TÉCNICAS DE ANÁLISE DE DADOS

PLANEJAMENTO DE EXPERIMENTOS

– observação = previsível + aleatória

– aleatória obedece algum modelo de probabilidade

– ferramenta: análise de variância

TÉCNICAS DE ANÁLISE DE DADOS

PLANEJAMENTO DE EXPERIMENTOS

– “identificar fatores, controláveis, que expliquem o fenômeno ou alterem a característica de interesse”

– “identificar estruturas nos dados, permite conhecer melhor o fenômeno”

TÉCNICAS DE ANÁLISE DE DADOS

PLANEJAMENTO DE EXPERIMENTOS

- fator *versus* variável
  - fator: variável experimental que está sendo investigada para se determinar seu efeito na resposta
    - é controlado (níveis podem ser pré-estabelecidos)
  - variável resposta: resultado de um experimento
  - covariáveis: variáveis adicionais que afetam a resposta mas não podem ser controlados
- níveis do fator (tratamento)
- unidade experimental/unidade amostral
  - homogênea
- bloco: agrupa unidades experimentais de maneira a garantir a homogneidade

TÉCNICAS DE ANÁLISE DE DADOS

PLANEJAMENTO DE EXPERIMENTOS

- Efeito: medida da variação da resposta em função da ação do fator
- fator fixo *versus* fator aleatório
- fator cruzado *versus* fator hierárquico
- grupo controle ou testemunha (positivo/negativo)
- repetição (de medidas) e replicação (do experimento)
- casualização ou aleatorização (randomização)
  - todas as unidades tem a mesma probabilidade de serem escolhidas
- experimento cego (blind)/duplo cego: experimentador/participante

TÉCNICAS DE ANÁLISE DE DADOS

PLANEJAMENTO DE EXPERIMENTOS

- Definir:
  - a unidade experimental
  - a variável medida e como medir
  - os fatores e seus níveis
  - a forma como os fatores serão designados ás unidades experimentais
  - o número de unidades experimentais

TÉCNICAS DE ANÁLISE DE DADOS

PLANEJAMENTO DE EXPERIMENTOS

- Experimentos com um fator T com k níveis
  - Xerox (notação)
  - Plano experimental completamente casualizado:
    - unidades experimentais tem igual probabilidade de receber um tratamento (= estar num dado nível do fator)

TÉCNICAS DE ANÁLISE DE DADOS

PLANEJAMENTO DE EXPERIMENTOS

- Estudar a influência dos k níveis do fator T sobre uma variável resposta Y a partir de j observações
- Metodologia: comparar as k médias de Y
  - experimentos com um fator fixo e k níveis:

$$y_{ij} = \mu + T_i + e_{ij}$$

μ: média geral de todas as observações

T_i: efeito do i-ésimo nível do fator T (cte.)

μ_i: média no nível i (i= 1...k)

e_{ij}: erro casual não observável

$T_i = (\mu_i - \mu)$

TÉCNICAS DE ANÁLISE DE DADOS

PLANEJAMENTO DE EXPERIMENTOS

- Restrição do modelo
- Suposições do modelo

(lousa)

TÉCNICAS DE ANÁLISE DE DADOS

PLANEJAMENTO DE EXPERIMENTOS

– Hipótese H_0: T_1 = ... = T_k = 0
  - H_0: ausência de efeito => μ_i´s são iguais
  - H_a:ao menos um μ_i é diferente
– Teste de hipótese:

Decisão	Hipótese verdadeira	
	H_0 verdadeira	H_a verdadeira
não rejeito H_0	decisão correta	erro do tipo II (β)
rejeito H_0	erro do tipo I (α)	decisão correta

α = nível de significância do teste

(1- β) = poder do teste

(1- α) = nível de confiança do teste = p: rejeitar H_0: quando H_0 é verdadeira

TÉCNICAS DE ANÁLISE DE DADOS

PLANEJAMENTO DE EXPERIMENTOS (xerox)

