

Universidade Federal de Pernambuco
Centro de Ciências Exatas e da Natureza
Departamento de Estatística

Planos Experimentais Para Realização de Testes com Softwares

Roberto Ferreira Manghi

Recife, Dezembro de 2008

Sumário

1	Agradecimentos	2
2	Apresentação	3
3	Introdução	4
4	Planos Experimentais do Tipo Quadrado Latino	5
4.1	Metodologia	5
4.2	Realização do Experimento	9
4.3	Conclusões	13
5	Planos Experimentais do Tipo Split-Plot	14
5.1	Metodologia	14
5.2	Realização do Experimento	15
5.3	Conclusões	18
5.4	Comparação Entre Planos Experimentais	19
5.4.1	Plano Experimental Completamente Aleatorizado com Estrutura Fatorial Completa 2^4	19
5.4.2	Plano Experimental Completamente Aleatorizado com Estrutura Fatorial Fracionada $\frac{1}{2}2^4$ (Res. IV)	21
5.4.3	Plano Experimental Strip-Block em Dois Estágios (Completamente Aleatorizado, Completamente Aleatorizado)	22
5.4.4	Vantagens e Desvantagens dos Planos Experimentais Propostos	24
6	Conclusões Gerais	26
6.1	Conclusões Acerca dos Estudos Realizados	26
6.2	Conclusões Acerca da Experiência Adquirida	26
7	Referências	27
8	Anexo	28
8.1	Função que executa o Teste de Não-Aditividade de Tukey para experimentos do tipo Quadrado Latino com Replicação	28

1 Agradecimentos

À minha família, pelo incentivo e apoio durante todo o Curso de Estatística;

À Paula Tiyemi Shinozaki Mendes, pelo amor e dedicação constante;

Ao Professor Cristiano Ferraz, determinante na inclusão ao referente estágio;

Aos pesquisadores Eduardo Aranha e Rodrigo Bonifácio, pelo suporte, interação e pesquisa conjunta acerca dos temas aqui estudados;

Ao CIn – Centro de Informática da Universidade Federal de Pernambuco, e mais precisamente ao Centro de Testes da Motorola.

2 Apresentação

O presente relatório trata da conclusão de Estágio Supervisionado Obrigatório do aluno **Roberto Ferreira Manghi**, integrante do curso de Bacharelado em Estatística pela UFPE - Universidade Federal de Pernambuco. O estágio foi realizado no prédio do C.E.S.A.R. - Centro de Estudos e Sistemas Avançados do Recife, mais especificamente no Centro de Testes da Motorola, no período de agosto a novembro de 2008. As análises computacionais foram executadas através do software **R**, e a edição do presente relatório foi realizada através do software \LaTeX , na linguagem \TeX .

3 Introdução

Uma das importantes etapas na Engenharia de Software é a realização de testes que envolvem seus produtos. Estes testes têm por finalidade, dentre outras, a detecção de eventuais falhas nos programas computacionais elaborados, que uma vez detectadas e entendidas podem ser solucionadas mais facilmente. Métodos estatísticos são de grande valia neste processo, e neste caso específico envolvem o planejamento, a execução e a análise de experimentos. Através de uma metodologia bem definida, estes métodos permitem inferir sobre diversas características do produto. Tais inferências, por sua vez, podem ser determinantes nas conclusões a respeito da qualidade de um determinado software, dados os diversos fatores que venham a contribuir com este conceito neste cenário específico.

O primeiro experimento tinha como objetivo comparar duas técnicas de gerência de variação em requisitos de linhas de produto de software (LPS): **SVCM** e **PLUSS**. Diferentes variáveis de interesse foram investigadas, como o tempo necessário de evolução e o impacto das mudanças. Usou-se, então, um plano experimental do tipo Quadrado Latino.

O segundo experimento teve como objetivo detectar se determinados fatores são significativos no esforço necessário para a execução de testes manuais com celulares, e consequentemente no tempo gasto para a realização destes testes. Para este estudo foi utilizado um plano experimental do tipo Split-Plot.

Dadas as análises dos experimentos, foi feita também uma comparação entre planos experimentais dentro da realidade deste projeto. O objetivo, neste caso, foi o de encontrar planos experimentais que fossem mais viáveis na execução dos testes do segundo experimento considerado, no sentido de otimizar características como custo, complexidade, disponibilidade de material para execução, etc.

4 Planos Experimentais do Tipo Quadrado Latino

4.1 Metodologia

Os planos experimentais envolvendo quadrados latinos são utilizados quando deseja-se controlar duas fontes conhecidas de variabilidade, ou seja, permite-se utilizar o princípio do controle local sistematicamente em duas direções (Montgomery, 2001, pág. 144). Existem exemplos de experimentos com esta metodologia em diversas áreas, como na agricultura (Snedecor e Cochran, 1967, pág. 313) e setores industriais (Cochran e Cox, 1957, pág.119).

Em cada quadrado latino formado para um dado experimento, temos as linhas representando um variável de controle e as colunas representando outra variável também de controle, podendo ser chamadas também de variáveis de controle local. Deste modo, as linhas e as colunas de cada quadrado representam duas restrições na aleatorização dos tratamentos às unidades experimentais. Os níveis destas variáveis são dispostos, então, nas respectivas linhas e colunas de cada quadrado, não necessariamente via processo de aleatorização. Já a aleatorização dos tratamentos é feita dentro de cada quadrado, de forma que cada linha e cada coluna contenha exatamente um de cada tratamento considerado no experimento. Desta maneira as linhas e as colunas podem ser consideradas como ortogonais em relação aos tratamentos, ou seja, assume-se que não há efeito de interação entre a variável disposta nas linhas e a variável disposta nas colunas dos quadrados, com relação às respostas obtidas pelos tratamentos investigados no experimento. Considere o exemplo abaixo:

		COLUNAS		
		I	II	III
LINHAS	I	C	B	A
	II	B	A	C
	III	A	C	B

Figura 1: Exemplo de um quadrado latino de dimensão três.

