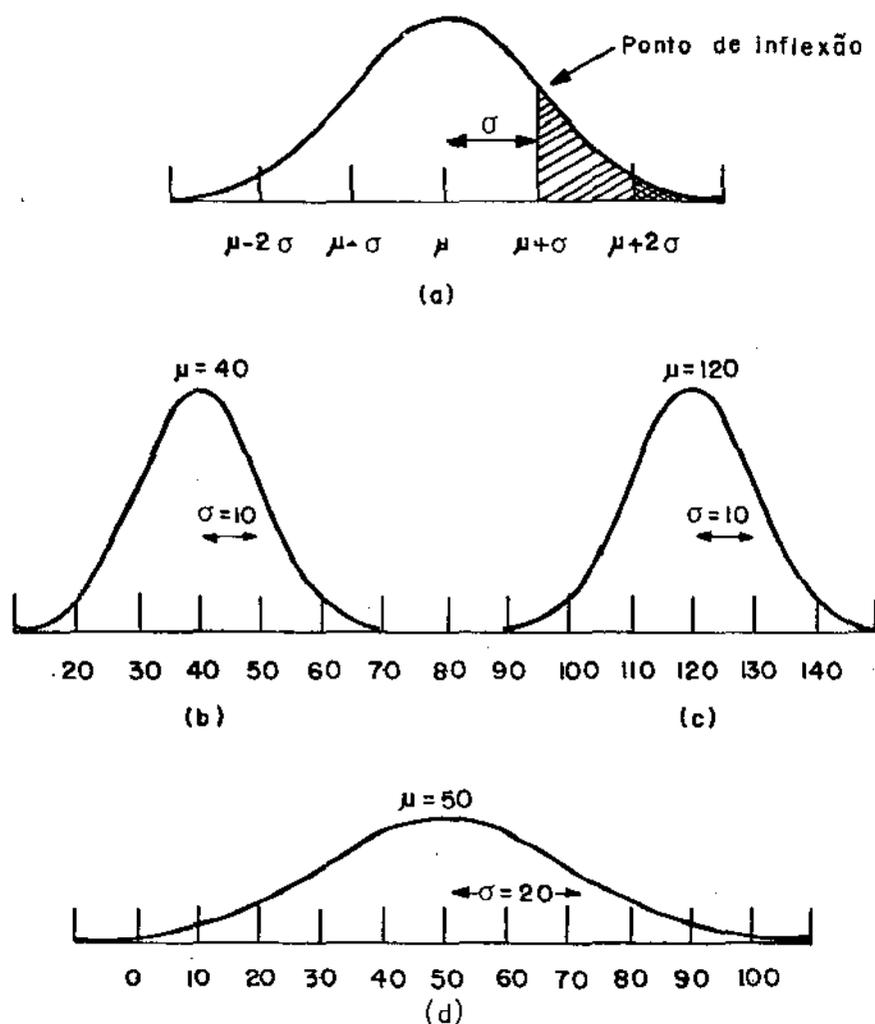


Fig. 2.8 Função Densidade Normal

ponto de inflexão é aquele onde a curva muda de convexa para côncava. As Figs. 2.8 b, 2.8 c e 2.8 d mostram densidades com médias e desvios padrões respectivamente dados por $\mu = 40, \sigma = 10$; $\mu = 120, \sigma = 10$; $\mu = 50, \sigma = 20$, e que tem a mesma forma.

Conforme temos visto, probabilidades podem ser medidas por áreas sob a curva da função densidade. Para a Normal, cerca de um sexto das observações na população excederá a média por um desvio padrão ou mais. Isso significa que a área sob a curva para a direita da linha vertical marcada com $\mu + \sigma$ contém cerca de um sexto da área total sob a curva (fig. 2.9 a).

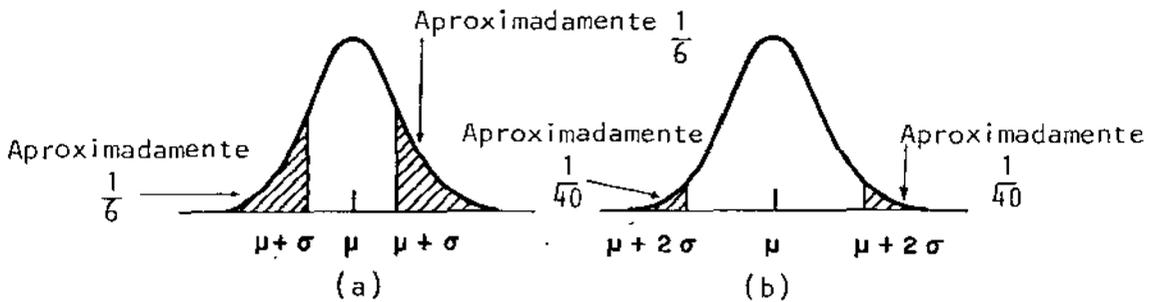


Fig. 2.9 Áreas sob a função densidade Normal.

Uma outra maneira de dizer isso é que a probabilidade de uma única observação y exceder o valor $\mu + \sigma$ é cerca de um sexto. Por causa da simetria, também é verdade que cerca de um sexto das observações tem um valor menor que $\mu - \sigma$, com afirmações inteiramente paralelas.

Aproximadamente $1/40$ das observações em uma população normal excedem a média por dois desvios padrões ou mais. Isso

nifica que a área sob a curva e para a direita da linha vertical marcada por $\mu + 2\sigma$ contam cerca de $1/40$ da área total sob a curva. Uma outra maneira de dizer isso é que a probabilidade de que qualquer observação simples y exceda o valor $\mu + 2\sigma$ é cerca de $1/40$. Então, para as funções de densidade (b), (c) e (d) da fig. 2.8, as probabilidades de se observar um valor que excede 60, 140 e 90, respectivamente, tem para cada uma delas um valor muito próximo a $1/40$ ou 2,5%.

Se nós considerarmos desvios da média em ambas as direções, nós podemos dizer que, para uma variável aleatória Normal, cerca de $1/3$ ($= 1/6 + 1/6$) das observações desviam da média, em uma ou outra direção, por um desvio padrão ou mais. Também cerca de $1/20$ ($= 1/40 + 1/40$) desviam da média em uma direção ou em outra por dois desvios padrões ou mais. Quando ambas as direções são consideradas, as probabilidades são ditas bi-caudais, caso contrário, mono-caudais.

Frequentemente nós precisamos de probabilidades associadas com desvios da média diferentes daqueles de um ou dois desvios padrões.

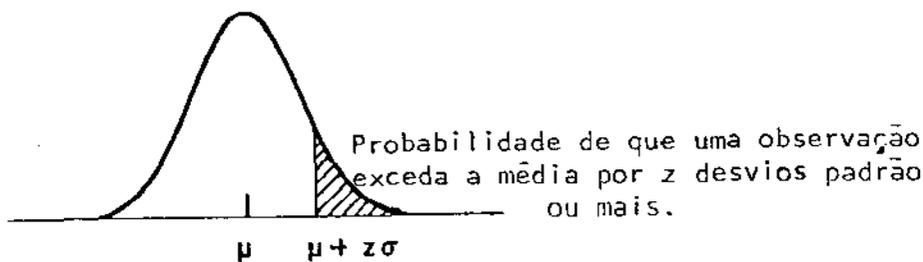


Fig. 2.10 Probabilidade mono-caudal observada na tabela 2.2

A tabela 2.2 mostra a probabilidade, expressa com um decimal, de uma observação individual y se afastar da média μ na direção positiva por z desvios padrões ou mais, para vários valores positivos de z . Então a tabela registra a área sombreada na fig. 2.10, e for

TABELA 2.2 - Distribuição Normal (mono-caudal)¹

Proporção (A) da área total que está à direita da ordenada $x = \mu + z\sigma$ [$z = (x - \mu) / \sigma$]

Desvio	Prefixo	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	Prefixo	Desvio
0.0	0.5	000	960	920	880	840	801	761	721	681	641	0.4	0.0
0.1	0.4	602	562	522	483	443	404	364	325	286	247	0.4	0.1
0.2	0.4	207	168	129	090	052	013	974	936	897	859	0.3	0.2
0.3	0.3	821	783	745	707	669	632	594	557	520	483	0.3	0.3
0.4		446	409	372	336	300	264	228	192	156	121	0.3	0.4
0.5	0.3	085	050	015	981	946	912	877	843	810	776	0.2	0.5
0.6	0.2	743	709	676	643	611	578	546	514	483	451	0.2	0.6
0.7		420	389	358	327	296	266	236	206	177	148	0.2	0.7
0.8	0.2	119	090	061	033	005	977	949	922	894	867	0.1	0.8
0.9	0.1	841	814	788	762	736	711	685	660	635	611		0.9
1.0		587	562	539	515	492	469	446	423	401	379		1.0
1.1		357	335	314	292	271	251	230	210	190	170	0.1	1.1
1.2	0.1	151	131	112	093	075	056	038	020	003	985	0.0	1.2
1.3	0.0	968	951	934	918	901	885	869	853	838	823		1.3
1.4		808	793	778	764	749	735	721	708	694	681		1.4
1.5		668	655	643	630	618	606	594	582	571	559		1.5
1.6		548	537	526	516	505	495	485	475	465	455		1.6
1.7		446	436	427	418	409	401	392	384	375	367		1.7
1.8		359	351	344	336	329	322	314	307	301	294		1.8
1.9		287	281	274	268	262	256	250	244	239	233		1.9
2.0		228	222	217	212	207	202	197	192	188	183		2.0
2.1		179	174	170	166	162	158	154	150	146	143		2.1
2.2		139	136	132	129	125	122	119	116	113	110	0.0	2.2
2.3	0.0	107	104	102	990	964	939	914	889	866	842	0.00	2.3
2.4	0.00	820	798	776	755	734	714	695	676	657	639		2.4
2.5		621	604	587	570	554	539	523	508	494	480		2.5
2.6		466	453	440	427	415	402	391	379	368	357		2.6
2.7		347	336	326	317	307	298	289	280	272	264		2.7
2.8		256	248	240	233	226	219	212	205	199	193		2.8
2.9	0.00	187	181	175	169	164	159	154	149	144	139	0.00	2.9

1. Adaptado de Davies (1956), condensado e adaptado de Pearson e Hartley (1954).

nece uma probabilidade mono-caudal. Nota-se que quando $z=1$, nos temos $p = 0,1587$, que é aproximadamente um sexto, e quando $z=2$, nós obtemos $p = 0,0228$, que é cerca de $1/40$, conforme dito anteriormente.

A tabela pode ser usada para outros propósitos; por exemplo, se nós precisarmos da probabilidade de que uma observação caia entre z_1 e z_2 desvios padrões da média, nós queremos a área sombreada na fig. 2.11 a. Essa área pode ser encontrada na tabela da Normal, como a diferença entre a área sombreada na fig. 2.11 b e a área sombreada na fig. 2.11 c. Neste exemplo, os desvios são ambos positivos. Quando são negativos, o mesmo cálculo pode ser feito como se ambos fossem positivos.

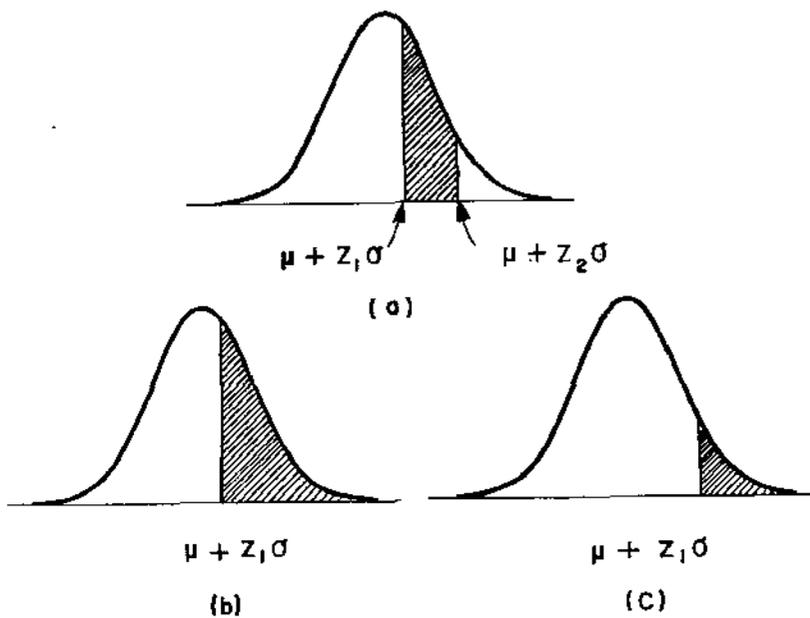


Fig. 2.11 Área (a) como a diferença entre áreas (b) e (c).

A probabilidade de que uma observação caia entre $\mu - z_1\sigma$ e $\mu + z_2\sigma$ é obtido somando as áreas obtidas da tabela, como ilustrado na fig. 2.12 b e 2.12 c.

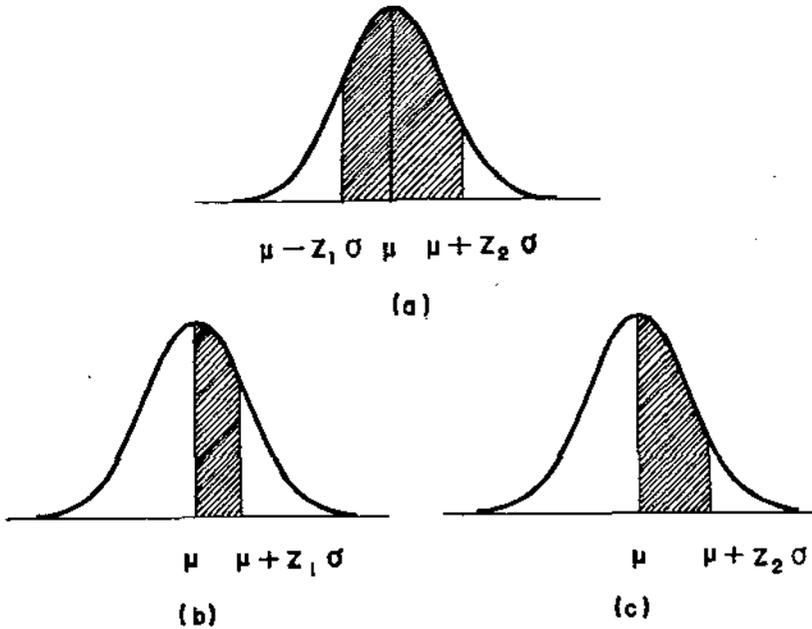


Fig. 2.12 A área sombreada em (a) é a soma das áreas sombreadas em (b) e (c).

Estimativa da média e desvio padrão - Vemos que para ter completo conhecimento de uma população com densidade Normal, nós temos de conhecer somente sua média μ e seu desvio padrão σ . Também chamamos a atenção para o fato de que na prática, nunca temos disponível a população dos resultados e assim nós nunca conhecemos μ e σ . O que nós temos é uma amostra de umas poucas observações com as quais nós gostaríamos de fazer afirmações probabilísticas acerca

da população ou equivalentemente (se assumida normal) acerca de μ e σ . Nós podemos fazer isso calculando estimativas de μ e σ através de nossa amostra limitada de observações.

Suponhamos que nós tenhamos uma pequena amostra de n observações Y_1, Y_2, \dots, Y_n . Então, uma estimativa da média μ da população é fornecida pela média amostral.

$$\bar{y} = \frac{Y_1 + Y_2 + \dots + Y_n}{n} = \frac{1}{n} \sum_{i=1}^n Y_i$$

E uma estimativa da variância populacional σ^2 é fornecida pela variância amostral

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (Y_i - \bar{y})^2$$

O desvio padrão amostral, chamado S , é a raiz quadrada positiva da variância amostral.

Cálculo de \bar{y} e S - Para ilustrar os cálculos da média amostral e desvio padrão amostral, nós usamos os dados mostrados na tabela 2.3.

	y_i	\bar{y}	$y_i - \bar{y}$	$(y_i - \bar{y})^2$
	66.1	65.2	0.9	0.81
	63.7	65.2	-1.5	2.25
	65.8	65.2	0.6	0.36
	64.8	65.2	-0.4	0.16
	65.7	65.2	0.5	0.25
	66.3	65.2	1.1	1.21
	64.8	65.2	-0.4	0.16
	65.2	65.2	0.0	0.00
	64.3	65.2	-0.9	0.81
	65.3	65.2	0.1	0.01
Soma	652.0	652.0	0.0	$6.02 = \sum_{i=1}^{10} (y_i - \bar{y})^2$

TABELA 2.3 - Cálculo de \bar{y} e s^2 .

Para a média amostral, nós temos

$$n = 10 \quad \sum y_i = 652$$

de forma que $g = \frac{652}{10} = 65.2$

e, para a variância amostral,

$$n-1 = 9 \quad \sum (y_i - \bar{y})^2 = 6.02 \quad s^2 = \frac{6.02}{9} \approx 0,67$$

e portanto $s \approx 0,82$.

Um cálculo curto para estimar o desvio Padrão - Com OPEV, deseja-se simplificar os cálculos ao máximo. Cálculos de \bar{y} são bastante simples, e, embora o cálculo para S não seja difícil, é mais fácil, e suficiente para o nosso propósito, usar o método baseado na amplitude (a diferença numérica entre a maior e a menor observação na amostra). Por exemplo, a amplitude da amostra

20, 4, 37, 15, 18, 39

é $39 - 4 = 35$. (Ver fig. 2.13a). Quando os números negativos ocorrem, o sinal já estará sendo levado em conta. Por exemplo, a amplitude da amostra 20, -4, 37, -15, 17, -39 é $37 - (-39) = 76$ (Ver fig 2.13b).

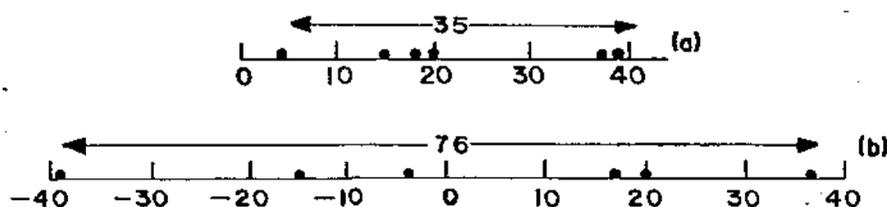


Fig. 2.13 As amplitudes de duas amostras.

Supondo que as observações provêm de uma Normal, nós podemos estimar σ multiplicando a amplitude por um fator w que varia de acordo com o número de observações na amostra. Os valores de w para número de observações entre 2 e 12 são dados na tab.2.4.

Tabela 2.4. Para estimar o desvio padrão de uma Normal, multiplique a amplitude da amostra de n observações pelo correspondente valor de w

n	w^1
2	0.886
3	0.591
4	0.486
5	0.430
6	0.395
7	0.370
8	0.351
9	0.337
10	0.325
11	0.315
12	0.307

¹Note que, para $2 \leq n \leq 10$, w é aproximadamente igual a $1/\sqrt{n}$. É útil tal lembrança, pois permite que os cálculos sejam feitos quando a tabela de w não está disponível.

Nós usamos esse método para as 10 observações da tabela 2.3. A maior observação é 66.3 e a menor é 63.7; a amplitude da amostra é então $66.3 - 63.7 = 2.6$. Da tabela 2.4, $w = 0,325$ quando $n = 10$.

Uma estimativa do desvio padrão da população é, então

$$2.6 (0,325) = 0,84$$

que está bem próximo da estimativa $s = 0.82$ obtido pelo uso de fórmula de s^2 .

Função densidade de Probabilidade para uma média amostral - Quando são calculadas as médias de resultados sujeitos a erro, nós esperamos que a precisão da média aumente quando também é aumentado o número de observações em que se baseia o cálculo. Supondo que as observações são estatisticamente independentes, a maneira como a precisão deve aumentar pode ser estudada considerando a densidade da média amostral. Para compreender o significado da densidade de uma média amostral, vamos mais uma vez imaginar que os resultados de uma população hipotética são escritos em bilhetes de uma urna. Se existe independência para as observações, a tiragem de uma amostra de n observações é simulada pela retirada aleatória dos bilhetes da urna. Suponhamos que nós estejamos interessados na densidade da média de amostras de tamanho $n = 4$.

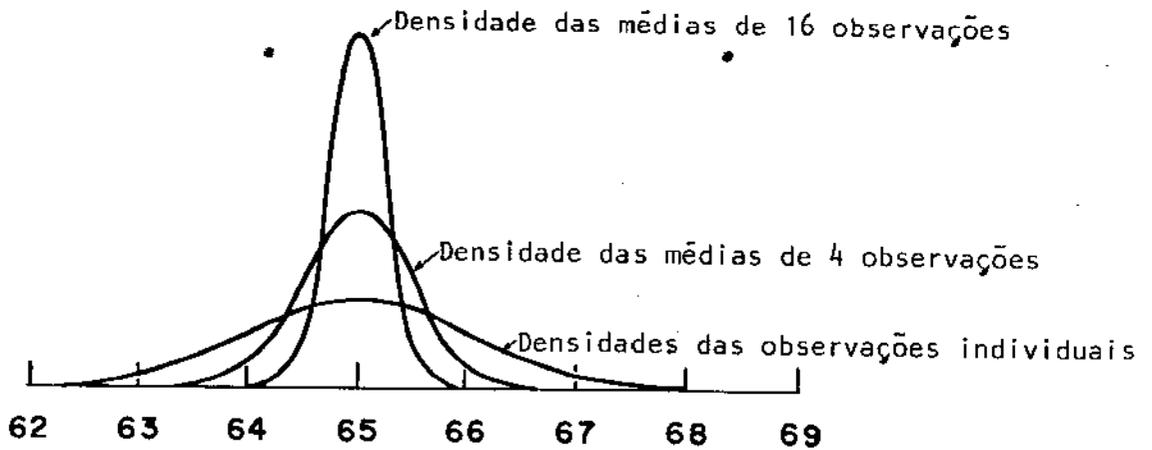
Nós podemos imaginar que tiramos uma amostra de 4 bilhetes, calculamos a média, recolocamos a amostra na população (os quatro bilhetes na urna), tiramos outra amostra de 4 bilhetes, calculamos a média dessa amostra, e assim por diante, até que tenha sido tirado um número muito grande de amostras de tamanho quatro. Suponhamos agora que à medida que cada média amostral é obtida, seu valor é escrito em outro bilhete, e que esses bilhetes com as médias são colocadas em uma segunda urna. Esses bilhetes representam a população das médias de quatro retiradas da primeira população. Que tipo de densidade teria essa segunda população? Como as características dessa densidade amostral dependeriam da original? Vamos supor que a média da população original fosse μ e a sua variância σ^2 . Então, independente da forma da densidade das observações ori

- 1 - A média da função densidade das médias amostrais seria μ , a mesma que a média da população original.
- 2 - A variância da função densidade de probabilidade das médias amostrais seria σ^2/n , um n -ésimo da variância da população original (no caso, $\sigma^2/4$). Em símbolos, se $V(y) = \sigma^2$, então $V(\bar{y}) = \sigma^2/n$, onde V denota variância, y denota uma única observação e \bar{y} denota a média de n observações y_1, y_2, \dots, y_n . Resulta, naturalmente, que o desvio padrão de \bar{y} (a raiz quadrada da variância) será σ/\sqrt{n} .

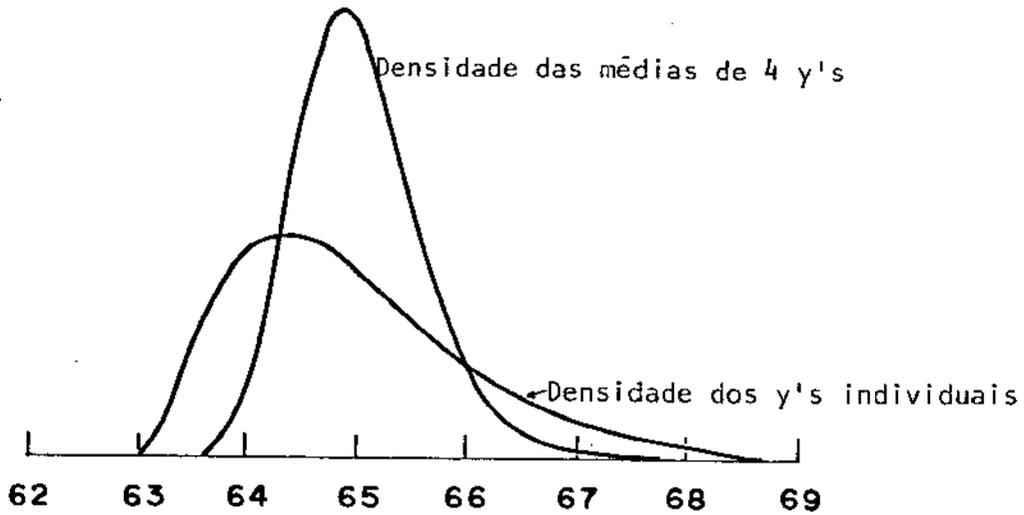
Se a população original tivesse densidade Normal, a média amostral também teria. A medida que é aumentado o número de observações independentes na amostra, a média amostral utilizada como estimativa de μ tem a sua precisão aumentada. Vejamos, como exemplo, o caso em que o rendimento individual em algum processo é uma variável aleatória Normal, com média $\mu = 65$ e desvio padrão $\sigma = 1$.

A fig. 2.14a mostra a função densidade Normal original para os resultados individuais, a função densidade Normal resultante da média de 4 observações (com desvio padrão igual a $1/2$) e da média de 16 observações (com desvio padrão igual a $1/4$).

Na ilustração anterior, a densidade original era assumida Normal. Suponhamos agora que não seja! Então a densidade de \bar{y} , embora tenha média μ e desvio padrão σ/\sqrt{n} , não será exatamente Normal. Entretanto, tenderá rapidamente à forma Normal com o aumento do tamanho da amostra. A razão é que:



(a)



(b)

Fig 2.14

- a) Função densidade Normal para observações individuais, médias de 4 e de 8 observações de um processo de produção.
- b) Densidade de observações individuais e médias de 4 observações.

$$\bar{y} = \frac{y_1 + y_2 + \dots + y_n}{n} = K_1 y_1 + K_2 y_2 + \dots + K_n y_n$$

onde todos os K_i são iguais a $1/n$. Então, o teorema central do limite anteriormente mencionado garante que a densidade de \bar{y} se aproxima da forma Normal com o aumento do tamanho da amostra. Isso é ilustrado na fig. 2.14b. A densidade das observações individuais

(com média $\mu = 65$ e desvio padrão $\sigma = 1$) não é Normal. A correspondente densidade das médias de quatro observações tem a mesma média com metade do desvio padrão, conforme esperado, e sua forma já é bastante próxima à de uma Normal.

Felizmente, em nossas aplicações, nós quase sempre queremos fazer afirmações probabilísticas a respeito de médias e não de resultados individuais. Por causa do efeito observado no teorema central do limite mesmo nos casos em que população dos resultados não é exatamente Normal, sua tabela dá boas aproximações para a densidade das médias.

Média e Variância de Variáveis Aleatórias - Essas médias não são fixas, mas variam aleatoriamente de um ciclo para outro. Quantidades com esse comportamento tem uma função de densidade e são chamadas "variáveis aleatórias". Vamos supor que haja p variáveis aleatórias y_1, y_2, \dots, y_p e p constantes l_1, l_2, \dots, l_p . A expressão

$$L = l_1 y_1 + l_2 y_2 + \dots + l_p y_p \quad 2.1$$

é chamada "combinação linear" das variáveis aleatórias y_1, y_2, \dots, y_p , sendo que L também é uma variável aleatória.

