

RAE – CEA – 07P27

**RELATÓRIO DE ANÁLISE ESTATÍSTICA SOBRE O
PROJETO: “EFEITO DA MODULAÇÃO
DOPAMINÉRGICA NO COMPORTAMENTO DE
RATOS NO JOGO *MATCHING PENNIES*”**

**Adilson Simonis
Marcelo Henrique do Amaral**

São Paulo, dezembro de 2007

CENTRO DE ESTATÍSTICA APLICADA – CEA – USP
RELATÓRIO DE ANÁLISE ESTATÍSTICA – CÓDIGO 07P27

TÍTULO: Efeito da modulação dopaminérgica no comportamento de ratos no jogo “Matching pennies”.

PESQUISADOR: Luiz Eduardo Tassi.

INSTITUIÇÃO: Instituto de Biociências da Universidade de São Paulo (IB-USP).

FINALIDADE: Mestrado.

RESPONSÁVEIS PELA ANÁLISE: Adilson Simonis e Marcelo Henrique do Amaral.

REFERÊNCIA DESTE TRABALHO: Amaral, M. H., Simonis, A. **Relatório de Análise Estatística sobre o projeto efeito da modulação dopaminérgica no comportamento de ratos no jogo “Matching pennies”**. São Paulo, IME-USP, 2007 (RAE – CEA – 07P27).

FICHA TÉCNICA

REFERÊNCIAS BIBLIOGRÁFICAS

Akaike H. (1973). Information theory and an extension of the maximum likelihood principle. **Second International Symposium on Information Theory**. 267-281.

BarracloUGH, D.J., Conroy, M.L. and Lee, D., (2004). Prefrontal cortex and decision making in a mixed-strategy game. **Nature Neuroscience**, **7**, 404 – 410.

Bortolossi, H., Garbaggio, G. and Sartini, B., (2007). **Uma introdução à teoria econômica dos jogos**. Rio de Janeiro: IMPA, 185p.

Cohen, M.X. and Ranganath, C., (2007). Reinforcement learning predict future decisions. **The Journal of Neuroscience**, **10**, 371 – 378.

Curi, E. J. M. (2003) **Estimação do alcance de cadeias de Markov**, São Paulo. 95p. Dissertação (Mestrado). Instituto de Matemática e Estatística - USP.

Dorea, C., Galves, A., Kira, E. e A. Pereira Alencar (1997). Markovian modeling of the stress contours of Brazilian and European Portuguese. **REBRAPE**, **11**, 161-173.

Lee, D., Conroy, M.L., McGreevy, B.P. and BarracloUGH, D.J., (2004). Reinforcement learning and decision making in monkeys during a competitive game. **Cognitive Brain Research**, **22**, 45 – 58.

Lehmann, E. and Casella, G. (1998). **Theory of point estimation**. 2.ed. New York: Springer-Verlag. 589p.

Katz, R.W. (1981) On some criteria for estimating the order of a Markov chain. **Technometrics**, **23**, 243 – 249.

Kullback, S. (1959). **Information theory and Statistics**. New York: Wiley. 395p.

Miller, G.A., (1955) Note on the bias of information estimates. **Information theory on Psychology**. 95 – 100.

Montague, P.R., Hyman, S.E. and Cohen, J.D., (2004) Computational roles for dopamine in behavioural control. **Nature**, **431**, 760 – 766.

Petrielli, A. (2004). **Reamostragem em blocos para cadeias de Markov**. São Paulo. 63p. Dissertação (Mestrado). Instituto de Matemática e Estatística - USP.

Schultz, W., Dayan, P. and Montague, P.R., (1997). A neural substrate of prediction and reward. **Science**, **275**. 1593 – 1599.

Schwarz, G. (1978). Estimating the dimension of a model. **Annals of Statistics**, **6**, 461 – 464.

Soltani, A., Lee, D. and Wang, X.J., (2006). Neural mechanism for stochastic behaviour during a competitive game. **Neural Networks**, **19**, 1075 – 1090.

PROGRAMAS COMPUTACIONAIS UTILIZADOS

Microsoft Excel (versão 2007);

R 2.6.0;

Python 2.5;

TÉCNICAS ESTATÍSTICAS UTILIZADAS

Análise Descritiva Unidimensional (03:010)

Testes de Hipóteses Paramétricas (05:010)

Processos Estocásticos (11:000)

ÁREA DE APLICAÇÃO

Bioestatística (14:030)

1. INTRODUÇÃO	1
2. DESCRIÇÃO DO ESTUDO	2
2.1. ESQUEMATIZAÇÃO DO ESTUDO	2
2.2. ALGORITMOS COMPUTACIONAIS	3
2.2.1. ALGORITMO 0	3
2.2.2. ALGORITMO 1	3
2.2.3. ALGORITMO 2	3
3. DESCRIÇÃO DAS VARIÁVEIS	4
4. ANÁLISE DESCRITIVA	4
4.1. BLOCOS DE OBSERVAÇÃO	4
4.2. ANÁLISES DE SIMILARIDADE E DIFERENÇA	4
4.3. ANÁLISE DE DEPENDÊNCIA E INDEPENDÊNCIA	4
5. ANÁLISE INFERENCIAL	5
5.1. ANÁLISE DAS ESCOLHAS	5
5.2. TESTES DE HOMOGENEIDADE	5
5.3. EFEITO DA ANTEFAZ DA MONTAGEM EM APENAS UMA PASTORA	5
6. CONCLUSÃO	6
APÊNDICE A	6
APÊNDICE B	6
APÊNDICE C	6
APÊNDICE D	6
APÊNDICE E	6
APÊNDICE F	6
APÊNDICE G	6

ÍNDICE

RESUMO	6
1. INTRODUÇÃO	7
2. DESCRIÇÃO DO ESTUDO.....	8
2.1. PROCEDIMENTO EXPERIMENTAL.....	9
2.2. ALGORITMOS COMPUTACIONAIS	10
2.2.1 ALGORITMO 0	10
2.2.2 ALGORITMO 1	10
2.2.3. ALGORITMO 2	11
3. DESCRIÇÃO DAS VARIÁVEIS.....	12
4. ANÁLISE DESCRITIVA.....	12
4.1. BLOCOS DE OBSERVAÇÕES.....	12
4.2. ANÁLISES DE ESCOLHAS E RECOMPENSAS.....	13
4.3. ANÁLISE DE DEPENDÊNCIA E ESTRATÉGIA.....	15
5. ANÁLISE INFERENCIAL	16
5.1. ANÁLISE DAS ESCOLHAS.....	17
5.2. TESTES DE HOMOGENEIDADE.....	19
5.3. EFEITO DA ANFETAMINA NO DESEMPENHO DOS RATOS	20
6. CONCLUSÃO	20
APÊNDICE A	22
APÊNDICE B	24
APÊNDICE C	30
APÊNDICE D	38
APÊNDICE E	46
APÊNDICE F.....	54
APÊNDICE G.....	56

RESUMO

Para estudar o processo de tomada de decisão em ratos em uma situação em que, de acordo com suas escolhas, eles podem ser recompensados ou não, foi realizado um experimento em que ratos “jogam um jogo” conhecido como “matching pennies”, contra um computador.

Neste jogo, os ratos devem escolher uma, e somente uma, entre duas opções possíveis. Para maximizar suas chances de vitória, os ratos tendem a fazer suas escolhas de forma aleatória e com igual probabilidade para as duas opções.

Neste experimento, três ratos jogaram uma série de vezes antes e depois de receberem uma dose de 0,5 ml/kg de anfetamina. O objetivo foi o de se estudar o efeito dessa substância sobre o processo de tomada de decisão.

Utilizando cadeias de Markov para descrever o processo de tomada de decisão dos ratos, concluímos que, estes fazem suas escolhas baseados no máximo em duas de suas últimas escolhas. A anfetamina diminui a influência de escolhas observadas na escolha futura.

1. INTRODUÇÃO

A tomada de decisão em um determinado processo consiste em selecionar uma particular ação de um conjunto de alternativas. Quando é fixada a relação entre cada ação e sua utilidade (resultados possíveis), a decisão de escolher a ação com a máxima utilidade é considerada ótima ou racional [Barraclough et al., 2004]. Recentemente, as bases neurais de indivíduos submetidos a este tipo de processo estão sendo estudadas em situações que simulam diferentes cenários. A literatura neste caso é extensa [Lee et al., 2004 e Soltani et al., 2006]. Em particular, ambientes onde cada indivíduo de um grupo toma decisões com base em experiências anteriores, dele e dos demais integrantes do grupo, estão tornando-se freqüentes nas pesquisas. É comum encontrarmos experimentos em que cada decisão atribui uma recompensa ou punição (utilidade positiva ou negativa) ao indivíduo ou ao grupo de indivíduos do estudo.

A teoria dos jogos é uma ferramenta adequada para estudar o processo de tomada de decisão e pode ser entendida como a base teórica dos modelos estatísticos que descrevem o processo da escolha de decisões ótimas sob condições de conflito. O estudo do processo de aprendizado de cada indivíduo e do grupo de indivíduos a fim de adotar a melhor estratégia (decisão ótima), está no escopo da teoria dos jogos.

Diversos estudos [Schultz et al., 1997, Cohen et al., 2007 e Waelti et al., 2001] sugerem que sinais liberados pelo sistema dopaminérgico estão envolvidos com o processo de aprendizado por recompensa. De fato, a teoria de aprendizado por recompensa busca explicar como indivíduos aprendem a organizar seu comportamento em uma dada situação sob a influência de recompensas [Montague et al., 2004].

Os sinais dopaminérgicos mais importantes para esta teoria são o fásico, que informa sobre a discrepância entre a expectativa de recompensa e a recompensa encontrada, e o tônico, que entre outras coisas, codifica uma "taxa média de recompensas" [Montague et al., 2004].

A quantidade de sinais liberados e os tempos de efeito destes podem ser ampliados através de drogas como a cocaína e a anfetamina. A consequência é uma alteração da capacidade do indivíduo em julgar o valor da recompensa recebida por

cada decisão tomada, e assim, prejudicando ou potencializando seu aprendizado por recompensa [Montague et al., 2004].

Neste relatório propomos uma análise dos processos de tomada de decisão de ratos enquanto estes jogam uma série de vezes, contra algoritmos computacionais, um jogo simples conhecido como “matching pennies”. Este é um jogo que exige tomada de decisão e distribui recompensa ao vencedor.

O jogo não possui uma ação ótima *pura* (ação que sempre atribui maior recompensa do que as demais, independente da ação realizada pelo oponente) [Bortolossi et al., 2007]. Neste caso, o jogo possui uma decisão ótima em *estratégia mista*, que define uma distribuição de probabilidade para a escolha das possíveis ações, o que pode exigir uma aleatoriedade na escolha da ação.

No “matching pennies”, a estratégia mista ideal (decisão ótima) é a que escolhe uma das ações de forma aleatória e com igual probabilidade. Deseja-se verificar então, se os ratos são capazes de aprender a jogar de forma ideal com base nas recompensas recebidas em cada vitória. Em outras palavras, deseja-se estudar o processo de aprendizado por recompensa em ratos neste jogo.

Além das tarefas descritas acima, cada rato também realizou uma série de jogos após ingerir 0,5 ml/kg de anfetamina. O objetivo é o de se estudar o efeito da maior liberação e maior tempo de duração dos sinais dopaminérgicos sob o processo de aprendizado por recompensa, de tal forma a alterar a proporção de recompensas dos ratos.

O principal objetivo deste trabalho é o estudo e modelagem da estocasticidade das escolhas dos ratos, pois com isso, poderemos comparar o processo de tomada de decisão dos ratos em diferentes situações.

As seqüências de escolhas dos ratos foram analisadas por cadeias de Markov, que consideram a possível dependência existente entre sucessivas escolhas.

2. DESCRIÇÃO DO ESTUDO

Na próxima subseção descrevemos o procedimento experimental e na subseção seguinte são descritos os algoritmos computacionais.

2.1. PROCEDIMENTO EXPERIMENTAL

Três ratos da linhagem Wistar foram utilizados no estudo, codificados como Rato 12, Rato 22 e Rato 23. Para realizar o experimento, foi utilizada uma caixa com dois compartimentos, um acima do outro. Um rato era colocado no compartimento superior da caixa na posição central. Este compartimento possui três buracos, um à frente, outro à esquerda e outro à direita do local onde o rato foi colocado, e cada buraco contém uma lâmpada em seu interior.

Além disso, esse compartimento também contém uma campainha e uma cuia. Já no compartimento inferior há um reservatório de água açucarada.

O procedimento experimental é descrito a seguir:

O rato, no centro do compartimento superior, observa que a lâmpada do buraco a sua frente era acesa e então fucinhava este buraco. Após isso, o rato ouvia a campainha, e devia fucinhar um dos buracos laterais. Caso em alguma realização do experimento o rato não o executasse nesta ordem, o experimento era cancelado e o resultado não era computado.

Enquanto o rato fucinhava um dos buracos laterais, um programa computacional escolhia um dos dois lados (direita ou esquerda) através de um algoritmo. Caso a escolha feita pelo computador fosse a mesma feita pelo rato, a cuia era acionada para buscar água doce no reservatório e retornava ao compartimento superior onde era ingerida pelo rato. A lâmpada do buraco escolhido pelo computador sempre era acesa, assim, o rato podia associar a recompensa com a lâmpada acesa.

Esta é uma versão do jogo “matching pennies”, onde dois oponentes, o rato e o computador devem, fazer uma escolha entre duas alternativas: lado direito e esquerdo. Caso façam a mesma escolha um dos jogadores vence, no caso o rato que recebe água com açúcar (recompensa). Quando as escolhas são diferentes, o outro jogador (computador) vence.

Devido a limitações dos ratos quanto à concentração para realizar os testes, cada um realizou 200 testes por dia.

Os ratos jogaram contra três algoritmos computacionais diferentes, que estão descritos na próxima seção.

Os três algoritmos utilizados no estudo foram codificados como algoritmos 0, 1 e 2. Os três ratos iniciaram suas participações no estudo jogando centenas de vezes contra o algoritmo 0, que é o algoritmo mais simples e que serviu de treinamento. Após esse período, os ratos jogaram milhares de vezes contra o algoritmo 1, e depois, jogaram mais milhares de vezes contra o algoritmo 2. Por fim, os ratos receberam uma dosagem de 0,5 ml/kg de anfetamina e jogaram outras centenas de vezes contra o algoritmo 2, aquele mais elaborado.

A razão dos ratos jogarem apenas centenas de vezes após receberem a dose de anfetamina, é o efeito da droga durar um período curto de tempo e o rato não poder receber muitas doses.

A quantidade exata de testes (válidos) realizados por cada rato pode ser visualizada na Tabela A.1.

2.2. ALGORITMOS COMPUTACIONAIS

2.2.1 ALGORITMO 0

De acordo com este algoritmo, o computador seleciona um dos dois lados de forma aleatória e com igual probabilidade, o que corresponde neste jogo, ao equilíbrio de Nash em estratégias mistas [Bortolossi et al., 2007], ou seja, independente da escolha adotada pelo rato, a probabilidade de vitória em cada teste é igual a 50% para cada jogador. Este algoritmo foi utilizado para estudar o padrão de escolhas dos ratos quando estes não têm estímulo para eventualmente alterar o procedimento de decisão.

2.2.2 ALGORITMO 1

Este algoritmo inicia cada sessão diária de jogos contra um determinado rato com a informação sobre as 200 escolhas feitas por este rato na sessão anterior. Conforme eles jogam na sessão atual, as escolhas do rato são armazenadas, assim, no

final da sessão o computador possui informação sobre as 400 últimas escolhas do rato, mas armazena somente as últimas 200 escolhas para iniciar a próxima sessão.

