

**DELINEAMENTOS D-ÓTIMOS PARA OS MODELOS DE  
MICHAELIS-MENTEN E DE HILL**

**Iuri Emmanuel de Paula Ferreira**

Texto apresentada à Universidade Estadual Paulista “Júlio de Mesquita Filho” para a obtenção do título de Mestre em Biometria.

BOTUCATU  
São Paulo - SP  
Outubro – 2010

**DELINEAMENTOS D-ÓTIMOS PARA OS MODELOS DE  
MICHAELIS-MENTEN E DE HILL**

**Iuri Emmanuel de Paula Ferreira**

Orientadora: Prof. Dr. **Luzia Aparecida Trinca**

Texto apresentada à Universidade Estadual Paulista “Júlio de Mesquita Filho” para a obtenção do título de Mestre em Biometria.

BOTUCATU  
São Paulo - SP  
Outubro – 2010

## Ficha Catalográfica

Ficha catalográfica elaborada pela Seção Técnica de Aquisição e Tratamento da Informação  
Divisão Técnica de Biblioteca e Documentação - Campus De Botucatu - UNESP  
Bibliotecária responsável: *Sulamita Selma Clemente Colnago* - CRB 8/4716

Ferreira, Iuri Emmanuel de Paula.

Delineamentos D-ótimos para os modelos de Michaelis-Menten e de Hill /  
Iuri Emmanuel de Paula Ferreira. - Botucatu, 2011

Dissertação (mestrado) - Instituto de Biociências de Botucatu,  
Universidade Estadual Paulista, 2011

Orientador: Luzia Aparecida Trinca

Capes: 10202072

1. Biologia - Experimentos. 2. Biometria.

Palavras-chave: Delineamentos Localmente Ótimos; Delineamentos Ótimos  
Pseudo-Bayesianos; Critério D; Modelo Não-Linear

## **Dedicatória**

Dedico este trabalho às Professoras Luzia A. Trinca e Cláudia P. Ferreira.

## **Agradecimentos**

Agradeço à Professora Luzia A. Trinca, pela orientação e por todo apoio durante o curso. Agradeço também à Professora Cláudia P. Ferreira, pela ajuda e pelos conselhos que me foram muito úteis durante este período. Também devo muito a minha família por toda compreensão e carinho, em especial, a minha avó, a minha mãe e a minha tia Elza. Agradeço a minha namorada Carol, por todo cuidado e ajuda dedicados a mim. Agradeço também ao companheirismo do pessoal da Biometria, alunos, funcionários e professores. Além disso, agradeço ao apoio financeiro fornecido pela CAPES, que tem sido muito útil.

# Sumário

	Página
<b>LISTA DE FIGURAS</b>	<b>viii</b>
<b>LISTA DE TABELAS</b>	<b>xi</b>
<b>RESUMO</b>	<b>xiii</b>
<b>SUMMARY</b>	<b>xv</b>
<b>1 INTRODUÇÃO</b>	<b>1</b>
<b>2 MOTIVAÇÃO E OBJETIVOS</b>	<b>3</b>
<b>3 FUNDAMENTOS TEÓRICOS</b>	<b>6</b>
3.1 Um pouco sobre experimentos . . . . .	6
3.2 Desafios na fase de delineamento . . . . .	7
3.3 Delineamentos experimentais ótimos . . . . .	9
3.4 Modelos de regressão e estimação de seus parâmetros . . . . .	10
3.5 Delineamentos exatos e contínuos . . . . .	17
3.6 Principais critérios para delineamento ótimo . . . . .	18
3.7 Abordagens para a construção de delineamentos ótimos para modelos não- lineares . . . . .	20
3.8 O critério D em situações não-lineares . . . . .	21
3.9 Método de Monte Carlo para obter o valor aproximado da função-critério	22
3.10 A variância padronizada . . . . .	23

	vii
3.11 Algoritmos para busca de delineamentos ótimos exatos . . . . .	24
3.11.1 Algoritmos genéticos . . . . .	24
3.11.2 Algoritmos tipo <i>exchange</i> . . . . .	27
<b>4 METODOLOGIA</b>	<b>29</b>
<b>5 RESULTADOS E DISCUSSÃO</b>	<b>36</b>
5.1 Comparação dos algoritmos genético e <i>exchange</i> . . . . .	36
5.2 Delineamentos localmente D-ótimos . . . . .	41
5.3 Delineamentos D-ótimos pseudo-Bayesianos . . . . .	48
5.4 Delineamentos ótimos segundo critérios compostos . . . . .	57
<b>6 CONCLUSÃO E PERSPECTIVAS FUTURAS</b>	<b>61</b>
<b>7 ANEXO A</b>	<b>63</b>
<b>8 ANEXO B</b>	<b>72</b>
<b>REFERÊNCIAS BIBLIOGRÁFICAS</b>	<b>81</b>



## Lista de Figuras

	Página	
1	Equação de Michaelis-Menten para valores diferentes de $\kappa$ e $\nu$ : em (a), $\nu$ é fixado em 3 e varia-se $\kappa$ ; em (b), $\kappa$ é fixado em 5 e varia-se $\nu$ . . . . .	16
2	Equação de Hill com $\kappa = 5$ , $\nu = 3$ e valores diferentes (0, 5, 1, 0 e 1, 5) para o coeficiente de Hill. . . . .	17
3	Desenho esquemático utilizado para exemplificar o procedimento realizado pelo algoritmo genético em cada iteração. . . . .	30
4	Determinante da matriz de informação para o melhor delineamento encontrado pelos algoritmos genético e <i>exchange</i> em função do número de delineamentos avaliados. . . . .	37
5	Tempo de execução dos algoritmos genético e <i>exchange</i> na busca de delineamentos D-ótimos exatos pseudo-Bayesianos com $N = 8$ em função do tamanho de amostras para representar o parâmetro $\kappa$ . . . . .	39
6	Tempos de execução dos algoritmos genético e <i>exchange</i> em função do tamanho de amostra. Dados e ajuste linear para a escala logarítmica. . .	41
7	Razão dos tempos de execução Razão( $a_i$ ) dos algoritmos genético e <i>exchange</i> em função do tamanho de amostra $a_i$ , calculada a partir dos valores de tempo ajustados pelo modelo (50). . . . .	42
8	Valores para $x^*$ em função da informação <i>a priori</i> para $\kappa$ . . . . .	43

9	Variância padronizada para os delineamentos localmente D-ótimos exatos (linha sólida) para o modelo de Michaelis-Menten a partir da informação <i>a priori</i> contida no trabalho de Condomina et al. (2002), sendo referente à parte proximal em (a), à parte medial em (b) e à parte distal em (c). Os círculos indicam a variância padronizada para o delineamento D-ótimo contínuo avaliada em seus dois pontos de suporte. A linha pontilhada é utilizada como referência para a D-otimalidade. . . . .	45
10	Variância padronizada para os delineamentos localmente D-ótimos exatos (linha sólida) para o modelo de Hill a partir da informação <i>a priori</i> contida no trabalho de Condomina et al. (2002), sendo referente à parte proximal em (a), à parte medial em (b) e à parte distal em (c). A linha pontilhada é utilizada como referência para a D-otimalidade. . . . .	47
11	Pontos de suporte inferiores dos delineamentos D-ótimos pseudo-Bayesianos para o modelo de Michaelis-Menten em função do coeficiente de variação adotado para a <i>a priori</i> do parâmetro $\kappa$ . . . . .	49
12	Distribuições discretas <i>a priori</i> utilizadas para representar o parâmetro $\gamma$ , o coeficiente de Hill. Em (i) tem-se uma distribuição simétrica com média 1; em (ii) tem-se uma distribuição assimétrica à direita com média 0,805; em (iii) tem-se uma distribuição assimétrica à esquerda com média 1,195. . . . .	51
13	Gráficos da variância padronizada para os delineamentos D-ótimos exatos pseudo-Bayesianos para o modelo de Michaelis-Menten. . . . .	53
14	Gráficos da variância padronizada para os delineamentos D-ótimos exatos pseudo-Bayesianos para o modelo de Hill quando utiliza-se uma distribuição discreta simétrica com média 1 para representar $\gamma$ . . . . .	54
15	Gráficos da variância padronizada para os delineamentos D-ótimos exatos pseudo-Bayesianos para o modelo de Hill quando utiliza-se uma distribuição discreta assimétrica à direita com média 0,805 para representar $\gamma$ . . . . .	55

16	Gráficos da variância padronizada para os delineamentos D-ótimos exatos pseudo-Bayesianos para o modelo de Hill quando utiliza-se uma distribuição discreta assimétrica à esquerda com média 1,195 para representar $\gamma$ . . . . .	56
17	Pontos de suporte de delineamentos localmente D-ótimos exatos a partir do critério composto de D-otimalidade em função de diferentes valores de $\lambda$ . . . . .	59

## Lista de Tabelas

	Página
1 Ajustes do modelo (50) para os tempos de execução (em segundos) dos algoritmos <i>exchange</i> e genético. . . . .	40
2 Estimativas dos parâmetros do modelo de Michaelis-Menten obtidas por Condomina et al. (2002). . . . .	43
3 Delineamentos localmente D-ótimos exatos (número de réplicas entre parênteses) para o ajuste do modelo de Michaelis-Menten utilizando como informação <i>a priori</i> as estimativas obtidas por Condomina et al. (2002). Seja $x^*$ o ponto de suporte inferior do delineamento D-ótimo contínuo. . . . .	44
4 Delineamentos localmente D-ótimos exatos (número de réplicas entre parênteses) para o modelo de Hill, utilizando-se as estimativas obtidas por Condomina et al. (2002) como informação <i>a priori</i> para representar $\kappa$ e $\nu$ . Para $\gamma$ foi suposto o valor de 1. . . . .	47
5 Delineamentos D-ótimos exatos pseudo-Bayesianos (número de réplicas entre parênteses) para o modelo de Michaelis-Menten, utilizando-se como informação <i>a priori</i> distribuições Gama com diferentes coeficientes de Variação (C.V.) baseadas nas estimativas obtidas por Condomina et al. (2002). . . . .	50

- 6 Delineamentos D-ótimos exatos pseudo-Bayesianos (número de réplicas entre parênteses) para o modelo de Hill, utilizando-se distribuições Gama com diferentes Coeficientes de Variação (C.V.) como informação *a priori* para representar  $\kappa$  e distribuições discretas para representar  $\gamma$ . Os valores esperados utilizados para os parâmetros  $\kappa$  e  $\nu$  foram baseados nas estimativas apresentadas por Condomina et al. (2002). . . . . 52
- 7 Delineamentos localmente D-ótimos exatos (número de réplicas entre parênteses) a partir do critério composto com diferentes valores de  $\lambda$  e utilizando-se como informação *a priori* para  $\kappa$  e  $\nu$  as estimativas obtidas por Condomina et al. (2002) para a porção proximal do intestino. Para  $\gamma$  foi suposto o valor de 1. . . . . 58
- 8 Delineamentos D-ótimos exatos pseudo-Bayesianos (número de réplicas entre parênteses) a partir do critério D composto com valor de  $\lambda = 0,5$  e diferentes coeficientes de variação para a *priori* de  $\kappa$ , baseando-se nas estimativas obtidas por Condomina et al. (2002). . . . . 60

# DELINEAMENTOS D-ÓTIMOS PARA OS MODELOS DE MICHAELIS-MENTEN E DE HILL

Autor: IURI EMMANUEL DE PAULA FERREIRA  
Orientadora: Prof. Dr. LUZIA APARECIDA TRINCA

## RESUMO

Os resultados de muitos experimentos em áreas da biologia, como a farmacologia, a bioquímica e a agronomia, geralmente são analisados por ajustes de modelos não-lineares através dos quais pretende-se explicar a resposta através dos fatores pré-especificados no experimento. As estimações dos parâmetros ou das funções de interesse podem ser imprecisas se os níveis dos fatores não forem adequadamente escolhidos, impossibilitando ao pesquisador a obtenção da informação desejada sobre o objeto de estudo. A construção de um delineamento ótimo, que maximize a informação sobre algum aspecto de interesse, é crucial para o sucesso da prática experimental. O objetivo deste trabalho foi a obtenção de delineamentos D-ótimos exatos para modelos não-lineares utilizados para estudar cinética enzimática e transporte de minerais no organismo, como o de Michaelis-Menten e o de Hill. Para este fim, duas abordagens foram consideradas, a saber, a de delineamentos localmente

ótimos e a pseudo-Bayesiana. Com o auxílio dos algoritmos genético e *exchange* foi possível obter delineamentos D-ótimos exatos para o modelo de Michaelis-Menten, para o modelo de Hill e para ambos, considerando-se valores diferentes e distribuições com diversos coeficientes de variação como informação *a priori*.

# D-OPTIMAL DESIGNS FOR THE MICHAELIS-MENTEN AND HILL MODELS

Author: IURI EMMANUEL DE PAULA FERREIRA

Adviser: Prof. Dr. LUZIA APARECIDA TRINCA

## SUMMARY

The results of many experiments in biological fields, as pharmacology, biochemistry and agriculture, usually are analyzed by fitting nonlinear models, which are supposed to describe well the response to the pre-specified factors in the experiment. The estimates of the parameters or of their functions of interest could be imprecise if the factor levels are not adequately chosen. The construction of an optimum design, which maximizes the information about some aspect of interest, is crucial for the success of the experimental practice. The aim of this work was constructing exact D-optimal designs for nonlinear models usually used in studies of enzyme kinetics and mineral transport in organisms, such as the Michaelis-Menten and Hill models. Two approaches were considered, the locally optimal and pseudo-Bayesian designs. Genetic and Exchange algorithms were used for getting exact designs aiming at the Michaelis-Menten model, aiming at the Hill model, each one



separately, and aiming at both models when considering a composite criterion. Different values and probability distributions with several variation coefficients were considered as prior information.

# 1 INTRODUÇÃO

Inúmeros processos, nas mais variadas áreas da ciência, que adquirem conhecimento através da experimentação, são compreendidos a partir do ajuste de curvas não-lineares. Na maioria das aplicações estes modelos são fundamentados nas leis físicas do processo mas, às vezes, eles são usados também como modelos empíricos. A obtenção dos dados para a estimação dos parâmetros do modelo envolve a realização de um experimento que, por sua vez, deve ser planejado cuidadosamente para o sucesso do estudo. Princípios gerais que norteiam o planejamento de experimentos se solidificaram com os trabalhos de Fisher e colaboradores a quase 100 anos atrás. No entanto, tais diretrizes focalizaram experimentos cujos resultados são analisados por modelos lineares e ademais, de certa forma, esbanjam recursos e tempo para sua execução. O desenvolvimento científico e a sofisticação tecnológica têm permitido a observação de informações cada vez mais detalhadas mas nem sempre com baixo custo, exigindo a realização de experimentos eficientes mas de tamanhos razoáveis. Um aspecto do sucesso do experimento pode ser traduzido como alta precisão das estimativas dos efeitos investigados, ou seja, pequena margem de erro, que depende dos níveis dos fatores quantitativos que são alterados com o objetivo de avaliar suas consequências sobre uma resposta de interesse.

A teoria de delineamentos ótimos propõe a busca destes níveis e respectivas réplicas de forma a maximizar a precisão das estimativas. Tal perspectiva é muito atrativa na situação não-linear já que, diferentemente do caso linear, não é o óbvio quais níveis fornecem a máxima informação. Uma grande dificuldade na construção de delineamentos eficientes é que a precisão das estimativas dos efeitos dependem dos verdadeiros valores destes, que, obviamente, na fase de planejamento

são desconhecidos. Esta dificuldade tem sido contornada principalmente através da incorporação de (1) valores pontuais obtidos da informação existente sobre os parâmetros na fase de planejamento, resultando em delineamentos localmente ótimos, ou (2) distribuições de probabilidade *a priori* para os parâmetros, resultando em delineamentos conhecidos como pseudo-Bayesianos. Vários outros detalhes compõem e dificultam o problema como a escolha da função-critério relacionada com a precisão, pois existem várias propriedades que podem ser de interesse, assim como a escolha do método de otimização e da forma de implementá-lo.

Tendo como foco os modelos de Michaelis-Menten e de Hill, amplamente aplicados em estudos de cinética enzimática, farmacocinéticos e bioquímicos, este trabalho tem como objetivos: o estudo da problemática de construção de delineamentos ótimos para modelos não-lineares segundo o critério de D-otimalidade; a avaliação de alguns aspectos relacionados ao desempenho de dois tipos de algoritmos, genético e *exchange*, na busca de delineamentos ótimos; e a avaliação da sensibilidade dos delineamentos ótimos com respeito à quantidade de informação incorporada *a priori*.

Para inserir o estudo realizado num contexto prático, utilizou-se o trabalho de Condomina et al. (2002). Desta maneira, visou-se a construção de delineamentos eficientes para conduzir experimentos futuros partindo-se das estimativas apresentadas pelos autores no referido trabalho.

Na seção 2, uma breve descrição do trabalho de Condomina et al. (2002) é apresentada, juntamente com os objetivos estabelecidos para este estudo. Já na seção 3 encontra-se a teoria na qual este trabalho está fundamentado. Os métodos utilizados para obtenção e otimização da função-critério de D-otimalidade são descritos na seção 4. Por sua vez, a seção 5 traz os resultados e sua discussão. Por fim, a seção 6 apresenta a conclusão deste estudo.

## 2 MOTIVAÇÃO E OBJETIVOS

Para desenvolver um estudo sobre delineamentos D-ótimos e sua aplicabilidade em ensaios biológicos, mantendo em foco os modelos não-lineares de Michaelis-Menten e de Hill, utilizou-se como motivação prática o trabalho de Condomina et al. (2002). Este trabalho teve como finalidade investigar o transporte de zinco em diferentes porções do intestino delgado e no cólon de ratos albinos. Além disso, os autores consideraram a interação do cobre no processo. Uma breve descrição deste trabalho é apresentada a seguir.

Primeiramente, seis concentrações diferentes de zinco foram ministradas como tratamentos às unidades experimentais (segmentos do trato gastrointestinal preparados *in vitro*) oriundas de diferentes porções do intestino delgado (proximal, medial e distal) e do cólon e, como resposta, o influxo de zinco foi mensurado. Os valores de concentração de zinco utilizados no experimento foram arbitrariamente escolhidos no intervalo de  $7 \cdot 10^{-4}$  a  $11 \text{ mM}$ . Um dos objetivos do trabalho foi encontrar o modelo que melhor representa o processo de transporte de zinco através da membrana de cada uma das porções estudadas, sendo que três possibilidades foram consideradas, a saber:

1. Processo de transporte não-saturável, no qual o influxo de zinco é diretamente proporcional a sua concentração externa e à permeabilidade difusiva;
2. Processo de transporte saturável no qual a relação entre a concentração externa e o influxo de zinco é modelada pela equação de Michaelis-Menten;
3. Processos de transporte saturável (item 1) e não-saturável (item 2) agindo em conjunto.

Os modelos foram ajustados e comparados através do Critério de Informação de Akaike, Akaike (1986). O modelo de Michaelis-Menten apresentou um ajuste melhor para o intestino delgado, porém, os valores estimados para os parâmetros mostraram uma diferença notável entre o transporte de zinco na porção proximal quando comparado ao que ocorre nas porções medial e distal. Para os autores, esta é uma evidência da existência de transportadores diferentes que atuam de forma cooperativa.

Além disso, os autores estudaram a influência do cobre no transporte de zinco. Para tanto, os pesquisadores fixaram a concentração de zinco em  $0.037 \text{ mM}$  e variaram a concentração de cobre em quatro níveis (contando com o controle) e, por fim, mediram o influxo de zinco em cada porção estudada. Então, foi verificado que o cobre é capaz de inibir o transporte de zinco, de maneira dependente da concentração. Dessa forma, os autores formularam a hipótese de existência de transportadores capazes de se ligarem aos dois íons num processo em que haja a cooperação negativa entre os ligantes.

Os autores concluíram que o transporte de zinco no intestino delgado de ratos albinos é um processo saturável que talvez seja afetado por cooperatividade positiva ou negativa entre ligantes e, desta maneira, o modelo de Hill é uma alternativa a ser estudada, mesmo que o trabalho de Condomina et al. (2002) não o tenha mencionado.

Para conduzir os estudos, foi proposta uma situação hipotética na qual seja interesse de pesquisa estimar com maior precisão os parâmetros da equação de Michaelis-Menten para modelar o transporte de zinco no intestino delgado a partir de dados provenientes de um experimento futuro. Ademais, foi suposto que os pesquisadores desejam ajustar o modelo de Hill para averiguarem se a descrição do processo pode ser mais completa com a inclusão do coeficiente  $\gamma$ . Além disso, foi suposto que o material utilizado é caro, que o processo de medida é trabalhoso e demorado e que há uma quantidade limitada de cobaias a serem utilizadas, resultando num número máximo  $N$  de unidades experimentais disponíveis.

Então, como motivação prática, objetivou-se delinear este experimento futuro de maneira a maximizar a informação esperada a respeito dos parâmetros e, assim, obter ajustes mais fidedignos. Com esta finalidade, foram utilizadas as abordagens localmente ótima e pseudo-Bayesiana de D-otimalidade.

## 3 FUNDAMENTOS TEÓRICOS

### 3.1 Um pouco sobre experimentos

Experimentos são conduzidos a fim de se estudar como um conjunto de variáveis controladas afeta os valores observados de outras variáveis. Em suma, um experimento consiste na observação das possíveis alterações de uma resposta mediante perturbações deliberadamente provocadas no sistema de estudo. Cada conjunto de perturbações é denominado tratamento, que pode ser categoria de variáveis explanatórias qualitativas, níveis de variáveis explanatórias quantitativas ou combinação de ambos. Por sua vez, os tratamentos são ministrados às unidades básicas do estudo, denominadas unidades experimentais.

A resposta observada também pode estar sujeita às variações ocasionadas por fatores que não são alvos de investigação. Uma fonte de variação sempre presente é aquela oriunda da variabilidade natural existente entre réplicas similares, chamada de erro experimental ou erro aleatório, ou seja, a variabilidade nas respostas de unidades experimentais que não pode ser explicada pelos tratamentos ou por fatores externos importantes. Se existirem outros fatores de heterogeneidade no experimento, estes devem ser controlados, numa forma de evitar que influências externas aumentem o erro aleatório, provoquem estimativas viciadas ou confundimentos, que ocorrem quando não é possível determinar se o efeito observado foi decorrente de um tratamento ou de outra causa de variação.

Geralmente, um experimento tem vários objetivos, sendo os mais comuns a estimação de efeitos de tratamentos e a elucidação do comportamento de um sistema através de uma curva ou modelo ajustado. Porém, um experimento só

atinge seus objetivos quando é conduzido seguindo os princípios básicos da experimentação e, desta maneira, apresenta resultados não viciados, eficientes e passíveis de inferência. Os três princípios básicos da experimentação advogados por Fisher (1935) são aleatorização das unidades experimentais aos tratamentos, controle de variações externas e replicagem de tratamentos idênticos. A probabilidade de existir um vício acentuado nos resultados é reduzida pelo processo de aleatorização, conduzida esta que garante a validade do processo de inferência, no qual as conclusões obtidas em um experimento são estendidas à população alvo de estudo. Além disso, a blocagem permite separar os efeitos decorrentes dos tratamentos daqueles ocasionados pela heterogeneidade no material experimental e, assim, aumenta o poder de testes de hipóteses. Por sua vez, a replicação é necessária para garantir uma medida de variabilidade dentro dos tratamentos, produzindo estimativa válida da variabilidade casual e aumentando a eficiência dos estimadores. Assim, o planejamento experimental adequado é crucial para o sucesso da prática, sendo uma das fases mais importantes em todo o processo de experimentação.