Devemos nos lembrar da importância da suposição de independência estatística. As variáveis aleatórias y_1, y_2, \dots, y_p são estatisticamente independentes se a probabilidade de que y_i exceda algum valor particular não depende dos valores dos outros y 's!.

A média de uma variável aleatória que seja combinação linear de outras é a mesma combinação linear das médias das respectivas variáveis aleatórias, mesmo que elas não sejam independentes. Se

$$L = \ell_1 y_1 + \ell_2 y_2 + \dots + \ell_p y_p$$

então

$$\mu_L = \ell_1 \mu_1 + \ell_2 \mu_2 + \dots + \ell_p \mu_p \quad 2.2$$

Em muitas de nossas aplicações de OPEV nós podemos assumir que os y_i são independentes. Se for possível tal simplificação, então a variância da combinação linear

$$L = \ell_1 y_1 + \ell_2 y_2 + \dots + \ell_p y_p$$

é dada por

$$V(L) = \ell_1^2 \sigma_1^2 + \ell_2^2 \sigma_2^2 + \dots + \ell_p^2 \sigma_p^2 \quad 2.3$$

Frequentemente nós precisamos conhecer a média e a variância da soma ou diferença de duas quantidades que variam aleatoriamente. Sejam y e z duas variáveis aleatórias independentes, com médias μ_y e μ_z e com variâncias σ_y^2 e σ_z^2 . Então,

Soma	Diferença	
$\mu_{Y+Z} = \mu_Y + \mu_Z$	$\mu_{Y-Z} = \mu_Y - \mu_Z$	
$\sigma_{Y+Z}^2 = \sigma_Y^2 + \sigma_Z^2$	$\sigma_{Y-Z}^2 = \sigma_Y^2 + \sigma_Z^2$	2.4

Essas fórmulas são casos especiais da combinação linear das variáveis Y e Z .

$$y - z = (1)y + (-1)z, \text{ tal que}$$

$$\mu_{Y-Z} = (1)\mu_Y + (-1)\mu_Z = \mu_Y - \mu_Z$$

$$V(Y-Z) = \sigma_{Y-Z}^2 = (1)^2\sigma_Y^2 + (-1)^2\sigma_Z^2 = \sigma_Y^2 + \sigma_Z^2 \quad 2.5$$

Erro padrão - Em particular, se as variâncias dos y_i são todos iguais,

$$\sigma_L = (\ell_1^2 + \ell_2^2 + \dots + \ell_p^2)^{1/2} \sigma_y \quad 2.6$$

Se σ_y não é precisamente conhecida, podemos substituí-lo por uma estimativa s_y e nós nos referimos então ao desvio padrão estimado de L

$$\hat{\sigma}_L = (\ell_1^2 + \ell_2^2 + \dots + \ell_p^2)^{1/2} s_y \quad 2.7$$

com o erro padrão e nós o escrevemos como E.P.(L). Para escrever uma expressão para o erro padrão de uma dada quantidade de interes

se, nós procedemos como segue:

- 1 - Escrevemos o desvio padrão da quantidade de interesse em termos do desvio padrão σ de observações individuais (usando 2.3)
- 2 - Substituímos a estimativa de σ nessa expressão. Por exemplo, vamos supor que precisamos do erro padrão da diferença de duas médias amostrais $\bar{y}_2 - \bar{y}_1$, cada qual baseada em n observações independentes, e suponhamos que nós temos uma estimativa s de σ . De 2.3 ou 2.5

$$\sigma_{\bar{y}_2 - \bar{y}_1}^2 = V(\bar{y}_2 - \bar{y}_1) = \frac{\sigma^2}{n} + \frac{\sigma^2}{n} = \frac{2\sigma^2}{n}$$

$$\sigma_{\bar{y}_2 - \bar{y}_1} = \frac{1.414\sigma}{\sqrt{n}} \quad 2.8$$

O erro padrão de $\bar{y}_2 - \bar{y}_1$ é, portanto, $1.414s / \sqrt{n}$, e os limites de dois erros padrão para $\bar{y}_2 - \bar{y}_1$ são $\pm 2 \frac{1.414s}{\sqrt{n}}$.

Testes de significância - Vamos supor que nós tenhamos examinado um produto que há bastante tempo vem sendo fabricado sob condições padrões. Essa experiência anterior nos revela que as observações individuais do rendimento y variam independentemente com uma densidade Normal com média $\mu_0 = 65$ e com desvio padrão $\sigma = 1$.

Vamos admitir que a fábrica paralizou por um período durante o qual é feita uma modificação, na esperança de que o rendimento seja aumentado. Após o reinício da fabricação, são feitas algumas observações para questionar se a respectiva densidade tem

a mesma média $\mu_0 = 65$ como anteriormente ou o valor verdadeiro da nova média μ_1 maior que $\mu_0 = 65$? Vamos supor podermos esperar que o desvio padrão seja essencialmente o mesmo.

Um estatístico expressaria o problema dizendo que há duas suposições acerca do processo modificado: uma hipótese onde a densidade das novas observações é Normal com média $\mu = 65$ e desvio padrão $\sigma = 1$, e uma alternativa, onde a função de densidade se deslocou para uma nova média μ maior que $\mu_0 = 65$, mas com o mesmo desvio padrão $\sigma = 1$. Vamos supor que nós temos uma amostra de $n=4$ observações após a modificação, e que a média dos rendimentos da amostra é $\bar{y} = 65.4$. Nós usaremos essas observações para testar a hipótese de que a média μ da densidade das novas observações era μ_0 , contra a alternativa de que mudou para algum outro valor μ_1 maior que $\mu_0 = 65.0$. Para testar a hipótese, nós a supomos verdadeira. Se, assim procedendo, os resultados se tornarem razoavelmente coerentes, não há nenhuma razão para duvidar da hipótese e nós dizemos que $\mu_1 > \mu_0$ não é estatisticamente significativa. Se, por outro lado, a adoção da hipótese conduzir a resultados pouco aceitáveis, deveremos abandonar a hipótese em favor da alternativa, e nós dizemos que $\mu_1 > \mu_0$ é estatisticamente significativa.

Se a hipótese é olhada como verdadeira, então as quatro observações feitas após a modificação têm cada uma, densidade Normal com média $\mu_0 = 65.0$ e um desvio padrão $\sigma = 1$. A média amostral das $n=4$ observações tem portanto uma densidade Normal com média $\mu_0 = 65.0$ e desvio padrão $\sigma = 0.5$ (já que $\sigma/\sqrt{n} = \sigma/\sqrt{4} = 0.5$).

A fig. 2.15 a mostra a média amostral observada de 65.4 (indicada por um ponto) em relação à função de densidade Normal de referência, centrada em 65.0 e com desvio padrão 0.5. Seria razoável considerar $\bar{y} = 65.4$, com um membro típico dessa população, sob a hipótese? Desde que a média $\bar{y} = 65.4$ caia sob a parte mais central da densidade de referência, nossa resposta deve ser "sim". A divergência aparente de $\bar{y} = 65.4$ de nossa média não é estatisticamente significativa. Se a média \bar{y} fosse igual a 67.0, por exemplo, nossa resposta deveria ser "não", e nós poderíamos ter dito que houve uma discrepância significativa, ou equivalentemente, que a hipótese foi rejeitada.

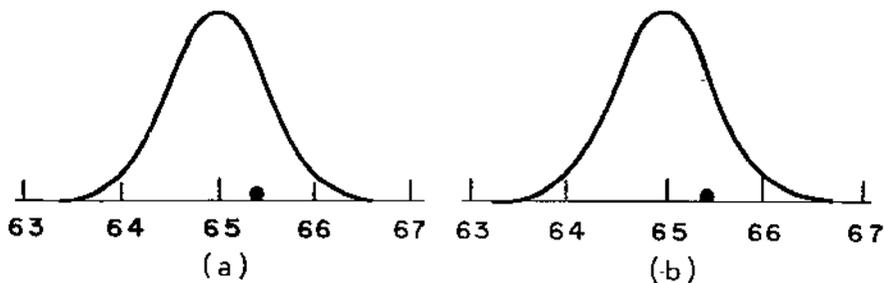


Fig. 2.15 a) densidade de referência
 b) densidade de referência mostrando área de 2.5% na cauda superior.

Para tornar essas idéias mais quantitativas, podem ser introduzidos níveis formais de significância. Isso é geralmente feito considerando a probabilidade específica de conseguir desviar \bar{y} acima de μ_0 por z desvios padrão. Do nosso conhecimento da

Normal, nós sabemos, por exemplo, que as discrepâncias maiores que 1.96 desvios padrão ocorrem 2.5% das vezes. Na fig. 2.15 b, o valor $\bar{y} = 65.98$ (onde $(65.98 - 65)/0.5 = 1.96$) separa na cauda superior uma área de 2.5%, e então corresponde a um valor que é significativamente alto a um nível de significância de 2.5%. O valor $\bar{y} = 66.29$ onde $[(66.29 - 65)/0.5 = 2.58]$ seria significativa ao nível de 0.5%.

Intervalos de Confiança - Testes de significância são ferramentas estatísticas muito usadas. O molde formal do teste de hipótese esboçado anteriormente não é o único que se ajusta a muitas situações reais. O investigador, atento como ele é a possibilidade de sua média \bar{y} observada estar errada, pode estar mais interessado no intervalo em torno de \bar{y} onde se espera estar incluída a média verdadeira μ . O problema de encontrar um tal intervalo pode ser abordado como segue. Nós discutimos anteriormente a questão de testar a hipótese de que a média μ seja igual a $\mu_0 = 65.0$ com um dado nível de significância de, digamos, 2.5%, contra uma alternativa monocaudal de que $\mu > \mu_0$. Equivalentemente, poderíamos ter testado a hipótese de que a média verdadeira μ seja igual a um outro valor hipotético μ_0 contra uma alternativa bi-caudal de que $\mu > \mu_0$ ou $\mu < \mu_0$.

Extendendo a idéia ilustrada na fig. 2.16, nós podemos imaginar a função densidade usada como referência se deslocando para a frente e para trás no eixo horizontal, com sua média μ tomando vários valores hipotéticos. Haveria um intervalo de valores de μ

para os quais pareceria que \bar{y} poderia aceitavelmente provir da referência centrada em μ . Se nós considerarmos como aceitável aquela amplitude de valores de μ que falharam ao fazer o valor de \bar{y} significativo ao nível de 2.5% em cada lado, então essa amplitude é chamada um intervalo de confiança de 95% para μ e seus extremos são chamados limites de confiança 95% para μ .

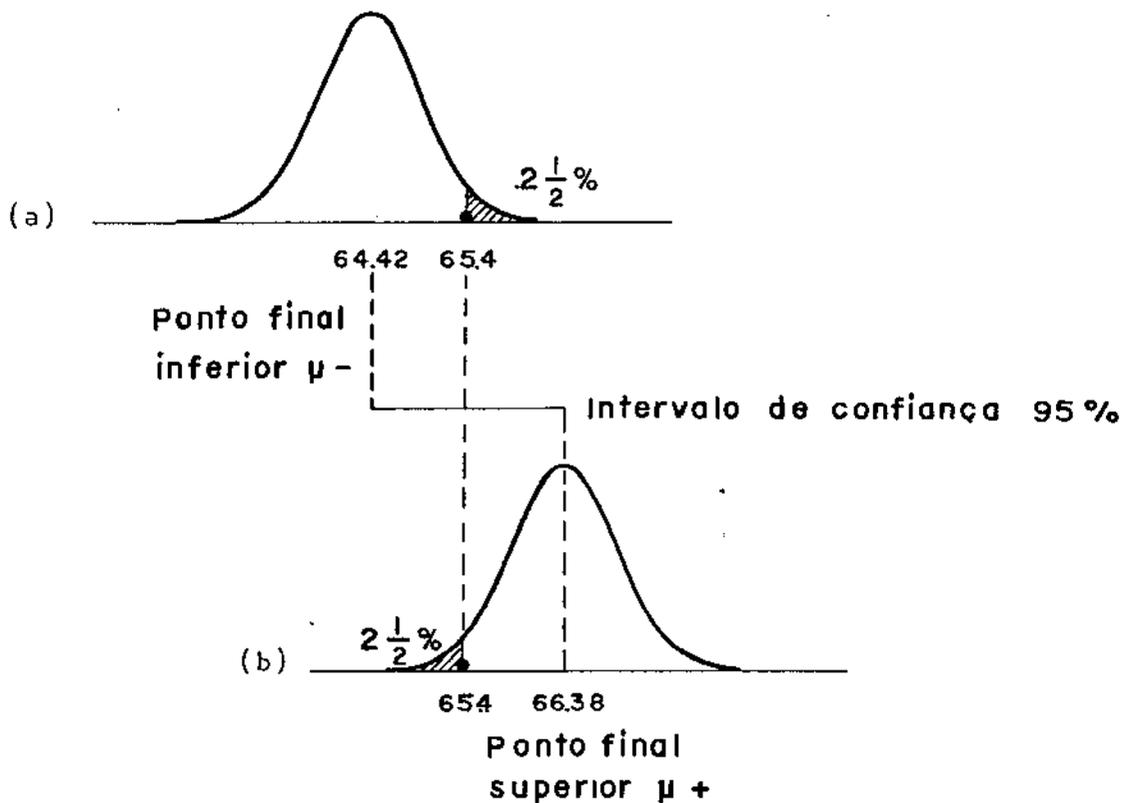


Fig. 2.16 Intervalo de confiança de 95% para μ .

Da fig. 2.16 a e b nós vemos que os limites de um intervalo de confiança poderiam ser os valores μ^- e μ^+ , para os quais

$$\frac{\bar{y} - \mu^-}{\sigma_{\bar{y}}} = +1.96 \quad \text{e} \quad \frac{\bar{y} - \mu^+}{\sigma_{\bar{y}}} = -1.96$$

Então μ^- e μ^+ são os valores

$$\bar{y} \pm 1.96 \sigma_{\bar{y}} .$$

Para nosso exemplo, esses limites são $65.4 \pm 1.96 \times 0.5$, tal que o intervalo de confiança de 95% se estende desde 64.42 até 66.38. Um intervalo de confiança de 99% seria dado por $65.4 \pm 2.58 \times 0.5$.

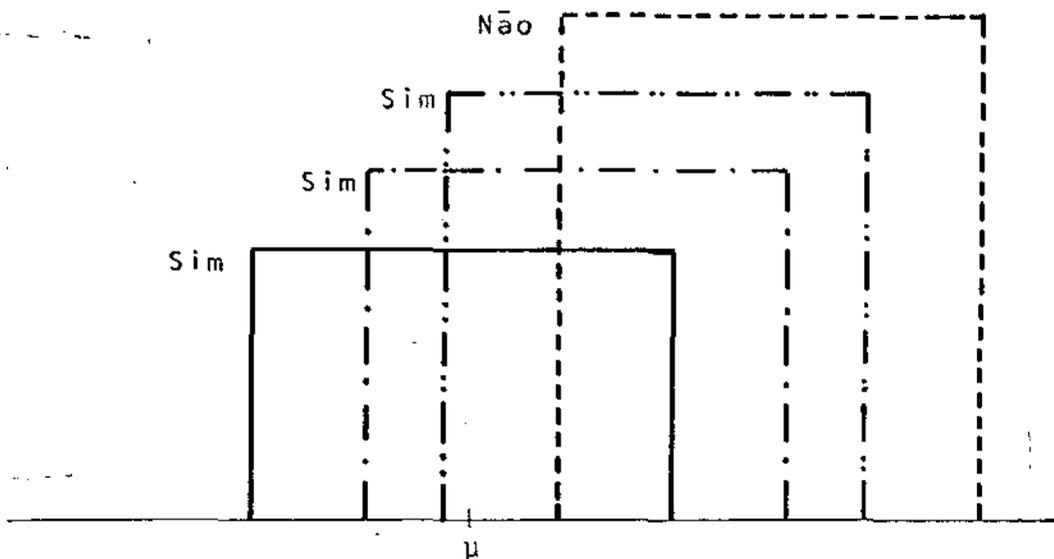


Fig. 2.17 A proporção do tempo em que μ não está contido no intervalo de confiança $\bar{y} \pm z_{(1-\frac{\alpha}{2})} \frac{\sigma}{\sqrt{n}}$ é α (probabilidade de não = α).

Em geral, o intervalo de confiança $100(1-\alpha)\%$ para μ seria dado por

$$\bar{y} \pm z_{(1-\alpha/2)} \sigma/\sqrt{n} ,$$

onde $z_{(1-\alpha/2)}$ é o desvio da Normal que corta a área do lado direito da cauda em $1/2 \alpha\%$. Um tal intervalo de confiança tem a propriedade de que, em repetidas tentativas, cobrirá o valor μ numa proporção $1-\alpha$ do tempo.

3 - OPERAÇÃO EVOLUTIVA

Operação Evolutiva (OPEV) foi proposta por Box(1957) [1] e Box e Draper (1969) [2] como uma técnica de melhoria para rotina de processos industriais. Essa técnica usa um ciclo de pequenas variações num processo de trabalho cuidadosamente planejado; o procedimento de rotina consiste em alterar, por turnos, cada uma das variáveis e, continuamente, repetir o ciclo. Quase sempre, os efeitos dessas mudanças deliberadas nas variáveis são mascarados pelos grandes erros inerentes às unidades de produção em larga escala. Entretanto, desde que a produção continue, um ciclo de variações que não afete a produção de modo significativo pode ser operado seguidamente e, por causa da constante repetição, o efeito de pequenas mudanças pode ser detectado.

As condições do processo que podem ser alteradas são de finidas por variáveis (ou fatores) de duas espécies - qualitativa e quantitativa. Variáveis qualitativas são relacionados com atributos e não podem sofrer variações em uma escala contínua. Temos, como exemplo: tipo de catalizador, tipo de matéria prima, operador, etc. Variáveis quantitativas, como temperatura, pressão, concentração, velocidade de agitação, etc., podem sofrer variações de forma contínua.

Delineamentos fatoriais são configurações particularmente úteis e que podem ser utilizadas tanto para variáveis qualitativas quanto quantitativas. Em um delineamento fatorial 2×2 ou

2^2 examinamos duas variáveis, cada uma delas em dois níveis. Num delineamento fatorial 2^3 , examinamos três variáveis, cada uma delas em dois níveis.

Consideremos o seguinte delineamento fatorial 2×2 :

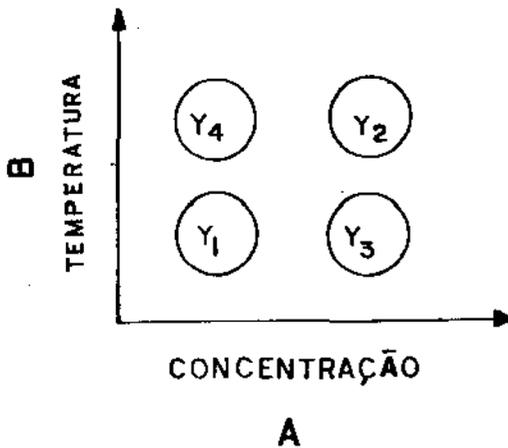


Fig. 3.1 Um delineamento fatorial em variáveis concentração e temperatura

Um efeito simples é a diferença em resposta entre dois níveis distintos de um fator quando o outro fator é mantido constante. No delineamento da fig. 3.1., $y_3 - y_1$ é o efeito simples da concentração no nível mais baixo de temperatura. Um efeito principal é a medida de dois possíveis efeitos simples. Isto é, o efeito de concentração é:

$$A = \frac{1}{2} (y_2 - y_4) + \frac{1}{2} (y_3 - y_1).$$

O efeito principal de temperatura é:

$$B = \frac{1}{2} (y_2 - y_3) + \frac{1}{2} (y_4 - y_1).$$

Interação é a medida da dependência da resposta com a variação si

multânea dos valores de ambos os fatores A e B. É definida como a metade da diferença entre as somas diagonais:

$$AB = \frac{1}{2} (y_1 + y_2) - \frac{1}{2} (y_3 + y_4).$$

O seguinte exemplo ilustra a importância da interação. Suponha que nós obtivemos os seguintes dados de custos unitários para nosso exemplo de concentração e temperatura:

	concentração 1	concentração 2	média
temperatura 2	0.50	0.25	0.375
temperatura 1	0.25	0.50	0.375
média	0.375	0.375	0.375

Se nós limitarmos nossa análise somente aos efeitos principais, teremos:

$$\text{Efeito de temperatura} = \frac{1}{2}(y_2 + y_4 - y_1 - y_3) = 0$$

$$\text{Efeito de concentração} = \frac{1}{2}(y_2 + y_3 - y_1 - y_4) = 0$$

Entretanto:

$$\text{Efeito de Interação} = \frac{1}{2}(y_1 + y_2 - y_3 - y_4) = 0.25$$

Em outras palavras, os efeitos principais são zero devido à média de efeitos iguais de sinais opostos. Neste caso, somente a interaç

ção indica o efeito de dois fatores nos custos unitários.

Erros Padrão para os Efeitos - Os efeitos principais e de interação são sujeitos a erros provenientes da variabilidade dos valores amostrais. Podemos obter o erro padrão para essas quantidades. Por exemplo, a variância do efeito principal é:

$$V(A) = V\left(\frac{1}{2}\bar{y}_2 + \frac{1}{2}\bar{y}_3 - \frac{1}{2}\bar{y}_1 - \frac{1}{2}\bar{y}_4\right)$$

Usando 2.3.

$$V(A) = \left(\frac{1}{2}\right)^2 V(\bar{y}_2) + \left(\frac{1}{2}\right)^2 V(\bar{y}_3) + \left(-\frac{1}{2}\right)^2 V(\bar{y}_1) + \left(-\frac{1}{2}\right)^2 V(\bar{y}_4)$$

Se as amostras são de n observações,

$$V(\bar{y}) = \sigma^2/n,$$

de forma que

$$V(A) = \frac{1}{4} \frac{\sigma^2}{n} + \frac{1}{4} \frac{\sigma^2}{n} + \frac{1}{4} \frac{\sigma^2}{n} + \frac{1}{4} \frac{\sigma^2}{n} = \frac{\sigma^2}{n} .$$

Se nós temos uma estimativa S de σ , então

$$E.P.(A) = S/\sqrt{n}$$

Na prática, interpretamos os efeitos aparentes apenas se seus erros padrões indicarem que tal interpretação é válida. Haverá, por

exemplo, dificuldade em considerar um possível significado da interação $AB = 5.0$ se os limites de 2 E.P. forem ± 7.0 .

Os efeitos principais são obtidos tirando as médias dos efeitos simples. Quando esses efeitos simples indicam fenômenos marcadamente diferentes (quando há uma forte interação), suas médias consideradas sozinhas tem pequeno significado, e devem ser interpretadas conjuntamente com a interação relevante. Um procedimento seguro para interpretar uma análise fatorial é conforme segue: Suponha que tenhamos um efeito principal $A = 5.0 + 2.0$, onde ± 2.0 se refere aos limites de 2 erros padrão. Nós imediatamente inspecionamos todos os efeitos de interação a dois fatores que envolvem A. (Para um modelo fatorial 2^2 haveria somente a interação AB, mas para um fatorial 2^3 , por exemplo, nós deveríamos considerar tanto AB como AC). Se todos os efeitos de interação envolvendo A são desprezíveis ou estão absorvidos pelos respectivos erros padrão, então podemos interpretar o efeito principal diretamente. Podemos dizer que, para B, C, etc., dentro das amplitudes testadas, o efeito de mudar A de um nível baixo para um alto é 5.0 ± 2.0 havendo um efeito substancial de interação $AB=4.0 \pm 2.0$, por exemplo, devemos dizer que o efeito do fator A depende do nível de B.

Referência Periódica às Condições Atuais de Operação - Uma modificação do delineamento fatorial é a inclusão de um 5º ponto experimental. Esse ponto é usado para representar as atuais condições do processo como condições de referência. No início do programa,

tais condições serão aquelas dadas nas especificações de operação. À medida que as melhorias são incorporadas, as condições de operação deverão mudar. Potencialmente, cada nova fase pode estabelecer um novo conjunto de condições. Ao iniciarmos uma nova fase, é útil podermos comparar os resultados diretamente com uma referência apropriada. Isso pode ser feito retornando à referência uma vez por ciclo. Em alguns casos isso significará que um ponto adicional é incluído no delineamento, e em outros casos poderá ocorrer que as condições atuais de operação correspondem a um dos pontos padrão do delineamento. A idéia é ilustrada na fig. 3.2

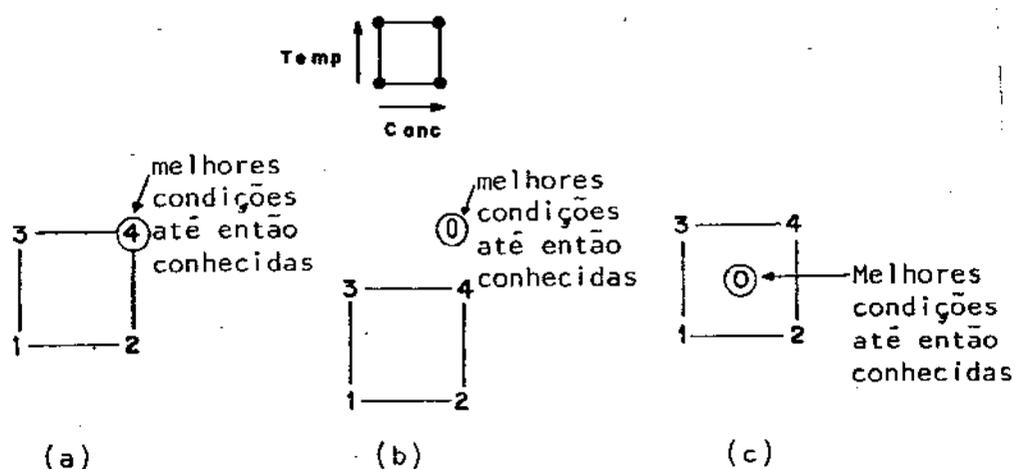
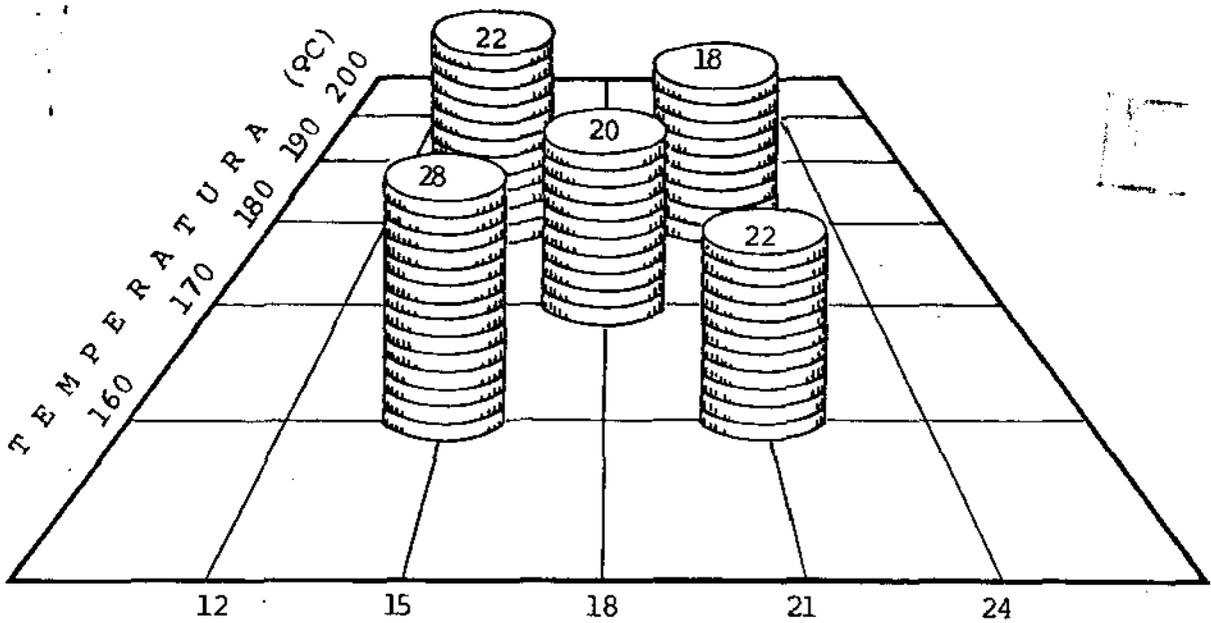


Fig. 3.2 Um delineamento fatorial 2^2 em relação às melhores condições conhecidas.

