

# *Simulação de sistema de caixa de supermercado: validação das postulações\**

**Carlos Alberto Agostini**  
Aluno do programa de pós-graduação em  
Administração da UFRGS

**João Luiz Becker**  
Professor do programa de pós-graduação em  
Administração da UFRGS

---

## *Resumo*

O desenvolvimento de sistemas simuladores, para processos de qualquer natureza, pressupõe a execução de alguns passos já há muito estabelecidos na teoria básica de simulação e modelagem. Destacam-se as fases de definição do problema, concepção do sistema, bem como representação, execução e validação do modelo. Dentre estas etapas, sem minorar a importância das demais, salienta-se, em dificuldade de acesso às informações e em complexidade de análise, a estimação dos parâmetros de entrada, atividade preponderante na fase de representação do modelo. Este trabalho mostra uma série de problemas encontrados na estimação dos parâmetros de entrada de um simulador do sistema de caixa de um supermercado, em processo de automação, e as alternativas encontradas para a sua solução.

## **Palavras-chave:**

- simulação de sistemas
- sistemas de filas
- automação comercial
- pontos-de-venda

---

\* Este artigo tem como origem um segmento de pesquisa que resultou na tese de mestrado de Carlos Alberto Agostini junto ao PPGA/UFRGS, tendo sido apresentado ao IX Encontro Nacional de Engenharia de Produção, realizado em Porto Alegre-RS em setembro de 1989.

## INTRODUÇÃO

O desenvolvimento de sistemas simuladores, para processos de qualquer natureza, pressupõe a execução de alguns passos já há muito estabelecidos na teoria básica de simulação e modelagem. Destacam-se as fases de definição do problema, concepção do sistema, bem como representação, execução e avaliação do modelo. Dentre estas etapas, sem minorar a importância das demais, salienta-se, em dificuldade de acesso às informações e em complexidade de análise, a estimação dos parâmetros de entrada, atividade preponderante na fase de representação do modelo. Uma boa estimação de parâmetros de entrada pressupõe:

- um conhecimento aprofundado do sistema objeto;
- a existência de condições para segregação, no universo de eventos reais, daqueles importantes e indispensáveis para composição do sistema virtual;
- a identificação de relações causais entre variáveis, bem como o estabelecimento correto da ordem das dependências; e
- a disponibilidade para utilização de ferramentas e técnicas de ajustamento de distribuições de probabilidades.

Este artigo mostra uma série de problemas encontrados na estimação dos parâmetros de entrada de um simulador do sistema de caixas de um supermercado, em processo de automação, e as alternativas encontradas para a sua solução. O simulador foi construído dentro de uma perspectiva evolutiva, no que podemos denominar processo de evolução e refinamento do sistema. Este estudo organiza-se da seguinte forma: no próximo tópico descreve-se, brevemente, o sistema objeto da simulação e seus objetivos; no seguinte discute-se a concepção do simulador e seus diversos parâmetros de entrada; o item **Estudo dos Parâmetros de Entrada do Sistema** apresenta a alternativa proposta e alguns testes de validação; finalmente a conclusão do trabalho, colocando-o em perspectiva.

## SISTEMA A SER SIMULADO

A empresa Supermercados Real S/A desponta no Rio Grande do Sul em automação no ramo de supermercados, devendo, a curto prazo, ser a primeira empresa no Brasil a operar com *scanners*. A empresa tem implantados, em uma de suas lojas, terminais ponto-de-venda com digitação de código próprio para controle da movimentação das mercadorias. Estes terminais estão ligados a microcomputadores que controlam o sistema.

A empresa desejava uma metodologia de avaliação e validação de determinadas técnicas que serão, possivelmente, incorporadas a essa nova tecnologia. Questões tradicionais de um sistema de caixa de supermercado eram colocadas lado a lado com outras recém-surgidas, em nível de Brasil, no que concerne à escolha de alguns caminhos a serem seguidos pela automação comercial no intento de adaptar-se à realidade do país. Existem dúvidas quanto ao ganho relativo de tempo com a utilização de *scanners* para leitura de código de barras, contra a digitação do correspondente código numérico, assim como quanto ao ganho relativo de tempo

com o aperfeiçoamento na forma de pagamento (eventual automação na emissão de cheques, por exemplo). Por outro lado, dificuldades operacionais (novos gargalos no sistema, por exemplo) podem ser provocadas pelo eventual ganho de tempo conseguido.