Neste caso, teríamos três tratamentos, denotados pelas letras **A**, **B** e **C**. Pode-se notar que em cada linha e em cada coluna existe uma e somente uma de cada letra referente a cada um dos tratamentos. O modelo de efeitos utilizado na análise de um experimento segundo um plano desta natureza, envolvendo p tratamentos e apenas um quadrado é dado por:

$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + \epsilon_{ijk}, \quad i, j, k = 1, 2, \dots, p, \text{ onde:}$$

y_{ijk} – variável de interesse obtida na k -ésima coluna, j -ésimo tratamento e i -ésima linha;

μ – média geral;

α_i – efeito da i -ésima linha;

β_j – efeito do j-ésimo tratamento;

γ_k – efeito da k-ésima coluna;

ϵ_{ijk} – erro aleatório.

Temos ainda a seguinte suposição:

$$\epsilon_{ijk} \sim N(0, \sigma^2),$$

e as restrições

$$\sum_{i=1}^p \alpha_i = 0, \quad \sum_{j=1}^p \beta_j = 0, \quad \sum_{k=1}^p \gamma_k = 0.$$

Ao adotar este modelo, supõe-se que não há interação entre os fatores relativos às linhas, colunas e tratamentos. Se usarmos replicações, (mais de um quadrado no experimento) podemos proceder de quatro formas:

Tipo I – Utilizando os mesmos níveis das variáveis referentes às linhas e às colunas em todas as replicações;

Tipo II – Mantendo os mesmos níveis da variável referente às linhas e alterando os níveis da variável respectiva às colunas ao considerar outra replicação;

Tipo III – Mantendo os mesmos níveis da variável referente às colunas e alterando os níveis da variável respectiva às linhas ao considerar outra replicação;

Tipo IV – Alterando os níveis das variáveis referentes às linhas e colunas, com relação a cada replicação.

Os modelos para cada situação, assumindo que n replicações são executadas, bem como suas respectivas análises de variância, são ilustrados logo a seguir:

Modelo com replicação do tipo I:

$$y_{lijk} = \mu + \tau_l + \alpha_i + \beta_j + \gamma_k + \epsilon_{lijk}, \quad l = 1, 2, \dots, n, \quad i, j, k = 1, 2, \dots, p, \text{ onde:}$$

y_{lijk} – variável de interesse obtida na l-ésima replicação, k-ésima coluna, j-ésimo tratamento e i-ésima linha;

μ – média geral;

τ_l – efeito da l-ésima replicação;

α_i – efeito da i-ésima linha;

β_j – efeito do j-ésimo tratamento;

γ_k – efeito da k-ésima coluna;

ϵ_{lijk} – erro aleatório,

onde

$$\epsilon_{lijk} \sim N(0, \sigma^2), \quad \sum_{l=1}^n \tau_l = 0, \quad \sum_{i=1}^p \alpha_i = 0, \quad \sum_{j=1}^p \beta_j = 0, \quad \sum_{k=1}^p \gamma_k = 0.$$

Fontes de variação	Graus de liberdade
Replicação	$n - 1$
Linha	$p - 1$
Coluna	$p - 1$
Tratamento	$p - 1$
Erro	$(p - 1)(n(p + 1) - 3)$
Total	$np^2 - 1$

Tabela 1: Análise de variância - experimento do tipo Quadrado Latino com n replicações (Tipo I).

Modelo com replicação do tipo II:

$$y_{lijk} = \mu + \tau_l + \alpha_i + \beta_j + \tau\gamma_{lk} + \epsilon_{lijk}, \quad l = 1, 2, \dots, n, \quad i, j, k = 1, 2, \dots, p, \text{ onde:}$$

y_{lijk} – variável de interesse obtida na k -ésima coluna, j -ésimo tratamento e i -ésima linha;

μ – média geral;

τ_l – efeito da l -ésima replicação;

α_i – efeito da i -ésima linha;

β_j – efeito do j -ésimo tratamento;

$\tau\gamma_{lk}$ – efeito de interação entre a l -ésima replicação e a k -ésima coluna;

ϵ_{lijk} – erro aleatório,

onde

$$\epsilon_{lijk} \sim N(0, \sigma^2), \quad \sum_{l=1}^n \tau_l = 0, \quad \sum_{i=1}^p \alpha_i = 0, \quad \sum_{j=1}^p \beta_j = 0,$$

$$\sum_{k=1}^p \tau\gamma_{lk} = 0 \text{ para } l \text{ fixo e } \sum_{l=1}^n \tau\gamma_{lk} = 0 \text{ para } k \text{ fixo.}$$

Fontes de variação	Graus de liberdade
Replicação	$n - 1$
Linha	$p - 1$
Coluna	$n(p - 1)$
Tratamento	$p - 1$
Erro	$(p - 1)(np - 1)$
Total	$np^2 - 1$

Tabela 2: Análise de variância - experimento do tipo Quadrado Latino com n replicações (Tipo II).

Modelo com replicação do tipo III:

$$y_{lijk} = \mu + \tau_l + \tau\alpha_{li} + \beta_j + \gamma_k + \epsilon_{lijk}, \quad l = 1, 2, \dots, n, \quad i, j, k = 1, 2, \dots, p, \text{ onde:}$$

y_{lijk} – variável de interesse obtida na l -ésima replicação, k -ésima coluna, j -ésimo tratamento e i -ésima linha;

μ – média geral;

τ_l – efeito da l -ésima replicação;

$\tau\alpha_{li}$ – efeito de interação entre a l -ésima replicação e a i -ésima linha;

β_j – efeito do j -ésimo tratamento;

γ_k – efeito da k -ésima coluna;

ϵ_{lijk} – erro aleatório,

onde

$$\epsilon_{lijk} \sim N(0, \sigma^2), \quad \sum_{i=1}^p \tau\alpha_{li} = 0 \text{ para } l \text{ fixo, } \sum_{l=1}^n \tau\alpha_{li} = 0 \text{ para } i \text{ fixo,}$$

$$\sum_{j=1}^p \beta_j = 0 \text{ e } \sum_{k=1}^p \gamma_k = 0.$$

Fontes de variação	Graus de liberdade
Replicação	$n - 1$
Linha	$n(p - 1)$
Coluna	$p - 1$
Tratamento	$p - 1$
Erro	$(p - 1)(np - 1)$
Total	$np^2 - 1$

Tabela 3: Análise de variância - experimento do tipo Quadrado Latino com n replicações (Tipo III).

