

---

# Introdução aos Planos Experimentais em Quadrados Latinos e suas Potencialidades de Aplicação em Engenharia de Software

---

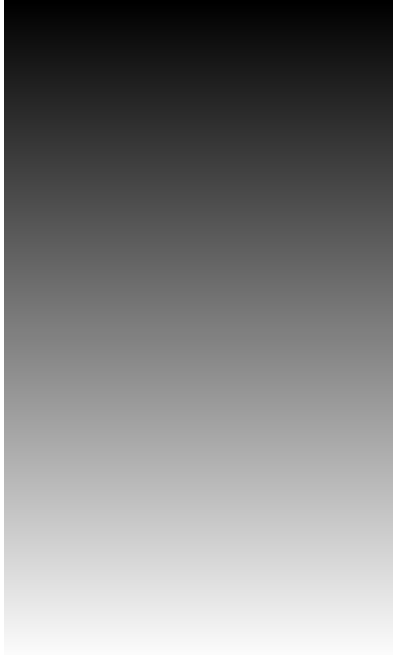
**Cristiano Ferraz**

*Departamento de Estatística - UFPE*

*Email: cferraz@de.ufpe.br*

*Fone: 2126-7433*

Recife, 30 de abril de 2008.



# 1. Objetivos

- Revisar os princípios estatísticos de experimentos;
- Apresentar o plano experimental aleatorizado em blocos e sua análise;
- Apresentar o plano experimental em quadrados latinos e sua análise;
- Discutir as dificuldades de planejamento de um experimento fatorial  $2^k$  em blocos;
- Avançar na discussão sobre as potencialidades de aplicação do método em Engenharia de Software.

## 2. Princípios Estatísticos de Experimentos

### O que é um EXPERIMENTO?

No contexto científico, um experimento pode ser definido da seguinte forma:

Um EXPERIMENTO consiste em um conjunto de procedimentos realizados com o objetivo de levantar evidências sobre a plausibilidade de uma hipótese idealizada.

## Tipos de EXPERIMENTOS:

### Experimentos ABSOLUTOS:

São experimentos empregados para avaliação de quantidades (ou atributos) constantes. Um exemplo típico é o tipo de experimento executado para se estimar a velocidade da luz.

### Experimentos COMPARATIVOS:

São experimentos onde o efeito de dois ou mais *tratamentos*, medido através de uma função de uma variável resposta, são comparados em uma certa população.

## Terminologia e Definições Básicas

Fator: Variável cujo efeito (observado em uma variável resposta) é de interesse para a investigação.

Níveis de um fator: Os níveis de um fator são os valores que este assume no experimento.

Fatores podem ser  
QUALITATIVOS ou  
QUANTITATIVOS.

Tratamento: O termo tratamento é dado aos níveis de um fator ou à combinação de níveis de fatores sobre os quais há interesse em comparar o desempenho de uma ou mais variáveis de resposta.

Fator de ruído: Um fator cujo efeito na variável-resposta seja conhecido e que não seja de interesse comparar o seu desempenho é dito fator de ruído.

## Unidade Experimental (U.E.):

É a unidade física/material experimental à qual os tratamentos são designados e aplicados. Em várias situações não é possível dissociar unidade experimental de tratamento. Nesses casos, dizendo de forma superficial, sem rigor, a ordem de execução do experimento “assume o papel” das unidades experimentais.

## Erro Experimental:

O erro experimental diz respeito às unidades experimentais. A existência de tal erro está associada ao fato de que não existem na natureza unidades experimentais perfeitamente idênticas.

Parte da variação observada na variável-resposta é explicada então pelas diferenças naturais entre unidades experimentais, que contribuem para o chamado erro experimental. O erro experimental também é reflexo de falha humana em reproduzir com exatidão um tratamento.

### Unidade Observacional (U.O.):

É a unidade física em que as medidas (mensurações) da variável resposta são avaliadas, após o experimento ser realizado.



Em muitas situações, as unidades experimentais e observacionais coincidem. Porém, pode ocorrer de não serem as mesmas.

### Erro Observacional:

O erro observacional diz respeito à unidade observacional. Tal erro deve-se, por exemplo, aos erros de medida (mensuração) cometidos e aos erros amostrais (se o experimento fosse repetido, provavelmente outras U.O. seriam utilizadas).

Replicação: Quando um tratamento é designado e aplicado a mais de uma U.E., sob as mesmas condições experimentais, diz-se que houve replicação.

Quando há replicação em um experimento, é possível estimar a variância do erro experimental a partir das variações observadas entre as unidades experimentais identificadas como réplicas.

