

UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA DE PRODUÇÃO

A Importância da Rapidez de Atendimento nos Caixas de Supermercados: um estudo de caso utilizando um modelo analítico de filas com trocas

José Gilberto Spasiani Rinaldi

São Carlos – SP
2007

UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA DE PRODUÇÃO

A Importância da Rapidez de Atendimento nos Caixas de Supermercados: um estudo de caso utilizando um modelo analítico de filas com trocas

José Gilberto Spasiani Rinaldi

Tese de doutorado apresentada ao Programa de Pós-Graduação em Engenharia de Produção da Universidade Federal de São Carlos, como parte dos requisitos para a obtenção do título de doutor em Engenharia de Produção, sob a orientação do Prof. Dr. Reinaldo Morabito.

São Carlos – SP
2007

**Ficha catalográfica elaborada pelo DePT da
Biblioteca Comunitária/UFSCar**

R578ir

Rinaldi, José Gilberto Spasiani.

A importância da rapidez de atendimento nos caixas de supermercados : um estudo de caso utilizando um modelo analítico de filas com trocas / José Gilberto Spasiani Rinaldi. -- São Carlos : UFSCar, 2007.
175 f.

Tese (Doutorado) -- Universidade Federal de São Carlos, 2007.

1. Pesquisa operacional. 2. Supermercados. 3. Qualidade de atendimento. 4. Análise de correspondência. 5. Sistemas de filas com trocas. I. Título.

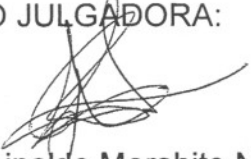
CDD: 658.4034 (20^a)




FOLHA DE APROVAÇÃO


Aluno(a): José Gilberto Spasiani Rinaldi


TESE DE DOUTORADO DEFENDIDA E APROVADA EM 29/06/2007 PELA
COMISSÃO JULGADORA:

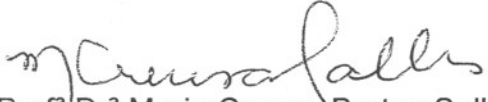

Prof. Dr. Reinaldo Morabito Neto
Orientador(a) PPGE/UFSCar

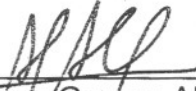

Prof. Dr. Fernando Yassuo Chiyoshi
COPPE/UFRJ


Prof. Dr. Solon Venâncio de Carvalho
INPE/LAC


Prof. Dr. Francisco Louzada Neto
DEs/UFSCar


Profª Drª Vilma Mayumi Tachibana
DEMEC-FCT/UNESP


Profª/Drª Maria Creusa Bretas Salles
ICMC/USP


Prof. Dr. Alceu Gomes Alves Filho
Coordenador do PPGE

DEDICATÓRIA

À RENATA

Companheira de todos os momentos... com quem compartilho aprendizagens diárias, principalmente, lições de superação e amor!

À MEUS PAIS (GILBERTO E ELZA - *in memoriam*)

Que tudo fizeram para que eu pudesse trilhar mais esse caminho...

AGRADECIMENTOS

Ao Prof. Dr. Reinaldo Morabito pela parceria (pessoal e profissional) estabelecida ao desenvolver este trabalho. Obrigado pela condução segura e atenciosa desse processo tão significativo em meu percurso profissional.

Aos professores Fernando Yassuo Chiyoshi, Francisco Louzada Neto, Maria Creusa Bretas Salles, Solon Venâncio de Carvalho e Vilma Mayumi Tachibana pelas contribuições neste trabalho de tese.

À empresa que forneceu condições para realização deste trabalho.

À UFSCar e ao Programa de Pós-Graduação em Engenharia de Produção.

A FCT/UNESP de Presidente Prudente pela oportunidade e, em especial, aos colegas do Departamento de Matemática, Estatística e Computação.

Aos colegas Ana Paula Ianoni, João Eduardo Ramos e Randal Farago pela contribuição neste trabalho. Aos colegas de convívio no Programa de Pós-Graduação da Engenharia de Produção.

A toda minha família, em especial meus pais Gilberto e Elza (*in memoriam*) e meus irmãos Glória, Márcio e Cilene que sempre estiveram presentes em minha vida.

À Renata, pelo apoio e amor incondicional....

Finalmente, agradeço a Deus por me permitir trilhar mais esse caminho e se fazer presente a todo o momento em minha vida!

Muito obrigado!

RESUMO

A Importância da Rapidez de Atendimento nos Caixas de Supermercados: um estudo de caso utilizando um modelo analítico de filas com trocas

As empresas de auto-serviço, em geral, têm se preocupado em melhorar seu desempenho em relação à concorrência, seja para ganhar mercado ou mesmo para sobreviver. É essencial que apresentem boa qualidade nos produtos oferecidos e/ou serviços prestados. Assim, neste trabalho, entre outros, um dos objetivos, foi mostrar que o tempo de espera na fila de caixas de supermercados assume grande importância na rapidez de atendimento e, portanto, para o nível de serviço dos clientes. Outros objetivos também foram considerados. As causas pelas quais o cliente frequenta aquele supermercado e, nestas, a relevância de cada uma em contraste com as outras, e sua importância relativa. Assim, pôde-se contextualizar o cenário no qual os clientes enxergam a importância da fila e de seu tempo de espera. Também, associou-se a escolaridade do cliente a fatores pelos quais este frequenta o local, sendo esta uma contribuição interessante devido a existir, no Brasil, uma relação entre renda e escolaridade. Também, estudaram-se modelos de filas que podem representar satisfatoriamente o sistema de filas nos caixas de supermercados e identificou-se, entre eles, o que melhor representou o sistema. Para o modelo de filas com trocas, foi verificado que os clientes trocam de fila quando percebem outra fila com menos clientes, sendo que o tamanho desta diferença foi avaliado. Este modelo foi o que melhor representou a situação comumente encontrada nos supermercados podendo fornecer subsídios para melhoria de atendimento. Para a viabilização da pesquisa foram contatadas algumas empresas da cidade de São Carlos/SP, sendo que, daquelas que permitiram a coleta de dados, uma foi selecionada. Desta forma, este trabalho se trata de um estudo de caso, tendo limitações inferenciais.

Palavras-chave: supermercados, rapidez de atendimento, análise de correspondência, sistemas de filas com trocas.

ABSTRACT

The importance of quick responsiveness in supermarket checkouts: a case study using an analytical queueing model with jockeying

Self-service companies have been concerned about improving their performance in face of competition, either to conquer new markets or to keep in business. It is essential for them to present good quality products and/or services. Thus, in this study, one of the aims was to show that the waiting time in supermarket checkout lines is very important for service speed and, therefore, for the service level of customers. Other aims include the reasons why the customer goes regularly to a specific supermarket and, among these reasons, the relevance of each one in contrast with the others, and its relative importance. As a result, it was possible to contextualize the scenario in which customers view the importance of queues and their waiting time. Also, the level of schooling of the customers was associated with the factors that lead them to go regularly to the place. This is an interesting contribution because in Brazil there is a relation between income and schooling. Besides, the queueing models that may satisfactorily represent the supermarket checkouts were studied and the one that best represents the system was identified. It was verified that the customers change lines when they notice another line with fewer customers, and the size of this difference was assessed. The queueing model with jockeying was the one to best represent the situation commonly found in supermarkets, and it can offer subsidies for service improvement. In order to carry out this research some companies in the city of São Carlos/SP were contacted, and one among the ones that allowed the data collection was selected. Therefore, this is a case study, thus presenting inferential limitations.

Key words: supermarkets, quick responsiveness, correspondence analysis, queueing model with jockeying.

SUMÁRIO

Resumo.....	vi
Abstract.....	vii
Lista de figuras.....	xi
Lista de tabelas.....	xiv
Lista dos principais símbolos e siglas.....	xvii
Capítulo 1. Introdução.....	1
1.1 Objetivos deste estudo.....	6
1.2 Organização do texto.....	8
Capítulo 2. Supermercados.....	9
2.1 Introdução.....	9
2.2 Tecnologias aplicáveis a supermercados.....	11
2.3 Mão de obra e aspectos legais.....	14
2.4 Gerenciamento das percepções nas filas de espera.....	16
2.5 A troca de filas.....	18
Capítulo 3. A realização do experimento e da coleta de dados.....	20
3.1 Introdução.....	20
3.2 A percepção de alguns gerentes.....	20
3.3 Instrumentos e materiais para a coleta de dados.....	21
3.3.1 O questionário elaborado.....	21
3.3.2 Realização da amostra piloto.....	21
3.3.3 O planejamento da coleta de dados.....	22
3.3.4 A planilha para coleta nas filas.....	23
3.3.5 A filmadora utilizada.....	24
3.3.6 A caixa para ocultar a filmagem.....	24
3.3.7 Cronometragem com computador e monitores.....	25
3.3.8 Recursos humanos empregados.....	25
3.4 O experimento.....	26
3.5 Considerações adicionais.....	29
Capítulo 4. A Análise dos questionários utilizando técnicas multivariadas.....	31

4.1	Introdução.....	31
4.2	Limite de espera e importância das características analisadas na frequência ao supermercado.....	32
4.3	O perfil das características pela análise de agrupamentos.....	36
4.3.1	Algumas considerações sobre o método de agrupamentos.....	40
4.3.2	Análise das características do supermercado por meio de agrupamentos hierárquicos de variáveis.....	44
4.4	O número de itens na compra, a percepção de diferença de preços e a escolaridade dos clientes entrevistados.....	53
4.5	O cruzamento das características pela análise de correspondência.....	58
4.5.1	A análise das características utilizando análise de correspondência simples.....	58
4.5.2	Desenvolvimento algébrico.....	60
4.5.3	Resultados das análises de correspondência simples para as características.....	64
4.5.4	A análise das características utilizando análise de correspondência múltipla.....	86
4.6	Considerações adicionais.....	95
Capítulo 5.	Sistemas de filas com trocas em supermercados.....	99
5.1	Introdução.....	99
5.2	As modelagens frequentemente utilizadas.....	103
5.3	O modelo com trocas.....	108
5.3.1	A mudança no instante de chegada.....	109
5.3.2	A mudança no instante de saída.....	111
5.4	As suposições dos modelos.....	116
5.5	Algumas considerações e verificação das suposições.....	119
5.6	Aplicação dos modelos frequentemente utilizados.....	128
5.7	Aplicação do modelo com trocas.....	132
5.8	Comportamento do modelo com trocas em função do número de filas para a primeira semana de coleta.....	142
5.9	A simulação do modelo com trocas.....	145
5.9.1	O método de replicações.....	149
5.9.2	Sistemas terminais e não-terminais.....	151
5.9.3	Simulando o modelo com trocas.....	152

Capítulo 6. Conclusões.....	158
Referências.....	162
Apêndice A.....	167
Apêndice B.....	168
Apêndice C.....	169
Apêndice D.....	170
Apêndice E.....	171

LISTA DE FIGURAS

Figura 1.1. Sistema M/M/m de fila única.....	4
Figura 1.2. Sistemas de filas M/M/1 paralelas.....	4
Figura 2.1. Visão geral dos caixas e empacotadores no supermercado.....	10
Figura 2.2. No supermercado do futuro, na Alemanha, carrinhos identificam o cliente e seu perfil, indicando promoções personalizadas.....	12
Figura 3.1. Visão frontal da caixa para a filmadora.....	25
Figura 3.2. Visão interna da caixa com a filmadora.....	25
Figura 3.3. Representação do supermercado: posição dos caixas, entrevistadores, coletores e monitores.....	28
Figura 4.1 Gráfico do tempo máximo de espera (x_i) pelo número máximo de clientes na fila (y_i).....	33
Figura 4.2. Visão geral dos métodos de agrupamentos.....	41
Figura 4.3. Representação gráfica dos agrupamentos utilizando correlação simples e ligação média para todos os questionários.....	47
Figura 4.4. Representação gráfica dos agrupamentos e similaridades, utilizando correlação simples e ligação média, para todos os questionários.....	48
Figura 4.5. Representação gráfica dos agrupamentos utilizando correlação transformada e ligação média para todos os questionários.....	49
Figura 4.6. Representação gráfica dos agrupamentos utilizando-se correlação simples e ligação média, para a primeira semana de coleta.....	50
Figura 4.7. Representação gráfica dos agrupamentos utilizando-se correlação transformada e ligação média, para a primeira semana de coleta.....	50
Figura 4.8. Representação gráfica dos agrupamentos utilizando-se correlação simples e ligação média, para a segunda semana de coleta.....	51
Figura 4.9. Representação gráfica dos agrupamentos utilizando-se correlação transformada e ligação média, para a segunda semana de coleta.....	51
Figura 4.10. <i>Scree plot</i> dos autovalores para preço, variedade, rapidez, qualidade, localização, estacionamento e outras, utilizando análise fatorial, sem rotação e com matriz de correlação.....	52
Figura 4.11. Representação gráfica do número de itens por cliente.....	54
Figura 4.12. Representação gráfica das classes para o número de itens por cliente.....	54
Figura 4.13. Representação gráfica dos diferentes níveis de escolaridade para os dias de coleta.....	55
Figura 4.14. Representação gráfica para a junção dos diferentes níveis de escolaridade, para os dias de coleta.....	56
Figura 4.15. Renda média de um trabalhador associada ao nível de escolaridade para o Brasil em 2004.....	57

Figura 4.16. Análise de Correspondência Simples considerando-se Preço e Escolaridade.....	68
Figura 4.17. Análise de Correspondência Simples considerando-se Localização e Escolaridade.....	70
Figura 4.18. Análise de Correspondência Simples considerando-se Rapidez e Escolaridade.....	71
Figura 4.19. Análise de Correspondência Simples considerando-se Qualidade e Escolaridade.....	73
Figura 4.20. Análise de Correspondência Simples considerando-se Variedade e Escolaridade.....	74
Figura 4.21. Análise de Correspondência Simples considerando-se Estacionamento e Escolaridade.....	76
Figura 4.22. Análise de Correspondência Simples considerando-se Outras e Escolaridade.....	77
Figura 4.23. Análise de Correspondência Simples considerando-se a primeira opção..	79
Figura 4.24. Análise de Correspondência Simples considerando-se a quarta opção.....	81
Figura 4.25. Gráfico da porcentagem de não respostas para os diferentes níveis de escolaridade.....	82
Figura 4.26. Análise de Correspondência Simples considerando-se Diferença de Preços e Escolaridade.....	83
Figura 4.27. Análise de Correspondência Simples considerando-se as classes para o número de itens e escolaridade.....	85
Figura 4.28. Análise de Correspondência Múltipla considerando-se preço, variedade e escolaridade.....	89
Figura 4.29. Análise de Correspondência Múltipla considerando-se rapidez, qualidade e escolaridade.....	92
Figura 4.30. Análise de Correspondência Múltipla considerando-se classes de itens, diferença de preços e escolaridade.....	94
Figura 4.31. Esquema das análises e associações entre as características apontadas pelos clientes como motivo de frequência ao supermercado.....	97
Figura 5.1. Três realizações de simulação da fila $M/M/1$ considerando o tempo gasto na fila com $\rho=0,9$	100
Figura 5.2. Autocorrelações teóricas do tempo gasto no sistema de fila $M/M/1$ com $\rho=0,9$	101
Figura 5.3. Esquema para possíveis mudanças no instante de chegada.....	110
Figura 5.4. Esquema para possíveis mudanças no instante de saída com um caixa vazio.....	112
Figura 5.5. Esquema para mudança instantânea, no momento de saída do cliente do caixa 1, sem caixa vazio.....	112
Figura 5.6. Diagrama parcial das possíveis mudanças nos estados, com $m=4$, $k=1$ e $N=3$	114
Figura 5.7. Ajuste para os tempos entre chegadas de todos os clientes nos caixas normais no dia 5 de dezembro.....	122

Figura 5.8 Ajuste uniforme para o número de clientes que entram nos caixas normais no dia 5 de dezembro.....	122
Figura 5.9 Ajuste realizado para os tempos de serviço do caixa normal 4, no dia 5 de dezembro, obtendo uma distribuição exponencial.....	123
Figura 5.10 Ajuste realizado para todos os tempos de serviço dos caixas rápidos, no dia 5 de dezembro, obtendo uma distribuição exponencial.....	124
Figura 5.11 Ajuste uniforme para o número de clientes que entram nos caixas rápidos, no dia 5 de dezembro.....	124
Figura 5.12 Ajuste exponencial dos tempos de serviço no caixa rápido 13, no dia 5 de dezembro.....	125
Figura 5.13 Variação da taxa de entrada para os caixas normais nos períodos de coleta.....	126
Figura 5.14 Variação da taxa de entrada para os caixas rápidos nos períodos de coleta.....	126
Figura 5.15 Gráfico mostrando a estabilização do tempo estimado de espera conforme aumento do número de filas mantendo $\lambda_E/m\mu$ fixo (amostra do dia 5/12/04).....	144
Figura 5.16 Gráfico mostrando a estabilização do tempo estimado de espera conforme aumento do número de filas mantendo $\lambda_E/m\mu$ fixo (amostra do dia 4/12/04).....	145
Figura 5.17 Autocorrelações em função dos <i>lags</i> para uma replicação de caixa normal para o dia 12 de dezembro.....	155
Figura 5.18 Autocorrelações em função dos <i>lags</i> para uma replicação de caixa rápido para o dia 11 de dezembro.....	155

LISTA DE TABELAS

Tabela 3.1. Informações gerais sobre a coleta de dados do experimento.....	23
Tabela 4.1. Intervalos de confiança simultâneos com $\alpha = 0,005$	35
Tabela 4.2. Níveis de similaridade e distância.....	47
Tabela 4.3. Variabilidade total explicada pelos 3 fatores.....	53
Tabela 4.4. Cruzamento entre as variáveis Preço e Escolaridade.....	59
Tabela 4.5. Porcentagem para o cruzamento entre Preço e Escolaridade.....	65
Tabela 4.6. Diferença entre as frequências observada e esperada.....	65
Tabela 4.7. Distâncias Qui-Quadrado para Preço e Escolaridade.....	66
Tabela 4.8. Inércias relativas aos componentes.....	67
Tabela 4.9. Contribuições do perfil linha para os componentes.....	67
Tabela 4.10. Contribuições do perfil coluna para os componentes.....	68
Tabela 4.11. Contribuições dos perfis linha e coluna para Localização.....	69
Tabela 4.12. Contribuições dos perfis linha e coluna para Rapidez.....	71
Tabela 4.13. Contribuições dos perfis linha e coluna para Qualidade.....	72
Tabela 4.14. Contribuições dos perfis linha e coluna para Variedade.....	74
Tabela 4.15. Contribuições dos perfis linha e coluna para Estacionamento.....	75
Tabela 4.16. Contribuições dos perfis linha e coluna para Outros.....	75
Tabela 4.17. Médias das características e suas inércias para não respostas.....	78
Tabela 4.18. Contribuições dos perfis linha e coluna para a primeira opção.....	79
Tabela 4.19. Contribuições dos perfis linha e coluna para a quarta opção.....	80
Tabela 4.20. Contribuições dos perfis linha e coluna para Diferença de Preços.....	83
Tabela 4.21. Contribuições dos perfis linha e coluna para Classes de Itens.....	84
Tabela 4.22 Decomposição da inércia para preço, variedade e escolaridade.....	87
Tabela 4.23. Contribuições do perfil coluna para os componentes.....	88
Tabela 4.24. Decomposição da inércia para rapidez, qualidade.....	90
Tabela 4.25. Contribuição do perfil coluna para as componentes.....	91
Tabela 4.26. Decomposição da inércia para classe de itens e diferença de preços.....	93
Tabela 4.27. Contribuição do perfil coluna para os componentes.....	93
Tabela 5.1 Quantidades calculadas para a ANOVA.....	118

Tabela 5.2 Taxas de entradas e serviço, a proporção de servidores ocupados, tamanho amostral e tempo médio de espera na fila para caixas normais e rápidos, nas duas semanas de coleta.....	120
Tabela 5.3 Quantidades calculadas para a ANOVA para as médias de tempos de serviço, no dia 5 de dezembro.....	125
Tabela 5.4 Porcentagens de trocas de filas observadas segundo o tipo de caixa e dias de coleta.....	127
Tabela 5.5 Resultados para as filas $M/M/m$ e $[M/M/1]^m$, para a primeira semana de coleta.....	129
Tabela 5.6 Resultados para as filas $M/M/m$ e $[M/M/1]^m$, para a segunda semana de coleta.....	129
Tabela 5.7 Resultados para as filas $M/M/m/mN$ e $[M/M/1/N]^m$ para a primeira semana de coleta.....	130
Tabela 5.8 Resultados para as filas $M/M/m/mN$ e $[M/M/1/N]^m$ para a segunda semana de coleta.....	131
Tabela 5.9 Resultados para as filas $M/G/m$ e $[M/G/1]^m$ para os dias de coleta.....	132
Tabela 5.10 Medidas para o modelo com trocas, com 8 filas de caixa normal sendo $\lambda_E = 1,6039$, $\mu = 0,2384$, com tempo observado na amostra de 1,80 minutos e $\lambda_E / m\mu = 0,841$, para o dia 3 de dezembro.....	133
Tabela 5.11 Medidas para o modelo com trocas, com 4 filas de caixa rápido sendo $\lambda_E = 1,8196$, $\mu = 0,6250$, com tempo observado na amostra de 1,43 minutos e $\lambda_E / m\mu = 0,728$, para o dia 3 de dezembro.....	134
Tabela 5.12 Medidas para o modelo com trocas, com 8 filas de caixa normal sendo $\lambda_E = 1,3182$, $\mu = 0,1880$, com tempo observado na amostra de 3,36 minutos e $\lambda_E / m\mu = 0,878$, para o dia 4 de dezembro.....	135
Tabela 5.13 Medidas para o modelo com trocas, com 4 filas de caixa rápido sendo $\lambda_E = 2,3785$, $\mu = 0,6693$, com tempo observado na amostra de 1,82 minutos e $\lambda_E / m\mu = 0,888$, para o dia 4 de dezembro.....	135
Tabela 5.14 Medidas para o modelo com trocas, com 8 filas de caixa normal sendo $\lambda_E = 1,6058$, $\mu = 0,2336$, com tempo observado na amostra de 1,85 minutos e $\lambda_E / m\mu = 0,859$, para o dia 5 de dezembro.....	136
Tabela 5.15 Medidas para o modelo com trocas, com 4 filas de caixa rápido sendo $\lambda_E = 2,0803$, $\mu = 0,6228$, com tempo observado na amostra de 1,61 minutos e $\lambda_E / m\mu = 0,835$, para o dia 5 de dezembro.....	137
Tabela 5.16 Medidas para o modelo com trocas, com 8 filas de caixa normal sendo $\lambda_E = 1,4172$, $\mu = 0,2182$, com tempo observado na amostra de 1,39 minutos e $\lambda_E / m\mu = 0,812$, para o dia 10 de dezembro.....	137
Tabela 5.17. Medidas para o modelo com trocas, com 4 filas de caixa rápido sendo $\lambda_E = 1,7311$, $\mu = 0,6130$, com tempo observado na amostra de 1,24 minutos e $\lambda_E / m\mu = 0,706$, para o dia 10 de dezembro.....	138

Tabela 5.18 Medidas para o modelo com trocas, com 8 filas de caixa normal sendo $\lambda_E = 1,3429$, $\mu = 0,1901$, com tempo observado na amostra de 3,63 minutos e $\lambda_E / m\mu = 0,883$, para o dia 11 de dezembro.....	138
Tabela 5.19 Medidas para o modelo com trocas, com 4 filas de caixa rápido sendo $\lambda_E = 2,4239$, $\mu = 0,6718$, com tempo observado na amostra de 2,19 minutos e $\lambda_E / m\mu = 0,902$, para o dia 11 de dezembro.....	139
Tabela 5.20 Medidas para o modelo com trocas, com 8 filas de caixa normal sendo $\lambda_E = 1,6099$, $\mu = 0,2254$, com tempo observado na amostra de 2,92 minutos e $\lambda_E / m\mu = 0,893$, para o dia 12 de dezembro.....	139
Tabela 5.21 Medidas para o modelo com trocas, com 4 filas de caixa rápido sendo $\lambda_E = 2,2399$, $\mu = 0,6272$, com tempo observado na amostra de 2,64 minutos e $\lambda_E / m\mu = 0,893$, para o dia 12 de dezembro.....	139
Tabela 5.22 Comparação entre os modelos $M/M/m$, $[M/M/1]^m$ e o melhor resultado obtido pelo modelo com trocas, para caixas normais e rápidos, na primeira semana de coleta.....	140
Tabela 5.23 Comparação entre os modelos $M/M/m$, $[M/M/1]^m$ e o melhor resultado obtido pelo modelo com trocas, para caixas normais e rápidos, na segunda semana de coleta.....	141
Tabela 5.24 Tempo estimado (W_q) para o modelo com trocas, com m filas de caixa normal, sendo λ_E variável, $\mu = 0,2336$, com tempo observado na amostra de 1,85 minutos e fixando-se $\lambda_E / m\mu = 0,859$, para o dia 5 de dezembro.....	143
Tabela 5.25 Tempo estimado (W_q) para o modelo com trocas, com m filas de caixa normal, sendo λ_E variável, $\mu = 0,1880$, com tempo observado na amostra de 3,36 minutos e fixando-se $\lambda_E / m\mu = 0,878$, para o dia 4 de dezembro.....	144
Tabela 5.26 Resultados da simulação, do modelo com trocas e das amostras para o tempo médio gasto na fila nos caixas normais.....	156
Tabela 5.27 Resultados da simulação, do modelo com trocas e das amostras para o tempo médio gasto na fila nos caixas rápidos.....	156

LISTA DOS PRINCIPAIS SÍMBOLOS E SIGLAS

ANOVA	análise de variância
$d(\mathbf{x}, \mathbf{y})$	distância entre dois itens
D	matriz de distâncias
E	característica estacionamento
E_i	categoria i de estacionamento
K	diferença de clientes para troca de filas
k	número de replicações (ou <i>batches</i>)
lag_j	atraso ou distância de j unidades entre duas observações
L	característica localização
L_i	categoria i de localização
L	número médio de usuários no sistema
L_i	número médio de usuários no caixa i
L_q	número médio de usuários na fila
$\Lambda_{(J-1) \times (J-1)}$	matriz diagonal de autovalores
λ	taxa de chegadas
λ_i	taxa de chegadas do caixa i
λ_E	taxa de entradas
$\lambda_E / m\mu$	proporção de servidores ocupados
m	número de filas no sistema
m	tamanho das replicações (ou <i>batches</i>)
N	número máximo de clientes na fila
n_i	número de clientes no caixa i
n	tamanho da rodada de simulação
(n_1, n_2, \dots, n_m)	espaço de estados do sistema de filas
O	característica outras
O_i	categoria i de outras
P	característica preço

P_i	categoria i de preço
p_N	probabilidade de perda para o sistema $[M / M / 1 / N]^m$
p_{mN}	probabilidade de perda para o sistema $M/M/m/mN$
$\mathbf{P}_{(I \times J)}$	matriz de correspondência
$p(0)$	probabilidade do sistema estar vazio
$p(n_1, n_2, \dots, n_m)$	probabilidade do sistema estar no estado (n_1, n_2, \dots, n_m)
Q	característica qualidade de atendimento
Q_i	categoria i de qualidade de atendimento
R	característica rapidez de atendimento
R_i	categoria i de rapidez de atendimento
R^2	coeficiente de explicação
S	matriz amostral de covariâncias
$\hat{\sigma}^2(\bar{X}(n'))$	estimador não viciado para a variância de $\bar{X}(n')$
Σ	matriz de covariâncias
\mathbf{R}	matriz de correlação amostral
ρ	matriz de correlação
ρ	fator de utilização
ρ_i	fator de utilização do caixa i
ρ_j^*	autocorrelação de lag j
ρ_{ij}	correlação entre duas variáveis aleatórias i e j
γ	medida a ser estimada por simulação
μ	taxa de serviços
v_i	autovalor i
$\mathbf{X}_{(I \times J)}$	matriz de frequências
$\bar{X}_i(m)$	média do <i>batch</i> i
$\bar{X}(n')$	média geral na simulação
χ^2	medida de associação Qui-quadrado

χ_{ij}^2	inércias das células
χ^2 / n	inércia total
W	tempo médio de espera no sistema
W_q	tempo médio de espera na fila
W_{q_i}	tempo médio de espera na fila i

Capítulo 1 - Introdução

As empresas de auto-serviço, em geral, têm se preocupado em melhorar seu desempenho em relação à concorrência, seja para ganhar mercado ou mesmo para sobreviver. Com este objetivo é essencial que estas apresentem boa qualidade nos produtos oferecidos ou serviços prestados. Para tanto, deve-se conduzir melhorias na formação de competências de seus funcionários e da organização, com objetivo de manter a competitividade da empresa.

Quando os clientes realizam uma compra, não estão simplesmente adquirindo um produto. Estão comprando um conjunto de benefícios que, esperam eles, atendam suas necessidades. Desta forma, é importante a percepção que os clientes têm sobre os benefícios do que foi comprado. Igualmente, é importante para a empresa saber o que seus clientes estão comprando dela. Esses benefícios, além de qualidades de bens físicos e serviços, podem incluir, por exemplo, um ambiente agradável ou uma atmosfera relaxante no local onde se realiza a compra. Em algumas situações esses benefícios podem garantir aos clientes da empresa uma agregação de valor. Este raciocínio se aplica ainda mais fortemente ao setor de serviços.

A competição entre as empresas pode causar uma procura atualizada por estratégias eficientes para enfrentar mudanças resultantes da busca por consolidação e ampliação do mercado. Para algumas empresas, como é caso dos supermercados, é de muita importância desenvolver estratégias orientadas para serviços.

Os supermercados são empresas varejistas, ou seja, representam o último elo do canal entre um produto e o consumidor, vendem proeminentemente alimentos perecíveis dispostos em formato para auto-atendimento (*self-service*) e com caixas (*checkouts*) na saída, tratando-se, portanto, de auto-serviço.

Estas empresas têm se preocupado em melhorar o nível de serviço ao consumidor para fidelizá-lo e aumentar o volume de vendas. Pode-se observar uma forte tendência de grandes redes de supermercados em fazer uso de tecnologias mais avançadas que permitam utilizar menor número de empregados, maior número de caixas para atendimento, maior flexibilidade de horário, maior variedade de produtos e disposição organizada dos mesmos.

De modo geral, o consumidor pode estar fidelizado ao supermercado onde compra. A frequência ao estabelecimento faz com que o usuário se acostume ao local, conhecendo melhor o espaço físico, a variedade encontrada e a disposição dos produtos, além de ter a percepção dos horários menos congestionados, facilitando suas compras.

Os clientes podem escolher o local de suas compras por fatores como rápido atendimento, funcionários educados e prestativos em atender, a localização, preço e estacionamento disponível, entre outros. Contudo, parece que preços altos e filas exageradas são fatores que podem levar o estabelecimento a uma diminuição da frequência dos clientes e, conseqüentemente, a uma queda no volume de vendas.

O cliente que tem como primeira opção o preço dos produtos geralmente faz pesquisa antes de realizar uma compra, o que pode ser muito facilitado pela distribuição de encartes de propaganda fornecidos pelos supermercados. Este cliente, pelo tipo de conduta de compra, é difícil de ser fidelizado. Então, a empresa que apostar como estratégia principal no fator preço pode encontrar dificuldades para fidelizar seus clientes e, assim, não alcançar os resultados desejados.

Consumidores podem relacionar o alto tempo total gasto no supermercado para realizar compras com falta de organização geral deste. No tempo total de permanência do cliente no supermercado, um fator determinante pode ser o tempo de espera na fila do caixa. Enquanto circula pelo local escolhendo produtos a serem adquiridos, ele está executando algo de seu interesse e cujo tempo pode ser razoavelmente controlado. Contudo, permanecer na fila do caixa é considerado pelo cliente um tempo praticamente perdido, onde ele não executa qualquer tarefa que contribua com sua compra.

Uma área de estudo que procura minimizar os efeitos da espera na fila é o gerenciamento das percepções nas filas de espera (COGAN, 1998). Em vários locais, técnicas aplicadas a esta área têm obtido resultados satisfatórios. Muito se tem pesquisado para melhoria de atendimento nos supermercados, contudo, os trabalhos com atenção essencial para filas são poucos.

Alguns estudos procuram relacionar a capacidade de serviço (ou atendimento) com o aumento da demanda (ITTIG, 2002). Assim, pode-se obter um *tradeoff* (uma compensação) entre capacidade de atendimento e volume de vendas, auxiliando na tomada de decisões.

Outros estudos levam em consideração até a justiça na prioridade de atendimento e psicologia das filas. Algumas filas podem ser injustas, ou seja, não necessariamente os clientes que chegam primeiro serão atendidos antes. Dependendo das circunstâncias em que ocorre, isto pode enfurecê-los com sérias conseqüências para a empresa (LARSON, 1987). Desta forma, quando se utiliza uma única fila de espera, a tendência é de que o cliente possa ficar mais satisfeito com o serviço prestado. Os supermercados, por problemas relacionados a espaço físico, se utilizam freqüentemente de filas paralelas nos caixas. Algumas vezes uma fila única é utilizada para caixas rápidos, por exemplo.

Entender a motivação e comportamento do consumidor, frente a novas tecnologias em auto-atendimento também tem sido pesquisados (DABHOLKAR, BOBBITT e LEE, 2003). Fatores situacionais e características do consumidor foram estudados e concluiu-se que uma variável relevante para qualquer serviço é a espera (DABHOLKAR e BAGOZZI, 2002). Até mesmo esperar para baixar informação (*download*) pela Internet já preocupa gerentes em relação ao comportamento do usuário. Também, falhas e sucessos de compra pela Internet têm sido investigados para entender e prever tecnologia em auto-atendimento (BOBBITT e DABHOLKAR, 2001).

Percebe-se que a espera, de um modo geral, sensibiliza o cliente em seu futuro comportamento. Um consumidor insatisfeito com seu tempo de espera talvez possa não desistir de sua compra naquele momento, contudo, pode não voltar mais àquele estabelecimento devido a imagem negativa percebida. Esta é uma perda difícil de ser quantificada e, em longo prazo permanecendo o problema, o estabelecimento terá queda de demanda possivelmente sem descobrir a razão disto.

As filas dos caixas em um supermercado, sendo responsáveis por grande parte da espera do cliente, devem ser objeto de estudo quanto ao nível de atendimento desejado em todo sistema prestador de serviço que a possua. Representá-la, modelando-a adequadamente, é fase importante para propor melhorias de desempenho.

Devido à existência de uma variedade de filas, houve a necessidade de desenvolver uma notação que as identificasse. Uma fila, de forma geral, possui seis características, denotadas como $A/B/m/C/N/K$ (GROSS e HARRIS, 1974), nas quais A denota a distribuição do tempo entre chegadas sucessivas, B denota a distribuição do tempo

de serviço, m denota o número de servidores em paralelo, C denota a disciplina da fila, N denota o número máximo de usuários no sistema (entre fila e atendimento) e K denota o tamanho da população, referente a usuários disponíveis. As notações A e B , por exemplo, podem indicar distribuição exponencial (denotada por M , de Markov, sem memória), ou Erlang com parâmetro n (denotada por E_n), entre outras. A disciplina pode ser FCFS (quem chega primeiro é atendido primeiro, situação usual), LCFS (quem chega por último é atendido primeiro), entre outras. Também, K e N podem ser considerados variando de 1 a ∞ , conforme as características do modelo a ser denotado. Quando a notação C está ausente, isto significa FCFS. Também, quando há a ausência de K e/ou N , pode ser considerado como ∞ .

Filas paralelas têm sido freqüentemente objetos de estudo, como, por exemplo, o sistema $M/M/m$ com fila única (MITRANI, 1998), (vide figura 1.1). Estas podem ser comparadas a m filas $M/M/1$ paralelas (também denotadas por $[M/M/1]^m$), situação encontrada em supermercados (vide figura 1.2), quando é razoável considerar as distribuições de chegada e serviço dos usuários como exponenciais negativas.

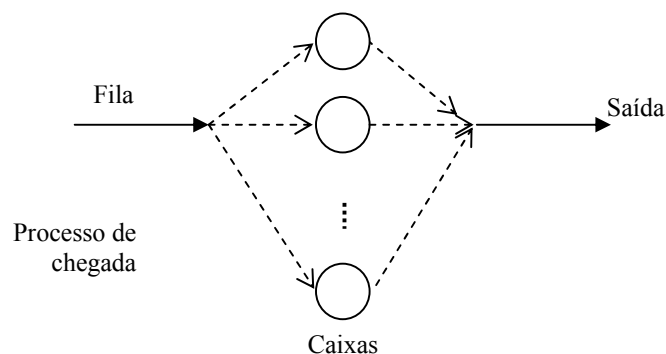


Figura 1.1. Sistema $M/M/m$ de fila única.

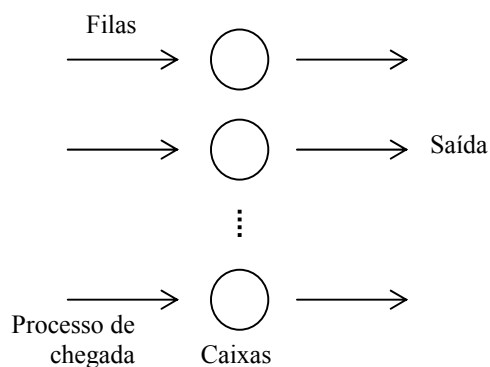


Figura 1.2. Sistemas de filas $M/M/1$ paralelas.

Em algumas ocasiões, as distribuições de chegada e/ou de serviço podem não ter distribuição exponencial e, então, os modelos acima não poderiam adequadamente ser aplicados. Quando isto ocorre, utilizam-se algumas aproximações referentes ao sistema com m filas $G/G/1$ (análogo à figura 1.2) ou a uma fila $G/G/m$ (análogo à figura 1.1), onde G denota uma distribuição genérica. Ainda, duas outras situações podem ser consideradas: o modelo de fila $M/M/m/FCFS/N$ (ou simplesmente $M/M/m/N$) e o modelo de fila $[M/M/1/FCFS/N]^m$ (ou simplesmente $[M/M/1/N]^m$). Estes modelos consideram a impossibilidade física de se alocar muitas pessoas na fila do supermercado. É natural então que, quando a fila está muito grande, as pessoas desistam das compras, tornando os modelos com capacidade limitada.

Uma outra situação é considerar as m filas $M/M/1/N$ paralelas, podendo haver a troca de fila (*jockeying*) por parte dos clientes, toda vez que houver uma outra fila com menor número de usuários do que os que estão na fila à sua frente (MORABITO e LIMA, 2004).

O espaço físico é fator importante quando da escolha de um sistema de filas. Como exemplo, se um supermercado opera com muitos caixas, pode haver grande limitação em considerar uma única fila. Há também estudos que foram realizados com filas paralelas e com *jockeying* probabilístico (FERRARI, 2002).

A aplicação e avaliação de modelos com fila única, e com filas em paralelo, com ou sem trocas, em uma situação real, se proporciona como interessante problema de pesquisa. Algumas empresas (supermercados) da cidade de São Carlos/SP foram contatadas quanto ao interesse em permitir a coleta de dados e realização de estudos. Apesar de algumas recepções positivas, verificou-se um comportamento de resistência por parte de alguns gerentes. Estes apostam, preferencialmente, em preços como uma causa extremamente importante para atração de clientes, alguns deles alegando que a fila não constitui um fator tão importante para a satisfação dos mesmos. Admitindo-se que as filas têm importância, duas hipóteses para este comportamento de resistência podem ser levantadas: a primeira seria a falta de conhecimento da real importância das filas para os clientes e de suas implicações; e a segunda, poderia constituir-se em certa insegurança para uma possível aplicação dos resultados obtidos no local analisado.

Para melhorar o sistema de filas é preciso estudá-lo, saber sua real importância, contextualizá-lo na empresa onde este se apresenta, principalmente considerando as opiniões dos clientes. A realização de troca de filas é um procedimento do cliente que procura minimizar seu tempo de permanência na fila. Questões importantes de pesquisa são obtidas, tais como que informações o cliente avaliou quando fez a troca, e assim qual seria a melhor maneira para representar o sistema de fila do supermercado estudado englobando este novo conhecimento.

1.1 Objetivos deste estudo

O objetivo geral deste estudo é mostrar que o tempo de espera na fila de caixas de supermercados desempenha um importante papel para o nível de serviço dos clientes. De forma mais específica, estudaram-se as causas pelas quais o cliente frequenta um supermercado e, nestas, a relevância de cada uma em contraste com as outras, e sua importância relativa. Assim, pôde-se contextualizar o cenário no qual os clientes enxergam a importância da fila e de seu tempo de espera. Associar a escolaridade do cliente aos fatores pelos quais este frequenta o local foi uma contribuição interessante para supermercados, pois, no Brasil, existe uma forte relação entre renda e escolaridade. Conhecer sua clientela é desejo de qualquer empresa que quer fidelizá-la. Desta forma, esta questão também foi pesquisada.

Outro objetivo é estudar, propor e analisar modelos analíticos de filas eficazes para mais adequadamente estimar o tempo de espera na fila em supermercados. Em particular, para o modelo de filas com trocas, foram avaliadas as causas que determinaram as trocas realizadas pelos clientes. Uma questão interessante que foi observada se refere a se o cliente avaliou somente o número de pessoas na fila para efetuar a troca. Isto, de fato, ocorreu, pois pouquíssimos clientes revelaram ter observado a quantidade de trabalho a ser processada a sua frente.

A sensibilidade do usuário em relação ao tempo máximo esperado na fila e no tamanho desta também foi investigada neste trabalho. Relacionar tempo e tamanho de fila não é tão simples, pois neste caso, a percepção do tempo esperado pelo cliente depende de vários fatores que muitas vezes não pode ser controlado. Por meio do coeficiente de

determinação aproximado notou-se que os clientes não conseguem relacionar de forma coerente tempo e tamanho de fila.

Para a coleta dos dados nas filas dos supermercados, foram utilizadas planilhas contendo informações como: tempos entre chegadas nas filas, tempos de serviço nos caixas, instantes das trocas de fila e as condições em que foram realizadas.

Com o objetivo de coletar informações relevantes sobre a sensibilidade do usuário em relação ao tempo na fila e as causas pelas quais ele frequenta o supermercado, foi planejado e construído um questionário (*survey*) (REA e PARKER, 2000). Este foi aplicado utilizando-se amostragem sistemática na saída dos clientes após o atendimento no caixa. Para realizar a análise dos dados fornecidos pelos questionários, foram necessárias técnicas estatísticas como análise de regressão, análise de agrupamento hierárquico e análise de correspondências (JOHNSON e WICHERN, 2002), entre outras. Os instrumentos principais de pesquisa, então utilizados, foram constituídos de um questionário, do *software* estatístico Minitab e da planilha para coleta dos tempos nos caixas, entre outros.

Para comparar os modelos analíticos de filas foram utilizadas rotinas que fornecem algumas medidas de desempenho (WOLFF, 1989; KLEINROCK, 1975), que permitem analisá-los e, assim, originando subsídios que apontaram aquele que melhor representou a fila real investigada.

Ao se realizar uma pesquisa, é importante que o método de procedimento assumido seja o mais adequado possível à situação do problema. Nas pesquisas organizacionais que se realizam atualmente, os métodos utilizados têm grande influência em pesquisas na área de Ciências Sociais. Contudo, nas organizações existem características próprias que conduzem a utilização mais freqüente de alguns métodos, sendo que o estudo de caso é o método de procedimento mais utilizado pela gestão da produção. Trata-se da análise minuciosa de alguns casos que freqüentemente são encontrados em organizações, focalizando acontecimentos contemporâneos e tentando facilitar decisões (YIN, 2001). Desta forma, no presente trabalho, o estudo de caso foi aplicado, existindo até mesmo a preocupação com a percepção das pessoas, o que é típico deste método de procedimento.

1.2 Organização do texto

O presente texto é organizado em seis capítulos e estruturado da forma a seguir.

O segundo capítulo trata de referenciais sobre supermercados, a percepção dos clientes em relação aos serviços oferecidos pelos supermercados e os aspectos legais sobre filas atualmente. Procura ainda discutir sobre gerenciamento das percepções nas filas de espera e a troca de filas nos supermercados.

O terceiro capítulo procura discutir a percepção de alguns gerentes sobre a importância das filas para supermercados, motivo pelo qual ampliou-se a pesquisa inicialmente proposta. Trata também, dos instrumentos e materiais utilizados na coleta de dados, bem como de seu planejamento e da amostra piloto. Refere-se, ainda, à forma com que os recursos humanos empregados foram necessários e como o experimento transcorreu no supermercado. Finalmente, fornece uma impressão geral sobre as contribuições que o contato direto com os clientes trouxe ao trabalho.

O quarto capítulo trata da análise do questionário. Empregaram-se técnicas multivariadas para posicionar as categorias pelas quais os clientes freqüentam o supermercado. Também, procura associá-las à escolaridade, ao tamanho da compra e à percepção da existência de diferença de preços. Estas associações fornecem subsídios a decisões gerenciais, visto que, renda está associada à escolaridade, formando um cenário sobre o comportamento dos clientes.

O quinto capítulo trata dos modelos analíticos de filas que podem ser aplicados aos supermercados. Alguns desses modelos foram comparados com a intenção de verificar aquele que melhor representou a situação real encontrada. Também, considerando o modelo com trocas, foram investigados os motivos que fazem com que os clientes troquem de fila, podendo contribuir com a melhoria do modelo.

Por último, são apresentadas as conclusões deste trabalho bem como as perspectivas para pesquisas futuras.

Capítulo 2 – Supermercados

2.1 Introdução

Os supermercados e hipermercados podem ser classificados por suas características ou mesmo pelo número de caixas (*checkouts*) existentes nas lojas (BRASIL, 1996). Os hipermercados possuem grandes áreas com diversificada gama de produtos, incluindo alimentos, bebidas, roupas, artigos para residências, artigos para carros, etc. Geralmente possuem mais de cinco mil metros quadrados. Estes podem ser classificados também desta forma por possuírem mais de cinquenta *checkouts*.

Os supermercados constituem-se, preferencialmente, de itens alimentares. Em suas classificações, por tamanho, constam lojas de cinco a quarenta e nove *checkouts*, ou seja, possuem grande flexibilidade de tamanho. Ainda, os estabelecimentos que possuem até quatro *checkouts* são denominados minimercados.

A empresa onde foram coletados os dados possui um total de dezesseis *checkouts* condizendo, segundo a definição, como supermercado, tanto por tamanho quanto por características. Como boa parte das empresas brasileiras, esta, também, caracteriza-se pelo controle familiar. Sendo fundada há algumas décadas, hoje se encontra sendo administrada por descendentes, atuando na cidade de São Carlos/SP há mais de quarenta anos. A empresa tem seis lojas nesta cidade e também em quinze cidades da região, com um total de vinte e oito lojas. Embora tenha controle familiar, pode-se observar uma tendência de elevar a profissionalização das administrações e de *marketing* demonstrando um bom conhecimento de gerenciamento da empresa.

Neste setor tem existido um forte enfoque à terceirização da mão de obra, principalmente de operadores de caixa. Contudo, o supermercado em questão não tem operado assim, havendo inclusive funcionários trabalhando há mais de dez anos na empresa. Um fato interessante deste setor é que geralmente os únicos contatos diretos da empresa com os clientes são obtidos por meio dos operadores de caixa. Desta forma, os supermercados que se utilizam de terceirização de mão de obra com grande rotatividade têm estes contatos perdidos, diminuindo a importância dos funcionários nas estratégias da organização, e não se utilizando de informações que poderiam ser valiosas para a empresa.

Deve-se ressaltar que, no supermercado onde a pesquisa foi realizada, alguns operadores de caixa conhecem vários clientes pelo nome, possibilitando um estreitamento do relacionamento e fornecendo um tratamento diferenciado, o que pode ocasionar inclusive em fidelização, característica muito desejável neste setor.

Outro diferencial observado na empresa é a presença de empacotadores (vide figura 2.1). Em vários supermercados têm se observado que o cliente e/ou o caixa ficam responsáveis por este serviço, podendo haver duas implicações negativas: desprestígio do cliente e seguramente um atraso no tempo de serviço. No supermercado em questão isto não ocorre, pois enquanto o cliente efetua o pagamento tem as compras devidamente embaladas, agilizando a fila, e obtendo uma atenção especial. Ainda, do contato dos operadores de caixa com os clientes, podem surgir várias informações sobre a percepção que estes têm em relação ao supermercado, quanto à qualidade e rapidez de atendimento e, mesmo, a falta de alguns produtos. Quando há rotatividade de mão de obra, muito desta informação é perdida.



Figura 2.1. Visão geral dos caixas e empacotadores no supermercado.

Tais características vêm sendo, também, ressaltadas por alguns autores (QUEIROZ, GOLVINHAS e SOUZA, 2004), enfatizando a importância dos funcionários para a empresa como um serviço diferenciado. Enfim, conhecer os seus clientes pode ser fundamental para uma empresa conhecer o seu negócio.

2.2 Tecnologias aplicáveis a supermercados

Para os supermercados da década de 90, automação poderia ser praticamente sinônimo de possuir *scanner*. Hoje estudam-se vários tipos de tecnologias aplicáveis ao setor, sendo que algumas já estão disponíveis.

Em meados da década passada a consultora *Price Waterhouse* divulgou que em 2005 os supermercados não teriam mais operadores de caixas e carrinhos, os depósitos seriam operados por robôs que entregariam as compras já embaladas. Seriam reduzidos drasticamente o pessoal empregado e o tempo de atendimento, os *checkouts* seriam acionados por cartões inteligentes e até mesmo as gôndolas só teriam rótulos e não as mercadorias (BADIN, 1997).

Como pode-se perceber estes fatos não ocorreram. As novas tecnologias ainda enfrentam problemas para tais pretensões e também há um outro fator, o preço para utilizá-las. Contudo, a automação e o uso de tecnologia têm permitido algumas inovações em supermercados.

Por exemplo, nos supermercados Winn Dixie (EUA) os caixas são inexistentes. O cliente fica responsável por passar seus itens selecionados pela leitora ótica, passar seu cartão de crédito e assinar a leitora ótica, indo embora imediatamente. Ainda, recebe descontos por ter utilizado este tipo de serviço. Observa-se que, por não possuir operadores, este tipo de método tem custo menor e pode permitir que se disponha de um grande número desses dispositivos, denominados caixas automáticos (*automatic checkouts*) (GUERRA, 2006).

Algumas lanchonetes da rede McDonald's (EUA) estão disponibilizando telefones instalados nas mesas, permitindo realizar os pedidos sem a presença de um funcionário. Junto com os telefones estão acopladas leitoras de cartões de crédito, possibilitando também o pagamento de forma simples e rápida. Toda esta tecnologia

aplicada reduz custos, eleva a produtividade, dá mais liberdade e autonomia ao cliente e é rápida, desde que confiável, levando maior satisfação ao usuário. Entretanto, um sistema de reserva é sempre necessário para garantir o bom desenvolvimento do serviço.

Alguns supermercados brasileiros têm oferecido a seus clientes o serviço de compra por Internet. Essa estratégia ainda representa uma fatia muito pequena do mercado, pois o consumidor tem reservas a este tipo de compra. As razões são que muitos não compram o produto sem examiná-lo, há dificuldade de atender toda a compra, a segurança na Internet, equipamentos incompatíveis, entre outros. Contudo, se a adoção desse serviço aumentar muito, os supermercados poderão ter informações imediatas sobre estoques, vendas, etc. Também, poderão identificar o perfil do seu cliente, melhorando o serviço prestado.

Um exemplo de toda esta tecnologia é apresentado pela Metro, terceiro maior varejista da Europa (ESTADÃO, 2005). Uma de suas lojas de supermercados, em Rheinberg (Alemanha), é considerada a mais moderna do mundo, exibindo até mesmo tecnologias experimentais. Os carrinhos (vide figura 2.2) são dotados de visores, que por meio do cartão do cliente, o identifica, informando até mesmo promoções focadas ao seu perfil. As prateleiras têm leitoras de etiquetas eletrônicas e identificam a retirada de um produto, informando o setor de estoque. As etiquetas dos produtos permitem informar até a data de validade e, se foram alocadas em local errado, a prateleira avisará. Ainda, os carrinhos são capacitados para passar todos os produtos de uma só vez e o cliente só pára para efetuar o pagamento por cartão de débito automático ou crédito. Finalmente, na saída da loja, os clientes desativam as etiquetas eletrônicas, evitando rastreamento fora do supermercado.



