Análise do problema de Monty Hall: um enfoque bayesiano

Paulo Henrique Borba Florêncio (PUC-GO) – phenrique3103@gmail.com

Agenor Sousa Santos Neto (PUC-GO) – agenor07@hotmail.com

Maria José Pereira Dantas (PUC-GO) – mjpdantas@gmail.com

Resumo: No programa de auditório "Let's Make a Deal", o anfitrião Monty Hall apresentava ao participante três portas. Atrás de uma das portas existia um carro e por trás das outras duas portas encontravam-se bodes. O participante escolhia uma porta, depois Monty abria uma das outras portas e mostrava uma bode. Monty oferecia ao participante a oportunidade de continuar com a mesma porta ou mudar para a outra porta fechada. O participante deveria mudar? Para compreender o Problema de Monty Hall, esta pesquisa relata sob um enfoque bayesiano a solução desse jogo de auditório utilizando o software Excel®, aliado a simulação de dados através de modelos probabilísticos. A utilização destes recursos acelera a compreensão e melhora o entendimento de grande parte do problema, bem como elimina a maior parte do trabalho braçal. Ganha-se qualidade com a substituição de boa parte da exposição oral tradicional por um ensino interativo onde qualquer pessoa pode experimentar e manipular grandes conjuntos de dados diretamente no computador.

Palavras-chave: Monty Hall, Enfoque bayesiano, Simulação, Modelos probabilísticos.

1. Introdução

Vitor e Lemos (2014) afirmam que a origem da probabilidade foi através dos jogos de azar e, por muitos séculos, matemáticos deram contribuições importantes para o desenvolvimento desse conteúdo. Todavia, a concepção usual de modelos estatísticos que dependem de parâmetros arbitrários, embora muito utilizada na prática, é muito polemizada. Wechsler (1993) expõe que para um modelo frequentista, a ideia de probabilidades fixas, porém desconhecidas, direciona à suposição de existência de parâmetros.

Quando se tem um problema com duas alternativas possíveis, primeiro pode-se ponderar que cada uma tem 50% de chance de acontecer, como se fosse um jogo de cara e coroa. Porém nem sempre é assim que as coisas funcionam, isso porque mesmo duas soluções com 50% de chance de acontecer, podem ter pesos diferentes de acordo com o problema. No caso de um governo de um país, restringindo as hipóteses das pessoas que aprovam e das pessoas que não aprovam, podemos ter, por exemplo, 30% e 70%, mas estes dados podem alterar de acordo com as notícias e acontecimentos no país. Por isso, é importante ressaltar que as probabilidades devem ser revistas à medida que as informações são alteradas ou chegam novas informações (SILVER, 2013).

Nesse contexto apresenta-se um jogo conhecido como Problema de Monty Hall no qual há três portas e que atrás de uma delas tenha um carro como prêmio. O apresentador do jogo pede que escolha uma porta. Se escolher a porta que tem o carro, ganha o carro. Entretanto, após escolher a porta, o apresentador abre outra porta do jogo e mostra que aquela porta está vazia. Atrás dela não está o prêmio. O apresentador, então, pergunta: "Quer mudar de porta?". Há apenas duas portas sobrando, aquela que você escolheu primeiro e aquela que não foi aberta. Mudaria de porta? Qual a probabilidade de ganhar o prêmio, se mudar de porta?

Nessa linha, o objetivo da pesquisa é construir um modelo em *software* Excel que caracterize o Problema de Monty Hall por meio da simulação de escolha das portas. Para tanto será aplicada estatística bayesiana integrada à planilha eletrônica como ferramenta para tratar o problema, determinando as consequências lógicas que se têm para cada conjunto de suposições.

O presente artigo encontra-se organizado em cinco sessões principais. A segunda sessão traz a revisão bibliográfica, parte na qual é feita uma descrição geral dos conceitos de Inferência Bayesiana, Simulação de Monte Carlo e Problema de Monty Hall. A terceira sessão evidencia a metodologia utilizada para atingir o objetivo proposto. Na quarta sessão têm-se os principais resultados obtidos a partir da simulação de escolha das três portas. Na sessão cinco são relatadas as conclusões a respeito do tema e a avaliação crítica a respeito dos resultados. Finalmente, tem se as referências bibliográficas, onde é mencionado todo acervo utilizado para embasar a pesquisa sobre o assunto.

2. Referencial teórico

2.1 Inferência Bayesiana

O teorema de Bayes relaciona informações, com a probabilidade de ocorrência, para gerar uma nova probabilidade quando os fatos acontecem de maneira relacionada ou são dependentes. Também é importante ressaltar que essas probabilidades podem e devem ser revistas à medida que são observados novos fatos que podem alterar a probabilidade de cada um dos acontecimentos envolvidos (SILVER, 2013).

Silver (2013) menciona ainda que o teorema de Bayes utiliza dois tipos de inferências estatísticas: as intuitivas e as experimentais. A primeira é adquirida através do conhecimento prévio em relação a situações passadas, desta maneira polde-se fazer uma conjectura sobre a situação e formular uma possível probabilidade. Já a outra é obtida à medida através da experimentação, a partir dos dados são feitos procedimentos estatísticos com os quais se calcula a probabilidade, à medida que novos dados coletados são refeitos os cálculos estatísticos e calculadas as novas probabilidades.

Wheaton (2009) explica que o teorema de Bayes é o fundamento da inteligência competitiva. Pode ser aplicado na análise de decisão, assim quando um produto é lançado, é necessária a formatação correta e adequada do teorema de Bayes para que este mostre probabilidades de modo condizente. Em situações de decisão é altamente recomendável o uso do teorema de Bayes, pois ele contribui para geração do cenário em conjunto com as probabilidades.