Os conceitos fundamentais para o desenvolvimento da Teoria Clássica de Experimentos Planejados surgiram em meados do século XX e são abordados em uma vasta bibliografia como, por exemplo, nos trabalhos de Fisher (1935), Box et al. (1978) e Cochran & Cox (1957).

### **3.2 Desafios na fase de delineamento**

A formulação do delineamento experimental não é uma tarefa fácil, pois envolve muitos aspectos distintos, tais como as hipóteses a serem testadas (que estão relacionadas com o objetivo do estudo), os fatores que influenciam a resposta, a região experimental, o tipo e a quantidade do material disponível, entre outros. Para Atkinson et al. (2007), o sucesso do delineamento de um experimento requer o cálculo e o uso de toda a informação *a priori*, mesmo que seja de uma maneira informal. O que é conhecido deve ser resumido e questões a serem respondidas devem ser claramente formuladas. Além disso, fatores que podem afetar a resposta devem



ser listados.

Mas um experimento é proposto somente quando não se compreende de maneira satisfatória um processo e, portanto, as variáveis que de fato afetam a resposta não são conhecidas *a priori*. Geralmente é tido como ponto de partida um grande número de fatores que podem, conjunta ou separadamente, explicar a variabilidade na resposta. Portanto, no início da fase de delineamento o pesquisador deve definir cuidadosamente com quais fatores deseja trabalhar, sendo que uma escolha errada pode implicar num modelo que não é capaz de descrever o comportamento da variável resposta e, então, resultar num experimento mal sucedido. O pesquisador deve também garantir que os fatores escolhidos possam ser avaliados ou medidos com facilidade, a partir de um esforço plausível para a prática. Todas as variáveis importantes que não são alvos de investigação devem ser controladas durante a execução do experimento. Por exemplo, se a temperatura não é um fator de interesse mas pode afetar a resposta, deverá ser mantida constante. Além disso, a blocagem é exigida quando não é possível assegurar a homogeneidade de tais influências sobre todas as unidades experimentais.

Outro detalhe é que a escolha da região experimental para os fatores quantitativos requer o máximo de cuidado. Se a região especificada for pequena demais, o efeito da variável expanatória não será observado devido ao erro aleatório, mesmo que exista e seja importante. Se a região for ampla demais, haverá a possibilidade do comportamento não ser descrito adequadamente pelo modelo através do qual pretende-se explicar os dados, levando a conclusões errôneas. Aliás, a estratégia de delineamento a ser seguida depende do modelo proposto e este, por sua vez, deve representar os dados de forma parcimoniosa, resumindo bem a informação que eles carregam a partir dos efeitos principais e sem apresentar uma complexidade desnecessária que poderia incluir um número excessivo de parâmetros no modelo, requerendo mais unidades experimentais ou produzindo estimativas dos parâmetros pouco precisas.

Outra questão levantada nesta fase é o tamanho do experimento, ou

seja, o número de unidades experimentais a serem utilizadas. Para defini-lo, deve-se ponderar sobre os custos do material experimental e dos processos de medida, sobre o tempo de execução do experimento, sobre a precisão exigida para as estimativas de interesse a fim de garantir a testabilidade das hipóteses de interesse, dentre muitos outros aspectos. Um tamanho de experimento pequeno pode ocasionar a obtenção de resultados não conclusivos por falta de poder do teste, já utilizar um tamanho muito grande significa desperdiçar material e tempo. Na maioria dos casos, os pesquisadores indicam um tamanho de experimento plausível para a prática e a construção do delineamento procura definir os tratamentos a serem designados às unidades experimentais, de forma a maximizar a precisão das estimativas dos efeitos de interesse.

Delineamentos que atendem a tais critérios são ditos ótimos. Os delineamentos clássicos para modelos lineares, seja com fatores qualitativos ou fatores quantitativos, são delineamentos ótimos segundo alguns aspectos e podem ser encontrados na bibliografia básica de planejamentos de experimentos, inclusive a partir de muitos exemplos, basta ver Fisher (1971) ou Box et al. (1978). Porém, quando o objetivo é ajustar modelos não-lineares, os delineamentos clássicos não são adequados e precisamos de métodos mais específicos para resolver o problema. Neste contexto é que surgiu a Teoria de Delineamentos Ótimos, construída por Kiefer (1959) com a finalidade de encontrar delineamentos que maximizem a informação a partir da otimização de propriedades ligadas aos estimadores de interesse. Uma lista de referências básicas que muito contribui nessa área é Fedorov (1972), Pukelsheim (1993) e Atkinson et al. (2007).

### **3.3 Delineamentos experimentais ótimos**

A partir da teoria de delineamentos ótimos é possível planejar experimentos que proporcionam a máxima informação a respeito dos aspectos que deseja-se investigar no problema. Conforme Atkinson et al. (2007), seu objetivo é garantir planejamentos experimentais eficientes para um conjunto de propósitos especificados *a*

*priori* pelo pesquisador. Com seu auxílio é possível selecionar os tratamentos que serão alocados às unidades experimentais, de maneira a maximizar a informação esperada sobre algum aspecto de interesse, possibilitando resultados conclusivos até em casos em que o material experimental é limitado. Porém, é importante notar que a determinação da quantidade de material experimental a ser utilizada geralmente está fora do seu escopo.

Além do tamanho do experimento, a teoria de delineamentos ótimos exige que a região experimental esteja bem definida e sua escolha, por sua vez, depende do conhecimento do pesquisador a respeito do processo a ser estudado e das possíveis limitações físicas no uso de todos os tratamentos que em teoria seriam plausíveis.

A necessidade de especificação do modelo é outro ponto a ser notado. Para a teoria de delineamento ótimo, o modelo especificado *a priori* é suposto como correto e a designação de unidades experimentais adicionais para avaliar a falta de ajuste fica a critério do pesquisador.

Além disso, as técnicas de delineamento ótimo determinam os níveis (para variáveis quantitativas) e o número de réplicas a serem utilizados em um experimento, dado um grau de liberdade fixo para a estimação de  $\sigma^2$ . Em sua maior parte, os critérios para a construção de delineamentos ótimos consideram  $\sigma^2$  conhecido e buscam maximizar a informação sobre os efeitos dos tratamentos.

Para entender melhor o processo de delineamento ótimo, é preciso compreender como ocorre a estimação dos parâmetros e como a precisão de seus estimadores pode ser afetada pelo conjunto de tratamentos a serem utilizados. Neste trabalho foram abordados os estimadores de mínimos quadrados, apresentados brevemente na seção 3.4 .

### **3.4 Modelos de regressão e estimação de seus parâmetros**

Os resultados de um experimento com  $N$  unidades podem ser representados, sem perda de generalidade, por um modelo estocástico ou estatístico no qual

o conjunto de  $N$  respostas observadas  $\mathbf{y}_{N \times 1}$  é suposto ser o resultado da soma entre o vetor de valores esperados para as respostas dadas as  $m$  variáveis explanatórias  $(\mathbf{x}_1, \dots, \mathbf{x}_m)$ , denotado por  $\boldsymbol{\eta}_{N \times 1}(\mathbf{x}_1, \dots, \mathbf{x}_m)$ , e um vetor de erros aleatórios  $\boldsymbol{\epsilon}_{N \times 1}$ , ou seja,

$$\mathbf{y} = \boldsymbol{\eta}(\mathbf{x}_1, \dots, \mathbf{x}_m) + \boldsymbol{\epsilon}. \quad (1)$$

A presença de erros aditivos é uma suposição que pode ser violada, porém, em grande parte dos casos é possível realizar transformações que fazem com que erros de outros tipos se tornem aditivos. Veja, por exemplo, Box & Cox (1964). Outras suposições costumeiramente utilizadas para inferências são as de independência e de homocedasticidade dos erros. Além disso, é razoável assumir que o valor esperado para os mesmos seja zero, pois, em caso contrário, este valor poderia ser incorporado pela parte fixa do modelo, ou seja, pela função  $\boldsymbol{\eta}(\mathbf{x}_1, \dots, \mathbf{x}_m)$ .

Tais modelos também são conhecidos como modelos de regressão e podem ser divididos em duas classes, a saber, a dos paramétricos e a dos não paramétricos, como colocado por Fox (2005). Nos paramétricos, a função  $\boldsymbol{\eta}(\mathbf{x}_1, \dots, \mathbf{x}_m)$  depende, além das variáveis explanatórias, de um conjunto de parâmetros que definem a forma da função a ser ajustada como, por exemplo, retas e planos, curvas ou superfícies de resposta dadas por polinômios ou modelos não-lineares. Sendo assim, tal função fica mais adequadamente especificada por  $\boldsymbol{\eta}(\mathbf{x}_1, \dots, \mathbf{x}_m, \boldsymbol{\theta})$ , em que  $\boldsymbol{\theta}$  é o vetor de parâmetros do modelo. Por sua vez, os modelos não paramétricos não necessitam da imposição, *a priori*, de formas para representar relações entre variáveis, sendo estas determinadas unicamente pelos dados. Neste trabalho serão considerados modelos com apenas uma variável explanatória e não-lineares nos parâmetros, ditos modelos de regressão não-lineares simples e expressos por

$$\mathbf{y} = \boldsymbol{\eta}(\mathbf{x}, \boldsymbol{\theta}) + \boldsymbol{\epsilon}, \quad (2)$$

em que  $\boldsymbol{\theta}$  é um vetor coluna de  $p$  parâmetros e  $\boldsymbol{\eta}(\mathbf{x}, \boldsymbol{\theta})$  é uma função não-linear em  $\boldsymbol{\theta}$ .

A fim de exemplificar a estimação pelo método de mínimos quadrados,

considera-se, primeiramente, um modelo de regressão linear expresso por

$$y_i = \theta_1 + \theta_2 x_i + \theta_2 x_i^2 + \dots + \theta_p x_i^{p-1} + \epsilon_i, \quad (3)$$

com  $i = \{1, 2, \dots, N\}$  e em que  $y_i$  é a resposta obtida, dada pelo seu valor esperado no ponto  $x_i$  mais um erro aleatório  $\epsilon_i$ , e  $\theta_k$ , com  $k = \{1, 2, \dots, p\}$  e  $p$  conhecido, é o  $k$ -ésimo parâmetro do modelo. Para desenvolver o método, também são assumido erros independentes e homocedásticos. Este modelo pode ser escrito na forma matricial da seguinte maneira:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\theta} + \boldsymbol{\epsilon}, \quad (4)$$

em que  $\mathbf{y}_{N \times 1}$  é o vetor de respostas observadas,  $\mathbf{X}_{N \times p}$  é a matriz de delineamento (definida pelo modelo e pelo delineamento),  $\boldsymbol{\theta}_{p \times 1}$  é o vetor de parâmetros e  $\boldsymbol{\epsilon}_{N \times 1}$  é o vetor de erros aleatórios.

No método de mínimos quadrados as estimativas dos parâmetros são obtidas através da minimização da soma de quadrados dos erros, expressa por

$$S(\mathbf{X}, \boldsymbol{\theta}) = \boldsymbol{\epsilon}^T \boldsymbol{\epsilon} = (\mathbf{y} - \mathbf{X}\boldsymbol{\theta})^T (\mathbf{y} - \mathbf{X}\boldsymbol{\theta}). \quad (5)$$

É possível mostrar que o valor mínimo de  $S(\mathbf{X}, \boldsymbol{\theta})$  ocorre quando verifica-se a igualdade

$$\frac{\partial S(\mathbf{X}, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = 0, \quad (6)$$

que implica no sistema de equações normais, isto é,

$$\mathbf{X}^T \mathbf{X} \hat{\boldsymbol{\theta}} = \mathbf{X}^T \mathbf{y}, \quad (7)$$

em que  $\hat{\boldsymbol{\theta}}$  é o estimador de mínimos quadrados dos parâmetros. Por sua vez, a solução para o sistema existe e é única somente quando  $\mathbf{X}^T \mathbf{X}$  admite inversa. Para tanto, a matriz de delineamento  $\mathbf{X}$  tem que ter posto completo, ou seja, ser constituída por pelo menos  $p$  linhas ou colunas linearmente independentes. Uma condição necessária para isto é  $p \leq N$ . Além disso, o determinante de  $\mathbf{X}^T \mathbf{X}$  deve ser diferente de zero. Caso contrário, somente funções dos parâmetros poderão ser estimadas. Quando tais condições são satisfeitas, o estimador dos parâmetros é expresso por

$$\hat{\boldsymbol{\theta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} \quad (8)$$

e a sua matriz de variâncias e covariâncias é dada por

$$\text{Var}(\hat{\boldsymbol{\theta}}) = \sigma^2(\mathbf{X}^T \mathbf{X})^{-1}. \quad (9)$$

em que  $\sigma^2$  é a variância de  $\epsilon_i$  para todo  $i \in \{1, \dots, N\}$ . Por sua vez, sob normalidade dos erros, a região com  $(1 - \alpha)100\%$  de confiança para os parâmetros do modelo satisfaz

$$(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})^T \mathbf{X}^T \mathbf{X} (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}) \leq ps^2 F_{p,v,\alpha}, \quad (10)$$

em que  $s^2$  é uma estimativa de  $\sigma^2$  com um valor  $v$  fixo para os graus de liberdade e  $F_{p,v,\alpha}$  é o quantil de ordem  $(1 - \alpha)$  da distribuição  $F$  com  $p$  e  $v$  graus de liberdade.

Logo, dados  $s^2$  e  $v$ , todos os aspectos relativos à precisão dos estimadores estão ligados à matriz de delineamento. Além disso, é fácil verificar que ela resume-se a um conjunto de funções dos níveis das variáveis explanatórias e, portanto, depende do delineamento adotado  $\xi$ , ou seja,

$$\mathbf{X} = \mathbf{X}(\xi) = \begin{bmatrix} f_1(x_1) & f_2(x_1) & \cdots & f_p(x_1) \\ f_1(x_2) & f_2(x_2) & \cdots & f_p(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(x_N) & f_2(x_N) & \cdots & f_p(x_N) \end{bmatrix}, \quad (11)$$

em que  $f_j(x)$ , para todo  $j \in \{1, \dots, p\}$ , é especificada de acordo com o modelo. Por exemplo, para o modelo em (3) tem-se que  $f_1(x) = 1$ ,  $f_2(x) = x$ ,  $f_3(x) = x^2$ , ...,  $f_p(x) = x^{p-1}$ . Assim, o delineamento experimental é o fator determinante para a quantidade de informação que um experimento pode fornecer, dada uma estimativa *a priori* de  $\sigma^2$ .

O processo de obtenção dos estimadores de mínimos quadrados para modelos não-lineares segue os mesmos princípios vistos quando modelos lineares são considerados. Seja o modelo de regressão não-linear expresso por

$$\mathbf{y} = \boldsymbol{\eta}(\mathbf{x}, \boldsymbol{\theta}) + \boldsymbol{\epsilon}, \quad (12)$$

para o qual  $\boldsymbol{\eta}(\mathbf{x}, \boldsymbol{\theta})$  é o vetor esperado para a resposta dados o vetor de tratamentos  $\mathbf{x}$  e o conjunto de parâmetros  $\boldsymbol{\theta}$  e  $\boldsymbol{\epsilon}$  é o vetor de erros aleatórios, independentes e

homocedásticos com média zero. Primeiramente, é requerida a linearização da parte fixa do modelo, que pode ser realizada a partir de aproximação por séries de Taylor em torno de um conjunto de valores pontuais *a priori* para o vetor de parâmetros  $\boldsymbol{\theta}_0$ . Desconsiderando os termos de alta ordem e fazendo algumas manipulações, a expansão resulta em

$$\mathbf{y} - \boldsymbol{\eta}(\mathbf{x}, \boldsymbol{\theta}_0) = \mathbf{X}(\xi, \boldsymbol{\theta}_0)(\boldsymbol{\theta} - \boldsymbol{\theta}_0) + \boldsymbol{\epsilon}, \quad (13)$$

em que  $\mathbf{X}(\xi, \boldsymbol{\theta}_0)$  é dada pela matriz de derivadas parciais de  $\boldsymbol{\eta}(\mathbf{x}, \boldsymbol{\theta})$  em relação a  $\boldsymbol{\theta}$ , conhecida como sensibilidade dos parâmetros, sendo expressa por

$$\mathbf{X}(\xi, \boldsymbol{\theta}_0) = \left[ \left( \frac{\partial \boldsymbol{\eta}(\mathbf{x}, \boldsymbol{\theta})}{\partial \theta_1} \right)_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \quad \left( \frac{\partial \boldsymbol{\eta}(\mathbf{x}, \boldsymbol{\theta})}{\partial \theta_2} \right)_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \quad \dots \quad \left( \frac{\partial \boldsymbol{\eta}(\mathbf{x}, \boldsymbol{\theta})}{\partial \theta_p} \right)_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \right]. \quad (14)$$

Desta forma, a expressão em (13) pode ser reescrita como

$$\mathbf{y}^* = \mathbf{X}(\xi, \boldsymbol{\theta}_0)\boldsymbol{\theta}^* + \boldsymbol{\epsilon}, \quad (15)$$

em que  $\mathbf{y}^* = \mathbf{y} - \boldsymbol{\eta}(\mathbf{x}, \boldsymbol{\theta}_0)$  e  $\boldsymbol{\theta}^* = \boldsymbol{\theta} - \boldsymbol{\theta}_0$ . Então, a solução para  $\hat{\boldsymbol{\theta}}$  é obtida através do sistema

$$\hat{\boldsymbol{\theta}}^* = (\mathbf{X}^T(\xi, \boldsymbol{\theta}_0)\mathbf{X}(\xi, \boldsymbol{\theta}_0))^{-1}\mathbf{X}^T(\xi, \boldsymbol{\theta}_0)\mathbf{y}^*. \quad (16)$$

A solução do sistema de equações normais para modelos não-lineares exige o emprego de métodos iterativos, visto que não existe uma forma fechada para  $\hat{\boldsymbol{\theta}}$ . Portanto, a obtenção de estimativas dos parâmetros também depende do conhecimento *a priori* a respeito deles. Ainda, se um experimento é mal planejado, pode gerar dados com os quais o ajuste do modelo não-linear desejado seja impossível, devido a problemas de convergência.

Outro fato interessante é que, dado  $\boldsymbol{\theta}_0$ , a expressão em (13) é a de um modelo linear para o qual  $\mathbf{X}(\xi, \boldsymbol{\theta}_0)$  pode ser verificada como a matriz de delineamento e, assim, a matriz de informação normalizada por  $N$  pode ser expressa por

$$\mathbf{M}(\xi, \boldsymbol{\theta}_0) = \frac{\mathbf{X}^T(\xi, \boldsymbol{\theta}_0)\mathbf{X}(\xi, \boldsymbol{\theta}_0)}{N}. \quad (17)$$

Assintoticamente, a matriz de variâncias e covariâncias de  $\hat{\boldsymbol{\theta}}$  é proporcional ao inverso da matriz de informação, ou seja,

$$\mathbf{Var}(\hat{\boldsymbol{\theta}}) \propto \mathbf{M}^{-1}(\xi, \boldsymbol{\theta}_0). \quad (18)$$

Um aspecto importante da matriz de variâncias e covariâncias para os estimadores dos parâmetros no caso não-linear é a sua dependência em relação a  $\boldsymbol{\theta}_0$ . Enquanto nos modelos lineares as eficiências dos delineamentos não dependem dos valores dos parâmetros, no caso não-linear é necessária a utilização de todo conhecimento *a priori* a respeito deles para que a fase de delineamento seja bem sucedida.

A especificação do modelo com o qual o pesquisador deseja descrever os dados também é de grande importância para a busca de delineamentos experimentais ótimos, devido ao fato de que as técnicas mais usuais otimizam funções da matriz de variâncias e covariâncias para os estimadores de mínimos quadrados, como aquelas vistas em Dette et al. (2008) e Atkinson et al. (2007). No entanto, há técnicas de delineamento que visam modelos de regressão não paramétricos, como as apresentadas por Chan (1991), mas estão fora do escopo deste trabalho.

Neste trabalho foram considerados os modelos de Michaelis-Menten e de Hill, que foram escolhidos devido à ampla gama de aplicações que eles apresentam. Segundo Dette et al. (2008), o modelo de Michaelis-Menten é usado em um grande número de disciplinas e, dentre elas, pode-se citar a agronomia, a bioquímica, a biologia, a microbiologia, a toxicologia, a nutrição e a farmacologia como exemplos. A principal razão para isto é sua simplicidade e a habilidade de prover informação útil na descrição de sistemas biológicos complexos, como no estudo de fenômenos saturáveis em cinética enzimática. Sua aplicação mais comum é ajustar dados oriundos de algumas reações catalizadas por enzimas, descrevendo o comportamento da velocidade de reação  $\eta(x)$  em face à concentração de substrato  $x$ . À medida que a concentração de substrato aumenta, a taxa de crescimento da velocidade de reação diminui, até que a velocidade atinge seu valor máximo,  $\nu$ . A equação pode ser



expressa por

$$\eta(x) = \frac{\nu x}{\kappa + x}, \quad (19)$$

com  $x \geq 0$  e os parâmetros  $\kappa, \nu > 0$ , em que  $\kappa$  é dita constante de Michaelis-Menten e indica o valor de  $x$  para o qual a velocidade de reação esperada é a metade do valor da velocidade máxima. Para maior compreensão do leitor, diversas situações para este modelo são apresentadas na Figura 1.

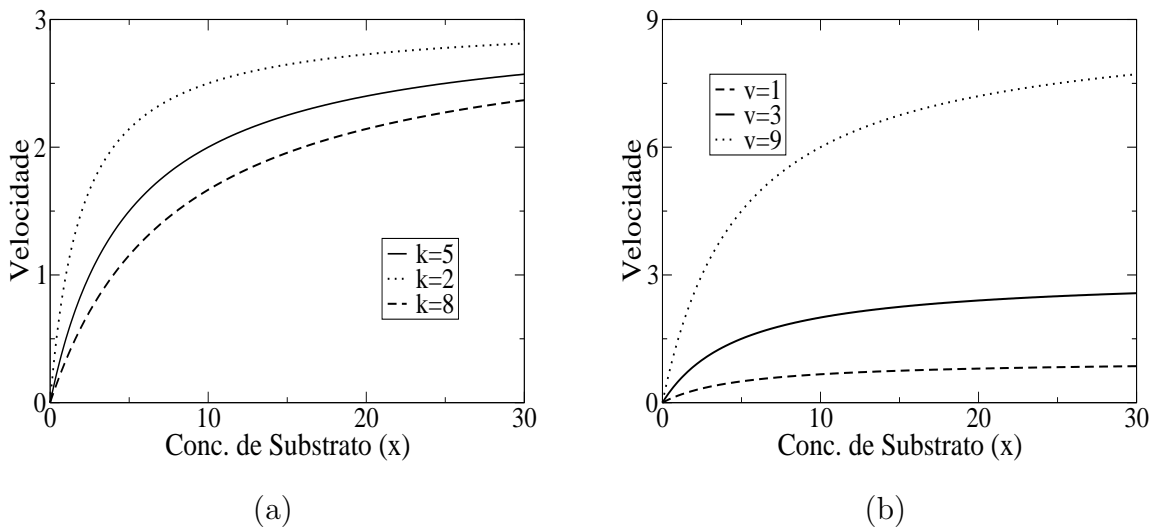


Figura 1: Equação de Michaelis-Menten para valores diferentes de  $\kappa$  e  $\nu$ : em (a),  $\nu$  é fixado em 3 e varia-se  $\kappa$ ; em (b),  $\kappa$  é fixado em 5 e varia-se  $\nu$ .