F.V.	gl	SQ	QM	F_0
entre	k-1	SQE	QME	QME/QMR
dentro	n-k	SQR	QMR	
Total	n-1	SQT		

TÉCNICAS DE ANÁLISE DE DADOS

PLANEJAMENTO DE EXPERIMENTOS

SQE, SQR, SQT:
lousa

TÉCNICAS DE ANÁLISE DE DADOS

PLANEJAMENTO DE EXPERIMENTOS

– Decisão:

rejeita-se H_0 se $F_0 > F_{k-1,\ n-k,\ \alpha}$

TÉCNICAS DE ANÁLISE DE DADOS

PLANEJAMENTO DE EXPERIMENTOS

- Estudar a influência dos k níveis do fator T sobre uma variável resposta Y a partir de j observações
- Metodologia: comparar as k médias de Y
  - experimentos com um fator aleatório e k níveis:

$$y_{ij} = \mu + T_i + e_{ij}$$

μ: média geral de todas as observações = $E(\mu_i)$

T_i: efeito do i-ésimo nível do fator T (cte.)

μ_i: média no nível i (i= 1...k)

e_{ij}: erro casual não observável

$T_i = (\mu_i - \mu)$

TÉCNICAS DE ANÁLISE DE DADOS

PLANEJAMENTO DE EXPERIMENTOS

- Restrição do modelo
- Suposições do modelo
- Conseqüências
- Hipótese

(lousa)

TÉCNICAS DE ANÁLISE DE DADOS

PLANEJAMENTO DE EXPERIMENTOS

– completamente aleatórios (completely randomized design)
  - número diferente de repetições

– blocos casualizados
  - poucas unidades similares
  - blocos completos, quadrado latino

TÉCNICAS DE ANÁLISE DE DADOS

PLANEJAMENTO DE EXPERIMENTOS

– experimentos com mais de 2 fatores

(lousa)

TÉCNICAS DE ANÁLISE DE DADOS

PLANEJAMENTO DE EXPERIMENTOS

- experimentos mais complexos (múltiplos fatores, fatores cruzados e hierárquicos, split-plot)
- comparações múltiplas

TÉCNICAS DE ANÁLISE DE DADOS

Análise de Variância

Objetivo: testar se existe diferenças nas médias de absorbância para os a=5 tipos (níveis) de solventes.

Tabela 1-2 Dados gerais de um experimento com um único fator

Tratamentos (níveis)	Observações						Totais	Médias
1	y_{11}	y_{12}	.	.	.	y_{1n}	$y_{1.}$	$\overline{y}_1$
2	y_{21}	y_{22}	.	.	.	y_{2n}	$y_{2.}$	$\overline{y}_2$
. .	. .	. .	. .	. .	. .	. .	. .	. .
a	y_{a1}	y_{a2}	.	.	.	y_{an}	$y_{a.}$	$\overline{y}_a$

TÉCNICAS DE ANÁLISE DE DADOS

Modelo estatístico (one-way):

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij} \quad \begin{cases} i=1,2,...,a \\ j=1,2,...,n \end{cases}$$

y_{ij}= é a ij-ésima observação;

μ é uma constante para todas as observações (média geral);

τ_i é o efeito do i-ésimo tratamento;

ε_{ij} é o erro aleatório.

Pressuposições: 1) os erros aleatórios são independentes;

2) os erros aleatórios são *normalmente* distribuídos;

3) os erros aleatórios tem média 0 (zero) e variância σ^2;

4) a *variância, σ^2, deve ser constante* para todos os níveis do fator.