Para uma resposta em lucro por unidade de material produzido, e considerando os fatores concentração e temperatura conforme a tabela 3.1, teríamos:



C O N C E N T R A Ç Ã O (%)

Fig. 3.3 - Amostras de pontos de uma superfície de resposta (lucro), onde as variáveis (concentração e temperatura) são fixadas em níveis segundo um esquema fatorial 2^2 com ponto central de referência.

Condição número	concentração %	temperatura °C	Lucro por unidade produzida \$
0	18	170	20
1	15	160	28
2	21	180	18
3	21	160	22
4	15	180	22

Tab. 3.1, Lucro (\$), para um delineamento fatorial 2^2 , com condição adicional de referência.

uma configuração conforme a fig. 3.3. Essa configuração nos dá uma

idéia de como se apresenta a superfície de resposta. Não obstante, esse ciclo de variações deve ser repetido um número suficiente de vezes para detectar variabilidade das medições em larga escala. Para se decidir que a condição mais adequada de operação para uma nova fase, é necessário investigar os efeitos de mudanças nos níveis das variáveis.

Quando as condições atuais de operação são incluídos em cada ciclo como uma referência, é instrutivo comparar os rendimento médio encontrado durante a operação do esquema OPEV, com o rendimento que haveria sido obtido se todas as corridas tivessem sido feitas nas condições de referência. Independente de onde o ponto de referência esteja localizado, uma tal comparação é sempre fornecida pelo contraste:

"Resposta média sobre todas as corridas de um esquema OPEV"

menos

"Resposta média nas condições de referência"

Esse contraste é chamado efeito de mudança na média. Suponhamos que n ciclos tenham sido completados tal que, para cada conjunto de condições, podemos calcular uma resposta média baseada em n resultados. Então, como o número de pontos mostrados na fig. 3.2, temos, para o arranjo (a) da fig 3.2.

$$\begin{aligned} \text{Mudança na média} &= \frac{1}{4}(\bar{Y}_1 + \bar{Y}_2 + \bar{Y}_3 + \bar{Y}_4) - \bar{Y}_4 \\ &= \frac{1}{4}(\bar{Y}_1 + \bar{Y}_2 + \bar{Y}_3 - 3\bar{Y}_4) \end{aligned}$$

Para os arranjos b e c da fig. 3.2

$$\begin{aligned} \text{mudança na média} &= \frac{1}{5}(\bar{y}_1 + \bar{y}_2 + \bar{y}_3 + \bar{y}_4 + \bar{y}_0) - \bar{y}_0 \\ &= \frac{1}{5}(\bar{y}_1 + \bar{y}_2 + \bar{y}_3 + \bar{y}_4 - 4\bar{y}_0) \end{aligned}$$

A mudança na média fornece uma medida de custo direto acarretado por obtermos informação em qualquer fase particular. Ela mede se as condições operadas na presente fase renderam um resultado médio melhor ou pior que as condições de referência. Supondo que não tivéssemos operado em um esquema OPEV, e que estivéssemos operando estatisticamente nas condições de referência, a mudança na média avalia o custo direto temporário, se houver, de operar o esquema e obter informação durante a presente fase.

Média na fase - Uma quantidade adicional que será útil ao superintendente do processo é a média na fase. É a resposta média sobre todas as condições operadas na presente fase, sendo calculada como a média dos resultados em todos os conjuntos de condições. Então, para um esquema fatorial 2^2 sem condição adicional de referência, a estimativa da média na fase é $(\bar{y}_1 + \bar{y}_2 + \bar{y}_3 + \bar{y}_4)/4$ e, para o delineamento com condição adicional de referência, a estimativa da média na fase é $(\bar{y}_0 + \bar{y}_1 + \bar{y}_2 + \bar{y}_3 + \bar{y}_4)/5$. A mudança na média é, em todos os casos, igual a "média na fase" - "média nas condições de referência".

Mudança na média para delineamento fatorial 2^2 com ponto central de referência - Quando as condições de referência estão no centro de um esquema fatorial, como mostrado na fig.3.2 C, há um pequeno custo direto. A razão para isso pode ser vista ao se observar a fig. 3.4. Ela mostra a seção de uma possível superfície de custo para a variação na temperatura T. Por causa de sua concavidade, o custo nas condições de referência é ligeiramente mais baixo que a média entre os níveis mais baixo e mais alto de temperatura. Então, operar nos três níveis custa um pouco mais do que manter as condições de referência.

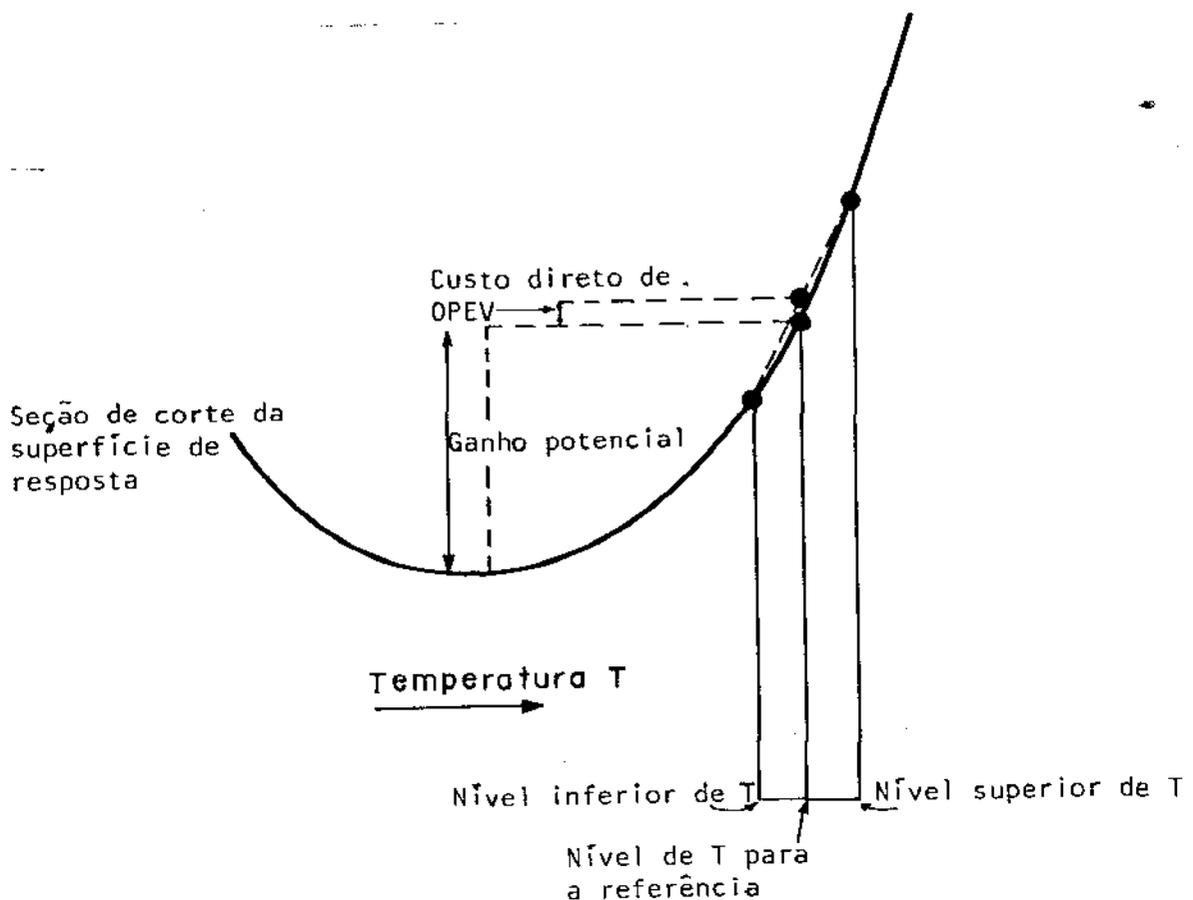
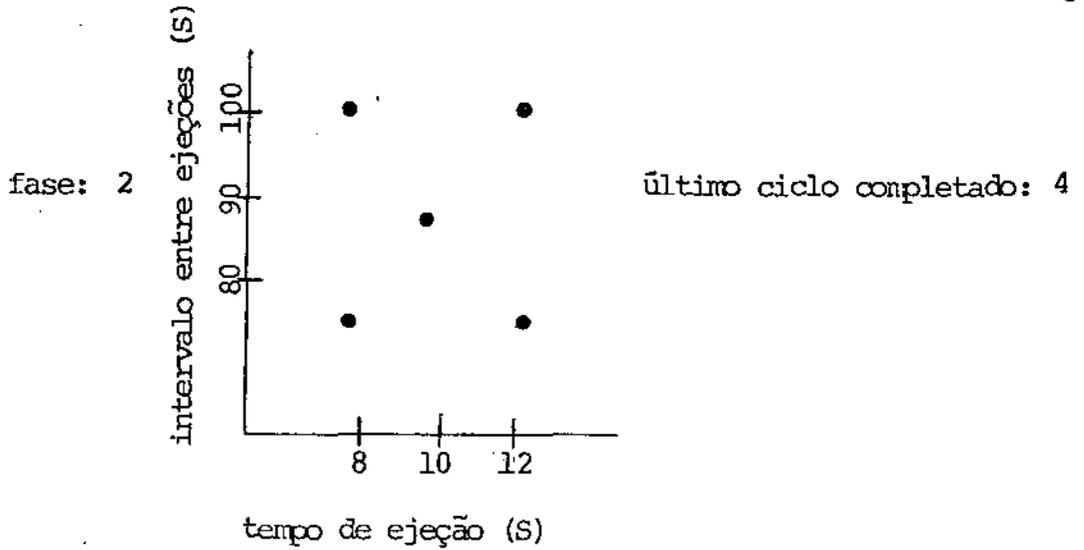


Fig. 3.4 Custo direto de OPEV; seção de uma superfície de resposta.

Como vemos na fig. 3.4, um ganho potencial permanente, via de regra, excederá em muito esse custo temporário. Por causa da concavidade de uma tal superfície, a mudança na média será normalmente positiva e estimará um pequeno acréscimo local no custo. Semelhantemente, por causa da convexidade de uma superfície típica de rendimento, a mudança na média seria quase sempre negativo e estimaria uma pequena perda local no rendimento.

Em geral, para um esquema fatorial 2^2 , com ponto central de referência, a mudança na média é uma medida da curvatura da superfície, à parte do efeito do erro experimental, será positiva próxima a um mínimo e negativa próximo a um máximo. Um efeito de mudança na média que exceda seus dois limites de erro e numericamente seja do mesmo tamanho que os efeitos principais, pode indicar que o máximo ou o mínimo estejam bem próximos.

Quadro de Informações - Com a variação controlada introduzida no processo de fabricação, o efeito de seleção é introduzido agrupando os resultados, que são apresentados continuamente ao superintendente do processo numa forma facilmente compreensível. O fluxo de informações relativo às várias condições de fabricação é sintetizado em um "Quadro de Informações". Tal quadro, cuja disposição é do tipo apresentado na fig. 3.5 " pode ser preenchido com giz em um quadro negro, com lápis de cera em um quadro plástico branco, letras e números magnéticos em um quadro de ferro, etc. O essencial é ser fácil substituir suas informações pelas mais recentes. O superintendente do processo pode ver que peso de evidência existe pa



		Concentrado		Clarificado	
Médias Atuais		8.60	6.89	0.05	0.05
		6.74		0.05	
		7.20	5.73	0.05	0.05
Limites de 2. E.P.		± 0.738		0.007	
Médias na fase		7.033		0.052	
Efeitos com limites		Tempo de Ejeção		-1.592	
para 2. E.P.		Intervalo entre Ejeções		1.281	
		T × I		-0.116	
		Mudança na Média		0.295	
Desvio padrão para as observações individuais		0.738		0.007	

Fig. 3.5 Aparência do Quadro de Informações no fim de 4 ciclos

ra mover o esquema de variáveis, quais tipos de mudanças são indesejáveis pela produção de material com qualidade inferior, quanto custa o esquema etc.

Necessidade de uma comissão OPEV - No capítulo 1 nós vimos como OPEV pode ser comparada com um processo evolutivo natural. Na natureza, as variações ocorrem espontaneamente, mas em nosso processo evolutivo artificial temos de introduzi-las artificialmente. Para o processo se tornar realmente eficiente, mais uma circunstância se faz necessária. Nós devemos estabelecer uma situação real em que estejam sempre surgindo novas idéias úteis. A melhor maneira de induzir uma atmosfera para a geração de tais idéias é reunir, com intervalos convenientes, pessoas com formação técnica diferente, mas que possam opinar sobre peculiaridades do processo. Tais pessoas formariam uma comissão OPEV. A existência de tal comissão garante que os resultados sejam revistos de maneira adequada. Também garante que a relação das idéias para estudo adicionais esteja disponível e seja sempre atualizada. A própria revisão gera novas idéias.

Para uma fábrica química, por exemplo, além do pessoal da planta e do superintendente do processo, candidatos a membros de tal comissão poderiam ser:

- 1) químico ou engenheiro químico que possuam algum conhecimento especializado do tipo particular do processo em estudo. Na revisão de um programa OPEV, devem ser levados em conta propriedades físicas e químicas dos produtos principais e subprodutos. A aná-

lise estatística do "Quadro de informações" revela as mudanças que ocorrem em todas essas propriedades quando são feitas alterações nas variáveis da operação. Tais mudanças podem sugerir, para uma pessoa familiar com a química básica do processo, uma explicação do que está ocorrendo. Essas conjeturas podem gerar idéias para serem exploradas em fases posteriores. Essas idéias podem envolver novas maneiras de manipular variáveis já consideradas ou completamente novas. Às vezes acontece que um único indício precede toda uma sequência de investigações produtivas ao longo de linhas não previstas anteriormente.

- 2) Representante do departamento do controle de qualidade, que terá a responsabilidade de ver quais mudanças incorrem, potencial ou efetivamente, em produção de material fora das especificações oficialmente instituídas.
- 3) Um estatístico, ou pelo menos uma pessoa com algum conhecimento de estatística, pode ajudar a interpretar questões estatísticas mais difíceis. Algumas vezes, quando vários fatores produzem interações complicadas, ele pode sugerir o uso de um programa estatístico mais complexo, empregando as idéias da metodologia de superfície de resposta. Pela sua formação matemática, também pode sugerir direções de melhoria, baseado na medida de curvatura da superfície indicada pela mudança na média e respectivo erro padrão.

Difícilmente, sem uma comissão OPEV atuante, a técnica será continuada e levada a bom termo. É a existência dessa comissão

que faz com que o pessoal de gerência e supervisão estabeleça uma linguagem comum, dissipe desentendimento, e aceite de forma natural a continuação da aplicação OPEV. Por mais ênfase que se dê ela nunca será suficiente para realmente traduzir quão essencial é o bom andamento dessa comissão para o sucesso do uso da Operação Evolutiva.

FOLHA DE TRABALHO PARA PROGRAMAS OPEV COM DUAS VARIÁVEIS

Após o estudo do quadro de informações ao final de um ciclo, o superintendente do processo pode tomar duas decisões básicas:

- 1) Permitir que o esquema prossiga sem mudança e esperar informação adicional do próximo ciclo.
- 2) Modificar a operação de alguma maneira, começando então uma nova fase do esquema OPEV.

O procedimento dos cálculos que são utilizados para esquemas com duas e três variáveis usa o método da máxima amplitude para fornecer uma estimativa facilmente calculável do desvio padrão σ . A seguinte convenção é empregada quando se usa o delineamento 2^2 . Depois do primeiro e segundo ciclo de cada nova fase é empregada uma estimativa a priori de σ obtida das fases anteriores. Após completar o terceiro ciclo, interrompe-se o uso da estimativa a priori e o desvio padrão é estimado dos dados da fase em curso.

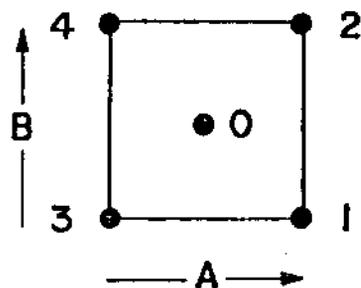
No início da primeira fase, não há nenhuma estimativa "a priori" disponível. Para começar, é calculada uma estimativa do desvio padrão a partir de dados históricos.

Os cálculos diferirão ligeiramente caso um dos conjuntos seja uma condição de referência. Nesse caso, o cálculo da mudança na média se faz necessário, caso contrário não. Se uma condição de referência for incluída, então fará alguma diferença se essa condição é uma do esquema fatorial ou se é adicional.

Vamos passar aos cálculos feitos na folha de trabalho para um delineamento 2^2 com condição adicional de referência. Para ilustração, mostramos nas fig. 3.6, 3.7 e 3.8 os cálculos feitos na 2^a fase de um esquema OPEV aplicado a um processo de obtenção de levedura de cerveja. Nesse processo é investigado o efeito de variar o tempo em que a centrífuga libera líquido clarificado e tempo em que a centrífuga libera líquido concentrado, usando um delineamento fatorial 2^2 com condições adicionais centrais de referência, conforme mostrado no topo da fig. 3.6*.

Nas folhas de trabalho as condições de referência são mostradas no centro do esquema, e no exemplo utilizado para ilustração, as condições de referência estavam assim localizadas. Entretanto, a mesma folha de trabalho e cálculos idênticos poderiam ser usados independentes da localização de um ponto adicional de referência.

* O objetivo é maximizar a proporção entre a concentração do centrifugado e a concentração original da suspensão de células.

FATORIAL 2^2 COM CONDIÇÃO ADICIONAL DE REFERÊNCIA

Ciclo $n = 1$
Resposta *concentrado*

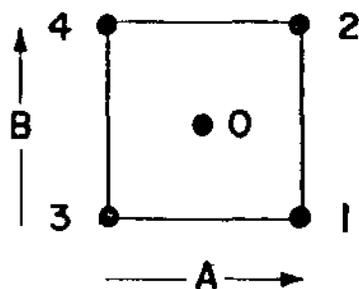
Projeto *Centrifuga*
Fase *2*
Data *6/6/78*

CÁLCULO DAS MÉDIAS					CÁLCULO DO DESVIO PADRÃO	
Condições de Operação	(0)	(1)	(2)	(3)	(4)	Estimativa a priori de $\sigma = *0.362$
(i) Soma do ciclo anterior						Soma anterior S =
(ii) Média do ciclo anterior						Média anterior S =
(iii) Novas observações	6.27	6.89	6.89	5.65	8.14	Novos = amplitudes $\times f_{5,n}$ =
(iv) Diferenças (ii) - (iii)						Amplitude =
(v) Novas somas	6.27	6.89	6.89	5.65	8.14	Nova soma S =
(vi) Novas Médias: \bar{y}_i	6.27	6.89	6.89	5.65	8.14	Nova médias = $\frac{\text{nova soma S}}{n-1}$ =
CÁLCULO DOS EFEITOS					CÁLCULO DOS LIMITES DE 2 ERROS PADRÃO	
Média na Fase = $1/5(\bar{y}_0 + \bar{y}_1 + \bar{y}_2 + \bar{y}_3 + \bar{y}_4)$	= 6.769				Para nova média S: $\pm \frac{2}{\sqrt{n}} S = \pm 0.724^*$	
(A) Efeito = $1/2(\bar{y}_2 + \bar{y}_3 - \bar{y}_1 - \bar{y}_4)$	= -1.243				Para novos efeitos: $\pm \frac{2}{\sqrt{n}} S = \pm 0.724^*$	
(B) Efeito = $1/2(\bar{y}_2 + \bar{y}_4 - \bar{y}_1 - \bar{y}_3)$	= 1.243				Para mudança na média: $\pm \frac{1.79}{\sqrt{n}} S = \pm 0.648^*$	
(AB) Efeito = $1/2(\bar{y}_1 + \bar{y}_2 - \bar{y}_3 - \bar{y}_4)$	= —					
Efeito de mudança na média = Média na fase - \bar{y}_0	= 0.497					

TABELA DE FATORES DE MULTIPLICAÇÃO

n	1	2	3	4	5	6	7	8	9	10
$f_{5,n}$		0.30	0.35	0.37	0.38	0.39	0.40	0.40	0.40	0.41
$1/n$	1.00	0.50	0.33	0.25	0.20	0.17	0.14	0.12	0.11	0.10
$1/(n-1)$		1.00	0.50	0.33	0.25	0.20	0.17	0.14	0.12	0.11
$2/\sqrt{n}$	2.00	1.41	1.15	1.00	0.89	0.82	0.76	0.71	0.67	0.63
$1.79/\sqrt{n}$	1.79	1.26	1.03	0.89	0.80	0.73	0.68	0.63	0.60	0.57

Fig. 3.6 - Folha de Trabalho para um programa OPEV com 2 variáveis.
Cálculos para o ciclo 1.

FATORIAL 2^2 COM CONDIÇÃO ADICIONAL DE REFERÊNCIA

Ciclo $n = 2$
Resposta *concentrado*

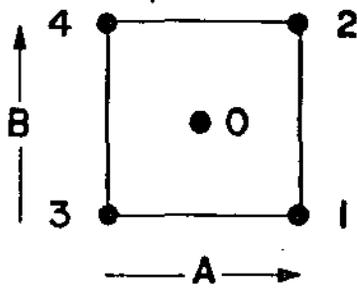
Projeto *Centrifuga*
Fase *2*
Data *8/6/78*

CÁLCULO DAS MÉDIAS					CÁLCULO DO DESVIO PADRÃO	
Condições de Operação	(0)	(1)	(2)	(3)	(4)	Estimativa a priori de σ = <i>*0.362</i>
(i) Soma do ciclo anterior	6.27	6.89	6.89	5.65	8.14	Soma anterior S =
(ii) Média do ciclo anterior	6.27	6.89	6.89	5.65	8.14	Média anterior S =
(iii) Novas observações	7.83	8.45	8.45	5.65	9.07	Novos = amplitudes $\times f_{5,n}$ = <i>0.466</i>
(iv) Diferenças (ii) - (iii)	-1.56	-1.56	-1.56	-	-0.93	Amplitude = <i>1.55</i>
(v) Novas somas	14.10	15.34	15.34	11.30	17.10	Nova soma S = <i>2.466</i>
(vi) Novas Médias: \bar{y}_i	7.05	7.67	7.67	5.65	8.60	Nova médias = $\frac{\text{nova soma S}}{n-1}$ = <i>0.466</i>
CÁLCULO DOS EFEITOS					CÁLCULO DOS LIMITES DE 2 ERROS PADRÃO	
Média na Fase = $1/5(\bar{y}_0 + \bar{y}_1 + \bar{y}_2 + \bar{y}_3 + \bar{y}_4)$	= <i>7.329</i>				Para nova média S: $\pm \frac{2}{\sqrt{n}} S$	= <i>$\pm 0.512^*$</i>
(A) Efeito = $1/2(\bar{y}_2 + \bar{y}_3 - \bar{y}_1 - \bar{y}_4)$	= <i>-1.475</i>				Para novos efeitos: $\pm \frac{2}{\sqrt{n}} S$	= <i>$\pm 0.512^*$</i>
(B) Efeito = $1/2(\bar{y}_2 + \bar{y}_4 - \bar{y}_1 - \bar{y}_3)$	= <i>1.475</i>				Para mudança na média: $\pm \frac{1.79}{\sqrt{n}} S$	= <i>$\pm 0.458^*$</i>
(AB) Efeito = $1/2(\bar{y}_1 + \bar{y}_2 - \bar{y}_3 - \bar{y}_4)$	= <i>0.544</i>					
Efeito de mudança na média = Média na fase - \bar{y}_0	= <i>0.280</i>					

TABELA DE FATORES DE MULTIPLICAÇÃO

n	1	2	3	4	5	6	7	8	9	10
$f_{5,n}$		0.30	0.35	0.37	0.38	0.39	0.40	0.40	0.40	0.41
$1/n$	1.00	0.50	0.33	0.25	0.20	0.17	0.14	0.12	0.11	0.10
$i/(n-1)$		1.00	0.50	0.33	0.25	0.20	0.17	0.14	0.12	0.11
$2/\sqrt{n}$	2.00	1.41	1.15	1.00	0.89	0.82	0.76	0.71	0.67	0.63
$1.79/\sqrt{n}$	1.79	1.26	1.03	0.89	0.80	0.73	0.68	0.63	0.60	0.57

Figura 3.7 - Folha de Trabalho para um programa OPEV com 2 variáveis
Cálculo para o ciclo 2 .

FATORIAL 2^2 COM CONDIÇÃO ADICIONAL DE REFERÊNCIA

Ciclo $n = 3$
Resposta *concentrado*

Projeto *Centrifuga*
Fase *2*
Data *13/06/78*

CÁLCULO DAS MÉDIAS						CÁLCULO DO DESVIO PADRÃO	
Condições de Operação	(0)	(1)	(2)	(3)	(4)	Estimativa a priori de σ	=* 0.362
(i) Soma do ciclo anterior	14.10	15.34	15.34	11.30	17.20	Soma anterior S	= 0.466
(ii) Média do ciclo anterior	7.05	7.67	7.67	5.65	8.60	Média anterior S	= 0.466
(iii) Novas observações	7.20	6.27	6.58	5.96	7.51	Novos = amplitudes x $f_{5,n}$	= 0.598
(iv) Diferenças (ii) - (iii)	-0.15	1.40	1.09	-0.31	1.09	Amplitude	= 1.708
(v) Novas somas	21.30	21.61	21.92	17.26	24.72	Nova soma S	= 1.064
(vi) Novas Médias: \bar{y}_i	7.10	7.20	7.31	5.75	8.24	Nova médias = $\frac{\text{nova soma S}}{n-1}$	= 0.532
CÁLCULO DOS EFEITOS						CÁLCULO DOS LIMITES DE 2 ERROS PADRÃO	
Média na Fase = $1/5(\bar{y}_0 + \bar{y}_1 + \bar{y}_2 + \bar{y}_3 + \bar{y}_4)$	= 7.121					Para nova média S: $\pm \frac{2}{\sqrt{n}} S$	= ± 0.614
(A) Efeito = $1/2(\bar{y}_2 + \bar{y}_3 - \bar{y}_1 - \bar{y}_4)$	= -1.191					Para novos efeitos: $\pm \frac{2}{\sqrt{n}} S$	= ± 0.614
(B) Efeito = $1/2(\bar{y}_2 + \bar{y}_4 - \bar{y}_1 - \bar{y}_3)$	= 1.294					Para mudança na média: $\pm \frac{1.79}{\sqrt{n}} S$	= ± 0.550
(AB) Efeito = $1/2(\bar{y}_1 + \bar{y}_2 - \bar{y}_3 - \bar{y}_4)$	= 0.259						
Efeito de mudança na média = Média na fase - \bar{y}_0	= 0.021						

TABELA DE FATORES DE MULTIPLICAÇÃO

n	1	2	3	4	5	6	7	8	9	10
$f_{5,n}$		0.30	0.35	0.37	0.38	0.39	0.40	0.40	0.40	0.41
$1/n$	1.00	0.50	0.33	0.25	0.20	0.17	0.14	0.12	0.11	0.10
$1/(n-1)$		1.00	0.50	0.33	0.25	0.20	0.17	0.14	0.12	0.11
$2/\sqrt{n}$	2.00	1.41	1.15	1.00	0.89	0.82	0.76	0.71	0.67	0.63
$1.79/\sqrt{n}$	1.79	1.26	1.03	0.89	0.80	0.73	0.68	0.63	0.60	0.57

Figura 3.8 - Folha de Trabalho para um programa OPEV com 2 variáveis
Cálculos para o ciclo 3.