Uma suposição para o modelo de Michaelis-Menten é a não existência de cooperatividade entre os ligantes que participam da dinâmica. O modelo de Hill é uma alternativa em casos que tal suposição não é plausível, pois trata-se de uma generalização do modelo de Michaelis-Menten que permite a introdução da cooperação a partir da inclusão de um coeficiente empírico  $\gamma$ , que assume valores maiores que zero e é denominado coeficiente de Hill. Mais detalhes podem ser encontrados em Motulsky & Christopoulos (2004). Este modelo é expresso por

$$\eta(x) = \frac{\nu x^\gamma}{\kappa^\gamma + x^\gamma}. \quad (20)$$

Quando  $\gamma < 1$ , a cooperação é negativa, ou seja, há uma inibição na ação de um ligante devido ao aumento da atividade de outro. Quando  $\gamma = 1$ , o modelo se reduz

ao de Michaelis-Menten, ou seja, o processo não envolve inibição nem cooperação. Já se  $\gamma > 1$ , ocorre cooperatividade positiva. Neste caso, a curva assume a forma sigmóide. Uma breve apresentação deste modelo pode ser vista na Figura 2.

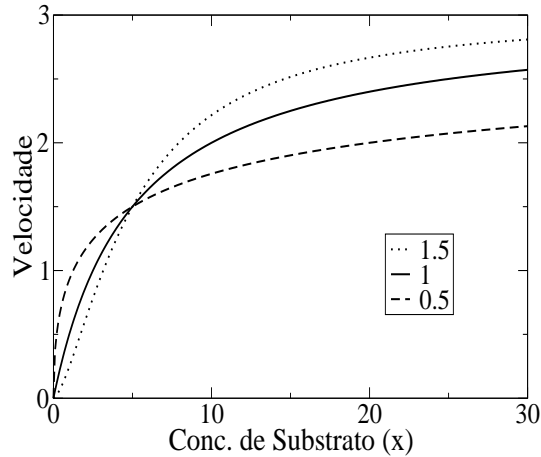


Figura 2: Equação de Hill com  $\kappa = 5$ ,  $\nu = 3$  e valores diferentes (0,5, 1,0 e 1,5) para o coeficiente de Hill.

### 3.5 Delineamentos exatos e contínuos

No estudo de delineamentos ótimos existem duas abordagens, a contínua ou aproximada e a exata ou discreta. Seja  $x_i$  com  $i \in \{1, \dots, n\}$  níveis distintos da variável explanatória, isto é, os tratamentos, aplicados a  $r_i$  unidades experimentais tais que  $\sum_{i=1}^n r_i = N$ , com  $x_i$  pertencente à região experimental  $\chi$  e  $N, r_i \in \mathbf{R}$ . Então, o delineamento experimental representado por

$$\xi = \left\{ \begin{array}{cccc} x_1 & x_2 & \dots & x_n \\ r_1/N & r_2/N & \dots & r_n/N \end{array} \right\},$$

é denominado delineamento exato ou discreto. Dado  $N$ , o delineamento ótimo é formado pelos pontos  $x_i$ 's e  $r_i$ 's que otimizam alguma propriedade de interesse sobre os estimadores dos parâmetros do modelo considerado. No entanto, tal problema de otimização é bastante difícil devido à natureza discreta das réplicas. Matematicamente,

o problema pode ser simplificado especificando o delineamento experimental  $\xi$  como contínuo ou aproximado, ou seja, uma medida de probabilidade que associa a cada ponto  $x_i$  da região experimental  $\chi$  um valor  $w_i$  pertencente ao intervalo  $[0, 1]$ . Por sua vez,  $w_i$  determina a probabilidade de que uma unidade experimental escolhida ao acaso receba  $x_i$  como tratamento e é denominado peso associado a este ponto. O delineamento experimental contínuo é representado por

$$\xi = \left\{ \begin{array}{cccc} x_1 & x_2 & \dots & x_n \\ w_1 & w_2 & \dots & w_n \end{array} \right\},$$

em que  $x_i$ , com  $i = \{1, 2, \dots, n\}$ , são os  $n$  pontos da região experimental para os quais  $w_i > 0$ , ditos pontos de suporte do delineamento. Além disso, a igualdade  $\sum_{i=1}^n w_i = 1$  é satisfeita.

Para a prática, todo delineamento é exato. Porém, em muitos casos é preferível realizar a busca por um delineamento ótimo contínuo e depois aproximá-lo para exato. A vantagem da abordagem contínua é a possibilidade de encontrar delineamentos ótimos de maneira analítica, a partir do teorema geral da equivalência de Kiefer & Wolfowitz (1960). No entanto, delineamentos exatos muito próximos do ótimo (ou ótimos) podem ser encontrados a partir de algoritmos heurísticos e métodos de programação linear e não-linear. Também é possível obter delineamentos exatos a partir de delineamentos ótimos contínuos. Basta fixar  $N$  e aproximar cada  $w_i N$  para o valor inteiro mais próximo, obtendo os  $r_i$ 's. Contudo, esta alternativa nem sempre produz delineamentos ótimos exatos.

### 3.6 Principais critérios para delineamento ótimo

Os critérios de otimalidade para delineamentos, genericamente denotados por  $\Psi\{\mathbf{M}(\xi)\}$ , são funções da matriz de informação ou de sua inversa e influenciam as medidas de precisão das estimativas dos parâmetros do modelo assumido. Por exemplo, um dos critérios mais utilizados é o D, que minimiza o volume do elipsóide de confiança dado um grau de liberdade fixo para a estimação da variância

do erro ou uma estimativa independente de  $\sigma^2$  a partir da minimização da variância generalizada dos estimadores dos parâmetros, ou seja, a partir da minimização do determinante da matriz de variâncias e covariâncias dos mesmos e, desta forma,

$$|\mathbf{M}^{-1}(\xi^*)| = \min |\mathbf{M}^{-1}(\xi)| \quad \forall \xi \in \Xi, \quad (21)$$

na qual  $\xi^*$  é o delineamento D-ótimo e  $\Xi$  é espaço de delineamento. Como o determinante da matriz de variâncias e covariâncias é inversamente proporcional ao determinante da matriz de informação, uma forma alternativa para o critério D é

$$|\mathbf{M}(\xi^*)| = \max |\mathbf{M}(\xi)| \quad \forall \xi \in \Xi, \quad (22)$$

desta maneira pode-se evitar o processo de inversão de matrizes, que é computacionalmente custoso. Em geral, maximiza-se

$$\Psi\{\mathbf{M}(\xi)\} = \ln |\mathbf{M}(\xi)|. \quad (23)$$

A razão para tal é que a introdução do logaritmo natural transforma a função-critério numa função côncava num espaço compacto e convexo  $\Xi$  e, desta maneira, qualquer ótimo encontrado será global ao invés de local.

Quando o interesse é concentrado apenas nas variâncias dos estimadores dos parâmetros e não importam muito inferências globais a respeito deles, um critério de otimalidade mais adequado é o A. Tal critério minimiza o traço da matriz de variâncias e covariâncias, ou seja, minimiza a soma das variâncias, equivalente a

$$\text{tr} \mathbf{M}^{-1}(\xi^*) = \min \text{tr} \mathbf{M}^{-1}(\xi) \quad \forall \xi \in \Xi, \quad (24)$$

em  $\xi^*$  é o delineamento A-ótimo. Uma característica deste critério que tem contribuído para a sua baixa popularidade é a dependência do delineamento ótimo em relação à escala de medida das variáveis quantitativas.

Outro critério muito importante é o G, no qual minimiza-se o máximo da variância da resposta predita sobre a região experimental. Na verdade, há vários outros critérios e um bom delineamento não deve atender a um somente, mas, pelo

contrário, deve atender de modo eficiente todos os interesses do pesquisador. No entanto, em geral, um delineamento ótimo segundo um critério pode ser sub-ótimo com relação a outro. Uma alternativa seria a construção de delineamentos transigentes, ou seja, comprometidos com várias propriedades de interesse. Para tanto, pode-se utilizar um critério composto, ou seja, uma combinação linear convexa de critérios individuais, cada um com seu peso, sendo tais pesos determinados pelas prioridades da pesquisa. Dessa maneira, o critério composto é descrito por

$$\Psi(\xi) = \sum_{i=1}^h \lambda_i \Psi_i\{\mathbf{M}_i(\xi)\}, \quad (25)$$

em que  $h$  é o número de propriedades de interesse relativas à(s) matriz(es) de informação,  $\lambda_i$  é um peso não-negativo associado a cada propriedade a partir das prioridades de pesquisa e  $\Psi_i\{\mathbf{M}_i(\xi)\}$  é a função-critério associada a cada uma delas.

Existe um grande número de critérios propostos na bibliografia, como os que podem ser vistos em Atkinson et al. (2007), Pinto & Ponce de Leon (2006), Pukelsheim (1993) e Pronzato & Walter (1985).

### 3.7 Abordagens para a construção de delineamentos ótimos para modelos não-lineares

Atkinson et al. (2007) discute que a necessidade de incorporação de valores *a priori* para representar os parâmetros na construção de delineamentos ótimos possibilita o desenvolvimento de diferentes abordagens. Caso seja adotado um vetor pontual  $\boldsymbol{\theta}_0$  como informação *a priori*, serão obtidos delineamentos que são ótimos apenas se  $\boldsymbol{\theta}$  coincidir com  $\boldsymbol{\theta}_0$ . Portanto, tais delineamentos são ditos localmente ótimos.

Uma abordagem mais robusta é obtida quando, ao invés de um valor pontual, é admitida uma distribuição de probabilidade ou uma função de densidade *a priori*, denotada por  $p(\boldsymbol{\theta})$ , para representar os parâmetros na construção do delineamento. Assim, uma região maior do espaço paramétrico  $\Theta$  é contemplada. Tais delineamentos são ditos bayesianos quando propriedades da matriz de variâncias e

covariâncias *a posteriori* são otimizadas. Quando, no entanto, a informação *a priori* é usada na fase de planejamento mas a matriz de variâncias é obtida pela versão frequentista de análise de dados, eles são ditos pseudo-bayesianos.

Outra abordagem é a *maximin*, que visa obter o melhor delineamento para o pior caso. Em resumo, trata-se da busca do ótimo para o ponto do espaço paramétrico no qual é verificada as estimativas mais imprecisas. Geralmente produz delineamentos muito eficientes para os limites da região experimental e bem diferentes dos oriundos de outras abordagens.

Neste trabalho, apenas os delineamentos localmente ótimos e ótimos pseudo-bayesianos foram considerados.

### 3.8 O critério D em situações não-lineares

Segundo Atkinson et al. (2007), o critério de delineamento mais importante é o D, devido a imensa quantidade de aplicações, mas, também, devido a facilidade de trabalhar com ele. Para o caso não-linear, pode-se descrevê-lo da seguinte maneira: seja

$$\mathbf{X}(\xi, \boldsymbol{\theta}_0) = [\mathbf{f}_1(\mathbf{x}, \boldsymbol{\theta}_0) \quad \mathbf{f}_2(\mathbf{x}, \boldsymbol{\theta}_0) \quad \cdots \quad \mathbf{f}_p(\mathbf{x}, \boldsymbol{\theta}_0)], \quad (26)$$

em que  $\mathbf{x}^T = [x_1 \dots x_N]$ ; seja  $N$  o número de unidades experimentais e seja

$$\mathbf{f}_j = \mathbf{f}_j(\mathbf{x}, \boldsymbol{\theta}_0) = \left[ \frac{\partial \boldsymbol{\eta}(\mathbf{x}, \boldsymbol{\theta})}{\partial \theta_j} \right]_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \quad (27)$$

para  $j = \{1, 2, \dots, p\}$ . Então, a função-critério de D-otimalidade é expressa por

$$\Psi\{\mathcal{M}(\xi, \boldsymbol{\theta}_0)\} = \ln \left\{ \left| \begin{array}{cccc} \frac{\mathbf{f}_1^T \mathbf{f}_1}{N} & \frac{\mathbf{f}_1^T \mathbf{f}_2}{N} & \cdots & \frac{\mathbf{f}_1^T \mathbf{f}_p}{N} \\ \frac{\mathbf{f}_2^T \mathbf{f}_1}{N} & \frac{\mathbf{f}_2^T \mathbf{f}_2}{N} & \cdots & \frac{\mathbf{f}_2^T \mathbf{f}_p}{N} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\mathbf{f}_p^T \mathbf{f}_1}{N} & \frac{\mathbf{f}_p^T \mathbf{f}_2}{N} & \cdots & \frac{\mathbf{f}_p^T \mathbf{f}_p}{N} \end{array} \right| \right\}. \quad (28)$$

A introdução de incerteza sobre os valores dos parâmetros permite o uso de várias definições diferentes para a função-critério de D-otimalidade. No

caso pseudo-Bayesiano, o delineamento que minimiza o valor esperado do volume do elipsóide de confiança para os parâmetros não é necessariamente o mesmo que maximiza o valor esperado do determinante da matriz de informação que, por sua vez, não é necessariamente o mesmo que maximiza a informação esperada a respeito dos parâmetros.

Segundo Pronzato & Walter (1985), para obter o menor volume esperado para o elipsóide de confiança deve-se minimizar  $E_{\Theta} |\mathbf{M}^{-1}(\xi, \boldsymbol{\theta})|$ , dado o espaço paramétrico  $\Theta \subset \mathcal{R}^p$ . Por sua vez, para se obter o maior valor esperado para o determinante da matriz de informação, basta maximizar  $E_{\Theta} |\mathbf{M}(\xi, \boldsymbol{\theta})|$ . Pronzato & Walter (1985) denominam esta última abordagem como ED-otimalidade. Outra possibilidade, que pode ser vista em Atkinson et al. (2007) e em Chaloner & Verdinelli (1995), é maximizar a informação de Shannon esperada a partir da otimização de  $E_{\Theta} \ln |\mathbf{M}(\xi, \boldsymbol{\theta})|$ , isto é,

$$\Psi\{\mathbf{M}(\xi)\} = \int_{\Theta} \Psi\{\mathbf{M}(\xi, \boldsymbol{\theta})\} p(\boldsymbol{\theta}) d\boldsymbol{\theta}. \quad (29)$$

Apenas a definição de D-otimalidade expressa em (29) é abordada neste trabalho.

### 3.9 Método de Monte Carlo para obter o valor aproximado da função-critério

Muitas vezes não é possível calcular analiticamente o valor da função-critério para as abordagens bayesiana e pseudo-bayesiana devido à complexidade da integral envolvida. Há vários métodos para obter um valor aproximado para tais integrais, sendo que o mais comumente utilizado é o método de Monte Carlo, como verifica-se em Atkinson et al. (2007) e Graham et al. (2005). Neste caso trata-se de amostrar valores da distribuição *a priori* de  $\boldsymbol{\theta}$  e calcular a média da função-critério para os conjuntos paramétricos amostrados, ou seja,

$$\Psi\{\mathbf{M}(\xi)\} \approx \frac{1}{a} \sum_{i=1}^a \Psi\{\mathbf{M}(\xi, \boldsymbol{\theta}_i)\}, \quad (30)$$

em que  $a$  é o número de conjuntos paramétricos amostrados (ou tamanho da amostra) e  $\theta_i$  é o  $i$ -ésimo destes conjuntos.

### 3.10 A variância padronizada

A variância para a predição da resposta, por uma equação ou modelo não-linear, em um ponto  $x \in \mathcal{X}$  e para um delineamento  $\xi$  é expressa por

$$\text{var}\{\hat{y}(x, \theta_0)\} = \sigma^2 \mathbf{f}^T(x, \theta_0) \mathbf{M}^{-1}(\xi, \theta_0) \mathbf{f}(x, \theta_0), \quad (31)$$

em que  $\mathbf{f}^T(x, \theta_0) = \{f_j(x, \theta_0)\}_{j=1}^p$ . Por sua vez, a variância padronizada para a resposta predita desconsidera a influência da constante  $\sigma^2$  e é dada por

$$d(x, \xi, \theta_0) = \mathbf{f}^T(x, \theta_0) \mathbf{M}^{-1}(\xi, \theta_0) \mathbf{f}(x, \theta_0). \quad (32)$$

Sua importância reside no fato de que, pelo Teorema geral da equivalência de Kiefer & Wolfowitz (1960), para a D-otimalidade e no contexto contínuo é verificada a desigualdade

$$d(x, \xi^*, \theta_0) \leq p, \quad (33)$$

sendo que  $d(x, \xi^*, \theta_0) = p$  ocorre quando  $x$  é ponto de suporte de  $\xi^*$ . Dessa maneira, a variância padronizada é uma ferramenta útil para a demonstração da D-otimalidade. Tal relação não é verdadeira para delineamentos D-ótimos exatos, mas os resultados obtidos pelas duas abordagens serão próximos se  $N$  for razoavelmente grande. Agora, se o D-ótimo contínuo for balanceado - isto é, apresentar o mesmo peso para todos os  $n$  pontos de suporte - e  $N$  for múltiplo  $n$ , então, os resultados devem coincidir.

No caso pseudo-Bayesiano a variância padronizada é expressa por

$$d(x, \xi) = \int_{\Theta} d(x, \xi, \theta) p(\theta) d\theta \quad (34)$$

e, portanto, métodos de aproximação também são usuais para seu cálculo.



### 3.11 Algoritmos para busca de delineamentos ótimos exatos

Algoritmos desenvolvidos para encontrar  $\xi^*$  avaliam um grande número de delineamentos antes de seu ponto de parada e, portanto, uma computação de alta performance é necessária. É por isso que, apesar de ser proposta por Kiefer em 1959, a teoria de delineamentos ótimos demorou a adquirir adeptos e só teve um desenvolvimento adequado após o grande avanço computacional ocorrido nas últimas décadas.

Encontrar delineamentos ótimos de maneira analítica é possível, porém, é muito difícil. Quando encontra-se na bibliografia um delineamento ótimo obtido de maneira analítica, geralmente este restringe-se a uma quantidade de pontos de suporte fixa, como os que mostram Box & Lucas (1959) e Dette et al. (2008). Ainda, se exato, tem seu tamanho como múltiplo do número de pontos de suporte de um caso contínuo. Além disso, delineamentos ótimos bayesianos ou pseudo-bayesianos encontrados analiticamente são raríssimos, obtidos em situações que envolvem poucos parâmetros e distribuições *a priori* discretas ou Uniformes para facilitar o cálculo do critério.

Desta forma, métodos de programação linear e não-linear, bem como algoritmos heurísticos, são empregados em larga escala na otimização de delineamentos exatos. A maioria deles não garante a otimalidade dos resultados, mas produz delineamentos ótimos ou muito próximos do ótimo com grande reprodutibilidade.

Neste trabalho, os algoritmos genético e *exchange* foram implementados e comparados. Tais algoritmos são métodos computacionais heurísticos amplamente aplicados na busca de delineamentos ótimos exatos. Uma breve introdução a respeito deles é apresentada nas próximas seções.

#### 3.11.1 Algoritmos genéticos

Numa visão bem geral, os algoritmos genéticos (AG's) são uma classe de modelos computacionais baseados nas leis da seleção natural e da genética, que buscam soluções ótimas a partir de uma população de soluções potenciais. Foram

inicialmente propostos por Holland (1975), mas ficaram popularizados com o auxílio de Goldberg (1989). Podemos descrevê-los através dos seguintes passos:

1. Primeiramente, uma população de soluções potenciais (cromossomos) para o problema é criada, geralmente de maneira aleatória, em que cada cromossomo representa um ponto dentro do espaço de soluções. Eles são codificados por sequências de valores binários, inteiros ou reais, que guardam sua informação e permitem que lhes sejam atribuídos valores de aptidão, através dos quais os cromossomos serão selecionados para constituírem uma população intermediária ou para serem descartados.
2. A população intermediária, por sua vez, aloca os cromossomos aptos a participarem do processo de reprodução, ou seja, abriga um percentual dos melhores cromossomos da população inicial que darão origem a uma nova geração. A nova geração é composta por cromossomos filhos da população intermediária, originados a partir da influência da mutação e da recombinação gênica. A frequência de filhos que cada cromossomo originará pode depender do valor da sua função de aptidão ou da sua posição numa classificação ou *rank*.

Em alguns casos, é desejado manter intactos os melhores indivíduos da população intermediária, visto que o processo de reprodução pode construir soluções melhores, mas pode também destruir soluções que já são boas. Esta conduta recebe o nome de elitismo e é utilizada para melhorar o tempo de convergência dos algoritmos, contudo, pode conduzir os resultados a ótimos locais ou ocasionar uma parada antecipada do procedimento devido à diminuição da diversidade gênica apresentada pela população.

3. Quando selecionado para a reprodução, o cromossomo tem sua informação combinada com a de outro cromossomo (recombinação gênica da reprodução sexuada) para gerar cromossomos filhos. A recombinação pode ocorrer com a troca de toda a informação a partir de um ponto aleatório, ou a troca da informação contida em sítios específicos e de muitas outras maneiras.

4. Além disso, com uma probabilidade pequena, sítios dos filhos (genes) podem ser trocados com valores candidatos (valores potenciais para sítios da solução), representando o processo de mutação. O papel da mutação é fundamental para o algoritmo, pois ela possibilita a introdução de valores que não estavam presentes na população inicial ou que foram eliminados prematuramente devido à seleção e, dessa maneira, assegura a possibilidade de ocorrência de qualquer ponto do espaço de soluções.
5. Ao final da reprodução, verifica-se uma população de mesmo tamanho da inicial, mas, quase sempre, com indivíduos mais aptos que os da anterior. Então, o processo é repetido inúmeras vezes, até que um critério de parada seja satisfeito. Critérios de parada comuns são: parar quando um número pré-fixado de gerações apresentarem o mesmo cromossomo como o mais apto; parar quando um determinado percentual da população apresentar o mesmo valor de aptidão; parar quando não houver grande diferença entre alguma estatística (média ou mediana, por exemplo) calculada para a aptidão dos cromossomos de duas gerações subsequentes; ou parar após um número de gerações pré-determinado.

A grande vantagem dos algoritmos genéticos em relação a outros métodos de otimização é o uso de uma população de soluções potenciais no início do processo, ao invés de um único ponto de partida. Isto evita que as soluções convirjam para ótimos locais e tornam-no um método mais robusto.

Porém, para utilizá-lo é importante estar atento a algumas características do problema. Primeiro, a região de busca deve estar limitada a uma faixa de valores. Segundo, é necessária a determinação de uma função de aptidão capaz de servir como métrica para indicar quão boa é uma solução. E, por fim, o pesquisador deve encontrar uma codificação simples do problema, que possibilite a implementação computacional.

Além disso, seu desempenho está fortemente ligado à escolha dos parâmetros genéticos. Por exemplo, a escolha de uma população inicial pequena faz

com que o método não garanta a cobertura do espaço de busca, levando a soluções locais. Por sua vez, uma população inicial grande melhora o processo de busca e traz mais credibilidade aos resultados, porém, exige mais recursos computacionais e prolonga o tempo de convergência.

Outro parâmetro é a taxa de recombinação. Se alta, faz com que estruturas novas apareçam com maior rapidez, mas servirá para destruir soluções boas ao final do processo e impedirá que o algoritmo satisfaça o critério de parada. Se baixa, toda introdução de informação nova ficará a cargo da mutação e o processo de busca será extremamente lento.

Já a taxa de mutação é importante para impedir a parada antecipada. Se isto ocorrer, teremos por ótimos apenas soluções que podem ser construídas com valores presentes na população, mas, por maior que ela seja, na prática nenhuma é capaz de cobrir todo o espaço de busca. No entanto, se a taxa de mutação for alta, não poderemos verificar uma homogeneidade na população que caracterize um critério de parada e soluções boas serão destruídas a todo momento.

Por último, a escolha do intervalo de geração necessita de cuidado. O intervalo de geração é o percentual de cada geração escolhido para a fase de reprodução. Se este percentual for alto, cromossomos com baixa performance serão replicados e passarão sua informação adiante, atrasando o processo de convergência. Porém, se for baixo, selecionará poucos cromossomos e diminuirá a multiplicidade de valores dentro do sistema, necessária para a construção de soluções melhores.