Segundo Droguett et al. (2004), uma simples visão do Teorema de Bayes é apresentada através da teoria de probabilidade. Considere que estamos interessados em um determinado evento ***A***, como por exemplo, o nível de confiabilidade a ser atingido por um equipamento, e que ***E*** represente alguma nova informação relevante à avaliação de ***A***, como o resultado de um teste acelerado. Então, o Teorema de Bayes estabelece que a probabilidade do evento ***A*** dado a nova evidência (dados) ***E*** é proporcional ao produto entre a probabilidade do evento ***A*** antesde obtermos a nova informação ***E,*** e a probabilidade de observar a evidência ***E*** caso o evento ***A*** ocorresse.

Ou seja,

$$P(A|E) = \frac{\mathrm{P(E|A)P(A)}}{\mathrm{P(E)}} \qquad (1),$$

onde $P(A)$ é conhecida como a *probabilidade a priori de A*, i.e., antes de tomarmos conhecimento de $\boldsymbol{E}$; $P(E \mid A)$ é a probabilidade de que a evidência $\boldsymbol{E}$ seja observada se $\boldsymbol{A}$ é realmente verdadeiro (ocorre), e é conhecida como *função de verossimilhança* e $P(A| E)$ é a probabilidade *a posteriori* de A, isto é, após termos obtido a nova informação representada por $\boldsymbol{E}$. Logo, $P(A| E)$ representa a nossa probabilidade atualizada sobre o evento $\boldsymbol{A}$ uma vez que obtemos a informação adicional $\boldsymbol{E}$ relevante a $\boldsymbol{A}$.

Para Silver (2013), ao se tratar de problemas buscam-se alternativas conhecidas para solucioná-las, porém também existe o caso em que as alternativas desconhecidas não são imaginadas, por isso muitas vezes não são questionadas outras hipóteses. Porém, dessa maneira podem-se estimar probabilidades de maneira errada, pois se descartam hipóteses por simplesmente desconhecer sua existência.

Gelman (2008) relata que existem duas objeções fundamentais para os métodos bayesianos:

- Primeiro, métodos bayesianos são apresentados como um motor de inferência automática, o que levanta a suspeita de alguém com experiência aplicada, que percebe que diferentes métodos funcionam bem em diferentes configurações. Promove a ideia de que uma multiplicidade de parâmetros pode ser tratada através de modelos hierárquicos, geralmente permutáveis, mas parece improvável que isso poderia realmente funcionar automaticamente. Em contraste, a maior parte dos trabalhos modernos sobre estatísticas Bayesianas são focados no desenvolvimento de métodos que dão respostas razoáveis utilizando-se premissas mínimas.
- Segundo, a objeção à Bayes vem do sentido oposto, por abordar a vertente subjetiva de inferência Bayesiana: a ideia de que as distribuições anteriores e posteriores representam estados subjetivos de conhecimento. Aqui a preocupação principal é que devemos nos preocupar com o conhecimento objetivo, em vez de crença subjetiva, e segundo, que não é clara a forma de avaliar o conhecimento subjetivo em qualquer caso.

Silver (2013) complementa que um problema grave é que nem sempre são mostradas as incertezas de um cálculo, levando as pessoas a pensarem que o número bruto é o valor real e ele acontecerá desta maneira. Para Wheaton (2009) o teorema de Bayes é algo que exige treinamento, pois se necessita estipular as probabilidades iniciais e quanto mais familiarizados com o assunto mais precisam será esta avaliação. Assim, para melhor compreensão do Teorema de Bayes é preciso compreender como o mesmo é calculado.

2.2 Problema de Monty Hall

Trata-se de um programa de auditório onde há três portas iguais, uma das quais escondendo um prêmio. Você precisa escolher a porta com o prêmio ou nada ganha. Diz o problema que, após você ter escolhido uma porta e antes de ser revelado seu conteúdo, o apresentador abre, como de costume, uma das duas outras portas, mostrando-lhe que o prêmio não se encontra ali. Você é, então, perguntado se deseja trocar sua porta inicial por aquela que permanece fechada. A questão é se a troca da porta aumentaria, diminuiria ou em nada alteraria sua chance de ganhar o prêmio escondido (SÁ; SÁ, 2008).

Se o participante acertar a porta onde se encontra o carro, ele será seu. Cada uma dessas portas tem exatamente probabilidade de $\frac{1}{3}$. A grande dúvida acontece quando o apresentador, ciente da porta onde está o carro, abre uma porta que esconde um bode. E agora, qual será a probabilidade de acertar a porta em que está o carro? Será que com duas portas apenas, a probabilidade passa a ser de $\frac{1}{2}$? Segundo Mlodinow (2011), o problema de Monty Hall é difícil de entender, pois não é facilmente percebida a participação do apresentador. Um probleminha maravilhosamente confuso em nenhum outro ramo da matemática é tão fácil para um especialista cometer erros como na teoria da probabilidade. A Figura 1 mostra as probabilidades mencionadas no Problema de Monty Hall.