Modelo com replicação do tipo IV:

$$y_{lijk} = \mu + \tau_l + \tau\alpha_{li} + \beta_j + \tau\gamma_{lk} + \epsilon_{lijk}, \quad l = 1, 2, \dots, n, \quad i, j, k = 1, 2, \dots, p, \text{ onde:}$$

y_{lijk} – variável de interesse obtida na l -ésima replicação, k -ésima coluna, j -ésimo tratamento e i -ésima linha;

μ – média geral;

τ_l – efeito da l -ésima replicação;

$\tau\alpha_{li}$ – efeito de interação entre a l -ésima replicação e a i -ésima linha;

β_j – efeito do j -ésimo tratamento;

$\tau\gamma_{lk}$ – efeito de interação entre a l-ésima replicação e a k-ésima coluna;

ϵ_{lijk} – erro aleatório,

onde

$$\epsilon_{lijk} \sim N(0, \sigma^2), \quad \sum_{i=1}^p \tau\alpha_{li} = 0 \text{ para } l \text{ fixo}, \quad \sum_{l=1}^n \tau\alpha_{li} = 0 \text{ para } i \text{ fixo},$$

$$\sum_{j=1}^p \beta_j = 0, \quad \sum_{k=1}^p \tau\gamma_{lk} = 0 \text{ para } l \text{ fixo e } \sum_{l=1}^n \tau\gamma_{lk} = 0 \text{ para } k \text{ fixo}.$$

Fontes de variação	Graus de liberdade
Replicação	$n - 1$
Linha	$n(p - 1)$
Coluna	$n(p - 1)$
Tratamento	$p - 1$
Erro	$(p - 1)(n(p - 1) - 1)$
Total	$np^2 - 1$

Tabela 4: Análise de variância - experimento do tipo Quadrado Latino com n replicações (Tipo IV).

4.2 Realização do Experimento

O experimento consistiu primeiramente na definição de 6 quadrados latinos de dimensão dois. As variáveis de controle local em cada quadrado foram **Aluno** e **Domínio**. Participaram do experimento 16 alunos do CIn – Centro de Informática da Universidade Federal de Pernambuco, para a execução dos procedimentos que utilizaram as técnicas em estudo. Definidos os quadrados, foram aleatorizados os tratamentos para cada um dos dois alunos em cada replicação, segundo os domínios pré-determinados, de forma que, considerado um quadrado específico, a alocação de uma determinada técnica a um determinado aluno segundo um domínio, determinou as outras três alocações de tratamentos naquele quadrado, de forma que, sob um dado domínio, cada aluno de cada quadrado só realizou a tarefa através de uma das duas técnicas, e sob cada domínio as duas técnicas foram utilizadas. O tempo de execução de cada tarefa por cada aluno das seis replicações (quadrados), segundo os respectivos tratamentos alocados foi então registrado. Temos, portanto, as observações da variável **Resposta**:

Replicação		Domínio		Replicação		Domínio	
	Aluno	eCommerce	cyber		Aluno	eCommerce	cyber
1	1	55(a)	83(b)	4	1	61(a)	46(b)
	2	50(a)	55(b)		2	69(a)	40(b)
2	1	71(a)	97(b)	5	1	55(a)	40(b)
	2	80(a)	130(b)		2	70(a)	60(b)
3	1	60(a)	76(b)	6	1	63(a)	42(b)
	2	82(a)	57(b)		2	60(a)	60(b)

Tabela 5: Observações do primeiro experimento (Quadrados Latinos): técnicas (a) SVCVM e (b) PLUS.

Para realizarmos a análise computacional no software **R**, começamos com os comandos

```
dados.ql<-read.table("C:/Estat/Estagio1/Exp1/dados.txt",header=TRUE)
attach(dados.ql)
```

tornando os dados do experimento disponíveis para uso. A seguir, definimos os fatores,

```
Replicacao<-factor(Replicacao)
Aluno<-factor(Aluno)
Dominio<-factor(Dominio)
Tecnica<-factor(Tecnica)
```

e posteriormente geramos gráficos de box-plot da variável resposta com relação a cada um dos dois tratamentos considerados no experimento, através do comando

```
plot(Resposta~Tecnica,col="gray",xlab="Técnica",ylab="Tempo")
```

Segue, então o gráfico:

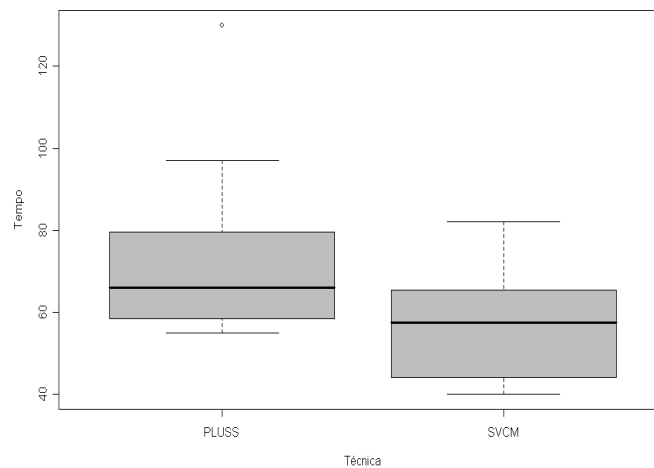


Figura 2: Box-plots dos tempos de execução dos testes segundo cada técnica utilizada.

Pode-se notar que as medianas das observações segundo cada técnica apresentam valores diferentes. O valor mediano dos tempos segundo a técnica **PLUSS** é maior que o valor mediano dos dados referentes à técnica **SVCM**. Existe, ainda, um *outlier* no conjunto de observações referente à primeira técnica citada. Através do comando

```
anova.q1<-aov(Resposta~Replicacao+Aluno:Replicacao+Dominio+Tecnica)
```

ajustou-se um modelo de efeitos (Tipo III) para a posterior análise do experimento. Note que o fator **Aluno** está associado ao fator **Replicação**, pois para cada replicação foi utilizado um par diferente de alunos. Antes de visualizar o ajuste, foram usados os comandos

```
library(MASS)
boxcox(anova.q1,lambda = seq(-3, 5, 1/10))
```

para verificar se a suposição de variância constante dos erros foi atendida. O método de Box-Cox consiste em encontrar possíveis potências para a variável resposta do experimento, tais que estabilizem a variância desta mesma. Para tal, eleva-se a variável resposta a várias potências num dado intervalo, e calcula-se a log-verossimilhança das observações elevadas a tais potências. A potência mais recomendável é aquela que maximiza a função de log-verossimilhança. Foi gerado, então, o seguinte gráfico:

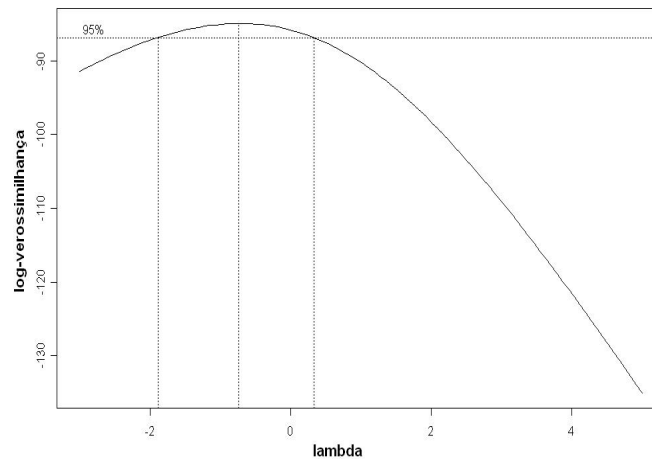


Figura 3: Gráfico de Box-Cox para transformação sugerida da variável resposta.

Segundo o gráfico, dado um nível de confiança de 95%, temos uma provável transformação da variável *Resposta* dada por $Resposta^{-1}$. Assim, temos o modelo novamente ajustado já com esta transformação:

```
anova2.q1<-aov(Resposta^(-1)~Replicacao+Aluno:Replicacao+Dominio+Tecnica)
```

e novamente, a construção do gráfico de Box-Cox

```
boxcox(anova2.q1,lambda = seq(-3, 5, 1/10))
```

gerando o seguinte gráfico:

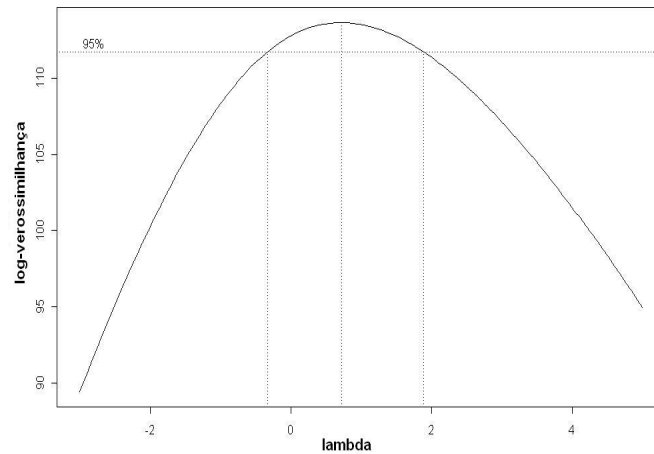


Figura 4: Novo gráfico de Box-Cox a partir do segundo modelo ajustado.

Desta vez, um nível de confiança de 95%, temos que o valor 1 está contido no intervalo de possíveis potências às quais deve ser submetida a variável resposta, ou seja, a variância neste caso já está estabilizada. A suposição de que o modelo está adequado, ou seja, que o modelo é aditivo (a menos da interação entre os fatores Replicação e Aluno) foi feita através do Teste de Não-Aditividade de Tukey, através do programa disponível na seção Anexos deste relatório:

```
TukeyNADD.QL.REP(anova2.q1)
```

As hipóteses deste teste são:

- H_0 : o modelo é aditivo.
- H_1 : H_0 é falsa.

Obteve-se o valor p de 0.29, o que não nos dá evidências suficientes para rejeitarmos a hipótese nula, ou seja, considera-se que o modelo é adequado na modelagem dos dados do experimento. Através do comando

```
plot(anova2.q1)
```

geramos o gráfico quantil-quantil dos resíduos gerados pelo ajuste do modelo. Temos então este gráfico:

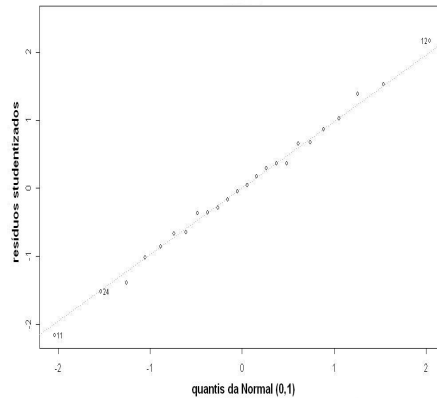


Figura 5: Quantil-Quantil Normal dos resíduos studentizados do primeiro experimento (Quadrados Latinos).

Os pontos do gráfico estão dispostos ao longo da linha pontilhada, indicando que podemos considerar os resíduos seguindo uma distribuição Normal. Já a análise de variância é obtida pelo comando

```
summary(anova2.q1)
```

e descrita logo a seguir:

Fontes de variação	GL	SQ	QM	Estatística F	valor p
Replicação	5	0.00018901	0.000037802	4.8472	0.016431
Domínio	1	0.000013944	0.000013944	1.7880	0.210785
Técnica	1	0.000094738	0.000094738	12.1479	0.005868
Replicação x Aluno	6	0.000068153	0.000011359	1.4565	0.285374
Erro	10	0.000077987	0.0000077990	—	—
Total	23	0.000443832	—	—	—

Tabela 6: Análise de variância do primeiro experimento (Quadrados Latinos).

A estatística do teste de hipóteses realizado para verificar se os efeitos das duas técnicas são iguais é dada pelo quociente entre o quadrado médio do tratamento (Técnica) e o quadrado médio do Erro. Sob a hipótese de que os efeitos dos tratamentos são iguais, este quociente segue uma distribuição F, com um grau de liberdade no numerador e dez graus de liberdade no denominador. A estatística é então comparada com esta distribuição. Assim, o valor p referente a este teste indica que existem fortes evidências de diferença significativa entre a técnicas **PLUSS** e **SVCM**.

4.3 Conclusões

A partir da análise do experimento, conclue-se que as técnicas **PLUSS** e **SVCM** apresentam efeitos significativamente diferentes no tempo de execução da determinada tarefa computacional. Neste caso, a técnica **SVCM** proporciona tempos de execução da tarefa, em média, menores que os tempos de execução da tarefa ao usar a técnica **PLUSS**.