Repetição: O termo repetição é aplicado quando mais de uma U.O. é utilizada numa situação em que as definições de U.E. e U.O. não coincidem.

Quando há repetição em um experimento, variações entre as unidades observacionais permitem estimar a qualidade do processo de mensuração utilizado no experimento.

Note as relações entre terminologias motivadas pela diferença conceitual entre unidade experimental e observacional:

U.E. ⇨ Replicação ⇨ Erro Experimental

U.O. ⇨ Repetição ⇨ Erro Observacional

E quando um experimento é executado novamente, em uma outra ocasião, digamos, por um outro grupo de pesquisadores?

Nesse caso, poderemos denominar de *uma nova execução, ou rodada* do experimento. Note que uma nova rodada do experimento logo em seguida ao término de outra gera uma replicação. Mas essa não é a única forma de gerar réplicas genuínas.

## **Etapas de uma Investigação Científica via condução de Experimentos Planejados.**

1. Descrição do Problema
2. Identificação de Fontes de Variação/Fatores que exercem influência sobre o fenômeno sob estudo
3. Identificação de um Plano Experimental e Modelo Linear Apropriados
4. Execução do Experimento
5. Desenvolvimento de Análise dos Dados

Esquema de modelo estatístico para as observações de um experimento planejado:

$$\begin{aligned} \text{Observação} = & \text{Efeito do Plano} \\ & + \text{Efeito do Tratamento} \\ & + \text{Erro Experimental} \\ & + \text{Erro Observacional} \end{aligned}$$

### **3. Planos Experimentais Aleatorizados em Blocos**

Em algumas situações é possível identificar diferenças (heterogeneidade) entre unidades experimentais disponibilizadas para um experimento, devido a uma fonte de variação sistemática. Quando essa fonte de variação sistemática é atribuível a um fator de ruído controlável, é possível reduzir a variância do erro experimental através de um plano aleatorizado em blocos, não obstante o nível de heterogeneidade observado entre unidades experimentais.

## Ilustrando via Exemplos:

Um experimento para comparar o desempenho de três linguagens de programação no tempo de execução de um programa básico de busca, num aparelho celular, deve ser rodado. Imagine, então as seguintes situações e tente identificar se o experimento foi “bem executado”:

### Situação 1:

#### “Material disponível”

- 6 aparelhos celulares de mesmo modelo;
- 1 testador (pessoa para executar o programa);



## “Forma como o experimento foi executado”

Os celulares foram numerados de 1 a 6 e, de uma urna contendo seis bolas também assim numeradas, foram sorteados, um a um, os números dos celulares que receberiam o software com a linguagem A; os que receberiam o da linguagem B e, logo depois, os da C. Em seguida, a ordem com que os celulares serão usados pelo testador foi aleatorizada.

Perguntas: U.E.?  
Houve réplica?  
O experimento foi bem executado?

## Situação 2:

### “Material disponível”

- 6 aparelhos celulares de mesmo modelo;
- 1 testador (pessoa para executar o programa);

### “Forma como o experimento foi executado”

Os aparelhos foram separados em três grupos de dois celulares cada. Em seguida, sorteou-se o grupo que receberia o software com a linguagem A, o que receberia o da linguagem B e o da C. O testador escolheu a ordem dos celulares que ia utilizar ao acaso.

Perguntas: U.E.?  
Houve réplica?  
O experimento foi bem executado?

### **Situação 3:**

#### **“Material disponível”**

- 6 aparelhos celulares de mesmo modelo;
- 2 testadores;

## “Forma como o experimento foi executado”

Os aparelhos foram separados em três grupos de dois celulares cada. Em seguida, sorteou-se o grupo que receberia o software com a linguagem A, o que receberia o da linguagem B e o da C. O testador escolheu a ordem dos celulares que ía utilizar ao acaso.

Perguntas: U.E.?  
Houve réplica?  
O experimento foi bem executado?

## Situação 4:

### “Material disponível”

- 6 aparelhos celulares de mesmo modelo;
- 2 testadores;



*Como o experimento deve ser executado?*

## Descrição Formal do Plano Experimental Aleatorizado em Blocos

Suponha que  $N = tb$  unidades experimentais estejam disponíveis para um experimento com o objetivo de comparar os efeitos de  $t$  tratamentos. Suponha ainda que seja possível identificar, entre elas,  $b$  conjuntos de  $t$  unidades experimentais cada, tais que:

- i) Dentro de cada conjunto, as unidades sejam o mais homogêneas possível;
- ii) Exista heterogeneidade entre unidades experimentais de conjuntos distintos, atribuível a um fator de ruído qualitativo.