Figura 2.2. No supermercado do futuro, na Alemanha, carrinhos identificam o cliente e seu perfil, indicando promoções personalizadas.

As limitações do emprego deste tipo de tecnologia são o custo e algumas barreiras físicas. O investimento ainda é muito alto, pois cada etiqueta eletrônica custa entre US\$ 0,05 a US\$ 0,40, o que para muitos produtos poderia representar um acréscimo considerável de preço e, como consequência, haver uma diminuição significativa de seu consumo, inviabilizando a adoção da etiqueta. Também as leitoras (ou sensores) custam entre US\$ 1000,00 a US\$ 4000,00, o que é um valor bastante alto considerando a quantidade delas que teria que se colocar em um supermercado.

Há também alguns problemas físicos que não foram resolvidos. A água absorve ondas de rádio e o metal, ao contrário, as reflete. Desta forma, fica difícil colocar a etiqueta em uma lata de refrigerante, por exemplo. Ainda, há problemas típicos de sistemas sem fio, como falha de cobertura. As etiquetas precisam estar ao alcance dos sensores e, nos supermercados com grande número de gôndolas, precisariam haver muitos deles.

Ainda que com as limitações citadas, nos EUA, a rede Walt-Mart tem exigido de seus maiores fornecedores que adotem a tecnologia, o que tem colocado em alerta constante os diversos fabricantes de etiquetas eletrônicas. A Gillette tem pago US\$ 0,30 por etiqueta, utilizando-as em pacotes de lâminas de barbear devido ao grande volume de mercadorias extraviado. Diferentes utilizações têm sido encontradas para as etiquetas eletrônicas, como registrar a chegada de alunos em uma escola, identificação de animais, anti-sequestro (sob a pele), resultados imediatos para atletas, pedágios sem-parar em rodovias, etc. Essa tecnologia veio para permanecer; os problemas devem ser solucionados e o custo deve diminuir com a popularização dessa tecnologia.

Deve-se considerar, contudo, que a Walt-Mart está adotando, desde o ano de 2006, previsões a cada quinze minutos de quantos caixas são necessários a cada momento, baseadas na série histórica de vendas de anos anteriores (ABASE, 2006). Nota-se, portanto, que apesar de apostar nas etiquetas eletrônicas, ainda investe em tecnologia baseada em caixas com operadores.

Uma questão a ser abordada é sobre os problemas que as pessoas enfrentam para utilizar serviços que não são interpessoais (MEUTER et al, 2000), como transações bancárias por Internet ou instruções em ensino a distância. Isto pode ser um problema enfrentado pelos supermercados quando das mudanças futuras, especialmente no Brasil

onde existe uma forte cultura de inter-relações humanas e, no qual, algumas pessoas ainda têm dificuldades com o código de barras.

Os custos desta tecnologia também devem representar um problema maior para o Brasil, pois o acréscimo de preço nas mercadorias seria mais difícil de assimilar, devido ao baixo poder aquisitivo da população. Anexe a isto os baixos salários pagos por este setor no país, deixando ainda mais lenta a velocidade de emprego desta tecnologia. A tendência é que esta mudança vá ocorrendo mais devagar e que o código de barras ainda conviva muito tempo com a etiqueta eletrônica.

Uma última questão é relativa à preocupação com a privacidade dos clientes. Nos EUA há o temor de que as etiquetas possam ser utilizadas para rastrear pessoas fora do supermercado. Também, por meio de cartões, que houvesse o fim do consumo com privacidade, ou seja, os supermercados teriam o perfil do cliente não só em relação ao consumo de mercadorias, mas que, por meio delas, tivesse um perfil mais geral (de hábitos, costumes, comportamentos, etc) de seus clientes. Esta discussão parece estar apenas no início.

2.3 Mão de obra e aspectos legais

Para o setor de serviços o relacionamento humano tem sido fundamental. Para supermercados isto não é diferente, uma vez que há contato direto com os clientes e a avaliação destes em relação ao serviço prestado é muita subjetiva. Neste setor, os funcionários que atendam às expectativas da empresa são bens que ela deve saber administrar, oferecendo treinamento, oportunidades de especialização, etc.

No Brasil, apesar dos salários ofertados pelos supermercados serem considerados baixos, existe muita procura por empregos toda vez que um deles veicula contratação. Geralmente são pessoas carentes de especialização e treinamento, dificultando a situação tanto para o pretendente à vaga como para a empresa. Ainda que esta situação não seja favorável, a fidelização de clientes em supermercados passa pela relação entre os próprios clientes e os funcionários responsáveis pela qualidade e rapidez de atendimento, por exemplo.

Como um fenômeno mundial, os trabalhadores têm experimentado pressão para aumento da jornada de trabalho, situação agravada pelo desemprego. Assim, esperas, como na fila de um supermercado, têm se tornado menos suportável e propiciado aumento crescente de reclamações. Alguns setores da sociedade brasileira têm percebido a insatisfação com as esperas e propõe como solução contratação de pessoal para agilização de serviços, apoiando-se inclusive na questão da falta de empregos.

Em junho de 2005 foi aprovado, pela Câmara Municipal da cidade de São Paulo, um projeto de lei no qual o tempo máximo de espera em filas de supermercados não poderia ultrapassar vinte minutos, inclusive para caixas rápidos. A fiscalização se daria por bilhetes ou senhas com o horário do início de atendimento, e previa multas progressivas para reincidência. Em dezembro de 2005, o então prefeito da cidade de São Paulo vetou na íntegra o texto aprovado, alegando principalmente que o tempo estipulado era desprovido de elementos essenciais em sua eficácia (DOM/SP, 2005, p.73). Contudo, o vereador proponente do projeto prometeu reformulá-lo e reenviá-lo.

Este projeto foi realizado ancorado na aprovação da lei que limita, na cidade de São Paulo, em quinze minutos o tempo máximo de espera em filas de banco, sancionada pelo prefeito em 30 de setembro de 2005. Nas razões de veto da limitação para os supermercados está que a matéria não guarda semelhança com a situação das agências bancárias.

Quase que ao mesmo tempo do projeto anterior, outro vereador da cidade de São Paulo traz uma outra proposta de lei que obriga os supermercados a contratar um empacotador para cada um dos caixas. Alega que a rapidez aumentaria a satisfação dos clientes e ao mesmo tempo criaria novos empregos. Salienta, ainda, que se poderia realizar um vínculo com o programa primeiro emprego do governo federal, o qual ficaria responsável por metade do salário.

A Associação Paulista de Supermercados (APAS) considera que ambos os projetos estariam indo contra a prática da livre iniciativa e que o custo seria repassado para o consumidor. Em alguns outros estados fatos semelhantes ocorreram, como em Santa Catarina (JORNAL DA CÂMARA, 2006) e Bahia, parecendo que os supermercados atraíram a atenção para o problema de espera em filas. Esta situação poderá levar ao caminho contrário à automação, contratando mais mão-de-obra não especializada e não

investindo em tecnologia ou, ao contrário, acelerando o processo de automação para reduzir a espera na fila.

Independentemente de leis, os supermercados que pretendam permanecer no mercado devem investir em capacitação e treinamento de pessoal (qualidade de atendimento), bem como minimizar o tempo de espera (rapidez de atendimento).

2.4 Gerenciamento das percepções nas filas de espera

O gerenciamento das percepções nas filas de espera basicamente trata de procedimentos possíveis para melhorar a satisfação dos clientes em relação aos serviços por eles recebidos. Existe uma diferenciação entre tempo efetivo de espera e tempo percebido de espera. O primeiro é aquele aferido por relógio e o segundo é o tempo que o cliente acha (percebe) que esperou. Se o cliente teve uma espera agradável, ele tende a subestimar este tempo, caso contrário, a tendência é de haver superestimação.

Maister (apud COGAN, 1998, p.13) desenvolveu teorias sobre a psicologia da espera focalizando gerenciamento das percepções e expectativas. Segundo ele:

$$\text{Satisfação} = \text{Percepção} - \text{Expectativa}$$

ou seja, a satisfação depende do que o cliente percebe e do que ele espera.

Larson (1987) em seu trabalho foi um dos pioneiros a propor gerenciar o tempo de fila, ocupando a atenção do cliente com algo que lhe pudesse interessar. Ele abordou a questão do que chama “justiça social da fila”, ou seja, isto deve ser entendido como “quem chega primeiro deve ser atendido antes”. Quando isto não ocorrer, ou seja, quando o sistema de fila é injusto (não obedece a esta prioridade) deste ponto de vista, o cliente pode ficar frustrado ou mesmo furioso. Ele propôs a eliminação do tempo não ocupado e o reconhecimento das filas como uma oportunidade para ocupar os clientes, inclusive com algo rentável para a empresa, usando o tempo que seria perdido.

O autor cita o exemplo do supermercado, no qual na abertura de um novo caixa, clientes rapidamente se deslocam de suas filas e, em algumas vezes, o último de alguma fila se torna o primeiro em outra. Apesar da intenção do supermercado ter sido reduzir as filas,

clientes desfavorecidos podem ficar enfurecidos. Finalmente, em seu trabalho, conclui ainda que algumas classes de usuários valorizam seu tempo mais que outras, e irão pagar para evitar ou reduzir esperas em filas.

Katz, Larson e Larson (1991) realizaram uma pesquisa em um banco nos EUA sobre as consequências de dois tipos de mudanças nas filas: na primeira, um visor com notícias para entreter os clientes na fila e na segunda, um marcador para medir os tempos de fila. Deste estudo surgiram algumas sugestões para gerenciar percepções dos clientes em espera, tais como, determinar o tempo tolerado de espera, captar a atenção dos clientes com distrações, se possível colocá-los fora da fila (sentados, por exemplo), manter recursos não utilizados para atendimento fora da visão dos clientes, etc. Esta pesquisa teve total enfoque em gerenciamento de percepções e isto fica mais claro quando os autores afirmam que “se os usuários pensam que estão satisfeitos, então eles estão satisfeitos” (p.44).

Geralmente o gerenciamento das percepções nas filas de espera tem sido utilizado quando não foi possível obter redução satisfatória do tempo de espera aplicando-se medidas operacionais tradicionais. Também, quando não se pretende despendere recursos para solucionar o problema, como documentado por Sasser et al (apud COGAN, 1998, p.13), no qual espelhos foram instalados em áreas de espera de elevadores em prédios. A finalidade era prender a atenção dos clientes que, enquanto se observavam nos espelhos, não notavam a demora, ao invés de investir no aumento de velocidade dos elevadores, sendo esta uma solução tecnicamente mais correta, mas de custo muito mais elevado.

Ainda que para os bancos a legislação determine que o tempo de atendimento não pode ultrapassar 30 minutos, estes têm conseguido aplicar o gerenciamento de percepções utilizando-se de senhas com prioridades, cadeiras de espera, aparelhos de TV, entre outros.

No caso dos supermercados, é um desafio a questão do gerenciamento de percepções de clientes. Algumas estratégias como espelhos, música e produtos próximos das filas em geral não se mostraram eficientes. Atualmente, o fator tempo está cada vez mais valorizado, desta forma, o consumidor tem se preocupado mais com isto. Também, depois dos produtos selecionados (mesmo que para isto se gaste muito tempo), o cliente em geral deseja retirar-se dali o mais breve possível.

O desconforto em empurrar o carrinho carregado até a fila e esperar em pé, são motivos suficientes para abalar sua satisfação com o estabelecimento, se o tempo gasto na fila for maior que o esperado. Ocorrências ainda não previstas como problemas no caixa e o sistema de fila injusto podem incomodá-lo ainda mais. Deve-se observar ainda que alguns produtos refrigerados são perecíveis e perdem suas características com a demora para acondicioná-los novamente em local apropriado.

De forma geral, quando o tempo que se dispõe é restrito, fato comum atualmente, gerenciar percepções pode não ajudar o suficiente, portanto, mais satisfação e fidelização poderão ser obtidas com a redução do tempo gasto em filas. Geralmente, a primeira idéia que surge para diminuição de tempo gasto é contratação de pessoal, contudo, um bom conhecimento e gerenciamento sobre como as filas ocorrem no estabelecimento podem propiciar melhorias com menor investimento. Uma modelagem adequada é tarefa bastante atrativa nestes casos.

2.5 A troca de filas

Quando um cliente entra em (ou escolhe) uma fila ele provavelmente fez alguma avaliação na qual vislumbra a hipótese de que levará um tempo menor para ser atendido. Quando esta expectativa se frustra, ou seja, a fila demora mais do que ele imaginou, é possível que reconsidere sua decisão e possa mudar para outra fila que, em sua visão, tenha maior possibilidade de ser atendido mais rapidamente. Na conduta de *trocas* por parte do cliente, duas questões interessantes surgem.

A primeira é que quando sua expectativa é frustrada, ele não deixa de analisar a causa e natureza da ocorrência. Esta pode ser externa ou interna à loja. Como exemplo, um cliente excessivamente lento à sua frente na fila é uma causa externa e a responsabilidade não recai sobre a empresa. Contudo, quando a causa é interna, por exemplo, um operador de caixa lento ou um preço de mercadoria não cadastrado no sistema, o cliente responsabiliza a empresa e pode ficar com uma imagem negativa da mesma. Esta imagem é percebida em graus de gravidade variados, dependendo da especificidade da causa averiguada pelo cliente.

A segunda questão é que, se a expectativa do cliente é frustrada, e um outro cliente que chegou depois dele, em outra fila, vai ser atendido antes, o seu grau de descontentamento tende a aumentar, pois ele percebe que aquele sistema de atendimento pode levar a situações injustas.

Quando então um cliente troca de fila, geralmente ele está buscando minimizar seu tempo de permanência em espera. Também, pode estar tentando minimizar alguma expectativa de tempo frustrada durante a espera e/ou ainda em busca da justiça que o sistema deveria realizar, em atender primeiro quem chega antes. Finalmente, ainda que por nenhum desses motivos, podem haver *trocas* simplesmente por uma oportunidade que se apresenta na fila, que é percebida e aproveitada pelo cliente com o simples objetivo de abreviar sua espera.

O desafio de melhorar o atendimento em supermercados reduzindo-se a espera em filas é um problema antigo. Em 1996, segundo a consultora Nielsen (apud BNDES 1996, p.33), os principais motivos de queixas dos clientes em supermercados referiam-se a filas no caixa (56%), supermercado lotado (41%), falta de empacotador (35%) e falta de preço nos produtos (25%). Nota-se, portanto, que todos estes motivos afetam o tempo de permanência do consumidor na fila.

Estratégias podem ser combinadas para minimizar o problema, contudo, deve haver empenho em inovação de idéias para um gerenciamento que efetivamente reduza os tempos de espera na fila, com o menor investimento e/ou custo possível.

Capítulo 3 – A realização do experimento e da coleta de dados

3.1 Introdução

Na busca por um local adequado onde se pudesse desenvolver um estudo de caso e realizar a coleta de dados, deparou-se com algumas dificuldades inerentes aos gerentes com relação ao estudo de filas dentro de um contexto do gerenciamento dos supermercados. A proposta de estudo pressupunha uma contribuição futura à diminuição do tempo de fila dos clientes e, portanto, uma melhoria no nível de satisfação dos mesmos. Contudo, como é descrito abaixo, pôde-se perceber que, devido talvez ao desconhecimento do potencial de aplicação da área, houve resistência na inserção ao estudo proposto. Este fato um pouco inesperado acabou sendo absorvido de modo a ampliar e enriquecer o desenvolvimento da pesquisa.

3.2 A percepção de alguns gerentes

Foram contatadas quatro empresas para a realização deste trabalho. As três primeiras empresas possuíam características semelhantes, em estrutura, gerenciamento e posicionamento quanto à questão de otimizar a rapidez de atendimento e do tempo de espera nas filas. Nestas, a pesquisa não pôde ser realizada por razões apontadas como: não permitir filmagem da operação, não haver problemas com filas, criação de incômodo aos clientes, já possuir algum planejamento implantado para gerenciamento das filas, não revelar as estratégias que utilizavam para o atendimento dos clientes em filas, etc.

A quarta empresa contatada é uma rede de supermercados onde já havia sido realizado um estudo anterior sobre filas em supermercados (MORABITO e LIMA, 2004). Após ser consultada, a empresa decidiu pela realização da pesquisa por meio de um sócio-proprietário que demonstrou interesse, entretanto, colocou como condição para a coleta dos dados, que após o término desta, todas as fitas de filmagens fossem entregues à empresa, por motivos de confiabilidade. Optou-se, assim, por desenvolver o estudo de caso nesta empresa. Esta rede de supermercados atua na região de São Carlos com lojas espalhadas em algumas cidades da região. A empresa está estabelecida em São Carlos há mais de quarenta anos, com algumas lojas localizadas em

diferentes bairros. O supermercado em que foi realizado o estudo é a maior loja da rede, tendo um total de dezesseis caixas, entre rápidos, preferencial e normais.

3.3 Instrumentos e materiais para a coleta de dados

3.3.1 O questionário elaborado

Elaborou-se um questionário (apêndice A) com o intuito de comparar a importância de algumas características pelas quais clientes frequentam o supermercado. No questionário procurou-se avaliar o tempo máximo que um cliente espera na fila e o número máximo de pessoas na fila que ainda não o fazem desistir das compras. Esses dois limites serão relacionados, permitindo estabelecer uma impressão dos clientes entre o número de pessoas e o tempo de espera na fila. Uma outra questão procura estabelecer uma ordem de importância para os clientes em relação a características como: preços praticados, localização do estabelecimento, qualidade e rapidez de atendimento, entre outras. Estas foram informadas aos clientes, que decidiam como desejavam ordená-las. Foram também solicitadas informações a respeito de escolaridade dos clientes, número de itens comprados e opinião sobre a existência de diferença de preço entre estabelecimentos distintos. Finalmente, foram solicitadas informações sobre se o cliente mudou de fila e os motivos que o levaram a realizar a troca.

3.3.2 Realização da amostra piloto

A finalidade de uma amostra piloto, geralmente, é de testar os instrumentos de coleta de dados e verificar se o tamanho amostral proposto é adequado. Desta forma, foi realizada uma amostra piloto em uma sexta-feira, dia 19/11/2004, durante um período de duas horas (das 18h às 20h). Assim, pôde-se observar que, para o referido dia e horário, havia (como antecipado pelo gerente) uma boa frequência de clientes, com formação de filas nos caixas. Uma filmadora foi localizada diagonalmente com o objetivo de filmar todas as doze filas, captando chegadas, término de serviços e trocas de filas realizadas pelos clientes. Constatou-se, contudo, que apenas uma filmadora era insuficiente para todas as filas e, então, propôs-se ao supermercado utilizar duas filmadoras, na filmagem, o que foi recusado sob a alegação de que poderia haver cerceamento de liberdade e constrangimento aos clientes. Disposto isto, decidiu-se

manter uma filmadora como auxílio à coleta das mesmas informações, contudo, utilizando agora planilhas a serem preenchidas por pessoal treinado, o que foi aceito pela empresa.

Outra contribuição da amostra piloto foi a abertura de mais uma característica de frequência dos clientes no supermercado, o estacionamento, que foi apontado significativamente. Também, decidiu-se por não coletar idade, pois, pela amostra piloto, percebeu-se que esta questão causava constrangimento aos clientes. Assim, esta informação foi retirada do questionário.

Quanto ao tamanho amostral, pelas características do questionário, decidiu-se que, durante todo o período de coleta para as filas, os clientes seriam ininterruptamente entrevistados. Tal decisão se pautou em tomar o maior tamanho amostral possível, pois se despendia, em média, cerca de cinco minutos por entrevista para preenchimento do questionário. Desta forma, no citado período seriam coletadas aproximadamente vinte e quatro questionários, o que, ainda, foi considerado pouco. Decidiu-se, então, realizar a coleta com dois entrevistadores, dobrando o tamanho amostral. Assim, com este procedimento, fica caracterizada uma amostragem sistemática na qual um cliente é entrevistado, em média, a cada 2,5 minutos. Desta forma, pelos cálculos obtidos para a amostra piloto, seriam entrevistados cerca de 10% dos clientes que passariam pelos caixas quando da coleta de dados.

3.3.3 O planejamento da coleta de dados

A coleta de dados foi realizada em dias pré-determinados e no período de 2 horas por dia, por ser considerado (baseado na experiência dos gerentes) como um período de grande fluxo de clientes e de frequência aproximadamente estável. Este fato é importante para a suposição do processo de chegada de usuários, no qual a taxa de chegada deve ser aproximadamente estacionária, caso contrário, poder-se-ia considerar um processo não estacionário (por exemplo, um processo de Poisson heterogêneo), mas tornando a análise bem mais complexa.

Tendo sido realizada a amostra piloto no dia 19/11/2004 (sexta-feira) e com a aproximação do Natal, as coletas ficaram para os dias 03, 04 e 05/12 e 10, 11 e 12/12/2004, pois assim haveria cerca de duas semanas para confeccionar as planilhas, planejar a filmagem, selecionar e capacitar os entrevistadores e coletores. Veiculou-se a

possibilidade de realizar uma terceira etapa de coleta (nos dias 17, 18 e 19/12), o que não foi permitido pela empresa devido a proximidade do Natal. Avaliou-se também que poderia não ser adequado, pois poderia haver um movimento atípico ao andamento normal do supermercado, viciando os dados.

Na tabela 3.1 encontram-se discriminados os dias e horários de coleta bem como a quantidade de questionários coletados e o tempo médio de entrevista por questionário.

Tabela 3.1. Informações gerais sobre a coleta de dados do experimento

Data	Horário	Questionários coletados	Tempo médio por questionário
03/12/2004 (sexta-feira)	18h às 20h	55	2,18 min./quest.
04/12/2004 (sábado)	10h30min. às 12h30min.	55	2,18 min./quest.
05/12/2004 (domingo)	10h30min. às 12h30min.	47 (1)	2,50 min./quest.
10/12/2004 (sexta-feira)	18h às 20h	49 (1)	2,40 min./quest.
11/12/2004 (sábado)	10h30min. às 12h30min.	47 (2)	2,45 min./quest.
12/12/2004 (domingo)	10h30min. às 12h30min.	50	2,40 min./quest.

Observando-se a tabela, pode-se notar que os números estão bastante coesos sendo que, para a terceira coluna, os números entre parênteses representam os questionários retirados da amostra referentes a clientes sem escolaridade.

3.3.4 A planilha para coleta nas filas

Devido à filmagem demonstrar-se na amostra piloto como insuficiente, para obter as informações necessárias foi preciso utilizar-se uma planilha preenchida manualmente. Para que esta (apêndice B) pudesse atender às necessidades da pesquisa, inicialmente, ela deveria conter o instante de chegada dos clientes na fila, o instante de início do serviço nos caixas e o instante de término do serviço, quando o cliente se retira do sistema. Dever-se-ia também anotar o instante de duas outras ocorrências, que eram

a desistência (denotada por D) e a troca (*jockeying*) de filas. Nesta última, dever-se-ia também anotar, além do instante de troca, a fila para a qual o cliente se dirigiu. Outra informação anotada era referente à carga de trabalho (ou à carga relativa ao carrinho de compra) demonstrada para cada um dos clientes, sendo que isto era aferido pela quantidade aproximada de itens para cada compra. Assim, a carga poderia ser leve (L, até 10 itens), moderada (M, de 11 até 50 itens) e pesada (P, acima de 50 itens). Ainda, qualquer fato que pudesse afetar o andamento do experimento e, portanto, os resultados da planilha, deveriam ser anotados, como, por exemplo, se o operador de caixa parasse o atendimento.

Outras informações da planilha eram referentes a posição da fila (ou filas) do caixa, a data e horário de coleta e o nome do coletor. Por meio de todos esses dados acredita-se que as informações coletadas são suficientes para as análises propostas.

3.3.5 A filmadora utilizada

A filmagem acabou tornando-se um auxílio para checagem das informações coletadas por meio das planilhas. Era necessário que a filmadora pudesse registrar as imagens com cronometragem também em segundos e que tivesse duração de pelo menos 2 horas, possuindo ainda controle remoto para ser ligada a distância. Foi utilizada uma filmadora com fitas VHSc JVC EHG40, que em modo SLP tinha duração em torno de 2 horas e 15 minutos. Desta forma, o equipamento foi capaz de atender as necessidades do experimento.

3.3.6 A caixa para ocultar a filmagem

Para efetuar a filmagem de forma a não constranger os clientes do supermercado, foi construída uma caixa de madeira para a filmadora não ficar aparente. Esta caixa era constituída por uma base quadrada de 32 centímetros e com 35 centímetros de altura, o interno pintado de preto fosco para não propagar luz e a frente com vidro especial espelhado não permitindo visão de seu conteúdo. A tampa se encontra atrás com fechamento por velcro colado (vide figuras 3.1 e 3.2). Não foi notado qualquer cliente observando a caixa durante as filmagens. Esta ficava alocada no alto de uma das prateleiras, quase imperceptível. Aos operadores de caixa foi informado

que a filmagem era relativa somente aos clientes na fila e não aos caixas, despreocupando-os de qualquer problema que estes pudessem imaginar.

3.3.7 Cronometragem com computador e monitores

Como cronômetro foi utilizado um computador ligado a dois monitores, nos quais estava aparente a passagem do tempo da coleta. O computador utilizado era bastante comum, cuja única informação utilizada era um relógio com horas, minutos e segundos. A ele estava acoplado um cabo de monitor que dividia o sinal em duas saídas (um divisor de frequências) para utilização de dois monitores, cada um distante cerca de 5 metros por meio de cabos extensores interligados, com 1 metro de comprimento cada. Não houve prejuízo de sinal devido as extensões e a imagem do cronômetro nos monitores apresentava excelente definição, podendo ser vista perfeitamente a distância.



Figura 3.1. Visão frontal da caixa para a filmadora.



Figura 3.2. Visão interna da caixa com a filmadora.

3.3.8 Recursos humanos empregados

Devido à impossibilidade de se utilizar duas ou mais filmadoras para a coleta de dados, conforme descrito anteriormente houve a necessidade de contratação e treinamento de pessoal para a realização de tal tarefa. Foram, então, contratadas dez pessoas, sendo duas delas responsáveis por realizar as entrevistas, cuja coleta aconteceu por meio do questionário, e oito pessoas para atuarem como coletores, ou seja, eram responsáveis pelo preenchimento dos dados da planilha.

Os coletores selecionados tinham experiência de coleta de dados realizada por meio de planilhas em empresas do setor privado. Os requisitos solicitados eram de que estas pessoas fossem rápidas, atentas, tivessem disciplina no cumprimento de horários, conduta e postura ética, soubessem utilizar planilhas estatísticas e, principalmente, conhecimento sobre a importância que os dados representam para uma pesquisa desta natureza. Assim, foram contratados alunos do último ano do curso de Estatística da UFSCar, por preencherem a todos os requisitos desejados e com a finalidade de garantir a eficácia e fidedignidade na coleta de dados.

Cada coletor tinha sob sua responsabilidade uma das filas de caixa rápido ou da preferencial, sendo então alocados quatro deles para esta finalidade. Para os oito caixas normais, utilizaram-se quatro coletores, ou seja, cada um era responsável simultaneamente por duas filas adjacentes.

Os entrevistadores que ficaram responsáveis pelos questionários tinham formação superior completa e experiência em atendimento ao público. Este perfil se fez necessário devido a diversidade dos clientes que freqüentavam o supermercado e a especificidade das questões a serem respondidas. Os entrevistadores necessitavam de flexibilidade na condução das entrevistas com o propósito de obter as informações de modo que estas expressassem verdadeiramente a opinião de quem as prestavam. Ainda, pelo plano amostral utilizado, o fator da rapidez era importante, pois implicava diretamente no tamanho amostral obtido.

Cabe ressaltar que todos os envolvidos assinaram um termo de compromisso no qual assumiram total e irrestrito compromisso de manter sigilo quanto as informações obtidas por meio dos instrumentos de coleta (planilhas) e cumprimento dos horários determinados para tal, conforme pode-se observar em detalhes no apêndice C.

3.4 O experimento

Para a realização do experimento, o computador foi ligado a um *nobreak* e colocado em funcionamento 20 minutos antes do seu início. Para poder acoplar o segundo monitor, era necessário esperar todo o processo de inicialização do computador, caso contrário não havia sinal (imagem), pois os monitores tinham marcas e configurações diferentes, o que gerava incompatibilidade. A filmadora era disparada por controle remoto imediatamente quando o cronômetro do computador chegava ao

horário especificado para início da coleta. A filmagem se dava em formato longo, com gravação na fita de tempo em horas, minutos e segundos.

Procedendo assim, os cronômetros apresentavam pouquíssima diferença e quando algo estranho ocorria, como um erro de anotação na planilha ou alguma observação sobre algum fato, podia-se ser auxiliado pela imagem da fita efetuando-se a correção, se necessária. O objetivo final desta preparação foi a de que todos os coletores utilizassem o mesmo cronômetro, evitando possíveis discrepâncias que poderiam ocorrer se utilizadas várias unidades (um cronômetro para cada coletor, por exemplo). Também, evitava-se qualquer manuseio deste equipamento, facilitando o decorrer do experimento.

Ao iniciar a coleta de dados os coletores começavam o preenchimento da planilha com o próximo cliente a entrar na fila, com seu instante de chegada. A partir de então, este era acompanhado e os próximos clientes que entrassem na fila tinham suas informações anotadas. Devido a este fato, o cliente em atendimento e os que já se encontravam na fila antes do horário previsto para o início da coleta não poderiam pertencer à amostra, assim como não deveriam ser entrevistados. Desta forma, as entrevistas eram iniciadas com um tempo estipulado de cinco minutos após o início da coleta e encerradas também cinco minutos depois do término desta.

Pela figura 3.3 pode-se obter uma melhor compreensão do experimento. Os clientes entravam no supermercado por meio de dois acessos, por uma rampa do estacionamento coberto e por uma entrada no mesmo nível do piso do supermercado (setas verdes), em seguida, se dirigiam às gôndolas onde selecionavam os itens a serem comprados. Posteriormente, dirigiam-se para as filas do caixa segundo o número de itens a serem adquiridos. Após a compra se retiravam (setas laranja) pelos mesmos locais de acesso.

Nos dias de coleta estavam em funcionamento três caixas rápidos (R), um preferencial (P) e oito normais (N), sendo que havia ainda a possibilidade de abertura de mais quatro caixas, o que não ocorreu durante os períodos de coleta. Os coletores se posicionavam próximo às gôndolas e em frente aos caixas sem obstruir a fila (losangos em amarelos, denotados por Ci). Os entrevistadores se colocavam no final do corredor de saída dos clientes que efetuavam compras (losangos amarelos, denotados por E1 e E2). Finalmente, a caixa com a filmadora foi posicionada em ângulo sobre uma gôndola em uma parede lateral do supermercado (vide figura 3.3).

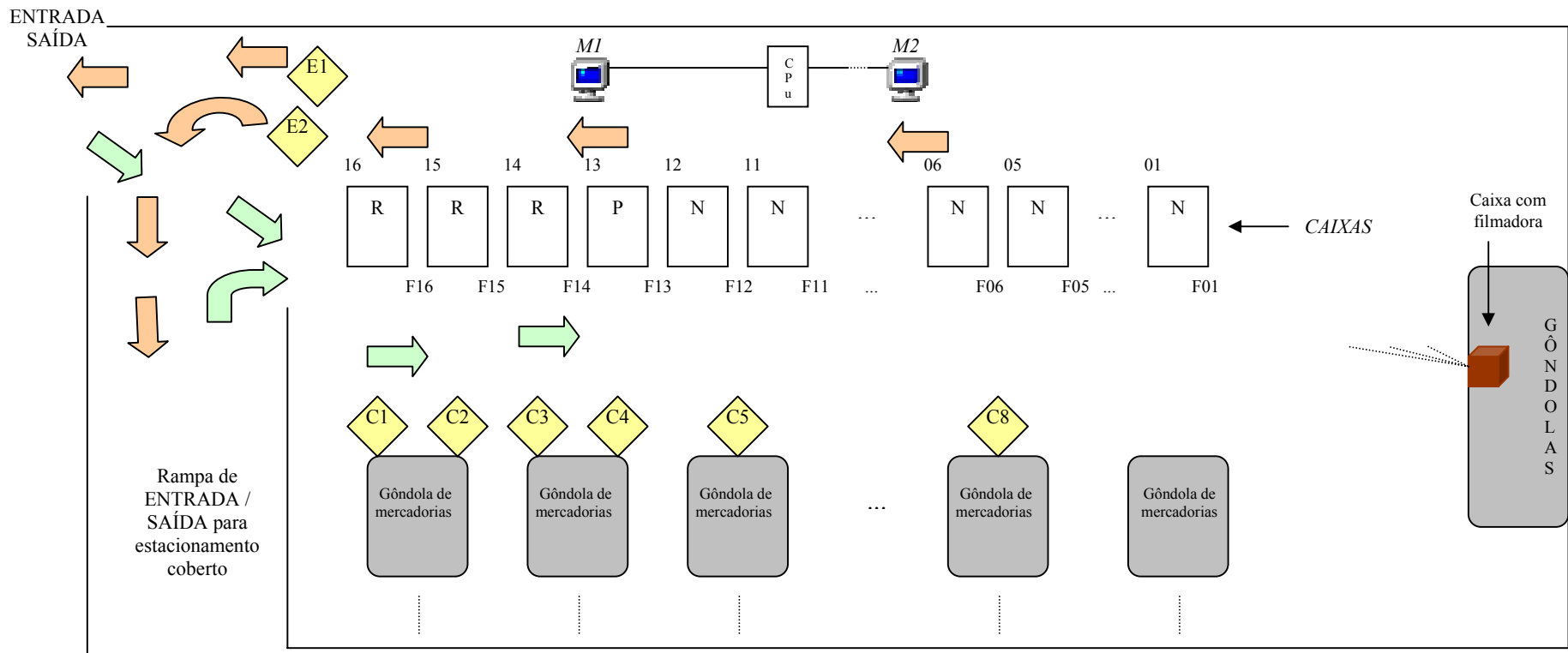


Figura 3.3. Representação do supermercado: posição dos caixas, entrevistadores, coletores e monitores

- Legendas: M1: (:monitor 1 com cronômetro); M2 (monitor 2 com cronômetro).
 CPU (relógio com precisão em segundos).
 E1 (entrevistador 1); E2 (entrevistador 2).
 R (caixa rápido); P (caixa preferencial); N (caixa normal).
 Fi (fila i; i = 1, 2, ..., 16).
 Ci (coletor i; i = 1, 2, ..., 8).

3.5 Considerações adicionais

Apesar de os coletores e entrevistadores estarem o mais imperceptível possível, ao iniciar a coleta de dados, a presença de algumas pessoas com pranchetas nas mãos provocou a incidência de alguns episódios em que os clientes se reportavam aos coletores e entrevistadores com o intuito de saber o que estava ocorrendo.

Alguns poucos clientes paravam próximo aos coletores e entrevistadores para verificar o que eles anotavam, havendo aqueles que perguntavam se era uma avaliação dos operadores de caixa e se estes poderiam perder o emprego, evidenciando visível preocupação com os mesmos. Estas ocorrências eram imediatamente levadas ao responsável pela pesquisa, que esclarecia o motivo pelo qual o trabalho estava sendo realizado e que seu objetivo não era avaliar os funcionários, mas possibilitar a melhoria do atendimento geral do supermercado e, proeminentemente, das filas dos caixas. Cabe ressaltar que alguns operadores trabalham no local há muitos anos, inclusive conhecendo clientes pelo nome, e revelando o motivo da preocupação dos mesmos.

Houve pessoas que ao serem abordadas para a entrevista não queriam responder ao questionário, pois alegavam não ter tempo; estas eram imediatamente liberadas e logo na sequência outro cliente era convidado pelos entrevistadores.

O supermercado analisado dispunha de empacotadores que permitiam a agilização do tempo de serviço dos caixas e comodidade dos clientes que, quando terminavam de efetuar o pagamento, tinham prontos os pacotes de compras.

Dependendo do volume adquirido, os empacotadores prontamente colocavam os pacotes em carrinhos e conduziam estes até o veículo do cliente. Por vezes, quando o cliente residia nas redondezas, esse carrinho era conduzido até sua residência. Ainda, o supermercado dispunha de serviços de entrega, abrangendo todas as alternativas possíveis de atendimento.

Fatos bastante interessantes podem ser relatados a respeito da operação dos caixas. Alguns supermercados, devido ao crescimento de demanda, deveriam abrir mais caixas, contudo, muitos não possuem sequer espaço físico para isso. No supermercado analisado, havia um total de dezesseis caixas, mas, segundo o próprio gerente, a abertura de todos eles raramente ocorria por não haver necessidade. Este fato sugere que houve um bom planejamento da parte de estrutura física do estabelecimento, possibilitando até mesmo a ampliação do número de caixas. Outro fato relevante diz

respeito ao caixa preferencial. Este deveria atender com prioridade pessoas idosas, gestantes, lactantes (com crianças de colo) e deficientes, contudo, isto era pouco observado. O que muitas vezes ocorria era que as pessoas que tinham prioridade não a utilizavam, colocando-se em outras filas, notadamente em caixas rápidos. O que se pôde compreender desta situação foi que, devido ao pequeno tempo de espera nas filas, estes não faziam muita questão deste benefício. Deste modo, o caixa preferencial acabou por funcionar de maneira muito semelhante aos outros três caixas rápidos.

O experimento transcorreu conforme o planejamento realizado, sem ocorrências que pudessem prejudicá-lo e possibilitando inclusive uma maior proximidade com a clientela, que era bastante acessível. Isto fez com que se conseguisse contemplar detalhes além do que a planilha ou questionário foi capaz de captar.

Capítulo 4 – A análise dos questionários utilizando técnicas multivariadas

4.1 Introdução

O questionário aplicado no supermercado foi realizado com a intenção de levantar características importantes para os clientes, ou seja, que estimula ou mesmo determina suas frequências ao local. Por outro lado, pretende também levantar possíveis problemas ou falhas que os clientes possam apontar em relação ao atendimento recebido. Estes problemas devem ser considerados como oportunidades na melhoria constante da qualidade de atendimento, que deve ocorrer em uma empresa que seja competitiva no seu mercado de atuação.

Características importantes, como a tolerância máxima dos clientes em relação ao tamanho e tempo de espera na fila, são abordadas. Ainda, com o objetivo de comparação, algumas características desejáveis são ordenadas pelos consumidores por ordem de preferência, permitindo compará-las, tais como preço, localização, rapidez e qualidade de atendimento, variedade e ainda outras, apontadas com menor frequência. Das comparações pôde-se obter as predileções (ou *rank* de causas) dos clientes em freqüentar o estabelecimento. Outras informações são analisadas como a escolaridade dos entrevistados, o número de itens por eles comprados e ainda se acreditam que haja diferença de preço entre supermercados distintos.

Apesar da riqueza que todas estas informações trazem isoladamente, é ainda mais interessante analisá-las conjuntamente, ou seja, relacioná-las pode levar a descobertas úteis para um melhor gerenciamento da qualidade de atendimento. Pode-se, por exemplo, associar níveis de escolaridade com as características citadas anteriormente.

Para realizar as análises pretendidas, técnicas estatísticas multivariadas podem ser empregadas, tais como análise de agrupamentos e análise de correspondência. Estas técnicas têm sido freqüentemente utilizadas em áreas como gerenciamento, *marketing*, medicina, sociologia, economia e mercado, entre várias outras. As técnicas multivariadas permitem analisar o comportamento de várias variáveis simultaneamente. Sua compreensão deve auxiliar bastante nas particularidades exibidas por dados multivariados, como ocorre aqui.

4.2 Limite de espera e importância das características analisadas na frequência ao supermercado

A idéia básica para as duas primeiras perguntas realizadas aos clientes (apêndice A) é estabelecer uma relação entre o número máximo de clientes em fila (pergunta 1) e o máximo tempo de espera em minutos na fila (pergunta 2) que estes clientes suportariam. Uma regressão linear (sem constante) entre o tempo e o número de clientes pode ser obtida e comparada com a conhecida lei de Little, na qual $L_q = \lambda W_q$, ou seja, o número esperado de clientes na fila L_q é igual a taxa média de chegada λ multiplicada pelo tempo médio de espera na fila W_q (KLEINROCK, 1975). Deve-se ressaltar que a lei de Little é para médias e a relação aqui estabelecida é para máximos de tempo e número de clientes na fila. Pesquisas indicam que os usuários sabem avaliar mais facilmente o número de clientes na fila, no entanto, não conseguem estabelecer acertadamente o tempo de espera destes.

Uma equação de regressão linear foi ajustada utilizando-se a resposta de 172 clientes dos 303 entrevistados, sendo que os demais 131 responderam que não abandonariam as compras e entrariam na fila em qualquer situação que a encontrasse. A regressão simples tem equação dada por

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \text{ para } i = 1, 2, \dots, n \quad (4.1)$$

nos quais y_i é o predito, x_i é o preditor e ε_i é um erro associado ao modelo. Para os dados em questão, o número máximo de clientes (da segunda questão) é o predito e o tempo máximo de espera em minutos (da primeira questão) é o preditor. Admite-se ainda que $\beta_0 = 0$ para comparação com a lei de Little. Desta forma, utilizando-se os dados coletados na amostra, foi obtida a equação de regressão

$$(\text{número máximo de clientes}) = 0,42719 \cdot (\text{tempo máximo de espera}).$$

Na teoria de regressão, um primórdio básico para que o modelo seja válido é avaliar o coeficiente de determinação (R^2). Este pode ser entendido como sendo a proporção de variação explicada pelo preditor, sendo que $0 \leq R^2 \leq 1$. Valores perto de 1 implicam que grande parte da variabilidade é explicada pelo modelo de regressão e, ao

contrário, quando perto de 0, que grande parte da variabilidade está nos erros. Para o modelo ajustado, uma aproximação para R^2 é dada por (MONTGOMERY e PECK, 1992, p.48):

$$R^2 = \frac{\sum_{i=1}^n \hat{y}_i^2}{\sum_{i=1}^n y_i^2} \quad (4.2)$$

na qual \hat{y}_i é o valor estimado pelo modelo para y_i . Contudo, conforme (4.2), foi obtido $R^2 = 0,52$ para o modelo ajustado e, sendo considerado baixo, pode-se dizer que não é suficiente para prever os valores y_i . Pode-se observar na figura 4.1 que os valores apontados pelos 172 clientes estão muito dispersos, explicando o baixo valor de R^2 .

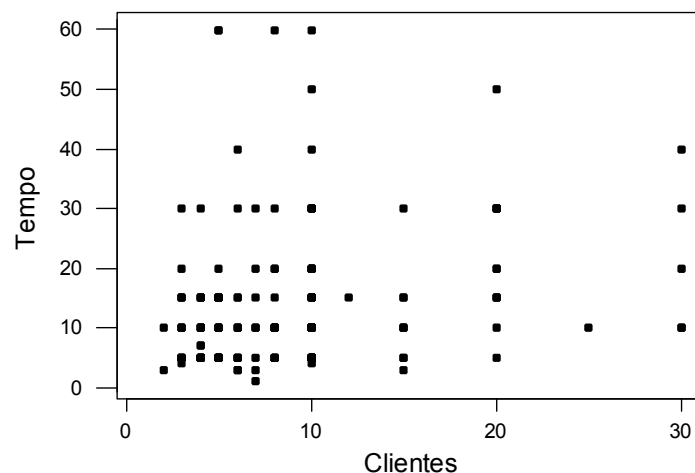


Figura 4.1. Gráfico do tempo máximo de espera (x_i) pelo número máximo de clientes na fila (y_i).

Nota-se na figura uma grande dispersão e, ainda, uma tendência na repetição das opiniões verificada pela sobreposição de vários pontos. Fica realmente difícil aproximar os valores por uma reta passando pela origem, como ocorre na lei de Little. Alguns clientes admitem entrar em fila com até 30 pessoas, contudo a espera máxima apontada é de 10 minutos, o que é praticamente impossível de ocorrer. Assim, aparentemente, grande parte dos respondentes não é capaz de relacionar efetivamente o número de usuários em fila com seus tempos de espera na fila.

Um aspecto importante para este capítulo está na análise da terceira questão formulada aos clientes, onde estes alocam em ordem de preferência os motivos pelos quais freqüentam o supermercado (apêndice A). Os dados para esta questão foram introduzidos segundo a ordem com que cada cliente manifestou sua preferência. Como exemplo, se um cliente colocou em primeira posição de preferência localização, em segundo preço, em terceiro qualidade de atendimento e não apontou mais nenhuma outra, o preenchimento da planilha ficou:

Preço	Localização	Rapidez	Qualidade	Variedade	Estacionamento	Outros
6	7	0	5	0	0	0

ou seja, 7 para a característica mais importante, 6 para a segunda, e assim consecutivamente (repare-se que há 7 alternativas), sendo 0 para as ausentes na resposta (que não são consideradas de alguma importância para o consumidor).

A suposição de normalidade para testes de hipótese e regiões de confiança para um vetor de médias é de difícil constatação e poucas vezes verificada, mesmo para pequenas dimensões. Algumas alternativas podem ser conseguidas quando se possuem amostras relativamente grandes, contornando a necessidade de normalidade. Um resultado bastante útil é enunciado. Seja $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ uma amostra aleatória de uma população com média $\boldsymbol{\mu}$ e matriz de covariâncias positiva definida $\boldsymbol{\Sigma}$. Se $n-p$ é grande, então:

$$\mathbf{a}'\bar{\mathbf{X}} \pm \sqrt{\chi_p^2(\alpha)} \sqrt{\frac{\mathbf{a}'\mathbf{S}\mathbf{a}}{n}} \quad (4.3)$$

irá conter $\mathbf{a}'\boldsymbol{\mu}$ para todo $\boldsymbol{\mu}$, aproximadamente com probabilidade $1-\alpha$, sendo p o número de variáveis aleatórias, \mathbf{a} um vetor de constantes, $\bar{\mathbf{X}}$ o vetor de médias amostrais, $\chi_p^2(\alpha)$ é o valor crítico da distribuição Qui-Quadrado com p graus de liberdade e nível de significância α e \mathbf{S} a matriz amostral de covariâncias (JOHNSON e WICHERN, 2002). Conseqüentemente, do resultado, obtêm-se os respectivos intervalos simultâneos de confiança $100(1-\alpha)\%$ para $\mu_1, \mu_2, \dots, \mu_p$

$$\bar{x}_1 \pm \sqrt{\chi_p^2(\alpha)} \sqrt{\frac{s_{11}}{n}};$$

$$\begin{aligned}
 & \bar{x}_2 \pm \sqrt{\chi_p^2(\alpha)} \sqrt{\frac{s_{22}}{n}} ; \\
 & \quad \vdots \\
 & \bar{x}_p \pm \sqrt{\chi_p^2(\alpha)} \sqrt{\frac{s_{pp}}{n}} ,
 \end{aligned} \tag{4.4}$$

nos quais s_{ii} é a variância amostral da variável aleatória i . Aqui admite-se que $\mathbf{a}=[0,0,\dots,1,0,\dots,0]$, colocando-se a constante 1 na posição adequada para o intervalo pretendido. Considerando-se os valores coletados na amostra para as variáveis preço, localização, rapidez, qualidade, variedade, estacionamento e outras, foram obtidos os intervalos simultâneos de confiança com $\alpha=0,005$, conforme (4.4), dados na tabela 4.1.

Analisando-se estes intervalos simultâneos de confiança, com $\alpha=0,005$, observa-se que a variável qualidade de atendimento é considerada a característica mais importante pela qual o cliente frequenta o supermercado, pois seu intervalo [4,11562, 4,98009] não é sobreposto por nenhum dos outros. Também, a variável rapidez de atendimento é mais importante que preço, variedade, estacionamento e outras, contudo, não se pode dizer que rapidez é mais importante que localização. Ainda, preço, localização e variedade são mais importantes que estacionamento e outras, sendo estas duas últimas de menor importância. Fazendo-se $\alpha=0,05$, a única mudança constatada é que localização se torna significativamente mais importante do que variedade para os clientes. Este fato ressalta a consistência que esta análise traz quanto às características estudadas. Aqui é possível expressar a força de cada característica na frequência do cliente ao supermercado.

Tabela 4.1. Intervalos de confiança simultâneos com $\alpha=0,005$.

Variável	Média	s_{ii}	Limite Inferior	Limite Superior
Preço	2,41584	2,83521	1,98022	2,85146
Localização	3,03630	3,13314	2,57837	3,49424
Rapidez	3,35974	2,90112	2,91908	3,80039
Qualidade	4,54785	2,79134	4,11562	4,98009
Variedade	2,28713	2,73318	1,85942	2,71484
Estacionamento	0,21122	0,99748	-0,04716	0,46961
Outras	0,56106	1,70568	0,22318	0,89894

Resumindo, pelos intervalos de confiança, é interessante observar que apostar em preço, como é comum segundo os gerentes de supermercados, pode ser um engano dependendo da clientela. Neste supermercado, pelo menos, clientes buscam prioritariamente qualidade e rapidez de atendimento. A próxima técnica empregada dá informações sobre o perfil de conduta dos clientes, ou seja, explicita a relação de proximidade entre essas características conforme as respostas dos clientes.

4.3 O perfil das características pela análise de agrupamentos

A análise de agrupamentos (ou conglomerados) difere de métodos de classificação, pois nestes últimos o número de grupos pode ser conhecido. Este tipo de análise não pressupõe número nem uma estrutura de conglomerados. Esta técnica toma como base similaridades (ou proximidades) entre itens (objetos, indivíduos ou variáveis) que podem ser de alguma forma medidas. Como resultado, obtém-se um certo número de conglomerados no qual se alocam, devido a sua proximidade, os itens, inicialmente dispersos (JOHNSON e WICHERN, 2002).

Os procedimentos de conglomerados podem ser divididos em hierárquicos e não-hierárquicos. Os procedimentos hierárquicos geralmente partem de um número N' total de itens, formando inicialmente $N' - 1$ conglomerados, consecutivamente $N' - 2$, e assim consecutivamente, até que um único conglomerado contendo todos os itens seja formado, de modo aglomerativo. Uma vez que dois itens são inicialmente agrupados, no decorrer do processo, não serão mais separados. Tal procedimento assemelha-se a construção em forma de árvore, unindo primeiro os conglomerados com maior similaridade. Medidas de similaridade são utilizadas inicialmente como critério para a junção dos conglomerados. Estas medidas podem ser caracterizadas como medidas de distância ou também como medidas correlacionais. As medidas de distância são freqüentemente utilizadas quando objetos ou indivíduos são agrupados (JOHNSON e WICHERN, 2002).

Uma das principais medidas de distância é a distância Euclidiana. Seja $d(\mathbf{x}, \mathbf{y})$ a distância entre dois itens de dimensão p , no qual $\mathbf{x} = [x_1, x_2, \dots, x_p]$ e $\mathbf{y} = [y_1, y_2, \dots, y_p]$. A distância Euclidiana é dada por:

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_p - y_p)^2} = \sqrt{(\mathbf{x} - \mathbf{y})'(\mathbf{x} - \mathbf{y})}. \quad (4.5)$$

Esta distância é uma das mais conhecidas e utilizadas para análise de agrupamentos. É comum, também, que seja utilizada como distância $d(\mathbf{x}, \mathbf{y})^2$. Uma outra medida de distância bastante aplicada é a distância estatística. Esta distância entre os mesmos itens citados anteriormente é dada por:

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{(\mathbf{x} - \mathbf{y})' \mathbf{A}(\mathbf{x} - \mathbf{y})} \quad (4.6)$$

na qual, geralmente, tem-se que a matriz $\mathbf{A} = \mathbf{S}^{-1}$, sendo \mathbf{S} a matriz amostral de variâncias e covariâncias. Nota-se que é necessário que se conheça *a priori* os agrupamentos para calcular \mathbf{S} , o que pode inviabilizar a aplicação desta distância. Ainda, quando $\mathbf{A} = \mathbf{S}^{-1}$ em (4.6), esta medida é conhecida como distância de Mahalanobis. Uma outra distância que pode ser utilizada é a distância de Minkowski. Esta é dada por:

$$d(\mathbf{x}, \mathbf{y}) = \left[\sum_{i=1}^p |x_i - y_i|^l \right]^{1/l}.$$

Quando $l = 1$ a medida resultante é a distância absoluta (ou *city-block*). Quando $l = 2$ é obtida a distância Euclidiana. Variando-se l é propiciada uma mudança de pesos para maiores e menores diferenças. Alguns autores (REIS, 2001) recomendam que, sempre que possível, sejam utilizadas distâncias que respeitem as seguintes propriedades métricas:

Simetria: $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$;

Diferenciabilidade: $d(\mathbf{x}, \mathbf{y}) > 0$ se $\mathbf{x} \neq \mathbf{y}$;

Indiferenciabilidade: $d(\mathbf{x}, \mathbf{y}) = 0$ se e somente se $\mathbf{x} = \mathbf{y}$;

Desigualdade triangular: $d(\mathbf{x}, \mathbf{y}) < d(\mathbf{x}, \mathbf{z}) + d(\mathbf{z}, \mathbf{y})$, em que \mathbf{z} é um ponto intermediário.