5 Planos Experimentais do Tipo Split-Plot

5.1 Metodologia

Segundo Montgomery (2001, pág. 573), os planos experimentais do tipo Split-Plot podem ser considerados como uma generalização dos planos experimentais fatoriais. Em geral, este plano contém mais de um estágio, onde em cada estágio é realizado um processo de aleatorização, envolvendo tratamentos formados por apenas uma parte do conjunto total de fatores. Ao final de todos os estágios envolvidos no processo, todos os tratamentos são aleatorizados às unidades experimentais envolvidas no experimento. As áreas industriais e agrícolas também podem ser consideradas como campos nos quais este plano pode ser aplicado (Snedecor e Cochran, 1967, pág. 369).

A idéia deste plano é determinar unidades experimentais maiores (*whole-plots*) e dentro de cada uma delas, unidades experimentais menores (*sub-plots*). A aleatorização dos tratamentos considerados no primeiro estágio se dá em relação aos *whole-plots*, e então, dentro de cada *whole-plot*, são aleatorizados tratamentos para os *sub-plots*, uma vez que todos os *whole-plots* já receberam a indicação de seus respectivos tratamentos.

RÉPLICA 1			RÉPLICA 2		
a ₂	a ₁	a ₃	a ₃	a ₁	a ₂
b2	b2	b1	b2	b1	b2
-----	-----	-----	-----	-----	-----
b1	b1	b2	b1	b2	b1

Figura 6: Exemplo de um experimento segundo um plano experimental Split-Plot.

Temos na figura os tratamentos referentes ao primeiro estágio do processo de aleatorização dados por **a1**, **a2** e **a3**. Já os tratamentos referentes ao segundo estágio do processo de aleatorização são representados pelas letras **b1** e **b2**. Este exemplo pode ser visto como a superposição de um Plano Experimental Completamente Aleatorizado no primeiro estágio e um Plano Experimental Aleatorizado em Blocos no segundo estágio. O modelo de efeitos usado nesta estrutura específica de superposição de planos dentro de um plano Split-Plot, considerando **a** tratamentos relativos ao primeiro estágio, **b** tratamentos relativos ao segundo estágio e **n** replicações é o seguinte:

$$y_{ijk} = \mu + \alpha_i + \beta_j + \delta_{ij} + \gamma_k + \beta\gamma_{jk} + \epsilon_{ijk},$$

$$i = 1, 2, \dots, n; \quad j = 1, 2, \dots, a; \quad k = 1, 2, \dots, b, \text{ onde:}$$

y_{ijk} – variável de interesse obtida na i -ésima replicação, j -ésimo tratamento referente ao primeiro estágio e k -ésimo tratamento referente ao segundo estágio;

μ – média geral;

α_i – efeito da i -ésima replicação;

β_j – efeito do j-ésimo tratamento referente ao primeiro estágio;

δ_{ij} – erro aleatório proveniente do primeiro estágio;

γ_k – efeito do k-ésimo tratamento referente ao segundo estágio;

$\beta\gamma_{jk}$ – efeito da interação entre o j-ésimo tratamento referente ao primeiro estágio e o k-ésimo tratamento referente ao segundo estágio;

ϵ_{ijk} – erro aleatório referente ao segundo estágio.

Temos as seguintes suposições:

$$\delta_{ij} \sim N(0, \sigma_1^2), \quad \epsilon_{ijk} \sim N(0, \sigma_2^2),$$

e as seguintes restrições:

$$\sum_{i=1}^n \alpha_i = 0; \quad \sum_{j=1}^a \beta_j = 0; \quad \sum_{k=1}^b \gamma_k = 0.$$

A análise de variância é descrita abaixo:

Fontes de variação	Graus de liberdade
Replicação	$n - 1$
Trat. (estágio 1)	$a - 1$
Erro (estágio 1)	$(n - 1)(a - 1)$
Trat. (estágio 2)	$b - 1$
Trat. (estágio 1) x Trat. (estágio 2)	$(a - 1)(b - 1)$
Erro (estágio 2)	$a(n(b - 1) - b + 1)$
Total	$nab - 1$

Tabela 7: Análise de variância - experimento Split-Plot com n replicações.

5.2 Realização do Experimento

O conjunto de unidades experimentais foi formado por 32 celulares com características semelhantes. Os fatores inerentes ao primeiro estágio do experimento foram dois, referentes ao testador: Experiência em Testes (D) e Nível de Inglês (E). Os fatores inerentes ao segundo estágio foram também dois, referentes ao teste: Experiência no Teste (F) e Tamanho do Teste (G). Cada um dos quatro fatores possuíram dois níveis, denotados por + e -. Temos, então, a disposição dos tratamentos nos dois estágios:

Primeiro Estágio		
Tratamento	Fatores	
	D	E
(1)	–	–
d	+	–
e	–	+
de	+	+
Segundo Estágio		
Tratamento	Fatores	
	F	G
(1)	–	–
f	+	–
g	–	+
fg	+	+

Tabela 8: Disposição dos tratamentos no experimento Split-Plot (primeiro e segundo estágios).

As unidades experimentais referentes ao primeiro estágio (whole-plots) foram formadas por quatro grupos de celulares, cada grupo contendo quatro aparelhos. Os tratamentos referentes ao primeiro estágio foram então aleatorizados a estes quatro grupos, caracterizando um experimento Completamente Aleatorizado. Uma vez alocados estes tratamentos, dentro de cada grupo de quatro celulares foram aleatorizados os tratamentos referentes ao segundo estágio para cada celular (sub-plots) caracterizando um experimento Completamente Aleatorizado em Blocos. Terminados os processos de aleatorização, os quatro testadores executaram os testes nos celulares aos quais eles foram alocados no primeiro estágio, e na ordem de execução dada pela aleatorização do segundo estágio. Os tempos de execução dos testes foram, então, registrados. Temos, assim, estes tempos:

Primeira Replicação				
Teste	Testador			
	(1)	f	g	fg
(1)	221(4)	268(1)	467(2)	409(3)
d	288(3)	196(2)	337(4)	311(1)
e	180(3)	103(1)	528(4)	239(2)
de	181(4)	114(1)	553(2)	301(3)
Segunda Replicação				
Teste	Testador			
	(1)	f	g	fg
(1)	241(3)	626(2)	97(1)	471(4)
d	183(2)	589(4)	126(3)	731(1)
e	240(3)	179(1)	120(4)	436(2)
de	142(2)	299(3)	294(1)	616(4)

Tabela 9: Observações do experimento Split-Plot (os parênteses ao lado das observações indicam a ordem na qual cada testador realizou os testes).