Assim, em cada teste, o computador usa a informação sobre as escolhas anteriores do rato para estimar as probabilidades condicionais de selecionar o lado direito no próximo teste, dado às últimas N escolhas do rato ($N=0, \dots, 4$). A seguir, para cada probabilidade condicional considerada, um teste de proporções é realizado com a hipótese nula igual a 0,5. Um P-valor (aproximado) de 5% é tomado como critério para rejeição da hipótese nula, pois em estudos similares a este, [Lee, et. al., 2004] e [Barracough, et. al., 2004], adotaram este nível de significância. Caso uma ou mais hipóteses nulas sejam rejeitadas, o algoritmo escolhe à direita com probabilidade $1-p^*$, onde p^* é a estimativa da probabilidade condicional utilizada no teste de hipótese, mais distante do valor 0,5.

Este algoritmo simula (ou provoca) uma situação, em que os ratos são induzidos a escolher qualquer um dos lados com igual probabilidade, e tanto quanto possível, de maneira independente das escolhas anteriores, visando maximizar o total de recompensas a serem obtidas.

2.2.3. ALGORITMO 2

Este algoritmo é mais elaborado do que o anterior. Além de considerar a informação sobre as escolhas anteriores do rato, ele também guarda os resultados dos testes.

Com base nessas informações, antes de cada decisão do rato, são estimados as probabilidades condicionais de o rato selecionar a direita, dada às últimas N escolhas e os respectivos resultados dos testes ($N=1, 2, 3$ e 4). Note que aqui, o caso $N=0$ (cenário de independência entre as escolhas e seus resultados) não é considerado.

O restante do procedimento, segue de maneira similar ao algoritmo 1.

De acordo com a regra de procedimento, induzida por este algoritmo, o rato deve selecionar as escolhas, para maximizar as recompensas, de maneira similar ao algoritmo 1. Isto é, escolher qualquer um dos lados com igual probabilidade, e tanto

quanto possível, de maneira independente das escolhas anteriores dele (o rato) e de seu oponente (o computador).

3. DESCRIÇÃO DAS VARIÁVEIS

Cada ensaio realizado retorna valores para as seguintes variáveis:

- Escolha do rato (1=Direita, 0=Esquerda)
- Escolha do computador (1=Direita, 0=Esquerda)
- Resultado do teste (1=se o rato recebe a recompensa, 0=caso contrário)

4. ANÁLISE DESCRITIVA

Foram tabulados os resultados para três ratos e para cada um destes, uma quantidade grande de ensaios foi realizada (Tabela A.1). A análise descritiva para cada rato foi realizada, e assim, foi possível identificar se alguns comportamentos padrões nos ensaios.

Inicialmente, para cada rato interagindo com cada algoritmo, foram calculadas as proporções de escolhas do lado direito, de recompensas, de repetição da escolha anterior e de realizações da estratégia “win stay / lose shift” (seções 4.2 e 4.3). Depois, essas medidas foram analisadas em blocos de observações com tamanhos aleatórios (Apêndice B).

Por último (Apêndice C), foram estimadas entropias e informações mútuas nas seqüências de escolhas dos ratos e do computador, com o objetivo de mensurar o grau de aleatoriedade nas escolhas feitas pelos ratos, neste apêndice também são apresentados os estimadores utilizados.

Na próxima seção será detalhada a regra que definiu a formação dos blocos com tamanhos aleatórios.

4.1. BLOCOS DE OBSERVAÇÕES

Para descrevermos o comportamento dos ratos no experimento, foram feitas análises descritivas em blocos sucessivos de observações. O intuito aqui é o de sumarizar os dados observados. Para cada rato, os blocos foram definidos com base nas transições da seqüência de escolhas, onde existem duas possibilidades em cada ensaio, direita ou esquerda, e onde existem quatro tipos de transição, a saber: direita-direita, direita-esquerda, esquerda-direita e esquerda-esquerda.

Todas as transições foram contadas, e o bloco é formado no instante em que observamos pelo menos 30 de cada tipo. Consideramos 30 transições de cada tipo como sendo um valor razoável, devido ao número de ensaios, e também, pois historicamente, considera-se 30 como o número mínimo de observações para o uso do Teorema Central do Limite.

4.2. ANÁLISES DE ESCOLHAS E RECOMPENSAS

Na terceira coluna da Tabela A.2, podemos notar que, quando jogaram contra o algoritmo 0 (treinamento), todos os ratos tiveram a tendência de escolher um dos lados com mais freqüência, sendo que, o rato 12 apresentou uma grande proporção de escolhas do lado direito, enquanto que, os ratos 22 e 23 escolheram mais a esquerda. Apesar disso, nota-se também na Tabela A.2 (quarta coluna), que os ratos foram recompensados em cerca de 50% dos testes, o que era esperado devido ao procedimento utilizado pelo computador para fazer suas escolhas. Nota-se nos Gráficos C.1, C.2 e C.3 valores baixos para as entropias estimadas a partir de seqüências de três escolhas sucessivas para os três ratos, o que também pode ser observado nos Gráficos C.4, C.5 e C.6 para seqüências de três escolhas sucessivas dos ratos e duas do computador. O que indica que os ratos não fizeram suas escolhas de forma equiprovável. Nos Gráficos C.7, C.8 e C.9 pode-se notar que para os três ratos as informações mútuas ficaram afastadas de zero, indicando assim, que as escolhas dos ratos eram previsíveis dados os três testes anteriores.

Após a introdução do algoritmo 1, a preferência por um dos lados diminuiu para os três ratos (Gráficos B.1, B.2 e B.3), com a proporção de escolhas do lado direito se aproximando de 0,5, sendo que, para o rato 12 (Gráfico B.1) esta proporção esteve

acima de 0,5 para quase todos os blocos. Na terceira coluna da Tabela A.2 nota-se que o rato 12 escolheu mais a direita do que os demais ratos contra o algoritmo 1.

Podemos notar nos Gráficos B.4, B.5 e B.6 que nos primeiros blocos com o algoritmo 1, as proporções de recompensas foram baixas, mas rapidamente cresceram para todos os ratos convergindo para 0,5, indicando assim, que conforme os ratos começaram a escolher os lados com igual probabilidade, as proporções de recompensas cresceram. O que evidencia uma capacidade de adaptação dos ratos ao jogo contra este algoritmo.

Nos Gráficos C.1, C.2 e C.3 nota-se que os valores das entropias rapidamente convergiram para 3 bits, e nos Gráficos C.4, C.5 e C.6 nota-se também o crescimento dos valores estimados, ou seja, as escolhas foram mais aleatórias. As informações mútuas também se aproximaram de zero, indicando imprevisibilidade nas escolhas dos ratos.

Com a introdução do algoritmo 2, não se notou grande variação nas medidas analisadas para os três ratos em relação às obtidas com o algoritmo 1, o que pode demonstrar ou a rápida adaptação dos ratos a este algoritmo ou que os ratos já jogaram contra o algoritmo 1 adotando a estratégia adequada ao jogo contra o algoritmo 2.

Com os ratos jogando com o efeito da dose de anfetamina contra o algoritmo 2, podemos notar (Gráficos B.1, B.2 e B.3) que as proporções de escolhas do lado direito por bloco ficaram mais afastadas de 0,5, sendo as do rato 22 as que mais se afastaram, o que também pode ser verificado na terceira coluna da Tabela A.2. Nos Gráficos B.4, B.5 e B.6 nota-se que as proporções de recompensas ficaram abaixo de 0,5 para todos os blocos de todos os ratos. Quanto às medidas de aleatoriedade (Apêndice C), não notamos grande alteração nas medidas obtidas para os ratos 12 e 23 em relação às obtidas antes da dose de anfetamina. Já para o rato 22 (Gráficos C.2, C.5 e C.8), notamos que as estimativas obtidas ficaram parecidas com as obtidas no algoritmo 0 (treinamento), indicando assim, que a dose de anfetamina alterou o comportamento desse rato no jogo, mas, isso não resultou em piores resultados para ele.

4.3. ANÁLISE DE DEPENDÊNCIA E ESTRATÉGIA

Se um rato selecionasse o lado direito em um dado teste com probabilidade p_D e se os sucessivos testes fossem independentes, a probabilidade de escolher o mesmo lado em dois testes consecutivos seria igual a $p_D^2 + (1-p_D)^2$.

Então, para cada rato, calculamos a proporção de vezes que eles fizeram a mesma escolha em dois testes consecutivos (quinta coluna da Tabela A.2), e com base nas proporções de escolha do lado direito (terceira coluna da Tabela A.2), estimamos a probabilidade de cada rato fazer a mesma escolha em dois testes consecutivos sob a hipótese de independência (sexta coluna da Tabela A.2).

Comparando os resultados encontrados, podemos notar que há evidência da falta de independência nas escolhas dos ratos em relação às anteriores contra os três algoritmos e quando está sob o efeito da anfetamina, essa evidência é sustentada pelo grande número de testes realizados por cada rato (Tabela A.1).

Em processos de tomada de decisão, a estratégia “win stay lose shift” (WSLS) - “ganha permanece perde troca” em uma tradução livre - é muito usada [Bush e Mosteller, 1955]. Essa estratégia consiste em escolher no próximo teste a escolha vencedora no teste anterior, o que no caso dos ratos neste estudo, é equivalente a fazer no próximo teste a mesma escolha feita pelo computador no teste anterior, ou seja, escolher o lado da caixa que a lâmpada correspondente acendeu no teste anterior.

Na sétima coluna da Tabela A.2 nota-se que o rato 12 utilizou essa estratégia com mais frequência contra o algoritmo 0, já com os outros algoritmos a frequência diminuiu e se tornou constante. No Gráfico B.7 notamos que as proporções estiveram sempre próximas de 0,5, e após a dose de anfetamina parece haver uma tendência de crescimento do uso dessa estratégia, apesar de que para um dos blocos a proporção foi bem abaixo de 0,5, o que pode ser apenas uma flutuação ou algum fator fora de controle.

O rato 22 foi o que mais utilizou essa estratégia contra o algoritmo 1 (sétima coluna da Tabela A.2), mas, no Gráfico B.8 nota-se que após a introdução do algoritmo 2 o uso dessa estratégia diminui gradativamente. O que caracteriza o aprendizado por recompensa, pois o algoritmo 1 não tem informação sobre os resultados dos jogos, logo, é possível adotar a estratégia WSLS e não ser prejudicado, já o algoritmo 2 detecta facilmente o uso dessa estratégia, e conseqüentemente, diminui a proporção de recompensas do rato. Após a dose de anfetamina, não se nota grande variação no uso dessa estratégia.

O rato 23 foi o que apresentou a menor variação no uso da estratégia WSLS durante o experimento (sétima coluna da Tabela A.2 e Gráfico B.9), sendo que, há um aparente crescimento do uso nos blocos durante o algoritmo 2 (Gráfico B.9), mas após a dose de anfetamina houve uma aparente queda neste uso.

5. ANÁLISE INFERENCIAL

Para estudar o processo de tomada de decisão dos ratos durante o experimento, analisamos as seqüências de escolhas feitas por eles nos jogos contra os diferentes algoritmos computacionais e sob o efeito da anfetamina. De acordo com resultados apresentados na subseção anterior, devemos considerar em nossa análise, a existência de dependência entre sucessivas escolhas.

No Apêndice D é apresentada a metodologia de cadeias de Markov que é adequada para analisar seqüências de valores aleatórios com dependência. Com base nesta metodologia, foram analisadas as seqüências de escolhas dos ratos nas diferentes situações impostas pelo experimento.

Os resultados obtidos são apresentados na próxima subseção e no Apêndice E. Na subseção 5.2 são comparados os processos de escolha de um mesmo rato em situações distintas e os processos de escolha de diferentes ratos em uma mesma situação. Já na subseção 5.3 são comparados os desempenhos dos ratos, quanto à proporção de recompensas recebidas, nos jogos contra o algoritmo 2 antes e depois da ingestão de anfetamina, os resultados são apresentados no Apêndice F.

No Apêndice G é apresentado um método para obter estimativas intervalares para as probabilidades de transição.

5.1. ANÁLISE DAS ESCOLHAS

A definição de Cadeias de Markov de ordem k ($k \geq 0$) é apresentada no Apêndice D, assim como os seguintes critérios para estimar a ordem da cadeia: Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) e Critério da Mínima Entropia (MEC).

Para cada rato em cada período de testes, a saber: treinamento, testes contra o algoritmo 1, contra o algoritmo 2 e sob o efeito de anfetamina contra o algoritmo 2; foram ajustadas cadeias de Markov de diversas ordens e a qualidade de seus ajustes foram comparadas com base nos critérios AIC, BIC e MEC para escolher a ordem mais apropriada. Como cada critério nos indica uma estimativa da verdadeira ordem da cadeia, a estimativa escolhida foi a menor entre as estimadas por cada critério, pois a complexidade da cadeia cresce com o aumento da ordem. No geral, o critério BIC estimou as menores ordens para os processos.

No Apêndice E (Tabelas E.1.a – E.12.a) são apresentados os valores do AIC, BIC e P-valor do MEC para as cadeias de Markov de ordem k ($0 \leq k \leq 5$) ajustadas às seqüências de escolhas de cada rato em cada período. Os valores destacados em amarelo nas tabelas representam os melhores valores encontrados para cada critério. Nas Tabelas E.1.b – E.12.b são mostradas as probabilidades de transição para as cadeias ajustadas segundo as ordens estimadas nas Tabelas E.1.a – E.12.a.

Na Tabela E.1.a podemos notar que o BIC estimou a ordem zero para o processo de escolhas do Rato 12 durante o treinamento, essa ordem corresponde à independência entre as escolhas do rato, ou seja, durante o treinamento esse rato fez suas escolhas de forma independente. Na Tabela E.1.b nota-se que o Rato 12 escolheu o lado direito com uma frequência maior.

Nos jogos contra o algoritmo 1, segundo os três critérios considerados, as escolhas do Rato 12 são dependentes das duas escolhas anteriores (Tabela E.2.a), na Tabela E.2.b podemos notar que a probabilidade da próxima escolha do rato ser o lado

direito dado que em pelo menos um dos dois testes anteriores ele optou pelo lado direito é maior do que 0,5.

Na Tabela E.3.a nota-se que, a exemplo dos jogos contra o algoritmo 1, a ordem estimada do processo de escolhas do rato 12 contra o algoritmo 2 é igual a 2 segundo os três critérios. Novamente, a probabilidade da próxima escolha ser o lado direito é superior a ser do lado esquerdo, caso o rato tenha escolhido o lado direito em algum dos dois testes anteriores (Tabela E.3.b).

Após o rato 12 receber a dose de anfetamina, suas escolhas passaram a depender somente da escolha anterior (Tabela E.4.a), ou seja, a anfetamina diminui em uma unidade o número de escolhas anteriores que influenciam as escolhas do rato 12 contra o algoritmo 2. Além disso, a Tabela E.4.b indica que o rato 12 tendeu a repetir a escolha anterior.

A Tabela E.5.a indica que, segundo o BIC, as escolhas do rato 22 durante o treinamento são influenciadas somente pela escolha anterior. De acordo com a Tabela E.5.b, o rato 22 tendeu a escolher a esquerda caso tivesse feito essa escolha no jogo anterior.