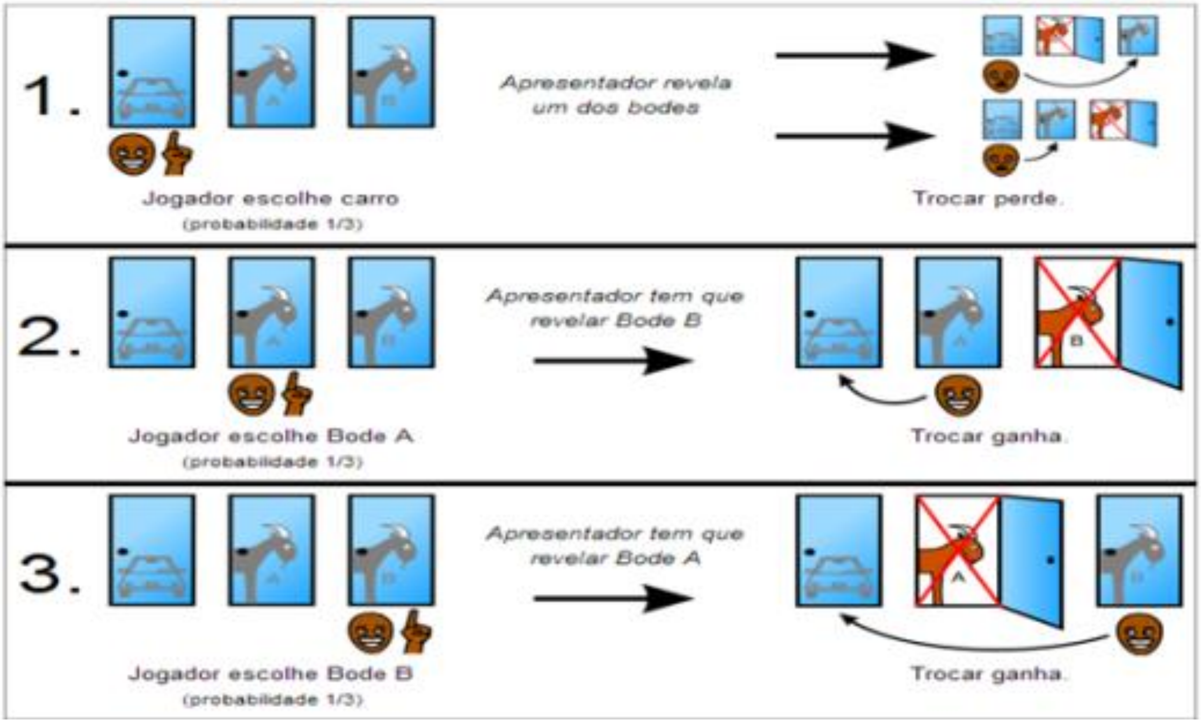

Figura 1- Probabilidade de Acerto para o Problema de Monty Hall.Fonte:Economist (1999)

Este problema de troca de porta pode não ser tão simples. E se o participante tivesse 1.000 portas para escolher, com apenas uma porta que esconde um carro e todos os outros ocultando bodes? Inicialmente, o participante tem uma chance entre em mil de acertar onde está o carro. A partir daqui, o anfitrião sabe onde o carro está e revela 998 bodes, deixando apenas duas portas fechadas. Se o participante não conhece o Problema de Monty Hall, ele pode pensar que as chances de acertar em qual porta está o carro é de 50%. Na verdade o participante deve alternar as portas sempre. Se o participante escolheu o carro inicialmente, ele o perderá caso mude de porta com apenas uma chance em mil. É muito mais provável (999 de 1.000) que o concorrente escolheu inicialmente um bode. Quando o anfitrião revela 998 bodes, o concorrente deve sabiamente mudar da porta inicialmente escolhida (JOHNSON, 2009).

Sá e Sá (2008) demonstram a resolução formal do Problema de Monty Hall através de probabilidades condicionais. Sejam os seguintes eventos:

– *PA*: o prêmio está em *A*
– *PB*: o prêmio está em *B*
– *PC*: o prêmio está em *C*
– *RA*: o apresentador revela o conteúdo (vazio) de *A*
– *RB*: o apresentador revela o conteúdo (vazio) de *B*
– *RC*: o apresentador revela o conteúdo (vazio) de *C*

Se mantivermos o rótulo A para a porta que você escolhe inicialmente e B para a porta vazia aberta pelo apresentador, estamos interessados em $\mathbf{Pr}[PA \mid RB]$. Pela "Regra de Bayes", temos:

$$\mathbf{Pr}[PA \mid RB] = \frac{\mathbf{Pr}[RB \mid PA]\,.\,\mathbf{Pr}[PA]}{\mathbf{Pr}[RB \mid PA]\,.\,\mathbf{Pr}[PA] + \mathbf{Pr}[RB \mid PB]\,.\,\mathbf{Pr}[PB] + \mathbf{Pr}[RB \mid PC]\,.\,\mathbf{Pr}[PC]} \quad (2)$$

$$\mathbf{Pr}[PA \mid RB] = \frac{\left(\frac{1}{2}\right).\left(\frac{1}{3}\right)}{\left(\frac{1}{2}\right).\left(\frac{1}{3}\right) + 0\,.\left(\frac{1}{3}\right) + 1\,.\left(\frac{1}{3}\right)} \quad (3)$$

$$\mathbf{Pr}[PA \mid RB] = 1/3 \quad (4)$$

Como, $\mathbf{Pr}[PA \mid RB] + \mathbf{Pr}[PB \mid RB] + \mathbf{Pr}[PC \mid RB] = 1$ e $\mathbf{Pr}[PB \mid RB] = 0$, concluímos que:

$$\mathbf{Pr}[PC \mid RB] = 1 - \mathbf{Pr}[PA \mid RB] \quad (5)$$

$$\mathbf{Pr}[PC \mid RB] = 1 - 1/3 \quad (6)$$

$$\mathbf{Pr}[PC \mid RB] = 2/3 \quad (7)$$

Sá e Sá (2008) apontam Thomaz Bayes como um dos primeiros a escrever sobre o pensamento matemático na tomada de decisões. Em estudo publicado em 1763, lembra-nos que um dos principais componentes do processo decisório é interpretar corretamente as informações disponíveis e lidar com as incertezas, determinando as probabilidades associadas aos possíveis desfechos a partir de uma ou outra decisão tomada.