### 3.11.2 Algoritmos tipo *exchange*

O algoritmo *exchange* é um método heurístico que busca por soluções ótimas a partir de uma sistemática de trocas. Em resumo, consiste na construção de combinações ótimas tendo por base uma configuração inicial, que é melhorada iterativamente através de uma sequência de trocas de seus pontos por pontos de um conjunto de candidatos.

Segundo Woods (2005), os algoritmos tipo *exchange* constituem a

técnica mais comum de busca de delineamentos ótimos exatos. Para executá-los, o primeiro passo é discretizar a região experimental numa lista de pontos que formarão o conjunto de candidatos. Logo em seguida, um delineamento é escolhido como suporte para a busca e seus pontos começam a ser trocados por pontos do conjunto de candidatos.

Se a troca melhora o valor do critério, ela é aceita. Caso contrário, a troca é rejeitada. Todos os pontos do conjunto de candidatos são avaliados e, por fim, permanece no delineamento aquele que propicia o melhor valor de critério. O procedimento é repetido para todo ponto do delineamento corrente e o algoritmo cessa quando não é possível realizar mais trocas.

O sucesso ou fracasso do algoritmo *exchange* depende do delineamento inicial e do caminho que ele irá percorrer no espaço de busca. Ótimos locais podem ser obtidos como resultado e, por isso, o algoritmo deve ser executado diversas vezes e com delineamentos iniciais diferentes para aumentar a probabilidade de encontrar o ótimo global.

## 4 METODOLOGIA

A primeira etapa deste trabalho consistiu na construção e avaliação de métodos para a busca de delineamentos ótimos exatos em modelo não-lineares. Para tanto, os algoritmos genético e *exchange* foram implementados em linguagem de programação C e executados num microcomputador convencional com 2 GB de memória RAM e processador *Pentium* T4300.

O algoritmo *exchange* exige uma região experimental discretizada enquanto que para o genético é possível realizar a busca numa região experimental contínua. Porém, para tornar plausível a comparação entre métodos, ela foi discretizada em ambos algoritmos. Visando a obtenção de delineamentos para os modelos de Michaelis-Menten e de Hill, considerou-se uma grade com 600 valores igualmente espaçados, pertencentes ao intervalo de concentração indo de 0,05 *mM* a 30,00 *mM*. Portanto, o conjunto de pontos candidatos utilizado foi  $P_c = \{0,05; 0,10; 0,15; \dots; 30,0\}$ . A seguir é apresentada uma breve descrição de cada algoritmo implementado.

(a) Algoritmo genético:

Na implementação do algoritmo genético, utilizou-se como condição inicial uma população aleatória contendo 100 delineamentos, cujos pontos foram obtidos do conjunto de candidatos a partir de uma distribuição Uniforme discreta. Cada um deles foi avaliado e ordenado de acordo com o valor da função-critério de D-otimalidade, sendo que os 25% mais aptos foram alocados numa população intermediária para participarem do processo de reprodução. A cada delineamento da população intermediária foi atribuído um valor de classificação, ou

*rank*, da seguinte maneira: seja  $s$  o número de delineamentos que a população intermediária possui, então o melhor delineamento recebe o valor  $s$ ; já o segundo melhor recebe  $s - 1$ ; por sua vez, o terceiro melhor recebe  $s - 2$  e, assim, por diante.

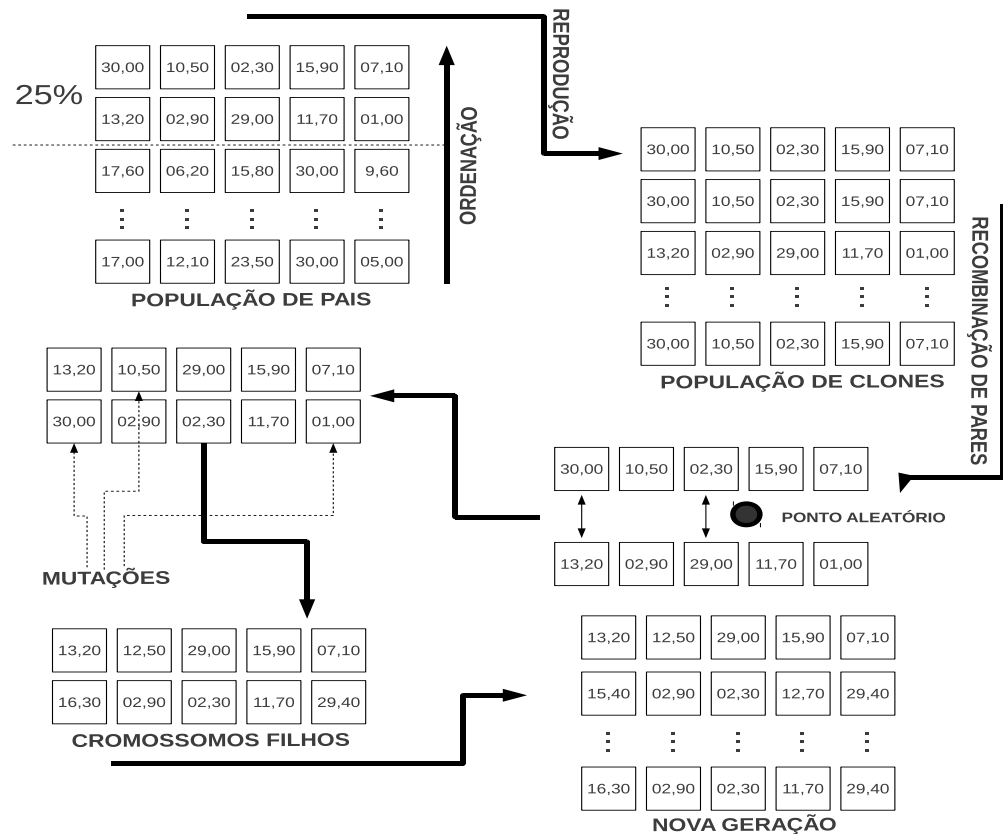


Figura 3: Desenho esquemático utilizado para exemplificar o procedimento realizado pelo algoritmo genético em cada iteração.

Antes de realizar os processos de mutação e recombinação, a população atual foi substituída por clones dos delineamentos da população intermediária. Para escolher qual delineamento seria replicado em cada instante, utilizou-se o valor de *rank* de cada delineamento dividido pela soma de todos os *rank*s como probabilidade de replicação.

A recombinação (ou cruzamento) foi realizada a partir de pares de deline-

mentos selecionados aleatoriamente (sem reposição) da população de clones e de pontos de troca (indo de 2 a  $N - 1$ ) escolhidos através de sorteio. Então, pares de genes (ou seja, pontos dos delineamentos) posicionados após o valor escolhido foram trocados com uma probabilidade de recombinação igual a 0,5. Além disso, cada gene poderia ter seu valor substituído por um ponto do conjunto candidatos  $P_c$ , obtido de maneira aleatória, com uma probabilidade de mutação igual a 0,1. Tal processo foi realizado para todo delineamento da população de clones e, como resultado, foi obtida uma nova geração. Para maior compreensão, veja a Figura 3.

Os operadores de seleção e reprodução foram repetidos até que a população atual apresentasse seus  $\delta\%$  ( $\delta = 60$  para  $N = 8$  e  $\delta = 45$  para  $N = 12$ ) melhores delineamentos praticamente empatados quanto ao valor do critério D. Para tanto, foram considerados empatados os delineamentos para os quais a diferença no valor do critério não excedeu  $10^{-15}$ . Por fim, o melhor delineamento da população final foi considerado como o delineamento D-ótimo exato para o caso estudado. Este algoritmo, implementado em linguagem C e incluindo comentários, está apresentado no anexo A.

(b) Algoritmo *exchange*:

O algoritmo *exchange* utilizado neste trabalho foi uma versão adaptada do algoritmo proposto por Trinca & Gilmour (2001), introduzido para encontrar delineamentos D-ótimos pseudo-Bayesianos para modelos não-lineares representados por polinômios fracionários.

Em sua implementação, o primeiro passo foi obter um delineamento inicial de maneira aleatória a partir do conjunto de pontos candidatos. Em seguida, os pontos deste delineamento foram ordenados do menor para o maior e o valor de seu critério D foi calculado.

Então, o primeiro ponto foi substituído, sequencialmente, por cada um dos pontos candidatos e o valor do critério D foi calculado para cada troca. Por



fim, permaneceu em sua posição o ponto que determinou o melhor valor do critério D.

O mesmo procedimento foi realizado para todos os outros pontos do delineamento, excetuando-se aqueles que possuíam o mesmo valor do que acabara de entrar. É fácil entender a razão para tal: como o valor havia sido aceito num passo anterior, garantia o melhor critério D quando comparado aos outros pontos candidatos, logo não poderia sair do delineamento.

Após avaliar todos os pontos do delineamento corrente, seus valores foram reordenados e todo o procedimento foi repetido. Porém, quando numa iteração nenhuma troca foi aceita, o algoritmo cessou. Sua implementação em linguagem C com comentários está apresentada no anexo B.

A outra etapa deste trabalho consistiu na busca por delineamentos localmente D-ótimos exatos e D-ótimos exatos pseudo-Bayesianos para os modelos de Michaelis-Menten e de Hill. O tamanho do experimento  $N$  foi fixado em 8 para o modelo de Michaelis-Menten e em 12 para o modelo de Hill. Também foram obtidos delineamentos compostos exatos com  $N = 12$ . Como informação *a priori* para representar os parâmetros do modelo foram utilizadas as estimativas pontuais obtidas por Condomina et al. (2002).

Para delinear possíveis experimentos futuros em aplicações relacionadas com os estudos de Condomina et al. (2002), considerando-se uma variedade de situações, executou-se por uma vez o algoritmo genético e por três vezes o algoritmo *exchange*. O algoritmo genético não foi repetido pelo fato de sua busca ser baseada numa população de soluções potenciais e, dessa maneira, ter menor possibilidade de atingir um ótimo local. Apenas o melhor delineamento encontrado em cada situação está apresentado na seção 5.

Um aspecto crucial para a implementação dos algoritmos foi a computação do valor da função-critério de D-otimalidade. Para a equação de Michaelis-Menten, que é um modelo de regressão não-linear cujo vetor resposta esperado dado

o vetor de parâmetros  $\boldsymbol{\theta}^T = [\kappa, \nu]$  é expresso por

$$\boldsymbol{\eta}(\mathbf{x}, \boldsymbol{\theta}) = \frac{\nu \mathbf{x}}{\kappa + \mathbf{x}}, \quad (35)$$

as derivadas parciais avaliadas no ponto  $\boldsymbol{\theta}_0^T = [\kappa_0 \ \nu_0]$  do espaço paramétrico são

$$\mathbf{f}_1 = \left( \frac{\partial \boldsymbol{\eta}(\mathbf{x}, \boldsymbol{\theta})}{\partial \nu} \right)_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} = \frac{\mathbf{x}}{(\kappa_0 + \mathbf{x})} \quad \text{e} \quad \mathbf{f}_2 = \left( \frac{\partial \boldsymbol{\eta}(\mathbf{x}, \boldsymbol{\theta})}{\partial \kappa} \right)_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} = -\frac{\nu_0 \mathbf{x}}{(\kappa_0 + \mathbf{x})^2}. \quad (36)$$

Logo, a função-critério de D-otimalidade local para delineamentos exatos é expressa por

$$\begin{aligned} \Psi\{\mathbf{M}(\xi, \boldsymbol{\theta}_0)\} &= \ln \left\{ \frac{\nu_0^2}{N^2} \left[ \left( \sum_{i=1}^N \frac{x_i^2}{(\kappa_0 + x_i)^2} \right) \left( \sum_{i=1}^N \frac{x_i^2}{(\kappa_0 + x_i)^4} \right) - \left( \sum_{i=1}^N \frac{x_i^2}{(\kappa_0 + x_i)^3} \right)^2 \right] \right\} \\ &= C_1 + \zeta(\xi, \kappa_0), \end{aligned} \quad (37)$$

em que  $C_1 = \ln(\nu_0^2/N^2) \in \Re$ . Já para a busca de delineamentos pseudo-Bayesianos, utiliza-se a expressão

$$\begin{aligned} \Psi\{\mathbf{M}(\xi)\} &= \int_{\boldsymbol{\Theta}} \ln \left\{ \frac{\nu^2}{N^2} \left[ \left( \sum_{i=1}^N \frac{x_i^2}{(\kappa + x_i)^2} \right) \left( \sum_{i=1}^N \frac{x_i^2}{(\kappa + x_i)^4} \right) - \left( \sum_{i=1}^N \frac{x_i^2}{(\kappa + x_i)^3} \right)^2 \right] \right\} p(\boldsymbol{\theta}) d\boldsymbol{\theta} \\ &= C_2 + \int_{\boldsymbol{\Theta}} \zeta(\xi, \kappa) p(\boldsymbol{\theta}) d\boldsymbol{\theta}, \end{aligned} \quad (38)$$

em que  $C_2 = \int_{\boldsymbol{\Theta}} \ln(\nu^2/N^2) p(\boldsymbol{\theta}) d\boldsymbol{\theta} \in \Re$ . Desta maneira, o argumento  $\xi^*$  que maximiza  $\Psi\{\mathbf{M}(\xi, \boldsymbol{\theta}_0)\}$  ou  $\Psi\{\mathbf{M}(\xi)\}$  sob o espaço de delineamento  $\Xi$  é o mesmo que maximiza  $\zeta(\xi, \kappa_0)$  ou  $\int_{\boldsymbol{\Theta}} \zeta(\xi, \kappa) p(\boldsymbol{\theta}) d\boldsymbol{\theta}$ . Portanto, a informação *a priori* a respeito do parâmetro  $\nu$  é necessária para o cálculo do critério D, porém, não é necessária para a busca do delineamento D-ótimo  $\xi^*$ .

Por sua vez, a equação de Hill é um modelo de regressão não-linear cujo vetor de resposta esperado dado o vetor de parâmetros  $\boldsymbol{\theta}^T = [\kappa, \nu, \gamma]$  é expresso por

$$\boldsymbol{\eta}(\mathbf{x}, \boldsymbol{\theta}) = \frac{\nu \mathbf{x}^\gamma}{\kappa^\gamma + \mathbf{x}^\gamma}. \quad (39)$$

Portanto, as derivadas parciais avaliadas no ponto  $\boldsymbol{\theta}_0^T = [\kappa_0 \ \nu_0 \ \gamma_0]$  do espaço paramétrico são

$$\mathbf{f}_1 = \left( \frac{\partial \boldsymbol{\eta}(\mathbf{x}, \boldsymbol{\theta})}{\partial \nu} \right)_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} = \frac{\mathbf{x}^{\gamma_0}}{(\kappa_0^{\gamma_0} + \mathbf{x}^{\gamma_0})},$$

$$\begin{aligned} \mathbf{f}_2 &= \left( \frac{\partial \eta(\mathbf{x}, \boldsymbol{\theta})}{\partial \kappa} \right)_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} = -\frac{\nu_0 \gamma_0 \kappa_0^{\gamma_0-1} \mathbf{x}^{\gamma_0}}{(\kappa_0^{\gamma_0} + \mathbf{x}^{\gamma_0})^2}, \\ \mathbf{f}_3 &= \left( \frac{\partial \eta(\mathbf{x}, \boldsymbol{\theta})}{\partial \gamma} \right)_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} = \frac{\nu_0 \kappa_0^{\gamma_0} \mathbf{x}^{\gamma_0} \ln(\mathbf{x}/\kappa_0)}{(\kappa_0^{\gamma_0} + \mathbf{x}^{\gamma_0})^2}. \end{aligned} \quad (40)$$

Logo, a função-critério de D-otimalidade para delineamentos exatos é expressa por

$$\Psi\{\mathbf{M}(\xi, \boldsymbol{\theta}_0)\} = \ln \left[ \frac{\nu^4 \gamma^2 \kappa^{4\gamma-2}}{N^3} (ADF + 2BCE - C^2D - B^2F - E^2A) \right]_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}, \quad (41)$$

em que

$$\begin{aligned} A &= \sum_{i=1}^N \frac{\mathbf{x}^{2\gamma}}{(\kappa^\gamma + \mathbf{x}^\gamma)^2}, \quad B = -\sum_{i=1}^N \frac{\mathbf{x}^{2\gamma}}{(\kappa^\gamma + \mathbf{x}^\gamma)^3}, \quad C = \sum_{i=1}^N \frac{\mathbf{x}^{2\gamma} \ln(\mathbf{x}/\kappa)}{(\kappa^\gamma + \mathbf{x}^\gamma)^3}, \\ D &= \sum_{i=1}^N \frac{\mathbf{x}^{2\gamma}}{(\kappa^\gamma + \mathbf{x}^\gamma)^4}, \quad E = -\sum_{i=1}^N \frac{\mathbf{x}^{2\gamma} \ln(\mathbf{x}/\kappa)}{(\kappa^\gamma + \mathbf{x}^\gamma)^4} \quad \text{e} \quad F = \sum_{i=1}^N \frac{\mathbf{x}^{2\gamma} \ln^2(\mathbf{x}/\kappa)}{(\kappa^\gamma + \mathbf{x}^\gamma)^4}. \end{aligned}$$

Para o caso pseudo-Bayesiano, a função-critério de D-otimalidade é expressa por

$$\Psi\{\mathbf{M}(\xi)\} = \int_{\Theta} \ln \left[ \frac{\nu^4 \gamma^2 \kappa^{4\gamma-2}}{N^3} (ADF + 2BCE - C^2D - B^2F - E^2A) \right] p(\boldsymbol{\theta}) d\boldsymbol{\theta}. \quad (42)$$

Novamente  $\nu$  (ou  $\nu_0$ ) aparece linearmente na função-critério, sendo influente em seu valor mas de nenhuma importância para a busca do delineamento D-ótimo.

Para a exploração de propriedades ótimas dos delineamentos encontrados foi relevante o cálculo da variância padronizada, expressa por

$$d(x, \xi, \boldsymbol{\theta}_0) = \mathbf{f}^T(x, \boldsymbol{\theta}_0) \mathbf{M}^{-1}(\xi, \boldsymbol{\theta}_0) \mathbf{f}(x, \boldsymbol{\theta}_0). \quad (43)$$

Para o modelo de Michalis-Menten, dada a matriz de informação

$$\mathbf{M}(\xi, \boldsymbol{\theta}_0) = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix}_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}, \quad (44)$$

em que  $M_{12} = M_{21}$ , a variância padronizada fica expressa como

$$d(x, \xi, \boldsymbol{\theta}_0) = \frac{M_{22}f_1^2(x, \boldsymbol{\theta}_0) - 2M_{12}f_1(x, \boldsymbol{\theta}_0)f_2(x, \boldsymbol{\theta}_0) + M_{11}f_2^2(x, \boldsymbol{\theta}_0)}{|\mathbf{M}(\xi, \boldsymbol{\theta}_0)|}. \quad (45)$$

Para o caso pseudo-Bayesiano, temos que  $d(x, \xi)$  é dada por

$$\int_{\Theta} \frac{M_{22}f_1^2(x, \boldsymbol{\theta}) - 2M_{12}f_1(x, \boldsymbol{\theta})f_2(x, \boldsymbol{\theta}) + M_{11}f_2^2(x, \boldsymbol{\theta})}{|\mathbf{M}(\xi, \boldsymbol{\theta})|} p(\boldsymbol{\theta}) d\boldsymbol{\theta}. \quad (46)$$

Já para o modelo de Hill, seja a matriz de informação dada por

$$\mathbf{M}(\xi, \boldsymbol{\theta}_0) = \begin{bmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{bmatrix}_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}, \quad (47)$$

em que  $M_{ij} = M_{ji}$  para todo  $i, j \in \{1, 2, 3\}$ , então a variância padronizada é expressa por

$$\begin{aligned} d(x, \xi, \boldsymbol{\theta}_0) &= \frac{1}{|\mathbf{M}(\xi, \boldsymbol{\theta}_0)|} [f_1^2(x, \boldsymbol{\theta}_0)(M_{33}M_{22} - M_{23}^2) + f_1(x, \boldsymbol{\theta}_0)f_2(x, \boldsymbol{\theta}_0)(M_{23}M_{13} \\ &- M_{33}M_{12} - f_1(x, \boldsymbol{\theta}_0)f_3(x, \boldsymbol{\theta}_0)(M_{13}M_{22} - M_{12}M_{23}) + f_2(x, \boldsymbol{\theta}_0)f_1(x, \boldsymbol{\theta}_0) \\ &\times (M_{23}M_{13} - M_{33}M_{12}) - f_2^2(x, \boldsymbol{\theta}_0)(M_{13}^2 - M_{33}M_{11}) + f_3(x, \boldsymbol{\theta}_0)(M_{12}M_{13} \\ &- M_{23}M_{11}) - f_3(x, \boldsymbol{\theta}_0)f_1(x, \boldsymbol{\theta}_0)(M_{13}M_{22} - M_{12}M_{23}) + f_3(x, \boldsymbol{\theta}_0)f_2(x, \boldsymbol{\theta}_0) \\ &\times (M_{12}M_{13} - M_{23}M_{11}) - f_3^2(x, \boldsymbol{\theta}_0)(M_{12}^2 - M_{22}M_{11})]. \end{aligned} \quad (48)$$

Para delineamentos pseudo-Bayesianos a variância padronizada é dada pelo valor esperado da expressão (48) sob o espaço paramétrico, ou seja,

$$d(x, \xi) = \int_{\Theta} d(x, \xi, \boldsymbol{\theta})p(\boldsymbol{\theta})d\boldsymbol{\theta}. \quad (49)$$

Para os pseudo-Bayesianos, a incerteza em  $\kappa$  foi representada pela distribuição Gama com esperanças dadas pelas estimativas de Condomina et al. (2002) e variâncias definidas pelos coeficientes de variação indo de 0,05 a 0,50. Para representar a incerteza a respeito de  $\gamma$  foram utilizadas três distribuições discretas *a priori* com formas diferentes, o que simplificou a integral com respeito a este parâmetro para uma soma. A fim de representar  $\nu$  foram usados apenas valores pontuais, visto que a incerteza sobre este parâmetro não influencia a busca do delineamento D-ótimo. Os cálculos das integrais em  $\kappa$ , tanto para o valor da função-critério de D-otimalidade como para a variância padronizada, foram obtidos por aproximação pelo método de Monte Carlo.

## 5 RESULTADOS E DISCUSSÃO

Neste capítulo são apresentados resultados relativos à comparação dos algoritmos genético e *exchange* (seção 5.1) na tarefa de obter delineamentos D-ótimos exatos, ao estudo da interferência da informação *a priori* na construção de delineamentos localmente D-ótimos exatos (seção 5.2) e D-ótimos exatos pseudo-Bayesianos (seção 5.3) para os modelos de Michaelis-Menten e de Hill, bem como de delineamentos compostos baseando-se em ambos modelos (seção 5.4).

### 5.1 Comparação dos algoritmos genético e *exchange*

A fim de comparar os algoritmos genético e *exchange* na tarefa de obter delineamentos localmente D-ótimos exatos, foi considerada uma situação hipotética na qual deseja-se delinear um experimento futuro cujo interesse de pesquisa seja ajustar o modelo de Michaelis-Menten para explicar o influxo de zinco na parte proximal do intestino de ratos. Para tanto, o número de unidades experimentais disponíveis foi fixado em  $N = 8$  e utilizou-se as estimativas obtidas por Condomina et al. (2002) como informação *a priori* para representar os parâmetros  $\kappa$  e  $\nu$  (veja a Tabela 2).