Para responder a estas e outras questões, optou-se pelo desenvolvimento de um simulador do sistema de caixas. O objetivo geral de tal simulador é reproduzir, em laboratório, todas as operações consideradas relevantes, para fins de tomada de decisão, realizadas nos caixas em uma loja do supermercado no período de um dia.

## CONCEPÇÃO DO SIMULADOR

O simulador do sistema de caixas foi dividido em três subsistemas:

- um conjunto de rotinas geradoras de eventos que têm como entrada os parâmetros que estabelecem o comportamento das diversas entidades que intervêm no sistema;
- um conjunto de rotinas para replicar o comportamento dinâmico da frente de caixa;
- um conjunto de rotinas para contabilizar as várias estatísticas de interesse.

Enquanto os dois últimos subsistemas não sofrem modificações significativas quando da migração de uma realidade para outra, o primeiro subsistema pode mudar, radicalmente, em função das características próprias dos usuários e dos gerenciadores do ambiente. Assim, reputa-se importante modelar com precisão as atividades da frente de caixas e isolar, corretamente, as variáveis de interesse. É fundamental observar e avaliar os parâmetros definidores de comportamento para uma perfeita identificação de cada meio ambiente.

Para satisfazer os propósitos do estudo, os seguintes parâmetros foram reconhecidos como importantes: taxa de chegada de clientes no sistema, quantidade de itens comprados pelos clientes, tempo de digitação, forma de pagamento utilizada pelos clientes e tempo de pagamento.

## ESTUDO DOS PARÂMETROS DE ENTRADA DO SISTEMA

Uma vez esboçado o simulador e decidido quais seriam os parâmetros de entrada do sistema, foram realizadas diversas análises exploratórias com o intuito de determinar possíveis relações de dependência entre os mesmos. Este tópico apresenta uma descrição do comportamento dos diversos parâmetros estudados, a modelagem utilizada e os resultados de diversos testes de ajustamento.

### Processo de chegada

#### Descrição

Por processo de chegada entende-se a forma pela qual os clientes chegam ao sistema. Diversas observações e ponderações foram realizadas para elucidação do comportamento deste parâmetro. Para uma perfeita identificação, seria ne-

cessária a percepção do momento exato em que cada cliente, após coletar os produtos nas gôndolas, chegasse ao sistema de caixas para ser atendido. Infelizmente, não dispomos de tão perfeita informação.

Com base em considerações, tanto teóricas como práticas, optou-se por conceber que os clientes chegam individualmente e independentemente uns dos outros. Em termos de modelagem, isto significa considerar que o processo de chegada é equivalente a um processo de *Poisson*, com tempos entre chegadas independentes e identicamente distribuídas segundo a distribuição exponencial. A taxa de chegada dos clientes pode ser avaliada com relativa precisão utilizando-se uma simples contagem do número de clientes que passam pelo sistema a intervalos regulares de tempo.

Uma dificuldade adicional, entretanto, é logo percebida: a taxa de chegada de clientes apresenta um comportamento essencialmente não-estacionário, como se depreende da figura 1. Ou seja, ela depende fundamentalmente do instante no tempo em que está sendo considerada. Trabalhos de cunho semi-científico ou operacional têm tratado este parâmetro como estacionário, quando não ao longo de todo, pelo menos em largos períodos do processo simulado. Tal tratamento produz resultados, senão desastrosos, no mínimo desprestigiadores do potencial resolutivo da técnica de modelagem e simulação(\*).

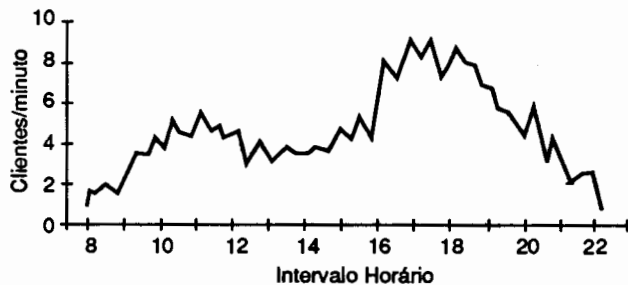


Figura 1: Taxas de Chegada por Intervalo Horário

Seguindo a intuição, bem como orientação contida em Law & Kelton (1982), modelou-se o processo de chegadas por um não-estacionário de *Poisson*(\*\*). Isto é, à medida que o relógio do simulador avança, a determinados intervalos de tempo (em geral de 15 minutos), reavalia-se a taxa de chegada para geração de futuras chegadas. Tentativas de implementar reavaliações a intervalos maiores (30 e 60 minutos) mostraram-se infrutíferas, em virtude da perda de controle sobre eventuais variações cíclicas importantes dentro do período.