DETALHES DOS CÁLCULOS NAS FOLHAS DE TRABALHO

Examinaremos uma folha de trabalho preenchida após completado o primeiro ciclo. Desde que não há observações disponíveis anteriormente, as linhas (i), (ii) e (iv) dessa folha são deixadas em branco. As "novas observações" do ciclo 1 para cada conjunto de condições de operação são inseridas na linha (iii). Esses mesmos números são também escritos nas linhas (v) e (vi) porque, como há uma única observação, "Novas Somas" e "Novas Médias" são as próprias observações. O desvio padrão σ não pode ser estimado dos dados de um único ciclo, e o que deveria estar no "Cálculo do Desvio Padrão" é deixado em branco, exceto a estimativa previa do desvio padrão (0.362), que é inserida na primeira linha. As entradas sob "Cálculo do Efeitos" podem ser feitas usando os números na linha (vi) que, para o primeiro ciclo, são as próprias observações. As entradas no "Cálculo dos limites de 2E.P." são obtidos substituindo a estimativa a priori de $\sigma = \underline{0.362}$ por S. (O asterísco na folha indica o uso da estimativa prévia). As constantes $2/\sqrt{n}$ e $1.79/\sqrt{n}$, necessárias para esses cálculos, são dadas em uma tabela no lado inferior da folha.

As médias, efeitos e os limites de 2 E.P. podem ser transferidos para o quadro de informações para exibir a situação em que se encontrava o programa OPEV no fim do ciclo 1.

Tão logo os dados do ciclo 1 estejam disponíveis, nós podemos formar uma primeira estimativa de σ baseado nos dados da fase atual. A folha de trabalho preenchida para o segundo ciclo se encon

tra na fig. 3.7. Os cálculos para todos os ciclos subsequentes seguem o mesmo algoritmo:

Passo 1: Os dados das linhas (v) e (vi) da folha de trabalho do ciclo anterior são transferidos para as linhas (i) e (ii) da folha de trabalho do presente ciclo.

Passo 2: As observações para o presente ciclo são inseridas na linha (iii).

Passo 3: Cada entrada na linha (iii) é subtraída da entrada na linha (ii) e o resultado com o respectivo sinal é inserido na linha (iv). Essas diferenças indicam quanto cada observação atual difere da média da experiência até então.

Passo 4: O resultado obtido pela soma da linha (i) com a linha (iii) entra na linha (v).

Passo 5: A "nova média" na linha (vi) é obtida dividindo o registro da linha (v) por n (número do ciclo).

Passo 6: O "cálculo dos efeitos" é realizado no espaço designado usando novas médias da linha (vi).

Passo 7: A maior e a menor diferença observada na linha (iv) são sublinhadas. A diferença entre esses valores sublinhados é a amplitude (1.55 para o ciclo 2) e entra no lado direito da linha (iv).

Passo 8: A amplitude na linha (iv) é multiplicada por um fator f_{5n}^* para obter a estimativa S de σ dada por esse ciclo. Essa estimativa entra na linha (iii) como "Novo S ".

Já que ao fim do ciclo 2 nenhuma estimativa prévia do desvio padrão era disponível a partir dos dados, os itens "soma anterior S " e "média anterior S " são deixados em branco e os registros para "Nova soma S " e "Nova média S " são idênticos àqueles da linha (iii).

Essa estimativa do desvio padrão após o segundo ciclo não é muito confiável e não é usada ainda. Ela é, registrada e combinada com uma estimativa semelhante da fase seguinte antes de ser utilizada para calcular os limites de 2. E.P. Os limites para o ciclo 2 são normalmente obtidos da estimativa a priori de $\sigma=0,362$, conforme indicado por asterisco.

As médias atualizadas, os efeitos e os limites de 2.E.P. mostrando as situações no fim do segundo ciclo são agora transferidos para o quadro de informações.

No fim do terceiro ciclo já podem ser efetuados todos os cálculos para registro na folha de trabalho. Os passos adicionais são como seguem:

Passo 9: O valor "Novo S " na linha (iii) é somado a "soma anterior S " na linha (i) e o resultado é registrado na linha (v), como "No-

* Ver cálculo rápido do desvio padrão amostral no capítulo 2. Em $f_{5,n}$ 5 se refere ao número de condições investigadas e n se refere ao número do ciclo.

va soma S'' . Essa quantidade é então dividida por $n - 1$ para dar uma "Nova média S " na linha (vi).

Passo 10: Os limites de 2.E.P. para as médias e os efeitos usando essa estimativa são obtidos substituindo S na equação pelo valor em "Nova soma S ". As médias, efeitos e seus limites de 2 E.P. são agora transferidos para o quadro de informações para mostrar a situação no fim do ciclo 3.

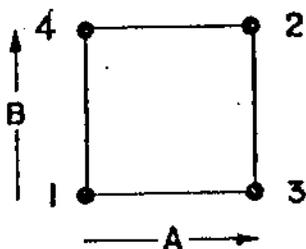
FOLHA DE TRABALHO PARA UM ESQUEMA FATORIAL 2^2
SEM CONDIÇÕES ADICIONAIS DE REFERÊNCIA.

Para um esquema fatorial 2^2 , o cálculo direto dos efeitos é tão simples, que o uso do algoritmo de Yates * deixa de ser vantajoso. Uma folha de trabalho para tal uso é mostrada na fig. 3.9. Nessa ilustração, os dados anteriores para o primeiro ciclo foram utilizados com a omissão das condições centrais, e supondo que as condições para o vértice 2 são as condições de referência. Os cálculos procedem essencialmente como antes. As únicas diferenças são:

- 1 - Os fatores $f_{4,n}$ são aqueles apropriados para conjunto de quatro condições, ao invés de cinco. Valores apropriados de $f_{4,n}$ são mostrados em uma tabela no canto superior esquerdo da folha de trabalho.
- 2 - Como sempre, o efeito de mudança na média é dado por "média sobre todas as condições" - "média nas condições de referência". No exemplo particular, isso poderia corresponder ao contraste $mm = \frac{1}{4}(\bar{y}_1 + \bar{y}_3 + \bar{y}_4 - 3\bar{y}_2)$. Os limites para o erro podem ser calculados como:

$$V(mm) = \frac{1}{16} (V(\bar{y}_1) + V(y_3) + V(\bar{y}_4) + 9V(\bar{y}_2)) =$$

* Ver capítulo 4.

FATORIAL 2^2 SEM CONDIÇÃO ADICIONAL DE REFERÊNCIA

Ciclo $n = 1$
Resposta *concentrado*

Projeto *Centrifuga*
Fase *2*
Data *6/6/78*

CÁLCULO DAS MÉDIAS					CÁLCULO DO DESVIO PADRÃO					
Condições de Operações	(1)	(2)	(3)	(4)	Estimativa a priori de σ	$= * 0.362$				
(i) Soma do ciclo anterior					Soma anterior S	=				
(ii) Média do ciclo anterior					Média anterior S	=				
(iii) Novas observações	6.89	6.89	5.65	8.14	Novo S = amplitude x $f_{4,n}$	=				
(iv) Diferenças (ii) - (iii)					Amplitude	=				
(v) Novas somas	6.89	6.89	5.65	8.14	Nova soma S	=				
(vi) Novas médias	6.89	6.89	5.65	8.14	Nova média S = $\frac{\text{nova soma S}}{n - 1}$	=				
CÁLCULO DOS EFEITOS USANDO O ALGORITMO DE YATES					CÁLCULO DOS LIMITES DE 2 ERROS PADRÃO					
Y	(i)	(ii)	Divisor	Efeito	Para nova média S: $\pm \frac{2}{\sqrt{n}} S$	$= \pm 0.724 *$				
(1) 6.89	12.54	2.57	4	6.89 (Média = Média na Fase)	Para novos efeitos: $\pm \frac{2}{\sqrt{n}} S$	$= \pm 0.724 *$				
(3) 5.65	15.03	-2.49	2	1.245(A)	Para mudança na média: $\pm \frac{1.73}{\sqrt{n}} S$	$= \pm 0.648 *$				
(4) 8.14	1.24	-2.49	2	-1.245(B)						
(2) 6.89	1.25	-0.01	2	— (AB)						
Verificação S de Q	193.1663	772.5052		Mudança na Média $6.89 - 6.89 = 0.$						
	(OK)									
TABELA DE FATORES MULTIPLICADORES										
n	1	2	3	4	5	6	7	8	9	10
$f_{4,n}$		0.34	0.40	0.42	0.43	0.44	0.45	0.45	0.46	0.46
$1/n$	1.00	0.50	0.33	0.25	0.20	0.17	0.14	0.12	0.11	0.10
$1/(n-1)$		1.00	0.50	0.33	0.25	0.20	0.17	0.14	0.12	0.11
$2/\sqrt{n}$	2.00	1.41	1.15	1.00	0.89	0.82	0.76	0.71	0.67	0.63
$1.73/\sqrt{n}$	1.73	1.22	1.00	0.87	0.77	0.71	0.65	0.61	0.58	0.55

Fig. 3.9. Folha de trabalho para um esquema fatorial 2×2 sem condição adicional de referência. Cálculo para o ciclo 1.

$$= \frac{1}{16} \left(\frac{\sigma^2}{n} + \frac{\sigma^2}{n} + \frac{\sigma^2}{n} + \frac{\sigma^2}{n} + 9 \frac{\sigma^2}{n} \right) =$$

$$= \frac{12}{16} \frac{\sigma^2}{n} = \frac{3}{4} \frac{\sigma^2}{n}$$

$$\bar{\alpha} \text{ mm} = \frac{1.73\sigma}{2\sqrt{n}} \quad . \quad \text{Se temos estimativa } S \text{ de } \sigma$$

$$\text{E.P. (mm)} = \frac{1.73\sigma}{2\sqrt{n}}$$

Para um intervalo de confiança 95%, temos que o Efeito de Mudança na média = mudança na média \pm 2 E.P. (mm). (ver Fig.3.9).

Quando as condições de referência são um dos pontos do delineamento fatorial, o efeito de "Mudança da Média" não será independente dos outros efeitos. É, de fato, uma combinação linear explícita daquelas quantidades, o que não significa ser desnecessária sua avaliação. De fato, tal avaliação permite que seja apreciado um aspecto diferente dos dados. A mesma folha de trabalho pode ser usada quando nenhum dos pontos é incluído como referência. Neste caso, nós simplesmente omitimos os cálculos de Mudança na Média e seus limites de 2 E.P.

4 - O DELINEAMENTO FATORIAL 2^3

A dificuldade em operar um esquema fatorial num programa OPEV aumenta com o número de fatores. Entretanto, é perfeitamente admissível um esquema que examine simultaneamente 3 variáveis. Como antes, os fatores envolvidos podem ser qualitativos e quantitativos. As respostas para um delineamento fatorial 2^3 são representados pelos oito vértices de um cubo, numerados de 1 a 8, como mostrado na fig. 4.1.

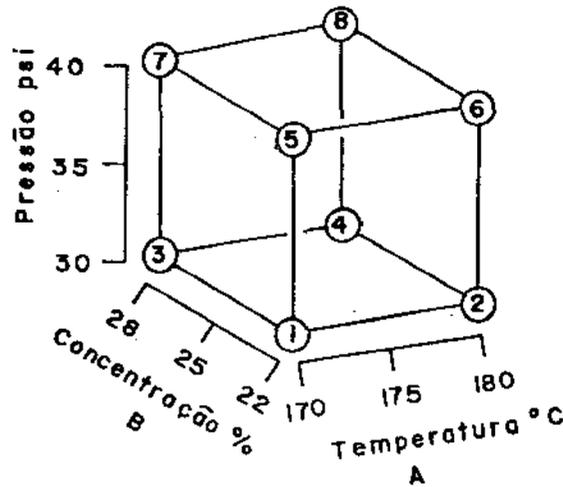


Fig. 4.1 Um delineamento fatorial 2^3

Estimativas dos efeitos principais - Para ilustração vamos considerar um esquema fatorial 2^3 usado para estudar os efeitos de temperatura A, concentração B e pressão C na resposta rendimento de produto. Os oito conjunto de condições, que incluem todas as combinações de pares de níveis, são mostrados na tabela 4.1.

Ordem Padrão	Temperatura A	Concentração B	Pressão C	Rendimento (após 3 ciclos)
1	170	22	30	3.9
2	180	22	30	4.8
3	170	28	30	2.9
4	180	28	30	3.6
5	170	22	40	3.5
6	180	22	40	5.1
7	170	28	40	2.8
8	180	28	40	4.4

TABELA 4.1 - Os oito conjuntos de condições de um delineamento fatorial 2^3 .

Suponha que nós queiramos obter o efeito de mudar a temperatura A de um nível baixo para um nível alto, sob condições fixas de outras variáveis. Uma tal medida é fornecida pela diferença entre os rendimentos da segunda para a primeira corrida,

$$\bar{Y}_2 - \bar{Y}_1 = 4.8 - 3.9 = 0.9.$$

Notar-se-á que nas corridas 1 e 2 tanto a concentração como a pressão são mantidas fixas em seus níveis baixos, e essas condições diferem apenas em seus níveis de temperatura. Duas outras corridas em que somente temperatura é mudada são as corridas 3 e 4, rendendo a segunda estimativa de temperatura,

$$\bar{Y}_4 - \bar{Y}_3 = 3.6 - 2.9 = 0.7.$$

Ao todo há 4 tais pares e esses conjuntos são mostrados na tabela 4.2. Como no caso de duas variáveis, a média desses efeitos simples é chamado efeito principal. Então, o

Medida do Efeito de Temperatura A	Valor	Nível de Concentração B	Nível de Pressão C
$\bar{Y}_2 - \bar{Y}_1$	4.8 - 3.9	0.9	Baixo
$\bar{Y}_4 - \bar{Y}_3$	3.6 - 2.9	0.7	Alto
$\bar{Y}_6 - \bar{Y}_5$	5.1 - 3.5	1.6	Baixo
$\bar{Y}_8 - \bar{Y}_7$	4.4 - 2.8	1.6	Alto

Média = efeito principal de temperatura A = 1.2

TABELA 4.2 - O efeito principal de temperatura A

efeito principal de temperatura é

$$A = \frac{1}{4} [(\bar{Y}_2 - \bar{Y}_1) + (\bar{Y}_4 - \bar{Y}_3) + (\bar{Y}_6 - \bar{Y}_5) + (\bar{Y}_8 - \bar{Y}_7)] = 1.2 \quad 4.1$$

De forma análoga, nós vemos que por comparação dos rendimentos para as corridas 3 e 1, 4 e 2, 7 e 5 e 8 e 6, nós obtemos quatro medidas separadas do efeito da concentração B. Cada uma dessas comparações é novamente feito sob condições idênticas dos outros dois fatores A e C. Especificamente, o efeito principal da concentração B é

$$B = \frac{1}{4} [(\bar{Y}_3 - \bar{Y}_1) + (\bar{Y}_4 - \bar{Y}_2) + (\bar{Y}_7 - \bar{Y}_5) + (\bar{Y}_8 - \bar{Y}_6)] = 0.9 \quad 4.2$$

Da mesma maneira o efeito principal da pressão C é

$$C = \frac{1}{4} [(\bar{y}_5 - \bar{y}_1) + (\bar{y}_6 - \bar{y}_2) + (\bar{y}_7 - \bar{y}_3) + (\bar{y}_8 - \bar{y}_4)] = 0.15 \quad 4.3$$

Interação de dois fatores AB, AC e BC - A interação entre quaisquer dois fatores tais como temperatura A e pressão C é dada por

$$AC = \frac{1}{2} [(\text{efeito médio do fator A no nível alto C}) - (\text{efeito médio do fator A no nível baixo C})]$$

ou, identicamente,

$$AC = \frac{1}{2} [(\text{efeito médio do fator C no nível alto A}) - (\text{efeito médio do fator C no nível baixo A})]$$

Efeito Médio de Temperatura (A)	Pressão C
$\frac{1}{2} [(\bar{y}_2 - \bar{y}_1) + (\bar{y}_4 - \bar{y}_3)] = 0.8$	Baixo
$\frac{1}{2} [(\bar{y}_6 - \bar{y}_5) + (\bar{y}_8 - \bar{y}_7)] = 1.6$	Alto
Interação AC = $\frac{1}{2}(1.6 - 0.8) = 0.4$	

TABELA 4.3 - Cálculo da interação AC entre temperatura e Pressão.

Genericamente, a interação entre dois fatores é definido como metade da diferença entre o efeito médio de um fator, do nível alto para o nível baixo do outro. (veja folha anterior, tabela 4.3).

Se nós considerarmos a manipulação das várias quantidades (tab. 4.3) para chegar a uma estimativa de AC, nós encontramos que

$$AC = \frac{1}{4} [(\bar{Y}_6 - \bar{Y}_5) + (\bar{Y}_8 - \bar{Y}_7) - (\bar{Y}_2 - \bar{Y}_1) - (\bar{Y}_4 - \bar{Y}_3)] = 0.4.$$

Argumento semelhante pode ser usado para definir as interações AB e BC de forma que as 3 interações a dois fatores são:

$$AB = \frac{1}{4} [(\bar{Y}_8 - \bar{Y}_7) + (\bar{Y}_4 - \bar{Y}_3) - (\bar{Y}_6 - \bar{Y}_5) - (\bar{Y}_2 - \bar{Y}_1)] = -0.05 \quad 4.4$$

$$AC = \frac{1}{4} [(\bar{Y}_6 - \bar{Y}_5) + (\bar{Y}_8 - \bar{Y}_7) - (\bar{Y}_2 - \bar{Y}_1) - (\bar{Y}_4 - \bar{Y}_3)] = 0.4 \quad 4.5$$

$$BC = \frac{1}{4} [(\bar{Y}_7 - \bar{Y}_5) + (\bar{Y}_8 - \bar{Y}_6) - (\bar{Y}_2 - \bar{Y}_1) - (\bar{Y}_4 - \bar{Y}_3)] = 0.2 \quad 4.6$$

Uma estimativa da interação a três fatores ABC - A interação a três fatores ABC pode ser definido como metade da mudança na interação AB resultante da mudança AC.

Considerando como os vários resultados contribuíram para os cálculos da interação ABC na tabela 4.4, será encontrado que essa interação pode ser escrita como o contraste entre duas médias de

quatro resultados.

Interação ABC	Nível de C
$\frac{1}{2}(\bar{y}_1 + \bar{y}_4) - \frac{1}{2}(\bar{y}_2 + \bar{y}_3) = -0.1$	Baixo
$\frac{1}{2}(\bar{y}_5 + \bar{y}_8) - \frac{1}{2}(\bar{y}_6 + \bar{y}_7) = 0.0$	Alto
Interação ABC = $\frac{1}{2}[0.0 - (-0.1)] = 0.05$	

TABELA 4.4 - Cálculo da Interação ABC

$$ABC = \frac{1}{4} (\bar{y}_2 + \bar{y}_3 + \bar{y}_5 + \bar{y}_8) - \frac{1}{4} (\bar{y}_1 + \bar{y}_4 + \bar{y}_6 + \bar{y}_7) = 0.05 \quad 4.7$$

Algoritmo de Yates - Os cálculos dos efeitos principais e interações diretamente das fórmulas anteriores pode ser tedioso. Um método simples que pode ser aplicado a qualquer delineamento 2^p foi desenvolvido por Frank Yates. Sua aplicação ao arranjo 2^3 em cada ciclo, está ilustrado na tabela 4.5.

Os 2^p conjuntos de condições para as variáveis A, B, são inicialmente escritas usando o sinal - para indicar o nível baixo e o símbolo + para indicar o nível alto. Para um fator qualitativo como catalizador, nós arbitrariamente assinalamos um sinal - para um e um sinal + para o outro.

Para empregar o algoritmo, devemos garantir que os resultados serão escritos em uma ordem particular. Para realizar essa

Linha	Fator de Combinação			Resposta Média	Coluna			Divisor	Efeito Coluna 3/divisor	Natureza do Efeito
	A	B	C		1	2	3			
1	-	-	-	3.9	8.7	15.2	31.0	8	3.875	Média
2	+	-	-	4.8	6.5	15.8	4.8	4	1.20	A
3	-	+	-	2.9	8.6	1.6	-3.6	4	-0.90	B
4	+	+	-	3.6	7.2	3.2	-0.2	4	-0.05	AB
5	-	-	+	3.5	0.9	-2.2	0.6	4	0.15	C
6	+	-	+	5.1	0.7	-1.4	1.6	4	0.40	AC
7	-	+	+	2.8	1.6	-0.2	0.8	4	0.20	BC
8	+	+	+	4.4	1.6	0.0	0.2	4	0.05	ABC
Verificação S de S				125.8	1000.64					

TABELA 4.5 - Cálculo dos efeitos para rendimento usando Algoritmo de Yates.

ordem, a coluna para o q^o fator é escrita alternando 2^{q-1} sinais negativos e 2^{q-1} sinais positivos até preencher as 2^p linhas. Isso é feito para todos os p fatores. (no nosso caso, $p = 3$ e $2^p = 8$). Então, são inseridas as respostas médias associadas a essas condições. A ordem assim obtida é chamada ordem padrão. As oito linhas da tabela são marcadas por linhas horizontais para formar quatro pares sucessivos da maneira mostrada. Os primeiros quatro números na coluna 1 são as somas dos pares (fig. 4.2). Então, $8.7 = 3.9 + 4.8$, $6.5 = 2.9 + 3.6$, $8.6 = 3.5 + 5.1$, $2.2 = 2.8 + 4.4$. Os últimos quatro números na coluna 1 são as diferenças dos pares. Então, $0.9 = 4.8 - 3.9$, $0.7 = 3.6 - 2.9$, $1.6 = 5.1 - 3.5$, $1.6 = 4.4 - 2.8$. A coluna 2 é agora formada usando os pares na coluna um e com um procedimento análogo aquele utilizado para formar a coluna 1. Final-

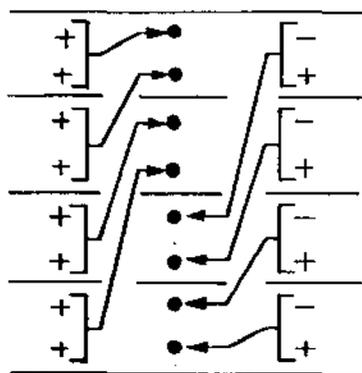


Fig. 4.2 Sequência de preenchimento das colunas, utilizando o algoritmo de Yates.

mente a coluna 3 é calculada da mesma maneira que a coluna 2^{*}. Os registros na coluna 3, divididos pelos divisores^{**} que são 8 para o primeiro registro e 4 para os remanescentes, fornecem os efeitos. Esses efeitos também ocorrem em uma ordem padrão. Para identificar os efeitos escreve-se na última coluna as letras que coincidem com sinal mais. Então, na linha 2 a combinação de fatores mostrada é (+, -) e o efeito correspondente indicado na última coluna é A. A combinação de fatores na linha 4 é AB, e assim por diante. O efeito na primeira linha correspondente a (-, -) é a média das oito respostas. Os efeitos então indicados são idênticos àqueles obtidos previamente.

Verificando o Algoritmo de Yates - a mais segura verificação dos cálculos de Yates é fornecida pelo que se chama Verificação da Soma dos Quadrados ou Verificação S de Q. Suponhamos que a soma de

* Genericamente, teremos tantas colunas quanto ao número de fatores.

** Se observarmos os resultados na 3^a coluna, vemos que na primeira linha temos um valor que é a soma de todas as respostas médias, de forma que um divisor 8 transforma esse valor na média de todas as observações até então. E as demais linhas são combinações de soma e subtração das respostas médias. Um divisor 4 transforma essas combinações nas mesmas que as das equações 4.1 a 4.7. Para um problema genérico com p fatores, os divisores serão 2^p para a primeira linha e 2^{p-1} para as demais.

quadrados dos elementos na coluna dos dados é S . Então, se os cálculos são feitos corretamente, as somas dos quadrados dos elementos das colunas 1, 2 e 3 são exatamente $2S$, $4S$ e $8S$, respectivamente.*

Na prática, nós geralmente procedemos calculando as somas dos quadrados da coluna de dados e da coluna 3 somente. Se a última é exatamente oito vezes a de dados, pode ser assumido que é o cálculo está correto. Caso contrário, a soma de quadrados para a coluna 2 e, se necessário, coluna 1, pode ser verificada para encontrar a coluna onde o erro está localizado. No exemplo anterior, $1000.64 = 8 \times 125,08$ e a verificação foi satisfeita.

Erro Padrão dos Efeitos para Fatoriais a dois Níveis - Se realizarmos n ciclos de um delineamento 2^P com p fatores, cada efeito principal e interação será um contraste entre a média da metade das observações e a média da outra metade.

Se σ^2 é a variância das observações individuais supostas independentemente distribuídas, então a variância de uma média de $\frac{1}{2}(n \times 2^P)$ é $2\sigma^2/n2^P$. Cada efeito principal e interação é a diferença de duas tais diferenças independentes e, portanto, tem variância

$$\frac{2\sigma^2}{n2^P} + \frac{2\sigma^2}{n2^P} = \frac{4\sigma^2}{n2^P} .$$

O erro padrão para cada efeito será obtido tirando a raiz quadrada dessa quantidade e substituindo o desvio padrão σ pela sua estimativa S , conforme os resultados mostrados na Tabela 4.6.

* Genericamente, para a p -ésima coluna, a soma dos quadrados seria $2^P * S$.

Delimitamento	2^P	2^2	2^3
Variância (efeito)	$\frac{4\sigma^2}{n2^P}$	$\frac{\sigma^2}{n}$	$\frac{\sigma^2}{2n}$
E.P. (efeito)	$\frac{2S}{\sqrt{n2^P}}$	$\frac{S}{\sqrt{n}}$	$\frac{S}{\sqrt{2n}}$

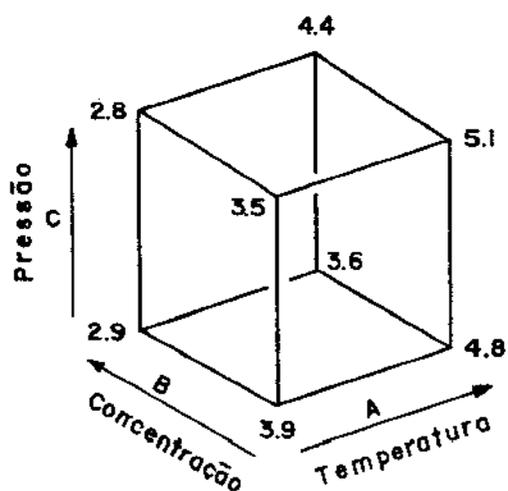
TABELA 4.6 - Variâncias e erros padrão para efeitos Principais e interações estimados de um delineamento fatorial 2^P , após n ciclos.

Interpretação dos Efeitos para o Fatorial 2^3 - Antes de olharmos o efeito principal de um dado fator, devemos ver se há qualquer interação apreciável que envolva esse fator. Se houver, então os fatores que interagem devem ser interpretados juntos ao invés de separadamente.