A Figura 4 apresenta o valor do determinante da matriz de informação do melhor resultado encontrado para cada algoritmo em função do número de delineamentos avaliados até encontrá-lo. Verificou-se que o algoritmo genético produziu delineamentos de alta eficiência logo no princípio de sua execução, porém, ficou estagnado por grande período até atingir o critério de parada. O algoritmo genético apresentou praticamente o mesmo valor para os 30.000 delineamentos que precede-

ram o fim de sua execução. Então, foi concluído que o critério de parada adotado foi muito rigoroso e que obter resultados tão eficientes com um esforço computacional menor é possível. Contudo, é importante notar que o algoritmo genético pode levar a resultados sub-ótimos quando adota-se um critério de parada pouco rigoroso. Por

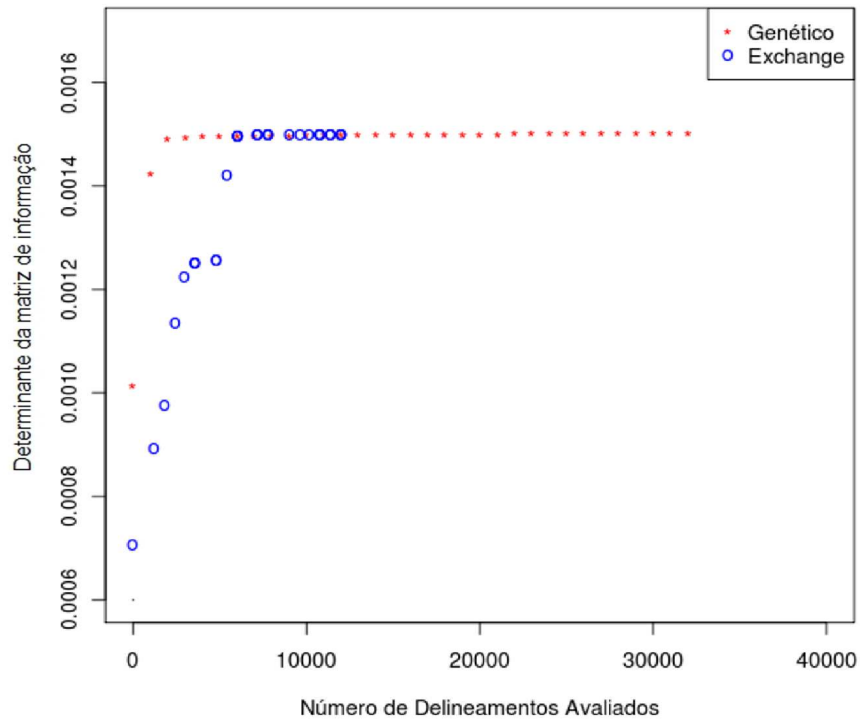


Figura 4: Determinante da matriz de informação para o melhor delineamento encontrado pelos algoritmos genético e *exchange* em função do número de delineamentos avaliados.

sua vez, o algoritmo *exchange* apresentou uma variação no valor do critério D mais lenta. No entanto, atingiu a convergência avaliando somente cerca de um terço do número de delineamentos avaliados pelo genético.

Apesar do algoritmo genético apresentar um esforço muito maior para convergir, possui a vantagem de ser baseado numa população de soluções potenciais, ao invés de utilizar um único ponto no espaço de busca. Portanto, dificilmente convergirá para ótimos locais. Agora, o algoritmo *exchange* apresenta este problema e precisa ser repetido. Se, por exemplo, for executado três vezes, exigirá um esforço

próximo ao requisitado pelo genético.

Não é possível afirmar que um algoritmo é melhor que o outro. Pode-se dizer que o genético leva vantagem no início da busca, porém, é complicado determinar um critério de parada adequado para ele, que garanta reprodutibilidade dos resultados e um esforço computacional pequeno, já que a população de soluções potenciais sofre a todo momento a influência estocástica da mutação e da recombinação. Deve-se reconhecer que o algoritmo *exchange* não permanece sob influência de fatores estocásticos, exceto pelo delineamento inicial, e, portanto, apresenta resultados mais refinados, com maior reprodutibilidade, sem permitir a perpetuação de cálculos desnecessários.

Talvez a conduta mais razoável seja a implementação de um algoritmo híbrido, que utilize a rapidez do genético no início da busca e, após um determinado número de delineamentos avaliados, refine a melhor solução a partir de algoritmo *exchange* modificado para tal.

Porém, este raciocínio introduz a necessidade de um ponto de corte, que determine quando deve-se utilizar o algoritmo genético e quando deve-se utilizar o algoritmo *exchange*. Logo, uma ampla gama de estudos futuros serão necessários se for desejada a implementação de uma versão híbrida.

Como o cálculo do valor da função-critério de D-otimalidade para delineamentos pseudo-Bayesianos é realizado neste trabalho com o auxílio do método de Monte Carlo, amostras de valores oriundos da distribuição *a priori* incorporada para representar os parâmetros do modelo são exigidas. Sabe-se que amostras maiores são mais representativas e aproximam melhor o valor do critério, contudo, podem ocasionar um aumento drástico no tempo de execução dos algoritmos.

Para avaliar como o tempo de execução dos algoritmos é influenciado pelo tamanho da amostra, considerou-se a busca de delineamentos D-ótimos exatos pseudo-Bayesianos com  $N = 8$  para o modelo de Michaelis-Menten. Foram utilizadas amostras com tamanho igual 10, 50, 100, 250 e 500 para representar o parâmetro  $\kappa$ , que foram obtidas a partir da distribuição Gama com média de 10,78 e coeficiente

de variação de 0,40. É importante mencionar que utilizou-se apenas uma amostra de cada tamanho para ambos algoritmos e suas repetições.

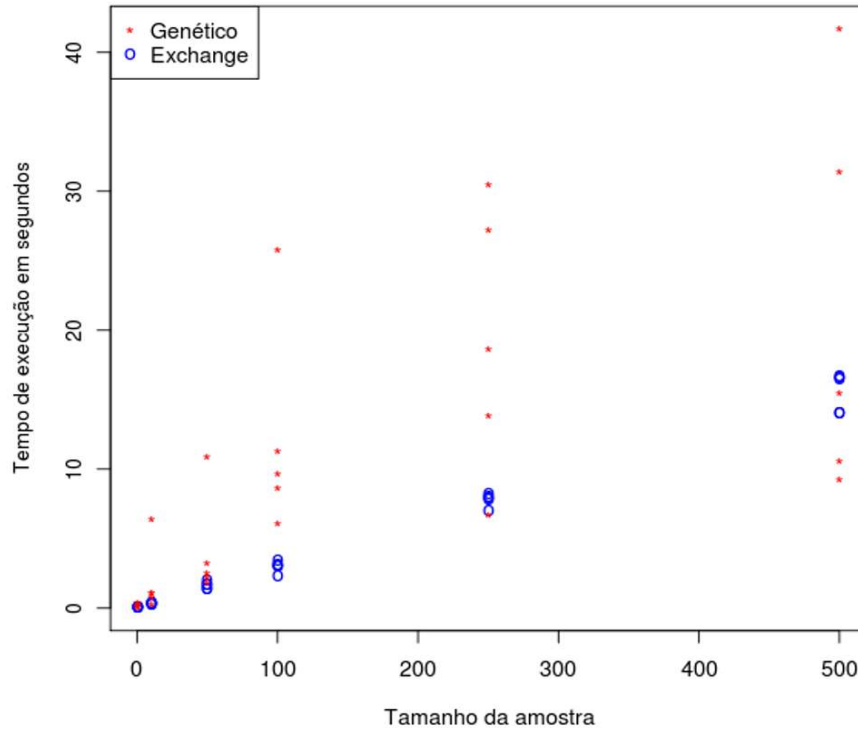


Figura 5: Tempo de execução dos algoritmos genético e *exchange* na busca de delineamentos D-ótimos exatos pseudo-Bayesianos com  $N = 8$  em função do tamanho de amostras para representar o parâmetro  $\kappa$ .

Os resultados apresentados na Figura 5 indicam heterocedasticidade para a resposta tempo de execução quando varia-se o tamanho da amostra. Para resumo e exploração destes resultados, a relação entre as variáveis tempo e tamanho da amostra foi linearizada pela transformação logarítmica, segundo o modelo

$$\ln t_i = \alpha_0 + \alpha_1 \ln a_i + e_i, \quad (50)$$

em que  $t_i$  é o tempo de execução,  $a_i$  é o tamanho da amostra e  $e_i$  é o erro aleatório, para o qual é suposto apresentar esperança igual a 0 e variância igual a  $\sigma^2$ , tendo em vista a  $i$ -ésima observação.



Retas foram ajustadas pelo método de mínimos quadrados considerando-se cada um dos algoritmos (veja a Figura 6). Para ambos ajustes, foram verificados resíduos que não violaram as pressuposições básicas de homocedasticidade e normalidade. As estimativas obtidas para os coeficientes do modelo estão apresentadas na Tabela 1.

Tabela 1: Ajustes do modelo (50) para os tempos de execução (em segundos) dos algoritmos *exchange* e genético.

Parâmetro	Estimativa	Erro Padrão
<i>Exchange</i>		
$\alpha_{e0}$	-3,19311	0,04771
$\alpha_{e1}$	0,94484	0,01110
Genético		
$\alpha_{g0}$	-2,05900	0,29562
$\alpha_{g1}$	0,86118	0,06876

O ajuste para o algoritmo *exchange* apresentou um alto coeficiente de determinação, a saber,  $R^2 = 0,996$ . Além disso, o quadrado médio do resíduo foi relativamente pequeno (0,016), produzindo estimativas precisas para os coeficientes do modelo. Por sua vez, o ajuste para o algoritmo genético apresentou um coeficiente de determinação um pouco menor,  $R^2 = 0,843$ . Já o quadrado médio do resíduo foi bem maior que o apresentado para o algoritmo *exchange*, a saber, com o valor de 0,618.

A razão estimada (a partir dos modelos ajustados) entre o tempo gasto pelo algoritmo genético e o tempo gasto pelo algoritmo *exchange* em função do tamanho da amostra é dada pela expressão

$$\text{Razão}(a_i) = \exp\{\hat{\alpha}_{g0} - \hat{\alpha}_{e0}\} a_i^{(\hat{\alpha}_{g1} - \hat{\alpha}_{e1})} = 3,1084 a_i^{-0,0837}, \quad (51)$$

em que  $\text{Razão}(a_i)$  é a razão dos tempos em função do tamanho da amostra  $a_i$ .

A razão entre tempos, apresentada na Figura 7, decresce exponencialmente quando aumenta-se o valor de tamanho de amostra, indicando que o tempo de execução do algoritmo genético torna-se cada vez mais próximo do exibido pelo algoritmo *exchange*. No entanto, verifica-se que o algoritmo *exchange* é o muito mais rápido quando são considerados tamanhos de amostra utilizados na prática.

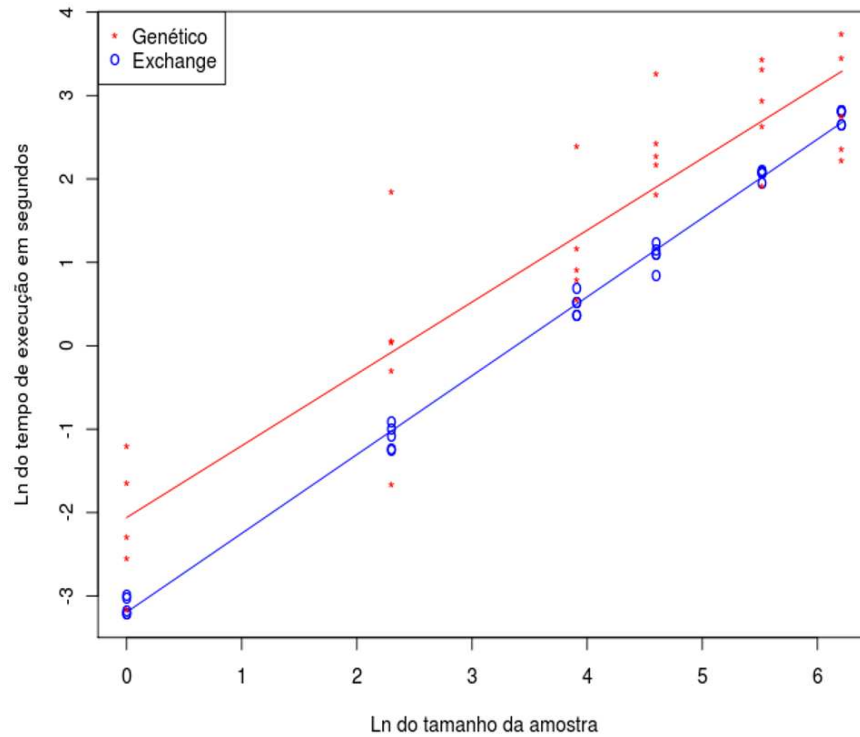


Figura 6: Tempos de execução dos algoritmos genético e *exchange* em função do tamanho de amostra. Dados e ajuste linear para a escala logarítmica.

## 5.2 Delineamentos localmente D-ótimos

Nesta seção são apresentados os delineamentos localmente D-ótimos exatos encontrados para os modelos de Michaelis-Menten e Hill, levando-se em conta as estimativas obtidas por Condomina et al. (2002) como informação *a priori* para representar os parâmetros  $\kappa$  e  $\nu$ . Para tanto, considerou-se como região experimental o conjunto  $\chi = [X_{inf}, X_{sup}]$ , com  $X_{inf}$  e  $X_{sup}$  positivos. Para esta região, Dette et al.

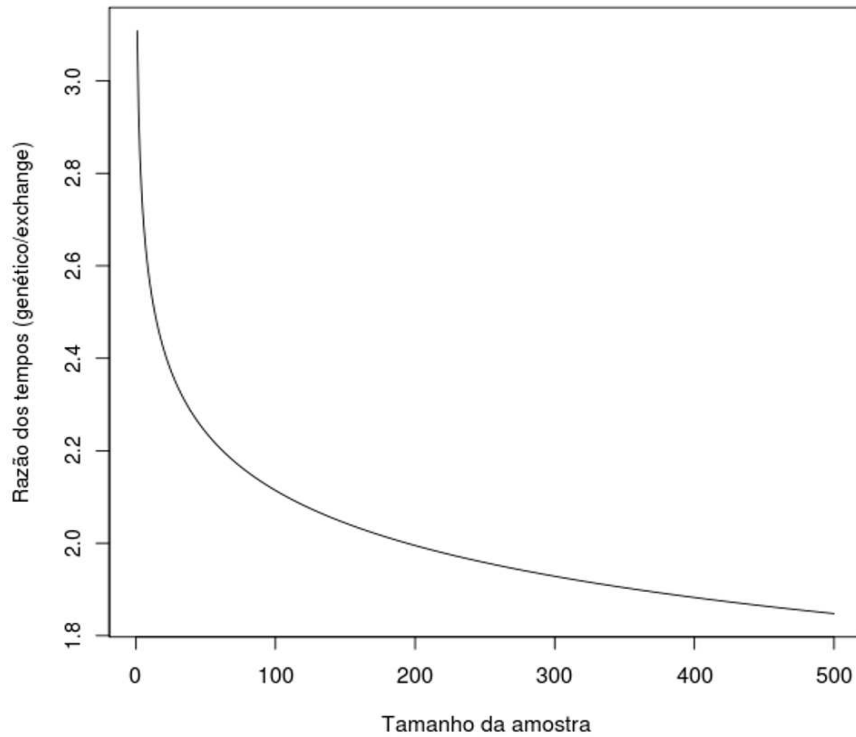


Figura 7: Razão dos tempos de execução  $Razão(a_i)$  dos algoritmos genético e *exchange* em função do tamanho de amostra  $a_i$ , calculada a partir dos valores de tempo ajustados pelo modelo (50).

(2008) mostraram que existe um delineamento localmente D-ótimo contínuo  $\xi^*$  para o ajuste do modelo de Michaelis-Menten com dois pontos de suporte e balanceado, ou seja,

$$\xi^* = \left\{ \begin{array}{cc} \max\{X_{inf}, x^*\} & X_{sup} \\ 0,5 & 0,5 \end{array} \right\}, \quad (52)$$

em que  $x^* = \kappa X_{sup} / (2\kappa + X_{sup})$ . Uma propriedade importante para o modelo de Michaelis-Menten é que  $\xi^*$  não depende do parâmetro  $\nu$ . Isto é decorrente da linearidade de  $\nu$  na parte fixa do modelo. Além disso, é possível notar que o extremo superior da região experimental  $X_{sup}$  está presente em  $\xi^*$ .

A Figura 8 apresenta a relação entre  $x^*$  e  $\kappa$  e mostra que o aumento de  $\kappa$  ocasiona um incremento no valor de  $x^*$ . Ainda, é possível mostrar que  $x^*$  tende

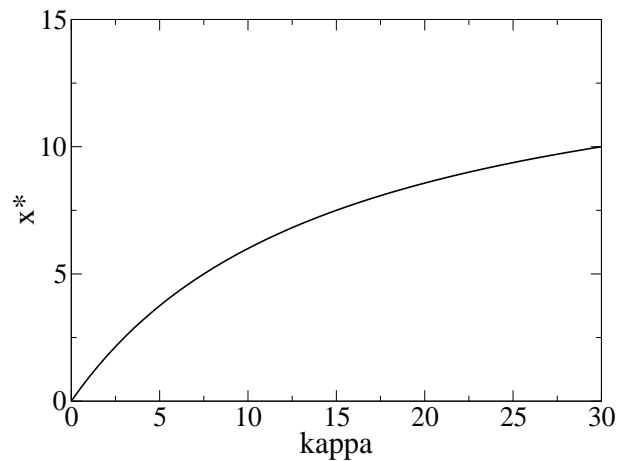


Figura 8: Valores para  $x^*$  em função da informação *a priori* para  $\kappa$ .

a  $X_{sup}/2$  quando  $\kappa$  tende ao infinito. Em geral, os pesquisadores utilizam extremos superiores  $X_{sup}$  maiores que o dobro do valor suposto para  $\kappa$ , com a finalidade de contemplar regiões próximas à assíntota (para a estimação da velocidade máxima) em seus estudos.

Para a obtenção de delineamentos exatos, a região experimental foi discretizada numa grade de 600 pontos igualmente espaçados, dados  $X_{inf} = 0,05 \text{ mM}$  e  $X_{sup} = 30,00 \text{ mM}$ . Foram obtidos delineamentos com tamanho  $N = 8$  para o modelo de Michaelis-Menten, utilizando-se como informação *a priori* as estimativas pontuais obtidas por Condomina et al. (2002), que são apresentadas na Tabela 2.

Tabela 2: Estimativas dos parâmetros do modelo de Michaelis-Menten obtidas por Condomina et al. (2002).

Porção do Intestino	$\kappa$ (em $mM$ )	$\nu$ ( $mmol/cm^2h$ ) $\cdot 10^3$
Proximal	10,78	8,39
Medial	1,94	1,62
Distal	3,04	3,42

Os delineamentos encontrados, dispostos na Tabela 3, apresentam apenas dois pontos de suporte, com 4 réplicas cada. Como esperado, é verificado em todos os resultados o extremo superior da região experimental. Além disso, o menor ponto de suporte de  $\xi^*$ , em cada caso, mostra-se proporcional ao valor considerado para representar  $\kappa$ .

Os delineamentos exatos encontrados são muito similares aos D-ótimo contínuos. Para a parte medial e distal, os delineamentos obtidos pelos algoritmos implementados coincidem com os delineamentos oriundos do arredondamento do D-ótimo contínuo, ou seja, para tais regiões a abordagem exata é equivalente à aproximada. Para a parte proximal do intestino, o delineamento aproximado tem como menor ponto de suporte o valor de 6,30 ao invés de 6,25. Então,  $\Psi\{\mathbf{M}(\xi^*, \boldsymbol{\theta}_0)\} = -6,502169$ , isto é, um valor ligeiramente menor que o encontrado pela abordagem exata. Logo, conclui-se que, para este caso estudado, as abordagens exata e aproximada são realmente muito próximas.

Tabela 3: Delineamentos localmente D-ótimos exatos (número de réplicas entre parênteses) para o ajuste do modelo de Michaelis-Menten utilizando como informação *a priori* as estimativas obtidas por Condomina et al. (2002). Seja  $x^*$  o ponto de suporte inferior do delineamento D-ótimo contínuo.

Porção do Intestino	$\kappa$ (em $mM$ )	$x^*$	$\xi^*$	$\Psi\{\mathbf{M}(\xi^*, \boldsymbol{\theta}_0)\}$
Proximal	10,78	6,2723	6,25 (4) 30,00 (4)	-6,502164
Medial	1,94	1,7178	1,70 (4) 30,00 (4)	-4,895438
Distal	3,04	2,5277	2,55 (4) 30,00 (4)	-4,502492

Uma maneira de verificar se delineamentos contínuos são D-ótimos é visualizar o gráfico de sua variância padronizada. O delineamento contínuo é D-ótimo se em seus pontos de suporte a variância padronizada assumir um valor igual ao número de parâmetros do modelo utilizado. Tal propriedade não é necessariamente

verdadeira para delineamentos ótimos exatos, contudo, para modelos simples como o de Michaelis-Menten e o de Hill, espera-se que os resultados sejam próximos àqueles verificados para delineamentos contínuos.

Na figura 9 são apresentados os gráficos da variância padronizada para os delineamentos encontrados. Verificou-se que eles apresentam picos (máximos) da variância padronizada em pontos muito próximos aos obtidos pelos delineamentos localmente D-ótimos contínuos. Além disso, a variância padronizada se aproxima de  $p$  em seus pontos de suporte, possibilitando serem D-ótimos. Ainda, a similaridade dos resultados verificada pelas duas abordagens indica que os algoritmos implementados, que são métodos heurísticos, estão convergindo de maneira satisfatória para o D-ótimo exato ou para pontos próximos a ele.

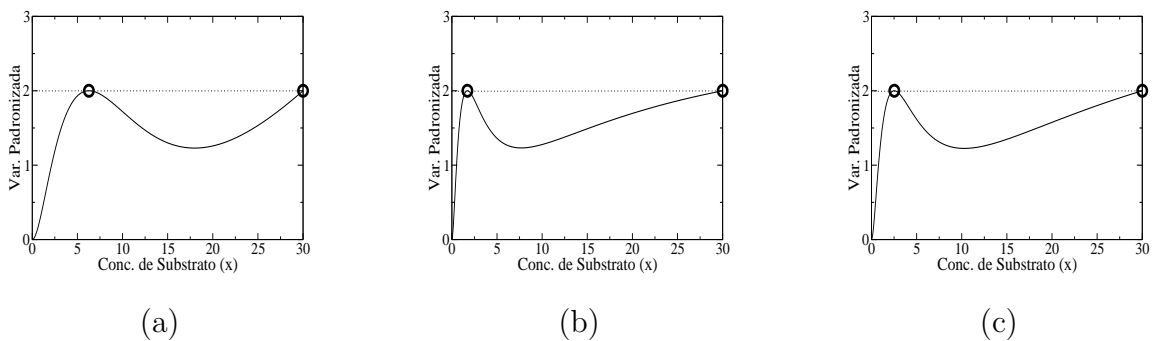


Figura 9: Variância padronizada para os delineamentos localmente D-ótimos exatos (linha sólida) para o modelo de Michaelis-Menten a partir da informação *a priori* contida no trabalho de Condomina et al. (2002), sendo referente à parte proximal em (a), à parte medial em (b) e à parte distal em (c). Os círculos indicam a variância padronizada para o delineamento D-ótimo contínuo avaliada em seus dois pontos de suporte. A linha pontilhada é utilizada como referência para a D-otimalidade.

O fato dos delineamentos D-ótimos para o modelo de Michaelis-Menten, tanto contínuos como exatos, apresentarem apenas dois pontos de suporte torna-os não muito atrativos na prática, pois mais pontos seriam necessários para checar se há falta de ajuste. Com o objetivo de obter delineamentos um pouco mais atrativos, considerou-se a generalização do modelo para situações em que a coope-

ratividade positiva ou negativa pode dominar o processo de transporte de zinco no intestino delgado. Isto é, visou-se obter delineamentos para o ajuste do modelo de Hill em experimentos futuros.