(\*) A utilização de valores estacionários para qualquer parâmetro deve ser precedida de uma rigorosa avaliação das disfunções que poderão resultar. Eventuais dificuldades de identificação da correta relação causal existente entre variáveis não servem como desculpa para uso de taxas estacionárias. Ao se proceder assim, corre-se o risco de, no processo de validação do modelo, ver perdido todo um esforço de pesquisa.

(\*\*) Uma decorrência, não tão facilmente perceptível desta formulação, é a dificuldade de obtenção de hipóteses testáveis, para fins de validação, uma vez que não se está postulando uma determinada (estacionária) distribuição de probabilidades para o tempo entre chegadas. Restam, entretanto, testes de aderência ao processo como um todo, como relatamos no item *Validação das Postulações*.

## Validação das postulações

Foram realizadas 15 replicações independentes (isto é, simulações completas de um dia inteiro usando-se sementes aleatórias distintas) do processo modelado, com a intenção de comparar com a amostra originalmente colhida. Agregou-se o número total de clientes chegados por intervalo de tempo e testou-se a semelhança de cada uma destas distribuições (não-estacionárias) com a distribuição empiricamente observada. Para isto fez-se uso do teste de Kolmogorov-Smirnov (K-S) para duas amostras (veja Siegel, 1981). A hipótese nula do teste K-S é que as duas amostras estudadas são provenientes da mesma população, o que não foi rejeitado nos 15 testes realizados, a um nível de significância de 5%. A figura 2 ilustra a aderência encontrada.

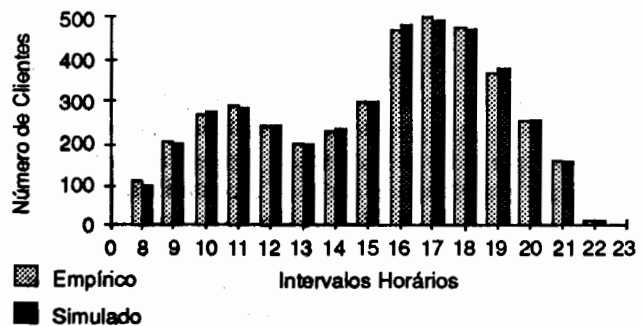


Figura 2: Número de Clientes por Intervalo Horário: Empírico X Modelado

## Quantidade de itens

### Descrição

A distribuição da quantidade de itens que cada cliente transporta é função do horário em que ele chega ao sistema de caixas, conforme se depreende dos gráficos apresentados na figura 3(\*). Cada curva representa a distribuição acumulada de probabilidades da quantidade de itens em um determinado intervalo horário. Por exemplo, a curva representada pelo símbolo □ fornece a probabilidade acumulada até cada quantidade de itens no intervalo das 11 às 12 horas. Assim sendo, a probabilidade de o cliente transportar até 6 itens é de 74%. Uma curva inicialmente mais baixa mostra que, naquele intervalo horário, a probabilidade de o cliente levar mais itens é maior, como é o caso da curva das 17 às 18 horas, representada pelo símbolo ■.

Tentativas de modelagem do sistema usando-se uma distribuição estacionária (média das distribuições acumuladas por intervalo horário) revelaram-se desastrosas. Optou-se, então, pela incorporação da não-estacionariedade ao modelo.

A não-estacionariedade é bastante clara até a quantidade de 30 itens. Acima de 30 itens o conjunto de dados coletados não permite extrair inferências por intervalo horário, em

(\*) A figura apresenta apenas um subconjunto de todos os gráficos gerados, por questões de clareza.