A tabela 4.7 mostra os resultados conforme deveriam estar dispostos em um quadro de informações. As médias dos rendimentos após 3 ciclos são mostrados nos vértices do cubo. Os efeitos principais e interações estimados, com seus respectivos limites de 2 E.P. são dados abaixo do cubo.

Nos vemos no caso da concentração B que, embora haja um grande efeito principal $B = -0,90$, as interações $AB = -0,05$ e $BC = 0,20$ que envolvam B são pequenas, ambas de magnitude comparável aos seus erros padrão. Por outro lado, embora haja um grande efeito principal, $A = 1.2$ associado à temperatura, há também uma

RENDIMENTO
ULTIMO CICLO COMPLETADO : 3



Temperatura	A	1.20 ± 0.31
Concentração	B	-0.90 ± 0.31
Pressão	C	0.15 ± 0.31
Temperatura \times Concentração	AB	-0.05 ± 0.31
Temperatura \times Pressão	AC	0.40 ± 0.31
Concentração \times Pressão	BC	0.20 ± 0.31
	ABC	0.05 ± 0.31

TABELA 4.7 - Os limites de 2 E.P. Para médias individuais;
 ± 0.44 .

apreciável interação AC = -0,5 entre a temperatura e a pressão. O efeito da temperatura e pressão deve, portanto, ser avaliado conjuntamente, ao invés de individualmente.

Dividindo o Esquema Fatorial 2^3 em 2 Blocos - o uso do delineamento 2^3 nos permite estudar os efeitos principais e interações entre três variáveis com a mesma precisão que seria obtida do mesmo número de corridas aplicadas a somente dois fatores. Por outro lado, obtemos a informação acerca do terceiro fator e suas interações como um "bonus". O mesmo argumento sugere que seria sempre mais vantajoso estudar quatro ou mais fatores simultaneamente, mas a experiência indica que o estudo de três variáveis simultaneamente representa o limite prático sob as circunstâncias de um programa OPEV genuíno.

O argumento de que três variáveis podem ser examinadas tão economicamente quanto duas variáveis, assume que o erro de uma corrida individual é o mesmo, tanto para o delineamento 2^2 quanto para o 2^3 . Essa suposição é menos "inócua" do que aparenta à primeira vista. Será lembrado que estamos operando duas ou três variáveis, diferenças sistemáticas entre ciclos são automaticamente eliminadas. O erro S que estimamos e que determina nossos limites de 2 E.P. é uma estimativa da variação que poderia ocorrer dentro de um ciclo se não fosse feita nenhuma mudança deliberada. Espera-se que tal oportunidade para variação ocorresse mais para um ciclo com $2^3=8$ corridas do que em um ciclo com $2^2=4$ corridas. Felizmente, podemos ganhar todas as vantagens do menor tamanho de ciclo com o delineamento maior 2^3 , operando-o em duas metades, isto é, em dois blocos de quatro corridas cada. Isso pode ser feito de tal maneira que todos os efeitos e interações de interesse sejam calculados das comparações feitas dentro de um bloco.

A idéia pode ser compreendida imaginando a situação onde pudéssemos fazer quatro das oito corridas em um dia, e quatro

corridas no dia seguinte. Podemos propor a questão: "como arranjamos as oito corridas de forma que as possíveis diferenças sistemáticas de um dia para o outro não afetassem as comparações de interações a dois fatores. A interação a três fatores ABC geralmente será pequena e a de menor interesse, sendo que podemos usá-la para acomodar a diferença dia a dia.

Nosso plano de formar blocos consiste, então, de correr 2, 3, 5 e 8 em um dia e 1, 4, 6 e 7 no dia seguinte (fig. 4.3). Então temos:

$$ABC = \frac{1}{4}(y_2 + y_3 + y_4 + y_5 + y_8) - \frac{1}{4}(y_1 + y_4 + y_6 + y_7)$$

$$\text{dia 1} - \text{dia 2} = \frac{1}{4}(y_2 + y_3 + y_5 + y_8) - \frac{1}{4}(y_1 + y_4 + y_6 + y_8).$$

Em outras palavras, o contraste que mede o efeito ABC também mede o efeito do dia 1 - dia 2. Diz-se que esses dois fatores são confundidos. Se pretendermos estimar a interação ABC, o que realmente obteremos é o efeito ABC mais o contraste entre o dia 1 e o dia 2, e será impossível saber quanto do efeito é produzido pela interação e quanto pela diferença de dia a dia. Por esse sacrifício de informação acerca da interação e três fatores, nós podemos manter todos os outros efeitos livres da diferença de um dia para outro. Isto pode ser visto da seguinte maneira: vamos supor que os valores das observações do dia 1 são, em média, x unidades mais altas que aquelas do dia 2. Então, se deixarmos Y_1, Y_2, \dots, Y_8 denotarem os resultados que seriam obtidos se todas

as corridas tivessem sido conduzidas no dia 2, os resultados seriam:

$$\begin{array}{ll} y_1' = y_1 & y_5' = y_5 + x \\ y_2' = y_2 + x & y_6' = y_6 \\ y_3' = y_3 + x & y_7' = y_7 \\ y_4' = y_4 & y_8' = y_8 + x . \end{array}$$

Agora calculando o efeito principal de A a partir desses resultados, teremos:

$$\begin{aligned} A &= \frac{1}{4} (y_2' + y_4' + y_6' + y_8') - \frac{1}{4} (y_1' + y_3' + y_5' + y_7') = \\ &= \frac{1}{4} [(y_2 + x) + y_4 + y_6 + (y_8 + x)] - \frac{1}{4} [y_1 + (y_3 + x) + (y_5 + x) + y_7] \end{aligned}$$

Vemos que dois dos x aparecem no primeiro colchete com sinal $+$ e os outros dois aparecem no segundo colchete com sinal $-$. O efeito de um dia para o outro é cancelado. Isso acontece não somente para o efeito A mas para todos os outros efeitos principais e interações a dois fatores, só não ocorrendo para a interação ABC onde todos os quatro x aparecerão no primeiro colchete. Essa propriedade provém da chamada ortogonalidade do delineamento. Poderíamos (deixar de considerar alguma) interação a dois fatores ou algum efeito principal sem afetar as outras estimativas? É, obviamente, mais sensato sacrificar a informação que se espera ser a menos

importante no uso, a interação ABC.

Embora nossa discussão tenha sido em termos de observações simples de um único ciclo, resultados semelhantes se aplicam caso sejam usadas as médias obtidas dos n ciclos de operação.

Inclusão das condições de refererência com esquemas 2^3 - Para um esquema 2^3 , da mesma maneira que para duas variáveis, é gratificante e útil voltar periodicamente ao conjunto de condições de referência. Como antes, seriam as melhores condições até então em curso. As condições correspondentes a um dos pontos do delineamento, a um ponto fora do delineamento, ou a um ponto interno ao delineamento. Também podemos voltar às condições de referência uma vez por ciclo, ou com maior ou menor frequência. Em qualquer caso, uma comparação de mudança na média, isto é, (média de todas as condições operadas na atual fase do esquema OPEV) menos (média nas condições de referência), pode ser calculada e será estimada como custo temporário, se houver, de informação obtida, em termos de material, qualidade, ou \$.

Num caso especial, onde as condições de referência estão no centro do delineamento, a mudança na média, adicionalmente, fornecerá uma medida da curvatura na superfície de resposta.

Um delineamento fatorial com 2 pontos centrais operados em dois blocos - Um delineamento desse tipo, de interesse particular, é aquele mostrado na fig. 4.3, na qual as oito corridas do delineamento 2^3 são separadas em dois blocos de quatro corridas, como an-

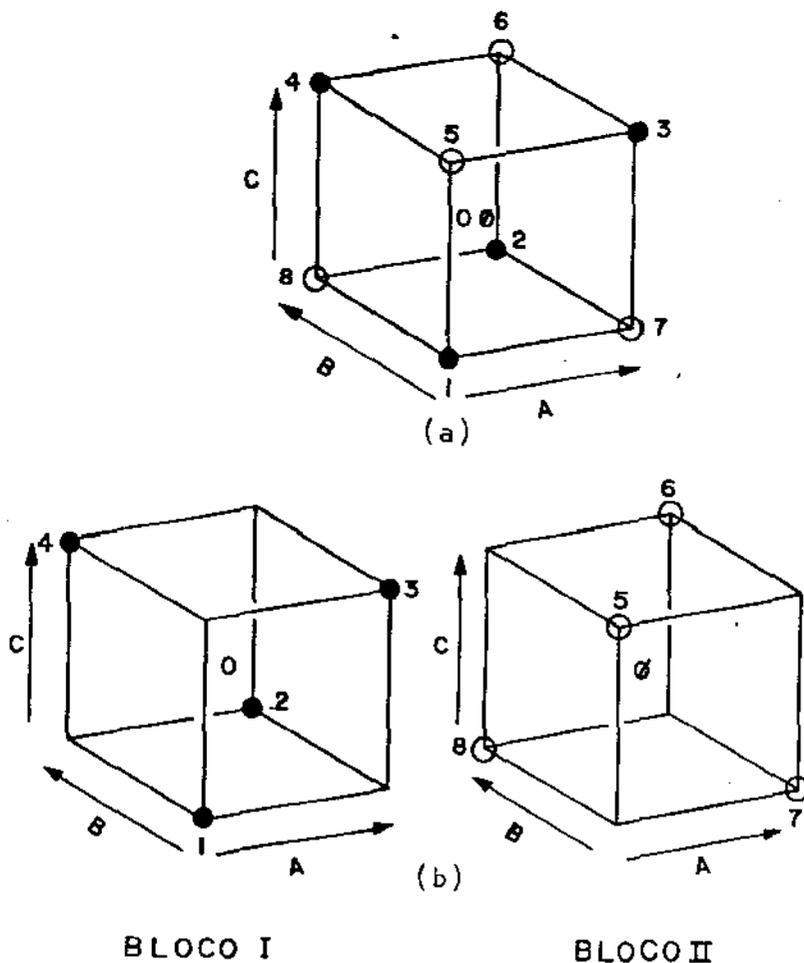


Fig. 4.3 Modelo de um esquema OPEV com três variáveis, e pontos centrais.

- a) um único ciclo em 2 blocos (numeração da sequência)
- b) Os dois blocos do delineamento (numeração da sequência)

teriormente descrito, e um ponto central é associado a cada bloco. Enfatizamos novamente que esse não é o único arranjo que pode ser usado, mas é aquele em que, aparentemente há mais conveniências práticas. As condições no centro do primeiro bloco são assinaladas por 0 (zero) e no segundo por \emptyset (zero barra).

Ainda que consideremos o efeito de possível correlação serial entre erros, muito pouco se perde por usar esse padrão sis-

temático ao invés de um esquema aleatório. Do delineamento da fig. 4.3, podem ser calculadas duas medidas de mudança média, sendo uma para cada bloco. Para o bloco I, em termo de observação simples,

$$\text{mudança na média} = \frac{1}{5} (y_1 + y_2 + y_3 + y_4 - 4y_0)$$

Para o bloco II ,

$$\text{mudança na média} = \frac{1}{5} (y_5 + y_6 + y_7 + y_8 - 4y_0)$$

A média

$$\frac{1}{10} (y_1 + y_2 + y_3 + y_4 - 4y_0 + y_5 + y_6 + y_7 + y_8 - 4y_0)$$

fornece uma medida da mudança global na média, livre dos efeitos dos blocos. É precisamente a mesma interpretação que no caso de duas variáveis e, como sempre, é igual à média na fase menos a média nas condições de referência.

Dois limites do erro padrão para a mudança global na média - Depois de n ciclos a mudança global na média seria

$$-\frac{4}{10} \bar{y}_0 + \frac{1}{10} \bar{y}_1 + \frac{1}{10} \bar{y}_2 + \frac{1}{10} \bar{y}_3 + \frac{1}{10} \bar{y}_4 - \frac{4}{10} \bar{y}_0 + \frac{1}{10} \bar{y}_5 + \frac{1}{10} \bar{y}_6 + \frac{1}{10} \bar{y}_7 + \frac{1}{10} \bar{y}_8$$

Consequentemente, usando equação 2.3, a variância da mudança global da média é

$$\left(\frac{16}{100} + \frac{1}{100} + \frac{1}{100} + \frac{1}{100} + \frac{1}{100} + \frac{16}{100} + \frac{1}{100} + \frac{1}{100} + \frac{1}{100} + \frac{1}{100}\right) \frac{\sigma^2}{n} = 0,4 \frac{\sigma^2}{n}$$

Segue que o desvio padrão para a mudança global na média é $0,632 \sigma/\sqrt{n}$. Se a estimativa de σ é S , os limites de 2 E.P. se-
rão mudança global na média $\pm 2(0,632) \frac{S}{\sqrt{n}}$.

FOLHA DE TRABALHO PARA PROGRAMA OPEV COM TRÊS VARIÁVEIS

Para eliminar as variações estranhas que ocorrem duran-
te a operação de um programa OPEV com três variáveis, a primeira e
a segunda metade de um dado ciclo são arranjadas de tal maneira que
correspondam aos dois blocos de um delineamento 2^3 , conforme já des-
crito.

As folhas de trabalho para os programas com três variá-
veis são muito semelhantes àquelas para programas com duas variá-
veis, mas serão utilizadas duas folhas de trabalho para cada ciclo:
^{uma} para cada bloco. Após o primeiro ciclo, uma nova estimativa do des-
vio padrão σ estará disponível ao final de cada novo bloco (depois
de cada meio ciclo). As novas estimativas dos efeitos estarão dis-
poníveis somente após completado cada ciclo. Como antes, as folhas
de trabalho e os cálculos diferirão ligeiramente dependendo de as
condições de referência estarem incluídas ou não no esquema e se
essas condições correspondem a uma corrida do delineamento ou a uma
corrida adicional.

Inicialmente, veremos com detalhe o uso da folha de tra

balho apropriado quando a condição de referência não é uma das corridas fatoriais. Em nosso esquema de modelo, a condição de referência é operada uma vez em cada bloco, (isto é, duas vezes por ciclo). Também, no exemplo particular usado, essa condição de referência ocorre no centro do delineamento fatorial. As mesmas folhas de trabalho e cálculos seriam, é claro, usados independente de onde se situa a condição adicional de referência. Depois será mostrada a pequena modificação que é necessária se não houver referência.

FOLHA DE TRABALHO PARA UM ESQUEMA FATORIAL DISPOSTO
EN DOIS BLOCOS, COM UMA CORRIDA DE REFERÊNCIA EM CADA
DA BLOCO.

A seguinte ilustração do uso das folhas de resposta é baseada nos dados provenientes da terceira fase de um programa OPEV aplicado à manufatura de um antibiótico. Nessa fase, estão sendo investigados tempo de residência, temperatura da reação e ph. A resposta seguida era rendimento, que era extremamente variável. A fig. 4.4 mostra a disposição do delineamento. Os números se referem à sequência em que as corridas eram efetuadas. As corridas 0, 1, 2, 3 e 4 mostradas por pontos cheios compreendem o bloco I e as corridas 5, 6, 7 e 8 mostradas por pontos abertos, compreendem o bloco II. Os dados na tabela 4.8 são os rendimentos resultados após os primeiros três ciclos.

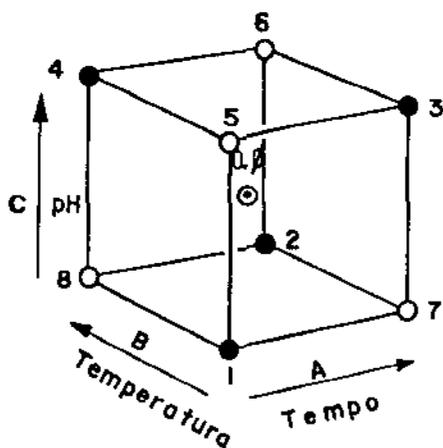
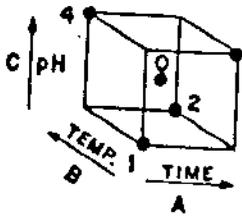


Fig. 4.4 Um delineamento 2^3 com pontos centrais, usados em um esquema OPEV com 3 variáveis.

Os cálculos incluídos nas folhas de resposta se assemelham bastante àqueles usados em um esquema com duas variáveis, e os registros nas folhas (fig. 4.5 a 4.7) são auto-explicativas. Deverá, entretanto, ser notado que desde que há 10 conjuntos de condições de operação (as condições de operação 0 e \emptyset ao centro do cubo são, de fato, idênticas, mas são tratadas em separado já que são efetuadas em blocos diferentes), devemos manter sob controle as 10 médias acumuladas tal que, por exemplo, a soma precedente para a condição 0 será encontrada na folha de trabalho precedente para o bloco I e não na folha de trabalho para o bloco II.

Do ciclo $n = 2$ em diante, é calculada uma estimativa do desvio padrão dos dados de cada bloco e é combinada com as estimativas dos blocos anteriores. Como antes, uma estimativa a priori de σ é usada até (ao fim do bloco II) duas estimativas atuais estarem disponíveis para se tirar a média. (Como antes, o asterisco indica uso da estimativa a priori no primeiro ciclo).



FATORIAL 2³ EM 2 BLOCOS COM OPERAÇÃO DE REFERÊNCIA EM CADA BLOCO

Bloco I

Ciclo n = 1

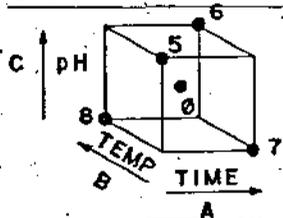
Resposta Rendimento

Projeto Antibiótico
 Fase 3
 Data 29/01

Cálculo das Médias						Cálculo do Desvio Padrão	
Condições de Operação	(0)	(1)	(2)	(3)	(4)	Estimativa a priori de σ	= 8*
(i) Soma anterior para o Bloco I						Soma anterior s (todos os Blocos)	=
(ii) Média anterior para o Bloco I							
(iii) Novas observações para o Bloco I	78	82	63	81	88	Novo s = amplitude x f5,n	=
(iv) Diferença (ii) menos (iii)						Amplitude	=
(v) Novas somas para o Bloco I	78	82	63	81	88	Novo soma s (todos os Blocos)	=
(vi) Novas médias para o Bloco I	78	82	63	81	88		

TABELA DE FATORES DE MULTIPLICAÇÃO

n	1	2	3	4	5	6
f5,n		0,30	0,35	0,37	0,38	0,39
1/n	1	0,50	0,33	0,25	0,20	0,17
1/(2(n-1))		0,50	0,25	0,17	0,12	0,10
2/√n	2	1,41	1,15	1,00	0,89	0,82
1,41/√n	1,41	1,00	0,82	0,71	0,63	0,58
1,26/√n	1,26	0,89	0,73	0,63	0,57	0,52



FATORIAL 2³ EM 2 BLOCOS COM OPERAÇÃO DE REFERÊNCIA EM CADA BLOCO

Bloco II

Ciclo n = 1

Resposta Rendimento

Projeto Antibiótico
 Fase 3
 Data 31/01

Cálculo das Médias						Cálculo do Desvio Padrão	
Condições de Operação	(0)	(5)	(6)	(7)	(8)	Estimativa a priori de σ	= 8*
(i) Soma anterior para o Bloco II						Soma anterior s (todos os Blocos)	=
(ii) Média anterior para o Bloco II							
(iii) Novas observações para o Bloco II	85	79	75	78	67	Novo s = amplitude x f5,n	=
(iv) Diferença (ii) menos (iii)						Amplitude	=
(v) Nova soma para o Bloco II	85	79	75	78	67	Novo soma s (todos os Blocos)	=
(vi) Novas médias para o Bloco II	85	79	75	78	67	Média s = 1/(2(n-1)) x Nova soma s =	

CÁLCULO DOS EFEITOS USANDO O ALGORITMO DE YATES

$\bar{y} - 60$	(i)	(ii)	(iii)	Multiplicador	Efeito	
(1) 22	40	50	133	0,125	36,6	Média fatorial - 60
(7) 18	10	23	-19	0,25	-4,75	A
(8) 7	40	-8	-27	0,25	-6,75	B
(2) 3	43	-11	-15	0,25	-3,75	AB
(5) 19	-4	-30	33	0,25	8,25	C
(3) 21	-4	3	-3	0,25	-0,75	AC
(4) 18	3	0	33	0,25	8,25	BC
(6) 15	-13	15	-15	0,25	-3,75	ABC + Blocos

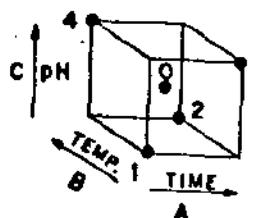
Verificação S de Q 2,677 22,416

CÁLCULO DOS LIMITES DE 2 ERROS PADRÃO

Para novas médias: $\pm \frac{2}{\sqrt{n}} s = \pm 16^*$
 Para novos efeitos: $\pm \frac{1,41}{\sqrt{n}} s = \pm 11,3^*$
 Para mudança na média: $\pm \frac{1,26}{\sqrt{n}} s = \pm 10,1^*$

	$\bar{y} - 60$	soma	multiplicador	média	
Condições de referência (0) (10)	38	43	0,5	21,5	Média das referências 81,5
Condições fatoriais		133	0,125	16,6	Média fatorial 76,6
Todas as condições		176	0,1	17,6	Média na fase 77,6
Mudança na média =		77,6	-81,5	-3,9	

Fig. 4.5. Folhas de trabalho para um programa OPEV com 3 variáveis, cálculos para o ciclo 1.



FATORIAL 2³ EM 2 BLOCOS COM OPERAÇÃO DE REFERÊNCIA EM CADA BLOCO

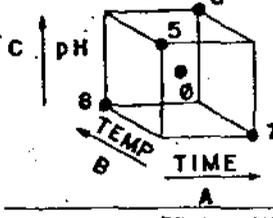
Bloco I
Ciclo n = 2
Resposta Rendimento

Projeto Antibiótico
Fase 3
Data 2/02

Cálculo das Médias					Cálculo do Desvio Padrão	
Condições de Operação	(0)	(1)	(2)	(3)	(4)	Estimativa a priori de σ
(i) Soma anterior para o Bloco I	78	82	63	81	88	= 8*
(ii) Média anterior para o Bloco I	78	82	63	81	88	
(iii) Novas observações para o Bloco I	82	95	79	96	77	Novo σ = amplitude x 15/n = 8.1
(iv) Diferença (ii) menos (iii)	-4	7	-16	-15	11	Amplitude = 2.7
(v) Nova soma para o Bloco I	160	157	142	177	165	Nova soma σ (todos os Blocos) = 8.1
(vi) Novas médias para o Bloco I	80.0	78.5	71.0	88.5	82.5	

TABELA DE FATORES DE MULTIPLICAÇÃO

n	1	2	3	4	5	6
15/n		0,30	0,35	0,37	0,38	0,39
1/n	1	0,50	0,33	0,25	0,20	0,17
1/2(n-1)		0,50	0,25	0,17	0,12	0,10
2/√n	2	1,41	1,15	1,00	0,89	0,82
1,41/√n	1,41	1,00	0,82	0,71	0,63	0,58
1,26/√n	1,26	0,89	0,73	0,63	0,57	0,52



FATORIAL 2³ EM 2 BLOCOS COM OPERAÇÃO DE REFERÊNCIA EM CADA BLOCO

Bloco II
Ciclo n = 2
Resposta Rendimento

Projeto Antibiótico
Fase 3
Data 4/02

Cálculo das Médias					Cálculo do Desvio Padrão	
Condições de Operação	(0)	(5)	(6)	(7)	(8)	Estimativa a priori de σ
(i) Soma anterior para o Bloco II	85	79	75	78	67	= 8*
(ii) Média anterior para o Bloco II	85	79	75	78	67	Soma anterior σ (todos os Blocos) = 8.1
(iii) Novas observações para o Bloco II	69	77	80	70	84	Novo σ = amplitude x 15/n = 9.9
(iv) Diferença (ii) menos (iii)	16	2	-5	8	-17	Amplitude = 3.3
(v) Nova soma para o Bloco II	154	156	155	148	151	Nova soma σ (todos os Blocos) = 18.0
(vi) Novas médias para o Bloco II	77.0	78.0	77.5	74.0	75.5	Média σ = 1/2(n-1) x Nova soma σ = 9.0

CÁLCULO DOS EFEITOS USANDO O ALGORITMO DE YATES

	∇ -70	(i)	(ii)	(iii)	Multiplificador	Efeito	
(1)	8.5	12.5	19.0	65.5	0.125	8.19	Média fatorial -70
(7)	4.0	6.5	46.5	-3.5	0.25	-0.88	A
(8)	5.5	26.5	-9.0	-12.5	0.25	-3.2	B
(2)	1.0	20.0	5.5	-15.5	0.25	-3.88	AB
(5)	8.0	-4.5	-6.0	27.5	0.25	6.88	C
(3)	18.5	-4.5	-6.5	14.5	0.25	3.62	AC
(4)	12.5	10.5	0.0	-0.5	0.25	-0.12	BC
(6)	7.5	-5.0	-15.5	-15.5	0.25	-3.88	ABC + Blocos

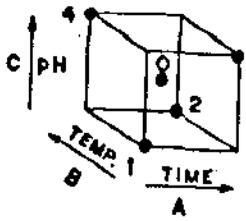
Verificação S de Q 738.25 5906.00

CÁLCULO DOS LIMITES DE 2 ERROS PADRÃO

Para novas médias: $\pm \frac{2}{\sqrt{n}} \sigma = \pm 12.7$
 Para novos efeitos: $\pm \frac{1.41}{\sqrt{n}} \sigma = \pm 9.0$
 Para mudança na média: $\pm \frac{1.26}{\sqrt{n}} \sigma = \pm 8.0$

	∇ -70	soma	multiplificador	média -70	
Condições de referência (0)	10.0 7.0	17.0	0.5	8.5	Média das referências 78.5
Condições fatoriais		65.5	0.125	8.19	Média fatorial 78.19
Todas as condições		82.5	0.1	8.25	Média na fase 78.25
Mudança na média =		78.25 - 78.5 = -0.25			

Fig.4.6. Folhas de trabalho para um programa OPEV com 3 variáveis, cálculos para o ciclo 2.