Tal alternativa será bem sucedida se, após ajustado o modelo, a estimativa para o coeficiente de Hill for diferente de 1, de modo significativo. Caso contrário, o modelo de Michaelis-Menten deverá ser ajustado, porém, fazendo-se uso de resultados provenientes de um experimento sub-ótimo para este fim.

A partir do uso das estimativas pontuais obtidas por Condomina et al. (2002) como informação *a priori* para representar  $\kappa$  e  $\nu$ , buscou-se por delineamentos localmente D-ótimos exatos com  $N = 12$  para o modelo de Hill. Para  $\gamma$  foi suposto o valor de 1, visto que há pouca informação para dizer se a cooperatividade positiva ou negativa é predominante.

Novamente, pode-se mostrar que os delineamentos obtidos não dependem do valor do parâmetro  $\nu$ , pois no modelo de Hill ele também aparece de maneira linear e pode ser fatorado na função-critério de D-otimalidade, não afetando, assim, o processo de otimização. No entanto, as suas estimativas pontuais, vistas em Condomina et al. (2002), foram admitidas como *prioris* no planejamento apenas para cálculo do valor do critério. Os melhores delineamentos encontrados são dispostos na Tabela 4.

Os delineamentos exatos encontrados para o modelo de Hill apresentam de 3 a 5 pontos de suporte. Em todos os casos, o extremo superior da região experimental ocupa 1/3 das unidades experimentais. Para as partes proximal e distal foram obtidos delineamentos exatos com 5 e 4 pontos de suporte, respectivamente. Contudo, pares de pontos próximos como 1, 80 e 1, 85; 10, 25 e 10, 30 ou 0, 75 e 0, 80 podem indicar arredondamentos oriundos da discretização da região experimental. Nestes casos, pontos de suporte de delineamentos localmente D-ótimos contínuos podem estar situados no interior de intervalos definidos por tais pares de pontos dos delineamentos exatos. Para a parte medial foi observado um delineamento exato com apenas 3 pontos de suporte, ou seja, com suporte mínimo, pois o número de pontos

de suporte é igual ao número de parâmetros do modelo.

Tabela 4: Delineamentos localmente D-ótimos exatos (número de réplicas entre parênteses) para o modelo de Hill, utilizando-se as estimativas obtidas por Condomina et al. (2002) como informação *a priori* para representar  $\kappa$  e  $\nu$ . Para  $\gamma$  foi suposto o valor de 1.

Porção do Intestino	$\xi^*$				$\Psi\{\mathbf{M}(\xi^*, \boldsymbol{\theta}_0)\}$	
Proximal	1,80 (2)	1,85 (2)	10,25 (2)	10,30 (2)	30,00 (4)	-8,084668
Medial		0,55 (4)	3,80 (4)	30,00 (4)		-8,693460
Distal	0,75 (3)	0,80 (1)	5,15 (4)	30,00 (4)		-6,996473

A D-otimalidade dos delineamentos encontrados foi explorada a partir dos gráficos de variância padronizada, conforme a Figura 10. Visivelmente, é possível determinar 3 máximos locais para a variância padronizada em cada um dos casos.

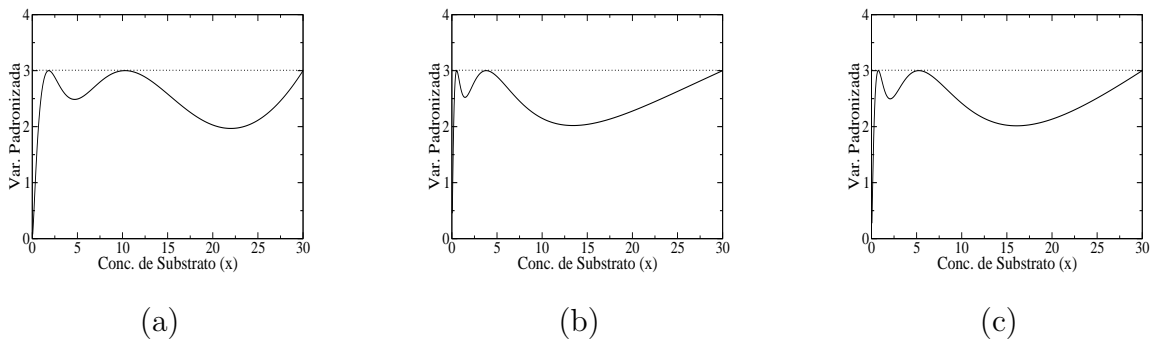


Figura 10: Variância padronizada para os delineamentos localmente D-ótimos exatos (linha sólida) para o modelo de Hill a partir da informação *a priori* contida no trabalho de Condomina et al. (2002), sendo referente à parte proximal em (a), à parte medial em (b) e à parte distal em (c). A linha pontilhada é utilizada como referência para a D-otimalidade.

Ademais, observa-se que as ordenadas de tais pontos situam-se



próximas do valor  $p$  e, desta forma, tem-se como conclusão que os delineamentos exatos encontrados apresentam propriedades muito similares aquelas observadas para os delineamentos localmente D-ótimos contínuos.

Verifica-se que os delineamentos encontrados para o modelo de Michaelis-Menten são bem distintos dos obtidos para o modelo de Hill. Como esperado, a introdução do parâmetro  $\gamma$  exige mais pontos de suporte. Além disso, enquanto os delineamentos exatos para o modelo de Michaelis-Menten apresentam metade das unidades experimentais ao nível  $X_{sup}$ , no modelo de Hill esta proporção é de apenas um terço.

### 5.3 Delineamentos D-ótimos pseudo-Bayesianos

Com a finalidade de obter delineamentos mais robustos, ou seja, menos sensíveis aos valores pontuais introduzidos como informação *a priori*, buscou-se por delineamentos D-ótimos exatos pseudo-Bayesianos. Primeiramente, considerou-se o modelo de Michaelis-Menten com  $N = 8$ . Nesta etapa, o cálculo do critério D foi realizado com o auxílio do método de Monte Carlo, a partir de 500 valores amostrados da distribuição Gama para representar o parâmetro  $\kappa$ . Os valores apresentados por Condomina et al. (2002) foram fixados como média da distribuição *a priori* e utilizou-se diferentes coeficientes de variação. Os delineamentos obtidos para a porção proximal, medial e distal do intestino são apresentados na Tabela 5.

Como esperado, o extremo superior da região experimental é ponto de suporte para todos os delineamentos obtidos. Além disso, em todos os casos ele ocupa metade das unidades experimentais. Também foi verificado que os  $N/2$  pontos remanescentes variam de acordo com os valores introduzidos para representar  $\kappa$ , com uma associação positiva e, aparentemente, não-linear. Ademais, foi verificado que o aumento do coeficiente de variação ocasiona a diminuição dos valores de tais pontos (Figura 11). Este fato pode ser explicado da seguinte maneira: com a introdução de incerteza sobre os valores de  $\kappa$ , deve-se contemplar uma região maior do espaço paramétrico e, desta forma, valores menores de  $x$  são necessários para estimar os

parâmetros de modo adequado. Pequenas variações a esta tendência são esperadas, já que utilizou-se o método de Monte Carlo para aproximar o valor do critério D.

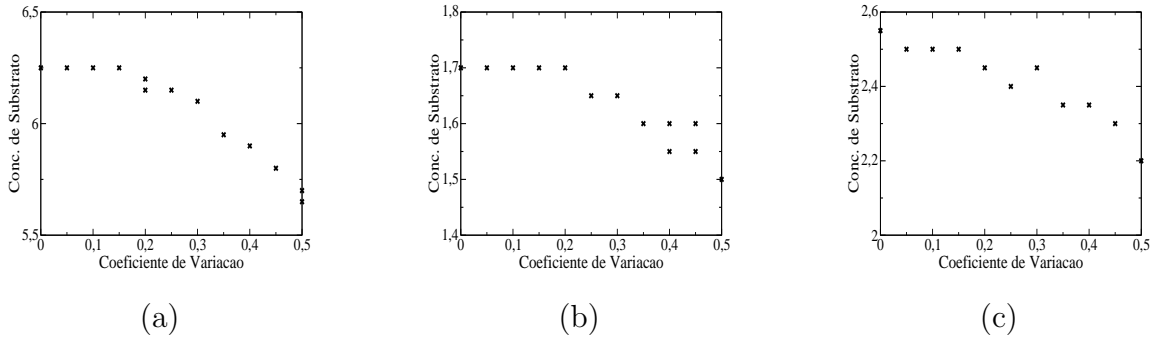


Figura 11: Pontos de suporte inferiores dos delineamentos D-ótimos pseudo-Bayesianos para o modelo de Michaelis-Menten em função do coeficiente de variação adotado para a *priori* do parâmetro  $\kappa$ .

Outro aspecto interessante é que os delineamentos encontrados apresentam apenas dois pontos de suporte, com exceção do relativo à parte proximal com coeficiente de variação de 0,50. Contudo, tal delineamento apresenta dois pontos de suporte muito próximos, 5,65 e 5,70, fato que pode ser proveniente de arredondamento ocasionado pela utilização de uma grade de valores para representar a região experimental. Intuitivamente, esperava-se que a incerteza introduzida a respeito de  $\kappa$  produzisse delineamentos D-ótimos pseudo-Bayesianos com maior número de pontos de suporte que os localmente D-ótimos, porém, isto não ocorreu.

Considerando agora o modelo de Hill, obteve-se delineamentos D-ótimos exatos pseudo-Bayesianos com tamanho  $N = 12$ . Para representar o parâmetro  $\kappa$ , foram utilizados 100 valores amostrados da distribuição Gama. As estimativas obtidas por Condomina et al. (2002) para  $\nu$  e  $\kappa$  foram fixadas como médias para a distribuição *a priori* e foram adotados diferentes coeficientes de variação.

Para representar  $\gamma$ , fez-se uso de distribuições *a priori* discretas, considerando-se três possibilidades: (i) simétrica com média 1 - situações em que acredita-se que há cooperação positiva ou negativa, porém, não há nenhum conheci-

Tabela 5: Delineamentos D-ótimos exatos pseudo-Bayesianos (número de réplicas entre parênteses) para o modelo de Michaelis-Menten, utilizando-se como informação *a priori* distribuições Gama com diferentes coeficientes de Variação (C.V.) baseadas nas estimativas obtidas por Condomina et al. (2002).

Porção do Intestino	C.V.	$\xi^*$	$\Psi\{\mathbf{M}(\xi^*)\}$
Proximal	0,05	6,25 (4) 30,00 (4)	-6,490417
	0,25	6,15 (4) 30,00 (4)	-6,456540
	0,50	5,65 (2) 5,70 (2) 30,00 (4)	-6,231684
Medial	0,05	1,70 (4) 30,00 (4)	-4,897111
	0,25	1,65 (4) 30,00 (4)	-4,832924
	0,50	1,50 (4) 30,00 (4)	-4,692446
Distal	0,05	2,50 (4) 30,00 (4)	-4,497628
	0,25	2,40 (4) 30,00 (4)	-4,380247
	0,50	2,20 (4) 30,00 (4)	-4,289565

mento *a priori* a respeito do tipo de fenômeno que domina o processo; (ii) assimétrica à esquerda com média menor que 1 - situações em que a cooperação negativa entre ligantes é mais provável; (iii) assimétrica à direita com média maior que 1 - situações em que a cooperação positiva entre ligantes é mais provável.

A escolha de uma distribuição discreta para representar  $\gamma$  deve-se ao fato do coeficiente de Hill ser determinado pelo número de sítios que são liberados ou ocupados para um ligante específico em uma macromolécula quando outra estrutura se liga a algum de seus sítios e, por isso, ele pode assumir, em teoria, apenas valores discretos. As distribuições *a priori* utilizadas para representar  $\gamma$  são apresentadas na Figura 12.

Os resultados obtidos para as várias situações mencionadas acima são apresentados na Tabela 6. Um fato notável é que em todos os casos os delinea-

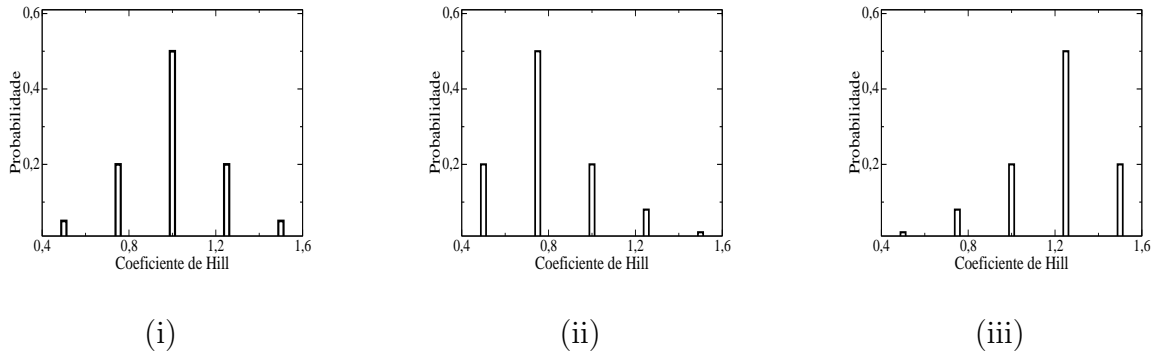


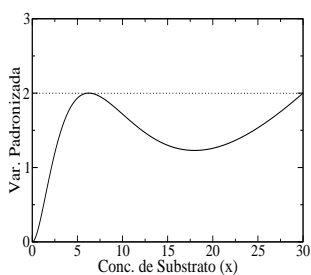
Figura 12: Distribuições discretas *a priori* utilizadas para representar o parâmetro  $\gamma$ , o coeficiente de Hill. Em (i) tem-se uma distribuição simétrica com média 1; em (ii) tem-se uma distribuição assimétrica à direita com média 0,805; em (iii) tem-se uma distribuição assimétrica à esquerda com média 1,195.

mentos encontrados apresentam 1/3 de seus pontos ao nível 30,00, ou seja, situados no extremo superior da região experimental. Além disso, observou-se que, conforme o esperado, o aumento da imprecisão na especificação da distribuição *a priori* para  $\kappa$  resultou em um aumento do número de pontos de suporte dos delineamentos encontrados. Com a distribuição assimétrica à direita para representar  $\gamma$ , com média igual a 0,805, são notados pontos de suporte inferiores menores que os obtidos quando utiliza-se a distribuição simétrica. Por sua vez, para a distribuição assimétrica à esquerda, com média igual a 1,195, são encontrados pontos de suporte inferiores maiores que os obtidos para as demais situações. Em suma, parece haver uma correlação positiva entre  $\gamma$  e os valores dos pontos de suporte inferiores dos delineamentos encontrados.

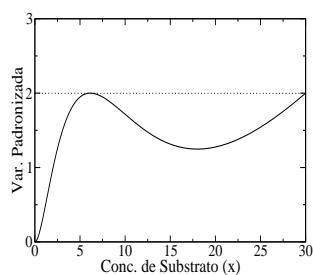
A fim de verificar a proximidade entre as abordagens exata e contínua, tendo em vista os delineamentos D-ótimos pseudo-Bayesianos, foram construídos gráficos da variância padronizada para cada um dos casos estudados. A Figura 13 apresenta os resultados obtidos para o modelo de Michaelis-Menten enquanto as Figuras 14, 15 e 16 mostram os resultados obtidos para o modelo de Hill quando considerou-se a distribuição simétrica, assimétrica à direita e assimétrica à esquerda como *priori* de  $\gamma$ , respectivamente.

Tabela 6: Delineamentos D-ótimos exatos pseudo-Bayesianos (número de réplicas entre parênteses) para o modelo de Hill, utilizando-se distribuições Gama com diferentes Coeficientes de Variação (C.V.) como informação *a priori* para representar  $\kappa$  e distribuições discretas para representar  $\gamma$ . Os valores esperados utilizados para os parâmetros  $\kappa$  e  $\nu$  foram baseados nas estimativas apresentadas por Condomina et al. (2002).

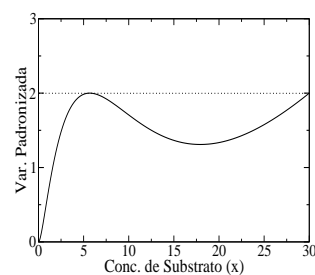
Distribuição de $\gamma$	Porção	C.V. para $p(\kappa)$	$\xi^*$	$\Psi\{M(\xi^*)\}$
Simétrica				
	Proximal	0,05	1,85 (4) 10,30 (4) 30,00 (4)	-8,304525
		0,25	1,80 (4) 10,10 10,15 (3) 30,00 (4)	-8,206643
		0,50	1,65 (2) 1,70 (2) 9,80 (4) 30,00 (4)	-8,133024
	Medial	0,05	0,55 (4) 3,70 3,75 (3) 30,00 (4)	-8,919067
		0,25	0,50 (3) 0,55 3,55 3,60 (3) 30,00 (4)	-8,824269
		0,50	0,45 0,50 (3) 3,45 3,50 (2) 3,75 30,00 (4)	-8,940705
	Distal	0,05	0,75 (2) 0,80 (2) 5,10 (4) 30,00 (4)	-7,671944
		0,25	0,75 (4) 5,00 (4) 30,00 (4)	-7,636434
		0,50	0,55 0,60 (2) 0,80 4,40 (2) 4,45 4,50 30,00 (4)	-7,234589
Assimétrica à esquerda				
	Proximal	0,05	2,65 (4) 11,40 (4) 30,00 (4)	-7,795537
		0,25	2,55 (4) 10,15 11,20 (3) 30,00 (4)	-7,720015
		0,50	2,35 (2) 2,40 (2) 10,80 10,85 (3) 30,00 (4)	-7,718657
	Medial	0,05	0,70 (4) 3,65 3,70 (3) 30,00 (4)	-8,555263
		0,25	0,65 (2) 0,70 (2) 3,55 (3) 3,60 30,00 (4)	-8,486217
		0,50	0,45 0,55 0,80 (2) 2,40 4,05 (3) 30,00 (4)	-8,669185
	Distal	0,05	1,05 (4) 5,25 (4) 30,00 (4)	-7,238994
		0,25	1,00 (4) 5,10 (2) 5,15 (2) 30,00 (4)	-7,230194
		0,50	0,55 0,95 (2) 1,00 4,55 (2) 4,60 4,80 30,00 (4)	-6,915425
Assimétrica à direita				
	Proximal	0,05	1,05 (2) 1,10 (2) 8,75 8,80 (3) 30,00 (4)	-8,932629
		0,25	1,05 (4) 8,65 (4) 30,00 (4)	-8,819830
		0,50	0,90 0,95 (2) 1,15 8,35 (3) 8,50 30,00 (4)	-8,690278
	Medial	0,05	0,35 (4) 3,55 (4) 30,00 (4)	-9,508709
		0,25	0,30 0,35 (3) 3,40 3,45 (3) 30,00 (4)	-9,390721
		0,50	0,30 (2) 0,35 (2) 3,35 (2) 3,40 3,50 30,00 (4)	-9,442649
	Distal	0,05	0,45 0,50 (3) 4,70 (4) 30,00 (4)	-8,321463
		0,25	0,45 (2) 0,50 (2) 4,60 (4) 30,00 (4)	-8,263408
		0,50	0,35 (2) 0,45 (2) 4,05 4,10 (2) 4,15 30,00 (4)	-7,778941



$$C.V. = 0,05$$

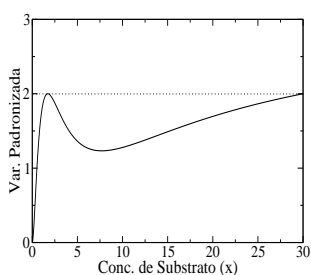


$$C.V. = 0,25$$

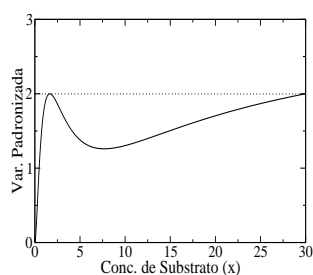


$$C.V. = 0,50$$

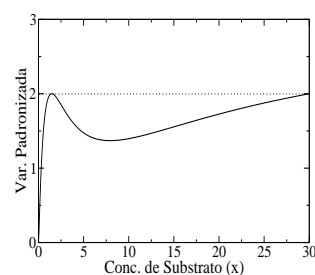
### Parte Proximal do Intestino Delgado



$$C.V. = 0,05$$

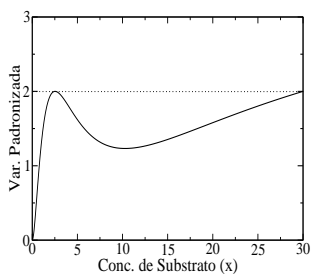


$$C.V. = 0,25$$

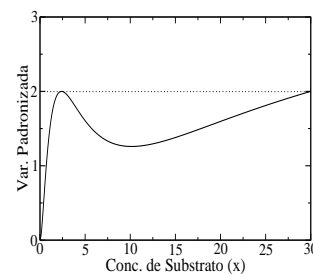


$$C.V. = 0,50$$

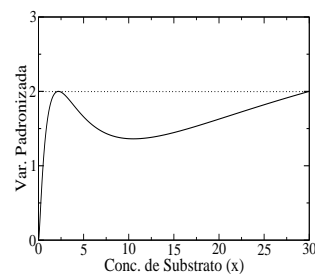
### Parte Medial do Intestino Delgado



$$C.V. = 0,05$$



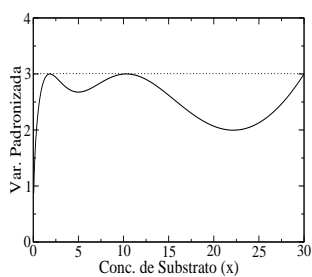
$$C.V. = 0,25$$



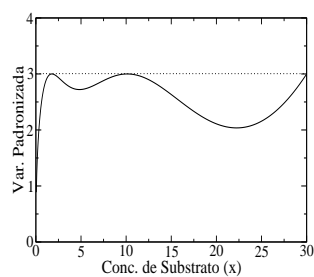
$$C.V. = 0,50$$

### Parte Distal do Intestino Delgado

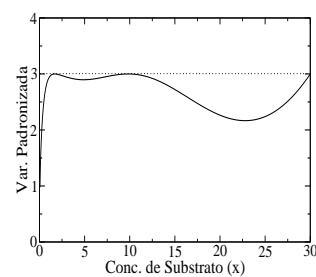
Figura 13: Gráficos da variância padronizada para os delineamentos D-ótimos exatos pseudo-Bayesianos para o modelo de Michaelis-Menten.