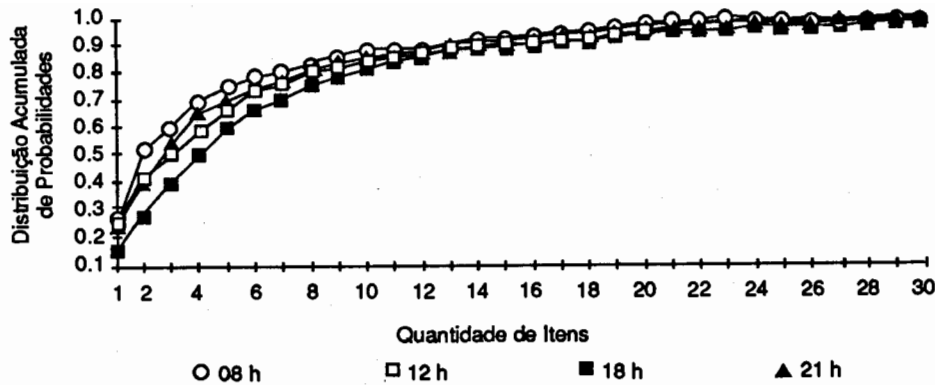


Figura 3: Probabilidades Acumuladas de Quantidades de Itens: Até 30 Itens

função da dispersão que os dados apresentam. Como solução para este caso, optou-se por uma análise estacionária em relação ao tempo, projetando uma distribuição de probabilidade única para todos os horários. A figura 4 apresenta esta distribuição. Em 90 itens a curva de probabilidades acumuladas atinge a marca de 0,9798. A área excedente, de 0,0202, é atingida pelo modelo com uma cauda exponencial passando pelo ponto superior da curva. Com isto oferece-se a possibilidade, ainda que pequena, de algum cliente levar consigo mais de 90 itens. Esta possibilidade, efetivamente, deve existir em teoria, ainda que não tenha sido observada durante a coleta de informações.

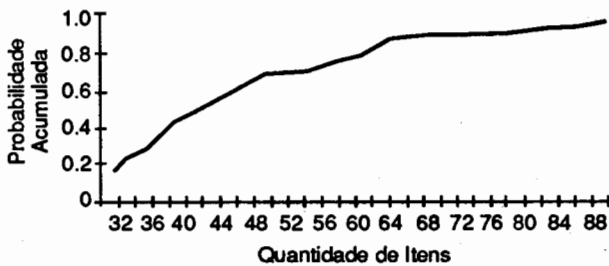


Figura 4: Probabilidades Acumuladas de Quantidades de Itens: Mais de 30 Itens

Foi examinada, ainda, a hipótese de dependência, direta ou supletiva, da forma de pagamento, parâmetro estudado mais adiante. A análise realizada indicou ser melhor inverter-se a relação de dependência. Ou seja, o parâmetro forma de pagamento é que será modelado como dependente, supletivamente, da quantidade de itens transportada pelo cliente.

#### Validação das postulações

Para testar a validade das pressuposições realizadas, foram efetuados 2 conjuntos de 15 replicações independentes do processo modelado. O primeiro conjunto de replicações foi usado para testar a semelhança de cada uma das distribuições do número de clientes por quantidade de itens, com a distribuição empiricamente observada. Novamente, utilizou-se o teste K-S para duas amostras e os resultados encontrados

não permitem rejeitar, a um nível de significância de 5%, a hipótese nula. A figura 5 ilustra a aderência encontrada.

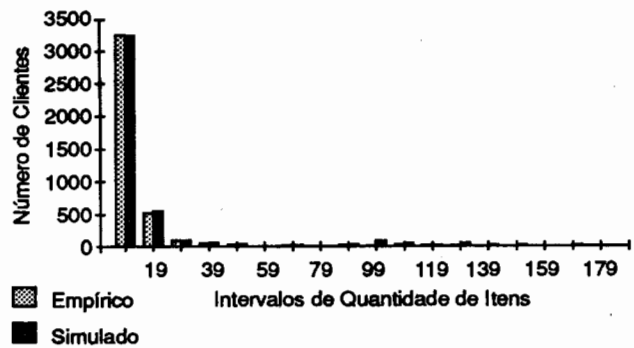


Figura 5: Número de Clientes por Quantidade de Itens: Empírico X Modelado

O segundo conjunto de replicações foi usado para testar a semelhança de cada uma das distribuições do número total de itens agregados por intervalo horário (não-estacionárias), com a distribuição empiricamente observada. Também neste caso utilizou-se o teste K-S para duas amostras e os resultados encontrados não permitem rejeitar, a um nível de significância de 5%, a hipótese nula. A figura 6 ilustra a aderência encontrada.