FATORIAL 2³ EM 2 BLOCOS COM OPERAÇÃO DE REFERÊNCIA EM CADA BLOCO

Bloco I

Ciclo n = 3

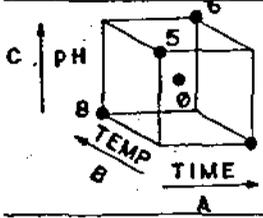
Resposta Rendimento

Projeto Antibiótico
 Fase 3
 Data 7/02

Cálculo das Médias						Cálculo do Desvio Padrão	
Condições de Operação	(0)	(1)	(2)	(3)	(4)	Estimativa a priori de σ	= 8*
(i) Soma anterior para o Bloco I	160	157	142	177	165	Soma anterior s (todos os Blocos)	= 18
(ii) Média anterior para o Bloco I	80	78.5	71	88.5	82.5	Novo σ = amplitude x 1/5,n	= 6.5
(iii) Novas observações para o Bloco I	65	82	68	85	79	Amplitude	= 18.5
(iv) Diferença (ii) menos (iii)	15	-3.5	3	3.5	3.5	Novo soma s (todos os Blocos)	= 24.5
(v) Novas somas para o Bloco I	225	239	210	262	244		
(vi) Novas médias para o Bloco I	75	79.3	70	87.3	81.3		

TABELA DE FATORES DE MULTIPLICAÇÃO

n	1	2	3	4	5	6
1/5,n		0,30	0,35	0,37	0,38	0,39
1/n	1	0,50	0,33	0,25	0,20	0,17
1/12(n-1)		0,50	0,25	0,17	0,12	0,10
2/n	2	1,41	1,15	1,00	0,89	0,82
1,41/n	1,41	1,00	0,82	0,71	0,63	0,58
1,26/n	1,26	0,89	0,73	0,63	0,57	0,52



FATORIAL 2³ EM 2 BLOCOS COM OPERAÇÃO DE REFERÊNCIA EM CADA BLOCO

Bloco II

Ciclo n = 3

Resposta Rendimento

Projeto Antibiótico
 Fase 3
 Data 9/02

Cálculo das Médias						Cálculo do Desvio Padrão	
Condições de Operação	(5)	(6)	(7)	(8)	Estimativa a priori de σ	= 8*	
(i) Soma anterior para o Bloco II	154	156	155	148	151	Soma anterior s (todos os Blocos)	= 24.5
(ii) Média anterior para o Bloco II	77	78	77.5	74	75.5	Novo σ = amplitude x 1/5,n	= 10.3
(iii) Novas observações para o Bloco II	87	96	66	82	72	Amplitude	= 29.5
(iv) Diferença (ii) menos (iii)	-10	-18	11.5	-8	3.5	Novo soma s (todos os Blocos)	= 34.8
(v) Nova soma para o Bloco II	241	252	221	230	223	Média σ = 1/12(n-1) x Nova soma s =	8.7
(vi) Novas médias para o Bloco II	80.3	84.0	73.7	76.7	74.3		

CÁLCULO DOS EFEITOS USANDO O ALGORITMO DE YATES

$y - \bar{y}_0$	(i)	(ii)	(iii)	Multiplicador	Efeito
(1) 7.7	16.4	20.7	67.0	0.125	8.38 Média fatorial ~ 70
(7) 6.7	4.3	46.3	-11.6	0.25	-2.9 A
(8) 4.3	31.3	-7.3	-38.4	0.25	-7.10 B
(2) 0.0	15.0	-4.3	-12.2	0.25	-3.05 AB
(5) 14.0	-3.0	-12.2	25.2	0.25	6.40 C
(3) 17.3	-4.3	-46.3	3.0	0.25	0.75 AC
(4) 11.3	3.3	-3.3	-4.2	0.25	-1.05 BC
(6) 2.2	-7.6	-10.9	-9.6	0.25	-2.40 ABC + Blocos

Verificação: 794.14 6353.12

CÁLCULO DOS LIMITES DE 2 ERROS PADRÃO

Para novas médias: $\pm \frac{2}{\sqrt{n}} s = \pm 10.0$
 Para novos efeitos: $\pm \frac{1.41}{\sqrt{n}} s = \pm 7.1$
 Para mudança na média: $\pm \frac{1.35}{\sqrt{n}} s = \pm 6.3$

	$\bar{y} - \bar{y}_0$	soma	multiplicador	média - 70	
Condições de referência (0)	5	45.3	0.5	7.65	Média das referências 77.65
Condições fatoriais	67	0.125	8.38		Média fatorial 78.38
Todas as condições	82.3	0.1	8.23		Média na fase 78.23
Mudança na média =	78.23	-77.65	-0.58		

Fig. 4.7. Folhas de trabalho para um programa OPEV com 3 variáveis, cálculos para o ciclo 3.

Condições	BLOCO I					BLOCO II				
	0	1	2	3	4	Ø	5	6	7	8
Ciclo 1	78	82	63	81	88	85	79	75	78	67
Ciclo 2	82	75	79	96	77	69	77	80	70	84
Ciclo 3	65	82	68	85	79	87	96	66	82	72

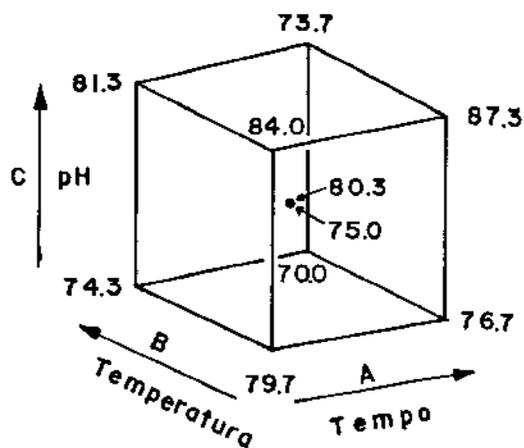
TABELA 4.8 - Rendimentos de corridas individuais para três ciclos OPEV.

Completando cada ciclo (cada par de blocos), os efeitos podem ser recalculados usando o algoritmo de Yates. A folha de trabalho para o bloco II fornece um formato adequado para esse cálculo. Os vários multiplicadores $f_{5,n}$ $1/n$, $1/[2(n-1)]$, $2/\sqrt{n}$, $1.41/\sqrt{n}$ e $1.26/\sqrt{n}$, que são necessários em vários estágios, são listados para $n = 1, 2, \dots, 6$, nas folhas para o bloco I. Então, por exemplo, a quantidade "Média S" é obtida multiplicando "Nova Soma S" pelo multiplicador $1/[2(n-1)]$ que está listado na terceira linha da tabela de fatores multiplicadores.

No extremo inferior direito do bloco III estão expostos os cálculos para a média na fase, média de referência e a mudança na média. Novamente a apresentação é auto-explanatória.

O quadro de informações, como deveria estar no fim do ciclo 3, depois de transferir as várias quantidades calculadas através da folha de resposta apropriada, é mostrado na fig. 4.8 .

ULTIMO CICLO COMPLETADO : 3

Limites de 2 E.P. para as médias individuais: ± 10.0

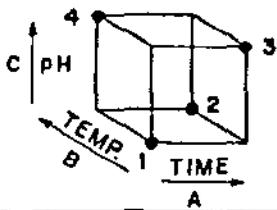
Objetivos	<u>Rendimento</u> Maximizar	<u>Outras Respostas</u> ...
Média na fase	78.2	
Efeitos com limites de 2 E.P.		
Tempo	-2.90 ± 7.1	
Temperatura	-7.10 ± 7.1	
ph	6.40 ± 7.1	
tempo \times temperatura	-3.05 ± 7.1	
tempo \times ph	0.75 ± 7.1	
temperatura \times ph	-1.05 ± 7.1	
mudança na média	0.58 ± 6.3	

Fig. 4.8 - Parte de um quadro de informações para um esquema OPEV, após 3 ciclos.

Verificações - Uma filosofia válida é tratar todos os cálculos como provavelmente errados, até que se prove o contrário. Essas verificações feitas podem ser formais como aquelas da soma dos quadrados no algoritmo de Yates ou apenas informais. Como um exemplo de uma verificação informal, consideremos o quadro de informações após os três ciclos mostrados na fig. 4.8. De acordo com os registros mostrados aqui, um efeito negativo de temperatura e um efeito positivo de ph estão começando a se distinguir do erro. Eles deveriam, entretanto, ser evidentes nas médias dispostas nos vértices do cubo imediatamente acima para análise. Uma inspeção dos valores apropriados no cubo confirma isso. Por outro lado, se algum grande efeito tivesse sido encontrado, que não fosse evidente na inspeção dos dados dispostos no cubo, deveríamos suspeitar de um erro nos cálculos.

FOLHAS DE TRABALHO PARA UM MODELO FATORIAL 2^3 COM DISPOSIÇÃO EM DOIS BLOCOS, SEM CONDIÇÃO ADICIONAL DE REFERÊNCIA.

Folhas de trabalho para um delineamento 2^3 sem corridas adicionais de referência são um pouco diferentes daquelas mostradas anteriormente. Um formato conveniente é mostrado na fig.4.9 onde estão mostrados os cálculos para o primeiro ciclo, usando dados anteriores. A diferença principal é o uso do fator $f_{4,n}$ ao invés de $f_{5,n}$ na estimação do desvio padrão pela amplitude. Também não foi tomada nenhuma iniciativa no sentido de calcular a média na fase. Isso pode ser adaptado, mas devemos notar o seguinte:



FATORIAL 2³ SEM CONDIÇÃO ADICIONAL DE REFERÊNCIA

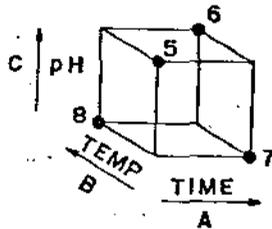
Bloco I
Ciclo n = 1
Resposta: *Rendimento*

Projeto: *Antibiótico*
Fase: 3
Data: 29/01

CÁLCULO DAS MÉDIAS					CÁLCULO DO DESVIO PADRÃO	
Condições de Operação	(1)	(2)	(3)	(4)	Estimativa a priori de σ	= 8*
(i) Soma anterior para o Bloco I					Soma anterior S (todos os Blocos)	=
(ii) Média anterior para o Bloco I						
(iii) Novas observações para o Bloco I	82	63	81	88	Novo S = amplitude x f4,n	=
(iv) Diferença (ii) menos (iii)					Amplitude	=
(v) Novas somas para o Bloco I	82	63	81	88	Novo soma S (todos os Blocos)	=
(vi) Novas médias para o Bloco I	82	63	81	88		

TABELA DE FATORES DE MULTIPLICAÇÃO

n	1	2	3	4	5	6
f4,n		0.34	0.40	0.42	0.43	0.44
1/n	1	0.50	0.33	0.25	0.20	0.17
1/[2(n-1)]		0.50	0.25	0.17	0.12	0.10
2/√n	2	1.41	1.15	1.00	0.89	0.82
1.41/√n	1.41	1.90	0.82	0.71	0.63	0.58
1.26/√n	1.26	0.89	0.73	0.63	0.57	0.52



FATORIAL 2³ SEM CONDIÇÃO ADICIONAL DE REFERÊNCIA

Bloco II
Ciclo n = 1
Resposta: *Rendimento*

Projeto: *Antibiótico*
Fase: 3
Data: 31/01

CÁLCULO DAS MÉDIAS					CÁLCULO DO DESVIO PADRÃO	
Condições de Operação	(5)	(6)	(7)	(8)	Estimativa a priori de σ	= 8*
(i) Soma anterior para o Bloco II					Soma anterior S (todos os Blocos)	=
(ii) Média anterior para o Bloco II						
(iii) Novas observações para o Bloco II	79	75	78	67	Novo S = amplitude x f4,n	=
(iv) Diferença (ii) menos (iii)					Amplitude	=
(v) Nova soma para o Bloco II	79	75	78	67	Novo soma S (todos os Blocos)	=
(vi) Novas médias para o Bloco II	79	75	78	67	Média S = 1/[2(n-1)] x Nova soma S	=

CÁLCULO DOS EFEITOS USANDO O ALGORITMO DE YATES						CÁLCULO DOS LIMITES DE 2 ERROS PADRÃO	
	$\bar{y} - 60$ (i)	(ii)	(iii)	Multiplicador Efeitos			
(1)	22	40	50	0.125	366	Média Fatorial - 60	
(7)	18	10	83	0.25	-478	A	Para novas médias: $\pm \frac{3}{\sqrt{n}} S$ = ± 16 *
(8)	7	40	-8	0.25	-678	B	
(2)	3	43	-11	0.25	-378	AB	Para novos efeitos: $\pm \frac{1.41}{\sqrt{n}} S$ = ± 11.3 *
(5)	19	-4	-30	0.25	825	C	
(3)	21	-4	3	0.25	-975	AC	
(4)	28	2	0	0.125	825	BC	
(6)	15	-13	15	0.25	-375	ABC + Blocos	Média na fase = Média fatorial = 77.6
Verificação S de Q	2.167		21.456				

Fig. 4.9 Folhas de trabalho para um esquema OPEV com três variáveis, sem ponto adicional de referência. Cálculos para o ciclo 1.

se uma das corridas fatoriais é uma condição de referência e é operada em somente um bloco, a mudança global na média pode ser calculada como de costume, mas é influenciada pela variação de bloco para bloco. Se uma das corridas fatoriais é uma condição de referência e é operada em todos os blocos, a mudança global na média é livre da variação de bloco para bloco, mas a corrida "extra" de referência, isto é, aquela que está em um bloco a que ela não pertence, seria ignorada quando os efeitos fatoriais fossem calculados. Embora essas variações OPEV sejam um pouco mais complexas do que as anteriores, elas ainda podem ser manipuladas usando os princípios descritos até aqui.

5 - OTIMIZAÇÃO DE UMA CENTRÍFUGA

MOTIVAÇÃO - Separador Centrífugo é bastante empregado nas indústrias químicas (Usina de açúcar e de álcool, fábrica de cerveja, fábrica de fermento, etc), e de alimentos (laticínios [13], etc). Com o semeio de levedura em uma cultura apropriada, a fermentação se propaga formando um "mosto". A separação das células desse "mosto" pode ser feita por decantação, o que na maioria das vezes é desaconselhado, pela lentidão do processo.

Na utilização do separador centrífugo, vários fatores afetam as características do(s) produto(s) de interesse.

Com o propósito de observar as influências de tais fatores, realizamos experiências em nível de laboratório, utilizando um separador de laboratório Alfa Laval LAPX 202.

MATÉRIAS E MÉTODOS - Realizamos o experimento com os equipamentos abaixo relacionados, segundo o diagrama da fig. 5.1.

A → Tanque de Aço Inox, capacidade 100 l + agitador Vibratório CHEMAPEC.

B → Separador Centrífugo com ejeção de sólidos, α -laval LAPX 202.

C → Bomba Mono SB-15 e painel → bomba de deslocamento positivo.

D e E → Frascos para o recolhimento das amostras.

1 → Concentrado Original

4 → Corrente de Clarificado

6 → Corrente de Concentrado

8 → Dreno

15 e 16 → Líquido de Operação

- 22 → Conexão Elétrica
 50 → Válvula de Controle
 53 → Soquete
 60 → Válvula de Fechamento
 74 → Coador
 79 → Filtro
 91 → Válvula de Controle de Fluxo
 T → Transformador
 U → Relé de Tempo
 V → Válvula Solenóide

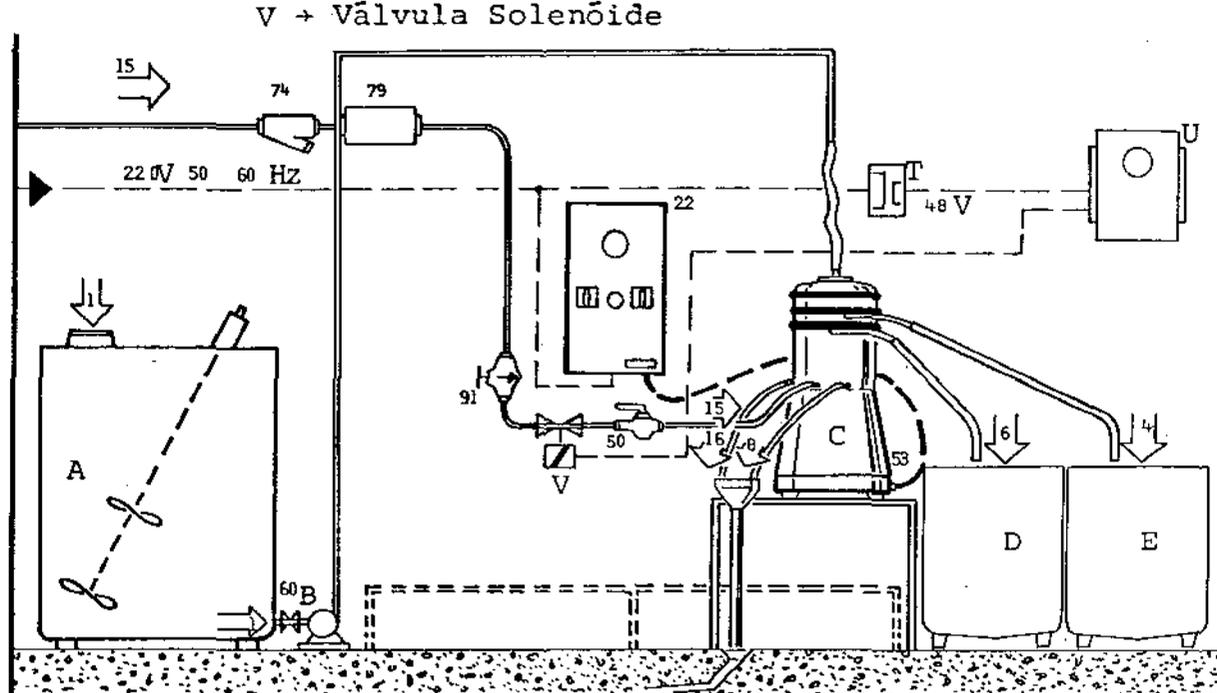


Fig. 5.1 DIAGRAMA DO PROCESSO

Diluímos um pacote de levedura (200 g base úmida) no tanque (A) contendo 70 litros de água. Em constante homogeneização pelo agitador, a suspensão de células de levedura é bombeada (B) até a centrífuga (C). Tal suspensão chega à centrífuga por cima e flue, através de um distribuidor, para os intervalos entre os discos do tambor da centrífuga. Pela ação da "força centrífuga", as partículas sólidas se movem ao longo e para fora dos discos, em direção à periferia do tambor, ali sedimentando. A "corrente de clarificado" prossegue em direção ao centro do tambor e é descarrega

da através de orifícios na zona próxima ao eixo da centrífuga, para o frasco E. Os sólidos (a corrente de concentrado de células) são expelidos por uma abertura na parede do tambor, de maneira intermitente para o frasco D, pela operação do sistema de ejeção de sólidos. Entre as ejeções, o tambor é mantido fechado. O anel de vedação cobre uma pequena fenda entre as bordas do tambor e da tampa. Tal anel é expandido contra a tampa, pela pressão do líquido de operação que atua em seu lado inferior, sendo erguido durante a rotação por causa da "força centrífuga". O líquido de operação atua segundo uma "válvula operada por piloto interno" (fig. 5-3 a e 5-3 b). Essa válvula é equipada com um orifício piloto e um de

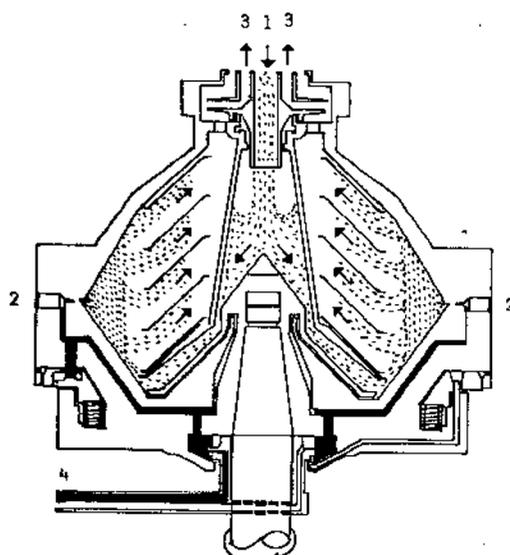


Fig. 5.2 Corte longitudinal do rotor (modelo básico) em posição fechada, mostrando, em preto, o líquido de manobra que mantém o fundo móvel premido contra a capa. 4 indica a entrada do líquido de manobra para abertura e fechamento. No modelo estanque ao gás, o líquido de manobra forma um selo hidráulico rotativo, o qual impede a penetração de gases na caixa de engrenagens. A capa do rotor é construída para suportar uma pressão de gás inerte de 0.7 kg/cm^2 (10 psig).

1. Alimentação
2. Saída para o líquido concentrado
3. Saída para o líquido clarificado

sangramento, e utiliza a própria linha de pressão para operação. Quando o solenóide está ativado, abre o orifício piloto e libera pressão do topo do pistão da válvula ou diafragma para o lado da saída da válvula. Isso resulta em uma pressão desbalanceada que faz a pressão da linha erguer o pistão ou diafragma do orifício principal, abrindo portanto a válvula.

Quando o solenóide está desativado, o orifício piloto é fechado e a pressão da linha é aplicada no topo do pistão ou dia

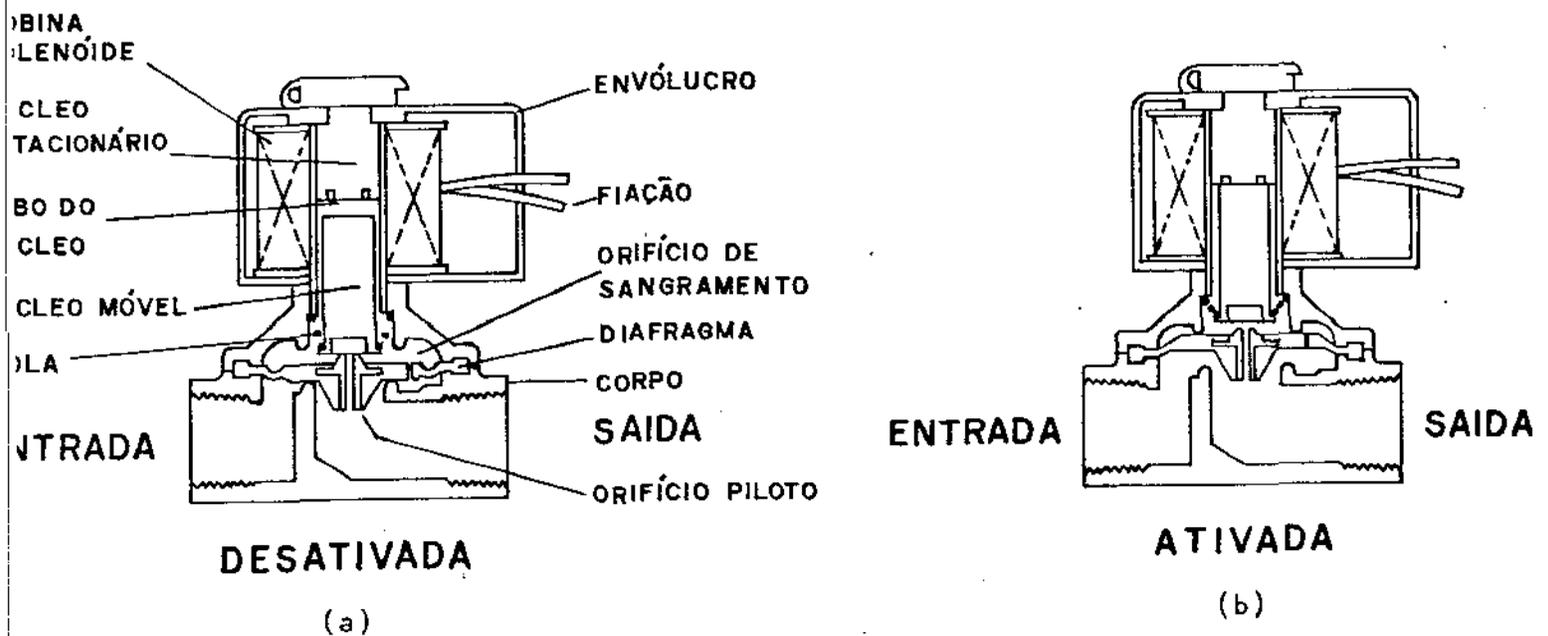


Fig. 5.3 Válvula solenóide: a) desativada
b) ativada

fragma através do orifício de sangramento, portanto fechando completamente a válvula. A ativação do solenóide é comandada por um relé de tempo, eletrônico, cíclico, com dois tempos ajustáveis, independentes e repetitivos. Após receber a alimentação inicia-se o primeiro ciclo; quando se esgota o tempo pré-ajustado, inicia-se novamente o primeiro ciclo, e assim sucessivamente.

MEDIDAS FÍSICAS:

RPM da Centrífuga: Cravados o número de RPM da centrífuga vazia. Estimamos esse número em 5000 através de um Estrolocópio Strobotac. Este ajuste permaneceu por toda a experiência.

Concentração das amostras: Utilizamos um Espectrofotômetro Spectronic 20 Baush Lamb. Com luz incidindo sobre água, ele é calibrado para zero, assumindo que nenhuma luz incidente é transmitida. Fizemos uma calibração da porcentagem de célula desidratada para se obter concentrado em mg de célula seca /100 ml. As células foram desidratadas a vácuo (28" de Hg), a 60°C. Obtivemos então, pelo método dos Mínimos Quadrados, a reta

$$y = -1.54214 + 85.32091 x,$$

onde

$$x = \log (\text{absorvância})$$

$$y = \text{concentração (mg célula seca (100 ml))}$$

"e o ajuste foi de boa qualidade".

Medida de Vazão: Com a bomba mono SB15 entre 400 e 1200 RPM, estimamos a vazão da suspensão de células para a centrífuga em intervalos de 100 RPM. Ajustamos os dados pelo Método dos Mínimos Quadrados, a

$$y = -17.29563 + 0.0599X,$$

X = RPM da bomba

Y = vazão (ml/S) ,

"também com bom ajuste".

Delineamento do Experimento: Utilizamos um delineamento fatorial 2x2 com condições centrais de referência. Nas três primeiras fases, os fatores foram tempo de ejeções e tempo entre ejeções (S). A vazão foi calibrada para 36.7 ml/S. Os fatores estavam dispostos conforme fig. 5.4

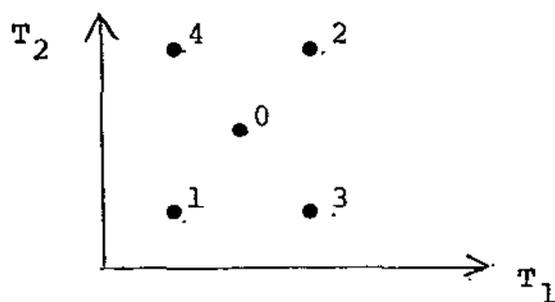


Fig. 5.4 - Delineamento do experimento de otimização de centrífuga que foi utilizado para as 3 primeiras fases.

onde T_1 é o tempo de ejeção e T_2 o tempo entre ejeções.

Como a vazão de concentrado de células para a centrífuga estava fixado, certamente seria obtida uma combinação entre os tempos que maximizasse a concentração da corrente de concentrado colhida no frasco 6. Para as fases sucessivas, fixamos o tempo entre ejeções e o fator introduzido foi vazão da suspensão de células de levedura para a centrífuga (ml/S).

A disposição dos fatores era conforme a fig. 5.5

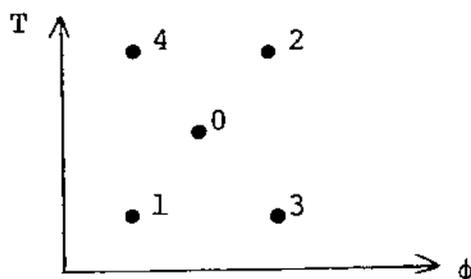


Fig. 5.5 - Delineamento do experimento de Otimização da centrífuga que foi utilizado para as fases 4, 5 e 6 .

onde ϕ é a vazão em ml/S e T o tempo entre ejeções.(S)

PROCEDIMENTOS DA OPERAÇÃO

O método de absorvância para estimar a concentração só é válido para células lavadas. Usamos Fermento Itaiquara, onde as células são lavadas e relativamente uniformes, facilitando a centrifugação. Para cada fase, com o tanque tendo aproximadamente 70 l

de água, utilizamos um pacote de 200 g (base úmida) de levedura ; após a homogeneização, eram retiradas duas amostras do tanque, diluídas a 1/10 e medidas suas absorvâncias.

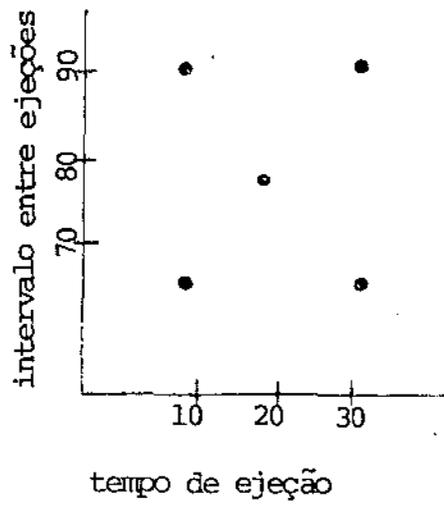
Em cada conjunto de condições de operações eram tiradas amostras de concentrado e clarificado. Este não necessitava diluição para se poder medir sua absorvância, mas o concentrado necessitava diluição 1/100. As medidas de interesse eram a proporção entre concentração das amostras coletadas nos frascos D e E e a concentração da suspensão de células no tanque.