$C.V. = 0,05$

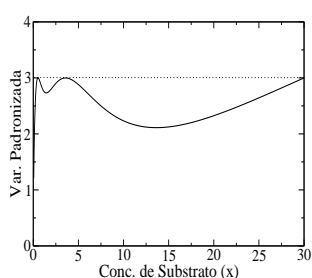


$C.V. = 0,25$

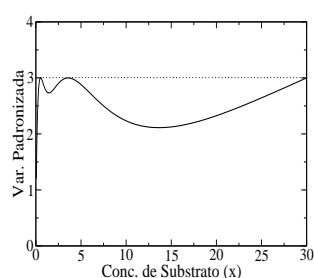


$C.V. = 0,50$

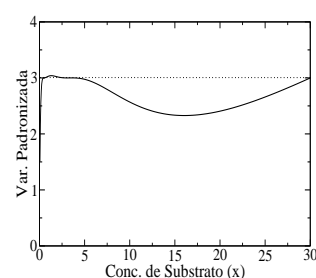
### Parte Proximal do Intestino Delgado



$C.V. = 0,05$

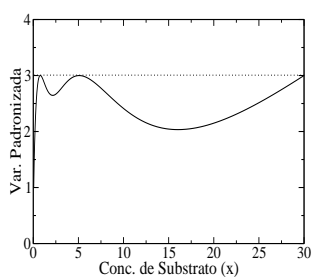


$C.V. = 0,25$

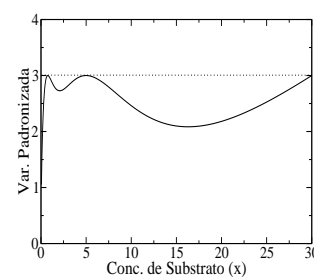


$C.V. = 0,50$

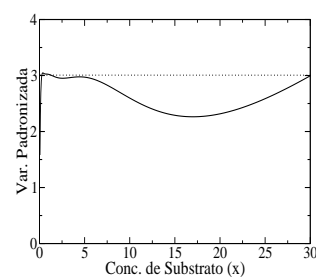
### Parte Medial do Intestino Delgado



$C.V. = 0,05$



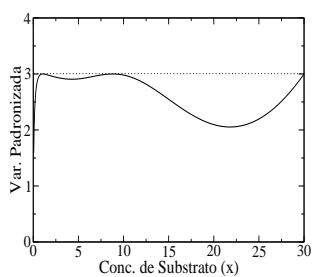
$C.V. = 0,25$



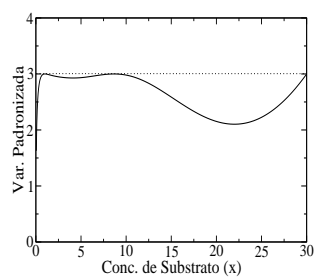
$C.V. = 0,50$

### Parte Distal do Intestino Delgado

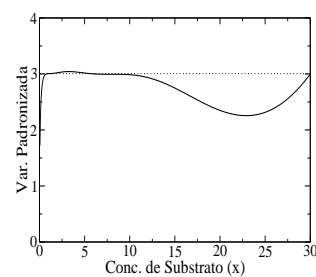
Figura 14: Gráficos da variância padronizada para os delineamentos D-ótimos exatos pseudo-Bayesianos para o modelo de Hill quando utiliza-se uma distribuição discreta simétrica com média 1 para representar  $\gamma$ .



$$C.V. = 0,05$$

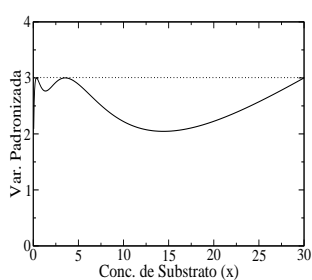


$$C.V. = 0,25$$

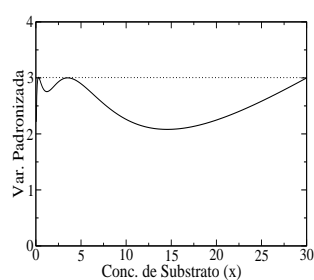


$$C.V. = 0,50$$

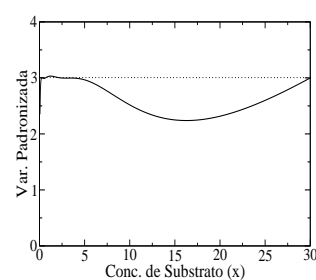
### Parte Proximal do Intestino Delgado



$$C.V. = 0,05$$

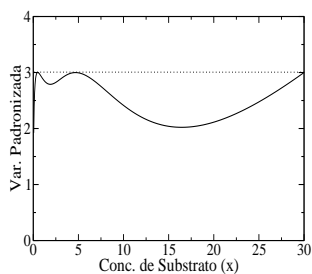


$$C.V. = 0,25$$

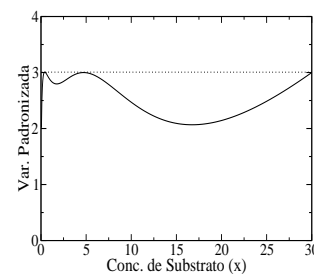


$$C.V. = 0,50$$

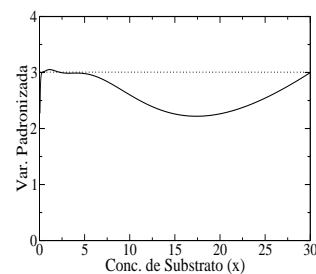
### Parte Medial do Intestino Delgado



$$C.V. = 0,05$$



$$C.V. = 0,25$$

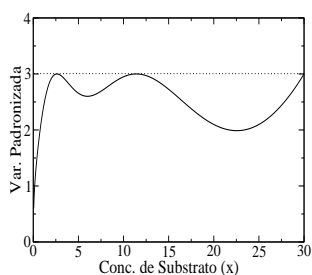


$$C.V. = 0,50$$

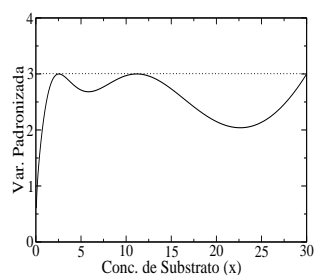
### Parte Distal do Intestino Delgado

Figura 15: Gráficos da variância padronizada para os delineamentos D-ótimos exatos pseudo-Bayesianos para o modelo de Hill quando utiliza-se uma distribuição discreta assimétrica à direita com média 0,805 para representar  $\gamma$ .

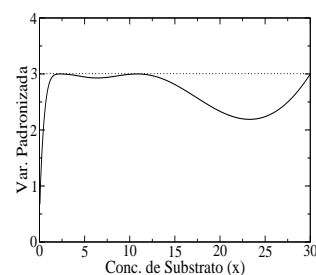




$$C.V. = 0,05$$

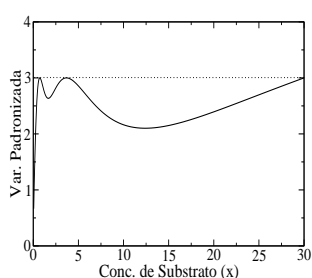


$$C.V. = 0,25$$

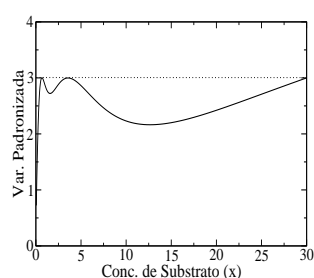


$$C.V. = 0,50$$

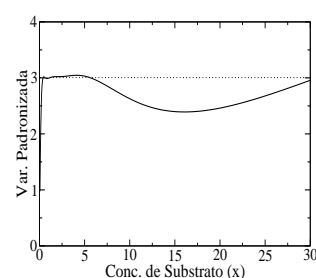
### Parte Proximal do Intestino Delgado



$$C.V. = 0,05$$

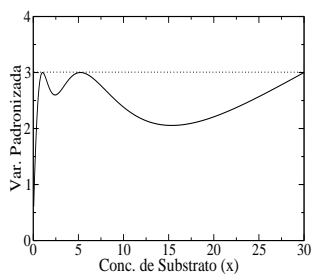


$$C.V. = 0,25$$

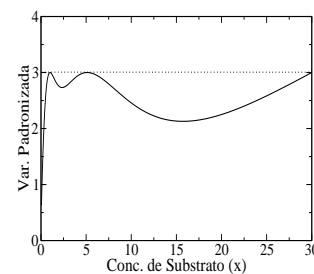


$$C.V. = 0,50$$

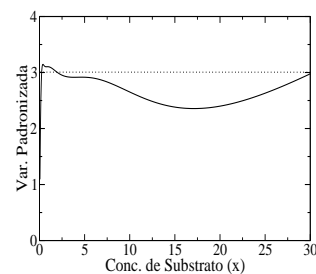
### Parte Medial do Intestino Delgado



$$C.V. = 0,05$$



$$C.V. = 0,25$$



$$C.V. = 0,50$$

### Parte Distal do Intestino Delgado

Figura 16: Gráficos da variância padronizada para os delineamentos D-ótimos exatos pseudo-Bayesianos para o modelo de Hill quando utiliza-se uma distribuição discreta assimétrica à esquerda com média 1,195 para representar  $\gamma$ .

Em geral, verificou-se que a variância padronizada apresenta valores próximos de  $p$  nas imediações dos pontos de suporte dos delineamentos encontrados. Apenas alguns desvios moderados podem ser vistos, isto quando considera-se um coeficiente de variação alto para a distribuição *a priori* de  $\kappa$ .

Pela Figura 15 verifica-se um resultado atípico para a parte distal, pois o valor da variância padronizada excede em muito a linha pontilhada nos menores pontos de suporte e, após, para os pontos medianos do delineamento, fica bem abaixo da referência. Neste caso há duas possibilidades: as abordagens exata e contínua apresentam resultados muito diferentes para esta situação ou o delineamento encontrado não é o D-ótimo exato pseudo-Bayesiano. A última possibilidade é mencionada devido à natureza heurística dos algoritmos e ao pequeno número de repetições realizadas para a obtenção dos resultados. Os demais delineamentos apresentam-se adequados e, em tais situações, verifica-se que as abordagens exata e contínua apresentam-se próximas.

## 5.4 Delineamentos ótimos segundo critérios compostos

Nesta seção serão abordados delineamentos compostos para os modelos de Michaelis-Menten e de Hill. A razão para tal é encontrar delineamentos bons para o ajuste de ambos. Assim, a cooperação entre ligantes pode ser explorada a partir do ajuste do modelo de Hill e, caso  $\gamma$  não seja diferente de 1, ainda pode-se obter um bom ajuste para o modelo de Michaelis-Menten.

Critérios que consideram a incerteza sobre modelos de regressão dado um número limitado de modelos plausíveis podem ser encontrados na bibliografia, como os vistos em Ponce de Leon (1993), e seriam uma alternativa a ser pensada para este caso, porém, não foram abordados neste trabalho.

Nesta etapa foram obtidos delineamentos localmente D-ótimos exatos com  $N = 12$  a partir do critério de D-otimalidade composto, dado pela combinação linear das funções-critérios oriundas de ambos modelos. Para tanto, foi introduzido o parâmetro  $\lambda$  com o objetivo de especificar o peso com o qual cada modelo contri-

bui no planejamento experimental. Se  $\lambda$  é igual a 1, todo o esforço concentra-se em estimar os parâmetros para o modelo de Michaelis-Menten e o delineamento obtido pode ser, até mesmo, singular para o modelo de Hill. Já se  $\lambda$  for igual a 0, todo esforço concentra-se no ajuste do modelo de Hill, sendo o de Michaelis-Menten negligenciado. Para a obtenção dos resultados, as estimativas de  $\kappa$  e  $\nu$  apresentadas por Condomina et al. (2002) foram utilizadas e, a fim de simplificação, considerou-se apenas a porção proximal do intestino. Para  $\gamma$  foi suposto o valor de 1. Tais resultados são apresentados na Tabela 7.

Tabela 7: Delineamentos localmente D-ótimos exatos (número de réplicas entre parênteses) a partir do critério composto com diferentes valores de  $\lambda$  e utilizando-se como informação *a priori* para  $\kappa$  e  $\nu$  as estimativas obtidas por Condomina et al. (2002) para a porção proximal do intestino. Para  $\gamma$  foi suposto o valor de 1.

$\lambda$	$\xi^*$	$\Psi\{\mathbf{M}(\xi^*, \boldsymbol{\theta}_0)\}$
1,0	6,25 (6) 30,00 (6)	-6,502164
0,8	2,55 (2) 7,95 (4) 8,00 30,00 (5)	-7,064889
0,6	2,40 (3) 9,15 (3) 9,20 30,00 (5)	-7,384013
0,5	2,20 (3) 9,35 (3) 9,40 30,00 (5)	-7,523808
0,4	2,10 (3) 9,60 (4) 30,00 (5)	-7,656444
0,2	1,95 (4) 10,05 (4) 30,00 (4)	-7,880106
0,0	1,80 (2) 1,85 (2) 10,25 (2) 10,30 (2) 30,00 (4)	-8,084668

Os delineamentos para  $\lambda$  igual a 0 e 1 são os mesmos que foram encontrados quando procurou-se por delineamentos localmente D-ótimos exatos para cada um dos modelos, separadamente. Também pode ser visto que apenas o primeiro delineamento apresenta 2 pontos de suporte, sendo singular para o modelo de Hill. Os delineamentos restantes possuem pelo menos 3 pontos de suporte, já que é o mínimo exigido pelo modelo de Hill para que a matriz de informação tenha posto

completo e seu determinante não seja nulo. Ademais, vê-se que, desconsiderando arredondamentos causados pela discretização da região experimental, os delineamentos encontrados apresentam 3 pontos de suporte para todos os valores de  $\lambda$  menores que 1. Ou seja, a utilização do critério composto não acarreta no aumento do número de pontos de suporte para localmente D-ótimos.

Pela Figura 17 é possível observar que os dois pontos de suporte inferiores, necessários para estimar os parâmetros no modelo de Hill, parecem tender a um único valor (6, 25) conforme  $\lambda$  aumenta, sendo este valor referente ao menor ponto de suporte do delineamento localmente D-ótimo para o modelo de Michaelis-Menten.

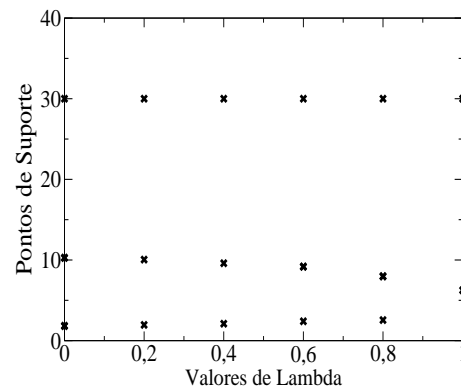


Figura 17: Pontos de suporte de delineamentos localmente D-ótimos exatos a partir do critério composto de D-otimalidade em função de diferentes valores de  $\lambda$ .

Posteriormente, buscou-se por delineamentos D-ótimos exatos pseudo-Bayesianos para  $\lambda$  igual a 0,5 e diferentes coeficientes de variação para a *priori* de  $\kappa$ . Para tanto, foram utilizadas 100 amostras dos valores de  $\kappa$  para o cálculo do critério. Os resultados são apresentados na tabela 8.

Os delineamentos encontrados apresentam uma grande concentração de pontos ao nível 30,00. Além disso, é possível verificar que o aumento no coeficiente de variação de  $\kappa$  ocasiona a diminuição dos valores dos pontos de suporte inferiores dos delineamentos encontrados. Outro fato interessante é que, apesar de ser esperado um maior número de pontos para o caso pseudo-Bayesiano, vê-se apenas

Tabela 8: Delineamentos D-ótimos exatos pseudo-Bayesianos (número de réplicas entre parênteses) a partir do critério D composto com valor de  $\lambda = 0,5$  e diferentes coeficientes de variação para a *priori* de  $\kappa$ , baseando-se nas estimativas obtidas por Condomina et al. (2002).

$C.V.$	$\xi^*$	$\Psi\{\mathbf{M}(\xi^*)\}$
0,05	2,25 (3) 9,40 9,45 (3) 30,00 (5)	-7,636163
0,25	2,15 (2) 2,20 9,25 9,30 (3) 30,00 (5)	-7,539911
0,50	2,00 (2) 2,05 9,10 (4) 30,00 (5)	-7,445280

3 pontos de suporte para cada delineamento quando arredondamentos causados pela discretização da região experimental são desconsiderados.

## 6 CONCLUSÃO E PERSPECTIVAS FUTURAS

Neste trabalho, os algoritmos genético e *exchange* foram comparados em relação à tarefa de construir delineamentos D-ótimos exatos e verificou-se que o *exchange* apresentou um tempo médio de execução menor. No entanto, tal algoritmo deve ser repetido para evitar a obtenção de ótimos locais. Já o genético não precisa ser repetido porque leva em conta uma população de soluções potenciais, fato que diminui a probabilidade de se encontrar um ótimo local. Ainda, os algoritmos apresentaram soluções muito próximas quanto ao valor da função-critério de D-otimalidade. Portanto, não é possível afirmar que um dos algoritmos é o melhor que o outro. Ademais, observou-se que a utilização de ambos propiciou aos resultados maior credibilidade e, dessa maneira, foi de grande importância para a prática.

Assim como proposto, com os algoritmos implementados em linguagem de programação C foi possível obter delineamentos exatos localmente D-ótimos e D-ótimos pseudo-Bayesianos para o modelo de Michaelis-Menten, para o modelo de Hill e para um critério composto visando o ajuste de ambos, considerando-se as estimativas apresentadas por Condomina et al. (2002) para cada uma das porções do intestino e distribuições com diversos valores de coeficiente de variação como informação *a priori* para representar os parâmetros dos modelos.

Em geral, os delineamentos encontrados apresentam número mínimo de pontos de suporte e, desta maneira, impossibilitam o cheque de falta de ajuste. Esta é uma desvantagem bem conhecida de grande parte dos delineamentos D-ótimos apresentados na bibliografia. Algumas alternativas são disponibilizar unidades experimentais a mais com a finalidade de verificar se os modelos ajustados estão coerentes ou, então, construir critérios compostos que incluam outras propriedades de interesse

numa forma de se obter delineamentos com mais pontos e, portanto, mais transigentes.

Ainda, observou-se que grande parte das unidades experimentais pertencentes aos delineamentos encontrados situam-se no extremo superior da região experimental. Tal fato ressalta a importância que medidas a este nível têm na estimação adequada dos parâmetros. Além disso, indica que a escolha da região experimental deve ser realizada com cuidado, pois seu extremo superior tem grande influência na precisão esperada para os estimadores de interesse.

Ademais, verificou-se que os delineamentos encontrados são afetados de maneira considerável pela informação introduzida *a priori*, pois tanto os valores esperados como os coeficientes de variação (para o caso pseudo-Bayesiano) mostram-se correlacionados com os seus pontos de suporte. Assim, conclui-se que a escolha cuidadosa de valores pontuais e distribuições que representem os parâmetros durante a fase de planejamento é fundamental para o sucesso da prática.

Devido à aplicabilidade dos modelos estudados nas mais diversas áreas da biologia, deve-se enfatizar a importância de trabalhos futuros neste contexto. Pode-se considerar, por exemplo, diferentes parametrizações dos modelos e variados tipos de critérios, como por exemplo, a ED-otimalidade, outros critérios compostos, etc. Já para os pseudo-Bayesianos, uma possibilidade que pode tornar o processo de busca mais eficiente em termos de custo computacional é a aproximação para a função-critério a partir de quadraturas.

## 7 ANEXO A

```
/* Este programa obtém o delineamento D-ótimo exato para uma região experimental
 $\chi = [0,05; 30,00]$  discretizada em 600 pontos candidatos fazendo uso do algoritmo
genético. */

/* Bibliotecas utilizadas */
#include<stdio.h>
#include<stdlib.h>
#include<math.h>
#include<time.h>

/* Declaração de constantes */
#define TX 8 // Tamanho do delineamento
#define TY 1 // Número de amostras utilizadas para o método de Monte Carlo (se
igual a 1 - delineamentos locais)
#define PONTOS 600 // Número de pontos candidatos
#define pc 0.5 // Probabilidade de recombinação
#define pp 0.60 // Critério de parada
#define ps 0.25 // Proporção da população atual selecionado para a população
intermediária
#define del 0.0000000000000001 // Diferença utilizada para a comparação
#define modelo 3 // Michaelis-Menten - 1 // Hill - 2 // 3 - Composto
#define alpha 0.5 // Valor de  $\lambda$ 

/* Declaração de funções */
void criterio(int TAM); // Calcula o critério - ver anexo C
```



```

void delinear(int TAM); // Cria a população inicial
void ordenarpopulacao(int TAM); // Ordena a população de acordo com o valor do
critério
void ordenarvetor(int TAM); // Ordena o delineamento para melhor visualização
void reproduzir(int TAM, int SEL); // Cria uma população de clones
void recombinar(double MUT, int TAM); // Recombina e muta os delineamentos
void imprimir(); // Imprime o resultado
/* Declaração de variáveis globais */
double P[PONTOS],K1[20000],X[20000][TX];
double crit[20000],yy[20000],yk[20000][TX],sel[20000];
/* Função principal */
int main(){
int w,t,n,cont,i,j,TAM,SEL;
double auxX1,auxX2,aux1,aux2,xe,MUT;
char nome[100];
srand(85479425); // semente do gerador de números aleatórios
/* Vetor de pontos candidatos */
for(w=0; w<PONTOS; w++){ P[w] = 0.05 + w*0.05;}
TAM=100; // Tamanho da população
SEL=(int)(TAM*ps); // Número de delineamentos selecionados para a reprodução
MUT=0.1; // Probabilidade de mutação
/* Abrindo arquivo para a leitura de valores a priori */
FILE *ar;
sprintf(nome,"nomedoarquivo.dat",nome);
ar=fopen(nome,"r");
for(i=0; i<TY; i++){ fscanf(ar,"%lf",&K1[i]); }
delinear(TAM);
cont=0;
while(cont==0){

```

```

criterio(TAM);
ordenarpopulacao(TAM);
if(crit[0]-crit[(int)(TAM*pp)]<=del){
cont=1;
for(i=0; i<TAM; i++){
ordenarvetor(TAM);
}
}
imprimir();
reproduzir(TAM,SEL);
recombinar(MUT,TAM);
}
fclose(ar); // fechando arquivo
} // Fim função principal
/* Função para gerar a população inicial */
void delinear(int TAM){
double xe;
int aux,i,j;
for(i=0; i<TAM; i++){
for(j=0; j<TX; j++){
xe=PONTOS*(rand()/(RAND_MAX+1.));
aux=(int)xe;
X[i][j]=P[aux];
}}
}
/* Função para ordenar a população */
void ordenarpopulacao(int TAM){
int i,j,k;
double aux1,auxX1;
for(i=0; i<TAM; i++){

```

```

for(k=i+1; k<TAM; k++){
if(crit[i]<crit[k]){
aux1=crit[i];
crit[i]=crit[k];
crit[k]=aux1;
for(j=0; j<TX; j++){
auxX1=X[i][j];
X[i][j]=X[k][j];
X[k][j]=auxX1;
}}}}
/* Cálculo do valor da Função-Critério de D-otimalidade */
void criterio(int TAM){
int i,j,k,w;
double aux,aux2,A[TY],B[TY],C[TY],D[TY],E[TY],F[TY],DET[TY],V;
double WGAMA[4],GAMA[4],nw,soma;
V=8.39; // Esperanca de  $\nu$ 
if(modelo==1){
for(k=0;k<TAM;k++){
if(TY==1){ A[0]=B[0]=C[0]=0.0; K1[0]=10.78;} // Esperanca de  $\kappa$ 
else{
for(i=0;i<TY;i++){
A[i]=B[i]=C[i]=0.0;
}}
aux=0;
for(i=0;i<TY;i++){
for(j=0;j<TX;j++){
A[i]+= pow(X[k][j],2.)/(pow((K1[i]+X[k][j]),2.)*TX);
B[i]+= pow(X[k][j],2.)/(pow((K1[i]+X[k][j]),4.)*TX);
C[i]+= pow(X[k][j],2.)/(pow((K1[i]+X[k][j]),3.)*TX);