Quando o objetivo é agrupar variáveis as medidas de similaridade geralmente empregadas são medidas correlacionais. Destas, a principal medida é a

correlação, muito utilizada por ser de fácil interpretação. Sejam duas variáveis aleatórias i e j , a correlação entre elas é definida por:

$$\rho_{ij} = \frac{\sum_{k=1}^n (x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j)}{\sqrt{\sum_{k=1}^n (x_{ki} - \bar{x}_i)^2} \sqrt{\sum_{k=1}^n (x_{kj} - \bar{x}_j)^2}} \quad (4.7)$$

na qual n é o número de itens. Utilizando-se como distâncias $d_{ij} = 1 - \rho_{ij}$, obtêm-se menor distância (maior similaridade) quanto maior a correlação.

Pode-se também empregar a correlação absoluta, ou seja, a distância fica sendo $d_{ij} = 1 - |\rho_{ij}|$. A escolha entre correlação ou correlação absoluta deve ser realizada considerando-se a importância que cada forma tem em relação a natureza dos dados a serem analisados. Contudo, estas duas formas não obedecem as propriedades métricas citadas anteriormente, pois não satisfazem a propriedade de desigualdade triangular. Para tanto, uma transformação do coeficiente de correlação pode ser considerada. Fazendo-se:

$$d_{ij} = [0,5(1 - \rho_{ij})]^{1/2} \quad (4.8)$$

a nova distância obtida obedece as propriedades métricas. Novamente, quanto maior a correlação, menor a distância, que varia entre 0 e 1 neste caso. Uma terceira possibilidade que pode ser empregada como medida de similaridade são os chamados coeficientes de similaridade (JOHNSON e WICHERN, 2002), contudo, são utilizados somente para variáveis binárias, o que restringe bastante sua aplicação.

Quando se deseja fazer agrupamentos com padrões similares, estes devem ser realizados com medidas correlacionais, podendo não haver valores similares, mas obtendo padrões similares, como é o caso de variáveis. Quando se deseja fazer agrupamentos com valores similares, deve-se empregar medidas de distância, podendo obter padrões diferentes, mas havendo valores mais similares, como é o caso de objetos e indivíduos.

Os conglomerados hierárquicos possuem vários métodos que determinam como as medidas de similaridade, entre os conglomerados, é definida. Inicialmente, quando cada observação constitui um conglomerado, a distância entre conglomerados é simplesmente a distância entre os itens. Posteriormente, quando os itens vão sendo

alocados, uma regra de ligação se faz necessária para calcular a similaridade entre os agrupamentos, sendo que em pelo menos um deles há mais de um item (HAIR et al., 2005).

Uma regra de ligação freqüentemente utilizada é denominada ligação simples (*single linkage*). Ela define como critério de ligação (semelhança) a menor distância entre dois itens quaisquer pertencentes a agrupamentos distintos. São calculadas as menores distâncias assim definidas entre todos os agrupamentos existentes, e a distância mínima destas apontará quais agrupamentos serão agregados. Este método tende a ligar itens (grupos unitários), preferencialmente, a grupos já definidos, ao invés de formar novos agrupamentos.

Outro método bastante utilizado é o de ligação completa (*complete linkage*), sendo que este é definido como o inverso do anterior. Deste modo, são calculadas as maiores distâncias entre todos os agrupamentos existentes e a menor delas apontará quais agrupamentos serão agregados. Este método tende a obter agrupamentos cujos itens são muito semelhantes entre si, contudo, algumas vezes não apresentam concordância elevada com a estrutura inicial dos dados.

Um método intermediário entre os anteriores que também pode ser utilizado é o de ligação média. Este método tem início como os demais tomando a menor distância para formar o primeiro agrupamento. Após isto, são calculadas médias de distâncias entre todos os itens de agrupamentos distintos, para todos eles. A menor distância média de todas indicará a junção a ser realizada. Ao contrário dos dois métodos anteriores, este não é afetado por valores extremos, a maior ou menor distância, considerando todos os itens dos agrupamentos.

Ainda, um outro método que pode ser considerado é o método do centróide. Neste, a distância entre dois agrupamentos é a distância entre seus centróides. Os centróides são definidos como as médias de todas as variáveis que caracterizam os itens de cada agrupamento. A menor distância indica qual junção deve ser realizada. Neste método pode ocorrer que as características de pequenos agrupamentos se percam quando agregados a grandes agrupamentos.

Por último, um método que pode ser aplicado é o método Ward, que tem como objetivo minimizar a variação interna dos agrupamentos conforme junções vão sendo realizadas. Este método tem tendência de formar grupos de tamanhos semelhantes e, em alguns casos, impor uma estrutura aos dados que não lhe é própria.

Todos estes métodos podem ser aplicados e comparados, obtendo-se um resultado mais consistente na revelação da estrutura dos dados analisados. Deve-se ressaltar que existem outros métodos a serem considerados, porém com utilização pouco freqüente.

Os conglomerados hierárquicos podem ainda ser classificados como aglomerativos, como na situação descrita acima, ou divisivos, quando partem de um único grupo contendo todos os elementos e vão sucessivamente sendo divididos até que cada elemento forme um grupo.

Técnicas de conglomerados não-hierárquicos são preferencialmente designadas para agrupar itens do que variáveis (JOHNSON e WICHERN, 2002). O número de conglomerados pode ser especificado *a priori* ou determinado durante o próprio procedimento a ser realizado para agrupamento. Neste tipo de conglomerado é permitido que os itens mudem de um conglomerado para outro enquanto o procedimento se realiza. Inicialmente um conglomerado central é selecionado e todos os itens em um limiar pré-definido estão nele contidos. Posteriormente, da mesma forma, um outro conglomerado é escolhido e todos os itens limiares são alocados, repetindo a operação até que todos estejam em algum conglomerado.

Os itens podem ser realocados se estiverem mais próximos de um outro conglomerado do que daquele que originalmente estavam. Nos conglomerados não-hierárquicos geralmente é utilizado um método denominado *k-means*. Este é inicializado particionando-se os itens em k conglomerados e, posteriormente, esses itens são analisados, sendo que aquele que estiver mais próximo de um centróide de outro conglomerado para lá é deslocado. Distâncias e centróides são recalculados até que nenhuma realocação seja realizada. Percebe-se aqui a dependência deste processo no valor de k , que é fixado arbitrariamente. Uma representação geral do método de agrupamentos está esquematizada na figura 4.2.

4.3.1 Algumas considerações sobre o método de agrupamentos

Uma crítica ao agrupamento não-hierárquico é que o pesquisador pode conseguir soluções diferentes quando inicialmente são designados conglomerados centrais distintos. Desta forma a dúvida seria qual a melhor representação de estrutura da questão investigada.

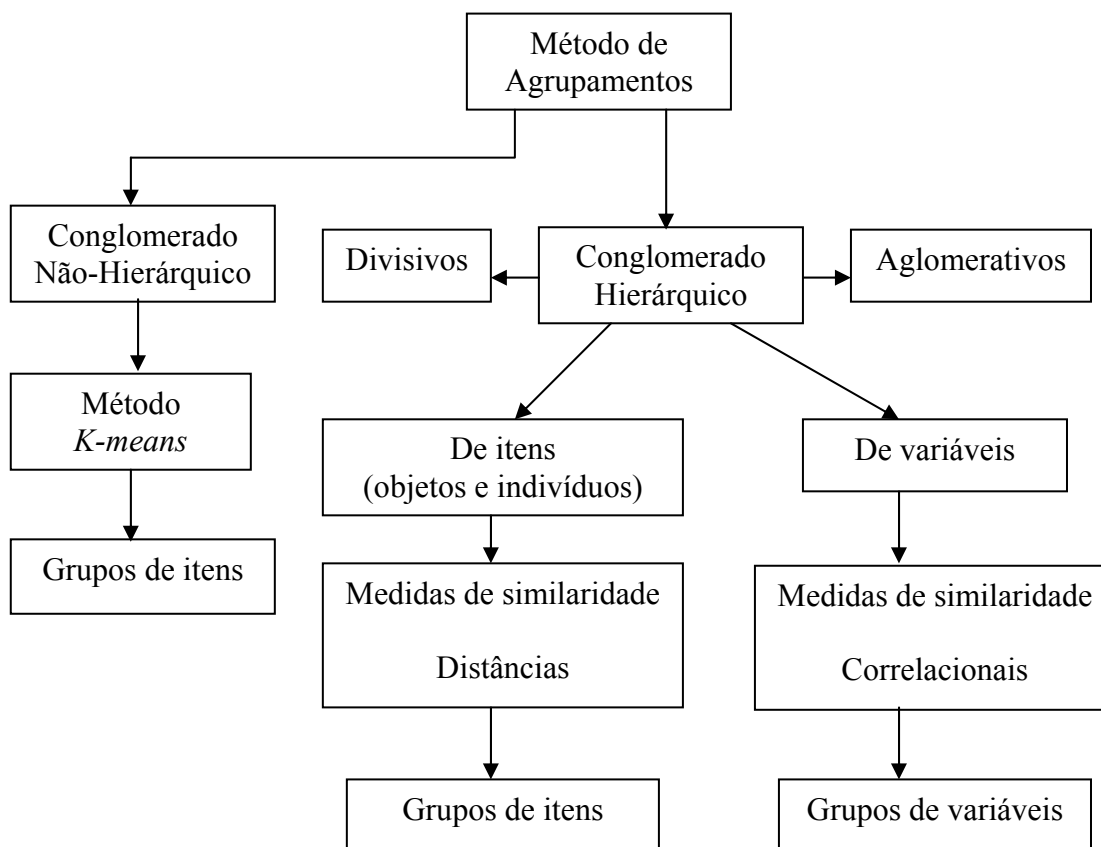


Figura 4.2. Visão geral dos métodos de agrupamentos.

Para o agrupamento hierárquico, uma crítica é quanto a sua instabilidade quando uma primeira combinação de itens (ou variáveis) pode ter ocorrido com pequena diferença de critérios. Para tanto, aconselha-se uma divisão da amostra em pelo menos dois grupos, realizando o mesmo procedimento para verificar se são obtidos conglomerados similares.

Na análise hierárquica para variáveis, a matriz de correlação amostral, denotada por \mathbf{R} , tem fundamental importância, pois dará origem a matriz de distâncias, denotada por \mathbf{D} . Se as correlações, representando as similaridades, não forem significantes, este problema será transferido para suas respectivas distâncias. As junções entre as variáveis são obtidas a partir de diferenças de similaridades, logo as correlações devem ser estatisticamente distintas para que não sejam gerados grupos espúrios de variáveis. Devem-se então testar as hipóteses (REIS, 2001):

$$H_0: \boldsymbol{\rho} = \begin{bmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \cdots & 1 \end{bmatrix}$$

versus

$$H_1: \boldsymbol{\rho} \neq \begin{bmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \cdots & 1 \end{bmatrix}$$

na qual $\boldsymbol{\rho}$ é a matriz de correlação. Um teste é obtido utilizando-se somente os elementos fora da diagonal da matriz \mathbf{R} , usada para estimar $\boldsymbol{\rho}$. São calculadas as seguintes quantidades:

$$\bar{r}_k = \frac{1}{p-1} \sum_{\substack{i=1 \\ i \neq k}}^p r_{ik}, \text{ para } k = 1, 2, \dots, p; \quad (4.9)$$

$$\bar{r} = \frac{2}{p(p-1)} \left(\sum_{i < k} r_{ik} \right); \quad (4.10)$$

$$\hat{\gamma} = \frac{(p-1)^2 [1 - (1 - \bar{r})^2]}{p - (p-2)(1 - \bar{r})^2}. \quad (4.11)$$

Como se pode observar, \bar{r}_k é a média dos elementos fora da diagonal da k -ésima coluna de \mathbf{R} e \bar{r} é a média de todos os elementos fora da diagonal. A estatística de teste é dada por:

$$T = \frac{(n-1)}{(1 - \bar{r})^2} \left[\sum_{i < k} (r_{ik} - \bar{r})^2 - \bar{\gamma} \sum_{k=1}^p (\bar{r}_k - \bar{r})^2 \right] \quad (4.12)$$

sendo que a rejeição de H_0 ocorre quando $T > \chi_{(p+1)(p-2)/2}^2(\alpha)$, na qual $\chi_{(p+1)(p-2)/2}^2(\alpha)$ é o (100α) -ésimo percentil da distribuição Qui-quadrado com $(p+1)(p-2)/2$ graus de liberdade.

Outra questão importante é relativa a com quantos agrupamentos a análise final deve suportar. Vários critérios foram elaborados, contudo nenhum deles pode garantir com absoluta certeza o número ótimo de agrupamentos.

Um primeiro critério é avaliar a curva de decrescimento do nível de similaridade, conforme os grupos forem se formando. Quando o gráfico de similaridade com número de grupos tende a sofrer estabilização, existe indício de que deve-se parar de agrupar os elementos, pois daí em diante a diferença de similaridade entre os grupos pode ser insignificante e estes fariam a junção ao acaso.

Um outro critério bastante utilizado é realizar uma análise fatorial com o objetivo de redução do número de fatores (itens ou variáveis) sem perda de representatividade (ou de informação). Na análise fatorial os valores dos autovalores (vide apêndice D) de cada eixo (ou fator) representam o quanto da variância amostral total (ou da informação total) neles ficaram retidas, ou seja, o quanto da variância original seu respectivo fator explica. A variância total original, por sua vez, pode ser obtida somando-se todos os autovalores. Estes podem então ser somados do maior para o menor até que se atinja um nível satisfatório de informação, desprezando os menores, conseqüentemente, reduzindo a dimensão e, possivelmente, obtendo uma estrutura mais fácil de ser explicada. Pode ser utilizada tanto a matriz amostral de covariâncias, como a matriz amostral de correlações, para obter-se os autovalores. Desta forma, a proporção da variância amostral total explicada pelo j -ésimo fator é dada por:

$$\frac{\hat{\nu}_j}{s_{11} + s_{22} + \dots + s_{pp}}, \text{ para análise fatorial de } \mathbf{S}; \quad (4.13)$$

$$\frac{\hat{\nu}_j}{p}, \text{ para análise fatorial de } \mathbf{R}, \quad (4.14)$$

sendo que $\hat{\nu}_j$ é o j -ésimo autovalor (vide apêndice D) ordenado de forma decrescente e obtido de uma matriz amostral de covariâncias \mathbf{S} ou a matriz de correlação amostral \mathbf{R} .

Observa-se, para a análise fatorial de \mathbf{R} , que $\sum_{j=1}^p \hat{\nu}_j = p$, ou seja, em média cada autovalor teria peso 1, contudo alguns têm valores maiores que 1 e outros não. Pensando assim, um critério denominado raiz latente (HAIR et al, 2005), determina que aqueles autovalores que são menores que 1 nem deveriam ser mantidos, pois sequer explicam a variabilidade de uma única variável.

Utilizando-se as expressões (4.13) ou (4.14), pode-se fazer um gráfico (conhecido como *scree plot*) do número de fatores com os autovalores em ordem decrescente, sendo que seu comportamento é utilizado para determinar o número de fatores a ser mantido, análogo ao gráfico de similaridades.

4.3.2 Análise das características do supermercado por meio de agrupamentos hierárquicos de variáveis

Inicialmente deve-se testar a estrutura presente nos elementos fora da diagonal da matriz \mathbf{R} , usada para estimar $\boldsymbol{\rho}$. Aplicando-se os dados coletados na amostra, a matriz amostral de correlações para preço (P), localização (L), rapidez (R), qualidade (Q), variedade (V), estacionamento (E) e outras (O) é dada por:

$$\mathbf{R} = \begin{bmatrix} 1 & & & & & & & \\ 0,064 & 1 & & & & & & \\ 0,061 & -0,101 & 1 & & & & & \\ -0,094 & -0,131 & 0,333 & 1 & & & & \\ 0,150 & 0,082 & 0,129 & -0,104 & 1 & & & \\ -0,024 & -0,061 & 0,050 & 0,057 & -0,076 & 1 & & \\ -0,184 & -0,184 & -0,038 & -0,214 & -0,133 & -0,002 & 1 & \end{bmatrix}.$$

São obtidas as quantidades abaixo, conforme as expressões (4.9) a (4.11) e a matriz \mathbf{R} :

$$\bar{r}_k = \begin{bmatrix} -0,004497 \\ -0,055120 \\ 0,072350 \\ -0,025260 \\ 0,007983 \\ -0,009178 \\ -0,125632 \end{bmatrix};$$

$$\bar{r} = -0,0199077;$$

$$\hat{\gamma} = -0,804706;$$

$$T = 107,102;$$

$$\chi_{(p+1)(p-2)/2}^2(\alpha) = \chi_{20}^2(0,005) = 40.$$

Logo, como $T > \chi_{20}^2(0,005)$, ocorre a rejeição de H_0 , ou seja, existe diferenças significativas nos elementos da diagonal da matriz \mathbf{R} e a análise de agrupamentos pode ser aplicada. Inicialmente, para efeito de ilustração do método, é utilizada como distância $d_{ij} = 1 - \rho_{ij}$, ligação média para todos os 303 questionários. A matriz \mathbf{D} é:

$$\mathbf{D} = \begin{matrix} & \text{P} & \text{L} & \text{R} & \text{Q} & \text{V} & \text{E} & \text{O} \\ \begin{matrix} 0 \\ 0,936 & 0 \\ 0,939 & 1,101 & 0 \\ 1,094 & 1,131 & \mathbf{0,667} & 0 \\ 0,850 & 0,918 & 0,872 & 1,104 & 0 \\ 1,024 & 1,061 & 0,950 & 0,943 & 1,076 & 0 \\ 1,184 & 1,184 & 1,038 & 1,214 & 1,133 & 1,002 & 0 \end{matrix} \end{matrix}.$$

O menor valor das distâncias (maior similaridade) é 0,667, que liga rapidez (coluna 3) e qualidade (coluna 4). Logo, compondo as colunas 3 e 4 pela média, a nova matriz de distâncias fica

$$\mathbf{D} = \begin{matrix} & \text{P} & \text{L} & \text{R,Q} & \text{V} & \text{E} & \text{O} \\ \begin{matrix} 0 \\ 0,936 & 0 \\ 1,016 & 1,116 & 0 \\ \mathbf{0,850} & 0,918 & 0,988 & 0 \\ 1,024 & 1,061 & 0,946 & 1,076 & 0 \\ 1,184 & 1,184 & 1,126 & 1,133 & 1,002 & 0 \end{matrix} \end{matrix}$$

A nova junção é realizada entre as colunas 1 e 5, preço e variedade, cuja distância é 0,850. Como exemplo, a nova distância entre o agrupamento preço e variedade com o agrupamento rapidez e qualidade fica:

$$d_{(1,5)(3,4)} = \frac{d_{13} + d_{14} + d_{53} + d_{54}}{4} \approx 1,002.$$

A nova matriz \mathbf{D} fica:

$$\begin{array}{ccccc}
 & P,V & L & R,Q & E & O \\
 \mathbf{D} = & \begin{bmatrix} 0 & & & & & \\ \mathbf{0,927} & 0 & & & & \\ 1,002 & 1,116 & 0 & & & \\ 1,050 & 1,061 & 0,946 & 0 & & \\ 1,158 & 1,184 & 1,126 & 1,002 & 0 & \end{bmatrix} & .
 \end{array}$$

Devido a menor distância encontrada (0,927), a nova junção é dada por preço e variedade (coluna 1,5) com localização (coluna 2). Repetindo o processo é obtida:

$$\begin{array}{ccccc}
 & P,V,L & R,Q & E & O \\
 \mathbf{D} = & \begin{bmatrix} 0 & & & & \\ 1,040 & 0 & & & \\ 1,054 & \mathbf{0,946} & 0 & & \\ 1,667 & 1,126 & 1,002 & 0 & \end{bmatrix} & .
 \end{array}$$

Realizando a junção 3,4 e 6, a nova matriz fica:

$$\begin{array}{ccccc}
 & P,V,L & R,Q,E & O \\
 \mathbf{D} = & \begin{bmatrix} 0 & & & \\ \mathbf{1,044} & 0 & & \\ 1,167 & 1,126 & 0 & \end{bmatrix} & .
 \end{array}$$

Finalmente,

$$\begin{array}{ccccc}
 & P,V,L,R,Q,E & O \\
 \mathbf{D} = & \begin{bmatrix} 0 & \\ 1,126 & 0 \end{bmatrix} & .
 \end{array}$$

A figura 4.3, juntamente com a tabela 4.2, pode representar muito bem as junções obtidas com seus níveis de similaridade, utilizando correlação e ligação média.

Como exemplo, a primeira junção realizada na matriz **D** é rapidez com qualidade, exatamente como ocorre na figura 4.3. O nível de distância na tabela 4.2 é obtido com os menores valores das matrizes de distância no decorrer do processo de agrupamento, sendo o nível inicial zero. O nível de similaridade aqui exibido é obtido fazendo-se a interpolação de seu nível inicial sem junções (100) e o nível de distância correspondente.

Pela figura 4.3 e pela tabela 4.2 pode-se observar que algumas junções ocorrem em um nível de similaridade muito baixo (parte de cima da figura 4.3) e estas junções não devem ser consideradas. Para melhor avaliação, é confeccionado um gráfico dos níveis de similaridade com o número de agrupamentos.

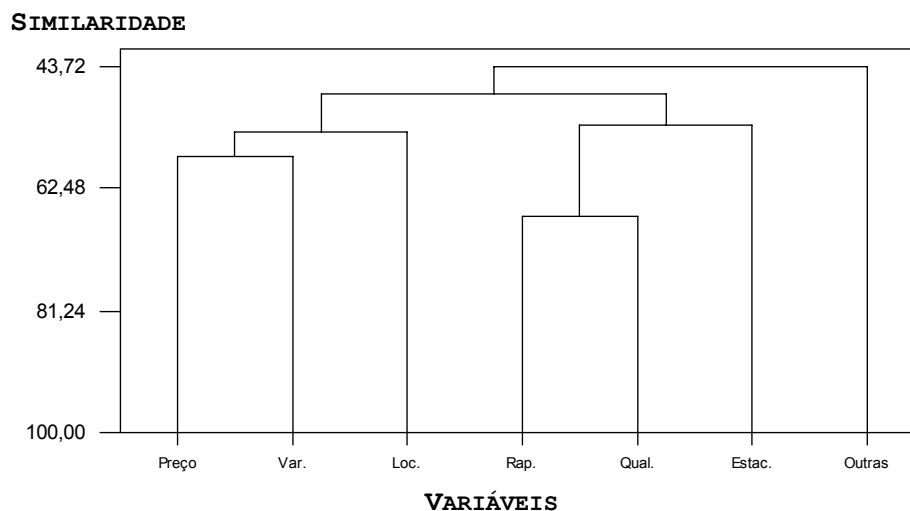


Figura 4.3. Representação gráfica dos agrupamentos utilizando correlação simples e ligação média para todos os questionários.

Tabela 4.2. Níveis de similaridade e distância.

Passo	Nível de Similaridade	Nível de Distância
1	66,67	0,667
2	57,50	0,850
3	53,64	0,927
4	52,68	0,946
5	47,78	1,044
6	43,72	1,126

Pela figura 4.4 a seguir percebe-se que o nível de similaridade decresce muito com menos de 3 grupos (agrupamentos), o que não deve ser considerado. Observa-se também que há um grande declive logo na primeira junção (de 7 grupos para 6), referentes a rapidez e qualidade. As junções devem deixar de ser realizadas quando o nível de similaridade estiver muito baixo, por exemplo, quando houver três grupos. Contudo, deve-se considerar uma análise mais elaborada com autovalores para decidir quantos agrupamentos formar.

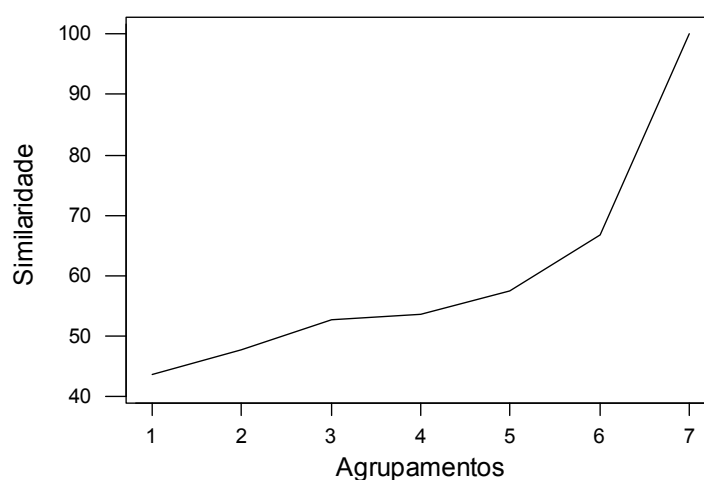


Figura 4.4. Representação gráfica dos agrupamentos e similaridades, utilizando correlação simples e ligação média, para todos os questionários.

Para melhorar a análise de agrupamentos, pode-se proceder uma comparação assumindo a ligação média como critério para todas as análises, visando evitar os problemas que podem ocorrer com a ligação simples e a ligação completa, como comentado anteriormente. Assim, utilizando a divisão da amostra total em duas amostras, relativas a primeira semana de coleta (157 questionários) e a segunda semana (146 questionários) e, variando também o tipo de correlação, sendo aplicadas a correlação simples e a correlação transformada (expressão (4.8)), que obedecem as propriedades métricas, pode-se verificar a consistência da análise final obtida.

Para iniciar a comparação, é realizada a mesma análise que a anterior com a correlação transformada. Pode-se observar pela figura 4.5 que preço, variedade e localização estão ligados, bem como rapidez, qualidade e estacionamento, ficando outras como a última ligação a ser obtida. Nota-se, contudo, pela figura, que o nível de

similaridade com que as ligações foram ocorrendo estão abaixo das obtidas com a análise anterior, na qual utilizava-se correlação simples. Isto provavelmente deve ter ocorrido como efeito realizado pela transformação, a qual a correlação original foi submetida.

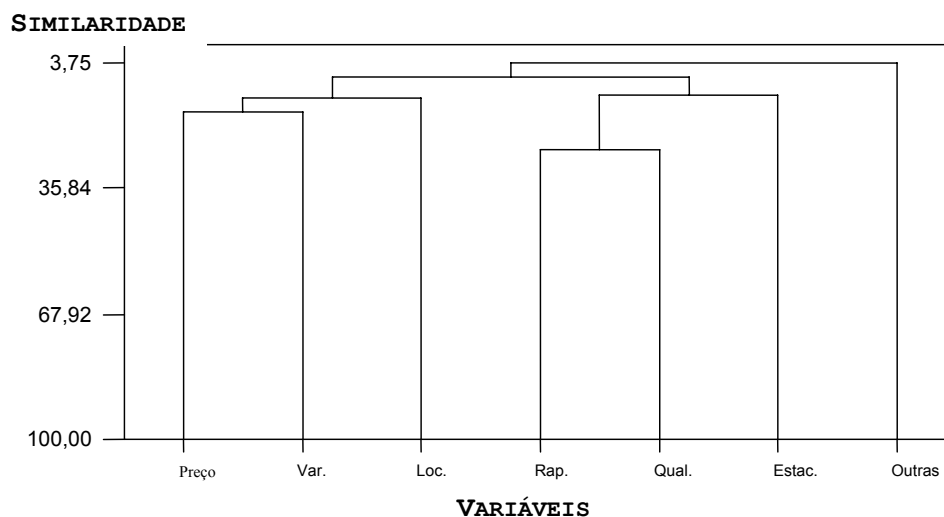


Figura 4.5. Representação gráfica dos agrupamentos utilizando correlação transformada e ligação média para todos os questionários.

A seguir, a figura 4.6, utilizando-se correlação simples e ligação média para a primeira semana de coleta de questionários. Também, as figuras 4.7, 4.8 e 4.9 representam, respectivamente, análise com correlação transformada e ligação média para a primeira semana de coleta de questionários, correlação simples e ligação média para a segunda semana de coleta de questionários e, por último, correlação transformada e ligação média para a segunda semana de coleta de questionários, respectivamente.

Comparando-se as análises, verifica-se que, com exceção da figura 4.7, os agrupamentos formados basicamente são compostos por preço, variedade e localização com rapidez, qualidade e estacionamento, mantendo-se isolado outras. Na figura 4.7 observa-se que a ligação ocorrida entre o agrupamento preço e variedade com outro agrupamento em vez de localização ocorre com pouquíssima diferença de similaridade. Desta forma, pode-se dizer que, apesar disto, o agrupamento apontado pelas outras análises pode ser considerado consistente com a primeira delas, da figura 4.3.

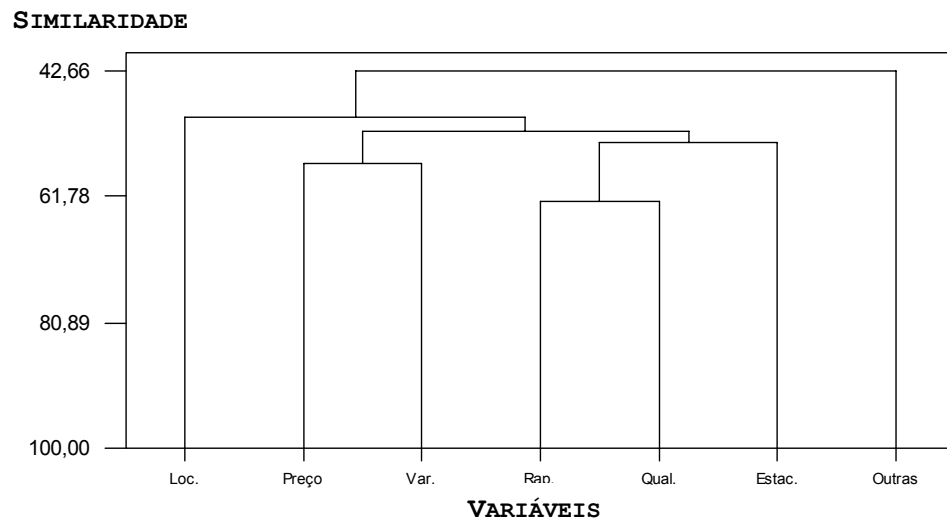


Figura 4.6. Representação gráfica dos agrupamentos utilizando-se correlação simples e ligação média, para a primeira semana de coleta.

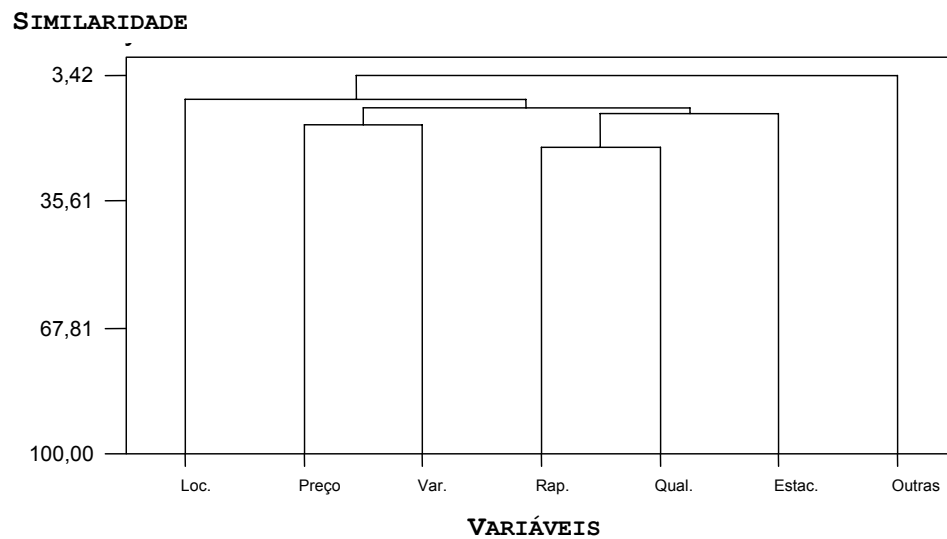


Figura 4.7. Representação gráfica dos agrupamentos utilizando-se correlação transformada e ligação média, para a primeira semana de coleta.

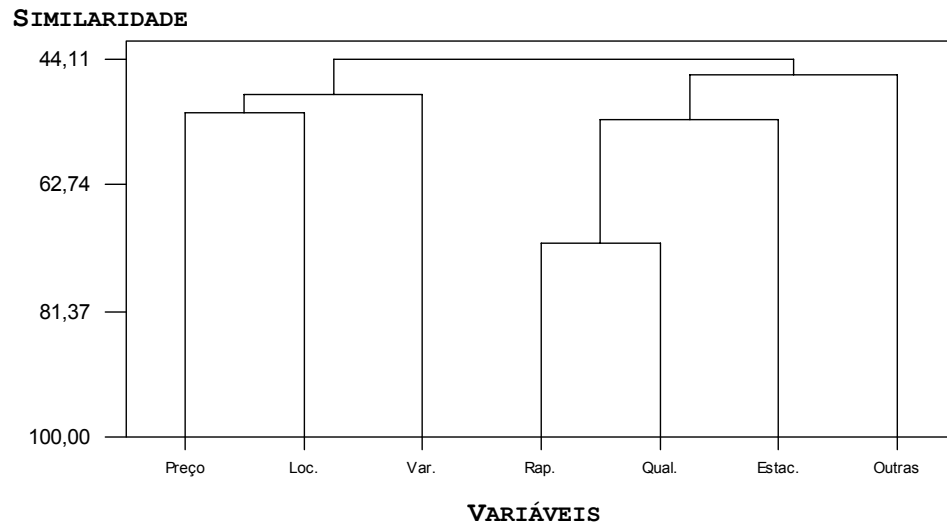


Figura 4.8. Representação gráfica dos agrupamentos utilizando-se correlação simples e ligação média, para a segunda semana de coleta.

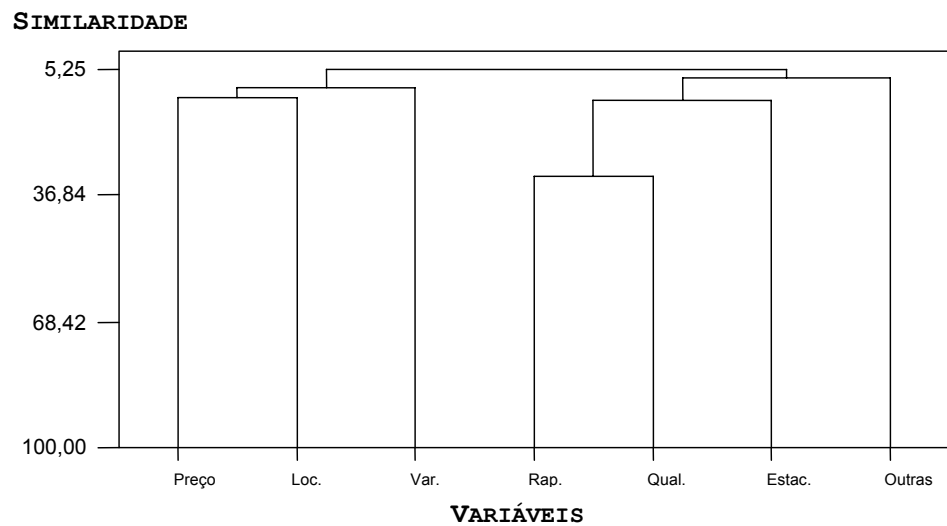


Figura 4.9. Representação gráfica dos agrupamentos utilizando-se correlação transformada e ligação média, para a segunda semana de coleta.

Apesar da análise realizada entre similaridades e número de agrupamentos, deve-se proceder também uma análise fatorial (AAKER et al., 2004), utilizando-se os autovalores na determinação do número de fatores a ser mantido no final. Também, não deve ser descartado o entendimento que os agrupamentos fornecem, sua possível explicação no que se refere a maior facilidade de análise da situação e da informação que se pode conseguir. Desta forma, não se deve simplesmente olhar valores de autovalores, mas também haver coerência no número de fatores a ser assumido com a representação obtida das informações dos dados.

A figura 4.10 representa o *scree plot* da análise fatorial utilizando-se, para as sete variáveis, de seus autovalores ordenados em forma decrescente, com matriz de correlação e sem rotação. Observa-se que 3 como número de fatores pode ser considerado satisfatório, pois há um grande decréscimo aí e, ainda, os fatores restantes têm autovalores menores que um, ou seja, tem pouca absorção de informação (ou de variabilidade), podendo então ser descartados por ter menor importância. Também, acaba sendo propiciada a análise até então explicitada pelas análises anteriores, ou seja, o agrupamento preço, variedade e localização, o agrupamento rapidez, qualidade e estacionamento, e outras isoladamente. Deve-se ressaltar ainda que não havia outra estrutura de agrupamentos concorrente a esta se fossem considerados outros números de fatores, o que acaba corroborando ainda mais a análise final obtida.

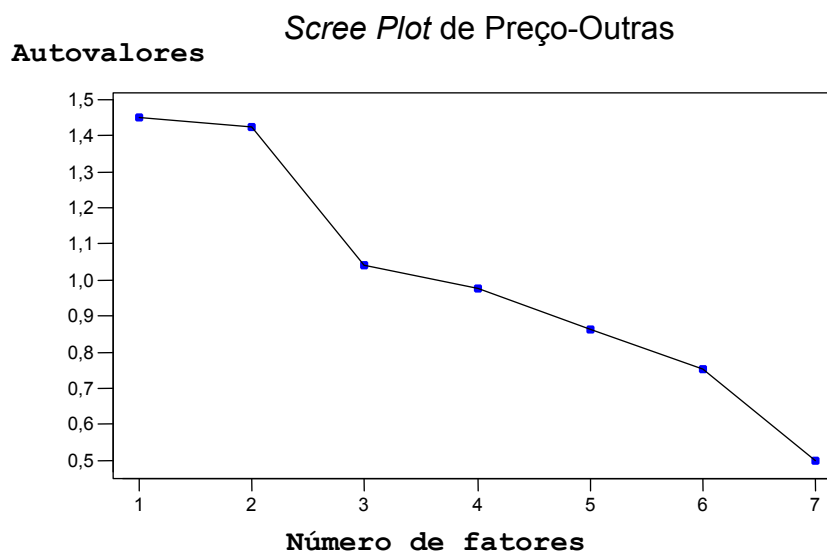


Figura 4.10. *Scree plot* dos autovalores para preço, variedade, rapidez, qualidade, localização, estacionamento e outras, utilizando análise fatorial, sem rotação e com matriz de correlação.

Analisando-se ainda a tabela 4.3, pode-se observar que o total da variância explicada pelos fatores obtidos não é alto, mas, como comentado, a análise de agrupamentos parece ser bem consistente. Segundo Hair et al. (2005), para este tipo de informação, em torno de 0,6, pode ser considerado satisfatório.

Tabela 4.3. Variabilidade total explicada pelos 3 fatores.

	Fator 1	Fator 2	Fator 3	Total
Variância	1,4496	1,4238	1,0410	3,9144
% da variância	0,207	0,203	0,149	0,559

De forma geral, pode-se notar que ficaram estabelecidos dois grupos bastante interessantes: no primeiro, rapidez, qualidade e estacionamento e em outro preço, variedade e localização. Nota-se que um deles privilegia atendimento enquanto o outro basicamente privilegia preço em mercadorias variadas, portanto bastante distintos.

4.4 O Número de itens na compra, a percepção de diferença de preços e a escolaridade dos clientes entrevistados.

Outra questão importante é referente ao número de itens que os clientes costumam comprar no supermercado. Este dado pode ser utilizado para dimensionar o tamanho dos carrinhos para melhor atender os clientes. Na figura 4.11 está a distribuição do número de itens para todos os dias de coleta. Observa-se que existe um decaimento praticamente exponencial, conforme o número de itens vai crescendo. Uma outra forma de analisar os dados pode ser definindo classes em função do número de itens, por exemplo, classe A para quem adquire até 15 itens, B até 70 itens e C acima disto, conforme figura 4.12.

Esta divisão de classes é utilizada posteriormente em outras análises. Contudo, observa-se que se alocados caixas que atendem até 15 itens, estes são responsáveis por praticamente dois terços de todas as compras realizadas no supermercado, enquanto cerca de 11% compram acima de 70 itens. Poderia-se estudar então a implementação de três tipos de caixas, relativos as classes A (até 15 itens), B (de 16 a 70 itens) e C (acima de 70 itens), proporcional ao número total de caixas e a frequência encontrada nas classes. O supermercado em questão tem três tipos de

carrinhos e poderia haver um tratamento diferenciado aos clientes que realizam compras grandes, sem incomodar aqueles com compras menores.

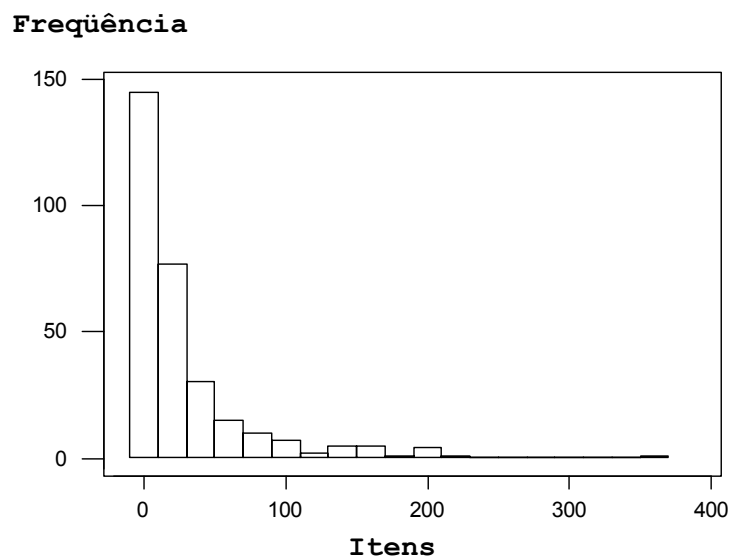


Figura 4.11. Representação gráfica do número de itens por cliente.

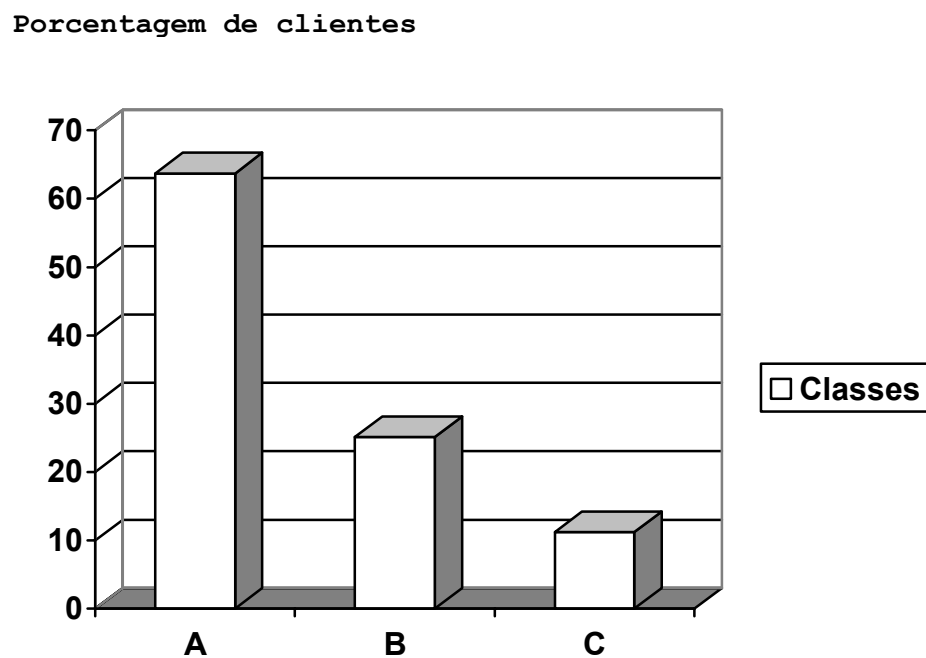


Figura 4.12. Representação gráfica das classes para o número de itens por cliente.

Quanto a questão relativa à impressão do cliente em haver diferença de preços entre supermercados distintos, dos entrevistados, 29,7% disseram não acreditar que exista e a maioria, 66,7%, afirma que há diferença entre supermercados. Um dado importante foi obtido com 11 clientes (3,6%) que não sabiam responder. A razão apontada foi que não faziam compras em outros estabelecimentos, sendo fiéis ao supermercado analisado.

A escolaridade dos clientes é uma informação bastante importante, pois está relacionada à renda e, conseqüentemente, ao poder de compra (como é visto posteriormente). Espera-se que, para diferentes níveis de escolaridade, haja comportamentos e necessidades distintas, portanto, quantificar esta informação, ainda que de forma amostral, é importante.

Nos seis dias de coleta foram entrevistados 307 clientes, sendo que apenas quatro não tinham escolaridade alguma. Estes, até porque não tiveram segurança em responder as questões, foram retirados das análises estatísticas realizadas, sendo então utilizadas 303 respostas. Pela figura 4.13 pode-se verificar as porcentagens para as seguintes categorias de escolaridade: sem escolaridade (s/ escol.), ensino fundamental incompleto (E.F.I.), ensino fundamental completo (E.F.C.), ensino médio incompleto (E.M.I.), ensino médio completo (E.M.C.), ensino superior incompleto (E.S.I.) e ensino superior completo (E.S.C.). Observa-se que há pouca porcentagem para ensino médio e superior incompletos, diferente do ensino fundamental incompleto.

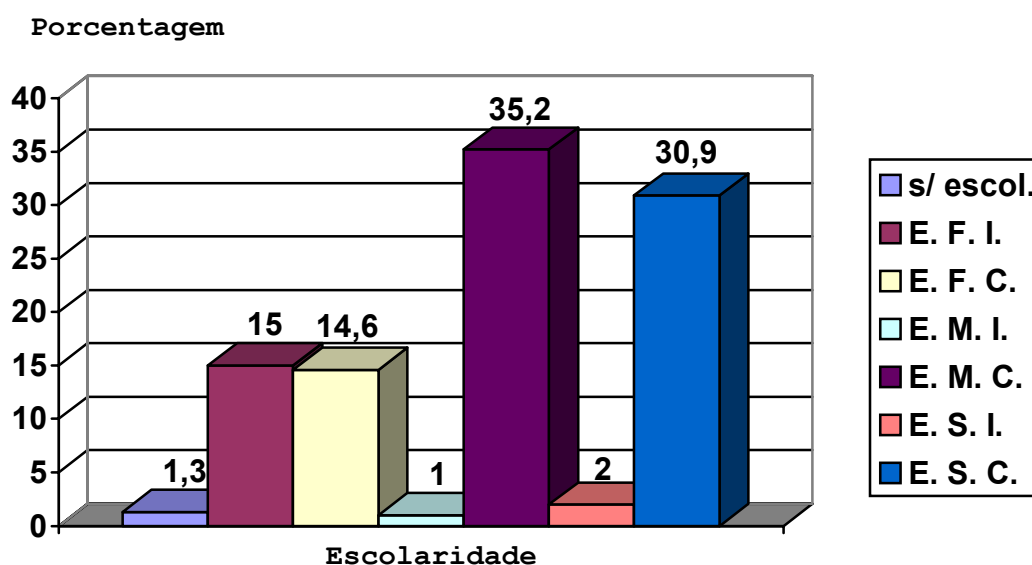


Figura 4.13. Representação gráfica dos diferentes níveis de escolaridade para os dias de coleta.

Para evitar categorias com pouca representatividade, resolveu-se juntar os ensinos completos e incompletos, ficando com as categorias sem escolaridade (s/ escol.), ensino fundamental (E.F.), ensino médio (E.M.) e ensino superior (E.S.), cuja representação gráfica pode ser vista na figura 4.14. Ressalta-se ainda que a junção encontra justificativa devido ao fato de, somente no ensino fundamental, haver equilíbrio entre os clientes que o completaram e aqueles que não têm esta escolaridade completa.

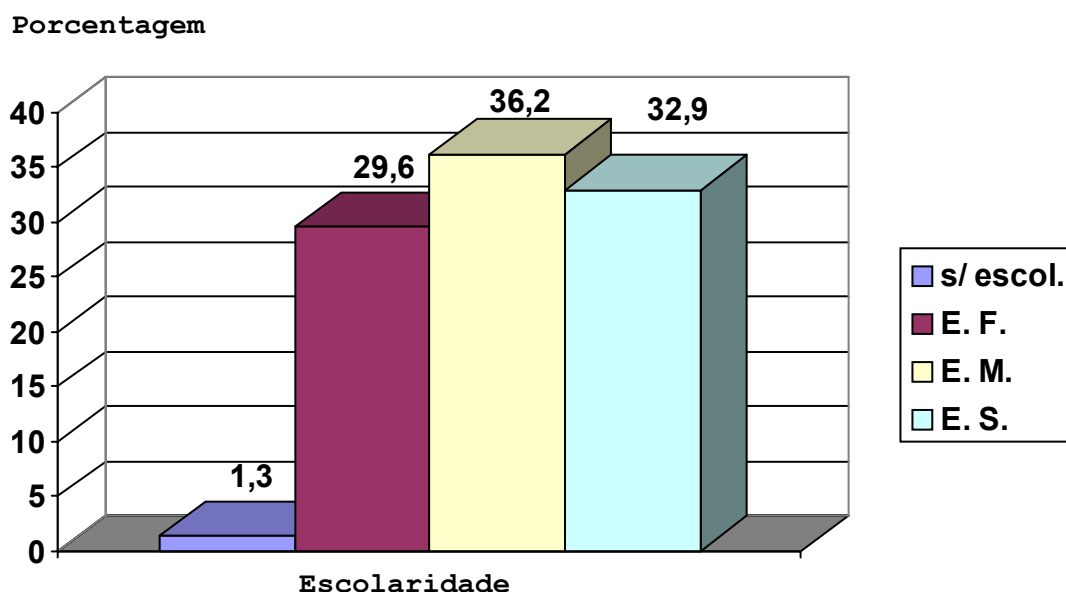


Figura 4.14. Representação gráfica para a junção dos diferentes níveis de escolaridade, para os dias de coleta.

Segundo o IBGE, a cidade de São Carlos, para o censo de 2.000, tinha 5,8% de sua população sem escolaridade, 43,1% com ensino fundamental (completo ou não), 20,3% com ensino médio (completo ou não) e 30,8% com mais ensino que as classes anteriores. Deve-se ressaltar que as diferenças obtidas na amostra podem ter duas hipóteses distintas: a primeira é que provavelmente deve-se à coleta de dados ter sido realizada quatro anos após o ano do censo, quando então as porcentagens deveriam estar mais coesas em razão da crescente escolaridade da população em geral. A segunda hipótese é que a clientela do supermercado pode naturalmente ser de uma parte da população que tem um nível de escolaridade mais alto que a média da cidade onde está o supermercado; nesta questão pode ser importante também a localização do estabelecimento, e a população que o circunda.