RESULTADOS E DISCUSSÃO

Para a primeira fase investigamos o tempo de ejeção nos níveis (10, 20, 30) e o tempo entre ejeções nos níveis (70, 80, 90). O quadro de Informações para essas condições, após quatro ciclos, é mostrado na Tab. I. Até as casas decimais consideradas os níveis dos tempos utilizados não influíram na concentração do clarificado. A célula de levedura tem, em média, 10 micra (partícula considerada grande). Como é baixa a viscosidade da água, o arrasto não é suficiente para reter as células, havendo elevada velocidade de sedimentação. O campo centrífugo aplicado é ocioso e a separação das células não é afetada pelo tempo de residência no tambor da centrífuga. Para o concentrado, existe um acentuado efeito negativo de T_1 , e um pequeno efeito positivo de T_2 . Verificamos então que, quando T_1 é maior que o necessário para expelir células sedimentadas, o líquido começa a ser expelido com as células dimi

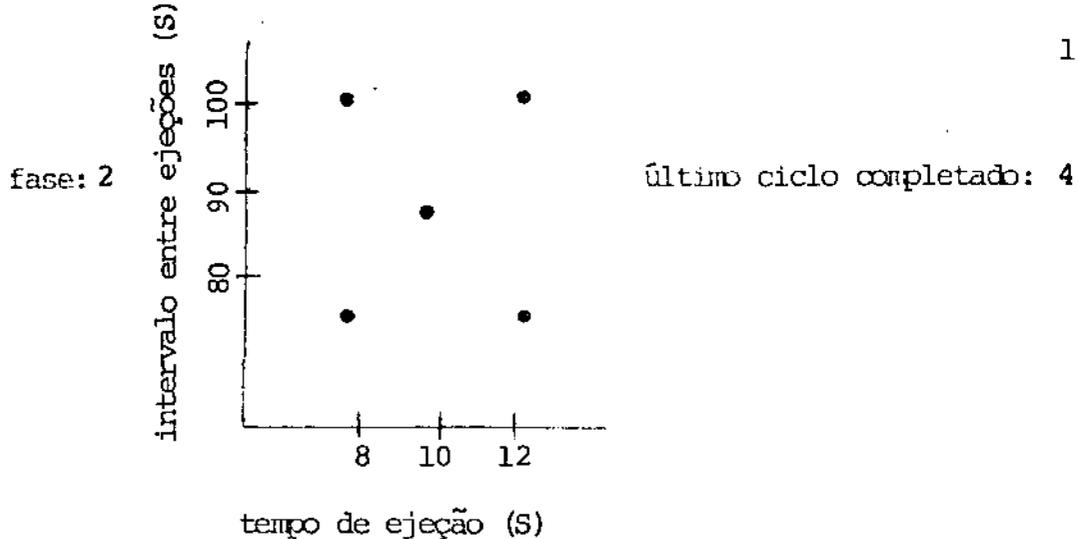
fase: 1

último ciclo completado: 4



		Concentrado	Clarificado
Médias Atuais		5.98 3.45	0.04 0.04
		3.93	0.04
Limites de 2. E.P.		5.03 3.14	0.04 0.04
		± 0.362	± 0.001
Médias na fase		4.305	0.042
Efeitos com limites para 2. E.P.	Tempo de Ejeção	-2.209 ± 0.362	—
	Intervalo entre Ejeções	0.631 ± 0.362	—
	T x I	-0.316 ± 0.362	—
	Mudança na Média	0.379 ± 0.323	0.001 ± 0.001
S		1.131	0.002
Desvio padrão para as observações individuais		0.362	0.001

Tab. I - Aparência do Quadro de Informações no fim de 4 ciclos

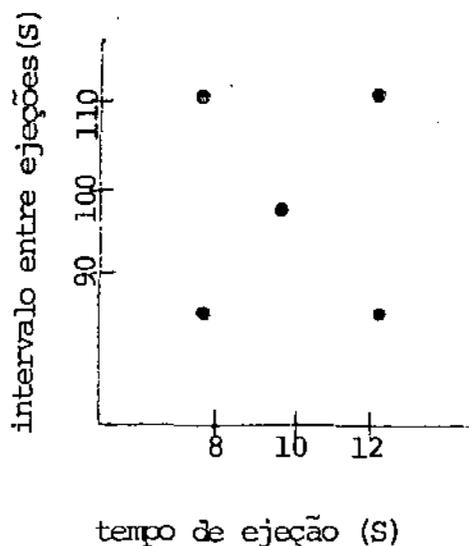


		Concentrado	Clarificado
Médias Atuais		8.60 6.89	0.04 0.03
		6.74	0.05
		7.20 5.73	0.05 0.04
Limites de 2. E.P.		± 0.738	± 0.014
Médias na fase		7.033	0.052
Efeitos com limites para 2. E.P.	Tempo de Ejeção	-1.592 ± 0.738	0.002 ± 0.014
	Intervalo entre Ejeções	1.281 ± 0.738	-0.003 ± 0.014
	T x I	-0.116 ± 0.738	-0.001 ± 0.014
	Mudança na Média	0.295 ± 0.660	0.001 ± 0.012
S		0.362	0.001
Desvio padrão para as observações individuais		0.738	0.001

Tab. II - Aparência do Quadro de Informações no fim de 4 ciclos

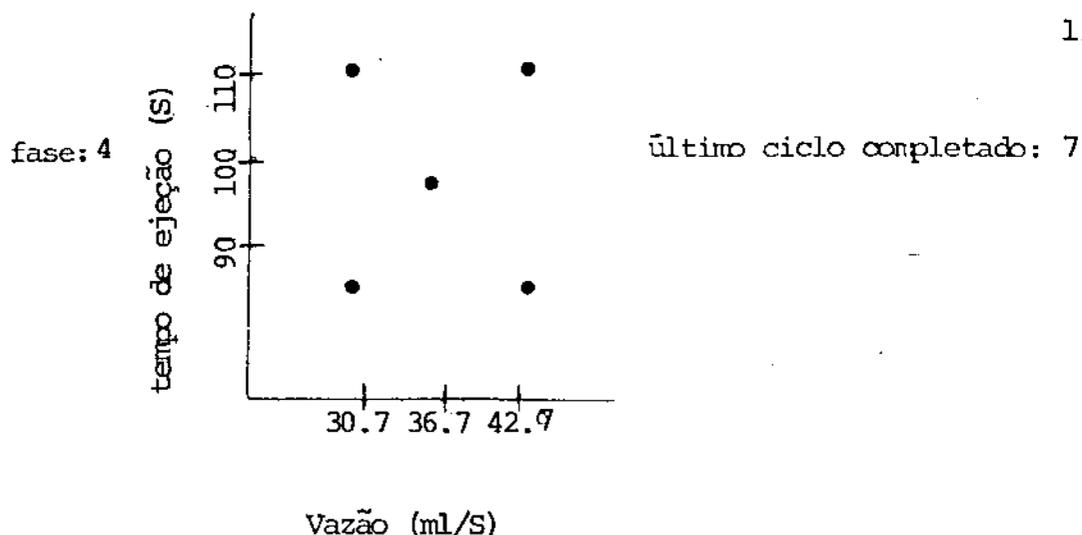
fase: 3

último ciclo completado: 4



		Concentrado	Clarificado
Médias Atuais		7.92 8.00	0.04 0.04
		7.40	0.04
		8.74 6.14	0.03 0.04
Limites de 2. E.P.		± 1.372	± 0.008
Médias na fase		7.640	0.038
Efeitos com limites para 2. E.P.	Tempo de Ejeção	-1.261 ± 1.372	—
	Intervalo entre Ejeções	0.519 ± 1.372	0.004 ± 0.008
	T x I	1.336 ± 1.372	-0.005 ± 0.008
	Mudança na Média	0.237 ± 1.228	-0.002 ± 0.008
S		0.138	0.001
Desvio padrão para as observações individuais		1.372	0.008

Tab. III - Aparência do Quadro de Informações no fim de 4 ciclos.



		Concentrado	Clarificado
Médias Atuais		9.21 12.17 10.09 7.58 9.84	0.04 0.06 0.04 0.04 0.05
Limites de 2. E.P.		± 1.154	± 0.002
Médias na fase		9.78	0.045
Efeitos com limites para 2. E.P.	Vazão	2.613 ± 1.514	0.018 ± 0.002
	Intervalo entre Ejeções	1.983 ± 1.514	0.001 ± 0.002
	T x I	0.346 ± 1.514	0.001 ± 0.002
	Mudança na Média	-0.315 ± 1.355	0.002 ± 0.002
S		1.372	0.008
Desvio padrão para as observações individuais		2.003	0.003

Tab. IV - Aparência do Quadro de Informações no fim de 7 ciclos.

nuindo a sua concentração. Para a vazão utilizada, T_1 deve ser diminuído e T_2 deve ser aumentado.

Para a segunda fase, investigamos T_1 nos níveis (8,10, 12) e T_2 nos níveis (80, 90, 100). O quadro de informações para o ciclo 4 dessa fase é mostrado na tab. II. A concentração do clarificado ainda não depende do nível dos fatores utilizados. T_1 ainda deve ser diminuído, e T_2 aumentado. Como o Relé de tempo utilizado não conseguia comandar tempos abaixo de 8 S, para a fase seguinte: utilizamos os mesmos níveis para T_1 aumentando os níveis de T_2 .

Para a terceira fase investigamos os fatores nos seguintes níveis: T_1 (8, 10, 12) e T_2 (90, 100, 110). O quadro de informações para essa fase é mostrado na tabela III.

A clarificação é boa para todas as combinações de tempo até então utilizados. Para a sedimentação, os intervalos de confiança dos efeitos principais, de interação e de mudança na média já contêm o zero. As médias se confundem pelo erro padrão a elas associados. Para essa vazão, o processo deve ser realizado nas atuais condições de referência.

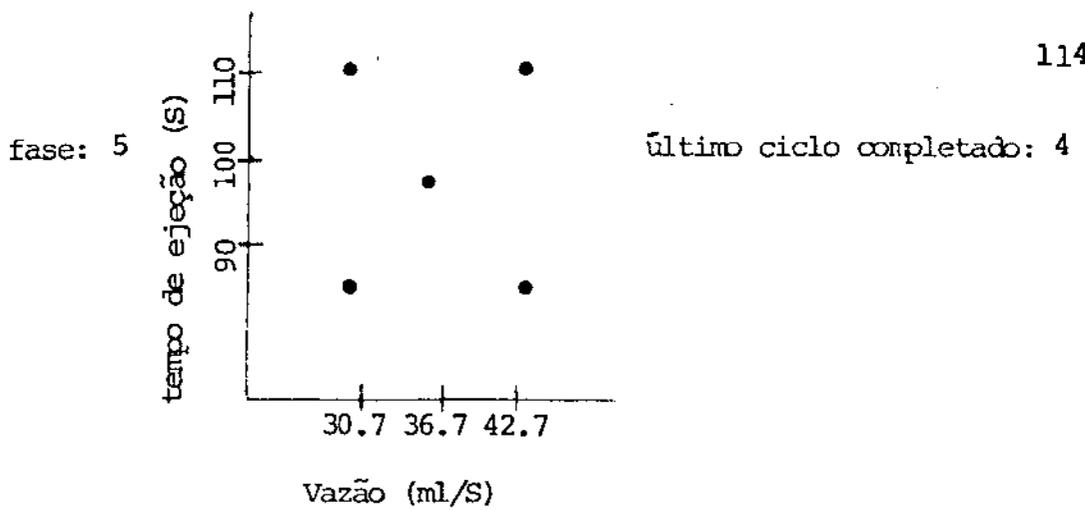
Para a quarta fase, decidimos fixar o tempo de ejeção no mínimo permitido pelo Relé de tempo, e investigar os efeitos de mudança em níveis de: intervalo de tempo entre ejeções (T) e vazão nos níveis (30.7, 36.7 e 42.7) e T nos níveis (90, 100, 110). Foram necessários 7 ciclos para obtermos uma indicação dos efeitos dessas variáveis. A tabela IV mos-

tra o Quadro de Informações para o 7º ciclo dessa fase. Pela grande flutuação dos resultados, observada durante toda a fase, suspeitamos que alguma variável influente no processo estivesse fora de controle. Constatamos a existência de um transitório nas mudanças dos níveis da vazão. Seria necessário um tempo médio de 10 min para haver uma estabilização na vazão. Optamos por repetir o de lineamento, na próxima fase, sem mudança nos níveis. Para tal, um intervalo de 15 minutos seria observado entre as medidas onde houvesse mudança de nível para vazão.

A tabela V mostra o 4º ciclo da 5ª fase.

Observamos que houve um aumento na concentração do clarificado, para as vazões 36.7 e 42.7. O efeito de T é "pouco significativo". Com o aumento da vazão, o tempo médio de residência das células no tambor da centrífuga é diminuído, caindo o rendimento da clarificação. Também, para maior vazão, mais células são sedimentadas nas imediações da parede do tambor e se obtém uma maior densidade de células na "corrente de concentrado". Porém, o espaço entre a borda dos discos e a parede é gradativamente diminuído. Nessas condições, temos mais células que chegam ao tambor na unidade de tempo, e o tempo de estadia no tambor é menor, o que torna a clarificação menos eficiente. Mais células conseguem escapar com a corrente de clarificado.

Para o concentrado existe um pequeno efeito positivo do intervalo de tempo entre ejeções, e um grande efeito decorrente do aumento da vazão. Com o aumento do número de células pelo aumento da vazão, houve grande sedimentação, mesmo com a queda de



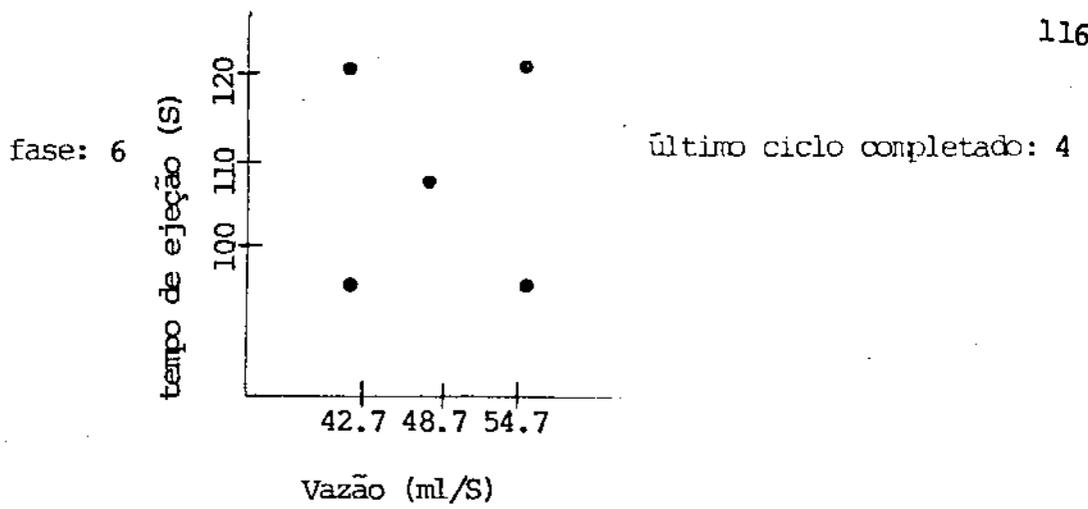
		Concentrado	Clarificado
Médias Atuais		9.09 17.82	0.03 0.06
		12.47	0.06
		8.18 12.87	0.03 0.06
Limites de 2. E.P.		± 2.717	± 0.012
Médias na fase		12.084	0.048
Efeitos com limites para 2. E.P.	Vazão	6.078 ± 2.717	0.026 ± 0.012
	Intervalo entre Ejeções	2.93 ± 2.717	—
	T x I	2.019 ± 2.717	—
	Mudança na Média	-0.395 ± 2.432	-0.007 ± 0.011
S		2.003	0.003
Desvio padrão para as observações individuais		2.717	0.012

Tab. V - Aparência do Quadro de Informações no fim de 4 ciclos

eficiência da clarificação.

Para a sexta fase, decidimos investigar a vazão nos níveis (42.7, 48.7, 54.7) e T nos níveis (100, 110, 12). O quadro de Informações para essa fase está na tabela 6. Com o clarificado verificamos o mesmo que na 5^a fase. Existe evidência de perda de eficiência na clarificação com aumento da vazão. Para o concentrado, um efeito positivo da vazão indicou um aumento da sedimentação. Aumentar a vazão, entretanto, não seria possível devido à impossibilidade de aumentar acima de 1 200 o número de RPM da bomba. Não há "efeito significativo" para T e interação, e o efeito de mudança na média pode ser empregado para se avaliar a forma da superfície de resposta nessas imediações. Para o concentrado, apesar das características de otimalidade da região investigada, a sexta fase foi, em números absolutos, inferior à 5^a fase. Certamente existe mais alguma variável influente no processo e que está fora de controle. Isso vem evidenciar a necessidade de uma pessoa com conhecimento sobre os vários fatores que influem no processo, para ponderar sobre a condução da experiência e opinar sobre as decisões.

Como não houve padronização do número de RPM da centrífuga, sua eficiência pode diminuir. Além disso, o transitório para a centrífuga é bem maior que o considerado para a bomba. Outra variável fora de controle, que influiu nos resultados é a concentração original da suspensão de células. Para a 5^a fase era 1,63 g/l e para a 6^a fase 2.23 g/l . Como a característica medida era



		Concentrado	Clarificado
Médias Atuais		8.47 11.43	0.03 0.06
		11.62	0.05
		7.80 10.48	0.03 0.06
Limites de 2. E.P.		± 1.663	± 0.007
Médias na fase		9.962	0.045
Efeitos com limites para 2. E.P.	Vazão	2.816 ± 1.663	0.028 ± 0.007
	Intervalo entre Ejeções	0.81 ± 1.663	—
	T x I	0.143 ± 1.663	-0.003 ± 0.007
	Mudança na Média	-1.661 ± 1.448	-0.003 ± 0.006
S		2.717	0.012
Desvio padrão para as observações individuais		1.663	0.007

Tab. VI - Aparência do Quadro de Informações no fim de 4 ciclos.

a proporção entre a concentração da "corrente de concentrado" e a concentração da suspensão de células original, a eficiência da sedimantação para a 5^a fase esteve bastante aumentada.

O fato de termos tido melhores resultados na fase 5 com respeito à fase 6 se deve ao seguinte:

O tempo entre as ejeções é suficiente para sedimentar um número muito grande de células. A corrente de concentrado teve para ambas as fases e nos níveis comparáveis, a mesma concentração. Mas, a quantidade medida era a proporção entre a concentração dessa corrente e da suspensão de células original. Como, para a 5^a fase, a concentração da suspensão original foi bem inferior à da 6^a fase, os resultados aparentaram contraditórios.

CONCLUSÕES

Para a clarificação, a variável dentre as investigadas, que influe decisivamente é a vazão. O tempo médio de residência (no tambor da centrífuga) da suspensão original também é afetado principalmente pela vazão.

Para o concentrado, influíram todas as variáveis. O tempo de ejeção deve ser o suficiente para expelir as células sedimentadas com um mínimo de líquido. A vazão deve ser conjugada com o tempo entre as ejeções de forma a evitar a formação de turbulência, que rouba células para o clarificado, reduzindo a densidade de células na corrente de concentrado.

Em cada mudança de vazão, deve-se calibrar o número de RPM da centrífuga, e então esperar tempo suficiente para sua estabilização. A menos que a concentração da suspensão original no tanque possa ser controlada, os valores absolutos das quantidades de interesse não devem ser olhadas como indesejáveis.

Para que a OPEV possa ser aplicada com sucesso em uma indústria, é necessário que as pessoas envolvidas em tal aplicação possam ser devidamente treinadas. Em um curso de OPEV, que poderia ser de curta duração, as pessoas aprenderiam os fundamentos básicos da estatística necessária à OPEV (já que um maior conhecimento poderia ficar a cargo do estatístico membro da comissão). Aquelas pessoas aprenderiam algo acerca de variação, dos 2 limites de E.P., delineamentos fatoriais 2^2 e 2^3 . Após esse curso seria interessante que as pessoas pudessem ser treinadas em um processo simulado. Para esse treinamento é que propomos o jogo da OPEV.

Como proponentes, sabemos a forma analítica de uma função objetivo desconhecido pelos jogadores. Tal função representa um processo que descrevemos detalhadamente antes do jogo. Descrevemos os fatores com os respectivos intervalos de variação e também as características do produto obtido pela operação do processo em determinadas condições. Perturbamos a função objetivo por meio da geração de amostras aleatórias $N(0, \sigma^2)$, e tal função perturbada será uma amostra da característica de interesse em cada condição de operação.

Para a realização desse jogo, elaboramos dois programas (OPEVTY e OPEVWR). Os programas devem ser executados em terminal impressor com 80 caracteres. Tais programas, bem como o manual de uso dos mesmos, estão disponíveis na Biblioteca de programas do

CCUEC (Centro de Computação da Universidade Estadual de Campinas). O manual está contido no Apêndice II. Descrevemos a seguir, dois processos de produção que podem ser utilizados durante o jogo.

PROCESSO DE EXTRUSÃO DE FARINHA

Extrusão de alimentos é um processo de importância crescente. O diagrama da fig. 6.1 ilustra uma planta para produção de

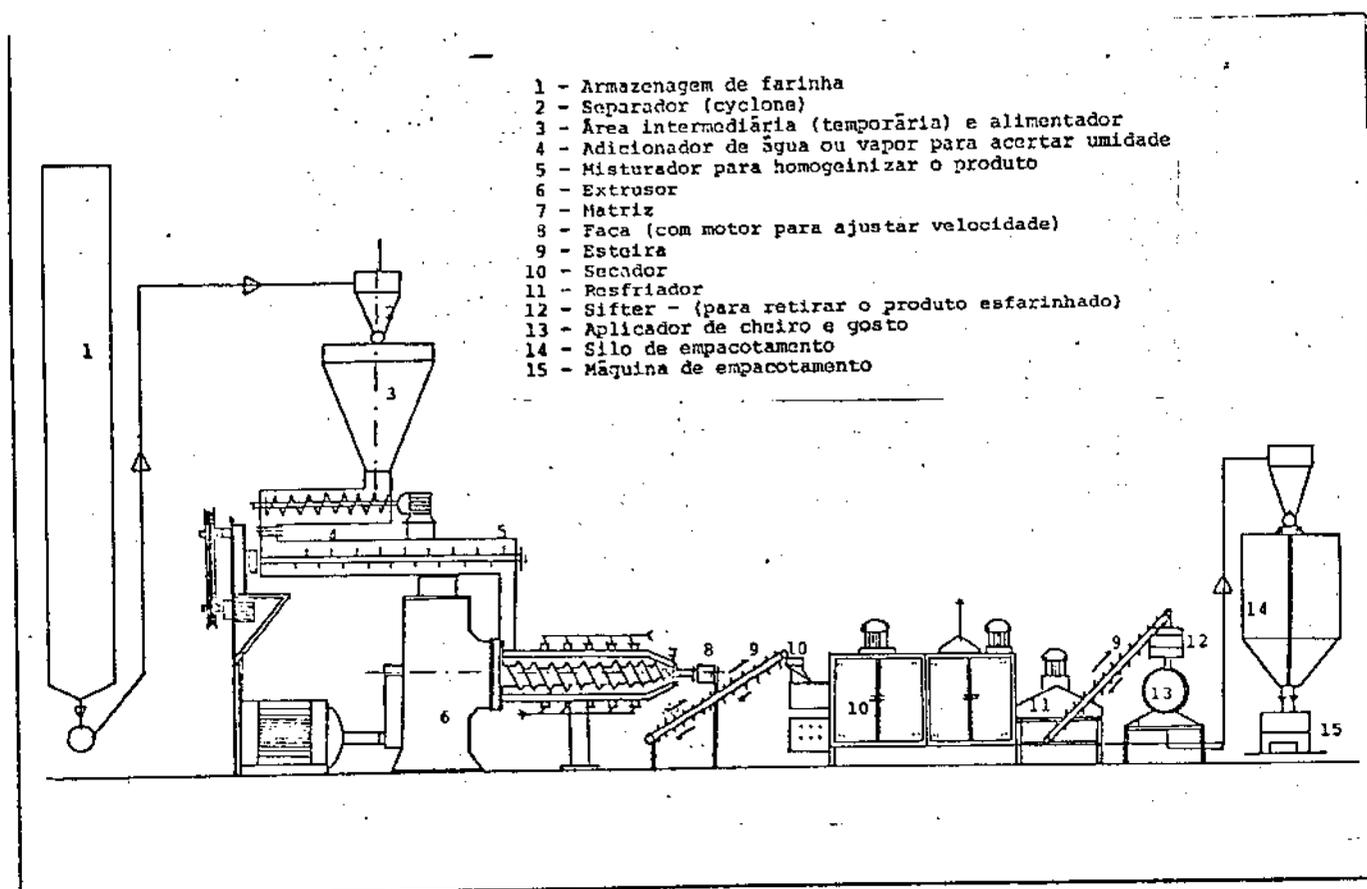


Figura 6.1 - Layout de planta para produção de "puffed snacks".

Prof. Dr. Ahmed A. El-Dash
 FTAA - UNICAMP
 Fevereiro 1978

"puffed snacks". As tres variáveis de controle que afetam as características do produto após a extrusão são:

- a) Umidade da farinha antes da extrusão X_1
- b) Temperatura interna no extrusor X_2
- c) Abertura da matriz X_3

Para o jogo são observadas duas características do produto extrudado:

- 1) *Textura* (valor entre zero e dez).

Os coeficientes da função objetivo são:

β_0	β_1	β_2	β_3	β_{22}	β_{12}	β_{13}	β_{23}	β_{123}	β_{122}	β_{322}	β_{1223}
2.275	0.425	0.375	-0.875	.0825	0.25	-0.425	0.025	-0.25	-0.3	1.375	0.8

e a perturbação é $N(0,0.2)$.

- 2) *Umidade após a extrusão* (%).

Os coeficientes da função objetivo são:

β_0	β_1	β_2	β_3	β_{22}	β_{12}	β_{13}	β_{23}	β_{123}	β_{122}	β_{322}	β_{1223}
7.85	1.05	-2.075	1.35	-0.025	-0.225	0.55	-0.175	-0.175	-0.425	1.175	-0.425

e a perturbação é $N(0,0.4)$.

O modelo empregado é:

$$\begin{aligned} \hat{y} = & \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_{22} (X_2)^2 + \beta_{12} X_1 X_2 + \beta_{13} X_1 X_3 \\ & + \beta_{23} X_2 X_3 + \beta_{123} X_1 X_2 X_3 + \beta_{122} X_1 (X_2)^2 + \beta_{322} X_3 (X_2)^2 \\ & + \beta_{1223} X_1 X_3 (X_2)^2 \end{aligned}$$

onde $X_1 = (\% H_{2O} - 17.5)/2.5$;

$X_2 = (T(^{\circ}F) - 400)/50$;

$X_3 = (\text{Diâmetro (mm)} - 2)$.