Já nos jogos contra o algoritmo 1, os três critérios estimaram ordem dois para o processo de escolhas do rato 22 (Tabela E.6.a). Nota-se na Tabela E.6.b que o rato 22 tendeu a repetir o lado escolhido no teste anterior, principalmente se esse lado foi o esquerdo. A Tabela E.7.a mostra que nos jogos contra o algoritmo 2 a próxima escolha do rato 22 dependeu da escolha anterior, além disso, segundo a Tabela E.7.b ele apresentou a tendência de repetir a escolha anterior.

Quando jogou sob o efeito da anfetamina, o rato 22 realizou escolhas independentes das anteriores (Tabela E.8.a), ou seja, essa substância eliminou a dependência da escolha anterior que existia nas escolhas contra o algoritmo 2 antes da intervenção com anfetamina. Podemos notar na Tabela E.8.b que esse rato teve preferência pelo lado esquerdo nessa situação.

As escolhas do rato 23 durante o treinamento e nos jogos contra os algoritmos 1 e 2 foram dependentes das escolhas feitas nos dois jogos anteriores (Tabelas E.9.a, E.10.a e E.11.a). Nota-se nas Tabelas E.9.b, E.10.b e E.11.b que as proporções de escolhas desse rato foram mais equilibradas entre direita e esquerda.

Mas após ingerir a anfetamina, o processo de escolhas do rato 23 foi de ordem 0 (Tabela E12.a), ou seja, suas escolhas passaram a ser independentes das anteriores. Na Tabela E.12.b nota-se que a distribuição das escolhas foi quase equiprovável.

Na Tabela E.13 nota-se para todos os ratos, a anfetamina diminuiu a ordem do processo de escolhas, ou seja, a anfetamina diminuiu a “memória” dos ratos quanto às escolhas anteriores.

5.2. TESTES DE HOMOGENEIDADE

Foram feitos testes de homogeneidade para comparar os processos de escolhas de um mesmo rato em situações diferentes e para comparar os processos de escolhas dos ratos que tiveram a mesma ordem estimada. O teste utilizado foi o teste da razão de verossimilhança com a hipótese nula sendo a igualdade dos processos e a hipótese alternativa sendo a diferença dos processos.

Na Tabela F.1 é testada a igualdade dos processos de escolhas do rato 12 nos jogos contra os algoritmos 1 e 2, nota-se que o teste não rejeita a hipótese de igualdade dos processos, isto é, não há evidências de que este rato alterou seu comportamento quanto às escolhas nos jogos contra o algoritmo 2 em relação ao processo de escolhas nos jogos contra o algoritmo 1.

A Tabela F.2 apresenta o resultado do teste de homogeneidade entre os processos de escolhas do rato 23 durante o treinamento e nos jogos contra os algoritmos 1 e 2, nota-se que não há evidência de que os processos sejam diferentes, ou seja, este rato não alterou seu comportamento quanto às escolhas desde o treinamento até os jogos contra o algoritmo 2.

Os processos de escolha do rato 12 nos jogos contra os algoritmos 1 e 2 (que são iguais segundo a Tabela F.1) foram comparados com os processos de escolha do rato 23 nos jogos contra os mesmos algoritmos, a Tabela F.3 apresenta o resultado do teste que indica não haver evidência de que os processos sejam diferentes, ou seja, os ratos 12 e 23 realizaram o mesmo processo de escolhas (de ordem 2) contra os algoritmos 1 e 2.

5.3. EFEITO DA ANFETAMINA NO DESEMPENHO DOS RATOS

Como os três ratos jogaram contra o algoritmo 2 antes e depois de receberem a dose de anfetamina, comparamos a proporção de recompensas recebidas por cada rato nestes dois períodos. Para isso, inicialmente verificamos se existe dependência entre os resultados dos jogos dos três ratos nestes períodos. As Tabelas E.14.a, E.15.a e E.16.a indicam que os resultados dos jogos dos três ratos contra o algoritmo 2 antes da ingestão de anfetamina são independentes entre si. O mesmo pode ser concluído para os jogos depois da ingestão de anfetamina (Tabelas E.14.b, E.15.b e E.16.b).

Como a suposição de independência dos resultados dos jogos é válida, foram feitos testes de comparação de proporções entre as proporções de recompensas recebidas por cada rato antes e depois da ingestão de anfetamina.

Na Tabela F.4 é apresentado o teste para o rato 12, nota-se que há diferença significativa entre as proporções de recompensas recebidas nos jogos contra o algoritmo 2 antes e depois da anfetamina. Este mesmo resultado foi obtido para o rato 23 (Tabela F.6). Já para o rato 22, não foi detectada diferença significativa no desempenho quanto à proporção de recompensas (Tabela F.5).

6. CONCLUSÃO

Pelas análises realizadas, nota-se que quando os ratos jogam o “matching pennies” contra um adversário que adota a estratégia ideal (algoritmo 0), eles não se sentem estimulados a variar suas escolhas, pois são recompensados em cerca de 50% dos jogos.

Quando começaram a jogar contra o algoritmo 1, os ratos se sentiram forçados a mudar a forma de jogar, fazendo suas escolhas de forma aleatória, segundo as entropias e informações mútuas estimadas, e com igual probabilidade para os dois lados, para assim, continuarem sendo recompensados em 50% dos jogos.

Verificamos que os ratos se adaptaram rapidamente ao algoritmo 2, não sendo possível notar uma queda inicial na proporção de recompensas, como foi possível notar quando os ratos começaram a jogar contra o algoritmo 1. Isso aconteceu, pois os ratos

adotaram a estratégia ideal para jogar contra o algoritmo 2 ainda nos jogos contra o algoritmo 1.

Para o rato 22 foi possível notar uma perda de rendimento no jogo após a ingestão da dosagem de anfetamina, pois este rato passou a escolher a esquerda com uma frequência maior do que vinha escolhendo, além disso, segundo as estimativas das entropias e informações mútuas, suas escolhas ficaram mais previsíveis, entretanto, este rato não foi prejudicado na quantidade de recompensas.

Os ratos não utilizaram a estratégia WSLS com muita frequência mesmo quando era possível utilizá-la sem ser prejudicado (contra os algoritmos 0 e 1), além disso, as escolhas dos ratos não são independentes de suas escolhas anteriores.

Utilizamos cadeias de Markov para analisar os processos de escolhas dos ratos, observamos que a maior ordem estimada para as cadeias de Markov é a ordem 2, ou seja, no máximo, a próxima escolha depende das duas últimas escolhas.

Os processos de escolhas do rato 12 contra o algoritmos 1 e 2 são iguais, o mesmo ocorre para o rato 23 durante o treinamento e contra os algoritmos 1 e 2. Além disso, o processo aleatório de escolhas do rato 12 é igual ao do rato 23 contra os algoritmos 1 e 2.

No geral, a anfetamina diminuiu a ordem das cadeias de markov estimadas, ou seja, essa substância tornou as escolhas dos ratos mais independentes das escolhas anteriores.

Além disso, houve diferença significativa entre as proporções de recompensas recebidas pelos ratos 12 e 23 contra o algoritmo 2 antes e depois da ingestão de anfetamina.

APÊNDICE A

(Tabelas)

APÊNDICE A

(Tabelas)

Tabela A.1. Número de testes por animal.

Algoritmos	Animal	Número de testes
0	12	200
	22	500
	23	600
1	12	2410
	22	2410
	23	2400
2	12	6000
	22	5600
	23	6000
2 (anfetamina)	12	400
	22	400
	23	400

Tabela A.2. Proporção de características de interesse do projeto.

Algoritmos	Animal	Escolhas da direita	Recompensas	Repetição da escolha anterior	Repetição da escolha anterior sob independência	WSLS
0	12	0,94	0,55	0,91	0,89	0,58
	22	0,35	0,49	0,63	0,55	0,58
	23	0,35	0,48	0,68	0,55	0,51
1	12	0,59	0,48	0,56	0,52	0,55
	22	0,5	0,47	0,61	0,50	0,59
	23	0,49	0,5	0,57	0,50	0,55
2	12	0,6	0,48	0,55	0,52	0,54
	22	0,5	0,46	0,55	0,50	0,56
	23	0,51	0,48	0,53	0,50	0,55
2 (anfetamina)	12	0,59	0,44	0,61	0,51	0,55
	22	0,40	0,46	0,5	0,52	0,53
	23	0,51	0,47	0,54	0,50	0,55

Gráfico 6.1: Evolução da produtividade (kg/ha) da cultura de milho em função da aplicação de adubo nitrogenado (kg/ha) em 2010 e 2011, em função da aplicação de adubo nitrogenado (kg/ha).

Fig. 6.1: Evolução da produtividade (kg/ha) da cultura de milho em função da aplicação de adubo nitrogenado (kg/ha) em 2010 e 2011, em função da aplicação de adubo nitrogenado (kg/ha).



APÊNDICE B

(Gráficos por animal)

Gráfico 3.1: Evolução da produtividade (kg/ha) da cultura de milho em função da aplicação de adubo nitrogenado (kg/ha) em 2010 e 2011, em função da aplicação de adubo nitrogenado (kg/ha).

Fig. 3.1: Evolução da produtividade (kg/ha) da cultura de milho em função da aplicação de adubo nitrogenado (kg/ha) em 2010 e 2011, em função da aplicação de adubo nitrogenado (kg/ha).



Gráfico B.1. Evolução da proporção de escolhas da direita para o rato 12 por blocos durante o experimento.

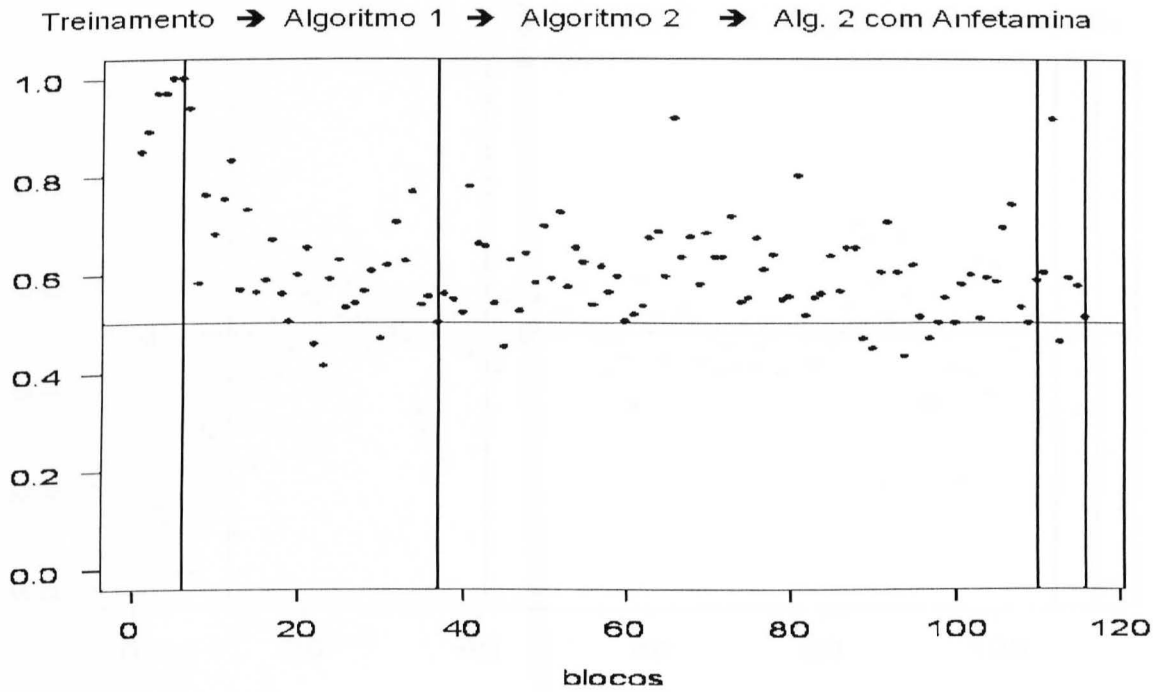


Gráfico B.2. Evolução da proporção de escolhas da direita para o rato 22 por blocos durante o experimento.

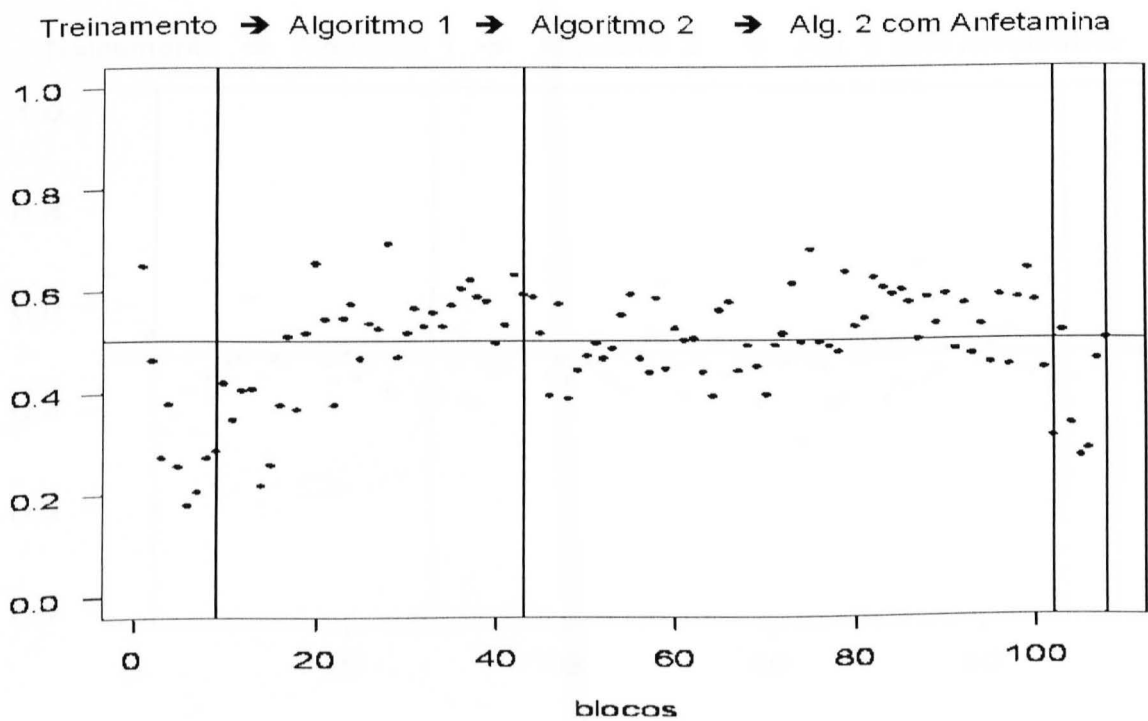


Gráfico B.3. Evolução da proporção de escolhas da direita para o rato 23 por blocos durante o experimento.