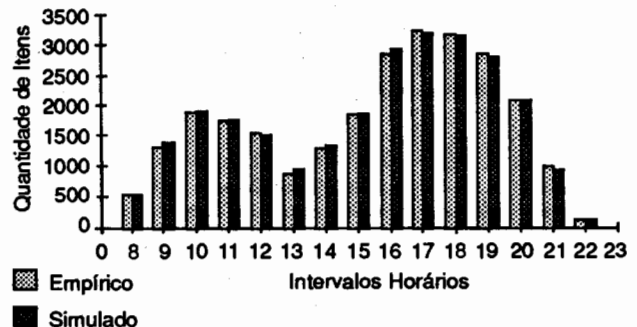


Figura 6: Total de Itens por Intervalo Horário: Empírico X Modelado

Conclui-se, assim, que o modelo consegue simular, razoavelmente, a realidade, no que diz respeito ao parâmetro quantidade de itens.

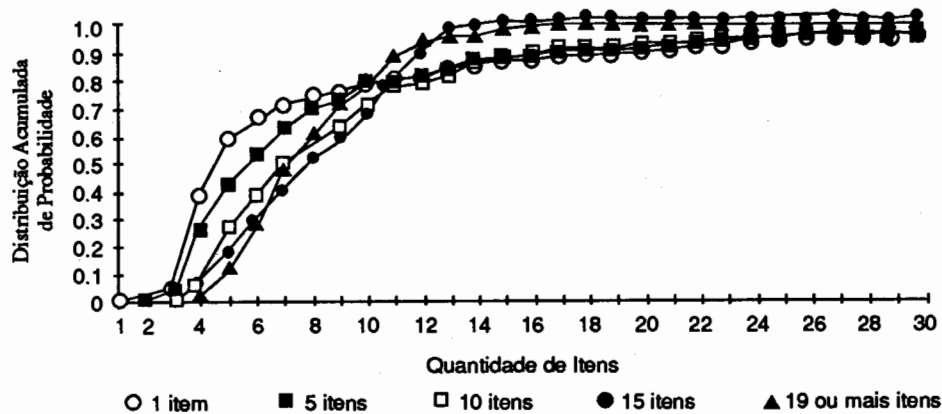


Figura 7: Probabilidades Acumuladas de Tempos Médios de Digitação

## Tempo de digitação

### Descrição

Tempo de digitação é o tempo que o operador do caixa gasta para digitar a totalidade dos itens transportados pelo cliente. Como parâmetro, é medido em termos de média por item digitado. Esta média apresenta correlação com a quantidade de itens que o cliente transporta. Subsidiariamente, mantém também relação de dependência com a hora em que é iniciado o serviço. Esta segunda relação, entretanto, foi considerada como de efeito secundário e desprezada em termos de sistema.

A figura 7 apresenta algumas distribuições acumuladas de probabilidade de tempos médios de digitação por quantidade de itens. A relação de dependência com a quantidade de itens está bem evidenciada. Também aqui, as tentativas de modelagem, independentemente da quantidade de itens, revelaram-se infrutíferas, tendo-se optado pela incorporação das mesmas ao modelo.

É interessante observar, ao contrário do que poder-se-ia esperar, a redução de probabilidade de baixos tempos médios de digitação quando se observam grandes quantidades de itens digitados. Por exemplo, a probabilidade de que o tempo médio de digitação seja menor ou igual a 5 segundos é de 59% para clientes que carregam consigo apenas 1 item e de 52% para quem leva 2 itens, caindo para apenas 18% para clientes que levam consigo 18 ou 19 itens. Também digno de nota é o fato de as distribuições associadas a grandes volumes de itens apresentarem menores variâncias. Tal fato, entretanto, não causa surpresa, em virtude da lei dos grandes números.