Para a textura, os valores que aparecem aos jogadores são transladados por $\hat{\tilde{y}} = 10. - 2x \hat{y}$

Dependendo dos níveis das variáveis, o produto tem outras características: cor, produto quebradiço, produto chamuscado, etc. Por exemplo, se a abertura da matriz (graduada entre .1 mm e 3 mm) for pequena (próxima de 1 mm) e a temperatura interna (variável entre 350^oF e 450^oF) for próxima de 450^oF, deve ocorrer o seguinte: quando o produto está próximo da matriz fica sob alta pressão e alta temperatura. A água do produto fica sob a forma líquida e quando ele passa pela matriz, a pressão cai à ambiente rapidamente. Então possivelmente teremos uma situação de produto quebradi-

ço, leve, dourado.

Na descrição do processo para os jogadores, devemos estabelecer que tipo de produto queremos obter. E as características textura e umidade após a extrusão devem indicar combinações das tres variáveis que possam favorecer o produto a ser obtido.

OBTENÇÃO DE PÃO COM ELEVADO TEOR PROTEICO

Para a fabricação de pão, com elevado teor proteico e boas características (nutricionais, visuais, gustativas) podemos adicionar 3 tipos de proteínas:

- a) Proteína de Leite X_1
- b) Proteína de Soja X_2
- c) Proteína de Peixe X_3

Dependendo da combinação das percentagens desses tres tipos de proteínas adicionados à massa do pão, podemos obter produto com as seguintes características:

1) *Volume (litro)/Kg*

Os coeficientes da função objetivo são:

β_0	β_1	β_2	β_3	β_{11}	β_{22}	β_{33}	β_{12}	β_{13}	β_{23}
5.508	-0.4492	-0.6354	-0.338	0.3364	-0.1062	0.01564	0.1	-0.05166	-0.0291

e a perturbação é $N(0, 0.15)$

2) *Impressões gerais* (entre zero e dez).

Os coeficientes da função objetivo são:

β_0	β_1	β_2	β_3	β_{11}	β_{22}	β_{33}	β_{12}	β_{13}	β_{23}
5.033	0	-0.75	-1.13	0.046	0.046	0.0458	0.125	0.125	0.125

e a perturbação é $N(0,0.1)$.

O modelo empregado é:

$$\hat{Y} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_{11} (X_1)^2 + \beta_{22} (X_2)^2 + \beta_{33} (X_3)^2 \\ + \beta_{12} X_1 X_2 + \beta_{13} X_1 X_3 + \beta_{23} X_2 X_3 ,$$

onde:

$$X_1 = (\% \text{ proteína de leite } - 4) / 4$$

$$X_2 = (\% \text{ proteína de Soja } - 4) / 4$$

$$X_3 = (\% \text{ proteína de Peixe } - 4) / 4$$

Cada tipo de proteína pode ser adicionado em até 8%. O mínimo teor de proteína admissível é 12%. O objetivo dos jogadores é maximizar a soma (volume + impressões gerais), sujeita às restrições de mínimo teor total e máximo teor individual. Quando os jogadores admitirem que o processo está otimizado, o programa se utiliza da subrotina TESTE, que otimiza analiticamente a função objetivo (volume + impressões gerais) e compara com aquela otimizada pelos alunos.

FLUXOGRAMA DO JOGO

Um fluxograma do jogo está detalhado na fig. 6.2. O programa pode ser empregado em várias etapas do jogo, e com finalidades diversas. Na etapa em que os jogadores estão aprendendo a preencher folha de resposta e quadro de informações, dados conhecidos de outros trabalhos podem ser introduzidos pelo terminal. (Caso alguém já saiba como fazer o preenchimento da folha de respostas e quadro de informações, mas comodamente quer se valer do computador para efetuar e dispor os cálculos, pode utilizar essa etapa do programa OPEVWR).

Quando os jogadores já puderem tentar decidir (como se fossem superintendente do processo ou membro de uma comissão OPEV) olhando o quadro de informações e/ou folha de resposta, devem se valer dos dados gerados pelo programa, e investigar os efeitos de mudança nos níveis das variáveis de qualquer dos processos propostos.

APÊNDICE I

COMENTÁRIOS

Box [1] recomendou que EVOP fosse conduzido em ciclos de duas ou três variáveis e propos esquemas fatoriais 2^2 e 2^3 , completos, com a introdução de um ponto central para referência. Tem sido introduzidos arranjos alternativos como OPEV SIMPLEX [3] ROVOP e REVOP [4]. Esses esquemas foram descritos brevemente por Box e Draper [2] e em grande detalhe por Lowe [4]. Lowe caracteriza vantagens e desvantagens dos vários métodos. Por exemplo, a OPEV de Box pode investigar efeitos que envolvam tanto variáveis quantitativas quanto qualitativas, mas apresenta dificuldades em considerar delineamentos com mais de três variáveis. Já a OPEV Simplex poderia considerar qualquer quantidade de fatores, não podendo, entretanto, considerar em seu delineamento fatores de natureza qualitativa.

Hunter et al. [5], em 1966, fez uma "coletânea" da extensiva literatura relativa ao OPEV durante seus 10 primeiros anos.

Após a publicação de *Evolution Operation* [2] em 1969, OPEV tem sido mais largamente empregada, e vários trabalhos tem mostrado exemplos de aplicação [6], [7], avaliação de resultados de aplicação através de "pesquisa" entre os usuários [8] ou ainda comentários sobre vantagens e desvantagens das diferentes técnicas [7].

Nesse trabalho não se tentou fazer um levantamento bibliográfico completo sobre Operação Evolutiva, o que pode ser encontrado nas referências [2], [5], [7] e [8].

APENDICE II

* MANUAL DE USO PARA OS PROGRAMAS OPEVTY E OPEVWR *

PARA O PRESENTE MANUAL, CONSIDERAMOS AS PERGUNTAS DIRIGIDAS AO USUARIO ENTRE ASPAS (""); PARA O PROGRAMA OPEVWR, SOMENTE ESSAS PERGUNTAS SERAO VISTAS PELO USUARIO, QUE JA DEVE CONHECER O FUNCIONAMENTO DO OPEVTY. AQUELE PROGRAMA IMPRIME TUDO MAIS RAPIDAMENTE EM UM ARQUIVO COM EXTENSAO LPT. ANTES DA EXECUCAO OS ARQUIVOS COM TAL EXTENSAO DEVEM SER DELETADOS OU RENOMADOS.

O PROGRAMA ARMAZENA, INICIALMENTE, OS VALORES F5,N QUE SAO UTILIZADOS PARA ESTIMAR O DESVIO PADRAO PELO METODO DA AMPLITUDE. DEPOIS, O PROGRAMA ARMAZENA OS COEFICIENTES DAS FUNCOES OBJETIVO PROPOSTAS COMO ALTERNATIVAS DE PROCESSOS DE PRODUCAO. EXISTEM 4 FUNCOES OBJETIVO, DUAS PARA CADA PROCESSO:

==> PRODUCAO DE PAO COM ALTO TEOR PROTEICO, ONDE AS CARACTERISTICAS DE INTERESSE SAO:

- 1 ==> VOLUME/KG
- 2 ==> IMPRESSOES GERAIS

2 ==> EXTRUSAO DE TRITICALE, ONDE AS CARACTERISTICAS DE INTERESSE SAO:

- 1 ==> TEXTURA
- 2 ==> UMIDADE APOS A EXTRUSAO

A MATRIZ GARCEZ(I,J,K) E UTILIZADA PARA ARMAZENAR OS COEFICIENTES DESSAS FUNCOES OBJETIVO, ONDE:

- K=1..10 SE REFERE AOS COEFICIENTES
J=1, 2 SE REFERE AS CARACTERISTICAS DE INTERESSE EM CADA PROCESSO
I=1, 2 SE REFERE AOS PROCESSOS

DEPOIS OS DESVIOS PADRAO PARA CADA FUNCAO OBJETIVO SAO ARMAZENADOS EM SIGMA(I,J), ONDE I E J SAO CONFORME EMPREGADOS ANTERIORMENTE.

JA NA SUBROTINA TESTE, SAO LIDAS AS RESTRICOES PARA AS VARIAVEIS DE UM DOS PROCESSOS (O DE FABRICACAO DE PAO COM ELEVADO TEOR PROTEICO). NESTA ETAPA, O PROGRAMA OTIMIZA ANALITICAMENTE A FUNCAO OBJETIVO PROPOSTA NO JOGO (MAXIMIZA A SOMA (VOLUME/KG + IMPRESSOES GERAIS)) E COMPARA COM AQUELA OTIMIZADA PELOS ALUNOS.

APÓS EXEC OPEVTY(OPEVWR), O PROGRAMA INICIA UMA SÉRIE DE PERGUNTAS QUE DESCREVEREMOS A SEGUIR; O FORMATO E NATUREZA DAS RESPOSTAS ESTÁ DETALHADO E, ENQUANTO O USUÁRIO RESPONDER DE UMA MANEIRA INCORRETA, O PROGRAMA GRACEJA E REPETE A PERGUNTA (SALVO SE A RESPOSTA IMPLICAR EM ERRO DE EXECUÇÃO; NESSE CASO A EXECUÇÃO É INTERROMPIDA).

"QUAL O NÚMERO MÁXIMO DE CICLOS PREVISTO PARA A FASE A SEGUIR?"
 MAX<CR> (FORMATO LIVRE PARA INTEIROS); MAX DEVE SER UM INTEIRO ENTRE DOIS E DEZ.

"VOCE CONHECE AS OPCOES?"

EXISTEM OITO OPCOES, QUE NÃO NECESSARIAMENTE PRECISAM SER INFORMADAS PELO PROGRAMA A CADA EXECUÇÃO. SE O USUÁRIO AS CONHECE, DEVE RESPONDER:

SIM<CR>(FORMATO A3)

SE FOR ESSA A RESPOSTA, O PROGRAMA NÃO DESCREVE AS OPCOES E PASSA A PRÓXIMA PERGUNTA.

CASO O USUÁRIO DESCONHEÇA AS OPCOES E/OU QUER QUE AS MESMAS SEJAM IMPRESSAS, DEVE RESPONDER:

NAO<CR>(FORMATO A3)

NESSE CASO, O PROGRAMA IMPRIME AS OITO OPCOES.

- OPCAO UM: VOCE ESTA NA PRIMEIRA FASE
 VOCE QUER QUE O PROGRAMA IMPRIMA APENAS OS REGISTROS
 VOCE QUER INSERIR OS DADOS PELO TERMINAL
- OPCAO DOIS: VOCE ESTA NA PRIMEIRA FASE
 VOCE QUER QUE O PROGRAMA IMPRIMA APENAS OS REGISTROS
 VOCE QUER QUE OS DADOS SEJAM GERADOS PELO PROGRAMA
- OPCAO TRES: VOCE ESTA NA PRIMEIRA FASE
 VOCE QUER QUE O PROGRAMA IMPRIMA AS FOLHAS DE RESPOSTA COM OS RESPECTIVOS REGISTROS
 VOCE QUER INSERIR OS DADOS PELO TERMINAL
- OPCAO QUATRO: VOCE ESTA NA PRIMEIRA FASE
 VOCE QUER QUE O PROGRAMA IMPRIMA AS FOLHAS DE RESPOSTA COM OS RESPECTIVOS REGISTROS
 VOCE QUER QUE OS DADOS SEJAM GERADOS PELO PROGRAMA
- OPCAO CINCO: VOCE NAO ESTA MAIS NA PRIMEIRA FASE
 VOCE QUER QUE O PROGRAMA IMPRIMA APENAS OS REGISTROS
 VOCE QUER INSERIR OS DADOS PELO TERMINAL
- OPCAO SEIS: VOCE NAO ESTA MAIS NA PRIMEIRA FASE
 VOCE QUER QUE O PROGRAMA IMPRIMA APENAS OS REGISTROS
 VOCE QUER QUE OS DADOS SEJAM GERADOS PELO PROGRAMA
- OPCAO SETE: VOCE NAO ESTA MAIS NA PRIMEIRA FASE
 VOCE QUER QUE O PROGRAMA IMPRIMA AS FOLHAS DE RESPOSTA COM OS RESPECTIVOS REGISTROS
 VOCE QUER INSERIR OS DADOS PELO TERMINAL
- OPCAO OITO: VOCE NAO ESTA MAIS NA PRIMEIRA FASE
 VOCE QUER QUE O PROGRAMA IMPRIMA AS FOLHAS DE RESPOSTA COM OS RESPECTIVOS REGISTROS
 VOCE QUER QUE OS DADOS SEJAM GERADOS PELO PROGRAMA

"QUAL A OPCAO QUE PREFERE?"

K<CR> (FORMATO LIVRE PARA INTEIROS); K DEVE SER INTEIRO ENTRE UM OITO

*** DADOS INSERIDOS PELO TERMINAL *** (OPCAO IMPAR)
 (AQUI O PROGRAMA INICIALIZA TUDO. TORNA JT=0;M=1;IA=0;TORNA
 ZERO TODAS AS VARIAVEIS UTILIZADAS COMO AUXILIAR NAS
 INCREMENTACOES).

SE A FASE FOR A PRIMEIRA(OPCAO 1 OU 3)

"QUANTOS SAO OS DADOS HISTORICOS?"

NUM<CR> (FORMATO LIVRE PARA INTEIROS); NUM DEVE SER INTEIRO <= 10

"QUAIS SAO ELLES?"

(F(I),I=1,NUM)<CR> (FORMATO LIVRE PARA REAIS)

====> LABEL(0)

"QUAIS SAO OS DADOS PARA O CICLO M?"

A SEQUENCIA NO DELINEAMENTO E SUPOSTA CONHECIDA, E O
 USUARIO DEVE INSERIR:

(DADOS(M,I),I=0,4)<CR> (FORMATO LIVRE PARA REAIS) (*)

SE A FASE NAO FOR A PRIMEIRA(OPCAO 5 OU 7).

"QUAL FASE ESTAMOS?"

FASE<CR> (FORMATO LIVRE PARA INTEIROS)

"QUAL O VALOR DA ESTIMATIVA A PRIORI?"

PRIORI(JT)

"QUAIS SAO OS DADOS PARA O CICLO M?"

(DADOS(M,I),I=0,4)<CR> (FORMATO LIVRE PARA REAIS)

SE A OPCAO TIVER SIDO 3 OU 7, O PROGRAMA IMPRIME A FOLHA
 DE RESPOSTAS PREENCHIDA PARA O CICLO ATUAL. APOS ISSO,
 IMPRIME O QUADRO DE INFORMACOES.

SE A OPCAO TIVER SIDO 1 OU 5, O PROGRAMA IMPRIME APENAS
 O QUADRO DE INFORMACOES.

*** DADOS GERADOS PELO PROGRAMA *** (OPCAO PAR)

"VAMOS TRABALHAR COM O EXTRUSOR?"

SE

SI<CR> (FORMATO A3)

(*) -----

NO PREENCHIMENTO DA FOLHA DE RESPOSTA, OS DADOS SAO
 IMPRESSOS NUM FORMATO F5.1. OS DADOS DEVEM SER DIMENSIONADOS
 DE FORMA QUE POSSAM SER ESCRITOS EM TAL FORMATO.

ENTAO IA=2

SE

NAO<CR> (FORMATO A3)

ENTAO IA=1

SE IA=1, TEMOS O PROCESSO DE FABRICACAO DE PAO COM ALTO TEOR PROTETICO. NESSE CASO, TEMOS DUAS CARACTERISTICAS DE INTERESSE:

- 1) - VOLUME(LITRO)/KG, ONDE JT=1
- 2) - IMPRESSOES GERAIS (NOTAS ENTRE ZERO E DEZ), ONDE JT=2

PARA IA=2, VAMOS TRABALHAR COM O EXTRUSOR, ONDE TEMOS TAMBEM DUAS CARACTERISTICAS DE INTERESSE:

- 1) - TEXTURA(MEDIDA ENTRE ZERO E DEZ), ONDE JT=1
- 2) - UNIDADE APOS A EXTRUSAO(%), ONDE JT=2

*** JT=1 ***

SE A FASE FOR A PRIMEIRA(OPCAO 2 OU 4), O PROGRAMA, ATRAVES DA SUBROTINA PREVIA PERGUNTA:

"QUANTOS DADOS HISTORICOS VOCE QUER CONSIDERAR?"

NUM<CR> (FORMATO LIVRE PARA INTEIROS)

A SEGUIR, O PROGRAMA GERA NUM DADOS HISTORICOS E OS IMPRIME. APOS ISSO, IMPRIME A ESTIMATIVA A PRIORI CALCULADA.

SE A FASE NAO FOR A PRIMEIRA:

"EM QUAL FASE ESTAMOS?"

FASE<CR>(FORMATO LIVRE PARA INTEIROS)

"QUAL O VALOR DA ESTIMATIVA A PRIORI?"

PRIORI(JT)<CR>(FORMATO LIVRE PARA REAIS)

=====> LABEL(1) =====> LABEL(2)

DEPOIS O PROGRAMA PROPOE TRES VARIABEIS PARA O PROCESSO ESCOLHIDO.

SE EXTRUSAO:

"TEMOS TRES VARIABEIS:

- 1) - UNIDADE INICIAL DO TRITICALE
- 2) - TEMPERATURA INTERNA
- 3) - DIAMETRO DO BOCAL"

SE PAD DE ALTO TEOR PROTEICO:
 "TEMOS TRES VARIABEIS:

- 1) - PROTEINA DE LEITE
- 2) - PROTEINA DE SOJA
- 3) - PROTEINA DE PEIXE"

ENTAO, PARA QUALQUER DOS PROCESSOS:
 "QUAL DELAS VOCE QUER MANTER FIXA?
 ESCREVA O RESPECTIVO NUMERO"
 IJ<CR>(FORMATO LIVRE PARA INTEIROS)

O PROGRAMA FORNECE OS LIMITES ENTRE OS QUAIS DEVEM ESTAR
 CONTIDAS TODAS AS VARIABEIS

SE EXTRUSAO:
 A UMIDADE INICIAL DO TRITICALE DEVE ESTAR ENTRE 15% E 22%
 A TEMPERATURA INTERNA DEVE ESTAR ENTRE 350F E 450F
 A ABERTURA DO BUCAL DEVE ESTAR GRADUADA ENTRE 1MM E 3MM"

SE PAD DE ALTO TEOR PROTEICO:
 "A PORCENTAGEM DE CADA TIPO DE PROTEINA DEVE ESTAR ENTRE 0% E 8%"

ENTAO, PARA QUALQUER DOS PROCESSOS:
 "QUAL O VALOR DA VARIABEL A SER FIXADA?"
 X(IJ)<CR> (FORMATO LIVRE PARA REAIS)

TENDO FIXADA UMA VARIABEL, AS DUAS RESTANTES SERAO "A" E "B"
 NA SEQUENCIA EM QUE FORAM DESCRITAS. POR EXEMPLO, SE O
 PROCESSO FOR O DE PRODUCAO DE PAD COM ELEVADO TEOR PROTEICO,
 E A VARIABEL FIXADA FOR PROTEINA DE SOJA, A CONDICAO "A" SERA
 UM NIVEL PARA PROTEINA DE LEITE E A CONDICAO "B" SERA UM
 NIVEL PARA A PROTEINA DE PEIXE. A SEQUENCIA NO DELINEAMENTO
 SERA:

*(4)

*(2)

*(0)

*(1)

*(2)

SUPONDO CONHECIDA A SEQUENCIA NO DELINEAMENTO, ATRAVES
 DA SUBROTINA GERA, O PROGRAMA GERA AMOSTRAS DE VALORES DAS
 CARACTERISTICAS DE INTERESSE (INICIALMENTE PARA JT=1), DEPOIS
 QUE LHE FOREM FORNECIDAS AS SEGUINTE S INFORMACOES:

I=0,4 <=====1

"QUAIS AS CONDIÇÕES PARA O PONTO 1?"

"ESCREVA A CONDIÇÃO "A""

X(IM)<CR>(FORMATO LIVRE PARA REAIS)

"AGORA ESCREVA A CONDIÇÃO "B""

X(IM)<CR> (FORMATO LIVRE PARA REAIS)

CONTINUE=====1

DEPOIS O PROGRAMA IMPRIME OS CINCO PONTOS GERADOS. SE A OPCAO FOR 4 OU 8, O PROGRAMA IMPRIME A FOLHA DE RESPOSTAS E O QUADRO DE INFORMACOES.

SE A OPCAO FOR 2 OU 6, O PROGRAMA IMPRIME APENAS O QUADRO DE INFORMACOES.

*** JT=2 ***

SE A FASE FOR A PRIMEIRA (OPCAO 2 OU 4), O PROGRAMA ATRAVES DA SUBROTINA PREVIA INDAGA O NUMERO DE DADOS HISTORICOS PARA A CARACTERISTICA JT=2. (AQUI SE PROCESSA EXATAMENTE COMO NA ETAPA EQUIVALENTE PARA JT=1).

O PROGRAMA, COM O DELINEAMENTO JA "MEMORIZADO", IMPRIME AS CINCO AMOSTRAS GERADAS PARA A CARACTERISTICA JT=2. DEPENDENDO DA OPCAO, IMPRIME ENTAO SO O QUADRO DE INFORMACOES OU FOLHA DE RESPOSTA E QUADRO DE INFORMACOES, DA MESMA MANEIRA QUE NA ETAPA EQUIVALENTE PARA JT=1.

APOS O PREENCHIMENTO DO QUADRO DE INFORMACOES (AGORA VOLTAMOS A GENERALIZAR TANTO PARA AS OPCOES DE DADOS INSERIDOS PELO TERMINAL COMO PARA AQUELES DADOS GERADOS PELO PROGRAMA):

"COM M CICLOS E SUFICIENTE?"

SE A RESPOSTA FOR
NAO<CR>(FORMATO A3)

O PROGRAMA COMPARA M (O NUMERO DE CICLOS ATE ENTAO REALIZADOS) COM MAX (DEFINIDO NO INICIO DA FASE). SE M<MAX, INCREMENTA 1 A M E GERA NOVOS DADOS, REALIZANDO NOVO CICLO. SE M=MAX, O PROGRAMA DA A MENSAGEM DE:

"PASSAREMOS DO MAXIMO DE CICLOS PREVISTOS. PENSEM QUE FAZER E VOLTEM A FALAR COMIGO"
E SUPRIME A EXECUCAO.

SE A RESPOSTA FOR
SIM<CR>(FORMATO A3)

O USUARIO PODE ESTAR TERMINANDO A FASE POR DOIS MOTIVOS:
1)VAI REINICIAR NOVA FASE, UTILIZANDO OUTROS NIVEIS DAS
VARIABLES PARA OS DELINEAMENTOS
2)CONSIDERA O PROCESSO OTIMIZADO.

PARA SABER ISSO, O PROGRAMA PERGUNTA:
"VOCE ACHA QUE JA E SUFICIENTE?"

SE
NAO<CR>(FORMATO A3)

O MOTIVO E O PRIMEIRO. ENTAO, O PROGRAMA PERGUNTA:
"OTIMO. VAMOS PARA UMA NOVA FASE AINDA HOJE?"

SE
NAO<CR> (FORMATO A3)

O PROGRAMA ESCREVE
"GUARDE NOVA MEDIAS.ELA SERA A ESTIMATIVA A
PRIORI PARA A PROXIMA FASE.
E SUPRIME A EXECUCAO

SE
SIM<CR> (FORMATO A3)

O PROGRAMA ATUALIZA SUAS INFORMACOES (FAZ
CICLO=M=1; INCREMENTA 1 EM FASE;
ATRIBUE NOVA MEDIA S A PRIORI(JT),ETC.),
E VOLTA PARA LABEL(IA). SE IA=2,
IMPRIME CARACTERISTICAS DO PRODUTO
EXTRUDADO NAS CONDICIONES DE CADA PONTO
DO DELINEAMENTO.

SE
SIM<CR>(FORMATO A3)

O PROCESSO E CONSIDERADO OTIMIZADO.

SE IA=0 (DADOS INSERIDOS PELO TERMINAL), MENSAGEM
"PARABENS PELO SUCESSO COM OPEV"

E SUPRIME A EXECUCAO

SE IA=1 O PROGRAMA ATRAVES DA SUBROTINA TESTE (QUE POR
SUA VEZ SE UTILIZA DE INVERT E DE MULT) EMPREGA O ALGORITMO
DO GRADIENTE PROJETADO PARA VERIFICAR SE O OTIMO ENCONTRADO
ANALITICAMENTE ESTA PROXIMO DO OTIMO ENCONTRADO POR OPEV. DA
O SEU BECADO E SUPRIME A EXECUCAO.

SE IA=2, IMPRIME CARACTERISTICAS DO PRODUTO EXTRUDADO EM
CADA PONTO DO DELINEAMENTO, E FAZ VOTOS DE FELIZ NEGOCIOS COM
O PRODUTO EXTRUDADO.

F I M

REFERÊNCIA BIBLIOGRÁFICA

- [1] BOX, G.E.P. (1957) - "Evolutionary Operation - A Method for Increasing Industrial Productivity", Appl. Statist., 6, p.p. 81-101.
- [2] BOX, G.E.P. e DRAPPER, N.R. (1969) - "Evolutionary Operation - A Method for Increasing Industrial Productivity" Wiley.
- [3] SPENDLEY, W., HEST, G.R. and HIMSWORTH, F.R. (1962) - "Sequential Application of Simplex designs in Optimization and EVOP" - Technometrics, 62, 4, pp. 441-461.
- [4] LOWE, C.W. (1964) - "Some Techniques of Evolutionary Operation", Trans. Inst. Chem. Engrs, Vol. 42 pp. T334-T344.
- [5] HUNTER, G.W. e KITTRELL, J.R. (1966) - "Evolutionary Operation: A Review", Technometrics", Vol. 8 nº 3, pp. 389 - 397.
- [6] MORGAN, S.L. e DEMING, S.N. (1974) - "Simplex Optimization of Analytical Chemical Methods" - Analytical Chemistry, Vol. 46, Nº 9, pp. 1170-1181.
- [7] LOWE, C.M. (1974) - "Evolutionary Operation in Action" - Applied Statistics, 23, nº 2, pp. 218-226.

- [8] HAHN, G.J. e DERSHOWITZ, A.F. (1964) - "Evolutionary Operation To-Day, Some Survey Results and Observations" Applied Statistics, 23, Nº 2, pp. 214-217.

De alguma forma também foram consultadas (figuras, definições, algoritmos, etc) as seguintes referências:

- [9] KOEHLER, T.L. (1959) - "Evolutionary Operation", Chemical Engineering Progress, Vol. 55, nº 10, pp. 76-79.
- [10] COSTA, ELIEZER ARANTES - Notas de aulas do curso "Tópicos em Sistemas", 1/78.
- [11] RETCHSCHAFFEN, EDGAR - Notas de aulas do curso "Análise Matricial", 1/77.
- [12] EL-DASH, AHMED A. - Comunicação Verbal (Figuras e explicações do processo de extrusão).
- [13] SARRI FOIX, D. ANDRÉS. (1971) - "El empleo de la fuerza centrífuga en la Indústria Láctea" - Separata de Via Láctea: Revista Técnica de Informaciones Lactológicas - Vol. 3, nº 9.
- [14] ALFA-LAVAL - INSTRUCTION BOOK
ALFA-LAVAL - USER'S GUIDE
Laboratory Separator LAPX 202
End: Separadores ALFA-LAVAL S.A.
Av. das Nações Unidas, 3115
São Paulo - S.P.

- [15] AOKI, MASNAO - (1971) "Introduction to OPTIMIZATIONS TECHNIQUES - The Mcmillan Company.
- [16] LORENZ, K. e WELSH, J. e NORMANN, R. e BEETNER, G. e FREY, A. (1974) - "Extrusion Processing of Triticale", Journal of Food Science, Vol. 39, pp. 572-576.
- [17] ASCO RED/HAT
"Solenoid Values"
Catálogos nº 29
Automatic Switch Co.
Florhan Park, New Jersey.