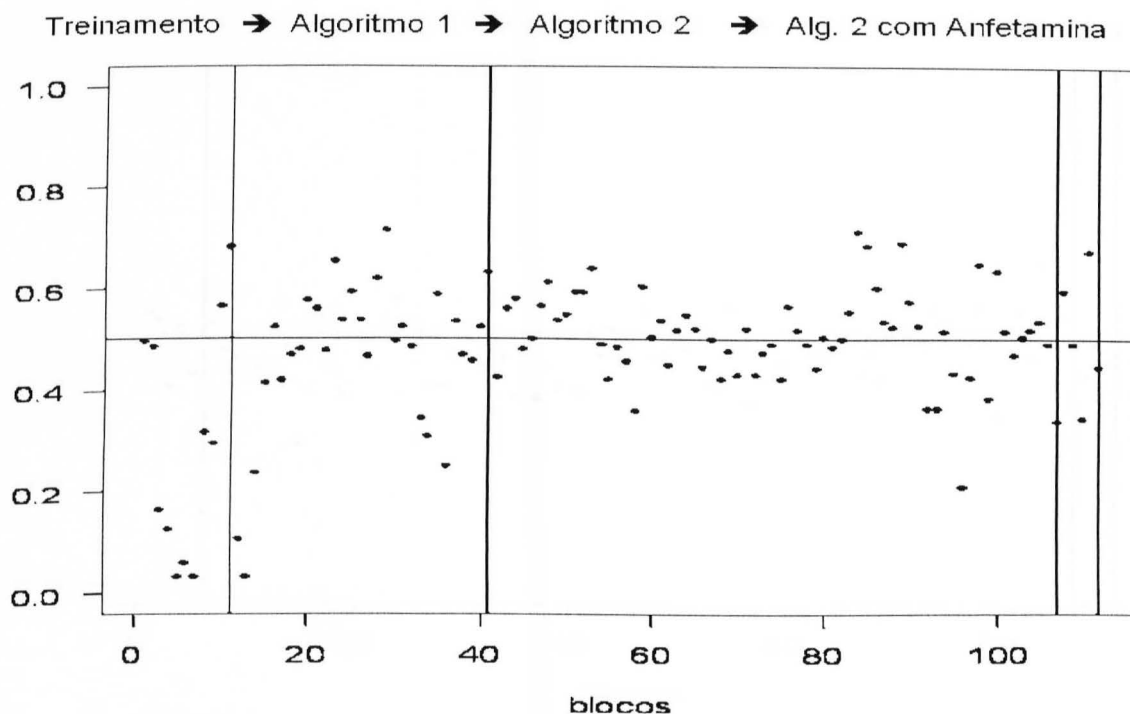


Gráfico B.4. Evolução da proporção de recompensas para o rato 12 por blocos durante o experimento.

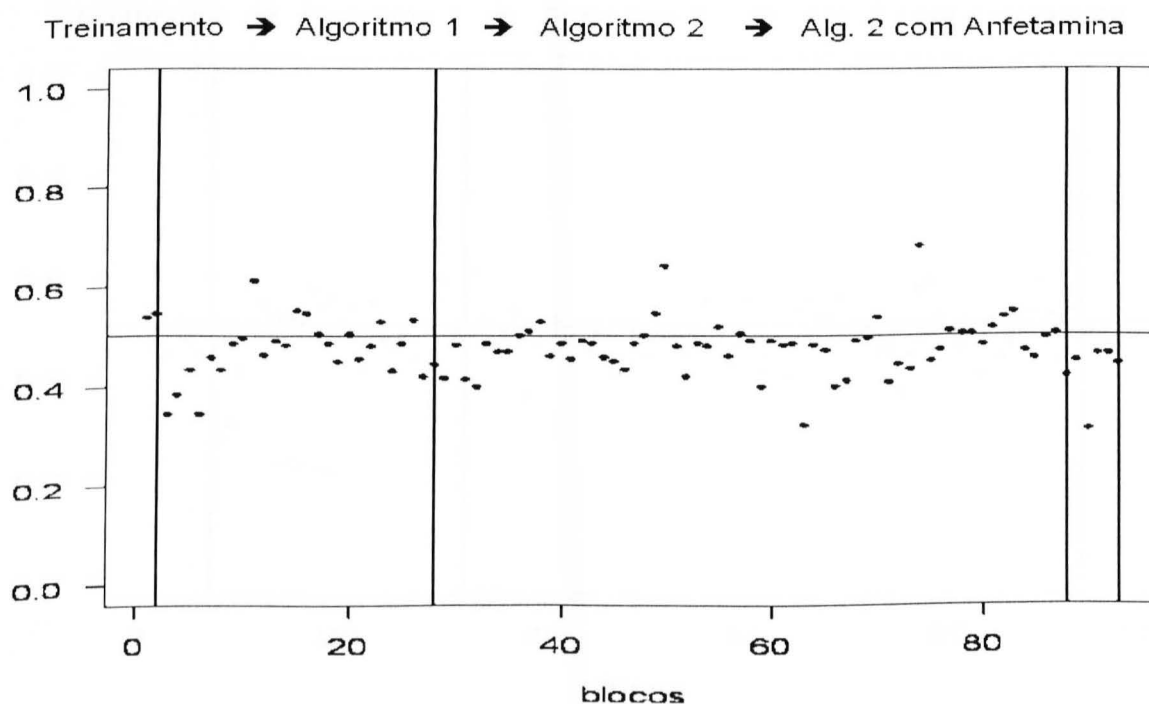


Gráfico B.5. Evolução da proporção de recompensas para o rato 22 por blocos durante o experimento.

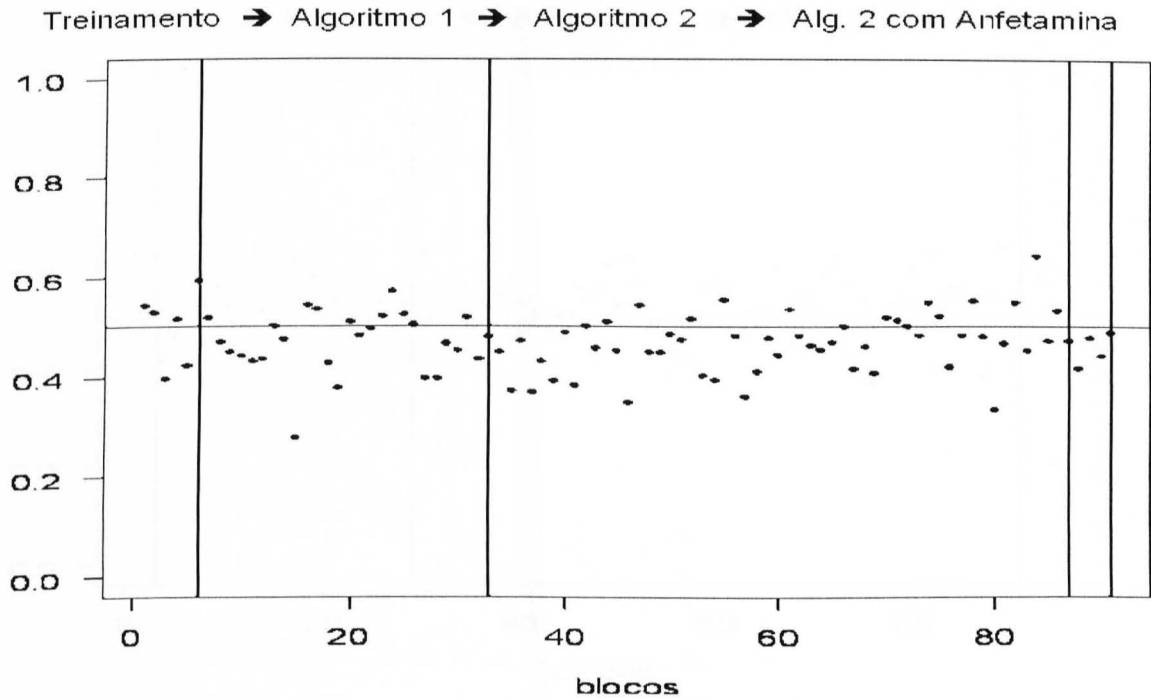


Gráfico B.6. Evolução da proporção de recompensas para o rato 23 por blocos durante o experimento.

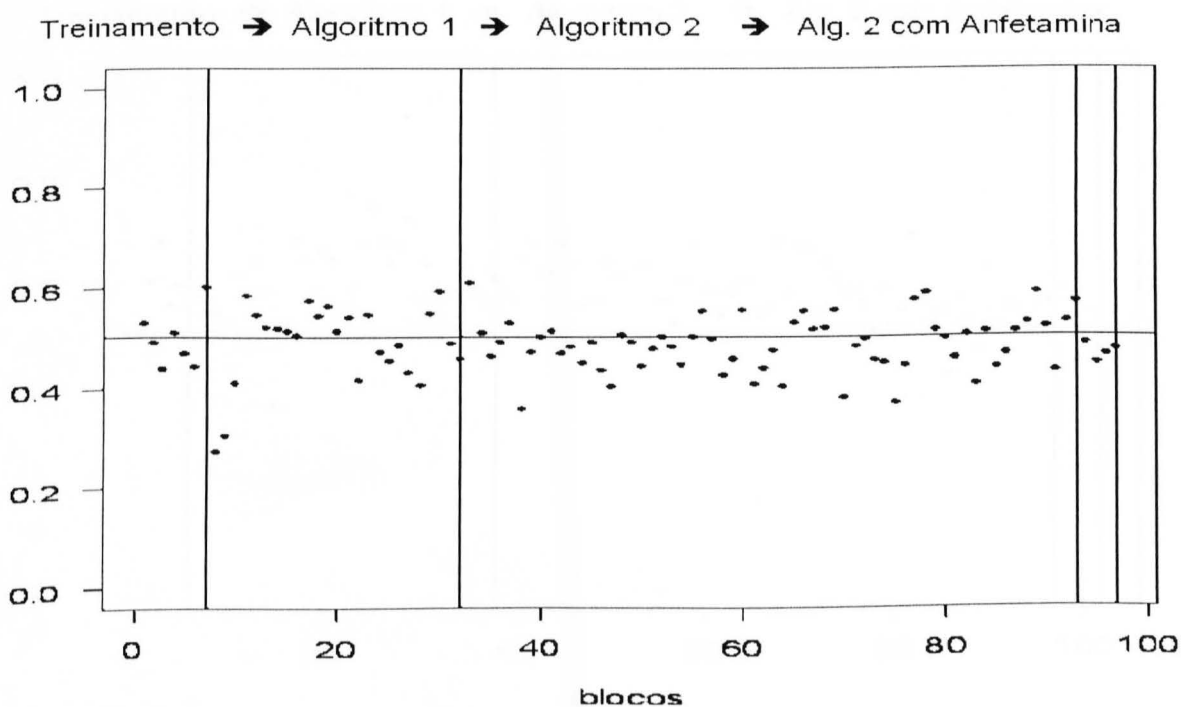


Gráfico B.7. Evolução da proporção de realizações da estratégia WSLS para o rato 12 por blocos durante o experimento.

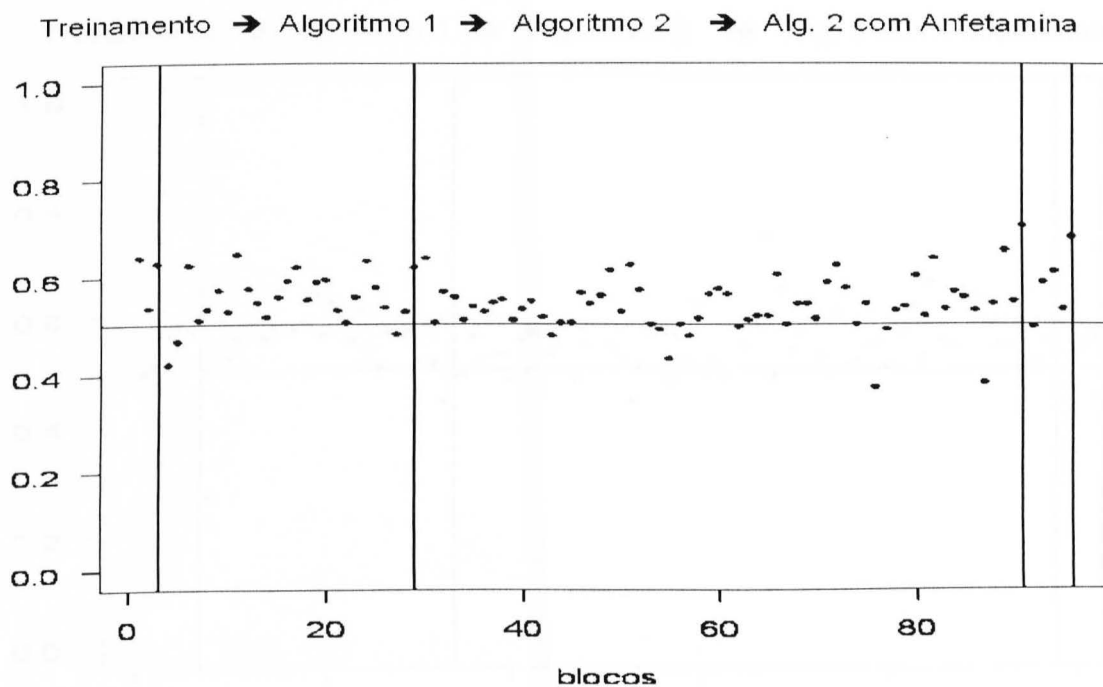


Gráfico B.8. Evolução da proporção de realizações da estratégia WSLS para o rato 22 por blocos durante o experimento.