A probabilidade de tempos (médios por item) de digitação elevados (acima de 30 segundos) é praticamente nula para grandes quantidades de itens. Entretanto, esta probabilidade não é negligenciável para pequenas quantidades de itens. Devido ao escasso número de observações com tempos médios de digitação superiores a 30 segundos, optou-se neste caso por modelar a distribuição acumulada de tempos independentemente da quantidade de itens. Tempos médios de digitação superiores a 90 segundos podem ser consegui-

dos pelo modelo usando-se o artifício de ajustamento de uma cauda exponencial.

### Validação das postulações

Para testar a validade das pressuposições realizadas, foram efetuados 2 conjuntos de 15 replicações independentes do processo modelado. O primeiro conjunto de replicações foi usado para testar a semelhança de cada uma das distribuições do número de clientes por tempo médio de digitação, com a distribuição empiricamente observada. Novamente, utilizou-se o teste K-S para duas amostras e os resultados encontrados não permitem rejeitar, a um nível de significância de 5%, a hipótese nula. A figura 8 ilustra a aderência encontrada.

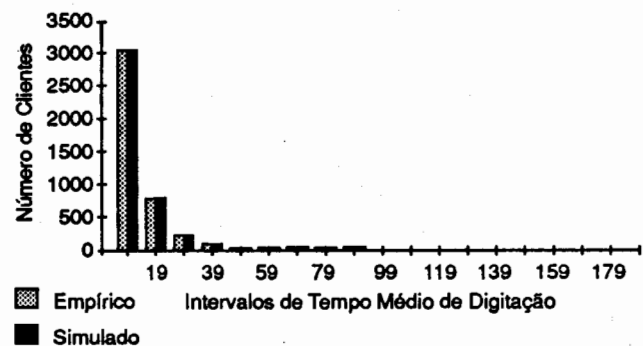


Figura 8: Número de Clientes por Tempo de Digitação: Empírico X Modelado

O segundo conjunto de replicações foi usado para testar a semelhança de cada uma das distribuições do tempo total de digitação agregadas por intervalo horário (não-estacionárias), com a distribuição empiricamente observada. Também neste caso utilizou-se o teste K-S para duas amostras e os resultados encontrados não permitem rejeitar, a um nível de significância de 5%, a hipótese nula. A figura 9 ilustra a aderência encontrada.

Conclui-se, assim, que o modelo consegue simular, razoavelmente, a realidade, no que diz respeito ao parâmetro tempo de digitação.

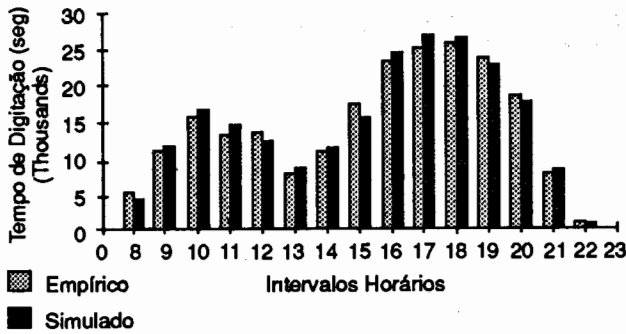


Figura 9: Total de Digitação por Intervalo Horário: Empírico X Modelado

### Forma de pagamento

#### Descrição

Este parâmetro trata da decisão sobre o meio utilizado pelo cliente para saldar seu débito. O meio utilizado pode ser qualquer um dos admitidos pela administração do estabelecimento ou prescritos em lei. As técnicas do meio escolhido são analisadas, posteriormente, de forma individual.

Os meios estabelecidos pela administração, na data da amostragem, eram: dinheiro e cheque. Cabe salientar, entretanto, que apesar de as informações colhidas não revelarem, especificamente, houve diferentes submodalidades dentro das duas especificadas. No pagamento em cheque os procedimentos de conferência foram segregados por valor (recebimento autorizado pelo próprio caixa ou pelo fiscal) e no pagamento em dinheiro estão incluídas outras formas de pagamento que lhe são similares, tais como: cupons e vales.