Para realizarmos a análise computacional no software **R**, começamos com os comandos

```
dados.sp<-read.table("C:/Estat/Estagio1/Exp2/dados.txt",header=TRUE)
attach(dados.sp)
```

tornando os dados do experimento disponíveis para uso. A seguir, definimos os fatores,

```
Rep<-factor(Rep)
Est1<-factor(Est1)
Est2<-factor(Est2)
D<-factor(D)
E<-factor(E)
F<-factor(F)
G<-factor(G)
```

e posteriormente geramos gráficos de box-plot da variável resposta com relação aos tratamentos do primeiro estágio e com relação aos tratamentos do segundo estágio considerados no experimento, através dos comandos

```
par(mfrow=c(1,2))
plot(Resposta~Est1,col="gray",xlab="Tratamento (estágio 1)",ylab="Tempo")
plot(Resposta~Est2,col="gray",xlab="Tratamento (estágio 2)",ylab="Tempo")
```

Segue, então o gráfico:

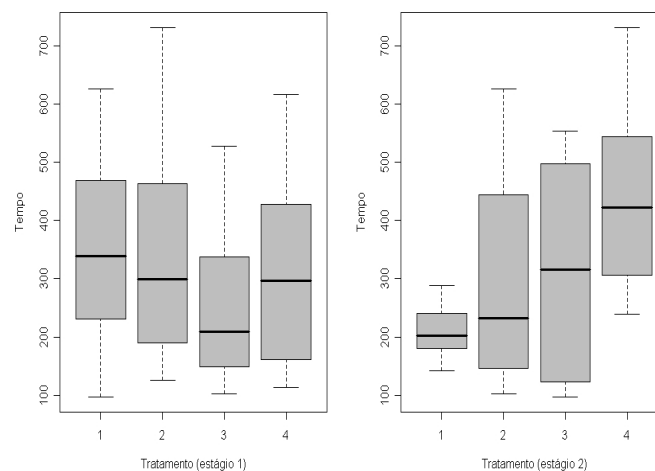


Figura 7: Box-plots dos tempos de execução dos testes segundo os tratamentos dos primeiro e segundo estágios.

Com relação aos tratamentos referentes ao primeiro estágio, os valores obtidos através da combinação dos níveis dos fatores D(-) e E(+) produziram menor tempo mediano, enquanto que, com relação aos tratamentos do segundo estágio, os níveis dos fatores F(-) e G(-) produziram o menor tempo mediano, sugerindo que talvez uma combinação D(-), E(+), F(-) e G(-) produza os menores tempos de execução de testes. Essa hipótese precisaria ainda ser testada, pois as interações entre os fatores poderiam produzir resultados diferentes de uma análise em separado com relação a cada estágio do experimento. A construção da ANOVA no **R** foi realizada da seguinte forma: através dos comandos

```
anova.sp<-aov(Resposta~D*E*F*G+Error(Rep/(D*E)))
summary(anova.sp)
```

ajustou-se o modelo de efeitos para as observações do experimento, obtendo, assim a ANOVA:

Primeiro estágio					
Fontes de variação	GL	SQ	QM	Estatística F	valor p
Replicação	1	15051	15051		
D	1	5941	5941	0.80	0.43
E	1	33540	33540	4.52	0.12
D x E	1	8256	8256	1.11	0.37
Erro	3	22241	7413.67		
Segundo estágio					
Fontes de variação	GL	SQ	QM	Estatística F	valor p
F	1	89253	89253	2.1468	0.1686
G	1	123256	123256	2.9646	0.1108
D x F	1	5408	5408	0.1301	0.7246
E x F	1	79202	79202	1.9050	0.1927
D x G	1	10082	10082	0.2425	0.6313
E x G	1	53792	53792	1.2938	0.2775
F x G	1	2701	2701	0.0650	0.8031
D x E x F	1	496	496	0.0119	0.9148
Dx E x G	1	1891	1891	0.0455	0.8347
D x F x G	1	1200	1200	0.0289	0.8679
E x F x G	1	0.5	0.5	<0.001	0.9973
D x E x Fx G	1	10011	10011	0.2408	0.6325
Erro	12	498910	41576		

Tabela 10: Análise de variância do segundo experimento (Split-Plot).

Os testes de hipóteses para cada fator e as respectivas interações são realizados através do quociente entre o quadrado médio referente a cada fator e o quadrado médio do Erro associado ao estágio onde se encontra este fator ou a respectiva interação. Compara-se, então, este valor com uma distribuição F com os graus referentes ao fator (ou interação) no numerador e referentes ao Erro no denominador, de acordo com o estágio considerado. De acordo com estes testes nota-se que, dado um nível de significância de 5%, não há fatores (ou interações) com efeitos significativos.

5.3 Conclusões

A partir dos resultados obtidos, tem-se que o plano experimental utilizado não detectou, dentre os fatores analisados, nenhum significativamente influente no esforço gasto na execução de testes manuais em celulares. Esta conclusão leva à discussão acerca do uso de outros planos experimentais, ou até mesmo à investigação acerca de outros fatores que venham ter maior influência no respectivo estudo.

5.4 Comparação Entre Planos Experimentais

A seguir temos uma comparação entre diversas propostas de planos experimentais alternativos ao experimento split-plot descrito anteriormente. Como objetivo, tem-se o de encontrar uma metodologia que mais se adeque à realidade do projeto, enquadrando os diversos fatores que envolvem a execução de um experimento dentro desse cenário aqui descrito. Para tanto, descreve-se inicialmente algumas propostas de planos experimentais, e ao final realiza-se uma exibição de vantagens e desvantagens de cada abordagem.

5.4.1 Plano Experimental Completamente Aleatorizado com Estrutura Fatorial Completa 2^4

Seguindo a metodologia de um plano experimental completamente aleatorizado, segundo uma estrutura fatorial completa com 4 fatores, com 2 níveis associados a cada fator, tem-se a estrutura de aleatorização dos tratamentos (indicados pelas setas) às unidades experimentais (celulares) ilustrada abaixo:

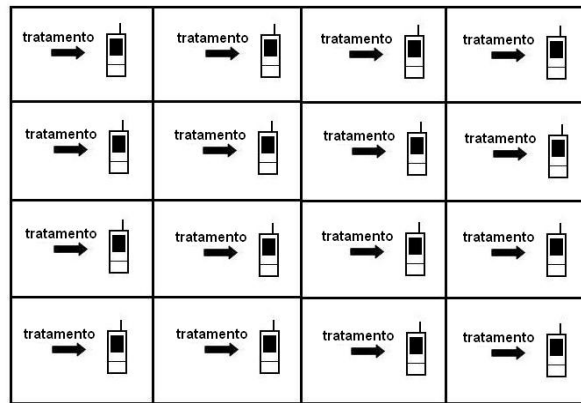


Figura 8: Plano Experimental Completamente Aleatorizado com Estrutura Fatorial Completa 2^4 .