Nesse cenário, cada conjunto de unidades experimentais é chamado de *bloco* e as unidades experimentais pertencentes a cada bloco tem em comum o mesmo nível do fator de ruído.

Dentro de cada bloco, os  $t$  tratamentos são designados aleatoriamente às  $t$  unidades experimentais, cada tratamento sendo observado uma única vez em cada bloco.

O termo “aleatoriamente” significa “através de um processo físico de aleatorização que garanta igual probabilidade de ocorrência para cada par tratamento-unidade experimental”.

Os processos de aleatorização são empregados independentemente em cada bloco.

## O Modelo Linear Associado ao Plano

$$y_{ik} = \mu + \beta_i + \tau_k + \varepsilon_{ik} \quad (1)$$

No modelo acima, o termo aleatório de resíduo  $\varepsilon_{ik}$  tem como propriedades (parcialmente deduzidas e parcialmente assumidas):

$$E(\varepsilon_{ik}) = 0;$$

$$Var(\varepsilon_{ik}) = \sigma^2;$$

$$Cov(\varepsilon_{ik}, \varepsilon_{i'k}) = 0, \text{ para } i \neq i';$$

$$Cov(\varepsilon_{ik}, \varepsilon_{ik'}) = 0, \text{ para } k \neq k';$$



## Análise de Variância

Identidade Algébrica Associada ao Experimento:

$$y_{ik} = \bar{y}_{..} + (\bar{y}_{i.} - \bar{y}_{..}) + (\bar{y}_{.k} - \bar{y}_{..}) \\ + (y_{ik} - \bar{y}_{i.} - \bar{y}_{.k} + \bar{y}_{..})$$

Tem-se então

$$(y_{ik} - \bar{y}_{..}) = (\bar{y}_{i.} - \bar{y}_{..}) + (\bar{y}_{.k} - \bar{y}_{..}) \\ + (y_{ij} - \bar{y}_{i.} - \bar{y}_{.k} + \bar{y}_{..})$$

$$\sum_{ik} (y_{ik} - \bar{y}_{..})^2 = t \sum_i (\bar{y}_{i.} - \bar{y}_{..})^2 + b \sum_k (\bar{y}_{.k} - \bar{y}_{..})^2 + \sum_{ik} (y_{ik} - \bar{y}_{i.} - \bar{y}_{.k} + \bar{y}_{..})^2$$

$$\text{SQTotal} = \text{SQBloco} + \text{SQTratamento} + \text{SQResíduo}$$

## ANOVA

### Plano Aleatorizado em Blocos

Fonte	G.L.	S.Q.	E(QM)
<i>Bloco</i>	$b-1$	$t \sum_i (\bar{y}_{i.} - \bar{y}_{..})^2$	$g(\sigma^2) + \frac{t}{b-1} \sum_{i=1}^t \beta_i^2$
<i>Trat.</i>	$t-1$	$b \sum_k (\bar{y}_{.k} - \bar{y}_{..})^2$	$\sigma^2 + \frac{b}{t-1} \sum_{k=1}^t \tau_k^2$
<i>Resíduo</i>	$(t-1)(b-1)$	$\sum_{ik} (y_{ik} - \bar{y}_{i.} - \bar{y}_{.k})^2$	$\sigma^2$
<i>Total</i>	$tb-1$	$\sum_{ik} (y_{ik} - \bar{y}_{..})^2$	

## 4. Planos Experimentais em Quadrados Latinos

Os planos experimentais aleatorizados em blocos, como visto, permitem controlar uma fonte de variação sistemática atribuível a um fator de ruído. Imagine, no entanto, que seja necessário controlar fontes de variação provenientes de dois fatores de ruído. Por exemplo, nos experimentos considerados na seção anterior, imagine que estivessem disponíveis:

- Três aparelhos: um do modelo M1, outro do modelo M2 e um terceiro, do modelo M3; e
- Três testadores;

O que fazer nessa situação, se acreditarmos que tanto testador quanto modelo de aparelho sejam fatores de ruído?

**Usar blocos?** Seria possível. Porém, nesse caso, cada combinação de modelo e testador seria um bloco: **ao todo seriam  $3 \times 3 = 9$  blocos!** Como são 3 os tratamentos a serem aplicados em cada bloco, **precisaríamos de no mínimo  $3 \times 9 = 27$  aparelhos celulares!**

Exitem situações, porém, em que é impossível conceber fisicamente a formação dos blocos enquanto cruzamento de níveis de dois fatores de ruído.