2.3 Previsão baseada em simulação

A Simulação é um dos métodos utilizados para previsão de demanda. Taha (2008) define Simulação como uma imitação virtual do comportamento aleatório de um sistema com a finalidade de estimar suas medidas de desempenho. Harrel (2002) afirma que as respostas desse sistema só podem ser determinadas por meio de um modelo detalhado.

Segundo Moreira (2010), a Simulação baseia-se na criação de um modelo que se aproxima da realidade, realizando sucessivas análises dos resultados para que se possa manipular, controlar e compreender o as possibilidades de ocorrência de um resultado específico, geralmente por meio do uso de um computador.

Andrade (2009) evidencia o uso da simulação como um primeiro teste para o esboço de novas políticas e princípios de decisão. Para formular um problema é preciso coletar dados que sejam suficientes, qualitativamente garantidos e significativos ao processo de tomada de decisão. Ao se estabelecer o modelo, torna-se necessário incorporar aos objetivos da simulação, os testes de dados de maneira que se possa constatar sua coerência.

Para Taha (2008) a Simulação pode ser classificada em dois tipos:

a) Modelos Contínuos: estudam sistemas cujo comportamento muda de forma contínua ao longo do tempo. Esses modelos geralmente fazem uso de equações diferenciais para descrever as interações entre os diferentes elementos de um sistema;
b) Modelos Discretos: estudam basicamente filas de espera, com a intenção de estabelecer medidas como o tempo médio de espera e o comprimento da fila.

O uso da simulação tem sido cada dia mais aceito e empregado como uma técnica que propicia aos analistas dos mais diversos segmentos verificarem ou encaminharem soluções com a profundidade desejada, aos problemas com os quais se deparam diariamente. A simulação computacional permite que estudos sejam realizados a partir de sistemas que ainda não existem, levando ao desenvolvimento de projetos eficientes antes que qualquer mudança física tenha sido estabelecida (RYAN, 2006).

Andrade (2009) lista alguns dos benefícios que justificam o uso da simulação, entre eles, estão:

a) Previsão de resultados na execução de uma determinada ação;
b) Redução de riscos na tomada decisão;
c) Identificação de problemas antes mesmo de suas ocorrências;
d) Eliminação de procedimentos em arranjos industriais que não agregam valor a produção;
e) Realização de análises de sensibilidade;
f) Redução de custos com o emprego de recursos (mão de obra, energia, água e estrutura física);
g) Revelação da integridade e viabilidade de um determinado projeto em termos técnicos e econômicos.

Por sua vez, Freitas Filho (2008) elenca algumas das desvantagens do emprego da simulação, as quais se destacam:

a) Treinamento especial para a construção de modelos;
b) Difícil interpretação dos resultados de simulação em razão dos processos aleatórios incluídos no modelo;
c) Consumo de muitos recursos, principalmente tempo, por parte da modelagem e experimentação associadas a modelos de simulação.

Freitas Filho (2008) discorre ainda sobre as situações onde o uso da simulação deve ser considerado:

a) Em momentos onde não exista formulação matemática completa para o problema;
b) Quando não houver solução analítica para o problema;
c) A obtenção de resultados torna-se mais fácil de ser alcançada por meio da simulação do que com o modelo analítico;
d) Habilidade pessoal não deve ser considerada para a resolução de modelos matemáticos por técnicas analíticas ou numéricas;
e) Quando se faz necessária a observação do processo do começo até o seu final, excetuando detalhes mais específicos;
f) Para situações de dificuldade ou mesmo de impossibilidade de experimentação no sistema real;
g) Quando deseja-se observar longos períodos de tempo ou alternativas não apresentadas pelos sistemas reais.

Porém, segundo Law e Kelton (2000), a construção de modelos exige treinamento e experiência prévia e nem sempre a variabilidade de um sistema é bem captada e modelada, podendo levar a resultados equivocados.

Segundo Taha (2008), uma das técnicas mais utilizadas para a construção de modelos de Simulação diz respeito à técnica de Simulação de Monte Carlo.

2.4 Simulação de Monte Carlo

Segundo Oliveira, Barros e Reis (2006), os modelos de simulação podem ser definidos em determinísticos ou probabilísticos. No primeiro caso, os dados são obtidos com maior certeza. Já no segundo, utiliza-se técnica estatística para analisar o comportamento probabilístico das variáveis no sistema. O método de Monte Carlo teve sua origem diante do modelo probabilístico com foco em simulações de fatos aleatórios.

Morais (2010) ressalta que o objetivo da Simulação de Monte é descrever as características possíveis de uma variável dependente *y*, para então estabelecer os possíveis valores das variáveis independentes *x1, x2,..., xn*. Caso em algum modelo a variável *x* apresente variáveis aleatórias, por certo a variável dependente *y* também apresentará variáveis aleatórias.

O nome Monte Carlo teve sua origem em uma roleta de cassino localizado no principado de Mônaco. A criação sistemática do método ocorreu por volta de 1944 quando Von Neumenn nomeou a técnica matemática para resolução de problemas de física nuclear, a qual foi utilizada para a criação da bomba nuclear (LOESCH; HEIN, 2009). Ainda segundo Loesch e Hein (2009), a técnica de Monte Carlo é caracterizada por uma roleta hipotética que leva a resultados aleatórios, de maneira que alguns resultados conduzam a uma determinada interpretação e outros resultem em outras interpretações.