```

```

}
DET[i]= pow(V,2.)*(A[i]*B[i]-pow(C[i],2.));
aux+= log(DET[i]);
}
crit[k]=aux/TY;
}}
else if(modelo==2){
for(k=0;k<TAM;k++){
if(TY==1){A[0]=B[0]=C[0]=D[0]=E[0]=F[0]=0.0;
GAMA[0]=1.0; WGAMA[0]=1; K1[0]=10.78; nw=1;}
else{
nw=4;
WGAMA[0]=0.25; WGAMA[1]=0.25; WGAMA[2]=0.25; WGAMA[3]=0.25;
GAMA[0]=1; GAMA[1]=2; GAMA[2]=3; GAMA[3]=4;
}
aux=0.;
for(w=0;w<nw;w++){
soma=0;
for(i=0;i<TY;i++){
A[i]=B[i]=C[i]=D[i]=E[i]=F[i]=0.0;
for(j=0;j<TX;j++){
A[i]+= pow(X[k][j],2.*GAMA[w])/(pow((pow(K1[i],GAMA[w])
+pow(X[k][j],GAMA[w])),2.)*TX);
B[i]+= -GAMA[w]*V*pow(K1[i],GAMA[w]-1)*pow(X[k][j],2.*GAMA[w])
/(pow((pow(K1[i],GAMA[w])+pow(X[k][j],GAMA[w])),3.)*TX);
C[i]+= V*pow(K1[i],GAMA[w])*pow(X[k][j],2.*GAMA[w])*(log(X[k][j])
-log(K1[i]))/(pow((pow(K1[i],GAMA[w])+pow(X[k][j],GAMA[w])),3.)*TX);
D[i]+= pow(GAMA[w],2.)*pow(V,2.)*pow(K1[i],2*GAMA[w]-2)*pow(X[k][j],
2.*GAMA[w])/(pow((pow(K1[i],GAMA[w])+pow(X[k][j], GAMA[w])),4.)*TX);

```

```

E[i]+=-GAMA[w]*pow(V,2.)*pow(K1[i],2*GAMA[w]-1)*pow(X[k][j],2.*GAMA[w])
(log(X[k][j])-log(K1[i]))/(pow((pow(K1[i],GAMA[w])+pow(X[k][j],GAMA[w])),4.)*TX);
F[i]+= pow(V,2.)*pow(K1[i],2*GAMA[w])*pow(X[k][j],2.*GAMA[w])*pow(log(X[k][j])-
log(K1[i]),2.)/(pow((pow(K1[i],GAMA[w])+pow(X[k][j],GAMA[w])),4.)*TX);
}
DET[i]= A[i]*D[i]*F[i]+2*B[i]*C[i]*E[i]-pow(C[i],2.)*D[i]
-pow(B[i],2.)*F[i]-pow(E[i],2.)*A[i]; soma += log(DET[i]);
}
aux+=WGAMA[w]*soma;
}
crit[k]= aux/TY;
}}
else{
for(k=0;k<TAM;k++){
if(TY==1){ A[0]=B[0]=C[0]=0.0; K1[0]=10.78;} // Esperanca de  $\kappa$ 
else{
for(i=0;i<TY;i++){
A[i]=B[i]=C[i]=0.0;
}}
aux=0;
for(i=0;i<TY;i++){
for(j=0;j<TX;j++){
A[i]+= pow(X[k][j],2.)/(pow((K1[i]+X[k][j]),2.)*TX);
B[i]+= pow(X[k][j],2.)/(pow((K1[i]+X[k][j]),4.)*TX);
C[i]+= pow(X[k][j],2.)/(pow((K1[i]+X[k][j]),3.)*TX);
}
DET[i]= pow(V,2.)*(A[i]*B[i]-pow(C[i],2.));
aux+= log(DET[i]);
}
}

```

```

if(TY==1){A[0]=B[0]=C[0]=D[0]=E[0]=F[0]=0.0;
GAMA[0]=1.0; WGAMA[0]=1; K1[0]=10.78; nw=1;}
else{
nw=4;
WGAMA[0]=0.25; WGAMA[1]=0.25; WGAMA[2]=0.25; WGAMA[3]=0.25;
GAMA[0]=1; GAMA[1]=2; GAMA[2]=3; GAMA[3]=4;
}
aux2=0.;
for(w=0;w<nw;w++){
soma=0;
for(i=0;i<TY;i++){
A[i]=B[i]=C[i]=D[i]=E[i]=F[i]=0.0;
for(j=0;j<TX;j++){
A[i]+= pow(X[k][j],2.*GAMA[w])/(pow((pow(K1[i],GAMA[w])+
pow(X[k][j],GAMA[w])),2.)*TX);
B[i]+=-GAMA[w]*V*pow(K1[i],GAMA[w]-1)*pow(X[k][j],2.*GAMA[w])
/(pow((pow(K1[i],GAMA[w])+pow(X[k][j],GAMA[w])),3.)*TX);
C[i]+= V*pow(K1[i],GAMA[w])*pow(X[k][j],2.*GAMA[w])*(log(X[k][j])
-log(K1[i]))/(pow((pow(K1[i],GAMA[w])+pow(X[k][j],GAMA[w])),3.)*TX);
D[i]+= pow(GAMA[w],2.)*pow(V,2.)*pow(K1[i],2.*GAMA[w]-2)*pow(X[k][j],
2.*GAMA[w])/(pow((pow(K1[i],GAMA[w]) +pow(X[k][j], GAMA[w])),4.)*TX);
E[i]+=-GAMA[w]*pow(V,2.)*pow(K1[i],2.*GAMA[w]-
1)*pow(X[k][j],2.*GAMA[w])*(log(X[k][j])
-log(K1[i]))/(pow((pow(K1[i],GAMA[w])+pow(X[k][j],GAMA[w])),4.)*TX);
F[i]+= pow(V,2.)*pow(K1[i],2.*GAMA[w])*pow(X[k][j],2.*GAMA[w])*pow(log(X[k][j])-
log(K1[i]),2.)/(pow((pow(K1[i],GAMA[w])+pow(X[k][j],GAMA[w])),4.)*TX);
}
DET[i]=A[i]*D[i]*F[i]+2*B[i]*C[i]*E[i]-pow(C[i],2.)*D[i]-
pow(B[i],2.)*F[i]-pow(E[i],2.)*A[i];

```

```

soma += log(DET[i]);
}
aux2+=WGAMA[w]*soma;
}
crit[k]=alpha*(aux/TY)+(1-alpha)*(aux2/TY);
}}

/* Função para ordenar o delineamento e melhorar a visualização */
void ordenarvetor(int TAM){
int i,k,g;
double aux1;
for(g=0; g<TAM; g++){
for(i=0; i<TX; i++){
for(k=i+1; k<TX; k++){
if(X[g][i]>X[g][k]){
aux1=X[g][i];
X[g][i]=X[g][k];
X[g][k]=aux1;
}}}}
}

/* Cria uma populacao de clones dos pais de acordo com o valor de rank */
void reproduzir(int TAM, int SEL){
int i,k,j,cont;
double aux1,xe,sum,aux2;
sum=0;
for(k=0; k<SEL; k++){
for(j=0; j<TX; j++){
yk[k][j]=X[k][j]; // Alocando delineamentos na população intermediária
}
yy[k]=crit[k]; // Alocando critério num vetor

```

```

sel[k]=(SEL-k); // Rank para os delineamentos
sum+=(SEL-k); // Soma dos ranks
}
for(j=0; j<TAM; j++){
xe=(1.*rand()/(RAND_MAX));
aux1=0; aux2=0; cont=0;
for(k=0; k<SEL; k++){ aux2=aux1;
aux1+=(sel[k]/sum);
if((xe<=aux1)&&(xe>aux2)){
cont++;
crit[j]=yy[k];
for(i=0; i<TX; i++){ X[j][i]=yk[k][i];
} break;} }}
/* função que executa a recombinação e mutação dos delineamentos */
void recombinar(double MUT, int TAM){
int i,kk,j,aux,a,cont;
double xe,auxX1;
/* Recombinação */
for(i=0; i<TAM; i+=2){
xe = rand()/(RAND_MAX+1.);
kk=xe*TX; // Ponto do delineamento escolhido para troca
for(j=0; j<TX; j++){
xe =rand()/(RAND_MAX+1.); // Probabilidade de cruzar os genes
if((j>kk)&&(xe<pc)){
auxX1=X[i][j]; X[i][j]=X[i+1][j]; X[i+1][j]=auxX1;
}
}
/* Mutação */
xe=rand()/(RAND_MAX+1.);
if(xe<MUT){

```



```

for(a=i; a<i+2; a++){
xe=rand()/(RAND_MAX+1.);
aux=(int)(xe*(PONTOS+1.));
X[a][j]=P[aux]; // Valor atribuído ao gene trocado
}}}}
/* Impressão do delineamento D-ótimo e do critério D */
void imprimir(){
int i;
for(i=0; i<TX; i++){
printf("%.2lf",X[0][i]);
}
printf(" crit=-%lf",crit[0]);
}

```

## 8 ANEXO B

```

/* Este programa obtém o delineamento D-ótimo exato para uma região experimental
 $\chi = [0, 05; 30, 00]$  discretizada em 600 pontos candidatos fazendo uso do algoritmo
exchange. */
/* Bibliotecas utilizadas */
#include<stdio.h>
#include<stdlib.h>
#include<math.h>
#include<time.h>
#define modelo 3 // Michaelis-Menten - 1 // Hill - 2 // 3 - Composto
#define alpha 0.5 // Valor de  $\lambda$ 

```

```

/* Declaração de constantes */
#define PONTOS 600 // Número de pontos candidatos
#define TX 8 // Tamanho do delineamento
#define TY 100 // Tamanho da amostra de prioris
/* Declaração de funções */
void leitura(); // Para ler as amostras
void pontoscandidatos(); // Cria um conjunto de pontos candidatos
void delineamentoinicial(); // Cria um delineamento atual (corrente)
void ordena(); // Ordena o valores do delineamento
double criterio(); // Calcula o critério
void impressao(double crit); // Imprime o delineamento obtido e o critério
/* Declaração de variáveis globais */
double K1[5000],P[PONTOS],X[TX],CRITE[PONTOS];
/* Função principal */
main(){
int cont,i,j,aux,k,t,pos;
double dif,xe,faux,fk1,crit,aux2;
/* Realiza a leitura dos dados */
leitura(TY);
/* Cria o vetor de pontos candidatos */
pontoscandidatos();
/* Cria um delineamento inicial */
delineamentoinicial();
/* Calcula o critério para o delineamento corrente */
fk1=criterio();
/* O contador é usado para avaliar se houve algum ponto do delineamento trocado
durante a simulação */
cont=1;
while(cont!=0){

```

```

cont=0;
/* Variável recebe o primeiro ponto do delineamento corrente */
faux=X[0];
/* Todos os pontos candidatos são avaliados e aquele que propiciar maior aumento
na função critério, se houver tal ponto, é aceito */
for(k=0; k<PONTOS; k++){
if(faux!=P[k]){X[0]=P[k]; CRITE[k]=criterio(); }
else{CRITE[k]=fk1;}
}
pos=0;
for(k=1; k<PONTOS; k++){
if(CRITE[k]>CRITE[pos]){pos=k;}
}
if(CRITE[pos]<=fk1){X[0]=faux; aux2=faux; }
else{X[0]=P[pos]; aux2=P[pos]; cont=1; fk1=CRITE[pos]; }
/* O processo é repetido para todos os outros pontos do delineamento */
for(t=1; t<TX; t++){
if(X[t]!=aux2){ /* Mas se o ponto for igual ao que foi aceito anteriormente, ele não
será retirado */
faux=X[t];
for(k=0; k<PONTOS; k++){
if(faux!=P[k]){X[t]=P[k]; CRITE[k]=criterio();}
else{CRITE[k]=fk1;}
}
pos=0;
for(k=1; k<PONTOS; k++){
if(CRITE[k]>CRITE[pos]){pos=k;}
}
if(CRITE[pos]<=fk1){X[t]=faux; aux2=faux;}

```

```

else{X[t]=P[pos]; aux2=P[pos]; cont=1; fk1=CRITE[pos]; }
}
/* ordena o vetor após uma sequência de trocas */
ordena();
}}
crit=criterio();
impressao(crit);
}
/* Função para ler os dados */
void leitura(){
int i;
FILE *ar;
ar=fopen("nomedoarquivo.dat","r");
for(i=0; i<TY; i++){
fscanf(ar,"%lf",&K1[i]);
}}
/* Cálculo do valor da Função-Critério de D-otimalidade */
double criterio(){
int i,j,k,w;
double CV,VarV,V,criterio1,criterio2,composto;
double A[TY],B[TY],C[TY],D[TY],E[TY],F[TY],DET[TY];
double WGAMA[5],GAMA[5],nw,soma;
V=8.39; // Esperanca de  $\nu$ 
if(modelo==1){
criterio1=0.;
if(TY==1){A[0]=B[0]=C[0]=0.0; K1[0]=10.78;}
else{
for(k=0;k<TY;k++){
A[k]=B[k]=C[k]=0.0;

```

```

}}
for(i=0;i<TY;i++){
for(j=0;j<TX;j++){
A[i]+= pow(X[j],2.)/(pow((K1[i]+X[j]),2.)*TX);
B[i]+= pow(X[j],2.)/(pow((K1[i]+X[j]),4.)*TX);
C[i]+= pow(X[j],2.)/(pow((K1[i]+X[j]),3.)*TX);
}
DET[i]= pow(V,2.)*(A[i]*B[i]-pow(C[i],2.));
criterio1 += log(DET[i]);
}
return criterio1/TY;
}
else if(modelo==2){
if(TY==1){A[0]=B[0]=C[0]=D[0]=E[0]=F[0]=0.0;
GAMA[0]=1.0; WGAMA[0]=1; K1[0]=10.78; nw=1;}
else{
nw=5; WGAMA[0]=0.05; WGAMA[1]=0.2; WGAMA[2]=0.5; WGAMA[3]=0.2;
WGAMA[4]=0.05;
GAMA[0]=0.5; GAMA[1]=0.75; GAMA[2]=1.0; GAMA[3]=1.25; GAMA[4]=1.5;
}
criterio1=0.;
for(w=0;w<nw;w++){
soma=0;
for(i=0;i<TY;i++){
A[i]=B[i]=C[i]=D[i]=E[i]=F[i]=0.0;
for(j=0;j<TX;j++){
A[i]+= pow(X[j],2.*GAMA[w])/(pow((pow(K1[i],GAMA[w])+pow(X[j],GAMA[w])),2.)*TX);
B[i]+= -GAMA[w]*V*pow(K1[i],GAMA[w]-1)*pow(X[j],2.*GAMA[w])/
(pow((pow(K1[i],GAMA[w]) +pow(X[j],GAMA[w])),3.)*TX);

```

```

C[i]+=V*pow(K1[i],GAMA[w])*pow(X[j],2.*GAMA[w])*(log(X[j])-log(K1[i]))/
(pow((pow(K1[i],GAMA[w])+pow(X[j],GAMA[w])),3.)*TX);
D[i]+=pow(GAMA[w],2.)*pow(V,2.)*pow(K1[i],2*GAMA[w]-
2)*pow(X[j],2.*GAMA[w])/(pow((pow(K1[i],
GAMA[w])+pow(X[j], GAMA[w])),4.)*TX);
E[i]+=-GAMA[w]*pow(V,2.)*pow(K1[i],2*GAMA[w]-
1)*pow(X[j],2.*GAMA[w])*(log(X[j])
-log(K1[i]))/(pow((pow(K1[i],GAMA[w])+pow(X[j],GAMA[w])),4.)*TX);
F[i]+= pow(V,2.)*pow(K1[i],2*GAMA[w])*pow(X[j],2.*GAMA[w])*pow(log(X[j])-
log(K1[i]),2.)/(pow((pow(K1[i],GAMA[w])+pow(X[j],GAMA[w])),4.)*TX);
}
DET[i]= A[i]*D[i]*F[i]+2*B[i]*C[i]*E[i]-pow(C[i],2.)*D[i]
-pow(B[i],2.)*F[i]-pow(E[i],2.)*A[i];
soma += log(DET[i]);
}
criterio1+=WGAMA[w]*soma;
}
return criterio1/TY;
}
else if(modelo==3){
criterio1=0.;
if(TY==1){A[0]=B[0]=C[0]=0.0; K1[0]=10.78;}
else{
for(k=0;k<TY;k++){
A[k]=B[k]=C[k]=0.0;
}}
for(i=0;i<TY;i++){
for(j=0;j<TX;j++){
A[i]+= pow(X[j],2.)/(pow((K1[i]+X[j]),2.)*TX);

```

```

B[i] += pow(X[j],2.)/(pow((K1[i]+X[j]),4.)*TX);
C[i] += pow(X[j],2.)/(pow((K1[i]+X[j]),3.)*TX);
}
DET[i] = pow(V,2.)*(A[i]*B[i]-pow(C[i],2.));
criterio1 += log(DET[i]);
}
if(TY==1){A[0]=B[0]=C[0]=D[0]=E[0]=F[0]=0.0;
GAMA[0]=1.0; WGAMA[0]=1; K1[0]=10.78; nw=1;}
else{
nw=5; WGAMA[0]=0.05; WGAMA[1]=0.2; WGAMA[2]=0.5; WGAMA[3]=0.2;
WGAMA[4]=0.05;
GAMA[0]=0.5; GAMA[1]=0.75; GAMA[2]=1.0; GAMA[3]=1.25; GAMA[4]=1.5;
}
criterio2=0.;
for(w=0;w<nw;w++){
soma=0;
for(i=0;i<TY;i++){
A[i]=B[i]=C[i]=D[i]=E[i]=F[i]=0.0;
for(j=0;j<TX;j++){
A[i] += pow(X[j],2.*GAMA[w])/ (pow((pow(K1[i],GAMA[w])+pow(X[j],GAMA[w])),2.)*TX);
B[i] += -GAMA[w]*V*pow(K1[i],GAMA[w]-1)*pow(X[j],2.*GAMA[w])
/(pow((pow(K1[i],GAMA[w])+pow(X[j],GAMA[w])),3.)*TX);
C[i] += V*pow(K1[i],GAMA[w])*pow(X[j],2.*GAMA[w])*(log(X[j])-log(K1[i]))
/(pow((pow(K1[i],GAMA[w])+pow(X[j],GAMA[w])),3.)*TX);
D[i] += pow(GAMA[w],2.)*pow(V,2.)*pow(K1[i],2*GAMA[w]-
2)*pow(X[j],2.*GAMA[w])
/(pow((pow(K1[i],GAMA[w])+pow(X[j], GAMA[w])),4.)*TX);
E[i] += -GAMA[w]*pow(V,2.)*pow(K1[i],2*GAMA[w]-
1)*pow(X[j],2.*GAMA[w])*(log(X[j])

```

```

-log(K1[i]))/(pow((pow(K1[i],GAMA[w])+pow(X[j],GAMA[w])),4.)*TX);
F[i]+= pow(V,2.)*pow(K1[i],2*GAMA[w])*pow(X[j],2.*GAMA[w])*pow(log(X[j])-
log(K1[i]),2.)/(pow((pow(K1[i],GAMA[w])+pow(X[j],GAMA[w])),4.)*TX);
}
DET[i]=A[i]*D[i]*F[i]+2*B[i]*C[i]*E[i]-pow(C[i],2.)*D[i]-pow(B[i],2.)*F[i]-
pow(E[i],2.)*A[i]; soma += log(DET[i]);
}
criterio2+=WGAMA[w]*soma;
}
composto= alpha*(criterio1/TY)+(1-alpha)*(criterio2/TY);
return composto;
}}
/* Função para criar o conjunto de pontos candidatos */
void pontoscandidatos(){
int i;
for(i=0; i<PONTOS; i++){
P[i]=0.05+i*0.05;
}}
/* Função para criar o delineamento atual ou corrente, que será melhorado durante
o processo */
void delineamentoinicial(){
double xe;
int aux,j;
for(j=0; j<TX; j++){
xe=1.*rand()/(RAND_MAX+1.);
aux=(int)((PONTOS)*xe);
X[j]=P[aux];
}}
/* Ordena o vetor para ser realizada uma nova sequência de trocas. Faz com que o

```



```
algoritmo convirja mais rápido */
void ordena(){
int i,k;
double auxX1;
for(i=0; i<TX; i++){
for(k=i+1; k<TX; k++){
if(X[i]>X[k]){
auxX1=X[i]; X[i]=X[k]; X[k]=auxX1; }}}
/* Esta função imprime o resultado */
void impressao(double crit){
int i;
for(i=0; i<TX; i++){
printf("%.2lf ",X[i]); }
printf("%.2lf ",crit);
}
```

## REFERÊNCIAS BIBLIOGRÁFICAS

- AKAIKE, H. An information criterion, AIC. **Mathem. Sci.**, v.14, p.5–9, 1986.
- ATKINSON, A. C.; DONEV, A. N.; TOBIAS, R. D. **Optimum Experimental Designs, with SAS**. Oxford Press, 2007.
- BOX, G. E. P.; COX, D. R. An Analysis of Transformations. **Journal of the Royal Statistical Society**, v.26, p.211–252, 1964.
- BOX, G. E. P.; HUNTER, J. S.; HUNTER, W. G. **Statistics for Experimenters**. New York: Wiley, 1978.
- BOX, G. E. P.; LUCAS, H. L. Design of Experiments in Non-Linear Situations. **Biometrika**, v.46, p.77–90, 1959.
- CHALONER, K.; VERDINELLI, I. Bayesian experimental design - a review. **Statistical Science**, v.10, n.3, p.273–304, 1995.
- CHAN, L.-Y. Optimal design for estimation of variance in nonparametric regression using first order differences. **Biometrika**, v.78, n.4, p.926–9, 1991.
- COCHRAN, W. G.; COX, G. M. **Experimental Designs**. New York: Wiley, 1957.
- CONDOMINA, J.; ZORNOZA-SABINA, T.; GRANERO, L.; POLACHE, A. Kinetics of zinc transport in vitro in rat small intestine and colon: interaction with copper. **European Journal of Pharmaceutical Sciences**, v.16, p.289–295, 2002.
- DETTE, H.; KISS, C.; WONG, W. K. Robustness of Optimal Designs for the Michaelis-Menten Model under a Variation of Criteria. **Fakultat Statistik, technische universität dortmund, Germany**, 2008.

FEDOROV, V. V. **Theory of Optimal Experiments**. New York: Academic Press, 1972.

FISHER, R. A. **The Design of Experiments**. Edinburgh: Oliver & Boyd, 1935.

FISHER, R. A. **The Design of Experiments**. 9. ed. Hafner Press, 1971.

FOX, J. **Introduction to Nonparametric Regression**. ESRC Oxford Spring School, 2005.

GOLDBERG, D. E. **Genetic Algorithms in Search, Optimization and Machine Learning**. Massachusetts: Addison-Wesley Co, 1989.

GRAHAM, G.; GUEORGUIEVA, I.; DICKENS, K. A program for the optimum design of pharmacokinetic, pharmacodynamic, drug metabolism and drug-drug interaction models. **Computer Methods and Programs in Biomedicine**, v.78, p.237–249, 2005.

HOLLAND, J. H. **Adaptation in Natural and Artificial Systems**. Ann Arbor: University of Michigan Press, 1975.

KIEFER, J. Optimal experimental designs (with discussion). **Journal of the Royal Statistical Society B**, v.21, n.21, p.272–319, 1959.

KIEFER, J.; WOLFOWITZ, J. The Equivalence of two extremum problems. **Canadian Journal of Statistics**, v.12, p.363–366, 1960.

MOTULSKY, H.; CHRISTOPOULOS, A. **Fitting Models to Biological Data Using Linear and Nonlinear Regression**. Oxford University Press, 2004.

PINTO, E. D.; PONCE DE LEON, A. C. M. **Planejamento Ótimo de Experimentos**. ABE, 2006.

PONCE DE LEON, A. C. M. Optimum Experimental Design for Model Discrimination and Generalized Linear Models. London, England, 1993. Thesis(Ph.D.) - London School of Economics.

PRONZATO, L.; WALTER, E. Robust Experimental Designs via Stochastic Approximation. **Mathematical Biosciences**, v.75, p.103–120, 1985.

PUKELSHEIM, F. **Optimal Design of Experiments**. New York: Wiley, 1993.

TRINCA, L. A.; GILMOUR, S. G. Optimal designs for fitting fractional polynomial models. **Rel. Téc.**, 2001.

WOODS, D. C. An Exchange Algorithm to find Designs under Generalized Linear Models. **Southampton Statistical Sciences Research Institute University of Southampton**, 2005.