5) as observações são *adequadamente descritas pelo modelo*

TÉCNICAS DE ANÁLISE DE DADOS

Duas situações:

1) modelo de efeito fixo (níveis selecionados pelo pesquisador);

2) modelo de efeito aleatório (amostra aleatória). Neste caso, vamos estimar e testar hipóteses sobre a variabilidade de τ_i

TÉCNICAS DE ANÁLISE DE DADOS

Análise de Variância do Modelo de Efeito Fixo – 1 fator fixo

Hipóteses: H_0: $\mu_1 = \mu_2 = ... = \mu_a$

H_a: $\mu_i \neq \mu_j$ para pelo menos um par (i,j)

1-3.1 Decomposição da soma de quadrados total

$$\underbrace{\sum_{i=1}^{a}\sum_{j=1}^{n}\left(y_{ij}-\bar{y}_{..}\right)^2}_{\text{Corrigida para a média}} = n\sum_{i=1}^{a}\left(\bar{y}_{i.}-\bar{y}_{..}\right)^2 + \sum_{i=1}^{a}\sum_{j=1}^{n}\left(y_{ij}-\bar{y}_{i.}\right)^2$$

TÉCNICAS DE ANÁLISE DE DADOS

Análise Estatística

F_0 = QMTratamentos / QMErro

Critério para rejeição de H_0: $F_0 > F_{\alpha,a-1,N-a}$. Pode-se usar o nível descritivo (em inglês: *p-value: É o menor valor de α para o qual rejeitamos a hipótese nula. Exemplo: para α=5%, assim, se o nível descritivo < do que 0,05 $\Rightarrow$ rejeitar H_0, caso contrário, $\Rightarrow$ aceitar H_0.*

TÉCNICAS DE ANÁLISE DE DADOS

Fórmulas para o cálculo das somas de quadrados:

$$SS_T = \sum_{i=1}^{a}\sum_{j=1}^{n} y_{ij}^2 - \frac{y_{..}^2}{N}$$

$$SS_{Tratamentos} = \frac{1}{n}\sum_{i=1}^{a} y_{i.}^2 - \frac{y_{..}^2}{N}$$

$$SS_{Erro} = SS_T - SS_{Tratamentos}$$

TÉCNICAS DE ANÁLISE DE DADOS

$$SS_T = SS_{Tratamentos} + SS_E$$

Graus de liberdade:

SS_T tem *an-1* graus de liberdade; $SS_{Tratamentos}$ tem *a-1* g.l. e SS_{erro} tem *a(n-1)* g.l.

Quadrados médios: $QMTrat = \frac{SQTratamentos}{a-1}$ $QMErro = \frac{SQErro}{a(n-1)}$

TÉCNICAS DE ANÁLISE DE DADOS

Esperanças dos quadrados médios:

$$E(QMErro) = \sigma^2$$

$$E(QMTratamentos) = \sigma^2 + \frac{n\sum_{i=1}^{a}\tau_i^2}{a-1}$$

Teste de hipótese: QMTratamentos/QMErro

TÉCNICAS DE ANÁLISE DE DADOS

Tabela da análise de variância de um experimento com um fator.

Causas de variação	Soma de quadrados	Graus de liberdade	Quadrados médios	F_0
Entre tratamentos	$SS_{Tratamentos}$	a-1	$QM_{Tratamentos}$	$\frac{QM_{Tratamentos}}{QM_{Erro}}$
Erro (dentro de trata/os)	SS_{Erro}	N-a	QM_{Erro}	
Total	SS_T	N-1		

N=an

TÉCNICAS DE ANÁLISE DE DADOS

Estimação dos parâmetros do modelo

Estimativas da média geral e dos efeitos dos tratamentos:

$$\hat{\mu} = \overline{y}_{..}$$
$$\hat{\tau}_i = \overline{y}_{i.} - \overline{y}_{..}$$

Estimativa pontual de μ_i:

$$\hat{\mu}_i = \hat{\mu} + \hat{\tau}_i = \overline{y}_{i.}$$

Um intervalo de confiança para μ_i é dado por:

$$\overline{y}_{i.} \pm t_{\alpha/2,N-a}\sqrt{QM_{Erro}/n}$$

TÉCNICAS DE ANÁLISE DE DADOS

PLANEJAMENTO DE EXPERIMENTOS

Referências para a aula:

Statistics as a catalyst to learning by scientific method Part 1 – an example - George Box e Patrick Y. T. Liu - Journal of Quality Technology, vol. 31, no.1, Jan. 1999, pág. 1-15.