- ❏ Os planos experimentais em quadrados latinos foram desenvolvidos para lidar com essas situações. Através deles, é possível controlar duas fontes de variação sistemáticas simultaneamente, de forma mais eficiente que os planos aleatorizados em blocos (no sentido de exigir uma quantidade menor de unidades experimentais), ou mesmo em situações nas quais o uso de blocos seria impossível.

No exemplo do experimento para comparar as linguagens de programação, representando os tratamentos pelas letras A, B e C, o plano em quadrados latinos seria aplicado usando 9 aparelhos celulares, como sugere o esquema do Quadro 1 mostrado a seguir.

### Quadro 1.

Arranjo dos tratamentos “A”, “B” e “C” do experimento, em um quadrado latino.

Testadores (Experiência)	Modelos de aparelhos		
	M1	M2	M3
T1	A	B	C
T2	C	A	B
T3	B	C	A

 NOTE:

Observando o Quadro 1, percebe-se que cada tratamento foi aplicado em um único aparelho celular de cada modelo. Simultaneamente, cada tratamento foi testado uma única vez por cada testador. Dizendo de outra forma, cada tratamento aparece uma única vez em cada linha e em cada coluna, simultaneamente.

## Descrição Formal do Plano Experimental em Quadrado Latino

O plano experimental em quadrado latino é usado para comparar  $t$  tratamentos, controlando  $t$  níveis de um fator de ruído “linha” e  $t$  níveis de um fator de ruído “coluna”.

Para a aplicação do plano, são necessárias  $t \times t$  unidades experimentais com características diferenciadas pela presença de um nível de cada fator de ruído considerado.



Para o caso em que  $t = 3$ , por exemplo, são necessárias 9 unidades. Denotando por  $u_{ij}$  a u.e. com nível  $i$  do fator “linha” e nível  $j$  do fator “coluna”, e dispondo as unidades experimentais em um arranjo quadricular, elas seriam representadas por:

$$\begin{array}{l} u_{11} , u_{12} , u_{13} \\ u_{21} , u_{22} , u_{23} \\ u_{31} , u_{32} , u_{33} \end{array} \quad (A1)$$

Do ponto de vista matemático, um quadrado latino de ordem  $t$  é um arranjo de  $t$  letras latinas em um quadrado com  $t$  linhas e  $t$  colunas, de tal forma que cada letra aparece apenas uma vez em cada linha e em cada coluna. No contexto de um plano

experimental, cada letra representa um tratamento. Superpondo um quadrado latino ao arranjo de unidades experimentais disponíveis, tem-se a designação de cada tratamento a cada unidade experimental. O Quadro 1 é um exemplo de superposição de um quadrado latino de ordem 3,

$$\begin{array}{ccc} A & B & C \\ C & A & B \\ B & C & A \end{array}$$

no arranjo de unidades experimentais (A1), da página anterior.

## O Processo de Aleatorização

O processo de aleatorização de um experimento em quadrado latino foi proposto originalmente por Yates (1933). A sua execução, para quadrados de ordem 3, 4 e 5, é dada por etapas, da seguinte forma:

Passo 1: Selecione aleatoriamente um dentre todos os quadrados latinos reduzidos existentes;

Passo 2: Proceda à permutação aleatória de todas as linhas do quadrado selecionado no passo 1, exceto a primeira;

Passo 3: Proceda à permutação aleatória de todas as colunas;

Passo 4: Atribua aleatoriamente os tratamentos às letras utilizadas (A, B, C...).

Uma forma equivalente de aleatorização seria, no passo 2, proceder à permutação de todas as linhas e, no passo 3, proceder à permutação de todas as colunas, exceto a primeira.

Nos casos de quadrados latinos de ordem superior a 5, o processo é o mesmo, permutano-se todas as linhas e colunas. Para quadrados latinos de ordens superiores a 6, considera-se suficiente escolher qualquer quadrado antes de proceder os passos seguintes.

## O Modelo Linear Associado ao Plano

$$y_{ijk} = \mu + \rho_i + \gamma_j + \tau_k + \varepsilon_{ijk}$$

onde:

$y_{ijk}$  é o valor observado do tratamento  $k$  aplicado sob o  $j$ -ésimo nível do fator coluna e  $i$ -ésimo nível do fator linha;

$\rho_i$  é o efeito (fixo) do  $i$ -ésimo nível do fator linha;

$\gamma_j$  é o efeito (fixo) do  $j$ -ésimo nível do fator coluna;

$\tau_k$  é o efeito (fixo) do tratamento  $k$ ;