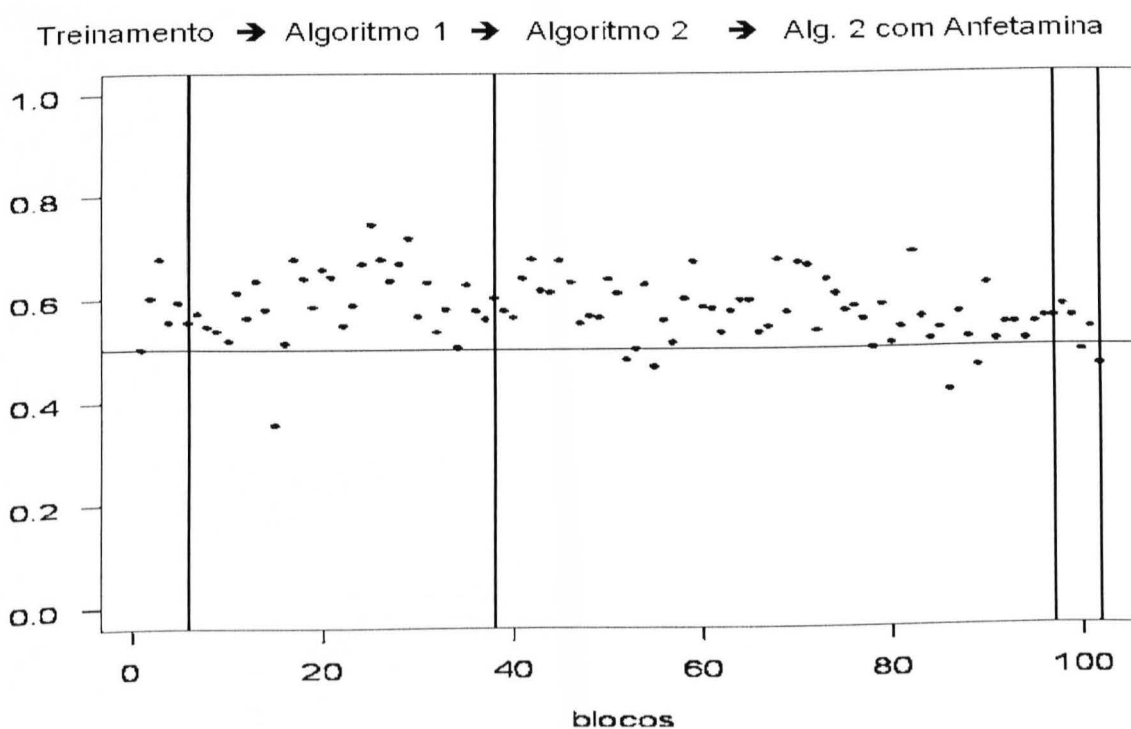
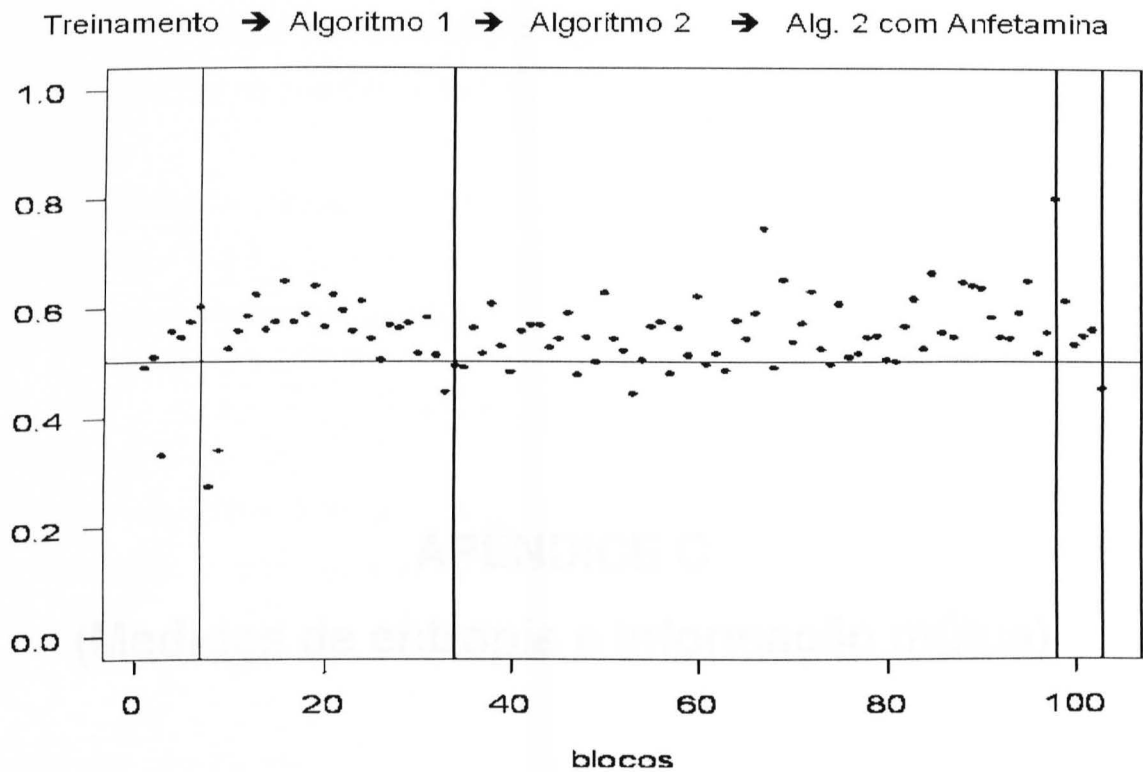


Gráfico B.9. Evolução da proporção de realizações da estratégia WSLS para o rato 23 por blocos durante o experimento.



APÊNDICE C

(Medidas de entropia e informação mútua)

O grau de aleatoriedade na seqüência de escolhas de cada rato foi quantificado, nesta parte do trabalho, através de estimativas de entropia e informação mútua. Essas medidas foram obtidas para seqüências de três escolhas sucessivas de ambos os jogadores, pois segundo estudos anteriores [Lee et al., 2004 e Soltani et al., 2006], com seqüências maiores os resultados são similares aos obtidos para seqüências com três escolhas.

Se um experimento tem k possíveis resultados, cada um com probabilidade p_i de ocorrência, com $1 \leq i \leq k$, a entropia da distribuição de probabilidade dos possíveis resultados do experimento é definida como:

$$H = -\sum_{i=1}^k p_i \log_2 p_i \text{ (bits)}$$

A fórmula acima é viesada para estimar a entropia a partir de uma amostra, sendo necessário utilizar a seguinte correção [Miller, 1955]:

$$\hat{H} = -\sum_{i=1}^k \hat{p}_i \log_2 \hat{p}_i + \frac{k-1}{1,3863N} \text{ (bits)}$$

Onde \hat{p}_i é o estimador de máxima verossimilhança (proporção amostral) de p_i e N é o tamanho da amostra.

Para cada rato, foram estimadas entropias a partir de seqüências de três escolhas sucessivas do rato e as duas últimas do computador. Considerando seqüências de três escolhas sucessivas dos ratos, há $k=2^3=8$ possíveis combinações e o valor máximo que a entropia pode assumir é igual a 3 bits, já para três escolhas do rato e duas do computador, há $k=2^5=32$ possíveis combinações e o valor máximo são 5 bits. Quanto mais próxima à entropia estiver do valor máximo, mais uniforme é a distribuição de probabilidade.

A informação mútua foi estimada para seqüências de escolhas do rato e do computador em três testes sucessivos (entrada) mais a próxima escolha do rato (saída). Ela foi estimada com a fórmula abaixo que já inclui a correção para tirar o viés.

$$\hat{I} = -\sum_{i=1}^r \sum_{j=1}^c \hat{p}_{ij} \frac{\hat{p}_{ij}}{\hat{p}_i \hat{p}_j} - \frac{(r-1)(c-1)}{1,3863N} \text{ (bits)}$$

Onde \hat{p}_i é a estimativa da probabilidade (proporção amostral) de ocorrer a i -ésima entrada ($r=2^6=64$), \hat{p}_j é a estimativa da probabilidade (proporção amostral) de ocorrer a j -ésima saída ($c=2$), e \hat{p}_{ij} é a estimativa da probabilidade (proporção amostral) conjunta de ocorrer a i -ésima entrada e a j -ésima saída. Quanto mais próxima do valor zero for a informação mútua, mais imprevisível é a próxima escolha do rato dadas as escolhas dos dois jogadores (rato e computador) nos três testes anteriores.

Gráfico C.1. Evolução da entropia de três escolhas sucessivas para o rato 12 por blocos durante o experimento.

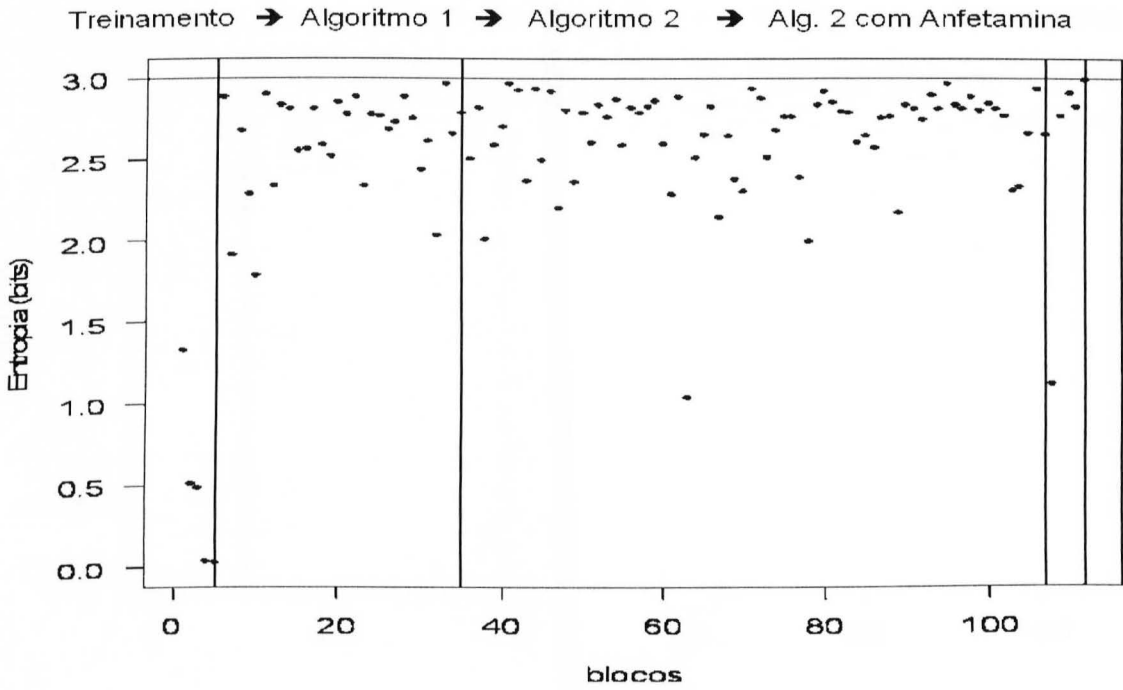


Gráfico C.2. Evolução da entropia de três escolhas sucessivas para o rato 22 por blocos durante o experimento.

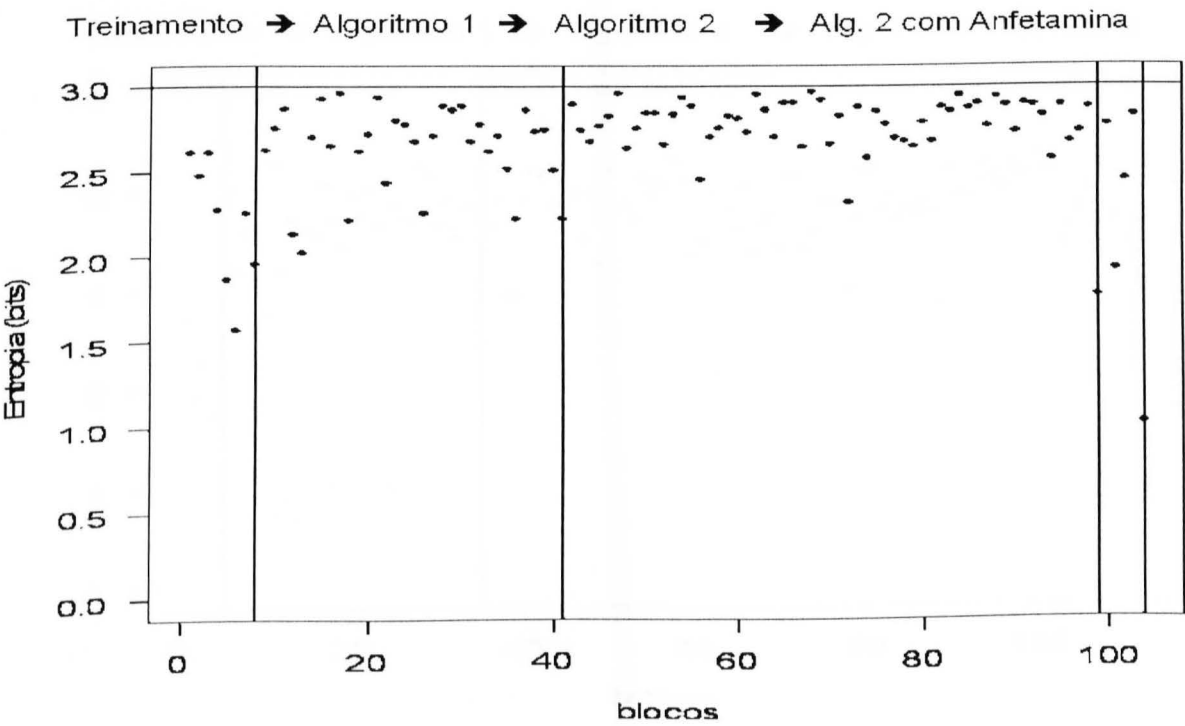


Gráfico C.3. Evolução da entropia de três escolhas sucessivas para o rato 23 por blocos durante o experimento.

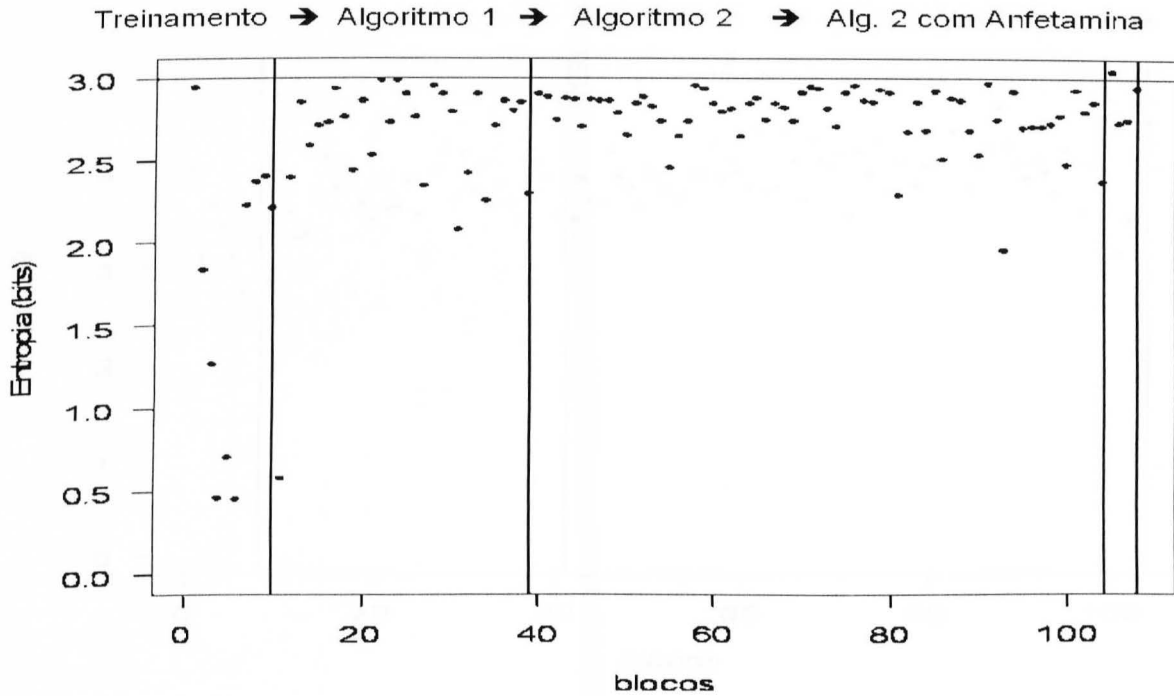


Gráfico C.4. Evolução da entropia de três escolhas sucessivas para o rato 12 e duas do computador por blocos durante o experimento.

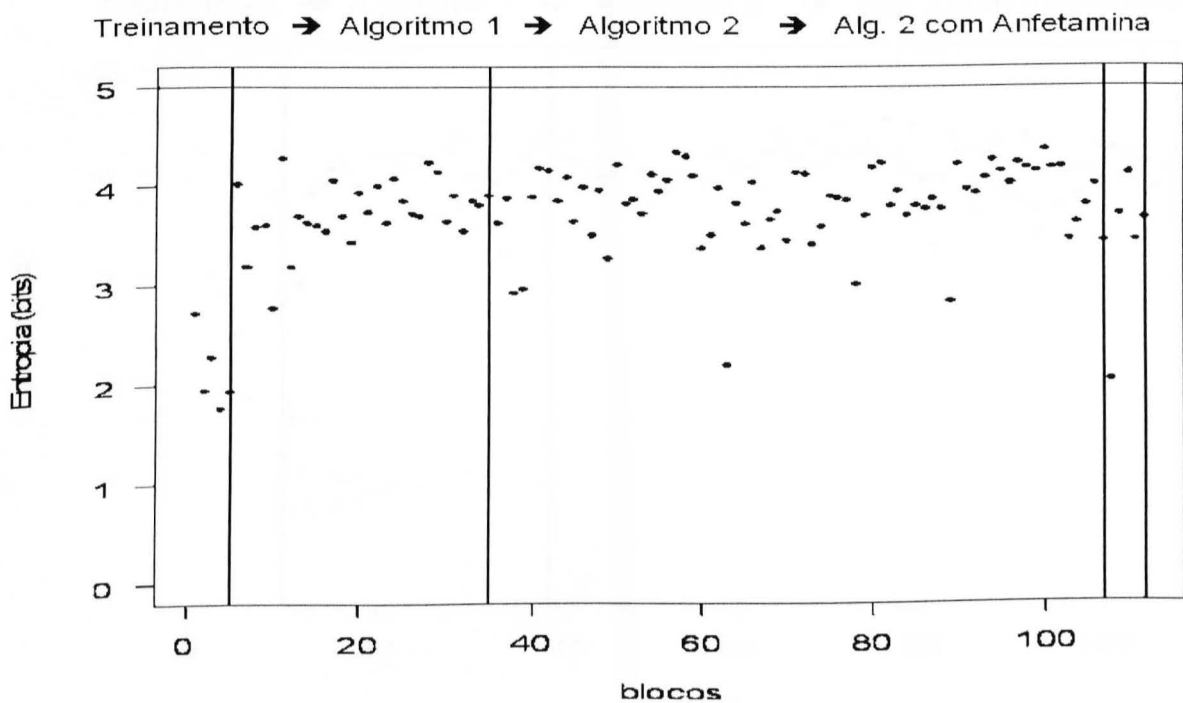


Gráfico C.5. Evolução da entropia de três escolhas sucessivas para o rato 22 e duas do computador por blocos durante o experimento.