Introdução ao controle estatístico de qualidade – 4a. Ed. – Douglas C. Montgomery. LTC, 2004. Cap. 12: Experimentos fatorial e fatorial fracionado para planejamento e melhoria do processo (pág. 365 – 406).

Brincando com papel – M. Kanegae e A. Haga – EDART – 1983, pág.17.

TÉCNICAS DE ANÁLISE DE DADOS

PLANEJAMENTO DE EXPERIMENTOS

Mesmo quando os dados estão sujeitos a erros observacionais, considerações sobre probabilidade e algumas conclusões podem ser obtidas.

Diferentes planejadores podem formular diferentes experimentos:

- Considerar diferentes fatores
- Escolher diferentes intervalos de variação dos fatores
- Usar diferentes transformações dos fatores
- Considerar diferentes modelos

TÉCNICAS DE ANÁLISE DE DADOS

PLANEJAMENTO DE EXPERIMENTOS

Tais mudanças podem influenciar as conclusões mais que os erros observacionais.

Observações sobre o experimento que será desenvolvido:

- Na prática:
- não apenas uma mas algumas variáveis resposta serão medidas, tabuladas e consideradas de forma conjunta;
- outros fatores serão incorporados por especialistas da área (fabricantes de helicóptero).

TÉCNICAS DE ANÁLISE DE DADOS

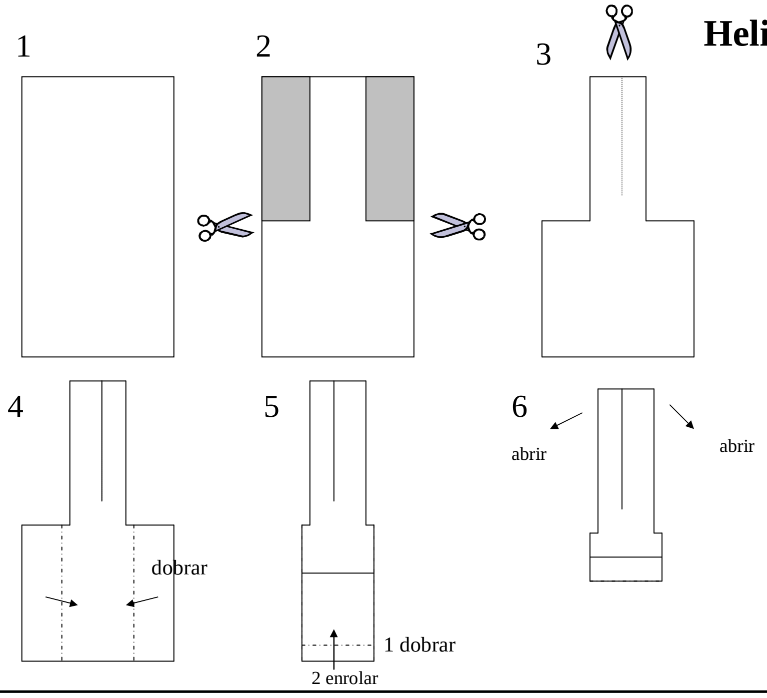
PLANEJAMENTO DE EXPERIMENTOS

Experimento: propor um protótipo para um helicóptero de papel.

Objetivo: obter o melhor design do helicóptero de maneira a permitir o maior tempo de vôo.

Restrições: considerar fatores que possam ser avaliados na sala de aula (altura da sala); todos os tempos de vôo serão medidos pelo mesmo tipo de instrumento; serão utilizados 2 tipos de papel.

Helicóptero



TÉCNICAS DE ANÁLISE DE DADOS

PLANEJAMENTO DE EXPERIMENTOS

Experimento:

grupos de 2 ou 3;

propor fatores a serem considerados (incluindo o tipo de papel) e a variável resposta a ser medida;

cada medida será repetida 3 vezes (3 repetições);

restrição: cada fator será testado em 2 níveis (+ e -).