No caso do pagamento em dinheiro não foram percebidas quaisquer anomalias, o que permite deduzir que os procedimentos para vales, cupons e dinheiro são iguais (ou desiguais de forma não significativa). Quanto ao pagamento em cheque, entretanto, estabeleceu-se uma clara bimodalidade na curva de freqüência do tempo de pagamento. Tendo em vista que a amostra não segrega as submodalidades, não foi possível estabelecê-las individualmente. Contudo, ela está avaliada de forma agregada. Assim, quando se fala em modalidade cheque, fica desde já entendido que se trata da união de dois conjuntos de procedimentos: o do cheque autorizado pelo caixa e o do cheque autorizado pelo fiscal de frente de caixas. Pesquisas posteriores devem atentar para as submodalidades e procurar distinguir os comportamentos individuais.

A escolha da forma de pagamento é crucial para a avaliação do tempo de pagamento, componente importante na determinação do tempo de serviço no sistema de caixas. Parece claro que o simulador deverá realizar esta escolha com base na probabilidade (estimada pela proporção) de utilização das diversas formas de pagamento permitidas. Algumas dificuldades são previsíveis.

A figura 10 apresenta a proporção amostral da forma de pagamento (dinheiro ou cheque), onde ficam estabelecidas as principais relações de dependência.

A partir de 30 itens a avaliação da proporção fica prejudicada pela exigüidade das amostras individuais. A forma de pagamento, além de ser dependente da quantidade de itens,

mantém uma segunda dependência, de intensidade menor, porém muito significativa, em relação ao horário em que o cliente inicia a utilização do serviço. Este múltiplo relacionamento só pode ser avaliado tridimensionalmente até a quantidade de 5 itens, pois acima deste volume a amostragem individualizada não é representativa.

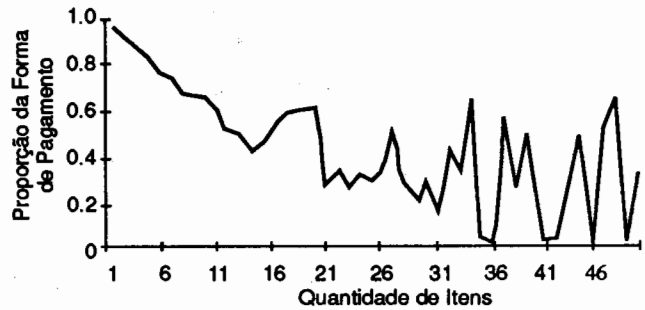


Figura 10: Proporção de Pagamento em Dinheiro por Quantidade de Itens

Para superar este impasse, para quantidades acima de 5 itens, utiliza-se um modelo multiplicativo, como segue. Projetou-se, empiricamente, uma curva de proporções por quantidades de 6 a 16 itens e uma função estacionária para 17 ou mais itens. Posteriormente, extraiu-se da amostra a proporção horária para quantidades maiores de 5 itens. Assim, para estimar-se a probabilidade de pagamento de uma determinada forma, primeiro obtém-se a proporção em função do volume de itens comprados para, em seguida, aplicar-se a esta proporção o multiplicador do respectivo horário.

#### Validação das postulações

Para testar a validade das pressuposições realizadas, foram efetuados 2 conjuntos de 15 replicações independentes do processo modelado. O primeiro conjunto de replicações foi usado para testar a semelhança de cada uma das distribuições do número de clientes que pagaram em dinheiro por quantidade de itens carregados, com a distribuição empiricamente observada. Novamente, utilizou-se o teste K-S para duas amostras e os resultados encontrados não permitem rejeitar, a um nível de significância de 5%, a hipótese nula. A figura 11 ilustra a aderência encontrada.

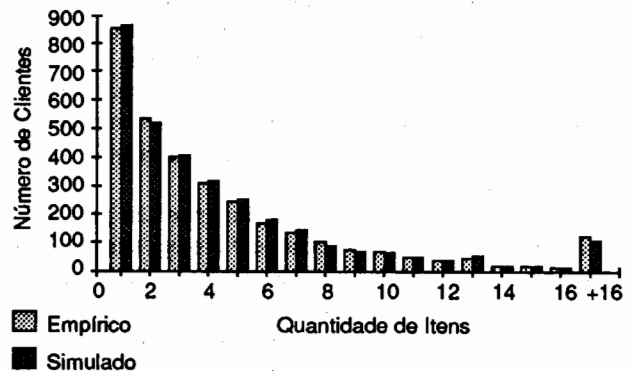


Figura 11: Pagamento em Dinheiro por Quantidade de Itens: Empírico X Modelado

O segundo conjunto de replicações foi usado para testar

a semelhança de cada uma das distribuições do número de clientes que pagaram em dinheiro por intervalo horário (não-estacionárias), com a distribuição empiricamente observada. Também neste caso utilizou-se o teste K-S para duas amostras e os resultados encontrados não permitem rejeitar, a um nível de significância de 5%, a hipótese nula. A figura 12 ilustra a aderência encontrada.