Moreira (2010) afirma que a Simulação de Monte Carlo trabalha a partir da geração artificial de valores das variáveis de interesse, tendo como auxílio números ao acaso ou números aleatórios. Números aleatórios são qualquer sequência numérica em que os números são completamente independentes entre si, sendo gerados por um gerador de número aleatório.

Oliveira, Barros e Reis (2006) enfatizam que deve-se seguir algumas etapas para a execução do Método de Monte Carlo, tais como:

a) Definição das variáveis envolvidas por meio de dados passados,
b) Identificação da distribuição e possibilidade das variáveis aleatórias referentes ao estudo;
c) Construção da distribuição de probabilidades para as variáveis definidas;
d) Definição dos intervalos dos números aleatórios;
e) Geração de números aleatórios;
f) Simulação do experimento.

2.5 Aplicações da Simulação de Monte Carlo

A aplicação de Simulação de Monte Carlo mostra-se bastante eficiente para os sistemas quando se trata de sistemas complexos. A seguir seguem artigos que representam o estado da arte no que diz respeito à simulação aplicada.

Souto (2006) com sua monografia entitulada “Uso de Redes Neurais Artificiais na Simulação Monte Carlo Aplicado ao Problema de Dobramento de Proteínas”, propõe um novo método de otimização do método de Monte Carlo (MC) aplicado ao doramento de proteínas. Este método baseia-se em informações oriundas de Redes Neurais Artificiais (RNAs) treinadas para prever a estrutura secundária de proteínas.

Guimarães (2012) com sua dissertação “Análise do Valor do Negócio de Concessões Rodoviárias Federais: Demanda Simulada pelo Método de Monte Carlo”, propõe uma metodologia usando o método de Monte Carlo para mensurar riscos de demanda de tráfego assumidos no momento da licitação de concessões rodoviárias.

Miranda (2009) com sua pesquisa nomeada “Efeitos de Desordem e Correlação Eletrônica numa Abordagem Local”, estuda os efeitos de desordem nas proximidades da transição metal-isolante de Mott por meio da aplicação do Método de Monte Carlo Quântico, algoritmo Hirsch-Fye, que faz cálculos em temperatura finita.

Santos (2011) em sua tese “Dimensionamento e Análise do Ciclo de Vida de Pavimentos Rodoviários: uma Abordagem Probabilística”, utiliza a Simulação de Monte Carlo para fazer uma avaliação econômica de pavimentos com o propósito de fornecer apoio a decisão quanto à seleção de alternativas de construção mais viáveis.

Aguiar (2013) em seu trabalho “Um Agente Jogador de Go com Busca em Árvore Monte-Carlo Aprimorada por Memória Esparsamente Distribuída”, implementa o SDM-Go, um agente jogador de Go competitivo que procura reduzir a utilização de supervisão no processo de busca pelo melhor movimento. O SDM-Go emprega o modelo de memória esparsamente distribuída como recurso adicional à busca em árvore Monte-Carlo.

Mena e Andrade Filho (2002) por meio de seu artigo “Séries Temporais para Modelos com Parâmetros Aleatórios: uma Abordagem Bayesiana” apresenta uma abordagem bayesiana para fazer inferência sobre os parâmetros de modelos auto-regressivos. Os modelos foram analisados usando técnicas de Simulação de Monte Carlo e a geração de amostras das densidades *a posteriori* permitiram fazer previsões de séries por intermédio das densidades.

3. Metodologia

Para esta pesquisa foi utilizada a estrutura indicada por Montevechi et al. (2010) que sugere a utilização de três etapas para o processo de construção do modelo, são elas: concepção, implementação e análise.

As etapas do projeto de simulação são descritas a seguir:

a) Concepção: Segundo Robinson (2008), o início de qualquer estudo de simulação é compreender o problema e traçar objetivos claros, pois eles guiarão a modelagem, servirão como referencial para a validação do modelo e fornecerá auxílio para a experimentação, bem como, uma métrica para avaliar o sucesso do estudo. Outra questão fundamental é expor as limitações e o nível de detalhe do modelo.
b) Implementação: Chwif e Medina (2007) recomendam que na segunda etapa o modelo conceitual seja transformado em um modelo computacional. A simulação computacional se apresenta como uma poderosa ferramenta, uma vez que favorece a visualização do processo e o acompanhamento dos seus resultados.
c) Análise: A terceira etapa é o momento onde o modelo está pronto para os experimentos, permitindo assim, o início do modelo operacional. Os resultados são analisados e documentados após inúmeras rodadas serem realizadas (CHWIF E MEDINA, 2007).

A pesquisa de Viali (2001) é utilizada como referência para construção da simulação no *software* Excel® referente ao Problema de Monty Hall. Segundo ele, os recursos de simulação Monte Carlo do Excel podem ser divididos em duas categorias. Os que são acionados através do ícone ***fx*** de "**Inserir Função**" e os que fazem parte da biblioteca "**Ferramentas de análise**". Esta biblioteca pode ser acionada através do menu "**Dados**", sub-menu "**Análise de dados ...**". No primeiro caso existem duas funções que desempenham papéis importantes para simular distribuições que o Excel não apresenta no recurso análise de dados. Estas funções são a **ALEATÓRIO()** e a **ALEATÓRIOENTRE(inferior; superior)**.