Lembrete: o que está sendo testado é o planejamento e não o particular helicóptero “construído” segundo o planejamento proposto.

TÉCNICAS DE ANÁLISE DE DADOS

PLANEJAMENTO DE EXPERIMENTOS

Diretrizes (pág. 369 do Montgomery):

- Reconhecimento e relato do problema;
- Escolha dos fatores e dos níveis;
- Seleção da variável resposta;
- Escolha do planejamento experimental;
- Realização do experimento;
- Análise dos dados;
- Conclusões e recomendações.

TÉCNICAS DE ANÁLISE DE DADOS

PLANEJAMENTO DE EXPERIMENTOS

Experimentos fatoriais

- Há vários fatores de interesse;
- Os fatores variam juntos;
- Ex. 2 fatores:A com *a* níveis e B com *b* níveis, cada replicação contém todas as *ab* combinações possíveis - interação;
- Alternativa usada na prática: mudar os fatores um de cada vez, ao invés de variá-los simultaneamente (ex.: dietas);
- Análise: via ANOVA.

TÉCNICAS DE ANÁLISE DE DADOS

PLANEJAMENTO DE EXPERIMENTOS

Experimentos fatoriais

Modelo (2 fatores):

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \varepsilon_{ijk}$$

$i=1,\ldots, a$; $j=1,\ldots,b$; $k=1,..n$

Coleta de dados: *abn* selecionadas em ordem aleatória.

Resíduos: $\varepsilon_{ijk} = y_{ijk} - y_{ijk(estimado)}$.

TÉCNICAS DE ANÁLISE DE DADOS

PLANEJAMENTO DE EXPERIMENTOS

Experimento fatorial 2^k

K fatores, cada um com 2 níveis;

cada replicação completa tem 2^k experiências (realizações);

em geral, níveis: alto e baixo ou + e -.

TÉCNICAS DE ANÁLISE DE DADOS

PLANEJAMENTO DE EXPERIMENTOS

Experimentos fatoriais: matriz de planejamento

3 fatores: 2^3 = 8 experiências

	A	B	C	AB	AC	BC	ABC	y_{ijk}
1	+	+	+	+	+	+	+	
2	+	+	-	+	-	-	-	
3	+	-	+	-	+	-	-	
4	+	-	-	-	-	+	+	
5	-	+	+	-	-	+	-	
6	-	+	-	-	+	-	+	
7	-	-	+	+	-	-	+	
8	-	-	-	+	+	+	+	

TÉCNICAS DE ANÁLISE DE DADOS

PLANEJAMENTO DE EXPERIMENTOS

Experimentos fatoriais: 3 fatores: 2^3 = 8 experiências

Só disponho de 4 experiências:

	A	B	C	AB	y_{ijk}
1	+	+		+	
2	+	-		-	
3	-	+		-	
4	-	-		+	

Confundimento

Tabela de aliases

TÉCNICAS DE ANÁLISE DE DADOS

PLANEJAMENTO DE EXPERIMENTOS

Experimento fatorial fracionado 2^{k-p}

- Aumentando o número de fatores em um experimento 2^k, aumenta o número experiências;
- 2^5 exige 32 experiências, onde 5 são dos efeitos principais e 10 das interações de 2 fatores;
- Suposição: efeitos das interações maiores são desprezíveis;
- Planejamento fatorial fracionado 2^{k-p}: exige $1/(2^p)$ experiências.

TÉCNICAS DE ANÁLISE DE DADOS

- **PLANEJAMENTO DE EXPERIMENTOS**

Experimento fatorial fracionado 2^{k-p}

fatores: tipo de papel; tamanho da asa; largura do corpo; tamanho do corpo; (número de) dobras

repetições: 3 => 96 experiências;

planejamento fatorial fracionado 2^{5-2}: exige $1/(2^2)$ menos experiências =>8, como são 3 repetições => 24 experiências.