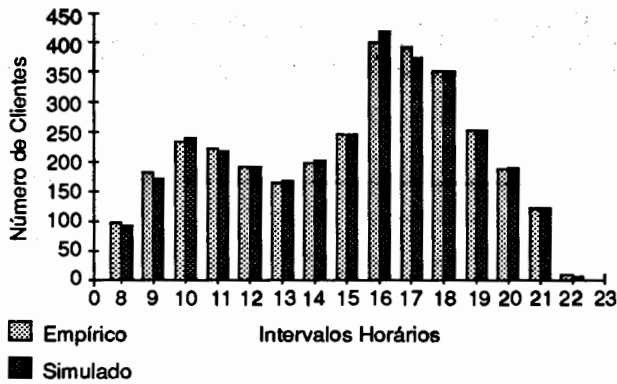


Figura 12: Pagamento em Dinheiro por Intervalo Horário: Empírico X Modelado

Conclui-se, assim, que o modelo consegue simular, razoavelmente, a realidade, no que diz respeito ao parâmetro forma de pagamento.

#### Tempo de pagamento

Tanto o tempo de pagamento em dinheiro como o em cheque apresentam correlações com o intervalo horário no qual o cliente inicia a utilização do serviço de caixas (não-estacionariedade). A metodologia utilizada para isolar e medir estas dependências é idêntica à aplicada para determinar o comportamento da quantidade de itens. Em ambos os casos,

os testes realizados apontam para uma adequação razoável do modelo à realidade.

## CONCLUSÕES

Este trabalho apresentou as principais pressuposições admitidas para os diversos parâmetros de entrada de um simulador do sistema de caixas de um supermercado, bem como os esforços de validação realizados. A definição clara dos parâmetros de entrada de qualquer sistema, além de estar estritamente ligada à qualidade dos resultados produzidos, constitui-se em um campo extraordinário para compreensão e segregação das relações que, posteriormente, irão formar o modelo lógico do simulador.

As funções clássicas de probabilidades, que tão bem replicam comportamentos em variadas circunstâncias de pesquisa, mostraram-se pouco produtivas neste estudo, apesar do grande esforço aplicado na tentativa de obtenção de algum ajustamento. Conseqüentemente, postularam-se distribuições empíricas de probabilidades, com ajustamento exponencial de suas caudas. Grande número de relações de dependência foi, também, postulado. Em geral, os testes de validação comprovam o acerto das pressuposições.

O trabalho de análise e segregação de parâmetros, em qualquer sistema econômico-social, é uma arte que requer tempo e paciência, com constantes revisões do caminho percorrido. Cada nova hipótese que se deseje estabelecer como verdade deve, no mínimo, não ir de encontro às já, anteriormente, verificadas. Após um trabalho estafante, temos a satisfação de apresentar um conjunto harmônico de verdades que nos proporcionam um sistema imaginário simulador do correspondente sistema real, onde a significância do todo depende do esforço aplicado na depuração de cada uma das importantes partes.

### Abstract

A typical sound simulation study usually comprises several basic steps, such as problem formulation, system conception and model representation, along with its validation and execution. During the phase of model representation, a great deal of effort is normally made to estimate input parameters, which is specially noticeable because the difficulty in the access of information and complexity of analysis. This paper shows a number of problems faced in the input parameter estimation stage of a simulation study of the check-out system of a supermarket going through a process of automation, and the alternatives generated for its solution.

### Uniterms:

- system simulation
- queuing systems
- automation in the retail industry
- check-out points

### Referências Bibliográficas

LAW, A. M. & W. D. KELTON. *Simulation modeling and analysis*. New York,

McGraw-Hill, 1982.  
SIEGEL, S. *Estatística não-paramétrica para as ciências do*

*comportamento*. São Paulo, McGraw-Hill do Brasil, 1981.

Recebido em outubro/89