De acordo com a figura, cada unidade experimental da réplica receberia um tratamento aleatoriamente, num único processo de aleatorização, de modo que cada possível configuração de alocação de tratamentos às unidades experimentais teria igual probabilidade de ocorrência. Tem-se, posteriormente, uma tabela com as observações por tratamento geradas em uma replicação do plano experimental proposto:

Tratamento	Fatores				Observação
	A	B	C	D	
(1)	-	-	-	-	X
a	+	-	-	-	X
b	-	+	-	-	X
ab	+	+	-	-	X
c	-	-	+	-	X
ac	+	-	+	-	X
bc	-	+	+	-	X
abc	+	+	+	-	X
d	-	-	-	+	X
ad	+	-	-	+	X
bd	-	+	-	+	X
abd	+	+	-	+	X
cd	-	-	+	+	X
acd	+	-	+	+	X
bcd	-	+	+	+	X
abcd	+	+	+	+	X

Tabela 11: Observações do experimento completamente aleatorizado com estrutura fatorial completa 2^4 .

As análises de variância do experimento sem replicação, e também considerando duas replicações são dadas a seguir:

ANOVA (a)		ANOVA (b)	
Fonte de variação	Graus de liberdade	Fonte de variação	Graus de liberdade
Tratamento	15	Réplica	1
Erro	0	Tratamento	15
Total	15	Erro	15
		Total	31

Tabela 12: Análises de variância do experimento (a) sem replicação e (b) com duas réplicas.

5.4.2 Plano Experimental Completamente Aleatorizado com Estrutura Fatorial Fracionada $\frac{1}{2}2^4$ (Res. IV)

Com um plano experimental completamente aleatorizado, segundo uma estrutura fatorial fracionada de resolução IV, temos a seguinte disposição de aleatorização dos tratamentos às unidades experimentais:

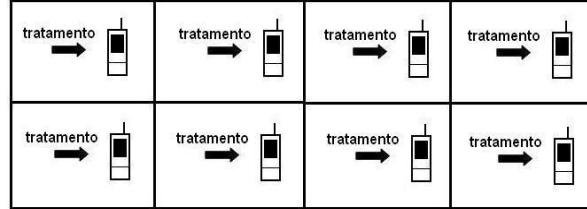


Figura 9: Plano Experimental Completamente Aleatorizado com Estrutura Fatorial Fracionada $\frac{1}{2}2^4$ (Resolução IV).

Em relação ao plano anterior, há uma diminuição no número de tratamentos utilizados, como também no número de unidades experimentais em cada possível réplica. A tabela logo a seguir apresenta as observações por tratamento geradas em uma replicação desta proposta de plano experimental:

Tratamento	Fatores				Identidade	
	A	B	C	D	I=ABCD	Observação
(1)	-	-	-	-	+	X
a	+	-	-	-	-	
b	-	+	-	-	-	
ab	+	+	-	-	+	X
c	-	-	+	-	-	
ac	+	-	+	-	+	X
bc	-	+	+	-	+	X
abc	+	+	+	-	-	
d	-	-	-	+	-	
ad	+	-	-	+	+	X
bd	-	+	-	+	+	X
abd	+	+	-	+	-	
cd	-	-	+	+	+	X
acd	+	-	+	+	-	
bcd	-	+	+	+	-	
abcd	+	+	+	+	+	X

Tabela 13: Observações do experimento completamente aleatorizado com estrutura fatorial fracionada $\frac{1}{2}2^4$ (Resolução IV)

Ao usar este formato de aleatorização, o número de tratamentos, com relação ao plano experimental sugerido anteriormente, é reduzido pela metade. Além disso, os efeitos dos fatores são confundidos com efeitos de interações, de acordo com a resolução utilizada.

Esta resolução define o nível de uma dada interação a ser considerado, e apenas este nível é utilizado na definição dos níveis individuais e de interação dos tratamentos. Temos, então, as análises de variância desta proposta tanto no caso sem replicação como o caso de duas replicações:

ANOVA (a)		ANOVA (b)	
Fonte de variação	Graus de liberdade	Fonte de variação	Graus de liberdade
Tratamento	7	Réplica	1
Erro	0	Tratamento	7
Total	7	Erro	7
		Total	15

Tabela 14: Análises de variância do experimento (a) sem replicação e (b) com duas réplicas.

5.4.3 Plano Experimental Strip-Block em Dois Estágios (Completamente Aleatorizado, Completamente Aleatorizado)

Considerado um plano experimental do tipo strip-block, temos os tratamentos aleatorizados em dois estágios, da forma abaixo:

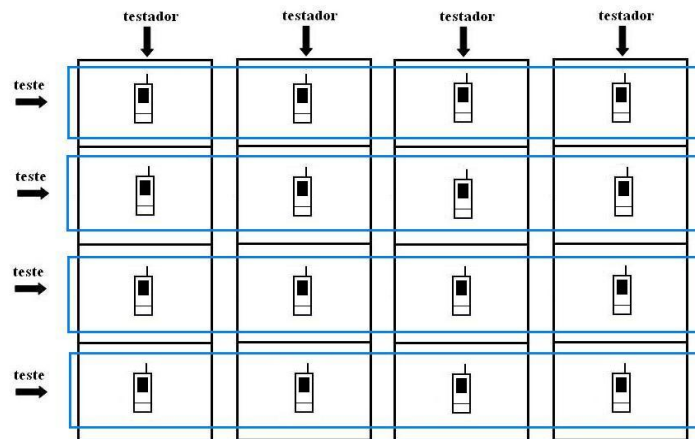


Figura 10: Plano Experimental Strip-Block.