Segundo a PNAD (Pesquisa Nacional por Amostra de Domicílio) do IBGE, um fato interessante que ocorre no Brasil é que a renda média de um trabalhador está intimamente ligada a sua escolaridade, como revela a figura 4.15. Pessoas com escolaridades maiores tendem a ter rendas significativamente mais elevadas que aquelas com escolaridades mais baixas.

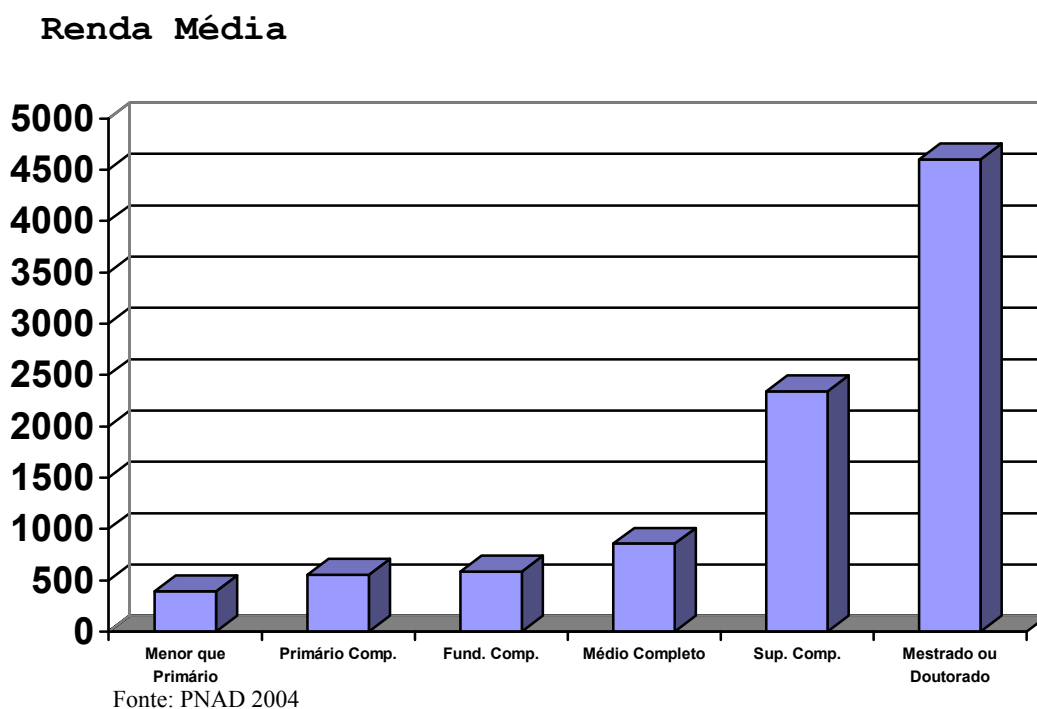


Figura 4.15. Renda média de um trabalhador associada ao nível de escolaridade para o Brasil em 2004.

Os níveis de escolaridade são importantes quando forem analisados seus cruzamentos com as características de frequência ao supermercado, com o número de itens comprados, entre outras. Para tanto, a análise de correspondência é empregada, conforme é visto na próxima seção. Para as análises subseqüentes, devido à maior facilidade de tratamento, os níveis de escolaridade são considerados como sendo a categoria Esc1 para o ensino fundamental (E. F.), a categoria Esc2 para o ensino médio (E. M.) e a categoria Esc3 para o ensino superior (E. S.).

4.5 O cruzamento das características pela análise de correspondência

Técnica exploratória desenvolvida pelos franceses, a análise de correspondência é freqüentemente utilizada para análise de cruzamentos entre variáveis. Resumidamente, é usualmente empregada para estudar as relações entre variáveis categóricas nominais, como também as categorias destas mesmas variáveis. Por meio dessa análise, pode-se reconhecer as relações entre as variáveis e/ou suas categorias, utilizando suas proximidades em um sistema de projeção plana. Estas proximidades são representadas por distâncias χ^2 projetadas em um plano e avaliadas segundo seu posicionamento, estabelecendo ou revelando possíveis associações. O plano resultante da análise é freqüentemente denominado de mapa de percepções.

O principal objetivo da análise de correspondência é a representação de dados categóricos em um espaço de dimensão menor identificando seus aspectos. As variáveis consideradas podem ser nominais e ordinais, com categorias mutuamente exclusivas ou não. Como produto final é obtido um gráfico, representando associações em uma tabela de freqüência ou contagem. A análise de correspondência simples pode ser encontrada em alguns livros, tais como Hair et al. (2005), Aaker et al (2004), Carvalho (2004), Pereira (2004), Johnson e Wichern (2002), entre outros.

4.5.1 A análise das características utilizando análise de correspondência simples

Nesta seção está contemplada a análise de correspondência simples, ou seja, obtida de uma tabela de dupla entrada referente ao cruzamento de duas variáveis. O gráfico vai então apresentar dois conjuntos de pontos: I pontos correspondendo às linhas e J pontos correspondendo às colunas. As associações podem ser obtidas por meio dos pontos (categorias) linha que estão próximas dos pontos (categorias) coluna, e que representam combinações que ocorrem com maior freqüência do que seria esperado por um modelo de independência, ou seja, no qual as categorias linha não estariam correlacionadas com as categorias coluna.

Como resultados da análise de correspondência simples, é obtida a melhor representação bidimensional dos dados e uma medida denominada inércia, que representa a quantidade de informação retida em cada dimensão.

Como exemplo, pode-se utilizar a característica preço (com sete categorias) realizando um cruzamento com a escolaridade (com três categorias). As categorias para preços devem ser entendidas como, por exemplo, P0 sendo a categoria na qual os clientes não apontaram esta característica. Já, para P7, os clientes apontaram em primeira ordem, para P6, em segunda ordem, e assim consecutivamente. Este procedimento também é seguido para todas as outras características. A tabela 4.4 de dupla entrada, representando essa situação, pode ser vista abaixo.

Tabela 4.4. Cruzamento entre as variáveis Preço e Escolaridade.

Preço	Escolaridade			Total
	Esc1	Esc2	Esc3	
P0	40	58	65	163
P2	0	0	1	1
P3	12	14	9	35
P4	3	3	8	14
P5	4	8	6	18
P6	7	9	9	25
P7	25	19	3	47
Total	91	111	101	303

Analisando-se a tabela 4.4 de forma simplificada, pode-se observar que a frequência, com a qual os clientes apontam preço como a primeira característica pela qual frequentam o supermercado (categoria P7), decresce conforme a escolaridade aumenta. Por outro lado, os clientes que simplesmente não apontaram preço (categoria P0) cresce conforme aumenta-se a escolaridade. A análise de correspondência então pode ser utilizada na tentativa de estabelecer o quão próximo as categorias dessas variáveis se encontram, ou seja, permite avaliar de forma mais fidedigna se a suposta relação pode ser considerada como autêntica.

A tabela de dupla entrada dá origem a matriz $\mathbf{X}_{(I \times J)}$ de frequências encontradas. Para o caso acima, $\mathbf{X}_{(I \times J)}$ é dada por:

$$\mathbf{X}_{(I \times J)} = \begin{bmatrix} 40 & 58 & 65 \\ 0 & 0 & 1 \\ 12 & 14 & 9 \\ 3 & 3 & 8 \\ 4 & 8 & 6 \\ 7 & 9 & 9 \\ 25 & 19 & 3 \end{bmatrix}.$$

Esta matriz dá início ao desenvolvimento algébrico da análise de correspondência.

4.5.2 Desenvolvimento algébrico

O desenvolvimento algébrico da análise de correspondência é importante, pois permite avaliar a forma com que as distâncias são utilizadas para as associações de categorias das variáveis. O tópico aqui presente está baseado em Johnson, Wichern (1998). Seja a matriz $\mathbf{X}_{(I \times J)}$, na qual x_{ij} são as frequências encontradas nas categorias de duas características distintas a serem pesquisadas. Sem perda de generalidade, é considerado que $I > J$ e que $\text{rank}(\mathbf{X}) = J$, ou seja, o número de linhas linearmente independentes de \mathbf{X} é J . Também, seja n o total de frequências em $\mathbf{X}_{(I \times J)}$. A matriz de correspondência, denotada por $\mathbf{P}_{(I \times J)} = \{p_{ij}\}$ (também conhecida por matriz de proporções), é a matriz $\mathbf{X}_{(I \times J)}$ dividida por n , ou seja, $p_{ij} = \frac{x_{ij}}{n}$, então:

$$\mathbf{P}_{(I \times J)} = \frac{1}{n} \mathbf{X}_{(I \times J)}.$$

Agora faz-se a centralização da matriz $\mathbf{P}_{(I \times J)}$, denotada por $\tilde{\mathbf{P}}$, e dada por:

$$\tilde{\mathbf{P}}_{(I \times J)} = \mathbf{P} - \mathbf{rc},$$

na qual $\mathbf{r} = \mathbf{P}_{(I \times J)} \mathbf{1}_{(J \times 1)}$; $\mathbf{c} = \mathbf{P}_{(J \times I)} \mathbf{1}_{(I \times 1)}$ e $\mathbf{1}$ é o vetor formado pelo número 1 com dimensão adequada. Desta forma, sendo $r_i = \sum_{j=1}^J p_{ij}$ para $i = 1, 2, \dots, I$, então $\mathbf{r}_{I \times 1}$ é o vetor soma dos elementos das linhas de $\mathbf{P}_{(I \times J)}$, e sendo $c_j = \sum_{i=1}^I p_{ij}$ para $j = 1, 2, \dots, J$, então $\mathbf{c}_{J \times 1}$ é o vetor soma dos elementos das colunas de $\mathbf{P}_{(I \times J)}$.

Observa-se então que a centralização é obtida fazendo-se a subtração em $\mathbf{P}_{(I \times J)}$ do produto do total de linhas e total de colunas para cada entrada de $\mathbf{X}_{(I \times J)}$.

Definem-se as matrizes diagonais:

$$\mathbf{D}_r = \text{diag}(r_1, r_2, \dots, r_I) \text{ e}$$

$$\mathbf{D}_c = \text{diag}(c_1, c_2, \dots, c_J).$$

Após a centralização, deve-se obter a padronização de $\mathbf{P}_{(I \times J)}$, cuja matriz é denotada por \mathbf{P}^* e dada por:

$$\mathbf{P}^*_{(I \times J)} = \mathbf{D}_r^{-1/2} \tilde{\mathbf{P}} \mathbf{D}_c^{-1/2}. \quad (4.15)$$

Desta forma, a (i, j) -ésima entrada da matriz \mathbf{P}^* é dada por $p_{ij}^* = \frac{p_{ij} - r_i c_j}{\sqrt{r_i c_j}}$.

Daqui percebe-se a centralização e a padronização efetuada na matriz $\mathbf{P}_{(I \times J)}$.

Havendo calculado a matriz \mathbf{P}^* , deve-se verificar a associação para a tabela de dupla entrada executando as fases seguintes. Inicialmente é encontrada a decomposição do valor singular de \mathbf{P}^* . Esta é dada por:

$$\mathbf{P}^*_{(I \times J)} = \mathbf{U}_{I \times (J-1)} \mathbf{\Lambda}_{(J-1) \times (J-1)} \mathbf{V}_{(J-1) \times J}, \quad (4.16)$$

na qual a matriz $\mathbf{U}_{I \times (J-1)}$ é formada pelos autovetores (vide apêndice D) de $\mathbf{P}^*_{(I \times J)} \mathbf{P}^*_{(J \times I)}$,

$\mathbf{V}_{J \times (J-1)}$ é formada pelos autovetores de $\mathbf{P}^*_{(J \times I)} \mathbf{P}^*_{(I \times J)}$ e $\mathbf{\Lambda}_{(J-1) \times (J-1)} = \text{diag}(\hat{v}_1, \hat{v}_2, \dots, \hat{v}_{J-1})$, ou seja, uma matriz diagonal formada pelos autovalores com $\hat{v}_1 > \hat{v}_2 > \dots > \hat{v}_{J-1}$, sendo que

$$\mathbf{U} \cdot \mathbf{U} = \mathbf{V} \cdot \mathbf{V} = \mathbf{I}.$$

Deve-se agora obter os eixos de coordenadas a serem utilizadas na representação gráfica para a análise de correspondência. Sejam:

$$\tilde{\mathbf{U}}_{I_x(J-1)} = \mathbf{D}_r^{1/2} \mathbf{U}$$

$$\tilde{\mathbf{V}}_{J_x(J-1)} = \mathbf{D}_c^{1/2} \mathbf{V}$$

Utilizando-se (4.15) e (4.16) é obtido que:

$$\tilde{\mathbf{P}}_{J_x J} = \mathbf{P} - \mathbf{r}\mathbf{c}' = \tilde{\mathbf{U}}\tilde{\mathbf{\Lambda}}\tilde{\mathbf{V}}' = \sum_{j=1}^{J-1} \hat{\nu}_j \tilde{\mathbf{u}}_j \tilde{\mathbf{v}}_j',$$

na qual $\tilde{\mathbf{u}}_j$ é o j -ésimo vetor coluna de $\tilde{\mathbf{U}}$ e

$\tilde{\mathbf{v}}_j$ é o j -ésimo vetor coluna de $\tilde{\mathbf{V}}$.

Da forma representada acima, pode-se verificar que:

$$\tilde{\mathbf{U}}' \mathbf{D}_r^{-1} \tilde{\mathbf{U}} = \tilde{\mathbf{V}}' \mathbf{D}_c^{-1} \tilde{\mathbf{V}} = \mathbf{I}_{(J-1) \times (J-1)}.$$

Agora, as colunas de $\tilde{\mathbf{U}}$ definem os eixos de coordenadas para os pontos representando os perfis coluna de \mathbf{P} , e as colunas de $\tilde{\mathbf{V}}$ definem os eixos de coordenadas para os pontos representando os perfis linha de \mathbf{P} .

Uma vez obtidos os eixos, deve-se calcular as coordenadas para os perfis linha e coluna. As coordenadas dos perfis linha, alocadas em $\mathbf{Y}_{I_x(J-1)}$, são obtidas como:

$$\mathbf{Y}_{I_x(J-1)} = \mathbf{D}_r^{-1} \tilde{\mathbf{U}}_{I_x(J-1)} \tilde{\mathbf{\Lambda}}_{(J-1)(J-1)}.$$

As coordenadas dos perfis coluna, alocadas em $\mathbf{Z}_{J_x(J-1)}$, são obtidas por:

$$\mathbf{Z}_{J_x(J-1)} = \mathbf{D}_c^{-1} \tilde{\mathbf{V}}_{J_x(J-1)} \tilde{\mathbf{\Lambda}}_{(J-1)(J-1)}.$$

As duas primeiras colunas de \mathbf{Y} contém os pares de coordenadas de pontos linha na melhor representação bidimensional dos dados e, da mesma forma, as duas primeiras colunas de \mathbf{Z} contém os pares de coordenadas de pontos coluna na melhor representação bidimensional dos dados.

O procedimento a ser realizado agora é colocar os dois conjuntos de pontos obtidos acima sobrepostos no mesmo gráfico. Ainda que não haja distância direta entre um ponto representando perfil coluna e um outro representando perfil linha, o gráfico obtido

pode ser interpretado como um cenário de aproximação (ou de correspondência) entre as categorias (ou variáveis) ali representadas.

Como uma última etapa, deve-se obter uma medida, denominada inércia, de grande importância para avaliar a análise de correspondência resultante. A inércia total é uma medida de variação dos dados e pode ser definida por:

$$\text{Inércia total} = \text{tr}(P^*P^{*\prime}) = \sum_{i=1}^k \hat{\nu}_i^2, \quad (4.17)$$

na qual $(\hat{\nu}_1 \geq \hat{\nu}_2 \geq \dots \geq \hat{\nu}_k > 0)$ são os valores não nulos da diagonal de $\Lambda_{(J-1) \times (J-1)}$. Desta forma, havendo k autovalores não nulos, então, também, $k = \text{rank}(\tilde{\mathbf{P}}) = \min(I-1, J-1)$.

As inércias associadas a alguma dimensão são definidas como o quadrado dos valores singulares (autovalores) não nulos, ou seja, $\hat{\nu}_i^2$. A interpretação dos valores de ν_i indicam (como é próprio de autovalores) quanta informação é retida em sua respectiva dimensão. A inércia associada com a melhor aproximação de *rank* reduzido $K_b < J$ tem inércia $\sum_{i=1}^{K_b} \nu_i^2$. Assim a inércia residual seria dada por $\sum_{i=K_b+1}^{J-1} \nu_i^2$ e uma boa aproximação teria este valor pequeno em relação a inércia total.

Nota-se que deve haver perda de informação quando $k \geq 3$, ou seja, as variáveis têm pelo menos quatro categorias. Apesar disto, deve-se avaliar quanto da informação é perdida, o que pode ser realizado pelas medidas de inércia, quando comparadas a inércia total verificando a viabilidade (confiança) de aplicação aos resultados obtidos. Desde que a inércia é uma medida da variação total dos dados, pode-se dizer que se $(\nu_1^2 + \nu_2^2) / \sum_{i=1}^{J-1} \nu_i^2$ tiver um valor alto, grande parte da informação estará representada em um plano, obtendo-se então do gráfico uma boa aproximação.

Das expressões (4.15) e (4.17) verifica-se que:

$$\text{tr}(P^*P^{*\prime}) = \text{tr}[(\mathbf{D}_r^{-1/2}(\mathbf{P} - \mathbf{rc}'))\mathbf{D}_c^{-1/2})(\mathbf{D}_r^{-1/2}(\mathbf{P} - \mathbf{rc}'))\mathbf{D}_c^{-1/2}'] =$$

$$= \sum_{i=1}^I \sum_{j=1}^J \frac{(p_{ij} - r_i c_j)^2}{r_i c_j}.$$

É interessante observar a relação entre a medida de associação qui-quadrado (χ^2) para uma tabela de dupla entrada. Esta é dada por:

$$\chi^2 = \sum_{i,j} \frac{(O_{ij} - E_{ij})^2}{E_{ij}},$$

na qual O_{ij} é a frequência observada e E_{ij} é a frequência esperada para ij -ésima célula. Adaptando-se esta situação, à análise de correspondência, pode-se fazer $O_{ij} = np_{ij}$ e $E_{ij} = nr_i c_j$, e assim:

$$\text{Inércia total} = \sum_{i=1}^I \sum_{j=1}^J \frac{(p_{ij} - r_i c_j)^2}{r_i c_j} = \chi^2 / n.$$

4.5.3 Resultados das análises de correspondência simples para as características

As análises de todas as características, apresentadas a seguir, foram consideradas para todos os 303 questionários, em todos os dias de coleta. Inicialmente, como um exemplo expondo os cálculos realizados, é retomada a análise da característica preço associada a escolaridade, apresentando vários resultados e o gráfico para interpretação das associações.

Em se tratando do cruzamento com a escolaridade, para as características em questão, não há perda de informação, pois $J-1=2$, ou seja, o gráfico contém toda a informação dos dados. Esta situação vai ocorrer em outras análises que são realizadas ainda neste capítulo.

A partir da tabela 4.4 foram calculadas as porcentagens para o cruzamento entre preço e escolaridade, apresentadas na tabela 4.5 a seguir.

Tabela 4.5. Porcentagem para o cruzamento entre Preço e Escolaridade.

Preço	Escolaridade			
	Esc1	Esc2	Esc3	Total
P0	13,20	19,14	21,45	53,80
P2	0	0	0,33	0,33
P3	3,96	4,62	2,97	11,55
P4	0,99	0,99	2,64	4,62
P5	1,32	2,64	1,98	5,94
P6	2,31	2,97	2,97	8,25
P7	8,25	6,27	0,99	15,51
Total	30,03	36,63	33,33	100,00

Observa-se da tabela 4.5 que mais da metade dos clientes não optou pela característica em questão, contudo existe uma informação importante quando se considera que não há uniformidade entre os clientes que o fizeram. Este fato é levantado pela análise de correspondência. A diferença entre as frequências observada e esperada pode ser dada por $np_{ij} - nr_i c_j$ para $I=1,\dots,7$ e $J=1,2,3$. Como exemplo, $np_{11} - nr_1 c_1 = 40 - 48,95 = -8,95$. Da tabela 4.6 observam-se valores não condizentes para o modelo de independência para as categorias P0 e P7, revelando uma suposta tendência de associação.

Tabela 4.6. Diferença entre as frequências observada e esperada para preço e escolaridade.

Preço	Escolaridade		
	Esc1	Esc2	Esc3
P0	-8,95	-1,71	10,67
P2	-0,30	-0,37	0,67
P3	1,49	1,18	-2,67
P4	-1,20	-2,13	3,33
P5	-1,41	1,41	0,00
P6	-0,51	-0,16	0,67
P7	10,88	1,78	-12,67

As distâncias χ_{ij}^2 (ou inércias das células) podem ser dadas por $\frac{(np_{ij} - nr_i c_j)^2}{nr_i c_j}$

para $I=1, \dots, 7$ e $J=1, 2, 3$. Como exemplo, $\chi_{11}^2 = \frac{(np_{11} - nr_1 c_1)^2}{nr_1 c_1} = \frac{(-8,95)^2}{48,95} \approx 1,64$. Percebe-

se pela tabela 4.7 que, pela ponderação ocorrida, P7 tem valores bastante interessantes em contraste à independência, ainda que só 15,51% dos clientes tenham apontado preço como principal característica. Isto ocorre pelo desbalanceamento em relação à escolaridade, e também pela soma de todos os valores, $\chi^2 = 29,894$.

Tabela 4.7. Distâncias Qui-Quadrado para Preço e Escolaridade.

Preço	Escolaridade			
	Esc1	Esc2	Esc3	Total
P0	1,638	0,049	2,094	3,781
P2	0,300	0,366	1,333	2,000
P3	0,211	0,108	0,610	0,929
P4	0,345	0,884	2,381	3,610
P5	0,366	0,300	0,000	0,665
P6	0,034	0,003	0,053	0,090
P7	8,393	0,184	10,241	18,819
Total	11,287	1,894	16,712	29,894

A tabela 4.8 traz as inércias relativas aos eixos (componentes) que irão compor o gráfico para análise. Ela revela que o componente 1 acumula quase toda a informação dos dados. A inércia total é obtida fazendo-se $\chi^2 / n \approx 0,0987$, dos quais 0,0929 (ou 94%) estão no componente 1. Esta construção se realiza sempre de forma decrescente, ou seja, as maiores inércias (ou autovalores) vão compondo os primeiros eixos, permitindo avaliar o quanto está contido no plano bidimensional. As tabelas 4.9 e 4.10, respectivamente das contribuições do perfil linha e das contribuições do perfil coluna, ajudam a interpretar os componentes obtidos. A tabela de contribuição do perfil linha é agora analisada considerando-se alguns conceitos.

Tabela 4.8. Inércias relativas aos componentes.

Componentes	Inércia	Proporção
1	0,0929	0,9417
2	0,0058	0,0583
Total	0,0987	1

Tabela 4.9. Contribuições do perfil linha para os componentes.

Preço	Inércia	Componente 1		Componente 2	
		Coordenadas	Contribuição	Coordenadas	Contribuição
P0	0,126	0,152	0,994	-0,011	0,006
P2	0,067	1,329	0,883	0,483	0,117
P3	0,031	-0,159	0,951	-0,036	0,049
P4	0,121	0,457	0,809	0,222	0,191
P5	0,022	0,066	0,117	-0,181	0,883
P6	0,003	0,060	1,000	0,000	0,000
P7	0,630	-0,630	0,991	0,060	0,009

A coluna inércia na tabela 4.9 é a proporção da inércia total contribuída por cada uma das linhas. Nota-se que P7 tem grande contribuição para o total. Observando-se a tabela 4.7, das distâncias Qui-Quadrado, vê-se que a contribuição de P7 é $\sum_{J=1}^3 \chi_{7,J}^2 / \chi^2 = 18,819/29,894 \approx 0,630$. As colunas de coordenadas são relativas aos componentes 1 e 2 a serem alocadas no gráfico bidimensional. São obtidas respectivamente por $\mathbf{Y}_{I_x(J-1)}$ e $\mathbf{Z}_{J_x(J-1)}$, como descrito anteriormente.

A coluna denominada contribuição refere-se a contribuição do componente para a inércia das linhas. Percebe-se desses valores porque 94% de toda informação está no componente 1. Somente P5 tem um valor alto (0,883) para o componente 2. Observa-se ainda que as categorias P0, P6 e P7 detêm 75,9% da inércia total do perfil linha.

A análise da tabela do perfil coluna (tabela 4.10) ocorre de forma análoga a anterior. Nota-se que as categorias Esc1 e Esc3 explicam grande parte da inércia total,

sendo que o componente 1 tem grande contribuição para a inércia destas duas categorias, ao contrário da categoria Esc2.

Tabela 4.10. Contribuições do perfil coluna para os componentes.

Escol.	Inércia	Componente 1		Componente 2	
		Coordenadas	Contribuição	Coordenadas	Contribuição
Esc1	0,378	-0,343	0,951	0,078	0,049
Esc2	0,063	-0,087	0,444	-0,097	0,556
Esc3	0,559	0,405	0,992	0,037	0,008

Por toda a análise do preço percebe-se que a categoria P2 (com frequência 1) não é importante e foi retirada para o resultado final, ou seja, para o gráfico bidimensional (figura 4.16) que, devido não haver perda de informações representa perfeitamente preço relacionado a escolaridade.

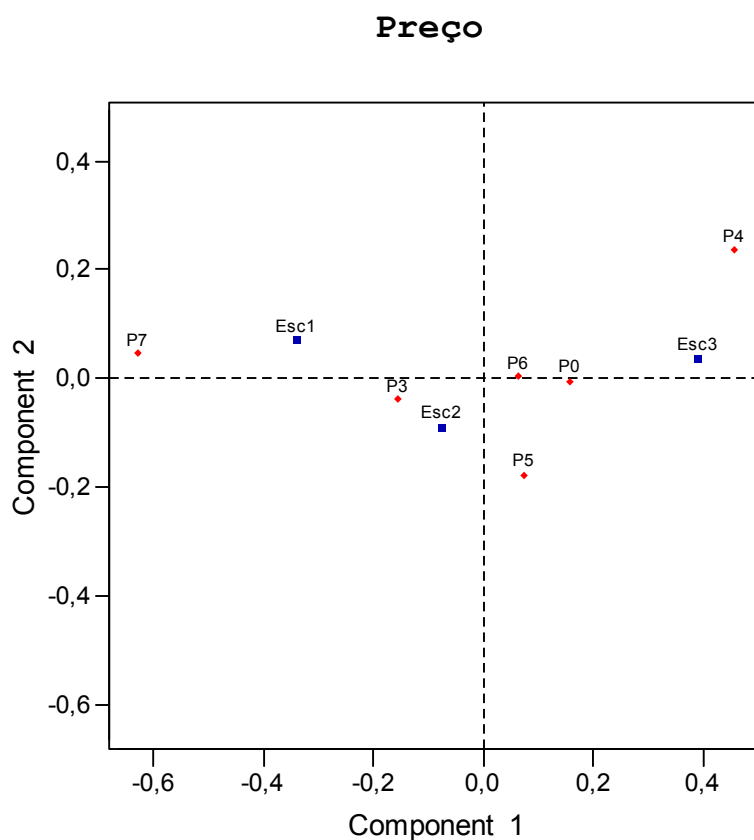


Figura 4.16. Análise de Correspondência Simples considerando-se Preço e Escolaridade.

Pela figura 4.16 pode-se dizer que a escolaridade 1 (Esc1) é a que mais se aproxima de P7, ou seja, para esta classe o preço deve ter grande importância nos motivos pelos quais frequenta o supermercado. Por outro lado, a escolaridade 3 (Esc3) se aproxima de P0 e P4, representando baixa importância para esta categoria, sendo ainda que escolaridade 2 (Esc2) está alocada em um plano intermediário (de transição). Desta forma, pode-se observar uma tendência inversa entre preço e escolaridade.

Para as outras características os resultados das análises estão sintetizados em suas respectivas tabelas e gráficos bidimensionais.

A próxima característica a ser analisada é a localização associada a escolaridade. A inércia total obtida foi $\chi^2 = 12,943$, sendo que 77% da informação se encontra no componente 1. Pela tabela 4.11 nota-se que a categoria L7 tem a maior contribuição para a inércia total do perfil linha, enquanto para o perfil coluna as inércias estão todas próximas. Observe-se ainda que 78,3% de toda a informação do perfil linha está nas categorias L7, L6 e L0.

Tabela 4.11. Contribuições dos perfis linha e coluna para Localização.

Perfis	Categorias	Inércias	Contribuição do Componente 1	Contribuição do Componente 2
L	L0	0,137	0,501	0,499
I	L3	0,060	1,000	0,000
N	L4	0,140	0,983	0,017
H	L5	0,017	0,743	0,257
A	L6	0,268	0,511	0,489
	L7	0,378	0,940	0,060
C	Esc1	0,300	0,135	0,865
O	Esc2	0,367	0,806	0,194
L.	Esc3	0,333	0,978	0,022

Pela figura 4.17 observa-se que a escolaridade 3 está muito próxima a L7 e L3, revelando haver considerável prioridade por esta característica. Ao contrário, a escolaridade 1 está próximo a L0 e a escolaridade 2 em uma posição intermediária. Desta forma, pode-se

observar uma tendência de aumento da prioridade da localização conforme existe um aumento de escolaridade. Deve-se observar ainda o distanciamento entre as categorias de escolaridade, revelando diferenças para seus posicionamentos em relação à localização.

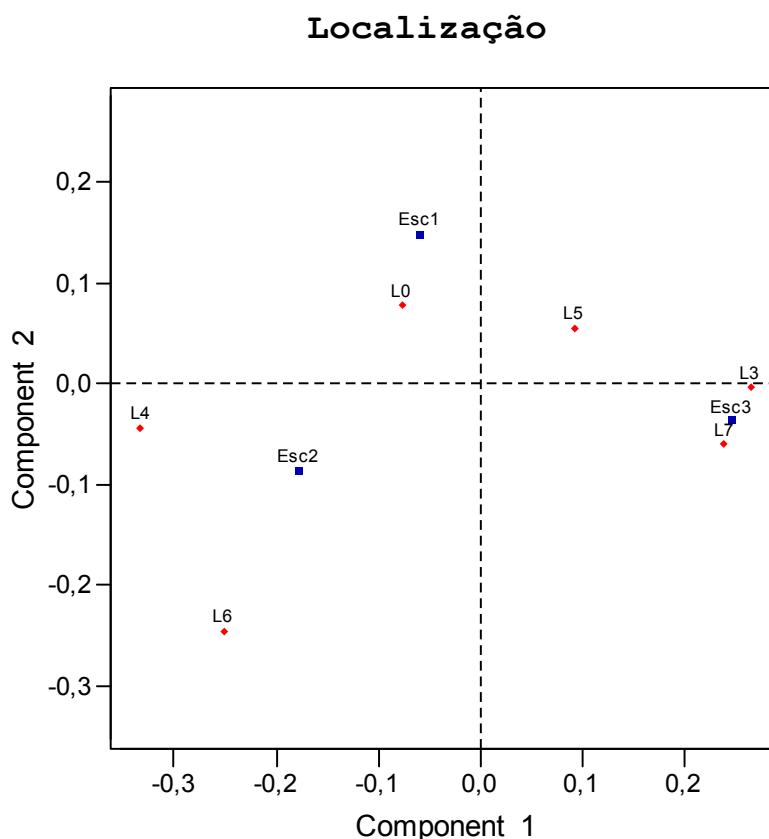


Figura 4.17. Análise de Correspondência Simples considerando-se Localização e Escolaridade.

A característica rapidez obteve inércia total $\chi^2 = 18,853$, sendo que 99,35% de toda informação está no componente 1. Pela tabela 4.12 observa-se que as categorias R7 tem a maior contribuição para a inércia total do perfil linha, enquanto no perfil coluna quase nenhuma informação está na categoria escolaridade 2. Finalmente, 95,1% de toda informação do perfil linha está contida nas categorias R7, R6 e R0.

Pela figura 4.18 nota-se alta prioridade de rapidez (categorias R6 e R7) para escolaridade 3. A escolaridade 1 apresenta baixa prioridade, pois está próxima a R0 e R4, sendo que a escolaridade 2 novamente se aloca em uma posição intermediária.

Tabela 4.12. Contribuições dos perfis linha e coluna para Rapidez.

Perfis	Categorias	Inércias	Contribuição do Componente 1	Contribuição do Componente 2
L	R0	0,336	0,997	0,003
I	R3	0,003	0,211	0,789
N	R4	0,034	1,000	0,000
H	R5	0,012	0,883	0,117
A	R6	0,211	0,998	0,002
	R7	0,404	0,997	0,003
C	Esc1	0,467	0,997	0,003
O	Esc2	0,009	0,552	0,448
L.	Esc3	0,524	0,998	0,002

Rapidez

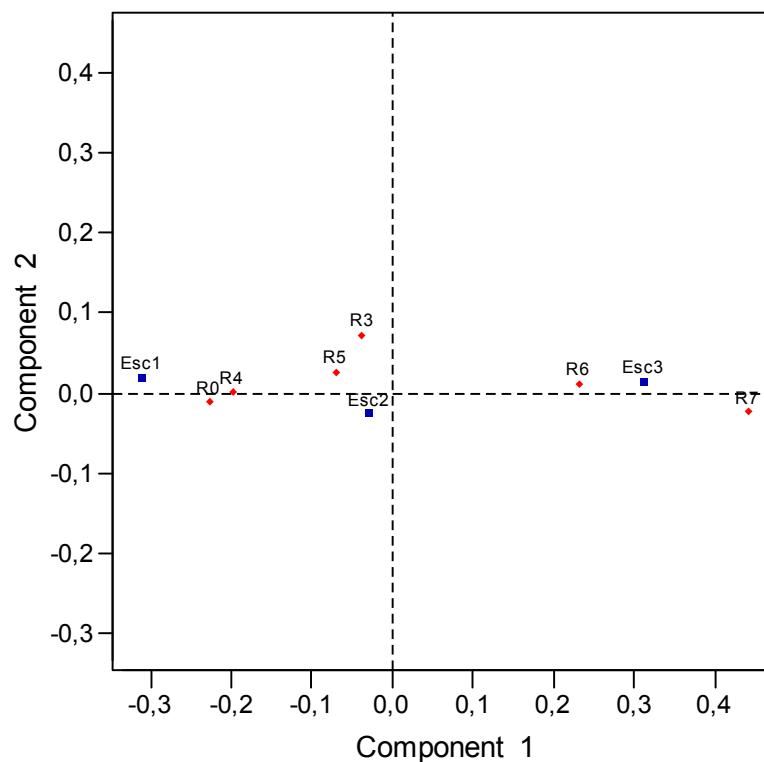


Figura 4.18. Análise de Correspondência Simples considerando-se Rapidez e Escolaridade.

Disposto isto, pode-se dizer que a prioridade de rapidez cresce conforme se aumenta a escolaridade. Observe-se ainda o isolamento de cada categoria de escolaridade com as citadas categorias de rapidez, evidenciando a associação entre elas.

A característica qualidade obteve inércia total $\chi^2 = 9,947$, sendo que 79,9% de toda informação está presente no componente 1. Pela tabela 4.13 pode-se notar que a categoria Q0 tem mais que a metade da contribuição para a inércia total do perfil linha, enquanto a escolaridade 3 tem mais que a metade da contribuição para a inércia total do perfil coluna. Similarmente às análises anteriores, as categorias Q0, Q6 e Q7 detêm juntas 75,8% de toda informação do perfil linha.

Tabela 4.13. Contribuições dos perfis linha e coluna para Qualidade.

Perfis	Categorias	Inércias	Contribuição do Componente 1	Contribuição do Componente 2
L	Q0	0,551	0,999	0,001
I	Q3	0,165	0,008	0,992
N	Q4	0,003	0,943	0,057
H	Q5	0,074	0,665	0,335
A	Q6	0,150	0,990	0,010
	Q7	0,057	0,815	0,185
C	Esc1	0,203	0,408	0,592
O	Esc2	0,267	0,701	0,299
L.	Esc3	0,530	0,998	0,002

Observa-se na figura 4.19 a seguir, que a categoria escolaridade 3 está próxima às categorias Q6, Q5 e Q7, ou seja, tem alta prioridade em qualidade. As categorias escolaridade 1 e escolaridade 2 estão mais próximas a Q0 e Q4, indicando baixa prioridade. Note-se, ainda, que a escolaridade 3 está bem distanciada de escolaridade 1 e escolaridade 2, parecendo atrair as três categorias mais altas de qualidade, ou seja, as categorias Q5, Q6 e Q7.

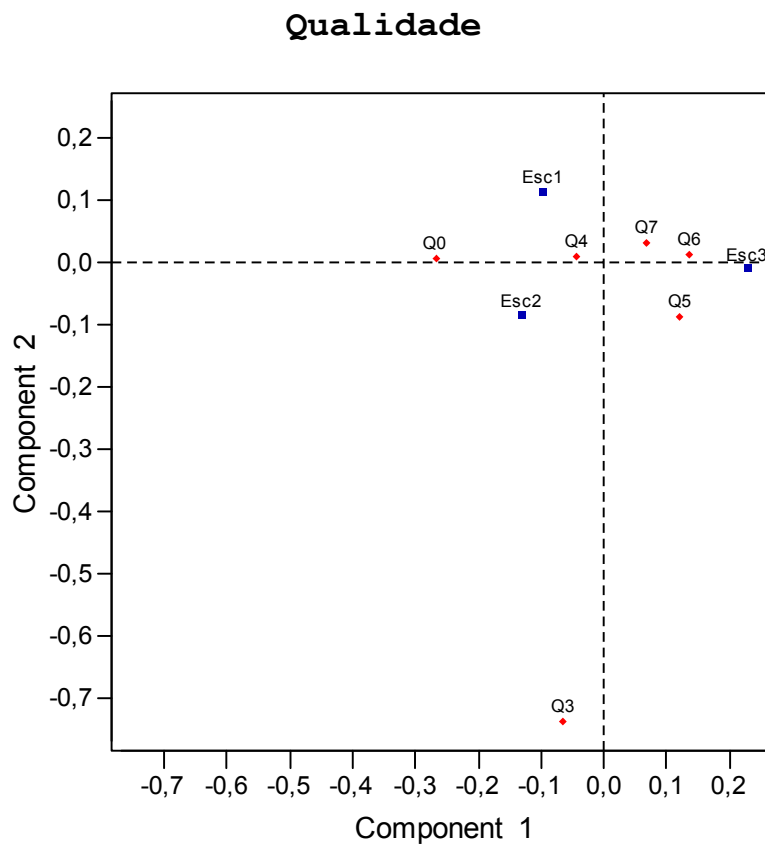


Figura 4.19. Análise de Correspondência Simples considerando-se Qualidade e Escolaridade.

A característica variedade obteve inércia total $\chi^2 = 7,830$, sendo que no componente 1 está 81,9% de toda informação. Observando-se a tabela 4.14, pode-se perceber que a categoria V7 tem praticamente dois terços da contribuição para a inércia total do perfil linha. Também, a categoria escolaridade 2 tem pouco mais da metade da contribuição para a inércia total do perfil coluna. Deve-se considerar ainda que as categorias V7, V6 e V0 juntas possuem 75,6% de toda a informação do perfil linha.

A análise gráfica da característica variedade (figura 4.20) revela que, para a categoria V7, a escolaridade mais próxima é a 2, ainda que distante. A escolaridade 3 está mais próxima a V5 e V6 e a escolaridade 1 mais próxima a V3 e V0. Pode-se pensar, de forma geral, que as categorias escolaridade 2 e escolaridade 3 priorizam mais a característica variedade que a escolaridade 1.

Tabela 4.14. Contribuições dos perfis linha e coluna para Variedade.

Perfis	Categorias	Inércias	Contribuição do Componente 1	Contribuição do Componente 2
L	V0	0,083	0,998	0,002
I	V2	0,110	0,327	0,673
N	V3	0,088	0,256	0,744
H	V4	0,022	0,528	0,472
A	V5	0,024	0,665	0,335
	V6	0,009	0,042	0,958
	V7	0,664	0,980	0,020
C	Esc1	0,310	0,759	0,241
O	Esc2	0,507	0,993	0,007
L.	Esc3	0,183	0,440	0,560

Variedade

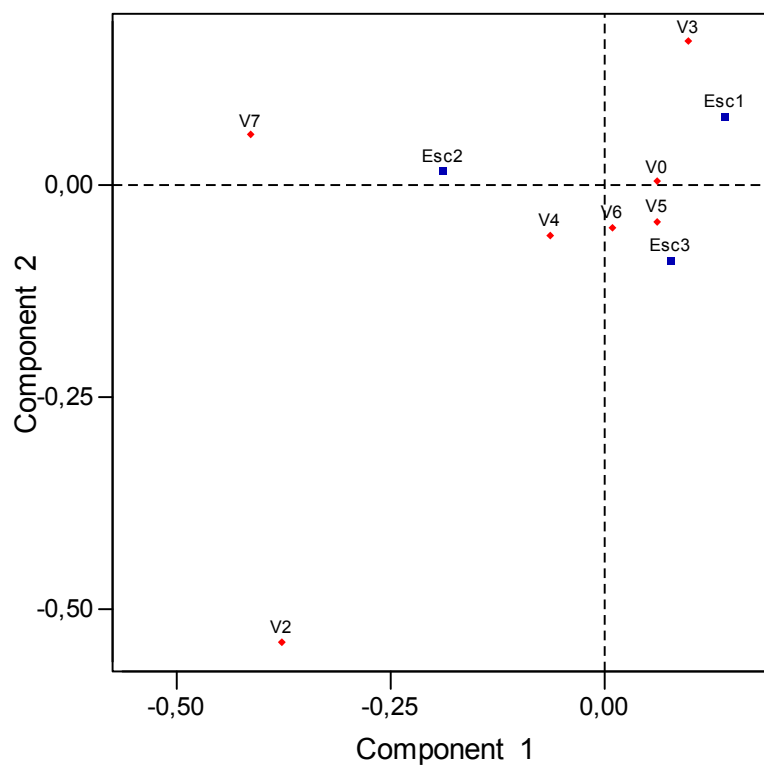


Figura 4.20. Análise de Correspondência Simples considerando-se Variedade e Escolaridade.

Para as categorias estacionamento ($\chi^2 = 11,472$) e outras ($\chi^2 = 16,788$) houve pouca frequência nos apontamentos realizados pelos clientes. Assim, as categorias E0 e O0 (tabelas 4.15 e 4.16) tiveram inércias praticamente iguais a zero. Deve-se observar

Tabela 4.15. Contribuições dos perfis linha e coluna para Estacionamento.

Perfis	Categorias	Inércias	Contribuição do Componente 1	Contribuição do Componente 2
L	E0	0,009	0,074	0,926
I	E2	0,045	0,500	0,500
N	E4	0,066	0,737	0,263
H	E5	0,299	0,906	0,094
A	E6	0,175	0,667	0,333
	E7	0,406	0,792	0,208
C	Esc1	0,465	0,931	0,069
O	Esc2	0,140	0,008	0,992
L.	Esc3	0,395	0,877	0,123

Tabela 4.16. Contribuições dos perfis linha e coluna para Outros.

Perfis	Categorias	Inércias	Contribuição do Componente 1	Contribuição do Componente 2
	O0	0,003	0,531	0,469
L	O2	0,061	0,959	0,041
I	O3	0,052	0,260	0,740
N	O4	0,070	0,371	0,629
H	O5	0,273	1,000	0,000
A	O6	0,273	1,000	0,000
	O7	0,268	1,000	0,000
C	Esc1	0,169	0,710	0,290
O	Esc2	0,222	0,832	0,168
L.	Esc3	0,609	1,000	0,000

por suas respectivas tabelas que as informações ficaram concentradas nas três primeiras categorias, E7, E6 e E5 com 88% e O7, O6 e O5 com 81,4%. As categorias estacionamento e outras tiveram respectivamente 78,1% e 91,4% de suas informações contidas no componente 1.

A análise gráfica para estacionamento e outras (figuras 4.21 e 4.22) revela que as categorias de escolaridade estão em torno do centro do gráfico, que está bem próximo de E0 e O0. Isto ocorre devido a baixa frequência de apontamentos para essas categorias, sendo as análises gráficas então, de forma geral, muito pouco informativas.

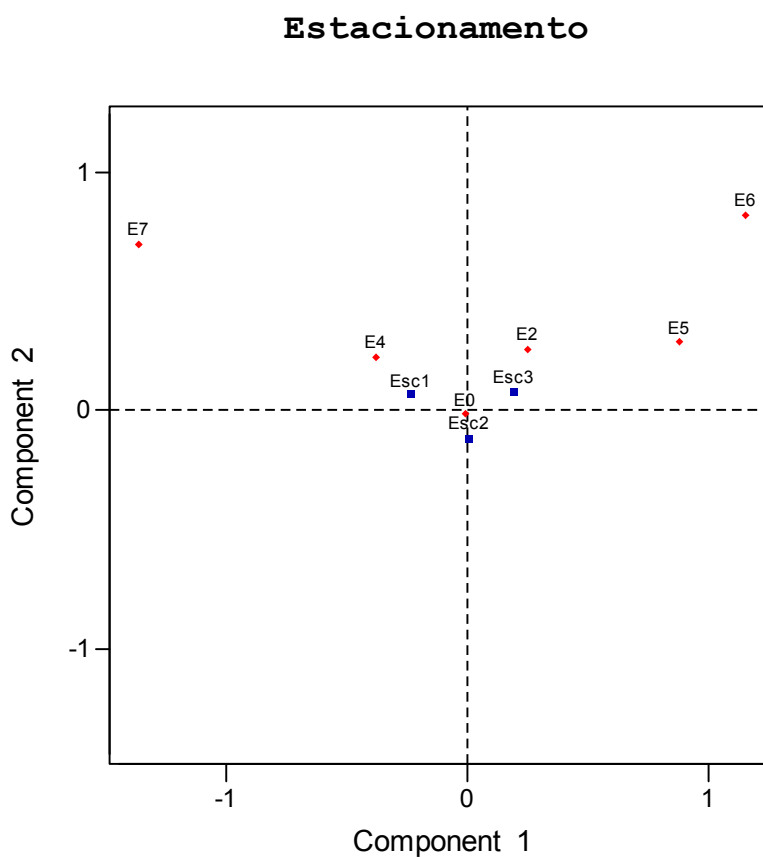


Figura 4.21. Análise de Correspondência Simples considerando-se Estacionamento e Escolaridade.

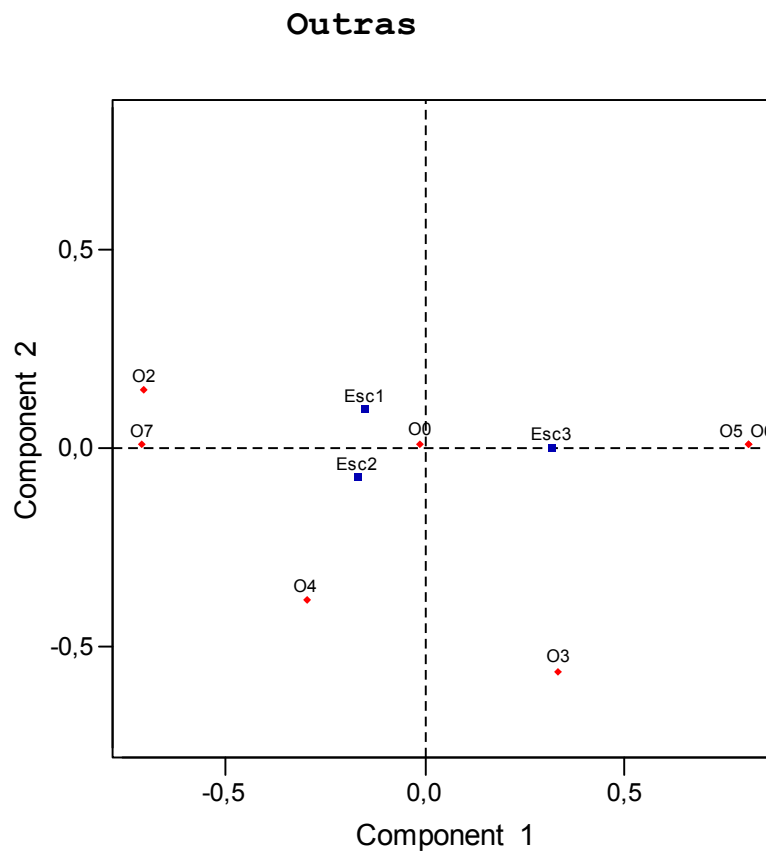


Figura 4.22. Análise de Correspondência Simples considerando-se Outras e Escolaridade.

Várias informações significativas podem ser retiradas das análises anteriores. Associando-se às médias das características analisadas (vide tabela 4.1) com as inércias das categorias P0, L0, R0, Q0, V0, E0 e O0, pode-se verificar, pela tabela 4.17, que conforme as médias das características decrescem, as inércias das respectivas características também decrescem, com exceção da característica outras.

O aumento da frequência de não respostas, as quais é atribuído valor zero, faz com que as médias decresçam juntamente com as inércias. Este fato pode ser explicado pelas distâncias χ^2 que diminuem (menor disparidade) conforme a frequência de zeros aumenta, causando diminuição da inércia das categorias P0, L0, R0, Q0, V0, E0 e O0.

Tabela 4.17. Médias das características e suas inércias para não respostas.

Característica	Média	Inércia
Qualidade	4,55	Q0 = 0,551
Rapidez	3,36	R0 = 0,336
Localização	3,04	L0 = 0,137
Preço	2,42	P0 = 0,126
Variedade	2,29	V0 = 0,083
Estacionamento	0,21	E0 = 0,009
Outras	0,56	O0 = 0,003

Utilizando-se do fato acima, pode-se pensar em realizar uma análise de correspondência simples pela ordem de opção associada à escolaridade. Esta análise pode servir ao estudo sobre quantas opções o cliente aponta, de modo a se confiar nas respostas de forma geral e, ainda, categorizando por escolaridade. A análise é iniciada pela primeira opção na qual as sete características foram apontadas, ou seja, têm frequências maiores que zero. Desta forma, para a primeira opção, agora as categorias são estacionamento (denotado por E), localização (por L), outras (por O), preço (por P), qualidade (por Q), rapidez (por R) e variedade (por V).

A primeira opção obteve inércia $\chi^2 = 46,264$, sendo que no componente 1 está 80,4% de toda a informação. Os resultados gerais podem ser vistos na tabela 4.18. Percebe-se, como era esperado, que qualidade tem a menor contribuição para a inércia total do perfil linha e escolaridade 3 tem quase metade da contribuição para a inércia total do perfil coluna.

Analisando-se a figura 4.23, observa-se que o posicionamento das categorias de escolaridade estão compatíveis com as análises anteriores para as características: a escolaridade 1 próxima de preço (aqui denotado por P), compatível com escolaridade 1 próximo de P7 (na figura 4.16), a escolaridade 2 próxima de variedade (V), compatível com escolaridade 2 sendo a mais próxima de V7 (na figura 4.20) e, ainda, a escolaridade 3 próxima de localização (L), rapidez (R) e qualidade (Q), compatível respectivamente com as análises anteriores para localização (na figura 4.17), rapidez (na figura 4.18) e qualidade (na figura 4.19).

Tabela 4.18. Contribuições dos perfis linha e coluna para a primeira opção.