$\varepsilon_{ijk}$  é o termo aleatório de resíduo,  
cujas propriedades  
(parcialmente deduzidas) são:

$$E(\varepsilon_{ijk}) = 0;$$

$$Var(\varepsilon_{ijk}) = \sigma^2;$$

$$Cov(\varepsilon_{ijk}, \varepsilon_{ij'k'}) = 0;$$

$$Cov(\varepsilon_{ijk}, \varepsilon_{i'jk'}) = 0;$$

$$Cov(\varepsilon_{ijk}, \varepsilon_{i'j'k}) = 0;$$

$$Cov(\varepsilon_{ijk}, \varepsilon_{i'j'k'}) = 0.$$

# Análise de Variância

## ANOVA

### *Plano Experimental em Quadrado Latino de Ordem $t$*

F.V.	G.L.	E(Q.M.)
<i>Linhas</i>	$t-1$	$g(\sigma^2) + \frac{t}{t-1} \sum_i \rho_i^2$
<i>Colunas</i>	$t-1$	$g(\sigma^2) + \frac{t}{t-1} \sum_j \gamma_j^2$
<i>Tratamentos</i>	$t-1$	$\sigma^2 + \frac{t}{t-1} \sum_k \tau_k^2$
<i>Resíduo</i>	$(t-1)(t-2)$	$\sigma^2$
<i>Total</i>	$t^2-1$	

O plano em quadrado latino tem duas limitações:

- O número de tratamentos, níveis do fator “linha” e do fator “coluna” tem que ser os mesmos;
- Quadrados latinos de baixa ordem disponibilizam poucos graus de liberdade para o resíduo (ver quadro abaixo).

$t$	G.L. Resíduo
2	0
3	2
4	6
5	12
6	20



## Replicações em Quadrados Latinos

Para aumentar os graus de liberdade do resíduo é desejável que o experimento seja replicado um certo número de vezes. Essas replicações implicam na existência de mais de um quadrado latino, onde a aleatorização é feita, independentemente, em cada quadrado. O aumento no número de graus de liberdade do resíduo, porém, varia de acordo com a situação experimental encontrada. Para motivar a discussão sobre o tema, o seguinte experimento é apresentado:

Três tipos diferentes de linguagens de programação (A, B, C) devem ser comparadas com respeito ao tempo para executar um programa. Para isso, estão disponíveis três testadores, bem como três

modelos de aparelhos. A fim de controlar variações existentes entre os testadores e entre os modelos de aparelhos, um plano experimental em quadrado latino deve ser empregado. O arranjo abaixo sugere a alocação dos tratamentos às unidades experimentais, após a aleatorização. Os modelos são denotados por M1, M2 e M3 e os testadores, por T1, T2 e T3.

	T1	T2	T3
M1	B	C	A
M2	C	A	B
M3	A	B	C

A replicação do experimento pode ocorrer de diversas formas, dependendo da disponibilidade do material experimental:

**Cenário 1:**

*Sem disponibilidade de testadores ou modelos extras.*

Nessa situação, os mesmos testadores e modelos de aparelhos são aproveitados.

O experimento é executado novamente em outra ocasião, como sugere a ilustração a seguir.

	T1	T2	T3
M1	B	C	A
M2	C	A	B
M3	A	B	C

☞ Réplica 1

Réplica 2 ☞

	T1	T2	T3
M1	C	B	A
M2	B	A	C
M3	A	C	B

## **Cenário 2:**

*Com disponibilidade de testadores ou modelos extras.*

Suponha que testadores extras estejam disponíveis e que os mesmos modelos de aparelhos sejam utilizados (ou vice-versa).

O experimento é executado novamente em outra ocasião, com outros testadores, como sugere a ilustração a seguir.

	T1	T2	T3
M1	B	C	A
M2	C	A	B
M3	A	B	C

Réplica 1

	T4	T5	T6
M1	A	B	C
M2	C	A	B
M3	B	C	A

Réplica 2

### Cenário 3:

*Com disponibilidade de testadores e modelos extras.*

Suponha que testadores e modelos extras de aparelhos estejam disponíveis. O experimento é executado como sugere a ilustração a seguir.

	C1	C2	C3	C4	C5	C6
M1	B	C	A			
M2	C	A	B			
M3	A	B	C			
M4				A	B	C
M5				B	C	A
M6				C	A	B
	Réplica 1			Réplica 2		

## 5. Discussão Geral sobre Potencialidades de Aplicação em Engenharia de Software

Algumas questões:

- Quais as variáveis respostas típicas?
- Em que consistem tipicamente os tratamentos? Há espaço para o uso de estruturas fatoriais?
- Qual o cenário de disponibilidade de material experimental mais típico?