Neste plano experimental, teríamos os tratamentos devido aos fatores provenientes do testador aleatorizados num primeiro estágio, e os tratamentos devidos aos fatores provenientes do teste num segundo estágio. Tanto o primeiro como o segundo estágio podem ser vistos como experimentos completamente aleatorizados, onde as unidades experimentais utilizadas em ambos os estágios são formadas por grupos de quatro celulares numa dada replicação. Dada a disposição de tratamentos do primeiro estágio, os tratamentos devidos ao teste no segundo estágio são aleatorizados de forma que todos os testadores executarão os testes na mesma ordem, devido ao fato de que o processo de aleatorização do segundo estágio determina um tratamento devido ao teste para quatro celulares em cada alocação. Estes celulares, por sua vez, já estão com seus tratamentos devido aos fatores do testador determinados no primeiro estágio. Temos, portanto, as observações do tempo de execução dos testes dadas da seguinte forma:

Estágio 1		Estágio 2				C
		-	+	-	+	
A	B	-	-	+	+	D
-	-	X	X	X	X	
+	-	X	X	X	X	
-	+	X	X	X	X	
+	+	X	X	X	X	

Tabela 15: Observações do experimento Strip-Block.

A análise de variância é descrita abaixo:

Fontes de variação	Graus de liberdade
Réplica	1
Testador	3
Erro (Testador)	3
Teste	3
Erro (Teste)	3
Testador x Teste	9
Erro (Interação)	9
Total	31

Tabela 16: Análise de variância - experimento Strip-Block com duas replicações.

5.4.4 Vantagens e Desvantagens dos Planos Experimentais Propostos

Na tabela exibida abaixo temos as vantagens e desvantagens de cada plano experimental proposto:

Plano experimental	Vantagens	Desvantagens
Completamente Aleatorizado (Estrutura Fatorial Completa 2^4)	Utiliza todos os tratamentos possíveis	Possibilidade de aprendizado Réplicas requerem muitas unidades experimentais Sem replicação não pode haver todas as interações
Completamente Aleatorizado (Estrutura Fatorial Fracionada $\frac{1}{2}2^4$)	Réplicas requerem muitas unidades experimentais	Possibilidade de aprendizado Não usa todos os tratamentos possíveis Os efeitos dos fatores estão confundidos com efeitos de interação
Strip-Block (Completamente Aleatorizado, Completamente Aleatorizado)	Utiliza todos os tratamentos possíveis Fácil processo de aleatorização	Réplicas requerem muitas unidades experimentais Possibilidade de aprendizado Sem replicação há dificuldade na obtenção de conclusões

Tabela 17: Vantagens e desvantagens dos planos experimentais propostos.

A próxima tabela descreve as quantidades de testes, testadores e unidades experimentais utilizadas segundo cada proposta de plano experimental anteriormente sugerida:

Plano experimental	Cel. utilizados	Testadores utilizados	Testes executados
Completamente Aleatorizado (Estrutura Fatorial Completa 2^4)	Sem replicação: 16 Com duas replicações: 32	Sem replicação: 4 Com duas replicações: 4	Sem replicação: 16 Com duas replicações: 32
Completamente Aleatorizado (Estrutura Fatorial Fracionada $\frac{1}{2}2^4$)	Sem replicação: 8 Com duas replicações: 16	Sem replicação: 4 Com duas replicações: 4	Sem replicação: 8 Com duas replicações: 16
Strip-Block (Completamente Aleatorizado, Completamente Aleatorizado)	Com duas replicações: 32	Com duas replicações: 4	Com duas replicações: 32

Tabela 18: Número de unidades experimentais, testadores e testes utilizados nos planos experimentais propostos.

6 Conclusões Gerais

6.1 Conclusões Acerca dos Estudos Realizados

Com relação aos estudos realizados, conclue-se que houve enorme ganho de conhecimentos, no sentido de abranger novas metodologias até então não adquiridas pelo estagiário Roberto Ferreira Manghi durante o curso de Bacharelado em Estatística. A análise de experimentos é uma ferramenta poderosa na busca de explicações acerca de fenômenos de interesse em praticamente todas as áreas científicas, repercutindo sensivelmente nos direcionamentos da vida prática em geral. Portanto, o conhecimento adquirido durante todo o período de estágio foi e será de extrema importância daqui por diante para o estudante.

6.2 Conclusões Acerca da Experiência Adquirida

No que diz respeito à experiência pessoal adquirida, o estagiário Roberto Ferreira Manghi pôde utilizar seus conhecimentos teóricos em situações reais, necessitando assim que a busca de soluções para problemas fosse realizada em um ambiente diferente do meio acadêmico ao qual está vinculado. O contato com problemas de uma área até então desconhecida pelo estagiário proporcionou uma interação entre conhecimentos de ambas as partes. Esta interação foi possível, principalmente, através do direcionamento dos pesquisadores Eduardo Aranha e Rodrigo Bonifácio, que proporcionaram ao estagiário o acesso tanto aos problemas relativos ao estágio dentro da área de Engenharia de Software como a possibilidade de uma troca de conhecimentos entre esta área e a área Estatística. Logo, conclue-se que o estágio do aluno Roberto Ferreira Manghi foi de grande êxito.

7 Referências

- COCHRAN, William Gemmel; COX, Gertrude Mary. Experimental designs. New York: J. Wiley ; London : Chapman and Hall, 1957.
- MONTGOMERY, Douglas C. Design and analysis of experiments. New York: John Wiley, 2001.
- SNEDECOR, George Waddel; COCHRAN, William G. Statistical methods. Ames: The Iowa State University Press, 1967.

8 Anexo

8.1 Função que executa o Teste de Não-Aditividade de Tukey para experimentos do tipo Quadrado Latino com Replicação

```
TukeyNADD.QL.REP<-function(objeto1)
{
y1<-NULL
y2<-NULL
y1<- fitted(objeto1)
y2<- y1^2
objeto2<- aov(y2 ~ objeto1[13]$model[,2] +
objeto1[13]$model[,3]:objeto1[13]$model[,2]
+ objeto1[13]$model[,4]+ objeto1[13]$model[,5])
ynew <- resid(objeto1)
xnew <- resid(objeto2)
objeto3 <- lm(ynew ~ xnew)
M <- anova(objeto3)
MSN <- M[1,3]
MSErr <- M[2,2]/(objeto1[8]$df.residual-1)
F0 <- MSN/MSErr
p.val <- 1 - pf(F0, 1,objeto1[8]$df.residual-1)
p.val
}
```