Perfis	Categorias	Inércias	Contribuição do Componente 1	Contribuição do Componente 2
L I N H A	E	0,101	0,582	0,418
	L	0,106	0,842	0,158
	O	0,097	0,843	0,157
	P	0,407	0,997	0,003
	Q	0,012	0,369	0,631
	R	0,165	0,979	0,021
	V	0,112	0,022	0,978
C	Esc1	0,385	0,851	0,149
O	Esc2	0,137	0,121	0,879
L.	Esc3	0,478	0,961	0,039

Primeira Opção

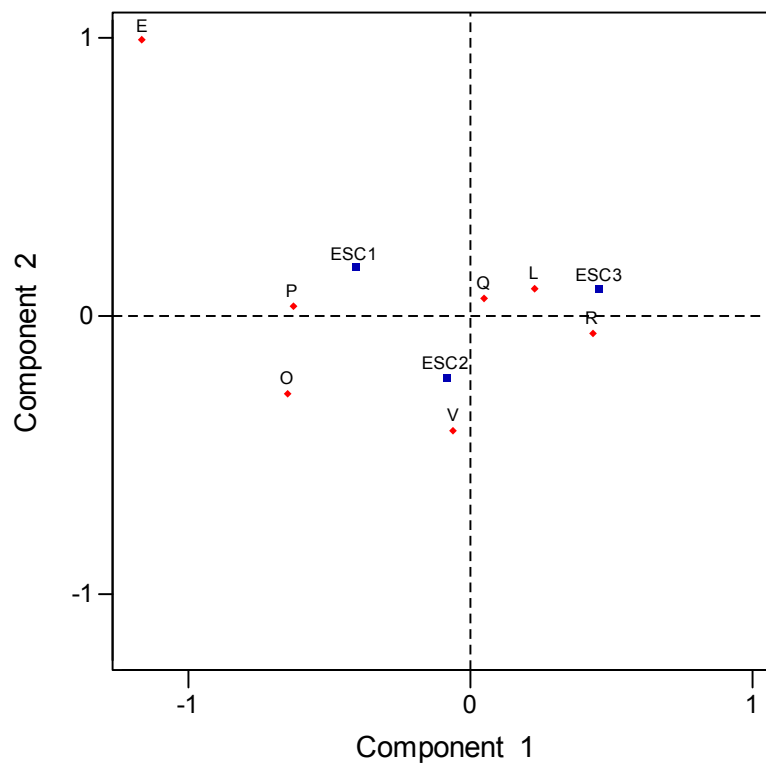


Figura 4.23. Análise de Correspondência Simples considerando-se a primeira opção.

Para a análise das demais opções, deve-se considerar uma nova categoria, aqui denominada simplesmente como *zero*. Esta se refere aos clientes que pararam de optar em alguma das categorias anteriores atribuindo-se valor zero para todas as restantes. Desta forma, conforme se diminui o grau de prioridade aumenta-se a frequência de *zeros*. Como exemplo, na análise para a quarta opção foi obtida uma inércia total $\chi^2 = 9,356$, sendo que no componente 1 está 76,7% de toda informação. Observando-se a tabela 4.19 percebe-se que a categoria preço tem 38,6% da contribuição para a inércia total do perfil linha. Novamente, a categoria escolaridade 3 tem quase a metade (47,3%) da contribuição para a inércia total do perfil coluna.

Tabela 4.19. Contribuições dos perfis linha e coluna para a quarta opção.

Perfis	Categorias	Inércias	Contribuição do Componente 1	Contribuição do Componente 2
L I N H A	Zero	0,006	0,917	0,083
	E	0,081	0,007	0,993
	L	0,193	0,995	0,005
	O	0,249	0,649	0,351
	P	0,386	0,948	0,052
	Q	0,005	0,780	0,220
	R	0,068	0,474	0,526
	V	0,012	0,461	0,539
C	Esc1	0,171	0,068	0,932
O	Esc2	0,356	0,840	0,160
L.	Esc3	0,473	0,965	0,035

Analisando-se a figura 4.24, verifica-se que agora todas as categorias de escolaridade se encontram próximas e com distância equivalentes em relação à categoria zero, sendo que esta se aloca bem perto ao centro do gráfico. Esta situação se deve ao aumento do número de não respostas, ou seja, aumento da frequência do valor zero. Nota-se então que a escolaridade 3 perde sua característica de grande diferenciação em relação às escolaridades 1 e 2. Agora, para a categoria preço a escolaridade mais próxima é a 3,

enquanto para a rapidez é a escolaridade 1. Percebe-se então que parece ter havido uma inversão de posicionamento das escolaridades 1 e 3 em relação a análise da primeira e quarta opção, nessas categorias.

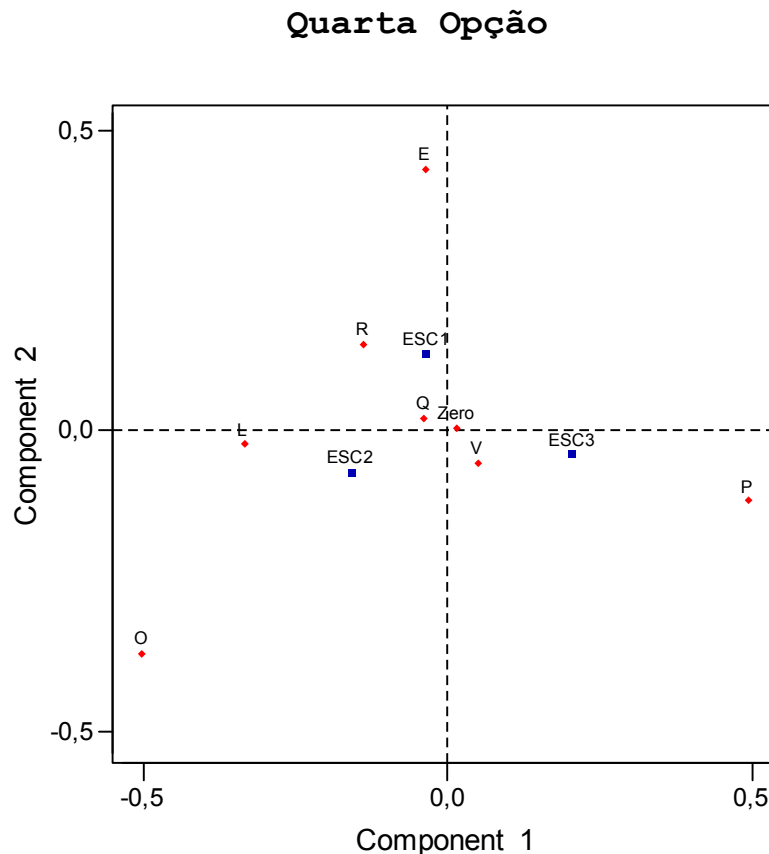


Figura 4.24. Análise de Correspondência Simples considerando-se a quarta opção.

O aumento de zeros devido a não apontamentos (não respostas) para as características, da segunda opção à sétima, vai conduzindo a categoria zero para o centro do gráfico, levando consigo as categorias de escolaridade. Lembrando que a quantidade de informação está na disparidade, percebe-se que, conforme há o aumento do grau da opção a categoria zero, também aumenta sua frequência e vai perdendo inércia. Desta forma pode-se pensar que não seja necessário colocar mais de três opções ordenadas de resposta para os clientes, pelo menos para o local no qual os dados foram coletados. Esta idéia pode ser melhor entendida observando-se a figura 4.25, na qual está representada a evolução de não respostas para as diferentes escolaridades. Pode-se notar que até a terceira opção a categoria escolaridade 3 tem uma porcentagem maior de respostas que as escolaridades 1 e

2, mas todas se tornam praticamente equivalentes para a opção 4. Observa-se ainda, de forma geral, que quanto maior a escolaridade, existe uma tendência de mais respostas obtidas, contudo há um limite para este comportamento.

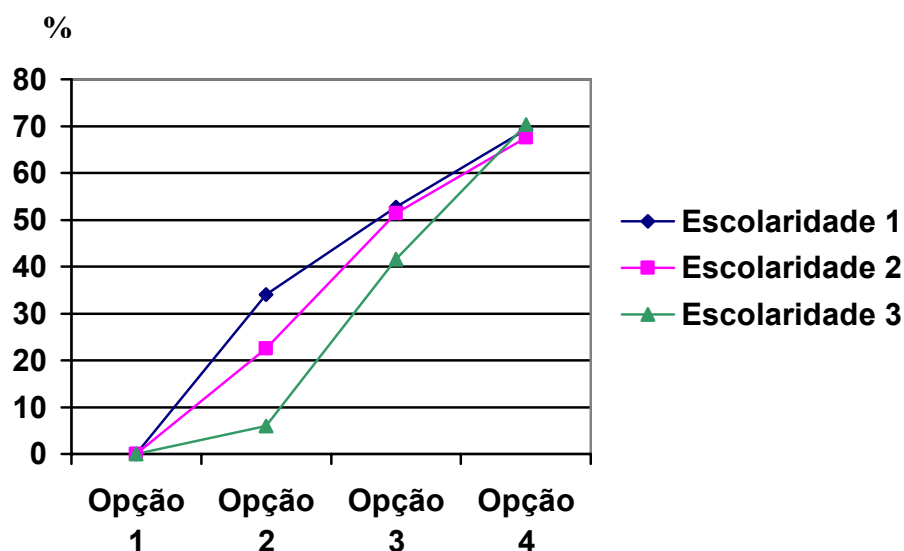


Figura 4.25. Gráfico da porcentagem de não respostas para os diferentes níveis de escolaridade.

Foi também realizada a análise quanto a impressão do cliente de diferença de preços entre supermercados distintos, associada a escolaridade do cliente. As três categorias foram alocadas como não existindo diferenças de preços (N), existindo (S) e não sabendo se existe ou não (NS). A análise de correspondência simples para diferença de preço obteve inércia total $\chi^2 = 5,340$ e com 92,8% de toda informação contida no componente 1. Pela tabela 4.20, pode-se observar que a categoria N tem praticamente dois terços da contribuição para a inércia total do perfil linha. Agora, as escolaridades 1 e 3 têm juntas praticamente toda a contribuição para a inércia total do perfil coluna.

Analisando-se a figura 4.26, observa-se que a categoria escolaridade 1 está mais próxima da categoria S, enquanto a escolaridade 2 está mais próxima ao centro, em posição intermediária entre S e NS. Claramente, a escolaridade 3 está muito mais próxima a N e isolada das demais categorias, ou seja, esta categoria está associada a não haver diferenças de preços entre supermercados, ao contrário da escolaridade 1.

Tabela 4.20. Contribuições dos perfis linha e coluna para Diferença de Preços.

Perfis	Categorias	Inércias	Contribuição do Componente 1	Contribuição do Componente 2
LINHA	N	0,648	1,000	0,000
	S	0,280	0,991	0,009
	NS	0,072	0,028	0,972
C	Esc1	0,514	0,978	0,022
O	Esc2	0,046	0,008	0,992
L.	Esc3	0,440	0,966	0,034

Diferença de Preços

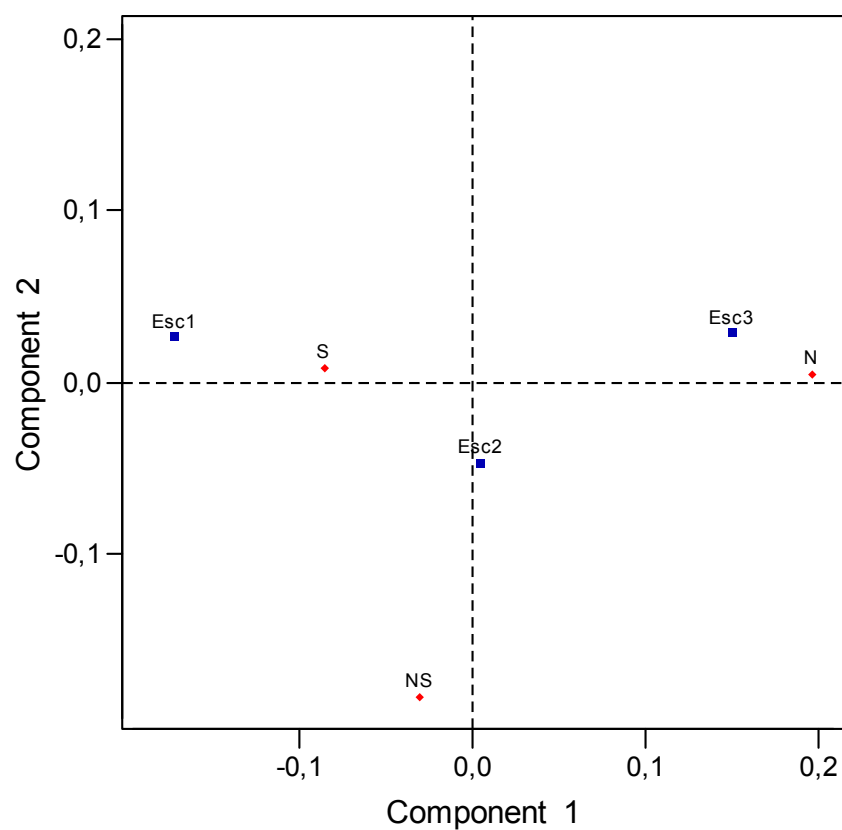


Figura 4.26. Análise de Correspondência Simples considerando-se Diferença de Preços e Escolaridade.

Outra análise realizada foi das classes relativas ao número de itens adquiridos por cliente, associado à escolaridade do cliente. Desta análise foi obtida uma inércia total $\chi^2 = 0,927$, sendo que 93% de toda informação está no componente 1. Observando-se a tabela 4.21 nota-se que a categoria B (de 16 a 70 itens) tem 68,4% da contribuição para a inércia total do perfil linha. Também, a categoria escolaridade 1 tem 64,5% da contribuição para a inércia total do perfil coluna.

Tabela 4.21. Contribuições dos perfis linha e coluna para Classes de Itens.

Perfis	Categorias	Inércias	Contribuição do Componente 1	Contribuição do Componente 2
LINHA	A	0,251	0,972	0,028
	B	0,684	0,998	0,002
	C	0,065	0,050	0,950
C	Esc 1	0,645	0,999	0,001
O	Esc 2	0,230	0,873	0,127
L.	Esc 3	0,125	0,680	0,320

Analisando-se a figura 4.27, pode-se notar claramente que a escolaridade 1 está associada à categoria B (compras médias), pois se encontra bastante distante das demais e de forma bem isolada. Nota-se também que a escolaridade 2 está associada à categoria A (compras pequenas, até 15 itens), pois estão bem próximas. Finalmente, a escolaridade 3 é a que está mais próxima da isolada categoria C (compras grandes com mais de 70 itens), contudo, esta escolaridade parece se encontrar em uma posição intermediária entre as categorias A e C. Pode-se pensar, então, que a escolaridade 3 pode estar realizando compras cujo número de itens é relativo às duas classes em questão.

As análises de correspondência simples realizadas contribuem para um melhor entendimento de algumas características importantes, pelas quais os clientes freqüentam o supermercado, associadas com os graus de escolaridade estipulados. Também pôde-se realizar as análises pelo nível de opção decrescente, de modo a ter uma visão das mudanças ocorridas durante este processo, associadas à escolaridade do cliente. Dos muitos fatores que poderiam ser considerados para expandir o conhecimento sobre o perfil dos clientes, além das características já citadas, foram também analisadas a percepção de diferença de

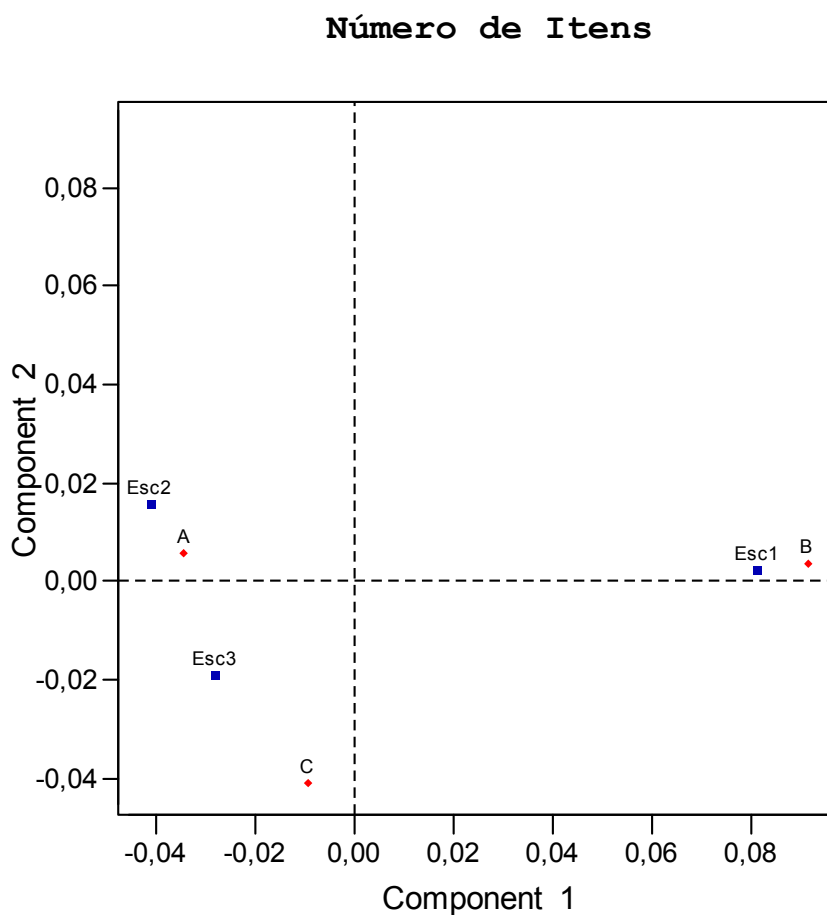


Figura 4.27. Análise de Correspondência Simples considerando-se as classes para o número de itens e escolaridade.

preços e as classes sobre a quantidade de itens adquiridos associados com a escolaridade, obtendo resultados interessantes.

Todas as análises anteriores dão suporte e direcionamento a análises mais abrangentes. Desta forma, a análise de correspondência múltipla, na qual mais de duas variáveis são cruzadas, permite a associação de características como preço e variedade com escolaridade. Cruzamentos desse tipo podem ser obtidos ampliando o alcance dos resultados até aqui encontrados, tornando os resultados mais completos. Contudo, deve-se ressaltar que a aplicação dessa técnica acarreta perda de informações que devem ser avaliadas.

4.5.4 A análise das características utilizando análise de correspondência múltipla

A análise de correspondência, de forma geral, é uma técnica ainda pouco difundida e conhecida apesar de sua ampla possibilidade de utilização. Como comentado, a análise de correspondência simples pode ser encontrada em Hair et al (2005), Aaker et al (2004), Carvalho (2004), Pereira (2004), Johnson; Wichern (2002), entre outros. Contudo, nenhum deles traz informações relevantes sobre a análise de correspondência múltipla, sendo apenas citada como existente.

A análise de correspondência múltipla é uma extensão da análise de correspondência simples para o caso de três ou mais variáveis categóricas. Este procedimento é muito parecido com o anterior (GREENACRE e BLASIUS, 2006, p. 49); as variáveis com suas categorias são ajustadas em uma seqüência de forma que fiquem todas no mesmo espaço multidimensional. Desta forma, a análise de correspondência múltipla realiza uma análise de correspondência simples na qual cada coluna da matriz corresponde a um nível da variável categórica. Anteriormente, uma tabela de dupla entrada era analisada, enquanto que aqui, uma tabela de múltiplas entradas é projetada dentro de uma dimensão.

O procedimento da análise de correspondência múltipla traz algumas alterações importantes para a análise dos dados. Utilizando-se apenas uma dimensão (ou uma variável), obtém-se apenas contribuições para um perfil, o das colunas. Não se pode desprezar da análise o número de categorias para cada variável original, ainda que todas tenham sido alocadas como uma única variável. Então, se os números de categorias em cada bloco (das variáveis originais) de categorias são dados por C_1, C_2, \dots, C_j , o número de dimensões é dado por

$$\sum_{i=1}^j (C_i - 1), \text{ para } i = 1, 2, \dots, j. \quad (4.18)$$

Da mesma forma como ocorria na análise de correspondência simples, a análise múltipla considera partições da estatística χ^2 . Assim sendo, o número de dimensões (eixos ou componentes) determina também o número de inércias (e de autovalores), tornando válida a mesma idéia sobre perda de informações citada anteriormente.

Apesar da análise de correspondência múltipla se tornar uma poderosa ferramenta para analisar a associação de um grande número de variáveis, percebe-se, pela expressão (4.18), que pode haver mais facilmente perda de informação. Existe então uma relação direta entre esta perda e a dimensão desejada (geralmente pequena).

Como exemplo, para melhor entendimento da técnica, é considerada a análise das características preço e variedade, juntamente com escolaridade. Percebe-se que o número de dimensões (novos componentes gerados por autovalores decrescentes) é dado por $(7-1) + (7-1) + (3-1) = 14$, já que as categorias P1 e V1 tiveram frequência zero. A tabela 4.22 fornece uma síntese da decomposição dos componentes para a análise das características preço, variedade e escolaridade. A coluna denominada inércia, como ocorria na análise de correspondência simples, é o valor de χ^2/n calculado para cada um dos componentes. Na terceira coluna estão os valores proporcionais de cada uma das inércias e, na última coluna, a respectiva proporção acumulada.

Tabela 4.22. Decomposição da inércia para preço, variedade e escolaridade.

Componentes	Inércia	Proporção	Acumulada
1	0,5246	0,1124	0,1124
2	0,4655	0,0998	0,2122
3	0,4376	0,0938	0,3059
4	0,4069	0,0872	0,3931
5	0,3678	0,0788	0,4719
6	0,3529	0,0756	0,5476
7	0,3422	0,0733	0,6209
8	0,3300	0,0707	0,6916
9	0,3068	0,0657	0,7574
10	0,2908	0,0623	0,8197
11	0,2786	0,0597	0,8794
12	0,2154	0,0461	0,9255
13	0,2101	0,0450	0,9705
14	0,1375	0,0295	1,0000
Total	4,6667		

Observa-se da tabela 4.22 que, para não haver grande perda de informações, seria necessário utilizar-se um grande número de componentes, contudo, esta análise está também amparada nas análises de correspondências simples para estas duas características. Na tabela 4.23, podem ser observadas as contribuições do perfil coluna para os componentes. A interpretação desta tabela é análoga à que foi realizada na análise de correspondência simples. Nota-se pelos valores das inércias que nenhuma das categorias se destaca das outras. Também, pode-se perceber, pelas contribuições dos componentes, que em raras exceções dois componentes seriam suficientes.

Tabela 4.23. Contribuições do perfil coluna para os componentes.

Categorias	Inércias	Contribuição do Componente 1	Contribuição do Componente 2
P0	0,033	0,447	0,230
P2	0,071	0,006	0,016
P3	0,063	0,578	0,134
P4	0,068	0,005	0,067
P5	0,067	0,016	0,045
P6	0,066	0,004	0,038
P7	0,060	0,054	0,322
V0	0,032	0,555	0,000
V2	0,071	0,031	0,000
V3	0,067	0,014	0,523
V4	0,066	0,332	0,031
V5	0,064	0,128	0,021
V6	0,065	0,035	0,010
V7	0,064	0,001	0,019
Esc1	0,050	0,014	0,148
Esc2	0,045	0,006	0,002
Esc3	0,048	0,039	0,107

Observando-se a figura 4.28, nota-se que a escolaridade 3 está perto da categoria P0, como ocorreu para a análise de correspondência simples, da mesma forma a escolaridade 1 está perto das categorias P6, P5, P4 e P7, o que está condizente com a análise anterior conforme a figura 4.16. Para a característica variedade, a escolaridade 1 se encontra mais próxima de V0, como ocorria anteriormente (vide figura 4.20), e a escolaridade 3 se encontra agora próxima de V7, sendo que, na análise anterior estava próxima de V5 e V6 e distante de V7. A escolaridade 2 agora está próxima a V7 e também a V6, tendo análise parecida com a anterior. Ainda para a escolaridade 2, na análise anterior se encontrava próxima às categorias P3, P5, P6 e P0 e, agora, as categorias são P0, P6 e P5, portanto, ainda condizente com a análise realizada.

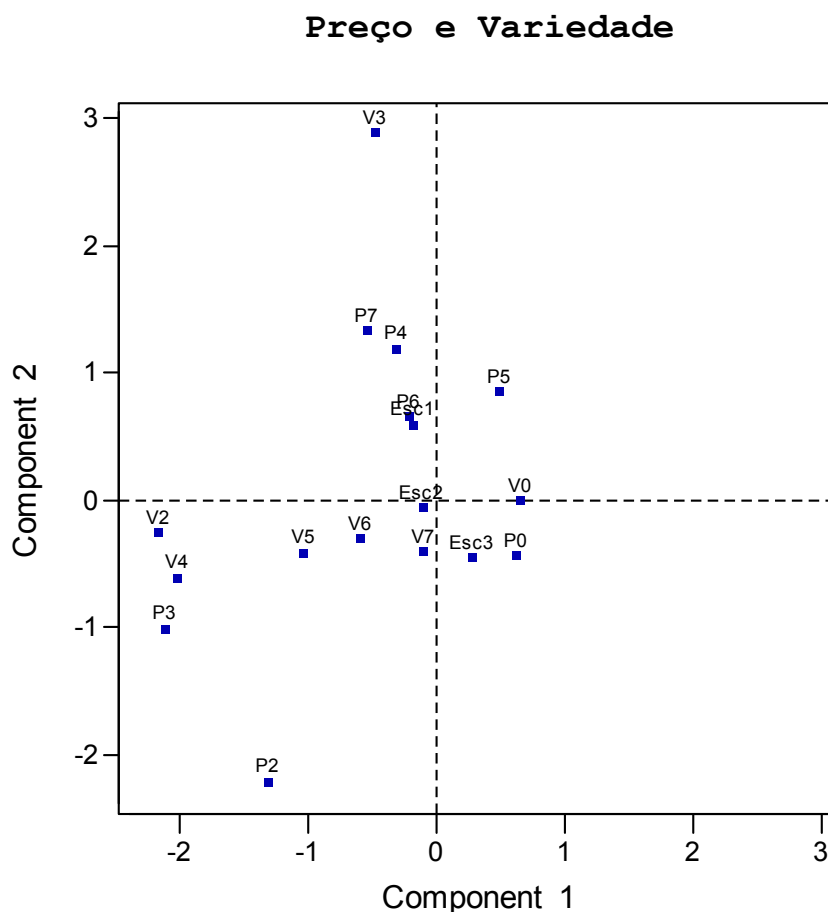


Figura 4.28. Análise de Correspondência Múltipla considerando-se preço, variedade e escolaridade.

Apesar da análise de correspondência múltipla apontar necessidade de muitos componentes, sua análise comparada às efetuadas com a análise de correspondência simples revelam a não existência de conflito entre elas, ao contrário, parece existir grande concordância e similaridade nos resultados. Desta forma, pela coerência encontrada, pode-se considerar a validade do resultado obtido.

Outra análise realizada foi das características rapidez, qualidade e escolaridade. Na tabela 4.24 podem ser observadas as inércias relativas aos componentes. Nota-se um quadro muito semelhante à análise anterior no qual, considerando a perda de informações, um grande número de componentes deveria ser considerado.

Tabela 4.24. Decomposição da inércia para rapidez, qualidade.

Componentes	Inércia	Proporção	Acumulado
1	0,5416	0,1354	0,1354
2	0,5194	0,1298	0,2652
3	0,4290	0,1072	0,3725
4	0,4138	0,1034	0,4759
5	0,3934	0,0983	0,5743
6	0,3586	0,0897	0,6639
7	0,3009	0,0752	0,7391
8	0,2682	0,0670	0,8062
9	0,2460	0,0615	0,8677
10	0,2374	0,0593	0,9270
11	0,1498	0,0375	0,9645
12	0,1421	0,0355	1,0000
Total	4,0000		

Também, pelas contribuições do perfil coluna para os componentes, na tabela 4.25, é observado também um quadro muito semelhante à análise anterior. Deve-se então observar o gráfico bidimensional para verificar se existem contradições em relação as análises de correspondência simples.

Tabela 4.25. Contribuição do perfil coluna para as componentes.

Categorias	Inércias	Contribuição do Componente 1	Contribuição do Componente 2
R0	0,050	0,644	0,020
R3	0,081	0,067	0,024
R4	0,079	0,031	0,016
R5	0,072	0,021	0,119
R6	0,063	0,328	0,198
R7	0,073	0,002	0,121
Q0	0,062	0,536	0,010
Q3	0,083	0,016	0,001
Q4	0,079	0,049	0,083
Q5	0,074	0,246	0,008
Q6	0,062	0,001	0,227
Q7	0,057	0,062	0,170
Esc1	0,058	0,073	0,001
Esc2	0,053	0,014	0,000
Esc3	0,056	0,145	0,002

Analisando-se a figura 4.29, pode-se evidenciar que as escolaridades 1 e 2 têm baixa prioridade em qualidade e rapidez, acentuada para a categoria escolaridade 1. Este resultado está bastante coerente com as análises de correspondência simples obtidas para rapidez e qualidade (vide figuras 4.18 e 4.19). Da mesma forma, esta coerência é mantida observando-se que existe alta prioridade em rapidez e qualidade para a categoria escolaridade 3. Um fato interessante é notado do lado direito da figura 4.29 na qual as categorias Q6 e R7 se encontram no lado superior do gráfico, enquanto as categorias Q7 e R6 estão no lado inferior, se contrapondo, parecendo que os clientes optavam alternando primeira e segunda opção entre estas categorias. Ainda, em torno da categoria escolaridade 3, estão várias categorias de alta e média prioridade para rapidez e qualidade, como ocorria com as análises de correspondência simples.

Novamente, apesar da análise de correspondência múltipla apontar necessidade de muitos componentes, a não existência de conflito entre as análises de correspondência simples e múltipla permite então considerar a validade do resultado obtido.

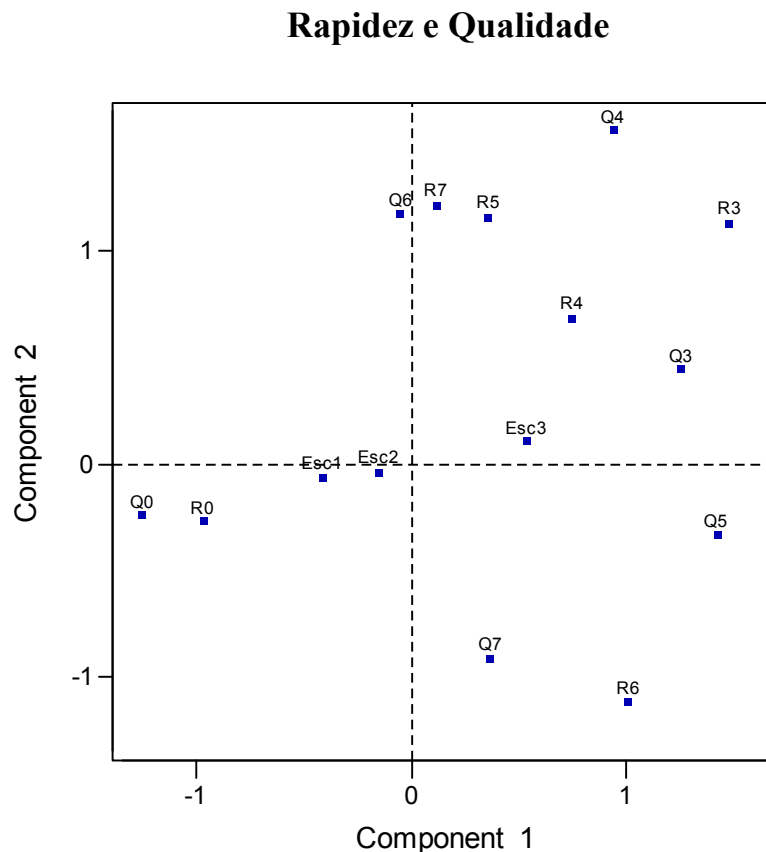


Figura 4.29. Análise de Correspondência Múltipla considerando-se rapidez, qualidade e escolaridade.

A última análise realizada neste capítulo é referente ao cruzamento das classes de itens, diferença de preços e escolaridade. Aqui, o número de componentes é $(3-1) + (3-1) + (3-1) = 6$, referentes a classes de itens, a diferença de preços e escolaridade, respectivamente. Pela tabela 4.26, observa-se que as inércias referentes aos componentes têm valores muito próximos, denotando um quadro parecido com as duas análises anteriores. Agora, contudo, o número de componentes é bem menor, mas deveria se considerar pelo menos uma inércia acumulada de 0,7, ou seja, quatro componentes.

Tabela 4.26. Decomposição da inércia para classe de itens e diferença de preços.

Eixos	Inércia	Proporção	Acumulada
1	0,3797	0,1898	0,1898
2	0,3509	0,1755	0,3653
3	0,3417	0,1708	0,5361
4	0,3296	0,1648	0,7009
5	0,3184	0,1592	0,8601
6	0,2797	0,1399	1,0000
Total	2,0000		

Considerando-se as contribuições do perfil coluna para os componentes (tabela 4.27), nota-se novamente que nenhuma das inércias se destaca muito das outras, estando bem equilibradas.

Tabela 4.27. Contribuição do perfil coluna para os componentes.

Categorias	Inércias	Contribuição do Componente 1	Contribuição do Componente 2
A	0,061	0,039	0,406
B	0,125	0,000	0,313
C	0,148	0,085	0,041
N	0,117	0,566	0,011
S	0,056	0,495	0,002
NS	0,161	0,004	0,136
Esc 1	0,117	0,347	0,309
Esc 2	0,106	0,001	0,424
Esc 3	0,111	0,359	0,016

Como ocorreu nas duas análises anteriores, compara-se a análise múltipla com as análises simples para classes de itens e diferença de preços. Pela figura 4.30 observa-se que a escolaridade 3 está bem próxima às categorias N e C, enquanto a escolaridade 1 está mais próxima às categorias S e B. A escolaridade 2 está mais próxima à categoria A e entre

as categorias S e NS. Este cenário pode ser quase que totalmente corroborado pelas análises de correlação simples, se observadas as figuras 4.26 e 4.27. O fato de não haver quase nenhuma mudança para a análise de correspondência múltipla pode ser devido ao fato do número bem menor de categorias. Além disso, deve-se lembrar que, para todas as análises, os gráficos bidimensionais estão em sua melhor representação, ou seja, com as duas maiores inércias.

Resumindo, clientes com escolaridade 3 estão associados com maiores compras e em não acreditar em diferença de preços, clientes com escolaridade 1 estão associados em acreditar em diferença de preços e realizar compras médias. A escolaridade 2 está associada a compras pequenas e tem uma posição intermediária em relação a diferença de preços.

Classes de Itens e Diferença de Preços

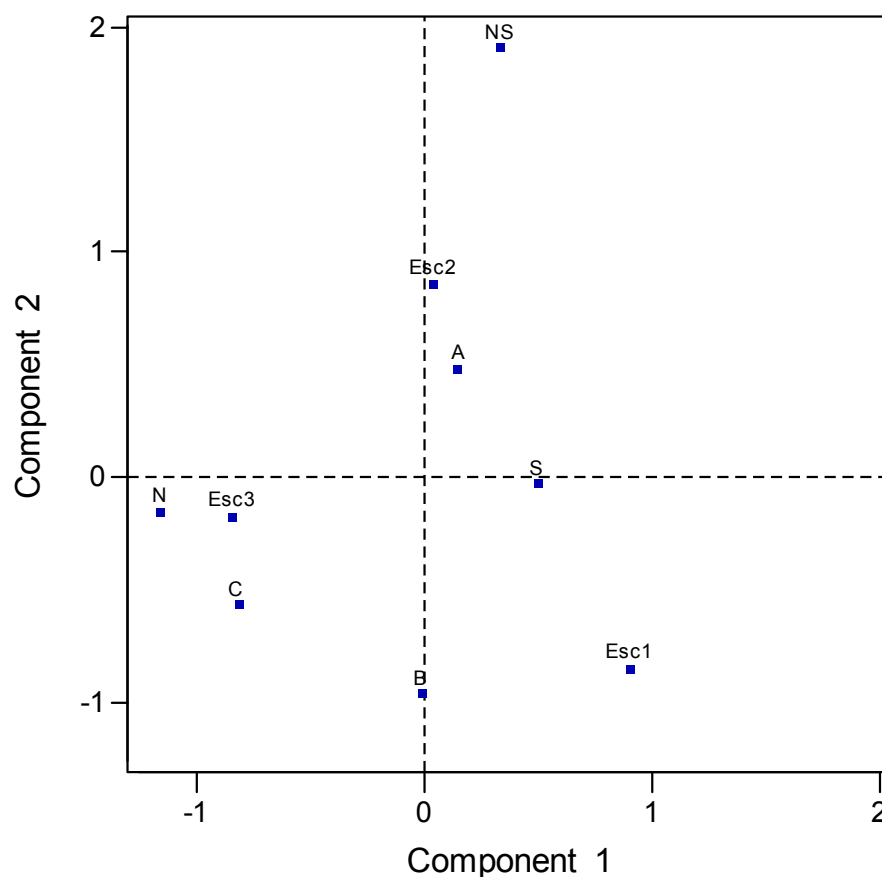


Figura 4.30. Análise de Correspondência Múltipla considerando-se classes de itens, diferença de preços e escolaridade.

Deve-se ressaltar que todas as análises múltiplas estão amparadas pelas análises de correspondências simples respectivas, contudo, levando-se em consideração as inércias acumuladas nos dois primeiros eixos e, utilizando-se de rigor estatístico, estas análises não estariam válidas.

4.6 Considerações adicionais

Algumas características percebidas pelos clientes em supermercados influenciam suas decisões, determinando suas preferências. Neste capítulo, foram analisadas as características preço, localização, rapidez, qualidade, variedade, estacionamento e outras.

Os preços estipulados pelos supermercados são influenciados pelos concorrentes, tanto que uma prática comum é a pesquisa de preços. A localização é apontada pelas empresas desse setor como sendo um fator prioritário para o bom desempenho do negócio. A esse fator, deve-se considerar condições de acesso, concorrência e um estudo sobre as condições sócio-econômicas do local. A rapidez tem assumido um papel cada vez mais importante neste setor, sendo que uma das queixas mais freqüentes nos supermercados refere-se a filas nos caixas, supermercado lotado e falta de empacotador. A melhoria da qualidade de serviços de atendimento é importante, pois pode proporcionar ao cliente boa impressão em termos de percepção do valor atribuído à relação custo/benefício da compra. A variedade é uma característica importante, pois permite ao cliente a escolha em quantidade e qualidade. Além disso, para redução da dependência de fornecedores, algumas redes têm investido em produtos com marcas próprias. A característica estacionamento pode ter se tornado importante nos últimos tempos. Isto porque existem clientes que utilizam veículos para realizar compras e querem guardá-los em local apropriado para sua comodidade. Finalmente, existem outras características que, com menor freqüência, devem ser o motivo pelo qual o cliente freqüenta o supermercado.

Destas características apresentadas muito já se apostou em preço, contudo, com a intensificação da disputa pelo cliente, cada vez mais informado e exigente, o preço, muito provavelmente, não é o único e nem o principal fator que determina uma compra. Observa-se que, por meio dos intervalos de confiança simultâneos obtidos (seção 4.2), pode-se notar

que qualidade e rapidez de atendimento formam as características mais apontadas pelos clientes como motivo de frequência ao supermercado. Estas duas características podem ser consideradas estatisticamente superiores a característica preço para frequentar o supermercado, na opinião dos clientes da empresa estudada.

Qualidade e rapidez de atendimento são características que foram associadas por meio de análise de agrupamentos hierárquicos. Então, os clientes podem ter revelado, segundo os dados coletados, que estas características estão interligadas e pertencem a um conceito mais global, de satisfação com o atendimento recebido. Assim, se a rapidez de atendimento estiver afetada negativamente, isto será refletido na satisfação do cliente, podendo, além disso, afetar sua percepção quanto à qualidade de atendimento recebida.

Os cruzamentos dessas características com a escolaridade dos clientes, por meio de análise de correspondência, permitem realizar algumas considerações. A característica preço está associada às escolaridades mais baixas, enquanto localização, rapidez e qualidade estão associadas a escolaridades mais altas. Cruzando-se rapidez, qualidade e escolaridade, pode-se perceber que escolaridades mais baixas têm pouca prioridade com relação a estas características, ao contrário de escolaridades mais elevadas.

Finalmente, o cruzamento de classes de itens, diferença de preços e escolaridade permite considerar que clientes com escolaridades mais altas parecem efetuar compras com maior número de itens e não acreditar haver diferença de preços entre supermercados. Ao contrário, escolaridades mais baixas parecem realizar compras menores e acreditar haver diferença de preços. Isto se torna mais coerente quando é considerada a pesquisa da PNAD (2004), na qual clientes com maior escolaridade provavelmente têm maior renda. Para estes clientes, não haver diferenças de preços deve ser interpretado como não haver diferença significativa de preços, o que é inerente ao seu poder aquisitivo. Além disso, efetuar compras possivelmente maiores também depende de sua possibilidade financeira.

Deve-se ressaltar que para os dados deste trabalho, 32,9% dos clientes são da categoria escolaridade 3, dos quais somente 2% não completaram o ensino superior, e 30,9% tem, no mínimo, o ensino superior completo. Além disso, outro fato conhecido é que a escolaridade no Brasil, em geral, tem aumentado ano a ano e como consequência os clientes devem ficar cada vez mais exigentes.

Quando se pensa em rapidez de atendimento em supermercados é inevitável relacioná-la ao tempo de fila. Pequenas esperas dentro do supermercado, enquanto se faz compras, parecem bem suportadas pelos clientes. O que poderia afetar negativamente a rapidez de atendimento é um tempo de fila alto, quando os clientes dificilmente se ocupam com alguma tarefa, dado que suas compras terminaram. Uma vez modificada a rapidez de atendimento, pode-se obter, como consequência, modificar a possível fidelização de algumas categorias de clientes (vide figura 4.31).

Desta forma, é possível que o principal atrativo para a frequência aos supermercados, do ponto de vista dos clientes, não sejam mercadorias variadas e com preço acessível. Estes clientes podem até mesmo pagar um pouco mais por qualidade e rapidez no atendimento recebido. Desta forma, preço pode não ter toda a importância que em geral lhe é atribuída, desde que seja similar à concorrência, pelo menos é o que sugere a análise no estudo de caso do supermercado pesquisado. Ainda, a satisfação ao atendimento recebido pode ser um fator determinante para a desejada fidelização, o que não ocorre com preço.

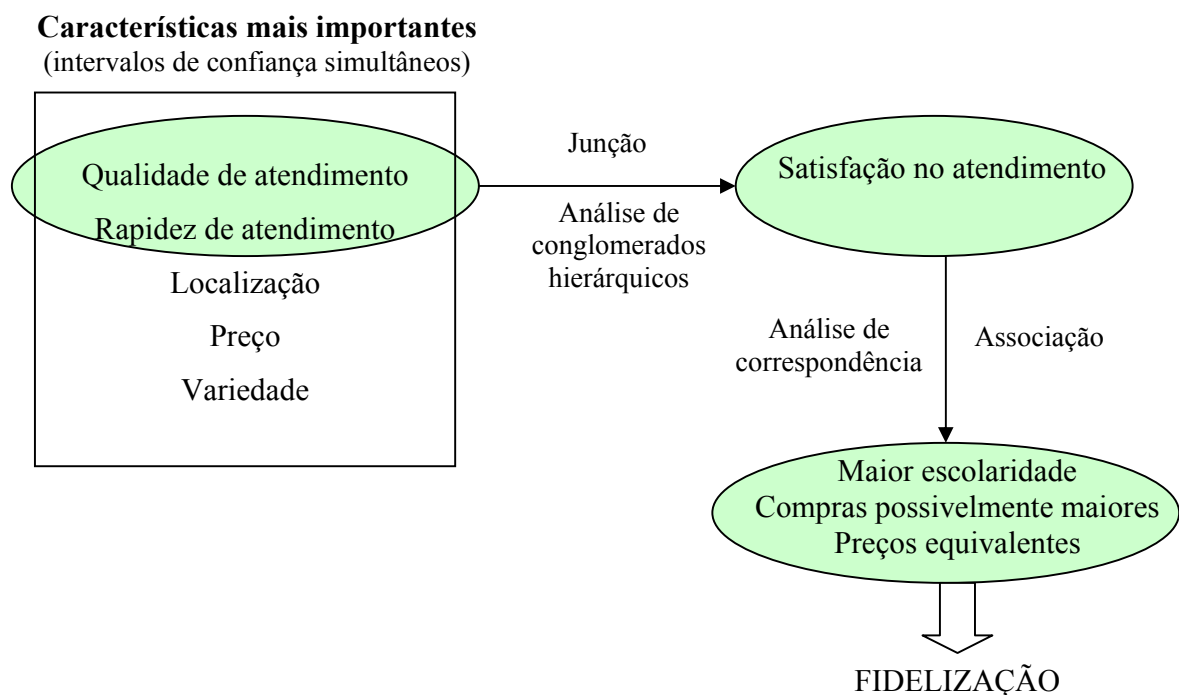


Figura 4.31. Esquema das análises e associações entre as características apontadas pelos clientes como motivo de frequência ao supermercado.

Pode-se considerar que, devido ao crescente nível de escolaridade dos brasileiros, os supermercados devem esperar também um aumento do nível de exigência dos consumidores. Seria então estratégico antecipar-se quanto à satisfação no atendimento e assim possibilitando a fidelização de seus clientes.

Com os resultados obtidos, nota-se que preço não é a característica mais importante para os clientes, principalmente aos de maior escolaridade. Quando se pensa em rapidez de atendimento, logo se associa ao tempo de espera na fila que, quando elevado, gera muitas reclamações. O cliente pode também associar grandes compras a altos tempos de espera, evitando assim supermercados em que haja demora nas filas.

5 - Sistemas de filas com trocas em supermercados

Neste capítulo são apresentados, aplicados e comparados modelos analíticos de filas para representar sistemas de caixas comumente encontrados em supermercados.

5.1 Introdução

A teoria das filas foi desenvolvida com intuito de fornecer modelos analíticos para se obter previsões sobre o comportamento dos sistemas de atendimentos. Atualmente a teoria das filas é aplicada em várias áreas distintas como manufatura, telefonia, tráfego de aeronaves, reparo de máquinas, administração hospitalar, entre outras, podendo ser citados como exemplo as filas para pagamento de compras em supermercados ou para se obter atendimento de um serviço bancário desejado. Um dos principais objetivos ao se estudar um sistema de filas é obter-se algumas medidas para conhecer seu nível de eficiência e, se necessário e possível, propor algumas políticas ou configurações alternativas para otimizar esse sistema.

Observando-se os sistemas de caixas freqüentemente encontrados em supermercados, pode-se pensar, em primeira análise, que se tratam apenas de um conjunto de filas $G/G/1$ em paralelo. Se os processos de chegada e serviço dos usuários puderem ser representados por distribuições exponenciais, então têm-se o sistema de fila $M/M/1$ que, em sua forma mais simples, tem sala de espera com capacidade infinita e apenas um servidor apresentando algumas características importantes no que diz respeito ao seu comportamento. Este sistema tem a propriedade de se esvaziar infinitas vezes com probabilidade 1, já que é regenerativo para $\rho < 1$, e pode ficar muito cheio rapidamente, ou seja, apresenta alta variabilidade no número de usuários presentes no sistema e tempo gasto pelos mesmos.

Como ilustração, foram feitas três realizações independentes da simulação da fila $M/M/1$, com 700 observações cada uma, considerando a variável tempo gasto na fila com taxa de chegadas $\lambda=0,9$ e taxa de serviços $\mu=1$ e representadas na figura 5.1 (RINALDI, 1999). Deve-se ressaltar que estas realizações partiram em situação de

igualdade, ou seja, foram iniciadas com tempo médio gasto (ou de espera) na fila, obtido como

$$W_q = \frac{\lambda}{\mu(\mu - \lambda)} = 9 \text{ unidades de tempo}$$

o que corresponde ao tempo médio gasto na fila $M/M/1$ com as taxas consideradas.

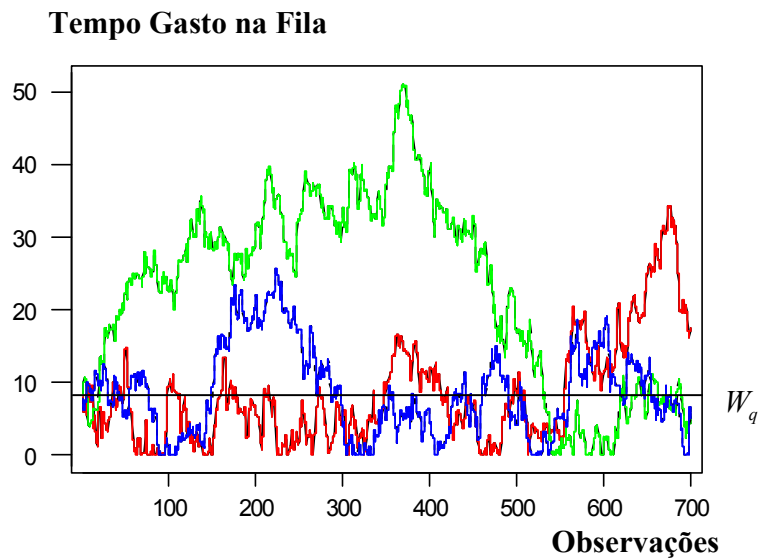


Figura 5.1. Três realizações de simulação da fila $M/M/1$ considerando o tempo gasto na fila com $\rho=0,9$.

Apesar de partir nesta situação (no mesmo ponto), pela figura 5.1, observam-se comportamentos distintos em relação às observações de tempo gasto na fila para as realizações consideradas (se analisadas suas médias para este período, aquela da realização representada em verde é visivelmente maior que as outras), bem como no número de passagens para o estado vazio que cada uma delas obtém. Evidentemente esta situação é agravada pelo alto valor do fator de utilização considerado ($\rho = \lambda / \mu = 0,9$). Deve-se ressaltar que quando $\rho \rightarrow 1$, como conseqüência, $W_q \rightarrow \infty$.

De forma geral, para a fila $M/M/1$, quando o valor do fator de utilização é alto, o sistema tende a se tornar instável, ou seja, apresenta alta variabilidade em suas medidas. Se fossem três caixas de supermercados, para este período, este sistema não seria justo, pois notam-se grandes diferenças de tempos de espera nas filas. Evidentemente, em uma

situação real, um novo cliente que chega irá entrar na menor fila encontrada, evitando que as filas fiquem tão discrepantes. Também, após estar em alguma fila, o cliente provavelmente não desprezará a oportunidade de mudar-se para outra que se apresente com menos clientes. Desta forma, pode-se dizer que o sistema é equilibrado (ou auto-regulado) por meio do próprio comportamento dos clientes e espera-se alguma uniformidade para as chegadas nas filas.

As trocas de filas têm outras conseqüências quando se considera um sistema composto de filas $M/M/1$ independentes e em paralelo. Em cada fila, a função de autocorrelação (denotada por ρ_j^*) entre as observações do tempo gasto na fila revela que existe forte correlação, mesmo entre observações mais distantes (vide figura 5.2), sendo que estas distâncias podem ser medidas pelos *lags*, ou atrasos. Como exemplo, a autocorrelação entre duas observações x_i e x_{i+j} é dada por $\rho_j^* = \text{corr}(x_i, x_{i+j})$, ou seja, a correlação entre duas observações que distam j observações uma da outra (ou de *lag* j). Pela figura, observa-se que mesmo com *lag* 100, a autocorrelação está acima de 0,5, forte para este *lag*, explicando possivelmente o comportamento de altas e baixas alongadas para as três realizações representadas na figura 5.1. Quando clientes trocam de fila, rompem esta forte estrutura de autocorrelação, deixando falhas em sua seqüência, e alterando o comportamento anterior.

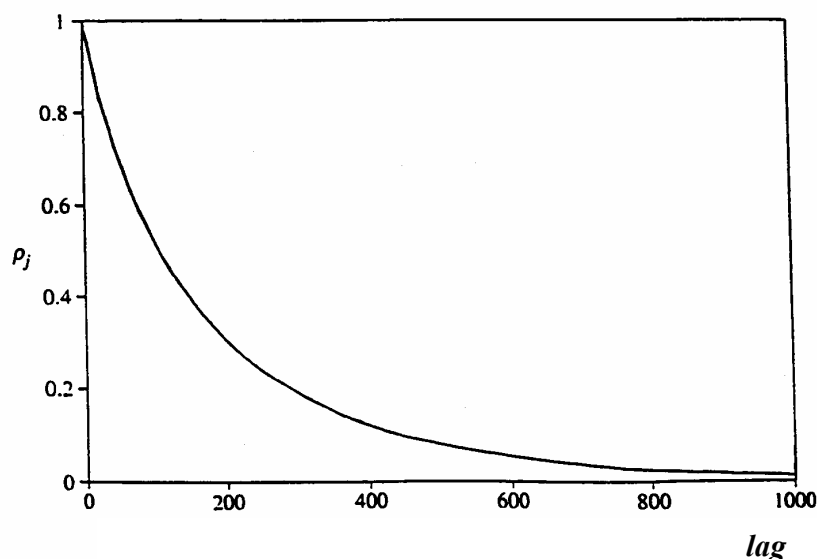


Figura 5.2. Autocorrelações teóricas do tempo gasto no sistema de fila $M/M/1$ com $\rho=0,9$ (FISHMAN, 1996, p. 550).