A função **ALEATÓRIO()** não possui argumentos e gera valores de uma distribuição uniforme no intervalo [0; 1]. Estes números são denominados de pseudo-aleatórios. Esta função é denominada pelo Excel de volátil, isto é, ela será recalculada toda vez que uma célula da planilha for calculada. Para evitar que isto ocorra pode-se transformar a fórmula em um número aleatório e para tanto deve-se selecionar a célula que contém a fórmula, clicar na barra de fórmulas e digitar F9. Com este procedimento, ao se clicar na célula de interesse não mais aparecerá uma fórmula e sim um número que não mais se alterará.

Esta função é a mais importante, pois pode ser utilizada como base para gerar valores de qualquer outra distribuição de probabilidade, como, por exemplo, simular a escolha de portas. Supondo que sejam três portas, pode-se simular a variável "porta escolhida" a partir de: **INT(ALEATÓRIO()*3)**, onde **INT** é o arredondamento para baixo de um número compreendido entre 0 e 1 até o número inteiro mais próximo. A multiplicação por “3” resulta três números possíveis, 0, 1 ou 2, que é justamente a quantidade de portas presentes no jogo.

4. Resultados e discussões

O programa foi criado para gerar amostras de mil jogos. Em 15 simulações, o percentual médio de vezes que a estratégia de troca revelou bem sucedida foi de 67,11%, com um desvio padrão de 0,033%. Em um conjunto separado de 30 simulações, a estratégia de sempre ficar com a escolha original obteve sucesso em 32,89% das vezes.

Na Tabela 1 demonstra-se o modelo computacional para desenvolvimento da planilha de simulação dos mil jogos propostos.

Tabela 1 - Modelo Computacional

	A	B
1	Opções / Qtd. Simulações ==>>	1
2	Porta Premiada	=INT(ALEATÓRIO()*3)
3	Porta Escolhida	=INT(ALEATÓRIO()*3)
4	Porta à direitra da escolhida	=MOD(B3+1;3)
5	Porta à esquerda da escolhida	=MOD(B3-1;3)
6	Porta não premiada (1)	=SE(B4<>B2; B4; B5)
7	Porta não premiada (2)	=SE(B5<>B2;B5;B4)
8	Porta aberta pelo apresentador	=SE(ALEATÓRIO()<choice_prob;B6;B7)
9	Porta escolhida foi aberta?	=SE(B3=B8;1;0)
10	Nenhuma estratagégia levou a ganhar?	=ABS(1-B11-B12)
11	O Candidato não trocou de porta	=SE(B2=B3;1;0)
12	O candidato trocou de porta	=SE(MOD(3-B3-B8;3)=B2;1;0)

Fonte: Autor (2014)

Na Tabela 2 mostra-se o resultado para os 20 primeiros jogos como exemplo dos resultados obtidos em todos os mil propostos.

Tabela 2 - Exemplo do Resultados de 20 Simulações

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	Opções / Qtd. Jogos ==>>	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2	Porta Premiada	1	0	1	0	0	2	2	2	2	1	2	1	2	2	1	2	1	2	2	0
3	Porta Escolhida	2	1	0	2	0	2	1	0	0	0	1	1	2	2	2	0	0	0	0	1
4	Porta à direitra da escolhida	0	2	1	0	1	0	2	1	1	1	2	2	0	0	0	1	1	1	1	2
5	Porta à esquerda da escolhida	1	0	2	1	2	1	0	2	2	2	0	0	1	1	1	2	2	2	2	0
6	Porta não premiada (1)	0	2	2	1	1	0	0	1	1	2	0	2	0	0	0	1	2	1	1	2
7	Porta não premiada (2)	0	2	2	1	2	1	0	1	1	2	0	0	1	1	0	1	2	1	1	2
8	Porta aberta pelo apresentador	0	2	2	1	2	1	0	1	1	2	0	0	1	1	0	1	2	1	1	2
9	Porta escolhida foi aberta?	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	Nenhuma estratagégia levou a ganhar?	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	O Candidato não trocou de porta	0	0	0	0	1	1	0	0	0	0	0	1	1	1	0	0	0	0	0	0
12	O candidato trocou de porta	1	1	1	1	0	0	1	1	1	1	1	0	0	0	1	1	1	1	1	1

Fonte: Autor (2014)

Para melhor compreensão da Tabela 2, considere o seguinte:

- As portas têm os números 0, 1 e 2.
- A "Porta Premiada" é a que tem o prêmio.
- A "Porta Escolhida" é a porta que o candidato escolheu inicialmente.
- A "Porta à direita da escolhida" é a porta que se encontra a direita da que o candidato escolheu.
- A "Porta à esquerda da escolhida" e a porta que se encontra a esquerda da que o candidato escolheu.
- A "Porta não premiada (1)" é uma das portas não escolhidas pelo candidato e que não tem o prêmio.
- A "Porta não premiada (2)" é uma das portas não escolhidas pelo candidato e que não tem o prêmio.
- "O candidato não trocou de porta" tem valor 1 se ele ganhou com essa estratégia e 0 se ele não ganhou com essa estratégia.
- "O candidato trocou de porta" tem valor 1 se ele ganhou com essa estratégia e 0 se ele não ganhou com essa estratégia.

Como exemplo, após executar 15 simulações, cada uma contendo 30 jogos, chegou-se aos seguintes resultados, conforme demonstrado na Tabela 3.