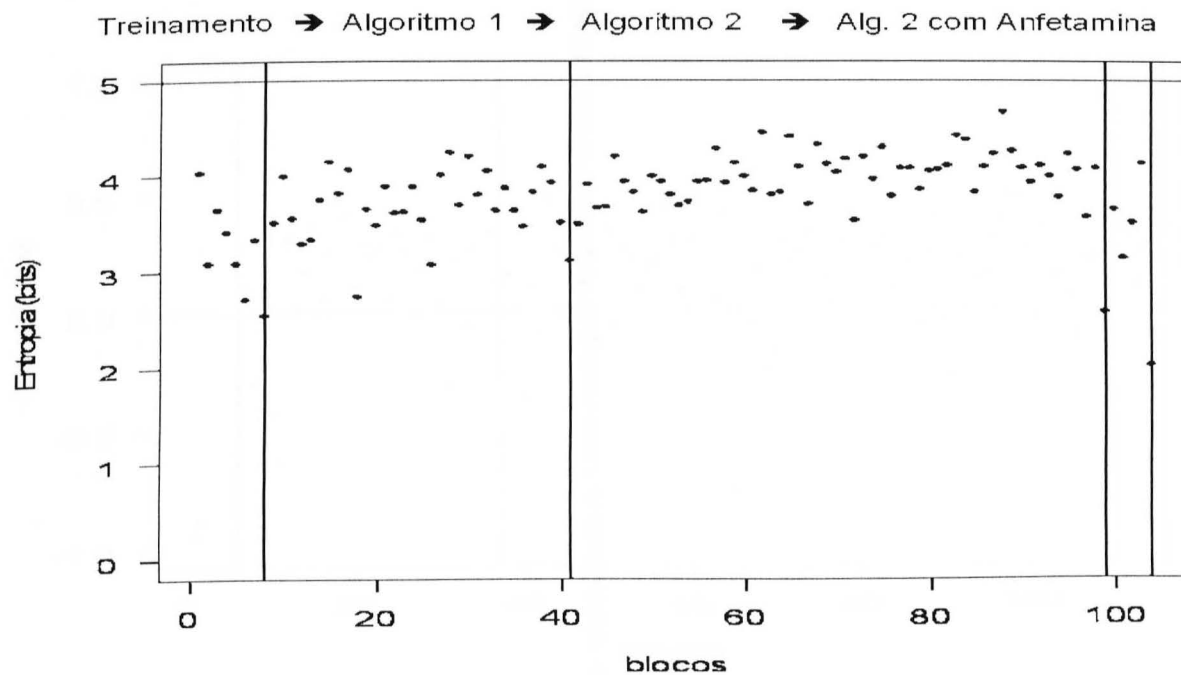


Gráfico C.6. Evolução da entropia de três escolhas sucessivas para o rato 23 e duas do computador por blocos durante o experimento.

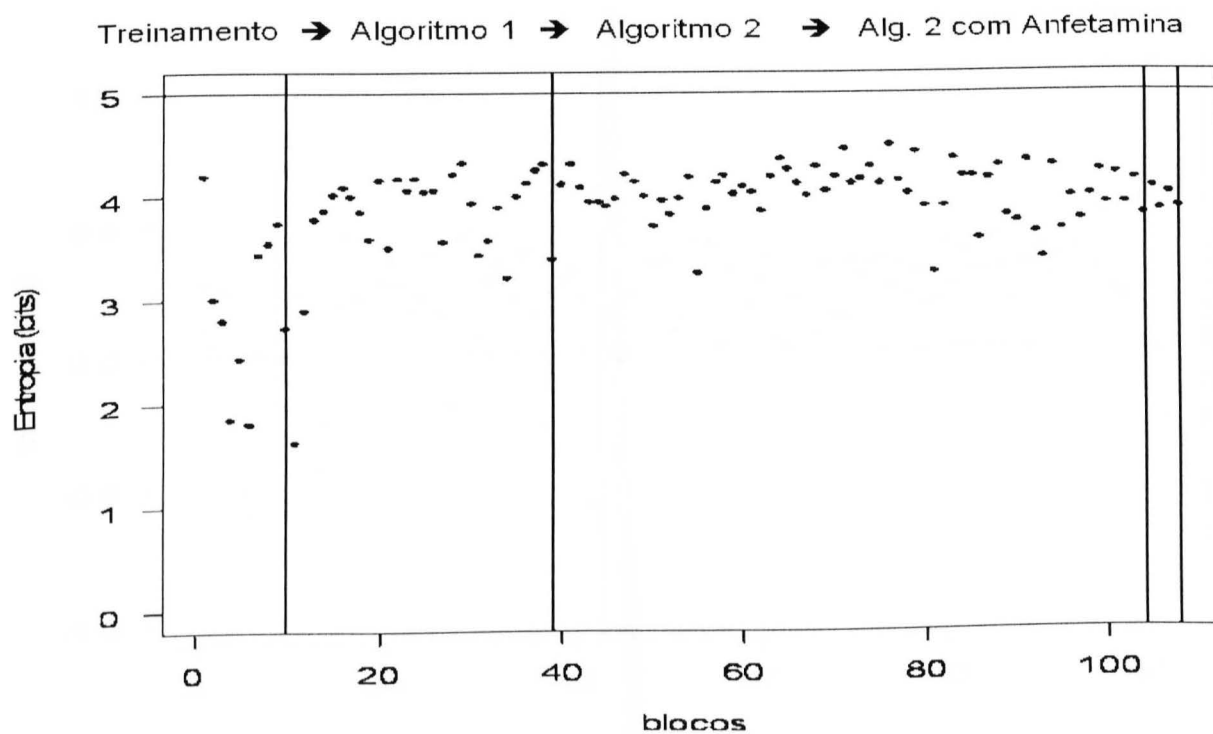


Gráfico C.7. Evolução da informação mútua o rato 12 e por blocos durante o experimento.

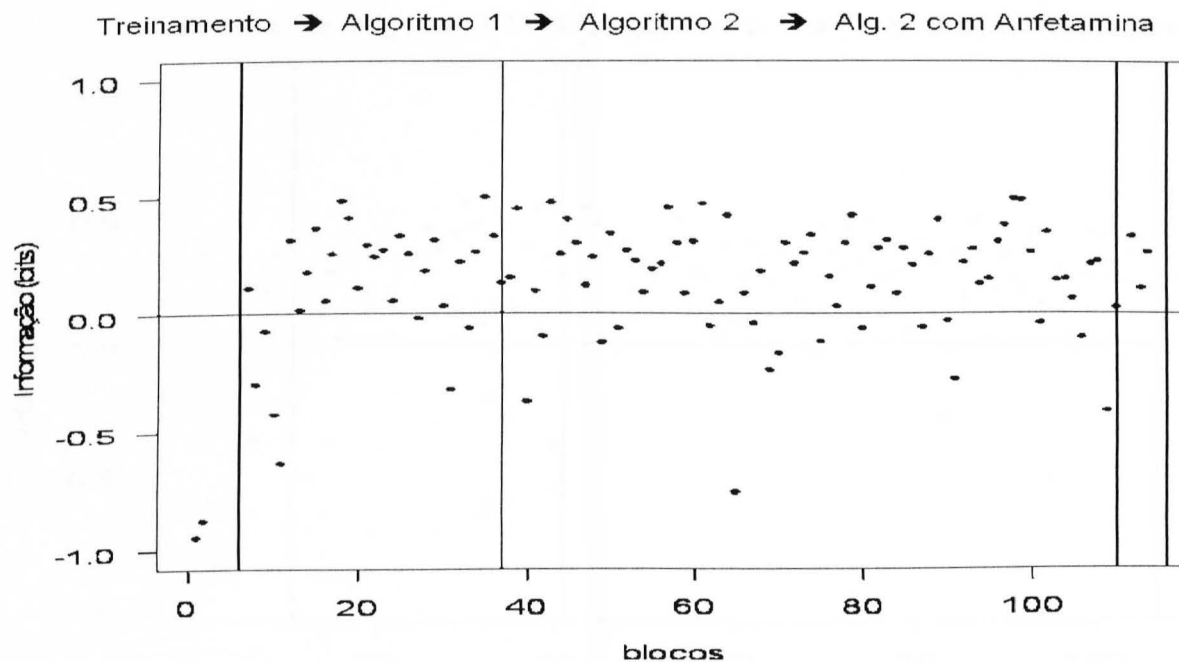


Gráfico C.8. Evolução da informação mútua o rato 22 e por blocos durante o experimento.

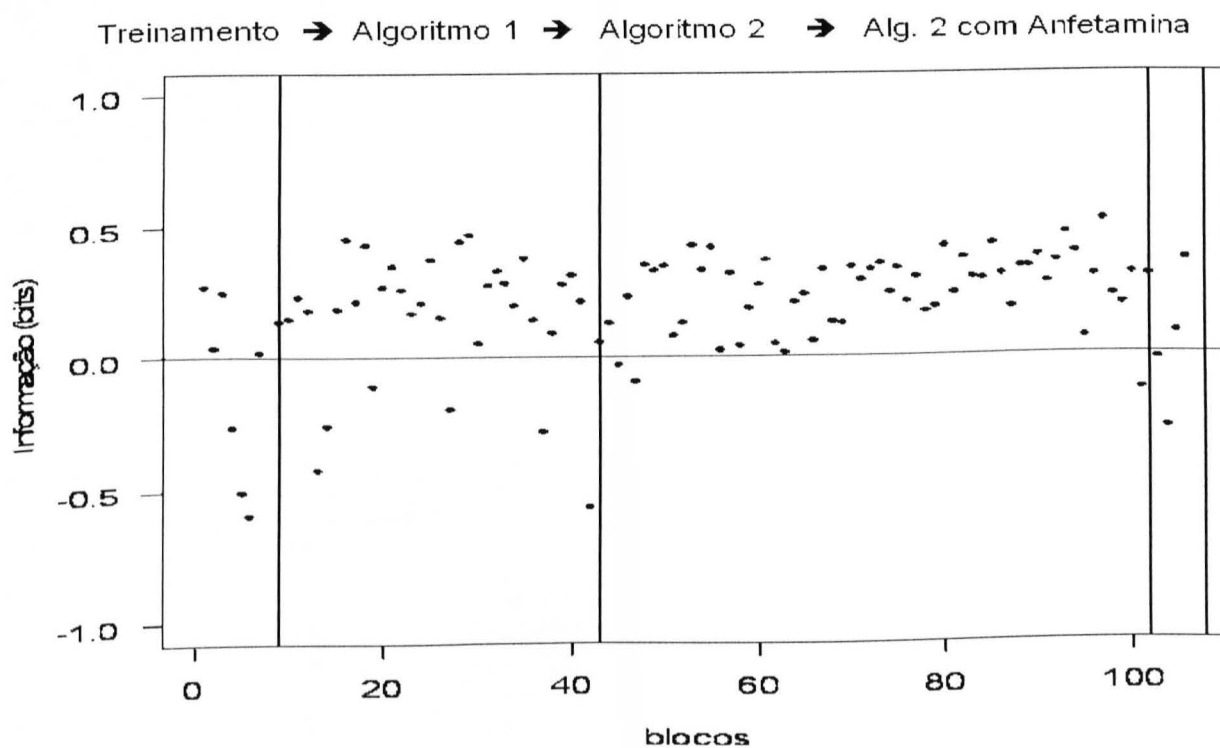
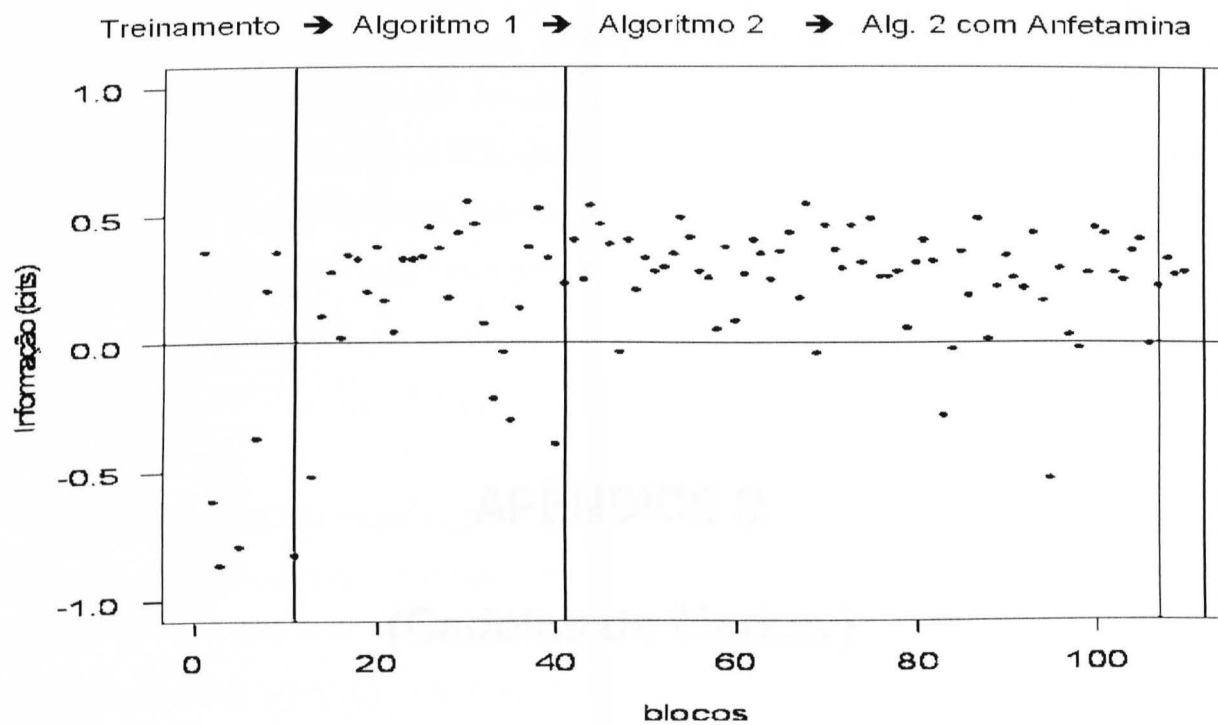


Gráfico C.9. Evolução da informação mútua o rato 23 e por blocos durante o experimento.