Utilizando-se modelos de filas que permitam realizar trocas, espera-se uma aproximação da realidade dos supermercados e obter melhorias, como menor tempo de espera na fila, e propiciar um sistema mais justo ao usuário. Algumas pesquisas realmente apontam nessa direção, como em Zhao e Grassmann (1995), na qual analisam um modelo com filas paralelas de servidores heterogêneos com uma fila para cada um, serviço com distribuição exponencial, tempos entre chegadas com uma distribuição arbitrária, sendo permitidas trocas para filas com menos clientes. Por meio de matriz-geométrica solucionaram o modelo, obtendo significativa vantagem utilizando-se trocas em relação ao mesmo modelo sem trocas.

Ferrari (2002) analisou um sistema de filas com três servidores com distribuição de tempos de serviço exponenciais, heterogêneos, em paralelo e permitindo trocas probabilísticas (não obrigatórias) entre as filas após o usuário examinar o número de clientes nas filas e seu grau de vizinhança. Em suas conclusões ressalta que o sistema com troca probabilística apresenta menor ociosidade que o sistema tradicional $M/M_i/3$ com fila única, no qual M_i denota que os três servidores operam de forma independente entre si.

Koenigsberg (1966) analisou várias estratégias para troca entre duas filas paralelas. A troca denominada instantânea tem como característica obrigar o cliente a executá-la. Outro tipo analisado por ele foi a probabilística, na qual o usuário opta ou não pela troca, podendo a troca ainda ser regida pelas probabilidades de uma distribuição exponencial. Finalmente, analisou a possibilidade de realizar a troca em duas direções, da fila maior para a menor, ou vice versa.

Um modelo denominado por Koenigsberg como “*Teller’s Windows with Jockeying*” é composto por dois servidores paralelos e independentes com filas próprias. As chegadas ocorrem segundo um processo de Poisson e os tempos de serviço são distribuídos exponencialmente. O usuário que chega escolhe a fila de menos clientes, se houver igualdade de condições, entra em qualquer uma, aleatoriamente. Estando em uma fila, se ele observa a existência de uma outra menor com uma diferença maior que um usuário, instantaneamente ele se transfere para ela. Este modelo pode ser interpretado como um caso particular do modelo analisado no presente trabalho, quando a diferença entre filas for maior que dois usuários e/ou com número maior de filas.

5.2 As modelagens freqüentemente utilizadas

Uma primeira modelagem que pode ser pensada é com o sistema de fila Markoviano no qual há m caixas atendendo em paralelo e com uma única fila, tornando este justo para os clientes. Trata-se então do modelo $M/M/m$ com fila única e disciplina de atendimento FCFS (*First Come, First Served*). Deve-se considerar a taxa média de chegadas λ (clientes por unidade de tempo) e que os caixas são idênticos e independentes, ou seja, tem a mesma capacidade de atendimento, com taxa média μ (clientes por unidade de tempo) e o funcionamento de um não interfere nos outros.

Este sistema de filas é considerado de simples análise e, em equilíbrio, várias medidas de desempenho podem ser obtidas (KLEINROCK, 1975). A medida de maior importância para este estudo é o tempo médio de espera na fila, denotada por W_q , e determinada pela expressão:

$$W_q = \frac{(\lambda / \mu)^m p_0}{\mu m (1 - \rho)^2 m!} \quad (5.1)$$

na qual o fator de utilização, denotado por ρ , e a probabilidade do sistema estar vazio, denotada por p_0 , são obtidos respectivamente pelas expressões:

$$\rho = \frac{\lambda}{m\mu} \quad \text{e} \quad (5.2)$$

$$p_0 = \left\{ \sum_{i=0}^{m-1} \frac{(\lambda / \mu)^i}{i!} + \frac{(\lambda / \mu)^m}{(1 - \rho) m!} \right\}^{-1}. \quad (5.3)$$

Apesar deste modelo ter análise simples e ser justo, é pouco realista, pois os supermercados geralmente não possuem espaço físico para que seja formada uma única fila, pelo menos considerando os caixas normais.

A situação que comumente se observa é a utilização de filas paralelas, ou seja, uma fila para cada caixa. Assim, uma outra modelagem que pode ser pensada é admitir que os caixas são supostamente idênticos e independentes com taxa de serviço μ , e agora se consideram a taxa de chegada $\lambda' = \lambda / m$ para cada uma das m filas. Neste caso supõe-se que, uma vez que o cliente entrou em uma fila, ele não mudará para outra. Este sistema descrito é o de m filas $M/M/1$ paralelas e independentes. Assim, como o anterior, é de simples análise e várias medidas de desempenho podem ser obtidas, dentre as quais, o tempo médio de espera na fila, determinado pela expressão (KLEINROCK, 1975):

$$W_q = \frac{\lambda' / m}{\mu (\mu - \lambda' / m)}. \quad (5.4)$$

Este modelo, denotado por $[M/M/1]^m$, parece ser mais realista para representar sistemas de caixas de supermercado do que o primeiro, contudo, além de não ser socialmente justo, não prevê que o cliente possa escolher em que fila entrar e realizar a troca quando julgar vantajoso, o que é comum de ocorrer na prática. Deve-se ressaltar ainda que nenhum dos dois modelos considera que o cliente possa simplesmente desistir das compras, o que acontece quando este julga que as filas estão maiores do que ele toleraria.

Pensando-se assim, poderiam-se considerar os dois modelos anteriores de forma truncada, ou seja, aceitar novos clientes que chegam até que as filas tenham um número máximo de N pessoas (entre serviço e fila) e, a partir daí, os clientes seriam perdidos, pois desistiriam das compras. Desta forma, podem ser considerados também o modelo $M/M/m/mN$ com fila única e o modelo $[M/M/1/N]^m$. Deve-se salientar que, deste modo, os dois modelos levam ao mesmo número máximo de clientes no sistema, ou seja, mN , e, ainda, é suposto aqui e para o presente trabalho que $\rho < 1$.

As medidas de desempenho para esses modelos também podem ser facilmente obtidas (GROSS e HARRIS, 1974). Para ambos os modelos, a medida de maior interesse é W_q . Considerando-se o modelo $M/M/m/mN$ com fila única, o tempo médio de espera na fila é obtido pela expressão:

$$W_q = L_q / \lambda_E, \quad (5.5)$$

na qual λ_E (taxa de entradas, referente aos usuários que efetivamente entram no sistema) e o número médio de usuários na fila (L_q) são dados respectivamente por:

$$\lambda_E = \lambda(1 - p_{mN}) \text{ e} \quad (5.6)$$

$$L_q = \frac{p_0(m\rho)^m}{m!(1-\rho)^2} \left[1 - \rho^{mN-m+1} - (1-\rho)(mN-m+1)\rho^{mN-m} \right],$$

nos quais a probabilidade de perda (p_{mN}) e a probabilidade de o sistema estar vazio (p_0) são dadas respectivamente por:

$$p_{mN} = \frac{1}{m^{mN} m!} \left(\frac{\lambda}{\mu} \right)^{mN} p_0 \text{ e}$$

$$p_0 = \left[\sum_{n=0}^{m-1} \frac{1}{n!} \left(\frac{\lambda}{\mu} \right)^n + \frac{(\lambda/\mu)^m}{m!} \frac{1 - (\lambda/m\mu)^{mN-m+1}}{1 - (\lambda/m\mu)} \right]^{-1}.$$

Da expressão (5.6) observa-se que a taxa de chegadas λ deve ser corrigida, pois se trata de um sistema com perda, medida por p_{mN} . Para este sistema, por haver perda de clientes, o fator de utilização não corresponde à utilização média do sistema, que é então a proporção de servidores ocupados. Aqui, esta proporção é dada por $\lambda_E/m\mu$.

De forma análoga ao sistema anterior, pode-se considerar o sistema de fila formado por m filas $M/M/1/N$, ou seja, $[M/M/1/N]^m$, com taxa de chegadas $\lambda = \lambda/m$ em cada uma das filas. O tempo médio de espera na fila é obtido pela expressão:

$$W_q = W - 1/\mu, \quad (5.7)$$

na qual o tempo médio de espera no sistema é dado por:

$$W = L/\lambda_E,$$

em que o número médio de usuários no sistema (L), λ_E e ρ são obtidos como:

$$L = \frac{\rho[1 - (N+1)\rho^N + N\rho^{N+1}]}{(1 - \rho^{N+1})(1 - \rho)},$$

$$\lambda_E = \lambda(1 - p_N) \text{ e}$$

$$\rho = \lambda / \mu,$$

sendo ainda que

$$p_N = \frac{(1 - \rho)\rho^N}{1 - \rho^{N+1}}.$$

Ainda, como informação adicional para este sistema, a probabilidade de o sistema estar vazio é dada por $p_0 = (1 - \rho)/(1 - \rho^{N+1})$. Novamente, como discutido na fila $M/M/m/mN$, a taxa de chegadas deve ser corrigida, pois se trata de um sistema com perda, medida por p_N para cada fila. Observa-se que o tempo médio de espera no sistema é o tempo médio de espera em qualquer das m filas e que o número máximo de clientes no sistema é dado por mN . Deve-se ressaltar que há outras formas de modelar a desistência do usuário. Como exemplo, podem existir desistências na chegada (o usuário, ao chegar, desiste porque observa que a fila está muito grande) representadas por uma função decrescente da taxa de chegadas com o número de usuários no sistema. Em geral, quando o sistema está com n usuários, chegadas são desencorajadas. A desistência é representada por

$$C_n = \left(\frac{1}{n+1}\right)^a \text{ para } n=0,1,2,\dots$$

onde a é uma constante tal que $a \geq 0$, a ser determinada. Também, abandonos na espera (o usuário, ao esperar muito tempo na fila, abandona o sistema) pode ser obtido como

$$D_n = 1 + \frac{g_n}{n}$$

na qual g_n é uma função de abandono a ser determinada em cada caso. A função g_n e o parâmetro a podem ser obtidos, por exemplo, aplicando-se técnicas de regressão.

Uma outra possibilidade para representar filas em supermercados é a aplicação de modelos de filas com distribuições genéricas. Inicialmente pode-se pensar em m filas $M/G/1$, com taxa de chegadas $\lambda = \lambda/m$ para cada uma das filas, sendo que o número médio de usuários no sistema pode ser obtido pela fórmula de Pollaczek e Khintchine (ARENALES et al, 2006), cuja expressão é

$$L = \rho + \frac{\rho^2 + \lambda^2 \sigma_s^2}{2(1 - \rho)} \quad (5.8)$$

na qual $\rho = \lambda / \mu$ e σ_s^2 é a variância do tempo de serviço. As demais medidas de desempenho podem ser obtidas utilizando-se a fórmula de Little. Da mesma forma que para os demais modelos, pode-se pensar agora no modelo $M/G/m$ com taxa de chegadas λ . Uma aproximação para o número médio de usuários no sistema é baseada na fórmula de Kraemer e Lagenbach-Belz (BITRAN e MORABITO, 1996), dada por

$$L = \rho m + \frac{(1 + C_s^2)}{2} (L_q)_{M/M/m} \quad (5.9)$$

na qual $(L_q)_{M/M/m}$ é o número médio de usuários no sistema $M/M/m$, $\rho = \lambda / \mu$ e C_s é o coeficiente de variação do tempo de serviço. Novamente, as demais medidas de desempenho podem ser obtidas utilizando-se a fórmula de Little.

Um outro modelo, mais completo que os modelos markovianos $M/M/m/mN$ $[M/M/1/N]^m$ anteriores, foi proposto por Morabito e Lima (2004), considerando a desistência do serviço e a troca de filas pelos clientes nas m filas $M/M/1/N$ paralelas. Desta

forma, em relação aos modelos markovianos anteriores, este modelo se torna mais próximo da situação na qual, geralmente, os supermercados operam. Apesar de ainda não ser completamente justo permite, como realmente ocorre, que o cliente troque de filas na tentativa de abreviar seu tempo de espera. Este modelo é detalhado a seguir.

5.3 O modelo com trocas

A proposta deste modelo, em sua formulação, é englobar algumas características que ocorrem na prática em filas de supermercados, que alguns modelos abordam de forma isolada. Inicialmente, algumas considerações devem ser realizadas. Em supermercados, conforme depoimentos de alguns gerentes, é bastante comum que, quando os clientes encontram um certo número de pessoas na fila, desistam de comprar. Desta forma, considera-se que a partir de N pessoas, contadas entre fila e serviço, os clientes que chegam se retiram do estabelecimento sem efetuar suas compras. Também, se a diferença de usuários entre duas filas estiver maior que um certo número, diga-se k , os consumidores trocam de fila na tentativa de abreviar sua espera. Ainda, se o consumidor que chega decidir entrar na fila, o fará onde considera que irá demorar menos.

Existe algum subjetivismo em fixar um valor para N , pois o comportamento para desistência pode variar de um cliente para o outro. Também, o valor de k pode variar de acordo com a expectativa e a quantidade de tempo disponível do cliente no instante de decisão. Contudo, para o modelo a seguir, será fixado um valor para N , a ser definido em situação prática, e quando houver uma diferença maior que a estipulada (k) entre duas filas, o modelo considera que o cliente necessariamente efetuará a mudança para a fila menor.

O espaço de estados do sistema, denotado por (n_1, n_2, \dots, n_m) , é referente ao número de clientes encontrados em cada um dos m caixas em um dado instante. Seja então o espaço de estados do sistema com m filas, (n_1, n_2, \dots, n_m) , na qual n_i é o número de clientes no caixa i , para $i=1, 2, \dots, m$. A partir da forma com que está sendo considerado o modelo, pode-se observar que, para todo caixa i , $0 \leq n_i \leq N$, e o espaço de estados pode ser definido por:

$$(n_1, n_2, \dots, n_m)$$

no qual $0 \leq n_i \leq N$ e $|n_i - n_j| \leq k$ para $i \neq j$ e $i, j = 1, 2, \dots, m$.

Observa-se que, das duas condições acima, não se permite que nenhum caixa tenha mais que N clientes (em fila e em serviço) e, que nenhum dos caixas possa ter, entre si, uma diferença no número de clientes maior que k . Como exemplo, se $m=4$, $k=1$ e $N=3$, um possível estado é $(1, 2, 2, 1)$ e um impossível é $(1, 2, 3, 1)$.

No modelo de fila em questão, podem ocorrer mudanças do estado atual de duas formas. Estas se referem a dois instantes, nos quais é possível transições entre os estados: no instante de chegada de um novo cliente e no instante em que algum cliente termina de ser servido e deixa o caixa. No primeiro, o cliente chega e, se não houver saturação do sistema, entra em qualquer uma das filas menores, com probabilidades iguais. No segundo, o cliente deixa o sistema e, dependendo do estado em que o sistema fica quando da saída, isto pode envolver também alguma troca de clientes nas filas.

5.3.1 A mudança no instante de chegada

Para haver mudança no instante de chegada, é necessário que, para um presente estado $(n_1, n_2, \dots, n_i, \dots, n_m)$, algum $n_i < N$, ou seja, o sistema não está saturado. Então, um novo estado $(n_1, n_2, \dots, n_i + 1, \dots, n_m)$ pode ser obtido, com taxa de transição λ / m' . Aqui, m' é tal que $m' \leq m$ e se refere ao número de possíveis transições a partir de $(n_1, n_2, \dots, n_i, \dots, n_m)$, pois alguns caixas podem estar lotados. Outra condição para que uma mudança ocorra é que $|n_i + 1 - n_j| \leq k$ para todo $i \neq j$ e $j = 1, 2, \dots, m$, ou seja, é necessário também que a diferença entre o número de clientes nos caixas não seja superior ao que foi estipulado (o valor de k). De modo geral, m' é definido pela expressão

$$m' = \sum_{i=1}^m \delta(|n_i + 1 - n_j| \leq k) \quad (5.10)$$

na qual

$$\delta(|n_i + 1 - n_j| \leq k) = 1 \text{ se } |n_i + 1 - n_j| \leq k \text{ para todo } i \neq j \text{ e } j = 1, 2, \dots, m;$$

$$\delta(|n_i + 1 - n_j| \leq k) = 0, \text{ caso contrário.}$$

Deve-se então, a partir do estado presente, contar as possíveis transições para balancear a taxa de transição λ/m . Como exemplo, sejam $m=4$, $k=1$ e $N=4$ e um estado atual $(1,0,1,0)$. Uma transição de chegada pode ser realizada para o estado $(1,1,1,0)$, pois

$$|n_2 + 1 - n_1| = 0 \leq k;$$

$$|n_2 + 1 - n_3| = 0 \leq k;$$

$$|n_2 + 1 - n_4| = 1 \leq k$$

mas não pode ser realizada para o estado $(2,0,1,0)$, pois

$$|n_1 + 1 - n_2| = 2 > k.$$

Outra transição também poderia ser para o estado $(1,0,1,1)$ e, assim sendo, percebe-se duas possibilidades para efetuar a mudança de estado por meio de uma chegada. Desta forma, tendo duas possibilidades, pode-se obter que $m=2$ e, logo, a taxa de transição para os referidos estados é dada por $\lambda/2$. A figura 5.3 representa um esquema das possíveis mudanças para esta situação.

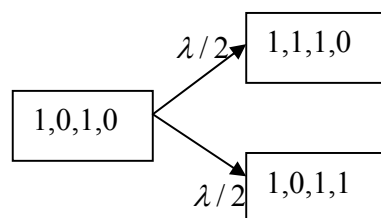


Figura 5.3. Esquema para possíveis mudanças no instante de chegada.

5.3.2 A mudança no instante de saída

Para a mudança no instante de saída, há dois subcasos que podem ser considerados: o primeiro se refere a quando há pelo menos um caixa vazio e o segundo, a quando todos as caixas estão ocupados.

i- Quando há pelo menos um caixa vazio:

A partir da função indicadora:

$$\phi(n_j)=1, \text{ se } n_j > 0$$

$$\phi(n_j)=0, \text{ se } n_j = 0, \text{ para } j=1,2,\dots,m,$$

pode-se obter que:

$$\sum_{j=1}^m \phi(n_j) < m$$

em razão de haver pelo menos um $n_j = 0$. Desta forma, se $|n_i - 1 - n_j| \leq k$ para todo $i \neq j$ e $j = 1, 2, \dots, m$, então, a partir de um estado $(n_1, n_2, \dots, n_i, \dots, n_m)$ com $n_i > 0$, muda-se para um novo estado $(n_1, n_2, \dots, n_i - 1, \dots, n_m)$ com taxa de transição μ . Como exemplo, se $m=4$, $k=1$ e $N=4$, então se o estado atual é $(1, 0, 1, 1)$, pode-se mudar somente para $(0, 0, 1, 1)$, $(1, 0, 0, 1)$ ou $(1, 0, 1, 0)$ com taxa de transição μ (vide figura 5.4).

ii- Quando não há caixas vazios:

Na situação em que todos os caixas estão ocupados, pode-se observar que

$$\sum_{j=1}^m \phi(n_j) = m.$$

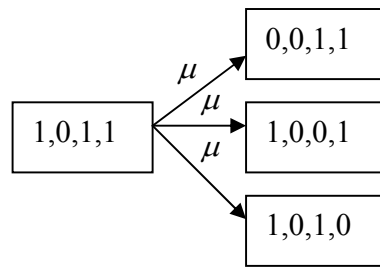


Figura 5.4. Esquema para possíveis mudanças no instante de saída com um caixa vazio.

Assim como ocorreu com as mudanças no instante de chegada, agora também é necessário conhecer o número de possíveis transições do estado presente. Como exemplo, seja o estado $(1,2,1,1)$ com $m=4$, $k=1$ e $N=4$ e que, devido ao término do serviço, o cliente do caixa 1 saia do sistema. Instantaneamente, um cliente do caixa 2 se desloca para o caixa 1, deixando o sistema no estado $(1,1,1,1)$ (vide figura 5.5, a seguir). Observa-se então que, como $k=1$, qualquer uma das quatro possíveis saídas só podem levar a um único estado, que é $(1,1,1,1)$, então com taxa 4μ .

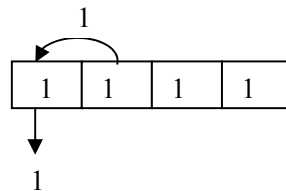


Figura 5.5. Esquema para mudança instantânea, no momento de saída do cliente do caixa 1, sem caixa vazio.

De forma geral, a partir do estado $(n_1, n_2, \dots, n_i, \dots, n_m)$, a mudança para $(n_1, n_2, \dots, n_i - 1, \dots, n_m)$ ocorre se $|n_i - 1 - n_j| \leq k$ para todo $i \neq j$ e $j = 1, 2, \dots, m$ e com taxa de transição $m\mu/m''$ (com $m'' \leq m$), sendo m'' dado pela expressão:

$$m'' = \sum_{i=1}^m \delta(|n_i - 1 - n_j| \leq k) \quad (5.11)$$

na qual

$$\delta(|n_i - 1 - n_j| \leq k) = 1 \text{ se } |n_i - 1 - n_j| \leq k \text{ para todo } i \neq j \text{ e } j = 1, 2, \dots, m;$$

$$\delta(|n_i - 1 - n_j| \leq k) = 0, \text{ caso contrário.}$$

Desta forma, por meio da razão m/m'' , pode-se balancear a taxa de transição de acordo com os possíveis novos estados. Nota-se que, no exemplo anterior da figura 5.5, tem-se $m'' = 1$, logo, $m\mu/m'' = 4\mu$.

Para solução do problema, é preciso montar cada uma das equações de equilíbrio (ou de balanço) para as probabilidades de cada possível estado (n_1, n_2, \dots, n_m) do sistema de fila, obtendo-se um sistema de equações lineares nas probabilidades de equilíbrio. Para que não haja indeterminação, é suficiente substituir uma das equações pela que representa a soma das probabilidades de todos os possíveis estados como sendo igual a 1. As equações do sistema a ser resolvido devem restringir-se às condições $n_i \leq N$ e $|n_i - n_j| \leq k$ para $i \neq j$ e $j = 1, 2, \dots, m$, conforme descrito anteriormente.

Se nenhuma condição fosse imposta sobre o valor de k , o sistema linear gerado pelo problema teria $(N+1)^m$ incógnitas, o que revela uma tendência em crescer exponencialmente com o número de caixas m . Para melhor entendimento sobre o funcionamento do sistema de fila, fez-se uma ilustração de parte de todas as possíveis mudanças de estados, de um exemplo com um sistema no qual $m=4$, $k=1$ e $N=3$ (vide figura 5.6). Observando-se o diagrama, pode-se montar as equações de equilíbrio para cada um dos estados possíveis do sistema. Como exemplo, a equação do estado $(1,1,2,1)$ fica

$$(4\mu + \frac{\lambda}{3} + \frac{\lambda}{3} + \frac{\lambda}{3})p(1,1,2,1) = \frac{\lambda}{4}p(1,1,1,1) + (\mu + \frac{\mu}{2} + \frac{\mu}{2})p(2,1,2,1) +$$

$$+ (\frac{\mu}{2} + \mu + \frac{\mu}{2})p(1,2,2,1) + (\frac{\mu}{2} + \frac{\mu}{2} + \mu)p(1,1,2,2).$$

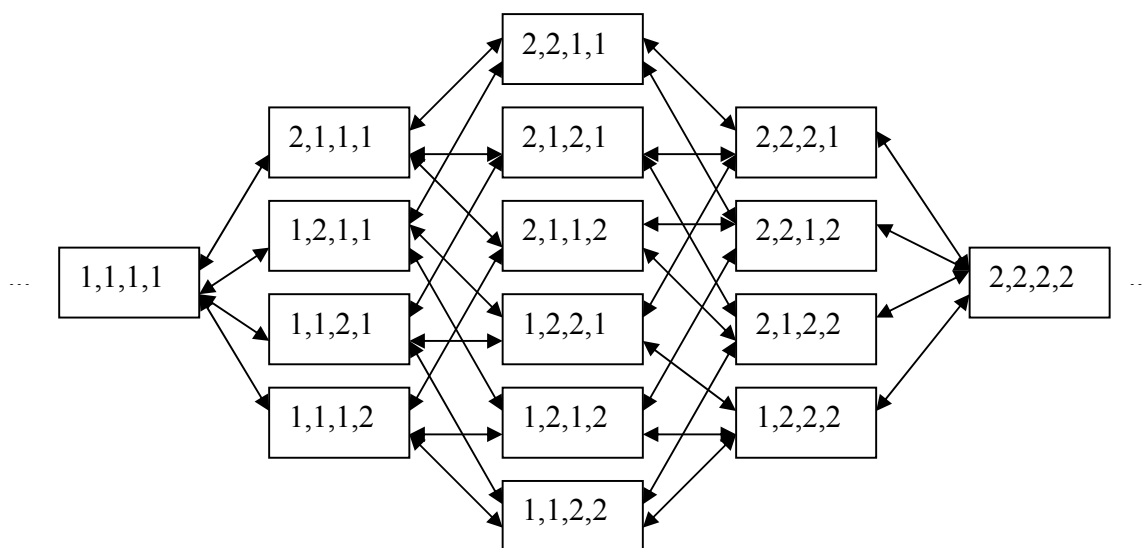


Figura 5.6. Diagrama parcial das possíveis mudanças nos estados, com $m=4$, $k=1$ e $N=3$.

No exemplo acima, a partir do estado $(1,1,2,1)$ pode-se alcançar os seguintes estados: $(1,1,1,1)$ com taxa 4μ (relativo a uma saída; note que $m^* = 1$), $(2,1,2,1)$ com taxa $\lambda/3$, $(1,2,2,1)$ com taxa $\lambda/3$ e $(1,1,2,2)$ com taxa $\lambda/3$ (relativos a uma chegada; note que $m^* = 3$), explicando o lado esquerdo da equação. Pelo lado direito, o estado $(1,1,2,1)$ pode ser obtido a partir do estado $(1,1,1,1)$ com taxa $\lambda/4$ (relativo a uma chegada; note que $m^* = 4$) e pelos estados $(2,1,2,1)$, $(1,2,2,1)$ e $(1,1,2,2)$, com as respectivas taxas $\mu + \mu/2 + \mu/2$, $\mu/2 + \mu + \mu/2$ e $\mu/2 + \mu/2 + \mu$ (ou 2μ , relativos a uma saída; note que $m^* = 2$). A igualdade estabelece o equilíbrio que deve existir no sistema.

Continuando este procedimento, são obtidas as respectivas equações (vide exemplo no apêndice E) para todos os estados possíveis, considerando-se ainda que

$\sum_{\substack{n_1, n_2, n_3, n_4 \\ \text{possíveis}}} p(n_1, n_2, n_3, n_4) = 1$. Com a resolução do sistema linear em $p(n_1, n_2, n_3, n_4)$, são

obtidas estas probabilidades e, assim, a distribuição de equilíbrio do sistema. Com esta, podem ser calculadas várias medidas de desempenho para o sistema. Por exemplo, o fator de utilização para o caixa i é obtido pela expressão:

$$\rho_i = \sum_{(n_1, n_2, \dots, n_m) \text{ para } n_i > 0} p(n_1, n_2, \dots, n_m) \quad (5.12)$$

ou seja, a soma de todas as probabilidades onde $n_i > 0$ (ocupado). Como exemplo, seja a situação em que $m=4$, $k=1$ e $N=2$. Pode-se calcular ρ_1 (o fator de utilização do primeiro caixa) somando-se todas as probabilidades $p(n_1, n_2, n_3, n_4)$ nas quais $n_1 > 0$. Alternativamente, para simplificação, pode-se também realizar este cálculo de forma complementar, ou seja, subtrair de 1 (a probabilidade total) todas as probabilidades $p(n_1, n_2, n_3, n_4)$ nas quais $n_1 = 0$. Então é obtido

$$\rho_1 = 1 - [p(0,0,0,0) + p(0,0,0,1) + p(0,1,0,0) + p(0,0,1,0) + p(0,0,1,1) + p(0,1,0,1) + p(0,1,1,0) + p(0,1,1,1)].$$

Também, o tempo médio de espera na fila do caixa i é obtido por:

$$W_{q_i} = \sum_{(n_1, n_2, \dots, n_i, \dots, n_m) \text{ para } n_i > 0} n_i p(n_1, n_2, \dots, n_i, \dots, n_m) / (\lambda_E / m) - \frac{1}{\mu} \quad (5.13)$$

Na expressão acima, o termo $\sum_{(n_1, n_2, \dots, n_i, \dots, n_m) \text{ para } n_i > 0} n_i p(n_1, n_2, \dots, n_i, \dots, n_m)$ fornece o número médio de usuários do caixa i (L_i), para o modelo em questão. A taxa de entrada $\lambda_E = \lambda(1 - p(n_1, n_2, \dots, n_m))$ é dividida por m , pois os clientes que chegam se distribuem pelos m caixas equitativamente. Pela lei de Little, a razão entre este número médio de usuários (L_i) e a taxa λ_E / m fornece o tempo médio de espera no caixa i (W_i) que, subtraído do tempo médio de serviço ($1/\mu$), leva a obter o tempo médio de espera na fila do caixa i (W_{q_i}).

O modelo com trocas mostra-se bastante flexível, sendo que o modelo $M/M/m$ com fila única e o modelo $[M/M/1]^m$ são casos particulares dele. Fazendo-se $k=1$ e $N \rightarrow \infty$ para o modelo com trocas, este se torna o modelo $M/M/m$ (ou $M/M/m/\infty$) com fila

única. Intuitivamente, se $N \rightarrow \infty$ então não há desistências e com $k=1$ as mudanças ocorrem com diferença mínima de clientes na fila, ocasionando maior agilidade e funcionamento muito próximo ao de uma fila única, inclusive tornando mais justo o sistema.

Também, fazendo-se $k \rightarrow \infty$ e $N \rightarrow \infty$ para o modelo com trocas e supondo que cada chegada escolhe aleatoriamente uma das filas para entrar, este se torna o modelo $[M/M/1]^m$. Intuitivamente, se $N \rightarrow \infty$ não há desistências e se $k \rightarrow \infty$ não há trocas, portanto, o sistema se torna igual a m filas $M/M/1$ paralelas e independentes. Poderia-se pensar ainda no modelo com trocas, retirando as trocas (com $k \rightarrow \infty$). Este modelo se torna então um sistema com m filas $M/M/1/N$ paralelas e independentes. De forma geral, pode-se observar que o modelo com trocas é bem abrangente, conforme as situações de k e N .

Para a utilização dos modelos de filas apresentados, é preciso que as suposições que os suportam sejam averiguadas. Para tanto, são apresentados alguns resultados sobre processo de Poisson e distribuições de probabilidade que devem ser verificadas. Também, é apresentada a análise de variância (ou ANOVA), com a finalidade de utilizá-la para a investigação da igualdade de médias de tempo de serviço para os distintos caixas considerados. A verificação de todas as suposições se fazem necessárias para aplicação dos modelos a serem comparados e aos resultados obtidos.

5.4 As suposições dos modelos

1- Para o processo de Poisson e suas partições

É conhecido que para um processo de Poisson com taxa λ , os correspondentes tempos entre chegadas são variáveis aleatórias exponenciais independentes e identicamente distribuídas com média $1/\lambda$. De maneira inversa, pode-se dizer que, se os tempos entre chegadas para um processo de chegadas são variáveis aleatórias exponenciais independentemente e identicamente distribuídas com média $1/\lambda$, então o processo em questão é Poisson com taxa λ (LAW e KELTON, 1991, p.406; ARENALES et al., 2006). Assim, averiguando-se a distribuição exponencial e supondo-se independência, pode-se aproximar a um processo de Poisson.

Um outro resultado importante é sobre a partição de um processo de Poisson. Inicialmente, seja o número de eventos que ocorrem em um dado período de tempo uma variável aleatória Poisson com parâmetro λ . Se cada evento é classificado como sendo do tipo i com probabilidade p_i , para $i=1,2,\dots,n$, com $\sum_{i=1}^n p_i=1$, e supondo que são independentes entre si, então o número de eventos do tipo i que ocorrem, para $i=1,2,\dots,n$, é uma variável aleatória Poisson com parâmetro respectivo λp_i , para $i=1,2,\dots,n$.

Como consequência deste fato, é obtido um resultado importante sobre a partição de um processo de Poisson. Seja um fluxo de chegadas com um processo de Poisson com taxa λ , e seja uma partição de forma supostamente independente em n categorias de acordo com as probabilidades p_1, p_2, \dots, p_n e $\sum_{i=1}^n p_i=1$. Então, para uma dada categoria i , com $i=1,2,\dots,n$, seu respectivo fluxo de chegadas é um processo de Poisson com taxa $p_i \lambda$ e independente dos outros $n-1$ processos (WOLFF, 1989, p.74). Desta forma, verificando que o fluxo de chegadas a ser distribuído (particionado) pelas filas é um processo de Poisson, todos os outros resultantes de partições independentes também o serão. Contudo, a dificuldade é verificar a independência, o que geralmente é suposto.

2- Para uniformidade das chegadas

Especialmente para o modelo com trocas é necessário testar se a distribuição dos clientes pelos caixas é equitativa. Desta forma, deve-se testar se ocorre uniformidade de chegadas nos caixas, tanto para os caixas rápidos como para os caixas normais analisados. Uma vez constatada a uniformidade, pode-se considerar que os clientes entram de forma relativamente equitativa nas filas, resultando que seus processos de Poisson devem ter taxas aproximadamente equivalentes.

3- Para distribuição exponencial dos tempos de serviço

Outra hipótese importante que deve ser testada é relativa à distribuição exponencial para os tempos de serviço dos caixas. Além disso, deve-se testar a igualdade de

suas médias, o que também é suposto pelos modelos.

4- Para igualdade de médias dos tempos de serviço

Pode-se utilizar a análise de variância (ANOVA) para testar a igualdade de médias dos tempos de serviço. O teste pode ser sumariado da forma a seguir. Seja $X_{l1}, X_{l2}, \dots, X_{ln_l}$ uma amostra aleatória de uma distribuição $N(\mu_l, \sigma^2)$, com $l = 1, 2, \dots, g$, com g amostras aleatórias independentes. Testa-se a hipótese $H_0: \mu_1 = \mu_2 = \dots = \mu_g$ com uma estatística F e com as quantidades presentes na tabela 5.1, a seguir. Rejeita-se H_0 a um nível de significância α se

$$F = \frac{SS_F / (g - 1)}{SS_{Res} / (\sum_{l=1}^g n_l - g)} > F_{g-1, \sum_{l=1}^g n_l - g}(\alpha) \quad (5.14)$$

Tabela 5.1. Quantidades calculadas para a ANOVA.

Fonte de variação	Soma de Quadrados	Graus de Liberdade
Fator	$SS_F = \sum_{l=1}^g n_l (\bar{x}_l - \bar{x})^2$	$g-1$
Resíduos	$SS_{Res} = \sum_{l=1}^g \sum_{j=1}^{n_l} (x_{lj} - \bar{x}_l)^2$	$\sum_{l=1}^g n_l - g$
Total	$SS_{Cor} = \sum_{l=1}^g \sum_{j=1}^{n_l} (x_{lj} - \bar{x})^2$	$\sum_{l=1}^g n_l - 1$

A condição de normalidade pode ser relaxada quando os valores de n_l são relativamente grandes. No caso dos caixas normais, como exemplo, há oito amostras aleatórias independentes e a hipótese testada é $H_0: \mu_1 = \mu_2 = \dots = \mu_8$, ou seja, se as médias dos tempos de serviço para cada caixa podem ser consideradas como iguais.

Para testar as suposições que envolvem distribuições de probabilidade (como para a distribuição de probabilidade dos tempos de serviço), foi utilizado o *software* BestFit 4.0 da Palisade Tools (BESTFIT, 2000). Este *software* fornece a distribuição que melhor se ajusta aos dados com base nos testes não-paramétricos Qui-Quadrado, Kolmogorov-Smirnov e Anderson-Darling (CONOVER, 1999). Desta forma, os testes podem ser realizados obtendo-se relativa segurança nas conclusões sobre as distribuições de probabilidade admitidas nos modelos.

5.5 Algumas considerações e verificação das suposições

Como descrito no capítulo 3, seções 3.3 e 3.4, os dados utilizados se referem a seis amostras, em dois finais de semana, para um supermercado na cidade de São Carlos. São aqui analisadas algumas características das amostras relativas ao primeiro final de semana, nos dias 03, 04 e 05 de dezembro de 2004, conforme horários especificados na tabela 3.1. Também, são relatadas as principais ocorrências em relação à verificação das suposições para a segunda semana de coleta, nos dias 10, 11 e 12 de dezembro. Deve-se lembrar que foram considerados oito caixas normais e quatro caixas rápidos, sendo que um destes últimos era preferencial, mas, como observado anteriormente, funcionou praticamente como um caixa rápido (vide figura 3.3).

Por meio da planilha (apêndice B) preenchida pelos coletores, algumas informações importantes foram obtidas, das quais algumas essenciais para a aplicação dos modelos em questão. Na tabela 5.2 pode-se observar, além das taxas de serviço, a taxa de entradas no sistema (λ_E), a proporção de servidores ocupados ($\lambda_E / m\mu$), o tamanho amostral e o tempo médio de espera na fila (denominado como valor observado), para caixas normais e rápidos, nas duas semanas de coleta. A taxa de chegadas é formada pelos clientes que entram nas filas e pelos clientes que desistem das compras devido ao número de clientes nas filas, sendo que estes últimos não puderam ser observados. Nota-se ainda, que apesar dos valores altos de $\lambda_E / m\mu$, o tempo médio de espera na fila não é grande para os dois tipos de caixa considerados.

É possível estimar a taxa de chegadas (λ) por meio de um processo iterativo utilizando-se a taxa de entradas (λ_E) e a probabilidade de perda (p_{mN}), quando se utiliza

um modelo com número de clientes na fila limitado (N), como é o caso do modelo com trocas. Esta probabilidade se refere ao evento do sistema estar lotado, podendo haver então perda e é calculada de forma relativa ao modelo que está se utilizando. Também, se a probabilidade de perda é muito pequena, este processo não se faz necessário.

Tabela 5.2. Taxas de entradas e serviço, a proporção de servidores ocupados, tamanho amostral e tempo médio de espera na fila para caixas normais e rápidos, nas duas semanas de coleta.

Dia	Tipo de caixa	λ_E clientes/min	μ clientes/min	$\lambda_E / m\mu$	Total das amostras	Valor observado (min.)
03/12	Normal	1,6039	0,2384	0,841	194	1,80
	Rápido	1,8196	0,6250	0,728	220	1,43
04/12	Normal	1,3182	0,1880	0,878	159	3,36
	Rápido	2,3785	0,6693	0,888	287	1,82
05/12	Normal	1,6058	0,2336	0,859	195	1,85
	Rápido	2,0803	0,6228	0,835	253	1,61
10/12	Normal	1,4174	0,2182	0,812	179	1,39
	Rápido	1,7311	0,6130	0,706	209	1,24
11/12	Normal	1,3429	0,1901	0,883	168	3,63
	Rápido	2,4239	0,6718	0,902	297	2,19
12/12	Normal	1,6099	0,2254	0,893	202	2,92
	Rápido	2,2399	0,6272	0,893	266	2,64

O processo iterativo pode ser descrito em três passos (FONTANELLA e MORABITO, 2002):

Passo 1: Faça $\lambda^0 = \lambda_E$ como sendo a taxa de chegadas média inicial e faça $n=1$.

Passo 2: Calcule a probabilidade de perda p_{mN}^n e a taxa $\lambda_E^n = \lambda^{n-1}(1 - p_{mN}^n)$, onde p_{mN}^n é obtido do modelo considerado, utilizando-se λ^{n-1} , m , μ e N .

Passo 3: Se λ_E e λ_E^n estiverem suficientemente próximos, então pare com $\lambda = \lambda^{n-1}$. Caso contrário, faça $\lambda^n = \lambda^{n-1} + (\lambda_E - \lambda_E^n)$, $n = n+1$, e volte para o passo 2.

Observa-se que a taxa de chegadas é obtida em função da taxa de entradas e da probabilidade de perda relativa ao modelo utilizado, incorporando os clientes que devem ter sido perdidos pela limitação da fila. É realizada uma análise mais detalhada para a amostra coletada no dia 05 de dezembro, para os caixas normais, principalmente no que for relativa a verificação das suposições para aplicação de todos os modelos considerados. Inicialmente é necessário que os tempos entre chegadas para todos os clientes, antes de se distribuírem pelos caixas, tenha uma distribuição exponencial. Utilizou-se o *software* BestFit 4.0 e este apontou, para os testes Qui-Quadrado, Kolmogorov-Smirnov e Anderson-Darling, com pelo menos 5% de significância, não rejeitar a hipótese de distribuição exponencial.

A figura 5.7 ilustra os dados observados (em colunas) e a curva obtida do ajuste da distribuição exponencial. Observando-a nota-se grande semelhança dos dados observados com a distribuição exponencial. Desta forma, pode-se assumir um processo de Poisson para o processo de chegadas e, conseqüentemente, para os processos derivados para cada um dos oito caixas normais, conforme resultados explanados anteriormente.

Pode-se agora testar a uniformidade para o número de clientes que chegam em cada um dos caixas. Novamente, utilizando-se o BestFit e o teste Qui-Quadrado, não se pode rejeitar, com 5% de significância, esta hipótese. A figura 5.8 denota uma uniforme ajustada ao número de chegadas para os oito caixas normais (com os caixas de 12 a 05 renomeados respectivamente como de 01 a 08). Assim, pode-se considerar que os clientes entram de forma relativamente equitativa pelas filas dos caixas normais.

Observando-se ainda a figura 5.8, nota-se que a fila onze (renomeada na figura como 2) tem uma frequência bem abaixo do esperado (em vermelho). Ainda que este fato não tenha levado a rejeição da uniformidade, sua causa foi investigada. Isto ocorreu porque um cliente fez uma compra muito grande, com tempo de serviço de aproximadamente 28 minutos, fazendo com que outros clientes evitassem entrar nesta fila. Também, isto fez com que houvesse a rejeição da igualdade das médias dos tempos de serviço (testada pela ANOVA), bem como da distribuição exponencial (testada da mesma forma que dos tempos entre chegadas) para este caixa.

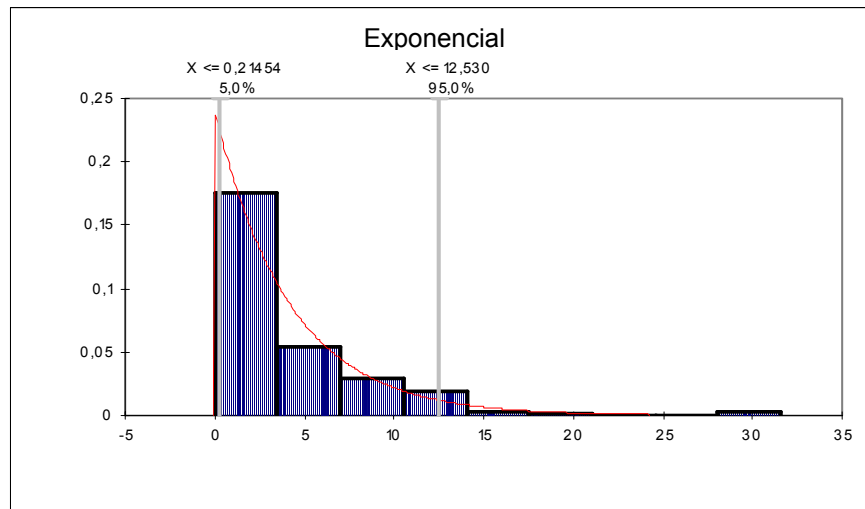


Figura 5.7. Ajuste para os tempos entre chegadas de todos os clientes nos caixas normais no dia 5 de dezembro.

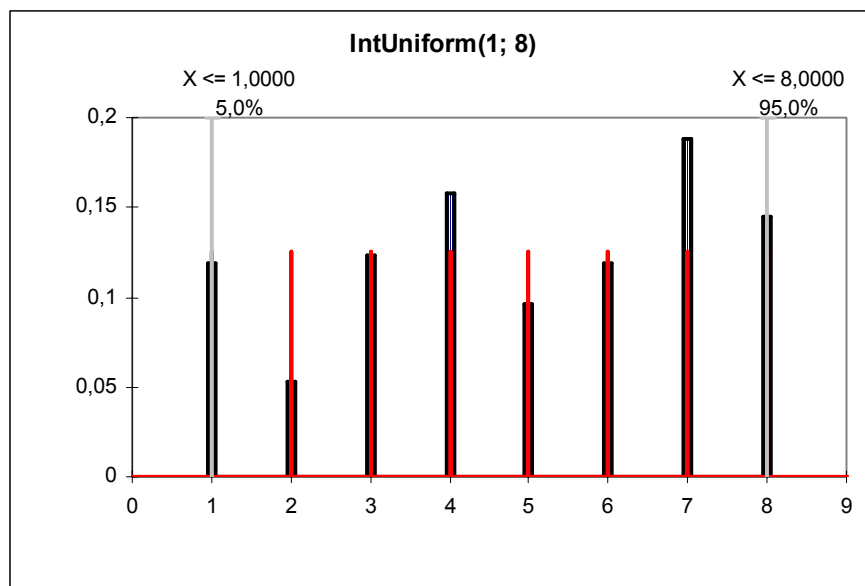


Figura 5.8. Ajuste uniforme para o número de clientes que entram nos caixas normais no dia 5 de dezembro.

Retirando-se o ponto da amostra (considerado um *outlier*), ambas as hipóteses não foram mais rejeitadas, considerando um nível de significância de 5%. Como ilustração, a figura 5.9 apresenta o ajuste realizado para a distribuição exponencial para os tempos de serviço do caixa normal 4, no dia 5 de dezembro.

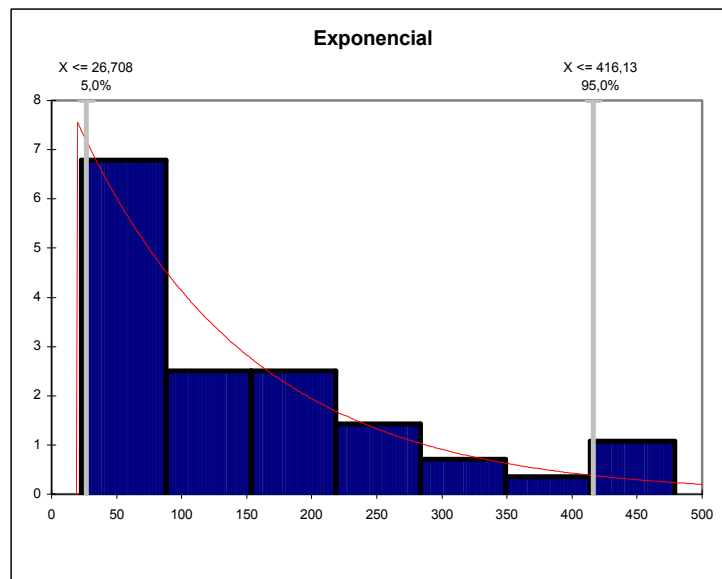


Figura 5.9. Ajuste realizado para os tempos de serviço do caixa normal 4, no dia 5 de dezembro, obtendo uma distribuição exponencial.

Apesar do ponto (*outlier*), quando retirado, deixar todas as suposições satisfeitas, este foi mantido na amostra, pois esta é uma situação que de fato ocorre em se tratando de filas em supermercados. Também, foi possível com isto testar se os modelos são robustos a estas ocorrências.

O total das amostras para os caixas normais e rápidos (tabela 5.2) revela que o número de clientes que passam pelos caixas rápidos é sensivelmente maior que para os caixas normais. Estes então se distribuem por quatro caixas e, com poucos itens, experimentam pequenos tempos de fila e serviço. Nos caixas rápidos, com alto fluxo de clientes e tempos coesos, pode-se verificar que o processo de chegadas é Poisson (vide figura 5.10, para a distribuição exponencial), o número de clientes pelos caixas é uniforme (vide figura 5.11) e que os tempos de serviço são exponenciais (como exemplo, vide figura 5.12), com médias consideradas idênticas pela ANOVA (como exemplo, vide tabela 5.3),

todos com 5% de significância. Igualmente, estas conclusões são obtidas para os três dias de coleta. Provavelmente isto ocorre devido ao tamanho amostral e a limitação do número de itens em dez unidades, tornando os dados bem consistentes.

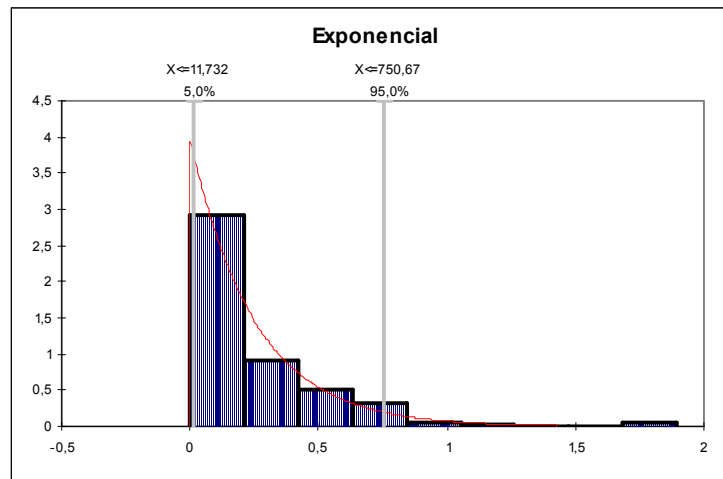


Figura 5.10. Ajuste realizado para todos os tempos entre chegadas nos caixas rápidos, no dia 5 de dezembro, obtendo uma distribuição exponencial.

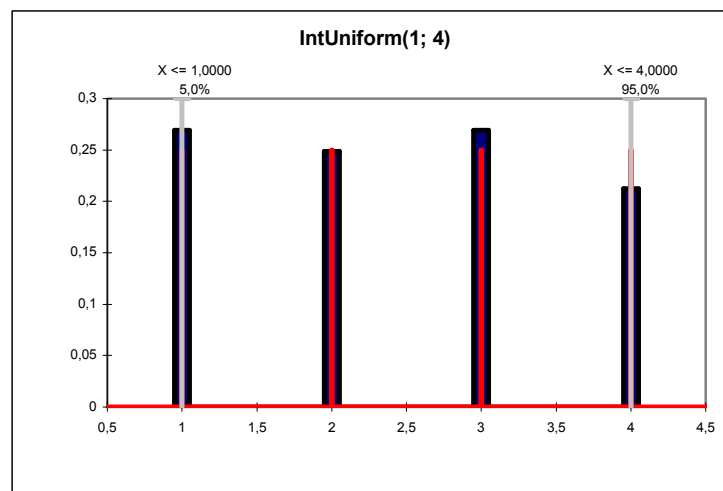


Figura 5.11. Ajuste uniforme para o número de clientes que entram nos caixas rápidos, no dia 5 de dezembro.

Deve-se investigar ainda o comportamento das taxas de entrada no que se refere a sua estabilização no decorrer do tempo em que a amostra foi coletada, tanto para os caixas normais como para os rápidos. O período de coleta de 2 horas foi dividido em quatro

partes e as taxas foram recalculadas, para cada parte, e dispostos em dois gráficos (figuras 5.13 e 5.14), para analisar seus comportamentos.