Tabela 3 – Médias para 15 Simulações

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
15	Simulação ======>	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
16	Não Trocou de Porta	4	8	12	5	14	9	9	10	11	12	13	5	9	13	14
17	Trocou de Porta	26	22	18	25	16	21	21	20	19	18	17	25	21	17	16
18	% de acerto quando não trocou de porta	13,33%	26,67%	40,00%	16,67%	46,67%	30,00%	30,00%	33,33%	36,67%	40,00%	43,33%	16,67%	30,00%	43,33%	46,67%
19	% de acerto quando trocou de porta	86,67%	73,33%	60,00%	83,33%	53,33%	70,00%	70,00%	66,67%	63,33%	60,00%	56,67%	83,33%	70,00%	56,67%	53,33%
20	Média quando não trocou de porta	13,33%	20,00%	26,67%	24,17%	28,67%	28,89%	29,05%	29,58%	30,37%	31,33%	32,42%	31,11%	31,03%	31,90%	32,89%
21	Média quando trocou de porta	86,67%	80,00%	73,33%	75,83%	71,33%	71,11%	70,95%	70,42%	69,63%	68,67%	67,58%	68,89%	68,97%	68,10%	67,11%

Fonte: Autor (2014)

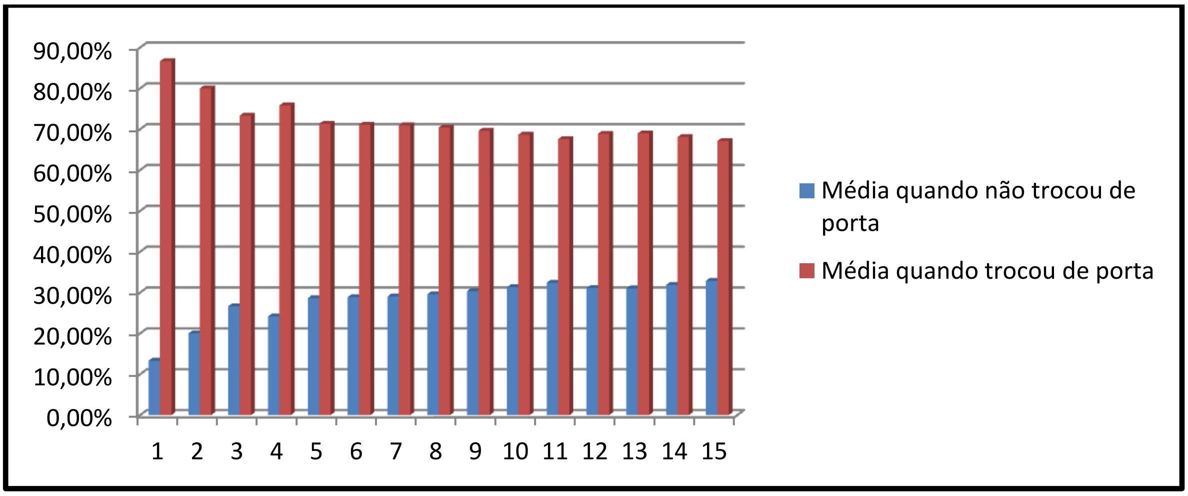

Figura 2 – Comparativo Das Médias. Fonte: Autor (2014)

Com os resultados apresentados na Tabela 3 e apresentados de forma gráfica na Figura 2, pode-se notar que em todas as 15 simulações o percentual de acerto ao se alterar a estratégia, ou seja, trocou de porta após o apresentador abrir uma delas, é maior do que permanecer na escolha inicial. Esse resulta comprova a troca da porta aumenta a chance de ganhar o prêmio escondido.

5. Considerações finais

O argumento falacioso descrito na introdução sugere que, mostrando-nos um bode, Monty não nos deu nenhuma informação nova, já que sabia que havia um bode atrás de uma das portas. Este raciocínio está errado. Simplesmente listando as possibilidades verifica-se que a estratégia de comutação funciona sempre que a estratégia de ficar falha. Ficar claramente sucede uma vez em cada três, de modo que a comutação funciona duas vezes em três, combinando com a evidência experimental. Diferente do que se imagina, que o apresentador ao dar a opção de troca da porta está tentando lhe induzir ao erro, ao aceitar a sugestão dele aumenta-se de 33% para 67% de chance de acerto médio.

Para tal, a aplicação de Simulação de Monte Carlo mostrou-se bastante eficiente na resolução da Problema de Monty Hall. Nota-se que ao simular o jogo *n* vezes a média tende a se equalizar próximo do 1/3 e 2/3 conforme proposto no referencial teórico. Este estudo pode ser utilizado para comprovação de qualquer problema de predição de sequência de variáveis aleatórias binárias (0 ou 1), sob a perspectiva bayesiana, demonstrando a inexistência de permutabilidade entre os resultados obtidos.

Com este artigo explicamos o problema de Monty Hall e, dado as provas com base em um programa computacional, encontramos a resposta correta para o enigma, além de fornecer um tratamento matemático, sugerimos que o conceito intuitivo de escolha restrita é a chave para a compreensão do problema Monty Hall e situações semelhantes.

Referências

AGUIAR, M. A. *Um Agente Jogador de Go com Busca em Árvore Monte-Carlo Aprimorada por Memória Esparsamente Distribuída.* Dissertação de Mestrado em Ciência da Computação. Universidade Federal de Uberlândia, Uberlândia, MG, 2013.

ALMEIDA, A. N. *Teoria dos Jogos: As Origens e os Fundamentos da Teoria dos Jogos.* Centro Universitário Metropolitano de São Paulo, 2006.

ANDRADE, E. L. Introdução à Pesquisa Operacional: Métodos e Modelos para Análise de Decisões. 4 ed. Rio de Janeiro: LTC, 2009.

BERTRAND, J. W. M.; FRANSOO, J. C**.** *Modeling and Simulation: Operations Management Research Methodologies Using Quantitative Modeling.* International Journal of Operation & Production Management**,** v. 22, pg. 241-264, 2002.