01. CADEIAS DE MARKOV

Uma cadeia de Markov é um processo estocástico com um conjunto finito $S = \{1, 2, \dots, N\}$ de estados possíveis em cada instante. O tempo decorre em passos de tempo e, para cada instante t , o estado da cadeia é denotado por X_t . A probabilidade de transição de um estado i para um estado j em um único passo de tempo é denotada por P_{ij} . A matriz de transição P é uma matriz $N \times N$ com elementos P_{ij} e satisfaz a condição $\sum_{j=1}^N P_{ij} = 1$ para todo i . A cadeia de Markov é dita homogênea se as probabilidades de transição não dependem do instante t .

$$P(X_{t+1} = j | X_t = i) = P_{ij} \quad \text{para todo } i, j \in S \text{ e } t \geq 0.$$

APÊNDICE D

(Cadeias de Markov)

D.1. CADEIAS DE MARKOV

Uma seqüência $(Y_n)_{n \geq 0}$ assumindo valores em um conjunto finito $S = \{0, 1, \dots, s\}$ conhecido como espaço de estados, é uma cadeia de Markov de ordem k ($k \geq 0$), caso satisfaça a seguinte propriedade chamada de propriedade markoviana: A probabilidade condicional de Y_n ser igual à $y_n \in S$ dada toda a seqüência de termos anteriores (de Y_0 até Y_{n-1}) é igual à probabilidade condicional de Y_n ser igual à $y_n \in S$ dado os k últimos termos da seqüência, em expressões matemáticas temos:

$$P(Y_n = y_n \mid Y_{n-1} = y_{n-1}, Y_{n-2} = y_{n-2}, \dots, Y_0 = y_0) = P(Y_n = y_n \mid Y_{n-1} = y_{n-1}, \dots, Y_{n-k} = y_{n-k})$$

No caso das escolhas dos ratos, temos que a probabilidade do rato escolher a esquerda no próximo teste, por exemplo, sabendo todo o histórico de escolhas do rato é igual à probabilidade do rato escolher a esquerda no próximo teste dado o conhecimento das últimas k escolhas feitas por ele.

O valor de k representa a ordem da cadeia e pode ser interpretado como o número de escolhas anteriores que influenciam a próxima escolha do rato, por exemplo, uma ordem dois indica que dada toda a seqüência de escolhas do rato, a próxima escolha depende somente das duas últimas.

Como em cada teste o rato pode optar pela direita ou pela esquerda, o espaço de estados é constituído por dois estados, a saber: 0=Esquerda e 1=Direita. A probabilidade condicional de o rato escolher em um dado teste o lado y_k ($y_k=0$ ou $y_k=1$) dado as k escolhas anteriores foi estimada da seguinte forma (estimador de máxima verossimilhança):

$$\hat{p}(y_k \mid y_{k-1}, y_{k-2}, \dots, y_0) = \frac{f_{y_0, \dots, y_{k-1}, y_k}}{f_{y_0, \dots, y_{k-1}}}$$

Sendo que:

$f_{y_0, \dots, y_{k-1}, y_k}$: representa a frequência de ocorrências da seqüência de escolhas sucessivas y_0, \dots, y_k

Assim, a estimativa da verossimilhança é dada por:

$$\hat{L}_k = \prod_D \hat{p}(y_k | y_{k-1}, y_{k-2}, \dots, y_0)^{f_0 \dots f_k}$$

Sendo:

$$D = \{ (y_0, \dots, y_k) ; \hat{p}(y_k | y_{k-1}, y_{k-2}, \dots, y_0) > 0 \}$$

Além disso, temos que:

$$2 \log \frac{\hat{L}_{k+1}}{\hat{L}_k} \rightarrow \chi^2$$

Com $m^k(m-1)^2$ graus de liberdade, onde m é o número de estados do processo (neste caso $m=2$).

D.2. CRITÉRIOS PARA A ESTIMAÇÃO DA ORDEM DE UMA CADEIA DE MARKOV

A seguir são apresentados os três critérios utilizados para a estimação das ordens das cadeias de Markov.

D.2.1. AKAIKE INFORMATION CRITERION (AIC)

Dada uma seqüência de observações assumindo valores em um conjunto finito, podemos ajustar modelos de cadeias de Markov de diversas ordens a esta seqüência.

No entanto, existe uma ordem k_0 tal que a probabilidade condicional da próxima escolha dada as últimas k_0 escolhas é igual à probabilidade condicional da próxima escolha dada toda a seqüência de escolhas anteriores, essa ordem k_0 é a verdadeira ordem da cadeia.

Para comparar uma cadeia de Markov de ordem k com densidade de probabilidade $f(\cdot|k)$ com a cadeia que possui a verdadeira ordem k_0 e densidade de probabilidade $f(\cdot|k_0)$, pode-se utilizar a estatística suficiente minimal [Lehmann and Casella, 1998]:

$$\tau(\cdot) = \frac{f(\cdot|k)}{f(\cdot|k_0)}$$

[Akaike, 1973] propõe uma distância baseada na informação de Kullback-Leibler:

$$\begin{aligned} D(k_0, k) &= 2E_{k_0} \left[\log \left(\frac{f(\cdot|k_0)}{f(\cdot|k)} \right) \right] \\ &= 2E_{k_0} [\log f(\cdot|k_0)] - 2E_{k_0} [\log f(\cdot|k)] \end{aligned}$$

Pela propriedade da Informação de Kullback-Leibler [Kullback, 1959] temos que $D(k_0, k) \geq 0$, com igualdade para $f(\cdot|k) = f(\cdot|k_0)$. Logo, o valor k que maximiza $2E_{k_0} [\log f(\cdot|k)]$ corresponde ao valor que minimiza a distância $D(k_0, k)$ entre as cadeias de ordem k e k_0 (ordem verdadeira).

Assim, para encontrar a verdadeira ordem da cadeia, deve-se minimizar $-E_{k_0} [\log f(\cdot|k)]$ em relação à k . Como não dispomos a princípio de informação sobre $E_{k_0} [\log f(\cdot|k)]$, esta quantidade deve ser estimada a partir de uma amostra aleatória.

[Akaike, 1973] mostra que o estimador de momentos $\frac{1}{N} \sum_{j=1}^N \log \hat{L}_k$ é viesado com viés aproximadamente igual a $-\frac{\gamma}{N}$ para N suficientemente grande, com γ sendo o número de parâmetros da cadeia de Markov de ordem k .

Assim, um estimador assintoticamente não viesado para $-E_{k_0} [\log f(\cdot|k)]$ é dado por:

$$-\frac{2}{N} \sum_{j=1}^N \log \hat{L}_k + \frac{2\gamma}{N}$$

Dessa forma, deve-se escolher a ordem k que minimize a expressão acima, ou simplesmente minimize:

$$AIC(k) = -2 \sum_{j=1}^N \log \hat{L}_k + 2\gamma$$

Com $\gamma = m^k(m-1)$, onde m é o número de estados da cadeia.

Assim, para cadeias de Markov de ordem k , o $AIC(k)$ é dado por:

$$AIC(k) = -2 \sum_D f_{y_0, \dots, y_k} \log \hat{p}(y_k | y_0, \dots, y_{k-1}) + 2m^k(m-1)$$

Sendo:

$$D = \{ (y_0, \dots, y_k) ; \hat{p}(y_k | y_{k-1}, y_{k-2}, \dots, y_0) > 0 \}$$

O critério AIC penaliza o ajuste obtido na verossimilhança com a complexidade do modelo. Cadeias com ordens maiores terão penalização maior.

D.2.2. BAYESIAN INFORMATION CRITERION (BIC)

[Katz, 1981] demonstra que o critério AIC é inconsistente para estimar a ordem da cadeia de Markov, pois superestima a verdadeira ordem mesmo para um grande número de observações. Para contornar esse problema, [Katz, 1981] propõe o critério consistente BIC. Este critério usa um argumento bayesiano, e além de ser consistente é também assintoticamente ótimo no sentido de minimizar a função de perda considerada. O critério BIC, definido abaixo, é similar ao AIC no sentido de penalizar um número excessivo de parâmetros.

$$BIC(k) = -2 \sum_{j=1}^N \log \hat{L}_k + \gamma \log N$$

Com $\gamma = m^k(m-1)$, onde m é o número de estados da cadeia.

O valor \hat{k}_{BIC} que minimiza $BIC(k)$ é o estimador da ordem da cadeia de Markov.

Da expressão do $BIC(k)$, nota-se que $AIC(k) \geq BIC(k)$ para $n \geq 8$

Em [Katz, 1981] o seguinte teorema é demonstrado:

Teorema: A estimativa \hat{k}_{BIC} é consistente para estimar k_0 (verdadeira ordem da cadeia), isto é:

$$\lim_{n \rightarrow \infty} P(\hat{k}_{BIC} = k_0) = 1$$

A idéia do critério BIC parte de considerar uma distribuição a priori α_k para cada possível ordem k da cadeia de Markov ($k \geq 0$), além disso, considera-se a distribuição conjunta a priori μ_k das probabilidades de transição dado que a cadeia é de ordem k . Por simplicidade, consideram-se priores não informativas, por exemplo, para cada linha da matriz de probabilidades de transição é definida uma distribuição Dirichlet com todos os parâmetros iguais a um.

Para uma função de perda fixada, o procedimento consiste em selecionar a ordem \hat{k}_{BIC} que maximiza a probabilidade a posteriori de k (que possui distribuição a priori α_k) dado o conjunto de observações.

Em [Katz, 1981] é demonstrado que para N suficientemente grande, o logaritmo da distribuição a posteriori de k dado o conjunto de observações é aproximadamente igual ao logaritmo da função de verossimilhança de k menos uma quantidade que depende do número de parâmetros e do tamanho da amostra, ou seja:

Para N grande, temos:

$$\log \pi(k | \mathbf{x}) \cong \log \hat{L}_k - \frac{1}{2} m^k (m-1) \log N$$

Sendo que $\pi(k | \mathbf{x})$ é distribuição a posteriori de k dado o conjunto de observações.

Assim, para N grande, maximizar a distribuição a posteriori corresponde a minimizar a expressão do $BIC(k)$ em relação à k . O $BIC(k)$ é muito semelhante ao $AIC(k)$ com a diferença que a penalização do $BIC(k)$ cresce à medida que N cresce.

Quanto mais não informativas forem as distribuições a priori, melhor será a aproximação do resultado anterior, dessa forma, esse critério poderia ser considerado “quasi-bayesiano” ou “pseudo-bayesiano” no sentido de que a distribuições a priori não influencia o resultado obtido.

[Dorea et. Al, 1997] sugere a seguinte interpretação heurística para o BIC:

Uma interpretação heurística do critério BIC que concilia os pontos de vista bayesiano e freqüentista consiste em assumir uma distribuição a priori para μ_k igual ao produto de distribuições uniformes definidas em partições de tamanho u do intervalo $[0,1]$. Segundo o teorema central do limite, para uma amostra de tamanho n , esperamos uma precisão da ordem de $\frac{1}{\sqrt{n}}$ das estimativas. Assim, considerando $u = \frac{1}{\sqrt{n}}$, temos a seguinte distribuição a priori para as probabilidades de transição dado que a ordem da cadeia é igual a k .

$$\mu_k = \left(\frac{1}{\sqrt{n}} \right)^\gamma$$

Onde γ é o número de probabilidades de transição independentes do modelo.

Então, utilizando o teorema de Bayes obtemos a distribuição a posteriori, e tomando o logaritmo da posteriori e multiplicando o resultado obtido por -2, obtém-se a expressão do BIC(k). Logo, maximizar a posteriori corresponde a minimizar o BIC(k).

D.2.3. CRITÉRIO DA MÍNIMA ENTROPIA (MEC)

Considerando $(Y_n)_{n \geq 0}$ uma cadeia de Markov de ordem k . A entropia da cadeia é definida por:

$$H(k) = -E[\log P(Y_n | Y_{n-1}, \dots, Y_{n-k})]$$

Seja x_0, \dots, x_N uma amostra de uma cadeia de Markov de ordem k . Para j fixado definimos:

$$\hat{H}(j) = \sum_{y_0, \dots, y_{j-1} \in S} \hat{p}(y_0, \dots, y_{j-1}) \log \hat{p}(y_j | y_0, \dots, y_{j-1})$$

Caso $j=k$, a expressão acima é um estimador da entropia da Cadeia de Markov.

O critério da mínima entropia toma como ordem da cadeia o valor j que minimiza $\hat{H}(j)$. Esse critério baseia-se na seguinte proposição:

Proposição: Considere $(Y_n)_{n \geq 0}$ um processo estacionário assumindo valores em um conjunto finito S . Definimos $H(0) = -E[\log P(Y_0)]$ e para $j > 0$, $H(j) = -E[\log P(Y_j | Y_0, \dots, Y_{j-1})]$,

então a seqüência $(H_j)_{j \geq 0}$ é monótona decrescente. Ainda $H(k_0-1) > H(k_0) = H(i)$ para algum k_0 e para todo $i \geq k_0$, se e somente se, $(Y_n)_{n \geq 0}$ é uma cadeia de Markov de ordem k_0 .

Segundo a proposição acima, a seqüência $(H_j)_{j \geq 0}$ é decrescente até o termo k_0 , após esse termo, a seqüência passa a ser constante, dessa forma, o termo k_0 é a ordem da cadeia de Markov, ou seja, a ordem de uma cadeia de Markov é dada por:

$$k_0 = \min\{j \geq 0 : H(j+1) = H(j)\}$$

Demonstra-se também que, se $(Y_n)_{n \geq 0}$ é uma cadeia de Markov de ordem k e x_0, \dots, x_N é uma amostra dessa cadeia. Então:

$$2N \left[\hat{H}(k) - \hat{H}(k-1) \right] \rightarrow \chi^2_{m^k (m-1)^k}$$

Esse resultado é demonstrado verificando que a expressão acima é equivalente a estatística do teste assintótico da razão de verossimilhanças.

Com base nesse critério, o seguinte algoritmo deve ser utilizado para escolher a ordem da cadeia:

Considerando as hipóteses a seguir:

H_0 : A cadeia é de ordem k

H_1 : A cadeia é de ordem $k+1$

Os seguintes passos devem ser seguidos:

- 1 – fixar $k=0$
- 2 – testar a ordem k contra $k+1$
- 3 – se o teste não rejeita H_0 , então há evidência de k é a ordem da cadeia.
caso contrário, faça $k \leftarrow k+1$ e volte ao passo 2.

Esse critério tem a desvantagem de não garantir a estimação da verdadeira ordem para seqüências curtas.

APÊNDICE E

(Ajustes das cadeias de Markov)

Tabela E.1.a. Estimativas da ordem da cadeia de Markov das escolhas do Rato 12 durante treinamento.

Ordem	AIC	BIC	Log-verossimilhança	Razão de verossimilhança	P-valor	graus de liberdade
0	92,79	96,09	-45,4	5,12	0,023	1
1	89,67	96,25	-42,83	2,75	0,25	2
2	90,92	104,07	-41,46			

Tabela E.1.b. Probabilidades estimadas das escolhas do Rato 12 durante o treinamento.

Probabilidade da próxima escolha	
Esquerda	Direita
0,06	0,94

Tabela E.2.a. Estimativas da ordem da cadeia de Markov das escolhas do Rato 12 contra o Algoritmo 1.