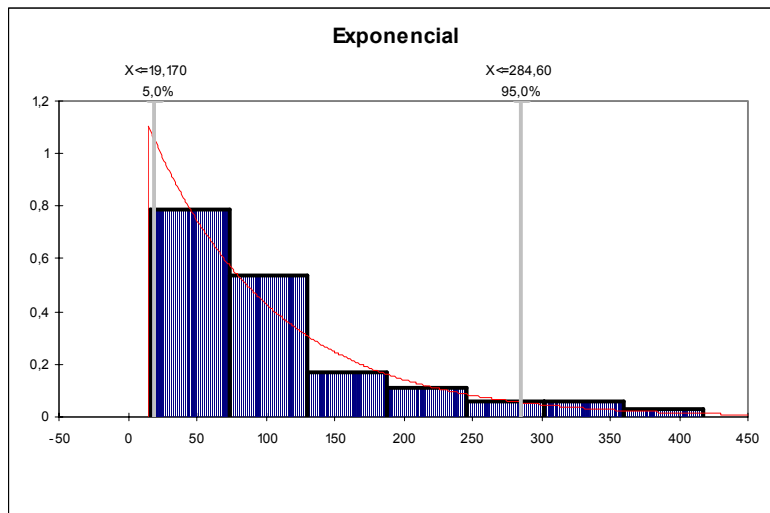


Figura 5.12. Ajuste exponencial dos tempos de serviço no caixa rápido 13, no dia 5 de dezembro.

Tabela 5.3. Quantidades calculadas para a ANOVA para as médias de tempos de serviço, no dia 5 de dezembro.

Fonte de variação	Soma de Quadrados	Graus de Liberdade
Fator	$SS_F = 19974,61592$	$g-1=3$
Resíduos	$SS_{Res} = 29847,39348$	$\sum_{l=1}^g n_l - g = 191$
Total	$SS_{Cor} = 49822,0094$	$\sum_{l=1}^g n_l - 1 = 194$
Conclusão	$F = \frac{SS_F / 3}{SS_{Res} / 191} = 2,06 < F_{3,191}(5\%) = 2,60$	Não se rejeita $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$

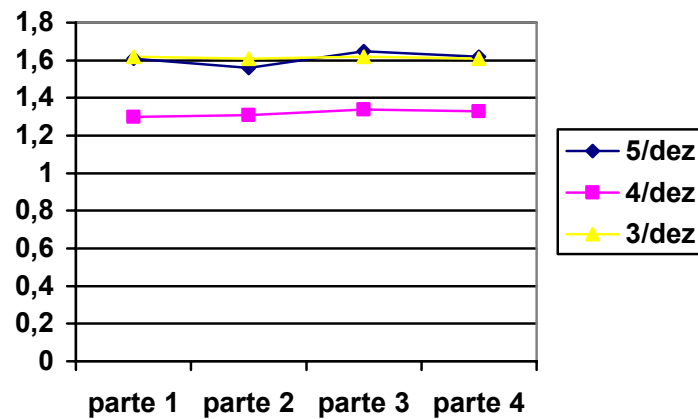


Figura 5.13. Variação da taxa de entrada para os caixas normais nos períodos de coleta.

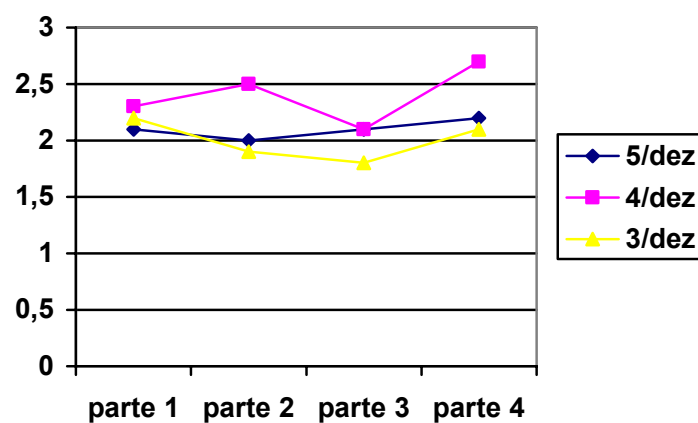


Figura 5.14. Variação da taxa de entrada para os caixas rápidos nos períodos de coleta.

Pela figura 5.13 pode-se perceber que existe um comportamento um pouco distinto para o dia 4 de dezembro, pois este apresenta uma taxa de entrada relativamente inferior aos demais dias. Também, analisando-se o comportamento no decorrer no período de duas horas em suas quatro divisões (partes 1,2,3 e 4 na figura 5.13), pode-se considerar que não houve variação significativa das taxas durante o período de coleta. Caso contrário, poderia-se tratar de um processo não estacionário, onde a taxa seria da forma $\lambda_E(t)$. Observando-se a figura 5.14 nota-se que existe um comportamento com relativa variação,

para a taxa de entrada dos caixas rápidos, somente para o dia 4 de dezembro, um sábado. Apesar disto, esta parece não ter tendência em estar apenas crescendo ou apenas declinando no decorrer do período. Desta forma, pode-se considerar as taxas relativamente estáveis para o período de coleta.

As ocorrências havidas na segunda semana de coleta, referentes aos dias 10, 11 e 12 de dezembro, foram com menor intensidade de freqüência que na primeira. Em um dos caixas normais, no dia 11, foi atendido um cliente com muitos itens, o que levou a um tempo de serviço alto, ocasionando o mesmo problema que para um caixa normal do dia 5. Também, todos os caixas rápidos tiveram um bom comportamento em relação às suposições e as taxas de entrada apresentaram pouca variação no período de coleta. Em resumo, uma situação muito parecida com a primeira semana.

Poder-se-ia pensar na questão de analisar o impacto que as trocas de filas têm sobre o comportamento do cliente no supermercado, como também nos modelos que as utilizam como um fator diferenciador. É necessário verificar o quão freqüente o cliente realiza a troca, e se este realmente é um fator que deve ser englobado nas análises como importante. No supermercado estudado isto foi obtido de duas formas: perguntando ao cliente em um questionário, e observando e anotando as trocas nas filas dos caixas, sendo esta última apresentada na tabela 5.4.

Tabela 5.4. Porcentagens de trocas de filas observadas segundo o tipo de caixa e dias de coleta.

Dia	Caixas Normais	Caixas Rápidos
3/12	4,2%	9,4%
4/12	7%	12,1%
5/12	8,5%	10,5%

O resultado para as porcentagens do número de trocas de filas, em primeira análise, não parece muito consistente com o tipo de caixa na qual está o cliente. Nos caixas normais onde, em média, os tempos de fila e em serviço são maiores que nos caixas rápidos, os clientes poderiam obter considerável diminuição do tempo total de espera efetuando a troca para uma fila menor. Contudo, as filas dos caixas rápidos têm um menor tempo médio de espera, ou seja, existe mais dinamismo (mais trocas de estados) em suas

filas e as oportunidades de trocas podem aparecer, com maior frequência, sendo uma possibilidade de explicação ao resultado obtido.

Considerando-se o total de trocas (para todos os dias), a porcentagem aferida na coleta de dados observando-se as filas foi de 8,3%. No questionário aplicado aos clientes (apêndice A), também foi perguntado se estes haviam realizado alguma troca de filas e a porcentagem obtida foi de 8,2%, revelando muita coerência nas informações prestadas e as observadas nas filas. Desta forma, pode-se considerar que para o cliente, pelo fato de lembrar-se devidamente das trocas, estas podem ser interessantes. Acredita-se, então, que as trocas são importantes como reguladoras do tamanho das filas, uniformizando-as. No momento que o cliente chega e se junta à fila mais conveniente e/ou quando realiza uma troca, ele faz com que as filas não tenham tamanhos tão distintos, diferenciando o modelo que as engloba dos demais.

5.6 Aplicação dos modelos freqüentemente utilizados

Apesar da importância que há em tratar as trocas entre filas, podem ser considerados outros modelos de filas que não privilegiam esta ocorrência. Alguns deles foram citados anteriormente e são agora aplicados para os dados coletados. Nas tabelas 5.5 e 5.6 encontram-se os resultados para os modelos de filas $M/M/m$ e $[M/M/1]^m$, para os mesmos dias e parâmetros dos dias de coleta em duas semanas, constando do tempo médio de espera na fila (W_q) para cada modelo com seus respectivos desvios (desvio da razão entre o valor modelado e o observado, em porcentagem) em relação ao valor observado na amostra. Pode-se notar a diferença existente nas estimativas entre a fila $M/M/m$ (com fila única) e a fila $[M/M/1]^m$ (com filas paralelas independentes). Neste caso, desconsiderar a dependência existente nas filas traz grande vício na estimação dos tempos de espera na fila. Desta forma, ainda que com estimativas alternantes em sua qualidade, o modelo de fila única tem melhores resultados. Nota-se nas tabelas 5.5 e 5.6 que o modelo $[M/M/1]^m$ em geral superestima o tempo de espera na fila. Deve-se ressaltar ainda que este modelo ($M/M/m$) é um caso particular do modelo com trocas quando $k=1$ e $N \rightarrow \infty$ para o primeiro momento (em média) (KOENIGSBERG, 1966).

Tabela 5.5. Resultados para as filas $M/M/m$ e $[M / M / 1]^m$, para a primeira semana de coleta.

ρ	Dia	Caixa	$M/M/m$	$[M / M / 1]^m$	Valor Observado (minutos)
			W_q desvio	W_q desvio	
0,841	03/12	N	1,8159 0,9%	22,18 1232,3%	1,80
0,728	03/12	R	0,6957 48,7%	4,2824 299,5%	1,43
0,878	04/12	N	3,4397 2,4%	37,7379 1123,2%	3,36
0,888	04/12	R	2,5594 40,6%	11,8973 653,7%	1,82
0,859	05/12	N	2,2640 22,4%	26,1374 1412,8%	1,85
0,835	05/12	R	1,6086 0,1%	8,1291 504,9%	1,61

Tabela 5.6. Resultados para as filas $M/M/m$ e $[M / M / 1]^m$, para a segunda semana de coleta.

ρ	Dia	Caixa	$M/M/m$	$[M / M / 1]^m$	Valor Observado (minutos)
			W_q desvio	W_q desvio	
0,812	10/12	N	1,4744 6,1%	19,7924 1423,9%	1,39
0,706	10/12	R	0,6076 -51,0%	3,9173 315,9%	1,24
0,883	11/12	N	3,6885 1,6%	39,7087 1093,9%	3,63
0,902	11/12	R	3,0074 37,3%	13,7033 625,7%	2,19
0,893	12/12	N	3,5288 20,8%	36,9499 1265,4%	2,92
0,893	12/12	R	2,8757 8,9%	13,2810 503,1%	2,64

Uma outra possibilidade é realizar o mesmo experimento considerando-se uma capacidade máxima para as mesmas filas em questão. É natural, então, que esta capacidade seja equivalente ao que é considerado no modelo com trocas, para efeito de comparação. É importante ressaltar que existe certo subjetivismo, para o modelo com trocas, em fixar o

valor de N , ou seja, em considerar a capacidade máxima das filas. No caso do supermercado, como discutido anteriormente, os clientes ficam em pé com suas compras em carrinhos ocupando um espaço físico considerável, dependendo do tamanho da fila. Também, de acordo com informações obtidas do gerente daquele supermercado, nos caixas normais, os clientes reclamam quando as filas têm mais de três pessoas. Observando-se as planilhas e as filmagens realizadas para a coleta de amostras, pôde-se notar que o número máximo de pessoas nos caixas rápidos e normais, ainda que em raras oportunidades, foram respectivamente 7 e 5. Os resultados podem ser vistos nas tabelas 5.7 e 5.8. Ali constam os valores de W_q e de seu desvio, a probabilidade do sistema estar saturado (p_{mN} , ou perda) e de estar completamente vazio (p_0 , ou vazio).

Para o modelo $M/M/m/mN$ foram utilizadas as taxas de entrada (λ_E), pois as probabilidades de perda são muito pequenas (vide tabelas 5.7 e 5.8). Desta forma, o procedimento iterativo descrito anteriormente foi desnecessário. No modelo $[M/M/1/N]^m$, as probabilidades de perda não são pequenas, contudo, como pode ser notado, todas as estimativas para este modelo estão bem acima do valor observado. Então, aplicar o procedimento iterativo só iria aumentá-las.

Tabela 5.7. Resultados para as filas $M/M/m/mN$ e $[M/M/1/N]^m$ para a primeira semana de coleta.

$\frac{\lambda_E}{m\mu}$	Dia	Caixa	$M/M/m/mN$			$[M/M/1/N]^m$			Valor Observado (minutos)
			W_q desvio	Perda p_{mN}	Vazio p_0	W_q desvio	Perda p_{mN}	Vazio p_0	
0,841	03/12	N	1,57 -12,8%	0,00034	0,00084	6,95 286%	0,10352	0,24606	1,80
0,728	03/12	R	0,69 -51,7%	0,00006	0,04296	2,92 104%	0,03196	0,29546	1,43
0,878	04/12	N	3,27 -2,7%	0,00123	0,00054	9,24 175%	0,11689	0,22596	3,36
0,888	04/12	R	2,32 27,5%	0,00502	0,01309	4,44 144%	0,07965	0,18236	1,82
0,859	05/12	N	2,17 17,3%	0,00065	0,00068	7,26 292%	0,11004	0,23597	1,85
0,835	05/12	R	1,51 -6,2%	0,00145	0,02118	3,69 129%	0,06118	0,21599	1,61

Tabela 5.8. Resultados para as filas $M/M/m/mN$ e $[M/M/1/N]^m$ para a segunda semana de coleta.

$\frac{\lambda_E}{m\mu}$	Dia	Caixa	$M/M/m/mN$			$[M/M/1/N]^m$			Valor Observado (minutos)
			W_q desvio	Perda p_{mN}	Vazio p_0	W_q desvio	Perda p_{mN}	Vazio p_0	
0,812	10/12	N	1,46 5,3%	0,00116	0,00116	7,29 524,5%	0,09303	0,26354	1,39
0,706	10/12	R	0,61 -50,8%	0,00003	0,04860	2,82 127,4%	0,02739	0,31334	1,24
0,883	11/12	N	3,40 -6,3%	0,00145	0,00050	9,22 154,0%	0,11940	0,22243	3,63
0,902	11/12	R	2,62 19,6%	0,00670	0,01129	3,86 76,2%	0,08474	0,17443	2,19
0,893	12/12	N	3,17 8,6%	0,00199	0,00044	7,87 169,5%	0,12320	0,21720	2,92
0,893	12/12	R	2,58 -2,3%	0,00574	0,01279	4,07 54,2%	0,08128	0,17977	2,64

Novamente, como para o caso anterior, o modelo que trata o sistema como com filas paralelas independentes tem o problema de superestimar o tempo de espera na fila. Também o modelo $M/M/m/mN$ tem alternância de estimativas (ainda que mais suaves), assim como ocorria com o modelo $M/M/m$. Contudo, nota-se que os desvios para o modelo $[M/M/1/N]^m$ estão bem menores que para o modelo $[M/M/1]^m$, ainda que sejam grandes.

Uma outra possibilidade de abordagem do problema é utilizar filas com distribuições genéricas. No presente caso pode-se pensar em m filas $M/G/1$ ($[M/G/1]^m$) e uma fila $M/G/m$. Com as expressões 5.8 e 5.9 e a fórmula de Little, foram obtidos os tempos médios de espera na fila para os dias de coleta, que podem ser vistos na tabela 5.9 a seguir. Pelos resultados obtidos, percebe-se que os modelos produziram estimativas muito ruins, principalmente para as filas paralelas e independentes, pois todas elas superestimaram em muito o valor observado, sendo que o menor desvio obtido foi de 10,4%. Assim, pode-se pensar que, neste caso, estes modelos podem ser descartados como uma alternativa em aproximar a situação das filas no supermercado.

Tabela 5.9. Resultados para as filas $M/G/m$ e $[M/G/1]^m$ para os dias de coleta.

Dia	ρ	Caixa	$M/G/m$ desvio	$[M/G/1]^m$ desvio	Valor observado (min.)
03/12	0,841	N	3,99 121,7%	30,38 1587,8%	1,80
	0,728	R	1,58 10,5%	5,34 273,4%	1,43
04/12	0,878	N	5,13 52,7%	43,24 1186,9%	3,36
	0,888	R	6,21 241,2%	11,22 516,5%	1,82
05/12	0,859	N	4,03 117,8%	24,61 1230,3%	1,85
	0,835	R	3,16 96,3%	8,30 415,5%	1,61
10/12	0,812	N	2,19 57,6%	19,69 1316,5%	1,39
	0,706	R	1,43 15,3%	5,32 329,0%	1,24
11/12	0,883	N	4,35 19,8%	34,86 860,3%	3,63
	0,902	R	8,26 277,2%	15,36 601,4%	2,19
12/12	0,893	N	8,96 206,8%	39,88 1265,7%	2,92
	0,893	R	5,98 126,5%	9,25 250,4%	2,64

5.7 Aplicação do modelo com trocas

Após admitir como satisfeitas todas as suposições para a aplicação do modelo com trocas, deve-se investigar se este é capaz de estimar satisfatoriamente o tempo médio de espera na fila, observado na amostra. O modelo foi implementado em linguagem Pascal e compilador Delphi, utilizando-se um computador Athlon 3500 de 64 bits com 2 Gbytes de memória. Os sistemas lineares gerados são resolvidos por meio do método de Gauss-Jordan e podem ter milhares de equações. Por exemplo, o modelo com $m=8$, $k=2$ e $N=3$ envolve 12866 equações, cuja geração e resolução consumiu mais de 3 horas. Por outro lado, sistemas com cerca de 6000 equações foram resolvidos em 29 minutos, e sistemas menores, com 1000 equações, consumiram apenas 10 segundos para serem resolvidos.

Para o modelo com trocas foi utilizado o processo iterativo para estimar as taxas de chegadas (λ) a partir das taxas de entradas, para que o resultado final pudesse ser mais fidedigno.

A amostra coletada no dia 3 (vide tabela 5.2) apresentou dois pontos que podem ser considerados como *outliers*: um tempo de serviço de aproximadamente 22,6 minutos no caixa de número 13 e outro, também em tempo de serviço, aproximadamente de 22,5 minutos no caixa 10. Nota-se que se tratam de caixas normais, onde algumas vezes se apresentam compras com grande número de itens. Quando os pontos são retirados, as suposições que haviam sido rejeitadas (igualdade das médias e distribuição exponencial para os tempos de serviço) são verificadas, podendo-se aplicar os modelos. Também, como realizado anteriormente, os pontos foram mantidos com a finalidade de avaliar a robustez dos modelos a variações nos parâmetros (taxa de serviços). As demais suposições, como processo de chegadas e uniformidade de entradas nos caixas, não apresentaram problemas de rejeição.

Na tabela 5.10 pode-se observar que, para os caixas normais, o desvio observado obtido quando $k=1$ e $N=5$ (valor de N como comentado anteriormente) é de apenas -1,2%, o que é um resultado bastante interessante e equivalente ao modelo $M/M/m$ que obteve um desvio de 0,9%. Poder-se-ia pensar em $k=2$, contudo, pela última linha da tabela pode-se observar que o resultado fica superestimado, mesmo para $N=3$.

Tabela 5.10. Medidas para o modelo com trocas, com 8 filas de caixa normal sendo $\lambda_E = 1,6039$, $\mu = 0,2384$, com tempo observado de espera na fila na amostra, de 1,80 minutos e $\lambda_E / m\mu = 0,841$, para o dia 3 de dezembro.

λ	k	N	W_q	Desvio	Perda	Vazio	Ocioso
1,6061	1	4	1,7028	-5,4%	0,0014	0,0008	0,4499
1,6044	1	5	1,7790	-1,2%	0,0003	0,0008	0,4494
1,6041	1	6	1,8052	0,3%	0,0001	0,0008	0,4492
1,6232	2	3	3,1639	75,8%	0,0119	0,0001	0,5312

Diferentemente dos caixas normais, uma conclusão distinta é obtida para os caixas rápidos (tabela 5.11 a seguir), com $k=2$ e $N=7$ (valor de N como comentado

anteriormente). Observa-se que os clientes têm maior tolerância nestes caixas ($k=2$) que a averiguada nos caixas normais. Contudo, o desvio aqui obtido é -11,9%, um valor razoável como estimativa do observado. Ainda assim, é equivalente ao obtido pelo modelo $M/G/m$ (10,5%) e bem menor que para os modelos $M/M/m$ e $M/M/m/mN$.

Nota-se que os desvios são, em quase todos os casos, menores que os obtidos nas tabelas 5.5, 5.7 e 5.9, ainda que para o modelo $M/M/m$ o desvio para os caixas normais seja de apenas 0,9%. Observa-se também que, devido aos valores menores de $\lambda_E / m\mu$ neste dia, existe menor sensibilidade quando N é alterado para $N-1$ ou $N+1$, tanto para os caixas normais quanto para os rápidos.

Tabela 5.11. Medidas para o modelo com trocas, com 4 filas de caixa rápido sendo $\lambda_E = 1,8196$, $\mu = 0,6250$, com tempo observado de espera na fila na amostra, de 1,43 minutos e $\lambda_E / m\mu = 0,728$, para o dia 3 de dezembro.

λ	k	N	W_q	Desvio	Perda	Vazio	Ocioso
1,8200	1	6	0,6880	-51,9%	0,0002	0,0430	0,5272
1,8198	1	7	0,6924	-51,5%	0,0001	0,0430	0,5272
1,8196	1	8	0,6941	-51,5%	0,0000	0,0430	0,5272
1,8201	2	6	1,2531	-12,4%	0,0003	0,0225	0,5982
1,8198	2	7	1,2600	-11,9%	0,0001	0,0225	0,5980
1,8196	2	8	1,2620	-11,7%	0,0001	0,0225	0,5980
1,8196	3	7	1,7656	23,5%	0,0001	0,0147	0,6401

Analisando-se a amostra do dia 04/12 (vide tabela 5.2), ocorreu *outlier*, assim como no dia 03/12. Houve um tempo de serviço de aproximadamente 33 minutos no caixa normal de número 10. Da mesma forma anterior, as suposições de igualdade das médias dos tempos de serviço entre os caixas e da distribuição exponencial do caixa 10 foram rejeitadas em função deste ponto. Uma vez retirado, as suposições foram verificadas com 5% de significância. Contudo, o ponto foi mantido na amostra, pelos mesmos motivos citados anteriormente. Quanto ao processo de chegadas e a uniformidade de entrada nos caixas, as conclusões são as mesmas da amostra anterior. Para os caixas rápidos, também como citado anteriormente, todas as suposições foram verificadas.

Na tabela 5.12 a seguir, pode-se observar que, para os caixas normais do dia 4 de dezembro, o desvio do observado obtido quando $k=1$ e $N=5$ é de apenas -3,5%, portanto um resultado bastante interessante apesar dos desvios dos modelos $M/M/m$ (2,4%) e $M/M/m/mN$ (-2,7%) estarem um pouco melhor.

. Contudo, a mesma conclusão não é obtida para os caixas rápidos (tabela 5.13), com $k=1$ e $N=7$, pois seu desvio foi de 18,8%, um resultado pouco razoável, apesar de ser o menor para todos os modelos utilizados. Neste dia, um sábado, foram verificadas filas maiores, o que é corroborado pelos valores de $\lambda_E / m\mu$ estarem mais elevados. Assim sendo, os clientes, de forma geral, estavam com pouca tolerância à diferença de tamanho de filas. Nota-se ainda que, para caixas normais e rápidos, existe alguma sensibilidade quando N é alterado para $N-1$ ou $N+1$, provavelmente devido ao alto valor de $\lambda_E / m\mu$.

Tabela 5.12. Medidas para o modelo com trocas, com 8 filas de caixa normal sendo $\lambda_E = 1,3182$, $\mu = 0,1880$, com tempo observado de espera na fila na amostra, de 3,36 minutos e $\lambda_E / m\mu = 0,878$, para o dia 4 de dezembro.

λ	k	N	W_q	Desvio	Perda	Vazio	Ocioso
1,3231	1	4	2,9971	-10,8%	0,0037	0,0005	0,3620
1,3198	1	5	3,2440	-3,5%	0,0012	0,0005	0,3613
1,3187	1	6	3,3548	-0,2%	0,0004	0,0005	0,3611
1,3402	2	3	4,3796	30,3%	0,0200	0,0001	0,4429

Tabela 5.13. Medidas para o modelo com trocas, com 4 filas de caixa rápido sendo $\lambda_E = 2,3785$, $\mu = 0,6693$, com tempo observado de espera na fila na amostra, de 1,82 minutos e $\lambda_E / m\mu = 0,888$, para o dia 4 de dezembro.

λ	k	N	W_q	Desvio	Perda	Vazio	Ocioso
2,4020	1	6	1,9961	9,7%	0,0100	0,0126	0,2369
2,3921	1	7	2,1615	18,8%	0,0058	0,0127	0,2363
2,3866	1	8	2,2802	25,3%	0,0034	0,0128	0,2359
2,3930	2	7	2,6320	44,6%	0,0069	0,0054	0,2765

Na tabela 5.14 abaixo pode-se notar que, para os caixas normais do dia 5 de dezembro, o desvio do tempo observado obtido é de 18,1% (de 1,85 para 2,19 minutos) quando o número máximo de clientes na fila (N) é 5 e a diferença entre filas (k) é 1. As probabilidades de haver perda no sistema e do sistema ficar completamente vazio são pequenas e a ociosidade (probabilidade de pelo menos um caixa vazio, denotada por ocioso) está em cerca de 0,4. Como de fato ocorreu neste dia, em raras situações a fila era maior que três pessoas, pois, como ressaltado, isto poderia gerar reclamações pelos clientes. Também, pelos resultados do modelo, estes parecem ter pouca tolerância à diferença de tamanho de filas, efetuando a troca ainda que vá diminuir apenas uma posição à sua frente. Os resultados com $k=2$ e $N=3$ ampliam os desvios obtidos aqui, podendo então ser uma situação descartada para os caixas normais neste dia, já que, com $k=2$ e conforme N cresce, os desvios tendem a aumentar.

Observa-se também que existe alguma sensibilidade do modelo quando $N=5$ é alterado para 6 ou 4, contudo, provavelmente isto se deve ao alto valor de $\lambda_E / m\mu$. Apesar do desvio ser pouco razoável (18,1%), nenhum outro modelo apresentou um resultado melhor.

Tabela 5.14. Medidas para o modelo com trocas, com 8 filas de caixa normal sendo $\lambda_E = 1,6058$, $\mu = 0,2336$, com tempo observado de espera na fila na amostra, de 1,85 minutos e $\lambda_E / m\mu = 0,859$, para o dia 5 de dezembro.

λ	k	N	W_q	Desvio	Perda	Vazio	Ocioso
1,6095	1	4	2,0549	11,1%	0,0023	0,0007	0,4057
1,6069	1	5	2,1855	18,1%	0,0007	0,0007	0,4048
1,6061	1	6	2,2348	20,8%	0,0002	0,0007	0,4047
1,6268	2	3	3,3720	82,3%	0,0155	0,0001	0,4881

Para os caixas rápidos (tabela 5.15), duas situações poderiam concorrer como satisfatórias: uma com $k=1$ e $N=7$ e a outra com $k=2$ e $N=7$. Isto poderia ser pensado, pois com $k=1$ há subestimação do tempo de espera e observa-se uma tendência em crescer o valor estimado pelo modelo. Contudo, como se pôde notar, para $k=2$ e $N=7$ o desvio aumentou bastante, descartando-se esta possibilidade. Desta forma, pode-se, então, para os

caixas rápidos do dia 5 de dezembro, assumir que o cliente tem pouca tolerância à espera, como ocorre para os caixas normais. Também, como ocorre para os caixas normais, existe alguma sensibilidade quando $N=7$ é alterado para 6 ou 8, provavelmente devido ao valor de $\lambda_E / m\mu$. Ainda deve-se ressaltar que o desvio obtido aqui, apesar de satisfatório, é maior que para a fila $M/M/m$ (0,1%), contudo menor que para todos os outros modelos.

Tabela 5.15. Medidas para o modelo com trocas, com 4 filas de caixa rápido sendo $\lambda_E = 2,0803$, $\mu = 0,6228$, com tempo observado de espera na fila na amostra, de 1,61 minutos e $\lambda_E / m\mu = 0,835$, para o dia 5 de dezembro.

λ	k	N	W_q	Desvio	Perda	Vazio	Ocioso
2,0870	1	6	1,4636	-9,1%	0,0032	0,0209	0,3395
2,0834	1	7	1,5259	-5,2%	0,0015	0,0210	0,3393
2,0818	1	8	1,5623	-3,0%	0,0007	0,0210	0,3391
2,0840	2	7	2,0808	29,2%	0,0019	0,0096	0,3921

São apresentados agora sucintamente os resultados obtidos na coleta da segunda semana, referente aos dias 10, 11 e 12 de dezembro. No dia 10/12 (uma sexta-feira) ocorreu, no caixa rápido, a menor proporção de servidores ocupados em toda a coleta realizada. O comportamento para este dia ficou muito parecido com a sexta-feira anterior (03/12) levando a pensar que, talvez, haja menor movimento neste dia da semana para aquele horário. Também, o comportamento para o valor de k foi idêntico ao da semana anterior, contudo, as estimativas apresentaram melhores resultados (vide tabelas 5.16 e 5.17 a seguir).

Tabela 5.16. Medidas para o modelo com trocas, com 8 filas de caixa normal sendo $\lambda_E = 1,4172$, $\mu = 0,2182$, com tempo observado de espera na fila na amostra, de 1,39 minutos e $\lambda_E / m\mu = 0,812$, para o dia 10 de dezembro.

λ	k	N	W_q	Desvio	Perda	Vazio	Ocioso
1,4183	1	4	1,4271	2,7%	0,0006	0,0012	0,5163
1,4175	1	5	1,4622	5,2%	0,0001	0,0012	0,5162
1,4191	1	6	1,4714	5,9%	0,0000	0,0012	0,5162
1,4172	2	3	3,1649	127,7%	0,0155	0,0001	0,4881

Tabela 5.17. Medidas para o modelo com trocas, com 4 filas de caixa rápido sendo $\lambda_E = 1,7311$, $\mu = 0,6130$, com tempo observado de espera na fila na amostra, de 1,24 minutos e $\lambda_E / m\mu = 0,706$, para o dia 10 de dezembro.

λ	k	N	W_q	Desvio	Perda	Vazio	Ocioso
1,7313	1	6	0,6039	-51,3%	0,0001	0,0486	0,5620
1,7313	1	7	0,6068	-51,1%	0,0000	0,0486	0,5619
1,7311	1	8	0,6072	-51,0%	0,0000	0,0486	0,5620
1,7314	2	6	1,1794	-4,9%	0,0002	0,0262	0,6349
1,7311	2	7	1,1830	-4,6%	0,0000	0,0262	0,6349
1,7311	2	8	1,1843	-4,5%	0,0000	0,0262	0,6348
1,7311	3	7	1,6885	36,2%	0,0001	0,0175	0,6772

Para o dia 11 de dezembro foram obtidas conclusões semelhantes com as do dia 04 de dezembro (ambos os dias um sábado) no que se refere tanto a estimativas quanto ao comportamento dos clientes para troca de filas. Nos caixas normais foi obtido $k=1$ enquanto, para os caixas rápidos, $k=2$, revelando assim mais tolerância dos clientes que utilizaram este serviço (vide tabelas 5.18 e 5.19, a seguir).

Tabela 5.18. Medidas para o modelo com trocas, com 8 filas de caixa normal sendo $\lambda_E = 1,3429$, $\mu = 0,1901$, com tempo observado de espera na fila na amostra, de 3,63 minutos e $\lambda_E / m\mu = 0,883$, para o dia 11 de dezembro.

λ	k	N	W_q	Desvio	Perda	Vazio	Ocioso
1,3487	1	4	3,1475	-13,3%	0,0044	0,0005	0,3452
1,3449	1	5	3,4363	-5,3%	0,0015	0,0005	0,3443
1,3436	1	6	3,5737	-1,6%	0,0005	0,0005	0,3440
1,3429	2	3	4,2576	17,3%	0,0181	0,0001	0,4606

Finalmente, para o dia 12 de dezembro, obteve-se um cenário também muito parecido ao domingo anterior (dia 05/12). Novamente, houve mais tolerância nos caixas rápidos, com $k=2$ (vide tabelas 5.20 e 5.21 abaixo).

Tabela 5.19. Medidas para o modelo com trocas, com 4 filas de caixa rápido sendo $\lambda_E = 2,4239$, $\mu = 0,6718$, com tempo observado de espera na fila na amostra, de 2,19 minutos e $\lambda_E / m\mu = 0,902$, para o dia 11 de dezembro.

λ	k	N	W_q	Desvio	Perda	Vazio	Ocioso
2,4553	1	6	2,1917	0,1%	0,0132	0,0108	0,2101
2,4432	1	7	2,4036	9,8%	0,0081	0,0109	0,2092
2,4348	1	8	2,5527	16,6%	0,0050	0,0110	0,2095
2,4435	2	7	2,8474	30,0%	0,0093	0,0046	0,2463

Tabela 5.20. Medidas para o modelo com trocas, com 8 filas de caixa normal sendo $\lambda_E = 1,6099$, $\mu = 0,2254$, com tempo observado de espera na fila na amostra, de 2,92 minutos e $\lambda_E / m\mu = 0,893$, para o dia 12 de dezembro.

λ	k	N	W_q	Desvio	Perda	Vazio	Ocioso
1,6179	1	4	2,8936	-0,9%	0,0055	0,0004	0,3208
1,6131	1	5	3,2128	10,0%	0,0021	0,0004	0,3187
1,6112	1	6	3,3745	15,6%	0,0008	0,0004	0,3181
1,6232	2	3	3,6661	25,6%	0,0205	0,0001	0,4392

Tabela 5.21. Medidas para o modelo com trocas, com 4 filas de caixa rápido sendo $\lambda_E = 2,2399$, $\mu = 0,6272$, com tempo observado de espera na fila na amostra, de 2,64 minutos e $\lambda_E / m\mu = 0,893$, para o dia 12 de dezembro.

λ	k	N	W_q	Desvio	Perda	Vazio	Ocioso
2,2605	1	6	2,1795	-17,4%	0,0107	0,0122	0,2306
2,2527	1	7	2,3798	-9,9%	0,0064	0,0122	0,2286
2,2480	1	8	2,5246	-4,8%	0,0039	0,0122	0,2276
2,2643	2	6	2,6595	0,7%	0,0128	0,0051	0,2697
2,2549	2	7	2,8842	9,3%	0,0076	0,0051	0,2667
2,2493	2	8	3,0466	15,4%	0,0046	0,0051	0,2652
2,2576	3	7	3,3750	27,8%	0,0090	0,0027	0,2921

Nota-se, para a segunda semana de coleta, muita semelhança com os resultados obtidos na primeira, não só em relação aos tempos esperados na fila, mas também em relação aos valores de k , aos desvios, às probabilidades de perda, do sistema estar completamente vazio e de haver ociosidade (de haver pelo menos um caixa desocupado). Ainda, as flutuações nas estimativas de W_q causadas pelos diferentes valores de N têm comportamentos muito coesos para as duas semanas.

Nas tabelas 5.22 e 5.23 foi realizado um resumo das comparações entre os modelos apresentados e aplicados. Foram destacados, em negrito, em todas as situações e para todos os modelos comparados, os menores desvios obtidos, como também os desvios aceitáveis, ou seja, que não são os menores, mas que são considerados toleráveis. Nota-se que o modelo com trocas tem, de modo geral, melhores estimativas que os demais modelos.

Tabela 5.22. Comparação entre os modelos $M/M/m$, $[M/M/1]^m$ e o resultado obtido pelo modelo com trocas, para caixas normais e rápidos, na primeira semana de coleta.

Dia	Tipo de caixa	$M/M/m$ W_q (desvio)	$[M/M/1]^m$ W_q (desvio)	$M/M/m/mN$ W_q (desvio)	$[M/M/1/N]^m$ W_q (desvio)	Trocas W_q (desvio)	Valor observado (min.)
03	Normal $m=8$	1,8159 (0,9%)	22,18 (1232,3%)	1,57 (-12,8%)	6,95 (286%)	1,7790 (-1,2%)	1,80
	Rápido $m=4$	0,6957 (48,7%)	4,2824 (299,5%)	0,69 (-51,7%)	2,92 (104%)	1,2600 (-11,9%)	1,43
04	Normal $m=8$	3,4397 (2,4%)	37,7379 (1123,2%)	3,27 (-2,7%)	9,24 (175%)	3,2440 (-3,5%)	3,36
	Rápido $m=4$	2,5594 (40,6%)	11,8973 (653,7%)	2,32 (27,5%)	4,44 (144%)	2,1615 (18,8%)	1,82
05	Normal $m=8$	2,2640 (22,4%)	26,1374 (1412,8%)	2,17 (17,3%)	7,26 (292%)	2,1855 (18,1%)	1,85
	Rápido $m=4$	1,6086 (0,1%)	8,1291 (504,9%)	1,51 (-6,2%)	3,69 (129%)	1,5259 (-5,2%)	1,61

Ainda que em seis (das doze) situações, seu comportamento não tenha sido melhor, em duas delas os desvios foram de apenas -1,2% e -3,5%, podendo ser considerados quase que equivalente aos desvios da $M/M/m$, respectivamente de 0,9% e de 2,4%, relativos aos caixas normais dos dias 3 e 4 de dezembro.

É notório o impacto que as trocas de filas surtem nas estimativas quando se consideram as filas paralelas independentes, como no caso do modelo de fila $[M/M/1]^m$. Os desvios obtidos denotam o quanto pode ser prejudicial às análises dos sistemas desconsiderar a dependência, quando esta existe. Também, pode parecer, em análise inicial, que os modelos $M/M/m$ e $M/M/m/mN$ são concorrentes ao modelo com trocas. Deve-se lembrar que, de fato, o modelo $M/M/m$ pode ser visto como o caso particular do modelo com trocas com $k=1$ e $N \rightarrow \infty$. Como as probabilidades de perda do modelo com trocas com $k=1$ e $N=5$ (para caixas normais) e com $k=1$ e $N=7$ (para caixas rápidos) são pequenas, esperam-se resultados similares aos obtidos pelo modelo com trocas.

Tabela 5.23. Comparação entre os modelos $M/M/m$, $[M/M/1]^m$ e o resultado obtido pelo modelo com trocas, para caixas normais e rápidos, na segunda semana de coleta.

Dia	Tipo de caixa	$M/M/m$ W_q (desvio)	$[M/M/1]^m$ W_q (desvio)	$M/M/m/mN$ W_q (desvio)	$[M/M/1/N]^m$ W_q (desvio)	Trocas W_q (desvio)	Valor observado (min.)
10	Normal $m=8$	1,4744 (6,1%)	19,7924 (1423,9%)	1,46 (5,0%)	7,29 (524,5%)	1,4622 (5,2%)	1,39
	Rápido $m=4$	0,6076 (-51,0%)	3,9173 (315,9%)	0,61 (-50,8%)	2,82 (127,4%)	1,1830 (-4,6%)	1,24
11	Normal $m=8$	3,6885 (1,6%)	39,7087 (1093,9%)	3,40 (-6,3%)	9,22 (154,0%)	3,4363 (-5,3%)	3,63
	Rápido $m=4$	3,0074 (37,3%)	13,7033 (625,7%)	2,62 (19,6%)	3,86 (76,2%)	2,4036 (9,8%)	2,19
12	Normal $m=8$	3,5288 (20,8%)	36,9499 (1265,4%)	3,17 (8,6%)	7,87 (169,5%)	3,2128 (10,0%)	2,92
	Rápido $m=4$	2,8757 (8,9%)	13,2810 (503,1%)	2,58 (-2,3%)	4,07 (54,2%)	2,8842 (9,3%)	2,64

Para melhor comparação entre os modelos foi realizado o cálculo das médias dos valores absolutos dos seus respectivos desvios. Assim, ao realizar-se um cálculo simples, observa-se que esta média para o modelo $M/M/m$ é 18,2% para a primeira semana e 21% para a segunda semana (valores próximos aos obtidos para o modelo $M/M/m/mN$, de 19,7% e 15,4%, respectivamente), sendo que para o modelo com trocas foram 9,8% e 7,4%, respectivamente.

O comportamento de k parece estar ligado ao tipo de caixa e ao valor de $\lambda_E / m\mu$. Como exemplo, nos caixas rápidos foram obtidos resultados com $k=2$, ou seja, os clientes trocam de fila para ganhar duas posições e não uma (com $k=1$). Como ali, o tempo gasto na fila deve ser pequeno, a compensação da troca deve ser maior. Por outro lado, mesmo em caixas rápidos, são obtidos resultados com $k=1$ mas com valor de $\lambda_E / m\mu$ mais alto (deve-se lembrar que, para os caixas rápidos do dia 12, os desvios para $k=1$ e $k=2$ ficaram praticamente iguais).

Concluindo, em geral, o modelo de filas com trocas mostra-se mais eficiente (com desvios menores) e mais consistente em suas medidas, ou seja, seu comportamento é mais harmônico que os outros modelos considerados. Neste modelo, devido a se conhecer a distribuição de equilíbrio dos estados do sistema, várias medidas de desempenho podem ser calculadas, contudo, o tempo de espera na fila é a mais importante para este estudo.

De forma geral, para o supermercado analisado e nos dias de coleta realizados, pode-se dizer que as trocas de filas, quando englobadas ao modelo, tendem a melhorar seu desempenho em estimar o tempo de espera observado na fila. Também, deixam as estimativas mais consistentes, ou seja, com menor variabilidade que os outros modelos. Este fato é importante para avaliar e propor melhorias em sistemas de filas comumente encontradas em supermercados.

5.8 Comportamento do modelo com trocas em função do número de filas para a primeira semana de coleta

Um experimento foi realizado para observar o comportamento das estimativas de tempo de espera na fila em função do número de filas (mantendo-se $\lambda_E / m\mu$ fixo) para algumas possibilidades dos caixas normais.

Este estudo foi realizado exatamente pelo aumento exponencial do número de equações em função de N . O objetivo era verificar se seria razoável simplesmente decompor um sistema de filas em subsistemas menores e independentes, que poderiam ser analisados isoladamente. Então, para algumas situações, o número de filas (m) foi sendo elevado, mantendo-se a carga de trabalho ($\lambda_E / m\mu$) constante.

Pelas tabelas 5.24 e 5.25 e figuras 5.15 e 5.16 (a seguir), nota-se que, para um mesmo N e para $k=1$, o decaimento dos valores parece ser relativamente acentuado e harmônico, conforme o número de filas (m) vai aumentando.

Observa-se que para $N=4$ e $k=1$, quando $m=4$ é obtido que $W_q=3,2664$ minutos e, quando $m=8$, é obtido que $W_q=1,9958$, valores muito diferentes. Analisando-se este fato pode-se dizer que não é razoável simplesmente decompor um sistema com $m=8$ em dois subsistemas com $m=4$ (supostamente independentes). Assim, com base nos resultados obtidos, esta idéia pode ser desconsiderada para este modelo com trocas.

Na tabela 5.25, apresentada a seguir, não foi possível obter os cálculos relativos aos modelos quando um grande número de filas foi considerado (denotados com o símbolo *). Como exemplo, quando se utilizam $m=8$, $N=3$ e $k=2$ é necessário resolver um sistema com 12866 equações. O máximo que se consegue resolver, pelo menos com essa implementação computacional, é $2^{14} - 1$ equações, ou seja, 16.383 equações.

Tabela 5.24. Tempo estimado (W_q) para o modelo com trocas, com m filas de caixa normal, sendo λ_E variável, $\mu = 0,2336$, com tempo observado na amostra de 1,85 minutos e fixando-se $\lambda_E / m\mu = 0,859$, para o dia 5 de dezembro.

m (n. de filas); λ_E	$N = 3, k = 1$	$N = 4, k = 1$
2; 0,401544	2,4241	4,0095
3; 0,602316	2,3734	3,6532
4; 0,803088	2,2498	3,2664
5; 1,00386	2,0997	2,8956
6; 1,204632	1,9425	2,5577
7; 1,405404	1,7877	2,2578
8; 1,606176	1,6401	1,9958
9; 1,806948	1,5022	1,7691

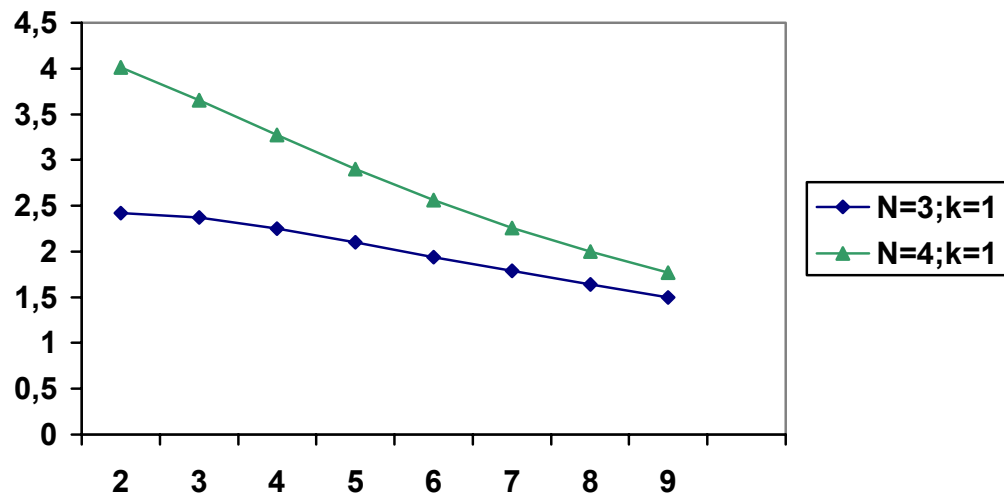


Figura 5.15. Gráfico mostrando a estabilização do tempo estimado de espera conforme aumento do número de filas mantendo $\lambda_E/m\mu$ fixo (amostra do dia 5/12/04).

Tabela 5.25. Tempo estimado (W_q) para o modelo com trocas, com m filas de caixa normal, sendo λ_E variável, $\mu = 0,1880$, com tempo observado na amostra de 3,36 minutos e fixando-se $\lambda_E/m\mu = 0,878$, para o dia 4 de dezembro.

M (n. de filas); λ_E	$N = 5, k = 1$	$N = 6, k = 1$
2; 0,3295	6,9768	8,5885
3; 0,4943	6,2179	7,3601
4; 0,6591	5,4668	6,2487
5; 0,8239	4,7807	5,3001
6; 0,9886	4,1737	4,5095
7; 1,1534	3,6532	3,8663
8; 1,3182	3,2094	3,3425
9; 1,4830	*	*

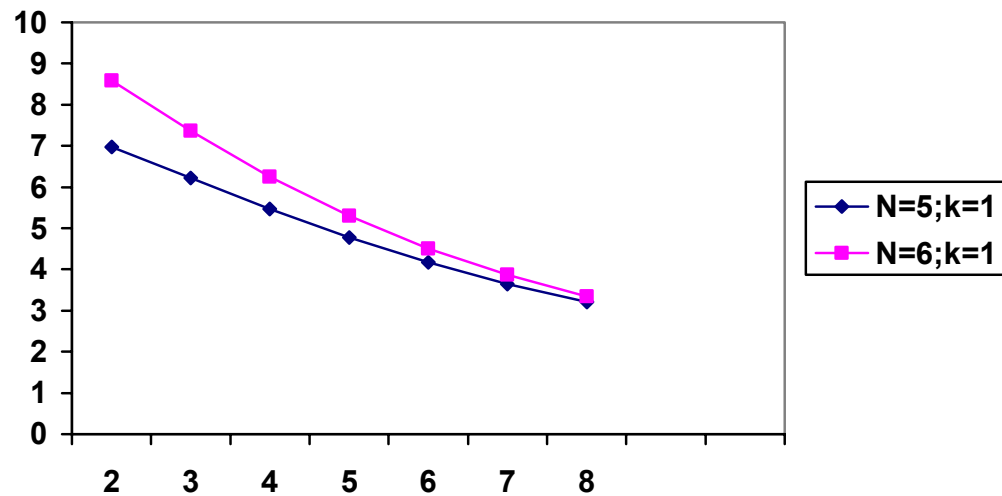


Figura 5.16. Gráfico mostrando a estabilização do tempo estimado de espera conforme aumento do número de filas mantendo $\lambda_E/m\mu$ fixo (amostra do dia 4/12/04).

Observando-se a tabela 5.25, pode-se obter a mesma conclusão que para a tabela 5.24, onde a divisão de um sistema com $m=8$ em outro com $m=4$ não deve ser considerada razoável. Analisando-se os resultados, nota-se que os valores estimados para $m=4$ estão acima dos valores estimados quando $m=8$ pelo menos em 38%, portanto uma diferença muito grande.

Como pode ser observado, o decaimento das curvas ocorrem de forma acentuada, desencorajando a estratégia proposta.

5.9 A simulação do modelo com trocas

A simulação é uma forma de representar algum sistema real de interesse através de um modelo matemático implementado em computador. Para que esta representação seja válida é necessário que todas as características do sistema a ser simulado sejam amplamente conhecidas e muito bem avaliadas. É freqüentemente utilizada para avaliar um novo sistema proposto, modificações que se pretenda introduzir em algum já existente, ou ainda para se obter informações sobre algum sistema cujas medidas teóricas não puderam ser obtidas analiticamente. Pode-se ainda utilizá-la com o objetivo de validar algum sistema

estudado ou comparar dois ou mais sistemas de forma a inferir qual deles terá melhor desempenho. Atualmente observa-se um crescente uso da simulação em várias áreas das engenharias, entre outras. Isto ocorre porque sua utilização é simples e pode fornecer resultados quase imediatos.

Para que se tenham resultados representativos é necessário que se realizem muitas rodadas de simulação, ou seja, repeti-la várias vezes, pois para cada rodada é obtido apenas um valor da estatística. A diferença de resultados entre as rodadas é devido ao fato de que em cada uma delas é utilizada uma semente aleatória distinta. Com essas repetições, obtêm-se resultados consistentes para que se possa inferir sobre o sistema de interesse. Também, se forem conhecidas algumas medidas de desempenho deste, pode-se compará-las com os resultados obtidos pelo simulador, verificando se o sistema está representado satisfatoriamente. Por outro lado, a quantidade de observações em cada rodada deve ser suficiente para que o valor da estatística de interesse seja representativo.

Uma vez que a simulação represente bem o sistema de interesse, geralmente o resultado que se busca é estimar adequadamente a média (e seu intervalo de confiança) ou então alguma medida teórica γ desconhecida. Como se está analisando sistemas estocásticos estacionários, onde os dados não são independentes, geralmente enfrenta-se, entre outros, dois problemas que têm reflexos diretos na qualidade dos resultados obtidos para posterior análise estatística.

O primeiro problema trata da redução da influência (ou vício) das condições iniciais da simulação nas observações coletadas devido à presença de correlação entre as observações de um processo estocástico. Se a simulação é iniciada com o sistema vazio e ocioso, até que este adquira equilíbrio (estacionariedade) é decorrido um período, chamado período transitório inicial. Por exemplo, em um experimento descrito (LAW, 1977), depara-se com este problema quando se calcula a média dos tempos de espera em uma fila $M/M/1$ com taxa de ocupação 0,9 e iniciada vazia. Após coletar 320 tempos de espera, concluiu-se que a média com valor 6,01 ainda não estava suficientemente perto do seu valor teórico de 8,1, de modo a considerar-se que esta fila pudesse ter entrado em equilíbrio, se esta média pudesse ser tomada como um critério de decisão. Se quer-se obter informações sobre o sistema estacionário, é necessário saber aproximadamente o instante a

partir do qual esta condição está satisfeita para, a partir daí, coletar as observações que efetivamente serão analisadas estatisticamente.

Com o objetivo de localizar o instante a partir do qual o sistema estará aproximadamente estacionário e, portanto, conseguir reduzir o vício devido as condições iniciais à um nível desprezível, muitas regras de inicialização foram propostas (PAWLIKOWSKI, 1990). Contudo, não se conseguiu uma regra única que detectasse adequadamente este período para uma grande quantidade e variedade de sistemas estocásticos, pois as regras têm diferentes critérios para tratar deste problema, mesmo que se restrinjam a verificar estacionariedade de valores médios. Desta forma pode haver necessidade de se pesquisar qual regra trará melhores resultados do sistema estocástico que está sendo simulado.