CHWIF, L.; MEDINA, A. C. *Modelagem e Simulação de Eventos Discretos: Teoria e Aplicações*. 2 ed. rev. São Paulo: Ed. dos Autores, 2007.

DROGUETT, E. L.; GROEN, F.; MOSLEH, A. *The Combined Use of Data and Expert Estimates in Population Variability Analysis.* Reliability Engineering and System Safety, v. 83, pg. 311-321, 2004.

ECONOMIST. *The Monty Hall Puzzle.* The Economist Newspaper, v. 350, 1999.

FREITAS FILHO, P. J. *Introdução à Modelagem e Simulação de Sistemas: com Aplicações em Arena*. 2 ed. Florianópolis: Visual Books, 2008.

GELMAN, A. Objections to Bayesian Statistics. Journal International Society for Bayesian Analysis, v. 3, pg. 445-450, 2008.

GUIMARÃES, A. G. *Análise Do Valor Do Negócio De Concessões Rodoviárias Federais: Demanda Simulada Pelo Método Monte Carlo.* Dissertação de Mestrado em Transportes. Universidade de Brasília, Brasília, DF, 2012.

HARRELL, C.;GHOSH, B.; BOWDEN, R. *Simulation using ProModel***.** McGraw-Hill, 2000.

JOHNSON, B. *The Monty Hall Problem.* 2009.

LAW, A. M.; KELTON, W. D. *Simulation Modeling and Analysis*. 3 ed. New York: McGraw-Hill, 2000.

LOESCH, C.; HEIN, N. *Pesquisa Operacional: Fundamentos e Métodos*. São Paulo: Saraiva, 2009.

MENA, L.; ANDRADE FILHO, M. G. *Séries Temporais para Modelos com Parâmetros Aleatórios: uma Abordagem Bayesiana.* Acta Scientiarum, v. 24, pg. 1745-1753, 2002.

MLODINOW, L. *O Andar do Bêbado.* Rio de Janeiro: Zahar, 2011.

MONTEVECHI, J. A. B.; PINHO, COSTA, R. F. S.; OLIVEIRA, M. L; SILVA, A. L. F. *Conceptual Modeling in Simulation Projects by Mean Adapted IDEF: an Application a Brazilian Tech Company.* In: Winter Simulation Conference, Baltimore, MD, USA, 2010.

MORAIS, M. F. *Pesquisa Operacional Aplicada - Apostila.* Curso de Engenharia de Produção Agroindustrial. Fecilcam: Campo Mourão, 2010.

MOREIRA, D. A. *Pesquisa Operacional: Curso Introdutório.* 2 ed. São Paulo: Cengage Learning, 2010.

MIRANDA, D. C. B. *Efeitos de Desordem e Correlação Eletrônica numa Abordagem Local.* Dissertação de Mestrado em Física. Universidade Estadual de Campinas, Campinas, SP, 2009.

OLIVEIRA, P. H. D.; BARROS, N. R.; REIS, S. G.. *Aplicabilidade do Método da Simulação de Monte Carlo na Previsão dos Custos de Produção de Companhias Industriais: O Caso Companhia Vale do Rio Doce*. Congresso de Iniciação Científica em Contabilidade, Universidade de São Paulo, VII, São Paulo, 2007.

ROBINSON; S. *Conceptual Modeling for Simulation Part II: a Framework for Conceptual Modeling.* Journal of the Operational Research Society, v. 59, pg. 291-304, 2008.

RYAN, J.; HEAVEY, C. *Process Modeling for Simulation.* Computers in Industry, v. 57, pg. 437- 450, 2006.

SÁ, I. P.; SÁ, V. G. P. *Desafio: A Porta dos Desesperados.* GEPEM, v. 52, 2008.

SANTOS, C. R. G. *Dimensionamento e Análise do Ciclo de Vida de Pavimentos Rodoviários: uma Abordagem Probabilística.* Tese de Doutorado em Engenharia. Universidade São Paulo, São Paulo, SP, 2009.

SILVER, N. *O Sinal e o Ruído: Porque Tantas Previsões Falham e Outras Não.* Rio de Janeiro: Intrínseca, 2013.

SOUTO, A. C. S. *Uso de Redes Neurais Artificiais na Simulação Monte Carlo Aplicado ao Problema de Dobramento de Proteínas.* Dissertação de Mestrado em Computação Aplicada. Universidade do Vale do Rio dos Sinos, São Leopoldo, RS, 2014.

TAHA, H. A. *Pesquisa Operacional.* São Paulo: Pearson, 2008.

VIALI, L. *Utilizando as Planilhas e Simulação para Modernizar o Ensino de Probabilidade e Estatística para os Cursos de Engenharia.* In: Anais do XXIX COBENGE – Congresso Brasileiro de Educação em Engenharia, Porto Alegre, RS, 2001.

VITOR, D. P.; LEMOS, H. C. F. *O Teorema de Monty Hall e o Cálculo da Aproximação de Como Fator de Estímulo ao Ensino da Probabilidade.* Dissertação de Mestrado Profissional em Matemática – PROFMAT. Universidade Federal de São João del-Rei – UFSJ, 2014.

WECHSLER, S. *Exchangeability and Predictivism.* Erkenntnis, v. 38, pg. 243-350, 1993.

WHEATON, K. J.;LEE, J.; DESHMUKH, H. *Teaching Bayesian Statistics to Intelligence Analists: Lessons Learned.* Journal of Strategic Security, v. 2, pg. 39-58, 2009.