Ordem	AIC	BIC	Log-verossimilhança	Razão de verossimilhança	P-valor	graus de liberdade
0	3260,7	3266,49	-1629,35	15,99	<0,001	1
1	3246,71	3258,28	-1621,35	32,48	<0,001	2
2	3218,23	3241,37	-1605,11	7,88	0,096	4
3	3218,34	3264,63	-1601,17			

Tabela E.2.b. Probabilidades estimadas das escolhas do Rato 12 contra o Algoritmo 1.

Penúltima escolha	Última escolha	Probabilidade da próxima escolha	
		Esquerda	Direita
Esquerda	Esquerda	0,55	0,45
Esquerda	Direita	0,394	0,606
Direita	Esquerda	0,375	0,625
Direita	Direita	0,365	0,634

Tabela E.3.a. Estimativas da ordem da cadeia de Markov das escolhas do Rato 12 contra o Algoritmo 2.

Ordem	AIC	BIC	Log-verossimilhança	Razão de verossimilhança	P-valor	graus de liberdade
0	8155	8161,7	-4076,5	27,46	<0,001	1
1	8129,54	8142,93	-4062,77	34,28	<0,001	2
2	8099,26	8126,06	-4045,63	3,64	0,457	4
3	8103,62	8157,21	-4043,81			

Tabela E.3.b. Probabilidades estimadas das escolhas do Rato 12 contra o Algoritmo 2.

Penúltima escolha	Última escolha	Probabilidade da próxima escolha	
		Esquerda	Direita
Esquerda	Esquerda	0,515	0,485
Esquerda	Direita	0,41	0,59
Direita	Esquerda	0,407	0,593
Direita	Direita	0,377	0,623

Tabela E.4.a. Estimativas da ordem da cadeia de Markov das escolhas do Rato 12 sob o efeito de anfetamina contra o Algoritmo 2.

Ordem	AIC	BIC	Log-verossimilhança	Razão de verossimilhança	P-valor	graus de liberdade
0	544,9	548,89	-271,45	16,087	<0,001	1
1	530,81	538,79	-263,407	2,91	0,233	2
2	531,91	547,85	-261,95			

Tabela E.4.b. Probabilidades estimadas das escolhas do Rato 12 sob o efeito de anfetamina contra o Algoritmo 2.

Última escolha	Probabilidade da próxima escolha	
	Esquerda	Direita
Esquerda	0,524	0,476
Direita	0,335	0,665

Tabela E.5.a. Estimativas da ordem da cadeia de Markov das escolhas do Rato 22 durante treinamento.

Ordem	AIC	BIC	Log-verossimilhança	Razão de verossimilhança	P-valor	graus de liberdade
0	651,888	656,102	-324,94	20,306	<0,001	1
1	633,581	642,007	-314,79	6,8	0,033	2
2	630,773	647,615	-311,39	6,43	0,169	4
3	632,342	666,011	-308,17			

Tabela E.5.b. Probabilidades estimadas das escolhas do Rato 22 durante o treinamento.

Última escolha	Probabilidade da próxima escolha	
	Esquerda	Direita
Esquerda	0,715	0,285
Direita	0,523	0,477

Tabela E.6.a. Estimativas da ordem da cadeia de Markov das escolhas do Rato 22 contra o Algoritmo 1.

Ordem	AIC	BIC	Log-verossimilhança	Razão de verossimilhança	P-valor	graus de liberdade
0	3620,17	3626,04	-1809,09	137,33	<0,001	1
1	3484,84	3496,58	-1740,42	36,83	<0,001	2
2	3452,01	3475,48	-1722,01	1,58	0,81	4
3	3458,43	3505,36	-1721,21	7,21	0,51	8

Tabela E.6.b. Probabilidades estimadas das escolhas do Rato 22 contra o Algoritmo 1.

Penúltima escolha	Última escolha	Probabilidade da próxima escolha	
		Esquerda	Direita
Esquerda	Esquerda	0,589	0,411
Esquerda	Direita	0,48	0,52
Direita	Esquerda	0,65	0,35
Direita	Direita	0,325	0,675

Tabela E.7.a. Estimativas da ordem da cadeia de Markov das escolhas do Rato 22 contra o Algoritmo 2.

Ordem	AIC	BIC	Log-verossimilhança	Razão de verossimilhança	P-valor	graus de liberdade
0	7206,05	7212,61	-3602,02	54,9	<0,001	1
1	7153,15	7166,26	-3574,57	7,96	0,019	2
2	7149,19	7175,41	-3570,6	6,44	0,1686	4
3	7150,75	7203,2	-3567,37			

Tabela E.7.b. Probabilidades estimadas das escolhas do Rato 22 contra o Algoritmo 2.

Última escolha	Probabilidade da próxima escolha	
	Esquerda	Direita
Esquerda	0,537	0,463
Direita	0,436	0,564

Tabela E.8.a. Estimativas da ordem da cadeia de Markov das escolhas do Rato 22 sob o efeito de anfetamina contra o Algoritmo 2.

Ordem	AIC	BIC	Log-verossimilhança	Razão de verossimilhança	P-valor	graus de liberdade
0	541,21	545,2	-269,6	1,41	0,23	1
1	541,8	549,78	-268,9	13,96	<0,001	2
2	531,83	547,78	-261,9			

Tabela E.8.b. Probabilidades estimadas das escolhas do Rato 22 sob o efeito de anfetamina contra o Algoritmo 2.

Probabilidade da próxima escolha	
Esquerda	Direita
0,5975	0,4025

Tabela E.9.a. Estimativas da ordem da cadeia de Markov das escolhas do Rato 23 durante treinamento.

Ordem	AIC	BIC	Log-verossimilhança	Razão de verossimilhança	P-valor	graus de liberdade
0	780,17	784,56	-389,08	51,32	<0,0001	1
1	730,85	739,64	-363,42	15,24	0,0004	2
2	719,61	737,18	-355,8	14,94	0,005	4
3	712,67	747,8	-348,33	19,21	0,014	8
4	709,46	779,7	-338,73	18,28	0,31	16
5	723,18	863,61	-329,59			

Tabela E.9.b. Probabilidades estimadas das escolhas do Rato 23 durante o treinamento.

		Probabilidade da próxima escolha	
Penúltima escolha	Última escolha	Esquerda	Direita
Esquerda	Esquerda	0,791	0,209
Esquerda	Direita	0,542	0,458
Direita	Esquerda	0,629	0,371
Direita	Direita	0,398	0,602

Tabela E.10.a. Estimativas da ordem da cadeia de Markov das escolhas do Rato 23 contra o Algoritmo 1.

Ordem	AIC	BIC	Log-verossimilhança	Razão de verossimilhança	P-valor	graus de liberdade
0	3327,89	3333,67	-1662,95	44,69	<0,001	1
1	3285,21	3296,77	-1640,6	23,56	<0,001	2
2	3265,65	3288,78	-1628,82	11,86	0,018	4
3	3261,78	3308,04	-1622,89	7,68	0,465	8
4	3270,1	3362,61	-1619,05			

Tabela E.10.b. Probabilidades estimadas das escolhas do Rato 23 contra o Algoritmo1.

Penúltima escolha	Última escolha	Probabilidade da próxima escolha	
		Esquerda	Direita
Esquerda	Esquerda	0,615	0,385
Esquerda	Direita	0,5	0,5
Direita	Esquerda	0,523	0,477
Direita	Direita	0,398	0,602

Tabela E.11.a. Estimativas da ordem da cadeia de Markov das escolhas do Rato 23 contra o Algoritmo 2.

Ordem	AIC	BIC	Log-verossimilhança	Razão de verossimilhança	P-valor	graus de liberdade
0	8155	8161,7	-4076,5	27,46	<0,001	1
1	8129,54	8142,93	-4062,77	34,28	<0,001	2
2	8099,26	8126,06	-4045,63	3,64	0,457	4
3	8103,62	8157,21	-4043,81			

Tabela E.11.b. Probabilidades estimadas das escolhas do Rato 23 contra o Algoritmo2.

Penúltima escolha	Última escolha	Probabilidade da próxima escolha	
		Esquerda	Direita
Esquerda	Esquerda	0,565	0,435
Esquerda	Direita	0,498	0,502
Direita	Esquerda	0,486	0,514
Direita	Direita	0,43	0,57

Tabela E.12.a. Estimativas da ordem da cadeia de Markov das escolhas do Rato 23 sob o efeito de anfetamina contra o Algoritmo 2.

Ordem	AIC	BIC	Log-verossimilhança	Razão de verossimilhança	P-valor	graus de liberdade
0	556,48	560,47	-277,24	3,81	0,051	1
1	554,67	562,64	-275,33			

Tabela E.12.b. Probabilidades estimadas das escolhas do Rato 23 sob o efeito de anfetamina contra o Algoritmo 2.

Probabilidade da próxima escolha	
Esquerda	Direita
0,495	0,505

Tabela E.13. Ordens estimadas das cadeias de Markov das escolhas.

	Treinamento	Algoritmo 1	Algoritmo 2	Alg. 2 (anfetamina)
Rato 12	0	2	2	1
Rato 22	1	2	1	0
Rato 23	2	2	2	0

Tabela E.14.a. Estimativas da ordem da cadeia de Markov dos resultados dos jogos do Rato 12 contra o Algoritmo 2 antes da ingestão de anfetamina.

Ordem	AIC	BIC	Log-verossimilhança	Razão de verossimilhança	P-valor	graus de liberdade
0	8309,51	8311,84	-4153,76	13,08	<0,001	1
1	8298,44	8316,21	-4147,22	2,12	0,35	2
2	8300,32	8327,12	-4146,16			

Tabela E.14.b. Estimativas da ordem da cadeia de Markov dos resultados dos jogos do Rato 12 contra o Algoritmo 2 depois da ingestão de anfetamina.

Ordem	AIC	BIC	Log-verossimilhança	Razão de verossimilhança	P-valor	graus de liberdade
0	550,25	554,24	-274,13	1,37	0,24	1
1	550,87	558,85	-273,44			

Tabela E.15.a. Estimativas da ordem da cadeia de Markov dos resultados dos jogos do Rato 22 contra o Algoritmo 2 antes da ingestão de anfetamina.

Ordem	AIC	BIC	Log-verossimilhança	Razão de verossimilhança	P-valor	graus de liberdade
0	7191,76	7198,31	-3594,88	5,74	0,02	1
1	7188,02	7201,13	-3592,01	3,4	0,18	2
2	7188,63	7214,85	-3590,31			

Tabela E.15.b. Estimativas da ordem da cadeia de Markov dos resultados dos jogos do Rato 22 contra o Algoritmo 2 depois da ingestão de anfetamina.

Ordem	AIC	BIC	Log-verossimilhança	Razão de verossimilhança	P-valor	graus de liberdade
0	553,62	557,62	-275,81	1,57	0,21	1
1	554,05	562,03	-275,02			

Tabela E.16.a. Estimativas da ordem da cadeia de Markov dos resultados dos jogos do Rato 23 contra o Algoritmo 2 antes da ingestão de anfetamina.

Ordem	AIC	BIC	Log-verossimilhança	Razão de verossimilhança	P-valor	graus de liberdade
0	8315,39	8322,1	-4156,7	4,41	0,03	1
1	8312,98	8326,38	-4154,49	3,41	0,18	2
2	8313,57	8340,36	-4152,78			

Tabela E.16.b. Estimativas da ordem da cadeia de Markov dos resultados dos jogos do Rato 23 contra o Algoritmo 2 depois da ingestão de anfetamina.

Ordem	AIC	BIC	Log-verossimilhança	Razão de verossimilhança	P-valor	graus de liberdade
0	554,56	558,55	-276,28	1,95	0,16	1
1	554,6	562,57	-275,3			

APÊNDICE F

(Testes de Homogeneidade)

Tabela F.1. Teste de homogeneidade entre as cadeias de Markov do Rato 12 nos jogos contra os Algoritmos 1 e 2.

Estatística do teste	graus de liberdade	P-valor
0,064	4	0,999

Tabela F.2. Teste de homogeneidade entre as cadeias de Markov do Rato 23 durante o treinamento e nos jogos contra os Algoritmos 1 e 2.

Estatística do teste	graus de liberdade	P-valor
0,91	8	0,999

Tabela F.3. Teste de homogeneidade entre as cadeias de Markov dos Ratos 12 e 23 nos jogos contra os Algoritmos 1 e 2.

Estatística do teste	graus de liberdade	P-valor
0,09	4	0,999

Tabela F.4. Teste de diferença de proporções de recompensas do Rato 12 contra o Algoritmo 2 antes e depois da ingestão de anfetamina.

Estatística do teste	P-valor
6,7	<0,001

Tabela F.5. Teste de diferença de proporções de recompensas do Rato 22 contra o Algoritmo 2 antes e depois da ingestão de anfetamina.

Estatística do teste	P-valor
1,85	0,064

Tabela F.6. Teste de diferença de proporções de recompensas do Rato 23 contra o Algoritmo 2 antes e depois da ingestão de anfetamina.

Estatística do teste	P-valor
3,44	<0,001

APÊNDICE G

(Método para obter estimativas intervalares para as probabilidades de transição)

Segundo os critérios AIC e BIC a ordem 1 é apropriada para a cadeia de Markov que define o processo de escolhas do Rato 22 quando joga contra o Algoritmo 2. Escolhemos então para estimar, a probabilidade de transição p_{00} que corresponde à probabilidade do rato escolher a esquerda no próximo jogo dado que escolheu a esquerda no jogo anterior, e utilizamos o método de reamostragem em blocos para estimar a distribuição de probabilidade dessa probabilidade de transição.

Os blocos de observações foram definidos da seguinte forma:

Foram contadas todas as quatro transições possíveis das escolhas, a saber: esquerda→esquerda, esquerda→direita, direita→esquerda e direita→direita, até a ocorrência de pelo menos 30 de cada tipo, com isso, era definido um bloco. Dessa forma, os blocos têm tamanhos aleatórios, e pela Lei Forte de Markov, as estatísticas de cada bloco são independentes.

Para as 5200 escolhas do Rato 22 contra o Algoritmo 2, foram definidos 59 blocos. Obtendo uma matriz de probabilidades de transição para cada bloco, obtivemos a seguinte distribuição de freqüências absoluta e relativa para p_{00} :

	$p_{00} < 0,3$	$0,3 < p_{00} < 0,4$	$0,4 < p_{00} < 0,5$	$0,5 < p_{00} < 0,6$	$0,6 < p_{00} < 0,7$	$p_{00} > 0,7$
Freq. absoluta	1	10	11	23	11	3
Freq. relativa	0,017	0,169	0,186	0,390	0,186	0,051

Com as seguintes medidas resumo:

Mínimo	1º quartil	Mediana	Média	3º quartil	Máximo
0,2308	0,4545	0,5406	0,5259	0,5882	0,75

Considerando a distribuição de freqüências relativa como sendo o estimador da distribuição de probabilidade de p_{00} , foram realizados os passos a seguir:

1 – Foram obtidas 3000 amostras com 59 observações cada uma com reposição da distribuição de freqüências relativas com reposição e para cada amostra foi obtida a média amostral.

2 – Calculamos a média das 3000 médias amostrais, o valor obtido é a estimativa da probabilidade p_{00}

3 – Calculamos o erro padrão do estimador, ou seja, o desvio padrão das 3000 médias dividido por 3000.

O resultado obtido foi o seguinte:

$$\hat{p}_{00} = 0,5248$$
$$EP(\hat{p}_{00}) = 0,0003$$

Logo, um intervalo de confiança com 95% de confiança é dado por:

$$[0,5242; 0,5254]$$