O segundo problema diz respeito à obtenção de estatísticas apropriadas devido à correlação existente entre as observações de um processo estocástico. A presença desta correlação dificulta bastante que se obtenha um estimador da variância da média amostral que faça uma boa aproximação desta e, conseqüentemente, que se calcule um intervalo de confiança que cubra o parâmetro verdadeiro com o nível de confiança e precisão desejados.

Por outro lado, existem vários métodos de coleta de dados de simulação. Cada método apresenta diferentes condições de aplicação, portanto, fornecendo diferentes estatísticas para análise dos dados coletados. Também, por possuírem estruturas distintas, apresentam problemas que se diferem (LAW e KELTON, 2000; PAWLIKOWSKI, 1990; FREITAS, 2001). Contudo, a correlação entre as observações de um processo estocástico se apresenta como um grande problema a ser considerado nos métodos que utilizam uma única grande rodada de simulação. Há métodos que escapam deste problema, mas enfrentam outros, como o método de replicações, que é visto mais adiante.

Descreve-se agora o método de coleta de dados *batch means*. Suponha uma longa rodada de simulação de tamanho n de um processo estocástico estacionário, dividida em k *batches* (blocos) consecutivos e não sobrepostos de tamanho m , tal que $m = n/k$, com m inteiro. Neste método, é calculada a média de cada um dos *batches*, totalizando k médias, e usa-se como estimador para uma medida de interesse γ do processo estacionário, a média geral dessas médias. Neste caso a definição do tamanho dos

batches, m' , é importante, pois se este for suficientemente grande, tem-se que as médias dos *batches* são assintoticamente não correlacionadas entre si. Além disso, se o tamanho de m' for suficiente, essas médias são assintoticamente normalmente distribuídas e assintoticamente independentes. Sendo assim, pode-se calcular um intervalo de confiança para γ usando-se como seu estimador a média geral (LAW e KELTON, 2000). Contudo, tomar um valor para m' maior que o necessário (para algum n' fixado), para satisfazer as suposições de normalidade e independência, implica, por exemplo, em aumento no meio comprimento esperado do intervalo de confiança para γ , havendo um acréscimo na probabilidade de se cobrir pontos que são estatisticamente diferentes de γ , ou seja, um teste estatístico rejeitaria a hipótese de igualdade. Se, porém, m' for menor que o necessário, as médias dos *batches* poderiam estar altamente correlacionadas, tornando-se a maior fonte de erro deste método, devido ao vício induzido no estimador da variância da média geral.

No método de *batch means*, para que haja representatividade, devem ser realizadas várias rodadas de simulação com sementes distintas onde, para todas as rodadas, um mesmo tamanho de *batch* é previamente determinado. Desta forma há perda de precisão nos resultados, pois cada rodada têm resultados diferentes, obtendo, portanto, resultados de correlações diferentes. Assim, haveria necessidade de tamanhos de *batches* distintos e adequados para dissipá-las.

Uma variante do método *batch means*, chamado *batch means* espaçado, procura diminuir o efeito da correlação entre as médias dos *batches* inserindo espaços (observações não utilizadas nas estatísticas) entre *batches* consecutivos, obtendo assim estimadores menos viciados que o método anterior e melhorando a cobertura do intervalo de confiança para a média do processo analisado (FOX et al., 1991). Além disso, para este método é possível determinar um estimador fortemente consistente para a variância do processo. Contudo, existe a perda das observações referentes aos espaços inseridos entre os *batches*.

Existem ainda outros métodos, como o *batch means* sobreposto que, a cada observação, da primeira até a $(n' - m' + 1)$ -ésima, um *batch* de tamanho m' é coletado. A idéia é oposta a do *batch means* espaçado e as médias dos *batches* são extremamente correlacionadas. Existem ainda outros métodos (PAWLIKOWSKI, 1990; LAW e

KELTON, 2000) como o método regenerativo, cujos *batches* têm ciclos (comprimentos) aleatórios determinados pelos instantes sucessivos de tempo em que o processo se regenera (inicia novamente).

5.9.1 O método de replicações

No método de replicações (LAW e KELTON, 2000; FREITAS, 2001), a simulação é repetida um número pré-fixado de vezes onde, para cada repetição, é usada uma seqüência de números pseudo-aleatórios diferente. Conseqüentemente, essas repetições, assim tomadas, são estatisticamente independentes. Para cada repetição é calculada a média das observações que foram coletadas, fornecendo uma nova seqüência de médias, ainda independente umas das outras. O inconveniente deste método é que a cada rodada deve ser determinado o período transitório inicial, cujo vício é sua maior fonte de erro, podendo também haver, dependendo do sistema, considerável perda de informações. Nos métodos que utilizam uma grande rodada de simulação, como no método *batch means*, o problema de vício de inicialização é considerado uma única vez a cada rodada, porém não se pode utilizar da independência entre os *batches*, como a obtida pelo método de replicações.

Fazendo-se agora k replicações com m observações cada uma, haverá um total de $n = mk$ observações e uma seqüência de k variáveis aleatórias $\overline{X}_i(m)$, $i=1,2,\dots, k$ (média da replicação i), que são as médias das k replicações, calculadas por

$$\overline{X}_i(m) = \frac{1}{m} \sum_{j=1}^m X_{ij}, \text{ para } i = 1, 2, \dots, k$$

onde X_{ij} é a j -ésima variável aleatória da seqüência de variáveis aleatórias na replicação i . Pode-se visualizar a estrutura usual de uma particular replicação do seguinte modo:

Replicação	Descarte	Observações Utilizadas	Média
i	d	$x_{i1}, x_{i2}, \dots, x_{im}$	$\bar{x}_i(m)$

onde observa-se a i -ésima replicação com d observações descartadas e sua média, respectivamente. Usa-se, como estimador pontual para a média γ do processo estacionário, a seguinte estatística (média geral)

$$\bar{\bar{X}}(n') = \frac{1}{k} \sum_{i=1}^k \bar{X}_i(m).$$

Sabendo-se que este estimador é não viciado para γ e que a seqüência de médias são independentemente e identicamente distribuídas, pode-se obter um estimador não viciado para a variância de $\bar{\bar{X}}(n')$, dado por

$$\hat{\sigma}^2(\bar{\bar{X}}(n')) = \frac{1}{(k'-1)} \sum_{j=1}^{k'} (\bar{X}_j(m) - \bar{\bar{X}}(n'))^2$$

Sabendo-se que $\bar{\bar{X}}(n')$ é uma soma de variáveis aleatórias independente e identicamente distribuídas, pode-se considerar a estatística

$$T_{k'-1} = (\bar{\bar{X}}(n') - \gamma) / \hat{\sigma}(\bar{\bar{X}}(n'))$$

que tem uma distribuição t-Student padrão com $k'-1$ graus de liberdade. Usando o resultado obtido pode-se facilmente calcular um intervalo de confiança para γ . Supondo o nível de significância α conhecido, este intervalo é dado por

$$\bar{\bar{X}}(n') - t_{k'-1, 1-\alpha/2} \hat{\sigma}(\bar{\bar{X}}(n')) \leq \gamma \leq \bar{\bar{X}}(n') + t_{k'-1, 1-\alpha/2} \hat{\sigma}(\bar{\bar{X}}(n')) \quad (5.13)$$

na qual $t_{k'-1, 1-\alpha/2}$ é o valor da distribuição t-Student com $k'-1$ graus de liberdade. Quando houver mais de 30 replicações pode-se substituir a t-Student pela Normal padrão obtendo uma boa aproximação para o intervalo considerado. Fazendo-se uma análise detalhada do período transitório inicial a ser considerado e não havendo a possibilidade de uma grande perda de observações pode-se considerar o método de replicações uma boa alternativa ao método de *batch means* devido aos problemas mencionados anteriormente.

5.9.2 Sistemas terminais e não-terminais

Os sistemas a serem simulados dividem-se em terminais e não-terminais. Os sistemas terminais apresentam condições iniciais fixas e alguma ocorrência que determina o fim da simulação. Alguns exemplos desses sistemas podem ser verificados em bancos, restaurantes, escolas, etc. Geralmente estes sistemas se apresentam vazios como condição inicial e retornam a este estado quando no fim de seu funcionamento. Nos sistemas não-terminais, em várias ocasiões, não se conhece as condições iniciais e nem há uma ocorrência que determine o fim da simulação (ou o final de seu funcionamento). Alguns exemplos são sistemas de segurança, comunicações, hospitais, entre outros. Pode-se perceber que nem todos os sistemas são facilmente classificáveis, apresentando características próprias pertencentes a sistemas terminais e não-terminais.

Nos sistemas terminais, uma prática comum é determinar o tamanho da amostra (ou número de replicações) para um nível de confiança (ou de precisão) anteriormente definido para a média de uma medida de desempenho avaliada. Desta forma, geralmente é simulado todo o período de funcionamento do sistema terminal a ser estudado.

Nos sistemas não-terminais, as condições iniciais não são determinadas e, geralmente, se deseja estudá-los em equilíbrio. Para tanto é necessário retirar da análise as observações do período transitório inicial. Também, como não há uma ocorrência para o final da simulação, deve-se determinar um período de forma a não alongá-lo desnecessariamente. Pode-se observar então que para este tipo de sistema estes dois problemas devem ser enfrentados.

5.9.3 Simulando o modelo com trocas

O modelo com trocas utilizado nas seções anteriores tem algumas características bem específicas (no processo de chegadas e condições de troca) cujo modelo de simulação deve representar com bastante fidelidade. Com esta finalidade, os resultados obtidos por simulação são comparados aos obtidos pela aplicação do modelo analítico. Deve-se lembrar que o próprio modelo com trocas faz algumas aproximações em relação ao que na realidade ocorre no supermercado e, assim sendo, os resultados da simulação devem se aproximar menos aos obtidos na amostra do que aos obtidos pelo modelo.

Uma situação que deve ser considerada com cuidado é referente às condições de local e período de onde foram retirados os dados em relação a simulação a ser realizada. Pela tabela 3.1 podem ser observados os horários de coleta; em alguns dias das 10h30 min. às 12h30min. e em outros das 18h as 20h. Pode-se observar que se trata de uma parte de um período de tempo inserido em um sistema terminal. Contudo, observa-se que não se conhece as condições iniciais do sistema, o que é uma característica de sistemas não-terminais, inclusive, não retornando ao estado inicial quando do final da coleta.

Law e Kelton (2000, p.518) sugerem duas aproximações heurísticas para este caso: em uma delas deve-se coletar dados sobre o número de usuários presentes no local, no mesmo período considerado, para vários dias e iniciar a simulação aleatoriamente, de acordo com a proporção de clientes encontrados. Assim, para várias rodadas de simulação existe relativa independência, pois diferentes condições de inicialização seriam obtidas (de acordo com a proporção observada). A outra regra sugerida é que seja simulado o período todo, desde a abertura do local até o fim do período de coleta, retirando-se depois o período inicial, relativo ao início de todo o processo até o começo da coleta.

Para a segunda aproximação, onde ocorre a retirada das observações da simulação no período citado, percebe-se que ocorre uma representação fiel do cenário que seria encontrado no supermercado para os clientes que dele se utilizam no horário da coleta de dados. Também, é mais complexa a aplicação da primeira aproximação apresentada, pois haveriam vários dias de observação das condições iniciais, o que para este estudo não foi realizado.

Uma próxima etapa a ser enfrentada diz respeito ao método de coleta de dados a ser utilizado. Poderia ser pensado no método *batch means*, contudo, duas horas pode ser um período pequeno, e assim, poderia haver correlação entre os *batches*, maior fonte de vício deste método de coleta de dados. O método de replicações pode ser indicado desde que a avaliação do período transitório inicial não aponte que este seja muito longo. Law e Kelton (2000, p.213) afirmam que “um método estatístico para construir intervalos de confiança para a média geral deve ser fácil e fornecer bons resultados estatísticos. Em consideração a isto, o método de replicação é definitivamente superior”. Resultados de Law e Kelton (1984) mostram que é melhor realizar replicações mais longas do que mais replicações, pois assim obtém-se maior cobertura de γ nos intervalos de confiança gerados na simulação. Contudo, no presente caso, o período de cada replicação é fixo pois deve corresponder à situação real da coleta realizada no supermercado e, além disso, considerar um número de replicações pequeno poderia aumentar a cobertura de γ nos intervalos de confiança, mas com aumento na amplitude dos mesmos. Fishman (1996) sugere que, para cada replicação, sejam utilizadas no mínimo cem observações, sem contar as descartadas, referentes ao período transitório inicial. Pode-se considerar ainda que, com $k' = 30$, os intervalos de confiança já podem utilizar a distribuição normal, obtendo boa aproximação. Considerando todos estes aspectos, pode-se realizar a simulação do modelo com trocas de forma conveniente e obter resultados confiáveis.

A simulação foi realizada utilizando-se o *software* ARENA, com ênfase em buscar informações sobre o tempo médio gasto na fila com trocas, enfoque principal deste trabalho. Deve-se observar que o simulador foi construído considerando-se que todas as distribuições de tempo de serviço eram exponenciais, como o modelo com trocas considera, contudo, como foi descrito anteriormente, em alguns caixas, para alguns dias da coleta, isto não ocorreu. Desta forma, o simulador representa o modelo com trocas, mas com aproximações em relação aos dados obtidos na coleta. Inicialmente deve-se verificar se não existe um vício considerável referente ao período transitório inicial. Sabe-se que este vício ocorre quando ainda há autocorrelação significativamente diferente de zero, influenciando as observações coletadas após considerar o referido período como terminado. Desta forma, analisar o decaimento das autocorrelações aumentando-se os *lags*, para este período, é uma maneira de verificação da condição necessária.

Para determinar aproximadamente quando termina o período, Freitas (2001, p.235) sugere que se verifique em que *lag* a autocorrelação se torna próxima de zero. Desta maneira, a partir de então, pode-se pressupor que a influência será praticamente nula, ou seja, supondo-se que a autocorrelação está próxima de zero no *lag* j , a observação x_i não mais influencia significativamente a observação x_{i+j} . Ainda, nos procedimentos práticos, costuma-se multiplicar o número de observações referentes a este período por dez como forma de, praticamente, garantir o término do período transitório inicial. Para este trabalho as situações menos favoráveis ocorrem onde este período é de 2,5 horas, como nos dias 11 e 12/12. Nos outros dias restantes, utilizando-se deste critério, seus respectivos períodos são todos menores que dos dias 11 e 12/12, de 2,5 horas.

Decidiu-se então realizar trinta replicações, para cada dia e tipo de caixa, coletando inclusive seus respectivos períodos transitórios iniciais. Para exemplificar são apresentados os gráficos (vide figuras 5.17 e 5.18) referentes as autocorrelações para os caixas normais do dia 12/12 e para os caixas rápidos do dia 11/12, sendo que, nestes dias, ocorreram as maiores proporções de servidores ocupados. As duas figuras revelam o comportamento dos dois cenários descritos acima. Para o caixa normal, obteve-se que $\rho_{20}^* \approx 0$ e então deveria-se descartar 200 observações; para o caixa rápido, obteve-se $\rho_{32}^* = 0,09$ e então deveria-se descartar 320 observações, o que poderia parecer muito, contudo, deve-se ressaltar que os respectivos períodos transitórios iniciais contém mais observações que o solicitados por esta regra. Desta forma, considera-se o vício desprezível e o método de replicações pode ser assumido para a simulação. Deve-se ressaltar que, para todas as replicações realizadas e em todos os cenários considerados, encontraram-se resultados semelhantes aos obtidos aqui.

Autocorrelação típica para caixa normal

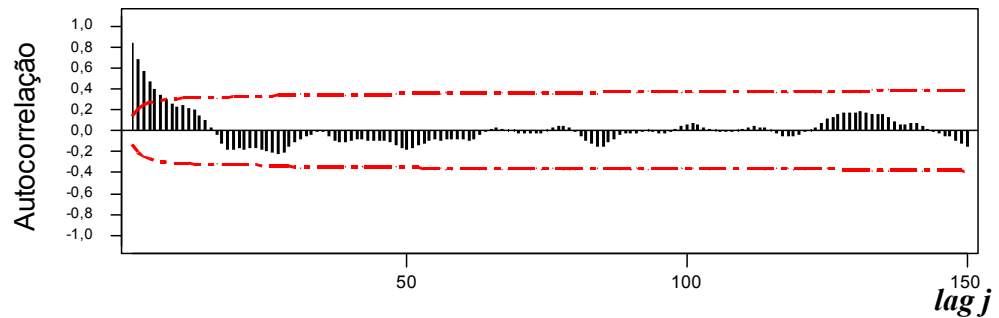


Figura 5.17. Autocorrelações em função dos *lags* para uma replicação de caixa normal para o dia 12 de dezembro.

Autocorrelação típica para caixa rápido

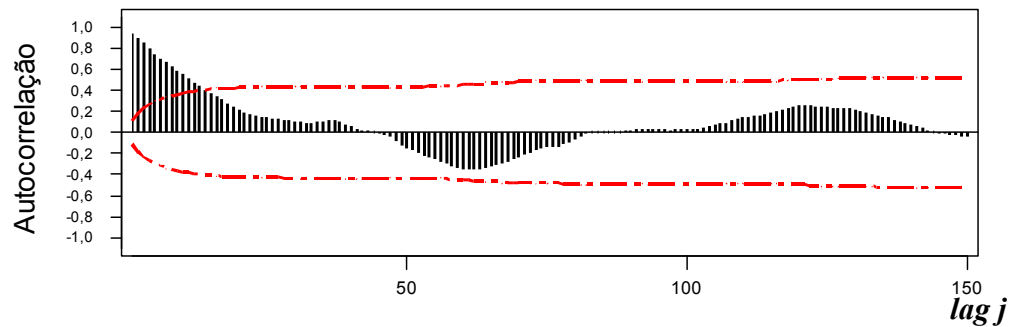


Figura 5.18. Autocorrelações em função dos *lags* para uma replicação de caixa rápido para o dia 11 de dezembro.

Observando-se as figuras, nota-se que o decaimento das autocorrelações para os caixas normais é mais acentuada e de menor magnitude que para as autocorrelações para os caixas rápidos. Isto, provavelmente, se deve ao fato de que, nos caixas rápidos, os tempos de fila são mais coesos pelo fato da limitação do número de itens a ser processado. De modo oposto, os caixas normais aceitam qualquer número de itens a ser processado, obtendo menor coesão e, portanto, um decaimento mais rápido. Deve-se observar ainda que isto ocorreu de forma geral para os dados analisados.

Em cada rodada de simulação realizada foram obtidas 30 replicações independentes, cujo tempo de execução foi menor que um segundo. Utilizando-se a expressão 5.13, foram calculados os limites dos intervalos de confiança com $\alpha=0,05$, e as

médias dos tempos de espera na fila. Nas tabelas 5.26 e 5.27 podem ser vistos os resultados obtidos pela simulação, pelo modelo com trocas e pela amostra coletada.

Tabela 5.26. Resultados da simulação, do modelo com trocas e das amostras para o tempo médio gasto na fila nos caixas normais.

Dia de coleta	$\lambda_E / m\mu$	Resultados da Simulação			Resultado Modelado	Resultado Observado
		Média	L.I.	L.S.		
03/12	0,841	2,09	1,75	2,43	1,78	1,80
04/12	0,878	3,07	2,73	3,41	3,2	3,36
05/12	0,859	2,21	1,87	2,55	2,19	1,85
10/12	0,812	1,70	1,36	2,04	1,46	1,32
11/12	0,883	3,67	3,33	4,01	3,44	3,63
12/12	0,893	2,56	2,22	2,90	3,21	2,92

Tabela 5.27. Resultados da simulação, do modelo com trocas e das amostras para o tempo médio gasto na fila nos caixas rápidos.

Dia de coleta	$\lambda_E / m\mu$	Resultados da Simulação			Resultado Modelado	Resultado Observado
		Média	L.I.	L.S.		
03/12	0,728	0,95	0,61	1,29	1,26	1,43
04/12	0,888	1,83	1,49	2,17	2,16	1,82
05/12	0,835	1,47	1,13	1,81	1,53	1,61
10/12	0,706	0,93	0,59	1,27	1,18	1,24
11/12	0,902	2,11	1,77	2,45	2,40	2,19
12/12	0,893	2,81	2,47	3,15	2,88	2,64

Pela tabela 5.26 pode-se observar que quase todos os intervalos de confiança cobrem os resultados obtidos pelo modelo com trocas, com exceção do dia 12/12, onde o resultado obtido por simulação apresentou uma razoável subestimação não capturando o resultado obtido por modelagem. Todas as condições para esse dia foram verificadas e nada que divergisse significativamente dos outros dias foi encontrado. Apesar deste fato, em geral, os resultados da simulação podem ser considerados satisfatórios para os caixas normais em relação aos resultados modelados.

Para os caixas rápidos, deve-se observar na tabela 5.27 que todos os resultados simulados (médias) estão subestimando os resultados obtidos pelo modelo. Uma possibilidade de explicação é que o período transitório inicial ainda possa estar influenciando o período subsequente, onde coletaram-se os dados referentes a tabela. Esta idéia pode ser reforçada pela magnitude das autocorrelações para os caixas rápidos, como pode ser notado na figura 5.18. Contudo, todos os intervalos de confiança obtidos por simulação estão cobrindo os resultados obtidos pelo modelo com trocas, com $\alpha=0,05$ de confiança, fato esse muito importante para que se possam utilizar os resultados obtidos por meio de simulação. Deve-se ressaltar ainda que o número de trocas observadas na simulação foi bem maior que as observadas nas amostras coletadas no supermercado. Estas eram de adjacências mais próximas à fila na qual originalmente o cliente se encontrava, o que não ocorre na simulação. De modo geral, os resultados alcançados pela simulação representam bem o modelo analítico com trocas, apesar das aproximações realizadas para as distribuições de tempos de serviço e do período de simulação ter sido fixado de acordo com a situação real.

6 - Conclusões

Os supermercados são empresas de auto-serviço e como tal têm se preocupado em melhorar continuamente seus desempenhos em nível de serviço para conseguirem fidelização de seus clientes. Para alcançar este objetivo é necessário que apresentem boa qualidade nos produtos oferecidos e nos serviços prestados. O supermercado pesquisado neste estudo possui dezesseis caixas, trabalha essencialmente com itens alimentares, caracteriza-se pelo controle familiar e atua na cidade de São Carlos há mais de quarenta anos. Desta forma, é uma empresa como grande parte das empresas brasileiras desse segmento.

Um dos objetivos deste trabalho foi o de revelar a importância relativa da rapidez de atendimento na opinião dos clientes que freqüentam o supermercado. Desta forma, esta importância recai principalmente sobre o tempo de espera na fila e suas conseqüências para os usuários. Outro objetivo deste trabalho foi o de analisar modelos analíticos de filas que pudessem ser mais eficazes para estimar o tempo de espera na fila em supermercados.

O gerenciamento de percepções em filas de espera em supermercados parece algo difícil de ser aplicado, motivando a pesquisa em modelagem para melhor gerenciar os tempos de espera. Contudo, como visto, aparentemente os clientes não são capazes de relacionar corretamente o tamanho de uma fila com seu tempo de espera. Este fato é interessante, pois esta relação (Lei de Little) existe, mas a dificuldade dos clientes em enxergá-la pode ser devido a conseguir medir (estimar) corretamente o tempo para atendimento, relativo ao número de pessoas que estão na fila.

Utilizando-se intervalos de confiança simultâneos foi possível observar que o preço não é a principal característica pela qual os clientes freqüentam o supermercado. Esta característica, na opinião dos clientes, ficou em quarto lugar de preferência, sendo que qualidade e rapidez de atendimento foram estatisticamente as mais apontadas. Por meio de análise de conglomerados hierárquicos de variáveis, pôde-se notar que qualidade e rapidez de atendimento foram associadas com diferentes tipos de correlações e ligações e, desta forma, compõem o que se poderia denominar de satisfação no atendimento, um possível fator de fidelização de clientes.

Foram levantadas características como, número de itens nas compras, percepção de diferença de preços em relação a outros estabelecimentos e a escolaridade dos clientes entrevistados. Estas informações foram associadas às características de frequência ao supermercado. Como resultado das análises de correspondência simples e múltiplas, as duas importantes características qualidade e rapidez revelaram-se associadas a escolaridades mais altas que, possivelmente, realizam compras maiores e acreditam não haver diferença significativa de preços. Otimizar estas características pode ser essencial para conseguir a fidelização dos clientes que não optam por preço como principal motivo de realizar ali suas compras.

Foram apresentados e empregados alguns sistemas de filas que poderiam ser utilizados para aproximar o sistema de filas comumente encontrado em supermercados, tais como os sistemas $M/M/m$, $M/M/m/mN$, $[M/M/1]^m$, $[M/M/1/N]^m$ e, ainda, os sistemas mais genéricos $M/G/m$ e $[M/G/1]^m$. O que se pode observar é que as filas paralelas e independentes, em geral, superestimaram os tempos de espera na fila, enquanto as restantes têm muita variabilidade em suas estimativas.

Um modelo analítico de filas com trocas foi empregado sendo que este foi capaz de reduzir para cerca da metade a média dos desvios absolutos de seus principais concorrentes, os modelos $M/M/m$ e $M/M/m/mN$. Também, demonstrou alguma robustez em absorver os efeitos causados por *outliers*, que foram mantidos na análise pois estes ocorrem no sistema de filas do supermercado. Deve-se ressaltar que, quando retirados estes pouco influenciaram nos resultados devido a estar-se analisando o tempo médio de espera na fila, ou seja, valores médios com boa quantidade de pontos. Ainda, o modelo, devido a permitir a realização de trocas de filas, se aproxima do comportamento de um sistema de fila única (apesar de ter filas em paralelo), tornando então o sistema possivelmente mais justo para o cliente que o utilize. Outro aspecto importante é em que situações os clientes efetuam as trocas. Parece que estas ocorrem em função do número de clientes nas filas, não importando a carga de trabalho (número de itens dos carrinhos à sua frente). Também, o número ganho de posições na troca de fila (k) parece estar relacionado ao fator de utilização da fila e do tipo de caixa.

Algumas perspectivas de pesquisa futura podem ser consideradas. Uma questão interessante é analisar o comportamento dos clientes no que se refere às trocas de filas

existentes no sistema. A diferença entre filas (k) é importante, pois revela, de certa forma, a impaciência dos clientes (isto é, quando $k=1$, por exemplo). Pode-se pensar que quanto maior o movimento no supermercado (isto é, quanto maior o valor de ρ), menor a disposição de esperar do cliente. Desta forma, pode-se investigar se k é uma função de ρ , ou seja, se é plausível supor que $k(\rho)$, e/ou se é uma função dos parâmetros, ou seja, $k(\lambda, \mu)$. Também, uma outra questão interessante que pode ser estudada é analisar o comportamento do tempo de espera na fila estimada pelo modelo, variando-se os valores dos parâmetros dentro de uma faixa de utilização plausível para sistemas reais. Pode-se, por exemplo, fazer uma superfície de resposta com as estimativas, e, em seguida, otimizá-la para se obter menores esperas.

Uma extensão do modelo seria obtida considerando-se classes de usuários, por exemplo, pelo número de itens comprados e em três categorias de carrinhos: leve, moderado e pesado. O espaço de estados do modelo de filas com trocas cresceria muito e, com isso, as dificuldades computacionais seriam bem maiores. Uma idéia seria considerar que os tempos de atendimento para os diferentes tipos de carrinhos são iguais, reduzindo novamente o espaço de estados. Contudo, ao se testar três filas de caixas normais do dia 5 de dezembro (pela ANOVA), em todas elas resultaram haver diferença de tempos de serviço para distintos tipos de carrinhos. Este resultado é interessante, pois na amostra coletada, por meio de entrevistas com os clientes, apenas 0,6% apontaram ter trocado de fila por avaliar haver carrinhos mais vazios em outra fila. Este fato tem implicações para o modelo com trocas: ele justifica a política adotada em trocas baseadas em número de clientes, e não em número de itens, e, com isso, o modelo com trocas consegue representar (ou capturar) o comportamento dos clientes. Este comportamento pode mudar se o cliente puder ser avisado quanto à sua percepção (errada), por exemplo, no supermercado do futuro (vide figura 2.2).

A aceleração (ou *speed-up*) do servidor também pode ser considerada nos modelos de filas. Para se verificar se isto ocorreu, pode-se testar se os tempos de serviço dos usuários atendidos sem fila são iguais aos dos usuários atendidos com fila. Uma questão a ser analisada é se o fato do valor de ρ ser alto pode provocar a aceleração do serviço a ponto de haver significância estatística. Pode-se pensar em utilizar também o modelo de filas com desistência na chegada dos clientes, contudo, nestes casos, geralmente,

existe grande dificuldade em estimar a taxa de chegadas (λ) em função do estado do sistema na prática.

Uma outra estratégia que pode ser considerada para o modelo de filas com trocas é a simulação discreta para problemas que apresentem espaços de estados muito grandes. Modelos de simulação podem ser úteis para validar as aproximações dos modelos analíticos para problemas maiores.

Utilizar modelos de filas que representem bem a situação encontrada em supermercados parece ser o caminho para obter a melhor política de gerenciamento dos tempos de espera na fila. Com este objetivo, as trocas de filas, comuns e já incorporadas ao comportamento dos clientes, não deve ser desprezada nestes modelos.

REFERÊNCIAS

- AAKER, D.A.; KUMAR, V.; DAY, G.S. *Pesquisa de marketing*. Tradução Reynaldo Cavalheiro Marcondes. 2. ed. São Paulo: Atlas, 2004.
- ARENALES, M.; ARMENTANO, V.; MORABITO, R.; YANASSE, H. *Pesquisa Operacional*. Rio de Janeiro: Elsevier, 2007.
- BADIN, N.T. *Avaliação da produtividade de supermercados e seu benchmarking*. 1997. 130f. Dissertação (Mestrado em Engenharia de Produção) – Universidade Federal de Santa Catarina, Florianópolis, 1997.
- BESTFIT - Probability distribution fitting for windows, versão 4. New York: Palisade Tools, 2000.
- BITRAN, G.; MORABITO, R. Open queueing networks: optimization and performance evaluation models for discrete manufacturing systems. *Production and Operations Management*, n.5, p.163-193, 1996.
- BOBBITT, L.M.; DABHOLKAR, P.A. Integrating attitudinal theories to understand and predict use of technology-based self-service: the internet as an illustration. *International Journal of Service Industry Management*, v.12, n.5, p.423-450, 2001.
- BRASIL. Comércio varejista: supermercados. In: *BNDES: Banco Nacional de Desenvolvimento*, Brasília, nov. 1996. Disponível em <<http://www.bndes.gov/conhecimento/relato/supermerca.pdf>>. Acesso em 11 jul. 2006.
- CARVALHO, H. *Análise multivariada de dados qualitativos: utilização da HOMALS com o SPSS*. Lisboa: Silabo, 2004.
- COGAN, S. *Gerenciando as percepções nas filas de espera: para aumentar o nível de satisfação dos clientes*. Rio de Janeiro: Qualitymark, 1998.
- CONOVER, W.J. *Practical nonparametric statistics*. 3.ed. New York: John Wiley, 1999.

DABHOLKAR, P.A.; BOBBITT, L.M.; LEE, E.J. Understanding consumer motivation and behavior related to self-scanning in retailing: implications for strategy and research on technology-based self-service. *International Journal of Service Industry Management*, v.14, n.1, p.59-95, 2003.

DABHOLKAR, P.A.; BAGOZZI, R.P. An attitudinal model of technology-based self-service: moderating effects of consumer traits and situational factors. *Journal of the Academy of Marketing Science*, v.30, n.3, p. 184-201, 2002.

Dispositivo substitui o código de barras no supermercado do futuro. *ESTADÃO - O Estado de São Paulo*, São Paulo, 21 mar. 2005. Disponível em: <<http://www.estadao.com.br>> Acesso em: 11 jul. 2006.

Empresas apostam em tecnologia para reduzir filas nas lojas. In: *ABASE – Associação Bahiana de Supermercados*. Bahia, 14 mar. 2006. Disponível em: <http://www.abase-ba.org.br/noticias/folder_noticias.2006-03-02.0126144930/view> Acesso em: 11 jul. 2006.

FERRARI, S.C. *Filas paralelas com servidores heterogêneos e jockeying probabilístico*. 2002. 76f. Dissertação (Mestrado em Ciências de Computação e Matemática Computacional - ICMC) – Universidade de São Paulo, São Carlos, 2002.

FISHMAN, G.S. *Monte Carlo: concepts, algorithms, and applications*. New York: Springer-Verlag, 1996.

FONTANELLA, G.; MORABITO, R. Analyzing the tradeoff between investing in service channels and satisfying the target user service for Brazilian Internet service providers. *International Transactions in Operational Research*, v.9, n.3, p.247-260, 2002.

FOX, B.L.; GOLDSMAN, D.; SWAIN, J.J. Spaced batch means. *Opns. Res. Letters*, v.10, p.255-263, 1991.

FREITAS, P. *Introdução à modelagem e simulação de sistemas - com aplicações em Arena*. Florianópolis: Visual Books, 2001.

GREENACRE, M.; BLASIUS, J. *Multiple correspondence analysis and related methods*. Londres: Chapman & Hall/CRC, 2006.

GROSS, D; HARRIS, C. M. *Fundamentals of Queueing Theory*. New York: John Wiley & Sons, 1974.

GUERRA, S. Mais por menos: produtividade para resultados. *Associação dos analistas e profissionais de investimentos do mercado de capitais*, jun.2006. Disponível em: <www.apimecmg.com.br>. Acesso em: 11 de jul. 2006.

HAIR, J.F.Jr. et al. *Análise multivariada de dados*. Trad. Adonai Schlup Sant'Anna e Anselmo Chaves Neto. 5.ed. Porto Alegre: Bookman, 2005.

ITTIG, P.T. The real cost of making customers wait. *International Journal of Service Industry Management*, v.13, n.3, p.231-241, 2002.

JOHNSON, R.A.; WICHERN, D.W. *Applied multivariate statistical analysis*. 5.ed. New Jersey: Prattice Hall, 2002.

JOHNSON, R.A.; WICHERN, D.W. *Applied multivariate statistical analysis*. 4.ed. New Jersey: Prattice Hall, 1998.

KATZ, K.; LARSON, B.M.; LARSON, R.C. Prescription for the waiting-in-line blues: entertain, enlighten, and engage. *Sloan Management Review*, v.32, n.2, p.44-53, 1991.

KLEINROCK, L. *Queueing systems*. New York: John Wiley & Sons, 1975.

KOENIGSBERG, E. On jockeying in queues. *Management Science*, v.12, p.412-436, 1966.

LARSON, R.C. Perspectives on queues: social justice and the psychology of queueing. *Operations Research*, v.35, n. 6, p.895-905, nov-dec, 1987.

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

_____, Confidence intervals for steady-state simulations, I: a survey of fixed sample size procedures. *Operations Res.*, n.32, p.1221-1239, 1984.

LAW, A.M. Confidence intervals in discrete event simulation: a comparison of replication and batch means. *Naval Res. Logist. Quart.*, v.24, p.669-678, 1977.

MEUTER, M.L. et al. Self-service technologies: understanding customer satisfaction with technology-based service encounters. *Journal of Marketing*, v.64,p.50-64, jul., 2000.

MITRANI, I. *Probabilistic modelling*. Cambridge: University Press, 1998.

MONTGOMERY, D.C.; PECK, E.A. *Introduction to linear regression analysis*. New York: John Wiley, 1992.

MORABITO, R.; LIMA, F.C.R. A markovian queueing model for the analysis of user waiting times in supermarket checkouts. *International Journal of Operations and Quantitative Management*, v.10, n, 2, p.165-177, jun., 2004.

PAWLIKOWSKI, K. Steady-state simulation of queueing processes: a survey of problems and solutions. *ACM Computing Surveys*, v.22, p.123-169, 1990.

PEREIRA, J.C.R. *Análise de dados qualitativos: estratégias metodológicas para as ciências da Saúde, Humanas e Sociais*. 3.ed. São Paulo: Edusp, 2004.

QUEIROZ, T.S.; GOLVINHAS, R.P.; SOUZA, T. Marketing de varejo: percepção dos clientes em relação aos serviços oferecidos pelo supermercado de bairro. In: *ENEGEP – ENCONTRO NACIONAL DOS ESTUDANTES DE GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO*, 24, 2004. Florianópolis.

REA, L.M.; PARKER, R.A. *Metodologia de pesquisa*. São Paulo: Pioneira, 2000.

REIS, E. *Estatística multivariada aplicada*. 2.ed. Lisboa: Silabo, 2001 .

Representantes de supermercados discutem sobre filas. *Jornal da Câmara*, Blumenau, 12 jun. 2006. Disponível em:
<<http://www.camarablumenal.sc.gov.br/noticias/novojournal/gerais/pag109.htm>> Acesso em: 11 jul. 2006.

RINALDI, J.G.S. *Aplicação de métodos ABATCH e LBATCH na simulação da fila M/M/1*. 1999. 67f. Dissertação (Mestrado em Ciências Matemáticas) – Universidade de São Paulo, São Carlos, 1999.

SÃO PAULO. Projeto de Lei nº. 372, de 2 de dezembro de 2005. Dispõe sobre o tempo máximo de espera para atendimento nos caixas de supermercados, hipermercados e estabelecimentos congêneres instalados no município de São Paulo e dá outras providências. *Diário Oficial do Município de São Paulo*, São Paulo, SP, ano 50, n.226, 2 dez. 2005. Disponível em: <<http://www.fecomercio.com.br/sistema/downloads/arquivos/PLM372jun05>>. Acesso em: 11jun. 2006.

YIN, R.K. *Estudo de caso: planejamento e métodos*. Trad. Daniel Grassi. 2.ed. Porto Alegre: Bookman, 2001.

WOLFF, R.W. *Stochastic modeling and the theory of queues*. New Jersey: Prentice Hall, 1989.

ZHAO, Y.; GRASSMANN, W.K. Queueing analysis of a jockeying model. *Operations Research*, v.43, n.3, p.520-529, 1995.

APÊNDICE A

QUESTIONÁRIO

Questões

Qual o tempo máximo aceitável de espera em uma fila de caixa? _____ minutos

Você (Sr., Sra., ...) abandonaria suas compras se observasse quantos clientes nas filas?

Ordenar quais as características mais importantes pelas quais frequenta este supermercado:

- () preço dos produtos
- () localização do supermercado
- () rapidez no atendimento dos caixas (tempo de fila, etc)
- () qualidade de atendimento nos caixas (cortesia, etc)
- () variedade de mercadorias no supermercado
- () estacionamento
- () outras; _____

Escolaridade do entrevistado _____

Número de itens comprados: _____

Você (Sr, Sra,...) acredita que para uma compra exista diferença significativa de preços entre supermercados distintos? _____

Para os clientes que realizaram *jockeying*:

Qual o motivo pelo qual houve a troca de fila:

- () menor número de clientes na outra fila
- () carrinhos mais vazios (menos carregados) na outra fila
- () os dois motivos anteriores
- () considerou que o outro caixa estava mais rápido
- () predileção pelo outro caixa
- () outros motivos _____

APÊNDICE B

PLANILHA

Filas: ____ e ____ Dia: _____ (_____ Horas) N.: _____

Tempos: Horas (H): Minutos (M): Segundos (S) (H/M/S)

Carga: Leve (L); Moderada (M); Pesada (P) Desistência: D → (H/M/S)

Jockeying na fila i: Fila (i-n) ←(H/M/S)← Fila (i) →(H/M/S)→ Fila (i+n) p/n=1,2,...

N.	Carga	Chegada na fila	Entrada em serviço	Sai do serviço
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				
17				
18				
19				
20				
21				
22				
23				
24				
25				
26				
27				
28				
29				

N.	Carga	Chegada na fila	Entrada em serviço	Sai do serviço
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				
17				
18				
19				
20				
21				
22				
23				
24				
25				
26				
27				
28				
29				

APÊNDICE C

TERMO DE COMPROMISSO

Eu _____, RG _____, estarei realizando prestação de serviços à **José Gilberto Spasiani Rinaldi**, aluno do Programa de Pós-Graduação em Engenharia de Produção, nível Doutorado, da Universidade Federal de São Carlos, referente a coleta de dados em uma Rede de Supermercados na cidade de São Carlos nos dias 03, 04, 05/12/2004 e 10, 11, 12/12/2004.

Torna-se declarado no presente Termo de Compromisso que estou ciente:

- 1- de que o Processo de Coleta de Dados se pauta em padrões ético-profissionais (sigilo quanto as informações obtidas por meio dos instrumentos de coleta: planilhas);
- 2- de que todo material utilizado para coleta dos dados (planilhas originais manuscritas e digitadas, respectivamente junto com o disquete) será entregue ao responsável José Gilberto Spasiani Rinaldi;
- 3- de que estarei no Supermercado por um período de tempo necessário para desenvolvimento da coleta;
- 4- de que em hipótese alguma serão veiculadas informações obtidas pela coleta de dados em qualquer ambiente ou meio de comunicação;

São Carlos, 2 de dezembro de 2004

(prestador de serviço)

RG:

RECIBO

Eu _____, RG _____, recebi de **José Gilberto Spasiani Rinaldi** a quantia de R\$ _____, referente à coleta de dados nos dias 03, 04, 05/12/2004 e 10, 11, 12/12/2004 totalizando ___h de serviços prestados e para clareza firmo o presente.

São Carlos, 14 de dezembro de 2004.

(prestador de serviço)

RG:

APÊNDICE D

DEFINIÇÃO DE AUTOVALORES E AUTOVETORES

Definição de autovalor:

Seja \mathbf{A} uma matriz quadrada $k \times k$ e \mathbf{I} a matriz identidade $k \times k$. Então, os escalares $\lambda_1, \lambda_2, \dots, \lambda_k$ satisfazendo a equação polinomial $|\mathbf{A} - \lambda \mathbf{I}| = 0$ são denominados autovalores da matriz \mathbf{A} .

Definição de autovetor:

Seja \mathbf{A} uma matriz quadrada $k \times k$ e λ_i um autovalor de \mathbf{A} . Se $\mathbf{x}_{k \times 1}$ é um vetor tal que $\mathbf{x}_{k \times 1} \neq \mathbf{0}_{k \times 1}$ (não é um vetor com elementos iguais a zero), e também

$$\mathbf{Ax} = \lambda \mathbf{x}$$

então $\mathbf{x}_{k \times 1}$ é dito ser um autovetor da matriz \mathbf{A} associado ao autovalor λ_i .

APÊNDICE E

Exemplo com todas as respectivas equações do sistema de fila com trocas, sendo $N=3$, $k=1$, $m=4$, $\lambda_E=2,0803$ e $\mu=0,6228$.

estado 0000:

$$(0.52 + 0.52 + 0.52 + 0.52) p[0000] = 0.62 p[1000] + 0.62 p[0100] + 0.62 p[0010] + 0.62 p[0001]$$

estado 0001:

$$(0.69 + 0.69 + 0.69 + 0.62) p[0001] = 0.52 p[0000] + 0.62 p[1001] + 0.62 p[0101] + 0.62 p[0011]$$

estado 0010:

$$(0.69 + 0.69 + 0.69 + 0.62) p[0010] = 0.52 p[0000] + 0.62 p[1010] + 0.62 p[0110] + 0.62 p[0011]$$

estado 0011:

$$(1.04 + 1.04 + 0.62 + 0.62) p[0011] = 0.69 p[0001] + 0.69 p[0010] + 0.62 p[1011] + 0.62 p[0111]$$

estado 0100:

$$(0.69 + 0.69 + 0.69 + 0.62) p[0100] = 0.52 p[0000] + 0.62 p[1100] + 0.62 p[0110] + 0.62 p[0101]$$

estado 0101:

$$(1.04 + 1.04 + 0.62 + 0.62) p[0101] = 0.69 p[0001] + 0.69 p[0100] + 0.62 p[1101] + 0.62 p[0111]$$

estado 0110:

$$(1.04 + 1.04 + 0.62 + 0.62) p[0110] = 0.69 p[0010] + 0.69 p[0100] + 0.62 p[1110] + 0.62 p[0111]$$

estado 0111:

$$(2.08 + 0.62 + 0.62 + 0.62) p[0111] = 1.04 p[0011] + 1.04 p[0101] + 1.04 p[0110] + 0.62 p[1111]$$

estado 1000:

$$(0.69 + 0.69 + 0.69 + 0.62) p[1000] = 0.52 p[0000] + 0.62 p[1100] + 0.62 p[1010] + 0.62 p[1001]$$

estado 1001:

$$(1.04 + 1.04 + 0.62 + 0.62) p[1001] = 0.69 p[0001] + 0.69 p[1000] + 0.62 p[1101] + 0.62 p[1011]$$

estado 1010:

$$(1.04 + 1.04 + 0.62 + 0.62) p[1010] = 0.69 p[0010] + 0.69 p[1000] + 0.62 p[1110] + 0.62 p[1011]$$

estado 1011:

$$(2.08 + 0.62 + 0.62 + 0.62) p[1011] = 1.04 p[0011] + 1.04 p[1001] + 1.04 p[1010] + 0.62 p[1111]$$

estado 1100:

$$(1.04 + 1.04 + 0.62 + 0.62) p[1100] = 0.69 p[0100] + 0.69 p[1000] + 0.62 p[1110] + 0.62 p[1101]$$

estado 1101:

$$(2.08 + 0.62 + 0.62 + 0.62) p[1101] = 1.04 p[0101] + 1.04 p[1001] + 1.04 p[1100] + 0.62 p[1111]$$

estado 1110:

$$(2.08 + 0.62 + 0.62 + 0.62) p[1110] = 1.04 p[0110] + 1.04 p[1010] + 1.04 p[1100] + 0.62 p[1111]$$

estado 1111:

$$(0.52 + 0.52 + 0.52 + 0.52 + 0.62 + 0.62 + 0.62 + 0.62) p[1111] = 2.08 p[0111] + 2.08 p[1011] + 2.08 p[1101] + 2.08 p[1110] + 2.49 p[2111] + 2.49 p[1211] + 2.49 p[1121] + 2.49 p[1112]$$

estado 1112:

$$(0.69 + 0.69 + 0.69 + 2.49) p[1112] = 0.52 p[1111] + 1.25 p[2112] + 1.25 p[1212] + 1.25 p[1122]$$

estado 1121:

$$(0.69 + 0.69 + 0.69 + 2.49) p[1121] = 0.52 p[1111] + 1.25 p[2121] + 1.25 p[1221] + 1.25 p[1122]$$

estado 1122:

$$(1.04 + 1.04 + 1.25 + 1.25) p[1122] = 0.69 p[1112] + 0.69 p[1121] + 0.83 p[2122] + 0.83 p[1222]$$

estado 1211:

$$(0.69 + 0.69 + 0.69 + 2.49) p[1211] = 0.52 p[1111] + 1.25 p[2211] + 1.25 p[1221] + 1.25 p[1212]$$

estado 1212:

$$(1.04 + 1.04 + 1.25 + 1.25) p[1212] = 0.69 p[1112] + 0.69 p[1211] + 0.83 p[2212] + 0.83 p[1222]$$

estado 1221:

$$(1.04 + 1.04 + 1.25 + 1.25) p[1221] = 0.69 p[1121] + 0.69 p[1211] + 0.83 p[2221] + 0.83 p[1222]$$

estado 1222:

$$(2.08 + 0.83 + 0.83 + 0.83) p[1222] = 1.04 p[1122] + 1.04 p[1212] + 1.04 p[1221] + 0.62 p[2222]$$

estado 2111:

$$(0.69 + 0.69 + 0.69 + 2.49) p[2111] = 0.52 p[1111] + 1.25 p[2211] + 1.25 p[2121] + 1.25 p[2112]$$

estado 2112:

$$(1.04 + 1.04 + 1.25 + 1.25) p[2112] = 0.69 p[1112] + 0.69 p[2111] + 0.83 p[2212] + 0.83 p[2122]$$

estado 2121:

$$(1.04 + 1.04 + 1.25 + 1.25) p[2121] = 0.69 p[1121] + 0.69 p[2111] + 0.83 p[2221] + 0.83 p[2122]$$

estado 2122:

$$(2.08 + 0.83 + 0.83 + 0.83) p[2122] = 1.04 p[1122] + 1.04 p[2112] + 1.04 p[2121] + 0.62 p[2222]$$

estado 2211:

$$(1.04 + 1.04 + 1.25 + 1.25) p[2211] = 0.69 p[1211] + 0.69 p[2111] + 0.83 p[2221] + 0.83 p[2212]$$

estado 2212:

$$(2.08 + 0.83 + 0.83 + 0.83) p[2212] = 1.04 p[1212] + 1.04 p[2112] + 1.04 p[2211] + 0.62 p[2222]$$

estado 2221:

$$(2.08 + 0.83 + 0.83 + 0.83) p[2221] = 1.04 p[1221] + 1.04 p[2121] + 1.04 p[2211] + 0.62 p[2222]$$

estado 2222:

$$(0.52 + 0.52 + 0.52 + 0.52 + 0.62 + 0.62 + 0.62 + 0.62) p[2222] = 2.08 p[1222] + 2.08 p[2122] + 2.08 p[2212] + 2.08 p[2221] + 2.49 p[3222] + 2.49 p[2322] + 2.49 p[2232] + 2.49 p[2223]$$

estado 2223:

$$(0.69 + 0.69 + 0.69 + 2.49) p[2223] = 0.52 p[2222] + 1.25 p[3223] + 1.25 p[2323] + 1.25 p[2233]$$

estado 2232:

$$(0.69 + 0.69 + 0.69 + 2.49) p[2232] = 0.52 p[2222] + 1.25 p[3232] + 1.25 p[2332] + 1.25 p[2233]$$

estado 2233:

$$(1.04 + 1.04 + 1.25 + 1.25) p[2233] = 0.69 p[2223] + 0.69 p[2232] + 0.83 p[3233] + 0.83 p[2333]$$

estado 2322:

$$(0.69 + 0.69 + 0.69 + 2.49) p[2322] = 0.52 p[2222] + 1.25 p[3322] + 1.25 p[2332] + 1.25 p[2323]$$

estado 2323:

$$(1.04 + 1.04 + 1.25 + 1.25) p[2323] = 0.69 p[2223] + 0.69 p[2322] + 0.83 p[3323] + 0.83 p[2333]$$

estado 2332:

$$(1.04 + 1.04 + 1.25 + 1.25) p[2332] = 0.69 p[2232] + 0.69 p[2322] + 0.83 p[3332] + 0.83 p[2333]$$

estado 2333:

$$(2.08 + 0.83 + 0.83 + 0.83) p[2333] = 1.04 p[2233] + 1.04 p[2323] + 1.04 p[2332] + 0.62 p[3333]$$

estado 3222:

$$(0.69 + 0.69 + 0.69 + 2.49) p[3222] = 0.52 p[2222] + 1.25 p[3322] + 1.25 p[3232] + 1.25 p[3223]$$

estado 3223:

$$(1.04 + 1.04 + 1.25 + 1.25) p[3223] = 0.69 p[2223] + 0.69 p[3222] + 0.83 p[3323] + 0.83 p[3233]$$

estado 3232:

$$(1.04 + 1.04 + 1.25 + 1.25) p[3232] = 0.69 p[2232] + 0.69 p[3222] + 0.83 p[3332] + 0.83 p[3233]$$

estado 3233:

$$(2.08 + 0.83 + 0.83 + 0.83) p[3233] = 1.04 p[2233] + 1.04 p[3223] + 1.04 p[3232] + 0.62 p[3333]$$

estado 3322:

$$(1.04 + 1.04 + 1.25 + 1.25) p[3322] = 0.69 p[2322] + 0.69 p[3222] + 0.83 p[3332] + 0.83 p[3323]$$

estado 3323:

$$(2.08 + 0.83 + 0.83 + 0.83) p[3323] = 1.04 p[2323] + 1.04 p[3223] + 1.04 p[3322] + 0.62 p[3333]$$

estado 3332:

$$(2.08 + 0.83 + 0.83 + 0.83) p[3332] = 1.04 p[2332] + 1.04 p[3232] + 1.04 p[3322] + 0.62 p[3333]$$

estado 3333:

$$(0.62 + 0.62 + 0.62 + 0.62) p[3333] = 2.08 p[2333] + 2.08 p[3233] + 2.08 p[3323] + 2.08 p[3332]$$